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

Peng Ren, Armando Barreto, Jian Huang, Ying Gao, Francisco R. Ortega & Malek Adjouadi

ABSTRACT

Physiological signals, which are controlled by the autonomic nervous system (ANS), could detect the affective state of the human beings and therefore provide the application in medicine and engineering. The Pupil Diameter (PD) seems to be an indication of the affective state, which has been found by the previous research but has not been investigate fully yet. In this study, new approaches based on monitoring and processing the PD signal for offline and online affective assessment (“relaxation” vs. “stress”) are proposed. For the offline affective assessment, wavelet denoising and Kalman filtering are implemented first in order to remove the noise of the PD signal. Then normalization and feature extraction are utilized. In order to select the more relevant and convinced data for further analysis, two types of data selection methods are applied, which are based on paired t-test and questionnaire, achieving the average accuracy up to 86.43% and 87.20% respectively. For the online affective assessment, the moving average window method is implemented first in order to remove the random noise and obtain the representative value for every one second. There are three main steps for online affective assessment algorithm, which are preparation, decision-based feature voting and affective determination three phases. The final results show that the accuracies are 72.30% and 73.55% respectively for two types of data selection methods. In order to further analyze the efficiency of affective sensing through the PD signal, the Galvanic Skin Response (GSR) was also monitored and processed, whose highest affective assessment classification rate is only 63.57% (based on offline processing algorithm). The overall results confirm that PD signal should be considered as one of the most powerful physiological signals to involve in the future automated real time affective recognition systems, especially for detecting the “relaxation” vs. “stress” states.

Key word
Autonomic Nervous System (ANS), Pupil Diameter (PD), Galvanic Skin Response (GSR), Wavelet denoising, Kalman filtering, Moving Average Window, Backward Differentiation, Mathematical Morphology
1 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, and many aspects of emotional behavior [1]. Two branches of the ANS are the Sympathetic Nervous System (SNS) and Parasympathetic Nervous System (PSNS). The SNS activities that occur together might include increased heart rate, blood pressure, sweating, cardiac output and respiration changes that enhance a flight-or-flight reaction [1], [2], [3]. 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. The functions caused by the stimulation of PSNS may slow the heart rate, promote peristalsis, and increase salivary secretions and so on. Although the SNS and the PSNS have contrasting functions, the activities are integrated and not antagonistic [1], [2], [3].

The affective state of the human beings ("relaxation" vs. "stress"), which have been studied by the previous researchers [4], [5], [6], [7], [8], [9], [10], can be regarded as the best two kinds of the emotions representing the actions of the SNS vs. PSNS. In addition, 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 [11]. Furthermore, automated stress detection may enhance the social security (for example, bank account access or border crossing frontiers in airports) or help the criminal investigation progression (for example, polygraph based on the stress experienced by someone who lies [12]). Therefore, it is sensible and reasonable to investigate and accurately detect the "relaxation" vs. "stress" affective state of the human beings.

Until nowadays many physiological signals, which can reflect the ANS reactions [13], [14], [15], have been used to assess the affective state ("relaxation" vs. "stress") of the human beings such as Electroencephalogram (EEG), the Electrocardiogram (ECG), Blood Pressure (BP), Blood Volume Pulse (BVP), Skin Temperature (ST), Galvanic Skin Response (GSR), etc. More recently, pupil size was verified by Partala and Surakka as an evident indication of affective state in their auditory emotional stimulation experiment [16]. However, the efficiency and the robustness of the Pupil Diameter (PD) variations for affective sensing ("relaxation" vs. "stress") have not been fully explored yet.

The pupil, which can constrict to 1.5mm in diameter or dilate to about 8 to 9 mm, is the opening through which light enters the eye and begins the process of visual perception [17], [18]. 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 [18]. More specifically, neurons of the PNS innervate circular fibers of the iris, causing pupillary constriction, whereas excitation by SNS neurons causes the radial fibers of the iris to produce dilation of the pupil [19]. Therefore, if the sympathetic division of the ANS is activated (e.g., due to stress), the size of pupil diameter tends to increase; whereas if the parasympathetic division of the ANS dominates (e.g., during relaxation), the pupil diameter will remain small [19].

In our study, the PD signal, obtained with the eye gaze tracking system, was chosen to be monitored for affective sensing in a non-invasive way. Nowadays eye gaze tracking system has become a robust and intuitive tool for human- computer interaction and continues to extend its accuracy and affordability [20], [21]. 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 [22], [23], [24], [25]. Therefore GSR signal was also measured and analyzed in order to compare the efficiency of our proposed offline and online approaches for the affective assessment ("relaxation" vs. "stress") of a human subject based on the PD signal.

2 EXPERIMENT SETUP

In our experiment, PD and GSR signals were monitored in order to assess the affective state change ("relaxation" vs. "stress") of a human subject. In addition, the intensity of light is the other primary factor affecting the pupillary constriction and dilation, therefore the Illumination Intensity (IL) in the environment was also measured and recorded.
2.1 Software Setup

The “Stroop Color-Word Interference Test” (SCWT) [26] was used to elicit mild mental stress in the experimental subjects during controlled intervals with the aim of recognizing the “relaxation” and “stress” affective state of the subject based on identifying PD and GSR signal variation. Until nowadays, the SCWT has been applied by several research groups as the way to elicit stress emotion for experimental subjects. Tulen et al. [27] used the SCWT to study the stress-induced sympathetic effects, based on physiological, psychological, and biochemical responses. They demonstrated that “the SCWT can induce increases in plasma and urinary adrenaline, heart rate, respiration rate, electrophysiological activity, electromyography, feelings of anxiety, and decreased finger pulse amplitude”. In addition, Hjemdahl et al. [28] studied the sympatho-adrenal and hemodynamic responses to mental stress elicited by the SCWT and demonstrated that the SCWT can increase heart rate and blood pressures by 28 beats/min and 29/14 mmHg for the subjects, on average. Furthermore, the SCWT was utilized as the psychological or cognitive stressor to introduce an emotional response by Feng-Tso Sun et al. [29], and simultaneously they measured the GSR and ECG data to assess the “stress” affective state of the subject.

In the test, a word with the font color that may (“Congruent”) or not (“Incongruent”) match its meaning was presented to the subject, which is shown in Fig.1. The subjects needed to read the presented word aloud first and then were required to click one of the five screen buttons to indicate the font color of the word within three seconds, otherwise the system will automatically display the next word.

Fig. 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. Because each section was purposely designed to provide the opportunity to observe a transition from “relaxation” to “stress”, we could determine whether our approaches (include offline and online methods) are able to signal the emergence of stress.

