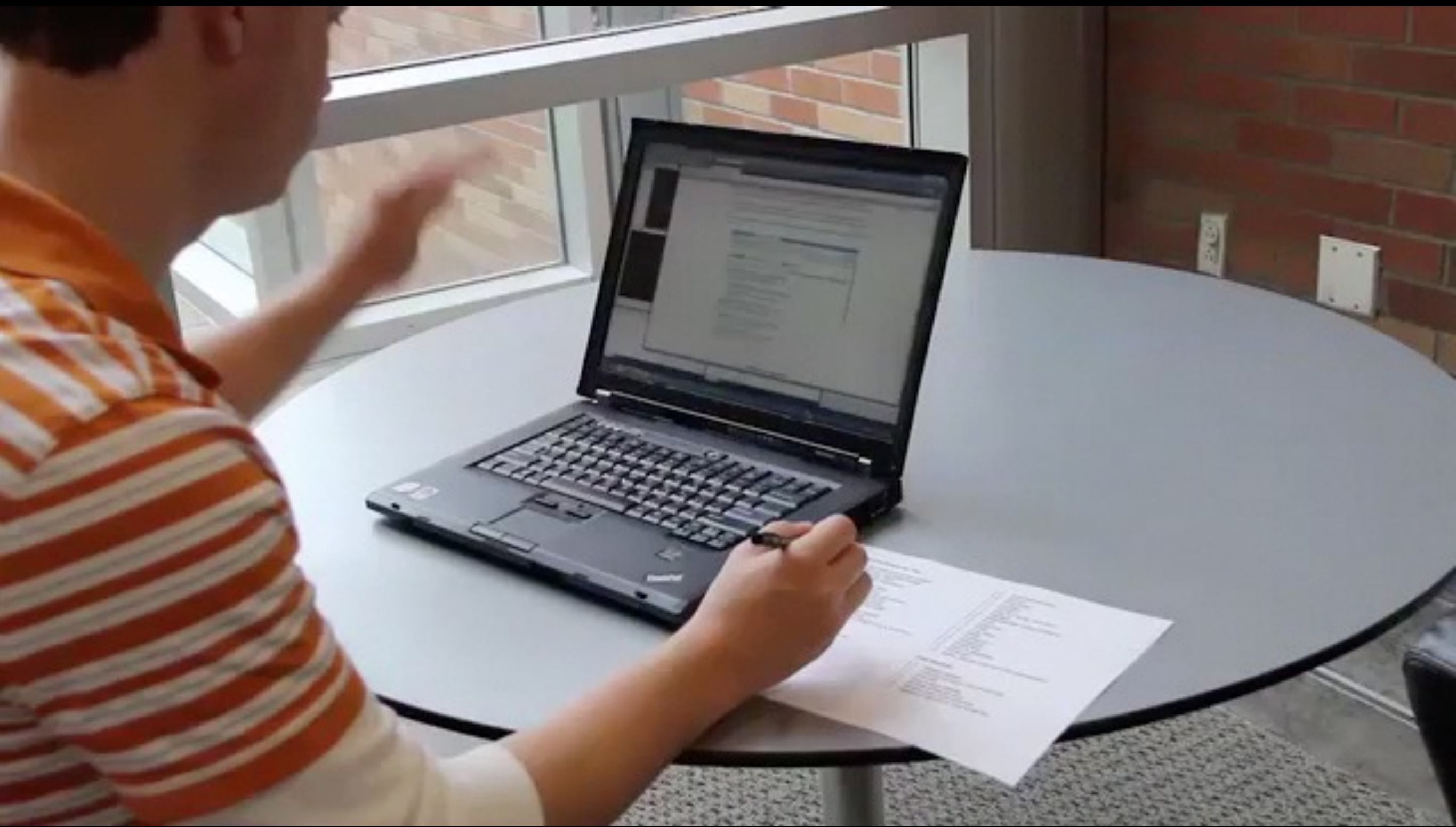
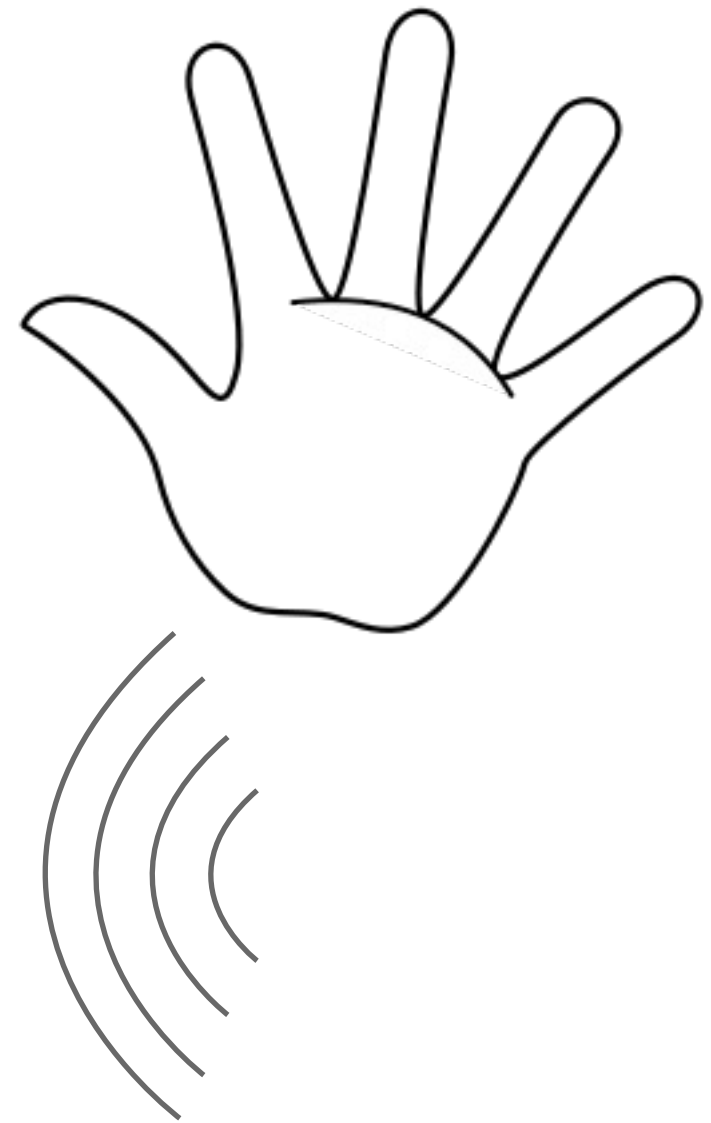


