

Automated Sleep Quality Measurement using EEG Signal

-First Step Towards a Domain Specific Music Recommendation System

Wei Zhao, Xinxi Wang and Ye Wang
School of Computing, National University of Singapore
{zhaowei,wangxinx,wangye}@comp.nus.edu.sg

ABSTRACT

With the rapid pace of modern life, millions of people suffer from sleep problems. Music therapy, as a non-medication approach to mitigating sleep problems, has attracted increasing attention recently. However the adaptability of music therapy is limited by the time consuming task of choosing suitable music for users. Inspired by this observation, we discuss the concept of a domain specific music recommendation system, which automatically recommends music for users according to their sleep quality. The proposed system requires multidisciplinary efforts including automated sleep quality measurement and content-based music similarity measure. As a first step, we focus on the automated sleep quality measurement in this paper. An EEG-based approach is proposed to measure user's sleep quality. The advantages of our approach over standard Polysomnography (PSG) method are: 1) it measures sleep quality by recognizing three sleep categories rather than six sleep stages, thus higher accuracy can be expected; 2) three sleep categories are recognized by analyzing Electroencephalography (EEG) signal only, so the user experience is improved because he is attached with fewer sensors during sleep. We conduct experiments based on a standard data set. Our approach achieves high accuracy and shows promising potential for the music recommendation system.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Relevance feedback

General Terms

Algorithms, Design, Experimentation, Human Factors

Keywords

Music Recommendation, Music Therapy, Sleep Disorder, Sleep Quality Analysis

1. INTRODUCTION

Nowadays millions of people are affected by sleep problems, many of whom remain undiagnosed. At present mu-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM'10, October 25–29, 2010, Firenze, Italy.

Copyright 2010 ACM 978-1-60558-933-6/10/10 ...\$10.00.

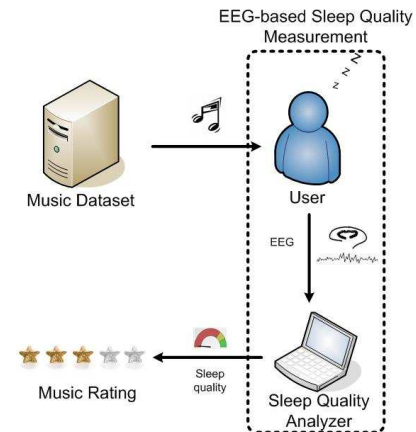


Figure 1: EEG-based music rating in the proposed recommendation system

sic therapy offers an alternative healing method, which improves sleep quality by playing back music at bed time. Music therapy research has indicated that music does have beneficial effects on sleep for children [17], young people [8] and older adults [10]. During the process of music therapy, people are asked to listen to a list of music, which is pre-selected by music therapist. In spite of its clear benefits towards sleep quality, current approach is difficult to be widely used because it is a time consuming task for music therapist to produce a personalized music list. Based on this observation, we introduce a domain specific music recommendation system, which automatically recommends music for user according to his sleep quality.

The proposed recommendation system consists of two main components: **EEG-based music rating** and **content-based music recommendation**. After music is played back at bed time, the former component will monitor user's sleep quality. Music items will be rated according to user's sleep quality, as figure 1 shows. The music items associated with good sleep quality are used as queries in the music recommendation component. The music which are similar to the query items will be recommended to the user by conducting music similarity analysis. As the first step towards the proposed recommendation system, we focus on automated sleep quality measurement which is the major task in the EEG-based music rating component.

Normal sleep physiology consists of Non Rapid Eye Movement (NREM) and Rapid Eye Movement (REM). NREM is subdivided into four stages: stages 1 (S1), stage 2 (S2), stage 3 (S3) and stage 4 (S4). S3 and S4 are also called deep sleep. A healthy person will experience about 5 complete sleep cy-

cles per night. [1]. S1 is the first stage in a sleep cycle. S2 is the second stage. Then S3 and S4 will occur in the cycle consecutively. After the completion of four stages of NREM, these four stages will reverse rapidly and then followed by REM, as Figure 2 shows.

