AUTOMATIC DETECTION OF SLOW EYE MOVEMENTS BY DISCRETE WAVELET TRANSFORM

E. Magosso*, F. Provini**, P. Montagna** and M. Ursino*

* Department of Electronics, Computer Science and Systems, University of Bologna, Bologna, Italy
** Department of Neurological Sciences, Polysomnographic Laboratory, University of Bologna, Bologna, Italy

emagosso@deis.unibo.it

Abstract: Electro-oculographic (EOG) activity during wake-sleep transition is characterized by the appearance of slow eye movements (SEM). In the present work, an algorithm for automatic identification of SEM episodes from EOG signal is developed. The algorithm is based on a 10-level wavelet decomposition of the EOG recording by using the Daubechies order 4 wavelet. Energies at each scale are used to build a discriminant function, which expresses the ratio between energies of high-scale details and energies of both low-scale and high-scale details. The main assumption is that the value of the discriminant function increases above a given threshold during SEM episodes because of energy redistribution towards higher scales. Parameters of the algorithm were tuned on three EOG recordings (training set), previously classified by human experts. Then, the algorithm was validated against visual scoring on other seven recordings (testing set). Performances are as follows: agreement = 80%; sensitivity = 67%; selectivity = 84%. However, most errors are not imputable to inability of the system to detect intervals with SEM activity against non-SEM intervals, but are merely due to a different collocation of the beginning and termination of some SEM episodes. The proposed method may represent a valuable support in the field of computerized EOG analysis.

Introduction

Detection of drowsiness is becoming increasingly important not only in normal humans but also for the diagnosis and research of several pathological conditions [1]. However, criteria for assessing the transition from wakefulness to sleep have not been standardized yet. Electro-oculographic activity (EOG) is considered a possible indicator of sleep onset [2-5]: indeed, EOG, during the transition from wakefulness to sleep, is characterized by the appearance of rolling slow eye movements (SEMs). SEM activity starts before the onset of stage 1 sleep, continues through stage 1, then it declines progressively during the first minutes of stage 2.

Visual inspection of EOG tracings for scoring SEMs is routinely performed in the clinical practice. However, visual scoring of SEMs suffers of inherent problems of inter-scorer variability: decisions of human scorers are unavoidably affected by individual bias. Moreover, visual inspection of EOG tracing during sleep requires much time and a high level of expertise.

Despite these disadvantages of the visual scoring, only few works can be found in the literature facing with the automatic identification of slow eye movements [6]. Moreover, the previously proposed algorithms, based on filtering techniques, showed poor performances in detecting SEMs. The reason for this difficulty is that some SEM events have a transient nature being characterized by one oscillation period or less, hence showing a highly non-stationary behaviour. This makes their study by conventional methods (such as Fourier analysis) problematic. Therefore, automatic detection of SEMs may benefit from the use of the wavelet transform, which is a method specialized for the analysis of non-stationary signals [7-9].

The purpose of the present study was to realize a wavelet-based method for automatic detection of SEM waveforms, with their start instant and duration. Wavelet transform is applied to bipolar electro-oculograms, acquired in the context of polygraphic recordings. The EOG signal is decomposed at different scales, each scale associated with a time resolution and frequency band. The automatic detector is based on a discriminant function, which expresses the ratio between energy associated to low-frequency scales and energy associated to both low-frequency and high-frequency scales. The algorithm was tuned on the basis of a training set of EOG recordings, on which two experts had previously marked the SEM events. Then, another set of previously classified recordings (testing set) was used to validate the method.

Materials and Methods

Clinical Material: 10 EOG recordings (indicated in the following by numbers from 1 to 10) acquired in 10 male subjects during overnight polysomnography, were extracted from the database of the Centre for Sleep Disorders (Department of Neurological Sciences, Bologna). Two EOG channels (E1-A1, E2-A1) were recorded from electrodes placed according to the recommendation in the manual of Rechtschaffen and Kales [10]. Sampling rate was 128 Hz. Only a portion of

ISSN: 1727-1983

each overnight recording session was considered for the analysis, from pre-sleep wakefulness to stage 2 sleep.

