International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2022)
R. Kondo and T. Teo (Editors)

Poster

Sign Language Learning System with Concurrent Shared Avatar Hand in a Virtual Environment: Psychological Evaluation

Y. Katsumata[1], H. Ishimoto[1], Y. Inoue[1], and M. Kitazaki[1]

[1]Department of Computer Science and Engineering, Toyohashi University of Technology, Aichi, Japan

Abstract

We aimed to develop a sign language learning system using virtual reality to improve learning motivation. Hand movements for twenty words consisted of three letters were recorded with a hand motion capture system (model hand). In the learning system, the participant was asked to mimic the model hand movement while looking at both the model hand and the "own hand" in a head mounted display (HMD) with the hand motion capture. The "own hand" avatar was either of the real own hand or the shared hand motion, which was made by averaging the participant's hand and the model hand movements. The model hand was presented either in the opposite or same direction as the participant. Participants rated the usability of the system in 2 x 2 (own/shared hand x opposite/same direction) experimental blocked design. We found that the shared hand avatar and the same direction presentation were better than the own hand and the opposite direction presentation, respectively. Thus, the proposed shared hand avatar system with the HMD and hand motion capture could improve sign language learning.

CCS Concepts

• ***Human-centered computing*** → ***Virtual reality;*** • ***Applied computing*** → *Computer-assisted instruction;*

1. Introduction

Sign languages are visual languages with manual articulation. American sign language is used in United States, and learned by many students in high school, colleges, and universities [QP11]. Video materials are often used in addition to textbooks and in-person lectures. They enable students to learn without teachers, and with changing speed. A sign language learning system based on two-dimensional image sampling with convolutional neural network has been developed [JKL17]. This learning system can be implemented with low-cost cameras like as a web camera. A game-based learning system has been developed for American sign language, and it is shown that the learning system is better than the traditional face-to-face learning method [KHKW17]. The game-based learning system captures student's hand movements as three-dimensional data by Kinect motion capture, compares them with teachers' stored data, and feedbacks similarity scores [KHKW17]. Thus, the students can learn sign language by alone with appropriate feedback.

Recently, virtual co-embodiment has been proposed [FOH*21] [HGS*20] [HK22]. In the virtual co-embodiment, two or more people use an avatar in the first-person perspective in a virtual environment. Fribourg et al. (2020) showed that the sense of agency is higher than the actual controlling ratio of individuals who control co-embodiment avatar if the goal is common. Hagiwara et al (2020) showed that the shared avatar of which movements are averaged by two individuals' movements performs straighter and smoother (less jerk) movement than the solo avatar [HGS*20]. The co-embodiment or shared avatar can be applied for a learning system if an expert and a novice co-embody an avatar. The purpose of the present study is to develop a sign language learning system with the shared hand avatar, of which movements are the average movements of the student and the teacher. We hypothesized that the shared hand system could improve usability and learning because the participants could mimic model/teacher's movements intuitively with embodiment.

2. Preparing the model data

Thirty nonsense syllables/words consisted of three letters [Hul33] were prepared for the stimuli (for example, YOS, WEF, HUJ). We used these stimuli to avoid the effect of knowledge. We captured 3-D hand movements of 60 words from a person with American sign language skill using a motion capture system (Manus Prime II Haptic) with a tracker (HTC Vive Tracker 3.0). An expert evaluated the captured hand motions of sign languages, and we chose best 20 words for the experiments based on the evaluation.

3. Methods of Evaluation

Twenty-six healthy adults (three females, mean 21.69, SD 1.44 years old, all right-handed) participated in the experiment. They provided written informed consent before the experiment. The

DOI: 10.2312/egve.20221300

methods of the experiment were approved by the Ethical Committee at Toyohashi University of Technology, and all methods were carried out in accordance with the relevant guidelines and regulations.

Participants wore a head-mounted display (HMD, HTC Vive Pro EYE) and a hand motion capture system (Manus Prime II Haptic) with a tracker (HTC Vive Tracker 3.0) on the right hand. The experiment was controlled by a computer.

The "own hand" avatar and the model/teacher hand avatar were presented side-by-side. The participants were asked to mimic the model hand motion as much as possible. The model hand was either opposite (face-to-face; Figure 1 Left) or same direction (Figure 1 Center) as the participant. The "own hand" was either of the real own hand motion or the shared hand avatar (Figure 1 Right), which was made by averaging the participant's hand movement and the model hand movement. Thus, the conditions were 2 x 2 (own/shared hand x opposite/same direction). Blocked design was used; Five randomly chosen words were used for each block in one of 4 conditions. In the end of each block/condition, the participants rated the usability of the system in 7 levels (How easy is this learning system to use?: 1: very difficult, 7: very easy). All conditions were repeated 5 times (sessions) in random order.

