

# Stress Detection in Computer Users Based on Digital Signal Processing of Noninvasive Physiological Variables

Jing Zhai, *Student Member, IEEE* and Armando Barreto, *Member, IEEE*

**Abstract**— A stress detection system is developed based on the physiological signals monitored by non-invasive and non-intrusive sensors. The development of this emotion recognition system involved three stages: experiment setup for physiological sensing, signal preprocessing for the extraction of affective features and affective recognition using a learning system. Four signals: Galvanic Skin Response (GSR), Blood Volume Pulse (BVP), Pupil Diameter (PD) and Skin Temperature (ST) are monitored and analyzed to differentiate affective states in a computer user. A Support Vector Machine is used to perform the supervised classification of affective states between “stress” and “relaxed”. Results indicate that the physiological signals monitored do, in fact, have a strong correlation with the changes in emotional state of our experimental subjects when stress stimuli are applied to the interaction environment. It was also found that the pupil diameter was the most significant affective state indicator, compared to the other three physiological signals monitored.

## I. INTRODUCTION

Affective Computing is a new area of computing research described as “computing which relates to, arises from, or deliberately influences emotions” [1]. It emphasizes the importance of adding new emotional features to the human-computer interaction [2]. Affective sensing, i.e., the development of mechanisms to make the computer aware of the user’s affective state, is one key prerequisite to the development of affective computing systems. Our research attempts to explore a way to visualize and evaluate the emotional state identified with ‘stress’ of the users, through several physiological signals that can be measured non-invasively and non-obtrusively, and its results may come in support of the concept of Affective Computing.

## II. METHODOLOGY AND RESULTS

### A. Experimental Setup

Thirty two healthy subjects recruited from the student body at Florida International University (ages 21–42) participated in this study. The subjects were asked to sit comfortably and keep their left hand still when the experiment started. As a preliminary stage, 30 still, emotionally-neutral pictures were

presented to the subject in order to have him/her rest for about 5 minutes. Then the subject took a “Paced Stroop Test” while his/her Galvanic Skin Response (GSR), Blood Volume Pulse (BVP) and Skin Temperature (ST) were recorded from sensors placed in his/her hand, and while his/her Pupil Diameter (PD) was recorded with an eye gaze tracking instrument. In this test the subject was asked to click on the button of a graphic user interface with the label that matched the font color of a word shown on the screen, which named a color. As explained below, the computer game is arranged so that, during selected, identifiable segments of this process, subjects will experience the “Stroop Effect” which has been shown to elicit mental stress in humans. To assure reliability of the changes of PD and ST, the lighting and the temperature of the environment were kept at a constant level during the experiments and the illumination of the eye gaze tracking system was the same for all the subjects.

The computerized “Paced Stroop Test” was developed to mimic a real world stress environment. In the context of human-computer interaction, the stress experienced by users is most likely to be mental (as opposed to physical), and moderate in intensity. Physical stressors occur far less frequently [3]. The Stroop Color-Word Interference Test[4], in its classical version, demands that the color of a word designating a different color be named. In our research, the classical Stroop test was adapted into an interactive version that requires the subject to click on the correct answer rather than stating it verbally. Each trial word is presented to the subject for only 3 seconds. This additional task pacing of the Stroop test might intensify the physiological response [5]. If the subject cannot make a decision within 3 seconds, the screen automatically changes to the next trial. 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, in order to time-stamp the recorded signals at those critical instants [6].

The complete experiment comprises three consecutive sections. In each section, we have an Introductory Segment (IS); a Congruent segment (C), with matching meaning and font color in the Stroop words; an Incongruent segment (IC), with mismatched meaning and font color in the Stroop words; and a Resting Section (RS). Four physiological signals were simultaneously recorded and synchronized by the hardware and software setup throughout the whole experiment to analyze potential concurrent changes that may be due to sympathetic activation associated with ‘stress’ during the

Jing Zhai and Armando Barreto are with the Digital Signal Processing Laboratory at Florida International University, 10555 W. Flagler Street, Miami, Florida, 33174. Phone (305) 348-3711; Fax (305) 348-3707, Emails: jzhai002@fiu.edu, barretoa@fiu.edu

ACKNOWLEDGMENT - This work was sponsored by NSF grants IIS-0308155, CNS-0520811, HRD-0317692 and CNS-0426125. Ms Zhai is the recipient of an FIU Dissertation Year Fellowship.

incongruent Stroop segments. The galvanic skin resistance (GSR), blood volume pulse (BVP), pupil diameter (PD) and skin temperature (ST) were selected for the study because they can be conveniently monitored by non-invasive means (figure 1). The detailed description of the GSR, BVP and PD sensors and the signal collection and associated synchronization methods can be found in our previous report [6]. The skin temperature has the same recording scheme as the GSR and BVP signals.

