Estimation of respiratory rate in various environments using microphones embedded in face masks

Chhayly Lim[1] · Jungyeon Kim[2] · Jeongseok Kim[1] · Byeong-Gwon Kang[3] · Yunyoung Nam[4]

Abstract

Wearable health devices and respiratory rates (RRs) have drawn attention to the healthcare domain as it helps healthcare workers monitor patients' health status continuously and in a non-invasive manner. However, to monitor health status outside healthcare professional settings, the reliability of this wearable device needs to be evaluated in complex environments (i.e., public street and transportation). Therefore, this study proposes a method to estimate RR from breathing sounds recorded by a microphone placed inside three types of masks: surgical, a respirator mask (Korean Filter 94), and reusable masks. The Welch periodogram method was used to estimate the power spectral density of the breathing signals to measure the RR. We evaluated the proposed method by collecting data from 10 healthy participants in four different environments: indoor (office) and outdoor (public street, public bus, and subway). The results obtained errors as low as 0% for accuracy and repeatability in most cases. This research demonstrated that the acoustic-based method could be employed as a wearable device to monitor RR continuously, even outside the hospital environment.

Keywords Respiratory rate · Power spectral density · Microphone · Face mask · Wearable health device

1 Introduction

Respiratory rate (RR) is a vital sign that plays an important role in healthcare institutions for identifying various abnormal events including cardiac arrest, hypoxia, hyper-capnia, metabolic, and respiratory acidosis [1–4]. Furthermore, evidence shows that RR, along with other vital signs helps clinics lower fatality rates of

✉ Yunyoung Nam
   ynam@sch.ac.kr

Extended author information available on the last page of the article

coronavirus 2019 (COVID-19) patients by detecting critically ill patients with RR over 30 breaths/min at resting state [5, 6]. However, RR remains neglected in clinical records since most clinics continue to rely on trained staff to measure it by manually counting chest wall movements, measuring breath sounds through the stethoscope, and self-reports by posing a question as follow: "Is your breathing faster, slower, or the same as normal?". These methods are unreliable and time-consuming [7, 8]. Conversely, a few sensors showed the potential for an automated RR assessment, which depended on respiratory airflow, including differential flowmeters, turbine flowmeters, and hot wire anemometers. However, these instruments were reported to be inconvenient for patients and costly for general public use [9, 10].

Wearable health devices have become popular as they help people track their vital signs continuously and non-invasively, even outside the hospital. Moreover, they can provide health status data for healthcare workers to make early diagnoses and treatments [11]. Therefore, numerous studies have investigated various sensors that can be used to extract RR, including photoplethysmogram (PPG), electrocardiogram (ECG), piezoresistive pulse transducer (PZO), accelerometer, camera (vision-based), and microphone (acoustic-based). Previous studies [12–18] showed that RR can be extracted from PPG and ECG signals by a physiological synchronization of heart rate (HR) variation and respiration system, which is called respiratory sinus arrhythmia, through HR increase during inspiration and decrease during expiration. The respiratory signal can be obtained from PPG and ECG signals by either frequency modulation or amplitude modulation. Moreover, nowadays people carrying a smart device are exploding exponentially, this method could easily implement on either a smartphone or smartwatch to estimate the RR without any external device needed. However, such methods perform well when subjects remain still. Otherwise, the quality of RR estimation is negatively affected by noises associated with motion artifacts and baseline drifting noises [10, 17]. The accelerometer approach was depended on measure the chest wall movement. This approach is sensitive to movement artifacts which allows the measurement of the RR in minor movements. Nonetheless, the signal is easily affected by large movement artifacts, which makes it difficult to extract respiratory signals from the accelerometer signal [10, 17]. Meanwhile, motion artifact also can be of an issue for PZO, it requires close contact with the skin to prevent a large baseline drift. Thus, the wearable PZO bands must be tight with the user's body to achieve good results, perhaps making the user feel uncomfortable [10, 17]. A vision-based method was proposed to extract RR from the image based on tracking the thorax, face, or neck movement. Nonetheless, this non-contact method needs to be implemented in a static position, such as a bedroom with a fixed camera position. Therefore, this approach could be implemented in a clinical setting to monitor the RR of multiple patients at once. However, this type of approach is sensitive to lighting, the influence of a thick blanket, and movement of target people, which can have negative impacts [19, 20]. On the other hand, the acoustic-based approach which used breathing sound signals to estimate RR has shown more accurate estimation even during movement and sitting far away from the recording device [21, 22]. Nonetheless, this approach is more suitable and comfortable as a daily wearable device than other sensors such as PZO, PPG, and ECG

which need to be attached to the skin directly. However, this approach needs further investigation in different environments.

