Robust 2-D Object Recognition Using Bispectrum and LVQ Neural Classifier

Han, Soowhan* **and Woo, Young-woon****

* *Dept. of Multimedia Eng., Dongeui University*
** *Dept. of Computer Eng., Dongeui University*

Abstract

This paper presents a translation, rotation and scale invariant methodology for the recognition of closed planar shape images using the bispectrum of a contour sequence and the learning vector quantization(LVQ) neural classifier. The contour sequences obtained from the closed 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 higher order spectra based on third order cumulants is applied to this contour sample to extract fifteen bispectral feature vectors for each planar image. These feature vectors, which are invariant to shape translation, rotation and scale transformation, can be used to represent two-dimensional planar images and are fed into a neural network classifier. The LVQ architecture is chosen as a neural network classifier because the network is easy and fast to train, the structure is relatively simple. The experimental recognition processes with eight different shapes of aircraft images are presented to illustrate the high performance of this proposed method even the target images are significantly corrupted by noise.

I. Introduction

In this study, an investigation of the classification of objects by the use of bispectral feature vectors that represent the shapes of boundaries detected in the digitized images is presented. The developed recognition process is independent of object size and planar orientation.

The studies on two-dimensional object recognition problem have broad applications such as satellite image identification, the 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-varying. Additionally, the systems are required to be tolerant to noisy shapes results from the segmentation of objects in varying backgrounds as well as non-ideal imaging conditions. There have been over a dozen prior research efforts including Fourier descriptors[1], autoregressive modeling method[2]-[4], dynamic alignment process of contour sequences[5], and neural network approach[6][7].

The neural network classifier in the classification tasks is widely used. A neural network can perform the necessary and suitable transformation and clustering operations automatically and simultaneously.

The recognition accuracy on neural network approach for pattern recognition problems, while keeping simplicity of the overall system, depends on two important factors. One is to extract feature vectors representing a 2-D object image. The feature vectors should have the small dimensionality for real-time process, the similarity between intraclass. In this study, the boundary of a closed planar shape is characterized by an ordered sequence that represents the Euclidean distance between the centroid 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 the same starting boundary pixel remains unchanged even the image is shifted. Then, the contour sequence in normalized with respect to the size of image. This normalization includes the amplitude and the duration of the contour sequence. Next the bispectrum based on third order cumulants is applied to this normalized contour sequence as a means of feature selection. Higher order spectra(bispectrum, trispectrum) play an important role in digital signal processing due to their ability of preserving nonminimum phase information, as well as information due to deviations from Gaussianity and degrees of nonlinearities in time series[8]. In the last few years, bispectral analysis has been an active research area. The applications of the bispectrum extend over several disciplines. These applications include ARMA modeling, analysis of bilinear models, detection of phase coupling, signal reconstruction, image processing, radar signal detection, and so on [9]-[11]. Therefore, in this investigation of 2-D object classification, the bispectral components of the normalized contour sequence of the object image are utilized as feature vectors. These bispectral feature vectors have enough shape information to represent each 2-D object, a property to be invariant in size, shift, and rotation, and are used as the input of neural network classifier.

Another factor is to select an appropriate neural network architecture for this particular recognition task. In this paper, the learning vector quantization(LVQ) is utilized as an artificial neural network(ANN) based classifier because of its simples structure and fast training procedure[12]-[14]. It is a pattern classification neural algorithm in which each output units represents a particular class or category. The weight vector for an output unit is often referred to as a reference vector for the class that the unit represents. During training, the output units are positioned (by adjusting their weights through supervised training) to approximate the decision surface for each class. After the learning process is completed, an LVQ classifies an input vector by assigning it to the same class as the output unit that has its weight vector closest to the input vector. In the experimental procedure of this study, both of LVQ1 and its improved versions (LVQ2, LVQ3) are tested.

