3-29-2013

Off-line and On-line Affective Recognition of a Computer User through A Biosignal Processing Approach

Peng Ren
Florida International University, pren001@fiu.edu

DOI: 10.25148/etd.FI13042211
Follow this and additional works at: https://digitalcommons.fiu.edu/etd

Recommended Citation
https://digitalcommons.fiu.edu/etd/838

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.
A dissertation submitted in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY
in
BIOMEDICAL ENGINEERING
by
Peng Ren
2013
To: Dean Amir Mirmiran  
College of Engineering and Computing  

This dissertation, written by Peng Ren, and entitled Off-line and On-line Affective Recognition of a Computer User through A Biosignal Processing Approach, having been approved in respect to style and intellectual content, is referred to you for judgment.  

We have read this dissertation and recommend that it be approved.  

________________________________________  
Malek Adjouadi  

________________________________________  
Jean Andrian  

________________________________________  
Wei-Chiang Lin  

________________________________________  
Armando Barreto, Major Professor  

Date of Defense: March 29, 2013  

The dissertation of Peng Ren is approved.  

________________________________________  
Dean Amir Mirmiran  
College of Engineering and Computing  

________________________________________  
Dean Lakshmi N. Reddi  
University Graduate School  

Florida International University, 2013
DEDICATION

I dedicate this Ph. D. dissertation to my dear parents, other family members and friends, for without their patience, understanding, support, and love, the completion of this endeavor would never have been possible. I thank them for their support, especially, when I were suffered from anxiety disorder and panic attack, during Ph.D. study.
ACKNOWLEDGMENTS

I would like to thank my major Professor, Dr. Armando Barreto, who is a very responsible, excellent and nice advisor. The National Science Foundation supported the research described in this dissertation through grants CNS-0520811, HRD-0833093, CNS-0426125, IIS-0308155 and HRD-0317692. I would also like to thank the Biomedical Engineering Department at Florida International University for their financial support and encouragement during my five years of study. Finally, I would like to thank my committee members Dr. Malek Adjouadi, Dr. Jean Andrian and Dr. Wei-Chiang Lin for their expertise and advice.
Physiological signals, which are controlled by the autonomic nervous system (ANS), could be used to detect the affective state of computer users and therefore find applications in medicine and engineering. The Pupil Diameter (PD) seems to provide a strong indication of the affective state, 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 affective assessment ("relaxation" vs. "stress") are proposed. Wavelet denoising and Kalman filtering methods are first used to remove abrupt changes in the raw Pupil Diameter (PD) signal. Then three features (PDmean, PDmax and PDWalsh) are extracted from the preprocessed PD signal for the affective state classification. In order to select more relevant and reliable physiological data for further analysis, two types of data selection methods are applied, which are based on the paired $t$-test and subject self-evaluation, respectively. In addition, five different kinds of the classifiers are implemented on the selected data, which achieve average accuracies up to 86.43% and 87.20%, respectively. Finally, the receiver operating characteristic (ROC)
curve is utilized to investigate the discriminating potential of each individual feature by evaluation of the area under the ROC curve, which reaches values above 0.90.

For the on-line affective assessment, a hard threshold is implemented first in order to remove the eye blinks from the PD signal and then a moving average window is utilized to obtain the representative value PDₚ for every one-second time interval of PD. There are three main steps for the on-line affective assessment algorithm, which are preparation, feature-based decision voting and affective determination. The final results show that the accuracies are 72.30% and 73.55% for the data subsets, which were respectively chosen using two types of data selection methods (paired t-test and subject self-evaluation).

In order to further analyze the efficiency of affective recognition through the PD signal, the Galvanic Skin Response (GSR) was also monitored and processed. The highest affective assessment classification rate obtained from GSR processing is only 63.57% (based on the off-line processing algorithm). The overall results confirm that the PD signal should be considered as one of the most powerful physiological signals to involve in future automated real-time affective recognition systems, especially for detecting the “relaxation” vs. “stress” states.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Problem Statement</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Objective of Research</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Significance of this Research</td>
<td>3</td>
</tr>
<tr>
<td>1.4 Structure of the Dissertation</td>
<td>4</td>
</tr>
<tr>
<td>2 BACKGROUND</td>
<td>6</td>
</tr>
<tr>
<td>2.1 Overview of Affective Computing Concepts</td>
<td>6</td>
</tr>
<tr>
<td>2.2 Models and the Applications of the Affective Computing</td>
<td>7</td>
</tr>
<tr>
<td>2.3 Autonomic Nervous System</td>
<td>9</td>
</tr>
<tr>
<td>2.4 Physiological Signals</td>
<td>12</td>
</tr>
<tr>
<td>2.4.1 Pupil Diameter</td>
<td>13</td>
</tr>
<tr>
<td>2.4.2 Galvanic Skin Response</td>
<td>14</td>
</tr>
<tr>
<td>2.5 Overview of Stress</td>
<td>16</td>
</tr>
<tr>
<td>2.5.1 The Concept of Stress</td>
<td>16</td>
</tr>
<tr>
<td>2.5.2 Literature Review of Stress Detection</td>
<td>17</td>
</tr>
<tr>
<td>2.6 Summary</td>
<td>18</td>
</tr>
<tr>
<td>3 EXPERIMENT DESIGN</td>
<td>19</td>
</tr>
<tr>
<td>3.1 Software Development</td>
<td>19</td>
</tr>
<tr>
<td>3.2 Hardware Development</td>
<td>23</td>
</tr>
<tr>
<td>3.2.1 PD Measurement System</td>
<td>25</td>
</tr>
<tr>
<td>3.2.2 GSR Measurement System</td>
<td>27</td>
</tr>
<tr>
<td>3.2.3 IL Measurement System</td>
<td>28</td>
</tr>
<tr>
<td>3.2.4 Synchronization of All the Measured Signals</td>
<td>29</td>
</tr>
<tr>
<td>3.2.5 Overall Setup of the Experiment</td>
<td>30</td>
</tr>
<tr>
<td>3.3 Experimental Procedure</td>
<td>33</td>
</tr>
<tr>
<td>3.4 Summary</td>
<td>33</td>
</tr>
<tr>
<td>4 METHODOLOGY</td>
<td>35</td>
</tr>
<tr>
<td>4.1 Off-line Affective Assessment</td>
<td>36</td>
</tr>
<tr>
<td>4.1.1 Physiological Signal Preprocessing</td>
<td>36</td>
</tr>
<tr>
<td>4.1.2 Data Normalization and Feature Extraction</td>
<td>48</td>
</tr>
<tr>
<td>4.1.3 Classification Algorithm</td>
<td>51</td>
</tr>
<tr>
<td>4.1.4 Receiver Operating Characteristic (ROC) Curve</td>
<td>54</td>
</tr>
<tr>
<td>4.2 On-line Affective Assessment</td>
<td>56</td>
</tr>
<tr>
<td>4.2.1 Physiological Signal Preprocessing</td>
<td>56</td>
</tr>
<tr>
<td>4.2.2 Development of On-line Affective Assessment Algorithm</td>
<td>57</td>
</tr>
<tr>
<td>4.2.2.1 Modified Backward Differential Method for Sudden PD Signal Change Detection</td>
<td>58</td>
</tr>
<tr>
<td>4.2.2.2 Shape Information Detection for PD Signal</td>
<td>61</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3-1: Four kinds of experimental situation</td>
<td>23</td>
</tr>
<tr>
<td>Table 3-2: Three steps for eye tracking</td>
<td>26</td>
</tr>
<tr>
<td>Table 4-1: Features extracted from PD and GSR signals</td>
<td>51</td>
</tr>
<tr>
<td>Table 4-2: Three different conditions studied in the classification phase of this study</td>
<td>54</td>
</tr>
<tr>
<td>Table 5-1: Values of all 6 features from one subject</td>
<td>74</td>
</tr>
<tr>
<td>Table 5-2: Results of affective assessment by five types of classification algorithms (Selected data based on the paired ( t )-test)</td>
<td>77</td>
</tr>
<tr>
<td>Table 5-3: Results of affective assessment by five types of classification algorithms (Selected data based on the questionnaire)</td>
<td>79</td>
</tr>
<tr>
<td>Table 5-4: AUROC results for all 6 features derived from PD and GSR</td>
<td>80</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2-1: The model of affective computing</td>
<td>8</td>
</tr>
<tr>
<td>Figure 2-2: Composition of the nervous system</td>
<td>9</td>
</tr>
<tr>
<td>Figure 2-3: Pathway of the autonomic nervous system</td>
<td>12</td>
</tr>
<tr>
<td>Figure 2-4: The dual innervations of the iris by the ANS</td>
<td>14</td>
</tr>
<tr>
<td>Figure 3-1: 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).</td>
<td>21</td>
</tr>
<tr>
<td>Figure 3-2: Stimuli schedule of the experimental protocol.</td>
<td>21</td>
</tr>
<tr>
<td>Figure 3-3: The testing signal tracing in TracerDAQ®</td>
<td>24</td>
</tr>
<tr>
<td>Figure 3-4: TOBII T60</td>
<td>25</td>
</tr>
<tr>
<td>Figure 3-5: TOBII T60 eye tracking interface</td>
<td>27</td>
</tr>
<tr>
<td>Figure 3-6: GSR2 module for GSR measurement system</td>
<td>28</td>
</tr>
<tr>
<td>Figure 3-7: Circuit of luminance meter</td>
<td>29</td>
</tr>
<tr>
<td>Figure 3-8: “Start” and “Stop” button in the “Track” section of TOBII control program</td>
<td>30</td>
</tr>
<tr>
<td>Figure 3-9: Complete instrumental setup</td>
<td>31</td>
</tr>
<tr>
<td>Figure 3-10: An example of participant in the experiment</td>
<td>32</td>
</tr>
<tr>
<td>Figure 3-11: A set of signals after synchronization. From top to bottom: the raw PD, the raw GSR and IL</td>
<td>32</td>
</tr>
<tr>
<td>Figure 4-1: The nested vector space for MRA</td>
<td>38</td>
</tr>
<tr>
<td>Figure 4-2: Multi-resolution wavelet decomposition</td>
<td>42</td>
</tr>
<tr>
<td>Figure 4-3: Multi-resolution wavelet reconstruction</td>
<td>43</td>
</tr>
<tr>
<td>Figure 4-4: Model of Kalman filter.</td>
<td>45</td>
</tr>
<tr>
<td>Figure 4-5: Model of one step recursive Kalman filter</td>
<td>46</td>
</tr>
</tbody>
</table>
Figure 4-6: Kalman filter recursive algorithm .................................................................48

