HAMBURG 2016

Low Delay Wind Noise Cancellation for Binaural Hearing Aids

Nobuhiko HIRUMA[1]; Ryousuke KOUYAMA[2]; Hidetoshi NAKASHIMA[3]; Yoh-ichi FUJISAKA[4]

[1, 4] Rion Co., Ltd, Japan

[2, 3] National Institute of Technology, Kumamoto College, Japan

ABSTRACT

Wind noise annoys hearing aid users, and it is undesirable to attach a windscreen to a hearing aid microphone for cosmetic reasons. Accordingly, some hearing aid devices reduce the low frequency components of input signals by using high-pass filters to suppress wind noise. Although wind noise can be attenuated by using this approach, the perceived binaural information of the desired signal is also simultaneously degraded; therefore, localization cues are partially lost. We previously proposed a FFT-based binaural wind noise cancellation algorithm to preserve the cues. This algorithm required a 32 ms frame length in order to maintain a high-frequency resolution. However, it is known that the tolerable delay time for mild hearing loss or better should be below about 5 ms in the high frequency region. In this study, we propose a low delay binaural wind noise cancellation algorithm using a frequency warping filter. The delay time caused by this algorithm is shorter than the tolerable one. The objective evaluation results, SNR and PESQ scores, were improved while maintaining a low delay time. As a result of subjective experiments, the proposed method produces almost the same score as our conventional FFT-based method for the directionality of output signals.

Keywords: Hearing aids, Wind noise, Binaural processing, Low delay, Frequency warping filter, HRTF
I-INCE Classification of Subjects Number(s): 01.4

1. INTRODUCTION

With an increasing number of active hearing aid users in recent years, opportunities for speech communication have increased even in outdoor situations. Therefore, the development of optimal hearing aid processing appropriate to various environments is desired. According to a survey by MarkeTrak (1), wind noise still remains as a major cause of dissatisfaction for hearing aid users in listening environments. Wind noise in hearing aids is generated by the air-flow around the head, by the flow over and around the pinna, and by the flow past the tragus and helix(2,3). Generally, a windproof is utilized to reduce the noise by attaching it to the microphone. However, for hearing aids, wind noise is reduced by signal processing for cosmetic reasons. Thus, several algorithms have been proposed using digital signal processing to reduce the discomfort of wind noise (4–7). In single channel wind noise suppression in hearing aids, high-pass filter processing has been utilized to reduce wind noise since the main components of wind noise exist in the low frequency region. Although wind noise can be attenuated by using a high-pass filter, the desired signal is also degraded simultaneously. Furthermore, when using hearing aids on both the left and right ears that work independently, the perception of the arrival direction and distance of the desired signal will be uncertain.

Therefore, we previously proposed a binaural wind noise cancellation algorithm using the characteristics of head related transfer functions (HRTFs) in order to construct a future binaural type of hearing assistant system (8,9) to avoid the above issue. In our previous study, it was shown that the spatial information (directionality and externality) was maintained, compared with the conventional high-pass filter method by performing psychoacoustic experiments.

Incidentally, open-fit styles (10) have been applied to reduce occlusion effect, so that especially mild hearing loss often uses it. However, a processed sound of hearing aids is heard as echo if the

[1] n-hiruma@rion.co.jp
[2] ae15kouyama@g.kumamoto-nct.ac.jp
[3] nakashi@kumamoto-nct.ac.jp
[4] y-fujisaka@rion.co.jp

sound has some long delay, since a direct sound is also heard through a large vent in hearing aids for open-fit styles.

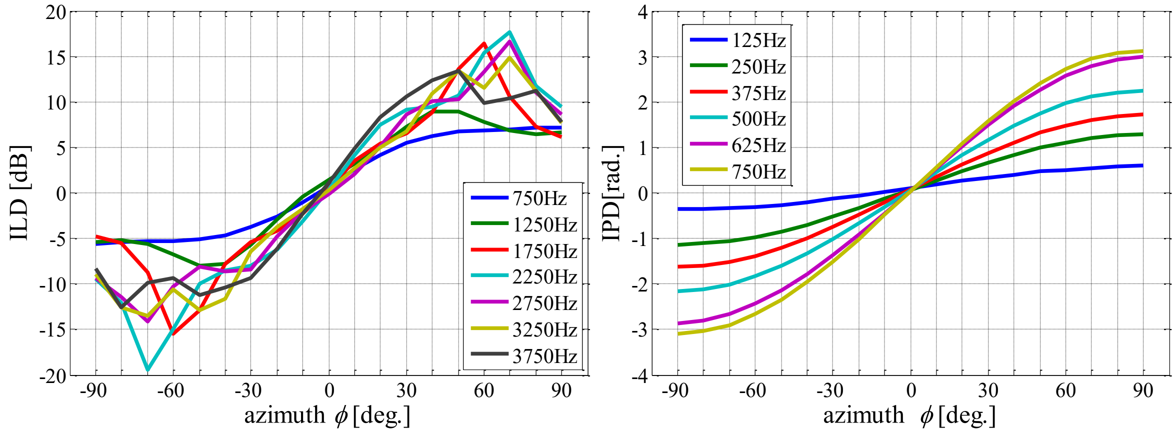
Our proposed method has a required frame processing of 32 ms to maintain a high frequency resolution in the low frequency region using a Fourier transform. However, the delay time with this method is intolerable for hearing aid users (10). Therefore, the processing delay must be as short as possible.

