35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Automatic SLEEP staging: From young adults to elderly patients using multi-class support vector machine

Jacob Kempfner [a,d], Poul Jennum [b,c,d], Helge B. D. Sorensen [a], Julie A. E. Christensen [a] and Miki Nikolic [c]

Abstract— **Aging is a process that is inevitable, and makes our body vulnerable to age-related diseases. Age is the most consistent factor affecting the sleep structure. Therefore, new automatic sleep staging methods, to be used in both of young and elderly patients, are needed. This study proposes an automatic sleep stage detector, which can separate wakefulness, rapid-eye-movement (REM) sleep and non-REM (NREM) sleep using only EEG and EOG. Most sleep events, which define the sleep stages, are reduced with age. This is addressed by focusing on the amplitude of the clinical EEG bands, and not the affected sleep events. The age-related influences are then reduced by robust subject-specific scaling. The classification of the three sleep stages are achieved by a multi-class support vector machine using the one-versus-rest scheme. It was possible to obtain a high classification accuracy of 0.91. Validation of the sleep stage detector in other sleep disorders, such as apnea and narcolepsy, should be considered in future work.**

I. Introduction

The temporal structure of sleep is described by dividing the sleep into so-called sleep stages, in this case wakefulness, NREM and REM sleep. This is based on different sleep events as described in [1]. Age is the most consistent factor affecting the sleep architecture. The amount of time spent in deep-sleep, which is described as NREM3 sleep, decreases with age, and more wakefulness is also observed. Additionally, the amount of time spent in REM sleep also tends to decrease in elderly when comparing with young adults [2]. Major changes in the EEG during NREM sleep in elderly are well known. The amplitude of the delta-band, which includes slow wave activity, is lowered in elderly. Additionally, the numbers of sleep spindles and K-complexes become fewer compared to young adults [3], and the amplitude of the sleep EEG in elderly is suggested reduced [4]. Elderly also tend to have increased EMG activity in form of periodic limb movements and restless legs during NREM sleep [5], [6], and sleep related respiratory irregularities are also more prevalent in elderly [7], [8]. The volume of short-lasting awakenings during NREM sleep, also known as arousals, increases with age, but is stable across age in REM sleep [9]. Few changes during REM sleep have been reported. The occurrence of REMs during REM sleep is substantially reduced in the elderly [10] and increased EMG activity has also been reported [11]. These changes are a challenge for automatic sleep staging.

Automatic sleep-staging methods are requested in the sleep-clinics, since manual scoring is time-consuming and it introduces inter-rater variability. This study is based on our previous study [12], where only the REM sleep stage in elderly has been detected. This method has been extended, such that is capable of detecting REM sleep, NREM sleep and wakefulness in young and elderly patients. Furthermore, the data used in this study has been increased by including other sleep disorders and young healthy adults.

II. Data

A. Subjects and demographics

A total of forty subjects from the Danish Center for Sleep Medicine, Department of Clinical Neurophysiology, Glostrup University Hospital, Denmark, were enrolled. Their diagnosis and demographics are summarized in Table 1. The young and elderly groups are both healthy and have no known sleep disorders. Subjects with periodic limb movements disorder (PLM) and idiopathic REM sleep behavior disorder (iRBD) are experiencing increased amount movements during sleep [13].

Table 1: Demographics

Group	№ (♀, ♂)	Age μ ± σ (min, max, median) *years* [A]
Young	10 (6, 4)	28.6 ± 5.2 (25, 41, 26)
Elderly	10 (8, 2)	55.2 ± 8.2 (45, 73, 53)
PLM	10 (5, 5)	58.6 ± 10.3 (43, 75, 59)
iRBD	10 (2, 8)	61.7 ± 8.0 (51, 76, 61)

[A] μ and σ corresponds to the mean and standard deviation.

