

Clasificación bayesiana de sujetos roncadores con Síndrome de Apnea Hipoapnea del Sueño mediante un método Kernel

Jordi Solà-Soler^{1,2}; José A. Fiz^{2,3}; José Morera³ y Raimon Jané^{1,2}

¹Departament ESAIL, Universitat Politècnica de Catalunya, ²Institut de Bioenginyeria de Catalunya (IBEC), Barcelona, España; CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España ({jordi.sola, raimon.jane}@upc.edu, ³Servicio de Neumología, Hospital Universitari Germans Trias i Pujol de Badalona, España, jafiz@msn.com, josepmorera.germanstrias@gencat.net

Resumen

El gold standard para diagnosticar el Síndrome de Apnea Hipoapnea del Sueño (SAHS) es la polisomnografía (PSG), un procedimiento caro y laborioso. Resultaría de mucha ayuda disponer de un método de screening sencillo que permitiera determinar la severidad de un sujeto antes de realizarle una PSG. En la literatura se han reportado diferencias significativas en las características acústicas del ronquido entre roncadores simples y pacientes con SAHS. La mayoría de estudios suele clasificar los sujetos en dos grupos a partir de un umbral de Índice de Apnea-Hipoapnea (AHI). Recientemente se ha propuesto un clasificador multigrupo bayesiano con estimación Gaussiana de la función de densidad de probabilidad (PDF), usando características del ronquido pero también información de las apneas. En este trabajo mostramos que un clasificador bayesiano con estimación Kernel de la PDF mejora la aproximación Gaussiana, y permite la clasificación de sujetos con SAHS en tres grupos en función de su severidad. El clasificador utiliza sólo información obtenida de los ronquidos. El método propuesto podría ser la base de un procedimiento de screening de SAHS con un solo canal basado en el ronquido.

1. Introducción

El Síndrome de Apnea Hipoapnea del Sueño (SAHS) es una patología ampliamente extendida cuyo síntoma más temprano suele ser el ronquido. Las repercusiones del ronquido pueden ir desde la no interrupción del sueño hasta una disrupción continua del mismo [1]. El *gold standard* para diagnosticar SAHS es la polisomnografía (PSG), un procedimiento laborioso y caro. Resultaría de mucha utilidad disponer de un método de screening que ayudara a los neumólogos a determinar a priori la severidad de un sujeto para priorizar mejor la programación de PSG.

Numerosos estudios han hallado diferencias significativas en las características acústicas del ronquido entre sujetos sanos y pacientes con SAHS [2]-[4]. Esos estudios suelen clasificar a los sujetos roncadores en dos grupos mediante un umbral de Índice de Apnea-Hipoapnea (AHI). Sin embargo, no proporcionan más detalle sobre la severidad del sujeto. Un trabajo reciente de nuestro grupo ha analizado a los sujetos roncadores

con SAHS en varios grupos [5]. Otros autores han usado un clasificador de Bayes basado en parámetros del ronquido y las apneas, con aproximación Gaussiana de la función de densidad de probabilidad (PDF) [6]. Pero en general los datos no siguen una distribución normal.

Nuestro estudio está basado en un solo canal, el sonido respiratorio. Basándonos únicamente en la información extraída de los ronquidos, investigamos las prestaciones de un clasificador de Bayes con estimación Kernel de la PDF comparándolo con la estimación Gaussiana. En cada caso, el conjunto de parámetros óptimo se obtiene mediante un algoritmo de selección automático.

2. Material y Métodos

2.1 Adquisición de señales y Base de datos

El sonido respiratorio se adquirió mediante un micrófono electret unidireccional encapsulado, acoplado a la piel a través de un encapsulado cónico, y sujetado al nivel del cartílago cricoides mediante una banda elástica. Los registros se llevaron a cabo durante las sesiones de polisomnografía en el Laboratorio del Sueño del *Hospital Universitari Germans Trias i Pujol* de Badalona. El sonido se amplificó y filtró paso-banda entre 70 y 2000 Hz con un filtro analógico Butterworth, y se digitalizó a 5000 muestras/s mediante un conversor A/D de 12 bits.