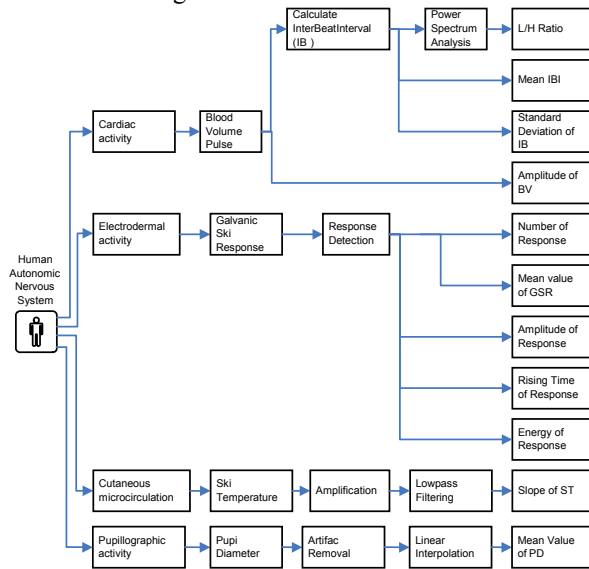


Fig. 1. Physiological features extracted

### B. Feature Extraction and Data Normalization

One key research problem in the Affective Computing area is the mapping between affective states and physiological states [7]. Using multiple features of the physiological signals to indicate the correlations between a set of raw data and the internal stress state in each segment (C or IC) of our experiment is one promising solution to this challenge. In our research a total of 11 features were extracted from each segment of the physiological signals monitored (Figure 1). Our previous report [8] has the detailed descriptions on how this 10 features were extracted from the GSR, BVP and PD recording signals.

For the skin temperature signal, it was expected that the temperature in the finger surface would display transient decreases when the stressor stimuli occur. To extract this information, the amplified ST signal was first filtered to remove recording noise. The slope of the filtered skin temperature in each segment was then used as a feature element of this signal. We found that the patterns of temperature slope gave more meaningful information (resulting in improved classification accuracy) than the patterns of mean value of this signal. One possible explanation for this finding is that the skin temperature seems to obey much longer time constants in its variation, and, as such, does not necessarily reflect well the “current” affective status of the subject, at any given time. In a protocol that included alternation between two types of stimuli (congruent and incongruent Stroop), part of the ST response during one

given interval may still reflect the response to the previous interval. However, the derivative of the changing signal showed an interesting pattern. When the stressor stimuli occur, the slope of the temperature signal generally displayed a negative trend. The slope of the ST signal was estimated using the digital low pass differentiation algorithm  $\frac{1}{T} \frac{dT}{dt}$ , from [9].

The next stage of processing involved application of feature extraction techniques to the congruent (C) and incongruent (IC) segments in each of the 3 sections for each subject, and also to the resting signal recorded when the subjects were watching the still pictures before the experiment started. To account for differences in the initial arousal levels due to individual differences, normalization of the data was needed prior to using the features in the learning classification system. We performed three stages of normalization for all three C segments and three IC segments and also for the preliminary stage signals. After this preprocessing, all features were normalized to the range of [0, 1]. Details of the normalization procedure can be found in our previous report [8].

### C. Support Vector Machine Classification

After all the features were extracted from each Stroop segment (C or IC), they were provided as input 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. A Support Vector Machine (SVM) was used for this learning and classification process. Support Vector Machines are computational machine learning systems that use a hypothesis space of linear functions in a high dimensional feature space to perform supervised classification [10][11][12]. The support vector machine constructs a 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.

The practical interest of this method is the capability to predict the class of previously unseen samples (test set). To verify this capability, the strategy of dividing input samples into training and test sets is used in k-fold cross validation techniques [13]. This strategy eliminates the need to test on unknown physiological signal samples whose labels (targets) may be uncertain. We used the Weka software, that contains a collection of machine learning algorithms for data mining tasks, to implement the SVM used for this research [14].

In this study we were particularly interested in determining the added recognition capability that was achieved by including the pupil diameter measurements as part of the signals monitored, since it had not been studied in combination with other physiological signals. Assessment of the overall usefulness of the pupil diameter signal is particularly important because of the significant increase in complexity required in an affective sensing system that will also monitor this variable. Similarly, we were also interested in gauging the classification value of the other three

monitored signals. As such, we proceeded to perform the classification phase of our project under five conditions: using all the features extracted from all four signals monitored (BVP, GSR, ST and PD); excluding the feature derived from the PD signal, excluding the features from the GSR signal, excluding the features from the BVP signal and excluding the feature from the ST signal. It should be noted that when each individual signal was removed, the learning classifier was retrained and retested using the same cross validation scheme.

