MADRID
inter.noise 2019
June 16 - 19
NOISE CONTROL FOR A BETTER ENVIRONMENT

Adaptive noise cancellation for improving speech recognition

Lee, Nokhaeng[1]
Samsung Electronics Co., Ltd.
129, Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do 16677, Korea

Seo, Eung Ryeol[2]
Samsung Electronics Co., Ltd.
129, Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do 16677, Korea

ABSTRACT

The purpose of this paper is improvement of speech recognition in noisy device itself. The performance of speech recognition deteriorates in noisy device because it depends on the level of signal to noise. Normally, noise reduction techniques are applied in postprocessing with recorded voice signal with noise. But, it is effective only when the level of voice is higher than those of noise. The condition is not satisfied in heavy noisy devices such as robot vacuum cleaner and so on. Under this background, a method of adaptive noise cancellation is proposed for improving speech recognition in noisy device itself. As a case study, a robot vacuum cleaner is investigated with respect to what the main noise source is and how the noise propagates. An adaptive noise cancellation is applied with the placement of reference sensors and development of adaptive algorithm. Numerical simulation shows that the proposed scheme has a potential application.

Keywords: Adaptive noise cancellation, Speech recognition, Voice recognition
I-INCE Classification of Subject Number: 74

1. INTRODUCTION

Voice or speech recognition is a useful technology to help machines to communicate with people. It is the process of capturing spoken words with microphones and converting them into some digitally stored set of words. Recently, it has been more powerful technology based on deep learning with massive data of human voices. The performance is measured by accuracy and speed. Among them, the accuracy can deteriorate in noisy environment. Normally, signal processing techniques are applied with several signals measured by using microphone array. They are effective when the voice signal is generated in a particular direction and the level is higher than background noise. But, they are not enough in extremely noisy environment even if the intense noise steadily exists. Besides, methods of spectral noise reduction are applied for the accuracy. But, they are not enough when the frequency range is similar with those of the voice.

---

[1] norukang@gmail.com

In this paper, a pre-processing technique is applied to overcome the problem through a case study of speech recognition in robot vacuum cleaner. To determine which noise is dominant, the power spectral density of the noise signals measured at microphone array is compared to those of voice signals. An adaptive noise cancellation is proposed to remove the target noise in measured signal at microphone array. Noise generation process is investigated and additional microphones are placed and used as reference signals. An adaptive algorithm is developed to continually estimate the target noise. Numerical simulation shows that the proposed scheme has a potential application.

2. Problem definition

One of robot vacuum cleaner with two microphones installed to capture spoken words is investigated. It is placed in semi-anechoic chamber and adjusted to keep operating without moving. The microphone signals are sampled by 16 kHz which is enough to cover the frequency range of interest for the speech recognition. The top of Figure 1 shows the time-domain signals measured at two microphones and the periodicity is easily observed in both signals. It is observed that the fundamental frequency of about 330 Hz and its harmonics are dominant in the frequency domain as the down-left in Figure 1. The fundamental frequency is corresponding to the blade passing frequency of the fan for the suction. The component steadily exists while operating as the down-right in Figure 1.

