FUZZ-IEEE2004 and IJCNN2004, No.1335
Budapest, Hungary, July 25-29, 2004

A Formulation of Receptive Field Type Input Layer for TAM Network Using Gabor Function

Isao Hayashi
Department of Informatics,
Kansai University
Takatsuki, Osaka 569-1095, Japan
E-mail: ihaya@res.kutc.kansai-u.ac.jp

Hiromasa Maeda
Information Science Research Center Corp.
Minatoku, Tokyo 108-0075, Japan
E-mail: CBC10360@nifty.com

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

***Abstract*— The TAM (Topographic Attentive Mapping) network is a biologically-motivated neural network. Gabor function is a model for receptive field. In this paper, we formulate a new input layer using Gabor function for TAM network to realize receptive field of human visual cortex.**

I. Introduction

Concerning signal processing in the human visual cortex, three kinds of cells which are simple cell, complex cell and hypercomplex cell, play important roles. Especially, a hypercomplex cell is powerful to detect contour orientation by slit type field. The slit is a narrow visual field from 0° to 5°. We call the field Receptive Field, which is represented by adjacent ON and OFF regions filled, respectively, with plus and minus signs. The sixteen orientations from 0° to 337.5° per 22.5° are detected there. Plenty models imitate the visual cortex have been proposed[1]-[6], and the receptive field models are also in [7]-[14]. An approach for modeling receptive field is Gabor function[10], [11]. Marčelja defined Gabor function as a oscillator which is a complex sinusoidal plane wave of some frequency and orientation within a Gaussian envelope and sine/cosine function. Daugman[11] extended it to two-dimensional format.

On the other hand, TAM (Topographic Attentive Mapping) network is a biologically-motivated neural network[6], [12]-[14]. The TAM network is composed of three layers where feature layer imitates the retina, category layer imitates the lateral geniculate nucleus and in the class layer, the output is given by the name of object grouping.

In this paper, we formulate receptive field type input layer for TAM network using Gabor filtering. The biological motivation for Gabor filtering to the TAM network lies in constructing like the receptive field in human visual cortex. In order to constructing receptive field type input layer, we define a shape of receptive field using Gabor function and formulate convolution process for Gabor filtering in the input layer. By the Gabor filtering, feature map is compound from density values at each receptive field for the sixteen orientations. The feature map is incorporated in the input layer of TAM network. We formulate here the receptive field type input layer in the TAM network, and show the usefulness of TAM network through some examples.

II. Gabor Function

Let x and y denotes axis on two-dimensions, respectively. The two-dimensional Gabor function, $G(x,y)$, is represented as follows;

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

where, K is amplitude, (μ_x, μ_y) is the center coordinate of Gabor function, σ_x and σ_y are standard deviations, and f_x and f_y are frequencies. In the case of $\phi = \pi/2$, the sine form is equal to the cosine form of Gabor function.

An example of Gabor function is shown in Figure 1, where $(\mu_x, \mu_y) = (0.0, 0.0)$, $\sigma_x = 2.0$, $\sigma_y = 2.0$, $f_x = 1.0\ Hz$, $f_y = 1.0\ Hz$, $\theta = 0^\circ$ and $\phi = 0.0$.

