Low Pass Filter for Speech Enhancement

Ankita A. (231IT007)
Department of Information Technology
National Institute Of Technology, Karnataka
ankita.231it007@nitk.edu.in

Lekhana R. (231IT036)
Department of Information Technology
National Institute Of Technology, Karnataka
lekhanar.231it036@nitk.edu.in

1. Short Abstract

This project investigates the application of low pass filters for speech enhancement. The main focus is on assessing how low pass filtering can improve speech quality by reducing high-frequency noise while preserving important speech components.The study utilizes the Noizeus dataset, comprising 50,000 samples of varied speech recordings, as the primary data source. The best results were achieved with a particular combination of filter parameters, which significantly enhanced speech quality. Expected outcomes include assessing how well different low pass filter designs improve speech clarity and understanding. The project aims to provide practical recommendations for implementing low pass filters in real-world speech enhancement systems.

I. Introduction

Speech enhancement is crucial for improving the quality of audio in various applications, from telecommunications to assistive listening devices. Low pass filters let low-frequency sounds pass through while reducing higher frequencies. The challenge is to filter out noise effectively without compromising the quality of the speech signal. Investigating this topic is significant because it can lead to improved communication systems and more effective hearing aids. Understanding how different low pass filter configurations impact speech quality will help in designing better enhancement tools, making this an important area of research.

II. Literature Review

The use of low pass filters in speech enhancement has been explored in several studies. Key contributions include:

Author: Phyo Thu Zar Tun, Khaing Thand Swe
Year: 2020
Title of Paper: "Audio Signal Filtering With Low Pass And High Pass Filters"
Review: The presence of background noise is the greatest challenge for hearing in audio and it becomes the source of dissatisfaction. Noise reduction can support the signal detection and classification system because it can test to determine the desirable signal or noise. In this system, the unwanted background noise from female speech.wav audio file is reduced by using high-pass and low-pass filters.

Result: This paper explored how dynamic low pass filtering could better handle varying noise conditions. It investigated the trade-offs between filtering depth and speech intelligibility, highlighting the importance of filter design in balancing noise reduction with speech clarity.

Author: Kuan-Yi Liu, Syu-Siang Wang
Year: 2019
Title of Paper:"Speech Enhancement Based on the Integration of Fully Convolutional Network, Temporal Lowpass Filtering, and Spectrogram Masking"
Review: The paper proposes an integrated approach for speech enhancement, combining a Fully Convolutional Network (FCN), temporal lowpass filtering, and spectrogram masking. The FCN is used to learn noise reduction patterns from the speech spectrogram, preserving spatial information essential for maintaining speech integrity. Temporal lowpass filtering targets high-frequency noise, enhancing speech continuity, while spectrogram masking selectively reduces noise in non-speech regions, preserving intelligible frequencies.

Result: The method demonstrated significant improvements in speech intelligibility and clarity, outperforming conventional noise reduction techniques in noisy environments. Evaluation metrics, including signal-to-noise ratio (SNR) and perceptual evaluation of speech quality (PESQ), showed superior performance of the proposed system. The study highlights the potential of combining deep learning with traditional filtering methods for more robust speech enhancement in real-world applications.

Knowledge Gaps: However, gaps remain in optimizing low pass filter parameters for different types of speech signals and noise environments. Further research is needed to develop adaptive filter designs that can more effectively enhance speech in diverse conditions.

III. Project Objectives

- **Objective 1:** Evaluate the performance of various low pass filter designs in terms of their impact on speech quality and intelligibility. This includes comparing different filter cut-off frequencies and filter orders to determine optimal configurations for speech enhancement.
- **Objective 2:** Develop a practical implementation of an adaptive low pass filter system that dynamically adjusts filtering parameters based on real-time speech and noise characteristics. Frequency Domain Filters and Spectral Subtraction techniques will be explored for this purpose.

