How to Detect and Classify Stress Using Wearable Sensors to Recommend Task

Naoya Wakabayashi[1], Hiromitsu Shimakawa[2], and Fumiko Harada[3]
[1] Graduate School of Information Science and Engineering, Ritsumeikan University, Japan
[2] College of Information Science and Engineering, Ritsumeikan University, Japan
[3] Connect Dot Ltd., Kyoto, Japan
Email: {wakaba.naoya, simakawa, harada}@de.is.ritsumei.ac.jp

Abstract—**To improve work productivity, this study proposes a method to detect stress, which is classified as positive or negative after the detection. The classification contributes to recommending suitable tasks under every stress condition. The method uses less invasive wearable sensors to acquire biological data. It is checked whether the characteristics of the wearable sensors would affect the biological data. Important variables are also explored to estimate stress. This paper presents the results and discussions to confirm through experiments whether the proposed method can detect and classify stress.**

Index Terms—**Classification of stress, detection of stress, positive stress, negative stress, task recommendation, wearable sensor**

I. Introduction

In Japan, an aging population has currently caused a serious shortage of workers in the future. Against this background, "Work Style Reform" promulgated in 2018 is a policy to efficiently run the Japanese economy even with a small working population. One of the key policies in the work style reform is to improve an individual's productivity who engaged in work. To improve work productivity, it is necessary to carry out tasks that are appropriate to the individual's stress state.

It is impossible to prevent stress from occurring in the workplace. However, by optimizing the assignment of a task according to a stress state, we can improve work efficiency. Eventually, we can achieve the improvement of work productivity.

Using less invasive wearable sensors, this paper proposes a method not only to detect whether a person working on a task is under stress but also to classify the type of the stress, if there. The proposed method brings, a mechanism to recommend what task the person should engage in next under the detected stress.

The paper addresses the following questions to put the proposed method into practice.

- Does the attachment of the wearable sensors themselves affect biological data?
- What kind of variables are important to estimate the stress of many persons?

When persons wear sensors, they often present physiological reactions such as sweating at the part where the wearable sensors are attached. In other words, we cannot deny the possibility that the characteristics of the wearable sensor affect the biological data. The effects of the sensor attachment on the biological data should be examined to accurately estimate stress. Furthermore, it is necessary to find out variables playing an important role to estimate stress. The variables should be common to many humans in the estimation of their stress. It is useless to measure unimportant variables. The classification of variables in terms of their importance leads to reducing the number of sensors that need to be attached. It would make the proposed method less invasive.

II. Existing Works

Stress is inseparable from engagement in work. People cannot avoid various stress while they engaged in their work. Though it is said that stress reduces work performance [1], it is a prejudiced view of stress.

The law of Yerkes-Dodson [2] indicates moderate stress, which is neither too high nor too low, evokes optimal performance. Moderate stress increases work productivity, while excessive stress decreases it. It means there are different types of stress, each of which will affects work productivity. We can maximize the work productivity of people, if we engage them in important works under stress to evoke optimal performance.

Some works [3]-[5] find out the presence or the absence of stress, succeeding the prediction of its magnitude. However, they take prejudiced views for stress. To assess stress in an unprejudiced way, we should examine the relation of its magnitude with work productivity. Nevertheless, to the authors' best knowledge, no study takes the relation into account.

We should also note that existing studies addressing stress acquire biological data through sensors such as EEG sensors, which are difficult to wear. Sensors difficult to wear are undesirable for estimating the stress state of people while they are engaging in their works. We should use less invasive sensors as well as have target persons wear as few sensors as possible.

Manuscript received August 30, 2021; revised November 28, 2021.

doi: 10.18178/ijeee.9.4.83-92

III. Task Recommendation According to Stress

A. Positive Stress and Negative Stress

Stress is an external stimulus to a person. It is diverse. This type of stress can affect work productivity. In this study, stress will be classified into positive and negative stress with reference to the law of Yerkes-Dodson.

Positive stress refers to a state in which a moderate load is placed on the body and mind to promote a subject's physical and mental functioning. On the other hand, negative stress refers to a state in which an excessive load is placed on the body and mind. It is likely to lead to a decline in physical and mental functions. Specifically, it is a state in which a subject has frustrated or dissatisfied.

Whether stress is benign or malignant may be determined by the amount of capacity the subjects can afford to accomplish the tasks they are engaged in. Even if the amount of stress is high, it is likely to be benign stress if the subject has the capacity to accomplish their tasks. Even less stress may result in malignant stress when the subject cannot afford the tasks. The classification of stress depends on the condition of the subject.

The categorization of stress reflecting states of target persons allows us to know whether the current task should be maintained to engage in or a new task is to be introduced.

B. Types of Task

The study assumes that work can be categorized into three types: Creation, To-Do, and Meeting. It also assumes that all jobs will always belong to one of the categories.

The Creation refers to creative work where the quality of a task is required. Thinking is dominant in it. It requires the issue of different ideas repeatedly. The To-Do refers to relatively simple tasks that can be done mechanically by moving your hands. In carrying out the work, we do not need to think continuously, but we must process it quickly while preventing mistakes. The Meeting refers to work that requires a partner, such as discussion. It contains an element of refreshing the mood. However, since it requires a partner, it should not be encouraged when a person is in a state that makes the other uncomfortable.

C. Task Recommendation with Machine Learning

There are several types of tasks. A task to be carried out differs depending on the type of stress. Since there is a task suitable for the type of stress a person has, we should assign an optimal task to the person. To estimate the stress state of a target person, the proposed method not only detects stress but also examines whether it is a positive or negative one.

