A Hybrid SFANC-FxNLMS Algorithm for Active Noise Control based on Deep Learning

Zhengding Luo, Student Member, IEEE, Dongyuan Shi, Member, IEEE, and Woon-Seng Gan, Senior Member, IEEE

Abstract—The selective fixed-filter active noise control (SFANC) method selecting the best pre-trained control filters for various types of noise can achieve a fast response time. However, it may lead to large steady-state errors due to inaccurate filter selection and the lack of adaptability. In comparison, the filtered-X normalized least-mean-square (FxNLMS) algorithm can obtain lower steady-state errors through adaptive optimization. Nonetheless, its slow convergence has a detrimental effect on dynamic noise attenuation. Therefore, this paper proposes a hybrid SFANC-FxNLMS approach to overcome the adaptive algorithm’s slow convergence and provide a better noise reduction level than the SFANC method. A lightweight one-dimensional convolutional neural network (1D CNN) is designed to automatically select the most suitable pre-trained control filter for each frame of the primary noise. Meanwhile, the FxNLMS algorithm continues to update the coefficients of the chosen pre-trained control filter at the sampling rate. Owing to the effective combination of the two algorithms, experimental results show that the hybrid SFANC-FxNLMS algorithm can achieve a rapid response time, a low noise reduction error, and a high degree of robustness.

Index Terms—Active noise control, selective fixed-filter ANC, deep learning, hearables.

I. INTRODUCTION

Active noise control (ANC) involves the generation of an anti-noise of equal amplitude and opposite phase to suppress the unwanted noise [1]–[7]. Owing to its compact size and high efficiency in attenuating low frequency noises, ANC has been widely used in industrial operations and consumer products such as headphones [8]–[14]. Traditional ANC systems typically employ adaptive algorithms to attenuate noises [15]–[17]. Filtered-X normalized least-mean-square (FxNLMS), being an adaptive ANC algorithm, not only can achieve low noise reduction errors through adaptive optimization but also does not require prior knowledge of the input data [18]–[20]. However, due to the least mean square (LMS) based algorithms’ inherent slow convergence and poor tracking ability [21]. FxNLMS is less capable of dealing with rapidly varying or non-stationary noises. Its slow response to these noises may impact customers’ perceptions of noise reduction effect [22].

In industrial applications and mature ANC products, fixed-filter ANC methods [23], [24] are adopted to tackle the problems of adaptive algorithms, where the control filter coefficients are fixed rather than adaptively updated to avoid slow convergence and instability. However, the coefficients of the fixed-filter algorithm are pre-trained for a specific noise type, resulting in the degradation of noise reduction performance for other types of noises [25]. Therefore, for flexible selection of different pre-trained control filters in response to different noise types, Shi et al. [26] proposed a selective fixed-filter active noise control (SFANC) method based on the frequency band matching. Nevertheless, several critical parameters of the method can only be determined through trials and errors. Also, its filter-selection module consumes extra computations in the real-time processor [27].

Thereby, automatic learning of the SFANC algorithm’s critical parameters is a significant issue that would promote its uses in real-world scenarios and industrial products. Under this consideration, deep learning techniques, particularly convolutional neural networks (CNNs), achieving significant success in classification [28]–[31] appear to be an excellent approach for improving SFANC. By replacing the filter-selection module with a CNN model, the SFANC algorithm can automatically learn its parameters from noise datasets and select the best control filter given different noise types without resorting to extra-human efforts [32]. Additionally, a CNN model implemented on a co-processor can decouple the computational load from the real-time noise controller [32]. Though its superiority on response speed has been demonstrated, the CNN-based SFANC algorithm is limited by some wrong noise classifications and the lack of adaptive optimization, which may result in large steady-state errors.

As discussed above, using either SFANC or FxNLMS is difficult to constantly obtain satisfactory noise reduction performance for various noises. The SFANC algorithm and FxNLMS algorithm can be effectively combined to overcome the limitations of the individual ANC algorithms. Therefore, a hybrid SFANC-FxNLMS algorithm is proposed in this paper. In the proposed algorithm, a lightweight one-dimensional CNN (1D CNN) implemented on a co-processor can automatically select the most suitable control filter for different noise types. For each frame of the noise waveform, the control filter coefficients are decided by the 1D CNN and continue to be updated adaptively via the FxNLMS algorithm at the sampling rate. Simulation results show that the proposed algorithm not only achieves faster noise reduction responses and lower steady-state errors but also exhibits good tracking and robustness. Thus, it can be used for attenuating dynamic noises such as traffic noises and urban noises etc.
The remainder of this paper is organized as follows: Section II introduces the hybrid SFANC-FxNLMS algorithm in detail. Section III presents a series of numerical simulations to evaluate the effectiveness of the proposed algorithm. Finally, the conclusion is given in Section IV.