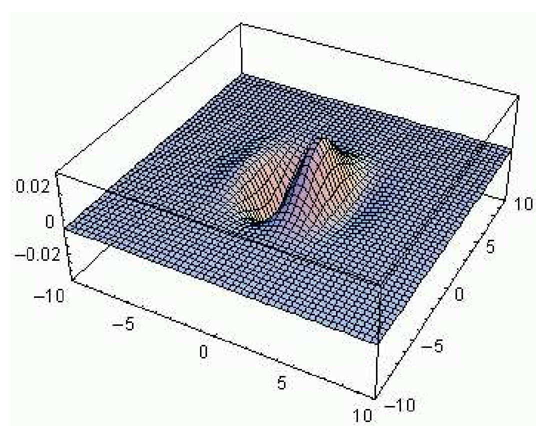

Fig. 1. Gabor Function

Gabor filtering is a well known method to detect the contour orientations using the Fourier transform of Gabor function. Given an image, the sixteen orientations of contour in the image are detected. Let the horizontal and vertical scale of the input image denote as R_H and R_V, respectively. In the discrete type Gabor filtering algorithm, the contour image, $C_i(x,y),\ i = 1, 2, \cdots, M,\ 0 \leq x \leq R_H,\ 0 \leq y \leq R_V$, for the i-th orientation is calculated by the following convolution process between the Gabor function and the original input image.

$$C_i(x,y) = \sum_{q=1}^{R_V}\sum_{p=1}^{R_H} G_i(x-p, y-q) \times I(p,q) \qquad (2)$$

where, G_i is the Gabor function for detecting the i-th orientation, and $I(p,q),\ 0 \leq p \leq R_H,\ 0 \leq q \leq R_V$ is the input image.

III. TAM Network with Receptive Field Type Input Layer

The structure of the TAM Network is shown in Figure 2. The Gabor filtering process is represented in Figure 3. The Gabor filtering is incorporated in the input layer of the TAM network. First, an image is clipped by the scale of $R_H\ pixel \times R_V\ pixel$. The contour images for sixteen orientations of the input image is next detected at each receptive field by the Gabor filtering. The density from 0 to 255, i.e., black color is 0 and white color is 255 in the PGM format, is then piled as the orientations of the contours. Since a receptive field is here shifting one pixel by one, the center coordinate position of receptive field is equal to a pixel location and the number of receptive fields is totally $R_H \times R_V$. Finally, $R_H \times R_V$ density values are normalized and feature map, f_{ih}, for the i-th feature in the TAM network is constructed as follows;

$$f_{ih} = \frac{\sum_{\{x,y|O_i(x,y)=h\}} C_i(x,y)}{\sum_{y=1}^{R_V}\sum_{x=1}^{R_H} C_i(x,y)}, \quad h = 1,2,\cdots,L \qquad (3)$$

where, $L = 256$.

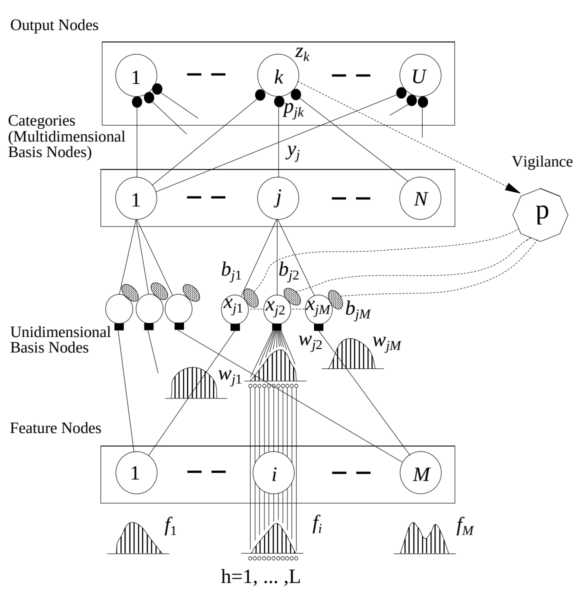

Fig. 2. TAM Network

When feature maps, $f_{ih},\ i = 1,2,\cdots,M$, are calculated, the output signal to the category layer, y_j, are calculated using

