2003 International Symposium on Advanced Intelligent Systems
September 25-28, 2003, Jeju, Korea

A Formulation of Fuzzy TAM Network with Gabor Type Receptive Fields

Isao Hayashi
Graduate School of Corporate Information
Hannan University
5-4-33, Amami-higashi, Matsubara,
Osaka 580-8502, Japan
ihaya@kcn.hannan-u.ac.jp

Hiromasa Maeda
Graduate School of Corporate Information
Hannan University
5-4-33, Amami-higashi, Matsubara,
Osaka 580-8502, Japan
mc02017@hannan-u.ac.jp

***Abstract*— The TAM (Topographic Attentive Mapping) network is a biologically-motivated neural network. Fuzzy rules are acquired from the TAM network by the pruning algorithm. 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

Several models which translate the human visual cortex into neural model have been proposed[1]-[6]. Recently, the model for the receptive field on the retinal-ganglion cell and visual cortex cell have been proposed[7]-[11]. A cell on the visual cortex has visual field where responds to slit singnals with visual angles varied from 0° to 5°. We call the visual field as receptive field. The receptive fields are represented by adjacent ON and OFF regions filled, respectively, with plus and minus signs, and they have the orientation selectivity to make a response of eight directions from 0° to 315° per 45°. More than one hundred receptive fields are overlapping at the same visual field. An approach for modeling receptive field is using Gabor function[12], [13]. Gabor function is defined by a oscillator which is a complex sinusoidal plane wave of some frequency and orientation within a Gaussian envelope and sine/cosine function, and they were extended to two-dimension by Daugman[13].

On the other hand, TAM (Topographic Attentive Mapping) network is a biologically-motivated neural network[6], [14], [15]. 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 Gabor type input layer for TAM network. The biological motivation for Gabor functions lies in their goodness of fit to receptive field in human visual system. One of the advantages of using Gabor filters is that they achieve optimal resolution in both edge filtering and orientation selectivity. By the advantages, we expect an approach of image processing using TAM network is achieved. In the formulation of Gabor type input layer for TAM network, eight orientations are detected in the visual field depending on the filtering size of Gabor functions and the size of receptive fields. The features in the input layer are defined as the strength of eight orientations. The strength of orientations is passed to the category layer of TAM network and the image processing is achieved at over the category layer. We formulate here first the Gabor type input layer. The consistency between Gabor function and TAM network is next discussed and finally show the usefulness of TAM network through some examples.

II. Gabor Function

Gabor function is a model for realizing the receptive field of visual cortex. Let x, $0 \leq x \leq F$ and y, $0 \leq y \leq F$ denote axes on two-dimension x and y, respectively. We call the size of $F \times F$ as filter size. The general form of two-dimensional Gabor functions, G, is represented as follows;

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

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

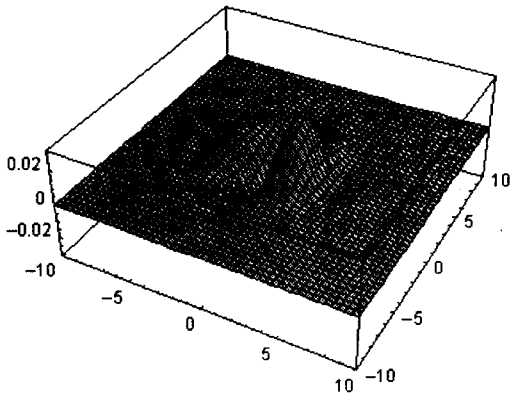

Fig. 1. Gabor Function

An example of Gabor function is shown in Figure 1. The filter's size in Figure 1 is 20×20. Given an image, the general orientation of the image is detected by the center coordinate of

Gabor function scanning some receptive fields corresponding to the whole image. We call the scale of receptive field as receptive field size and denote it as $R \times R$.

III. Fuzzy TAM Network with Gabor Type Receptive Fields

The structure of the TAM Network is shown in Figure 2. The Gabor filtering is explained in Figure 3. The Gabor function is incorporated into the input layer of TAM network to detect the general orientation of the given image. Feature maps, f_{ih}, are defined as the follows;

$$f_{ih} = G_{ih}(x, y), \quad i = 1, 2, \cdots M, \quad h = 1, 2, \cdots L \qquad (2)$$

where, i denotes the feature number and M is the product of the number of receptive fields and eight orientations. The size of receptive field is denoted $L = R \times R$.

