In-Ear Audio Wearable: Measurement of Heart and Breathing Rates for Health and Safety Monitoring

Alexis Martin, *Student Member, IEEE*, and Jérémie Voix, *Member, IEEE*

***Abstract*—*Objective:* This paper examines the integration of a noninvasive vital sign monitoring feature into the workers' hearing protection devices (HPDs) by using a microphone positioned within the earcanal under the HPD. *Methods:* 25 test-subjects were asked to breathe at various rhythms and intensities and these realistic sound events were recorded in the earcanal. Digital signal processing algorithms were then developed to assess heart and breathing rates. Finally, to test the robustness of theses algorithms in noisy work environments, industrial noise was added to the in-ear recorded signals and an adaptive denoising filter was used. *Results:* The developed algorithms show an absolute mean error of 4.3 beats per minute (BPM) and 2.7 cycles per minute (CPM). The mean difference estimate is –0.44 BPM with a limit of agreement (LoA) interval of –14.3 to 13.4 BPM and 2.40 CPM with a LoA interval of –2.62 to 7.48 CPM. Excellent denoising is achieved with the adaptive filter, able to cope with ambient sound pressure levels of up to 110 dB SPL, resulting in a small error for heart rate detection, but a much larger error for breathing rate detection. *Conclusion:* Extraction of the heart and breathing rates from an acoustical measurement in the occluded earcanal under an HPD is possible and can even be conducted in the presence of a high level of ambient noise. *Significance:* This proof of concept enables the development of a wide range of noninvasive health and safety monitoring audio *wearables* for industrial workplaces and life-critical applications where HPDs are used.**

***Index Terms*—Acoustic signal processing, biosignals, in-ear wearables, health and safety monitoring, heart and breathing rates.**

I. Introduction

HEAVY industries and confined spaces such as mines are hazardous and extremely noisy work environments in which work accidents and sudden ailments are more likely to occur. Improvements in occupational health and safety policies over the last decades and the number of mining fatalities has decreased significantly, from an average of 141 deaths in the 70's to an average of 35 deaths per year in the USA from 2006 to 2010 [1]. The remote monitoring of workers' vital signs could enable efficient paramedic interventions and further reduce fatalities for these industrial workers but also for other workers, including armed forces, first responders, firefighters and the like.

Non-invasive health monitoring methods already exist and are widely used in clinical applications to monitor physiological parameters such as heart rate or breathing rate. Systems such as electrocardiography, stethoscopy, plethysmography, and spirography are extremely accurate but often cumbersome and only used in controlled environments. More recently, several connected wearable devices have been developed to monitor heart rate or breathing rate: watches, shirts, wrist bands or belts. However, those *wearables* are not appropriate for monitoring the vital signs of workers in the industries mentioned. Most wearables are not designed for the rough environments these workers are exposed to and are not designed to be compatible with the personal protection equipment (PPE) that these workers are likely to already wear. Interestingly, one of the most used PPE is the hearing protection device (HPD), which is used to protect workers from toxic noise and also sometimes, in its electronic version, to communicate over personal radio in very noisy environments.

Therefore, a promising solution to monitor the vital signs of the workers mentioned would be to directly integrate such biosignal monitoring system within an electronic HPD, by taking advantage that these HPDs are often already equipped with external and internal microphones [2]. Using such hardware and some digital signal processing (DSP) the envisioned in-ear audio wearable device could capture the physiological noise resulting from cardiovascular activity (and highly correlated with heartbeats) as well as from respiration using a small microphone positioned inside the occluded earcanal. This approach simplifies hardware design and reduces related costs compared to the use of another type of sensor used in the past, such as piezoelectric film [3], accelerometers or electrodes [4], IR emitter and receiver [5], or a microphone worn on a different part of the body [6]–[8].

Past research on acoustical measurements of heartbeats and respiration has been comprehensively covered by [6], [7], [9]. The developed algorithms seem to perform very well to detect heartbeats and respiration. However, the existing literature

Manuscript received October 6, 2016; revised June 5, 2017; accepted June 13, 2017. Date of publication June 27, 2017; date of current version May 18, 2018. This work was supported in part by the EERS Technologies, Inc., and in part by the NSERC-EERS Industrial Research Chair in In-Ear Technologies. *(Corresponding author: Jérémie Voix.)*

A. Martin is with the Department of Electrical Engineering, École de technologie supérieure.

J. Voix is with the Department of Mechanical Engineering, École de technologie supérieure, Montreal, QC H3C 1K3, Canada (e-mail: jeremie.voix@etsmtl.ca).

Digital Object Identifier 10.1109/TBME.2017.2720463

shows limitations: either the validation has been conducted on very limited data, or the experimental protocol is not explicit, or the recorded signals are not representative of real-world conditions. Furthermore, the microphones used in these studies are located on the trachea or on the chest. Consequently, the measurements are exposed to ambient noise, which may alter the signal being measured. To address this problem, the microphone can be located behind an earplug to take advantage of its passive acoustic attenuation.

Only two studies have been conducted with acoustic measurements in the earcanal. First, Kusche *et al.* created an experimental prototype consisting of a sound pressure sensor (microphone) located outside the ear with a tube passing through an earplug with a microphone [10]. An algorithm was developed to detect heartbeats to calculate the transient time of the wave between the heart and the ear. Unfortunately, this was validated using a very small data set and the protocol was not explicit. It is possible that the subject breathed very lightly to make heartbeat detection easier given that breathing generates low frequency artefacts. Second, Pressler *et al.* measured respiration sounds on 20 subjects with a microphone located inside an earplug and showed that the ear is a promising location to detect respiration [11]. However, no specific algorithms were developed to extract the breathing rate. Moreover, the total duration of the recorded signals is short and the hardware used for acquisition is now obsolete. To the authors' knowledge, no other research has been conducted to extract heart and breathing rates from an in-ear acoustic measurement. Therefore, no database currently exists of sounds measured in the earcanal with signals sufficiently long in duration and representative of real-world conditions.