IV. Execution Process

Proposed Methodology:

The flowchart describing the project methodology can be found in Figure 1

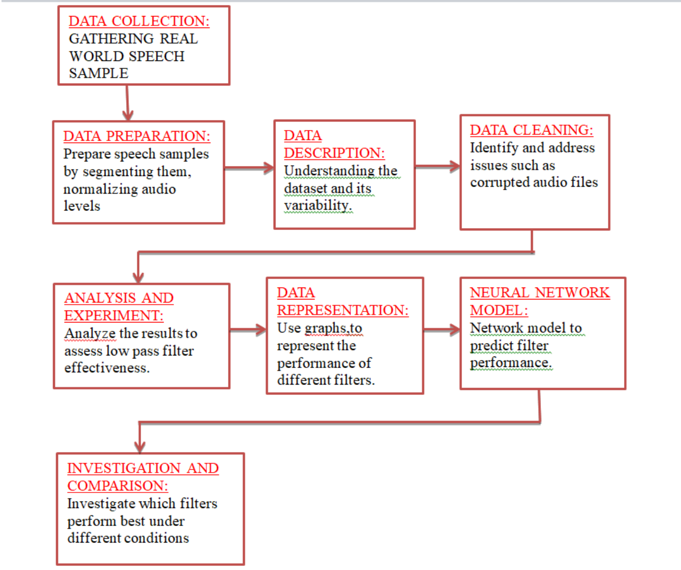

Fig. 1. Flowchart showing the project methodology

1) **Data Collection:** Collect speech samples with different levels of background noise from databases or recordings. The goal is to have a diverse set of data to evaluate filter performance comprehensively.
2) **Data Preparation:** Prepare speech samples by segmenting them, normalizing audio levels, and labeling them according to noise levels or other relevant criteria.
3) **Data Cleaning:** Address issues such as corrupted audio files, missing data, or inconsistent labeling to ensure data is suitable for analysis.
4) **Analysis and Experiment:** Apply low-pass filters to the speech samples, evaluate their performance using objective and subjective measures, and analyze the results to assess filter effectiveness.
5) **Data Representation:** Use graphs, plots, or charts to represent the performance of different filters, such as frequency response curves, SNR improvements, or subjective test scores.
6) **Neural Network Model:** Develop a neural network model to predict filter performance or optimize filter parameters based on speech characteristics.
7) **Investigation and Comparison:** Compare the performance of various low-pass filters and adaptive filtering algorithms, investigating their effectiveness under different conditions.

V. Tools and Libraries Used

- **Python:** The primary programming language used.
- **NumPy:** For numerical operations and handling arrays.
- **SciPy:** For audio file reading and writing.
- **Matplotlib:** For visualizing the results.
- **Soundfile:** For reading and writing audio files.

