IJEE Volume-5 • Number-2 December 2013 pp. 6-12 ISSN: 0973-7383

Correlation between Physiological Parameters of Automobile Drivers and Traffic Conditions

Mandeep Singh, Abdullah Bin Queyam
Department of Electrical & Instrumentation Engineering, Thapar University, Patiala, INDIA
mandy_tiet@yahoo.com, abdullahbinqueyam@gmail.com

Abstract
Driving a car is a complex cognitive process in which even a small lack of attention can have disastrous consequences. Various studies have been conducted in the past focusing mainly on the driver's internal state (physical and emotional) like drowsiness, fatigue and mental stress as these are the major cause behind large number of fatal road accidents worldwide. The studies show that in most of the automobile driver's Galvanic Skin Response (GSR) and Heart Rate (HR) parameters are more closely related to drivers stress level. The authors accessed raw physiological signals available at PHYSIONET website and then extract useful statistical features from these. The most relevant features were then selected using open source software. Removal of irrelevant features makes the stress detection model much faster and helps to gain a deeper insight into the underlying processes of stress detection and classification. Correlation analysis on the selected features showed that Mean HR and Mean Hand GSR are the two statistical features that have a very strong correlation with changing traffic conditions.
Keywords: Automobile driver, stress, correlation, physiological parameters, features extraction, traffic conditions.

1. Introduction

Stress can be recognized by the physiological information of human being. One of the most relevant studies in this field was conducted by Healey and Picard, in which they presented methods for collecting and analysing physiological data during real world driving tasks. They continuously recorded ElectroCardioGram (ECG), ElectroMyoGram (EMG), Skin Conductance (SC) also known as Galvanic Skin Response (GSR) and Respiration Rate (RR) signals of drivers. They showed that physiological measurements can predict mental stress with high accuracy. Healey and Picard's work, however, lacked the procedure of feature selection, which may result in higher accuracy and performance in real time stress recognition. The database created by Healey and Picard mentions the drive time through highway and through city conditions, along with the period during which the driver is resting [1].

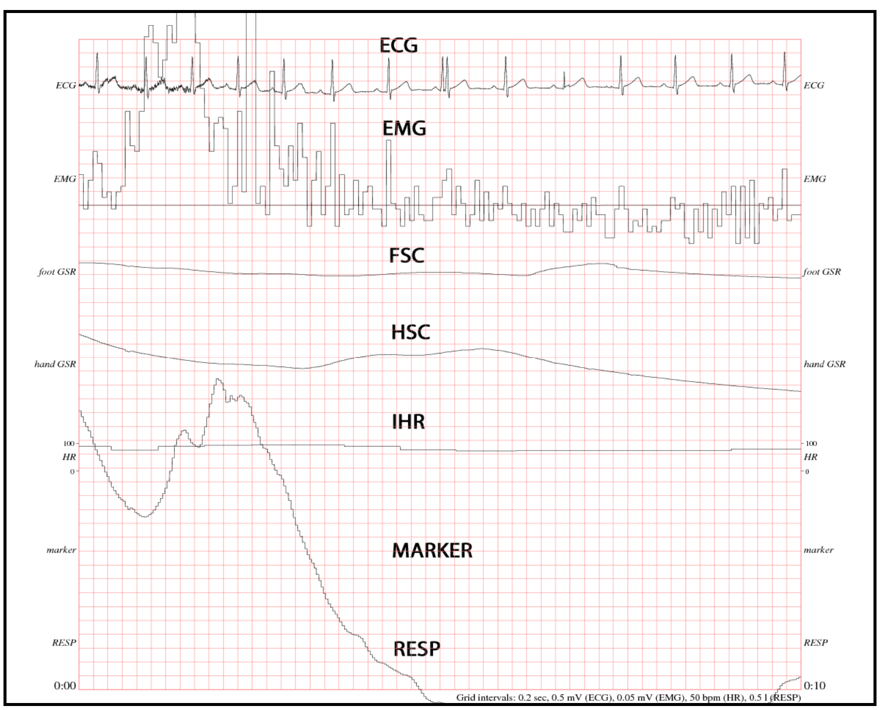

Fig. 1. A 10-s close up of original 'drive05' bio-signal dataset.

