

Classification of Sleep Stages Using Multi-wavelet Time Frequency Entropy and LDA

L. Fraiwan¹; K. Lweesy¹; N. Khasawneh²; M. Fraiwan²; H. Wenz³; H. Dickhaus⁴

¹Biomedical Engineering Department, Jordan University of Science and Technology, Irbid, Jordan;

²Computer Engineering Department, Jordan University of Science and Technology, Irbid, Jordan;

³Thoracic Clinic, University of Heidelberg, Heidelberg, Germany;

⁴Medical Informatics Department, University of Heidelberg, Heidelberg, Germany

Keywords

Sleep stage scoring, multi-wavelets, time frequency entropy, linear discriminant analysis

Summary

Background: The process of automatic sleep stage scoring consists of two major parts: feature extraction and classification. Features are normally extracted from the polysomnographic recordings, mainly electroencephalograph (EEG) signals. The EEG is considered a non-stationary signal which increases the complexity of the detection of different waves in it.

Objectives: This work presents a new technique for automatic sleep stage scoring based on employing continuous wavelet transform (CWT) and linear discriminant analysis (LDA) using different mother wavelets to detect different waves embedded in the EEG signal.

Methods: The use of different mother wavelets increases the ability to detect waves in the EEG signal. The extracted features were

formed based on CWT time frequency entropy using three mother wavelets, and the classification was performed using the linear discriminant analysis. Thirty-two data sets from the MIT-BIH database were used to evaluate the performance of the proposed method.

Results: Features of a single EEG signal were extracted successfully based on the time frequency entropy using the continuous wavelet transform with three mother wavelets. The proposed method has shown to outperform the classification based on a CWT using a single mother wavelet. The accuracy was found to be 0.84, while the kappa coefficient was 0.78.

Conclusions: This work has shown that wavelet time frequency entropy provides a powerful tool for feature extraction for the non-stationary EEG signal; the accuracy of the classification procedure improved when using multiple wavelets compared to the use of single wavelet time frequency entropy.

to R&K, the state of sleep (score) is determined in a time epoch of 30 seconds based not only on the characteristics of the recorded EEG signal, but also EMG and EOG signals are necessary. During sleep, the EEG signal is characterized by the waves and events (bands) listed in ► Table 1 according to their frequency bands [4].

Sleep stages are identified by the presence and the duration of these waves (time and frequency). Stage awake is characterized by the presence of continuous alpha waves. In stage 1, low beta and theta activities exist. Stage 2 is identified by the presence of delta waves in less than 20% of the epoch and the presence of K-complexes and sleep spindles with time duration of more than 0.5 seconds. Stage 3 is scored when there are low-frequency waves with frequency less than 2 Hz, also sleep spindles and K-complexes may occur. The deepest sleep stage, 4, is characterized by the existence of delta waves in more than 50% of the epoch. The REM is similar to the awake stage, but REM has low amplitude alpha activity. ► Figure 1 shows the EEG signals for the above-mentioned sleep stages.

The sleep scoring procedure is usually done visually by experts, which is a time-consuming and tedious process as the scoring needs to be done for an entire night recording (8 hours). Furthermore, the scoring procedure is difficult to perform due to similarities between different stage properties. Automatic sleep stage scoring has been developed for more than 25 years [4–14]. It consists mainly of two parts: feature extraction from the polysomnographic recording and classification method. The feature extraction methods were based on the estimation of power of the frequency bands [4, 11, 12, 15, 16] and time frequency distribution such as wave-

Correspondence to:

Luay Fraiwan
Biomedical Engineering Department
Jordan University of Science and Technology
Irbid 22110
Jordan
E-mail: fraiwan@just.edu.jo

Methods Inf Med 2010; 49: 230–237

doi: 10.3414/ME09-01-0054

received: June 19, 2009

accepted: November 4, 2009

prepublished: January 20, 2010

1. Introduction

Human sleep can be split into two states: the rapid eye movement (REM) and the non-rapid eye movement (NREM). The NREM state is further divided into four stages (1–4). Therefore, beside the awake state, the state of human sleep is classified

into five stages [1], and is interpreted based on the characteristics of the following bio-signals (polysomnographic recordings): electroencephalograph (EEG), electrooculograph (EOG), and electromyograph (EMG). In practice, the state of sleep is determined using Rechtschaffen and Kales (R&K) recommendations [2, 3]. According

let transforms [17–19]. The classification methods were performed using, mainly, neural networks and fuzzy logic [12, 15, 18].

In this study, the continuous wavelet transform time frequency distribution was used to represent the EEG signal. Three mother wavelets with different center frequencies and different shapes were used to evaluate the entropy for the frequency bands of EEG waves that characterize each sleep stage. The classification procedure was done using the linear discriminant analysis for categorical variables.