Recent studies have shown that wearing face masks could help prevent the spread of COVID-19, and researchers have suggested that people wear masks whenever they go outside or meet at a crowded place [23, 24]. However, different types of masks have different effectiveness. The KF94 mask, equivalent to the N95 mask, effectively prevent airborne infection through droplets [25, 26]. KF-94 is a class of face masks used in the Republic of Korea. KF stands for Korean Filter whereas 94 refers to the filtering percentage of particles indicting that KF-94 face masks are designed to filter 94% of particles, instead of 95% (N-95). Nonetheless, owing to the shortage of mask production, researchers advise to use alternative masks, such as surgical masks and reusable fabric mask, to slow down the spread of the virus [27, 28]. In addition, even fully vaccinated people in South Korea and the USA no longer need to wear masks in a public place, there are exceptions where they still require masks, such as crowded places, public transportation, and airplanes [29, 30]. Therefore, an opportunity arises for embedding a sensor in a face mask that can continuously monitor vital signs (i.e., respiration). It is recommended that people wear face masks in crowded area even when they are outdoors. However, most previous studies conducted experiments under structured indoor environments (i.e., office rooms and intensive care units), while few studies have been conducted in outdoor environments (i.e., sports activities) [9, 10]. To the best of our knowledge, no studies have investigated interactions between mask types and surroundings (i.e., public streets and transportation) even though different structure or materials of such masks may influence RR estimation. Therefore, in this study, we describe the development of three different types of face masks embedding a microphone for RR estimation and report the RR estimation performance on the masks under three different environmental conditions, such as indoor environments (sitting in an office) and outdoor environments (walking on a public street, taking a public bus, and subway).

The remainder of this paper is organized as follows: Sect. 2 provides the materials and methods of study. The experimental results are presented in Sect. 3 In Sect. 4, we discuss the finding. Finally, we conclude the paper in Sect. 5.

2 Materials and methods

2.1 Prototype of the mask

The proposed device was developed using a low-cost ESP-32 microcontroller (MCU) from Espressif system Co., LTD [31]. The MCU contains two central processing unit cores that can perform at based frequency of 80 MHz and boost up to 240 MHz. The SPH0645LM4H (Knowles Elec-tronic, LLC) microelectromechanical system (MEMS) integrated interchip-sound (I2S) microphone breakout board from Adafruit was used for audio recording [32]. The MCU receives an audio signal from the MEMS microphone via the I2S protocol. Subsequently, the data signal is converted to waveform audio file format and stored in a secure digital card with a
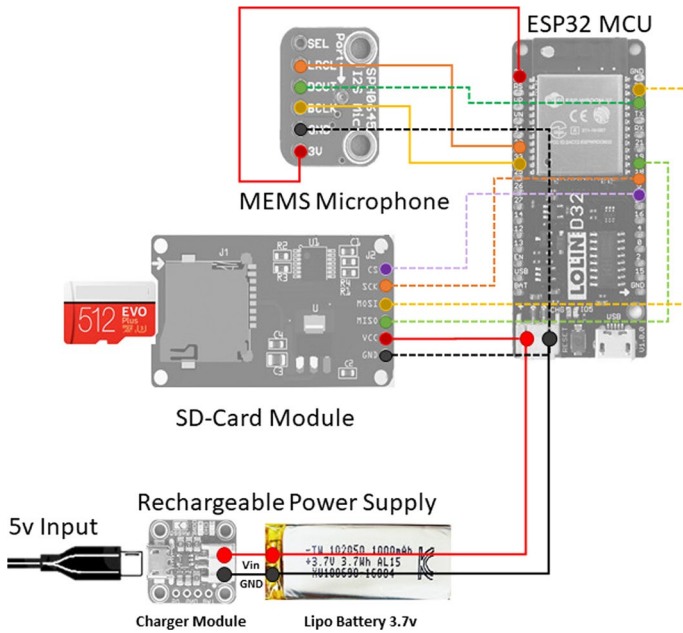
Fig. 1 Configuration circuits of the developed device

sampling rate of 44.1 kHz. Figure 1 shows the overall circuit configuration of the proposed device. A smartphone application was used as a remote for the proposed device. It can start and stop recording, and trigger a beeping sound, as explained in the next section.

2.2 Data acquisition

In this study, ten volunteers were enrolled as participants (five men, five women) with ages ranging from 22 to 29. Figure 2 shows the overall experimental protocols
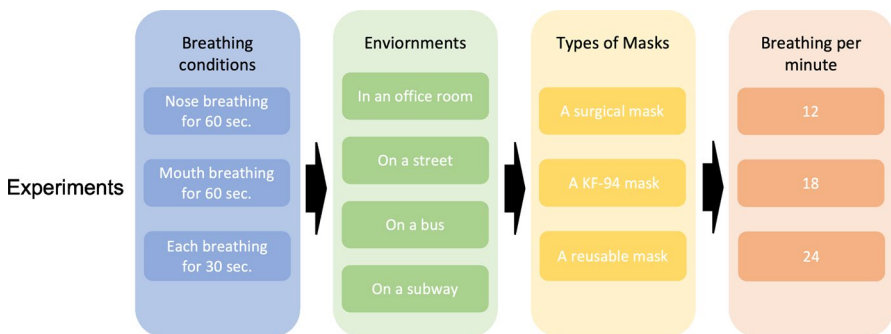

Fig. 2 Experiment design

(a) Surgical　　　　　　　**(b)** KF94　　　　　　　**(c)** Reusable

Fig. 3 Three types of masks used for the experiment

and three types of masks (surgical, KF94, and reusable fabric masks) were used for each experiment (see Fig. 3). The participants were trained to breathe according to the given beeping sound under different breathing conditions before the experiments. Once the training session of each participant was completed, they enrolled experiments. Each participant was asked to wear three different types of face masks and perform three types of breathing for 60 seconds (nose breathing, mouth breathing, and each type of breathing for 30 seconds in four different environments: in an office, on a street, on a bus, and on a subway. The participants were instructed to perform controlled breathing at a frequency ranging from 0.2 to 0.4 Hz with an increase of 0.1 Hz with each breath, which corresponds to 12 to 24 breaths/min with the step of 6 breaths. The data were collected for 1-minute and the participants were given a break until they felt comfortable starting the next experiment.

