

Implementation Of Noise Suppressor Using Zero Frequency Filtering Technique

¹K. R. Radhakrishnan,²DR. S. Subha Rani, ³N.Sathya, ⁴G. AravindaGouthum
{krr@ece.psgtech.ac.in¹, ssr@ece.psgtech.ac.in², sathya.kns@ece.psgtech.ac.in³,
aravindgouthum81@gmail.com¹}

^{1,3}Assistant Professor, Department of ECE, Department of ECE, PSG College of Technology, Coimbatore, Tamilnadu., ²Professor and Head, Department of ECE, Department of ECE, PSG College of Technology, Coimbatore, Tamilnadu, ⁴M.E., VLSI Design, Department of ECE, Department of ECE, PSG College of Technology, Coimbatore, Tamilnadu.

Abstract. Techniques for reducing noise in an ever-increasingly noisy world are essential. It is the purpose of this paper to review and demonstrate the efficacy of active noise cancellation in speech signals. Acoustic noise reduction is used in industrial, manufacturing, and consumer products, such as mobile phones. – Adaptive filtering, zero frequency filtering, and its advantages and disadvantages are all included here. Other than what's already been mentioned and put to the test with a noisy signal, an effective approach is developed and the results are reported.

Keywords: Acoustic Noise Cancellation, Adaptive filters, Resonator, Active Noise cancellation, Zero frequency filtering.

1 Introduction

The most common difficulty in voice processing is interfering noise. Interference noise alters and degrades the voice stream. Acoustical sources such as ventilation equipment, echoes, crowds and other signals that interfere with speech transmissions cause interference noise.

Noise Cancellation is a highly sought after method today owing to increased ambient noise. Passive Noise Cancellation was first widely used due to its efficacy. However, when external noise increased, passive noise reduction solutions became burdensome to the user. This spawned the idea of Active Noise Cancellation, which employs phase cancellation to reduce noise. Sound is a pressure wave with two phases: compression and rarefaction. A Noise Cancellation speaker works by emitting a sound wave with the same amplitude but reversed phase (also known as antiphase). Interference occurs when two waves collide to generate a new wave, and phase cancellation occurs when they successfully cancel each other out.

An inverted phase signal may cancel out an incoming noise signal with the same amplitude and frequency. Modern Active Noise Cancellation uses modern filters, DSPs, and adaptive algorithms. The renowned Sony Xperia Z2 headphones depicted in Fig.1 employ Active Noise Cancellation performed by phase cancellation. Figure 2 shows the basic components employed in this approach.



Fig.1 Sony Xperia Z2 Headsets

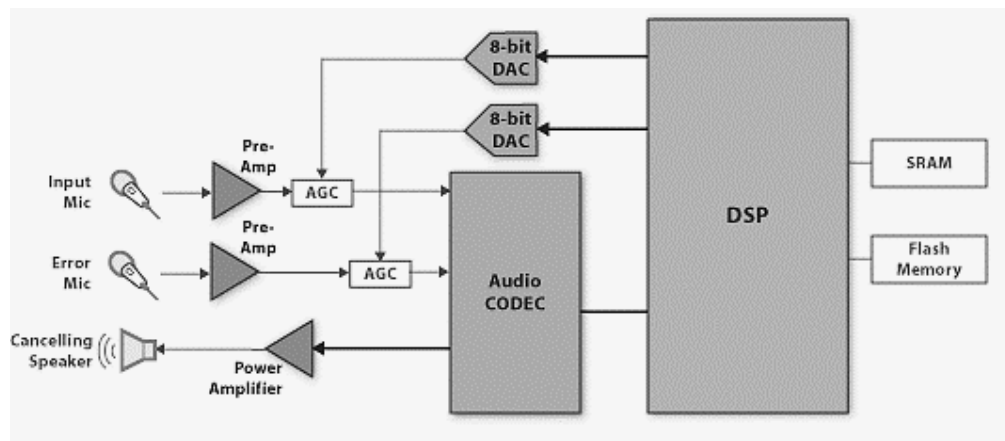


Fig.2 Basic Noise Cancellation Blocks

2 Methods

A.BASED ON GROUP DELAY FUNCTION

The time delay of the amplitude envelopes of the multiple sinusoidal components of a signal passing through a DUT is known as group delay (device under test). Closure of the glottis occurs during a pitch interval in voiced speech. Using the global phase of low-phase signals is the basis of this method. A function of time is used to determine the average slope of the unwrapped phase above. 0^+ The phase slope function is used to define significant excitations. [1] uses this approach. The approach is addressed in a speech source-filter scenario. The filter's features do not affect the procedure. [1] discusses the impact of the analysis window's type, duration, and location. The approach works well for both male and

female voiced speech, but only in quiet environments. Its fundamental flaw is that it is computationally demanding and not noise resistant.