An overview of this method is shown in Fig. 1. More than one wearable sensor is attached to the target person engaging in a task. Biological data from the wearable sensors are attained as an indicator to calculate stress. The values calculated from the biological data are analyzed to detect whether the target person is stressed or not. Only if it turns out the person has stress, the stress should be a positive or negative one. Based on the information, the paper proposes a method to recommend optimal tasks.

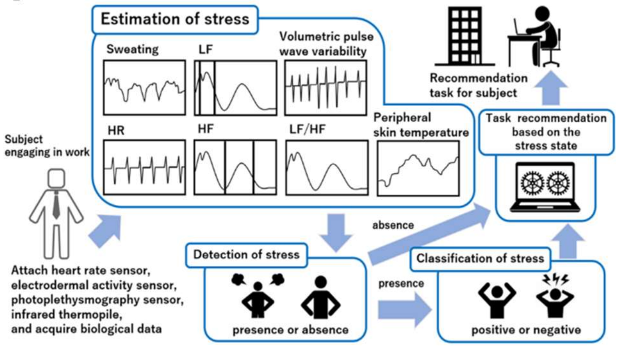

Figure 1. Schematic diagram of the method.

D. Physiological Signals

To estimate the stress state of a target person, the method analyzes data acquired with wearable sensors to detect stress as well as to examine whether it is a positive or negative one. The target person wears an electrodermal activity sensor that can measure sweating, a photoplethysmogram sensor that can measure volumetric pulse wave variability, and an infrared thermopile that can measure peripheral skin temperature. In addition, the person wears a heart rate sensor to the chest.

The electrodermal activity sensor constantly measures the fluctuation of electrical properties of the subject's skin. They involve not only thermal sweating but also psychogenic sweating. They can be measured electrically. Skin conductance can be used as an indicator of arousal of the autonomic nervous system. It is said that emotional stimuli received by humans have a correlation with sweating [6]. Stress can be predicted with electrodermal activity sensors [7], [8].

Furthermore, human skin temperature varies with stress [7], [9]. Skin temperature measured with the infrared thermopile can be one of the indicators to estimate stress.

The human autonomic nervous system consists of sympathetic and parasympathetic nerves. The sympathetic nerves get active in the state of fight, while the parasympathetic nerves work the state of relaxation. In other words, the sympathetic nervous system is more active when we are under stress. When the sympathetic nervous system gets active, the blood pressure is risen, peripheral blood vessels are constricted, and heart rate is accelerated [10]. The photoplethysmogram sensor can predict stress, measuring changes in the volume of blood flow. It analyzes the Blood Volume Pulse components of the blood flow to measure changes in heart rate variability and peripheral vascular blood flow.

The heart rate during hyperactivity of the sympathetic is significantly influenced by the type of stress [11]. The study uses a heart rate sensor to obtain the heart rate interval of the target person. The activities of the sympathetic nerves and the parasympathetic nerves fluctuate due to the interval between heart rates [12]. The largest wave in the heart rate is called the *R wave*. Stress causes fluctuation in the interval between the R wave and the R wave, which is generally called the RR interval. Stress also changes in the ratio of the low frequency and the high frequency. The former and the latter are hereinafter referred to as HF and LF, respectively [13], [14]. The frequency components of the RR interval can be used for the estimation of stress, as the heart rate interval can.

E. Instantaneous Values v.s. Trend Values

When biological information is collected using a wearable sensor, the attachment of the sensor itself may affect the measurement. Especially in the case of a sensor that measures stress based on sweating, it is necessary to investigate the effect of stuffiness caused by sweating from the skin because of the attachment of the sensor. The study confirms whether the stuffiness of wearable sensors affects stress discrimination.

The effects of wearable sensors for biological data are examined in two cases. Namely, the study compares instantaneous values with trend values for skin properties.

- In the case that there is almost no effect of steaming by the wearable sensors, the instantaneous values are used for skin properties. For other signals, the instantaneous values and frequency analysis values are used as explanatory variables.
- In the case the effect of steaming by the wearable sensors is not negligible, we should adopt trend values for the skin properties. Even in this case, we should adopt the instantaneous values and frequency analysis values for the other signals.

In this study, the trend refers to the difference of values in the current time from ones at a time point a certain period ago. The study builds a model in each case to compare the performance and important variables.

IV. EXPERIMENT

A. Procedure in Experiment

The study creates a mechanism that estimates the stress state of the target person to recommend a suitable task. It has conducted an experiment, where biological data necessary to estimate the stress state is obtained under various stress tasks.

The subjects were 22 people, 13 undergraduate students and 9 graduate students of Ritsumeikan University in their 20s. Every subject wore an Empatica E4 wristband sensor on the wrist opposite to their dominant hand and a Polar H10 sensor on the chest. Five of these subjects had missing sensor data due to connection failures to receivers. The analysis was conducted on the data from a total of 17 students whose complete data set is available. They involve 10 undergraduate students and 7 graduate students.

The experiment assumes that every subject stayed in one of the three states: unstress, positive stress, and negative stress. All subjects used a laptop computer of the same specification. The experiment was conducted as follows.