Before the experiment began, the participants were asked to wear earphones and listen to metronome beeping sounds produced by the smartphone application at a given frequency. Furthermore, each participant was instructed to inhale at each beeping sound and exhale before the next beeping sound occurred. Participants were advised to give instant feedback during the experiment if they failed to follow instructions, such as breathing out of sync with the beeping sound, so that the experiment could begin all over again. Every participant was trained to breath according to given breathing interval as operated by beeping sound before enrolling experiments. Breathing rate (BR) tested in the experiments were within the normal range of RR in healthy adults and would not cause any harm to the participants in this study [33, 34]. The experiment protocols including the three different breathing rates, which do not cause the symptoms of hyperventilation, were screened and approved by the Institute Review Board on Human Subject Research and Ethics Committees of Soonchunhyang University (Approval No. 1040875-202102-SB-019).

2.3 Data preprocessing and analysis

Figure 4 shows the work-flow of the proposed method. The breathing audio was collected at a 44.1 kHz sampling rate with 32-bit per sample. The data consisted of various background noises, including sounds from crowded areas, transportations sound, and random noise. DC component removal with bandpass filter within the
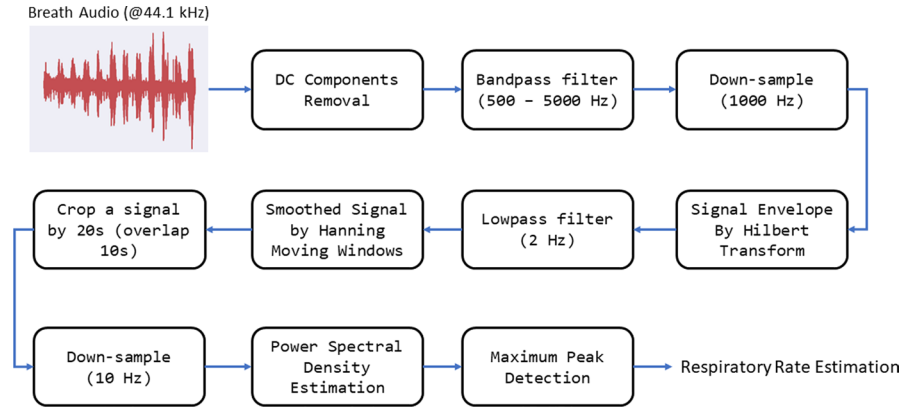
Breath Audio (@44.1 kHz)

Fig. 4 Overall experimental flowchart procedures

range of 500-5000 Hz was applied to eliminate noise from the original signal. Afterward, to reduce computational time complexity, the filtered signal was down-sampled to 1 kHz. The signal envelope of the breathing sound was extracted using the Hilbert transform which defined as follows:

$$H(u(t)) = \frac{1}{\pi} p \int_{-\infty}^{\infty} \frac{u(\tau)}{t - \tau} d\tau \tag{1}$$

where p denotes the Cauchy principal value, and u(t) is a continuous-time signal. Then, a low-pass filter at 2 Hz with a Hanning moving window was applied to smoothen the signal envelope [22, 35]. Near-real-time estimation was obtained by cropping the 1-min smoothed envelope signal into 20 s intervals with overlapped of 10 s and then down-sampling to 10 Hz. In total, five segments with a 1-min smoothed envelope signal were obtained.

In this study, the Welch-periodogram method was used for power spectral density (PSD) estimation from the breathing signal envelope. The frequency below 0.15 Hz was excluded in this study. Previous researches suggests that a frequency lower than 0.15 Hz is likely to be a sympathetic tone that can be active in a frequency ranging from 0.04 to 0.15 Hz. For example, if breathing frequency 0.2 Hz consists of a sympathetic tone, then the largest peak of PSD could show up within range between 0.04 and 0.15 Hz instead of the actual breathing frequency [12]. Raw audio signals were preprocessed before estimating RR, and the resulting signals of each preprocessing procedure are presented in Fig. 5. A segment of raw signals, a smoothed signal envelop and power spectral density transformed from the segment are shown as shown in Fig. 6. Estimation of RR was achieved by finding the largest PSD peak obtained from segments of 20 seconds preprocessed signals.
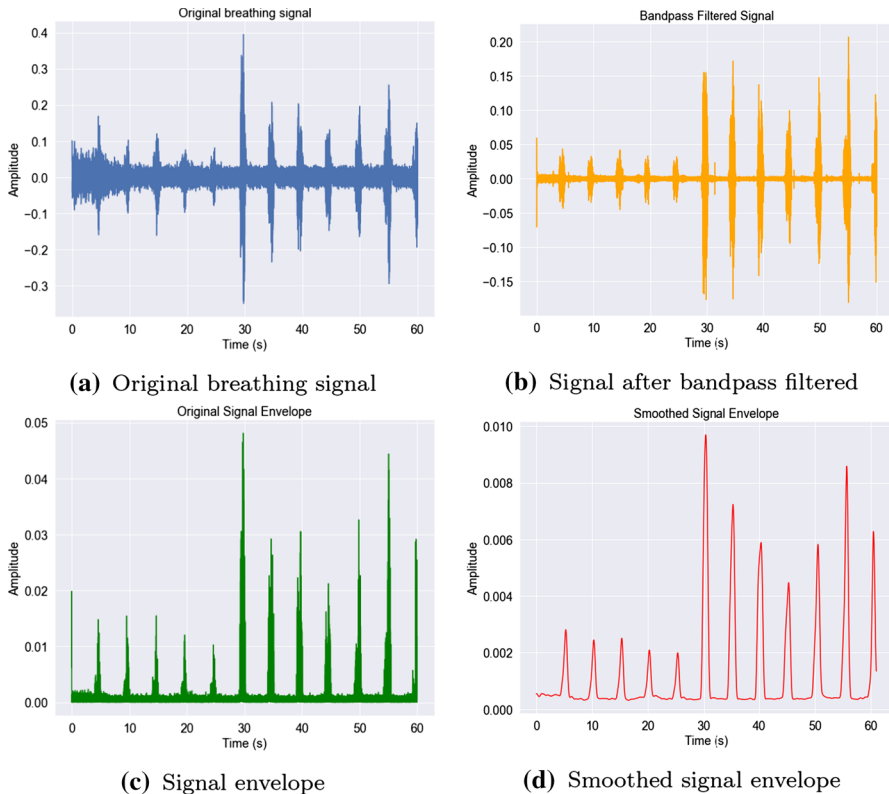
(a) Original breathing signal

