Orientation Selectivity of TAM Network with Extensive Receptive Field

Isao Hayashi
Department of Informatics,
Kansai University
2-1-1, Ryozenji-cho, Takatsuki,
Osaka 569-1095, Japan
ihaya@kcn.res.kutc.kansai-u.ac.jp

James R. Williamson
Theater Missile Defense,
Lockheed Martin Corp.
MIT, Lexington,
MA 02420, U.S.A.
jrw@ll.mit.edu

***Abstract*— TAM (Topographic Attentive Mapping) network is a biologically-motivated neural network with Gabor function type receptive fields. However, the structure of receptive fields is a mono-layer, and there is a lack of performance for rotating images. In this paper, we formulate a new TAM network with multilayer structure of extensive receptive fields. We also show the usefulness of TAM network using some examples of character recognition.**

I. Introduction

In the human visual system, the visual modalities, e.g., image's shape, orientation, color and brightness are detected at the primary visual cortex (V1) through the lateral geniculate nucleus (LGN). In the higher visual cortex, object's rotation and symmetry are detected using the orientation selectivity[1]. In the retina, the receptive field by simple cell, complex cell and hypercomplex cell play important roles for detecting contours of objects. As a model representing the receptive field, Marčelja[2] defined Gabor function using a complex sinusoidal plane wave of some frequency and orientation. Daugman[3] extended the Gabor function to two-dimensional model. Other many receptive field models [4], [5], [6] have been also proposed.

On the other hand, many powerful neural networks [7], [8], [9], [10], [11], [12] based on Hubel-Wiesel's architecture have been proposed. We have already formulated TAM network as a visual neural network, which incorporates Gabor function type receptive field in the input architecture of the neural network. The TAM (Topographic Attentive Mapping) network has been proposed by Williamson[11], where four layers, the feature layer, the basis layer, the category layer and the class layer, represents the input-output relationship between the distributed input data and the categorized output classes, after learning of TAM network when the supervised data sets is obtained. Our TAM network is an advanced model of the original TAM network, which has six layers including retina and LGN layers adding them to the input layer of the original TAM network[13], [14], [15], [16]. Therefore, the TAM network should be a biologically-motivated neural network. However, the performance of TAM network for object's rotation is not so well since the receptive field in the retina architecture is single layer, and a mechanism of orientation selectivity covering wider visual area doesn't exist in the receptive field layer.

In this paper, we formulate two layers' structure at the receptive field architecture, which are the filtering process at the first layer detects the sharp orientation selectivity using the narrow space of Gabor function as well as simple cell, and the filtering process at the second layer obtains the whole rotation selectivity using more extensive receptive field as well as complex cell. In our model, first, the sixteen orientation selectivity of the original object's contour is detected using narrow space of Gabor function. Next, the total orientation selectivity for more wider contour image is detected using wider space of Gabor function, and the rotation selectivity of objects is last calculated. The feature map of TAM network is compound from density values of the sixteen orientations in the first layer. At the upper layers of TAM network, the input-output relationship of data sets is acquired by two kind of signal streams, which are the bottom-up stream from feature map to the class layer, and the top-down learning to category layer from class layer, when the supervised data sets are obtained. We discuss here the mechanism of the rotation selectivity using the extensive receptive field, and show the usefulness of the TAM network through some examples of character recognition.

II. TAM Network with Extensive Receptive Field

The structure of TAM network with receptive field is shown in Figure 1. Sixteen orientation selectivity of the original images is detected using Gabor filtering in the first retina layer, and the rotation selectivity of the contour images is calculated at the second retina layer. The feature map is compound from density values of the sixteen orientations in the first layer. The object images are recognized by learning mechanism of the TAM network at the upper feature maps, shown in Figure 2.

