Development of Active Hear-Through Equalization Algorithm for Earphones

Chong-Rui Huang, Cheng-Yuan Chang and Sen M. Kuo
Chung Yuan Christian University, Taoyuan, Taiwan
E-mail: rayhuang@cycu.org.tw Tel: +886-981659930
E-mail: ccy@cycu.edu.tw Tel: +886-3-2654838
E-mail: kuo1065@gmail.com Tel: +886-3-2654851

Abstract— **This research developed an active hear-through equalization (AHE) algorithm for earphones. Using a digital signal processor as the computing core, this proposed work put a microphone on the housing of the earphone to receive ambient sound and then created a pseudo-ambient sound in the eardrum of the human ear to achieve the hear-through (HT) function. In addition, this research analyzed the dependence on the direction-of-arrival (DOA) for the HT function. Finally, we compared the performance of the proposed work with the leading commercial earphones to verify its enhancement.**

I. INTRODUCTION

In recent years, advances in hardware and software technologies have facilitated the ear-worn devices boom in acoustic experiences. Today, the headphone market in terms of audio technology has made active noise control (ANC), hear-through (HT) mode, and spatial audio standard functions in headphones, providing users with diverse listening experiences [1], [2]. However, to make a better life, many acoustical scientists and scholars continue to develop new headphones with the associated techniques. Where include the multi-reference control strategies [3], [4], the comb-effect [5], the virtual sensing techniques [6], the directions of sound arrival [7], [13], and the occlusion effect [8], etc.

The HT function is an audio technology of augmented reality (AR). It receives the real sound from the external microphone to produce the virtual sounds by the HT filter, then creates pseudo-ambient sound at the eardrum to compensate for the passive noise isolation caused by the earphones. To provide users with the feel of natural listening experience as if there were without earphones over the ear [9]. Many scholars have presented research results on augmented reality audio [2], [10]-[12]. In recent years, Gan et al. described an approach for natural and augmented listening for virtual reality (VR), augmented reality (AR), and mixed reality (MR) [1], and then analyzed the characteristics of AR audio applied to different types of headphones [9]. Patel et al. proposed a selective reduction in environmental noise while providing the HT capability in the looking direction [14]. Liebich et al. proposed active occlusion cancellation (AOC) system for headphones and hearing aids to reduce the occlusion effect [8]. From the above research, it can find the importance of earphones to be equipped with assistive listening.

This research implements the HT function with earphones and then presents an AHE algorithm to train the HT filter. Simulations were used to verify the proposed HT function and analyzed the effects of the DOA dependence on the HT function. Finally, the effectiveness of the proposed system was assured by real-time experiments.

II. SYSTEM DESIGN

This section first describes the proposed AHE algorithm and the experimental setup. This research conducts experiments in an anechoic chamber to conduct a series of subsequent analyses and comparisons to verify the effectiveness of the proposed system.

A. AHE Algorithm

This research proposes a single-channel AHE algorithm based on the modified filtered-x least mean square (FxLMS) algorithm to implement the augmented sound in earphones, as shown in Fig. 1. The AHE algorithm has two stages, the tuning and equalization stages, to equalize the real and virtual sounds by implementing the pseudo-ambient sound.

Fig. 1a shows the tuning stage. For obtaining the equalization filter $\hat{W}(z)$ to generate the virtual signal $y_T(n)$, the proposed algorithm will first model the target path $T(z)$ and secondary path $S(z)$. Obtain a reasonable estimate of the target path $\hat{T}(z)$ and secondary path $\hat{S}(z)$ is the key to success during the tuning stage. This research uses adaptive system identification using the LMS algorithm [15] to estimate their transfer functions. The random noise and the pink noise are used as the excitation signals to estimate the secondary path $S(z)$ and the target path $T(z)$. The target path $T(z)$ refers to the acoustic path without earphone between the reference and the error microphones. The secondary path $S(z)$ refers to the secondary speaker to the error microphone, which includes the digital-to-analog converter (DAC), reconstruction filter, power amplifier, acoustic path from secondary speaker to error microphone, preamplifier, anti-aliasing filter, and analog-to-digital converter (ADC). We designed a band-pass filter $\hat{F}_{BP}(z)$ according to the different DOA to attenuate the undesired target signal $d_T(n)$ to produce desired target signal $d'_T(n)$. In addition, the delay $z^{-\Delta}$ is used to ensures the causality of the AHE algorithm. The error signal $e_T(n)$ is expressed as