One simple approach in classification of stress level is to consider the rest period as low stress condition, highway driving as moderate stress condition and city drive as high stress condition. Neglecting all other factor towards the contribution of stress, an analysis of correlation of various physiological parameters with the three mentioned states of mental stress has been made in this paper.

Ahmet Akbas concluded that by effectively measuring the physiological signals like ECG, EMG, SC and RR the stress level of the automobile drivers can be evaluated. A 10 sec of database is shown in Fig.1. He also stated that mean of the Instantaneous Heart Rate (IHR), Hand based Skin Conductance (HSC), Foot based Skin Conductance (FSC), EMG signals, Instantaneous Respiratory Rate (IRR) and average number of Contractions/Minutes (CPM) metrics related to the measurements can be used to indicate dynamic stress level of drivers [2].

2. Database Collection

A lot of work has already been done, but only a few physiological sensor data sets which can be used for detection of stress have been made available for further exploration. Fortunately, there is one available from the PHYSIONET website [3] which contains sensors signal data directly obtained from different drivers. The driver data set is originally collected by **Healey & Picard** from MIT Media Lab. It provides sufficient data for us to do some further work about the feature selection and stress identification [1].

PHYSIONET is an on-line forum for the dissemination and exchange of recorded biomedical signals and open-source software for analyzing them. It provides facilities for the cooperative analysis of data and the evaluation of proposed new algorithms. It also provides free electronic access to PhysioBank data and PhysioToolkit software via the World Wide Web (http://www.physionet.org).

We have obtained all our raw physiological signals from the database of PHYSIONET website located at http://www.physionet.org/physiobank/database/drivedb/. This database, contributed to PHYSIONET by its creator, **Jennifer Healey**, contains a collection of multi-parameter recordings from healthy volunteers, taken while they were driving on a prescribed route including city streets and highways in and around Boston, Massachusetts. However these data sets actually are not as complete as those in the experiment of **Healey and Picard** [1]. In their original work, a total of seventeen drivers participated in the experiment and for each driver eight types (Time Stamp, ECG, EMG, Foot GSR, Hand GSR, HR, Marker and Respiration Rate) of raw data are acquired from the sensors that the drivers wear. On carefully inspecting the data, it was found that amongst the seventeen data sets, only ten drives data sets (drive 05, 06, 07, 08, 09, 10, 11, 12, 15 and 16) are complete, which includes all the sensor information as well as have clear mark identification. The remaining seven drive data sets (drivers 01, 02, 03, 04, 13, 14, 17a and 17b) do not contain all the sensor information and the mark of different driving period is not clear. The work conducted by **Yong Deng** *et al.* [4] also helped us to clearly understand the selection of different drivers' data sets.

We have downloaded our entire database from PHYSIONET Website in the form of .csv files which is in the form of comma-separated values and can be directly opened in MS Excel Software [5]. But one major problem that we encountered while opening our downloaded .csv files is that they are very long columns of data and not within the permissible range of MS Excel which is 65,536 rows by 256 columns. To overcome this problem, two software's viz; CSV Splitter v1.1 by Sopheap Ly and Merge Excel Sheets v 4.0 © 2010 by Abdel Yezza were used for opening these .csv files. We divided single .csv files into multiple file of one minute duration each for the whole one hour long drive using CSV Splitter and then merged them column-wise using Merge Excel Sheets. By this method we were able to download and save all ten different drivers data set in MS Excel.

Each signal acquired from PHYSIONET was sampled at a rate appropriate for capturing the information contained in the signal constrained by the sampling rates available on the FlexComp system. The ECG was sampled at 496 Hz, the skin conductivity (SC) and RESP signals were sampled at 31 Hz, and the EMG was sampled at 15.5 Hz after first passing through a 0.5 s averaging filter. The signals were collected by embedded computer in the testing car [1].

The data acquired from **Healey and Picard's** [1] experiment lacks the information regarding the duration of each Rest, City and Highway driving task, but the same durations were mentioned in **Ahmet Akbas** [2] research and we found it to be very useful in our research for stress detection in automobile drivers. List of the time intervals for different driving segment for each of the 10 selected Drive's i.e. Drive 05, 06, 07, 08, 09, 10, 11, 12, 15 and 16 are shown in Table 1.