Each of the ten EOG recordings included in the present project was first inspected separately by two human experts, to score SEMs, and then revised jointly to solve the disagreement. According to the clinical literature [2-5], the following scoring criteria for visual classification of a SEM event were established: i) slow sinusoidal excursion (0.2-0.6 Hz) lasting more than 1 s with amplitude above 20 μV; ii) binocular synchrony with opposed-phase deflections in the two channels; iii) absence of artefacts (such as blinks, EEG/EMG artifacts). Each event scored by the experts was filed in the computer memory, with its starting instant and duration.

Detection algorithm: Goal of the algorithm is to identify automatically the start instant and the duration of each SEM event. The software program processes the data from the two EOG channels and consists of four main stages (see fig.1).

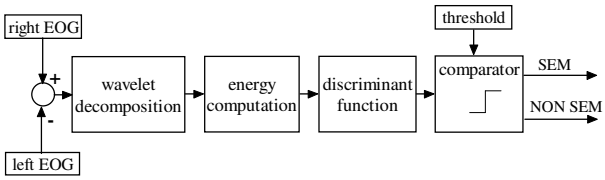

Figure 1: Block diagram of the automatic detector.

In the first stage, the discrete wavelet transform is applied to the overall length of the EOG recording.

For a given signal, $s(t)$, initially represented by means of its coefficients at resolution 0, the wavelet decomposition can be written as follows:

$$s(t)=\sum_{k=-\infty}^{+\infty} c_{0,k}\varphi(t-k)=\sum_{k=-\infty}^{+\infty} c_{N,k}\varphi(2^{-N}t-k)+\sum_{j=1}^{N}\sum_{k=-\infty}^{+\infty} d_{j,k}\,2^{-j/2}\psi(2^{-j}t-k)=A_N(t)+\sum_{j=1}^{N} D_j(t) \quad (1)$$

$c_{N,k}$ represent approximation coefficients at level N, while $d_{j,k}$ $j=1,\ldots N$ represent detail coefficients or wavelet coefficients at level j. $\psi(t)$ is the wavelet function, while $\varphi(t)$ is a companion function, named scaling function. Wavelet functions describe signal high frequency components, while scaling function describes smooth components of the signal.

Eq. (1) implements a multiresolution analysis of the signal, that is the signal is decomposed into N details (D_j, $j=1\ldots N$) and one approximation (A_N) at level N. Approximation is the residual function representative of all the higher scales of the signal, after the detail functions are computed.

Extending decomposition over all resolutions levels, the complete wavelet expansion can be obtained:

$$s(t)=\sum_{j=1}^{+\infty}\sum_{k=-\infty}^{+\infty} d_{j,k}\,2^{-j/2}\psi\left(2^{-j}t-k\right) \quad (2)$$

According to Eq. (2), wavelet coefficients $d_{j,k}$ represent weighting factors for synthesizing a given signal from translated and dilated versions of a mother wavelet, with scale parameter $a=2^j$ and translation parameter $b=k2^j$.

When the family $\{\psi_{j,k}(t) = \psi(2^{-j}t-k)\}$ forms an orthonormal basis, then wavelet coefficients can be used to compute energy associated with the detail j. The energy series associated to coefficient series $d_{j,k}$ is:

$$E_{j,k}=\left|d_{j,k}\right|^2, \quad (3)$$

and the energy associated to the entire signal is:

$$E=\|s(t)\|^2=\sum_{j=1}^{\infty}\sum_{k=-\infty}^{k=+\infty}\left|d_{j,k}\right|^2 \quad (4)$$

In our application to EOG signals, we applied a 10 level wavelet decomposition to the difference among the right and left EOG, by using Daubechies order 4 wavelet, as mother wavelet. Daubechies order 4 wavelet was found to be the most appropriate because of its likeness to the event under analysis. The ten level decomposition is necessary to extract the frequency range of slow eye movements (typically 0.2-0.6 Hz).

Figure 2 shows the 10 level decomposition of a small section of EOG (recording 2) containing an SEM event from t = 174 s to t = 234 s, according to the two experts. The 10 detail functions ('D') and the approximation function ('A') are reconstructed from wavelet coefficients and scaling coefficients, respectively. The original signal is the superimposition of details D_1-D_{10} and approximation A_{10}.