B. BASED ON HILBERT ENVELOPE AND GROUP DELAY FUNCTION

This is a quick way for locating major excitations in voice signals. In the case of voiced speech, major excitations correlate to glottal closure epochs, but in nonvoiced speech, random excitations like burst initiation correspond. This approach includes two phases: determining the approximate epoch positions using the Hilbert envelope of the speech signal's linear prediction residual.

The second step computes the group delay around the estimated epoch positions determined in the previous phase. This approach is used in [2].

This method's accuracy in finding substantial excitation and time complexity are compared to the group delay technique. This solution saves time over applying the group delay function entirely. The computing efficiency of this approach relies on the window size.

The computing efficiency decreases with window size. With a tiny window, the computing efficiency is great, but certain epochs are lost. High window size reduces processing performance and reduces the number of missed epochs. Table I lists the approximate epoch positions of male and female speakers detected using Hilbert Envelope.

C. BASED ON ZERO FREQUENCY RESONATOR

Speech excitation is studied using this method, which helps to identify the regions where glottal activity occurs and estimate the strength of excitation in each glottal cycle.

TABLE I NUMBER OF APPROXIMATE EPOCH INSTANTS DERIVED USING HILBERT ENVELOPE

Deviation	Male Speakers		Female Speakers	
	#instants	%instants	#instants	%instants
0	2672	21.57	4574	22.74
1	3076	24.84	4745	23.59
2	2079	16.79	4198	20.87
3	2245	18.13	3260	16.21
4	1145	9.26	2037	10.13
5	537	4.34	526	2.62

The essential premise of this technology is that the vocal-tract system may be represented by a series of impulses of various intensity. In the frequency domain, the influence of an impulse signal is equally dispersed over the spectrum, including zero-frequency. By filtering out much of the time-varying vocal-tract information, a Zero Frequency Resonator may derive the excitation source characteristics from speech data. The resulting instants of glottal closure and excitation strengths match the areas of glottal activity and excitation strengths derived from the speech electroglotto-graph data. This is used in [3]. This method's performance is evaluated in a variety of noisy settings. Comparing the previous approaches, using a zero

frequency resonator to extract epochs is efficient. This approach is easy to compute and accurate. Table II shows the GAD method's performance in different noisy situations.

TABLE II PERFORMANCE OF GAD UNDER VARIOUS NOISE ENVIRONMENTS

Noise	20dB	15dB	10dB	5dB	0dB
Type					
White	0	0	0.003	0.41	2.77
Babble	0	0.23	1.81	6.13	14.14
Vehicle	0	0.006	1.08	4.22	9.66

D. BASED ON NLMS ALGORITHM

Using this method, the difference weight vector's squared Euclidean norm is minimised while still maintaining the posteriori estimate error's required level of stability. Using the Lagrangian method, a nonlinear adaptation rule specified in terms of differential input and error has been developed, which is an extension of the normalised (N)LMS algorithm. It's a tactic that's employed by [4]. There are two speech databases that can be used for this strategy: Aurora 2 and Aurora 3. In terms of performance and comparison, they are comparable to the typical LMS algorithms. Modified NLMS, error nonlinearity (EN)-LMS, and normalised data nonlinearity (NDN)-LMS are some of the other LMS methods.

