

Respiratory Inductance Plethysmography for Automated Swallowing Detection

Eric Tatulli, Univ. Grenoble Alpes, CNRS, CHU Grenoble Alpes, Grenoble INP, TIMC-IMAG, La Tronche, France

Julie Fontecave-Jallon, Univ. Grenoble Alpes, CNRS, CHU Grenoble Alpes, Grenoble INP, TIMC-IMAG, La Tronche, France

Pascale Calabrese, Univ. Grenoble Alpes, CNRS, CHU Grenoble Alpes, Grenoble INP, TIMC-IMAG, La Tronche, France

Pierre-Yves Gumery, Univ. Grenoble Alpes, CNRS, CHU Grenoble Alpes, Grenoble INP, TIMC-IMAG, La Tronche, France

ABSTRACT

In the context of remote medical monitoring of swallowing, the authors investigate the potential of non-invasive respiratory inductance plethysmography (RIP) technique to automatically detect swallow events in a wide context of respiration and phonation. Signal acquisitions were carried out on 5 healthy volunteers equipped with RIP and electroglottograph as swallowing reference. They were asked for spontaneous breathing, speaking and diverse bolus ingesting. The RIP signal was then segmented into cycles, each cycle being annotated according to one of the three class of interest, respectively ventilation i.e. spontaneous breathing (1257 cycles), swallowing (221) and phonation (216). Automated classification was performed using quadratic discriminant analysis. Focusing on swallowing class, the authors achieve an accuracy of 79% from the full wide protocol. It increases up to 86% with prior removal of vocalizations. These preliminary results in healthy subjects make RIP a promising candidate as a non-invasive and convenient technology for medical follow-up of swallowing.

KEYWORDS

Automatic Classification, Deglutition, Discriminant Analysis, Healthy Subjects, Non-Invasive Sensors, Remote Monitoring, Signal Processing, Speech, Ventilation

INTRODUCTION

Swallowing disorders exist at any age. However, its prevalence increases in elderly population associated with an increase of morbidity and mortality (Schindler & Kelly, 2002). Consequences of dysphagia are multiple: pneumonia due to aspiration, dehydration, but also psychological problems such as anxiety at mealtimes and social isolation (Tibbling & Gustafsson, 1991). Common techniques used to detect swallowing disorders are based on clinical evaluation such as cervical auscultation and gag reflex in combination with water/food swallow clinical tests and X-ray tests (Mathers-Schmidt & Kurlinski, 2003). The gold standard to characterize swallowing disorders and especially oropharyngeal dysphagia is a method based on irradiation called videofluoroscopy (Fattori et al., 2016). However, the usual clinical evaluations of cervical auscultation and gag reflex have been documented as having limitations in swallow assessment (Leder, 1996; Leslie, Drinnan, Finn, Ford,

DOI: 10.4018/IJEHMC.2020040104

This article, originally published under IGI Global's copyright on April 1, 2020 will proceed with publication as an Open Access article starting on January 25, 2021 in the gold Open Access journal, International Journal of E-Health and Medical Communications (converted to gold Open Access January 1, 2021), and will be distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

& Wilson, 2004). In addition, videofluoroscopy, as well as other instrumental methods (endoscopy, scintigraphy), besides being invasive, focuses on the examination of the deglutition of one bolus and does not allow a daily monitoring of swallowing. Therefore, even if videofluoroscopy is essential in the diagnosis and planning of rehabilitation treatment for dysphagia and remains the test of choice for assessing pre-swallowing spillage and tracheobronchial inhalation, it may be of interest to investigate user-friendly and non-invasive techniques which may allow a complementary longitudinal monitoring of swallowing. The potential of non-invasive approach in this context must also be balanced by the possible shortcomings; the phenomenon measure is less well located and less targeted. Data obtained are more equivocal and therefore more complex to analyze. The first question is therefore to assess whether a non-invasive method is competitive enough to detect deglutition occurrences in peripheral signals. Longitudinal monitoring of parameters from the non-invasively detected swallows, such as the numbers or durations of occurrences, can be a first analysis of swallowing disorders.

Swallow detection from non-invasive techniques have been investigated in many studies, considering electromyography (Vaiman & Eviatar, 2009), microphone (Walker & Bhatia, 2011; Golabbakhsh et al., 2014), a combination of both systems (Amft & Troster, 2006), electroglottography (EGG) (Farooq, Fontana, & Sazonov, 2014) and accelerometry (Sejdic, Steele, & Chau, 2009; Steele, Sejdic, & Chau, 2013; Mohammadi, Steele, & Chau, 2016). However, placing sensors on the face or neck can lead to problems of comfort and use, which makes complicated a prolonged use. Therefore, other types of comfortably wearable sensors have also been considered, such as piezo belts (Dong & Biswas, 2016) or jackets using Respiratory Inductive Plethysmography (RIP) (Moreau-Gaudry, Sabil, Baconnier, Benchetrit, & Franco, 2005; Kenny, Casas, & McPherson, 1989; Casas, Kenny, & McPherson, 1994; Geddes, Chadwick, Kent, Garbin, & Hartmann, 2010).

RIP is a non-invasive method, usually dedicated to respiratory monitoring, that can be investigated to analyze deglutition, by studying the impact of deglutition on RIP signal. RIP monitors thoracic and abdominal cross-sectional area changes during ventilation, thanks to 2 sinusoidal wires positioned around the thorax above the line of the nipple and around the abdomen at the umbilical level. After calibration, the combination of thoracic and abdominal signals allows obtaining the ventilatory volume. It has been shown that RIP derivative is a proper estimation of the airflow signal (Calabrese, Besleaga, Eberhard, Vovc, & Baconnier, 2007; Eberhard, Calabrese, Baconnier, & Benchetrit, 2001). A preliminary work on elderly people (Moreau-Gaudry et al., 2005) has demonstrated the interest of using RIP to detect swallowing. Moreover, several works have studied the coordination of infants' suck-swallow-breathing patterns using RIP (Kenny et al., 1989; Casas et al., 1994; Geddes et al., 2010). In Geddes et al. (2010), the authors showed that ultrasound imaging of swallowing correlates well with swallow apnea manually detected on RIP signal. When studying the relationship between breathing and deglutition, it was observed that a physiological apnea occurs as a brief closure of the larynx to protect the airway from the aspiration of the bolus ingested. This swallow apnea can be detected as zero segments on the respiratory airflow (Moreau-Gaudry et al., 2005; Hardemark Cedborg et al., 2010). As RIP volume signal corresponds to airflow integral, the swallow apnea will result on RIP signal in plateaus during ventilatory cycles (Smith, Wolkove, Colacone, & Kreisman, 1989; Preiksaitis & Mills, 1996; Tarrant, Ellis, Flack, & Selley, 1997; Terzi et al., 2007; Camargo, Ono, Park, Caruso, & Carvalho, 2010; McFarland, Martin-Harris, Fortin, Hill, & Armeson, 2016; Fontecave-Jallon & Baconnier, 2016). The induced modifications of the breathing pattern are then mainly characterized by differences on RIP signal shapes (Smith et al., 1989; Tarrant et al., 1997; Matsuo & Palmer, 2009) and by an increase of the respiratory cycle duration. Breathing pattern may also be changed by speech. It has been observed on RIP signal (Binazzi et al., 2006; Rochet-Capellan & Fuchs, 2013), that when speaking, breaths consist of a rapid inspiration with a slower expiration performed with essentially constant flow. This will therefore influence the signal shapes and the physiological respiratory variables extracted from RIP signal.

