Respiratory Rate Estimation from Tracheal Sound Analysis Using a Mobile Application for Smartphones

Alison Contreras-Rodríguez[1]([✉]) [iD], Nemecio Olvera-Montes[2] [iD],
Carlos Mariaca-Gaspar[2] [iD], Abel Mosco-Vargas[1] [iD], Victor Maya-Venegas[1] [iD],
Sonia Charleston-Villalobos[3] [iD], Tomas Aljama-Corrales[3] [iD], and Bersain Reyes[1] [iD]

[1] School of Sciences, Universidad Autónoma de San Luis Potosí, 78295 SLP, Mexico
A310209@alumnos.uaslp.mx

[2] Mechatronics, Mechanical and Industrial Engineering Division, Tecnológico de Estudios Superiores de Ecatepec, 55210 Ecatepec de Morelos, Mexico

[3] Electrical Engineering Department, Universidad Autónoma Metropolitana Iztapalapa, 09340 Mexico City, Mexico

Abstract. Diagnosing several lung diseases is challenging and usually requires various methods and tests, including a patient's clinical history, auscultation, spirometry, pulmonary function tests, and other methods using more specialized medical devices. For its part, the pulmonary auscultation with the mechanic stethoscope represents an early approach to the disease. However, it is highly subjective. Therefore, acquiring and analyzing respiratory sounds through mobile computerized devices, such as smartphones, has been an attractive alternative for the estimation of physiological parameters, including respiratory rate (RR). This study explored the estimation of RR performed completely on a single smartphone device, from the tracheal sound acquisition, signal conditioning and processing, and results report. To this end, a mobile application was developed for the Android system, and acquisitions were made in ten ($N = 10$) healthy volunteers while breathing at different metronome RR. The results obtained with the app were compared with the ones obtained from a respiratory reference signal. Mean absolute errors of 0.06, 0.18, 0.66 and 0.54 bpm were found for RR of 6, 12, 18 and 24 bpm, respectively. The promising results point out to test the mobile-developed system in breathing maneuvers that include temporal changes in RR.

Keywords: tracheal sounds · respiratory rate · smartphone

1 Introduction

Many years ago, the World Health Organization ranked cardiorespiratory diseases among the top ten causes of death worldwide, independently of the income level of the countries. A vital sign that allows their detection and monitoring is respiratory rate (RR). Among simple methods used to estimate RR are human observation and palpation, as well as auscultation using a mechanical or digital stethoscope [1]. The latter allows the respiratory acoustic signals to be amplified and filtered, which facilitates their hearing that sometimes cannot be achieved through the classical mechanical stethoscope. This study focused on respiratory tracheal sounds (TS) whose vibrations are originated in the trachea and upper lung and transmitted through the surrounding tissue up to the neck surface and its vicinity, where they are heard, analyzed, and related to physiological characteristics that represent health or disease. Therefore, TS and RR are relevant to learn about the respiratory system [2].

Computerized respiratory sound analysis systems (CORSA) have helped overcome traditional auscultation's limitations, making it possible to acquire, store, reproduce, analyze, and display information on various respiratory sounds, including TS. Nowadays, CORSA systems have begun to permeate clinical settings, and commercial medical devices currently exist, e.g., the MASIMO Rainbow equipment, which allows the estimation of RR using an acoustical approach based on TS [3]. Alternatively, efforts have recently been made to use smartphones to develop mobile CORSA systems (mCORSA), given the characteristics of ubiquity, mobility, and cost-effective sensors of these devices. This approach has allowed the implementation of digital signal processing algorithms on smartphones to facilitate the application of CORSA systems in different locations, without traveling to more specialized sites, and even to perform the acquisition and analysis of respiratory sounds directly at the patient's bedside [4].

