STRESS DETECTION IN COMPUTER USERS THROUGH NON-INVASIVE MONITORING OF PHYSIOLOGICAL SIGNALS

JING ZHAI[1], ARMANDO BARRETO [1,2]
[1]Electrical and Computer Engineering Department, [2]Biomedical Engineering Department
Florida International University

KEYWORDS

Physiological Monitoring, Sympathetic Activation, Blood Volume Pulse, Galvanic Skin Resistance, Pupil Diameter, Skin Temperature, Stroop Test, Stress Detection, Affective Computing, Human-Computer Interaction

ABSTRACT

The emerging discipline of Affective Computing pursues the development of computers that could interact with their users taking their affective states into account. For example, if a computer could detect when its user is experiencing stress, it could change the colors and sounds of its user interface to try to calm him/her down. Similarly, the pace of instruction in a computer-based training system could be adapted according to the stress level sensed in the pupil. The research described in this paper aims at the development of a stress detection approach based on automatic monitoring of physiological signals in the computer user. The paper describes the three main aspects of our work: experiment setup for physiological sensing, signal processing to detect the affective state 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 the user, in a non-invasive fashion. 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.

INTRODUCTION

Affective Computing [1] strives to provide computers with the ability to tailor the interaction with their users dynamically, according to the their affective states. Its emergence has highlighted the importance of adding new emotional features to the human-computer interaction [2]. Fulfillment of that promise, however, requires the availability of a robust, practical implementation of emotion detection, i.e., the automatic assessment of the user's affective state. Our research attempts to monitor four physiological signals: Galvanic Skin Response (GSR), Blood Volume Pulse (BVP), Pupil Diameter (PD) and Skin Temperature (ST) to differentiate the affective state associated with "stress" in the user, through non-invasive and non-intrusive techniques. This paper outlines the approach we have followed towards that goal, using digital signal processing and machine learning tools to analyze the physiological signals monitored.

METHODS

1. Experimental Process

Our aim in this research is the detection of mental stress, as physical stressors occur far less frequently in the context of human-computer interaction [3]. Therefore, in order to elicit mental stress at controlled intervals a computerized "Paced Stroop Test" was used. The Stroop Color-Word Interference Test [4], in its classical version, requires 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. Since adding task pacing to the Stroop Test might intensify the physiological responses [5], each trial was designed to only wait 3 seconds for a user response. If the subject could not make a decision within 3 seconds, the screen automatically changed to the next trial. This modified version was implemented with

Macromedia Flash® and also programmed to output bursts of sinusoidal tones 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. The complete experiment comprises three consecutive sections. In each section, we have an Introductory Segment "IS", a Congruent Segment "C" (where the font color and the meaning of the words displayed match), an Incongruent Segment "IC" (where the font color and the meaning of the words displayed differ) and a Resting Section "RS". Four physiological signals were simultaneously recorded and synchronized by the hardware and software integration throughout the whole experiment to analyze potential concurrent changes that may be due to sympathetic activation associated with 'stress'. These signals were: Galvanic skin resistance (GSR), blood volume pulse (BVP), pupil diameter (PD), and skin temperature (ST). We selected these four signals based on the fact that they could be conveniently monitored by non-invasive means. GSR, BVP and ST sensors were attached to the left hand of the subjects. The pupil diameter was recorded through an infrared eye-gaze tracking system. The detailed description of the GSR, BVP and PD sensors and the signal collection and associated synchronization method can be found in our previous report [6]. The skin temperature has the same recording scheme as the GSR and the BVP signals.

Thirty-two healthy subjects (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 played a computer game where he/she repeatedly clicked 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 above, the computer game is arranged so that, during selected, identifiable "IC" 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. Figure 1 shows a complete recording of the four signals after synchronization (From the beginning to the end of the Stroop Test game).

