2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Study on Optimal Position and Covering Pressure of Wearable Neck Microphone for Continuous Voice Monitoring

Yonghun Song, Yunsik Kim, Inyeol Yun, *Student Member, IEEE*, Jinpyeo Jeung, Jiwon Kang and Yoonyoung Chung, *Member, IEEE*

***Abstract*— Vocal cord disorder is one of the important health problems, especially in noisy industrial sites where excessive voice is required. A convenient and reliable communication method is required in a noisy environment to prevent the related disorders. However, the signal sensitivity of previous neck microphones is still insufficient to accurately convey the voice. In this study, we developed a skin-attachable neck microphone with a lightweight and flexible form factor. Also, we optimized the attachment position and covering pressure to maximize the signal sensitivity. As a result, we obtained the optimal position near the thyroid cartilage and confirmed that the signal sensitivity is the highest when the covering pressure is approximately 4 mmHg.**

***Clinical Relevance*— People can measure the voice status using a wearable neck microphone at the optimal position and covering pressure. It provides a solution to keep the vocal cords in good health even in a noisy environment.**

I. Introduction

Speaking is one of the important biological signals of humans. Through the voice, people communicate and control the voice recognition system [1]. A voice can be measured through a microphone, which converts the sound wave pressure into electrical signals. However, conventional microphones have a disadvantage that it not only transmits the voice signal but also the ambient noise. Especially in some severe surroundings near factories and construction areas, an accurate voice signal is difficult to be delivered. In this manner, people speak loudly to communicate and use their voices excessively. This continuous phonation leads to abuse and misuse of the voice, which causes vocal cord disorders, such as laryngitis and vocal cord nodule [2-4]. To prevent this problem, numerous research activities were conducted to reduce the noise measured by the microphone. The noise signal can be filtered based on the signal interference of multiple microphones [5]. However, this signal processing cannot completely remove the noise, and a novel approach is required for noise-free voice delivery in a noisy environment.

The neck microphone has been widely studied to reduce the effects of noise [6-8]. It measures the acceleration of the skin near the vocal cords, and the acceleration is known to be proportional to the sound pressure level [9]. The neck microphone can be conveniently used in various environments where noise and wind are severe.

The signal sensitivity of neck microphone is highly dependent on the usage. For example, the position where the neck microphone is placed and its covering pressure on the skin significantly affect the output signal. A previous study on attachment position of the neck microphone observed the signal intensity at the fundamental frequency and designated optimal position [10]. Since the voice contains various elements with natural overtones of the fundamental frequency [11], the previous study shows a certain limitation in finding the optimal attachment position. Moreover, the relationship between the covering pressure of neck microphone and the output signal quality has not yet been reported.

In this paper, the position and covering pressure of the wearable neck microphone are optimized by analyzing the frequency components of measured signals. We introduce a new methodology to find the similarity between voice signals: a combination of the Short-Time Fourier Transform (STFT) and Root Mean Square Error (RMSE) model. The quantitative similarity was compared at nine positions on the neck skin with ten different covering pressures. The optimal attachment position was found to be near the thyroid cartilage, and the optimal covering pressure was approximately 4 mmHg. With optimized position and covering pressure, the neck microphone can detect the voice accurately to communicate in a noisy environment and provide information for diagnosing vocal cord disorders.

II. Device and Methods

A. Wearable Neck Microphone System

Our neck microphone consists of signal measurement, processing, and transmission modules, as shown in Fig. 1. The signal measurement module included an accelerometer (STMicroelectronics, LIS2DS12), and it was attached onto the skin near the vocal cords. It measured the acceleration of skin vibration when speaking. An accelerometer sensor must be conformably attached onto the skin for measuring the acceleration accurately; therefore, a flexible polyimide substrate was used instead of a rigid printed circuit board (PCB). The skin acceleration on the vocal cords contains small-amplitude. In this experiment, the accelerometer was prepared with a dynamic range and sampling frequency (f_s) of ±2 g and 6,400 Hz, respectively. In the signal processing module, the measured accelerometer signal was passed through a built-in low-pass filter with a cut-off frequency of

This work was supported by the Joint Research on Smart Solution project of POSCO and by the Sports Promotion Fund of Seoul Olympic Sports Promotion Foundation from the Ministry of Culture, Sports and Tourism.

