Studies on Skill Level and Dialogue Satisfaction for Achieving Mind-to-Mind Communications Technology

Ryohei Saijo, Yoko Tokunaga, Daichi Yamaguchi, Lidwina Andarini, Shohei Matsuo, Iwaki Toshima, Takao Kurahashi, and Shiro Ozawa

Abstract

We have undertaken the development of mind-to-mind communications technology to enable a uniform means of communication that can be mutually understood by all humans. As the first step in this effort, we focused on (1) how to convey information tailored to the other person by studying technologies for presenting information tailored to work skills and (2) how to evaluate communication by evaluating and estimating satisfaction on the basis of positivity and impact of a dialogue participant. As the next step, we will refine the technologies presented in this article to improve understanding of utterance intention through expressions that match one another's sensitivities.

Keywords: mind-to-mind communications, skill level, satisfaction level

1. Activities toward mind-to-mind communications

In human communication, information uttered by a person will generally not be conveyed to the other person with 100% accuracy. It is not unusual for some misunderstanding (miscommunication) to occur, and real dialogue may never become established (discommunication). As a means of communication to avoid these problems, we are developing mind-to-mind communications technology to help convey the feelings (sensitivities) that one wants to communicate to another person. By conveying sensitivities, mind-to-mind communications technology conveys what a person wants to convey accurately, which makes it possible to find a new solution or reach a better agreement while maximizing communication results and the satisfaction level of the parties concerned. This requires that sensitivity be considered from two viewpoints, i.e., sensitivity of the sender and sensitivity of the receiver. In other words,

the task is to determine how the sensitivity that a person wants to convey will be expressed depending on the sensitivity of the sender and predict how that will be interpreted depending on the sensitivity of the receiver. Studies are needed to determine what is necessary to satisfy the parties concerned in actual communication. The reason for this is as follows. In a situation in which communication aims to solve a problem or build a consensus, it is not enough to simply reach a better consensus from on objective point of view—it is also essential that the parties concerned be satisfied and reach an agreement in a subjective manner (on the basis of the sensitivities of each party).

We present the following two example studies related to these two problems:

(1) How can the sensitivity that a person wants to convey be conveyed in a manner tailored to the other person? We investigated an interface that can determine a person's level of skill from biological signals and switch the

information to be presented accordingly.

(2) What effect does the attitude of the dialogue participants have on the degree of satisfaction with respect to the results of that dialogue? We examined the results of a basic experiment on the degree to which positive dialogue participation can contribute to satisfaction.

We hope that these studies will serve to stimulate discussion on mind-to-mind communications technology.

2. Interface technology for presenting information tailored to work skills

Given a scenario of collaborative work in which a variety of people come together to accomplish the same task while engaging in communication, it is extremely important that information be conveyed in a manner tailored to the other person. We have been aiming for some time to make collaborative work even smoother by achieving information communication tailored to the other person through the use of human digital twins. In terms of technologies and services in support of collaborative work, there have been many proposals on how best to communicate work instructions or supplementary information from one user to another. Most of these technologies and services provide such support in a uniform manner to each user (for example, displaying helpful marks at specific locations within an online collaborative work space).

However, the support needed during work differs according to the user's skill level, and the cognitive load may also differ in accordance with the information being given (or not being given). For example, we can consider a game of pair *shogi* (Japanese chess played in pairs) in which one pair exchanges opinions on their next move. For expert and intermediate players, a simple information display, such as displaying marks at important locations on the *shogi* board, is considered sufficient. However, for beginners, there may be a need for a more detailed explanation such as showing how pieces are moved using a drawing or illustration and explaining the merits of each move. When targeting users of various levels, there is a need for technology that would enable a system to understand those differences and provide appropriate support tailored to each type of user.

This technology would estimate a user's skill level from biometric information obtained from the user while working and from information related to past work experience and reproduce those aspects in a human digital twin corresponding to that user. That human digital twin could then be used to infer the information needed by the user and switch the information display accordingly without obstructing work (**Fig. 1**). This technology will enable collaborative work support promoting smooth communication even among users having different sensitivities, such as intuition and ability of understanding, that affect work skills.

Referring to a previous study [1] stating that eye movement (gaze) during work differs according to skill level with respect to that work, we have undertaken the use of sensing data centered around eye tracking*. Our aim is to conduct machine learning that combines the amount of eye movement, movement patterns, user work experience, difficulty of work, etc. to estimate the user's skill level. We are pursuing the design of an information-presentation method tailored to the user through wide-ranging studies on information displays including the modal to be used (graphics, voice, etc.), granularity of information, and placement of information in space.

