Influence of network delay on ensemble application

Isamu Arimoto[1], Kenji Hikichi[2], Kaoru Sezaki[3],Yasuda Yasuhiko [2]
[1]Institute of Industrial Science, University of Tokyo
4-6-1, Komaba, Meguro-ku, Tokyo 153-85, Japan
[2]School of Science Engineering, Waseda University
3-4-1, Okubo, Shinjuku-ku, Tokyo 169-8555, Japan
[3]Center for Spatial Information Science , University of Tokyo.
4-6-1, Komaba, Meguro-ku, Tokyo 153-85, Japan

Abstract – Haptic device has been brought to public attention .We are developing networked music ensemble system with haptic devices. Pairs of user connected network play drums. In such systems, the impairment of network QoS such as network delay or packet loss cause degradation of collaboration performance among users. The purpose of this study is to examine the influence of network delay and to examine the effect of distributed synchronization and media synchronization, and to investigate role of haptics for this application in such haptic audio communication system.

Keywords – Haptic, Audio, Ensemble, Collaboration, Multimodal

I. INTRODUCE

Haptic device has been brought to public attention.

In collaboration between users with haptic device, degradation of network QoS(Quality of service) such as network delay or packet loss influences QoS of haptic , visual and audio media.

We are developing haptic audio application. In this application, contact between haptic cursors and virtual drums make sound. User can play a drum in concert over the network. We are investigating influence of network delay on such application, especially more sensitive music ensemble application. In this paper, we show how to influence synchronization control and haptic media for ensemble application.

II. RELATION WORK

Some systems that the user in real time plays in concert mutually on the network are proposed. The influence of the delay is measured in [2]. There are two users in a separate room. The task of beating the hand while mutually hearing other user's clapping with the headphone was done. The microphone is set up near the hand. The speed chosen beforehand at random is remembered with a metronome. It plays in concert so that it is presented in the experiment and the rhythm should not shift like keeping the rhythm though other user's sound is heard. The gap between the rhythm of the performance when experimenting and the rhythm felt before it experiments is evaluated.

The tempo of the performance quickens when the delay is 11.5ms or less, and the performance slows in a delay that is bigger than it. This is heard in the performance by a real space delaying the sound because of the spread delay. When the delay of the network becomes small more than spreading late, it hears of the sounds earlier than the performances by a real space.

In virtual environment, the method of haptic rendering and generating the sound can be controlled adaptive. For this reason, ensemble application is optimized and it is likely to be able to optimize it to a delay different from a real space. First of all, the influence of the network delay in this application was examined in this paper.

I explained the synchronization control. To measure the influence, I explain the system that made it this time. Next, the experiment method of measuring the influence of the network delay in the system is described. It experimented the evaluation of the influence that the network delay and the synchronization control gave the user at the end and the influence by haptic presentation.

III. APPLICATION

In this section, we explain haptic audio ensemble application we developed. There are a drum and a cursor in Virtual Environment. User can move this cursor throw the haptic device PHANToM. User can beat the drum throw this cursor. When this cursor contacts this drum, sound is emitted. Loudness of sound of the drum is proportional to magnitude of reaction force, which apply the user. The contact information includes timestamp and magnitude of reaction force send to another host. At receiver host, sound is emitted when contact information is received. Only sending contact information, information of position of the drum and position of the cursor does not send to another host. Thus, there is no haptic interaction between users and visual of another host does not display. For synchronization control to say later, there is possibility that user misunderstand own sound as another sound, two sound sources is prepared.

IV. SYSTEM OVERVIEW

This section we show system overview. Network architecture of this system applies peer-to-peer architecture. Each hosts equipment visual display, headphone and PHANToM as haptic interfaces. Host manages information of VE, include cursor's position same as PHANToM position

and drum position. When the cursor contacts the drum, reaction force is calculated and reaction force is applied to user. Collision detection process at each client, only collision information includes magnitude of sound send to another host. Information send every 1ms interval in time with PHANToM's servo loop. When each hosts received the packet, sound of another host present user immediately, but display of own sound and visual is waiting for queued access method.
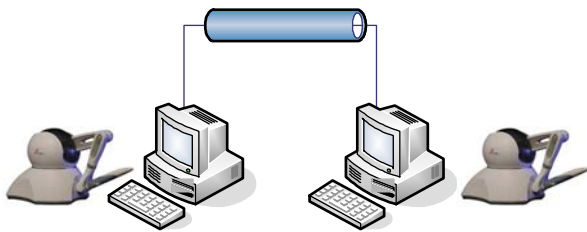

Figure 1. System overview

V. INFLUENCE OF HAPTIC MEDIA

In the ensemble application, when the sound from other hosts is delayed, the user will feel the uncomfortable. In addition, because the sound is presented by contact with the cursor that the drum and the user operate, timing to which the user acknowledges coming in contact becomes important in the drum application. Therefore, the user will feel sense of discomfort by generating the gap of the audio, visual, and haptics.

