Automated Sleep Spindle Detection Using Novel EEG Features and Mixture Models

Chanakya Reddy Patti, Ramiro Chaparro-Vargas, *Member* and Dean Cvetkovic, *Member*

***Abstract*— Research in automated Sleep Spindle detection has been highly explored in the past few years. Although a number of automated techniques were developed, many of them were based on using fixed parameters or thresholds which do not consider subject specific differences. In this research study, we introduce a novel method of sleep spindle detection using Gaussian Mixture Models with no fixed parameters or thresholds. The algorithm was tested on an online public spindles database consisting of six 30 minute sleep excerpts extracted from whole night recordings of 6 subjects. The results obtained were better when compared with other methods. We obtained an overall sensitivity of 74.9% at a 28% False Positive proportion.**

I. Introduction

Sleep scoring involves the identification of numerous transient events from Electroencephalogram (EEG), Electrooculogram (EOG) and Electromyogram (EMG) signals associated with sleep stages and events, such as sleep spindles. They are characteristic waves present in the EEG signal during sleep stages 2 and 3. According to the latest standard for sleep scoring (i.e the American Academy of Sleep Medicine (AASM) sleep standard [1]), spindles are defined as bursts of activity in the 11-16Hz frequency range with a minimum duration of 0.5 seconds [1] observed in EEG.

Previous work in the area of spindle detection has led to various techniques being developed. Many of these techniques involve band pass filtering or using frequency related features to separate spindles from non spindles [2]. Frequency and amplitude features calculated using Short Time Fourier Transform (STFT) were used as inputs to neural networks by Nurettin et al. [3] and Gorur et al. [4]. Spindle detection with band pass filtering and amplitude thresholding was also widely used [2]. Due to differences in ideal amplitude threshold in subjects, various recording-specific threshold calculation methods were developed. Hupponean et al. [5] and Ray et al. [6] have proposed methods which calculate recording specific thresholds based on spectral features. Other methods used were Matching Pursuit (MP) techniques [7] and Discrete Wavelet Transform methods [8].

Although many past works reported significant results, there had been no common database to evaluate and compare the various techniques against each other. Devuyst et al. [9] proposed a standard assessment method which involves testing algorithms on their sleep spindles database which was published online [10]. The algorithm used by Devyust and colleagues was based on band pass filtering and level detection using recording specific thresholds. The recording specific threshold calculation was based on the work done by Hupponean et al. [4] which involves calculating a Bayes threshold to distinguish spindles from non spindles. The algorithm also used a non-recording specific threshold of 0.22 for the ratio of the power in the spindle region (11- 15 Hz) to that of total power in a 1 second moving-window.

The aim of this paper is to implement automated detection of sleep spindles with the highest possible accuracy without the use of any non-recording specific thresholds.

II. Methodology

A. Database Details-

The Polysomnography (PSG) sleep data containing spindles was obtained online from the Dreams Sleep Spindle Database [10]. The database contains six 30 minute sleep excerpts extracted from whole night recordings of 6 subjects, aged between 30 and 55 years. The visual scoring of spindles was undertaken by two independent scorers and the scoring data was available as part of the Dreams Database. All subject excerpts were sampled at 200 Hz except for subject 1(100Hz) and subject 3(50Hz).

B. Detection Method Overview

The overview of the detection method in Fig. 1 shows the

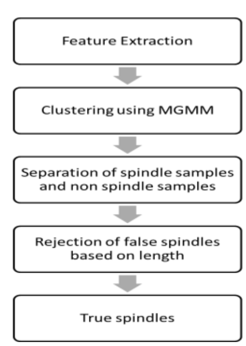

Figure 1. Overview of the Sleep Spindle Detection Method.

C.R. Patti is a doctoral student at the Royal Melbourne Institute of Technology,Melbourne,VIC 3001, Australia. (s3304790@student.rmit.edu.au)
R. Chaparro-Vargas is a doctoral student at the Royal Melbourne Institute of Technology, Melbourne, VIC 3001, Australia. (s3361953@student.rmit.edu.au)
D. Cvetkovic is with the School of Electrical and Computer Engineering, Royal Melbourne Institute of Technology, Melbourne, VIC 3001, Australia (dean.cvetkovic@rmit.edu.au)

sequential stages of sleep spindle detection starting with the extraction of relevant features followed by separation of data using a Multivariate Gaussian Mixture Model (MGMM) and the removal of false spindles.