As an example of an activity related to skill level, we introduced interface technology for estimating skill level using gaze data and for displaying information on the basis of that estimation. Going forward, we will deepen our studies on information-presentation methods that incorporate user-interface/user-experience knowledge [2] accumulated at NTT Digital Twin Computing Research Center while expanding our work toward the development of a more human-centric system.

3. Technology for estimating satisfaction level based on positivity and impact of a dialogue participant

When building a consensus through dialogue, whether the dialogue that took place was satisfactory or the results achieved are acceptable may differ from one participant to another. For example, there are people who place importance on having a positive attitude as in "I was able to talk a lot and enjoyed the discussion" and who evaluate their level of satisfaction with the dialogue on the basis of the number of times they got to speak. However, there are also people who place importance on the degree to which they were able to contribute as in "My opinion helped to build a consensus" and who evaluate their level of

---

* Eye tracking: A biometric technique that analyzes the motion of a person's eye to clarify visual attention and other characteristics.
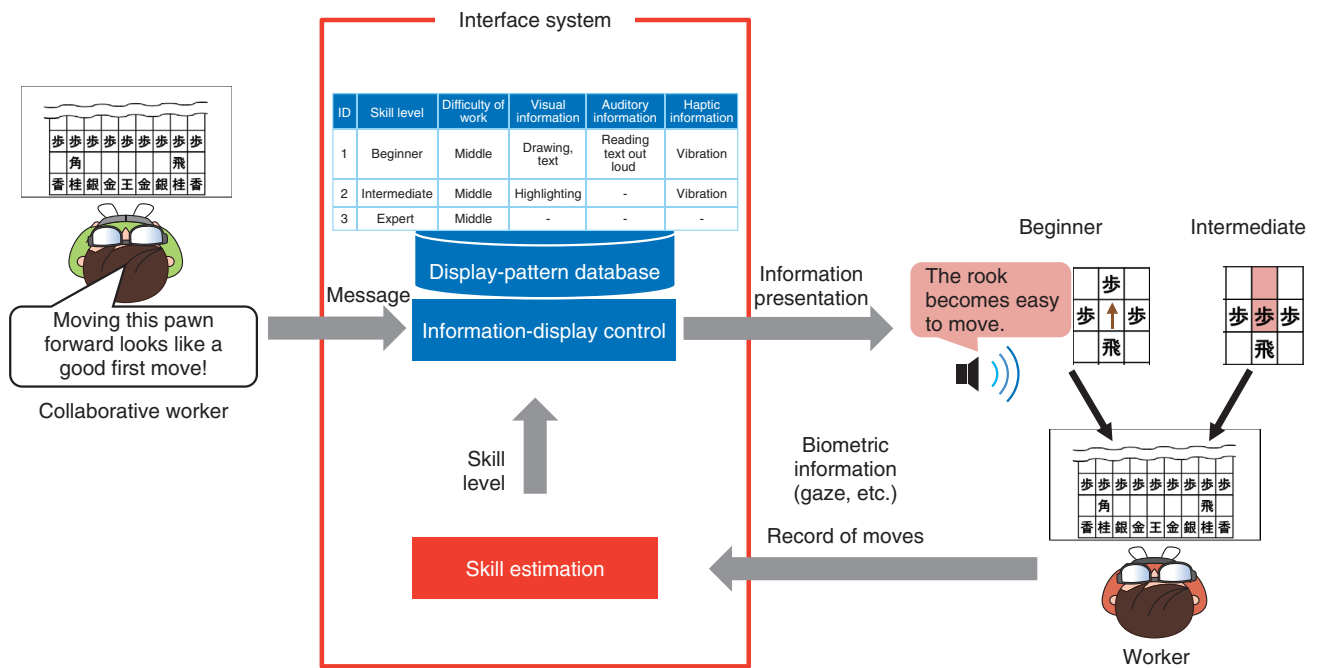
Fig. 1. Overview of interface system for displaying information tailored to skill level.

satisfaction on the basis of whether their opinion is reflected in the results achieved. It is thought that such differences are influenced by the values that participants hold with respect to dialogue. That is to say, differences arise according to what individual participants consider essential for a dialogue and to what they consider to be the factors changing the levels of satisfaction and acceptance [3]. We are testing the reproduction of such values in human digital twins as the inner state of humans to estimate satisfaction level for individual dialogue participants. With this technology, we should be able to estimate the levels of satisfaction and acceptance of each dialogue participant and the reasons for those estimations, follow-up on each participant using expressions tailored to the sensitivities of that person, and provide assistance for improving creativity in consensus building such as through teaming techniques or dialogue intervention to raise the level of satisfaction of all participants. Among the various factors that can be considered to affect the degree of dialogue satisfaction, we describe an estimation technique that focuses on participant positivity and the effect of participant remarks (**Fig. 2**).

