Title	Sleepy watch : towards predicting daytime sleepiness based on body temperature
Sub Title	
Author	包, 婕(Bao, Jie) 加藤, 朗(Katō, Akira)
Publisher	慶應義塾大学大学院メディアデザイン研究科
Publication year	2020
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2020年度メディアデザイン学 第791号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002020-0791

Master's Thesis
Academic Year 2020

Sleepy Watch: Towards Predicting Daytime Sleepiness Based on Body Temperature

Keio University
Graduate School of Media Design

Jie Bao

A Master's Thesis
submitted to Keio University Graduate School of Media Design
in partial fulfillment of the requirements for the degree of
Master of Media Design

Jie Bao

Master's Thesis Advisory Committee:

Professor Akira Kato	(Main Research Supervisor)
Professor Kai Kunze	(Sub Research Supervisor)

Master's Thesis Review Committee:

Professor Akira Kato	(Chair)
Professor Kai Kunze	(Co-Reviewer)
Senior Assistant Professor Junichi Yamaoka	(Co-Reviewer)

Abstract of Master's Thesis of Academic Year 2020

Sleepy Watch: Towards Predicting Daytime Sleepiness Based on Body Temperature

Category: Science / Engineering

Summary

In our society, daytime sleepiness has become an epidemic because of lack of sleep. Daytime sleepiness means difficulty in maintaining an alert waking state. It causes vehicle accidents and has adverse effects on well-being, productivity.

Predicting daytime sleepiness is a direct way to solve the issue and help people to manage their energy better. There are considerable research trying to solve drivers' sleepiness. This study aims to figure out using wrist temperature to predict objective daytime sleepiness in wild-life. Experimental results of this study present that the wrist temperature has the potential to play a valuable part in better understanding daytime sleepiness.

Keywords:

Daytime Sleepiness, Objective Sleepiness, PVT, Wrist Temperature, Body Temperatures

Keio University Graduate School of Media Design

Jie Bao

Contents

List of Figures

List of Tables

Acknowledgements

First of all, I would like to thank my supervisor Prof. Akira Kato for his sincere help and guidance on both research and life during my KMD life. I am grateful for his support. I would also like to thank my Sub supervisor Prof. Kai Kunze and Prof. Hideki Sunahara for giving me invaluable advice and informative suggestions. I am also indebted to my friend Jiawen Han for helping me with data analysis and discussion when I met trouble during busy times. Without their help, encouragement, and guidance, I could not complete this paper.

My thanks also go to all the professors, teachers, and schoolmates who have taught me and communicate with me during the past two and a half years. I also want to thank all of my friends, without you guys I can not collect any data, especially in this COVID-19 period.

Last but not least, thanks to my parents, who always support me and encourage me during my study

Chapter 1
Introduction

Keio University Graduate School of Media Design (KMD)[1] was established

1.1. Research Question

1.1.1 Background of the Research Questions

In our society, lack of sleep has become an epidemic. Insufficient sleep is normally associated with disturbances in cognitive and psychomotor function which would gradually develop significant impact on mood, thinking, concentration, memory, learning, vigilance and reaction time. [1]These disturbances also have adverse effects on wellbeing, productivity and personal safety. Insufficient sleep is a direct contributor to injury and death from motor vehicle and workplace accidents. [2] According to Psychiatry Research [3], in general, the sleep duration of adult population in Japan is relatively short. Sleep problems are common and comparable to those reported in Western countries. There are various causes for daytime sleepiness , but lack of sleep at night is usually the most common reason. Moreover, so far, the outbreak of COVID-19 has been almost 6 months around world. Lockdown life force people to stay at home and work from home, with minimal social interactions allowed.People have to adjust themselves mentally and physically, in order to adapt to the disruption of regular daily routine. These disruptions, have no doubt, also affect their sleeping hours. Research shows, during COVID-19, the general public has developed poor sleep hygiene habits, and sleep problem is more severe in people who are female, or young [4].

Daytime sleepiness means difficulty in maintaining an alert waking state. It

1 http://www.kmd.keio.ac.jp

has been operationally defined as a physiological need [5]. Sleepiness and hours of sleep are inversely associated, this can be simply explained as less sleep leads to more sleepiness [6]. As mentioned previously, in certain activities, such as driving, sleepiness is considered as a significant risk factor that substantially contributes to the increasing number of motor vehicle accidents each year [7]. Critical aspects of driving impairments is normally associated with drowsiness, slow reaction times, reduced vigilance, and deficits in information processing that could possibly all lead to an abnormal driving behavior [2, 8]. Besides traffic accidents, sleepiness also leads to decline of productivity, accidents and health deterioration. Huge economic loss caused by sleepiness was reported in the recent studies [9, 10]. For detecting drivers drowsiness, there have been numerous attempts to resolve the issue with sensible solutions. However, studies on sleepiness in daily life are still very limited. If daytime sleepiness can be rigorously measured and predicted automatically, it would bring a huge impact to our daily life.

1.1.2 Definition of Research Questions

I would like to investigate and discuss two main research questions in this thesis for predicting people's daytime sleepiness by wrist temperature. The first and the most important one is to explore if there is any correlation between body temperatures and daytime drowsiness. To establish the baseline for the study, we need to understand the most suitable body temperature for further experiments. Various thermometers are used. All experiment tools will be introduced and compared in detail in the next chapter. In addition, it is also important to compare different methods of measuring sleepiness which includes objective way and subjective way. After comparing these methods, Psychomotor Vigilance Task(PVT) was selected to measure sleepiness in this research.

Predicting daytime sleepiness is the second research question. According to the result of main question, it is important to find a method to predict daytime sleepiness.

1.2. Contribution

In this thesis, possible tools and devices that could be used to solve the core problem will be introduced and discussed, which are those that could measure and record body temperature, and those which are usable but not suitable in some way. Those would be introduced in the second chapter which would talk about those related works.

Then to present our data set along with PVT measurements and body temperature, which amount to 19 days of wrist temperature data and ground truth assessments from 8 participants collected in the wild.

1.3. Thesis Structure

This thesis consists of 6 Chapters.

- This chapter depicts the introduction of the thesis.
- Chapter 2 introduces an overview of how sleepiness are measured in previous works and the motivation of this work.
- Chapter 3 presents the approach of this study and introduces research materials used in this study
- Chapter 4 expresses the experiment set up and the results of the collected data
- Chapter 5 presents the analysis of the data result from chapter 4 and lists the possible limitations of this study.
- Chapter 6 concludes this work by summarizing all the works has been done for this research topic and discussing about possible future plan.

1.4. Key terms

PVT - psychomotor vigilance task
WT - Wrist Temperature

Chapter 2
Related Works

2.1. Sleepiness measurement

With the increasing complexities of life tasks and high speed of life paces in modern life, it is not difficult to image that sleepiness has become a big issue nowadays. Studies suggest sleepiness is a problem reported by 10% to 25% of the population. It occurs more frequently in young adults and in older adults [5]. There are many different kinds of scales designed to measure sleepiness. Most of them can be grouped to the subjective measurement methods and the objective measurement methods. The subjective measurements usually involve validated self-rated scales, such as the Stanford Sleepiness Scale(SSS) [11] and the Epworth Sleepiness Scale (ESS) [12]. While some objective measurements method use physiologic measures, for example, the Multiple Sleep Latency Test [13] and Psychomotor Vigilance Task (PVT). Both of these two measurements are used to assess the presence and degree of sleepiness. We will discuss more in the following sessions.

2.1.1 Subjective sleepiness

Various subjective measurement methods are carried out to evaluate sleepiness. For example, the the visual analog scale(VAS) [14], the Stanford sleepiness scale(SSS), the Epworth sleepiness scale(ESS)and the Karolinska Sleepiness Scale(KSS) [15]. The most well known subjective sleepiness scale are the Stanford Sleepiness Scale(SSS) and the Epworth Sleepiness Scale(ESS).

The Stanford sleepiness scale (SSS) is a quick and simple test. It involves the subject's own reports of symptoms and feelings at a particular time. However, this test does not attempt to measure the general level of daytime sleepiness, as distinct from feelings of sleepiness at a particular time [16]. Scores on the SSS of sleepiness are not significantly correlated with sleep latency in the MSLT, even when measured at virtually the same time [17]. These subjective reports may be related more to tiredness and fatigue than to sleep propensity, as manifested by the tendency to fall asleep.

For measuring sleep propensity in a simple, standardized way, Murray W. Johns designed the the Epworth sleepiness scale (ESS). The scale covers the whole range of sleep propensities, from the highest to the lowest [18].

However, the best screening tool is still matter of debate. One example of contesting the reliability of the ESS would be the study "Is the Epworth Sleepiness Scale a useful tool for screening excessive daytime sleepiness in commercial drivers?" [19]. In this study, Simone Baiardi got results from 221 commercial drivers, only ten (4.5%) had Epworth Sleepiness Scale scores indicative of excessive daytime sleepiness. These findings and the lack of concordance in estimating excessive daytime sleepiness among commercial drivers in previous studies using the same psychometric measure indicate that the Epworth Sleepiness Scale is not a reliable tool. This may be due to the low internal consistency of the scale in non-clinical samples and the possible intentional underscoring of sleepiness due to a perceived threat of driver ' s license suspension. Moreover, the reliability of the Epworth Sleepiness Scale results may be strongly influenced by the administration setting. Some other studies also showed that measures have limitations. Individuals are not always aware of their degree of sleepiness or their susceptibility to impairment, with significant inter-individual differences in performance described following sleep deprivation [20–22]. The clinical application of inexpensive less time-consuming new tools like performance tests should be considered for the objective evaluation of excessive daytime sleepiness in commercial drivers.