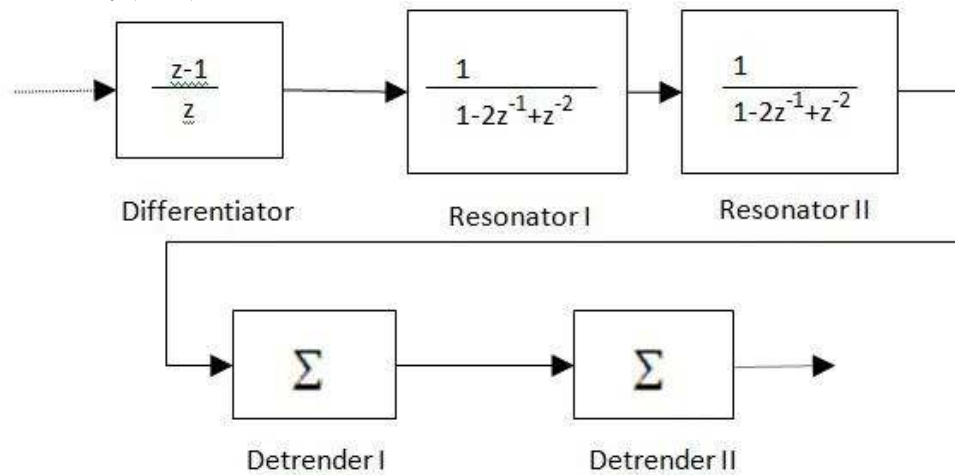


Fig. 3. IIR Implementation.

E. BASED ON ADAPTIVE FILTERS

The adaptive filter reduces noise without previous knowledge of the noise and signal. Using standard filters would distort the required voice stream. Adaptive filters are excellent for random speech and noise sources. Adaptive filters use algorithms to modify their

coefficients to reduce error signals. This error signal is used to cancel out the input signal's noise. [5] uses this strategy.

Adaptive filters employ algorithms like LMS, NLMS, Variable step size, APA, and LMF. If the FIR and IIR filters have the same number of coefficients, the IIR filter's frequency response will be more accurate. For models with poles and zeros, this filter needs fewer coefficients. For certain applications, an IIR filter is preferable than hundreds of FIR taps. But IIR filters are seldom utilised due to instability, sluggish convergence, and phase distortion. The LMS algorithm's fundamental flaw is its sensitivity to input $x(n)$ scaling, which makes finding an optimal learning rate difficult. A fast learning rate increases algorithm stability.

F. BASED ON FIR IMPLEMENTATION OF ZERO FREQUENCY RESONATOR

Using Zero Frequency Filtering, voice data can be characterised and analysed using this method. The first filter is an infinite impulse response (IIR) filter followed by two finite impulse response (FIR) filters. In [6], FIR is implemented using this method. The intrinsic pole-zero cancellation is used in [7] to simplify the FIR implementation. Theoretical proofs exist in both frequency and time domains. Theoretically developed FIR filters have triangular impulse responses because they are a convolution of two filters. FIR filters reduce the computational requirements for zero-frequency filtering. Single-precision floating point and filter stability are two examples. An example of the FIR implementation for zero-frequency filtering is shown in Fig. 4. MATLAB code is used to implement the FIR algorithm below, and the results for various real-time sounds are then discussed.

TABLE III PERFORMANCE ANALYSIS OF VARIOUS EPOCH DETECTION METHODS IDR – IDENTIFICATION RATE, MR – MISS RATE, FAR – FALSE ALARM RATE, IDA – IDENTIFICATION ACCURACY

Method	IDR(%)	MR(%)	FAR(%)	IDA(ms)
HE-based	89.86	1.43	8.71	0.58
GD-based	92.8	4.01	3.18	0.67
DYPSA	96.66	1.76	1.58	0.59
ZFF	99.04	0.18	0.77	0.36