First, we considered the extent to which a participant was able to positively participate in a dialogue to be one factor affecting that participant's level of sat-

isfaction and quantified that as a positivity score. We first classified remarks made by a participant as either content-rich or content-poor. Content-rich remarks are those that include content related to the topic of the dialogue. They are relatively long utterances that include many nouns, verbs, and adjectives. Content-poor remarks, on the other hand, are those consisting of backchannel. They are relatively short utterances that include hardly any nouns, verbs, or adjectives. We then defined the positivity score by dividing dialogue data into specific time frames and using the amount of content-rich and content-poor remarks made by a specific participant in each time interval.

Next, considering that the degree to which remarks made by a participant has stimulated the dialogue and helped build consensus can contribute to the satisfaction level of that participant, we quantified that as an impact score. This score is based on how long the dialogue continued on the topic initially proposed by the target participant. We first prepared a list of important words on the basis of a document compiling the results of the consensus reached by the group conducting the dialogue. In the event that such important words appeared in the participant's utterances, we then determined the number of turns that the participant took in uttering those words including related words within the group. We also added weights
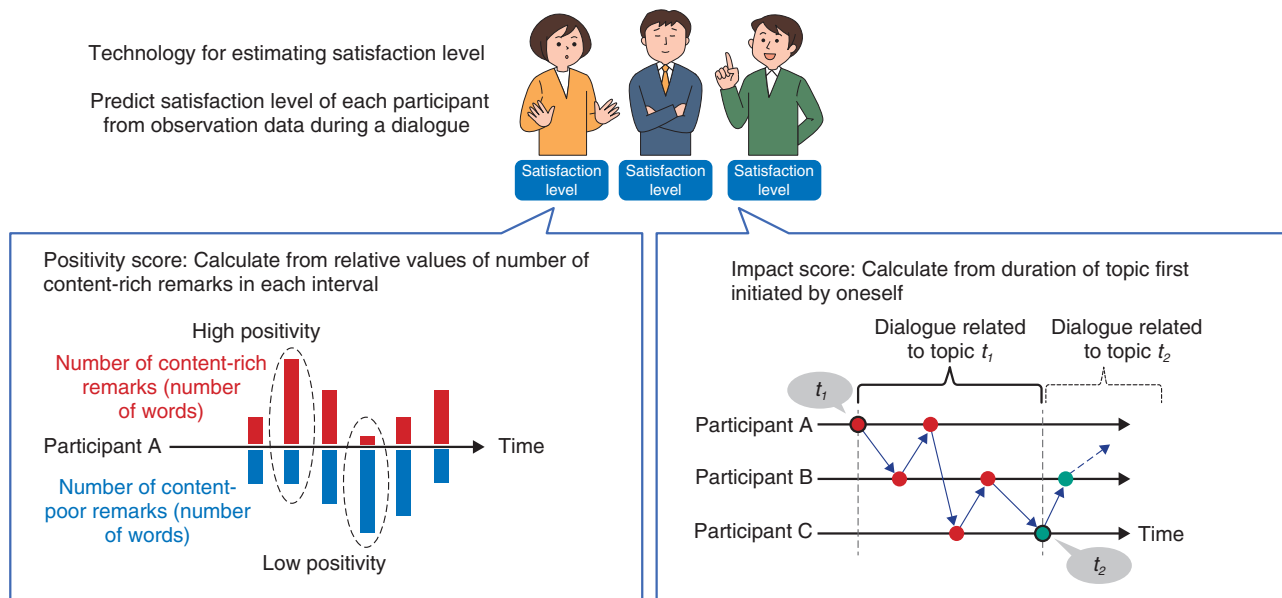
Fig. 2.   Overview of dialogue-satisfaction estimation using positivity toward dialogue and impact of remarks.

depending on whether the person who first used those important words during the dialogue was that participant or another participant and calculated the impact score on the basis of the above characteristics.

We conducted a simulated dialogue among four participants meeting for the first time and estimated the level of satisfaction using the two scores described above. For comparison, we treated the answers given in a questionnaire on satisfaction level as correct. Results revealed that estimation accuracy improved using the positivity score when each topic was created during the beginning portion of a dialogue. It was also found that the effects of the positivity score were minimal in brainstorming and consensus building during the middle and end of the dialogue but that estimation accuracy improved using the impact score. These results indicate that the factors affecting satisfaction level may differ with the time elapsed from the beginning of the dialogue and with the stage of consensus building. In future research, we plan to study the factors affecting satisfaction level for not only dialogue among strangers but also among participants who have already built a personal relationship.