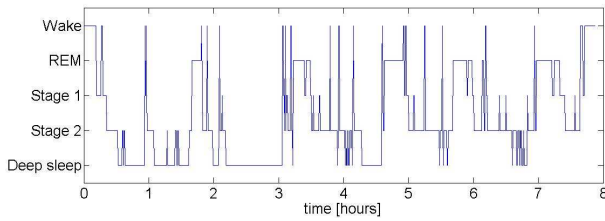


Figure 2: Typical sleep cycles of one night

Given the sleep cycles over night, three main parameters can be calculated to measure sleep quality: sleep latency, sleep efficiency and percentage of deep sleep. Specifically, **sleep latency** is the time that it takes to finish the transition from wakefulness to the first sleep stage. **Sleep efficiency** is the ratio of time spent asleep to the time spent in bed. **Percentage of deep sleep** is the ratio of deep sleep to the all sleep stages. To calculate these parameters, we do not need to recognize every sleep stages in sleep cycles. It is enough if three sleep categories can be distinguished: wakefulness, deep sleep (S3,S4), and other sleep stages (S1,S2,REM). Consequently, the problem of sleep quality measurement is converted into how to recognize these three sleep categories.

Based on the analysis of current PSG techniques in Section 2, we proposed to recognize the three sleep categories from EEG signal only. The details of our approach is presented in Section 3. Section 4 discusses the experiment results. Conclusions are given in Section 5.

2. RELATED WORKS

Now Polysomnography (PSG) technique is widely used in hospital to monitor patients' sleep cycles. First developed in 1960s [15], this method, also known as sleep scoring, has become a golden standard in sleep studies. While users are sleeping, three physiological signals are monitored: Electroencephalography (EEG), Electrooculography (EOG), and Electromyography (EMG). Based on the analysis of these three signals, six sleep stages can be recognized by human expert: wakefulness, S1, S2, S3, S4 and REM [16]. These stages are usually scored on 30-second epoch.

The standard PSG approach utilizes 9 sensors to monitor the required signals: EEG, EOG and EMG [3]. In this setup, users may feel uncomfortable during sleep because so many sensors are attached on their face and scalp. To improve the user experience, we use fewer signal channels to recognize sleep stages. As discussed in Section 1, to calculate sleep quality parameters, we just need to recognize three categories: wakefulness, deep sleep and other stages. On the account of the fact that EOG and EMG are mainly used to distinguish between REM and S1, these two signals may not contribute much in the calculation of sleep quality. Consequently, we recognize the three sleep categories from EEG signal only. This approach is also approved by Sleep Doctors of Hospital.

Two nice survey papers were published about computerized PSG analysis that is how to automatically recognize sleep stages from physiological signals [12, 13]. Based on

literature, automated PSG analysis can achieve reasonable accuracy compare to human scoring. Current automated scoring systems can be mainly categorized into two families: rule-based expert system and classifier-based system.

Peter Anderer et al. [2] developed an expert system to recognize six sleep stages (W,S1,S2,S3,S4,REM) from three signals (EEG,EMG,EOG). A decision tree is built based on 10 linear discriminant analysis classifiers. The 80% agreement between expert system scoring and human scoring was achieved. The commercial package, BioSleep of Oxford BioSignals Company, is a typical classifier-based system [11]. Auto regressive coefficients are extracted as feature vectors from EEG segment, then a neural network classifier is used to classify each EEG segment into different sleep stages. The BioSleep package obtains reasonable results with the comparison of human scoring in the third-part evaluation [9].

Based on the survey in literature, most of existing works intend to provide a service of sleep scoring for clinic application. The main differences between our approach and these works are that our approach is specifically designed for a multimedia application rather than clinic one. In particular, we balance the user experience and system accuracy, a EEG-based approach is proposed to recognize three sleep categories.

3. EEG-BASED SLEEP QUALITY MEASUREMENT

To measure sleep quality, we need to recognize three sleep stage categories: wakefulness, deep sleep, and other sleep stages. To recognize these three sleep categories, we extract spectral power feature from each 30-seconds EEG epoch. LibSVM package [4] is used to build a SVM classifier to classify each EEG epoch into one of the three categories. Additionally, as discussed in Section 1, the transition of sleep stages follows the trend of sleep cycles. For example, if current epoch belongs to deep sleep, then the next epoch probably also belongs to deep sleep. However, SVM classifier treats each 30-second epoch independently, thus it does not take the advantage of correlations between nearby epochs. Consequently, we further model the sleep transition as a Markov chain. A matrix will be learned from training data to indicate the transition probability between different sleep stages, as Figure 3 illustrates. Based on this matrix the classification results are further refined. The probability estimated by SVM classifier for each epoch is processed by a dynamic programming algorithm.

