Recognizing highly variable American Sign Language in virtual reality

-Md Shahinur Alam (Educational Neuroscience & VL2 Center, Gallaudet University, Washington, DC)
-Myles De Bastion (CymaSpace, Portland, Oregon)
-Melissa Malzkuhn (VL2 Center, Gallaudet University, Washington, DC)
-Lorna C. Quandt (Educational Neuroscience & VL2 Center, Gallaudet University, Washington, DC)

Paper ID# 6971

Signing Avatars& Immersive Learning

GU

Introduction

- Over 5% (430 million) of the world's population has some form of hearing loss, which is projected to increase to 2.5 billion by 2050[1].
- American Sign Language (ASL) recognition in 2D/3D is a relatively mature research area than virtual reality (VR).
- As immersive technology grows, ASL interactions in VR are more relevant and timely.

1. https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

Objectives

- The objectives of this research are to teach ASL in VR and make the learning process fun and entertaining [2,3].
- State-of-the-art ASL recognition research mostly focused on 2D or RGB-D-based cameras where users cannot feel the real-world 3D experience. Our focus is to develop more interactive learning environment.
- Provide real-time feedback to the users.

2. Quandt, L. C., Lamberton, J., Leannah, C., Willis, A., & Malzkuhn, M. (2022). Signing avatars in a new dimension: Challenges and opportunities in virtual reality. In *Proceedings of the 7th International Workshop on Sign Language Translation and Avatar Technology (SLTAT)*

3. Quandt, L. C., Lamberton, J., Willis, A. S., Wang, J., Weeks, K., Kubicek, E., & Malzkuhn, M. (2020). Teaching ASL signs using signing avatars and immersive learning in virtual reality. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20), October 26–28, Virtual Event, Greece*.

Concept

- Interactive learning system in VR
- Student
  - User of the system (ASL learner).
  - Will get feedback from the system about sign accuracy.
- Teacher
  - Animated avatar who teaches ASL.
  - Mocap data from a native signer.
  - “Decides” whether the student's sign is correct or not.

Student

Teacher

Signing Avatars& Immersive Learning

GU

Signing avatars in a 3D environment

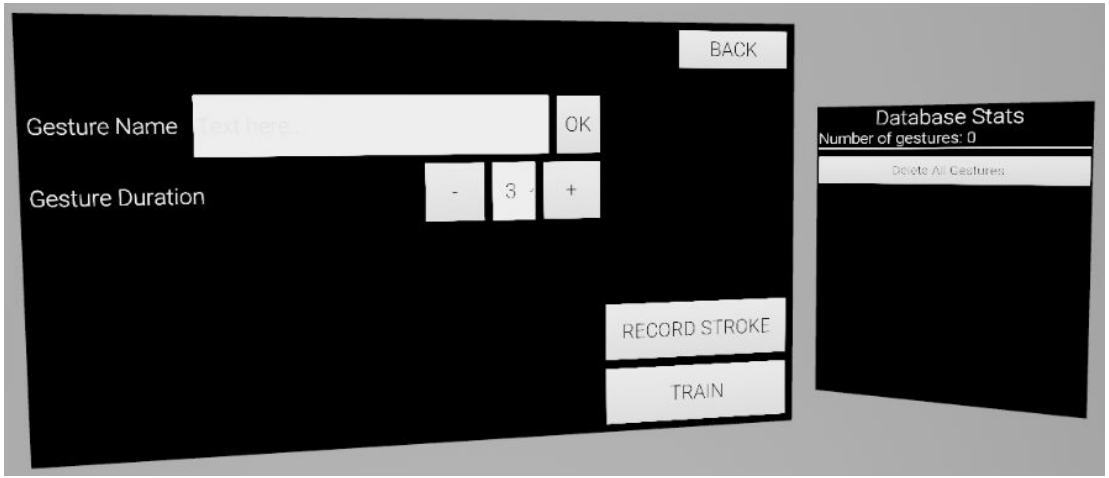
- We designed a 3D interactive coffee shop environment
- The avatar acts as the Teacher and shows ASL signs
- A motion capture system (Vicon) was used to animate the 3D avatar
- We have nine different coffee shop sign in our sign vocabulary