C. Feature Extractions-

Four spindle related features were extracted and then classified using a MGMM approach. To extract the four features, a STFT moving window of a given set length was used. During analysis of results this moving window length was changed over a range of values to examine the effect of change on the results. A (N-1) overlap moving window was used, i.e the moving window was shifted by each sample. For purposes of demonstration and explanation a 1.5 second moving window size (no zero padding) is used in the following descriptions.

For each window, the following features were calculated:

1) *Sigma Power* was calculated as the energy in the 12.5 – 15Hz frequency band in the moving window. The narrow band produced the best results, compared to the 10 - 16Hz band and the 11 - 15Hz band.
2) *Sigma Index* was calculated as the ratio of energy in the 12.5 - 15Hz band to the total energy (0.5Hz - 40Hz) in the moving window.

$$\textit{Sigma Index (t)} = \frac{\int_{12.5}^{15} S(f,t)\,df}{\int_{0.5}^{40} S(f,t)\,df} \qquad (1)$$

3) *Sigma Power 2* was calculated as the ratio of spindle power of a 1.5 second window to the spindle power in the two adjacent moving windows to the current window. This is illustrated in Fig. 2. Since spindles are limited to the 0.5 -2 second time length, it was assumed that if a spindle exists in a moving window, there are no spindles in the adjacent moving windows (i.e. separated by the length of the moving window, please refer to Fig. 2). A high Relative Spindle Power 2 indicates a high probability of a spindle existing in the current moving window. This feature can also be stated as the ratio of *Sigma Power (feature 1)* in the window of interest to the sum of *Sigma Power* in the two adjacent windows.
4) *Sigma Index 2* like the Sigma Power 2 feature, this is the ratio of the relative spindle power in a moving window with the sum of relative spindle powers in the two adjacent moving windows. This feature can also be stated as the ratio of *Sigma Index (feature 2)* in the window of interest to the sum of Sigma index in the two adjacent windows.

The four features are shown in Fig. 2 for a 15 second epoch.

D. Clustering using Gaussian Mixture Modelling

The four input features were mapped into two clusters. Cluster one consists of non-spindle segments and cluster two consists of spindle segments. The clustering was based on the Expectation Maximization (EM) algorithm [13]. In order to separate spindles from non spindles a MGMM was used to cluster the four features into a spindle and non-spindle clusters. The EM algorithm is based on the probability density function of a multivariate Gaussian mixture model which is defined in the following

$$p(x) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu) \right\} \qquad (2)$$

where d is the dimension of x (feature vector, in this case d = 4), μ is a vector consisting of mean of individual features and $\sum$ is the covariance matrix of the four features.

The result of Gaussian clustering will produce two multivariate Gaussian clusters. One cluster defines the non-spindle segments and the other cluster defines the spindle segments.

