Human Interface of Robots or Agents via Facial and Word Expression

Kaoru Sumi[1]

[1]Future University Hakodate, Hakodate, Hokkaido 041-8655, Japan
kaoru.sumi@acm.org

Abstract: To design an intelligent interactive system, it is necessary to consider how humans feel about the system and establish a good relationship with them. In human robot interaction or human agent interaction, to establish a fifty-fifty relationship between a technical artifact (such as a robot or an agent system) and a human, the power of conviction or influence of the artifact over the human is very important. To develop an intelligent system using a robot or an agent such as a system that proactively interacts with a user and even changes the user's intention according to the user's circumstances, our project investigated reactions with the user under several situations, considering human robot interaction and human agent interaction using facial and word expressions. Accordingly, we established some rules for making the agent's reaction favorable to the user on the basis of facial expressions and words, and gained some insights into the differences between human robot interaction and human agent interaction. In this talk, I introduce the possibility of human persuasion by a robot or an agent using facial expressions and emotion words, based on the experimental results.

Keywords: Facial expression, word expression, persuasive technology

1 INTRODUCTION

Nowadays, there are many systems that use virtual agents to mediate between a user and the system. Character agents make the user feel the presence of an assistant to access the system and also give him or her feeling of affinity with the system. According to Media Equation [1], people treat computers, television, and new media as real people and places, thereby making the users uncomfortable if an agent behaves in a disagreeable manner. In the field of persuasive technology research [2] it is said that if a user recognizes the presence of something in a computer, he or she will respond to it according to the normal social rules. However, there are still many things that we do not know about how an agent's response affects a user during their interaction.

In the development of intelligent systems, it is important to consider how best a feeling of affinity with the system and show the presence of the system that has human-like intelligent functions such as recommendation or persuasion. Therefore, evaluating the interpersonal impressions conveyed by agents is very important.

Our research group performed an experiment to evaluate how the facial expressions of an agent and the words used by the agent affected users during agent-user interaction [3]. In this paper, I introduce our evaluation of the user's impression of agents in emotion-arousing scenarios set up to see how users react to various patterns of agent reactions. In particular, after setting up situations that evoke feelings of "joy", "anger", "sadness", "disgust", "fright", and "surprise" and matching up the agent's reaction with the combination of facial and verbal expressions, we asked users about their impressions of the agent.

There have been some studies using character agents as the interface of a system. However, it is not known how and there has been no evaluation of how the agents affect the users. Interactive patterns of the combination of the agents' facial expressions and words have not yet been evaluated.

The situation in which a message and its contradiction are both presented simultaneously is called a double-bind situation [4]. There has been research on information processing by the memory and the feelings perceived by humans in double bind situation. However, it is not clear what impression the speaker conveying the messages of double-bind situation gives to the receiver.

2 EXPERIMENT ON IMPRESSIONS OF REPLIES FROM THE AGENT

We chose six kinds of feelings. From the total of 216 combinations, covering multiple feelings that the user felt (6 patterns) and the facial expressions for the agent's interaction with the user (6 patterns) and word expressions used by the agent (6 patterns), we selected 96 patterns in this experiment. These covered 16 patterns in each feeling: empathetic words and consistent facial expressions, nonempathetic words and consistent facial expressions, word consistent and facial inconsistent, and word inconsistent and facial consistent. This is because the conditions of nonempathetic and both inconsistent word and facial expressions are nonsensical in normal communication. This is the condition where the word and

facial expressions are inconsistent, which is the condition for double bind communication, but it can be considered as either word or facial being empathetic to the user. The case of nonempathetic condition and inconsistent word and facial expressions can be considered as pathological. A total of 1236 people, 568 male and 668 female (AV. age 38.0, SD age 11.5), were assigned 96 contents. More than ten users were assigned to each content.

➢ Experimental Materials

"Joy", "anger", "sadness", "disgust", "fright", and "surprise" scenarios were selected as scenarios with a high concordance rate in a preliminary experiment as emotion-arousing scenarios for each emotion. These scenarios were described by a male reader reading in a neutral manner.

A female character agent was used to react to the user's emotion. Faces representing "joy", "anger", "sadness", "disgust", "fright", and "surprise" were selected as faces with a high concordance rate in a preliminary experiment as emotional faces.

The agent dialog was read by a female reader with emotions conveyed. At first, as empathetic dialogue, "I think so, too" or "I don't think so" as nonempathetic dialog were spoken. Then, emotionally, the dialogue of "that's nice", "that's aggravating", "that's sad", "that's disgusting", "that's scary", and "that's a big surprise" were spoken.

In this experiment, we use nine factors: three factors for interpersonal impression evaluation [5] were "affable-inaffable","serious-unserious","conversable-inconversable" and original six factors.

" reliable-unreliable", "gentle-bitter", "egotistic-humble", "empathetic-unempathetic", "authoritative-unauthoritative", and "offensive-inoffensive".

We prepared 96 contents to cover the combination of emotions that a user feels, the facial expressions of the agent, and the word expressions used by the agent. These contents were developed using the Bot3D Engine [6], which displays an agent on web pages. The Bot3D Engine is an embedded engine for developing software using the Web3D plug-in and ActiveX component. Users can use the 3D agent program only to access web pages that have the program embedded.

➢ Procedure

The examination was conducted in the form of a questionnaire on the Web. The content was displayed on user's own PC monitor after the user accessed the target URL.

The users were asked about their sex, age, marital status, occupation, intended purpose of using the PC, and for how long they had been using a PC.

Next, the following teaching sentences were presented.

