

DETECCIÓN AUTOMÁTICA DE AROUSALS EN REGISTROS POLISOMNOGRÁFICOS

Ignacio Santamaría, David Gutiérrez,
Raquel Blanco, Blanca Cuesta
Departamento de Ingeniería de Comunicaciones
Universidad de Cantabria
e-mail: nacho@gtas.dicom.unican.es

Antonio Jiménez, Rosario Carpizo
Unidad de Trastornos del Sueño
Hospital Universitario Marqués de Valdecilla
e-mail: nmljga@humv.es

Abstract—The study and analysis of polysomnographic recordings is one of the main tools used for the diagnosis and treatment of sleep illnesses and disorders. One step of that analysis is the manual scoring by a specialist of different events that happen while the patient sleeps. The arousals, which are defined as abrupt changes in the EEG frequency, are one of those events. In this paper we describe a procedure for the automatic detection and scoring of arousals, which consists of three stages: feature extraction, detection and post-processing. In particular, we compare three different detectors: Bayesian, multilayer perceptron (MLP) and support vector machines (SVM). The performance of these detectors is evaluated using polysomnographic recordings taken from several patients.

I. INTRODUCCIÓN

Los arousals son breves alertamientos del cerebro que interrumpen la estructura habitual del sueño pero sin llegar a despertar completamente al sujeto que los padece [1]. La presencia de arousals es habitual en multitud de enfermedades y trastornos del sueño, por lo que habitualmente se suelen marcar y contar en los estudios polisomnográficos [2]. El marcaje manual de los arousals es una tarea tediosa, ya que pueden llegar a aparecer un elevado número de ellos en una sola noche. Esta situación hace que sea de gran interés el desarrollo de técnicas automáticas que, a partir de los registros polisomnográficos obtenidos del paciente, sea capaz de reconocer automáticamente la presencia de arousals y marcarlos correctamente para su posterior recuento y estudio.

En este artículo se propone un procedimiento para la detección automática de estos eventos que consiste en tres etapas: en la primera, a partir de la señal del electroencefalograma (EEG) se extraen, para cada instante de tiempo, una serie de índices que reflejan los cambios que caracterizan un arousal. En la segunda fase, un detector procesa los índices obtenidos decidiendo la presencia o ausencia de arousal (el detector toma una decisión por cada muestra del EEG). Por último, se aplica un postprocesado que consiste en una serie de reglas que permiten agrupar las decisiones individuales, eliminando además falsas alarmas. Para la segunda etapa se comparan las prestaciones de un detector Bayesiano, un MLP y una máquina de vectores soporte (SVM). La probabilidad de detección y falsa alarma para estos detectores es estimada a partir de un conjunto de arousals marcados por especialistas.

II. ETAPA I: EXTRACCIÓN DE CARACTERÍSTICAS

La *American Sleep Disorders Association* [1] establece que un arousal es un cambio abrupto en la frecuencia del EEG, que puede incluir actividad alfa, theta y/o frecuencias superiores a 16 Hz pero no “spindles”, sujeto a un conjunto de reglas y condiciones.

Este trabajo se centra en la detección de arousals en fases del sueño No REM, para lo cual únicamente es necesario medir de forma cuantitativa el cambio en la señal del EEG (que a partir de ahora llamaremos $x[n]$) de frecuencias lentas (principalmente delta) a frecuencias más rápidas (principalmente alfa). Un estudio detallado de los índices que reflejan estos cambios se hace en [3]. De entre ellos, en este estudio hemos encontrado relevantes los siguientes: la potencia total de la señal, la potencia en la banda alfa (7-12 Hz), la tasa de cruces por cero, la movilidad (definida más adelante) y la frecuencia media. También se han incorporado otros tres índices que no aparecen en [3], pero son de utilidad: la potencia en la banda low-delta (0.5-2 Hz), la frecuencia media en la banda alfa y la frecuencia media en la banda de 4 a 25 Hz.

