

Human Perception of Group Synchronization Error in Remote Learning: Dependencies of Voice and Video Contents in One-Way Communication

Hay Mar Mo Mo Lwin¹, Yutaka Ishibashi², Khin Than Mya¹

¹Faculty of Computer Systems and Technologies, University of Computer Studies Yangon, Yangon, Myanmar

²Graduate School of Engineering, Nagoya Institute of Technology, Nagoya, Japan

Email: ishibasi@nitech.ac.jp

How to cite this paper: Lwin, H.M.M.M., Ishibashi, Y. and Mya, K.T. (2022) Human Perception of Group Synchronization Error in Remote Learning: Dependencies of Voice and Video Contents in One-Way Communication. *Int. J. Communications, Network and System Sciences*, 15, 31-42.

<https://doi.org/10.4236/ijcns.2022.153003>

Received: February 3, 2022

Accepted: March 6, 2022

Published: March 9, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This paper examines dependencies of voice and video contents on human perception of group (or inter-destination) synchronization error in remote learning by Quality of Experience (QoE) assessment. In our assessment, we use two videos and three voices (two voices for one video and one voice for the other video). We also investigate influences of silence periods in the voices and temporal relations between the voices and videos (called the tightly-coupled and loosely-coupled contents here). The voices are spoken by a teacher according to the videos. Each subject as a student assesses the group synchronization quality by watching each lecture video and the corresponding explanation voice, and then the subject answers whether he/she perceives the group synchronization error or not. As a result, assessment results illustrate that silence periods mitigate the perception rate of the error, and we can also find that we can more easily perceive the error for tightly-coupled contents than loosely-coupled ones.

Keywords

Remote Learning, Voice, Video, Group Synchronization Error, Human Perception, QoE Assessment

1. Introduction

Media synchronization is important in networked multimedia applications, such as video conferencing, networked games, and remote learning [1] [2] [3] [4]. In such applications, the temporal relationships among multiple media streams (for

example, voice and video) are important [5] [6] [7] [8] [9].

The applications sometimes need to output multiple media streams synchronously at all the terminals (*i.e.*, group (or inter-destination) synchronization) [10] [11] [12] [13]. If the output timings of each *media unit* (*MU*), which is an information unit such as a video picture and a voice packet for media synchronization, among multiple terminals are different from each other, the quality of experience (QoE) [14] [15] may be damaged seriously. This may influence the learning effect in remote learning, for example.

To solve the problem, it is necessary to carry out group synchronization control [16]-[21], which adjusts the output timing of *MU* among multiple terminals (or destinations). In [22] and [23], two types of error ranges are employed under media synchronization control. One is the imperceptible range, in which users cannot perceive the error, and the other is the allowable range, in which users feel that the synchronization error is allowable. Thus, it is important to clarify human perception of group synchronization errors. However, the perception has not sufficiently been clarified so far.

Some papers clarify the human perception of media synchronization errors such as lip synchronization and pointer synchronization [24] [25] [26]. In [24], Steinmetz clarifies the human perception of lip synchronization error and pointer synchronization error. He concludes that the lip synchronization errors within about ± 80 ms are hardly perceivable, and almost everyone perceives the errors beyond around ± 160 ms; also, the pointer synchronization errors within about -750 ms and $+500$ ms are hardly perceivable, and almost everyone perceives the errors less than around -1000 ms or larger than about $+1250$ ms. The results mean that the human perception depends on voice and video contents. In [25], Staelens *et al.* investigate the influence of lip synchronization error on the ability to perform real-time language interpretation during video conferencing. Younkin and Corriveau obtain the minimum amount of audio-visual synchronization error that can be detected by users [26]. However, in [24] [25] [26], they do not handle the human perception of group synchronization error.

In this paper, therefore, we clarify the human perception of group synchronization error in remote learning. To investigate the dependencies of voice and video contents on the human perception, we use two types of video contents and the corresponding voice contents with/without silence periods; that is, how tightly the voice and video contents are related to each other.

The remainder of this paper is organized as follows. Section 2 describes the group synchronization in remote learning. Section 3 explains the assessment method. Section 4 presents and discusses assessment results. Section 5 concludes the paper.

2. Group Synchronization in Remote Learning

In remote learning, it is necessary to perform the group synchronization control [15] [16] [17] [18] [19], which tries to output each *MU* simultaneously at all the

different terminals in multicast communication. If the control is not carried out, the MU cannot be outputted at the same time at the terminals; that is, the group synchronization error occurs.

The configuration of our remote learning system is shown in **Figure 1**. The system consists of a single teacher terminal, $N (\geq 1)$ student terminals, and a file server. The teacher terminal uses a microphone, and each student terminal employs a headset. The file server multicasts a video stream to the teacher terminal and all the student terminals. The teacher orally explains the video contents while watching the video. The voice stream of the teacher captured via the microphone is multicast from the teacher terminal to all the student terminals. Each student listens to the teacher's voice while watching the same video.

