A constrained optimal hear-through filter design approach for earphones

Yongjie Zhuang[1]
Ray W. Herrick Laboratories, Mechanical Engineering, Purdue University
177 S, Russell Street, West Lafayette
IN 47907-2099, USA

Yangfan Liu[2]
Ray W. Herrick Laboratories, Mechanical Engineering, Purdue University
177 S, Russell Street, West Lafayette
IN 47907-2099, USA

ABSTRACT

Signal characteristics can be altered when sound from environment transmits through earphones to the ear canal. A hear-through filter is usually implemented in an earphone to create a more natural hearing experience. Hear-through filter is also important in augmented reality audio applications. In this paper, a constrained optimal hear-through filter design approach is proposed, where the hear-through filter is designed using a formulation similar to that for a constrained active noise control filter design. One advantage of such a filter design approach is that, compared with the commonly used direct plant response inversion method, the leakage sound around and through earphone will be attenuated in the proposed method, so that the comb-filtering effect is alleviated. Another advantage of the proposed method is that multiple practical constraints can be applied conveniently by formulating a constrained optimization problem and it can be solved efficiently. The proposed design approach can specify the desired delays of reproduced sound in each earphone channel, if a spatial sound impression is desired. The designed hear-through filter can be directly implemented in an active noise control system framework so that the requirement for additional electronic hardware and software components can be minimal for an active noise control earphone.

1. INTRODUCTION

In recent years, earphones are being used more and more frequently in daily life because of the fast development of the digital music and smartphones. Various functions were incorporated into earphone design to improve the user experience. Among those functions, active noise control (ANC), equalization techniques, and a hear-through function have become popular functions of many earphones [1].

When wearing a earphone, the signal characteristics can be altered when sound from outside environment transmits through the earphone to the ear canal. This makes the person wearing an

[1]zhuang32@purdue.edu

[2]yangfan@purdue.edu

earphone hears an unnatural environment sound. A hear-through filter is usually implemented to process the sound signals measured at the microphones at the exterior side of an earphone so that appropriate sound can be reproduced by the earphone speaker to create a more natural hearing experience. A well-performing hear-through function, also referred to as ambient mode or transparent mode of an earphone, allows one to hear the ambient sound more clearly and realistically while wearing an earphone. The hear-through filter can also be an important technique in achieving better augmented reality audio (ARA) performance [2–4].

Various approaches have been proposed to design a hear-through filter in previous literature. They can be classified into two categories: design of a direct inverse, and filter design using an ANC structure. For the direct inverse filter approach, a hear-through filter can be designed by flattening the attenuation curve caused by the earphone and (or) the ear canal, e.g., the allpass filter design [5]. However, a certain level of ambient sound will usually be transmitted into the ear canal as leakage through the earphone which deteriorates the performance of hear-through filter and causes a comb-filtering effect [4, 6], i.e., the resulting sound signal will be osculating in the spectrum due to the different delay between reproduced sound and leakage sound. The hear-through filter can also be designed by using an ANC structure, e.g., the use of LMS based algorithm [2], and directional hear-through design can also be incorporated [7]. Since an ANC filter can cancel the sound leakage, the comb-filtering effect can, in principle, be attenuated. However, some practical limitations of designed hear-through filter have not been considered in previous studies yet. For example, when acoustic feedback path exists, i.e., the sound reproduced by the speaker can propagate back to the reference microphone outside the earphone, the control system becomes a closed loop which means robust stability is required to be considered. This is more important for earphones that are designed not to fully cover the ear.

In this paper, the proposed hear-through filter design approach is based on an ANC structure. Instead of considering the attenuation of sound leakage into the ear canal and the reproduction of ambient sound separately, the proposed method formulates them as a single optimization problem such that the hear-through filter can be optimal while the leakage sound is attenuated at the same time. An H_2/H_∞ framework that was previously used in ANC structure [8, 9] is adopted to formulate the hear-through filter design problem. The robust stability and hear-through filter response amplitude limitation are considered by formulating the design problem as a constrained optimization problem. The results show that comb-filtering effect can be significantly attenuated by the proposed method compared with the method that designs an inverse filter directly. Also, the constraints ensure that the stability requirement is satisfied.