Estos índices se calculan para cada una de las muestras de la señal del EEG ($f_s = 50$ Hz), utilizando una ventana deslizante de 128 muestras (≈ 2.5 segundos), centrada en la muestra de interés. Para calcular la potencia total y las potencias en bandas alfa y low-delta, se estima la densidad espectral de potencia de la señal inventanada, y posteriormente se integra dicha DEP en la banda de interés: 0-25 Hz. para la potencia total, 7-12 Hz. para la potencia alfa y 0.5-2 Hz. para la potencia low-delta.

Para la estima de la DEP, se han estudiado varios estimadores espectrales (periodograma, Blackman-Tuckey, Welch, etc.), obteniéndose resultados similares con todos ellos, por lo que al final se optó por utilizar el periodograma, dada su mayor sencillez.

La frecuencia media en la banda (f_1, f_2) se estima como el primer momento de la DEP en dicha banda

$$\mu_f = \left\langle \hat{\phi}(f) \right\rangle_{f_1}^{f_2} = \sum_{f_i=f_1}^{f_2} f_i \hat{\phi}(f_i), \quad (1)$$

siendo $\hat{\phi}(f)$ la estima de la DEP.

Otro de los índices empleados es la movilidad, que se define como el cociente de las desviaciones estándar de la amplitud del EEG y su primera derivada

$$M_b = \frac{\sigma_{x'}}{\sigma_x}, \quad (2)$$

la derivada de $x[n]$ se estima como la diferencia primera: $x'[n] = x[n] - x[n-1]$.

Por último, la tasa de cruces por cero se calcula como el número de veces que la señal del EEG pasa por cero por unidad de tiempo. Para incrementar la robustez frente al ruido se puede incluir un umbral del 5% de la desviación estándar de $x[n]$.

$$Z_c = \frac{f_s}{N} \sum_{n=1}^{N-1} \xi(n), \quad (3)$$

donde f_s es la frecuencia de muestreo, N es el número total de muestras del segmento de señal $x[n]$ sobre el que se está calculando este índice ($N = 128$), $\xi(n)$ es una función indicador que toma los siguientes valores

$$\xi(n) = \begin{cases} 1, & \text{si } x[n] > x_{th} \text{ y } x[n+1] < -x_{th} \\ 1, & \text{si } x[n] < -x_{th} \text{ y } x[n+1] > x_{th} \\ 0, & \text{resto de los casos;} \end{cases} \quad (4)$$

y $x_{th} = 0.05 \cdot \sigma_x$.

De esta forma, se obtiene una estima instantánea de los 8 índices seleccionados $I_j[n]$, para $j = 1, \dots, 8$. Aquellos en los que interviene la DEP, presentan una elevada variabilidad; para eliminarla se ha filtrado cada señal índice obtenida con un filtro paso bajo aplicando un factor de olvido K

$$I_j^s[n] = K \cdot I_j^s[n-1] + (1-K) \cdot I_j[n] \quad (5)$$

donde $I_j[n]$ es el índice original e $I_j^s[n]$ es el índice ya filtrado. El factor de olvido empleado ha sido $K = 0.98$.

Asimismo, existe gran variabilidad en los valores de los índices anteriores obtenidos para distintos pacientes [4]. Para eliminar esta variabilidad entre sujetos se trabajará con la desviación de los índices respecto a su media, en tanto por ciento. Así damos más importancia a los cambios en los valores de los índices que a su valor absoluto. En particular, se obtienen finalmente unos índices normalizados

$$I_j^{norm}[n] = \frac{I_j[n] - \hat{I}_j[n]}{\hat{I}_j[n]} \cdot 100, \quad (6)$$

donde $\hat{I}_j[n]$ es la media a largo plazo de dicho índice, que se ha calculado como la media de las N muestras del índice en torno a la muestra n

$$\hat{I}_j[n] = \frac{1}{N} \sum_{k=n-N/2+1}^{n+N/2} I_j[k], \quad (7)$$

El número de muestras utilizado para calcular la media a largo plazo es $N = 1000$, es decir, 20 segundos.