In **Figure 1**, the video delay from the file server to the teacher terminal is denoted by D_{ft} , and the video delay from the file server to student terminal i ($1 \leq i \leq N$) is denoted by D_{fsi} . Also, the voice delay from the teacher terminal to student terminal i is denoted by D_{tsi} . We can define the two types of group synchronization errors as the differences among D_{ft} , D_{fs1} , ..., D_{fsN} and those among D_{ts1} , ..., D_{tsN} in **Figure 1**. We assumed here that the global clocks (that is, clock ticks at the sources and destinations have the same advancement, and the current location times are also the same [10] [11]) are used at all the terminals and the file server.

3. Assessment Method

In our assessment system, we set $N = 1$, and $D_{ts1} = 0$ for simplicity. In this case, the group synchronization error about the video is expressed by $D_{ft} - D_{fs1}$. If the

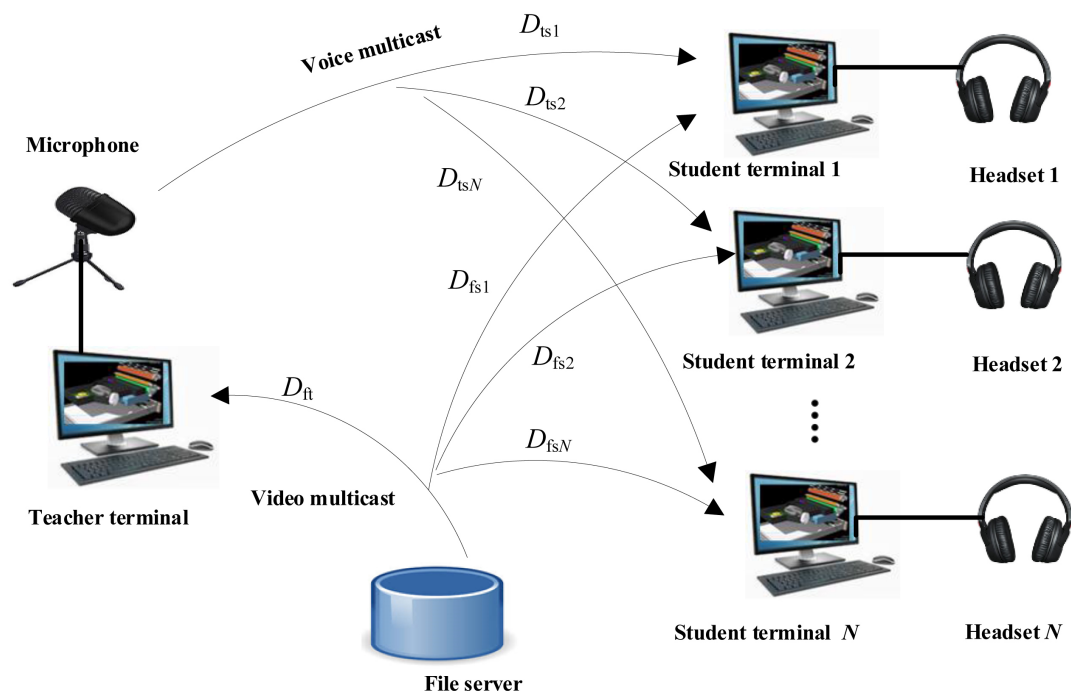


Figure 1. System configuration of remote learning.

voice delay exists (*i.e.*, $D_{ts1} \neq 0$), the student can start to hear the voice at $D_{ft} + D_{ts1}$. Therefore, if $D_{ft} + D_{ts1} - D_{ts1} = 0$, the student does not perceive any synchronization error. We examined the influence of D_{ts1} on the human perception of group synchronization error [26]. As a result, we found that the perception rate of the error depends on the group synchronization error plus the voice delay. Thus, we can set $D_{ts1} = 0$ without losing generality in this paper. Because the group synchronization error is $D_{ft} - D_{ts1}$, we can produce the error by making the difference $D_{ft} - D_{ts1}$ in starting time between the voice and video at the student terminal as shown in **Figure 2**.

In **Figure 2**, student terminal 1 saves video files as the video server. Also, the terminal stores voice files which have been recorded in advance to speak always in the same way in the assessment; one of the authors played the teacher's role and saved her voice. Each subject used the headset at student terminal 1. We produced the group synchronization error by changing the start times of voice and video outputs at student terminal 1. At the beginning of assessment, we presented the perfect situation (*i.e.*, the group synchronization error is zero) to each subject; that is, we started to output the voice and video files simultaneously at the student terminal. We used the single stimulus method [27] for QoE subjective assessment. However, when the subject requested the perfect one during the assessment, we showed it. The subject did not know the value of the error presented in the assessment.

