Does the invariance in multi-modalities represent the body scheme? - a case study with vision and proprioception -

Yuichiro Yoshikawa[*], **Koh Hosoda**[*†] **and Minoru Asada**[*†]

[*] Dept. of Adaptive Machine Systems, Graduate School of Engineering, Osaka University.

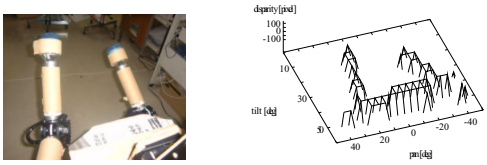
[†] HANDAI Frontier Research Center, Osaka University.

{yoshikawa, hosoda, asada}@er.ams.eng.osaka-u.ac.jp

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. In order for the robot to adapt its body representation to the changes, the robot should have acquired its own body representation by itself. Although it is a formidable problem for robots, biological agents seem to acquire their body representation, called *body scheme* or *body image* [1], without any difficulty. Therefore, a constructive approach by building a real robot which imitates the cognitive developmental processes of biological agents seems a promising way to design intelligent robots which acquire its body representation.

How to find out the body representation in the receptive field without any interpretation by the designer is one of the most fundamental problems of acquiring the body scheme. Previous methods were based on correlation between optic flow and its motor commands [2, 3]. However, instant correlation does not seem to represent the robot body because these methods could not discriminate the robot body and static environment without a prior knowledge about its DOFs. We suggest that the body can be defined by the invariance in the multi-modal sensory data caused by the fact that the sensors are embedded in the rigid robot body while the motion plays a roll of leading experiences to perceive the invariance. Although Yoshikawa et al. proposed a cross modal map which learns to represent the invariance [4], they assumed that the visual patterns were segmented by the designer.

In this paper, we begin with the problem how to find the body surface in the vision. When the robot body is captured in its view, a kind of relation between the vision and the proprioception is invariant with the environmental changes since the body structure usually does not change in a certain period. On the other hand, when the captured areas are not the robot body, the relations between them depend on the environmental changes. Therefore, the robot can find its body by judging whether the multi-modal relation is invariant or not. In order for a robot to learn the invariance, we introduce a cross modal map which consists of a fully-connected network of the sensor nodes with a Hebbian based learning rule and implement it to a robot with the stereo cameras and the arms (Fig. 1(a)). After learning process, the robot acquired a body representation in the cross modal map (Fig. 1(b)) by which it can judge whether the fixating area is its body or not.



(a) (b)

Figure 1: An egocentric view of the robot (a) and acquired body representation in a cross modal map.

References

[1] A. Iriki, M. Tanaka, and Y. Iwamura, "Coding of modified body schema during tool use by macaque postcentral neurons", *Neuroreport*, Vol.7, pp.2325-2330, 1996.

[2] M. Asada, E. Uchibe, and K. Hosoda, "Cooperative behavior acquisition for mobile robots in dynamically changing real worlds via vision-based reinforcement learning and development", *Artificial Intelligence*, Vol.110, pp.275-292, 1999.

[3] P. M. Fitzpatrick and G. Metta, "Toward Manipulation-Driven Vision", In *Proc. of the 2002 IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, pp.43-48, 2002.

[4] Y. Yoshikawa, H. Kawanish, M. Asada, and K. Hosoda, "Body Scheme Acquisition by Cross Map Learning among Tactile, Image, and Proprioceptive Spaces", In *Proc. of the 2nd Intl. Workshop on Epigenetic Robotics*, pp.181-184, 2002.