Classification of Awake, REM, and NREM from EEG via Singular Spectrum Analysis

Sara Mahvash Mohammadi[1], Shirin Enshaeifar[2], Mohammad Ghavami[1], and Saeid Sanei[2]

Abstract— In this study, a single-channel electroencephalography (EEG) analysis method has been proposed for automated 3-state-sleep classification to discriminate Awake, NREM (non-rapid eye movement) and REM (rapid eye movement). For this purpose, singular spectrum analysis (SSA) is applied to automatically extract four brain rhythms: delta, theta, alpha, and beta. These subbands are then used to generate the appropriate features for sleep classification using a multi class support vector machine (M-SVM). The proposed method provided 0.79 agreement between the manual and automatic scores.

I. INTRODUCTION

Sleep analysis plays a significant role in clinical and physiological sciences. Objective assessment of sleep is often based on the monitoring of sleep and wake states throughout the entire night's sleep. An important factor in sleep analysis is discrimination of sleep–wake states. To this end, polysomnography (PSG) signals are recorded and analyzed according to the standard rules. The Rechtschaffen and Kales standard (R&K) [1] and American Academy of Sleep Medicine (AASM) [2] are the most common guidelines to govern the standards for sleep classification. According to R&K, standard sleep stages are divided into: Awake, rapid eye movement (REM), and non-rapid eye movement (NREM) including stages 1, 2, 3, and 4 [3]. The study of sleep and wakefulness can be accomplished through PSG including electroencephalography (EEG), electromyogram (EMG), electrooculogram (EOG), respiratory effort and an electrocardiogram (ECG) [4].

Awake stage occurs at the beginning of sleep and is characterised by alpha rhythm (8-12 Hz), eye movements and high muscle tone. Almost 75% to 80% of the total sleep cycle is formed by NREM sleep. Each stage of the four stages of NREM corresponds to discrete brain activity and physiology. Stage 1 is referred to as a shift stage from wakefulness to sleep with a considerable representation of theta-wave activity (4-7.5 Hz). Stage 2 can be identified by the incidence of sleep spindles and K-complexes. Stages 3&4 are associated with delta activity which are combined and called slow wave sleep (SWS). Following the NREM stages, the REM stage, which contributes 20%-25% of total sleep, is characterized by incidence of rapid movements of the eye muscles under closed eyelids [5]. Fig. 1 indicates ten second segments of a single channel (C3–A2) of the EEG

[1]Sara Mahvash Mohammadi and Mohammad Ghavami with Department of Engineering and Design, London South Bank University, London, UK mahvashs@lsbu.ac.uk

[2]Shirin Enshaeifar and Saeid Sanei are with Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, UK s.sanei@surrey.ac.uk
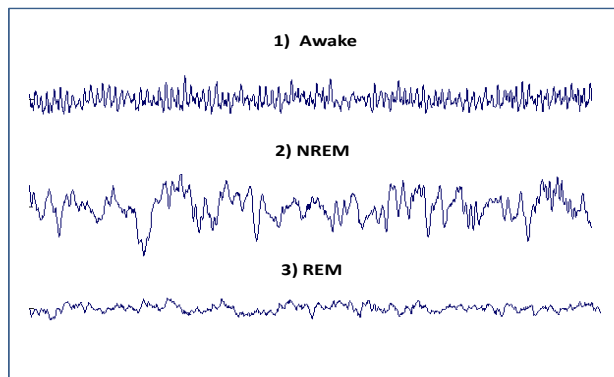
Fig. 1. Ten second segments of single channel EEG signal in three sleep states (Awake, NREM, and REM)

signal in three sleep states (Awake, NREM, and REM).

Since manual sleep classification is very time consuming, several methods have been developed for automatic sleep staging. Generally, the research of automated sleep analysis is basically focused on three basic tasks: artifact removal, feature extraction and stage classification.

In the same spirit, this study proposes an automated sleep classification method using singular spectrum analysis (SSA). SSA is a widely used method of a well-established time series analysis. In contrast to the previous time-series methods, the basic SSA method is non-parametric and we do not have to make prior statistical assumptions about the data. SSA has been utilized in biomedical signal processing area for different applications including detection of murmur from heart sounds [6], localizing localizing heart sounds in respiratory signals [7], and separation of ECG and EMG [8]. In this paper, SSA is applied to extract the desired EEG subcomponents, such as delta, theta, alpha and beta. These subbands are then utilized to construct features for sleep classification.

