Annals of Biomedical Engineering, Vol. 42, No. 1, January 2014 (© *2013*) pp. 162–176
DOI: 10.1007/s10439-013-0880-9

Off-line and On-line Stress Detection Through Processing of the Pupil Diameter Signal

PENG REN,[1] ARMANDO BARRETO,[2] JIAN HUANG,[2] YING GAO,[3] FRANCISCO R. ORTEGA,[4] and MALEK ADJOUADI[2]

[1]Biomedical Engineering Department, Florida International University, Miami, FL 33174, USA; [2]Electrical and Computer Engineering Department, Florida International University, Miami, FL 33174, USA; [3]Electrical Engineering Department, University of Wisconsin-Platteville, Platteville, WI 53818, USA; and [4]School of Computing and Information Sciences, Florida International University, Miami, FL 33199, USA

(*Received 21 March 2013; accepted 24 July 2013; published online 1 August 2013*)

Associate Editor Leonidas D Iasemidis oversaw the review of this article.

Abstract—The pupil diameter (PD), controlled by the autonomic nervous system, seems to provide a strong indication of affective arousal, as found by previous research, but it has not been investigated fully yet. In this study, new approaches based on monitoring and processing the PD signal for off-line and on-line ''relaxation'' vs. ''stress'' differentiation are proposed. For the off-line approach, wavelet denoising, Kalman filtering, data normalization, and feature extraction are sequentially utilized. For the on-line approach, a hard threshold, a moving average window and three stress detection steps are implemented. In order to use only the most reliable data, two types of data selection methods (paired *t* test based on galvanic skin response (GSR) data and subject self-evaluation) are applied, achieving average classification accuracies up to 86.43 and 87.20% for off-line and 72.30 and 73.55% for on-line algorithms, with each set of selected data, respectively. The GSR was also monitored and processed in our experiments for comparison purposes, with the highest classification rate achieved being only 63.57% (based on the off-line processing algorithm). The overall results show that the PD signal is more effective and robust for differentiating ''relaxation'' vs. ''stress,'' in comparison with the traditionally used GSR signal.

Keywords—Autonomic nervous system (ANS), Pupil diameter (PD), Galvanic skin response (GSR), Wavelet denoising, Kalman filtering, Moving average window, Backward differentiation, Mathematical morphology.

INTRODUCTION

In the field of psychophysiology, some relevant physiological signals, which are controlled by the autonomic nervous system (ANS), have been chosen to reflect the inherent activity of the nervous system. The ANS is the regulator and coordinator of important bodily activities below the level of consciousness, including digestion, body temperature, blood pressure (BP), and many aspects of emotional behavior.[6] Two branches of the ANS are the sympathetic nervous system (SNS) and the parasympathetic nervous system (PSNS). The SNS activities may include increased heart rate, BP, sweating, cardiac output, and respiration changes that enhance a fight-or-flight reaction.[6,18] On the other hand, the function of the PSNS is associated with the relaxation of the body and its activation promotes a return of several organs to regular function. Activation of the PSNS may slow the heart rate, promote peristalsis, increase salivary secretions, and so on. Although the SNS and the PSNS have contrasting functions, their activities are integrated and not necessarily antagonistic.[6,18]

''Relaxation'' vs. ''stress'' of the human beings, which have been studied by previous researchers,[2,4,5,21,25,26] involve observable manifestations of activations of the SNS and the PSNS. In the context of human–computer interaction stress is often found defined along the lines of H. Selye's conceptualization of stress. For example, Healey and Picard[13] indicate, citing Selye: ''Historically, stress has been defined as a reaction from a calm state to an excited state for the purpose of preserving the integrity of the organism. For an organism as highly developed and independent of the natural environment as socialized man, most stressors are intellectual, emotional, and perceptual.'' This proposed definition of stress reveals the close relationship that stress may have with cognitive processes, as an increase in cognitive

Address correspondence to Peng Ren, Biomedical Engineering Department, Florida International University, Miami, FL 33174, USA. Electronic mail: pren001@fiu.edu

demands may cause an individual to become stressed. In contrast, relaxation means a tension-free state in which internal conflicts and disturbing feelings of anxiety, anger, and fear are eased and a state of tranquility prevails.[22] In any case, stress can be damaging to the well being of the individual. Recent research has demonstrated that not only severe stress, lasting weeks or months, but also the short-term stress (as little as few hours) can impair cell communication in the brain,[36] and therefore, would be important to detect early on. Furthermore, automated stress detection may enhance security (for example, in banks or airports) or help in criminal investigations (for example, use of the polygraph, based on the stress experienced by someone who lies).

In the past many physiological signals, which can reflect ANS reactions,[12] have been used to assess the ''relaxation'' vs. ''stress'' of the human beings, such as electrocardiogram (ECG), BP, skin temperature (ST), galvanic skin response (GSR), blood volume pulse (BVP), *etc*. More recently, the pupil size was verified by Partala and Surakka[20] as an evident indication of affective arousal in their auditory emotional stimulation experiment. Similarly, our own group explored the feasibility of monitoring multiple physiological signals, such as GSR, BVP, ST, including pupil diameter (PD), for non-invasive stress detection in computer users,[39] developing a system that combined 11 features derived from the 4 signals. This system achieved an average accuracy of 85.59%, across three different classification algorithms, on a strictly off-line basis. These results began to indicate the high potential of PD as an indicator of arousal. However, the efficiency and the robustness of the PD variation monitoring for differentiation between ''relaxation'' and ''stress'' have not been fully explored yet. The pupil, which can constrict to 1.5 mm in diameter or dilate to about 8–9 mm, is the opening through which light enters the eye and begins the process of visual perception.[3] The diameter of the human pupil is controlled by two opposing sets of muscles in the iris, the sphincter, and dilator pupillae, which are governed by the SNS and the PSNS of the ANS.[3] If the SNS is activated (e.g., due to stress), the size of PD tends to increase; whereas if the PSNS dominates (e.g., during relaxation), the PD will remain small.[1,10,28,29,37]

In our study, the PD signal, obtained with an eye gaze tracking system, which has become a robust and intuitive tool for human–computer interaction,[15] was chosen to be monitored for affective sensing in a non-invasive way. The GSR is one of the commonly used physiological signals to detect stress, fear, lying, anxiety, and arousal as these events tend to make the sweat glands more active and this lowers the skin's resistance.[17,23] Therefore the GSR signal was also measured and analyzed in order to compare the efficiency of our proposed off-line and on-line approaches for stress detection of a human subject based on the PD signal.

METHODS

Experiment Setup

In our experiment, PD and GSR signals were monitored in order to detect the emotional change (''relaxation'' vs. ''stress'') of a human subject. In addition, as the intensity of light is another primary factor affecting the pupillary constriction and dilation, the Illumination Intensity (IL) in the environment was also measured and recorded.

The ''Stroop color-word interference test'' (SCWT)[30] was used to elicit mild mental stress in the experimental subjects during controlled intervals with the aim of recognizing the ''relaxation'' and ''stress'' of the subject based on identifying PD and GSR signal variations. Previously, the SCWT has been applied by several research groups as the way to elicit stress in experimental subjects.[14,31,35,39] Tulen *et al.*[35] evaluated the SCWT as a test for the study of stress-induced sympathetic effects, on the basis of psychological, physiological, and biochemical responses. They demonstrated that the SCWT can induce increases in plasma and urinary adrenaline, heart rate, respiration rate, electrodermal activity, electromyography, feelings of anxiety, and decreased finger pulse amplitude. In addition, Hjemdahl *et al.*[14] studied the sympatho-adrenal and hemodynamic responses to mental stress induced by the SCWT, which increased heart rate and BP by 28 beats/min and 29/14 mmHg for the subjects, on average. Further, Sun *et al.*[31] utilized the SCWT as the psychological or cognitive stressor to introduce an emotional response, and simultaneously measured the GSR and ECG data to assess the ''stress'' affective state of the subject. These studies have used the SCWT to elicit stress in experimental subjects, which was also the goal in our study. However, as noted, earlier, the specific form of stress elicited by SCWT may not be separate from a simultaneous increase in the cognitive demands placed on the subject. Therefore, the observations made in our work must be understood in the specific frame of responses to intellectual stressors, to which SCWT belongs.

In the test, a word with the font color that may (''congruent'') or not (''incongruent'') match its meaning was presented to the subject, as shown in Fig. 1. The subjects needed to read aloud the word presented first and then were required to click one of the five screen buttons to indicate the font color of the word within 3 s, otherwise the system would automatically display the next word. There were 45 consecutive word presentations in each congruent segment and 30 consecutive word presentations in each incongruent segment.

