A Study of Orientation Selectivity of TAM Network Incorporated Receptive Field Structure

Isao Hayashi[1] and James R. Williamson[2]

[1] Kansai University, Takatsuki, Osaka, Japan, ihaya@res.kutc.kansai-u.ac.jp
[2] Lockheed Martin Corporation, Lexington, MA, U.S.A., jrw@ll.mit.edu

TAM (Topographic Attentive Mapping) network is a biologically-motivated neural network. In this paper, we define the Gabor function type receptive field and incorporate it to the TAM network's feature layer. We also discuss the orientation selectivity of the receptive field through some examples of character recognition.

1 Introduction

In the visual cortex, the orientation selectivity of simple cell, complex cell and hypercomplex cell play important roles for detecting contours of visual objects. Simple cell has the slit type receptive fields which are represented by adjacent ON and OFF regions filled with plus and minus signs and sixteen orientations. Many models of receptive fields have been proposed[1, 2]. Marčelja[1] defined Gabor function as a oscillator which is a complex sinusoidal plane wave of some frequency and orientation. Daugman[2] extended it to two-dimensional format.

Many neural networks imitating human visual system have been also proposed[3, 4, 5]. The TAM network[5, 6] is a biologically-motivated neural network. The TAM network has four layers, the feature layer, the basis layer, the category layer and the class layer, and achieves the input-output relationship between the distributed input data and the categorized output classes using the bottom-up and the top-down learning.

In this paper, we formulate a new TAM network with feature layer of the Gabor function type receptive field. By Gabor filtering, the sixteen oriention selectivities for visual object's contour are detected. The feature map in the feature layer is compound from density values accumulated the sixteen orientations. We especially discuss here the orientation selectivity of the feature layer and show the usefulness of the TAM network through some examples of character recognition.

2 TAM Network

The structure of the TAM network is shown in Figure 1. In the category layer, the activities, $x_{ji},\ i = 1, 2, \cdots, M,\ j = 1, 2, \cdots, N$, are described by products between feature maps, f_{ih} and node's weights, $w_{jih},\ h = 1, 2, \cdots, L$. The output signal of the category layer, y_j, are then calculated.

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

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

The output prediction, K, is calculated as follows:

$$K = \{k | \max_k z_k\} = \{k | \max_k \sum_{j=1}^{N} y_j p_{jk}\} \tag{2}$$

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

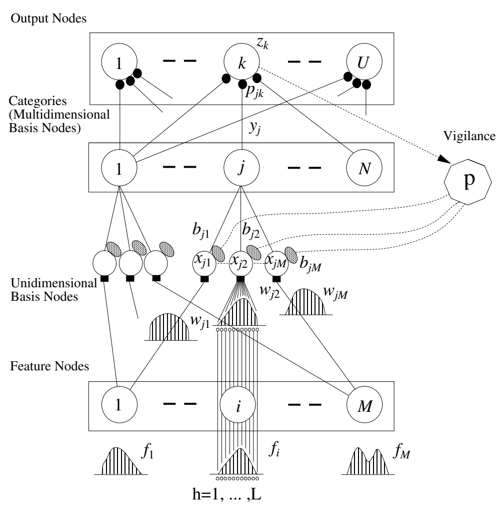

Fig. 1. TAM Network

Let K^* and K denote the supervised output class and the network's output prediction, respectively. If $K \neq K^*$, 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. When the vigilance parameter reaches the maximal level, one more node is added to the category layer. When $z_{K^*}/z_K \geq OC$, the learning parameters, $w_{jih},\ p_{jk}$ and b_{ji}, are renewed as follows:

$$\triangle w_{jih} = \frac{\alpha y_j^*(1-\lambda^{1/M})(f_{ih}-w_{jih})}{(\alpha-1)\lambda^{1/M}+n_j}, \qquad \lambda \in (0,1) \tag{3}$$

$$\triangle p_{jk} = \frac{\alpha y_j^*(z_k^*-p_{jk})}{\alpha+n_j} \tag{4}$$

$$\triangle b_{ji} = b_j^{(rate)} y_j^*(x_{ji}-b_{ji}) \tag{5}$$

where, α, λ, n_j and $b_j^{(rate)}$ are parameters, and y_j^* is the feedback signal.

3 TAM Network Incorporated Receptive Fields

Figure 2 shows the receptive field type feature layer. In Ganglion cell, sixteen orientations of visual image on the retina are detected through Gabor filtering. The sixteen orientations are accumulated as the feature maps in the LGN and input to the feature layer of TAM network.

