2017 26th IEEE International Symposium on
Robot and Human Interactive Communication (RO-MAN)
Lisbon, Portugal, Aug 28 - Sept 1, 2017.

Entropy-based Eye-Tracking Analysis when a User Watches a PRVA's Recommendations

Tetsuya Matsui[1] and Seiji Yamada[1,2,3]

Abstract— **We conducted three experiments to discover the effect of a virtual agent's state transition on a user's eye gaze. Many previous studies showed that an agent's state transition affects a user's state. We focused on two kinds of transitions, the internal state transition and appearance state transition. In this research, we used a product recommendation virtual agent (PRVA) and aimed to discover the effect of its state transitions on users' eye gaze as it made recommendations. We used entropy-based analysis to visualise the deviation of a user's fixations. In experiment 1, the PRVA made recommendations without state transitions. In experiment 2, the amount of the PRVA's knowledge transitioned from low to high during the recommendations. This is an internal state transition. In experiment 3, the PRVA's facial expressions and gestures transitioned from a neutral to positive emotion during the recommendations. This is an appearance state transition. As a result, both the entropy-based analysis and fixation duration based analysis showed significant differences in experiment 3. These results show that an agent's appearance state transitions cause a user's eye gaze to transition.**

I. Introduction

In the field of human-robot interaction (HRI) and human-agent interaction (HAI), the effects of an agent's state transitioning are widely researched [1]. We classified these studies into research on internal state transitions and appearance state transitions.

In the case of internal state transitions, Moon showed that self-disclosure was effective for increasing trustworthiness through electronic commerce [2]. This effect was shown in the case of human-robot communication [3]. Ho and Watson showed that disclosing a virtual agent's autobiographic knowledge increased trust toward agents [4]. These studies show that disclosing personal knowledge causes users' internal states to transition; however, most agents and robots do not have much impersonal general knowledge.

In the case of appearance state transitions, de Melo et al. showed that virtual agents' facial expression transitions caused users to concede [5]. Tsai et al. showed that a virtual agent's positive emotion caused users' emotion to transition to a positive one [6]. In the case of robots, Xu et al. showed that robots can present their mood to users by showing gestures [7].

These studies show ed that an agent's state transition affected a user's emotion, decision-making, and impressions of the agent and robot. In this research, we focused on disclosing impersonal general knowledge and showing smiles and cute gestures related with a positive emotion.

We experimented with a product recommendation virtual agent (PRVA) to evaluate the social practicability of agent state transitions. PRVAs are agents developed to recommend products to users on-line [8]. Some research showed that a PRVA's state impacted the effectiveness of recommendations. Terada et al. showed that the appearance of PRVAs was one factor in buying motivation [9]. Matsui and Yamada showed that PRVAs' internal and appearance state transitions caused users' internal states to transition in terms of emotion, perceived knowledge, and trust in the PRVAs [10]. This previous paper was based on self-reports given by participants. We will show the results of an eye-tracking analysis for the same experiments already reported in this paper [10]. Kamei et al. researched with a robot clerk as a type of PRVA in the real world [11]. They showed that the customers felt more familiarity with the robot clerk when they were invited into the store by it. This showed that the robot clerk can built trust worthy through interactions.

Eye-tracking analysis is widely used in the field of human-agent interaction or human-robot interaction. Strait et al. showed that users tended to gaze at an agent's or robot's eyes when they felt familiarity with the agent or robot [12]. Prendinger et al. evaluated the effect of a PRVA's recommendations by gaze-tracking and showed that the recommendations induced users to gaze at products [13].

We focused on area of interest (AOI) analysis in this paper. An AOI is a divided stimuli area. Ponsoda et al. introduced a transition matrix constructed by transition probability to analyze the direction of a saccade [14]. Goldberg and Kotval discussed constructing a transition matrix from the fixation transition probability between AOIs [15].

The Markov chain model is used in eye-tracking analysisespecially AOI analysisbecause the fixation transition satisfies the Markov property [16]. Liechty et al. suggested a transition model including hidden states, and it is based on the Bayesian hidden Markov model for user-web interactions [17]. He et al. suggested a partially observable Markov model with duration (POMD), a hidden Markov model adding a partially observable state, to analyze a user's eye gaze during web searches [18]. Krejtz et al. analyzed the fixation transition between AOIs when users were looking at pictures [16]. They used entropy derived by the Markov chain model as described in the following section. This research showed the effectiveness of using the Markov model to analyze eye gaze. However, we could not find any research that applied the Markov model to the research of PRVAs.