(b) Signal after bandpass filtered

(c) Signal envelope

(d) Smoothed signal envelope

Fig. 5 Examples of preprocessed signals from breathing at 12 breaths/min during taking a bus

2.4 Evaluation metric

The accuracy and repeatability of the methods were measured by calculating the percentage error of estimated RR from the 20 s cropped signal as follows:

$$\%E = \frac{R_{est} - R_{act}}{R_{act}} \times 100 \tag{2}$$

where R_{est} and R_{act} represent the estimated and actual RRs, respectively. Further, we observed that the errors of five estimations had a non-normal distribution. Not only are median and interquartile range preferred measures when there are outliers in the dataset, but also they are frequently used when reporting respiration rate estimation in previous studies. Therefore, median and interquartile range (IQR), calculated as the difference between the 25th and 75th percentiles, were quantified as accuracy and repeatability instead of mean and standard deviation [12, 13]. Not only are median and interquartile range preferred measures when there are outliers in the dataset, but also they are frequently used measures when reporting respiration rate estimation in previous studies. Therefore, errors of median and interquartile ranges
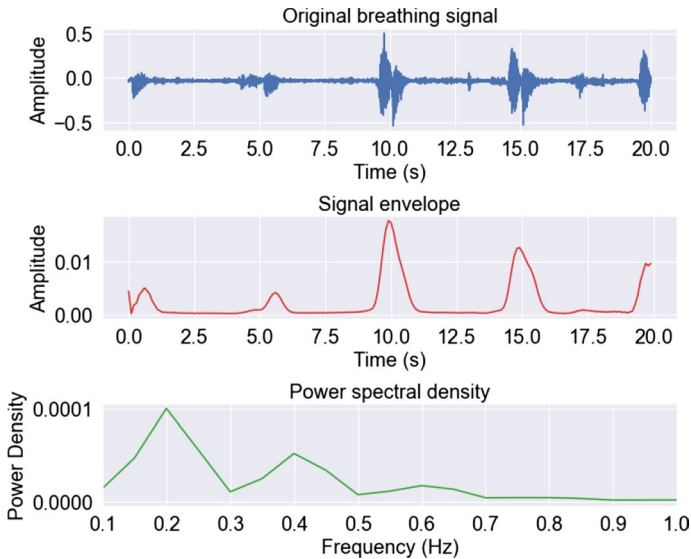
Fig. 6 Example of PSD estimation from a breathing signal envelope at 0.2 Hz (12 breaths/min) that contained both breathing through a nose and mouth signals

are reported in this study. The estimation errors in accuracy and repeatability closer to zero indicate a better performance of the method. Furthermore, to perform statistically significant testing of masks and environments in the median of accuracy and repeatability, Kruskal-Wallis and Mann-Whitney tests were used. If there were any significant differences ($\alpha = 0.05$) found in the Kruskal-Wallis test, pair-wise comparisons were conducted using Mann-Whitney test was done.

3 Results

3.1 Experiments under nose breathing condition

Table 1 summarizes the results of accuracy, which represent median estimation errors by comparing three types of masks in different environments. The results yielded a 0% error for median and IQR of accuracy for every breathing frequency ranging from 0.2 to 0.4 Hz (12 to 24 breaths/min). There was no significant difference ($p > 0.05$) in accuracy between the three types of masks for the same breathing rate and environment. Similarly, there was no significant difference ($p > 0.05$) in the accuracy among the four environments for the same type of mask and breathing rate.

A summary of the results of repeatability, which represent the IQR estimation errors, is shown in Table 2. Most of the results in the median and IQR of repeatability were yielded as low as 0%. However, the IQR of repeatability for a frequency of 0.4 Hz (24 breaths/min) when using a reusable fabric mask on the bus increased to

Table 1 The median and IQR of accuracy using three types of masks in four different environments (Nose breathing condition)

Environment	Beep rate (Hz)	Accuracy error (%)	Types of masks			p value
			Surgical	KF94	Reusable	
Indoor	0.2	Median	0	0	0	1.0
		IQR	0	0	0	
	0.3	Median	0	0	0	1.0
		IQR	0	0	0	
	0.4	Median	0	0	0	1.0
		IQR	0	0	0	
Street	0.2	Median	0	0	0	1.0
		IQR	0	0	0	
	0.3	Median	0	0	0	1.0
		IQR	0	0	0	
	0.4	Median	0	0	0	1.0
		IQR	0	0	0	
Bus	0.2	Median	0	0	0	0.59
		IQR	0	0	0	
	0.3	Median	0	0	0	1.0
		IQR	0	0	0	
	0.4	Median	0	0	0	0.36
		IQR	0	0	0	
Train	0.2	Median	0	0	0	0.23
		IQR	0	0	0	
	0.3	Median	0	0	0	0.1
		IQR	0	0	0	
	0.4	Median	0	0	0	0.32
		IQR	0	0	0	

18.75%. There was no significant difference (p > 0.05) in repeatability between the three types of masks for the same breathing rate and environment. Similarly, there was no significant ($p > 0.05$) in the repeatability among the four environments for the same type of mask and breathing rate.