#### D. Results

Signals from 32 experimental subjects were collected and divided into 192 data entries, since each participant generated data under three non-stress (Congruent Stroop) segments and three stress (Incongruent Stroop) segments. Eleven attributes ( $GSR_{mean}$ ,  $IBI_{mean}$ ,  $IBI_{sd}$ , etc.) were determined for each data entry. After the feature extraction and normalization stages, the data set from each segment had the structure shown in Figure 1. The classification performance was evaluated using 20-fold cross validation: 20 samples were pulled out as the test samples, and the remaining samples were used to train the classifiers. The goal was to develop and train a system that accepts the various physiological variables as input and indicates the participant's affective state for an unlabeled Stroop segment. The SVM was trained to build a model, which could be used later to identify unlabeled affective states (from the corresponding feature vectors). Performance of such systems is commonly summarized through a confusion matrix which contains information about actual and classified affective states. A confusion matrix summarizing the results of our experiments is shown in Table 1. In this table the two affective state labels are "Stress" (incongruent Stroop segments) and "NS", i.e., non-stressed (congruent Stroop segments).

TABLE 1  
CONFUSION MATRIX FOR STRESS DETECTION USING A SUPPORT VECTOR MACHINE (SHADED CELLS ARE THE NUMBER OF CORRECTLY RECOGNIZED AFFECTIVE STATES)

			Classification	
			NS	Stress
Actual	11 Features	NS	85	11
		Stress	8	88
	PD Removed	NS	51	45
		Stress	29	67
	GSR Removed	NS	86	10
		Stress	10	86
	BVP Removed	NS	88	8
		Stress	12	84
	ST Removed	NS	86	10
		Stress	11	85

To evaluate the classification ability of the system under the different classification conditions, the total classification accuracy, which is the number of correctly classified samples divided by the number of total samples, was calculated for

each classification condition. The overall accuracy reached in each case with the SVM classifier is listed in Table 2.

TABLE 2  
STRESS CLASSIFICATION ACCURACIES WITH SVM CLASSIFIERS USING PHYSIOLOGICAL FEATURES

	Support Vector Machine
11 Features	<b>90.1042%</b>
PD Removed	<b>61.4583%</b>
GSR Removed	<b>89.5833 %</b>
BVP Removed	<b>89.5833 %</b>
ST Removed	<b>89.0625%</b>

### III. DISCUSSION AND CONCLUSION

The results from the research outlined in this paper show a promising correlation between the emotional stress and the physiological signals monitored. These results reinforce the belief that, in the future, non-invasive, non-intrusive real-time assessment of the affective state of a computer user may be achievable. Our results show that, under controlled conditions, the simultaneous monitoring and concurrent processing of four physiological signals: BVP, GSR, PD and ST, yields acceptable levels (up to 90.10%) of differentiation between "relaxed" and "stressed" user states, as elicited by congruent and incongruent "Stroop" stimulation, respectively. The encouraging results obtained highlight the promise held by affective sensing systems which integrate digital signal processing and machine learning algorithms, allowing the affective state of the computer users to be detected with suitable monitoring techniques.

Our development adds to the progress made by groups such as the MIT Affective Computing research group who have developed a real-world driving stress detection system via recording physiological signals (EKG, EMG, GSR and Respiration), achieving an overall accuracy of up to 97% [15]. Their system, however, was developed for a single individual through a recording procedure that lasted 50 minutes. In contrast, the system presented in this paper is a subject-independent, multi-user system and the monitoring duration was less than 10 minutes. Therefore, both developments represent independent avenues of progress towards the establishment of robust affective sensing based on the monitoring of physiological signals.

Since the emotional responses reflect the complex ANS activities expressed by multiple physiological signals, simple separation methods do not seem sufficient to classify the different affective states of a computer user. The SVM classifier, which can deal with more complex pattern distributions, shows a good performance. In a preliminary comparison of the performance of three different classifiers: A Naïve Bayes classifier, a Decision Tree classifier, and an SVM classifier, using all 11 available features, the SVM classifier proved to be the most efficient mechanism for the identification of stress states [8].

Additionally, the implementation of the classification stage of the process with all the available features and with selected subsets of these features has revealed that some physiological

signals seem to provide more informative features to the classification process than others.

For example, it was observed that with the PD signal excluded, the recognition rate dramatically dropped to 61.45%. The interaction of the information provided by the PD signal with the remaining signals is particularly interesting. While the individual PD signal trace does not strictly mimic the temporal pattern of “C” and “IC” segments, its inclusion in the classification process makes an important difference in the results achieved. By involving the pupil dilation signal into the affective recognition, the system performs much better (up to 90.10%). The potential importance of the PD signal for affect recognition may prompt its inclusion in future studies, particularly at a point in time when “web cams” on PCs are becoming ubiquitous and may evolve to a stage in which they might be used for pupil diameter assessment in standard PCs.

