

Midsagittal Jaw Movements Analysis for the Scoring of Sleep Apneas and Hypopneas

Frédéric Senny, Jacques Destiné and Robert Poirrier

Abstract

Given the importance of the detection and classification of sleep apneas and hypopneas (SAHs) in the diagnosis and the characterization of the SAH syndrome, there is a need for a reliable non-invasive technique measuring respiratory effort. This paper proposes a new method for the scoring of SAHs based on the recording of the midsagittal jaw motion (mouth opening) and on a dedicated automatic analysis of this signal. Continuous wavelet transform is used to quantize respiratory effort from the jaw motion, to detect salient mandibular movements related to SAHs and to delineate events which are likely to contain the respiratory events. The classification of the delimited events is performed using multi-layer perceptrons which were trained and tested on sleep data from 34 recordings. Compared with SAHs scored manually by an expert, the sensitivity and specificity of the detection were 86.1% and 87.4% respectively. Moreover, the overall classification agreement in the recognition of obstructive, central and mixed respiratory events between the manual and automatic scorings was 73.1%. The midsagittal jaw motion signal is hence a reliable marker of respiratory effort and allows an accurate detection and classification of SAHs.

Index Terms

classification, esophageal pressure, jaw movement, respiratory effort, sleep apnea/hypopnea, wavelet.

Manuscript received November 29, 2006; revised February 19, 2007.

F. Senny and J. Destiné are with the Montefiore Department of Electrical Engineering and Computer Science, University of Liège, Sart-Tilman, Building B28, B-4000, Liège, Belgium. Contact: {F.Senny; jdestine}@ulg.ac.be.

R. Poirrier is with the Department of Neurology, Faculty of Medicine, University of Liège, Sart-Tilman, Building B35, B-4000, Liège, Belgium. Contact: Robert.Poirrier@chu.ulg.ac.be.

Copyright (c) 2006 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org

I. INTRODUCTION

More than 4% of adult males and 2% of adult females of the Western countries suffer from obstructive sleep apnea/hypopnea syndrome (OSAHS) [1]. Sleep apneas and hypopneas (SAHs) belong to a wider family known as sleep-disordered breathing (SDB) [2]. They have essentially two origins: partial or total occlusion of the upper airway causes obstructive apnea and hypopnea whereas lack of central moto-neural respiration drive causes central apnea. The OSAHS has been associated with cardiovascular disorders and severe damage to central nervous system [3]–[5].

A sleep apnea occurs when breathing stops while sleeping. The clinical definition of an apnea is a decrease in magnitude of the respiration movements lower than 20% of the reference value for at least 10 s. An apnea is either obstructive (OA), central (CA) or mixed (MA) depending on whether respiratory effort is completely present, absent or partially present during the apnea. During a hypopnea (H), the respiration movements decrease below 70% of the reference value and are associated with a SpO₂ fall of at least 4% or an arousal [2], [6]. The severity of the OSAHS is measured by the apnea-hypopnea index (AHI), the number of apneas and hypopneas per sleep hour.

SAHs induce changes in a number of physiological parameters, which can be used for the diagnosis of the syndrome. The obstruction leads to variations of nasal airflow and to an increase in the esophageal pressure (Pes) which is manifested as respiratory effort. Arterial blood oxygen saturation (SpO₂) falls because no gas is exchanged, heart rate and blood pressure increase leading to possible cardiovascular damage. Lastly, arousals, whose intensity and latency depend on the type of sleep disorder, appear at the end of more than 70% of apneas and fragment the sleep structure [7]–[10].

The signals commonly used to detect and classify apneas and hypopneas, both in full polysomnography and ambulatory screening methods, are SpO₂, plethypulse rate (both recorded by a pulse oxymeter), nasal airflow and thoraco-abdominal movements [11]. The Pes gives an accurate determination of intrathoracic pressure which becomes subatmospheric at inspiration and decreases further when the upper airway is partially or totally occluded [12]. Consequently, Pes is considered as the *gold standard* measure of respiratory effort to discriminate SDB. However, because the probe must be placed in the esophagus, the technique is relatively invasive and may alter sleep. Consequently, non-invasive Pes surrogates are preferred, such as nasal flow recorded by cannula (NAF), thoraco-abdominal movements (TAM), pulse transit time (PTT), or forced oscillation technique (FOT).

The nasal flow is obviously dedicated to SAHs detection. Automatic algorithms based on NAF provides good detection rate (see [13] for instance). Though the flattening of the inspiratory curve is a sensitive

marker of respiratory effort [14]–[16], the classification of apnea using NAF is impossible if complete cessation of airflow occurs. TAM, recorded using strain gauges belts, are usually used in addition to NAF to classify SAH by the visual inspection of these movements [11], [17] or by an automated method [18]. However, the signal quality is poor when belts are not held in position, as it is the case in obese patients. The PTT is defined as the time between the R wave on the electrocardiogram (ECG) to the pulse wave arrival at the finger [19]. The classification of SAH is feasible with PTT by visual inspection [20]. FOT is a measure of the upper airway impedance. This method consists in the application of a low amplitude pressure at a frequency of 5 to 20 Hz by a mask pressure. From the flow and the pressure both recorded, the FOT signal is extracted and associated with the flow signal to score SAHs using a breath-by-breath. Despite a cumbersome apparatus, the obstructive and central nature of a respiratory event is accurately found [21], [22]. Other automated methods for SAHs detection (without classification of respiratory events) have been recently developed to diagnose OSAHS using ambulatory measurement and the analysis of the pulse arterial tonometry (PAT) [23], [24] or ECG [25], [26]. Generally, the accuracy of automatic methods varies from 80% to 90%. Misclassification errors come from the intrinsic limitations in signal processing, in detection and also in classification methods, but also from the lack of agreement between scorers due to their own interpretation of rules, especially those applying to hypopneas.

The topic of this paper is the automatic detection and classification of SAHs through the *midsagittal jaw motion (MJM)*, a novel Pes surrogate. The rationale is based on the principle that when the upper airway collapses and when, in the meantime, the pressure in the lungs becomes more and more subatmospheric, the jaw is passively lowered by thoracic attraction called thoracic tugs [27]. The consequent opening of the mouth reflects, during sleep, a stronger effort to breath. In [28]–[30], oscillating jaw movements (breathing frequency band [0.15-0.33] Hz) were observed in correlation with the Pes during obstructive sleep events. This correlation confirmed that the jaw movements could reflect the effort to breath (cases (a), (b), (c), (d) and (f) in Fig. 1). In addition to the breathing jaw movements, vertical mandibular position is more open during sleep in patients with OSAHS than in healthy patients. Besides, significant mandibular opening during an obstructive apnea in NREM sleep was observed, while in REM sleep, the muscle atonia attenuates these movements [31]. The masseter (main muscle of the mandible) plays a major role in stabilizing the jaw during a SAH and in the closing of the mandible which terminates an obstructive sleep event [28], [30], [31]. In this paper, this closure of the mandible is referred to *salient jaw movement*. It appears during the breathing recovery and can be either a smooth or a sharp closing jaw movement. Fig. 1 depicts 8 patterns of the MJM signal in cases of (a) obstructive apnea, (b) central apnea, (c) mixed apnea, (d) hypopnea, (e) healthy sleep, (f) snoring, (g) awake and (h) isolated artefacts