Signing Avatars& Immersive Learning

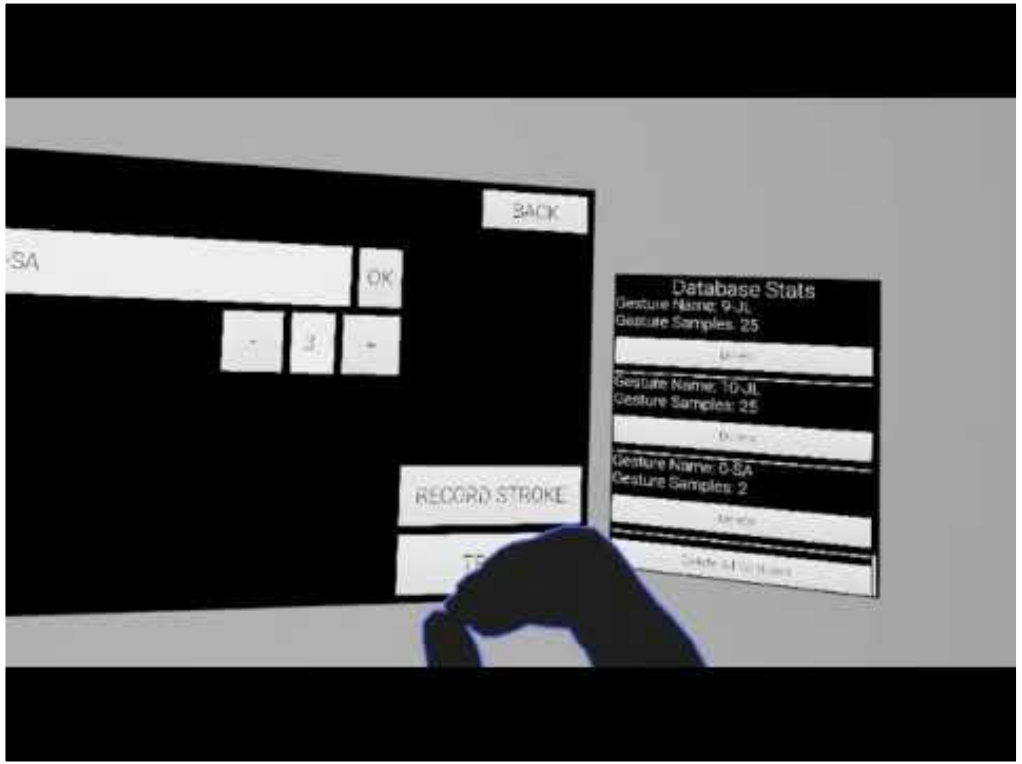
GU

Data Collection

- Data plays an important role in any machine learning and deep learning research.
- For the training process, we gathered ASL sign data from native signers.
- A user interface (UI) was designed where users could interact with some basic buttons and record ASL signs without much external intervention.
- User can see the existing and new sign in the right window (see next slide).



Data Collection - Video



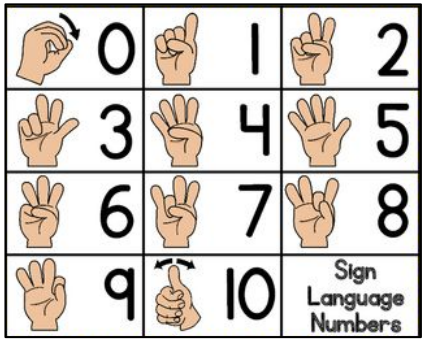
Signing Avatars& Immersive Learning

GU

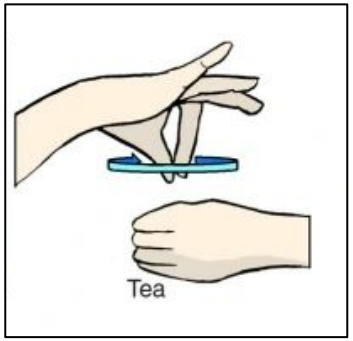
Datasets

- We trained two different models with different dataset-
  - ASL numbers 0-10
  - TEA sign

ASL Number sign



ASL TEA sign



Number Signs 0-10

- Total signs: 2500
- Participants: 10
  - Men: 4
  - Women: 6
  - Age range: 22-46
- Hearing status: D, H, HH

TEA Sign

- Total signs: 500
- Participants: 10
  - Men: 4
  - Women: 6
  - Age range: 22-46
- Hearing status: D, H, HH

Signing Avatars& Immersive Learning

GU

Experimental Setup

- VR Device: Oculus Quest 2 (software version 44.0.0.169.455).
- Environment: Unreal Engine v4.27
- AI Plugin: MiVRy v2.5
- PC:
  - Windows 11 pro 64 bit
  - Memory: 32GB
  - Processor: Core i9 3.50Ghz

Methods

- Sign recognition is performed by an AI model.
- MiVRy Unreal Engine plugin is used for sign recognition.
- The AI model is generated based on training data.
- The model provides similarity values in real time and this is the backbone of our feedback system.

Signing Avatars& Immersive Learning

GU

Results

- Each number was signed 10 times.
- Average accuracy was 46%.
- Best and worst result found for numbers 10 and 6, respectively.

*The recognition accuracy of the sign TEA is around 55%.

The model guesses (recognizes) this:

The signer signs this:

	0	1	2	3	4	5	6	7	8	9	10
0	4	1	0	0	0	0	0	0	0	0	3
1	1	7	0	0	0	0	0	0	0	0	2
2	0	3	3	0	0	0	0	0	0	0	4
3	0	0	4	3	0	0	0	0	0	0	3
4	0	0	0	0	4	3	0	0	0	0	3
5	0	0	0	0	3	6	0	0	0	0	1
6	0	0	0	7	0	0	2	0	0	0	1
7	1	0	5	0	0	0	0	3	0	0	1
8	0	4	0	0	0	0	0	0	3	0	2
9	4	1	0	0	0	1	0	0	0	3	1
10	2	0	0	0	0	0	0	0	0	0	8

Signing Avatars& Immersive Learning

GU

Conclusion

The ten signers in this initial study were *diverse* in age, sex, ASL proficiency, and hearing status, with most being deaf lifelong ASL users.

Next steps:

- Improve our recognition accuracy by modifying the AI model
- Add more content / signs
- Continue building the 3D environments
- Continue adding rich facial expressions to the avatars

Signing Avatars& Immersive Learning

GU

Thank you! Questions?

Action & Brain Lab

Motion Light Lab

Athena Willis
Kaitlyn Weeks
Carly Leannah
Joseph Palagano
Taylor Delorme
Jason Lamberton
Yiqiao Wang
Jianye Wang
Sarah Miller
Heather Smith

GALLAUDET UNIVERSITY

CREST

NSF

VISUAL LANGUAGE AND VISUAL LEARNING
A NSF Funded Science of Learning Center

www.motionlightlab.com
www.vl2.gallaudet.edu