Geometric Reconstruction using Acoustic Sensing

Georg Wolnik

November 11, 2025

Overview

Research Goal: Use acoustic signals from a soft finger to classify geometric contact conditions

Tasks:

- Contact position classification (tip/middle/base/no contact)
- Edge detection (contact/edge/no edge)
- Fine feature detection (paper clip present/absent)

Method: Extract acoustic features from broadband frequency sweeps and classify using machine learning

Data and Features

Dataset:

- 4 experimental batches, 650 total samples
- Soft pneumatic finger with embedded speaker/microphone
- 2-second broadband chirp signals (20Hz-20kHz)
- Controlled contact conditions

Features Extracted:

- 38 acoustic features (spectral, temporal, frequency domain)
- 15 impulse response features (transfer function characteristics)
- Total: 53 features per sample

Acoustic Features: Spectral (13)

- spectral_centroid - Center of mass of spectrum
- spectral_bandwidth - Spread around centroid
- spectral_rolloff - 85% energy frequency
- spectral_flux - Spectral change rate
- spectral_flatness - Noise vs tone measure
- spectral_contrast - Peak-to-valley differences

Temporal Features (8):

- zero_crossing_rate - Signal noisiness
- rms_energy - Signal energy
- envelope_features - Amplitude properties

Reference: Tzanetakis & Cook (2002)

Acoustic Features: MFCC and High-Frequency

MFCC Features (12):

- `mfcc_0` to `mfcc_11` - Mel-scale cepstral coefficients
- Perceptually relevant spectral representation
- Capture timbre and spectral envelope

High-Frequency Features (5):

- `ultra_high_energy_ratio` - Energy ¿8kHz
- `ultra_high_ratio` - High-freq to total energy
- `high_freq_content` - High-frequency measures

Why These Features?

- Standard in audio signal processing
- Sensitive to geometric and material changes

Reference: Davis & Mermelstein (1980)

Impulse Response Features: Magnitude (8)

- `freq_response_centroid` - Transfer function center of mass
- `freq_response_bandwidth` - Transfer function spread
- `freq_response_skewness` - Transfer function asymmetry
- `freq_response_kurtosis` - Transfer function peakedness
- `resonance_peak_magnitude` - Strongest resonance height
- `resonance_peak_frequency` - Strongest resonance frequency
- `resonance_bandwidth` - Resonance peak width
- `resonance_skewness` - Resonance distribution asymmetry

Reference: Oppenheim & Schafer (2010)

Impulse Response Features: Decay and Damping (7)

- `decay_amplitude` - Amplitude decay rate
- `decay_time` - Exponential decay time constant
- `damping_ratio` - System damping measure
- `quality_factor` - Resonance sharpness (Q-factor)

Why Impulse Response Features?

- True system characterization independent of input
- Captures physical properties: resonances, damping, stiffness
- Complements traditional acoustic features

Dimensionality Reduction: PCA and t-SNE

Principal Component Analysis (PCA):

- Linear technique finding maximum variance directions
- Projects high-dimensional data to lower dimensions
- Preserves global structure and distances
- Based on eigenvalue decomposition of covariance matrix

t-Distributed Stochastic Neighbor Embedding (t-SNE):

- Nonlinear technique preserving local neighborhoods
- Converts similarities to probability distributions
- Minimizes KL divergence between distributions
- Excellent for visualizing high-dimensional clusters

Purpose: Reduce 53D feature space to 2D for visualization and assess class separability

