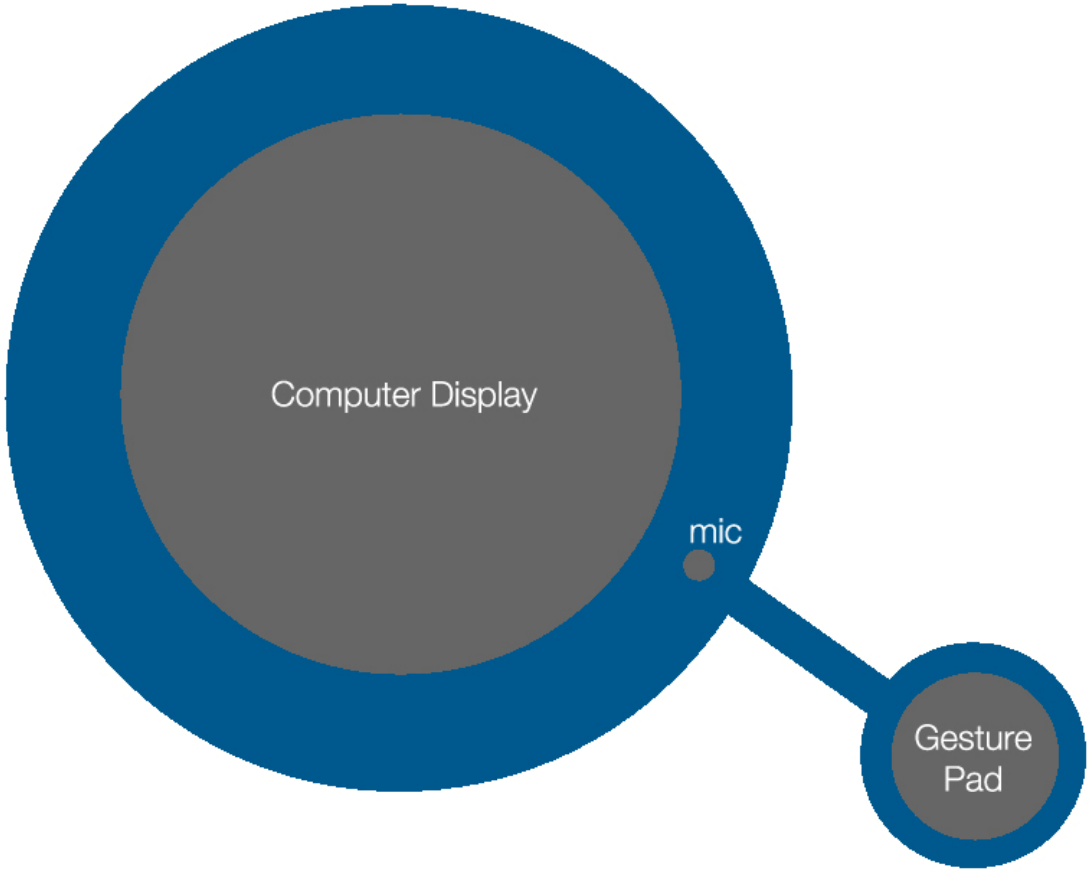
wallcomputing

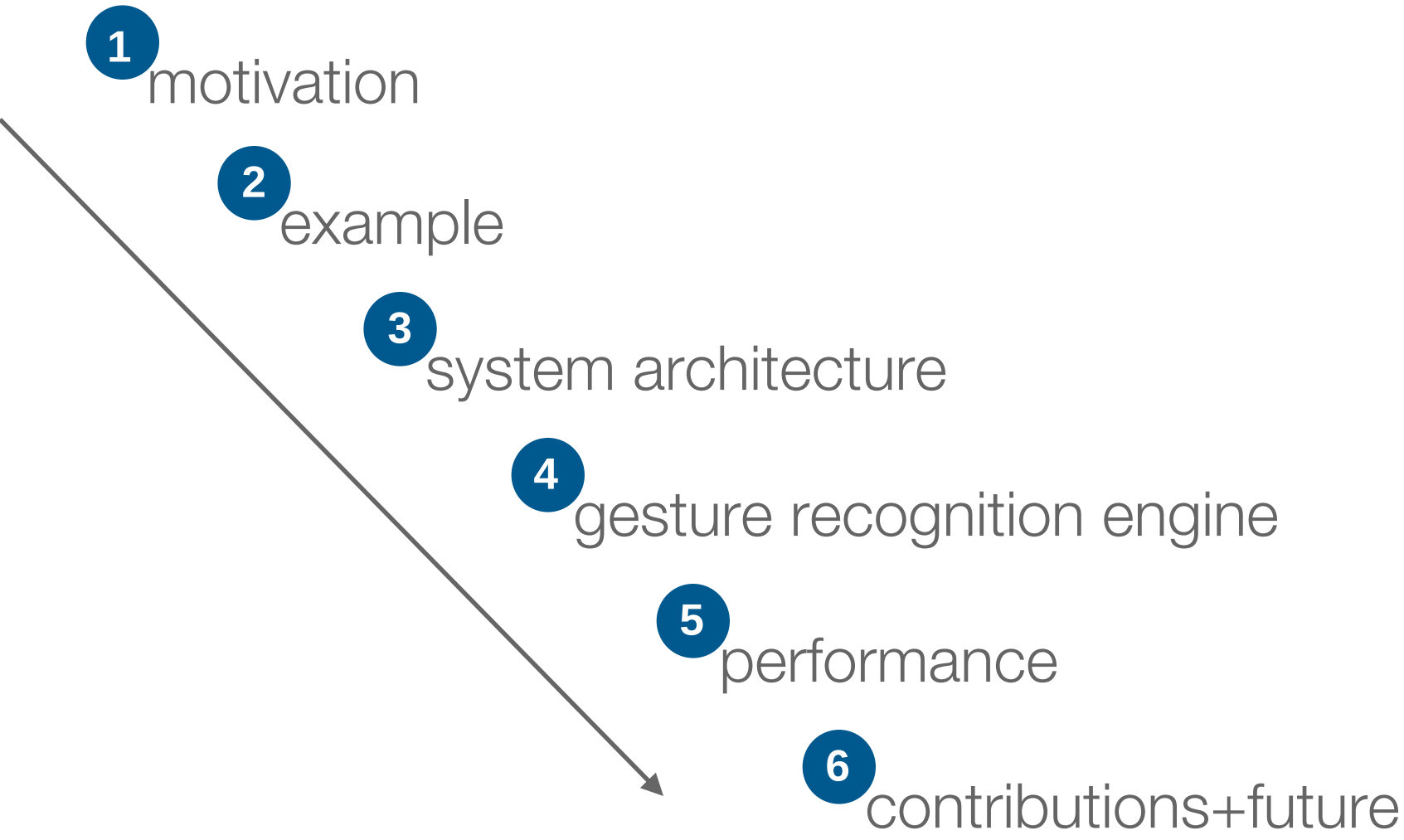
IMMERSIVE MULTIMODAL INFORMATION MACHINE