(swallowing,...). To sum up the previous findings [27]–[31], the three main features of the midsagittal jaw motion signal to suspect and score sleep apneas and hypopneas are 1) a significant opening of the mouth, 2) a salient jaw movement occurring after the sleep event (pointed out by an arrow in Fig. 1) and 3) the oscillating jaw movements related to respiratory effort which are expected to be present during an obstructive apnea. This last feature is the core asset of this signal, which means that the jaw motion sensor is an *effort sensor* and not a respiration sensor (like cannula). Recent studies presented a rule-based approach to detect SAHs [32] that predicted OSAHS accurately [33], in which the jaw motion is recorded at home with an easy-to-use ambulatory device [34]: the sensitivity and specificity of SAHs detection were respectively 71.8% and 76.8% when considering the whole recording, awakenings included, while OSAHS was diagnosed with a sensitivity and a specificity greater than 90%. The purpose of this paper is to propose an off-line automatic analysis of the jaw motion to improve SAHs detection and to classify them. The proposed method combines wavelet analysis to detect the salient mandibular movements (section II-B1), to delineate events which are likely to contain the SAHs, to quantize respiratory effort (section II-B2), and multi-layer perceptrons to classify these delimited events using local and contextual features (sections II-B3 and II-B4). The perceptrons were trained and tested using only sleep data (data during wake were discarded to simplify the task). The ability of jaw motion analysis to differentiate sleep from wake is out of the scope of this paper and will be assessed in a further study.

II. MATERIALS AND METHODS

A. Subjects and recordings

Data from 34 recordings between 10pm and 3am were studied. For all of them, a standard polysomnography (PSG) was recorded in a hospital setting. PSG is a multichannel recording composed of at least the electroencephalogram, electrooculogram, electromyogram, electrocardiogram, nasal flow, arterial blood oxygen saturation, thoraco-abdominal movements, snoring sounds and body position. Sleep stages, and sleep apneas/hypopneas were visually screened according to standard rules [2], [35].

The mandibular opening was measured by the JAWSENS sensor supplied by NOMICS [36]. The sensor (Fig. 2) is made up of two coupled resonant circuits, placed perpendicularly to the medium line of the face on the forehead and below the lower lip. The output voltage is a cubic function of the distance [37]. The voltage is sampled at 10 Hz, digitally linearized and stored on the computer synchronously with the other PSG channels in a EDF file [38].

The mandibular movements signal can be expressed either in absolute value (millimeter) or in normalized value (percentage of mouth opening). Normalizing the signal is a two-steps process. Firstly, patient

is asked to widely open then fully close his/her mouth just at the very beginning of the night recording. Secondly, the user must choose the two values v_o "wide open mouth" and v_c "fully closed mouth" through a graphical user interface displaying the first minutes of the signal. Data are then normalized

$$d_n(\%) = 100 \frac{d - v_c}{v_o - v_c}. \quad (1)$$

100% corresponds to a wide open mouth and 0% to a fully closed mouth.

In this paper, the mandibular movements signal is displayed in percentage of mouth opening for convenience, but analyzed in millimeter to avoid user's measure errors occurring in the normalization step.

B. Signal processing

The scoring of sleep apneas and hypopneas by the analysis of jaw motion relied on the detection of salient jaw movements, which occur after the sleep events, and on the quantization of both the opening of the mandible and the respiratory effort (Fig. 1). Each movement of the mandible was localized and characterized by the continuous wavelet transform. High amplitude or discontinuous movements were assumed to contain the salient jaw movements which allowed the delineation of time courses likely to contain the sleep events (such assumption was confirmed by the results, see section III). The relative temporal positions of movements were used to track the effort patterns. The multi-layer perceptrons, trained with local and contextual features, rejected non-apneic events at first, and then classified the detected apneic events into obstructive, central or mixed apneas or hypopneas. Wavelet transform and features computation were performed under MATLAB, while the selection of features, their normalization and the training of perceptrons were performed under LNKnet using a cross-entropy cost function, the back-propagation training algorithm and the standard sigmoidal activation function [39] (a C code of a trained perceptrons can be generated by LNKnet and included in a MATLAB routine). The signal processing block diagram is sketched at Fig. 3 and each block will be detailed further.

1) *Block 1 - Continuous Wavelet Transform (CWT)*: Wavelet transform [40] is a widely used tool for the analysis of biosignals [41]. The wavelet coefficient W of the continuous wavelet transform is defined as

$$W(u, s) = f(t) * \psi_s(t) = \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-u}{s}\right) dt, \quad (2)$$

with

$$\psi(t) = -\frac{1}{\sqrt{2\pi}} t \exp\left(-\frac{t^2}{2}\right), \quad (3)$$

the first derivative of a Gaussian window in our application, t the time and s the scale. Dilating ψ_s (big s) tracks low frequency changes whereas compressing ψ_s (small s) tracks high frequency changes. The CWT of the discretized signal $f(n)$ was computed in the Fourier domain using the fast Fourier transform (FFT) and its inverse (IFFT) on a N ($= 1024$) points sliding window (overlap= $\frac{4}{5}$ and border coefficients were ignored), scales s covering 5 octaves from $s_1(= 1)$ to $s_2(= 2^5)$ with 20 points per octave, and the following relations

$$W(n, s) = IFFT[\hat{f}(w) \hat{\psi}_s(w)](n), \quad (4)$$

$$\hat{f}(w) = FFT[f(n)], \quad (5)$$

$$\hat{\psi}_s(w) = jws e^{-\frac{(ws)^2}{2}}, \quad (6)$$

$$w = \frac{2\pi k}{N}, k = 0, \dots, N - 1. \quad (7)$$

Two properties of the first derivative of a Gaussian wavelet are to remove the average component and to associate a curve in function of the scale with each edge in the signal (*i.e.* the closing or opening jaw movements): the curve, noted $|W(s)|$, is a series (or a chain) of wavelet transform modulus maxima (local maxima of the wavelet transform close in the time but at different scale) [40]. Finding such maxima chains allowed the detection of jaw movements, while the area under the curves was used to characterize the jaw movements, instead of a parametric study of $|W(s)|$ as in [40], [42], [43]: if the area under four curves related to four edges of a synthetic signal was compared (edges noted A, B, C and D in Fig. 4), the area associated with sharp or high amplitude edges (edges A and B) was obviously bigger than the area associated with smooth or low amplitude edges (edges C and D). Consequently, the measure of the area contained information about amplitude and regularity such that the bigger the area was, the bigger the amplitude was or the more irregular (or discontinuous) the edge was. Hence, the measure of the area under the maxima chains was sufficient to detect salient jaw movements (SJMs) and relevant movements related to respiratory effort (RE), because, in our application, the two important features of an edge were its amplitude and its regularity (a discontinuous structure in a signal appears as an irregular edge).