Examples of Sound as a Signal in Research

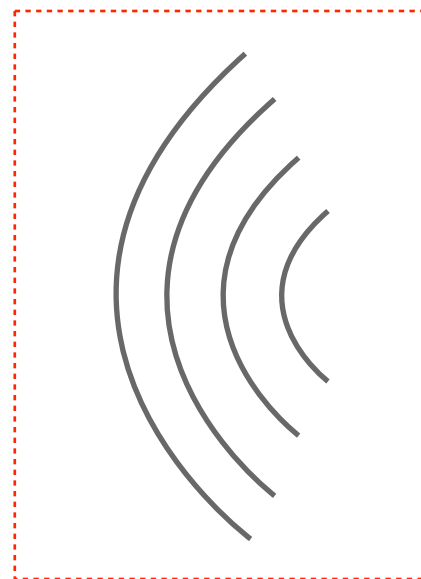
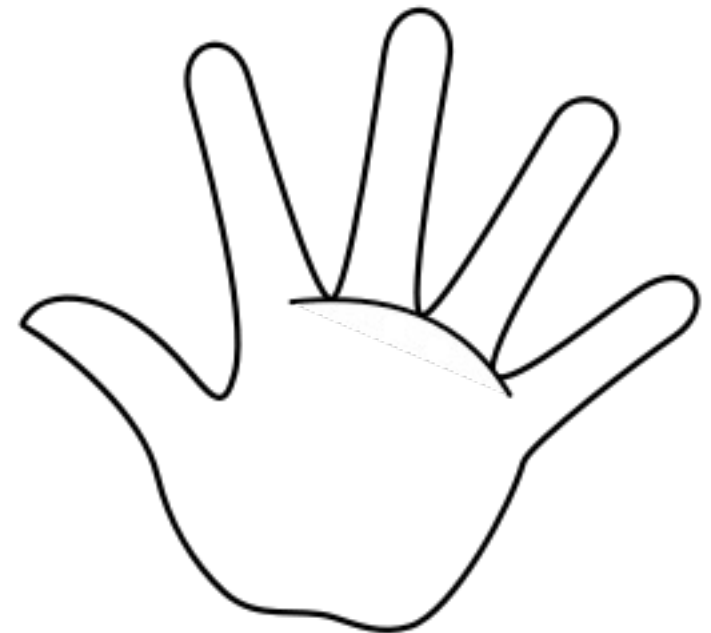




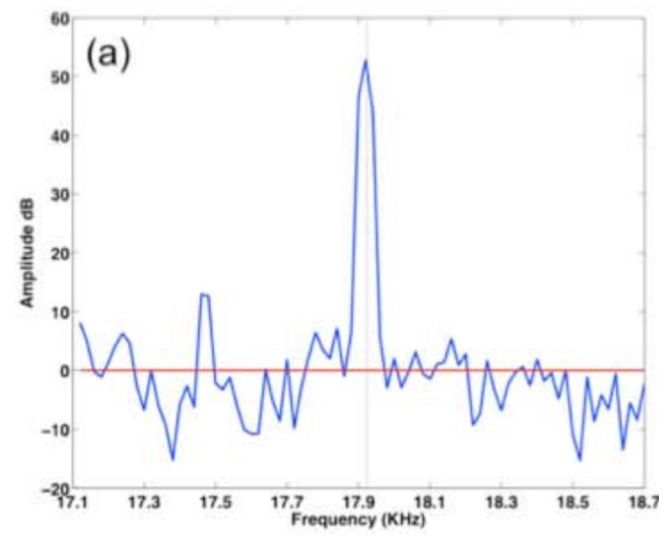
Speaker produces and **ultrasonic tone** of 18-22 kHz



Speaker produces and **ultrasonic tone** of 18-22 kHz

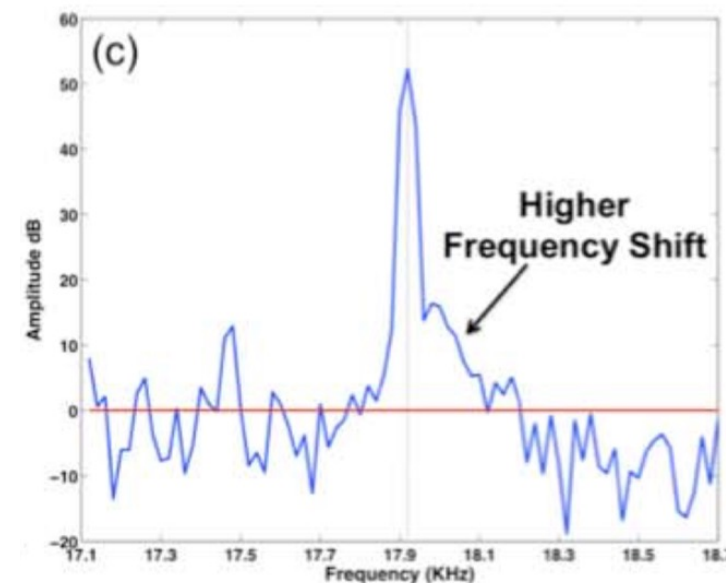


reflected
sound
waves



Features Used

- 2048-point FFT
 - 1024 valid points
- 0 to 22 kHz
- **Then completely heuristics based**



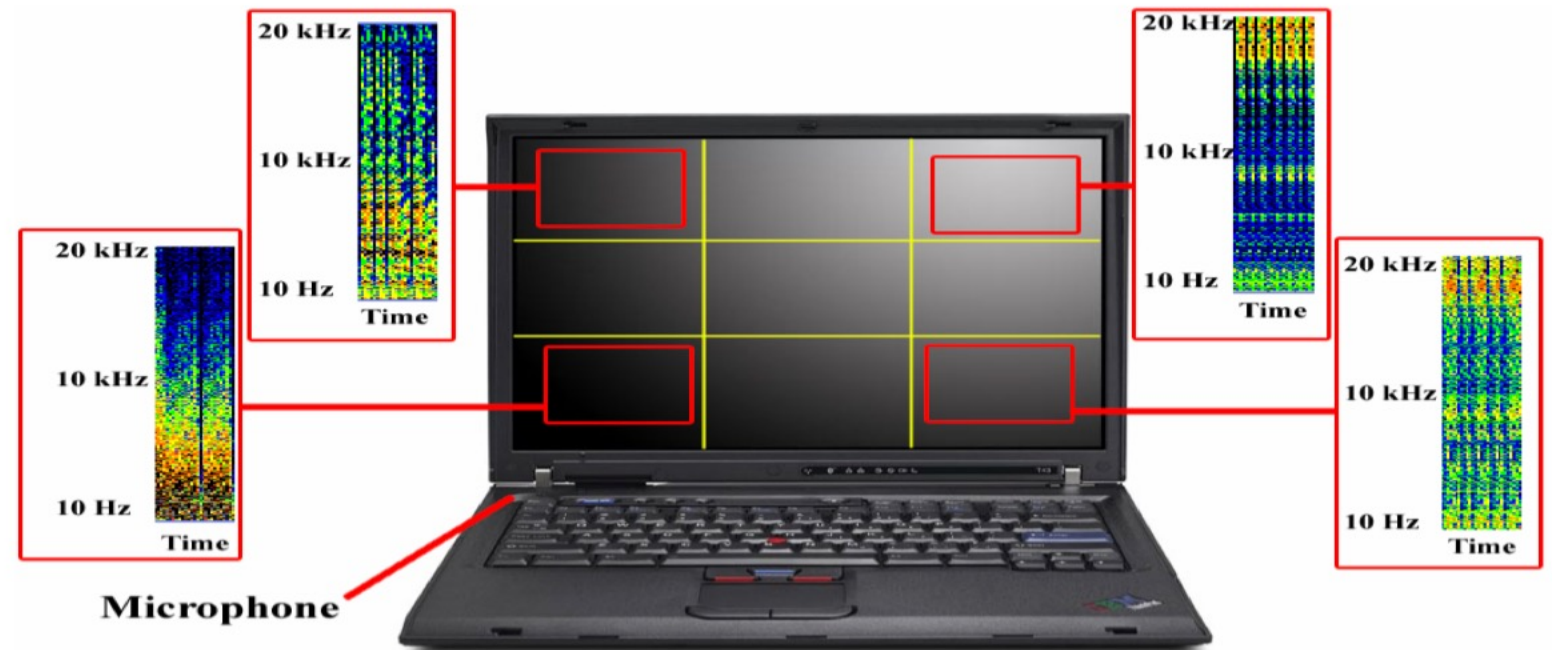
BLUI: Low-cost Localized Blowable User Interfaces