$$e_T(n) = d'_T(n-\Delta) - y_T(n), \quad (1)$$

where $d'_T(n)$ is the desired target signal; $y_T(n)$ is the virtual signal; n is the time index with sampling rate f_S ; Δ is the number of delay samples. The adaptive finite impulse response (FIR) equalization filter $\hat{W}(z)$ generates virtual signal $y_T(n)$ using the relevant reference signal $\mathbf{x}'_T(n)$ expressed as

$$y_T(n) = \hat{\mathbf{w}}^T(n)\mathbf{x}'_T(n), \quad (2)$$

where the T in the upper right denotes the transpose operation. $\hat{\mathbf{w}}(n) \equiv [w_0(n)\ w_1(n)\ \cdots\ w_{L-1}(n)]^T$ is the weight vector; and $w_l(n)\ (l = 0,...,L\text{-}1)$ is the l^{th} coefficients of equalization filter $\hat{W}(z)$ with filter length L . $\mathbf{x}'_T(n) \equiv [x'_T(n)\ x'_T(n-1)\ \cdots\ x'_T(n-L+1)]^T$ is the relevant reference signal vector. The equalization filter $\hat{W}(z)$ updated by

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu\ [\hat{\mathbf{s}}(n) * \mathbf{x}_T(n)]\ e_T(n), \quad (3)$$

where $*$ denotes the linear convolution, $\hat{\mathbf{s}}(n)$ is the impulse responses of the secondary path $\hat{S}(z)$, and the μ is the step size. With sufficient step size, the $\hat{\mathbf{w}}(n)$ will converge to the optimal value to minimize the error signal is $e_T(n)$.

Next, Fig. 1b, shows the equalization stage, where the primary path $P(z)$ refers to the acoustic path with earphones between the reference and the error microphones. Here the reference microphone is used to capture the external ambient signal $x_P(n)$. The equalization filter $\hat{W}(z)$ estimated in the tuning stage is used to generate the virtual signal $y_P(n)$, which is then acoustically superimposed on the real signal $d_P(n)$ through the secondary path $S(z)$ to compensate for the loss caused by the passive noise isolation of the earphones. Consequently, it gets clear that the feedforward equalization system results in the performance expressed as

$$\frac{E_P(z)}{X_P(z)} = P(z) + S(z)\hat{W}(z). \quad (4)$$

B. Experimental Setup

To verify that the proposed AHE algorithm can be effectively applied to earphones. This research conducted all experiments in an anechoic chamber (Internal dimensions: 4.45m (L) x 4.0m (W) x 2.07m (H)) that complies with ISO 3745 (2003) and ASTM E336 (2009), as shown in Fig. 2a.

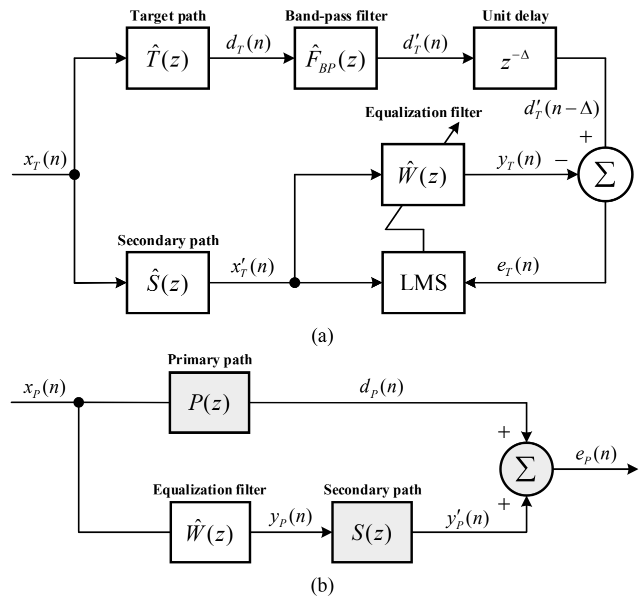


Fig. 1 Block diagram of the single-channel AHE system based on the modified FxLMS algorithm, (a) tuning stage, (b) equalization stage.