Wavelet transform modulus maxima were chained and two sets of chains (S1 and S2) were computed

- 1) S1 contains chains Ch_i such that $I_i \geq Th_{RE}$,
- 2) S2 contains chains Ch_j such that $I_j \geq Th_{SJM}$,

where

$$I_k = \int_{\log_2(s_1)}^{\log_2(s_2)} |W| d(\log_2(s)) \approx \sum_{n=1}^{100} |W| h, \quad (8)$$

is the area under the curve formed by the maxima chain Ch_k (graph (c) in Fig. 4) and h is the constant step in LOG2 of scales ($h = \frac{1}{20} = 0.05$). Thresholds Th_{RE} and Th_{SJM} were two fixed percentages of all the integrals of the chains computed on the sliding window. Thus, the chains of sets S1 and S2 on a window were independent whether or not the signal had been normalized, and depended only on the content of that window. Sleep events longer than the length of the sliding window ($\frac{1024}{10\text{Hz}} = 102.4$ sec) could not be detected, but the sleep events of interest usually last less than 60 seconds. Th_{RE} was arbitrarily fixed at 0.005 so that almost all movements were tracked. Several values of Th_{SJM} were tested (see the results section) and the best value found maximized the sensitivity with a reasonable specificity.

Each chain Ch_k had thereby three features:

- I_k , the integral under the chain of maxima characterizing the edge;
- t_{Ch_k} , the instant the edge occurred (time of the smallest scale of the chain);
- S1 or S2, the set it belonged to.

2) *Block 2 - Respiratory effort and delimitation of time courses*: In order to assess quantitatively respiratory effort cycles, the temporal distribution of chains from set S1 had to meet the next criteria derived from our normal breathing observations:

- i. a cycle of effort was composed of inspiration and expiration,
- ii. the frequency varied between 0.15 and 0.33 Hz,
- iii. the inspiration time was usually less than the expiration time.

Time differences of successive chains of S1 were computed, rules were applied to these time differences. A rule is a look-up table whose output O varies from 0 to 1. Finally, outputs of all rules were combined using a combination function F (average, geometric distance or Dempster-Shafer's rule for instance) to quantify respiratory effort (RE) using

$$RE(t_k) = I_k F(O_1, O_2, \dots, O_n), \quad (9)$$

where I_k was weighted by the combination function of the n rules $F(O_1, O_2, \dots, O_n) \in [0 - 1]$ and $t_k \in \left[\frac{t_{Ch_{k-1}} + t_{Ch_k}}{2}, \frac{t_{Ch_k} + t_{Ch_{k+1}}}{2} \right]$.

Salient jaw movements were tracked with the set of chains S2 which contained high amplitude or discontinuous jaw movements. An apneic event was assumed to last more than 8 seconds (instead of 10 seconds, by definition), because its onset and offset could not be found as precisely with the jaw motion as with nasal flow. Thus, successive chains from set S2 with time difference less than 8 seconds were grouped, and each group was (or at least contained) the desired salient jaw movement. A time course, a potential respiratory event, was defined between two successive salient jaw movements.

3) *Block 3 - Computation of features*: The features were computed on each time course (TC), the related salient jaw movement (SJM) and the neighborhood. The Table I lists these features and gives a short description. These features were normalized (zero mean and unit variance) and selected using a nearest-neighbor leave-one-out cross-validation technique under LNKnet to reduce the complexity of the multi-layer perceptrons (a decision tree trained on the same data also selected most of these features). Most features are self-explanatory, except the mean variations ($\mu_1, \Delta\mu_{21}, \Delta\mu_{31}, \Delta\mu_{41}, \Delta\mu_{51}$), the durations without effort ($d_{1/2}(no RE_{TC})$ and $d_{2/2}(no RE_{TC})$) and the index of periodicity (IOP). The five features from local means of the signal were computed as follows

- the time course was decomposed into 5 subregions of equal range;
- the mean of the signal on each subregion ($\mu_i, i = 1, \dots, 5$) was computed ;
- μ_1 was a feature as it is, the other four were the differences $\Delta\mu_{j1} = \mu_j - \mu_1, j = 2, \dots, 5$

The two durations $d_1(no RE_{TC})$ and $d_2(no RE_{TC})$ were defined as the number of samples in the first (and the second) half of the time course for which no respiratory effort (RE) was present, that is the effort was below the threshold Th_{noRE} . These two durations were computed for different values of Th_{noRE} , the features selection methods found $\frac{3h}{10}$ (h is the constant step in LOG2 of scales) as the most relevant value.

The index of periodicity IOP was a temporal and a local measure of periodicity in successive time courses. It was computed from the durations of the salient jaw movement (dSJM) and of the time course (dTC) as follows

- normalize i^{th} sleep event dSJM and dTC using the durations from the whole night recording

$$d_n X(i) = \frac{dX(i) - \mu(dX)}{\sigma(dX)}, X = \{SJM, TC\};$$

- cumulate upon five successive sleep events and compute the local Euclidian distance

$$led(i) = \sum_X \sum_j^{SJM, TC - 2, -1, 1, 2} (d_n X(i+j) - d_n X(i))^2;$$

- define a threshold Th_{led} as the 75th percentile of led , and apply a rule such that

$$IOP(i) = 1 \quad , \text{ if } led(i) \geq Th_{led},$$

$$IOP(i) = 0 \quad , \text{ if } led(i) < Th_{led},$$

to finally compute the feature $IOP(i)$.

TABLE I: List of features computed on each time course (TC) and on the related salient jaw movement (SJM). The features marked with a ^d or ^c were dedicated respectively to SAHs detection or classification.

FEATURES	DESCRIPTION
$dSJM$ ^d	duration of SJM (in samples)
$\mu(I_{SJM})$ ^{dc}	mean of $I_k \in \text{SJM}$
$\max(I_{SJM})$ ^d	max of $I_k \in \text{SJM}$
$pRESJM$ ^d	percentage of RE in SJM
$\mu(I_{TC})$ ^d	mean of $I_k \in \text{TC}$
$\sigma(I_{TC})$ ^d	standard deviation of $I_k \in \text{TC}$
dTC ^{dc}	duration of TC (in samples)
$\frac{\mu(I_{TC})}{\max(I_{SJM})}$ ^d	increase between TC and SJM
μ_1 ^d	mean of signal of 1 st subregion
$\Delta\mu_{21}$ ^{dc}	mean variation between 2 nd and 1 st subregions = $\mu_2 - \mu_1$
$\Delta\mu_{31}$ ^{dc}	mean variation between 3 rd and 1 st subregions
$\Delta\mu_{41}$ ^d	mean variation between 4 th and 1 st subregions
$\Delta\mu_{51}$ ^d	mean variation between 5 th and 1 st subregions
IOP ^d	index of periodicity
$p_{25}(RE_{TC})$ ^c	25th percentile of RE
$d_1(\text{no}RE_{TC})$ ^c	duration for which no effort were found in the first half of the TC
$d_2(\text{no}RE_{TC})$ ^c	duration for which no effort were found in the second half of the TC

