Application Informed Motion Signal Processing for Finger Motion Tracking Using Wearable Sensors

Yilin Liu
Fengyang Jiang
Mahanth Gowda

ICASSP 2020

PennState

Motivation

Finger motion tracking has a lot of useful applications in user-interfaces, sports analytics, medical rehabilitation and sign language translation.

Whole finger tracking is difficult, but application context can fill the gap

Deaf and hard of hearing community (DHH) use Sign Language (SL) for communication, but other people do not understand

We want to create an accurate, convenient and non-intrusive translator for the application of American Sign Language (ASL) recognition

Examples of ASL words

Hello

My

Friend

PennState

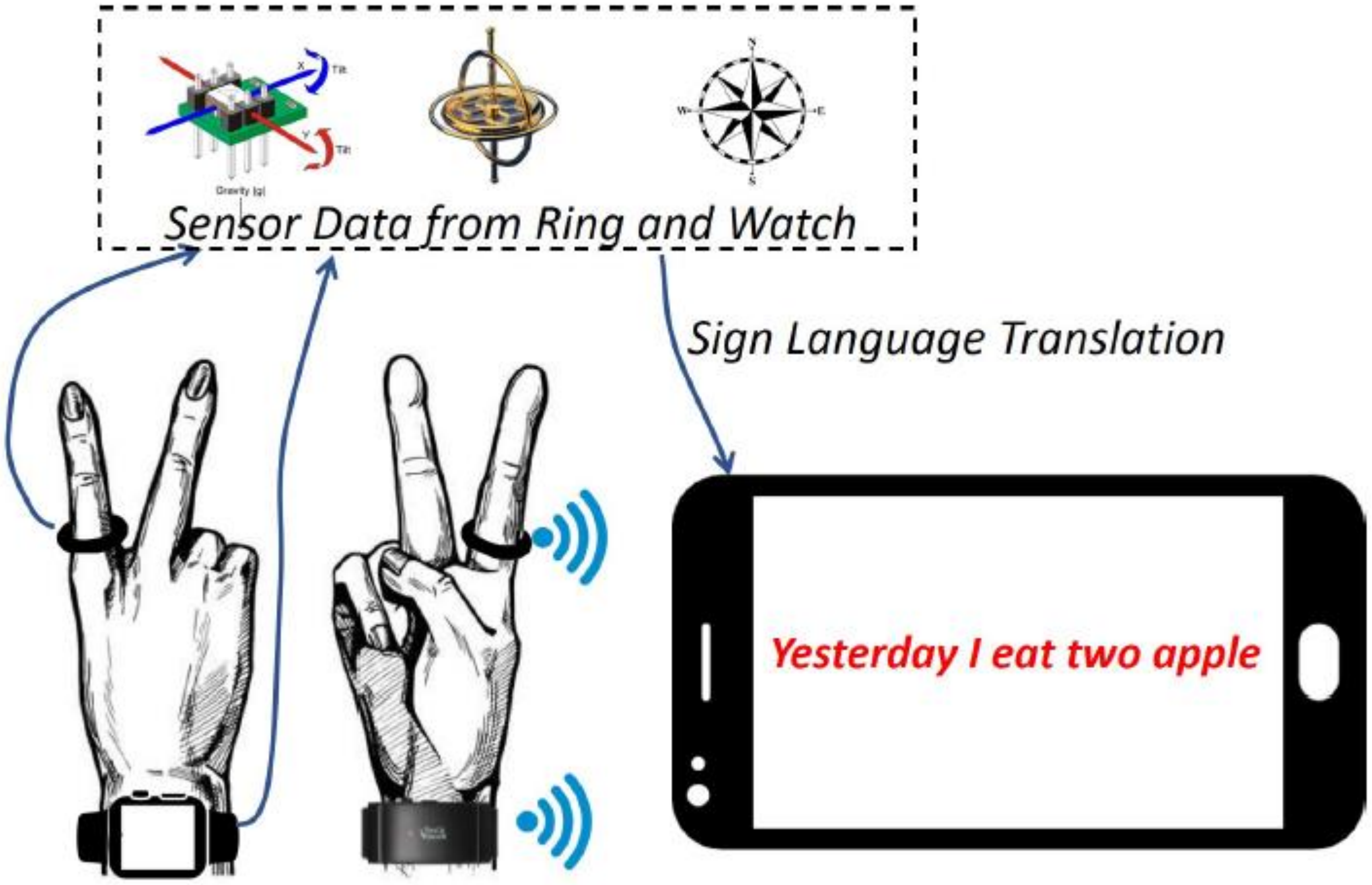
Existing Works

Only classify a few word gestures

Too many sensors -- intrusive

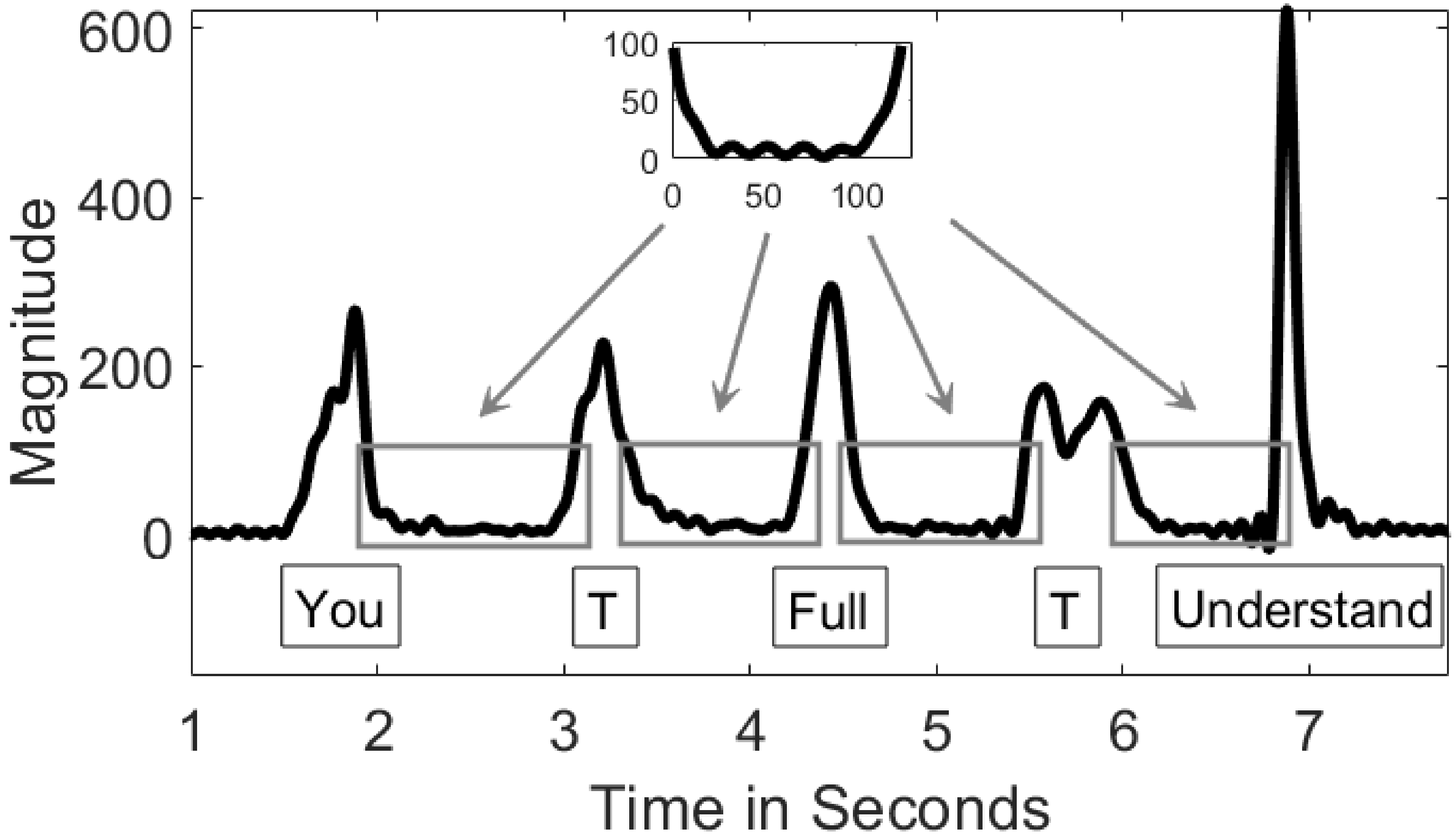
Cameras are not ubiquitous and need good lighting

