Inferring Low-level Mental States of Mobile Users from Plethysmogram Features by Regression Models based on Kernel Method

Toshiki Iso
Research Laboratories, NTT DOCOMO, INC., Kanagawa, Japan

Keywords: Earlobe Plethysmogram, Low-level Mental States, Gaussian Process Regression, Support Vector Regression, Feature Extraction, Automatic Relevance Determination, LF/HF, Lyapunov Coefficient.

Abstract: To infer user's response when using mobile services without direct interrogation, we propose an algorithm that analyzes earlobe plethysmograms to determine low-level mental states such as 'relax', 'concentration', 'awake'. We use subject's responses acquired in a subjective evaluation as indicative of low-level mental states when subjects use some mobile contents. In order to draw an inference of low-level mental states based on plethysmogram features, our proposed algorithm uses a kernel-based regression model such as Gaussian Process Regression (GPR) or Support Vector Regression (SVR). Our evaluations show that features effective for inferring user's low-level mental states can be extracted from plethysmograms by using regression and Automatic Relevance Determination (ARD); the regression performance of GPR and SVR are described.

1 INTRODUCTION

Service providers want to be able to detect the user's low-level mental states such as 'concentration' and 'vagueness'. Because plethysmogram can be easily detected by non-invasive means (Allen, 2007), low-level mental states based on plethysmogram are being used for medical applications such as the diagnosis of dementia in aged people(Oyama-Higa et al., 2006), analysis of physio-psychological status(Oyama-Higa and Miao, 2005), and detecting snooze driving(Vicente et al., 2016).Recently, Internet service providers also want to know the user's low-level mental states in order to discover their responses for improving the services. In more detail, the desired user responses are not high-level mental states such as positive or negative feelings about the content, but whether the services enhance or weaken the user's interest. Information on whether the user is actively searching for desired contents or aimlessly accessing web pages allows the service provider to control the type of banner advertisements displayed.

While plethysmography appears attractive in estimating the user's low-level mental states, there is no commonly accepted method for detecting the states with minimal user interference. Diagnosis of apnea syndrome(Loube et al., 1999),(Romem et al., 2014) and mental illness and detecting drunk driving(Murata et al., 2011) process plethysmograms gathered by attaching sensors to the subject's body. Fixing a sensor to the user's finger makes mobile services very cumbersome and unattractive. We need a method with minimal user constraints that allows context-aware services to realize effective control of advertisements depending on the user's low-level mental states. Service providers currently are conducting questionnaires to determine the user's responses. While overall service impressions can be gathered, details are difficult to acquire because users have difficulty remembering and verbalizing their impressions. One method detects the user's responses by analyzing Internet access history, but it is impossible to distinguish between focused and non-focused browsing. Another proposed method analyzes user eye movements(Wong et al., 2014). Unfortunately, combining eye-tracking with smartphone device use in mobile environments is not realistic.

Y. Cho et al. proposed to infer mental stress from photoplethysmographic data acquired from RGB and thermal images(Cho et al., 2019),(Alafeef, 2017)This scheme is difficult to use in mobile environments because of its constraints such as assuming the user's finger and face occupy fixed positions. More importantly, as photoplethysmography extracts only heart rate features, it can detect few mental states. However, if plethysmograms can be detected without undue user constraints, it would be useful for inferencing low-level mental states because apnea syndrome,

Iso, T.
Inferring Low-level Mental States of Mobile Users from Plethysmogram Features by Regression Models based on Kernel Method.
DOI: 10.5220/0008977002500257
In *Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2020) - Volume 4: BIOSIGNALS*, pages 250-257
ISBN: 978-989-758-398-8; ISSN: 2184-4305

Figure 1: Wearable device mounted on earlobe for plethysmogram capture.

mental illness, and detecting snooze driving, use features extracted from plethysmogram data such as LF/HF and Lyapunov by chaos analysis(Oyama-Higa and Miao, 2005),(Sumida et al., 2000). Low-level mental states can be detected by analyzing plethysmograms as heart rate control is strongly influenced by sympathetic nerves. Though plethysmograms are being used for the analysis of low-level mental states, it is not clear how to extract valid features from plethysmograms nor which statistical analysis models are valid for recognizing the mental states.