II. Contour Representation and Bispectral Feature Measurement

A careful examination of the boundary of a closed planar shape reveals that similar shapes have similar boundaries. That is, essential shape information is contained in the boundary of the shape. Thus, in this portion of the study, the boundary of a closed planar shape is characterized by an ordered sequence that represents the Euclidean distance between the centroid and all contour pixels of the digitized shape. First, the boundary pixels are extracted by using contour following algorithm from the planar image and the centroid is derived[15]-[16]. 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).

$$b(N+i)=b(i) \qquad i=1,2,3,......,N \qquad (1)$$

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, 256 points fixed, is obtained by resampling operation and function approximation. This is shown in equation (2).

$$c(k)=b(k*N/256) \quad k=1,2,3......,256 \qquad (2)$$

where N is the total number of boundary pixels. After duration normalization, the amplitude is divided by the sum of contour sequence and removed the mean. It is shown in equation (3) and (4).

$$d(k)=c(k)/s \quad k=1,2,3,.....,256 \tag{3}$$
$$d(k)=d(k)-\mathrm{mean}(d(k)) \tag{4}$$
where $s=c(1)+c(2)+c(3)+........+c(256)$

This sequence d(k) is invariant to translation and scaling. In a forth, bispectral feature measurement is taken into the sequence d(k). Higher-order spectra can address noise suppression, and preserve nonminimum phase information as well as the information due to degrees of nonlinearities[17]. Thus, it has been widely used in the area of the identification of nonminimum phase systems, detection of phase coupling and ARMA modeling[18]-[19]. The *n*th order moment spectrum of contour sequence d(k) is defined as

$$\begin{aligned} &H_n(\omega_1, \ldots, \omega_{n-1}) \\ &= F(\omega_1) \cdot \cdot \cdot F(\omega_{n-1}) F^*(\omega_1 + \cdot \cdot \cdot + \omega_{n-1}) \end{aligned} \tag{5}$$

where $F(\omega)$ is the Fourier transform of the sequence d(k). For the special case where n=3 (bispectrum) :

$$H_3(\omega_1, \omega_2) = F(\omega_1) F(\omega_2) F^*(\omega_1 + \omega_2) \tag{6}$$

where $|\omega_1| \leq \pi$, $|\omega_2| \leq \pi$, $|\omega_1 + \omega_2| \leq \pi$. If n=2, equation (5) becomes the power spectrum of contour sequence d(k). The magnitude of bispectrum, $|H_3(\omega_1, \omega_2)|$, is unchanged even after the sequence d(k) is circular shifted because the magnitude of Fourier transform, $|F(\omega)|$, is not changed[20]. Thus $|H_3(\omega_1, \omega_2)|$ is invariant to the rotation of an image. Finally, the two dimensional bispectral magnitude (256 by 256) is projected to vertical axis (ω_1) by taking mean value of each column for feature extraction. It is shown in equation (7).

$$h[k]=[\mathrm{mean}(k\textit{th} \text{ column of } |H_3(\omega_1, \omega_2)|)] \tag{7}$$

where k=1,2,....256. The first column and the row in the magnitude of bispectrum contain all zero value because the normalized contour sequence has a zero mean. It means h(1) is always zero. And the projected bispectral components exceed to the sixteenth have very small values (near zero). Thus, for fast recognition process with reliable accuracy, the projected bispectral components from the second to the sixteenth (h(2), h(3),.....h(16)) are chosen to be used as feature vectors to represent each image shape, which are fed into LVQ neural network classifier for recognition process. The process for feature extraction is shown in figure 1.

