Automatic classification of infant sleep based on instantaneous frequencies in a single-channel EEG signal

Maja Čić[a], Joško Šoda[b], Mirjana Bonković[a]

[a]*Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, R. Boškovića 32, 21000 Split, Croatia*

[b]*Faculty of Maritime Studies, University of Split, Z. Frankopanska 38, 21000 Split, Croatia*

Abstract:

This study presents a novel approach for electroencephalogram (EEG) signal quantification in which the empirical mode decomposition method, a time-frequency method designated for nonlinear and non-stationary signals, decomposes the EEG signal into intrinsic mode functions (IMF) with corresponding frequency ranges that characterize the appropriate oscillatory modes embedded in the brain neural activity acquired using EEG.

To calculate the instantaneous frequency of IMFs, an algorithm was developed using the Generalized Zero Crossing method. From the resulting frequencies, two different novel features were generated: the median instantaneous frequencies and the number of instantaneous frequency changes during a 30 s segment for seven IMFs. The sleep stage classification for the daytime sleep of 20 healthy babies was determined using the Support Vector Machine classification algorithm. The results were evaluated using the cross-validation method to achieve an approximately 90% accuracy and with new examinee data to achieve 80% average accuracy of classification. The obtained results were higher than the human experts' agreement and were statistically significant, which positioned the method, based on the proposed features, as an efficient procedure for automatic sleep stage classification.

The uniqueness of this study arises from newly proposed features of the time-frequency domain, which bind characteristics of the sleep signals to the oscillation modes of brain activity, reflecting the

physical characteristics of sleep, and thus have the potential to highlight the congruency of twin pairs with potential implications for the genetic determination of sleep.

Keywords: EEG quantification, sleep classification, Empirical Mode Decomposition (EMD), Intrinsic Mode Function (IMF), Generalized Zero Crossing (GZC), Support Vector Machine (SVM)

1. Introduction

Successful automatic sleep stage classification based on polysomnography (PSG) signals has been a constant focus of scientific research due to the practical needs of physicians, who must perform time-exhaustive procedures of visual classification when studying the functional states of the brain.

Algorithms for automatic classification generally follow the typical steps of the pattern recognition process, including data pre-processing, characteristic feature extraction and classification on the basis of extracted features. In the case of automatic sleep classification, the features are extracted from PSG data, which has been transformed from analog to digital sequences of the samples. Many studies have been performed to identify suitable features for sleep stage classification [1], [2], [3] [4]. The most exhaustive study was published by Šušmakova and Krakovska [5], where 73 characteristics for various channels and channel combinations were analyzed.

All sleep stage-related features can be divided into four main groups [6]: time domain features, frequency domain features, time-frequency features, and nonlinear features. Features from the time and frequency domains are based on a known kernel function that determines signal decomposition and appropriate feature values, which are applicable for linear and stationary signals. A widely used method is Fourier transform (FT), which requires a time window in which the signal is assumed to be stationary. For many natural phenomena with approximation using linear systems, the Fourier

components might make mathematical sense but will make little physical sense [7]. The FT provides the same frequency resolution over the entire window function. With FT, the time evolution of the frequency patterns is lost. Thus, to overcome these limitations, time-frequency methods were developed.

A powerful time-frequency method is the wavelet transform (WT), which decomposes the electroencephalogram (EEG) signal into frequency bands that match particular sleep waves, such as delta, theta, alpha and beta bands. Usage of the wavelet transform with a specific kernel function generates different statistical features that represent the time–frequency distribution of EEG signals. These generated features include the wavelet packet coefficients [8], mean quadratic value of wavelet packet coefficients for each frequency band [9], center frequency, bandwidth and values at the center frequency [10], mean of the absolute values of the coefficients in each sub-band, standard deviation of the coefficients in each sub-band [11], and Renyi's entropy [12], [13].

The WT is suitable for the processing of non-stationary signals; however, this method is linear. Moreover, the wavelet method cannot have a fine frequency resolution if it has a fine time resolution. It provides good frequency resolution and poor time resolution at low frequencies. An additional difficulty of wavelet analysis is its non-adaptive nature. Once the basic wavelet is selected, it is used to analyze all the data.

Considering the nonlinearity of an EEG signal, there are features extracted on the basis of nonlinear approaches that suggest nonlinear dynamic and chaos theory: the largest Lyapunov entropy, approximate entropy, Hurst exponent [14] and fractal dimensions [15]. Chaos methods are suitable for the processing of nonlinear signals, but have problems with the stationarity of the signal examined.

Our hypothesis was that features that might better describe sleep stage specificities could be extracted from an EEG signal if it was decomposed into components that reflected the true physical

processing. Thus, the kernel was not known in advance and the nature of the signal was acquired by applying the appropriate decomposition mode.

For the time-frequency method, which was designated for nonlinear and non-stationary signals, the empirical mode decomposition method (EMD) was chosen for the basic decomposition procedure because it exhibits the basic idea of the proposed hypothesis [16]. It decomposes the signal into Intrinsic Mode Functions (IMF), with corresponding frequency ranges that characterize the appropriate oscillatory modes embedded in the brain neural activity acquired using EEG [17]. The EMD method has been used in the literature [10, 12, 13, 18, 19] for automatic sleep stage scoring, with varied success related to the classification accuracy achieved using features extracted from the IMF functions alone. Such features include the center frequency, bandwidth and value at the center frequency of EEG signals [10]; the energy distribution of the IMFs and ratio of the derived energies [18]; total and relative power of the different frequency bands and ratio of the power in two different spectral bands [19]; and time-frequency entropy [12], [13].

The new approach proposed in this study includes the introduction of new EEG features derived from IMFs based on instantaneous frequencies calculated using the Generalized Zero Crossing (GZC) method.

2. Materials and methods

2.1. Subjects

Twenty healthy Croatian babies aged three months were recruited from the well-child nursery in the Clinic of Gynecology and Obstetrics at the Clinical Hospital of Split, Croatia. Informed parental consent was obtained after a detailed explanation of the method and purpose of the study. The procedure was approved by the ethics committee of the Clinical Hospital Split. The study was performed according to the Helsinki Declaration.

