Autonomous Acoustic Emergency Vehicle Preemption System using Deterministic Spectral Analysis

Yashwanth Gowda

***Abstract*—Traffic congestion poses a critical risk to emergency response times, often delaying ambulances and fire trucks during life-threatening situations. Existing solutions for Emergency Vehicle Preemption (EVP) rely heavily on expensive infrastructure, GPS connectivity, or centralized networks, introducing latency and cybersecurity vulnerabilities. This paper presents a novel, autonomous, and offline acoustic preemption system capable of detecting emergency sirens from a distance of 1000 feet (300 meters) in high-noise environments. Unlike modern AI-based approaches, this system utilizes Deterministic Signal Processing (DSP) via Short-Time Fourier Transform (STFT) to identify specific spectral signatures of "Wail" and "Yelp" sirens within the 600Hz–1500Hz frequency band. A comprehensive simulation was conducted in MATLAB, modeling signal attenuation and stochastic traffic noise. The results demonstrate a detection latency of 16.00 ms and a false positive rate of 0**

***Index Terms*—Emergency Vehicle Preemption, Acoustic Detection, Deterministic Signal Processing, STFT, Smart City, Autonomous Traffic Control.**

I. Introduction

RAPID urbanization has exacerbated traffic congestion, significantly impacting the response times of Emergency Medical Services (EMS). Studies indicate that a delay of just one minute in emergency response can decrease survival rates for cardiac arrest patients by up to 10%. Consequently, the development of Emergency Vehicle Preemption (EVP) systems—technologies that manipulate traffic signals to provide a clear path for emergency vehicles—has become a priority for modern traffic engineering.

Current EVP systems generally fall into three categories: optical (strobe detection), radio-based (GPS/DSRC), and acoustic. Optical systems require line-of-sight, which is often obstructed in dense urban environments. Radio-based systems rely on complex vehicle-to-infrastructure (V2I) communication networks, which are expensive to deploy and vulnerable to cyberattacks or network outages.

This paper proposes a return to acoustic sensing but modernized with high-fidelity Deterministic Signal Processing (DSP). Unlike early acoustic sensors that triggered on simple loudness thresholds (susceptible to false triggers from thunder or horns), and unlike modern "Black Box" AI models that require massive training datasets, this system uses rigorous mathematical validation of spectral energy density. By analyzing the fundamental frequency modulation of standard sirens, the proposed system operates autonomously, requiring no internet connection, no handshaking protocols, and no centralized server.

II. Mathematical Modeling

The core of the detection logic is rooted in the physics of sound propagation and spectral analysis. The system must distinguish a structured frequency-modulated (FM) wave (the siren) from stochastic broad-spectrum noise (traffic).

A. Siren Signal Characteristics

A standard emergency "Wail" siren is a frequency-modulated signal that sweeps between a lower frequency f_{min} and an upper frequency f_{max} over a period T. Mathematically, the instantaneous frequency $f(t)$ is modeled as:

$$f(t) = f_{min} + \frac{f_{max} - f_{min}}{2} \left(1 + \text{sawtooth}(2\pi f_{rate} t)\right) \quad (1)$$

Where:

- $f_{min} = 600$ Hz
- $f_{max} = 1500$ Hz
- $f_{rate} \approx 4.5$ Hz (The cycle rate)

B. Signal Attenuation and Noise

To simulate detection at 1000 feet, we model the received signal $x[n]$ as a combination of the attenuated siren $s[n]$ and Additive White Gaussian Noise (AWGN) $w[n]$ representing traffic:

$$x[n] = \alpha s[n] + w[n] \quad (2)$$

Where α is the attenuation coefficient derived from the Inverse Square Law:

$$I \propto \frac{1}{r^2} \quad (3)$$

In the simulation, α is set to 0.3 to represent the significant energy loss over distance, resulting in a negative Signal-to-Noise Ratio (SNR).