![Fig. 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).](image)

![Figure 2: Stimuli schedule of the experimental protocol](image)

The segments in each section are:
- ‘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) [30];
‘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 state, not having to perform any action, during one minute. The physiological signals are also monitored.

The aim of our study is to analyze GSR and PD signals and their changes from a state of “relaxation” to a state of “stress”, therefore it is 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), which are shown in Fig.2. These sinusoidal bursts, recorded on two additional channel of the data acquisition system, served as time-stamps to the physiological signals recorded at those critical instants. The luminance intensity remains constant except in the segments IC2 and C3, where the illumination will be temporarily increased (marked as “VI” in Fig. 2).

2.2 Hardware Setup

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. The TOBII system measured the pupil diameter values 60 times per second. The purpose of our study is to offline and online detect the affective state change (from “relaxation” to “stress”) of the subject based on physiological signals. For offline study use, 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, which, in turn, can be read into MATLAB®. At the same time, the average PD value of both eyes was simultaneously filed into MATLAB® for online study use.

While performing the Stroop Test, the subject has the GSR sensor (GSR 2, from Thought Technology) attached to his/her left hand and the IL sensor (BS500B0F photo-diode, from Sharp) on his/her forehead, above the eyes. All these two signals, together with the left and right audio output (to provide the corresponding time stamping in the experiment) are recorded and converted to a MATLAB®-readable data file directly at rate 360 samples/second, using a multi-channel MCC DAQ system (PCI-DAS6023 board). Later for offline study use, these data were downsampled by six, to share the same sampling rate at 60 samples/second for all measured signals.

2.3 Experimental Procedure

In our experiment, 42 individuals, with ages ranging from 20 to 50 (Mean: 27.5; Std. dev.: 5.14), from diverse professional and ethnic backgrounds.

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 minutes. All the normal lights in the room were kept ON, but an additional level of illumination provided by a desk lamp placed above the eye level of the subject was switched ON during the IC2 and C3 segments (as shown in Fig. 2) in order to judge the robustness of our affective assessment algorithm with the interference of the illumination intensity variation in the environment. In the IC2 segment the increase of the light intensity would inhibit the dilation of the pupil diameter caused by stress, whereas in the C3 segment the increase of the light intensity would further promote the contraction of the pupil diameter caused by the relaxed affective state.

3 METHODS

3.1 Development of Offline Physiological Signal Processing Algorithm
3.1.1 Physiological Signal Preprocessing
In our study, the combination implementation of wavelet denoising and Kalman filtering is the main signal processing method applied to PD data. However, on account of the fact that the presence of the eye blinks is an inevitable factor appearing in the raw PD signal, the eye blinks must be removed before further signal processing. In our TOBII system, eye blinks are signaled by a value of “4” in the validity code and identified as sudden transitions to a false PD value of zero, therefore linear interpolation was implemented first to compensate the interruption in the PD data by each blink. Figure 3 shows a set of signals (include the raw PD signal, the IL signal and the raw GSR signal) recorded from one subject during the experiment. The vertical lines are the segment transition boundaries, whose most important three ones respectively separate each congruent Stroop segment (C) from the incongruent Stroop segment (IC) that follows. The first plot of Fig. 4 shows the PD signal after blink & artifact removal, which is called the original PD signal in this paper. It should be noted that though 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. In addition, previous research [31] has also indicated that these fast variations in the PD signal could be due to quantization noise in the pupil diameter measurement.

Wavelet denoising seems to be a proper approach to remove the abrupt changes in the original PD signal. There are main three sub-steps for wavelet denoising [32], [33]. The discrete wavelet transform (DWT) is first implemented to the original PD signal, which includes the fast variability seemed as the noise, to produce the noisy wavelet coefficients (approximation coefficients and detail coefficients) to a level where we can properly separate most of the noise. The multiresolution analysis (MRA) is designed for the practical use of DWT with the dilated and translated version of the mother wavelet. In MRA, scaling function is to create a series of approximations of the signal and wavelet is 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 downsampling operation to result in levels of approximation and detail coefficients. In our study, Daubechies wavelet and nine decomposition levels were applied to the original PD signal. The second sub-step is to select proper threshold method to best remove the noise by altering the values of 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 specific steps for the Birge-Massart strategy are described below [34]:

Let j be the decomposition level, M be the length of coarsest approximation coefficient over 2 and $\alpha$ be a real parameter greater than 1.

- At level j+1 (and coarser levels), everything is kept.
- For level i from 1 to j, the n(i) largest coefficients are kept with $n(i)=M/(j+2-i)^{\alpha}$

The last sub-step is to apply the inverse discrete wavelet transform on the approximation coefficients and the altered detail coefficients to obtain the denoised signal, whose result is shown in the second plot of fig. 4.

Although the original PD signal after wavelet denoising has significant improvement in clearing the signal, there are still some abrupt changes remained, which indicates the further signal processing approach should be applied. Kalman filtering is utilized as the second step to remove the noise that remained in the PD signal. The Kalman filtering, which is a recursive data processing algorithm, generates optimal estimate of desired quantities given the set of measurements. The algorithm works in two sub-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 sub-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 [35], [36]. The PD signal after Kalman filtering is illustrated in the lower plot of fig. 4, which can exemplify the ability of this method to be an effective approach to remove the remaining noise of the PD signal.
3.1.2 Data Normalization and Feature Extraction

The purpose of the work in this part is to do the preparation for performing the classification of the segments in the experiment. The PD signal is normalized to the range [-1, 1] first because different subject has different PD signal baseline. Then the features are extracted from the filtered PD signal, which is shown in the lower plot of fig. 4. It is obviously from the plot that the PD signal rise steeply at the beginning of each incongruent segments, which are considered to elicit the “stress” emotion for the subjects; whereas the PD signal has almost no significant increase at the beginning of each congruent segment. Therefore, it is feasible to take this characteristic to differentiate the “relaxation” vs. “stress” affective states of the human subjects, which the Walsh transform is applied in our study to achieve.
The 1D Walsh transform function implemented is defined as

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

(1)

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 our study, eight consecutive windows, with 100 samples each, from the beginning of each segment (both C and IC), are analyzed. A sequence, which is formed by the eight mean values of these windows, is used to represent the trend of the PD signal during the beginning of the C and IC segments. The Walsh transform decomposes the original signal into a set of orthogonal functions, and encodes the decomposition in the resulting Walsh coefficients. The first and the last several Walsh coefficients can respectively indicate the “low frequent” and the “high frequent” components of the PD signal to some extent. In our study, only the overall trend rather than the detailed information of each segment is needed, therefore the “lower frequency domain variability” is relevant. “PDWalsh”, the difference between the first and the second Walsh coefficient during the onset period of each Stroop segment, is extracted as one of the features from the filtered PD signal. The example is illustrated in Fig. 5 and the values of the “PDWalsh” feature for three Congruent Stroop segment are 0.2730, 0.4233 and 0.2956 respectively, whereas for the three Incongruent Stroop segments are -0.5070, -0.3951 and -0.8192.

