

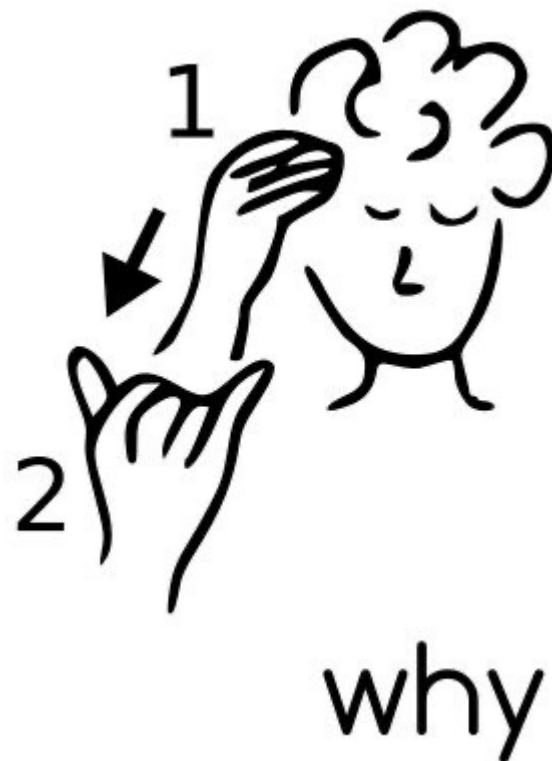
SignQuery: A Natural User Interface and Search Engine for Sign Languages with Wearable Sensors

Hao Zhou, Taiting Lu, Kristina McKinnie, Joseph Palagano
Kenneth DeHaan, Mahanth Gowda

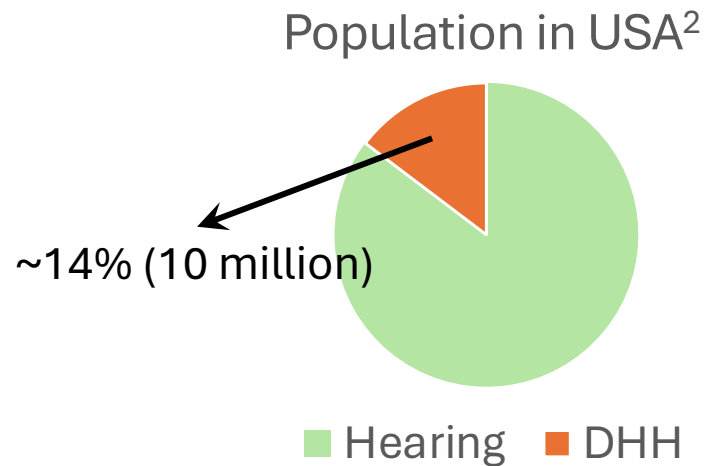
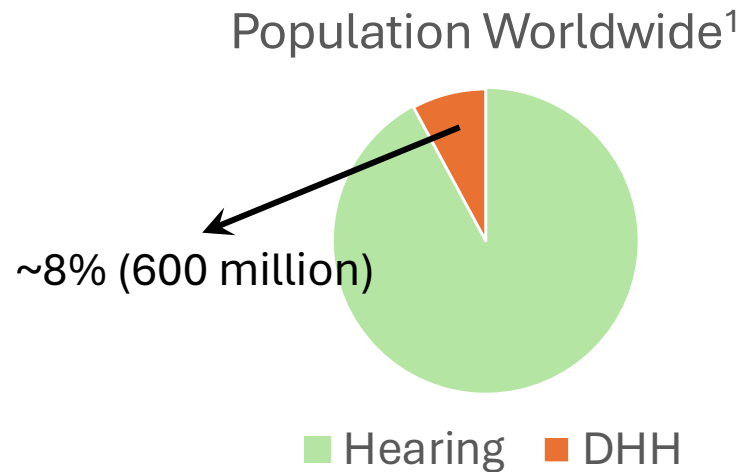


The 29th International Conference on Mobile Computing and Networking (MobiCom '23)

Why do we want to
build a search engine with
sign languages?