In order to infer low-level mental states, we propose a method that uses earlobe plethysmograms; it uses the already proposed wearable mobile device that places no undue constraints on user operations(Kimura et al., 2013). The interface sets plethysmogram sensors on the ear lobes areas of the user. We identify which plethysmogram features are valid for recognizing the key mental states and evaluate the performance of different statistical analysis models in inferring low-level mental states.

In the next section, we describe feature extraction from plethysmograms and subjective evaluation features for determining low-level mental states. We then explain our proposals for selecting plethysmogram features, regression analysis of plethysmograms, and deriving subjective evaluation features by kernel-based regression techniques. Finally, we explain our experiments on inferencing low-level mental states and future work.

2 FEATURE EXTRACTION

In this section, we describe feature extraction from plethysmograms and the subjective evaluation features needed for determining low-level mental states.

2.1 Subjective Evaluation Features of Low-level Mental States

In order to determine low-level mental states, 13 subjects wearing earlobe plethysmogram sensors were subjectively evaluated; they accessed eight types of mobile services (each for about 5 to 15 minutes) on a smartphone: The contents contained the subject's favorite contents such as movies and games and contents uncomfortable on a biological level such listening to strange noise. These contents were intended to force the subjects to experience known low-level mental states. Immediately after the experiments, videos of the services experienced were reviewed. The 13 subjects scored each subjective evaluation items (eight types) with one of seven numerical levels (seven degrees) every 5 minutes. The eight subjective evaluation items were 'boredom', 'concentration', 'satisfaction', 'interest', 'joy', 'enjoyment', 'anger', and 'sadness'; this was done because it is easier for subjects to indicate their feelings about service contents than their low-level mental states such as 'relax' and 'concentration'. However, in order to obtain reliable evaluation data, only scores that the subject felt confident in making were used. Figure 2 plots the relationships between the scores of the subjective evaluation items as indicated by a subject's watching a favorite video. We can see that these relationships have little direct correspondance other than the relationship between "satisfaction" and 'interest'. In other words, the subjective evaluation items are basically independent of each other. Then, under the assumption that all scores of the subjective evaluation items are indicative of low-level mental states such as 'relax' and 'concentration', we define low-level mental state $y(t)$ at time t as follows:

$$y(t) = \frac{1}{(Sr * M)} \sum_{k=1}^{M} Score_k(t) \qquad (1)$$

where $Score_k(t)$ represents the score of the kth evaluation item at time t by the subject. Sr and M represent number of judgements possible (seven degrees) and number of subjective evaluation items (eight types), respectively. Using just the direct summation of the subjective evaluation scores, not the items individually, fails to distinguish positive from negative low-level mental states. Note that $y(t)$ does offers service providers valuable information about the user's responses.

2.2 Plethysmogram Feature

As described above, we obtained earlobe plethysmograms of 13 subjects when using mobile services. In

