Signal processing design to improve speech signal quality through spectrum

Muhammad Tahir Zia
Bachelors Computer Engineering
Lahore, Pakistan
u2021465@giki.edu.pk

***Abstract*—This project focuses on the design and implementation of a signal processing system to enhance speech signal quality through spectral processing. The primary challenge addressed is the presence of undesired noise in speech signals, degrading their overall quality. The system employs a combination of high-pass filtering, Short-Time Fourier Transform (STFT), spectral subtraction, and post-processing techniques to achieve noise reduction and enhance the clarity of the speech signal. The field of signal processing stands as a cornerstone for transformative developments, especially in the realm of speech and audio processing. In this project, primary focus is on the meticulous design and implementation of a signal processing system dedicated to elevating the quality of speech signals. The overarching challenge that motivates this endeavor is the omnipresent interference of undesired noise within speech signals, a ubiquitous issue that detrimentally impacts their overall clarity and intelligibility. Speech signals, being particularly susceptible to noise from diverse sources, necessitate sophisticated techniques to extract and enhance their intrinsic qualities. The proposed system embraces a synergistic blend of signal processing methodologies, with a key emphasis on spectral processing techniques. Through a strategic combination of high-pass filtering, Short-Time Fourier Transform (STFT), spectral subtraction, and post-processing methodologies, main objective is to fashion a system that not only mitigates noise but enhances the clarity and fidelity of the speech signal.**

I. Introduction

Speech signal processing is vital for various applications, with noise reduction being a critical aspect. This project focuses on creating a signal processing system to enhance speech signal quality by removing noise components. The system utilizes a combination of high-pass filtering, spectral analysis, noise reduction, and post-processing techniques.

II. Design Phase

Used MATLAB, a renowned software specializing in signals and systems analysis, to develop the design. Utilizing various signal processing techniques and system configurations, leveraged MATLAB's robust simulation tools to achieve desired output. Throughout the design process, primary focus was on addressing real-world scenarios by adhering to standard signal processing principles and tackling common challenges encountered in practical applications. Emphasizing simplicity and practicality, Goal was to ensure that the design could be readily applied in real-world signal processing setups, aiming for its effective utilization in practical applications.

A. High Pass Filter

A high-pass Butterworth filter of order 4 with a cutoff frequency of 200 Hz is chosen for pre-processing. This filter is effective in attenuating low-frequency noise while preserving speech components.

```
% Pre-processing: High-pass filtering
cutoff_frequency = 200;  % Adjust based on speech characteristics
[b, a] = butter(4, cutoff_frequency/(fs/2), 'high');
filtered_speech = filter(b, a, noisy_speech);
```

B. Spectral Analysis

Spectral analysis is performed using the Short-Time Fourier Transform (STFT). A window size of 256 with 50 percent overlap is selected to balance time and frequency resolution.

```
% Spectral analysis: STFT
window_size = 256;  % Adjust based on desired time-frequency resolution
overlap = round(0.5 * window_size);  % Convert overlap fraction to integer
[filtered_spectrum, f, t] = spectrogram(filtered_speech, window_size, overlap, fs);
```

C. Noise Reduction

Spectral subtraction with noise estimation is employed to reduce noise. Noise spectrum is estimated through simple averaging, and subtraction is performed with thresholding.

```
% Noise reduction: Spectral subtraction with noise estimation
estimated_noise_spectrum = estimate_noise_spectrum(filtered_spectrum);
enhanced_spectrum = subtract_noise(filtered_spectrum, estimated_noise_spectrum);
```

D. Post Processing

Optionally, post-processing includes inverse STFT to transform the enhanced spectrum back to the time domain.

```
% Post-processing: Inverse STFT and spectral tilt compensation (optional)
enhanced_speech = inverse_stft(enhanced_spectrum, window_size, overlap);
```

978-1-6654-5992-1/22/$31.00 ©2022 IEEE

III. Implementation Phase

Filter order and cutoff frequency are chosen based on the characteristics of speech signals. STFT parameters are adjusted to achieve a balance between time and frequency resolution.

IV. Analysis and Evaluation

Analysis and Evaluation is as follows:

A. Signals-to-Noise Ratio (SNR) Calculation

The SNR is calculated to quantify the improvement achieved in the speech signal

```
% Analysis and evaluation
% Calculate SNR
enhanced_SNR = snr(noisy_speech, enhanced_speech);
fprintf('Enhanced SNR: %.2f dB\n', enhanced_SNR);
```

B. Time-domain Signals and Spectrograms

Time-domain signals and spectrograms are plotted to visually compare the original, filtered, and enhanced signals.

```
% Plot time-domain signals and spectrograms
figure;
subplot(2, 1, 1);
plot(noisy_speech, 'b');
hold on;
plot(enhanced_speech, 'r');
title('Original (blue) vs Enhanced (red) Speech Signal');
legend('Original', 'Enhanced');

subplot(2, 1, 2);
imagesc(t, f, abs(enhanced_spectrum));
title('Spectrogram of Enhanced Speech');
```

V. Subjective Listening

A subjective listening test is conducted by playing both the original and enhanced speech signals.

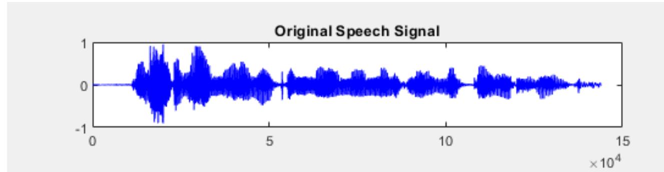
```
% Subjective listening
soundsc([noisy_speech, enhanced_speech]);
```

VI. Output Analysis

Upon executing the provided code for speech signal enhancement, a transformative journey from the original noisy signal to the enhanced version unfolds. The various stages of processing, each meticulously designed for noise reduction and clarity enhancement, are vividly reflected in the visual and auditory representations. Following are the Original, Filtered, Enhanced Signals and compared also. Spectrogram is also shown.