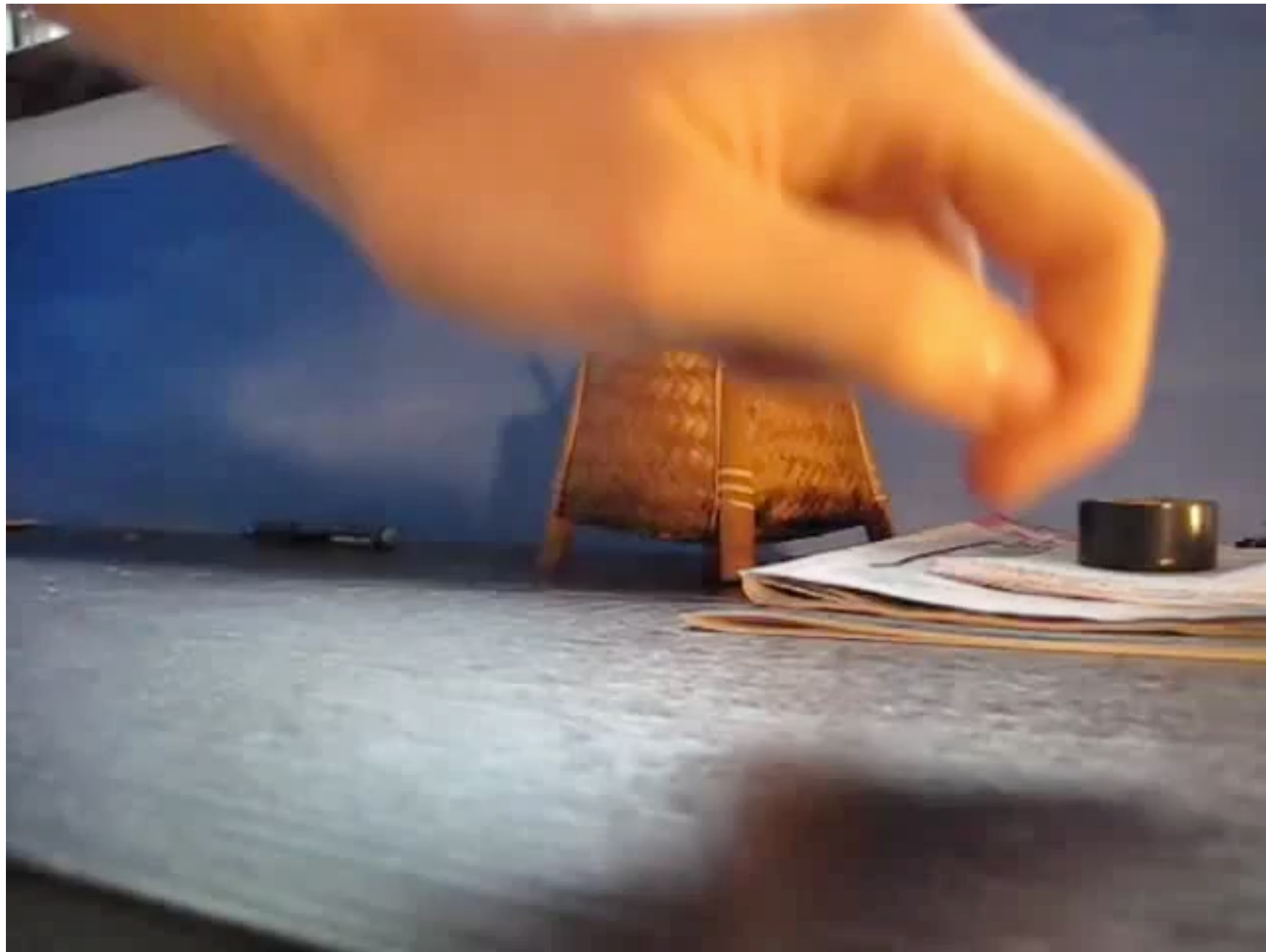
Shwetak N. Patel and Gregory D. Abowd

Georgia Institute of Technology

Features Used

- 0 to 22 kHz
- 2048-point FFT
 - 1024 valid points
- **Used the 1024 magnitude values as features (binned)**
- **PCA to reduce the feature space**





Scratch Input. Harrison and Hudson 2008

Features Used

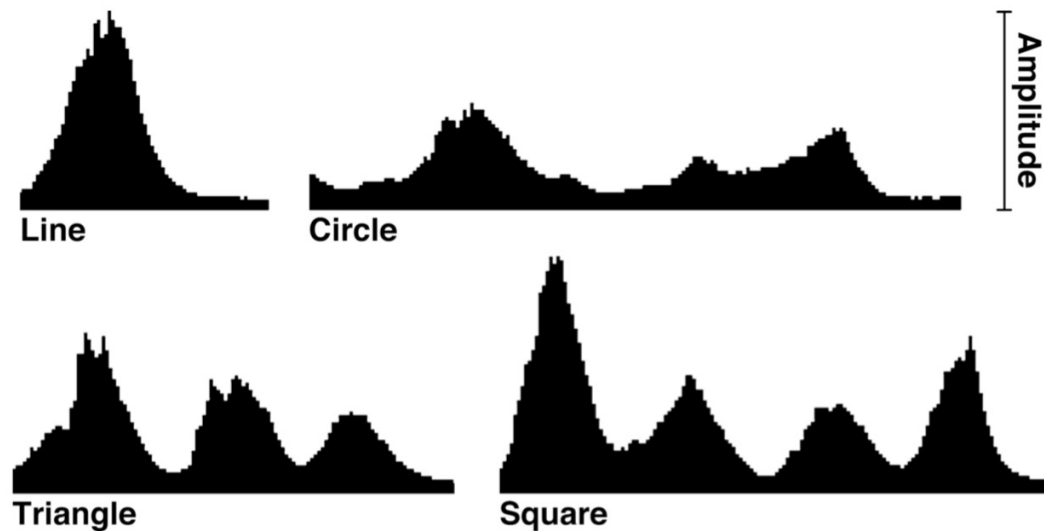


Figure 2. Amplitude profiles for different gestures.