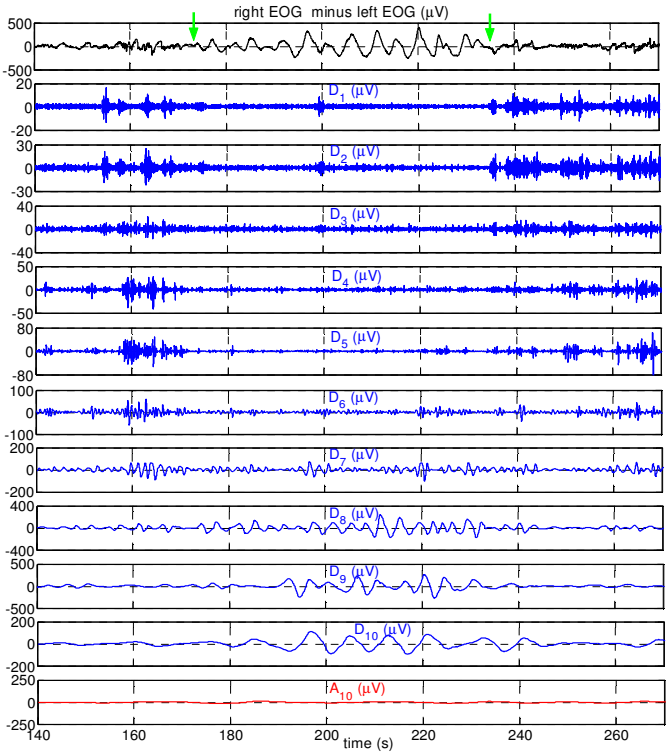

Figure 2: Portion of the EOG tracing belonging to patient 2 containing an SEM episode according to clinicians (between arrows), and its wavelet decomposition into 10 levels of details (D_1-D_{10}) and one approximation A_{10}. The decomposed signal is the difference between right and left EOG.

ISSN: 1727-1983

It is worth noticing, that in the figure, during SEM activity, details have maximum value in scales 8÷10.

In the second stage of the algorithm, energy coefficients are computed from wavelet coefficients according to Eq. (3). Indeed, as shown in Fig. 2, during SEM activity the energy of the higher details rises, while that of lower ones decreases. This suggests that an appropriate function of energy at different scales can be built, which assumes high values in presence of SEM, and lower values otherwise. However, this function cannot be developed by using energy series as computed by Eq. (3), since detail coefficients have different density and localization on the temporal axis. In fact, since we are using a dyadic multiresolution analysis (i.e., the scale and translation parameter are based on powers of two), at every level of decomposition we get a half of the wavelet coefficients with respect to the previous level. This drawback of the dyadic scheme is overcome by processing energy coefficients (obtained by Eq. 3), in order to achieve uniformly distributed 'atoms' of energy throughout all the scales, with each 'atom' of energy positioned every 0.5 seconds. A time resolution of 0.5 s is suitable for detection of SEM events, since a slow eye movement lasts not less than 1 s. Obviously, this procedure has been applied so that total energy associated to each level j is preserved. At the end of this process, we obtain energy series, $E_j(n)$ (j = 1÷10), each representing the energy of detail j sampled every 0.5 s (i.e., at the instants $t_n = n \cdot 0.5$ s).

In the third stage, the energy series $E_j(n)$ are used to define a discriminant function, which has the following structure:

$$f(n) = \frac{N(n)}{D(n)} = \frac{\sum_{j=7}^{10} W_j \cdot E_j(n)}{\sum_{j=3}^{5} W_j \cdot E_j(n) + N(n)} \qquad (5)$$

where W_j are constant weights. The numerator is the weighted sum of energies associated to low frequency scales (7÷10); the denominator is composed by the numerator plus the weighted sum of energies associated to lower scales (3÷5). Hence, the discriminant function expresses a ratio of specific energy combinations at lower frequencies with respect to both lower and higher frequency components. The choice of considering scales 7÷10 at numerator and 3÷5 at denominator was suggested by visual inspection of energy series obtained by wavelet decomposition of recording 1. When an SEM occurs (as marked by the two experts), energy associated to scales 7÷10 increases, while energy associated to scales 3÷5 decreases; the opposite occurs in absence of SEMs. Behaviour of energy series at scale 6 is more ambiguous, hence it was not included in the discriminant function. Finally, scales 1 and 2, which capture high frequency components, were not introduced in the denominator, since high frequency noise can be superimposed over slow movements.

Finally, the discriminant function $f(n)$ obtained by Eq. (5) has been smoothed trough a 3 s moving average, to eliminate sudden variations, then compared with a threshold *th:* an SEM is marked by the computer when the smoothed discriminant function remains above the threshold for at least 1 s .