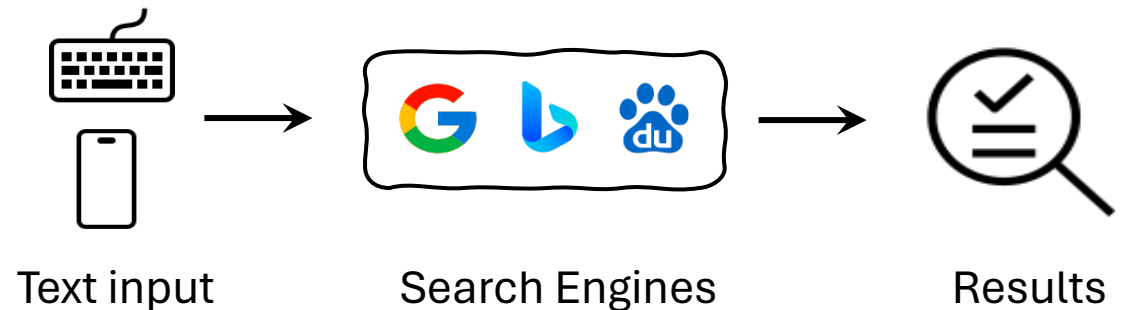
The Deaf and Hard of Hearing community is a part of the world.



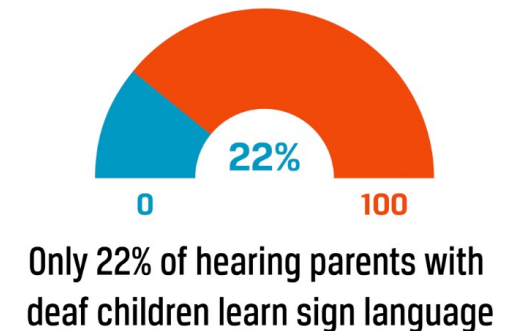
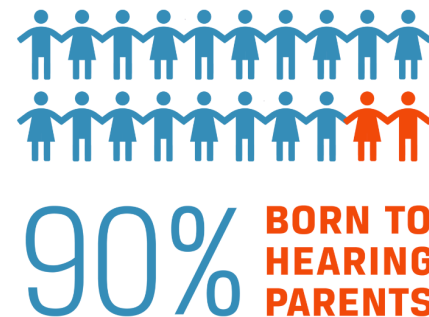
Accessibility and Inclusion: To create a more inclusive digital environment for the Deaf and Hard of Hearing (DHH) community so that they communicate in their native language on the internet and beyond.

Text search is indeed an option, but...

we ignored one important fact...
Language Deprivation and Literacy Challenges³



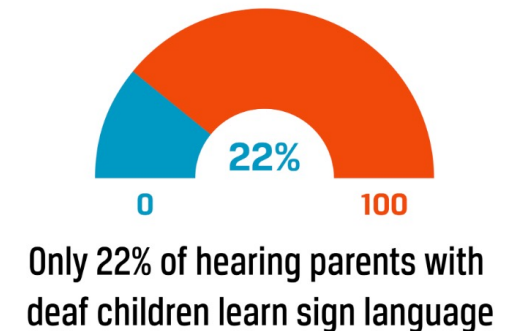
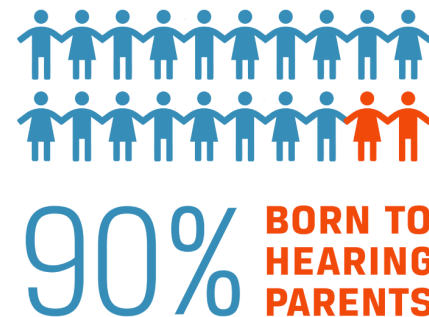
- 90% of deaf children are born into a hearing family.
- deaf children may not naturally and stably achieve fluency in any language without a frequent and accessible environment in their native language, sign language, before the age of four.



Signing is more natural to the DHH community.

Natural and Intuitive Interaction: signing is a natural and intuitive process. Allowing queries through signs can make the search experience more fluid and instinctive for them.

- 90% of deaf children are born into a hearing family.
- deaf children may not naturally and stably achieve fluency in any language without a frequent and accessible environment in their native language, sign language, before the age of four.

