

Visual Feedback to Reduce Influence of Delay on Video Chatting

Kazuyoshi Murata, Masatsugu Hattori, and Yu Shibuya

Kyoto Institute of Technology

Matsugasaki, Sakyo-ku, Kyoto 606-8585 Japan

{kmurata, shibuya}@kit.ac.jp, hattori@hi.cis.kit.ac.jp

Abstract. When there is a certain delay in video chatting, participants often misunderstand other partners' response and make unintended interruptions. In this paper, to overcome these problems, we present two kinds of visual feedback: Scroll Wave Indicator and Afterimage Indicator. An experiment was conducted to confirm effectiveness of these indicators. The result of experiment showed that the Scroll Wave Indicator helped participants to understand the remote participant's response timing and decreased unintended interruptions even if there was 2 [sec] round-trip delay.

Keywords: visual feedback, delay time, video chatting.

1 Introduction

Video chat systems have been popular to support communication among spatially distributed friends, couples, or families. However, there are unavoidable problems caused by a delay in such video systems. For example, Ruhleider and Jordan [1] showed unintended interruptions as one of the particular problems due to delay for video-based remote communication.

Fig.1 illustrates an example of unintended interruptions in delayed situation. If there is not a delay, an utterance of participant A is transmitted to participant B's site right away and participant B can receive the utterance immediately. However, if there is a certain delay, each participant receives delayed utterance of other participant. In this example, participant A utters something (A1), and then waits responses from participant B. However, there is no response for a while because of delay. After a certain time, participant A starts another utterance (A2) because he/she assumes that participant B does not want to respond. Then participant A is interrupted by utterance of participant B (B1') suddenly and he/she is confused. At participant B's site, participant B receives a delayed utterance (A1'). Participant B starts responding to participant A (B1) after a certain thinking time, and then he/she is interrupted by next utterance of participant A suddenly (A2'). Participant B is also confused because the utterance of A2 is not related to A1' and B1. These interruptions are called unintended interruptions.

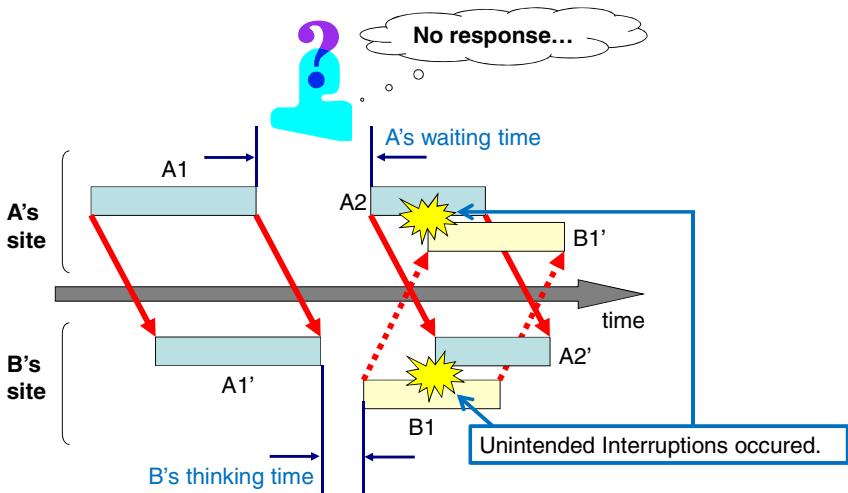


Fig. 1. An example of unintended interruption

Thus, when there is a certain delay in video chatting, participants often misunderstand partners' response and make unintended interruptions. These misunderstandings and unintended interruptions disrupt conversation and make participants confused. In this situation, because participant A cannot know if there is some delay or not, he/she cannot distinguish the waiting time due to the delay from the thinking time of participant B. However, if the participant can recognize the relationship between the timing of the playback of his/her utterance at remote site and the timing of other participant's response to it, he/she can understand the response timing of other participant and avoid these unintended interruptions.

In this paper, we propose a visual feedback method to overcome these problems. The proposed method consists of two kinds of indicators: Scroll Wave Indicator and Afterimage Indicator. There are many studies aimed to reduce delay of video-based telecommunication. However, it is impossible to get rid of delay completely. The purpose of our proposed method is to reduce the influence of delay by improving user interface of video chat system instead of removing the delay itself.