Table 1
Time intervals of the 7 driving segments of available bio-signal datasets.

Drive No.	Driving period (min)							Total rec. time (min)
	Initial Rest	City 1	Highway 1	City 2	Highway 2	City 3	Final Rest	
Drive05	15.13	16	7.74	6.06	7.56	14.96	15.78	83.23
Drive06	15.05	14.49	7.32	6.53	7.64	12.29	15.05	78.38
Drive07	15.04	16.23	10.96	9.83	7.64	10.15	15.03	84.87
Drive08	15	12.31	7.23	9.51	7.64	13.43	15.07	80.19
Drive09	15.66	19.21	8.47	5.2	7.06	13.21	NA	68.82
Drive10	15.04	15.3	8.66	5.27	7.04	12.06	14.79	78.15
Drive11	15.02	15.81	7.43	7.15	6.96	11.72	14.99	79.08
Drive12	15.01	13.41	7.56	6.5	8.06	11.68	15.01	77.23
Drive15	15	12.54	7.24	5.99	6.82	12.12	15	74.7
Drive16	15.01	16.12	7.14	5.12	6.81	13.91	NA	64.1

NA: Not available

3. Feature Extraction

In our previous work we have done a literature survey on the topic of stress detection and selected four physiological parameters like ECG, EMG, GSR and RR that will help in detecting stress of automobile driver's easily. These four signals were selected based on their properties regarding non-invasivity when being acquired and because their variation is strongly related to stress stimuli [6]. Overall 8 statistical features were extracted from the 4 available physiological parameters.

Heart rate variability (HRV) analysis is commonly used as a quantitative marker depicting the activity of autonomic nervous system (ANS) that may be related to mental stress [7]. Although HRV features can be extracted by detecting QRS complexes from electrocardiogram (ECG) signals, the signals acquired from PHYSIONET website already contained HR signal along with ECG signal. So, we have used available HR signal directly to extract features like Mean HR, HRV and normalised HRV (NHRV).

GSR is a bio-electric physiological signal controlled by the sympathetic nervous system. The GSR signal comprises of two components: tonic (Frequency Range: 0.0Hz to 0.16Hz) and phasic (Frequency Range: 0.16Hz and above) [8]. Features extracted from GSR can give significant emotion information about an individual like surprise, fear, disgust, grief, and happy [9]. In our work we have extracted Mean hand GSR and Mean foot GSR features. All mean values are taken over a period of 60 seconds. Thus minutes to minute mean values are used for analysis.

The stress changes amplitude and temporal features of the EMG of the upper trapezius muscle. A statistically significant relation was found between stress level and features extracted from EMG signals of the upper trapezius muscle. Amplitude of the EMG signal is reported to be significantly higher during stress conditions than during rest [10], [11]. We used Mean EMG and rms EMG features, again on minute to minute basis.

Respiration is primarily regulated for metabolic and homeostatic purposes in the brainstem. However, breathing can also change in response to changes in emotions, such as sadness, happiness, anxiety or fear [12]. We used Respiration Rate per minute (RR) as a feature for stress detection. Table 2 gathers a list of the physiological signals and the corresponding statistical features extracted from them. These extracted features can be categorized in the following groups:

Table 2
List of the physiological signals and the corresponding features extracted from them.

Physiological Signals	Features Selected
GSR (Galvanic Skin Response)	Mean hand GSR, Mean foot GSR
EMG (Electromyogram)	Mean EMG, rms EMG
Respiration Rate (RR)	Respiration Rate per minute (RR)
ECG (Electrocardiogram)	Mean Heart Rate (HR), Heart Rate Variability (HRV), Normalised HRV (NHRV)

We have calculated 8 statistical features for every minute of the driving period for each of the 10 selected driver's data set i.e. drive 5, 6, 7, 8, 9, 10, 11, 12, 15 and 16 as shown in Table 3. All these features are then accumulated to form a feature matrix and correlation is evaluated to check the strength of relationship between these features.