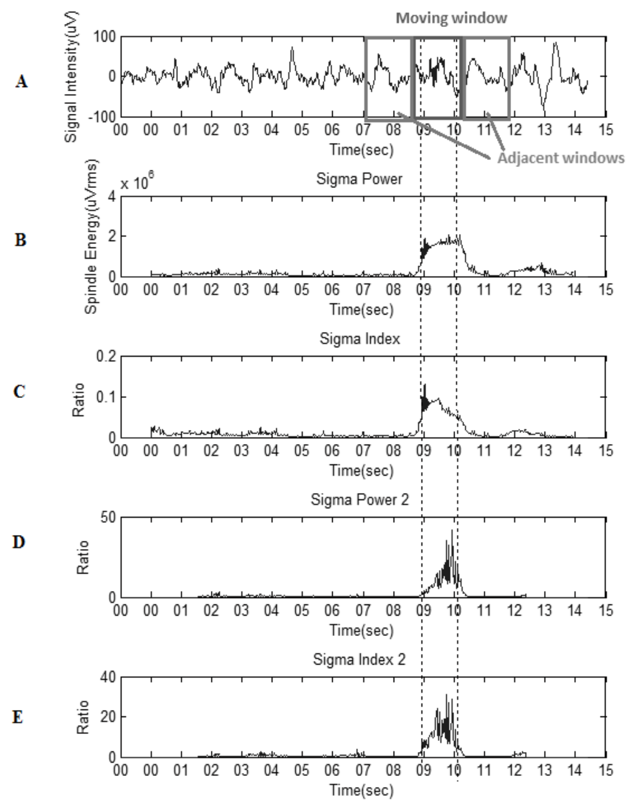

Figure 2. Four Features for a 15 second epoch containing a 0.5 second spindle. A- Original 15 sec signal, B- Feature 1(Sigma Power),C- Feature 2(Sigma Index),D- Feature 3(Sigma Power 2), E- Feature 4(Sigma Index 2)

All features were normalised to have total energy between 1 and 0, where 1 is the maximum and 0 is the minimum.

E. Expectation Maximisation (EM) algorithm and Clustering

EM is an iterative algorithm applied to clustering data when using mixture models. The EM algorithm iteratively tries to decrease the maximum likelihood of the required number of clusters by altering the cluster means and cluster covariance matrices [13].

A MATLAB function from its statistics toolbox ('gmdistribution.fit') was used for calculating the cluster parameters. It was observed that the resulting clustering always converged to similar values, hence ensuring that there were no multiple solutions to the clustering problem.

F. Spindle Length Check to Separate False Detections

A spindle of *x* samples length will last inside a window of length *k* for *(k-x)* samples completely and partially for the whole window length. To ensure that only true spindles were classified as spindles, a spindle needed to last for majority part of the window length. The ideal length to check for was established empirically and was calculated to be 90% of window length. Fig. 3 below shows Receiver Operator Characteristics (ROC) calculated by varying the spindle length check (from 70% of window length to 100% of window length) to classify spindles. The intersection of the line connecting the top left corner and bottom right corner with the curves gave the ideal false detection check length. The length approximately corresponds to be 90% of moving window size.

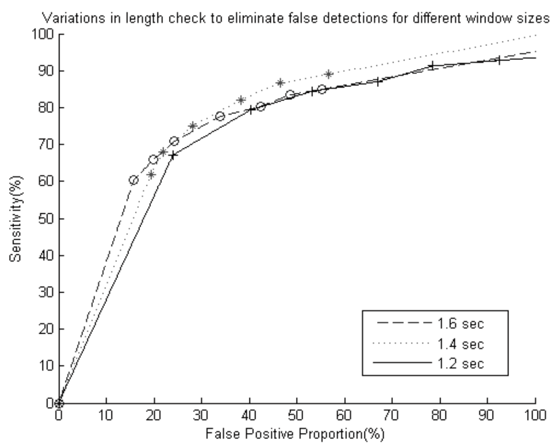

Figure 3. ROC curves for various length check to remove false spindle detections.

III. RESULTS

The spindles classified by the clustering algorithm were compared to the spindles scored by two individual visual scorers. A standardised method of assessment developed by Devuyst et al [8] was used to assess the clustering algorithm. The results were also compared to spindle detection method developed by Devuyst et al. Table 1. shows the results of our method on each of the 6 subjects data using a 1.4 second STFT moving window.

The parameters estimated are - Tp – True positives, Fn – False Negatives, Fp – False Postives, Tn – True negatives, TDDS–total duration of database in seconds.

$$\text{Sensitivity} = \frac{Tp}{Tp + Fn} \quad (3)$$

$$\text{False positive proportion} = \frac{Fp}{Tp + Fn} \quad (4)$$

$$\text{Specificity} = \frac{Tn}{Tn + Fp} \quad (5)$$

True negatives were calculated using equation 6, which is based on the approximation given by Devyust et al [8].

