WriteHear

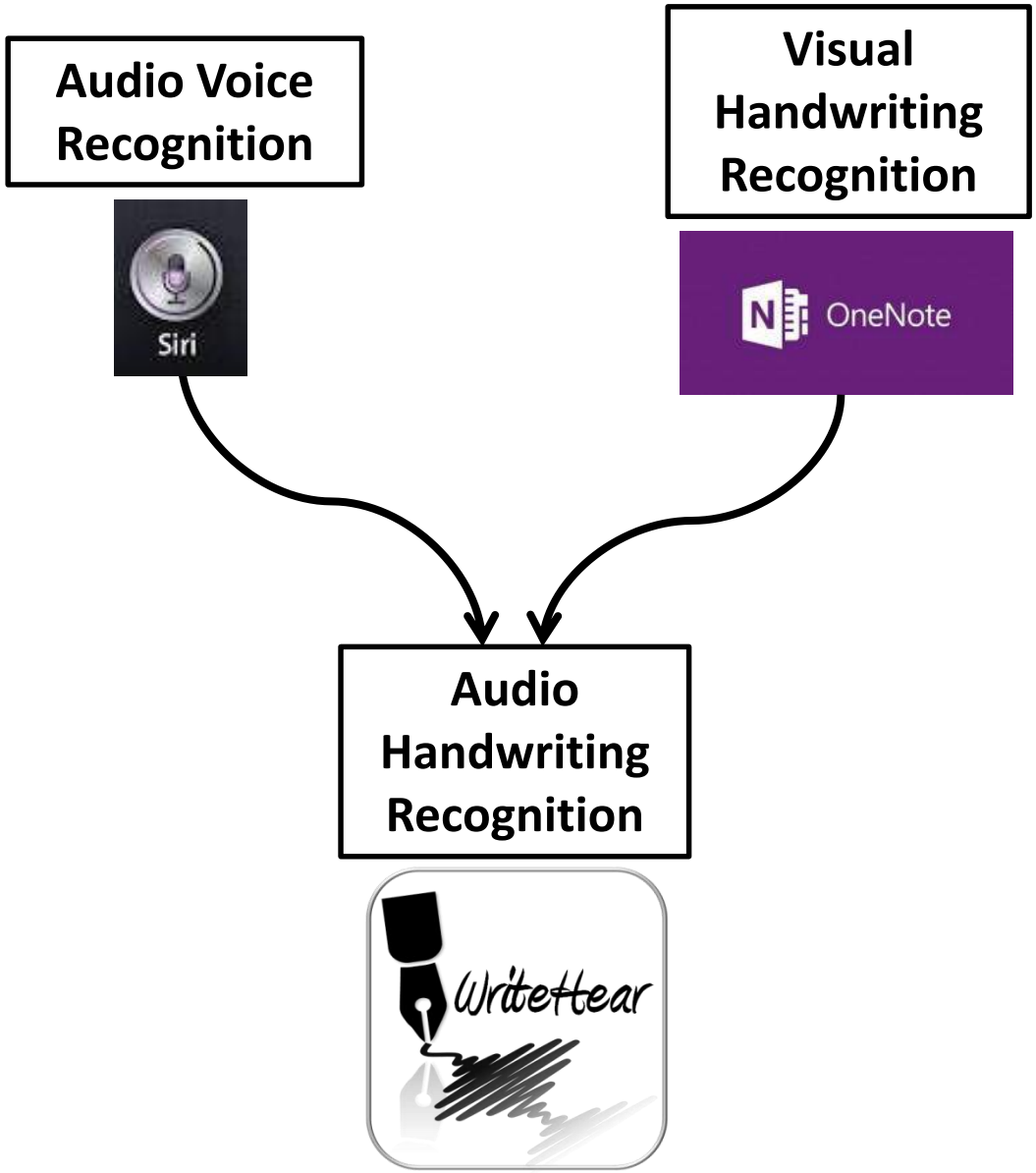
Jesse Miller
Dayton, OH
ISSI

Overview

I. What is WriteHear?
   A. Concept
   B. How it works
      1. Software
      2. Hardware
II. Testing and analysis
   A. Hardware
   B. Software

What is WriteHear?

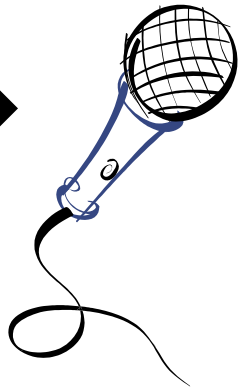
Concept


Audio Voice Recognition
Siri
Visual Handwriting Recognition
OneNote
Audio Handwriting Recognition
WriteHear

Concept

Sounds produced by handwriting

Recorded by microphone



Transcribed by software

Motivation

Using audio alone:

- Transcribe handwriting
- Verify signatures
- Send encrypted messages

Past Research

Sketch Recognition Lab at Texas A&M

Did Not

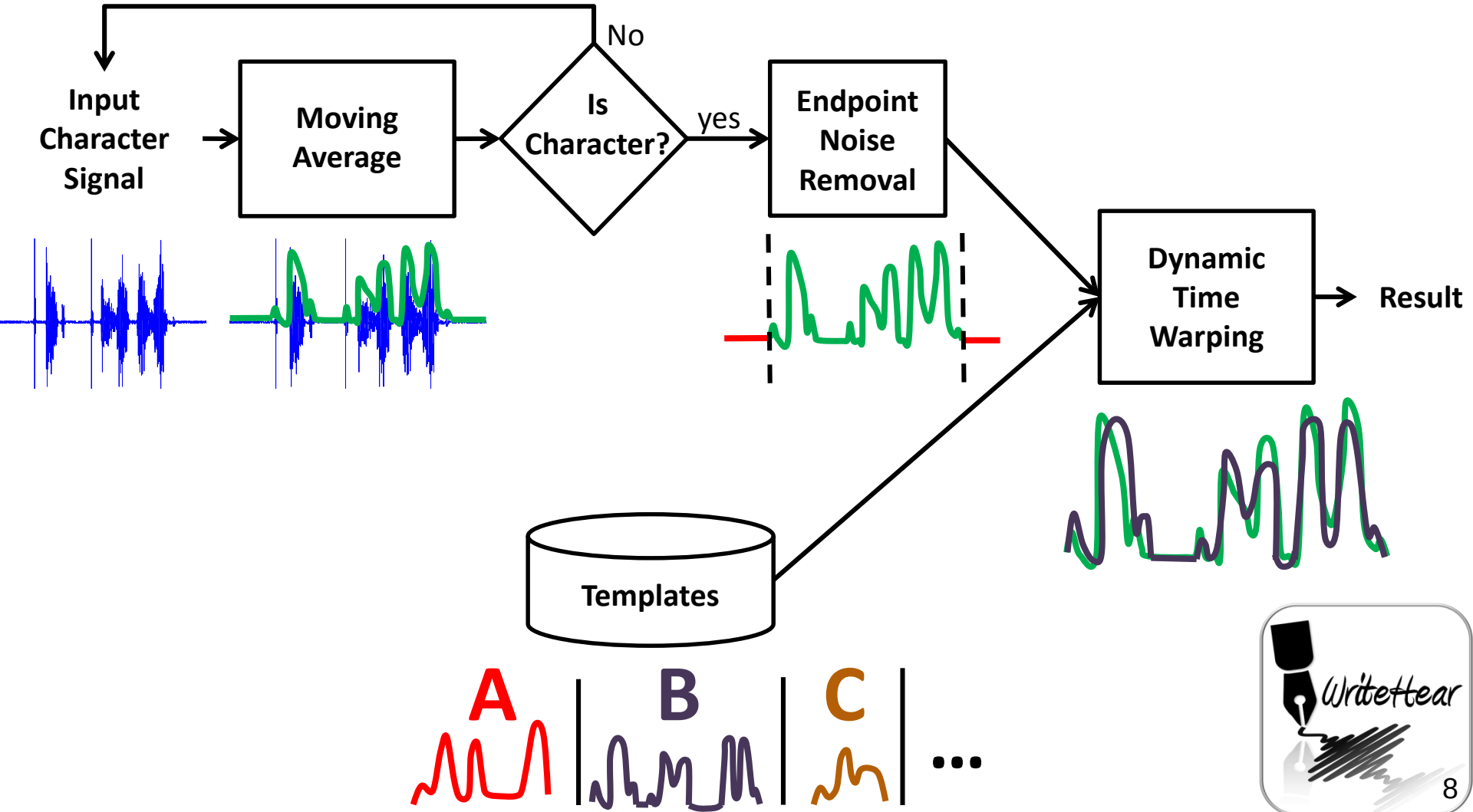
- Continuous character recognition
- Noise reduction
- Test various sets of characters
- Test different hardware setups
- Signature verification