2.2. Procedures

An electroencephalogram (EEG), two channels of electrooculogram (EOG), one channel of electromyogram (EMG), electrocardiogram (ECG) and respirogram were recorded during the babies' daytime sleep. The PSG records were performed during only one sleep cycle, in the absence of the wake phase. The EEG channels were sampled at 256 Hz. A high-pass filter of 0.3 Hz, a low-pass filter of 70 Hz, a notch filter of 50 Hz and a sensitivity of 7 μV/mm were used to record the EEG. EEG caps with 12 gold electrodes, placed according to the International '10-20' system were used, with bipolar derivations [20]. In [21], the authors reported that quiet sleep was perceptible mainly on the C3 and C4 electrodes and other states on the T3 and T4 electrodes. Thus, the measured signals were analyzed only from channels Fp1-C3, Fp2-C4, Fp1- T3 and Fp2-T4. The non-cerebral measures were used with the EEG by a sleep specialist to visually score the PSG records using standard Guilleminault and Soquet criteria of sleep classification [22] as recommended by the Pediatric Task Force of AASM (American Academy of Sleep Medicine). In some recorded EEGs, sleep spindles and K-complexes emerged, whereas in other EEGs, such features were not visible even the babies were of the same gestation age because the maturity of the brain varied individually. Thus, NREM stages could not be separated, so sleep has been scored in two stages, namely REM and NREM sleep. The artifact epochs were manually removed during the preprocessing step.

2.3. Empirical mode decomposition

The main purpose of this study was to present the efficiency of the automatic sleep stage classification based on features of only the EEG signals. The aim was to find a feature that represents the time sequence of a signal according to the physical meaning of the sleep stage to identify with the signal desired.

The EMD method is a decomposition method that separates the component of the investigated signal demonstrating the highest local instantaneous frequency into a separate intrinsic mode function. The

obtained IMFs characterize the appropriate oscillatory modes embedded in the brain neural activity acquired using EEG [17].

An IMF function has the following two properties:

1) The number of extrema and number of zero crossings are either equal or differ at most by one,
2) The mean value of two envelopes associated with the local maxima and minima is zero.

To decompose the EEG signal into IMF functions, a procedure known as sifting is performed. The sifting process includes four iterative steps:

1) To identify all local maximums and all local minimums of the input signal,
2) To create the upper envelope by interpolation between the maximums and lower envelope by interpolation of the minimums,
3) To calculate the mean value of the upper and lower envelopes,
4) To subtract the signal obtained after step three from the input signal.

The resulting signal is an IMF candidate. Next, the IMF candidate is confirmed to determine if it satisfies the two properties previously mentioned above for the function to be IMF. If it does not satisfy the properties of an IMF, then the process is repeated with the resulting signal as the input. After the IMF properties have been fulfilled, the obtained function is declared the IMF and it is subtracted from the input signal. The residue of the subtraction constitutes a new input signal, and the sifting process begins again. Because each subsequent IMF contains lower frequencies than the previous IMF functions, the EMD procedure can be continued until all IMFs are found, or stopped when a satisfied number of IMFs is obtained. In this study, the EMD procedure was stopped when an EEG signal segment had been decomposed into seven IMFs, discarding the IMFs of frequencies less than 0.3 Hz.

2.4. Generalized zero crossing

After the EMD procedure had been performed, the desired numbers of IMF functions were obtained. To create the new features for automatic sleep stage classification, it was necessary to calculate the instantaneous frequencies of each IMF function. The Generalized Zero Crossing (GZC) method, which was patented by Huang in 2006 [23], was employed to calculate the instantaneous frequencies based on extrema points. The GZC method is appropriate for nonlinear non-stationary signals, based on full and partial periods around the observed point. Zero crossings are the points in which the voltage polarity alternated. The GZC method defines all zero crossings and local extrema points as critical points. Their number and position are used for signal analysis and decision-making as follows:

For each series of signal points between two successive critical points, seven different period values are calculated:

- Time between two consecutive zero crossings of the same type (positive to negative or negative to positive) or between two consecutive maximum or two consecutive minimum values can be treated as a single period. For any point on the time axis, four different values of this type of period can be determined and marked as P1.
- The time between two consecutive zero crossings or two consecutive extrema points (maximum to minimum or vice versa) are considered half-periods. For any point on the time axis, two different values of the half-period can be calculated and marked as P2.
- Time between the extrema point and zero crossing (or vice versa) can be considered a quarter-period. For any point on the time axis, the quarter-period value can be calculated and marked as P4.

For a given series of signal points between two critical points, the instantaneous frequency can be computed as:

$$\varpi = \frac{1}{12}\left\{\frac{1}{P_4} + \left(\frac{1}{P_2^1} + \frac{1}{P_2^2}\right) + \left(\frac{1}{P_1^1} + \frac{1}{P_1^2} + \frac{1}{P_1^3} + \frac{1}{P_1^4}\right)\right\}. \quad (1)$$

It is argued in [14] that this approach provides localized and the most accurate results. In addition, this method can provide a statistical measure of frequency value scattering. Localization is rough, down to the quarter-period value, which is considered sufficiently good for most applications [23].