In our study, the features obtained from GSR signal are also analyzed in order to evaluate the affective assessment performance of the features derived from the PD signal. There are three characteristic types of skin conductance variations (tonic, spontaneous and phasic) according to the physiological studies. However, only the phasic skin conductance responses are regarded as a noticeable reaction to the emotional stimuli, which can be quantified by parameters such as amplitude, latency, rising time, and half recovery time [37]. Table 1 shows three features respectively extracted from the PD and GSR Signals. Table 2 illustrates three different classification phases for the purpose of evaluating the efficiency of the PD signal with the performance of the GSR signal.
### Table 1: Features Obtained from the PD and GSR Signals

<table>
<thead>
<tr>
<th>Signal</th>
<th>Features</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>PDmean</td>
<td>Average value of the PD signal in a segment</td>
</tr>
<tr>
<td></td>
<td>PDmax</td>
<td>Maximum value of the PD signal in a segment</td>
</tr>
<tr>
<td></td>
<td>PDWalsh</td>
<td>Difference value between the first and the second Walsh coefficient based on</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the PD signal during the onset of each Stroop segment</td>
</tr>
<tr>
<td>GSR</td>
<td>GSRmean</td>
<td>Mean value of the amplitude of each GSR response in a segment</td>
</tr>
<tr>
<td></td>
<td>GSRrisingTime</td>
<td>(Average) Rising time of each GSR response in a segment</td>
</tr>
<tr>
<td></td>
<td>GSRnum</td>
<td>Number of the GSR responses in a segment</td>
</tr>
</tbody>
</table>

### Table 2: Classification Phases (Different Conditions)

<table>
<thead>
<tr>
<th>Phase</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>Using all features extracted from the monitored PD and GSR signals (all 6 features are used for classification)</td>
</tr>
<tr>
<td>P2</td>
<td>Excluding the feature extracted from the PD signal (only 3 features from GSR)</td>
</tr>
<tr>
<td>P3</td>
<td>Excluding the feature extracted from the GSR signal (only 3 features from PD)</td>
</tr>
</tbody>
</table>

### 3.2 Development of Online Physiological Signal Processing Algorithm

#### 3.2.1 Physiological Signal Preprocessing

The purpose of this section is to derive the representative point for every short time interval of the preprocessed raw PD signal, labeled as PD_r, for further online affective assessment. In our online processing algorithm, the hard threshold setting for eye blink removal, which is denoted as “Threshold_{blink}”, and the moving average window method are applied in order to obtain the PD_r. The moving average is the most common method for reducing noise in digital signal processing (DSP) filed, because it is not only simple but also optimal for retaining a sharp step response [38], which is suitable for the shape of PD signal change from Congruent (C) to Incongruent segment (IC) (see the first plot of Fig. 3). In addition, the sampling rate of the raw PD signal is 60 Hz. It is not necessary and practical to classify every PD signal sample for affective assessment. Therefore, 60 samples, namely 1 second duration, were selected as window length to achieve one PD_r value per interval. However, because of the presence of eye blinks, hard threshold must be set, which is denoted as “Threshold_{blink}”. If the computer detects the value of current PD sample is smaller than the “Threshold_{blink}”, the present PD sampling data is not going to be counted as the valid value in the current time interval. Otherwise, the present PD sampling data will contribute to achieving one PD_r value in the current time interval. The equations are illustrated below:

\[
m = 0
\]

\[\text{If } \text{PD}(i) \geq \text{Threshold}_\text{blink} \quad X(i) = \text{PD}(i); \quad m = m + 1; \quad (60*n-59 \leq i \leq 60*n \quad n=1,2,...)\]

\[\text{If } \text{PD}(i) < \text{Threshold}_\text{blink} \quad X(i) = 0; \quad (60*n-59 \leq i \leq 60*n \quad n=1,2,...)\]

\[
\text{PD}_r(n) = \frac{X(60*n-59) + X(60*n-58) + ... + X(60*n)}{m}
\]

Fig. 6 shows the preprocessed PD signal after hard threshold setting and the moving average window methods. It has been shown in Fig. 6 that the length of the original PD signal is about 28800 samples, whereas the preprocessed signal is just about 480 data points, which are derived from the 60 samples length sliding window for averaging. In addition, it is apparent that the new derived 480 data points can considerably represent the variations of the original PD signal while simultaneously removing the noise to some extent.
3.2.2 Online Affective Assessment Algorithm

After capturing the subject’s PD signal and averaging valid numbers of interval samples to achieving PD, value per second, the next step is the identification of affective state of the human subject. The most important part of waveforms in the PD signal is the large sudden increase and the large sudden decrease variation, which appears in the affective state transition from the “relaxation” to the “stress” and from the “stress” to the “relaxation” respectively. In addition, unlike some other emotion related physiological signals, such as ECG and BVP [39], it is almost impossible to examine the difference between the frequency domain of “relaxation” and the “stress” affective state for the PD signal. Therefore, the accurate detection of the large PD signal increase or the large decrease variation, which implies the subject’s affective state transition, seems like the most encouraging and feasible way for the realization of online PD based affective recognition system.

3.2.2.1 Modified Backward Differential Method for Sudden PD Signal Change Detection

A finite difference is a mathematical expression of the form \( f(x + b) - f(x + a) \), which has three different types of forms (forward, backward and central difference respectively). If a finite difference is divided by \( b - a \), one gets a difference quotient [40]. A forward difference is an expression of the form:

\[
\Delta f(x) = \frac{f(x + h) - f(x)}{h}
\]  

(6)

A backward difference uses the values at \( x \) and \( x-h \), instead of the values at \( x+h \) and \( x \):

\[
\nabla f(x) = f(x) - f(x - h)
\]  

(7)

Finally, the central difference is presented by:

\[
\delta f(x) = f(x + h) - f(x - h)
\]  

(8)

\( h \) may be variable or constant, which is depended on the specific application. The finite difference methods are usually applied to numerically approximate the differential equations at the given equally spaced discrete points because we can not make assumptions about the differential properties for some problems such as digital signal or digital image processing. The first order differential expression \( \frac{df(x)}{dx} \) for a function \( f(x) = \rho \) can be approximately presented by the forward, the backward or the central finite difference equivalence, which are shown respectively as follows:

\[
\frac{\hat{f}(x)}{\hat{x}} \approx \frac{\rho_{i+1} - \rho_i}{\Delta x}
\]  

(9)
Differential operator has been successfully used in the digital signal and image processing areas with the aim of detecting the significant changes. For example, N. Paivinen etc. utilized first and second derivative of EEG signal to extract time domain features for automatic seizure detection [41]. S. Suppappol etc. [42] and M. Paoletti etc. [43] achieved real-time QRS complex detection methods based on first derivative since this approach not only requires few computations but also can avoid a delay in the detection. In addition, the Sobel operator is a classic first order edge detection operator that finds contrast by a process of approximation of the gradient of the digital image intensity function [44].

