Toward the Next-Generation Sleep Monitoring / Evaluation by Human Body Vibration Analysis

Takahiro Komine
The University of Electro-Communications, Tokyo
tkomine@cas.hc.uec.ac.jp

Keiki Takadama
The University of Electro-Communications, Tokyo
keiki@inf.uec.ac.jp

Seiji Nishino
Stanford University School of Medicine
nishino@stanford.edu

Abstract

This paper describes one of the future images of the sleep monitoring system. The new technology should satisfy the following requirements: (1)noninvasive, (2)low cost and (3)long-term monitoring. What we propose here is the sleep monitoring system based on the human body vibrations sensed by the mattress type pressure sensors that gradually improves its estimation performance to the particular user by learning collected data and reconstructing its classifier. In order to learn the data, however, the system needs the vibration data mapped to the appropriate sleep stages. As the solution to the problem, we use the existing approximate sleep stage estimation method. The experimental results reveal that (1)there is only a slightly difference between the accuracies of the two classifiers; the one trained the original dataset plus PSG based sleep stage labeled data; the other one trained the original dataset plus approximate sleep stage labeled data; (2)Adding a particular user's several days data to the training data improves the accuracy of the original classifiers. The REM estimation accuracy is 87 % in maximum. From those results, the contribution of this research is suggesting the way to personalize sleep estimation, and proving the effectiveness.

Introduction

Recently, the number of the people suffered from the sleep disorders is increasing. There is a research report that one fifth people in Japan have some problems in their sleep(Kim et al. 2000). The ministry of Health, Labour and Welfare of Japan also reports nearly half of people in Japan is unsatisfied to their sleep quality(Ministry of Health, Labour and Welfare of Japan 2013). Despite such a condition, the number of the sleep quality quantification methods which are available in medical diagnosis is very limited. The major sleep evaluation method in the medical front is PSG (polysomnography). Some medical doctors mentioned the problems of PSG as follows;

1. PSG is invasive to the users, so it is difficult to use this method for children or some people feel discomfort of wearing electrodes;
2. PSG analysis require several professionals, so expensive cost and time are needed;
3. PSG is detectable to the users, in concrete, it requires a special environment (including devices and rooms) bringing *unusual* atmosphere to the patients, and such situation hinders the patients' natural sleep.
4. The criteria of the sleep depths are defined as world standard(Rechtschaffen and Kales 1968), however, the details are slightly different by each user. This is the reason why the PSG analysis is still performed by hands.

Now, the new sleep evaluation, in other words, sleep stage estimation system is required.

Our Motivations: Through Monitoring, Improving Estimation Accuracy

Our basic motivation is to develop new sleep stage estimation system which is (1)noninvasive, (2)low cost, and (3)undetectable to users. Considering those conditions, we select the mattress type sensor based system. Not from brain wave, but from body vibrations our system evaluate the depth of sleep. In addition, what we focus is that those four problems of PSG make it more difficult to monitor (or evaluation) the everyday's sleeping at home continuously, despite of the fact that such monitoring should be essential to evaluate the effectiveness of sleep disorder treatments. Furthermore, we can also imagine that it enables to prevent any disease by sensing some change of sleep. For those reasons, not only the viewpoints of medical treatments, but also of our daily life, sleep monitoring might improve our QOL.
Using AI, what can we do for the sleep monitoring? The idea is that the machine might gradually improve its estimation performance to the particular user by using collected

data. We aim at developing the user-adaptive personalized sleep stage estimation system based on the human body vibrations. Once the system begins operating, it goes on collecting the characteristics of him/her. Here, the word "characteristics" means, for example, the average rhythm of respirations and of heart beat variability. In other words, the system *learns*, and gradually be personalized to the user.

Awating Solution

The idea is very simple, however, it has a potential problem. In this research, we suppose that the sleep stage estimation system improves by supervised learning technique. To improve itself, the system needs training data; the sensed vibration mapped to the appropriate sleep stages. As the true sleep stages are acquired by using PSG analysis only, the approximate sleep stage estimation is needed.
Then, the questions are; (1)How to estimate the approximate sleep stages? (2)How much the estimation accuracy improves using such approximate sleep stages?
In this paper, we provide one of the solutions of the question(1), and reveal the experimental results for (2).

