Pitch Estimation Using Linear Predictive Coding (LPC)

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

---

1. Introduction

Linear Predictive Coding (LPC) is a popular method in speech processing and audio analysis for modeling the spectral envelope of a digital signal. It is widely used for pitch estimation, which involves determining the fundamental frequency of a sound. This report discusses the LPC-based pitch estimation experiment performed on an audio signal, outlining the methodology, results, and analysis.

2. Objective

The primary objective of this experiment is to estimate the pitch (fundamental frequency) of an audio signal using Linear Predictive Coding (LPC). The focus is on analyzing how well LPC can identify pitch variations over time, particularly in the context of vocal sounds.

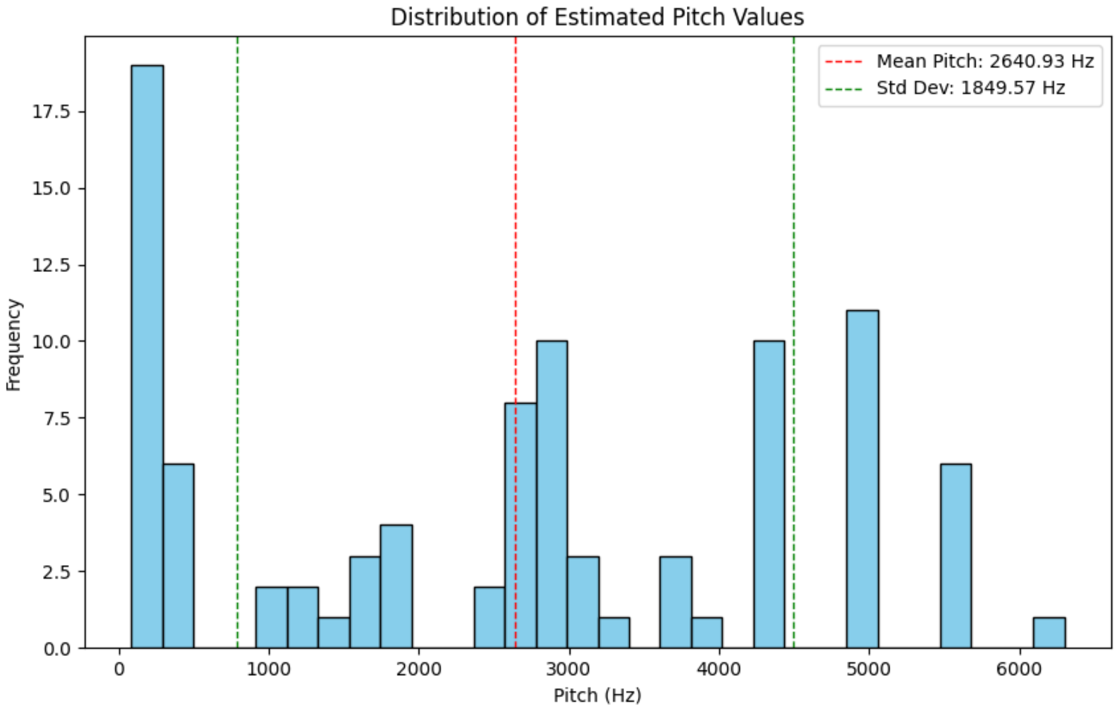
3. Methodology

1. **Audio File Preparation:**
    - The audio file used in this experiment (`aud2.wav`) was read using `scipy.io.wavfile`. The audio was converted to mono if it was initially stereo to simplify the analysis.
2. **Audio Slicing:**
    - A specific segment of the audio, from 0.4 to 1.5 seconds, was extracted to focus on a portion with distinct pitch characteristics.
3. **LPC Analysis:**
    - LPC analysis was performed using the autocorrelation method. The autocorrelation of the signal was computed to form the Yule-Walker equations, which were solved to obtain LPC coefficients.
    - The LPC order was set to 12, which is commonly used for speech signals, to balance the accuracy and computational complexity.
4. **Pitch Estimation:**
    - The residual signal was obtained by filtering the original signal using the LPC coefficients. The pitch was estimated by finding the peaks in the autocorrelation of the residual signal. The distance between these peaks corresponds to the period of the signal, allowing for pitch calculation.
5. **Plotting and Analysis:**
    - The estimated pitch values were plotted over time to visualize pitch variations throughout the analyzed segment.