In our study, we attempt to implement the algorithm on the online affective recognition system, so it is evident that only the present and the previous PD signal data can be used. Therefore, only equation (10), which is called backward differentiation, is feasible. In this case, the numerator of the differential is defined as the difference between the actual value $\rho_i$ and its previous neighbor $\rho_{i-1}$. When proceeding from a low value of $\rho_{i-1}$ to a high value, the $\nabla f_i$ operator $(\rho_i - \rho_{i-1})$ takes a high positive value; whereas if $\rho_{i-1}$ is high and $\rho_i$ is low, the $\nabla f_i$ operator $(\rho_i - \rho_{i-1})$ takes a high negative value. Therefore, we could infer the direction and the extent of the sudden change of the PD signal according to the result calculated from (10), which can further indicate the affective state change of the human subject. In digital signal or image processing field, we usually set $\Delta x$ equal to 1, which denotes the distance between neighboring pixels or samples. Therefore, the equation (10) can be written as

$$\frac{\partial f(x)}{\partial x} \approx \frac{\rho_i - \rho_{i-1}}{\Delta x}$$

A kernel for backward differentiation in the x-direction can be given as:

$$K_{Backward} = [-1 \ 1]$$

Fig. 7 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 signal. In our study, we attempt to arouse “stress” affective state of the human subject during the Incongruent segment (IC) while keep the “relaxation” during the Congruent segment (C). It is conceivable that during the subject’s affective state variation period, there should be significant PD signal value changes (increase or decrease). In Fig. 7, the second vertical line denotes that the affective state of the human subject is from the “relaxation” to the “stress”; whereas the third vertical line indicates that the affective state of the human subject is from the “stress” to the “relaxation”. Therefore, the short time periods after these two vertical lines are the most important moments needed to be noticed. As we can see from Fig.7, in these two short time periods, which are the duration after the second and the third vertical lines, just a few points show a little difference compared with other time periods. In other words, the traditional first order backward differentiation operator can not effectively identify the sudden change of the PD signal.
In order to better detect the significant change of the PD signal at the beginning of (from “relaxation” to “stress”) and after (from “stress” to “relaxation”) the Incongruent segment, the modified backward differential method is implemented, which is given by

$$\frac{\delta f(x)}{\delta x}_{\text{Modified}} \approx \frac{\left(\rho_i - \rho_{i-1}\right) + \left(\rho_{i-2} - \rho_{i-3}\right) + \left(\rho_{i-3} - \rho_{i-4}\right)}{4\Delta x}$$

(14)

where $\rho_i$ is the present PD signal value; $\rho_{i-1}$, $\rho_{i-2}$, $\rho_{i-3}$ and $\rho_{i-4}$ are the four neighboring previous PD signal values. In the strict sense, equation (14) can not accurately present the mathematical defined derivative but estimate the overall tendency of the change of the neighboring points. A kernel for the modified backward differentiation in the x-direction can be given as:

$$K_{\text{Modified Backward}} = [-1 \ 0 \ 0 \ 0 \ 1]$$

(15)

Fig. 8 shows the results after the convolution of $K_{\text{Modified Backward}}$ with the same pair of segments (one Congruent segment and one following Incongruent segment) of the PD signal as Fig. 7. It is evident from the results that at the beginning of the Incongruent segment (a short time period after the second vertical line in Fig. 8), there are four consecutive points with large positive values (nearly or larger than 0.3); after the Incongruent segment (a short time period after the third vertical line in Fig. 8), there are six consecutive points with large negative values (nearly or less than -0.3). However, in other time durations, the calculated results seldom or not 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 convoluted by $K_{\text{Modified Backward}}$, it could be inferred that the PD signal has the significant increase or decrease at the certain moment, which implies that the computer user’s affective state are varying from the “relaxation” to the “stress” or from the “stress” to the “relaxation” at present.
Fig. 8. The results after the convolution of $K_{Modified\ Backward}$ with a pair of segments of the PD signal

### 3.2.2.2 Shape Information Detection for PD Signal

The shape information is a common used feature in digital signal and digital image processing field, whose most famous approach is the mathematical morphology [45],[46]. Fei Zhang and Yong Lian used multiscale mathematical morphology for GRS detection in wearable ECG devices in body area network [47]. Tomonari Yamaguchi etc. implemented the morphological multiresolution analysis to extract the features of EEG waves in order to discriminate the EEG signals recorded during left and right hand motor imagery and oddball task [48]. Chee-Hung Henry Chu and Edward J. Delp provided a new approach to impulsive noise suppression of ECG signals by using mathematical morphological operators [49].

In mathematical morphological operations, a structuring element, which is also called morphological operator, is operated on the original signal or image in order to extract the shape information. In a morphological operation, the value of each point (pixel) in the output signal (image) is based on a comparison of the corresponding point (pixel) in the input signal (image) with its neighbors. By choosing the size and shape of the structuring element, the researcher can construct the different neighborhood. There are two types of basic morphological operators, which are erosion and dilation. Erosion is the operation that outputs the maximum value of all the points (pixels) in the input value’s (pixel’s) neighborhood; whereas dilation is the operation that outputs the minimum value of all the points (pixels) in the input point’s (pixel’s) neighborhood.

The purpose of our study is to implement the developed algorithm to the online system for affective assessment of the human subject. Therefore, a modified approach, which is just based on the current and the previous neighboring values of the PD signal, is presented. The elementary single-scale mathematical morphology operators, which include modified dilation and modified erosion operators, for length N signal $f(n)$ are listed below:

**Modified Dilation:**

$$ f \oplus g(n) = \max_i (f(n+1-i) + g(i)) \quad n = L, \ldots, N \quad i = 1, \ldots, L \quad (16) $$

**Modified Erosion:**

$$ f \ominus g(n) = \min_i (f(n+1-i) + g(i)) \quad n = L, \ldots, N \quad i = 1, \ldots, L \quad (17) $$

where $i$ denotes the $i^{th}$ element in a length L structure element and $g(i)$ is a predefined structure element. The size and the value of the structure element can be generated according to the specific application. In some common cases, a constant value is chosen for the structure element. In our study, in order to timely detect the shape variation of the PD signal of the human subject without much delay, the length of the structure element of 5 and a constant value of 0 are implemented.