Our System: FinGTrAC



PennState

Segmentation & Preprocessing

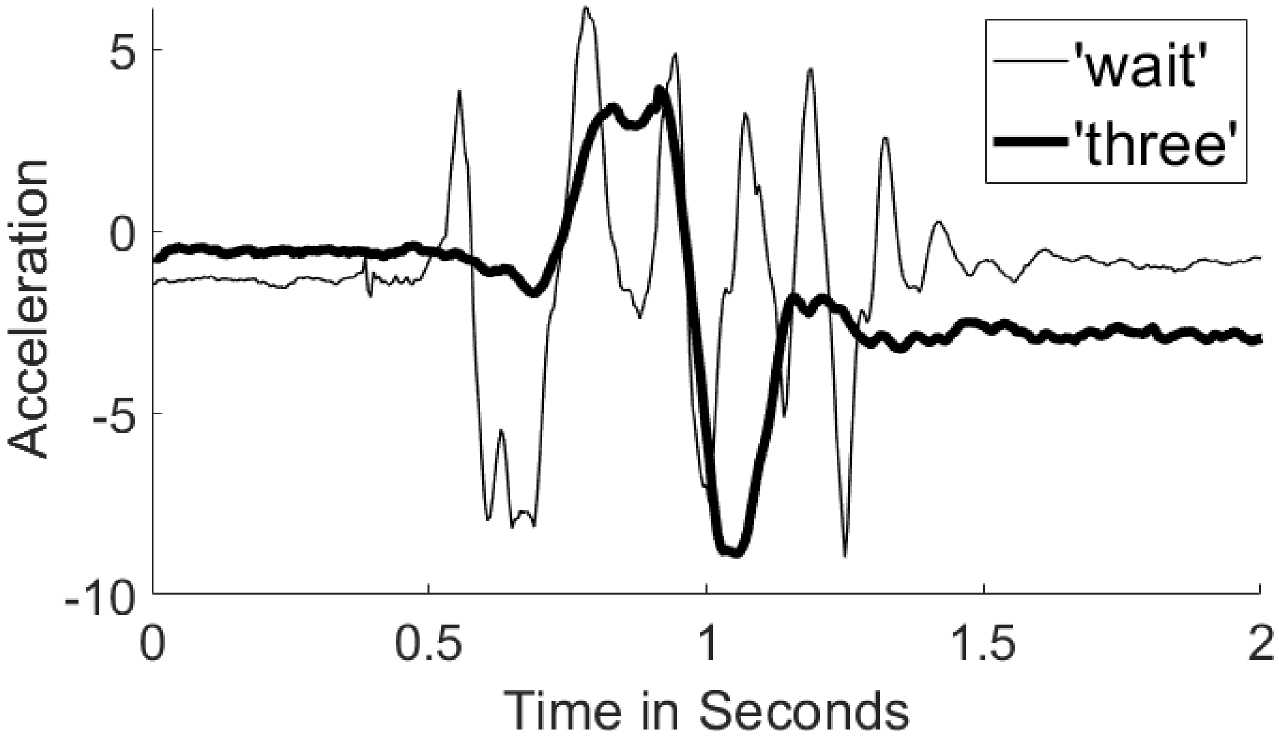


Segmentation & Preprocessing

Low pass filter

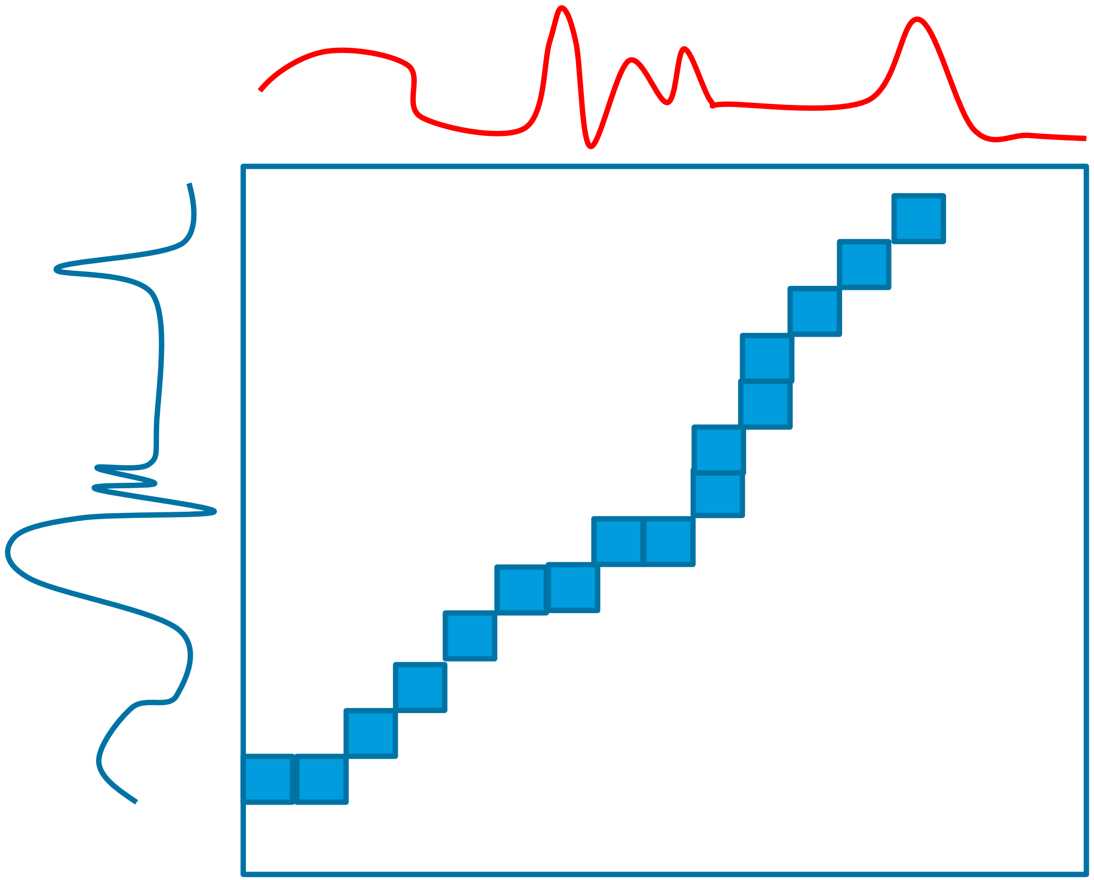
Peaks

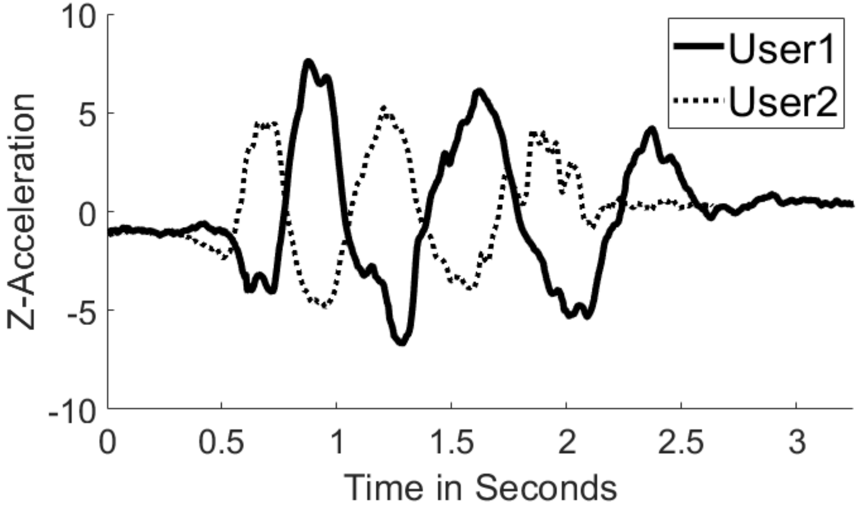
Orientations…

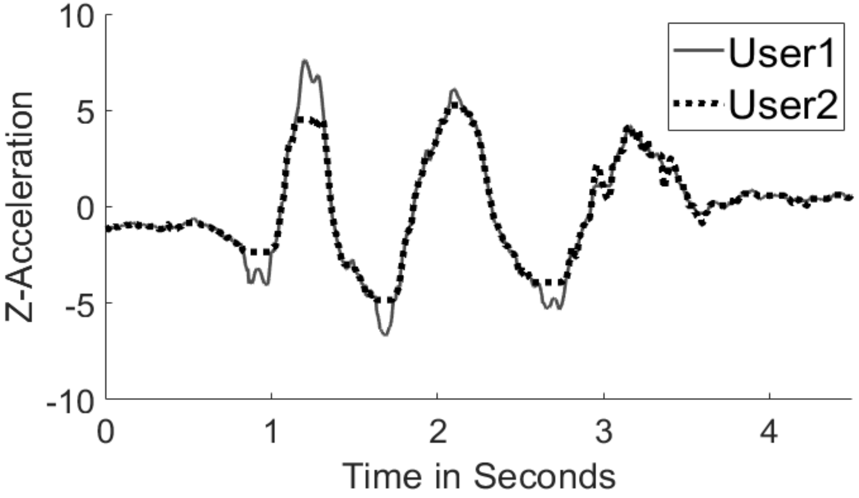


PennState

Dynamic Time Warping (DTW)



