A learning model of head-direction cells and grid cells by VQ layers connected via anti-Hebbian synapses

Naoki Oshiro, Koji Kurata and Tetsuhiko Yamamoto
Faculty of Engineering, University of the Ryukyus,
1 Senbaru, Nishihara, Okinawa. 903–0213
oshiro@mibai.tec.u-ryukyu.ac.jp

Abstract

In this paper, we proposed three layered self-organizing model to extract head direction and position of a moving object separetely. The model consists of three layers, each of which is a self-organizing vector quantization (VQ) model. The second layer receives inhibitory input from the first, and the thrid layer receives inhibitory input from the first and the second. The first layer is to detect head direction and the second and the third are to detect position. The information representation in the second and the third layers is shown to be multi-ary expression and the units in the third layer develop receptive field with grid structure as was observed in entorhinal cortex of a rat.

1 Introduction

The self-organizing map (SOM) and the vector quantization (VQ) are very famous and widely used among many self-organizing models. They are useful for application and also important as a computational models of neural systems.

SOM algorithm was proposed by Kohonen[1] as a model of the cerebral cortex and its self-organization. It was successful at reproducing a functional map of the visual cortex[2, 3], and was applied to many kinds of data as a statistical tool of nonlinear auto-regression[1]. This map can extract major informations from multi-dimensional data[4].

The vector quantization (VQ) is an on-line learning algorithm to generate reference vectors for a set of input vectors[1, 5, 6, 7]. For a given input vector, the closest reference vector is chosen to approximate it. VQ algorithm can generate a set of reference vectors which minimizes mean square of approximation error. The algorithm can be interpreted as a self-organizing neural network model adopting competitive Hebbian learning rule, i.e. a self-organizing map (SOM) lacking neighborhood learning. Thus, VQ and SOM are very similar models closely related to each other.

We studied a model consisting of two SOMs connected via anti-Hebbian connections, and showed that it can extract two different information components on the two SOMs each[8]. The model was applied to extract position and head direction from visual information. In this application, one of the two SOMs consists of two-dimensional array of cells so that it can represent two-dimensional positional information, and the other SOM consisting of one-dimensional array of cells is assigned for one-dimensional information of direction of the moving object. Simulational results showed that position-sensitive and direction-insensitive cells are formed on the two-dimensional SOM, and direction-sensitive and position-insensitive cells are formed on the one-dimensional SOM.

In recent study of cortical micro circuits, it was found that positional information are represented by ‘grid cell’ in the dorsocaudal medial entorhinal cortex (dMEC)[9]. Grid cells are a kind of position cells but their *receptive field* shows periodic grid structure in the area within which rats are allowed to walk around. A receptive field of a position cell is the set consisting of all positions at which the cell fires. It is striking that the receptive field of the grid cell in dMEC was not connected but distributed periodically in hexagonal lattice, for in most of the self-organizing models receptive fields of the cells tend to be connected. We, however, succeeded in reproducing grid-cell-like position cells by a self-organizing model consisting of two VQ layers connected by anti-Hebbian inhibitory synapses[10]. In the study we sipmplified the problem by using input carring positional information only. Here, we propose a more natural model consisting of three VQ layers all of which are trained by 4D inputs carrying both of positional and directional information. Computer simulation showed that also in

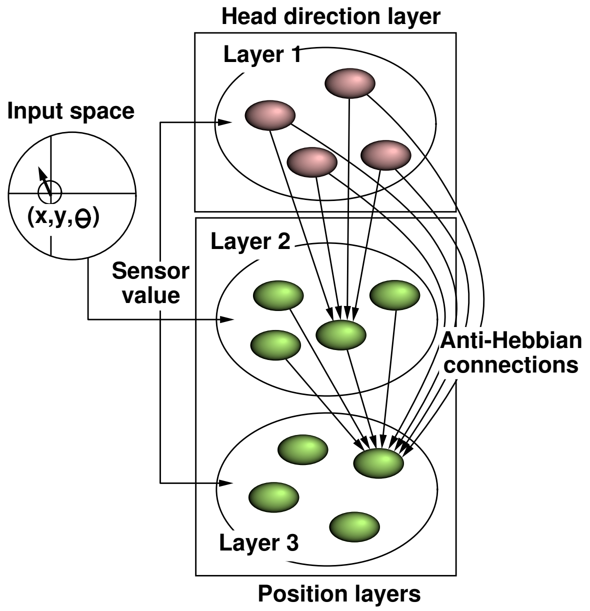


Figure 1: Structure of the proposed model. Three layers receive a same input at a time. Input vector contains two-dimensional positional information and one-dimensional directional information. Layer 1 is for directional information and Layers 2 and 3 are for positional information.

this model grid-cell-like position cells were generated in Layer 3.

