Metrol. Meas. Syst., Vol. 24 (2017), No. 2, pp. 229–240.

PAN POLSKA AKADEMIA NAUK

METROLOGY AND MEASUREMENT SYSTEMS

Index 330930, ISSN 0860-8229

www.metrology.pg.gda.pl

M&MS

EFFECT OF FEATURE EXTRACTION ON AUTOMATIC SLEEP STAGE CLASSIFICATION BY ARTIFICIAL NEURAL NETWORK

Monika Prucnal, Adam G. Polak

Wroclaw University of Science and Technology, Faculty of Electronics, B. Prusa 53/55, 50-317 Wroclaw, Poland
(✉ monika.kaduk-prucnal@pwr.edu.pl, +48 71 320 6247, adam.polak@pwr.edu.pl)

Abstract

EEG signal-based sleep stage classification facilitates an initial diagnosis of sleep disorders. The aim of this study was to compare the efficiency of three methods for feature extraction: *power spectral density* (PSD), *discrete wavelet transform* (DWT) and *empirical mode decomposition* (EMD) in the automatic classification of sleep stages by an *artificial neural network* (ANN). 13650 30-second EEG epochs from the *PhysioNet* database, representing five sleep stages (W, N1-N3 and REM), were transformed into feature vectors using the aforementioned methods and *principal component analysis* (PCA). Three feed-forward ANNs with the same optimal structure (12 input neurons, 23 + 22 neurons in two hidden layers and 5 output neurons) were trained using three sets of features, obtained with one of the compared methods each. Calculating PSD from EEG epochs in frequency sub-bands corresponding to the brain waves (81.1% accuracy for the testing set, comparing with 74.2% for DWT and 57.6% for EMD) appeared to be the most effective feature extraction method in the analysed problem.

Keywords: sleep stage classification, EEG signal, power spectral density, discrete wavelet transform, empirical mode decomposition, artificial neural network.

1. Introduction

Sleep is one of the basic modes of human brain activity. It is a recurring state of mind and body characterised by altered consciousness and body stillness. It is well known that adults spend about 1/3 of their life in sleeping, therefore a right quality and amount of sleep have a significant impact on human mood and health.

There is a large group of sleep disorders related to the respiratory, nervous and other physiological systems, typically monitored by polysomnography [1]. Among them, hyperventilation and sleep apnea are the most prevalent ones. Interrelations between pathology, system functions and recorded biophysical signals are generally complex [2]. One of approaches to improve at-home patient care is the use of tele-monitoring [3].

Loomis at al. were the first observing that the pattern of brain potentials alters systematically in a sleeping person [4]. These cyclic shifts of brain waves are known as the sleep phases. Asernisky and Kleitman observed that normal, healthy sleep is divided into two main phases: REM (*Rapid Eye Movement*) and NREM (*Non-Rapid Eye Movement*) [5]. REM sleep is also known as paradoxical or active sleep, which generally occurs about from 90 to 120 minutes during sleep in adults [6]. The remaining time of sleep is NREM sleep and night awakenings. There has been reported recently that the sleep macrostructure is strongly associated with apnea episodes [7].

The most important signal for the classification of sleep stages is the *electroencephalogram* (EEG), one of signals recorded during *polysomnography* (PSG) [8]. It is used for distinguishing the wake and sleep phases [6]. The first EEG was recorded by Hans Berger. He was also the first who observed the brain waves and described two of them: the alpha (8–14 Hz) and beta

Article history: received on Mar. 01, 2017; accepted on Apr. 26, 2017; available online on Jun. 30, 2017; DOI: 10.1515/mms-2017-0036.

(14–30 Hz) waves [9]. The other brain waves are delta (0.5–4 Hz), theta (4–8 Hz) and gamma (30–80 Hz) ones [6]. Other brain activities, besides the waves, are artefacts like: saw-tooth waves, sleep spindles and K complexes [6].

Because of EEG signal complexity, the traditional manual sleep stage classification is time-consuming and depends on knowledge and experience of the expert. Therefore, an automatic sleep stage classification is expected to be more objective, faster and more efficient. There are two approaches to scoring the sleep stages [6]. The first one follows the standardised scoring systems introduced by Rechtschaffen and Kales [10], where the following phases are distinguished: *wakefulness* (W), *rapid eye movement* (REM), *non-rapid eye movement* (NREM) and *movement time* (MT). The NREM phase is additionally divided into light sleep (S1 and S2 stages) and deep sleep (S3 and S4 stages) [10]. Currently, a new method proposed by the American Academy of Sleep Medicine is used [1]. The main difference in stage definition is that the S1-S4 stages are replaced by N1, N2 and N3 (joined S3 and S4) ones and the MT stage is no longer distinguished [1].

