ELEC5305 Project Proposal

Real-Time speech Enhancement Using FFT-Based Noise Suppression

Joshua Chao-Hsu Chen

510656681

JoshuaCHChen

https://github.com/JoshuaCHChen/elec5305-project-510656681

Project Overview

The aim of this project is to design and implement a MATLAB-based speech enhancement system using FFT-based noise suppression techniques. The primary problem that we're addressing is the degradation of speech intelligibility in noisy environments, such as public transport, lecture halls, and crowded streets. Existing communication platforms often suffer from reduced clarity when background noise overwhelms the speech signal.

The proposed solution is to develop a lightweight, real-time noise reduction pipeline that operates in the frequency domain. By leveraging the FFT, spectral subtraction, and adaptive filtering, the system will attenuate noise components while preserving the underlying speech. The project focuses on methods that are computationally efficient enough to be deployed on mobile or embedded platforms.

Background and Motivation

Background noise significantly reduces the intelligibility and quality of speech. While simple FIR filters can attenuate broadband noise, they often distort speech and are less effective when noise overlaps with the signal in frequency. To address this, frequency-domain approaches such as spectral subtraction have been developed.

Spectral subtraction was first proposed by Boll (1979) and remains a cornerstone of speech enhancement research. The method assumes noise can be estimated during non-speech segments and then subtracted from noisy speech in the spectral domain. The ELEC5305 labs already introduced the core building blocks for such a system: FFT/DFT computation, noise simulation, FIR filtering, and STFT spectrogram analysis.

The motivation for this project is to extend these lab exercises into a complete noise suppression system that not only demonstrates theoretical principles but also produces tangible improvements in speech quality. It is also particularly relevant to resource-constrained applications such as mobile communication, low-power IoT devices, and hearing aids, where transparency and efficiency are more important than black-box performance.

This project provides a direct link between theory and practice, and being able to apply these techniques to real-world recordings highlights its practical value. By working with noisy recordings and reconstructing enhanced speech, we will be able to emphasize how signal processing fundamentals serve as the backbone of modern communication systems, reinforcing the concepts of lab materials while addressing an issue that is immediately relatable to everyday experience.

Proposed Methodology

1. Data Acquisition
    - Clean speech samples from open datasets (e.g. TIMIT, LibriSpeech)
    - Additional recordings captured via microphone to simulate real-world conditions
2. Noise Simulation
    - Additive white Gaussian noise introduced using MATLAB's awgn() function
    - Environmental noise samples (traffic, restaurant noise) mixed with clean speech
3. STFT Analysis
    - The noisy signal is transformed using STFT to provide a time-frequency representation
    - The STFT is defined as:

      $$X(m,k) = \sum_{n=-\infty}^{\infty} x[n]\varpi[n-mR]e^{-j\frac{2\pi}{N}kn}$$

      where $\varpi[n]$ is a window function, R is the hop size, and k indexes frequency bins
4. Noise Spectrum Estimation
    - Noise estimated during silent segments or using minimum statistics
    - A running average of spectral magnitudes used for smoother noise estimation
5. Spectral Subtraction
    - The enhanced magnitude spectrum is computed as:
      $|Y(m,k)| = \max(|x(m,k)| - |N(k)|, 0)$
    - Negative values floored to avoid musical noise artifacts
6. Reconstruction

    - Enhanced magnitude combined with original phase
    - Inverse STFT applied to reconstruct the time-domain enhanced speech
7. Evaluation
    - Objective metrics: Signal-to-Noise Ratio (SNR) improvement before and after enhancement
    - Subjective metrics: Spectrogram inspection and informal listening tests

Tools: MATLAB Live Script

Expected Outcomes

- A working MATLAB prototype that performs spectral subtraction-based noise suppression on speech recordings
- Enhanced speech signals with measurable improvement in SNR relative to noisy inputs
- Spectrograms comparing clean, noisy, and enhanced speech
- GitHub repository including:
    - MATLAB code
    - Documentation of methodology and results
    - Sample audio files demonstrating input vs. output

Timeline (Weeks 6-13)

Weeks	
6-7	Literature review on spectral subtraction, dataset collection
8-9	Implement STFT-based enhancement pipeline; initial testing on synthetic noisy data
10-11	Optimize noise estimation; evaluate with real-world noise samples
12-13	Finalize report, complete GitHub repository with audio demos, documentation, and plots

References

1. Boll, S.F. (1979). “Suppression of Acoustic Noise in Speech Using Spectral Subtraction.” *IEEE Transactions on Acoustics, Speech, and Signal Processing*.
2. Loizou, P. C. (2013). “Speech Enhancement: Theory and Practice.” *CRC Press*.
3. Ephraim, Y., & Malah, D. (1984). “Speech enhancement using a minimum-square error short-time spectral amplitude estimator.” *IEEE Transactions on Acoustics, Speech, and Signal Processing*.
4. Vary, P. (1985). “Noise suppression by spectral magnitude estimation –mechanism and theoretical limits.” *Signal Processing*.