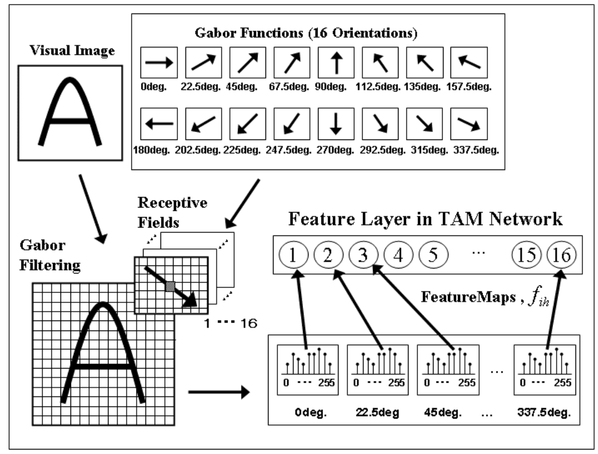

Fig. 3. Gabor Filtering Process

the node's weights, w_{jih}.

$$\begin{aligned} y_j &= \prod_{i=1}^{M} x_{ji} \\ &= \prod_{i=1}^{M} \frac{\sum_{h=1}^{L} f_{ih} w_{jih}}{1 + \rho^2 b_{ji}} \end{aligned} \qquad (4)$$

where x_{ji} are activities, ρ represents the vigilance parameter and b_{ji} are inhibitory weights. Since the sixteen orientations are defined here as the selectivity, we set $M = 16$.

The output prediction, K, is calculated as follows:

$$\begin{aligned} K &= \{k | \max_k z_k\} \\ &= \{k | \max_k \sum_{j=1}^{N} y_j p_{jk}\} \end{aligned} \qquad (5)$$

where, z_k are the output at each node of output layer and p_{jk} are weighted connections.

Let K^* denote the index of the "correct" supervised output class. If the network's output prediction K is not similar enough to K^*, we do $\rho = \rho + \rho^{(step)}$ until either $z_{K^*}/z_K \geq OC$ or $\rho \geq \rho^{(max)}$, where OC is the maximal vigilance level. Once the subject of $z_{K^*}/z_K \geq OC$ is satisfied, the feedback signal y_j^* is calculated for the learning step.

$$y_j^* = \frac{\prod_{i=1}^{M} x_{ji} \times \sum_{k=1}^{U} z_k^* p_{jk}}{\sum_{j'=1}^{N} \prod_{i=1}^{M} x_{j'i} \times \sum_{k=1}^{U} z_k^* p_{j'k}} \qquad (6)$$

$$z_k^* = 1 \text{ if } k = K^*; \quad z_k^* = 0 \text{ otherwise} \qquad (7)$$

The learning parameters, $w_{jih},\ p_{jk},\ b_{ji}$, are obtained as follows:

$$\triangle w_{jih} = \frac{\alpha y_j^* (1 - \lambda^{1/M})(f_{ih} - w_{jih})}{(\alpha - 1)\lambda^{1/M} + n_j}, \lambda \in (0,1) \qquad (8)$$

$$\triangle p_{jk} = \frac{\alpha y_j^* (z_k^* - p_{jk})}{\alpha + n_j} \qquad (9)$$

$$\triangle b_{ji} = b_j^{(rate)} y_j^* (x_{ji} - b_{ji}) \qquad (10)$$

$$\triangle n_j = \alpha y_j^*(1-n_j) \tag{11}$$

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

The algorithm of the TAM network including learning steps and pruning steps is represented as follows:

[Step 1] Given an image, the sixteen orientations of contours are detected by Gabor filtering.
[Step 2] The feature map, f_{ih}, is calculated from the density values of the input image.
[Step 3] The output prediction, K, is calculated.
[Step 4] If K is not similar enough to K^*, we do $\rho = \rho + \rho^{(step)}$. When ρ reaches the maximal level, one node is added to categories.
[Step 5] If $z_{K^*}/z_K \geq OC$, the learning step starts. Parameters, w_{jih}, p_{jk} and b_{ji}, are updated.
[Step 6] Until $z_{K^*}/z_K \geq OC$, let the algorithm repeat from step 1 to step 3.
[Step 7] After learning, the pruning step starts. The data set in which $f_{si},\ s = 1, 2, \cdots, R$ is divided into learning data and checking data. The information entropy, $H(i)$, is calculated using the learning data for feature selections, where ψ_k is a set of the data of the class k.

