LAB – 3 REPORT

Name: Ozan Berk Gultegin

Submission Date:

# Purpose of the Work

The objective of this lab work is to analyze speech signals in the frequency domain using spectral and cepstral methods. The lab focuses on extracting fundamental frequency and identifying voiced, unvoiced, and background sound characteristics. By applying cepstral analysis and comparing it with autocorrelation results, deeper insights into pitch detection and speech structure are achieved.

# Work Results

II.I Time-Domain Waveforms

A graph with a blue line

AI-generated content may be incorrect.

II.II Frequency Spectra

A graph with blue lines

AI-generated content may be incorrect.

II.III Cepstral Analysis and Fundamental Frequency Estimation

A graph of a sound wave

AI-generated content may be incorrect.

II.IV Spectrogram and Sound Classification

A close-up of a colorful image

AI-generated content may be incorrect.

# Analysis

## III.I Time Domain Analysis

**Background Sound:**

* The signal appears dense and continuous throughout the whole duration.
* Lacks clear structure or silent intervals.
* The amplitude fluctuates within a consistent.

**Voiced Sound:**

* The waveform is centered within a shorter time window.
* The amplitude gradually increases to a peak and then decays.
* The central portion shows smoother, oscillating patterns.

**Unvoiced Sound:**

* The signal is also localized like Voiced sound.
* Compared to voiced sound, the waveform looks more irregular and less smooth.
* Although it has some energy, the signal appears noisier and sharper.

## III.II Frequency Domain Analysis

**Background Sound:**

* The spectrum is dominated by low frequency.
* The magnitude rapidly decreases as frequency increases, suggesting a broadband but low intensity.

**Voiced Sound:**

* Contains a strong low-frequency peak, followed by a series of descending harmonics.
* Peaks are spaced somewhat regularly.
* Energy is concentrated below 3 kHz.

**Unvoiced Sound:**

* The energy is more dispersed and irregular, with no clear harmonic structure.
* There are several peaks, but their spacing is inconsistent and their magnitudes are lower than those of the voiced signal.

## III.III Cepstral Analysis and Fundamental Frequency

**Background Sound:**

* The plot shows that the amplitude is constantly increasing and decreasing rapidly.
* There are two peaks in the quefrency. The first one at 0 ms, which is common, but the next one is at 10 ms.
* There is no periodicity which aligns with characteristics of background noise.
* The fundamental frequency is 100.46 Hz.

**Voiced Sound:**

* Amplitude changes are softer and less common than background sound.
* There are no clear peaks, so the amplitude is relatively similar.
* The fundamental frequency is 100.63 Hz.

**Unvoiced Sound:**

* Like Voiced sound, there are no peaks however the change on the amplitude is sharper.
* Unlike voiced sound however there are no harmonic or periodic structure to it.
* The fundamental frequency is the same Voiced sound, 100.63 Hz.

## III.IV Cepstral Analysis and Fundamental Frequency

**Background Sound:**

* Displays a consistent energy band, especially in the lower frequency range.
* Does not have any harmonic structures or pitch variations.
* This confirms the signal is non-speech background noise, aperiodic and spectrally flat.

**Voiced Sound:**

* Clear horizontal striations, especially in the lower frequencies.
* These horizontal bands correspond to harmonics of the fundamental frequency.
* Resonant frequency bands are clearly visible.

**Unvoiced Sound:**

* The unvoiced segment shows broad, noisy energy distribution, concentrated in lower frequencies.
* Lacks the periodicity and harmony of voiced signal.

**Classification:**

* **Voiced Sound:** Periodic, harmonically rich, concentrated at low to mid frequencies.
* **Unvoiced Sound:** Aperiodic, noisy, more energy in mid-range frequencies.
* **Background Sound:** Flat spectrum with high energy all over the place. Displays a much higher frequency range than both voiced and unvoiced sound.

# Summary and Conclusion

In this lab work, a comprehensive analysis of speech signals was performed across the time domain, frequency domain, and cepstral domain, complemented by spectrogram-based classification of voiced, unvoiced, and background noise segments.

**Time-domain analysis:**

Revealed distinct waveform characteristics for each signal type. Voiced sounds had a smoother and harmonic waveform, unvoiced sounds showed energy bursts, and background noise appeared as a dense, continuous signal.

**Frequency-domain analysis:**

Provided insight into energy distribution. Voiced signal did not show multiple peaks occasional increase and decrease while Unvoiced signal showed multiple peaks and frequent magnitude change. On the other hand, Background Noise signals magnitude decreased immediately to near 0 almost immediately.

**Cepstral analysis:**

Helped to estimate fundamental frequency. While voiced and unvoiced signal did not show any peaks and kept their amplitude like each other, background noise signals amplitude changed constantly and unpredictably while occasionally peaking on high values.

**Spectrogram analysis:**

Voiced signal showed strong harmonic and periodic bands, while unvoiced signal showed aperiodic characteristics with high noise on lower frequencies. Lastly, the background noise signal showed flat energy levels across the duration and frequency range.