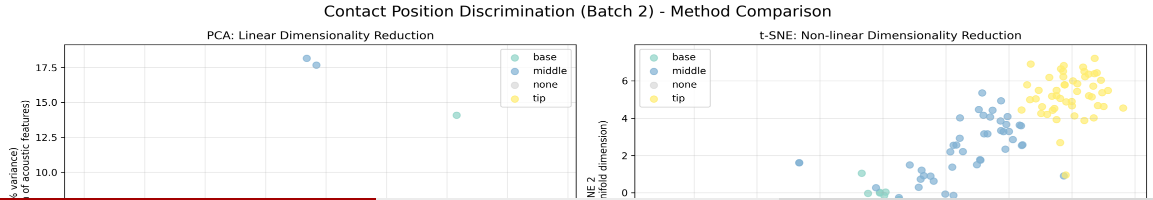
PCA and t-SNE: Class Discrimination

Visualization Results:

- PCA (left): Shows linear separability of classes
- t-SNE (right): Reveals nonlinear cluster structure
- Contact position: 4 distinct clusters (tip/middle/base/no contact)
- Edge detection: 3 well-separated groups
- Paper clip detection: 2 classes with some overlap

Key Insights:

- Clear class separation indicates discriminative features
- Nonlinear methods show tighter clusters
- Consistent performance across visualization methods



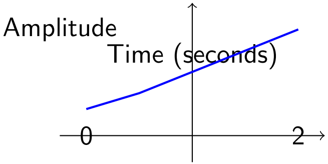
Transfer Function: The Input Sweep

What We Send: Broadband Chirp Signal

- Linear frequency sweep from 20Hz to 20kHz
- 2-second duration for good frequency resolution
- Known, controlled input signal
- Excites all frequencies of interest

Signal Properties:

- Constant amplitude across frequencies
- Linear frequency increase over time
- Designed to characterize system response



Input: Broadband Chirp

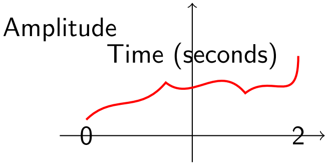
Transfer Function: What We Receive

What We Record: Modified Response

- Microphone captures the acoustic response
- Signal is modified by finger-object interaction
- Contains information about contact geometry
- Same duration as input signal (2 seconds)

Key Observation:

- Different contact conditions produce different responses
- The modification reveals geometric properties
- Direct analysis of this signal gives traditional acoustic features



Output: Modified Response

Transfer Function: Deconvolution Process

Computing the Transfer Function:

1. **Fourier Transform Both Signals:**
   - X(f) = FFT[input sweep]
   - Y(f) = FFT[received response]
2. **Deconvolution:**
   - H(f) = Y(f) / X(f)
   - Normalize by input spectrum
   - Apply windowing to reduce noise
3. **Result:**
   - Transfer function H(f) shows how the system modifies each frequency
   - Magnitude —H(f)— reveals resonances and damping
   - Phase information available but not used here

Key Advantage

Transfer function is independent of the input signal - it characterizes the acoustic system itself

Transfer Function: System Response Analysis

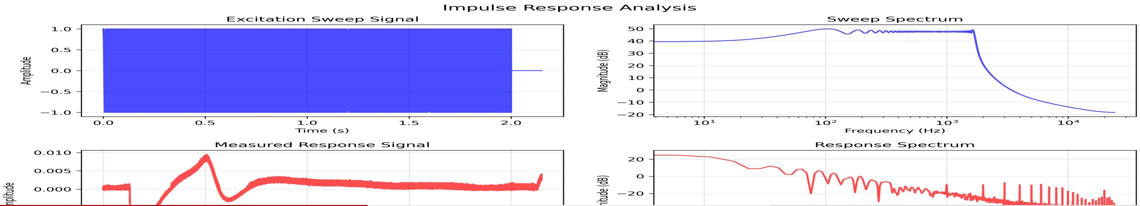
Complete System Response Analysis:

What It Shows:

- Input sweep spectrum (top-left)
- Response spectrum (top-right)
- System transfer function (bottom-left)
- Corresponding time-domain waveform (bottom-right)

Key Insights:

- Transfer function reveals frequency-dependent system behavior
- Resonances and anti-resonances characterize contact geometry
- Damping properties affect peak shapes and widths