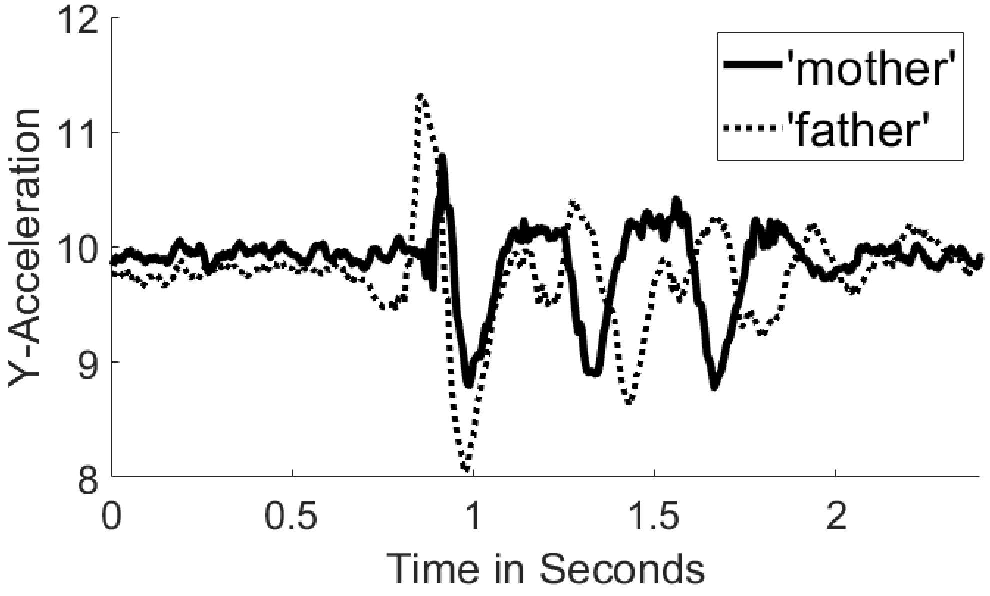


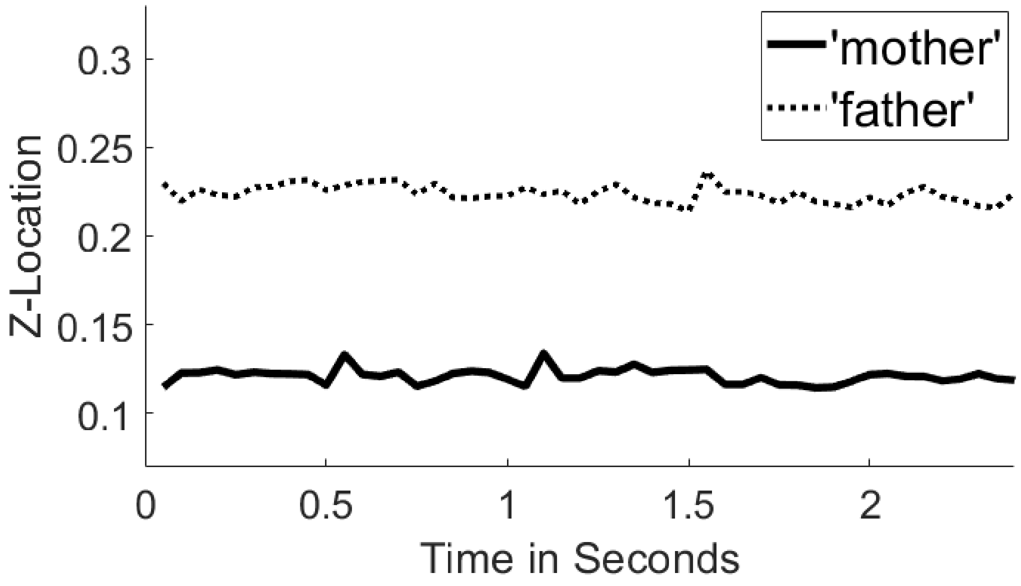


Failed Cases

However......

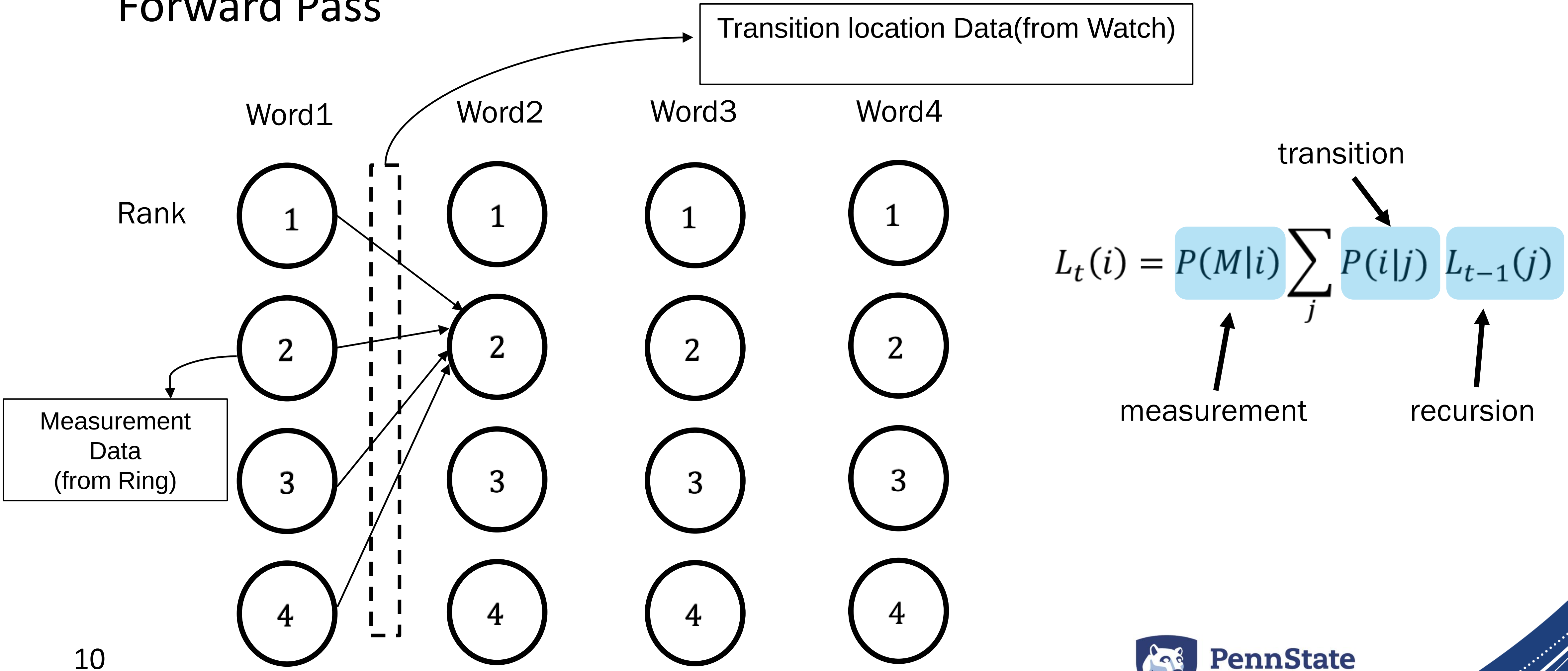
The accuracy is only 70.1%





Hidden Markov Model (HMM)

