**FUZZ-IEEE2002 in WCCI2002, pp.914-919,
Honolulu, U.S.A., May 12-17, 2002**

An Analysis of Aperture Problem Using Fuzzy Rules Acquired from TAM Network

Isao Hayashi * and James R. Williamson **
* Hannan University, Graduate School of Corporate Information
5-4-33, Amami-higashi, Matsubara, Osaka 580-8502, Japan, email: ihaya@hannan-u.ac.jp
** Lockheed Martin Corporation, Lincoln Laboratory, MIT
244 Wood Street, Lexington, MA 02420-9185, USA, email: jrw@ll.mit.edu

Abstract - The aperture problem is a significant experiment for discussing visual models. Three circle apertures has been displayed at a computer CRT and a line inside the middle circle is moving. Then, a subject perceives perceptual grouping changing orientation of the line. In this paper, we analysis aperture experiment data using fuzzy rules acquired from TAM Network (Topographic Attentive Mapping Network) based on a biologically-motivated neural network with folded feedback mechanism.

I. INTRODUCTION

In the human visual system, an image on retina enters to the lateral geniculate nucleus(LGN) through rod cells and pyramidal cells, and is sent to the primary visual cortex, V1 and V2. In V1 and V2, the contour, color and texture of the image are detected and the motion of direction is perceived. After going through visual cortex, the image signal is divided into two paths which are for the temporal lobe and the vertral lobe[1]. A lot of visual neural networks based on the human visual system have been proposed, e.g., BCS[2], Neocognitron[3], ARTMAP[4], fuzzy ARTMAP[5], TAM[6], Visual Cortex Model[7], V1-V2 model[8] and so on.

In the Visual Cortex Model, Grossberg has discussed the concept of visual system and showed a guideline for modeling visual systems. He especially mentioned the following two mechanisms are necessary subjects for visual models.

1. The folded feedback signals
2. The horizontal connections with excitatory and inhibitatory signals

Perceptual grouping[7] in aperture problem is a kind of visual phenomenon and be an good experiment data for constructing visual models. In this paper, we show a perceptual grouping in circle aperture problems in which a circle has been displayed at the computer CRT and subjects perceive perceptual grouping when a line inside circles is moving[9]-[14]. Nishina et al.[12] pointed out the perceptual grouping strongly depends on the display time long and discussed the neccessity to include the feedback mechanism in visual models. We report here we have gotten the similar result which supports the part of Nishina's insistences. Alternatively, the aperture experiment data is analyzed using fuzzy rules acquired from TAM (Topographic Attentive Mapping) network. The TAM network has three layers constructed by feature, category and class layers. The feature layer is constructed for imitating the retina and the feature node has a receptive field. The category layer imitates the LGN and the class layer imitates V1. The learning algorithm of TAM network provides a folded feedback mechanism and let the output prediction of TAM network be coincided with supervised output. The weak connections in TAM network are removed by pruning algorithm[14] and fuzzy rules are acquired from network structure since the data procedure of the TAM network is similar to fuzzy logic.

We here analysis the aperture experiment data using TAM network and show the necessity of the feedback mechanism of the TAM network.

II. APERTURE PROBLEM

The aperture problem is a kind of visual experiment. Figure 1 shows an outline of aperture problem. A circle aperture has been first displayed at the center of computer CRT and a line inside the circle moves from the right-bottom to the left-up. While the line is moving, other two circles at both sides of 45 degrees of the center circle appear and each line inside two circles is also moving from the bottom to the top at the vertical direction. If a subject would recognize those three lines might be connected each other, the center line is perceived to have changed the direction from the left-up to the top at the vertical direction as same as the lines at both sides. This phenomenon is called perceptual grouping. Nishina

0-7803-7280-8/02/$10.00

et al.[12] showed that the perceptual grouping strongly depends on the length of time a display is viewed. Alternatively, they discussed the neccessity of feedback mechanism of visual model.