Where the primary loudspeaker (Tang Band, W8-1808) was placed about 30 cm away from the torso simulator's right ear (G.R.A.S, KEMAR) to generate the broadband pink noise, where the sound pressure level of KEMAR's right ear is about 80 dB(A). A dynamic signal analyzer (Keysight, 35670A) is used to measure the spectrum from the right ear microphone of the KEMAR. In addition, a rotating disk under the KEMAR, which has the 360° rotation angle, is used to evaluate the DOA dependence (a total of 8 angles, each with a 45° difference), as shown in Fig. 2b. This experimental setup is used to conduct analysis and verify the performance of the proposed method. This research uses commercially available earphones (Sennheiser, IE 40 Pro) and sets an additional MEMS microphone (Knowles, SPU0414HR5H-SB) as the reference microphone on the shell to test the performance of the proposed work. The testing earphone is inserted inside the KEMAR’s right artificial ear. The right ear microphone of KEMAR was used as the error microphone in this experiment. This research uses a floating-point DSP (Texas Instruments, TMS320C6713 DSP Starter Kit) with the ADC and DAC converter interface card (HEG Company, DSK6713IF-B) in the experiments. The sampling rate is 48 kHz, and the cut-off frequency of the anti-aliasing filter and reconstruction filter is 20 kHz.

Finally, this experiment uses pink noise played at the same level in an anechoic chamber to compare commercially available earphones (Apple, AirPods Pro) and earphones by this research to perform a fair comparison. The validity of the proposed AHE algorithm is then verified by adopting the experiments.

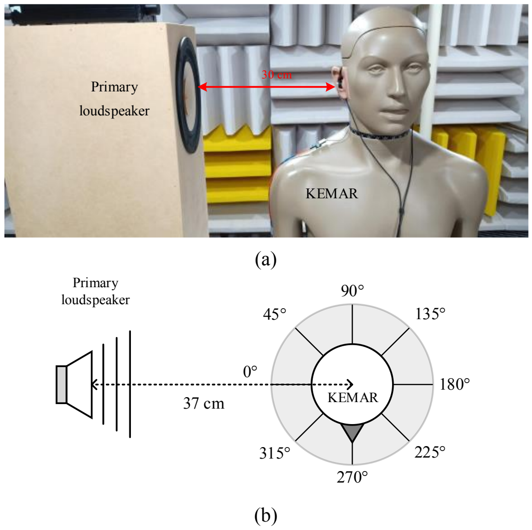


Fig. 2 (a) Experimental setup, (b) The setup for testing different DOA.

III. Implementation Results and Discussion

This section first analyzes the performance of earphones (Sennheiser, IE 40 Pro). Then uses simulations and real-time experiments to evaluate the effectiveness of the proposed AHE algorithm under different DOA. Also, we will make a comparison between the HT function performance of leading commercial earphones too.

A. Analysis and Measurements

Before implementing the proposed system, it was necessary to analyze and measure the earphones we purchased to understand whether it was necessary to implement the HT function in the earphones. Different types of earphones have different passive noise isolation characteristics. This research used earphones (Sennheiser, IE 40 Pro) with silicones tips were chosen since it strong passive noise isolation of external sounds. We analyze the passive attenuation performance of earphones with different sound DOA.

Fig. 3 shows the passive noise isolation of earphone, where light blue lines are the passive noise isolation measured for each of the 8 rotation angles, and blue line by averaging these 8 curves. It shows the exhibits strong isolation at high-frequency noise, at about 20 dB at 2 kHz. However, the passive material of the earphone cannot effectively reduce low-frequency noise below 500 Hz. We can find out the earphone attenuates the high-frequency information above 500 Hz from the passive noise isolation.

Fig. 4 shows that the magnitude response of the target path $T(z)$ is affected by the different sound DOA, where light gray lines are the target path measured for each of the 8 rotation angles, and red line by averaging these 8 curves. Obviously, as the frequency increases to 4 kHz, the magnitude response consistency gradually changes, which causes degrade the performance of HT equalization with a fixed equalization filter $\hat{W}(z)$. Finally, through the above analysis, we know that HT function is only effective in handling a bandwidth range between 100 Hz to 4 kHz because this range shows low DOA dependency, and the effect of DOA limits the high-frequency ranges. Therefore, this research introduces a bandpass filter $\hat{F}_{BP}(z)$ to reduce unwanted noise amplification. However, the sound with higher frequency is susceptible to DOA dependent. Therefore, this research used 4th-order infinite impulse response (IIR) band-pass filter $\hat{F}_{BP}(z)$, where the f_L as 100Hz (-3 dB); and f_H as 6 kHz (-3 dB), as shown in Fig. 5.