2. THEORY

2.1. Control System Description

The system block diagram of the proposed hear-through filter structure is shown in Fig. 1(a), which is similar to a typical active noise control system. This system includes one microphone at the exterior surface of an earphone as the reference sensor of the ANC system, one loudspeaker unit in the earphone as the secondary source, one microphone located in the ear canal as the error sensor. The error microphone is not needed in principle after the hear-through filter is designed.

In the control diagram, $\mathbf{x}$ denotes the incoming ambient sound signal from the environment measured at the reference sensor location. $\mathbf{r}$ is the reference signal measured by the reference microphone when the control system is activated, thus, it includes both the ambient sound and the reproduced sound at the reference microphone position. $\mathbf{y}$ is the output of the hear-through filter. $\mathbf{d}_e$ denotes the ambient sound that transmit around and through the earphone to the ear canal. $\mathbf{e}$ includes both the leakage sound and the reproduced sound at error sensor location. $\tilde{\mathbf{x}}$ is the signal of $\mathbf{x}$ with an application dependent delay. The total power of $\tilde{\mathbf{e}}$ is to be minimized in the hear-through filter design process, i.e., the reproduced sound in ear canal should resemble a delayed ambient signal $\mathbf{x}$.

The added delay $\Delta = e^{-j2\pi f\delta}$ is due to the physical limitation that, by adding electronic system and filtering process, additional delay will be inevitably introduced. An advantage of this filter design structure is that the expected delay of reproduced sound can be controlled easily to satisfy particular requirement in applications such as the creation a spatial impression of sound for augmented reality.

$\mathbf{G}_s$ is the acoustic feedback path that represents the acoustic responses of the secondary source at the reference sensor. It is noted that the reference signal $\mathbf{r}$ is a combination of the ambient sound signal, $\mathbf{x}$, and the reproduced signal from the acoustic feedback path $\mathbf{G}_s$. To estimate the ambient sound signal, $\mathbf{x}$, from reference signals $\mathbf{r}$, an internal model control (IMC) structure is implemented, similar to that in an ANC system [10]. In the controller $\mathbf{H}$, the $\hat{\mathbf{G}}_s$ is a model of the physical feedback path, $\mathbf{G}_s$; thus, this cancels the acoustic feedback path effect. In the current work, it is assumed that $\hat{\mathbf{G}}_s$ is a perfect model, i.e., $\hat{\mathbf{G}}_s = \mathbf{G}_s$, the estimated noise signals $\hat{\mathbf{x}}$ will be the same as the true primary noise signals $\mathbf{x}$ and this system becomes a standard feedforward system in Fig. 1(b). The modeling error in this process is treated as a plant uncertainty and is considered in the robustness constraint. $\mathbf{G}_e$ represents the acoustical responses of the secondary source at the error sensor position.

$\mathbf{W}_x$ is the frequency response of the hear-through FIR filters that need to be designed. The FIR filter coefficients of the hear-through filter, denoted as $\vec{\mathbf{w}}_F$, are the variables that need to be calculated in the filter design problem. The optimal solution $\vec{\mathbf{w}}_F^*$ obtained by solving the filter design problem can be directly implemented in the real-time signal processing controller via time-domain filtering. It is noted that the hear-through filter structure presented here is similar to an ANC filter structure which can be directly implemented in an ANC system in the earphone so that the requirement for additional electronic or software components is minimal if this hear-through function is to be added into an ANC earphone.