2. Materials and Methods

2.1 Wavelet Time Frequency Feature Extraction

EEG is considered a non-stationary signal as its properties change during each sleep stage; therefore classical spectral methods are not appropriate for feature extraction, as they provide description of the frequency contents of the signal, but not the timing of the EEG signal. Timing information is

Table 1 Different EEG waves and events embedded in the EEG signal and their frequency range

Waves and events	Frequency range (Hz)
Delta	0.5–4
Theta	4–8
Alpha	8–13
Beta1	13–22
Beta2	22–35
Sleep spindles	12–14
K-complex	0.5–1.5

required for EEG signal analysis since some of the EEG waves may exist in part of the epoch but not in the entire epoch. Time frequency distribution such as wavelet transform (WT) and short time Fourier transform (STFT) are better tools for analysis and feature extraction of the non-stationary EEG signal [20–23].

Wavelet transform has been widely used in signal processing and analysis in many fields over the past two decades. It is a

powerful tool as it provides localization both in time and frequency. The wavelet transform for a signal $s(t)$ can be written as:

$$W_{a,b} = \int s(t) \psi_{a,b}(t) dt \quad (1)$$

where $\psi_{a,b}(t)$ is the dilation and transformation of the mother wavelet defined as:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad (2)$$

where $\psi(t)$ is called the mother wavelet, b is the translation parameter (time shifting parameter) that provides the time domain characteristics of the signal $s(t)$, and a is the scale parameter ($a \neq 0$) that provides a dilation and compression of the mother wavelet function $\psi(t)$, which can be used to analyze the frequency characteristics of the signal $s(t)$. The scaling parameter values were used to cover the entire frequency range of the different EEG waves based on the following equation that relates the scaling parameter a to the frequency f :

