RESEARCH ARTICLE | MARCH 19 2024

Airborne ultrasound for the contactless mapping of surface thoracic vibrations during human vocalizations: A pilot study ¹

Frédéric Wintzenrieth [[]© ; Mathieu Couade **≥** [[]© ; Feizheun Lehanneur [[]© ; Pierantonio Laveneziana [[]© ; Marie-Cécile Niérat [[]© ; Nicolas Verger [[]© ; Mathias Fink [[]© ; Thomas Similowski [[]© ; Ros Kiri Ing [[]©

() Check for updates

AIP Advances 14, 035033 (2024) https://doi.org/10.1063/5.0187945



APL Energy



Latest Articles Online!





Read Now

ARTICLE

r٦٦

Export Citation

Airborne ultrasound for the contactless mapping of surface thoracic vibrations during human vocalizations: A pilot study ¹

Cite as: AIP Advances 14, 035033 (2024); doi: 10.1063/5.0187945 Submitted: 13 December 2023 • Accepted: 20 February 2024 • Published Online: 19 March 2024

Frédéric Wintzenrieth,^{1,a)} (D) Mathieu Couade,^{1,a),b)} (D) Feizheun Lehanneur,^{2,a)} (D) Pierantonio Laveneziana,^{2,3,a)} (D) Marie-Cécile Niérat,^{2,a)} (D) Nicolas Verger,^{2,a)} (D) Mathias Fink,^{4,a)} (D) Thomas Similowski,^{2,3,5,a)} (D) and Ros Kiri Ing^{4,a)} (D)

AFFILIATIONS

¹Austral Diagnostics, 1 Rue Jussieu, 75005 Paris, France

- ²Sorbonne Université, INSERM, UMRS1158 Neurophysiologie Respiratoire Expérimentale et Clinique, 75005 Paris, France
- ³Groupe Hospitalier Universitaire APHP-Sorbonne Université, Département R3S, Service des Explorations Fonctionnelles de La Respiration, de L'Exercice et de La Dyspnée, Site Pitié-Salpêtrière, 75013 Paris, France
- ⁴Institut Langevin, ESPCI Paris, PSL University, 1 Rue Jussieu, 75005 Paris, France
- ⁵Groupe Hospitalier Universitaire APHP-Sorbonne Université, Département R3S, Site Pitié-Salpêtrière, 75013 Paris, France

^{a)}Postal address: Institut Langevin, 1 rue Jussieu, 75005 Paris, France. ^{b)}Author to whom correspondence should be addressed: mcouade@australdx.fr

ABSTRACT

Physical examination of the thorax is key to the clinical diagnosis of respiratory diseases. Among other examination techniques, palpation evaluates the transmission of high-frequency vibrations produced by vocalizations (tactile fremitus), which helps the physicians to identify abnormalities within the respiratory system. We propose the use of an airborne ultrasound surface motion camera (AUSMC) to quantitatively map the vibrations induced by subject vocalization. This approach could make the examination of vocal fremitus quantifiable, reproducible, and archivable. Massive data collection of vocal fremitus could allow using artificial intelligence algorithms to isolate vibration patterns that could help disease identification. Until now, in contrast, the interpretation of vocal fremitus has been subject to the physician's experience and remains subjective. In the present work, we demonstrate the capabilities of the AUSMC to measure vocal fremitus thoracic vibration maps on 77 healthy volunteers. We have observed a spatial dependence of vibration maps on vocalization frequency. We observed that the left lung generates fewer surface vibrations than the right one, which was expected according to their respective dimensions. We also discuss the implications of our findings.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/5.0187945

INTRODUCTION

Physical examination of the thorax is key to the clinical diagnosis of respiratory diseases. It includes the assessment of the respiratory system's ability to transmit vibrations of various frequencies and origin. Auscultation evaluates the transmission of sound vibrations produced by the airflow induced within the bronchial tree and the lungs during inspiration and expiration. Palpation evaluates the transmission of sound vibrations produced by the larynx during vocalizations and propagating through the entire respiratory tree (tactile fremitus). Physicians compare the sensory elements perceived during physical examination with a personal repertoire resulting from initial learning and accumulated practice. This allows them to identify abnormalities deriving from disease-related changes in the structural characteristics of the respiratory system. For example, localized pneumonia results in the concerned region of the lungs becoming retracted and denser than the surrounding normal tissue. On auscultation, this translates into a modified breathing sound (tubal murmur). On palpation, this translates into augmented tactile fremitus. $^{\rm l}$

The value of physical examination of the thorax depends on the skills and experience of the operator but is generally low.² Indeed, the performances of the human senses limit it, as does its fragmentary nature (localized sampling). In addition, the collected data cannot be recorded for transmission and external assessment and used for future reference. Direct contact is required between the examiner and the patient.