En la Figura 1 se muestra un segmento de señal del EEG de un minuto de duración, con un arousal típico marcado por un especialista. Se puede ver cómo durante el arousal aumenta la potencia alfa (P_α) y disminuye la potencia delta (P_δ). También aparecen otros artefactos que hacen que aumente P_δ , pero que no llevan asociados cambios en P_α .

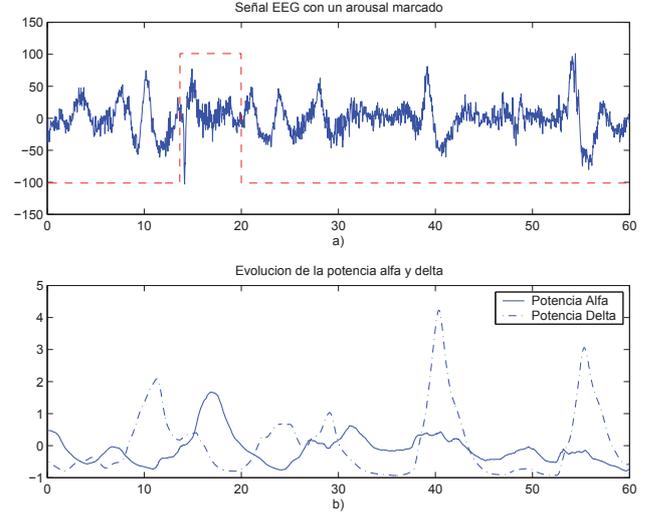


Fig. 1. a) EEG con un arousal típico y b) evolución de los índices potencia alfa y delta en dicho arousal

Por último, con objeto de introducir la dependencia temporal que existe en los índices seleccionados durante un arousal, el vector de características con el que entrenar los detectores se forma agrupando los índices obtenidos en n , $n-N$ y $n-2N$. De esta manera se forma un vector de características con 24 componentes: $\mathbf{x}_n = (I_1[n], \dots, I_8[n], I_1[n-N], \dots, I_8[n-N], I_1[n-2N], \dots, I_8[n-2N])^T$. En concreto, se ha usado $N = 25$ (muestras retrasadas medio segundo y un segundo).

III. ETAPA II: DETECCIÓN

En la siguiente etapa, cada vector de características asociado con cada una de las muestras del EEG se clasificará como "arousal" o "no arousal". En esta sección se describen los tres esquemas de detección estudiados: detección Bayesiana, perceptrón multicapa (MLP) y máquina de vectores soporte (SVM).

A. Detección Bayesiana

Para cada muestra n de la señal del EEG la etapa de extracción de características proporciona un vector \mathbf{x}_n , que tendremos que clasificar en una de las dos clases posibles: C_0 (no existe arousal) y C_1 (existe arousal)

$$H_0 : \mathbf{x}_n \in C_0$$

$$H_1 : \mathbf{x}_n \in C_1$$

Para resolver este problema, emplearemos en primer lugar el detector óptimo bajo el criterio MAP, es decir, máxima probabilidad a posteriori. El detector óptimo MAP decide la clase C_1 si la probabilidad a posteriori de que una observación

\mathbf{x}_n pertenezca a C_1 es mayor que la probabilidad a posteriori de pertenecer a C_0

$$P(C_1|\mathbf{x}_n) \underset{H_0}{\overset{H_1}{\geq}} P(C_0|\mathbf{x}_n). \quad (8)$$

Aplicando el teorema de Bayes, el criterio de decisión queda

$$f(\mathbf{x}_n|C_1)P(C_1) \underset{H_0}{\overset{H_1}{\geq}} f(\mathbf{x}_n|C_0)P(C_0). \quad (9)$$