Figure 4-7: The algorithm for off-line affective assessment based on PD signal ............56

Figure 4-8: The algorithm for on-line affective assessment based on PD signal ..........67

Figure 5-1: A set of raw signals in the experiment. From top to bottom: the raw PD, the raw GSR and IL .................................................................................................................70

Figure 5-2: Original PD signal (upper), signal after wavelet denoising (middle) and signal after Kalman filtering (lower) ........................................................................................................72

Figure 5-3: The Walsh transform of the average onset value of the PD signal for six Stroop segments of one subject .............................................................................................................73

Figure 5-4: The ROC/ROCCH of PDmean ........................................................................80

Figure 5-5: The ROC/ROCCH of PDmax ..........................................................................81

Figure 5-6: The ROC/ROCCH of PDMwalsh ..................................................................81

Figure 5-7: The ROC/ROCCH of GSRmean ....................................................................82

Figure 5-8: The ROC/ROCCH of GSRrisingTime ............................................................82

Figure 5-9: The ROC/ROCCH of GSRnum ....................................................................83

Figure 5-10: The preprocessed raw PD signal in the on-line algorithm.........................84

Figure 5-11: The results after the convolution of $K_{\text{Backward}}$ with a pair of segments of the PD$_t$ signal .................................................................85

Figure 5-12: The results after the convolution of $K_{\text{Modified Backward}}$ with a pair of segments of the PD$_t$ signal .................................................................86

Figure 5-13: The results after processing with modified dilation and modified erosion operation ................................................................................................................................88

Figure 5-14: The on-line affective assessment result for one subject .........................89
## LIST OF SYMBOLS

<table>
<thead>
<tr>
<th>SYMBOLS</th>
<th>DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A:</strong></td>
<td></td>
</tr>
<tr>
<td>( a_k )</td>
<td>Wavelet coefficient in the subspace of scale ( k ).</td>
</tr>
<tr>
<td>( A(k) )</td>
<td>The state transition matrix of the process from state ( k-1 ) to state ( k ), ((n \times m)).</td>
</tr>
<tr>
<td>( A^j )</td>
<td>The approximation coefficient of the original signal at levels ( j ).</td>
</tr>
<tr>
<td>( A^j' )</td>
<td>The processed or non processed approximation coefficients of the original signal at levels ( j ).</td>
</tr>
<tr>
<td>( A_j )</td>
<td>The approximation of the original signal at level ( j ).</td>
</tr>
<tr>
<td><strong>B:</strong></td>
<td></td>
</tr>
<tr>
<td>( b<a href="u">k</a> )</td>
<td>The ( k^{th} ) bit in the binary representation of the index ( u ) in Walsh transform.</td>
</tr>
<tr>
<td><strong>C:</strong></td>
<td></td>
</tr>
<tr>
<td>( C(k) )</td>
<td>The noiseless connection between the state vector and the measurement vector, ((m \times n)).</td>
</tr>
<tr>
<td>( C(s, \tau) )</td>
<td>Continuous wavelet transform of the signal.</td>
</tr>
<tr>
<td>( C_1 )</td>
<td>The constant used to define the threshold for relaxation detection.</td>
</tr>
<tr>
<td>( C_2 )</td>
<td>The constant used to define the threshold for stress detection.</td>
</tr>
<tr>
<td><strong>D:</strong></td>
<td></td>
</tr>
<tr>
<td>( \text{DWT}(m,n) )</td>
<td>Discrete Wavelet Transform of the signal.</td>
</tr>
<tr>
<td>( \text{DP}_{MB} )</td>
<td>Differentiated PD(_r) value after processing with modified backward differential method.</td>
</tr>
</tbody>
</table>
\( D^j \) The detail coefficients of the original signal at level \( j \).

\( D^j' \) The processed or non processed detail coefficients at level \( j \).

\( D_j \) The detail of the signal at level \( j \).

\[ F: \]
\[ f \] The input signal.

\[ G: \]
\[ g \] High-pass decomposition filter for multiresolution analysis.

\[ g' \] High-pass reconstruction filter for multiresolution analysis.

\[ GSRmean \] Mean value of the amplitude of each GSR samples.

\[ GSRnum \] Number of the GSR in a segment.

\[ GSRrisingTime \] (Average) Rising time of each GSR in a segment.

\[ H: \]
\[ h \] Low-pass decomposition filter for multiresolution analysis.

\[ h' \] Low-pass reconstruction filter for multiresolution analysis.

\[ H(k) \] The correction matrix.

\[ I: \]
\[ IC \] The Incongruent segment in the Stroop Test (the font color and the meaning of the 30 words are presented differently).

\[ IC1 \] The first Incongruent segment in the Stroop Test.
IC2 The second Incongruent segment in the Stroop Test.

IC3 The third Incongruent segment in the Stroop Test.

IL The Illumination signal.

IS The Introductory Segment of the Stroop Test with the purpose of making the subject get used to the task environment and establishing an appropriate initial level for his/her psychological state.

**K:**

K\text{Backward} The backward differentiation operator \([-1 \ 1]\).

K\text{Modified Backward} The modified backward differentiation operator \([-1 \ 0 \ 0 \ 0 \ 1]\).

**L:**

L^2(\mathbb{R}) The space of square-summable sequences, which is a Hilbert space.

**M:**

M The length of coarsest approximation coefficient over 2 in multi-resolution wavelet analysis.

Max Maximum value in all of the values.

Min Minimum value in all of the values.

**P:**

PD Pupil diameter signal and the sampling rate is 60 Hz.

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 based on the PD signal during the onset of each Stroop segment.

PD_{r} The average value of valid PD samples in every one second interval. (The valid PD samples mean their amplitudes are greater or equal to the threshold sat for eye blink removal).

PDReference Value The mean value of twenty PD_{r} points during introductory section 1 calculated for the further threshold setting.

