

Evaluating the use of line length for automatic sleep spindle detection

Syed Anas Imtiaz and Esther Rodriguez-Villegas

Abstract—Sleep spindles are transient waveforms observed on the electroencephalogram (EEG) during the N2 stage of sleep. In this paper we evaluate the use of *line length*, an efficient and low-complexity time domain feature, for automatic detection of sleep spindles. We use this feature with a simple algorithm to detect spindles achieving sensitivity of 83.6% and specificity of 87.9%. We also present a comparison of these results with other spindle detection methods evaluated on the same dataset. Further, we implemented the algorithm on a MSP430 microcontroller achieving a power consumption of 56.7 μ W. The overall detection performance, combined with the low power consumption show that line length could be a useful feature for detecting sleep spindles in wearable and resource-constrained systems.

I. INTRODUCTION

Human sleep is divided into two distinct phases known as Rapid Eye Movement (REM) and Non-Rapid Eye Movement (NREM). The NREM phase is further classified into three stages namely N1, N2 and N3 with the former being the lightest and latter being the deepest stage of sleep [1]. The beginning of N2 stage is marked by the presence of distinct burst of high frequency transients, which are observed on the sleep electroencephalogram (EEG), occurring for a duration of at least 0.5 seconds. These transient waveforms are known as *sleep spindles*. Their rate of occurrence is highest during N2 stage and they can also be seen during N3 stage of sleep. Identification of sleep spindles are of particular interest because of their role in classifying N2 stage of sleep. Further they are also linked with melatonin secretion aiding circadian rhythms [2], early stage development of central nervous system (CNS) [3], indication of intellectual ability [4] and memory consolidation during sleep [5]. An example of a sleep spindles during N2 stage is shown in Fig. 1a. The number of sleep spindles observed during an overnight sleep is in the range of 200-1000 [6]. Hence, their manual identification is a laborious and error-prone task. There may be subjective differences in multiple spindles scored by a single scorer and between different scorers. A previous study showed the variability between different scorers to be around 20% [7]. Hence automatic detection of spindles is desirable in order to save time and reduce variability.

Several methods of automatic sleep spindle detection using various features already exist in literature. Gorur et al. [8]

S. A. Imtiaz and E. Rodriguez-Villegas are with the Circuits and Systems Group, Electrical and Electronic Engineering Department, Imperial College London, United Kingdom. Email: ({anas.imtiaz,e.rodriquez}@imperial.ac.uk).

The research leading to these results has received funding from the European Research Council under the European Community's 7th Framework Programme (FP7/2007-2013) / ERC grant agreement no. 239749.

extracted spectral features of the signal with Fourier transform and used the resulting coefficients as input to artificial neural network (ANN) and support vector machine (SVM) classifiers. [9] also demonstrated the use of both ANN and SVM classifiers for spindle detected but used coefficients extracted from adaptive autoregressive (AR) modelling as feature input. [10] also used AR modelling with relative power and filtered signal thresholding. A data-driven probabilistic approach was used by Babadi et al. [11] that showed high detection accuracy. For detection of spindles in children [12] presented a method using Hilbert-Huang transform while [13] used amplitude-frequency normal modelling to detect spindles in both children and adults. In another method, Schönwald et al. [14] evaluated the use of matching pursuit (MP) for spindle detection and achieved good results. Duman et al. [15] used Teager energy, maximum frequency and harmonic decomposition with a decision tree classifier to mark the presence of spindles. In our previous work [16] we also used Teager energy of the signal as a feature together with spectral edge frequency to score spindles in a two stage algorithm.

Most of the methods cited above achieved good performance with sensitivity of at least 70%. However direct comparison between them is difficult because of the varying dataset and performance metrics used in each. To avoid this problem, [10] proposed a set of performance metrics and also published a spindle database online to allow fair comparison of methods assessed using the same set of signals and metrics.

In this paper we introduce the use of *line length* [17] as a potential feature for automatic spindle detection and evaluate its detection performance with a simple algorithm. We also implement this low-complexity algorithm on a resource-constrained hardware to demonstrate its applicability within a wearable setting. In the following section we briefly describe the spindle dataset used in this study and define line length. We then explain the spindle detection algorithm used with line length as the only feature and describe its hardware implementation on a MSP430 microcontroller. Section III assesses the spindle detection performance of the algorithm and its overall power consumption. Section IV discusses these results and compares the detection ability of our algorithm against other methods that have reported performance on the same dataset.