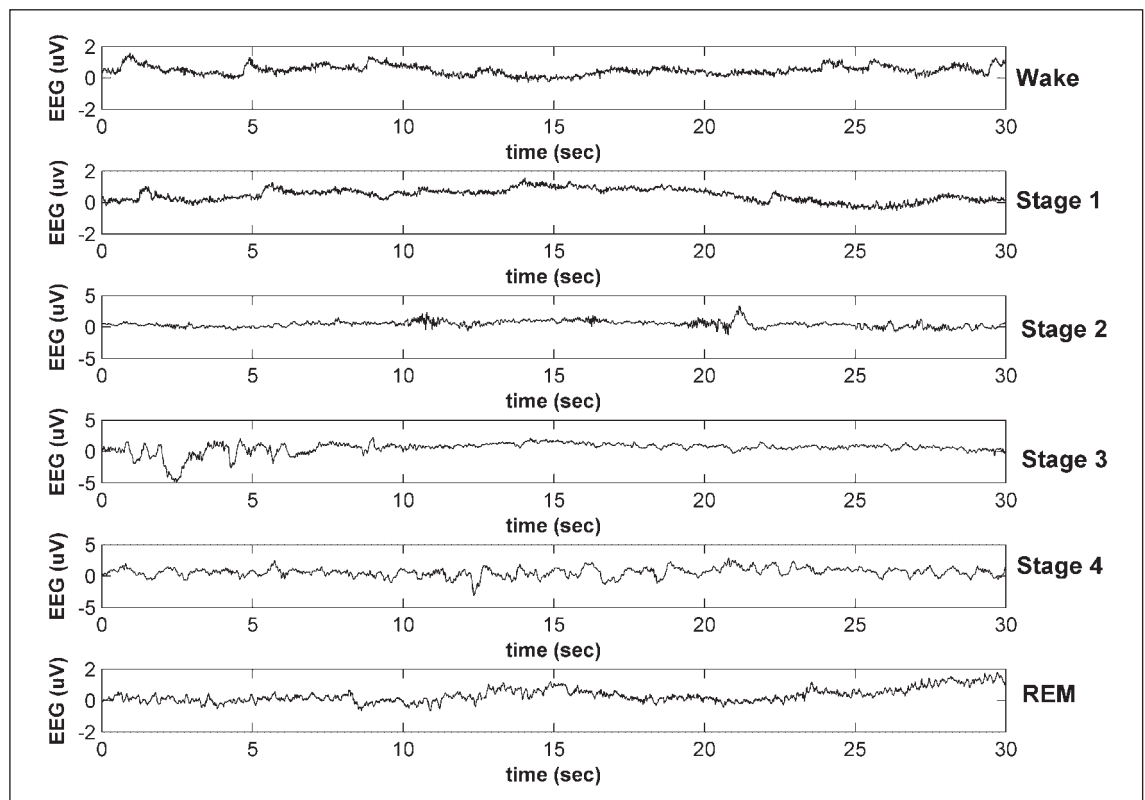


Fig. 1 EEG signals for different sleep stages

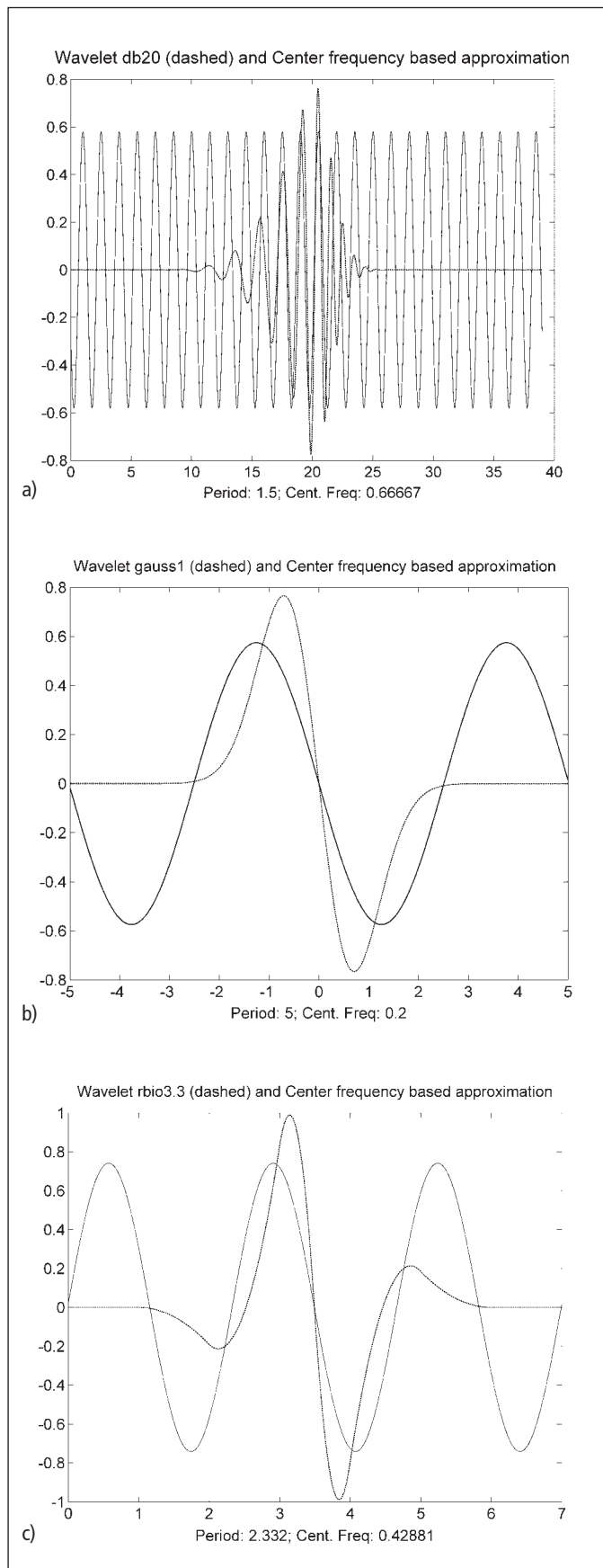


Fig. 2 Plot of the mother wavelets and the center frequency approximation for a) Daubechies (db20), b) Gauss (gauss1), and c) reverse bio-orthogonal (rbio 3.3) wavelets

$$f = \frac{f_c}{aT_s} \quad (3)$$

where f_c is the center frequency of the mother wavelet, which represents the associated sine wave that maximizes the Fourier transform of the mother wavelet modulus, and T_s is the sampling period.

Many mother wavelets can be used to find the wavelet transform of the signal $s(t)$ such as Daubechies, Coiflets, Morlet, Gaussian, Symlets, Biorthogonal, and reverse Biorthogonal [20]. The choice of the mother wavelet normally depends on the shape of the signal $s(t)$, as the WT is a measure of similarity between both the signal $s(t)$ and the mother wavelet function $\psi(t)$. As can be concluded from Equation 1, the CWT consists of calculating a resemblance index between the signal and the wavelet. If the resemblance is high, the CWT has large coefficients $W_{a,b}$, while the CWT coefficients are small if there is a slight resemblance. The EEG signal contains many waveforms and there is no single wavelet that is suitable to analyze its time frequency characteristics. Thus, three different mother wavelets were arbitrary chosen to do the analysis: reverse bio-orthogonal (rbio3.3), Daubechies (db20), and Gauss of order 1 (gauss1), with center frequencies of 0.67, 0.43, and 0.20, respectively. Moreover, the three mother wavelets were chosen to be visually different in shape to maximize the detection ability of the different EEG bands that have different shapes and frequencies. Figure 2 shows the three chosen mother wavelets and their associated center frequencies.

2.2 Time Frequency Entropy

The classification of the EEG signal depends on the localization of the different EEG waves both in time and frequency. For this reason, features were extracted using entropy estimation at different scales (frequency bands). The entropy is defined as:

$$E = -\sum_{i=1}^n p_i \log p_i \quad (4)$$

where p is the probability mass function of the wavelet coefficients $W_{a,b}$ in the band of