In this paper, we propose a binaural wind noise cancellation algorithm using a frequency warped filter as a substitute for FFT. With this approach, it is possible to keep the delay time tolerable while maintaining a high frequency resolution in the low-frequency region.

2. BINAURAL WIND NOISE CANCELLATION SYSTEM

2.1 Characteristics of HRTFs (IPDs and ILDs)

The characteristics of the frequency response from a sound source to both ears can be described as HRTFs. Figure 1 shows (a) interaural phase differences (IPDs) and (b) interaural level differences (ILDs) calculated by using HRTFs (11), where the horizontal axis shows the direction of arrival (DOA) of a sound source. It is apparent that IPD and ILD vary depending on the DOA and frequency of the sound. These characteristics are stored to memory as a database (12) for wind noise detection and cancellation. Thus, DOA can be estimated by comparison between the database and the characteristics that are derived from the observed binaural signals. The database of each frequency band is $\theta_{DB}(k,\varphi)$ for IPDs and $\xi_{DB}(k,\varphi)$ for ILDs, where φ and k represent the azimuth angle of a sound source and frequency index, respectively.



(a) (b)

Figure 1 –Database of proposed system. (a) IPDs and (b) ILDs (13).

2.2 System Block Diagram

Figure 2 shows a block diagram of the proposed system. In our system, the use of binaural hearing aids is assumed; therefore, binaural input signals are utilized for signal processing. Binaural hearing aids have two microphones put on each ear. The observed signals on both the left and right sides, $l(n)$ and $r(n)$, are defined as

$$l(n) = s(n) * h_{l,\phi}(n) + wn_l(n), \tag{1}$$

$$r(n) = s(n) * h_{r,\phi}(n) + wn_r(n), \tag{2}$$

where $h_{l,\varphi}(n)$ and $h_{r,\varphi}(n)$ denote HRTFs from the direction φ, and $wn_l(n)$ and $wn_r(n)$ denote the wind noise signals. The asterisk symbol [*] represents the convolution operation. We assume that $wn_l(n)$ and $wn_r(n)$ are uncorrelated and that the IPD and ILD of the noise will be random in the time domain.

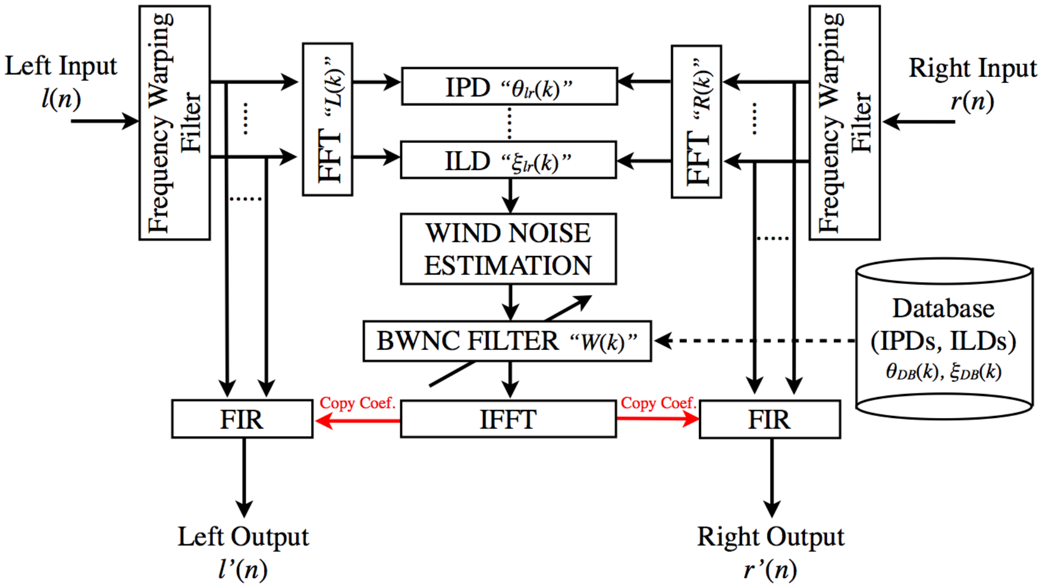


Figure 2 – Block diagram of proposed system

2.3 Frequency Warping Filter

In our previous wind noise cancellation algorithm (8,9), a 32-ms frame length was required for FFT analysis with high resolution. In the case of 16-kHz sampling rates, it needs 512 samples for frequency analysis in order to satisfy the above frequency resolution Δf of 31.25 Hz, but the processing delay time, which means the group delay of the system, is 32 ms.

Stone and Moore (13) showed that the tolerable hearing aid delay time must be less than about 5 ms. Therefore, the group delay of our system should also be shorter than the tolerable processing time, defined as tolerable delay, for the analysis and the synthesis of audio signals. In our FFT based algorithm, it is necessary that the frame length must be shorter than the tolerable group delay. However, when the number of signal samples is reduced in FFT analysis, the frequency resolution is degraded. As a result, effective wind noise suppression cannot be possible because the analysis frequency bandwidth is larger than the fine harmonic structure of speech.

