Assignment Report 1 & 2

Ankit KumarBhakar EE 679

Speech Synthesis using Linear Predictive (LP) Analysis

1 Introduction

Speech synthesis is a fascinating field that aims to create artificial speech signals that mimic human speech patterns. Linear Predictive Coding (LPC) analysis is a widely used technique in speech processing to model and synthesize speech signals. In this report, we explore the application of LPC analysis for synthe- sizing vowels, focusing on the use of a periodic triangular pulse train to excite the vocal tract's all-pole filter model.

2 Synthesizing Vowels with Periodic Pulse Train

The process of synthesizing vowels involves exciting the vocal tract's all-pole fil- ter model using a periodic triangular pulse train. This excitation source closely resembles the glottal excitation observed in natural speech production. By ap- plying Linear Predictive Analysis with orders ranging from 2 to 10, we observed reduced errors in signal energy with higher orders. This suggests that higher- order LPC models capture finer details of the vocal tract's characteristics.

3 Spectral Analysis and Pre-Emphasis

To capture formant shaping and lip radiation effects, we conducted spectral analysis of pre-emphasized signals. Pre-emphasis involves boosting higher fre- quency components to enhance the prominence of formants. We observed that pre-emphasized signals exhibit reduced spectral tilt, indicating enhanced clarity of formant frequencies.

3.1 Pre-Emphasis Filter

We applied a pre-emphasis filter to the original speech signal to enhance the higher frequency components. The difference equation for the pre-emphasis

filter is given by:

$y[n] = x[n] - \alpha \cdot x[n-1]$
where $x[n]$ is the original signal, $y[n]$ is the pre-emphasized signal, and α is the

pre-emphasis coefficient.

3.2 Observations

The pre-emphasized signal exhibited a boosted spectral tilt, where higher fre- quency components were accentuated compared to the original signal. This tilt reduction enhances the intelligibility of speech signals by emphasizing formants that carry crucial phonetic information.

4 Levinson-Durbin Recursion and LPC Coeffi- cients

We employed Levinson-Durbin recursion to estimate Linear Prediction Coeffi- cients (LPC) for various orders. The autocorrelation function played a vital role in computing the LPC coefficients. We observed that the error energy de- creases as the order of the LPC model increases, indicating improved modeling accuracy.

4.1 Observations

The error energy decreased with increasing LPC order, reaching a plateau after a certain order (e.g., 8-10). This observation highlights the trade-off between modeling accuracy and complexity. Higher-order LPC models capture more details of the speech signal but may not yield significant performance gains beyond a certain point.

5 Inverse Filtering and Residual Signal

By employing inverse filtering, we extracted the residual error signal that repre- sented the difference between the original signal and the LPC model's output. This residual signal contains the fine details that the LPC model couldn't cap- ture.

5.1 Observations

The analysis of the residual error signal revealed the fundamental frequency of the speech signal. Peaks in the autocorrelation function of the residual signal indicated the fundamental pitch period, which was used to reconstruct the orig- inal signal. The spectral content of the residual signal highlighted the presence of harmonics and the removal of formant shaping.