Tn = TDDS - Tp- Fn - Fp (6)

The results for a moving window length of 1.4 seconds are shown in Table 1 below

TABLE I. RESULTS OBTAINED USING A 1.4 SECOND STFT WINDOW

Subjects	Sex	Age	Sensitivity %	False Positive proportion %	Total number of spindles scored by visual scorers	True positives	False positives
S1	F	51	75.37	17.16	134	101	23
S2	M	40	77.92	28.57	77	60	22
S3	M	46	90.90	63.63	44	40	28
S4	M	31	39.68	31.74	63	25	20
S5	F	53	77.66	30.09	103	80	31
S6	F	53	82.9	23.07	117	97	27
All Male subjects			67.93	38.04	184	125	70
All Female Subjects			78.53	22.88	354	278	81
All subjects			74.90	28.06	538	403	151

The results were also obtained for various STFT window sizes ranging from 0.4 seconds to 2.8 seconds in increments. The resulting ROC curve is shown in Fig.4

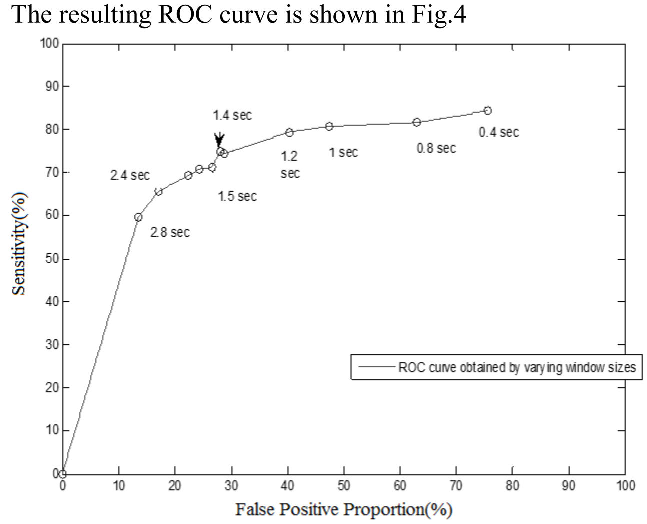

Figure 4. ROC curve obtained by varying STFT moving window size.

Based on the ROC curve in Fig. 4, the ideal results were obtained using a 1.4 second window. Out of the total of 538 spindles, 403 spindles were correctly and 151 falsely identified. The sensitivity achieved was 74.9% and the false positive proportion was 28.07%. Specificity was calculated using the method described in Devyust et al [8] was 98.53%.

The above results are based on using the narrow band frequency of 12.5 – 15Hz for feature calculation. This band width was used since it produced ideal results. This may be due to the lack of interference from the alpha band or the beta bands. The results for different band-widths using different window sizes are shown in Table 2.

TABLE II. RESULTS OBTAIED USING DIFFERENT BANDWIDHTS

Window size	10Hz- 16Hz		11Hz-15Hz		12Hz-14Hz		12.5Hz-15Hz	
seconds	S (%)	FP (%)	S (%)	FP (%)	S (%)	FP (%)	S (%)	FP (%)
1.4	57.8	35.3	62.4	33.2	63.9	32.1	**74.9**	**28.0**
1.5	55.3	33.6	63.7	28.9	57.6	27.8	71.1	26.5
1.6	54.8	28.2	62.4	25.0	51.6	28.6	70.8	24.3
1.7	52.0	26.5	59.6	24.3	55.7	25.6	68.9	22.3

S- Sensitivity, FP – False positive proportion

IV. DISCUSSION AND CONCLUSION

The total number of spindles scored by Visual scorer no. 1 and scorer no. 2 were 289 and 409, respectively. The mutual agreement rate between scorers was therefore only 55%. Standard agreement rates between scorers as reported by Huppueon et al[12] is expected to be 81%.

The overall sensitivity of the spindle detection algorithm was 74.91% and the false positive proportion was 28% with a 1.4 second STFT moving window. A difference in sensitivity and false positive proportion was observed between the male and female subjects. The difference may be explained as due to lower number of spindles present in male subjects compared to female subjects. The sensitivity and accuracy of the algorithm were better in female subjects (Sensitivity – 75.14, FP proportion – 23.15%) compared to male subjects (Sensitivity – 63.58%, FP proportion – 33%). Devuyst and colleagues [3] reported a sensitivity of 70.1% and a false positive proportion of 26.44% based on their spindle detection algorithm and the same assessment method as it was developed by them [8].

