Echo cancellation using Artificial Intelligence technique

Ziyang Lin[1,5,6,†], Jincheng Xu[2,†], Yuxuan Yang[3,†], Yichen Zeng[4,†]

[1]Tianfu No.7 High School, Chengdu, China
[2]American School of Warsaw, Warsaw, Poland
[3]Georgia Institute of Technology, Atlanta, United States
[4]Qingdao No.9 High School, Qingdao, China

[5]1910821233@mail.sit.edu.cn
[6]Corresponding author
[†]These authors contributed equally.

Abstract. To improve the audio quality in telecommunication, echo cancellation was proposed to prevent signals from re-appearing. Generally speaking, the echo canceler first recognizes the original transmitted signal that appears again due to delay in the transmitted or recieved signal. Then, the echo canceler subtracts the echo in the signals. Commonly, echo cancelers are calledacoustic echo cancellation (AEC) in general, whether they aim for acoustic echo, line echo, or both. However, in modern times, new problems have been emerging which require more innovative, more advanced methods to tackle the echo issue. Thus, the paper reviews methods regarding Artificial Intelligence (AI) to improve the current means for echo cancellation. In this paper, we mainly discuss the approaches to echo cancellation, the identification of echo using adaptive AI and the usage of adaptive filters and neural network for echo cancellation. For instance, deep learning approaches can facilitate the issues of AEC (active echo control) and ANC (active noise control),the optimization of LMS algorithm can be achieved by combining Swarm Intelligence (SI), combining adaptive digital filter and recurrent neural network for acoustic echo cancellation etc. Using AI in echo cancellation can help us achieve high-quality, full-duplex telecommunication in modern telephony system.

Keywords: echo cancellation, AI, SI, algorithm, adaptive filter, neural network.

1. Introduction

Acoustic echo has been a persistent problem for telecommunication systems ever since the early days of the telephone. In the late 19th and early 20th centuries, people tried to deal with echo by using hardware solutions like cloth cushions and padding and speaking techniques like adjusting the distance from the mouth to the microphone and holding the handset differently during talking and listening. As technology has improved, telecommunication devices have become smaller than ever and are expected to work well in the most challenging situations. Engineers have developed new echo cancellation techniques for the changes in the devices' hardware, the most prevalent being adaptive filtering techniques. However, these techniques have their downsides. While filtering techniques are sufficient at cancelling linear echo, microspeakers in modern hands-free devices can produce significant nonlinear distortion during playback, and the acoustic echo cancellation stage must be able to operate when nonlinear echo is present. Furthermore, the digital signal processing algorithm should always strive for

less computational complexity to ensure power efficiency on small portable devices. In recent years, teams worldwide have been implementing novel machine learning-based echo cancellation algorithms to improve the current echo cancellation approaches.

This paper aims to find and discuss possible solutions to noise and echo cancellation using artificial intelligence. The reason for using artificial intelligence is that conventional methods were never able to completely solve the problem of noise and echo cancellation problem, as they would always leave some sounds unattended. This is because conventional methods, such as tightly fitted earphones, use rigid solutions. This is a problem because, although this might help with the problem, it won't solve it. This is where artificial intelligence and other smart methods described in the paper come into play, where they are more capable due to their ability to adapt to different noisy situations and identify and differentiate between wanted and unwanted sounds. This allows artificial intelligence and other smart methods to be more dynamic than conventional methods and filter out more noise more effectively.

The paper focuses on three of the more important factors in niose cancelling using smart methods or AI. First, different approaches to noise cancellation Second, identification of noise and echos using adaptive filtering and AI Finally, in the third section, echo and noise cancellation using adaptive filters and neural networks In addition, the third section also includes a little bit on the audio processing that goes on behind the scenes of the adaptive filters mentioned in the section.

The research in this paper is of great value to the research and application of AI in echo and noise cancellation because it provides new methods such as adaptive filtering, algorithms, artificial intelligence, swarm intelligence, neural networks, etc. to achieve more effective echo and noise cancellation.

2. Noise Suppression Based on AI

Ritika Thakur and a couple other researchers said that one of the ways artificial intelligence can be applied to the context of noise cancelation is through adaptive noise cancelation. There are many ways in which this can be achieved. According to a study in 2014 that tried to use optimizing algorithms such as Artificial Bee Colony, Cuckoo Search, or Particle Swarm Optimization to improve adaptive filtering and/or noise cancellation. In the study, All of the proposed methods were simulated in Matlab, and it was found that the ABC algorithm and Cuckoo Search were superior to the "conventional and state-of-art speech signal denoising techniques" [1]. ABC and CS consistently yielded lower MSEs and higher SNRs. Finally, according to the images provided, it is clear that ABC and CS optimization methods were superior to that of PSO.