P^{*}(C|m) Given a test instance m whose class is unknown, its probability of being in category C is calculated by summing the probabilities from m to each instance that is a member of C.

Q: Q Process noise w_{1} is drawn from N(0,Q), with covariance matrix Q.

R: R Measurement error w is drawn from N(0, R), with covariance matrix R.

ROC Receiver Operating Characteristic Curve.

RS The resting segment in the Stroop Test, which makes the subject back to a baseline affective state.

S: s The scale parameter in continuous wavelet transform, which is the measurement of the frequency.

\frac{1}{\sqrt{s}} Normalization constant in continuous wavelet transforms.

S(k) The state vector of the process at time k,(n×1).

\hat{S}(k) The optimal estimation of the state vector.
\( \hat{S}(k) \)  
The prior estimate of \( \hat{S}(k) \) at time \( k \).

\( \hat{S}(k) \)  
The difference of real value \( S(k) \) and its estimation \( \hat{S}(k) \).

**T:**

Threshold\textsubscript{blink}  
The threshold set to remove the eye blink.

Threshold\textsubscript{relaxation}  
The upper limit of the PD\textsubscript{r} signal amplitude fluctuation during the “relaxation” state.

Threshold\textsubscript{stress}  
The lower limit of the PD\textsubscript{r} signal amplitude fluctuation during the “stress” state.

Threshold\textsubscript{MB}  
The threshold that has to be implemented for DP\textsubscript{MB}.

**U:**

\( u(n) \)  
The PD signal after wavelet denoising and Kalman filtering.

**V:**

\( V_j \)  
Vector spaces for multiresolution analysis (MRA) at the resolution of \( j \).

**W:**

\( w(k) \):  
The associated measurement error.

\( w_1(k-1) \)  
The associated white noise process with known variance, \( (n \times 1) \).

\( W(u) \)  
The \( u \)-th Walsh coefficient derived from the Walsh transforms.

\( W_j \)  
The difference (in the subspace sense) between adjacent space \( V_j \) and \( V_{j+1} \).

Weight\textsubscript{stress}  
Weight score for stress detection.
Weight\textsuperscript{relaxation}  Weight score for relaxation detection.

**X:**

$X(k)$  The actual measurement of $S$ at time $k$, $(m \times 1)$.

$\hat{X}'(k)$  The prior estimation of $\hat{X}(k)$ at time $k$.

$\hat{X}(k)$  The innovation or measurement residual during Kalman filtering.

**Y:**

$y(m)$  The one-dimensional sequence to be Walsh-transformed.

**Z:**

$z(x)$  A scalar function

**Others:**

$\psi_k(t)$  Mother wavelet, used as basis functions to represent other functions.

$\tau$  The translation parameter in continuous wavelet transform, which is the measurement of time.

$\phi(t)$  Scaling function.

$\varepsilon(k)$  The error covariance matrix.

$\varepsilon'(k)$  The prior estimation of $\varepsilon(k)$ at time $k$.

$\Delta$  Forward difference.

$\nabla$  Backward difference.

$\delta$  Central difference.
\( \lambda \)  
The step size used to define the finite difference, is assumed to be small, \( |\lambda| \ll 1 \)

\( \rho_i \)  
The current PD\(_i\) signal value.

\( \rho_{i-k} \)  
The previous \( k^{th} \) PD\(_i\) signal value.
CHAPTER 1
INTRODUCTION

1.1 Problem Statement

The concept of affective computing was first proposed by Rosalind Picard (in 1995) from MIT as the tool to improve the experience of human-computer interaction by recognizing, interpreting and responding to the affect of the computer user. In addition, affective computing is an interdisciplinary research area, which includes computer science, electrical engineering, biology and psychology.

Emotions can be recognized as discrete and consistent responses to internal or external events which have a particular significance for the organism [1]. In human beings, emotions play an important role in many aspects: (1) perception; (2) organization of memory; (3) physical resources; (4) categorization and preference; (5) goal monitoring and activation; (5) decision making; (6) planning; (7) intention; (8) focus and attention; (9) motivation and performance; (10) communication; (11) learning; etc. Therefore, affective computing can not only improve the computer’s performance but also the user’s productivity as well.

For a computer to be affective should have at least one of the following functions: (1) recognizing emotion; (2) expressing emotion; (3) having emotion. A critically important issue is how the emotions can be recognized by the computers. A variety of methods for measuring affective states of the users have been tried, such as the identification of facial expressions, speech understanding and body gesture recognition [2]. However, these methods for identifying the emotions of the computer user are
susceptible to environmental interference or voluntary masking. Therefore, alternative approaches, which analyze a variety of autonomic activities such as Electroencephalogram (EEG), Electrocardiogram (ECG), Galvanic Skin Response (GSR), Blood Volume Pulse (BVP), Blood Pressure (BP), Skin Temperature (ST), have been chosen. These physiological signals represent the activity of the autonomic nervous system (ANS), whose functions are involuntary. In our study, we monitored the Pupil Diameter (PD) signal and proposed new approaches off-line and on-line to detect the “stress” affective state, which is associated with activation of the sympathetic division of the ANS.

1.2 Objective of Research

This dissertation focuses on the development of off-line and on-line approaches to visualize and recognize the affective state (“relaxation” vs. “stress”) of a computer user from analysis of his/her pupil diameter variations caused by sympathetic activation, through non-invasive and non-obtrusive PD monitoring.

For the off-line affective assessment algorithm, wavelet denoising and Kalman filtering methods are first used to remove abrupt changes in the raw Pupil Diameter (PD) signal, which seems to be effective way to isolate the variation of the PD based on the affective changes of the computer user. Then three features, which are denoted as “PDmean”, “PDmax”, and “PDWalsh”, are extracted from the preprocessed PD signal for the affective state classification. Finally, five different types of classifiers are implemented in order to evaluate the average classification accuracy and the robustness of our developed algorithm. For comparison purposes, the Galvanic Skin Response (GSR)
signal has also been measured and analyzed as the traditionally leading indicator of the sympathetic activation that accompanies “stress”, when it is experienced by a computer user. Three features, which are “GSRmean”, “GSRrising Time” and “GSRnum”, are extracted from the preprocessed GSR signal and then processed by the same five types of classifiers as the PD signal.

For the on-line affective assessment algorithm, a hard threshold and a moving average window are implemented to derive the representative data denoted as PD, to reflect the overall PD sampling values within one second. Then three steps for on-line affective assessment, which are preparation, decision based-feature voting and affective state determination, are performed in order. In this on-line processing algorithm, the extracted features need to satisfy 3 detection criteria, which are, respectively based on the current values of PD, the modified backward differential method and the method for shape information detection.

1.3 Significance of this Research

The purpose of this research is to develop new approaches (off-line and on-line) to recognize the affective state of the computer users, which can, in turn, make the computer much more useful. The computer with the affective assessment ability could select proper strategies such as praise, encouragement or intervention as the response to the users. While this ability could seem somewhat unnecessary or impractical in some computer systems, it may be useful or even critical to fulfill the main purpose of some systems. For example, the emotional interaction in child learning; the emotional-social intelligence prosthesis for the emotionally-impaired; combating frustration during the
learning process and so on.

In addition, the enhancement of the experience of the human-computer interaction is another contribution of the implementation of affective computing system. For example, it is easy for the computer user, who works for long periods in front of a computer, to become fatigued. In some cases, the computer user even feels stressed or depressed. Friedman and Booth-Kewley (1987) [3] cited evidence of disease-prone personality traits, which include depression, anger, hostility and anxiety. Hans Eysenck (1988) [4] demonstrated that long term negative emotional states may lead to cancer or heart disease. Candace B. Pert (2003) [5] demonstrated that relaxation may be able to increase the effectiveness of the immune system of the body. Therefore, if computers can recognize the affect of the users and be aware of the emergence of the negative emotions such as stress, they may warn or adjust their interaction patterns in order to reduce the possible negative effects.

