Development of Culture-specific Gaze Behaviours of Virtual Agents

Tomoko Koda[1,2], Taku Hirano[1] and Takuto Ishioh[2]

[1]*Faculty of Information Science and Technology, Osaka Institute of Technology, Osaka, Japan*
[2]*Graduate School of Information Science and Technology, Osaka Institute of Technology, Osaka, Japan*

Abstract: Gaze plays an important role in human-human communication. Adequate gaze control of a virtual agent is also essential for successful and believable human-agent interaction. Researchers in intelligent virtual agents have developed gaze control models by taking account of gaze duration, frequency and timing of gaze aversion. However, none of them have considered cultural differences in gaze behaviours. We aim to investigate cultural differences in gaze behaviours and their perception, by developing virtual agents with Japanese gaze behaviours, western gaze behaviours, their hybrid gaze behaviours, and full gaze behaviours, and compare their effects on the impressions of the agents and interactions. This position paper proposes our research agenda, describes the implemented gaze models, and our experimental design.

1 INTRODUCTION

Intelligent virtual agents (IVAs) that interact face-to-face with humans are beginning to spread to general users across cultures, and IVA research is being actively pursued. IVAs require both verbal and nonverbal communication abilities to achieve natural interaction with humans. Among those non-verbal behaviours, gaze plays an important role in our social interactions such as controlling the flow of a conversation, indicating interest and intentions, and improving listener's attention and comprehension (Argyle and Cook, 1976; Bayliss et al., 2006).

As in humans, virtual agent's gaze behaviour is also important to provide natural interaction. Previous research on modelling gaze behaviour of virtual agents were conducted to make appropriate turn management (Pelachaud and Bilvi, 2003), to figure out where to look at (Lee et al., 2007), to make idle gaze movements (Cafaro et al., 2009), to express social dominance by gaze (Bee, 2010), to compare sensitivity to amount of gaze according to evaluators' shyness level (Koda et al., 2016), and what the adequate amount of gaze is to facilitate interaction (Ishii et al., 2006; Ishii et al., 2008), all of which report modelling realistic human gaze behaviour to an agent resulted in more natural and smooth interaction.

However, none of the above IVA research has addressed and implemented cultural difference in gaze behaviours, while researchers in psychology report cultural difference in gaze behaviours and their perception. We believe there is a strong need to develop enculturated agents by making them exhibit culture-specific non-verbal behaviours such as gaze.

In terms of culture-specific gaze behaviours, there are findings from observation and video analysis of human-human and human-agent interactions that show cultural differences. Mayo indicated gaze patterns differ according to the culture of the conversant by analysing gaze behaviours in video recordings of human-human conversations (Mayo and La France, 1978). Elzinga reported that Japanese had "more frequent and shorter lasting other directed gazes" than Australian participants. He also found that English-speaking participants looked at the other person to signal turns, while Japanese did not (Elzinga, 1978). Argyle found that Swedes gaze at their conversation partner more than English (50% vs. 38% of the time) (Argyle and Cook, 1976).

In terms of perception of gaze behaviours, there are studies that indicates cultural preferences of gaze amount that one receives. According to Cook, favourableness of impression would be a linear function of amount of gaze a person receives, and the 50% of gaze amount gave the most favourable impression toward the human gazer in the experiment conducted in UK (Cook and Smith, 1975). Fukayama et al., changed the amount of gaze from a virtual agent

by 25%, 50%, 75% and 100% and compared the agent's impressions by Japanese evaluators. The results showed 50% amount of gaze was perceived as most friendly, followed by 75%, then 25% of gaze amount, but friendly impression plummeted in full gaze (100%) condition (Fukayama et al., 2002).

If there are cultural differences in performing gaze behaviours, there should be cultural differences in perceiving gaze behaviours of other cultures. We aim to investigate cultural differences in gaze behaviours and their perception by developing virtual agents with Japanese gaze behaviours, western gaze behaviours, their hybrid gaze behaviours, and full gaze behaviours, and compare their effects on the impressions of the agents and interactions in this study. This position paper proposes our research agenda, describes the implemented gaze models, and our experimental design.

2 GAZE MODELS

We implemented "western gaze behaviours (WG hereafter)", "Japanese gaze behaviours (JG hereafter)", "hybrid gaze behaviours (HG hereafter)", and "full gaze behaviours (FG hereafter)" to our virtual agent in order to compare the impression of different cultural gaze behaviours.