Fig. 9 shows the results of a pair of segments (one Congruent segment and one Incongruent segment) of the PD signal after processing with the modified dilation and erosion operators. Each dot point in the figure is the PD value, which represents the general value of the PD signal within one second. In order to better illustrate the effect of
the morphological transformation, the line rather than the dot point is used in the plot to show the results after modified dilation and erosion operations. As shown in Fig. 9, the solid line, the upper edge of the cluster of PD signal points, is the results after modified dilation operator processing; whereas the dash line, the lower edge of the cluster of PD signal points, is the results after modified erosion operator processing. These two lines can well describe the overall tendency of the PD signal, especially in amplitude change duration. It is apparent from Fig. 9 that when the PD signal has significant increases, the dot points are almost consecutively on the solid line; whereas when the PD signal has significant decreases, the dot points are almost consecutively on the dash line. Therefore, the evidence suggests that the relatively large number of the dot points within short period of time, which are on the solid line or the dash line, can indicate the obvious amplitude increase or the decrease variation of the PD signal to some extent.

![Fig. 9. The results after processing with modified dilation and modified erosion operator](image)

### 3.2.2.3 Algorithm for Online Affective Assessment of the Computer User

In our algorithm, there are mainly three steps for affective assessment of the human subject, which are preparation, decision based-feature voting and affective state determination. For the affective assessment of one certain PD$_i$ point, the current and the previous eight PD$_i$ points are needed to be considered.

**Step1:** The pupil diameter can constrict to 1.5mm or dilate to about 8 to 9 mm. In addition, different people has different pupil diameter values even under the same internal and external conditions. Hence, it is reasonable and necessary for each human subject to calculate the mean value of pupil diameter during certain period of time as the reference baseline when the subject is relaxed. In our experiment, during introductory section 1 (see Fig. 2), the mean value of twenty PD$_i$ points is calculated as the Threshold$_{\text{Reference value}}$ for the further threshold setting. In addition, two thresholds are needed for the affective state assessment, which are respectively regarded as the upper limit of the PD$_i$ signal amplitude fluctuation during the “relaxation” state and the lower limit of the PD$_i$ signal amplitude fluctuation during the “stress” state. These two thresholds are calculated based on Threshold$_{\text{Reference Value}}$ and two constants (c1 and c2, c1>1, c2>1, c2>c1; eg. c1=1.02, c2=1.07) are also required. The equations are shown below:

\[
\begin{align*}
\text{Threshold}_{\text{relaxation}} & = \text{PD}_{\text{Reference Value}} \times c1 \quad (c1>1) \\
\text{Threshold}_{\text{stress}} & = \text{PD}_{\text{Reference Value}} \times c2 \quad (c2>1)
\end{align*}
\]
Step2: In our algorithm, three features are extracted in order to accurately and timely identify the PDs signal change (from “relaxation” to “stress” or from “stress” to “relaxation”). Different feature has different weight score. If the current PDs signal value satisfies the criterion of one feature, the corresponding weight score will be added.

The first criterion for the first feature detection is the PD signal amplitude testing.

\[ \text{If } PDr(n) \geq \text{Threshold}_{\text{stress}} \quad \text{Weight}_{\text{stress}}(n) = \text{Weight}_{\text{stress}}(n) + 3 \]  
\[ \text{If } PDr(n) \leq \text{Threshold}_{\text{relaxation}} \quad \text{Weight}_{\text{relaxation}}(n) = \text{Weight}_{\text{relaxation}}(n) + 3 \]

The second criterion for the second feature detection is based on the modified backward differential method, which has been mentioned earlier. The derived values, which are stemmed from the convolution of \( K_{\text{Modified Backward}} \) with a period of the previous neighboring PDs signals, are denoted as \( \text{DP}_{MB} \) (Differentiated PDs value after processing with modified backward differential method). From fig. 8, it is obvious that during the transition of the affective state of the human subject, the \( \text{DP}_{MB} \) has almost consecutively large absolute values. Therefore, in order to measure the amplitude of the \( \text{DP}_{MB} \), the threshold has to be implemented, which is denoted as \( \text{Threshold}_{MB} \). In addition, it is also necessary to count the number of \( \text{DP}_{MB} \) greater than \( \text{Threshold}_{MB} \) within a short time interval so as to detect whether \( \text{DP}_{MB} \) values are almost consecutively large. The detailed approach for this criterion testing is as follows:

1. **Calculate the \( \text{DP}_{MB} \)**  
   \[ \text{DP}_{MB}(i) = 1*PDr(n+1-i) - 1*PDr(n-i-3) \]  
   (i=1,2,3,4,5)

2. **Calculate the number of \( \text{DP}_{MB} \) greater than the absolute value of \( \text{Threshold}_{MB} \)**
   \[ \text{Num1} = 0; \]
   \[ \text{If } \text{DP}_{MB}(i) > \text{Threshold}_{MB} \quad \text{Num1} = \text{Num1} + 1; \]  
   (i=1,2,3,4,5)

3. **Modify the weight for “relaxation” or “stress” affective assessment**
   \[ \text{If } \text{Num1} \geq 3 \quad \text{Weight}_{\text{stress}}(n) = \text{Weight}_{\text{stress}}(n) + 1 \]
   \[ \text{If } \text{Num2} \geq 3 \quad \text{Weight}_{\text{relaxation}}(n) = \text{Weight}_{\text{relaxation}}(n) + 1 \]

The third criterion is based on the shape information detection method which has also been discussed formerly. For one certain PDs signal decision making, the current PDs value and its previous neighboring PDs values are all considered, which can reveal the shape of PDs signal within short period of time to some extent. When the PDs signal is in its uptrend or downtrend (see fig. 9), we call these points “morphological matched points”. The details for this criterion testing are proposed below:

1. **Calculate the number of “morphological matched points”**
   \[ \text{Num}_{\text{dilate}}(i) = 0; \]  
   (29)
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=0,1,2,3,4) \((30)\)

\[
\text{Num_dilate} = \text{Num_dilate} + 1
\]

(31)

\[
\text{Num_erosion} = 0;
\]

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=0,1,2,3,4) \((33)\)

\[
\text{Num_erosion} = \text{Num_erosion} + 1
\]

(34)

(2) **Modify the weight for “relaxation” or “stress” affective assessment**

If \( \text{Num_dilate} \geq 3 \)

\[
\text{Weight}_{\text{stress}}(n) = \text{Weight}_{\text{stress}}(n) + 1
\]

(35)

If \( \text{Num_erosion} \geq 3 \)

\[
\text{Weight}_{\text{relaxation}}(n) = \text{Weight}_{\text{relaxation}}(n) + 1
\]

(36)

**Step3:** The purpose of this step is to determine the final affective state of the human subject. The processed PD\(_r\) value may have three statuses: from “relaxation” to “stress”, from “stress” to “relaxation” or no affect change.