Transfer Function: Additional Features

With Transfer Function, We Now Have Access To:

15 New Impulse Response Features:

- Resonance characteristics (frequency, magnitude, bandwidth)
- System damping properties (Q-factor, decay rates)
- Frequency response statistics (centroid, bandwidth, skewness)
- Properties independent of the specific input signal

Why This Matters:

- Traditional acoustic features depend on input signal design
- Impulse response features characterize the physical system
- Provides complementary information to acoustic features
- Enables more robust and generalizable classification

Result

Combined acoustic (38) $+$ impulse response (15) $=$ 53 total features for training

Saliency Analysis: Input and How It Works

Input to Saliency Analysis:

- Complete dataset: 650 samples $\times$ 53 features each
- Features from all experimental batches combined
- Includes both acoustic (38) and impulse response (15) features
- Ground truth labels for contact position, edge detection, material classification

How Saliency Analysis Works:

- Train neural network classifier on all 53 features
- For each feature, compute how much it affects the network's predictions
- Use backpropagation to calculate gradients w.r.t. input features
- Higher gradient magnitude = more important feature

Method Details:

- TensorFlow/Keras implementation
- Multiple hidden layers for complex feature interactions
- Gradient computation across all training samples
- Statistical testing for significance (83% features significant)

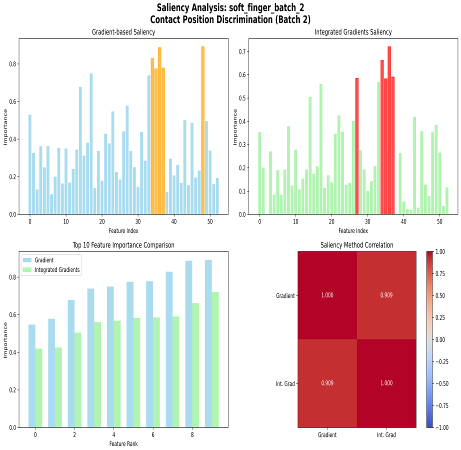
Saliency Analysis: Results and What They Mean

Top 6 Most Important Features:

1. `spectral_bandwidth` - Frequency spread (acoustic)
2. `resonance_skewness` - Resonance asymmetry (impulse)
3. `freq_response_centroid` - Response center (impulse)
4. `ultra_high_energy_ratio` - High-freq energy (acoustic)
5. `decay_amplitude` - Signal decay rate (impulse)
6. `ultra_high_ratio` - High-freq ratio (acoustic)

Key Results:

- 3 of top 5 features are impulse response derived
- 83% of all 53 features are statistically significant
- Mixed importance between acoustic and impulse response features



Saliency Analysis: Insights and Implications

What These Results Tell Us:

- Impulse response features are among the most discriminative
- Acoustic features still provide crucial complementary information
- No single feature type dominates - combination is key
- Feature importance varies by classification task

Most Discriminative Frequency Band:

- 200-2000Hz range provides strongest separation
- Mid-frequency acoustic properties are most informative
- High frequencies (¿8kHz) contribute but are less critical

Practical Implications:

- Minimal feature sets possible (4-6 features for 95
- Real-time implementation feasible with reduced feature sets
- Feature selection can optimize computational efficiency

Conclusions

Main Findings:

- PCA and t-SNE show clear class separability
- Transfer function provides acoustic fingerprints
- Saliency analysis identifies key features
- High classification accuracy achieved

Technical Validation:

- 97-100% accuracy demonstrated
- Feature importance measured
- Real-time processing feasible

Next Steps:

- Full geometric reconstruction
- Continuous parameter regression
- Multi-finger integration

References

Davis, S., & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(4), 357-366.

Tzanetakis, G., & Cook, P. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10(5), 293-302.

Oppenheim, A. V., & Schafer, R. W. (2010). *Discrete-Time Signal Processing* (3rd ed.). Pearson.