An automatic sleep stage classification usually takes the following steps: dataset preparation, signal pre-processing, feature extraction and final classification [11, 12]. The dataset preparation includes splitting the EEG signal into 30-second epochs and organising subsets of epochs from the same sleep stages: W, REM, N1, N2 and N3. The pre-processing consists mainly of filtering and normalisation of the signals. The crucial step is, however, the extraction of discriminative features form the prepared epochs, simultaneously reducing the number of data for further processing. It is usually performed in the time, frequency or time-frequency/scale domains. Particularly the analyses in the frequency domain are very fruitful [13, 14]. The time domain methods include statistical analyses [12, 15–20], the Hjorth approach focused on activity, mobility and complexity [12, 16, 20, 21], and *singular spectrum analysis* (SSA) [22]. The methods in the frequency or time-frequency/scale domains describe the EEG spectral or scale properties using the *Fourier transform* (FT) [11, 23, 24], *power spectral density* (PSD) [12, 25], *short-time Fourier transform* (STFT) [12, 26], *adaptive directional time-frequency distribution* (ADTFD) [27], *Wigner-Ville* distribution (WV) [12, 16], *matching pursuit* (MP) [28], *wavelet transform* (WT) [12, 15, 16, 18, 29, 30], *empirical mode decomposition* (EMD) [19, 31–35] or *Hilbert-Huang transform* (HHT) [36]. The methods most commonly used for classification are *artificial neural networks* (ANN) [11, 12, 15, 16, 18, 20, 23–25, 30, 32, 37], *support vector machine* (SVM) [12, 16, 17, 22, 23, 38], *decision trees* (DT) [16, 19, 33], *random forest* algorithm (RF) [12, 16, 29], and *fuzzy systems* [39].

The best reported accuracy of sleep stage classification exceeded 90%, *e.g.* 97.03% [16], 96.75% [40], 95.42% [18], 93.93% [41], 93.84% [42], 93.0% [15] and 90.11% [33]. In these works, the discriminative features were extracted using, among others, such methods as *power spectral density* (PSD) [41], *discrete wavelet transform* (DWT) [15, 16], *complex wavelet transform* [18, 42] and *empirical mode decomposition* (EMD) [19, 33, 34]. Simultaneously, ANNs were used as classifiers in some of these approaches, *e.g.* [15, 16, 18, 40, 42].

It follows from the literature survey that often different classifiers of sleep stages were compared using only one feature extraction method [16, 19, 23]. There are only a few works analysing combinations of some feature extractors and classifiers [22, 40], therefore comparing effects of the most promising feature extraction methods on the automatic sleep stage classification results is desirable. For this reason, in this paper we examine three of such methods: *power spectral density* (PSD), *discrete wavelet transform* (DWT) and *empirical mode decomposition* (EMD) applied to signals from a single EEG channel, using a *feedforward multilayer neural network* (FFNN) as the automatic sleep stage classifier.

2. Materials and methods

An automatic classification of sleep phases proposed in this work assumes the following steps: preparation of a database, signal pre-processing, feature extraction and final classification. All the above processes were performed using MATLAB software (*The MathWorks, USA*).

2.1. Data

Data from the Sleep-EDF Database, available at the PhysioBank, were used. This dataset contains *polysomnographic* (PSG) recordings from 10 healthy females and 10 males (25–34 years old) without any medication, registered during two subsequent day-night periods (about 20 hours in total each). One of these 40 records had been destroyed, so we have analysed all remaining 39 files. The sleep recordings include signals from two EEG channels (Fpz-Cz and Pz-Oz) and the horizontal EEG, sampled with 100 Hz. All hypnograms were manually scored by well-trained technicians according to the Rechtschaffen and Kales manual [10] (based, however, on Fpz-Cz/Pz-Oz instead of C4-A1/C3-A2 EEGs). The signals were divided into 30-second epochs and each epoch was assigned to one of the following sleep stages: W, S1, S2, S3, S4 and REM. From these sets we selected 13650 epochs of a single Pz-Oz EEG channel (2730 epochs for each sleep stage) to prepare a maximally large and evenly distributed database. Finally, the selected epochs were organized into 5 classes: W, N1, N2, N3 (combined S3 and S4) and REM, according to the AASM scoring system [1].

2.2. Signal pre-processing

At the beginning, the linear trend was removed from each of EEG epochs to eliminate the effect of a slow drift of electrode potential, amending the low frequency spectrum of a signal [44]. Then the epochs were normalised into a range between –1 and 1, aligning the energy of signals coming from different subjects, electrodes and periods [44].

2.3. Feature extraction

The aim of feature extraction is to transform the 3000-sample epochs into much smaller, yet still containing maximally discriminative information, vectors – *i.e.* into the feature vectors (FVs). The main idea of this work is to compare the three popular data processing methods used for extraction of features from an EEG signal, which are: power spectral density, wavelet transform and empirical mode decomposition.

2.3.1. Power spectral density

Power spectral density (PSD) describes the distribution of average signal power in the frequency domain. The used Welch method is one of the most popular approaches to calculating PSD from the fast Fourier transform [45]. It averages signal spectra from succeeding, overlapping time intervals, returning the estimate called a periodogram. In this work, the analysed epochs consisting of 3000 samples were split up into 512-sample segments, overlapped by 50%, and then windowed using the Hanning window. As a result, 257 PSD values were received in a range from 0 Hz to 50 Hz with a resolution of approximately 0.19 Hz.

Admittedly, all the brain waves (delta, theta, alpha, beta and gamma) can be observed in each of the sleep stages, yet in a given stage some of them are dominant. Since the periodograms characterising diverse sleep stages are different [6, 8], they can be used to generate discriminative features. The available frequency range of PSD was divided into five bands

corresponding to the five brain waves spectra. Power density in each of the bands was integrated over three equal intervals (with midpoints at 0.59, 1.56, 2.54, 3.81, 5.37, 7.03, 8.69, 10.35, 12.11, 14.94, 18.95, 23.05, 29.20, 37.4 and 45.70 Hz), resulting in 15-element FVs characterising the analysed epochs.

