Digital Filtering of Pupil Diameter Variations for the Detection of Stress in Computer Users

Ying Gao
ygao002@fiu.edu, (305) 348-6072
Electrical and Computer Engineering, Florida International University
Miami, Florida 33174, USA

Armando Barreto
barretoa@fiu.edu, (305) 348-3711
Biomedical and Electrical and Computer Engineering, Florida International University
Miami, Florida 33174, USA

Jing Zhai
jzhai002@fiu.edu, (305) 348-6072
Electrical and Computer Engineering, Florida International University
Miami, Florida 33174, USA

Naphtali Rishe
rishen@cis.fiu.edu, (305) 348-2025
School of Computing and Information Sciences, Florida International University
Miami, Florida 33199, USA

ABSTRACT

There is increasing interest in developing means to estimate, in real-time, the level of stress of computer users, particularly for applications such as computer-based tutoring. This real-time stress recognition has been attempted through the processing of a variety of biosignals measured from the computer user, such as the Galvanic Skin Response (GSR), the Blood Volume Pulse (BVP), etc. Recent reports in the literature have strengthen the notion that the Pupil Diameter (PD) can also play an important role in the affective assessment of stress in humans, and its differentiation from a baseline state of relaxation. This paper studies different approaches to perform digital pre-processing of the raw PD data towards the detection of stress states in the computer user. The pre-processing consists of removal, by filtering, of abrupt changes in the PD signal that are not likely to correspond to actual pupillary reactions. This study also summarizes preliminary results for stress detection obtained by imposing a threshold on the filtered PD signal.

1. INTRODUCTION

Previous studies have proposed that the interaction between computing systems and their users would be greatly enhanced if computers had an awareness of the user's affective state. Recent scientific findings have indicated that emotions play an essential role in rational decision making, perception learning, and various cognitive tasks [1]. Therefore, giving computers the capability to be aware of the user's affect can permit more meaningful, natural and productive human-machine interaction.

It is known, from studies in psychophysiology, that the balance of the sympathetic and parasympathetic divisions of the autonomic nervous system (ANS) changes when emotions are elicited [2]. It is also known that the ANS has a significant impact on the regulation of many physiological variables. Therefore, observation of the physiological variables is a critical mechanism to detect a user's affective state. Being aware of the user's affective state could enable the computers to select appropriate strategies to adapt the interface to the user. Some relevant physiological signals that have been chosen in previous research to recognize emotions are the Electroencephalogram (EEG), the Electrocardiogram (ECG), the Electromyogram (EMG), Blood Pressure (BP), Blood Volume Pulse (BVP), Skin Temperature (ST), Galvanic Skin Response (GSR), Heart Rate Variability (HRV), etc. Many of the mechanisms to measure and evaluate these signals have been directly inherited from the "lie detector" evolution. However, the pupil diameter is a physiological variable that has not been fully investigated in its potential for real-time assessment of the affective state of a human subject. This may be due to the fact that the instrumentation required for this kind of real-time measurement did not become available until recently. Nonetheless, it is now established that if the sympathetic division of the ANS is activated (e.g., due to stress) , the pupil is enlarged, and if the parasympathetic division predominates (e.g., during relaxation), the pupil shrinks [3, 4]. In an isolated fashion, it has been verified that the variations of the Pupil Diameter (PD) reflect the emotional change driven by auditory emotional stimulation [5].

In general, the human pupil can constrict to 1.5mm, and dilate to as much as 9mm. Typically, the pupil can react to stimuli within 2 seconds. These constriction and dilation are controlled mainly by the ANS. Janisse [6] pointed out that the relation of pupillary response to ANS activity appears to have been accepted as early as the 1850s by investigators such as Claude Bernard. Charles Darwin related pupil dilation to fear and other emotions in animals in his book [7]. Partala and Surakka have found, using auditory emotional stimulation, that the pupil size variation can be seen as an indication of affective processing [5]. All these known facts prompted us to consider the possibility of using pupil size variation to detect affective state changes during the human-computer interaction.