2 Layer composition of connected VQs

Our model consists of three layers. Layer 1 is to detect head direction of the object and Layers 2 and 3 to detect position. Any unit in Layer 2 receives inhibitory input from all units in Layer 1 via anti-Hebbian synapses, and any unit in Layer 3 receives inhibitory input from all units in Layers 1 and 2 via anti-Hebbian synapses.

All the cells in the three layers receive a same input vector $\boldsymbol{x}$ which is a function of position (x, y) and head direction θ of the object moving randomly in a room (Figure 1). Layer 1 receives no inputs from Layer 2 or 3, so it works just as an usual VQ model for the input data. On the other hand, the learning processes of Layers 2 and 3 are influenced by the cells of Layer 1 through the anti-Hebbian synapses.

The learning rule is described as follows. In this description, we refer to our algorithm as 'VQ-AH':

(VQ-AH1) Assign random values for reference vectors of all units in the three layers $\boldsymbol{m}_i^{(1)}, \boldsymbol{m}_j^{(2)}, \boldsymbol{m}_k^{(3)}$, where $i = 1, ..., N^{(1)}$, $j = 1, ..., N^{(2)}$, $k = 1, ..., N^{(3)}$, and superscript $^{(1)}$, $^{(2)}$ and $^{(3)}$ stand for the Layers 1, 2 and 3, respectively. Initialize all inhibitory connections $s_{j,i}^{(2)(1)}$, $s_{k,i}^{(3)(1)}$, $s_{k,j}^{(3)(2)}$ to zero.

(VQ-AH2) Set the position (x, y), $(x^2 + y^2 \leq 1)$ and the head direction θ $(-\pi \leq \theta < \pi)$ of the object randomly w.r.t the uniform distribution. Calculate input vector $\boldsymbol{x} = (x, y, \cos\theta, \sin\theta)$ to the VQ layers.

(VQ-AH3) Find the winner, $c^{(1)}$ in Layer 1 for the input $\boldsymbol{x}$:

$$c^{(1)} = \underset{i}{\operatorname{argmin}} \left\{ \left|\left| \boldsymbol{m}_i^{(1)} - \boldsymbol{x} \right|\right|^2 + \theta_i \right\} \tag{1}$$

where θ_i is the threshold or handicap of cell i in Layer 1. Then find the winner $c^{(2)}$ in Layer 2, considering the inhibitory input $s_{j,c^{(1)}}^{(2)(1)}$ from the winner in the superior layer:

$$c^{(2)} = \underset{j}{\operatorname{argmin}} \left\{ \left|\left| \boldsymbol{m}_j^{(2)} - \boldsymbol{x} \right|\right|^2 + s_{j,c^{(1)}}^{(2)(1)} \right\} \tag{2}$$

Finally, find the winner, $c^{(3)}$ in Layer 3 with inhibitory inputs $s_{k,c^{(1)}}^{(3)(1)}$, $s_{k,c^{(2)}}^{(3)(2)}$ from the other two layers

$$c^{(3)} = \underset{j}{\operatorname{argmin}} \left\{ \left|\left| \boldsymbol{m}_k^{(3)} - \boldsymbol{x} \right|\right|^2 + s_{k,c^{(1)}}^{(3)(1)} + s_{k,c^{(2)}}^{(3)(2)} \right\}. \tag{3}$$

(VQ-AH4) After finding three winners $c^{(1)}$, $c^{(2)}$ and $c^{(3)}$, we assign winners' information to variable $y_i^{(1)}$, $y_j^{(2)}$ and $y_k^{(3)}$, respectively:

$$y_i^{(l)} = \begin{cases} 1, & i = c^{(l)}, \\ 0, & \text{otherwise}, \end{cases} \quad (l = 1, 2, 3) \tag{4}$$

(VQ-AH5) Update the reference vectors of the winners in the three layers.

$$\boldsymbol{m}_i^{(l)} := \boldsymbol{m}_i^{(l)} + y_i^{(l)} \alpha^{(l)} \left(\boldsymbol{x} - \boldsymbol{m}_i^{(l)} \right), \quad (l = 1, 2, 3) \tag{5}$$

where $\alpha^{(l)}$ is the learning parameter of the reference vectors.