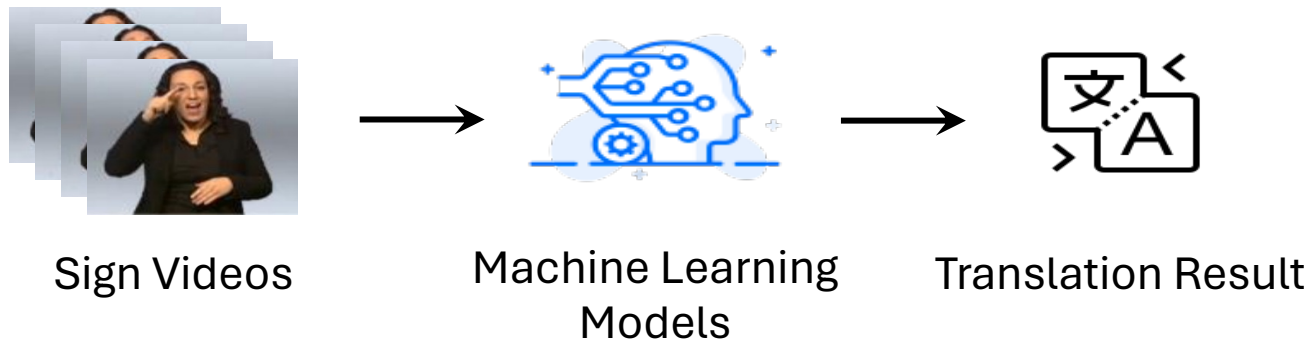


How to enable direct
sign search?



Related work: Translating signs into text before searching

Video based sign language translation systems^{4,5} as an example



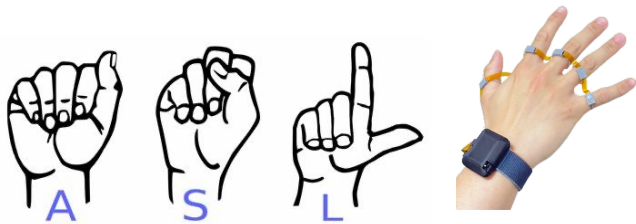
These systems require video inputs, which

- exposes private personal information,
- is sensitive to lightness, occlusions, etc.
- is not ubiquitous.

Importantly, they are still text-based.
(recall existing Language Deprivation
and Literacy Challenges)

SignQuery

Signs are captured
by our IMU device.

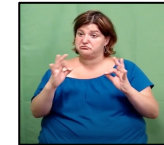


Storage for all documents



Outcomes

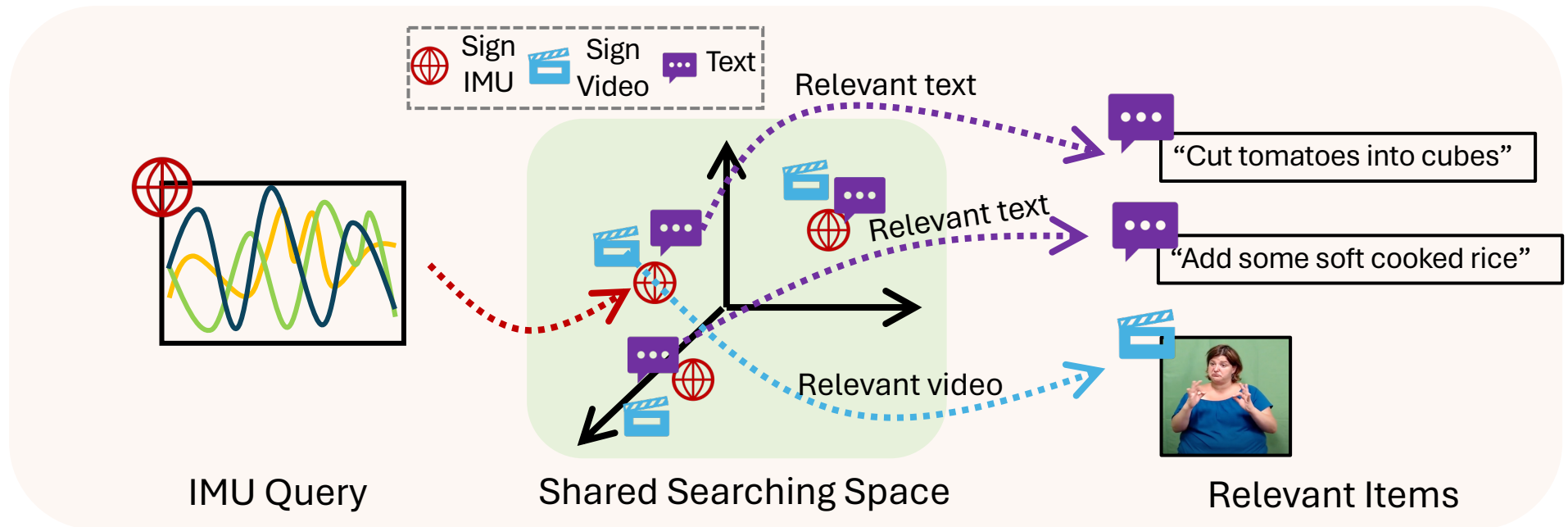
“Cut tomatoes into cubes”
“Add some soft cooked
rice”