2.5. Relative power spectral density feature

In studies of the discrimination ability of individual PSG measures used for automatic sleep stage classification [1], [9], [24], the best single performing measure for distinguishing between sleep stages was based on the power spectral density. In studies [9] and [25], the relative power spectral density (PSD) of the EEG signal at characteristic frequency bands was selected as the feature that had the highest effect on sleep stage classification. Thus, to validate the efficiency of the sleep stage classification of our proposed features, the relative PSD was chosen as the reference point. The power of the EEG sleep signal could be obtained from either the time or frequency domain using the following equation:

$$P = \frac{1}{N^2}\sum_{i=1}^{N}|x_i|^2 = \frac{1}{M \cdot N} \cdot \sum_{k=1}^{M}\left|X_k(e^{j\omega})\right|^2 \quad (2)$$

where x_i is the samples of EEG sleep signal in time, N the number of samples of EEG sleep signal in time domain of one epoch, M the number of points for fast Fourier transform (FFT) algorithm, and $X_k(e^{j\omega})$ samples of EEG signal in the frequency domain, i.e., spectrum frequency of the EEG signal. For each epoch of the EEG signal, the relative spectral density (RSD) was calculated as $RSD_b = P_b/P_{total}$, where P_b was the energy from one of frequency bands: 0.5 – 4 Hz (delta waves), 4 – 8 Hz (theta waves), 8 – 12 Hz (alpha waves), 12 – 16 Hz (sigma waves) and 16 – 20 Hz (beta waves), which were calculated using Eq. (2). $P_{total}=\sum P_b$ for all five bands.

Based on the idea in [9], to reduce the effect of extreme values that were often observed on the physiological variables, each value of RSD in the database was transformed using the following nonlinear transformation:

$$\text{arc sin}(\text{sqrt}(x)), \text{ for } \delta \text{ and } \theta \text{ frequency bands, and} \quad (3)$$

$$\log(1+x/1-x), \text{ for } \alpha, \sigma \text{ and } \beta \text{ frequency bands.} \quad (4)$$

After this transformation, each transformed feature x was normalized using z-score normalization:

$$z = \frac{x-\mu}{SD} \quad (5)$$

where μ was the signal recording mean value of the transformed feature x, and SD was its standard deviation.

2.6. Newly proposed features

Observing the rules defined for the visual scoring of sleep stages, we concluded that the REM sleep stages were 'more active' than the NREM sleep stages; there were more changes in the EEG signal during REM than during NREM sleep stage. The criteria for visual scoring cannot isolate any classical frequency band of the EEG signal (delta, theta, alpha, beta or sigma) as a unique feature to distinguish between stages of sleep. In this study, we searched for a feature that revealed the physical description of the sleep stages. Thus, we focused on the instantaneous frequency of oscillatory modes of brain activity gained from EMD decomposition.

The algorithm based on the GZC method was developed to calculate the instantaneous frequencies. It was applied to the chosen first seven obtained IMFs as functions of higher order – lower frequencies were discarded. The median values of the instantaneous frequencies, as observed in the form of a vector with seven elements, had higher values for ‘more active’ sleep stages than for ‘more quiet’ sleep stages.

An additional new feature, which experiments confirmed had an effect on the classification of sleep stages, showed that during one 30 s epoch, the EEG signal changed its instantaneous frequency more often if it belonged to the REM sleep stage than if it was a NREM sleep stage epoch. Considering this frequency dynamic, for each IMF, the number of instantaneous frequency changes during one 30 s epoch was quantified. Thus, both of the new feature vectors were generated with seven elements.

To achieve a more accurate automatic classification, a hybrid feature was created by combining two different types of EEG features, the spectral feature and time-frequency feature. The newly defined hybrid feature vector consisted of 12 elements. Five elements were RSD values for five classical frequency bands and seven elements consisted of the number of changes in values of the instantaneous frequencies for the first seven IMFs, for a 30 s EEG signal epoch.

2.7. Stage classification using Support vector machines

After the feature vectors had been generated, the visually scored class label was attached to each epoch feature vector. Considering the nature of this study to classify sleep into two stages, the Support Vector Machines (SVM) method was chosen due to its binary classification output. The SVM is a nonlinear, nonparametric classification technique that has previously shown good results in sleep stage classification [2], [10].

The SVM is an algorithm in which the cases of training data belonging to different classes are mapped into a higher dimensional space and then separated using a hyperplane chosen to have maximal margins from boundary cases [26]. The boundary training instances are called support vectors. The hyperplane may be defined by different kernel functions. Basic kernels are linear, polynomial and Gaussian radial basis function (RBF). The RBF kernel induces boundaries by placing weighted Gaussian functions upon key training instances, defined with input parameter *gamma*. Parameter *C* defines a measure of misclassified samples, thereby determining the regularity (smoothness) and complexity of the classification boundaries. We considered the visual scoring

provided by the neurologist as the correct reference class for evaluating the proper function of the classifier. The software package LIBSVM [27] (available at http://www.csie.ntu.edu.tw/~cjlin/libsvm) was used as a classification algorithm. The RBF kernel was chosen and parameter selection was performed on a random half of the entire data set containing the same number of epochs for each class. A grid search of the best parameters, using the 10-fold cross-validation method, was completed as follows: the chosen data set was divided into ten subsets of equal size. The *gamma* parameter was seriatim taken from the interval 2^{-5} to 2^{10}. For each *gamma* value, the value of parameter *C* was taken across the same interval 2^{-5} to 2^{10}. For each (*gamma*, *C*) pair of parameters, the sequential classification accuracy was obtained for the one data subset considered unknown, while the classifier was trained on nine remaining data subsets. After ten sequences, the average classification result was calculated. From numerous (*gamma*, *C*) pairs of parameters, the pair with the best cross-validation accuracy was selected for classification of the entire data set.

3. Experimental results

The performance of the classification system with the newly proposed features was evaluated on a database of infant EEG sleep recordings containing 1414 epochs per EEG channel, all with constant length of 30 s.

The shape of the relative PSD feature is shown in **Fig. 1**. Five feature vector elements are represented as non-linear transformation of relative PSD values in five frequency bands: 1. delta (0.5–4 Hz), 2. theta (4-8 Hz), 3. alpha (8-12 Hz), 4. sigma (12-16 Hz) and 5. beta (16-32 Hz), from Fp2-T4 EEG channel, for REM and NREM sleep stages.

