RF-Sensing for ASL-Based Human Computer Interactions

Dr. Sevgi Zubeyde Gurbuz
Lab for Computational Intelligence in Radar (CI4R)
Dept. of Electrical and Computer Engineering
The University of Alabama, Tuscaloosa, AL
szgurbuz@ua.edu
http://ci4r.ua.edu

August 19, 2022

Our Team

Sevgi Z. Gurbuz
Electrical Eng.
Univ. Alabama

Chris Crawford
Computer Sci.
Univ. Alabama

Evie Malaia
Neurolinguistics
Univ. Alabama

Darrin J. Griffin
Communication
Univ. Alabama

Ali C. Gurbuz
Electrical Eng.
Miss. State Univ.

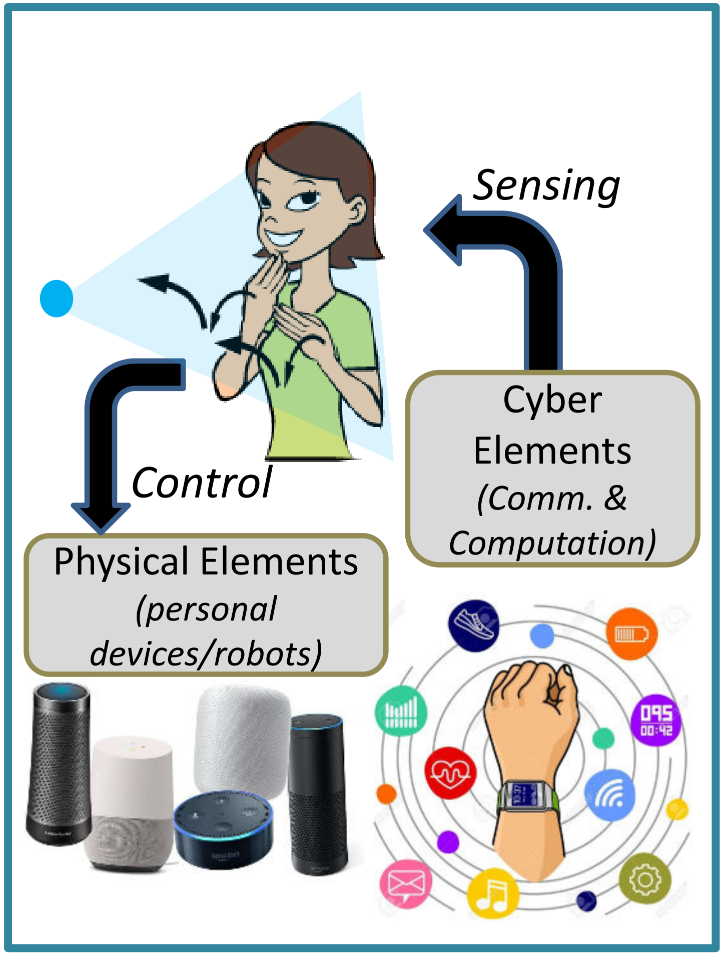
Collaborators @ Gallaudet:

Dr. Caroline
Kobek-Pezzarossi

Dr. Kenneth
De Haan

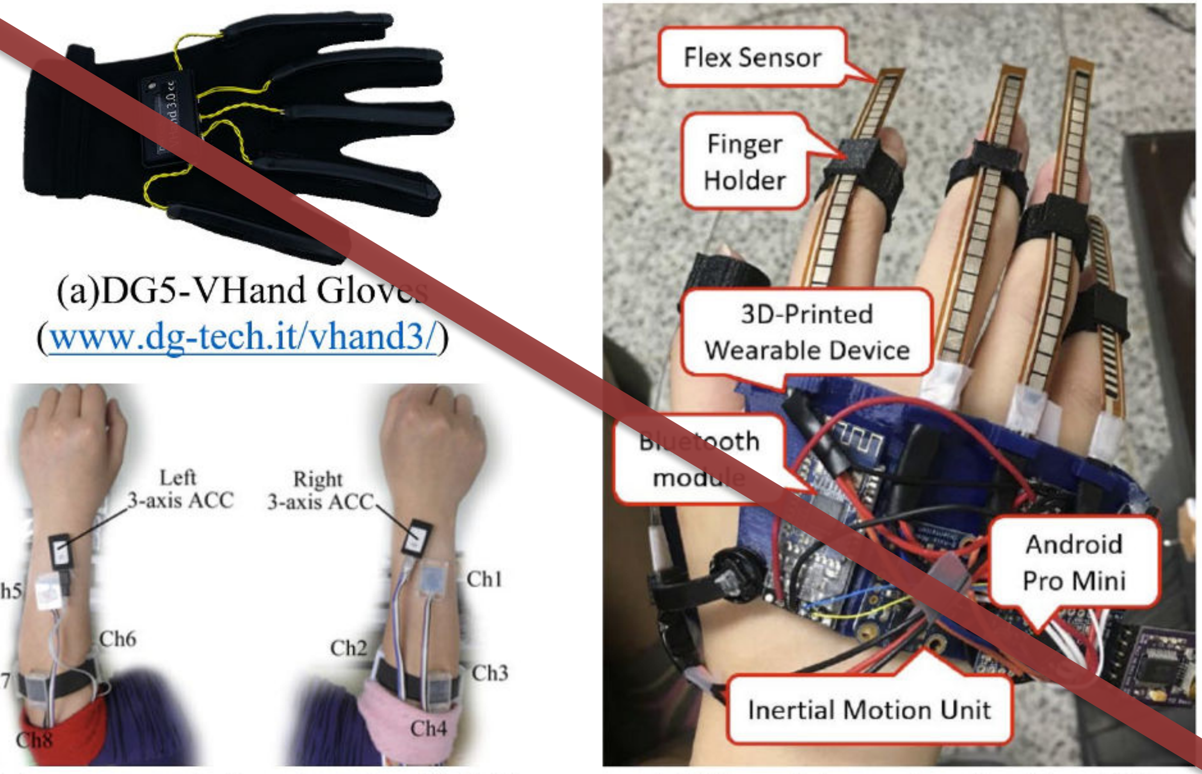
Objective

How can design technology to interact with humans via American Sign Language?