II. MATERIAL & METHODS

A. Database

EEG signals from the DREAMS Sleep Spindles database available online courtesy University of MONS - TCTS Labo-

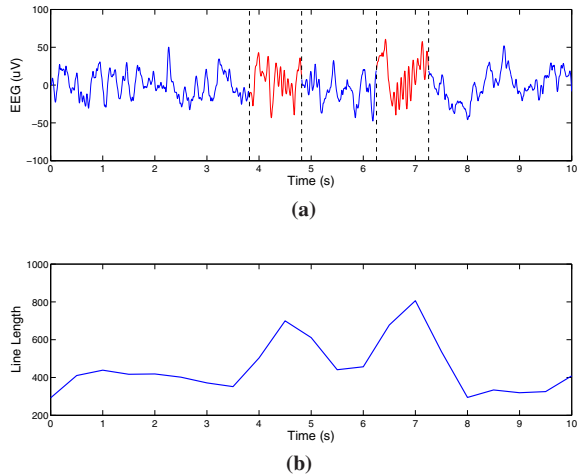


Fig. 1: (a) EEG signal with two sleep spindles marked between vertical lines; (b) Line length of the EEG signal showing higher values during spindle occurrence

ratory and Universite Libre de Bruxelles - CHU de Charleroi Sleep Laboratory [18] were used to test the algorithm. This is the recommended dataset to use to be able to fairly compare different methods [10]. It consists of 30-minute recording each from six subjects including 3 males and 3 females. The spindles have been marked visually by two scorers and the union of their markings is taken as the reference set. All the subjects have channel CZ-A1 annotated with data sampled at 200 Hz except subjects 1 and 3, for whom the annotated channel is C3-A1 and sampling frequencies are 100 Hz and 50 Hz respectively.

B. Line length

Line length was introduced as a low-complexity feature for seizure onset detection by [17]. It is the sum of absolute differences between subsequent samples and is defined by the following equation:

$$LL = \sum_{n=1}^N |x(n-1) - x(n)| \quad (1)$$

where LL is the line length, x is the input signal and N is the number of samples in the signal (or a block of signal under analysis).

Fig. 1a shows an EEG signal with two sleep spindles. The line length corresponding to this signal, calculated in blocks of 1 s (with 50% overlap) is shown in Fig. 1b. It can be seen on the figure that the occurrence of a spindle in the original signal leads to a rise in the line length. This lasts approximately until the end of spindle duration returns to a lower level thereafter. This effect of line length having higher values during spindle occurrence will be used in an algorithm to test its utility for detecting spindles.

C. Algorithm

A block diagram of the algorithm with line length as the analysis feature is shown in Fig. 2. EEG signal for a single channel is used as input to the algorithm. At the first