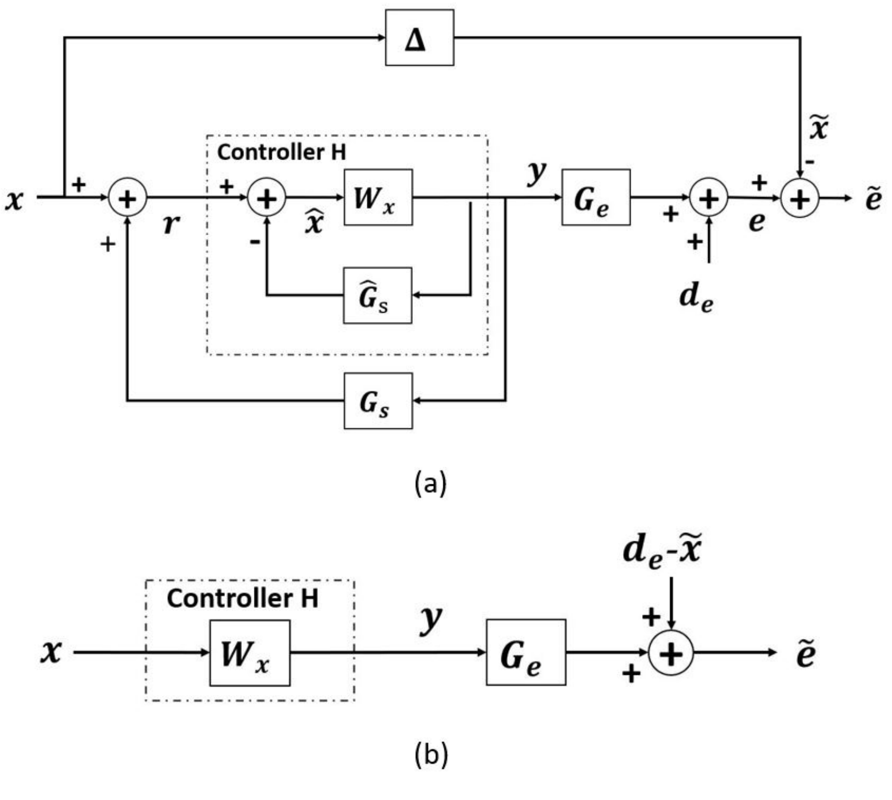

Figure 1: Block diagram of a hear-through filter structure: (a) system with an acoustic feedback path and an internal model control structure, (b) an equivalent feedforward system when assuming a perfect cancellation of the acoustic feedback path effect.

2.2. H_2/H_∞ Formulation

Firstly, frequency response of the FIR filter can be expressed as:

$$\mathbf{W}_x(f) = \vec{\mathbf{F}}_z^{\mathrm{T}}(f)\vec{\mathbf{w}}_F, \tag{1}$$

where

$$\vec{\mathbf{F}}_z(f) = \begin{bmatrix} 1 & e^{-j2\pi f \frac{1}{f_s}} & e^{-j2\pi f \frac{2}{f_s}} & \ldots & e^{-j2\pi f \frac{N_t-1}{f_s}} \end{bmatrix}^{\mathrm{T}};$$

f denotes the frequency; f_s denotes the sampling frequency.

The objective function $J_0(\vec{\mathbf{w}}_F)$ to be minimized in the filter design process is the total power of the error signal, $\tilde{\mathbf{e}}$, across all desired frequencies:

$$J_0(\mathbf{w}_F) = \sum_{k=1}^{N_f} J(f_k), \tag{2}$$

where the $J(f_k)$ is the power of $\tilde{\mathbf{e}}$ at k-th frequency f_k. N_f denotes the total number of frequency points in the desired hear-through band. $J(f_k)$ can be expressed as:

$$\begin{aligned} J(f_k) &= E\left[\tilde{\mathbf{e}}(f_k)\tilde{\mathbf{e}}^*(f_k)\right] \\ &= E\left[\left(\mathbf{d}_e(f_k) - e^{-j2\pi f_k \delta}\mathbf{x}(f_k) + \mathbf{G}_e(f_k)\mathbf{W}_x(f_k)\mathbf{x}(f_k)\right)\left(\mathbf{d}_e(f_k) - e^{-j2\pi f_k \delta}\mathbf{x}(f_k) + \mathbf{G}_e(f_k)\mathbf{W}_x(f_k)\mathbf{x}(f_k)\right)^*\right] \\ &= \left|\mathbf{G}_e(f_k)\mathbf{W}_x(f_k) - e^{-j2\pi f_k \delta}\right|^2 \mathbf{S}_{xx}(f_k) + 2\Re\{\mathbf{G}_e(f_k)\mathbf{W}_x(f_k)\mathbf{S}_{xd_e}(f_k)\} + \mathbf{S}_{d_e d_e}(f_k), \end{aligned} \tag{3}$$

where E is the expectation operator; $*$ denotes the complex conjugate operation; $\Re()$ denotes the real part of a complex number; $\mathbf{S}_{xx}(f_k)$ and $\mathbf{S}_{d_e d_e}(f_k)$ are the power spectral density of $\mathbf{x}$ and $\mathbf{d}_e$ respectively at frequency f_k; $\mathbf{S}_{xd_e}(f_k)$ is the cross spectral density between $\mathbf{x}$ and $\mathbf{d}_e$ at frequency f_k, which is defined as $\mathbf{S}_{xd_e} = \lim_{T\to\infty} \frac{1}{T}\mathrm{E}[\mathfrak{F}(\mathbf{x})\mathfrak{F}(\mathbf{d}_e)^*]$ where $\mathfrak{F}$ denotes the Fourier transform and T is the segment length of the signals. By minimizing this objective function, the designed hear-through filter can be optimal in terms of hear-through function while attenuating the leakage sound simultaneously.