The two-dimensional Gabor function type receptive field, $G(x, y)$, is represented as follows:

$$\begin{aligned} G(x,y) &= Ke^{-\frac{1}{2}\left(\frac{(x-\mu x)^2}{\sigma_x^2}+\frac{(y-\mu y)^2}{\sigma_y^2}\right)} \\ &\quad \times \sin\left(2\pi f_x x\cos\theta + 2\pi f_y y \sin\theta + \phi\right) \end{aligned}$$

where, K is amplitude, (μ_x, μ_y) is the center coordinate of Gabor function, σ_x and σ_y are standard deviations of axis X and Y, respectively, and f_x and f_y are frequencies of axis X and Y, respectively. Figure 3 shows an example of Gabor function.

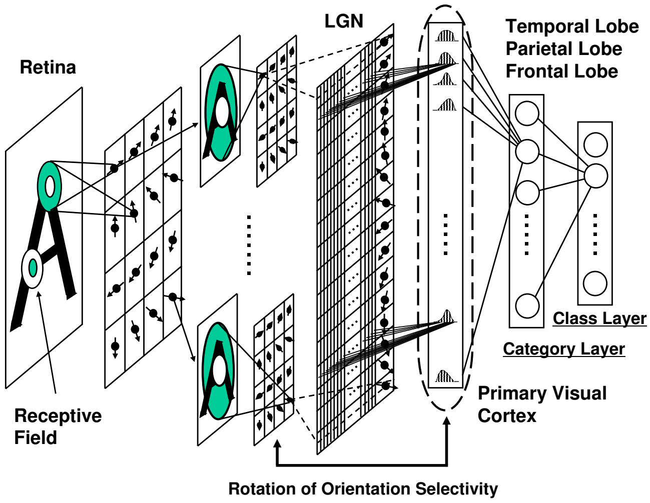

Fig. 1. TAM Network with Extensive Receptive Field

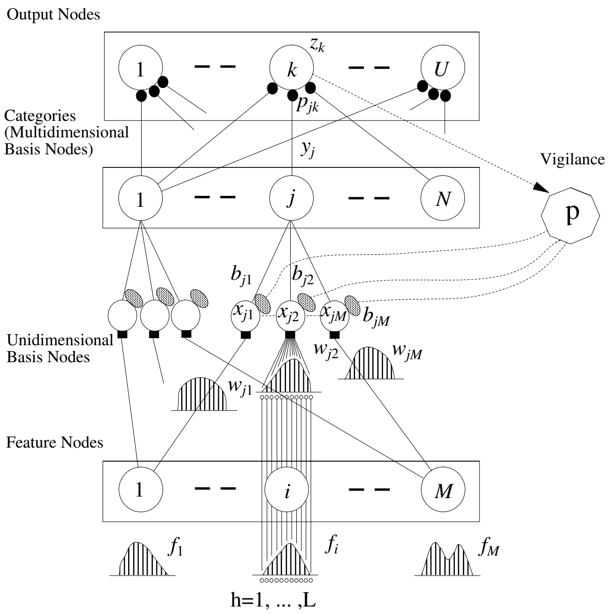

Fig. 2. Structure at the Upper Feature Maps

The process for detecting orientation selectivity is shown in Figure 4. Let the horizontal scale and vertical scale of the original image denote as R_H $pixels$ and R_V $pixels$, respectively. In the first retinal layer, the i-th orientation selectivity of contour image, $C_i^1(x,y)$, $i = 1, 2, \cdots 16$, for

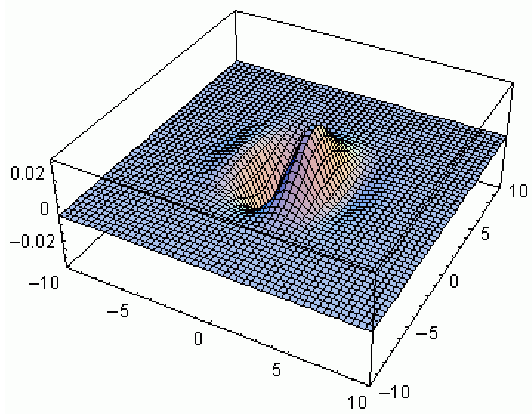

Fig. 3. Gabor Function

the original image, $O(p,q)$, $0 \le p \le R_H$, $0 \le q \le R_V$, is calculated by the following convolution process:

$$C_i^1(x,y) = \sum_{q=1}^{R_V} \sum_{p=1}^{R_H} G_i(x-p, y-q) \times O(p,q).$$

After the I-th orientation selectivity for contour images is calculated at the first retina layer, the I-th orientation selectivity, $C_I^2(x,y)$, at the second layer is calculated using more extensive receptive field again:

$$C_I^2(x,y) = \sum_{q=1}^{R_V} \sum_{p=1}^{R_H} G_I(x-p, y-q) \times C_I^1(p,q). \qquad (1)$$

The i-th feature map, f_{ih}, at the feature layer of TAM network is calculated by the following normalization:

$$f_{ih} = \frac{\sum_{\{x,y|O_i(x,y)=h\}} C_i^1(x,y)}{\sum_{y=1}^{R_V} \sum_{x=1}^{R_H} C_i^1(x,y)}, \quad h = 1, 2, \cdots, 256.$$

We should notice that a couple set of the feature map and the rotation selectivity by Equation (1) is inputted to the feature layer of TAM network.

We utilize here the rotation selectivity to estimate the rotation angle of checking image comparing to the training image in TAM network. When an orientation selectivity, $C_{i,CHD}^1(x,y)$, of the checking image is obtained, the feature map, f_{ih}, of the checking image then is exchanged itself by the minimized rotation angle, r, comparing with the orientation selectivity of the training image $C_{I,TRD}^2(x,y)$.

$$r = \min_i \left(\frac{\sum_x \sum_y C_{i,CHD}^2(x,y)}{R_H \times R_V} - \frac{\sum_x \sum_y C_{I,TRD}^2(x,y)^2}{R_H \times R_V} \right)$$

At the upper feature layers of TAM network, the activities, x_{ji}, $i = 1, 2, \cdots, M$, $j = 1, 2, \cdots, N$, are calculated in the basis layer by the following equation with feature maps, f_{ih}, and node's weights, w_{jih}, $h = 1, 2, \cdots, L$. The output signal of the category layer, y_j, is then calculated by products with

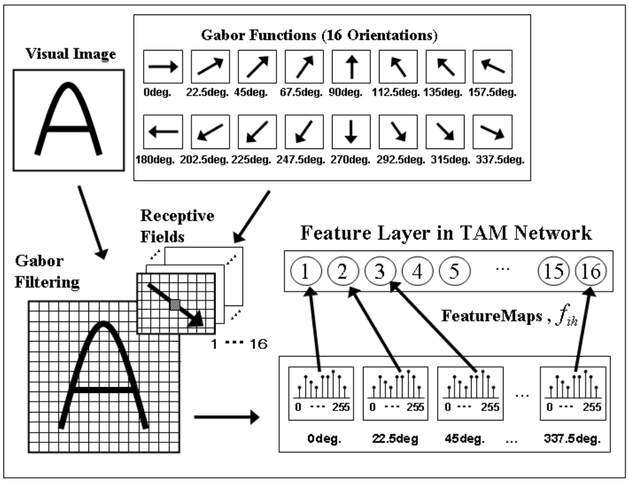

Fig. 4. Gabor Filtering Process

activities x_{ji}.

$$
\begin{aligned}
x_{ji} &= \frac{\sum_{h=1}^{L} f_{ih} w_{jih}}{1+\rho^2 b_{ji}} \\
y_j &= \prod_{i=1}^{M} x_{ji}
\end{aligned}
$$

where, ρ represents the vigilance parameter, and b_{ji} are inhibitory weights.

The output prediction, K, at the class layer is calculated as follows:

$$
\begin{aligned}
z_k &= \sum_{j=1}^{N} y_j p_{jk} \\
K &= \{k | \max_k z_k\}
\end{aligned}
$$

where, z_k are outputs of the k-th class layer, and p_{jk} are weighted connections.