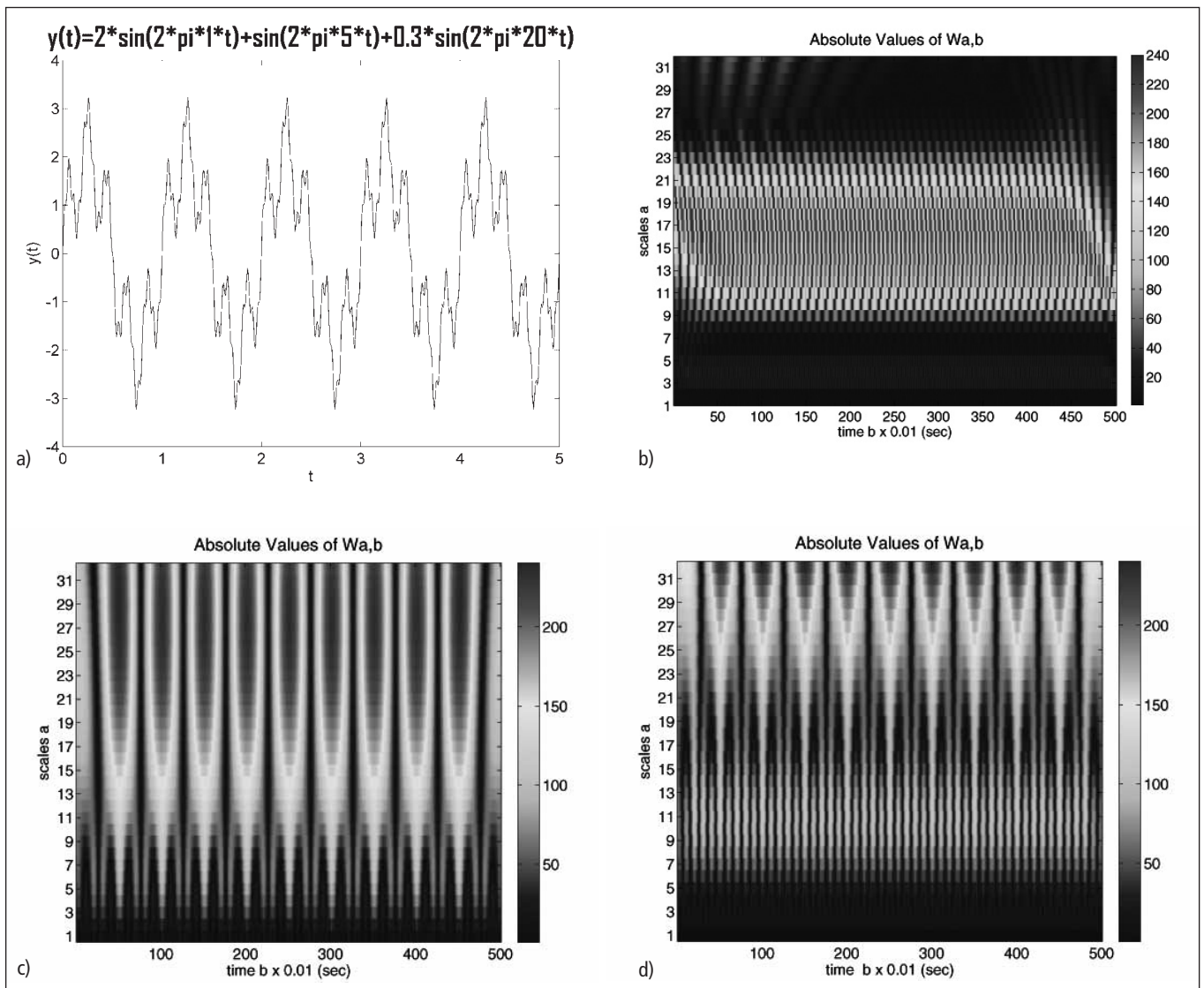


Fig. 3 a) Test signal analyzed using CWT based on three different mother wavelets; b) CWT using db20 mother wavelet; c) CWT using gauss1 mother wavelet; d) CWT using rbio3.3 mother wavelet

interest represented by their histogram with n bins.

To illustrate the idea of using multiple wavelets time frequency entropy, consider the signal shown in ►Figure 3a, which is composed of three sine waves added together, with three different frequencies: 2, 5, and 20 Hz. The signal was sampled at a sampling rate of 100 Hz and the CWT using the three mother wavelets db20, gauss1, and rbio3.3 was calculated and shown in ►Figures 3b, 3c, and 3d. As it is shown the figure, the CWT of the signal was different for each mother wavelet due to the difference in shape and center fre-

quency of the used mother wavelets. Each CWT produced higher coefficients in different scale ranges (frequency ranges), which makes each mother wavelet able to detect different frequency ranges from the others. Furthermore, the shape of the distribution is different due to the difference in shape between the mother wavelets and the signal. If the mother wavelet has the same exact shape as the signal, the CWT coefficients have maximum value at the frequency band of the signal. To implement and validate the above argument, the entropy (a measure of information) of three frequency (scale) bands, each containing a

single frequency of the test signal, was calculated and tabulated in ►Table 2. The calculated entropies were found different for each of the used mother wavelet.