Related Works

Sleep Stage

The world standard sleep criteria is defined as **Sleep Stage**(Rechtschaffen and Kales 1968), based on PSG; the states of brain waves, eye movements and activities of chin muscles. The criteria classifis human sleep conditions into 6 stages; **WAKE, REM, NREM-1, 2, 3 and 4**. In this work we cannot use those brain consition data directly, because the mattress sensor could get only the vibration waves from our body (chest, strictly).
Indirectly, however, the body activities responce the brain consitions. What we should do is to catch the well-known features which appear as body activities during REM sleep as follows;

- The variablity of heart and respiration rate are increasing than those of NREM sleep.
- The antigravity muscles relax.
- Twitch movements appear.

Mattress Sensor based Sleep Stage Estimation

Some researcher develop the mattress sensor based sleep stage estimation methods. Watanabe et al. (Watanabe and Watanabe 2001) focused on the macro change of heart rate variability (HRV) which is acquired from the filtered viblation waves. In the papers they reports that their method successfully extract the macro change of HRV, and that is correration to the sleep stage transition. This method has, however, a potential problem. That can extract the macro change (or wave) of HRV, but there is not a clear criterion that relates the amplitude of the wave and the depth of sleep stages. Tomura et al. are inspired from Watanabe's works, and they propose the new criteria for REM sleep and WAKE detection (Tomura et al. 2015) as follows;

NREM stages estimation criteria: Suppose the coefficient of variation$CV(x) = Std(x)/Ave(x)$. Calculating CV every 5 minutes of respiration. During the CV is over the threshold θ_1, such epochs are estimated to stage NREM-1. During CV is under the threshold θ_3, such epochs are estimated to stage NREM-3, otherwise NREM-2. Tomura mentions those thresholds should be $\theta_1 = CV^* \times 0.6$, and $\theta_3 = CV^* \times 0.2$, where CV^* indicates the value of CV of whole respiration data.

REM stage estimation criteria: Calculating averages every 5 minutes of *the macro HRV wave:* W, then estimate the REM sleep when the average is above the threshold θ_{REM}, calculated as the following equation; $\theta_{REM} = Ave(W) + Std(W) \times 0.6$.

WAKE stage estimation criteria: Calculating every minute of *Activity*, then estimate the WAKE stage when the average is above the threshold θ_{Act}, calculated as the following equation; $\theta_{Act} = Ave(Activity) + Std(Activity) \times 0.6$.

Note that *the macro HRV wave:* W is the waves filtered from HRV wave which extracted from body vibration waves. Tomura uses the filter which transmits the waves of 1.23×10^{-4} to 2.5[Hz]. Also here the *Activity* indicates the standard deviation of the sensor values during 1second. $Ave(x)$ and $Std(x)$ indicate the average and the standard deviation of an array x, respectively.

Proposed Method

Selecting the Type of the Sensors

The mattress type pressure sensors are (1) noninvasive, (2) low cost. Generally, such sensors are undetectable to users when they are set under the mattress. This time, we use the TANITA Sleep Scan; SL-511(Noh et al. 2009) (Fig.1), which is one of the commercially available pressure sensor in Japan. The sensor outputs the sensing values 16 times per second. Note that the sensor covers only the area of the chest.

Figure 1: TANITA Sleep Scan SL-511.
(http:\\www.tanita.co.jp\product\g_TSL511WF\)

Inputs, Internal Processing and Outputs

From the pressure sensor, the input information is vibrations from the patient's body only. The sensor values of 30 seconds are shown in Fig.2. Here, the vibration includes respirations, heart beats, body motions and noise. In summary, the proposed method senses vibrations of the user through the sensor, and outputs the 6 levels of the labels; sleep stages. The estimation is designed based on frequency domain analysis by the short-time Fourier transform (STFT). In concrete, every second the system applies Fourier transform to

2^N (where N is a user-defined natural number) sensed values. We focus on the power spectrums calculated by such process. Fig.3 shows the spectrogram of the sensor value which received during all the night ($30000sec \simeq 8.5hours$). Despite from the brief observation, we can recognize some patterns of the spectrums, and would estimate the condition of the users sleep.