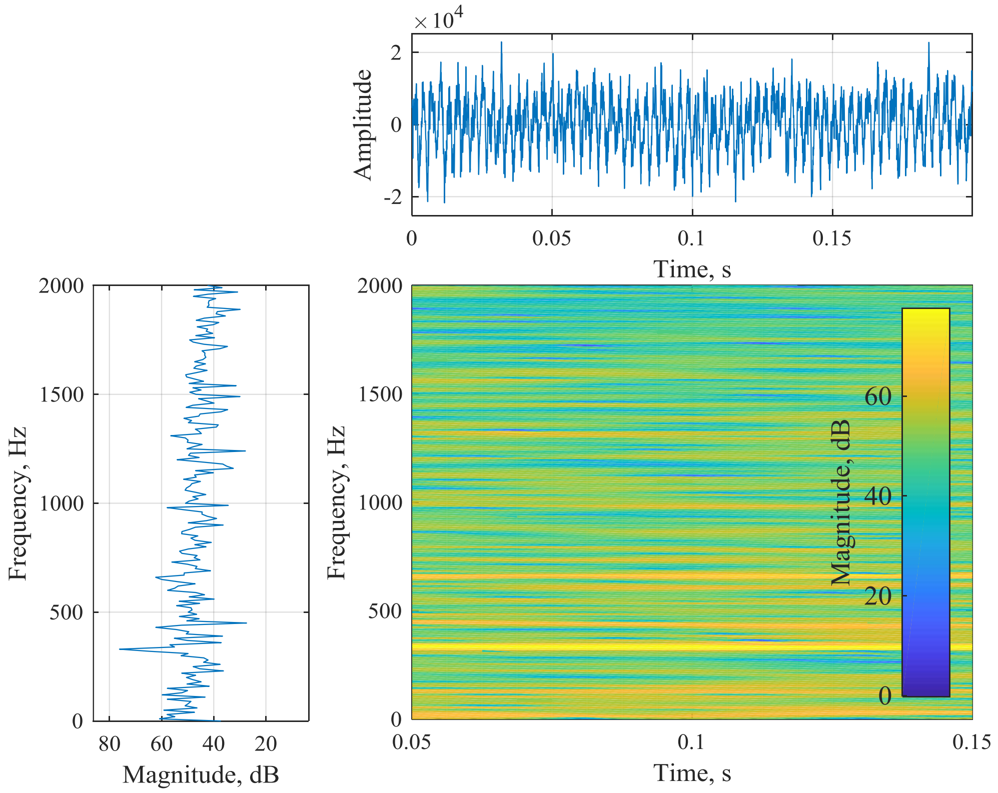

Figure 1 Time-frequency domain analysis of the noise signal measured by a voice microphone in robot vacuum cleaner.

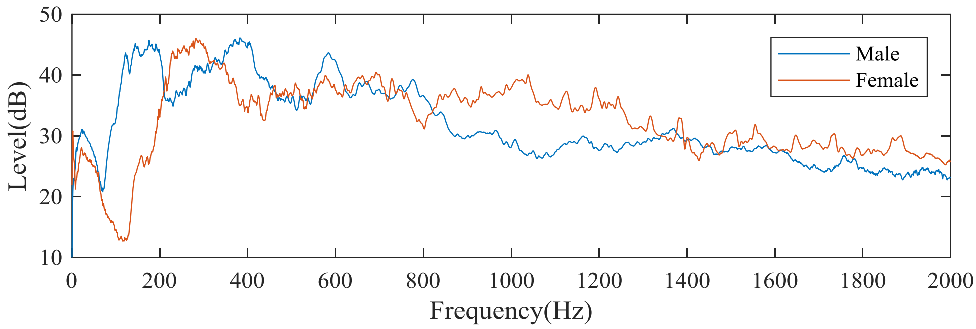

Figure 2 Power spectral density of measured signal of male and female voice.