Unfortunately, the instruments that are typically used for capturing pupil diameter variations in human-computer interaction are not necessarily designed with that primary goal in mind. In many instances, and also in our own case, pupil diameter measurements are obtained as a "byproduct" of the process of monitoring eye movements with an Eye Gaze Tracking (EGT) system. These instruments are optimized to provide estimated point-of-gaze (POG) coordinates, but only provide a sub-optimal pupil diameter signal for the purpose of affective assessment.

The aim of this study is to investigate alternative digital filtering techniques that can be applied to the PD signal obtained from an EGT instrument to facilitate its use in the detection of computer user stress. Further, a thresholding mechanism is proposed and implemented on experimental PD signals filtered by a median filter to evaluate the performance of a stress detector based on them. The paper first introduces the instrumental setup used to record PD signals while stress stimuli were applied to the subjects and the alternative filtering approaches that were initially considered. This is followed by the implementation of a proposed thresholding algorithm on the median-filtered PD signals, and a report of the classification performance obtained by these means in the analysis of our experimental data. The results are discussed and some conclusions are drawn. These results are also considered to guide the directions of future work that might be undertaken towards the ultimate goal of accurate, real-time measurement of pupil diameter changes and their utilization for affective assessment of computer users.

2. METHODOLOGY

Experiment Description

The complete instrumental setup used in our experiments is shown in Figure 1. This instrumental setup was designed for a broader study on affective sensing that included measurement of Blood Volume Pulse (BVP), Galvanic Skin response (GSR) and Skin Temperature (ST), in addition to Pupil Diameter (PD) signals. The part of the instrumentation related to PD recording is enclosed in the dotted-line box.

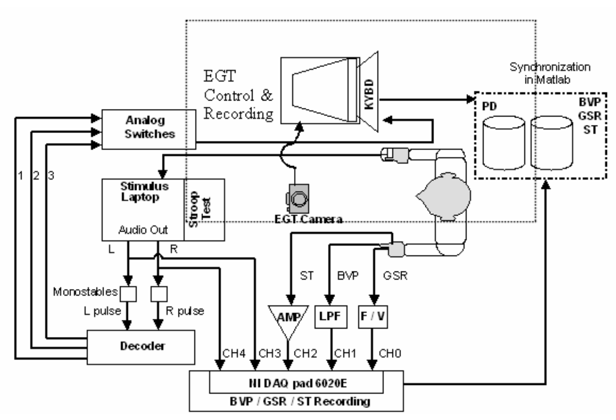

Figure. 1. Instrumental setup.

As indicated in Figure 1, the PD signal was obtained from a desk-mounted Eye Gaze Tracking instrument. Further details on this instrument and the PD signal obtained from it are given in the next section.

In order to observe the changes of pupil diameter due to stress in the computer user, a program was created to elicit mild mental stress in the participating subjects at known times during the experiment. This was accomplished through implementation of a "Paced Stroop Test". Figure 2 shows the stimuli schedule in this experiment from the beginning of the session to its end. In total, the experiment is composed of three consecutive sections. In each section, there were four segments. They are:

1) 'IS' – the Introductory Segment to let the subject get used to the task environment;
2) 'C' – is a Congruent segment of the Stroop Test, in which the subject was asked to click the on-screen button naming the font color of a word that spelled the actual font color being displayed;
3) 'IC' – is an Incongruent segment of the Stroop Test in which the subject was asked to click the on-screen button naming the font color of a word that spelled the name of a *different color*;
4) 'RS' – is a Resting Segment to let the subject relax for some time.

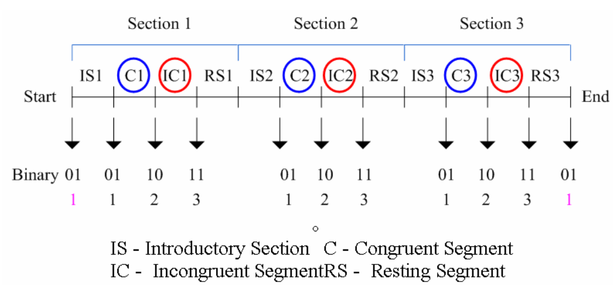

Figure. 2. Stimuli schedule.