4) *Block 4 - Classification of time courses:* The classification was performed in two steps to reduce the complexity of the problem: rejection of non-SAH time courses was performed by two multi-layer perceptrons (MLPs) and the classification of respiratory events by a third one. The first MLP was fed with 14 features computed on the time course, the related salient event and the neighborhood (features marked with a ^d in Table I). The second MLP modelled the neighborhood of the time course to be classified by adding to the 14 features the two outputs of the first MLP from n successive time courses where the one to be classified was the c^{th} ($c \in 1, \dots, n$) (Fig. 5). The outputs of the first MLP were the estimated posterior probabilities (EPPs) using the cross-entropy cost function during the training step. Depending

on n and c , the added features (the EPPs) contained contextual information about past, future or both: if $n = 5$ and $c = 3$, a time course was classified by the second MLP when the current, the previous two and the next two time courses were classified by the first MLP. Results are shown for different references in three and five successive time courses: the reference is the center one ($n = 5$ and $c = 3$) and the last one ($n = 3$ and $c = 3$).

Once SAHs were detected, they were classified using the 7 features marked with a c in Table I. Two approaches were studied for this purpose:

- 1) The first approach was the basic one: one classifier (7 inputs, 4 outputs and no hidden layer) to classify a sleep event into one of the four classes: hypopnea, obstructive apnea, central apnea and mixed apnea (Fig. 6 (a)).
- 2) The second approach used three MLPs (7 inputs, 2 outputs and no hidden layer), which had a specific task and were applied in a priority order (Fig. 6 (b)). The task of the first MLP was to find central apneas among all the detected SAHs. If the event was not central, the second MLP had to recognize mixed apneas among totally or partially obstructive events. The last MLP tried to distinguish obstructive apneas from hypopneas. The greatest priority was given to the detection of central apneas, followed by the classification of mixed events and finally the separation between obstructive apneas and hypopneas.

The second approach was motivated by the decomposition of this complex classification problem into smaller and easier binary classification problems. Moreover, because the jaw movements signal is not a respiration signal, the classification of obstructive apneic and hypopneic events was expected to be difficult, thus of less priority than the recognition of central events.

III. RESULTS

Delimitation of time courses depended on the selection of significant edges, which depended on the integral under the maxima chains and the threshold Th_{SJM} . Table II shows that the use of the integral according to the LOG2 of scales with $Th_{SJM} = 2$ gives the best trade-off between good sensitivity and a reduced false alarm rate. 3127 SAHs were manually scored in the 170 hours of recording ($= 5 * 34$). The method detected $\frac{2990}{3127} = 95.6\%$ of the SAHs.

For both local and contextual classification, several MLPs structures had been trained (3200 non-SAHs and 1830 SAHs) and tested (2200 non-SAHs and 1160 SAHs). The sensitivity and specificity

TABLE II: Number of delineated non-SAH and SAH time courses in function of Th_{SJM} and the formula used to compute I_k .

NON-SAH	SAH	% OF SAHS	Th_{SJM}	COMPUTE I_k USING
5400	2990	95.6	2	$\int W d(\log_2(s))$
7111	3034	97.0	2	$\int W ds$
5949	2887	92.3	2.2	$\int W ds$

were computed

$$Sensitivity \quad Se(\%) = 100 \frac{TP}{TP + FN}, \quad (10)$$

$$Specificity \quad Sp(\%) = 100 \frac{TN}{TN + FP}, \quad (11)$$

where TP were the true positives, FN the false negatives, TN the true negatives and FP the false positives. Table III sums up the most interesting experiments. The first MLP had 14 inputs listed in Table I. Three structures are presented in Table III : one without any hidden layer, one with 6 nodes and another one with 11 nodes in a single hidden layer. The first structure is the simplest model which leads to the best sensitivity but the worst specificity. The performance of the two other ones was similar but the complexity of the third model was obviously higher. Thus, the second model was chosen ($Se = 86.0\%$ and $Sp = 84.2\%$).

The number of inputs of the second MLP was equal to $14 + 2 * n$ (the 14 features of the current time course to be classified and the EPPs of neighboring time courses) and two outputs *SAH* or *non-SAH*. Results of two cases are shown in Table III. The added features of the first case contained past, present and future information ($n = 5$ and $c = 3$), while the second case contained only past and present information ($n = 3$ and $c = 3$). The EPPs were computed by the MLP whose structure was 14-6-2. Specificity was increased up to at least 87.4% in both cases, but the increase of sensitivity was higher when more information about the neighborhood was added (90.1% versus 88.5%).

To sum up, when taking the first MLP with structure 14-6-2 and the second MLP with structure 24-2, 90.1% of the 95.6% of detected SAHs were correctly scored. Therefore, the global sensitivity for the detection step was 86.1% with a specificity of 87.4%. In order to assess the capability of the detection method in full night recording, it was applied to the whole sleep data (awakenings were discarded and the total sleep time was given). The apnea-hypopnea indices (AHIs) from the 34 recordings were

TABLE III: Sensitivity and specificity of the recognition of detected SAH computed on the test set. Comparison of several structures of the first two MLPs.

MLPs STRUCTURE	TP	FN	FP	TN	SE	SP
14-2	1016	144	432	1768	87.6	80.4
14-6-2	998	162	347	1853	86.0	84.2
14-11-2	1007	153	345	1855	86.8	84.3
24-2*	1093	115	277	1923	90.1	87.4
20-2**	1027	133	275	1925	88.5	87.5

* $n = 5$ and $c = 3$, EPP computed by MLP 14-6-2.

** $n = 3$ and $c = 3$, EPP computed by MLP 14-6-2.

computed and compared with the AHIs from the polysomnography: no significant difference was found between them (through T-student and Fischer-Snedecor tests, $p < 0.01$, the mean and standard deviation of their difference were 0.62 ± 6.67 and the slope of the linear regression was 0.85 with a coefficient of determination R^2 of 0.89).