Many real signals particularly physiological signals such as EEG are nonstationary. However, SSA comprises the elements of multivariate geometry and statistics, classical time series analysis, signal processing, linear algebra, and dynamical systems which can dominate the nonstationarity of the signal [9]. Another remarkable benefit of SSA is that during the decomposition stage the noise component can be eliminated from the signal.

In this paper for automatic features extraction a new constrained SSA has been proposed. These features are then classified using a multi class support vector machine

(M-SVM) and compared to visual analysis. The result showed a high level of agreement between the automatic and manual analysis.

The paper is organised as follows: section 2 explains the fundamentals of the employed method. Section 3 reveals the result of implementing the proposed method to real data and the last section draws the concluding points.

II. METHOD

A. Singular Spectrum Analysis

SSA is a time-series technique which can be used in decomposing a signal into a sum of interpretable components such as trend, periodic, quasi-periodic and noise, and can be applied to any time-series with a complex structure [6] [8]. The basic SSA technique comprises two stages which complement each other; decomposition and reconstruction. Both stages, in turn, entail two discrete stages.

1) Decomposition: This stage includes embedding operation accompanied by singular value decomposition (SVD).
a) Embedding: In the embedding stage, an original signal presented as a one-dimensional vector $\mathbf{f}_s = (f_1,\ldots,f_s)^T$ of length s is chosen. At this stage, the vector should be mapped into a matrix $\mathbf{X} \in \mathbb{R}^{l \times n}$ known as a multidimensional trajectory matrix where $n = s - l + 1$ and l is the window length $(1 < l < s)$. The embedding method starts by producing n lagged vectors $(\mathbf{x_i})$ and continues by entering them as columns of the desired trajectory matrix [9]:

$$
\begin{aligned}
\mathbf{x}_i &= (f_i, f_{i-2}, \ldots, f_{i+l-1}) \\
\mathbf{X} &= [\mathbf{x}_1, \ldots, \mathbf{x}_n] \\
&= (x_{ij})_{i,j=1}^{l,n} \\
&= \begin{pmatrix}
f_1 & f_2 & f_3 & \ldots & f_n \\
f_2 & f_3 & f_4 & \ldots & f_{n+1} \\
\vdots & \vdots & \ddots & & \vdots \\
f_l & f_{l+1} & f_{l+2} & \ldots & f_s
\end{pmatrix}
\end{aligned} \quad (1)
$$

$\mathbf{X}$ is a Hankel matrix as the entries along the diagonals $(i + j = \text{const})$ are equal. Remember that, in order to keep the information about the data variation, the window length l should be adequately large [9], [6]. *b) Singular Value decomposition*: In this step, SVD is applied to the produced trajectory matrix. j-th eigenvalue (λ_j) and eigenvector $(\mathbf{q}_j)$ of covariance matrix $\mathbf{C_x} = \mathbf{X}\mathbf{X}^T$ define the j-th component of SVD. $\mathbf{C_x}$ can be diagonalised as $\mathbf{Q}\Sigma\mathbf{Q}^T$ since it is a symmetric-positive matrix, where Σ is a diagonal matrix of eigenvalues arranged in decreasing order $(\lambda_1 > \lambda_2 > \ldots > \lambda_l)$ and $\mathbf{Q}$ is the corresponding eigenvectors [9]. Therefore, SVD of the trajectory matrix can be rewritten as:

$$
\begin{aligned}
\mathbf{X} &= \sum_{j=1}^{r} \mathbf{X}_j = \sum_{j=1}^{r} \sqrt{\lambda_j}\mathbf{q}_j\mathbf{v}_j^T \\
\mathbf{v}_j &= \mathbf{X}^T\mathbf{q}_j/\sqrt{\lambda_j}
\end{aligned} \quad (2)
$$

in which $\sqrt{\lambda_j}$ is known as the singular value of matrix $\mathbf{X}$ and

$$
r = \max\left(j, \text{ such that } \lambda_j > 0\right) = \text{rank}(\mathbf{X}) \quad (3)
$$

The set $(\lambda_j, \mathbf{q}_j, \mathbf{v}_j)$ is named the j-th eigentriple of the matrix $\mathbf{X}$. The definition of $\mathbf{X}_j$ is equivalent to the elementary matrix.