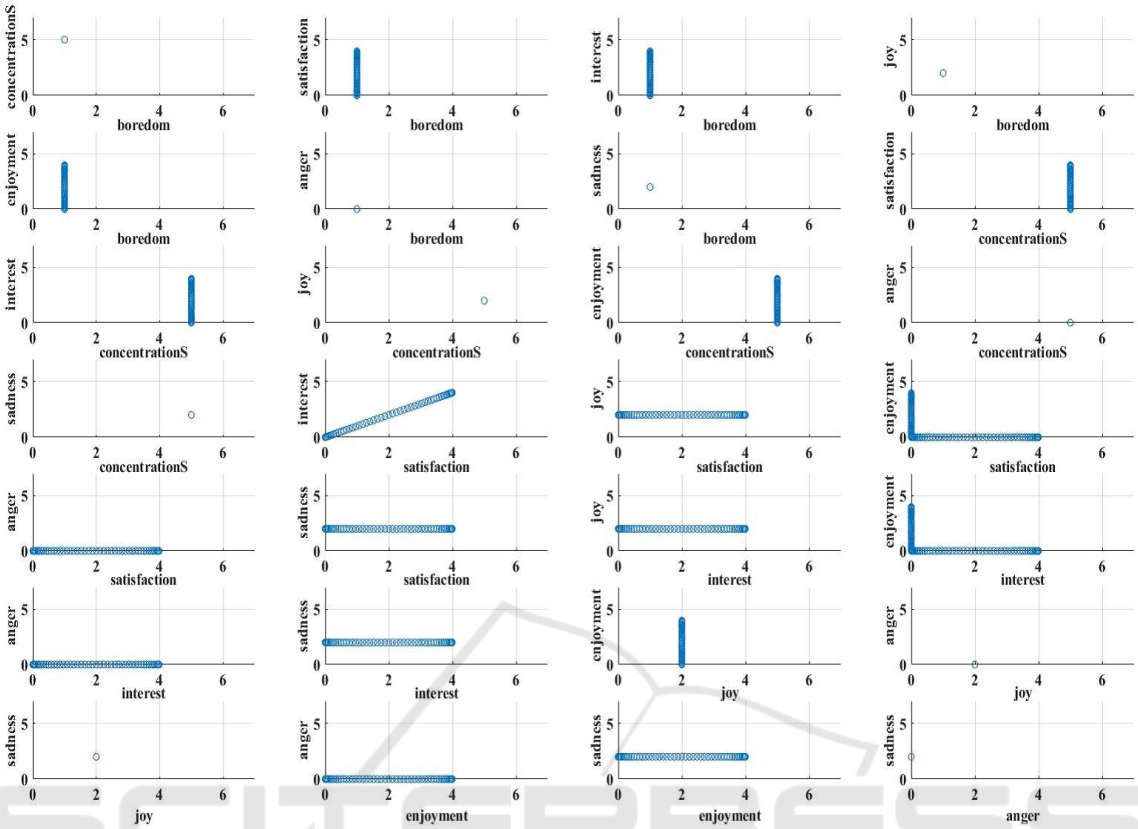

Figure 2: Linearity relations between subjective evaluation items.

Figure 3: The earlobe plethysmogram device made by TAOS Institute, Inc.

the experiments, we used the earlobe plethysmogram device made by TAOS Institute, Inc. (see Figure 3) because more accurate data can be extracted by tight coupling to the subject's earlobe.

The upper part of Figure 4 shows typical time-series data from the earlobe plethysmogram device. We extract four features from the data.

Feature Type1: Second Derivative

The upper part of Figure 4 plots the second derivative of the plethysmogram(Mohamed, 2012). By detecting five extreme values such as from a-wave to e-wave, we calculate SDPTGAI (second derivative of plethysmogram aging index) at time t as follows:

$$x_1(t) = \frac{(b-c-d-e)}{a} \tag{2}$$

The SDPTGAI feature represents a measure of vascular stiffness.

Feature Type2: Heart Rate Variability

$x_2(t)$ at time t can be calculated as peak-to-peak interval times; that is, the second derivative of plethysmogram. Thus heart rate $x_3(t)$ at time t can be calculated from PPI $x_2(t)$ as follows:

$$x_3(t) = \frac{60}{x_2(t)} \tag{3}$$

Feature Type3: Frequency Spectrum

After removing outliers from PPI $x_2(t)$, we calculate the power spectral density $PSD_{all}(t)$ from $t - ws_{fs}/2$ to $t + ws_{fs}/2$ where window size is set at $30[sec]$. We then extract the low-frequency component $PSD_{LF}(t)$ in the frequency range from $0.04[Hz]$ to $0.15[Hz]$ and the high-frequency component $PSD_{HF(t)}$ in the frequency range from $0.15[Hz]$ to $0.40[Hz]$ from PSD_{all} (see Figure 5).The low-frequency component represents the responses of blood pressure fluctuation (Mayer wave) and sympathetic activity(Shaffer

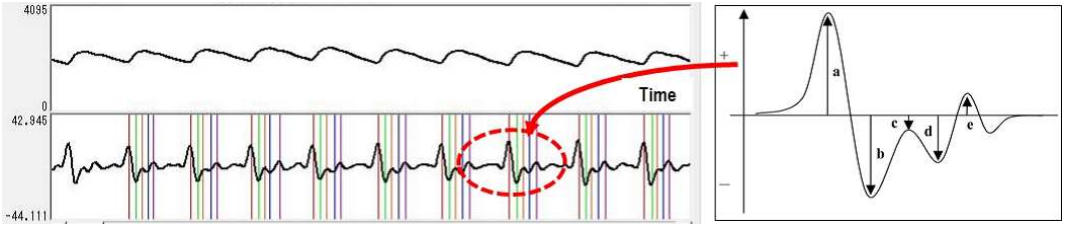

Figure 4: Plethysmogram and second derivative of plethysmogram.

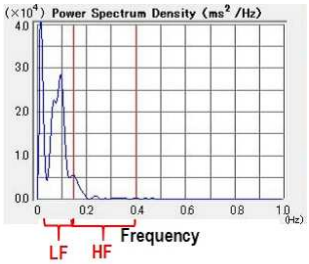

Figure 5: Low-frequency component and high-frequency component.