Training of the Algorithm and Validation Procedure: Parameter values of the discriminant function (weights and threshold) were tuned by means of three EOG recordings (recordings 1-3, representing the training set). Then the tuned program was applied to the 7 EOG recordings (recordings 4-10) of the testing set, for evaluation of its performances.

Both the training process and the performance evaluation are based on 2-sec epochs. Each EOG signal was split into 2-sec epochs. A 2-sec epoch is classified as SEM epoch according to the experts or to the computer analysis, if at least the 50% of this epoch is covered by an SEM event marked by the human experts or by the algorithm, otherwise it is classified as a NON-SEM epoch. Then, as to the algorithm, all epochs were assigned to four categories: TP (true positive), TN (true negative), FP (false positive), FN (false negative). The categories TP and TN designate epochs for which the computer analysis has the same outcome as the human experts (TP for epochs classified as SEM and TN for epochs classified as NON-SEM); FP and FN correspond to mismatches (FN if SEM for the experts only, FP if NON-SEM for the experts only). Then, the following indices have been defined:

$$agreeement = \frac{TP + TN}{total\ number\ of\ epochs} \cdot 100$$

$$disagreement = \frac{FP + FN}{total\ number\ of\ epochs} \cdot 100 \quad ;$$

$$= 100 - agreement$$

$$sensitivity = \frac{TP}{TP + FN} \cdot 100 \, ;$$

$$selectivity = \frac{TP}{TP + FP} \cdot 100 \, .$$

It is worth noticing that the 2-sec epochs have been used just to perform a comparison between the performance of the algorithm and the human experts. Computation of wavelet coefficients, of the energy series, $E_j(n)$, and of the discriminate function, $f(n)$, have always been performed on the entire signal, in order to avoid edge or boundary effects.

The training of the algorithm, that is estimation of the weights and threshold of the discriminant function, has been obtained by optimizing the values of these coefficients on the training set, in order to minimize the overall disagreement between computer detection and visual detection. The overall disagreement is defined as the ratio between the total number of false epochs (throughout all the three tracings) and the total number of epochs composing the three tracings. The minimisation method adopted was the Nelder-Mead simplex algorithm. Table 1 lists the values of the weights and threshold.

Performances of the algorithm have been quantified on the testing set by computing the agreement, sensitivity and selectivity indices for each of the seven tracings.

ISSN: 1727-1983

Table 1: Values for the weights (W_3-W_{10}) in the discriminant function and for the threshold *th*, obtained with the minimization procedure of the disagreement index on the training set (records 1-3)

$W_3 = 0.51$	$W_4 = 1.04\ 1$	$W_5 = 1.07$
$W_7 = 0.1$	$W_8 = 0.15$	$W_9 = 0.05$
$W_{10} = 0.05$	$th = 0.85$	

The entire algorithm was implemented using Matlab and the Wavelet Toolbox (The Mathworks, Inc., Natick, MA).

Results

Application of the tuned program to the testing data set produces the following performance indices, averaged on all the seven EOG tracings: agreement = 80.44 ± 4.09 %; sensitivity = 67.2 ± 7.37 % and selectivity = 83.93 ± 8.65 %.

Care must be taken in interpreting the values obtained for these indices. Indeed, goal of the algorithm is to identify the intervals of SEM episodes with their start instant and duration. Therefore, in the validation procedure, a simple partial overlapping between a visually-detected SEM event and a computer-detected event was not admitted and it gave rise to errors. More specifically, it is admitted only a temporal shift not longer than 1 sec between the visual and computerized detection. Most of the occurring FP and FN epochs are merely due to a different collocation in time of the beginning and ending of SEM episodes, whereas entire FP intervals or totally missed SEM events are very rare.

An example of perfect functioning of the algorithm is presented in Fig. 3. The figure shows a 30-sec portion of left and right EOG belonging to patient 10, containing an SEM event according to the experts, and the corresponding pattern of the discriminant function.