Figure 2 shows the stimuli schedule in this experiment from the beginning of the session to its end. In total, the experiment includes three consecutive sections. The aim of our study was to analyze PD and GSR signals and their changes from ''relaxation'' to ''stress,'' therefore it was essential to mark the boundaries of the C and IC segments in the data. For this purpose, when one of the C, IC, or RS (respectively) segments started, bursts of sinusoidal tones were output through the sound system of the computer following a binary encoding (01, 10, or 11), as shown in Fig. 2. The luminance intensity remained constant except in the segments IC2 and C3, where the illumination was temporarily increased (marked as ''VI'' in Fig. 2). The experimental setup included measurement of the PD, GSR, and IL signals. The visual stimuli for the subject (Stroop color word presentations) were displayed on the TOBII T60 eye tracker monitor. TOBII T60 is robust, being able to operate in a wide range of head movement. This provides a distraction-free test environment that ensures natural behavior, and therefore realistic results. The high level of performance of this type of eye gaze trackers has been evaluated by other research groups.[19,38] Specifically, Klingner[16] has found that a similar TOBII remote eye tracker (TOBII model 1750) measures mean binocular PD with precision of 0.10 mm. In addition, before starting the experiment, each subject must go through two preliminary procedures for confirming eye tracking functionality (verify that the eye tracking system can detect both eyes of the subject) and calibration (9-point calibration procedure, which can allow certain level of tolerance for small head movements of the subject). The TOBII system measured the PD values 60 times per second. For use in the off-line study, the relevant variables from the eye tracking system (in this case, the PD of both eyes and their validity code) were stored at the frequency of 60 Hz and, later, read into MATLAB®. In addition, the subject had the GSR sensor (GSR 2, from Thought Technology) attached to his/her left hand and the IL sensor (BS500B0F photodiode, from Sharp) on his/her forehead, above the eyes. These signals, together with the left and right audio output (to provide the corresponding time stamping in the experiment) were recorded and converted to a MATLAB-readable data file directly at a rate of 360 samples/s, using a multi-channel MCC DAQ system (PCI-DAS6023 board). Later, for use in the off-line study, these data were down-sampled by six, to establish the same sampling rate of 60 samples/s for all measured signals. For the on-line stress detection algorithm, a buffer of data was captured while, simultaneously, the previous buffer of data was processed.

The experimental sequence shown in Fig. 2 attempts to expose each subject to alternating states of absence and then presence of our stressor stimulus (the incongruence of SCWT word presentations in the ''IC'' segments of the protocol), keeping all other parameters of the test environment (e.g., position of the subject and the display, overall brightness of the images displayed in the monitor, *etc.*) and the functional characteristics of the subject (his/her age, visual performance, *etc.*) constant. We also tried to minimize changes in those functional characteristics that could be introduced by habituation or fatigue. These precautions represent our best effort to minimize the impact of other causes of PD change (e.g., light and dark adaptation, fixation, accommodation, *etc.*) That is, while we were not able to ensure the elimination of these other events, we have attempted to minimize their potential impact in our study by fostering their (approximately) equal presence (or absence) during the exposure of each of our subjects to the two conditions that we are studying: congruent (C) and incongruent (IC) Stroop stimulation. Further, we have attempted to minimize the impact of inter-subject variability of

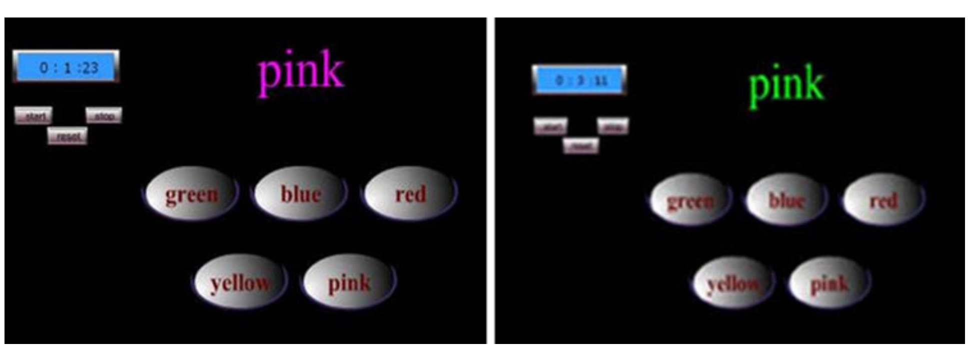

FIGURE 1. Samples of the Stroop test interface. The left panel shows a ''congruent'' word presentation (word ''pink'' in pink font). The right panel shows an ''incongruent'' word presentation (word ''pink'' in green font).

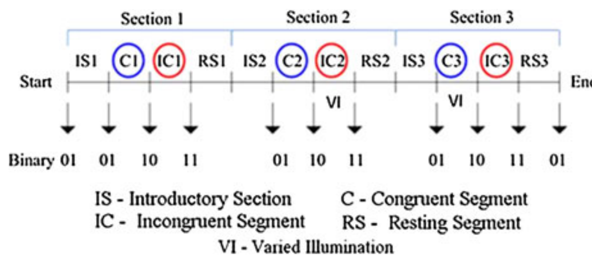

FIGURE 2. Stimuli schedule of the experimental protocol. ''IS''—the introductory segment to let the subject get used to the task environment, in order to establish an appropriate initial level for his/her psychological state, according to the law of initial values (LIV); ''C''—the congruent segment, comprising 45 Stroop congruent color 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 differ, which is expected to induce stress in the subject; ''RS''—a resting segment to let the subject return to a baseline affective arousal, not having to perform any action, during 1 min. The physiological signals are also monitored.

those extraneous factors in our off-line processing approach by normalizing the preprocessed PD signal to the same range of [−1 1], before feature extraction.

The experiment involved 42 individuals (23 male, 19 female), with ages ranging from 20 to 50 (mean 27.5; SD 5.14), and from diverse professional and ethnic backgrounds. Subjects were asked if they had any problems with color vision and none of them reported any. It should be noted that our experiment was approved by the Institutional Review Board of Florida International University. In the experiments performed for this study, the participant was asked to remain seated in front of the TOBII screen, interacting with the Stroop Test program for about 30 min. All the normal lights in the room were kept on, and an additional level of illumination provided by a desk lamp placed above the eye level of the subject was also switched on during the IC2 and C3 segments (as shown in Fig. 2) in order to explore the robustness of our stress detection algorithm to illumination intensity variations in the environment. In the IC2 segment the increase of the light intensity would inhibit the dilation of the PD caused by stress, whereas in the C3 segment the increase of the light intensity would further promote the contraction of the PD caused by the relaxation of the human subjects.

Development of Off-line Physiological Signal Processing Algorithm

Previous research efforts in the area of pupillometry have resulted in traditional processing approaches applied to the PD signal.[24] However, we sought to develop alternative processing approaches meant to preserve specifically the aspects of the PD signal that we found to be more relevant to the differentiation between stress and relaxation, in order to facilitate the feature extraction step of our method. In our study, the combined implementation of wavelet denoising and Kalman filtering is the main signal processing method applied to the original PD data, after the removal of the eye blinks by linear interpolation (eye blinks are signaled by a value of ''4'' in the validity code from the TOBII system). Figure 3 shows a set of signals (including the raw PD signal, the IL signal, and the raw GSR signal) recorded from one subject during the experiment. The first plot of Fig. 4 (upper trace) shows the PD signal after blink and artifact removal, which is called the ''original PD signal'' in this paper. It should be noted that, although the linear interpolation removes the blinking artifacts, a substantial amount of fast variability still remains, which is not likely to represent pupil size changes due to affective variations. Furthermore, previous research[9] has indicated that these fast variations in the PD signal could be due to quantization noise in the PD measurement.

Wavelet denoising seems to be a proper approach to remove the abrupt changes in the original PD signal. There are three main steps for wavelet denoising.[34] The multiresolution analysis (MRA) makes use of the discrete wavelet transform (DWT) with the dilated and translated versions of the mother wavelet. In MRA, function scaling is performed to create a series of approximations of the signal and the versions of the wavelet are used to encode the difference in information between different approximations. The subspaces spanned by the scaling function at low scales are nested within those spanned at higher scales. Therefore, we can use the recursive algorithm for wavelet signal decomposition by implementing low/high-pass filters and the down-sampling operation to result in levels of approximation and detail coefficients. In our study, the Daubechies wavelet and nine decomposition levels were applied to the original PD signal. The second step is to select a proper threshold method to best remove the noise by altering the values of the detail coefficients. In our research, the Birge–Massart strategy, a method based on adaptive functional estimation in regression or density contexts, is used to set the level-dependent threshold for denoising. The last step is to apply the inverse DWT on the approximation coefficients and the altered detail coefficients to obtain the denoised signal, shown in the second plot of Fig. 4.

