

An Approach of Sleep Stage Classification Based on Time–frequency Analysis and Random Forest on a Single Channel

Siyuan Bi, Qingmin Liao and Zongqing Lu

Department of Electronic Engineering, Tsinghua University, Shenzhen, China

Abstract— An approach of sleep stage classification based on Time–Frequency analysis and Random Forest (TFRF) on a single channel is presented in the paper. Before classifying sleep stages, representative features are extracted by feature extraction method, such as FFT in the frequency domain of EEG signal and Hilbert transform in time domain, to reduce dimensionality and channel numbers. Other parameters are also used, e.g. Hjorth parameters. Then a Random Forest is trained by classifying the sleep stages. The TFRF is evaluated by means of Dreams database provided by University of MONS – TCTS Laboratory. The new standard of the American Academy of Sleep Medicine about the sleep stage classification is obeyed. 16 objects are considered as the training samples, while the other 4 objects are regarded as the test samples. A best result of 90.4% sensitivity is got with the signals from only one channel. Its application prospect is very extensive.

Keywords—sleep stage classification; EEG; time-frequency analysis; Random Forest

I. INTRODUCTION

As living standards improve, people's health consciousness constantly improve, especially, awareness of sleep health. Thus, sleep health evaluation and sleep problem diagnosis has become a concerned subject. Apart from the assessment of sleep health being performed in hospitals and clinics, people are expecting to evaluate their sleep health at home and by themselves, which are much more convenient compared to the whole process conducted in hospitals. So, In addition to making classification standards, developing wearable, structure compact and low energy consumption devices, more important is to be able to provide accurate technical data to determine the sleep quality good or not. Therefore, the method and algorithm to be discussed in the paper take both the effectiveness and efficiency into consideration.

To examine one's sleeping condition, the analysis of polysomnogram (PSG) is the most common way. PSG is the test result of polysomnography and it can be used to diagnose or assess many types of sleep disorders including idiopathic hypersomnia, narcolepsy, periodic limb movement disorder, REM behavior disorder, parasomnias, and sleep apnea. PSG will typically record a minimum of twelve channels, including a minimum of three channels for the electroencephalographic (EEG) records, two for electrooculographic (EOG) records, and electromyographic (EMG) records. Among them the most important signal is EEG signal, which can represent most features of one's whole night's sleeping state. The traditional method to analyze one's EEG is artificial analysis, which takes

plenty of time. So the approaches of classifying sleep stage were automatically developed fast in recent years.

The sleep stages were classified by many methods proposed by previous researchers and better results were reported in scientific publications. Features was extracted by the Genetic Algorithm (GA) proposed by Sheng-fu Liang and sleep stages were classified by Fuzzy Vector Quantization (FVQ) to model HMM [3]. A set of significant features was selected by minimum-redundancy maximum-relevance (mRMR) proposed by Sirvan Khalighi and importance weighted kernel logistic regression (IWKLR) as a classifier for the whole five stages [4]. The whole five stages were classified by a state Machine-controlled Decision Trees proposed by Syed Anas Imtiaz [5].

In order to evaluate the healthiness of one's sleep in an efficient and effective way, TFRF in the paper mainly discuss the method to differentiate three stages, namely wake, N2, and N3. On the other hand, to reduce the power consumption and minimize size of devices, only one channel of EEG records is selected for analysis and sleep stage classification. Features of EEG can be extracted by time-frequency analysis method of TFRF and Random Forest is considered as a classification method.

II. METHODOLOGY

A. Basis

Up till now, two standards of sleep stage classification have been proposed, namely, Rechtschaffen and Kales standard (R&K) [1] and the American Academy of Sleep Medicine (AASM) standard [2]. The R&K standard was proposed in 1968, which divides the whole night's sleep into six stages: wake, none rapid eye movement (NREM) 1 to 4 and rapid eye movement (REM). The AASM standard was proposed in 2007. It shares some similarities with the R&K standard, except for the divisions of NREM state. AASM standard divides the NREM state into 3 stages: NREM 1 to 3. The AASM standard is adopted for TFRF discussed throughout the paper. According to the AASM standard, N2 and N3 states account for a much larger percentage of one's whole night's sleep compared to the other states. Besides, to identify the wake stage is critical for the classification of sleep stage, as it is used to decide whether one has fall into sleep or not.