Los episodios de ronquido se identificaron mediante un detector automático previamente entrenado y validado, desarrollado por nuestro grupo [5],[7]. El detector se diseñó para identificar los episodios de ronquido en sujetos con y sin SAHS, y para rechazar el resto de sonidos respiratorios como son la respiración normal, la tos, la voz y otros artefactos. Para la clasificación de sujetos se usaron todos los ronquidos detectados durante la noche. En la Tabla 1 se muestran las características de los 36 sujetos de la base de datos (Número total de ronquidos=65625).

La American Academy of Sleep Medicine propone estratificar la severidad del SAHS en cuatro niveles de acuerdo a los umbrales $AHI=5h^{-1}$, $15h^{-1}$ y $30h^{-1}$ [8]. En este artículo abordamos la clasificación de sujetos roncadores en tres grupos de severidad de SAHS dados por los umbrales $AHI=5h^{-1}$ y $AHI=30h^{-1}$. De esta forma cada grupo tendrá un tamaño comparable dados los

TABLA 1
CARACTERÍSTICAS DE LA BASE DE DATOS

		G1	G2	G3	p ₁₂	p ₁₃	p ₂₃
Sujetos (N=36)	H/M	10/3	7/4	8/4	---	---	---
Edad (años)	m	45	48	52	0.706	0.096	0.460
	s	11	12	10			
BMI (kg/m ²)	m	27.1	28.9	32.9	0.339	0.012	0.085
	s	4.4	4.1	6.0			
AHI (1/h)	m	1.8	8.7	44.1	<0.001	<0.001	<0.001
	s	1.5	2.5	20.2			
Num.	m	1226	2243	2084			
Ronq.	s	1289	1177	937	0.068	0.044	0.580
	T	15942	24672	25011			

H=Hombres, M=mujeres, m=media, s= desviación estándar, BMI=Índice de masa corporal, AHI=Índice de apnea-hipoapnea, T=Número total de ronquidos, G1=AHI<5, G2=5≤AHI<30, G3=AHI≥30, p_{xy}=Significación estadística del Test U de Mann-Whitney U entre los grupos x e y.

sujetos disponibles en nuestra base de datos. Los grupos son G1 (no SAHS, AHI<5h⁻¹), G2 (SAHS ligero a moderado, 5h⁻¹≤AHI<30h⁻¹) y G3 (SAHS severo, AHI≥30h⁻¹). Todos los sujetos son roncadores (el número total de ronquidos por sujeto oscila entre 117 a 4214).

2.2 Caracterización del ronquido

En nuestros estudios previos hemos desarrollado diversas técnicas para el análisis y caracterización del ronquido en los dominios temporal y frecuencial. En el dominio temporal los ronquidos se caracterizan por el periodo de las vibraciones sonoras o *pitch*. La curva de *pitch* de cada ronquido se parametriza por: su valor medio (P_m); su desviación estándar (P_s); su rango intercuartílico (P_{iqr}); la densidad de *pitch* (P_{dens}), definida como la fracción de tiempo con *pitch* respecto a la duración total del ronquido; y el número de intervalos con *pitch* dentro de un ronquido (P_{ints}) [9].

El contenido frecuencial del ronquido se calcula a través de la Densidad Espectral de Potencia (PSD). La forma de la PSD se caracteriza mediante un conjunto de parámetros [10]: las frecuencias media, mediana, pico y máxima (F_{mean} , F_{med} , F_{peak} , F_{max}); la desviación estándar en frecuencia ($StdDev$); y los coeficientes de simetría y aplanamiento (C_{symm} , C_{Flatn}). La distribución de energía dentro de la PSD se mide a través de los ratios de potencia en tres bandas de frecuencias de interés: $B=(0,500)Hz$, $B=(100,500)Hz$ y $B=(0,800)Hz$. La energía en cada banda se calcula respecto a la energía total (RW_B) y respecto a la energía fuera de esta banda ($Rout_B$).

