WORKPLACE STRESS ESTIMATION METHOD BASED ON MULTIVARIATE ANALYSIS OF PHYSIOLOGICAL INDICES

Hirohito Ide, Guillaume Lopez, Masaki Shuzo, Shunji Mitsuyoshi, Jean-Jacques Delaunay and Ichiro Yamada
The University of Tokyo, School of Engineering, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

Keywords: Stress monitoring, Wearable sensors, Multivariate analysis, Virtual healthcare.

Abstract: In this research, we have been developing a new integrated analysis method of multiple physiological signals to estimate stress in daily life, which is important in depression screening and life-style related diseases prevention. Experiments have been carried out on 100 participants, measuring electrocardiogram, pulse wave, breath rhythm, and skin temperature in four patterns of psychological states; relax state, normal stress state, monotonous stress state, and nervous state. The newly developed stress state estimation method relies on the integrated analysis of nine physiological indices related to stress that have been extracted from the four measured physiological signals. Because variation range of each index is different between individuals and types of stress, we divided estimation process into three steps. For each step, we performed cross-validation using various classification schemes to select the most relevant set of indices that enable estimation of stress state with few influences of individual variations. Through this method we could achieve 87% accuracy for stress detection, and 63% accuracy for stress type classification. Finally a validation study was performed to confirm this method can be an effective solution to estimate various types of stress state regardless of individuals.

1 INTRODUCTION

Nowadays, most developed countries are facing a serious problem with the increasing number of diseases caused by excessive stress, not only mental disorder diseases (depression, etc.), but also lifestyle-related diseases (hypertension, metabolic syndrome, etc.). Indeed, when we are subject to excessive stress, we tend to overeating, drinking alcohol, smoking and such lifestyle-related disease risk factors. In this introduction, we define our field of study, describe the existing approaches and their issues, and present our approach to address them in the work reported in this paper.

1.1 Background and Definition of Stress

Current stress detection methods, when not an afterwards conclusion, rely on inquiry sheets or interviews with a medical specialist. Though, because stress is so pervasive in our social activities, there is an inherent need to be able to monitor stress continuously in daily life, in a seamless way, and over extended periods. It is important to propose such new system for personal continuous stress monitoring, which would enable prevention of serious stress-related health disorders, through a seamless and regular screening of stressful experiences an individual is exposed during his daily life activities. Such system would benefit both individuals by providing regular feedback about their stress, as well as physicians by supporting patient status monitoring and evaluation with objective and in context information.

We hear a lot about stress, but what is it? Taber’s Cyclopedic Medical Dictionary defines stress as “the result produced when a structure, system or organism is acted upon by forces that disrupt equilibrium or produce strain”. In simpler terms, we will consider stress as the result of any emotional, physical, social, economic, or other factors that require a response or change. It is generally believed that some stress is okay (sometimes referred to as “normal stress” or “positive stress”), but when it occurs in amounts that cannot be handled, both mental and physical changes may occur.

In our study, we focus on “Workplace stress,” which we define as the physiological responses that can happen when there is a conflict between job demands on the person and the amount of control this person has over meeting these demands. In gen-
eral, the combination of high demands in a job and a low amount of control over the situation can lead to stress. Regarding the amount of oxygen consumption, work strongly stimulates heart activity (Carroll et al., 2009), and both qualitative and quantitative augmentations are reported to have a strong correlation with dissatisfaction, decline in self-evaluation, which are risk factors of mental disorders (Araki and Kawakami, 1993). That is why on both mental disorder diseases and lifestyle-related diseases point of views, there is a considerably strong need in workplace stress monitoring.

As stated by the Canadian Mental Health Association (http://www.cmha.ca/), stress in the workplace can have many origins: Fear of job redundancy, pressure to perform, increased demands for overtime due to staff cutbacks, layoffs due to an uncertain economy act as negative stressors. Though, among these origins, we chose to remove economical and non-work-related stressors, to focus on the following three workplace stress categories as partly proposed by Shimono et al. (Shimono et al., 1998).

- Monotonous stress: stress accompanied by a tedious feeling when repeating a work with little content changes for a long continuous time (job redundancy, frequent overtimes, etc.).
- Nervous stress: stress accompanied by a feeling of tension when performing a work that cannot afford any miss (pressure to perform, speech, meeting with hierarchical superiors, etc.).
- Normal stress: stress accompanied by any feeling different from above described, when performing a basic work (in other words basic work that does not generate extra stress).