In this paper, sounds in the occluded earcanal are measured by an in-ear microphone (IEM) located in an instrumented earpiece. A database of in-ear audio recordings was created using a sample of 25 individuals. During the experiment, subjects were asked to breathe at various rhythms and intensities through the mouth or nose to achieve realistic recordings. A total of 16 hours and 40 minutes of sounds in the earcanal were recorded. The features of sounds correlated with heartbeats and respiration are investigated as recorded at this specific location to develop signal processing algorithms to assess the user's heart and breathing rates. Results from the algorithms are then compared to the numerical values obtained by a commercial reference device used during the measurement. Finally, noise is added numerically to the IEM signal to assess the robustness of the algorithms against ambient noise for further applications, such as monitoring workers' health.

This paper is organized as follows: Section II, on materials, describes the database that was created. Section III presents the proposed methods used to develop algorithms and for denoising. The results are presented in Section IV and are followed by the discussion and conclusion in Sections V and VI, respectively.

II. Materials

This section presents the hardware used for data acquisition, the experimental protocol and the characteristics of the database. Sounds in the left and right ear were recorded simultaneously on 25 subjects (19 males and 6 females, aged between 21 and 53, with an average of 28). This study was approved by the *Comité d'éthique pour la recherche*, the internal review board of École de technologie supérieure.

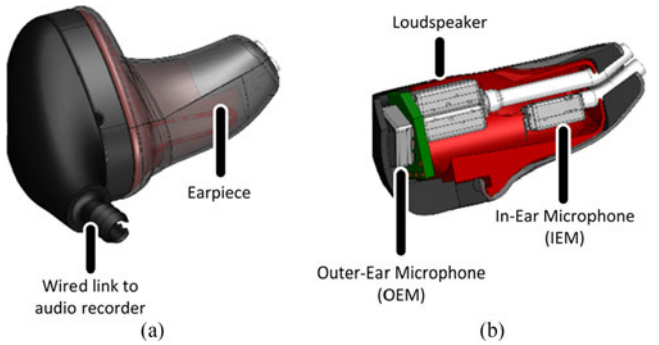

Fig. 1. Instrumented earpiece worn by each subject in each ear: (a) Intra-aural earpiece, (b) electroacoustic components.

A. Data Acquisition

Each participant was equipped with two instrumented earpieces developed by EERS Technologies Inc. (Montreal, Canada) [12] with disposable foam tips (Tx200, Comply, Oakdale, USA). Each earpiece has two microphones and one loudspeaker as illustrated in Fig. 1. An in-ear microphone (IEM) (GA-38, Knowles, Itasca, USA) was used to record sounds in the earcanal, with a usable frequency response ranging from approximately 5 Hz to 20 kHz (slope of -6 dB per octave from 20 to 50 Hz) [13]. No analog filters were used for the signal conditioning. Audio data was recorded with a multichannel digital audio recorder (H4n, Zoom Corporation, Tokyo, Japan), that shows a usable frequency response between 5 Hz and 24 kHz, at a sample rate of 48 kHz and a resolution of 24 bits.

A wearable chest belt (Bioharness 3, Zephyr, Annapolis, USA) was used as a reference system. It was shown that the Bioharness gives very accurate measurements of heartbeats acquired by ECG and respiration when used on a test-subject at rest [14]. The data acquisition framework developed by MuSAE Lab was used to record and display raw data from the Bioharness [15]. A Python routine was developed to synchronize audio data and Bioharness data. A small synchronisation error between audio data and Bioharness data might exist, but it results in a delay of less than 125 ms and thus does not affect the process of extraction of the heart and breathing rates.

B. Experimental Protocol

To ensure that the recordings are as realistic as possible, the test-subjects were asked to breathe at various rhythms and intensities. The test-subjects were seated in an audiometric double-wall sound booth during the recordings. Table I presents the experimental protocol, which was conducted once with nasal breathing and once with mouth breathing, resulting in ten different respiration types. To assess whether the subject's earpieces were well positioned within the earcanal and provided a good attenuation of the ambient noises, the transfer functions between

TABLE I
EXPERIMENTAL PROTOCOL SHOWING REAL-LIFE RECORDINGS, DONE ONCE WITH NASAL BREATHING AND ONCE WITH MOUTH BREATHING

Action	Acronym	Duration (seconds)
Deep slow Breathing	DB	90
Apnea	A1	20
Normal Breathing	NB	240
Apnea	A2	20
Fast Breathing	FB	30
Apnea	A3	20
Exercise on a bike	–	*90*
Apnea after exercise	AE	10
Normal Breathing after Exercise	NBE	180

TABLE II
HEART AND BREATHING RATE VALUES RECORDED WITH BIOHARNESS WITH AVERAGE, STANDARD DEVIATION, MINIMUM AND MAXIMUM VALUES ACROSS ALL CONDITIONS FOR 20 SUBJECTS

	Average	STD	Min	Max
Heart Rate (Beats Per Minute, BPM)	78.6	12.6	50.7	120.7
Breathing Rate (Cycles Per Minute, CPM)	22.8	10.7	3.3	138.0

A wide range of heart and breathing rate values were recorded.

the outer-ear microphone (OEM) and the IEM were assessed for each subject's ear by playing white noise in the sound booth at 85 dB SPL: 30 seconds at the beginning and 30 seconds at the end of the recording session.