In this work we treat the human body wave in the frequency domains and treat the sleep stage estimation as a classification task that classify the power spectrums into the 6 categories appropriately. Note that the all power spectrum are standardized such that the summation of the all frequency components equals to 1. In addition, because of the small scale power of the high frequency regions($\geq$ 5.0Hz), we focus on only the frequency band under 5.0Hz.

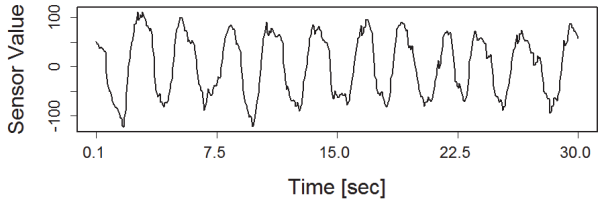


Figure 2: Sensor values of 30 sec.

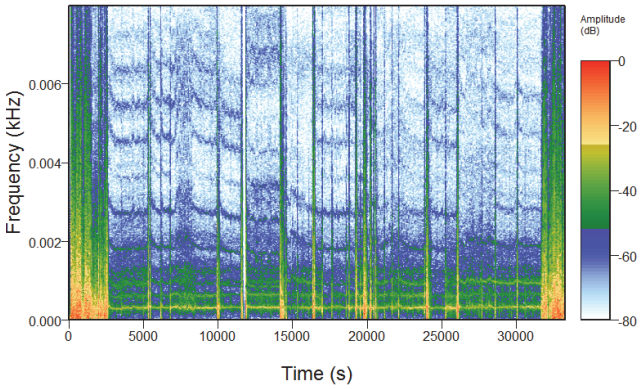


Figure 3: Spectrogram of the human subject's one night.

Classifier for Sleep Stage Estimation

As the previous subsection mentions, our proposed system treats the power spectrums which are consisted of 320 variables. To tackle such large scale data sets, a machine learning method must be implemented as the classifier of the system. Here we select a **random forest** classifier model(Breiman 2001) because of its (1) readability, (2) robustness to noise, and (3) robustness to class imbalance. **Random forest (RF)** is a method which combines some decision trees that generated by randomly selected features and samples. RF proceeding is as follows;

1. Generate N sets of bootstrapped samples S_1, S_2,...,S_N.
2. Using each S_i, grow trees T_i. Note that randomly sampled m_{try} variables are only used for branching. The growing is end when the terminal nodes appears or the depth reaches the threshold set in advance.
3. To the input, the trees T_1, T_2,...,T_N vote for the most popular class.

For the details, please refer to a thesis (Breiman 2001). RF could be used as multi-class ($i.e.$ more than 3 classes) classifier. This time, however, we combine several two-class RF classifier models as shown in Fig.4.

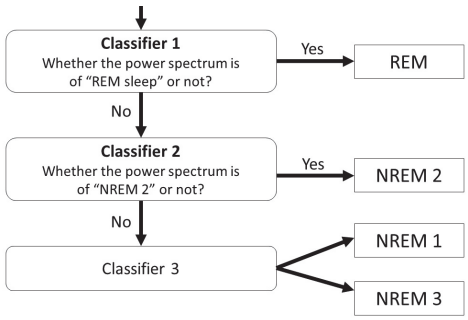


Figure 4: The proposed classification proceeding for sleep stage estimation. Each classifier is a random forest model.

Improving Classifier

Suppose that the system has the classifiers, explained in the previous subsection, and the data set S_0, initially. S_0 is the training data set of the classifiers, and is consisted of some human subjects' data. Each row is a pair of the frequency components and corresponding sleep stage which is estimated by PSG analysis. The system improves itself through reconstructing the classifiers using the S_0 which is added the user's data as a training data.

