Realization of Stress Detection using Psychophysiological Signals for Improvement of Human-Computer Interactions

JING ZHAI[1], ARMANDO B. BARRETO [1,2], CRAIG CHIN[1], CHAO LI[1]
[1]Electrical and Computer Engineering Department, [2]Biomedical Engineering Department
Florida International University
{jzhai002, barretoa, cchin006, cli006}@fiu.edu

Abstract

It has been suggested that effectively detecting the stress level of a computer user could possibly develop the computers' ability to respond intelligently and help the user relax from negative emotional states during human-computer interaction. Our research focuses on the use of three physiological signals: Blood Volume Pulse (BVP), Galvanic Skin Response (GSR) and Pupil Diameter (PD), to automatically monitor the stress state of computer users. This paper reports on the hardware and software instrumentation development and signal processing approach used to detect changes in the stress level of a subject interacting with a computer, within the framework of a specific experimental task. For this experiment a computer game was implemented on the basis of a clinical mental stress test, called the 'Stroop Test', adapted to make the subject experience two different levels of stress, while his/her BVP, GSR and PD signals were continuously recorded. Several data processing techniques were applied to extract effective attributes of the 'stress' state of the subjects. Current results indicate that there exists a strong correlation among changes in those three signals and the shift in the emotional states when stress stimuli are applied to the interaction environment.

1. Introduction

Computers are used during an incredible amount of work and leisure time of human beings. Accordingly, ideal human-computer interactions should involve a two-way exchange, with each participant aware of the other and responding appropriately. Numerous studies have focused on ways to make human-computer interaction more intelligent and natural. Many scientific findings indicate that emotions play an essential role in rational decision-making, perception, learning and various cognitive tasks[1]. Therefore, making machines aware of the emotional state of the user would yield a more natural human-machine interaction. Our research attempts to visualize and evaluate the emotional state identified with 'stress' of the computer users, through several physiological signals that can be measured non-invasively and non-intrusively, and its preliminary results may be projected to support the emerging field of Affective Computing.

2. Physiological aspects of stress detection

In the past, researchers have investigated many aspects of the responses of the human body to emotions. Many attempts, including the identification of facial expressions and the combination with speech understanding techniques, have been made to assess the affective state of the user, towards enriching the communication between humans and computers. In the future, human-computer interactions should become more inherently natural and social [2]. Physiological signals, which are not voluntarily manipulated ('faked') by the computer users, can be seen as a truly reliable indicator of the emotional state of the subject.

Many publications in the psychophysiological literature, e.g. [3, 4], indicate that changes in the emotional state of a subject are accompanied by many bodily reactions, such as modifications in the heat rate, blood pressure, blood volume, electrical properties of the skin, brain waves, temperature and pupil size. Most previous reports from the literature on Affective Computing use Blood Volume Pulse (BVP), Galvanic Skin Response (GSR) and Electrocardiography (EKG) to monitor the state of the subject. Although, the variation of pupil diameter (PD) has been reported to have a potential relationship with the stress state during and after auditory emotional stimulations[5], the joint analysis of pupil diameter changes with other physiological signals has not been fully investigated. Our study sought to coalesce the information derived from three physiological signals: GSR, BVP and PD, to detect the stress state of the users when they are interacting with the computer.

From human physiology studies, it is well known that the sympathetic division of the human Autonomic Nervous System (ANS) significantly influences several physiological variables. The heart rate, skin resistance, blood volume and pupil diameter are all affected by branches of the sympathetic division of the ANS. We collected these three physiological variables (GSR, BVP, PD) simultaneously to analyze potential concurrent changes that may be due to sympathetic activation associated with a multi-faced emotional state – 'stress'.

