PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Estimation of Cross-Talk Compensation Filter using Bone Conduction Ear Microphone

IRWANSYAH[+,*]; Tsuyoshi USAGAWA[+]

[+]Kumamoto University, Japan

ABSTRACT

Having a bone conduction (BC) transducer anywhere on the head will result in a BC sound reaching the cochlea in both ears. This "cross-talk" phenomenon may limit a listener's ability with a pair of BC transducer to sense sound direction. In this paper, we discuss a way to minimize cross-talks in BC sound reproduction using a method called "cross-talk cancellation." Ideally, the method requires transfer functions (TFs) from each of the BC transducers to each of the listener's cochleae to accurately synthesize cross-talk compensation (CTC) filters; however, a direct measurement of the TFs at the cochlea is not possible. Since the ear canal is the closest to the cochlea, we thus hypothesize that the TFs measured at the ear canal might be used to achieve cross-talk cancellation at the cochlea in the inner ear. Therefore, we utilized a BC ear microphone to capture the vibration of the bony ear-canal caused by vibrating transducers on the mastoid. The filtered-x least mean square (FxLMS) algorithm was then used to estimate the CTC filter. Experiments with and without cross-talk cancellation were done to determine the effective frequency range that could achieve cancellation in the ear canal of three normal-hearing participants.

Keywords: Cross-Talk Cancellation, Bone Conduction Ear Microphone, FxLMS

1. INTRODUCTION

Binaural hearing refers to our ability to perceive a sound with two ears. It enables us to use the relative differences in the sound as it is transmitted by air conduction (AC) to both ears as cues to sense the sound direction. The AC-sound-arrival time and intensity differences between the two ears, known as binaural cues, provide us robust information regarding the sound direction. Besides, when a target speech source has been localized and is in a different direction from interferers or noise, the speech can be more easily understood even in the presence of multiple interferers (1). Such benefits are usually found in normal individuals who hear with air conduction. However, the benefits of binaural hearing for sound heard through bone conduction (BC) are less than for that heard through air conduction (2); such as in the case of patients with conductive hearing loss who may require bone conduction hearing aids/devices to perceive sound.

Conductive hearing loss (CHL) is a decrease in perceived sound due to a problem in the outer or middle ear that prevents the sound from reaching the cochlea in the inner ear. For individuals with permanent CHL, a bone conduction device is an efficient alternative to regular AC-based hearing aids. It works by delivering sound directly to the cochlea in the inner ear by bone vibration; however, unlike AC sound, BC sound from a single BC transducer can reach the cochlea in both ears as illustrated in Fig. 1. As a result, binaural cues tend to be distorted by the BC pathway. This "cross-talk" or "cross-hearing" phenomenon is then considered by researchers as one of the factors responsible for limiting the benefits of binaural processing for patients fitted bilaterally with BC devices (2, 3, 4).

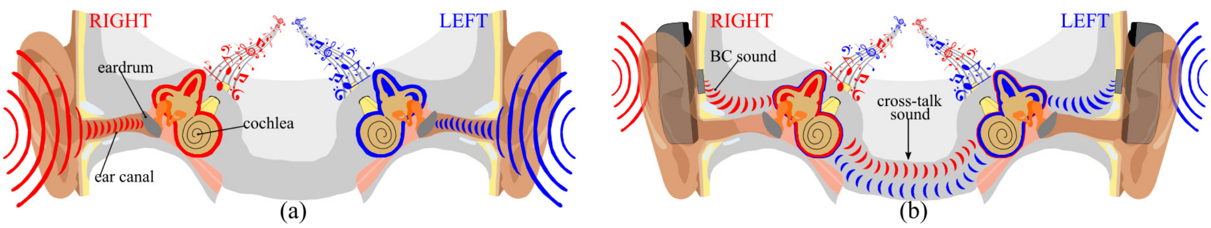

Figure 1 – (a) Air conduction hearing; (b) Vibrating transducers on the mastoid transmitting BC sounds.