As the first section mentiones, the user's data does not have "true" sleep stage labels. Thereupon we use the Tomura's method as the approximate estimation and label the sleep stages to the additional data.

Experimental Conditions

Research Participants

The information of the participants is shown in table. No human subject has any sleep disorder and any unsatisfuction to own sleep. Each slept three nights. Some of them slept consecutively and the others did in a same week. During their sleep, the sensor was on the bed under the low repulsion mattress. At the same time we acquired each PSG, and the sleep stages were estimated by the medical technologists we offered.

Construction of Initial Data Set and Classifiers

Initial data set S_0 is constructed by the other nine human subjects. Eight of them are in their 20s and the other is in

40s. They are all male, and no sleep disorder they have. Using this S_0, initial classifiers are constructed. The parameters of each classifier ($i.e.$ random forest) are; $ntrees = 300$, $m_{try} = 24$.

Experiments

In the following experiments, we observe the classifiers be personalized. As explained in the previous section, we have three nights data sets for each human subject. Therefore, it is possible to simulate the situation as follows — *Our proposed system monitored two nights of one person, and learned him/her. Now we evaluate the estimation accuracy of the system to the 3ed night*.
We experiment the following five cases of each human subject's three-night data sets;

- **Case 1:** Test the classifiers which are trained by original data set S_0.
- **Case 2(a):** Test the classifiers which are trained by modified S_0 which is added **one day** data. The sleep stages of the additional data are results of PSG, that is, the labels are "true".
- **Case 2(b):** Test the classifiers which are trained by modified S_0 which is added **one day** data. The sleep stages of the additional data are estimated by Tomura's method, that is, the labels are"approximately true".
- **Case 3(a):** Test the classifiers which are trained by modified S_0 which is added **two days** data. The sleep stages of the additional data are results of PSG, that is, the labels are "true".
- **Case 3(b):** Test the classifiers which are trained by modified S_0 which is added **two days** data. The sleep stages of the additional data are estimated by Tomura's method, that is, the labels are"approximately true".

Considering the randomness of random forests, each case is ran three times. The results we discuss are the average.

Evaluation Criteria

As evaluation criteria, we focus on *accuracy* (Metz 1978). The measurement is defined as a ratio of the numbers of epochs which the classifier answered correctly to the whole. *Accuracy* to all sleep stages is mainly used to evaluation. Also we use *Accuracy* to REM sleep supplementally.

Table 1: The information of the human subjects.

Subject	Age	Sex
A	20	M
B	40	F
C	60	F
D	40	M
E	50	F
F	30	F
G	60	M

When calculating accuracies, PSG's WAKE estimated epochs are excluded.

Results and Discussions

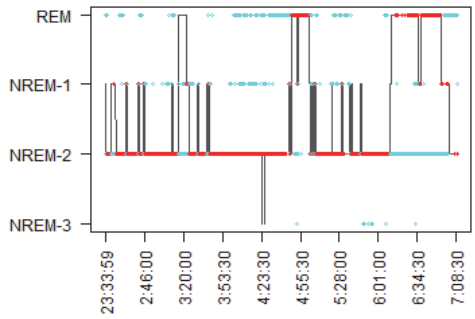
The reults are summarized in Fig.6. The left half columns of the figure indicates the accuracies to all sleep stages of each test data, the right half indicates that accuracies to REM sleep. Let us look at the first row as an example. The test data is subject A's first night. The results of Case 1 is accuracy of the classifiers which learned the original dataset S_0. Case 2 (a) is that of accuracy which learned S_0 plus subject A's one data. There are two data of A; day 2 and day 3, except of the test data. So, these results shown here is the average of the two ways. And the same things can be mentioned for the other results. The difference between (a) and (b) is the sleep stage labels of the additional data. In case (a), the labels are the results of PSG, that is, "true" sleep stages. On the other hand, in case (b), the sleep stages are labeled by the Tomura's "approximately true" estimation. In Case 3, the number of additional data is two. The difference between (a) and (b) is the same thing of case 2.