TAM network has a mechanism of adding a new node to category layer when a deviation between the output of class layer and the correct supervised output exists. Let K^* denote the index of the "correct" supervised output class. In the case of $K \neq K^*$, the vigilance parameter, ρ, is rising from the initial value, ρ^{init}, until either $\rho \geq \rho^{max}$ or $z_{K^*}/z_K \geq OC$, where OC is the maximal vigilance level. When the vigilance parameter reaches the maximal level, ρ^{max}, a node is added to the category layer.

On the other hand, when the subject of $z_{K^*}/z_K \geq OC$ is satisfied, the following feedback signal y_j^* is calculated, and the learning parameters, w_{jih}, p_{jk}, and b_{ji}, are renewed.

$$
z_k^* = \begin{cases} 1 & ; \text{ if } k = K^* \\ 0 & ; \text{ otherwise} \end{cases}
$$

$$
\begin{aligned}
y_j^* &= \frac{\prod_{i=1}^{M} x_{ji} \times \sum_{k=1}^{U} z_k^* p_{jk}}{\sum_{j'=1}^{N} \prod_{i=1}^{M} x_{j'i} \times \sum_{k=1}^{U} z_k^* p_{j'k}} \\
\triangle w_{jih} &= \frac{\alpha y_j^* (1-\lambda^{1/M})(f_{ih} - w_{jih})}{(\alpha - 1)\lambda^{1/M} + n_j}, \quad \lambda \in (0,1) \\
\triangle p_{jk} &= \frac{\alpha y_j^* (z_k^* - p_{jk})}{\alpha + n_j} \\
\triangle b_{ji} &= b_j^{(rate)} y_j^* (x_{ji} - b_{ji}) \\
\triangle n_j &= \alpha y_j^* (1 - n_j)
\end{aligned}
$$

where, α, λ are parameters, and $b_j^{(rate)}$ are learning parameters.

III. Character Recognitions

In order to show the usefulness of TAM network with extensive receptive field, some examples of character recognition are illustrated here. As an input image, all 26 alphabets are filled in the electronic pad whose size is $15\ pixels \times 15\ pixels$. First, Gabor filtering makes a contour images for the alphabets, and feature maps with rotation selectivity of Equation (1) are calculated. Next, the training data set between these feature maps and alphabet's classes is constructed, and TAM network acquires the input-output relationship of the training data set by the learning mechanism. After terminated the learning of TAM network, the TAM network estimates alphabet class of the obtained checking images. The following three kinds of experiments are discussed here.

1) Calculating recognition rates of all 26 kinds of alphabets using the original TAM network without estimating the rotation.
2) Calculating recognition rates of characters using the original TAM network when the alphabets are rotated, and declined.
3) Calculating recognition rates of rotation characters using a new TAM network with extensive receptive field.

The learning parameters are set as follows:

$$
\begin{array}{lllllll}
L & = & 255 & & \rho^{init} & = & 0.0 \\
OC & = & 0.8 & & \rho^{step} & = & 0.1 \\
\alpha & = & 0.01 & & \rho^{max} & = & 100.0 \\
\lambda & = & 0.33 & & b_j^{(rate)} & = & 0.0000001 \\
\mu_x & = & 0.0 & & \mu_y & = & 0.0 \\
\sigma_x & = & 1.99 & & \sigma_y & = & 1.92 \\
f_x & = & 0.127 & & f_y & = & 0.127 \\
\phi & = & 90.0 & & K & = & 1.0.
\end{array}
$$

First, let show the performance of the original TAM network without estimating rotation. The training images are shown in Figure 5. The contour images of the left upper side alphabet 'A' in Figure 5 are shown in Figure 6. In the Figure 6, the left-upper side image indicates 0°, the right side of the image of 0° shows 22.5°, and the next image shows 45°, 67.5°, $\cdots$, 337.5°, respectively. From those orientation images, we should notice that Gabor filtering certainly detects the orientation selectivity.

