Holographic Sign Language Interpreters

Fu-Chia Yang
yang1684@purdue.edu
Purdue University
West Lafayette, Indiana, USA

Christos Mousas
cmousas@purdue.edu
Purdue University
West Lafayette, Indiana, USA

Nicoletta Adamo
nadamovi@purdue.edu
Purdue University
West Lafayette, Indiana, USA

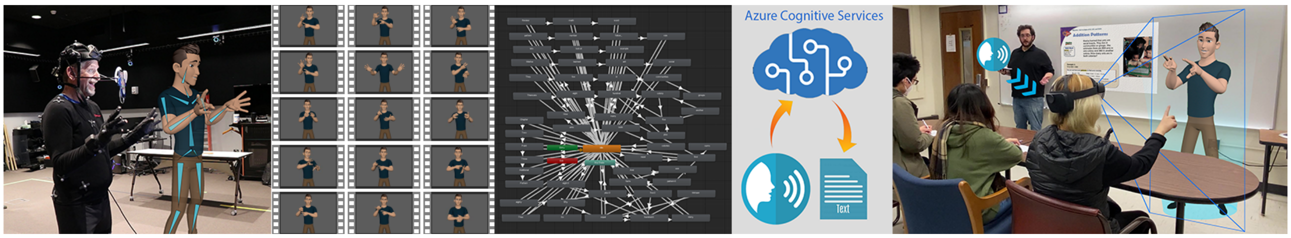

Figure 1: The pipeline of our implementation. From left to right: Motion Capture session; Sign animation dataset; Unity animation controller; Azure Speech-to-Text; and MR classroom.

ABSTRACT

We describe the implementation of a prototype system of 3D holographic sign language interpreters. The signing avatars, observed through wearable Mixed Reality (MR) smartglasses (e.g., Microsoft HoloLens), translate speech to Signed Exact English (SEE) in real-time. Such a system can be used by deaf and hard of hearing students in the classroom or other contexts to remove current accessibility barriers.

CCS CONCEPTS

• **Applied computing → Computer-assisted instruction**.

KEYWORDS

mixed reality, signing avatar, sign language, deaf and hard of hearing education, assistive technology

ACM Reference Format:

Fu-Chia Yang, Christos Mousas, and Nicoletta Adamo. 2022. Holographic Sign Language Interpreters. In *SIGGRAPH '22 Educator's Forum.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3532724.3535593

1 INTRODUCTION

Human sign language interpreters can be costly and not always available. For this reason, we can consider the use of computer assistive technology, and more specifically computer-generated virtual characters interpreters, as a promising alternative. Such virtual characters provide three main benefits compared to human interpreters: scalability, cost-effectiveness, and character control.

Sign language discourse can be decomposed into signs and sign-to-sign transitions; signs can be decomposed into manual and non-manual components and component-to-component transitions. Seamless transitions can be computed automatically by animation algorithms, which makes sign components and signs powerful building blocks that can be linked to form new signs and new discourse. Once the domain-specific sign language lexicon, such as mathematics, has been encoded into an animation database, a virtually unlimited set of virtual sign language interpreters of math lessons can be generated, and scalability to additional domains becomes progressively easier and cost-effective. Signing avatars can be designed to represent different human ages, gender, ethnicities as well as different levels of stylization. Hence the possibility of creating specialized characters for different types of learners while using the same system and the same database of signs. A substantial body of research supports the value of signing avatars for deaf education [De Martino et al. 2017; Quandt et al. 2022].

The integration of technology in the learning processes of deaf and hard of hearing (DHH) students has the potential to improve their abilities in different domains [Eden 2020]. However, the specific benefits of mixed reality (MR) technology for deaf education are still unclear, and further research on its real-world use and utility is needed [Constantinou et al. 2020]. A few examples of MR for DHH education can be found in the literature. Miller et al. [Miller et al. 2017] examined whether MR smartglasses would reduce visual dispersion and improve lecture comprehension for DHH students. While results showed no comprehension gains with the interpreter in the same visual field, participants commented that the lectures were easier to follow with the binocular presentation. Glass Vision 3D is a Google Glass application that allows DHH children to look at the QR code of an object in the classroom and view an augmented reality (AR) projection displaying a related sign language video. A preliminary study showed all participants' engagement and enthusiasm, a steep learning curve, and hardware issues [Parton 2017]. HoloHear, a prototype developed at the Microsoft Holographic Hackathon in 2016, recognizes speech and outputs sign language translation on a holographic animated mannequin [Dachis 2016]. In the proof-of-concept AR system developed by Adamo and Anasingaraju [Adamo-Villani and Anasingaraju 2017] a stylized 3D animated avatar translates from spoken to Signed English (SE) using a limited set of signs. The results of a preliminary

SIGGRAPH '22 Educator's Forum, August 08–11, 2022, Vancouver, BC

ACM ISBN 978-1-4503-9366-9/22/08.
https://doi.org/10.1145/3532724.3535593

study showed the system was acceptable, and children enjoyed the virtual signer.

Although we build on prior work, our project is highly innovative. We use the latest advancement in holographic wearable MR and life-like signing avatar technology to further investigate the potential of MR for deaf education. We consider the strong need for solutions that allow deaf students to communicate and interact with other people in an environment free of prejudice, stigma, and other obstacles. Thus, this work contributes toward filling this need.