Accuracy Comparing

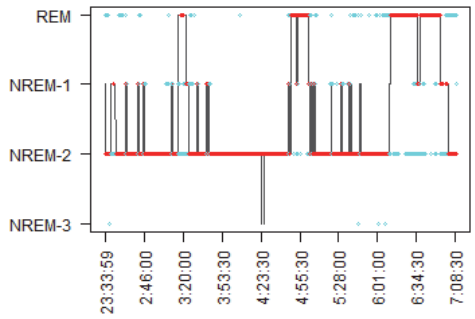
Focus on the average in the bottom row, the accuracies of the Case 2 and 3 are above that of case 1. Furthermore, according to the Wilcoxon signal rank test(Wilcoxon 1945), the differences between the mean of Case1 and of Case2, and of Case1 and Case 3 are ststistically significant ($p<0.01$). Those results indicate that the accuracy of the personalized classifiers would better performed than the original classifiers. Next we focus on the Case 3(a) and (b). The difference between those two are not statistically significant. This result reveals that there are only slightly difference between the classifiers which learned "true" sleep stages estimated data and those of "approximately true" one. In other words, Tomura's estimation method is available as supporting the classifier improvement, as in this case.
One pair of the estimated sleep stages in Case1 and 3(b) are shown in Fig.5. Comparing those two firures, we can see the wrong estimated epochs to REM sleep and NREM-2 are decreased after the two data added. Also focus on the red dots, it can be said that the accuracy depends on REM accuracy and NREM-2 accuracy. Our classifiers mostly failed to detect NREM-1 and NREM-3. The mechanism of the classifiers shown in Fig4 should be reconsidered.

Conclusion and Future Works

The aim of our research is developing the sleep stage estimation system based on the human body vibrations sensed by the commercially available mattress type in bed pressure sensors, toward the next-generation sleep monitoring technique which would supplant the current standard method, PSG. In this paper we focus mainly on improving and personalization of the system through everyday's sleep monitoring. Technically, the system adds the new sleep data by day, and modifies its classifier using that data as training



[1] The result of Case 1.



[2] The result of Case 3(b).

Figure 5: The estimated sleep stages of subject A's day 2. In those figures, the PSG sleep stages are indicated as gray lines. The red and blue dots indicates correct and wrong estimated epochs, respectively. Note that the epochs which PSG estimated as WAKE are excluded.