⋮

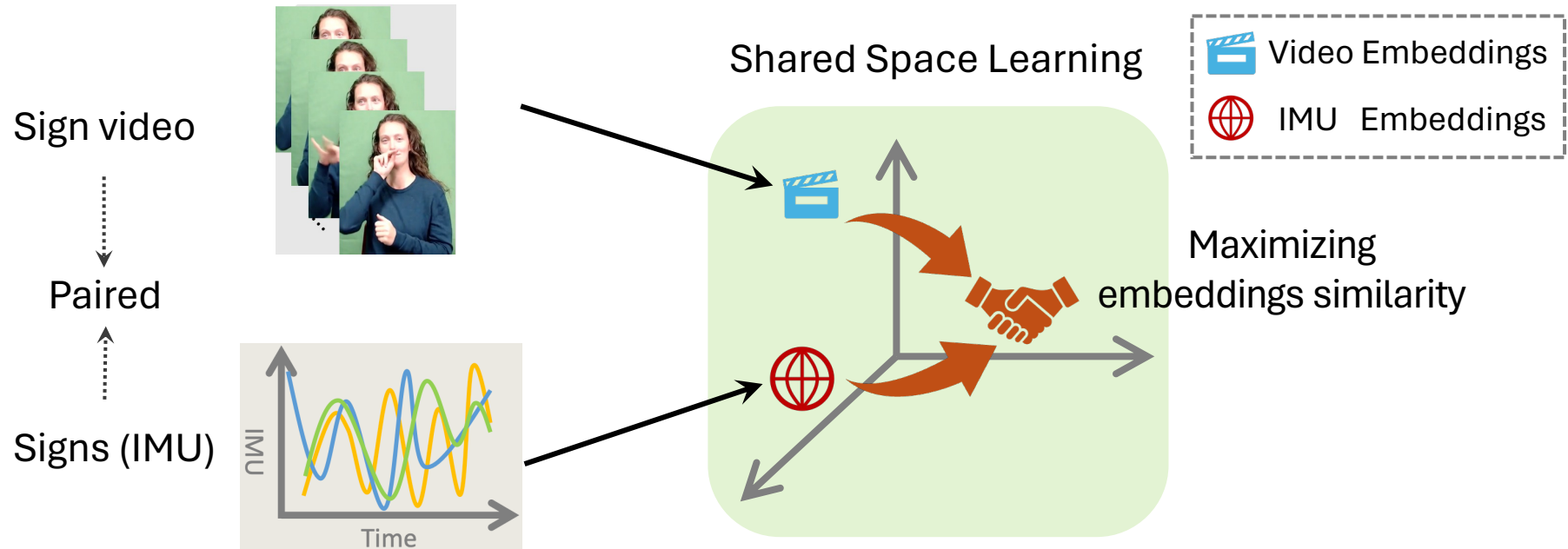
A recommendation system that supports a more natural way of searching (signing vs texting) for the DHH community, thus promoting accessibility and ensuring equitable access principles for all Deaf users to search online.

Key Idea: Encoding different modalities into a shared search space



SignQuery supports direct sign search (in form of IMU) directly in a shared space and returns relevant items.

How to build shared search space with different modalities?