2.1.2 Objective sleepiness

The Psychomotor Vigilance Task (PVT) is widely used in assessing behavioral alertness and sustained attention related to sleepiness induced through sleep deprivation [23, 24]. The psychomotor vigilance test (PVT) and divided attention driving task (DADT) focus on neurocognitive function. This is especially relevant among commercial drivers who undertake the complex task of driving [25, 26]. Researchers also found that single Divided Attention Driving Task(DADT) and

PVT administrations are reliable measures of sleepiness. These results support the use of a single administration of some objective tests of sleepiness when performed under controlled conditions in routine clinical care. Another study indicated that PVT performance significantly predicted specific aspects of simulated driving performance. Thus, psychomotor vigilance impairment may be a key cognitive component of driving impairment when sleep deprived [27]. Above all, while not direct test of physiological sleepiness, the PVT has been shown to be sensitive to sleepiness.

2.2. Automatic Sleepiness Detection

For detecting drowsiness, there are many studies are physiological based, vehicle based, and behavioural based [28, 29]. Physiological methods such as heartbeat, pulse rate, and Electrocardiogram(ECG)etc. are used to detect fatigue level [30, 31]. Vehicle based methods include accelerator pattern, acceleration and steering movements. Behavioural methods [28,29]include yawn, Eye Closure, Eye Blinking, etc. Most of the traditional methods are based on behavioural aspects while some are intrusive and may distract drivers, while some require expensive sensors.

2.2.1 Sleepiness Application Used on Android System

It is important to develop a light-weight, real time driver's sleepiness detection system. One example of Detecting sleepiness by android application would be the study" Real-Time Driver Drowsiness Detection System Using Eye Aspect Ratio and Eye Closure Ratio" [32]. The study focused on developed a system which is able to detect driver's facial landmarks(Figure 2.1), computes Eye Aspect Ratio and Eye Closure Ratio. It can detect driver's drowsiness based on adaptive thresholding. The result showed that it is useful in situations when the drivers are used to strenuous workload and drive continuously for long distances. The facial landmarks captured by the system are stored and machine learning algorithms have been employed for classification. The system gives best case accuracy of 84% for random forest classifier.

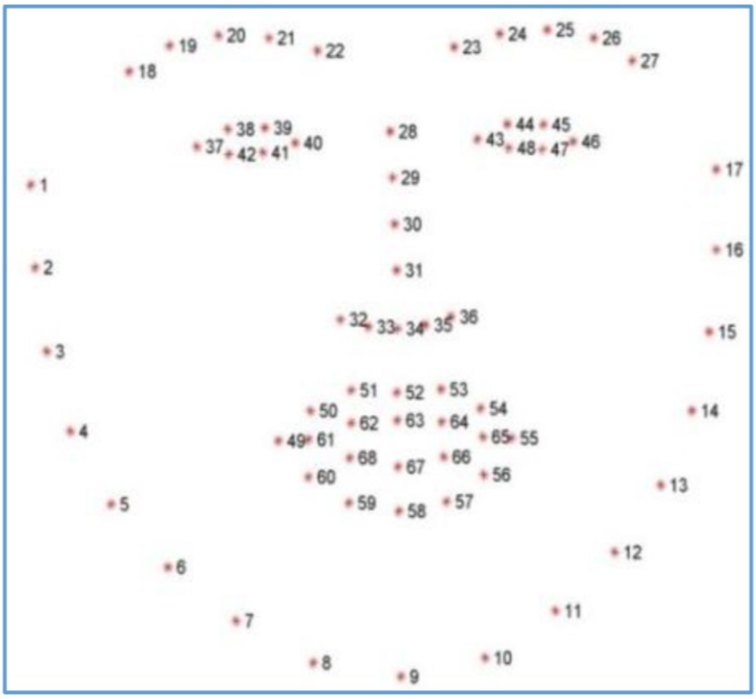

Figure 2.1 Facial Landmark Points the study used according to Dlib library

Figure 2.2 The wearable respiration sensor

2.2.2 Daytime Sleepiness Level Prediction

The study" Daytime Sleepiness Level Prediction Using Respiratory Information" [33] is an example of continuous sleepiness tracking in daily living situation. The study focused on predicting subjective sleepiness levels utilizing respiration and acceleration data obtained for a novel wearable sensor [34](Figure 2.2). The study contained two main parts with a five-minute rest in between. Two main parts are typing task and video watching task. They collected the physiological data of the subjects doing both active and passive tasks which represent activities in everyday life, contrast the past studies which concentrated on driver's sleepiness. Their dataset and derived models reflected the actual association of the physiological information and sleepiness in daily living more precisely.

2.3. Circadian Rhythm Monitoring by Temperature sensors

Recently skin temperature circadian rhythms have also been explored. The research "Relationships between the Circadian Rhythms of Finger Temperature, Core Temperature, Sleep Latency, and Subjective Sleepiness" [35] is an invesitgation about the circadian finger and core temperature rhythms in conjunction with the circadian rhythms of subjective and objective sleepiness. It(Figure 2.3) presents the core body temperature(CBT) minimum temperature adjusted group mean(SEM) curves of CBT, finger temperature, rectified SOL, and rectified Stanford Sleepiness Scale(SSS). They found that the maximum 5 possible cross-correlation curves between palmar finger temperature, rectal temperature, subjective sleepiness and objective sleep latency suggested that finger temperature preceded core temperature by 3h(r=-0.22), and subjective sleepiness followed core temperature by 0.5h(r= -0.33) and objective sleepiness by 2h(r= 0.39). Although these data are correlational, they are consistent with the notion that core temperature changes are driven by finger temperature changes, which determine changes of subjective and objective sleepiness. The research indicated distal skin temperature helps to reduce CBT, which in turn promotes subjective and objective sleepiness.

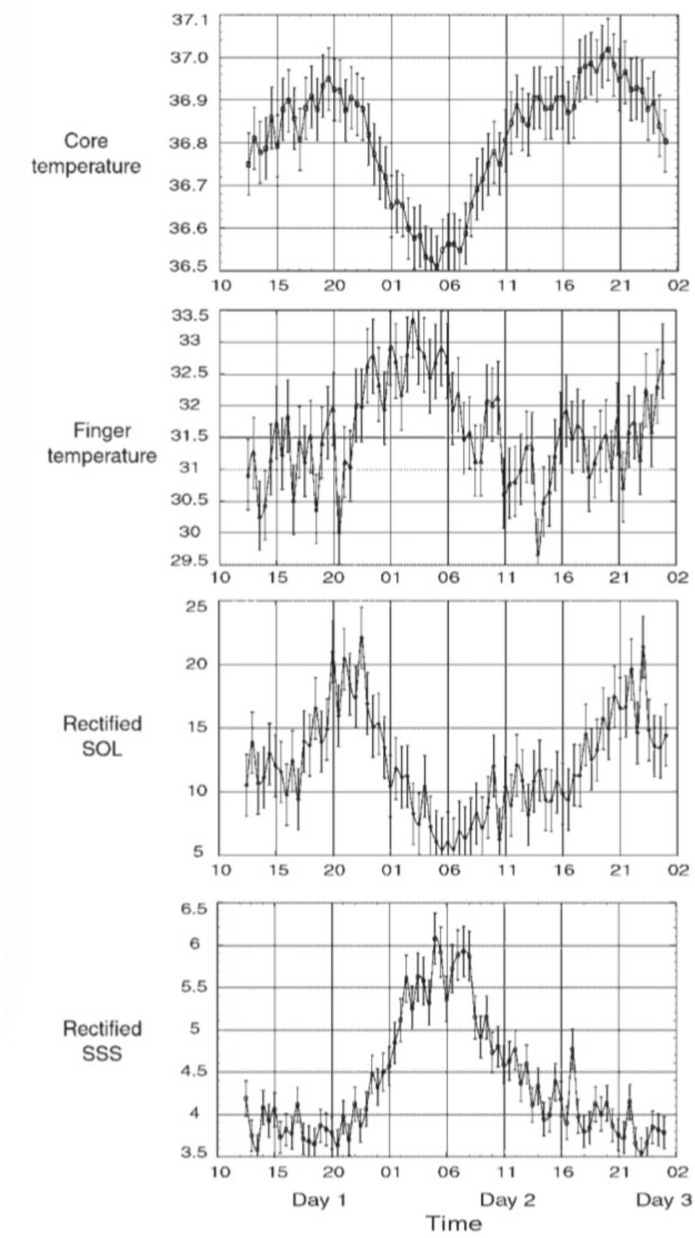

Figure 2.3 Circadian rhythms of core temperature (˚ C), finger temperature (˚ C), rectified sleep onset latency (SOL) (min), and rectified Stanford Sleepiness Scale (SSS) presented as means ± 1 SEM per 30 min

Another research of exploring the wrist temperature patterns and their underlying implications for both circadian rhythms and sleep patterns for people of different ages and cognition is "Monitoring Circadian Rhythm and Sleep Patterns Using Wrist-worn Temperature and 3-axis Accelerometer Sensors: A Study with Healthy Younger Adults, Healthy Older Adults, and People Living with Dementia" [36].Their participants wore a customized wristband with a temperature sensor and a three-axis accelerometer sensor(Figure 2.4) along with a commercial wristband(Mi Band 2)(Figure 2.5) for 14days. They analyzed wrist temperature rhythms and compared for the three groups and found differences in daytime value and variation. As a conclusion, they found while the wrist temperature alone algorithm performed better than the Mi Band for older adults with dementia, using both data sources showed increases in sleep detection accuracy for all participants. That shows the wrist temperature has the potential it to play a valuable role in better identification and understanding of sleep including for people with movement-related sleep disorders.

