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

Classification of iRBD and Parkinson's Disease patients based on eye movements during sleep

Julie A. E. Christensen[a,b,c], Henriette Koch[a], Rune Frandsen[c], Jacob Kempfner[a,c], Lars Arvastson[b], Soren R. Christensen[b], Helge B. D. Sorensen[a] and Poul Jennum[c]

Abstract— **Patients suffering from the sleep disorder idiopathic rapid-eye-movement sleep behavior disorder (iRBD) have been observed to be in high risk of developing Parkinson's disease (PD). This makes it essential to analyze them in the search for PD biomarkers. This study aims at classifying patients suffering from iRBD or PD based on features reflecting eye movements (EMs) during sleep. A Latent Dirichlet Allocation (LDA) topic model was developed based on features extracted from two electrooculographic (EOG) signals measured as parts in full night polysomnographic (PSG) recordings from ten control subjects. The trained model was tested on ten other control subjects, ten iRBD patients and ten PD patients, obtaining a EM topic mixture diagram for each subject in the test dataset. Three features were extracted from the topic mixture diagrams, reflecting "certainty","fragmentation" and "stability" in the timely distribution of the EM topics. Using a Naive Bayes (NB) classifier and the features "certainty" and "stability" yielded the best classification result and the subjects were classified with a sensitivity of 95 %, a specificity of 80 % and an accuracy of 90 %. This study demonstrates in a data-driven approach, that iRBD and PD patients may exhibit abnorm form and/or timely distribution of EMs during sleep.**

I. INTRODUCTION

Patients suffering from the sleep disorder idiopathic rapid-eye-movement sleep behavior disorder (iRBD) are at high risk of developing Parkinson's disease (PD) [1]. In consequence, some studies focus on sleep data in the search for PD biomarkers, where polysomnographic (PSG) data are analyzed either manually or automatic [2] [3]. Supportively, many different attempts to automatic score sleep stages, both in control subjects as well as in sleep disorder patients, have been developed [4] [5]. In [4], a data-driven method was developed, where a topic model with five topics were conducted for each subject based on their sleep electroencephalography (EEG). The method was subject-specific, as it was aimed at providing a complementary approach to sleep analysis by presenting each sleep epoch as a mixture of stages. This study raised the idea of developing a data-driven topic model with the aim of using it to analyze and automatic classify control subjects and patients suffering from either iRBD or PD.

Corresponding author: Julie AEC: jaec@elektro.dtu.dk
Contact information: HBDS: hbs@elektro.dtu.dk, HK: henriette_koch@hotmail.com, RF: RFRA0011@regionh.dk, JK: jkem@elektro.dtu.dk, LA: larv@lundbeck.com, SRC: srac@lundbeck.com and PJ: poje@glo.regionh.dk
[a]DTU Electrical Eng., Ørsteds Plads, bldg. 349, DK-2800 Kgs. Lyngby
[b]H. Lundbeck A/S, Ottiliavej 9, DK-2500 Valby
[c]Danish Center for Sleep Medicine, Department of Clinical Neurophysiology, Center for Healthy Aging, Faculty of Health Sciences, University of Copenhagen, Glostrup Hospital, DK-2600 Glostrup

During sleep, eye movements (EMs) are among other structures controlled by neurons located in the brain stem, and in [6], it was found that EMs during sleep hold the possibility of being a PD biomarker. In [6], the obtained performance was based on features reflecting EMs as well as features reflecting electromyography (EMG) measured at the EOG site. Also, the features were computed as the means and standard deviations in energy measures across all sleep epochs during a whole night of sleep. They thereby only reflected the overall differences in EMs between control subjects and iRBD/PD patients. In this study, the focus is on EMs alone, and a general data-driven topic model will be developed illustrating the timely distribution of EMs. A topic model is a statistical model revealing "topics" or "themes", which describe the latent structure behind the generation of a collection of documents. Here, a topic model is applied on data desribing EMs during sleep, and each sleep epoch will be represented as a mixture of three different states for EMs. The three states are thought to be related to slow EMs (SEMs), rapid EMs (REMs) and no EMs (NEMs). By applying the topic model on three test groups of ten control subjects, ten iRBD patients and ten PD patients, it will be analyzed how well the EMs from the patients fall into the normal states for EMs during sleep. By extracting three features from the topic models reflecting "certainty", "fragmentation" and "stability", the test subjects will be classified as "control" or "patient" by use of a Naive Bayes (NB) classifier.