No target rhythms or expiration volumes were imposed. The subjects were free to interpret how to breathe according to the type of respiration and therefore, the obtained database contains a wide range of real-life signals. Apnea recordings were used for spectral characterization only.

C. Database Analysis

Five subjects were removed from the analysis, because of a synchronization issue between the audio and Bioharness signals. Table II shows information about the values of heart and breathing rates recorded for the 20 subjects.

An illustrative recording of sound with normal breathing in the occluded earcanal is shown in Fig. 2(a) with the Bioharness reference signals in Fig. 2(b) and (c). The pressure variation measured in the earcanal due to heartbeats can be caused by the combination of several effects: 1) the blood vessels around the earcanal expand due to the pulse wave [4], [10], 2) the indirect bone and tissue conduction of the heartbeat sound from the heart to the occluded earcanal (occlusion effect), 3) movements of the body in response to the heartbeat [16], [17]. The exact origin of the measured sounds is not known yet but appears to be highly correlated with heartbeats and therefore seems to be caused by cardiovascular activity: the measured audio signal is perfectly correlated to an ECG signal. Two events correlated with heartbeat sounds (S1 and S2) are clearly discernible. S1 corresponds to the closure of the tricuspid and mitral valves and occurs after the RS segment of an ECG signal. S2 corresponds to the closure of the aortic and pulmonary valves and occurs during the T wave of an ECG signal [18]. The respiration sounds result

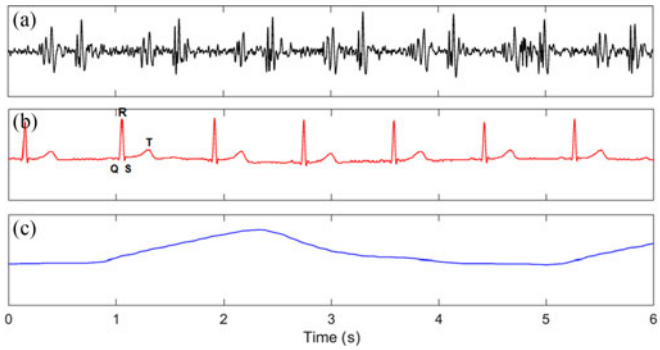

Fig. 2. (a) Time signal recorded by the IEM, showing events correlated to S1 and S2. Respiration is weak and almost non-measurable by the IEM. (b) Heartbeat reference recorded by the Bioharness, which is an electrocardiogram. A delay delay of less than 125 ms between audio data and Bioharness data may exist. (c) Respiration reference recorded by the Bioharness, which is a respiratory inductance plethysmography showing inspiration (ascending phase) and expiration (descending phase).

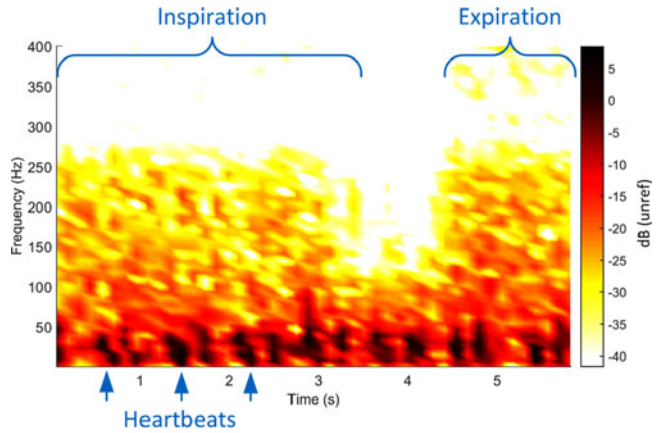

Fig. 3. Illustrative spectrogram of deep mouth-breathing for one subject, showing that both breathing and heartbeats can be measured with the IEM. The inspiration is followed by a short pause and then an expiration.

from turbulence in the nasal and oral cavities and respiratory conduits [19]. In the case of normal breathing, as shown in Fig. 2, the amplitude of respiration sounds is very small and almost non-measurable by the IEM. Moreover, it was observed that in the case of fast breathing, the respiration's sound amplitude generally exceeds the measured sound amplitude correlated with heart activity (S1, S2 or both).

Fig. 3 shows the spectrogram of the audio signal with deep mouth-breathing for one subject. Sounds resulting from cardiac activity are identified by black spots below 60 Hz. The end of an inspiration phase is shown, followed by a short pause and then expiration. Physiological noise at frequencies below 150 Hz is observable due to body activity (muscle activity, blood flow, etc.).

Illustrative spectra of four types of respiration (apnea, normal breathing, fast breathing, and deep breathing) are shown in Fig. 4 for one subject. The energy of sounds correlated with heartbeats is maximized in the frequency band between 10 to 50 Hz. The respiration sounds appear to have low frequency components in the same band as the sound pressure variation

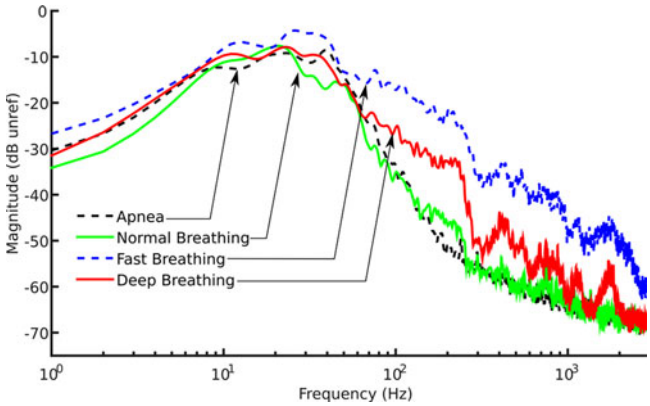

Fig. 4. Spectra of four types of nose respiration for one subject, showing different spectral characteristics to distinguish respiration types.

