様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科研費 KAKENHI

科学研究費助成事業　　研究成果報告書

平成 ３０ 年　 ５ 月 ３１ 日現在

機関番号：１２６０１
研究種目：研究活動スタート支援
研究期間：2016～2017
課題番号：１６Ｈ０７４６９

研究課題名（和文）Research and Development of a Novel Approach for Accurate Personal Sleep Tracking

研究課題名（英文）Research and Development of a Novel Approach for Accurate Personal Sleep Tracking

研究代表者
梁　滋路（LIANG, ZILU）

東京大学・大学院工学系研究科（工学部）・特任助教

研究者番号：１０７８２８０７

交付決定額（研究期間全体）：（直接経費）　　2,300,000円

研究成果の概要（和文）：本研究の目的は、高い精度の睡眠追跡ウェアラブルデバイスの開発である。研究の結果は、以下の３点にまとめる。１．市販のデバイスは、総睡眠時間と睡眠効率の測定は医療機器によく一致する。しかし、睡眠段階の測定精度はまだ高くではない。２．エンドユーザーは睡眠および睡眠追跡装置の理解が限られている。３．一般化線形混合モデルに基づく新しい睡眠段階アルゴリズムを開発した。Fitbitの独自のアルゴリズムと比較すると、我々の新しいアルゴリズムは、より良い感度とバランスの取れた精度を達成することが判明した。これらの結果は、今後ウェアラブル睡眠追跡技術に関する研究の推進に新たな視点を与えることが期待できる。

研究成果の概要（英文）：This study aimed to develop accurate home sleep tracking technology. Our main findings are listed below. (1) Consumer sleep trackers agreed well to medical devices in measuring total sleep duration and sleep efficiency, but classifying sleep stages remains challenging for both devices. (2) In terms of user experience with sleep tracking devices, people had incomplete and faulty mental models of sleep, and there are two blind spots in people’s mental models of sleep tracking devices. In addition, they encountered challenges in reconciling device differences for assessing their sleep quality. (3) We developed a new sleep staging algorithm based on Generalized Linear Mixed Model to count in fixed and random effects during sleep tracking. Compared to the proprietary algorithm of Fitbit, our new algorithm achieved similar specificity, better sensitivity and thus higher balanced accuracy. This study provided rich implications to the design of accurate home sleep tracking technologies.

研究分野：Sleep

キーワード：Sleep　Personal informatics　Wearable computing　Quantified Self　Health informatics　HCI　Fitbit　EEG

１．研究開始当初の背景

Sleep disturbances can cause severe problems such as depression, anxiety, cancer and even morbidity. In Japan, 1 out of 5 people is suffering from bad sleep, but very few consider visiting sleep clinics, because sleep test and treatment in clinical settings is expensive and uncomfortable. Under this background, sleep research has been undergoing a paradigm shift from clinics-centered approach to self-guided approach. With the development of wearable and mobile computing technologies, several types of personal sleep tracking tools were developed, such as wearable wrist band (e.g., Fitbit), mobile apps (e.g., SleepAsAndroid), and embedded systems (e.g., Beddit). Currently, active research is going on in this field worldwide, and researchers from a diverse academic background are interested in not only developing personal sleep-tracking technologies but also evaluating the efficacy of such technologies on improving sleep.

My previous study found that personal sleep tracking plays an important role in helping individuals understand, reflect, and eventually improve their sleep. However, accuracy of consumer devices has become a bottleneck for any research effort attempting to establish values out of personal sleep tracking. For one thing, it was not clear how accurate the popular consumer devices were compared to medical devices. For another, there was a lack of understanding on how users experience the accuracy issue of consumer devices. Therefore, this study approached the accuracy issue of consumer sleep tracking devices from three aspects: the validity of popular consumer devices, users’ experience comparing consumer and medical devices, and development of new algorithm for better accuracy.