### 1.2 Current Technological Solutions and their Issues

Traditionally, personal medical monitors have been used only to perform data acquisition. Typical examples are holter monitors that are routinely used for electrocardiogram (ECG) and electroencephalogram (EEG) monitoring. Recently, with the miniaturization and improved performances of micro-sensors, wearable computing, and wireless communication technologies (Fukuda et al., 2001; Itao, 2007), a new generation of wearable intelligent sensors have been developed (Jovanov et al., 2000). Such devices can significantly decrease the number of hospitalizations and nursing visits (Heidenreich et al., 1999) by acting as a personal quotation: virtual health adviser that can warn the user of a medical emergency or contact a specialized medical response service. A wearable health-monitoring device using a personal area network (PAN) or BAN can be integrated into a user’s clothing (Park and Jayaraman, 2003), though such system organization is unsuitable for lengthy, continuous monitoring, particularly during normal activity.

We can classify prior research related to stress study using wearable physiological sensing into the following three categories.

1. Studies that demonstrate the causal relationship between stress and changes in physiological indices (Kim et al., 2008; Ohsuga et al., 2001; Schubert et al., 2009)
2. Studies that evaluate qualitatively and/or quantitatively the stress issued by an external stimulus (Kotlyar et al., 2008; Watanabe et al., 2008)
3. Studies that estimate the occurrence or not of stress based on the observation of changes in physiological indices (Aasa et al., 2006; Fukuda et al., 2001; Itao et al., 2008)

Aiming at stress monitoring during daily life activities, our research corresponds to the third category. This category is composed of two groups of methodologies, being methodologies to retrieve stress changes based on the observation of long-term evolution for a single physiological index, and methodologies that build models for stress status output from input of physiological indices, based on multivariate analysis. We consider these groups of methodologies has four big issues that need to be addressed.

1. As physiological indices are strongly influenced by individual differences, their values on stress occurrence are different depending on each individual (Miyake, 1997). Therefore, it is necessary to pick-up physiological indices that are less prone to individual differences when estimating stress.
2. Depending on the type of stress (in other words the type of emotion), reacting physiological indices are different, so that it is difficult to estimate stress in detail from a single physiological index (Miyake, 2001; Ohsuga et al., 2001; Shimono et al., 1998).
3. Stress status output models application is often limited to only one specific individual, and cannot output stress status correctly for a different person (Ohsuga et al., 2001; Soda and Narumi, 2007).
4. Models are often limited to an output of having or not stress, and do not estimate stress status in details (i.e. stress type) (Shin et al., 1998; ?).

According to above statements, the study we present here aims at addressing the described issues.
by establishing a detailed and high-generality stress estimation method, in other words a stress type estimation method using physiological indices less prone to individual differences.

The remainder of the article is structured as follows. Section 2 described the physiological signals and indices we selected as an input to our proposed multivariate analysis method and the experimental set-up used to measure these signals. Section 3 presents the experimental procedure and the pre-processing of collected data we executed to build-up a reliable database of physiological indices corresponding to targeted three types of workplace stress. Section 4 describes our proposed method, evaluates its efficiency using experimental data, and validates its reliability regarding the database. Finally, Section 5 sum-up our findings, raises remaining issues, and opens a short view on future implementations we plan to pursue.

## 2 PHYSIOLOGICAL INFORMATION USEFUL TO STRESS ESTIMATION

For monitoring stress, we focus on autonomic nervous system activity, though we don’t use EEG due to its difficult processing that makes it difficult to use for a real-time stress monitoring solution.

It is known that the autonomic nervous system influences the activity of the heart, the breath, the lung, and the skin activities. If there is any change on the autonomic nervous system due to stress, it should be detectable through the activity of these physiological elements. Typical studies of the autonomic nervous system activity monitoring consist in ECG’s heart beats $R$ peak time interval variations (RRV: $R-R$ interval variations) frequencies analysis, in which strength of low frequencies zone (LF: 0.04-0.15Hz) reflects sympathetic nerve’s activity, and strength of high frequencies zone (HF: 0.15-0.4Hz) reflects parasympathetic nerve’s activity. Then, $LF/HF$ power ratio is an indicator of activity dominant nervous system (large: sympathetic nerve is dominant, small: parasympathetic nerve is dominant).