Sevgi Z. Gurbuz (szgurbuz@ua.edu)

Wearables?



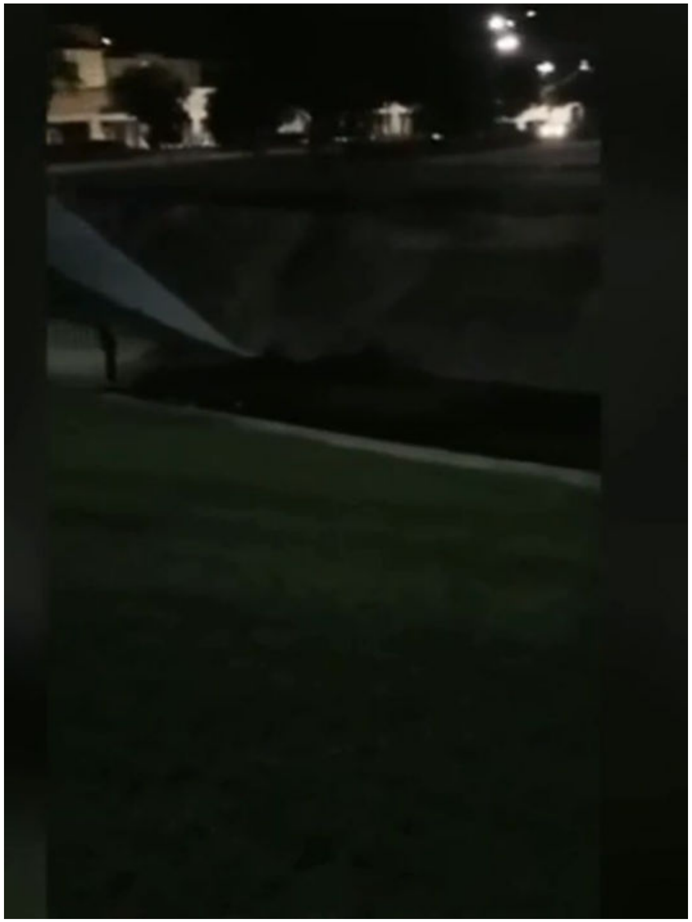
(a)DG5-VHand Gloves
(www.dg-tech.it/vhand3/)

(b)Arm-mounted system by (Li12)

(c)Wearable system by (Lee18)

Video?

Limitations of Video



- Not effective in the dark
- Narrow field of view
- Can be affected by skin tone or the color of the clothes you wear

Non-Contact Sensing

- Radio Frequency (RF) Sensing

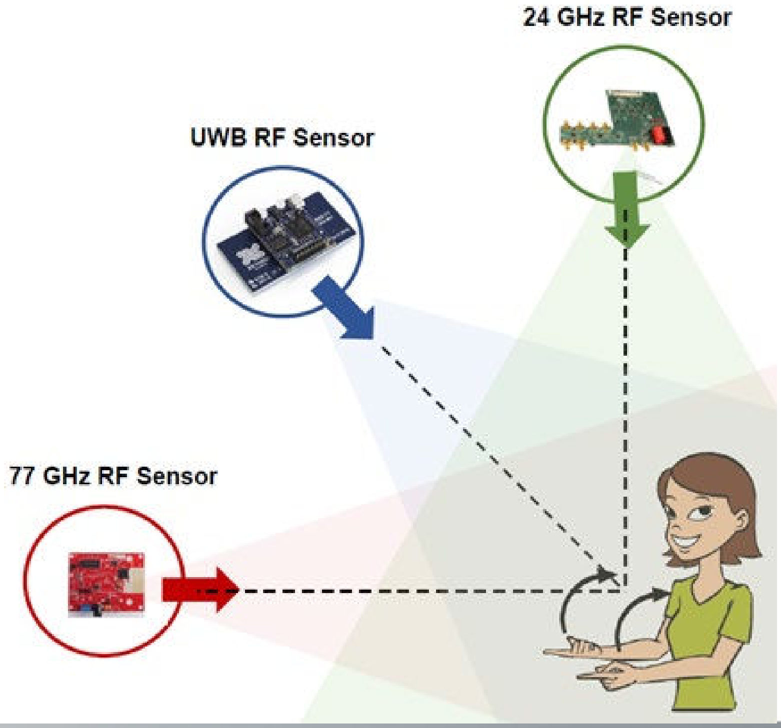
Non-Invasive, remote

Works day and night

Protects privacy

Wide field-of-view

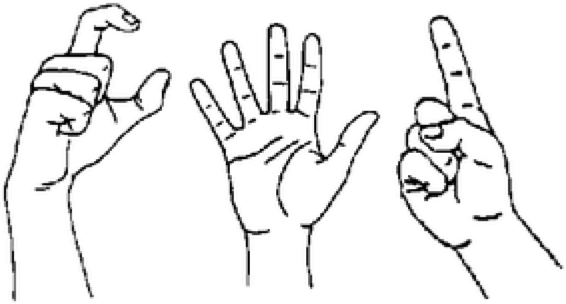
Color-blind and
not affected by clothes



Sevgi Z. Gurbuz (szgurbuz@ua.edu)