Due to the existence of acoustic feedback cancellation path $\hat{\mathbf{G}}_s$ in the controller $\mathbf{H}$, the stability and robustness constraints are required to be considered. It is noted that the stability and robustness constraints can be formulated exactly the same as in the single-input single-output (SISO) ANC optimal design problem when the proposed hear-through structure is used. The controller stability is ensured by limiting the open-loop response trajectory of controller $\mathbf{H}$ to be at the right hand side of the Nyquist point in the Laplace domain [10], which can be expressed as:

$$\Re\{\mathbf{W}_x(f_k)\hat{\mathbf{G}}_s(f_k)\} > -1. \tag{4}$$

For robustness constraint, the $\mathbf{M}$-$\boldsymbol{\Delta}$ structure and the small-gain theory is applied [10]. The robustness constraint can be expressed as:

$$\left|\mathbf{W}_x(f_k)\hat{\mathbf{G}}_s(f_k)\right| B(f_k) \le 1, \tag{5}$$

where $B(f_k)$ is the upper bound on the output multiplicative plant uncertainty at frequency f_k.

In practical applications, a constraint on the amplitude of the hear-through filter response, $\mathbf{W}_x(f_k)$, is usually needed to ensure that the loudspeaker is operating in its linear response range. Another reason is that, in some applications, the power consumption of the loudspeaker needs to be limited. The filter response amplitude constraint can be expressed as:

$$|\mathbf{W}_x(f_k)| \le C(f_k) \tag{6}$$

where $C(f_k)$ is the required upper bound on the frequency response amplitude of filter $\mathbf{w}_F$ at frequency f_k. Note that, if $C(f_k)$ is small enough and Eq. (6) is satisfied at some frequencies, Eqs. (4), (5) are also satisfied at those frequencies. So for those frequency bands where $C(f_k)$ is small, Eq. (6) can be used to replace Eqs. (4), (5) to reduce the calculation effort of solving the optimization problem. Also, it is found that when the energy of the disturbance signal is small at some frequency bands, e.g., near Nyquist frequency due to anti-aliasing filtering, the use of response limits, Eq. (6), to replace Eqs. (4), (5) on these frequency bands can help improve the numerical behavior during solving the optimization problem.

The optimization problem for the hear-through filter design can be constructed by using Eq. (2) as the objective function, and Eqs. (4), (5), (6) as constraints:

$$\begin{aligned} \min_{\mathbf{w}_F}. \quad & \sum_{k=1}^{N_f} J(f_k) \\ \text{s.t.} \quad & \Re\{\mathbf{W}_x(f_k)\hat{\mathbf{G}}_s(f_k)\} > -1, \quad \text{for all } f_k, \\ & \left|\mathbf{W}_x(f_k)\hat{\mathbf{G}}_s(f_k)\right| B(f_k) \le 1, \quad \text{for all } f_k, \\ & |\mathbf{W}_x(f_k)| \le C(f_k), \quad \text{for all } f_k. \end{aligned} \tag{7}$$

The formulated optimization problem has the same form as the H_2/H_∞ formulation in a constrained ANC filter design problem [8, 9]. By using the same technique and process in the work of Zhuang and Liu's [8, 9], the problem specified in Eq. (7) can be convexified and reformulated as cone programming, which is a type of convex programming that can be solved efficiently using primal-dual interior-point method.

3. RESULTS

3.1. Experimental Setup Description

To investigate the proposed method for designing a hear-through filter. An earphone is used in an experiment to collect the required system response data. There are one reference microphone and one speaker in the earphone. The error microphone used in this experiment is in an artificial ear (Hangzhou Aihua AWA6162 (IEC711)). The pictures of the earphone mounted on the artificial ear is shown in Fig. 2. It is noted that a non-adaptive controller is implemented in the current work using nominal system characteristics $\mathbf{G}_e$ measured using the setup shown in Fig. 2.