[*]irwansyah@ieee.org

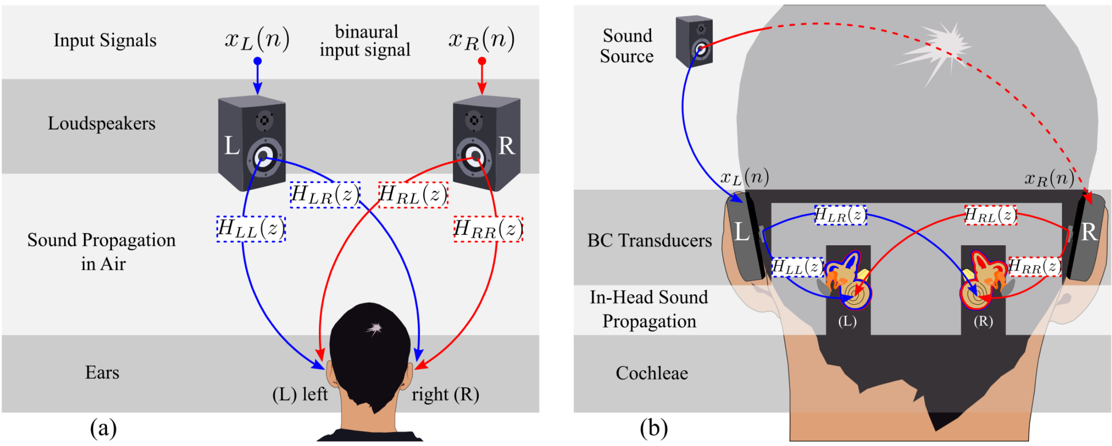

Figure 2 – Cross-talk in binaural sound reproduction through (a) a two-loudspeaker system analogous to that using (b) a pair of BC transducers.

Sound from a BC transducer is transmitted not only to the ipsilateral cochlea but also to the contralateral cochlea as "cross-talk" sound. For this reason, researchers sometimes argue whether the binaural application of BC devices is beneficial or not. See Fig. 2(b); if the greater the intensity difference between the transfer functions of $H_{RR}(z)$ and $H_{RL}(z)$ (from the right BC transducer to the ipsilateral and contralateral cochleae), the more beneficial the binaural hearing will be. Over the past three decades, many attempts have been made to study this "cross-talk" issue. Nolan and Lyon (1981) (5) and Stenfelt (2012) (6) used subjective measures to investigate the transcranial attenuation, which is defined as the difference in sensitivity between BC sounds transmitted from one BC transducer to the ipsilateral and contralateral cochleae. Stenfelt found the median transcranial attenuation to be 3–5 dB at frequency up to 500 Hz with BC stimulation at the mastoid; it is about 0 dB at 0.5–1.8 kHz, and close to 10 dB at the frequency between 3 and 5 kHz (6). Even though the attenuation is rather small, depending on the individual, but it can also indicate that listeners with two BC devices might be able to preserve binaural cues to some degree in the presence of cross-talk sound.

It is well known that binaural cues play an important role not only in sound localization but also in the separation of speech in noise. When target speech and noise have distinct binaural cues or they are spatially separated, normal-hearing individuals can understand speech better at a low signal-to-noise ratio (SNR) (7). Stenfelt and Zeitooni (2013) showed in their speech-based tests that the binaural benefit with BC stimulation for normal-hearing participants was only about half of the benefit with AC stimulation in terms of SNR (8). On the other hand, localization tests conducted with three experienced bone-anchored hearing aid (BAHA) users in (3) and with twelve normal-hearing individuals in (4) revealed that listeners show relatively good localization accuracy only for angles less than 45°. Although binaural benefits are present as reported in (3, 4, 8), improvements are still necessary in order to assist listeners in better localizing and segregating a desired target speech from the interfering speech and/or noise.

Cross-talk that occurs between a two-loudspeaker system and the listener's ears as illustrated in Fig. 2 is basically analogous to that in the human head during BC sound reproduction. So, a "cross-talk cancellation" method which is a solution to minimize cross-talks in reproduction of binaural sounds via loudspeakers (9), might also be applied to BC sound reproduction. Liao (2010) showed that the cross-talk cancellation method can be applied to a dry human skull, achieved by placing accelerometers at the locations of the cochleae to estimate BC transfer functions (TFs) (10). In the skull, it works at the frequency between 0.25 and 1.5 kHz with the cross-talk attenuation of up to 12 dB, depending on frequency (10). However, applying the method to humans is an still issue since it is not possible to directly observe BC sound at the cochlea. Therefore, a non-invasive way to estimate TFs is required and preferable.