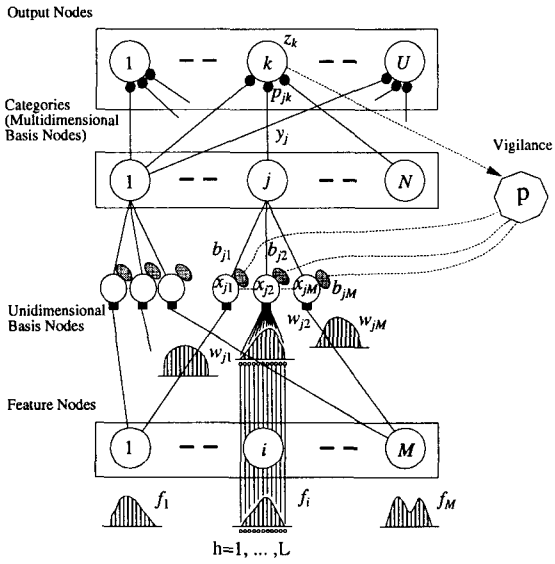

Fig. 2. TAM Network

When feature maps, f_{ih}, are calculated, the output signal to the category layer, y_j, are calculated using the node's weights, w_{jih}.

$$\begin{aligned} y_j &= \prod_{i=1}^{M} x_{ji} \\ &= \prod_{i=1}^{M} \frac{\sum_{h=1}^{L} f_{ih} w_{jih}}{1 + \rho^2 b_{ji}} \end{aligned} \qquad (3)$$

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

The output prediction, K, is calculated as follows:

$$\begin{aligned} K &= \{k | \max_k z_k\} \\ &= \{k | \max_k \sum_{j=1}^{N} y_j p_{jk}\} \end{aligned} \qquad (4)$$

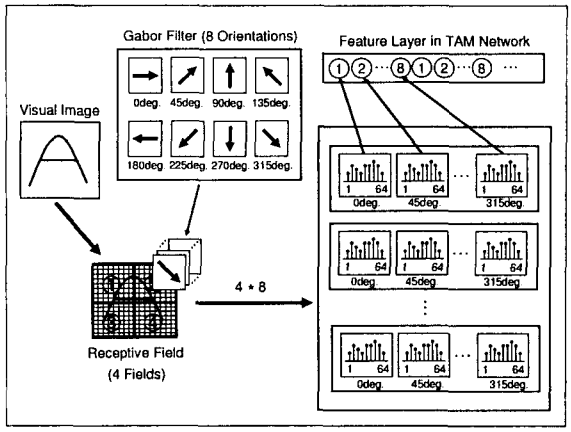

Fig. 3. Gabor Filtering

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.

$$y_j^* = \frac{\prod_{i=1}^{M} x_{ji} \times \sum_{k=1}^{U} z_k^* p_{jk}}{\sum_{j'=1}^{N} \prod_{i=1}^{M} x_{j'i} \times \sum_{k=1}^{U} z_k^* p_{j'k}} \qquad (5)$$

$$z_k^* = 1 \text{ if } k = K^*; \quad z_k^* = 0 \text{ otherwise} \qquad (6)$$

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

$$\Delta w_{jih} = \frac{\alpha y_j^* (1 - \lambda^{1/M})(f_{ih} - w_{jih})}{(\alpha - 1)\lambda^{1/M} + n_j}, \lambda \in (0, 1) \qquad (7)$$

$$\Delta p_{jk} = \frac{\alpha y_j^* (z_k^* - p_{jk})}{\alpha + n_j} \qquad (8)$$

$$\Delta b_{ji} = b_j^{(rate)} y_j^* (x_{ji} - b_{ji}) \qquad (9)$$

$$\Delta n_j = \alpha y_j^* (1 - n_j) \qquad (10)$$

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] The output prediction, K, is calculated.
[Step 2] If K is not similar enough to K^*, we do $\rho = \rho + \rho^{(step)}$. When ρ reachs the maximal level, one node is added to categories.
[Step 3] If $z_{K^*}/z_K \geq OC$, the learning step starts. Parameters, w_{jih}, p_{jk} and b_{ji}, are updated.
[Step 4] Until $z_{K^*}/z_K \geq OC$, let the algorithm repeat from step 1 to step 3.
[Step 5] 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.

