33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Affective Assessment of Computer Users Based on Processing the Pupil Diameter Signal

Peng Ren[1], Armando Barreto[2], Ying Gao[3], and Malek Adjouadi[2]

[1]Biomedical Engineering Department, Florida International University, Miami, FL, USA 33174
[2] Electrical and Computer Engineering Department, Florida International University, Miami, FL, USA 33174
[3]Electrical Engineering Department, University of Wisconsin – Platteville, Platteville, WI, USA 53818

Abstract-- Detecting affective changes of computer users is a current challenge in human-computer interaction which is being addressed with the help of biomedical engineering concepts. This article presents a new approach to recognize the affective state ("relaxation" vs. "stress") of a computer user from analysis of his/her pupil diameter variations caused by sympathetic activation. Wavelet denoising and Kalman filtering methods are first used to remove abrupt changes in the raw Pupil Diameter (PD) signal. Then three features are extracted from the preprocessed PD signal for the affective state classification. Finally, a random tree classifier is implemented, achieving an accuracy of 86.78%. In these experiments the Eye Blink Frequency (EBF), is also recorded and used for affective state classification, but the results show that the PD is a more promising physiological signal for affective assessment.

***Keywords*-Affective Computing, Pupil Diameter, Wavelet denoising, Kalman filter, Walsh transform, Random tree**

I. INTRODUCTION

Affective computing, a relatively new branch of human-computer interaction, attempts to recognize, process, interpret and respond to the affective states of the computer user. It aims to close the communication gap between the human and the machine so as to help people learn perceptively and enhance a variety of other cognitive functions, while interacting with computers [1]. The development of affective computing systems will benefit the future practice of medicine in many ways: "From advisory systems that understand emotional attitudes toward medical outcomes … to computer simulations of emotions and their disorders" [2]. In medical applications of computers, as in many other areas of application, knowledge of a user's affect can provide useful feedback regarding the degree to which a user's goals are being met, enabling dynamic and intelligent adaptation [3].

A variety of methods for measuring affective states in the users have been tried, such as the identification of facial expressions, in isolation, or in combination with speech understanding and body gesture recognition [4]. However, these methods for identifying the emotions of the computer user are susceptible to environmental interference or voluntary masking. Therefore, alternative approaches, which analyze a variety of autonomic activities such as Electroencephalogram (EEG), Electrocardiogram (ECG), Blood Volume Pulse (BVP), Skin Temperature (ST), Galvanic Skin Response (GSR), etc., have been chosen.

From human physiology studies, the autonomic nervous system (ANS) is considered to include three separate systems or divisions: the sympathetic, the parasympathetic and the enteric system (now generally accepted as a separate system whose main function is to innervate the gut region of the body) [5]. The sympathetic nervous system (SNS) originates in the thoracic and lumbar regions of the spinal cord. When fully activated, this division readies the body for a crisis that may require sudden, intense physical activity, which is known as the "fight or flight" response [3]. Generally an increase in sympathetic activity dilates the pupil diameter, causes the sweat glands to secrete copious sweat, increases the heart rate, and so on. On the other hand, the parasympathetic nervous system (PSNS) originates in the brain stem and the lower part of the spinal cord, and typically functions in opposition to the SNS. The parasympathetic division of the ANS stimulates visceral activity and is generally considered to cause a relaxation of the body [3]. These two divisions may work independently or jointly to control a given stage of a complex process. In this study, we monitored the Pupil Diameter (PD) signal and tracked the rate of eye blinks, and proposed a new approach to detect the sympathetic activation associated with a multifaceted emotional state – "stress".

The human pupil is a circular aperture at the center of the iris of the eye, through which light passes to the retina. The range of the pupil diameter is from 1.5mm to 9mm. The diameter of the pupil is controlled by two opposing sets of muscles in the iris, the sphincter and dilator pupillae, which are governed by the parasympathetic and sympathetic divisions of the ANS [6]. The sympathetic ANS division, mediated by posterior hypothalamic nuclei, produces enlargement of the pupil by direct stimulation of the dilator muscles, which causes them to contract [7]. Pupil constriction is caused by excitation of the circular pupillary constriction muscles innervated by fibers from the parasympathetic division. The motor nucleus for these muscles is the Edinger-Westphal nucleus located in the midbrain [8]. The human pupil dilations and constrictions, governed by the ANS, have close relationships with emotions. In fact, previous research showed that the pupil size variation can be an indication of affective processing when using auditory emotional stimulation [9]. Those previous findings prompted us to attempt to use the pupil size variation for the detection of affective changes during human-computer interaction.