We propose a binaural wind noise cancellation algorithm using a frequency warping filter (FWF) to preserve a low group delay and the intended frequency resolution. A block diagram of analysis and synthesis using FWF is shown in Figure 3, where M represents the number of all-pass filters (APFs). The signal vector, $p_l(n,0)$ to $p_l(n,M)$ in Figure 3, is a warped signal, of which the frequency components are warped from low to high frequency region with an appropriate ratio decided by the coefficient γ in equation (3) (14,15).

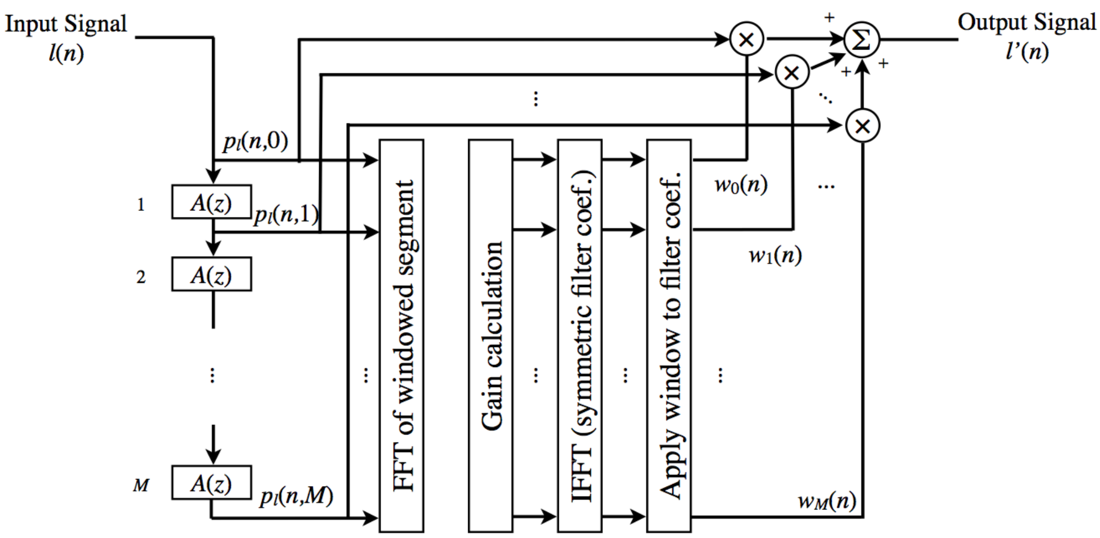


Figure 3 – Block diagram of portion of wind noise cancellation system using frequency warping for left channel

The all-pass filter is given by

$$A(z) = (z^{-1} - \gamma)/(1 - \gamma z^{-1}), \tag{3}$$

where γ is the warping parameter.

The frequency warped sequence of the left channel is obtained by

$$p_l(n,0) = l(n), \tag{4}$$

$$p_l(n,m) = \gamma\{p_l(n-1,m-1) - p_l(n,m-1) + p_l(n-1,m-1)\quad , where\ m = 1 \sim M, \tag{5}$$

and that of the right channel is obtained with the same calculation. The frequency warped output of the kth all-pass stage is given by $\{p_l(n,m), p_r(n,m)\}$ in both the left and right channels.

Figure 4 shows an example of the group delay characteristics of FWF compared with the tolerable delay. The solid line represents the tolerable group delay time, derived from reference (13), as a function of frequency. The broken line represents the FWF group delay time in the case that γ equals 0.8. We can see that the FWF group delay time is shorter than the tolerable group delay time.

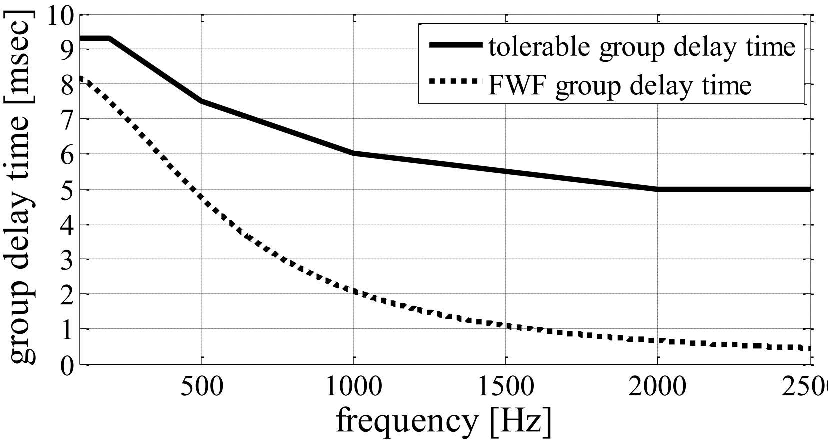


Figure 4 – Group delay of proposed system (γ = 0.8) and tolerable delay time (10)

2.4 IPD and ILD Calculation of Input Signal

The IPD of the input signals, $\theta_{lr}(k)$, is calculated by using the cross spectrum of each input signal, $C_{lr}(k)$, represented as

$$\theta_{lr}(k) = \tan^{-1}\left[\mathrm{Im}\{C_{lr}(k)\}/\mathrm{Re}\{C_{ll}(k)\}\right], \tag{6}$$

where

$$C_{lr}(k) = L(k)R(k)^*. \tag{7}$$

Note that the operator [*] represents a complex conjugate. The same as IPD calculation, the ILD of the input signals, $\xi_{lr}(k)$, is calculated as

$$\xi_{lr}(k) = 20\log\left|C_{lr}(k)/C_{ll}(k)\right|, \tag{8}$$

where $C_{ll}(k)$ denotes the power spectrum of the input signal $l(n)$.