Although the application of wavelet denoising to the original PD signal achieves significant improvement in clearing the signal, there are still some abrupt changes that remain. Kalman filtering is utilized as the second step to remove the noise that remained in the PD signal. Kalman filtering, which is a recursive data

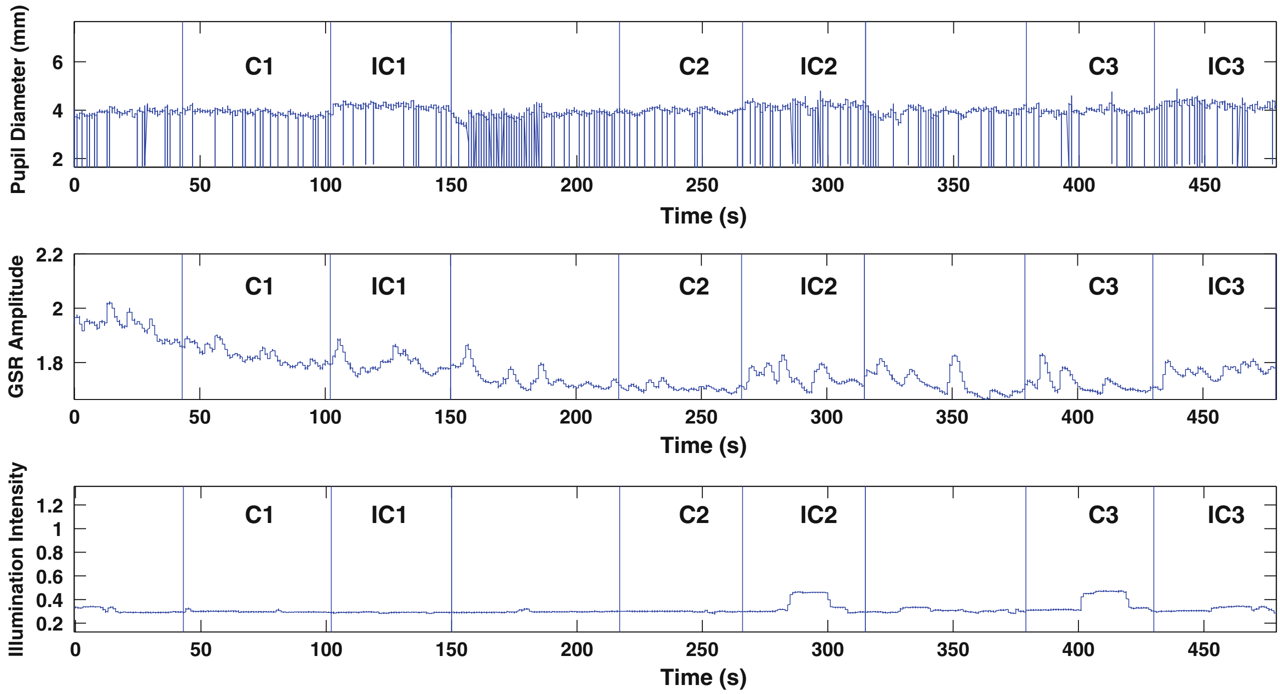

FIGURE 3. A set of signals after synchronization. From top to bottom: the raw PD, the raw GSR, and IL. The vertical lines are the segment transition boundaries. The most important three boundaries are those that separate each congruent Stroop segment (C) from the incongruent Stroop segment (IC) that follows.

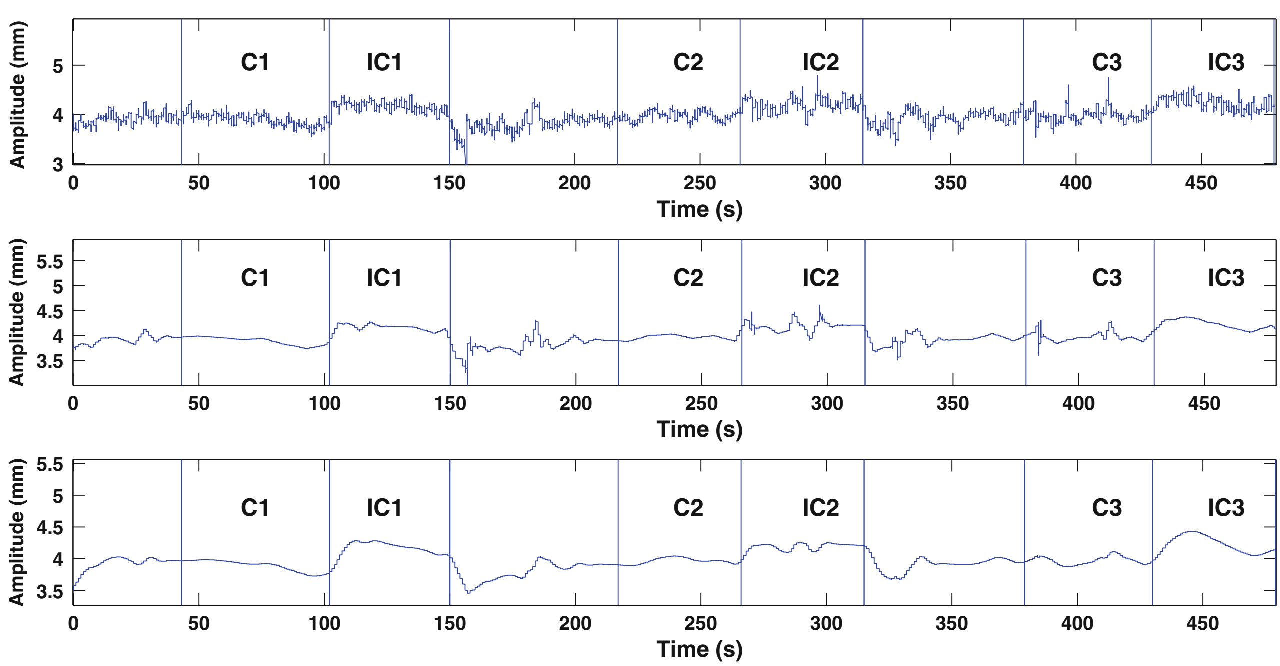

FIGURE 4. Original PD signal (upper), signal after wavelet denoising (middle) and signal after Kalman filtering (lower).

processing algorithm, generates optimal estimates of desired quantities given a set of measurements. The algorithm works in two steps: in the prediction part, the Kalman filtering uses initial conditions and models to produce estimates of the current state variables, along with their uncertainties. After observation of the next measurement, the correction step is implemented. The estimated variables are updated based on constructing a mean squared error minimizer in order to ensure that the prediction variances are minimized.[11]

The PD signal after Kalman filtering is illustrated in the lower plot of Fig. 4.

The preprocessed PD signal is then normalized to the range [−1 1] because different subjects have different baseline PD signal levels. Then key features are extracted from the filtered PD signal. The GSR signal recorded during the protocol was preprocessed by a 64-order low-pass FIR filter (1.066 s buffer) with a cutoff frequency at 2 Hz. for the purpose of removing recording noise and artifacts.[8] Table 1 shows three features respectively extracted from the PD and GSR Signals. Table 2 illustrates three different classification phases (P1, P2, P3) for the purpose of evaluating the efficiency of the PD signal, in comparison to the performance of the GSR signal.

Development of On-line Physiological Signal Processing Algorithm

The on-line system derives a representative point for every short time interval of the preprocessed raw PD signal, labeled as PD_r, to achieve stress detection. In the on-line processing algorithm, a hard threshold setting for eye blink removal, which is denoted as "$\text{Threshold}_{\text{blink}}$," and a moving average window method are applied in order to obtain the PD_r. This threshold is applied to the PD signal at its original scale, since we do not have the benefit of knowledge of the complete data set (in the on-line algorithm we do not involve the value of future samples) and we cannot, therefore, normalize the signal. The moving average, one of the most common signal processing methods, is not only simple but it also produces "the lowest noise possible for a given edge sharpness,"[27] which is suitable for the shape of the PD signal when it changes from a congruent (C) to an incongruent segment (IC). Originally, the sampling rate of the raw PD signal is 60 Hz. It is neither necessary nor practical to classify every PD signal sample for stress detection. Therefore, 60 samples, namely a time span of 1 s, were selected as window length to achieve one PD_r value per interval. Due to the presence of eye blinks, a hard threshold must be set ("$Threshold_{\text{blink}}$"). The detailed algorithm is presented below:

$$m = 0, \tag{1}$$

$$\text{If} \quad PD(i) \geq Threshold_{\text{blink}} \quad X(i) = PD(i); \;\; m = m + 1 \quad (60 \times n - 59 \leq i \leq 60 \times n \quad n = 1, 2, \ldots), \tag{2}$$

$$\text{If} \quad PD(i) < Threshold_{\text{blink}} \quad X(i) = 0 \quad (60 \times n - 59 \leq i \leq 60 \times n \quad n = 1, 2, \ldots), \tag{3}$$

$$PD_r(n) = [X(60 \times n - 59) + X(60 \times n - 58) + \cdots + X(60 \times n)]/m, \tag{4}$$

where n is the number of the 1-s window being considered (which will be assigned as the index of the resulting PD_r sample), i is the original PD sample number counted from the first sample of the first 1-s window involved in the analysis, and m counts the number of PD samples with value greater than or equal to $Threshold_{\text{blink}}$, in a given 1-s window. Figure 5 shows the preprocessed PD signal after use of a hard threshold (for the removal of the eye blinks) and the application of a moving average window. It can be seen in Fig. 5 that while the length of the original PD signal in a typical record is about 28,800 samples, the preprocessed signal is just about 480 data points (PD_r).

