Semantic rewiring mechanism of neural cross-modal integration based on spatial and temporal properties of attention

Takamasa Koshizen, So Yamada, Hiroshi Tsujino

Wako Research Centre, Honda R & D Co. Ltd., 1-4-1 Chuo Wako-Shi, Saitama 351-0193 Japan,
koshiz@f.rd.honda.co.jp

Abstract. This paper presents cross-modal integration system which is hypothetically dedicated by a neuronal network where prefrontal cortex is interacted with hippocampus. The computational meaning of the network allows to have its bidirectional structure as well as its rewiring mechanism. We conclude that the rewiring mechanism may particularly be crucial as consolidating spatial and temporal attention induced by the bidirectional neuronal network mainly coordinated by PFC. Seemingly, this is because the semantic role of the rewiring mechanism is to engender proper behavior sequences using the spatiotemporal attention.

1 Introduction

Cross modal integration is crucial in terms of brain computations because it allows us to contribute to endanger adaptive and flexible behavior in various forthcoming events of our daily life. It is however still not elucidated about the neural mechanism of how underlying cross-modal integration can be taken in account of attention mechanisms. With respect to the neuronal network of attention mechanisms, [1,2] showed neurobiologically the attention demanding and modulation process. In CNS*2001, we suggested the computational model of Anterior Cingulate (AC)-Prefrontal (PF)-Posterior Parietal (PP) Cortex in relation to attention demanding and modulation mechanism [3]. Specifically, PP is important for encoding object locations while Temporal Cortex (TC) is for object shape. The computational work also suggested such attention mechanism could be processed by bottom-up and top-down mechansims of attention in order to enhance visual responses and of guiding proper behavioral sequences at any given time. Fig 1. shows presumably a neuronal network which is related to spatial and temporal characteristic of attention. We infer the role of the network centered by PF might consolidate both spatial and temporal characteristic of attention bound at Hippocampus (HC), where the appropriate relationship between behavior sequences and different modalities is learned to come out indelible memory. Indeed, HC encodes both the spatiotemporal relationship by involving PP and PF cortex for example to see [4]. We in this paper aim to provide the computational model of spatiotemporal attention in accordance with the cross-modal integration. Although Sharma et al, suggested neurophysiologically the relationship between the rewiring mechanism and cross-modal integration in ferret [5], we suggest also a self-supervised learning scheme of the rewiring mechanism as well as its meaning presumably. In summary, we speculated the rewiring mechanism may play the crucial role to train the nonlinearity of modalities by a bidirectional neuronal network with spatiotemporal attention, which is dedicated by PF-HC connection.

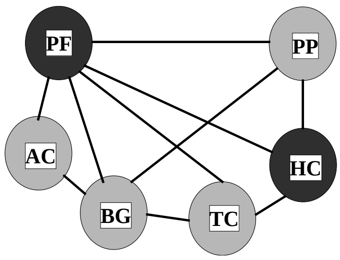

Fig. 1. Schematic of prefrontal-based network

2 Computational model across cross-modal integration

In this section, we elucidate the computation model based on the cross-modal integration. Fig. 2 (right) illustrates the learning architecture of the cross-modal integration system. The system is mainly comprised of the three parts, namely rewiring mechanism, attentive reinforcement learning and behavior control. Essencially, this is designed by the bayesian selective attention (BSA) with sensorimotor system [7]. That is, the system as shown in Fig. 2 (right) is derived from the bidirectional neural network where PF presumably generates the top-down information processing for selecting the attention classes Ω, which are computed by the bottom-up information processing based on the reinforcement learning at basal ganglia (BG). To deal with the cross-modal integration, the rewiring mechanism is especially added with the BSA's sensorimotor system. In the rewiring mechanism each sensory modality allows to be weighted by

Fig. 2. Cross-modal integration system (Left) and BSA model (Right)

the (symmeric or asymmetric) parameter ω through the bidirectional neural network (Fig.2, (left)). In addition, the cross-modal integration system employs the attentive reinforcement learning part of the BSA mechanism where parameters of each attention class across the visual motion and shape modalities, are trained.

Now, we assume the attention class of visual modalities Ω_i^t is mathematically modelled as follows,

$$\Omega_i^t \propto \rho_i(\mathbf{x}^t;\ \theta_i) \tag{1}$$

Where, $\mathbf{x}$ denotes the modality vector belonging in $\mathcal{R}^M$ and the dimensionality of the vector is M. The probability density function $|\rho|$ quantifies the number of population of $\mathbf{x}^t$ which is parameterized by θ_i, where $\rho_i(\mathbf{x}^t) = 1/(2\pi)^{\frac{M}{2}}|\sigma_i|^{\frac{1}{2}} exp{-}0.5(\mathbf{x}^t - \mu_i)^T\sigma_i^{-1}(\mathbf{x}^t - \mu_i)$. In addition, θ_i is comprised of the mean μ_i and standard deviation σ_i of density ρ_i. In our case, θ_i $(i = -1 : shape, +1 : motion)$ is computed by Expectation Maximization (EM) algorithm [7].

$$\hat{\mathcal{O}}_j^t(\mathbf{x}^t) = {\omega_{-1j}}^t\rho_{-1}(\mathbf{x}^t;\ \theta_{-1}) + {\omega_{+1j}}^t\rho_{+1}(\mathbf{x}^t;\ \theta_{+1}) \tag{2}$$