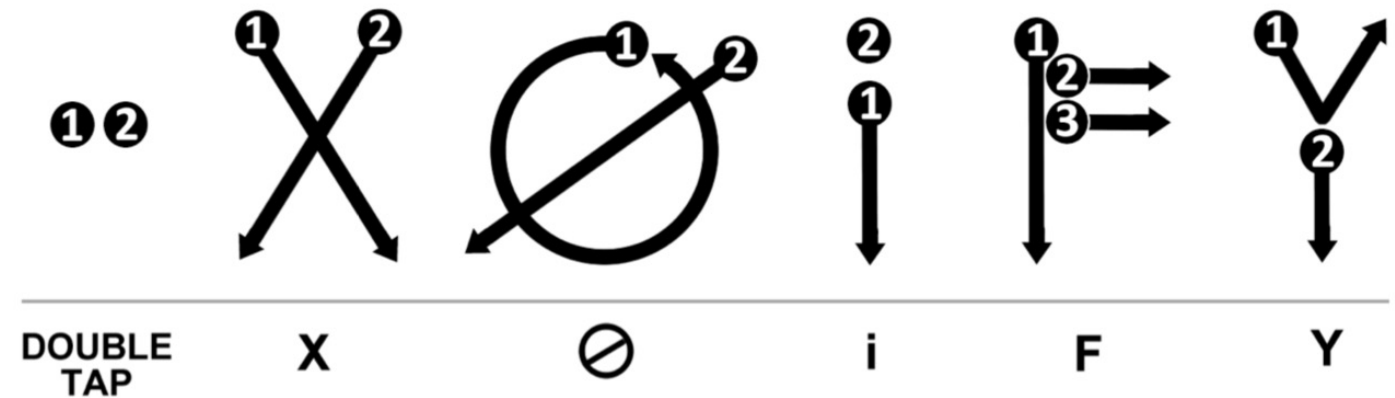
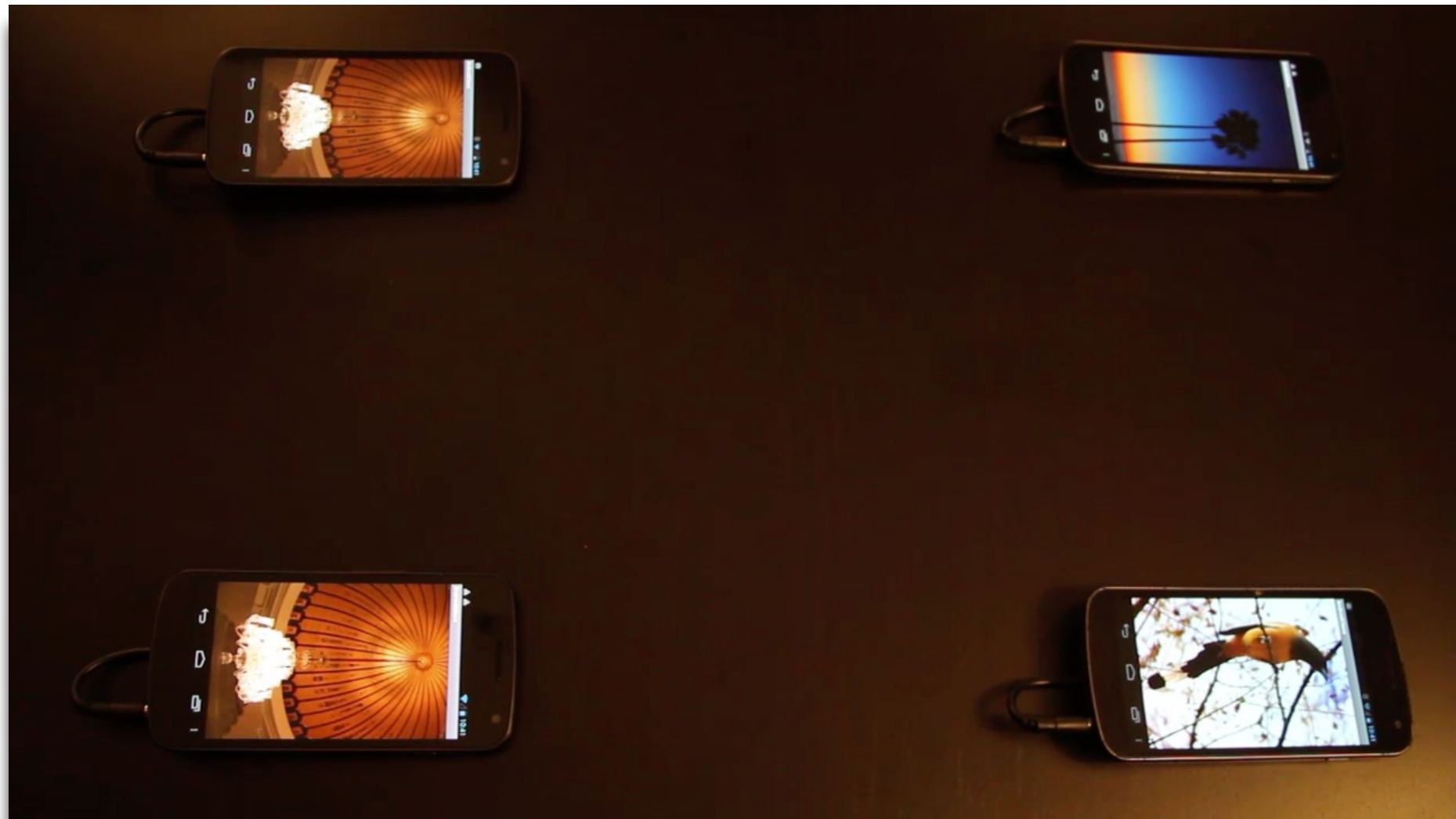


Figure 3. Distinct multi-part gestures composed of taps, lines and circles.

