Cultural Behaviors of Virtual Agents in an Augmented Reality Environment

Mohammad Obaid[1,2], Ionut Damian[1], Felix Kistler[1], Birgit Endrass[1], Johannes Wagner[1], and Elisabeth André[1]

[1] Human Centered Multimedia, University of Augsburg, Augsburg, Germany
{damian,kistler,endrass,wagner,andre}@informatik.uni-augsburg.de
[2] Human Interface Technology Lab. New Zealand, Christchurch, New Zealand
mohammad.obaid@hitlabnz.org

Abstract. This paper presents a pilot evaluation study that investigates the physiological response of users when interacting with virtual agents that resemble cultural behaviors in an Augmented Reality environment. In particular, we analyze users from the Arab and German cultural backgrounds. The initial results of our analysis are promising and show that users tend to have a higher physiological arousal towards virtual agents that do not exhibit behaviors of their cultural background.

1 Introduction

In human-human communications, the majority of our communicative behaviors are demonstrated non-verbally based on the individuals' personality, emotion and cultural background. Researchers have demonstrated that social behaviors in different cultures are addressed differently; therefore, social behavioral differences between communicating individuals can be misunderstood, which might result in a negative perception of each other.

With the increased use of 3D virtual agents, researchers have identified the importance of integrating social behaviors in virtual agents and demonstrated that they have an impact on the user-agents social interaction, e.g. Jan et al. [1]. While, more recently, several researchers measured the influence on the user's perception by the deployment of enculturated virtual characters [2].

To understand human-agent social behaviors, researchers are usually conducting studies in traditional virtual reality (VR) environments displayed on a screen [1]. In this case, the digital medium has an influence on (1) what the user considers a real experience and (2) how the user responds socially when interacting with a virtual agent within the digital medium.

This paper presents the integration of culturally aware virtual agents in the real-space of the user using the Augmented Reality (AR) Technologies; which may have an impact on the users' responses. In recent years, only few studies explore the use of virtual agents in AR environments such as the work in [3,4].

In this paper, we focus on two kinds of signals that play an important role in communicating interest in a social interaction and creating intimacy: interpersonal distance and gaze. Interpersonal distance and gaze are closely coupled. In combination, they convey a certain level of intimacy while people tend to compensate for an increased amount of eye gaze that they consider as socially inappropriate by a higher interpersonal distance. Our paper aims at studying how the response of people is influenced by the cultural background such agents reflect. In particular, we study the physiological arousal response of German and Arab users towards agents reflecting these cultures.

2 Proxemics and Eye Gaze in Virtual Agents

In the literature, a variety of approaches have been presented to simulate proxemic behaviors (the use of space as a form of nonverbal communication) in virtual agents. For example, [5] and [6] presented work based on reinforcement learning and on crowd simulation respectively. Studies have been conducted to investigate the effect of an agent's proxemic behaviors on the users' perception of the agent and their psychological state. Pedica et al. [7] showed that a simulation of human social territoriality, in simulated conversations, contributes to the believability of agent groups. Llobera et al. [8] found that people showed increased physiological arousal the closer they were approached by virtual agents.

Vice versa, a number of researchers investigated how the behavior of a virtual agent influences the proxemics behavior of human users. Bailenson et al. [9] observed that human participants gave more personal space to virtual agents who establish mutual gaze with them.

The simulation of proxemics behaviors has also attracted a significant amount of attention in the robotics community. The work by Mumm and Mutlu [10] is of particular interest to us because it includes a combined analysis of gaze and proxemics behaviors. In a study, they observed that people who disliked a robot compensated for an increase amount of gaze, from the robot, by maintaining a higher physical distance from the robot. While people who liked the robot did not change their proxemics behaviors across various gaze conditions.

In most culture related experiments, researchers extract data only from subjective sources and few attempts were made to study the cultural differences on an objective level, such as the work by Obaid et al. [3] and Llobera et al. [8].

3 Theoretical Background

The integration of cultural factors into the behavior of virtual agents came into focus in recent years. Researchers often investigate the user's perception of a group of virtual agents that show culture-related differences in perception studies. Endrass et al. [2] showed that a culturally prototypical performance of gestures and body postures can enhance the user's perception of an agent conversation in the German and Japaneses cultures. Jan et al. [1] took into account culture-related gaze, proxemics and turn-taking behaviors representing

the Arab and the US American cultures. Their results reveal that participants perceived differences between behaviors that are in line with their own cultural background, and behaviors from different cultural backgrounds.