In most studies previously cited (Preiksaitis & Mills, 1996; Terzi et al., 2007; Camargo et al., 2010; McFarland et al., 2016), RIP signal is generally used to locate swallowing in the

ventilatory cycle and to observe if the swallow event occurs during in- halation or exhalation. In most cases, none specific processing is carried out on RIP signal. A few works have proposed to detect swallowing on RIP signal, but it is manually realized (Kenny et al., 1989; Casas et al., 1994; Geddes et al., 2010). A first attempt of automatic detection has been carried out by (Moreau-Gaudry et al., 2005) on the derivative of RIP signal. Furthermore, automatic detection of swallowing has been investigated for some other non-invasive techniques, such as accelerometry (Steele et al., 2013; Mohammadi et al., 2016), chest belt piezo-device (Dong & Biswas, 2016) or acoustic measurements (Golabbakhsh et al., 2014).

Therefore, the purpose of the present work is to propose an automatic swallowing detection on RIP signal, in order to investigate RIP technology as a non-invasive automatic remote medical monitoring of swallowing events. The present work is a preliminary study on healthy volunteers, which main purpose is to highlight if swallowing events can be automatically discriminated from ventilation and speech. Based on a wide range of experimental swallow and speech tests, a discriminant analysis is considered for an automatic classification of breathing, swallowing and speech events detected on RIP signal.

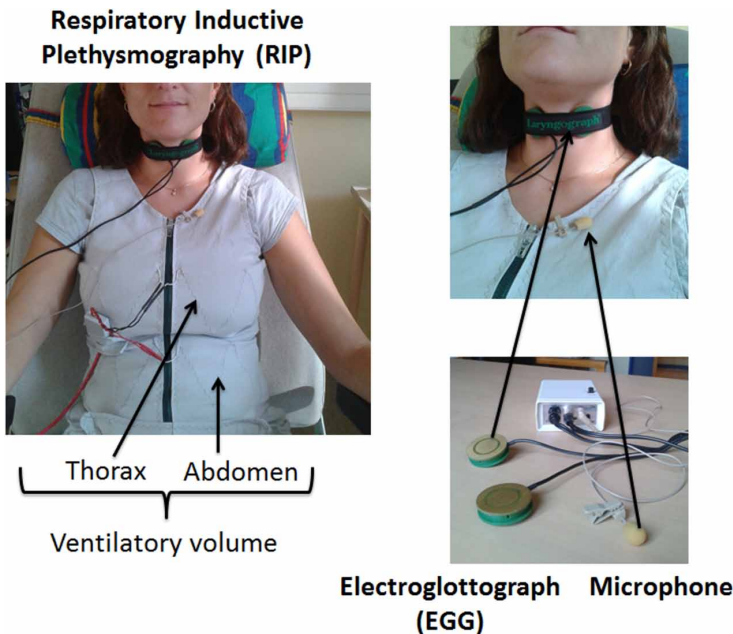
METHODOLOGY

Equipment and Protocol

Data acquisitions were conducted in TIMC-IMAG Laboratory (La Tronche, France), in a dedicated place of biomedical research. The study was approved by the relevant ethics committee (CHU Grenoble Alpes). Five healthy volunteers (2 men and 3 women, from 22 to 64 years old) participated after providing informed consent. This number of subjects is comparable to (Dong & Biswas, 2016).

Equipment: Each subject was wearing a connected Respiratory Inductive Plethysmography (RIP) jacket above its clothes and was equipped with a microphone on the jacket and two electrodes positioned on the neck either side of the thyroid cartilage and connected to an electroglottograph device (Figure 1). Participants were in a meal-time seated position and were asked not to move.

Figure 1. Experimental protocol: Subjects are equipped with a RIP jacket, an electroglottograph neckband and a microphone



Thorax (THO) and abdomen (ABD) cross sectional area changes were recorded with the computer-assisted RIP vest (Visuresp, RBI, Meylan, France). During 3-4 minutes at the beginning of each recording, breathing was also simultaneously recorded with a flowmeter (Fleish head no.1, Emka Technologies, Paris, France) and a differential transducer (163PC01D36, Micro Switch, Honeywell, United States) placed on a face mask. The Laryngograph EGG-A100 device recorded the variations in the conductance between the electrodes, produced by changes in the nature and area of vocal fold contact. The acquired signal is noted EGG signal. The acoustic signal was recorded according to an audio microphone available with the EGG-A100 device. All signals were synchronously digitized at a rate of 10000 Hz with a PowerLab acquisition system and LabChart Soft-ware V7 (ADInstruments Pty Ltd).

Experimental protocol: Subjects were asked to follow an experimental protocol made of several steps. Indeed, many studies in the literature (Martin-Harris et al., 2005; Wheeler Hegland, Huber, Pitts, Davenport, & Sapienza, 2011) have documented the variation in the swallowing pattern based on bolus volume and presentation. The experimental swallow tests have deliberately been largely chosen (bolus nature and size, and types of container) and the protocol also included phonation events.