- **If** \( \text{Weight}_{\text{stress}}(n) \geq 4 \) & the previous PD\(_r\) value in “relaxation” affective state

At present, PD\(_r\) (n) is the key point, which indicates that the affective state of the human subject from “relaxation” to “stress”

- **If** \( \text{Weight}_{\text{relaxation}}(n) \geq 4 \) & the previous PD\(_r\) value in “stress” affective state

At present, PD\(_r\) (n) is the key point, which indicates that the affective state of the human subject is from “stress” to “relaxation”

- **Otherwiese**

PD\(_r\) (n) is not the key point, which indicates that the human subject has the same affective state as the previous PD\(_r\) point.

4 **RELEVANT DATA SELECTION**

While the SCWT was implemented in our experiment for the purpose of eliciting mild mental stress of the human subjects during controlled intervals and analyzing the relevance of PD and GSR signal variation with affective state change, it is important to make the further verification for each subject or each pair of C and IC segments (denoted C/IC) causing a significant stress-related change, as hypothesized. In our study, two types of approaches were applied. The one is the selection of relevant C/IC segments pairs based on paired t-test analysis. The other method is to have traditionally used questionnaire to assess emotional state.

4.1 **Selection of Relevant C/IC Segment Pairs Based on Paired T-Test**

GSR signal is one of the most commonly used physiological signals which can gauge some degree of stress affect of the human subject [9], [13], [50], [51]. Therefore, the mean GSR value is respectively calculated in each C segment and in its following IC segment. Then a paired t-test is utilized to determine whether these C/IC segments pairs have significant increase change from “Congruent” to “Incongruent” SCWT presentations.

In our study, 42 subjects (mean: 27.52; std: 5.14) are 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). In order to better show the variation of GSR signal with the affective state change, the filtering and the normalization are preprocessed prior as mentioned above. The difference of the mean GSR values between an 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 [52], with the significance values of 0.200 and 0.446 respectively. Therefore, it is evident to confirm that the values of “DiffGSRmean” were normally distributed. Then the paired \( t \)-test (2-tailed) was implemented, resulting in the mean value and the standard deviation of “DiffGSRmean” about 0.093 and 0.326 respectively, and the significant value of 0.002, which can verify that the SCWT is generally an effective approach to elicit mild stress during Incongruent segment and remain relax during congruent segment. In addition, the 90% confidence interval of 126 “DiffGSRmean” values is [0.0453, 0.1414]. In order to remove the C/IC segments pairs which are probably insufficient, the C/IC segment pairs whose “DiffGSRmean” was less than 0.0453 were discarded from subsequent analysis. Hence, the total of 70 pairs of C/IC segment (i.e., 140 segments) is as the data to be used in the rest of offline and online algorithms study.

### 4.2 Selection of Relevant Individual Data Based on Questionnaire

In order to better understand the responses of the human subjects for the SCWT experiment, we conducted an evaluation that represents the two questions quantified with the rating scales of affective state to obtain a self report immediately after the whole SCWT experiment. Because emotion is very subjective and only the subject has the epistemic authority to express it (e.g., [53]), it has been widely used in psychological and affective computing research filed [54], [55], [56].

In our questionnaire, two questions were provided: “How do you feel in Congruent segment?” and “How do you feel in Incongruent segment?”. Five different affective rating scales (from “1” to “5”) are listed under each question. “1” represents that the subject feels relaxed just like in the normal environment. “5” represents that the subject feels very stressful just like he/she can not find the key of his/her house. “2”, “3” and “4” represents the increasing intermediate affect between “1” and “5”. The final results show that compared with the feeling in the Congruent segment, all the subjects feel more stressful when in the Incongruent segment, which indicates that the SCWT is generally an effective approach to elicit mild stress in some controlled intervals. In addition, in order to make the selected individual data much more convinced, the subject with the difference evaluation score between Incongruent and Congruent segment greater than or equal to two will be retained. Therefore, the total data of 31 subjects (include 93 Congruent and 93 Incongruent segments) are the valid data for offline and online affective assessment.

### 5 RESULTS

#### 5.1 Affective Assessment Based on Offline Processing

In this section, five different classification algorithms are going to be implemented through the features derived from the physiological signals recorded in each segment (see Table 1) in order to offline assess the affective state (relaxation vs. stress) of the human subjects. Specifically, in our work, Waikato Environment for Knowledge Analysis (WEKA) software, which can be freely downloaded from http://www.cs.waikato.ac.nz/ml/weka/, was utilized. The WEKA includes a collection of visualization tools and algorithms for data mining (containing preprocessing, classification, regression and clustering), together with graphical user interfaces for easy access [57], [58]. In our study, the five different classification algorithms are all from the WEKA for the purpose of observing the classification effectiveness and robustness of the features we extracted (see Table 1). In addition, the three classification phases were also performed, as illustrated in Table 2, in order to compare the classification efficiency of PD and GSR signal for affective assessment.

For the selected data (the pairs of C/IC segments) based on the paired \( t \)-test, the totals of 70 congruent segments and 70 incongruent segments were considered as viable used for training and testing the five types of classification algorithms for affective assessment. In addition, in order to obtain a more accurate and realistic assessment of the classifiers, a 10-fold cross validation method [59] was used. The accuracy rates from the experiments are shown in Table 3.
TABLE 3
RESULTS OF AFFECTIVE ASSESSMENT BY FIVE TYPES OF CLASSIFICATION ALGORITHMS
(SELECTED DATA BASED ON THE PAIRED T-TEST)

<table>
<thead>
<tr>
<th>Phase of Classification</th>
<th>Classification Algorithm</th>
<th>K*</th>
<th>Multilayer Perceptron</th>
<th>Naïve Bayes</th>
<th>Random Forest</th>
<th>JRip</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1.6 features from PD and GSR</td>
<td></td>
<td>80.00%</td>
<td>85.71%</td>
<td>87.86%</td>
<td>87.14%</td>
<td>86.43%</td>
<td>85.43%</td>
<td>9.85%</td>
</tr>
<tr>
<td>P2.3 features from GSR (No PD)</td>
<td></td>
<td>60.00%</td>
<td>63.57%</td>
<td>60.71%</td>
<td>63.57%</td>
<td>63.57%</td>
<td>62.29%</td>
<td>3.16%</td>
</tr>
<tr>
<td>P3.3 features from PD (No GSR)</td>
<td></td>
<td>85.71%</td>
<td>87.86%</td>
<td>88.57%</td>
<td>84.29%</td>
<td>85.71%</td>
<td>86.43%</td>
<td>3.06%</td>
</tr>
</tbody>
</table>

For the selected data (the individual subject data) based on the questionnaire, the totals of 93 congruent segments and 93 incongruent segments were considered as viable used for training and testing the five types of classification algorithms for affective assessment. In addition, in order to obtain a more accurate and realistic assessment of the classifiers, a 6-fold cross validation method was used. The accuracy rates from the experiments are shown in Table 4.