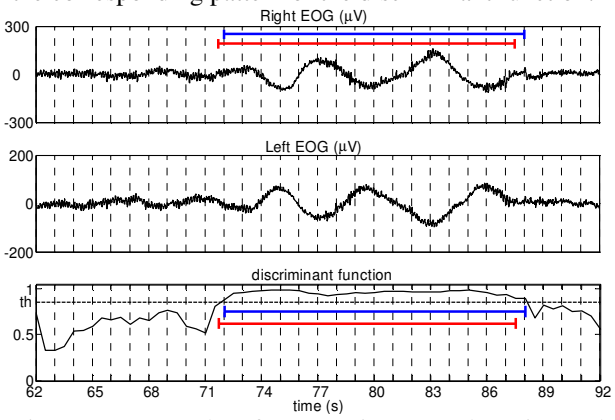

Figure 3: Example of automatic SEM detection. Blu line: computer detected SEM; red line: visually detected SEM.

In this case, a perfect agreement between visual scorers and computer detection occurs: indeed, the time shift between the visual and computer recognition is less than 1 second, both at the beginning and at the end of the event. Hence, nor false positive epochs, neither false negative epochs are generated. The computer detected SEM corresponds to the interval during which the discriminant function overcomes the threshold (indicated by the horizontal dashed line in the bottom panel).

False negative and false positive detections arise from five different type of errors. *Type I error*: a brief SEM event detected by visual scorers is recognized as a longer event by the computer. This error originates FP epochs at the boundaries of the SEM episode. *Type II error*: an SEM event lasting several seconds is recognized as a shorter event by the algorithm, thus generating FN epochs. *Type III error*: Beginning and ending of the SEM event detected by the computer are both shifted in advance or in late with respect to visual detection. In this case, both FP and FN epochs originate. *Type IV error*: a prolonged SEM event for visual scorers is split into two or more brief SEM episodes by the algorithm. In this case, FN epochs arise. *Type V error*: Two or more consecutive SEM events separated by a short lag are recognized as a single prolonged event by the computer, thus producing FP epochs.

Fig. 4 shows an example of type II and type III errors. Two SEM events are marked by the experts (red line) on recording 5. Both events are recognized by the automatic system (blue line) but with two different error types. The first SEM event is recognized only in the intermediated portion (type II error), producing false negative epochs (880÷882 s, 894÷896 s); the beginning and the end of the second SEM event are both anticipated by the computer (type III error), hence both a FP epoch (898÷900 s) and a FN epoch (904÷906 s) occur.

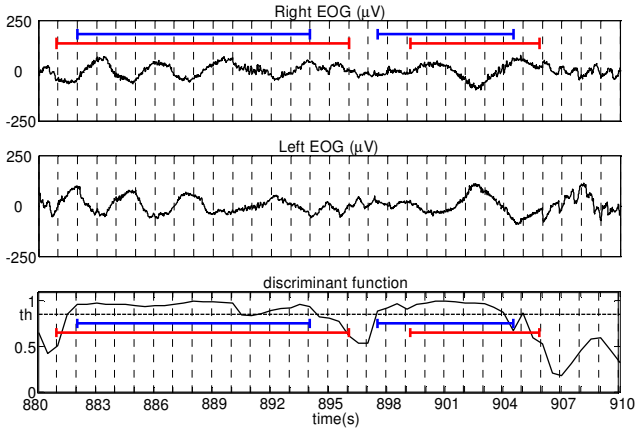

Figure 4: Example of type II and type III errors. Type II error is committed in detection of the first SEM event; while recognition of the second SEM event occurs with a type III error.

Fig. 5 shows an example of type IV error. A 30-sec segment of recording 10 is displayed, containing a train of SEM waveforms lasting several seconds, according to the visual scoring. On the contrary, the algorithm identifies three separated and shorter SEM events: the unrecognized portions generate FN epochs (1098÷1100 s, 1104÷1106 s, 1108÷1110 s, 1122÷1124 s).

ISSN: 1727-1983

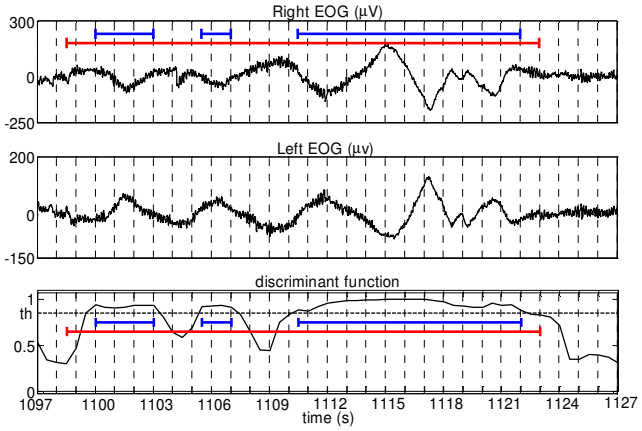

Figure 5: Example of type IV error. A single SEM event lasting several seconds is recognized as three separated SEM events by the computerized system. Hence, several epochs are classified as false negative epochs.