II. HYBRID SFANC-FxNLMS ALGORITHM

The overall framework of the hybrid SFANC-FxNLMS algorithm is presented in Fig. 1. Although learning acoustic models directly from the raw waveform data is challenging [33], the input of the proposed ANC system is raw waveform instead of frequency spectrogram.

A. A concise explanation of SFANC

An ANC progress can be abstracted as a first-order Markov chain [34], [35], as shown in Fig. 2, where \( w_o(n) \) represents the optimal control filter to deal with the disturbance \( d(n) \). If we assumed that the optimal control filter has \( C \) discrete states as \( \{ w_i \}_{i=1}^C \), the control filter selection process can be considered as determining the best control filter from a pre-trained filter set \( \{ w_i \}_{i=1}^C \). Hence, the SFANC strategy can be summarized as follows:

\[
\mathbf{w}_o = \arg\min_{\mathbf{w} \in \{\mathbf{w}_i\}_{i=1}^C} E \left\{ \left[ d(n) - \mathbf{x}^T(n) \mathbf{w} * s(n) \right]^2 \right\}, \tag{1}
\]

where \( \mathbf{w} \) returns the input value for minimum output; \( * \), \( \mathbf{x}(n) \), and \( s(n) \) represent the linear convolution, the reference signal, and the impulse response of the secondary path, respectively. The reference signal is assumed to be the same as the primary noise.

In practice, \( d(n) \) is typically regarded as the linear combination of \( x(n) \). Thus, (1) is equivalent to

\[
\mathbf{w}_o = \arg\max_{\mathbf{w} \in \{\mathbf{w}_i\}_{i=1}^C} P[\mathbf{w}|d(n)] = \arg\max_{\mathbf{w} \in \{\mathbf{w}_i\}_{i=1}^C} P[\mathbf{w}|x(n)], \tag{2}
\]

which refers to the process of selecting a control filter from the pre-trained filter set that maximizes its posterior probability in the presence of reference signal \( x(n) \). Moreover, according to Bayes’ theorem [36], the posterior probability in (2) can be replaced with a conditional probability as

\[
\mathbf{w}_o = \arg\max_{\mathbf{w} \in \{\mathbf{w}_i\}_{i=1}^C} P[x(n)|\mathbf{w}], \tag{3}
\]

which predicts the most suitable control filter straight from the reference signal \( x(n) \).

To achieve a satisfactory noise reduction for the SFANC algorithm, we should find a classifier model \( P[\mathbf{x}(n)|\mathbf{w}, \Theta] \) to approximate \( P[\mathbf{x}(n)|\mathbf{w}] \) from the pre-recorded sampling set \( \{ \mathbf{x}(n), \mathbf{w} \}_j^N \). The \( \Theta \) denotes the parameters of the classifier and can be obtained through maximum likelihood estimation (MLE) [37] as

\[
\Theta = \arg\max_{\Theta} \frac{1}{N} \sum_{j=1}^N \log \hat{P}[\mathbf{x}(n)|\mathbf{w}, \Theta]. \tag{4}
\]

Hence, we can utilize the deep learning approach to extract this classifier model from the training set \( \{ \mathbf{x}(n), \mathbf{w} \}_j^N \).

B. CNN-based SFANC algorithm

Motivated by the work [32], this paper develops a 1D CNN capable of dealing with noise waveforms end-to-end. The input of the network is the normalized noise waveform with a 1-second duration. The min-max operation is defined as:

\[
\hat{x}(n) = \frac{x(n)}{\max[x(n)] - \min[x(n)]}, \tag{5}
\]

where \( \max[\cdot] \) and \( \min[\cdot] \) obtain the maximum and minimum value of the input vector \( x(n) \). It aims to rescale the input range into \((-1, 1)\) and retain the signal’s negative part that contains phase information.

1) Pre-training control filters: The primary and secondary paths used in the training stage of the control filters are band-pass filters with a frequency range of 20Hz-7.980Hz. Broadband noises with 15 frequency ranges shown in Fig. 3 are used to pre-train 15 control filters. The filtered reference least mean square (FxLMS) algorithm is adopted to obtain the optimal control filters for these broadband noises due to its low computational complexity [38]. Subsequently, these 15 pre-trained control filters are saved in the control filter database.