En este estudio, modelamos las verosimilitudes $f(\mathbf{x}_n|C_i)$ como una variable aleatoria Gaussiana multidimensional

$$f(\mathbf{x}_n|C_i) = \frac{1}{(2\pi)^{\frac{M}{2}} \sqrt{|\Sigma_i|}} e^{\left(-\frac{1}{2}(\mathbf{x}_n - \mu_i)^T \Sigma_i^{-1} (\mathbf{x}_n - \mu_i)\right)}, \quad (10)$$

siendo $\mu_i = E[\mathbf{x}_n^i]$ y

$$\Sigma_i = E[(\mathbf{x}_n^i - \mu_i)(\mathbf{x}_n^i - \mu_i)^T],$$

donde \mathbf{x}_n^i denota los vectores pertenecientes a la clase i -ésima. La aplicación del detector Bayesiano requiere estimar los valores de medias y covarianzas para cada clase (μ , Σ), así como las probabilidades a priori $P(C_i)$. Estos valores pueden estimarse de manera sencilla en una fase previa de entrenamiento del detector, a partir de un conjunto de datos etiquetados.

B. El Perceptrón Multicapa (MLP)

Las redes neuronales (en concreto, en este estudio hemos empleado el perceptrón multicapa) presentan la ventaja de que no es necesario hacer ninguna suposición acerca de la distribución de los datos de entrada. Un MLP es capaz de “aprender” la distribución de dichos datos, así como su relación con las salidas deseadas (la presencia o no de arousals), durante una fase de entrenamiento [5].

Tras un estudio exhaustivo, se ha seleccionado finalmente un MLP con estructura [24, 60 1], es decir, 24 entradas, 60 neuronas en la capa oculta y una única neurona en la capa de salida. El entrenamiento se ha realizado mediante el algoritmo de “retropropagación” sobre un conjunto de patrones de entrada previamente etiquetados como -1 para el caso “no arousal” y $+1$ para el caso “arousal”. Como funciones de activación se han empleado tangentes hiperbólicas en la capa oculta y en la capa de salida, que nos proporcionan valores de salida entre -1 y $+1$.

C. Máquinas de Vectores Soporte (SVM)

Las máquinas de vectores soporte (SVM) permiten la resolución de problemas de clasificación, regresión o estimación [6]. En un problema de clasificación binario como el que nos ocupa, una SVM proporciona el hiperplano óptimo que maximiza el margen de separación entre las dos clases. Típicamente, el hiperplano de máximo margen se obtiene en un espacio de dimensión mayor (incluso de dimensión infinita) que el de entrada. Bajo ciertas condiciones (condición de

Mercer), el producto escalar en el espacio de salida se puede escribir a través de un cierto núcleo $K(\mathbf{x}, \mathbf{y})$

$$(\phi(\mathbf{x}_1) \cdot \phi(\mathbf{x}_2)) = K(\mathbf{x}_1, \mathbf{x}_2). \quad (11)$$

En estas condiciones sólo es necesario conocer el núcleo $K(\mathbf{x}_i, \mathbf{x}_j)$, y no la función de transformación $\phi(\mathbf{x}_i)$.

La salida de la SVM viene dada por

$$f(\mathbf{x}) = \text{sign} \left(\sum_n \alpha_n y_n K(\mathbf{x}, \mathbf{x}_n) + b \right). \quad (12)$$

donde \mathbf{x}_n son los patrones de entrenamiento, y_n las salidas deseadas y α_n son los multiplicadores de Lagrange, muchos de los cuales son cero. Los vectores asociados a los multiplicadores no nulos se denominan “vectores soporte”, y son los que determinan el hiperplano de decisión. El resto son irrelevantes a los efectos de clasificación. Los multiplicadores de Lagrange α_n y el término de umbral b se calculan en la fase de entrenamiento de la SVM a partir de un conjunto de datos de entrenamiento (\mathbf{x}_n, y_n) , resolviendo para ello un problema de programación cuadrática (QP), para el que existen implementaciones eficientes.