2.5 Wind Noise Detection Algorithm

Wind noise detection is performed by comparing databases $\theta_{DB}(k)$ and $\xi_{DB}(k)$. $\Delta\theta(k)$ and $\Delta\xi(k)$ are the absolute difference between them, defined as

$$\Delta\theta(k) = \left|\theta_{DB}(k,\psi) - \theta_{lr}(k)\right|, \tag{9}$$

$$\Delta\xi(k) = \left|\xi_{DB}(k,\psi) - \xi_{lr}(k)\right|. \tag{10}$$

Here, ψ denotes a speech signal direction. Normally, this direction ψ is in front of a hearing aid listener. However, listeners can select the direction freely for speech enhancement in our system.

The minimum value of $\Delta\theta(k)$ and $\Delta\xi(k)$ is 0, whereas the maximum value varies depending on the frequency, so normalized differences are introduced as follows.

$$D_k = \begin{cases} \Delta\theta(k)P(k)^{-1} & \text{(for low frequency region)} \\ \Delta\xi(k)L(k)^{-1} & \text{(for high frequency region)} \end{cases}, \tag{11}$$

where $P(k)$ is the maximum difference of $\theta_{DB}(k,\phi)$ when ϕ varies, defined as

$$P(k) = \max\left[\theta_{DB}(k,\phi)\right] - \min\left[\theta_{DB}(k,\phi)\right]. \tag{12}$$

$L(k)$ is also defined as

$$L(k) = \max\left[\xi_{DB}(k,\phi)\right] - \min\left[\xi_{DB}(k,\phi)\right]. \tag{13}$$

We assume that wind noise exists in the frequency region k if $D_k(n)>Tw$, where Tw is a predetermined wind detection threshold. This comparison procedure is performed for whole frequency bands, and $c(n)$ is the wind counter for determining the amount of suppression regarding wind noise, given by

$$c(n) = \sum_{k=0}^{K} u_k(n), \tag{14}$$

where $u_k(n)$ represents the detection flag of wind noise related to each frequency band. The flag is described by

$$u_k(n) = \begin{cases} 1 & (if\ D_k(n) > Tw) \\ 0 & (otherwise) \end{cases}. \tag{15}$$

2.6 Binaural Wind Noise Cancellation Filter

The instantaneous binaural wind noise cancellation (BWNC) filter $W^i{}_k(n)$ is defined as follows.

$$W^i{}_k(n) = \left[1 + \exp(\beta_k(n)D_k(n))\right]^{-1}, \tag{16}$$

where $\beta_k(n)$ is a gain control parameter that varies depending on the wind counter $c(n)$ and is appropriately determined.

$$\beta_k(n) \propto c(n). \tag{17}$$

Thus, $c(n)$ and $\beta_k(n)$ will be large under windy conditions, and $W_k(n)$ approaches 0. On the contrary, $W_k(n)$ approaches 1 under calm conditions. To avoid sudden changes in the filter gain, the actual BWNC filter $W_k(n)$ is defined as

$$W_k(n) \Leftarrow \lambda \cdot W_k(n-1) + (1-\lambda) \cdot W^i{}_k(n), \tag{18}$$

where λ is a forgetting factor. The impulse response form of the time domain is obtained by the inverse Fourier transform of the wind noise suppression filter coefficient,

$$w_k(n) = \mathrm{IFFT}\left[W_k(n)\right]. \tag{19}$$

The processed output signals, $l'(n)$ and $r'(n)$, are obtained by convoluting both warped signals and the noise suppression filter $w_k(n)$.

$$l'(n) = \sum_{m,k=0}^{M} w_k(n) p_l(n,m), \tag{20}$$

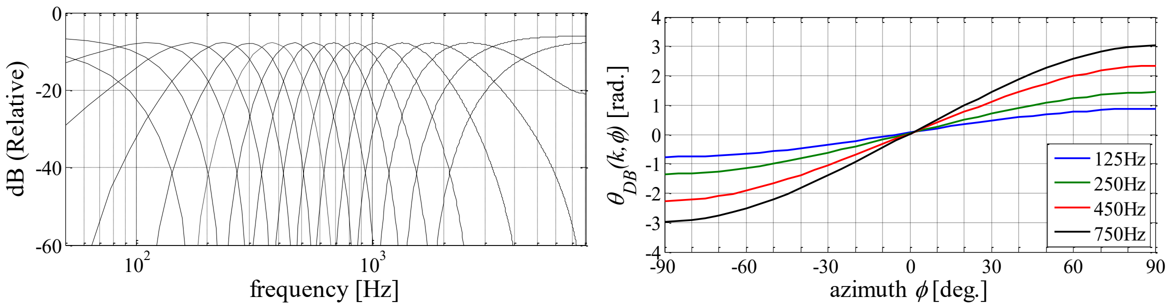
$$r'(n) = \sum_{m,k=0}^{M} w_k(n) p_r(n,m). \tag{21}$$

3. NUMERICAL SIMULATION

3.1 Numerical Configuration

A numerical test was performed to confirm the ability of the wind noise detection and the control of the filter gain $w_k(n)$. In this simulation, the sampling frequency was 16 kHz, and 32-point FFT was utilized to analyze warped signals. The warping filter coefficient γ in equation (3) was set to 0.8 in order to analyze the wind noise components for the high resolution in low frequency. The filter bank derived from FWF in this simulation is shown in Figure 5 (a). The frequency characteristics of the filter bank between 100 to 1000 Hz are similar to that of 1/3 octave filter bank.