1) Each subject took behavior to enter a resting state in their arbitrary way. The arbitrary way means that it is up to the subject to decide how to get into a resting state.
2) The subject was notified in advance of the competition with other subjects. After that, the subject performed a typing test within the time limit of 90 seconds, while disclosing the remaining time to the subject.
3) Each subject entered their arbitrary resting state.
4) The subject typed 35 English words common to all subjects with an editor. On completion of the typing task, the subject pressed a submit button.
5) For the English words entered by the subject, an error message is presented to suggest a typo in one of the words. However, the English words that cause the error are not disclosed. The subject searches for the correction while looking at the English words entered in 4. The subject presses the send button after correction.
6) A second error message is presented on a screen. The input buffer is reset to blank for typing English words. The subject types the 35 English words again. On the completion, the subject presses a submit button.
7) The third error message is presented on a screen. The input buffer is reset to blank. The subject types the 35 English words again to press a submit button after the completion.
8) A message telling the input completion is displayed on the screen on the laptop.
9) Each subject takes behavior to enter a resting state in their way.
10) The subject performs a 1-digit + 1-digit calculation problem without any announcement of the time limit and the elapsed time. The actual time of calculation is 90 seconds. The subject provides answer to the displayed question in an answer field on the same screen. When the subject presses a submit button, the following question is displayed.
11) Each subject enters a resting state in their way.

The subjects recorded a stress state they feel during the experiment. A free-drawing questionnaire is used to evaluate what kind of stress is imposed by the experiment task. The free-drawing questionnaire sheet used in this experiment is shown in Fig. 2. The free-drawing questionnaire was created based on the existing study [15]. The vertical axis indicates stress states while the horizontal axis corresponds to the experiment numbers the subjects engage in the course of the time. The numbers at the bottom correspond to the experiment IDs.

At the end of the experiment, the subject fills out the evaluation form shown in Fig. 2, which indicates the stress states while performing each task. The stress states are represented with a single continuous line. It is used as a correct label for the data set. Since the length of the time required to complete all the tasks in the experiment varies with each subject, the start time and the end time of each task are shown to the subject. The stress state and the time for each task are annotated based on what the subject filled out in the sheet.

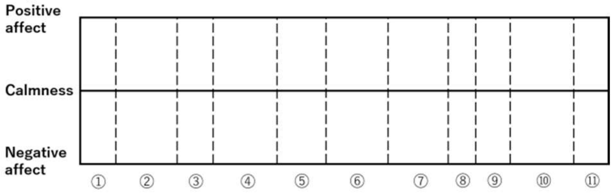

Figure 2. Free-drawing questionnaire.

B. Data Preprocessing

The E4 wristband sensor incorporates an electrodermal activity sensor, a photoplethysmogram sensor, and an infrared thermopile [16]. It is possible to acquire the psychogenic sweating rate, the volumetric pulse wave variability, and the peripheral skin temperature from subjects. The heart rate interval of a subject can be obtained from the H10 sensor attaches to the chest [17].

When time series of data acquired with the sensors is fed to a machine learning model as explanatory variables, it is necessary to make every sample arrange in a specific interval. However, each sensor has its sampling rate to acquire data. The study resamples all data in 1-second intervals as the following.

The electrodermal activity sensor acquires sweating data at 4 Hz with a pair of electrodes attached to the band of the E4 wristband sensor. The photoplethysmogram sensor samples the heart rate at 64 Hz based on the degree of reflection of the green LED output from the E4 wristband sensor. The peripheral skin temperature is obtained at 4 Hz from the absorption of the infrared thermopile. These time series are smoothed using the cubic spline interpolation. Succeedingly, outlier values are treated based on the Smirnov-Grubbs test. With the significance level of 5%, values that turn out outliers are substituted with the most frequent value. Eventually, the study averages the values that appear within every second. It achieves down-sampling, regarding the average as the representative value in the second.

The LF and the HF are attained through frequency analysis of the heart rate interval interpolated with a cubic spline function. The frequency analysis method uses the FFT (Fast-Fourier-Transformation) to calculate the frequency analysis values at 1-second intervals for all data. The Hanning window is adopted as a parameter in the FFT. The window size is 10 sec, while the overlap is 50%. The band of 0.00 to 0.04 Hz in the frequency is regarded as the LF component, while 0.15 to 0.4 Hz is the HF component [18], [19]. Needless to say, the LF/HF, the ratio of the LF component to the HF one, is calculated through the division of the magnitude of the LF component by the one of the HF components.

Each subject has its biological characteristics. Though the amplitude of sampled values under a specific kind of stress is different from person to person, all subjects should be compared in the same value range. In this study, the sampled values are standardized for each subject when they are fed to a machine learning model as its explanatory variables.

V. Estimation of Stress State

A. Stress Estimation Model

The proposed method detects stress. It also examines whether it is positive or negative. This chapter shows the results in each case. The proposed method uses either instantaneous values or trend values for skin properties. In addition, the chapter compares the degree of important variables for each variable of every subject.

To estimate stress, the method uses a Random Forest model, which is supervised learning. Data acquired from each wearable sensor are resampled at 1-second intervals to be used as explanatory variables. The data sets of instantaneous values are used as the resampled values. The data sets of trend values are calculated using the difference of current values from values 90 seconds before.

The ratio of training data and testing data for a Random-Forest model used in the estimation is determined referring to the default value of the train-test-split function implemented in Python's Scikit-learn. The ratio of training data to testing data is 3:1 for the estimation. Taking the possibility of overfitting for each subject's model into account, the mean value was calculated through cross-validation. The number of divisions for the cross-validation is 5. The AUC is also calculated as an assessment index. The AUC is an area under the curve drawn by the Recover Operating Characteristic Curve, which is an evaluation index for binary classification. It has an advantage that enables appropriate evaluation even when the number of stressed samples is biased compared with that of non-stressed ones [20].