Did

- Distinguish between uppercase characters with %86.8 accuracy

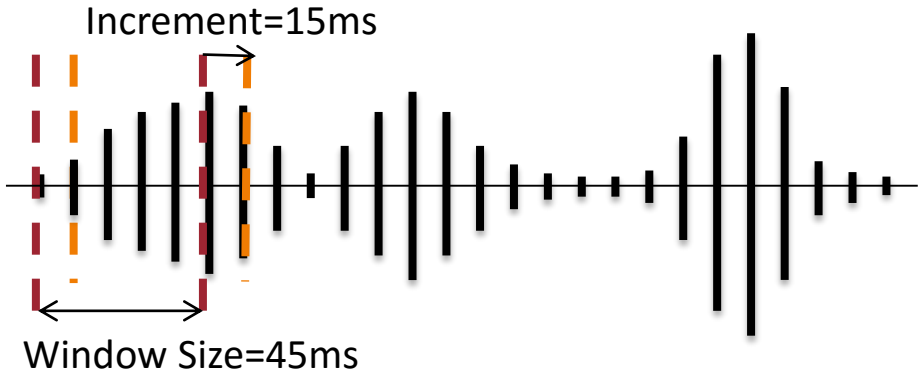
Software

System Architecture

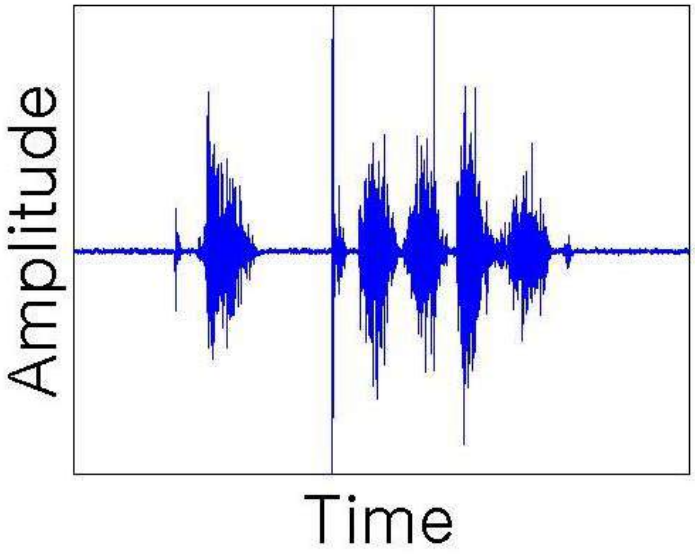


Software

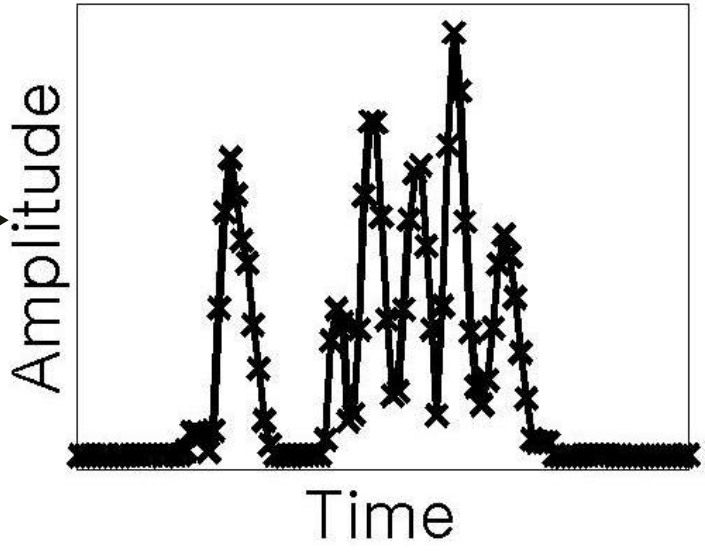
Moving Average



Input Signal



Moving Average



Software

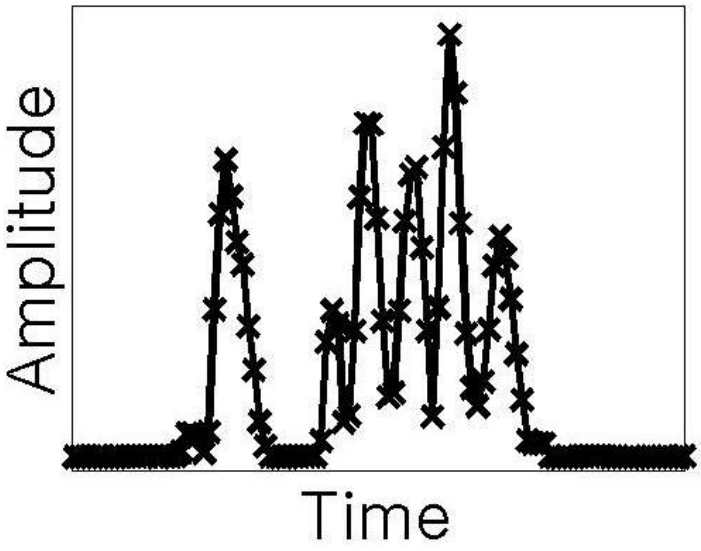
Endpoint Noise Removal

$$E = signal\ energy = \frac{1}{n}\sum_{i=1}^{n} {S_i}^2$$

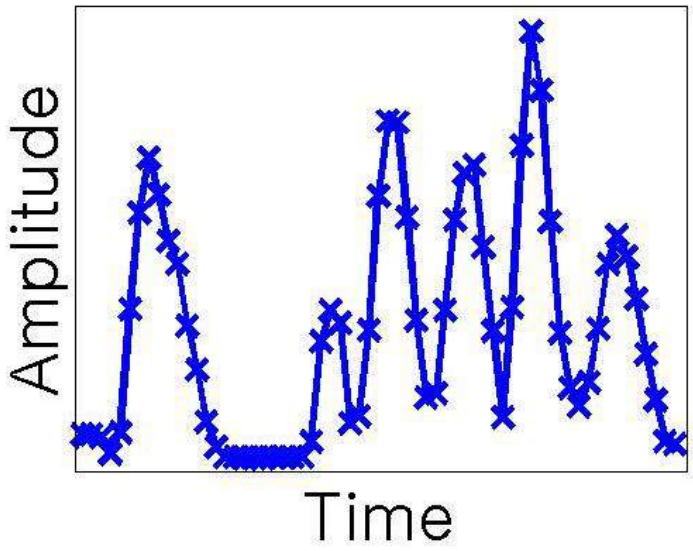
$$E_{character} > T \times E_{noise}$$

T is a fixed threshold

Moving Average



Noise Removal

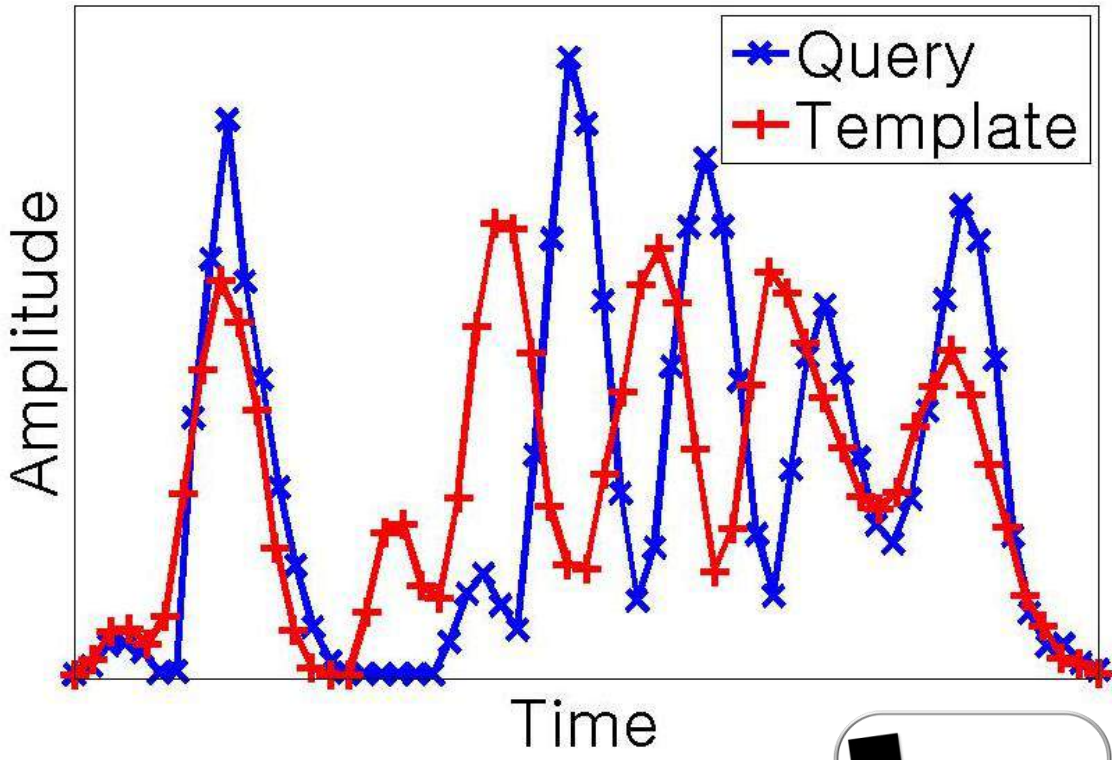


Software

Dynamic Time Warping

Method for Comparison:

- Stretches signal along time axis
- Independent of signal length



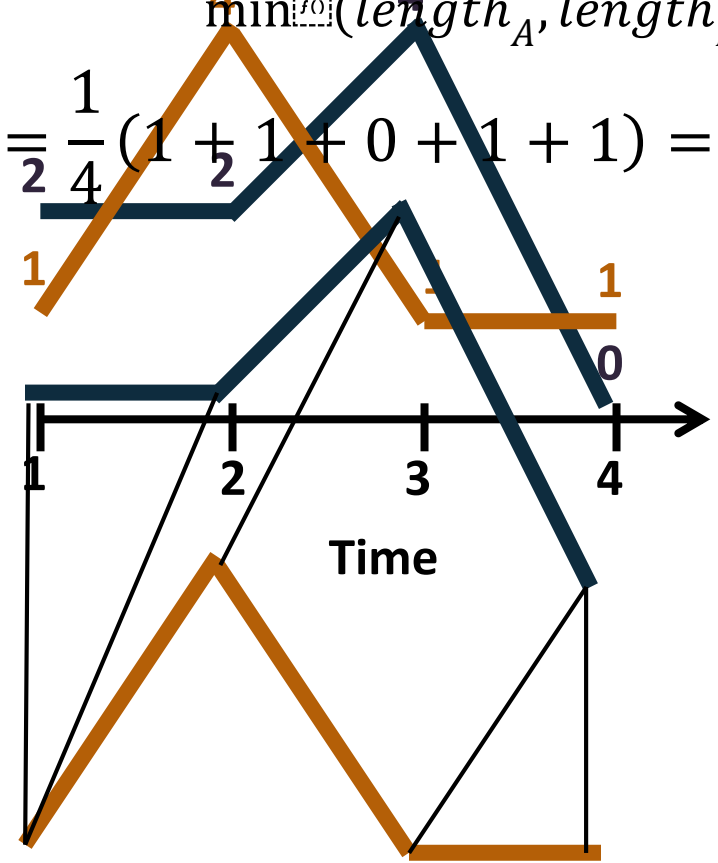
Software

Dynamic Time Warping

$$d_{r,c} = |A(r) - B(c)|$$

$$Score = \frac{1}{\min(length_A, length_B)} \sum_{i=1}^{n} p(i)$$

$$= \frac{1}{4}(1 + 1 + 0 + 1 + 1) = 1$$



B

A

start 1	1	3	1
2	2	0	4
1	1	3	1
1	1	3	1 end

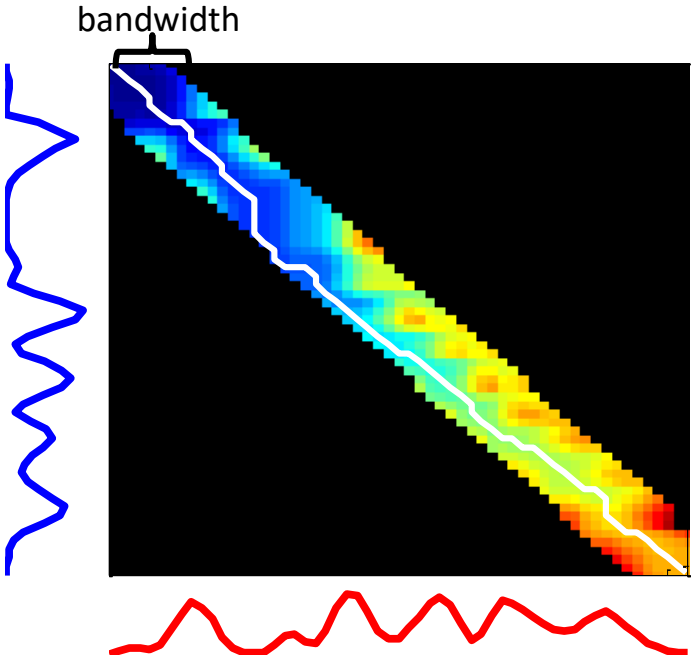
p(i)

p=path of minimal accumulated distance

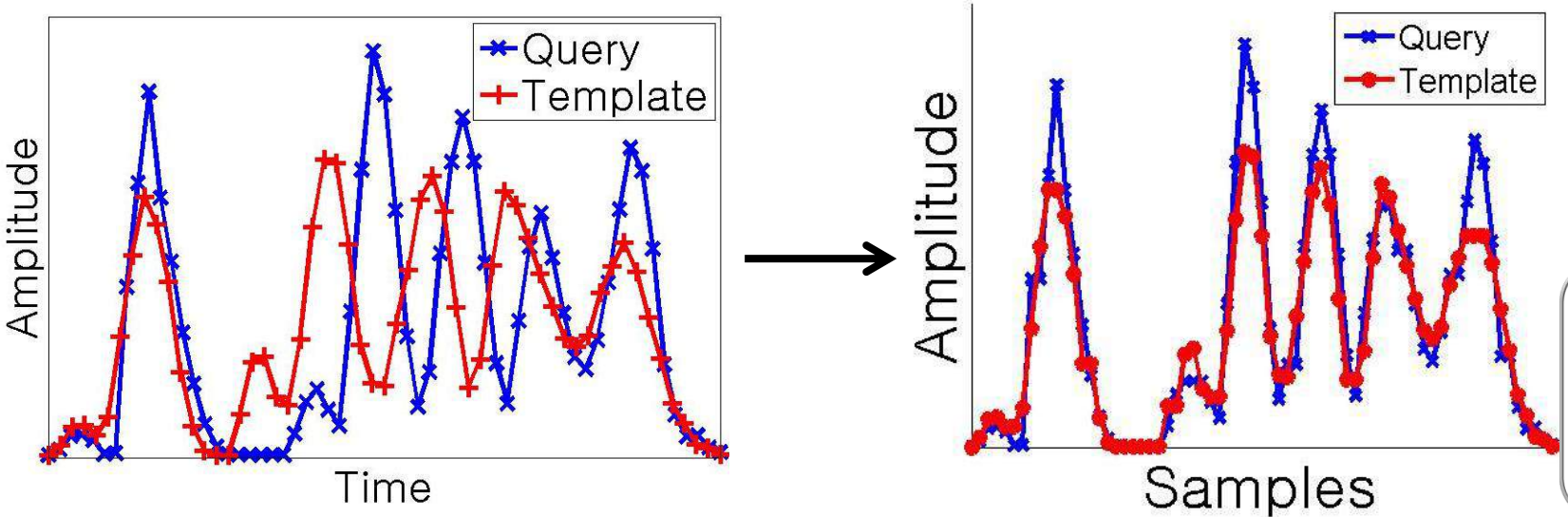
Software

Dynamic Time Warping

- Sakoe-Chiba band width of 8% used to restrict warping and maximize accuracy
- Square root of the signals are used for comparison to maximize accuracy