1.4 Structure of the Dissertation

This dissertation is structured in the following manner. Chapter 2 introduces the background information of this dissertation including the concept, model and application of affective computing. An overview of autonomic nervous system concepts and its related physiological signals are also described. Finally, this chapter covers the concept of stress and summarizes the ways to elicit stress used in previous research. Chapter 3 describes the experimental setup for this project, which includes the software and hardware implementations. Chapter 4 states the framework and the methodologies of this research. It contains two parts, which are off-line and on-line affective assessment
algorithms respectively. The 4 sections describing the off-line algorithm present the physiological signal preprocessing, data normalization and feature extraction, classification, and receiver operating characteristic curves. The 4 sections describing the on-line algorithm present the physiological signal preprocessing, preparation, decision based-feature voting and affective state determination four sections. Chapter 5 shows the results and then provides a detailed and comprehensive discussion. Chapter 6 concludes this dissertation and proposes the potential directions for future work.
CHAPTER 2
BACKGROUND

This chapter describes the background information for this dissertation topic. First, the overview concept of affective computing is introduced. Second, the models and the applications of affective computing are illustrated. Third, the physiological background relative to the Autonomic Nervous System (ANS) is provided. Forth, two different kinds of physiological signal, which are the Pupil Diameter (PD) and the Galvanic Skin Response (GSR), are presented. Finally, an overview of the concept of stress is introduced and the review of previous research about stress detection is also summarized.

2.1 Overview of Affective Computing Concepts

Affective computing is an emerging branch of human-computer interaction which seeks to close the communication gap between the human and the machine – as Rosalind Picard wrote: “computers do not need affective abilities for the fanciful goal of becoming humanoids, they need them for a meeker and more practical goal: to function with intelligence and sensitivity towards humans” [6]. Affect and emotion play an important role in our everyday lives: key scientific findings indicate that emotions play an essential role in rational decision making, perception learning, and a variety of other cognitive functions [6], [7]. Similarly, Damasio argues that the body and our emotions have a key role in the way we think and in rational decision-making [8]. In addition, Reeves and Nass investigated a number of scenarios that demonstrated that we treat computers as if they were social actors – i.e., as if they had emotions, feelings and intelligence too, even if we know that we are dealing with a computer [9].
Emotions are both physical and cognitive. “Today we know that not only thoughts can lead to emotions, but emotions can occur without obvious cognitive evaluation, such as by changes in bodily chemistry [6]”. In the physical aspects of emotions, Manfred Clynes wrote: “The emotional character is expressed by a specific subtle modulation of the motor action involved which corresponds precisely to the demands of the sentient state” [10]. The apparent sentient modulations easily perceived by other people are as follows: facial expression, gesture, posture, speech and motor form of expression. Other physiological measures which are less apparent to other people include physiological signals, self-conscious emotions (guilt and shame) marked with postural differences and so on. Emotions are also cognitive. Not only science study but also personal experience indicate that cognitive appraisal can contribute to the generation of emotions. In addition, the appraisal, comparison, categorization, inference, attribution or judgment can be considered to be “cognitive”, which influences the emotions.

2.2 The Models and the Applications of the Affective Computing

There are three main key components in affective computing, which are initially proposed by Hudlicka [11] and identified by Barreto [12]:

(i) Affect Sensing and Recognition;

(ii) User Affect Modeling / Machine Affect Modeling;

(iii) Machine Affect Expression.

Figure 2-1 shows a generalized diagram for the affective computing system.
There are many applications for affective computing:

(i) Virtual characters (Health interventions; Entertainment; Intelligent training system; Commercial agent; and so on)

(ii) Measure the affect of the computer user (Detection of negative emotions of the computer user such as frustration, stress; Health interventions; and so on)

From the previous descriptions, it is clear that measuring, understanding and responding appropriately to the computer user’s affect are meaningful work to influence our daily life. This research proposes new affective sensing approaches to evaluate the computer user’s affective states (“relaxation” vs. “stress”) through processing the physiological signals, which are inherently controlled by their Autonomic Nervous System (ANS).
2.3 Autonomic Nervous System

The nervous system is an organ system containing neurons that coordinate the actions of the human beings and transmit signals between different parts of their bodies. The nervous system consists of two parts: central and peripheral, which are illustrated in Figure 2-2.

![Figure 2-2: The compositions of nervous system](image)

The peripheral nervous system handles the input and output of the central nervous system and includes all the portions of the nervous system outside the brain and spinal cord. The peripheral nervous system consists of sensory nerves and motor nerves. “The sensory nerves carry messages from receptors in the skin, muscles and other internal and external sense organs to the spinal cord and then to the brain.” “The motor nerves carry orders from central nervous system to muscles, glands to contract and produce chemical messages [13].”

The ANS is the subdivision of the peripheral nervous system transmitting impulses from the central nervous system to peripheral organ systems, which regulates
the body activities that are generally not under conscious control. “The ANS contains motor neurons that innervate smooth and cardiac muscles and glands and its function is to make adjustments to ensure optimal support for body activities. [13]” The ANS differs from the Somatic Nervous System (SNS) mainly in the three following areas: effectors, efferent pathways and target organ responses. The detailed differences between the ANS and the SNS are shown in Table 2-1.

The ANS is classically divided into two subsystems: the Parasympathetic Nervous System (PSNS) and the Sympathetic Nervous System (SNS) [14]. “The Sympathetic Nervous System (SNS) originates in the thoracic and lumbar regions of the spinal cord, and its function is to mobilize the body's resources under stress [14].” The sympathetic fibers, which innervate tissues in many organs, provide regulatory functions, such as increasing heart rate, respiratory rates, sweating, cardiac output, pupil size and inhibiting digestion. The SNS is the “fight-or-flight” system, which involves E activities (exercise, excitement, emergency, and embarrassment). On the other hand, “the PSNS originates in the brain stem and the lower part of the spinal cord, and its function is associated with the relaxation of the body and its activation promotes a return of several organs to regular function [14].” The functions caused by the stimulation of PSNS may slow the heart rate, respiratory rate, promote peristalsis, increase salivary secretions, and constrict the pupil size and so on. The PSNS involves D activities including digestion, defecation, and dieresis.
Most of the viscera receive nerve fibers from both parasympathetic (PSNS) and sympathetic (SNS) divisions, which do not innervate an organ equally. The SNS and the PSNS systems may work jointly or independently to control a given stage of a complex process. The SNS and the PSNS may present dual innervations acting on certain effectors including antagonistic effects, which means the two nerve divisions oppose each other such as heart rate increase (sympathetic) and heart rate decrease (parasympathetic), and unified effect such as salivary mucous cell secretion (both sympathetic and parasympathetic). In addition, the SNS and the PSNS may also act on certain effectors

Table 2-1: The difference between the ANS and the SNS

<table>
<thead>
<tr>
<th>Somatic Nervous System</th>
<th>Automatic Nervous System</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effectors</strong></td>
<td>Skeletal muscle</td>
</tr>
<tr>
<td><strong>Efferent Pathways</strong></td>
<td>Heavily myelinated axons of the somatic motor neurons extend from the central nervous system to the effectors</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Target Organ Responses</strong></td>
<td>All somatic motor neurons release Acetylcholine (ACh), which has an excitatory effect</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
without dual innervations. For example, the adrenal medulla, the arrector pili muscles, the sweat glands and many blood vessels only receive sympathetic division influences. The SNS and PSNS systems are shown in the Figure 2-3 [15].

2.4 Physiological Signals

Half a century ago, John Lacey (1959) [16] envisioned psychophysiological assessments as providing an objective means for evaluating psychotherapeutic process and outcome. In 1990, a model for the relationship between physiological events and psychological or behavioral processes was proposed by Cacioppo and Tassinary [17], which maps physiological variables to an emotion space. Modeling of the relationship between physiology and emotions is essential for the affective assessment based on the ANS patterns.

Figure 2-3: Pathway of the autonomic nervous system
In our study, the PD signal was chosen to be monitored in a non-invasive way for affective sensing. The PD measurements were obtained with an eye gaze tracking system, which has become a robust and intuitive tool for human-computer interaction. Technological advances continue to extend the accuracy and affordability of these systems [18, 19]. The GSR signal was also measured and analyzed simultaneously for comparison purposes. GSR is one of the commonly used physiological signals to detect stress, lying, fear, anxiety and arousal as these events tend to make the sweat glands more active, which leads to lowering the skin's resistance [20 -23]. Hence, the changes of the GSR signal, which are caused by emotional sweating, can indicate the affective variation of the user.