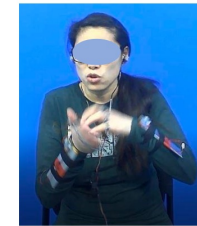
Observation: video contents are more complex and richer than that of IMU signals
➤ made the learned shared space ineffective

Video vs. IMU signal

- Video contents tend to be more complex and richer
 - lots of redundancy
 - insignificant information from background
- Users might have different facial expressions even they were signing for the same thing
- While IMU signals contain information from hands only



User 1

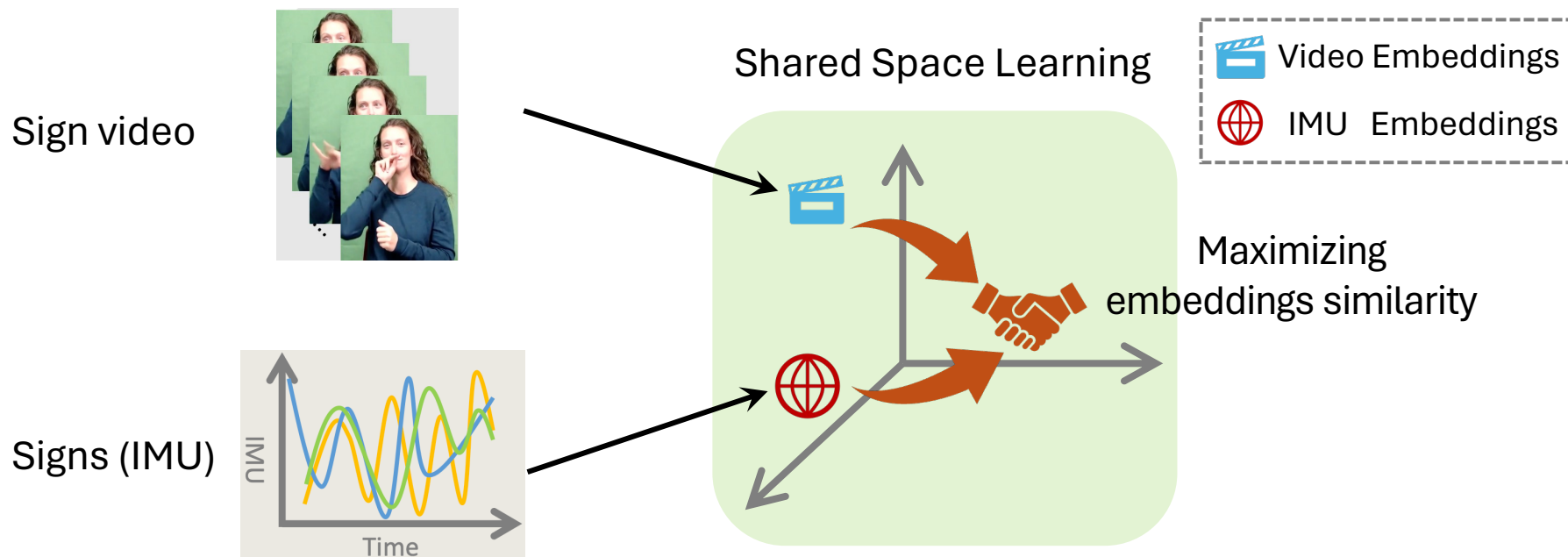


User 2



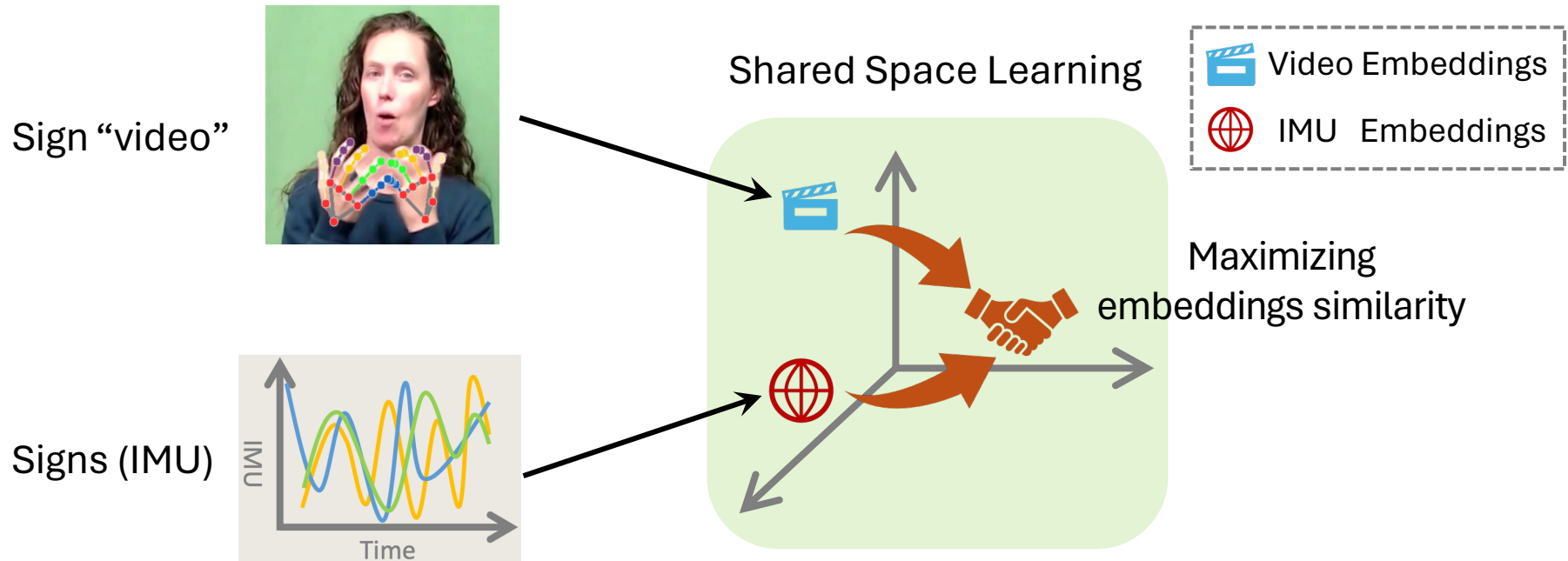
We need to simplify videos to make the learning process efficient.

