Remote Photoplethysmography-Triggered Signal Averaging for Non-Contact Heart Sound Measurement

Shun Muramatsu, *Member, IEEE,* and Kosuke Kurihara, *Member, IEEE*

Abstract—Remote photoplethysmography- (rPPG-) triggered signal averaging, a novel approach for the non-contact heart sound measurement, was proposed. Recently, there has been a growing need for a heart sound monitoring method for the early detection of cardiovascular diseases. Therefore, we proposed a non-contact heart sound measurement approach using a signal averaging with rPPG peaks as triggers. This approach can emphasize the heart sound, the only sound synchronized in time with the heartbeat, and reduces other ambient noises. Herein, we evaluated the feasibility of our approach. As a result, it was indicated that our approach can measure the heart sound without contact using the appropriate triggers. The accuracy required for the triggers was approximately 0.04 s as the standard deviation of the R-R interval estimation error.
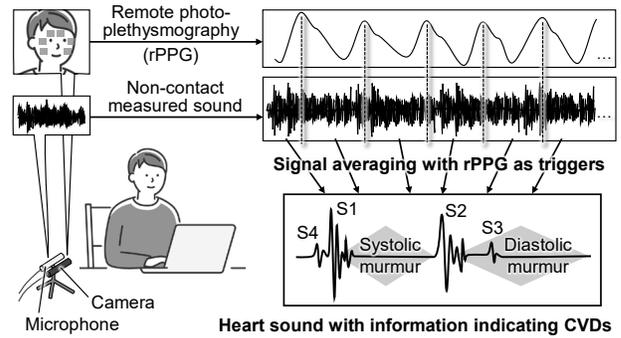
Fig. 1. Outline of the rPPG-triggered signal averaging, our method for non-contact heart sound measurement.

I. INTRODUCTION

Cardiovascular diseases (CVDs) are the most common causes of death worldwide. These accounted for 32% of all cases of death in 2019 [1], and this number was reported to be increasing further [2]. Most of CVDs are preventable because these are caused by unhealthy lifestyle habits. Therefore, to monitor and detect them early is very important [3].

Heart sound monitoring is a popular approach for early detection of CVDs. The heart sound can indicate some heart valve diseases because it is produced when the heart valves close. In fact, many studies for automated classification and analysis of heart sounds have been reported [4, 5]. Therefore, there have also been many studies to measure heart sounds using wearable devices [6, 7], and some of them have successfully achieved heart sound monitoring. However, these devices are not yet widely used in the world. One of the reasons for this is that these wearable devices have disadvantages such as discomfort due to constant wear and the need for intermittent charging. In other words, a novel heart sound monitoring method that does not impose any burden at all on the user is required.

Therefore, we have attempted to measure the heart sound without contact. This approach has advantages for daily continuous health monitoring by eliminating the need for physical constraints, wearing discomfort, batteries, and dedicated devices compared with the existing methods. However, the non-contact heart sound measurement is very difficult. The non-contact measured sound is very noisy, and the heart sound is buried in the noise. This is because the power of the airborne heart sound is lower than other ambient noises. On the other hand, a narrow bandwidth filter cannot be used because it also removes some components of the heart sound or murmur. Therefore, another way to extract the heart sound is required. So far, we have reported the method to extract the airborne heart sound using the array signal processing [8]. However, another problem, the impossibility of determining if the output sound is a heart sound or noise, appeared.

To overcome this problem, we propose a novel non-contact approach using a microphone and camera fusion. Specifically, we attempt to extract the heart sound from a non-contact measured sound by using signal averaging with a remote photoplethysmography (rPPG) as a trigger. In this study, we aim to evaluate the feasibility of our approach.

Shun Muramatsu is with the Dept. Precision Engineering, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan (corresponding author to provide email: smuramatsu@g.ecc.u-tokyo.ac.jp).

Kosuke Kurihara is with the Dept. Electrical Engineering, Faculty of Engineering, Tokyo University of Science, 6-3-1 Niijuku, Katsushika-ku, Tokyo, Japan (email: kuriharak@rs.tus.ac.jp).