Though these studies reported that RRV spectral analysis was effective to evaluate the physical and mental loads by quantifying respectively the activity level of sympathetic and parasympathetic nerves (Akselrod et al., 1981; Itao et al., 2008), this index is known to be different according to the age, sex and the individual variation (Miyake, 2001). The physiological indices should meet the following two conditions: the first is that they can reflect the categories of stress, and the second is that individual differences are not large. In this study, we decided to measure simultaneously electrocardiogram (ECG), pulse wave by photoplethysmography (PPG), breath, and temperature of finger’s skin.

From these four physiological signals, we extract the following nine physiological indices, which we adopted as the basic information for stress type estimation.

- From ECG: $HR$ (Heart rate), $RRV$, $LF/HF$
- From PPG: $t_{PAT}$ (pulse arrival time)
- From breath: $f_C$ (respiratory central frequency), $|f_P-f_C|$ (absolute value of difference between $f_C$ and peak frequency), $t_B$ (breath time), $std_{t_B}$ (derivation of breath time)
- From finger’s skin temperature: $T_{F}$ (average temperature of the finger’s skin)

To collect above selected physiological signals, we used multi-channels biological amplifier (Polymate, Digitex lab. Co. ltd.) that is basically composed of an electrocardiograph (ECG) and an ear clip type photoplethysmograph (PPG), but to which optional sensors such as belt-type breath sensor, and temperature sensor needed for our study can be connected.

## 3 STRESS CORRELATED PHYSIOLOGICAL INDICES DATABASE BUILD-UP

For predicting stress by using psycho-physiological indices, it is necessary to build a database based on these indices at the situation when people faced targeted different types of stressor and effectively get the expected stress reaction. The following paragraphs will present the experimental procedure we defined to stimulate efficiently the three types of stress reactions targeted, and how we extract a sufficient number of high quality data sets among the whole measurement.

### 3.1 Experimental Procedure for Physiological Indices Data-sets Measurements

In this study, we focus on normal stress, monotonous stress, and nervous stress, the most usual stresses that may occur at a workplace. To collect data that will populate our database of a person under these different types of stress we used the Paced Auditory Serial
Addition Test (PASAT, see Fig. 1), which has an acknowledged authority among scientific community as to having high reproducibility (Al’Absi et al., 2005; Willemsen et al., 1998). Based on former research work about monotonous stress (Yamada and Miyake, 2007) and nervous stress (Al’Absi et al., 1997), we defined the following three types of PASAT tasks.

- **PASAT1**: 5 minutes PASAT task to stimulate a normal stress reaction
- **PASAT2**: 60 minutes PASAT task to stimulate a monotonous stress reaction
- **PASAT3**: 5 minutes PASAT task combined with reward cutting on miss, to stimulate a nervous stress reaction

Figure 1: Scheme presenting the principle of PC based PASAT. PASAT task consists in adding consecutive single digit numbers presented by voice continuously.

We have built a dedicated black room to avoid any environmental light and noise disturbance, and a PC interface to execute PASAT and answer to it in a simple and quick way. The experimental protocol flow is described on Figure 2. We have performed above described experimental for each PASAT task with respectively 48, 18, and 46 participants, whose ages were from 15 to 47 years old.

However, to be sure that before performing the assigned PASAT task each participant was not already in a psychological status that may influence expected stress reaction, due to the lack of sleep, overwork, alcohol and such, subjective assessments were conducted before and after PASAT task. Subjective assessment is done on dedicated PC interface composed of a list of short questions to which the participant can answer using a mouse to set the Visual Analog Scale (VAS) for each question corresponding to scores from 0% to 100%. To evaluate the effective reaction to each specific stress type stimulation, we used the scores of related questions, among which the average score of questions about feeling exhausted, ineffective, and depressed, for monotonous stress, and questions about feeling strained and palpitating for nervous stress.

### 3.2 Data Quality Evaluation Results

To verify if each expected stress type effectively occurred or not, we performed a test to check the significant difference (t-test) among the different PASAT tasks, using the subjective stress type score variation between subjective assessment before and after the PASAT task corresponding to this stress type. Figure 3 shows the results significant difference examination among the three PASAT tasks. Considering only the group of participants with more than 60% of accuracy rate in PASAT1 and PASAT3 tasks (PASAT1: 19, PASAT2: 18, PASAT3: 21), we could verify high significant differences. From these results, we can consider that expected stress reaction has effectively occurred for experimental participants, respectively corresponding to the stress type stimulated by each PASAT task.

Figure 2: Experimental procedure flow. A five minutes-long relaxing time was set before each test to be able to evaluate physiological indices in a relax situation.

Figure 3: Subjective assessment scores in each PASAT. Compared to PASAT1, “boredom” score (monotonous stress) is higher with PASAT2, while “high-tension” score (nervous stress) is higher with PASAT3.