B. Data acquisition

All subjects underwent one full night polysomnography in accordance with American Academy of Sleep Medicine [1]. Approximately seven to eight hours of sleep per subject were available when using data from light-off to light-on. Only the left and right EOG channels combined with the F_3-A_2, C_3-A_2 and O_1-A_2 EEG channels were used. A_2 denotes the right mastoid. All recordings were visually inspected by experienced specialists to ensure their quality. The sampling frequency of the analyzed sleep data was 256 Hz. The proposed algorithm was programed in MATLAB.

[a] Department of Electrical Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark.
[b] Danish Center for Sleep Medicine, Glostrup University Hospital, Denmark.
[c] Department of Clinical Neurophysiology, Glostrup University Hospital, Denmark.
[d] Center for Healthy Ageing, University of Copenhagen, Copenhagen, Denmark.

III. METHOD

A. Preprocessing of the EEG and EOG

In [12], [14] a fourth-order Butterworth (BW) bandpass filter with 3 dB cutoff frequencies at 1 and 5 Hz, respectively, was successfully used to separate the REM from slow-eye-movement (SEM), baseline drift and EMG activity. However, in this study a total of eight bands, including the abovementioned, were used to separate REMs from everything else. One band could not separate SEMs from REMs properly in our new increased data set. The number of bands, eight in this case, was estimated by trial-and-error. The left and right EOG channels were both band-pass filtered with fourth-order BW bandpass filters according to [15]. Furthermore, each EEG channel (F_3-A_2, C_3-A_2, O_1-A_2) was filtered into the traditional five clinical bands using fourth-order BW bandpass filters [15]. Finally, the power-line noise was reduced by a fourth-order BW notch filter with the cutoff frequencies (3 dB) 48 Hz and 52 Hz, respectively. The frequency bands are defined in Table 2.

Table 2: Cutoff frequencies

Channels	Band	Type	Low [Hz]	High [Hz]
EOG	ψ_1	bandpass	0.25	5.00
EOG	ψ_2	bandpass	0.50	5.00
EOG	ψ_3	bandpass	0.75	5.00
EOG	ψ_4	bandpass	1.00	5.00
EOG	ψ_5	bandpass	1.25	5.00
EOG	ψ_6	bandpass	1.50	5.00
EOG	ψ_7	bandpass	1.75	5.00
EOG	ψ_8	bandpass	2.00	5.00
EEG	δ	bandpass	1.00	4.00
EEG	θ	bandpass	4.00	8.00
EEG	α	bandpass	8.00	13.0
EEG	β	bandpass	13.0	30.0
EEG	γ	bandpass	30.0	65.0
EEG		bandstop	48.0	52.0

B. Partition of data

The preprocessed sleep data were then partitioned into 3-second mini-epochs, which is widely used in the sleep community [16]. This was achieved by using a sliding window, as illustrated in Fig. 1.

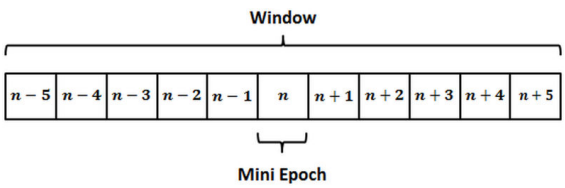

Fig. 1: The analysis window consists of 11 mini-epochs, corresponding to 33 seconds [12].

Ten mini-epochs surrounding the center mini-epoch (n) were included in the reference window. The total reference window duration was therefore 11 mini-epochs, which corresponds to 33 seconds, and the step size of the sliding window was 1 mini-epoch (3 seconds). Furthermore, the partitioned signals were expanded at the beginning and ending by repeating the 5 first mini-epochs at the beginning and the 5 last mini-epochs at the ending.