The closest and easily accessible position that might help to estimate BC sounds at the cochlea in live humans is in the ear canal. Reinfeldt (2013) reported that the vibration of the cochlea due to BC stimulation at the mastoid is relatively similar to the vibration of the bony ear-canal (11). We thus hypothesize that the TFs measured from each of the BC transducers to each of the listener's ear canals might be used to achieve cross-talk cancellation at the cochlea in the inner ear. We then utilized a BC ear microphone to capture the vibration of the bony ear-canal caused by BC transducers on the mastoid; the time-stretched pulse (TSP) is used as a stimulus to estimate the TFs (12). The filtered-x least mean square (FxLMS) algorithm was then used to estimate a cross-talk compensation (CTC) filter, which is used for generating a cancellation signal. Experiments with and without cancellation were also done to determine the effective frequency range that could achieve cancellation in the ear canal of three normal-hearing participants.

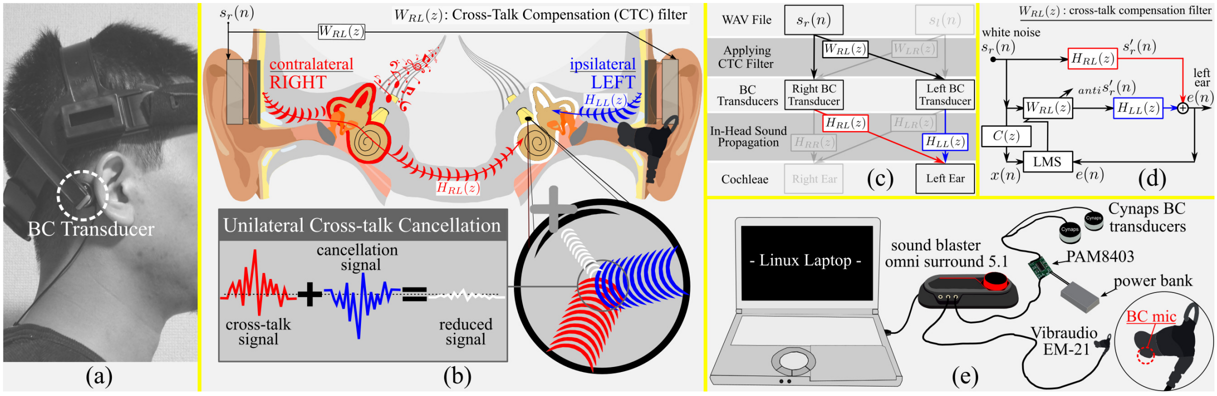

Figure 3 – (a) Image of the right side view of a participant wearing a helmet headband with attached BC transducers. (b) Illustration of the unilateral cross-talk cancellation for the left ear. (c) Block diagram of the unilateral cross-talk cancellation. (d) Block diagram of the cross-talk compensation (CTC) filter estimation using the FxLMS algorithm. (e) Diagram of the measurement system setup.

2. UNILATERAL CROSS-TALK CANCELLATION

In this paper, we mainly focus on implementing a one-sided or unilateral cross-talk cancellation system in order to confirm whether or not BC sounds observed in the ear canal can be used to achieve cancellation at the cochlea in the inner ear. To do this, participants required to wear a helmet headband with attached BC transducers to help maintain a stable acoustic coupling between the transducers and the head, see Fig. 3(a). BC transducers were placed on the left and right mastoid, and a BC ear microphone was inserted only into the left ear canal; the left and right sides are referred to as the ipsilateral and contralateral sides, respectively. Figure 3(b) illustrates the unilateral cross-talk cancellation system (for the case of the left ear) in which a cross-talk sound from a contralateral BC transducer is canceled by an anti-sound from the ipsilateral one. To achieve such cancellation, the cross-talk compensation (CTC) filter $W_{RL}(z)$ is introduced between the right ear input signal $s_r(n)$ and the ipsilateral (left) BC transducer as shown in Fig. 3(c). The input signal $s_r(n)$ is then passed to the CTC filter to create the anti-sound (or cancellation signal). Details about how the CTC filter, $W_{RL}(z)$, is estimated are explained in the sub-section below.