Subjects were first asked to breathe calmly and naturally with the face mask; this is considered as the calibration phase necessary in the pre-processing to calibrate RIP according to the ventilatory airflow. Then the mask was removed and 3 minutes of spontaneous breathing was recorded. Considering swallowing protocol, they were then asked to repeat 6 times a forced swallow of saliva. Intervals of 30 seconds were kept between each deglutition. Then the subject drank different volumes of water (5, 10, 15, or 100 ml) in random order through a straw or by glass. Different boluses of stewed fruit were also ingested in a random order with a spoon (5 ml) and a flask (90 ml). Each type of bolus was repeated 3 times. The subjects were instructed to breath in a regular manner between each bolus. As mentioned in introduction, the proposed protocol included a phonation part. Subjects were then asked to pronounce and repeat 3 times short sentences and longer sequences. The subjects were instructed to breathe in a regular manner between each phonation.

The average length of recording was 35 minutes. During each recording session, the starting times of all events were manually annotated by the investigator. Each event was annotated with its detailed description but also with a more macroscopic one among the 3 main categories (ventilation, swallowing or phonation).

Data Pre-Processing and Labeling

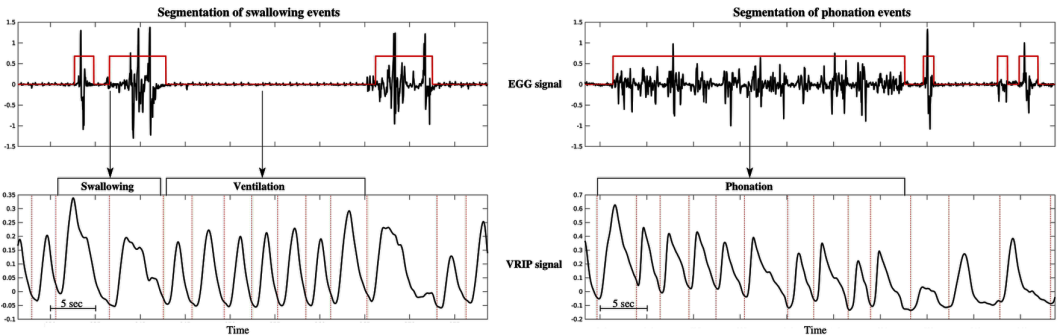
Events segmentation from EGG data: In this paper, the segmentation of specific events was processed from EGG data as illustrated in Figure 2 (top), the microphone being only used as further confirmation in case of dubious origin (e.g. vocalization or not) of a given detection. The annotation was performed under a strict conservative approach, in order not to feed the classifier with wrong input information:

- The EGG events were segmented using signal-to-noise criterion (set to 4 in the following), the noise variance being estimated from calibration data, in which a regular breathing without swallowing is insured;
- Each segmented event was annotated according to the manual comments recorded during the experiment. Non-attributable detection and unwanted perturbations (strong movement, coughing, etc.) of the EGG signal were subsequently discarded.

Note that an automated segmentation algorithm (see also (Sejdic et al., 2009; Wang & Willett, 2003)) has also been developed but is out of the scope of this paper and will be presented in a dedicated article.

RIP volume reconstruction (VRIP): Thorax and Abdomen signals were linearly combined to compute the ventilatory volume, noted VRIP, the coefficients being estimated from a least square fit of the ventilatory flow recorded with the mask during the calibration phase previously described

Figure 2. Schematic description of the proposed approach. Top: for each subject independently, events are manually segmented from EGG data. Bottom: by matching time frames, each cycle on RIP volume signal (VRIP) is associated to one of the three defined class: ventilation (no detection on the EGG signal), swallowing or phonation. The set of annotated VRIP cycles is then used as input for classification.

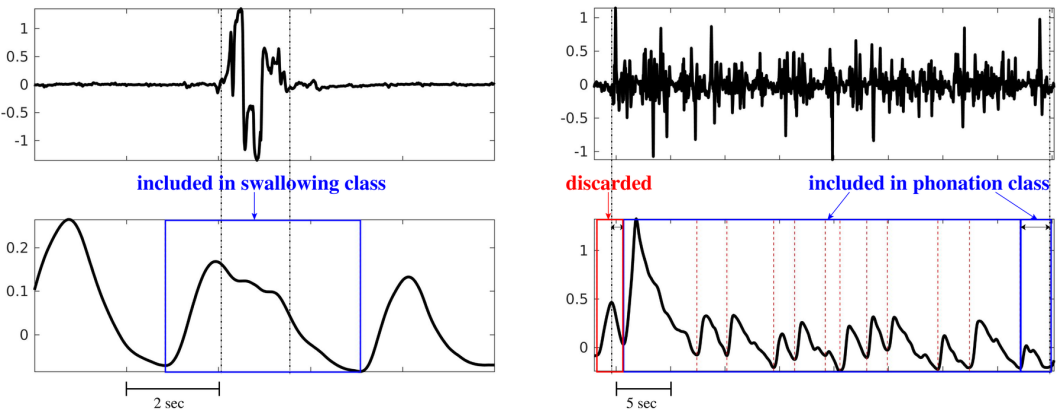


in the experimental protocol (Eberhard et al., 2001; Calabrese et al., 2007). The obtained signal was then divided into cycles where cycles starts are estimated from negative to positive zero-crossing of the airflow (numerical derivative of the signal).

Cycle class eligibility rule: As shown in Figure 2 (bottom), each VRIP cycle was then associated to one of the corresponding three main classes, respecting the following rules of eligibility criterion illustrated in Figure 3:

- If the time interval of a given event lies within one VRIP cycle, this event is associated to the corresponding event class;
- If the event time interval crosses several VRIP cycles, cycles that are fully included in the interval were affiliated to the matching class. Considering both edge cycles, if the part of the event time interval that lies within the VRIP cycle is greater than 50% of the cycle time length, then the same class was applied. If not, these cycles were removed from the classification process.

Figure 3. Example of eligibility criterion for VRIP cycles. Left: the EGG event time interval lies within one VRIP cycle which is included in the corresponding class (here swallowing). Right: the EGG event time interval crosses several VRIP cycles. Inside cycles are included in the class of interest (here phonation). Right edge VRIP cycle is also included because the EGG event occurs during most of the cycle duration (as shown by the horizontal arrow). On the contrary left edge VRIP cycle is discarded because the EGG event is present in less than 50% of the cycle time length.