The next step is differentiating between relaxation and stress. For this analysis, the most important parts of the waveform in the PD_r signal are the large sudden increases and the large sudden decreases, which appear in the emotional transitions from "relaxation" to "stress" and from "stress" to "relaxation," respectively. In this step, there are three sub-steps for identifying "relaxation" and "stress," which are preparation, feature-based decision voting and "relaxation"/"stress" indication.

Sub-step 1 (Preparation) The PD can constrict to 1.5 mm or dilate to about 8–9 mm. In addition, different individuals have different PD values even under the same experimental conditions. Hence, it is reasonable and necessary to calculate the mean value of PD during a certain period of time when each subject is relaxed as the reference baseline for that subject. In our

TABLE 1. Features obtained from the PD and GSR signals.

Signal	Features	Definition
PD (3 features)	PDmean	Average value of the PD signal in a segment
	PDmax	Maximum value of the PD signal in a segment
	PDWalsh	Difference value between the first and the second Walsh coefficient after Walsh transform based on the PD signal during the onset of each Stroop segment
GSR (3 features)	GSRmean	Mean value of the amplitude of each GSR response in a segment
	GSRrisingTime	(Average) rising time of each GSR response in a segment
	GSRnum	Number of the GSR responses in a segment

TABLE 2. Classification phases (different conditions).

Phase	Description
P1	Using all features extracted from the monitored PD and GSR signals (all six features are used for classification)
P2	Excluding the features extracted from the PD signal (only three features from GSR)
P3	Excluding the features extracted from the GSR signal (only three features from PD)

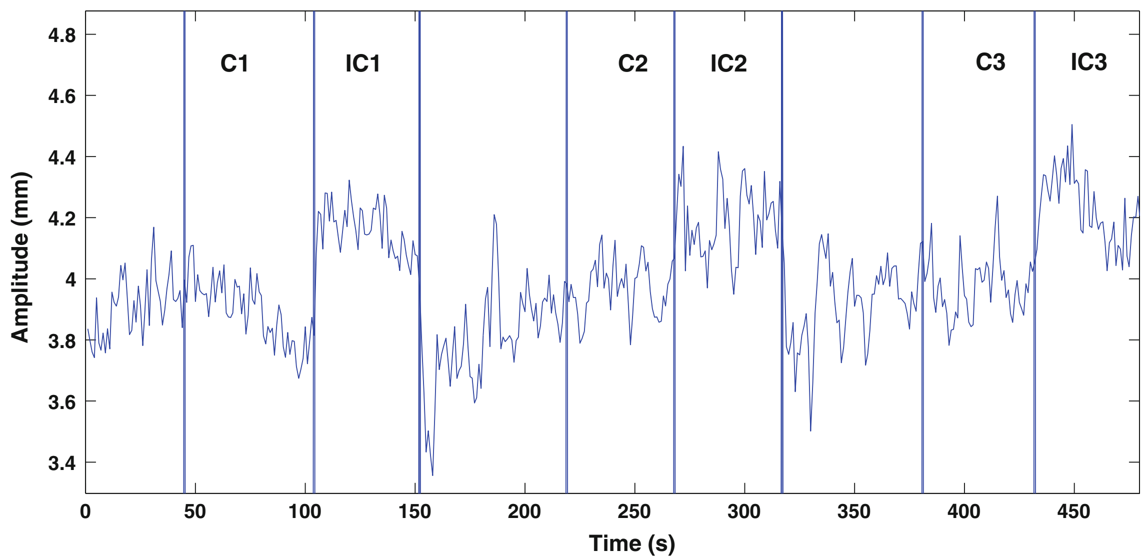

FIGURE 5. The preprocessed raw PD signal (PD_r) in the online algorithm.

experiment, during the first introductory section (see Fig. 2), the mean value of twenty PD_r points is calculated as the $PD_{\text{Reference Value}}$ for the definition of appropriate thresholds. In addition, two thresholds are needed for stress detection, which are respectively regarded as the upper limit of the PD_r signal amplitude fluctuation during the "relaxation" intervals and the lower limit of the PD_r signal amplitude fluctuation during the "stress" intervals. These are denoted as "$\text{Threshold}_{\text{relaxation}}$" and "$\text{Threshold}_{\text{stress}}$" respectively. These two thresholds are calculated based on $PD_{\text{Reference Value}}$, and two constants ($k_1$ and k_2, $k_1 > 1$, $k_2 > 1$, $k_2 > k_1$. E.g., values defined empirically and used were $k_1 = 1.02$, $k_2 = 1.07$) are also required. The equations are shown below:

$$Threshold_{\text{relaxation}} = PD_{\text{Reference Value}} \times k_1 \quad (k_1 > 1), \tag{5}$$

$$Threshold_{\text{stress}} = PD_{\text{Reference Value}} \times k2 \quad (k2 > 1). \tag{6}$$

Sub-step 2 (Feature-based decision voting) Three features are extracted to identify the PD_r signal change. Each feature has a different level of importance. If the current PD_r signal value satisfies the criterion of one feature, then the corresponding weight score ("weight stress" or "weight relaxation") will be updated as indicated below.

The *first criterion* for the first feature detected is the PD_r signal amplitude testing.

$$\text{If } PD_r(n) \geq Threshold_{\text{stress}}\ Weight_{\text{stress}}(n) = Weight_{\text{stress}}(n) + 3, \tag{7}$$

$$\text{If } PD_r \leq Threshold_{\text{relaxation}}\ Weight_{\text{relaxation}}\ (n) = Weight_{\text{relaxation}}\ (n) + 3. \tag{8}$$

This means that the "vote count" for one or the other classification outcomes (stress or relaxation) will be increased by 3 (voting) points according to this criterion.

The *second criterion* for the second feature detected is based on the modified backward difference method. The kernel for traditional backward differentiation of a sequence can be given as [−1 1], which is denoted as K_{Backward}. Figure 6 illustrates the results after the convolution of K_{Backward} with a pair of segments (one congruent segment and one incongruent segment) of one subject's PD_r signal. In Fig. 6, the second vertical line denotes the expected affective transition of the human subject from "relaxation" to "stress"; whereas the third vertical line indicates an expected affective transition of the human subject from "stress" to "relaxation." The short time periods immediately after affective transitions are critical for the performance of the system. However, as we can see from Fig. 6, in these two short time periods, right after the second and the third vertical lines, just a few points show a slightly different pattern compared with the rest of the values. In other words, the traditional first order backward differentiation operator can not effectively identify the sudden change of the $\mathrm{PD_r}$ signal. Therefore, in our study, the modified backward differential kernel $K_{\mathrm{ModifiedBackward}}$ is implemented, which is given as [−1 0 0 0 1]. Figure 7 shows the results after the convolution of $K_{\mathrm{Modified\ Backward}}$ with the same pair of segments of the $\mathrm{PD_r}$ signal as Fig. 6. It is evident from the results that at the beginning of the incongruent segment, there are four consecutive points with large positive values (slightly larger than 0.3), whereas after the incongruent segment, there are six consecutive points with large negative values (slightly less than −0.3). However, in other time intervals, the calculated results are not often consecutively larger than 0.3 or less than −0.3, which shows the noticeable contrast with the relatively large calculated values mentioned above. Therefore, according to the amplitude and the continuity of these calculated values obtained by convolution with $K_{\mathrm{Modified\ Backward}}$, it could be inferred that the $\mathrm{PD_r}$ signal has significant changes during those time intervals, which implies that the computer user's state is varying from "relaxation" to "stress" or from "stress" to "relaxation" at those times. These changes are detected, algorithmically, performing the following calculations.