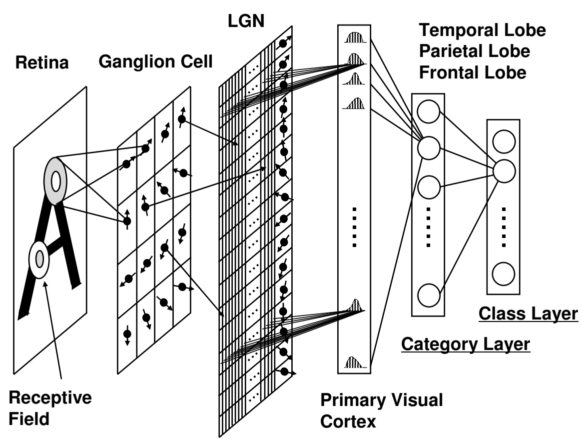

Fig. 2. TAM Network with Receptive Field

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

$$G(x,y) = Ke^{-\frac{1}{2}(\frac{(x-\mu x)^2}{\sigma_x^2}+\frac{(y-\mu y)^2}{\sigma_y^2})} \times \sin(2\pi f_x x\cos\theta + 2\pi f_y y \sin\theta + \phi) \tag{6}$$

where, K is the amplitude, (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.

Let the horizontal and vertical scale of the original visual image, $I(p, q)$, denote as R_H and R_V, respectively. In the discrete type Gabor filtering algorithm, the contour image, $C_i(x, y),\ i = 1, 2, \cdots, M,\ 0 \le x \le R_H,\ 0 \le y \le$

R_V, of the i-th orientation selectivity is calculated by the convolution process between the Gabor function and the original visual image as follows:

$$C_i(x, y) = \sum_{q=1}^{R_V} \sum_{p=1}^{R_H} G_i(x - p, y - q) \times I(p, q). \tag{7}$$

The orientation selectivities of $R_H \times R_V$ are obtained by shifting the receptive field one pixel in the visual image. The i-th feature map, f_{ih}, is described as a following normalized format:

$$f_{ih} = \frac{\sum_{\{x,y|O_i(x,y)=h\}} C_i(x, y)}{\sum_{y=1}^{R_V} \sum_{x=1}^{R_H} C_i(x, y)}, \qquad h = 1, 2, \cdots, 256. \tag{8}$$

4 Character Recognition

In order to show the usefulness of TAM network, some examples of character recognition are illustrated. The visual image size of alphabets is 15 $pixels \times$ 15 $pixels$.

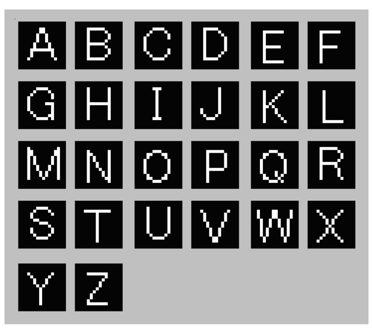

Fig. 3. Training Images

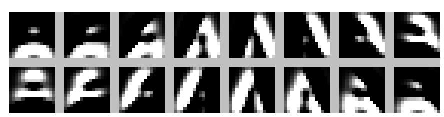

Fig. 4. Orientation Selectivities after Gabor Filtering

The training images are shown in Figure 3. The contour images of sixteen orientation selectivities of the alphabet, 'A', are shown in Figure 4. The left-upper side image indicates 0 degrees, and the next images moving the right side show 22.5 degrees, 45 degrees, 67.5 degrees, $\cdots$ 337.5 degrees, respectively. The recognition rates are shown in Figure 5. We set here the trials of the experiment as 30 times, and compared the recognition rates of two epochs with five epochs. We obtained at least the higher recognition rates than 60%.

Next, in order to discuss the robustness of TAM network for reducing pixels, the checking images with reducing three types of pixels, 10%, 20% and

