Head Tracking and Gesture Recognition in Museum Guide Robots for Multiparty Settings

Yoshinori Kobayashi*+, Takashi Shibata*, Yosuke Hoshi*,
Yoshinori Kuno*, Mai Okada*, Keiichi Yamazaki*, Akiko Yamazaki**
*Saitama University, Japan, **Tokyo University of Technology, Japan
+*yosinori@cv.ics.saitama-u.ac.jp*

Abstract. In this paper, we describe head tracking and gesture recognition techniques in a museum guide robot. The vision system we propose has been developed to deal with challenges we discovered through experiments using our robot guide in an actual art museum. Based on these observations, we developed a prototype guide robot that can explain an exhibit to multiple visitors.

Recently, several museum guide robots that engage visitors in their talk in one-on-one settings have been proposed. In particular, K. Yamazaki et al. (2009) reported that the guide robot increases visitor engagement when it turns its head towards the visitor at interactionally relevant places (or TRPs, transition relevance places) in its talk. Through experiments using our guide robot in an actual museum, we found that we need to consider not only one-on-one settings but also multiparty settings, as our guide robot often attracted multiple visitors at the same time. In this study, we propose a new vision system that enables the robot to explain exhibits to multiple visitors. In our previous experiment at an actual museum, there were some cases when visitors appeared to lose interest, such as when the robot provided the same explanation to the same visitor or changed the addressee (target person) during the explanation. Based on ethnographic observations of human museum guides, we found that facework plays an important role in engaging multiple visitors in their talk. We also found that the visitor who is the addressee of the human guide may avoid eye contact to display a negative answer. So it is important for a guide robot to be able to track and identify multiple visitors and recognize their implicitly displayed intentions.

In addressing this issue, we propose a new vision system that enables a museum guide robot to interact with multiple visitors (Figure.1). This system does principally three things: it tracks and identifies visitors, and recognizes their displayed intentions. 1) In tracking visitors, the system first detects faces using the frontal face classifier based on AdaBoost and Haar-like features. The system then starts to track each face using a particle filter framework. When the detected faces have been tracked, the system does not start to track these faces. We divide face region into 4x4 blocks and calculate the optical flow vector for each block between consecutive frames. We use these optical flow vectors for the motion prediction in the particle filter framework. Hypotheses are evaluated using multiple classifiers, each of which is trained to detect human heads in a particular direction. The most suitable classifier is selected adaptively by considering face direction with respect to the camera. In this way, we track not only visitor head position but also head direction simultaneously. 2) In identifying visitors, the system records color histograms of each image region corresponding to the clothes and motion histories of each visitor. The robot can continue to explain to the same visitor even when the system temporarily loses track of the head due to occlusion. 3) In recognizing visitors' intentions, as displayed in head nodding and head shaking, we calculate the mean value of the vertical and horizontal components of the optical flow vectors and use them to recognize these actions. The system evaluates the score of positive intentions by considering the context of the robot's talk especially focused on synchronization with TRPs.

We developed a prototype robot that can explain paintings to multiple persons. The robot turns its head towards the person who displays positive intention. We conducted experiment in which the robot explains a painting to 15 participants and asked the participants at the end to answer the question, "Did you feel that the robot was looking at you and other participants?" The participants' answers suggest that the robot gazes towards participants in an effective way.

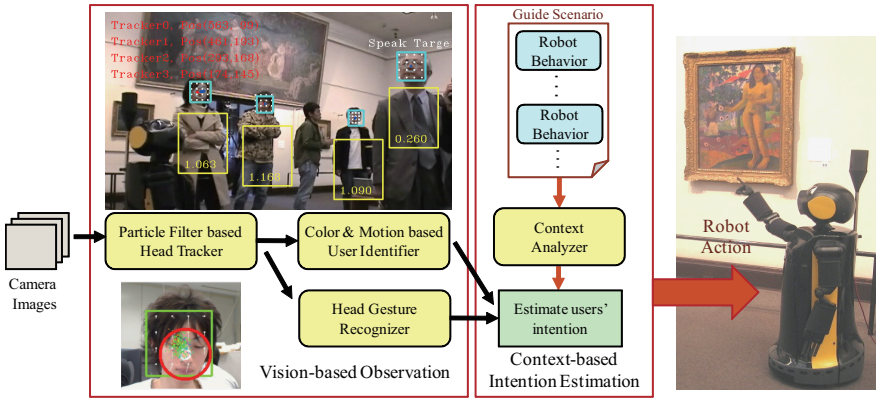

Figure 1. Vision system for museum guide robot.

References

K. Yamazaki, A. Yamazaki, M. Okada, Y. Kuno, Y. Kobayashi, Y. Hoshi, K. Pitsch, P. Luff, D. Lehn, and C. Heath (2009): 'Revealing Gauguin: Engaging Visitors in Robot Guide's Explanation in an Art Museum'. In: *CHI2009*. pp. 1437–1446.