2) The Proposed 1D CNN: The detailed architecture of the 1D CNN is illustrated in Fig. 4. Every residual block in the network is composed of two convolutional layers, subsequent batch normalization [39], and ReLU non-linearity [40]. Note that a shortcut connection is adopted to add the input with
Table I

| Network                  | Test Accuracy | #Layers | #Parameters |
|--------------------------|---------------|---------|-------------|
| Proposed Network         | 99.35%        | 6       | 6.21M       |
| M3 Network               | 99.15%        | 3       | 0.22M       |
| M5 Network               | 99.10%        | 5       | 0.56M       |
| M11 Network              | 99.15%        | 11      | 1.79M       |
| M18 Network              | 98.50%        | 18      | 3.69M       |
| M34-res Network          | 98.35%        | 34      | 3.99M       |

Notably, the co-processor operates at the frame rate, whereas the real-time controller performs at the sampling rate.

The hybrid SFANC-FxNLMS algorithm can achieve faster responses to different noises with the assistance of the SFANC. Meanwhile, the FxNLMS algorithm continues to update the control filter’s coefficients in the SFANC-FxNLMS algorithm, which contributes to obtain lower noise reduction errors.

III. Experiments

In the experiments, we trained the 1D CNN model on a synthesized noise dataset and then evaluated the hybrid SFANC-FxNLMS algorithm on real-record noises. The primary path and secondary path are chosen as band-pass filters. Moreover, the step size of the FxNLMS algorithm is set to 0.002, and the control filter length is 1,024 taps.

A. Training and testing of 1D CNN

The synthetic noise dataset was subdivided into three subsets: 80,000 noise tracks for training, 2,000 noise tracks for validation, and 2,000 noise tracks for testing. During training, the Adam algorithm [43] was used for optimization. The training epoch was set at 30. To avoid exploding or vanishing gradients, the glorot initialization [44] was used. To avoid overfitting due to the scarcity of training data, the 1D network was built as a light-weighted network with weights all subjected to $\ell_2$ regularization with a coefficient of 0.0001.

The proposed 1D CNN is compared to several different 1D networks proposed in [33] utilizing raw acoustic waveforms: M3, M5, M11, M18, and M34-res. Table I summarizes the comparison results. The number of network layers includes the number of convolutional layers and fully connected layers. Among these networks, the proposed 1D network obtains the highest classification accuracy of 99.35% on the test dataset, which means it can provide the most suitable pre-trained control filters for different noises. Also, the proposed 1D network is lightweight, utilizing the fewest network parameters possible due to its efficient architecture.

B. Real-world noise cancellation

This section presents using the SFANC, FxNLMS, and hybrid SFANC-FxNLMS algorithm to attenuate real-record noises. The real-world noises are not included in the training datasets. We created two distinct noises by cascading and mixing an aircraft noise with a frequency range of 50Hz-1,400Hz and a traffic noise with a frequency range of 40Hz-1,400Hz.

1) Cascaded noise attenuation: The varying noise in this simulation is provided by cascading aircraft noise and traffic noise. The noise reduction results using different ANC methods on the cascaded noise are shown in Fig. 5.
of dealing with new noise during the first 1s, since it only updates the control filter index after the first 1s. However, owing to combining with FxNLMS, the hybrid SFANC-FxNLMS algorithm can slightly attenuate the noise during the first 1s. Furthermore, we can see that the averaged noise reduction level of hybrid SFANC-FxNLMS algorithm is about 10dB higher than that of SFANC and FxNLMS algorithms in 12s-13s. Therefore, the results on the mixed noise also confirm the superiority of the hybrid SFANC-FxNLMS algorithm over the SFANC and FxNLMS algorithm.

C. Noise reduction performance on the varying primary path

In this section, the robustness of different ANC algorithms are examined when dealing with a primary path that changes every 13 seconds. The impulse response of the primary path is mixed with different levels of white noise to obtain additional signal-to-noise ratios (SNR) of 30dB and 10dB respectively as shown in Fig. 7. The primary noise is the traffic noise.

According to Fig. 7 in the initial period, the SFANC and hybrid SFANC-FxNLMS algorithm have better noise reduction performances over the FxNLMS algorithm. However, when the SNR of the primary path varies, SFANC can’t perform as well as FxNLMS and hybrid SFANC-FxNLMS algorithm. Since the primary noise type remains the same, the selected pre-trained control filter is fixed in the SFANC algorithm, which may not be suitable on the varying primary path. The results indicate that the hybrid SFANC-FxNLMS algorithm is sufficiently robust to deal with the varying primary path.

IV. CONCLUSION

The proposed hybrid SFANC-FxNLMS algorithm effectively combines SFANC and FxNLMS to achieve faster noise reduction responses and lower steady-state errors. A lightweight 1D CNN operating at frame rate is used to automatically determine control filter coefficients. Also, the FxNLMS algorithm continues to update the coefficients adaptively at the sampling rate. Simulation results validate the good tracking capability and robustness of the hybrid SFANC-FxNLMS algorithm when dealing with varying noises and primary paths in dynamic noise control environments. The proposed hybrid algorithm is promising to be further explored for combining other fixed-filter and adaptive ANC algorithms.