Examining medical literature allows one to identify several attempts to overcome these limitations. Likewise, multi-point electronic stethoscopes can give access to detailed descriptions of breath sounds and their topographical distribution.^{3,4} For example, "vibration response imaging" improves upon the diagnostic performance of physical examination in a range of clinical circumstances,^{5,6} with or without the aid of artificial intelligence.⁷ Electronic stethoscopes focus on breathing sounds and therefore relate to auscultation. In contrast, literature searches fail to identify devices that would characterize surface vibrations produced by vocalizations and therefore improve upon tactile fremitus.

We previously described the possibility of imaging thoracic surface displacements due to respiratory and cardiac activities using airborne ultrasounds, a contactless method generally relying on the physical principles governing sonar system operation.^{8,9}

Here, we hypothesized that, adequately adapted, this technology would detect the low amplitude movements produced by vocalizations at the surface of the chest and therefore provide a substitute for tactile fremitus. We further hypothesized that it would be possible to map these vibrations and follow the evolution of their characteristics over a short period of time.

To test this hypothesis, we developed a multi-point airborne ultrasound vibrometer capable of mapping surface vibrations over 30 \times 30 cm² at a rate up to 1 kHz, with a spatial resolution of 30 mm and an instantaneous sensitivity to surface displacement of 100 μ m/s.¹⁰ The good harmonicity of the fremitus signals allows averaging-equivalent sensitivity improvements. We then performed pilot measurements in healthy volunteers. The study's objectives were to determine the feasibility of an objective measure of tactile vocal fremitus and to provide pilot normative data, in the perspective of using the technology for an early detection of respiratory system disease-related structural changes.

Using the same device, we demonstrate here its capacity to map the small amplitude periodic vibrations induced by single vocalization over the entire thorax.

MATERIALS AND METHODS

Airborne ultrasound surface motion camera (AUSMC)

The theory of operation of the AUSMC system has been described in Ref. 10. In brief, its principle consists in continuously sonicating the observation surface with a 40 kHz central frequency ultrasonic chirp using a set of panels embedding matrix arrays of piezoelectric transmitters. Emitted chirps are designed to increase the effective bandwidth while defining a specific Doppler period repetition frequency (PRF). The latter ranges from 600 Hz to about 1 kHz so that the Doppler Nyquist frequency lies between 300 and 500 Hz. The emitted ultrasonic waves are reflected by the

surface toward a receive matrix array spatially interleaved within the transmitting matrix array. The ultrasonic signal received by each microphone is amplified and digitized. Then, a software beamforming operation is applied to focus on each resolution cell of the observation surface. The displacement information is retrieved by measuring the phase shift between two successive Doppler frames of signals beamformed for a given resolution cell. Compared to previous work,¹⁰ the minimal PRF has been raised up to 600 Hz to correctly sample the human subject's vocalizations, of which the fundamental frequency typically lies between 100 and 300 Hz.

Study population

Seventy-seven participants were recruited in the study (37 men, 37 women, and three unspecified; median age 22, minimum 18, maximum 57; median body mass index 22 kg·m⁻², minimum 17, maximum 28). The inclusion criteria were:

- no history of chronic disease of any sort;
- absence of acute disease of any sort at the time of the study;
- non-smokers or smokers with less than two pack-years;
- normal physical examination of the heart and respiratory system;
- normal vital capacity and forced expiratory volume in one second on spirometry performed the day of the study;
- mastery of the French language;
- being affiliated to the French social security system.