2.3 Linear Discriminant Analysis Classification

Discriminant analysis has been widely used in classification problems involving categorical parameters [21, 22]. The discriminant analysis has a similar approach to the analysis of variances (ANOVA) method. It classifies the features by means of sepa-

Frequency band (Hz)	Entropy		
	rbio3.3	db20	gauss1
0–4	3.84	3.79	3.15
4–10	2.25	0.24	4.26
10–50	0.43	0.02	2.84

rating surfaces defined by discriminant functions, which may be linear or non-linear. The linear discriminant analysis (LDA) function is defined as:

$$d_k(x) = w_{k,1}x_1 + w_{k,2}x_2 + \dots + w_{k,n}x_n + w_{k,n+1} \quad (5)$$

which can be written in matrix form as:

$$d_x(X) = W_k^T X \quad (6)$$

where W is a constant coefficient vector, and X is the feature vector.

The decision rule is based on finding the minimum distance defined by:

$$X \in Stage_j \text{ if } D(X, Z) = \min_k D(X, Z_k) \quad (7)$$

where $D(\cdot)$ is the Euclidian distance of the unknown feature vector X from Z_k , and Z_k is the training set class center for the sleep stages from $Stage_k$.

2.4 Automated Sleep Stage Scoring Process

Thirty-two polysomnographic recordings from different subjects, taken from the MIT-BIH database, were used in this study with durations of 8–10 hours [23]. The polysomnographic recordings had their associated manual sleep stage score with the following vital signals: 2 EEG, horizontal EOG, oro-nasal respiration, and sub-mental EMG. In this work, only single-channel EEG signals (24 records C3-A2 and 8 records Fpz-Cz) were used for sleep stage scoring. ▶ Figure 4 shows the entire procedure used in this work. The pre-processing stage was used to smooth the EEG signal using a Savitzky-Golay filter [24] and to divide the EEG signal into epochs of 30 seconds duration. Then the CWT of the EEG signal was computed using three different mother

wavelets: reverse bio-orthogonal (rbio3.3), Daubechies (db20), and Gauss of order 1 (gauss1). The center frequencies of these three mother wavelets were 0.43, 0.67, and 0.20, respectively. The features of the EEG signal were calculated using the time frequency entropy of the signal frequency bands that correspond to the EEG bands mentioned earlier in Section 1. A feature vector of 21 parameters for each epoch in the record was formed. Features from all recordings were extracted and put together with a total of 41,778 epochs for the 32 records used in this study. Training and testing of the classifier were done using 10-fold cross-validation [25]. The proposed method, including feature extraction and linear discriminant analysis, was

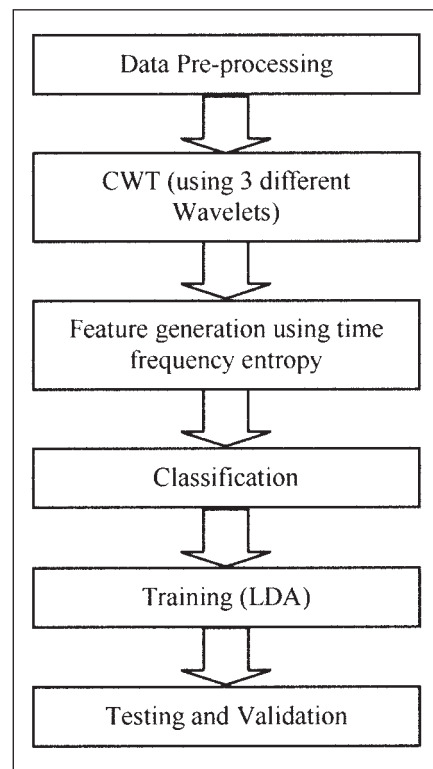


Fig. 4 The EEG automated sleep stage scoring process

Table 2

Frequency entropy for different frequency bands using the mother wavelets rbio3.3, db20, and gauss1

implemented under Matlab platform (Mathworks-USA).

2.5 Performance Measures

The performance of the algorithm was evaluated by computing the values of sensitivity (SE), specificity (SP), and accuracy (AC) of classification. The equations are as follows, respectively:

$$SE = \frac{TP}{TP + FN} \quad (8)$$

$$SP = \frac{TN}{TN + FP} \quad (9)$$

$$AC = \frac{TN + TP}{TN + TP + FP + FN} \quad (10)$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives.

Kappa statistical analysis was also used in this study to assess the agreement between the expert's score and the proposed method [26, 27]. Kappa is considered as a measure of the true agreement between the two scores by taking into account the agreement that would be expected purely by chance. The kappa coefficient is defined as:

$$kappa = \frac{P_o - P_c}{1 - P_c} \quad (11)$$

where P_o is the proportion of the observed agreements and P_c is the proportion of agreements expected by chance.

3. Results

The procedure of the proposed automatic sleep stage scoring process starts (after the pre-processing step) with the computation of the CWT using the three mother wavelets mentioned earlier. Then, the time frequency entropy is calculated for the frequency bands that correspond to the EEG waveforms and events.

The performance of the proposed automatic sleep stage scoring system was evaluated using features from a single mother

wavelet for each of the mother wavelets mentioned earlier and a combination of the three mother wavelets (multi-wavelet). The algorithm was tested against a manual score by experts available with the 32 records (subjects). ► Figure 5 shows the average accuracy and kappa coefficient using the single wavelet (db20, gauss1, rbio3.3) features and the multi-wavelet features. The multi-wavelet classification-based procedure was found to outperform the single-wavelet classification with an average accuracy of 0.84. The kappa coefficient, which measures the agreement between the expert's score and the proposed methods score, was found to be 0.78 (the maximum possible value is 0.96), which indicates a substantial agreement [25].