In contrast to the performance difference observed when the PD information was removed from the analysis, suppression of the features extracted from GSR, BVP and ST only resulted in small changes in classification performance. This phenomenon makes the pupil diameter prominent amongst physiological signals used for affective state assessment, confirming results from previous research which indicated that the pupil diameter could be an effective indicator of the affective states during tests that used the standard International Affective Digital Sounds (IADS) system [16] for auditory affective stimulation. The recognition of the potential affective sensing value of pupil diameter variations may be particularly relevant since this is not a variable that is commonly incorporated in affective sensing or “detection of deception” (“lie detector”) instrumentation.

An emphasis of the system we presented in this paper was the use of signals that can be collected under normal conditions of computer usage. All the sensors we used are non-invasive and non-intrusive. We expect that, eventually, all sensors could be built into standard pieces of hardware, such as a mouse with BVP, GSR and ST sensors, or integrated in the display unit of the computer, such as a dedicated infrared camera to measure PD. Several groups, (e.g., [1]), have made progress in the development of this type of sensing devices through the last decade.

While it is clear that the conditions for the testing reported here are not typical of human-computer interaction (controlled environmental conditions, prolonged blocks of congruent and incongruent stimulation, etc.), this research confirms the potential of integrated digital signal processing and machine learning algorithms to differentiate key affective states of the computer users from a suitable set of physiological responses.

There are multiple applications that will benefit from advances on the development of practical affective sensing systems, like the one we have described here. For example, a computer tutor capable of recognizing the user’s stress state could individualize its teaching strategy and adapt to the user’s needs. Prendinger [17] has proposed a virtual

quizmaster in a mathematical game to illustrate how to encourage computer users improving their problem-solving capabilities. Rani [18] has proposed an anxiety detection robotic system to enhance collaboration between human and robots. Lisetti [19] has built a multimodal intelligent affective interface for tele-home health care. The “affective mirror” [1] gives a person awareness of how he/she appears to others in based on the integration of physiological signals and other external manifestations, such as facial expression.

## REFERENCES

- [1] R. W. Picard, *Affective computing*. Cambridge, Mass.: MIT Press, 1997.
- [2] E. Hudlicka, "To feel or not to feel: The role of affect in human-computer interaction," *International Journal of Human-Computer Studies*, vol. 59, pp. 1-32, 2003.
- [3] H. Selye, *Selye's Guide to Stress Research*: Van Nostrand Reinhold Company, 1980.
- [4] J. R. Stroop, "Interference in serial verbal reactions," *Journal of Experimental Psychology*, vol. 18, pp. 643-661, 1935.
- [5] 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.
- [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] R. Fernandez and R. W. Picard, "Signal Processing for Recognition of Human Frustration," *M.I.T 447*, 1998.
- [8] J. Zhai, A. Barreto, C. Chin, and L. Chao, "Realization of Stress Detection Using Psychophysiological Signals for Improvement of Human-Computer Interaction," presented at IEEE SouthEastCon, Fort Lauderdale, 2005.
- [9] S. Usui and I. Amidror, "Digital Low-Pass Differentiation for Biological Signal Processing," *IEEE Transactions on Biomedical Engineering*, vol. 29, pp. 686-693, 1982.
- [10] T. Joachims, "Text categorization with support vector machines: learning with many relevant features," presented at Proc of the 10th European Conference on Machine Learning, ECML-98, 1997.
- [11] M. Adjouadi and M. Ayala, "Introducing Neural Studio: An Artificial Neural Networks Simulator for Educational Purposes," *Computers in Education Journal*, vol. 14, pp. 33-40, 2004.
- [12] M. Adjouadi and A. Zong, "Multidimensional Pattern Recognition and Classification of White Blood Cells Using Support Vector Machines," *Journal of Particle and Particle Systems Characterization*, vol. 22, pp. 107-118, 2005.
- [13] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*. New York, London: Chapman and Hall, 1993.
- [14] I. H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*, 2nd ed. San Francisco: Morgan Kaufmann, 2005.
- [15] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 6, pp. 156-166, 2005.
- [16] 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.
- [17] H. Prendinger, J. Mori, and M. Ishizuka, "Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game," *International Journal of Human-Computer Studies*, vol. 62, pp. 231-245, 2005.
- [18] P. Rani, N. Sarkar, C. A. Smith, and L. D. Kirby, "Anxiety Detecting Robotic System-Towards Implicit Human-Robot Collaboration," *Robotica*, vol. 22, pp. 85-95, 2004.
- [19] C. Lisetti, F. Nasoz, C. LeRouge, O. Ozyer, and K. Alvarez, "Developing multimodal intelligent affective interfaces for tele-home health care," *International Journal of Human-Computer Studies*, vol. 59, pp. 245-255, 2003.