Eventually it ended up for the 5 subjects with a total VRIP cycle number of 221 swallows (comparable e.g. to (Sejdic, Falk, Steele, & Chau, 2010)), 216 phonations, and 1257 breathing cycles. Figure 4 displays for each class random examples of VRIP cycle profiles extracted from the continuous RIP signal of a single subject (namely JF). One can see that within each class VRIP profiles are showing strong similarities whereas exhibiting noticeable differences between two separate groups. As illustrated in Figure 5, prototypical VRIP profiles are extracted for each class respectively: (i) a smooth “Gaussian-like” shape in case of regular ventilation, (ii) the presence of a full or partial plateau witnessing apnea, which is expected during (non-pathological) swallowing events, and (iii) a specific linear slope during the breathing out part of the vocalization (Binazzi et al., 2006). In the following, the authors explore how these profiles can provide a robust basis to perform classification and automatic detection of swallowing occurrences.

PROPOSED CLASSIFICATION APPROACH

Features Extraction

Before properly apply classification, each VRIP cycle needs to be described with a vector (so-called feature) as representative as possible, and which dimension remains reasonable with respect to the number of samples. The authors propose here to combine both physiological and morphological parameters:

Figure 4. Random subsets of VRIP ventilatory (left), swallowing (center) and phonation (right) cycles (taken from JF dataset). The vertical axis is fixed for all little squares and corresponds to 0.5 liters. The horizontal axis (in time) is normalized.

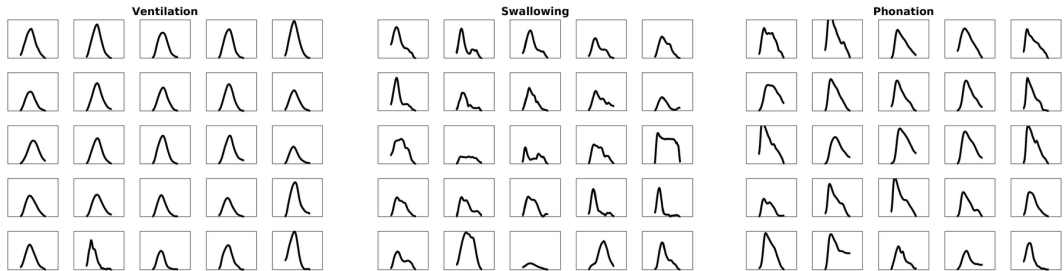
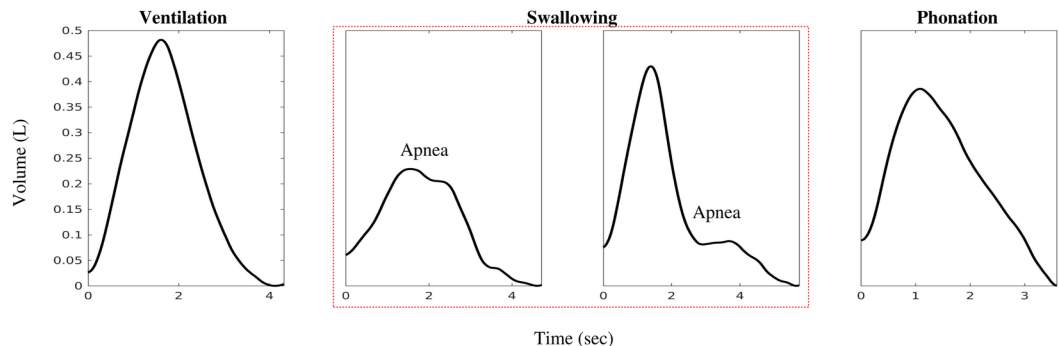


Figure 5. Prototypical VRIP profiles of respectively ventilatory, swallowing and phonation events. Two profiles of swallowing are shown, depending on the swallow onset (during inhalation or during exhalation). One can notice a plateau associated to apnea during swallowing, while phonation displays a typical linear slope during breathing out.



- They have extracted from each VRIP cycle 3 classical physiological variables associated to ventilatory activity: the tidal volume defined as the cycle breathed-in air volume, as well as inspiratory and expiratory durations;
- Moreover, a good descriptor of ventilatory cycles morphology is given by the so-called ASTER vector (Bachy, Eberhard, Baconnier, & Benchetrit, 1986) which is computed by the four first harmonics of the Fourier transform of the airflow cycles (i.e. derivative of the VRIP cycles). This makes up 8 morphological parameters (real and imaginary parts of each harmonic). Another parameter is given by the fraction of power contained in the four harmonics. It represents how much a given cycle spreads out the canonical respiratory profile. Note that in place of ASTER features, a mere sampling of the normalized profiles has also been tried, however leading to slightly lower classification performance.

Furthermore, it has been shown in the case of resistive loading (Laouani, Rouatbi, Saguem, & Calabrese, 2016) that shape differences between rib cage and abdomen signals could be a good indicator of changes in the level of bronchial obstruction in patients with moderate obstructive syndrome.

The authors have therefore added a last descriptor, computed from the quadratic distance between rib cage and abdomen normalized profiles, as followed by (Laouani et al., 2016).

In total, VRIP cycles are then featured by 13-dimensioned vectors, used to feed the classifier.

Classification Method

VRIP cycles are categorized with discriminant analysis technique (Klecka, 1980) where input features of each class are assumed to follow a multivariate normal probability density. So-called within-class covariance matrices of the normal density functions are estimated empirically from the data, enabling to calculate for each sample the probability of belonging to a given class and attributing the class for which this probability is maximal.

Similar to (Steele et al., 2013; Dong & Biswas, 2016) classification performance is calculated using leave-one-out validation (Jain, Duin, & Mao, 2000): a dataset of N samples is divided into a training set of $N - 1$ samples (from which the probability function is computed) and the remaining one is used for test. The class of this latter, supposedly unknown, is estimated by the classifier (i.e. by discriminant analysis). Finally, the estimated class is compared to the true class for validation. This process is repeated until all samples have been tested. Given the idiosyncrasies exhibited during the various breathing sequences of the protocol, the validation has been conducted for each subject separately.

Metrics

For each class, the numbers of true positives (TP, that is a sample correctly attributed to the class of interest), false positives (FP, that is a sample attributed by mistake to the class of interest) and false negatives (FN, a sample that should have been credited to the class of interest but was not) are counted. From these numbers metrics that are commonly used for classification analysis are computed:

- Sensitivity (also recall) $S_e = \frac{TP}{TP + FN}$. It measures the fraction of samples that are correctly assigned to a given class. Note that the denominator corresponds to the total number of sampling truly belonging to this class;
- Precision (also specificity) $P_r = 1 - \frac{FP}{TP + FP}$, where the second term is the rate of false positive, for each class.