3.2 Experiments under mouth breathing condition

Table 3 summarizes the results of accuracy by comparing the three types of masks in different environments. The overall results yielded a 0% error for median and IQR of accuracy in every frequency ranging from 0.2 to 0.4 Hz (12 to 24 breaths/min). There was no significant difference ($p > 0.05$) in accuracy between three types of masks for the same frequency and environment. Furthermore, no significant difference ($p > 0.05$) in accuracy by comparing the four different environment results for the same mask type and breathing frequency.

Table 2 The median and IQR of repeatability using three types of masks in four different environments (Nose breathing condition)

Environment	Beep rate (Hz)	Repeatability error (%)	Types of masks			p value
			Surgical	KF94	Reusable	
Indoor	0.2	Median	0	0	0	1.0
		IQR	0	0	0	
	0.3	Median	0	0	0	0.59
		IQR	0	0	0	
	0.4	Median	0	0	0	1.0
		IQR	0	0	0	
Street	0.2	Median	0	0	0	0.36
		IQR	0	0	0	
	0.3	Median	0	0	0	1.0
		IQR	0	0	0	
	0.4	Median	0	0	0	1.0
		IQR	0	0	0	
Bus	0.2	Median	0	0	0	0.81
		IQR	0	0	0	
	0.3	Median	0	0	0	0.36
		IQR	0	0	0	
	0.4	Median	0	0	0	0.18
		IQR	0	0	18.75	
Train	0.2	Median	0	0	0	0.83
		IQR	0	0	0	
	0.3	Median	0	0	0	0.31
		IQR	0	0	0	
	0.4	Median	0	0	0	0.71
		IQR	0	0	0	

A summary of the results of repeatability is shown in Table 4. Most of the median and IQR of repeatability obtained results as low as 0%. However, the IQR of repeatability increased to an error of 18.75% at a frequency of 0.2 Hz (12 breaths/min) when using a surgical mask on the train. There was no significant difference ($p > 0.05$) in repeatability between three types of masks for the same frequency and environment. Further, no significant difference ($p > 0.05$) in repeatability for the same mask type and breathing frequency.

3.3 Experiments under nose/mouth breathing condition

Table 5 summarized the results of accuracy by comparing three types of masks in different environments. Most of results yielded as low as 0% error for median and IQR of accuracy in every frequency ranging from 0.2 to 0.4 Hz (12 to 24 breaths/min). However, the IQR of the accuracy was increased to an error of 9.37% when

Table 3 The median and IQR of accuracy using three types of masks in four different environments (Mouth breathing condition)

Environment	Beep rate (Hz)	Accuracy error (%)	Types of masks			p value
			Surgical	KF94	Reusable	
Indoor	0.2	Median	0	0	0	1.0
		IQR	0	0	0	
	0.3	Median	0	0	0	1.0
		IQR	0	0	0	
	0.4	Median	0	0	0	1.0
		IQR	0	0	0	
Street	0.2	Median	0	0	0	1.0
		IQR	0	0	0	
	0.3	Median	0	0	0	1.0
		IQR	0	0	0	
	0.4	Median	0	0	0	1.0
		IQR	0	0	0	
Bus	0.2	Median	0	0	0	0.36
		IQR	0	0	0	
	0.3	Median	0	0	0	1.0
		IQR	0	0	0	
	0.4	Median	0	0	0	0.59
		IQR	0	0	0	
Train	0.2	Median	0	0	0	1.0
		IQR	0	0	0	
	0.3	Median	0	0	0	0.36
		IQR	0	0	0	
	0.4	Median	0	0	0	0.23
		IQR	0	0	0	

using reusable fabric mask at 0.4 Hz (24 breaths/min) in the bus environment. There was no significant difference ($p > 0.05$) in accuracy between three types of masks for the same frequency and environment. Further, no significant difference ($p > 0.05$) in accuracy by comparing the four different environment results for the same mask type and breathing frequency.

A summary of the results of repeatability is shown in Table 6. The results achieved a 0% error for the median repeatability for all experimental cases. However, the IQR of repeatability was found to increase in the bus and train environments. For the experiment that used a surgical mask on the bus, the errors yielded 25%, 16.66%, and 9.37% for a breathing frequency of 0.2, 0.3, and 0.4 Hz (12, 18, and 24 breaths/min), respectively. For the experiment that used a reusable fabric mask on the bus, the errors increased to 43.75%, 12.49%, and 12.5% for breathing frequencies of 0.2, 0.3, and 0.4 Hz (12, 18, and 24 breaths/min), respectively. Further, the experiment on the train that used a reusable mask yielded an error of

Table 4 The median and IQR of repeatability using three types of masks in four different environments (Mouth breathing condition)