The non-inclusion criteria mirrored the inclusion criteria, with the addition of legal guardianship, known pregnancy, or lactation. One protocol violation consisted in the inclusion of a smoker. The study was approved by the appropriate regulatory and ethical authorities (CPP SUD-OUEST ET OUTRE-MER II, decision reference: 2-18-2; March 08, 2018). The participants were informed of the purpose of the study and the methods used and gave written consent to participate. One protocol violation consisted in the inclusion of a smoker.

Acquisition protocol

The participants were first trained to reproducibly inhale up to their total lung capacity (TLC) with the help of spirometrybased feedback. During the ultrasound acquisitions, they stood at about 70 cm away from the ultrasonic array, turning their back on the array. They were asked to inflate their lungs to TLC, from which they had to produce a steady vocalization for a few seconds (/a/, /o/, or /z/sound). Meanwhile, the AUSMC system was set to record the surface motion for about 3 s. No specific instruction was given regarding voice pitch, assuming the participants would vocalize at their "natural" pitch. Most acquisitions were performed with arms crossed for scapula to be located on the sides of the back.

Detection of vocalization

The vibration signal corresponding to the vocalization is partially shadowed by different noises. This background noise includes low frequency motion of the thorax (breathing and cardiac motion), acoustic clutter (array beamforming artifacts), acoustic background

noise, thermal noise (electric), and random phase noise (electronic). First, the fundamental frequency of the vocalization f_v was detected by maximizing the spatial variance of the detrended Fourier domain cross spectral density $\Gamma_{va}(f)$, crossing the Doppler velocity fields V_t with the raw signal of the receive array in the audio band A_t . This is made possible as the microphones are sensitive to both the audio (voice) and the ultrasound frequency bands (20–60 kHz). While the power spectral density may be defined as the Fourier transform of the autocorrelation function γ_{xx} , the cross spectral density is defined as the Fourier transform of the cross-correlation function γ_{vv} ,

$$\Gamma_{xy}(f) = \mathcal{F}\{\gamma_{xy}\}(f) = \sum_{\tau=-\infty}^{+\infty} \gamma_{xy}(\tau)e^{-2\pi i\tau f}$$

where

$$\gamma_{xy}(\tau) = \mathbf{E}[(x_t - \mu_x)(y_{t+\tau} - \mu_y)].$$

The latter are computed by using a readily available discrete time implementation (e.g., MATLAB[®] or PythonTM SciPy).

Second, the power spectrum of the thorax vibrations was computed over the time dimension. The Fourier vector corresponding to f_v was isolated, and the corresponding energy map was displayed on top of the body surface profile as captured by a 3D video camera.

Spectral analysis is performed using Welch's method with a time window of about half a second. The corresponding signal-to-noise ratio (SNR) gain is about 27 dB, leading to a harmonic sensitivity of about $5 \,\mu$ m/s.

Spatial analysis

Once the vocal frequency is identified, the vibrational energy map at the vocal frequency is analyzed. To simplify the analysis of the spatial distribution across the cohort, the center of mass (i.e., barycenter) of the vibrational energy map at f_v is computed for each

dataset, which allows us to easily represent and analyze the interindividual statistics in a three-dimensional space, i.e., center of mass and vocalized central frequency f_v .

RESULTS

The mapping of vocalization was feasible for all enrolled subjects. Some acquisitions had to be discarded because of sampling issues such as temporal spikes in the per channel ultrasonic data. Figure 1 shows an example of the vocal spectrum (audio signal), the vibration spectrum, and the cross-spectrum. It shows the interest of analyzing the cross-spectrum rather than the raw Doppler spectrum to enhance the corresponding vocalization signal at the expense of other source of vibrations, for instance, cardiac motion, or colored noises. It also allows us to select the fundamental frequency as opposed to higher harmonics, which may show higher maxima if we were to rely solely on the audio spectrum.

The fundamental frequency of vocalization is clearly above the noise floor (typically 37 ± 30 dB), but some spurious noises are polluting the spectra.

As it can be expected, we found a significant difference between men and women's pitch. The distribution of frequency is presented for both men and women in Fig. 2.