2.1 Cross-Talk Compensation Filter

In this paper, the CTC filter is a linear filter whose coefficients are adjusted based on the FxLMS algorithm to achieve destructive interference in the target ear. Figure 3(c) presents the diagram of the unilateral cross-talk cancellation for the left ear, where $W_{RL}(z)$ is the CTC filter. Because it is not possible to directly observe a BC sound at the cochlea, we then define $H_{RL}(z)$ and $H_{LL}(z)$ as the transfer functions (TFs) from the contralateral and ipsilateral BC transducers to the BC ear microphone, respectively. In our case, the CTC filter $\mathbf{w}_{RL}(n)$ is not estimated in real time, but we first measure the impulse responses $\mathbf{h}_{RL}(n)$ and $\mathbf{h}_{LL}(n)$ – $H_{RL}(z)$ and $H_{LL}(z)$ in the frequency domain – using TSP sound. Then, the obtained impulse responses are used to estimate the CTC filter by computer simulation.

The diagram of the CTC filter estimation using the FxLMS algorithm is shown in Fig. 3(d). The weight vector of the CTC filter at time n is defined as $\mathbf{w}_{RL}(n) = [w_{RL_0}(n), w_{RL_1}(n), \cdots, w_{RL_{N_w-1}}(n)]^T$ where N_w is the filter tap-length. Since the CTC filter is estimated by simulation, a generated white noise is used as the reference signal $s_r(n)$ and then filtered by the obtained impulse response $\mathbf{h}_{RL}(n)$ to simulate the cross-talk sound $s'_r(n)$. On the other hand, the cancellation signal $_{anti}s'_r(n)$ is obtained by filtering the white noise $s_r(n)$ with the CTC filter $\mathbf{w}_{RL}(n)$.

$$_{anti}s'_r(n) = \mathbf{s}_r^T(n)\ \mathbf{w}_{RL}(n) \tag{1}$$

where $\mathbf{s}_r(n) = [s_r(n), s_r(n-1), \cdots, s_r(n-N_w+1)]^T$. The error signal $e(n)$ observed by the BC microphone in the left ear canal is defined as

$$e(n) = s'_r(n) + \left({}_{anti}\mathbf{s}'^T_r(n)\ \mathbf{h}_{LL}(n)\right) \tag{2}$$

where $_{anti}\mathbf{s}'^T_r(n) = [\ _{anti}s'_r(n),\ _{anti}s'_r(n-1), \cdots,\ _{anti}s'_r(n-N_{LL}+1)]^T$, and $\mathbf{h}_{LL}(n) = [h_{LL_0}, h_{LL_1}, \cdots, h_{LL_{N_{LL}-1}}]^T$ with a length of N_{LL}. The error signal $e(n)$ needs to be minimized by updating the weight vector of the CTC filter $\mathbf{w}_{RL}(n)$ using the FxLMS algorithm to achieve the cross-talk cancellation at the location of the BC ear microphone. The update algorithm is defined as follows

$$\mathbf{w}_{RL}(n+1) = \mathbf{w}_{RL}(n) - \mu\ \mathbf{x}(n)\ e(n) \tag{3}$$

where $\mathbf{x}(n) = [x(n), x(n-1), \cdots, x(n-N_{LL}+1)]^T$, and μ denotes the step size that determines the convergence of the algorithm. $\mathbf{x}(n)$ is the filtered version of $\mathbf{s}_r(n)$ defined as

$$x(n) = \mathbf{s}_r^T(n)\ \mathbf{c}(n) \tag{4}$$

where $\mathbf{c}(n) = \mathbf{h}_{LL}(n)$. Once the error signal $e(n)$ converges, the final CTC filter coefficients $\mathbf{w}_{RL}$ will be used to estimate cancellation signals in order to confirm whether or not the cross-talk sound can be canceled in the left ear canal and its attenuation can also be felt by participants.

2.2 Instrumentation

The measurement system setup using BC ear microphone is shown in Fig. 3(e). For observing BC sounds, the BC ear microphone – Vibraaudio EM-21 (13) – was used. Sound stimuli were presented by Cynaps BC transducers (14). A 3W stereo amplifier, PAM8403, with power bank was used to drive the BC transducers. Playback and recording were done through an external sound card, Sound Blaster Omni Surround 5.1, which has a USB computer connection; and recorded signals were stored in a laptop.

