Attention Control System Considering the Target Person's Attention Level

Dipankar Das[1], Mohammed Moshiul Hoque[1], Yoshinori Kobayashi[1,2] and Yoshinori Kuno[1]
[1]Graduate School of Science and Engineering, Saitama University
255 Shimo-Okubo, Sakura, Saitama 338-8570, Japan
[2]JST, PRESTO, 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan
Email: {dipankar, moshiul, yosinori, kuno}@cv.ics.saitama-u.ac.jp

***Abstract*—In this paper, we propose an attention control system for social robots that attracts and controls the attention of a target person depending on his/her current attentional focus. The system recognizes the current task of the target person and estimates its level of focus by using the "task related behavior pattern" of the target human. The attention level is used to determine the suitable cues to attract the target person's attention toward the robot. The robot detects the interest or willingness of the target person to interact with it. Then, depending on the level of interest, the robot displays an awareness signal and shifts his/her attention to an intended goal direction.**

***Keywords*—attention control; attention level; human-robot interaction**

I. Introduction

In HRI, attention control refers to the robots' capability of gaining the target person's attention toward it and shift his/her attention to an intended goal direction [3]. To gain or attract the target person's attention, the robot must show some sign of intentionality through verbal or nonverbal actions based on the urgency of the situation. If it urgently needs to control the attention, the robot cannot help using some strong cues (such as voice, reference term, touch etc.) without considering the target persons situation. However, if not so urgent, the robot should try not to disturb the persons current work as far as time allows. It should observe what s/he is doing and try to attract his/her attention when his/her attention level to the current work becomes low and shifts his/her attention to a goal direction. In this paper, we propose such an attention control system consisting of four active phases: attention attraction, interest/willingness detection, awareness generation and eye contact, and attention shift.

II. Attention Control System

Fig. 1 shows the states of the proposed attention control system. When the target person (TP) is involved in a task, the robot determines the current attention level on the task (ALT) of the person. Then depending on the level (low or high) of current focus, the robot generates an appropriate attention attraction (AA) signal (weak or strong) to attract the attention of the target person toward it. If the attention level toward the robot (ALR) is low, the robot re-starts the process from the beginning or end the process. On the other hand, if ALR is medium, the robot generates an awareness signal and waits for ALRs turning high. When ALR of the target person becomes high, the robot makes eye contact with him/her and shifts his/her attention to an intended goal direction.

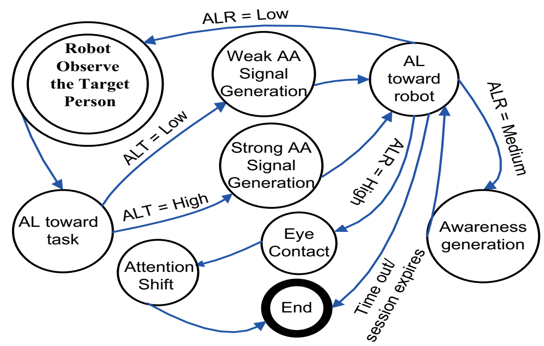

Fig. 1: States of the attention control system.

Attention Attraction. The robot sends control signals to attract the target person's attention based on his/her level of current focus. To measure the current focus and its level, we use the task dependent behavioral pattern analysis. Given a video sequence, we extract the histogram of orientation gradient (HOG) feature for each frame. The HOG features are combined for 10 consecutive frames to build a HOG feature pattern: $HOG_P = F_0 + \sum_{i=1}^{9} |F_{i-1} - F_i|$, where F_0 and F_i are the HOG features of the first and ith frames, respectively. A multi-class support vector machine (SVM) with HOG_P feature is used to recognize the current task.

After detecting the task, the robot observes the person to detect his/her loss of current attentional focus, $A_L \leftarrow$ {Changes in current focus}. In the experiment, we focused on three types of tasks and their related change of pattern in attentional focus: $A_{L,read} \leftarrow H_M \vee P_T$, $A_{L,write} \leftarrow H_M \wedge W_S$, and $A_{L,browse} \leftarrow H_M$, where H_M, P_T, and W_S are 'head movement', 'page turn over', and 'stop writing' behavior patterns, respectively. We use the resultant magnitude of the optical flow pattern in the changed area to detect these patterns. For each task the level of sustained attention is classified into two categories: low and high. If any of above three conditions is true and this situation continues for greater than or equal to 3 frames, then the level of attention is considered low for the corresponding task. On the other hand, if no cues for loss of attention are detected, and the current attentional focus remains on the task, then the attention is considered high. From the survey of psychology, HRI literature, and our preliminary experiment, we choose a head turning action as the weak signal, which is used when the attention level is low.

TABLE I: Task detection performance

Task	Browsing	Reading	Writing
Browsing	**3101**	3	0
Reading	28	**2448**	428
Writing	0	4	**2728**

TABLE II: Attention level detection performance

Task	Browsing	Reading	Writing
#of test low attention	14	16	13
#of detected attention	12	15	10
#of false positive	10	8	12
#of false negative	2	1	3

We use a head shaking action when the attention level is high and the robot needs to attract the attention of the target person, because abrupt object motion draws people's attention [4].