2.4.1 Pupil Diameter

The pupil is a circular aperture at the center of the iris of the eye through which light passes towards the retina. The pupil of the human eye can constrict to 1.5 mm in diameter or dilate to about 8 to 9 mm [24]. “These constriction and dilation are controlled by two opposing sets of muscles in the iris, the constrictor and dilator pupillae, which are governed by the sympathetic and parasympathetic divisions of the ANS [24].” “The sympathetic ANS division, which is mediated by posterior hypothalamic nuclei, causes the enlargement of the pupil through direct stimulation of the dilator muscles, which produces them to contract [25].” The excitation of the circular pupillary constrictor muscles innervated by fibers from the parasympathetic division causes pupil constriction [26].
Therefore, if the sympathetic division of the ANS is activated (e.g., due to stress), the pupil diameter tends to increase; whereas if the parasympathetic division of the ANS dominates (e.g., during relaxation), the pupil diameter will remain comparatively small [27]. Figure 2-4 shows the dual innervations of the iris by the ANS [28].

2.4.2 Galvanic Skin Response

Galvanic skin response (GSR) is the method of measuring the electrical conductance of the skin, which can represent the variation of the moisture levels of the body. The skin contains three layered adaptive organs, which are epidermis, dermis and subdermis, helping maintain the body’s water balance and the core temperature. Blood vessels and the eccrine sweat glands carry out these functions. The constriction and the dilation of the blood vessels regulate the blood flow to the surface of the body for heat conservation and loss. The eccrine sweat glands cool the body by depositing moisture on
the skin surface through the ducts in the dermis and sweat pores in the epidermis. According to the Edelberg’s model, when the sweat ducts in the epidermis fill, phasic changes in the skin conductance occur. When the moisture is reabsorbed by the sweat glands or deposited on the skin, skin conductance returns to its baseline levels [29]. Sweat gland activation is a physiological survival mechanism, which is controlled by the brain via the sympathetic division of the ANS [29].

While the main focus of this research is the use of features derived from the PD signal for affective assessment, the GSR signal was also recorded and analyzed during the experiment, and was considered as a comparative information source for the classification of segments. So, this study evaluated the affective assessment performance of three features derived from the pupil diameter signal, with respect to three features obtained from the GSR signal. To derive meaningful features from the GSR signals recorded, the three characteristic types of skin conductance variations were considered: Tonic, spontaneous and phasic. Tonic skin response is the basal activity, which represents the baseline level of skin conductance in the absence of any particular discrete environmental event. Spontaneous activity of GSR is in reality a change in physiological activity that occurs in the absence of any known stimuli. The typical frequency of spontaneous activity response is between one and three transient GSR increases or “responses” per minute. Phasic skin conductance responses are generated as a reaction to stimuli and can be quantified by parameters such as amplitude, latency, rising time, and half recovery time [29-30].
2.5 Overview of Stress

2.5.1 The Concept of Stress

“Stress” is a common used term today both in biology and psychology, which describes the negative concept that can have an impact on one’s mental and physical well-beings. When the person is under stress, the sympathetic nervous system becomes primarily active and regulates the body’s physiological functions to help it adapt to its environment. The threat, event or change promoting the adaptation is commonly called stressors. Stressors can be internal (thoughts, beliefs, attitude) or external (loss, tragedy and change). The physiological changes caused by stress are needed for the “fight or flight” response. In order to be able to either attack (fight) or run away (flight), energy is required. Therefore, nutrients in the blood broke down, so that the blood glucose level increases. For this to happen the heart has to pump with increased force and the blood pressure is raised. To make more blood available to active areas, other changes take place in the blood vessels in the skin, kidneys, and digestive tract, which change diameter. The pupil of the eye dilates allowing more clear vision. If the person is under repeated stress situations, this can have serious health consequences for example, the hormone levels will change and the immune system will be weakened, which will increase the chance of heart attack and stroke, anxiety disorders and so on. 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 [31].

Lazarus and Folkman’s model the stress response includes three components: physiological component (arousal and hormone secretion), emotional component (anxiety,
fear, grief and resentment) and behavioral component (coping strategies) [32]. Therefore, the stress can not only arouse physiological change but also psychological change and the two changes interact with each other.

There are many applications for stress detection such as satisfying the demand of security for commerce, electronic voting, banking and so on. In addition, automated stress detection from analysis of physiological signals may help advance the enhancement of the polygraph (lie detector), based on the stress experienced by someone who lies [33] (To date, the operation of the polygraph typically requires a human expert).

2.5.2 Literature Review of Stress Detection

There are many different approaches to detect the stress affect in the computer user. Andren & Funk analyzed the manner and rhythm in which a person types characters on a keyboard to estimate the stress level [34]. Dinges et al. proposed a solution to the problem of stress detection based on facial recognition [35]. Begum et al. studied stress detection based exclusively on Finger Temperature (FT) [36]. Picard & Healey used Electrocardiogram (ECG), Breath Rate (BR), Electromyogram (EMG) and Electroencephalogram (EEG) to detect stress [37]. In our study, Galvanic Skin Response (GSR) was also simultaneously monitored and processed for comparison purposes with PD. Picard & Healey (2000), Moore & Dua (2004), Zhai et al. (2005), Lin et al. (2005) and Alberto de Santos Sierra (2011) all used the GSR signal to detect stress in the subjects during their experiments [37-41]. Therefore, from consideration of previous research, it is clear that GSR is one kind of widely used physiological signals to detect the affective state of the computer user.
In addition, it is also of importance to develop proper approaches to elicit stress in individuals. Partala and Surakka [42] presented the study of affective sensing, via a specific psychophysiological measure, using the driver and pilot simulation to elicit different affective status including stress. Lin et al. implemented video games to provoke different emotions (including stress) and found that the physiological data correlated with task performance data in a video game [39]. Alberto De Santos Sierra at al. used talk preparation and hyperventilation to provoke the alternation in physiological parameters together with different emotional experiences [43].

In our study, the “Stroop Color-Word Test” (SCWT) [44] was used to elicit mild mental stress in the experimental subjects during controlled intervals at known times. The approval of the FIU Internal Review Board (IRB) for experimentation with human subjects was sought and obtained (shown in Appendix A). Renaud and Blondin [45] have demonstrated the efficiency of this stress elicitation method. The implementation of this stimulus program and any other details are described in the following chapter.

2.6 Summary

In this chapter, the concept of the affective computing is presented first and then its general models and applications are outlined. In addition, the autonomic nervous system and its potential role in affective computing are also described. Moreover, the physiological signal background used for affect assessment is discussed and a detailed description of the signals monitored in the project (PD and GSR) is provided as well. Finally, the literature about the concept of stress and the methods of stress detection was reviewed.
CHAPTER 3
EXPERIMENT DESIGN

The main purpose of this research is to develop off-line and on-line algorithms to achieve the affective assessment of the computer user (“stress” vs. “relaxation”) by the analysis of measured PD signals from the subject. In order to evaluate the efficiency and the robustness of the PD signal to distinguish the affective state of “stress” from “relaxation”, GSR and IL (illumination of the environment) signals are also simultaneously measured. In summary, the designed experimental setup and protocols must satisfy the following criteria:

(i) An appropriate stimulus must be provide in order to elicit “stress” affect in the participating subjects.
(ii) All the signals must be collected simultaneously in order to make the comparisons of the different physiological signals for affective assessment
(iii) Appropriate illumination (IL) changes, which affect the size of pupil, are also necessary

The whole experimental implementation, including the software and the hardware, is described below.

3.1 Software Development

The “Stroop Color-Word Test” (SCWT) was used to elicit mild mental stress in the experimental subjects during controlled intervals for the purpose of observing the changes in the PD and GSR signals and their correlation to the affective states of “stress” and “relaxation”. The SCWT has been accepted by several research groups as a viable
stress stimulus. Tulen et al. [46] 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. [47] studied the sympatho-adrenal and hemodynamic responses to mental stress induced by the SCWT, which increased heart rate and blood pressure by 28 beats/min and 29/14 mmHg for the subjects, on average. Further, Feng-Tso Sun et al. [48] 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.

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 Figure 3-1. 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 would automatically display the next word. Figure 3-2 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 (including both the off-line and on-line methods) are able to signal the emergence of stress.
Figure 3-1: 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 3-2: Stimuli schedule of the experimental protocol.

