In D. Reitter & F. E. Ritter (Eds.), Proceedings of the 14th International Conference on Cognitive Modeling (ICCM 2016). University Park, PA: Penn State.

Eye-tracking Analysis for Product Recommendation Virtual Agent with Markov Chain Model

Tetsuya Matsui (tmatsui@nii.ac.jp)
National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda Tokyo 101-8430, Japan

Seiji Yamada (seiji@nii.ac.jp)
National Institute of Informatics/Sokendai/Tokyo Institute of Technology
2-1-2 Hitotsubashi, Chiyoda, Tokyo 101-8430, Japan

Keywords: eye-tracking, product recommendation virtual agents, Markov chain, Shanon entropy, transition matrix, stationary distribution, fixation duration

Introduction

PRVAs, product recommendation virtual agents, are agents that are designed for virtual clerks in online shopping. Prendinger et al. investigated the effect of virtual clerks by eye tracking analysis (Prendinger, Ma, & Ishizuka, 2007). In their experiment, participants were introduced real estate properties by text, speech, and an animated agent. They showed that the agent's use of deictic gestures had the effect of attracting a participant's gaze. Terada et al. studied what appearance was the most suitable for PRVAs (Terada, Jing, & Yamada, 2015). They showed that one of the most effective appearances were dog, robot, and young woman. In this paper, we investigated the effect of PRVA's emotion transition to user's gaze by eye tracking analysis.

A Markov chain model is widely used for constructing a model of eye tracking transition. Liechty et al. showed local and global covert visual attention by adapting a Bayesian hidden Markov model (Liechty, Pieters, & Wedel, 2003). He et al. suggested investigating hidden user behaviors that occur when a user is using a search site by using a partially observable Markov model with duration (POMD) (He & Wang, 2011). This model is derived from the hidden Markov model (HMM). The difference was that POMD contained a partially observable event. He et al. suggested that only seeing without clicking links was the hidden user behavior.

In this paper, our goal was to improve the PRVA design methodology by analyzing user eye-tracking data. We focused on transition-based analysis. In prior research on human-agent interaction, eye-tracking data were mainly analyzed on the basis of fixation durations. This is the most important method in this paper.

Markov chain

In our research, we used the Markov chain model for analyzing the fixation transitions between areas of interest (AOI sequence). The Markov chain satisfies the following equation, where X_n is a random variable and n means time step (Brooks, Gelman, Jones, & Meng, 2011).

$$\begin{aligned}\mathbf{P}(X_{n+1} = x_{n+1} \mid X_n = x_n, \cdots, X_0 = x_0) \\ = \mathbf{P}(X_{n+1} = x_{n+1} \mid X_n = x_n)\end{aligned} \quad (1)$$

In this research, our goal was to compare the transition entropy and the stationary entropy of the AOI sequence

Experiment

Participants

Fifteen Japanese participants joined in the experiment. There was eight males and seven females, and they were aged between 20 and 39, for an average of 29.3 (SD = 6.9). Due to not getting sufficient gaze data, we omitted the data of one male participant.

Task

The PRVA recommended 10 package tours to Japanese castles. These castles were built in the Japanese Middle Ages, from about the 13th to 16th century. The PRVA made recommendations successively, and the recommendation order was random. For the first half of the recommendations, the PRVA kept a poker face without making any gestures. We defined this agent as the apathy agent. In the latter half, the PRVA smiled and made cute gestures. We defined this agent as the positive agent. This change in facial expressions and gestures expressed the agent's emotion transition, and we aimed for the agent's positive emotion to infect participants.

The PRVAs were executed with MMDAgent[1]. This is a free toolkit for constructing agent systems with speech. It contains the agent character "Mei" and is distributed by the Nagoya Institute of Technology. We also used the text to speech software VOCELOID+ Yuduki Yulari EX2[2] for the agent's voice.

Apparatus

We carried out experiments with Tobii Pro X2-60 and a 30-inch LCD monitor (1920 × 1200 resolution). Eye movements were recorded at a 60-Hz sampling rate. All participants were requested to sit down in a chair at a 60-cm distance from the monitor during the experiment. All stimuli

[1] http://www.mmdagent.jp/

[2] http://www.ah-soft.com/voiceroid/yukari/

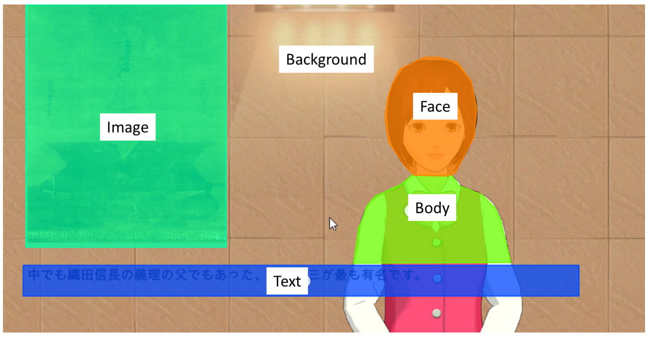

Figure 1: Defined AOIs

were presented on the monitor, and all participants listened to the recommendations with headphones. To construct a transition matrix and stationary distribution, we used the R package "markovchain" (Spedicato, Kang, & Yalamanchi, n.d.).