$$H(i) = -\sum_{j=1}^{N} g_j \sum_{k=1}^{U} G_{jk} \log_2 G_{jk} \tag{12}$$

$$g_j = \frac{\sum_{s=1}^{R} x_{jis}}{\sum_{j=1}^{N}\sum_{s=1}^{R} x_{jis}} \tag{13}$$

$$G_{jk} = \frac{\sum_{s\in\psi_k} \gamma_{js} \times p_{jk}}{\sum_{s=1}^{R} \gamma_{js} \times p_{jk}} \tag{14}$$

$$\gamma_{js} = \prod_{i\in I^*} x_{jis} \times x_{jis} \tag{15}$$

[Step 8] The following feature i^* is extracted as an important feature and we set $I^* = \{i^*\}$.

$$i^* = \{i | \max_i H(i)\} \tag{16}$$

[Step 9] If the following condition is satisfied for checking data at a category j, the link connections between j and outputs $k',\ k' = 1, 2, \cdots, U,\ k' \neq k$, are removed. Simultaneously, the connections between j and features $i' \notin I^*$, are removed, where η is a threshold.

$$G_{jk} \geq \eta \tag{17}$$

[Step 10] If the following condition is satisfied for checking data at the category j, the link connections between j and i, and $i' \notin I^*$, are removed, where θ is a threshold.

$$\frac{1}{R}\sum_{s=1}^{R} \gamma_{js} < \theta \tag{18}$$

[Step 11] If the following condition is satisfied for checking data at K, the link connections between K and categories, $j',\ j' = 1, 2, \cdots, N,\ j' \neq j$, are removed, where ξ is a threshold.

$$\varphi_{jK} = \frac{\sum_{s\in\Gamma_K} \gamma_{js} \times p_{jK}}{\sum_{j=1}^{N}\sum_{s\in\Gamma_K} \gamma_{js} \times p_{jK}} \geq \xi \tag{19}$$

[Step 12] When a category has lost connections to all outputs or features, the category is removed. Any output and feature which has been disconnected from all categories is also removed.
[Step 13] Until all features are selected at step 6, let the algorithm repeat from step 7 to step 12.

When the algorithm is terminated, the neural network whose needless connections and nodes are pruned is obtained. We should notice that the algorithm is a kind of fuzzy tuning methods since the data procedure is the same as that of fuzzy logic. Thus, we can acquire fuzzy rules from the TAM network as a knowledge representation.

IV. Character Recognitions

In order to show the efficiency of the receptive field type input layer of the TAM network, some examples are here illustrated. The alphabets, “A and B”, are filled in the electronic pad corresponding to the input image. The size of the electronic pad is $15\ pixels \times\ 15\ pixels$. The normal alphabets, “A and B” as training image are first filtered and translated to feature maps at the input layer. After the learning of the TAM network, the recognition rate of the checking images is calculated and estimated per changing parameters of the Gabor function. To estimate the usefulness of the TAM network, the following three views are pointed.

- Analyzing the orientation selectivity of the input images after Gabor filtering.
- Estimating the recognition rate after the learning of the TAM network.
- Comparing robustness of four types of checking images.

