Evaluation of a wind noise reduction method using DNN for Bicycle Audio Augmented Reality Systems

Toya Kitagawa, Kazuhiro Kondo, Graduate School of Science and Engineering Yamagata University, Japan
e-mail: klabkitagawa@yahoo.co.jp, kkondo@yz.yamagata-u.ac.jp

***Abstract*--In a previous study, we studied wind noise reduction methods to use with navigation systems for bicycles. As a result, the Wiener (iteration) filter was superior in wind noise elimination performance, and it was able to give a very high navigation speech intelligibility. However, not only the wind noise but also the environmental sounds were overly attenuated. Therefore, in this paper, DNN was introduced to wind noise reduction, Wind noise reduction and environmental sound performance were evaluated. As a result, it was concluded that DNN has high wind noise removal performance while environmental sound can be preserved.**

I. INTRODUCTION

In recent years, navigation applications such as GoogleMap have been used due to the spread of smartphones. However, this is supposed to be used on a fixed terminal in a car, or used by a pedestrian. When using it on a bicycle, there is a danger that it will lead to an accident by forcing the user to gaze at a small screen. Therefore, we propose a speech navigation system using Audio Augmented Reality (AAR). This eliminates the need to expose the terminal, so it is not affected by the climate, and the user does not have to stare at the screen. However, since it is navigation with only voice information without visual information, it is difficult to accurately convey instructions. In order to navigate only with sound without looking at the screen, we intuitively indicate the direction to go by adding direction information using a sign signal localized towards the target direction as guidance. In order to use stereophonic sound, a presentation device such as an earphone is necessary. However, by blocking the ear, the environmental sound is also blocked out. Thus, there is a danger of not noticing, for example, an approaching car. Therefore, hear-through earphones, capable of acquiring ambient sound from the integrated microphone and feedback through the earphones, are effective. When using hear-through earphones, not only environmental sounds but also wind noise is acquired at a significant level. This makes the navigation speech and environmental sounds difficult to hear. Accordingly, we are developing a system capable of delivering environmental sound while attenuating wind noise. The overall configuration of the system is shown in Figure 1. The environmental sound is taken from the microphone portion of the hear-through earphone and output together with the navigation sound. At this time, we want to reduce the wind noise taken in along with the environmental sound. In previous research, we aimed to evaluate and compare several methods to identify the optimum wind noise countermeasure by actually evaluating it under the condition used on a bicycle in real time [1]. As a result, the Wiener (iteration) filter was able to eliminate almost all wind noise and obtained high intelligibility. However, not only the wind noise but also the environmental sound was also almost completely attenuated. This is problematic since the user will not be able to grasp potential hazards in the environment. In this paper, we aimed to reduce wind noise by using DNN and evaluated wind noise reduction and environmental sound acquisition performance.

II. EVALUATION OF WIND NOISE REDUCTION USING DNN

We configured and trained a Deep Neural Network (DNN) to attenuate only the wind noise while preserving the environmental sound (horn from an automobile in this paper). The conditions for the DNN training are shown in Table 1, with mixed wind noise and horn as training, and horn only as the supervisory data. The input and the output of the DNN is the magnitude component of the Short-Term Fourier Transform (STFT) of the signal. The phase component of the input was used along with the DNN output to generate the processed signal. In order to generate a localized horn signal, a monaural signal, obtained from auto manufacturer's Web page, was convolved with the HRTF of the specified azimuth (0, ±45, ±90, ±135 deg) on the horizontal plane. The localized horn signal was mixed with pre-recorded wind noise (recorded using the hear-through earphone while riding a bicycle).

The trained DNN was applied to an unknown mixed wind noise and horn sample to attenuate the wind noise while preserving the horn. Fig. 2 shows the spectrogram of the unknown mixed sound, and Fig. 3 shows the spectrogram of the DNN output. As can be seen, most of the wind noise is attenuated, while the horn is mostly preserved. We were able to archive a 10.7 dB reduction of the wind noise using DNN. The SNR of the horn was about 4.41 dB. However, in extremely noisy conditions, where the SNR between the horn signal and the wind noise is below -20 dB, we did see some instances where the horn signal was completely attenuated. We believe this is caused by insufficient training data under extremely noisy conditions. We plan to improve this situation by adding more extremely noisy conditions as training data.

III. EVALUATION OF ENVIRONMENTAL SOUND PERCEPTION ACCURACY

It was confirmed in the previous chapter that there is almost no attenuation of the environmental sound when DNN is applied. Therefore, in this chapter, we evaluated whether the direction of the environmental sound can be grasped correctly even if DNN is applied.

We applied DNN to horn sounds recorded with hear-through earphones, and wind noise was added at SNR = 0 dB (horn and wind noise set at equal levels). We prepared three different horn sounds. Ten subjects selected the direction of the horn sounds. In this test, we regarded the error within a range of ± 45 ° as correct since we can assume that the general direction of the horn was perceived correctly in this case. The results are shown in Table 2. Regardless of the type of horn, the mean of direction recognition rate was over 90% correct. This is a significant improvement over Wiener (iteration) filter.