(1) Calculate the DP_{MB}s. Five DP_{MB} values are based on the current and the previous PD_{r} values, which can consequently represent the overall PD_{r} signal amplitude variation in a short period.

$$DP_{\mathrm{MB}}(i) = 1 \times PD_{\mathrm{r}}(n+1-i) - 1 \times PD_{\mathrm{r}}(n-i-3) \quad (i = 1,2,3,4,5). \tag{9}$$

(2) Calculate the number of DP_{MB} greater than the absolute value of $Threshold_{\mathrm{MB}}$

$$Num1 = 0, \tag{10}$$

$$\text{If } DP_{\mathrm{MB}}(i) > Threshold_{\mathrm{MB}}\ Num1 = Num1 + 1\ (i = 1,2,3,4,5), \tag{11}$$

$$Num2 = 0, \tag{12}$$

$$\text{If } DP_{\mathrm{MB}}(i) < -Threshold_{\mathrm{MB}}\ Num2 = Num2 + 1\ (i = 1,2,3,4,5). \tag{13}$$

(3) Modify the weight for "relaxation" or "stress" affective assessment

$$\text{If } Num1 \geq 3\ Weight_{\mathrm{stress}}(n) = Weight_{\mathrm{stress}}(n) + 1, \tag{14}$$

$$\text{If } Num2 \geq 3\ Weight_{\mathrm{relaxation}}(n) = Weight_{\mathrm{relaxation}}(n) + 1. \tag{15}$$

The *third criterion* is based on shape information detection methods, supported on concepts from mathematical morphology. In mathematical morphological operations, a structuring element is operated on the original signal or image in order to extract the shape information. In a morphological operation, the

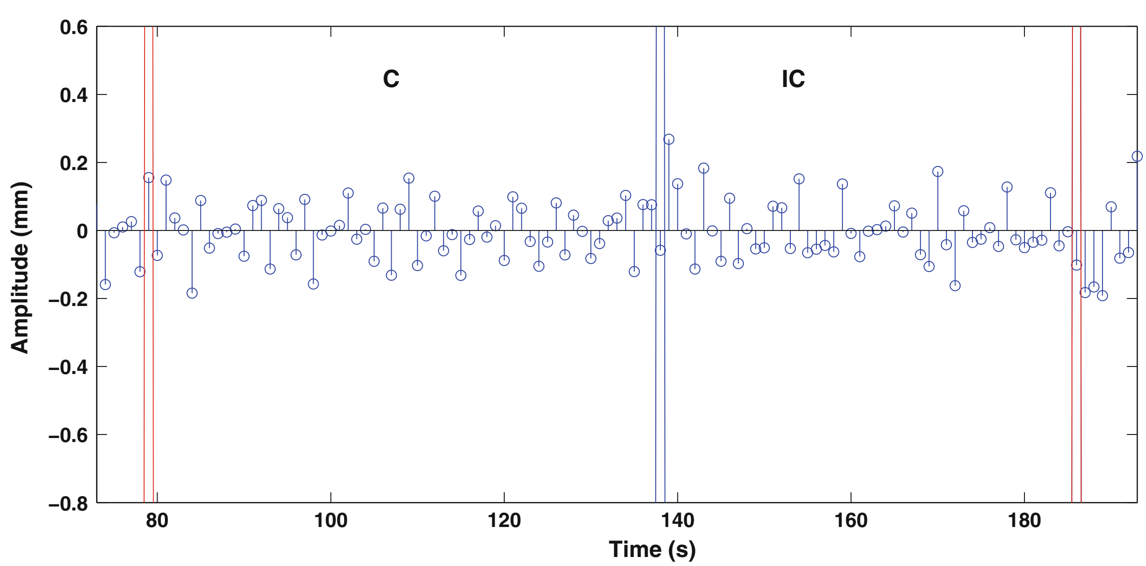

FIGURE 6. The results after the convolution of K_{Backward} with a pair of segments of the $\mathrm{PD_r}$ signal.

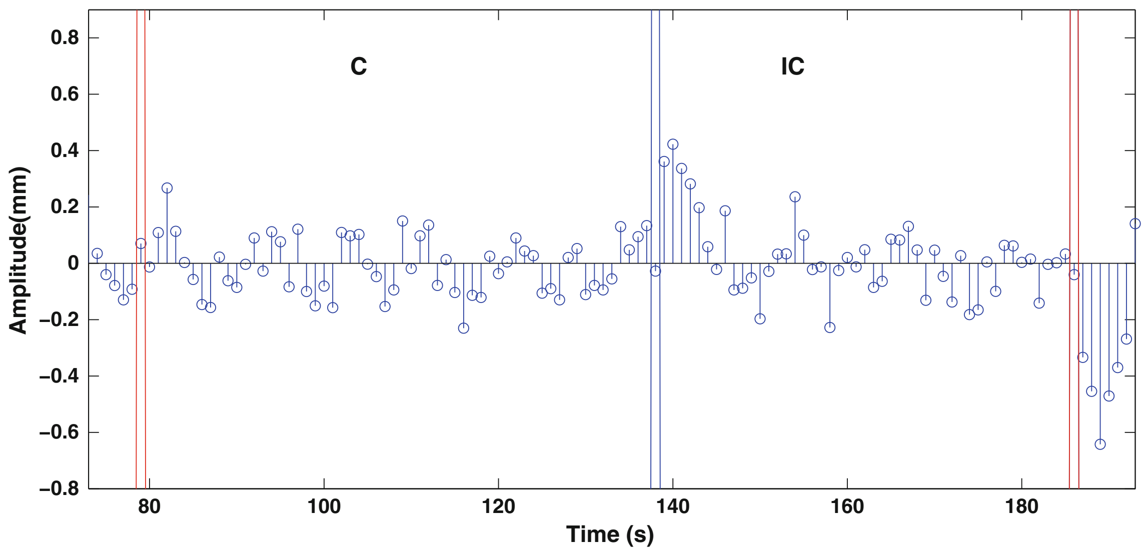

FIGURE 7. The results after the convolution of $K_{\text{Modified Backward}}$ with the same pair of segments of the PD_r signal as in Fig. 6.

value of each point (pixel) in the output signal (image) is based on a logical transformation, which depends on the comparison of the point (pixel) neighborhood in the input signal (image) with a pattern. By choosing the size and the shape of a structuring element, the researcher can construct a different pattern. In addition, there are two types of basic morphological operations, which are erosion and dilation. Erosion is the operation that outputs the maximum value of all the points (pixels) in the neighborhood of the input value (pixel) after the transformation; whereas dilation is the operation that outputs the minimum value of all the points (pixels) in the neighborhood of the input value (pixel) after the transformation.

Figure 8 shows the results of the same pair of segments of the PD_r signal as Fig. 7 after processing with the dilation and erosion operations. As shown in Fig. 8, the solid line, the upper edge of the cluster of PD_r signal points, is the result after modified dilation operation processing; whereas the dashed line, the lower edge of the cluster of PD_r signal points, is the result after modified erosion operation processing. It is apparent from Fig. 8 that when the PD_r signal has significant increases, the stem points are almost consecutively on the solid line; whereas when the PD_r signal has significant decreases, the stems are almost consecutively on the dash line. Therefore, these observations suggest that the location of clusters of stems on the solid line or on the dashed line could indicate a significant amplitude increase or decrease in the PD_r signal. For a certain PD_r signal decision making process, the current PD_r value and its previous neighboring PD_r values are all considered, which can reveal the shape of PD_r signal within a short period of time. The detailed algorithm is proposed below:

(1) Calculate the number of "morphologically matched points"

$$Num_{\text{dilation}} = 0, \tag{16}$$

If $PD_r(n-i) = \max[PD_r(n-i), PD_r(n-i-1), PD_r(n-i-2), PD_r(n-i-3), PD_r(n-i-4)]$ $(i = 1, 2, 3, 4, 5)$, (17)

$$Num_{\text{dilation}} = Num_{\text{dilation}} + 1. \tag{18}$$

Similarly,

$$Num_{\text{erosion}} = 0, \tag{19}$$

If $PD_r(n-i) = \min[PD_r(n-i),\ PD_r(n-i-1),\ PD_r(n-i-2),\ PD_r(n-i-3),\ PD_r(n-i-4)]$ $(i = 1, 2, 3, 4, 5)$, (20)

$$Num_{\text{erosion}} = Num_{\text{erosion}} + 1. \tag{21}$$

(2) Modify the weight for "relaxation" or "stress" affective assessment

$$\text{If } Num_{\text{dilation}} \geq 3 \quad Weight_{\text{stress}}(n) = Weight_{\text{stress}}(n) + 1, \tag{22}$$

$$\text{If } Num_{\text{erosion}} \geq 3 \quad Weight_{\text{relaxation}}(n) = Weight_{\text{relaxation}}(n) + 1. \tag{23}$$

Sub-step 3 ("relaxation"/"stress" indication) The purpose of this sub-step is to provide an indication of the state of stress or relaxation of the subject for each PD_r point.