PERFORMANCE AND DISCUSSION

Results

Table 1 shows the classification performance for each subject individually. First of all, one can see that individual results are achieved with small (a few percent) inter-subject dispersion, both for sensitivity and precision values and for the three class of interest. This assesses the robustness of the proposed approach despite idiosyncratic breathing behaviors. Note that the variation of performance across subjects is comparable to (Dong & Biswas, 2016) and results from the different ways people speak, drink or eat.

Focusing on swallowing events, average sensitivity and precision of 86% and 71% are respectively obtained, leading to an accuracy (average of sensitivity and precision) of $Acc = 79\%$. These results lie within the same magnitude than recently published studies making use of different detection techniques. They are whether comparable ($Se = 87\%$, $Pr = 75\%$ with the chest belt piezo-device (Dong & Biswas, 2016), or $Acc = 82\%$ from acoustic measurements (Golabbakhsh et al., 2014)), whether in the lower range ($Acc = 86\%$, 89% , 90% from accelerometry (Mohammadi et al., 2016), previous RIP experiment (Moreau-Gaudry et al., 2005) and EGG data (Farooq et al., 2014) respectively).

Although it remains interesting to confront these approaches in terms of rough estimates, it is however not possible to undertake a deeper comparative analysis since each protocol is unique, that is following specific experimental requirements whether in terms of age (e.g. focusing on the elderly for Moreau-Gaudry et al. (2005)), pathology (targeting Parkinson's disease in Golabbakhsh et al. (2014)), type of bolus (thin liquid barium in Mohammadi et al. (2016)) or food consistency (solid food ingested in Dong & Biswas (2016)). In the present study, the authors' prejudgment was to scan a rather large variety of swallowing conditions as well as a richer environment by including vocalizations. This last point is of high importance because it adds a third class that has to be retrieved by automated classification, thus greatly increasing the complexity of the detection process whereas most of the above cited experiments are restricting people from talking or are considering vocalization as disturbances that are discarded from their dataset (Steele et al., 2013; Mohammadi et al., 2016).

This issue is explicitly tackled in Sejdic et al. (2010) in the context of accelerometry, where they demonstrate that removing vocalizations enables to improve the accuracy of their technique by 55%. Reprocessing the results with prior removal of phonation events (assuming e.g. that the RIP jacket is equipped with a microphone from which the vocalization sequences can be segmented), the performance significantly increases, as summarized in Table 2. Noticeably, by narrowing the study to two class of interest, the authors straightforward achieve top-ranking performance in sensitivity,

Table 1. Classification performance for each subject individually trained and tested with left-one-out validation method. V, S and P denote respectively ventilation, swallowing and phonation class.

Subjects						
	JF	PC	PB	TF	NR	Average $\pm$ SD
Sensitivity (%)						
V	97	95	95	95	95	95 $\pm$ 1
S	91	84	90	77	89	86 $\pm$ 5
P	93	86	83	83	91	87 $\pm$ 4
Precision (%)						
V	100	97	100	97	97	98 $\pm$ 1
S	79	69	65	69	75	71 $\pm$ 5
P	90	97	93	86	100	93 $\pm$ 5

Table 2. Classification performance with prior removal of phonation. Only 2 classes are remaining: ventilation and swallowing.

Subjects						
	JF	PC	PB	TF	NR	Average±SD
Sensitivity (%)						
V	96	96	95	95	94	95±1
S	98	92	97	88	98	95±4
Precision (%)						
V	100	98	100	98	100	99±1
S	79	82	76	77	71	77±4

with an average of $95\% \pm 4\%$ for the detection of swallows. The precision of the swallowing class – though slightly improved –, keeps nonetheless a comparatively lower ranking. The authors explain this behavior by their deliberated choice of involving a large variety of combined bolus types (water, stewed fruits) containers (glass, straw, spoon, flask) and volumes (5, 10, 15, 100 ml), thus driving to large variances of swallowing features as discussed in more details below.

Variability of Swallowing Class

From inductive reasoning, the authors postulate that spanning various swallowing conditions in their protocol drives the classifier to predominantly favor this latter class when dealing with outliers. If such an assumption is true, two joint effects are expected and indeed observed: (i) a fairly good sensitivity because outliers belonging to swallowing class are mostly correctly classified (i.e. few swallowing FN) and (ii) a relatively low precision because outliers not belonging to swallowing class are mostly wrongly attributed to this latter (i.e. many swallowing FP).

Verifying this hypothesis requires to compare the variability – i.e. the covariance matrices – of each class. In multivariate statistics, this is commonly done by performing a simultaneous diagonalization (Fukunaga, 2013) – or equivalently by solving the relative eigenvalue problem (Welling, 2005) – of the two matrices of interest: in the current case the covariance matrix of the swallowing class with respect to that of the ventilation or the phonation class, respectively. This process defines a new vector basis where each so-called relative eigenvalue can be interpreted as the maximum ratio of variances (here swallowing vs. ventilation/phonation variance ratio) in the corresponding eigenvector direction (Flury, 1985). Figure 6 (left) shows for PC subject the sorted eigenvalues when comparing swallowing features to ventilation (top) and phonation (bottom) features respectively. For sake of clarity, the inverse value for eigenvalues below one (i.e. when the maximum variance ratio stands in favor of the ventilation/phonation class) has been plotted with white color. It obviously exhibits a greater amount of variability in the swallowing class, both relatively to ventilation and phonation groups (though in a less pronounced fashion for the latter). This is quantitatively confirmed by computing the respective generalized variance ratios – defined as the product of the relative eigenvalues (Bookstein & Mitteroecker, 2014) – of roughly 4.107 and 103, that is larger than one by several magnitudes.

To experimentally confirm this analysis, a new protocol has been set-up involving four subjects among which the authors were able to record two participants of the first experiment (JF and PC). The main concern is here to restrict swallowing (and phonation) events to a single type which duration lies within a single VRIP cycle: 5ml water spoons and one short repeated sentence respectively, in the purpose of reducing intra-class variability as much as possible. Attention has been carried out to produce datasets with roughly the same number of cycles in each class than in the first protocol (~ 40 swallows, 40 phonations and 250 breathings by subject). Table 3 reports the performance achieved in this narrow configuration. Focusing on the precision value for the swallowing class,

Figure 6. Illustration of the variability of the swallowing class, for PC subject. Each plot displays a class-to-class comparison of experimental covariance matrices by solving the relative eigenvalue problem. Top: swallowing vs. ventilation covariance matrices, bottom: swallowing vs. phonation. PC subject was recorded under full (left) and narrow (right) protocol. The dashed line delineates the eigenvalues for which the maximum ratio of (class to class) variance is greater than one whether for swallowing (left side, filled bars) or ventilation/phonation (right side, empty bars).