In [7] is a general topic model built on EEG developed based on the same training data as in this study. The number of topics were set to five reflecting the five sleep stages stated in 2004 by the American Academy of Sleep Medicine (AASM) [8]. Features were extracted from the topic models obtained from the same test groups as in this study, and thereby the same subjects as in this study were classified. In this way, this study and [7] reveal how well these patient groups can be classified by application of EOG and EEG, respectively.

II. DATA ACQUISITION

Fourty subjects were enrolled in this study. They were all evaluated at the Danish Center for Sleep Medicine at Glostrup Hospital in Denmark, and the evaluation of the patients included PSG, multiple sleep latency test and a comprehensive medical history and medication. The control subjects included have no history of movement disorder, dream enacting behavior or other former diagnosed sleep

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

disorders. The quality of the PSG data was individually evaluated, and recordings were excluded if the analyzed channels were disconnected or continuously contaminated with artifacts. The demographic data for the groups is seen in Table I.

TABLE I
DEMOGRAPHIC DATA FOR THE FOUR SUBJECT GROUPS

Patient groups	Total No.	Male / Female	Age ($\mu \pm \sigma$) [years]
Controls (for train)	10	5 / 5	57.2 ± 8.1
Controls (for test)	10	5 / 5	59.8 ± 8.4
iRBD (for test)	10	8 / 2	59.0 ± 14.2
PD (for test)	10	6 / 4	63.2 ± 8.4

All subjects underwent at least one full night PSG according to AASM standards by use of different amplifier systems, where the lowest anti-aliasing filter cut-off frequency was 70 Hz. The EOG electrodes were placed one cm out and up (left) or down (right) from the outer canthus with reference to the right and left mastoid, respectively. The sampling frequency of the analyzed sleep data was 256 Hz.

III. METHODOLOGY

The overall methodology of this study is presented schematically in Fig. 1. Ten control subjects selected to best match the patient groups in age were used to develop a general topic model. As input to the topic model, features extracted from bandpass filtered EOG signals were given. By use of the general topic model, 30 topic mixture diagrams were obtained from ten control test subjects, ten iRBD test patients and ten PD test patients. Three features were extracted from these mixture diagrams, and by use of a standard NB classifier, the test subjects were classified as being either "patient" or "control". Below follows a more detailed description of the steps seen in Fig. 1

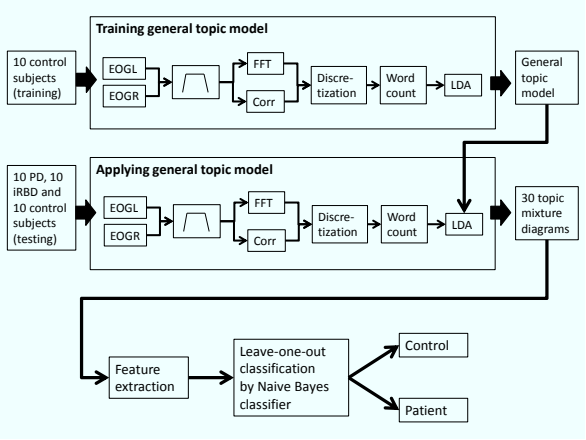

Fig. 1. A schematical overview of the methodology of this study. A general topic model was trained using ten control subjects. The general topic model was applied on ten other control subjects, ten iRBD patients and ten PD patients obtaining 30 topic mixture diagrams. Features were extracted from the topic mixture diagrams, and the subjects were classified as "control" or "patient" using an NB classifier.

A. Generating topic model

Initially, both EOG signals were bandpass filtered by a 4th order Butterworth filter with cut-off frequencies (3 dB) at 0.3 Hz and 10 Hz. These cut-off frequencies were chosen to focus the topic model on EMs by suppressing the influence of the baseline drift, the EMG activity as well as some EEG activity measured at the EOG sites. Both EOG signals were divided into non-overlapping segments of length L, and for each of these segments, three features were computed, yielding a feature vector $f(n)$ expressed as,

$$f(n) = \begin{bmatrix} S_{ll}(n) \\ S_{rr}(n) \\ R_{lr}(n) \end{bmatrix} \quad (1)$$

where n denotes the segment index, S_{ll} and S_{rr} represents the spectral power computed by the fast Fourier Transform (FFT) below 5 Hz in the left and right EOG signal segment, respectively. Any EMs, whether it be SEMs, REMs or a combination of the two, are assumed to be in the range of 0-5 Hz [9]. The R_{lr} represents the normalized cross-correlation coefficient between the left and right EOG signal segment given by,

$$R_{lr}(n) = \frac{\sigma_{lr}(n)}{\sqrt{\sigma_{ll}(n)\sigma_{rr}(n)}} \quad (2)$$

where σ_{ll} and σ_{rr} denotes the variance of the left and right EOG signal segment, respectively, and σ_{lr} denotes the covariance of the left and right EOG signal segment. As the EOG signals appear anticorrelated during EMs, it is assumed that R_{lr} will obtain negative values when REMs occur during REM sleep or wakefulness and when SEMs occur during N1 sleep. Background EOG should appear almost uncorrelated, and the high-amplitude EEG artifacts which can occur during deep sleep should appear correlated. The subject-specific median of the cross-correlation features was subtracted to align the values around zero.