Therefore, even if DNN is employed, it can be said that the direction of the environmental sound can be perceived correctly.

IV. SUMMARY

In this paper, the wind noise reduction by DNN was examined, and the environmental sound attenuation, which was a problem with the Wiener (iteration) filter, was evaluated. From both the wind noise reduction and the environmental sound direction perception accuracy, we found that using DNN is more effective than the Wiener (iteration) filter. However, in some low SNR conditions, the horn sound could not be separated correctly. Thus we need to increase training data and are corresponded to large wind noise and unknown data in the future.

TABLE I
TRAINING CONDITION USING DNN

	Items	Specification
NN	Input scaling	Standard scaler
	Output scaling	Standard scaler
	Intermediate layers	3
	Intermediate layer activation function	PReLU
	Output layer activation function	Linear
	Units	2000
	Batch size	500
	Drop out rate	0.5
	Epochs	1000
	Loss function	Squared error
	Optimization algorithm	Adam
	Learning rate	0.001
STFT Front-end	Frame length	320 samples
	Window	Hann
	Overlap	50%
Training data	Sampling frequency	8 kHz
	Quantization bits	16 bits
	Channels	Monaural
	Wind noise	5 patterns
	Horn	10 patterns
	Direction	8 every 45 deg

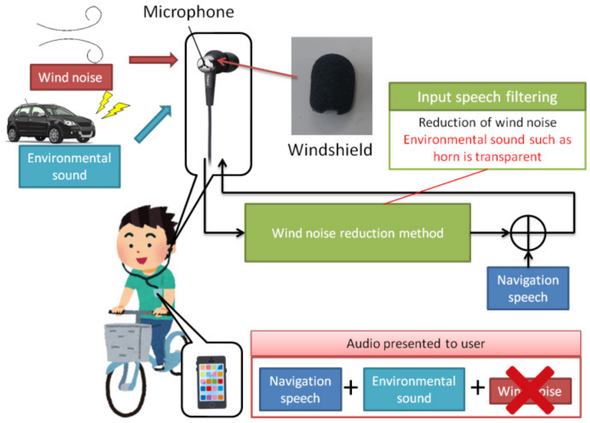

Fig 1. System Configuration

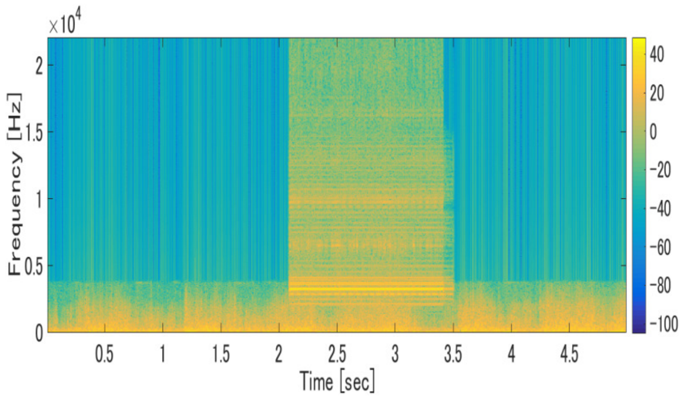

Fig 2. Spectrogram of unknown data (wind noise + horn)

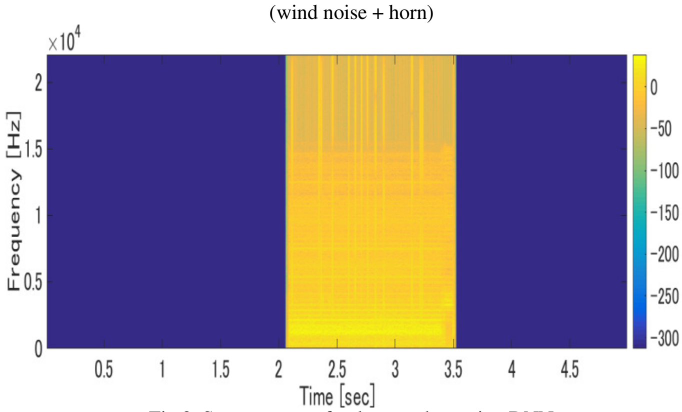

Fig 3. Spectrogram of unknown data using DNN

TABLE II
EVALUATION OF DIRECTION ACQUISITION PERFORMANCE OF DNN

Wind noise reduction	Horn1	Horn2	Horn3
DNN	95.6	91.3	95.6
Wiener (iteration)	66.7	82.2	73.0

REFERENCE

[1] T. Kitagawa, K. Kondo, "On a Wind Noise Countermeasure for Bicycle Audio Augmented Reality Systems", 2017 IEEE 6th Global Conference on Consumer Electronics (GCCE 2017), 2017.10.