Waveform Wars – DSP Design Challenge Report

Team Name: Ep
Team Leader: Pothireddy Nishma Reddy
Team Members: Sharath

---

1. Introduction

Sound surrounds modern digital systems, from hands-free communication devices to music and multimedia platforms. Digital Signal Processing (DSP) provides the mathematical and algorithmic foundation required to analyze, model, and manipulate audio signals in real time. As described in the *Waveform Wars* challenge filecite turn0file0 , this project addresses two real-world DSP problems using *pure signal-processing techniques*, without relying on machine learning or pretrained models.

The two problem statements tackled in this report are: 1. **Acoustic Echo Cancellation (AEC)** in hands-free systems 2. **Audio Source Separation (BSS)** for overlapping sound sources in different frequency bands

Both solutions are implemented in MATLAB and emphasize adaptive filtering, time–frequency analysis, and robust signal reconstruction.

---

2. Problem Statement 1: Acoustic Echo Cancellation

2.1 Algorithm Description

In hands-free audio systems, a loudspeaker output (far-end signal) is partially captured by the microphone along with the near-end speech. This undesired coupling creates acoustic echo, degrading speech quality. The goal is to estimate the acoustic echo path and subtract the echo component from the microphone signal in real time.

Our implementation follows a **Normalized Least Mean Squares (NLMS)**-based adaptive filtering approach:

1. **Signal Acquisition**
   • Far-end reference signal: music played through the speaker
   • Microphone signal: near-end speech + echoed far-end signal

2. **Adaptive Filter Structure**
   • FIR adaptive filter with 1024 taps to model the room impulse response
   • Weight vector initialized to zeros

3. **Adaptive Update Rule**
   The filter coefficients are updated sample-by-sample using:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\|\mathbf{x}(n)\|^2 + \delta}\, e(n)\, \mathbf{x}(n)$$

where: • $\mathbf{x}(n)$ is the reference signal buffer
• $e(n) = d(n) - y(n)$ is the error signal
• μ is the adaptive step size
• δ ensures numerical stability

1. **Double-Talk Robustness**
   Instead of aggressive coefficient freezing, a conservative step-size and energy-based smoothing strategy is adopted to maintain stability during near-end speech activity.

2. **Post-Processing**
   A soft-gating mechanism attenuates low-energy residual noise, improving perceptual quality without distorting speech.

2.2 Results and Visualizations

• The adaptive filter successfully converges to the acoustic echo path.
• The processed output shows strong attenuation of the music echo while preserving near-end speech.
• Time-domain waveforms demonstrate a clear reduction in background echo energy.

Key output: - **removed_echo.wav** – Echo-suppressed microphone signal

---

2.3 Discussion

The NLMS-based approach provides fast convergence and numerical stability, even with long filter lengths. Increasing the number of taps improves echo path modeling for room acoustics, at the cost of computational load. While explicit double-talk detection was not implemented, the algorithm remains stable under simultaneous speech and music, satisfying the problem constraints.

---

3. Problem Statement 2: Audio Source Separation

3.1 Algorithm Description

This problem focuses on separating two overlapping audio sources (speech and tonal interference) using *pure signal-processing methods*. The key assumption is partial separation in the frequency domain.

The proposed solution uses **Short-Time Fourier Transform (STFT)-based spectral masking**:

1. **Time–Frequency Analysis**
   • Input signal transformed using STFT (Hamming window, overlap)
   • Magnitude and phase components extracted

2. **Noise Profile Estimation**
   • Initial frames assumed to contain dominant interference
   • Mean and standard deviation of spectral magnitudes computed per frequency bin

3. **Spectral Masking Strategy**
   • Tonal/constant interference identified using adaptive thresholds
   • Two masks constructed:

      ◦ Speech mask (broadband, non-tonal energy)
      ◦ Interference mask (horizontal spectral bands)

4. **Source Reconstruction**
   • Masked magnitudes combined with original phase
   • Inverse STFT applied to recover time-domain signals

5. **Artifact Reduction**
   • Median filtering applied to speech magnitude spectrum
   • Normalization prevents clipping

3.2 Results and Visualizations

• Clear separation of speech and tonal interference achieved
• Spectrograms show strong localization of interference in narrow frequency bands
• Speech output retains intelligibility with minimal musical noise

Key outputs: - **EXTRACTED_SPEECH.wav**
- **EXTRACTED_INTERFERENCE.wav**

3.3 Discussion

This approach demonstrates that classical DSP techniques can effectively solve source separation problems without machine learning. The method works best when sources occupy distinct spectral regions. Limitations arise when frequency overlap increases, but within the challenge constraints, the performance is robust and interpretable.

---

4. Conclusion

This project successfully addresses both challenge problem statements using principled DSP techniques. For Acoustic Echo Cancellation, an NLMS-based adaptive filter effectively models and suppresses room echoes in a realistic hands-free scenario. For Audio Source Separation, STFT-based spectral masking enables clean extraction of overlapping sources without prior training data.

Overall, the solutions highlight the power of adaptive filtering and time–frequency analysis in real-world audio systems. The project reinforces that well-designed classical DSP algorithms remain highly relevant, efficient, and explainable in modern signal-processing applications.

End of Report