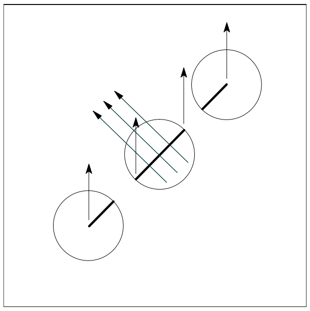

Figure 1: APERTURE EXPERIMENT

In this paper, we assume the perceptual grouping at aperture experiments would depend on the relationship among the distance between subject and a computer display, the radius of circles and the distance between circles. In order to specify the relationship, we let the distance between a subject and the computer display be 50cm, and conducted two kinds of experiments whose ratio of the radius to the distance between circles is constant and be not.

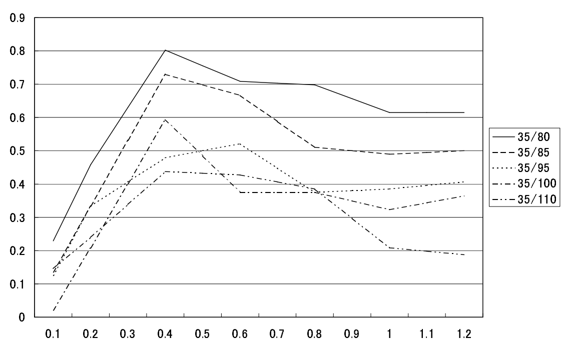

Figure 2: REACTION OF SUBJECTS (INCONSTANT CASE)

Figure 2 shows the reaction of subjects when the ratio of radius (mm) to the distance (mm) is not constant. Five experiments, $(35, 80)$, $(35, 85)$, $(35, 95)$, $(35, 100)$ and $(35, 110)$, where (a,b) means that (radius, distance between circles), are conducted by four subjects who are in twenties. The horizontal axis represents the display time (sec) and the vertical axis represent the perceptual rate that subjects perceived the motion of direction of lines. Each subject conducted 420 experiments double which means totally 3,360 experiments for four subjects. We found that Figure 2 supported the part of Nishina's paper in terms of dependency of perceptual rate to display time. Alternatively, the perceptual rate is lower when the distance between circles is longer since the radius was fixed. Additionally, we found there were the maximal values of the perceptual rate and the function of perceptual rate was convex. That means that subjects possess the most visiable display time.

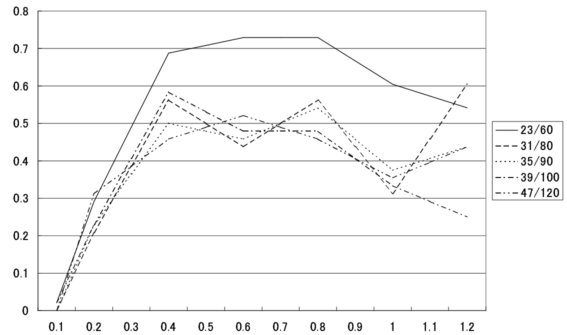

Figure 3: REACTION OF SUBJECTS (CONSTANT CASE)

Figure 3 shows the second experiments in which the ratio of radius to the distance is constant. Five experiments, $(23, 60)$, $(31, 90)$, $(35, 90)$, $(39, 100)$ and $(47, 120)$ are conducted when the ratio was settled as 0.38. Each subject conducted 420 experiments single which means totally 1,680 experiments by four subjects. Excepting (23,60), the perceptual rates are approximately the same shape. That means that subjects perceived the same kind of perceptual grouping among the ratio is constant.

III. TAM NETWORK

The structure of TAM Network is shown in Figure 4. When feature maps, f_i, are given, the basis nodes are activated by the match between the activity distribution in a feature map, f_{ih}, and the distribution of the node's weights, w_{jih}. The outputs to output nodes, y_j, are calculated as follows:

$$x_{ji} = \frac{\sum_{h=1}^{L} f_{ih} w_{jih}}{1 + \rho^2 b_{ji}} \qquad (1)$$

$$y_j = \prod_{i=1}^{M} x_{ji} \qquad (2)$$

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

The class prediction, K, is the index of the maximally activated output node:

$$K = \{k | \max_k z_k\} \qquad (3)$$

$$z_k = \sum_{j=1}^{N} y_j p_{jk} \qquad (4)$$

where, p_{jk} are weighted connections which represent the probability of each output given the category.