The database used for the classification problem was composed of 1715 respiratory events (149 Hs, 876 OAs, 157 CAs and 533 MAs) from the 8 recordings for which a mild or severe sleep apnea syndrome was diagnosed through the polysomnography. Because the database did not contain many hypopneas or central apneas, the training set was limited to 100 patterns per class, but it was large enough to train the small MLPs (7 inputs, 2 or 4 outputs and no hidden layer). The remaining patterns formed the test set (49 Hs, 776 OAs, 57 CAs and 433 MAs). The same training and test sets were used to assess the two approaches (one MLP and three MLPs in cascade, Fig. 6). The comparison of the classes attributed to each sleep event by the expert and by the automatic scoring allowed to construct the two confusion matrices, which are illustrated in Tables IV and V. The diagonal elements of a confusion matrix are well-classified patterns while the others are errors: in Table IV for instance, 238 MAs were correctly classified by both the expert and the automatic method, while 58 events were classified as MA by the expert but as CA by the automatic method (numbers in bold). The global agreement of the first approach was 50.1% and, for the second one, 61.4% which was thereby better but barely enough. Despite the small amount of CAs, 71.9% of the CAs were found. The CAs found as MAs were mainly due to one oscillation of the jaw, considered as effort, included in the event delineated automatically. MAs were classified satisfactorily (69.0% of agreement), while OAs were distributed into H, OA and MA events

(57.7% of agreement). However, if Hs and OAs were grouped as obstructive respiratory events (that is if the italic numbers in the confusion matrix were summed), 73.9% of obstructive events were correctly classified. Moreover, in this case, 73.1% of respiratory events were correctly classified into obstructive, central or mixed events using the classification approach in cascade. Fig. 7 (a) depicts good classification of obstructive, mixed and central apneas. Fig. 7 (b) shows the difficulty to distinguish obstructive apneas from hypopneas.

TABLE IV: Confusion matrix of the classification of SAHs into hypopnea (H), obstructive apnea (OA), central apnea (CA) or mixed apnea (MA) by the first approach. The MLP structure was 7-4. The global agreement was 50.1%, and increased up to 65.3% if obstructive events (H and OA) were grouped.

		<i>automatic scoring</i>			
		H	OA	CA	MA
<i>visual scoring</i>	H	20	13	6	10
	OA	190	315	61	210
	CA	2	0	38	17
	MA	36	56	58	238

TABLE V: Confusion matrix of the classification of SAHs into hypopnea (H), obstructive apnea (OA), central apnea (CA) or mixed apnea (MA) by the second approach. The structure of the three MLPs was 7-2 each. The global agreement was 61.4%, and increased up to 73.1% if obstructive events (H and OA) were grouped.

		<i>automatic scoring</i>			
		H	OA	CA	MA
<i>visual scoring</i>	H	20	20	4	5
	OA	134	448	41	155
	CA	1	3	41	12
	MA	12	91	32	299

IV. DISCUSSION AND CONCLUSIONS

The esophageal pressure (Pes) should be the favorite signal for the screening of sleep-disordered breathing (SDB) in hospital settings or in ambulatory methods, as the Pes signal is the best clinical marker of respiratory effort. However, the sensor must be placed by an expert, it is invasive and alters the patient's sleep. Instead of using Pes to score sleep apneas and hypopneas (SAHs), non-invasive Pes surrogates were studied or are still under study, such as pulse transit time (PTT) [20], forced oscillation technique (FOT) [21], [22], nasal airflow (NAF) [13]–[16] and thoraco-abdominal movements (TAM) [11], [17], [18]. This paper proposed a novel non-invasive easy-to-use Pes surrogate, the midsagittal jaw movements, and a dedicated automatic method to score SAHs without the use of the nasal flow signal. Continuous wavelet transform was used to delineate SAHs, through the detection and the characterization of jaw movements using the measure of the area under the curve formed by wavelet transform modulus maxima chain, and to quantize respiratory effort. Classification of SAHs was performed using MLPs with local and contextual features. The proposed method reached 86.1% of sensitivity with a specificity of 87.4% for SAHs detection when wake data were discarded. False negative patterns mainly came from isolated SAHs which were difficult to detect with a contextual approach favoring the detection of salvo of SAHs which are more relevant, while respiratory effort related arousal (RERA) and sleep events close to hypopnea definition mainly contributed to false positive patterns. The classification of SAHs by jaw motion analysis is quite accurate with an agreement of 73.1%. Such results were achieved 1) by decomposing the four-classes problem into easier binary sub-problems, 2) by ordering their application in descending priority (central apnea recognition is more important than the separation between hypopnea and obstructive apnea, Fig. 6) and 3) by viewing obstructive hypopneic and apneic events as one kind of obstructive respiratory events. This last condition is intelligible because of the inherent capacity of the maxillo-mandibular movements to assess effort and not respiration (like the nasal airflow). Therefore, the behaviour of the jaw movements are expected to be quite the same whether the decrease in amplitude of airflow signal is 75% (defining a hypopnea) or 85% (defining an apnea) of the reference value (Fig. 7 (b)). In addition to this limitation about classification, two other origins of misclassification errors should be pointed out: First, the computed features could not reflect the real content of the sleep event because delineated sleep events with jaw motion did not correspond exactly to those delineated with the nasal flow signal by a human expert. Second, scorer's interpretation of the standard rules is known to play a role in classification results, that is the boundaries between apneas and hypopneas or between obstructive and mixed apneas are somewhat fuzzy. Nonetheless, sleep apneas and hypopneas were correctly detected

and delimited. Besides, the obstructive or central nature of a respiratory event was found with a fairly good accuracy.

Compared with other methods presented in the literature, the proposed method is competitive when only sleep data are considered (data in wake state were excluded). In [18], the nasal flow was used to delimit apneas, while wavelet coefficients, computed on the thoracic movements signal, were the inputs of a feedforward neural network. 90.0% of apneas were detected with a 76.6% of specificity. The classification agreement was 83.8% using a cross-validation technique and dropped to 74.0% on a validation set composed of new data. The FOT signal could be used instead of thoracic movements to classify apneas [22]. The global agreement of the dedicated algorithm was 73.5% on 12 recordings. The results obtained with our method based on jaw movements analysis were better than the previous one, since this latter was only able to detect apneas and hypopneas with a sensitivity of 71.8% and a specificity of 76.8% [32], and were similar to the methods cited above. Moreover, an advantage of our method is that the detection and the classification of respiratory events is feasible with only one signal, while nasal flow, PAT or ECG based methods require another physiological parameter, which measures respiratory effort (like TAM or FOT). This asset is valuable especially in ambulatory diagnosis methods.

The proposed method could be improved by refining the delimitation of suspected SAHs which the features should better characterize. Moreover, time series classification tools, such hidden markov models, should be tested and compared with the proposed one. A multi physiological signals analysis will certainly improve the scoring of SAHs. For example, SpO₂ is a simple and reliable indicator of SAHs in most cases and its use would increase both sensitivity and specificity. The classification of SAHs may be improved by making use of the respiratory events delimited by the nasal flow signal, especially in cases of hypopneas and RERA. Nevertheless, the jaw motion is a reliable signal to score SAHs and may be a very useful physiological signal to screen SDB (SAH, RERA or snoring). Further studies should focus on the relationship between salient jaw movements and arousals, and on the discrimination between sleep and wake. Consequently, the jaw motion, on its own, should be able to diagnose and characterize the sleep apnea syndrome in a easy and convenient manner, well suited for home sleep studies.