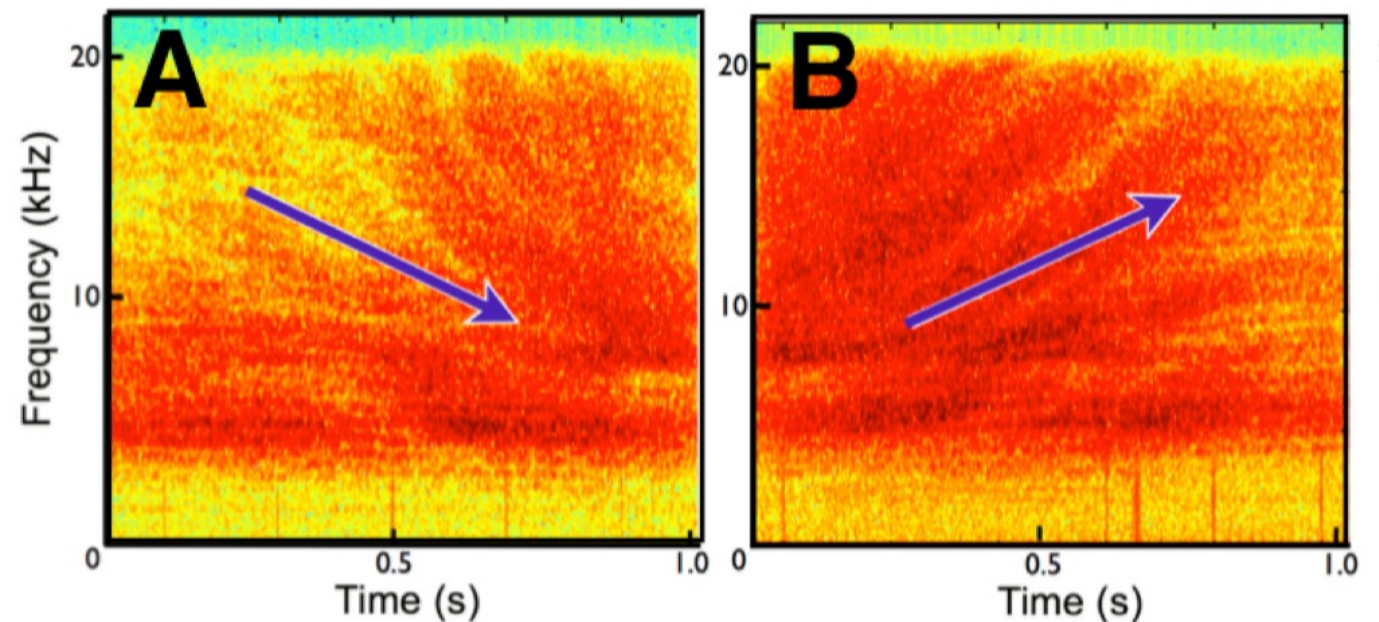




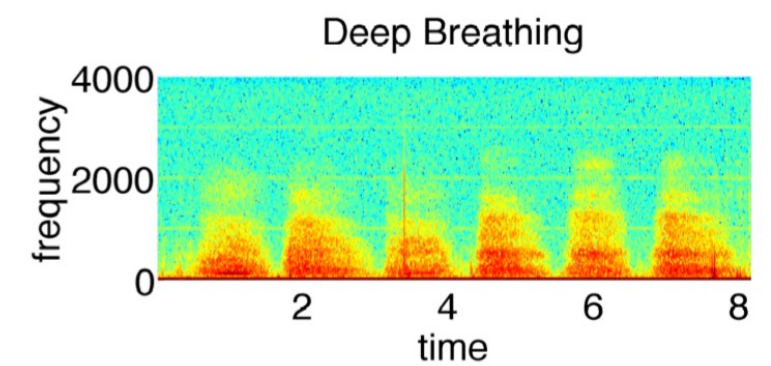
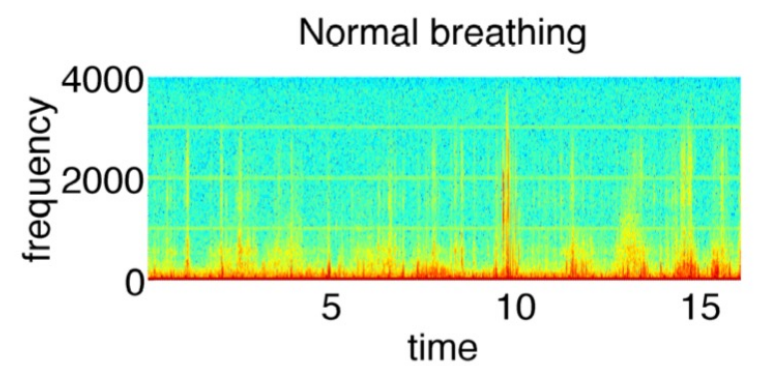
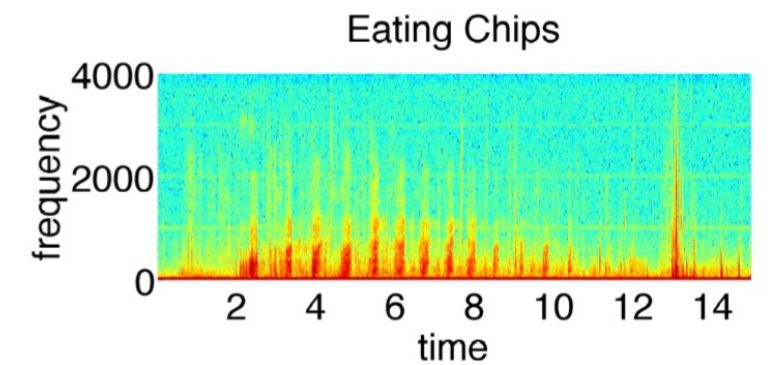
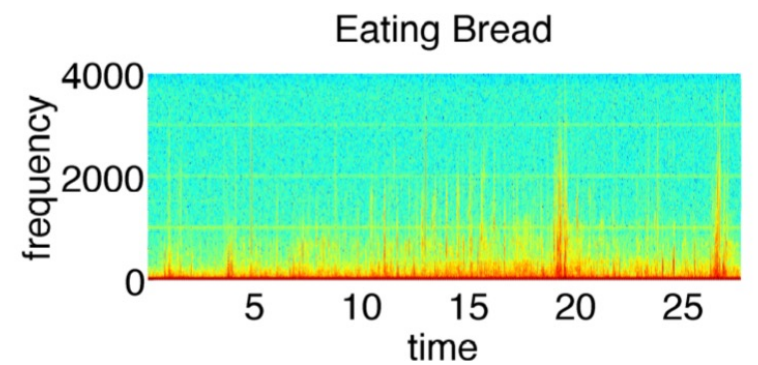
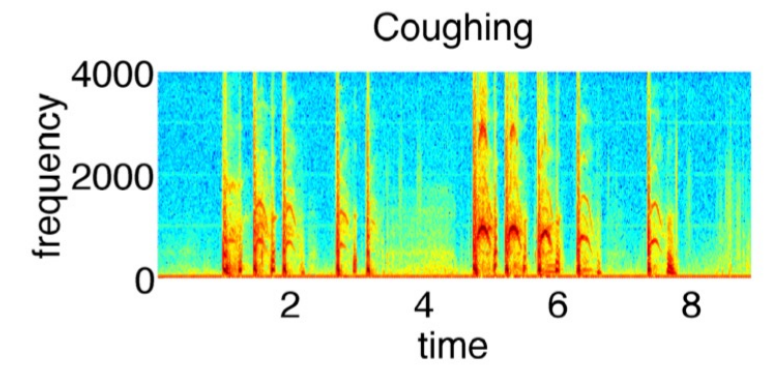
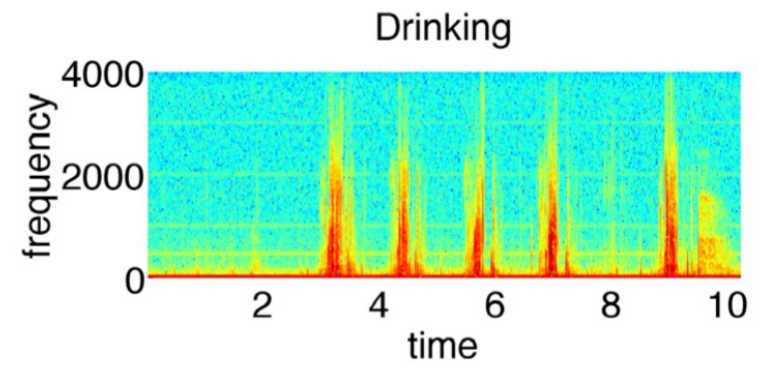
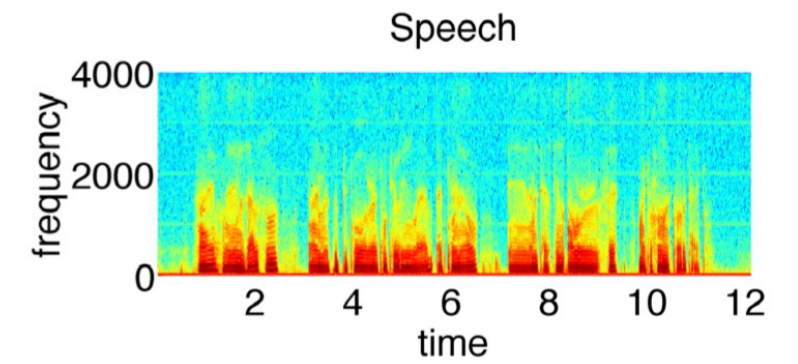
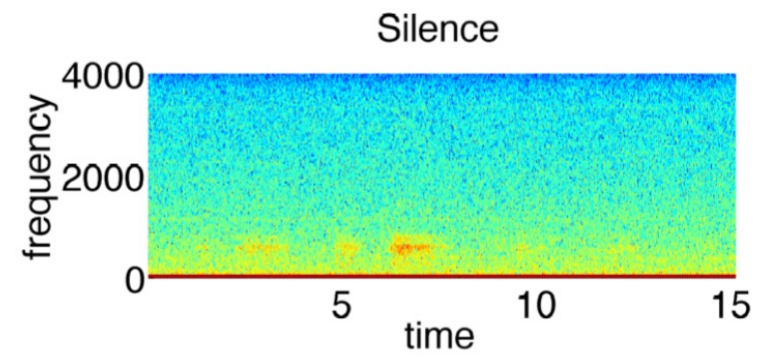
Surface Link. Goel et al. 2014