2) Reconstruction: This stage is composed of two steps: grouping and diagonal averaging.
a) Grouping: In this stage, the elementary matrices from the previous stage are split into several $\tilde{\mathbf{X}}_g$ where:

$$
\mathbf{X} = \sum_{g=1}^{g_t} \tilde{\mathbf{X}}_g \quad (4)
$$

in which g_t shows the total number of groups, index g assigns the g-th subgroup of eigentriples, and $\tilde{\mathbf{X}}_g$ specifies the sum of $\mathbf{X}_j$ within group g.
b) Diagonal averaging: In the final stage, a specific $(\tilde{\mathbf{X}}_g)$ is selected and transformed into the form of a Hankel matrix. This matrix is then converted into a process using the diagonal averaging. To do this, the k-th element of the signal is obtained by averaging the matrix entries $\tilde{x}_{ij}$ over all i, j where $(i + j = k + 1)$.

B. Constrained SSA

Grouping the desired subspaces is a challenging issue for the SSA-based applications. To solve this, generally, some a priori knowledge is used to define heuristic criteria for noise rejection and subspace extraction. Similarly, sleep EEG signals contain both desired components (meaningful brain rhythms) and noise (artifacts). Thus, SSA is applied to remove the noise and separate the desired components such as delta, theta, alpha and beta. The fundamental periodic components are localized by analyzing the relevant subspace. Thus, it is first required to define an appropriate signal space so that it covers all essential subspaces. To this end, noise space is associated with eigenvalues below 90% of the total variance of the signal. Thus, eigenvalue λ_j with index j is rejected if $j > \mathscr{L}$ [7].

$$
\mathscr{L} = \min\left\{h, \text{ such that } \frac{\sum_{i=1}^{h} \lambda_i}{\sum_{i=1}^{l} \lambda_i} > 0.9\right\} \quad (5)
$$

In addition, the desired frequency band is used as a priori knowledge for the grouping stage to reconstruct the required brain subbands. Generally, in SSA periodic components are represented by a pair of eigenvalues with similar amplitudes. However, the following points should be considered for the eigenvalue pairs selection: 1) two eigenvalues may not be exactly equal and 2) the spurious pairs can be created due to the noise. Therefore, in order to acquire the actual periodic pair, the eigenvalue pairs λ_j and λ_i are selected as a pair only if both of the following conditions are applied:

1) The corresponding pair of indexes are within the signal subspace (index i and j are less than $\mathscr{L}$, where

$\mathscr{L}$ is defined in (5))

2) $|1 - \frac{\lambda_j}{\lambda_i}| < 0.05$

Rejection in selecting the eigenvalue pairs means that no component is found in the desired time domain signal [10]. However, if some eigenvalue pairs are selected, the highest peak in the Fourier transform of the corresponding eigenvectors is relevant to the frequency of the periodic component [11]. Thus, the spectral power is utilized to compute the relevant frequency. The peak in the Fourier transform is needed to fall into the bandwidth of interest delta (1-4 Hz), theta (4-7.5 Hz), alpha (7.5-12 Hz), beta1 (16-18 Hz), and beta2 (18-40 Hz).

C. Feature Extraction

After reconstructing the desired components, the following features are obtained for individual EEG epochs, i.e. window of 30-seconds.

- The mean and standard deviation (SD) of different subbands including delta (up to 4 Hz), theta (4-7.5 Hz), alpha (7.5-12 Hz), lower beta or beta1 (16-18 Hz), and higher beta or beta2 (18-40 Hz).
- Sum of the power in all five frequency bands.
- The mean and SD of various ratios $\frac{alpha}{sum}$, $\frac{beta1}{sum}$, adding the other 4 features.
- Since the resting wakefulness is usually determined by plentiful alpha frequencies and infrequent delta frequencies, the mean and SD of the ratio $\frac{alpha \times beta1}{delta}$ are calculated as features the for awake state [12].

Using the above features, each 30-second EEG data epoch is represented by a 17-dimensional input vector which is utilized in a classification algorithm.

D. Multiclass SVM classification

SVM is a supervised learning method used for regression and classification in various applications [13]. One hallmark of SVM is that, it can minimize the empirical classification error and at the same time maximize the geometric margin. SVM looks for an optimal hyperplane to classify through different classes exploiting a few features with the maximum distance between the decision boundaries and training set [14].