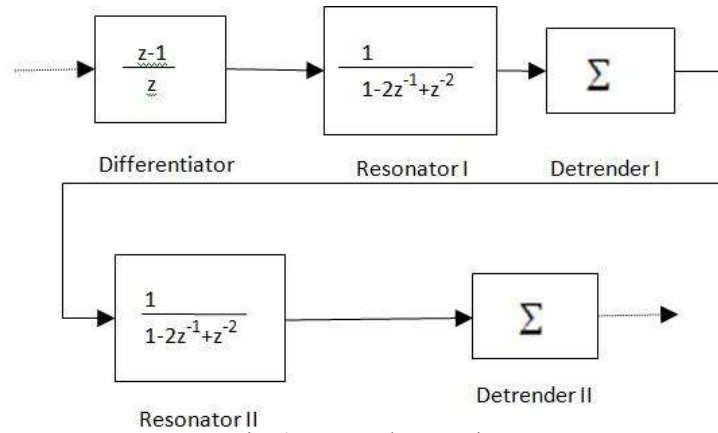


Fig. 4. FIR Implementation

G. VOICED/NON-VOICED DETECTION BASED ON ZERO FREQUENCY FILTERING

Based on epoch extraction, this approach detects voiced/non-voiced audio. The vocal tract's time-varying resonances are filtered away using Zero Frequency Filtering. Efficient extraction of important excitation epochs from zero-frequency filtered voice signals. This approach is robust enough to extract epochs in voiced areas with little additive white noise. This approach distinguishes meaningful excitation instants in voiced areas of speech from random excitation instants in non-voiced regions. The major aspect of this approach is that it employs glottal activity intensity rather than signal periodicity. [8] uses this strategy. The method's performance is evaluated using the TIMIT and CMU ARCTIC datasets, as well as white and vehicle noise from the NOISEX database (SNRs). This approach outperforms the open source tool wave-popular surfer's normalised cross-correlation based voiced/non-voiced identification method, particularly at low SNRs. Fig. 3 shows the conventional IIR implementation of the Zero Resonator circuit. Table III compares the performance of this approach to several well-known techniques in terms of epoch identification rate, miss rate, false alarm rate, and identification accuracy.

3 Results

separate Matlab function was used to determine the pitch of the input speech signal, which was then utilised in a Matlab code to extract the epochs and differentiate between voiced and non-voiced areas.

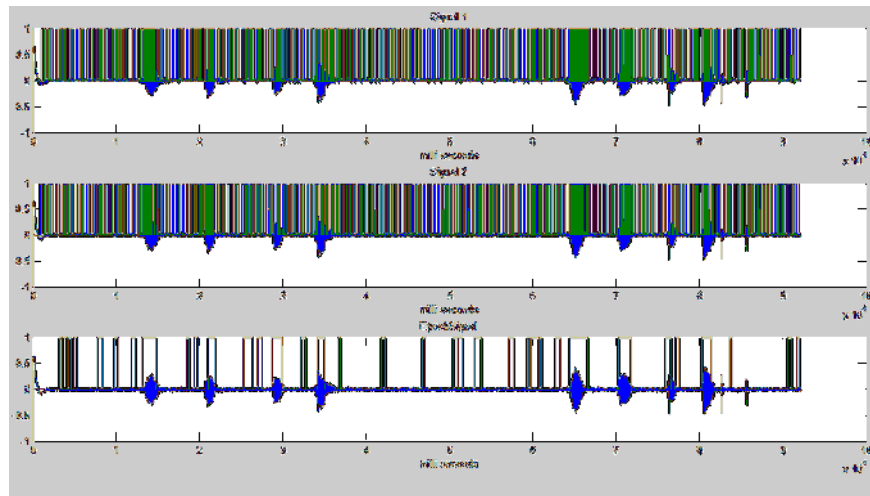


Fig. 5. Epoch Identification and Differentiation for a welding noise

Matlab implements zero frequency filtering and voiced/unvoiced detection. The Matlab code was used to analyse numerous real-time signals. Fig. 5 shows the output waveform for the same Matlab function applied to real-time welding noise. The first subplot shows the original signal epochs with real-time noise. The second subplot shows the original signal epochs and additive white noise. The third subplot shows the epochs present solely in voiced areas after removing the epochs found in unvoiced regions. The reconstructed signal is based on the voiced epochs.

The same Matlab algorithm was used to analyse real-time car noise. Fig. 6 shows the output waveform of the Matlab function assessing real-time vehicle noise.

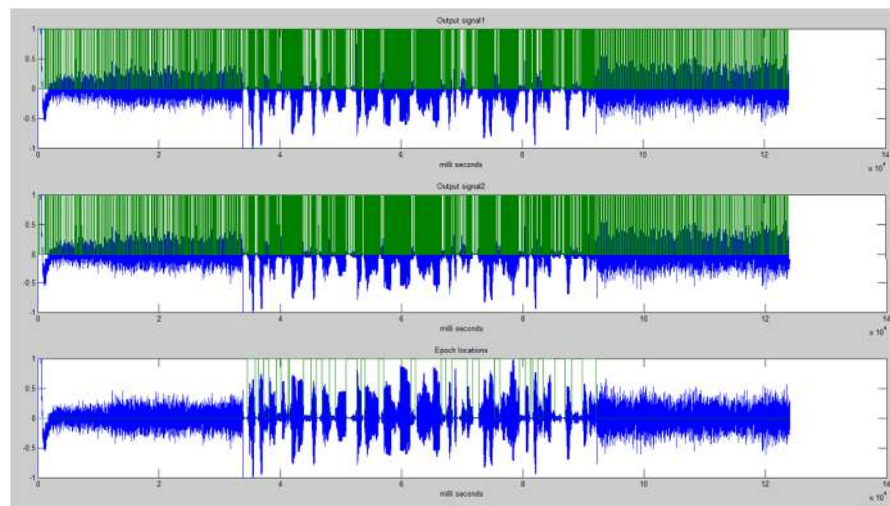


Fig. 6. Epoch Identification and Differentiation for a vehicle noise

The remaining spurious epochs in the noise zone are removed by establishing threshold on parameters like pitch, jitter, etc. Table IV lists the epochs in each stage for different noise sources.