2 Proposed Method

Our proposed method provides visual feedbacks which show the status of voice and video playback of remote site. Fig.2 illustrates an example of display timing of these feedbacks. The display timings of these feedbacks are delayed for a length of a one-way delay from actual start timing and stop timing of the voice and video playback. That is, every start timing and stop timing of the visual feedbacks is delayed for the length of a round-trip delay from these timings of the participant's own utterances. The length of his/her partner's response timing at remote site (T_R) is equal to the length of it at local site (T_L) consequently.

The participant who sees these feedbacks can recognize when his/her partner responded to his/her utterance. Consequently, misunderstandings of the response timing of his/her partner and occurrences of unintended interruptions are prevented.

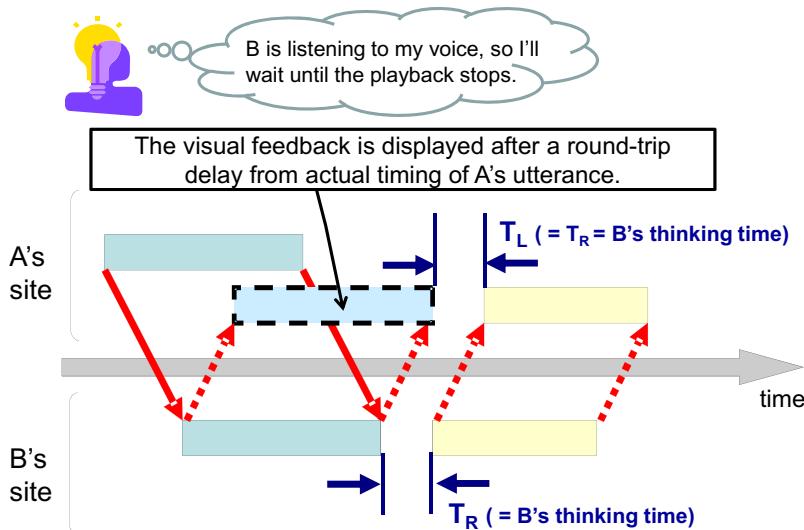


Fig. 2. An example of display timing of proposed feedbacks

We introduced the Progress Bar Indicator to decrease unintended interruptions of voice chatting in our previous study [2]. However, it is difficult to apply this indicator to video chatting. Kawashima et al. proposed Visual Filler to solve these problems [3]. However, the Visual Filler requires inferring the end timing of the user's utterance automatically. This inferring process is not needed for our proposed method.

2.1 Scroll Wave Indicator

Fig. 3 shows an appearance of our proposed method. The Scroll Wave Indicator shows a state of a voice playback of a user and his/her partner. A waveform of upper area means an amplitude of the partner's voice playback. This waveform appears from left end of a black rectangle area and scrolls toward left side of the indicator at a constant rate as the voice of the partner is played at the user's site. That is, the waveform is displayed after a length of a one-way delay from actual start and end timing of an utterance of the partner.

A waveform of lower area means an amplitude of the user's voice. The waveform also displayed after a length of a round-trip delay. The waveform appears from the black line, and then it scrolls toward left side of this indicator at a constant rate. This black line is shifted to the right from the center of the lower area. At the center of the lower area, a dashed line is drawn. The distance between the black line and the dashed line means the length of the round-trip delay. That is, when the waveform reaches the halfway between the black line and the dashed line after it scrolled for the length of a

one-way delay, the user's voice starts to play at the partner's site. Subsequently the waveform reaches the dashed line after it scrolled for the length of a round-trip delay. The waveform over the dashed line means the voice of the user which was being played at his/her partner's site when the voice of the partner which is being played at the user's site was uttered by the partner.

By comparing the positional relation between the partner's waveform and the user's waveform, the user is able to understand the timing when the partner's utterance which is being played at the present moment was uttered at his/her partner's site. In addition, the user is able to infer which part of his/her utterance was played at the partner's site from the figure of the waveform. The user is able to refrain from unnecessary utterance, e.g. "Can you hear me?"