Como núcleo se pueden emplear funciones polinómicas, funciones gaussianas de base radial o sigmoides. En este trabajo en concreto se han empleado el siguiente kernel gaussiano: $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-(\mathbf{x}_i - \mathbf{x}_j)^2 / \sigma^2)$, con $\sigma = 1$. Se obtiene así una máquina con 3800 vectores soporte.

IV. ETAPA III: POSTPROCESADO

A la salida de cada detector se toman decisiones individuales que es necesario someter a un postprocesado antes de compararla con los arousals marcados por el especialista. En primer lugar, el postprocesado pretende reducir la variabilidad que existe a la salida del detector (Bayesiano, MLP o SVM). Para ello, se realiza un promediado integrando la señal de salida del detector (antes de tomar la función signo) en una ventana deslizante de tres segundos de longitud (150 muestras, duración mínima de un arousal) alrededor de cada muestra. Posteriormente se aplica un umbral duro que nos proporciona a la salida dos únicos valores: -1 para la clase C_0 y $+1$ para la clase C_1 .

En la segunda fase del postprocesado se implementan algunas reglas básicas para el marcaje de arousals, tal y como se explica en [1] y [2]. Básicamente, las reglas implementadas consisten en eliminar los arousals de una duración menor de 3 segundos y superior a 20 segundos. Además, la separación mínima entre arousals ha de ser de 10 segundos; si no es así se eliminará el arousal más corto de los dos contiguos.

V. RESULTADOS

Para evaluar las prestaciones de los sistemas aquí propuestos se han utilizado registros polisomnográficos procedentes de la Unidad de Trastornos del Sueño del Hospital Universitario Marqués de Valdecilla. De estos registros se ha extraído el canal del EEG y se ha diezmado la señal para trabajar con 50 muestras por segundo. Se han seleccionado segmentos de señal con gran cantidad de arousals marcados por especialistas. Para

TABLE I
RESULTADOS OBTENIDOS SOBRE DISTINTOS CONJUNTOS DE TEST

Conjunto de test	Bayesiano		MLP		SVM	
	P_d	P_{fa}	P_d	P_{fa}	P_d	P_{fa}
Test1(c1)	0.373	0.099	0.675	0.052	0.423	0.060
Test1(c2)	0.868	0.554	0.967	0.144	0.750	0.240
Test2(c1)	0.289	0.082	0.476	0.199	0.432	0.134
Test2(c2)	0.785	0.565	0.846	0.374	0.846	0.214

el conjunto de datos de entrenamiento se han seleccionado segmentos del EEG de varios pacientes, con una duración total de 132 minutos: 398541 muestras o patrones, de los que 55213 corresponden a arousal y el resto a no arousal. A continuación se han realizado dos test: en el Test 1, a partir de los mismos pacientes usados en el entrenamiento, se han seleccionado nuevos segmentos no utilizados anteriormente, que en total suman 130 minutos de señal del EEG (390016 muestras o patrones) y un número total de 153 arousals. En el Test2 se han utilizado segmentos de nuevos pacientes no utilizados durante la fase de entrenamiento, con una duración total de 120 minutos (360429 muestras o patrones) y un número total de 130 arousals.

Para comparar el comportamiento de los distintos detectores se han utilizado dos criterios de detección distintos. El primero (c1) considera la señal directamente a la salida de cada detector, antes de aplicar el postprocesado, evaluando la probabilidades de detección y falsa alarma sobre las decisiones individuales (muestra a muestra). Empleando este criterio se estima la probabilidad de detección y falsa alarma a partir del conjunto de Test 1 ó 2.

El segundo criterio (c2) estima las probabilidades de detección y falsa alarma una vez realizado el postprocesado. En este caso se estiman las probabilidades considerando los eventos arousals marcados correcta o incorrectamente (no las decisiones individuales). Se considera que un arousal ha sido correctamente detectado cuando se solapa con uno marcado por el especialista; de lo contrario, se cuenta ese arousal como una falsa alarma. Se estima la probabilidad de detección P_d como el número de arousals detectados correctamente entre el número total de arousals marcados por el especialista. A su vez, la probabilidad de falsa alarma P_{fa} se estima como el número de falsas alarmas entre el número total de arousals detectados por nuestro sistema.