"This examination aims to discover what emotions people feel in various cases. There is no correct answer, so please say exactly what you think and feel. This examination is not a test of your personal abilities. The answers will be analyzed statistically and private information will not be released. First, please consider the given scenario and then select from the alternatives the emotions that you feel. Next, an animated character will respond to your selected answer. Please answer the question by giving your impression of the character. Your answer should relate only to this scenario. Please do not include feeling from previous scenarios, but think scenario by scenario."

Each user was presented with one of 96 contents. One of the emotion-arousing scenarios was read by a male voice. It then asked: "What kind of emotion do you feel?" and prompted the user to select from the alternatives "joy", "anger", "sadness", "disgust", "fright", and "surprise". On the other hand, the female agent on the screen responded with facial and word expressions.

Users were asked: "How do you feel about this person? Please answer using the degrees listed in the questionnaire." Five conditional moods in nine answers, "conversable-inconversable", "reliable-unreliable", "gentle-bitter", "egotistic-humble", "empathetic-unempathetic", "authoritative-unauthoritative", "offensive-inoffensive", "serious-unserious", "affable-inaffable" were given and the user selected a suitable answer. The order of the terms was kept constant throughout the questions.

3 Result and Discussion

We conducted the experiment on how the user felt about the agent's reaction by setting up an emotion-arousing scenario for the user. Four factors were extracted by using factor analysis. Eight clusters were indicated by cluster analysis by using four factors as the dependent variable. Favorable, intermediate, and unfavorable impressions fell in category of higher-level clusters among these eight clusters. Therefore, we focused on the relationship between agent reaction and the higher clusters.

The synchronization of the agent's words with the user's emotion has a major impact on the impression of the agent as perceived by the user. However, the synchronization of facial expressions of the agent with the user's emotion does not have a major impact on the creation of an impression.

First, we predicted that words and facial expressions reflected on the emotions aroused by the scenario would lead to the most favorable impression, so we set these data as the control group. In fact, there were more favorable impressions than those obtained for the control group. For example, the words and facial expressions were "joy" when the user's emotion was "joy" for the control group. It is very interesting that when the user's emotion was "joy", the agent's words for "joy" with facial expressions of "surprise", "sadness", or "fright" were most favorable. On the other hand, when the user's emotion was "fright", the agent's words for "fright" with facial expressions of "disgust" or "sadness" were the most favorable.

These facial expressions were recognized as the emotion conveyed by the words and were more empathetic and somewhat meaningful emotions. For example, when the user's emotion was "joy", the agent's words of "joy" with facial expressions of "surprise" or "fright" might have been recognized as the agent being exaggeratedly surprised at the "joy" scenario. When the user's emotion was "joy", the agent's words of "joy" with facial expressions of "sadness" might have been recognized as the agent being highly pleased from the heart at the "joy" scenario. When the user's emotion was "fright", the agent's words of "fright" with facial expressions of "sadness" might have been recognized as the agent grieving deeply at the user's "fright" scenario. When the user's emotion was "fright", the agent's words of "fright" with facial expressions of "disgust" might have been recognized as the agent feeling deep hate at the user's "fright" scenario.

Through these observations, we concluded that there is a rule for facial expressions: in a certain scenario, synchronizing foreseen emotion of the user caused by the situation will make a favorable impression. For example, when the user has the emotion of "joy", he/she wants someone to be surprised or highly pleased. Then, showing surprised or highly pleased face expression make the user feels favorable impression. When the user has the emotion of "fright", he/she wants someone to grieve deeply or disgust. Then, showing grieved or disgust face expression make the user feels favorable impression. Users want the agent to ooze synchronized their foreseen emotion by hearing the news instead of simply showing synchronized reaction according to emotion at present time.

The ability to do this is known as the emotional intelligence quotient (EQ) [7], which is a measure of the ability to understand the feelings of the partner and maintain human relations well.

This facial expression rule is a kind of EQ rules as we can often see service-minded persons show their sympathy with very sad face when they on hearing bad news. In this case, the emotional situation was "disgust", their word is "disgust" with facial expressions of "sad". These patterns are consistent with the cluster of favorable impressions in the result of the experiment. These persons favorably impress, as we often see them employed as salesmen having some technical know-how in order to make themselves look good. We often see this type of person in our country and they are accepted as favorable. However, this facial expression might be considered as a specific feature of Japanese culture. It needs more examination, taking into account diverse nationalities.

This facial expression rule is a technique of foreseeing the other's emotion and the agent can behave proactively by reading the other's feelings.

5 CONSLUSION

As a purpose of developing intelligent system using virtual agent which interacts with a user proactively according to the user's circumstances, we evaluated the user's impression of agents by setting up an emotion-arousing scenario and observed how the users reacted to various patterns of agent reactions. The results of the experiment reveal the rule for creating an agent which reacts proactively using facial expressions.

For the next step of this research, our research group have evaluated human robot interaction using facial expression.

REFERENCES

[1] Reeves, Byron and Clifford Nass The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places, Cambridge University Press, 1996.

[2] B.J.Fogg: Persuasive Technology –Using Computers to Change What We Think and Do-, Elsevier, 2003.

[3] Kaoru Sumi and Mizue Nagata: Evaluating a Virtual Agent as Persuasive Technology, Psychology of Persuasion, Janos Csapó and Andor Magyar eds., Nova Science Publishers (2010).

[4] Bateson, G., Jackson, D. D., Jay Haley and Weakland, J., "Toward a Theory of Schizophrenia", Behavioral Science, vol.1, 1956, 251-264, 1956

[5] Ikuo Daibo: Sansya kan communication ni okeru taijininsyo to genngo katudousei, jikkenn shinri gaku kenkyu,18, 21-34, 1978 (in Japanese)

[6] Bot3D: http://www.atom.co.jp/bot/

[7] Daniel Goleman: Emotional Intelligence, Bantam Dell Pub Group, 1995.