3.1 Cultural Profiles

In the scope of this paper, culture-related aspects of interpersonal distance and gaze behavior are investigated for the Arab and German cultures. We chose these two cultural groups because clear distinctions can be derived from the research literature, thus promising large differences in participants' perception of nonverbal behavior. A very well known dimensional model of culture is introduced by Hofstede et al. [11], whose theory categorizes national cultures into a six dimensional model.

The Individualism dimension is of special interest for our purpose since it is strongly related to nonverbal behavioral norms. The dimension describes the degree to which individuals are integrated into a group. On the individualist side, ties between individuals are loose, and everybody is expected to take care for him or herself. On the collectivist side, people are integrated into strong, cohesive in-groups. Germany is considered to be an individualistic culture, while the Arab world is considered collectivistic.

Hall introduces a dichotomy [12] that distinguishes so-called *high-* and *low-contact* cultures, which is mainly related to space and appropriate interpersonal distances in social situations. According to Ting-Toomey [13], Germany belongs to the medium-contact group while Arabia belongs to the high-contact group.

3.2 Expectations on Behavioral Differences

The cultural dimension and dichotomy, described in Section 3.1, influence stereotypical behavior for interpersonal distance and eye gaze, which should differ vastly in the Arab and German cultures.

Interpersonal Distance: in [14], Hofstede et al. investigate cultural dimensions in isolation and describe prototypical behavior for the extreme sides. Individualistic cultures, such as Germany, are described to stand free in groups, while collectivistic cultures, such as Arabs tend to be physically close, especially to in-groups. This suggests that interpersonal distance should be higher in German conversations compared to Arab ones. This idea is supported by the categorization into high- and low-contact cultures. In high-contact cultures (such as Arabia) close interactions are common, while in low-contact cultures wider interpersonal space is appropriate.

Eye Contact: the scores on Hofstede's individualism dimension give some interesting insights. In [14], members of individualistic cultures are described as making eye contact freely which suggests that this should hold true for the German culture. Taking into account Hall's categorization into high- and low-contact cultures, members of Arab countries (low-contact) use direct facing and frequent direct eye contact. Since Germany is located on a medium-low contact level, direct facing should occur less frequently.

Table 1. Culture-related behaviors

Aspect	Arab	Neutral	German
distance	low	medium	high
eye contact	high	medium	low

Table 2. Experimental Conditions

Condition	distance	eye contact
0	Neutral	Neutral
1	Arab	German
2	German	Arab
3	Arab	Arab
4	German	German

4 Experimental Evaluation

To evaluate the impact of culture-specific behavior on human users, we conducted a pilot study with Egyptian and German participants. Due to the similarity principle, the social attraction perceived by individuals is increased when interacting with someone who is perceived as being similar to oneself [15]. Therefore, participants should prefer agent behavior that was designed to resemble their own cultural background. Vice versa, culturally different behavior might lead to tension which should be measurable in physiological arousal. We therefore derive our ***Hypothesis:*** *AR agents that have a different culture behavior from the human user elicit a higher level of physiological arousal.*

4.1 Design

Agents: we use two virtual agents in our setup based on the AAA engine and functionalities [16], which includes the control of non-verbal behaviors conveyed by agents. The appearance of the agents are consistent and are chosen to be of dark-hair, no-culture specific clothing, and of mixed gender. Moreover, to allow for a focus on the nonverbal behaviors, the agents use Gibberish during dialog, a fantasy language without any specific meaning of the words but with the same statistical distribution of syllables as the words from the English language.

Conditions: The experimental conditions are designed based on the cultural profiles and expectations (Section 3). Therefore, we defined three prototypical performances (Arab, Neutral and German), where the Neutral behavior serves as the average of the behaviors for the German and Arab cultures. Neutral was used during a non-evaluated user familiarization phase (Tables 1 and 2). Since our work focuses on analyzing the impact of differences between the cultural groups and not measuring the exact effect of the conditions, therefore, the exact definition of the values for the behavior aspects was not a priority and are defined relative to each other (e.g. high gaze is higher than low gaze).