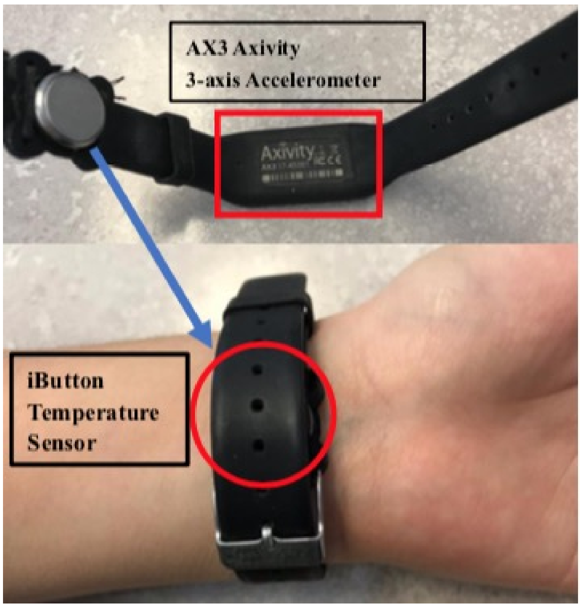

Figure 2.4 The study used prototype wristband

Figure 2.5 The Mi Band 2

Chapter 3
Proposal

I propose a sleepiness monitoring system for people who are lack of sleep in this thesis. This chapter will focus more on the reasoning behind the selected prototype decisions, and also introduce the main system.

As the most important issue to solve is to continuously record body temperature data for the whole day and to investigate the correlation between sleepiness level and body temperatures.

3.1. Approach

As introduced in the chapter Related Works, there are two kinds of sleepiness measurements which are objective sleepiness measurement and subjective sleepiness scales. The experiments were designed in objective way which is PVT. Besides sleepiness measurement, there are two part of body temperatures' data should be collected. The experiments included both and there are the process and conclusion below.

3.1.1 PVT

I'd like to introduce the most important tool be used first. According to previous study, objective and quantitative assessments are necessary to evaluate the presence of fatigue-related deficits, especially as self-reports of sleepiness and self-assessments of performance capability have been shown to be unreliable [37, 38]. In this study, we used 1-min PVT(Figure 3.1) to asses participant's sleepiness by their reaction time. The PVT application can work on both android and IOS system[1].

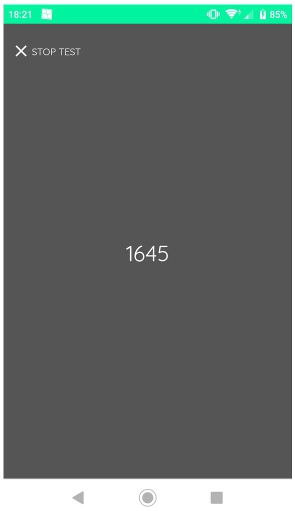

Figure 3.1 PVT test we used

A large number of performance tests have been developed to objectively assess the degree of cognitive performance deterioration related to sleep loss. The PVT is widely used among those [39], [24]. It is based on simple reaction time to stimuli that occur at random intervals and therefore measures vigilant attention [40]. Most commonly used version is the standard 10-min PVT with 2-10s inter-stimulus intervals, although both longer [41, 42]and shorter [43] duration versions have been evaluated. Test duration is an important aspect of the PVT because even severely sleep deprived subjects may be able to perform normally for a short time by increasing compensatory effort. However, in a systematic analysis of PVT duration, the study "Validity and sensitivity of a brief psychomotor vigilance test (PVT-B) to total and partial sleep deprivation" [44] showed that the ability of the PVT to differentiate alert and sleepy subjects was, depending on the outcome variable, only marginally lower (and at times higher) for shorter than 10-min test durations [39].Therefore, optimal PVT duration may be shorter than 10-min for some outcome variables, demonstrating feasibility of shorter ver-

1 Vigilance Buddy 1.53
`https://researchbuddies.com`

sions of the PVT. Accordingly, a 5-min handheld version of the PVT already exists [43], [45], [46], [47], [48]. However, both 2-min [43] and 90s [46] versions of the PVT were supposed to be too insensitive to be used as valid tools for the detection of neuro behavioral effects of fatigue. In 2011, the study "Validity and sensitivity of a brief psychomotor vigilance test (PVT-B) to total and partial sleep deprivation" [44] showed that PVT(3min) tracked standard 10-min PVT performance and yielded medium to large effect sizes. 3-min version PVT may be a useful too for assessing behavioral alertness in settings where the duration of the 10-min PVT is considered impractical.

Because all of participants in this study had to do PVT test every hour during their awake time, it is challenging to do the standard PVT which lasts 10 minutes at least. The shortest PVT which has been assessed as a useful tool is 3-min version. We did a pre-test to explore if 1-min PVT can be handling in this study. There were 4 participants(3 female, 1 male) of all participants joined this pre-test. After they completed 5-min PVT, we compared their performance between first minute and the whole 5 minutes.

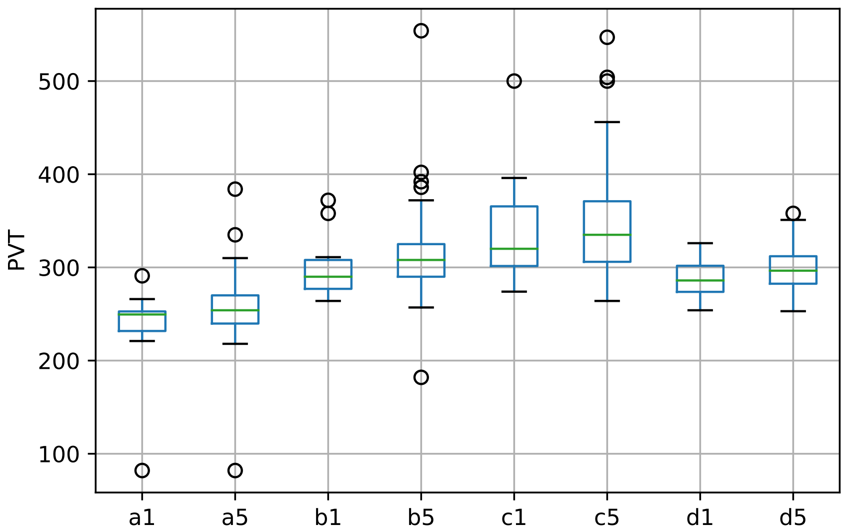

Figure 3.2 PVT result between the 1st minute and the whole 5 minutes

From figure above (Figure 3.2), we can find the mean value in the first minute is faster than the 5 minutes result a bit even included more outlier value in the whole 5 minutes.

Table 3.1 describe pvt in 1min and 5mins

	a1	a5	b1	b5
count	18.000000	92.000000	17.000000	85.000000
mean	238.500000	255.543478	296.235294	313.164706
std	42.380975	31.056635	29.752163	42.180690
min	82.000000	82.000000	264.000000	182.000000
25%	231.750000	239.750000	277.000000	290.000000
50%	249.500000	254.000000	290.000000	308.000000
75%	252.750000	270.000000	308.000000	325.000000
max	291.000000	384.000000	372.000000	554.000000

Table 3.2 describe for pvt in 1min and 5mins

	c1	c5	d1	d5
count	19.000000	89.000000	20.000000	92.000000
mean	337.947368	343.617978	290.050000	298.054348
std.	53.391087	50.040099	19.494871	21.129306
min	274.000000	264.000000	254.000000	253.000000
25%	301.500000	306.000000	273.750000	282.500000
50%	320.000000	335.000000	286.000000	296.500000
75%	365.500000	371.000000	301.750000	312.000000
max.	500.000000	547.000000	326.000000	358.000000

Table 3.3 T-test for pvt in 1min and 5mins

participant	statistic	pvalue
A	8.6097850191096070	0.003343617659587512
B	10.810245442292791	0.0010093991946557028
C	0.37998343088756237	0.5376121039897125
D	0.5859870861498245	0.44397509653377565

We can see more clearly in Figure3.1(Table 3.1)and Figure3.2(Table 3.2). The t-test result is as(Table 3.3). The result shows that 2 participants' p-value(C and D) >0.05, which means under these experimental conditions, a significant difference could not be detected. The left two p-value<0.05 mention that they are near-marginal significance. Two participants performed better in the first minute than the whole five minutes. By reason of that we don't want to distract participants redundantly, it is acceptable to use the 1-min PVT. The other setting is as below(Figure 3.3).

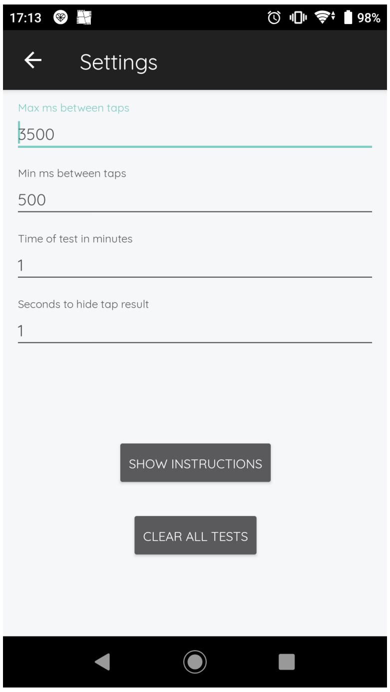

Figure 3.3 The settings of PVT

3.1.2 Temperature sensor

The other data we have to collected is body's temperature. According to the research" Relationships between the Circadian Rhythms of Finger Temperature, Core Temperature, Sleep Latency, and Subjective Sleepiness" [35], distal skin temperature helps to reduce CBT, following promotes sleepiness. We decided to collected skin temperature with core temperature.

For monitoring temperature continuously, we tried to use FeverScout(Figure 3.4)[2] at first. The sensor could be pasted to everywhere of body. Temperature data could be sent to its own application via bluetooth.

Figure 3.4 FeverScout

However, the body temperature data was measuring only while we were tapping

2 FeverScout
`http://www.vivalnk.com/feverscout`

the application on our smartphone (Figure 3.5).Therefore, we can see there are 3 temperature data from 7am to 2pm. The first time temperature was measuring after 7am, then the second time was around 8:30am. The third time was 11:30am. There was no data collected between the second time and third time. That means the Feverscout can not monitor body temperature automatically. Besides that, the temperature data can not show out if the temperature below 37°C. So the conclusion of this experiment is that the FeverScout is unusable in this study.