[1]National Institute of Informatics 2-1-2 Hitotsubashi Tokyo 101-8430, Japan {tmatsui,seiji}@nii.ac.jp
[2]The Graduate University for Advanced Studies (SOKENDAI) 2-1-2 Hitotsubashi, Tokyo 101-8430, Japan
[3]Tokyo Institute of Technology 2-12-1 Ookayama Meguro-ku, Tokyo, Japan

This paper presents our analysis of the users' eye gaze when they watched PRVAs making recommendations. This research has two novel points: a focus on PRVAs' state transitions and the use of entropy-based analysis. We focused on both internal state transitions and appearance state transitions. Internal state transitions were caused by demonstrating to users the PRVA's expertise when making a recommendation. Appearance state transitions were caused by expressing smiles and cute gestures. These displayed a positive emotion. We analyzed the effects of these two transitions on users' eye gaze by using entropy-based analysis.

We focused on the virtual agents in this research; however, the aforementioned problems are the same for human-robot interactions. The state transitions seem to be a very hopeful method for constructing familiarity or trustworthy , and eye gaze is an important factor.

II. Markov chain model and stationary entropy

We used two kinds of value, stationary entropy and total fixation duration, to evaluate the effect of PRVAs on eye-tracking. Stationary entropy is derived by using the Markov chain model [16]. The Markov chain satisfies the following equation, where n and $n+1$ mean that time step X_n is a random variable [19].

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

The AOI sequence satisfying the Markov property was proved by Krejtz et al. [16].

We used the definition of stationary entropy given by Krejtz et al. [16]. Suppose transition matrix $\mathbf{P}$, the stationary distribution π derived from $\mathbf{P}$, and the state space (AOIs) $\varphi = \{1,\cdots,s\}$, where $i \in \varphi$. We can get the entropy of the stationary distribution [20] as follows.

$$H_s = -\sum_{i\in\varphi} \pi_i \log \pi_i \tag{2}$$

If the value of H_s is high, it means that the fixations are uniform between all AOIs. If the value is low, fixations tend to be kept on certain AOIs [16].

In addition to this entropy, we used the stationary distribution itself. The stationary distribution was the vector means probabilities that the gaze converges on each AOI when the user's eye gaze transition diverges to infinity. This value can be calculated from the transition matrix, and it means what AOIs attracted a user's gaze.

We calculated the transition matrix from the probabilities of gaze transitions between each AOI.

We used this method to evaluate a user's eye-tracking when a PRVA was making recommendations. However, we also used the total fixation duration because stationary entropy does not refer to the duration of each or the total duration of fixations. The total fixation duration means the total fixation time for each AOI. Our goals in this research are as follows.

- To discover the effect of PRVA state transitions on a user's eye gaze
- To evaluate the utility of the analysis of stationary entropy

III. Experiment

A. Task

The experiments took 20 to 30 minutes for each participant. The participants were asked to watch movies in which a PRVA recommended a package tour to Japanese castles. The PRVA was executed with MMDAgent[1], and the agent's character was "Mei", a free model of an MMDAgent distributed by the Nagoya Institute of Technology. We used VOCELOID+ Yuzuki Yukari EX[2], which is text-to-speech software, for smooth utterances. Figure 1 shows a part of the experiment in action.

The PRVA recommended ten package tours to ten castles, and these recommendations were indicated by R1R10. All castles were built in the Japanese Middle Ages, and all of them have castle towers. These castles were recommended at random for counterbalancing. We conducted three experiments with this format. In experiment 1, the PRVA made recommendations with a neutral emotion and neutral knowledge during the ten recommendations. In experiment 2, the PRVA made recommendations with neutral knowledge for the first five recommendations and with high knowledge for the latter five recommendations. In experiment 3, the PRVA made recommendations with a neutral emotion in the first five recommendations and with a positive emotion in the latter five recommendations.

