Speech Signal Processing (EC5.408)

Assignment 2

January 22nd, 2025 - February 4th, 2025

Guidelines

- **Do not copy or plagiarise.** If you're caught for plagiarism, the penalty will range from zero in the assignment to an F grade in the course.
- **Always cite your sources** (be it images, papers, or existing libraries).
- **Mention clearly** if any assumptions are being considered.
- Only **MATLAB or Python** can be used for the coding part.
- For this assignment, you might use *Audacity* or *Wavesurfer* software.

Submission Format

Make a directory using the naming format `SSP_A2_RollNo`. The submission might include:

- **Codes** (`.py/.m`) to answer the coding problems.
- **Reports** (`.pdf`) to answer the theory questions.
- **Notebooks** (`.ipynb`) to answer both coding and theory questions together.

Place the files in their respective folders and zip the main directory using the naming format `SSP_A2_RollNo.zip` and upload this zip file to Moodle.

Example Directory Structure

```
SSP_A1_RollNo
├── Codes
│   ├── 1.py
│   └── 2.m
├── Audio
│   └── 1.wav
├── Report.ipynb
└── Report.pdf
```

Questions

I. Define the following:

1. Autocorrelation
2. Zero-Crossing Rate (ZCR)
3. Mel Spectrogram
4. Linear Prediction (LP) Spectrum

II. Explain the following:

1. Explain voiced and unvoiced speech? Explain any three different methods used for identifying them.
2. What is STFT? Explain the effect of window length and window shape.
3. Explain what is pitch contour. Do we need to do compute pitch at the frame level? If yes, are all frames considered? Explain.

III. Short-Time Fourier Transform

1. Plot the fourier transform of the entire signal and on a single frame of the signal. What are your observations based on the plots for each case?
2. Compute and plot the STFT of the given signal. You need to implement STFT using FFT from any suitable library.
3. Plot the STFT for different window lengths and window shapes. Analyze these plots and comment on the effect of window length and window shape on the resulting STFT.

IV. Voiced - Unvoiced Detection

1. Write a function for each of the methods described in above Section *II.1* for detecting voiced and unvoiced frames.
2. Plot the waveform of the given audio file and use the above functions to classify each frame of the audio file as either voiced or unvoiced.
3. Select one voiced frame and one unvoiced frame from the audio file. Plot their corresponding time-domain and frequency-domain spectra. Compare the voiced and unvoiced frames in both the time-domain and frequency-domain spectra.

—— ***End of Assignment*** ——