3.1 Feature Extraction

EEG is the summation of electrical signal generated by millions of neurons in the brain. It was first recorded by Richard in 1875 [5], now has been widely used in PSG studies. EEG signal is usually divided into five different frequency bands: delta (0.5-2 Hz), theta (2-7 Hz), alpha (8-12 Hz), beta (12-20 Hz) and fast beta (20-40 Hz). According to previous studies, spectral powers of EEG in different frequency bands highly relate to sleep stages. In particular, some sleep stages are recognized by the presence of EEG signal in the specific frequency band. Spectral power in specific frequency band of EEG is a well-known pattern in PSG studies [16]. For example, deep sleep is recognized by the present of high amplitude delta wave of EEG. One epoch will be scored as wakefulness when alpha wave presents in

EEG. Consequently, we use spectral power extracted from five frequency bands as the feature vector. As sleep stage is usually scored based on 30-second epoch, feature vector is generated for each 30-second EEG epoch.

3.2 Classification by SVM

Platt et al. [14] presented a work to convert the results of binary SVM into posterior probabilities. Later this work is extended to estimate multi-class SVM probabilities by combining the output of multiple one-against-one binary classifiers [18]. With this technique, we use a multi-class SVM classifier to estimate $P(x|F_t)$, the probability of epoch t belonging to sleep stage x . F_t is the feature vector extracted from epoch t . We also define $C(t, x) = P(x|F_t)$. $C(t, x)$ will be used in the following section.

In SVM classification, epoch t is scored as the stage which could maximize the probability $C(t, x)$, as following equation shows:

$$\psi(t) = \arg \max_x C(t, x),$$

where $\psi(t)$ is the label generated by classifier for epoch t .

3.3 Sleep Stage Transition Modeling

As SVM classifier treats each epoch independently, it does not utilize the correlation between epochs. To take the advantage of sequential information, we model the transition of sleep stages as a discrete time Markov chain. The effectiveness of this modeling has been proved by Gibellato [6]. The property of Markov chain can be formalized in equation 1. It is also called first-order Markov assumption, the stage of current epoch just depends on the stage of previous epoch.

$$\begin{aligned} Pr(X_t = x_t | X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \dots, X_1 = x_1) \\ = Pr(X_t = x_t | X_{t-1} = x_{t-1}) = P_{x_{t-1}x_t}, \end{aligned} \quad (1)$$

where X_t indicates the sleep stage of epoch t ; $P_{x_{t-1}x_t}$ indicates the transition probability from stage x_{t-1} to stage x_t .

As x has three possible stages, there are nine possible values for $P_{x_{t-1}x_t}$. These nine values construct a 3 by 3 matrix which indicates the transition probability between three stages. This matrix is learned from training data as follows:

$$P_{ij} = \frac{Count(i, j)}{\sum_{k=1}^3 Count(i, k)}$$

where $Count(i, j)$ counts the transitions from stage i to stage j . The calculation of this matrix is also illustrated in Figure 3.

3.4 Dynamic Programming for Post-processing

Based on the probability generated by SVM and the transition matrix learned from training data, a dynamic programming (DP) algorithm is designed to score each epoch in the way that optimum overall posterior probability can be obtained. The subproblem of dynamic programming is defined in Equation 2, which presents the maximum overall probability from the first epoch to current epoch t , where current epoch is scored as stage v .

$$\begin{aligned} L(t, v) = \max_{x_1, \dots, x_{t-1}} \\ \left\{ Pr(X_1 = x_1, \dots, X_{t-1} = x_{t-1}, X_t = v) \prod_{i=1}^t C(i, X_i) \right\} \end{aligned} \quad (2)$$

Table 1: Accuracy in 10-fold cross-validation

Recording	st7022j0	st7052j0	st7121j0	st7132j0
SVM	88.4%	93.9%	90.9%	92.7%

Because first-order Markov assumption is hold, Equation 2 can be simplified as follows:

$$L(t, v) = \max_u \{L(t-1, u)Pr(X_t = v|X_{t-1} = u)\} C(t, v)$$

Thus $L(t, v)$ can be calculated on top of $L(t-1, \cdot)$, and this forms an optimal substructure for the subproblem. Therefore, a dynamic programming algorithm can be designed to find the optimum solution.