Figure 5 (b) shows the IPD database of the warped HRTFs for each angle, which were calculated via FWF. The frequency bands under the 750 Hz center frequency were utilized for IPD calculation, so that wind noise estimation is mainly performed by using IPD.



(a) (b)

Figure 5 – Characteristics of frequency warp filter-bank (a) and warped domain IPDs (b)

Figure 6 shows an image of the relative locations among a listener, a desired speech source, and wind noise in our numerical simulation. The wind noise flow is located at 0 degrees, and the desired speech is located at -60 degrees. HRTFs donated by the MIT Media Lab (12) were utilized to locate the speech source at -60 degrees and the distance between the sound source and dummy head. In addition, wind noise sound that was pre-recorded by using the binaural microphones of the dummy head was used in a condition where wind velocity was approximately 4m/s. There were two silent intervals in the wind noise test sound. The intervals were set to time periods of 2 - 3 and 5 - 6 sec. The SNR of the input signal was set to -10 dB in front of the hearing aid listener. Here, the speech signal direction ψ was set to -60 degrees, assuming that the direction of the desired signal is known.

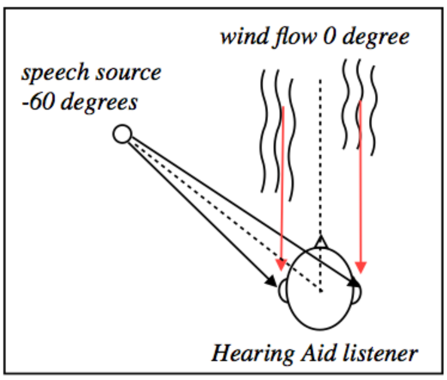


Figure 6 – Relative location among listener, desired speech source, and wind noise. Desired speech source was located at -60 degrees, while wind noise flowed from 0 degrees. SNR was set to 0 dB.

3.2 Result of Simulation

Figure 7 shows the waveforms and internal parameters obtained from the simulation in 3.1. Each figure shows the (a) left channel input signal and (b) right channel input signal. For the left channel, the original speech signal (represent by the black line) and the speech in wind signal (represented by the red line) are shown in (a), and the processed output signal is in (b). In the case of the right channel, the above mentioned waveforms are shown in Figures 7 (c) and (d), which are similar to the left channel. We can see that the wind noise in the processed output signal gradually decreased as described in (b) and (d) when wind noise existed.

Figure 7 (e) shows the wind counter value as a function of time. This value varies corresponding to wind noise strength. Figure 7 (f) is the variation of the filter gain $W_k(n)$ for the 3, 9, and 13th bands. It is clearly seen that the gain controller $W_k(n)$ for all frequencies was close to 1 when wind noise disappeared, and it controlled the gain to reduce the noise gradually when it appeared. These results reveal that the gain controller worked precisely according to wind noise conditions.