The segments in each section are:
• ‘IS’ – the Introductory Segment to let the subject get used to the task environment. According to the Law of Initial Values (LIV) [49], the purpose of this part is to establish an appropriate initial level for his/her psychological state;

• ‘C’ – the Congruent segment, containing 45 Stroop Congruent color word presentations, in which the font color matches the meaning of the word. These sections are not expected to elicit significant stress in the subject;

• ‘IC’ – the Incongruent segment, containing 30 presented words, in which the font color and the meaning differ. These sections are expected to induce stress in the subject;

• ‘RS’ – a Resting Segment, lasting one minute, which lets the subject return to a baseline affective state. In addition, the subject does not need to perform any action during this segment. The physiological signals are also monitored during this segment.

The binary numbers (01, 10 or 11, respectively at the beginning of each C, IC or RS segment) shown in Figure 3-2 represent the de-multiplexed two channel output used in the system to time-stamp the recorded physiological signals. The goal of this research was to investigate the variation of the PD signal for affective assessment, even in the presence of different light intensities in the environment surrounding the subject. Therefore, the illumination intensity was temporarily increased during the IC2 and C3 segments, which were marked as “VI” in Figure 3-2, in order to investigate the robustness of PD signal for affective assessment (“relaxation” vs. “stress”). Table 3-1 shows the four kinds of the experimental situations during the whole procedure.
3.2 Hardware Development

The purpose of the hardware setup designed for this experiment is to accurately, simultaneously and continuously to record the physiological signals (including PD and GSR) from the subjects and also the IL signal variation from the surroundings (the illumination is changed temporarily during IC2 and C3 segments). In addition, the necessary time markers must be recorded for identifying different segments or sections during the SCWT.

In our experiment, a multi-channel MCC DAQ system (PCI-DAS6023 board) is configured to sample the amplified analog GSR and IL signals at the rate of 360 samples/second. The PCI-DAS6023 provides 16 channels of 12-bit analog input, 8 digital I/O lines and two 16-bit counter timers. The analog inputs on each board can be configured as either eight differential or 16 single-ended channels. The input ranges are bipolar, in four ranges of ±10 V, ±5 V, ±500 mV, and ±50 mV. The ranges are software-selectable [50]. In addition, the PCI-DAS6023 also provides the software (TracerDAQ®) for acquiring and displaying data and generating signals. Before the experiment, the subjects are asked to test the GSR and IL sensors through the

Table 3-1: Four kinds of experimental situation

<table>
<thead>
<tr>
<th>Situation</th>
<th>Segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relaxed States &amp; Constant Illumination</td>
<td>C1&amp;C2</td>
</tr>
<tr>
<td>Stressed States &amp; Constant Illumination</td>
<td>IC1&amp;IC3</td>
</tr>
<tr>
<td>Relaxed States &amp; Varied Illumination</td>
<td>C3</td>
</tr>
<tr>
<td>Stressed States &amp; Varied Illumination</td>
<td>IC2</td>
</tr>
</tbody>
</table>
TracerDAQ® application in order to confirm that they are being recorded correctly. The testing signal tracing is illustrated in Figure 3-3.

![Figure 3-3: The testing signal tracing in TracerDAQ®](image)

CH0 and CH3 indicate the IL and the GSR signal variations, respectively. During the signal test, the surrounding light intensity is increased and decreased in succession, which results in the trace deflections seen in Figure 3-3 (CH0). In addition, the subject is required to tap the metal electrode of the GSR sensor frequently to observe the change of the signal, which can also be appreciated in Figure 3-3 (CH3). Therefore, the operation of both sensors (IL and GSR) can be verified before the actual recording session starts.
3.2.1 PD Measurement System

In our experiment, a desk-mounted eye tracking system (TOBII T60), which can measure the pupil diameter (PD) of the computer user in real-time, is utilized as the device that generates the PD signal. Eye gaze tracking systems have become robust and intuitive tools for human-computer interaction. Technological advances continue to extend the accuracy and affordability of these systems. Figure 3-4 shows the image of the TOBII T60 eye tracker [51].

![TOBII T60 Eye Tracker](image)

Figure 3-4: TOBII T60

The TOBII T60 Eye Tracker is integrated in a 19-inch TFT monitor and can be set up easily and run on most Windows computers, which provides a highly portable eye tracking solution. In addition, TOBII T60 is very 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 accuracy and precision of this device ensures that the research results are reliable.

In our experiment, custom software, based on Microsoft Visual Studio 2005 and the C# language, is used to continuously measure the PD data from the TOBII T60. The
original programming modules are provided in the TOBII software development kit (SDK), which includes three main steps shown in Table 3-2.

Table 3-2: Three steps for eye tracking

<table>
<thead>
<tr>
<th>Step</th>
<th>Name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Confirm Track Status</td>
<td>Verify that the eye tracking system can detect both eyes of the subject</td>
</tr>
<tr>
<td>2</td>
<td>Calibrate</td>
<td>TOBII provides a 9-point, 5-point or 2-point calibration procedure, which can allow certain level of tolerance for small head movements of the subject</td>
</tr>
<tr>
<td>3</td>
<td>Track</td>
<td>Measure and record the data, which contain the pupil diameter (PD) and validity code</td>
</tr>
</tbody>
</table>

Figure 3-5 shows the TOBII T60 eye tracking interface. In the “Track Status” box, the right side small window shows the word “Both”, and two white dots, which indicate that the system has already found the two eyes of the computer users. In the second step (“Calibrate” box), a 9-point calibration item is chosen, so all (9) points are used to verify the initial accuracy of the tracker. The last step, which is the “Track” box, starts to track the eye of the computer user. The final output data of the raw PD signal (including eye blinks) is the average value of the pupil diameters of both eyes. In Figure 3-5, the olive green square shows the position where the computer user is looking in the screen.
3.2.2 GSR Measurement System

In the experiment, the GSR sensor is used for detecting the moisture of the palm of the hand. This sensor is shown in Figure 3-6. It works by simply laying two fingers across the electrodes. The sensor is produced by Thought Technology LTD (West Chazy, New York), and is combined with a custom “frequency to voltage converter” integrated circuit (LM2917N).

The device measures the electrical conductance between the two fingers. Active measuring involves sending a small amount of current through the body. Because of the response of the skin and muscle tissue to external and internal stimuli, the conductance can vary by several microsiemens. Fear, anger, stress and some other responses can lead
to the variation of the electrical conductance of the body. Then the measured conductivity level can be converted to an oscillation frequency and, eventually, to a proportional to voltage.

![GSR2 module for GSR measurement system](image)

Figure 3-6: GSR2 module for GSR measurement system

3.2.3 IL Measurement System

In order to detect the light variation of the experimental surroundings, which is one of the important factors affecting the size of the pupil, the Illumination (IL) Measurement System is used. A BS500B0F photo-diode (Sharp), which is a luminance meter, is placed on the forehead of the computer user to measure the illumination in the area around his/her eyes. Gao et al. used the circuit [52] (shown in Figure 3-7), which can linearly convert the luminance variation to voltage changes in the analog output signal [53]. In order to acquire the digital signal of illumination in real time, the output analog voltage signal was digitized at a frequency of 360Hz by the DAQ System to obtain the IL signal.
3.2.4 Synchronization of All the Measured Signals

In our experiment, it is essential to simultaneously and independently obtain the data from the DAQ System (PCI-DAS6023 board from Measurement Computing Co.), including GSR, IL and audio time markers, and the PD signal from the TOBII T60 eye tracking system. In order to achieve the synchronization of acquiring these signals, the “Start” and the “Stop” buttons in the “Track” section of the TOBII control program were modified for the purpose of beginning and ending the recording of the PD values of the subject (for the on-line affective assessment algorithm, these two buttons can also control MATLAB® to real time process the PD signal), and simultaneously starting and stopping the collection of samples of the GSR, IL and audio markers from the DAQ. Figure 3-8 shows the “Start” and “Stop” buttons (marked with red circles) in the “Track” section of the TOBII control program.
3.2.5 Overall Setup of the Experiment

The complete instrumental setup used in our experiments is shown in Figure 3-9. The setup included measurement of the PD, GSR, IL signals and audio time markers. The 19” display integrated in the TOBII T60 eye tracker was used to display the words to the subjects. For the off-line affective assessment algorithm, the TOBII system measured the pupil diameter values 60 times per second and later they were read into MATLAB® for further analysis. For the on-line affective assessment algorithm, MATLAB® can real-time process the measured pupil diameter in the background processing.