Zhang Hao also proposed deep learning approaches to the issues of AEC(active echo control) and ANC(active noise control) since conventional methods have proven unable to address these issues as the quality of the electronic devices, such as amplifiers and speakers, limits them. The study first goes a bit into the background and the motivation of the study. The author talks about how humans perceive and hear two types of sounds, wanted and unwanted, where one is to "communicate meaning, delight, and soothe a person" [2], and the other "can be disruptive and affect a person's mental health" [2]. The study then went into different experiments conducted. The results from these experiments were that deep AEC and ANC methods proved promising and advantageous to traditional methods. However, the study states two problems with the proposed methods: generalizing and processing latency.

Vanus and a couple other researchers focus on implementing multiple techniques in areas of nonlinear noise suppression [3] called Adaptive Neuro-Fuzzy Inference System (ANFIS). This means phone calls where the elimination of background noise is required. The experiments conducted as a part of the study proved that ANFIS had the best overall performance in noise cancelation as it "efficiently canceled noise even in highly noise-degraded speech" [2]. It is concluded in the study that ANFIS is best used "when there is a characteristic of an unknown external interference source, and the background noise is similar tothe measured noise" [2].

3. Echo cancellation Using Optimized Algorithms and adaptive filters

Making the best use of a resource or circumstance is the definition of optimization. The major focus of the work is on the optimization of the Least Mean Square (LMS) algorithm using a combination of Swarm Intelligence (SI) techniques like Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO).

The current LMS algorithm is trapped at the system's local minima with multi-model error surface while taking into account its global minima, leading to non-optimized convergence. LMS algorithm will be improved by combining it with SI like PSO and ACO, which will aid in resolving major interference and noise problems.

The mechanism of SI such as Ant Colony Optimization and Particle Swarm Optimization is finding the global minima value of certain programs by recursive algorithms which resolves the problem that the current LMS algorithm is stuck at the local minima. [4]

The conventional normalized least mean squares (NLMS) algorithm is lack of efficiency. Therefore, a faster, more efficient adaptive filtering algorithm for network echo cancellers Proportionate normalized least mean squares + + (PNLMS++) was proposed. It converges much more quickly than NLMs when the echo path is sparse.

PNLMS++ uses NLMS and PNLMS type coefficient updates and provides improved convergence and tracking over PNLMS when the echo path impulse response disperses. In addition, PNLMS+ + is only 50% more computationally complex than NLMS and requires no additional memory. PNLMS and PNLMS++ have about the same complexity.[5]

An adaptive linear filter that may be used to approximate the real acoustic echo route and subtract the estimated echo from the microphone signal is required for a variety of applications and devices where echoes and noises are progressively becoming more complicated and mixed. But for various reasons, there is always a residual echo after the aforementioned linear adaptive subtraction.

Thus, the paper proposed an effective signal processing system that consists of a noise reduction (SBNR) layer, a joint perceptual subband residual echo suppression (SBRES) layer, and an adaptation-based nonlinear echo cancellation (NLEC) layer that is easily extendable to multiple-channel cases. Linear, nonlinear, and time-variant echoes can all be reduced using the SBRES and NLEC layers. The SBNR layer can reduce echoes and noises with similar statistical properties to noises in communication.

The proposed signal processing system can significantly improve noise reduction, echo suppression, full-duplex voice communication (FDVC) and automatic speech recognition (ASR) in emerging artificial intelligence speakers. [6]

4. Echo Cancellation Using Neural Networks

The reduction of acoustic coupling between a loudspeaker and a microphone has gained considerable attention over the past decades. In the present day, the growing demand for full-duplex human-machine interfaces in our daily lives serves as a strong motivating factor for improvements in echo cancellation performance. Simultaneously, this interface is challenged by nonlinear distortion in smaller devices such as smartphones and wearables. This paper introduced a neural network-driven acoustic echo canceller that draws inspiration from traditional NLAEC methods. This network is designed to capture the characteristics of the nonlinear echo path and undergoes initial training using the error backpropagation algorithm. Part of the network corresponds to the time-invariant components of the echo path and is used to transfer learning. While in operation, the network selectively adjusts a subset of its parameters by employing the significance-aware elitist resampling particle filter. This adaptation is intended to accommodate any differences between the conditions encountered during training and actual deployment. The research team evaluates the proposed approach using synthesized and real nonlinear distortions recorded with a commercial mobile phone. The potential applications of this approach in real-world scenarios include improving the quality of audio and speech signals in teleconferencing, voice-controlled systems, and hands-free communication systems. The paper also discusses the limitations of the proposed approach, such as the tendency towards overfitting when increasing the number of neurons and the need for longer training sequences to model the distortion characteristics at