Analysis method

We defined the AOIs as shown in Figure 1. We divided the presented stimuli area into five areas (*"background," "body," "face," "image,"* and *"text"*). We analyzed based on the fixation order. Fixation order meant the path of a participant's fixations, and we counted the number of transitions of the AOIs that the participants fixed on (including self transitions). We constructed the transition matrix and stationary distribution from this analysis. The minimum fixation duration was 60 ms, and transition advanced one step when fixation occurred.

Results

We constructed the transition matrix and stationary distribution from the fixation order of the first half of the recommendations. We calculated each transition matrix from each recommendation. We got 10 transition matrices from one participant and got 140 transition matrices in total. We calculated the average of all matrix elements. This was for the "transition matrix derived apathy agent".

Also, we calculated the stationary distribution from this matrix. This was for the "stationary distribution derived apathy agent" ($\pi_\mathbf{a} = (0.22, 0.11, 0.05, 0.26, 0.37)$). On the matrix, each coordinate means these AOIs: 1 = "background," 2 = "body," 3 = "face," 4 = "image," and 5 = "text." In the stationary distribution, the same coordinate means the same AOI .

We constructed the transition matrix and stationary distribution from the latter half of the recommendations in the same way. These were for the "transition matrix derived positive agent" and "stationary distribution derived agent" ($\pi_\mathbf{p} = (0.22, 0.16, 0.098, 0.21, 0.31)$). The same coordinate means the same AOI in $\pi_\mathbf{a}$.

Discussion

From $\pi_\mathbf{a}$ and $\pi_\mathbf{p}$, we can find few definite differences. The most different element was p_3 between these two matrices. In $\pi_\mathbf{a}$ this means 0.05, and in $\pi_\mathbf{p}$, this means0.098. This coordinate means the percentage of probability that fixation transitions to "face" when the fixation is on "face" one time-step before. This shows that implementing the positive emotion caused participants' fixations to stay on the agent's face. This phenomenon proves that the participants felt more human-likeness with the agent (Strait, Vujovic, Floerke, Scheutz, & Urry, 2015).

Conclusion

There is demand for PRVAs that have the ability to attract a user's attention to products or to themselves. This can be rephrased as the ability to attract and keep a user's fixation on the images of products or agents. We investigated the effect of implementing a positive emotion in a PRVA by analyzing eye-tracking and aimed to adapt the result to the model of designing PRVAs that attract a user's fixation. From our experiment, a positive emotion attracted participants' gaze to the agent's face. This suggest a methodology of attracting or keeping a user's gaze and buying motivations.

Acknowledgment

This study was partially supported by JSPS KAKENHI " Cognitive Interaction Design " (No.26118005).

References

Brooks, S., Gelman, A., Jones, G., & Meng, X.-L. (2011). *Handbook of Markov Chain Monte Carlo*. CRC press.

He, Y., & Wang, K. (2011). Inferring search behaviors using partially observable Markov model with duration (POMD). In *Proceedings of WSDM'11*. Hong Kong, China.

Liechty, J., Pieters, R., & Wedel, M. (2003). Global and local covert visual attention: Evidence from a bayesian hidden markov model. *Psychometrika*, *68*(4), 519–541.

Prendinger, H., Ma, C., & Ishizuka, M. (2007). Eye movements as indices for the utility of life-like interface agents: A pilot study. *Interacting with Computers*, *19*(2), 281–292.

Spedicato, G. A., Kang, T. S., & Yalamanchi, S. B. (n.d.). The markovchain Package: A Package for Easily Handling Discrete Markov Chains in R.

Strait, M., Vujovic, L., Floerke, V., Scheutz, M., & Urry, H. (2015). Too Much Humanness for Human-Robot Interaction: Exposure to Highly Humanlike Robots Elicits Aversive Responding in Observers. In *Proceedings of CHI '15*. Seoul, Korea.

Terada, K., Jing, L., & Yamada, S. (2015). Effects of Agent Appearance on Customer Buying Motivations on Online Shopping Sites. In *Proceedings of CHI '15*. Seoul, Korea.