C. EOG features

Sideway eye movements were described by the normalized covariance values of the pre-processed and partitioned EOG channels. The normalized covariance is given by:

$$r_k(n) = \frac{\sigma_{ab}(n)}{\sqrt{\sigma_{aa}(n)\,\sigma_{bb}(n)}} \tag{1}$$

where σ corresponds to the covariance and $a = \psi_k^{\mathrm{L}}$ and $b = \psi_k^{\mathrm{R}}$. The ψ_k^{L} and ψ_k^{R} correspond to the left and right EOG analysis window for the bands $k = 1,2,\ldots,8$ (Table 2). It is assumed that REMs in REM sleep, SEMs in NREM1 sleep and eye movements in wakefulness will yield a negative correlation, whereas background EOG would be uncorrelated

D. EEG features

Movement disorders, especially iRBD, may have increased muscle noise in the EEG channels. This is addressed by using the median as a robust amplitude measure:

$$m_\delta(n) = \text{median}\,|\delta(n)| \tag{2a}$$

$$m_\theta(n) = \text{median}\,|\theta(n)| \tag{2b}$$

$$m_\alpha(n) = \text{median}\,|\alpha(n)| \tag{2c}$$

$$m_\beta(n) = \text{median}\,|\beta(n)| \tag{2d}$$

$$m_\gamma(n) = \text{median}\,|\gamma(n)| \tag{2e}$$

In (2) the $\delta(n)$, $\theta(n)$, $\alpha(n)$, $\beta(n)$ and $\gamma(n)$ correspond to the clinical EEG bands, which is defined in Table 2, of the window illustrated in Fig. 1 at mini-epoch index *n*.

E. Feature scaling

The age-related influence is addressed by rescaling each feature into the range of approximately 0 to 1. In this study two modified min-max methods were used, one for the EOG and another for the EEG.

1) EOG feature scaling

The individual EOG features were rescaled by:

$$\mathrm{R}_k = \frac{\mathrm{r}_k - \min(\mathrm{r}_k)}{\max(\mathrm{r}_k) - \min(\mathrm{r}_k)} \tag{3}$$

where $k = 1,2,\ldots,8$, and r_k corresponds to the N normalized covariance values computed by (1) for a given subject and band k. The original min-max scaling method defined in (3) is not robust towards outliers. The minima and maxima were therefore estimated by:

$$\min(\mathrm{r}_k) \stackrel{\text{def}}{=} \text{median}\,(\mathrm{r}_k^-) \tag{4a}$$

$$\max(r_k) \stackrel{\text{def}}{=} \text{median}(r_k^+) \quad (4b)$$

where r_k^- corresponds to all the normalized covariance values between $-1.00 \le r_k^- < -0.25$, whereas r_k^+ corresponds to the all the normalized covariance values between $-0.25 \le r_k^+ \le +1.00$ for a given subject and band k [12].

2) EEG feature scaling

Furthermore, the EEG features were rescaled according to:

$$M_k = \frac{m_k - \min(m_k)}{\max(m_k) - \min(m_k)} \quad (5)$$

where $k = \delta, \theta, \alpha, \beta, \gamma$, and m_k corresponds to the N median values computed by (2) for a given subject and band k. The minima and maxima were estimated by:

$$\min(m_k) \stackrel{\text{def}}{=} P_{25}(m_k) \quad (6a)$$

$$\max(m_k) \stackrel{\text{def}}{=} P_{75}(m_k) \quad (6b)$$

where P_{25} and P_{75} correspond to the 25th and 75th percentile, respectively [12].

F. Feature merging

The scaled features were then merged into an [N x 23] feature matrix **F** defined in (8), where N corresponds to the number of feature samples for a given subject. The feature matrix is therefore given by:

$$\mathbf{F_{EOG}} = [R_1, R_2, R_3, R_4 R_5, R_6, R_7, R_8] \quad (7a)$$

$$\mathbf{F_{EEG}^{F_3}} = \left[M_\delta^{F_3}, M_\theta^{F_3}, M_\alpha^{F_3}, M_\beta^{F_3}, M_\gamma^{F_3}\right] \quad (7b)$$

$$\mathbf{F_{EEG}^{C_3}} = \left[M_\delta^{C_3}, M_\theta^{C_3}, M_\alpha^{C_3}, M_\beta^{C_3}, M_\gamma^{C_3}\right] \quad (7c)$$

$$\mathbf{F_{EEG}^{O_1}} = \left[M_\delta^{O_1}, M_\theta^{O_1}, M_\alpha^{O_1}, M_\beta^{O_1}, M_\gamma^{O_1}\right] \quad (7d)$$

where **F** is given by:

$$\mathbf{F} = \left[\mathbf{F_{EOG}}, \mathbf{F_{EEG}^{F_3}}, \mathbf{F_{EEG}^{C_3}}, \mathbf{F_{EEG}^{O_1}}\right] \quad (8)$$