The incongruent Stroop segments (IC) were expected to elicit mild mental stress in the subject, according to previous research found in the psychophysiological literature [8]. In contrast, the congruent Stroop segments (C) were expected to allow the subject to continue in a relaxed state. The binary numbers shown in Figure 2 represent the de-multiplexed output of the stimulus generator, which was used in the system to insert the corresponding values (1,2,3) in the event channel of the pupil diameter record obtained from the EGT system, along with the PD values. A previous report on this overall instrumental setup [9] provides more details on this scheme to label the boundaries between experimental segments.

In these experiments, signals from 32 adult, healthy subjects were collected. Three relaxed (congruent Stroop) and three stressed (Incongruent Stroop) segments were recorded from each subject. The lighting of the experimental environment and the brightness of the computer stimuli were kept constant through all the experimental segments and across all the subjects.

Eye Gaze Tracking (EGT) System

In our study, we used the Applied Science Laboratories series 5000 Eye Gaze Tracking system. This system determines first the line-of-gaze, and then the point-of-gaze (POG) from real-time analysis of video images of one of the subject's eyes. The system captures these images through a Sony EVI-D30 pan/tilt/zoom camera fitted with an infrared filter. In addition, the camera has been fitted with infrared illuminators that irradiate the face of the subject (Figure 3) . As a result, the image captures gray-scale images that exhibit two circular reflections: the "glint" or very bright and small "corneal reflection", and the back-reflection of the infrared radiation on the retina of the subject, which is perceivable only through the aperture of the pupil, or "pupil reflection".

Figure. 3. Eye Gaze Tracking camera

The EGT system performs real-time edge detection to isolate these two circular landmarks of each image in the video stream, also estimating their centers and diameters. Figure 4 shows the EGT system monitors and an example of the gray-scale images that the EGT system analyzes in real-time.

Figure. 4. EGT system showing eye image processed

In our case, the PD values calculated by the EGT system were first measured at 60 samples/second, but they were subsequently upsampled by interpolation to 360 samples/second, to synchronize them with the other variables measured during the experiment. The event markers created by the instrumental setup at each of the critical segment transitions (e.g., from a congruent segment to the following incongruent segment) were also stored.

PD Data Processing

In addition to the interpolation process performed to obtain an effective sampling rate of 360 samples/second, the PD signals were analyzed by an algorithm that detected the interruptions due to eye blinks (identified as sudden transitions to a false PD value of zero) and compensated them by interpolation of the lost values. Figure 5 below shows the results of the blink-elimination process (bottom panel), on a raw PD signal (top panel)

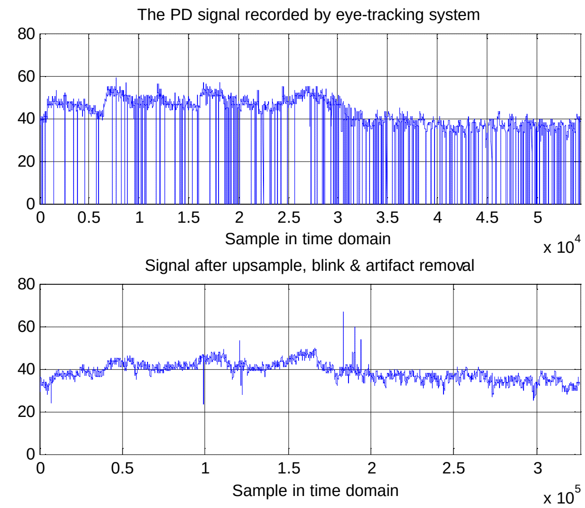

Figure 5. PD data record by the EGT system and signal after blink artifact removal.

It should be noted that, even after elimination of blinking artifacts the PD signal provided by the EGT system contains a significant amount of variability that is not likely to originate from actual pupil size variations, given the extremely fast rate of the changes. Furthermore, when shorter segments of PD data, such as the signal displayed in Figure 6 are analyzed, it becomes apparent that the high variability of the signal is in fact quantization noise, derived from the fact that the pupil only occupies a small portion of the camera's field of view (see Figure 4), which is, in turn, sampled at a relatively low resolution (768 H x 492 V pixels).