B. Database

The Dreams database [6] provided by University of MONS – TCTS Laboratory (Stéphanie Devuyt, Thierry Dutoit) is used

in the experiments. The database contains 20 healthy subjects' whole-night PSG recordings, 16 of them are female while 4 of them are male. The age of participants ranged between 20 and 65. All the PSG datasets were collected in a sleep laboratory of a Belgium hospital by a digital 32-channel polygraph. The EEG data was sampled at 200Hz. The standard European Data Format (EDF) was used for storing. The label of the each stage was stored as integer 1~5, representing N3, N2, N1, R and Wake stage, respectively.

C. Algorithms

The EEG signal can be modelled as a main signal consisted of several sub-bands with different frequency adding a Gauss white noise. The purpose is to filter the electrical noise and other bio-electricity signals from the EEG signal, and extract features which could represent difference between each sub-bands. The frequency that could not represent the feature of the sub-band is regarded as noise as well. The sleep stage signal can be represented as (1):

$$S(t) = \sum s_i(t) + n(t) \quad (1)$$

$S(t)$ represents the initial signal, $s_i(t)$ represents the sub-bands of the signal, and $n(t)$ represents the electrical noise and other bio-electricity signals.

EEG signal is a traditional bioelectricity signal, in which there are interfered signals, such as other bioelectricity signals and systemnoise. Thus, analyzing the signal only in time domain or frequency domain cannot obtain satisfactory results. The signal can be analyzed by the TFRF algorithm below in both time domain and frequency domain in the paper.

A pre-processing on the datasets is done as follows. As the TFRF algorithm discussed mainly classify one's whole night sleep into four stages, N2, N3, Wake, and a DS stage which is represented as N1 and R. As N1 and R are not differentiated throughout, in this regard, labels equal to 3 or 4 representing DS are denoted by 3. Then the data collected from the CZ-A3 channel are gotten, and the pre-processing is completed.

First, the signal is transformed to frequency domain by Fast Fourier Transform (FFT), and its features are simultaneously extracted [7]. Then, the sub-bands over 30Hz can be ignored, most of which only carries little effective information and is probably noise or other useless sub-bands. Finally, the frequency spectra can be transformed to magnitude spectra again. The process on extracting features from frequency domain is done.

Second, features are extracted from the time domain. According to the AASM, sleep stage classification depends on both frequency and characteristic waveform. Nevertheless, when determine an epoch to be one of the stages, doctors do not depend on frequency but waveform. So it is necessary to extract features from time domain. Hilbert transform is a typical time domain method in signal processing of sleep [8]. It can calculate the envelope and the instantaneous frequency of the signal, which are both necessary for the sleep stage classification. As the maximum frequency of effective EEG signals is 30Hz, when the training data has a sampling rate of 200Hz, the signal is processed by Hilbert transform and its envelop is calculated by

the TFRF algorithm. The transform result is gotten with 200Hz sampling rate.

Third, the Hjorth parameters of each epochs as feature were calculated [9]. Hjorth parameters include three: Activity, Mobility, and Complexity, defined as follow (2)-(5). In which $s(t)$ represents an epoch of the signal.

$$\text{var}(s(t)) = \int s^2(t) dt \quad (2)$$

$$\text{Activity} = \text{var}(s(t)) \quad (3)$$

$$\text{Mobility} = \sqrt{\frac{\text{var}\left(\frac{d(s(t))}{dt}\right)}{\text{var}(s(t))}} \quad (4)$$

$$\text{Complexity} = \frac{\text{Mobility}\left(\frac{d(s(t))}{dt}\right)}{\text{Mobility}(s(t))} \quad (5)$$

Hjorth Parameters were statistical parameters which can extract features from EEG signals. They can also process tactile signals, which is beyond the discussion. Hjorth parameters as features of time domain are used below.

It is generally recognized that 0~4Hz sub-bands, 4~8Hz sub-bands, 8~13Hz sub-bands and 13~30Hz sub-bands are special sub-bands which are the main frequency in sleep stage classification. However, because of the individual difference, a part of the frequency spectrum which represents the sub-bands above cannot distinguish different sleep stages well. Also, modeling EEG signals as a Gauss distribution faces the same problem. Thus, using traditional linear classifier such as Support Vector Machine directly on features above cannot receive better performance on sleep stage classification. Therefore, sleep stages is classified by Random Forest method [10].

Random Forest was improved to solve the over-fitting of the traditional decision tree [11]. It contains many decision trees, which are mutually independent. Although each tree has a weak performance on classification, by voting Random Forest can maximize the effect of correct classification of each trees. Because of this, Random Forest is widely adopted and has superiority in several types of situations [12]. According to previous research, Random Forest performs well on unbalanced datasets and on non-stationary process. It can also calculate the importance of each features automatically. All these superiorities of Random Forest mentioned above fit the problem discussed very well.