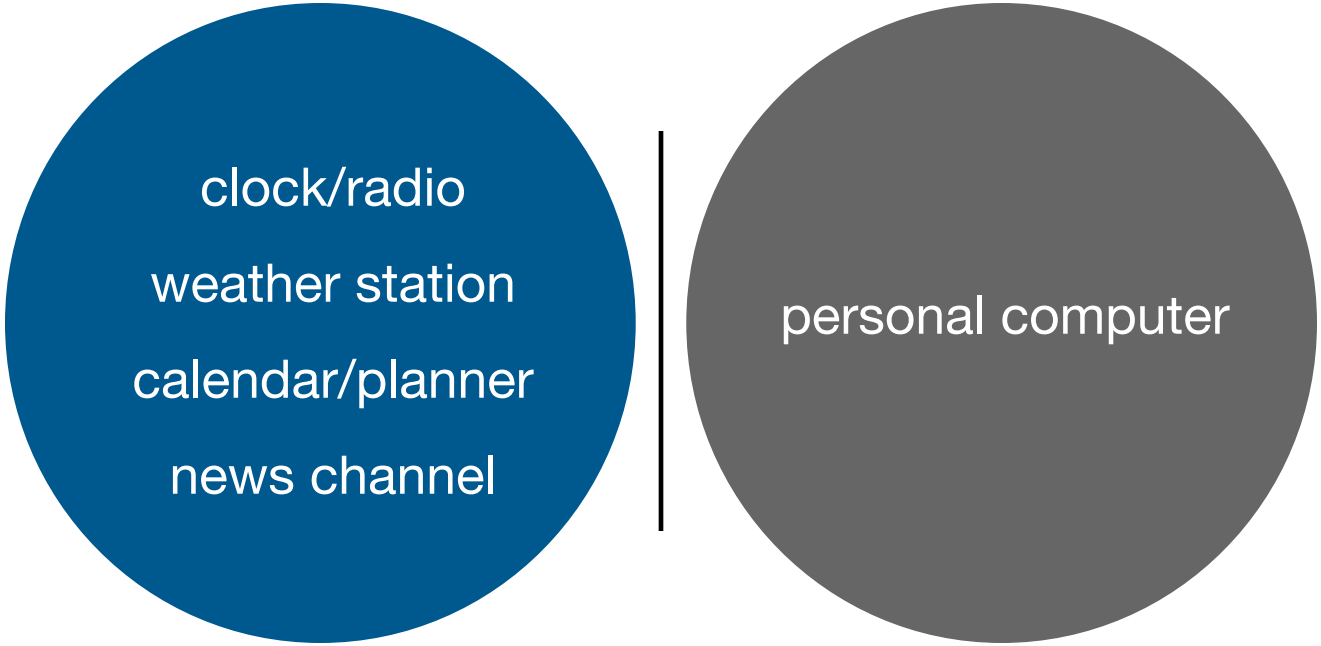
6.835 Multimodal Interfaces Final Presentation

Zack Anderson

Contents


1
motivation
2
example
3
system architecture
4
gesture recognition engine
5
performance
6
contributions+future