When acquiring measurement data, the sampling rate of data acquisition system was set to be 48000 Hz to prevent aliasing. Two million sampling points (about 42 seconds) for each channel were acquired for calculating the $\mathbf{G}_e$. A hamming window of 48000 points (1 second) is used for averaging with fifty percent overlapping (83 times averages). The length of designed hear-through FIR filter is 128. The sampling rate for hear-through filter is 24000 Hz. The desired hear-through band is below 6000 Hz.

3.2. Investigation of Comb-Filtering Effect

As mentioned in the introduction, if a direct inverse filter design approach is being used to flatten the response of $\mathbf{G}_e$, there will be comb-filtering effect due to the sum of leakage sound and reproduced sound. However, if the proposed method is used, it will reproduce the desired sound field while attenuating the leakage sound. Thus, the comb-filtering effect should be alleviate. To demonstrate this effect, the performance of using the direct inverse filtering design approach and the proposed method is compared. In this simulation, secondary path measured using the earphone in the previous section is used. The reference sound signal and leakage sound signal are generated in simulation to compared the performance of different methods. $\mathbf{x}$ is a white noise processed by a low-pass filter with a cut-off frequency of 6000 Hz. The leakage sound signal, $\mathbf{d}_e$ is around 6 dB lower than the ambient sound signal at the reference microphone, $\mathbf{x}$, and has a 1 ms lag with respect to $\mathbf{x}$. The desired sound $\tilde{\mathbf{x}}$ has a 2 ms lag with respect to $\mathbf{x}$, i.e., $\delta = 2 \times 10^{-3} s$. So the there is relative delay between the desired reproduced sound and the leakage sound.

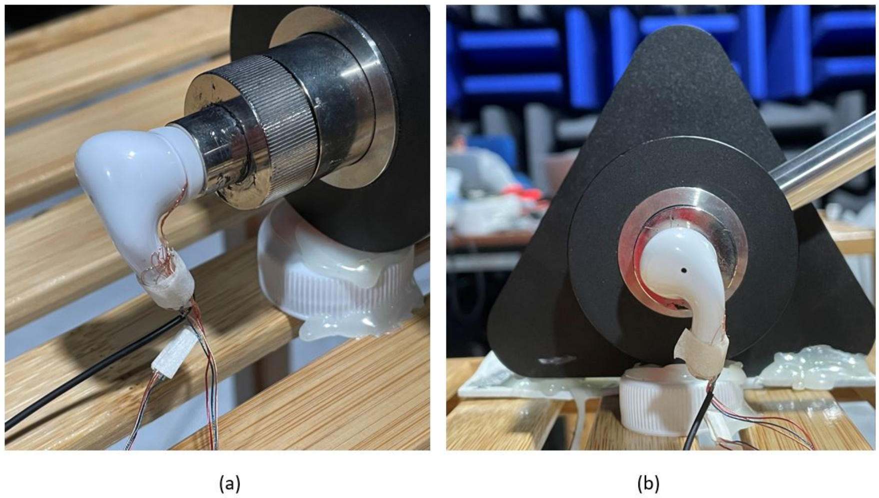

Figure 2: Pictures of the earphone mounted on AWA6162 (IEC711) for experimental data collection.