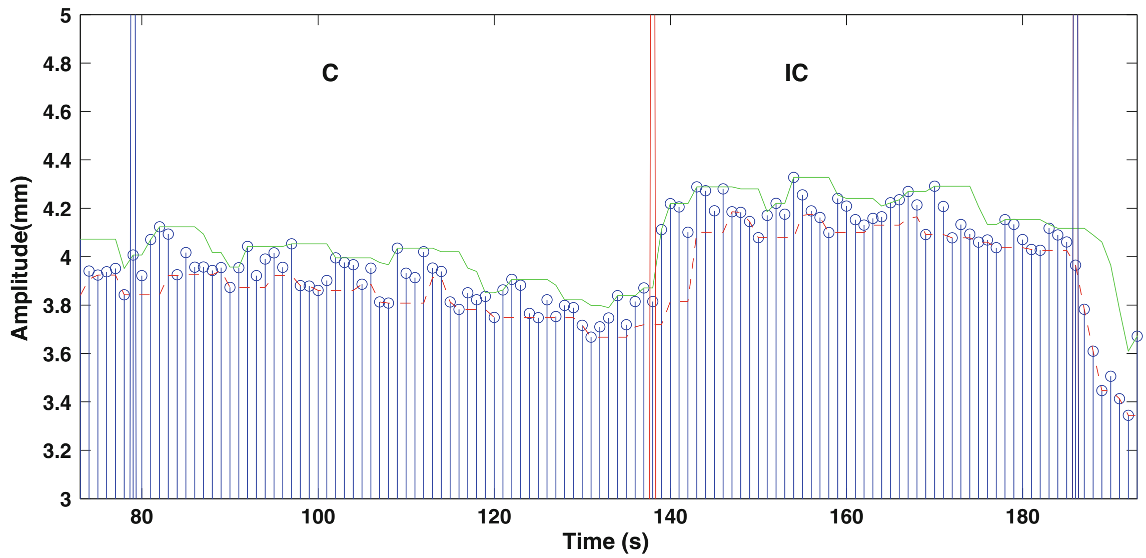

FIGURE 8. The results of PD_r after processing with dilation operation and erosion operation, respectively. The solid line, the upper edge of the cluster of PD_r signal points, is the result after modified dilation operation processing; whereas the dash line, the lower edge of the cluster of PD_r signal points, is the result after modified erosion operation processing.

- If $Weight_{\text{stress}}(n) \geq 4$ and the previous PD_r point is detected as "relaxation"

The current $PD_r(n)$ is the key (upward) transition point, which indicates that the subject is changing from "relaxation" to "stress."

- If $Weight_{\text{relaxation}}(n) \geq 4$ and the previous PD_r point is detected as "stress."

The current $PD_r(n)$ is the key (downward) transition point, which indicates that the subject is changing from "stress" to "relaxation."

- Otherwise.

$PD_r(n)$ is not a transition point, which indicates that the subject is at the same level of "relaxation" or "stress" as in the previous PD_r point.

All the data manipulations described for the on-line physiological signal processing algorithm were implemented in custom Matlab scripts, running in the main processor of a Windows personal computer.

Selection of Relevant Individual Data

While the SCWT was implemented in our experiment for the purpose of eliciting mild mental stress in the human subjects during controlled intervals, we considered that it would be interesting to verify for each subject and for each pair of C and IC segments (denoted C/IC), whether the stimuli were effective in producing a significant stress-related change, as hypothesized. It is known that different individuals may be more or less responsive to the application of stressor stimuli, and we wished to guide the development of our algorithms by using physiological data of instances where the subject was, indeed, stressed. However, a "golden standard" to define when a subject was really stressed is not readily available for an experimental environment like ours, and we had to estimate the success of our stress elicitation method by non-invasive, non-obtrusive methods that would not, themselves, create unplanned stress in our subjects.

Toward this goal, two types of approaches were applied. The first is the selection of relevant C/IC segments pairs based on paired *t*-test analysis. The GSR signal is one of the most commonly used physiological signals for the purpose of gauging some degree of stress affecting a human subject. Therefore, the mean GSR value is calculated in each C segment and in the following IC segment. Then a paired *t*-test is utilized to determine whether these C/IC segment pairs have significant increase change from "congruent" to "incongruent" SCWT presentations.

The second approach relied on the feedback provided by the subjects themselves, who were asked to describe their reactions in a questionnaire. In order to better understand the responses of the human subjects to the SCWT experiment, we conducted an evaluation that posed two questions, quantified with numerical rating scales, to capture the subject's self evaluation immediately after the conclusion of the SCWT experiment. Because emotion is very subjective and only the subject has the epistemic authority to express it,[32] the collection of self-reports from subjects has been widely used in the psychological and affective computing research fields.[33] In our questionnaire, two questions

TABLE 3. Results of stress detection by five types of classification algorithms (selected data based on the paired *t*-test).

Phase of classification	Classification algorithm (%)					Mean (%)	Variance (%)
	*K** (*K*-star)	Multilayer perceptron	Naïve Bayes	Random forest	JRip		
P1:6 features from PD and GSR	80.00	85.71	87.86	87.14	86.43	85.43	9.85
P2:3 features from GSR (no PD)	60.00	63.57	60.71	63.57	63.57	62.29	3.16
P3:3 features from PD (no GSR)	85.71	87.86	88.57	84.29	85.71	86.43	3.06

*K** (*K*-star): This method is a refinement of the *k*-nearest-neighbor rule. It measures the distance between two instances using the entropic theory, based on the probability of transforming one instance into another by randomly choosing between all possible transformations. The probabilities for each category are calculated and the category with the highest probability is considered as the classification of the new instance.

TABLE 4. Results of stress detection by five types of classification algorithms (selected data based on the questionnaire).

Phase of classification	Classification algorithm (%)					Mean (%)	Variance (%)
	*K**	Multilayer perceptron	Naïve Bayes	Random forest	JRip		
P1:6 features from PD and GSR	83.87	86.56	84.95	87.10	83.87	85.27	2.25
P2:3 features from GSR (no PD)	53.23	52.69	47.85	43.31	58.60	51.94	21.19
P3:3 features from PD (no GSR)	87.10	88.17	88.71	86.56	85.48	87.20	1.65

were posed: "How did you feel in the congruent segments?" and "How did you feel in the incongruent segments?" An affective rating scale with five levels (from "1" to "5") was listed under each question. The subjects were also told that "1" represents that the subject feels relaxed. "5" represents that the subject feels very stressed just as if he/she could not find the key of his/her house. "2," "3," and "4" represent increasing, intermediate levels between "1" and "5."

RESULTS

Result of Selection of Relevant Individual Data Based on Paired t-Test

In our study, 42 subjects volunteered to participate in the experiment, so there are 126 pairs of mean GSR values of data segments (3 "C" and 3 "IC" for each subject). The difference of the mean GSR values between each incongruent segment and the preceding congruent segment was calculated and denoted as "DiffGSRmean." The normality of these 126 "DiffGSRmean" values was first examined by the Kolmogorov–Smirnov (K–S) test and Shapiro–Wilk (S–W) test,[7] with significance values of 0.200 and 0.446 respectively. This confirmed that the values of "DiffGSRmean" were normally distributed. In fact, the "DiffGSRmean" population had a mean value of 0.093 with a standard deviation of 0.326. Then the paired *t*-test (2-tailed) was implemented and confirmed that, at a significance value of 0.002 (*t* value: 3.218, df value: 125), the SCWT, generally, produced a change in GSR mean levels, which has been linked to stress elicitation by previous research, as outlined in the "Experiment Setup" section. Therefore, we postulate that our protocol evoked more stress during incongruent segments than during congruent segments. In the distribution of "DiffGSRmean" values, the 90% confidence interval was [0.0453, 0.1414]. In order to remove the C/IC segments pairs, in which the stress elicitation may have been insufficient, the C/IC segment pairs whose "DiffGSRmean" was less than 0.0453 were discarded from subsequent analysis. Hence, a total of 70 pairs of C/IC segments (i.e., 140 segments) was kept (under this selection method) as the data to be used in the rest for the development of the off-line and on-line algorithms.