II. Materials and Methods

A. Our Proposed Method

The outline of our method is described in Fig. 1. As an overall flow, the heart sound is extracted from the non-contact measured sound by using signal averaging with the rPPG as a trigger. Signal averaging by the electrocardiogram (ECG) has been reported to extract the weak heart sound from the sound measured by the stethoscope [9]. The heart sound is a mixture of sounds produced when the heart valves close and related blood flow sounds. The valves close as a result of their contraction by the ECG. Thus, the timing of the heart sound is synchronized with the ECG. This is why the signal averaging by the ECG as a trigger can extract the weak heart sound. The n-times averaging can improve the signal-to-noise ratio (SNR) by $n^{1/2}$ [10]. This technique can be effective for non-contact heart sound measurement. However, the ECG cannot be used for non-contact heart sound measurement because it cannot be measured in a non-contact manner. Therefore, we attempt to utilize the rPPG instead of the ECG.

The rPPG is the blood volume pulse associated with the ECG and can be measured from a distance by a camera. We have reported some methods for heart rate estimation by the rPPG [11, 12]. However, there are no studies to use the rPPG for a trigger of signal averaging. In this study, the rPPG is measured simultaneously with the sound. Then, the sound is divided into segments according to the rPPG peaks, and all segments are added together. In other words, the segments corresponding to each cycle of a heartbeat are all synchronized and averaged. This allows the extraction of the heartbeat-synchronized components; the first, second, third, and fourth heart sounds (S1, S2, S3, S4), and some murmurs.

B. Experimental Method

The experiment was conducted to measure the sound, ECG, and rPPG simultaneously. The location was a quiet conference room. The sound was measured by a microphone and preamplifier (MI-1235 and MI-3111, respectively, Ono Sokki Co., Ltd., Yokohama, Japan). To clarify the effect of distance, the measurement distance, a distance from the chest wall to the microphone, varied from 10 mm to 160 mm. During the measurement, the subject (the first author) was shirtless, sat motionless in a chair, and breathed slowly. After the sound was measured at a sampling frequency of 48 kHz, it was converted from analog to digital at a quantization resolution of 16-bit and saved with a data logger (DR-7100, Ono Sokki Co., Ltd., Yokohama, Japan). The measured sound was preprocessed using a bandpass filter (10 Hz to 1000 Hz) to extract the frequency range of the heart sound.

The ECG was acquired simultaneously to use as a reference for the appropriate trigger. It was measured using a 3-point induction method with an ECG monitor (AD8232 SparkFun Single Lead Heart Rate Monitor, SparkFun Electronics, Niwot, Colorado, USA), and saved in the same data logger as the sound. The measured ECG was preprocessed, including the hum noise reduction and smoothing, to amplify the R wave.

To acquire the rPPG signal, an RGB video of the subject's chest wall was also recorded simultaneously. In this experiment, we used the green signal because it contains more rPPG components than the red and blue channels [13]. The RGB video was captured using a machine vision camera (AD-130GE, JAI A/S, Denmark). This camera can record 8-bit color information (RGB) with 1296 × 966 resolution at 30 fps. The peak and full width at half maximum (FWHM) of the spectral sensitivity of the green sensor are 541 nm and 93 nm, respectively. We captured the upper body and manually selected the chest region as a region of interest. We set the size of the region of interest as 40 × 40 px. We obtain a time-series signal by spatially averaging the green component within the region of interest by each frame. We next perform bandpass filtering (0.7 Hz to 4 Hz) to extract the rPPG signal. The bandwidth is determined based on the normal heart range [14].

C. Evaluation Method

To evaluate the feasibility of our method, signal averaging by the ECG trigger was evaluated before by the rPPG trigger. In the evaluation of signal averaging by the ECG trigger, the effects of three parameters were clarified. The three parameters are the number of the averaging (n), the measurement distance, and the trigger accuracy. The trigger accuracy is the correctness of the time direction index of the trigger, and it was defined as the standard deviation (SD) of the R-R interval (RRI) estimation error. The trigger accuracy was varied by adding normal random numbers to the true time direction index of the ECG trigger so that the trigger has an arbitrary SD of RRI estimation error. The performance of signal averaging with the parameter varied was calculated, and the requirement for the trigger accuracy was clarified.

The performance of signal averaging was evaluated with the SNR expressed as

$$\text{SNR} = 10\log_{10}\frac{\sum_t \left(s[t]\right)^2}{\sum_t \left(y[t]-s[t]\right)^2}. \tag{1}$$

Here, $s[t]$ denotes a reference heart sound. In this study, the heart sound extracted by signal averaging with the most ideal conditions—the measurement distance is 10 mm, n (the number of the averaging) is 386, and the trigger is ECG—was used as $s[t]$. $y[t]$ denotes an extracted heart sound with each varied parameter.