The classification of each sleep stage using the multi-wavelet procedure was further analyzed and performance measures were calculated for each sleep stage as shown in ► Figure 6. These statistics were calculated based on a one-versus-all classification (the analyzed stage is the positive and all other stages together are the negative). The results indicate that the awake stage has the highest performance compared to the other statistics with an accuracy of 0.93 and a kappa coefficient of 0.89, while stage 1 and stage 3 have the lowest classification performance compared to other stage with an accuracy of 0.69 and 0.62, respectively. The errors in each stage can be investigated by looking into the confusion matrix, shown in ► Table 3, which indicates the agreement between the proposed method and the experts' score. The high accuracy in classifying the awake stage can be explained by the high number of awake stages in the recordings. The low accuracy values reported for stage 1 and stage 3 can be attributed to the low number of epochs for this stage and the confusion between these stages and other similar stages due to their similar properties.

4. Discussion and Conclusions

This work presented a new approach for automated sleep stage scoring based on wavelet time frequency distribution, with

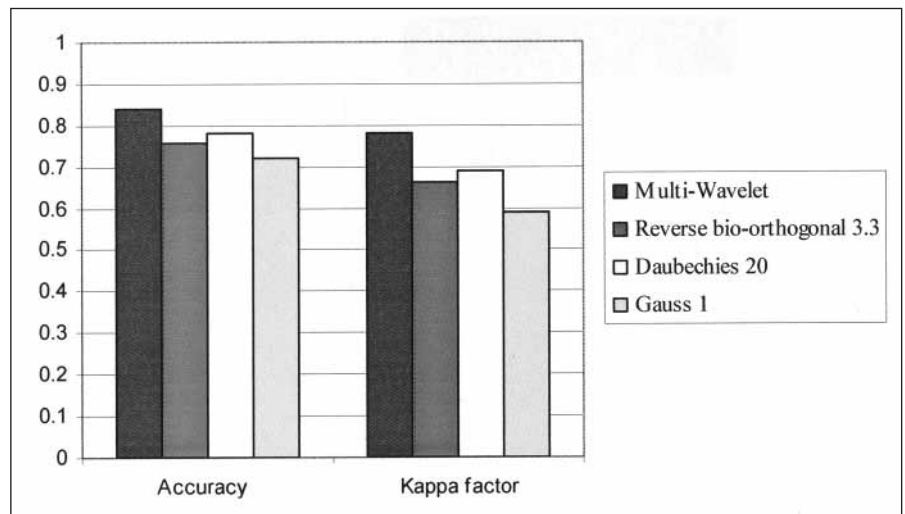


Fig. 5 The accuracy and kappa coefficient for the multi-wavelet-based classification and single wavelet-based classification using rbio3.3, db20, and gauss1

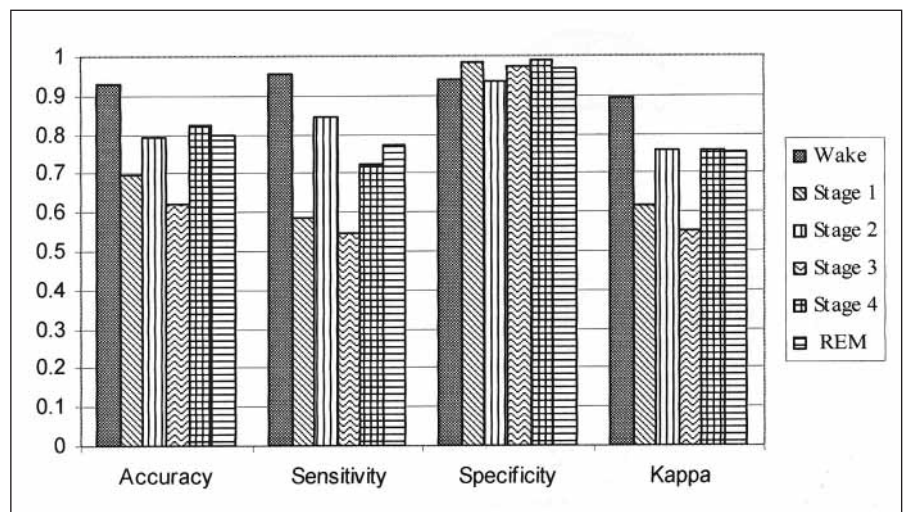


Fig. 6 Performance statistics for each sleep stage

Table 3 Scoring agreement between manual scoring and the proposed method (confusion matrix)