B. Detection of Stress

This section shows the results of detecting stress along with important variables for the detection.

The results for skin properties are compared in the two cases: one using instantaneous values and the other using trend values. Few subjects reported staying in unstressed states in the experiment. In addition, there is a large bias for the number of samples between stressed states and unstressed states. Here, the analysis focuses on samples acquired from only the 11 subjects who reported the absence of stress in more than 5% of the experiment period.

Table I shows the accuracy, the mean in the cross-validation, and the AUC of each model for every subject.

In both models, the accuracy of all subjects is high enough. It does not differ significantly from the mean in the cross-validation. The AUC of all subjects is over 0.83. The value is statistically significant. However, there is a difference in the AUC for some subjects. In particular, in cases of subjects 1, 2, 4, and 10, the AUC calculated using the instantaneous values is higher than that using trend values by more than 0.07. It implies we can construct a detection model of stress for each person with no relation to effects from the type of skin properties. We can also conclude instantaneous values for skin properties have a slight advantage in performance.

TABLE I. TYPE SIZES FOR CAMERA-READY PAPERS

	using instantaneous values for skin properties			using trend values of skin properties		
Subject	Accuracy	Mean in the cross-validation	AUC	Accuracy	Mean in the cross-validation	AUC
1	1.00	0.99	1.00	0.97	0.98	0.84
2	0.99	0.99	0.99	0.92	0.94	0.89
3	0.98	0.97	0.87	0.97	0.97	0.86
4	1.00	0.97	1.00	0.97	0.97	0.9
5	0.93	0.94	0.87	0.92	0.92	0.83
6	0.96	0.98	0.95	0.94	0.94	0.91
7	0.98	0.98	0.97	0.95	0.96	0.94
9	0.97	0.99	0.94	0.98	0.98	0.95
10	0.97	0.97	0.91	0.93	0.93	0.84
14	0.98	0.97	0.95	0.97	0.95	0.91
16	0.95	0.97	0.85	0.95	0.97	0.85

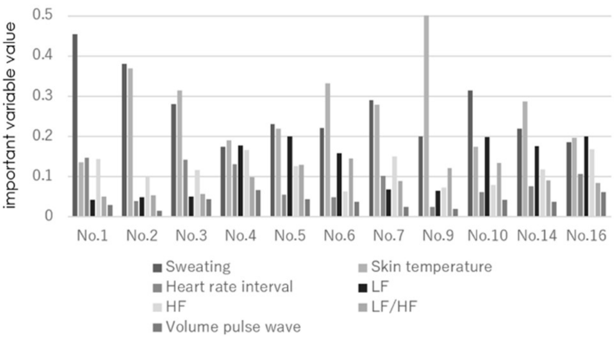

Figure 3. : Important variables for the detection model using instantaneous values for skin properties.

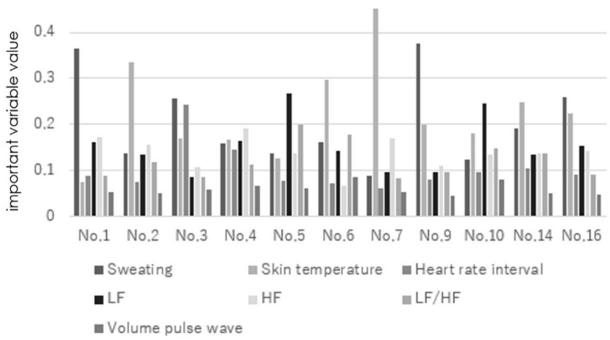

Figure 4. Important variables for the detection model using trend values for skin properties.

Fig. 3 shows the important variables for the detection model using instantaneous values for skin properties. Subject 1 has prominently high in sweating while subject 9 has prominently high in skin temperature. For the other subjects, both the sweating and the skin temperature are prominent, or all variables are much the same. For the cases where all variables are much the same, the values of each important variable get almost equally small in proportion to the rank of the important variable. For subjects 2 and 7, the sweating is the most important, followed by the skin temperature, and the other variables are similar in that HF has a slightly higher value. For subjects 4 and 16, the importance is much the same in all variables and the rankings of each variable are similar. Although there are differences in the ranking of the important variables for subjects 5 and 10, they are similar in that the most important variable is the sweating and the characteristics of the variables in the heart rate system are similar.

The sweating and the skin temperature are located within the top three important variables for all subjects. The volume pulse wave variability is the last important variable for all subjects. Likely, important variables to construct a stress detection model using instantaneous values for skin properties are sweating and skin temperature. On the other hand, volumetric pulse wave variability is least likely to be an important variable.

Fig. 4 presents the degree of important variables for the detection model using trend values for skin properties. Compared with Fig. 3, many subjects have higher importance for variables related to the heart rate system. In addition, for the subjects in which both the sweating and the skin temperature are prominent as important variables in the case of instantaneous values, only one of the variables is prominent. From these, in all subjects, one variable is important, or all variables are almost equal in their importance.