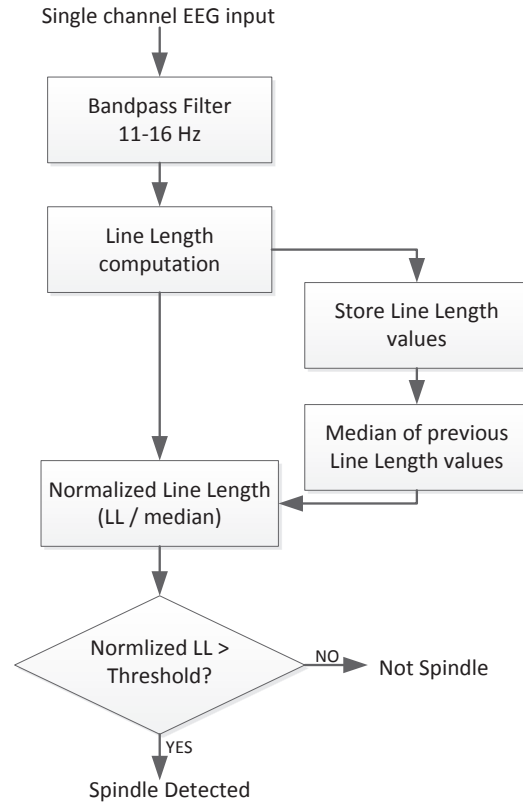


Fig. 2: Block diagram of sleep spindle detection algorithm

stage, a second order Butterworth bandpass filter is applied to the input signal to limit it in the 11-16 Hz frequency range, which is the spindle range of interest. The filtered signal is then partitioned into blocks of 1 s epochs with 50% overlap between subsequent epochs. The line length for each epoch is calculated using (1). It is normalized by a factor which is obtained by taking the median line length value of the last 80 epochs. This step obviates the need of any patient-specific adjustments. The number of epochs to use in the computation of the median was determined empirically by trying out various values. Initially when the number of epochs processed is less than 80, the median of all available previous values is taken. This normalized value is then compared against a detection threshold K . If the value is found to be greater than K , the epoch is marked as spindle. The detection threshold also controls the sensitivity of the algorithm, with higher values resulting in stricter classification criterion.

D. Hardware implementation

The algorithm was implemented on Texas Instruments MSP430F5438A microcontroller [19] to measure its online performance and power consumption. All arithmetic operations were performed in fixed point arithmetic to make use of the hardware multiplier on the chip and the micro-

controller was put to idle mode (LPM3) whenever there was no data to process. The coefficients of the bandpass filter were represented as fixed point numbers in Q15 format and the whole filtering operation required only three 16-bit multiplication and addition operations. The line length was calculated by taking the absolute difference between each new and previous filtered sample and accumulating the result in a register which is initialized to zero at the start of an epoch. This way it is updated with each new sample without the need to store all previous samples in an epoch. The median was then computed using a linked list method based on [20] that has a sorting complexity of N . The resulting value is used to normalize the line length. Finally, a detection flag is raised and time noted whenever the normalized line length is found to be greater than the threshold K .

III. RESULTS

A. Performance metrics

The algorithm performance was characterised using the following metrics, where *sensitivity* represents the fraction of spindles detected while *specificity* refers to the fraction of background EEG rejected as non-spindle.

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

A detection by the algorithm is considered to be a True Positive (TP) if there is at least an overlap of 0.5 s between the reference spindle (visually scored) and the detected one. A False Positive (FP) is scored when there is no such overlap or the overlap is smaller than 0.5 s. A False Negative (FN) represents spindles which have been scored visually but are not detected by the algorithm and True Negative (TN) counts the number of events successfully rejected as non-spindle. The number of TN was estimated using (4), as suggested in [10], where average detected spindle duration was calculated separately for each subject.

$$TN = \frac{\text{Total record duration}}{\text{Avg. detected spindle duration}} - TP - FP - FN \quad (4)$$

We used leave-one-out cross validation (LOOCV) to determine the detection performance of the algorithm on each subject while training it on the five remaining subjects. A receiver operating characteristic (ROC) curve was plotted for average sensitivity against specificity by varying the detection threshold K in small steps using five training subjects. The best performing threshold was determined from the curve as the point that maximised both sensitivity and specificity. This threshold was then used to test the performance of the remaining sixth subject. This procedure was carried out six times such that each subject becomes the test case once with other five being used for training.

The total number of spindles visually scored and those detected by the algorithm for each subject are shown in

Table I. The average sensitivity of the algorithm is 83.6% and its 95% confidence interval range is between 80.5% and 86.8%. The sensitivity results are consistently greater than 80% for all cases except subject 4 where less than half of reference spindles are automatically detected. In this case, the sensitivity is low because of the high proportion of Wake stages and artefacts present in the signal making detection difficult. Of the 63 spindles visually scored in subject 4, half of them are marked in Wake stage (according to the corresponding hypnogram). We have used all signals from the dataset *as is* in our analysis, including those with artefacts and noisy segments. The addition of any artefact removal preprocessing stage is likely to help improve the performance in such cases. The specificity in all the subjects is close to the overall average of 87.9% which shows that most of the non-spindle data has been successfully rejected.

TABLE I: Spindle detection algorithm performance

Subject	Total Spindles	True Pos.	Sens. (%)	Spec. (%)
1	134	121	90.3	83.4
2	77	62	80.5	94.0
3	44	44	100	85.8
4	63	28	44.4	88.1
5	103	87	84.5	88.2
6	117	108	92.3	86.5
All	538	450	83.6	87.9

A breakdown of the spindles detected in each sleep stage is shown in Table II. Of the 841 spindles detected by the algorithm (SS_{det}) 61.5% true positives were in N2, 58.2% in N2 and N3 combined and 53.5% across all sleep stages. The proportion of false positives was highest in subjects 3 and 4. From the total detections made by the algorithm, 90% were in N2 and N3 combined and less than 6% were recorded in the Wake stage.