3. Measurements

3.1 Participants

Three self-reported normal hearing individuals participated in the experiments. The age range is between 23 and 29 years old. They were wearing a helmet headband with attached BC transducers as illustrated in Fig. 3(a); and had the BC ear microphone in their left ear canal. The participant sat in a quiet room.

3.2 Parameters

16-bit playback and recording were used at the sampling rate of 16 kHz. The length of the TSP signal used for estimating $H_{RL}(z)$ and $H_{LL}(z)$ was 16383 samples (equal to about 1 s); see examples of a recorded TSP signal in Figs. 4(a) and (b). Both estimated BC transfer functions – see Fig. 4(c) – were then used to estimate the CTC filter by computer simulation. Estimation of the CTC filter was made by minimizing the error signal $e(n)$ – see Fig. 4(d) – using the FxLMS algorithm in which a 15-s white noise was used as the input signal. The filter tap-length was considered to be 256 samples (or 16 ms) such as shown in Fig. 4(e). Then, a 14-s pink noise as test signal for all participants was filtered by their own CTC filter, and its result was considered to be the "individualized" cancellation signal; see Fig. 4(f) as an example.

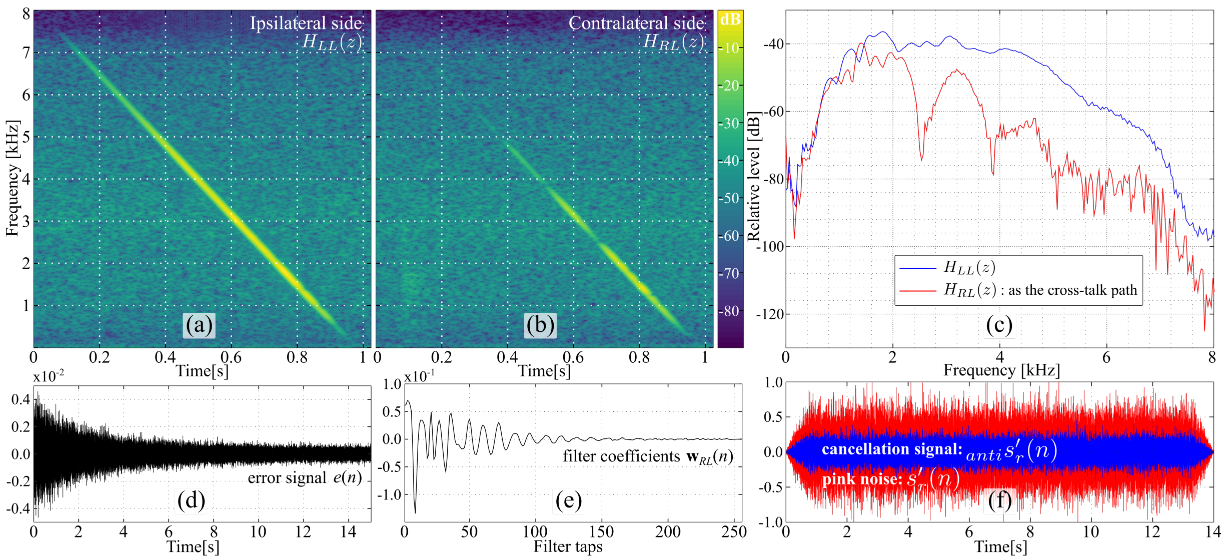

Figure 4 – Spectrograms of the TSP signal produced by (a) ipsilateral or (b) contralateral BC transducer and recorded by the BC ear microphone in the left ear of the participant 1. (c) Frequency responses of $H_{LL}(z)$ and $H_{RL}(z)$. (d) Error signal $e(n)$ converging in 15 seconds. (e) The estimated CTC filter coefficients $\mathbf{w}_{RL}(n)$. (f) Pink noise (for contralateral BC transducer) and its estimated cancellation signal (for ipsilateral one).