For subjects 1 and 3, the sweating is located in the first place in the importance of variables. For subjects 2, 6, and 7, the skin temperature is located in the first important variables. For subjects 4 and 10, the skin temperature is ranked within the top two important variables. For subjects 9, 14, and 16, sweating and skin temperature are the top two important variables. Only subject 5, the variables of the heart rate system occupy the top two important variables. Therefore, at least one of the sweating and the skin temperature is likely to be an important variable for every subject. On the other hand, values of the volumetric pulse wave variability are less than 0.10 in all subjects. It is placed in the last important variable with a high percentage. In conclusion, it is likely the variables that are important to construct a detection model of stress using trend values for skin properties are either or both sweating and skin temperature. In addition, volumetric pulse wave variability is least likely to be important.

C. Classification of Stress

This section presents the results of classifying stress which is positive or negative. The results are compared in the two cases regarding the skin properties: one using their instantaneous values and the other using their trend values. For all subjects whose complete data are available, their stress is classified into positive ones or negative ones.

Table II presents the accuracy, the mean in cross-validation, and the AUC of each model for every subject at the classification.

The model using instantaneous values has an accuracy that is higher than 0.95 for all subjects. It is not significantly different from the cross-validation accuracy. All values of AUC are also more than 0.9. Regarding the model using trend values, almost every subject has a lower accuracy than in the case of the classification model using instantaneous values. However, the values of accuracy are greater than 0.86. In addition, they are not significantly different from the average accuracy in the cross-validation. The AUC of the model for the trend values is 0.77 at the lowest, which is high for the value. On the contrary, the AUC of the trend value is much lower than the AUC of the instantaneous value. Therefore, it is better to use instantaneous values for skin properties to struct a classification model for each.

TABLE II. TYPE SIZES FOR CAMERA-READY PAPERS

	using instantaneous values for skin properties			using trend values of skin properties		
Subject	Accuracy	Mean in the cross-validation	AUC	Accuracy	Mean in the cross-validation	AUC
1	0.97	0.98	0.98	0.89	0.90	0.89
2	0.97	0.95	0.96	0.92	0.93	0.88
3	0.95	0.95	0.94	0.96	0.97	0.95
4	0.97	0.96	0.97	0.93	0.95	0.93
5	0.95	0.96	0.94	0.92	0.94	0.9
6	0.99	0.99	0.99	0.97	0.98	0.97
7	1.00	1.00	1.00	0.99	0.99	0.98
8	0.96	0.97	0.91	0.97	0.98	0.93
9	0.99	0.99	0.99	0.97	0.98	0.97
10	1.00	0.99	1.00	0.86	0.88	0.83
11	0.97	0.95	0.94	0.9	0.91	0.77
12	0.96	0.97	0.90	0.93	0.95	0.82
13	0.98	0.98	0.98	0.91	0.93	0.88
14	0.98	0.97	0.98	0.96	0.96	0.95
15	0.98	0.97	0.98	0.99	0.98	0.99
16	0.96	0.95	0.95	0.92	0.93	0.89
17	0.99	0.98	1.00	0.93	0.91	0.94

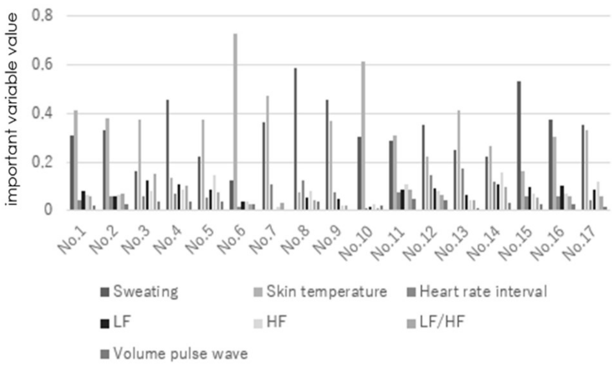

Figure 5. Important variables for the classification model using instantaneous values for skin properties.

The important variables for the classification model using instantaneous values are shown in Fig. 5. Except for subject 8, the top two important variables for each subject are the sweating and the skin temperature. Furthermore, for all subjects, these two variables account for more than 50% of the weight in classifying the stress. Not only the values of the volumetric pulse wave variability of all subjects are less than 0.05, but also it is ranked in the last with high percentages. We can conclude sweating and skin temperature are important variables to construct a classification model of stress using instantaneous values. We can also say volumetric pulse wave variability has the least possibility to be an important variable. Furthermore, by looking at the sweating and the skin temperature of each subject, it is possible to divide them into several groups.

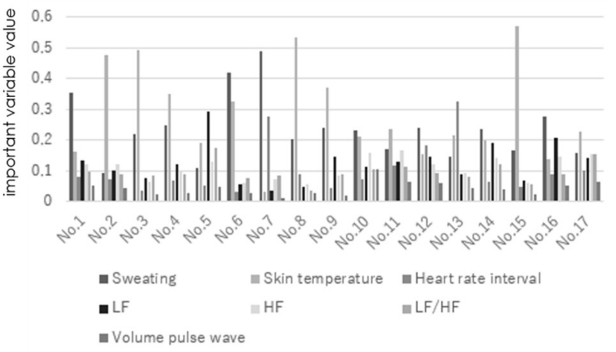

Figure 6. Important variables for the classification model using trend values for skin properties.

Fig. 6 shows the degree of important variables for the classification model using trend values. Compared to Fig. 5, the sweating and the skin temperature are less important in many subjects, and the importance of variables in the heart rate system is higher. This results in equal distribution of the importance of variables for many subjects. The values of the volumetric pulse wave variability of all subjects are less than 0.10. There is a high percentage for it to be ranked in the least important variable. In conclusion, sweating and skin temperature have high possibilities to be important variables for a classification model of stress using trend values. The variables of the heart rate system are also important indicators for classification. The conclusion includes volumetric pulse wave variability has the low possibility to be an important variable.