## 4 A NEW METHOD FOR WORKPLACE STRESS TYPE ESTIMATION

According to the purpose of our study described in introduction, we propose a new method for stress type estimation applicable with a high-generality. After describing the points of the method, we will present the results of its efficiency evaluation using the experimental data, and validate its reliability regarding the physiological indices database.

### 4.1 Estimation Method

The physiological indices that we should use for a good estimation should meet the following two con-
ditions: the first is that they are able to reflect workplace stress types, and the second is that they are less prone to individual differences. To specify the physiological indices set that achieve these conditions, we performed a cross-validation with the available data-sets, and identified the combination of physiological that result in the highest estimation accuracy. Here we used the leave-one-out cross-validation (LOOCV) since we don’t need to tackle actually computation performance issues and we still have a relatively small number of samples in the data-set. In our case, LOOCV involves using a single person data as sample data for validation, and the remaining persons’ data as the sample data for training, which is repeated such that each person’s data id used once as the validation data sample.

Accordingly, we have defined an original multi-steps logic as shown on Figure 4, to perform “individual-independent” stress-type estimation with high accuracy. The first step aim at discriminating with high accuracy stress status from and relax status, in other words the presence of any workplace stress. Once some stress reaction presence has been detected, the second step consists in discriminating normal stress and other workplace stress types, which means identifying the harmfulness of the stress. Finally, if physiological indices are identified as reflecting some extra stress, the third step consists in discriminating the physiological reaction between nervous stress and monotonous stress.

![Diagram](image-url)

Figure 4: Proposed multi-step estimation procedure with integrated indices selection. It is composed of three steps to gradually estimate the psychological status corresponding to input physiological indices.

According to the results of significant difference study between the three PASAT tasks, which requires good task performance to guarantee reliable subjective evaluations, we selected 39 persons among the whole participants (13 persons by PASAT task type). For each selected participant, we extracted five data-sets for each stress type, which contain the physiological indices during the five minutes of PASAT (the last five minutes for PASAT2) extracted with a one minute rolling window. Five data-sets for relax status for each selected participant were also extracted in the same way. As a result, we obtain a training database filled with 390 data-sets, each representing the physiological reaction to one minute exposure to relax (195 data-sets), normal stress (65 data-sets), monotonous stress (65 data-sets), and nervous stress (65 data-sets) situations.

Then, to evaluate our method in a first step independently from the classification scheme adopted, we tested 24 classification schemes among which Linear Discriminant Analysis (LDA), Support Vector Machines (SVM), and so forth. At each step of the estimation procedure and for each adopted classification scheme, we identified and used the best set of physiological indices (those less prone to individual differences).

4.2 Validation of the Method

We compared the following four methods combined with cross validation based on the selected 39 participants to validate our multi-steps logic combined with integrated physiological indices selection. The results of comparison with other conventional methods are shown on Figure 5, validating the efficiency of our proposed method.

- Method 1 proceeds to only one discrimination step, using only $\frac{LF}{HF}$ index
- Method 2 proceeds to only one discrimination step, using all provided physiological indices
- Method 3 proceeds to only one discrimination step, using identified best physiological indices
- Method 4 proceeds to multi-step discrimination with best physiological indices selection at each step (our method)

4.3 Reliability Study of the Proposed Stress Estimation Method

In previous paragraph we identified the physiological indices that are less prone to individual differences. However, as a condition to ensure reliability of the estimation using these physiological indices, it is necessary to validate its application to stress estimation for participants whose physiological indices were not included into the training database. Precisely, we raised the following three conditions as an assurance of reliability of the proposed workplace stress estimation method.
The selected set of physiological indices input into the stress classification scheme does not belong to the trained database (low dependence on database).

Accordingly, we selected three sets of physiological indices combinations with high accuracy, and used these three sets with the 24 classification schemes, resulting in 72 possible classification schemes. Among these we have investigated the classification schemes that meet the defined reliability conditions, with a particular focus on the third one, which corresponds to the dependence on the training database.

In the investigation process to index database dependence criteria, we increased gradually from 9 to 39 the number of participants whose data are included in the database. At each indentation, the maximum estimation accuracy is calculated for each classification scheme by leave-one-out cross-validation. For each database size, the smaller the difference between estimation accuracy using selected scheme and the maximum accuracy is, the more robust the scheme is. Indeed, a small difference means that selected scheme would keep high estimation accuracy even with an ever-growing database that will reflect more and more human diversity.

