Motion-triggered human-robot synchronization for autonomous acquisition of joint attention

Hidenobu Sumioka, Koh Hosoda[‡], Yuichiro Yoshikawa, and Minoru Asada[‡]
Department of Adaptive Machine Systems,
[‡]HANDAI Frontier Research Center,
Graduate School of Engineering, Osaka University
2-1 Yamadaoka, Suita, Osaka, 565-0871 Japan
{*sumioka, yoshikawa*}*@er.ams.eng.osaka-u.ac.jp*
{*hosoda, asada*}*@ams.eng.osaka-u.ac.jp*

***Index Terms*— joint attention, motion information, real time learning in the real world, asynchrony, bootstrap learning**

Joint attention, a behavior to attend to an object to which another person attends, is an important element not only for human-human communication but also human-robot communication. Building a robot that autonomously acquires the behavior is supposed to be a formidable issue both to establish the design principle of a robot communicating with humans and to understand the developmental process of human communication.

To accelerate learning of the behavior, the motion synchronization among the object, the caregiver, and the robot is important since it will ensure the information consistency between them. In the previous work, however, it was not taken into account [2], [3] or ensured by the human operation [1]. For realizing autonomous rapid learning of the robot, it should synchronize its motion with those of the object and the caregiver by itself.

In this paper, we extend the architecture proposed in [3] by taking the motion information into accout (see **Fig. 1**). We utilize motion information for two purposes: to segment motion sequence for synchronization and to determine the gaze direction. The robot will change the gaze direction when it observe motion in its view. If the caregiver picks up and look at an interesting object to attend to, the robot will change its gaze to the object because the caregiver's gaze is fixed and the object that is picked up moves. This enable it to acquire the correct relation between the caregiver's face pattern and position of the object. If the robot does not observe any motion in its view, it will change its gaze to the caregiver because the caregiver's face is very interesting. The gaze direction is determined based on the preference measures of objects and of the caregiver's face. The motion information enables the robot to take turns by itself to acquire consistent information.

We conducted several experiments on a real robot to test whether it could learn joint attention efficiently in the context that a caregiver moved an object autonomously. From **Fig. 2**, we can see that it needs only about 20 minutes to perform joint attention with high success rate. We might conclude that it could efficiently perform learning of joint attention due to the synchronization by the proposed method.

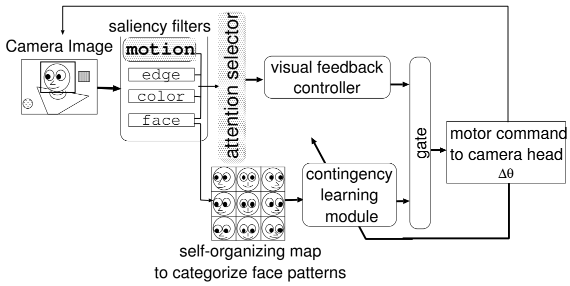

Fig. 1. The proposed mechanism for learning of joint attention.

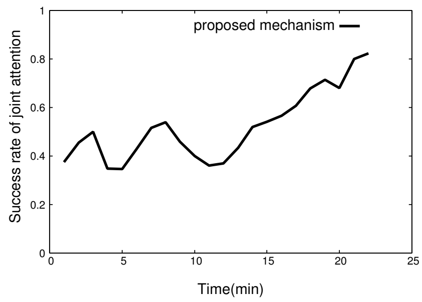

Fig. 2. The time course of success rate of joint attention.

References

[1] Y. Nagai, K. Hosoda, A. Morita, and M. Asada, "A constructive model for the development of joint attention," Connection Science, 15(4),pp211-229,12,2003.
[2] E. Carlson and J. Triesch, "A computational model of the emergence of gaze following," In H. Bowman and C. Labiouse, editors, Proc. of the 8th Neural Computation and Psychology Workshop, Progress in Neural Processing. World Scientific,2003.
[3] K. Hosoda, H. Sumioka, A. Morita, and M. Asada, "Acquisition of Human-Robot Joint Attention through Real-time Natural Interaction," the IEEE/RSJ International Conf. on Intelligent Robots and Systems,pp2867-2872,2004.