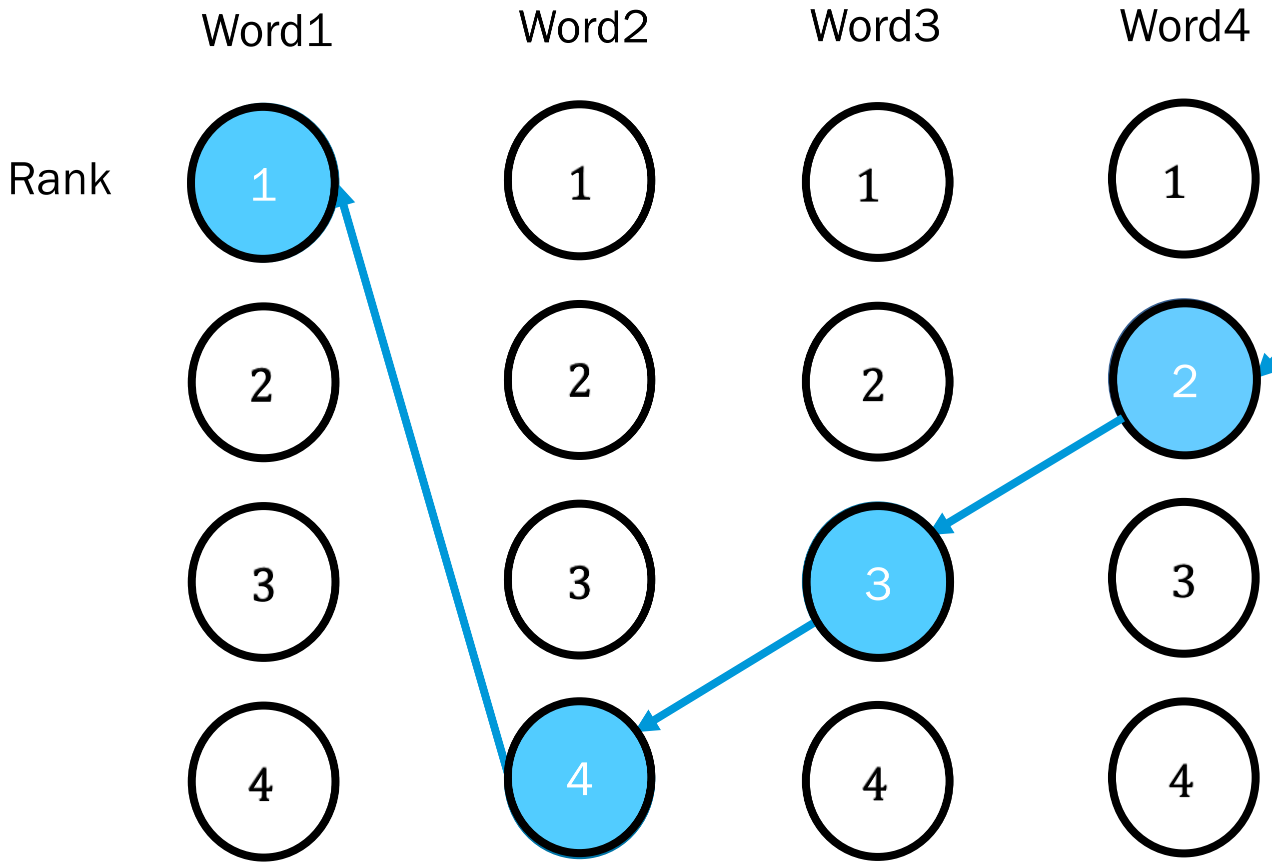
Forward Pass



$$L_t(i) = P(M|i) \sum_j P(i|j)\, L_{t-1}(j)$$

Hidden Markov Model (HMM)