In the experiment, each participant sat in front of the TOBII screen, interacting with the “Stroop Test” program for about 30 minutes, wearing the GSR sensor (GSR 2,
from Thought Technology) on the middle and ring fingers, and the IL sensor (BS500B0F photo-diode, from Sharp), on his/her forehead at the same time. These two signals (GSR and IL), together with the left and right audio output (to provide the corresponding time stamping in the experiment) were recorded and converted to a MATLAB® file at the rate of 360 samples/second, using a multi-channel MCC DAQ system (PCI-DAS6023 board). (The audio signals were not heard by the subjects). Later, GSR, IL and both audio output channels were down-sampled to 60 Hz, to match the PD sampling rate.

![Complete instrumental setup](image)

Figure 3-9: Complete instrumental setup

Figure 3-10 shows one subject wearing GSR and IL sensors and sitting in front of the TOBII T60 screen to participate in our experiment.

Figure 3-11 shows a set of signals (PD, GSR and IL) recorded from the experiment.
Figure 3-10: An example of participant in the experiment

Figure 3-11: A set of signals after synchronization. From top to bottom: the raw PD, the raw GSR and IL.
3.3 Experimental Procedure

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

Each participant first interacted with the “Stroop test” program under the normal lights in the room. In order to judge the robustness of our affective assessment algorithm with the interference of the illumination intensity variation in the environment, an additional level of the illumination provided by a desk lamp placed above the eye level of the subject was switched ON briefly during the IC2 and C3 segments (as shown in Figure 3-2). This introduced intervals of higher illumination in the experiment, which would trigger the pupillary light reflex (i.e., the pupil diameter changes in response to light intensity variations of the environment. In this case, the pupil diameter would tend to decrease in response to the increase of the light intensity). 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.

3.4 Summary

This chapter described the software and hardware development in the instrumental setup for eliciting stress in the experiment. In this instrumental setup, the PD, GSR and IL signal are simultaneously recorded for analysis. The PD signal is measured by a TOBII T60 eye tracking system, which is capable of accurately detecting the pupil diameter of both eyes of the subject. The GSR signal is also recorded, to enable a comparison of the stress recognition performance, in contrast with the PD signal. The IL
signal is also measured in order to investigate the robustness of the PD signal for affective assessment, with respect to illumination changes. Finally, the experimental procedure and the pool of the subjects were described.
CHAPTER 4

METHODOLOGY

The purpose of this research is to develop both an off-line and an on-line systems for detecting the stress affect of computer users. The Pupil Diameter (PD) seems to be an indication of the affective state, according to previous research, but this had not been investigated fully. In the off-line affective assessment algorithm, wavelet denoising and Kalman filtering were used to preprocess the PD signal. Then three features were extracted from it and five classification algorithms were used to evaluate the overall performance of the identification of “stress” states in the computer users. The Galvanic Skin Response (GSR) signal was also analyzed to study the comparative efficiency of affective sensing through the PD signal. The discriminating power of the three features derived from the preprocessed PD signal was compared to three features derived from the preprocessed GSR signal, in terms of their Receiver Operating Characteristic curves. For the on-line affective assessment, a hard threshold is implemented first in order to remove the eye blink of the PD and then a moving average window is utilized to obtain the representative value PD_r for every one second of the raw PD signal. There are three main steps in the on-line affective assessment algorithm, which are preparation, feature-based decision voting and affective determination, in three phases. The overall results confirm that the PD signal should be considered as one of the most powerful physiological signals to involve in the future automated affective recognition systems, especially for detecting the “relaxation” vs. “stress” states.
4.1 Off-line Affective Assessment

4.1.1 Physiological Signal Preprocessing

The main processing methods applied to the raw PD signal are wavelet denoising and Kalman filtering. However, because of the presence of eye blinks (signaled by a value of “4” in the validity code given by the TOBII system and identified as sudden transitions to a false PD value of zero), linear interpolation was implemented first to compensate the interruption in the PD data by each blink.

Although linear interpolation removes the blinking artifacts, it should be noted that there is still a substantial amount of fast variability that is not likely to represent pupil size changes due to affective variations. In fact, previous research [54] has indicated that these fast variations in the PD signal could be due to quantization noise in the pupil diameter measurement.

The complete PD signal recorded during an experiment is showed in Figure 3-11. The vertical lines are the segment transition boundaries. The most important boundaries separate each congruent Stroop segment (C) from the incongruent Stroop segment (IC) that follows. On account of the noise included in the raw PD signal, it is obvious that it would be difficult to successfully classify the segments by using a threshold on the raw PD data. Therefore, our goal in this study was to process the PD signal to retain primarily components that are derived from affective changes in the user, removing the noise components.
4.1.1.1 Wavelet Denoising

The first section of the two-stage signal filtering approach we propose uses the wavelet denoising method, which works better in preserving the shape of the real signal, especially for signals with abrupt changes [55], [56].

Wavelets are functions that “wave” above and below the x-axis, having the following properties: (1) varying frequency; (2) limited duration; (3) an average value of zero. Like the sines and cosines in Fourier Transform (FT), wavelets are used as the basis functions $\psi_k(t)$ in presenting other functions:

\[
 f(t) = \sum_k \alpha_k \psi_k(t) \quad (4.1)
\]

Once the mother wavelet $\psi(t)$ is defined, a basis can be formed from this fixed mother wavelet by applying translation and scaling:

\[
 \psi(s, \tau, t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right) \quad (4.2)
\]

If we take the special values for $s$ and $\tau$ in defining the wavelet basis: $s = 2^{-m}$ and $\tau = n \cdot 2^{-m}$:

\[
 \psi(s, \tau, t) = \frac{1}{\sqrt{2^{-m}}} \psi\left(\frac{t-n\cdot2^{-m}}{2^{-m}}\right) = 2^m \psi(2^m \cdot t - n) \quad (4.3)
\]

The Continuous Wavelet Transform (CWT) can be presented as

\[
 C(\tau, s) = \frac{1}{\sqrt{s}} \int f(t) \psi^*\left(\frac{t-\tau}{s}\right)dt \quad (4.4)
\]

$\tau$ is the translation parameter, which is the measurement of time and $s$ is the scale parameter, which is the measurement of the frequency. Therefore, the signal can be presented by the constituent wavelets of different scales and positions. The advantage of
the CWT is simultaneous localization in time and scale: the location of wavelet allows to explicitly represent the location events in time; the shape of the wavelet allows to explicitly represent different detail or resolution.

From equation 4.3 and 4.4, it is easy to infer the Discrete Wavelet Transform (DWT), which is defined as:

\[ DWT(m, n) = \int_{-\infty}^{\infty} f(t) \psi_{m,n}(t) dt \]  

(4.5)

Because the DWT has orthonormal properties, there is no information redundancy in the transform. In addition, the choice of \( m \) and \( n \) provides a way to decompose the signal into scales with different time and frequency resolution, to perform multiresolution analysis (MRA).

The key concept in MRA is that of vector spaces. For each vector space, there is another vector space of higher resolution, recursively, until the space with the highest possible resolution, which implies that each vector space contains all vector spaces that are of lower resolution. The nested vector space concept (Figure 4-1) is shown below:

![Figure 4-1: The nested vector space for MRA](image)

Define \( V_j \), which are nested inside each other, to be the set of the signal \( f \). These spaces have the relationship as follows:
Given \( f \in L^2(R) \), and if we do the multiresolution approximation of \( L^2(R) \), we can get the \( f_j \in V_j \). Another two relationships between \( V_j \) are as follows:

\[
\bigcup_{j=-\infty}^{\infty} V_j \quad \text{dense} \quad L^2(R) \tag{4.7}
\]

\[
\bigcap_{j=-\infty}^{\infty} V_j = \{0\} \tag{4.8}
\]

Define the space \( W_j \) as the difference (in the subspace sense) between adjacent spaces \( V_j \) and \( V_{j+1} \). Therefore, we can write:

\[
V_{j+1} = W_j + V_j \tag{4.9}
\]

Finally, we can state, for \( V_j \):

\[
V_j = V_0 \oplus W_0 \oplus W_1 \oplus W_2 \oplus \cdots \oplus W_{j-1} \tag{4.10}
\]

The wavelet, \( \psi(t) \), has a scaling function \( \phi(t) \), which can produce the multiresolution subspace \( V_j \) as follows:

\[
\phi_{m,n}(t) = \sqrt{2^m} \phi(2^m * t - n) \tag{4.11}
\]

It is clear that space \( V_j \) can be obtained from the \( V_0 \), by time compression or dilation by powers of 2, therefore, we only needs to check space \( V_0 \).