VI. Consideration

A. Calculation of Trend Values

This study calculated trend values for skin properties, assuming the use of wearable sensors may cause steaming. Stress was estimated for two datasets: one is using instantaneous values for skin properties and the other is using trend values for skin properties. This section discusses whether the calculation method is appropriate.

To examine whether the delay of 90 seconds to get trend values is appropriate, we should compare the AUC values when the trend values are calculated at other times. The constant times used for comparison in this section are 10, 15, 30, 45, 60, 90, and 100 seconds. For each of them, trend values are calculated to detect stress and to classify it as positive stress or negative stress.

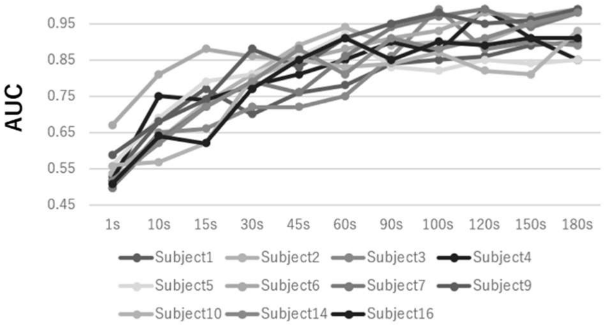

Figure 7. Transition of AUC for the detection model.

Fig. 7 shows the transition of AUC when stress was detected for each subject. Regarding the stress detection model, there is an upward transition to the right as a whole. In addition, there is a relatively parallel transition from 90 seconds to 180 seconds. In particular, the 90 second has the smallest AUC amplitude for every subject.

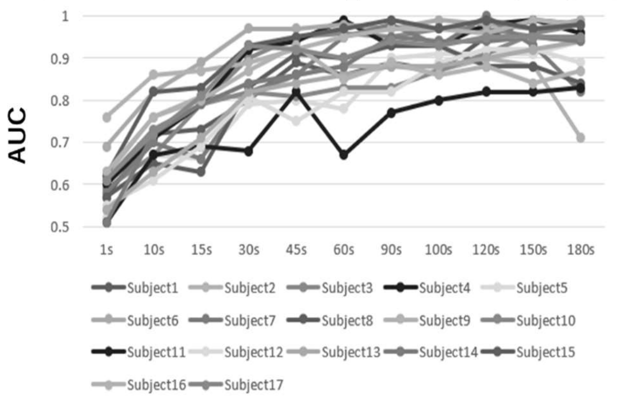

Figure 8. Transition of AUC for the classification model.

Fig. 8 shows the transition of AUC when stress was examined whether it is positive or negative for each subject. Regarding the stress classification model, there is also an overall upward transition to the right. Compared to the detection model in Fig. 7, the transition is relatively quick to become parallel. The AUCs for all subjects except subject 11 are relatively parallel from 30 seconds to 180 seconds. In the detection model, the AUC increases for some subjects while it decreases for others after 90 seconds. However, in the classification model, the AUC of most subjects increased after 30 seconds and decreased after 120 seconds.

The appropriate time varies with each subject. However, the criteria for calculating the trend values is preferable to be the same for all subjects. Therefore, the average value of the transition at each time was calculated to extract the overall trend of the transition.

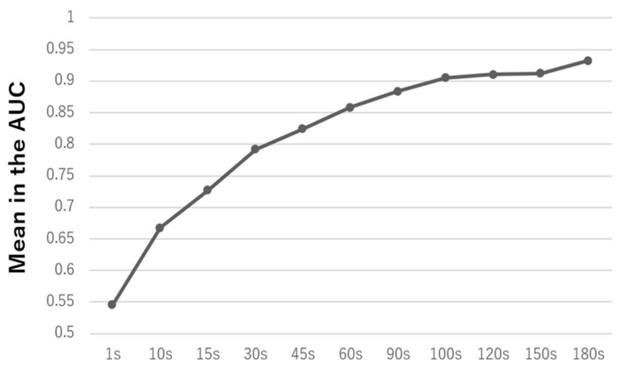

Figure 9. Transition of mean in the AUC for the detection model.

Fig. 9 shows the average transition of AUCs in the detection model. The transition of the mean values of the detection models is relatively horizontal between 100 seconds and 150 seconds, among an overall upward rightward transition. At 180 seconds, some subjects show a significant drop in AUC. Therefore, 100 seconds is desirable for the calculation of the trend value to detecting the presence or absence of stress.

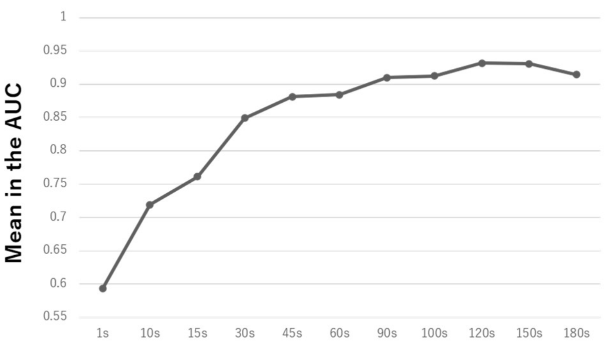

Figure 10. Transition of mean in the AUC for the classification model.