The training images are shown in Figure 4 and 5. The checking images are shown in Figure 6. The parameters of the TAM network are set as follows:

L	$=$	255	$rho\ init$	$=$	0.0
OC	$=$	0.8	$rho\ step$	$=$	0.1
α	$=$	0.0000001	$rho\ max$	$=$	100.0
λ	$=$	0.33	$b_j^{(rate)}$	$=$	0.01
η	$=$	0.8	θ	$=$	0.03
ξ	$=$	0.5	ϕ	$=$	0.0
μ_x	$=$	0.0	μ_y	$=$	0.0

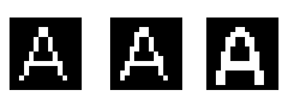

Fig. 4. Training Image of ‘A’

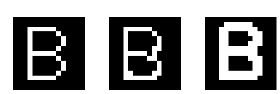

Fig. 5. Training Image of ‘B’

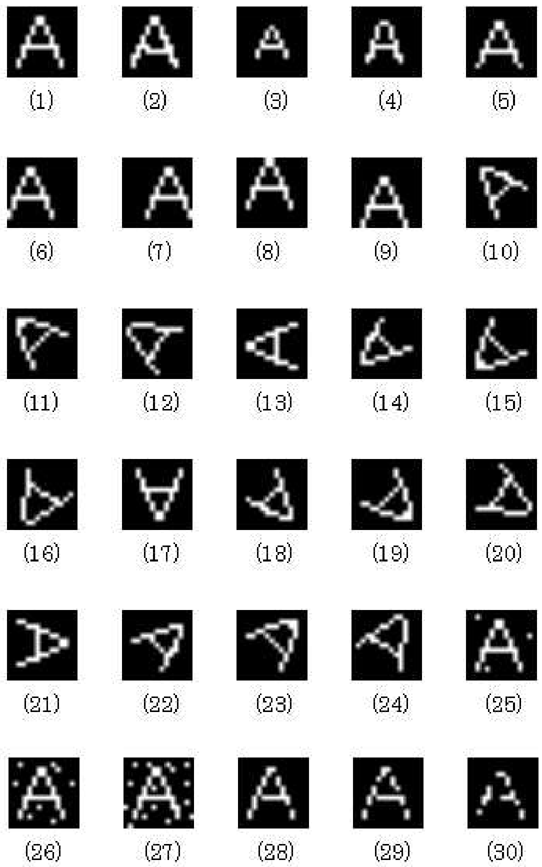

Fig. 6. Checking Image

The contour images of sixteen orientations are shown in Figure 7, which are the left side character 'A' in Figure 4. The orientation of the left-upper side indicates $0°$ and be moving the right side one by one to be $22.5°$, $45°$, $\cdots$, $337.5°$, respectively. We set here the parameters of Gabor function as $\sigma_x = 1.99$, $\sigma_y = 1.92$, $f_x = f_y = 0.127$, and $K = 1.0$. The contours according to each orientation are clearly detected.

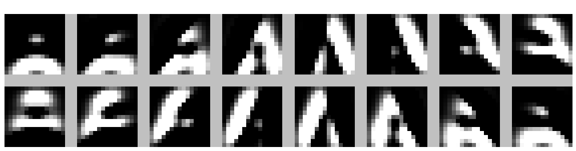

Fig. 7. Image after Gabor Filtering

For three points of views, the parameters of Gabor function are changing as follows;

$$\begin{aligned} \sigma_x &= 0.5, 0.75, \cdots, 4.0 \\ \sigma_y &= 0.5, 0.75, \cdots, 4.0 \\ f_x &= 0.01, 0.015, \cdots, 0.3 \\ f_y &= 0.01, 0.015, \cdots, 0.3 \\ K &= 0.1, 0.2, \cdots, 1.0. \end{aligned}$$

The contour images of the orientation, $135.0°$, for changing σ_x are shown in Figure 8. The Gabor functions are then shown in Figure 9. The value of σ_x of the left side image is 0.5 and be moving the right side one by one to be 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, respectively. Other parameters except for σ_x are set as $\sigma_y = 1.92$, $f_x = f_y = 0.127$, and $K = 1.0$. The larger σ_x is, the more vague the contour image is. However, the contour image at $\sigma_x = 0.5$ is not so correct. Therefore, we have to adjust the value of σ_x to get correct contour images. The orientation selectivity for changing σ_y and the amplitude, K, were sensitive as well as σ_x. However, the frequency, f_x and f_y, were not so sensible to orientation selectivity.