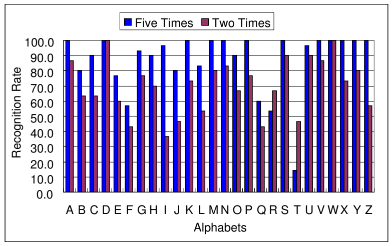

Fig. 5. R.R. of All Alphabets

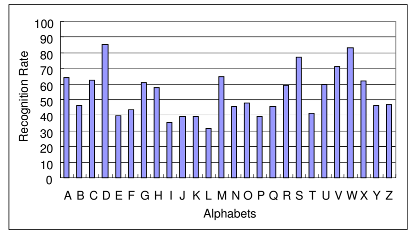

Fig. 6. R.R. of Reducing 10% Pixels

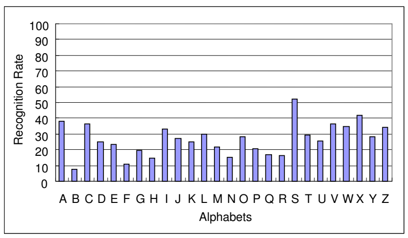

Fig. 7. R.R. of Reducing 20% Pixels

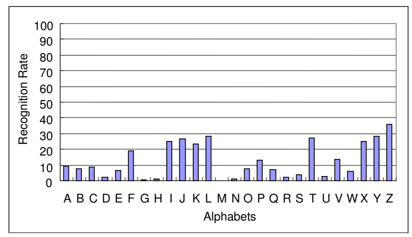

Fig. 8. R.R. of Reducing 30% Pixels

30%, are estimated. Figure 9 shows the checing images. The results are shown in Figure 6 to Figure 8. The recognition rates are obtained as the average of 30 trials. In Figure 6, the recognition rates of four alphabets, ‘D’, ‘S’, ‘V’, and ‘W’, are relatively higher than others, but the recognition rates in two groups of ‘E’, ‘F’, and ‘I’, ‘J’, ‘K’, ‘L’, ‘T’ are lower because of the similarity of alphabets in each group. In Figure 7, the almost recognition rates are under 40%, and all recogtion retes without ‘Z’ are extremely very low in Figure 8. However, the distribution shapes of recognition rates through Figure 6 to Figure 8 are very similar each other. Therefore, we should notice that the recogntion rates are strongly affected by the shape of alphabets.

We last discuss the recognition rates for rotation and inclining characters. The checking images are shown in Figure 10. The alphabets, ‘I’ and ‘L’, have relatively the robustness for rotation of 180 degrees and 270 degrees. The alphabets, ‘N’, ‘O’, ‘X’, and ‘Z’, are strongly robust for rotation of 180 degrees. Since these alphabets consist of the straight lines and simple connections, that result should be understandable. For the inclining characters, the recognition rates without the alphabet ‘Z’ are very low. That weakness of TAM network for rotation and inclining characters comes from too sensitive for the orientation selectivity of TAM network. Therefore, if the TAM network had the more wider orientation selectivity covering whole visual field with the multi-layers’ bottom-up and top-down learning, the recognition rate would be more increased.

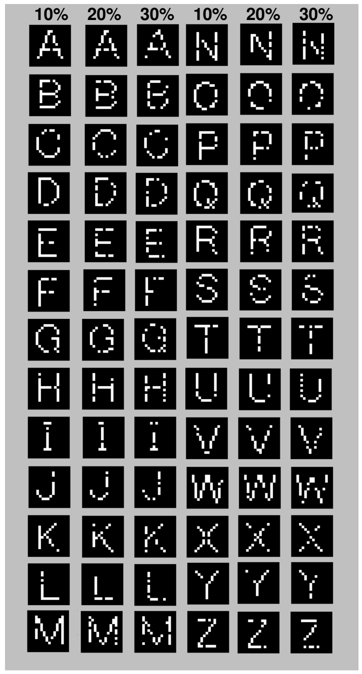

Fig. 9. Checking Images of Reducing Pixels

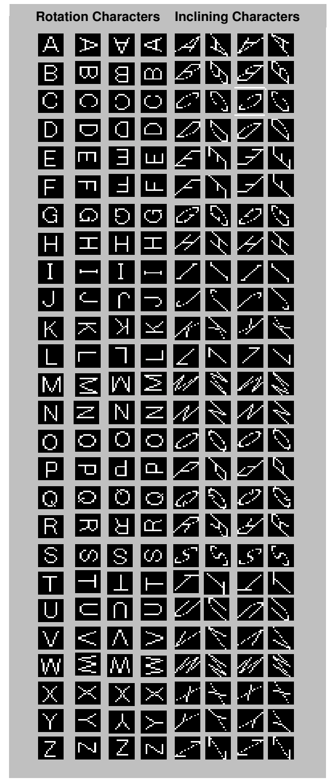

Fig. 10. Checking Images of Rotation and Inclining Characters

5 Conclusions

We formulated here the TAM network with Gabor function type receptive field and discuss its orientation selectivity through some examples of character recognition.

References

1. S.Marčelja (1980) Optical Society of America 70,11:1297-1300
2. J.Daugman (1985) Optical Society of America 2,7:1160-1169
3. S.Grossberg (1999) Spatial Vision 12,2:163-185
4. K.Fukushima (2001) Biological Cybernetics 84,4:251-259
5. J.R.Williamson (2001) Neural Computation 13:563-593
6. I.Hayashi and H.Maeda (2003) Proceedings of the 4th International Symposium on Advanced Intelligent Systems(ISIS2003) 620-623