Fig. 10 shows the average transition of AUCs in the classification model. The transition of the mean value of the classification model is upward to the right, but there are several times when the transition is relatively horizontal. Looking at the trends for each subject at the maximum value of 120 seconds, the AUCs of several subjects are decreasing. Also, after 150 seconds, the AUC of some subjects decreases while the AUC of others increases. At 90 seconds, many subjects' AUCs are higher or equal to the previous AUC. Therefore, 90 seconds is desirable for the calculation of the trend values to classify stress.

The maximum value in the mean of the detection model is 0.93. The value of the detection model at 90 seconds is 0.88. The degrade is less than 0.05, which is not a problem. Therefore, the 90 seconds used in this paper is appropriate to calculate the trend values.

B. Instantaneous Values and Trend Values

In this paper, we constructed two models, one is using instantaneous values for skin properties and the other is using trend values for skin properties, because the effect of steam from the sensor may not be negligible. Based on the performance results of the models compared in the chapters, the section discusses which one is more suitable for stress estimation.

As for the detection model, the results of 5.2 show that the accuracy of all subjects is high enough in both models. It does not differ significantly from the mean in the cross-validation. However, there are more subjects with higher AUCs when using instantaneous values for skin properties. They have higher AUC and slightly better performance.

As for the classification model, the results of 5.3 show that the model using instantaneous values has accuracy for all subjects is higher than 0.95. It is not significantly different from the cross-validation accuracy. Most of the subjects using trend values have lower accuracy than that of the classification model using instantaneous values. The AUC of the trend values is much lower than that of the instantaneous value. Therefore, it is better to use instantaneous values for skin properties to construct a classification model for each.

Therefore, to estimate human stress using wearable sensors, it is better to use instantaneous values for skin properties.

C. Important Variables for Stress Estimation

To estimate the stress of a subject engaged in work, it is desirable to measure it with a sensor that is as minimally invasive as possible. For this purpose, it is important to search for variables that are commonly important for each subject. Based on the experimental results, the section discusses the variables that are important for the estimation of stress.

Concerning the detection of stress, when the instantaneous values for skin properties are used, the subjects are divided into two groups: those for whom the importance of the sweating and the skin temperature is very high, and those for whom the sweating and the skin temperature are high and other variables also contribute to the model.

The volumetric pulse wave variability is the least important of all subjects. When the trend values are used for skin properties, either the sweating or the skin temperature is prominent, or the values of all variables are much the same in importance. The volumetric pulse wave variability is the least important for almost all subjects.

In the model using instantaneous values for skin properties, the sweating and the skin temperature are important for the classification, while other variables are relatively unimportant. In particular, the important variable of the volumetric pulse wave variation is very low. In the model using trend values for skin properties, there are two patterns of subjects: one in which the sweating and the skin temperature are prominently important, and one in which all variables are important.

Therefore, the variables of particular importance in detecting the presence of stress and classifying it are sweating and skin temperature. Variables related to the heart rate are not important for the classification of stress using instantaneous values. Furthermore, volumetric pulse wave variability is not important in estimating stress in humans.

D. Recommendation of Tasks

This study proposes a method to understand the stress state of the subject. When it detects stress, it classifies the stress into a positive one or negative one, to recommend tasks. Using qualitative considerations, the paper discusses tasks suitable for the stress state.

The stress state of an individual changes from time to time. However, it is not desirable to change the recommended task continuously. This paper proposes a mechanism that aggregates the stress states predicted every second for 10 minutes, to determine the recommended task for each period.

The subject stays in his best working state when he is predicted to be in a positive stress state for most of the period. The recommend task is creative work that is thought-centered and requires high quality in the work. On the other hand, if the subject is predicted to be in a negative stress state for most of the period, he needs a break because he is not in a state to engage in work. When the majority of the period is covered with absent stress, the method recommends To-do tasks that can be completed with the mechanical movement of the hands. The subject can prevent mistakes because less time is spent under stress.

When there is a mixture of positive stress and absent stress conditions in the period, a meeting requiring partners is recommended. In the meeting, it is preferable to avoid negative stress that may cause discomfort to the partner. Therefore, when the subject has no negative stress and relatively positive stress, meetings can be carried out smoothly. If there is a mixture of negative stress and absent stress conditions, the method recommends a task to investigate and collect information about the job. Although only negative stress might interfere with working on the job, a mixture of stress-free periods would contribute to the job, because it might bring refreshing.

When positive stress and negative stress are mixed, or when all stress states are mixed, the state change is intense. It is considered that the subject has changed his task or has encountered some factors that interfered with him during the process. However, it may be due to temporary factors. In such a situation, it is one way to maintain the task that is being done at moment to see how it changes thereafter.

VII. Conclusion

The method proposed in this paper uses less invasive wearable sensors to detect and classify the stress of a subject who engaged in work. Furthermore, a mechanism is discussed to recommend the next task to be carried out under the constraints of the stress on the subject.

The proposed method estimates stress from biological data acquired from wearable sensors. From the results of this experiment, it is possible to detect stress and classify it into positive or negative ones. It turned out, that we had better use instantaneous values for skin properties when estimating stress. In addition, sweating and skin temperature are found to be important variables for stress estimation, while volumetric pulse wave variability has turned out to be unimportant.

This paper also discusses a way to recommend suitable tasks for stress states, though it is based on qualitative consideration. In the future, we are going to conduct experiments to see if the recommended task is suitable. It is necessary to examine whether the clustering is proper.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Wakabayashi conducted the research and experiments, discussing with Shimakawa and Harada; Wakabayashi analyzed the data; Shimakawa and Harada gave advice on how to analyze the data; Wakabayashi and Shimakawa wrote the paper; all authors had approved the final version.