In a few cases, the subject voice pitch varies depending on the type of vocals (e.g., /a/, /o/, or /z/ sound). This highlights the dependence of the spatial distribution of vibrational energy induced by the vocal cords at the surface of the thorax with the frequency of the emitted vocalization. One of the subjects in the cohort (depicted on Fig. 3) managed to cover a range of pitch of almost 50 Hz, from about 100 Hz to about 150 Hz. This series of acquisitions shows vibratory energy move up to the top of the thorax as the frequency of the voice increases. When summing together the whole series into a unique average vibratory map, an "acoustic shadow" of the lungs is clearly identifiable (see Fig. 3 map on the far right). On this series, it can also be seen that the vibration energy is more prominent on the right than on the left.



FIG. 1. Sample of individual and cross frequency spectra of the audio signal and the thorax surface vibration Doppler signal recorded during the /z/-vocalization of a healthy subject.

04 April 2024 15:18:38



Healthy Volunteers' Vocal Frequency Distribution



In order to check whether observations of asymmetry on the latter single subject can be generalized to the whole cohort, let us consider now regional vibrational statistics (see Fig. 4). The region of interest is now divided into four quadrants. For each quadrant, the Fig. 4 box plot shows the first percentile, the first-to-second quartile boundary, the median, the third-to-fourth quartile boundary, and the 99th percentile boundary. Except for the first percentile representative of noisy acquisitions (i.e., low signal-to-noise ratio), all other boundaries increase in the following quadrant order: bottom left, bottom right, top left, top right. In other words, there is both a left-right and a top-bottom asymmetry. Overall, top quadrants vibrate more than bottom quadrants, while right quadrants vibrate more than left quadrants.

Finally, to facilitate the visualization of the whole cohort, let us consider the vibration barycenter (i.e., center of mass). It allows

us to summarize the spatial and spectral information of the whole cohort shown in Fig. 5. The figure shows barycenter locations on top of either an average male [Fig. 5(a)] or female [Fig. 5(c)]. The intermediate axis [i.e., Fig. 5(b)] shows all the acquisitions together in an ordinate barycenter location as a function of detected vocal frequency. Considering the distribution of vocalization frequencies shown in Fig. 2, the barycenter cluster below 200 Hz (from violet blue to light green) mostly corresponds to male subjects, whereas the barycenter cluster beyond 200 Hz (from light-green-yellow to red) mostly corresponds to female subjects. As far as the male subjects are concerned, the dependency of the barycenter location with frequency seems to match the one of the singled-out subjects depicted in Fig. 3. The observed vibration seems to undergo a rather abrupt transition around 120 Hz that corresponds to an upward displacement of the vibration barycenter. There also seems to be some



04 April 2024 15:18:38



FIG. 4. Regional statistics: box plot distributions of the median rms speed for the four quadrants across the cohort are represented.



FIG. 5. Vibration energy barycenter locations represented over (\mathcal{O}^* male and \mathcal{P} female) average thorax profiles [left (a) and right (c), respectively]. The ordinates of the vibration energy barycenter are also plotted as a function of frequency (b) for both populations. Male and female ordinate distributions are represented as boxplots [(a') and (c'), respectively]. Hence, the ordinate is shared across all axes (a, a', b, c, c'). Each data marker corresponds to a specific vocalization (i.e., acquisition).

transitional behavior on the female cohort around 240 Hz. As can be seen on the average back profile shown in Fig. 5(c), female subjects are shorter ($168 \pm 8 \text{ cm} < 178 \pm 7 \text{ cm}$). The female barycenter cluster is accordingly shifted to lower ordinates ($-3.3 \pm 2.6 \text{ cm}$). Note that the male subject barycenter cluster (from dark blue to light green) is not laterally centered on the thorax. Thus, it also lightly reflects the aforementioned (i.e., Fig. 4) left-right asymmetry.

DISCUSSION

This study shows that the vibrational energy mapping device that we developed and previously validated to measure respiratory movements and cardiac activity¹⁰ can detect vocalization-driven vibrations over the entire surface of the human thorax and produce a four-dimensional mapping of these vibrations. This indicates that our contactless device can record quantitative information similar in nature to that derived from the clinical palpation of tactile fremitus.