Backward Pass



Maximum Likelihood

transition

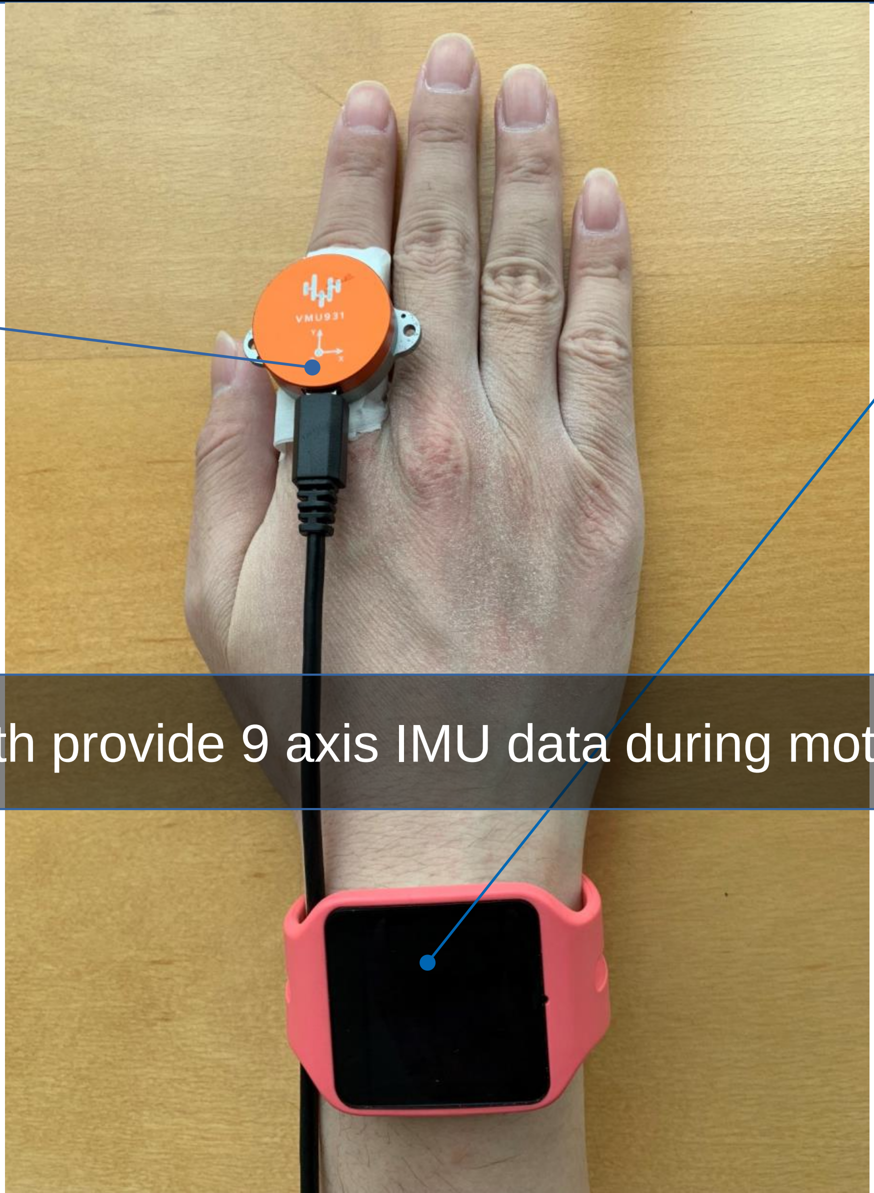
$$w(K-1) = max_j(L_j(K-1) \cdot p(w(K) \mid j))$$