Table 3
List of 8 statistical features extracted and their description

S. No.	Features	Description
1	Mean hand GSR	The mean of the hand GSR data samples for each minute
2	Mean foot GSR	The mean of the foot GSR data samples for each minute
3	Mean EMG	The mean of the EMG data samples for each minute
4	rms EMG	The root-mean-square of mean EMG for each minute
5	RR	The mean of respiration rate for each minute
6	HR	The mean of the heart rate data samples for each minute
7	HRV	The standard deviation of heart rate data for each minute
8	NHRV	The normalised standard deviation of heart rate data for each minute

4. Evaluating Correlation Matrix

The cross correlation function is a measure of the similarity between two data sets. Corresponding values of the two sets are multiplied together, and the products are summed to give the value of cross correlation without any phase lag. It is a standard method of estimating the degree to which two different series or signals are correlated. If the signals are identical, then the correlation coefficient is 1; if they are totally different, the correlation coefficient is 0, and if they are identical except that the phase is shifted by exactly 180^{o} (i.e. mirrored), then the correlation coefficient is -1. Strength of correlation between the features are evaluated based on the range shown in Table 4.

Table 4
Interpretation of the Strength of correlation results.

Correlation Coefficient Range	Strength of Correlation
|0.00 to 0.30|	Weak
|0.31 to 0.50|	Moderate
|0.51 to 0.80|	Strong
|0.81 to 1.00|	Very Strong

Correlation coefficient is also known as Pearson product-moment correlation coefficient, or "Pearson's correlation". If we have a series of n measurements of X and Y written as x_i and y_i where i = 1,2,...,n, then the sample correlation coefficient can be used to estimate the population Pearson correlation r_{xy} between X and Y. The sample correlation coefficient r_{xy} is written as:

$$r_{xy} = \frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{(n-1)s_x s_y} = \frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2 \sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2}}$$

Where x and y are the sample means of X and Y, and s_x and s_y are the sample standard deviations of X and Y.

From the duration of different driving segment obtained from Table 1 we have defined a stress class for our database so that the correlation analysis of the database can be done properly. Here we have considered the rest period as low stress condition, highway driving as moderate stress condition and city drive as high stress condition. Neglecting all other factor towards the contribution of stress, an analysis of correlation of various physiological parameters with the three mentioned states of mental stress has been made in this paper.

e.g. The states changes from moderate to high stress where the first three minutes are in moderate stress and the subsequent three minutes in high stress condition. Instead of making this change over six minutes as 3,3,3,5,5,5 the transition is recorded as 3,3,3,4,5,5.

We have calculated the correlation between all the 8 selected statistical features with respect to different stress class values as stated in Case 1, Case 2 and Case 3. These category and stress class are shown in Table 5. Case 1 and Case 2 shows abrupt change in stress level while Case 3 shows a gradual change in stress level of driver with respect to change driving segment.

Table 5
Driving Segments divided into different Category and Class

Driving Segments Used	Category	Stress Class		
		Case 1	Case 2	Case 3
Initial Rest + Final Rest	Relaxed or Low Stress (LS)	**1**	**1**	**1**
Transition State	-	**-**	**-**	**3**
City 1 + City 2 + City 3	High Stress (HS)	**10**	**5**	**5**
Transition State	-	**-**	**-**	**4**
Highway 1 + Highway 2	Medium Stress (MS)	**5**	**3**	**3**

We evaluated Pearson's coefficient *r* for each drive's dataset. These *r* values show the strength of relationship between the different features and the stress class value defined in this research. Correlation matrix obtained for Drive07 for three different cases as shown in Table 5, along with its corresponding *r* values are shown in Table 6.

Table 6
Correlation matrix obtained for Drive07 and its corresponding *r* values including all three cases Case 1, Case 2 and Case 3