Las cavidades nasal y bucal introducen resonancias en el sonido del ronquido. Estas se pueden medir a través de los picos de la Envolvente Espectral, los llamados formantes. Cada formante se caracteriza por su frecuencia F_i , su amplitud M_i con respecto al máximo, y su atenuación L_i [11].

Las variables independientes X_k del modelo de clasificación se seleccionan de forma automática entre todos los parámetros del ronquido derivados de la intensidad sonora, la PSD, la Envolvente Espectral y el *pitch*. El número de variables independientes se ha limitado a 10.

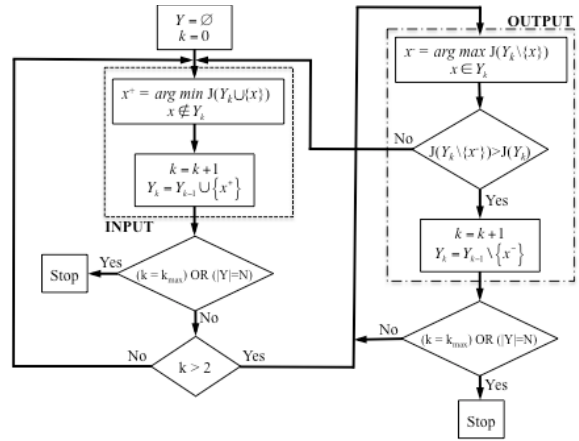


Fig. 1. Algoritmo de selección secuencial hacia adelante (SFFS) usado en el clasificador de Bayes. Las funciones de coste $J(\cdot)$ estudiadas son NPV_{G1} , ACC y TFN. Al finalizar, el modelo óptimo se considera $Y_{opt} = \arg \max \{J(Y_k)\}$.

2.3 Clasificador Bayes Naive

La regla de Bayes proporciona una forma directa de clasificación multigrupo de sujetos. Las variables independientes X_k se suponen independientes (aproximación Naive Bayes) de forma que su PDF conjunta pueda factorizarse como producto de las PDF's individuales. La aproximación Naive se ha reportado que da buenos resultados en muchos escenarios [12]. Para evitar cálculos con cantidades numéricas extremadamente pequeñas, se suele usar el logaritmo de las PDF's en lugar de las propias PDF's. De esta forma, en un modelo de Bayes con K variables independientes X_1, \dots, X_K , una observación dada (x_{j1}, \dots, x_{jK}) se clasificaría en el grupo G_{class} dado por:

$$\begin{aligned}
 G_{class} &= \underset{i=1:3}{\operatorname{argmax}} \{P(G_i) \cdot p_{X_1, \dots, X_K}(x_{j1}, \dots, x_{jK} | G_i)\} \approx \\
 &\approx \underset{i=1:3}{\operatorname{argmax}} \left\{ \hat{P}(G_i) \cdot \prod_{k=1}^K \hat{p}_{X_k}(x_{jk} | G_i) \right\} = \\
 &= \underset{i=1:3}{\operatorname{argmax}} \left\{ \log(\hat{P}(G_i)) + \sum_{k=1}^K \log(\hat{p}_{X_k}(x_{jk} | G_i)) \right\}
 \end{aligned} \quad (1)$$