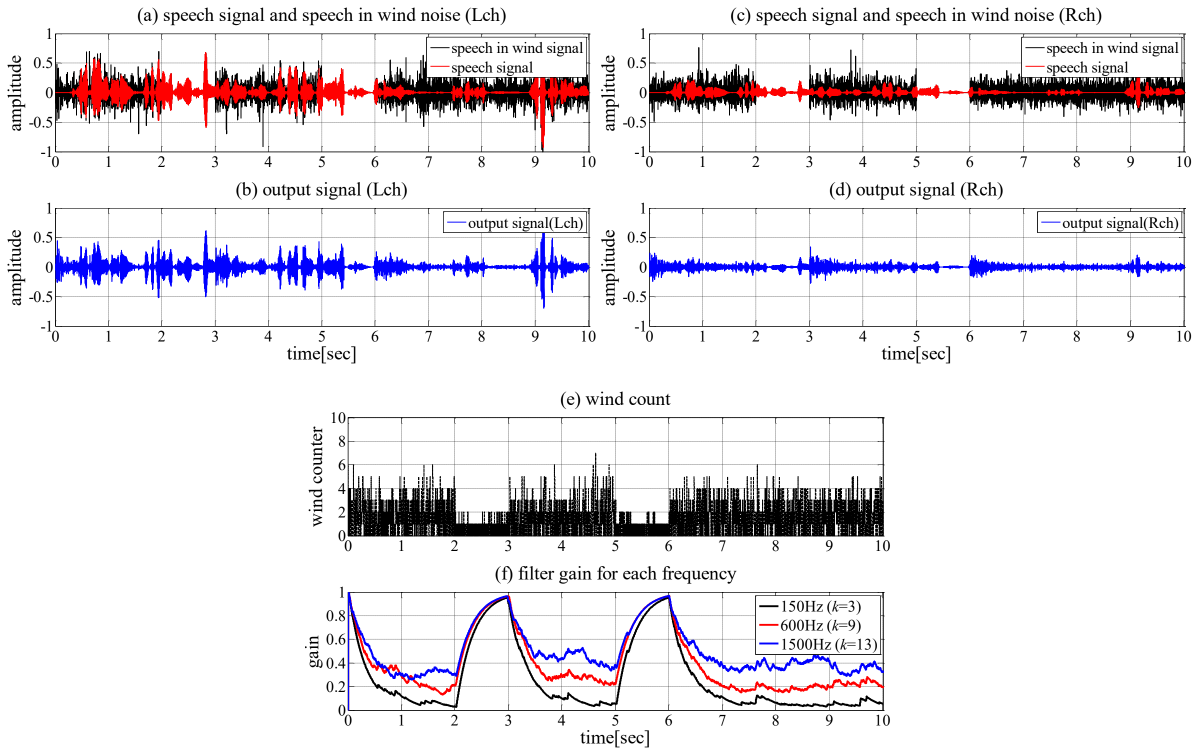


Figure 7 – Simulation results. (a) Input speech signal and speech in wind noise at Lch, (b) noise cancelled output signal at Lch, (c) input speech signal and speech in wind noise at Rch, (d) noise cancelled output signal at Rch, (e) wind count c(t), and (f) filter gain for each frequency band. SNR was set to -10 dB.

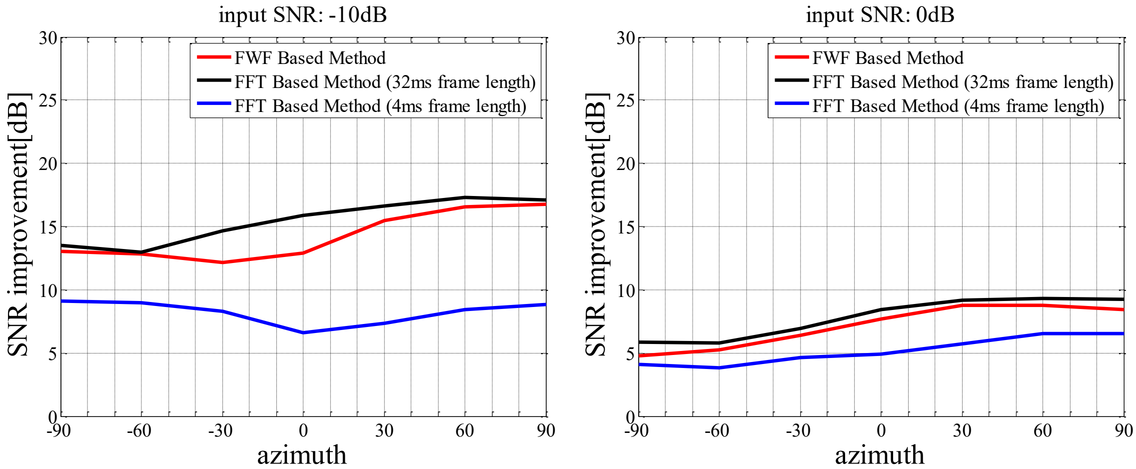
4. EVALUATIONS

4.1 Objective Evaluations

Two kinds of objective evaluations were performed to verify the effectiveness of the proposed method. The first one was SNR improvement, and the second one was an evaluation of the sound quality using the PESQ measure. The sound level of wind noise was set to SNRs of 0 dB and -10 dB for the left channel when the speech signal came from 0 degrees. The wind noise level was fixed, and the direction of the speech signal varied from -90 to +90 degrees in 30-degree steps in this simulation. To evaluate the SNR improvement and PESQ score, there was a comparison between the previous FFT method with a 32-ms frame length (FFT32 method), the FFT method with a 4-ms frame length (FFT4 method), and the proposed method (FWF method).

(I) SNR Improvement

Figure 8 shows the result of SNR improvement, which was obtained from the difference between the SNR of the input signal and its output signal. Figures 8 (a) and (b) represent the SNR improvement as a function of the degree for the speech arrival direction for the SNRs of -10 dB and 0 dB. The results show that the SNR of the signal after processing with our previous FFT32 method was significantly improved compared with the input signal, whereas it was slightly improved with the FFT4 method. The figure demonstrates that it is also indicated good performance in the case of using the FWF method since the curve was close to the FFT32 result.