Multi-Wavelet Score	Experts' Score						
	Stage	Awake	Stage 1	Stage 2	Stage 3	Stage 4	REM
Awake		18117	126	456	42	9	219
Stage 1		321	1392	207	66	42	351
Stage 2		549	120	8337	429	36	414
Stage 3		174	87	774	1635	246	72
Stage 4		48	96	102	360	1575	0
REM		300	183	624	102	9	4158
Accuracy		0.93	0.69	0.79	0.62	0.82	0.80

linear discriminant analysis being used as the classifier. The implementation included the use of three different mother wavelets and a combination of these wavelets. The performance measures indicate that the approach of using a single mother wavelet and a combination of the three mother wavelets was successfully implemented in automatic sleep stage scoring with multi-wavelet approach and outperforms the single mother wavelet approach. The reason is that the multiple-wavelet method provided better ability to detect the different EEG waves in terms of frequency and shape as the three mother wavelets have different shapes and frequencies. Furthermore, the use of wavelet time frequency entropy has shown to be a powerful tool

for feature extraction for non-stationary EEG signal, where the duration and the occurrence of EEG bands within the epoch are variable.

The use of linear discriminant analysis has shown to be accurate, simple, and fast compared to the traditional neural networks method. The accuracy of classifying some stages (stages 1 and 3) was found to be less than others. This was mainly due to the low number of these stages in the training data and the similarities with other stages. Furthermore, the recorded signals were shown to an expert and there was a discrepancy between the score reported by the expert and the associated score with the data which may be another source of error. In this work, single EEG signal was the only

information used for extracting features. Further improvement can be done to this method by including the EMG and EOG in the classification between the REM and NREM stages. This can be implemented in two steps: First the data can be classified into two groups: NREM group (awake and stages 1–4) and REM group. The second step is classifying the NREM stages.

The choice of the mother wavelet is also another way of improving the performance of the suggested algorithm. This can be done by studying the shapes of the different EEG bands and constructing a special mother wavelet(s) that better match their shapes.

The performance of the proposed procedure was compared with recent works available in the literature listed in ► Table 3

Table 4 Recent works for automatic sleep stage scoring

Authors	Method	Signals	Dataset	Performance
Hanaoka et al. [28]	Feature extraction: Detection of EEG waveforms and events. Muscle discharge level (EMG). Rise angle, the fall angle, and the angle, time, and amplitude until reaching the flat. Classification: decision-tree learning,	EEG, EOG, EMG	5 subjects	80% Accuracy
Tagluk et al [29]	Feature extraction: N/A Classification: Artificial neural networks	EEG, EMG, EOG	21 subjects	75%
Wang et al. [7]	Feature extraction: Duration of alpha activity, duration of slow wave activity, amplitude of mixed frequency activity, amplitude of high frequency activity, amounts of eyes movements, amount of muscle activity. Classification: Conditional probability of the knowledge base.	EEG, EOG, EMG	4 subjects	Accuracy: Awake 89% REM:68% S1/2: 82 % S3/4: 96% S1/2: sleep stages 1 and 2 were considered as one stage and S3/4 as well.
Park et al. [30]	Feature extraction: Spectral method by calculating the power spectral densities of different EEG frequency bands, state of the EOG by calculating the spectral power at 2 frequency bands and the correlation coefficient between left and right EOG, and tone of the EMG using the average of every half second variance between interquartile ranges. Classification: Case based reasoning and hybrid rule	EEG, EOG, EMG	6 subjects	Accuracy 87.5 %
Anderer et al. [31]	Feature extraction: Background activity of the EEG signal at five different frequency bands (delta: 0.5–2 Hz; theta: 2–7 Hz; alpha: 7–12 Hz; beta: 12–20 Hz and fast beta: 20–40 Hz) and the predominant activity (density) as well as the mean amplitude (intensity), mean frequency and its variability each frequency band and for the total band (0.5–40 Hz). EMG trimmed mean of square amplitudes. Classification: Rule-based and linear discriminant analysis	EEG EOG EMG	590 subjects	Accuracy 80% kappa: 0.72
Current study	Feature extraction: Time frequency entropy at different frequency bands Classification: Linear discriminant analysis	EEG	32 subjctcs	Accuracy: 84% kappa = 0.78

[7, 28–31]. Most of these works concentrated on the classification algorithm more than the feature extraction procedure where classical spectral and time approaches were used. In comparison, this study used time frequency distribution to represent the EEG signal. Except for the study of Anderer et al. [31], the studies shown had used small number of subjects. The performance statistics obtained for the proposed method is among the best compared to the studies tabulated in ► Table 4.

Acknowledgment

This work was supported by the German Research Foundation – DFG (Deutsche Forschungsgemeinschaft).