A GSR2 module, by Thought Technology LTD (West Chazy, New York) was used to sense the Galvanic Skin Response (GSR) and photoplethysmography (PPG) was used to measure the blood volume in the skin capillary bed, in the finger. The sampling rate for GSR and BVP recording was set to 360 samples/sec. To get an accurate and continuous pupil diameter signal, we used the ASL-504 eye gaze tracking system running at a different sampling rate of 60 samples/sec. The detailed description of these sensors and the signal collection and synchronization methods can be found in our report of the instrumentation designed for this study[6].

3. Experiment design

3.1. Stress elicitation

One of the most challenging points in this research is to obtain accurate physiological signals related to mental stress of a human subject. Design and implementation of experiments simulating mental stress require considerable insight into human psychology. In the context of human-computer interaction, the stress experienced by the user is most likely to be mental (as opposed to physical), and moderate in intensity. Therefore, our experimental protocol sought to instill moderate mental stress in the participating subjects, at pre-determined times. Accordingly, a computer game based on the well-known 'Stroop Test' was designed and adapted to elicit the mental stress while the subject was interacting with the computer.

The Stroop Color-Word Interference Test[7], in its classical version, demands that the color of a word designating a different color be named. Although there is controversy concerning the exact mechanisms responsible for the Stroop effect, this task has been widely utilized as a psychological or cognitive stressor to induce emotional responses and heightened levels of physiological, (especially autonomic) reactivity[8]. In our research, the interacting environment should be established to let the subjects experience a similar effect. To accomplish this, the classical Stroop Test was adapted into an interactive version that requires the subject to click the correct answer, which is one of the five buttons shown on the screen, rather than stating it verbally. One typical example of this test interface is shown on Figure 1. This modified version was implemented with Macromedia Flash® and also programmed to output bursts of a sinusoidal tone through the sound system of the laptop used for stimulation, at selected timing landmarks through the protocol, to time-stamp the recorded signals at those critical instants. Our previous report on the instrumental setup[6] provides more details on the audio schemes to achieve the desired time-stamping in the three recorded signals. Figure 2 is the audio output schedule in this experiment from the beginning of the game to its end. The complete experiment comprises three consecutive sections. In each section, we have 1) 'IS' - the Introductory Section to let subject get used to the game environment, in order to establish an appropriate initial level in the psychological experiment, according to the law of initial values (LIV)[9]. 2) 'C' – is a Congruent segment, in which the font color and the meaning of the words presented to the user match. 3) 'IC' – is an Incongruent segment of the Stroop Test in which the font color and the meaning of the words presented differ. 4) 'RS' – is a Resting Section to let the subject relax for a certain time. The binary number shown in Figure 2 is the de-multiplexed output of the audio signaling used in the system to time-stamp the three physiological signals, BVP, GSR and PD. '01' represents a burst in the left channel audio signal, '10' represents a burst in the right channel and '11' represents simultaneous bursts in both channels.