$$H(i) = -\sum_{j=1}^{N} g_j \sum_{k=1}^{U} G_{jk} \log_2 G_{jk} \quad (11)$$

$$g_j = \frac{\sum_{s=1}^{R} x_{jis}}{\sum_{j=1}^{N} \sum_{s=1}^{R} x_{jis}} \quad (12)$$

$$G_{jk} = \frac{\sum_{s \in \psi_k} \gamma_{js} \times p_{jk}}{\sum_{s=1}^{R} \gamma_{js} \times p_{jk}} \quad (13)$$

$$\gamma_{js} = \prod_{i \in I^*} x_{jis} \times x_{jis} \quad (14)$$

[Step 6] The following feature i^* is extracted as an important feature and we set $I^* = \{i^*\}$.

$$i^* = \{i | \max_i H(i)\} \quad (15)$$

[Step 7] 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.

$$G_{jk} \geq \eta \quad (16)$$

[Step 8] 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.

$$\frac{1}{R} \sum_{s=1}^{R} \gamma_{js} < \theta \quad (17)$$

[Step 9] 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.

$$\varphi_{jK} = \frac{\sum_{s \in \Gamma_K} \gamma_{js} \times p_{jK}}{\sum_{j=1}^{N} \sum_{s \in \Gamma_K} \gamma_{js} \times p_{jK}} \geq \xi \quad (18)$$

[Step 10] 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 11] Until all features are selected at step 6, let the algorithm repeat from step 5 to step 10.

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 representaion.

IV. Examples

In order to show the efficiency of the Gabor type input layer, some examples are here illustrated. The alphabets, "A and B", are filled in the electronic pad corresponding to the visual field. The size of the electronic pad is 16 × 16. The training image for TAM network is shown in Figure 4 and 5. The orientations for two alphabets, "A and B" are first detected by Gabor function and their orientations are next composed as some data sets. All data sets are piled ten times after shuffling them and be totally a data set of thirty. The parameters of the TAM network are set as follows:

L	$= R \times R$	$rho\ init$	$= 0.0$
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	$= 0.5$	N_{set}	$= 6.$
μ_x	$= 0.0$	μ_y	$= 0.0$
σ_x	$= 2.02$	σ_y	$= 1.87$
f_x	$= 0.123$	f_y	$= 0.123$
ϕ	$= 0$		

The image as the checking data is shown in Figure 6. The results for the training image and the checking image are shown in Table I. The correct rates of both the training image (training data: TRD) and the checking image (checking data: CHD) are shown by a combination of the filter size of $F \times F$ and the receptive field size of $R \times R$. The number of receptive fields is depending on the filter size and the receptive field size.

Fig. 4. Training Image of 'A'

Fig. 5. Training Image of 'B'

Fig. 6. Checking Image

In Table I, all correct rates for the training data are higher than 98.3%. All correct rates for the checking data are over 75.0%. The bigger filter size is, the better the correct rate is. The highest correct rate at each filter size is the case of the receptive field divided by 16 × 8. The division by 16 × 8 means to divide the whole visual field by two receptive fields with horizontal. We guess it is easier to recognize the difference between 'A' and 'B' than the virtical case. We also guess the number of nodes in the category layer is averagely two because the number is depending on two alphabets, 'A' and 'B'.