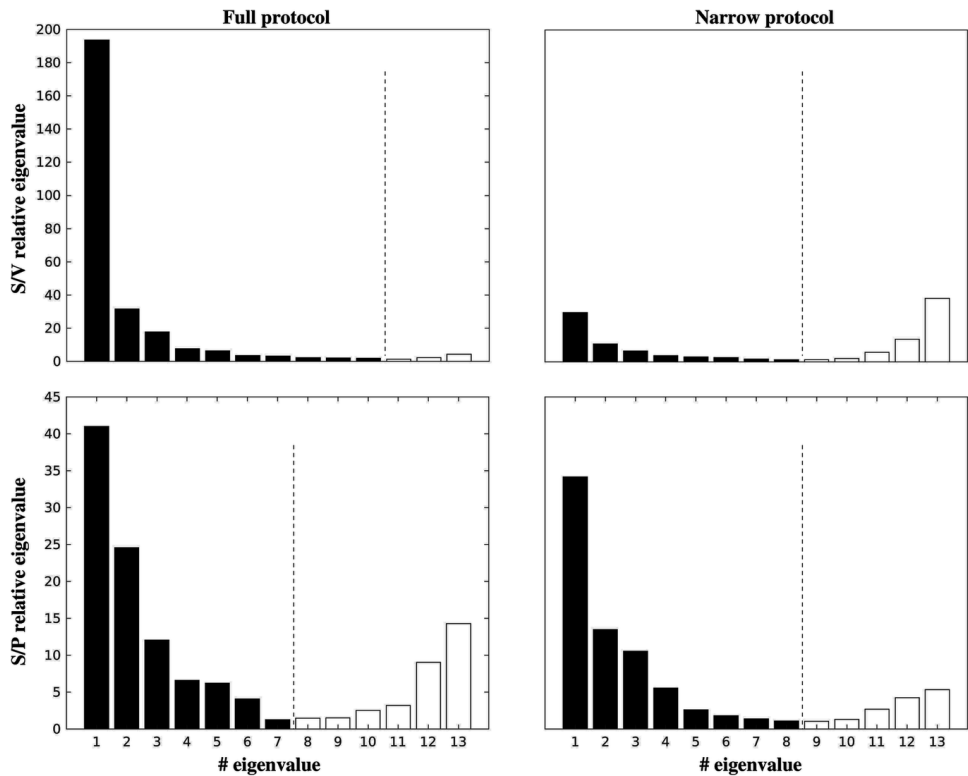


Table 3. Classification performance from narrow protocol (a single type of swallowing and phonation). V, S and P denote respectively Ventilation, Swallowing and Phonation class. Note that the first two subjects are the same than in Table 1, hence their individual performance can be quantitatively compared.

Subjects					
	JF-2	PC-2	CS	ST	Average±SD
Sensitivity (%)					
V	99	99	98	97	98±1
S	95	92	94	79	90±6
P	90	82	84	100	89±7
Precision (%)					
V	100	99	99	98	99±1
S	89	81	90	79	85±5
P	100	97	78	95	93±9

one can see, as expected, a significant improvement both individually for the two subjects involved in the two protocols and in average with a precision of 85%, that is reaching typical performance published in the papers discussed above. Figure 6 (right) compares again the variability of each group in the context of the narrow protocol for subject PC. The relative variance between swallowing and ventilation appears now more balanced, as witnessed by the value of the generalized variance ratio that has severely dropped to 15. Swallowing vs. phonation variance ratio is kept roughly unchanged however, explaining why the precision of the swallowing– though noticeably improved – remains the lowest among the three groups.

CONCLUSION

It has been demonstrated in this paper the potential of Respiratory Inductance Plethysmography for an automated detection of swallowing events. Within a protocol spanning a rich variety of swallows and phonations, the authors achieve a sensitivity of 86% and a precision of 71%. Assuming a prior removal of the vocalization sequences, the results increases up to 95% and 77% respectively, these performance being comparable to that of previously published techniques. A qualitative method has also been provided, which allows to investigate the variability of each class (and in particular that of the swallowing group) in different protocol environment.

This study was carried out in seated subjects with as less movement as possible. Indeed, movement artefacts make RIP signal more difficult to analyze. However, there are many daily life situations where subjects are quiet (meal time or sleep) and where the proposed method could be easily adapted.

The preliminary results in the healthy population investigated in the present study stands as a first promising step towards a non-invasive and convenient technology for a remote monitoring of swallowing. In case of dysphagic populations, greater variations of the swallowing profiles on RIP signal can be anticipated and further developments will be necessary to enable the use of the proposed approach.

Afterwards RIP could be of interest for swallowing investigation in frailty subjects (elderly (Moreau-Gaudry et al., 2005), infants (Geddes et al., 2010)), in a complementary way to the gold standard videofluoroscopy and other clinical evaluations.

ACKNOWLEDGMENT

Research supported by The French National Research Agency. This work is part of the e-SwallHome project (Swallowing and Respiration: Modelling and e-Health at Home, ANR 2014-2018). The authors thank Angelique Brouta for technical experimental assistance.