Principally, the design of SVMs are based on discrimination between two classes. However, our objective here is to automate sleep state classification for discriminating between three classes (Awake, NREM, and REM). For this purpose, application of an M-SVM classifier is more appropriate. For this classifier, a "one–against–all" approach is used [15].

III. APPLICATION TO SLEEP EEG

To evaluate the performance of the proposed method, we have applied it to sleep analysis. Sleep consists of two major states NREM and REM with different neural activities. In this study, we aim to propose a 3 state classification to discriminate the Awake-NREM-REM stages. To this end, we have used a subset of the dataset recorded in the Sleep Centre

TABLE I
SVM CLASSIFICATION RESULTS VIA SSA.

Manual \ Automatic	Awake	NREM	REM	Sensitivity
Awake	18	0	2	90.0%
NREM	15	281	12	91.2%
REM	2	0	52	96.2%
Specificity	51.4%	100.0%	78.8%	
Agreement				92.0%
Cohen's Kappa				0.79

TABLE II
THE RELATION BETWEEN LEVEL OF AGREEMENT AND KAPPA VALUES [16]

Agreement	Kappa
Substantial	0.61 - 0.80
Moderate	0.41 - 0.60
Fair	0.21 - 0.40
Slight	0 - 0.20

of the University of Surrey. PSG data of approximately 8 nights is collected. The dataset is recorded from subjects participating in an EEG recording for a baseline night (BL, 8 hours).

To start the EEG decomposition, the eigenvalue pairs were identified for individual brain rhythms (delta, theta, alpha, beta) using the aforementioned criteria. Then, each subband was reconstructed via the corresponding eigentriple and its Fourier transform was calculated. After generating the desired brain subbands, features were extracted for a single epoch and were fed to the multi class support vector machine (M-SVM) classifier, (see Section 2.3). As we have used M-SVM, it automatically constructed a model by following the dynamics of the training data.

A. Result

382 epochs from 6 subjects are selected. Half of the dataset are utilized for training set and the rest are considered for testing set. In order to evaluate the classification, the confusion matrix of 3 state Awake-NREM-REM manual scoring versus automated scoring via SSA is presented in Table 1.

To evaluate the performance of our classifier three statistical measures were obtained: (i) sensitivity (reflects the performance of the automatic classifier compared to the visual analysis) (ii) specificity and (iii) Cohen's Kappa coefficient of agreement [17] which provides a more insightful measure of the general performance of the classifier. A heuristic way of explaining the Kappa values is represented in Table 2 [16]. Utilizing single channel EEG signal and automatic analysis through SSA, the epoch-by-epoch agreement (percentage of epochs that were assigned to the same state) was 92% and Kappa coefficient was measured as 0.79 which represents substantial agreement between manual and automated scoring. As shown in Table 1, classification sensitivities of all three states exceeded 90% which indicates that the individual states have been well discerned via the features extracted
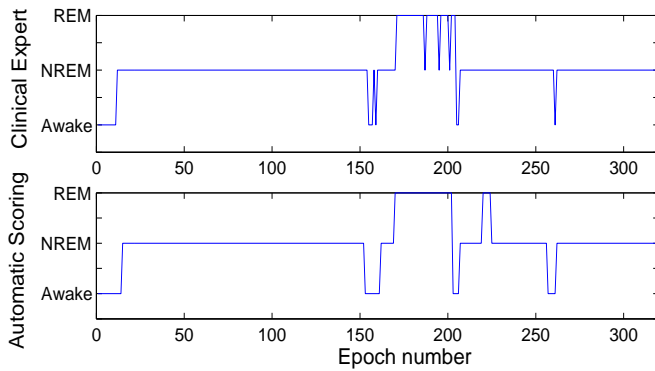
Fig. 2. Representative 3 state sleep classification of the first 320 epochs obtained using manual scoring (top) and SSA (bottom).

from SSA. As an example, Fig. 2 illustrates both manual and automatic scoring obtained for the first 320 epochs of a single subject.

Note that in most automatic sleep scoring systems both EMG and EOG are considered for differentiating the REM stage [18], [19]. However, in this study we have performed single channel (C3–A2) EEG analysis. As we defined specific features for Awake state and we applied 3 states classification, REM state is well distinguished from Awake and general NREM states.