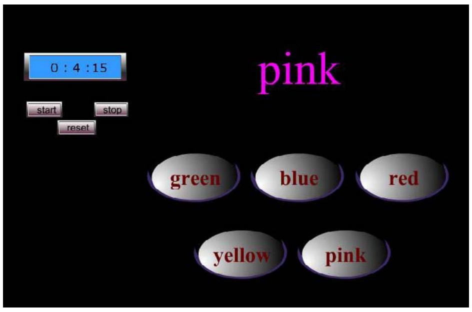

Figure1. **Sample Stroop Test interface**

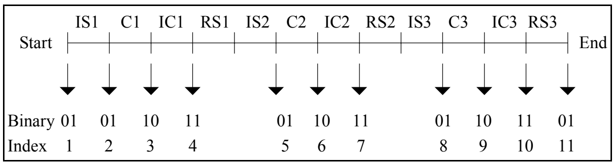

Figure2. **Audio output schedule**

3.2. Instrumentation Hardware

Six healthy male subjects (ages 21 to 35) participated in this study. They were recruited from the student body of our FIU College of engineering. To assure reliability of the changes of PD, the lighting of the environment was kept at a constant level during the experiment and the illumination of the system was the same for all the subjects. The complete instrumental setup used is illustrated in Figure 3. The stimulus program (interactive Stroop Test) described above ran in a laptop PC. While playing the Stroop Test, the subject had the GSR and BVP sensors attached to his/her left hand. Both GSR and BVP signals were converted, using a multi-channel data acquisition system, NI DAQPad-6020E for USB, a product of National Instrumentation Corp, to be read into Matlab® directly at rate 360 samples/sec. Additionally, the eye gaze tracking system (ASL-504) had been calibrated and recorded PD data to a file on its own interface PC, at rate of 60 samples/sec. The software for this system allows the extraction of selected variables (in this case the pupil diameter and the marker channel) to a smaller file, which in turn can be read into Matlab® also, where it can be aligned with the BVP and GSR signals, thanks to their common timing marks for the start and stop events. At this point the pupil diameter data can be upsampled (interpolated) by six, to achieve a common sampling rate of 360 samples/sec for all three measured signals.

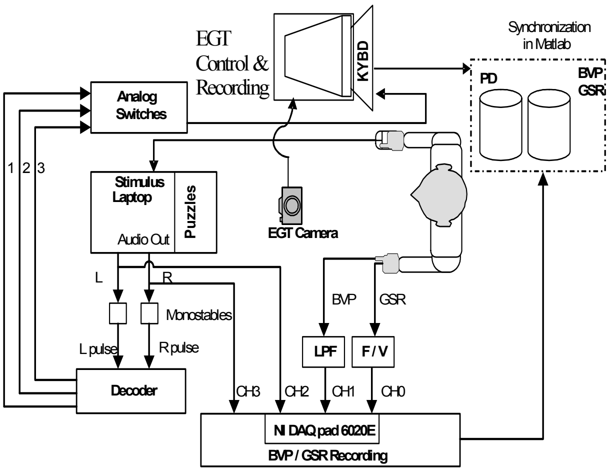

Figure3. **Instrumental Setup**

4. Feature extraction

From a set of raw data, with two audio signals as the markers, a set of significant feature signals that might have correlates with the internal stress state in each game section must be obtained. This is one of the most important research problems that exist in this area. The mapping between affective states and physiological states is still a research question being investigated in the psychophysiology community[10]. Multiple parameters were extracted from each physiological measure, as shown on Figure4.

Most psychophysiological recordings can be analyzed in terms of three types of activities: spontaneous, tonic(background), and phasic (evoked responses). Spontaneous activity is in reality a change in physiological activity that occurs in the absence of any known stimulus. Tonic activity is often referred to as the background level or resting level of activity of a particular physiological measure. Phasic activity is a discrete response to a specific stimulus – an evoked response[9]. In our protocol, the tonic activity is measured by letting the subjects relax for at least 5 minutes before the stimulus game starts.

The analysis of the GSR signal for stress assessment involved the evaluation of spontaneous and phasic activities. Since our stimulus aimed to induce the stress effect by the contradiction between font color and meaning of the words presented, responses that happened in the congruent Stroop segments can be treated as spontaneous responses and those that happened in the incongruent Stroop segments can be considered phasic responses.

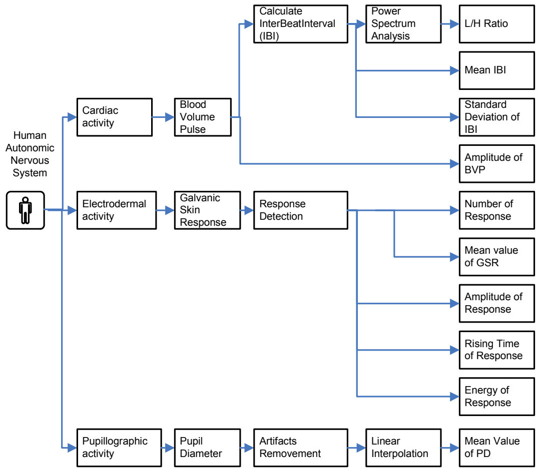

Figure4. **Physiological features extracted**

Both these responses were isolated by using Ktonas' 7-point Lagrangian interpolation algorithm[11] on the raw GSR signals. In his paper, he proposed the measurement of the second time derivative with a combination of the 3-point Lagrangian derivative and the 5-point second order data fit algorithm. Let g[n] represents the discrete time signals obtained by sampling the GSR signals. The first derivatives are computed using the 3-point Lagrangian interpolation approach:

$$g'[n] = \frac{g[n+1] - g[n-1]}{2h} \qquad (1)$$

where h is the sampling interval. Then, the second derivative is calculated with the 5-point second-order fit differentiation:

$$g''[n] = \frac{2g'[n+2]+g'[n+1]-g'[n-1]-2g'[n-2]}{10h} \qquad (2)$$

Substituting equation (1) into equation (2), yields:

$$g''[n] = \frac{2g[n+3]+g[n+2]-2g[n+1]-2g[n]-2g[n-1]+g[n-2]+2g[n-3]}{20h^2} \qquad (3)$$

Equation (3) can be considered as a 7-point operator. We multiplied $g''[n]$ by a constant value of 200 to keep the original signal scale. This operator includes the advantages provided by the use of interpolation and data-fit and makes the transient increase in the raw signal 'clearly distinguished from the background'. By thresholding, the responses in the GSR signal could be counted and localized. The number of the responses and the mean value in each section could then be calculated. A typical GSR response is shown on Figure 5. In our study, since each stimulus is triggered interactively by the subject, the appearance of a new stimulus may artificially shorten the half recovery time of the response to the previous stimulus. Therefore, only the amplitude and the rising time were recorded as features from each GSR response. Additionally, the total area under the rising time curve is treated as the GSR response energy.

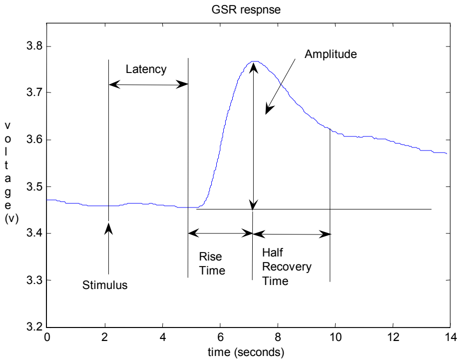

Figure5. **A typical GSR response**

Heart rate variability (HRV) has been extensively applied in understanding the function of ANS, and has shown a close connection to the emotional state of the subject[12, 13]. Clinically, HRV has been obtained through the electrocardiogram (EKG) signal, but in the context of human-computer interaction studies, photoplethysmographic (PPG) measurement of the Blood Volume Pulse (BVP) was preferred as a much less invasive and obtrusive monitoring method.

From the BVP signal, each heart beat was first separated and two parameters were generated, as shown in Figure 6. P is the BVP period, also called interbeat interval (IBI), defined as the time in milliseconds between two normal, consecutive peaks in the BVP signal. AM is the beat amplitude. The IBI is a very valuable index for measuring heart rate variability, which can be assessed either in the time domain or in the frequency domain. The time domain analysis has the limitation of needing very large sets of data to yield robust HRV estimates. For the short time duration data, the frequency domain analysis is more reliable than the time domain analysis. Since our research is focused on detecting stress during one computer task, the frequency domain analysis was preferred.

HRV was associated with three frequency bands: Very Low Frequency (VLF) (0.00-0.04Hz), Low Frequency (LF) (0.05-0.15Hz) and High Frequency (HF) (0.16-0.40Hz). The low frequency band reflects sympathetic activity with vagal modulation, and the high frequency band reflects parasympathetic activity. The very low frequency (VLF) domain was not analyzed in this study because VLF assessed from short-term recordings (≤ 5 min) has been shown to be an unreliable measure, according to the Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology[14].