Our research group has made some efforts regarding RR estimation from TS and the development of mobile applications (apps) for TS analysis. In a first effort, the estimation of RR was performed offline, i.e., the smartphone-acquired TS was analyzed on a regular computer, not in the mobile device itself [5]. In a posterior effort, an mCORSA system allowing the acquisition, processing, and display of results was developed to detect adventitious respiratory sounds [4]. In a recent effort, a mobile app was implemented to compute the Shannon entropy (SE), which quantifies the uncertainty of a stochastic signal of a previously uploaded file, sound or not, to the app [6]. It has been reported that the SE of TS provides a surrogate signal of the normalized respiratory airflow, providing the rationale for using them to estimate RR [7]. Hence, the development of an mCORSA system that allows RR estimation directly from TS acquired by the smartphone device was pending for our research group.

In this study, we addressed the end-to-end estimation of RR on a single smartphone using a tracheal sounds approach which, to the best of our knowledge, has not been addressed yet. We developed a mobile app to govern the acquisition of TS using an acoustical sensor attached to the smartphone. The app computes the RR based on the power spectral density (PSD) of the SE of the acquired TS. The results are displayed right on the mobile device. Tests were made with metronome breathings maneuvers while considering the RR derived from a piezoelectric respiratory band as a reference.

2   Methodology

2.1   Hardware

An acoustical sensor was employed to acquire TS, as described in our previous study [4]. Briefly, the sensor comprises a subminiature electret microphone (BT-2159000, Knowles Electronics, IL, USA) encapsulated in a plastic bell and connected to the 3.5 mm audio input of the smartphone. Regarding the smartphone, the Huawei Y6 2018 (Huawei, Shenzhen, China), which has 2.0 GB of RAM and runs an Android 8.0 operating system, was used. The mCORSA system was governed by the mobile app, whose design and implementation are described below.

2.2   Mobile App

The app was developed in Android Studio Flamingo 2022.2.1 (Google, CA, USA) using Java (Oracle Corp., CA, USA). The app oversaw the setting, starting, and stopping of the TS acquisition stage, as well as performed the digital signal processing required to estimate RR from the acquired sounds. Hence, the design and implementation were divided into two main parts: 1) the graphical user interface (GUI) and 2) the internal digital signal processing of the TS.
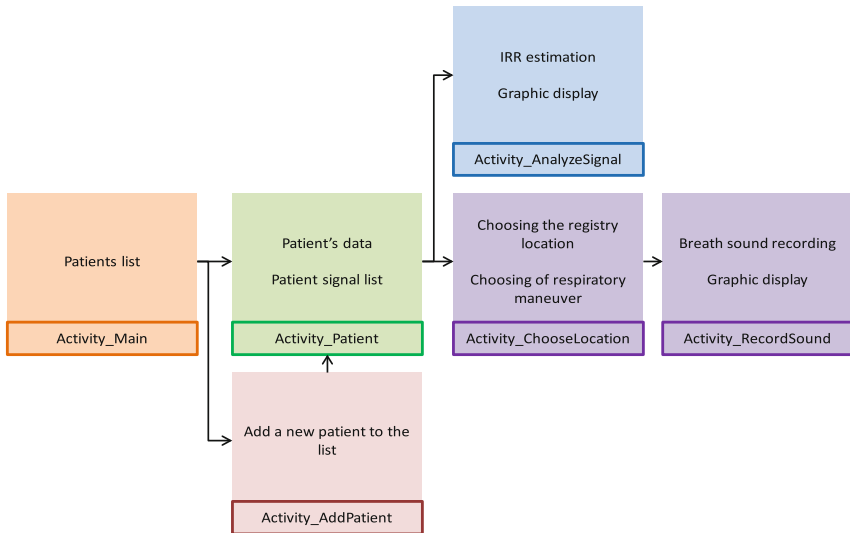
Fig. 1.   Main Flow of activities of the developed mobile app.

Several activities were implemented regarding the GUI, as shown in Fig. 1. First, the user can select a patient (volunteer) from an existing list in the main activity. In the add patient activity, the user is prompted to introduce the patient's data, including first and last name, age, and gender. Subsequently, the user can select among different breathing maneuvers and the location of the acquired signal. The neck location was always used

in this study, as only TS was acquired. Finally, the user can start recording the TS by clicking an acquisition signal button. After the acquisition is finalized and saved, the acquired data is displayed if not discarded. If the user selects an existing patient from the list, the app uploads its available data and recordings. Finally, the user can choose the "analyze signal" activity to perform the data analysis, resulting in the visualization of the TS, its normalized SE signal, its PSD, and the corresponding estimated RR.