Environment	Beep rate (Hz)	Accuracy error (%)	Types of masks			p value
			Surgical	KF94	Reusable	
Indoor	0.2	Median	0	0	0	1.0
		IQR	0	0	0	
	0.3	Median	0	0	0	1.01
		IQR	0	0	0	
	0.4	Median	0	0	0	0.36
		IQR	0	0	0	
Street	0.2	Median	0	0	0	1.0
		IQR	0	0	0	
	0.3	Median	0	0	0	1.0
		IQR	0	0	0	
	0.4	Median	0	0	0	1.0
		IQR	0	0	0	
Bus	0.2	Median	0	0	0	0.36
		IQR	0	0	0	
	0.3	Median	0	0	0	0.59
		IQR	0	0	0	
	0.4	Median	0	0	0	0.33
		IQR	0	0	0	
Train	0.2	Median	0	0	0	0.36
		IQR	18.75	0	0	
	0.3	Median	0	0	0	0.59
		IQR	0	0	0	
	0.4	Median	0	0	0	0.59
		IQR	0	0	0	

12.49% for a breathing frequency 0.3 Hz (18 breaths/min). There was no significant difference ($p > 0.05$) in repeatability for three types of masks for the same breathing frequency and environment. Nonetheless, significant difference ($p < 0.05$) was found in repeatability at a breathing frequency 0.2 Hz (12 breaths/min) for indoor vs. bus, and street vs. bus.

4 Discussion

This study investigated the possibility of estimating RR in outdoor environments by utilizing a proposed microphone device placed inside a face mask. Many previous researchers have investigated RR estimation based on data collected from various sensors, including PPG, ECG, camera, PZO, accelerometer, and microphone [10, 22]. However, these studies examined the methods by conducting experiments in an

Table 5 The median and IQR of accuracy using three types of masks in four different environments (Nose/Mouth breathing condition)

Environment	Beep rate (Hz)	Accuracy error (%)	Types of masks			p value
			Surgical	KF94	Reusable	
Indoor	0.2	Median	0	0	0	1.0
		IQR	0	0	0	
	0.3	Median	0	0	0	0.36
		IQR	0	0	0	
	0.4	Median	0	0	0	0.59
		IQR	0	0	0	
Street	0.2	Median	0	0	0	0.36
		IQR	0	0	0	
	0.3	Median	0	0	0	1.0
		IQR	0	0	0	
	0.4	Median	0	0	0	1.0
		IQR	0	0	0	
Bus	0.2	Median	0	0	0	0.75
		IQR	0	0	0	
	0.3	Median	0	0	0	0.75
		IQR	0	0	0	
	0.4	Median	0	0	0	0.96
		IQR	0	0	9.37	
Train	0.2	Median	0	0	0	0.8
		IQR	0	0	0	
	0.3	Median	0	0	0	1.0
		IQR	0	0	0	
	0.4	Median	0	0	0	0.77
		IQR	0	0	0	

indoor environment and static positions (i.e., office room in a supine position). This type of approach is well suited for hospital settings. Therefore, the purpose of this proposed device and method is to allow people to continue monitoring their health status even outside the hospital.

Accuracy and repeatability of RR estimation obtained from experiments under three different breathing conditions suggest that the proposed method effectively measures the RR in many cases. As described in [9, 22], acoustic-based methods performed extremely well in an indoor environment, which correlates to the results of the experiments in an office. In addition, we also found that the results of the experiment in an outdoor open area (walking on a public street) and indoors (sitting in an office room) achieved remarkably similar outcomes, and obtained 0% errors for every experiment case. Moreover, the results showed that the developed device performed extremely well with the KF94 mask, obtaining 0% errors in accuracy and repeatability for every environment. This could be due to the thickness of the mask,

Table 6 The median and IQR of repeatability using three types of masks in four different environments (Nose/Mouth breathing condition)

Environment	Beep rate (Hz)	Repeatability error (%)	Types of masks			p value
			Surgical	KF94	Reusable	
Indoor	0.2	Median	0	0	0	1.0
		IQR	0	0	0	
	0.3	Median	0	0	0	1.0
		IQR	0	0	0	
	0.4	Median	0	0	0	0.36
		IQR	0	0	0	
Street	0.2	Median	0	0	0	0.36
		IQR	0	0	0	
	0.3	Median	0	0	0	0.36
		IQR	0	0	0	
	0.4	Median	0	0	0	0.36
		IQR	0	0	0	
Bus	0.2	Median	0	0	0	0.26
		IQR	25	0	43.75	
	0.3	Median	0	0	0	0.26
		IQR	16.66	0	12.49	
	0.4	Median	0	0	0	0.4
		IQR	9.37	0	12.5	
Train	0.2	Median	0	0	0	0.26
		IQR	0	0	0	
	0.3	Median	0	0	0	0..26
		IQR	0	0	12.49	
	0.4	Median	0	0	0	0.40
		IQR	0	0	0	

which makes the airflow sound louder than the surrounding noises. However, there was no evidence of a significant difference ($p > 0.05$) found in the statistical testing between the three types of masks in experimental cases. According to the statistical tests of accuracy and repeatability, it is possible to conclude that influence of structure or materials on RR estimation achieved by detecting largest peaks in PSD estimation can be neglected. Nevertheless, this can be further investigated by recording the breathing airflow, which is beyond the scope of this study. Another benefit of our method is that the signal duration used for estimation was approximately 20 s, which is a real-time estimation and has been considered by many researchers [12, 18].