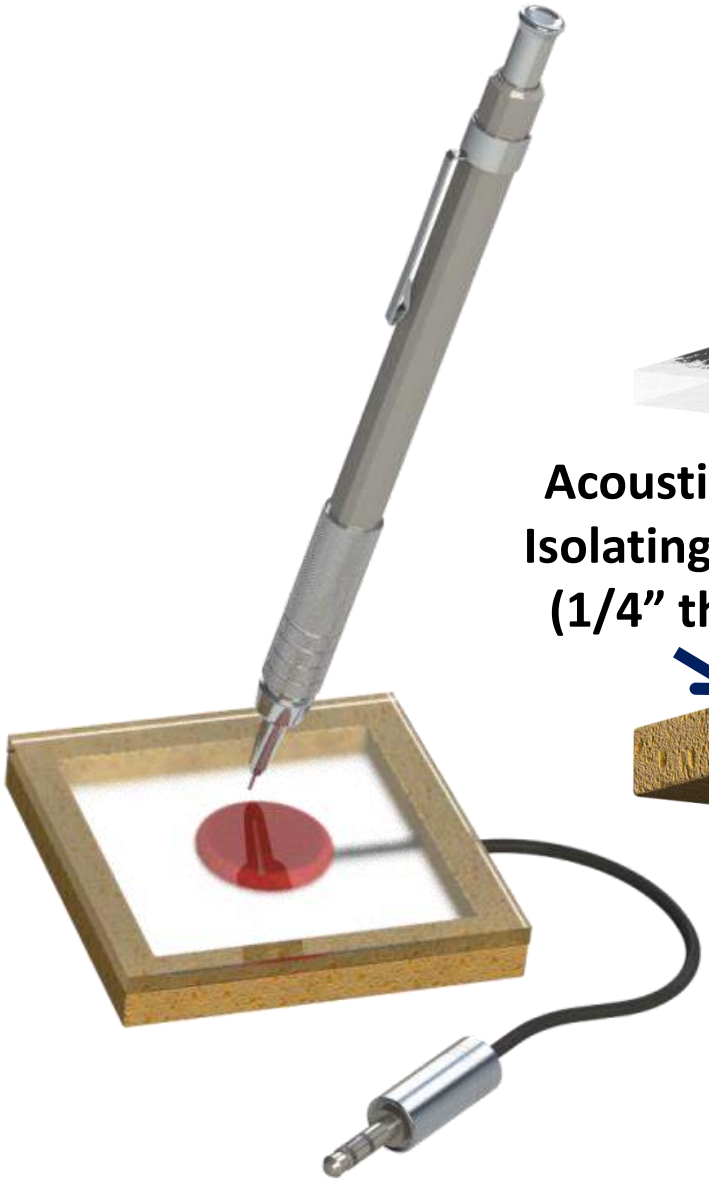
*see appendix for details

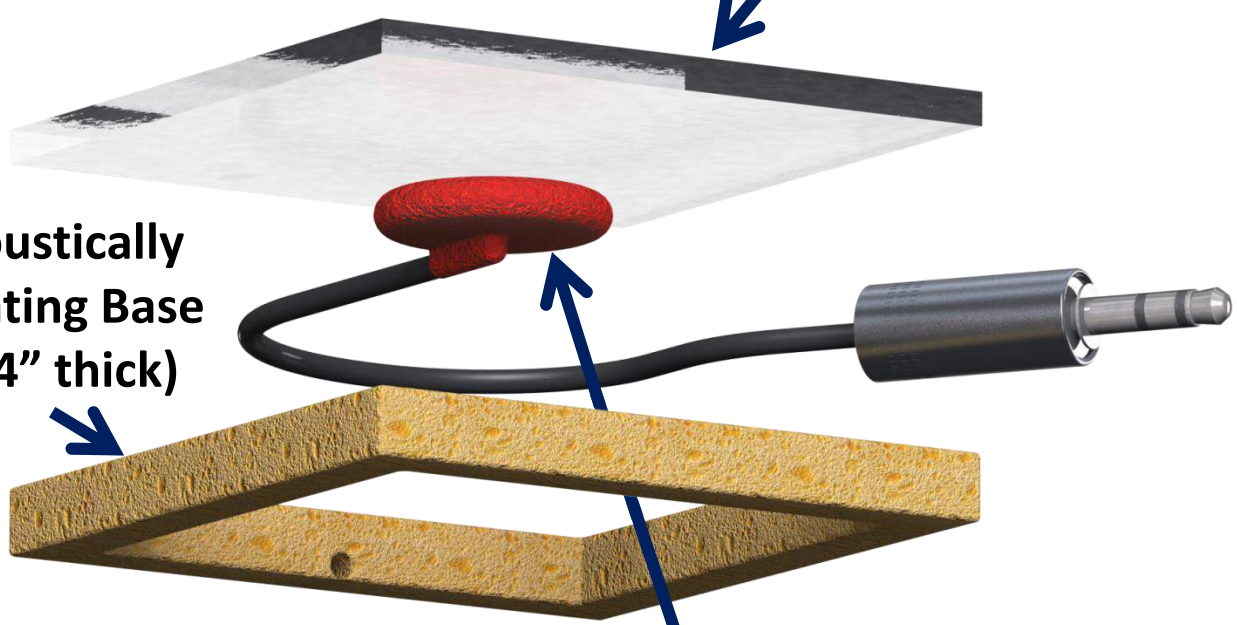


MATLAB GUI

Spellcheck database contains over 58,000 words

Hardware





Thin Writing Surface (2.5x2.5x1/8)”

Acoustically Isolating Base (1/4” thick)

Contact Microphone

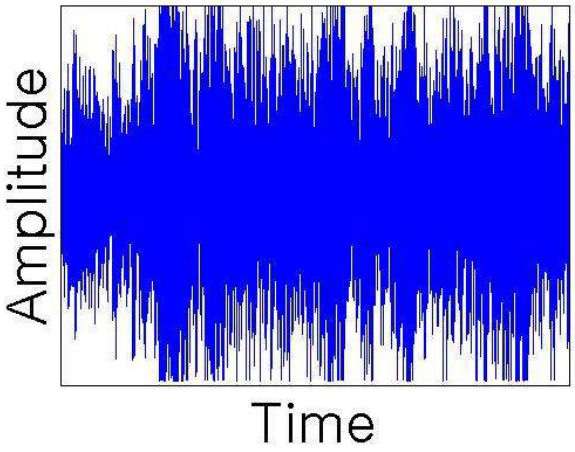
Testing and Analysis

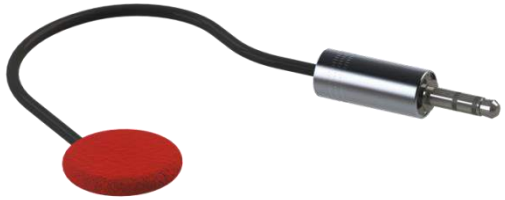
Type of Mic.

Dynamic Vs. Contact

- Electromagnetic induction
- Senses vibrations through air and solid objects
- Susceptible to noise

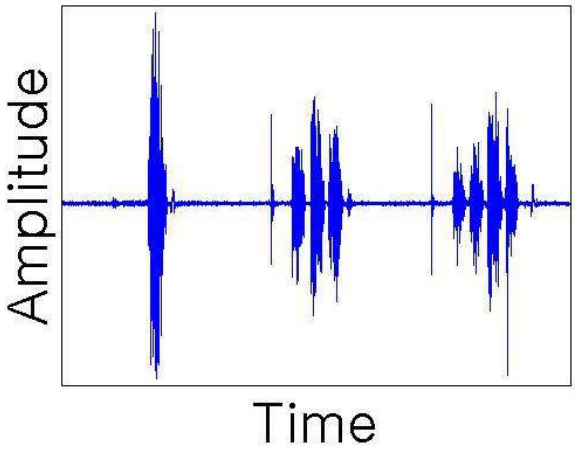
90 dB Noise Test





- Piezoelectric
- Senses vibrations through solid objects only
- Significantly less effected by noise