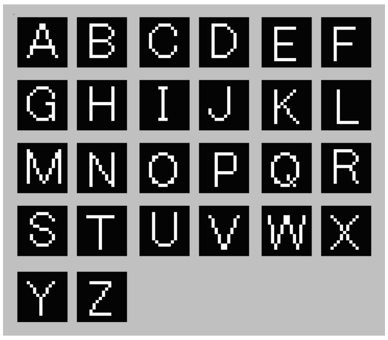

Fig. 5. Training Images

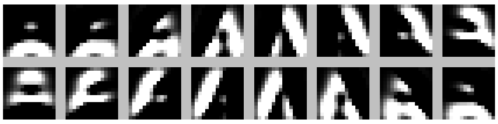

Fig. 6. Orientation Selectivity after Gabor Filtering

The resultant recognition rates of all 26 checking alphabets are shown in Figure 7. The recognition rates of both cases of two epochs and five epochs are compared, and the recognition rate is calculated as the average of 30 trials. In the case of the five epocks, the almost recognition rate is over the sixty percentages.

Next, we discuss recognition rates when the characters are rotated and declined. The training images are shown in Figure 5, and the checking images are shown in Figure 8. In Figure 8, three types of rotating charecters are prepared by rotating 90°, 180°, and 270°, and four types of declining characters are prepared. The resultant recognition rates are shown in Table I. The almost recognition rates are zero percentage, i.e., TAM network couldn't recognize the rotating and declining characters. However, the alphabets, 'I' and 'L', have relatively robustness for rotating 180° and 270°. The alphabets, 'N', 'O', 'X', and 'Z', are strongly robust for rotating 180°. Since these alphabets consist of straight lines and simple connections, the discussion should be understandable. We suppose that the result comes from the mono-layer in the TAM network, i.e., there is a lack of performance for rotation mechanism.