TS is acquired using the aforementioned acoustic sensor, where the corresponding electrical signal is digitized at a sampling frequency of 10 kHz and 16 bits-per-sample, to comply with Nyquist sampling theorem and international CORSA guideline recommendations. The TS is bandpass filtered using a digital FIR filter with a Hamming window of 500 samples and 100 and 3000 Hz cutoff frequencies. The filtered TS are normalized between 0 and 1, resulting in the signal $x_1[n]$ of the block diagram in Fig. 2. The files in this stage correspond to the raw audio signal and the text file with the amplitudes obtained from the TS. The block diagram shown in Fig. 2 expands the "analyze signal" activity block from Fig. 1.
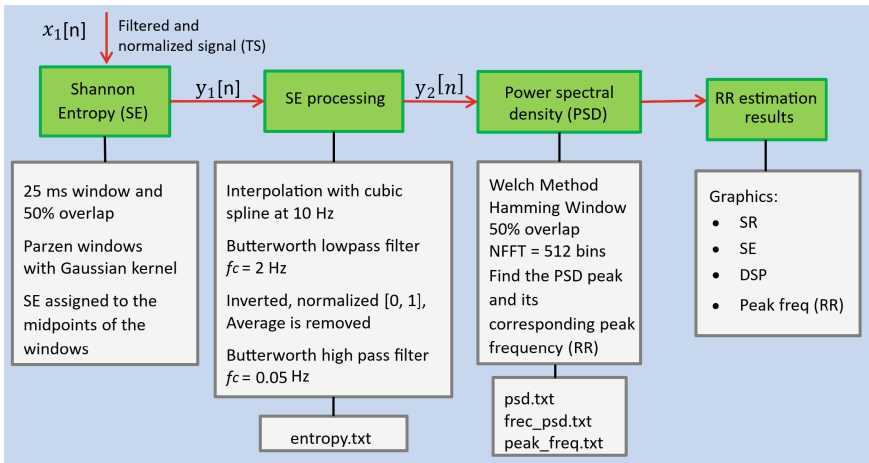
Fig. 2. Block diagram of the digital signal processing performed by the mobile app.

First, the SE signal of the acquired and preprocessed TS, $x_1[n]$, is estimated following the method described in our previous work [4, 5]. Briefly, TS is divided into successive windows of 25 ms with 50% overlap, and the probability density function (pdf) is calculated for each window using the Parzen window method with a Gaussian kernel. The SE value of each window is assigned to its midpoint, resulting in the time series $y_1[n]$ in Fig. 2. To achieve a uniform sampling frequency adequate for RR estimation accordingly to Nyquist sampling theorem, as well as to focus on the RR frequency range, $y_1[n]$ is interpolated using cubic splines at 10 Hz and filtered with an IIR Butterworth lowpass filter with a cutoff frequency of 2 Hz. To facilitate the subsequent signal processing, the SE signal is reversed, normalized between 0 and 1, its mean (DC component) is removed, and filtered again with an IIR Butterworth high pass filter with a cutoff frequency of 0.05 Hz, resulting in the time series $y_2[n]$.

After computing the processed SE signal, the RR is obtained via the frequency corresponding to the peak of the PSD of the SE signal. To this end, the PSD is estimated with the Welch periodogram method, using a Hamming window of 10 s, with 50% overlap between consecutive windows, and $NFFT = 512$ frequency bins.

Once all the calculations are completed, the TS signal, the SE signal, the PSD of the SE, and the corresponding RR are displayed in the mobile app. Finally, the related text files are stored for possible export to external programs like MATLAB or Python. It is worth mentioning that, in this study, the resulting text files with information about the SE signal, its PSD, and its corresponding peak frequency, i.e., estimated RR, were exported to MATLAB R2023a (The MathWorks, MA, USA) only for comparison purposes with the respiratory reference, i.e., no more signal processing was performed outside the developed mobile app.