2.3.2. Discrete wavelet transform

Discrete wavelet transform (DWT) enables the time-frequency (or time-scale) analysis of non-stationary signals and it is often used to study EEG [46]. This transform is similar to the Fourier one, but it applies wavelets as the basic functions instead of sinusoids. A single wavelet, discretely sampled at n, is given in a general form as:

$$\psi_{j,k}(n) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{n - k \cdot 2^j}{2^j} \right), \tag{1}$$

where ψ is a wavelet prototype (or mother-wavelet). Wavelets are localised in time (a shift parameter k) and frequency (a scale parameter j), have a limited duration, zero mean and normalised energy [46].

Using DWT, an original signal is decomposed by low-pass and high-pass filters, returning appropriate signal components together with approximation coefficients a and detailed coefficients d for a given level, respectively. Then the low-frequency component can be further processed at the next level of decomposition. Finally, the signal x is decomposed into a weighted sum of J-level series of basic wavelet functions ψ and a scaling function φ (covering all wavelets of higher levels) [15]:

$$x(n) = \sum_{j=1}^{J} \sum_{k} d_{j,k} \psi_{j,k}(n) + \sum_{k} a_{J,k} \varphi_{J,k}(n). \tag{2}$$

The sets of detailed coefficients $D_j = \{d_{j,k}\}$ and approximation coefficients $A_J = \{a_{J,k}\}$ are then commonly used to create the FVs.

Practical application of DWT requires identification of an appropriate wavelet type, which should be similar to the analysed signal [15]. EEG signals are usually decomposed using the Daubechies wavelets of order 2 or 4 [15, 16, 29, 47], so in this work the wavelet of order 3 (*db*3), particularly well resembling a local structure of this signal sampled with 100 Hz, has been applied. Additionally, it is necessary to determine the maximal level of decomposition, depending on the required frequency range. Because EEG carries important information in a range of 0.5–50 Hz, 5 levels of decomposition have been chosen, resulting in the following frequency sub-bands: D_1 (25–50 Hz), D_2 (12.5–25 Hz), D_3 (6.25–12.5 Hz), D_4 (3.125–6.25 Hz), D_5 (1.5625–3.125 Hz), and A_5 (0–1.5625).

The last step of extracting features using DWT is transforming the wavelet coefficients into numbers. Finally, the average powers P_j of D_1–D_5 and A_5 are calculated in each sub-band [15, 16, 18, 29, 30, 47], and expressed in dB (first six entries of the FV):

$$P_j = 10 \log_{10}\left(\frac{1}{N_j} \sum_{k=1}^{N_j} c_{j,k}^2 \right), \tag{3}$$

where $c_{j,k}$ denotes $d_{j,k}$ or $a_{J,k}$, and N_j is the number of coefficients in the respective set at level j, as well as standard deviations S_j of these coefficients [15, 16, 18, 47] (next six entries):

$$S_j = \sqrt{\frac{1}{N_j - 1} \sum_{k=1}^{N_j} \left(c_{j,k} - \bar{c}_j \right)^2}, \tag{4}$$

where $\bar{c_j}$ are appropriate means. Although both P_j and S_j are proportional to epoch energy, this energy is then such nonlinearly transformed, that P_j represents small differences between the features with higher resolution, improving the discriminative properties of the FV. Finally, each EEG epoch is represented by a 12-element vector of features extracted from the wavelet decomposition.

2.3.3. Empirical mode decomposition

Empirical mode decomposition (EMD) is a method used in analysing nonlinear and nonstationary signals, and it is often applied to EEGs [19, 31–35]. EMD is implemented as an efficient iterative algorithm, which decomposes a signal into a finite number of non-parametric *intrinsic mode functions* (IMFs), having two properties [31]: 1) the number of extrema is the same as the number of zero crossings (± 1); and 2) their envelopes are symmetrical in relation to the zero line. A signal $x(t)$ after decomposition is represented as:

$$x(t) = \sum_{j=1}^{p} c_j(t) + r_p(t), \tag{5}$$

where: p is the number of IMFs depending on signal complexity; $c_j(t)$ are IMFs and $r_p(t)$ is the final residue. The iterative procedure is automatically terminated when either c_j or r_p are negligible, or r_p becomes a monotonic function.

It is common to further apply the Hilbert transform to each of IMFs (the combined procedure known as the *Hilbert-Huang transform*, HHT) to compute the instantaneous frequencies and amplitudes of these signals [36]. In this work, however, the FVs of the EEG epochs are calculated directly from IMFs as the average powers of IMFs expressed in dB (according to (3)).

2.3.4. Alignment of feature vectors

Three methods used for feature extraction from EEG epochs were based on PSD, DWT and EMD. They returned, however, feature vectors of different lengths: 15, 12 and 12–22, respectively. To objectively compare efficiency of these methods in classification of sleep stages, the same classifier should be used, *i.e.* an ANN with a fixed structure, particularly with the same number of input neurons. Next, the *principal component analysis* (PCA) of the FVs obtained from PSD and EMD arranged as matrices was applied. This procedure transforms orthogonally an original feature matrix **F** into a matrix **S** of linearly uncorrelated columns called the principal components, ordered according to the decreasing variabilities of data (related to their discriminative abilities) [48]:

$$\mathbf{S} = \mathbf{FQ}, \tag{6}$$

where **Q** is a matrix constructed with eigenvectors of $\mathbf{F}^{\mathrm{T}}\mathbf{F}$. Finally, 12 first principal components of the transformed PSD and EMD features were chosen for classification purposes, after standardisation of relevant **F**s.