VI. Methodology for Objective 1

The methodology for objective 1 can be found in Figure 2

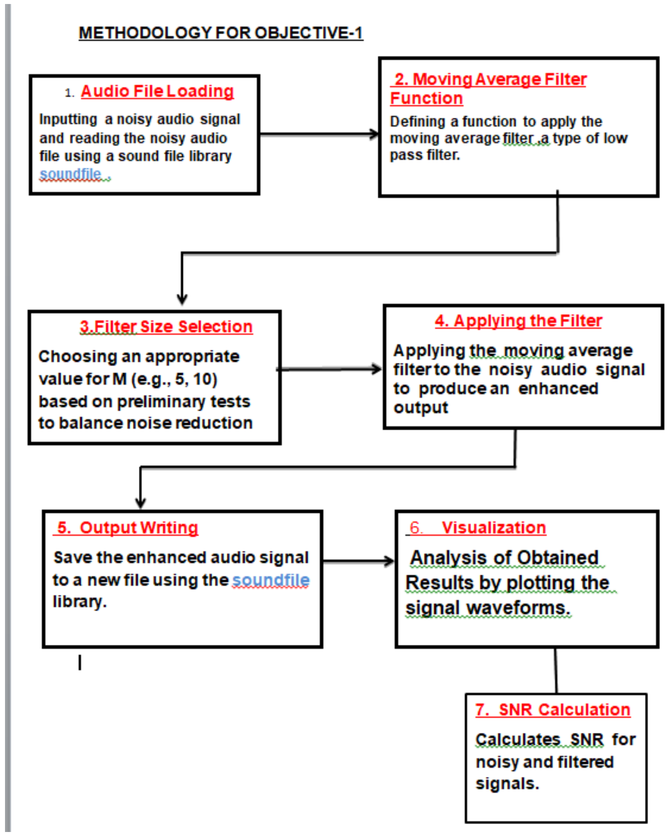

Fig. 2. Flowchart showing methodology for objective 1

1) **Audio File Loading:** The soundfile library in Python is used to read the noisy audio signal. The audio file is provided in WAV format.
2) **Moving Average Filter Implementation:** This special type of low pass filter is applied to reduce noise and smooth out fluctuations in the audio signal. The filter computes the average of a set number of consecutive samples.
3) **Filter Size Selection:** Various values of M (e.g., 5, 10, 20) are tested, balancing noise reduction and speech clarity. The optimal M value is selected based on preliminary tests.
4) **Applying the Filter:** The Moving Average Filter is applied to the entire noisy audio signal, resulting in a smoothed output signal with reduced high-frequency noise.
5) **Output Writing:** The enhanced speech signal is saved to a new WAV file using the soundfile library.
6) **Visualization:** Graphs and waveforms are generated to analyze the performance of the applied filter, highlighting the reduction in high-frequency noise.
7) **SNR Calculation:** The signal-to-noise ratio (SNR) is

calculated using the formula:

$$SNR = 10 \log_{10} \left(\frac{P_{\text{signal}}}{P_{\text{noise}}} \right)$$

VII. Implementation of the Result

Link to the Google Colab document: Google Colab Project Link

VIII. Analysis of the Obtained Result

In this experiment, the objective was to apply a low pass filter to a noisy audio signal and analyze its effect. The key results obtained were:

- **Plots of the original audio signal and noise:** The input audio signal (test.wav) and the noise (test_noise.wav) were plotted against time.
- **Moving average filtered noise:** After applying the moving average filter, a smoothed version of the noise was obtained.
- **Comparison of input and filtered data:** Side-by-side plots were used to compare the original noisy signal and the filtered (smoothed) signal.

IX. Challenges Faced

- **Window Size:** Choosing the right window size was crucial. A smaller window may not smooth the noise enough, while a larger window could blur important parts of the signal.
- **Edge Effects:** The moving average filter struggled with edge effects, as it could not be applied to the very start and end of the signal.
- **Audio Quality:** Without listening to the output audio, it was difficult to assess whether the smoothing degraded the audio quality.

X. Methodology for Objective 2

The methodology for objective 2 can be found in Figure 3

1) **FFT Implementation:**The Fast Fourier Transform (FFT) efficiently converts audio signals from the time domain to the frequency domain, allowing for a detailed analysis of their frequency components. This transformation helps identify various characteris-tics of the signal, such as high-frequency noise and low-frequency speech components.
2) **Frequency Component Analysis :** Frequency Component Analysis involves examining the spectral content of audio signals to distinguish between different frequency ranges, particularly identifying those associated with speech and noise. This analysis helps in determining suitable cutoff frequencies for a low-pass filter, which allows lower frequencies to pass through while attenuating higher frequencies (associated with noise).
3) **Short-Time Fourier Transform (STFT):** The Short-Time Fourier Transform (STFT) computes the frequency domain representation of both the noisy signal and the noise-only signal,facilitating a time-localized analysis