Motivation



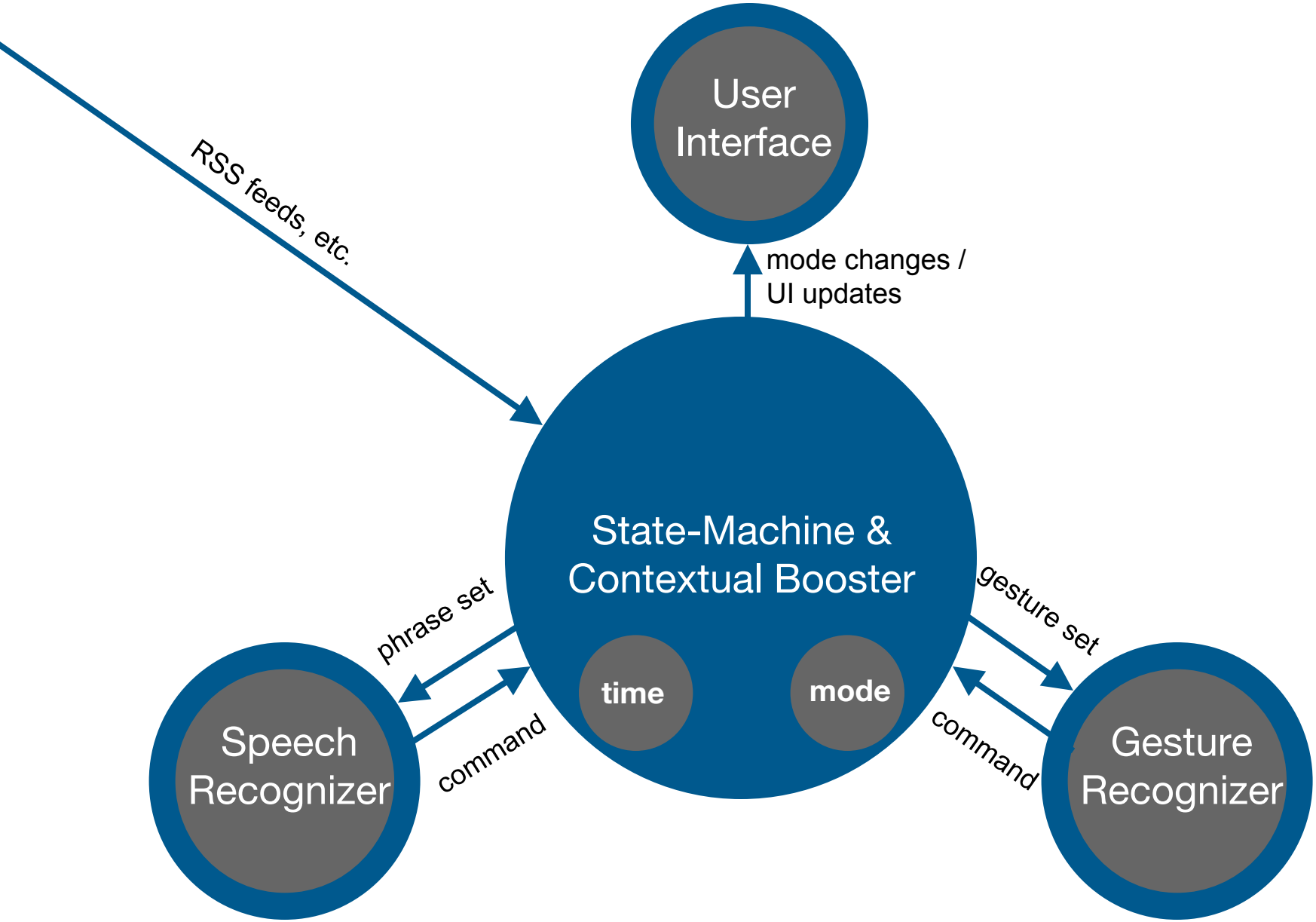
KEY OBSERVATION: Disconnect between two classes of devices. Single-purpose home devices are easy and efficient. PCs offer extensible interfaces to data.