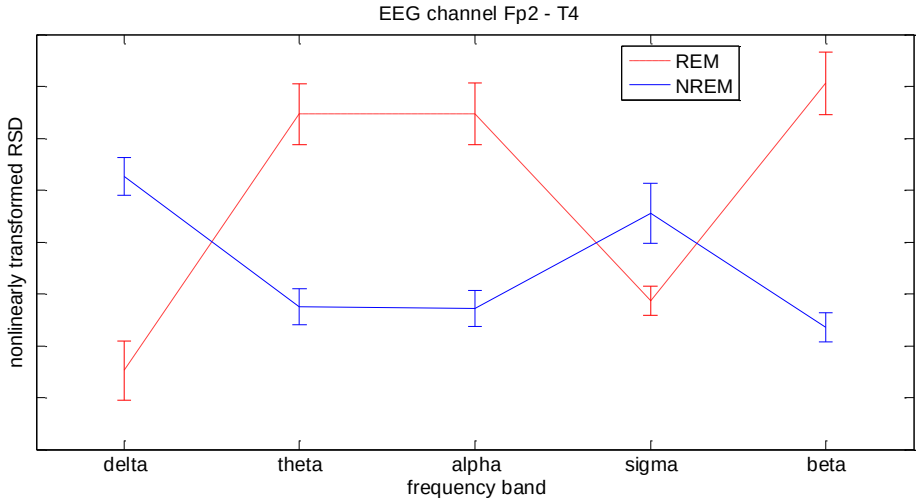

Fig. 1 Feature of relative PSD, where five feature vector elements are represented as non-linear transformation of relative PSD values in five frequency bands: 1. delta (0.5–4 Hz), 2. theta (4-8 Hz), 3. alpha (8-12 Hz), 4. sigma (12-16 Hz) and 5. beta (16-32 Hz), from the Fp2-T4 EEG channel for REM and NREM sleep stages with bars representing variances of each feature element.

The visual differences of the feature values for the REM and NREM sleep stages indicated the potential to distinguish between sleep stages using this feature. The accuracy of automatic sleep stage classification of infant EEG signals is provided in Tables 1 and 2. The obtained results were higher than the results reported in [9] for classification of adult sleep in five stages using only the RSD feature. The supposed reasons are the reduced number of sleep stages and the use of the SVM classification algorithm.

For the second feature type, the EMD process was implemented. A result of the applied technique that relates the measured EEG signal with oscillatory modes of brain activity is the decomposition of EEG signals, which was presented with the first seven IMFs. On the left side of **Fig. 2**, one sample of a 30 s EEG epoch is displayed as the original signal, and its seven IMFs are sorted below it, with an amplitude μV. On the right side of **Fig. 2** there are instantaneous frequencies over a 30 s epoch for each of seven IMFs, with a horizontal line marking the median frequency. Higher orders of IMFs, of lower frequency ranges (less than 0.3 Hz), were not considered in the calculations.

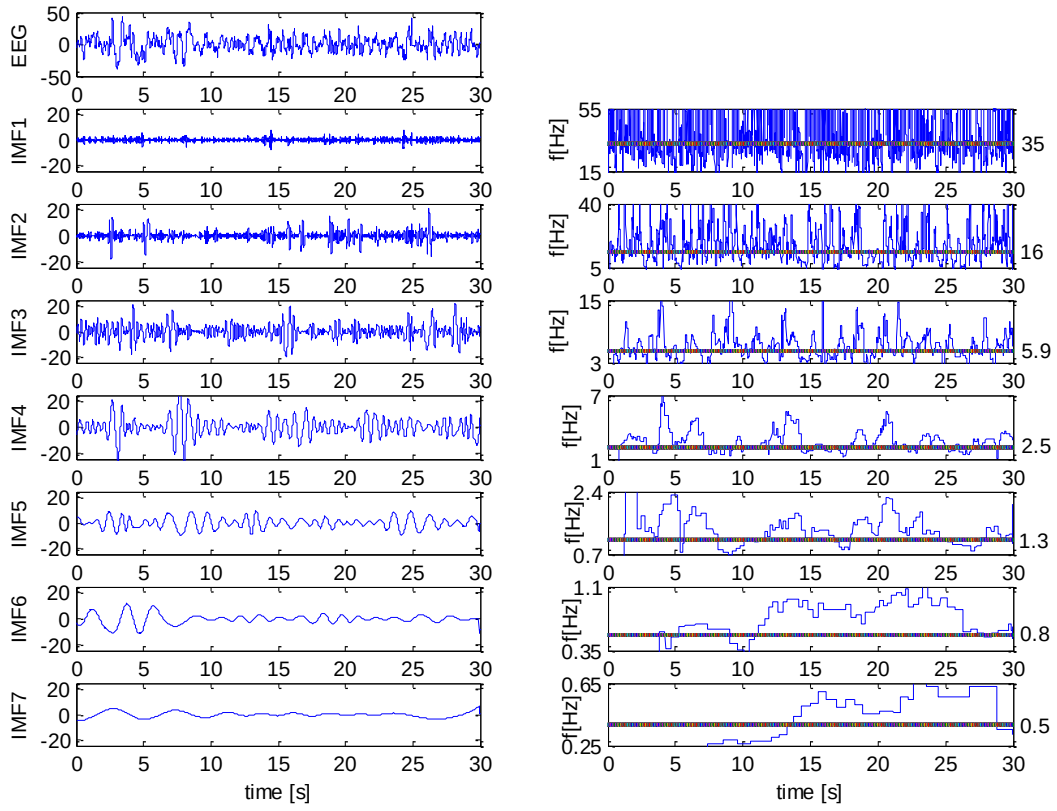

Fig. 2 Left: Analyzed EEG epoch lasting 30 s (from the Fp2-T4 EEG channel) and its seven IMFs with amplitude in μV. Right: corresponding instantaneous frequencies. Horizontal lines are median instantaneous frequencies with values in Hz printed at the end of each line.