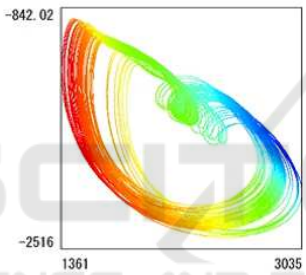

Figure 6: Chaos attractor based on plethysmogram (trajectory projected).

and Ginsberg, 2017). The high-frequency component represents parasympathetic activity and respiratory variation. While the high-frequency component is present only when the parasympathetic nerve is tense (autonomic nerve is quiescent as in 'relax'), the low-frequency component is present when both the parasympathetic and sympathetic nerve are tense. We define Feature Type3 as follows:

$$x_4(t) = \log \frac{PSD_{HF}(t)}{PSD_{all}(t)} \tag{4}$$

$$x_5(t) = \log \frac{PSD_{LF}(t)}{PSD_{all}(t)} \tag{5}$$

$$x_6(t) = \log \frac{PSD_{LF}(t)}{PSD_{HF}(t)} \tag{6}$$

MeanHR $x_7(t)$ is calculated as the mean heart rate from $t - ws_{fs}/2$ to $t + ws_{fs}/2$.

Feature Type4: Chaos Attractor

According to Tsuda et al. (Tsuda et al., 1992), plethysmograms can be represented as a function of four dimensional state variables, while the trajectory mirrors a four dimensional chaos attractor. Figure 6 shows a typical trajectory of a four-dimensional chaos attractor projected into two dimensions. Based on Sano and Sawada(Sano and Sawada, 1985), we calculate the Lyapunov exponent from the four dimensional chaos attractor as follows:

$$x_8(t) = \max_{d=1,2,3,4} \lambda_d \tag{7}$$

where window size is set at $17.5[sec]$. Then,

$$\lambda_d = \lim_{n\to\infty} \frac{1}{n\tau} \sum_{j=1}^{n} \ln A_j e_d^j \tag{8}$$

where d is chaos attractor dimension, n is the number of sampled data points on the attractor, A_j is a linear operator matrix of trajectory tangent vectors, τ is the evolution of time interval, and e_d^j is a basis vector of the tangent space. If λ is positive, the trajectory demonstrates chaotic behavior, but, if λ is not positive, it is characteristic of the turbulence of chaos, which has the potential to be a criterion for distinguishing low-level mental states. We calculate the entropy of the chaos attractor $x_9(t)$ as follows:

$$x_9(t) = -\sum_{i=1}^{N} p_i \log p_i \tag{9}$$

where p_i represents the percentage of all sampled points on the chaos attractor that go through the ith small area in phase space, and N is the number of small areas in phase space. As the entropy represents chaos turbulence, it also has the potential to be a valid criterion for low-level mental states. Feature Type1 to Feature Type4 are defined at time t. In particular, because Feature Type3 and Type4 are calculated using time windows, they can contain some characteristics evidenced in the periods around time t. The above nine feature are different from one another in terms of time sampling. Then, after normalizing each feature against maximum and minimum values, they are synchronized by using time interpolation using the time resolution (sampling frequency $100[Hz]$) of

$$\mathbf{X} = (x_1(t), x_2(t), \cdots, x_9(t))^T \tag{10}$$

To preprocess the time series data, we use the time difference of $\mathbf{X}$. This is necessary because $\mathbf{X}$ is