Y. Song, Y. Kim, I. Yun, J. Jeung, and Y. Chung are with the Department of Electrical Engineering, Pohang University of Science and Technology (POSTECH), Pohang, 37679 Korea. (e-mail: {yhsong, ys.kim, inyul1225, jpjeung, ychung*}@postech.ac.kr)

J. Kang is with the School of Electronic Engineering, Soongsil University, Seoul, 06978 Korea. (e-mail: primewisdom@soongsil.ac.kr)

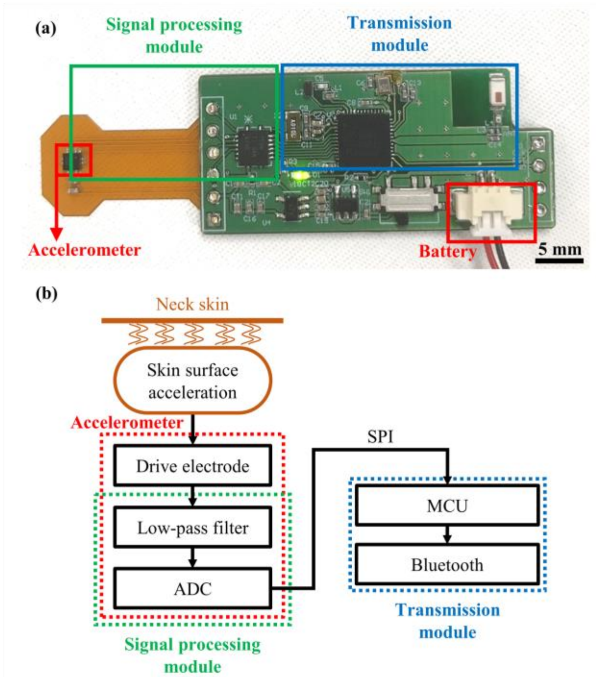

Figure 1. (a) Photograph and (b) block diagram of the wearable neck microphone. The wearable neck microphone includes three modules to detect the voice: signal measurement, processing, and transmission modules.

2,840 Hz. The signal was then converted into the digital domain by an analog-to-digital converter (ADC) and transmitted to a microcontroller unit (MCU, Nordic Semiconductor, nRF52832) by serial peripheral interface (SPI) communication. Finally, the signal was collected in a laptop or smartphone through Bluetooth.

B. Signal Measurement in Wearable Neck Microphone

The voice measurement setup is shown in Fig. 2. Five male subjects with no history of vocal cord disorder participated in this experiment without ambient noise. The neck microphone was attached onto the skin near the vocal cords with a skin tape (3M, Tegaderm). A reference microphone (BOYA, BY-PVM50), which is known to measure the sound wave pressure accurately, was placed 30 cm apart from the subject's lips. The subjects sat in a chair and read the first two sentences of the rainbow passage [12], which was written by a speech pathologist and commonly used for voice tests. When the subjects read the rainbow passage, the signals from the neck microphone and reference microphone were simultaneously recorded on the laptop.

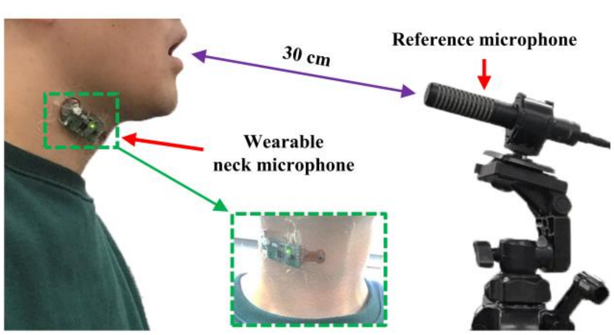

Figure 2. Experiment setup for measuring voice signal using wearable neck microphone and reference microphone, simultaneously.