In conclusion, the midsagittal jaw motion signal is a reliable marker of respiratory effort and its analysis using wavelet transform and multi-layer perceptrons allows an accurate detection and classification of sleep apneas and hypopneas.

ACKNOWLEDGMENT

The authors would like to thank L. Cambon and P. Maquet from the sleep laboratory of the University Hospital of Liège for their help in the scoring of the recorded data.

Special thanks to the NOMICS staff for the technical aid.

REFERENCES

- [1] T. Young, M. Palta, J. Dempsey, J. Skatrud, S. Weber, and S. Badr, "The occurrence of sleep-disordered breathing among middle-aged adults," *N Engl J Med*, vol. 328, no. 7, p. 1230-1235, 1993.
- [2] "Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. the report of an american academy of sleep medicine task force.," *Sleep*, vol. 22, no. 5, pp. 667-689, 1999.
- [3] J. Hung, E. Whitford, R. W. Parsons, and D. R. Hillman, "Association of sleep apnea with myocardial infarction in men," *Lancet*, vol. 336, pp. 261-264, 1990.
- [4] C. Bassetti and M. Aldrich, "Sleep apnea in acute cerebrovascular diseases: Final report on 128 patients," *Sleep*, vol. 22, pp. 217-223, 1999.
- [5] P. Peppard, T. Young, M. Palta, and J. Skatrud, "Prospective study of the association between sleep-disordered breathing and hypertension," *N Engl J Med*, vol. 342, pp. 1378-1384, 2000.
- [6] A. L. Meoli, K. R. Casey, R. W. Clark, J. A. Coleman, R. W. Fayle, R. J. Troell, and C. Iber, "Hypopnea in sleep-disordered breathing in adults," *Sleep*, vol. 24, pp. 469-470, 2001.
- [7] J. R. Stradling, D. J. Pitson, L. Bennet, C. Barbour, and R. J. O. Davies, "Variation in the arousal pattern after obstructive events in obstructive sleep apnea," *Am J Respir Crit Care Med*, vol. 159, pp. 130-136, 1999.
- [8] J. Krieger, C. Schröder, and C. Erhardt, "Cortical autonomous arousal. technical evaluations and clinical interest.," *Revue Neurologique (Paris)*, vol. 159, 2003.
- [9] K. Rees, D. P. Spence, J. E. Earis, and P. M. A. Calverley, "Arousal responses from apneic events during non-rapid-eye-movement sleep," *Am J Respir Crit Care Med*, vol. 152, pp. 1016-1021, 1995.
- [10] K. Dingli, I. Fietze, T. Assimakopoulos, S. Quispe-Bravo, C. Witt, and N. J. Douglas, "Arousability in sleep apnoea/hypopnoea syndrome patients," *Eur Respir J*, vol. 20, pp. 733-740, 2002.
- [11] M. H. Kryger, T. Roth, and W. C. Dement, *Principles and practice of sleep medicine*, ch. Monitoring respiratory and cardiac function, pp. 1217-1230. Philadelphia:W. B. Saunders Company, 2000.
- [12] S. Berg, J. C. Hybbinette, T. Gislason, and M. Hawke, "Continuous intrathoracic pressure monitoring with a new esophageal microchip catheter in sleep-related upper airway obstructions," *J. Otolaryngol*, vol. 24, no. 3, pp. 160-164, 1995.
- [13] P. Varady, T. Micsik, S. Benedek, and Z. Benyo, "A novel method for the detection of apnea and hypopnea events in respiration signals," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 9, pp. 936-942, 2002.
- [14] R. G. Norman, M. M. Ahmed, J. A. Walsleben, and D. M. Rapoport, "Detection of respiratory events during npsg: nasal cannula/pressure sensor versus thermistor," *Sleep*, vol. 20, no. 12, pp. 1175-1184, 1997.
- [15] I. Ayappa, R. G. Norman, A. C. Krieger, A. Rosen, R. L. O'Malley, and D. M. Rapoport, "Non-invasive detection of respiratory effort-related arousals (reras) by a nasal cannula/pressure transducer system," *Sleep*, vol. 23, no. 6, pp. 763-771, 2000.
- [16] C. Cracowski, J. Pépin, B. Wuyam, and P. Levy, "Characterization of obstructive nonapneic respiratory events in moderate sleep apnea syndrome," *Am J Respir Crit Care Med*, vol. 164, pp. 944-948, 2001.
- [17] R. Farré, J. Montserrat, and D. Navajas, "Noninvasive monitoring of respiratory mechanics during sleep," *Eur Respir J*, vol. 24, pp. 1052-1060, 2004.
- [18] O. Fontenla-Romero, B. Guijarro-Berdinas, A. Alonso-Betanzos, and V. Moret-Bonillo, "A new method for sleep apnea classification using wavelets and feedforward neural networks," *Artificial Intelligence in Medicine*, vol. 34, pp. 65-76, 2005.