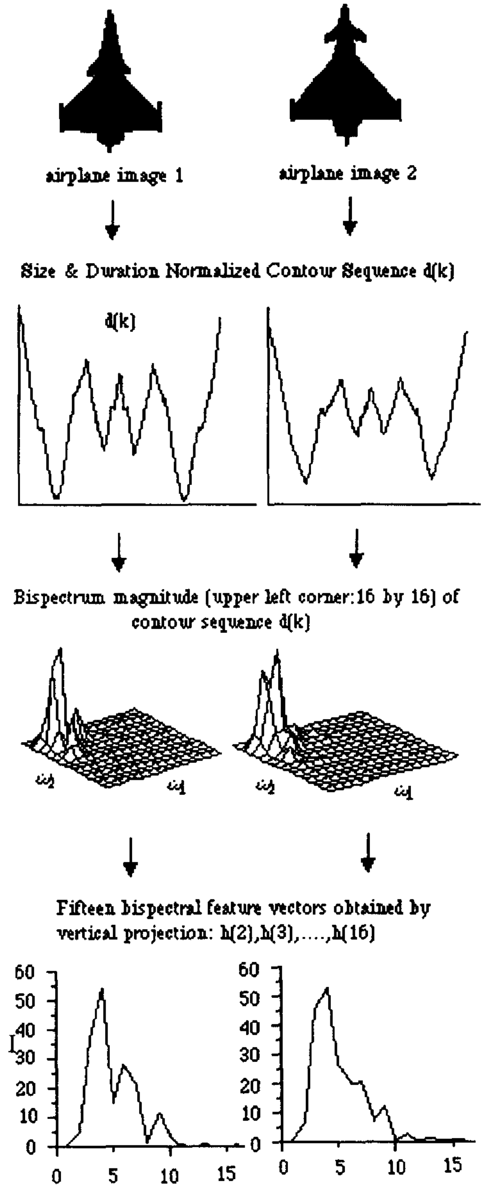

Figure 1. The process for feature extraction

III. Neural Classifier : Learning Vector Quantization (LVQ)

The recognition task being considered in this paper requires supervised training. That means, the classifier requires class labels to be specified during training. Among many different structures of neural networks, Learning Vector Quantization (LVQ) is chosen because of its simple structure and training procedure. LVQ is a pattern classification method in which each output unit represents a particular class or category [14]-[16], [23]. It is mainly intended to approximate input patterns, or their probability density function, by quantized "codebook" vectors that are localized in the input space to minimize a quantization error functional. On the other hand, if the patterns are to be classified into a finite number of categories, then several codebook vectors are usually made to represent each class. In LVQ algorithm, the weight vector is often referred to as a reference (or codebook) vector for the class that the unit represents. During training procedure, the output units are positioned by adjusting their weights through supervised training to approximate the decision surfaces of the theoretical Bayes classifier. It is assumed that a set of training pattern with known classification is provided, along with an initial distribution of reference vectors. After the learning procedure is completed, an LVQ net classifies an input vector by assigning it to the same class as the output unit that has its weight vector closest to the input vector. The architecture of an LVQ neural net, shown in figure 2, is essentially the same as that of a Kohonen self-organizing map without a topological structure being assumed for the output units. In addition, each output unit has a known class that it represents.

In the original LVQ algorithm, called LVQ1, only the reference vector that is closest to the input vector is updated. The direction it is moved depends on whether the winning reference vector belongs to the same class as the input vector. In the improved algorithms, called LVQ2 and LVQ3, two vectors (the winner and a runner-up) learn if several conditions are satisfied. The idea is that if the input is approximately the same distance from both the winner and the runner-up, then each of them should learn. More details concerning learning algorithm can be found in [12]-[14], [21] for LVQ1, [22] for LVQ2 and [23] for LVQ3. For the experimental results shown in the next section, all of LVQ algorithms (LVQ1, LVQ2, and LVQ3) were trained with a fixed learning rate=0.01 and the first m training vectors were utilized as the initial weight vectors for fast learning.

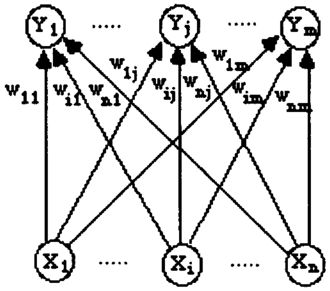

Figure 2. Learning vector quantization neural net

Ⅳ. Experimental Results and Performance Assessment

The methodology presented in this paper, for the recognition of closed planar shape, was evaluated with eight different shapes of aircrafts. They are shown in figure 3. From each reference shape of aircraft, 36 noisy-free patterns were generated by rotating the original image with 30 degree increment and scaling with three factor (1, 0.8 and 0,6). And forty noisy corrupted patterns were made by adding four different level of random gaussian noise (25db, 20db, 15db, 10db SNR : ten noisy patterns for each SNR) to 36 noisy-free patterns. Thus the data set for each reference aircraft image has 36 noisy-free patterns and 1440 (40×36) noisy corrupted patterns. The number of total test patterns becomes 11808 (1476 ×8 reference image).