Scenario: Within each condition, the user enters the AR environment with two agents engaged in a conversation. Starting from a position of 2.8 meters away from the agents, the user is instructed to approach the agents to a comfortable distance to join their conversation as a listener (Fig. 1). As the user approaches, the agents adjust their conversational formation with an interpersonal distance based on the condition. In addition, the agents' amount of eye contact with the

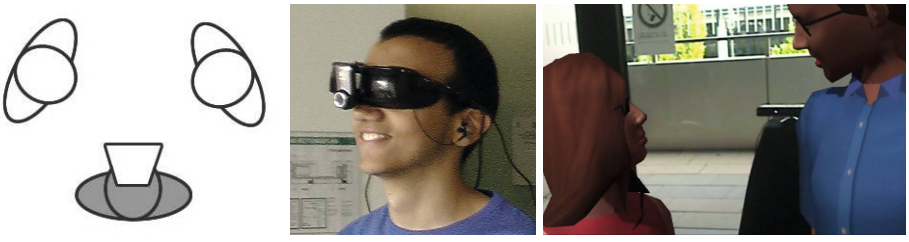

Fig. 1. Schematic setup of the experiment showing the position of the user and virtual agents (left), participant during interaction (middle), and an example of the user's view showing the virtual agents in the AR environment (right)

interlocutors is adapted to the condition. Other nonverbal behaviors, such as gestures and body postures, are left constant during all conditions.

Experimental Setup: The experiment is arranged in a room with five-meter width and six meter depth. The setup contains two main pieces of hardware equipment for tracking and AR display: (1) A Microsoft Kinect[1] is put on a pedestal of one meter height and placed at a centered position against the wall. We use the OpenNI framework and NITE middleware[2] to track the users in front of the Kinect and to get the user's position in relation to the sensor. (2) A Vuzix[3] Head Mounted Display (HMD) that includes earphones and a head tracker (measuring the head orientation via an accelerometer and compass). Using the tracking technologies, we incorporated the virtual agents into an AR environment. We apply the user's position (tracked by the Kinect) and head orientation (from the HMD's tracker) to the virtual camera used to compute the 3D transformations for rendering the virtual agents. To simulate the user's real world view, a webcam is attached to the HMD. The video image is captured, augmented with the virtual agents, and then presented back to the user's eyes. In addition, for accuracy, the head tracker is calibrated for each user, before the study, so that head orientation and virtual camera orientation are in line.

Throughout the study, we use the *NeXuS*[4] and the SSI framework [17] to measure and record four biosignals (electrocardiography, respiration, skin conductivity and blood-volume-pulse). Using these signals, we were then able to compute the user's level of physiological arousal during the study.

Procedure: The experiment begins by giving a description of the study and the equipment to the participant. After fitting the equipment to their body, they are asked to stand at the initial position in the room. Thereafter, the participant is acquainted with the environment and the gibberish dialog by allowing them to experience condition 0 of Table 2. Only then, the participant will randomly conduct each of the four different conditions, described in Table 2.

[1] http://www.xbox.com/kinect
[2] http://www.openni.org
[3] http://www.vuzix.com/consumer/products_wrap_1200vr.html
[4] http://www.humankarigar.com/wireless_nexus10.htm

5 Results and Discussion

For our evaluation study, we recruited 20 participants: 10 from Egypt (5 females, 5 males) and 10 from Germany (3 females, 7 males), while all of them had little or no experience with augmented reality environments.

The analysis of the physiological recordings yielded some interesting tendencies. From the ECG signals, we extracted the heart rate from the interval between successive R waves. Since the physical load was steady across the conditions an increase of the heart rate can be taken as a sign for an increased stress level of the participants. To observe differences between the four conditions (Table 2), we calculated for each participant, the mean heart rate over the period of each condition. In order to reduce the effect of individual differences we also subtracted the mean heart rate extracted from the whole session. Then we compared the average of the normalized values between Egyptian and German participants. For the conditions, where gaze and distance are adjusted to the same country (3+4), the values are within a range from $-1.1hz$ to $-0.6hz$, which is below the average heart rate. This indicates that participants generally were in a relaxed mood even when they were facing the conditions of the other culture, as long as eye gaze and distance are of the same culture. In case of the mixed conditions (1+2), for both groups, the normalized heart rate increased above zero when the eye gaze was adjusted to the other culture ($0.2hz$ for German and $0.4hz$ for Egyptians), but not in the case where the distance was adjusted to the other culture ($-1.1hz$ for German and $-1.4hz$ for Egyptians). This suggests that the eye gaze aspect has a higher impact on the level of physiological arousal than distance. However, we predict that a clearer picture on the statistical analysis of our results will reform when analyzing the results of additional participants. Therefore, we neither deny nor approve our hypothesis at this stage.