CHALLENGE: Design an easy and efficient interface to access time-sensitive data.

Example

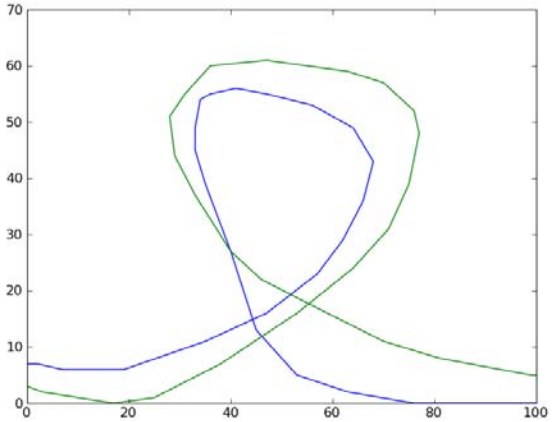
- Live demo

System Architecture



Gesture Recognition Engine

- Nearest neighbors classification



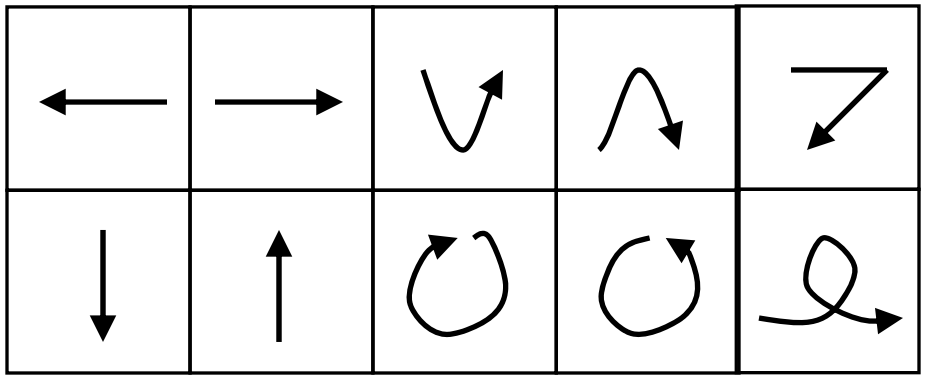
Gesture Recognition Engine

- Nearest neighbors classification
- Weighted Euclidian distance measures

$$\begin{bmatrix} a & b & c & d \end{bmatrix} \cdot \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta x\_dot \\ \Delta y\_dot \end{bmatrix}$$