To eliminate the float point overflow problem, logarithm probability is used: $S(t, v) = \ln L(t, v)$. For the first epoch, we define $S(1, v) = \ln C(1, v)$.

$$S(t, v) = \begin{cases} \ln C(1, v) & t = 1 \\ \max_u \{S(t-1, u) + \ln P_{uv}\} + \ln C(t, v) & t > 1 \end{cases}$$

In backtracing, variable $\alpha(t, v)$ is defined to record the optimum sleep stage of previous epoch.

$$\alpha(t, v) = \arg \max_u \{S(t-1, u) + \ln P_{uv} + \ln C(t, v)\}$$

The process of backtracing is shown in the following formulation:

$$\phi(t) = \begin{cases} \alpha(t+1, \phi(t+1)) & t < n \\ \arg \max_v S(n, v) & t = n \end{cases}$$

where ϕ is the refined classification labels which result in the maximum overall posterior probability.

4. EVALUATIONS

The automated sleep quality measurement consists of two phases. First, SVM classifier categorizes EEG epoch into one of three classes: wakefulness, deep sleep, and other sleep stages. Second, a dynamic programming algorithm is used to refine the results of SVM by utilizing the sequential information. Two experiments are respectively conducted to evaluate the classifier and the post-process algorithm.

Experiments are conducted based on the Sleep EDF Database, which is selected from the PhysioBank that is a large achieve of digital recording for biomedical research community [7]. Four recordings which were obtained from healthy subjects in the hospital are used. Three signals (EEG, EOG and EMG) were monitored during night with good signal quality. Only the EEG channel (Fpz-Cz) is utilized in our experiment. This data set is annotated by human expert according to PSG standard [15]. Human annotation is used as the ground truth in our experiments.

4.1 Sleep Stage Classification

In the first experiment, we evaluate the accuracy of SVM classifier. Four sleep recordings of Sleep EDF Database are divided into 30-second epochs, and 3874 epochs are generated in total. The spectral power feature is extracted from each epoch. 10-fold cross-validation is conducted based on the feature vectors of each recording. The classifier achieves the average accuracy up to 93.9%, as described in Table 1.

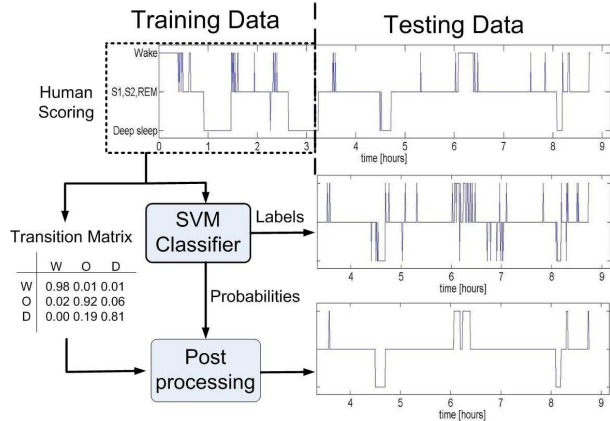
4.2 Post Processing

In the second experiment, we evaluate the performance of DP algorithm. As DP algorithm utilizes sequential information, it processes on continuous epochs. Consequently,

Table 2: Accuracy in experiment 2

Recording	st7022j0	st7052j0	st7121j0	st7132j0
SVM	87.5%	92.7%	93.4%	94.2%
SVM+DP	89.3%	95.8%	96.8%	99.3%

each recording is divided into two continuous parts, one for training and one for testing. The first 400 epochs of each recording are used to train the SVM model and learn the transition matrix. The rest epochs are testing data. The experiment conducted on one recording is demonstrated in Figure 3. Based on the experiment results showed in Table 2, DP algorithm stably improves the accuracy in each recording. Comparing the sleep cycle generated by classifier and DP algorithm, in spite of the improvement in accuracy, DP algorithm also generates a more smooth sleep cycle than the SVM classifier, as Figure 3 shows.