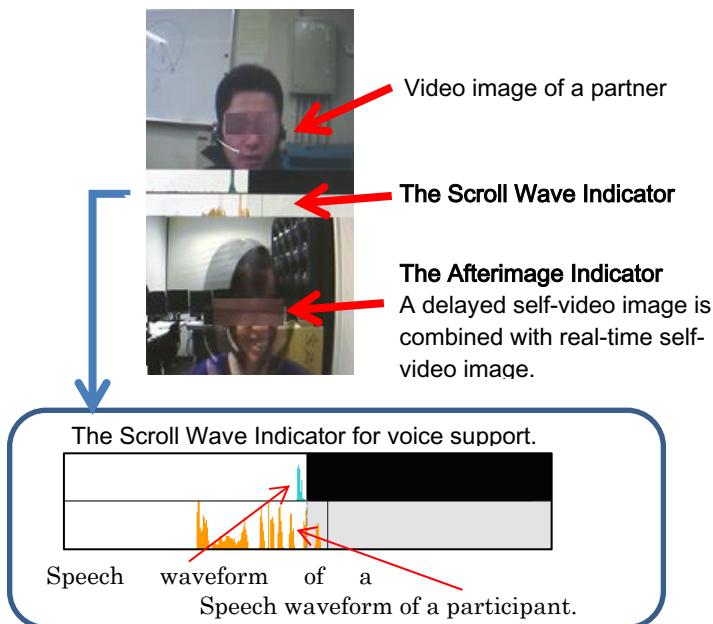


Fig. 3. An example of display timing of proposed feedbacks

2.2 Afterimage Indicator

Through video chatting, participants often communicate not only verbal messages but nonverbal messages, e.g. notations or gestures. Accordingly we propose the Afterimage Indicator as a visual feedback which presents the timing when such nonverbal messages are played through a video image at a remote site. An appearance of the Afterimage Indicator is also shown in Fig.3. In the Afterimage Indicator, a video image of a user which is delayed for a length of a round-trip delay is overlapped with a real time video image of the user. This overlapped image helps the user to

understand the timing when the user's nonverbal messages are played at his/her partner's site.

3 Experiment

We conducted an experiment to confirm the following effectiveness of our proposed method.

- The proposed method helps a user to understand the timing of his/her partner's response.
- The proposed method decreases unintended interruptions.

3.1 Procedure

Six pairs of volunteers, ranging in age 22 to 25 years, were participated in this experiment. They were four students, seven graduate students and one alumnus of our university. They had been known well each other and made daily face-to-face communication. Each participant of a pair was in separate rooms and both of them wore headsets. They talked to each other through the headsets. They were asked to make three minutes discussion. The discussion topic was to line up three candidates for certain goal. For example, "You are going to have lunch. What kind of lunch will you have?" They were asked to make the discussion with our proposed method. Moreover, they were asked to do it with usual video chat system (hereafter, "conventional method"). The video chat system had two video windows of the partner and the participant himself/herself, and they talked each other through headsets. This experiment was performed under three kinds of round-trip delay conditions: 0, 1 and 2 [sec].

3.2 Measures

We measured rates of unintended interruptions in this experiment. In addition, participants were asked to answer a questionnaire for subjective evaluation after each discussion.

Rate of unintended interruptions. The rate of unintended interruptions (hereafter, "RUI") was defined as a following equation.

$$\text{RUI} = \text{NIU} / \text{NUP}.$$

NIU: Number of utterances interrupted by a partner that were uttered by a participant successively.

NUP: Number of utterances of a participant.

Participants were able to communicate by using both of voice and a user's body motion through headset and video image on PC in this experiment. Therefore we considered that there were two kinds of unintended interruptions. One was unintended interruptions by a partner's voice through a headset, and the other was unintended interruptions by a partner's body motion through a video image. We defined the former as RUI by voice, and the latter as RUI by video. The Scroll Wave Indicator was expected to decrease RUI by voice and the Afterimage Indicator was expected to decrease RUI by video.

Questionnaire. The participants were asked to answer a questionnaire including following questions.

1. You could understand the timing when your partner responded to your utterance.
(1: disagree – 5: Agree)
2. How often did you watch the self-video image on PC? (1: didn't watch it – 5: always watched it)

4 Result and Discussion

Fig. 4 shows the result of RUI by voice. In the case of 2 [sec] delay condition, a t-test revealed a significant difference between the RUI by voice with proposed method and that with conventional method ($p<.05$). There was no significant difference in the case of 1 [sec] condition. Fig. 5 shows the average score of question #1. An ANOVA showed a significant main effect of delay condition in the case of using the conventional method ($F_{2,11}=4.696$, $p<.05$). A post-hoc test revealed that the score of 2 [sec] delay condition was significantly lower than that of 0 [sec] delay condition. However, there was no significant difference in the case of using the proposed method.