Limitations of RF Sensors

Cannot perceive facial expressions

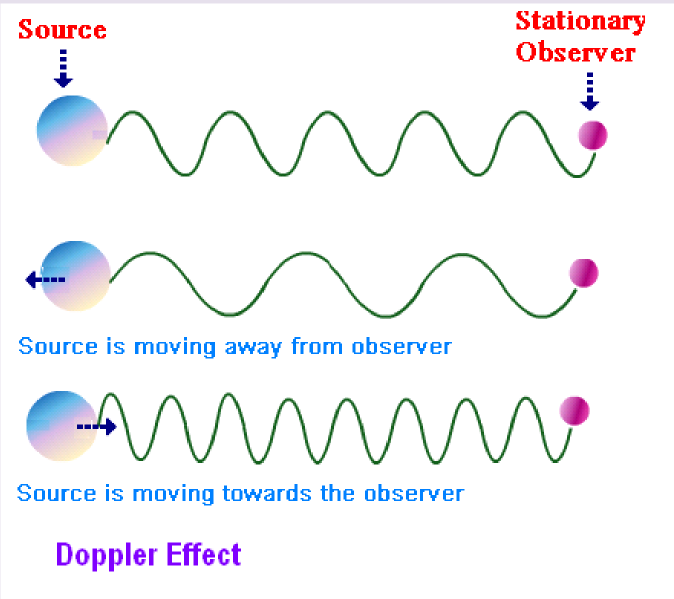


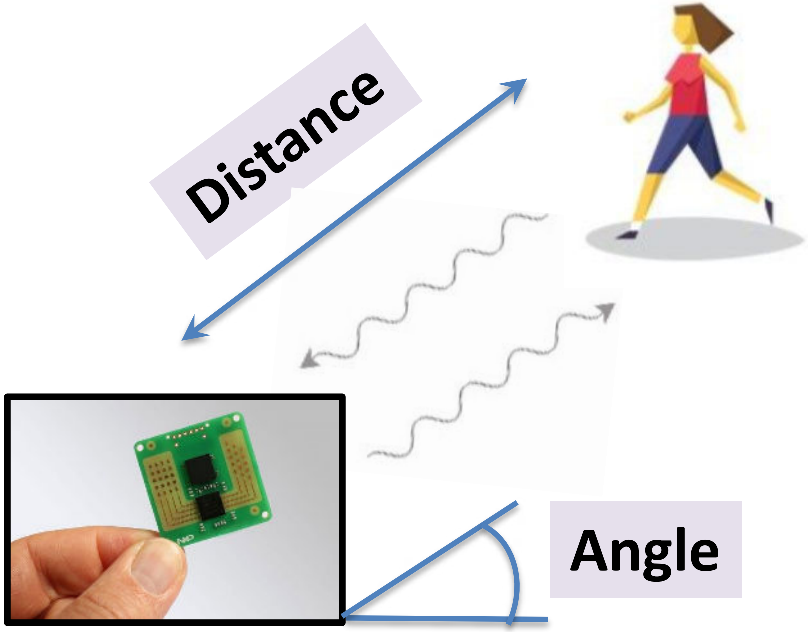
(Currently) cannot figure out hand shape

Sevgi Z. Gurbuz (szgurbuz@ua.edu)

What do RF sensors measure?

Velocity

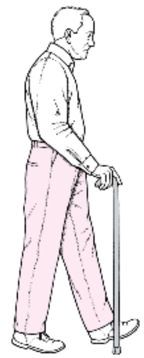




Sevgi Z. Gurbuz (szgurbuz@ua.edu)