II. SIGNAL MONITORING

In this work, we measured and analyzed the PD signal to determine the affective state of a computer user. At the same time, the Eye Blink Frequency (EBF), which can be obtained from the original PD signal, was also monitored, since Haak et al. [10] proposed that there is a strong correlation between eye blink frequency and emotional stress. Bacher and Allen [11] also showed that, during their experiments, participants who reported having experienced a "stressful event" exhibited more blinking than those who reported "no stressful events".

A. Software Used for Affective Stimulation

In our study, the "Stroop Color-Word Interference Test", which elicits mild mental stress in the experimental subjects during controlled intervals, is used in order to observe the changes in the PD signal and its correlation to the affective states of "stress" and "relaxation".

In the test, a word presented to the subject designates a color, and is written in a font color that may ("Congruent") or not ("Incongruent") match its meaning. The subjects were instructed automatically by the program to click one of the five screen buttons to indicate the font color of the word presented. Fig. 1 shows a typical (Incongruent) example of this test interface and Fig. 2 shows the complete experiment protocol, comprising three consecutive sections.

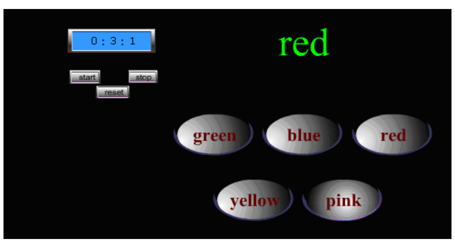

Fig. 1. Sample of the Stroop test interface.

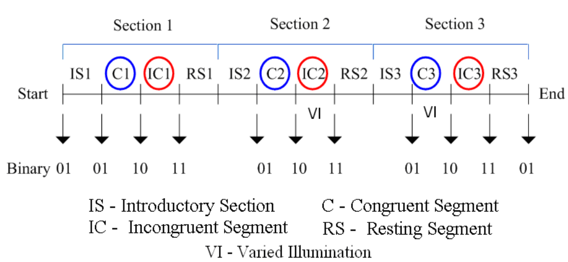

Fig. 2. Stimulus schedule of the experimental protocol.

In each section, there were four segments, including:

- 'IS' – the Introductory Segment to let the subject get used to the task environment, in order to establish an appropriate initial level for his/her psychological state, according to the Law of Initial Values (LIV) ;
- 'C' – the Congruent segment, comprising 45 Stroop congruent word presentations (font color matches the meaning of the word), which are not expected to elicit significant stress in the subject;
- 'IC' – the Incongruent segment, in which the font color and the meaning of the 30 words presented are different, which is expected to induce stress in the subject, according to previous research reported in the psychophysiological literature;
- 'RS' – a Resting Segment to let the subject relax for some time.

At the beginning of each C, IC or RS segments, the binary codes (01, 10 or 11, respectively) shown in Fig. 2, serve as time-stamps for the recorded physiological signals.

B. Hardware Setup

The visual stimuli for the subject (Stroop test) were displayed on the screen of the TOBII T60 eye tracker. The program developed for the eye tracking system allows the extraction of the PD measurements of both eyes and their validity codes at a frequency of 60 samples/second. The average ((left + right) / 2) PD signal, together with time stamp codes created through the experiment and recorded using a multi-channel DAQ system (MCC PCI-DAS6023 board) were saved to permanent storage files for off-line analysis.

III. SIGNAL PROCESSING

A. Physiological Signal Preprocessing

In this research, wavelet denoising and Kalman filtering methods were applied as the preprocessing procedures to remove the noise of the raw PD signal. However, as a first step, the disruptions of the PD signal caused by eye blinking had to be identified in order to compensate them by linear interpolation of the PD data and also to derive the Eye Blink Frequency (EBF), which was also considered for our study. Fig. 3 below shows the results of the blink removal and interpolation process (bottom) on a raw PD signal (top).

