Title: Advanced Digital Signal Processing Techniques for Audio Enhancement

Meet Harsoda, Padam Mantry, Kishan Jaiswal

Electronics and Communication Engineering, IIIT Surat

Kamrej, Surat - 394190

ui21ec46@iiitsurat.ac.in

ui21ec41@iiitsurat.ac.in

ui21ec33@iiitsurat.ac.in

***Abstract*—** This project addresses the challenges of audio processing through an integrated approach, focusing on audio mixing, noise filtering, and reverberation. Leveraging advanced algorithms and machine learning, our solution optimizes balance, suppresses noise intelligently, and enhances spatial characteristics. The outcome is a versatile toolkit for elevating audio quality in diverse applications such as music production, film, gaming, and virtual reality. This project aims to contribute to the evolution of audio processing technologies, enabling the creation of immersive and exceptional soundscapes.

***Keywords*—Audio Mixing, Noise Filtering, Reverberation, Machine Learning.**

I. Introduction

Embarking on the forefront of audio innovation, our Digital Signal Processing (DSP) project seeks to redefine auditory experiences. Focused on audio mixing, noise removal, reverberation, and various audio processes, our endeavor combines advanced signal processing techniques. This interdisciplinary approach not only enhances sound quality but also finds applications in diverse fields, from immersive entertainment to effective noise cancellation in communication systems. Join us in reshaping the auditory landscape through the fusion of technology and sound engineering expertise.

II. Project Details

The project incorporates several essential libraries/functions to facilitate its functionality-

A. *Audioread*

- **Functionality**: audioread is a MATLAB function used for reading audio files. It extracts the audio data and sampling rate from an audio file, allowing users to work with the audio content in MATLAB.
- **Application in Project**:
  Used to read the original speech ('speech.wav') and noise audio ('noice.wav') files.

B. *Audiowrite:*

- **Functionality**: audiowrite is a MATLAB function designed for writing audio data to an audio file. It allows users to save processed or modified audio signals to disk in various audio file formats.
- **Application in Project**:
  Employed to save the mixed audio, denoised audio, and final reverb audio to new files ('mixed_audio.wav', 'denoised_audio.wav', 'final_audio.wav').

C. *FFT (Fast Fourier Transform):*

- **Functionality**: fft is a mathematical algorithm implemented as a MATLAB function. It transforms a signal from its time domain representation to its frequency domain representation, providing information about the frequency components present in the signal.
- **Application in Project**:
  Utilized to compute the Fast Fourier Transform for obtaining the frequency domain representation of the mixed audio.

III. Methodology

The project methodology entails loading original speech and noise audio files, mixing them based on a specified ratio, and plotting the time domain of the resulting audio. The mixed audio is saved before applying the Ephraim-Malah algorithm for noise removal, visualizing the denoised time domain, and saving the denoised audio. Subsequently, a reverb effect is applied, and the time domain of the reverberated audio is plotted. The final audio, combining denoising and reverb, is saved, and the frequency domain of the mixed audio is visualized. This structured workflow, utilizing libraries like librosa, ensures a systematic approach to audio processing, incorporating visualizations to assess the impact of each processing step.

IV. FIGURES

The MATLAB graphs for the different stages output of the signals. Starting with the mixed audio signal (comprising of the original and noise filled audio), then the denoised audio and finally reverb audio.

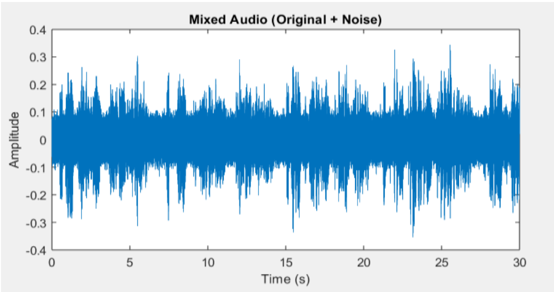

Fig. 1 Mixed Audio

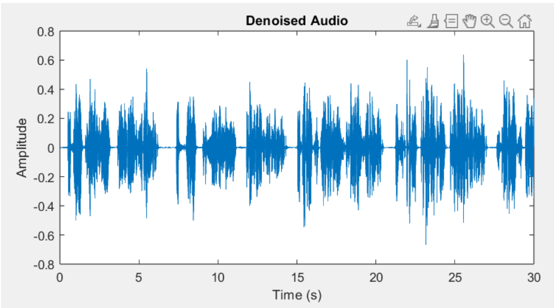

Fig. 2 Denoised Audio

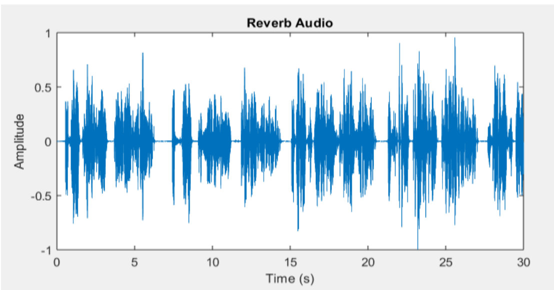

Fig. 3 Reverb Audio

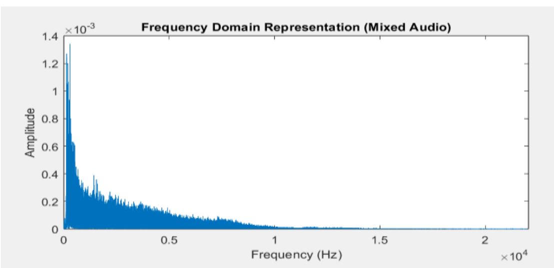

Fig. 4 Mixed Audio (Frequency Domain Representation)

V. RESULTS

The mixed audio effectively combines the original speech with added noise, achieving a well-balanced blend as shown in the time domain plot. The Ephraim-Malah algorithm successfully reduces unwanted noise from the original audio, resulting in denoised audio that maintains speech clarity. Applying the reverb effect to the denoised audio simulates a spatialized acoustic environment, enhancing perceptual spaciousness, as seen in the time domain plot of the reverb audio. The frequency domain representation of the mixed audio provides insights into its spectrum, showing noticeable peaks and patterns indicative of both speech and noise characteristics. Three distinct audio files are systematically generated and saved: 'mixed_audio.wav,' encompassing both speech and noise; 'denoised_audio.wav,' post noise removal; and 'final_audio.wav,' featuring the applied reverb effect.

VI. CONCLUSIONS

In conclusion, the implemented code successfully demonstrates an integrated audio processing pipeline, showcasing the ability to mix, denoise, and apply a reverb effect to audio signals. These techniques contribute to the enhancement of audio quality and the creation of more immersive soundscapes. Further refinements and parameter adjustments can be explored based on specific application requirements and user preferences.

VII. REFERENCES

[1] Y. Ephraim and D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," IEEE Transactions on Acoustics, Speech, and Signal Processing, 1984.

[2] John G. Proakis and Dimitris G. Manolakis, "Digital Signal Processing: Principles, Algorithms, and Applications"Pearson Education, 2013.

[3] Meribeth A. Dayme and Johann Sundberg, "Dynamics of the Singing Voice"1982