4.   Future outlook

We discussed how to achieve mind-to-mind communications from two important viewpoints. From the viewpoint of what should be conveyed in what manner to convey sensitivity, we described our study on measuring skill level with respect to the receiver's subject matter, on switching to information deemed useful to the receiver depending on skill level, and making it easy to convey the sensitivity that a person wants to convey. To measure skill level, we examined a mechanism that uses sensing data centered around eye tracking and used measurement results to switch between skill levels and the information needed instead of relying on indeterminate elements such as the person's self-reports. From the viewpoint of determining (predicting) how a person feels as an outcome to mind-to-mind communications, we analyzed how a positivity score of dialogue participants could affect the degree of satisfaction with dialogue results. Our initial hypothesis was that dialogue participants with a positive attitude would be satisfied with dialogue results, but the experimental results fell short in terms of clearly indicating the relationship between a positive attitude and satisfaction. However, we are just at the beginning of achieving mind-to-mind communications technology and feel that these studies are a necessary part of the trial-and-error process.

We will continue with the studies presented in this article to achieve mind-to-mind communications technology through expressions tailored to others' sensitivities to demonstrate the true value of diversity

and improve well-being. We look forward to more discussions and studies on mind-to-mind communications with all concerned.

References

[1]   H. Matsubara, "Toward a System That Learns Something Forever," Journal of the Japanese Society for Artificial Intelligence, Vol. 18, No. 5, pp. 564–567, 2003 (in Japanese).
[2]   R. Saijo, T. Sato, S. Eitoku, and M. Watanabe, "Method for Subtly Directing User's Gaze toward Information Breaking into Ongoing Activity," The Transactions of Human Interface Society, Vol. 23, No. 1, pp. 51–64, 2021 (in Japanese).
[3]   M. Nakatani, Y. Ishii, A. Nakane, C. Takayama, and T. Hayashi, "Improving Participant Satisfaction and Quality of Output in Group Dialogue," Correspondences on Human Interface, Vol. 22, pp. 67–74, 2020 (in Japanese).

Ryohei Saijo
Researcher, NTT Digital Twin Computing Research Center.
He received an M.E. from University of Tsukuba, Ibaraki, in 2018 and joined NTT the same year. His research interests include human-computer interaction, information presentation methods, and human digital twin computing.
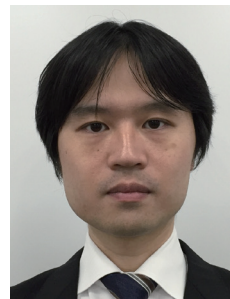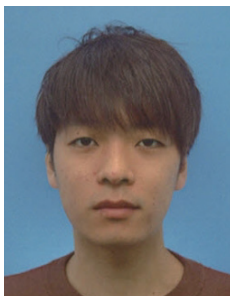
Shohei Matsuo
Senior Research Engineer, NTT Digital Twin Computing Research Center.
He received an M.E. from the graduate school of Global Information and Telecommunication Studies, Waseda University, Tokyo, in 2006. He joined NTT in the same year and engaged in research and development of video coding and its standardization activities. His current research interests include video coding technologies, image processing, human visual systems, and digital twin computing.

Yoko Tokunaga
Research Engineer, NTT Digital Twin Computing Research Center.
She received an M.E. from the graduate school of Kyoto University in 2010 and joined NTT the same year. Her research interests include conversation analysis and digital twin computing.

Iwaki Toshima
Senior Research Engineer, NTT Digital Twin Computing Research Center.
He received an M.E. from the University of Tokyo in 2002 and joined NTT the same year. He received a Ph.D. in engineering from Tokyo Institute of Technology in 2008. His research interests include human-computer interaction, robot audition, and digital twin.

Daichi Yamaguchi
Internship student, NTT Digital Twin Computing Research Center.
He is currently studying at the graduate school of Osaka University. His research interests include multimodal interaction and sensing.

Takao Kurahashi
Senior Research Engineer, NTT Digital Twin Computing Research Center.
He received an M.E. in industrial engineering from Hosei University, Tokyo, in 2002 and joined NTT the same year. His research interests include human-computer interaction and robotics.

Lidwina Andarini
Researcher, NTT Digital Twin Computing Research Center.
She received an M.E. from Nara Institute of Technology in 2016 and joined NTT the same year. Her research interests include the human digital twin platform and computer networking.

Shiro Ozawa
Senior Research Engineer, Supervisor, NTT Digital Twin Computing Research Center.
He received a B.E. and M.E. from Tokyo University of Mercantile Marine (now, Tokyo University of Marine Science and Technology) in 1997 and 1999. After joining NTT as a researcher, he worked for NTT Cyber Space Laboratories (now, NTT Human Informatics Laboratories). His fields of interest are computer vision, video communication, user interfaces, human-computer interaction, and 3D displays. He is a member of the Institute of Electronics, Information and Communication Engineers, the Institute of Image Information and Television Engineers, and the Institute of Image Electronics Engineers of Japan.