En la Tabla I se muestran los resultados obtenidos: el MLP y la SVM ofrecen resultados semejantes y notablemente superiores a los conseguidos por el detector Bayesiano. Los resultados se pueden considerar como buenos si se tiene en cuenta que entre dos especialistas distintos pueden existir discrepancias en el marcaje de los arousals de hasta el 10%.

En la Figura 2 se muestra un segmento de señal del EEG junto con el arousal marcado por el especialista y los marcados el esquema de detección empleando un MLP, antes y después del postprocesado.

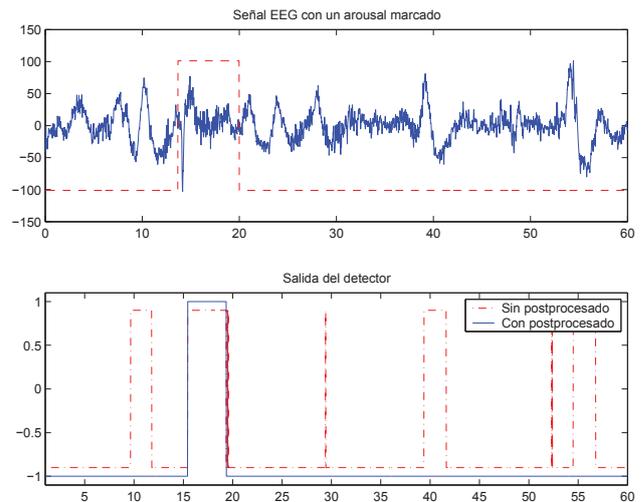


Fig. 2. Arousal marcado por el especialista y salida del detector (MLP).

VI. CONCLUSIONES

En este trabajo se ha propuesto un sistema para la detección automática de arousals que consta de tres etapas: extracción de características, detección muestra a muestra de los eventos y postprocesado. La primera etapa obtiene una serie de índices correspondientes a la potencia en distintas bandas, frecuencia media, etc; calculados en distintos instantes de tiempo. Tres esquemas de detección distintos han sido aplicados: detector Bayesiano, MLP y SVM. Una etapa de postprocesado final permite el marcaje de los eventos. Los distintos esquemas han sido evaluados sobre un conjunto amplio de datos reales con arousals marcados por especialistas. La hipótesis de gaussianidad asumida en el detector Bayesiano ocasiona una degradación en los resultados ofrecidos por este esquema. Los esquemas basados en estructuras de aprendizaje (MLP o SVM), que no realizan ninguna hipótesis sobre los datos de entrada, ofrecen claramente mejores resultados.

Es posible conseguir mejoras adicionales incorporando al proceso de detección información de otras señales del registro polisomnográfico: electromiograma, electrooculograma, etc. Esto permitiría, además, la detección de arousals en fase REM.

REFERENCES

- [1] Atlas Task Force of the American Sleep Disorders Association. *EEG Arousals: Scoring rules and examples*. Sleep, 15(2):174–184, 1992.
- [2] Rechtschaffen A, Kales A, eds. *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects*. Los Angeles: Brain Information Service/Brain Research Institute, University of California Los Angeles, 1968.
- [3] M. J. Drinnan, A. Murray, J. E. S. White, A. J. Smithson, C. J. Griffiths and G. J. Gibson. *Automated recognition of EEG changes accompanying arousal in respiratory sleep disorders*. Sleep, 19(4):296–303, 1992.
- [4] De Carli F, Nobili L, Gelcich, Ferrillo F. *A method for the automatic detection of arousals during sleep*. Sleep, vol. 22, no. 5, págs. 561–572, 1999.
- [5] Haykin S. *Neural Networks: A Comprehensive Foundation*. New York: Macmillan Publishing Company, 1994.
- [6] Vapnik V.N. *The nature of statistical learning theory*. New York: Springer-Verlag, 1999.