２．研究の目的

This study encompassed the following objectives: (1) to validate popular consumer sleep tracking devices against medical devices; (2) to examine end users’ experience comparing consumer sleep tracking devices with medical devices; (3) to develop new algorithms that re-engineer the data measured by consumer devices to produce more accurate estimation of sleep.

３．研究の方法

This study was conducted in three phases. In phase 1, we conducted a self-tracking experiment with 27 participants to collect sleep data using three devices concurrently: a Fitbit Charge 2 wristband, a Neuroon wearable EEG eye mask, and a portable single channel EEG Sleep Scope. The first two are consumer devices while the last is a clinical-grade sleep monitor. Each participant tracked their sleep for 3 nights in their homes and afterwards was invited for a 1-hour interview which was audio-recorded and transcribed.

In phase 2, we analyzed the quantitative sleep data and interview data collected in the sleep tracking experiment. The quantitative data was used to validate consumer devices against medical devices, and the qualitative interview data was analyzed through thematic analysis to examine end users’ experience with sleep tracking technologies.

In phase 3, we applied Generalized Linear Mixed Modelling (GLMM) to developed more accurate sleep staging algorithm based on multimodal data measured by Fitbit Charge 2.

４．研究成果

(1) Validity of Consumer Sleep Tracking Devices

Our analysis found good agreement between consumer sleep trackers and the medical device in measuring total sleep duration and sleep efficiency. In addition, Fitbit Charge 2 agreed well to the medical device on the number of awakenings, while Neuroon with good signal quality produced comparable measurements on total awake time and sleep onset latency. However, classifying sleep stages remains challenging for both devices. Both devices underestimated light sleep and overestimated deep sleep. Poor agreement was found on REM as well, which was overestimated by Fitbit but underestimated by Neuroon. As expected, Neuroon was able to accurately measure more sleep parameters than Fitbit. Since some of these parameters are important indicators of sleep disorders, Neuroon has the potential to be used for sleep disorder diagnosis in free living conditions. However, the performance of Neuroon may be significantly deteriorated by poor signal quality and disrupted sleep. Counting in other factors such as wearability and usability, Fitbit Charge 2 could be a good option for general-purpose sleep monitoring and tracking in home.

We also investigated the characteristics of the measuring errors by Fitbit Charge 2 and Neuroon under the influence of several user-specific factors. Age and sleep structure were significantly associated to the accuracy of consumer sleep trackers. Both devices had improved accuracy in measuring total sleep time and sleep efficiency for people above 26-year-old and for people with longer sleep duration, less fragmented and deeper sleep. In addition, measuring accuracy on wake time was negatively correlated to the total duration of wake, which may due to the tendency of misclassifying sleep epochs as wake. Notably, we also found that gender and subjective sleep quality measured by PSQI were not associated to the measuring errors of neither device. Our study suggested that consumer sleep trackers may be less accurate for young adults and for people with poor sleep (especially when accurate estimates of total sleep time and sleep efficiency are important.). These characteristics should be accounted for in selecting devices and in designing new sleep tracking technologies.

Results from this study advances the understanding of what consumer sleep tracking device can and cannot achieve. In the end, we highlighted three directions for future research: (1) to investigate the validity of the latest consumer devices for measuring sleep of clinical or elderly populations; (2) to clarify the classification performance of the devices epoch-by-epoch; and (3) to improve the accuracy of consumer sleep trackers on detecting sleep stage transitions.

(2) End-Users' Experience with Sleep Tracking Devices

Consumer sleep tracking technologies can help people to become more aware of their sleep habits, but these devices are also known to be inaccurate compared to medical devices. This study examined users' experience with the discrepancy between consumer devices and medical devices. Our findings provided insights into three important aspects of sleep tracking. First, we found that people had incomplete and faulty mental models of sleep. Though people understand the basic dimensions of sleep very well, their understanding of sleep stages significantly deviated from evidence-based sleep knowledge. The sleep stage fallacy may negatively impact how they set sleep goals, how they interpret sleep data, and what behavior change they plan to make for better sleep. It is therefore indispensable to help users construct a correct mental model of sleep. Second, we identified two blind spots in people's mental models of sleep tracking devices. Many participants understood some of the limitations of different types of sleep-tracking technologies. However, the devices remained black boxes for end users and led to blind spots of data collection and data processing. Third, people recognized that Fitbit was a good trade-off between accuracy and convenience. However, they encountered challenges in reconciling device differences for assessing their sleep quality.