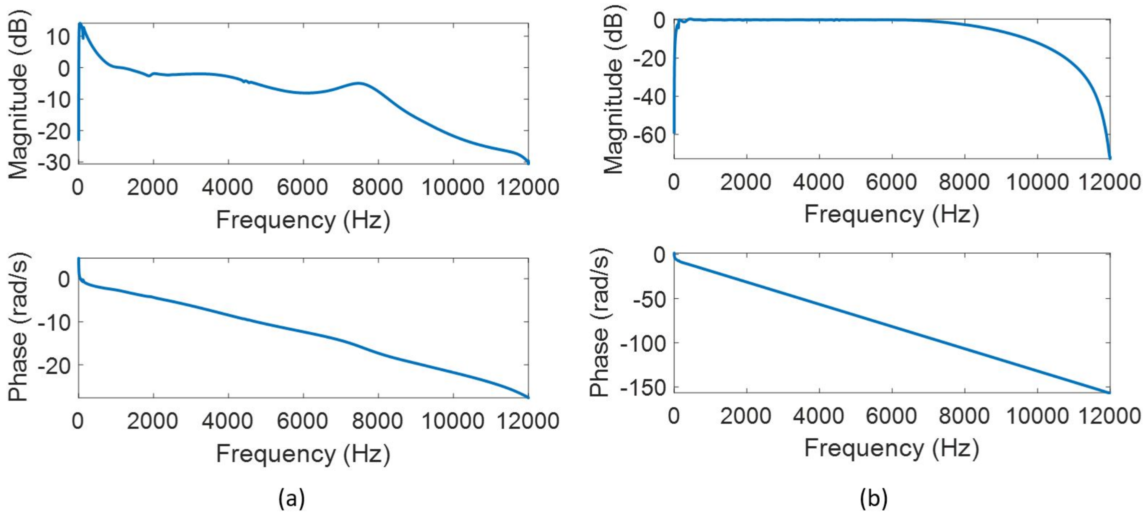

Figure 3: Frequency response of (a) $\mathbf{G}_e$, (b) $\mathbf{G}_e\mathbf{W}_x$, where $\mathbf{W}_x$ is designed using direct inverse filter method. The reference value for dB scale is 1.

The frequency response of measured $\mathbf{G}_e$ is shown in Fig. 3(a) which includes the attenuate due to earphone speaker unit response and the ear canal's acoustic response. Fig. 3(b) shows the frequency response of $\mathbf{G}_e\mathbf{W}_x$, where $\mathbf{W}_x$ is designed by using the direct inverse filter method (i.e., performing a least square estimation to fit $\frac{1}{\mathbf{G}_e}$). From Fig. 3(b), it demonstrates that the designed hear-through filter can flatten the attenuation curve caused by earphone and ear canal. However, observed from Fig. 4(a), due to the leakage sound around and through earphone, the contribution of the leakage sound and the reproduced sound leads to a comb-filtering effect in the total sound field, when a direct inverse filter method is used. On the other hand, the proposed method does not have the comb-filtering effect as shown in Fig. 4(a). By checking the impulse response of filters designed by two methods (shown in Fig. 4(b)), the filter designed by proposed method has significant response before 10^{-3} s. This is the filter response that actively reduces the leakage sound, since leakage sound has 1 ms delay with respect to the ambient sound signal while the reproduced sound requires 2 ms delay compared with ambient sound at reference microphone. This result shows that the proposed method can significantly alleviate the comb-filtering effect by attenuating the leakage sound.

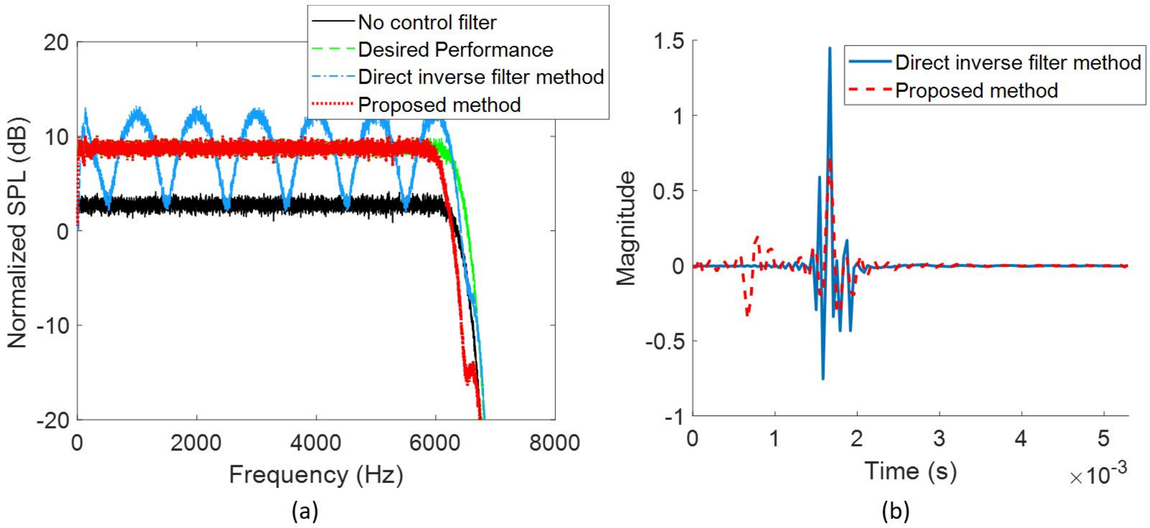

Figure 4: Comparison of (a) hear-through performance in normalized sound pressure level (SPL), and (b) impulse response for direct inverse filter design method and proposed method.