In [1], only audio was synchronized for distributed synchronization. At this time, some subjects shut eyes to match the synchronization of the sound with the other user and was doing the evaluation task from the appearance of the evaluation of subjects. A lot of users who were taking timing without valuing information on the sight to concentrate on audio and haptics were seen.

In the drum application that presents haptics from this, the gap of haptics and the audio has the possibility that the influence is large than the gap of the image and the voice. Then, it compared it with the case to present haptics and not doing in ensemble application.

VI. SYNCHRONIZATION CONTROL

The timing of the generation of the sound by own performance includes two timing. It is a method of generating the sound at the same time as own cursor's coming in contact with the drum. The other is a method of generating the assignation house sound only by the other user's delay at the time that comes in contact. The former is a method of valuing the synchronization between media of haptic and audio. The latter is a method of considering distributed synchronization of audio. The influence of the delay in each method is examined.

Media Synchronization

I explain the media synchronization in this application. In the network environment with a fixed delay, it plays in concert by host A and B. Host A and host B do the same performance while seeing the same score.

Each host starts performing at the same time as the clock synchronizes completely. When own drum is beaten, the drumbeat is generated in host A. But, it hears of other user's sound delaying only delay of the network. For host A, because I hear it delaying other user's sound, it is felt that other user's rhythm is late. Because it tries to match the following rhythm to the other user, time to beat the following drum is delayed. Even host B similarly hears it late of the sound of host A only at delay time. It is expected that the rhythm of the performance slows gradually if adjust the result of hearing delaying other user's sound both host A and host B.

Distributed Synchronization

I explain the case to do distributed synchronization. Audio and visual is synchronized on the condition of already-known fixed delay. It sounds it only delay time with the other user synchronously it not being generated at once by the sound when own drum comes in contact with the cursor.

As a result, after it meets only by delay time, the sound is generated in the user if the user forecasts timing where other party's sound is generated. Therefore, the delay of the performance by the influence of the delay might decrease. However, the sense of incompatibility will be caused for my performance with the reaction force where the generation of the sound is caused by contact with the drum because it shifts. When the delay is large, the user's correcting the gap with the other user's performance so that own operation and the generation of the sound may shift greatly might become difficult.

Synchronization Control

In [1], only audio media was synchronized between clients. In this technique, the gap of both haptic and visual media is caused for audio media. Not only the audio media but also the visual and haptic media can synchronize because it displays media of audio, visual and haptic media to the user in the haptic audio application. However, the haptic media present the user excessive power by synchronizing in proportion to the size of the delay. As a result, problems such as should being able to give the user a big sense of incompatibility and doing steady operation are caused. Therefore, haptic media cannot synchronize.[3].

In this paper we implemented the system with synchronization of audio and visual media without synchronization of haptic media. The comparison when this synchronization was done and when it did not synchronize was done.

VII. EVALUATION TASK

We explain evaluation task for this system. Subjects listen to rhythm of rhythm box (80BPM), and subjects play enough to remember this rhythm. In evaluation task, subjects play drum keeping this rhythm (80BPM) and beat simultaneously other user's sound. Each subjects beat drum in the same breathes. Beginning of this task, using timer without delay task will be begin simultaneously at each hosts. Then there is no delay of start timing. Subjects play drum at 20 second, afterward subjects will evaluate. Evaluation way adopts DSIS [4]. In first task, subjects play in reference condition without delay. In second task, subjects play in degradation condition with delay. After this, users evaluate degradation condition comparing reference condition. 8 kinds of delay was prepared, subjects play under such condition with randomly delay. Delay was known for system, length of queue is adjusted.
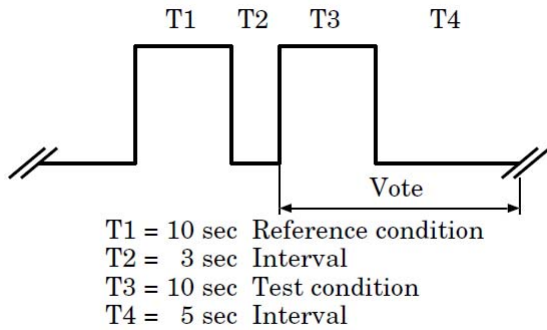

Figure 3. Subjective evaluation Result: with haptics

Figure 2. Evaluation order

```
T1 = 10 sec  Reference condition
T2 =  3 sec  Interval
T3 = 10 sec  Test condition
T4 =  5 sec  Interval
```

The result of the evaluation experiment is shown. A horizontal axis is a size of the delay, and a vertical axis is MOS(Meaning Opinion Score) value.