After presenting each error, we asked each subject (student) the following question: "Did you perceive the group synchronization error?" The subject answered either "Yes" or "No." He/she judged whether error was perceived or not by monitoring the temporal relation between the teacher's voice and displayed video contents.

To examine the dependency of voice and video contents, as shown in **Table 1** and **Figure 3**, we used three voices (called Voices 1, 2, and 3 here) and two videos (called Videos 1 and 2) in terms of the following two factors: Temporal

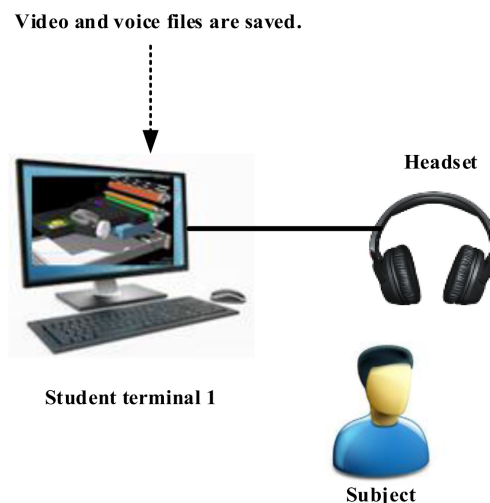


Figure 2. Assessment system.

Table 1. Voice and video contents.

	No silence period	Silence periods
Tightly-coupled	Voice 1, Video 1 (Figure 3(a))	Voice 2, Video 1 (Figure 3(b))
Loosely-coupled	-	Voice 3, Video 2 (Figure 3(c) , Figure 3(d) , Figure 3(e))

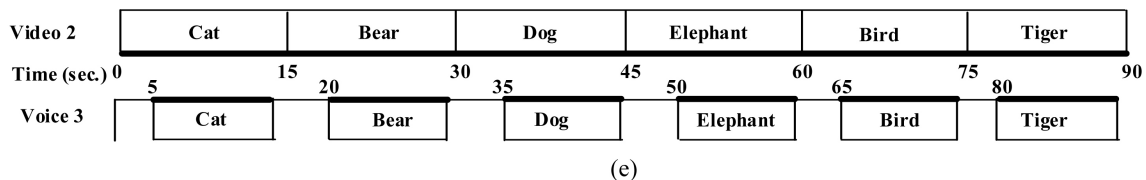
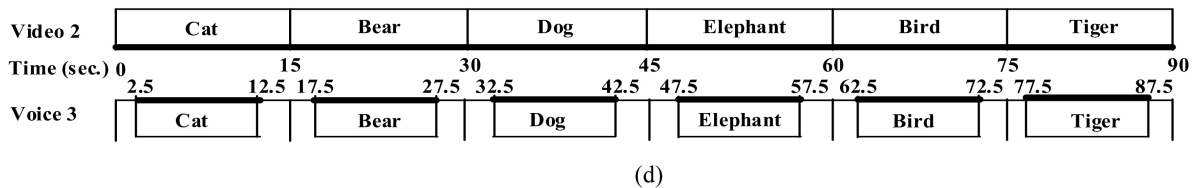
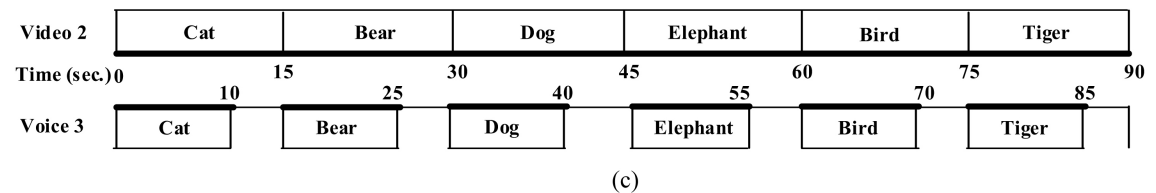
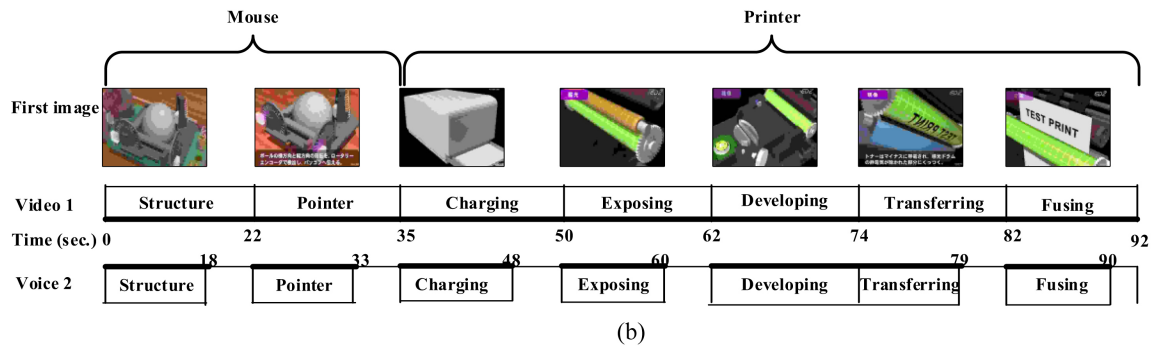
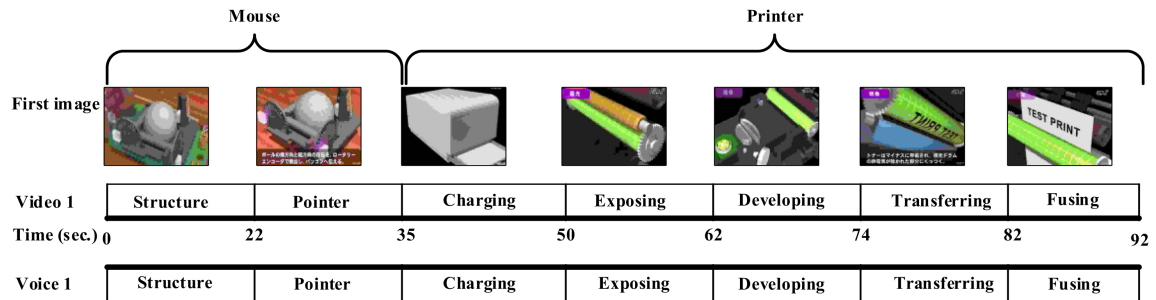