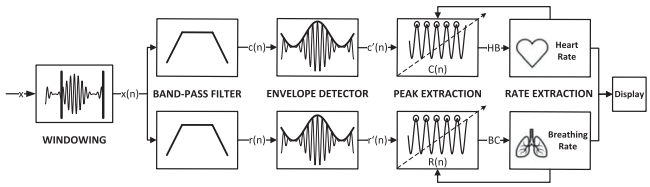

Fig. 5. Simplified block diagram of the heart and breathing rate extraction algorithms, showing the low complexity of the proposed algorithms.

correlated to heartbeats. The energy in the band of 100 to 400 Hz varies greatly depending on the subject and type of respiration. Respiration sound energy drops when frequency increases and sounds above 3000 Hz are not audible.

Sound pressure levels within the earcanal range between 45 dB and 65 dB SPL, dropping to 20 dBA–40 dBA when using A-weighting filters, because of low frequency components.

III. Methods

Digital audio signal processing algorithms were developed to extract heart and breathing rates from these audio recordings. To simulate a real-life situation, the audio recordings were sequentially added in this specific order: mouth breathing (normal, fast, normal after exercise, deep slow) then nasal breathing (same order). For breathing rate extraction, the fast breathing recordings were removed. The duration of analyzed signals was 36 minutes per subject with 8 respiration types for heart rate and 35 minutes per subject with 6 respiration types for breathing rate.

A. Algorithms for Extracting Heart and Breathing Rates

First, recorded audio signals were downsampled to 4800 Hz to reduce processing time, by applying a low-pass filter and removing samples to reduce the sampling rate. A block diagram of the algorithms is presented in Fig. 5. The first stage framed the input data *x* into windows of 10 seconds *x(n)*, where *n* ranges from 0 to $\mathrm{M}-1\,(\mathrm{M}\,=\,48000)$. Then, the signal was sent to two similar processes: one for heart rate extraction, one for breathing rate extraction. For heart rate extraction, *x(n)* was downsampled to 160 Hz, then band-pass filtered from 15 Hz to 45 Hz, to obtain *c(n)*. For breathing rate extraction, *x(n)* was

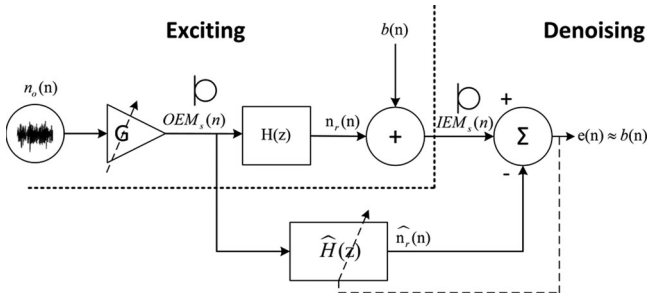

Fig. 6. Block diagram of adaptive filtering, showing the excitation portion, where industrial noise is added, and the denoising portion, where industrial noise is removed from the audio signal measured in the occluded earcanal.

downsampled to 1600 Hz, then band-pass filtered from 150 Hz to 400 Hz to obtain *r(n)*. Envelope extraction was done by applying the Hilbert Transform with a moving average. Each envelope was downsampled to 16 Hz to obtain *c'(n)* and *r'(n)*. Then, the peak extraction process included several steps, one of which was a band-pass filter with cut-off frequencies computed from the spectra of *c'(n)* and *r'(n)* to obtain *C(n)* and *R(n)*. Then, moving thresholds were applied to *C(n)* and *R(n)* to determine whether a beat or a respiration phase (inspiration or expiration) was detected. Heart and breathing rates were computed based on the number of heartbeats (HB) and breathing cycles (BC) detected. A minimum sample number between two detections was computed using previous values of the heart and breathing rates to avoid erroneous detection, assuming that these biosignals are somewhat stable over a couple of seconds.

To evaluate the performance of the algorithms, the absolute error (also called least absolute deviation) for one subject and one sequence of 18 minutes was computed with the following formula: $\varepsilon\ = \frac{1}{N}\sum_{i\,=\,1}^{N}|\mathrm{Ref}_i - A_i|$, where Ref_i is the value of the reference rhythm (BPM or CPM), A_i is the value of the rhythm computed by the algorithms (BPM or CPM) and N is the number of observations. Heart and breathing rates were computed during 5 seconds each using the current detections of HB and BC and two previous values of the heart and breathing rates. Also, the relative error was defined by the difference in percentage between the reference values and the algorithm output values.

B. Denoising of Biosignals From Ambient Noise

To simulate a noisy work environment such as a mine or a factory plant, realistic industrial noise recordings were added numerically (off-line) to the IEM signal and the performance of the developed algorithms were assessed in the presence of these disturbances.