Figure 3 shows subjective evaluation result without haptic media. Figure 4 shows subjective evaluation result with haptic media. Both figures compare synchronization of audio and visual media. From the result of which graph, the result of user's subjectivity evaluation has improved when there is no synchronization. Moreover, the difference was not seen so much by user's subjectivity evaluation result for the presentation of haptics. The difference was not seen so much by user's subjectivity evaluation result for haptic media. In the haptic audio application, distributed synchronization does not improve user QoS.
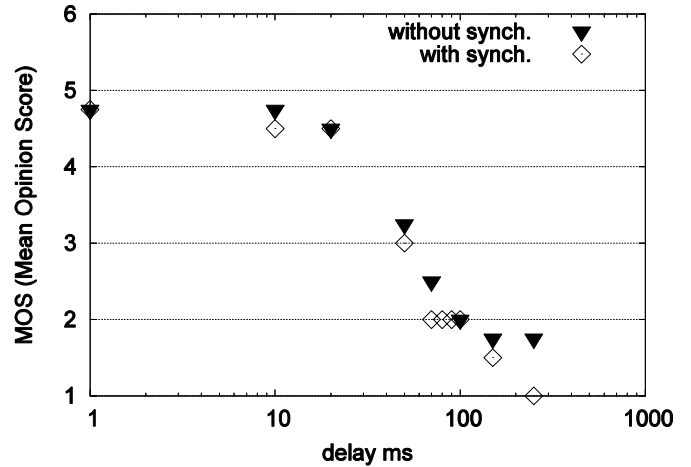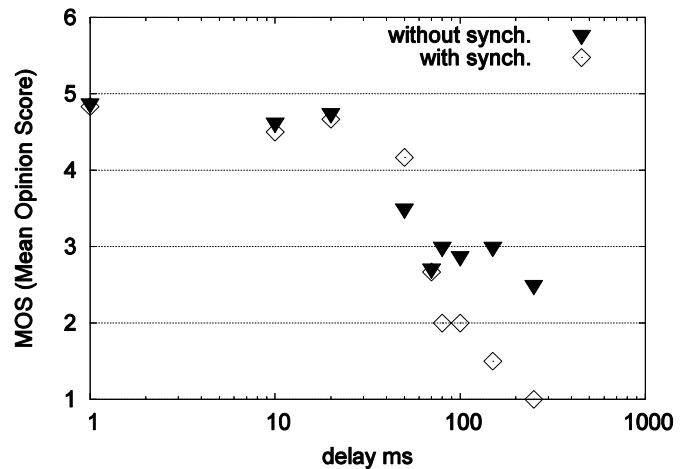

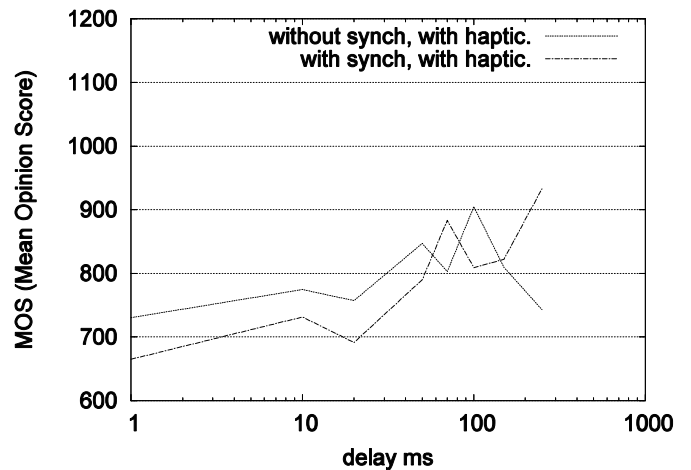Figure 4. Subjective evaluation Result: without haptics