In all experiments, the positive emotion was executed by the agent smiling and making cute gestures. High knowledge was executed by demonstrating historical expertise. Figure 2 shows movie snapshots of the agent with a neutral emotion and a positive one. Table 1 shows examples of two patterns (neutral/high knowledge) of speech text on recommending a trip to Odawara Castle.

We divided all stimuli into two AOIs, the agent area and background area. The agent area included the PRVA's face and the upper part of the chest. The lower part of the chest was included in the background area because text was presented there during the recommendations. These areas are shown in Figure 3. In the experiments, the independent variables were the PRVA's knowledge and appearance. The dependent variables were the stationary entropies and total fixation durations.

B. Apparatus

We carried out experiments with a 30-inch LCD monitor (1920 $\times$ 1200 resolution) to present the stimuli, and all participants were requested to sit at a 60-cm distance from the monitor and to listen to the recommendations using

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

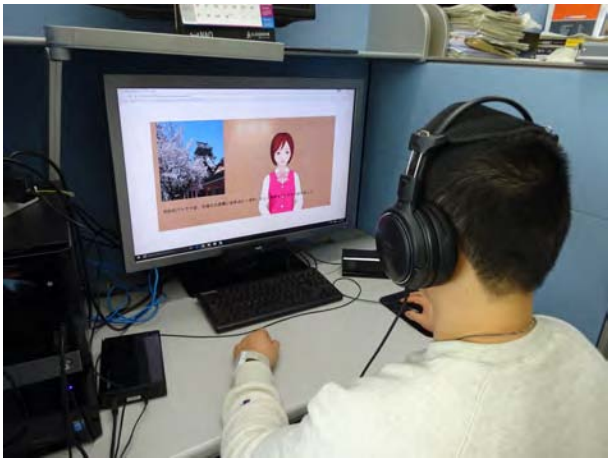

Fig. 1. Snapshot of experiment

Fig. 2. Snapshot of agent with and without smile

headphones. Eye movements were recorded with Tobii Pro X2-60 at a 60-Hz sampling rate. We used the R package "markovchain" [21] to construct a transition matrix to calculate stationary entropies.

IV. Experiment 1: No transition operators

In the first experiment, the PRVA made recommendations with a neutral emotion and neutral knowledge for ten recommendations. We recruited 15 Japanese participants for this experiment. The participants were eight men and seven women, between 21 and 39 years of age, with an average of 29.7 (*SD* = 6.5). They watched a movie series in which a neutral emotion and neutral knowledge were executed for all ten recommendations. However, the data on two of the men were removed because of machine trouble.

Fig. 3. Snapshot of divided AOIs

Neutral	Well, I recommend that you take a trip to Odawara Castle. Odawara Castle stands in Odawara City, Kanagawa Prefecture. We are currently selling a package tour to this castle for only 6440 yen, which includes the admission fee to the castle, the travel expenses, and lunch. Don't miss this opportunity!
High	Well, I recommend that you take a trip to Odawara Castle. Odawara Castle stands in Odawara City, Kanagawa Prefecture. During the Warring State Period, the Hojo family lived in the castle and reigned over the prefecture. When Toyotomi Hideyoshi attacked the castle, the family, and their feudatories wasted too much time discussing the situation and finally had to surrender. Thus, a vain discussion is called "an Odawara discussion". We are currently selling a package tour to this castle for only 6440 yen, which includes the admission fee to the castle. However, the package tour also contains travel expenses and lunch. If you miss this opportunity, you'll surely regret it!

TABLE I
RECOMMENDATION TEXT FOR ODAWARA CASTLE

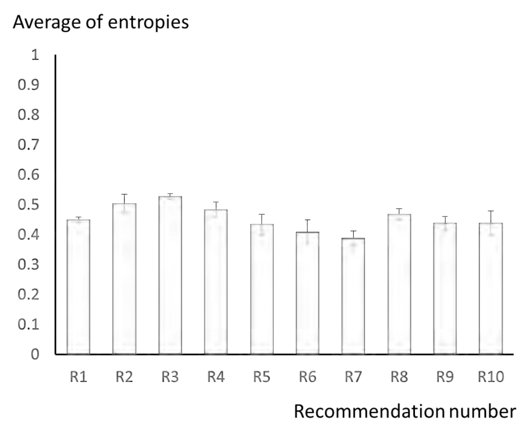

Fig. 4. Average of stationary entropies in experiment 1

Figure 4 shows the average of the stationary entropies for each recommendation, and Figure 6 shows the average of the total fixation durations for each recommendation.