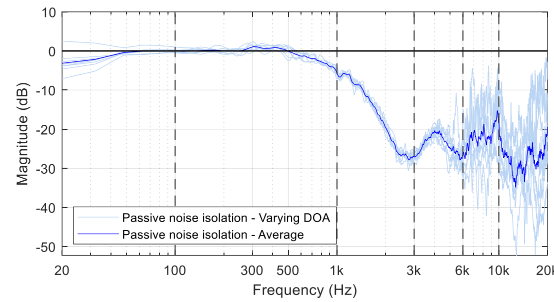


Fig. 3 The passive noise isolation of earphone with different sound DOA.

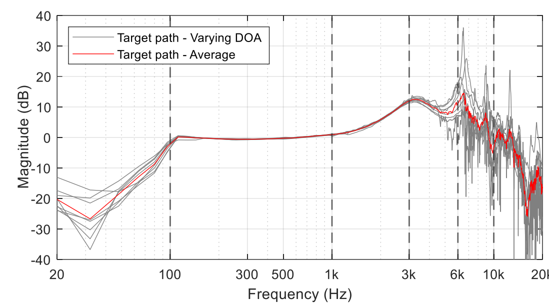


Fig. 4 Magnitude responses of the target path $T(z)$ with different sound DOA.

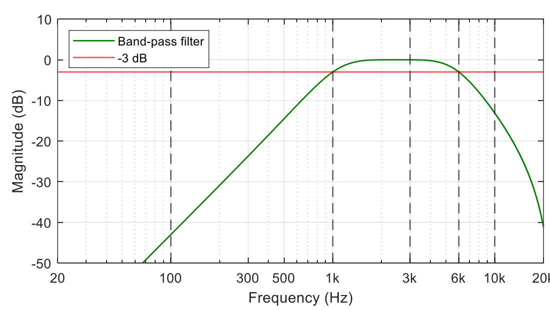
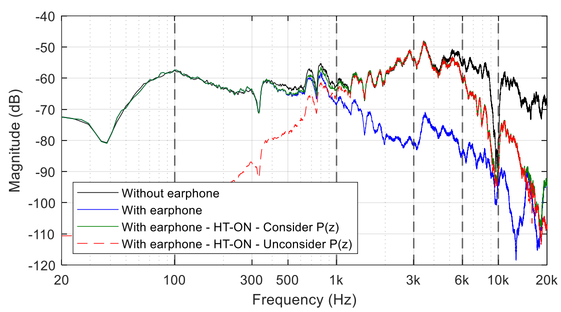


Fig. 5 Magnitude responses of the band-pass filter $\hat{F}_{BP}(z)$.

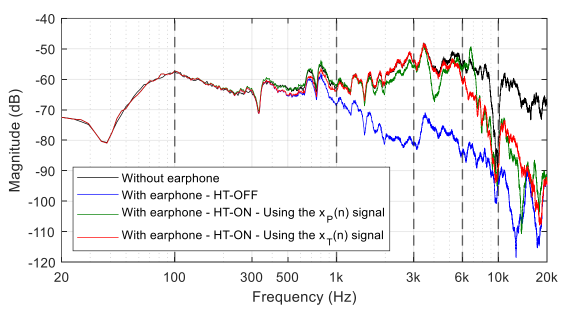
B. Simulation Results

In the simulations, this research used MATLAB software to analyze the performance of the proposed AHE algorithm for the earphones. First, we substitute the equalization filter trained during the tuning stage into the equalization stage to verify the effectiveness of the proposed AHE algorithm. The unit delay required for the modified FxLMS algorithm to converge was found to be 8 samples in the tuning stage, and the lengths of the designed FIR equalization filter $\hat{W}(z)$ is 180, and of the estimated FIR transfer function of the secondary path $\hat{S}(z)$ is 150 taps.