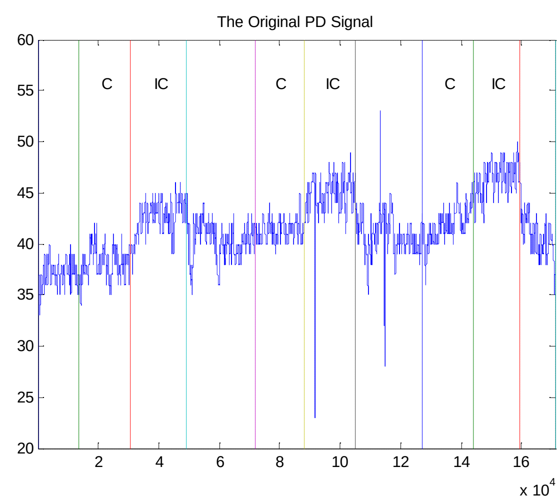

Figure. 6. Typical PD signal recorded during the experiment. Congruent ("C") and Incongruent ("IC") Stroop segments are indicated in the plot.

Figure 6 also shows the segment transition boundaries as vertical lines. The most important boundaries are the ones between congruent Stroop segments (C) and incongruent Stroop segments (IC). The PD signal is expected to be relatively low in the former and higher in the latter. In Figure 6, it is apparent that simple thresholding on the raw PD signal will not result in a very accurate detection result, due to the noise included in the raw PD data. This underscored the need to investigate alternatives for the removal of this noise, by filtering, prior to the application of a threshold.

A first filtering alternative consisted of the application of a digital low-pass Finite Impulse Response (FIR) filter. Several orders and cut-off frequencies were attempted, and the most promising results were obtained with a 42-order FIR low-pass filter with its cut-off set to ¼ of the sampling rate, which in this case was 90 Hz. Figure 7 shows the effect of filtering the same PD data sequence shown in Figure 6 with this low-pass filter. As this figure shows, some level of improvement was achieved, but the result was still not appropriate for a successful classification by thresholding.

Since our interest in the PD signal seemed to be focused on the general trend of its variations, without being concerned with the fast transitions that provide the details of the raw PD waveform, it seemed appropriate to attempt the separation of the components of interest using a wavelet-based approach. The specific denoising process used employed the Daubechies D4 wavelet, level 5, with a threshold value of 54600, established according to Stein's Unbiased Estimate of Risk (SURE) guidelines. An example of the results obtained with this approach is shown in Figure 8, for the same raw PD signal as displayed in Figure 6. It is clear that the success achieved by this method was also limited.

Since the noise observed in the PD signal seemed to have a structure similar to the "salt-and-pepper" noise of corrupted images, which is successfully addressed with a 2-dimensional median filter, we decided to also explore a median filter for this case. We applied a sliding window of 1900 samples, and, at each position of the window the median of the included samples was calculated and assigned as the result. Figure 9 shows the output from this filtering approach for the segment of raw PD data shown in Figure 6.

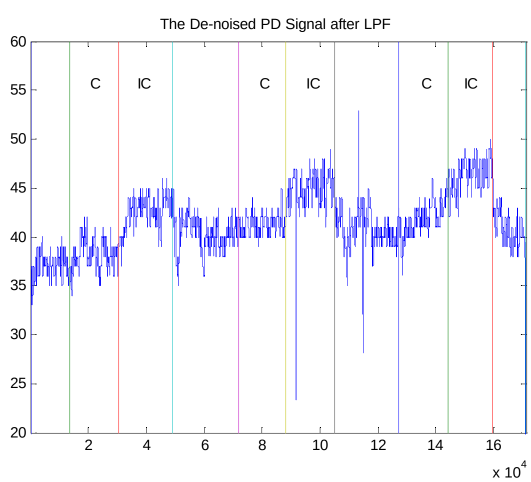

Figure 7. PD filtering using an LP FIR

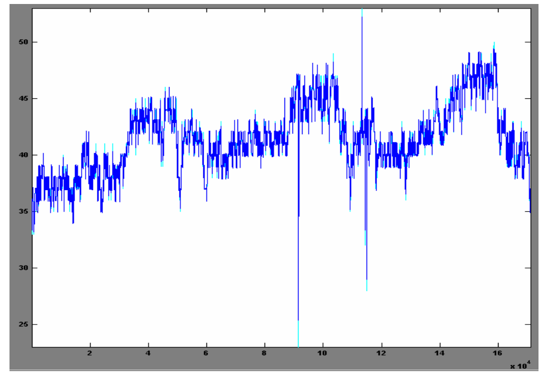

Figure 8. Noise removal output from a Wavelet de-noising approach.