Vibrational energy mapping could be achieved for all the participants, although some acquisitions had to be discarded. This issue is not a limitation of the technique itself but rather related to its prototypic implementation using off the shelf components. Another acquisition issue is related to a manual triggering delay that is too long with respect to the duration of the acquisition, resulting in some recordings being truncated. Finally, it became apparent over the duration of the study that factors related to the participants' vocalizations (strength, pitch, and stability in time), morphological characteristics, and posture could lead to an SNR insufficient to adequately detect the vocalization-induced vibrations. Technological improvements and protocol standardization will be needed to correct these limitations.

- The acquisition is triggered manually and only lasts a few seconds, while triggering delay is long (few seconds) and not precisely known, so some vocalizations are partially/completely missing.
- The overall acquisition signal depends on the specifics of the subject's vocalization including the posture of the subject; its strength, its pitch, and its time stability may lead to an insufficient SNR for the system to detect the induced vibrations.

Finally, some of these limitations will be lifted in the next version of the system so that the signal acquisition feasibility will be improved significantly.

The spatial distribution of vibrational energy was found to be asymmetric to the benefit of the right side of the chest. This is in line with the larger size of the right lungs compared to the left one. This observation represents an initial validation of the technology's ability to detect differences in the structural properties of the respiratory system. The spatial distribution of vibrational energy also appeared frequency dependent, with a shift of vibrational energy toward the upper part of the thorax with increasing vocal frequency. This could be due to the anisotropic nature of the lung parenchyma, known to be responsible for the basal parts of the lungs being denser than the upper parts. The corresponding vertical mechanical gradient is bound to translate into a similar gradient in acoustical properties. The frequency dependence of the vibrational energy distribution could also result from the geometrical changes inherent to the branched nature of the bronchial tree.

The present pilot observations open new perspectives regarding the understanding of the mechanisms underlying vocal fremitus as assessed during chest palpation. It should however be kept in mind that the AUSMC characterizes surface vibrations unopposed, whereas during physical examination, a counterpressure is exerted by the hands of the physician. It can be anticipated that the AUSMC should easily detect gross abnormalities in vibratory transmission, as in the case of pneumothorax or pleuritis where the interposition of air or liquid between the lungs and the chest wall completely abolishes vocal fremitus, or in the case of important atelectasis where it typically increases it. The right-to-left asymmetry of vibratory maps that we observed suggests that it should be the case. Beyond these relatively caricatural situations, we believe that the AUSMC should be able to detect focal abnormalities. If verified, such sensitivity would be invaluable in the context of the early, pre-radiological, orientation of patients presenting with respiratory symptoms.

We used vocalizations to produce surface chest vibration. This approach has the merit of replicating clinical practice. It does not resort to any extra apparatus, hence a very simple and safe usage. However, vocalizations produce complex frequency assemblages that can depend on many subject-related factors. Such factors are liable to contribute to the frequency dependence of the acoustic energy spatial distribution and most likely explain the marked gender differences that we observed. One possible way around this issue would be to produce surface chest vibrations not with vocalizations but with known vibratory inputs. Such inputs can be produced using the forced oscillation technique (FOT) that consists in applying a range of low frequency acoustic waves up to 40 Hz at the airway opening while measuring the corresponding flow and pressure responses, to estimate airway resistance and lungs impedance. Combined with a surface vibrometer, the FOT can be used to compute the transfer function between the acoustic source and the skin. Likewise, Royston et al. showed that combining laser scanning vibrometry and the FOT allowed them to detect the presence of pneumothorax.¹¹ A similar approach was also investigated by Aliverti et al. using photogrammetry.¹

In conclusion, we believe that the data that we present are sufficient to pursue the evaluation of the AUSMC as a putative future diagnostic tool. It will also be necessary to develop experimental study and physical models to understand how surface vibrations relate to structural and functional respiratory abnormalities, whether they are generated naturally (vocalization) or artificially (FOT).

ACKNOWLEDGMENTS

This work was partially supported by the "Association pour le Développement et l'Organisation de la Recherche en Pneumologie et sur le Sommeil (ADOREPS)."