Flowchart of TFRF algorithm is shown as figure 1.

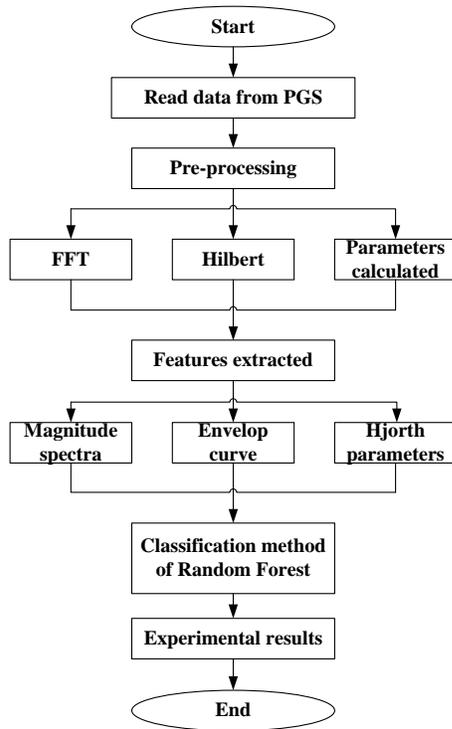


FIGURE I FLOWCHART OF TFRF

As the Dreams database provides an artificial sleep stage score for every 5 seconds for each subjects, to ensure the score is comparable, the data recorded firstly divided for each train subjects to several epochs, each of them has a length of 5 seconds. Then every 6 epochs is grouped into a sub-group and vote the classification results of each epoch inside the sub-group (30 seconds per sub-group). Each epoch is performed by FFT and the magnitude spectra as a feature are gotten [13]. The first 150 points of the magnitude spectrum are used which represent 0~30Hz. As discussed above, a part of the 150 points may miss out some critical features, thus all of the 150 points are selected as features. The Hilbert transform is performed on each epoch and the features generated are with 1000 dimensions [14]. The Hjorth parameters of each epochs are calculated afterwards. The last step of TFRF is completed by the Random Forest.

The performance of TFRF algorithm was evaluated by the four parameters below (6)-(9) [15]. They are defined as follow:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (6)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (7)$$

$$\text{Selectivity} = \frac{TP}{TP+FP} \quad (8)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+TN+FP} \quad (9)$$

Among which TP represents number of true positive, TN represents number of true negative, FP represents number of false positive, and FN represents number of false negative.

III. EXPERIMENTAL RESULTS

The experimental results based on TFRF are revealed in this section. TFRF has been tested on the DREAMS Subjects Database which contains twenty overnight PSG signals collected from 20 healthy subjects. Since the TFRF algorithm is based on the single channel, TFRF is tested by data from only one CZ-A3 channel.

16 overnight PSG signals is selected as the training-set, including 13 female subjects and 3 male subjects, and the rest 4 ones as the testing-set. As the N3, N2 and Wake stages is distinguished by TFRF, a pre-process is conducted on the database to combine label 3 and 4 into label 3. After that, TFRF is tested based on the data from the database. The performance of TFRF on the N3, N2, DS and the Wake stages is shown as Table I, which is the corresponding sensitivity and specificity. The sleep stages not belonging to the three stages are combined and represented as stage “DS”.

To compare the performance of the TFRF algorithm mentioned with those in existing research on sleep stages, some experimental results from other papers are listed here.

The Sensitivity and Specificity of four algorithm is shown as table I, including those of the algorithm proposed (TFRF), the algorithm proposed by Sheng-Fu Liang (HMM), the algorithm proposed by Sirvan Khalighi (IWKLR), and the algorithm proposed by Syed Anas Intiaz (DT). The Selectivity and Accuracy of four algorithm is shown as Table II. Some algorithm mentioned above do not reports the whole four parameters. The results not disclosed in their paper are omitted from the comparison.

TABLE I. COMPARISON OF SENSITIVITY AND SPECIFICITY

	Stage	TFRF	IWKLR	DT	HMM
Sensitivity/ %	W	81.6	74.8	72.1	80.6
	N3	89.6	88.3	79.6	90.9
	N2	90.4	73.9	88.2	80.4
	N1	69.8*	41.1	22.0	12.4
	R		77.6	84.1	97.3
Specificity/ %	W	98.6	96.1	96.9	\
	N3	97.2	94.2	97.7	\
	N2	84.6	89.0	85.3	\
	N1	95.4*	94.8	97.5	\
	R		93.2	92.1	\

* As the N3 and R staged are combined into DS, the percentage presented in this table represents the classification results for DS.