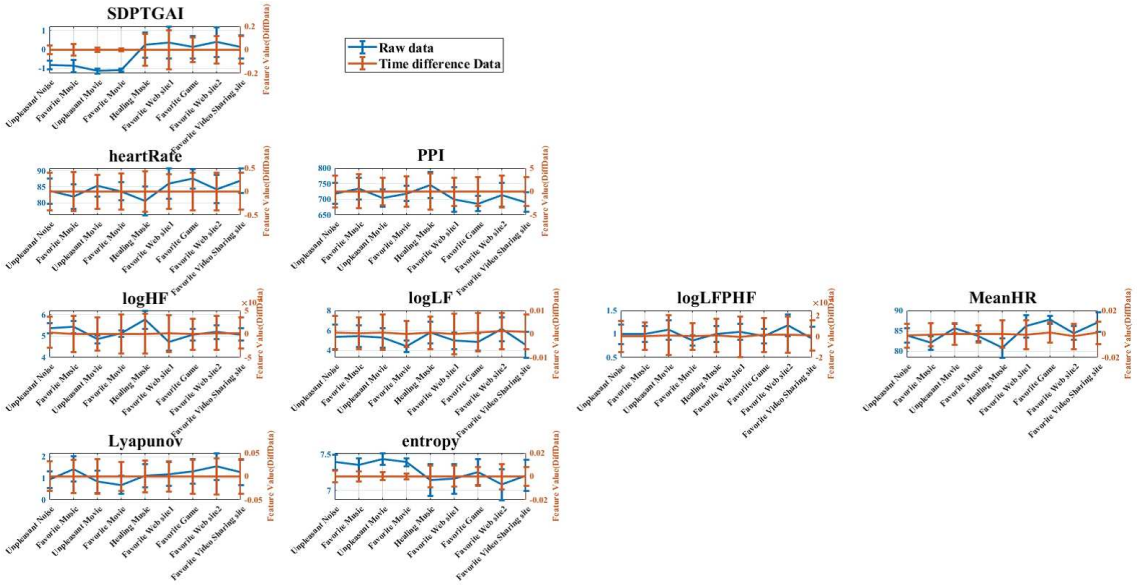

Figure 7: Time-Stationarity of feature vector based on time difference data and raw data.

likely to contain sensor noise and daily fluctuations such as the effects of 'human feelings'.

Time-Stationarity of Feature Vector
In order to extract more stable data for regression analysis, we calculate mean and standard deviation in local time for the above raw data and time difference data as basic descriptive statistical values. Figure 7 shows the results gained from each mobile service. We can confirm that the time difference data exhibits more time-stationarity than the raw data. Hereafter, the time differences on $\mathbf{X}$ are used as input data for inferencing low-level mental states.

3 INFERENCING LOW-LEVEL MENTAL STATES

3.1 Regression Models

Recently, deep learning methods have been applied in a variety of applications, but they need large amounts of high quality training data, which are very difficult to collect. Therefore, we need a regression method that can construct mapping feature vectors from plethysmogram data without over-fitting. To avoid over-fitting in constructing the regression model, we focus on kernel-function-based approaches with constraints. This is because it offers flexible regression with kernel function and suppresses over-fitting.

3.1.1 Gaussian Process Regression

We outline below the Gaussian Process Regression (GPR)(Rasmussen and Williams, 2006). When joint probability $p(y_1,\cdots,y_N)$ of output data points y_i $(i = 1,2,\cdots,N)$ corresponding to N-number of input data vectors $\mathbf{X}_i = (x_{i1},x_{i1},\cdots,x_{id})^T$ follows a multivariate Gaussian distribution; that is, the weight matrix $\mathbf{w}$ satisfies

$$\mathbf{w} \sim \mathcal{N}(0,\lambda^2\mathbf{I}) \tag{11}$$

the estimated output data $\hat{\mathbf{y}}(= (\hat{y}_1,\cdots,\hat{y}_N))$ is given as follows:

$$\hat{\mathbf{y}} \sim \mathcal{N}(0,\lambda^2\Phi\Phi^T) \tag{12}$$

where λ^2 and Φ are variance of w and a matrix of basis functions, respectively. $\mathbf{K} = \lambda^2\Phi\Phi^T$ is a covariance matrix of $\lambda\Phi(\mathbf{X}_i)$. Choosing a kernel function that ensures $\mathbf{K}$ is a positive definite matrix, yields the estimated output data $\hat{\mathbf{y}}$. GPR uses Bayesian estimation based on a kernel function, it avoids the over-fitting problem. GPR is used for not only regression but also Bayesian optimization for determining the optimal parameters in deep learning and feature selection as described below. We can analyze the fluctuations of output data $\hat{\mathbf{y}}$ to identify low-level mental states because standard deviations of Gaussian distributions can represent fluctuations in regression as GPR can basically inference the Gaussian distribution at each point of $\hat{\mathbf{y}}$.

3.1.2 Support Vector Regression

We outline Support Vector Regression (SVR)(Smola and Schölkopf, 2004). Unlike SVM (Support Vector

Machine), SVR uses criterion function S_r, which minimizes error function **h** between the estimated output data $\hat{\mathbf{y}}$ and the actual output data **y** (not the number of misclassified data), as follows:

$$S_r = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{d=1}^{D}\mathbf{h}(\hat{\mathbf{y}}-\mathbf{y}) \quad (13)$$

where D represents the dimension of **y**.