Since the effect of the primary path $P(z)$ is not considered in the tuning stage. However, we must take it into account in the real environment, the effects with and without the primary path $P(z)$ in the equalization stage are experimentally analyzed, as shown in Fig. 6a. We can be found that when the primary path $P(z)$ is unconsidered, the equalization filter $\hat{W}(z)$ designed by the band-pass filter $\hat{F}_{BP}(z)$ in the tuning stage can effectively compensate for the passive noise isolation caused by earphones in the equalization stage and only focus on the spectrum in the 1-6 kHz bandwidth range. In addition, we can find that when considering the primary path $P(z)$, the equalization filter $\hat{W}(z)$ is not needed to compensate for the low frequency below 1 kHz because the external sound of low-frequency is not attenuated by passive noise isolation. We can see that the transfer function of primary path does not affect the high frequency of HT function. The AHE algorithm proposed in this research assumes that the reference microphone signal of the target path $\hat{T}(z)$ and the primary path $P(z)$ must be the same (that is $x_T = x_P$). Therefore, we compare the different reference signal on the AHE system, and we can find that the HT function is better using the reference signal x_T of the target path. Unfortunately, the actual application uses the reference signal x_P from the primary path $P(z)$, as shown in Fig. 6b. Finally, the effectiveness of the AHE algorithm was verified by the simulation results.



(a) Analyze the effect of the primary path $P(z)$.



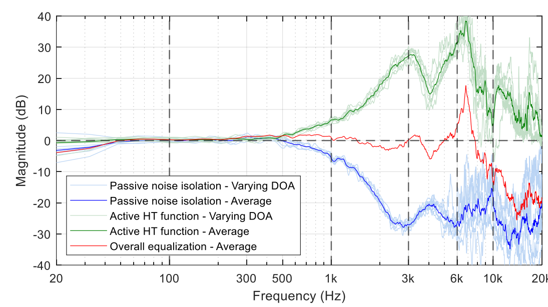
(b) Analyze the effect of the different reference signals $x_P(n)$ and $x_T(n)$.

Fig. 6 Simulation of the equalization stage of the AHE algorithm to verify HT function, where the spectrums are taken from the error signal $e_P(n)$.

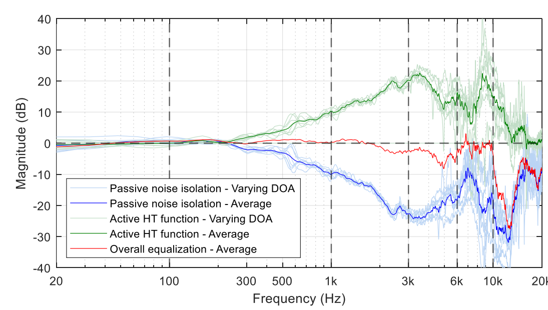
C. *Real-Time Experiments*

Real-time experiments were conducted to verify the active HT performance of the proposed AHE algorithm in real conditions. The length of the designed FIR equalization filter $\hat{W}(z)$ was 180 taps was the same as in the simulation. The pink noise was used as the excitation signal to analyze whether the proposed active HT function perfectly compensated for the passive noise isolation.

Fig. 7 shows the DOA dependence of passive noise isolation and active HT function of earphone, where the light blue lines and light green lines are the passive noise isolation and active HT function measured for each of the 8 rotation angles, and then the blue and green lines by averaging these 8 curves. The curve of the red line is the overall equalization of the earphones derived by the blue line equalization the green line, where show that the frequency-dependent gain in the hearing range (20 Hz-20 kHz) of a normal human hearing. Therefore, values below 0 dB correspond to sound attenuation, and values above 0 dB correspond to sound amplification. We can analyze whether the AHE algorithm compensates for passive noise isolation by the overall equalization curve (black line), where 0 dB means perfect compensation. Fig. 7 shows the experimental results of the two pairs of earphones that the high frequency above 4 kHz is easily affected by the different DOA. This research uses a fixed equalization filter $\hat{W}(z)$. Therefore, it is not possible to adapt to high-frequency different sound DOA. However, before 4 kHz, the proposed research can effectively compensate for passive noise isolation, and the results of HT function are close to the commercial earphones.



(a) Earphone by this research - Sennheiser, IE 40 Pro.



(b) Commercial earphones - Apple, AirPods Pro.