- [19] D. J. Pitson, A. Sandell, R. van den Hout, and J. R. Stradling, "Use of pulse transit time as a measure of inspiratory effort in patients with obstructive sleep apnea," *Eur Respir J*, vol. 8, pp. 1669–1674, 1995.
- [20] J. Argod, J. Pépin, R. P. Smith, and P. Levy, "Differentiating obstructive and central sleep respiratory events through pulse transit time," *Am J Respir Crit Care Med*, vol. 158, pp. 1778–1783, 1998.
- [21] S. Reisch, J. Daniuk, H. Steltner, K.-H. Rühle, J. Timmer, and J. Guttman, "Detection of sleep apnea with the forced oscillation technique compared to three standard polysomnographic signals," *Respiration*, vol. 67, pp. 518–525, 2000.
- [22] H. Steltner, R. Staats, J. Timmer, M. Vogel, J. Guttman, and H. Matthys, "Diagnosis of sleep apnea by automatic analysis of nasal pressure and forced oscillation impedance," *Am J Respir Crit Care Med*, vol. 165, pp. 940–944, 2002.
- [23] R. Schnall, A. Shlitner, J. Sheffy, R. Kedar, and P. Lavie, "Periodic, profound peripheral vasoconstriction – a new marker of obstructive sleep apnea," *Sleep*, vol. 22, no. 7, pp. 939–946, 1999.
- [24] A. Bar, G. Pillar, I. Dvir, J. Sheffy, A. Shlitner, and P. Lavie, "Evaluation of a portable device based on peripheral arterial tone for unattended home sleep studies," *Chest*, vol. 123, no. 3, pp. 695–703, 2003.
- [25] T. Penzel, J. McNames, P. de Chazal, B. Raymond, A. Murray, and G. Moody, "Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings," *Med Biol Eng Comput*, vol. 40, pp. 402–407, 2002.
- [26] P. de Chazal, T. Penzel, and C. Heneghan, "Automated detection of obstructive sleep apnoea at different time scales using the electrocardiogram," *Physiol Meas*, vol. 25, pp. 967–983, 2004.
- [27] W. B. van de Graaf, "Thoracic influence on upper airway patency," *J Appl Physiol*, vol. 65, pp. 2124–2131, 1988.
- [28] D. E. Hollowel and P. M. Surrat, "Mandible position and activation of submental and masseter muscles during sleep," *J Appl Physiol*, vol. 71, no. 6, pp. 2267–2273, 1991.
- [29] R. Poirrier, B. Chakar, A. Lacroix, D. Dive, I. Hansen, and G. Frank, "Relationship between mouth opening and intrathoracic pressure during sleep apnea," in *Somnologie suppl. of the 5th World Congress on Sleep Apnea*, no. 2, p. 37, Feb. 14-18 1997.
- [30] R. Poirrier, "Etude du comportement de la mandibule au cours des arythmies ventilatoires du sommeil. contribution la physiopathologie du syndrome des apnées obstructives et mise au point d'un système de dépistage." Thèse d'Agrégation de l'enseignement supérieur, Université de Liège, Belgique, 1998.
- [31] K. Miyamoto, M. M. Özbek, A. A. Lowe, T. T. Sjöholm, L. L. Love, J. A. Fleetham, and C. F. Ryan, "Mandibular posture during sleep in patients with obstructive sleep apnoea," *Archives of Oral Biology*, vol. 44, pp. 657–664, 1999.
- [32] F. Senny, J. Destiné, J. Verly, P. Ansay, and R. Poirrier, "Automatic scoring of sleep apnea and hypopnea by the analysis of mandibular movements," in *Proceedings of the 3rd IASTED international conference on Biomedical Engineering*, pp. 373–378, Feb. 14-18 2005.
- [33] F. Senny, J. Destiné, P. Ansay, L. Cambron, P. Maquet, and R. Poirrier, "Diagnosis of obstructive sleep apnoea and hypopnoea syndrome based on automatic analysis of mandibular movements," in *Proceedings of the World Association of Sleep Medicine 1st Congress*, pp. 73–77, Oct. 15-18 2005.
- [34] NOMICS. <http://www.nomics.be>.
- [35] A. Rechtschaffen and A. Kale, *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects*. Washington, DC: National Institutes of Health, 1968.
- [36] E. Graas, V. Remouchamps, D. Dive, and R. Poirrier, "A broad range of new clinical applications for a magnetic sensor measuring distance on human body," in *Proceedings of the 3rd IEEE-EMBS International Summer School and Symposium on Medical Devices and Biosensors*, pp. 73–76, Sept. 4-6 2006.

- [37] B. Beckers, R. Poirrier, and J. Destin e, "Screening of sleep-disordered breathing through the recording of mandibular movements," in *Proceedings of the First Annual International IEEE-EMBS - Special Topic Conference on Microtechnologies in Medicine and Biology*, pp. 593–596, Oct. 12-14 2000.
- [38] "European data format (edf)." <http://www.edfplus.info/>.
- [39] "Lnknet." <http://www.ll.mit.edu/SST/lnknet/>.
- [40] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, 1999.
- [41] A. Aldroubi and M. Unser, eds., *Wavelets in Medicine and Biology*. CRC Press, 1996.
- [42] C. Ducottet, T. Fournel, and C. Barat, "Scale-adaptive detection and local characterization of edges based on wavelet transform," *Signal Processing*, vol. 84, pp. 2115–2137, 2004.
- [43] Y. Chen, Z. Tian, and Z. Miao, "Detection of singularities in the pressure fluctuations of circulating fluidized beds based on wavelet modulus maximum method," *Chemical Engineering Science*, vol. 59, pp. 3569–3575, 2004.

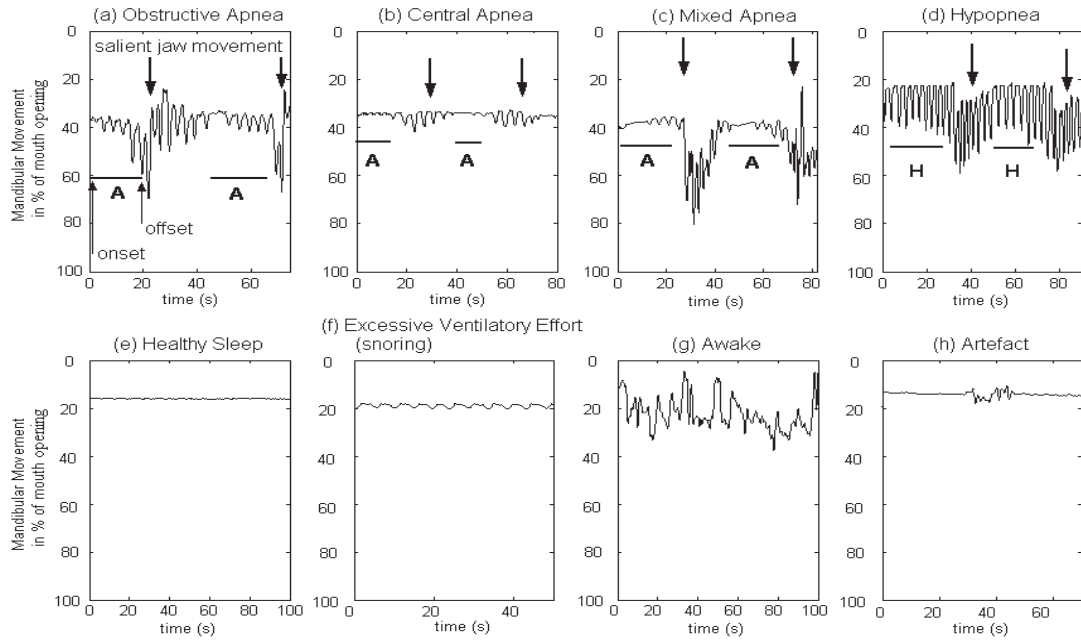


Fig. 1: .

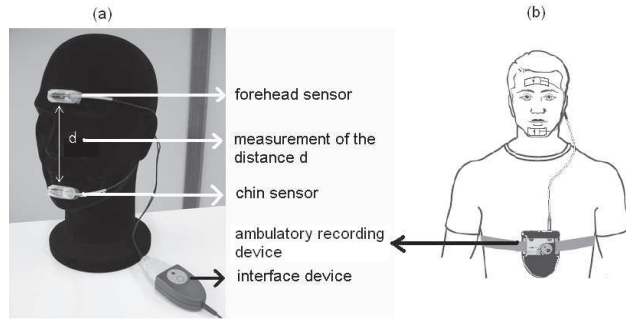


Fig. 2: .

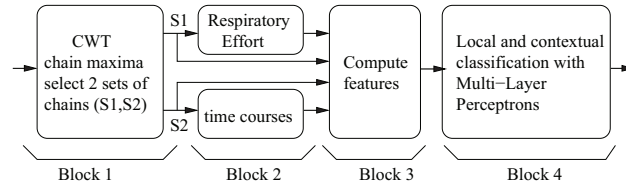


Fig. 3: .