To optimize the attachment position of the neck microphone, a mapping test was performed as shown in Fig. 3(a). Nine attachment positions were tested on the circle with a radius of 2 cm centered on the thyroid cartilage. In each measurement, the neck microphone was attached onto one of the test positions. Fig. 3(b) shows the measurement setup to study the performance of the neck microphone as a function of the covering pressure. A thin (2 mm) pressure sensor device (TT meditrade, Kikuhime) was mounted on the neck microphone and attached onto the skin using a Tegaderm film. The covering pressure was controlled through a compression bandage.

C. Quantitative Analysis between Wearable Neck Microphone and Reference Microphone

We compared the wearable neck microphone and the reference microphone in terms of the output signal sensitivity. In a previous study, the signal intensity from the microphones was compared [13]; however, it is not enough to quantify only the different intensities between the microphones. Because voice also has a frequency component. Fig. 4 shows our proposed quantitative similarity estimation process. First, we sampled the measured signal from the reference microphone with f_s of 6,400 Hz, which is equivalent to the neck microphone. Then, the two digital signals were normalized to have the same dynamic range and synchronized. We removed the high-frequency components by using a 4th-order Butter-worth low-pass filter with a cut-off frequency of 2,800 Hz to prevent the aliasing problem. A spectrogram was plotted by performing the STFT to analyze both frequency and intensity of the voice signal. The window frame width of STFT was 0.08 s, and the overlap width was 0.02 s. The intensity of each pixel in the spectrogram forms a matrix, and the RMSE was calculated from the corresponding elements of the two matrices. RMSE is defined as below:

$$RMSE = \sqrt{\frac{1}{M \times N} \sum_{j=1}^{N} \left(\sum_{i=1}^{M} (R_{ij} - N_{ij})^2 \right)} \quad (1)$$

where R_{ij} and N_{ij} are matrices of size M x N by STFT result of the reference microphone and neck microphone, respectively. As a result, the quantitative similarity of the two voice signals

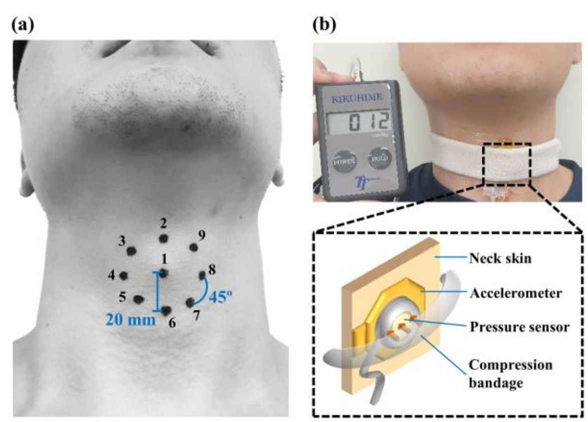

Figure 3. (a) Attachment position of wearable neck microphone. (b) Experiment setup for applying external pressure with compression bandage.

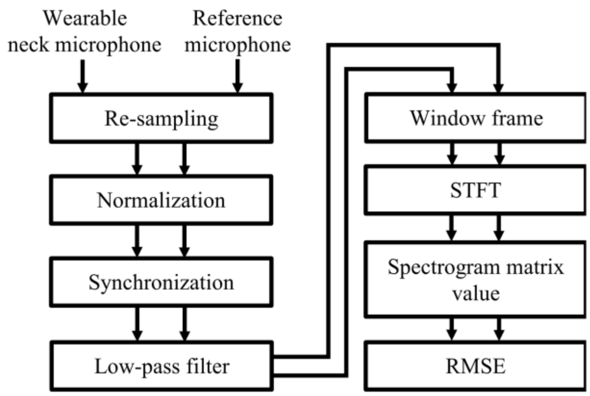

Figure 4. Block diagram of the similarity estimation methodology between wearable neck microphone and reference microphone.

was extracted. All signal analysis process was conducted with MATLAB R2019a.