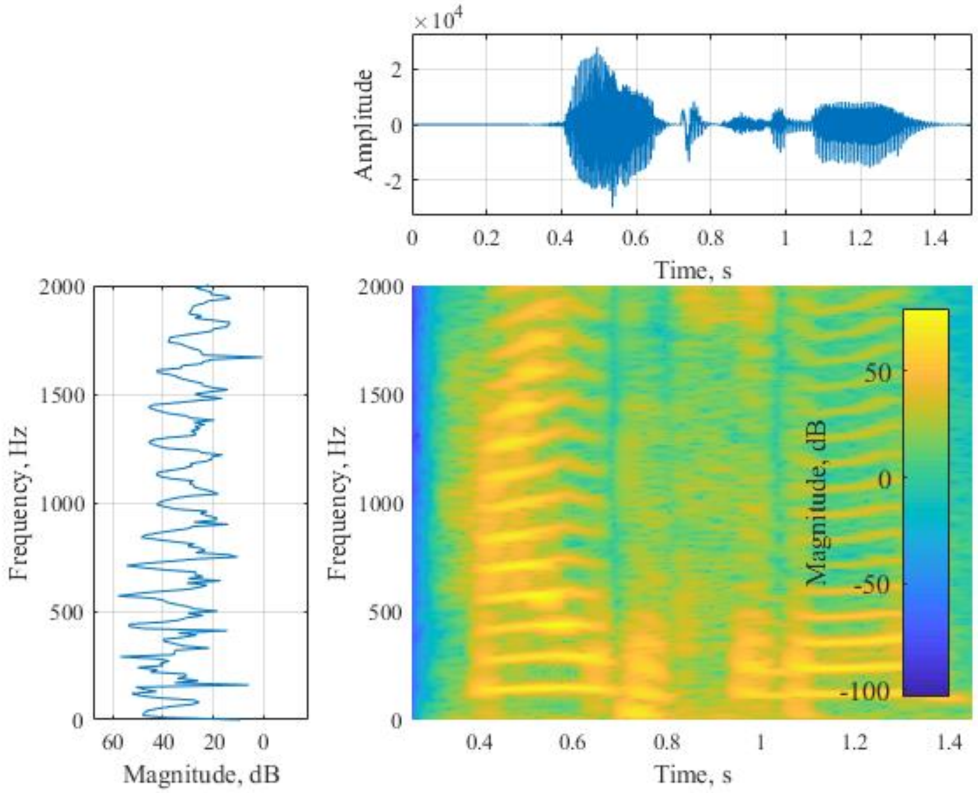

Figure 3 Time-frequency domain analysis of a sample signal of male voice.

Meanwhile, speech signals which are wanted to be clearly captured are also investigated. Total 20 people including 10 male and 10 female speak a particular wake-up-word fifteen times. The power spectral density is shown in Figure 2. The main frequency band of male voice is 150 to 1000 Hz and those of female voice is higher indeed as 200 to 1500 Hz. More specifically, a sample of spoken word by male is investigated in time and frequency domain as Figure 3. It is observed that the signal is quasi-periodic and has fundamental frequency and its harmonics in short time.

It is difficult to determine how the performance of speech recognition deteriorate in noisy environment exactly. But it is expected that it can deteriorate when the spectrum of noise exists around those of voice or overlaps it. So, we set the objective as cancellation of noise from voice signal in noisy environment and the frequency range of interest is from 150 Hz to 2000 Hz.

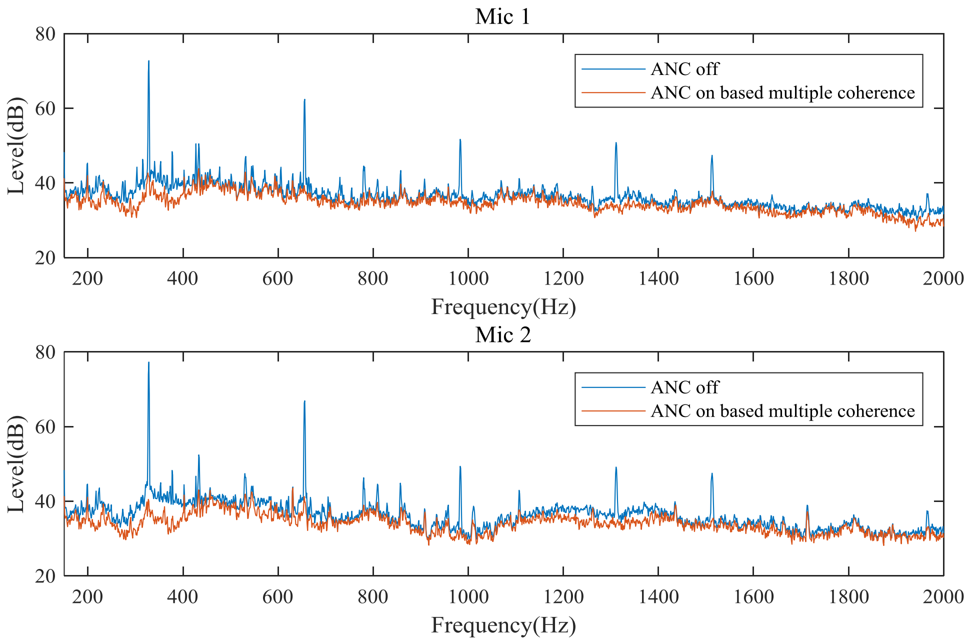

Figure 4 Expected performance of adaptive noise cancellation based on multiple coherence between reference signals and voice signals in noisy environment