The orientation selectivity of the left side of images in Figure 4 is shown in Figure 7. In Figure 7, the orientation

TABLE I

CORRECT RATE AND CATEGORIES

Filter Size (F × F): 4 × 4				
Number of RF	1	2	2	4
RF Size ($R \times R$)	16 × 16	8 × 16	16 × 8	8 × 8
Nodes of Category	2.1	2.4	2.2	2.0
Features	8	16	16	32
Correct Rate of TRD (%)	100.0	100.0	100.0	100.0
Correct Rate of CHD (%)	75.0	75.0	75.0	75.0
Filter Size (F × F): 8 × 8				
Number of RF	1	2	2	4
RF Size ($R \times R$)	16 × 16	8 × 16	16 × 8	8 × 8
Nodes of Category	2.2	2.1	2.1	2.4
Features	8	16	16	32
Correct Rate of TRD (%)	100.0	100.0	100.0	98.3
Correct Rate of CHD (%)	92.5	77.5	95.0	75.0
Filter Size (F × F): 16 × 16				
Number of RF	1	2	2	4
RF Size ($R \times R$)	16 × 16	8 × 16	16 × 8	8 × 8
Nodes of Category	2.4	2.3	2.2	2.3
Features	8	16	16	32
Correct Rate of TRD (%)	100.0	100.0	100.0	100.0
Correct Rate of CHD (%)	100.0	80.0	100.0	87.5

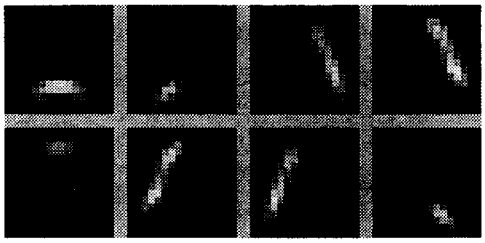

Fig. 7. Image after Filtering

angle of the left-upper side indicates 0° and be moving the right side one by one to be 45°, 90°, 135°, 180°, 225°, 270°, 315°, respectively. The good edge elements according to the orientation selectivity are detected.

Fig. 8. Checking Image for Robustness

Finally, we discuss the robustness of TAM network given the smaller alphabet 'A' compared with the original image. The checking images for robustness is shown in Figure 8. The result is shown in Table II. The correct rate in the case of the left-side is lower compared with Table I, but the correct rate in the case of the right-side is better than we expected. That means the TAM network is partly capable of robustness. All result support the usefulness of the TAM network with Gabor type input layer.

V. CONCLUSIONS

We fromulated here the TAM network with the Gabor type input layer and showed the usefulness through examples.

TABLE II

CORRECT RATE FOR ROBUSTNESS

Filter Size: 16 × 16 of the left-side				
Number of RF	1	2	2	4
RF Size	16 × 16	8 × 16	16 × 8	8 × 8
Nodes of Category	2.2	2.0	2.6	2.1
Features	8	16	16	32
Correct Rate of CHD (%)	0.0	0.0	0.0	60.0
Filter Size: 16 × 16 of the right-side				
Number of RF	1	2	2	4
RF Size	16 × 16	8 × 16	16 × 8	8 × 8
Nodes of Category	2.5	2.3	2.0	2.2
Features	8	16	16	32
Correct Rate of CHD (%)	100.0	100.0	100.0	100.0

However, we need more experiments and discussion in order to find characteristic of the Gabor type input.

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.

REFERENCES

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

[2] H.Neumann and W.Sepp, "Recurrent V1-V2 interaction in early visual boundary processing", *B*iological 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", *C*erebral Cortex, Vol.9, No.8, pp. 878–895, 1999.

[4] K.Fukushima, "Recognition of partly occluded patterns: a nural network model", *B*iological 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", *N*eural Networks, Vol.4, pp. 565–588, 1991.

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

[7] A.D.Pollen and S.F.Ronner, "Visual cortical neurons as localized spatial frequency filters", *I*EEE Transactions of System, Man and Cybernetics, Vol.SMC13, pp. 907–916, 1983.

[8] W.Urushihara and T.Nagano, "A model for the detection of second-order motion", *T*echnical report of IEICE, Vol.NC98-191, pp. 293–298, 1999.

[9] K.Okajima, "The Gabor function extracts the maximum information from input local signals", *N*eural Networks, Vol.11, pp. 435–439, 1998.

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

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

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

[13] J.Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientaion optimized by two-dimensional visual cortical filters", *O*ptical Society of America, Vol.2, No.7, pp. 1160–1169, 1985.

[14] 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.

[15] 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.