Our findings revealed that reconciling differences in sleep data from multiple sources is less a matter of accuracy but more about reaching a harmonious view of sleep. Based on these findings, we proposed the following five design recommendations for future sleep-tracking technologies to support users in constructing correct mental models and in assessing their sleep quality: ①Provide fundamental sleep knowledge; ②Support sleep goal setting in multiple dimensions; ③Give information on how devices work; ④Give complementary information on device limitations; ⑤Establish personalized benchmark of good sleep.

(3) Accurate Sleep Staging Algorithms

Analysis in previous phase revealed that Fitbit Charge 2 is a good trade-off between accuracy and usability and thus is a promising device for home sleep monitoring. However, Fitbit does not always agree well with medical devices. Therefore, we developed a sleep staging algorithm based on Generalized Linear Mixed Model to count in fixed and random effects. The algorithm takes 30s-epoch Fitbit sleep data and heart rate data as input, and produce sleep stages as output. The medical data was used as the ground truth. In total 3749 epochs from 5 subjects were used for parameter estimation.

Compared to the proprietary algorithm of Fitbit, our new algorithm achieved similar specificity (0.97 vs 0.95), better sensitivity (1.00 vs 0.48) and thus better balanced accuracy (97.4% vs 72.7%).

５．主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 3 件）

1. Liang, Z., Chapa-Martell, M. A. (2018) Validity of consumer activity wristbands and wearable EEG for measuring overall sleep parameters and sleep structure in free-living conditions. Journal of Healthcare Informatics Research.
2. Liang, Z., Liu, W., Ploderer, B., Bailey, J., Kulik, L., Li, Y. (2017). Designing intelligent sleep analysis systems for automated contextual exploration on personal sleep-tracking data. In M. Otake, S. Kurahashi, Y. Ohta, K. Satoh, & D. Bekki (eds.) New Frontiers in Artificial Intelligence, 367-379. Berlin: Springer.
3. Liang, Z., Ploderer, B., Liu, W., Nagata, Y., Bailey, J., Kulik, L., Li, Y. (2016). SleepExplorer: A visualization tool to make sense of correlations between personal sleep data and contextual factors. Personal and Ubiquitous Computing, vol.20, no.6, pp.985-1000.

〔学会発表〕（計 15 件）

1. Liang, Z., Chapa-Martell, M.A. (2017) Are Wearable EEG Devices More Accurate than

Fitness Wristbands for Home Sleep Tracking? Comparison of Consumer Sleep Trackers with Clinical Devices. IEEE 6th Global Conference on Consumer Electronics (GCCE 2017), Nagoya, Japan.