2.4. Classification

Artificial neural networks (ANNs) are widely applied to automatic classification of sleep stages using an EEG signal [11, 15, 16, 18, 20, 23–25, 30, 32, 37]. They are popular for their high classification efficiency and relatively simple implementation [25]. A very important task when creating an ANN is selecting a type and architecture of the network. Generally, an ANN consists of several layers of neurons: the input layer, one or more hidden layers and the output layer. The numbers of hidden layers and neurons within them influence the ANN classification

capability [25]. It is known that an ANN with two hidden layers can approximate any continuous mapping arbitrarily well. Also, most of classification problems can be solved by ANNs with only one hidden layer [25, 49].

In this paper, a *feedforward neural network* (FFNN) with the input layer consisting of 12 neurons (the size of FVs), two hidden layers with neurons characterised by a log-sigmoid transfer function and the output layer with 5 linear nodes (indicating the sleep stages: W, N1, N2, N3, and REM) was used as the classifier. The optimal number of hidden neurons depends on the numbers of input and output neurons, the volume of training data and information covered by the data. It is common to determine it empirically. Thus, the FFNN structure was selected by training FFNNs with different numbers of hidden neurons using the FVs obtained from PSD. Performance of each FFNN was assessed regarding the classification *mean squared error* (MSE) and classification accuracy (a percentage of properly identified sleep phases) [25].

The whole procedure of classification was carried out in the following steps. For the three examined feature extraction methods, the training, validation and testing sets were prepared by randomly selecting feature vectors in a proportion of 70%, 15% and 15%, respectively. In each of these sets, the classes were systematically mixed in the sequence: W, N1, N2, N3 and REM. Next, the PSD and EMD feature vectors were reduced to 12 principal components applying the PCA procedure (validation and testing matrices **F** were transformed into **S** according to (6), using standardisation parameters and matrices **Q** computed from the training sets). To find the optimal FFNN structure, the supervised training process with an increasing number of hidden neurons (until 10 consecutive MSEs were larger than the smallest one) was performed by the Levenberg-Marquardt algorithm, using the PSD features. It began with a random initialisation of neurons' biases and weights [25], and took into account the validation set. In each case this process was restarted 30 times to increase the chance of finding the global minimum. The best FFNN structure (returning the minimal MSE), found using the PSD data, was then used also for classifications based on the features obtained from DWT and EMD, repeating the training procedure with 30 random initialisations. Such an approach enabled to show the differences in discriminative potential of the three examined methods of feature extraction from EEG epochs.

3. Results

To analyse the feature extraction methods, 13650 30-second epochs of an EEG signal, suitable for this work and assigned by the experts to 5 sleep stages, were finally extracted from the *PhysioNet* database, with the same number of elements in each class (2730 epochs). These data were evenly and randomly divided into the training (9100 epochs), validation and testing sets (2275 epochs each).

PSD was used as the first feature extraction method of calculating the average signal power in 15 frequency intervals. The resulting feature vectors representing the N1, N2, N3, and REM sleep stages from the testing set (before PCA) are shown in Fig. 1.

The second method to concisely characterise the EEG epochs was DWT. According to the characteristic spectrum of EEG signal sampled with 100 Hz, 5 levels of decomposition were chosen, resulting in 6 vectors of detailed and approximation coefficients for each EEG epoch (Fig. 2), recalculated then into average powers and standard deviations.

Similarly, all EEG epochs were processed by EMD, returning from 12 to 22 intrinsic mode functions (Fig. 3), and then averaged powers and standard deviations were calculated from these components.

The next step of the study, where FFNNs with different numbers of neurons in two hidden layers were trained using the 12-element feature vectors obtained from PSD and PCA, yielded the optimal structure of this classifier, *i.e.* the FFNN with 23+22 hidden neurons (Fig. 4a),

characterised by the minimal MSE (0.0567) and the classification accuracy of 81.1% (Fig. 4b).

FFNNs with the same optimal architecture were further used to test the efficiency of sleep stage classification based on the features extracted from the EEG epochs also by DWT and EMD. The final results are summarised in Table 1.

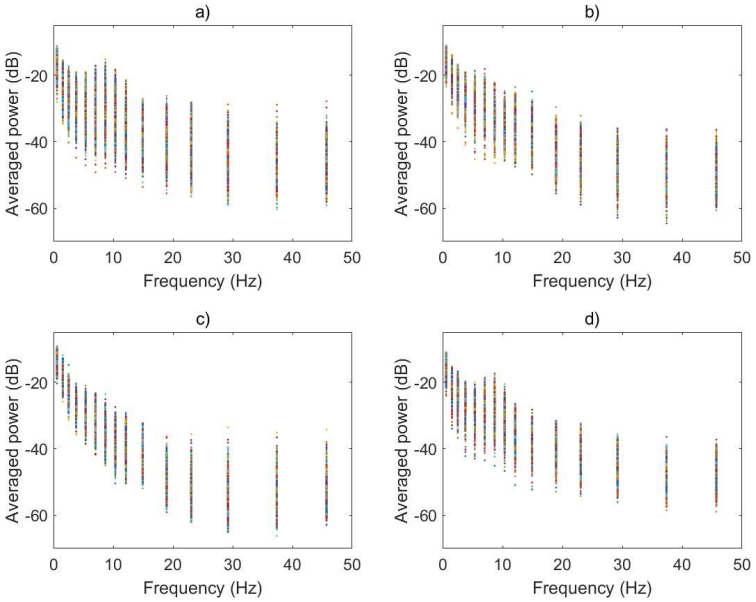

Fig. 1. Original features extracted by PSD for: N1(a); N2 (b); N3 (c) and REM sleep stages (d).