PennState

Evaluation platform

VMU931 IMU sensor

SONY SmartWatch 3 SWR50



Both provide 9 axis IMU data during motion

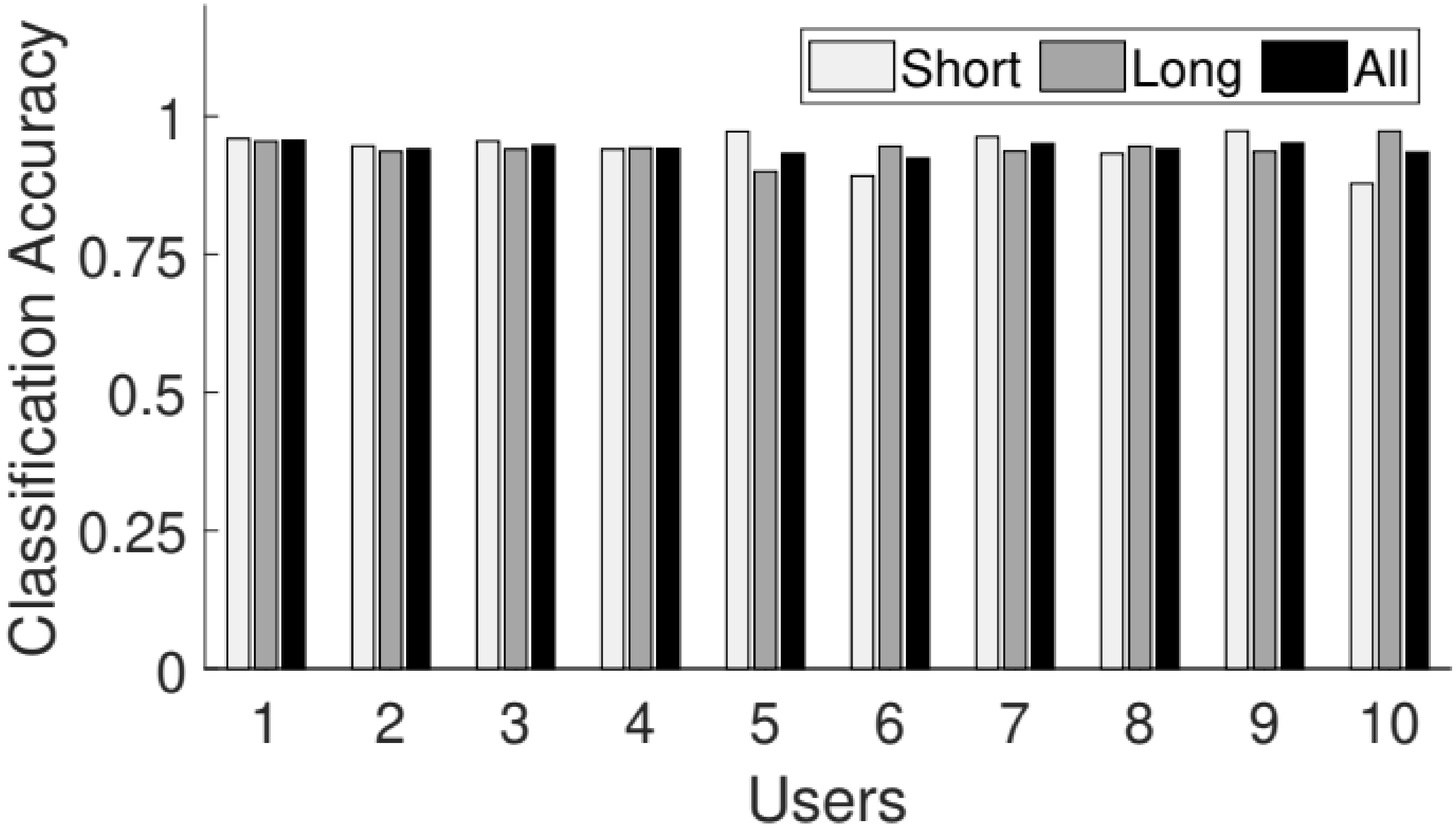
User Study

10 Users (7 male and 3 female)

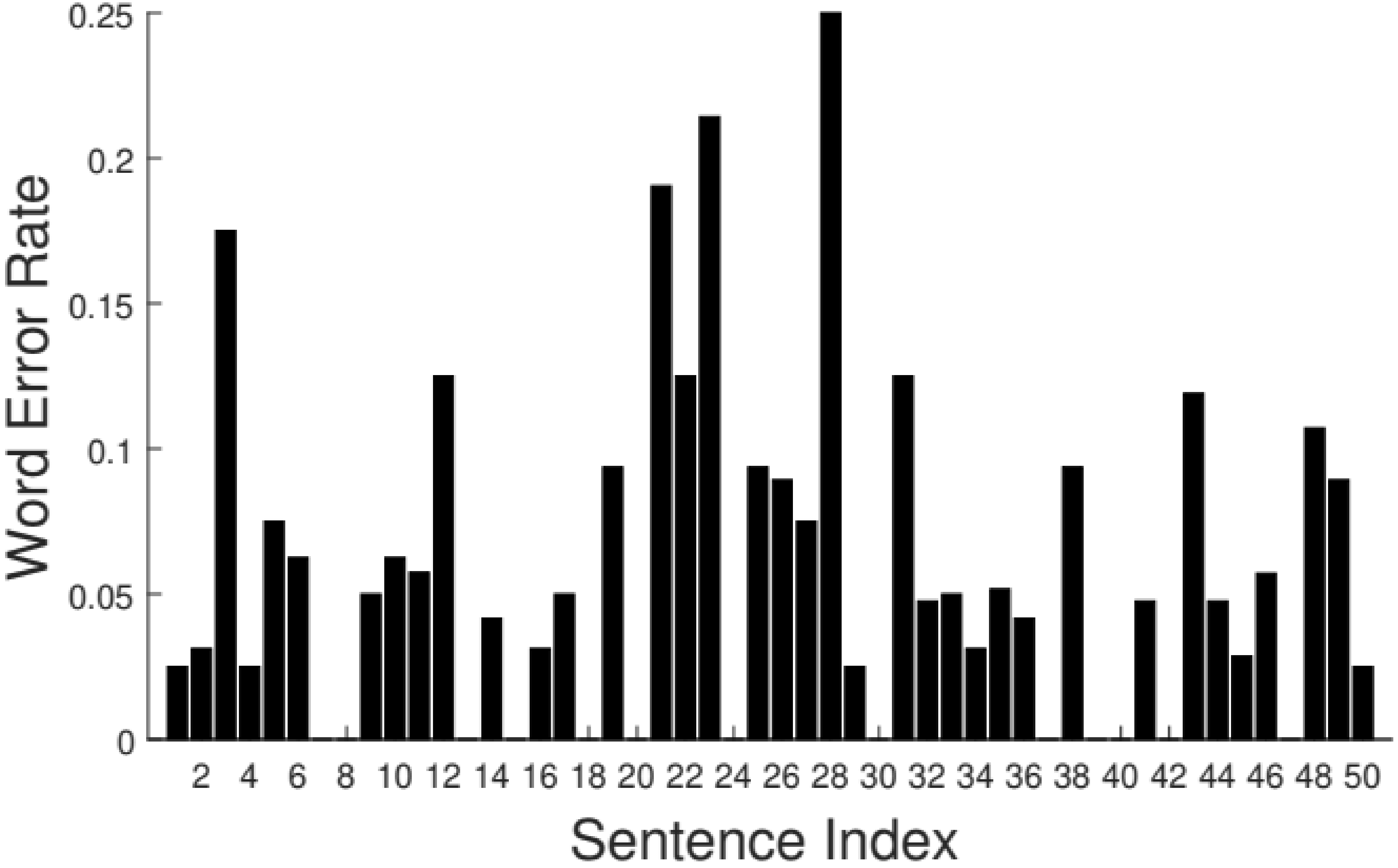