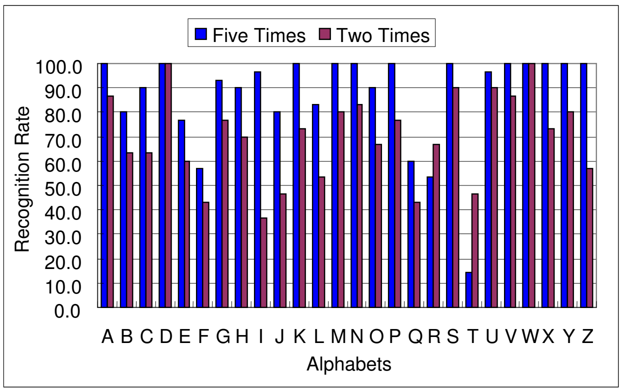

Fig. 7. Results of All Alphabets

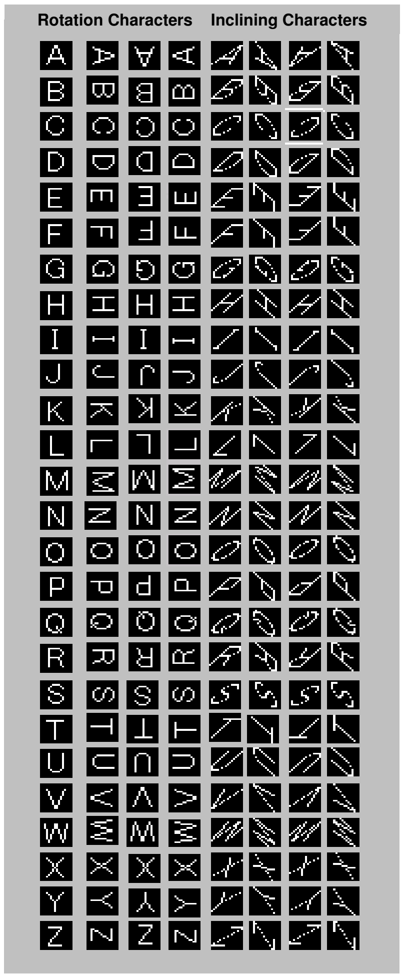

Fig. 8. Checking Images of Rotation and Declining Characters

We discuss last the performance of TAM network with extensive receptive field for two kinds of alphabets, 'A' and 'B', as the training images. The training images are shown in Figure 9. The checking images are shown in Figure 10. The resultant recognition rates are shown in Table II. The recognition rate for each character is shown in Figure 10. In the Figure 10, the upper value under each image represents recognition rate of TAM network without rotation mechanism, and the lower value represents the recognition rate of TAM network with rotation mechanism. The recognition rate is calculated as the average of 30 trials. In the Table II and Figure 10, we should notice that some recognition rate of 'B' by the new TAM network is better than the original TAM network. The other recognition rates are still extremely very

TABLE I

RECOGNITION RATES OF ROTATION AND DECLINING

Alpha.	Recognition Rate (%)							
	Rotation				Declining			
	0°	90°	180°	270°	45°	135°	225°	315°
A	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
B	80.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C	90.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
D	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
E	76.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
F	56.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
G	93.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H	90.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
I	96.7	60.0	96.7	73.3	20.0	0.0	80.0	3.3
J	80.0	0.0	0.0	0.0	6.7	0.0	0.0	0.0
K	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
L	83.3	0.0	20.0	50.0	6.7	10.0	10.0	10.0
M	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
N	100.0	0.0	93.3	0.0	0.0	0.0	0.0	0.0
O	90.0	0.0	80.0	0.0	0.0	0.0	0.0	0.0
P	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Q	60.0	0.0	20.0	0.0	0.0	0.0	0.0	0.0
R	53.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
S	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
T	100.0	3.3	0.0	0.0	3.3	0.0	13.3	3.3
U	96.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
V	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
W	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
X	100.0	13.3	100.0	0.0	0.0	0.0	0.0	0.0
Y	100.0	0.0	0.0	16.7	0.0	30.0	0.0	0.0
Z	100.0	0.0	93.3	16.7	16.7	0.0	80.0	76.7

low. In order to improve the recognition rate more, we should precisely adjust the parameters of Gabor function, and discuss how to modify the structure of TAM network to be better for rotating images. Additionaly, we recognize only two types of images of 'A' and 'B' is too small numbers of experiments. We should conduct more experiments and confirm a usefulness of new TAM network from their results. However, we found a new benefit of visual models in adding it to wider receptive fields and feedback route in retina structure.

TABLE II

RECOGNITION RATE

	Char.'A' (%)	Char.'B' (%)	Total (%)
TAM without Rotation	96.04	12.71	54.38
TAM with Rotation	98.33	30.42	64.38

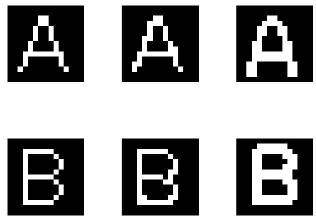

Fig. 9. Training Images of 'A' and 'B'