As in [4], the aim is to train a topic model by use of the Latent Dirichlet Allocation (LDA) model. To be able to use the features as input to such a topic model, the features were discretized on a per-subject basis. The spectral power features were given the values 1 to 4 based on boundaries set at each quartile for the full range of feature values for that specific subject. The cross-correlation features were discretized given values 1 to 4 based on boundaries set at [-0.7, 0, 0.7] for all subjects. These boundaries were set based on trial-and-error of best catching the EMs (at values below -0.7), and the EEG artifacts (values above 0.7) as well as the idea of having symmetric boundaries around zero.

The LDA method assumes that a "collection of documents" is derived from an underlying set of "topics", and that the topics are defined as a set of related "words" [10]. As the discretization in this study was done by symbols of 1 to 4, a word length of W is presented by either one of all combinations of W succeeding values of 1 to 4. The LDA assumes that each topic can be defined as a certain distribution over all of the available words. For each document in the collection of documents, a count is formed

of the number of occurrences of each word, and as an end result a topic-by-document matrix X is found, descriping the distribution over topics in each document [10].

As in [4], the document length in this study was set to 30 seconds (comparable with a sleep epoch), yielding that each sleep epoch consisted of a total of $3 \times \frac{30}{L}$ symbol instances. Different word lengths were tried ($W = 2,3,5$), giving that the total number of available words was 3×4^W. The number of topics was set to $T = 3$, in trying to reflect the different states (SEMs, REMs and NEMs) for EMs during sleep.

To train a general topic model, all the available sleep epochs in between lights off and lights on from ten control subjects were used as the collection of documents. By using data from control subjects only, a general "control topic model" was thereby trained. The topic model was applied on the three test groups (see Table I), yielding a topic mixture diagram X holding the distribution of the three "control topics" in each sleep epoch from each of the subjects in the test data.

B. Feature extraction and classification

The aim of this study is to classify the 30 test subjects into either "control" or "patient" based on the topic mixture diagrams obtained when using a general topic model. For each test subject, three features were computed. The features reflect "certainty", "fragmentation" and "stability", and are defined as:

Feature 1 - "Certainty": The amount of epochs with a dominating topic of a probability higher than a given threshold. Normalization was done by dividing the number with the subject-specific total number of epochs. Feature 1 is expressed as,

$$f_1^p = \frac{\sum_{k=1}^{K} logical\left(max(X_k^p) > th\right)}{K} \tag{3}$$

where K is the subject-specific total number of epochs and X_k^p is the EM topic mixture for epoch k in subject p. The threshold value th was defined as the one giving the highest mean Area Under Curve (AUC) when classifying the 30 test subjects using the leave-one-subject-out validation scheme.

Feature 2 - "Fragmentation": The amount of state shifts between topics when the dominating topic defines the state of an epoch. Normalization was done by dividing the number with the subject-specific total number of epochs. Feature 2 is expressed as,

$$f_2^p = \frac{\sum_{k=1}^{K-1} logical\left(max(X_k^p) \neq max(X_{k+1}^p)\right)}{K} \tag{4}$$

Feature 3 - "Stability": The normalized mean number of epochs kept in a certain state when the dominating topic defines the state of an epoch. Feature 3 is expressed as,

$$f_3^p = \frac{\sum_{m=1}^{M} e_m^{new}}{M} \quad \text{with} \quad e^{new} = \frac{e^{old} - \min(e^{old})}{\max(e^{old}) - \min(e^{old})} \tag{5}$$

where m is an index for a period, in where the epochs all have the same dominating topic, M is the subject-specific total number of such periods and e^{old} is a vector holding the M non-normalized numbers of epochs in each period.

As the topic mixture diagrams depend on the initialization of the LDA method, and as it was noticed that the feature values therefore slightly differed in between different runs on the same test subject, the three described features were computed for 20 different runs on the testdata. The mean of the 20 feature values were used as the final feature values. Using the leave-one-subject-out approach, a standard NB classifier was used to classify the subjects into either "control" or "patient". The classification were performed using all combinations of either one, two or all three feature values.