		Acuuracy to All Sleep Stages					Acuuracy to REM Sleep				
		Case:1	Case:2 (a)	Case:2 (b)	Case 3 (a)	Case 3 (b)	Case:1	Case:2 (a)	Case:2 (b)	Case 3 (a)	Case 3 (b)
Test data	Num. of add	0	1	1	2	2	0	1	1	2	2
	S.S. label	PSG	PSG	Approx.	PSG	Approx.	PSG	PSG	Approx.	PSG	Approx.
A : day 1		63.145%	70.163%	70.094%	73.292%	73.750%	72.587%	80.561%	80.483%	80.942%	81.102%
A : day 2		65.531%	68.673%	68.196%	77.606%	77.487%	72.366%	77.871%	77.729%	87.549%	87.411%
A : day 3		61.464%	67.609%	66.173%	67.144%	67.357%	72.577%	79.383%	78.677%	77.172%	77.184%
B : day 1		48.127%	61.992%	62.805%	71.164%	69.880%	66.622%	73.203%	74.433%	86.388%	85.564%
B : day 2		55.641%	58.771%	62.532%	65.439%	64.985%	67.318%	73.754%	77.725%	74.343%	73.562%
B : day 3		54.331%	66.193%	67.503%	65.493%	66.020%	67.888%	78.227%	80.336%	80.317%	80.425%
C : day 1		55.718%	76.721%	77.157%	46.785%	38.453%	72.113%	81.941%	82.346%	68.777%	68.917%
C : day 2		52.075%	70.659%	70.947%	76.170%	76.240%	72.337%	77.191%	77.341%	78.992%	79.094%
C : day 3		55.245%	67.170%	67.584%	78.935%	78.845%	72.542%	71.777%	72.268%	85.511%	85.337%
D : day 1		48.554%	61.266%	61.904%	53.926%	53.775%	73.012%	77.361%	76.874%	70.382%	70.789%
D : day 2		66.161%	53.834%	53.106%	68.846%	69.074%	73.724%	70.565%	70.148%	80.007%	80.393%
D : day 3		55.855%	58.662%	58.647%	55.811%	55.674%	73.741%	74.853%	74.622%	74.260%	74.285%
E : day 1		52.141%	61.284%	60.336%	56.929%	56.822%	74.573%	75.548%	75.236%	71.855%	72.185%
E : day 2		63.872%	58.034%	57.394%	69.136%	69.700%	72.428%	76.423%	76.867%	78.078%	78.268%
E : day 3		58.544%	62.692%	62.081%	59.995%	59.755%	73.552%	74.499%	74.449%	80.025%	80.478%
F : day 1		66.749%	71.061%	71.240%	79.454%	79.236%	72.086%	78.502%	78.412%	83.856%	83.768%
F : day 2		68.940%	80.410%	80.601%	66.906%	66.842%	72.387%	85.420%	85.656%	74.953%	74.527%
F : day 3		55.596%	70.604%	70.447%	81.325%	81.701%	72.122%	77.143%	77.076%	86.528%	86.708%
G : day 1		61.252%	51.706%	54.350%	55.255%	55.820%	76.527%	72.198%	74.798%	80.419%	80.776%
G : day 2		45.521%	52.604%	55.151%	53.535%	54.088%	76.391%	75.760%	79.008%	78.294%	78.620%
G : day 3		47.714%	50.939%	53.101%	54.152%	55.589%	77.859%	77.518%	79.547%	71.231%	71.892%
Average		57.247%	63.859%	64.350%	65.586%	65.290%	72.607%	76.652%	77.335%	78.566%	78.633%

Figure 6: The results summary of case1,2 and 3.

data sets. Although the approximate sleep stage estimation is needed for the system improving, we solve this problem using Tomura's estimation method. Experimental results indicates that the added data with Tomura's estimation has only a slightly difference to the data with true sleep stage. That is, effective as the improving support method. For those results and arguments, we conclude that this research provides one of the solutions for the personalized sleep stage estimation. In the next stage, (1)we have to do long-term experiments for researching the effectiveness of personalization; (2)the classification proceedings should be reconsidered; (3)we need to discover a criteria of constructing the initial data set S_0.

References

Breiman, L. 2001. Random forests. *Machine Learning* 45(1):5–32.

Kim, K.; Uchiyama, M.; Okawa, M.; Liu, X.; and Ogihara, R. 2000. An epidemiological study of insomnia among the Japanese general population. *Sleep* 23(1):41?47.

Metz, C. E. 1978. Basic principles of roc analysis. In *Seminars in nuclear medicine*, volume 8, 283–298. Elsevier.

Ministry of Health, Labour and Welfare of Japan. 2013. National health and nutrition survey.

Noh, T.; Serizawa, Y.; Kimura, T.; Yamazaki, K.; Hayasaka, Y.; Itoh, T.; Izumi, S.; and Sakai, T. 2009. The assessment of sleep stage utilizing body pressure fluctuation measured by water mat sensors. *Journal of Advanced Science* 21(1/2):27–30.

Rechtschaffen, A., and Kales, A. 1968. *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects*. Washington DC.

Tomura, S.; Harada, T.; Sato, H.; Takadama, K.; and Aoki, M. 2015. Estimating surrounding symptom level of dementia patient by sleep stage. In *Medical Information and Com-*

munication Technology (ISMICT), 2015 9th International Symposium on, 190–194.

Watanabe, T., and Watanabe, K. 2001. Estimation of the sleep stages by the non-restrictive air mattress sensor. *Transactions of the Society of Instrument and Control Engineers* 37(9):821–828.

Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics bulletin* 80–83.

Acknowledgments

The research reported here was supported in part by EMOOR Corporation and CAMPUS CREATE Corporation.