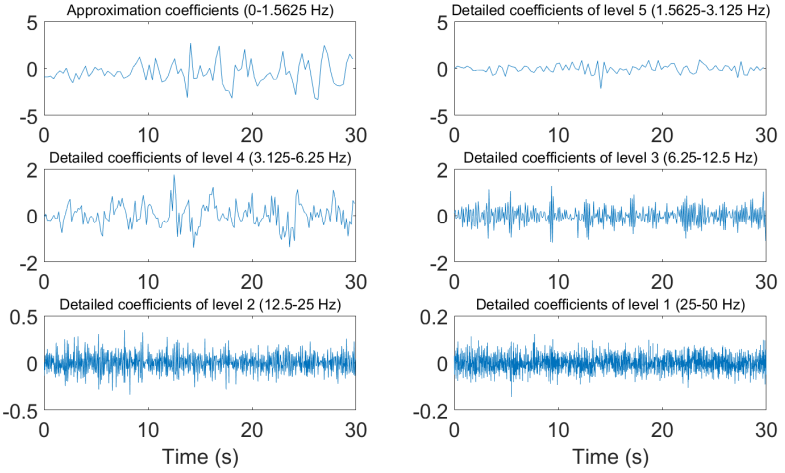

Fig. 2. Approximation and detailed coefficients from DWT (*db*3 wavelet) of an EEG epoch representing the REM sleep stage.

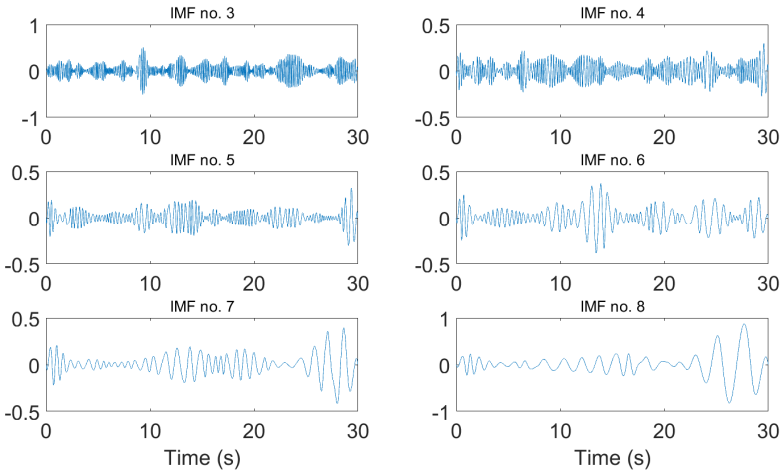

Fig. 3. Intrinsic mode functions (from 3 to 8 of 18 IMFs) derived by EMD from an EEG epoch representing the REM sleep stage.

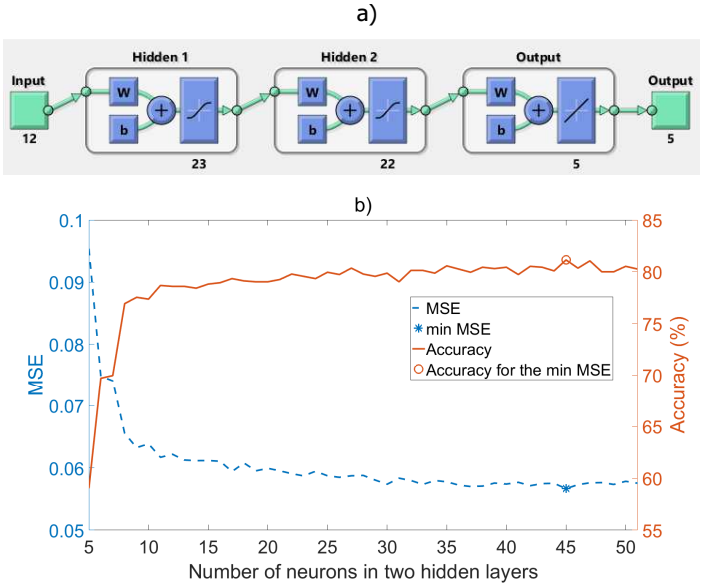

Fig. 4. The optimal structure of classifier –FFNN with 23+22 neurons in the hidden layers (a); dependencies of MSE and classification accuracy on the number of hidden neurons for the testing set (the best results from repetitions of training restarted 30 times) (b).

Table 1. Accuracy of sleep stage classification using the FFNN with 23+22 hidden neurons and the feature vectors extracted from EEG epochs by PSD, DWT and EMD.

Data	Classification accuracy (%)		
	PSD	DWT	EMD
Training set	81.2	74.2	58.7
Testing set	81.1	74.2	57.6

4. Discussion

The aim of this work was to compare the feature extraction efficiency of three methods: PSD, DWT and EMD in the automatic classification of sleep stages with the use of an ANN as the classifier.