This also avoids over-fitting because SVR also uses constraint condition $\frac{1}{2}\|\mathbf{w}\|^2$. Therefore, SVR can yield sparse regression because it applies L1-optimization to criterion function S_r. Note that, GPR is a generative model approach based on Ridge regression with a kernel trick, and so is more sensitive to data than SVR.

3.2 Selection of Plethysmogram Feature

As described in the Introduction, different applications use different plethysmogram features. We identify the valid plethysmogram features by Automatic Relevant Determination (ARD) with GPR(Wipf and Nagarajan, 2008). ARD can estimate the degree of contribution of explanatory variables to objective variables. In the above section, when we use GPR, we need to select the kernel function. For the purpose of feature selection, we use the squared exponent kernel function as follows:

$$\mathbf{K}(\mathbf{x_n}, \mathbf{x_{n'}}) = \exp\left(-\frac{1}{2}\sum_{d=1}^{D}\left(\frac{x_{nd}-x_{n'd}}{l_d}^2\right)\right) \quad (14)$$

where $\mathbf{x_n}$ and $\mathbf{x_{n'}}$ are explanatory variables vectors of the n-th and n'-th data, respectively. x_{nd} represents the d-th explanatory variable vectors of the n-th data. l_d and D represent characteristic length scale associated with d-th component in explanatory variables and dimension of explanatory variables vectors, respectively. Increasing the characteristic length scale, l_d, decreases the contribution of the d-th component in the explanatory variables. Therefore, we analyze which features extracted from plethysmograms are valid for inferencing low-level mental states.

4 EXPERIMENTAL RESULTS

We use GPR and SVR to make two regression models. Both use a Gaussian kernel as the kernel function. The regression deals with analog values of y as low-level mental states. Our regression evaluations employ k-fold cross validation($k = 5$). We combine the data from all subjects because some subjects exhibited many responses in their evaluation scores, while other subjects demonstrate scant responses.

4.1 Regression with GPR and SVR

Figures 8 and Figures 9 show the regression results yielded by GPR and SVR, respectively. Figures 10 and Figures 11 plot the regression coefficients for GPR and SVR, respectively. For GPR we show the confidence intervals based on standard deviation. Both regression models have only small differences between real and estimated values. As their regression coefficients are about 0.98, both methods are valid for inferencing low-level mental states. They use the kernel function based approach and their criterion contains a constraint for avoiding over-fitting. As the data set used did not have fine variations, SVR as a sparse regression is better than GPR. But, GPR can provide confidence scores for regression at each data point because of its Bayesian approach. The confidence scores are useful for real applications. As it is basically difficult to collect subjective evaluation data in real services, deep learning models are not suitable as they need big data sets of subjective evaluations to create adequate training for accurate parameter determination. The standard deviations of each point in $\hat{\mathbf{y}}$ yielded by GPR represent estimation variance. Therefore, to realize more accurate regression, we should use criteria that select high-quality data sets that yield strong correspondence between input data (plethysmogram feature) and output data (subjective evaluation).Data sets that contain big fluctuations in features extracted from plethysmograms or subjective evaluations will negatively influence the regression analyses by either GPR or SVR.

4.2 Feature Selection for Inferencing Low-level Mental States

Figure 12 shows feature selection by using ARD with GPR. It shows that Feature Type3 (frequency spectrum) and Type4 (chaos attractor) yield valid inferencing of low-level mental states. We consider that these features are difficult to extract with accuracy because Feature Type1 (second derivative of plethysmogram) and Type2 (heart rate variability) use information of second derivative of plethysmogram and so suffer from numerical error in detecting local maximum or minimum and calculating the derivative. In particular, Feature Type1 is not valid as each person has a different level of vascular flexibility. On the other hand, Feature Type3 (frequency spectrum) and Type4 offer stable feature extraction as they offer rich information in time windows.