The IMFs represent orthogonal functions, which indicates that they are independent, and thus, their frequency values in the feature vectors can be scaled using different factors. Because these values are input into the classification process, they have been scaled to the same range of values. The main advantage of scaling is to avoid elements in greater numeric ranges dominating those in smaller numeric ranges. Another advantage is to avoid numerical difficulties during the calculation. Because SVM kernel values are usually dependent on the inner products of the feature vectors, large element values might cause numerical problems. Due to the calculation of products, linear scaling of each element to the range [-1; +1] or [0, 1] is recommended [27]. The same method was used to scale training and testing data. To make the data more generalized, our procedure did not look for the

maximum value of all elements in the database, but took into consideration only the order of the values of the training feature elements.

Our first feature consisted of seven elements that were based on median frequencies from seven IMFs. The 7th IMF represented the lowest frequency oscillation band decomposed from EEG signals. It was observed throughout the available database that median frequencies of the 6th IMF were of the same order as the median frequencies of the 7th IMF. The median frequencies of the 3rd, 4th and 5th IMF were 10 times greater than those of the 7th IMF, and thus, their values were divided by scaling with a factor 10. Finally, the median frequencies of the 1st and 2nd IMF were 100 times greater than those of the 7th IMF, and thus these elements were divided by 100. The average values of the scaled median instantaneous frequencies of the 1st to 7th IMFs from the Fp2-T4 EEG channel for REM and NREM sleep stages with bars representing the variances of each feature element is shown in **Fig. 3**.

Another newly proposed feature contained seven numbers of instantaneous frequency changes during 30 s epochs for the first seven IMFs. A scaling, similar to the one previously described for median frequencies, was performed using this frequency dynamic feature vector, with the difference of two-order values in elements. Here, for the lowest frequency IMFs, the 6th and 7th, elements were divided by 100. Elements for the 3rd, 4th and 5th IMF were divided by 1000. Elements for the 1st and 2nd IMF were divided by 10000. The average values of the scaled number of instantaneous frequency changes in the 1st to 7th IMFs from Fp2-T4 EEG channel for REM and NREM sleep stages with bars representing variances of each feature element are shown in **Fig. 4**.

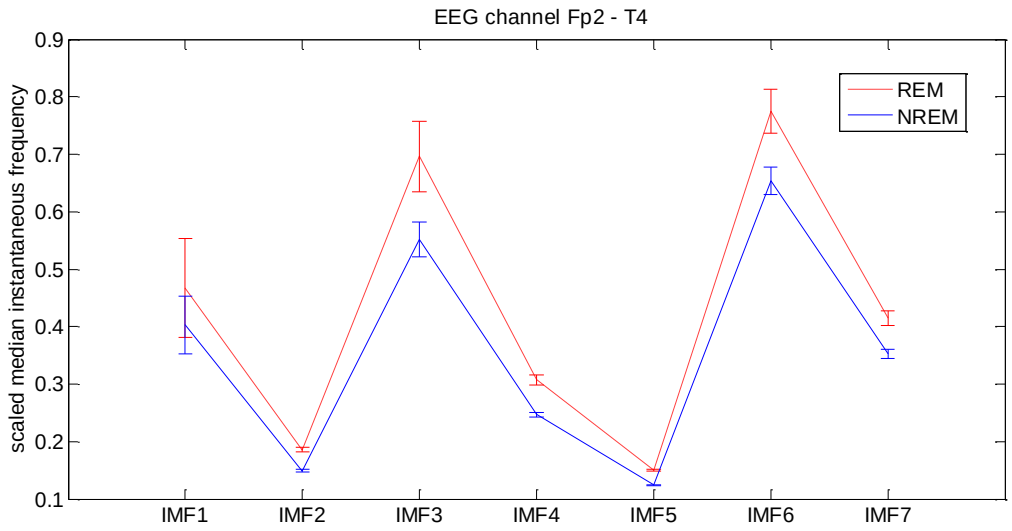

Fig. 3 New feature vector: Median instantaneous frequencies of the 1st to 7th IMFs (scaled, average values) from Fp2-T4 EEG channel for REM and NREM sleep stages with bars representing variances of each feature element.

From 1414 available epochs per EEG channel, only 476 epochs belonged to NREM sleep stage. Therefore, from the remaining 938 epochs that belonged to REM sleep stage, only 476 were chosen to avoid classification errors due to class misrepresentation. The REM sleep epochs have been chosen randomly and mixed with NREM epochs, generating a database of sleep epochs. Following the standard procedure of 10-fold cross validation, 90% of the base content had been used for training purposes, whereas the remaining 10% had been used for testing. The approach of random separation of the database into training and testing sets usually resulted in the appearance of the sample (epochs) belonging to the same person in both sets. Because it is known that every person produces a unique EEG signal that can be treated as an imprint of the person's brain [28], [29], epochs belonging to the same person are much more similar among themselves than of epochs belonging to a different person. Thus, a 10-fold cross validation procedure always generates better automatic classification results compared to those obtained using classifications of data of a new examinee.

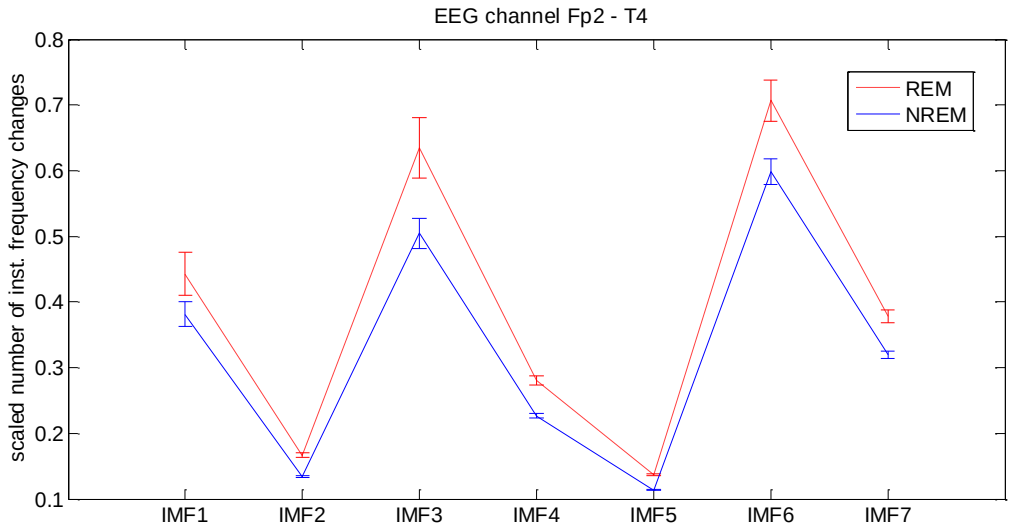

Fig. 4 New feature vector: number of instantaneous frequency changes for the 1st to 7th IMFs (scaled, average values) from the Fp2-T4 EEG channel for REM and NREM sleep stages with bars representing variances of each feature element.