90 dB Noise Test

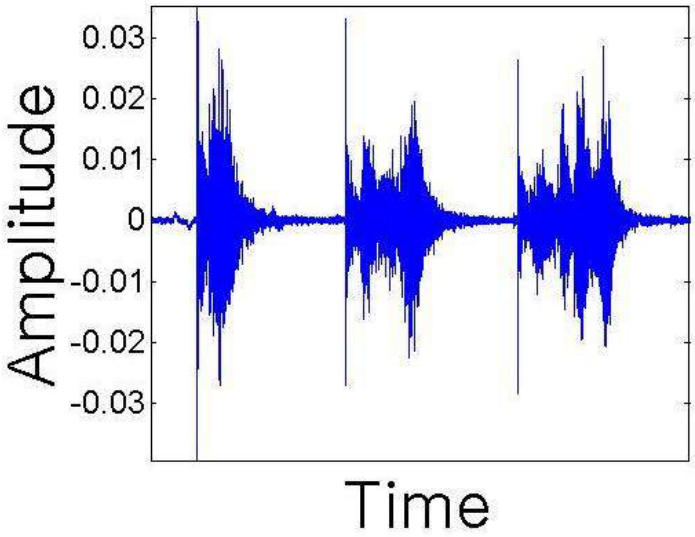


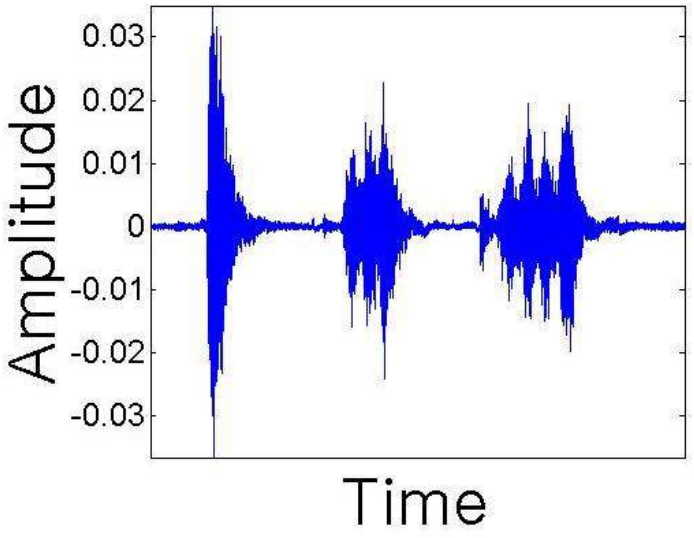
Distance From Mic.

1 inch ✓ Vs. 48 inches ✓

- Amplitude = 0.01-0.03
- Unaffected by distance

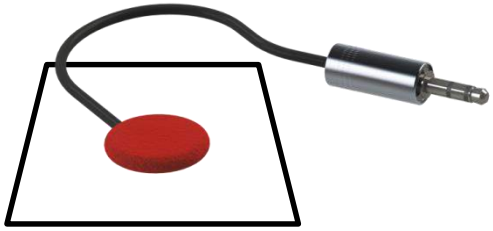
- Amplitude = 0.01-0.03
- Unaffected by distance

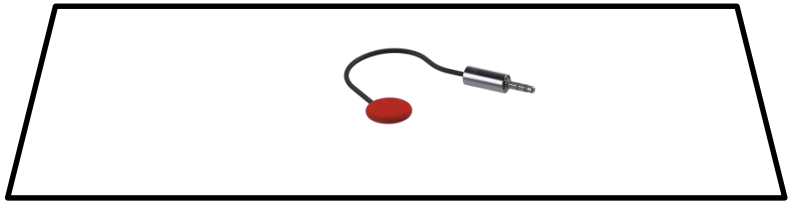




Surface Size

15 in^2 ✓ Vs. 720 in^2

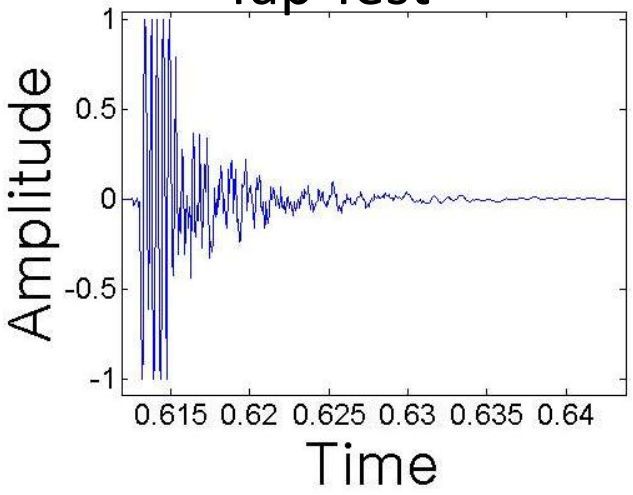




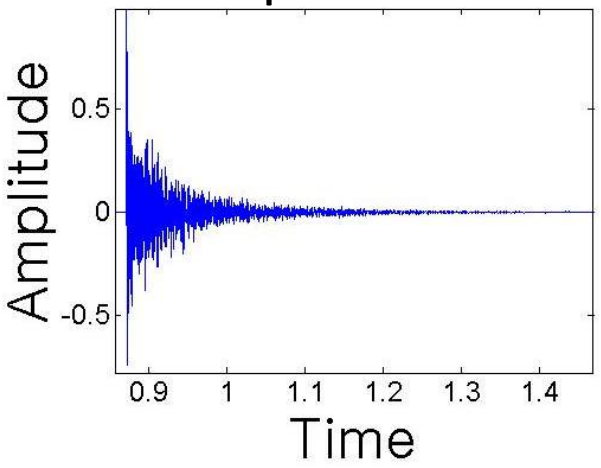
- Time to decay to noise level = 15ms
- More distinct features in writing
- Higher amplitude
- Less susceptible to noise

- Time to decay to noise level = 300ms
- Less distinct features in writing
- Lower amplitude
- More susceptible to noise