In all figures, the x-axis means the recommendation numbers. In Figure 4, the y-axis means the average of stationary entropies. In Figure 5, the y-axis means the average of probabilities that the gaze converges on an agent area when the gaze diverges to infinity. In Figure 6, the y-axis means the average of the total fixation durations. In all figures, the error bars mean standard deviations. These are the same for the other two experiments.

We conducted a one-way ANOVA on each result and found no significant differences.

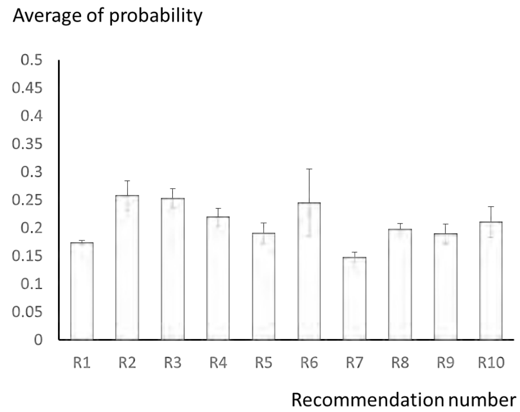

Fig. 5. Average of probabilities that gaze converges on agent area in experiment 1

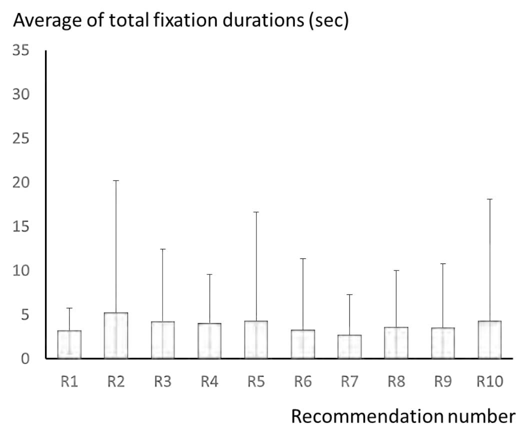

Fig. 6. Average of total fixation duration in experiment 1

V. EXPERIMENT 2: INTERNAL STATE TRANSITION

In the second experiment, the PRVA made recommendations with a positive emotion for all recommendations and also with high knowledge for the latter five recommendations. We recruited 15 Japanese participants for this experiment. The participants were seven men and eight women, between 20 and 39 years of age, with an average of 29.0 (*SD* = 5.7). However, one woman's data were removed because of machine trouble.

Figure 7 shows the average of stationary entropies for each recommendation, Figure 8 shows the average of probabilities that gaze converges on the agent area when the gaze diverges to infinity, and Figure 9 shows the average of the total fixation durations for each recommendation. We conducted a one-way ANOVA on each result and found no significant differences.

VI. EXPERIMENT 3: APPEARANCE STATE TRANSITION

In the third experiment, the PRVA made recommendations with high knowledge for all recommendations and also with a positive emotion for the latter five recommendations. We recruited 15 Japanese participants for this experiment. The participants were eight men and seven women, between 20

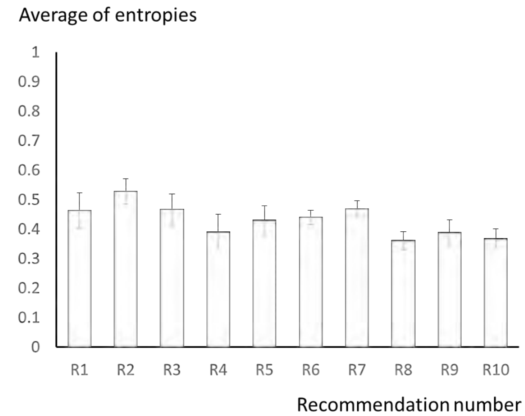

Fig. 7. Average of stationary entropies in experiment 2

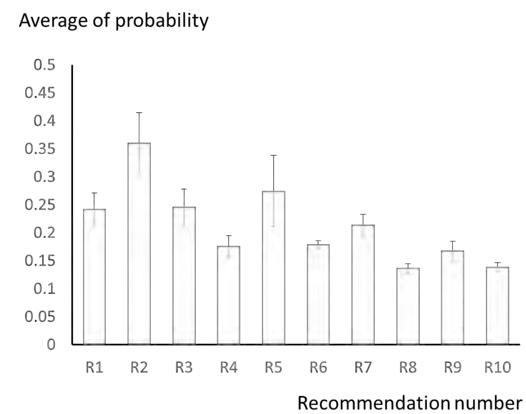

Fig. 8. Average of probabilities that gaze converges on agent area in experiment 2

