International Journal of Innovative Computing, Information and Control
Volume **5**, Number **11(B)**, November **2009**

ICIC International ©2009 ISSN 1349-4198
pp. **4167–4176**

MECHANISM OF FORMATION OF VISION BASED ON LEARNING OF CORRELATION BETWEEN SENSATION AND MOTION

Hirofumi Fukumoto and Takashi Toriu

Department of Physical Electronics and Informatics
Graduate School of Engineering
Osaka City University
3-3-138 Sugimoto, Sumiyoshi-ku, Osaka 558, Japan
fukumoto@ip.info.eng.osaka-cu.ac.jp; toriu@info.eng.osaka-cu.ac.jp

Received October 2008; revised February 2009

Abstract. *In this paper, we support the scheme that visual cognition is organized by natural neural networks based on experience. According to this scheme, networks learn correlations between temporal variations of visual stimulation and their own motion to organize visual cognition. We propose a neural network mechanism that learns correlation between two sensations of visual stimulation and self-motion without any prior visual-spatial knowledge. The main novelty of the present work is that the proposed neural network system does not use human's visual-spatial knowledge such as optics, or positional relationship of pixels, but uses only simple stimulations that have time-series structure obeying natural laws. Through experiments, we show that the proposed network can combine two sensations of visual stimulation and self-motion correctly, and can recall patterns of one side using the input patterns of the other side.*
Keywords: Vision, Self-motion, Visual-spatial, Learning, Organization, Cognition

1. **Introduction.** What is a conscious mind or what do you mean by cognition? This problem has been discussed since ancient times. In engineering, the problem has been addressed as a problem of artificially constructing functions of the conscious mind or cognition. In studies of AI, symbolism [1] and connectionism [2, 3, 4, 5] have been discussed. In the 70s and 80s, top-down approaches were attempted, were the external world represented in advance by symbols. This representation is based on human knowledge of the world. Since human knowledge is not complete, there are necessarily some gaps between the symbolized and the real world. In addition, a more fundamental problem exists, as Harnad [6] suggested. It is called the symbol grounding problem-a symbol cannot be completely defined by symbols. Some symbols must be dynamically connected to real objects in the real world. As a solution, Brooks [7] proposed behavior-based robotics in late 1980s. According to the approach, intelligent actions can certainly emerge based on the relation between a body and its environment [8]. However, the robot does not seem to have any representation of the external world, its behavior is too directly connected to environment. In fact, even if the robot (actor) successfully shoots a ball into a goal, it does not know anything about the ball, the goal, and the relation between them. No conscious mind or cognition seems to emerge from the approach.

In this paper, we focus on how visual cognition can be generated from the relationship between a robot 's actions and its sense of vision. In particular, we focus on the problem of how the relationship is represented in a neural net work. Specifically, we propose a primitive learning mechanism that organizes a vision system spontaneously by taking into account of correlations between the inputs of a sensor and its own motion.