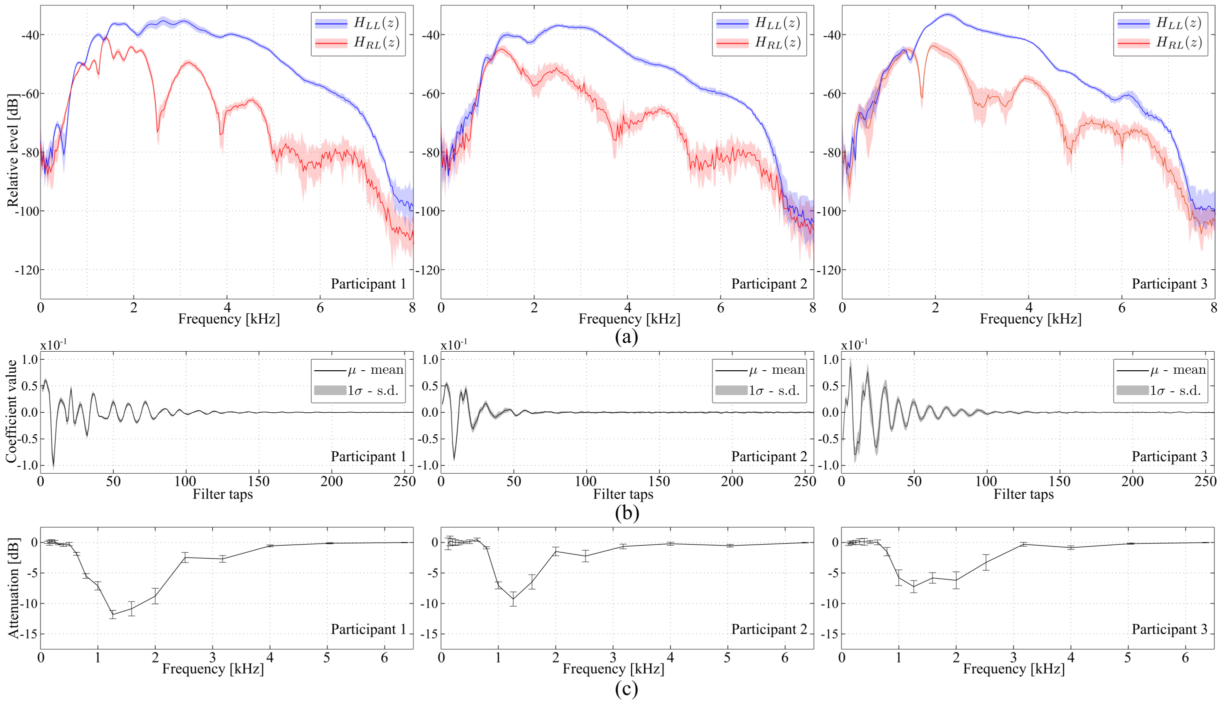

Figure 5 – (a) Frequency responses of ipsilateral BC transfer function $H_{LL}(z)$ (blue) and contralateral BC transfer function $H_{RL}(z)$ (red); means are shown as solid lines (—) while standard deviations are provided as band around the lines. (b) Impulse response of estimated 10 CTC filters shown as the mean (μ) with its standard deviation (σ). (c) Attenuation of the cross-talk sound observed by the BC ear microphone in the left ear canal of each participant.

3.3 Procedure

Single TSP signal (12) was presented alternately to the left and right BC transducers and simultaneously recorded by the BC ear microphone in the left ear canal of each participants in order to estimate $H_{LL}(z)$ and $H_{RL}(z)$. Once they were obtained, both were used to run an "offline" FxLMS algorithm (a simulation) with 15-s white noise as input to estimate the CTC filter which is unique for each participant. A 14-s pink pink noise as a test signal was then filtered by the CTC filter to estimate the "individualized" cancellation signal. After that, the pink noise and its cancellation signal were provided to the right and left BC transducers, respectively, and at the same time, the BC ear microphone picked up the interference of the two signals to confirm whether or not destructive interference occurs in the left ear canal. The picked-up signals with and without cross-talk cancellation were analyzed in each one-third octave band to obtain the attenuation of the cross-talk signal. All test sounds presented to the BC transducers were at a level corresponding to 60 dB sound pressure level (SPL) adjusted based on a loudness balancing technique (15). This measurement was repeated ten times in which the BC ear microphone was being reinserted into the left ear canal between repeated measures for each participant.