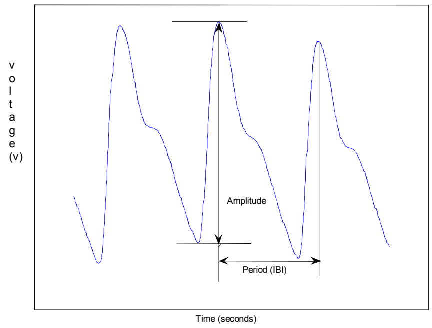

Figure6. **A typical BVP beat**

When a healthy person experiences mental stress, the sympathetic activity of his/her heart increases and the parasympathetic activity decreases, so the LF/HF ratio can be calculated as a single number estimate of mental stress[15]. In our study, we use this LF/HF ratio as one attribute to indicate the presence of stress when the subject is interacting with the computer. The mean value and standard deviation of the IBI sequence were also obtained as additional features of the cardiac response. Also, the amplitude of each BVP beat was calculated since its decrease may indicate that the subject is undergoing mental stress.

The raw pupil diameter (PD) signal was recorded separately, as previously described. The artifact gaps due to blinking have to be filled by interpolation. The feature extracted from the pupil diameter is a simple parameter: the mean value of PD. According to the introduction, we expect the mean PD should increase during the stress segments.

5. Stress recognition using a Support Vector Machine

After all the features were generated, they were provided to a learning system, to differentiate the stress state (incongruent Stroop segments) from the normal state (congruent Stroop segments) of a person working on a computer task. In particular, this project employed a Support Vector Machine (SVM) for this learning and classification process.

Support Vector Machines (SVMs) are the computational machine leaning systems that use a hypothesis space of linear functions in a high dimensional feature space to perform supervised classification. An SVM is trained with a collection of known, labeled feature vectors (which in this case are the features derived from the GSR, BVP and

DP signals). SVMs have been shown to perform well in a variety of research areas including pattern recognition [16], text categorization [17], face recognition [18], computer vision [19] and others.

Given a set of training examples, $X = \{x_i : x_i \in R^n\}$, with known labels (targets), $Y = \{y_i : y_i \in$ {possible types}}, a discriminant function, $f : R^n \rightarrow$ {possible types}, where n is the dimension of input vector, has to be learned. The number of misclassifications of f on the training set {X,Y} is minimized by the learning machine(algorithm) during the training phase. The practical interest of these methods is their capacity to predict the class of previously unseen samples (test set). The original data samples in any given data set are typically divided into a training set and a test set. This is done to have feature vectors available for testing that were never presented to the system during the training phase. Testing the system with totally new features vectors provides a more realistic evaluation of effective system performance. Such a strategy for dividing input samples into training and test sets is used in k-fold cross validation techniques[20]. This strategy allows us to train and test on different samples and obviates the need to test on unknown physiological signal samples whose labels (targets) may be uncertain.

The major problem of training a learning machine to perform supervised classification is to find a function (kernel function) that can not only capture the essential properties of the data distribution, but also prevent the over-fitting problem. The support vector machine tries to construct a (linear) discriminant function for the data points in feature space in such a way that the feature vectors of the training samples are separated into classes, while simultaneously maximizing the distance of the discriminant function from the nearest training set feature vector. Choosing a suitable kernel for a SVM is the key to the efficient use of high dimensional feature spaces. It is similar to choosing the architecture for a neural network application. SVM classifiers also allow for non-linear discriminant functions. This is achieved by mapping the input vectors into a different feature space using a mapping function $\Phi : x_i \rightarrow \Phi(x_i)$, and using the vectors, $\Phi(x_i)$, $x_i \in X$, as the feature vectors. The corresponding kernel function used by the SVM algorithm is $K(x_i, x_k) = \langle\Phi(x_i).\Phi(x_k)\rangle$. Standard kernel functions include

$$K(X,Y) = (X \bullet Y + 1)^d,\ d = 1 \qquad (4)$$
$$K(X,Y) = \exp(-\gamma\|X-Y\|^2) \qquad (5)$$
$$K(X,Y) = \tanh(\gamma\,(X \bullet Y)+\theta) \qquad (6)$$

Equation (4) is the polynomial kernel function of degree d which reverts to the linear function when d = 1. Equation (5) is the radial basis function (RBF) kernel with parameter γ, while equation (6) corresponds to a sigmoid kernel.