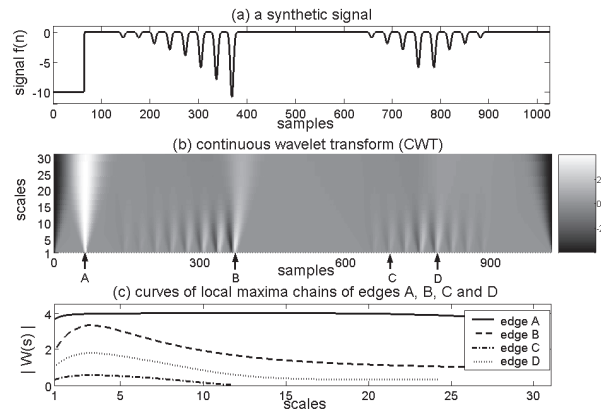


Fig. 4: .

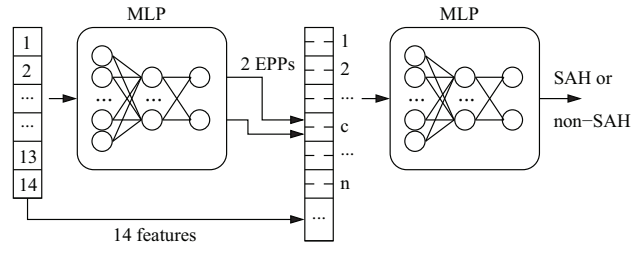


Fig. 5: .

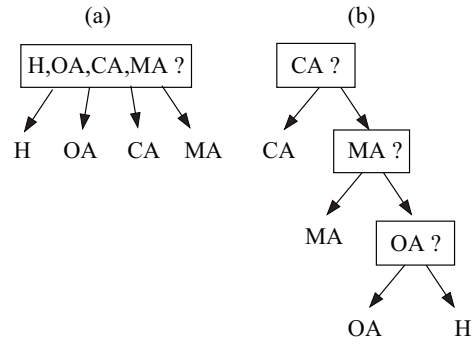


Fig. 6: .

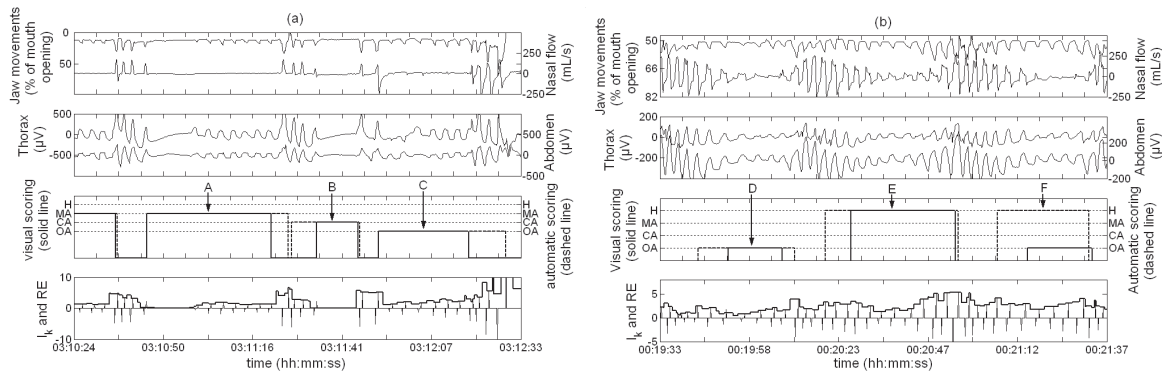


Fig. 7: .

Fig. 1: Typical patterns of sleep apneas and hypopneas (SAHs): (a) obstructive, (b) central, (c) mixed apneas and (d) hypopneas. A salient jaw movement (SJM, each arrow points at a SJM) appears after the offset of a SAH. Respiratory effort, correlated with esophageal pressure, is manifested as oscillations of the mandible. Typical non-SAH patterns: (e) healthy sleep, (f) snoring, (g) awake, (h) artefact (a swallowing). The mandibular movements signal is expressed in percentage of maximal mouth opening in function of time in seconds.

Fig. 2: (a) The midsagittal jaw movements sensor, the *JawSensTM* supplied by NOMICS, measures the distance d between two resonant circuits placed on the forehead and on the chin. The interface device allows to plug the sensor to any polysomnographic system. (b) The ambulatory recording device *BrizzyTM*.

Fig. 3: Block diagram of the analysis of midsagittal jaw movements.

Fig. 4: (a) A synthetic signal, (b) its continuous wavelet transform (CWT) and (c) the chains of the wavelet transform modulus maxima in function of scales of four edges noted A, B, C and D are illustrated. The computation of the area under the curves drawn by maxima chains was used to select and characterize relevant jaw movements related to salient jaw movements (edges A, B and D) and respiratory effort (edge C).

Fig. 5: Detection of SAHs using two multi-layer perceptrons (MLPs). The 14 features are listed in Table I and marked by a ^{d} , the others are the estimated posterior probabilities (EPPs) of n successive time courses, c is the current time course to be classified.

Fig. 6: Classification of hypopnea (H), obstructive apnea (OA), central apnea (CA) or mixed apnea (MA) using (a) one decision box or (b) three decision boxes in cascade. Each box is a multi-layer perceptron that has 7 inputs (see Table I) and no hidden layer.

Fig. 7: Illustration of the signal processing of jaw motion. Jaw movements (upper trace), nasal flow (lower trace), thoraco-abdominal movements, visual scoring (solid line) and automatic scoring (dashed line) and the quantization of respiratory effort (RE) based on jaw movements analysis are presented. (a) The three successive apneas labelled A, B, C (respectively a mixed, a central and an

obstructive apnea) were correctly classified and fairly well delineated. (b) These three obstructive events noted D, E and F (a hypopnea between two apneas) looked like each other when the jaw movements were considered, because these movements reflect effort and not respiration. Despite the difficulty to distinguish obstructive apneas from hypopneas, the obstructive feature was recognized.



Frédéric Senny received the civil Ing. and M. Eng. degree in electrical engineering from the University of Liège, Belgium, in 2002 and 2004. He is working towards the PhD degree at the department of electronic and computer science, Montefiore Institute, at the same University.

His main research interests are signal processing and medical informatics.



Jacques Destiné (S'70 - M '74) received the civil Ing. degree in electrical engineering and the Ph. D. degree from the University of Liège in 1972 and 1985 respectively.

He is now Professor in the electrical engineering and computer science department of the same university. His research and teaching interests lie in microsystems with applications ranging from DNA recognition to overhead power line monitoring.



Robert Poirrier received the PhD degree in neurology, with a thesis on cardio-vascular and breathing regulations during sleep, and the "qualified teacher" degree from the University of Liège in 1987 and 1998 respectively.

He is in charge of the Sleep Disorder Center of the University Hospital of Liège. His main research is devoted to human sleep and in particular to the Sleep Apnea Syndrome.