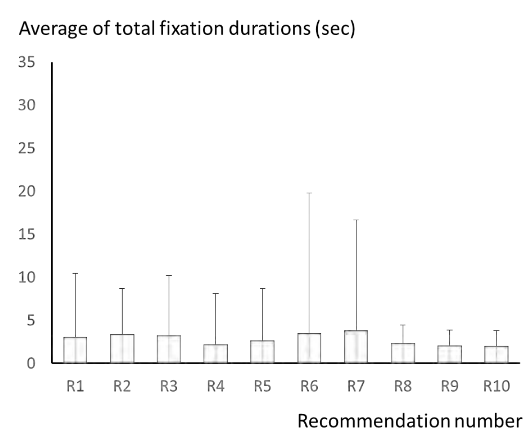

Fig. 9. Average of total fixation duration in experiment 2

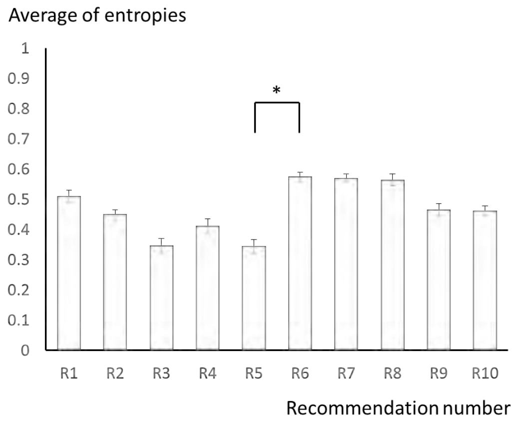

Fig. 10. Average of stationary entropies in experiment 3

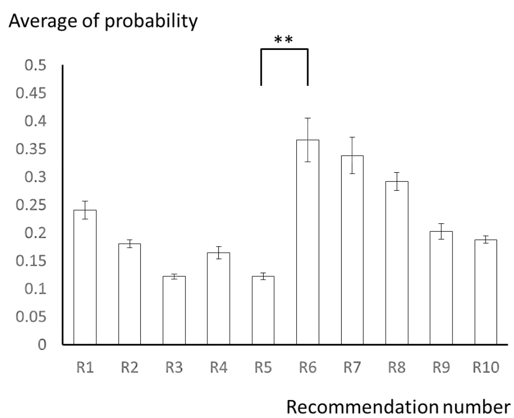

Fig. 11. Average of probabilities that gaze converges on agent area in experiment 3

and 39 years of age, with an average of 29.3 (*SD* = 6.9). However, one man's data were removed because of machine trouble.

Figure 10 shows the average of stationary entropies for each recommendation, Figure 11 shows the average of probabilities that the gaze converges on the agent area when the gaze diverges to infinity, and Figure 12 shows the average of total fixation durations for each recommendation.

We conducted a one-way ANOVA on each result. We observed significant differences in Figure 10 ($p < 0.05$), in Figure 11 ($p < 0.01$), and in Figure 12 ($p < 0.05$). We applied multiple comparisons with Bonferroni correction to reveal whether these significant differences were derived from the agent's state transition or not. The results were that significant differences in Figure 10 ($p < 0.05$), in Figure 11 ($p < 0.01$), and in Figure 12 ($p < 0.05$) were found.

VII. Discussion

In experiment 1, we could not observe significant differences. These results are consistent with our hypothesis because the PRVA's state did not transition in this experiment.