4. Results

- **Mean Pitch:** 2640.93 Hz
- **Standard Deviation of Pitch:** 1849.57 Hz
- **Minimum Pitch:** 85.30 Hz
- **Maximum Pitch:** 6300.00 Hz

The pitch estimation process resulted in a mean pitch value of 2640.93 Hz, with a standard deviation of 1849.57 Hz. The minimum and maximum pitch values were 85.30 Hz and 6300.00 Hz, respectively.

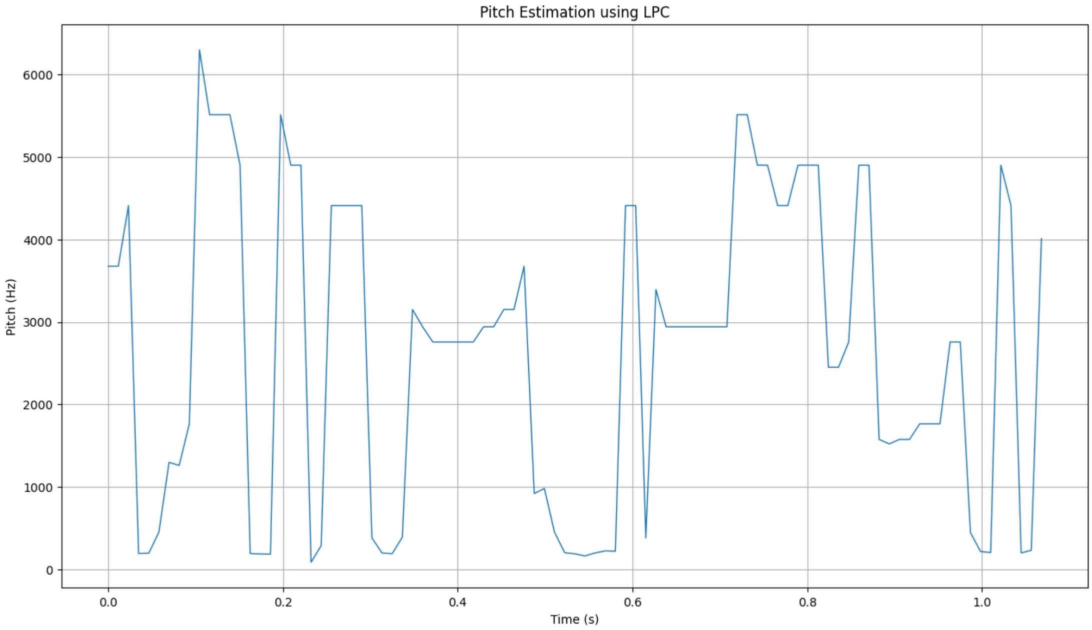


5. Analysis

- **Mean Pitch (2640.93 Hz):** The mean pitch represents the average fundamental frequency detected in the analyzed audio segment. A mean value of 2640.93 Hz indicates that the audio segment contained sounds that were relatively high-pitched.
- **Standard Deviation (1849.57 Hz):** The high standard deviation suggests a significant variation in the pitch throughout the analyzed segment. This could indicate the presence of both low and high frequencies, such as a combination of vocals and other sounds, or rapid pitch changes typical in speech or certain types of music.
- **Minimum Pitch (85.30 Hz):** This value is close to the lower range of human vocal pitch (e.g., a low male voice), suggesting that the LPC method successfully detected a low-frequency component in the signal.
- **Maximum Pitch (6300.00 Hz):** This high value suggests the presence of higher-frequency components, such as harmonics, higher-pitched notes, or noise elements in the audio segment.

6. Observations from the Plot

- The **pitch estimation plot** shows variations in pitch over time, corresponding to the changes in the sound's frequency content.



- The **peaks** in the plot represent moments of higher pitch, which could correspond to specific vocal sounds, musical notes, or other signal components.
- The **valleys** (lower pitch points) might indicate pauses, lower-frequency sounds, or intervals of less pronounced pitch content.

7. Conclusion

The experiment demonstrated that LPC-based pitch estimation can effectively identify variations in pitch over time. The method captured both low and high-frequency components, revealing the dynamic nature of the audio segment. However, the wide range and high standard deviation suggest that further refinement, such as noise filtering or segment selection, may improve the accuracy for specific applications.