WG is implemented in accordance with the gaze model proposed by Cassell et al., (1999). Their model shows western (mostly American) people's gaze patterns by analysing video recordings of human dyad conversations. The model shows probability of "looking away" at the beginning (44%) and the end (84%) of an utterance. Fig. 1 shows the state transition diagram of WG at the beginning of an utterance, and Fig. 2 shows the WG at the end of an utterance implemented in our WG model. Our WG also includes a gaze pattern at the end of a question, where the agent "gazes at" the user (human participant). The agent "looks away" from the user for 0.5 second 44% of the time at the beginning of an utterance, then shifts toward a "gaze-at" state. The agent keeps its gaze toward the user during the utterance. The agent "looks away" from the user for 2 seconds 84% of the time at the end of the utterance. The "looks away" timing at the end of the utterance is calculated by estimating the duration of the synthesized speech. The agent keeps its "gaze-at" state toward the user while listening.

JG is implanted in accordance with the gaze model proposed by Ishii et al., (2006; 2008). Their model shows Japanese gaze patterns by analysing video recordings of threesome human conversations.

We implemented JG by modifying their gaze model for dyad conversations by eliminating the state transitions to the third person. Fig.3 shows the state transition diagram of JG. The agent "gazes at" the user at the beginning of an utterance, keeps the gaze for 1.1 to 3.1 seconds, and then shifts its gaze to "vague gaze" (described in section 3) for 3.2-7.9 seconds. The agent shifts its gaze pattern to "gaze-at" state 67% of the time or to "averted gaze" for 2.0 seconds 33% of the time after the "vague gaze". The "gaze aversion" is continued 13% of the time or shifted to" gaze-at" state 87% of the time at the end of the gaze aversion. The agent follows the gaze transitions during its utterance and while it is listening.

HG is implemented by combining JG and WG. As gaze behaviours and patterns are culture dependent (Mayo and La France, 1978), we implemented HG as a culture-independent model, neither western nor Japanese. Fig. 4 shows the state transition diagram at the beginning and during an utterance. The agent follows the transition of WG at the beginning of an utterance, then follows JG during the utterance. The agent follows the state transition diagram of WG while listening.

In addition to WG, JG, and HG, we implemented FG, a full gaze model to the agent as a control gaze condition.
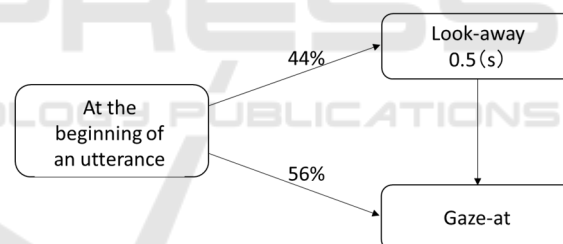

Figure 1: State Transition Diagram of the Western Gaze Behaviour at the Beginning of an Utterance.
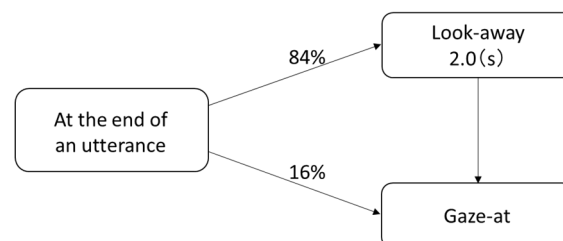

Figure 2: State Transition Diagram of the Western Gaze Behaviour at the End of an Utterance.
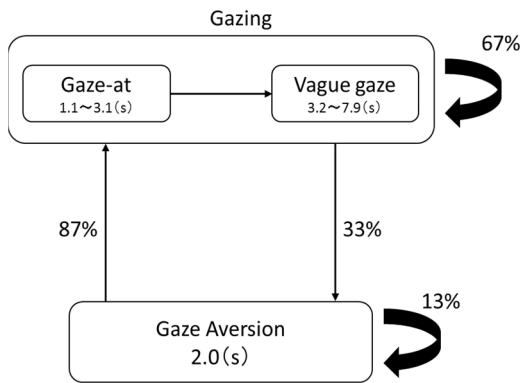
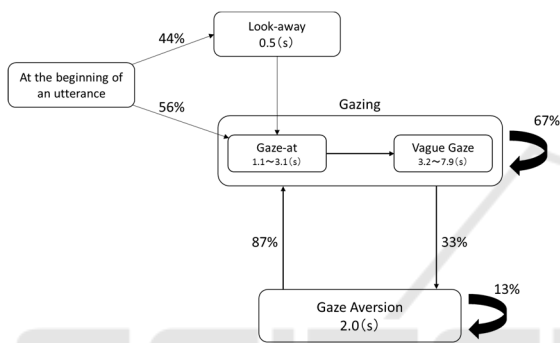Figure 3: State Transition Diagram of the Japanese Gaze Behaviour