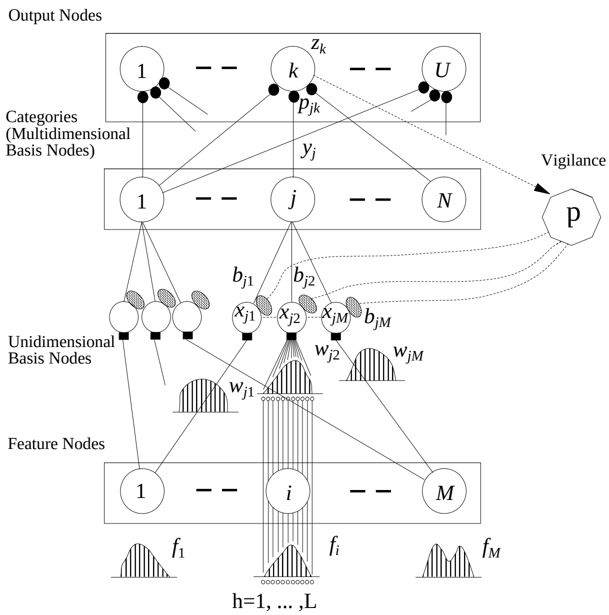

Figure 4: TAM NETWORK

Let K^* denote the index of the "correct" supervised output class. If the network's output prediction K is not similar enough to the supervised output K^*, the "attention" is invoked and the ρ is raised until either the following subject is satisfied or until the maximal vigilance level, OC, is reached. When the ρ reached the maximal level, one node is added to categories.

$$\begin{array}{l} \text{If } z_{K^*}/z_K < OC \text{ then repeat} \\ \quad (a)\ \rho = \rho + \rho^{(step)} \\ \quad (b)\ \text{equation } (1)-(4) \\ \text{until either } z_{K^*}/z_K \geq OC \text{ or } \rho \geq \rho^{(max)}. \end{array} \tag{5}$$

Once the subject of $z_{K^*}/z_K \geq OC$ is satisfied, the feedback signal y_j^* is calculated and the process moves to learning step.

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

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

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

$$\triangle w_{jih} = \frac{\alpha y_j^*(f_{ih} - w_{jih})}{\alpha\beta(M) + n_j} \tag{8}$$

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

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

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

$$\beta(M) = \frac{\lambda^{1/M}}{1 - \lambda^{1/M}}, \quad \lambda \in (0,1) \tag{12}$$

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

Whenever one datum put into the TAM network, the learning step is invoked and parameters are adjusted. The learning process is terminated when some epochs are achieved.

After learning, the process moves to pruning step. When the data set, D, in which the data, $f_{si},\ s = 1, 2, \cdots, R$ with the class k, are given, all data, R, are divided into two group, learning data for learning the parameters and checking data for estimating the result.

The information entropy[15], $H(i)$, is calculated for estimating the significance of feature maps, and all of features are lined in order of value by $H(i)$.

$$H(i) = -\sum_{j=1}^{N} g_j \sum_{k=1}^{U} G_{jk} \log_2 G_{jk} \tag{13}$$

$$g_j = \frac{\sum_{s=1}^{R} x_{jis}}{\sum_{j=1}^{N} \sum_{s=1}^{R} x_{jis}} \tag{14}$$

$$G_{jk} = \frac{\sum_{s\in\psi_k} \gamma_{js} \times p_{jk}}{\sum_{s=1}^{R} \gamma_{js} \times p_{jk}} \tag{15}$$

$$\gamma_{js} = \prod_{i\in I^*} x_{jis} \times x_{jis} \tag{16}$$

where, ψ_k is a set of the data with the class k.

The following feature i^* is extracted.

$$i^* = \{i | \max_i H(i)\} \tag{17}$$

In order to estimate the strength of link connections, the following three rules are defined.

