Tracking Gaze over Time in HRI as a Proxy for Engagement and Attribution of Social Agency

Paul Baxter, James Kennedy, Anna-Lisa Vollmer, Joachim de Greeff, Tony Belpaeme
Centre for Robotics and Neural Systems
Cognition Institute
Plymouth University, U.K.
paul.baxter@plymouth.ac.uk

ABSTRACT

In this contribution, we describe a method of analysing and interpreting the direction and timing of a human's gaze over time towards a robot whilst interacting. Based on annotated video recordings of the interactions, this *post-hoc* analysis can be used to determine how this gaze behaviour changes over the course of an interaction, following from the observation that humans change their behaviour towards the robot on the time-scale of individual interactions. We posit that given these circumstances, this measure may be used as *a* proxy (among others) for engagement in the interaction or the human's attribution of social agency to the robot. Application of this method to a sample of unstructured child-robot interactions demonstrates its use, and justifies its utilisation in future studies.

Categories and Subject Descriptors

H.1.2 [**Models and Principles**]: User/Machine Systems; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

General Terms

Experimentation, Measurement, Theory

Keywords

Child-Robot Interaction; Gaze Direction; Sandtray; Unstructured Social HRI

1. INTRODUCTION

In social human-robot interaction, an autonomous robot must gain and maintain the engagement of an interactant, whilst taking into account the context of the interaction (be it environmental, task, etc). There are many important facets to this maintenance of engagement; one central facet is the gaze of the interactant, both in terms of direction and timing, e.g. [4]. In addition to this, recent results have also emphasised the adaptation of human's behaviour in response to robot behaviour [5], indicating that this variability needs to be taken into account in analysis. These aspects taken together demonstrate the need to account for engagement over the course of an interaction; gaze provides a means of analysis of this when treated as a measurement of human behaviour. Given the difficulty in providing such a characterisation in real-time, *post-hoc* analysis nevertheless enables such an account to be formulated.

Figure 1: The 'Sandtray': a touchscreen device that provides a task context for social HRI to facilitate naturalistic human behaviour.

In this context, dyadic human-robot interactions with no explicit interaction structure imposed are of particular interest, with our studies centred on the use of the 'Sandtray' system [1] (figure 1). With no such imposed structure, it is hypothesised that more naturalistic interaction characteristics are more readily attainable [3], thus enabling the emergence of social interaction to be studied [2]. However, the resulting analysis is complicated by this lack of structure, adding to the difficulty in characterising the interaction (e.g. engagement [6]). In this contribution, we present a method of *post-hoc* analysis that can provide an objective characterisation of the gaze behaviour of the human towards the robot during an interaction, and explore how this analysis can shape the assessment of the human's behaviour, and the state of the interaction itself.

2. CHARACTERISING GAZE OVER TIME: A CASE STUDY

The basic premise of the method proposed here is that the change in the human gaze behaviour over the course of an interaction can provide useful information regarding the state of the interaction, and also the attitude of the human towards the robot. Assuming that from the recorded interac-

HRI'14, March 3–6, 2014, Bielefeld, Germany.
ACM 978-1-4503-2658-2/14/03.
http://dx.doi.org/10.1145/2559636.2559829.

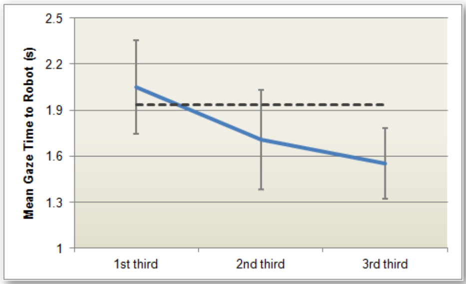

Figure 2: Mean length of each gaze to the robot in the three mean thirds of the interactions analysed. Dotted line shows mean gaze length over the whole interaction. Error bars show SEM.