References

[1] U. Yozgat, S. E. Yurtkoru, and E. Bilginoğlu, "Job stress and job performance among employees in public sector in Istanbul: Examining the moderating role of emotional intelligence," *Procedia - Social and Behavioral Sciences*, vol. 75, pp. 518-524, 2013.

[2] R. M. Yerkes and J. D. Dodson, "The relation of strength of stimulus to rapidity of habit-formation," *Journal of Comparative Neurology and Psychology*, vol. 18, no. 5, pp. 459-482, Nov. 1908.

[3] X. Hou, Y. Liu, O. Sourina, T. Y. R. Eileen, L. Wang, and W. Mueller-Wittig, "EEG based stress monitoring," presented at the 2015 IEEE International Conference on Systems, Man, and Cybernetics, Oct. 9-12, 2015.

[4] G. Jun and K. G. Smitha, "EEG based stress level identification," presented at the 2015 IEEE International Conference on Systems, Man, and Cybernetics. Oct. 9-12, 2015.

[5] Some works

[6] R. S. Lazarus, J. C. Speisman, and A. M. Mordkoff, "The relationship between autonomic indicators of psychological stress: Heart rate and skin conductance," *Psychosomatic Medicine*, vol. 25, no. 1, pp. 19-30, Jan. 1963.

[7] N. Sharma and T. Gedeon, "Objective measures, sensors and computational techniques for stress recognition and classification: A survey," *Computer Methods and Programs in Biomedicine*, vol. 108, no. 3, pp. 1287-1301, Dec. 2012.

[8] C. Setz, B. Arnrich, J. Schumm, R. L. Marca, G. Tröster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable EDA device," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 410-417, Mar. 2010.

[9] C. H. Vinkers, R. Penning, J. Hellhammer, J. C. Verster, J. H. G. M. Klaessens, B. Olivier, and C. J. Kalkman, "The effect of stress on core and peripheral body temperature in humans," *The International Journal on the Biology of Stress*, vol. 16, no. 5, pp. 520-530, Jul. 2013.

[10] G. L. Albright, J. L. Andressi, and A. L. Brockwell, "Effects of stress management on blood pressure and other cardiovascular variables," *International Journal of Psychophysiology*, vol. 11, no. 2, pp. 213-217, Aug. 1991.

[11] H. Takatsu, M. Munakata, O. Ozeki, K. Yokoyama, Y. Watanabe, and K. Tanaka, "An evaluation of the quantitative relationship between the subjective stress value and heart rate variability," *IEEJ Transactions on Electrical and Electronical Engineering*, vol. 120, no. 1, pp. 104-110, 2000. (in Japanese)

[12] George E. Billman, "The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance," Frontiers in Physiology, vol. 4, pp. 1-5, Fed. 2013.

[13] C. Schubert, M. Lamberta, R. A. Nelesen, W. Bardwell, J. B. Choi, and J. E. Dimsdale, "Effects of stress on heart rate complexity -- a comparison between short-term and chronic stress," *Biological Psychology*, vol. 80, no. 3, pp. 325-332, Mar. 2009.

[14] M. Kumar, M. Weippert, R. Vilbrandt, S. Kreuzfeld, and R. Stoll, "Fuzzy evaluation of heart rate signals for mental stress assessment," *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 5, pp. 791-808, Oct. 2007.

[15] D. Sugahara, Y. Araki, and M. Sugie, "Effect of different levels of arousal of positive emotion on inhibition of stress," *Bulletin of Hokuriku Psychological of Japan*, vol. 4, no. 1, pp. 11-20, Jul. 2015. (in Japanese)

[16] Empatica. (August 2021). E4 Wristband | Real-time physiological signals. [Online]. Available: https://www.empatica.com/research/e4/

[17] Polar. (August 2021). H10 | Accessories for heart rate monitors, fitness trackers and cycling sensors. [Online]. Available: https://www.polar.com/en

[18] F. Badilini, P. Maison-Blanche, and P. Coumel, "Heart rate variability in passive tilt test: Comparative evaluation of autoregressive and FFT spectral analyses," *Pacing and Clinical Electrophysiology*, vol. 21, no. 5, pp. 1122-1132, Jun. 2006.

[19] M. Malik, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *Circulation*, vol. 93, no. 5, pp. 1043-1065, Mar. 1996.

[20] J. Huang and C. X. Ling, "Using AUC and accuracy in evaluating learning algorithms," *IEEE Transaction on Knowledge and Data Engineering*, vol. 17, no. 3, pp. 299-310, Mar. 2005.

Naoya Wakbayashi is a student at the graduate school of information science and engineering in Ritsumeikan University. His research interests include data engineering and IoT.

Prof. Hiromitsu Shimakawa received Ph.D. degree from Kyoto Univ. in 1999. Since 2002, He has worked in Ritsumeikan Univ. as a professor. His research interests include data engineering, usability, and integration of psychology with IT. He is a member of IEEE and ACM.

Dr. Fumiko Harada received B.E., M.E, and Ph.D. degrees from Osaka University in 2003, 2004, and 2007, respectively. She joined Ritsumeikan University as an assistant professor in Ritsumeikan University in 2007, and is currently a counselor of Connectdot co. ltd. She engages in the research on real-time systems and data science. She is a member of IEEE.