G. Classification

In this study the multi-class support vector machine (MC-SVM), using the one-vs-rest approach, combined with the Radial Basic Function (RBF) as kernel, was used. Three SVM classifiers, which can separate each class from the rest, were constructed. They are denoted SVM_W (wake vs. rest), SVM_R (REM vs. rest) and SVM_N (NREM vs. rest). New objects are assigned to the class that has a positive vote and the largest distance to its hyperplane [17–19]. The optimization of each classifier was obtained by the *k*-fold stratified cross-validation scheme, where *k* was chosen as 5 in this study. The manual hypnogram was modified into a target vector by first labeling, e.g. the REM sleep epochs, '+1' and everything else '-1'. The target vector was then extended by successfully repeating each epoch a further nine times. This will increases the "sampling rate" from one score per epoch (30-seconds) to 10 scores per epoch, which is the same as one score per mini-epoch (3-seconds). The target vector of wakefulness and NREM was made similarly, yielding a total of three target vectors, one for each SVM.

H. Post-processing of the MC-SVM classifier

Normally, NREM and REM sleep tend to alternate cyclically through the night [20]. This trend was exploited by post-processing the MC-SVM output. This was achieved by converting the estimated MC-SVM output into smoothened "posterior-like probabilities", as illustrated in Fig. 2.

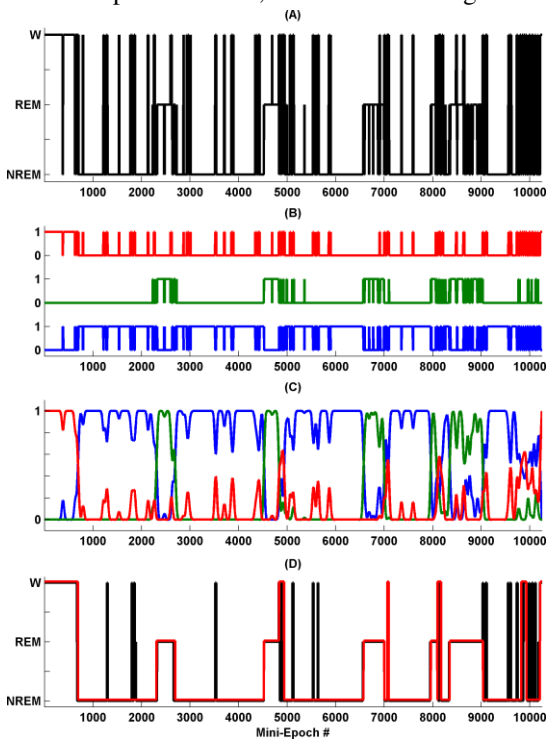

Fig. 2: Post-processing of the MC-SVM. Plot (A) corresponds to the raw MC-SVM output, while plot (B) shows the three converted binary outputs, where red, green and blue corresponds to W, REM and NREM, respectively. Plot (C) illustrates the smoothened "posterior-like probabilities", while plot (D) shows to the final prediction, where black and red corresponds to the true hypnogram and the estimated hypnogram, respectively.

The MC-SVM output consists of three classes (Fig. 2A), which was converted into three binary outputs as illustrated in Fig. 2B. In order to eliminate short-term transients, the three estimated binary outputs were individually smoothened by a normalized Blackman window (Fig. 2C) according to [15]. The class with the highest "posterior-like probability" in each mini-epoch was then the final outcome of prediction (Fig. 2D). The SVM parameters, which obtained the highest accuracy (agreement), are summarized in Table 3. The Blackman window duration was found best at 97.