**Figure 3: Experiment over the recording of st7052j0**

5. CONCLUSIONS

In this paper, we discussed the concept of a domain specific music recommendation system, which automatically recommends music for users according to their sleep quality. One important problem in the proposed recommendation system is how to automatically monitor users' sleep quality at night. To address this problem, we first investigate sleep physiology and traditional PSG approach in literature. Considering that standard PSG system may make users feel uncomfortable, we specifically design an approach to recognize three sleep categories from EEG signal only. Then three parameters, sleep latency, sleep efficiency and percentage of deep sleep, can be calculated to measure sleep quality. The experiment results demonstrates that our approach can achieve high accuracy though only the EEG channel is used. These results motivate us to further develop the rest of components in the music recommendation system and evaluate its full functions in the future.

6. REFERENCES

- [1] <http://en.wikipedia.org/wiki/sleep>, retrieved on 29 april 2010.
- [2] P. Anderer, G. Gruber, S. Parapatics, M. Woertz, T. Miazhyńskaia, G. Klosch, B. Saletu, J. Zeitlhofer, M. J. Barbanoj, H. Danker-Hopfe, S.-L. L. Himanen, B. Kemp, T. Penzel, M. Grozinger, D. Kunz, P. Rappelsberger, A. Schlogl, and G. Dorffner. An e-health solution for automatic sleep classification

according to rechtschaffen and kales: validation study of the somnolyzer 24 x 7 utilizing the siesta database. *Neuropsychobiology*, 51(3):115–133, 2005.

- [3] K. A. Carden. Recording sleep: The electrodes, 10/20 recording system, and sleep system specifications. *Sleep Medicine Clinics*, 4(3):333–341, September 2009.
- [4] C. C. Chang and C. J. Lin. *LIBSVM: a library for support vector machines*, 2001.
- [5] M. Deak and L. J. Epstein. The history of polysomnography. *Sleep Medicine Clinics*, 4(3):313–321, September 2009.
- [6] M. G. Gibellato. *Stochastic modeling of the sleep process*. PhD thesis, Ohio State University, 2005.
- [7] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000 (June 13).
- [8] Harmat, L. Takács, J. Bódizs, and Róbert. Music improves sleep quality in students. *Journal of Advanced Nursing*, 62:327–335, 2008.
- [9] C. Jennifer, G. G. John, H. J. Phil, J. GRIFFITHS Clive, and J. DRINNAN Michael. Comparison of manual sleep staging with automated neural network-based analysis in clinical practice. *Medical & biological engineering & computing*, pages 105–110, 2006.
- [10] Lai, H.-L. Good, and Marion. Music improves sleep quality in older adults. *Journal of Advanced Nursing*, volum 53, issue 1:134, 2006.
- [11] N. McGrogan, E. Braithwaite, and L. Tarassenko. Biosleep: a comprehensive sleep analysis system. In *Engineering in Medicine and Biology Society, Annual International Conference of the IEEE*, November 2002.
- [12] T. Penzel. Computer based sleep recording and analysis. *Sleep Medicine Reviews*, 4(2):131–148, April 2000.
- [13] T. Penzel, M. Hirshkowitz, J. Harsh, R. D. Chervin, N. Butkov, M. Kryger, B. Malow, M. V. Vitiello, M. H. Silber, C. A. Kushida, and A. L. Chesson. Digital analysis and technical specifications. *Journal of clinical sleep medicine*, 3(2):109–120, March 2007.
- [14] J. C. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in Large Margin Classifiers*, pages 61–74, 1999.
- [15] A. Rechtschaffen and A. Kales. *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects*. US Government Printing Office, US Public Health Service, Washington DC, 1968.
- [16] M. H. Silber. Staging sleep. *Sleep Medicine Clinics*, 4(3):343–352, September 2009.
- [17] L. P. Tan. The effects of background music on quality of sleep in elementary school children. *Journal of Music therapy*, pages 128–150, 2004.
- [18] T. F. Wu, C. J. Lin, and R. C. Weng. Probability estimates for multi-class classification by pairwise coupling. *J. Mach. Learn. Res.*, 5:975–1005, 2004.