Each user perform 50 ASL Sentences

The sentences were composed from a dictionary of 100 most frequently used ASL words

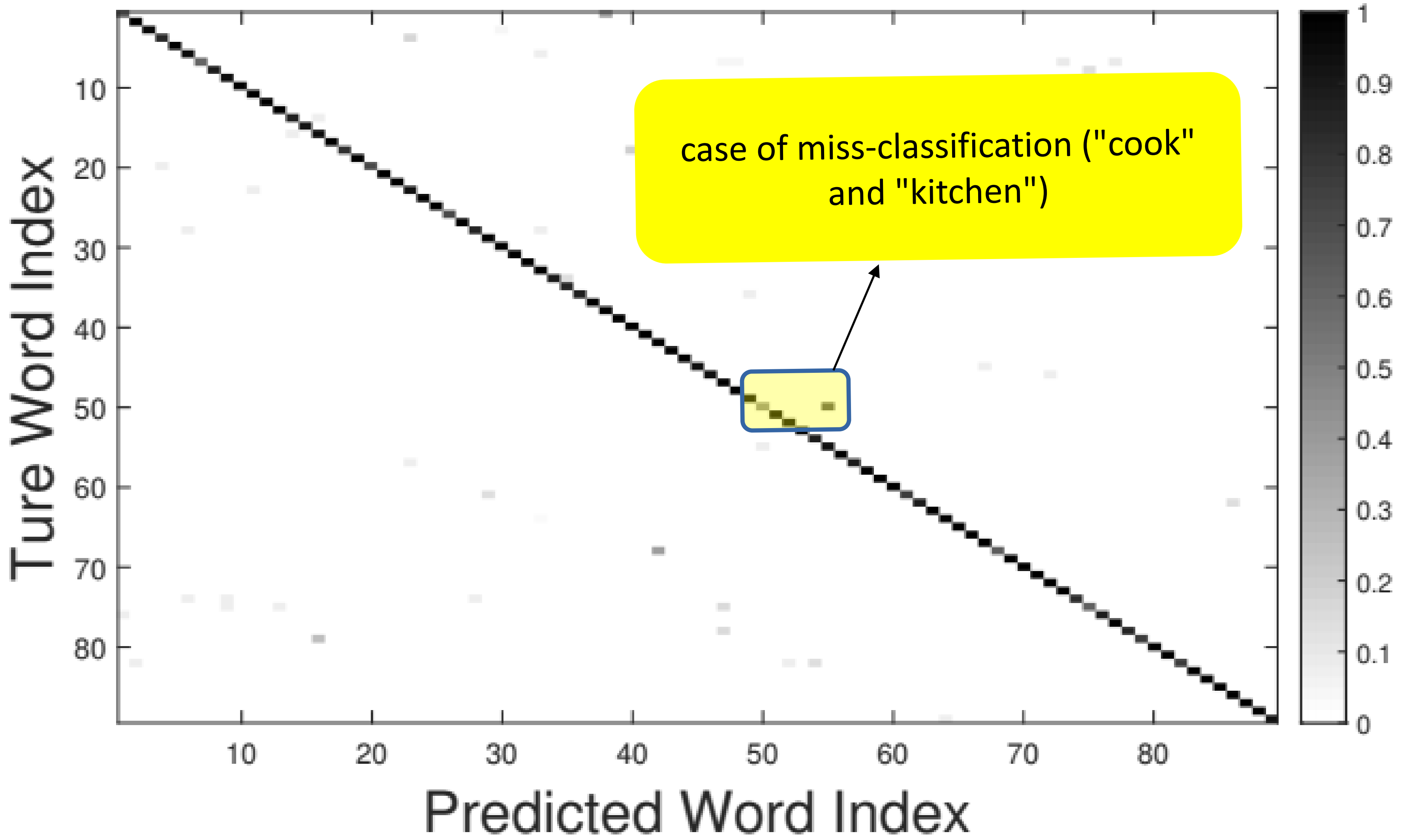
Classification accuracy is consistent across diverse users