2 IMPLEMENTATION DETAILS

The implementation of the holographic signing avatar system was twofold. First, we collected sign animation sequences through motion capture. Second, we developed a mixed reality system for Microsoft HoloLens that detects audio input and displays the corresponding sign animation. We provide implementation details of the system and how it works in the below sub-sections. We illustrate our implemented pipeline in Figure 1.

2.1 Motion Capture Animation Dataset

We prepared four short K-1 math lesson scripts and recruited a male signer to wear the motion capture (MoCap) suit and record sign animation for the math lessons. The signer was familiar with American Sign Language (ASL) and Signed Exact English (SEE). The MoCap system captured the signer's body, hands, and face through three platforms, OptiTrack, StretchSense, and FaceWare, respectively. We collected sign animation from four lessons in ASL (49 sentences) and one lesson in SEE (13 sentences).The data from the three tracking systems were transferred in real-time on a rigged character into Autodesk's MotionBuilder. The collected raw motion sequences contained mesh intersections and noise. Thus, the MoCap data were post-edited and cleaned in Autodesk's Maya to prepare them for the animation controller. The rationale behind generating sign animations through the MoCap system is to ensure that the animations displayed in our system are as natural and legible as possible.

The current prototype is built for SEE instead of ASL because SEE follows English sentence structure, whereas ASL has its own language structure and grammar, making it more difficult to process from text to sign. We manually segment every SEE animation clip into individual sign clips; each clip represents a word in the sentence. These sign segments are then saved into our sign animation dataset and mapped to a Unity animation controller, ready to be triggered through the system.

2.2 The System

The system takes input speech (audio) from the instructor and converts it to English text. The converted sentences are then analyzed to identify the corresponding signs, prosodic markers and prosodic modifiers. The system then triggers the sign animation segments in the dataset and renders a life-like holographic sign language interpreter who signs in SEE (see Figure 1 and the accompanying video).

The holographic sign interpreter application was built in Unity game engine version 2020.3.20f1, utilizing the Azure Speech-to-Text SDK version 1.20.0 and Microsoft Mixed Reality Toolkit (MRTK) version 2.7.3 packages. We used Unity's Mechanim animation controller to map out the logic behind switching sign animation segments. We blend the animation segments during the transition period, so the overall signing motions look realistic and smooth. We used MRTK to provide the user with the ability to control the signing avatar's position, orientation, and scaling.

3 CONCLUSION AND FUTURE WORK

The overall goal of our research is to improve DHH college students' accessibility to educational materials and DHH children's learning of math concepts through the application of holographic wearable MR. The specific objective of the work reported in the paper was to develop a system of 3D holographic sign language interpreters. The signing avatars, displayed as 3D holograms in a mixed reality environment, can be used by deaf learners in the classroom, at home, and while interacting with digital educational materials.

Currently, we are conducting a formative study with 12 ASL users to assess the quality of the MoCap animations with the goal of improving their accuracy and legibility in future iterations of the system. We are also conducting a usability study of the MR system with a group of DHH college students. In future work we plan to evaluate the usability, acceptability and pedagogical efficacy of the system with DHH children, their teachers and parents at the Indiana School for the Deaf (ISD).

While our initial objective is to target the educational audience, we believe that facilitating creation of sign language translation is important beyond the education domain. In the future, the holographic 3D signers could be used in many other contexts, e.g. day to day communications, social networking, entertainment to remove current accessibility barriers.

REFERENCES

N. Adamo-Villani and S. Anasingaraju. 2017. Holographic signing avatars for deaf education. In *E-learning, E-education, and online training*. 54–61.

V. Constantinou, A. Ioannou, I. Klironomos, M. Antona, and C. Stephanidis. 2020. Technology support for the inclusion of deaf students in mainstream schools: a summary of research from 2007 to 2017. *Universal Access in the Information Society* 19, 1 (2020), 195–200.

A. Dachis. 2016. HoloHear is like Google translate for deaf people.

J. M. De Martino, I. R. Silva, C. Z. Bolognini, P. D. P. Costa, K. M. O. Kumada, L. C. Coradine, P. H. da S. Brito, W. M. do Amaral, Â. B. Benetti, E. T. Poeta, L. M. G. Angare, C. M. Ferreira, and D. F. De Conti. 2017. Signing avatars: making education more inclusive. *Universal Access in the Information Society* 16, 3 (2017), 793–808.

S. Eden. 2020. Technology makes things possible: Improving the abilities of deaf and hard-of-hearing children with advanced technologies. *The Oxford Handbook of deaf studies in learning and cognition* (2020), 407–425.

A. Miller, J. Malasig, B. Castro, V. L. Hanson, H. Nicolau, and A. Brandão. 2017. The use of smart glasses for lecture comprehension by deaf and hard of hearing students. In *ACM CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 1909–1915.

B. S. Parton. 2017. Glass vision 3D: digital discovery for the deaf. *TechTrends* 61, 2 (2017), 141–146.

Lorna C. Quandt, Athena Willis, Melody Schwenk, Kaitlyn Weeks, and Ruthie Ferster. 2022. Attitudes Toward Signing Avatars Vary Depending on Hearing Status, Age of Signed Language Acquisition, and Avatar Type. *Frontiers in Psychology* 13 (2022).