La probabilidad $P(G_i)$ de pertenecer a cada clase se ha estimado por el número de sujetos en esa clase respecto al número total de sujetos. En estudios previos otros autores han usado una aproximación gaussiana para las PDF condicionales $p_{X_k}(X_k | G_i)$ de las variables independientes [5]. Si embargo, en general los datos disponibles no se distribuyen de forma normal, con lo cual es mejor utilizar una estimación de la PDF basada en Kernel. Dada una función de kernel simétrica $K(\cdot)$ y un conjunto de N observaciones $(x_{1k}, \dots, x_{Nk})^T$ de la variable independiente X_k , la PDF estimada viene dada por

$$\hat{p}_{X_k}(x) = \frac{1}{N} \sum_{j=1}^N K\left(\frac{x - x_{jk}}{h}\right) \quad (2)$$

donde h es el ancho de banda del kernel [12]. En este trabajo comparamos la estimación Kernel de la PDF con la estimación gaussiana. En ambos casos, el conjunto óptimo de variables independientes se selecciona de forma automática mediante el algoritmo de selección paso a paso hacia adelante (SFFS) descrito en la Fig.1. En cada paso, se seleccionan las variables que deben entrar o salir del modelo según una función de coste $J(\cdot)$ a optimizar.

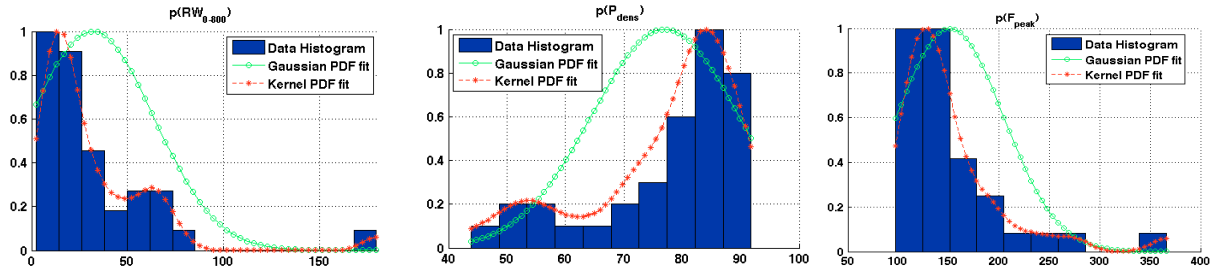


Fig. 2. Histograma de los datos, estimación de la PDF Gaussiana (círculos) y Kernel (asteriscos) para los parámetros del ronquido RW_{0-800} , P_{dens} y F_{peak} .

Considérese la siguiente matriz de confusión:

$$C = \begin{pmatrix} G1_{G1} & G2_{G1} & G3_{G1} \\ G1_{G2} & G2_{G2} & G3_{G2} \\ G1_{G3} & G2_{G3} & G3_{G3} \end{pmatrix} \quad (3)$$

donde G_{iG_j} es el número de sujetos del grupo G_j que han sido clasificados en el grupo G_i . Definimos tres funciones de coste de interés: el valor predictivo negativo (NPV) del grupo de sujetos sanos (G_1 , $AHI < 5h^{-1}$) dado por

$$NPV_{G1} = G1_{G1} / (G1_{G1} + G2_{G1} + G3_{G1}) \quad (4)$$

El número total de falsos negativos (TFN), eso es, el número de sujetos clasificados en un grupo con una severidad inferior a la que deberían:

$$TFN = G1_{G2} + G1_{G3} + G2_{G3} \quad (5)$$

Y la exactitud global en la clasificación (ACC), definida como

$$ACC = \frac{tr(C)}{sum(C)} \quad (6)$$

donde $sum(C)$ es la suma de todos los elementos y $tr(C)$ es la traza de la matriz de confusión. El algoritmo SFSS involucra una sola función de coste (principal) $J(\cdot)$. Si en una iteración dada el máximo o el mínimo de $J(\cdot)$ se alcanza para diversos conjuntos de variables, se utiliza una segunda función de coste $J2(\cdot)$ para seleccionar el mejor conjunto. Se ha estudiado las prestaciones del algoritmo SFSS con las 6 posibles combinaciones de funciones de coste (4)-(6) actuando como $J(\cdot)$ y $J2(\cdot)$.

Después de seleccionar el conjunto óptimo de variables independientes con el algoritmo SFSS, los modelos resultantes se evaluaron mediante la técnica de validación cruzada.