Figure 3.5 Body Temperature Data on the app of FeverScout

Here is another solution, iButton temperature sensor DS1922L[3], which was also used in the research "Monitoring Circadian Rhythm and Sleep Patterns Using Wrist-worn Temperature and 3-axis Accelerometer Sensors- A Study with Healthy Younger Adults, Healthy Older Adults, and People Living with Dementia" [36]. As seen in figure3.6(Figure 3.6), it was fixed at the inside of wrist by a tennis band. Temperature accuracy of ± 0.5 ° C is from -10 ° C to +65 ° C. The DS1922L is configured and communicates with a host computing device through the serial 1-Wire protocol[4]. In 2009, the research"The validity of wireless

3 iButton Temperature Loggers with 8KB Data-Log Memory `https://www.maximintegrated.com/en/products/ibutton-one-wire/data-loggers/DS1922L.html`

4 DS9490B USB to 1-Wire/iButton Adapters

iButtons® and thermistors for human skin temperature measurement" [49] also announced wireless iButton provide a valid alternative for human skin temperature measurement during laboratory and field investigations particularly when skin temperature measurement using other currently available methods may prove problematic. So we decided to use iButton to monitor wrist temperature. The setting in this study is monitoring wrist temperature every minute.

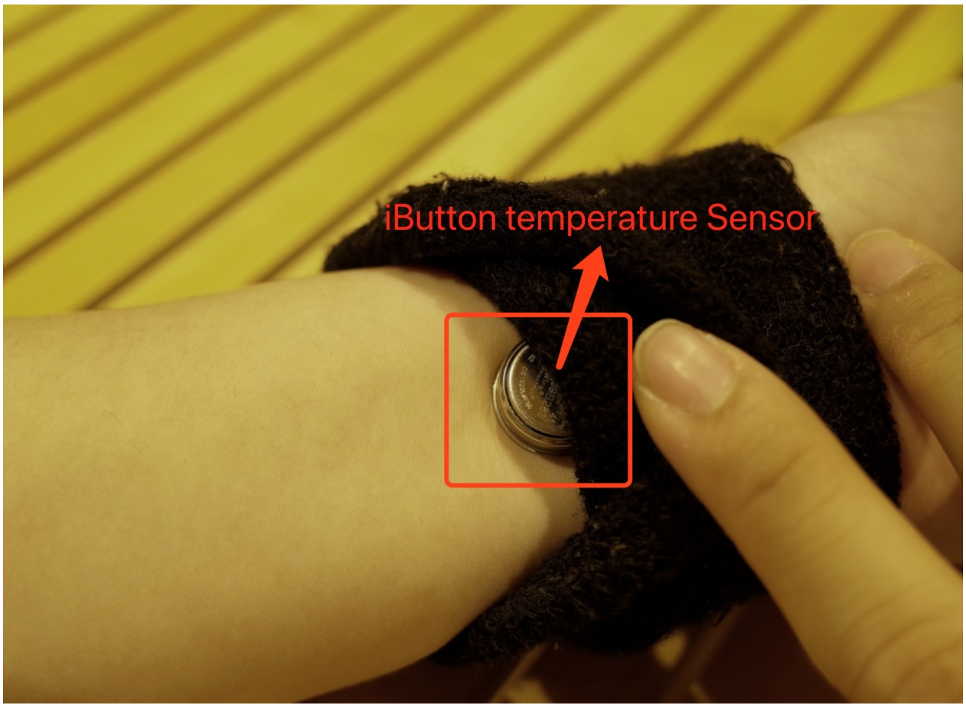

Figure 3.6 iButton Temperature sensor

Another thermometer[5] we used in this study is a commercial stuff. It is an infrared medical thermometer which measures temperature in the ear(Figure 3.7) and forehead(Figure 3.8). The body temperature display range is 32.0 ° C to 42.9 ° C, and the minimum unit is 0.1 ° C. All of participants measured their ear and forehead temperature every hour after they completed PVT.

https://datasheets.maximintegrated.com/en/ds/DS9490-DS9490R.pdf

5 Infrared Medical Thermometer
https://www.dretec.co.jp/product/thermometer-to-300/lang-en

Figure 3.7 Measuring ear temperature

Figure 3.8 Measuring forehead temperature

3.2. Target User and system scenario

The target users for this work can be generally classified as people who could greatly benefit from having insufficient sleep. The main target users that the study would be tailored to are college students and office workers. With productivity being the main focus of students' school life and office worker's workday, it is important to know when their sleepiness level rise to they can not concentrate their attention on work. Then the users can change their activity to get a short nap or get a cup of coffee. The age range for this study would be around 20-50 years old males and females.

As discussed in the second chapter, a lot of daytime sleepiness level detection systems have already existed, and most of them are based on subjective sleepiness scale or behavioural activity. The system that in this thesis proposes aim to be automatic, low cost and could be check real-time sleepiness. The system also aims to build a database for the users, to help them to have the ability to get the data of low sleepiness level time and high sleepiness level time every day, and if possible, also help them analyze and predict data to help the users to manage their energy better.

Chapter 4
Experiment

4.1. Methods

4.1.1 Participants

Eight participants(five females and three males) of this study were recruited from colleges. Because of limitation of sensor's storage, we have effective data from seven participants (five females and two males) finally. All participants with effective data were healthy young adults(23 - 32 years old). The mean of their age was 28.43 years old.

4.1.2 Experimental process

All participants had an one-on-one orientation session. After providing informed consent, all participants were asked to download PVT application. To avoid the effect by adjusting process, before the formal test, every participant tried it in 20s. Then all participants were introduced to the study sensors: (1) iButton Ds1922L temperature sensor(Maxim, Dallas, US), (2) infrared medical thermometer(Dretec, JP) and (3) Mi Band 4. Mi Band 4 is used for recording participants' sleep length before the test day. Participants were asked to wear iButton temperature sensor on their left wrist since they woke up to go to sleep which means wearing the sensor for a whole day. All participants kept on staying at home without any vigorous exercise or activities on test day. Wrist temperature measuring frequency is every minute.

Besides wearing iButton sensor on the left wrist by a tennis band, all participants were asked to do 1-min PVT once per hour from they woke up to get to sleep. After completed PVT, they were requested to measure and record their forehead temperature and ear temperature every hour by themselves. Participants were required to record their lunch, dinner time and other special status such as having a nap or drinking a cup of coffee.

The test started February 2020 to June 2020 in Tokyo. Because of the participants' schedule, most of the participants completed the test not in continuous days. In this experiment, we collected 19 days data in total. For calculating the individual's correlation, we used all the data we collected. In another analyzing process, when the data collected by participants is more than 2 days, we chose 2 days every participant to keep data volume same from different participants. The choosing standard is that the day with more data (sometimes participants forgot to measure temperature of ear or forehead after they completed PVT) has higher priority. If more than 2 days are having same data volume, we chose the date nearer to the other participants' to avoid the effect by season. While they are still similar, we chose the date randomly. Finally, we have 14 days data from seven participants to predict.

4.2. Results

A goal of this research was to investigate if there is a correlation between wrist temperature and objective sleepiness. This section shows the result collated with the research questions listed up in Chapter 1. The content is focused on the data measured from experiment. A visual qualitative analysis of the correlation is presented. In additionally, this section shows the predicting accuracy of sleepiness by wrist temperature data and all of body temperature data.

PVT has four features included which are pvt, pvt mean, pvt max and false(it means tap the screen before the signal appeals). Though most commonly used value is pvt mean. We did a Pearson correlation test between pvt mean with pvt max and pvt min. The results is as(Figure 4.1)(Figure 4.2). For pvt mean and pvt max, $r = 0.480$, $p = 0.000$. For pvt mean and pvt min, $r = 0.754$, $p = 0.000$. It indicated pvt mean has a strong significant correlation with pvt max and pvt min, even included outlier value. It is acceptable to use pvt mean as the pvt main feature in the followed process.

A Pearson correlation test was used to assess the correlations between wrist

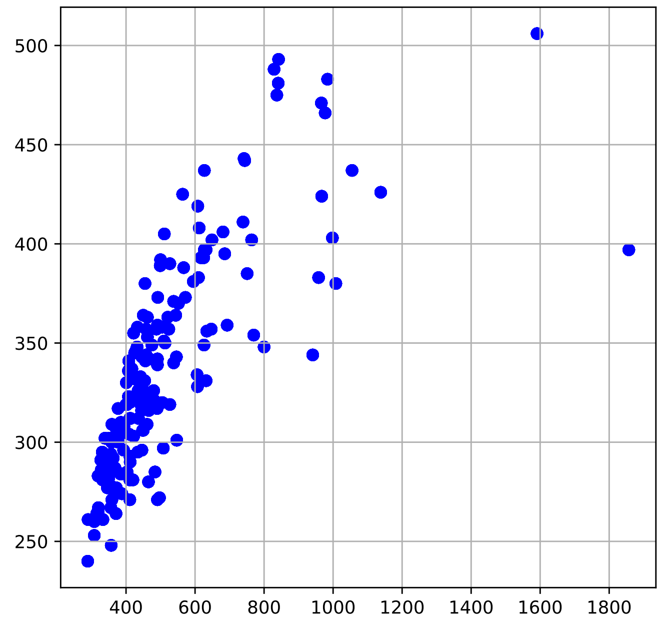

Figure 4.1 The correlation between PVT mean and PVT max

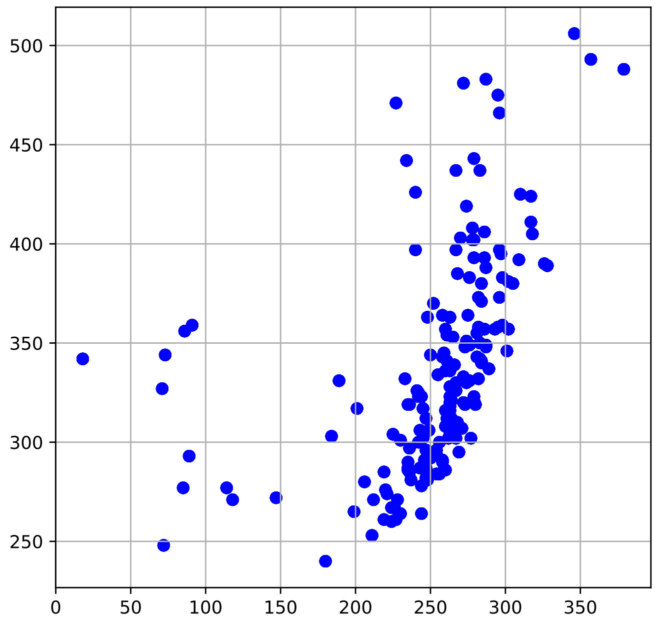

Figure 4.2 The correlation between PVT mean and PVT min

temperature and PVT, and a t-test was used to compare wrist temperature and PVT between different sleepiness level group. p<0.05 was considered to be significant.