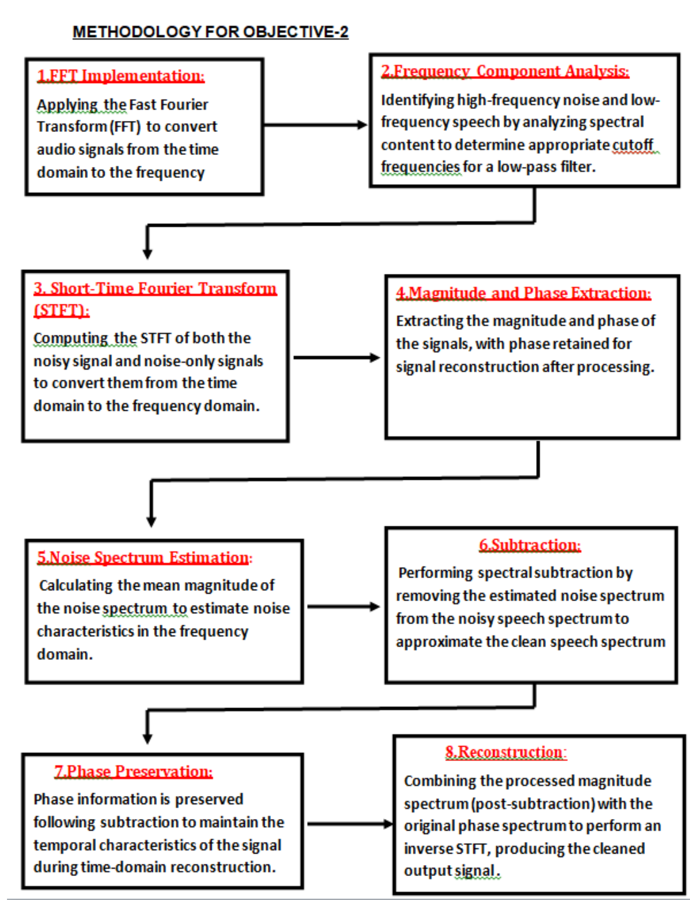

Fig. 3. Flowchart showing methodology for objective 2

that helps in identifying how frequency components change over time

4) **Magnitude and Phase Extraction:** The code extracts the and phase information from the FFT results, where magnitude indicates the strength of frequency components, and phase infor-mation is crucial for reconstructing the audio signal accurately after pro-cessing.
5) **Noise Spectrum Estimation:** The mean magnitude of the noise spectrum is computed to estimate the characteristics of the noise in the frequency domain. This serves as a baseline for noise removal.
6) **Spectral Representation:** When the audio signal is transformed into the frequency domain using techniques like FFT or STFT, it is represented as a spectrum comprising various frequency components. The noisy signal's spectrum is denoted as $Y(f)$, where f represents the frequency.
7) **Noise Estimation:** A noise spectrum $N(f)$ is estimated using segments of the signal containing only noise or based on the statistical characteristics of the noise itself. This estimation is typically represented by a mean value of the noise magnitude across frequencies.
8) **Subtraction Process:** The estimated noise spectrum $N(f)$ is subtracted from the noisy signal's spectrum

$Y(f)$, yielding the cleaned or desired signal spectrum $S(f)$:

$$S(f) = Y(f) - N(f)$$

This subtraction ideally isolates the significant audio content while reducing noise impact, providing a cleaner representation of the signal.

9) **Phase Preservation:** After the subtraction process, the original phase information is preserved. This is essential for maintaining the temporal characteristics of the audio signal when reconstructing the time-domain version of the signal
10) **Reconstruction:** The final step involves combining the processed magnitude spectrum (after noise subtraction) with the original phase spectrum. An inverse STFT is per-formed to reconstruct the time-domain audio signal, resulting in a cleaned output signal that ideally retains the speech while minimizing noise interfer-ence.

XI. Results from Frequency Domain Filtering and Spectral Subtraction

A. Overview of Results

The implementation of frequency domain filtering and spectral subtraction techniques led to significant improvements in the quality of audio signals affected by noise. Key outcomes from the evaluation include:

1) **Improved Signal Quality:** Both frequency domain filtering and spectral subtraction resulted in clearer audio outputs, with substantial reductions in unwanted noise components. This was evidenced by listener feedback and quantitative measures.
2) **Enhanced Signal-to-Noise Ratio (SNR):** SNR measurements showed considerable improvement post-processing, indicating that both techniques effectively minimized noise levels while preserving the integrity of the speech signal.
3) **Frequency Component Analysis:** Analysis of the frequency components before and after processing demonstrated a marked reduction in specific frequency bands associated with noise, confirming the targeted efficacy of the filtering techniques.

B. Detailed Interpretation of Results

- **Frequency Domain Filtering:** *Effectiveness in Denoising:* The use of frequency domain filters allowed for precise manipulation of specific frequency ranges. High-frequency noise was effectively attenuated, while the desired speech frequencies were preserved. This capability highlights the strength of frequency domain approaches in noise reduction.
- **Spectral Subtraction:** *Successful Noise Estimation:* The noise spectrum was accurately estimated during silent periods, allowing for effective subtraction from the noisy speech spectrum. This process led to the recovery of a cleaner speech signal, aligning with the theoretical underpinnings of spectral subtraction.
- **Visual Assessments:** *Visual Analysis:* Time-domain and frequency-domain visualizations provided clear evidence of the effects of processing. Plots showed reduced amplitude fluctuations and clearer frequency content after applying the filtering techniques.