3. Resultados

La Fig. 2 muestra una comparación de los dos métodos de estimación de la PDF en tres parámetros del ronquido. Podemos observar como en general los datos no siguen una distribución normal. En algunos casos incluso se observa una distribución bimodal, la cual se aproxima bien con la estimación de la PDF basada en Kernel.

La evolución del algoritmo SFSS para las diferentes combinaciones de las funciones de coste $J(\cdot)$ y $J2(\cdot)$ se muestra en la Fig.3 para la estimación PDF basada en Kernel. Las mejores prestaciones se obtienen para $J=ACC$ (Fig.3a) y $J=NPV_{G1}$ (Fig.3b). En estos casos, los resultados resultan ser independientes de la elección de la medida secundaria $J2(\cdot)$, así que sólo se muestra una de las combinaciones.

La matriz de confusión (3) de los modelos óptimos se

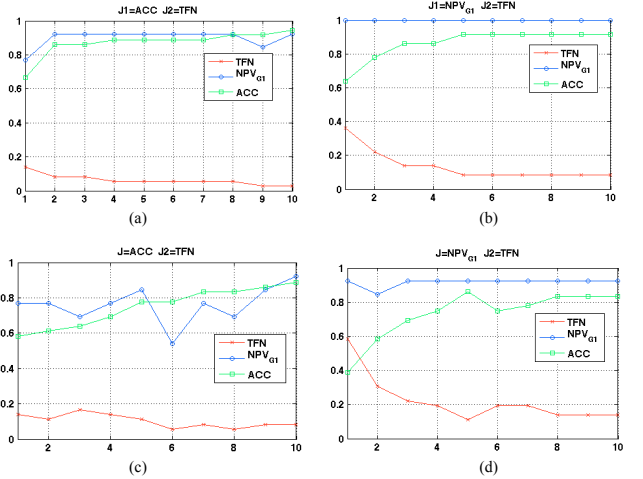


Fig. 3. Evolución del algoritmo SFSS para las diferentes combinaciones de las funciones de coste $J(\cdot)$ y $J2(\cdot)$ usando el método de estimación de la PDF (a,b) Kernel y (c,d) Gaussiano.

muestra en la Tabla 2 para la estimación de la PDF gaussiana, y en la Tabla 4 para la estimación de la PDF basada en Kernel. Se puede observar que el método basado en Kernel tiene mejores prestaciones que el método gaussiano en la mayoría de casos.

Las correspondientes matrices de confusión para la validación cruzada se muestran en las Tablas 3 y 5 para la estimación gaussiana y Kernel, respectivamente. En la validación, el modelo con PDF basada en Kernel usando $J=ACC$ proporciona el mejor balance de bajo TFN y alto NPV_{G1} (Tabla 5a,b).

4. Discusión y Conclusiones

La mayoría de estudios reportados suelen clasificar a los sujetos roncadores en dos grupos basándose en un solo umbral de AHI [2]-[4], o analizan parámetros en tres grupos de sujetos sin abordar la clasificación [5]. El clasificador multigrupo propuesto recientemente por Ben-Israel *et al.* [6] utiliza el método Bayes con información del ronquido, pero también con características relacionadas con las apneas, las cuales requieren sincronizar la señal acústica con las señales de la PSG. Esta complejidad adicional permite el uso de una estimación gaussiana sub-óptima de la PDF.