2.3   Data Acquisition

Data from ten ($N = 10$) respiratory healthy volunteers were acquired, with ages ranging from 19 to 23 years old, seven women and three men. Before the acquisition, the experimental protocol was explained to the volunteers, and their informed consent to participate in the study was obtained according to the Declaration of Helsinki.

To acquire the TS, the acoustic sensor was placed on the lateral surface of the neck, as shown in Fig. 3, using a two-sided adhesive ring. Simultaneously, the respiratory reference and ECG signals were acquired using the Biosignal Plux wireless system (PLUX Wireless Biosignal, Lisboa, Portugal), using a sampling frequency equal to 1000 Hz. It is worth mentioning that the ECG signal was not used in this study, but it was acquired for its future use by our research group when developing and comparing different RR estimation methods. The respiratory reference signal was acquired using a piezoelectric respiratory band, while the ECG was acquired following a configuration that emulated DI by placing three adhesive Ag/AgCl electrodes on the thoracic area.
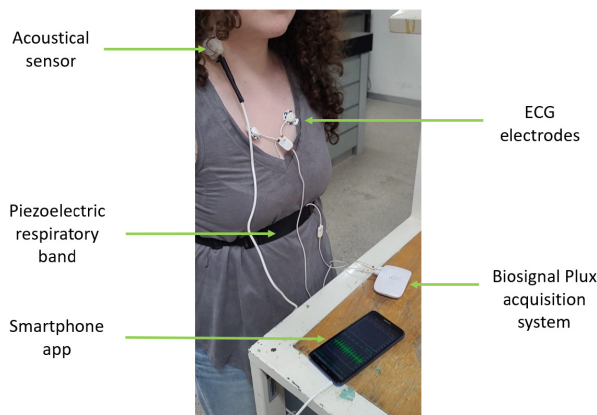
Fig. 3. Experimental setting showing a volunteer during the signal acquisition stage.

For this study, data from fixed RR maneuvers were acquired. To this end, each volunteer was instructed to perform four different metronome breathing maneuvers, each at a fixed RR equal to 0.1, 0.2, 0.3, and 0.4 Hz, equivalent to 6, 12, 18, and 24 breaths-per-minute (bpm). Each recording at a given fixed RR lasted one minute. Visual metronome feedback was provided to the volunteers through a display placed in front of them to facilitate them following the required breathing rhythm.

2.4 Data Analysis

The mCORSA-based RR estimates were compared to the ones obtained from the reference piezoelectric sensor. The reference RR estimates were automatically computed via the PSD of the respiratory reference signal by searching for the frequency corresponding to its maximum peak. Data analysis included descriptive statistics, boxplots, computing absolute errors (AE), in bpm and normalized percent, as well as Bland-Altman analysis and non-parametrical statistical tests using two-sided Wilcoxon signed-rank tests, considering a 5% significance level.

3 Results and Discussion

Screen captures of the implemented mobile app for the estimation of RR from TS are displayed in Fig. 4, where it can be seen the GUI developed for the "add patient" activity (left panel), for the "sound recording" activity (central panel), and the "results display" activity (right panel). The app lets us quickly introduce the patient's information and validate the data. The app allows the user to start and stop the recording of TS, as well as to save or discard the acquisition. After performing all the digital signal processing, from the TS to the RR estimation, the app displays several signals, including the TS alone, the TS together with its corresponding SE signal, and the PSD of the SE together with the detected maximum PSD peak and the frequency peak (RR estimate). It is possible to observe that, for the example shown in Fig. 4, the volunteer was asked to breathe at 12 bpm, and the calculated peak frequency using the app was 0.195 Hz, corresponding to a RR equal to 11.70 bpm.