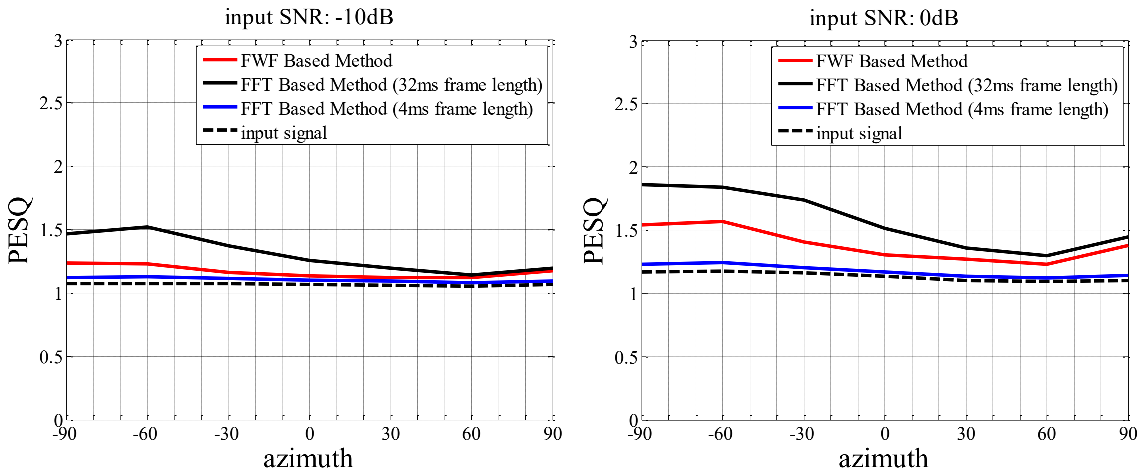


(a) (b)

Figure 8 – Result of SNR improvement. (a) Input SNR was -10 dB. (b) Input SNR was 0 dB.

4.1.2 PESQ Scores

The measurement results for PESQ are shown in Figures 9 (a) and (b), which represent the PESQ score as a function of the degree for the speech arrival direction for the SNRs of -10 dB and 0 dB. There was enough improvement in each direction in case of FWF compared with FFT4, while it was represented by the thinnest margin in the low SNR environment.



(a) (b)

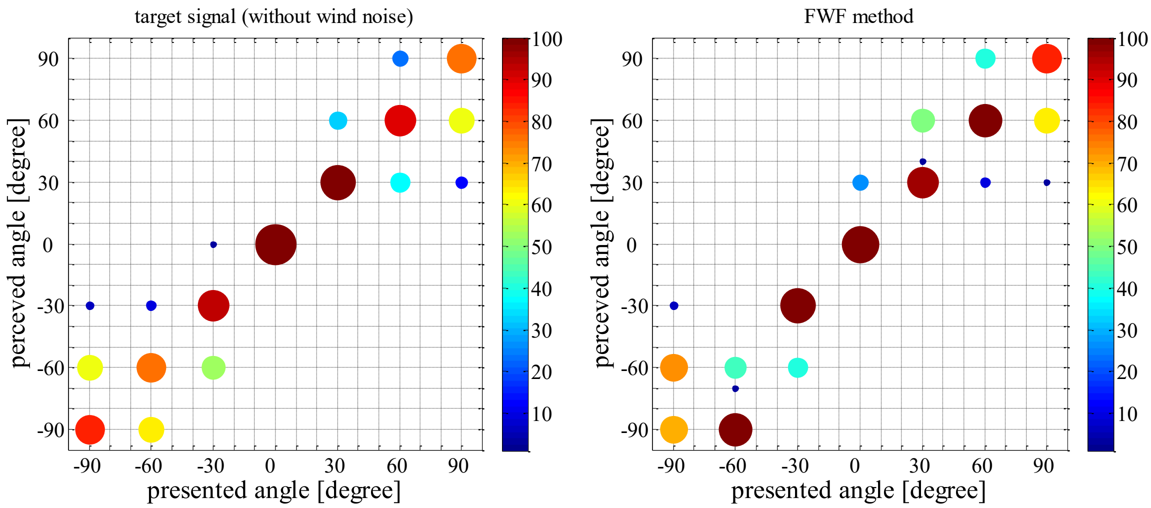
Figure 9 – Results of PESQ score for each method. (a) Input SNR: -10 dB. (b) Input SNR: 0 dB.

4.2 Subjective Evaluations

To evaluate the preservation of binaural information, we performed two types of subjective evaluation: “directionality” and “externality” tests. The directionality test is to examine the influence on the perception of the arrival direction of sound. The externality test compares the proposed FWF method with the conventional high-pass method in order to evaluate whether spatial information is retained. These tests were carried out in a way selected on in a way selected on the MATLAB GUI to present pre-processed sound from headphones (Sennheiser HD-600) in an anechoic chamber.

(I) Directionality Test

As stimulus, we used two types of the word. The first one is a 4-mora Japanese words to which seven kinds of directional characteristics were added by using HRTFs. The second one is the words in wind noise which were processed in the FWF proposed method. The seven directions were from -90 degrees to +90 degrees in 30-degree steps. In this experiment, three subjects from 19 - 21 years old were recruited. The stimuli were presented at random for each subject. The subjects were asked to answer the perceived direction of the words. Although the stimuli were presented in 30-degree steps, the subjects selected their perceived direction on a scale of 10-degree steps. Figure 10 shows the results of the directionality test for the (a) target signal and (b) FWF processed signal. It indicates that the perceived direction of unprocessed target signals was not affected. It also indicates that the signals output by the FWF process preserved the directional cues.



(a) (b)

Figure 10 – Result of directionality test, (a) target signal and (b) proposed method. Here, horizontal and vertical axes indicate presentation and perceived angles, respectively.

(II) Externality Test

The output signals obtained from FWF and conventional high-pass filter methods were used as stimuli, when the input signals of Japanese sentences have each seven directional characteristics with wind noise. In this experiment, five subjects from 19 - 21 years old were recruited. The subjects were presented randomly with two types of stimuli that were derived from the processed signals of these methods. The question to the subjects was, “which one is farther from your position?”. Then, the subjects had to select an answer from “first one,” “second one,” or “even.” The subjects had to answer “even” if he did not feel any difference. The experiments were performed for five subjects, and four kinds of sentences were used, so the number of results was twenty per angle. Figure 11 shows the result of the externality test. The horizontal and vertical axes show the angle of the stimuli and the percentage of the selected externality sound, respectively. The results of all directions imply that the proposed method was superior from an externality standpoint.

