Formant Detection in Audio Signals

Author: Thiksiga Ragulakaran (E/18/362)
Course: Audio Engineering
University: University of Peradeniya

Abstract

Formants are critical acoustic features in speech processing, representing the resonant frequencies of the vocal tract. This report explores the concept of formants, their significance in speech signals, and the procedure for detecting them using Linear Predictive Coding (LPC). The implementation details are discussed, including an observed case where only one formant was detected from two prominent peaks in the frequency response curve.

Introduction

Formants are prominent spectral peaks in the frequency domain of speech signals. They arise from the filtering effects of the vocal tract and play a significant role in defining vowel sounds and other speech characteristics. Accurate detection of formants is essential for various applications, including speech synthesis, recognition, and linguistic analysis.

Concept of Formants

1. **Definition**: Formants are resonant frequencies of the vocal tract that determine the acoustic quality of speech sounds. They are essential for distinguishing different vowel sounds and contribute to the timbre of speech.
2. **Formant Frequencies**:
   - **First Formant (F1)**: Related to vowel height (open vs. closed).
   - **Second Formant (F2)**: Associated with vowel backness (front vs. back).
   - **Higher Formants**: Provide additional nuances to vowel sounds.
3. **Speech Production**: The vocal tract acts as a resonator that amplifies certain frequencies produced by the vocal cords. The resulting peaks in the frequency spectrum are identified as formants.

Procedure for Formant Detection

1. **Preprocessing**:
   - **Load Audio Data**: Read the audio file and convert it to a mono signal if necessary.
   - **Normalization**: Normalize the audio signal to ensure consistency.
   - **Windowing**: Apply a window function (e.g., Hamming window) to minimize spectral leakage.
2. **Linear Predictive Coding (LPC)**:
   - **Autocorrelation Calculation**: Compute the autocorrelation of the windowed signal.
   - **Form LPC Coefficients**: Use the autocorrelation to construct a Toeplitz matrix and solve for LPC coefficients.
   - **Frequency Response Calculation**: Compute the frequency response of the LPC filter to obtain the magnitude spectrum.
3. **Peak Detection**:
   - **Identify Peaks**: Use peak detection algorithms to find prominent peaks in the magnitude spectrum.

    - **Formant Extraction**: Convert the peak indices to frequencies and filter out irrelevant peaks based on typical formant frequency ranges.
4. **Visualization and Analysis**:
    - **Plot Frequency Response**: Visualize the magnitude response and detected formants.
    - **Evaluate Detected Formants**: Analyze the detected formants and compare them with expected values.

Observations

In the analysis, two peaks were identified in the frequency response curve. However, only one of these peaks was recognized as a formant. The possible reasons for this discrepancy include:

- **Thresholding and Filtering**: The algorithm may have filtered out less significant peaks based on amplitude thresholds or frequency ranges.
- **Peak Selection**: The algorithm might prioritize the most prominent peak or one that aligns better with expected formant frequencies.
- **Resolution and Artifacts**: Resolution limitations or noise in the signal could affect peak detection.

Conclusion

Formant detection using LPC is a powerful technique for analyzing speech signals. While the implementation successfully identified one formant, the detection of multiple peaks requires careful tuning of parameters and thresholds. Further refinements in the algorithm and preprocessing steps can improve formant detection accuracy and reliability.