Video Simplification



Instead of embedding sign videos directly, we simplify videos by extracting hand joint locations, then embed “videos” with IMU signals.

Video Simplification



Instead of embedding sign videos directly, we simplify videos by extracting hand joint locations, then embed “videos” with IMU signals.

Evaluation

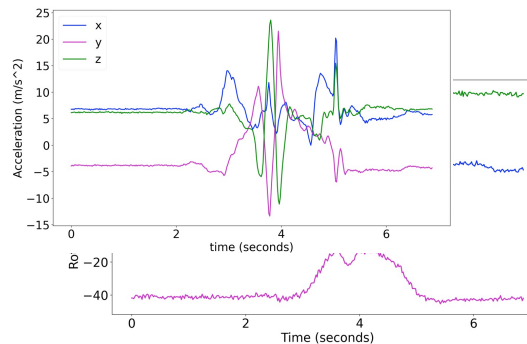
Training Data Details

- adopting from How2Sign⁶
- cover various of topics such as sports, arts, personal care, education.

Our User Study Details (Testing Data)

- 12 native ASL signers from D.C. area (mainly GU).

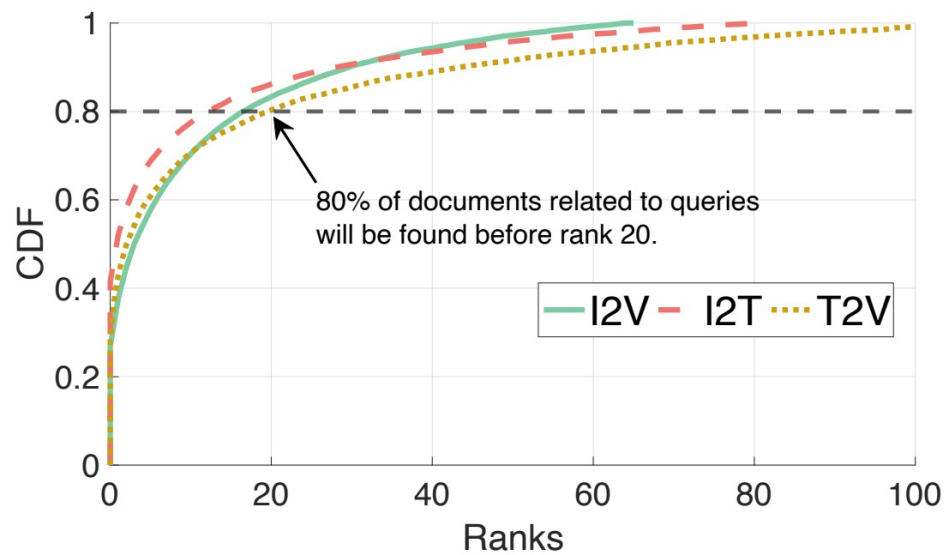
Properties	<i>How2Sign</i> [32] (Train/Evaluation)	<i>SignQuery</i> (Ours) (Test)
Language	ASL	ASL
Duration (h)	79	24
No. of Signers	11	12
No. of Sentences	24109 / 3178	2571
Modality	Video, Text	Video, Text, IMU