TABLE II. COMPARISON OF SELECTIVITY AND ACCURACY

	Stage	TFRF	IWKLR	DT	HMM
Selectivity/%	W	93.3	\	72.1	\
	N3	87.8	\	85.2	\
	N2	78.6	\	84.0	\
	N1	82.7*	\	45.3	\
	R		\	73.3	\
Accuracy/%	W	95.4	\	94.5	\
	N3	95.8	\	95.1	\
	N2	86.8	\	86.6	\
	N1	89.2*	\	91.1	\
	R		\	90.4	\

*. As the N3 and R staged are combined into DS, the percentage presented in this table represents the classification results for DS.

For the sake of comparison, Sensitivity, Specificity and selectivity about three methods is shown as figure 2~figure 4.

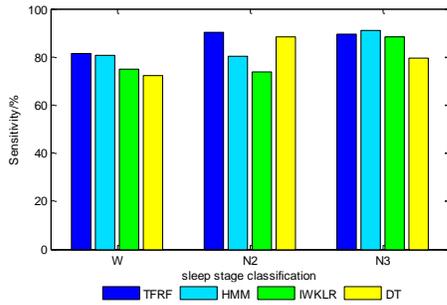


FIGURE II. COMPARISON OF SENSITIVITY ABOUT THREE METHODS

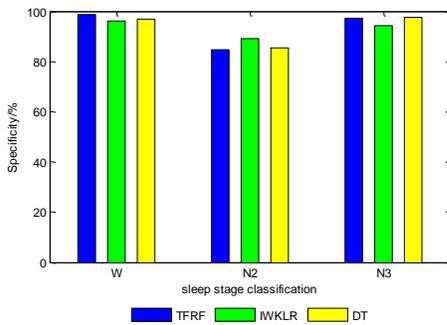


FIGURE III. COMPARISON OF SPECIFICITY ABOUT THREE METHODS

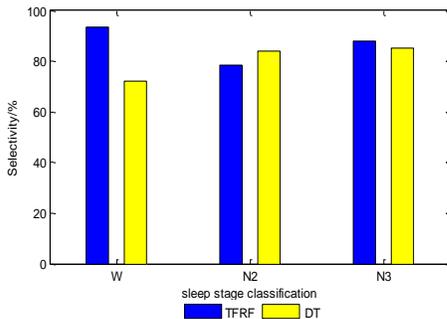


FIGURE IV. COMPARISON OF SELECTIVITY ABOUT TWOMETHODS

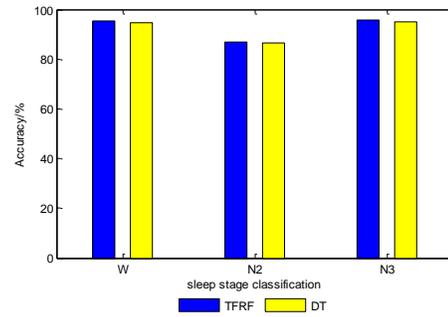


FIGURE V. COMPARISON OF ACCURACY ABOUT TWO METHODS

Results, 89.6% sensitivity of N3, 90.4% sensitivity of N2 and 81.6% sensitivity of W, are reported by the TFRF algorithm in tables. It shown that performances are better on the classification of sleep stage N3, N2 and Wake than many other algorithm proposed previously in both tables and figures.

Features was extracted by Genetic Algorithm and HMM model was modelled by Fuzzy Vector Quantization in the algorithm [15] proposed by Sheng-Fu Liang. Then, the sleep stages classified by the HMM model. The algorithm has a very good performance on REM stages, but the performance on other stages are not as good as the REM stages. This is because the waveform of other stages have obvious individual difference while REM stage has a special waveform with much less individual difference. TFRF is not sensitive to the individual difference, which is useful on the classification of N3, N2 and Wake stages.

Features were extracted from three channels: EEG, EOG and EMG signals by the maximum overlap discrete wavelet transform in the algorithm [16] proposed by Sirvan Khalighi. Features were selected by the minimum-redundancy maximum-relevance method. Then sleep stages were classified by the importance weighted kernel logistic regression (IWKLR). The algorithm has better performance on the classification between N1 and R, while its performance on the classification of the other three stages can be further improved, compared to TFRF. It is mainly due to its ignorance of the time domain features.