Signal averaging with the rPPG trigger, our proposed method, was also conducted. Specifically, the trigger accuracy, SD of the RRI estimation error, was evaluated by comparing the reference RRI calculated by the ECG. Then, the heart sound extraction by signal averaging with the rPPG trigger was demonstrated.

III. Results and Discussion

A. ECG-Triggered Signal Averaging

Fig. 2 shows the measured ECG and sound simultaneously. It indicates that the ECG was successfully acquired enough to detect R waves. On the other hand, the measured sound is so noisy that the heart sound cannot appear although it was measured at 10 mm from the chest wall. This represents the difficulty of non-contact heart sound measurement due to the low power of the heart sound.
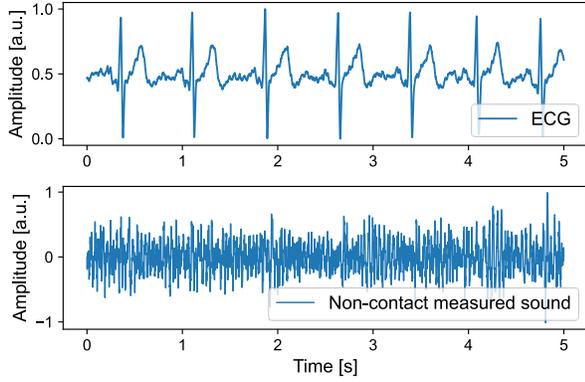
Fig. 2. ECG and sound measured simultaneously. The sound was measured at 10 mm from the chest wall.
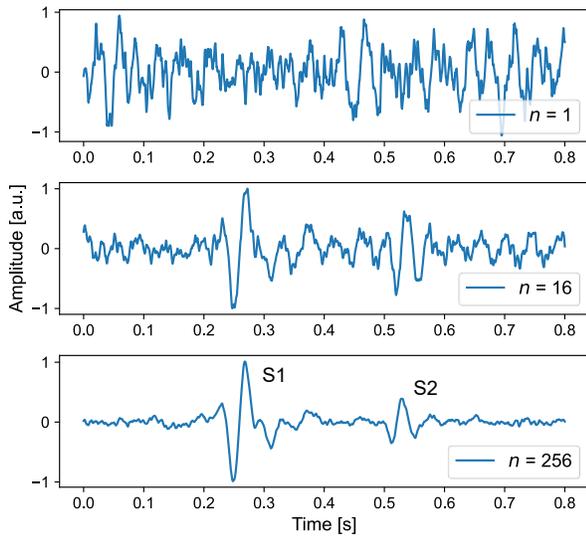

Fig. 3. Sounds generated by the ECG-triggered signal averaging. The numbers of averaging (n) are 1, 16, and 256, respectively.
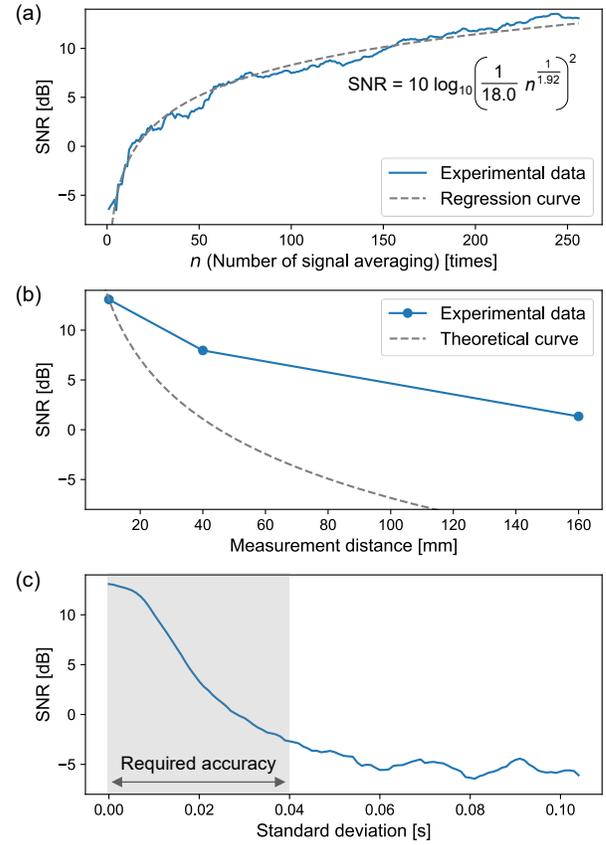

Fig. 4. Evaluation results of our approach. (a) SNR against n. (b) SNR against the measurement distance. (c) SNR against the SD of RRI estimation error.