White noise and industrial noise from NASA's steam plant database were used [20]. The block diagram of the process is presented in Fig. 6. First, the excitation consisted of multiplying the noise signal (white or industrial noise) $n_o(n)$ by a gain G, which is computed to obtain a calibrated noise level ranging from 50 to 110 dB SPL (in steps of 5 dB). Then, the normalized noisy signal $OEM_s(n)$ goes through $H(z)$ to obtain the residual noise inside the ear $n_r(n)$. $H(z)$ is the true transfer function of

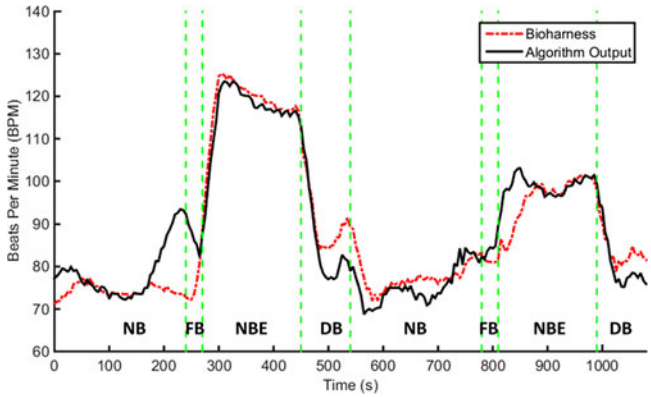

Fig. 7. Comparison between the reference heart rate and the output of the algorithm developed for 8 types of respiration over 18 minutes (in one ear), showing general good agreement between the two curves, with a mean absolute error of 4.0 BPM. Vertical lines indicate separations between the respiration types: four mouth breathing followed by four nose breathing.

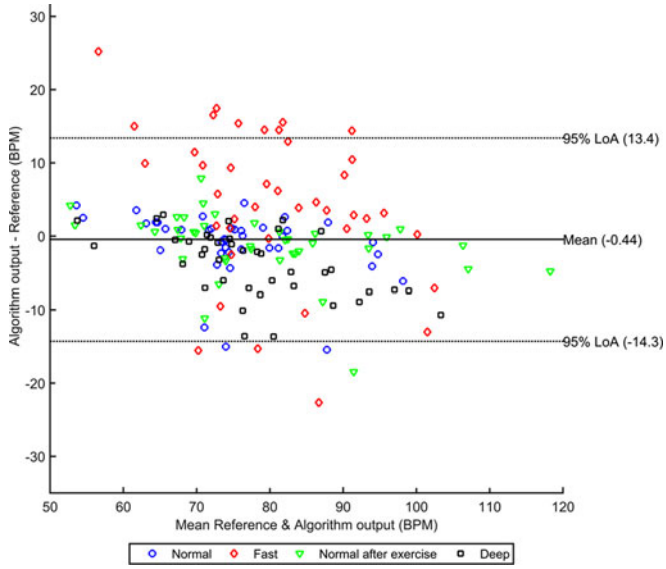

Fig. 8. Bland-Altman plot of heart rate extraction, showing good predictive capabilities of the algorithm on a wide range of BPM (mouth and nose breathing are both plotted but have the same symbol).

the subject's earplug computed from measurements made during the experimental protocol. Then, $n_r(n)$ was added to the biosignals *b(n)* measured by the IEM, which contains sounds resulting from cardiovascular and breathing activities, to obtain the noisy biosignals. Second, the denoising portion consisted of removing the residual noise from the noisy $IEM_s(n)$ signal originating from the ambient industrial noises and disturbances. Denoising was performed using a normalized least mean squared (nLMS) adaptive filter, originally developed by Bou Serhal *et al.* for denoising speech signals captured in an occluded earcanal with an IEM [21]. The structure of the adaptive filter is based on the structure described by [22] except that the signal of interest is the error signal *e(n)* [21]. $\widehat{H}(z)$ is the estimated transfer function of the earpiece (primary transfer function) and $\widehat{n}_r(n)$ is the estimated residual noise. Finally, the denoised biosignals were fed to the extraction algorithms previously developed.

IV. Results

Algorithms were implemented in Matlab (The MathWorks Inc., Natick, USA). This section presents the results of the extraction algorithms and the denoising from ambient noise for heart and breathing rates by computing the absolute and relative error defined in the "Methods" section.

A. Heart Rate Extraction Algorithm

The evolution of the heart rate over time is shown in Fig. 7 for one subject over a sequence of 18 minutes. The absolute error was computed as the difference between the reference rhythm and the algorithm output rhythm. The average absolute error over the entire 18 minutes of recording for this subject is 4.0 BPM. The routine for computing the heart rate from heart beats detected by the algorithm induces a delay on the curves when the respiration type changes. Also several artefacts were audibility identified afterwards, such as test-subject swallowing at 165 s., electrical interferences between 180 and 240 s. subject's body movement at 480 and 820 s.

For heart rate extraction, the mean absolute error for the 20 subjects, computed as the mean of all the individual absolute errors ε defined in the section "Methods", is 4.3 BPM, with a standard deviation of 2.2 BPM. This gives a mean relative error of 5.6%, with a relative standard deviation of 51.2%, computed as the radio in % between the standard deviation and the mean. Fig. 8 is a Bland-Altman plot of the heart rate obtained for each respiration condition for each subject. Algorithm output was close to the reference for most recordings. The mean difference estimate for heart rate estimation is -0.44 BPM with a 95% limit of agreement interval of -14.3 to 13.4.

B. Breathing Rate Extraction Algorithm

For breathing rate extraction, the fast breathing recordings were removed. The mean absolute error for the 20 subjects is 2.7 CPM, with a standard deviation of 1.6 CPM. This gives a mean relative error of 30.9%, with a relative standard deviation of 59%. Fig. 9 is a Bland-Altman plot of the breathing rate obtained for each respiration condition for each subject. Algorithm output is close to the reference for low values of the breathing rate under 30 CPM. The mean difference estimate for breathing rate estimation is 2.40 CPM with a 95% limit of agreement interval of -2.62 to 7.48.

C. Denoising of Biosignals From Ambient Noise

This section presents the results of the proposed extraction algorithms for heart and breathing rates when noise is added to the signal, as described in the "Methods" section.

