Report on LPC Coder

Introduction

In this report, I will discuss the design choices, calculations, and observations related to the LPC (Linear Predictive Coding) speech coding system.

Design Choices

1. Frame Length:
    - I chose a frame length of 180 samples (approximately 20 ms at a typical sampling rate of 8 kHz).
    - Short frames capture local spectral characteristics and allow for better modeling of speech.
2. Windowing:
    - I applied a Hamming window to each frame to reduce spectral leakage.
    - Windowing improves the accuracy of LPC analysis by emphasizing the central portion of the frame.
3. LPC Order:
    - I used an LPC order of 10.
    - Higher orders capture more details but increase the bit rate.
4. Quantization:
    - I quantized LPC coefficients, gain, and pitch period.
    - Uniform quantization was used for simplicity.
5. Voiced/Unvoiced Decision:
    - I used zero-crossing rate (ZCR) to determine voiced/unvoiced frames.
    - If ZCR exceeded a threshold, the frame was considered unvoiced.

Calculation of Bitrate

1. Gain Quantization:
    - 5 bits for gain (32 levels).
    - Bit rate for gain: 5 bits/frame.
2. Period Quantization:
    - 6 bits for pitch period (64 levels).
    - Bit rate for pitch: 6 bits/frame.
3. LPC Coefficient Quantization:
    - LAR coefficients: 5 bits each (32 levels).
    - PARCOR coefficients: 5 bits each (32 levels).

    - Total bit rate for LPC coefficients: 50 bits/frame.
4. Total Bitrate:
    - Total bitrate per frame: 5 (gain) + 6 (pitch) + 50 (LPC coefficients) = 61 bits/frame.

Observations and Modifications

1. Pitch Detection:
    - Accurate pitch detection is crucial for intelligibility.
    - Adjusted pitch period based on continuity between frames to avoid sudden changes.
2. Quantization Trade-offs:
    - Higher quantization levels improve quality but increase bitrate.
    - Balanced quantization to achieve a good trade-off.
3. Post Filtering:
    - Applied a post-filter to enhance perceptual quality.
    - Improved intelligibility by reducing spectral noise.

Judging the quality of reconstructed speech

I tried to listen to the generated .wav file manually but it was silent. I have calculated the residual in the encoding function and used it but it is not improving. I have also saved the encoded parameters in parameters.txt file and while decoding i passed this text file to decoder to reconstruct the speech. One reason for the speech could be silence is the way I have done the quantization.

Average processing delay

Average processing delay per frame is changing every time I run the encoder code but it is less than 50ms for sure.

Conclusion

LPC achieves good compression while maintaining intelligibility. Design choices, quantization, and pitch handling play critical roles. Continual adjustments and trade-offs are necessary to achieve optimal results.