From the previous results, two main considerations can be drawn: 1) The discriminant function by itself is an adequate detector of SEM zones: by looking only at the temporal course of the discriminant function, even without defining any threshold value, zones where slow movements appear can be in general well identified in all the 7 examined recordings, since in these regions, *f(n)* tends to increase and to assume higher values; 2) Definition of a threshold level is necessary only to achieve a precise demarcation of SEM events detected by computer, with a start instant and duration. The threshold value, however, seems to depend on individual variability. For example, by decreasing threshold value in Fig. 5 (recording 10), a larger portion of the SEM interval marked by the experts would be recognized by the computer, thus converting false negative epochs into true positive epochs. Hence, the performances of the detection algorithm can be ameliorated by choosing a different threshold level for each subject.

Therefore, performance indices were recalculated by using a different threshold level for each recording of the testing set, while maintaining structure of the discriminant function and weighting factors unchanged. For each subject, threshold level was chosen in order to minimize the percentage of disagreement between the computer classification and the visual scoring. Percentage of agreement and sensitivity increase respectively from 80% to almost 83% and from 67% to almost 81%, with selectivity remaining around 80%. Fig. 6 shows an example of detection improvement on recording 10 by using the optimized threshold value. The top panel shows the automatic detection when the threshold level is set at 0.85 (which is the value tuned on the training set). This result is the same as in figure 5, and several FN epochs occur. When the threshold value is adjusted on the patient (th = 0.76, bottom panel), longer portion of the SEM episode are recognized and the number of FN epochs decreases.

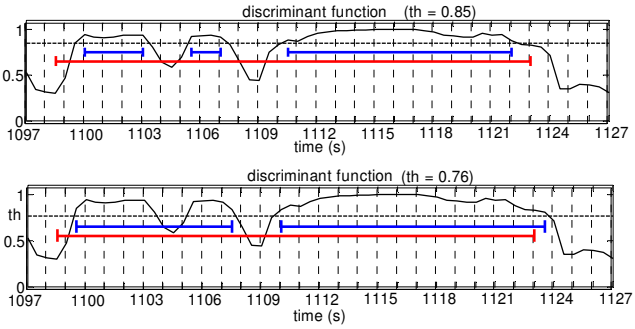

Figure 6: Example of improvement of the automatic detection by using threshold value optimized on the single patient.

Discussion

Results presented in this paper reveal the capability of the multiresolution wavelet analysis to characterize SEM activity in EOG recordings. Decision of using wavelet transform for detection of SEM events in EOG tracings was based on the fact that SEM episodes show a highly non-stationary behaviour, which makes their study by conventional methods (such as Fourier transform analysis) problematic. Indeed, slow eye movements may have different duration, ranging from tenth of seconds down to few seconds. Longer episodes may be detected quite well by using traditional Fourier techniques; on the contrary, shorter episodes cannot be identified with accuracy because of the intrinsic trade off between duration and bandwidth. Wavelet method, which is devoted to the analysis of non-stationary signals [7-9], is surely more appropriate for detection of transient phenomena that may be even very short (such as SEMs).

During SEM activity, a redistribution of signal energy takes place. Indeed, when SEMs occur, EOG shows an almost sinusoidal behaviour (0.2-0.6 Hz), with most of the energy confined in certain bands of low frequencies, corresponding to only some scales of the multiresolution analysis. Hence, we used a non-linear function of the energies associated to different resolution levels as discriminant function: such function is compared with a threshold level, above which detection of SEM events is accepted, and below which is refused. By using a unique value for the threshold, the method provides a percentage of agreement of about 80%, a 67% sensitivity and a 84% selectivity.