In our study, we defined the reliability criterion to be a difference in accuracy lower than 3%. The difference between stress estimation accuracy using below described method (Method A) and maximum stress estimation accuracy method (recomputed each time) reaches the required reliability criterion of 3% when database is composed by data-sets from 29 participants above 39 max participants (see Fig. 6).

- Classification scheme: SVM with Gaussian Kernel ($\sigma=2.75$)
- Physiological indices selected for step 1: $f_G$, $|f_F-f_G|$, $t_{PAT}$, $T_F$, $RRV$
- Physiological indices selected for step 2: $stdT$, $T_F$, $LF/HF$, $HR$
- Physiological indices selected for step 3: $|f_F-f_G|$, $i_E$, $stdT$, $t_{PAT}$, $RRV$, $LF/HF$

Then, we used this 29 participants’ database as the reference, and calculated for all possible classification schemes a reliability index $U$ using equation 1. In equation 1, $x_1$, $x_2$, and $x_3$, are respectively the average accuracy, the standard deviation of the accuracy, and the database dependency of the adopted stress classification schemes ($i=1,2,...,72$), which should result in positive values of $U$ index for reliable classification schemes, and small values of $U$ index for non-reliable schemes.

So, the classification scheme corresponding to the highest $U$ value is the one which characteristics are described above. This classification scheme enables an accuracy of 64% for the estimation of workplace stress type reaction, and 89% for occurrence or not of workplace stress reaction, standard deviation of 28%, and database dependence index of 29 (max 39). In the opposite, classification scheme with lowest $U$ value, which uses a fuzzy logic algorithm, also enables an accuracy of 64% for the estimation of workplace stress type reaction, but with a standard devia-
\[ U_i = \frac{x_{1i} - \bar{x}_1}{\sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{1i} - \bar{x}_1)^2}} - \frac{x_{2i} - \bar{x}_2}{\sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{2i} - \bar{x}_2)^2}} - \frac{x_{3i} - \bar{x}_3}{\sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{3i} - \bar{x}_3)^2}} \]  

(1)

Table 1: Best discrimination accuracy for stress presence and type depending on the method used.

<table>
<thead>
<tr>
<th>Physiological Indices Used</th>
<th>LF/HF only</th>
<th>All indices</th>
<th>Best Indices Selection</th>
<th>Best Indices Selection and Multi-steps Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress Types</td>
<td>26%±2%</td>
<td>48%±8%</td>
<td>56%±3%</td>
<td>63%±5%</td>
</tr>
<tr>
<td>Stress Presence</td>
<td>63%±1%</td>
<td>83%±7%</td>
<td>87%±3%</td>
<td>-</td>
</tr>
</tbody>
</table>

This last point confirms that in our reliability index \( U \), not only the average accuracy is dominant, but it is also strongly influenced by standard deviation and database dependency index.

5 CONCLUSIONS

In this research study we proposed a new method to address the problem of influence of both individual and stress type on physiological indices values. Based on the collection of a large amount of physiological data-sets under different stress types exposure, we evaluated the efficiency of our method for accurate stress type estimation by comparing it with other conventional methods. Table 1 presents a detailed result of efficiency analysis, showing that introducing the process of best fit physiological indices selection, has a great impact on stress reaction presence estimation accuracy, while the addition of intelligent multi-step discrimination process is essential to improve the accuracy of workplace stress type estimation. In addition, these results were issued with data from 39 different participants in age and sex, demonstrating our proposed method to be less prone to individual differences.

However, to achieve our goal of a system as shown on Figure 7 for personal continuous stress monitoring in daily life, we still have to tackle many issues among which the ones we consider in priority are listed below.

- Continue improving stress type discrimination accuracy while limiting at best the number of sensors worn
- Investigate the possibility to discriminate more types of stress
- Investigate a method that enables to evaluate quantitatively stress level
- Application of the proposed method to the estimation of chronic stress

Figure 7: Schematic of the virtual stress checker system that will implement our method. The wearable terminal collecting sensor information communicates through network connectivity with an online sub-system, at which detailed processing and various feedback can be performed (Faudot et al., 2010; Lopez et al., 2009).

ACKNOWLEDGEMENTS

This research was supported by Japan Science and Technology Agency’s (JST) strategic sector for creation of advanced integrated sensing technologies for realizing safe and secure societies: research project on “Development of a Physiological and Environmental Information Processing Platform and its Application to the Metabolic Syndrome Measures”.

REFERENCES