TABLE II: Sleep Spindles (SS) detected in each sleep stage

Sub	SS_{det}	SS_{Wake}	SS_{N1}	SS_{N2}	SS_{N3}	SS_{REM}
1	202	13	1	141	47	0
2	99	0	2	84	13	0
3	126	11	11	95	9	0
4	87	16	23	42	6	0
5	154	7	0	103	44	0
6	173	0	1	114	58	0
Total	841	47	38	579	177	0

B. Power consumption

The epoch size used for processing is 1 s and there are two epochs being processed during this time because of the overlap between subsequent epochs. The microcontroller was found to be active for only 10% of the time to perform all of the signal processing needed for the algorithm while spending the rest of the time in idle mode until a new sample arrives. Operating at a clock frequency of 1 MHz and supply voltage of 1.8 V the power consumption for the algorithm was found to be 56.7 μ W with one channel of EEG input.

IV. DISCUSSION

We compared the performance of this algorithm against other sleep detection methods that report their results on the same database (shown in Table III). In contrast to [10], our algorithm detected a higher number of spindles overall but also resulted in a higher proportion of false positives resulting in a lower specificity. Individual subject-specific performance was not reported in [10]. In another method, [13] reported a sensitivity of 75.1% after normal modelling however they were not able to detect any spindles in subject 4 and averaging the results including this case reduces their method's sensitivity to 62.6% only. However, they obtained specificity of more than 90% for other test subjects. Further, in our previous spindle detection algorithm [16] we reported sensitivity and specificity as 80.3% and 97.6% respectively. Overall the results using the algorithm based on line length in this work shows a higher sensitivity but a slightly reduced specificity. This reduction is because the average duration of detected spindles is higher in this work. Compared to our previous algorithm where the average duration was approximately 1 s it is now 2.7 s. This reduces the number of estimated true negatives (TN) in equation (4) and consequently specificity. However, using only one feature with a simple algorithm without any artefact rejection still demonstrates line length as a very useful feature for spindle detection. Future work involving a better classifier or coupling line length with other low complexity features can lead to further improvement in results.

TABLE III: Comparison of this work with other algorithms

Method	Sens. (%)	Spec. (%)
[10]	70.2	98.6
[13]	75.1	96.7
[16]	80.3	97.6
This work	83.6	87.9

We also implemented the line length based algorithm on a MSP430 microcontroller to measure its power consumption and determine its computational load. To the best of our knowledge, this is the very first hardware implementation of an automated spindle detection method. The overall system power consumption was 56.7 μ W with only 10% of microcontroller active time operating at clock frequency of 1 MHz. This shows that line length is not only useful for getting a good spindle detection performance but also a very efficient feature for use in resource-constrained wearable systems.

V. CONCLUSION

This paper presents an introductory evaluation of using line length as a discriminatory feature for automatic detection of sleep spindles. It is incorporated in a simple algorithm as a proof of concept to detect spindles in a test database with 538 spindles. It results in sensitivity and specificity of 84% and 88% respectively. This paper also presents the very first implementation of a spindle detection algorithm on a microcontroller where overall power consumption of the algorithm is only 57 μ W. Based on these results, line length

appears to be valuable feature with very low computational complexity for the detection of spindles and can also be used with other features for improved performance in existing methods.