Figure 3. Temporal relations between voice and video contents. (a) Mouse and printer (Video 1 and Voice 1, No silence period); (b) Mouse and printer (Video 1 and Voice 2, Silence period); (c) Animals (Video 2 and Voice 3, Starting time: 0 sec.); (d) Animals (Video 2 and Voice 3, Starting time: 2.5 sec.); (e) Animals (Video 2 and Voice 3, Starting time: 5.0 sec.).

relation (called *tightly-coupled* or *loosely-coupled* in this paper) and silence periods (with or without silence periods). Tightly-coupled contents have tighter relations between the voice and video temporally than loosely-coupled contents, in which the voice does not have such relations. As loosely-coupled contents, we did not use contents without silence periods (see **Table 1**) because voice contents change according to the scene change in video contents; that is, the voice and video contents are tightly-coupled in this case. Note that voice and video contents in lip synchronization are much tighter related to each other compared with the tightly-coupled contents handled in this paper.

3.1. No Silence Period

We used Video 1 and Voice 1 which teach I/O devices (a mouse and a printer). The video and voice contents explain the structure and pointer of the mouse, and the charging, exposing, developing, transferring, and fusing of the printer. Their output duration is 1 minute and 32 seconds as shown in **Figure 3(a)**. The voice does not include any silence period. In **Figure 3(a)**, the video has seven scenes, and we show the first image of each video scene at the beginning of the scene. In Video 1, the output duration of each scene is not the same.

During the subjective assessment, each subject watched Video 1 and listened to Voice 1. Group synchronization errors from -550 ms to $+550$ ms at intervals of 50 ms were presented to each subject in random order. Negative values denote the voice ahead of the video, and positive values do the voice behind of the video.

The total assessment time per subject was about 30 minutes. The number of subjects was 15 (10 females and 5 males), and their ages were between 28 and 35.

3.2. Silence Periods

3.2.1. Tightly-Coupled Contents

We employed Video 1 and Voice 2 explaining the I/O devices (see **Table 2**). The explanation of Voice 2 is almost the same as Voice 1, but Voice 2 is simplified from Voice 1 by reducing the number of words to produce silence periods as shown in **Figure 3(b)**. The number of words in Voice 2 is about 110, and that in Voice 1 is around 170. Note that Voice 2 starts to explain each video scene when the scene change occurs.

Table 2. Video outputted scenes.

Video 1	I/O devices : Structures of mouse and printer
Video 2	Six animals (moving of cat, bear, dog, elephant, bird, and tiger)
	Cat: Sitting in grass [28]
	Bear (panda bear): Playing on snowy day [29]
	Dog: Running on street [30]
	Elephant: Riding on wood [31]
	Bird (parrots): Talking about something [32]
	Tiger: Playing with teddy bear [33]

During the assessment, each subject watched Video 1 and listened to Voice 2. We changed the group synchronization error from -700 ms to $+550$ ms at intervals of 50 ms.

The total assessment time per subject was about 95 minutes including break times. The number of subjects was 15 females, and their ages were between 28 and 37.