Tap Test

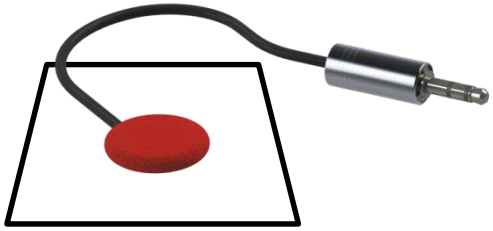


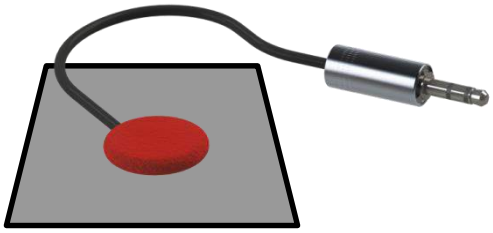
Tap Test



Surface Material

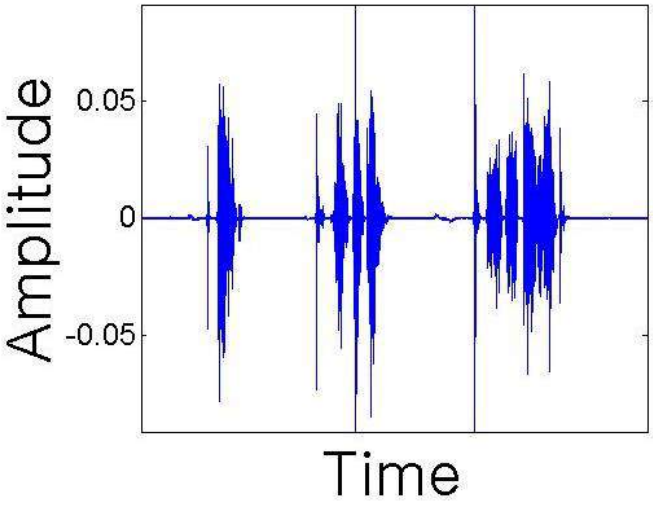
Plastic ✓ Vs. Metal ✓

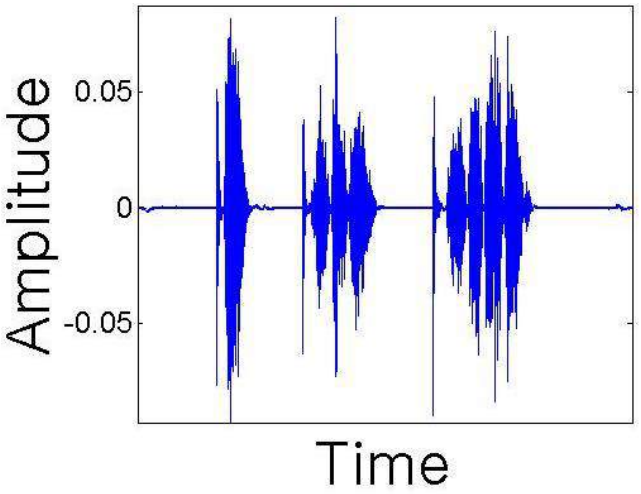




- Negligible effect on amplitude

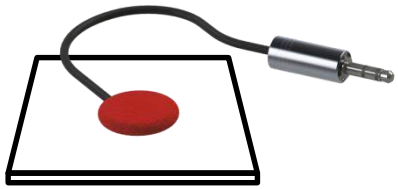
- Negligible effect on amplitude

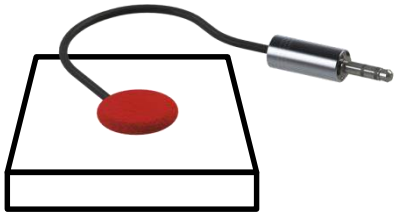




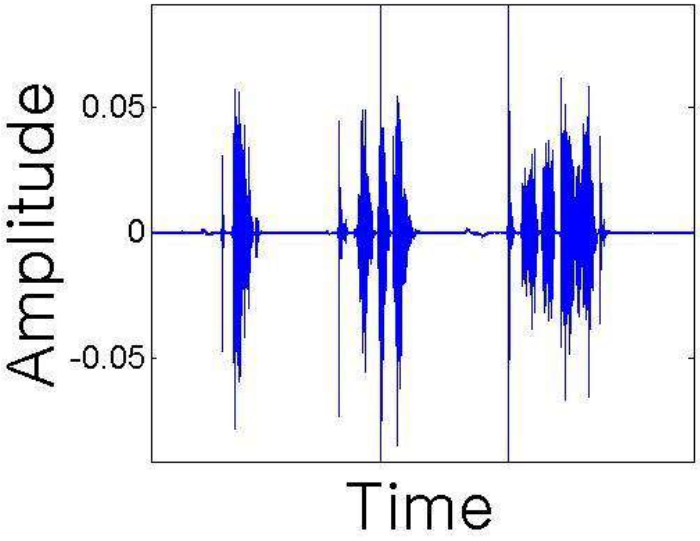
Surface Thickness

1/8” Thick ✓ Vs. 1/2” Thick

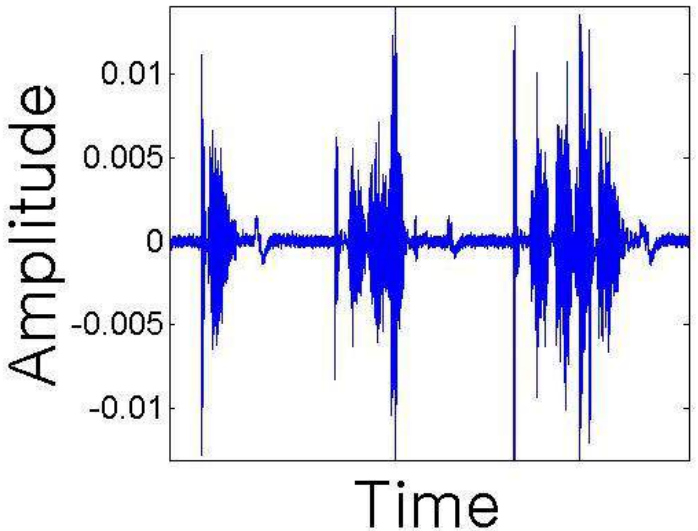




- Amplitude = 0.04-0.05



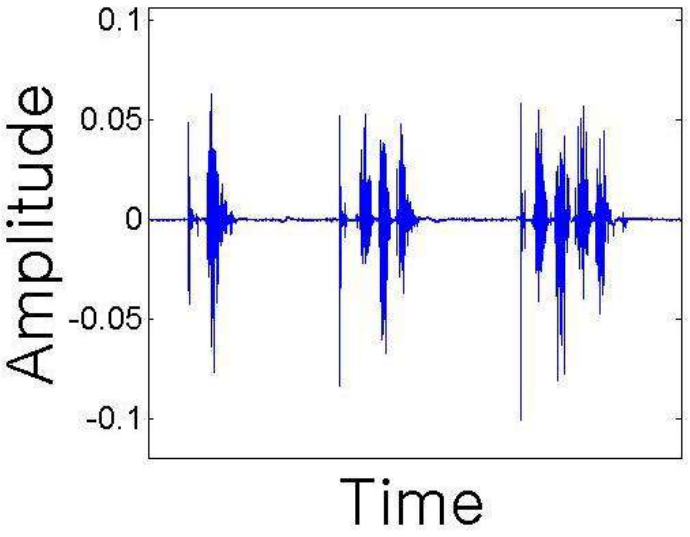
- Amplitude = 0.005-0.01 (1/8th the amplitude)