The feature vectors extracted from PSD of N1, N2, N3 and REM sleep stages are shown in Fig. 1. A pretty wide dispersion of values for particular features can be observed within a single sleep stage and similarities between FVs of N1 and REM, as well as FVs of N2 and N3. The differences represent the inter-subject variability following the fact that the analysed epochs are obtained from two all-night EEGs of 20 subjects (10 females and 10 males). The widest dispersion of averaged powers can be observed during the N1 stage. This is possible because N1 is the first stage of sleep, accompanying the process of falling asleep. N2 and N3 sleep stages characterise slow-wave sleep in which the delta waves (0.5–4 Hz) dominate, but there are also the theta waves (4–8 Hz) during the N2 stage – the longest part of sleep. The similarity between the N1 and REM stages is caused by a large variety of frequencies within them. Nevertheless, during the N1 stage the highest amplitude is in a range of 2–7 Hz.

An example of approximation and detailed coefficients from DWT of one EEG epoch is shown in Fig. 2. In this work, the EEG signal is transformed using the Daubechies wavelets of order 3 (*db*3) at 5 levels of decomposition. In the literature, the Daubechies wavelets of order 2 [15, 29, 47] or 4 [16] were often used. Moreover, they were analysed for 4 [47] to 7 levels [15]. An additional difference is that usually FVs were prepared using far more features, such as: energy of coefficients in selected sub-bands, total energy, ratio of different energy values, or standard deviation and mean of the absolute values of coefficients in each sub-band [15, 47]. In the work [29] also 5 levels of DWT were used, but the coefficients were transformed into a more rich FV by computing their variance, skewness and kurtosis.

Figure 3 presents intrinsic mode functions (from 3 to 8 of 18 IMFs in this case) derived by EMD from one EEG epoch. In this work, the FVs are created by calculating only average powers and standard deviations from all IMFs for each epoch, and then selecting the first 12 principal components using the PCA procedure. Because EMD yields different numbers of IMFs for different epochs, preselected quantities of IMFs are used in the literature to produce larger feature vectors. For example, the features based on statistical moments (mean, variance, skewness and kurtosis) were calculated from the first 4 IMFs [19], and from the first 7 IMFs [34].

The optimal structure of FFNN for the PSD feature vectors with 12 neurons in the input layer, 23 + 22 neurons in two hidden layers, and 5 neurons in the output layer has been found in this work (Fig. 4a). In the literature, other structures of FFNN for the PSD FVs were used. For example, a network with 30 input neurons (PSD for 30 spectral bands from 0.5 Hz to 30 Hz), 6 output neurons (W, S1, S2, S3, S4 and REM) and 11 neurons in one hidden layer revealed a classification accuracy of 76.7%, and an ANN with 4 output neurons (W, S1/S2, S3/S4 and REM) and 7+7 neurons in two hidden layers demonstrated a classification accuracy of 81.5% [25]. Hsu et al. [11] proposed an FFNN with 6 input neurons, 6 neurons in one hidden layer and 5 output neurons with a classification accuracy of 81.1% as the optimal structure from the three types of neuron classifiers: Elman Recurrent ANN, FFNN and Probabilistic ANN. In another work, a structure with 15 input neurons, 32 neurons in the hidden layer and 3 output neurons (alert, drowsy and sleep) was chosen, returning accuracies over 92% [47]. That work, however, was not focused on the classification of sleep phases.

The final classification results are presented in Table 1. The best accuracy (81.1% for the testing set) is obtained for extracting the features from EEG epochs by PSD and then calculating averaged powers in 15 sub-bands related to the brain waves spectra. This result is comparable to the former works using ANNs [11, 25]. The primary difference between the used feature

extraction methods based on PSD, DWT and EMD is that the first one takes directly into account the bounds of spectra of the brain waves, and the other two do not. Achieving higher accuracy for 5 classes using only the EEG signal is very difficult, because of the similarities between the N1 and REM, and the N2 and N3 sleep stages (compare Fig. 1). This is due to the fact that PSD presents information about the average spectral nature of signal in 30-second epochs. The classification results obtained with DWT could be probably better if the FVs were extended by either such features like energy of coefficients [15, 47] or statistical features: mean, variance, skewness and kurtosis [29], or by using DWT with the Daubechies wavelets of order other than 3 [16]. Especially the approach combining the decomposition coefficients related to the specific brain wave bounds seems to be very promising [15]. A classification accuracy with FVs from EMD is surprisingly low (57.6%). Moreover, this approach is computationally less efficient due to an iterative procedure of finding the intrinsic mode functions. Probably better results can be achieved if the Hilbert transform is applied to IMFs (the Hilbert-Huang Transform [31]) and then specific frequency sub-bands are selected to produce features [36], or if the statistical features of the IMFs are also taken into account [19, 34]. The best reported results of sleep stage classification from an EEG signal (*e.g.* [16, 18, 41, 42]) used mixed signals or methods of feature extraction and larger FVs, but such approaches are beyond the scope of this paper.

5. Conclusion

Three methods of feature extraction from EEG epochs: power spectral density, discrete wavelet transform and empirical mode decomposition, were tested for the purpose of sleep stage classification by artificial neural networks with the same structure. The best result, characterised by a classification accuracy of 81.1%, was yielded when the features were prepared using averaged powers from the frequency sub-bands of PSD and a feedforward neural network with 12 input neurons, 23 + 22 hidden neurons and 5 output neurons was applied as the classifier. Such an outcome shows that the efficiency of PSD is better than DWT and EMD in this specific classification problem. Also, it stresses the importance of using the frequency sub-bands characteristic for the brain waves to detect the sleep stages from EEG.