3.2.2. Loosely-Coupled Contents

We used Video 2 having six scenes of animals [28]–[33] (cat, bear, dog, elephant, bird, and tiger as shown in **Table 2**) and Voice 3 which teaches the English general vocabulary as the names of these six animals as follows: “This animal is xxx, it’s called yyy, and its spell is zzz.” In this explanation, “yyy” and “zzz” are English words, and the other parts are Myanmar words. Their output duration is 1 minute and 30 seconds (almost the same as the duration at Subsections 3.1 and 3.2.1). In Video 2, the duration of each scene is constant (*i.e.*, 15 sec.) to change the locations of silence periods easily.

During the assessment, each subject watched Video 2 and listened to Voice 3. We used the three different starting times for Voice 3 (0 sec., 2.5 sec., and 5.0 sec.) by changing the locations of silence periods as shown in **Figures 3(c)–(e)**. It should be noted that the locations of silence periods in **Figure 3(c)** are similar to those in **Figure 3(b)**. We selected the random order for the three different starting times for each subject. We changed the group synchronization error from -700 ms to -300 ms at intervals of 50 ms in random order for the starting time of 0 sec. We changed group synchronization error from -600 ms to $+600$ ms at intervals of 100 ms in random order for the starting time of 2.5 sec. Also, we changed the group synchronization error from 300 ms to 700 ms at intervals of 50 ms in random order for the starting time of 5.0 sec.

The total assessment time per subject was about 80 minutes including break times. The number of subjects was 13 females and 2 males. Their ages were between 33 and 39.

4. Assessment Results

In this section, we show assessment results for voice contents having no silence period (Voice 1) and silence periods (Voice 2). We also show the results of tightly-coupled and loosely-coupled contents for voices with silence periods (Voices 2 and 3, respectively).

4.1. No Silence Period

We plot the perception rate as a function of the group synchronization error for Voice 1 and Video 1 in **Figure 4** (the results of Voice 2 will be explained in Subsection 4.2.1). The perception rate is here defined as the ratio (expressed as a percentage) of the number of subjects who perceived group synchronization errors to the total number of subjects. Note that the group synchronization error is

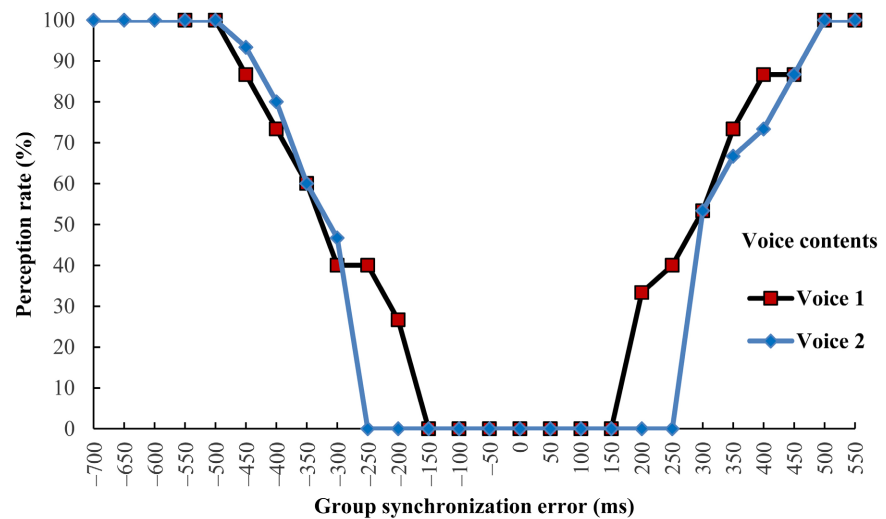


Figure 4. Perception rate versus group synchronization error for Video 1.

produced by changing the start times of voice and video at the student terminal as described in Section 3.

In **Figure 4**, we see that the perception rate is 0% when the group synchronization error is between about -150 ms and $+150$ ms (the results are almost the same as those in [34]). When the absolute error exceeds about 150 ms, the perception rate starts to increase up to 100% (at the absolute error of 500 ms). If we assume that the imperceptible range denotes a range in which the perception rate is less than or equal to 20% [34] [35], the range is between around -200 ms and $+200$ ms. If the allowable range is assumed to be a range in which the perception rate is greater than or equal to 60% [34] [35], the range is beyond the absolute error of about 300 ms. The group synchronization error in this paper is more easily perceived than the lip synchronization error [24], but it is more difficult to perceive the group synchronization error than the pointer synchronization error [24] (see the lip and pointer synchronization errors in Section 1).

