CREATING A PROGRAM RUNNING ON A VR HEADSET TO DETECT AND TRANSLATE SIGN LANGUAGE

100631213

Author: Miss Sky Stone, Email: Sky@skui.co.uk

Supervisor: Dr Patrick Merritt, Email: P.Merritt@derby.ac.uk

Introduction and Lit Review

While their are previous examples of sign recognition technology, there is a lack of examples that are built on consumer hardware, This projects aim is to fill that gap with a piece of software that can both run on a consumer grade headset while achieving equal or greater accuracy and latency.

In a project carried out by Zhihao Zhou (Zhou et al., 2020), they were able to create a system using gloves that contained stretchable sensor arrays and an AI model to achieve an accuracy of greater than 90% and latency of that isn't too imposing on the user. However their use of specialised gloves increases the cost to use this software making it less accessable.

In a study by A. Vaitkevičius (Vaitkevičius et al., 2019) they made use of a Leap motion tracker which generates a 3D model of the users hands. This model is sent to their cloud service where a machine learning model is use to decode sign. Their model managed to achieve an accuracy of 86.1%, allowing this model to be helpful for its intended users. However like the project above it makes use of specialized hardware making it less accessable to everyday users and its dependancy on a cloud service limits the software to use in an area with stable internet.

Aim

To create and test the accuracy of a Quest 3 Application in its ability capture and decode BSL to assist a hearing-impaired person with communicating without the use of a machine learning algorithm.

Objectives

· Make use of the Quest 3's internal tracking combine with the meta XR pose detection to detect a small group of simple hand poses.
· Create a Script that can use both the pose data and hand position to detect simple signs.
· Make use of Skeletal data to more accurately detect the position of each finger
· Create a system that can combine both the Skeletal data and the simple

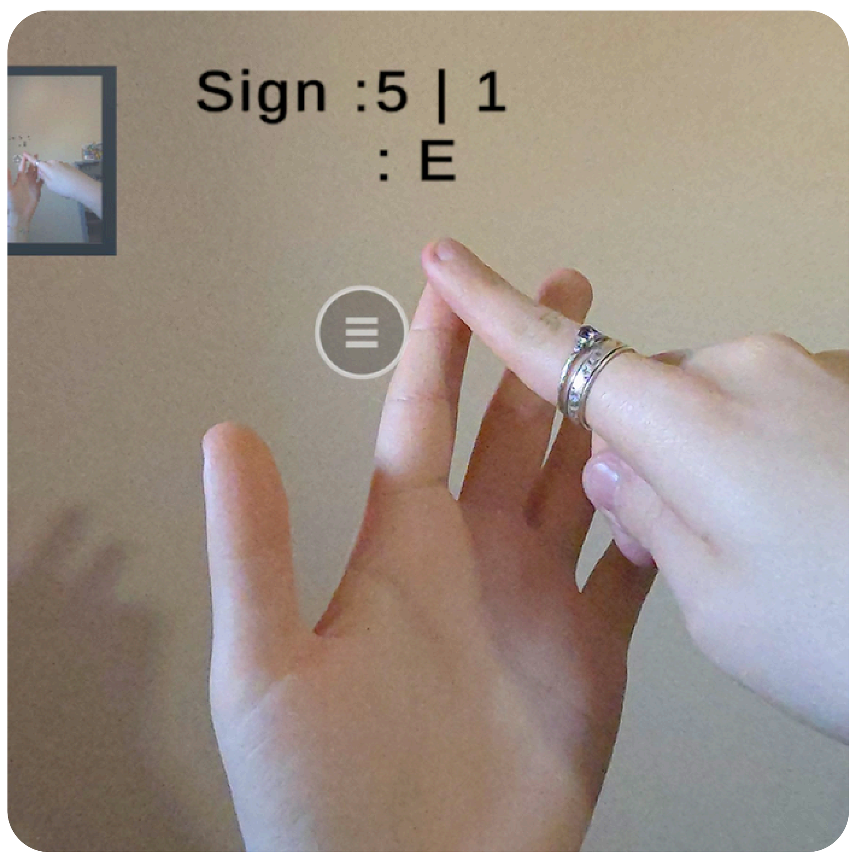
Methodology

The study will follow the Quantitative research method when looking at past papers to compare the accuracy of their algorithms and methods while bearing in mind the cost and accessibility of the hardware that is used. Whereas when looking at the hardware for this project, I will use Qualitative research to find the most effective and efficient ways to get the required output that, while it might not be the most accepted way, will hopefully help increase the efficiency of the program.

The final stage of the project will include the comparison of the accuracy of the product on average with the BSL fingerspelling alphabet against the accuracy of other models that have already been produced. Combined with the cost/accessibility of the hardware used in each method, this will assess whether the project:

- Is successfully able to detect and translate BSL with an appropriate level of accuracy.
- Uses hardware that is accessible to the target audience of this product.

To carry out the final test, I will use the results posted in the papers covering the already existing models. On average, these include an accuracy measurement in a percentage as well as the coverage of signs that the model can interpret. I plan to replicate this test data by signing each sign 3 times and repeating that 5 times. This should provide a good average on the accuracy of each sign in the model. I will then find the average of all signs, which I can then compare with the other models.



Prototype

The current prototype utilizes the Meta Interaction SDK and the Unity game engine to recognize and interpret signs by capturing pose data and hand locations. This system runs on a Quest 3 headset, enabling real-time sign decoding and interaction.

References

- Zhou, Z. et al. (2020) 'Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays'
- Vaitkevičius, A. et al. (2019) 'Recognition of American Sign Language Gestures in a Virtual Reality Using Leap Motion'