Figs. 10 and 11 show the evolution for the 20 subjects of the mean absolute error and the mean relative error respectively, when the signal is disturbed by an industrial noise at levels ranging from 50 to 110 dB SPL.

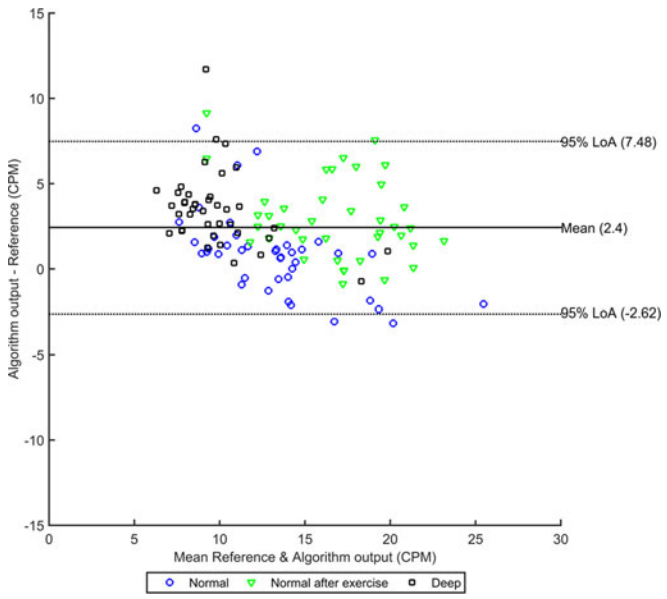

Fig. 9. Bland-Altman plot of the breathing rate extraction, showing good predictive capabilities of the algorithm on a breathing rate of fewer than 30 CPM (mouth and nose breathing are both plotted but are represented by the same symbol).

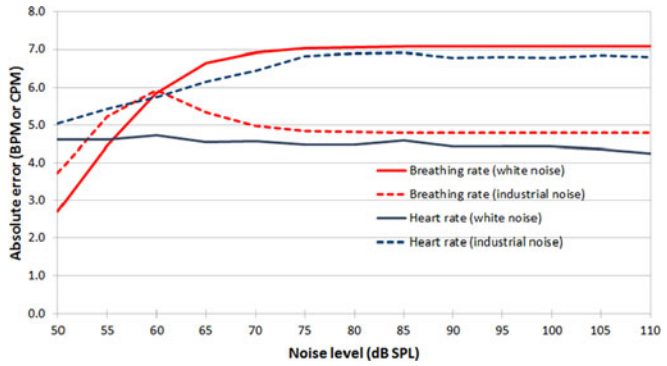

Fig. 10. Evolution of absolute error averaged for 20 subjects as a function of noise level, showing good performance for heart rate extraction with broadband white and industrial noise. The increase in the error is greater for breathing rate extraction with white and industrial noise.

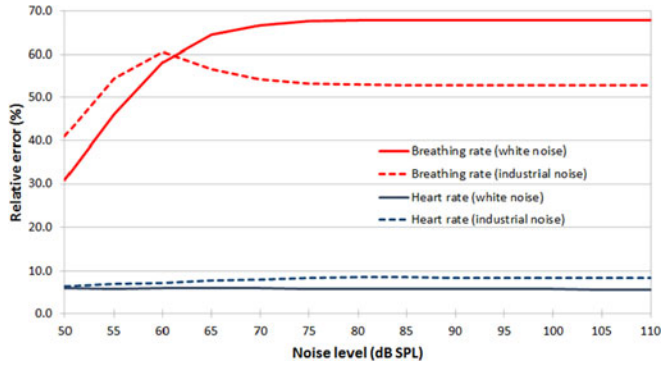

Fig. 11. Evolution of relative error averaged for 20 subjects as a function of noise level, showing good performance for heart rate extraction with white and industrial noise. The increase in the error is much greater for breathing rate extraction than for heart rate.

For the heart rate, the absolute error does not exceed 4.7 BPM for broadband white noise (6.1% relative error) and 7.6 BPM for industrial noise (9.1% relative error). With white noise, errors plateau from 50 dB up to 110 dB SPL. With industrial noise, errors increase up to 85 dB and then plateau up to 110 dB SPL.

For the breathing rate, the absolute error does not exceed 7.1 CPM for white noise (68% relative error) and 5.9 CPM for industrial noise (60.6% relative error). With white noise, errors increase until 75 dB and then plateau up to 110 dB SPL. With industrial noise, errors increase until 60 dB then plateau from 65 dB to 110 dB SPL.

V. Discussion

A. Heart Rate Extraction Algorithm

Experimental results showed an accurate extraction of the heart rate for 20 subjects. The mean absolute error is 4.3 BPM (5.6% relative error), as observed in Fig. 8. Failure to detect the heart rate is mostly due to four factors: short duration noise artefacts, respiration type, earplug fit quality, and weak sound amplitudes. Short duration noise artefacts include body movements and any type of noise generated by the person (swallowing, coughing, etc.). They alter the recorded signal and make heartbeat and breathing cycle detection harder by adding low frequency artefacts. This would explain the larger absolute errors observed in Fig. 7 at 180, 480 and 820 seconds. Also, strong breathing might generate low frequency artefacts that mask the sound correlated with heartbeats. The envelope of a fast breathing signal might have a similar fundamental frequency as that of the heartbeat, causing the algorithm to misinterpret one for the other. The measured signals are greatly affected by the quality of the fit of the earplug. Indeed, a small acoustic leak during a recording can result in energy drops in the lower frequencies. Even if transfer functions were checked before and after the measurement to detect a bad fit, it is possible that the recorded signals were corrupted by acoustic leaks. Finally, the measured physiological sounds were, for some specific subjects and specific respiration types, almost non-measurable by the IEM. Algorithm could output erroneous values for these specific recordings.