Features Used

- 1024-point FFT
 - 512 valid points
- **[6 x 10] bin of spectrogram**
- **[10] time bins for 5-15 kHz range of spectrogram**
- **Energy in the first half**
- **Energy in the second half**



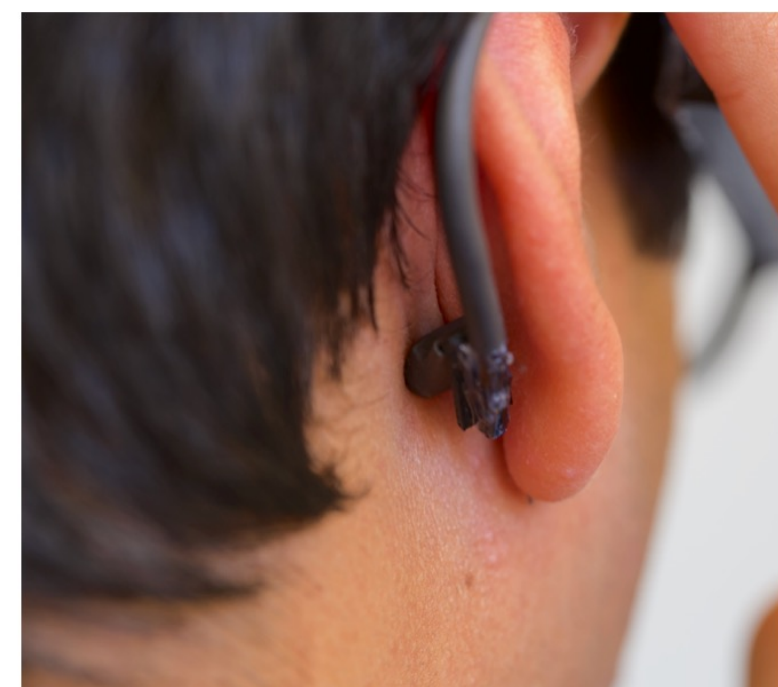
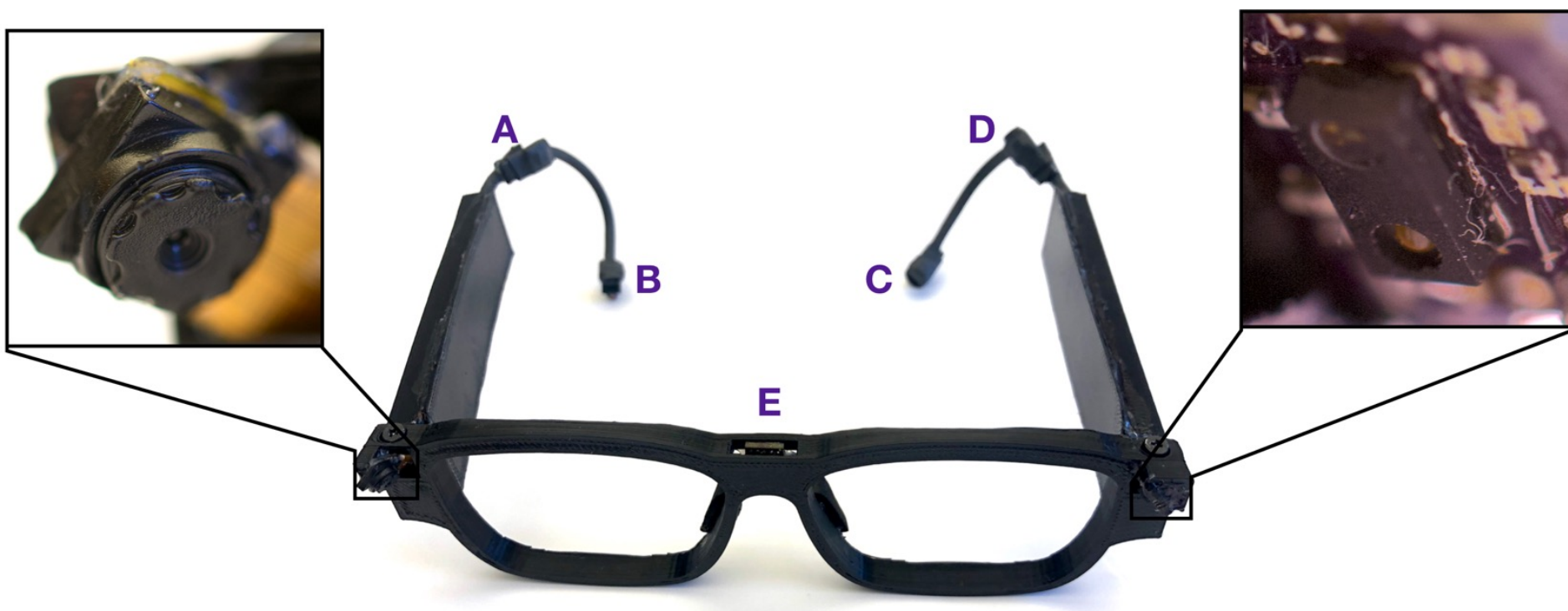
**A quick look at
some code**