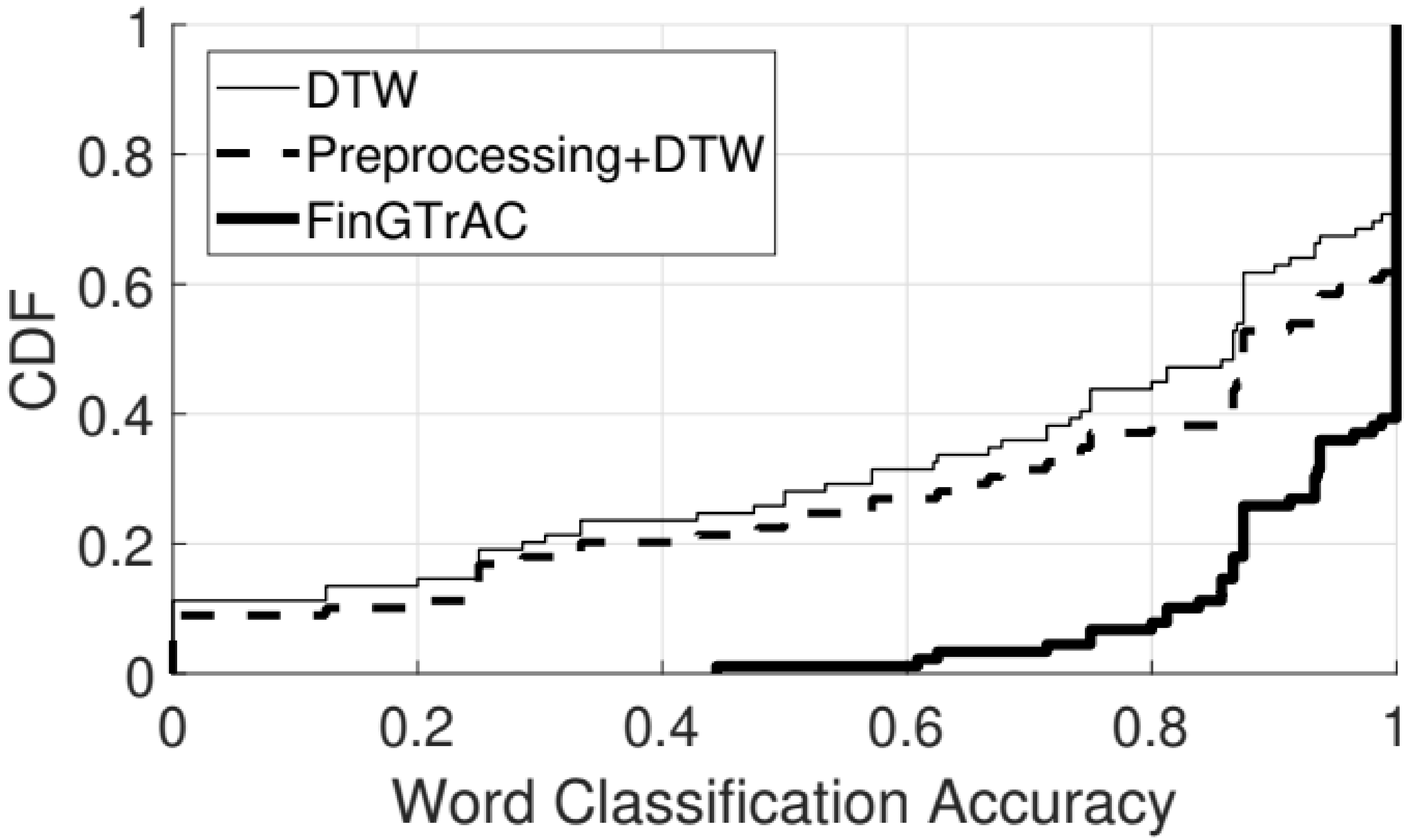
Word error rate across sentence



Most words was decoded with high classification accuracy



Classification accuracy improves with successive stages of optimization



Conclusion

Our system FinGTrAC shows the feasibility of finger gesture classification from only two non-intrusive and low-cost wearable sensors and very minimal training

We use a probabilistic framework incorporating the noisy motion sensor data, as well as contextual information between ASL words to decode the most likely sentence

We conduct systematic user study with 10 Users and show a word recognition accuracy of 94.2% over a dictionary of 100 most frequently used ASL words

Future work

- Plan to exploit transfer-learning approaches to train wearable sensor data from videos of online ASL tutorials

- Natural Language Processing (NLP) or Deep-Learning technique to expand the dictionary size of our system

- Facial expression detection using EMG (facial muscles), EEG (brain signals), and EOG (eye signals)

PennState