(VQ-AH6) Update the threshold values of Layer 1

$$\theta_i := \theta_i + \gamma \left(y_i^{(1)} - \frac{1}{N^{(1)}} \right) \tag{6}$$

where γ is the learning parameter of the threshold, and update the inhibitory connections

$$
\begin{aligned}
s_{j,i}^{(2)(1)} &:= s_{j,i} + \beta^{(2)(1)} \left(y_i^{(1)} y_j^{(2)} - \frac{1}{N^{(1)} N^{(2)}} \right), \\
s_{k,i}^{(3)(1)} &:= s_{k,i} + \beta^{(3)(1)} \left(y_i^{(1)} y_k^{(3)} - \frac{1}{N^{(1)} N^{(3)}} \right), \\
s_{k,j}^{(3)(2)} &:= s_{k,j} + \beta^{(3)(2)} \left(y_j^{(2)} y_k^{(3)} - \frac{1}{N^{(2)} N^{(3)}} \right)
\end{aligned} \tag{7}
$$

where $\beta^{(2)(1)}$, $\beta^{(3)(1)}$, $\beta^{(3)(2)}$ are the learning parameters of the inhibitory connections.

(VQ-AH7) Return to (VQ-AH2) and repeat (VQ-AH2)–(VQ-AH6) many times.

As iterative learning proceeds, the learning parameters $\alpha^{(l)}$ $(l = 1, 2, 3)$ and $\beta^{(2)(1)}$, $\beta^{(3)(1)}$, $\beta^{(3)(2)}$ are updated as follows:

$$
\begin{cases}
\alpha^{(l)} := \alpha_0^{(l)} (1 - t/t_{\max}), \quad (l = 1, 2, 3) \\
\beta^{(2)(1)} := \beta_0^{(2)(1)} (1 - t/t_{\max}), \\
\beta^{(3)(1)} := \beta_0^{(3)(1)} (1 - t/t_{\max}), \\
\beta^{(3)(2)} := \beta_0^{(3)(2)} (1 - t/t_{\max}).
\end{cases} \tag{8}
$$

where $\alpha_0^{(l)}$, $\beta_0^{(2)(1)}$, $\beta_0^{(3)(1)}$, $\beta_0^{(3)(2)}$ are initial values of $\alpha^{(l)}$ and correspondig β, and t, $t_{\max}$ are iterative learning times and maximum of iterative learning times, respectively. This update leads stability of learning.

3 Input Data

We need additional learning rules to assign directional information to Layer 1 and positional information to Layers 2 and 3. We divide the whole input sequence into short periods. There are two kinds of periods coming up alternatively. One is a *position-fix* period and the other *direction-fix* period. A direction-fix period consists of some sequential inputs which share a fixed direction θ but have different positions (x, y). A position-fix period consists of some sequential inputs which share a fixed position (x, y) but have different directions θ. For the first input of a direction-fix period the winner of Layer 1 is defined by (1), but for the rest of the inputs in the period the winner remains unchanged, while the winners of Layers 2 and 3 are defined for each inputs by (2) and (3), respectively. For the first input of a position-fix period the winners of Layers 2 and 3 are defined by (2), (3), but for the rest of the inputs in the period the winners remain unchanged, while the winner of Layer 1 is defined for each inputs by (1). In this simulation the either kind of period consists of 15 inputs.

4 Simulation results

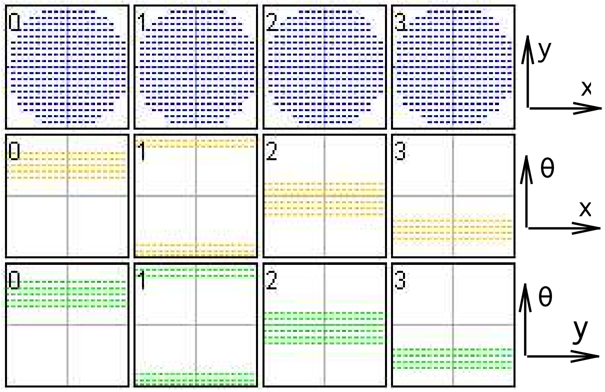
In the equilibrium of the learning process, the learning rule (6) assures that the distribution of the winner on Layer 1 should be uniform over the layer. Similarly, the learning rules (7) assure that the distribution of the two winners on two of the the three layers should be statistically pairwise-independent of each other or the joint probability of the two winners should be uniform. It should be noted this does not mean that the distributions of the three winners are independent.

Each of the three VQ layers consists of 4 units ($N^{(1)} = N^{(2)} = N^{(3)} = 4$). Learning parameters of this experiment are $\alpha^{(1)} = 0.01$, $\gamma = 0.01$; $\alpha^{(2)} = 0.0008$, $\beta^{(2)(1)} = 0.01$; $\alpha^{(3)} = 0.0005$, $\beta^{(3)(1)} = 0.05$, $\beta^{(3)(2)} = 0.005$. Also, $t_{\max} = 800,000$. Learning results are shown in Figure 2. For plots in this figure θ was calculated from Layer 3 and fourth components of the reference vector of each unit by $\theta = \arctan(m_{i3}, m_{i4})$. In Figure 2 (a), (b) and (c), each of the four columns corresponds to each unit. The figure 2 (a) shows that the units in Layer 1 are position insensitive and direction sensitive. The figure 2 (b) and (c) show that Layers 2 and 3 are position sensitive and direction insensitive. Each of the four positional receptive fields of the units in Layer 2 is connected. The whole disk area is divided into four receptive fields. In Layer 3, however, each of the receptive fields of the units consists of four disconnected sub-fields, each of which is contained in one of the four receptive fields of the units in Layer 2. The representation of the positional information in Layers 2 and 3 is similar to two-digit quaternary number system.