Example Applications of RF

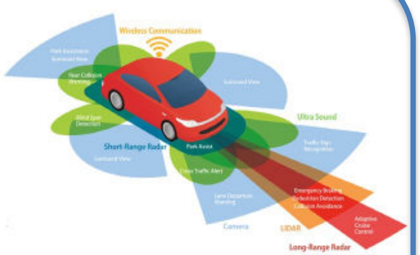
Fall Detection



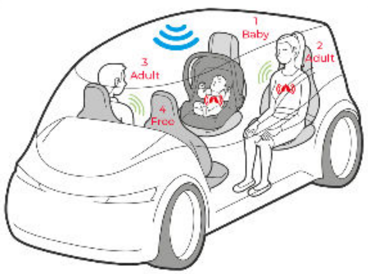
Gait Analysis

Activity Recognition

Vital Signs
Sleep State



Self-Driving Cars Safety



Passenger Monitoring

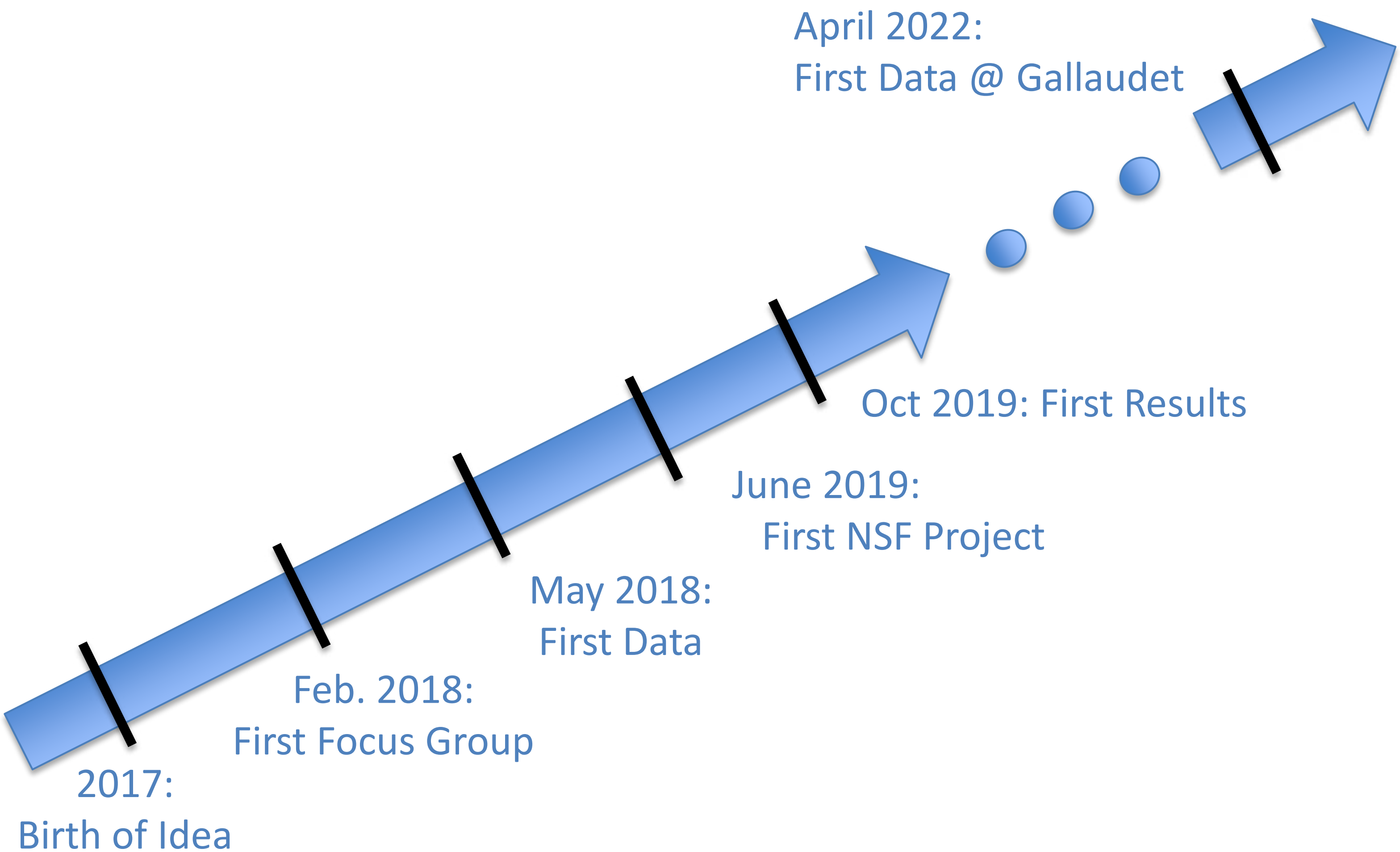
New RF Application

Sign Language Recognition

RF Only

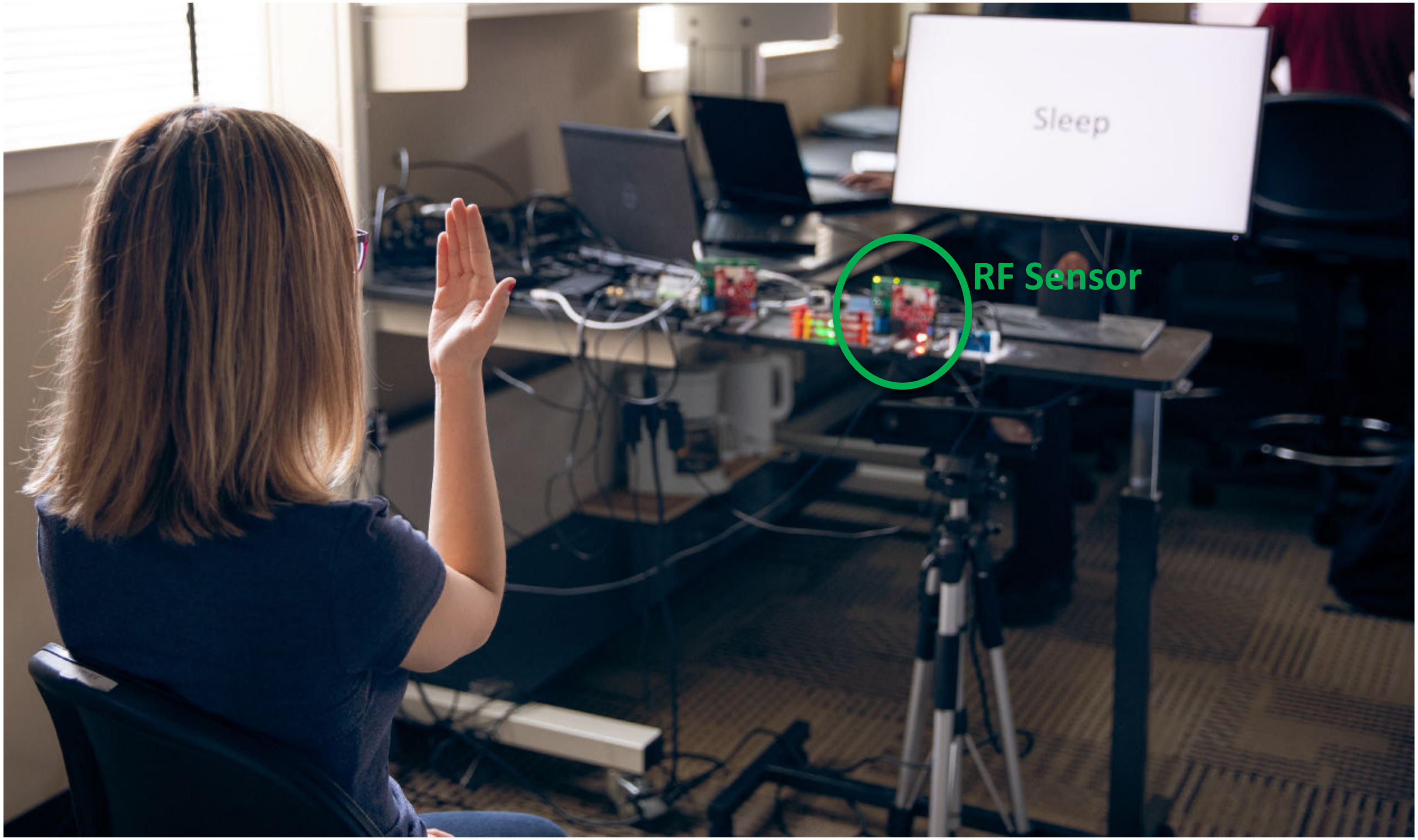
RF + Video

Evolution of Project



Sevgi Z. Gurbuz (szgurbuz@ua.edu)