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Frédéric Wintzenrieth: Data curation (lead); Software (lead); Writing - original draft (supporting); Writing - review & editing (supporting). Mathieu Couadecorr-auth: Data curation (equal); Software (equal); Writing - original draft (equal); Writing - review & editing (equal). Feizheun Lehanneur: Investigation (equal); Visualization (equal). Pierantonio Laveneziana: Investigation (equal); Writing - review & editing (supporting). Marie-Cécile Niérat: Investigation (equal); Writing - review & editing (supporting). Nicolas Verger: Investigation (equal); Writing - review & editing (supporting). Mathias Fink: Supervision (supporting); Writing review & editing (supporting). Thomas Similowski: Conceptualization (equal); Investigation (equal); Methodology (equal); Project administration (equal); Supervision (equal); Writing - review & editing (supporting). Ros Kiri Ing: Conceptualization (equal); Investigation (equal); Project administration (equal); Supervision (equal); Writing - review & editing (supporting).

DATA AVAILABILITY

The data underlying this study are not publicly available due to ethical considerations associated with the nature of the research involving human subjects. Participant privacy and confidentiality have been paramount throughout the research process. The data contain potentially identifying and sensitive patient information, and sharing this information could compromise the privacy of the participants. Ethical approval for this study, including data sharing restrictions, was obtained from CPP SUD-OUEST ET OUTRE-MER II, decision reference: 2-18-2; March 08, 2018.

REFERENCES

¹B. Maitre, T. Similowski, and J. Derenne, "Physical examination of the adult patient with respiratory diseases: Inspection and palpation," Eur. Respir. J. **8**(9), 1584–1593 (1995).

²J. E. Wipf *et al.*, "Diagnosing pneumonia by physical examination: Relevant or relic?," Arch. Intern. Med. **159**(10), 1082–1087 (1999).

³X. Zhang, D. Maddipatla, B. B. Narakathu, B. J. Bazuin, and M. Z. Atashbar, "Development of a novel wireless multi-channel stethograph system for monitoring cardiovascular and cardiopulmonary diseases," IEEE Access 9, 128951–128964 (2021). ⁴M. Yigla, M. Gat, J.-J. Meyer, P. J. Friedman, T. M. Maher, and J. M. Madison, "Vibration response imaging technology in healthy subjects," Am. J. Roentgenol. 191(3), 845–852 (2008).

⁵R. Morice, C. Jimenez, G. Eapen, D. Bethancourt, and L. Keus, "Vibration response imaging (VRI) for predicting postoperative lung function in patients with lung cancer," Chest **132**, 486A (2007).

⁶M. Blanco, R. Mor, A. Fraticelli, D. P. Breen, and H. Dutau, "Distribution of breath sound images in patients with pneumothoraces compared to healthy subjects," Respiration 77(2), 173–178 (2009).

⁷Y. Kim, Y. Hyon, S. Lee, S.-D. Woo, T. Ha, and C. Chung, "The coming era of a new auscultation system for analyzing respiratory sounds," BMC Pulm. Med. **22**, 119 (2022).

⁸N. Jeger-Madiot, "Conception d'un imageur ultrasonore dans l'air pour la mesure des ondes basses-fréquences en surface du corps humain," Ph.D. thesis, Université Pierre et Marie Curie - Paris VI, 2017. ⁹M.-C. Niérat, P. Laveneziana, B.-P. Dubé, P. Shirkovskiy, R.-K. Ing, and T. Similowski, "Physiological validation of an airborne ultrasound based surface motion camera for a contactless characterization of breathing pattern in humans," Front. Physiol. **10**, 680 (2019).

¹⁰P. Shirkovskiy, A. Laurin, N. Jeger-Madiot, D. Chapelle, M. Fink, and R. K. Ing, "Airborne ultrasound surface motion camera: Application to seismocardiography," Appl. Phys. Lett. **112**, 213702 (2018).

¹¹H. Palnitkar *et al.*, "Sound transmission in human thorax through airway insonification: An experimental and computational study with diagnostic applications," Med. Biol. Eng. Comput. **58**(10), 2239–2258 (2020).

¹²A. Aliverti, R. L. Dellacà, and A. Pedotti, "Transfer impedance of the respiratory system by forced oscillation technique and optoelectronic plethysmography," Ann. Biomed. Eng. **29**(1), 71–82 (2001).