As mentioned earlier, different values were tried for the word length W ($W = 2,3,5$) and for the segment length L ($L = 1,3$). The final topic model developed from the training dataset was chosen based on how well the NB classifier performed (according to accuracy) on the test dataset.

IV. RESULTS AND DISCUSSION

A. Interpreting the topic mixture diagrams

Fig. 2, 3 and 4 present an example of a topic mixture diagram from a control subject, an iRBD patient and a PD patient, respectively. Each vertical coloured bin presents a sleep epoch, and the amount of each color in a bin presents the individual topic probability. Remembering that the three topics are derived based on features reflecting EMs, it is seen, that the general topic model do recognize the characteristic temporal evolution of sleep. More specifically, the "blue" topic could be interpreted as having something to do with the REMs in REM sleep, whereas the "green" topic could be linked to SEMs and the "red" topic could be linked to NEMs. It is seen from the mixture topic diagrams in Fig. 3 and 4, that not as many sleep epochs show a high certainty of either topic as compared to the control mixture diagram in Fig. 2. Interpreting the topics as just described, this observation lead to the conception that the EMs (both the REMs and SEMs) in the patients are less pronounced or less alike the EMs in control subjects. Other observations include the more abrubt transitions in between topics as well as the less structured and more fragmented profiles for the iRBD and PD patients compared to the control subjects. These observations are tried captured in the features "certainty", "fragmentation" and "stability".

B. Classification

A standard NB classifier was used to classify the subjects by the leave-one-subject-out validation approach, and it was found that the model, which obtained the highest mean accuracy, had a segment length of $L = 1$ and a word length of $W = 3$. This model used the features "certainty" and "stability", and in Fig. 5 the decision boundary is illustrated by the colors gray (classified as "patient") and white (classified as "control"). The test subjects are marked by red (PD patient), green (iRBD patient) or blue (control subject) filled circles. It is seen that two control subjects and one iRBD patient are misclassified, yielding a sensitivity of 95 %, a specificity of 80 % and an accuracy of 90 %.

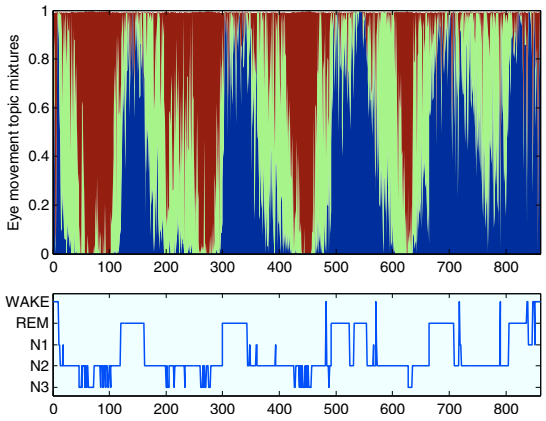

Fig. 2. A topic mixture diagram and the manually scored hypnogram for a control subject.

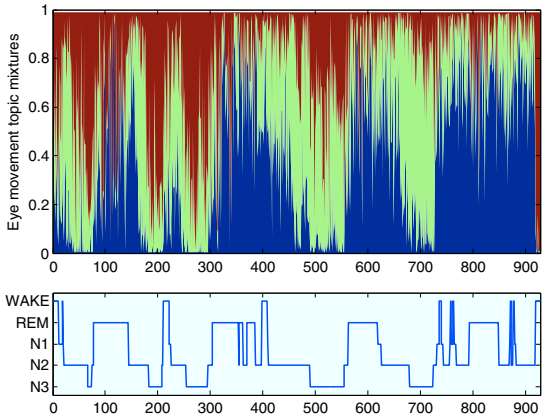

Fig. 3. A topic mixture diagram and the manually scored hypnogram for an iRBD patient.

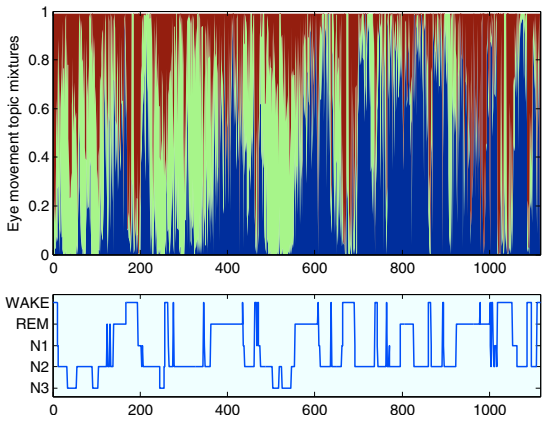

Fig. 4. A topic mixture diagram and the manually scored hypnogram for a PD patient.