[Pruning Rule 1]

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

$$G_{jk} \geq \eta \tag{18}$$

[Pruning Rule 2]

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

$$\frac{1}{R}\sum_{s=1}^{R} \gamma_{js} < \theta \tag{19}$$

[Pruning Rule 3]

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

$$\varphi_{jK} = \frac{\sum_{s \in \Gamma_K} \gamma_{js} \times p_{jK}}{\sum_{j=1}^{N} \sum_{s \in \Gamma_K} \gamma_{js} \times p_{jK}} \geq \xi \tag{20}$$

The learning of TAM network and pruning of nodes and link connections are achieved according to the following algorithm:

[Step 1] The output prediction, K, is calculated.
[Step 2] If K is not similar enough to the supervised output K^*, we do $\rho = \rho + \rho^{(step)}$. When the ρ reached the maximal level, one node is added to categories.
[Step 3] If $z_{K^*}/z_K \geq OC$ is satisfied, the learning step starts.
[Step 4] Until $z_{K^*}/z_K \geq OC$ is satisfied, let the algorithm repeat from step 1 to step 3.
[Step 5] After learning, the pruning step starts. The information entropy, $H(i)$, is calculated using the learning data.
[Step 6] The following feature i^* is extracted and we set $I^* = \{i^*\}$.

$$i^* = \{i | \max_i H(i)\} \tag{21}$$

[Step 7] If the following condition is satisfied for checking data, the link connections between the category j and classes $k' \neq k$ are removed. The link connections between the category j and features $i' \notin I^*$ are also removed.

$$G_{jk} \geq \eta \tag{22}$$

[Step 8] If the following condition is satisfied for checking data, the link connections between the category j and features i and $i' \notin I^*$ are removed.

$$\frac{1}{R} \sum_{s=1}^{R} \gamma_{js} < \theta \tag{23}$$

[Step 9] If the following condition is satisfied for checking data, the link connections between the class k and categories $j' \neq j$ are removed.

$$\varphi_{jk} \geq \xi \tag{24}$$

[Step 10] When a category has lost connections to all classes or features, the category is removed. Any class and feature has been disconnected from all categories are 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 optimum neural network pruned needless connections and nodes is obtained. We should notice that the algorithm is a kind of fuzzy tuning methods which adjust the number of features, classes and fuzzy rules since the data procedure is the same to the fuzzy logic.

IV. AN ANALYSIS OF APERTURE PROBLEMS USING TAM NETWORKS

We analysis here the aperture experiment data in Section II using the TAM network and show the availability of the feedback mechanism of the TAM network. The parameters for learning are set as follows:

$$\begin{array}{lll lll}
epoch & = & 1 & category\ init & = & 0 \\
L & = & 5 & rho\ init & = & 0.0 \\
OC & = & 0.8 & rho\ step & = & 0.1 \\
\alpha & = & 0.0000001 & rho\ max & = & 100.0 \\
\lambda & = & 0.33 & b_j^{(rate)} & = & 0.01 \\
\eta & = & 0.8 & \theta & = & 0.03 \\
\xi & = & 0.5. & & &
\end{array}$$

Table 1: CORRECTNESS AND THE NUMBER OF NODES OF TAM NETWORK

		TAM	Pruning TAM
Figure 2	Correctness (%)	62.4	58.8
	Number of Nodes	6	6
Figure 3	Correctness (%)	63.7	60.7
	Number of Nodes	6	1

Table 1 shows the correctness and the number of nodes at category layer of the TAM network before pruning and after pruning. The correctness is not so well. The reason that the correctness was not so well is that the data set includes many inconsistant supervised output that subjects answered "yes" and "no" for the responce of perceptual grouping in spite of the same feature values. We think that the low correctness come from the lack of horizontal connections with excitatory and inhibitory signals. On the other hand, the number of nodes for Figure 3 is reduced to one from six by pruning procedure. Since the nodes of category layer represent fuzzy rules, we interpret the number of fuzzy rules for representing Figure 3 is reduced.

Figure 5 shows the fuzzy rules acquired from the TAM network before pruning for the data in Figure 2. Figure 6 shows the fuzzy rules acquired from the TAM network before pruning for the data in Figure 3. The depth axis represents the fuzzy rule's number and the horizontal axis represents the discrete type fuzzy sets of features, which are the circle radius, the distance between circles, the dispaly time of antecedent part of fuzzy rules and the perceptual rate of conclusion part, where Figure 5 and Figure 6 only show the fuzzy rules representing perceptual grouping.

