Acoustic Signal Processing for Material Boundary Transition Detection

André Lopez[1], Ravi Prakash[1], Patrick Codd [1,2,3]

[1]Thomas Lord Department of Mechanical Engineering & Materials Science, Duke University
[2]Department of Neurosurgery, School of Medicine, Duke University
[3]Department of Head and Neck Surgery & Communication Sciences, School of Medicine, Duke University

Abstract

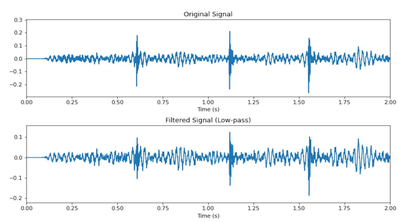
- This study explores the use of **sound signals** to discern material **boundaries**, with the goal of integrating this capability into a **controller** for a robotic tactile system.
- A variety of **sound features** have been studied including:
  - Short-time Fourier transform coefficients (STFTCs),
  - Wavelet packet transform coefficients (WPTCs),
  - Mel-frequency cepstrum coefficients (MFCCs),
  - Audio features: zero crossing rate, spectral centroid
- **Methods** for boundary identification include:
  - Mean Squared Energy (MSE)
  - Correlation Coefficients
  - Distance calculations
  - Dynamic Time Warping (DTW)

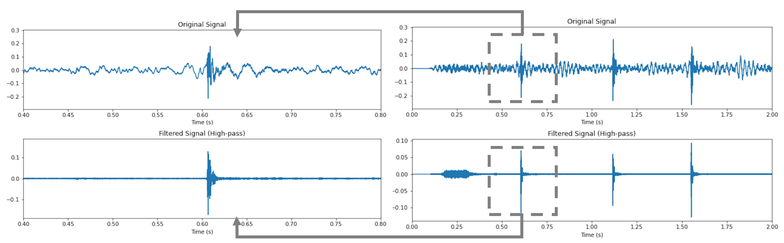
Signal Processing Pipeline

1. Preprocessing of signal:
   - Trimming
   - Filtering
   - Windowing [1]

2. Feature Extraction:
   - STFTCs
   - WPTCs

3. Similarity Calculations:
   - DTW
   - Correlation Coefficients
   - Cross correlation
   - MSE

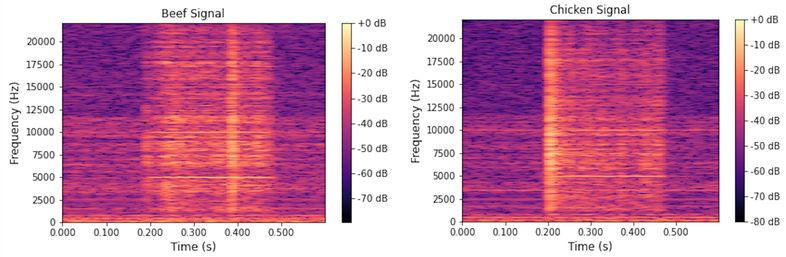


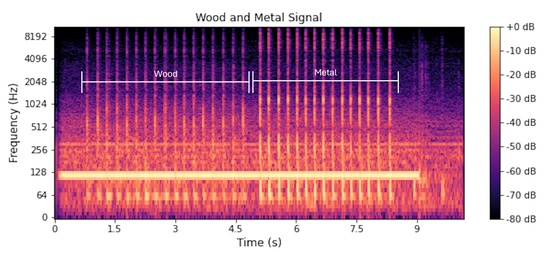


Signal before and after being passed through a low pass filter and a high pass filter. The cutoff frequency for both was 1000 Hz

Promising Feature: STFT Spectrograms

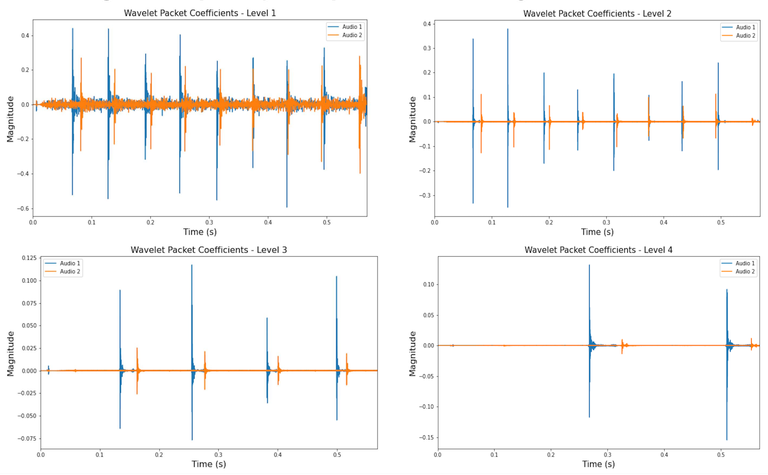
These reveal dynamic changes in the signal's frequency content. These are useful for audio signals because they are inherently time-varying.





Promising Feature: Wavelet Packet Transforms

Wavelet transform analysis splits the original signal into different levels. Each **level of decomposition** corresponds to a different scale or **frequency band**, with **higher levels** capturing finer details and **higher-frequency components** of the signal.

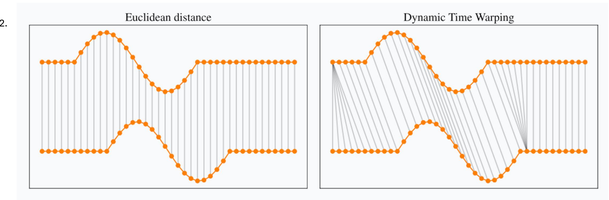


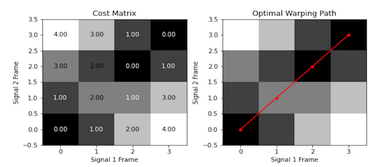
Results

Test data was produced by tapping on metal, tapping on wood, and a combination where metal was tapped on for four seconds and then wood was tapped on for three seconds. The STFT coefficients were calculated for each signal. These coefficients were then compared between all the signals, with the results displayed below.

Signals	MSE	Euclidean Distance	DTW
Metal-Wood	0.137	205.905	2.549
Metal-Combined(Wood)	0.123	188.502	2.575
Wood-Combined(Wood)	0.044	136.800	**1.773**
Metal-Combined(Metal)	0.156	228.153	**1.861**
Wood-Combined(Metal)	0.108	180.779	2.562

2.





Reference cost matrix generated from dummy data for visualization

Conclusions and Future Plans

Filtering and dynamic time warping provide the best signal similarity estimate. Moving forward, we plan to gather data in a more consistent way. This will be done by a linear actuator tapping on materials at a constant rate. The study then aims to utilize these results in real time to guide a controller for the robotic tactile system.

References + Acknowledgements

1.V. Zakeri and A. J. Hodgson, "Automatic Identification of Hard and Soft Bone Tissues by Analyzing Drilling Sounds," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 27, no. 2, pp. 404-414, Feb. 2019, doi: 10.1109/TASLP.2018.2880336.
2.Tavenard, Romain. "An Introduction to Dynamic Time Warping." 2021. Web. 15 Apr. 2024.
I would like to thank the Undergraduate Research Support Office for their financial support through the URS Assistantship. I would also like to thank the members of the Brain Tool Lab for their guidance throughout this project.