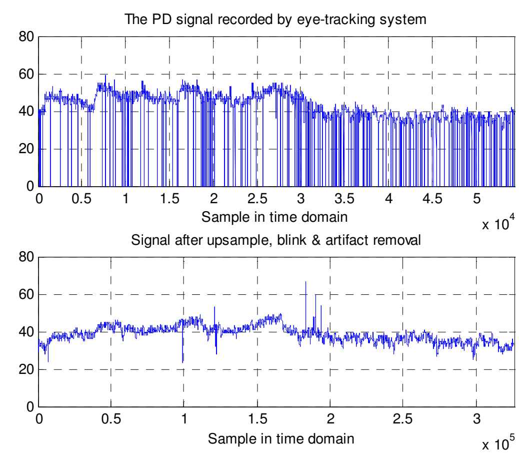

Figure 3. Raw PD data and signal after blinking artifact removal

From the appearance of the signal illustrated in Fig. 3, it is evident that, even after the elimination of blinking transients, the PD signals have a substantial amount of high frequency variability that is not likely caused by the actual size change of the pupil diameter. Therefore, the wavelet denoising and Kalman filtering methods were utilized to lessen this artifact.

Wavelet Denoising removes noise from signals using wavelet transforms, which are able to preserve the shape of the real signal being monitored (in this case, the pupil diameter).The general procedures are as follows [12]:

1. Apply the wavelet transform to the noisy signal to produce the noisy wavelet coefficients (approxima-

tion coefficients and detail coefficients) to a level in which the noise can be separated and removed.
2. Select an appropriate threshold at each level and apply a thresholding method to remove the noise by altering the values of the detail coefficients.
3. Perform inverse wavelet transform on the approximation coefficients and the altered detail coefficients to obtain a denoised signal.

Figure 4 shows the complete PD signal recorded during an experiment (top) and after wavelet denosing (middle). The vertical lines are the segment transition boundaries. The most important boundaries are those that separate each congruent Stroop segment (C) from the Incongruent Stroop segment (IC) that follows. It is obvious from Fig.4 (middle) that wavelet denoising helped remove most of the abrupt changes of the PD but there are still some artifactual sudden changes. Therefore, Kalman filtering is used as a follow-up procedure to further enhance the PD signal.

Kalman filtering is an iterative computational algorithm designed to improve noisy measurements, current state estimates and calculate the forecasts for time series models. The filter is constructed as a mean squared error minimizer, whose weights in the update rules are chosen to ensure that the forecast variances are minimized [13]. The output of the PD signal after Kalman filtering is illustrated in Fig. 4 (lower), which displays a noticeable improvement in the signal. As expected, the processed PD signal has significant increases during the Incongruent segments, compared with the signal in the Congruent segments.

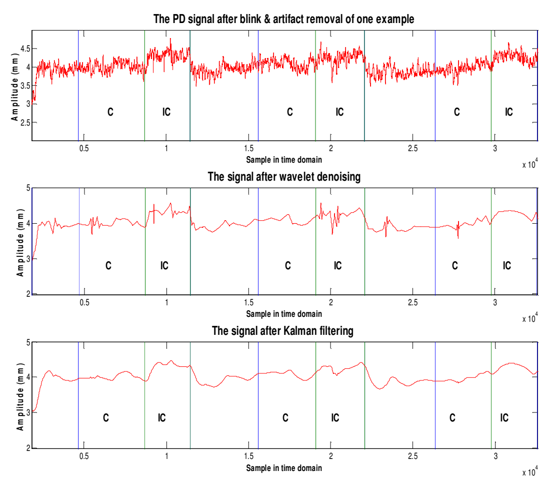

Fig. 4. The original signal(upper), the signal after wavelet denoising(middle) and the signal after Kalman filtering (lower)

B. Data Normalization and Feature Extraction

In order to perform the classification of the segments, the PD signal is normalized to the range [-1, 1] before extracting features from it, since its baseline for different subjects is different. The normalized signal is obtained as:

$$f^{'}(n) = 2 * \frac{f(n) - f_{\min}}{f_{\max} - f_{\min}} - 1 \qquad n=1,2,...,M \qquad (1)$$

where $f(n)$ and $f'(n)$ are the original and normalized signals, respectively.

Figure 4 shows that the filtered PD signal exhibits rapid increases at the beginning of the Incongruent segments, which are not observed at the beginning of the Congruent segments. We sought to obtain a feature of the PD signal that reflected this difference between the onset of "relaxed" (Congruent) and "stressed" (Incongruent) segments. We used the Walsh transform towards this end, since it has been employed successfully to detect interictal spikes in electroencephalogram (EEG) data [14] and the upstrokes in the carotid pulse wave [15], both of which are also characterized by abrupt increases.