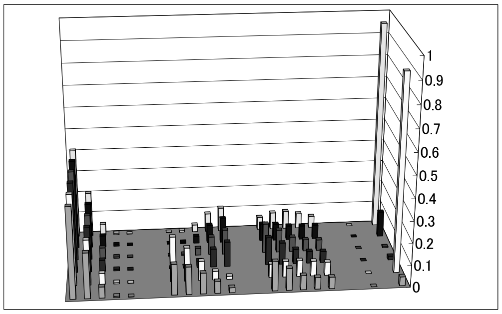

Figure 5: FUZZY RULES OF FIGURE 2

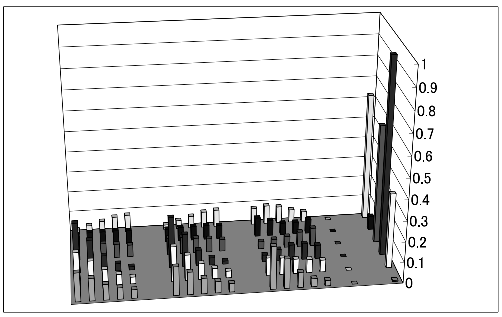

Figure 6: FUZZY RULES OF FIGURE 3

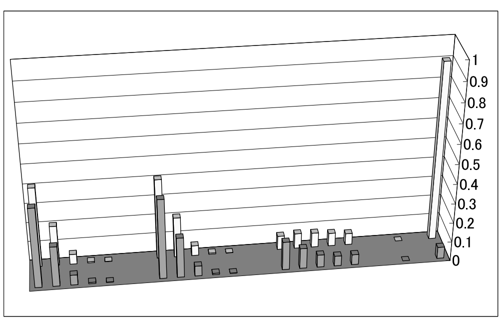

Figure 7: FUZZY RULES OF (35,80)

In the Figure 5, the perceptual rates of the second and the sixth fuzzy rules whose third feature's fuzzy sets represent middle, are higher than other rules. We interpret the reason as the higher perceptual rates around 0.4s and 0.6s at display time in Figure 2. In the Figure 6, the perceptual rate of the third fuzzy rules whose third feature's fuzzy sets is middle, is higher than others. We also interpret the reason as the higher perceptual rate around 0.6s in Figure 3.

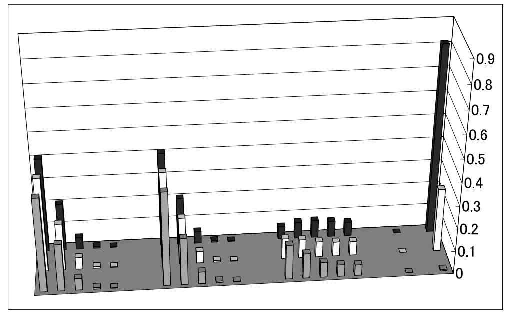

Figure 8: FUZZY RULES OF (35,95)

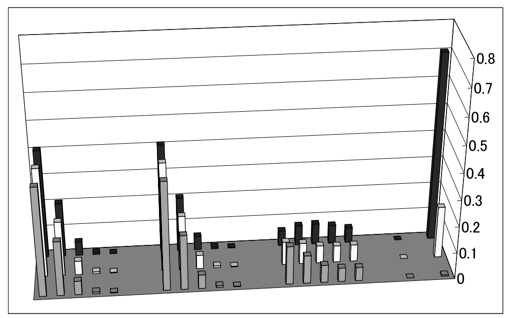

Figure 9: FUZZY RULES OF (35,110)

Next, we analysis the similarity of perceptual rates of (35,80), (35,95) and (35,110) in Figure 2 using the fuzzy rules. Figure 7 shows the fuzzy rules of (35,80). Figure 8 and 9 show the fuzzy rules of (35,95) and (35,110), respectively. Table 2 shows the number of fuzzy rules and correctness related to Figure 7 to Figure 9. In the Figure 7 to Figure 9, the membership functions of three features at antecedent part are approximately the same since the function of perceptual rates of (35,80), (35,95) and (35,110) in Figure 2 are the same shapes. On the other hand, the perceptual rate in Figure 7, the singleton of conclusion part, is the highest among Figure 7 to Figure 9. We interpret the reason as the highest perceptual rates of (35,80) among (35,80), (35,95) and (35,110) in Figure 2. Moreover, the perceptual rate in Figure 9 is the lowest. We also interpret the reason as the lowest perceptual rate of (35,110).