3.3. Investigation of Applied Constraints

Another advantage of proposed method is the capability of applying multiple constraints to satisfy practical industrial requirements. In this section, the capability of constraining the designed hear-through filter is investigated. Firstly, since the desired hear-through band is below 6000 Hz, the magnitude of hear-through filter response is constrained above 6000 Hz as shown in Fig. 5(a). Then stability constraint is focused on in this study. In this simulation, the acoustic feedback path $\mathbf{G}_s$ is assumed to have a low-pass type frequency response with 3 dB attenuation in the passband which is shown in Fig. 5(b). Fig. 6 shows the Nyquist plot of designed control system's open loop response, i.e., $\mathbf{W}_x\mathbf{G}_s$. It is clearly that if the stability constraint is not applied, the control system is unstable as shown in Fig. 6(a). When the stability constraint is applied, the control system becomes stable as shown in as shown in Fig. 6(b). Fig. 5(a) also shows that adding stability constraint will sacrifice the hear-through performance near the cut-off frequency of designed hear-through filter. The hear-through performance below 4000 Hz is still satisfactory.

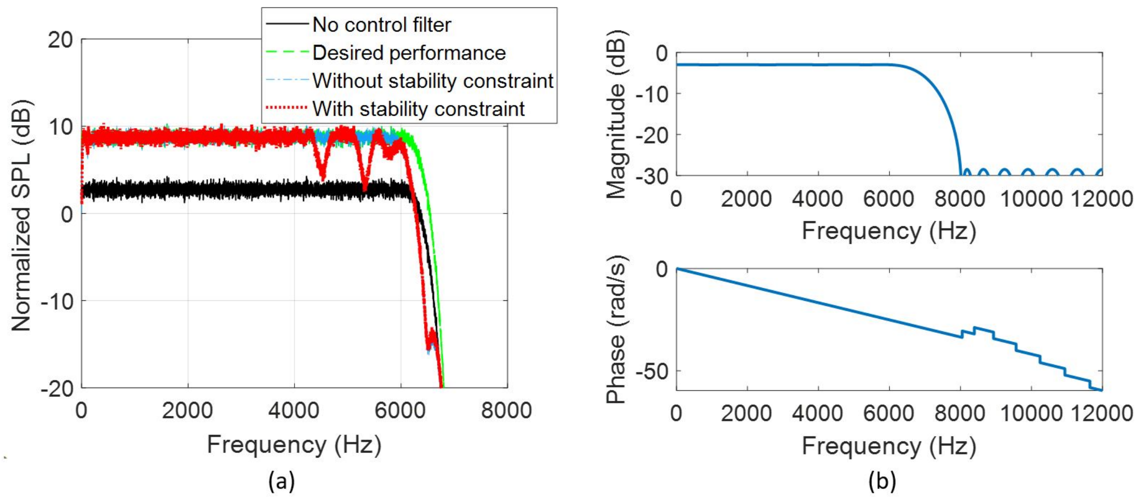

Figure 5: (a) Comparison of hear-through filter performance with and without applying stability constraint; (b) The frequency response of $\mathbf{G}_s$.