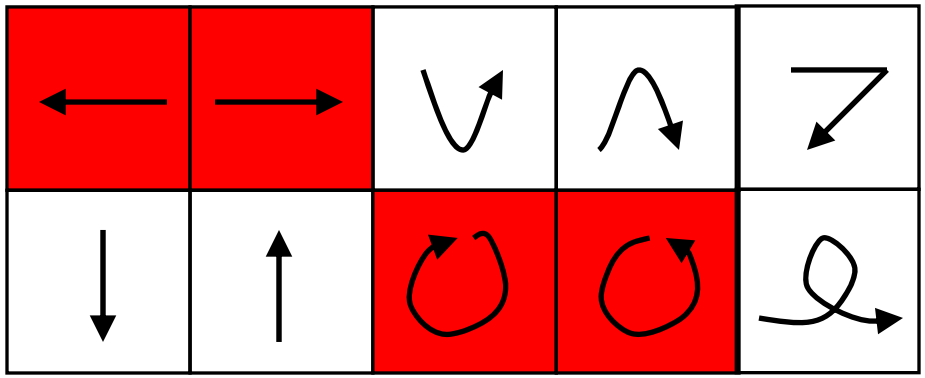
Gesture Recognition Engine

- Nearest neighbors classification
- Weighted Euclidian distance measures
- Dynamically-restricted gesture set for better performance



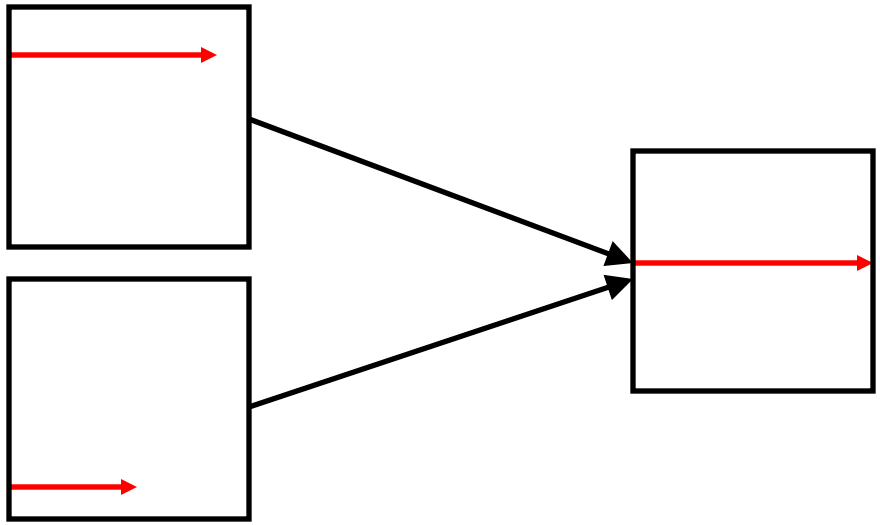
Gesture Recognition Engine

- Nearest neighbors classification
- Weighted Euclidian distance measures
- Dynamically-restricted gesture set for better performance