For this reason, two experimental approaches were utilized and presented in this study. In the first experiment, the 10-fold cross validation procedure was applied to the complete data set. In the second experiment, the leave-one-out validation procedure was applied, such that the data of 19 babies were used for system learning, and testing was performed with the data of the remaining baby. This process was repeated 20 times.

Table 1 Accuracy of the automatic sleep stage classification from 10-fold cross validation procedure

EEG channel	feature vector			
	Relative Spectral Density in 5 freq. bands	median inst. frequencies of 7 IMFs	numbers of change in inst. freq. value of 7 IMFs	5 RSD and numbers of change in inst. freq. of 7 IMFs
Fp2–T4	83.61	86.24	87.50	87.18
Fp1–T3	85.08	86.55	88.24	90.13
Fp2–C4	84.56	85.92	86.87	89.29
Fp1–C3	84.03	90.97	87.92	90.96
mean	84.32	87.42	87.63	89.39

Table 2 Accuracy of the automatic sleep stage classification of the new examinee signal (mean value for 20 new examinee classification ± SD, in %)

EEG channel	feature vector	
	RSD	new hybrid
Fp2–T4	78.96 ± 9.31	80.45 ± 13.41
Fp1–T3	76.34 ± 12.93	79.05 ± 14.88
Fp2–C4	77.71 ± 12.74	81.43 ± 13.02
Fp1–C3	78.04 ± 9.37	81.73 ± 12.17
mean	77.76 ± 11.05	80.66 ± 13.19

For the first type of feature validation, when the 10-fold cross validation procedure was applied, the classification results for all examinees (expressed as the percentage of agreement with visually scored epochs) for all described features are presented in **Table 1**.

For the second type of feature validation, the classification was performed using data of a new examinee, whose sleep data epochs had not been used during the training phase. Thus, the data belonging to the examinee was not affected by the training process, and the final results are presented in **Table 2**. For each of the four channels, the columns presented the classification results, which were expressed as the percentage of consistency with the visually scored epochs (mean value for 20 new examinees ± SD), based on the RSD for 5 frequency bands and classification results based on the new hybrid feature consisting of RSD with dynamics of IMF frequencies, respectively. In the last row of **Table 2,** there are mean values of the results for all four EEG channels.

4. Discussion and conclusions

Using the EMD method and applying the GZC method on obtained IMF functions, we generated time-frequency features related to the physical character of the oscillatory modes of brain activity recorded as EEG signals, namely the median instantaneous frequencies and the number of instantaneous frequency changes during 30 s segments for seven IMFs. The presented results were derived solely from the characteristics of the proposed features defined for one single epoch, in the

absence of smoothing solutions based on information obtained from adjacent epochs, which would enable the avoidance of sudden and inaccurate scoring into the next stage of sleep and thus improve classification accuracy.

This approach involving instantaneous frequency is worth studying because the IMFs relate sleep features to oscillatory modes of the EEG signals and thus follow individual characteristics of a subject well. Thus, the novelty features are relevant for sleep stage distinction and could be useful in the estimation of genetic influence on EEG in sleep.

To validate the efficiency of sleep stage classification, the results of our proposed features were compared with the results of the relative PSD, which is the best single performing measure for distinguishing between sleep stages [1], [9], [24]. The results in **Table 1** showed that the SVM classification algorithm provided distinctive discrimination of the sleep stages in infant EEG sleep recordings using the relative PSD in 5 frequency bands; however, better results were achieved using newly created feature vectors based on instantaneous frequencies in a single-channel EEG signal. The new hybrid feature, a combination of the frequency (spectral) feature and the time-frequency feature, increased the classification ability of the automatic sleep stage classifier. For both methods of testing the algorithm for automatic classification of sleep signals (cross-validation and for a new examinee), results of the classification accuracy were higher for the newly proposed hybrid feature (on average: 89.39% for cross-validation method and 80.66% for a new examinee) compared to the feature of the relative spectral density of the EEG signal in traditional frequency bands (on average: 84.32% for cross-validation method and 77.95% for new examinee, as shown in **Table 1** and **Table 2**, respectively).

The difference in the results shown in **Table 1** and **Table 2** confirmed the assumption that the accuracy of automatic classification was reduced if no part of the signal of the patient involved in the testing set of the classification method had participated in the training set of the classifier. Thus, to

correctly evaluate an automated classification system, it is necessary to pay attention to the way the dataset was divided into training and testing sets used to obtain accurate results.

Similarity in the results of the four EEG channels showed that for an acceptable automatic sleep staging it was sufficient to record the sleep signal at only one EEG channel. Such a simplification of the sleep data acquisition would bring more comfort for the examinee in sleep. This is the most important advantage of the proposed features compared with complex PSG recording [9], [21], [30]. The validation for the relevance of the channels analyzed has been explained in [31], where these four channels were found among five of the highest ranked channel combinations, resulting from optimal channel selection for analysis of full-term neonate sleep.