The RR estimation results obtained for all volunteers are summarized in Table 1 for each metronome maneuver, where RRref denotes the RR reference values based on the piezoelectric respiratory band, and RRapp indicates the RR estimated values computed with the developed mCORSA system. AE represents the absolute error of the estimate, in bpm units and normalized units, concerning the reference value. Each value is presented as mean ± standard deviation, median, and (minimum, maximum). It was found that the higher AE was 1.2 bpm, corresponding to 6.41% of the reference fixed RR of 18 bpm.

Figure 5 shows the normalized AE for each of the four metronome maneuvers (M1–M4), where it can be noted that data does not follow a normal distribution. No statistically significant differences were found between the median of RR estimates from the mCORSA system and those from the piezoelectric reference sensor, at the 5% significance level, for all maneuvers. A statistically significant bias of −0.27 bpm was found, and the 95% limits-of-agreement (LoA) were -1.07 and 0.53 bpm.
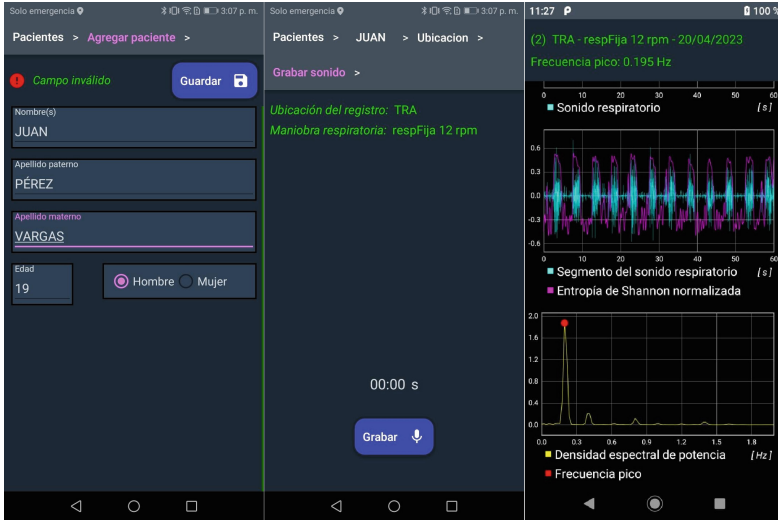
Fig. 4: Screenshots of the developed mobile app for RR estimation using TS. Left: Add patient activity. Center: Sound recording activity. Right: Results display activity.

The obtained results in this study are comparable to previous efforts reported in the literature, e.g., the bias and LoA were found to be 0.11 and -1.41 to 1.63 bpm in a study employing smartphone-acquired TS [5]. In contrast, the median error was less than 1% in a study using smartphone-acquired nasal sounds [6], but our system has the advantage of performing an end-to-end RR estimation in the smartphone.

Table 1: Results obtained for the metronome breathing maneuvers ($N = 10$).

Metronome breathing (bpm)	RRref (bpm)	RRapp (bpm)	AE (bpm)	AE (%)
6	5.76 ± 0.19 5.82 (5.22, 5.82)	$5.82 \pm 8.7\text{x}10{-16}$ 5.82 (5.82, 5.82)	0.06 ± 0.190 0.00 (0.00, 0.60)	1.15 ± 3.63 0.00 (0.00, 11.49)
12	11.88 ± 0.29 11.70 (11.70, 12.30)	$11.70 \pm 1.7\text{x}10^{-15}$ 11.70 (11.70,11.70)	0.18 ± 0.29 0.00 (0.00, 0.60)	1.46 ± 2.39 0.00 (0.00, 4.48)
18	18.18 ± 0.189 18.12 (18.12, 18.72)	$17.52 \pm 3.5\text{x}10{-15}$ 17.52 (17.52, 17.52)	0.66 ± 0.19 0.60 (0.60, 1.20)	3.62 ± 0.98 3.31 (3.31, 6.41)
24	24.06 ± 0.19 24.00 (24.00, 24.60)	23.76 ± 0.58 23.40 (23.40, 24.60)	0.54 ± 0.19 0.60 (0.00, 0.60)	2.25 ± 0.79 2.50 (0.00, 2.50)

Values presented as mean $\pm$ standard deviation, median, (minimum, maximum).