Where $\hat{\mathcal{O}}_j^t$ indicates j^{th} predictive behavioral sequences on $\mathcal{R}^N$. We assume the weight value ω_{ij}^t of the rewiring mechanism can be computed by Allen-Cahn (A-C) equation [8]. The A-C equation is basically applied for phase trasitions whose solution is provided by minimizing the following cost fucntion:

$$\Psi(\omega_{ij}^t) = \int_{B\subset\mathcal{R}^2} \phi(\omega_{ij}^t(\rho_i(\mathbf{x}^t;\ \theta_i)))d\mathbf{x}^t + \epsilon^2 \int_{B\subset\mathcal{R}^2} |\nabla\omega_{ij}^t(\rho_i(\mathbf{x}^t;\ \theta_i))|^2 d\mathbf{x}^t \tag{3}$$

The quantity Ψ can be derived from the double-well potential function ϕ, which is illustrated by Fig.3 (left). Here, we suggest the kind of potential function may be engaged at HC. Then, the updating equation of ω_{ij}^t is

$$\omega_{ij}^t(\rho_i(\mathbf{x}^t;\ \theta_i)) = \Delta\omega_{ij}^t(\rho_i(\mathbf{x}^t;\ \theta_i)) - \frac{1}{\epsilon^2}\frac{\partial\phi(\omega_{ij}^t(\rho_i(\mathbf{x}^t;\ \theta_i)))}{\partial\omega_{ij}} \tag{4}$$

Where Δ denotes the Laplacian. Note that the solution of A-C equation can become the tangential sigmoidal typed function (Fig.3 (right))

$$\omega_{ij}(\rho_i(\mathbf{x}^t;\ \theta_i)) \equiv tanh(\rho_i(\mathbf{x}^t;\ \theta_i)/\epsilon) \tag{5}$$

where ϵ is the perturbation parameter. Moreover, we attempt to im-

Fig. 3. Potential function (Left) and weight value (Right)

plement their mathematical formulas to compute the target location. Fig.4 (left) shows the target (H letter), which is decomposed by the gaussian receptive field (RF) relative to visual shape and motion. Fig.4 (right) illustrates also the learning parameters of the rewiring mechanism in the cross-modal integration system. Fig.5 shows the coherent relationship between densities ρ_i $(i = -1 or + 1))$ that is trained by weights ω_{ij} with our rewiring mechanism. Fig.6 shows

Fig. 4. Target location (Left) and parameters of Cross-modal integration (Right)

Fig. 5. Predicted weight values and coherent density modality

the computation of the cross-modal system across target location problem using Radio-Control (RC) helicopter. If the modality separation gains through the computation, the control state of RC-helicopter is modulated as represented by the visual sensation across target location in Fig.6 (left). Fig.6 (middle, right) represents $belief$ (posterior probability) of the target location relative to the learning stage shown in pattern A (*early*) or B (*late*), which is trained by the BSA mechanism. When the modality separation almost occurs (*e.g.,* $\epsilon = 0.01$), the belief can be most prominant that is resulted by shifting the belief A to B. Note ϵ can be obtained by the variance (uncertainty) of their overlapped beliefs.

Fig. 6. Visual sensor and belief of Target location

3 Conclusion

In this paper, our computational model suggest how the spatial and temporal properties of attention can be taken into account of cross-modal integration system where the two different modalities such as visual motion and shape, are self-supervisedly learned by the rewiring mechanism. Our result also suggests the computation of visual brain where the attention mechanism originates from object shape (what) and it allolws the feebback labeling to find out the exact position of the object (where.)

References

1. Cohen, J.D., Botvinick, M. and Carter C. S., (2000), *Anterior cingulate and prefrontal cortex: who's in control*, Nature Neuroscience, Vol.3, No.5.
2. Robertson, E.M., Tormos, J.M., Maeda, F. and Pascual-Leone, A., (2001), *The role of the dorsolateral prefrontal cortex during sequence learning is specific for spatial information*, Cerebral cortex, Vol.11, pp.628-635.
3. Koshizen, T., Akatsuka, K. and Tsujino, H., (2002), *A computational model of attentive visual system induced by cortical neural network*, Neurocomputing, (In press)
4. White, N.M.and M-C Ouellet, (1997), *Roles of movement and temporal factors in spatial learning*, Hippocampus, Vol.7, pp.501-510.
5. Sur, M. Pallas, S.L. and Roe, A.W., (1990), *Cross-modal plasticity in cortical development: differentiation and specification of sensory neocortex*, Trends Neurosci, Vol.13, pp.227-233.

6. Dempster A.P., Laird N.M. and Rubin, D.B., (1977), *Maximum likelihood from incomplete data via the EM algorithm*, J.Roy.Statist.Soc.Ser.B, Vol.39, pp.1-38.
7. Koshizen, T. and Tsujino, H., (2002), *The brain-like sensorimotor control system*, J. of int. and rob. sys., (In press)
8. Allen, S.M. and Cahn, J.W., (1979), *A macroscopic theory for antiphase boundary motion and its application to antiphase domain coarsening*, Acta. Metall, Vol.27, pp.1085-1095.