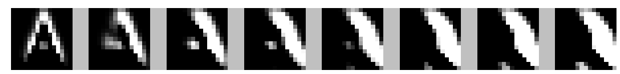

Fig. 8. Detecting of Contours of Input Images

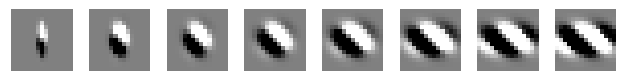

Fig. 9. Gabor Functions for Detecting Contours

Next, the recognition rates both of the training image and the checking image are shown in Figure 10 and in Table I. In Table I, TRI stands for the training image and CHI similarly stands for the checking image. The larger σ_x is, the lower the recognition rate of the checking image is. The recognition rate of the training image is constantly keeping in high. From Figure 8 and 10, we found that the contour orientation can be sharply detected after adjusted the optimal value of σ_x at relatively small region. That is why the narrow Gabor function can detect sharply the contour orientation since the value of σ_x means the width of Gabor function. On the other hand, the large σ_x can detect the orientations of the whole image, but the contour becomes obscure.

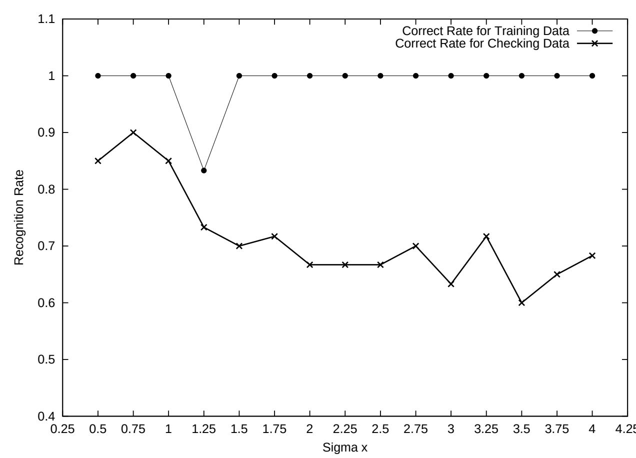

Fig. 10. Recognition Rate for Changing σ_x

Finally, the robustness of the TAM network is estimated using four types of the checking images. Figure 11 and 12 show the recognition rates for four types of checking images, which

TABLE I
RECOGNITION RATE AND CATEGORIES

σ_x	Recognition Rate of TRI (%)	Recognition Rate of CHI (%)	Nodes of Category
0.5	100.0	85.0	2
0.75	100.0	90.0	4
1.0	100.0	85.0	5
1.25	83.3	73.3	6
1.5	100.0	70.0	2
1.75	100.0	71.7	3
2.0	100.0	66.7	2
2.25	100.0	66.7	3
2.5	100.0	66.7	2
2.75	100.0	70.0	3
3.0	100.0	63.3	5
3.25	100.0	71.7	3
3.5	100.0	60.0	4
3.75	100.0	65.0	4
4.0	100.0	68.3	3

are shifting characters, rotations of characters, adding noise to characters, and reducing pixels in characters as divided in Figure 6. Figure 11 shows the recognition rates of image 'A' for changing σ_x and Figure 12 shows the recognition rates of image 'B' for changing f_y. The recognition rates of shifting and rotation of characters among four types are almost higher than 70.0%. In Figure 11, the larger σ_x is, the lower the recognition rate of characters added noise is. In Figure 12, the recognition rates both of characters added noise and reduced pixels are much low. Therefore, we found the Gabor filtering is not so efficient for noise images.