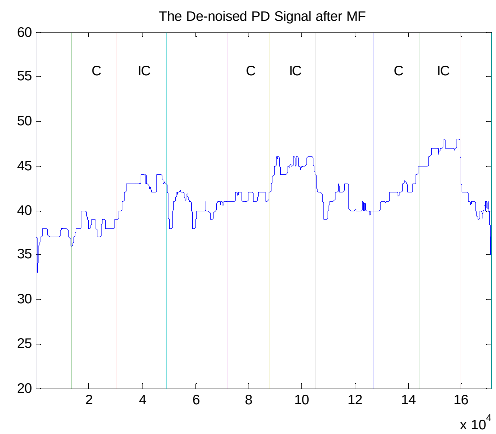

Figure 9. Noise removal output from Median Filter

By comparison of Figures 7, 8 and 9, one can conclude that the results obtained from the median filter (Figure 9) are the ones that are the most appropriate to attempt the detection of stress states (IC segments) through thresholding of the filtered PD signal.

As we proceeded with the application of the median filter to the data of all 32 subjects we noticed that, indeed, the incongruent Stroop segments (IC) displayed an increase in PD values, relative to the preceding congruent Stroop (C) segments. On the other hand, it also became evident that the IC and C PD levels were subject-dependent. This meant that the baseline size of the pupil (in a relaxed state) may be different from subject to subject. Therefore, the threshold for each subject must take into account this individual characteristic.

In particular, it was determined that the threshold could be set at a level that had been offset from the average value of PD in the segments of interest (congruent and incongruent Stroop segments). Specifically, we chose the threshold as

$$thr = m - 0.02 \times \mathrm{var} \quad (1)$$

where m indicates the mean value of the congruent and the incongruent segments, and var is the variance of the PD data within the congruent and incongruent Stroop segments .

3. RESULTS AND DISCUSSION

Selection of pupil diameter filtering method

Figures 7, 8 and 9 show the results obtained from filtering a sample raw PD data segment with an FIR low-pass digital filter, a wavelet de-noising approach, and a median filter. Inspection of these sample results, and other similar graphical outputs obtained in the analysis of data from other subjects, guided our selection of the use of the median filter as our chosen pre-processing block for the classification of PD data towards the detection of stress in the subject.

The selection of the median filter for this pre-processing task also seems appropriate taking into account the nature of the artifacts present in our recorded signals, which seem to have characteristics similar to impulse noise. Numerous reports from the image processing literature have indicated that non-linear filters, such as the median filter, tend to be particularly apt in the removal of this type of noise.

Selection of threshold definition

Our choice to use Equation (1) to define a customized threshold for each subject stems from the recognition that the baseline PD values of each individual in a relaxed state can vary considerably.

Sample results

Figure 10 shows the result of applying the median filter on the raw PD segment first presented in Figure 6. In addition, this figure shows graphically the threshold level determined by Equation 1. For this particular example, the threshold value was calculated as *thr* = 42.483 .

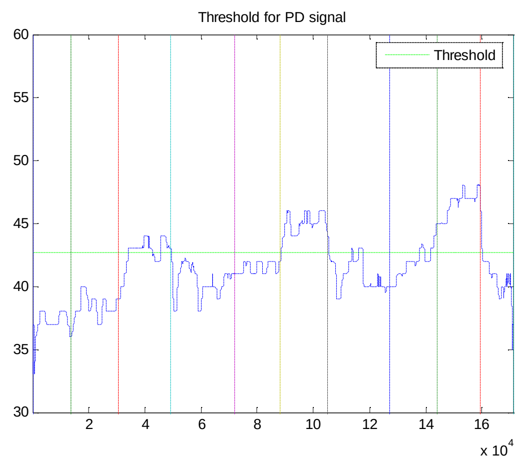

Figure.10. Threshold for the median-filtered PD signal

After the threshold is set, the congruent Stroop and the incongruent Stroop segments can be identified by the following algorithm:

Case 1:
If PD data< threshold, PD data ? Congruent Stroop
Case 2:
If PD data>threshold, PD data ? Incongruent Stroop

The identified congruent and incongruent Stroop sections are marked in Figure. 11.