TABLE 4
RESULTS OF AFFECTIVE ASSESSMENT BY FIVE TYPES OF CLASSIFICATION ALGORITHMS
(SELECTED DATA BASED ON THE QUESTIONNAIRE)

<table>
<thead>
<tr>
<th>Phase of Classification</th>
<th>Classification Algorithm</th>
<th>K*</th>
<th>Multilayer Perceptron</th>
<th>Naïve Bayes</th>
<th>Random Forest</th>
<th>JRip</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1.6 features from PD and GSR</td>
<td></td>
<td>83.87%</td>
<td>86.56%</td>
<td>84.95%</td>
<td>87.10%</td>
<td>83.87%</td>
<td>85.27%</td>
<td>2.25%</td>
</tr>
<tr>
<td>P2.3 features from GSR (No PD)</td>
<td></td>
<td>53.23%</td>
<td>52.69%</td>
<td>47.85%</td>
<td>43.31%</td>
<td>58.60%</td>
<td>51.94%</td>
<td>21.19%</td>
</tr>
<tr>
<td>P3.3 features from PD (No GSR)</td>
<td></td>
<td>87.10%</td>
<td>88.17%</td>
<td>88.71%</td>
<td>86.56%</td>
<td>85.48%</td>
<td>87.20%</td>
<td>1.65%</td>
</tr>
</tbody>
</table>

5.2 Affective Assessment Based on Online Processing

In this section, the online processing algorithm was implemented to every subject. Fig.10 shows the affective assessment result of one subject after online PD signal processing. There are two portions in this figure. The upper part is the preprocessed PD signal, whose every value is the mean value of intervals with sixty samples length each. The lower part is the online affective assessment result for the whole experiment. The negative points with value -1 denote at this second, the computer claims that the affective state of the subject is relaxed; whereas the positive points with value 1 indicate that the computer user is regarded as stress at this second. It should be noted that only the Congruent and the Incongruent segments are considered to be classified in our study, therefore, other time period denotes 0 (in fact, during the whole experiment, the computer also did the classification in these time periods.). The accuracy rate for this example is 98.69%, and for each C/IC segment pair, they are 100%, 95.92% and 100% respectively.
The final accuracies of affective assessment based on online processing are as follows: for the selected relevant 70 C/IC segments pairs data (based on pair $t$-test), the mean classification rate is 72.30%; for the selected relevant 31 subjects’ data (based on questionnaire), the mean classification rate is 73.55%.

6 DISCUSSION

In our experiment, the PD and GSR physiological signals were collected and analyzed in order to investigate their efficiency and robustness of making affective assessment (“relaxation” vs. “stress”) for the human subjects. We develop two affective assessment approaches based on the PD signal, which are offline and online processing respectively.

For offline affective assessment algorithm, the wavelet denoising and Kalman filtering methods were applied to the original PD signal (after interpolation of the eye blinks) to remove noise components that do not represent the affective state variations of the subjects. Daubechies wavelet and nine decomposition levels were applied to the raw PD signal, which can significantly remove the abrupt change of the original signal. Kalman filtering was applied as the second preprocessing step to further remove the remaining noise of the PD signal, which can make it much more distinguishable when the subject is stressed. In the feature extraction section of our study, the “PDWalsh”, which is the difference of the first and the second Walsh coefficients, was used as one of the features to differentiate the affective state of the human subjects. It should be noted that the “PDWalsh” is not only simple, involving only addition and subtraction operations, but also can represent the “low frequent” component of the PD signal. Finally, five different classifiers are implemented to investigate the robustness and the effectiveness of the PD and GSR signals to differentiate the affective state of the human subjects. The outcomes are illustrated in Table 3 (the selected data based on paired $t$-test) and Table 4 (the selected data based on questionnaire). Both tables show that the three features extracted from the PD signal reach the highest accuracy 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 have the medium classification rate, which implies that the classification capability of GSR signal seems not as good as GSR signals. The second phase of both tables demonstrates the assumption: the three features extracted from the GSR signals get the lowest average only about 62.29% and 51.94%. All these observations suggest that the PD is the robust and efficient physiological signal for affective assessment, especially in “relaxation” vs. “stress”. 

Fig. 10. The online affective assessment result for one subject
For online affective assessment algorithm, it needs to make the affective assessment (“relaxation” vs. “stress”) for the human subjects every second according to the current and the previous PD\(_r\) data. On the other hand, the GSR signal is also implemented as the contrast signal (offline analysis) for comparison. In the signal preprocessing section of the process, the moving average window method was applied to the original PD\(_r\) signal (after the removal of the eye blinks). For online signal processing algorithm, this method is not only simple but also can noticeably reduce the random noise of the original signal. Compared with other traditional filtering methods such as FIR and IIR filters, the moving average window is optimal for keeping the sharpest step response of the signal [47]. It is obviously from Fig. 6, when the affect of the human subjects changes from the “relaxation” state to the “stress” state (from Congruent segment (C) to Incongruent segment (IC)), the PD\(_r\) signal has significantly increase, which is one of the most important features for differentiating the affective state of the human subjects. Essentially identically, when the affect of the human subjects changes from the “stress” state to the “relaxation” state (from Incongruent segment (IC) to Congruent segment (C)), the PD\(_r\) signal has significantly decrease. Therefore, it seems highly probable that the moving average window method is the most suitable approach for maintaining the sudden PD\(_r\) signal change. Another important issue about moving average window method is the choice of window length. As the number of points in the window increases, the noise becomes lower; however, the edge becomes less sharp. In our experiment, the sampling rate of the PD\(_r\) signal is 60 Hz and in our proposed algorithm, the 60 sample length window is implemented, which can not only noticeably reduce the random noise but also keep the sudden change of the PD signal.