Although this preliminary study has unambiguously shown that PSD returns the best results in comparison with other tested methods of feature extraction, it is worth to continue this study focusing on some selected issues. First of all, a possibility of extracting the characteristic frequency sub-bands from DWT (by combining selected approximation and detailed coefficients) and EMD (by using the Hilbert transform) should be tested. Also, computing larger sets of features within these approaches to EEG processing, besides average powers and standard deviations used in this work, can be tried. And finally, other classification methods, *e.g.* the support vector machine or decision trees, may suit better dealing with this particular problem. Analysing all the above possibilities should lead to obtaining even better classification accuracy than that achieved in this study.

References

[1] Berry, R.B., Brooks, R., Gamaldo, C.E., Harding, S.M., Lloyd, R.M., Marcus, C.L., Vaughn, B.V. (2015). *The American Academy of Sleep Medicine Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, Version 2.2*. Darien, Illinois: American Academy of Sleep Medicine.

[2] Jabłoński, I. (2013). Modern methods for description of complex couplings in neurophysiology of respiration. *IEEE Sensors J.*, 13, 3182–3192.

[3] Polak, A.G., Głomb, G., Guszkowski, T., Jabłoński, I., Kasprzak, B., Pękala, J., Stępień, A.F., Świerczyński, Z., Mroczka, J. (2009). Development of a telemedical system for monitoring patients with chronic respiratory

diseases. In: O. Dössel and W.C. Schlegel (Eds): *World Congress on Medical Physics and Biomedical Engineering, IFMBE Proceedings*, Springer, 25/V, 51–54.

[4] Loomis, A.L., Harvey, E.N., Hobart, G. (1937). Cerebral states during sleep, as studied by human brain potentials. *J. Exp. Psychol.*, 21(2), 127–144.

[5] Kleitman, N., Asernisky, E. (1953). Regularly occurring periods of eye motility, and concomitant phenomena, during sleep. *Science*, 118(3062), 273–274.

[6] Chokroverty, S., Thomas, R., Bhatt, M. (2014). *Atlas of Sleep Medicine*. Philadelphia: Elsevier Saunders.

[7] Hwang, S.H., Lee, Y.J., Jeong, D.U., Park, K.S. (2016). Apnea-hypopnea index estimation using quantitative analysis of sleep macrostructure. *Physiol. Meas.*, 37, 554–563.

[8] Attarian, H.P., Undevia, N.S. (2012). *Atlas of Electroencephalography in Sleep Medicine*. New York: Springer.

[9] Berger, H. (1929). Uber das Elektrnkephalogramm des Menschen. *Arch Psychiat Nernenkr*, 87, 527–570.

[10] Rechtschaffen, A., Kales, A. (1968). *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. Los Angeles: Brain Information Service.

[11] Hsu, Y.L., Yang, Y.T., Wang, J.S., Hsu, Ch.Y. (2013). Automatic sleep stage recurrent neural classifier using energy features of EEG signals. *Neurocomputing*. 104, 105–114.

[12] Boostani, R., Karimzadeh, F., Nami, M. (2017). A comparative rewiev on sleep stage classification methods in patients and healthy individuals. *Comput. Methods Programs Biomed.,* 140, 77–91.

[13] Jabłoński, I., Mroczka, J. (2009). Frequency-domain identification of the respiratory system during airflow interruption. *Measurement*, 42, 390–398.

[14] Jabłoński, I., Polak A.G., Mroczka, J. (2011). A preliminary study on the accuracy of respiratory input measurement using the interrupter technique. *Comput. Methods Programs Biomed*., 101, 115–125.

[15] Ebrahimi, F., Mikaeili, M., Estrada, E., Nazeran, H. (2008). Automatic sleep stage classification based on EEG signals by using neural networks and wavelet packet coefficients. *Conf. Proc. IEEE Eng. Med. Biol. Soc.,* 1151–1154.

[16] Sen, B., Peker, M., Cavusoglu, A., Celebi, F. (2014). A Comparative Study on Classification of Sleep Stage Based on EEG Signals Using Feature Selection and Classification Algorithms. *J. Med. Syst.*, 38, 18.

[17] Diykh, M., Li, Y. (2016). Complex networks approach for EEG signal sleep stages classification. *Expert Syst. Appl.*, 63, 241–248.

[18] Peker, M. (2016). A new approach for automatic sleep scoring: Combining Taguchi based complex-valued neural network and complex wavelet transform. *Comput. Methods Programs Biomed*., 129, 203–216.

[19] Hassan, A.R., Bhuiyan, M.I.H. (2016). Computer-aided sleep staging using Complete Ensemble Empirical Mode Decomposition with Adaptive Noise and bootstrap aggregating. *Biomed. Signal Process. Control.*, 24, 1–10.

[20] Yucelbas, S., Ozsen, S., Yucelbas, C., Tezel, G., Kuccukturk, S., Yosunkaya, S. (2016). Effect of EEG Time Domain Features on the Classification of Sleep Stages. *Indian J. Sci. Technol.*, 9, 1–8.

[21] Oh, S.H., Lee, Y.R., Kim, H.N. (2014). A Novel EEG Feature Extraction Method Using Hjorth Parameter. *J. Electron. Electr. Eng.*, 2, 106–110.