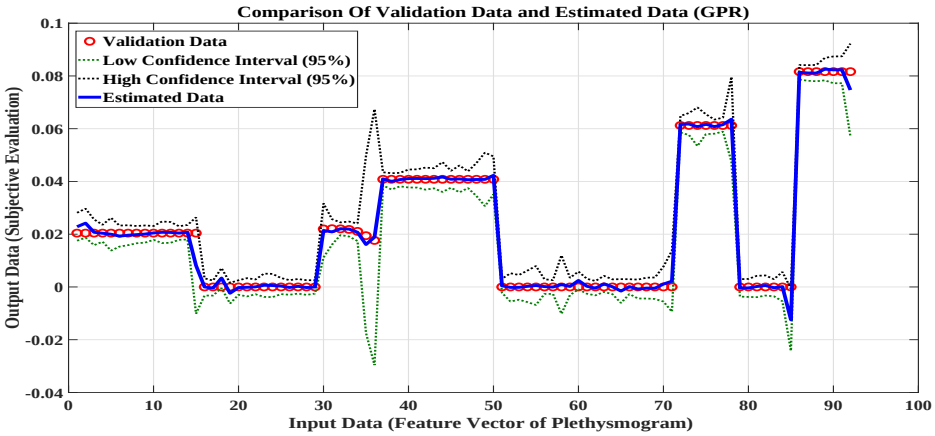

Figure 8: Regression by GPR.

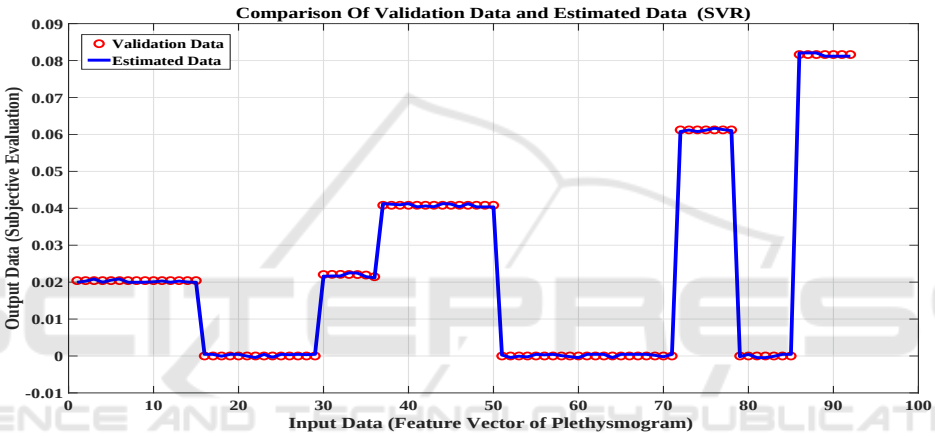

Figure 9: Regression by SVR.

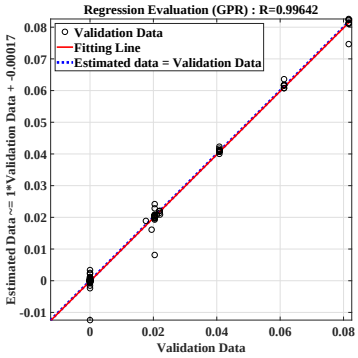

Figure 10: Evaluation of Regression by GPR.

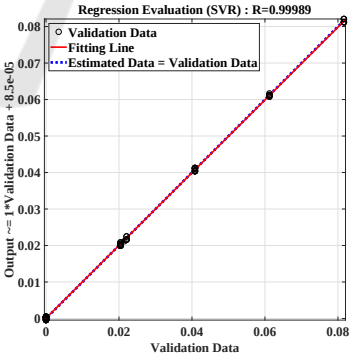

Figure 11: Evaluation of Regression by SVR.

5 CONCLUSION

Our experimental evaluation showed that both GPR and SVR offer good regression analyses of plethysmogram features and subjective evaluations; according to cross validation, the regression coefficients are about 0.98. As a result, we reveal the potential to detect low-level mental states by Earlobe Plethysmogram sensors mounted on wearable devices. Moreover, feature selection by ARD with GPR showed that Feature Type3 (frequency spectrum) and Type4 (chaos attractor) offer valid inferencing of low-level mental states. In future work, we will improve regres-

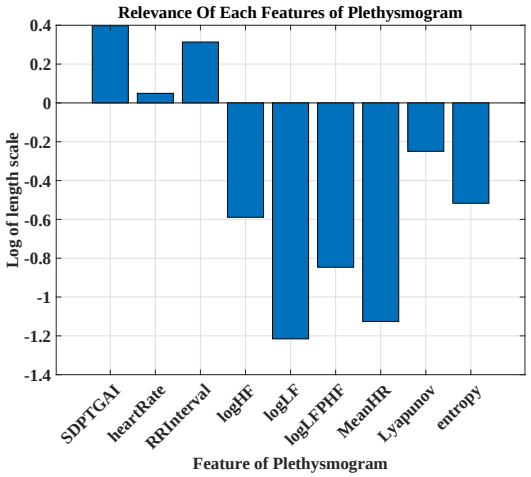

Figure 12: Feature Selection by using ARD with the GPR.

sion accuracy by selecting good quality data sets based on standard deviation from GPR or finding other criteria for discriminating low-level mental states.