4.2. Silence Periods

4.2.1. Tightly-Coupled Contents

As described earlier, the perception rate for Voice 2 and Video 1 is also shown in **Figure 4**. From the figure, we find that the perception rate for Voice 2 is 0% when the error is between about -250 ms and $+250$ ms. The imperceptible range is between around -250 ms and $+250$ ms, and the allowable range of the absolute error is larger than about 300 ms. The imperceptible range of Voice 2 is different from that of Voice 1 and the allowable ranges of Voices 1, and 2 are almost the same as each other. Therefore, we can conclude that the perception rate of group synchronization error depends on the voice contents.

4.2.2. Loosely-Coupled Contents

In **Figure 5**, we plot the perception rate versus the group synchronization error for Video 2 and Voice 3. The figure includes the results of the three different

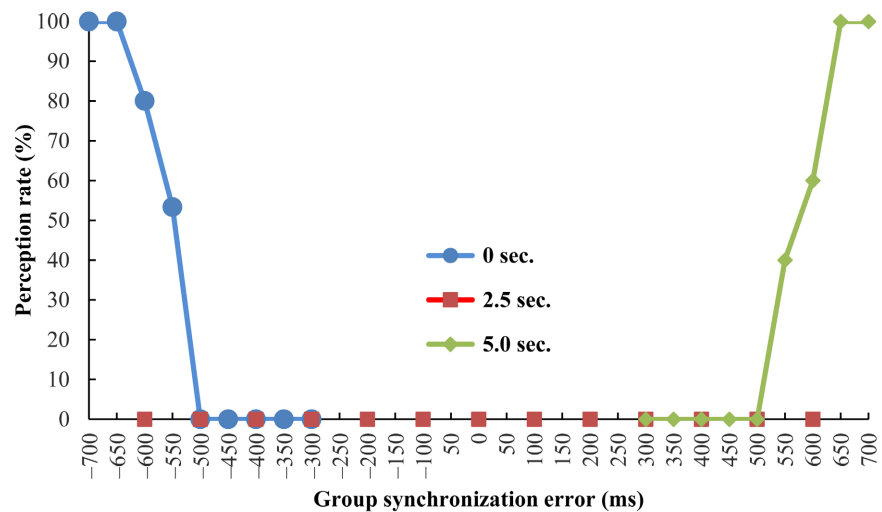


Figure 5. Perception rate versus group synchronization error for Voice 3 and Video 2.

starting times of 0 sec., 2.5 sec., and 5.0 sec. We find in the figure that when the starting time is 0 sec. and the group synchronization error is larger than about -500 ms, the perception rate is 0%. When the starting time is 2.5 sec., the perception rate is 0%; that is, no one finds the group synchronization error for the errors from around -600 ms to $+600$ ms. When the starting time of 5.0 sec. and the group synchronization error is less than about $+500$ ms, the perception rate is 0%.

From the above considerations, we can obtain the following results. Because the perception rate is 0% when the error is larger than about -500 ms and the starting time is 0 sec., and when the error is less than around $+500$ ms and the starting time is 5.0 sec., the perception rate is 0% when the error is between about -500 ms and $+500$ ms (*i.e.*, 5 sec. $+500$ ms) and the starting time is 0 sec. In the same way, the perception rate is 0% when the error is between about -3000 ms (*i.e.*, 2.5 sec. -500 ms) and 3000 ms (2.5 sec. $+500$ ms) and the starting time is 2.5 sec. Also, the perception rate is 0% when the error is between around -5500 ms (-5 sec. -500 ms) and $+500$ ms and the starting time is 5.0 sec. The ranges are much wider than those of the tightly-coupled contents. Therefore, we can obtain that the perception rate of group synchronization error depends on the starting time of voice as well as the temporal relations between voice and video contents.

From the above discussions, the human perception of group synchronization error depends on the voice and video contents.

5. Conclusions

In this paper, we investigated the dependencies of voice and video contents on human perception of group synchronization error for remote learning by carrying out QoE subjective assessment. Assessment results showed that we can more easily perceive the error for voice contents without silence period than those

with silence periods. Therefore, silence periods mitigate the perception rate of the error. We also found that we can more easily perceive the error for tightly-coupled contents than loosely-coupled ones. We further confirmed that the perception rate is dependent on the voice and video contents.