IV. CONCLUSIONS

In this paper, we have utilized a constrained SSA method to decompose a single-channel sleep EEG and extract the features. The proposed method not only extracted the periodic components automatically, but also removed the inherent noise subspace. Using the generated features, M-SVM performed 3 state classification to distinguish the Awake-NREM-REM stage. The proposed method paves the way for further analysis of sleep EEG to enable characterisation of sleep abnormalities and many mental and physical disorders. The novel approach presented here is also able to extract the sleep spindles and K-complexes as another avenue for more investigations.

V. ACKNOWLEDGMENT

REFERENCES

[1] A. Rechtschaffen and A. Kales, "A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects," 1968.

[2] C. Iber, *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*, American Academy of Sleep Medicine, 2007.

[3] G. Zhu, Y. Li, and P. P. Wen, "Analysis and classification of sleep stages based on difference visibility graphs from a single-channel EEG signal," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 6, pp. 1813–1821, Nov 2014.

[4] J. M. Kortelainen, M. O. Mendez, A. M. Bianchi, M. Matteucci, and S. Cerutti, "Sleep staging based on signals acquired through bed sensor," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 3, pp. 776–785, May 2010.

[5] S. Sanei, *"Sleep and Sleep Apnoea", Adaptive Processing of Brain Signals*, pp. 274–294, John Wiley & Sons, Ltd, 2013.

[6] S. Sanei, M. Ghodsi, and H. Hassani, "An adaptive singular spectrum analysis approach to murmur detection from heart sounds," *Elsevier Journal of Medical Engineering and Physics*, vol. 33, no. 3, pp. 362 – 367, 2011.

[7] F. Ghaderi, H. R. Mohseni, and S. Sanei, "Localizing heart sounds in respiratory signals using singular spectrum analysis," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 12, pp. 3360–3367, 2011.

[8] S. Sanei, T. K. M. Lee, and V. Abolghasemi, "A new adaptive line enhancer based on singular spectrum analysis," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 2, pp. 428–434, Feb 2012.

[9] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of time series structure: SSA and related techniques*, CRC Press, 2010.

[10] J. Mamou and E. J. Feleppa, "Singular spectrum analysis applied to ultrasonic detection and imaging of brachytherapy seeds," *Journal of the Acoustical Society of America*, vol. 121, no. 3, pp. 1790–1801, 2007.

[11] Y. Tao, E. C. M. Lam, and Y. Y. Tang, "Feature extraction using wavelet and fractal," *Elsevier Journal of Pattern Recognition Letters*, vol. 22, no. 3, pp. 271–287, 2001.

[12] C. Berthomier, X. Drouot, M. Herman-Stoïca, P. Berthomier, J. Prado, and O. Benoit D. Bokar-Thire, J. Mattout, and M. d'Ortho, "Automatic analysis of single-channel sleep EEG: validation in healthy individuals," *Sleep*, vol. 30, no. 11, pp. 1587, 2007.

[13] G. Zhu, Y. Li, and P. P. Wen, "An efficient visibility graph similarity algorithm and its application on sleep stages classification," in *Brain Informatics*, vol. 7670 of *Lecture Notes in Computer Science*, pp. 185–195. Springer Berlin Heidelberg, 2012.

[14] V. Vapnik, *Statistical learning theory*, vol. 1, Wile, NY, 1998.

[15] C-W. Hsu and C-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.

[16] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *biometrics*, pp. 159–174, 1977.

[17] J. Cohen, "A coefficient of agreement for nominal scales," *Educational Psyhol Meas*, vol. 20, pp. 37–46, 1960.

[18] E. Estrada, H. Nazeran, J. Barragan, J. R. Burk, E. A. Lucas, and K. Behbehani, "EOG and EMG: Two important switches in automatic sleep stage classification," in *Engineering in Medicine and Biology Society. EMBS'06. 28th Annual International Conference of the IEEE*, 2006, pp. 2458–2461.

[19] J. Kempfner, G. L. Sorensen, H. B. D. Sorensen, and P. Jennum, "Automatic REM sleep detection associated with idiopathic REM sleep behavior disorder," in *Engineering in Medicine and Biology Society, EMBC, Annual International Conference of the IEEE*, 2011, pp. 6063–6066.