Emergence of Joint Attention
based on Visual Attention and Self Learning

Yukie Nagai[1], Koh Hosoda[2], Akio Morita[3], and Minoru Asada[4]

Dept. of Adaptive Machine Systems, Graduate School of Engineering, [2,4]HANDAI Frontier Research Center,

Osaka University, Suita, Osaka, 565-0871 Japan

e-mail: {[1]yukie, [3]morita}@er.ams.eng.osaka-u.ac.jp, {[2]hosoda, [4]asada}@ams.eng.osaka-u.ac.jp

1.   Introduction

This paper presents a method which develops the ability of joint attention of a robot based on visual attention, a capability to find and attend to a salient object in the robot's view, and self learning, a capability to learn a sensorimotor coordination when the visual attention succeeds. Through the learning, the robot finds the correlation of the sensorimotor coordination when joint attention with a human caregiver is successfully realized. As the result, the correlation allows the robot to acquire the ability of joint attention even if multiple objects are set in the environment.

2.   Emergence of Joint Attention

Figure 1 shows the proposed method based on the visual attention and the self learning. It is supposed that a caregiver attends to an object, and the robot observes the caregiver. First, the robot captures the camera images $I_{L, R}$ and the camera angles $\theta_{pan, tilt}$, and extracts distinguishing image areas by the salient feature detector. The detected object's features are sent to the visual feedback controller, and the face-like stimuli of the caregiver are sent to the learning module. Then, the visual feedback controller outputs a motor command $^{VF}\Delta\theta$ to attend to an interesting object. At the
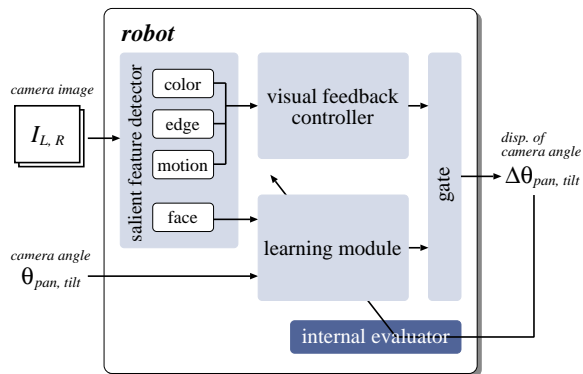
same time, the learning module outputs $^{LM}\Delta\theta$ based on the caregiver's face image and $\theta_{pan, tilt}$. Next, the gate arbitrates the motor command $\Delta\theta_{pan, tilt}$ between $^{VF}\Delta\theta$ and $^{LM}\Delta\theta$. As the result, if the robot can attend to an interesting object, the internal evaluator drives the learning of the sensorimotor coordination in the learning module. Through the learning process, the robot can find the correlation of the sensorimotor coordination only when joint attention succeeds.

3.   Experiment

The experiments are executed in a single-object situation, where an object is set to get away from the caregiver step by step, and in a multiple-objects situation, where some objects are set near the caregiver. The learned performance of the robot in the single-object situation is shown in Figure 2. This result indicates that the robot can acquire the ability of joint attention based on the proposed method.
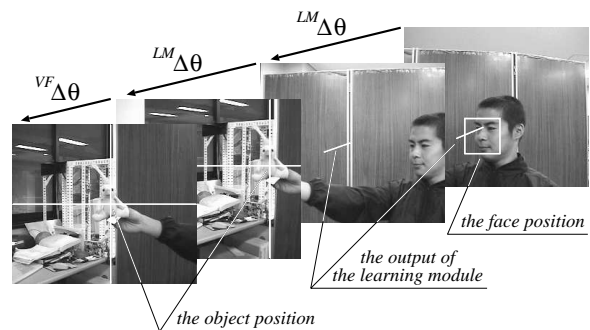

Figure 2: The learned performance of the robot

4.   Conclusion

The experimental results showed the validity of the proposed method in the simplified situation. Our future work is to examine whether the proposed method enables the robot to acquire the ability of joint attention in a never controlled situation.

Figure 1: The proposed method