REFERENCES

- Amft, O., & Troster, G. (2006). Methods for detection and classification of normal swallowing from muscle activation and sound. *Pervasive Health Conference and Workshops*, 1–10.
- Bachy, J.-P., Eberhard, A., Baconnier, P., & Benchetrit, G. (1986). A Program for Cycle-by-cycle Shape Analysis of Biological Rhythms. Application to Respiratory Rhythm. *Computer Methods and Programs in Biomedicine*, 23(3), 297–307. doi:10.1016/0169-2607(86)90064-7 PMID:3816177
- Binazzi, B., Lanini, B., Bianchi, R., Romagnoli, I., Nerini, M., Gigliotti, F., & Scano, G. et al. (2006). Breathing pattern and kinematics in normal subjects during speech, singing and loud whispering. *Acta Physiologica (Oxford, England)*, 186(3), 233–246. doi:10.1111/j.1748-1716.2006.01529.x PMID:16497202
- Bookstein, F. L., & Mitteroecker, P. (2014). Comparing covariance matrices by relative eigenanalysis, with applications to organismal biology. *Evolutionary Biology*, 41(2), 336–350. doi:10.1007/s11692-013-9260-5
- Calabrese, P., Besleaga, T., Eberhard, A., Vovc, V., & Baconnier, P. (2007). Respiratory Inductance Plethysmography is suitable for voluntary hyperventilation test. In *Proceedings of the 2007 Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBC* (pp. 1055-1057). IEEE.
- Camargo, F. P., Ono, J., Park, M., Caruso, P., & Carvalho, C. R. (2010). An evaluation of respiration and swallowing interaction after orotracheal intubation. *Clinics*, 65(9), 919–922. doi:10.1590/S1807-59322010000900015 PMID:21049220
- Casas, M. J., Kenny, D. J., & McPherson, K. A. (1994). 12). Swallowing/ventilation interactions during oral swallow in normal children and children with cerebral palsy. *Dysphagia*, 9(1), 40–46. doi:10.1007/BF00262758 PMID:8131424
- Dong, B., & Biswas, S. (2016). Analyzing Breathing Signals and Swallow Sequence Locality for Solid Food Intake Monitoring. *Journal of Medical and Biological Engineering*, 36(6), 765–775. doi:10.1007/s40846-016-0181-5 PMID:28111533
- Eberhard, A., Calabrese, P., Baconnier, P., & Benchetrit, G. (2001). Comparison between the respiratory inductance plethysmography signal derivative and the airflow signal. *Frontiers in Modeling and Control of Breathing*, 489-494.
- Farooq, M., Fontana, J. M., & Sazonov, E. (2014). A novel approach for food intake detection using electroglottography. *Physiological Measurement*, 35(5), 739–751. doi:10.1088/0967-3334/35/5/739 PMID:24671094
- Fattori, B., Giusti, P., Mancini, V., Grosso, M., Barillari, M. R., Bastiani, L., & Nacci, A. et al. (2016). 10). Comparison between videofluoroscopy, fiberoptic endoscopy and scintigraphy for diagnosis of oro-pharyngeal dysphagia. *Acta Otorhinolaryngologica Italica*, 36, 395–402. PMID:27958600
- Flury, B. N. (1985). Analysis of linear combinations with extreme ratios of variance. *Journal of the American Statistical Association*, 80(392), 915–922. doi:10.1080/01621459.1985.10478203
- Fontecave-Jallon, J., & Baconnier, P. (2016). A simple mathematical model of spontaneous swallow effects on breathing based on new experimental data. In *Proceedings of the 2016 Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBC* (pp. 4260-4263). IEEE.
- Fukunaga, K. (2013). *Introduction to statistical pattern recognition*. Academic press.
- Geddes, D. T., Chadwick, L. M., Kent, J. C., Garbin, C. P., & Hartmann, P. E. (2010). 9). Ultrasound Imaging of Infant Swallowing During Breast-Feeding. *Dysphagia*, 25(3), 183–191. doi:10.1007/s00455-009-9241-0 PMID:19626366
- Golabbakhsh, M., Rajaei, A., Derakhshan, M., Sadri, S., Taheri, M., & Adibi, P. (2014). Automated acoustic analysis in detection of spontaneous swallows in Parkinson's disease. *Dysphagia*, 29(5), 572–577. doi:10.1007/s00455-014-9547-4 PMID:24958599
- Hardemark Cedborg, A., Boden, K., Witt Hedstrom, H., Kuylensstierna, R., Ekberg, O., Eriksson, L., & Sundman, E. (2010). Breathing and swallowing in normal man – effects of changes in body position, bolus types, and respiratory drive. *Neurogastroenterology and Motility*, 10, 1365–2982. PMID:20618836