In the second step, the decision based-feature voting method is applied, which is the situation faced when individual features are brought together in a group to solve the problems. According to the idea of synergy, the decision is tended to be more effective and reliable than the decision made by single individual feature. In our algorithm, each feature has different weight score, which presents the importance of that feature. If one extracted feature satisfies its criterion, the corresponding weight score can be added. The consensus can be made when the final voting score is greater than the threshold score previously set. In this algorithm, three features are extracted. The fist one is the amplitude detection. If the current data point is greater than the “Threshold\(_{\text{stress}}\)”, it has the large possibility that the PD\(_r\) signal has significant increase, which implies that the affective state of the human subjects is in “stress” affective state; whereas if the current data point is lower than the “Threshold\(_{\text{relaxation}}\)”, it seems highly like that the PD\(_r\) signal has significant decrease, which suggests that the affective state of the human subjects is probably in “relaxation” affective state.

The second feature is derived from the modified backward differentiation operation. Traditional differential operators, which include forward, backward and central differential operators, have long been used in digital signal and image processing field with great success to detect significant changes. Because the purpose of our research is to develop the online PD\(_r\) signal based affective assessment (“relaxation” vs. “stress”) system for the human subjects, it is evident that only the present and the previous PD\(_r\) data can be used. Therefore, the traditional backward differentiation operator is implemented first, whose result is shown in Fig. 7. It is widely known that according to the sign and the amplitude of the value derived from differentiation operators, the significant change of the PD\(_r\) signal can be automatic identified. As we can see from Fig. 7, there are almost no values with large amplitude at the two boundaries of the Incongruent segment (IC), which implies that it seems there is no significant change during the affective transition. However, in fact, it is evident from Fig. 6 that the PD\(_r\) signal has obvious increase or decrease at the two boundaries of the Incongruent segment (IC). The reason why the traditional differentiation operator is not very effective may be small amplitude difference between two neighboring data points. Therefore, the modified differentiation operator [-1 0 0 0 1] is proposed. This operator detects the difference between the current and the previous fourth points. The proper interval between the two detected points can not only has the large difference between two points but also present the overall tendency of the PD\(_r\) signal. If the successive large value with the same sign is detected from the results derived from modified backward differentiation operator, it is highly believed that the PD\(_r\) signal has significant change with this period of time. In our algorithm, for one current PD\(_r\) data, a set of five differences are to be calculated, which are displayed in equation (22). In other words, the current and the previous eight PD\(_r\) data points are considered for voting, which can reliably outline the variation of the PD\(_r\) signal lately.

The third extracted feature is to detect the shape information of the PD\(_r\) signal based on the morphological processing. The morphological signal processing can quantify the shape, size and other aspects of the geometrical
structure of signal, in a rigorous way that also agrees with human intuition and perception [28]. In contrast, compared with traditional signal processing tools such as spectrum analysis, the morphological methods are more suitable for shape analysis. It should be noted that for the PD signal, it is almost impossible to differentiate the affective state of the human subjects based on frequency domain analysis. Therefore, shape detection seems an effective approach. In addition, the choice of structure element is another significant factor affecting the analysis results. The main purpose of this operation is to online detect the sudden change (“increase” or “decrease”) of the PD signal rather than noise suppression or background normalization. Therefore, considering the computation complexity, the length of the structure element of 5 and a constant value of 0 are implemented. In this operation, counting the number of “morphological matched points” (the points have the same values after the modified dilation or erosion processing with their originals) within one certain period of time can significantly represent the latest PD variation.

The last step for the PD signal based online affective assessment algorithm is the affective determination portion. In our study, we implemented two kinds of common used methods to select the more convinced data to do the affective assessment. One is the selection of relevant C/IC segment pairs based on paired t-test, which is based on the experimental statistical approach. We remain 70 out of 140 segment pairs to do the further analysis, which gets the average classification rate of 72.30%. The other method is the selection of relevant individual data based on questionnaire, which is the common used approach in psychology. 31 out of 42 subjects are remained and the average classification accuracy is up to 73.55% with the highest rate up to 98.69%. An additional important aspect to note in these results is that the encouraging level of online affective classification was achieved even in spite of the temporary illumination increases introduced during the IC2 and C3 segments of experimentation. This seems to confirm that the algorithms developed for PD signal based online affective assessment are very robust even under indoor ordinary light variation.

In order to study the comparative efficiency of affective assessment (“relaxation” vs. “stress”) of the PD signal, the GSR signal was also monitored as reference. GSR signal is a type of widely used physiological signal to detect the emotions of the human subjects. According to the physiological background and the previous research of GSR signal, “GSRmean”, “GSRnum” and “GSRrisingTime” three features are extracted from the whole Congruent (“C”) and Incongruent (“IC”) segments to offline differentiate between “relaxation” and “stress” state of the computer user. No matter which type of valid data selection method is applied, they are all significantly lower than the accuracy of online PD affective assessment algorithm (62.29% vs. 72.30%; 51.94% vs. 73.55%). Therefore, it is much more convinced to claim that the PD signal should be considered as one of the most powerful physiological signals to differentiate the affective state of the computer user, especially about “relaxation” vs. “stress”.

5 Conclusion

This paper presented a new offline and online affective assessment approach to differentiate the “stress” vs. “relaxation” states of the human subjects through detecting and processing the Pupil Diameter (PD) signal, which can be non-invasively monitored by eye gaze tracking system at present.

In the offline affective assessment algorithm, wavelet denoising and Kalman filtering were applied on the original PD signal (after interpolation of eye blinks) in order to remove the noise. Then three features (PDmean, PDmax and PDWalsh) are extracted and processed by five different types of classification algorithm. The final results show that using the features only derived from the PD signal can yield an average segment classification accuracy of 86.43% (see table 3) and 87.20% (see table 4) respectively based on two types of data selection approaches. However, using the features from the GSR signal (excluding the features from the PD signal) could only yield 63.57% (see table 3) and 58.60% (see table 4).

In the online affective assessment algorithm, the moving average window method was implemented on the original PD signal. In addition, sixty samples length window (the sampling rate of PD is 60Hz) is chosen to extract the representative data denoted as PD, to reflect the overall situation within one minute. There are mainly three steps for online affective assessment algorithm, which are preparation, decision based-feature voting and affective state determination. The final results are 72.30% and 73.55% for two different types of data selection methods, which are obviously higher than the accuracy for offline GSR affective assessment.

In sum, the results of our experiment are encouraging and confirm that the PD signal should be considered as one of the most powerful physiological signals to involve in the future automated real time affective recognition systems, especially for detecting the “relaxation” vs. “stress” states of a computer user.
Acknowledgment
This work was sponsored by NSF grants HRD-0833093 and CNS-0959985.

REFERENCES