Figure 4: State Transition Diagram of the Hybrid Gaze Model

3 VIRTUAL AGENT AND GAZE ANIMATIONS

The agent's appearance and gaze animations are developed by Unity 5.2.1fl (https://unity3d.com/) and Taichi Character Pack asset (https://www.assetstore.unity3d.com/jp/#!/content/1 5667). The agent's voice was synthesized with AITalk (http://www.ai-j.jp/english/). The gaze behaviours implemented to the agent are the four types described in section 2, namely, "gaze-at", "vague gaze", "look-away", and "gaze-aversion".

"Gaze-at" is a state where the agent keeps gazing at a user (shown in Fig. 5 top left and middle). "Vague gaze" is described by (Ishii, 2006, Ishii, 2008) "in order to express less-face-threatening eye-gaze in virtual space avatars", which is implemented as the agent looks at five degrees lower than the user's eye position (shown in Fig. 5 top right and bottom).

"Look-away" is implemented as an animation that the agent discontinues its gaze for 0.5 seconds and looks up, as in Gambi's agent (Gambi, 2015) which is implemented in accordance with the Cassell's western

game model (Cassell et al., 1999). The agent looks up (in "look-away" state for 0.5 second) before an utterance (shown in Fig. 6). "Gaze aversion" is implemented in two directions, to the right and the left, and each aversion lasts for 2 seconds as in (Ishii, 2008). Validation check for each gaze animation were conducted by 8 university students.
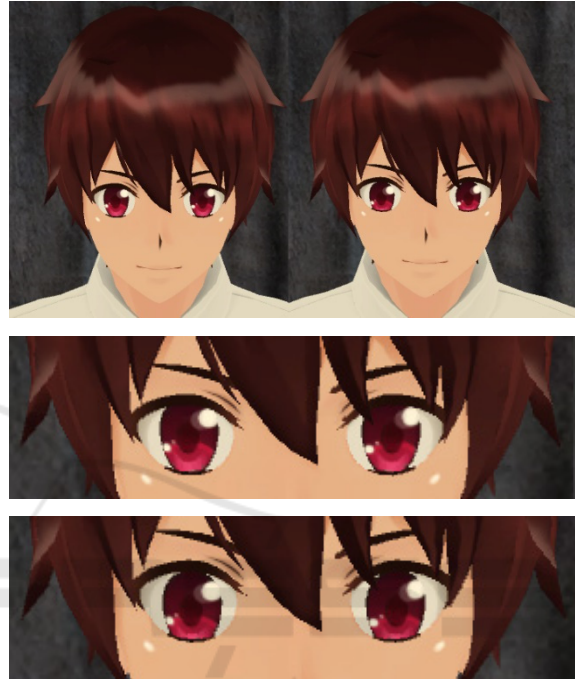

Figure 5: Agent's Gaze-at State (top left and middle) Vague Gaze State: 5 degrees lower than gaze-at (top right and bottom).
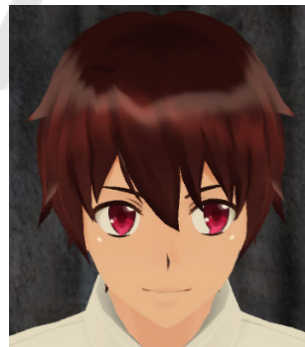

Figure 6: Agent's Look-away state.

Figure 7: Agent's Gaze Aversion States (to left and right, in either direction).

4 PROPOSED EXPERIMENT

4.1 Hypothesis

As described in Section 1, researchers in psychology report cultural differences in gaze behaviours and their perceptions. These facts and findings lead our assumption that gaze patterns perceived as friendly and comfortable are culture-dependent.

Our hypothesis is "If the culture of the gaze model implemented to the agent is the same as the participant's culture, the agent of the same cultural gaze model is perceived as most friendly and comfortable than other agents."

4.2 Experimental Procedure

The experiment is conducted as a Wizard of Oz experiment. Participants are asked to have conversations with a conversational virtual agent in four conversation sessions. The true purpose of the experiment is not explained to the participants during the experiment.