3. Adaptive noise cancellation system

To reduce the noise from mixed signal of voice and noise, it should be estimated. It is clear that the periodic component of blade passing frequency is dominant. So, reference sensors are placed in the path from the noise source to the microphone. Additional microphones are placed as candidates for the reference sensors. Two microphones are selected after comparison of the reduction performance based on multiple coherence. The expected reduction performance is calculated based on multiple coherence as Figure 4. The maximum peak of 330 Hz is reduced by more than 30 dB and the harmonics are reduced by more than 10 dB.

With the reference microphones, an adaptive noise cancellation algorithm is applied. The purpose of the algorithm is to estimate noise signal by using the transfer function of the path from reference microphones to voice microphones and to remove it from the voice signal with noise. The transfer function is affected by not only the direct path from reference to voice microphones but also indirect path including reflections by surroundings. Unfortunately, all robot vacuum cleaners keep moving while operating. It is almost impossible to expect all kinds of the indirect path in practical circumstances. So, the algorithm is continually updated to adapt the change of the transfer function. The block diagram of the algorithm is illustrated in Figure 5.

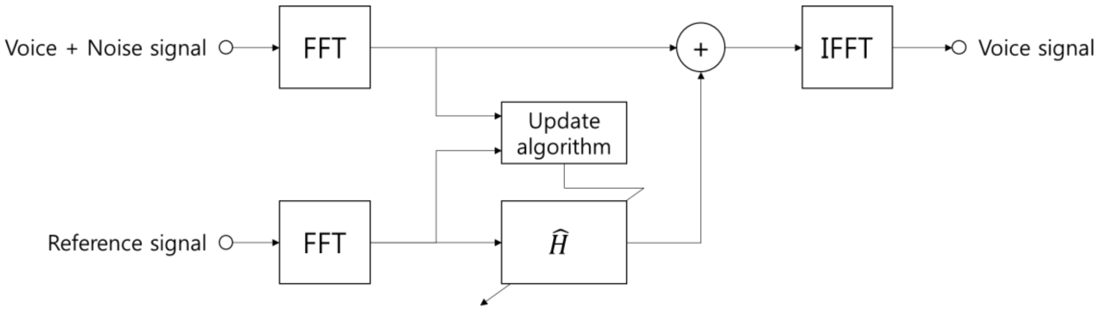

Figure 5 Block diagram of frequency-domain adaptive noise cancellation algorithm.

The voice and reference microphone signals are transformed to the frequency domain. To estimate the transfer function H, autocorrelation of reference signal and cross correlation of reference and voice signal are continually calculated and cumulated for the expectation effect. The transfer function H is achieved by multiplication of cross correlation and inverse of the autocorrelation. Then, noise signal is removed from the mixed signal of voice and noise signal in the frequency domain. Finally, the time-domain voice signal is achieved by inverse Fourier transform of the resultant frequency-domain signal.

Numerical simulation is performed with recorded signal of adaptive noise cancellation system. Robot vacuum cleaner is operated without moving and voice signal is repeated by loudspeaker at 2-m distance. The result is shown in Figure 6 and 7. When adaptive noise cancellation off, the periodic noise steadily captured at voice microphones at about 330 Hz and its harmonics as Figure 6. After adaptive noise cancellation on, the periodic noise is reduced down to the level of about 50 dB as Figure 7.