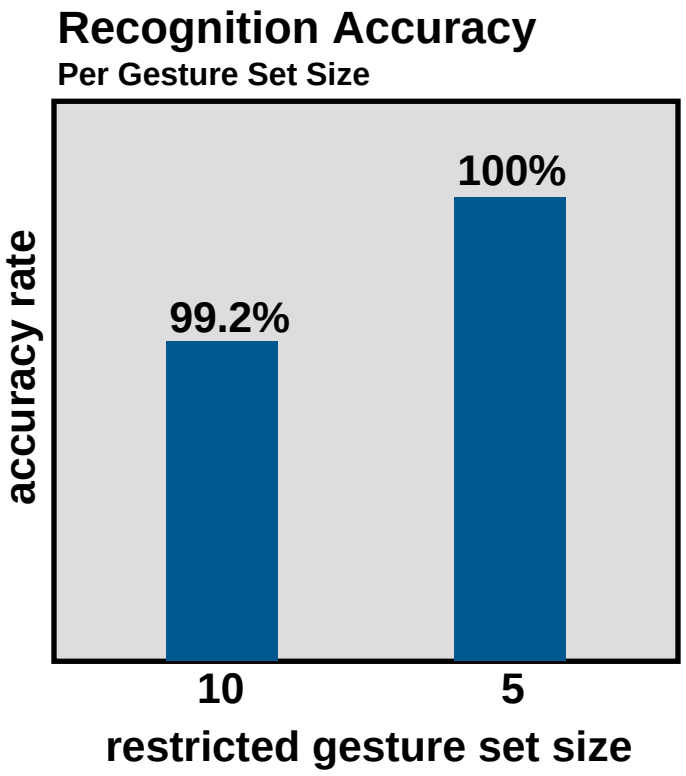


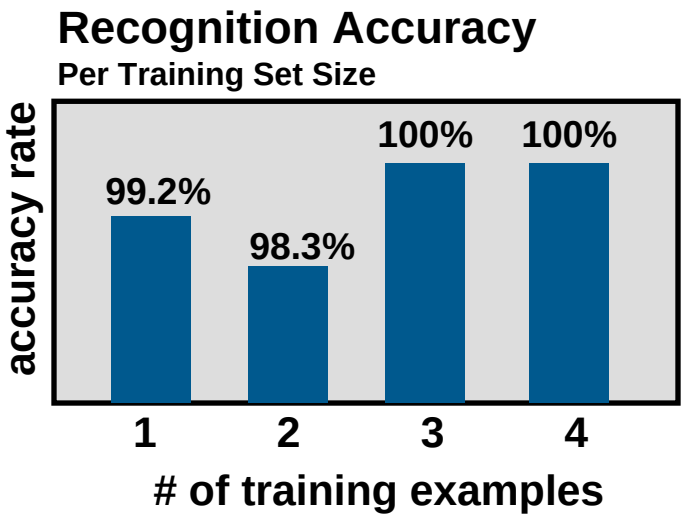
Gesture Recognition Engine

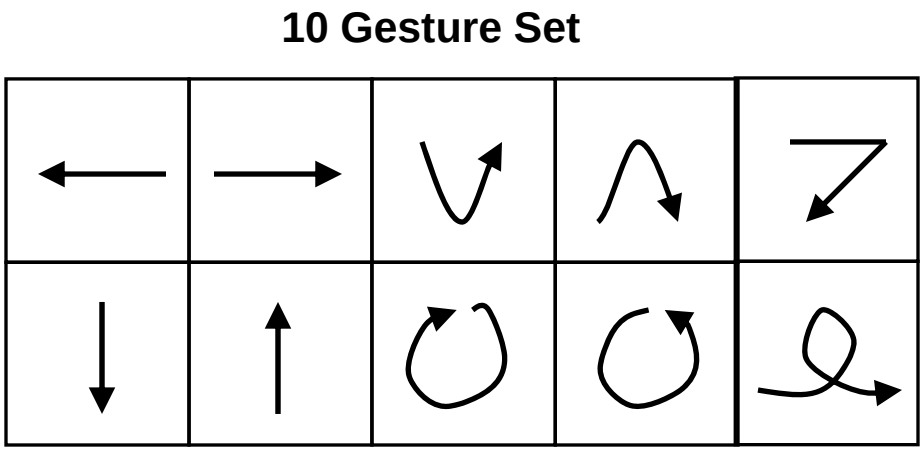
- Nearest neighbors classification
- Weighted Euclidian distance measures
- Dynamically-restricted gesture set for better performance
- Transforming-normalization algorithm to make temporally-similar gestures look the same