Besides the promising results of this study, some limitations should be addressed in future studies. First, the sensor is sensitive to other sounds coming from the respiratory tract, and for this reason the volunteers were asked not to swallow or talk during the acquisitions. Second, the sample size is small, and we are trying to increase it. Third, the piezoelectric sensor used as a reference is not as good as respiratory bands based on an inductive sensor, but unfortunately the latter is not available in our laboratory. Finally, more breathing maneuvers should be explored, including spontaneous breathing and abrupt changes in RR, where incorporating time-frequency analysis, e.g., the spectrogram, would be helpful to contend with the time-varying nature of the corresponding respiratory rates.
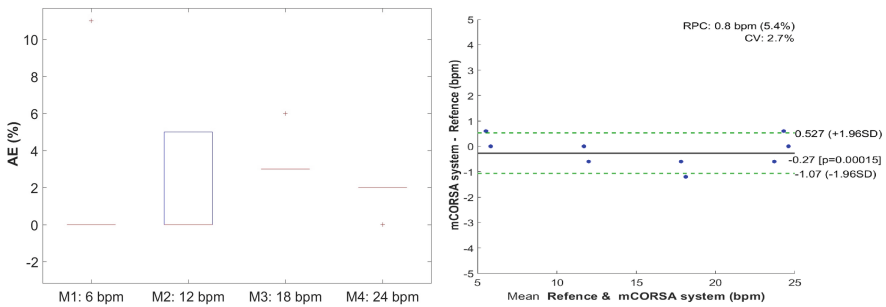
Fig. 5. Results of RR estimation via mCORSA system. Left: Boxplot of the absolute error for each metronome maneuver (M). Right: Bland-Altman graph.

4 Conclusions

The end-to-end estimation of respiratory rate using a smartphone-based system for tracheal sound analysis was explored in this study, obtaining results comparable to the ones reported in the literature for metronome breathing. It is worth mentioning that the recordings of this study were performed in a regular classroom, with the associated acoustical noise from the environment, pointing out the feasibility of extending the RR estimation beyond clinical or research settings.

References

1. Al-Khalidi, F.Q., Saatchi, R., Burke, D., Elphick, H., Tan, S.: Respiration rate monitoring methods: a review. Pediatr. Pulmonol. **46**(6), 523–529 (2011). https://doi.org/10.1002/ppul.21416
2. Beck, R., Rosenhouse, G., Mahagnah, M., Chow, R.M., Cugell, D.W., Gavriely, N.: Measurements and theory of normal tracheal breath sounds. Ann. Biomed. Eng. **33**(10), 1344–1351 (2005). https://doi.org/10.1007/s10439-005-5564-7
3. Acoustic Respiration Rate (RRa), MASIMO. https://www.masimo.com/evidence/pulse-co-oximetry/rra/. Accessed 28 June 2023

4. Reyes, B.A., Olvera-Montes, N., Charleston-Villalobos, S., González-Camarena, R., Mejía-Ávila, M., Aljama-Corrales, T.: A smartphone-based system for automated bedside detection of crackle sounds in diffuse interstitial pneumonia patients. Sensors **18**(11), 3813 (2018). https://doi.org/10.3390/s18113813
5. Reyes, B.A., Reljin, N., Chon, K.H.: Tracheal sounds acquisition using smartphones. Sensors **14**(8), 13830–13850 (2014). https://doi.org/10.3390/s140813830
6. Nam, Y., Reyes, B.A., Chon, K.H.: Estimation of respiratory rates using the built-in microphone of a smartphone or headset. IEEE J. Biomed. Health Inform. **20**(6), 1493–1501 (2015). https://doi.org/10.1109/JBHI.2015.2480838
7. Yadollahi, A., Moussavi, Z.M.: A robust method for estimating respiratory flow using tracheal sounds entropy. IEEE Trans. Biomed. Eng. **53**(4), 662–668 (2006). https://doi.org/10.1109/TBME.2006.870231