4.2.1 Correlation between body temperature and PVT

As the figure (Figure 4.3), correlation between wrist temperature and PVT mean value: pvalue = 0.497, r = -0.046, which means there is no linear relationship between wrist temperature and PVT.

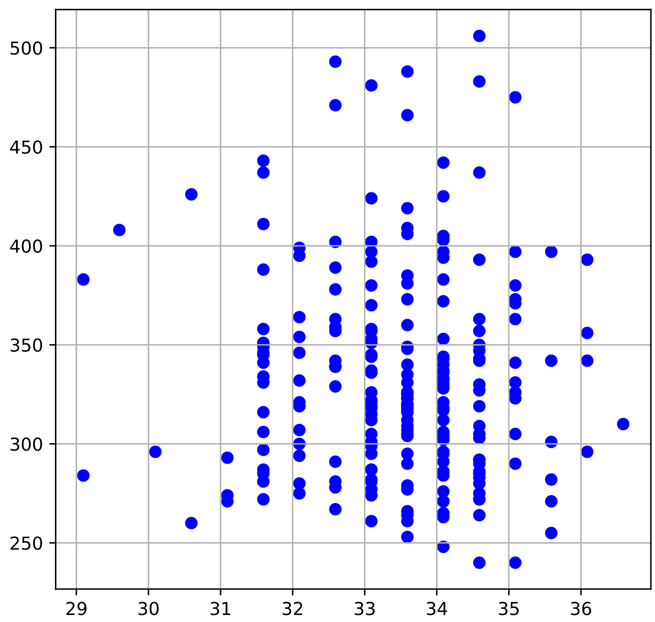

Figure 4.3 The correlation between PVT and WT

As the (Table 4.1), 3 of 7 participants' data indicate a moderate linear relationship between the variables. The visualization result is present as(Figure 4.2.1),(Figure 4.2.1),(Figure 4.2.1),(Figure 4.13). Their data is showed as below(Figure 4.4),(Figure 4.5),(Figure 4.6).Compare with in-ear and forehead temperature, the wrist temperature fluctuated more frequently and to a larger extent. Green dotted line is wrist value at the same time with in-ear and forehead which is measured every hour. Another line is wrist temperature data recorded every minute. We can find the trend of wrist value every minute is similar to every hour.

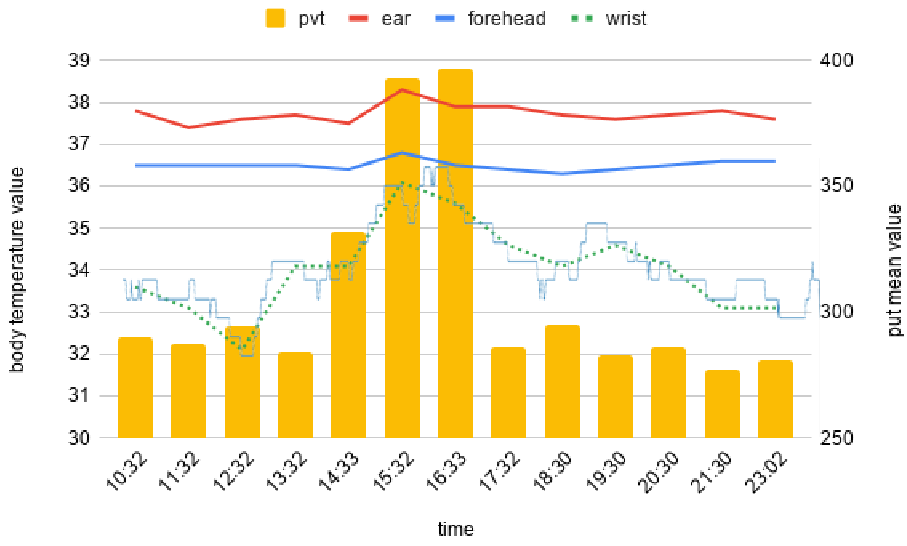

Figure 4.4 Body temperatures and pvt value for f2 on Feb.13th

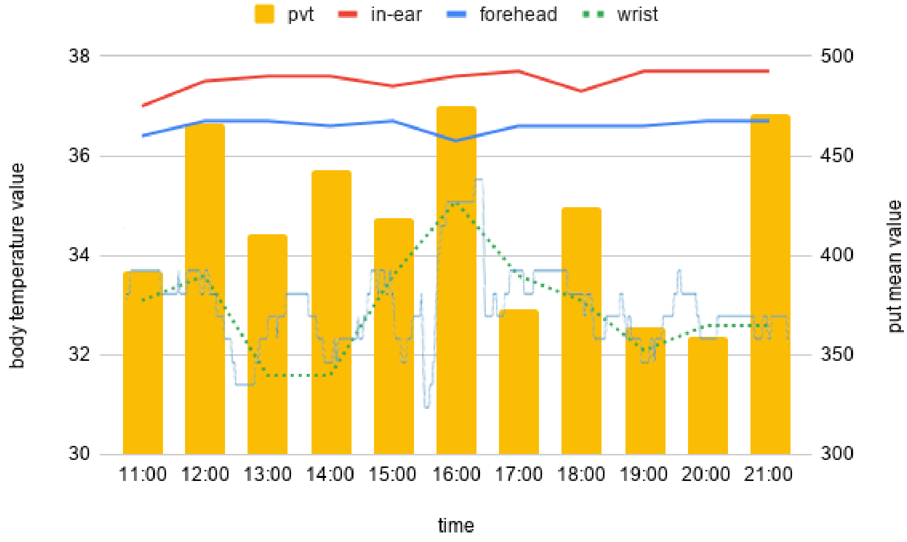

Figure 4.5 Body temperatures and pvt value for f2 on Apr.5th

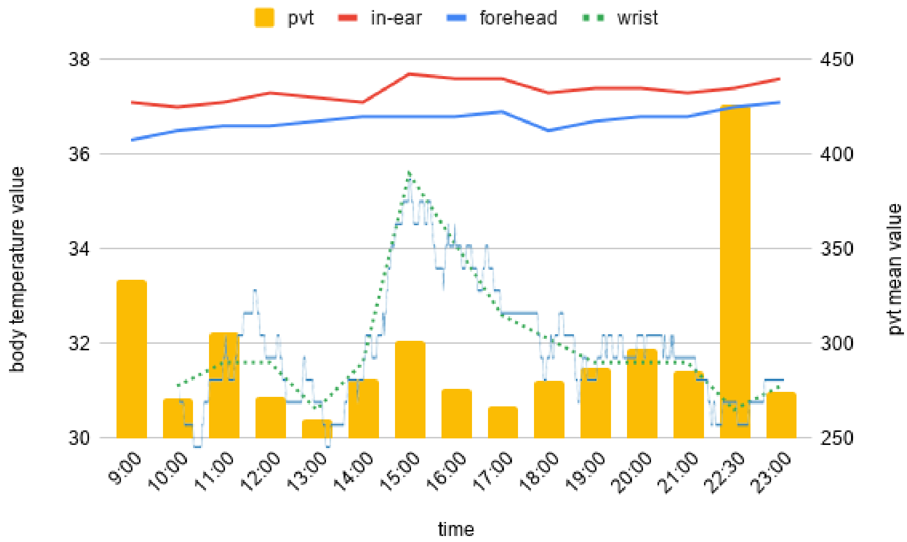

Figure 4.6 Body temperatures and pvt value for m1 on Apr.26th

participants	r	pvalue
f1	0.067	0.636
f2	0.402	0.009
f3	0.430	0.046
f4	0.094	0.662
f5	-0.290	0.135
m1	-0.316	0.057
m2	-0.009	0.971