Performance: Gesture Engine

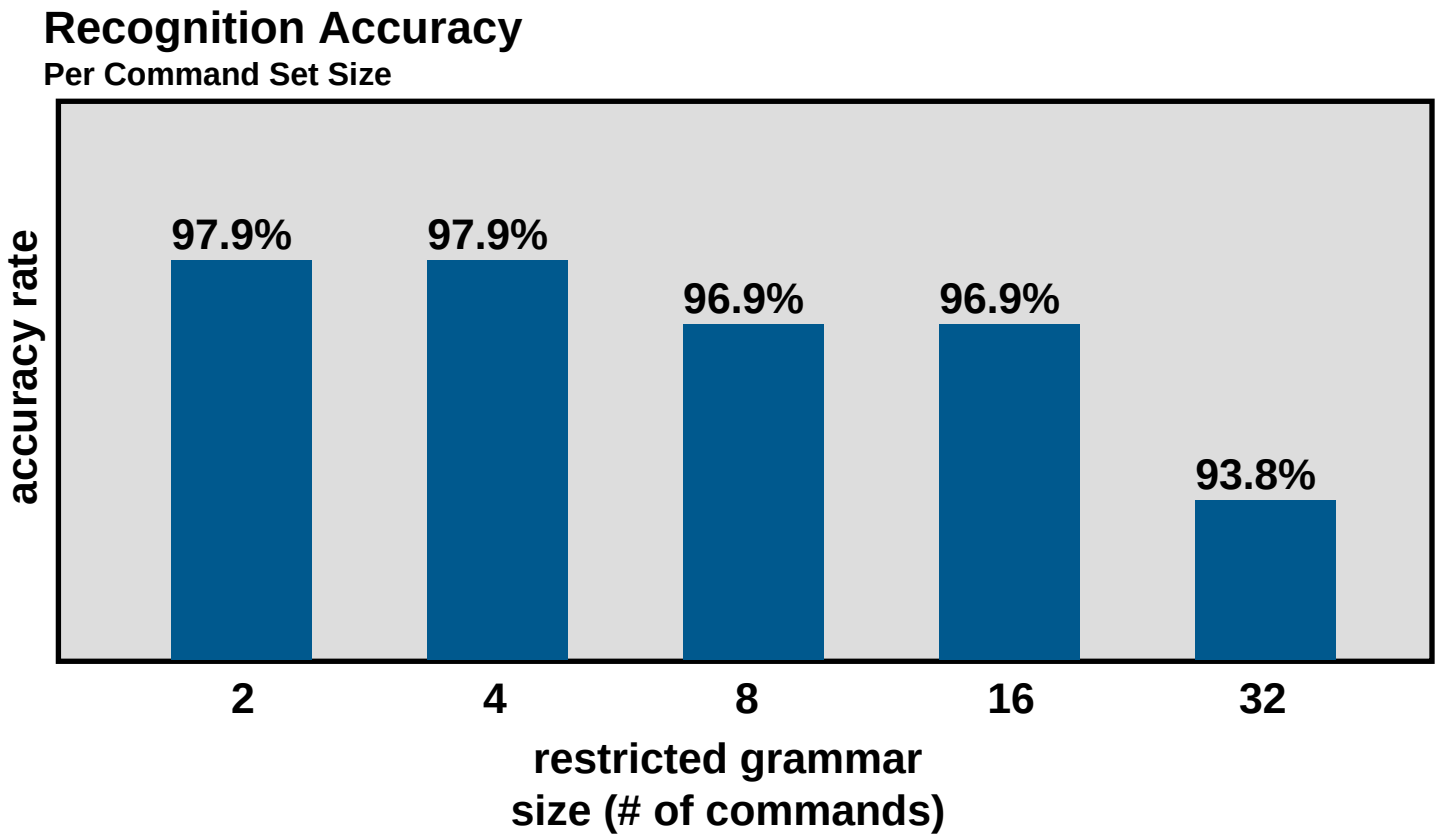






*Tests conducted on a total sample size of 300 gestures of 10 types input by 6 different people. Left chart used 1 training example per gesture.

Performance: Speech Engine



*Tests conducted using a custom python wrapper of the Microsoft Speech SDK. Grammars are dynamically-restricted. Microsoft Speech engine was trained before testing. Where possible, restricted grammars were kept within a domain. Non-recognitions are considered false recognitions.

Performance: **Usability**

“Gestures seem to **flow** with the UI, making the system very intuitive.”

“**Response time** needs to be faster to make the system seem seamless.”

“**Recognition accuracy** is surprisingly good, making the wallcomputer efficient, simple to learn, and pleasing to use.”

“System inputs are immersive and natural. It would be nice if the UI were more **tactile**.”

Contributions / Future

- **Designed** an accurate (>99%) gesture recognition system based on optimizations of a nearest-neighbors algorithm
- **Demonstrated** that multimodal, contextually-restricted UIs provide superior performance
- **Presented** a new paradigm of computer interaction that verges between ambient and full-PC capability
- **Built** a functional “wallcomputer”

future

- Add more modes (i.e. schedule, automation system control, stock quotes, etc.), integrate 3rd party APIs (i.e. gcalendar)
- Add more control modalities for greater user efficiency
- Incorporate tactile/auditory feedback