Even though the results from the experiments tended to achieve 0% errors for accuracy and repeatability, the IQR of repeatability for surgical and reusable masks was found to increase for every breathing condition we tested. There were three occasions where lower repeatability was observed. It was found that repeatability was lowered mostly in the experiments on a bus with reusable masks at 0.4Hz

breathing rate. Possible causes of this can be the nose breathing condition which is not as loud as the mouth breathing condition because lower repeatability was not observed with the experiments where involved with the mouth breathing condition or other types of masks. When manually inspecting audio recordings of breathing of the experiments , it was found that noises caused by the bus drove over road bumps that affected RR estimation, which may have interfered with nose breathing sounds In [9, 10], the authors mentioned that the acoustic could be susceptible to background noise and must be carefully investigated. Therefore, it is expected that future experiments by either our laboratory or others should consider recording the sound level from both breathing sounds and ambient sounds. Furthermore, in [22], the authors mentioned that using acoustic-based methods can achieve an RR estimation of up to 90 breaths/min. Thus, future research should consider breathing rates above the normal range in various environments so that it could indicate the abnormality in any situation. Finally, this study employed relatively small number of participants to estimate RR under three different breathing types in four different environments while wearing three frequently used types of masks. The result obtained from the experiments is difficult to generalize due to the small sample size. However, the errors of RR estimation observed in this study is little that it is reasonable to assume that the proposed method can be a robust method in the tested settings. Nevertheless, further investigation is required by recruiting more number of people.

5 Conclusion

In this study, we presented an accurate RR estimation method using breathing audio signals by placing a microphone inside three types of masks, including surgical, KF94, and reusable fabric masks. We evaluated the performance of the developed device and method under four different environmental conditions: indoor (sitting in an office) and outdoor (walking on a public street, taking a public bus, and subway). The results were significant and yielded as low as 0% errors of accuracy and repeatability in most cases. RR was measured using a Welch periodogram method to estimate the PSD of breathing signals. In addition, we believe that this research opens the door to new possibilities by taking the findings from this research and combining with various sensors to monitor important vital signs that could quickly indicate an abnormality, either in a public or a clinical setting.

References

1. Cretikos MA, Bellomo R, Hillman K, Chen J, Finfer S, Flabouris A (2008) Respiratory rate: the neglected vital sign. Med J Aust 188:657–659. https://doi.org/10.5694/j.1326-5377.2008.tb01825.x
2. Subbe C, Davies R, Williams E, Rutherford P, Gemmell L (2003) Effect of introducing the modified early warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions. Anaesthesia 58:797–802. https://doi.org/10.1046/j.1365-2044.2003.03258.x
3. Philip K, Richardson R, Cohen M (2013) Staff perceptions of respiratory rate measurement in a general hospital. Br J Nurs 22:570–574. https://doi.org/10.12968/bjon.2013.22.10.570
4. Rolfe S (2019) The importance of respiratory rate monitoring. Br J Nurs 28:504–508. https://doi.org/10.12968/bjon.2019.28.8.504
5. WHO (2021) COVID-19 Clinical management: living guidance. Available: https://www.who.int/publications/i/item/WHO-2019-nCoV-clinical-2021-1. Accessed 28 July 2021
6. Sun Q, Qiu H, Huang M, Yang Y (2020) Y. Lower mortality of COVID-19 by early recognition and intervention: experience from Jiangsu Province. Ann Intensive 10:1–4. https://doi.org/10.1186/s13613-020-00650-2
7. Loughlin PC, Sebat F, Kellett JG (2018) Respiratory rate: the forgotten vital sign-Make it count! Jt Comm J Qual Patient Saf 44:494–499. https://doi.org/10.1016/j.jcjq.2018.04.014
8. Greenhalgh T, Koh GCH, Car J (2020) Covid-19: a remote assessment in primary care. BMJ 368:m1182. https://doi.org/10.1136/bmj.m1182
9. Massaroni C, Nicolò A, Lo Presti D, Sacchetti M, Silvestri S, Schena E (2019) Contact-based methods for measuring respiratory rate. Sensors (Basel) 19:908. https://doi.org/10.3390/s19040908
10. Al-Khalidi FQ, Saatchi R, Burke D, Elphick H, Tan S (2011) Respiration rate monitoring methods: a review. Pediatr Pulmonol 46:523–529. https://doi.org/10.1002/ppul.21416
11. Khan Y, Ostfeld AE, Lochner CM, Pierre A, Arias AC (2016) Monitoring of vital signs with flexible and wearable medical devices. Adv Mater 28:4373–4395. https://doi.org/10.1002/adma.201504366
12. Chon KH, Dash S, Ju K (2009) Estimation of respiratory rate from photoplethysmogram data using time-frequency spectral estimation. IEEE Trans Biomed Eng 56:2054–2063. https://doi.org/10.1109/TBME.2009.2019766
13. Dash S, Shelley KH, Silverman DG, Chon KH (2010) Estimation of respiratory rate from ECG, photoplethysmogram, and piezoelectric pulse transducer signals: a comparative study of time-frequency methods. IEEE Trans Biomed Eng 57:1099–1107. https://doi.org/10.1109/TBME.2009.2038226
14. Hernando A, Peláez-Coca MD, Lozano MT, Lázaro J, Gil E (2019) Finger and forehead PPG signal comparison for respiratory rate estimation. Physiol Meas 40:905007. https://doi.org/10.1088/1361-6579/ab3be0
15. Karlen W, Raman S, Ansermino JM, Dumont GA (2013) Multiparameter respiratory rate estimation from the photoplethysmogram. IEEE Trans Biomed Eng 60:1946–1953. https://doi.org/10.1109/TBME.2013.2246160
16. Nam Y, Lee J, Chon KH (2013) Respiratory rate estimation from the built-in cameras of smartphones and tablets. Ann Biomed Eng 42:885–898. https://doi.org/10.1007/s10439-013-0944-x
17. Shen CL, Huang TH, Hsu PC, Ko YC, Chen FL, Wang WC, Kao T, Chan CT (2017) Respiratory rate estimation by using ECG, impedance, and motion sensing in smart clothing. J Med Biol Eng 37:826–842. https://doi.org/10.1007/s40846-017-0247-z
18. Fleming SG, Tarassenko L (2007) A comparison of signal processing techniques for the extraction of breathing rate from the photoplethysmogram. Int J Biol Med Sci 2:232–236
19. Li MH, Yadollahi A, Taati B (2014) A non-contact vision-based system for respiratory rate estimation. Annu Int Conf IEEE Eng Med Biol Soc 2014:2119–2122. https://doi.org/10.1109/EMBC.2014.6944035
20. Massaroni C, Lopes DS, Lo Presti D, Schena E, Silvestri S (2018) Contactless monitoring of breathing patterns and respiratory rate at the pit of the neck: a single camera approach. J Sens. https://doi.org/10.1155/2018/4567213
21. Gu F, Niu J, Das SK, He Z, Jin X (2017) Detecting breathing frequency and maintaining a proper running rhythm. Pervasive Mob Comput 42:498–512