The 1D Walsh transform function implemented is defined as

$$W(u)=\frac{1}{N}\sum_{m=0}^{N-1} y(m)\prod_{i=0}^{q-1}(-1)^{b[i](m)*b[q-1-i](u)} \qquad u=0,1,...,N-1 \qquad (2)$$

where y(m) is the one-dimensional sequence being transformed and b[k](u) is the k^{th} bit in the binary representation of u. W(u) are the Walsh coefficients, which define the signal in terms of the functions that serve as basis in the Walsh transform.

In this study, we analyze eight consecutive windows, with 100 samples each, from the beginning of each segment (both C and IC). The eight mean values of these windows are formed into a sequence to represent the trend of the PD signal during the beginning of the C and IC segments. Figure 5 shows the Walsh transform coefficients (bottom) obtained from these sets of 8 mean values for the 6 segments in the PD data (top) from one experimental subject.

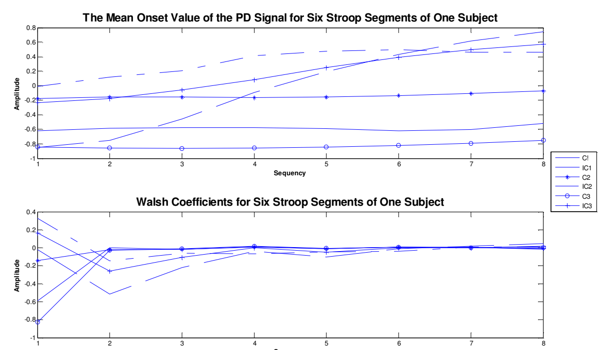

Fig.5. Walsh transform of mean onset values of the PD signal for six Stroop segments of one subject

Application of the Walsh transform decomposes each original signal into a set of orthogonal functions, and encodes the decomposition in the resulting Walsh coefficients, which are ordered in a progression called "sequency". The first and last few coefficients in this ordering are said to be the "low frequent" and "high frequent" components of the PD signal, respectively. In our study, only the "low frequent" components of the PD signal (which represent the overall trend of the signal, rather than the details) are of interest. Therefore the feature sought is extracted from only the first two Walsh coefficients. Specifically, only the difference between the first and the second Walsh coefficient during the onset period of each Stroop segment is utilized as a feature, and denoted as "PDWalsh". For the example illustrated in Fig. 5, the "PDWalsh" values for the three Congruent Stroop segments are 0.5848, 0.1162 and 0.802 whereas for the three Incongruent Stroop segments the "PDWalsh" values are -0.4943, -0.4726 and -0.4268. Similarly, it is expected that positive "PDWalsh" values will be observed in the Incongruent segments and not in the Congruent segments. In this study, three other features derived from PD and EBF signals are also used, which are described in Table I.

TABLE I
FEATURES EXTRACTED FROM THE PD AND EBF SIGNALS

Biosignal	Features	Definition
PD (3 features)	PDmean	Average value of the PD signal
	PDmax	Maximum value of the PD signal
	PDWalsh	Difference value between the first and the second Walsh coefficients based on PD signal during the onset of each Stroop segment
EBF(1 feature)	EBF	Eye Blink Frequency (blinks / time)

IV. AFFECTIVE ASSESSMENT

The "relaxed" vs. "stressed" classification of segments from PD and EBF features in our study was performed using the "Weka" software, which can be freely downloaded from http://www.cs.waikato.ac.nz/ml/weka/. The machine learning classification method utilized is called "Random Tree", and consists of using a decision tree classifier that considers a given number of random features at each node of the tree [16]. Due to the limited space for this paper, we direct the reader interested in details of the well-established decision tree classifier methodology to a comprehensive survey [17]. Classification was performed on 180 data segments (90 "C" and 90 "IC"), collected from 30 subjects.

To compare the significance of the PD and EBF signals for affective assessment ("relaxation" vs. "stress"), three different approaches for classification, involving different subsets of the features derived from PD and EBF, were tried. Table II shows these classification approaches and their corresponding results. The best accuracy was obtained using only the 3 PD features (Approach 2).