Result of Selection of Relevant Individual Data Based on Questionnaire

In our study, 42 subjects volunteered as participants and answered the 2 questions outlined in the previous section immediately after the experiment. The average subjective ratings and standard deviations for the C and IC segments were 1.1667 vs. 3.2857 (mean values) and 0.4371 vs. 0.8348 (standard deviations), respectively. It should be noted that, for all the participants, the subjective assessment was higher (subject felt more stressed) during the incongruent than during the congruent segments, by at least 1, in a scale with a maximum of 5. However, in order to remove from consideration C/IC pairs of experiments with potentially insufficient stress elicitation, only data from subjects whose difference evaluation score between incongruent and congruent segments was greater than or equal to two were retained. In this way, only data from 31 subjects (i.e., including 93 congruent and 93 incongruent segments) were kept for off-line and on-line affective processing, under this selection approach.

Stress Detection Based on Off-line Processing

In this section, five different classification algorithms were applied to the features extracted from PD and GSR (see Table 1) in order to assess the "relaxation" vs. "stress" of the human subjects. Specifically, for our work, we utilized the WEKA software, which can be freely downloaded from http://www.cs.waikato.ac.nz/ml/weka/. In addition, three classification phases (P1, P2, and P3) were also performed, as illustrated in Table 2, in order to compare the classification efficiency of PD and GSR for stress detection.

From the selection of data (the pairs of C/IC segments) made on the basis of the paired *t*-test, 70 congruent segments and 70 incongruent segments were kept for analysis. In addition, in order to obtain a more accurate and realistic assessment of the classifiers, a 10-fold cross validation method was used. The accuracy rates from the experiments are shown in Table 3. From the selection of data made on the basis of the questionnaire, 93 congruent segments and 93 incongruent segments were kept for analysis. In addition, a 6-fold cross validation method was used. The accuracy rates from the experiments are shown in Table 4.

Stress Detection Based on On-line Processing

The final accuracies of stress detection based on the on-line processing (for the data sets chosen by the two selection criteria we have outlined) are as follows: for the 70 C/IC segments pairs data selected on the basis of the paired *t*-test, the mean classification rate is 72.30%. For the data from the 31 subjects selected on the basis of the questionnaire or self-evaluation, the mean classification rate is 73.55%. Predictably, the mean classification accuracies from the simpler on-line classification algorithm are below their counterparts for the more elaborate off-line classification algorithm, presented in Table 3. Interestingly, if the on-line classification algorithm is used to process the complete set of congruent and incongruent segments recorded, i.e., without discarding any of the C and IC segments, in an emulation of the unconstrained real-time operation of this on-line system, the average classification accuracy remains about the same, with a value of 72.40%. As an example of how the method should work, Fig. 9 shows the stress detection result of one subject after on-line PD signal processing. It should be noted that the output of the system was only calculated for the congruent and incongruent segments. The accuracy rate for the whole example is 98.69%, and the accuracy values achieved for each C/IC segment pair, they are 100, 95.92, and 100%, respectively.

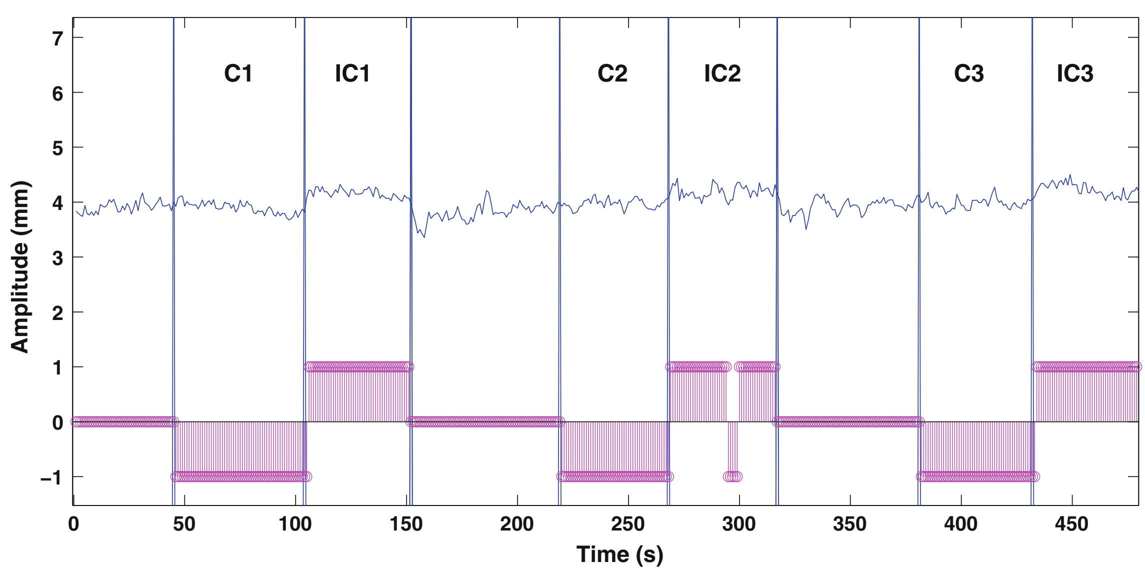

FIGURE 9. The on-line stress detection result for one subject. There are two traces in this figure. The upper trace is the pre-processed PD_r signal. The lower trace is the on-line stress detection result for the whole experiment. The negative outputs (stems) with value −1 denote that at that time, the system indicates that the subject was relaxed; whereas the positive stems with value 1 indicate that the computer user is regarded as stressed at that time.

DISCUSSION

In our experiment, the PD and GSR physiological signals were collected and analyzed in order to investigate their efficiency and robustness for detecting the "relaxation" vs. "stress" of human subjects. We developed off-line and on-line stress detection approaches based on the PD signal.

For our off-line stress detection algorithm, the outcomes are illustrated in Table 3 (for data selected on the basis of the paired *t*-test) and Table 4 (for data selected based on the questionnaire). Both tables show that the three features extracted from the PD signal reach highest accuracy levels of up to 86.43 and 87.20%, and minimum variance of 3.06 and 1.65%, respectively (Phase 3). However, in the first phase of both tables, the six features extracted from both PD and GSR signals achieve only limited classification success, which may imply that the classification capabilities of PD and GSR cannot be combined to achieve a higher detection rate, when using signal processing methods similar to those used in this experiment. The second phase of both tables further support this assumption: the three features extracted from the GSR signals reached the lowest classification averages (62.29 and 51.94%). All these observations suggest that the PD is a robust and efficient physiological signal for stress detection, especially in differentiating "relaxation" vs. "stress."

In the on-line stress detection algorithm, the feature-based decision voting method is applied, which is appropriate for a situation when individual features are brought together in a group to solve a classification problem. According to the idea of synergy, the global decision achieved tends to be more effective and reliable than the decision made based on a single individual feature. In our algorithm, each feature has different significance levels. The consensus can be achieved when the final voting score is greater than a threshold score previously set.

Since our interest is to detect significant changes in the PD_r signal, the traditional backward differentiation operator was implemented first, with the results shown in Fig. 6. However, this approach did not yield particularly strong (positive or negative) outputs at the boundaries of the IC segments, as we sought. Therefore, the modified differentiation operator $[-1\ 0\ 0\ 0\ 1]$ was used. This operator detects the difference between the current value and the value 4 time steps back, which reflects the local tendency of the PD_r signal. If large values with the same sign are detected successively in the results derived from the modified backward differentiation operator, it is believed that the PD_r signal has undergone a significant change in this interval of time. In an attempt to also consider shape change information in the analysis of the PD_r signal, morphological methods were applied as part of the on-line approach we implemented.

On the other hand, the GSR signal was also processed, off-line, for comparison. The GSR signal is a physiological signal widely used to detect emotional variations of the human subjects. However, GSR classification of both sets of selected data yields accuracy levels that are significantly lower than the accuracy of the on-line PD stress detection algorithm (62.29 vs. 72.30%; 51.94 vs. 73.55%).

It should be noted that the experimental protocol included temporary increases in the illumination in the field of view of the subjects during two of the six experimental segments (IC2 and C3 Stroop segments), to emulate the illumination variations that can occur during ordinary computer use. In order to better analyze the effect of illumination and affective changes on PD variation, the difference of the average value of PD signal for each incongruent Segment (IC1, IC2, and IC3) from its corresponding congruent segment (C1, C2, and C3) is calculated. These differences are denoted as "dPD1," "dPD2," and "dPD3," respectively. The average values of "dPD1," "dPD2," and "dPD3" for all 42 subjects were found to be 0.7377, 0.5406, and 0.6050, respectively. The initial reduction trend (from "dPD1" to "dPD2") could be explained in terms of possible habituation of each subject to the C/IC transitions. However, we observe that "dPD3" is larger than "dPD2," deviating from that trend. We propose as an explanation for this that the increased illumination in C3 further promotes the reduction of pupil size expected in a congruent Stroop segment, resulting in a larger differential for the third C/IC segment pair. The illumination increase in IC2, on the other hand, seems to have countered the increase of pupil size expected in an incongruent Stroop segment, yielding a small differential for the second C/IC segment pair.