As the next step of our study, we will investigate the two-way communication case in which a teacher and multiple students can interactively discuss with each other in a lecture. In addition, we need to handle a variety of contents, because there exist dependencies of contents on QoE.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Blakowski, G. and Steinmetz, R. (1996) A Media Synchronization Survey: Reference Model, Specification, and Case Studies. *IEEE Journal on Selected Areas in Communications*, **14**, 5-35. <https://doi.org/10.1109/49.481691>
- [2] Montagud, M., Cesar, P., Boronat, F. and Jansen, J. (2018) MediaSync: Handbook on Multimedia Synchronization. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-65840-7>
- [3] Huang, Z., Nahrstedt, K. and Steinmetz, R. (2013) Evolution of Temporal Multimedia Synchronization Principles: A Historical Viewpoint. *ACM Transactions on Multimedia Computing, Communications and Applications*, **9**, Article No. 34. <https://doi.org/10.1145/2490821>
- [4] Montagud, M., Jansen, J., Cesar, P. and Boronat, F. (2015) Review of Media Sync Reference Models: Advance and Open Issues. *Proceeding of MediaSync Workshop*, Brussels, 3 June 2015, 1-8.
- [5] Rop, K.V. and Bett, N.K. (2012) Video Conferencing and Its Application in Distance Learning. *Proceedings of Annual Interdisciplinary Conference*, Vol. 1, Nairobi, June 2012, 1-9.
- [6] Costa, R.M. and Santos, C.A.S. (2014) Systematic Review of Multiple Contents Synchronization in Interactive Television Scenario. *ISRN Communications and Networking*, **2014**, Article ID: 127142. <https://doi.org/10.1155/2014/127142>
- [7] Vakiliand, A. and Gregoire, J.C. (2013) QoE Management in a Video Conferencing Application. *Computer Networks*, **57**, 1726-1738. <https://doi.org/10.1016/j.comnet.2013.03.002>
- [8] Spyridonis, F., Daylamani-Zad, D. and O'Brien, M.P. (2018) Efficient In-Game Communication in Collaborative Online Multiplayer Games. *Proceedings of the 10th International Conference on Virtual Worlds and Games for Serious Application (VS-Games)*, Würzburg, 5-7 September 2018, 1-4. <https://doi.org/10.1109/VS-Games.2018.8493420>
- [9] Ishibashi, Y. and Tasaka, S. (1995) A Synchronization Mechanism for Continuous Media in Multimedia Communications. *Proceedings of IEEE INFOCOM*, Boston, 2-6 April 1995, 1010-1019. <https://doi.org/10.1109/INFCOM.1995.515977>
- [10] Ishibashi, Y., Tsuji, A. and Tasaka, S. (1997) A Group Synchronization Mechanism for Stored Media in Multicast Communications. *Proceedings of IEEE INFOCOM*, Kobe, 7-11 April 1997, 692-700. <https://doi.org/10.1109/INFCOM.1997.644522>

- [11] Ishibashi, Y. and Tasaka, S. (1997) A Group Synchronization Mechanism for Live Media in Multicast Communications. *GLOBECOM 97. IEEE Global Telecommunications Conference. Conference Record*, Phoenix, 3-8 November 1997, 746-752. <https://doi.org/10.1109/GLOCOM.1997.638431>
- [12] Ishibashi, Y. and Tasaka, S. (1999) A Distributed Control Scheme for Group Synchronization in Multicast Communications. *Proceedings of International Symposium on Communications*, Kaohsiung, November 1999, 317-323.
- [13] IETF RFC7272 (2014) Inter-Destination Media Synchronization (IDMS) Using the RTP Control Protocol (RTCP).
- [14] ITU-T Recommendation G.1070 (2018) Opinion Model for Video-Telephony Applications. International Telecommunication Union, Geneva.
- [15] Nawaz, O., Fielder, M., Moor, K.D. and Khatibi, S. (2020) Influence of Gender and Viewing Frequency on Quality of Experience. *Proceedings of 2020 20th International Conference on Quality of Multimedia Experience (QoMEX)*, Athlone, 26-28 May 2020, 1-4. <https://doi.org/10.1109/QoMEX48832.2020.9123106>
- [16] Miyashita, Y., Ishibashi, Y., Fukushima, N., Sugawara, S. and Psannis, K.E. (2011) QoE Assessment of Group Synchronization in Networked Chorus with Voice and Video. *TENCON2011: 2011 IEEE Region 10 Conference*, Bali, 21-24 November 2011, 192-196. <https://doi.org/10.1109/TENCON.2011.6129090>
- [17] Huang, P., Ishibashi, Y., Fukushima, N. and Sugawara, S. (2012) QoE Assessment of Group Synchronization Control Scheme with Prediction in Work Using Haptic Media. *International Journal of Communications, Network and System Science (IJCNS)*, **5**, 321-331. <https://doi.org/10.4236/ijcns.2012.56042>
- [18] Ishibashi, Y., Nagasaka, M. and Fujiyoshi, N. (2006) Subjective Assessment of Fairness among Users in Multipoint Communications. *Proceedings of ACM SIGCHI Advance in Computer Entertainment Technology (ACE)*, Hollywood, 14-16 June 2006. <https://doi.org/10.1145/1178823.1178905>
- [19] Hosoya, K., Ishibashi, Y., Sugawara, S. and Psannis, K.E. (2009) Group Synchronization Control Considering Difference of Conversation Roles. *Proceedings of the 13th IEEE International Symposium on Consumer Electronics (ISCE)*, Kyoto, 25-28 May 2009, 948-952. <https://doi.org/10.1109/ISCE.2009.5156876>
- [20] Kwon, D., Kim, H. and Ju, H. (2018) Play Sharing a Group Synchronization Scheme for Media Streaming Services in Hierarchical WLANs. *International Journal of Network Management*, **28**, Article ID: e2024. <https://doi.org/10.1002/nem.2024>
- [21] Ida, Y., Ishibashi, Y., Fukushima, N. and Sugawara, S. (2010) QoE Assessment of Interactivity and Fairness in First Person Shooting with Group Synchronization Control. *Proceedings the 9th Annual Workshop on Network and Systems Support for Games (NetGames)*, Taipei, 16-17 November 2010, 1-5. <https://doi.org/10.1109/NETGAMES.2010.5680283>
- [22] Huang, P., Ishibashi, Y. and Sithu, M. (2016) Enhancement of Simultaneous Output-Timing Control with Human Perception of Synchronization Errors among Multiple Destinations. *Proceedings of the 2nd IEEE International Conference on Computer and Communications (ICCC)*, Chengdu, 14-17 October 2016, 2099-2103. <https://doi.org/10.1109/CompComm.2016.7925070>
- [23] Ishibashi, Y., Kanbara, T. and Tasaka, S. (2004) Inter-Stream Synchronization between Haptic Media and Voice in Collaborative Virtual Environments. *Proceedings of 12th Annual ACM International Conference on Multimedia*, New York, 10-16 October 2004, 604-611. <https://doi.org/10.1145/1027527.1027670>
- [24] Steinmetz, R. (1996) Human Perception of Jitter and Media Synchronization. *IEEE*