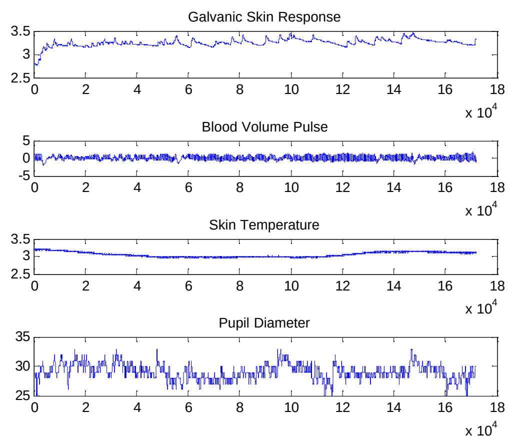

Figure1. Four physiological signals after synchronization

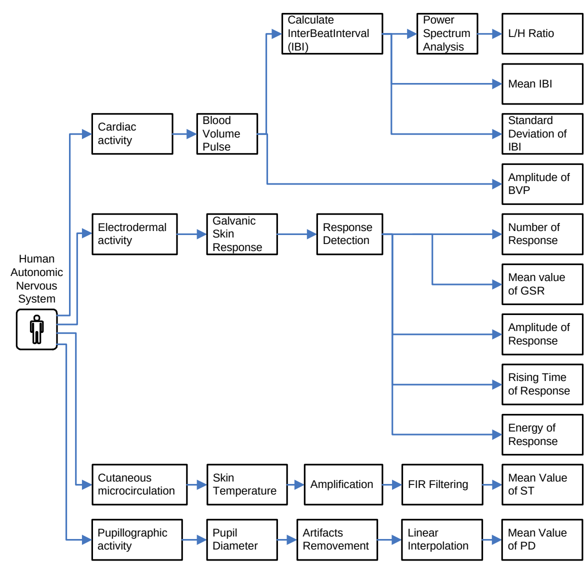

Figure2. Physiological features extracted

2. Feature Extraction and Data Preprocessing

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 game section is one promising solution to this challenge. In our research a total of 11 features are extracted from each segment of the physiological signals monitored (Figure 2). Our previous report [8] has the detailed descriptions on how to extract 10 features from GSR, BVP and PD recording signals. For the newly added skin temperature signal, the amplified ST was first filtered to remove noise components. The average value of the filtered skin temperature in each segment was then used as a feature element of this signal. It was expected that the temperature in the finger surface would display transient decreases when the stressor stimuli occur.

For each subject, the feature extraction techniques were applied to the congruent (C) and incongruent (IC) segments in each of the 3 sections of the complete experiment, and also to the resting signal recorded when the subjects were watching the still pictures before the experiment started. In order to account for the individual baseline levels in the features extracted from the signals monitored, Equation (1) was first applied to obtain corrected feature values for each of the subjects. In this equation Xs represents the feature value for any of the 11 features extracted from the signals recorded during congruent and incongruent segments of the computer game. Xr represents the corresponding feature value extracted from the signals recorded during the resting period, before the computer game started.

$$Ys = \frac{Xs}{Xr} \tag{1}$$

For each subject, there were three C segments and three IC segments. Therefore, six values of any of the features were obtained from the signals recorded during these segments. Equation (2) normalizes each feature value dividing it by the sum of all six segment values.

$$Y_s' = \frac{Ys_i}{\sum_{i=1}^{6} Ys_i} \tag{2}$$

These two stages of normalization proved essential for minimizing the impact of individual subject responses in the training of the learning systems used in our work. After this pre-processing, all features were normalized to the range of [0, 1] using max-min normalization, as shown in Equation (3), to be fed into three learning systems described in the following section.

$$Y_{norm} = \frac{Y'_s - Y'_{s\min}}{Y'_{s\max} - Y'_{s\min}} \tag{3}$$

3. Stress Recognition Using Learning Systems

After all the features were extracted, they were provided as input to learning systems, which were trained 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, three popular learning algorithms were employed: Naïve Bayes Classifier, Decision Tree Classifier, and Support Vector Machine (SVM) for this learning and classification process. We used the Weka software, that contains a collection of machine learning algorithms for data mining tasks, to implement all three learning systems [9].

3.1 Naïve Bayes Classifier: Naïve Bayes classifiers are based on probability models that incorporate class conditional independence assumptions [10]. We simply estimate the probabilities that an object from each class will fall in each cell of the discrete variables (each possible discrete value of the vector variable **X**), and then use Bayes theorem to produce a classification. This method computes the conditional probabilities of the different classes given the values of attributes of an unknown sample and then the classifier will predict that the sample belongs to the class having the highest posterior probability. If an instance is represented by an n-dimensional feature vector, $(x_1,x_2,\ldots,x_n)$, a sample is classified to a class c from a set of possible classes C according to the *maximum a posteriori (MAP)* decision rule:

$$classify(x_1, x_2, ..., x_n) = \underset{c \in C}{\arg\max}\ p(C=c)\prod_{i=1}^{n} p(x_i | C=c) \tag{4}$$

The conditional probability in the above formula is obtained from the estimates of the probability mass function using training data. Although the independence assumption may not be a realistic model of the probabilities involved, it may still permit relatively accurate classification performance.

3.2 Decision Tree Classifier: The decision tree classifier is a 'divide-and-conquer' approach. It has a flow-chart-like tree structure, where each internal node involves testing a particular attribute, each branch represents an outcome of the test, and leaf nodes represent classes or class distributions. The basic idea is to break up a complex decision into a union of several simpler decisions [11]. To classify an unknown sample, it is routed down the tree according to the values of the attributes tested in successive nodes, and when a leaf is reached the instance is classified according to the class assigned to the leaf. A path is traced from the root to a leaf node, which holds the class prediction for that sample. In summary, the general approach has the following steps:

(1) Create a separate tree branch for each value of the chosen attribute.
(2) Divide the instances into subgroups so as to reflect the attribute values of the chosen node.
(3) For each subgroup, terminate the attribute selection process if:
    a. All members of a subgroup have the same value for the output attribute, terminate the attribute selection process for the current path and label the branch on the current path with the specified value.
    b. The subgroup contains a single node or no further distinguishing attributes can be determined. Then, label the branch with the output value seen by the majority of remaining instances.
(4) For each subgroup created in (3) that has not been labeled as terminal, repeat the above process.

