Study on Wind Noise Reduction using Neural Network for Hearing Aids

ARIMOTO Katsuya^a, NAKATOH Yoshihisa^{a*}

^aKyushu Institute of Technology,
1-1 Sensui-cho, Tobata-ku, Kitakyushu-shi, Fukuoka 804-8550, Japan
*nakatoh@ecs.kyutech.ac.jp

Abstract

Currently, there are 3.4 million hearing aid wearers in Japan, and it is said that one-ear wearers are about 60% of them. In addition, wind noise is one of the reasons for interfering with listening when a hearing aid is used outdoors. In this study, in order to eliminate wind noise in a one-ear hearing aid, we proposed a method to estimate wind noise using GAN, and to eliminate wind noise every moment based on the spectral subtraction method. As a result of evaluating the effectiveness, we showed that by removing wind noise in In-the-Ear hearing aid and Behind-the-Ear hearing aid, speech with 70% or more correlation with ideal sound can be generated.

Keywords: wind noise, hearing aids, Neural Network.

structure of GAN is shown in Fig. 1. GAN is composed of a pair of neural networks called Generator (G) and Discriminator (D). Generator has Autoencoder structure, and performs convolution to dimensionally compress the input signal. Generator simulates real samples in the training set and generates pseudo samples from the input signal. Therefore, the purpose of Generator is to improve the similarity between real sample and pseudo sample. Discriminator classifies the input into real samples in the training set and pseudo samples generated by Generator. Therefore, the purpose of Discriminator is to improve classification accuracy of real sample and pseudo sample. In this manner, by performing learning with Generator and Discriminator in competition, it is possible to generate a signal with high accuracy similar to the real samples in the training set.

1. Introduction

Currently, it is said that there are about 3.4 million hearing aid wearers in Japan, and one-ear wearers are about 60% of them. Also, wind noise is one of the reasons to disturb hearing when hearing aids are used outdoors. Therefore, in this study, we aim to eliminate wind noise in one ear hearing aid. we propose the method to estimate wind noise using neural network and remove wind noise every moment based on spectral subtraction method, and evaluate the effective of the method.

2. Generative Adversarial Networks

In this study, wind noise was learned by using Generative Adversarial Networks (GAN) which is one type of machine learning method in the wind noise estimation part. GAN is a model that simulates real data distribution and generates new samples related to training sets. The

Fig.1 The structure of GAN

3. Outline of the proposed method

The flowchart of the proposed method is shown in Fig.2. The proposed method is divided into the wind noise estimation part and the wind noise reduction part.

3.1 The noise estimation part

In the proposed method, wind noise estimation is performed using GAN. Wind noise for training data and test data of GAN was recorded under the conditions as shown in Table 1 and Fig.3. The wind noise of the test data was recorded with an omnidirectional In-the-Ear hearing aid (ITE) and an omnidirectional Behind-the-Ear hearing aid (BTE) on the left ear of the dummy head, with the front angle set to 0° and counterclockwise by 180° every 15° .

Fig.2 The flowchart of the proposed method

	Training data	Test data
Condition	Soundproof room	
Wind Speed	2~5 m/s	4 m/s
# of data	11572	13
Microphone	Roland CS-10EL	ITE BTE
# of Channel	1	2

Fig.3 Recording Condition

Fig.4 Input data for GAN Training

In this research, training of GAN was carried out using input data like Fig.4. “Wind Noise” used the above Training data. “Speech + Wind Noise” used training data with “Noisy speech database” of train_data added. Generator and Discriminator were alternately subjected to $100 * 420$ epoch learning respectively. In this study, we used the same model as the previous study.

3.2 The noise reduction part

In this study, spectral subtraction method (the SS method) is used for wind noise elimination. The SS method is one of speech enhancement methods that subtracts the estimated noise spectrum in the frequency domain. Two methods, Fast Fourier Transform (FFT) and Discrete Cosine Transform (DCT), exist as typical methods for converting to the frequency domain, but in this study, in order to investigate which is suitable for wind noise cancellation, Study was carried out. Wind noise was applied

as noise to 10 kinds of test sounds, and the noise was converted by FFT and DCT and removed by the SS method. Next, the correlation between the output waveform and the original input waveform was subjectively evaluated objectively by STOI (Short-Time Objective Intelligibility measure). Table 2 shows the average value of the evaluation values of 10 kind of test sound. From Table 2, since the evaluation value of DCT is higher than the evaluation of FFT, DCT is used as the conversion method in the proposed method.

3.3 Wind noise reduction using estimated wind noise

In this study, estimated noise was generated from the input signal by Generator. In the SS method, the number of frames was 256 and the frame shift was 50% overlapped.

4. Evaluation of the proposed method

4.1 Evaluation method

As a previous study (SEGAN), it has been reported that efficient noise cancellation can be performed by letting GAN train sound. we compared the sound emphasis with this as a comparison object and removed the wind noise and objectively evaluated the correlation between the output speech and the ideal sound by STOI.

4.2 Experiment data

For experiments we used “Noisy speech database” of test_data with wind noise on hearing aid. 10 kinds of male and female voice each were used for evaluation speech data.

4.3 Experimental results

Fig.5,6 shows the correlation between the sound pressure level of wind noise and the wind angle in ITE and BTE. The average results of the evaluation values of the previous method and the proposed method are compared and shown in Fig.7~14. As a result, it was shown that the effectiveness of the proposed method is high in any microphone, wind angle (sound pressure level of wind noise), voice sex.

Table 2 STOI value between Clean Speech

	Noisy	After SS with FFT	After SS with DCT
Ave.	75.1%	88.6%	98.8%

Fig.5 SPL of Wind noise recorded by ITE

Fig.6 SPL of Wind noise recorded by BTE

Fig.7 Result of ITE_mic.1 and Male voice

Fig.8 Result of ITE_mic.1 and Female voice

Fig.9 Result of ITE_mic.2 and Male voice

Fig.10 Result of ITE_mic.2 and Female voice

Fig.11 Result of BTE_mic.1 and Male voice

Fig.12 Result of BTE_mic.1 and Female voice

Fig.13 Result of BTE_mic.2 and Male voice

Fig.14 Result of BTE_mic.2 and Female voice

5. Conclusions

Prior research generates and outputs speech signals, it is greatly affected by feature quantities such as gender, language, fundamental frequency, and sound pressure of speech in training data of GAN. However, in the proposed method, since the specific noise to be removed is used as training data of the GAN, it is considered that it is hardly affected by the characteristics of the voice signal. The results also showed that by removing wind noise by the proposed method, it is possible to generate speech with ideal voice and correlation of 70% or more. From the above, wind noise cancellation by the proposed method is hardly affected by wind angle (sound pressure level of wind noise), which is considered to be an efficient method with little effect on sound signal.

References

- (1) "JapanTrak2015", 17, 2015
- (2) Phonak Insight : "Speech in Wind", 2012
- (3) Valentini-Botinhao, Cassia. Noisy speech database for training speech enhancement algorithms and TTS models, [dataset]. University of Edinburgh. School of Informatics. Centre for Speech Technology Research (CSTR), 2016
- (4) Cees H. Taal, Richard C. Hendriks, Richard Heusdens, Jesper Jensen, "Short-Time Objective Intelligibility Measure for Time-Frequency Weighted Noisy Speech", in Proc. ICASSP, 4214-4217, 2010
- (5) SEGAN : Speech Enhancement Generative Adversarial Network: Santiago Pascual, Antonio Bonafonte¹, Joan Serrà¹ 1Universitat Politècnica de Catalunya, Barcelona, Spain Telefónica Research, Barcelona, Spain, 2017