Table 4.1 Correlation between wrist temperature and PVT mean of 7 participants

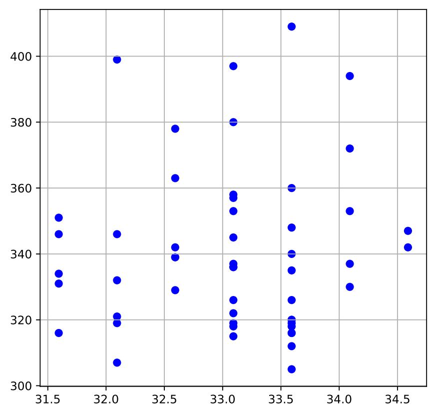

Figure 4.7 Pearson correlation of f1

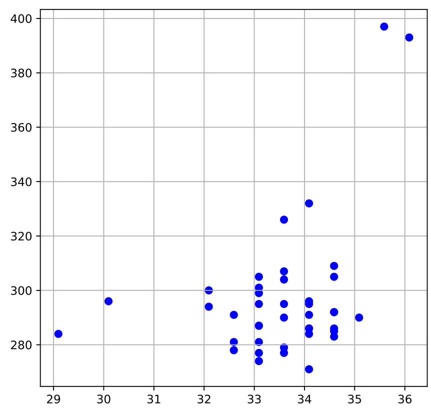

Figure 4.8 Pearson correlation of f2

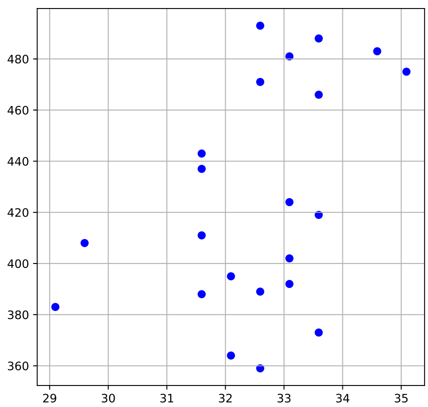

Figure 4.9 Pearson correlation of f3

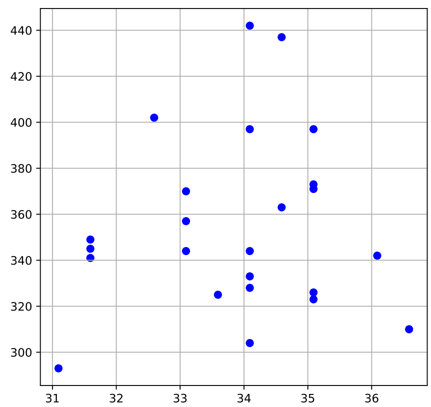

Figure 4.10 Pearson correlation of f4

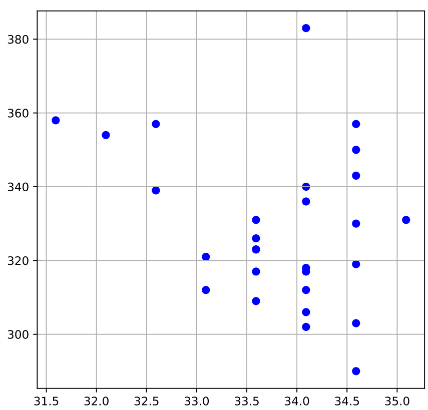

Figure 4.11 Pearson correlation of f5

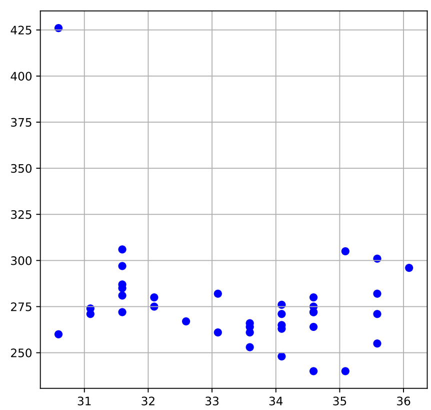

Figure 4.12 Pearson correlation of m1

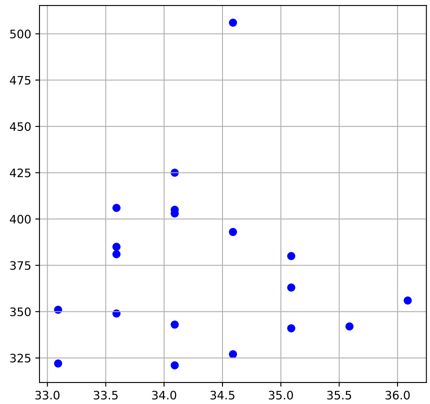

Figure 4.13 Pearson correlation of m2

Individual differences in PVT value and wrist temperatures is shown as Figure 4.2.1),(Figure 4.2.1),(Figure 4.2.1),(Figure 4.13). Then we did an One-Way ANOVA in every participant. First of all, we generate descriptive statistics for each participant. Next, we divided sleepiness level into 3 groups. If PVT mean<25%, the label is low. If PVT mean >75%, the label is high. The label for else is medium. For reducing individual differences of wrist temperature, we also get a relative value by mean value of temperature by participant in the same day. After that, we did an One-way ANOVA for body temperatures and pvt mean in different sleepiness level groups. Result is as below:

- f1
  'wrist': (statistic=1.8392586874878942, pvalue=0.1606166359042055),
  'wrist relative': (statistic=8.396771139848175, pvalue=0.0002793030888100379),
  'forehead': (statistic=8.97976784603084, pvalue=0.00016068030627613102),
  'ear': (statistic=21.843235480147175, pvalue=1.2863605525930523e-09)

- f2
  'wrist': (statistic=2.44241863761655, pvalue=0.08878132802586543),
  'wrist relative': (statistic=0.703359372662608, pvalue=0.4957829560536833),
  'forehead': (statistic=19.550516028406886, pvalue=1.114028140793258e-08),
  'ear': (statistic=36.9659972084366, pvalue=5.456766788733604e-15)

- f3
  'wrist': (statistic=51.67328512981918, pvalue=2.2380222274618613e-19),
  'wrist relative': (statistic=89.41942397166763, pvalue=1.0185813560804417e-29),
  'forehead': (statistic=2.5656876047571058, pvalue=0.07898220549896427),
  'ear': (statistic=7.716189195711445, pvalue=0.000565783180794665)

- f4
  'wrist': (statistic=9.23705532482748, pvalue=0.00013366720849767395),
  'wrist relative':(statistic=10.589010740036061, pvalue=3.8099981104373743e-05),
  'forehead':(statistic=33.349610655915896, pvalue=1.3280172711260402e-13),
  'ear':(statistic=33.05348992451568, pvalue=1.6803574449791867e-13)

- f5
  'wrist': (statistic=8.199801960326008, pvalue=0.0003429887030865057),
  'wrist relative': (statistic=9.94647436334277, pvalue=6.623462556936461e-05),
  'forehead':(statistic=0.3501771485383686, pvalue=0.7048587340434522),
  'ear': (statistic=29.598340857672707, pvalue=1.9748647315388137e-12)

- m1
  'wrist':(statistic=4.980377596600736, pvalue=0.007494883331162837),
  'wrist relative':(statistic=3.4199716446795105, pvalue=0.03409097568490512),
  'forehead':(statistic=1.7291661827104097, pvalue=0.1793283087560481),
  'ear':(statistic=2.192560092153081, pvalue=0.11355045557948831)

- m2
  'wrist':(statistic=5.168896481005451, pvalue=0.006513584868279087),
  'wrist relative':(statistic=5.172879335356285, pvalue=0.006489021168011806),
  'forehead':(statistic=2.753515110738544, pvalue=0.06623366080564128),
  'ear':(statistic=0.4522369412975546, pvalue=0.6368828600006322)

If pvalue<0.05, there were statistically significant differences between 3 sleepiness level group means as determined by on-way ANOVA. In contrast, when pvalue>0.05, that means there were no statistically significant differences. For wrist temperature, there are 5 participants' pvalue<0.05. For wrist temperature relative, there are 6 participants' pvalue<0.05. And the result of forehead is 2 participants, ear side is 3 participants. In summary, wrist temperature relative value and wrist temperature have a better performance than ear and forehead.

After individually analyzing, we combined all the participants data together by 3 sleepiness level groups which are low, medium, high. The dataframe describe included main features is as below(Figures 4.14),(Figures 4.15), (Figures 4.16).

	wrist	wrist relative value	forehead r	ear	ear r	pvt mean	pvtgroup
count	439.00000	439.00000	439.00000	439.00000	439.00000	439.00000	439.0
mean	33.517595	-0.013997	0.052645	37.268793	0.043533	306.26651	0.0
std	1.333375	1.066889	0.252005	0.326188	0.220653	35.752441	0.0
min	29.097000	-4.578000	-0.558000	36.600000	-0.518000	240.00000	0.0
25%	33.092000	-0.482300	-0.058000	37.000000	-0.079000	279.50000	0.0
50%	33.592000	0.046000	0.031000	37.300000	0.027000	305.00000	0.0
75%	34.091000	0.636000	0.109000	37.500000	0.113500	325.50000	0.0
max	36.586000	2.412000	1.042000	37.800000	0.783000	388.00000	0.0

Figure 4.14 Df low.describe()

	wrist	wrist relative value	forehead r	ear r	pvt mean	pvtgroup
count	1030.0000	1030.0000	1030.0000	1030.0000	1030.0000	1030.0
mean	33.332659	-0.049538	0.014921	-0.018647	332.56796	1.0
std	1.322762	1.060807	0.314968	0.238514	47.384546	0.0
min	26.098000	-6.858000	-2.955000	-1.355000	264.00000	1.0
25%	32.593000	-0.500000	-0.107000	-0.143000	291.00000	1.0
50%	33.592000	0.046000	0.000000	-0.017000	326.00000	1.0
75%	34.091000	0.532400	0.157000	0.100000	356.00000	1.0
max	36.586000	2.640000	0.842000	0.542000	471.00000	1.0

Figure 4.15 Df medium.describe()

	wrist	wrist relative value	forehead r	ear	ear r	pvt mean	pvtgroup
count	412.00000	412.00000	412.000000	412.00000	412.00000	412.00000	412.0
mean	33.671362	0.347592	-0.017626	37.219903	0.032126	384.49271	2.0
std	1.515661	1.395763	0.564354	0.549224	0.393959	59.820554	0.0
min	27.097000	-6.637000	-3.158000	35.000000	-1.758000	296.00000	2.0
25%	33.092000	-0.083000	-0.140000	37.100000	-0.017000	350.00000	2.0
50%	34.091000	0.420000	0.073000	37.200000	0.060000	380.00000	2.0
75%	34.590000	1.000000	0.242000	37.500000	0.182000	425.00000	2.0
max	36.586000	3.948000	0.545000	38.300000	0.742000	506.00000	2.0

Figure 4.16 Df high.describe()

We also used the Pearson correlation test to calculated the correlation between in-ear temperature and forehead temperature(Figures 4.17).($r = 0.592$, $p = 0.000$), which means there is a strong and significant correlation between in-ear temperature and forehead temperature. It also showed that forehead temperature can not be considered as the distal skin temperature.

Then we did the same test between in-ear temperature and wrist temperature(Figure 4.18)($r = 0.006$, $p = 0.803$), which shows there is no correlation between in-ear temperature and wrist temperature.