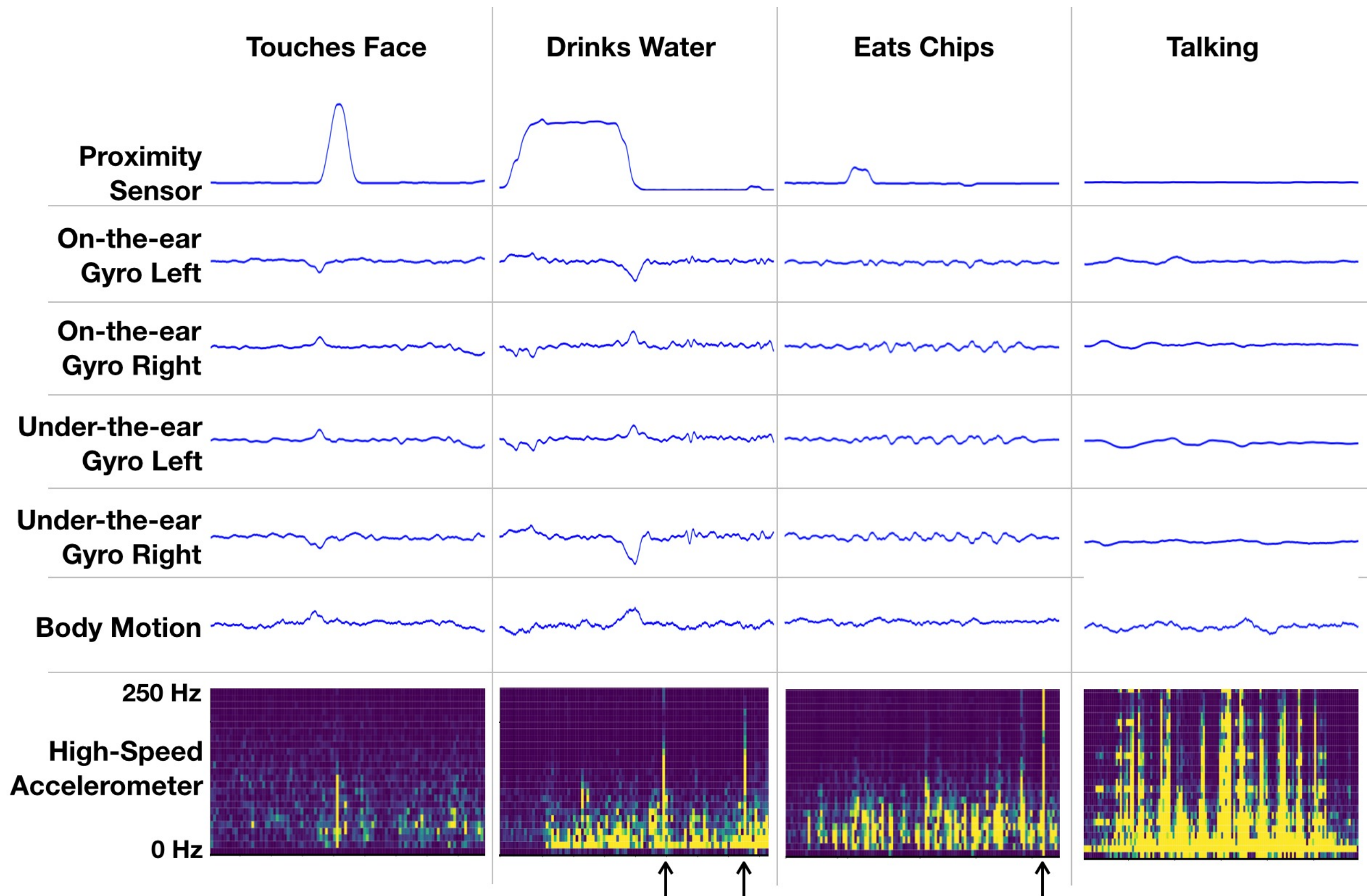


Features Used

| Group | Frame level descriptors | Acronym |
|---------------|--|---------------|
| Energy | log power of 8 subbands | LogSubband[i] |
| | Total RMS Energy | RMSenergy |
| Spectral | Spectral Centroid | SpectCent |
| | Spectral Flux | SpectFlux |
| | Spectral Variance | SpectVar |
| | Spectral Skewness | SpectSkew |
| | Spectral Kurtosis | SpectKurt |
| | Spectral Slope | SpectSlope |
| | Spectral Rolloff 25% | SpectROff25 |
| | Spectral Rolloff 50% | SpectROff50 |
| | Spectral Rolloff 75% | SpectROff75 |
| | Spectral Rolloff 90% | SpectROff90 |
| Crossing Rate | Zero Crossing Rate | ZCR |
| MFCC | 12 Mel Frequency Cepstral Coefficients | mfcc[i] |

| Type | Statistical Functions | Acronym |
|----------------|-------------------------|---------------|
| Extremes | Minimum | min |
| | Maximum | max |
| Average | Mean | mean |
| | Root Mean Square | RMS |
| | Median | median |
| Quartiles | 1st and 3rd Quartile | qrtl1, qrtl3 |
| | Interquartile Range | iql |
| Moments | Standard Deviation | std |
| | Skewness | skew |
| | Kurtosis | kurt |
| Peaks | Number of peaks | numOfPeaks |
| | Mean Distance of Peaks | meanDistPeaks |
| | Mean Amplitude of Peaks | meanAmpPeaks |
| Rate of Change | Mean Crossing Rate | mcr |
| Shape | Linear Regression Slope | slope |







Features Used

- 3 Axes data
- Combined FFT of the 3 axes by removing DC component and taking Max of 3 axes
- Average 20 consecutive FFT frames
 - mean, sum, min, max frequencies, standard deviation
 - ratios of different frequency regions
 - number of peaks