Fig. 3 shows the sounds generated by the ECG-triggered signal averaging, and n, the number of averaging, is 1, 16, and 256, respectively. The sound ($n = 1$) is equivalent to the sound with only bandpass filtering as preprocessing, and this indicates the heart sound is buried in the noise. Even though, as n increases, the heart sound gradually appears. Specifically, the sound ($n = 16$) has S1 and S2, the main components of the heart sound, and the sound ($n = 256$) has more clear them. This indicates that the S1 and S2 occur following the peaks of the ECG and are emphasized by signal averaging. In other words, this figure indicates that our proposed method is effective for non-contact heart sound measurement.

Fig. 4 shows the SNRs of the extracted heart sounds with the varied parameters. Each parameter was varied from a default value (256, 10 mm, and 0 s for n, the measurement distance, and the trigger accuracy, respectively). Fig. 4 (a) shows the SNR against n. The blue solid line depicts the calculated SNR. It indicates that the SNR improves as n increases. The gray dashed line depicts the regression curve, and its equation is placed beside it. This equation indicates that the SNR is proportional to $n^{1/1.92}$, and this is close to $n^{1/2}$, the theoretical value of signal averaging [10]. Fig. 4 (b) shows the SNR against the measurement distance. The blue solid line depicts the experimental result, and the gray dashed line depicts the curve based on the theory of attenuation of sound pressure from a point source over distance. These indicate that the experimental result is not as attenuated as the theory. There are two possible reasons for this. The first reason is that the heart, the source of the heart sound, is considered closer to a plane source than to a point source. The sound from the plane source attenuates more slowly than the point source. The second reason is that some ambient noises are generated by the human body. The noise from the body attenuates with the measurement distance as well as the heart sound. These indicate that our approach can be applied to more distant measurements. Fig. 4 (c) shows the SNR against the SD of the RRI estimation error. The SNR when the SD is larger than approximately 0.04 s is stagnant. It indicates that signal averaging for the heart sound using a trigger with an error larger than approximately 0.04 s as the SD does not work. On the other hand, the SNR when the SD is less than approximately 0.04 s emerges from the stagnation, and this indicates that signal averaging works properly. In other words, the requirement for the trigger accuracy is approximately 0.04
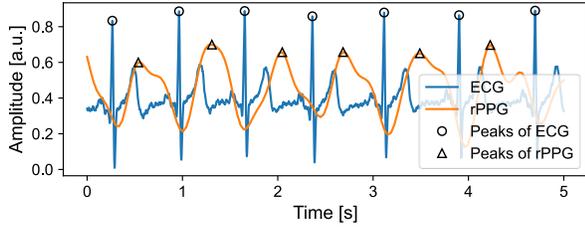
Fig. 5. ECG and rPPG signals measured simultaneously. Circles and triangles represent detected peaks.

Fig. 6. Error of RRI estimated by the rPPG, against the reference RRI.

s as the SD, and our method can extract the heart sound by using rPPG which satisfies this requirement.

B. rPPG-Triggered Signal Averaging

Fig. 5 shows an example of the ECG and rPPG measured simultaneously. It indicates that the rPPG was successfully acquired enough to detect the peaks. Specifically, the peaks of the ECG and rPPG appear alternately, and there are constant delays between the two. It looks accurate enough for heart rate variability estimation, a general application of rPPG.

Fig. 6 shows the error of RRI estimated by rPPG, against the reference RRI. The solid line depicts the mean of the errors, and the dashed lines depict the SD. The mean ± SD is 0.0021 ± 0.073 s. This value is considered reasonable accuracy for the application of heart rate estimation, however, it does not meet the requirement of the trigger accuracy (below approximately 0.04 s) for our method.

Fig. 7 shows the final demonstration result. It indicates that the rPPG-triggered signal averaging did not work because the SD was below the requirement. However, the results of this study indicate that our method is feasible by improving the accuracy of the rPPG measurement.

Fig. 7. Output sound by the rPPG-triggered signal averaging.