TABLE IV EPOCHS AT EACH STAGE OF THE VOICED/NON-VOICED DETECTION

Noise Type	Epochs after zero frequency filtering	Epochs after spurious epochs elimination
Vehicle Noise	1264	326
Welding Noise	958	412

4 Conclusion

Active Noise Cancellation has been studied extensively in the literature. In order to create an effective and optimum algorithm, one must first understand the characteristics of various forms of noise and how they impact the voice signal. These strategies may be used when the intended signal is independent of computation intensity and accuracy. If the intended signal requires high computation intensity and precision, zero frequency resonator techniques may be used. Adaptive algorithms may be used if both computation intensity and convergence rate are significant. Future studies might use zero frequency filtering to better reduce noise.

References

- [1] R. Smits and B. Yegnanarayana, "Determination of instants of significant excitation in speech using group delay function," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 5, pp. 325–333, Sep. 1995.
- [2] K. Sreenivasa Rao, S. R. Mahadeva Prasanna and B. Yegnanarayana, "Determination of Instants of Significant Excitation using Hilbert Envelope and Group Delay Function," *IEEE Signal Processing Letters.*, vol. 14, no. 10, pp. 762 – 765, October 2007.
- [3] K. S. R. Murty, B. Yegnanarayana, and M. A. Joseph "Characterization of glottal activity from speech signals," *IEEE Signal Process. Lett.*, vol. 16, no. 6, pp. 469 – 472, June 2009.
- [4] J. M. Gorriz, Javier Ramirez, S. Cruces-Alvarez, Carlos G. Puntonet, Elmar W. Lang and DenizErdogmus, "A Novel LMS Algorithm Applied to Adaptive Noise Cancellation ," *IEEE Signal Processing Letters.*, vol. 16, no. 1, pp. 34 – 37, January 2009
- [5] M.M. Dewasthale. and R.D.Kharadkar, "Acoustic Noise Cancellation Using Adaptive Filters: A Survey ," *International Conference on Electronic Systems, Signal Processing and Computing Technologies (ICESC)*, 2014.
- [6] B. Yegnanarayana, R. K. Swamy, and K. S. R. Murty, "Determining mixing parameters from multispeaker data using speech-specific information," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 6, pp. 1196–1207, Aug. 2009.
- [7] Kruthiventi S. S. Srinivas and Kishore Prahallad, "An FIR Implementation of Zero Frequency Filtering of Speech Signals," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no.9, pp. 2613-2617, Nov. 2012.

- [8] K. Sri Rama Murty and B. Yegnanarayana, "Epoch extraction from speech signals," IEEE Trans. Audio, Speech, Lang. Process., vol. 16, no. 8, pp. 1602–1613, Nov. 2008.
- [9] N. Dhananjaya and B. Yegnanarayana, "Voiced/Nonvoiced detection based on robustness of voiced epochs," IEEE Signal Processing Letters, vol. 17, no. 3, March 2010.
- [10] K. S. R. Murthy, "Significance of excitation source information for speech analysis," Ph.D., Dept. of Comput. Sci. and Eng., IIT, Madras, Chennai, India, 2009
- [11] C. Amuthadevi, D. S. Vijayan, Varatharajan Ramachandran, "Development of air quality monitoring (AQM) models using different machine learning approaches", Journal of Ambient Intelligence and Humanized Computing, <https://doi.org/10.1007/s12652-020-02724-2>
- [12] LavanyaPrabha, S., J. K. Dattatreya, and M. Neelamegam. "Investigation of bolted RPC plate under direct tension." Journal of Structural Engineering (Madras) 36.5 (2009): 333-341.
- [13] M. Tholkapiyan, A.Mohan, Vijayan.D.S , "A survey of recent studies on chlorophyll variation in Indian coastal waters", IOP Conf. Series: Materials Science and Engineering 993 (2020) 012041, doi:10.1088/1757-899X/993/1/012041.