Many implementations of SVMs are currently available, including mySVM [21], svmTorch [22], SVMLight [23], Gist [24], and LibSVM [25]. We used the LibSVM software package, which can be freely downloaded from http://www.csie.ntu.edu.tw/~cjlin/libsvm for academic use. The core optimization method in LibSVM is based on a decomposition method [23]. Once the SVM classifier is built, the prediction of unknown samples is efficient and rapid since the software only calculates the inner products between the unknown sample and a small subset of feature vectors known as the support vectors.

6. Results

Signals from six experimental subjects have been collected and divided into 36 data entries. Each participant generated data under three non-stress (congruent Stroop) segments and three stress (incongruent Stroop) segments. Ten attributes (GSR_{mean}, IBI_{mean}, IBI_{sd} , etc.) were determined for each data entry. After the feature extraction, the data set had the structure shown in Figure 4.

The prediction performance was evaluated using the jackknife test [20]; each sample was singled out in turn as test samples, and the remaining samples were used to train the classifiers. To evaluate the predictive ability of the classifiers, the total prediction accuracy, which is the number of correctly predicted samples divided by the number of total samples, was calculated for each class.

The goal was to develop and train a system that accepts the various physiological variables as input and predicts the participant's affective state. The SVM was trained to build the model, which could be used to predict the unknown affective state.

Table 1. **Stress prediction accuracies with SVM classifiers using physiological features.**

SVM Classifiers		
Linear Kernel	RBF Kernel	Sigmoid Kernel
Accuracy (%)	Accuracy (%)	Accuracy (%)
57.14%	**60%**	**80%**

Three standard kernel functions were tried on this affective data set. The overall accuracy reached in each case is listed in Table 1. Among the three approaches, the sigmod kernel gives the best prediction performance, 80% and the other two kernels show similar performance, 57.14% and 60%, respectively.

7. Conclusion

The results indicating the classification accuracy of the stress episodes show a promising correlation between the emotional stress and the physiological signals monitored. Additionally, the classifier developed in this project gives more flexibility in defining the feature vectors used for classification. The simplification of the BVP, GSR and PD data into the features derived form them has proven advantageous in providing only essential information to the classification system.

From the tests performed with SVM classifiers in this project, the sigmoid kernel function was more suitable for this physiological application. Although the SVM proposed in this work was able to reach 80% accuracy in differentiating the stress state from the normal working state, other classification methods should still be explored, and their potential to provide an even higher accuracy in the identification of stress episodes evaluated. Some alternative approaches that may be successful in this classification task include: dynamic neural networks, decision trees and fuzzy logic systems. Additionally, larger collections of experimental data should be gathered, to allow for the development of a stronger classifier.

8. Acknowledgements

This work was sponsored by NSF grants IIS-0308155, HRD-0317692 and CNS-0426125. Ms Zhai is the recipient of an FIU Presidential Fellowship. We also thank Dr. Chih-Jen Lin for the use of the LibSVM software.

9. References

[1] E. Hudlicka and M. D. McNeese, "Beyond Cognitive Systems Engineering: Assessing User Affect and Belief States to Implement Adaptive Pilot-Vehicle Interaction," STATE OF THE ART REPORTS (SOARS) 2002.

[2] B. Reeves and C. Nass, The Media Equation: Cambridge Univ. Press, 1996.

[3] W. W. Grings and M. E. Dawson, Emotions and Bodily Responses A psychophysiological Approach: Academic Press, INC, 1978.