Gloss **SEE IX CHAMP3**

English **I am happy to see you!**

Evaluation

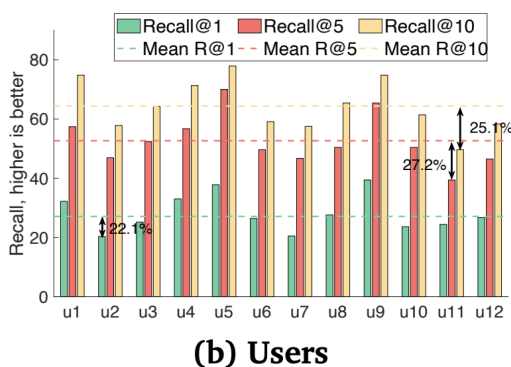


(a) Rank Distributions

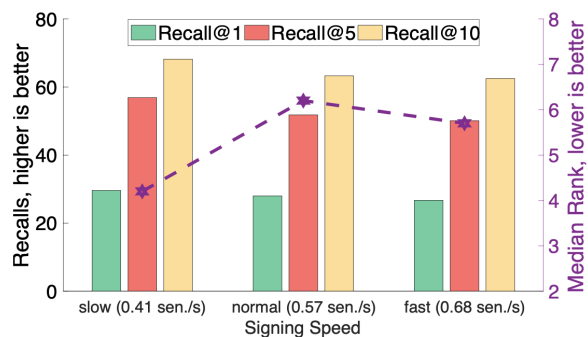
***I2V** is to use **IMU** as queries and returning **Videos** as results

More than 80% of documents that are related to queries can be found before rank 20 (from 24k).

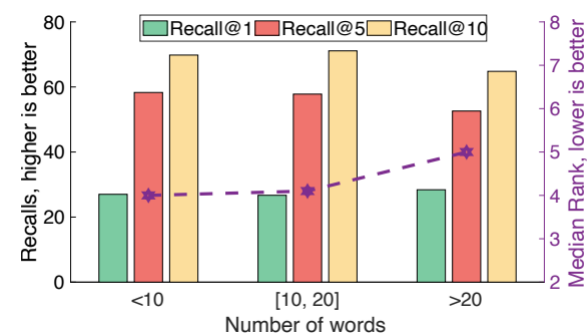
Evaluation



SignQuery performs well with 12 new users in our user study.