References

- Šušmáková K. Human Sleep. *Measurement Science Review* 2004; 4–2, Section 2: 59–74.
- Rechtschaffen A, Kales A. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Public Health Service 1968, U.S. Government Printing Office, Washington DC.
- Penzel T, Hirshkowitz M, Harsh J, Chervin R, Butkov N, Kryger M, Malow B, Vitiello M, Silber M, Kushida C, Chesson A. Digital analysis and technical specifications. *Journal of Clinical Sleep Medicine* 2007; 3 (2): 109–120.
- Hae-Jeong P, Jung-Su O, Do-Un J, Kwang-Suk P. Automated sleep stage scoring using hybrid rule and case-based reasoning. *Computers and Biomedical Research* 2000; 33: 330–349.
- Kubat M, Pfurtscheller G, Flotzinger D. AI-based approach to automatic sleep classification. *Biological Cybernetics* 1994; 70 (5): 443–448.
- Shimada T, Shiina T, Saito Y. Sleep stage diagnosis system with neural network analysis. *Engineering in Medicine and Biology Society* 1998; 4 (29): 2074–2047.
- Wang B, Xingyu W, Junzhong Z, Fusae K, Masatoshi N. Automatic determination of sleep stage through bio-neurological signals contaminated with artifacts by conditional probability of a knowledge base. *Artif Life Robotics* 2008; 12: 270–275.
- Fell J, Röschke J, Mann K, Schäffner C. Discrimination of sleep stages: a comparison between spectral and nonlinear EEG measures. *Electroencephalography and Clinical Neurophysiology* 1996; 98 (5): 401–410.
- Zhouyan F, Fei G. Power spectral analysis of recovery sleep of sleep deprivation and hypnotic drug induced sleep. *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China* 2005. pp 1–4.
- Agarwal R, Gotman J. Computer-assisted sleep staging. *IEEE Transactions on Biomedical Engineering* 2001; 48 (12): 1412–1423.
- Shimada T, Shiina T. Autodetection of characteristics of sleep EEG with integration of multichannel information by neural networks and fuzzy rules. *Systems and Computers in Japan* 1999; 30 (4): 1–10.
- Pimero P, Garcia P, Arco L, Alvarezc A, Garc M, Bonal R. Sleep stage classification using fuzzy sets and machine learning techniques. *Neurocomputing* 2004; 58: 1137 – 1143.
- Jansen B, Hasman A, Lenten R. Piece-wise EEG analysis: An objective evaluation. *International Journal of Bio-Medical Computing* 1981; 12 (1): 17–27.
- Bodenstein G, Praetorius H. Feature extraction from the EEG by adaptive segmentation. *Proc IEEE* 1977; 65: 642–653.
- Robert C, Guilpin C, Limoge A. Review of neural network applications in sleep research. *Journal of Neuroscience Methods* 1998; 79: 187–193.
- Schwab K, Eiselt M, Schelenz C, Witte H. Time-variant Parametric Estimation of Transient Quadratic Phase Couplings during Electroencephalographic Burst Activity. *Methods Inf Med* 2005; 44 (3): 374–383.
- Wei-Yen H, Chou-Ching L, Min-Shaung J, Yung-Nein S. Wavelet-based fractal features with active segment selection: Application to single-trial EEG data. *Journal of Neuroscience Methods* 2007; 163: 145–160.
- Bang-hua Y, Guo-zheng Y, Ting W, Rong-guo Y. Subject-based feature extraction using fuzzy wavelet packet in brain-computer interfaces. *Signal Processing* 2007; 87: 1569–1574.
- Bang-hua Y, Guo-zheng Y, Rong-guo Y, Ting W. Feature extraction for EEG-based brain-computer interfaces by wavelet packet best basis decomposition. *Journal of Neural Engineering* 2006; 3: 251–256.
- Walnut D. *Wavelet Analysis*. Birkhauser; 2002. ISBN 0-8176-3962-4
- Krzanowski WJ. *Principles of Multivariate Analysis*. Oxford University Press; 1988.
- Sing TZE Bow. *Pattern Recognition and Image Processing*. 2nd edition. Basel, Switzerland: Marcel Dekker; 2002.
- PhysioNet 2009. Research Resource for Complex Physiologic Signals. Available online at: <http://www.physionet.org/>.
- Orfanidis J. *Introduction to Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall; 1996.
- Subasi A, Ercelebi E. Classification of EEG signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine* 2005; 78: 87–99.
- Sim J, Wright C. The Kappa statistics in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy* 2005; 85 (3): 257–268.
- Cohen J. A coefficient of agreement for nominal scales. *Education and Psychological Measurements* 1960; 20: 37–46.
- Hanaoka M, Kobayashi M, Yamazaki H. Automatic sleep stage scoring based on waveform recognition method and decision tree. *Systems and Computers in Japan* 2002; 33 (11): 2672–2683.
- Tagluk M, Sezgin N, Akin M. Estimation of sleep stages by an artificial neural network employing EEG, EMG and EOG. *Journal of Medical Systems* 2009. Published online April 2009.
- Park H, Oh J, Jeong D, Park S. Automated sleep stage scoring using hybrid rule and cased based reasoning. *Computers and Biomedical Research* 2000; 33: 330–349.
- Anderer P, Gruber G, Parapatics S, Woertz M, Miazhynskaia T, Klösch G, Saletu B, Zeitlhofer J, Barbanoj M, Danker-Hopfe H, Kemp B, Penzel T, Grözinger M, Kunz D, Rappelsberger P, Schlögl A, Dorffne G. An E-Health solution for automatic sleep classification according to Rechtschaffen and Kales: validation study of the Somnolyzer 24 × 7 utilizing the Siesta database. *Neuropsychobiology* 2005; 51: 115–133.