different volume levels. Overall, the proposed approach shows promise in improving the quality of audio and speech signals in real-world scenarios [7]. Acoustic echo is a common problem among audio electronics ranging from communication to entertainment. Traditionally, many adaptive filter techniques have been used to eliminate echo in the acoustic path with varying levels of success. This paper presents a new method to reduce the amount of residual echo in the signal path by combining existing filtering methods and a recurrent neural network. The proposed scheme for acoustic echo cancellation combines adaptive filter and neural network, which can cancel echo in a large scale and suppress the residual echo to a considerable level. The experiments demonstrate that the proposed scheme, when applied, achieves superior echo suppression performance despite some spectrum damage. The response time is also increased when compared to other prevaling echo cancellation methods, but still within a reasonable range. The performance evaluation includes model training with 10 hours of speech and 5 hours of echo data, and the validation loss approaches zero, indicating a considerable model has been trained. The proposed RNN algorithm can greatly suppress residual echo compared with Speex and WebRTC, especially at speech gaps where only residual echo exists [8]. To achieve high quality full-duplex telecommunication, reliable, effective, and efficient echo cancellation is, and has been, an integral part of a modern telecommunication system. An Acoustic Echo Cancellation (AEC) system can face a few major challenges in its operation. For example, microspeakers in modern hands-free devices can produce significant nonlinear distortion during playback, and the AEC must be able to operate when nonlinear echo is present. This approach models the echo path in two parts. The first part uses a novel neural network that can learn the nonlinear distortions of the device's analog components (electronics and loudspeaker), and the second part uses a standard adaptive linear filter to subtract the echo signal out of the near-end microphone path. The neural network incorporates trainable memory length and nonlinear activation functions to accommodate variations across devices. These elements are not pre-parameterized but rather optimized during the training phase. The neural network is used as a reference and is not updated, while the linear filter is adapted to the time-varying acoustic paths. Test result shows that this approach demonstrates better stability, robustness and convergence time when compared to other methods [9].

5. Research on Audio Signal Based on Laser Sensor

Audio signal processing, or audio processing, is mainly used to adjust the audio signal's amplitude, frequency, waveform. Speech signal is the main communication component, and noise interference is one of the most influential factors in speech signal transmission. Effectively dealing with noise interference to audio signal is a very important research aspect in communication.

In system framework modeling, the real audio signal processing mostly converts the audio signal into a binary code sequence, then sends the code sequence into the digital channel transmission. In the digital channel, noise will interfere with the signal, so the corresponding filter should be designed at the receiving end of the signal for filtering processing. The noise encountered in the channel is generally thermal noise, and the most typical one is Gaussian white noise. Its one-dimensional probability density satisfies the Gaussian distribution, and its second moments are independent of each other.

This paper selects two commonly used physically realizable filters: Butterworth filter (a) and Chebyshev filter (b). Chebyshev filters decay faster than Butterworth filters in the transition band, but the amplitude-frequency characteristics of the frequency response are not as flat as the latter. The error between the frequency response curves of Chebyshev filter and ideal filter is minimal, but amplitude fluctuates in the passband. Chebyshev filter can be divided into passband Chebyshev filter and stop band Chebyshev filter according to the different positions of fluctuation of frequency response curve. [10]

The audio signal processing system of the laser sensor includes two parts: hardware subsystem and software subsystem, and its functional modules are divided into: transmitting signal module, signal receiving module, signal processing module, audio amplification module, audio signal recognition module and audio signal output module. The working principle of the audio signal processing system of the laser sensor is: first, the laser sensor is used to collect the signal, and the signal is sent to the audio signal processing module, and then the signal is amplified and classified in the audio signal processing

module, and finally the audio signal is output through the output module. Audio signal processing system directly affects audio signal quality, so it has always been the focus of attention. To solve some limitations of the current audio signal processing system and obtain better audio signal processing results, an audio signal processing system based on laser sensor is designed. First of all, the overall structure of the audio signal processing system is designed, and then the hardware and software of the system are designed respectively. Because the optical sensor is used in this paper to collect audio signals, the acquisition speed of audio signals is accelerated, and the audio signal processing results obtained are better than the comparison system, and have a wider range of applications.[11]

The system is designed with STM32F103C8T6 as the main control module chip, using semiconductor laser to emit laser signal, semiconductor laser receiving tube to receive laser signal, using TDA7297 audio power amplifier integrated circuit, receiving the command of the main control module to drive the speaker. Output notes DO, RE, MI, FA, SO, LA, SI sounds. The main control core STM32F103C8T6 is periodically cycled to collect data from the receiving module of the semiconductor laser. The receiving module of the semiconductor laser is designed as seven channels.