6 Conclusion and Future Work

In this paper we presented a pilot evaluation study to investigate the physiological arousal of users towards augmented reality virtual agents that resembles cultural behaviors. We demonstrate this by integrating two social signals (interpersonal distance and eye-gaze) for two different cultures (Arab and German) into virtual characters. The study is still in running process, however, we reported the analysis of the initial results from 20 participants (10 Arab and 10 German). The initial results, from the physiological responses of users, revealed interesting tendencies (such as a higher heart-rate average when interacting with agents of culturally inconsistent non-verbal behavior). The participants also perceived the conditions with a culturally appropriate eye gaze as more relaxing even if they are experiencing non-appropriate distance activity. Future work are directed towards exploring social attraction towards agents simulating cultural behaviors.

Acknowledgments. This work was partially funded by the European Commission within the 7th Framework Program under grant agreement eCute.

References

1. Jan, D., Herrera, D., Martinovski, B., Novick, D., Traum, D.: A Computational Model of Culture-Specific Conversational Behavior. In: Pelachaud, C., Martin, J.-C., André, E., Chollet, G., Karpouzis, K., Pelé, D. (eds.) IVA 2007. LNCS (LNAI), vol. 4722, pp. 45–56. Springer, Heidelberg (2007)
2. Endrass, B., André, E., Rehm, M., Lipi, A.A., Nakano, Y.: Culture-related differences in aspects of behavior for virtual characters across germany and japan. In: Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems, AAMAS, Richland, SC, vol. 2, pp. 441–448 (2011)
3. Obaid, M., Niewiadomski, R., Pelachaud, C.: Perception of Spatial Relations and of Coexistence with Virtual Agents. In: Vilhjálmsson, H.H., Kopp, S., Marsella, S., Thórisson, K.R. (eds.) IVA 2011. LNCS, vol. 6895, pp. 363–369. Springer, Heidelberg (2011)
4. Magnenat-Thalmann, N., Papagiannakis, G., Chaudhuri, P.: Applications of Interactive Virtual Humans in Mobile Augmented Reality. In: Furht, B. (ed.) Encyclopaedia of Multimedia, 2nd edn. Springer (2008)
5. Kastanis, I., Slater, M.: Reinforcement learning utilizes proxemics: An avatar learns to manipulate the position of people in immersive virtual reality. ACM Trans. Appl. Percept. 9(1), 3:1–3:15 (2012)
6. Jan, D., Traum, D.: Dynamic movement and positioning of embodied agents in multiparty conversations. In: Proc. of the Workshop on Embodied Language Processing, EmbodiedNLP 2007, Stroudsburg, PA, USA, pp. 59–66. Association for Computational Linguistics (2007)
7. Pedica, C., Högni Vilhjálmsson, H., Lárusdóttir, M.: Avatars in conversation: The importance of simulating territorial behavior. In: Safonova, A. (ed.) IVA 2010. LNCS, vol. 6356, pp. 336–342. Springer, Heidelberg (2010)
8. Llobera, J., Spanlang, B., Ruffini, G., Slater, M.: Proxemics with multiple dynamic characters in an immersive virtual environment. ACM Trans. Appl. Percept. 8(1), 3:1–3:12 (2010)
9. Bailenson, J.N., Blascovich, J., Beall, A.C., Loomis, J.M.: Interpersonal distance in immersive virtual environments. Personality and Social Psychology Bulletin 29(7), 819–833 (2003)
10. Mumm, J., Mutlu, B.: Human-robot proxemics: physical and psychological distancing in human-robot interaction. In: Proc. of the 6th Int. Conf. on Human-robot Interaction, HRI 2011, pp. 331–338. ACM, New York (2011)
11. Hofstede, G., Hofstede, G.J., Minkov, M.: Cultures and Organisations. Software of the Mind. Intercultural Cooperation and its Importance for Survival. McGraw Hill (2010)
12. Hall, E.T.: The Hidden Dimension. Doubleday (1966)
13. Ting-Toomey, S.: Communicating across cultures. The Guilford Press, New York (1999)
14. Hofstede, G.J., Pedersen, P.B., Hofstede, G.: Exploring Culture - Exercises, Stories and Synthetic Cultures. Intercultural Press, Yarmouth (2002)
15. Byrne, D.: The attraction paradigm. Academic Press, New York (1971)
16. Damian, I., Endrass, B., Huber, P., Bee, N., André, E.: Individualized Agent Interactions. In: Allbeck, J.M., Faloutsos, P. (eds.) MIG 2011. LNCS, vol. 7060, pp. 15–26. Springer, Heidelberg (2011)
17. Wagner, J., Lingenfelser, F., André, E.: The social signal interpretation framework (SSI) for real time signal processing and recognition. In: Proc. of Interspeech 2011, Florence, Italy, pp. 3245–3248 (2011)