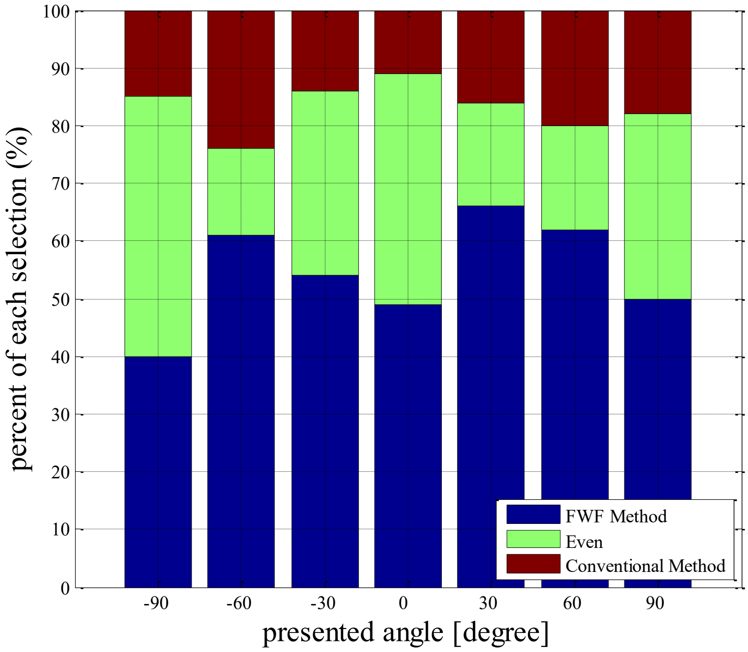


Figure 11 – Result of externality test. Horizontal and vertical axes show presentation angle and percentage of each selection.

5. CONCLUTIONS

We proposed a low delay wind noise cancellation algorithm for binaural hearing aids. The conflicting problem of delay time and frequency resolution was solved by using a frequency warping filter. The numerical simulation result showed that wind noise can be cancelled with accuracy by using a BWNC filter. The objective evaluation results reveal that the SNR and sound quality were improved while maintaining a low group delay time. As a result of subjective evaluation, a directionality test indicated that the BWNC filter did not influence the perceived direction. Furthermore, the perceived spatial impression was slightly superior to the conventional high-pass filter method from an externality standpoint.

The results of our research show that the processing delay time of the proposed BWNC system can be shorter than the tolerable group delay time using a frequency warping filter while maintaining the localization cue of the desired signal.

REFERENCES

1. S. Kochkin. MarkeTrak VIII: Consumer satisfaction with hearing aids is slowly increasing. Hearing Journal; 2010;63(1):19–32.
2. H. Dillon, Roe I, Katch R. The sources of wind noise in hearing aids. Lake Tahoe, USA: IHCON; August 2000.
3. Kates, J. M. Digital hearing aids. San Diego: Plural Publishing; 2008. p. 147-173
4. G. W. Elko. Reducing noise in audio systems. US Patent 7 171 008, 2007.
5. Kates, J. M. Hearing aid with suppression of wind noise. US Patent 8 019 103, 2007.
6. K. S. Petersen, G. Bogason, U. Kjems, T. B. Elmedyb. Device and method for detecting wind noise. US Patent 7 340 068, 2008.
7. J. A. Zakis, C. M. Tan. Robust wind noise detection. Florence, Italy: IEEE ICASSP; 2014. p. 3683–7.
8. Hiruma N, Nakashima H, Fujisaka Y. Binaural wind noise cancellation algorithm in hearing aids using HRTF cues. Lake Tahoe, USA: IHCON; August 2014.
9. Nakashima H, Kouyama R, Hiruma N, Fujisaka Y. Binaural wind noise detection, cancellation and its evaluation for hearing aids. IEEE IECON; 9-12 November 2015. Yokohama, Japan: p. 004896–9.
10. H. Dillon. Hearing Aids. Turramurra, Australia: Boomerang Press; 2001. p. 130-132
11. Gardner, B., Martin, K. HRTF measurements of a dummy-head KEMAR dummy-head microphone. 18 July 2000; Available from: http://sound.media.mit.edu/resources/KEMAR.html
12. Nakashima H, Chisaki Y, Usagawa T, Ebata M. Frequency domain binaural model based on interaural phase and level diffences. Acoust. Sci. & Tech. 2003;24(4): p. 172–8.

13. M. A. Stone, B. C. J. Moore, K. Meisenbacher, R. P. Derleth. Tolerable hearing aid delays. V. estimation of limits for open canal fittings. Ear Hear. 2008;29(4):601–17.
14. A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, J. Houpaniemi. Frequency-warped signal procesing for audio applications. J Audio Eng Soc. 2000;48(11):1011–31.
15. Kates, J. M., Arehart, K. H. Multichannel dynamic-range compression using digital frequency warping. EURASIP J Appl Signal Process. 2005;18:3003–14.