REFERENCES

- [1] C. Iber, S. Ancoli-Israel, A. Chesson, and S. Quan, Eds., *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. Westchester, IL: American Academy of Sleep Medicine, 2007.
- [2] L. D. Gennaro and M. Ferrara, "Sleep spindles: an overview," *Sleep Med. Rev.*, vol. 7, no. 5, pp. 423–440, 2003.
- [3] M. Schabus, K. Hödlmoser, G. Gruber, C. Sauter, P. Anderer, G. Klisch, S. Parapatics, B. Saletu, W. Klimesch, and J. Zeitlhofer, "Sleep spindle-related activity in the human eeg and its relation to general cognitive and learning abilities," *Eur. J. Neurosci.*, vol. 23, no. 7, pp. 1738–1746, 2006.
- [4] R. Nader and C. Smith, "The relationship between stage 2 sleep spindles and intelligence," *Sleep*, vol. 24, no. 1, p. A160, 2001.
- [5] S. M. Fogel and C. T. Smith, "The function of the sleep spindle: A physiological index of intelligence and a mechanism for sleep-dependent memory consolidation," *Neurosci. Biobehav. R.*, vol. 35, no. 5, pp. 1154–1165, 2011.
- [6] E. Huupponen, G. Gómez-Herrero, A. Saastamoinen, J. Vrri, A. Hasan, and S. Himanen, "Development and comparison of four sleep spindle detection methods," *Artif. Intell. Med.*, vol. 40, no. 3, pp. 157–170, 2007.
- [7] K. Campbell, A. Kumar, and W. Hofman, "Human and automatic validation of a phase-locked loop spindle detection system," *Electroencephalogr. and Clin. Neurophysiol.*, vol. 48, no. 5, pp. 602–5, 1980.
- [8] D. Gorur, U. Halici, H. Aydin, G. Ongun, F. Ozgen, and K. Leblebicioglu, "Sleep spindles detection using short time fourier transform and neural networks," in *IEEE IJCNN*, Honolulu, May 2002.
- [9] N. Acir and C. Güzeliz, "Automatic recognition of sleep spindles in eeg by using artificial neural networks," *Expert Sys. Appl.*, vol. 27, no. 3, pp. 451–458, 2004.
- [10] S. Devuyt, T. Dutoit, P. Stenuit, and M. Kerkhofs, "Automatic sleep spindles detection overview and development of a standard proposal assessment method," in *IEEE EMBC*, Boston, September 2011.
- [11] B. Babadi, S. M. McKinney, V. Tarokh, and J. M. Ellenbogen, "Diba: a data-driven bayesian algorithm for sleep spindle detection," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 2, pp. 483–93, 2012.
- [12] L. Causa, C. M. Held, J. Causa, P. A. Estévez, C. A. Perez, R. Chamorro, M. Garrido, C. Algarín, and P. Peirano, "Automated sleep-spindle detection in healthy children polysomnograms," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 9, pp. 2135–46, 2010.
- [13] A. Nonclercq, C. Urbain, D. Verheulpen, C. Decaestecker, P. Van Bogaert, and P. Peigneux, "Sleep spindle detection through amplitude-frequency normal modelling," *J. Neurosci. Methods*, vol. 214, no. 2, pp. 192–203, 2013.
- [14] S. V. Schönwald, E. L. de Santa-Helena, R. Rossatto, M. L. Chaves, and G. J. Gerhardt, "Benchmarking matching pursuit to find sleep spindles," *J. Neurosci. Methods*, vol. 156, no. 1-2, pp. 314–321, 2006.
- [15] F. Duman, A. Erdamar, O. Erogul, Z. Telatar, and S. Yetkin, "Efficient sleep spindle detection algorithm with decision tree," *Expert Sys. Appl.*, vol. 36, no. 6, pp. 9980–9985, 2009.
- [16] S. Imtiaz, S. Saremi-Yarahmadi, and E. Rodriguez-Villegas, "Automatic detection of sleep spindles using teager energy and spectral edge frequency," in *IEEE BioCAS*, Rotterdam, October 2013.
- [17] R. Esteller, J. Echaz, T. Tchong, B. Litt, and B. Pless, "Line length: an efficient feature for seizure onset detection," in *IEEE EMBC*, Istanbul, October 2001.
- [18] University of MONS - TCTS Laboratory. (2013) The DREAMS Sleep Spindles Database. [Online]. Available: <http://www.tcts.fpm.ac.be/~devuyt/Databases/DatabaseSpindles/>.
- [19] Texas Instruments. (2013) MSP430 ultra-low power 16-bit microcontrollers. [Online]. Available: <http://www.ti.com/msp430/>.
- [20] P. Ekstrom. (2000) Running median better than average. [Online]. Available: http://www.eetindia.co.in/STATIC/PDF/200011/EIOL_2000NOV03_EMS_EDA.TA.pdf