In experiment 2, we could not observe significant differences either. These results show that a PRVA's knowledge

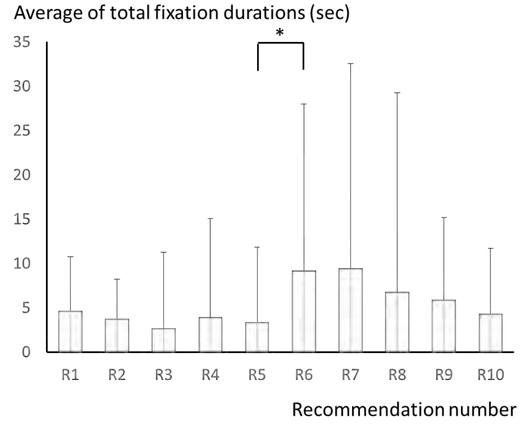

Fig. 12. Average of total fixation duration in experiment 3

state transition cannot cause user eye gaze transitions. The knowledge state seems to be important for users planning to buy products through the PRVA's recommendations. However, this is an internal state that cannot be observed at a glance. In experiment 2, the PRVA's internal state transition did not affect the user eye gaze. In the latter half of this experiment, the agents recommended with longer utterances than those in the first half. This seemed to attract the participants' gaze onto text; however, no significant differences between the first half and the letter half were found. These results show that the participants definitely gaze at text regardless of the text length or the content.

In experiment 3, significant differences emerged in stationary entropies, stationary distribution, and total fixation durations. These results show that the PRVA's emotion state transition caused users' eye gaze transitions. This shows that the PRVA's appearance state transition affects the users' eye gaze.

An increase in stationary entropies between R5 and R6 means the division of the fixations changed to uniform between two AOIs. Considering the difference in the area between the two AOIs, this change seems to show that the duration of fixations on the agent area increased between R5 and R6. The PRVA's smile and cute gestures caused users to fixate on it, and this fixation seems pertinent. The increase in probabilities of gaze converging on an agent area and the total fixation duration seemed to be caused by a similar mechanism.

It may seem to be natural that the appearance transition attracts the users' eye gaze. However, humans generally hardly notice the transition in the visual field [22]. This is called "change blindness". Because of this effect, we concluded that the user's eye gaze transition was not caused by only the agent's superficial change. The user's eye gaze transition seems to be caused by the user perceiving the agent's emotional state transition.

On the basis of these results, we conclude that a PRVA's appearance state transition can cause a user eye gaze transition and that an internal state transition cannot. Also, the stationary entropies and total fixation durations show significant differences between the same recommendations. These results reinforce the idea that stationary entropies are a valid parameter for evaluating changes in the user eye-tracking property.

Also, the transition of stationary entropies happened at the same time as the transition of total fixation durations. These two values are based on different sides of the eye gaze, the division of fixation, and the fixation time. These results support the idea that these two characteristics of eye gaze are correlated. However, we cannot provide a strong reason for using stationary entropy analysis with total fixation duration analysis. Providing it is our future work.

These results were gained from the research on virtual agents. When we apply these results to human-robot interactions, we should consider some limitations. Firstly, we executed emotion state transition with facial expressions. This method may be more difficult for robots than virtual agents. Secondly, the participants watched movies as one part of the experiments. Real store and robot clerk situations need to get the customers into the stores before recommendations can be made. However, in spite of these limitations, the importance of emotion state transition seems to be common in human-robot interactions. Lastly, Tobii T60XL pointed out the systematic delays and drifts [23]. However, we are not sure that this limitation is common with Tobii Pro X2-60, the model that we used in this research.

VIII. Conclusion

In this research, we investigated the effects of a PRVA's state transition on a user's eye gaze. In particular, we focused on the difference between internal state transition and appearance state transition. We used entropy-based analysis to evaluate the characteristics of eye gaze. We also used the total fixation duration. We conducted three experiments: without state transition, with internal state transition, and with appearance state transition. As a result, we observed significant differences in only the appearance state transition for stationary entropies, stationary distribution, and total fixation durations. These transitions occurred when the PRVA's appearance state transition happened. This means the appearance state transition affected the user's gaze, but the internal state transition did not affect it. These results suggest that this method is valid and that an agent's appearance state transition is important for attracting users' attention. This can be used in the human-robot interaction field.

IX. Acknowledgements

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

References

[1] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and autonomous systems*, vol. 42, no. 3, pp. 143–166, 2003.

[2] Y. Moon, "Intimate exchanges: Using computers to elicit self-disclosure from consumers," *Journal of Consumer Research*, vol. 26, no. 4, pp. 323–339, 2000.