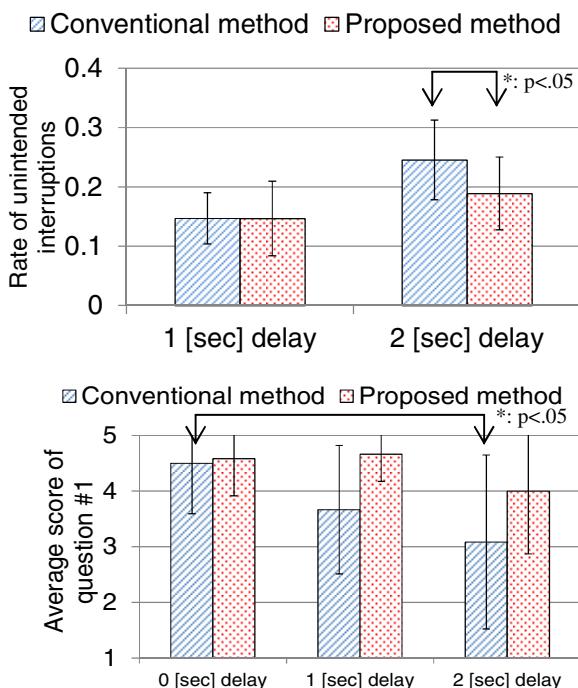


Fig. 4. The average score of question #1

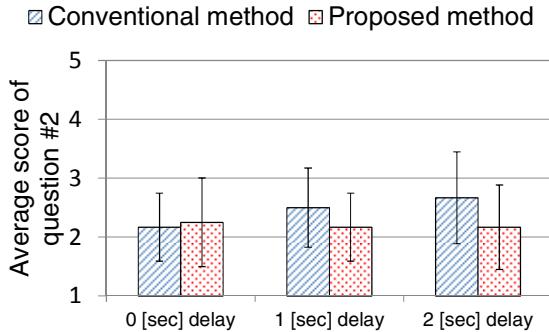


Fig. 5. The average score of question #2

Thus, as the length of the delay was extended, it became hard to understand the timing of a partner's response with the conventional method. On the other hand, even if the delay was 2 [sec], the participant with the proposed method was able to understand the timing easily. These results indicate that the Scroll Wave Indicator helps a participant to understand the timing of his/her partner's response. Furthermore, these results suggest that the Scroll Wave Indicator decreases unintended interruptions.

With regard to the Afterimage Indicator, the effect of the Afterimage Indicator was unable to be analyzed because there were few unintended interruptions by a participant's body motion. In this experiment, the task which performed by participants did not require them to use nonverbal expressions by body motion. They therefore needed not to use a video image on PC. The score of question #2 is shown in Fig. 6. There was no difference between the score of the proposed method and that of the conventional method. This result indicates that participants barely watch their self-video image. Some participants commented that the overlapped delayed self-video image was obstacle to see the usual self-video image. To solve this problem, we need to improve the Afterimage Indicator which combines the effectiveness of the indicator and visibility of the self-video image.

5 Conclusion

In this paper, the Scroll Wave Indicator and the Afterimage Indicator are proposed and evaluated experimentally. The result of evaluation showed that the Scroll Wave Indicator made participants understand the remote participant's response timing and decreased unintended interruptions even if there was 2 [sec] round-trip delay. However, the effectiveness of the Afterimage Indicator was not shown. Future work will involve an experiment using tasks that should be performed with both of verbal and nonverbal communications.

References

1. Ruhleder, K., Jordan, B.: Co-Constructing Non-Mutual Realities: Delay-Generated Trouble in Distributed Interaction. *Computer Supported Cooperative Work* 10(1), 113–138 (2000)
2. Murata, K., Nakamura, M., Shibuya, Y., Kuramoto, I., Tsujino, Y.: Visual feedback to reduce the negative effects of message transfer delay on voice chatting. In: Smith, M.J., Salvendy, G. (eds.) HCII 2007. LNCS, vol. 4558, pp. 95–101. Springer, Heidelberg (2007)
3. Kawashima, H., Nishikawa, T., Matsuyama, T.: Visual filler: Facilitating Smooth Turn-taking in Video Conferencing with Transmission Delay. In: CHI 2008 extended abstracts on Human factors in computing systems, pp. 3585–3590 (2008)