III. Analysis and Results

A. *Quantitative Similarity Estimation based on the Attachment Positions*

When voice is present in the vocal tract, the accurate voice is transmitted to the accelerometer through a transmission medium such as muscle, tissue, and cartilage. Transmission medium and presence or absence of voice differed depending on the attachment position, and the output signal can be varied. We evaluated the quantitative similarity while varying the attachment position of neck microphone. In Fig. 5, the similarity was estimated at nine positions. In all subjects, the similarity was high in the upper semicircle in position 1 (thyroid cartilage). As the glottis is located behind the thyroid cartilage, this result is consistent with the fact the voice from articulation organ is reflected at the mouth and transmitted to the supraglottis [14]. The reflected voice cannot be transmitted to the subglottis because the glottis becomes narrow during phonation. Therefore, below position 1, the similarity was estimated to be low. Subjects have different structures of the vocal cords, so the magnitude of RMSE can be varied. However, all subjects showed similar quantitative tendencies based on the attachment position. As a result, the optimal position for attaching the neck microphone was either position 1 or 2.

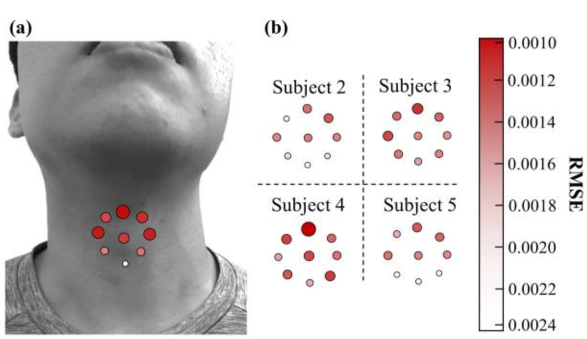

Figure 5. Quantitative similarity estimation in various positions: (a) subject 1 and (b) the other subjects. Similarity increases with the size and redness of the circle.

B. *Quantitative Similarity Estimation based on the Covering Pressures*

When attaching the neck microphone onto the skin, a certain force must be applied perpendicular to the skin on the vocal cords. As the pressure affects the transmitted vibration to the accelerometer, the measured signal can be varied. In Fig. 6, the quantitative similarity is correlated to the covering pressure. When the covering pressure was 4 mmHg, the signals from the two microphones were the most similar. When the covering pressure was less than 4 mmHg, the neck microphone did not contact conformally on the skin surface. In this case, the vibration of the skin surface cannot be measured accurately. In contrast, when the covering pressure exceeded 4 mmHg, the transmitted vibration was attenuated by the applied force, which is 180 degrees different from the direction of the skin surface acceleration. The attenuation was increased as the covering pressure was enlarged. The optimal covering pressure was measured to be 4 mmHg, which gave the highest signal similarity in the comparison between the neck microphone and the reference microphone.

C. *Reliability Verification of Wearable Neck Microphone*

We compared the waveforms of the neck microphone with the reference microphone to verify the proposed position and the pressure. Waveforms of the neck microphone and reference microphone were measured in a silent environment (Fig. 7a). The two signals were measured to be very similar. However, the output from the two microphones became different in a noisy environment (Fig. 7b). To set up a noisy environment, a speaker was placed at 30 cm from the two microphones, and white noise was reproduced at 70 dB_{SPL}. Since the neck microphone was not affected by sound pressure, which is the vibration of the air, it can measure voice without noise in a noisy environment. However, noise was measured by the reference microphone. Thus, these results demonstrate that the neck microphone detecting the authentic voice can be used to monitor the vocal cords condition and to allow conveniently communication with others even under noisy environments.