[3] J. Mumm and B. Mutlu, "Human-robot proxemics: physical and psychological distancing in human-robot interaction," in *Proceedings of the 6th international conference on Human-robot interaction(HRI'11)*, pp. 331–338, 2011.

[4] W. C. Ho and S. Watson, "Autobiographic knowledge for believable virtual characters," in *Proceedings of the International Workshop on Intelligent Virtual Agents(IVA'06)*, pp. 383–394, 2006.

[5] C. M. de Melo, J. Gratch, and P. J. Carnevale, "Humans versus computers: Impact of emotion expressions on people's decision making," *IEEE Transactions on Affective Computing*, vol. 6, no. 2, pp. 127–136, 2015.

[6] J. Tsai, E. Bowring, S. Marsella, W. Wood, and M. Tambe, "A study of emotional contagion with virtual characters," in *Proceedings of 12th International Conference on Intelligent Virtual Agents(IVA'12)*, pp. 81–88, 2012.

[7] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, "Robot mood is contagious: effects of robot body language in the imitation game," in *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, pp. 973–980, 2014.

[8] L. Qiu and I. Benbasat, "Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems," *Journal of Management Information Systems*, vol. 25, no. 4, pp. 145–182, 2009.

[9] K. Terada, L. Jing, and S. Yamada, "Effects of agent appearance on customer buying motivations on online shopping sites," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 929–934, 2015.

[10] T. Matsui and S. Yamada, "Building trust in prvas by user inner state transition through agent state transition," in *Proceedings of the Fourth International Conference on Human Agent Interaction(HAI'16)*, pp. 111–114, 2016.

[11] K. Kamei, T. Ikeda, M. Shiomi, H. Kidokoro, A. Utsumi, K. Shinozawa, T. Miyashita, and N. Hagita, "Cooperative customer navigation between robots outside and inside a retail shop–an implementation on the ubiquitous market platform," *annals of telecommunications-annales des télécommunications*, vol. 67, no. 7-8, pp. 329–340, 2012.

[12] M. Strait, L. Vujovic, V. Floerke, M. Scheutz, and H. Urry, "Too Much Humanness for Human-Robot Interaction: Exposure to Highly Humanlike Robots Elicits Aversive Responding in Observers," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI'15)*, 2015.

[13] H. Prendinger, C. Ma, and M. Ishizuka, "Eye movements as indices for the utility of life-like interface agents: A pilot study," *Interacting with Computers*, vol. 19, no. 2, pp. 281–292, 2007.

[14] V. Ponsoda, D. Scott, and J. M. Findlay, "A probability vector and transition matrix analysis of eye movements during visual search," *Acta Psychologica*, vol. 88, no. 2, pp. 167–185, 1995.

[15] J. H. Goldberg and X. P. Kotval, "Computer interface evaluation using eye movements: methods and constructs," *International Journal of Industrial Ergonomics*, vol. 24, no. 6, pp. 631–645, 1999.

[16] K. Krejtz, T. Szmidt, A. T. Duchowski, and I. Krejtz, "Entropy-based statistical analysis of eye movement transitions," in *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA'14)*, 2014.

[17] J. Liechty, R. Pieters, and M. Wedel, "Global and local covert visual attention: Evidence from a bayesian hidden markov model," *Psychometrika*, vol. 68, no. 4, pp. 519–541, 2003.

[18] Y. He and K. Wang, "Inferring search behaviors using partially observable Markov model with duration (POMD)," in *Proceedings of the fourth ACM international conference on Web search and data mining (WSDM'14)*, 2011.

[19] S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, *Handbook of Markov Chain Monte Carlo*. CRC press, 2011.

[20] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.

[21] G. Spedicato, "markovchain: an r package to easily handle discrete markov chains," *R package version 0.0*, vol. 8, 2014.

[22] D. J. Simons and D. T. Levin, "Change blindness," *Trends in cognitive sciences*, vol. 1, no. 7, pp. 261–267, 1997.

[23] J. D. Morgante, R. Zolfaghari, and S. P. Johnson, "A critical test of temporal and spatial accuracy of the Tobii T60XL eye tracker," *Infancy*, vol. 17, no. 1, pp. 9–32, 2012.