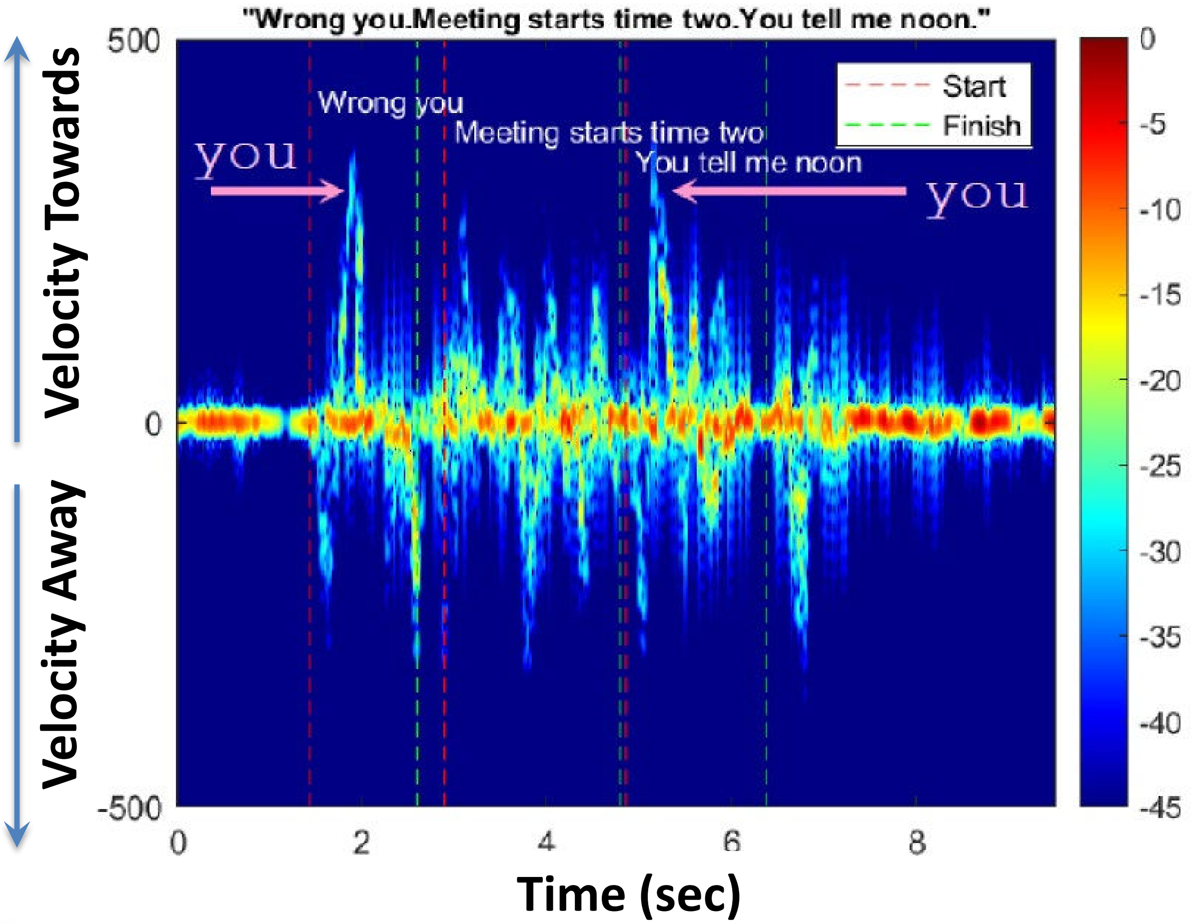
RF System



Sevgi Z. Gurbuz (szgurbuz@ua.edu)

What Does RF Data Look Like?

Micro-Doppler Signature

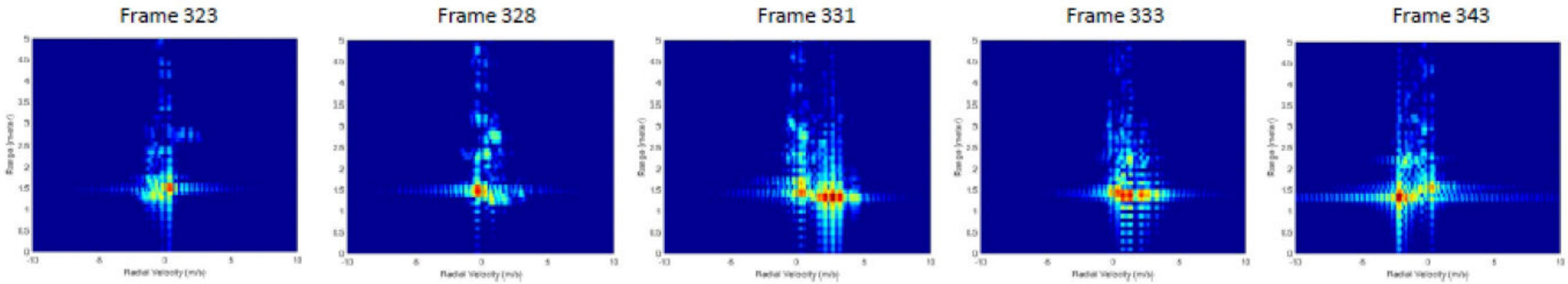


Sevgi Z. Gurbuz (szgurbuz@ua.edu)

Another Way to Visualize RF Data

<u>Distance-Velocity Maps</u>

- Snapshots of distance versus velocity at different times



Body at rest (0 m/s). Hands that were on knees, just beginning to move upward towards body, resulting in slightly negative velocity.

Hands brought just in front of chest, momentarily 0 velocity as hands prepare to reverse direction

Hands moving towards radar (positive Doppler)

Zero velocity point indicating farthest point hands moved towards radar

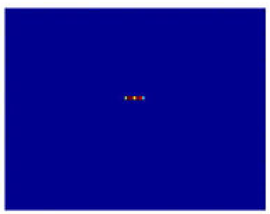
Largest negative velocity Indicates point at which hands move towards body

Samples of sequential frames from Range-Doppler video for the `breathe` sign.