The fifteen bispectral feature vectors extracted from the reference aircraft image are shown in figure 4. This plot shows that the bispectral feature vectors have the acceptable discrimination ability between each aircraft image.

In this study, the experimental process was performed under eight different simulation environments depend on the types of LVQ and the number of training patterns and output clusters.

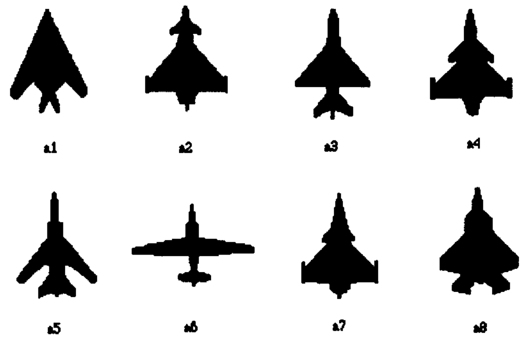

Figure 3. Eight different shapes of reference aircraft images

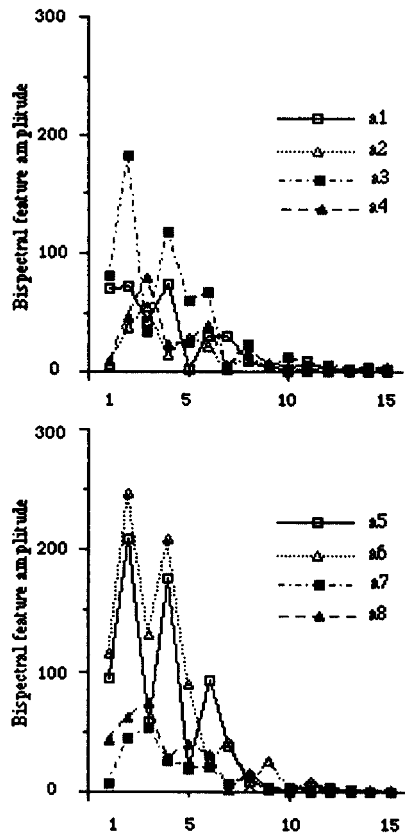

Figure 4. The fifteen bispectral feature vectors extracted from the eight reference aircraft images

These are as follows.

1. Neural classifier algorithm : LVQ1 with 8 output

clusters (one cluster for each reference shape).
Training data set : only the 8 reference aircraft images.
2. Neural classifier algorithm : LVQ1 with 8 output clusters (one cluster for each reference shape).
Training data set : 8 reference patterns + 32 noisy patterns (4 noisy patterns with 25db SNR generated from each of 8 reference images).
3. Neural classifier algorithm : LVQ1 with 8 output clusters (one cluster for each reference shape).
Training data set : 8 reference patterns +32 noisy patterns (4 noisy patterns with each of 25db, 20db, 15db and 10db SNR generated from each of 8 reference images).
4. Neural classifier algorithm : LVQ2 with 8 output clusters (one cluster for each reference shape).
Training data set : same as 3.
5. Neural classifier algorithm : LVQ3 with 8 output clusters (one cluster for each reference shape).
Training data set : same as 3.
6. Neural classifier algorithm : LVQ1 with 16 output cluster (two clusters for each reference shape).
Training data set : same as 3.
7. Neural classifier algorithm : LVQ2 with 16 output cluster (two clusters for each reference shape).
Training data set : same as 3.
8. Neural classifier algorithm : LVQ3 with 16 output clusters (two clusters for each reference shape).
Training data set : same as 3.

Under each of eight different simulation environments, 11808 of total test patterns (1476 patterns for each reference image) were evaluated. During training procedure, the only 120 ~ 170 epochs were needed for the weight vectors to be stable under simulation environments 2 - 8. It means LVQ whose weights were initialized with some of known training patterns is fast to learn the incoming patterns. In simulation 1, the initial weight vectors for 8 output clusters were taken from 8 training vectors. Thus training procedure was not necessary in simulation 1. The classification results under eight different simulation environments are summarized in the table 1 and 2.