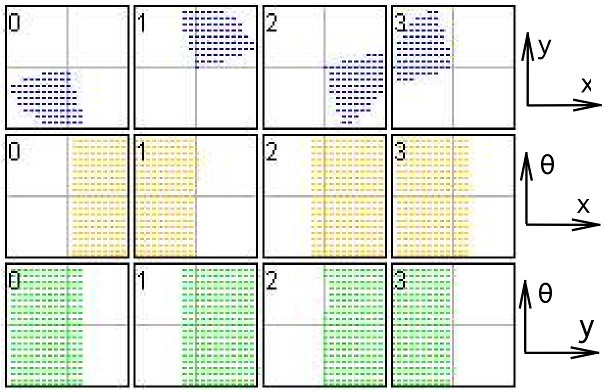
5 Conclusion

We showed that three VQ layers connected via anti-Hebbian synapses can extract directional and positional information separately, and the units in Layer 3 forms disconnected receptive fields, which is similar to what was observed in dMEC of rats by Hafting et al.[9]. Our model suggests that there should be some other position cells with larger and connected receptive fields in somewhere in rats' brain.

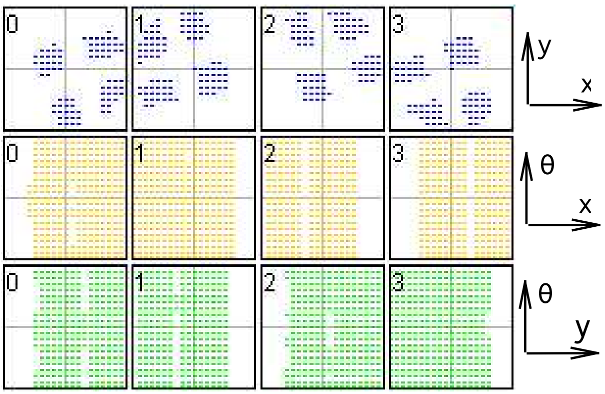
Unlike our preceding results [10] the sub-fields do not show clear hexagonal lattice pattern. This is because we used only four units for Layer 3. we assume



(a) Layer 1: The layer in head direction field



(b) Layer 2: The superior layer in position field



(c) Layer 3: The inferior in position field

Figure 2: Learning results: (a) Each row consisting of three small figures shows the property of each unit in Layer 1. The receptive field in 3D feature space (x, y, θ) is projected in three different ways. The units in Layer 1 show direction sensitivity only. (b) The units in Layer 2 show position sensitivity only. The positional receptive fields are connected. (c) The units in Layer 3 show position sensitivity only. The positional receptive fields are divided into four sub-fields.

that with more units in Layer 3 we can reproduce the hexagonal lattice pattern. The model should be tested with more realistic high-dimensional visual input.

References

[1] Kohonen T (1994) Self-organizing maps., Springer, Berlin

[2] Erwin E, Obermayer K, Schulten K (1995) Models of orientation and ocular dominance columns in the visual cortex: A critical comparison. Neural Comput., 7: 425–468

[3] Shouno H, Kurata K (2001) Formation of a direction map using Kohonen's self-organization map by projection learning. Biol Cybern, 85: 241–246

[4] Kurata K, Oshiro N (2004) Separating visual information into position and direction by SOM. Artificial Life and Robotics, 8 (1): 5–8

[5] Gersho A (1979) Asymptotically Optimal Block Quantization. IEEE Trans. on Inform. Theory, IT-25 (4): 373–380

[6] Linde Y, Buzo A, Gray R (1980) An algorithm for vector quantizer design. IEEE Trans. on Communications, COM-28: 84–95

[7] Gray R (1984) Vector quantization. IEEE ASSP Magazine: 4–29

[8] Oshiro N, Kurata K (2005) Separating visual information into position and direction by two inhibitory-connected SOMs. Artificial Life and Robotics, 9 (2): 86–89

[9] Hafting T, Fyhn M, Molden S, Moser M, Moser E (2005) Microstructure of a spatial map in the entorhinal cortex. Nature, 436: 801–806

[10] Oshiro N, Kurata K, Yamamoto T (2006) A self-organizing model of place cells with periodically distributed receptive fields. Proceedings of the International Symposium on Artificial Life and Robotics (AROB) 11, Beppu, Oita, Japan, Jan 23–25, 2006: OS10-4