The operation of semiconductor laser is affected by laser radiation, resonator and gain. Among them, when the laser radiation, the radiation light propagates in the PN junction plane, the monochromaticity is better, the intensity is larger. The resonator can improve the lasing efficiency and form oscillations in the resonator. Under the action of the injection current, the excited radiation in the activation region is enhanced continuously. Semiconductor lasers are stimulated by radiation and emit photons in the same direction and phase.

To realize the function of laser receiving end detection and sound frequency output, the program design includes main program, delay subroutine, laser receiving query subroutine, timer, interrupt subroutine and output program. The main program can realize the initialization of register address, variables, and the call of partial programs. Interrupt subroutines provide PWM frequencies and output different notes. When the laser receiver receives the laser beam, there is no output, and the program continues to monitor the laser receiver. [10]

6. Conclusion

This paper identifies the importance of noise suppression and echo cancellation performance in modern audio systems and presents many novel and effective approaches to ameliorate these problems with artificial intelligence. The paper first covers various topics related to modern audio signal processing techniques to establish background information. Some traditional approaches to achieve echo cancellation like the Least Mean Square methods, are summarized, and their deficiencies are explained. The paper then introduces how several noise cancellation and echo suppression methods using artificial intelligence, like the Adaptive Neuro-Fuzzy Inference System and Artificial Bee Colony, can work with traditional methods to improve system-level performance. The flexibility and adaptivity of the artificial intelligence layers can dramatically improve nonlinear echo cancellation performance over the traditional approaches, and it also allows the AI processing methods to work on a variety of mobile phones to smart home appliances. As the computing power in smart devices becomes cheaper and more plentiful, more algorithms based on artificial intelligence will certainly be implemented to enhance the echo cancellation performance in such devices. It is reasonable to expect continuous growth in echo cancellation performance in audio electronics soon. As a result, more clarity in full-duplex voice communication (FDVC) systems and higher accuracy in automatic speech recognition (ASR) systems will also be achieved.

References
[1] Thakur, Ritika, Papiya Dutta, and Dr GC Manna. "Analysis and comparison of evolutionary algorithms applied to adaptive noise cancellation for speech signal." International Journal of Recent Development in Engineering and Technology 3.1: pp. 172-178, 2014.

[2]   Vanus, Radek Martinek—Michal Kelnar—Jan, and Petr Bilik—Jan Zidek. "A robust approach for acoustic noise suppression in speech using ANFIS." Journal of electrical engineering 66.6: pp. 301-310, 2015.

[3]   Zhang, Hao. "Deep Learning for Acoustic Echo Cancellation and Active Noise Control." The Ohio State University, 2022.

[4]   Q. Ling, M.Ikbal, and P.Kumar, "Optimized LMS algorithm for system identification and noise cancellation." J.Intell.Syst., vol. 30(1), pp. 487-498, 2021.

[5]   S. L. Grant, "An Efficient, Fast Converging Adaptive Filter for Network Echo Cancellation," Conference Record of the 32nd Asilomar Conference on Signals, Systems & Computers, Vol. 1, pp. 394-398, November 1998.

[6]   J. Yang, "Multilayer adaptation based complex echo cancellation and voice enhancement" 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2131-2135, April 2018.

[7]   Halimeh, M. Modar, C. Huemmer, and W. Kellermann, "A neural network-based nonlinear acoustic echo canceller." IEEE Signal Process. Lett., vol. 26.12:, pp. 1827-1831, November 2019.

[8]   L. Ma, et al, "Acoustic echo cancellation by combining adaptive digital filter and recurrent neural network." arXiv preprint arXiv:2005.09237, May 2020.

[9]   A. Ivry, I. Cohen, and B. Berdugo. "Nonlinear acoustic echo cancellation with deep learning." arXiv preprint arXiv:2106.13754, Jun 2021.

[10]   Z. Yu "Research on Audio Signal Processing Based on MATLAB," ISSN.1672-7274.2023.02.015,   pp. 43-45, February 2023.

[11]   J. Zhang, "Design of audio signal processing system based on semiconductor laser," LASER JOURNAL, Vol. 42, No. 6, June 2021.