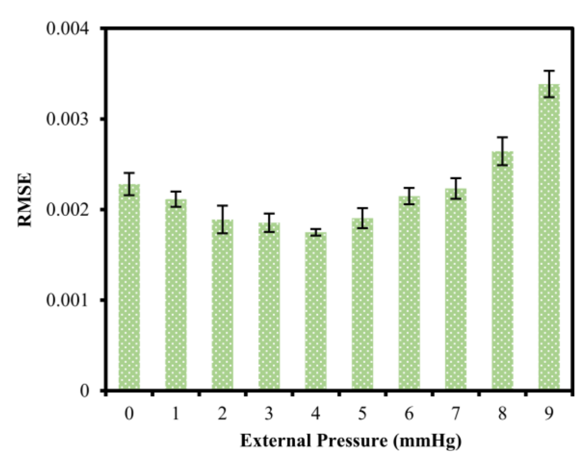

Figure 6. Quantitative similarity estimation under different covering pressures. Error bar means standard deviation.

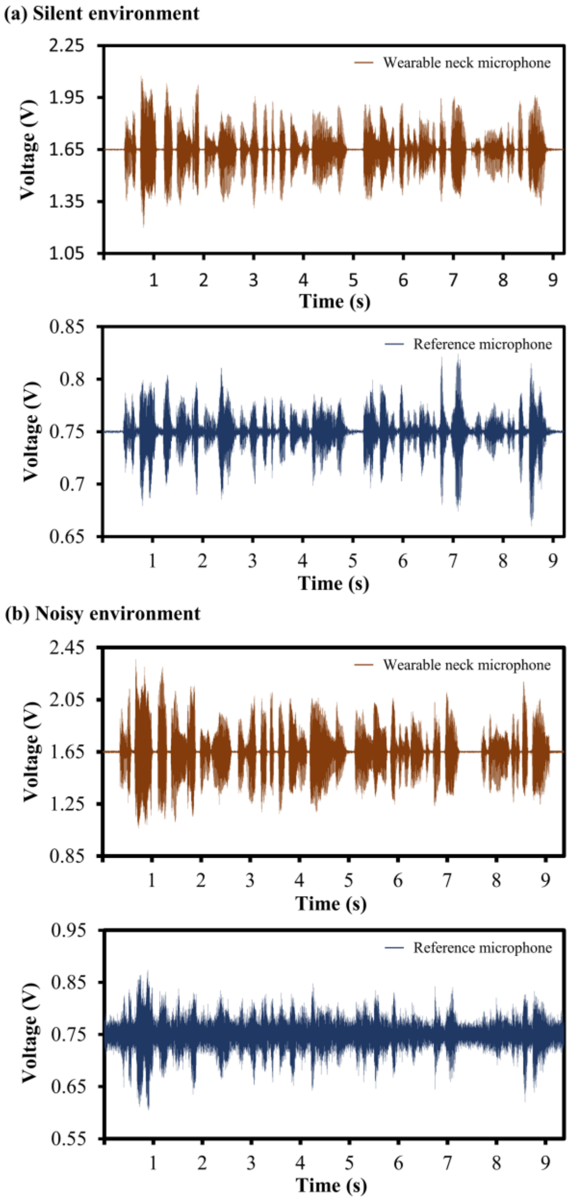

Figure 7. Measured voice signals from wearable neck microphone and reference microphone (a) in a silent environment and (b) in a noisy environment. The neck microphone is not affected by ambient noise.

IV. CONCLUSION

In this paper, we present a wireless wearable neck microphone that can be used in a noisy environment. The device is barely affected by the ambient noise. The optimal position and covering pressure of the neck microphone were studied using a quantitative similarity estimation methodology. As a result, the optimal position to attach the neck microphone was found to be near the thyroid cartilage, and the optimal covering pressure was 4 mmHg. Through the optimal position and covering pressure of the neck microphone, we confirmed that people can communicate with others even under noisy, windy, or other harsh environments and monitor the vocal cords condition with high reliability.