Drive07		Mean Hand GSR	Mean Foot GSR	Mean EMG	rms EMG	RR	Mean HR	HRV	NHRV	Stress Class
Mean Hand GSR		1	0.821	0.677	0.560	0.783	0.673	0.066	-0.085	0.614
Mean Foot GSR		0.821	1	0.523	0.478	0.603	0.353	-0.055	-0.149	0.417
Mean EMG		0.677	0.523	1	0.810	0.685	0.656	0.114	-0.012	0.500
rms EMG		0.560	0.478	0.810	1	0.561	0.449	0.081	-0.010	0.433
RR		0.783	0.603	0.685	0.561	1	0.623	-0.166	-0.306	0.552
Mean HR		0.673	0.353	0.656	0.449	0.623	1	0.296	0.093	0.346
HRV		0.066	-0.055	0.114	0.081	-0.166	0.296	1	0.977	-0.108
NHRV		-0.085	-0.149	-0.012	-0.010	-0.306	0.093	0.977	1	-0.182
Stress Class	**Case 1**	0.614	0.417	0.500	0.433	0.552	0.346	-0.108	-0.182	1
Stress Class	**Case 2**	0.622	0.432	0.501	0.437	0.571	0.343	-0.125	-0.201	1
Stress Class	**Case 3**	0.667	0.482	0.536	0.509	0.617	0.387	-0.078	-0.163	1

We have evaluated Correlation Matrix for all ten driver's data set and found that in most of the cases the Stress Class is strongly correlated with Mean Hand GSR. The correlation results of Stress Class with Mean EMG, rms EMG, RR, Mean HR also shows moderate to strong correlation but not as strong as Mean Hand GSR. On inspecting last three rows of Table 6, it is clear that Case 3 has stronger correlation of almost all physiological parameters with stress class. This makes Case 3 for stress class a probable candidate to be verified for the remaining drives as well. Table 7 summarises all the correlation coefficient values obtained between Stress Class and Mean Hand GSR in Case 1, Case 2, and Case 3 for all the 10 drives.

Table 7
Comparison of correlation coefficient obtained between Stress Class and Mean Hand GSR in Case 1, Case 2, and Case 3 for all 10 drives.

S No.	Drive No.	Correlation Coefficient between Stress Class and Mean Hand GSR		
		Case 1 (1-5-10)	Case 2 (1-3-5)	Case 3 (1~3~5)
1	Drive05	0.762	0.770	0.802
2	Drive06	0.291	0.263	0.257
3	Drive07	0.614	0.622	0.667
4	Drive08	0.745	0.753	0.772
5	Drive09	0.848	0.857	0.861
6	Drive10	0.329	0.323	0.357
7	Drive11	0.860	0.869	0.881
8	Drive12	0.570	0.586	0.610
9	Drive15	0.729	0.737	0.738
10	Drive16	0.793	0.803	0.804

The reason behind this very strong correlation between Stress and Mean Hand GSR is that the Autonomic Nervous System (ANS) of the human body get activated when people experience any stressful situation like driving in high traffic condition. In case of physical arousal sweat is produced in the eccrine glands, which measurably changes the conductivity of the skin. The sweat glands used for measurement are typically those in the palms of the hand or the soles of the feet. Thus, GSR is the most important indicator of stress. Also, GSR is one of the most straight-forward and low-cost psycho-physiological measures. Further, choice of Case 3 is most suitable way to classify the level of stress and hence will be used in the subsequent part of our study.

5. Results

We have divided the stress class in 3 different ways: Case 1, Case 2 and Case 3 based on the change in traffic condition. Case 1 and Case 2 includes step change in traffic condition while Case 3 includes gradual change in traffic condition. 9 drive's out of total 10 drive's, shows that the correlation results obtained in Case 3 was always greater than Case 1 and Case 2 as shown in Fig. 2. This is in contrast that when a driver goes from high traffic condition to low traffic condition or vice versa, there is a gradual change in stress level along with the physiological measures of the automobile drivers.