El método propuesto proporciona una forma de clasificar sujetos roncadores de acuerdo a la severidad de su SAHS, basándose exclusivamente en las características del ronquido obtenidas mediante un sistema de un solo canal. El método podría ser la base de un proceso de *screening* preliminar. El valor del NPV_{G1} debería ser lo más alto posible, ya que cada sujeto sano que se envía por

TABLA 2
MATRIZ DE CONFUSIÓN DE LOS MODELOS DE CLASIFICACIÓN ÓPTIMOS (ESTIMACIÓN PDF GAUSSIANA)

		J=ACC, J2=TFN NPV _{G1} =92.3% TFN=8.3%			J=ACC, J2=NPV _{G1} NPV _{G1} =76.9% TFN=5.5%			J=NPV _{G1} , J2=ACC NPV _{G1} =92.3% TFN=13.9%			J=NPV _{G1} , J2=TFN NPV _{G1} =100% TFN=13.9%		
		Grupo Clasificado (%)			Grupo Clasificado (%)			Grupo Clasificado (%)			Grupo Clasificado (%)		
		G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3
Grupo Original (%)	G1	92.3	0	7.7	76.9	7.7	15.4	92.3	0	7.7	100	0	0
	G2	27.3	72.7	0	9.1	90.9	0	9.1	90.9	0	0	100	0
	G3	0	0	100	0	8.3	91.7	8.3	25	66.7	0	25	75

(a)

(b)

(c)

(d)

TABLA 3
MATRIZ DE CONFUSIÓN DE LA VALIDACIÓN CRUZADA DE LOS MODELOS ÓPTIMOS (ESTIMACIÓN PDF GAUSSIANA)

		J=ACC, J2=TFN NPV _{G1} =46.1% TFN=36.1%			J=ACC, J2=NPV _{G1} NPV _{G1} =46.1% TFN=25.0%			J=NPV _{G1} , J2=ACC NPV _{G1} =61.5% TFN=33.3%			J=NPV _{G1} , J2=TFN NPV _{G1} =61.5% TFN=33.3%		
		Grupo Clasificado (%)			Grupo Clasificado (%)			Grupo Clasificado (%)			Grupo Clasificado (%)		
		G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3
Grupo Original (%)	G1	46.1	23.1	30.8	46.1	23.1	30.8	61.5	23.1	15.4	61.5	23.1	15.4
	G2	36.4	18.2	45.4	45.4	27.3	27.3	45.4	36.4	18.2	45.4	36.4	18.2
	G3	25.0	50.0	25.0	8.3	25.0	66.7	25.0	33.3	41.7	25.0	33.3	41.7

(a)

(b)

(c)

(d)

TABLA 4
MATRIZ DE CONFUSIÓN DE LOS MODELOS DE CLASIFICACIÓN ÓPTIMOS (ESTIMACIÓN PDF BASADA EN KERNEL)

		J=ACC, J2=TFN NPV _{G1} =92.3% TFN=2.8%			J=ACC, J2=NPV _{G1} NPV _{G1} =92.3% TFN=2.8%			J=NPV _{G1} , J2=ACC NPV _{G1} =100% TFN=8.3%			J=NPV _{G1} , J2=TFN NPV _{G1} =100% TFN=8.3%		
		Grupo Clasificado (%)			Grupo Clasificado (%)			Grupo Clasificado (%)			Grupo Clasificado (%)		
		G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3
Grupo Original (%)	G1	92.3	0	7.7	92.3	0	7.7	100	0	0	100	0	0
	G2	9.1	90.9	0	9.1	90.9	0	0	100	0	0	100	0
	G3	0	0	100	0	0	100	0	25	75	0	25	75

(a)

(b)

(c)

(d)

TABLA 5
MATRIZ DE CONFUSIÓN DE LA VALIDACIÓN CRUZADA DE LOS MODELOS ÓPTIMOS (ESTIMACIÓN PDF BASADA EN KERNEL)

		J=ACC, J2=TFN NPV _{G1} =53.8% TFN=30.6%			J=ACC, J2=NPV _{G1} NPV _{G1} =53.8% TFN=30.6%			J=NPV _{G1} , J2=ACC NPV _{G1} =53.8% TFN=41.7%			J=NPV _{G1} , J2=TFN NPV _{G1} =53.8% TFN=41.7%		
		Grupo Clasificado (%)			Grupo Clasificado (%)			Grupo Clasificado (%)			Grupo Clasificado (%)		
		G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3
Grupo Original (%)	G1	53.8	15.4	30.8	53.8	15.4	30.8	53.8	15.4	30.8	53.8	15.4	30.8
	G2	36.4	36.4	27.2	36.4	36.4	27.2	27.3	63.6	9.1	27.3	63.6	9.1
	G3	16.7	41.7	41.7	16.7	41.7	41.7	33.3	33.3	33.4	33.3	33.3	33.4