[22] Mohammadi, S.M., Kouchaki, S., Ghavami, M., Sanei, S. (2016). Improving time–frequency domain sleep EEG classification via singular spectrum analysis. *J. Neurosci. Methods*, 273, 96–106.

[23] Lee, J., Yoo, S. (2013). Electroencephalography Analysis Using Neural Network and Support Vector Machine during Sleep. *Engineering*, 5, 88–92.

[24] Dong, H., Supratak, A., Pan, W., Wu, Ch., Matthews, P., Guo, Y. (2016). Mixed neural network approach for temporal sleep stage classification. *arXiv preprint arXiv:1610.06421*.

[25] Ronzhina, M., Janousek, O., Kolarova, J., Novakova, M., Honzik, P., Provaznik, I. (2012). Sleep scoring using artificial neural networks. *Sleep Med. Rev.*, 16, 251–263.

[26] Sanders, T.H., McCurry, M., Clements, M.A. (2014). Sleep Stage Classification with Cross Frequency Coupling. *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, *2014*, 4579–82.

[27] Khan, N.A., Ali, S. (2016). Classification of EEG signal using adaptive time-frequency distributions. *Metrol. Meas. Syst.*, 2(23), 251–260.

[28] Malinowska, U., Durka, P., Blinowska, K.J., Szelenberger, W., Wakarow, A. (2006). Micro- and Macrostructure of Sleep EEG. *IEEE Eng. Med. Biol. Mag.*, 25, 26–31.

[29] Silveira, T.L.T., Kozakevicius, A.J., Rodrigues, C.R. (2017). Single-channel EEG sleep stage classification based on a streamlined set of statistical features in wavelet domain. *Med. Biol. Eng. Comput.*, 55, 343–352.

[30] Tsinalis, O., Matthews, P.M., Guo, Y., Zafeiriou, S. (2016). Automatic Sleep Stage Scoring with Single-Channel EEG Using Convolutional Neural Networks. *Ann. Biomed. Eng.*, 44, 1587–1597.

[31] Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.Ch., Tung, Ch.Ch., Liu, H.H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A*, 454, 903–995.

[32] Djemili, R., Bourouba, H., Korba, M.C.A. (2016). Application of empirical mode decomposition and artificial neural network for the classification of normal and epileptic EEG signals. *Biocybern. Biomed. Eng.*, 36, 285–291.

[33] Hassan, A.R., Bhuiyan, M.I.H. (2016). Automatic sleep scoring using statistical features in the EMD domain and ensemble methods. *Biocybern. Biomed. Eng.*, 36, 248–255.

[34] Hassan, A.R., Bhuiyan, M.I.H. (2017). Automated identification of sleep states from EEG signals by means of ensemble empirical mode decomposition and random under sampling boosting. *Comput. Methods Programs Biomed.*, 140, 201–210.

[35] Bajaj, V., Pachori, R.B. (2012). Classification of seizure and nonseizure EEG signals using empirical mode decomposition. *IEEE Trans. Inf. Technol. Biomed.*, 16, 1135–1142.

[36] Liu, Y., Yan, L., Zeng, B., Wang, W. (2010). Automatic Sleep Stage Scoring using Hilbert-Huang Transform with BP Neural Network. *Proceedings of ICBEE*, 1–4.

[37] Becq, G., Charbonnier, S., Chapotot, F., Buguet, a., Bourdon, L., Baconnier, P. (2005). Comparison Between Five Classifiers for Automatic Scoring of Human Sleep Recordings. *Stud. Comput. Intell.*, 4, 113–127.

[38] Wu, H.T., Talmon, R., Lo, Y.L. (2015). Assess Sleep Stage by Modern Signal Processing Techniques. *IEEE Trans. Biomed. Eng.*, 62, 1159–1168.

[39] Pinero, P., Garcia, P., Arco, L., Alvarez, A., Garcia, M.M., Bonal, R. (2004). Sleep stage classification using fuzzy sets and machine learning techniques. *Neurocomputing,* 58–60, 1137–1143.

[40] Yulita, I.N., Fanany, M.I., Arymurthy, A.M. (2016). Sequence-based sleep stage classification using conditional neural fields. *arXiv preprint arXiv:1610.01935.*

[41] Güneş, S., Polat, K., Yosunkaya, S., Dursun, M. (2009). A novel data pre-processing method on automatic determining of sleep stages: *K-means* clustering based feature weighting. *Complex Syst. Appl. ICCSA*, 112–117.

[42] Peker, M. (2016). An efficient sleep scoring system based on EEG signal using complex-valued machine learning algorithms. *Neurocomputing*, 207, 165–177.

[43] Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.Ch., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 101, 215–220.

[44] Varsavsky, A., Mareels, I., Cook, M. (2011). *Epileptic Seizures and the EEG: Measurement, Models, Detection and Prediction.* Boca Raton: CRC Press Taylor & Francis Group.

[45] Welch, P.D. (1967). The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms. *IEEE Trans. Audio Electroacoust.*, 15, 70–73.

[46] Adeli, H., Zhou, Z., Dadmehr, N. (2003). Analysis of EEG records in an epileptic patient using wavelet transform. *J. Neurosci. Methods*, 123, 69–87.

[47] Subasi, A. (2005). Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients. *Expert Syst. Appl.*, 28, 701–711.

[48] Abdi, H., Williams, L.J. (2010). Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.*, 2, 433–459.

[49] Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.*, 2, 303–314.