- Journal on Selected Areas in Communications*, **14**, 61-72.
<https://doi.org/10.1109/49.481694>
- [25] Staelens, N., De Meulenaere, J., Glenn, L.B., Wallendaal, V., Cock, J.D., Geeraert, K., Vercammen, N., Broeck, W.V., Vermeulen, B., Van de Walle, R. and Demeester, P. (2012) Assessing the Importance of Audio/Video Synchronization for Simultaneous Translation of Video Sequences. *Multimedia Systems*, **18**, 445-457.
<https://doi.org/10.1007/s00530-012-0262-4>
 - [26] Younkin, A.C. and Corriveau, P.J. (2008) Determining the Amount of Audio-Video Synchronization Errors Perceptible to the Average End-User. *IEEE Transactions on Broadcasting*, **54**, 623-627. <https://doi.org/10.1109/TBC.2008.2002102>
 - [27] ITU-R Recommendation BT.500. (2009) Methodology for the Subjective Assessment of the Quality of Television Pictures. International Telecommunication Union, Geneva.
 - [28] Cute Pets (2021, August 3) Cat in the Grass in the Sun #2.
<https://www.youtube.com/watch?v=zl8ZVbFX3RQ>
 - [29] iPanda (2019, February 21) Giant Pandas in the Snow Feb. 20, 2019.
<https://www.youtube.com/watch?v=pNff91GIUg>
 - [30] NHD-TV (2016, March 22) Cutes Dogs | Cutest Dog in the World | Cute Dogs Clips 2016. <https://www.youtube.com/watch?v=dQ5BAupEKvw>
 - [31] Funnyplox (2016, December 17) Most Funny and Cute Baby Elephant Videos Compilation. <https://www.youtube.com/watch?v=SNggmeilXDO>
 - [32] The Pet Collective (2016, July 27) Sh*t Birds Say Video Compilation 2016.
<https://www.youtube.com/watch?v=Wjvov2fo9vo>
 - [33] Kennys Wild Things (2014, December 5) Cute Tiger Cub Exploring, Playing & Talking. <https://www.youtube.com/watch?v=EuvHmfQMD30>
 - [34] Mo Mo Lwin, H.M., Ishibashi, Y. and Mya, K.T. (2019) Human Perception of Group Synchronization Error for Remote Learning: One-Way Communication Case. *Proceedings of IEEE International Conference on Consumer Electronics—Taiwan (ICCE-TW)*, Yilan, 20-22 May 2019, 1-2.
<https://doi.org/10.1109/ICCE-TW46550.2019.8991984>
 - [35] Mo Mo Lwin, H.M., Ishibashi, Y. and Mya, K.T. (2020) Influence of Voice Delay on Human Perception of Group Synchronization Error for Remote Learning: One-way Communication Case. *Proceedings of 2020 IEEE Conference on Computer Applications (ICCA)*, Yangon, 27-28 February 2020, 1-5.
<https://doi.org/10.1109/ICCA49400.2020.9022835>