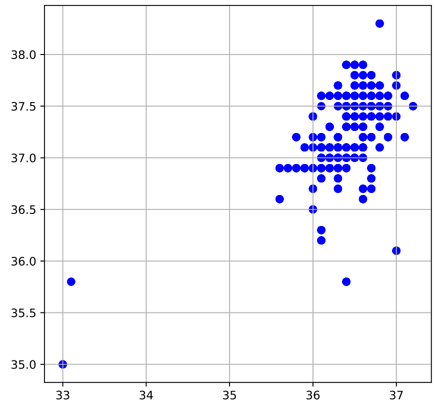

Figure 4.17 Pearson correlation of forehead vs ear

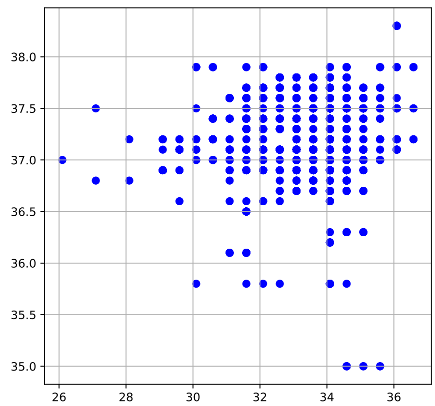

Figure 4.18 Pearson correlation of ear vs wrist

4.2.2 Algorithm

To predict sleepiness level, we used wrist temperature and wrist temperature relative value as two features. We set the row with low sleepiness level as 0, the row with high sleepiness level as 2, otherwise is 1. Then we shuffle split dataframe to training data and test data. Training data included training data and validate data. Our flowchart is a typical corss validation workflow in model training(Figures 4.19) [50]. The other parameters we used are (n splits = 1, training size = 0.8, test size = 0, random state = 0). Then we plot the scatter matrix for the wrist and wrist relative value of the dataframe and the pvt level of the dataframe which is presented as(Figures 4.20). In the next step, we used gridsearch to get the best accuracy: 75.74%, the best parameters:(C: 10000.0, gamma: 100, kernel: rbf). Then after filling into the best parameters, evaluated classification accuracy by confusion matrix. Classification report is as below(Tables 4.2). Confusion matrix is visually present as the following figure(Figures 4.21)

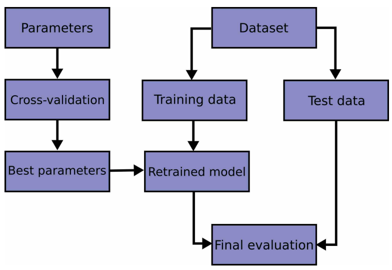

Figure 4.19 Workflow

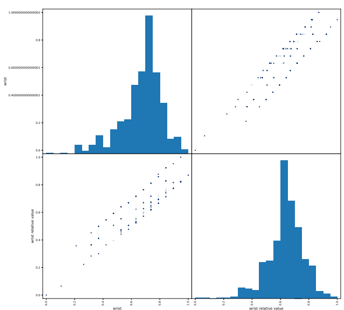

Figure 4.20 Scatter matrix for x train, y train

	precision	recall	f1-score	support
0	0.76	0.91	0.83	88
2	0.88	0.70	0.78	83
accuracy	0.81	171		
macro avg	0.82	0.80	0.80	171
weighted avg	0.82	0.81	0.80	171

Table 4.2 Classification report

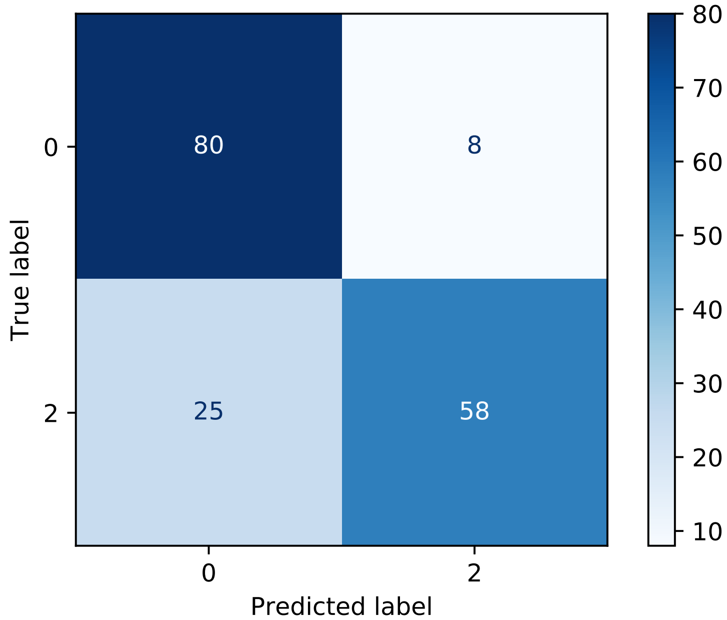

Figure 4.21 Confusion matrix

Chapter 5
Discussions

5.1. Discussion

The collected data of this study suggest that there is a medium strong correlation between wrist temperature and PVT among 3 out of 7 participants; however, we failed to find a linear relationship between wrist temperature and PVT when the combined all the data of the participants together. In this study, the body temperature data of the participants were measured through the wrist temperatures while the basic reaction time was measured through PVT. Simply combining all the data is illogical.

To avoid the individual differences, we calculated the mean wrist temperature for each participant every day, then received the relative body temperature. The Pearson result which is calculated by relative temperature is more positive than raw data. Wrist temperature has the best performance in wrist, ear and forehead. Ear and forehead temperature could potentially affected by different measuring methods. For example, when participants were measured from ear or forehead by a infrared thermometer, measuring angle is different every time, which leads to varied results.

Furthermore, our results also suggest that although there is no linear relationship between wrist temperature and PVT, there is a significant differences with wrist temperature value in 3 different sleepiness levels groups. Therefore, it shows that the wrist temperature has the potential to play a valuable role in better identification and understanding of daytime sleepiness.

5.2. Limitations

After discussion and experiment about the relation between body temperature and daytime sleepiness, we discovered the limitations in our study. There are two possible reasons. One is the number of participants taken this study is insufficient. Therefore we have limited features to divided data to different groups.

The other limitation could be the method we adopted to collect the core temperature more stable continuously and with high accuracy. Measuring temperature by participants themselves turned out to reduce data accuracy. Moreover, we need more objective tools to measure daytime sleepiness. Or in a follow up study we will prepare a training protocol on how to measure in-ear and forehead temperature for the participants. The continuous detection can help us collect more data to improve accuracy of prediction. The third one is it is necessary to find a wireless communication method to get real-time wrist temperature data.

Chapter 6
Conclusions

6.1. Conclusion

The target of this thesis is focused on two main research questions, firstly if there is a relationship between body temperature and objective sleepiness, and secondly which body temperature is the best to use for predicting daytime sleepiness. The experiments were carried out to verify these hypothesis. There is a medium strong linear correlation between wrist temperature and PVT existed in part of people. Preliminary results show that the wrist temperature has the potential to play a valuable part in better understanding daytime sleepiness. And there is a strong correlation between forehead temperature and in-ear temperature, which means forehead temperature can not be considered as distal skin temperature which can helps to reduce core body temperature.

6.2. Future plans

Since the sample size for this study is small and focused in certain group, it can only prove the linear correlation in these groups of people, and it is not representative to all adults in Japan. More data from participants of different age groups, ethnic backgrounds and professions will be needed to extract more features to further justify the hypothesis and improve the accuracy of prediction in the future.

From hardware perspective, in order to gather more accurate and less distorted data, we would like to improve the detection methods, including body temperature measuring tools and daytime sleepiness detection approach which does not affected participants' normal life. Wireless communication devices that can send data remotely would be preferred. Also we will use 3D printer to make a wrist band to fix DS1922L sensor instead of the tennis wrist band.

Another problem requires further study is integration of time series processing in predicting daytime sleepiness. According to the research"Daytime Sleepiness Level Prediction Using Respiratory Information" [33], it is possible to combine classification methods and hidden Markove model [51] to mitigate predict issues such as [52] who combined a neural network and hidden Markove model.

To consider the future prospects of this system. I am expecting it to fulfill the goal mentioned in the beginning of the thesis, that could help people manage their time and energy more effectively, have a healthy life work balance with higher productivity.

References

[1] HR Colten and BM Altevogt. Sleep disorders and sleep deprivation: an unmet public health problem. 2006. *Washington, DC: National Academy of Sciences*, 2007.

[2] David F Dinges, Frances Pack, Katherine Williams, Kelly A Gillen, John W Powell, Geoffrey E Ott, Caitlin Aptowicz, and Allan I Pack. Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4–5 hours per night. *Sleep*, 20(4):267–277, 1997.

[3] Xianchen Liu, Makoto Uchiyama, Keiko Kim, Masako Okawa, Kayo Shibui, Yoshihisa Kudo, Yuriko Doi, Masumi Minowa, and Ryuji Ogihara. Sleep loss and daytime sleepiness in the general adult population of japan. *Psychiatry research*, 93(1):1–11, 2000.

[4] Li-yu Lin, Jie Wang, Xiao-yong Ou-yang, Qing Miao, Rui Chen, Feng-xia Liang, Yang-pu Zhang, Qing Tang, and Ting Wang. The immediate impact of the 2019 novel coronavirus (covid-19) outbreak on subjective sleep status. *Sleep Medicine*, 2020.

[5] Timothy Roehrs. Daytime sleepiness and alertness. *Principles and practice of sleep medicine*, pages 39–50, 2005.

[6] Naomi Breslau, Thomas Roth, Leon Rosenthal, and Patricia Andreski. Daytime sleepiness: an epidemiological study of young adults. *American Journal of Public Health*, 87(10):1649–1653, 1997.

[7] Aurelie Campagne, Thierry Pebayle, and Alain Muzet. Correlation between driving errors and vigilance level: influence of the driver's age. *Physiology & behavior*, 80(4):515–524, 2004.

[8] David F Dinges and Nancy Barone Kribbs. Performing while sleepy: Effects of experimentally-induced sleepiness. 1991.