- Jain, A. K., Duin, R. P., & Mao, J. (2000). Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 4–37. doi:10.1109/34.824819
- Kenny, D. J., Casas, M. J., & McPherson, K. A. (1989). 6). Correlation of ultrasound imaging of oral swallow with ventilatory alterations in cerebral palsied and normal children: Preliminary observations. *Dysphagia*, 4(2), 112–117. doi:10.1007/BF02407155 PMID:2701093
- Klecka, W. R. (1980). *Discriminant analysis* (Vol. 19). Sage. doi:10.4135/9781412983938
- Laouani, A., Rouatbi, S., Saguem, S., & Calabrese, P. (2016). Thorax and abdomen motion analysis in patients with obstructive diseases. *Journal of Pulmonary & Respiratory Medicine*, 6(01), 2. doi:10.4172/2161-105X.1000313
- Leder, S. B. (1996). Gag reflex and dysphagia. *Head & Neck*, 18(2), 138–141. doi:10.1002/(SICI)1097-0347(199603/04)18:2<138::AID-HED5>3.0.CO;2-2 PMID:8647679
- Leslie, P., Drinnan, M. J., Finn, P., Ford, G. A., & Wilson, J. A. (2004). 11). Reliability and Validity of Cervical Auscultation: A Controlled Comparison Using Video fluoroscopy. *Dysphagia*, 19, 231–240. PMID:15667057
- Martin-Harris, B., Brodsky, M. B., Michel, Y., Ford, C. L., Walters, B., & Heffner, J. (2005). Breathing and swallowing dynamics across the adult lifespan. *Archives of Otolaryngology--Head & Neck Surgery*, 131(9), 762–770. doi:10.1001/archotol.131.9.762 PMID:16172351
- Mathers-Schmidt, B., & Kurlinski, M. (2003). Dysphagia Evaluation Practices: Inconsistencies in Clinical Assessment and Instrumental Examination Decision-Making. *Dysphagia*, 18(2), 114–125. doi:10.1007/s00455-002-0094-z PMID:12825905
- Matsuo, K., & Palmer, J. B. (2009). Coordination of mastication, swallowing and breathing. *Japanese Dental Science Review*, 45(1), 31–40. doi:10.1016/j.jdsr.2009.03.004 PMID:20161022
- McFarland, D. H., Martin-Harris, B., Fortin, K., Hill, E., & Armeson, K. (2016). Respiratory-swallowing coordination in normal subjects: Lung volume at swallowing initiation. *Respiratory Physiology & Neurobiology*, 234, 89–96. doi:10.1016/j.resp.2016.09.004 PMID:27612587
- Mohammadi, H., Steele, C., & Chau, T. (2016). Post-Segmentation Swallowing Accelerometry Signal Trimming and False Positive Reduction. *IEEE Signal Processing Letters*, 23(9), 1221–1225. doi:10.1109/LSP.2016.2566582
- Moreau-Gaudry, A., Sabil, A., Baconnier, P., Benchetrit, G., & Franco, A. (2005). Use of computer and respiratory inductance plethysmography for the automated detection of swallowing in the elderly. *Studies in Health Technology and Informatics*, 116, 95. PMID:16160242
- Preiksaitis, H. G., & Mills, C. A. (1996). Coordination of breathing and swallowing: Effects of bolus consistency and presentation in normal adults. *Journal of Applied Physiology*, 81(4), 1707–1714. doi:10.1152/jappl.1996.81.4.1707 PMID:8904590
- Rochet-Capellan, A., & Fuchs, S. (2013). The interplay of linguistic structure and breathing in German spontaneous speech. In *Proceedings of the 4th Annual Conference of the International Speech Communication Association (Interspeech 2013)* (pp. 1228-1232). Academic Press.
- Schindler, J., & Kelly, J. (2002). Swallowing disorders in the Elderly. *Laryngoscope*, 112(4), 589–602. doi:10.1097/00005537-200204000-00001 PMID:12150508
- Sejdic, E., Falk, T. H., Steele, C. M., & Chau, T. (2010). Vocalization removal for improved automatic segmentation of dual-axis swallowing accelerometry signals. *Medical Engineering & Physics*, 32(6), 668–672. doi:10.1016/j.medengphy.2010.04.008 PMID:20483652
- Sejdic, E., Steele, C. M., & Chau, T. (2009). Segmentation of dual-axis swallowing accelerometry signals in healthy usbjects with analysis of anthropometric effects on duration of swallowing activities. *IEEE Transactions on Biomedical Engineering*, 56(4), 1090–1097. doi:10.1109/TBME.2008.2010504 PMID:19171514
- Smith, J., Wolkove, N., Colacone, A., & Kreisman, H. (1989). Coordination of Eating, Drinking and Breathing in Adults. *Chest*, 96(3), 578–582. doi:10.1378/chest.96.3.578 PMID:2766816
- Steele, C. M., Sejdic, E., & Chau, T. (2013). Noninvasive detection of thin-liquid aspiration using dual-axis swallowing accelerometry. *Dysphagia*, 28(1), 105–112. doi:10.1007/s00455-012-9418-9 PMID:22842793

- Tarrant, S. C., Ellis, R. E., Flack, F. C., & Selley, W. G. (1997). 1). Comparative Review of Techniques for Recording Respiratory Events at Rest and during Deglutition. *Dysphagia*, 12(1), 24–38. doi:10.1007/PL00009515 PMID:8997830
- Terzi, N., Orlikowski, D., Aegerter, P., Lejaille, M., Ruquet, M., Zalcman, G., & Lofaso, F. et al. (2007). 2). *Breathing–Swallowing Interaction in Neuromuscular Patients.*, 175, 269–276.
- Tibbling, L., & Gustafsson, B. (1991). Dysphagia and its consequences in the Elderly. *Dysphagia*, 6(4), 200–202. doi:10.1007/BF02493526 PMID:1778095
- Vaiman, M., & Eviatar, E. (2009). Surface electromyography as a screening method for evaluation of dysphagia and odynophagia. *Head & Face Medicine*, 5(1), 9. doi:10.1186/1746-160X-5-9 PMID:19232090
- Walker, W. P., & Bhatia, D. (2011). Towards automated ingestion detection: Swallow sounds. In *Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBC* (pp. 7075-7078). IEEE.
- Wang, Z., & Willett, P. (2003). 3). Two algorithms to segment white Gaussian data with piecewise constant variances. *Signal Processing. IEEE Transactions on*, 51(2), 373–385. doi:10.1109/TSP.2002.806979
- Welling, M. (2005). Fisher linear discriminant analysis. Department of Computer Science, University of Toronto.
- Wheeler Hegland, K., Huber, J. E., Pitts, T., Davenport, P. W., & Sapienza, C. M. (2011). Lung Volume Measured During Sequential Swallowing in Healthy Young Adults. *Journal of Speech, Language, and Hearing Research: JSLHR*, 54(3), 777–786. doi:10.1044/1092-4388(2010/09-0237) PMID:20966381

Eric Tatulli received in 1999 an engineering degree in Physics from Ecole Nationale Supérieure de Grenoble (France) and in 2001 the M.Sc. in Signal, Images, Speech and Telecommunication of the Institut National Polytechnique de Grenoble (INPG). He worked towards his Ph.D. in Astronomy and Astrophysics at the Université Joseph Fourier of Grenoble in 2004. After ten years working as an astronomer in several observatories (IPAG-Grenoble/France, OAA-Florence/Italy, IUCAA-Pune/India), he graduated in the M.Sc. of Arts, Science and Technology of INPG in 2014. He has then joined the TIMC (Grenoble, France) since March 2017. His research interests are signal and image processing, machine learning techniques, applied mathematics and optics.

Julie Fontecave-Jallon received her engineer diploma in Telecommunications in 2002, her Master of Sciences and PhD in Signal Image Speech and Telecommunications in 2003 and 2006 from Grenoble INP institute of Engineering (France). She is an Associate Professor at University Grenoble Alpes (France) since 2011. She teaches at Polytech Grenoble in Health Information Technology Engineering formation. Her research realized at TIMC-IMAG laboratory, in PRETA team, focuses on the articulation of biomedical signals acquisition, mathematical modeling of cardio-respiratory system and physiological signal processing mainly based on time-scale approaches.

Pascale Calabrese received her PhD (Biological and medical engineering) from the Université Joseph Fourier (UJF, Grenoble, France) in 1998. She joined Université Joseph Fourier (UJF, Grenoble, France) in 2002 as an assistant professor in the PRETA Team (Applied, Theoretical and Experimental cardio-Respiratory Physiology) of the TIMC-IMAG lab (Techniques for biomedical engineering and complexity management - UJF/CNRS common unit). She teaches physiology, measurement and modeling in physiology. Her research interests concern respiratory physiology and cardio-respiratory interactions.