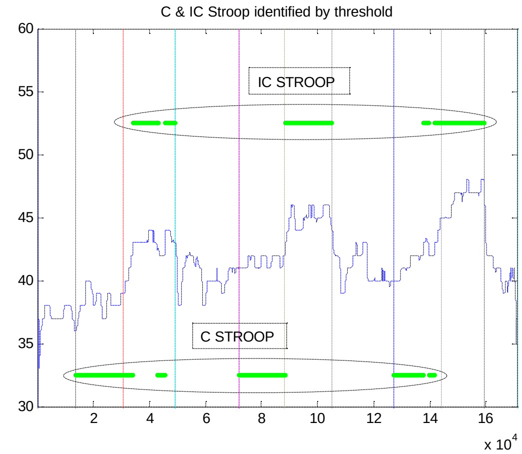

Figure 11. C & IC Stroop identified by threshold

Figure. 11 shows that the threshold we set can successfully distinguish the C and IC Stroop intervals. When the known boundaries of the incongruent Stroop segments are compared with the output of the thresholding operation, it is found that there is 96.42% coincidence between them (for the PD signal from this particular subject.)

Global results

Analyzing the PD time series obtained from all 32 subjects after median filtering in the same way as described in the previous paragraph, the results for agreement in the identification of samples corresponding to an incongruent Stroop ("stress") segment are distributed in the groups illustrated in Table 1.

AGREEMENT	NUM. OF SUBJECTS.
80%~85%	1
85%~90%	4
90%~95%	8
95%~100%	19

Table. 1. Accuracy analysis of the total 32 sets PD data

From 32 sets of data in total, there are 19 in which the agreement of recognized stress and incongruent Stroop is more than 95%. Furthermore, in 27 of the data sets the agreement is above 90%. This seems to indicate that the combination of median filtering performed with the type of adaptive thresholding implemented may hold significant promise for the use of pupil diameter data towards the identification of "stress" states as they occur in computer users.

4. CONCLUSION

This paper investigated the use of three filtering approaches for the partial removal of the impact of quantization noise in pupil diameter signals collected through an eye gaze tracking instrument. We concluded that the application of a median filter seems to be the most successful approach for this goal. In fact, the median-filtered pupil diameter data series were processed to detect the occurrence of incongruent Stroop segments which are expected to be associated with the elicitation of stress in the experimental subjects. Our results indicate that the implementation of a subject-adaptable threshold was able to generate a detection signal that was in good agreement with the time markers identifying the extent of those incongruent Stroop segments in the original data.

5. ACKNOWLEDGMENT

This work was sponsored by NSF grants IIS-0308155, CNS-0520811, HRD-0317692 and CNS-0426125. Ms. Ying Gao is the recipient of a Presidential Enhancement Assistantship from Florida International University.

6. REFERENCES

[1] A. R. Damasio, *Descartes' error*. New York: Grosset/Putnam, 1994.

[2] R. W. Levenson. Emotion and the autonomic nervious system: A prospectus for research on autonomic specificity. In H.L. Wagner, editor, social psychophysiology ans Emotion: Theory and Clinical Applications, page 17-42. John Wiley & Sons, Hoboken, NJ, 1988

[3] R. F. A. Eric Granholm, Andrew J. Sarkin, "Pupillary responses index cognitive resource limitations," *Psychophysiology,* vol. 33, pp. 457-461, 1996.

[4] B. W. Holger Ludtke, Martin Adler, Frank Schaeffel, "Mathematical procedures in data recording and processing of pupillary fatigue waves," *Vision Research,* vol. 38, pp. 2889-2896, 1998.

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

[6] M. P. Janisse, *Pupillometry*. Washington, DC: Hemispheric Publishing, 1977.

[7] C. Darwin, *The expression of the emotions in man and animals*. Chicago: University of Chicago Press, 1965.

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

[9] A. Barreto and J. Zhai, “Physiologic Instrumentation for Real-time Monitoring of Affective State of Computer Users”, WSEAS Transactions on Circuits and Systems, (2004), Vol. 3, Issue 3., pp. 496-501..