B. Breathing Rate Extraction Algorithm

Extraction of the breathing rate is less accurate than heart rate extraction, with an absolute error of 2.7 CPM (30.9% relative error). For some subjects, the extraction works very well but for others, there is a much larger error, which increases the group mean error, as observed in Fig. 9. Failure to accurately detect breathing signal is mostly due to five factors: the presence short-duration noise artefacts, the poor fit of the earpiece inside the earcanal, the weak breathing sound amplitudes, the lack of instructions to the test-subjects for the type of respiration, and the use of a default sub-optimal algorithm. Three of these limitations have already been discussed above, in Section A. To produce the widest range of real-life respiration intensities, no indication was given to the subjects. Consequently, subjects breathed very differently, with their own interpretation of rhythms and

intensities. This variety of conditions made the development of the algorithm more challenging. The "fast" breathing rates were removed from the results, as they ranged from values as low as 30 CPM to values reaching extremes as high as 138 CPM, making them too widely spread. Besides, the proposed algorithm was initially developed for heart rate extraction and then was adapted for breathing rate extraction. Promising results were achieved for some subjects, but the method show limitations for high breathing rate extraction and should be further optimized with a larger and more controlled dataset.

C. Originality and Overall Performance

The commercial reference device (Bioharness) used as reference also introduced a small bias compared to a gold standard clinical device [14]. The bias was computed with over 50 minutes of data from 10 subjects for static activity. For heart rate, the mean error was 0.78 BPM with a standard deviation of 2.21 BPM. For breathing rate, the mean error was 0.19 CPM with a standard deviation of 1.22 CPM. Furthermore, in their study [14], the heart rate ranges between 65.0 and 165.0 BPM and the breathing rate ranges between 6.0 and 65.0 CPM. In this study, the heart rate ranged between 50.7 and 120.7 BPM and the breathing rate ranged between 3.3 and 130.8 CPM. It is therefore possible that the Bioharness error could be larger for extreme values of breathing rate in particular.

The current study is the first to extract heart and breathing rates from acoustical measurements in the occluded earcanal. Comparing these results to those obtained in the literature is challenging. Indeed, studies conducted by Kaufmann *et al.* and Pressler *et al.* did not have the same research problem as that of this study. Also, signals measured at other locations such as phonocardiogram signals [23] are less noisy than earcanal signals. Besides, the database created in this paper has unique characteristics. First, its duration is consequential, with 12 hours analyzed. Second, the experimental protocol is explicitly stated: there are several conditions of respiration with the mouth and the nose, such wide range of conditions also tends to increase the number of artefacts. Third, subjects breathed very differently, with their own interpretation of "fast" and "deep" breathing with various rhythms and intensities. The resulting range of heart and breathing rates covers a wide range, making these signals representative of real-life conditions. In contrast, all these conditions are not seen in the existing literature covered by [6], [7], [9]–[11] Nevertheless, the present study does comprise certain inherent limitations. First, it was conducted on steady subjects, while it is clear that motion artifacts could corrupt the recorded signal: low frequency noise induced by the worker's footsteps can be captured by the microphone, and would probably affect heart rate extraction. However, such motion can easily be detected by the inertial measurement unit (IMU) and associated accelerometers that will be present in the envisioned version of the in-ear wearable device. Besides, the monitoring of these biosignals in industrial workplaces becomes life-critical, precisely when the worker is no longer moving, such as for a "man down" situation, where the vital signs need to be quickly assessed to possibly send immediate assistance.

D. Denoising of Biosignals From Ambient Noise

Denoising of the biosignals was conducted offline, in a Matlab simulation, without actually running the algorithms while the subject was exposed to the ambient noise source. While the results obtained are promising, they should also be field-validated in real-life environments. This study proposes an original approach that uses adaptive filtering to denoise acoustic biosignals measured inside the earcanal and a similar approach has never before been proposed in the literature.

For heart rate extraction, simulation results showed only a very small increase in the absolute error, as the adaptive denoising filter performed well in low frequencies. The absolute error does not exceed 4.7 BPM for white noise and 7.6 BPM for industrial noise, which still represent low relative errors (respectively 6.1% and 9.1%). Compared to the white noise, the industrial noise had a higher energy levels in the 20–120 Hz frequency range and consequently affected more significantly the detection in low frequency for heartbeats.

For the breathing rate extraction, simulation results showed a significant increase in the absolute error. The adaptive denoising filter does not perform well above 65 dB SPL. The absolute error increases until 7.1 CPM for white noise and 5.9 CPM for industrial noise, which represents a significantly high relative errors (68% and 60.6% respectively). But again, this group error comes from an average over 20 subjects, among which eight have a very high relative error (between 70% and 280%) which dramatically affect the group mean performance. Industrial noise has a spectrum with tonal components which clearly distinguishes it from the breathing spectrum in the band of interest, namely from 150 to 400 Hz. On the contrary, white noise, with its equi-level of energy at all frequencies, shows a spectrum somehow similar to a breathing spectrum for several subjects, in the frequency range of interest, making it harder to distinguish from a breathing spectrum.