22. Nam Y, Reyes BA, Chon KH (2016) Estimation of respiratory rates using the built-in microphone of a smartphone or headset. IEEE J Biomed Health Inform 20:1493–1501. https://doi.org/10.1109/JBHI.2015.2480838

23. Liu R, Fleisher LA (2020) Getting to a new normal: mandating that patients wear masks as hospitals fully reopen during the coronavirus pandemic. Anesthesiology 133:479–481. https://doi.org/10.1097/ALN.0000000000003386

24. Leung NHL, Chu DKW, Shiu EYC et al (2020) Respiratory virus shedding in exhaled breath and efficacy of face masks. Nat Med 26:676–680. https://doi.org/10.1038/s41591-020-0843-2

25. Rubio-Romero JC, Pardo-Ferreira MDC, Torrecilla-García JA, Calero-Castro S (2020) Disposable masks: disinfection and sterilization for reuse, and non-certified manufacturing, in the face of shortages during the COVID-19 pandemic. Saf Sci 129:104830. https://doi.org/10.1016/j.ssci.2020.104830

26. Kim MC, Bae S, Kim JY, Park SY, Lim JS, Sung M, Kim SH (2020) Effectiveness of surgical, KF94, and N95 respirator masks in blocking SARS-CoV-2: a controlled comparison in 7 patients. Infect Dis (Lond) 52:908–912. https://doi.org/10.1080/23744235.2020.1810858

27. Kirubarajan A, Khan S, Got T, Yau M, Bryan JM, Friedman SM (2020) Mask shortage during epidemics and pandemics: a scoping review of interventions to overcome limited supply. BMJ Open 10:e040547. https://doi.org/10.1136/bmjopen-2020-040547

28. Howard J, Huang A, Li Z, Tufekci Z et al (2021) An evidence review of face masks against COVID-19. Proc Natl Acad Sci U S A 118:e2014564118. https://doi.org/10.1073/pnas.2014564118

29. CDC (2021) COVID-19. Available: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/fully-vaccinated-guidance.html. Accessed 28 July 2021

30. CNA (2021) South Koreans no longer need masks outdoors if vaccinated against COVID-19. Available: https://www.channelnewsasia.com/news/asia/covid-19-malaysia-record-jul-28-17405-klang-valley-johor-mco-15317444. Accessed 28 July 2021

31. ESPREFFIF (2021) https://www.espressif.com/en/products/socs/esp32. Accessed 28 July

32. Adafruit shorturl.at/vCGHQ. Accessed 28 July 2021

33. Chourpiliadis C, Bhardwaj A (2019) Physiology, Respiratory Rate. In: StatPearls. StatPearls Publishing 2020

34. Lusardi MM (2012) Is walking speed a vital sign? Absolutely! Top Geriatr Rehabil 28:67–76. https://doi.org/10.1097/TGR.0b013e31824385a4

35. Sheen YT (2009) On the study of applying Morlet wavelet to the Hilbert transform for the envelope detection of bearing vibrations. Mech Syst Signal Pr 23:1518–1527. https://doi.org/10.1016/j.ymssp.2009.01.012

Authors and Affiliations

Chhayly Lim[1] · Jungyeon Kim[2] · Jeongseok Kim[1] · Byeong-Gwon Kang[3] · Yunyoung Nam[4]

Chhayly Lim
chaylylim@gmail.com

Jungyeon Kim
betterwithme@sch.ac.kr

Jeongseok Kim
vcfdregg3@sch.ac.kr

Byeong-Gwon Kang
bgkang@sch.ac.kr

[1] Department of ICT Convergence, Soonchunhyang University, Asan 31538, South Korea

2 ICT Convergence Research Center, Soonchunhyang University, Asan 31538, South Korea

3 Department of Information and Communication Engineering, Soonchunhyang University, Asan 31538, South Korea

4 Department of Computer Science and Engineering, Soonchunhyang University, Asan 31538, South Korea