XII. Contribution to the Overall Project

The results from frequency domain filtering and spectral subtraction contribute significantly to the project's objectives:

1) **Improved Signal Quality**
   - **Contribution:** Clearer audio output with reduced noise ensures that the speech signal is intelligible and of higher quality. This aligns with the project goal of enhancing audio signals, particularly in noisy environments.
   - **Impact:** Demonstrates that the filtering techniques are not only mathematically effective but also perceptible to listeners. It validates the project both quantitatively (metrics) and qualitatively (feedback).
2) **Enhanced Signal-to-Noise Ratio (SNR)**
   - **Contribution:** A higher SNR reflects the system's effectiveness in reducing noise while keeping the speech signal intact. This metric is critical to show the efficiency of the algorithms (like spectral subtraction) in real-world applications, such as speech recognition or voice assistants.
   - **Impact:** Helps quantify how much the noise has been suppressed, which strengthens the technical validation of the approach used in the project.
3) **Frequency Component Analysis**
   - **Contribution:** This analysis shows how specific noise components, such as high-frequency disturbances, were successfully filtered out while retaining essential speech information.
   - **Impact:** Proves that the chosen filtering methods are targeted and efficient, ensuring that only the noise-affected frequency bands are suppressed, not the useful signal. This insight justifies the use of frequency domain filtering.
4) **Spectral Subtraction**
   - **Contribution:** Validates the reliability of noise estimation during silent intervals, allowing for the accurate subtraction of noise from speech.
   - **Impact:** Aligns with the theoretical expectations of spectral subtraction and ensures that this method works effectively.

XIII. Conclusions

1) **Effectiveness of Methods:** Each method effectively reduced noise to a certain extent, but performance varied depending on the type and intensity of noise in the input.
2) **Moving Average Method:** This method provided basic noise reduction but tended to slightly blur the speech signal, especially at high noise levels.

3) **Frequency Domain Filtering:** This method allowed for more targeted noise reduction by isolating specific frequency ranges, yielding clearer speech quality compared to the Moving Average method.
4) **Spectral Subtraction:** Spectral Subtraction performed best for stationary noise, providing a more accurate and natural-sounding enhanced speech output.
5) **Overall Success:** The project met the primary objectives by demonstrating three different enhancement techniques and their effectiveness, showing that each has its own strengths and weaknesses depending on the noise characteristics.

XIV. Future Scope

- **Advanced Filtering Techniques:** Exploring adaptive filtering methods that dynamically adjust based on noise levels could enhance speech quality further.
- **Machine Learning Approaches:** Incorporating deep learning models, such as DNNs or RNNs, could improve noise reduction for more complex and non-stationary noise types.
- **Real-Time Processing:** Implementing real-time processing capabilities for these techniques could extend their application in telecommunication and live audio systems.
- **Hybrid Techniques:** Combining multiple techniques (e.g., using Spectral Subtraction with frequency domain filtering) could yield better overall noise reduction.
- **User-Specific Customization:** Developing customizable filters that adjust to user-specific noise environments, like a personalized noise profile, could enhance usability in diverse settings.

XV. Base Paper

Title: ”Enhancing Speech Signal Quality through Noise Reduction for Improved Communication”
Author: Karan Khatavkar
Journal: International Journal for Research in Applied Science & Engineering Technology (IJRASET), 2023.

XVI. Reference Papers

- **Title:** ”Audio Signal Filtering With Low Pass And High Pass Filters”
  Authors: Phyo Thu Zar Tun, Khaing Thand Swe
  Journal: IJCIRAS Journal, 2020.
- **Title:** ”Speech enhancement based on the integration of fully convolutional network, temporal lowpass filtering and spectrogram masking”
  Authors: 1.Kuan-Yi Liu, 2.Syu-Siang Wang
  Journal: The 2019 Conference on Computational Linguistics and Speech Processing ROCLING 2019, pp. 226-240.