Multimodal Representation of Hand Grasping based on Deep Belief Nets

Masaki Ogino* **Takanori Nagura**** **Minoru Asada*,****

*JST ERATO Asada Synergistic Intelligence Project
Yamadaoka 2-1, Suita, Osaka 565-0871, Japan

**Osaka University
Graduate School of Engineering, Department of Adaptive Machine Systems,
Yamadaoka 2-1, Suita, Osaka 565-0871, Japan

In human brain, different sensor information is thought to be processed in different area and integrated in parietal area. Fig. 1 shows a model of neural mechanism for grasping proposed by Oztop et al (Oztop et al., 2006). As shown in this figure, the information of hand and object is processed separately and important features for grasping are extracted in the hierarchical network.

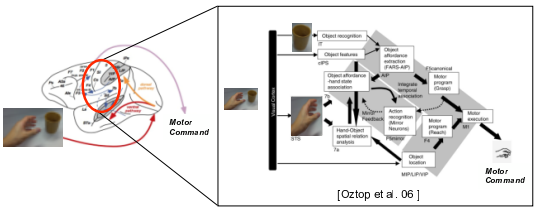

Figure 1: Neural mechanism for grasping proposed by Oztop et al. (Oztop et al., 2006)

In this paper, we aim to construct a hierarchical model for grasping like brain model. The hierarchical model is thought to be plausible as developmental model, because an infant learns its grasping skills gradually in the developing process (Case-Smith and Pehoski, 1992). For this purpose, we adopt deep belief network (DBN), proposed by Hinton, for representing the multimodal information in grasping, in which one modal information is self organized to extract statistical information of given data and different modal information are easily integrated in the hierarchical architecture (Hinton, 2007).

From grasping experiences, four kinds of multimodal information are extracted and input to neural networks for tactile sensing, joint angles, hand images, and object image, respectively. In each modal, raw sensor information are self organized using restricted Boltzmann machine (RBM) and input information is represented in tactile feature, hand feature and object image feature. In this model, it is assumed that information on hand posture such as joint angle and hand image are integrated as hand feature before integration for grasping. Total integration are processed using information during grasping objects. After learning of integration, tactile senses and hand information are recalled from the object image by virtue of DBN properties (Fig. 2).

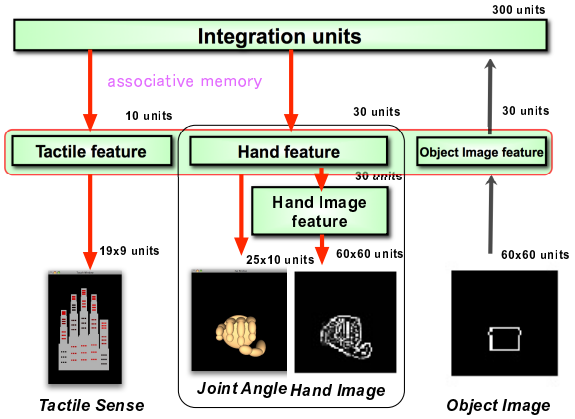

Figure 2: Reconstructing the tactile sensing and hand features from an object image

References

Case-Smith, J. and Pehoski, C. (1992). *Development of Hand Skills in the Child*. Amer Occupational Therapy Assn.

Hinton, G. (2007). To recognize shapes, first learn to generate images. In *Computational Neuroscience:Theoretical Insights into Brain Function.*

Oztop, E., Kawato, M., and Arbib, M. (2006). Mirror neurons and imitation : a computationally guided review. *Neural Networks*, 19:254–271.