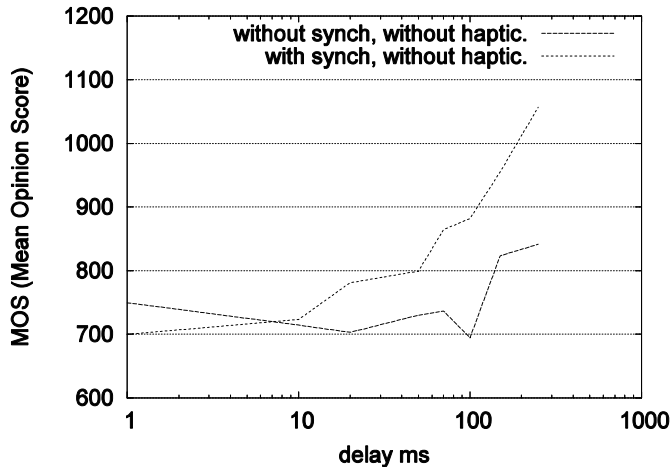Figure 5. Objective Evaluation Result: with haptics

21

Figure 6. Objective Evaluation Result: without haptics

Figure 5 shows objective evaluation result with haptic media. Figure 6 shows subjective evaluation result without haptic media. In both cases, the interval of time to pound on a drum becomes large increasing of network delay.

The interval of time beating a drum with synchronization has shortened regardless of the presence of haptics. It shifted while the case synchronously processed without presenting haptics was justifiable.

VIII. DISSCUSION

From subjective result, the result of user's subjectivity evaluation has improved when there is no synchronization in haptic audio application. Moreover, The interval of time beating a drum with synchronization has shortened regardless of the presence of haptics.

It waits only by delay time after the user operates it with PHANToM when synchronizing and the cursor in a VE moves. Moving the cursor as the user intended it as the delay grows becomes difficult. When the delay increases, the user's moving the cursor as intended becomes difficult. It is difficult and to understand timing that own cursor comes in contact with the drum as a result. Keeping own rhythm constant becomes difficult. It is expected that matching the timing of own performance and the other user's performance by this influence became difficult. Therefore, it is guessed that the gap of synchronization between media influenced result of distributed synchronization.

However, there was little increase in the delay at intervals of time that even user beats on a drum in the condition with delay of 70-100ms with synchronization .Own operation and sound are delayed with synchronization. Therefore, the user can expect the size of the delay, and can adjust it voluntarily. It is thought that the delay of the performance by the influence of the delay could be controlled. The influence on the performance could be suppressed by synchronizing though user QoS was depredated. Because the user can move it like having intended the movement of own cursor when not synchronizing, the user can surely control the sound by his

performance. Therefore, it is easy to keep the rhythm to be constant. The user can correct it to other party's rhythm even if there is a small delay. When network delay become large, suitable of the rhythm with the other user becomes under a difficult situation. Because own performance can be operated as thought, it is thought that the result of the subjectivity evaluation is not deteriorated easily.

The difference was hardly seen about the presence of haptics. If the user remembers the rhythm at the beginning of the task for the experiment only of carving the rhythm of equal intervals in this experiment, the rhythm can be kept accurate now. There was no difference like it if there was no sense of incompatibility for own operation. Moreover, because the timing of the generation of contact and the sound is understood more accurately from presenting haptics, it is thought that it becomes easy to keep the rhythm to be constant. On the contrary, the gap of the performance with the other party is also comprehensible. Those influences were not so seen in this evaluation result.

IX. CONCLUSION

In this paper, the influence that the network delay exerted on the user in the concert application was investigated.

I explain the influence of hapics and the influence of synchronization, and I report on the evaluation experiment to it.

IEEE HEADING (WITHOUT NUMBER)

[1] I. Arimoto, K. Hikichi, K. Sezaki, and Y. Yasuda. Study on media synchronization in haptic and audio application. 'Proc. of IEICE Society Conference 2005 B-11-1 (in Japanese), Mar. 2005.

[2] I. Arimoto, K. Hikichi, K. Sezaki, and Y. Yasuda. Study on media synchronization in haptic and audio application. 'Proc. of IEICE Society Conference 2005 B-11-1 (in Japanese), Mar. 2005.

[3] K.Hikichi, H.Morino, I.Arimoto, I.Fukuda, S.Matsumoto, M.Iijima, K.Sezaki, and Y.Yasuda. Architecture of Haptics Communication System for Adaptation to Network Environments. IEEE ICME2001, Aug. 2001.

[4] I.-R. Recommendations. Recommendation 500-10;. Methodology for the subjective assessment of the quality of. Television pictures. ITU-R Rec. BT.500-10, Mar. 2000.