The algorithm [17] proposed by Syed Anas Imtiaz adopts the following steps: i) features were extracted in frequency domain, ii) the power of 10 sub-bands were calculated, iii) 3 spectral ratio of powers were computed, iv) line length of the signal in 11-16 Hz range was analyzed, and v) the results calculated above were regarded as features. Then sleep stages were classified by a State Machine-controlled Decision Trees (DT). The algorithm focuses on the classification of N1 and R, while its performance on the classification of the other three stages is not as good as TFRF. One reason for the better performance of TFRF compared to this, is adopted Random Forest instead of Decision Tree.

IV. CONCLUSION

In the paper, a novel algorithm of sleep stages classification is proposed, i.e. an approach of sleep stage classification based on Time-Frequency analysis and RandomForest (TFRF), which it has advantages of EEG signals from only one channel and four parameters better. Features of the signals are extracted by FFT and the magnitude spectra represented the frequency of 0~30Hz

are gotten, the envelope is calculated as features of time domain by Hilbert transform and Hjorth parameters, and sleep stages is classified by Random Forest in TFRF, which its performance is very well. This method can be applied on many fields such as wearable devices, as the algorithm can be performed on single channel while obtains satisfying results. TFRF in the paper has advancement on the time-domain parameters and the classification of the R and N1 sleep stages, which will be future researching directions.

Society, EMBC, 2011 Annual International Conference of the IEEE, vol. no., pp.3306,3309, Aug. 30 2011-Sept. 3 2011.

REFERENCES

- [1] A. Rechtschaffen and A. Kales, *A Manual Of Standardized Terminology, Techniques And Scoring System For Sleep Stages Of Human Subjects*, Washington DC, 1968.
- [2] C. Iber, S. Ancoli-Israel, J. AL. Chesson, and S. Quan, *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Westchester, IL: AASM, 2007.
- [3] Sheng-Fu Liang, Ching-Fa Chen and Jian-Hong Zeng, Application of Genetic Algorithm and Fuzzy Vector Quantization on EEG-based automatic sleep staging by using Hidden Markov Model, 2014 International Conference on Machine Learning and Cybernetics (ICMLC)
- [4] Sirvan Khalighi, Teresa Sousa and Urbano Nunes, Adaptive Automatic Sleep Stage Classification under Covariate Shift. Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE
- [5] Syed Anas Imtiaz and Esther Rodriguez-Villegas, Automatic Sleep Staging Using State Machine-controlled Decision Trees, Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE
- [6] The DREAMS Subjects Database. [Online]. Available: <http://www.tcts.fpms.ac.be/~devuyt/Databases/DatabaseSubjects/>
- [7] Md. Riyasat Azim, Md. Shahedul Amin and Shah Ahsanul Haque, Feature extraction of human sleep EEG signals using wavelet transform and Fourier transform, 2010 2nd International Conference on Signal Processing Systems (ICSPS).
- [8] S.C. Kak, The discrete Hilbert transform, *Proceedings of the IEEE* (Volume: 58, Issue: 4, April 1970)
- [9] Hjorth, Bo; Elema-Schönander, AB (1970). "EEG analysis based on time domain properties". *Electroencephalography and Clinical Neurophysiology*. 2.2: 306–310
- [10] Tin Kam Ho, (1995). Random Decision Forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, Montreal, QC, 14–16 August 1995. pp.278–282.
- [11] Breiman, L. Random Forests. *Machine Learning*. (2001) 45:5. doi:10.1023/A:1010933404324
- [12] Chanakya Reddy Patti, Sobhan Salari Shahrababaki, Chamila Dissanayaka, Application of random forest classifier for automatic sleep spindle detection, *Biomedical Circuits and Systems Conference (BioCAS)*, 2015 IEEE
- [13] E. A. Vivaldi and A. Bassi, "Frequency domain analysis of sleep EEG for visualization and automated state detection," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 1, pp. 3740–3743, 2006.
- [14] P. Van Hese, W. Philips, J. De Koninck, R. Van de Walle, and I. Lemahieu, "Automatic detection of sleep stages using the EEG," in *Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2, 2001, pp. 1944 - 1947.
- [15] S. Imtiaz and E. Rodriguez-Villegas, "Recommendations for performance assessment of automatic sleep staging algorithms," in *IEEE EMBC*, Chicago, August 2014.
- [16] See, A.R.; Chih-Kuo Liang, "A study on sleep EEG Using sample entropy and power spectrum analysis," *Defense Science Research Conference and Expo (DSR)*, 2011, vol., no., pp.14, 3-5 Aug. 2011.
- [17] Khalighi, S.; Sousa, T.; Oliveira, D.; Pires, G.; Nunes, U., "Efficient feature selection for sleep staging based on maximal overlap discrete wavelet transform and SVM," *Engineering in Medicine and Biology*