Sleep was analyzed using recordings of only one EEG channel in several studies. In [32], Jo *et al.* used the normalized relative power of five frequency bands of a single channel EEG recording to show the intra- and inter-subject differences and similarities. An 84.6% accuracy of sleep classification was achieved, but the genetic fuzzy classifier was trained individually for each subject. In [2], Koley *et al.* selected 21 time, frequency and nonlinear features from 39 features extracted from a single EEG signal, and using the SVM classification algorithm, it reached 96.4% accuracy, but this approach imposed a large computational burden. In [1], Piryatinska *et al.* achieved a maximal accuracy of 87% using the delta power, the spectral entropy and fractional dimension of a single EEG signal. Taken together, more accurate results were achieved when different types of domain features were used together. This study differs from previous studies with its newly proposed features of the time-frequency domain, which bound characteristics of the sleep signals to the oscillation modes of brain activity, reflecting the physical characteristics of sleep and thus have the potential to highlight the congruence of twin pairs, indicating the genetic determination of sleep, which we plan to address in our future study.

The EMD method has been previously used for automatic sleep stage scoring in the literature. In [10], Tang *et al.* presented the center frequency, the bandwidth and the value at the center frequency

of EEG signals as features with 77.6% accuracy. In [18], Li *et al.* introduced the energy distribution of the IMFs and the ratio of the derived energies, with 81.7% accuracy. In [19], Liu *et al.* derived total and relative power of different frequency bands and the ratio of the power in two different spectral bands, with a mean accuracy of 87.3%. In [12] and [13], Fraiwan *et al.* used time-frequency entropy with less than 80% accuracy. These above-mentioned studies used the Hilbert Huang transformation (HHT) method; however, we employed the GZC method to calculate instantaneous frequencies. From studies that applied the EMD method as a part of the feature extraction algorithm, the introduced novelty feature of frequency dynamics reached the highest accuracy results of 87.63%.

The database in this study was of artifact-free EEG signals. These no-brain activities affect the EEG median frequency values, such that the success of the automatic sleep stage classification algorithm still must be validated with diverse EEG signals. The possibility for artifacts to affect the outcome of the classification may be reduced by generating the feature vector with values for the seven IMFs instead of a value for a single IMF.

For the newly proposed feature of frequency dynamic, no average value was calculated (except for the GZC method characteristic); thus it is expected that the effect of artifacts would be equal on the EEG signals of both phases of sleep. Taking into account that most of the artifacts caused by movement occur in REM sleep, we predict that the artifacts would only increase the value of the frequency dynamic feature vector of REM sleep, which has previously been expected to be higher compared to NREM sleep stages.

The obtained results are higher than the human experts' agreement [4], [33], which positions the proposed method as an efficient procedure for automatic sleep stage classification. However, all of the participants in this study ware babies of the same age, and thus the same ratio of feature values need to be reaffirmed at older ages. Thus, further research would be useful in that sense.

Acknowledgments

We are grateful to PhD Mirjana Vučinović, MD, the group leader at the Department of Neonatology, Clinical Hospital Split, Croatia for the visual classification of the sleep recordings.

Summary

There have been numerous studies related with sleep stage classification, which have confirmed that feature extraction plays an important role for success of the proposed methods. In addition, more than 70 different features have been tested for sleep stage classification [5]. The main purpose of this study was to present the efficiency of the automatic sleep stage classification based on features of only the electroencephalogram (EEG) signals. The aim was to find a feature that represents the time sequence of a signal according to the physical meaning of the sleep stage we wanted to identify the signal with. Our hypothesis was that features that better describe sleep stage specificities could be extracted from a signal if it was decomposed into components that reflected the true physical processing. Thus, the kernel is not known in advance and the nature of the signal is acquired by applying the appropriate decomposition mode.

As a time-frequency method, designated for nonlinear and non-stationary signals, the empirical mode decomposition method (EMD) was selected for the basic decomposition procedure because it embeds the basic idea of the proposed hypothesis [16]. It decomposes the signal into Intrinsic Mode Functions (IMF) with corresponding frequency ranges that characterize well the appropriate oscillatory modes embedded in the brain neural activity acquired by EEG [17].

The new approach presented in this study includes the introduction of new features derived from IMFs. To calculate the instantaneous frequency of IMFs, an algorithm was developed using the Generalized Zero Crossing method. With the resulting frequencies, we can demonstrated that during one 30 s epoch, the EEG signal changes its instantaneous frequency more often if it belonged to an active sleep stage than if it belonged to a quite sleep stage epoch. Considering this frequency

dynamic, we proposed the new feature, which describes how many times the instantaneous frequency changes during one 30 s epoch. An additional new feature affecting the classification of sleep stages was the median value of instantaneous frequency, which exhibited higher values in REM than in NREM sleep stages in the first seven IMFs contained in the EEG signals.

The sleep stage classification, based on the novel feature vectors of a single EEG channel, for the daytime sleep of 20 healthy babies, was performed using the Support Vector Machine (SVM) classification algorithm. The results were evaluated by applying the cross-validation method to achieve approximately 90% accuracy and with new examinee data achieving 80% average accuracy classification. The obtained results were higher than the human experts' agreement [4], [33], which positioned the method, based on the proposed features, as an efficient procedure for automatic sleep stage classification.

References

1. A. Piryatinska, G. Terdik, W. A. Woyczynski, K. A. Loparo, M. S. Scher, A. Zlotnik, Automated detection of neonate EEG sleep stages, Comput Methods Programs Biomed 95(1) (2009) 31-46.

2. B. Koley, D. Dey, An ensemble system for automatic sleep stage classification using single channel EEG signal, Comput Biol Med 42 (2012) 1186-1195.

3. S. Gudmundsson, T. P. Runarsson, S. Sigurdsson, Test-retest reliability and feature selection in physiological time series classification, Comput Methods Programs Biomed 105 (2012) 50-60.

4. S. Charbonnier, L. Zoubek, S. Lesecq, F. Chapotot, Self-evaluated automatic classifier as a decision-support tool for sleep/wake staging, Comput Biol Med 41(6) (2011) 380-389.

5. K. Šušmakova, A. Krakovska, Discrimination ability of individual measures used in sleep stages classification, Artif Intell Med 44 (3) (2008) 261-277.