2. Liang, Z., Chapa-Martell, M.A. (2017) Considering interpersonal differences in validating wearable sleep-tracking technologies. The 10th International Conference on Mobile Computing and Ubiquitous Networking, Toyama, Japan.
3. Liang, Z. (2017). Comparison of commercial wearable sleep trackers to subjective sleep quality. In Proceedings of the 42th Annual Meeting of Japanese Society of Sleep Research, Yokohama, Japan.
4. Liang, Z., Ploderer, B., Chapa-Martell, M. A. (2017) Is Fitbit fit for sleep-tracking? Sources of measurement errors and proposed countermeasures. In Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare, Barcelona, Spain.
5. Ploderer, B., Vyas, D., Nolan, S., Happer, C., Cunningham, T., Liang, Z. (2017). SleepBeta: engaging adolescents and their parents with sleep data. CHI workshop, Denver, USA.
6. Liang, Z., Chapa-Martell, M. A. (2017). Paradox in validating consumer wearable activity trackers on measuring sleep quality. The 16th Forum on Information Technology, Tokyo.
7. Liang, Z., Chapa-Martell, M. A. (2017). Mining wearable sleep data for detecting unusual sleep. IEICE General Conference, Nagoya.
8. Liang, Z., Chapa-Martell, M. A., Ploderer, B. (2017) Inter-individual differences in sleep quality: insights from mining wearable sleep-tracking data. IPSJ SIG Technical Report.
9. Liang, Z., Chapa-Martell, M. A. (2017) Paradox in validating consumer wearable activity trackers on measuring sleep quality. The 16th Forum on Information Technology, Tokyo.
10. Liang, Z., Chapa-Martell, M. A. (2017) Mining wearable sleep data for detecting unusual sleep. IEICE General Conference, Nagoya.
11. Liang, Z., Ploderer, B. (2016) Sleep tracking in the real world: a qualitative study into barriers for improving sleep. In Proceedings of OzCHI 2016. Tasmania, Australia.
12. Liang, Z., Chapa-Martell, M. A., Nishimura, T. (2016) A personalized approach for detecting unusual sleep from time series sleep-tracking data. In Proceedings of IEEE Int Conf on Healthcare Informatics (ICHI '16). Chicago, US.
13. Liang, Z., Nagata, Y., Chapa-Martell, M. A., Nishimura, T. (2016) Nurtuing wearable and mHealth technologies for self-care: mindset, tool set and skill set. In Proceedings of the 18th Int Conf on Networking, Applications & Services (HealthCom 2016), Munich, Germany.
14. Liang, Z., Chapa-Martell, M. A., Nishimura, T. (2016) Mining hidden dependencies between sleep and lifestyle factors from quantified-self data. In Proceedings of UbiComp '16. Heidelberg, Germany.
15. Liang, Z., Ploderer, B., Chapa-Martell, M. A., Nishimura, T. (2016). A Cloud-based Intelligent Computing System for Contextual Exploration on Personal Sleep-tracking Data Using Association Rule Mining. In Proc. of ISICS 2016, Merida, Mexico.

〔その他〕

Preprints

1. Liang, Z., Ploderer, B. (2018) "How Does Fitbit Detect Brainwaves?"-- A Qualitative Study into Sleep Tracking Experiences Comparing Consumer Devices and Medical Devices.
   https://www.researchgate.net/publication/325397519_How_Does_Fitbit_Detect_Brainwaves--_A_Qualitative_Study_into_Sleep_Tracking_Experiences_Comparing_Consumer_Devices_and_Medical_Devices
2. Liang, Z., Chapa-Martell, M.A. (2018). Not All Errors Are Created Equal: Influence of User-Specific Factors on Measuring Errors of Activity Wristbands and Wearable EEG for Home Sleep Monitoring.
   https://www.researchgate.net/publication/324389796_Not_All_Errors_Are_Created_Equal_Influence_of_User-Specific_Factors_on_Measuring_Errors_of_Activity_Wristbands_and_Wearable_EEG_for_Home_Sleep_Monitoring

ホームページ等

1. The URL of the developed web application SleepBeta
   https://sleepbeta.azurewebsites.net
2. The source code of SleepBeta
   https://www.researchgate.net/publication/324279907_Source_code_of_web_application_SleepBeta_httpssleepbetaazurewebsitesnet
3. Personal profile on ResearchGate
   https://www.researchgate.net/profile/Zilu_Liang
4. Personal profile on Google Scholar
   https://scholar.google.co.jp/citations?user=UlMDCkQAAAAJ&hl=ja

６．研究組織
(1)研究代表者
梁　滋路（LIANG, ZILU）

東京大学・大学院工学系研究科（工学部）・
特任助教

研究者番号：10782807