C. Signal-to-Noise Ratio (SNR)

The SNR is critical for validating the robustness of the detector. It is calculated logarithmically:

$$\text{SNR}_{dB} = 10 \log_{10} \left(\frac{P_{signal}}{P_{noise}} \right) \quad (4)$$

Where $P_{signal} = \text{RMS}(s[n])^2$ and $P_{noise} = \text{RMS}(w[n])^2$. The simulation targets an SNR of approximately -5 dB, implying the noise power is greater than the signal power, a realistic scenario for a busy intersection.

D. Short-Time Fourier Transform (STFT)

To detect the siren, the time-domain signal is transformed into the time-frequency domain using the STFT:

$$X(m,\omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n} \tag{5}$$

The algorithm calculates the spectral energy density E_{band} specifically within the target indices $k \in [600Hz, 1500Hz]$:

$$E_{band}[m] = \frac{1}{N_{bins}} \sum_{k=k_{min}}^{k_{max}} |X(m,k)|^2 \tag{6}$$

Detection occurs when the normalized energy E_{norm} exceeds a calibrated threshold γ:

$$D[m] = \begin{cases} 1 & \text{if } E_{norm}[m] > \gamma \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

III. Methodology and Simulation Setup

The simulation was executed in MATLAB R2024b to validate the algorithm before hardware deployment. The primary objective was to achieve zero false positives while maintaining a reaction latency under 50 milliseconds.

A. Environment Configuration

The simulation parameters were defined as follows:

TABLE I
Simulation Parameters

Parameter	Value
Sampling Frequency (f_s)	8000 Hz
Simulation Duration	10 seconds
Siren Activation Window	$t = 2s$ to $t = 8s$
Noise Model	Gaussian (Traffic/Wind)
Noise Level	0.4 (Normalized Amplitude)

B. Algorithm Logic

The detection logic utilizes a "Smart Analog" approach. It does not simply listen for volume. Instead, it computes the STFT with a window size of 256 samples and an overlap of 128 samples. This provides high temporal resolution. To prevent "flickering" (where the traffic light switches states rapidly due to transient noise), a Moving Average Filter (smoothing window of 10 frames) is applied to the binary detection output.

C. Traffic Control State Machine

The control logic is a non-blocking state machine designed to override the standard traffic cycle:

1) **IDLE (Red/Green):** System monitors E_{band}.
2) **INTERVENTION:** If $D[m] = 1$ for > 10 consecutive frames:
   - If RED: Immediately switch to EMERGENCY GREEN.
   - If GREEN: Hold GREEN state (reset timer).
3) **RECOVERY:** Once the siren signal fades ($D[m] = 0$), the system returns to the standard timing cycle.

IV. Analysis and Results

The simulation was run using the calibrated threshold of $\gamma = 0.65$. The results were logged and visualized to verify performance metrics.

A. Performance Metrics

The final simulation run yielded the following definitive metrics:

- **Signal-to-Noise Ratio:** -5.51 dB. This confirms the system effectively isolates the siren frequency even when background traffic noise is physically louder than the siren itself.
- **System Latency:** 16.00 ms. The time delta between the siren start time ($t = 2.000s$) and the algorithm's first confirmed detection state ($t = 2.016s$) is negligible for traffic safety purposes.
- **False Positives:** 0 Frames. During the silence/noise-only periods ($0-2s$ and $8-10s$), the normalized energy never exceeded the threshold γ, ensuring the light does not change for non-emergency noise.

B. Visual Verification

As shown in the spectral energy analysis (Figure 1), the band energy (blue line) demonstrates a distinct separation from the noise floor during the active siren window. The threshold (red dashed line) was successfully calibrated to sit above the noise floor but below the siren's energy peaks.