Finally, the results of the "relaxation" vs. "stress" differentiation based on PD signal processing (around 86–87% for off-line analysis; around 72–73% for on-line analysis), suggest that the methods developed to process the PD signal are promising and the PD signal is a valuable alternative for detecting "relaxation" vs. "stress" in comparison with the traditionally used GSR signal.

The results of this study, as well as the discussion offered in this section, reinforce the original speculation that emerged when our group reported the possibility of detecting affective changes through simultaneous monitoring of multiple physiological signals. That early work, however placed emphasis on the potential benefits of concurrent analysis of the multiple signals, whereas the results we present in Table 3 indicate that the PD signal may be a major contributor to the success of the classification performed on the basis of multiple signals, as processing with PD-derived features performed as well as a system that considers features of both PD and GSR. Two additional new steps forward from the previous work presented by Zhai and Barreto[39] are the experimentation with variable environmental illumination (Zhai and Barreto enforced a constant environmental lighting during those experiments), and the development of the on-line PD processing algorithm included in this report, as the previous work had only been develop for off-line application. Availability of an on-line algorithm, capable of operation without requiring knowledge of past and future samples of the signals, and simple enough for viable real-time implementation in a DSP board is a key consideration for the prospective use of PD-based affective sensing in standard computing systems.

ACKNOWLEDGMENTS

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

REFERENCES

[1]Andreassi, J. L. Psychophysiology: Human Behavior & Physiological Response. Mahwah, NJ: Lawrence Erlbaum Associates, 2007.
[2]Andren, J., and P. Funk. A case-based approach using behavioural biometrics to determine a user's stress level. In: ICCBR Workshops, 2005, pp. 9–17.
[3]Beatty, J., and B. Lucero-Wagoner. Handbook of Psychophysiology. Cambridge: Cambridge University Press, 2000.
[4]Begum, S., M. U. Ahmed, P. Funk, and N. Xiong. Using calibration and fuzzification of cases for improved diagnosis and treatment of stress. In: 8th European Conference on Case-Based Reasoning Workshop Proceedings, edited by M. Minor, 2006, pp. 113–122.
[5]Cano-Vindel, A., J. J. Miguel-Tobal, H. Gonzalez-Ordi, and I. Iruarrizaga-Diez. Hyperventilation and anxiety experience. *Anxiety Stress* 13(2–3):291–302, 2007.
[6]Fellous, M., and M. A. Arbib. Who Needs Emotions? The Brain Meets the Robot. New York: Oxford University Press, 2005.
[7]Field, A. Discovering Statistics Using SPSS (3rd ed.). New York: Sage, 2009.
[8]Gao, Y., A. Barreto, and M. Adjouadi. Monitoring and Processing of the Pupil Diameter Signal for Affective Assessment of a Computer User. Lecture Notes in Computer Science (LNCS). LNCS 5610, 2009, pp. 49–58.
[9]Gao, Y., A. Barreto, and M. Adjouadi. Affective assessment of a computer user through the processing of the pupil diameter signal. In: Innovations in Computing Sciences and Software Engineering, edited by T. Sobh, and K. Elleithy. New York: Springer, 2010, pp. 189–194.
[10]Granholm, E., and S. R. Steinhauer. Introduction: Pupillometric measures of cognitive and emotional processing. *Int. J. Psychophysiol.* 52:1–6, 2004.
[11]Grewal, M. S., and A. P. Andrews. Kalman Filtering: Theory and Practice Using MATLAB (3rd ed.). Hoboken, NJ: Wiley-IEEE Press, 2008.
[12]Healey, J. Wearable and Automotive Systems for Affect Recognition from Physiology. Ph.D. dissertation, MIT Media Lab, 2000.
[13]Healey, J. A., and R. W. Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intell. Transp. Syst.* 6(2):156–166, 2005.
[14]Hjemdahl, P., U. Freyschuss, A. Juhlin-Dannfelt, and B. Linde. Differentiated sympathetic activation during mental stress evoked by the Stroop test. *Acta Physiol. Scand. Suppl.* 527:25–29, 1984.
[15]Holmqvist, K., M. Nystrom, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Weijer. Eye Tracking: A Comprehensive Guide to Methods and Measures. Oxford: Oxford University Press, 2011.
[16]Klingner, J. The pupillometric precision of a remote video eye tracker. In: Proceedings of the 2010 Symposium on Eye-Tracking Research Applications (ETRA '10). New York: ACM, 2010, pp. 259–262.
[17]Lim, C. L., C. Rennie, R. J. Barry, H. Bahramali, I. Lazzaro, and B. Manor. Decomposing skin conductance into tonic and phasic components. *Int. J. Psychophysiol.* 24(2):97–109, 1997.
[18]Martini, F. H., and J. L. Nath. Fundamentals of Anatomy & Physiology (8th ed.). San Francisco: Benjamin Cummings, 2008.
[19]Morgante, J. D., R. Zolfaghari, and S. P. Johnson. A critical test of temporal and spatial accuracy of the Tobii T60XL eye tracker. *Infancy* 17:9–32, 2012.
[20]Partala, T., and V. Surakka. Pupil size variation as an indication of affective processing. *Int. J. Hum.–Comput. Stud.* 59:185–198, 2003.
[21]Picard, W., and J. A. Healey. Wearable and automotive systems for affect recognition from physiology. Technical report, MIT, 2000.
[22]Raymond, J. C. Dictionary of Psychology. New York: Routledge, 1999.
[23]Shi, Y., N. Ruiz, R. Taib, E. Choi, and F. Chen. Galvanic skin response as an index of cognitive load. In: Proceeding of Computer–Human Interaction conference on Human Factors in Computing System, 2007, pp. 2651–2656.
[24]Siegle, G. J., S. R. Steinhauer, and M. E. Thase. Pupillary assessment and computational modeling of early and sustained processing on the Stroop task in depression. *Int. J. Psychophysiol.* 52:63–76, 2004.
[25]Sierra, A. S., C. S. Ávila, J. G. Casanova, and G. B. Pozo. A stress-detection system based on physiological signals and fuzzy logic. *IEEE Trans. Ind. Electron.* 58(10):4857–4865, 2011.
[26]Sierra, A. S., C. S. Ávila, A. Mendaza-Ormaza, and J. G. Casanova. An approach to hand biometrics in mobile devices. *SIViP* 5(4):469–475, 2011.
[27]Smith, S. W. The Scientist and Engineer's Guide to Digital Signal Processing. San Diego, CA: California Technical Publishing, 1997.
[28]Steinhauer, S. R., G. J. Siegle, R. Condray, and M. Pless. Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *Int. J. Psychophysiol.* 52:77–86, 2004.
[29]Stern, R. M., W. J. Ray, and K. S. Quigley. Psychophysiological Recording. Oxford: Oxford University Press, 2000.

[30]Stroop, J. R. Studies of interference in serial verbal reactions. *J. Exp. Psychol*. 18:643–662, 1935.

[31]Sun, F., C. Kuo, H. Cheng, S. Buthpitiya, P. Collins, and M. L. Griss. Activity-Aware Mental Stress Detection Using Physiological Sensors. Silicon Valley Campus, Paper 23, 2010.

[32]Tognetti, S., M. Garbarino, M. Matteucci, and A. Bonarini. The affective triad: stimuli, questionnaires, and measurements. In: Proceeding of ACII'11, Vol. II, 2011, pp. 101–110.

[33]Tourangeau, R., L. J. Rips, and K. A. Rasinski. The Psychology of Survey Response. Cambridge: Cambridge University Press, 2000.

[34]Tsai, S. S. Power Transformer Partial Discharge (PD) Acoustic Signal Detection Using Fiber Sensors and Wavelet Analysis, Modeling, and Simulation. Master's Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, 2002.

[35]Tulen, J., H. P. Moleman, H. G. Steenis, and F. Boomsma. Characterization of stress reactions to the Stroop color word test. *Pharmacol. Biochem. Behav*. 32(1):9–15, 1989.

[36]University of California-Irvine. Short-term stress can affect learning and memory. *ScienceDaily*. March, 11, 2008.

[37]Verney, S. P., E. Granholm, and S. P. Marshall. Pupillary responses on the visual backward masking task reflect general cognitive ability. *Int. J. Psychophysiol*. 52:23–36, 2004.

[38]Weigle, C., and D. C. Banks. Analysis of eye-tracking experiments performed on a Tobii T60. In: Proceedings of the SPIE6809, Visualization and Data Analysis, January, 2008.

[39]Zhai, J., and A. Barreto. Stress detection in computer users through noninvasive monitoring of physiological signals. *Biomed. Sci. Instrum*. 42:495–500, 2006.