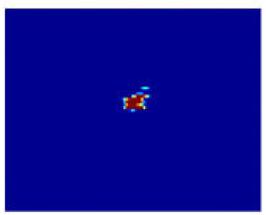
Other Ways to Visualize RF Data

Distance-Angle Maps

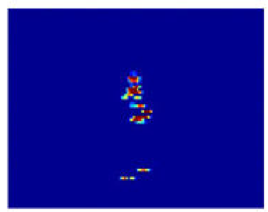
- Snapshots of distance versus angle at different times



Subject is sitting stationary

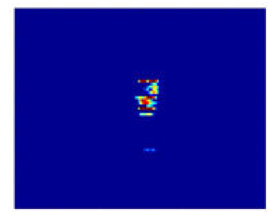


Hands start to move

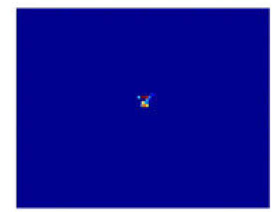


Hands moving towards the body

A short pause upon reaching the chest

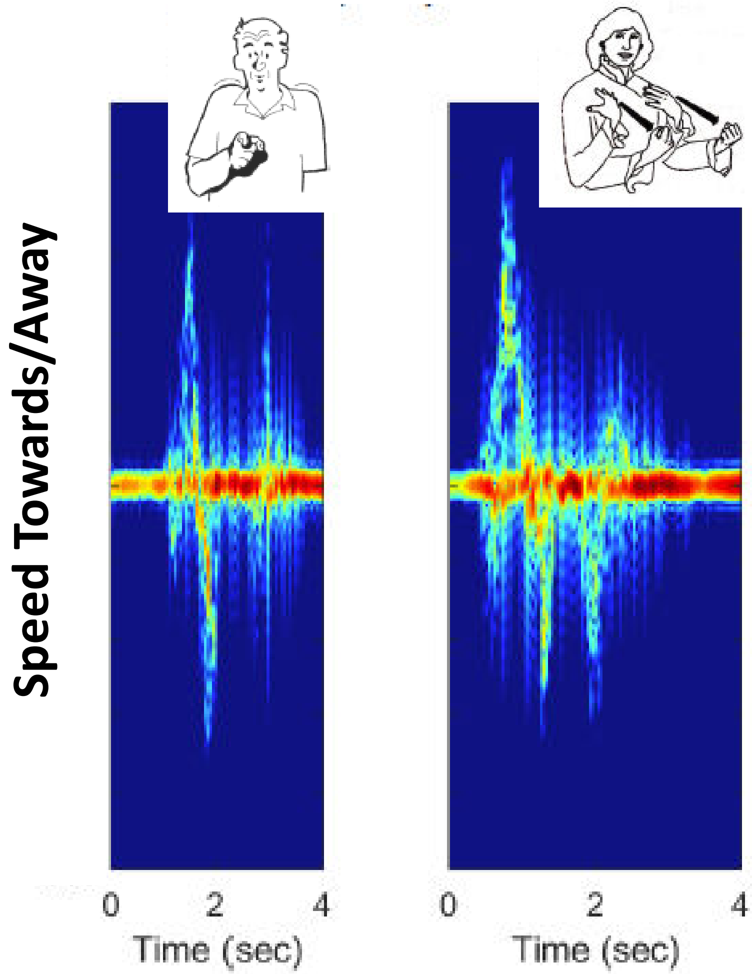


Retracting hands to original position



Sign completed

Sample ASL Signs in RF

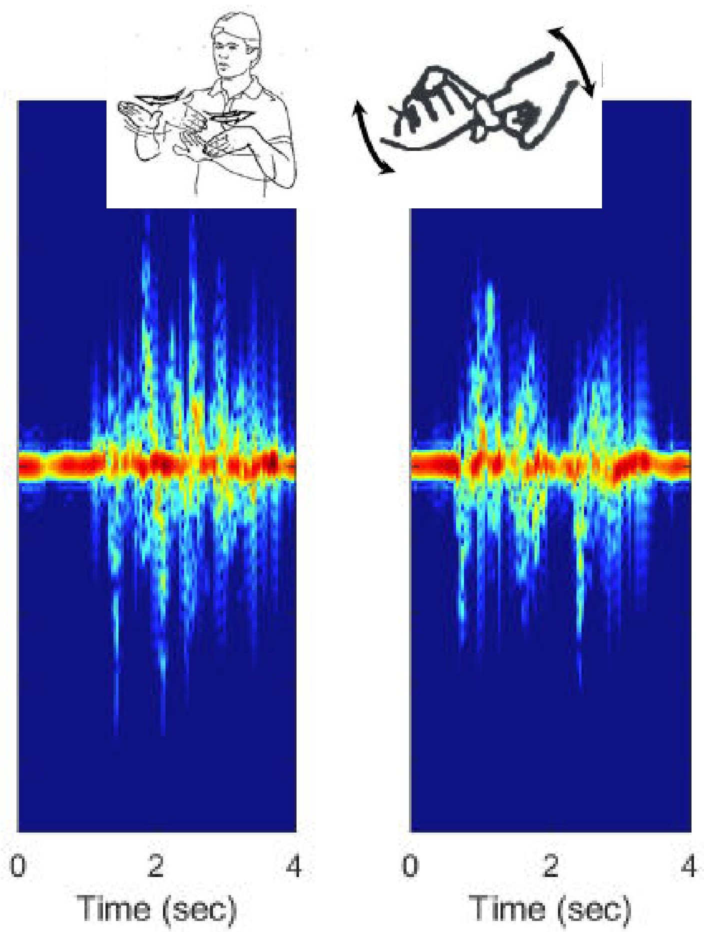


Single or Two-Handed Signs
(where hands do similar things)

- Can tell if hands moved towards or away by whether data peaks upward or downward
- Dominant direction at any given time

Sevgi Z. Gurbuz (szgurbuz@ua.edu)

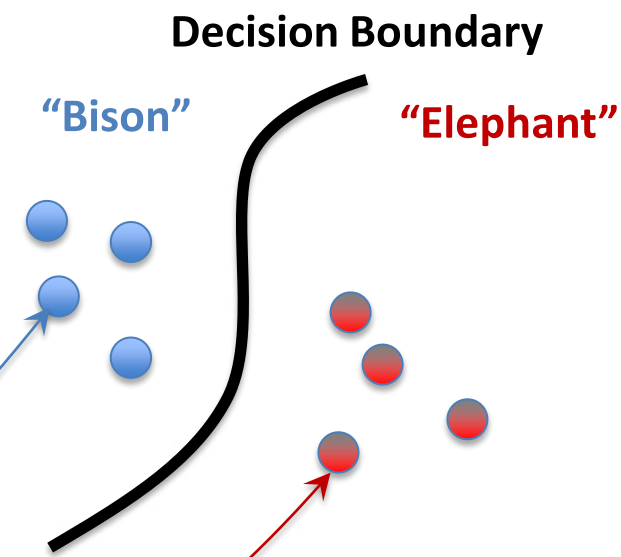
Sample ASL Signs in RF (cont.)