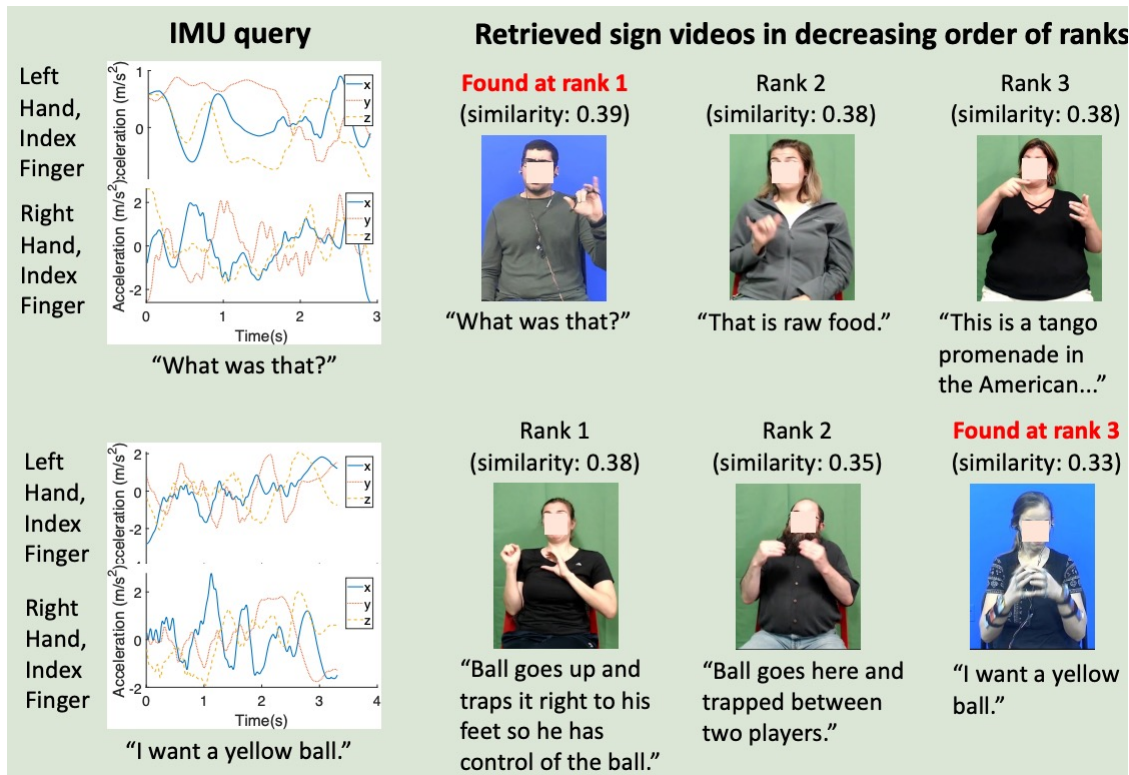


Stable across various signing speeds.



Stable across various sentence lengths

Evaluation



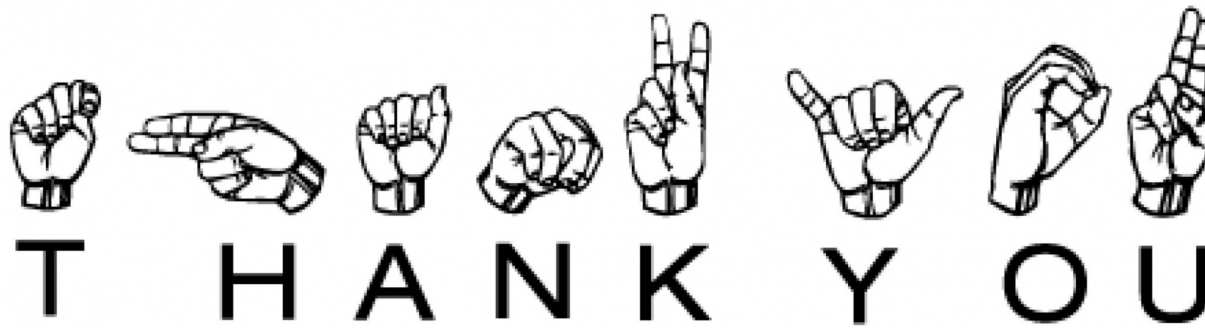
Using IMU as queries,
relevant documents (e.g.,
sign videos) can be retrieved
accurately.

Take-aways

Signing is always preferred in the DHH community. SignQuery is the first system to support native signs, thus ensuring equitable access principles for all Deaf users to search online.

Embedding all into one is our solution to eliminate the need of text that might not be available for all Deaf people.

SignQuery and its beyond. We believe current version of SignQuery only demonstrates a part of its potential. Other applications such as ASL animation, Sign Language Recognition can be extended on top of SignQuery.



Contact: **Hao Zhou** / hao.zhou@psu.edu



Reference

1. Bhuvaneswar C, Colon J, Matthews J, et al. Psychiatric care of deaf patients in the general hospital: an overview. 2015;56:1-11.
2. Centers for Disease Control and Prevention (CDC): National Center for Health Statistics. Summary Health Statistics: National Health Interview Survey, 2017 (pdf). Available at: https://ftp.cdc.gov/pub/Health_Statistics/NCHS/NHIS/SHS/2017_SHS_Table_A-6.pdf Accessed Dec 26, 2018.
3. Linguistic Neglect of Deaf Children in the United States. <https://ballardbrief.byu.edu/issue-briefs/linguistic-neglect-of-deaf-children-in-the-united-states>
4. Camgoz, N. C., Hadfield, S., Koller, O., Ney, H., & Bowden, R. (2018). Neural sign language translation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7784-7793).
5. Camgoz, N. C., Koller, O., Hadfield, S., & Bowden, R. (2020). Sign language transformers: Joint end-to-end sign language recognition and translation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10023-10033).
6. Duarte, A., Palaskar, S., Ventura, L., Ghadiyaram, D., DeHaan, K., Metze, F., ... & Giro-i-Nieto, X. (2021). How2sign: a large-scale multimodal dataset for continuous american sign language. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 2735-2744).