V. CONCLUSIONS

Training a general topic model based on sleep EOG from ten control subjects, revealed that the characteristic sleep cycles can be encompassed solely by use of features reflecting EMs. By applying the topic model on testdata from ten other control subjects, ten iRBD patients and ten PD patients, a topic mixture diagram was obtained for each subject. Features reflecting "certainty", "fragmentation" and "stability" of these diagrams were derived. It was found that by use of the features "certainty" and "stability", a simple NB classifier classified the subjects with a sensitivity of 95 %, a specificity of 80 % and an accuracy of 90 %. The separability of the individual features as well as new features derived from the topic mixture diagrams should be further investigated. Although more focused analyzes of the morphology of EMs are needed, this study demonstrates with a data-driven, unsupervised approach that PD and iRBD patients reflect abnorm form and/or timely distribution of EMs during sleep.

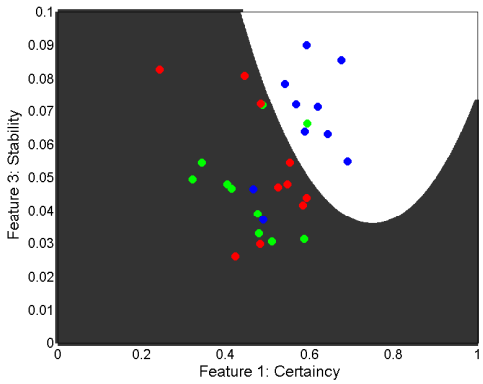

Fig. 5. The best NB classification result was based on two features. The decision boundary is illustrated by the colors white (control area) and gray (patient area), and the 30 test subjects are marked with blue (control subject), green (iRBD patient) or red (PD patient) filled circles.

REFERENCES

[1] C. Schenck, S. Bundlie, and M. Mahowald, "Delayed emergence of a parkinsonian disorder in 38% of 29 older men initially diagnosed with idiopathic rapid eye movement sleep behaviour disorder," *Neurology*, vol. 46, no. 2, pp. 388–393, 1996.

[2] R. B. Postuma, J. F. Gagnon, S. Rompré, and J. Y. Montplaisir, "Severity of REM atonia loss in idiopathic REM sleep behavior disorder predicts Parkinson disease," *Neurology*, vol. 74, pp. 239–244, Jan. 2010.

[3] V. Latreille, J. Carrier, J. Montplaisir, M. Lafortune, and J.-F. Gagnon, "Non-rapid eye movement sleep characteristics in idiopathic REM sleep behavior disorder," *Journal of the Neurological Sciences*, vol. 310, pp. 159–162, June 2011.

[4] A. Van Esbroeck and B. Westover, "Data-driven modeling of sleep states from EEG," in *34th Annual International Conference of the IEEE EMBS*, pp. 5090–5093, 2012.

[5] P. S. Jensen, H. B. D. Sorensen, H. L. Leonthin, and P. Jennum, "Automatic sleep scoring in normals and in individuals with neurodegenerative disorders according to new international sleep scoring criteria," *Journal of Clinical Neurophysiology*, vol. 27, pp. 296–302, Aug. 2010.

[6] J. A. E. Christensen, R. Frandsen, J. Kempfner, L. Arvastson, S. R. Christensen, P. Jennum, and H. B. D. Sorensen, "Separation of Parkinson's patients in early and mature stages from control subjects using one EOG channel," in *34th Annual International Conference of the IEEE EMBS*, pp. 2941–2944, 2012.

[7] H. Koch, J. A. E. Christensen, L. Arvastson, S. R. Christensen, P. Jennum, and H. B. D. Sorensen, "Classification of iRBD and Parkinson's patients using a general data-driven sleep staging model built on EEG," in *35th Annual International Conference of the IEEE EMBS [accepted]*, 2013.

[8] C. Iber, *The AASM Manual for the Scoring of Sleep and Associated Events*. American Academy of Sleep Medicine, 2007.

[9] R. Agarwal, T. Takeuchi, S. Laroche, and J. Gotman, "Detection of rapid-eye movements in sleep studies," *IEEE Transactions on Bio-Medical Engineering*, vol. 52, pp. 1390–1396, Aug. 2005.

[10] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.