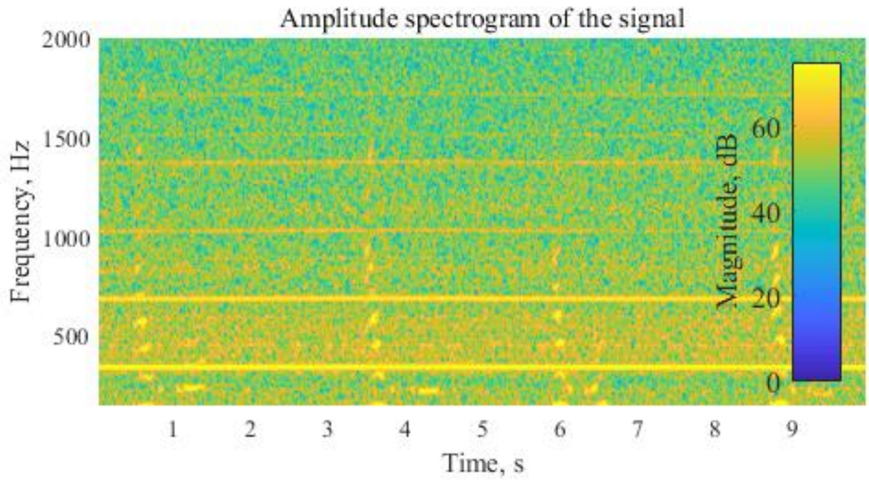

Figure 6 Time-frequency domain voice signal in noisy environment (adaptive noise cancellation off)

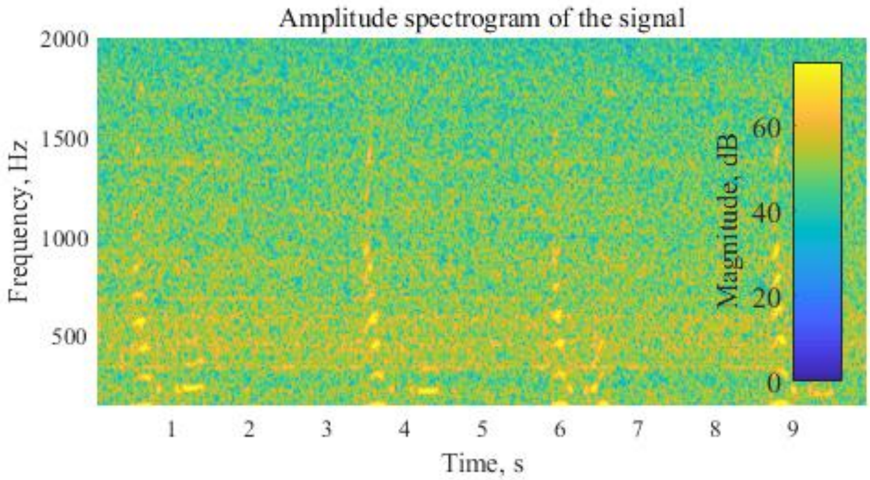

Figure 7 Time-frequency domain voice signal in noisy environment (adaptive noise cancellation on)

4. CONCLUSIONS

An adaptive noise cancellation system is proposed for the improvement of speech recognition in robot vacuum cleaner as a case study. The main noise source of the robot vacuum cleaner is generated by the blade passing of suction fan. Two additional microphones are placed in the path from the noise source to voice microphones and the signal is used as the reference. An adaptive algorithm is applied with measured reference and voice microphone signals. It continually estimates the transfer function from reference to voice microphone signal and extracts voice signal by removing estimated noise signal. Numerical simulation shows that the proposed scheme has a potential application.

5. REFERENCES

1. Van Compernolle, Dirk, et al. "Speech recognition in noisy environments with the aid of microphone arrays." *Speech Communication* 9.5-6 (1990): 433-442.
2. Kumatani, Kenichi, John McDonough, and Bhiksha Raj. "Microphone array processing for distant speech recognition: From close-talking microphones to far-field sensors." *IEEE Signal Processing Magazine* 29.6 (2012): 127-140.
3. Boll, Steven. "Suppression of acoustic noise in speech using spectral subtraction." *IEEE Transactions on acoustics, speech, and signal processing* 27.2 (1979): 113-120.
4. Fujimoto, Masakiyo, and Yasuo Ariki. "Noisy speech recognition using noise reduction method based on Kalman filter." *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 00CH37100)*. Vol. 3. IEEE, 2000.