The agent's gaze model and conversational topic are randomly assigned in each of the four conversation session in order to minimize the effect of conversational content. The topics include US Election, Senior Driving, Pokemon GO, and POP icons. The agent brings the above issue and asks opinions to the participants. The agent's reply is controlled by a Wizard. The participants are asked to answer a questionnaire in terms of the agent's likeability, perceived friendliness, comfortableness of the conversation, perceived stress of the conversation, and familiarity of the agent's behaviour after each session.

The experimental conditions are "gaze model (4 models)" and "participant's culture". We will start with Japanese participants in our first experiment, then to participants from western cultures. We will gather participants by asking our research communities in Japan, US and Europe, and ask their nationality, their country of origin, and the country they have stayed longest during the last 5 years to gather their demographics.

5 CONCLUSIONS

We aim to investigate cultural differences in gaze behaviours and their perception, by developing virtual agents with Japanese gaze behaviours, western gaze behaviours, their hybrid gaze behaviours, and full gaze behaviours, and compare their effects on the impressions of the agents and interactions. We hypothesize "If the culture of the gaze model implemented to the agent is the same as the participant's culture, the agent is perceived as most friendly and comfortable than other agents".

We believe the contribution of the research is to make us pay more attention to / to be more aware of cultural differences of gaze behaviours, which we usually control unconsciously. One of the applications of this research outcome would be a cultural training for adequate gaze behaviours of different cultures in order to facilitate mutual understandings and decrease possibility of misunderstandings by misinterpreting other cultures' gaze behaviours.

ACKNOWLEDGEMENTS

REFERENCES

Argyle, M., Cook, M. 1976. Gaze and mutual gaze. Cambridge University Press Cambridge.

Bayliss, A., Paul. M., Cannon, P, and Tipper, S. 2006. Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic bulletin & review* 13(6), pp. l06l-1066.

Bee, N., Pollock, C., André, E., Walker, M. 2010. Bossy or Wimpy: Expressing Social Dominance by Combining Gaze and Linguistic Behaviors. *Procs. of Intelligent Virtual Agents (IVA2010)*, Springer, pp 265-271.

Cafaro, A., Gaito, R., and Vilhjrilmsson, H. 2009. Animating idle gaze in public places. *Procs. of Intelligent Virtual Agents (IVA2009)*, Springer, pp. 250-256.

Cassell, J., Obed E. T., and Prevost, S. 1999. Turn taking versus discourse structure. *Machine conversations*. Springer US, pp.143-153.

Cook, M. and Smith, M.C. 1975. The Role of Gaze in Impression Formation, Br. J. *Clinical Psychology*, Vol.14, pp.19-25.

Elzinga, R.H. 1978. Temporal Organization of Conversation, *Sociolinguistics Newsletter*, 9, 2 (Summer), pp.29-31.

Fukayama, A. et al. 2002. Messages embedded in gaze of interface agents - impression management with agent's gaze. *Proc. of the SIGCHI (CHI '02)*. ACM, New York, pp.41-48.

Gambi, C., Staudte, M., and Torsten, J. 2015. The role of prosody and gaze in turn-end anticipation. *Procs. of the Annual Conference of the Cognitive Science Society.*

Ishii, R, et al. 2006. Avatar's gaze control to facilitate conversational turn-taking in virtual-space multi-user voice chat system. *Proc. of Intelligent Virtual Agents (IVA'06)*, Jonathan Gratch et al. (Eds.). Springer-Verlag, Berlin, pp.458-458.

Ishii, R, et al. 2008. Avatar's Gaze Control to Facilitate Conversation in Virtual-Space Multi-User Voice Chat System, *Journal of Human Interface Society*, 10 (1), pp. 87-94, 2008. (in Japanese).

Koda, T., Ogura, M, and Matsui, Y. 2016. Shyness Level and Sensitivity to Gaze from Agents? Are Shy People Sensitive to Agent's Gaze? *Proc. of Intelligent Virtual Agents (IVA 2016)*, 2016 (in press).

Lee, J., Marsella, S., Traum, D., Gratch, J., and Lance, B. 2007. The rickel gaze model: A window on the mind of a virtual human. *Proc. of Intelligent Virtual Agents (IVA 2007)*, Springer, pp. 296-303.

Mayo, C., and La France, M. 1978. Gaze Direction in Inter-racial Dyadic Communication. *Meeting of Eastan Psychological Association*. in Harper, R.G. et al.

Pelachaud, C., and Bilvi, M. 2003. Modelling gaze behavior for conversational agents. *Proc. of Intelligent Virtual Agents (IVA 2003)*, Springer, pp. 93-100.