In simulation 1, the weights for 8 output clusters were initialized and fixed by 8 training patterns. Thus the results from simulation 1 were made by measuring the Euclidean distance between the weight vectors and the incoming test pattern. The correct classification ratios for the 10db noisy patterns were increased by using the improved LVQ algorithm (LVQ2, LVQ3) and selecting some portion of training patterns from noisy corrupted images. They are shown in the results from simulation 2, 3, 4 and 5. And with the increasing number of output cluster for each pattern, the higher correct classification ratios were obtained in this experimental process (The

Table 2. Classification results with 10db SNR test patterns under simulation environments 1 through 8. (Classification results with noise-free, 25db, 20db, and 15db SNR under simulation environment 1-8 were exactly same as in table 1.)

	Classification results with 10db SNR	Total num. of correctly classified patterns (%)
simulation 1	2765 / 2880, (96.01%)	11693 / 11808, (99.03%)
simulation 2	2776 / 2880, (96.39%)	11704 / 11808, (99.12%)
simulation 3	2840 / 2880, (98.61%)	11768 / 11808, (99.66%)
simulation 4	2845 / 2880, (98.78%)	11773 / 11808, (99.70%)
simulation 5	2847 / 2880, (98.85%)	11775 / 11808, (99.72%)
simulation 6	2871 / 2880, (99.69%)	11799 / 11808, (99.92%)
simulation 7	2871 / 2880, (99.69%)	11799 / 11808, (99.92%)
simulation 8	2871 / 2880, (99.69%)	11799 / 11808, (99.92%)

Table 1. Classification results under simulation environment 1. (The numbers in tabels represent num. of correctly classified patterns / num. of test patterns.)

Noise-level / Aircraft image	Noise-free	25db SNR	20db SNR	15db SNR	10db SNR	Total num. of correctly classified patterns (%)
a1	36 / 36	360 / 360	360 / 360	360 / 360	360 / 360	1476 / 1476, (100 %)
a2	36 / 36	360 / 360	360 / 360	360 / 360	336 / 360	1452 / 1476, (98.37 %)
a3	36 / 36	360 / 360	360 / 360	360 / 360	360 / 360	1476 / 1476, (100 %)
a4	36 / 36	360 / 360	360 / 360	360 / 360	360 / 360	1476 / 1476, (100 %)
a5	36 / 36	360 / 360	360 / 360	360 / 360	360 / 360	1476 / 1476, (100 %)
a6	36 / 36	360 / 360	360 / 360	360 / 360	360 / 360	1476 / 1476, (100 %)
a7	36 / 36	360 / 360	360 / 360	360 / 360	269 / 360	1385 / 1476, (93.83 %)
a8	36 / 36	360 / 360	360 / 360	360 / 360	360 / 360	1476 / 1476, (100 %)
Total num. of correctly classified patterns (%)	288 / 288 (100%)	2880 / 2880 (100%)	2880 / 2880 (100%)	2880 / 2880 (100%)	2765 / 2880 (96.01%)	11693 / 11808 (99.03 %)

classification results from simulation 6 - 8 are higher than the results from simulation 2 - 4). Under simulation environments 6, 7 and 8, only nine patterns were misclassified among 11808 test patterns and the classification ratios were reached to almost 100%. From table 1, the overall classification results from simulation 1 through 8 were over 99 %. It means the bispectral feature vectors with LVQ neural classifier have the acceptable discrimination ability between each aircraft pattern even the patterns were significantly corrupted by noise.

V. Conclusion

The fine classification results in this paper show that the LVQ neural classifier, trained with the fifteen bispectral feature vectors extracted from the normalized contour sequences of planar shapes, performs well to recognize the different shapes of aircrafts even the aircraft images are rotated, scaled and corrupted by noise. The overall processing steps in this system is relatively simple ; extract the bispectral feature vectors from contour sequences of image by FFT algorithm and then input these vectors to LVQ neural classifier. The training process for LVQ is fast by initializing weight vectors with training patterns.