VI. Conclusion

To continuously monitor the vital signs of industrial workers, this paper proposed to measure physiological sounds in the occluded earcanal with a wearable audio device equipped with an in-ear microphone. A unique database of close to 17 hours of sounds measured in the occluded earcanal was created, and included a broad range of real-life types of respiration. A commercial device was used as a reference to simultaneously record heartbeats and respiration. The developed algorithms show an absolute mean error of 4.3 BPM and 2.7 CPM for heart and breathing rates extractions respectively. To assess the robustness of the algorithms against ambient noise, broad band white noise and industrial noise were added numerically to the in-ear signal. An nLMS adaptive filter was used to remove the unwanted noise disturbance. Excellent performance is achieved with noise levels up to 110 dB SPL with absolute errors below 7.6 BPM for heart rate detection. For breathing rate detection, the performance are definitely affected by the noise disturbance and result in absolute errors lower than 7.4 CPM. Nevertheless, this proof of concept now enables the development of in-ear technologies for a wide range of non-invasive vital signs monitoring appli-

cations, such as monitoring workers in hazardous and noisy environments through their HPDs and could be applicable to many fields from heavy industry and construction workers [14], to mining or firefighters [24]. Future work should focus on improving the performance of the algorithms, especially for high breathing rate detection, and optimizing the adaptive filtering. This could potentially be tackled by taking into account the unique characteristics of both the earplug attenuation and the spectrum of biosignals for each individual using this type of wearable device and individually tuning the proposed algorithm through a training or initialization phase.

Acknowledgment

The authors would like to thank MuSAE Lab for providing equipment for the experimental setups.

References

[1] Mine Safety and Health Administration—MSHA, "MSHA fact sheets—Injury trends in mining." [Online]. Available: http://arlweb.msha.gov/mshainfo/factsheets/mshafct2.htm. Accessed on: Dec. 13, 2016.

[2] J. Voix, "Did you say 'bionic' ear?" *Can. Acoust.*, vol. 42, no. 3, pp. 68–69, Aug. 2014.

[3] J.-H. Park *et al.*, "Wearable sensing of in-ear pressure for heart rate monitoring with a piezoelectric sensor," *Sensors*, vol. 15, no. 9, pp. 23402–23417, Sep. 2015.

[4] D. D. He *et al.*, "An ear-worn vital signs monitor," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 11, pp. 2547–2552, Nov. 2015.

[5] S. Vogel *et al.*, "In-ear vital signs monitoring using a novel microoptic reflective sensor," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 6, pp. 882–889, Nov. 2009.

[6] G. Chen *et al.*, "Algorithm for heart rate extraction in a novel wearable acoustic sensor," *Healthcare Technol. Lett.*, vol. 2, no. 1, pp. 28–33, Feb. 2015.

[7] P. Corbishley *et al.*, "Breathing detection: Towards a miniaturized, wearable, battery-operated monitoring system," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 1, pp. 196–204, Jan. 2008.

[8] T. T. Zhang *et al.*, "Sound based heart rate monitoring for wearable systems," in *Proc. 2010 Int. Conf. Body Sensor Netw.*, 2010, pp. 139–143.

[9] A. Yadollahi *et al.*, "A robust method for estimating respiratory flow using tracheal sounds entropy," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 4, pp. 662–668, Apr. 2006.

[10] R. Kusche *et al.*, "An in-ear pulse wave velocity measurement system using heart sounds as time reference," *Curr. Direction Biomed. Eng.*, vol. 1, no. 1, pp. 366–370, 2015.

[11] G. A. Pressler *et al.*, "Detection of respiratory sounds at the external ear," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 12, pp. 2089–2096, Dec. 2004.

[12] J. Voix *et al.*, "Advanced communication earpiece device and method," U.S. Patent Application 20160119732, 2011.

[13] Knowles, Itasca, IL, USA. *GA Series*. [Online]. Available: http://www.knowles.com/eng/Products/Microphones/Subminiature-performance/GA-series. Accessed on: Dec. 15, 2016.

[14] U. C. Gatti *et al.*, "Physiological condition monitoring of construction workers," *Autom. Construction*, vol. 44, pp. 227–233, Aug. 2014.

[15] R. Cassani *et al.*, "MuLES: An open source EEG acquisition and streaming server for quick and simple prototyping and recording," in *Proc. 20th Int. Conf. Intell. User Interfaces Companion*, New York, NY, USA, 2015, pp. 9–12.

[16] A. D. Wiens *et al.*, "Toward continuous, noninvasive assessment of ventricular function and hemodynamics: Wearable ballistocardiography," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 4, pp. 1435–1442, Jul. 2015.

[17] O. T. Inan *et al.*, "Ballistocardiography and seismocardiography: A review of recent advances," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 4, pp. 1414–1427, Jul. 2015.

[18] W. Kluwer, Ed., *Evaluating Heart and Breath Sounds*. Baltimore, MD, USA: Williams & Wilkins, 2009.

[19] Z. Moussavi, *Fundamentals of Respiratory Sounds and Analysis*. Williston, VT, USA: Morgan & Claypool, 2007.

[20] NASA/Glenn Research Center, *Auditory Demonstrations in Acoustics and Hearing Conservation*, vol. Multimedia Production #334, C-2004-786, 2004. [Online]. Available: https://archive.org/details/C-2004-786

[21] R. Bou Serhal *et al.*, "Improving the quality of in-ear microphone speech via adaptive filtering and artificial bandwidth extension," *J. Acoust. Soc. Amer.*, 2017. [Online]. Available: http://asa.scitation.org/doi/abs/10.1121/1.4976051

[22] D. G. Manolakis *et al.*, *Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing*. Boston, MA, USA: Artech House, 2005.

[23] S. E. Schmidt *et al.*, "Segmentation of heart sound recordings by a duration-dependent hidden Markov model," *Physiol. Meas.*, vol. 31, no. 4, pp. 513–529, Apr. 2010.

[24] S. A. Kahn *et al.*, "Line of duty firefighter fatalities: An evolving trend over time," *J. Burn Care Res.*, vol. 36, no. 1, pp. 218–224, Feb. 2015.