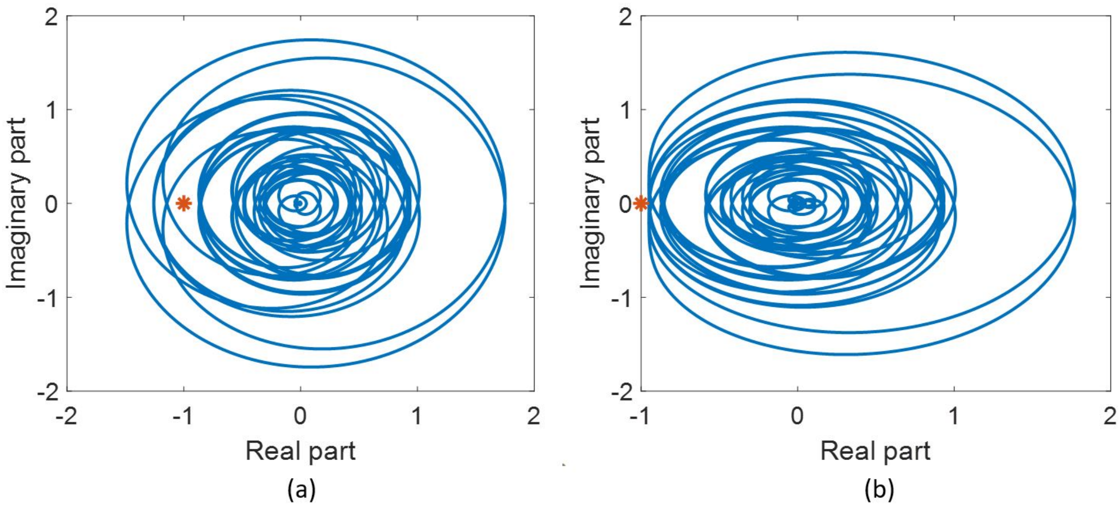

Figure 6: The Nyquist plot (i.e., frequency response of open loop $\mathbf{W}_x\mathbf{G}_s$) for hear-through filter $\mathbf{W}_x$ designed (a) without applying stability constraint, and (b) with stability constraints.

4. CONCLUSIONS

A constrained optimal hear-through filter design approach is proposed in this article. The hear-through control filter structure is combined with ANC structure to attenuate the leakage sound and alleviate the comb-filtering effect. The robust stability constraint required by the existence of the acoustic feedback path can be applied. It is shown that the design problem can be formulated as a constrained optimization problem similar to ANC filter design method. The results show that compared with the direct inverse filter design approach, the proposed method can effectively attenuate the comb-filtering effect. The applied filter magnitude constraint and stability constraint can ensure the corresponding practical requirements to be satisfied.

It is noted that the proposed method has the potential to be expanded to multichannel situations, which will have a wider applications of reproducing a desired sound field. For example, to provide hear-through function in an automobile when the ambient sound is desired to be heard clearly without opening window or door of the vehicle. In the future, the capability of multichannel hear-through filter design will be explored.

ACKNOWLEDGEMENTS

The authors thank Beijing Ancsonic Technology Co. Ltd for providing financial support and help on the experimental system setup and data collection related to the presented work.

REFERENCES

[1] Vesa Valimaki, Andreas Franck, Jussi Ramo, Hannes Gamper, and Lauri Savioja. Assisted listening using a headset: Enhancing audio perception in real, augmented, and virtual environments. *IEEE Signal Processing Magazine*, 32(2):92–99, 2015.

[2] Rishabh Gupta, Rishabh Ranjan, Jianjun He, and Gan Woon-Seng. On the use of closed-back headphones for active hear-through equalization in augmented reality applications. In *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*. Audio Engineering Society, 2018.

[3] Rishabh Gupta, Rishabh Ranjan, Jianjun He, and Woon Seng Gan. Parametric hear through equalization for augmented reality audio. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1587–1591. IEEE, 2019.

[4] Jussi Rämö and Vesa Välimäki. Digital augmented reality audio headset. *Journal of Electrical and Computer Engineering*, 2012, 2012.

[5] Jussi Rämö and Vesa Välimäki. An allpass hear-through headset. In *2014 22nd European Signal Processing Conference (EUSIPCO)*, pages 1123–1127. IEEE, 2014.

[6] Juho Liski. Adaptive hear-through headset. Master's thesis, School of Electrical Engineering, Aalto University, 2016.

[7] Vinal Patel, Jordan Cheer, and Simone Fontana. Design and implementation of an active noise control headphone with directional hear-through capability. *IEEE Transactions on Consumer Electronics*, 66(1):32–40, 2019.

[8] Yongjie Zhuang and Yangfan Liu. Study on the cone programming reformulation of active noise control filter design in the frequency domain. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 260, pages 126–136. Institute of Noise Control Engineering, 2019.

[9] Yongjie Zhuang and Yangfan Liu. Development and application of dual form conic formulation of multichannel active noise control filter design problem in frequency domain. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 261, pages 676–687. Institute of Noise Control Engineering, 2020.

[10] S.J. Elliott. *Signal Processing for Active Control*, chapter 6, pages 271–327. Signal Processing and its Applications. Academic Press, London, 2001.