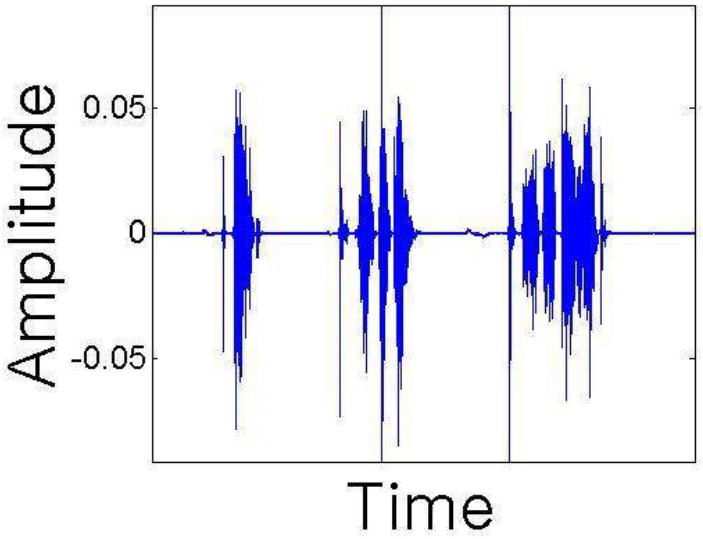
Writing Utensil

Pen ✓ Vs. Pencil ✓

- Amplitude = 0.04-0.05



- Amplitude = 0.04-0.05



Accuracy

A-Z	a-z	0-9	Spellcheck (A-Z)	Signature
87.5% Based on: 8 Individuals 832 samples	**77.9%** Based on: 2 Individuals 208 samples	**91.7%** Based on: 3 Individuals 120 samples	**85.9%** Based on: 3 Individuals 78 words	Type I Error: **4.8%** Type II Error: **6.3%** Based on: 3 Individuals 53 samples

- Similar sound profiles for different characters lowers accuracy
- Character accuracy calculated with a template database of 3 samples per character. A similar method was used by SRL at Texas A&M

Conclusions

- WriteHear is a robust and versatile software that can learn and understand different people's handwriting
- A Contact microphone setup greatly reduces background noise allowing WriteHear to work fine with 90dB (loud traffic)
- WriteHear works with various materials and writing utensils
- WriteHear can catch a forged signature with over 93% accuracy
- WriteHear recognized the upper case letters (A-Z) with 87.5% accuracy, matching the 86.8% accuracy achieved by the Sketch Recognition Lab at Texas A&M

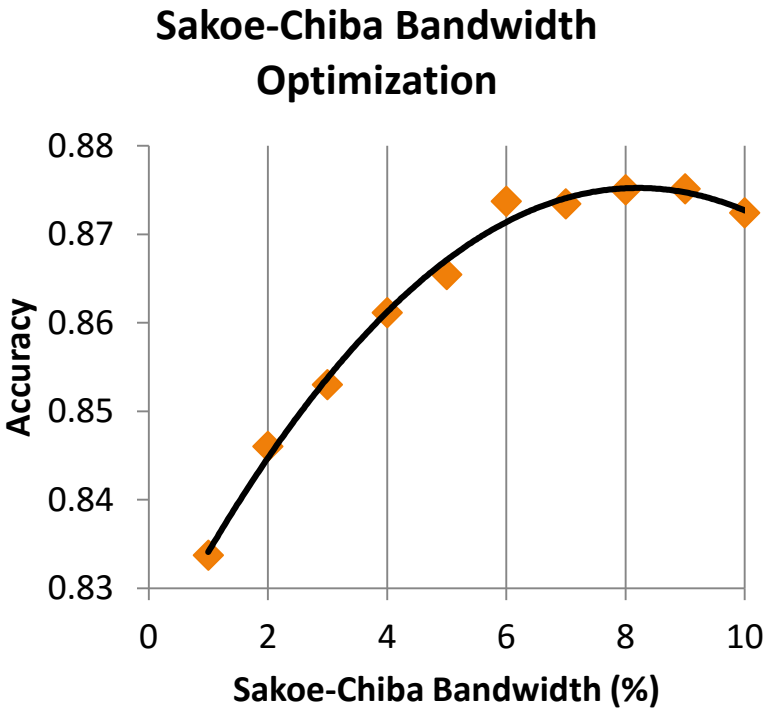
Future Work

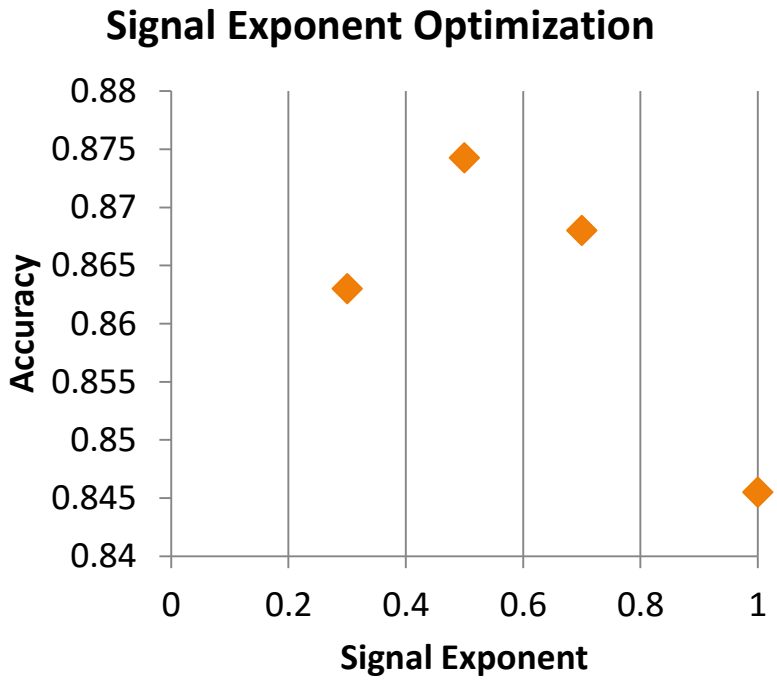
- Continue to improve and optimize software
- Test more conditions such as how a persons handwriting varies over time
- Create a user independent system
- Create a more fluid system where the user can write naturally without having to pause between characters

References

- Li, W., & Hammond, T. A. (2011, April). Recognizing Text Through Sound Alone. In *AAAI*.

Appendix





- Optimal bandwidth for dynamic time warping is 8%

- Optimal exponent on the signal is 0.5