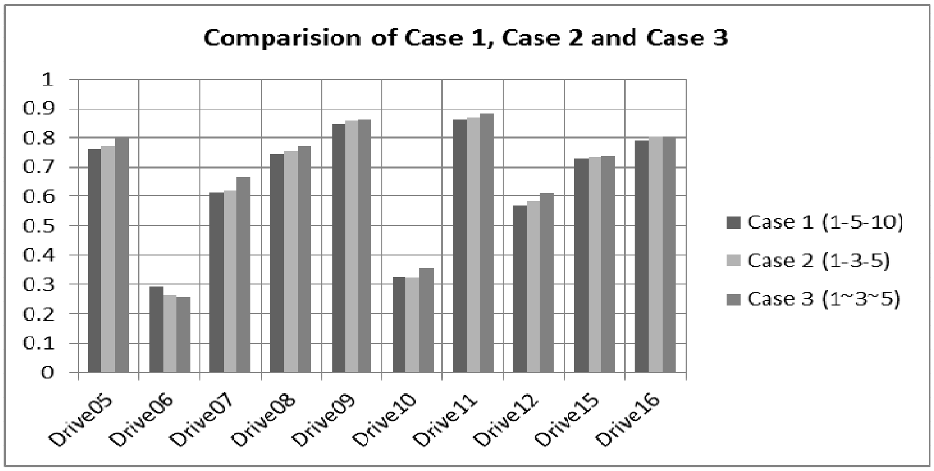

Fig. 2. Comparision of Case 1, Case 2 and Case 3 for choice of stress classification.

6. Conclusion and Future Work

Total 10 drivers' bio-signal datasets obtained from PHYSIONET website [3] were used to find correlation between stress class and statistical features obtained from physiological parameters. Out of these 10 drives, Drive06 and Drive10 database shows very weak correlation with stress may be due to some unavoidable problems related to sensors, results are shown in Fig. 2. Assuming that the psychological stress in driver is only on account of the traffic conditions which again depend on the terrain, correlation analysis shows very strong linear relationship between Stress and Mean Hand GSR. Stress class as indicated by Case 3 shows the best correlation result, thus we select Case 3 as our final stress class for further analysis. In future we want to build a function which will include a combination of different physiological features. This function will act as a direct indicator of stress level of automobile drivers. Further this function will try to improve correlation in case of problematic Drive06 and Drive10 as well. The aim of this function will be to design a direct reading stress level meter.

References

[1] J. A. Healey and R.W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, pp. 156-166, June. 2005.

[2] A. Akbas, "Evaluation of the Physiological Data Indicating the Dynamic Stress Level of Drivers," *Scientific Research and Essays*, vol.6(2): pp. 430-439, 18 Jan. 2011.

[3] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101(23):e215-e220 [Circulation Electronic Pages; http://circ.ahajournals.org/cgi/content/full/101/23/e215]; 2000 (June 13). PMID: 10851218; doi: 10.1161/01.CIR.101.23.e215
Available at: http://physionet.org/cgi-bin/atm/ATM/

[4] Y. Deng, D. F. Hsu, Z. Wu and C. H. Chu, "Combining Multiple Sensor Features for stress Detection using Combinatorial Fusion," *Journal of Interconnection Networks*, vol. 13, pp. 3-4, 2012.

[5] Microsoft. (2010). Microsoft Excel [computer software]. Redmond, Washington: Microsoft.

[6] Mandeep Singh and Abdullah Bin Queyam, "Stress Detection in Automobile Drivers using Physiological Parameters: A Review," *International Journal of Electronics Engineering*, vol. 5, no. 2, Dec 2013.

[7] L. Salahuddin and D. Kim, "Detection of Acute Stress by Heart Rate Variability Using a Prototype Mobile ECG Sensor," *International Conference on Hybrid Information Technology*, vol. 2; pp. 453-459, 2006.

[8] R.R. Singh, S. Conjeti and R. Banerjee, "Biosignal based on-road stress monitoring for automotive drivers," *Communications (NCC), 2012 National Conference on*, vol., no., pp. 1-5, 3-5 Feb. 2012.

[9] G. Wu, G. Liu and M. Hao, "The Analysis of Emotion Recognition from GSR Based on PSO," *Intelligence Information Processing and Trusted Computing (IPTC), 2010 International Symposium on* , vol., no., pp. 360-363, 28-29 Oct. 2010.

[10] J. Wijsman, B. Grundlehner, J. Penders, and H. Hermens, "Trapezius muscle EMG as predictor of mental stress," *Wireless Health* 2010 (WH '10). ACM, New York, NY, USA, pp. 155-163, 2010.

[11] J. A. Healey, "Wearable and automotive systems for affect recognition from physiology", MIT Dept. of Electrical Engineering and Computer Science, Ph.D. thesis, 2000.

[12] I. Homma and Y. Masaoka, "Breathing rhythms and emotions," *Experimental Physiology*, 93 (9), pp. 1011-1021, September 1, 2008.