Finally, in order to verify the necessity of feedback mechanism of TAM network, let the vigilance learning (and b_{ji}) be inactive, i.e., the feedback mechanism is stooped. Table 3 shows the comparision results between

Table 2: THE NUMBER OF FUZZY RULES AND CORRECTNESS OF TAM NETWORK

(radius, center distance)	Number of Fuzzy Rules	Correctness (%)
(35,80)	4	53.3
(35,95)	6	62.5
(35,110)	6	71.0

by TAM network with feedback mechanism and without feedback mechanism, where the data of (39,100) in Figure 3 is used for the checking data and other data are for the learning data. The correctness of TAM network without the feedback mechanism couldn't be measured because of the endless increasing of nodes at category layer. The correctness and the number of nodes of TAM network after pruning is slightly better than before pruning.

From these results, we insist the TAM network might be a useful model since acquired fuzzy rules are acceptable.

Table 3: COMPARISON BETWEEN WITH FEEDBACK AND WITHOUT FEEDBACK FOR TAM NETWORK

		TAM	Pruning TAM
with Feedback	Learning Data (%)	62.9	62.9
	Number of Nodes	6	6
	Checking Data (%)	55.4	66.1
	Number of Nodes	6	1
without Feedback	Learning Data (%)	–	–
	Number of Nodes	over 1000	over 1000
	Checking Data (%)	–	–
	Number of Nodes	–	–

V. CONCLUSIONS

We analysis the aperture experiments using fuzzy rules acquired from TAM network and discuss the usefulness of feedback mechanism of TAM network.

References

[1] R.L.Gregory, "Eye and Brain", Oxford University Press (1998)

[2] S.Grossberg and E.Mingolla: "Neural Dynamics of Perceptual Grouping: Texture, Boundaries, and Emergent Segmentation," *Percept Psychophys*, Vol.38, pp.141-171 (1985)

[3] K.Fukushima: "Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position," *Biological Cybernetics*, Vol.36, pp.193-202 (1980)

[4] G.A.Carpenter, 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)

[5] G.A.Carpenter, S.Grossberg, N.Markuzon, J.Reynolds and D.B.Rosen: "Fuzzy ARTMAP: A Neural Network Architecture for Incremental Learning of Analog Multidimensional Maps," *IEEE Transactions on Neural Networks*, Vol.3, pp.698-713 (1992)

[6] J.R.Williamson: " Self-organization of Topographic Mixture Networks Using Attentional Feedback," *Neural Computation*, Vol.13, pp.563-593 (2001)

[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] S.Grossberg and E.Mingolla: "Neural Dynamics of Form Perception: Boundary Completion, Illusory Figures, and Neon Color Spreading," *Psychological Review*, Vol.92, pp.173-211 (1985)

[10] M.Kawato, H.Hayakawa and T.Inui: "A Forward-Inverse Optics Model of Reciprocal Connections between Visual Areas," *Network: Comutation in Neural Systems*, Vol.4, pp.415-422 (1993)

[11] M.B.Ben-av and M.Shiffrar: "Disambiguating Velocity Estimates Across Image Space," *Vision Research*, Vol.35, No.20, pp.2889-2895 (1995)

[12] S.Nishina, M.Okada and M.Kawato: "Filling-in Process and Global Binding in Computation of Motion Direction," *in November, NC conference of IEICE*, (1998) in Japanese

[13] S.Grossberg, E.Mingolla and L.Viswanathan: "Neural Dynamics of Motion Integration and Segmentation within and Across Apertures," *Boston University Technical Report*, CAS/CNS-TR-2000-004 (2000)

[14] I.Hayashi and J.R.Williamson: "Acquisition of Fuzzy Knowledge from Topographic Mixture Networks with Attentional Feedback," *the International Joint Conference on Neural Networks (IJCNN'01)*, pp.1386-1391 (2001)

[15] I.Hayashi: "Acquisition of Fuzzy Rules Using Fuzzy ID3 with Ability of Learnining for AND/OR Operators," *1996 Australian New Zealand Conference on Intelligent Information Systems*, pp.187-190 (1996)