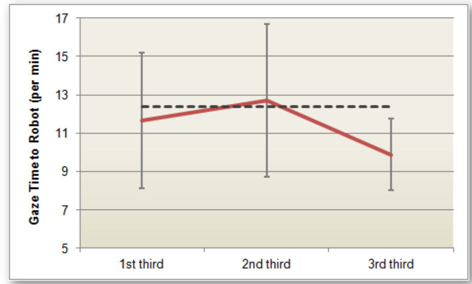

Figure 3: Length of gaze per minute to robot in the three mean thirds of the interactions analysed. Dotted line shows mean length of gaze (in seconds) per minute over the whole interaction. Error bars show SEM.

tions it is possible to distinguish and code the gaze direction of the human, the method essentially entails the calculation of the mean of a measurement of gaze within a time-window, such that the value can be compared across time-windows (e.g. normalised duration of gaze in each of the time-windows, enabling comparison across subjects). For relatively short interactions, one relevant comparison is between the beginning and end of the interaction by splitting the interaction into three phases, thus enabling standard statistical tests to be applied to the differences in means.

This method was applied to video data from unstructured child-robot interactions, in order to illustrate its potential as a means of assessing the child's interaction with the robot. Video data from a number of unstructured child-robot interactions in the context of the Sandtray were coded (see [1] for details); data from four interactions were analysed (mean length of interaction: 115s, *sd=16.4s*). Despite there being no significant difference between the values obtained for this case study, a decreasing profile is visible for both of the gaze measures analysed: mean length of gaze from human to robot (figure 2) decreases from the first to final third of the interactions, as does the length of gaze (normalised per minute, figure 3). It is instructive to compare these plots to the mean value for the interaction: in both cases there is an indication that there are changes within the interaction that are lost if only the average over the interaction were to be considered.

Previous work posited that the children in these interactions viewed the robot as a potentially social actor, resulting in the emergence of turn-taking despite the robot not using contingent behaviours [1]. The present results modify this perspective slightly as they indicate that over time, the children look less at the robot. This may be seen as a decrease in engagement with the robot, possibly as the limitations of the robot's social behaviour are experienced - indeed, the attribution of social agency may also decrease over time as a result. Further exploration of this effect is required, as it has fundamental consequences for child-robot interaction in particular [2]. However, the application of the proposed analysis method (or similar) will facilitate this process.

3. PERSPECTIVES

As described here, there are a number of limitations to the proposed *post-hoc* metric. Firstly, it assumes that the interactions analysed are of near-to-equal length, otherwise the three stages of individual interactions will be stretched or squashed (relatively). This can be circumvented by using constant-length time-windows instead of proportions of interaction lengths (although this is similarly problematic for comparing interactions of unequal length). Secondly, the data presented here is human-focused, rather than taking into account what the human's behaviour is contingent upon. While not conducted for the present example, this can be remedied by instead using measures of human gaze direction in response to robot behaviours (e.g. overlaps with movement or utterances).

Despite these limitations, the main point of the proposed analysis is that the change in gaze behaviours can be tracked over time, and that these changes can themselves provide a useful characterisation of the interaction, and aid in the interpretation of the human's behaviour. It may thus also be applied as a means of comparing the efficacy of different experimental conditions or systems in maintaining engagement over interaction time, as the profiles of the mean time-windowed gaze measure can be objectively compared.

4. ACKNOWLEDGMENT

This work is funded by the EU FP7 ALIZ-E project (grant 248116).

5. REFERENCES

[1] P. Baxter et al. Emergence of Turn-taking in Unstructured Child-Robot Social Interactions. In *HRI'13*, pages 77–78. ACM Press, 2013.

[2] T. Belpaeme et al. Child-robot interaction: Perspectives and challenges. In *ICSR'13*, volume 8239, pages 452–459. Springer, 2013.

[3] J. Kennedy et al. Constraining Content in Mediated Unstructured Social Interactions: Studies in the Wild. In *AFFINE'13*. IEEE Press, 2013.

[4] B. Mutlu et al. Footing In Human-Robot Conversations: How Robots Might Shape Participant Roles Using Gaze Cues. In *HRI'2009*, pages 61–68. ACM Press, 2009.

[5] M. Nalin et al. Children's adaptation in multi-session interaction. In *Ro-MAN 2012*. IEEE Press, 2012.

[6] C. Rich et al. Recognizing engagement in human-robot interaction. In *HRI'2010*, pages 375–382. IEEE Press, 2010.