6. M. Ronzhina, O. Janousek, J. Kolárová, M. Nováková, P. Honzík, I. Provazník, Sleep scoring using artificial neural networks, Sleep Med Rev 16 (2012) 251-263.

7. N. E. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. C. Yen, C. Tung, H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis, Proc. Royal Soc. A, vol. 454 (1998) 903–995.

8. R. Sinha, Artificial neural network and wavelet based automated detection of sleep spindles, REM sleep and wake states, J Med Syst 32(4) (2008) 291-299.

9. L. Zoubek, S. Charbonnier, S. Lesecq, A. Buguet, F. Chapotot, Feature selection for sleep/wake stages classification using data driven methods, Biomed Signal Process Control 2(3) (2007) 171–179.

10. W. C. Tang, S. W. Lu, C. M. Tsai, C. Y. Kao, H. H. Lee, Harmonic parameters with HHT and wavelet transform for automatic sleep stages scoring, World Acad Sci Eng Technol 33 (2007) 176-179.

11. L. Fraiwan, K. Lweesy, Automatic sleep stage scoring with wavelet packets based on single EEG recording, World Acad Sci Eng Technol 54 (2009) 485-488.

12. L. Fraiwan, K. Lweesy, N. Khasawneh, M. Fraiwan, H. Wenz, H. Dickhaus, Time frequency analysis for automated sleep stage identification in fullterm and preterm neonates, J Med Syst 35(4) (2011) 693-702.

13. L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz, H. Dickhaus, Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier, Comput Methods Programs Biomed 108(1) (2012) 10-19.

14. U. R. Acharya, O. Faust, N. Kannathal, T. Chua, S. Laxminarayan, Nonlinear analysis of EEG signals at various sleep stages, Comput Methods Programs Biomed 80(1) (2005) 37-45.

15. B. Weiss, Z. Clemens, R. Badizs, Comparison of fractal and power spectral EEG features: Effects of topography and sleep stages, Brain Res Bull 84(6) (2011) 359-375.

16. X. Li, Temporal structure of neuronal population oscillations with empirical model decomposition, Phys Lett A 356 (2006) 237–241.

17. C. Junsheng, Y. Dejie, Y. Yu, Research on the intrinsic mode function (IMF) criterion in EMD method, Mech Syst Signal Process 20 (2006) 817–824.

18. Y. Li, F. Yingle, L. Gu, T. Qinye, Sleep Stage Classification based on EEG Hilbert-Huang transform, in Proceedings of the IEEE International Conference of the Industrial Electronics and Applications (ICIEA) (2009) 3676-3681.

19. Y. Liu, L. Yan, B. Zeng, W. Wang, Automatic sleep stage scoring using Hilbert-Huang transform with BP neural network, , in Proceedings of the IEEE International Conference of the Bioinformatics and Biomedical Engineering (iCBBE) (2010) 90-93.

20. B. G. Boylan, Principles of EEG. In: Rennie JM, Hagman FC, Robertson NJ, editors. Neonatal cerebral investigation. 2th ed. Cambridge: Cambridge University press, (2008)

21. V. Gerla, K. Paul, L. Lhotska, V. Krajca, Multivariate analysis of full-term neonatal polysomnographic data, IEEE Trans Inf Technol Biomed. 13(1) (2009) 104-110.

22. C. Giuilleminault, M. Souquet, Sleep states and related pathology, Spectrum (1979) New York

23. N. E. Huang, Computing frequency by using generalized zero-crossing applied to intrinsic mode functions, United State Patent No: US 6,990,436 B1, Jan. 24, 2006.

24. M. E. Tagluk, N. Sezgin, M. Akin, Estimation of sleep stages by an artificial neural network employing EEG, EMG and EOG. J Med Syst 34(4) (2010) 717-725.

25. E. F. Behbahani, A. Bastani, EEG feature extraction using Hadamard transform for classification of sleep stages, European Journal of Scientific Research 50(2) (2011) 218-224.

26. V. Vapnik, Statistical learning theory, (1998) Wiley, New York

27. C. C. Chang, C. J. Lin, LIBSVM: a library for support vector machines, ACM Transactions on Intelligent Syst. and Technol. 2(3) (2011) Article no. 27.

28. M. Vucinovic, G. Kardum, M. Bonkovic, B. Resic, A. Ursic, J. Vukovic, Sleep EEG Composition in the First Three Months of Life in Monozygotic and Dizygotic Twins, Clin EEG Neurosci Dec. 9 (2013)

29. L. De Gennaro, C. Marzano, F. Fratello, F. Moroni, M. C. Pellicciari, F. Ferlazzo, S. Costa, A. Couyoumdjian, G. Curcio, E. Sforza, A. Malafosse, L. A. Finelli, P. Pasqualetti, M. Ferrara, M. Bertini, P. M. Rossini, The electroencephalographic fingerprint of sleep is genetically determined: a twin study, Ann Neurol. 64(4) (2008) 455-460.

30. P. A. Estévez, C. M. Held, C. A. Holzmann, C. A. Perez, J. P. Pérez, J. Heiss, M. Garrido, P. Peirano, Polysomnographic pattern recognition for automated classification of sleep-waking states in infants, Med Biol Eng Comput 40(1) (2002) 105-13.

31. A. Piryatinska, W. A. Woyczynski, M. S. Scher, K. A. Loparo, Optimal channel selection for analysis of EEG-sleep patterns of neonates, Comput Methods Programs Biomed 106 (2012) 14-26.

32. H. G. Jo, J. Y. Park, C. K. Lee, S. K. An, S. K. Yoo, Genetic fuzzy classifier for sleep stage identification, Comput Biol Med 40 (2010) 629-634.

33. P. Anderer, G. Gruber, S. Parapatics, M. Woertzc, T. Miazhynskaia, G. Klösch, *et al.* An E-Health Solution for Automatic Sleep Classification according to Rechtschaffen and Kales: Validation Study of the Somnolyzer 24x7 Utilizing the Siesta Database, Neuropsychobiology 51 (2005) 115–133.