However, the validation procedure of the algorithm is not merely based on a simple partial overlapping between the computer-detected event and the reference SEM event, that is mismatches in start instant and duration give rise to FP and FN epochs. Such a validation criterion may be considered excessively severe if the main target of the algorithm application is just the distinction between intervals with SEM activity and intervals without SEM activity, whereas a precise identification of the beginning and duration of each SEM episode is not required. In these circumstances, the performance indices, as computed here, are misleading, and the automated system works better than it appears by looking at the values of these indices only. Indeed,

ISSN: 1727-1983

the majority of disagreement between visual and computer decision occurs at the boundaries of SEM events visually detected (errors of types I-V in section Results). These kinds of errors do not avoid recognition of intervals with SEM events against intervals without SEM events. On the contrary, entire FP episodes or totally missed SEM events are very rare. Moreover, the discriminant function never falls during regular and high amplitude SEM activity, when it assumes values equal to 1 or very close to 1 (see Fig. 3-5); by contrast, mismatched epochs correspond to low-amplitude and less definite SEM activity.

Hence, the pattern of the discriminant function is surely more useful and informative than the result of a simple "all or nothing" classification, obtained by using a deterministic threshold. Indeed, the value of the discriminant function can be interpreted as a score, ranging between 0 and 1, which quantifies the similarity of the present portion of the EOG tracing to an "ideal" or "certain" SEM event.

However, a more definite identification of the start instant and duration of each SEM event may be important in some circumstances, for example to evaluate differences in distribution and duration of SEM intervals in different subjects or in different pathologies, or to trigger an alarm system in the case the automatic method operates in real time. If definition of a threshold level is necessary, threshold value could be optimized on every single patient by using a single EOG recording previously acquired on the subject, while maintaining the same discriminant function. Indeed, we observed that in same cases, performances of the automatic detector may ameliorate significantly if the threshold is adjusted on the particular subject.

Conclusions

The present work shows that wavelet decomposition is a suitable tool for detection of SEM episodes in EOG signals. Indeed, wavelets decompose signals at different scales distinguishing high-frequency behaviour from slower activity; hence they are particularly suitable in recognizing events with energy confined only in some scales of the multiresolution framework. By using the discriminant function, which expresses the ratio between energy associated to low-frequency scales and energy associated to both low-frequency and high-frequency scales, SEM activity is characterized rigorously in terms of energetic feature. The same approach (based on energy function) can be used to detect other repeated waveforms in EOG, such as REM or spindles, characterized by energy redistribution. Hence, the present work may furnish important contributions in the field of EOG computerized analysis.

References

[1] SANTAMARIA J., and CHIAPPA K. H. (1987): 'The EEG of drowsiness in normal adults', *J. Clin. Neurophysiol.*, **4**, pp. 327-382

[2] OGILVIE R. D., MCDONAGH D. M., STONE S. N., and WILKINSON R. T. (1988): 'Eye movements and the detection of sleep onset', *Psychophysiology*, **25**, pp. 81-91

[3] HYOKI K., SHIGETA M., TSUNO N., KAWAMURO Y., and TOSHIHIKO K. (1998): 'Quantitative electro-oculography and electroencephalography as indices of alertness', *Electroenceph. Clin. Neurophysiol.*, **106**, pp. 213-219

[4] DE GENNARO L., FERRARA M., FERLAZZO F., and BERTINI M. (2000): 'Slow eye movements and EEG power spectra during wake-sleep transition', *Clin. Neurophysiol.,* **111**, pp. 2107-15

[5] OGILVIE R. D. (2001): 'The process of falling asleep', *Sleep Medicine Reviews*, **5**, pp. 247-70

[6] VÄRRI A., KEMP B., ROSA C. A., NIELSEN K. D., GADE J., PENZEL T., HASAN J., HIRVONEN K., HÄKKINEN V., KAMPHUISEN A. C., and MOURTAZAEV M. S. (1995): 'Multi-centre comparison of five eye movement detection algorithms', *J. Sleep. Res.*, **4**, pp. 119-130

[7] DAUBECHIES I. (1990): 'The wavelet transform. Time-frequency localization and signal analysis.', *IEEE Tran.s Info. Theo.*, **36**, pp. 961-1005

[8] CHUI C. K. (1992): 'An introduction to wavelets', (Academic Press, London)

[9] RAO R. M. and BOPARDIKAR A. S. (1998): 'Wavelet transforms. Introduction to theory and applications', (Addison Wesley, Reading, Massachusetts)

[10] RECHTSCHAFFEN A. and KALES A. (1968): '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

ISSN: 1727-1983