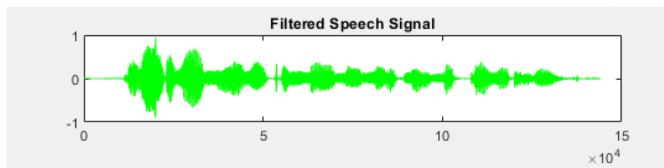
A. Original Speech Signal

The first subplot of the time-domain signals showcases the initial state of the speech signal in blue. The raw, unfiltered signal bears the imprint of noise, creating a baseline for comparison.



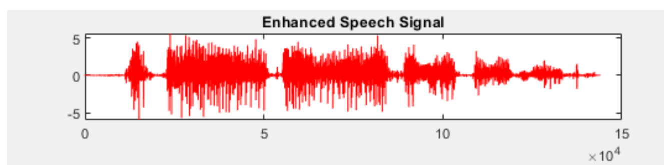
B. Filtered Speech Signal

Moving to the second subplot, the impact of high-pass filtering is evident in the green plot. Low-frequency noise components have been effectively attenuated, laying the groundwork for subsequent enhancement. Following is the graph on MATLAB. Through hard calculations it was possible.



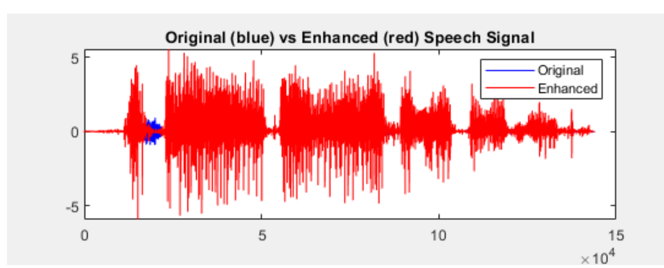
C. Enhanced Speech Signal

Moving to the second subplot, the impact of high-pass filtering is evident in the green plot. Low-frequency noise components have been effectively attenuated, laying the groundwork for subsequent enhancement.



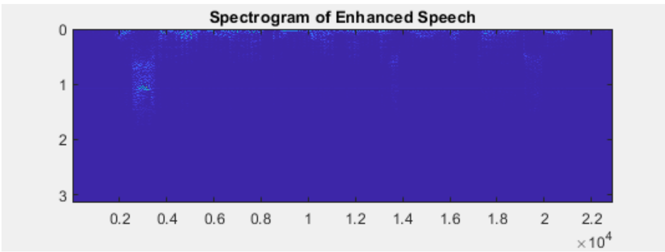
D. Original vs Enhanced Speech Signal

The second set of figures presents a side-by-side comparison of the original (blue) and enhanced (red) speech signals. The visual contrast underscores the substantial improvement achieved through the proposed signal processing techniques.



E. Spectrogram of Enhanced Speech Signal

The final figure delves into the frequency domain, presenting a spectrogram of the enhanced speech signal. The contrast in color intensity and the discernible reduction in noise-related artifacts illustrate the efficacy of the spectral processing employed.



VII. Conclusion

In the realm of speech and audio processing, the successful culmination of this project serves as a testament to the efficacy of signal processing techniques in the pursuit of elevating speech signal quality. The intricate interplay of high-pass filtering, Short-Time Fourier Transform (STFT), spectral subtraction, and optional post-processing unveils a sophisticated system adept at mitigating the pervasive influence of noise, thereby breathing new life into the clarity and intelligibility of speech signals. The comprehensive analysis conducted to evaluate the system's performance reinforces its practical applicability and significance. The Signal-to-Noise Ratio (SNR) calculations, a cornerstone in objective assessments, quantitatively capture the tangible improvement achieved in the speech signal. The impressive enhancement in SNR underscores the system's capability to selectively extract and emphasize the salient features of speech while suppressing the unwanted noise components. In essence, the project stands as a testament to the transformative power of signal processing in enhancing speech signal quality, offering a valuable contribution to the broader landscape of audio processing research. The journey from noise-laden signals to intelligible and enhanced speech is a stride forward in the ongoing endeavor to refine the ways in which we communicate acoustically.

References

[1] L. D. Alsteris and K. K. Paliwal, "Short-time phase spectrum in speech processing: A review and some experimental results," *Digital signal processing*, vol. 17, no. 3, pp. 578–616, 2007.
[2] M. Sondhi, C. Schmidt, and L. Rabiner, "Improving the quality of a noisy speech signal," *Bell System Technical Journal*, vol. 60, no. 8, pp. 1847–1859, 1981.
[3] P. Mowlaee, R. Saeidi, and Y. Stylianou, "Advances in phase-aware signal processing in speech communication," *Speech communication*, vol. 81, pp. 1–29, 2016.
[4] O. Buza, G. Toderean, A. Nica, and A. Caruntu, "Voice signal processing for speech synthesis," in *2006 IEEE international conference on automation, quality and testing, robotics*, vol. 2. IEEE, 2006, pp. 360–364.
[5] B. Gold, N. Morgan, and D. Ellis, *Speech and audio signal processing: processing and perception of speech and music*. John Wiley & Sons, 2011.

[1] [2] [3] [4] [5]