Two Hands Moving Independently

- Peaks in the data both upward and downward at the same time

How Do Computers Figure it Out?



Brown
Horns
Hoofs
White Shirt

Grey
Flappy Ears
Big Feet
Red Shirt

New Data:
Which side is it on?

Approach: Deep Learning

- Takes massive amounts of **DATA** for figuring out decision boundaries

- Roadmap for Rest of Talk:
  – Which data? The impact of fluency
  – Data Synthesis
  – Recognizing single words
  – Triggering and Controlling Devices
  – Future Directions

Interim Pause for Questions and Discussions

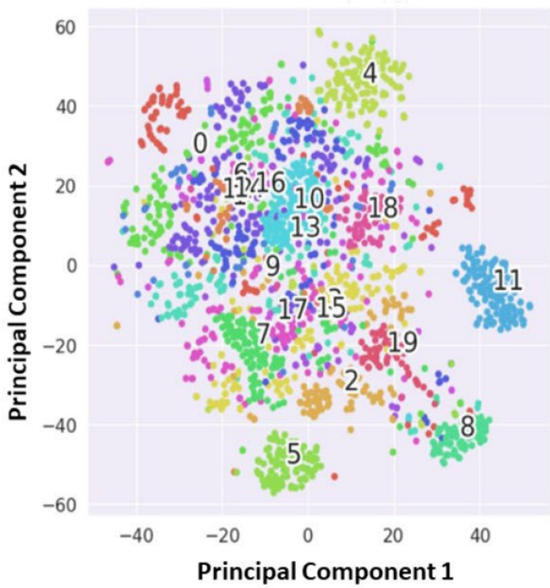
Let’s Keep Going ☺

Sevgi Z. Gurbuz (szgurbuz@ua.edu)

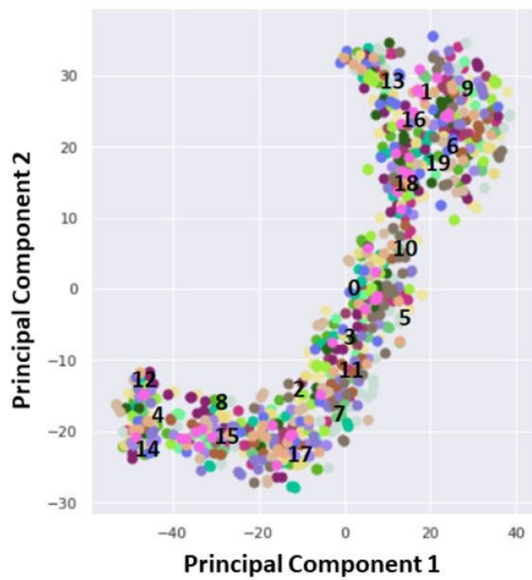
Imitation vs. Fluent

- Imitation Data:
  - Hearing person tries to sign after watching videos
- Fluent Signers:
  - Actually know ASL!

Imitation ≠ Fluent



Imitation



Fluent

Computers

can differentiate

imitation signing

from

fluent ASL signing

with **76%** accuracy

Can we effectively teach computers with imitation data?

NO

20 ASL Signs

Case 1:
- Teach a computer with imitation data
- Test with imitation data

→ **96%**

Case 2:
- Teach a computer with imitation data
- Test with fluent data

→ **24%**

Case 3:
- Teach a computer with fluent data
- Test with fluent data

→ **78%**

Alternative Approach

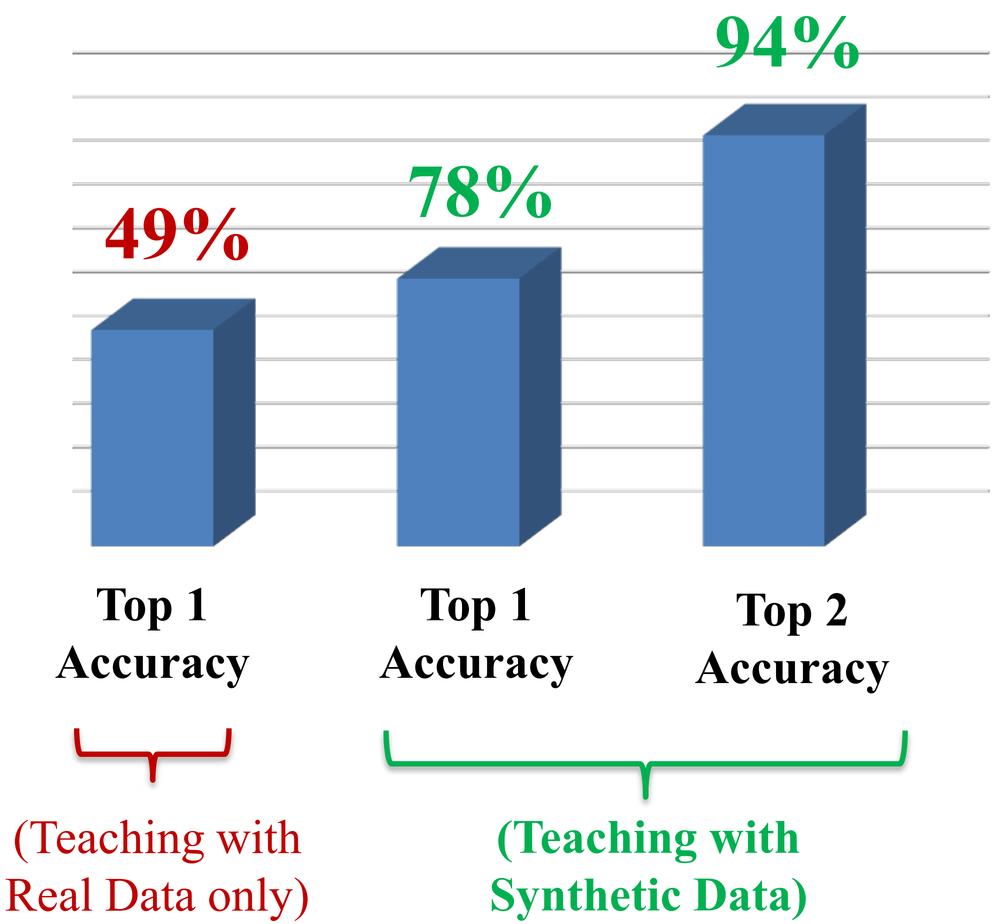
- Teach computers using synthetic data
  - "Generative Adversarial Networks" (GANs)
  - We have developed new types of "Physics-Aware GANs" (PhGAN) specific to RF ASL data