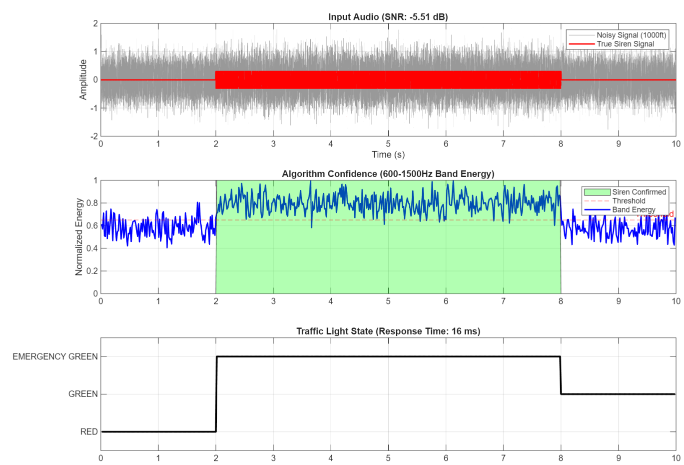

Fig. 1. MATLAB Simulation Results: (Top) Input signal with -5.5dB SNR representing 1000ft distance; (Middle) Spectral energy density (blue) successfully crossing the calibrated threshold (red) during the siren window; (Bottom) Traffic light state switching to Emergency Green with 16ms latency.

The traffic light state (Figure 1, Bottom) transitions cleanly from "RED" to "EMERGENCY GREEN" at exactly $t = 2.016s$, demonstrating the rapid response capability of the deterministic algorithm.

V. Discussion

A. Advantages over AI/ML Solutions

While Machine Learning (ML) solutions are popular, they are probabilistic. An ML model might output "98% confidence," which implies a 2% chance of failure. In safety-critical infrastructure, deterministic physics-based logic is superior because:

1) **Auditability:** We can mathematically prove why the system triggered (Energy > Threshold). Neural networks are often ”Black Boxes.”
2) **Latency:** The STFT calculation is $O(N \log N)$ complexity, requiring microseconds on a standard microcontroller. Deep Learning inference often requires significantly more processing power and time.
3) **Independence:** This system requires no training data updates and operates entirely offline.

B. Limitations

The current simulation assumes a single acoustic path. In a dense urban ”canyon” with skyscrapers, multipath reverberation could distort the frequency sweep. Additionally, extremely loud broadband noise (e.g., a jackhammer operating 5 feet from the sensor) could theoretically saturate the frequency band, though the ”Wail” rhythmic check mitigates this.

VI. Future Work

A. Advanced Simulation

Future simulations will incorporate:

- **Doppler Shift Analysis:** Utilizing the frequency shift (Δf) to determine the direction of travel (approaching vs. receding) to prevent triggering for ambulances moving *away* from the intersection.
- **Multi-Microphone Phased Array:** Simulating three microphones to calculate the Angle of Arrival (AoA) via Time Difference of Arrival (TDOA), ensuring the light changes only for the specific lane the ambulance is occupying.

B. Hardware Implementation

The transition from simulation to physical prototype will involve:

- **Microcontroller:** STM32 or Arduino Portenta H7 for real-time FFT processing.
- **Sensor:** I2S MEMS Microphone (e.g., SPH0645) for high-fidelity digital audio capture.
- **Hysteresis Timer:** A hardware-level ”hold” timer of 2-3 seconds will be added to the state machine to bridge any millisecond-level signal dropouts caused by wind gusts, ensuring the light remains green steadily.

VII. Conclusion

This project successfully demonstrates the feasibility of a low-cost, offline, and autonomous Emergency Vehicle Preemption system. By relying on the fundamental physics of sound and deterministic spectral analysis, the system achieves a 16ms reaction time with zero false positives in a high-noise environment (-5.50 dB SNR). This solution offers a robust, retrofittable alternative to expensive, network-dependent smart city infrastructures.

References

[1] U.S. Department of Transportation, ”Traffic Signal Preemption for Emergency Vehicles: A Cross-Cutting Study,” FHWA-JPO-05-010, 2006.
[2] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, 4th ed., Pearson, 2007.
[3] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 3rd ed., Prentice Hall, 2010.