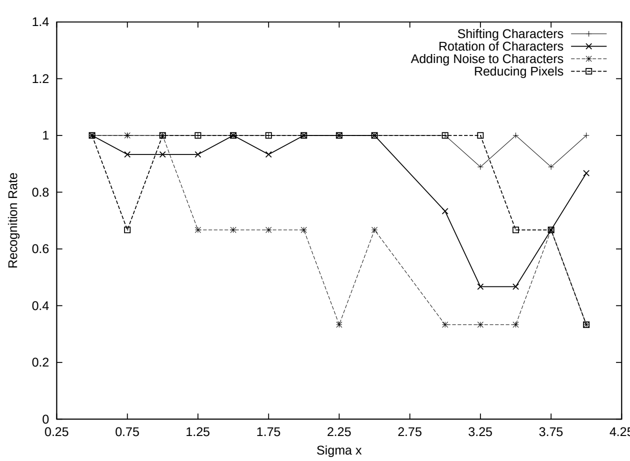

Fig. 11. Recognition Rate of Four Types of Image 'A' for Changing σ_x

V. CONCLUSIONS

We formulated here the TAM network with receptive field type input layer and showed the usefulness through character recognitions. However, we need more experiments and discussion in order to find characteristic of the Gabor filtering.

This research is partially supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan under Grant-in-Aid for Scientific Research number 14580433.

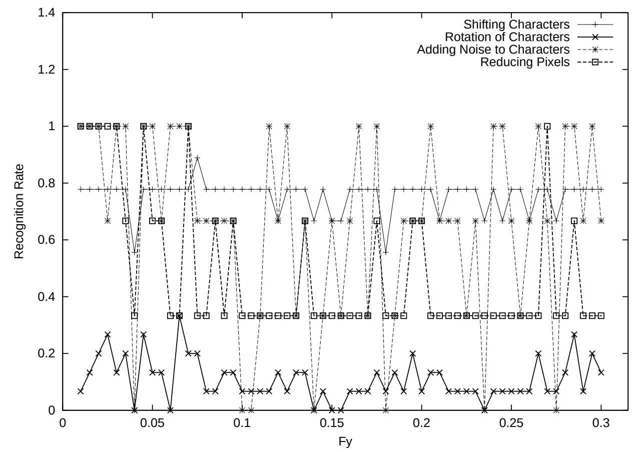

Fig. 12. Recognition Rate of Four Types of Image 'B' for Changing f_y

REFERENCES

[1] 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.

[2] H.Neumann and W.Sepp, "Recurrent V1-V2 interaction in early visual boundary processing", Biological Cybernetics, Vol.81, pp. 425–444, 1999.

[3] S.Grossberg and E.Mingolla and C.Pack, "A neural model of motion processing and visual navigation by cortical area MST", Cerebral Cortex, Vol.9, No.8, pp. 878–895, 1999.

[4] K.Fukushima, "Recognition of partly occluded patterns: a neural network model", Biological Cybernetics, Vol.84, No.4, pp. 251–259, 2001.

[5] G.A.Carpenter and S.Grossberg and J.Reynolds, "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network", Neural Networks, Vol.4, pp. 565–588, 1991.

[6] J.R.Williamson, "Self-organization of topographic mixture networks using attentional feedback", Neural Computation, Vol.13, pp. 563–593, 2001.

[7] 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.

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

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

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

[11] 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.

[12] I.Hayashi and J.R.Williamson, "Acquisition of fuzzy knowledge from topographic mixture networks with attentional feedback", Proc. The International Joint Conference on Neural Networks (IJCNN'01), pp. 1386–1391, 2001.

[13] I.Hayashi and J.R.Williamson, "An analysis of aperture problem using fuzzy rules acquired from TAM networks", Proc. IEEE International Conference on Fuzzy Systems (FUZZ-IEEE2002), pp. 914–919, 2002.

[14] I.Hayashi and H.Maeda, "A Formulation of Fuzzy TAM Network with Gabor Type Receptive Fields", Proc. of the 4th International Symposium on Advanced Intelligent Systems (ISIS2003), pp. 620–623, 2003.