Small amount of real fluent ASL data → PhGAN → **LARGE** amount of synthetic ASL data

Sevgi Z. Gurbuz (szgurbuz@ua.edu)

Single-Word ASL Results

100 ASL Signs



How does this compare with video?

	Top-1	Top-5
RF (PhGAN)	**78%**	**93%**
Video (HRNet)	**60%**	**90%**

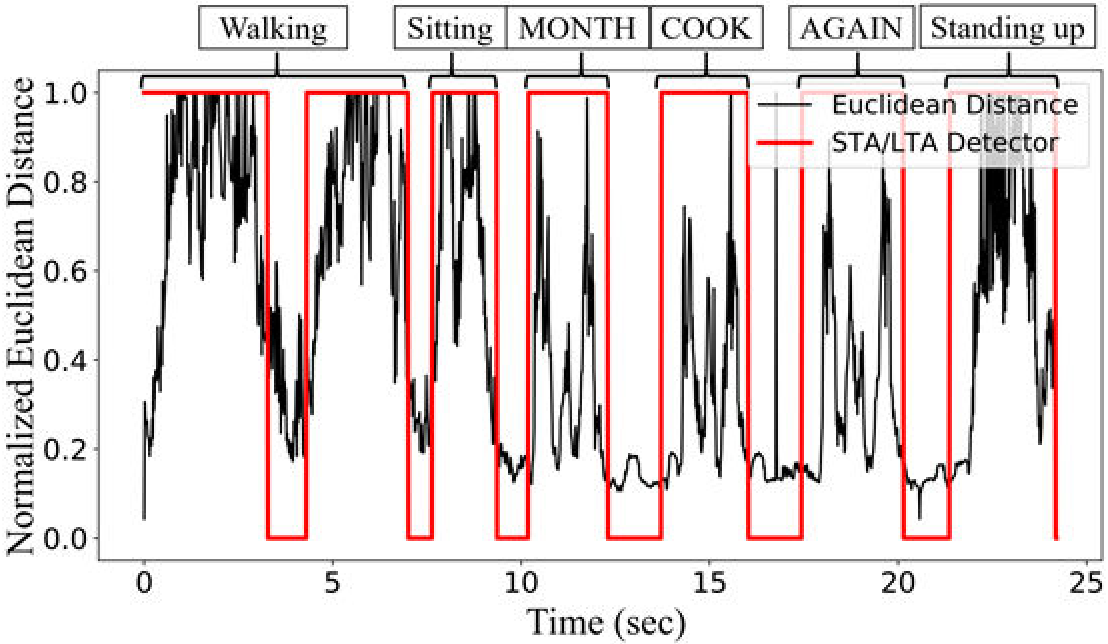
RF + Video

- RF and Video provide different information
  - RF: depth and azimuth distance + radial velocity (patterns of micro-Doppler)
  - Video: spatial information, expressions (weak in depth and temporal resolution)

Sevgi Z. Gurbuz (szgurbuz@ua.edu)

Triggering and Controlling Devices

- We can differentiate ASL from daily activity
  - Detect, segment, classify



(b)

What should trigger a device?

Initially explored signs more easily and consistently replicable

Words	Two-Handed	# of Strokes	Major Location	Movement
TIRED	✓	3	Body	Curved
BOOK	✓	3	Neutral	Curved
SLEEP	✓	4	Head	Straight
EVENING	✓	5+	Hand	Straight
READY	✓	4	Neutral	Straight
HOT	X	3	Head	Curved
MONTH	✓	3	Hand	Straight
COOK	✓	5+	Hand	Other
AGAIN	✓	4	Hand	Curved
SUMMON	✓	3	Hand	Back-and-Forth
MAYBE	✓	5+	Neutral	Straight
NIGHT	✓	3	Hand	Straight
SOMETHING	X	4	Neutral	Circular
TEACHER	✓	4	Head	Straight
TEACH	✓	3	Head	Straight

Sevgi Z. Gurbuz (szgurbuz@ua.edu)

Trigger Sign Design

- Trigger sign design considerations
  - Cultural
  - Not easily confused with daily activity
  - Not easily confused with other conversation
  - For RF: more movement the better

NEED COMMUNITY INPUT / COLLABORATION

Work in Progress

- Interactive Gaming
- Continuous Discourse:
  - Properties
  - Semantic segmentation
- ASL-Triggered/Controlled Interface Prototype
- Sensor Fusion (RF+Video, RF+Lidar, RF+?)

Community Outreach

- STEM Education with the Alabama Institute of the Deaf and Blind and NTID Regional STEM Center

Thank you!

My research was supported in part by the
National Science Foundation Cyber-Physical Systems Program
Award #: 1932547 / 1931861

Sevgi Zubeyde Gurbuz
Assistant Professor
Computationally Intelligence for Radar (CI4R) Lab
Dept. of Electrical and Computer Engineering
University of Alabama – Tuscaloosa
szgurbuz@ua.edu

Visit us at: http://ci4r.ua.edu