Table 3: SVM parameters

Parameters	SVM_W	SVM_R	SVM_N
C	2^5	2^{-2}	2^6
γ	2^{-3}	2^{-2}	2^{-3}

where C and γ are the SVM and RBF kernel paramters, respectively.

IV. Results and Discussion

The results are summarized in the confusion matrix in Table 4. The proposed sleep stage classification algorithm yields an overall accuracy of 0.91 in young healthy subjects and elderly patients. This is assumed to be acceptable when comparing with other promising sleep staging studies, where the accuracy range from 0.87 to 0.95 [21–23]. The number of enrolled subjects in those studies range from five to 39, mostly young healthy subjects. A sleep stage study of 28 healthy elderly subjects and Parkinson's disease patients (elderly) has been addressed in [24], where the obtained accuracy was 0.88 and 0.68 in the healthy and diseased group, respectively. Notice, some of those studies are also detecting the sub-classes of NREM sleep.

Table 4: Confusion matrix (mini-epochs)

		Manual		
		NREM	**REM**	**W**
Automatic	**NREM**	238,661	7,907	12,187
	REM	7,053	62,690	1,445
	W	6,496	423	45,778
	Sensitivity	0.95	0.88	0.77
	Specificity	0.85	0.97	0.98

The individual group accuracies of the young, elderly, PLM and iRBD are 0.91, 0.91, 0.90 and 0.91, respectively. The error primarily occurs in the transition regions, or when sleep is highly fragmented. Especially short periods of wakefulness gets misclassified, due to the smoothing effect of the post-processing scheme (Fig. 2D). The above-mentioned evaluation is based on mini-epochs (3-second), while the original hypnogram consists of epochs (30-seconds). The influence of this was tested by converting the mini-epochs back to epochs. Each epoch was classified by the majority vote between its 10 mini-epochs in the final prediction. This, however, did not have any influence on the performance, due to the size of the Blackman window D=97 (mini-epochs), which corresponds to a duration of 291 seconds.

V. Conclusion

The proposed algorithm is capable of estimating the sleep stages wakefulness, NREM and REM in young healthy subjects and elderly patients automatically, and is not affected by age related changes. It was possible to obtain a high overall accuracy of 0.91. This is comparable with other sleep stage studies, even those using only young healthy adults. Inclusions of other sleep disorders, such as apnea and narcolepsy, should be considered in future work.

References

[1] C. Iber, S. Ancoli-Israel, A. Chesson, and S. Quan, *The AASM manual for the scoring of sleep and associated events: rules, terminology, and technical specification*, 1st ed. Westchester, IL: American Academy of Sleep Medicine, 2007.

[2] P. Prinz, E. Peskind, P. Vitaliano, M. Raskind, C. Eisdorfer, N. Zemcuznikov, and C. Gerber, "Changes in the sleep and waking EEGs of nondemented and demented elderly subjects," *J Am Geriatr Soc*, vol. 86, no. 93, pp. 86–93, 1982.

[3] K. Crowley, J. Trinder, Y. Kim, M. Carrington, and I. Colrain, "The effects of normal aging on sleep spindle and K-complex production," *Clin Neurophysiol*, vol. 113, no. 10, pp. 1615–1622, 2002.

[4] M. Guazzelli, I. Feinberg, M. Aminoff, G. Fein, T. Floyd, and C. Maggini, "Sleep spindles in normal elderly: comparison with young adult patterns and relation to nocturnal awakening, cognitive function and brain atrophy," *Electroencephalogr Clin Neurophysiol*, vol. 63, no. 6, pp. 526–539, 1986.

[5] R. Coleman, L. Miles, C. Guilleminault, J. Zarcone VP, J. van den Hoed, and W. Dement, "Sleep-wake disorders in the elderly: polysomnographic analysis," *J Am Geriatr Soc*, vol. 29, no. 7, pp. 289–296, 1981.