The possible areas for further research could involve the investigation of feature extractions of a closed planar image which are insensitive to the random variation of the image and contain more detailed shape information for the characterization of the certain image pattern while still maintaining relatively small dimensionality for the inputs to the neural classifier. Also, the noise reduction process for significantly noisy corrupted image and the way to segment the object image from background for applications to real world environments should be investigated.

References

[1] C.C. Lin and R. Chellappa, "Classification of partial 2-D shapes using Fourier descriptors, "*IEEE Trans. on PAMI*, Vol. 9, No. 5, pp. 686-690, Sep., 1987.

[2] S.R. Dubois and F.H. Glanz, "Autoregressive Model Approach to Two-Dimensional Shape Classification," *IEEE Trans. on PAMI*, Vol. 8, No. 1, Jan., 1986

[3] I. Sekita, T. Kurita, and N. Otsu, "Complex Autoregressive Model for Shape Recognition," *IEEE Trans. on PAMI*, Vol. 14, No. 4, April

1992.
[4] M. Das, M. J. Paulik, and N. K. Loh, "A Bivariate Autoregressive Modeling Technique for Analysis and Classification of Planar Shapes," *IEEE Trans. on PAMI*, Vol. 12, No. 1, Jan., 1990.
[5] L. Gupta and M.D. Srinath, "Invariant Planar Shape Recognition Using Dynamic Alignment," *Pattern Recognition*, Vol.21, No. 3, pp. 235-239, 1988.
[6] L. Gupta and M.R.Sayeh and R. Tammana, "A neural network approach to robust shape classification," *Pattern Recognition*, Vol. 23, No. 6, pp. 563-568, 1990.
[7] 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.
[8] C.L. Nikias and M.R. Raghuveer, "Bispectrum estimation : A digital signal processing framework," *Proc. IEEE*, Vol. 75, No.7, pp.869-891, July 1987.
[9] T. Erdem and M. Tekalp, "Linear Bispectrum of Signals and Identification of Nonminimum Phase FIR Systems Driven by Colored Input," *IEEE Trans. Signal Processing*, Vol 40, No. 6, pp. 1469-1479,1992.
[10] I. Jouny, "Description of radar targets using the bispectrum," *IEEE Proc.-Radar, Sonar Navig.*, Vol. 141, No. 3, pp.159-163, 1994.
[11] E.K. Walton and I. Jouny, "Bispectral analysis of radar signature and application to target classification," *Radio Sci.*, Vol. 25, pp.101-113, 1990.
[12] T. Kohonen, *Self-organization and associative Memory (3rd ed.)*, Berlin : Springer-Verlag, 1989
[13] T. Kohogen, "Learning Vector Quantization," *Helsinki University of Technology, Lab. of Computer and Information Science*, Report TKK-F-A-601, 1986.
[14] L. Fausett, *Fundamentals of Neural Networks : Architectures, Alogrithm, and Applications*, Prentice Hall, 1994
[15] R.C. Gonzalez adn R.E. Woods, *Digital Image Processing*, Addison-Wesley Publishing Company, Inc., 1992.
[16] W.K. Patt, *Digital Image Processing, 2nd ed.*, Wiley Interscience, 1991.
[17] C.L. Nikias and J.M. Mendel, *Signal Processing with Higher-Order Spectra*, United Signals & Systems, Inc., 1990.
[18] J.K. Tugnait, "Identification of nonminimum phase linear stochastic system," *Automatica*, Vol. 22, pp.457-464, July 1986.
[19] C.L. Nikias, "ARMA bispectrum approach to nonminium phase system identification, " *IEEE Trans. Acoust., Speech, Signal Processing*, Vol.36, pp.513-524, April 1988.
[20] A.V. Oppenheim and R.W. Schafer, *Digital Signal Processing*, Prentice-Hall, N.J., 1975.
[21] T. Kohonen, "Learning Vetor Quantization," *Neural Networks*, Vol. 1, suppl. 1, pp. 303, 1988.
[22] T. Kohonen, "Improved Versions of Learning Vector Quantization," *International Joint Conference on Neural Networks*, San Diego, CA., U.S.A. pp. 545-550, 1990.
[23] T. Kohonen, "The Self-Organizing Map," *Proceedings fo the IEEE*, 78(9), pp. 1464-1480, 1990.