Fig. 7 Analyze the equalization performance of HT function.

The overcompensation between 7-8 kHz (red line) is caused by the bad design of the band-pass filter $\hat{F}_{BP}(z)$. We should reduce the cutoff frequency f_H to 4 kHz of the band-pass filter to avoid high-frequency overcompensation caused by the different DOA.

IV. Conclusion

This research implemented the AHE algorithm based on the modified FxLMS algorithm to derive the equalization filter for earphone to generate the pseudo-ambient sound of augmented sound. We verified the effectiveness of the proposed AHE algorithm through simulations and real-time experiments. It showed that the proposed system had good active HT performance before 4 kHz with the different sound DOA. The experimental results showed that the proposed earphone with HT function had similar results to the leading commercially available earphones.

References

[1] W. S. Gan, J. He, R. Ranjan, and R. Gupta, "Natual and augmented listening for VR/AR/MR," *ICASSP 2018 tutorial*, Calgary, Canada, Apr. 2018 [Online]. Available: https://sigport.org/documents/icassp-2018-tutorial-t11-natual-and-augmented-listening-vrarmr. Accessed: Jul. 2021.

[2] V. Valimaki, A. Franck, J. Ramo, H. Gamper, and L. Savioja, "Assisted listening using a headset: Enhancing audio perception in real, augmented, and virtual environments," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 92-99, Mar. 2015.

[3] V. Belyi, W. S. Gan, "A combined bilateral and binaural active noise control algorithm for closed-back headphones," *Applied Acoustics*, vol. 160, ID 107129, Mar. 2020.

[4] J. Cheer, V. Patel, S. Fontana, "The application of a multi-reference control strategy to noise cancelling headphones," *The Journal of the Acoustical Society of America*, vol. 145, no. 5, pp. 3095-3103, May. 2019.

[5] J. Rämö and V. Välimäki, "Digital augmented reality audio headset," *Journal of Electrical and Computer Engineering*, vol. 2012, pp. 1-13, Oct. 2012.

[6] C. Shi, R. Xie, N. Jiang, H. Li and Y. Kajikawa, "Selective Virtual Sensing Technique for Multi-channel Feedforward Active Noise Control Systems," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8489-8493, May. 2019.

[7] S. Liebich, J.-G. Richter, J. Fabry, C. Durand, J. Fels and P. Jax, "Direction-of-Arrival dependency of active noise cancellation headphones," *INTERNOISE 2018*, Aug. 2018.

[8] S. Liebich, R. Brandis, J. Fabry, P. Jax and P. Vary, "Active Occlusion Cancellation with Hear-Through Equalization for Headphones," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 241-245, Apr. 2018.

[9] R. Gupta, R. Ranjan, and J. He, W. S. Gan, "On the Use of Closed-Back Headphones for Active Hear-Through Equalization in Augmented Reality Applications," *AES International Conference on Audio for Virtual and Augmented Reality*, ID P8-2, Aug. 2018.

[10] R. Ranjan and W. S. Gan, "Natural listening over headphones in augmented reality using adaptive filtering techniques," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 11, pp. 1988-2002, Nov. 2015.

[11] S. Yao, "Headphone-based immersive audio for virtual reality headsets," *IEEE Transactions on Consumer Electronics*, vol. 63, no. 3, pp. 300-308, Aug. 2017.

[12] A. Härmä, J. Jakka, M. Tikander, M. Karjalainen, T. Lokki, J. Hiipakka, and G.Lorho, "Augmented reality audio for mobile and wearable appliances," *Journal of the Audio Engineering Society*, vol. 52, no. 6, pp. 618-639, June. 2004.

[13] R. Gupta, R. Ranjan, J. He and W. S. Gan, "Parametric Hear through Equalization for Augmented Reality Audio," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1587-1591, 2019.

[14] V. Patel, J. Cheer and S. Fontana, "Design and Implementation of an Active Noise Control Headphone With Directional Hear-Through Capability," *IEEE Transactions on Consumer Electronics*, vol. 66, no. 1, pp. 32-40, Feb. 2020.

[15] S. M. Kuo and D. R. Morgan, *Active Noise Control Systems - Algorithms and DSP Implementations*, John Wiley & Sons, New York, NY, 1996.