Interest/Willingness Detection. We use a vision-based technique proposed in our previous study [1] to detect the level of attention (interest/willingness) of the target person toward the robot. We use the term *attention level* to refer to the three levels of attention (low, medium, and high) that the robot perceives from the visual information of the target person.

Awareness Generation and Eye Contact. The robot observes the human as long as the attention level toward it is low. When the attention level becomes medium, the robot generates an awareness signal by raising its head toward the target person. The awareness signal indicates that the robot intends to communicate with the person. When the level of attention becomes high, the robot starts blinking its eyes to make eye contact with the target person.

Attention Shift. We use the combination of eyes and head turn actions to shift the attention to the target object because it is shown to be effective for attention shifting [2].

III. Experimental Result

We conducted a series of experiments to investigate the performance of the vision-based system to attract and shift the attention of the target person. To collect experimental data, we asked 14-non-paid participants (11 males and 3 females) for three different tasks: browsing, reading, and writing. All tasks were recorded using a video camera.

Task Detection Performance. The recorded task videos are divided into test and train sample videos. There were 6469 training and 8740 test sample videos. Among 8740 test samples 8277 task samples were correctly detected with 94.70% accuracy. Table I shows the task specific recognition performance and cross task confusion of the system.

Sustained Attention Level Detection. The level of sustained attention is detected using a threshold value for the resultant magnitude of the optical flow pattern. The low and high attention level is defined in the test video with starting and ending frame number. Table II shows the system performance to detect the attention level in the test videos.

Evaluation of the System on the Experimental Robot. To evaluate the effectiveness of our proposed method, we designed and implemented two types of robot behaviors:

TABLE III: Questionnaire results (7 is the most effective)

	ASR		ASRA	
	Mean	SD	Mean	SD
Q1	4.36	1.48	6	1.54
Q2	5.29	2.37	5.57	1.49

(i) *Attention Shift Robot (ASR)*: The robot makes eye contact by simply blinking the eyes and then shifts the participant's attention while his/her attention level toward the robot is high, (ii) *Attention Shift Robot with Awareness (ASRA)*: The robot generates the awareness signal when the attention level toward it is medium, makes eye contact and shifts attention.

For subjective evaluation, we asked participants to answer (rating on a Likert scale from 1 to 7) the following questions: **Q1:** Did you make eye contact with the robot when you shift your attention toward it? **Q2:** Did the robot shift your attention to the target object? Table III shows the results of the questionnaire measures which represent means, and standard deviations (SD) for each method. We conducted a repeated measure of analysis of variance (ANOVA) for both questions, Q1 and Q2. For Q1, the differences between the two methods are statistically significant ($F(1;13) = 23.15; p = 0.0003; n^2 = 0.3251$). The result reveals that ASRA is more effective at making eye contact than ASR. For Q2, the ANOVA on target participant's scores does not show significant differences between the two methods ($F(1;13) = 3.06; p = 0.1039; n^2 = 0.0112$). This is because, for attention shift, the robot exhibits the same behavior in both methods. From the videos, we examined the performance of the proposed system. The accuracy for attention level detection toward the robot is 89% (25 times among 28 samples). Among these 25 trials, attentions are successfully shifted 25 times with accuracy of 100%.

IV. Conclusion

In this paper, we proposed a vision based technique to attract and shift the target person's attention depending on his/her level of attention. The target person may be engaged in different tasks (such as, reading, writing, browsing etc.) and the robot observes his/her behavior. The robot chooses the suitable time to attract the target person's attention depending on the current attentional task and its level. We compared two types of robot behaviors to capture the attention of the person. The results indicate that our proposed robot (ASRA) is effective in capturing and holding the target person's attention.

Acknowledgment

This work was supported by JST, CREST.

References

[1] D. Das, M. M. Hoque, T. Onuki, Y. Kobayashi, and Y. Kuno, "Vision-based attention control system for socially interactive robots," in *RO-MAN*, Paris, France, Sep. 9-13, 2012 pp. 496-502.

[2] M. M. Hoque, T. Onuki, D. Das, Y. Kobayashi, and Y. Kuno, "Attracting and controlling human attention through robot's behaviors suited to the situation," in *HRI*, Boston, MA, USA, March 05 - 08 2012, pp. 149–150.

[3] M. M. Hoque, T. Onuki, E. Tsuburaya, Y. Kobayashi, Y. Kuno, T. Sato, and S. Kodama, "An empirical framework to control human attention by robot," in *ACCV Workshops (1)*, ser. LNCS, Queenstown, New Zealand: Springer, November 8-9 2010, pp. 430–439.

[4] W. James, *The Principles of Psychology*. Dover, New York, 1950.