IV. CONCLUSION

The rPPG-triggered signal averaging, a novel approach for the non-contact heart sound measurement, was proposed, and its feasibility was evaluated. As a result of the experiment using the ECG as an appropriate trigger, our approach improved the SNR of the non-contact measured heart sound by $n^{1/1.92}$ for n, the number of averaging. The SNR did not decrease with a measurement distance as much as the theory suggests, because some of the ambient noises are generated from the human body. The accuracy required for the trigger was approximately 0.04 s as the SD of the RRI estimation error. As a result of the demonstration using the rPPG as triggers, the SD of RRI estimated by the rPPG was 0.073 s. Although the rPPG-triggered signal averaging did not work because this SD was below the requirement, it was indicated that our method is feasible by improving the accuracy of the rPPG measurement. Our method after this improvement will have advantages for daily continuous health monitoring by eliminating the user burdens compared with the existing methods.
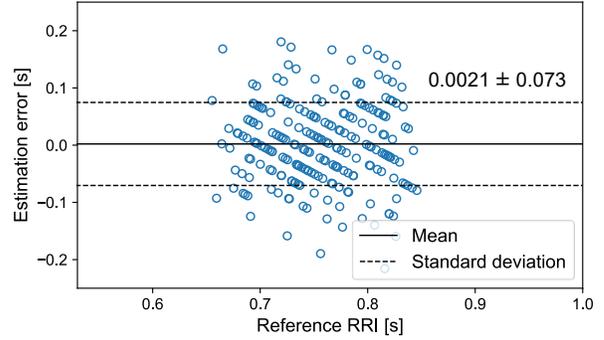
REFERENCES

[1] World Health Organization. "Cardiovascular diseases (CVDs)." https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (accessed Jan 19th, 2025).

[2] D. Mozaffarian *et al.*, "Heart disease and stroke statistics--2015 update: a report from the American Heart Association," *Circulation*, vol. 131, no. 4, pp. e29-e322, 2015.

[3] J. Hu *et al.*, "Portable microfluidic and smartphone-based devices for monitoring of cardiovascular diseases at the point of care," *Biotechnol. Adv.*, vol. 34, no. 3, pp. 305-320, 2016.

[4] A. K. Dwivedi, S. A. Imtiaz, and E. Rodriguez-Villegas, "Algorithms for Automatic Analysis and Classification of Heart Sounds-A Systematic Review," *IEEE Access*, vol. 7, pp. 8316-8345, 2019.

[5] H. Zhang, P. Zhang, Z. Wang, L. Chao, Y. Chen, and Q. Li, "Multi-Feature Decision Fusion Network for Heart Sound Abnormality Detection and Classification," *IEEE J. Biomed. Health Inform.*, vol. 28, no. 3, pp. 1386-1397, 2024.

[6] K. Lee *et al.*, "Mechano-acoustic sensing of physiological processes and body motions via a soft wireless device placed at the suprasternal notch," *Nat. Biomed. Eng.*, vol. 4, no. 2, pp. 148-158, 2020.

[7] A. Martin and J. Voix, "In-Ear Audio Wearable: Measurement of Heart and Breathing Rates for Health and Safety Monitoring," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 6, pp. 1256-1263, 2018.

[8] S. Muramatsu, M. Yamamoto, S. Takamatsu, and T. Itoh, "Non-Contact Heart Sound Measurement Using Independent Component Analysis," *IEEE Access*, vol. 10, pp. 98625-98632, 2022.

[9] T. Yanaga *et al.*, "Studies of the phonocardiogram by the signal averaging method: Quantitative analysis of the fourth heart sound," *Journal of Cardiology*, vol. 7, pp. 625-634, 1977.

[10] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, K. Michler, and S. Luo, "Comparing stress ECG enhancement algorithms," *IEEE Eng. Med. Biol.*, vol. 15, no. 3, pp. 37-44, 1996.

[11] K. Kurihara, D. Sugimura, and T. Hamamoto, "Non-Contact Heart Rate Estimation via Adaptive RGB/NIR Signal Fusion," *IEEE Trans. Image Process.*, vol. 30, pp. 6528-6543, 2021.

[12] K. Kurihara, Y. Maeda, D. Sugimura, and T. Hamamoto, "Spatio-Temporal Structure Extraction of Blood Volume Pulse Using Dynamic Mode Decomposition for Heart Rate Estimation," *IEEE Access*, vol. 11, pp. 59081-59096, 2023.

[13] F. C. M. Luis, P. Gonzalo, and S. Marija, "Optimal wavelength selection for noncontact reflection photoplethysmography," in *Proc. SPIE*, Nov 2011, vol. 8011, p. 801191.

[14] P. Palatini, "Need for a revision of the normal limits of resting heart rate," *Hypertension*, vol. 33, no. 2, pp. 622-625, 1999.