In our physiological signal classification task, a J48 decision tree [9] was employed. This is the Weka version of the C4.5 decision tree. The C4.5 is a software extension of the basic ID3 algorithm designed by Quinlan [12, 13] to address some issues not dealt with by ID3 such as avoiding overfitting the data, reduced error pruning and improving computational efficiency, etc.

3.3 Support Vector Machines: Support Vector Machines (SVMs) are the computational machine learning systems that use a hypothesis space of linear functions in a high dimensional feature space to perform supervised classification [14-16]. The support vector machine tries to construct 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. SVM classifiers also allow for non-linear discriminant functions by mapping the input vectors into a different feature space using a mapping function Φ: xi $\rightarrow$ Φ(xi), and using the vectors, Φ(xi), xi $\in$ X, as the feature vectors. The corresponding kernel function used by the SVM algorithm is

$$K(x_i, x_k) = < f(x_i) \cdot f(x_k) > \tag{5}$$

The SVM used in this classification implements John Platt's sequential minimal optimization algorithm (SMO) [17] for training a support vector classifier. Training a Support Vector Machine always requires the solution of a very large quadratic programming (QP) optimization problem. SMO breaks this large QP problem into a series of smallest possible QP problems. These small QP problems are solved analytically, which avoids using a time-consuming numerical QP optimization as an inner loop. The amount of memory required for SMO is linear in the training set size, which allows SMO to handle very large training sets. This new SVM learning algorithm is conceptually simpler, easier to implement, often faster, and has better scaling properties than a standard “chunking" algorithm that uses projected conjugate gradient (PCG) [18].

3.4 Training Sets and Test Sets 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 samples available for testing that were never presented to the system during the training phase. Such a strategy for dividing input samples into training and test sets is used in k-fold cross validation techniques [19]. 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.

RESULTS

Signals from 32 experimental subjects were collected and 192 feature vectors were extracted, 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 2. The prediction 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 predicts the participant’s affective state. The naïve Bayes classifier, decision tree classifier, and SVM were trained to build independent models, which could be used to identify an unknown affective state from a set of 11 features. Performance of such systems is commonly summarized in a confusion matrix, which contains information about the actual and the estimated classifications generated by each one of the systems. The results from our experiments are shown in that fashion in Table 1.

Table 1 **Confusion Matrix for the Three Classification Systems Used (Shaded cells are the number of correctly recognized samples)**

		Predict					
		Naïve Bayes		Decision Tree		Support Vector Machine	
		Rest	Stress	Rest	Stress	Rest	Stress
Actual	Rest	66	30	85	11	88	8
	Stress	11	85	12	84	11	85

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 system. The overall accuracy reached in each case is listed in Table 2. The SVM has the highest prediction accuracy of the three approaches.

Table 2 **Stress Prediction Accuracies With Three Classifiers Using Physiological Features**

Naïve Bayes	Decision Tree	Support Vector Machine
Accuracy (%)	Accuracy (%)	Accuracy (%)
78.65 %	88.02 %	90.10 %

DISCUSSION & CONCLUSIONS

The results from the research outlined in this paper show a promising correlation between the emotional stress and the physiological signals monitored. Accordingly, it seems likely that further refinement of our approach may lead to a real-time implementation of this non-invasive, non-intrusive assessment of the affective state of a computer user. In the tests performed with three independent classifiers, the support vector machine achieved the highest level of success in identifying user stress states, on the basis of the features extracted from the physiological signals monitored. These results have shown 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. Therefore, this work 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.

ACKNOWLEDGMENTS

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

REFERENCES

[1] R. W. Picard, *Affective Computing*. Cambridge,MA: M.I.T. 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 Frustraion," M.I.T 447, 1998.

[8] J. Zhai and A. Barreto, "Realization of Stress Detection using Psychophysiological Signals for Improvement of Human-Computer Interaction," presented at IEEE SouthEastCon, Fort Lauderdale, 2005.

[9] I. H. Witten and E. Frank, *Data Mining: Practical machine learning tools with Java implementations*: Morgan Kaufmann, 2000.

[10] G. H. John and P. Langley, "Estimating Continuous Distributions in Bayesian Classifiers," presented at Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence, 1995.

[11] S. R. Safavian and D. Landgrebe, "A survey of Decision Tree Classifier Methodology," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, pp. 660-674, 1991.

[12] J. R. Quinlan, "Simplifying decision trees," *International journal of Man-Machine Studies*, vol. 27, pp. 221-234, 1987.

[13] J. R. Quinlan, *C4.5: Programs for Machine Learning*: Morgan Kauffman, 1993.

[14] 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.

[15] 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.

[16] 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.

[17] J. C. Platt, "Fast training of support vector machines using sequential minimal optimization " in *Advances in kernel methods: support vector learning* MIT Press, 1999 pp. 185-208

[18] V. Vapnik, *Estimation of Dependences Based on Empirical Data*: Springer-Verlag, 1982.

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