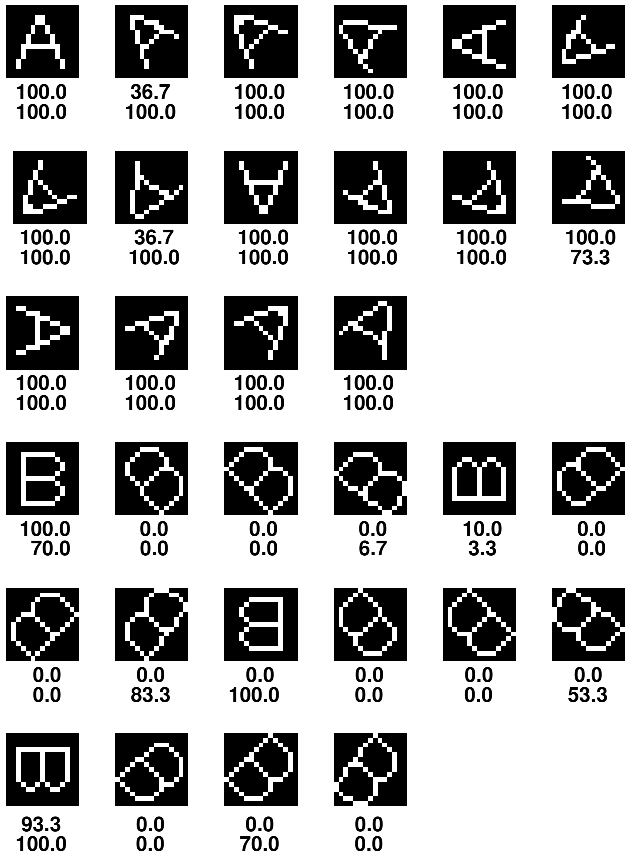

Fig. 10. Checking Images

IV. CONCLUSIONS

We formulated here a new TAM network with extensive receptive field, and discuss the usefulness of TAM network through some examples of character recognition.

REFERENCES

[1] S.Yantis, "Visual Perception: Essential Readings," East Sussex, UK: Psychology Press, 2000.

[2] S.Marčelja, "Mathematical description of the responses of simple cortical cells," Optical Society of America, Vol.70, No.11, pp.1297-1300, 1980.

[3] J.Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," Optical Society of America, Vol.2, No.7, pp.1160-1169, 1985.

[4] A.D.Pollen and S.F.Ronner, "Visual cortical neurons as localized spatial frequency filters," IEEE Transactions of System, Man and Cybernetics, Vol.SMC13, pp.907-916, 1983.

[5] D.C.Lee, "Adaptive processing for feature extraction: Application of two-dimensional Gabor function," Remote Sensing, Vol.17, No.4, pp.319-334, 2001.

[6] K.Okajima, "Two-dimensional Gabor-type RF as derived by mutual information maximization," Neural Networks, Vol.11, pp.441-447, 1998.

[7] S.Grossberg, "How does the cerebral cortex work? Learning, attention, and grouping by the laminar circuits of visual cortex," Spatial Vision, Vol.12, No.2, pp.163-185, 1999.

[8] H.Neumann and W.Sepp, "Recurrent V1-V2 interaction in early visual

boundary processing," Biological Cybernetics, Vol.81, pp.425-444, 1999.
[9] K.Fukushima, "Neocognitron: A hierarchical neural network capable of visual pattern recognition," Neural Networks, Vol.1, No.2, pp.119-130, 1988.
[10] K.Fukushima, "Recognition of partly occluded patterns: a neural network model," Biological Cybernetics, Vol.84, No.4, pp.251-259, 2001.
[11] J.R.Williamson, "Self-organization of topographic mixture networks using attentional feedback," Neural Computation, Vol.13, pp.563-593, 2001.
[12] K.Fukushima, "Use of non-uniform spatial blur for image comparison: Symmetry axis extraction," Neural Networks, Vol.18, No.1, pp.23-32, 2005.
[13] I.Hayashi and J.R.Williamson, "Acquisition of fuzzy knowledge from topographic mixture networks with attentional feedback," Proceedings of the International Joint Conference on Neural Networks (IJCNN'01), pp.1386-1391, 2001.
[14] I.Hayashi and H.Maeda, "A formulation of fuzzy TAM network with Gabor type receptive fields," Proceedings of the 4th International Symposium on Advanced Intelligent Systems (ISIS2003), pp.620-623, 2003.
[15] I.Hayashi, H.Maeda and J.R.Williamson, "A formulation of receptive filed type input layer for TAM network using Gabor function," Proceedings of 2004 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE2004), No.1335, 2004.
[16] I.Hayashi and J.R.Williamson, "Orientation selectivity by TAM network using Gabor function type receptive filed," Proceedings of the 5th International Conference on Recent Advances in Soft Computing (RASC2004), pp.122-127, 2004.