[4] K. Hugdahl, Psychophysiology The Mind-Body Perspective: Harvard University Press.

[5] T. Partala and V. Surakka, "Pupil size variation as an indication of affective processing," International Journal of Human-Computer Studies, vol. 59, pp. 185-198, 2003.

[6] A. Barreto and J. Zhai, "Physiological Instrumentation for Real-time Monitoring of Affective State of Computer Users," WSEAS Transactions on Circuits and Systems, vol. 3, pp. 496-501, 2003.

[7] J. R. Stroop, "Interference in serial verbal reactions," Journal of Experimental Psychology, vol. 18, pp. 643-661, 1935.

[8] P. Renaud and J.-P. Blondin, "The stress of Stroop performance: physiological and emotional responses to color-word interference, task pacing, and pacing speed," International Journal of Psychophysiology, vol. 27, pp. 87-97, 1997.

[9] R. M. Stern, W. J. Ray, and K. S. Quigley, Psychophysiological Recording, 2 ed: Oxford University Press, 2001.

[10] R. Fernandez and R. W.Picard, "Signal Processing for Recognition of Human Frustraion," M.I.T 447, 1998.

[11] P. Y. Ktonas, "Automated Spike and Sharp Wave (SSW) Detection," Methods of Analysis of Brain Electrical and Magnetic Signals, EEG Handbook (revised series), vol. 1, pp. 211-241, 1987.

[12] R. K. Dishman, Y. Nakamura, M. E. Garcia, R. W. Thompson, A. L. Dunn, and S. N. Blair, "Heart rate variability, trait anxiety, and perceived stress among physically fit men and women," International Journal of Psychophysiology, vol. 37, pp. 121-133, 2000.

[13] D. W. Rowe, J. Sibert, and D. Irwin, "Heart Rate Variability: Indicator of User State as an Aid to Human-Computer Interaction," CHI, pp. 18-23, 1998.

[14] T. F. o. t. E. S. o. C. t. N. A. S. o. P. Electrophysiology, "Heart Rate Variability : Standards of Measurement, Physiological Interpretation, and Clinical Use," Circulation, vol. 93, pp. 1043-1065, 1996.

[15] E. Kristal-Boneh, M. Raifel, P. Froom, and J. Ribak "Heart rate variability in health and disease," Scand J Work Environ Health, vol. 21, pp. 85-95, 1995.

[16] C. J. C. Burges, "A Tutorial on support vector machines for pattern recognition," Data Mining and Knowledge Discovery, vol. 2, pp. 121-167, 1998.

[17] T. Joachims, "Text categorization with support vector machines: learning with many relevant features," presented at Proc of the10th European Conference on Machine Learning, ECML-98, 1997.

[18] E. Osuna, R. Freund, and F. Girosi, "Training support vector machines: an application to face detection," presented at Proc of the IEEE Conf. Computer Vision and Pattern Recognition, 1997.

[19] B. Scholkopf, S. K.K., C. J. C. Burges, F. Girosi, P. Niyogi, T. Poggio, and V. Vapnik, "Comparing support vector machines with gaussian kernels to radial basis function classifiers," IEEE Transactions on Signal Processing, vol. 45, pp. 2758-2765, 1997.

[20] B. Efron and R. Tibshirani, An Introduction to the Bootstrap. New York, London: Chapman and Hall, 1993.

[21] S. Ruping, "mySVM," 2002.

[22] R. Collobert and S. Bengio, "Svmtorch: support vector machines for large-scale regression problems," J. Machine Learn. Res., vol. 1, pp. 143-160, 2001.

[23] T. Joachims, Making large-scale SVM learning practical: MIT-Press, 1999.

[24] P. Pavlidis, I. Wapinski, and W. S. Noble, "Support vector machine classification on the web," Bioinformatics, vol. 20, pp. 586-7, 2004.

[25] C.-C. Chang and C.-J. Lin, "LIBSVM: a Library for Support Vector Machines. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm," 2001.