(a)

(b)

(c)

(d)

error a una segunda etapa (PSG) supone un alto coste para la Sanidad Pública. El TFN también debería mantenerse al mínimo, para no confundir ningún paciente con SAHS (G2 o G3) con un sujeto sano. Actualmente se está mejorando el método para obtener mejores resultados de validación.

Referencias

- [1] D.J. Eckert, A. Malhotra, "Pathophysiology of Adult Obstructive Sleep Apnea", *Proc Am Thorac Soc*, vol. 5, pp. 144-153, 2008.
- [2] A. K. Ng, T. S. Koh, E. Baey, T. H. Lee, U. R. Abeyratne, K. Puvanendran, "Could formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea?", *Sleep Med*, vol. 9(8), pp. 894-898, 2008.
- [3] A. S. Karunajeewa, U. R. Abeyratne, C. Hukins, "Multi-feature snore sound analysis in obstructive sleep apnea-hypopnea syndrome", *Physiol Meas*, vol. 32(1), pp. 83-97, 2011.
- [4] D. Pevernagie, R. M. Aarts, M. De Meyer, "The acoustics of snoring", *Sleep Med Rev*, vol. 14(2), pp. 131-144, 2010.
- [5] Fiz, J. A., Jané, R., Solà-Soler, J., Abad, J., García, M. Á. and Morera, J., "Continuous analysis and monitoring of snores and their relationship to the apnea-hypopnea index". *The Laryngoscope*, 120: 854-862. doi: 10.1002/lary.20815, 2010.
- [6] N. Ben-Israel, A. Tarasiuk, Y. Zigel, "Nocturnal sound analysis for the diagnosis of obstructive sleep apnea", *32nd Ann Intl Conf of the IEEE EMBS*, 2010, pp. 6146-6149.
- [7] R. Jané, J.A. Fiz, J. Solà-Soler, S. Blanch, P. Artis, J. Morera, "Automatic snoring signal analysis in sleep studies", *25th Ann Intl Conf of the IEEE EMBS*, 2003, Vol.1, pp. 366-369.
- [8] W. Flemons et al., "Home diagnosis of sleep apnea: a systematic review of the literature. An evidence review cosponsored by the American Academy of Sleep Medicine, the American College of Chest Physicians, and the American Thoracic Society", *Chest*, vol. 124, n° 4, p. 1543-1579, Oct. 2003.
- [9] J. Solà-Soler, R. Jané, J.A. Fiz, J. Morera, "Pitch analysis in snoring signals from simple snorers and patients with obstructive sleep apnea", *24th Ann Intl Conf of the IEEE EMBS*, vol.2, 2002, pp. 1527-1528.
- [10] J. Solà-Soler, R. Jané, J.A. Fiz, J. Morera, "Variability of snore parameters in time and frequency domains in snoring subjects with and without Obstructive Sleep Apnea", *27th Ann Intl Conf of the IEEE EMBS*, 2005, pp. 2583-2586.
- [11] J. Solà-Soler, R. Jané, J.A. Fiz, J. Morera, "Spectral envelope analysis in snoring signals from simple snorers and patients with Obstructive Sleep Apnea", *25th Ann Intl Conf of the IEEE EMBS* vol.3, 2003, pp. 2527-30.
- [12] A.K. Jain, R.P.W. Duin, Jianchang Mao, "Statistical pattern recognition: a review", *IEEE Trans on Patt Anal Mach Intel*, vol. 22, pp. 4-37, 2000.