Carnegie
Mellon
University

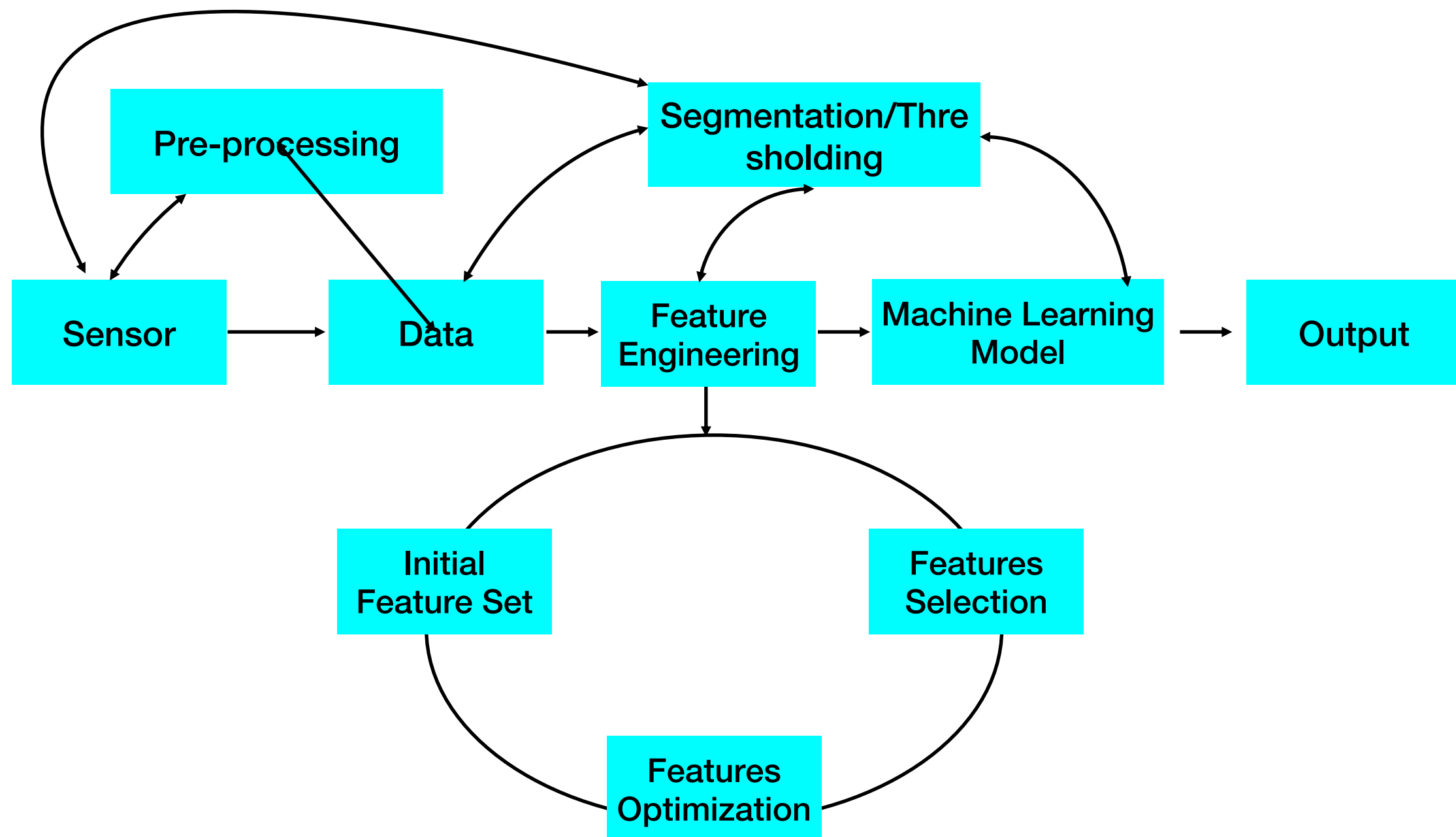
www.figlab.com



Assignment 2

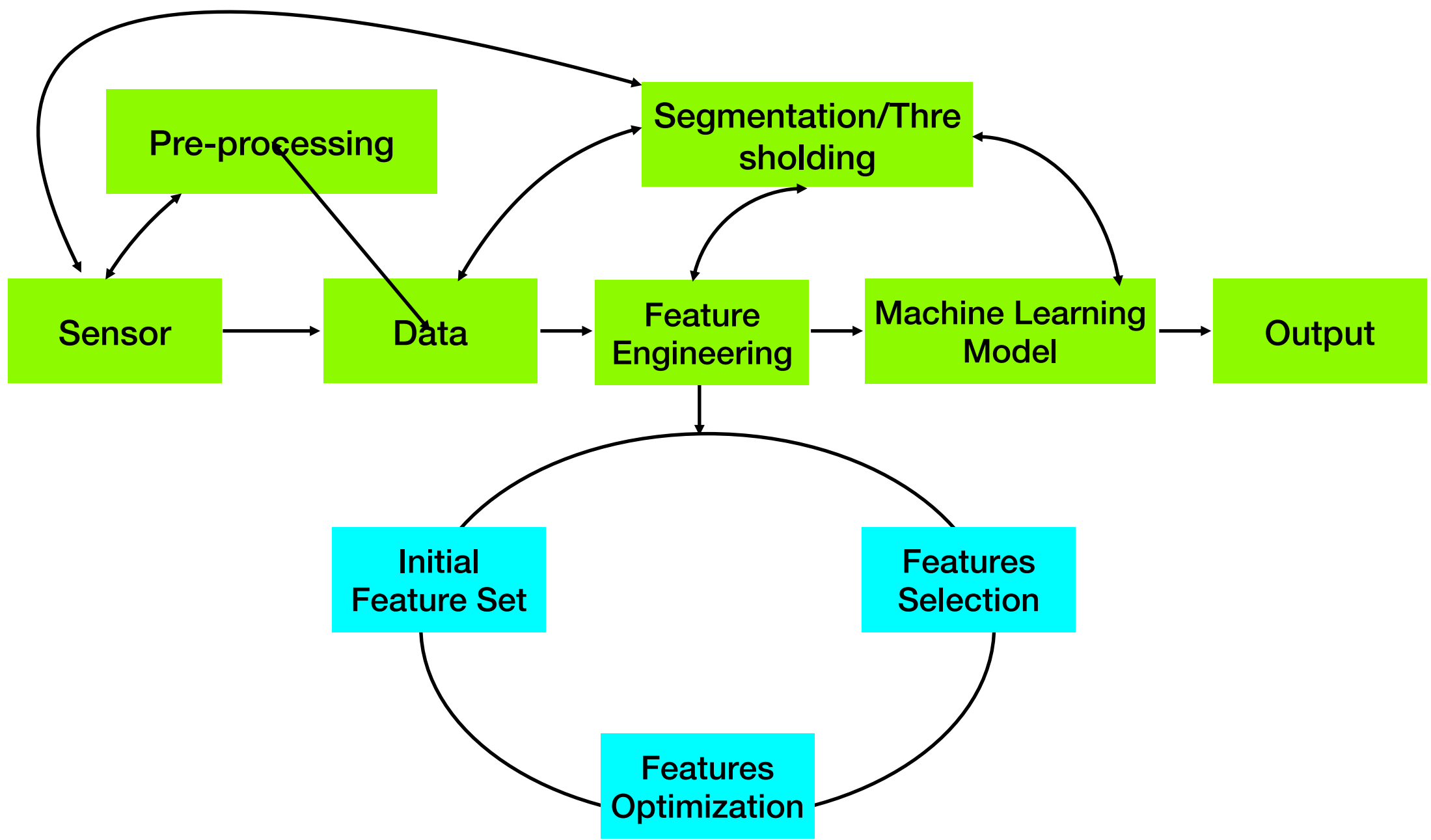
- Sound-based Activity/Event Detection
- Collect sound data for 5 classes (20 samples): **(5% grade)**
 - Microwave (run for approx. 30 seconds and record the whole session including opening and closing the door, and beeps)
 - Blender (run for approximately 30 seconds)
 - Fire-alarm or any kind of siren (can be from videos or apps)
 - Vacuum cleaner (run for approximately 30 seconds)
 - Music (each instance for approximately 30 seconds)

- Record blanks too (of approx. 30 seconds in length) to develop segmentation



Pipeline Components

- Signal conditioning **(5%)**
- Two approaches to feature engineering:
 - Binning **(15%)**
 - Domain-specific features **(15%)**
- Window and non-window approach **(15% each)**
 - Window length is up to you
 - In window approach, combine data/decisions from multiple windows
 - In non-window approach, assign one decision per “recording”
- 10-fold cross-validation
- Aim for above 80% performance in at least 3 cases, and above 90% in at least 1. *i.e.*, it is fine if the classification accuracy is below 80% for one of the case



Write-Up

- 30% value
- Describe:
 - Data collection process
 - Rationale for features
 - Graph results for different conditions