Recently developed methods by Imityaz et al. in 2013 used a Teager energy operator method for Sleep spindle detection [11]. The database used was the same database used in our study. Their best results were reported in terms of only sensitivity and specificity. No False positive proportion values were provided. According to Imtiyaz et al. , the results showed 80.3% sensitivity and a specificity of 97.6%. The results obtained in this study can be compared with Imtiyaz et al. by examining the ROC curve for a 1.4 second window. At a similar sensitivity of 82.16% using approximately 85% of window length to remove false detections, a specificity (calculated using the method in Devyust et al[8]) of 97.9% was obtained.

The proposed method in this study has been carefully evaluated for various parameters. The advantage of the proposed method is that there were no fixed thresholds used and it introduced the concept of comparing a spindle to its immediate neighborhood.

REFERENCES

[1] C. Iber, S. Ancoli-Israel, A. Chesson and SF. Quan, "The AASM manual for the scoring of sleep and associated events : rules, terminology and technical specifications, " American Academy of Sleep Medicine, Westchester, Illinois (IL), 2007.

[2] P. Schimicek , J. Zeitlhofer , P. Anderer, B. Saletu, " Automatic sleep-spindle detection procedure: aspects of reliability and validity", *Clin. Electroencephalogr*, vol.25, pp 26-29, 1994.

[3] A. Nurettin and G. Cüneyt, "Automatic recognition of sleep spindles in EEG by using artificial neural networks," Expert Systems with Applications, vol. 27, pp 451-458, 2004.

[4] D. Gorur, U. Halici, H. Aydin, G. Ongun, F. Ozgen, K. Leblebicioglu, "Sleep spindles detection using short time Fourier transform and neural networks," *Neural Networks, 2002. IJCNN '02. Proceedings of the 2002 International Joint Conference on*, vol.2, pp.1631-1636, 2002.

[5] E. Huupponen, A. Värri, SL. Himanen, J. Hasan, M. Lehtokangas and J. Saarinen, "Optimization of sigma amplitude threshold in sleep spindle detection," *J Sleep Res*, vol. 9, pp 327-334, 2000.

[6] LB. Ray, SM. Fogel, CT. Smith and KR. Peters, "Validating an automated sleep spindle detection algorithm using an individualized approach," *J Sleep Res*., vol. 19, no. 2, pp 374-378, 2010.

[7] P.J. Durka , K.J. Blinowska "Analysis of EEG Transients by Means of Matching Pursuit," *Annals of Biomedical. Engineering*, vol. 23, pp 608-611, 1995.

[8] B. Ahmed, A. Redissi, R. Tafreshi, "An automatic sleep spindle detector based on wavelets and the teager energy operator," *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE,* pp.2596-2599, 3-6 Sept. 2009.

[9] S. Devuyst, T. Dutoit, P. Stenuit, M. Kerkhofs, "Automatic sleep spindles detection — Overview and development of a standard proposal assessment method," *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE* , pp.1713-1716, Aug. 30 2011-Sept. 3 2011.

[10] University of MONS - TCTS Laboratory (Stéphanie Devuyst, Thierry Dutoit) and Université Libre de Bruxelles - CHU de Charleroi Sleep Laboratory (Myriam Kerkhofs). [Online]. Available: {Stéphanie DEVUYST}, {http://tcts.fpms.ac.be/~devuyst/#Databases}

[11] S.A Imtiaz, S. Saremi-Yarahmadi, E. Rodriguez-Villegas, "Automatic detection of sleep spindles using Teager energy and spectral edge frequency," *Biomedical Circuits and Systems Conference (BioCAS), 2013 IEEE* , pp.262-265, Oct. 31 2013-Nov. 2 2013.

[12] E. Huupponen, G. Gomez-Herrero, A. Saastamoinen, A. Varri, J. Hasan, S-L. Himanen, "Development and comparison of four sleep spindle detection methods," Artificial Intelligence in Medicine, vol. 40, pp 157-170, 2007.

[13] A.P. Dempster, N.M. Laird, & D.B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal statistical Society*, vol.39, no.1, pp.1-38, 1977.