[9] M Hafner et al. The economic costs of insufficient sleep: a cross-country comparative analysis. *Rand Health*, 2017.

[10] David Hillman, Scott Mitchell, Jared Streatfeild, Chloe Burns, Dorothy Bruck, and Lynne Pezzullo. The economic cost of inadequate sleep. *Sleep*, 41(8):zsy083, 2018.

[11] E Hoddes, Vincent Zarcone, Hugh Smythe, Roger Phillips, and William C Dement. Quantification of sleepiness: a new approach. *Psychophysiology*, 10(4):431–436, 1973.

[12] Murrayb W Johns. Reliability and factor analysis of the epworth sleepiness scale. *Sleep*, 15(4):376–381, 1992.

[13] Mary A Carskadon and William C Dement. Sleep studies on a 90-minute day. *Electroencephalography and clinical neurophysiology*, 39(2):145–155, 1975.

[14] Timothy H Monk. A visual analogue scale technique to measure global vigor and affect. *Psychiatry research*, 27(1):89–99, 1989.

[15] Torbjörn Åkerstedt and Mats Gillberg. Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, 52(1-2):29–37, 1990.

[16] Mark R Pressman and June M Fry. Relationship of autonomic nervous system activity to daytime sleepiness and prior sleep. *Sleep*, 12(3):239–245, 1989.

[17] Y Cook. The effect of nocturnal sleep, sleep disordered breathing and periodic movements of sleep on the objective and subjective assessment of daytime somnolence in healthy aged subjects. *Sleep Res*, 17:95, 1988.

[18] Murray W Johns. A new method for measuring daytime sleepiness: the epworth sleepiness scale. *sleep*, 14(6):540–545, 1991.

[19] Simone Baiardi, Chiara La Morgia, Lucia Sciamanna, Alberto Gerosa, Fabio Cirignotta, and Susanna Mondini. Is the epworth sleepiness scale a useful tool for screening excessive daytime sleepiness in commercial drivers? *Accident Analysis & Prevention*, 110:187–189, 2018.

[20] Merrill S Wise. Objective measures of sleepiness and wakefulness: application to the real world? *Journal of clinical neurophysiology*, 23(1):39–49, 2006.

[21] Timothy H Monk. *Sleep, sleepiness and performance.* John Wiley & Sons, 1991.

[22] Hans PA Van Dongen, Kristen M Vitellaro, and David F Dinges. Individual differences in adult human sleep and wakefulness: Leitmotif for a research agenda. *Sleep*, 28(4):479–498, 2005.

[23] David F Dinges and John W Powell. Microcomputer analyses of performance on a portable, simple visual rt task during sustained operations. *Behavior research methods, instruments, & computers*, 17(6):652–655, 1985.

[24] Jillian Dorrian, Naomi L Rogers, David F Dinges, et al. *Psychomotor vigilance performance: Neurocognitive assay sensitive to sleep loss.* PhD thesis, Marcel Dekker New York, NY, 2005.

[25] CFP George. Sleep· 5: Driving and automobile crashes in patients with obstructive sleep apnoea/hypopnoea syndrome. *Thorax*, 59(9):804–807, 2004.

[26] Charles FP George. Sleep apnea, alertness, and motor vehicle crashes. *American journal of respiratory and critical care medicine*, 176(10):954–956, 2007.

[27] Melinda L Jackson, Rodney J Croft, GA Kennedy, Katherine Owens, and Mark E Howard. Cognitive components of simulated driving performance: Sleep loss effects and predictors. *Accident Analysis & Prevention*, 50:438–444, 2013.

[28] Mr Shailesh Sangle, Bharat Rathore, Rishabh Rathod, Aakashkumar Yadav, and Abhishek Yadav. Real time drowsiness detection system. *IOSR Journal of Computer Engineering (IOSR-JCE)*, 2018.

[29] Ashish Kumar and Rusha Patra. Driver drowsiness monitoring system using visual behaviour and machine learning. In *2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, pages 339–344. IEEE, 2018.

[30] Taeho Hwang, Miyoung Kim, Seunghyeok Hong, and Kwang Suk Park. Driver drowsiness detection using the in-ear eeg. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4646–4649. IEEE, 2016.

[31] Muhammad Ramzan, Hikmat Ullah Khan, Shahid Mahmood Awan, Amina Ismail, Mahwish Ilyas, and Ahsan Mahmood. A survey on state-of-the-art drowsiness detection techniques. *IEEE Access*, 7:61904–61919, 2019.

[32] Sukrit Mehta, Sharad Dadhich, Sahil Gumber, and Arpita Jadhav Bhatt. Real-time driver drowsiness detection system using eye aspect ratio and eye closure ratio. *Available at SSRN 3356401*, 2019.

[33] Kazuhiko Shinoda, Masahiko Yoshii, Hayato Yamaguchi, and Hirotaka Kaji. Daytime sleepiness level prediction using respiratory information. In *IJCAI*, pages 5967–5974, 2019.

[34] Hirotaka Kaji, Hayato Yamaguchi, Kazuhide Shigeto, and Hirokazu Kikuchi. Wearable respiration sensor platform using ultrasound transducer. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, pages 86–89, 2018.

[35] Michael Gradisar and Leon Lack. Relationships between the circadian rhythms of finger temperature, core temperature, sleep latency, and subjective sleepiness. *Journal of biological rhythms*, 19(2):157–163, 2004.

[36] Jing Wei. Monitoring circadian rhythm and sleep patterns using wrist-worn temperature and 3-axis accelerometer sensors: A study with healthy younger adults, healthy older adults, and people living with dementia. Master's thesis, University of Waterloo, 2019.

[37] Danielle J Frey, Pietro Badia, and Kenneth P Wright Jr. Inter-and intra-individual variability in performance near the circadian nadir during sleep deprivation. *Journal of sleep research*, 13(4):305–315, 2004.

[38] PA Van Dongen, Maurice D Baynard, Greg Maislin, and David F Dinges. Systematic interindividual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. *Sleep*, 27(3):423–433, 2004.

[39] Mathias Basner and David F Dinges. Maximizing sensitivity of the psychomotor vigilance test (pvt) to sleep loss. *Sleep*, 34(5):581–591, 2011.

[40] Julian Lim and DavidF Dinges. Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences*, 1129(1):305, 2008.

[41] Julian Lim, Wen-chau Wu, Jiongjiong Wang, John A Detre, David F Dinges, and Hengyi Rao. Imaging brain fatigue from sustained mental workload: an asl perfusion study of the time-on-task effect. *Neuroimage*, 49(4):3426–3435, 2010.

[42] Clare Anderson, Alan WJ Wales, and James A Home. Pvt lapses differ according to eyes open, closed, or looking away. *Sleep*, 33(2):197–204, 2010.

[43] Sylvia Loh, Nicole Lamond, Jill Dorrian, Gregory Roach, and Drew Dawson. The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers*, 36(2):339–346, 2004.

[44] Mathias Basner, Daniel Mollicone, and David F Dinges. Validity and sensitivity of a brief psychomotor vigilance test (pvt-b) to total and partial sleep deprivation. *Acta astronautica*, 69(11-12):949–959, 2011.

[45] Nicole Lamond, Sarah M Jay, Jillian Dorrian, Sally A Ferguson, Gregory D Roach, and Drew Dawson. The sensitivity of a palm-based psychomotor vigilance task to severe sleep loss. *Behavior research methods*, 40(1):347–352, 2008.

[46] Gregory D Roach, Drew Dawson, and Nicole Lamond. Can a shorter psychomotor vigilance task be usedas a reasonable substitute for the ten-minute

psychomotor vigilance task? *Chronobiology international*, 23(6):1379–1387, 2006.

[47] Nicole Lamond, DREW DAWsON, and Gregory D Roach. Fatigue assessment in the field: validation of a hand-held electronic psychomotor vigilance task. *Aviation, space, and environmental medicine*, 76(5):486–489, 2005.

[48] David R Thorne, Dagny E Johnson, Daniel P Redmond, Helen C Sing, Gregory Belenky, and Jordan M Shapiro. The walter reed palm-held psychomotor vigilance test. *Behavior Research Methods*, 37(1):111–118, 2005.

[49] AD Harper Smith, DR Crabtree, JLJ Bilzon, and NP Walsh. The validity of wireless ibuttons® and thermistors for human skin temperature measurement. *Physiological measurement*, 31(1):95, 2009.

[50] scikit-learn developers (BSD License). Cross-validation: evaluating estimator performance. [EB/OL]. `https://scikit-learn.org/stable/modules/cross_validation.html#cross-validation/` Accessed April 4, 2010.

[51] Lawrence R Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.

[52] Kun Han and DeLiang Wang. Neural networks for supervised pitch tracking in noise. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1488–1492. IEEE, 2014.

Appendices

A. Example of Fault Detection Configuration File

Configuration File for Detecting Failures

```
<config>
<system>
<class>DefaultCompareClass</class>
</system>
<evaluate>
<compare_single_observation_point>
<function method="compareMax" recital="Temperature maximum threshold value"
type="Temperature"> <argument class="double">40.8</argument>
</function>
<function method="compareMin" recital="Temperature minimum threshold value"
type="Temperature"> <argument class="double">-41.0</argument>
</function>

<function method="compareChange" recital="Temperature change amount error"
type="Temperature"> <argument class="double">17.0</argument>
<argument class="int">1</argument>
</function>

<function method="compareConstant" recital="Temperature constant error"
type="Temperature"> <argument class="int">1</argument>
</function>
</compare_single_observation_point>

<compare_neighbor>
<function method="compareNeighbor" recital="Temperature neibor error"
```

```
type="Temperature"> <argument class="double">2.0</argument>
</function>
</compare_neighbor>

<compare_wide_area>
<function method="compareWide" recital="RainFall wide area error"
type="RainFall"> <argument class="double">10.0</argument>
</function>
</compare_wide_area>
</evaluate>
</config>
```