3.4 Result and Discussion

We hypothesized that the BC transfer function measured from the ipsilateral or contralateral BC transducer to the left ear canal might be used to achieve cross-talk cancellation at the cochlea in the inner ear. We thus estimated the BC transfer functions using a TSP sound presented alternately to the ipsilateral and contralateral BC transducers. Figure 5(a) shows the frequency responses of ipsilateral BC transfer function $H_{LL}(z)$ (blue) and contralateral BC transfer function $H_{RL}(z)$ (red) where means are shown as solid lines and standard deviations are provided as band around the lines. As expected, because the ipsilateral BC transducer is close to the BC ear microphone (see Fig. 3(b)), $H_{LL}(z)$ for all participants looks very stable, indicated by small changes in standard deviation across frequency. On the other hand, since the contralateral BC transducer is on the opposite side of the BC ear microphone (see Fig. 3(b)), $H_{RL}(z)$ at high frequencies were mostly attenuated by the participant head, and therefore, has large standard deviations. But, both $H_{LL}(z)$ and $H_{RL}(z)$ at low frequency below about 2 kHz were almost at the same level, which means that BC cross-talks in the three participant heads were quite dominant for the low-frequency range.

Using the transfer functions $H_{LL}(z)$ and $H_{RL}(z)$, see Fig. 5(a), estimated by TSP sounds, we then utilized a computer simulation to estimate the cross-talk compensation (CTC) filter based on the block diagram of the FxLMS algorithm shown in Fig. 3(d). Because we also intend to check how stable the estimated CTC filter is when the BC ear microphone was being reinserted between repeated measures, we therefore made ten times measurements (or ten times CTC filter estimations) to the three participants. Figure 5(b) presents the impulse response of estimated 10 CTC filters shown as the mean (μ) where the standard deviation (σ) is provided as gray band around the mean line. From the figure, we can see clearly that when the positions of the BC transducers remained fixed, small changes in the position of the BC ear microphone in the left ear canal have little influence on the impulse response of the CTC filter, indicated by small standard deviations for all participants. This means we may not need to frequently update the coefficient values of the CTC filter except when the positions of BC transducers change.

Estimation of the CTC filter was done ten times for each participant, and at the same time, using the estimated filter we also created ten "individualized" cancellation signals with a 14-s generated pink noise as the filter input signal. To perform unilateral cross-talk cancellation experiments, the pink noise was presented to the contralateral BC transducer (on the right-ear side) and recorded by the BC ear microphone in the left ear canal as the cross-talk. On the other hand, the cancellation signal was presented to the ipsilateral BC transducer (on the same side as the BC ear microphone). In this paper, the difference in level between the recorded BC sounds with and without cancellation signal is defined as the attenuation of the cross-talk sound. Figure 5(c) shows the attenuation result of the cross-talk sound observed by the BC ear microphone in the left ear canal of each participant. From the figure, it can be seen that BC cross-talks can be attenuated in the left ear canal of all participants. The attenuation mostly occurs in the low-frequency range (below about 2 kHz), but it looks ineffective for frequency below about 800 Hz. We suspect that this ineffectiveness might be related to the BC transducer's frequency response, which is 300 Hz to 1.9 kHz (14). Although the cancellation presented in Fig. 5(c) occurred in the ear canal, all participants reported they could feel a decrease in the perceived loudness due to the cross-talk cancellation. However, further experiment is still necessary in order to determine the effective frequency range that could achieve cancellation in both the ear canal and inner ear perceived by participants; such as using tone reception thresholds (TRT) (16).

4. CONCLUSION

This paper discusses the use of bone conduction (BC) ear microphone for estimating a cross-talk compensation (CTC) filter by which a cancellation signal with a similar amplitude but with inverted phase to a BC cross-talk sound can be estimated to achieve destructive interference in the target ear. We have demonstrated that it is potentially possible to achieve the cross-talk cancellation at the cochlea by canceling the cross-talk sound in the ear canal since it is the closest position to the cochlea. To estimate the CTC filter, we assumed that the vibration of the bony ear canal observed by BC ear microphone in the ear canal is relative to the vibration of the cochlea during BC stimulation. As a result, using a TSP sound, BC transfer functions can be estimated, and hence, the CTC filter can also be estimated using the filtered-x least mean square (FxLMS) algorithm. Experiment results show that low-frequency attenuation was observed by the BC ear microphone when a pink noise and its cancellation signal were presented through BC transducers. Also, all participants reported that they could feel a decrease in the perceived loudness. Because the BC ear microphone was placed in the ear canal, not exactly at the cochlea, further investigation is still necessary to determine the effective frequency range that could achieve cancellation in both the ear canal and inner ear perceived by participants; such as using tone reception thresholds (TRT) which remains as our future work.