TABLE II
RESULTS OF STRESS CLASSFICATION

Approach #	Condition	Accuracy
1	Using all features extracted from PD and EBF (4 features)	77.22%
2	Using features extracted from PD (3 features)	**86.78%**
3	Using the feature extracted from EBF (1feature)	61.67%

V. DISSCUSSION AND CONCLUSION

This paper proposes a new affective assessment approach to classify the "stress" vs. "relaxation" states of computer users through the monitoring of the Pupil Diameter (PD) signal. First, the wavelet denoising and Kalman filtering methods were used to remove abruptly-changing components of the raw PD signal. Then three features, PDmean, PDmax and PDWalsh, were extracted from the filtered PD signal, to emphasize the most meaningful characteristics of PD, for classification purposes. Finally, a random tree classifier was implemented to assess the "relaxation" vs. "stress" affective state of computer users, which achieved an accuracy of 86.78%. Since eye blinks can also be detected in the measured PD signal, we investigated the possibility of also incorporating the Eye Blink Frequency as an additional feature for segment classification. However, the resulting 4-feature classifier was less accurate (only 77.22%). We also observed that a classifier based on EBF alone recorded a much lower accuracy (61.67%). Therefore, all these observations suggest that the PD signal may be a particularly important physiological signal for affective assessment.

ACKNOWLEDGMENTS

This work was sponsored by NSF grants CNS-0959985, and HRD-0833093.

REFERENCES

[1] R.W. Picard, *Affective Computing*, MIT Press, Cambridge, Mass, 1997.

[2] C. Webster, "Medical affective computing: medical informatics meets affective computing". *Studies in Health Technology and Informatics*, vol. 52, pp. 1209-1212, 1998.

[3] J. Zhai and A. Barreto, "Significance of Pupil Diameter Measurements for the Assessment of Affective State in Computer Users", *Biomedical Science Instrumentation*, vol. 42, pp.495-500, 2006.

[4] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotion recognition in human-computer interaction," *Signal Processing Magazine, IEEE*, vol. 18, pp. 32-80, 2001.

[5] K. Hugdahl, *Psychophysiology,* The Mind-Body Perspective: Harvard University Press, 2001.

[6] J. Beatty and B. Lucero-Wagoner, "The pupillary system" in *Handbook of Psychophysiology*, 2 ed., Cacioppo, Tassinary, and Berntson, Eds.: Cambridge University Press, 2000, pp. 142-162.

[7] S. R. Steinhauer, G. J. Siegle, R. Condray, and M. Pless, "Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing," *International Journal of Psychophysiology*, vol. 52, pp. 77-86, 2004.

[8] P. C. Bressloff and C. V. Wood, "Spontaneous oscillations in a nonlinear delayed-feedback shunting model of the pupillary light reflex," *Physical Review*, vol. 58, pp. 3597-3605, 1998.

[9] T. Partala, V. Surakka, "Pupil size variation as an indication of affective processing", *Int. J. of Human-Computer Studies 59*:185-198

[10] M. Haak, S. Bos, S. Panic, L.J.M Rothkrantz, "Detecting Stress Using Eye Blinks and Brain Activity from EEG Signals", *Proceeding of the 1st Driver Car Interaction and Interface (DCII) 2008*, Chez Technical University, Prague, 2008.

[11] L.F. Bacher and K. J. Allen, "Sensitivity of spontaneous eye blinking to type of stimuli." Developmental Psychobiology, 51, 186-197.

[12] S-J S. Tsai, "Power Transformer Partial Discharge Acoustic Signal Detection using Fiber Sensors and Wavelet Analysis, Modeling, and Simulation", *Master's Thesis*, Virginia Technology University, 2002

[13] T. Lacey, "Tutorial: The Kalman Filter", *Computer Vision*, http://www.cc.gatech.edu/classes/cs7322_98_spring/PS/kf1.pdf

[14] M. Adjouadi, M. Cabrerizo, M. Ayala, D. Sanchez, I. Yaylali, P. Jayakar, A. Barreto, "Interictal Spike Detection Using the Walsh Transform", *IEEE Transactions on Biomedical Engineering,* Vol. 51, No. 5, pp.868-873.

[15] B. Meffert, O. Hochmuth, "The application of the Walsh Transform in Biosignal Processing.", *Probl. Techn. Med*., 1987

[16] I. H. Witten, E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques (Second Edition)*, Morgan Kaufmann, June, 2005

[17] 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.