ACKNOWLEDGEMENTS

The author would like to thank Mr. Gesshi Higashida, TAOS Institute, Inc. R & D Department for his technical support in the experiments.

REFERENCES

Alafeef, M. (2017). Smartphone-based photoplethysmographic imaging for heart rate monitoring. *Journal of Medical Engineering & Technology*, 41(5):387–395. PMID: 28300460.

Allen, J. (2007). Photoplethysmography and it application in clinical physiological measurement. *Physiological measurement*, 28:R1–39.

Cho, Y., Julier, S. J., and Bianchi-Berthouze, N. (2019). Instant stress: Detection of perceived mental stress through smartphone photoplethysmography and thermal imaging. *JMIR Ment Health*, 6(4):e10140.

Kimura, S., Fukuomoto, M., and Horikoshi, T. (2013). Eyeglass-based hands-free videophone. *Proceedings of the 2013 International Symposium on Wearable Computers*, pages 117–124.

Loube, D. I., Andrada, T., and Howard, R. S. (1999). Accuracy of respiratory inductive plethysmography for the diagnosis of upper airway resistance syndrome. *Chest*, 115(5):1333 – 1337.

Mohamed, E. (2012). On the analysis of fingertip photoplethysmogram signals. *Current cardiology reviews vol. 8,1 14-25*.

Murata, K., Fujita, E., Kojima, S., Maeda, S., Ogura, Y., Kamei, T., Tsuji, T., Kaneko, S., Yoshizumi, M., and Suzuki, N. (2011). Noninvasive biological sensor system for detection of drunk driving. *IEEE Transactions on Information Technology in Biomedicine*, 15(1):19–25.

Oyama-Higa, M. and Miao, T. (2005). Representation of a physio-psychological index through constellation graphs. In *ICNC (1)'05*, pages 811–817.

Oyama-Higa, M., Miao, T., and Mizuno-Matsumoto, Y. (2006). Analysis of dementia in aged subjects through chaos analysis of fingertip pulse waves. In *2006 IEEE International Conference on Systems, Man and Cybernetics*, volume 4, pages 2863–2867.

Rasmussen, B. C. E. and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. The MIT Press.

Romem, A., Romem, A., Koldobskiy, D., and Schar (2014). Diagnosis of obstructive sleep apnea using pulse oximeter derived photoplethysmographic signals. *Journal of Clinical Sleep Medicine*, 10(3):285–290.

Sano, M. and Sawada, Y. (1985). Measurement of the lyapunov spectrum from a chaotic time series. *Phys. Rev. Lett.*, 55:1082–1085.

Shaffer, F. and Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health*, 5:258.

Smola, A. J. and Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222.

Sumida, T., Arimitu, Y., Tahara, T., and Iwanaga, H. (2000). Mental conditions reflected by the chaos of pulsation in capillary vessels. *International Journal of Bifurcation and Chaos*, 10(09):2245–2255.

Tsuda, I., Tahara, T., and Iwanaga, H. (1992). Chaotic pulsation in human capillary vessels and its dependence on mental and physical conditions. *International Journal of Bifurcation and Chaos*, 02(02):313–324.

Vicente, J., Laguna, P., Bartra, A., and Bailón, R. (2016). Drowsiness detection using heart rate variability. *Medical & Biological Engineering & Computing*, 54(6):927–937.

Wipf, D. P. and Nagarajan, S. S. (2008). A new view of automatic relevance determination. In Platt, J. C., Koller, D., Singer, Y., and Roweis, S. T., editors, *Advances in Neural Information Processing Systems 20*, pages 1625–1632. Curran Associates, Inc.

Wong, W., Bartels, M., and Chrobot, N. (2014). Practical eye tracking of the ecommerce website user experience. In Stephanidis, C. and Antona, M., editors, *Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice*, pages 109–118, Cham. Springer International Publishing.