ACKNOWLEDGEMENTS

The authors thank Kamelia Rosiana, Fajar Purnama, Muhammad Bagus Andra, Zheng Chenyang and Akito Kawamura for their helpful contributions during measurements.

REFERENCES

1. Hawley ML, Litovsky RY, Culling JF. The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. J Acoust Soc Am. 2004;115(2):833–843. doi:10.1121/1.1639908
2. Stenfelt S. Bilateral fitting of BAHAs and BAHA® fitted in unilateral deaf persons: Acoustical aspects. International Journal of Audiology. 2005;44(3):178–189. doi:10.1080/14992020500031561
3. Snik AF, Beynon AJ, van der Pouw CT, et al. Binaural application of the bone-anchored hearing aid. Ann Otol Rhinol Laryngol. 1998;107(3):187–193. doi:10.1177/000348949810700301

4. Mejia J, Nguyen CV, Yeend I, Loi T, Cowan R, Dillon H. Bone-conduction hearing aids: Into the (near) future. The Hearing Journal. 2015;68(8):30–32. doi:10.1097/01.HJ.0000470894.83693.cf
5. Nolan M, Lyon DJ. Transcranial attenuation in bone conduction audiometry. Journal of Laryngology & Otology. 1981;95(6):597–608. doi:10.1017/S0022215100091155
6. Stenfelt S. Transcranial attenuation of bone-conducted sound when stimulation is at the mastoid and at the bone conduction hearing aid position. Otology & Neurotology. 2012;33(2):105–114. doi:10.1097/MAO.0b013e31823e28ab
7. Glyde H, Buchholz JM, Dillon H, Cameron S, Hickson L. The importance of interaural time differences and level differences in spatial release from masking. The Journal of the Acoustical Society of America. 2013;134(2):EL147–EL152. doi:10.1121/1.4812441
8. Stenfelt S, Zeitooni M. Binaural hearing ability with mastoid applied bilateral bone conduction stimulation in normal hearing subjects. J Acoust Soc Am. 2013;134(1):481–493. doi:10.1121/1.4807637
9. Winkler A, Brand T, Kollmeier B. Crosstalk cancellation in audiology. Proc NAG/DAGA; 23-26 March 2009; Rotterdam, Netherlands 2009. url:pub.dega-akustik.de/NAG_DAGA_2009/data/articles/000458.pdf
10. Liao CL. Application of cross-talk cancellation to the improvement of binaural directional properties for individuals using bone anchored hearing aids (BAHA). Thesis. Chalmers University of Technology. 2010. url:http://publications.lib.chalmers.se/records/fulltext/126958.pdf
11. Reinfeldt S, Stenfelt S, Håkansson. Estimation of bone conduction skull transmission by hearing thresholds and ear-canal sound pressure. Hearing Research. 2013;299:19–28. doi:10.1016/j.heares.2013.01.023
12. Suzuki Y, Asano F, et al. An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses. J Acoust Soc Am. 1995;97(2):1119–1123. doi:10.1121/1.412224
13. TEMCO, *Vibraaudio EM21*. url:https://www.temco-j.co.jp/pr_em20/. (visited on 05/23/2019)
14. MAX Virtual LLC, *New and improved Cynaps transducers!*, url:https://www.maxvirtual.com/#!/New-and-improved-transducers/p/53940038. (visited on 05/23/2019)
15. Qin X, Usagawa T. Frequency characteristics of bone conduction actuators–Measurements of loudness and acceleration. Applied Acoustics. 2017;126:19-25. doi:10.1016/j.apacoust.2017.05.007
16. McLeod Robert. Psychoacoustic Measurements of Bone Conducted Sound. PhD diss., Cardiff University, 2016. http://orca-mwe.cf.ac.uk/100559/2/2017mcleodphd.pdf