ACKNOWLEDGMENT

This work was supported by the Joint Research on Smart Solution project of POSCO and by the Sports Promotion Fund of Seoul Olympic Sports Promotion Foundation from the Ministry of Culture, Sports and Tourism.

REFERENCES

[1] P. R. Cohen, and S. L. Oviatt, "The role of voice input for human-machine communication," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 22, Oct. 1995, pp. 9921–9927.

[2] P. Carding, and A. Wade, "Managing dysphonia caused by misuse and overuse: Accurate diagnosis and treatment is essential when the working voice stops working," *British Medical Journal*, vol. 321, no. 7276, Dec. 2000, pp. 1544-1545.

[3] K. W. Altman, C. Atkinson, and C. Lazarus, "Current and emerging concepts in muscle tension dysphonia: a 30-month review," *Journal of Voice*, vol. 19, no. 2, Jun. 2005, pp. 261–267.

[4] M. D. Morrison, and L. A. Rammage, "Muscle misuse voice disorders: description and classification," *Acta Oto-Laryngologica*, vol. 113, no. 3, May. 1993, pp. 428–434.

[5] M. Togami, "Multichannel online speech dereverberation under noisy environments," *2015 23rd European Signal Processing Conference (EUSIPCO) IEEE*, Aug. 2015, pp. 1078–1082.

[6] B. Park, J. Kim, D. Kang, C. Jeong, K. S. Kim, J. U. Kim, P. J. Yoo, and T. Kim, "Dramatically enhanced mechanosensitivity and signal-to-noise ratio of nanoscale crack-based sensors: effect of crack depth," *Advanced Materials*, vol. 28, issue. 37, Oct. 2016, pp. 8130–8137.

[7] S. Kang, S. Cho, R. Shanker, H. Lee, J. Park, D. S. Um, Y. Lee, and H. Ko, "Transparent and conductive nanomembranes with orthogonal silver nanowire arrays for skin-attachable loudspeakers and microphones," *Science Advances*, vol. 4, no. 8, Aug. 2018.

[8] L. Q. Tao, H. Tian, Y. Liu, Z. Y. Ju, Y. Pang, Y. Q. Chen, D. Y. Wang, X. G. Tian, J. C. Yan, N. Q. Deng, Y. Yang, and T. L. Ren, "An intelligent artificial throat with sound-sensing ability based on laser induced graphene," *Nature Communications*, vol. 8, no. 14579, Feb. 2017.

[9] S. Lee, J. Kim, I. Yun, G. Y. Bae, D. Kim, S. Park, I. M. Yi, W. Moon, Y. Chung, and K. Cho, "An ultrathin conformable vibration-responsive electronic skin for quantitative vocal recognition," *Nature Communications*, vol. 10, no. 2468, Jun. 2019.

[10] M. Nolan, B. Madden, and E. Burke, "Accelerometer based measurement for the mapping of neck surface vibrations during vocalized speech," *31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Sep. 2009, pp. 4453–4456.

[11] I. R. Titze, "The human instrument," *Scientific American*, vol. 298, no. 1, Jan. 2008, pp. 94–101.

[12] G. Fairbanks, *Voice and articulation drillbook*, New York: Harper and Row, vol. 2, Jan. 1960.

[13] J. G. Švec, I. R. Titze, and P. S. Popolo, "Estimation of sound pressure levels of voiced speech from skin vibration of the neck," *The Journal of the Acoustical Society of America*, vol. 117, no. 3, Mar. 2005, pp. 1386–1394.

[14] N. Hanna, J. Smith, and J. Wolfe, "Low frequency response of the vocal tract: acoustic and mechanical resonances and their losses," *Proceedings of Acoustics, Fremantle, Australia*, Nov. 2012.