Binding visual and tactile sensations based on the invariance in double-touching with self-occlusion

Yuichiro Yoshikawa[1)], Koh Hosoda[1,2)], and Minoru Asada[1,2)]
([1]Dept. of Adaptive Machine Systems, Graduate School of Engineering, Osaka Univ.
[2]HANDAI Frontier Research Center)

Adaptability to the changes in the environment and the robot body itself fundamentally depends on the robot body representation, which is usually given by the designer and therefore fixed in many cases. To adapt its body representation to the changes, the robot should have acquired its own body representation by itself. Assuming that the designer does not give any explicit knowledge about the body, it should construct its body representation only from its uninterpreted sensory data. Yoshikawa et al. [1] addressed the issue of finding its body from its uninterpreted sensors as a first step of acquiring the body representation. According to the idea that the observation of the body is invariant with the changes in its environment, they built a robot that finds its body by a cross modal map to learn the invariance between the visual and the proprioceptive sensors. Although it can learn to judge whether the watched object is its body as well as whether the touched one is its body, it cannot match the tactile information with the watched part and the visual one with the touched part. It is difficult for the robot to bind the sensory data received in the different sensors because they do not always perceive the same phenomenon.

Recently, researchers in the other fields of study focus on such a binding problem that concerns human capability to integrate information of different attributes [2]. Although there are already some models of binding, for example based on attention, firing in synchrony, combinational coding, and so on, it is still not clear how to bind different sensor modalities. The purpose of this study is to propose a constructivist model of human binding problem by building a robot that can watch with its camera head and touch its body with its arm to learn to bind the visual and the tactile sensors.

According to the argument by Yoshikawa et al. [1] concerning the invariance of observation with the state of the robot, the robot can learn to judge whether the watched object is its trunk of the body or something that occludes it. In this study, based on an assumption that only its own body can occlude its sight when it watches objects judged as the trunk by the previous study, we try to build the robot that can bind the watching and the touching posture when the occlusion occurs. Note that it is difficult to learn the binding by applying the Hebbian learning rule simply based on the coactivation of nodes since there can be occlusions not only at the touched point. We propose an exclusive Hebbian rule that considers how many connections the coactivated node has and show a preliminary computer simulation of the robot with 1-DoF arm and 1-DoF camera head to confirm whether the proposed learning rule works.

Reference

[1] Y. Yoshikawa, K. Hosoda, and M. Asada. Does the invariance in multi-modalities represent the body scheme? – a case study with vision and proprioception -. In *Proc. of the 2nd Intl. Symp. on Adaptive Motion of Animals and Machines*, pp. CD-ROM, SaP-II-1, 2003.

[2] A. Treisman. Solutions to the binding problem: Progress through controversy and convergence. *Neuron*, Vol. 24, pp.105-110, 1999.