[6] S. Youngstedt, D. Kripke, M. Klauber, R. Sepulveda, and W. Mason, "Periodic leg movements during sleep and sleep disturbances in elders," *J Gerontol A Biol Sci Med Sci*, vol. 53, no. 5, pp. 391–394, 1998.

[7] S. Ancoli-Israel, D. Kripke, W. Mason, and S. Messin, "Sleep apnea and nocturnal myoclonus in a senior population," *Sleep*, vol. 4, no. 4, pp. 349–358, 1981.

[8] M. Carskadon and W. Dement, "Respiration during sleep in the aged human," *J Gerontol A Biol Sci Med Sci*, vol. 36, no. 4, pp. 420–423, 1981.

[9] M. Boselli, L. Parrino, A. Smerieri, and M. Terzano, "Effect of age on EEG arousals in normal sleep," *Sleep*, vol. 21, no. 4, pp. 351–357, 1998.

[10] N. Darchia, I. Campbell, and I. Feinberg, "Rapid eye movement density is reduced in the normal elderly," *Sleep*, vol. 26, no. 8, pp. 973–977, 2003.

[11] D. Bliwise, L. He, F. Ansari, and D. Rye, "Quantification of electromyographic activity during sleep: a phasic electromyographic metric," *J Clin Neurophysiol.*, vol. 23, no. 1, pp. 59–67, 2006.

[12] J. Kempfner, P. Jennum, M. Nikolic, and H. Sorensen, "Automatic detection of REM sleep in recordings with muscle noise present in patients with idiopathic REM sleep behavioral disorder," *Submitted to Computer Methods and Programs in Biomedicine*.

[13] American Academy of Sleep Medicine, *International classification of sleep disorders: diagnostic and coding manual*, 2nd ed. Westchester, Illinois: AASM, 2005, pp. 148–152.

[14] R. Agarwal, T. Takeuchi, S. Laroche, and J. Gotman, "Detection of rapid-eye movements in sleep studies," *IEEE Transactions on Signal Processing Biomedical Engineering*, vol. 52, no. 8, pp. 1390–1396, 2005.

[15] F. Gustafsson, "Determining the initial states in forward-backward filtering," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 988–992, 1996.

[16] C. Iber, S. Ancoli-Israel, A. Chesson, and S. Quan, *The AASM Manual for the scoring of sleep and associated events: rules, terminology and technical specification*. Westchester, Illinois: AASM, 2007.

[17] A. Statnikov, C. Aliferis, D. Hardin, and I. Guyon, *A Gentle Introduction to Support Vector Macnine in Biomedicine*, 1 st. 2011, pp. 91–96.

[18] C. Chang and C. Lin, "LIBSVM : A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, 2011.

[19] C. Hsu, C. Chang, and C. Lin, "A practical guide to support vector classification," 2010. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/.

[20] M. Kryger, T. Roth, and W. Dement, *Principles and Practice of SLEEP MEDICINE*, 4th ed. Elsevier Inc., 2005, pp. 13–23.

[21] S. Liang, C. Kuo, Y. Hu, and Y. Cheng, "A rule-based automatic sleep staging method," in *Conf Proc IEEE Eng Med Biol Soc*, 2011, pp. 6067–6070.

[22] U. Acharya, E. Chua, K. Chua, L. Min, and T. Tamura, "Analysis and automatic identification of sleep stages using higher order spectra," *International Journal of Neural Systems*, vol. 20, no. 6, pp. 509–521, 2010.

[23] R. Sinha, "Artificial neural network and wavelet based automated detection of sleep spindles, REM sleep and wake states," *Journal of Medical Sytems*, vol. 32, no. 4, pp. 291–309, 2008.

[24] P. Jensen, H. Sorensen, H. Leonthin, and P. Jennum, "Automatic sleep scoring in normals and in individuals with neurodegenerative disorders according to the new internationale sleep scoring criteria," *J Clin Neurophysiol*, vol. 27, no. 4, pp. 296–302, 2010.