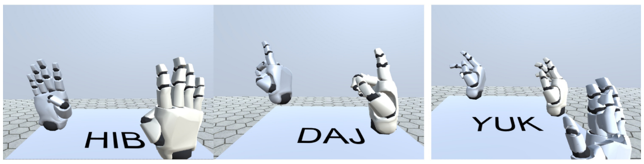

Figure 1: *Avatar-hand conditions (Left: opposite, Center: same directions; Right: An example of the shared hand. The right hand represents an actual posture of the participant (not shown in the experiment), and the middle hand is the shared avatar hand that is average of the left model hand and the right participant hand.).*

4. Results

We performed two-way repeated-measures ANOVAs (own/shared hand x opposite/same direction) with rated scores, and found main effects of own/shared hand avatar (F(1,25)=30.21, p<.001, ηp2=.547) and hand directions (F(1,25)=10.44, p=.003, ηp2=.295). Thus, the shared hand avatar and the same direction presentation were significantly better than the own hand and the opposite direction presentation, respectively (Figure 2 left). The interaction was not significant.

To investigate learning effects, we performed three-way measures ANOVAs (sessions x own/shared hand x opposite/same direction) with rated scores, and found a main effect of sessions (F(4,100)=5.04, p<.001, ηp2=.168), interactions of session x shared/own hand avatar (F(4,100)=4.07, p=.004, ηp2=.140) and sessions x direction (F(4,100)=3.20, p=.016, ηp2=.114), in addition to the main effects of own/shared hand avatar and hand directions. Thus, the usability of the shared hand avatar and the same direction presentation quickly increased more than the own hand and the opposite direction presentation, respectively (Figure 2 right).

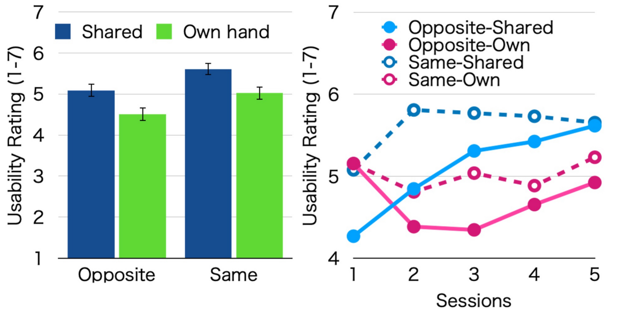

Figure 2: *Results of avatar type and direction (left), and effect of learning (right).*

5. Discussion

We demonstrated the virtual shared hand system for sign language learning. The psychological evaluation showed that the usability was better with the shared hand avatar and the same direction presentation than the own hand and the opposite direction presentation. Thus, the shared hand system may contribute to improving the sign language learning system by co-embodiment of hand motion. The same direction presentation could enable users to mimic the model hand motion without mirror reflection, so that it would be felt easy.

However, our data in the present study depends on only subjective scores. We should investigate behavioral efficiency of learning such as accuracy in the future study.

Acknowledgements

This research was supported by JST ERATO Grant Number JPMJER1701, and JSPS KAKENHI JP20H04489.

References

[FOH*21] FRIBOURG R., OGAWA N., HOYET L., ARGELAGUET F., NARUMI T., HIROSE M., LÉCUYER A.: Virtual co-embodiment: Evaluation of the sense of agency while sharing the control of a virtual body among two individuals. *IEEE Transactions on Visualization and Computer Graphics 27*, 10 (2021), 4023–4038. 1

[HGS*20] HAGIWARA T., GANESH G., SUGIMOTO M., INAMI M., KITAZAKI M.: Individuals prioritize the reach straightness and hand jerk of a shared avatar over their own. *iScience* (Nov. 2020), 101732. 1

[HK22] HAPUARACHCHI H., KITAZAKI M.: Knowing the intention behind limb movements of a partner increases embodiment towards the limb of joint avatar. *Sci. Rep. 12*, 1 (July 2022), 11453. 1

[Hul33] HULL C. L.: The meaningfulness of 320 selected nonsense syllables. *The American Journal of Psychology 45*, 4 (1933), 730–734. 1

[JKL17] JI Y., KIM S., LEE K.-B.: Sign language learning system with image sampling and convolutional neural network. In *2017 First IEEE International Conference on Robotic Computing (IRC)* (Apr. 2017), IEEE. 1

[KHKW17] KAMNARDSIRI T., HONGSIT L.-O., KHUWUTHYAKORN P., WONGTA N.: The effectiveness of the game-based learning system for the improvement of american sign language using kinect. *EJEL 15*, 4 (Aug. 2017), pp283-296. 1

[QP11] QUINTO-POZOS D.: Teaching american sign language to hearing adult learners. *Annu. Rev. Appl. Linguist. 31* (Mar. 2011), 137–158. 1