There is a significant equation, which is called the “two-scale equation”, connecting the scale function to itself at two different time scales, which gives rise to one of the filters. The discrete time filter coefficients are \( h(n) \) and the equation is as follows:

\[
\phi(t) = \sum_n h(n) \sqrt{2} \phi(2t - n) \tag{4.12}
\]
Because $W_0$ is also the subset of $V_1$, another two-scale equation for the wavelet can lead to another filter $g(n)$.

$$\psi(t) = \sum_n g(n)\sqrt{2}\phi(2t - n)$$  (4.13)

Since any signal can be approximately presented by the space $V_j$, if choosing a large enough value of $j=m$, and projecting the signal into $V_j$ using the basis $\phi_{m,n}(t)$, we can get the following equation:

$$A^0(n) = \int_{-\infty}^{\infty} f(t)\phi_{m,n}(t)dt$$  (4.14)

From Equation 4.14, we can recover the signal as:

$$f(t) = \sum_n A^0(n)\phi_{m,n}(t)$$  (4.15)

From Equation 4.10, it is easy to infer that the original signal in space $V_j$ can be decomposed into the subspaces $V_{j-k}$ and $W_{j-k}$ with their bases $\phi_{j-k,n}(t)$ and $\psi_{j-k,n}(t)$.

$$f(t) = \sum_n A^0(n)\phi_{m,n}(t)
= \sum_k A^1(k)\phi_{m-1,k}(t) + \sum_k D^1(k)\psi_{m-1,k}(t)
= A_1(t) + D_1(t)$$  (4.16)

We define $A_1(t)$ and $D_1(t)$ as the approximation and detail at level 1 and coefficients $A^l$ and $D^l$ are the approximation and detail coefficients at level 1.

$$f(t) = A_1(t) + D_1(t)
= \sum_l A^2(l)\phi_{m-2,l}(t) + \sum_l D^2(l)\psi_{m-2,l}(t) + \sum_k D^1(k)\psi_{m-2,k}(t)
= A_2(t) + D_2(t) + D_1(t)$$  (4.17)
We define $A_2(t)$ and $D_2(t)$ as the approximation and detail at level 2 and coefficients $A^2$ and $D^2$ are the approximation and detail coefficients at level 2.

Since the wavelets and the scales functions at each level are orthogonal, we can compute the coefficient $A^l$ and $D^l$ by the inner product formula:

$$A^l(n) = \langle f(t), \phi_{m-1,n}(t) \rangle$$
$$= \langle \sum_k A^0(k) \phi_{m,k}(t), \phi_{m-1,n}(t) \rangle$$
$$= \sum_k A^0(k) \langle \phi_{m,k}(t), \phi_{m-1,n}(t) \rangle$$ (4.18)

Because $\langle \phi_{m,k}(t), \phi_{m-1,n}(t) \rangle = h(k-2n)$ according to the orthogonal property, we can easily complete the computation:

$$A^l(n) = \sum_k h(k-2n) A^0(k)$$ (4.19)

The detail coefficients can be computed similarly, and we can get

$$D^l(n) = \sum_k g(k-2n) A^0(k)$$ (4.20)

There are mainly three steps for wavelet denoising. The first sub-step is to apply the wavelet transform to the noisy signal to produce the noisy wavelet coefficients (approximation coefficients and detail coefficients) to a level in which the noise can be separated and removed. The second sub-step is to select an appropriate threshold value at each level and apply a threshold method to remove the noise by altering the values of the detail coefficients. The last sub-step is to perform inverse wavelet transformation on the approximation coefficients and the altered detail coefficients to obtain a denoised signal [57].

First, the incoming signal $f(t)$ is passed through the low and high-pass filters and
its scale is changed by down sampling operations. The low-pass filter and the high-pass filter form a quadrature mirror filter, which satisfies the following relation:

\[ g(L - 1 - n) = (-1)^n \cdot h(n) \]  

(4.21)

where \( g(n) \) is the high-pass filter and \( h(n) \) is the low-pass filter. \( L \) is the filter length.

\[ g(n) \]
\[ h(n) \]
\[ L \]

\[ A^0 \]
\[ A^1 \]
\[ A^2 \]

\[ D^1 \]
\[ D^2 \]

Figure 4-2: Multi-resolution wavelet decomposition

In Figure 4-2, the identity of the processing blocks is: \( h \): low-pass decomposition filter; \( g \): high-pass decomposition filter; \( \downarrow \): down-sampling operation. \( A^1 \), \( A^2 \) are the approximation coefficients of the original signal at levels 1, 2. \( D^1 \), \( D^2 \) are the detail coefficients at level 1, 2.

The next step is to change the values of the detail coefficients by selecting an appropriate threshold limit at each level and the threshold method to best remove the noise. In our research, the Birge-Massart strategy, a method based on adaptive functional estimation in regression or density contexts, is applied as the wavelet coefficients selection rule to set the level-dependent threshold for denoising. The specific steps for the Birge-Massart strategy are described below [58]:

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 \( n(i) = M/(j+2-i)^\alpha \)

Furthermore, an appropriate choice of the threshold value is also necessary for the effectiveness of the described wavelet denoising procedure. A threshold that is too high might remove important parts of the underlying signals whereas a threshold that is too low may still retain part of the noise in the reconstruction [59]. Then, any coefficient of the \( i \)-th level that is less than or equal to the designed threshold is set to zero.

The last part of the procedure involves re-assembling the signal, following a reverse order of levels with respect to the decomposition, using the approximation coefficients and the altered detail coefficients to obtain the denoised signal, as shown in Figure 4-3.

![Figure 4-3: Multi-resolution wavelet reconstruction.](image)

In Figure 4-3, the identity of the processing blocks is: \( h' \): low-pass reconstruction filter; \( g' \): high-pass reconstruction filter; \( \oplus \): up-sampling operation. \( A'^1, A'^2 \) are the processed or non processed approximation coefficients of the original signal at levels 1, 2. \( D'^1, D'^2 \) are the processed or non-processed detail coefficients at level 1, 2.
4.1.1.2 Kalman filtering

Kalman filtering is an iterative computational algorithm designed to improve noisy measurements, current state estimate and calculate the forecasts for time series models [60-61]. The Kalman filter is constructed as a mean squared error minimizer, whose weights in the update rules are chosen to ensure that the forecast variances are minimized. Therefore, Kalman filter is an optimal estimator, which infers parameters of interests from the indirect, inaccurate and uncertain observations. The process includes prediction and correction applied recursively to a process represented by the equations:

\[
S(k) = A(k)S(k-1) + w_1(k-1)
\]  \hspace{1cm} (4.22)

where \(S(k)\) is the state vector of the process at time \(k, (n\times1)\); \(A(k)\) is the state transition matrix of the process from state \(k-1\) to state \(k, (n\times m)\); and \(w_1(k-1)\) is the associated white noise process with known variance,\((n\times1)\), and whose covariance matrix is denoted as \(Q\).

Observations on this variable can be modeled as:

\[
X(k) = C(k)S(k) + w(k)
\]  \hspace{1cm} (4.23)

where \(X(k)\) is the actual measurement of \(S\) at time \(k, (m\times1)\); \(C(k)\) is the noiseless connection between the state vector and the measurement vector, \((m\times n)\); and \(w(k)\) is the associated measurement error and its covariance matrix, denoted as \(R\). Figure 4-4 shows a model of the Kalman filter.
Define $\hat{S}(k)$ as the optimal estimation and $\hat{S}'(k)$ is the prior estimate of $\hat{S}(k)$ at time $k$. Therefore, we can get the following two equations:

$$\hat{S}'(k) = A(k)\hat{S}(k - 1) \quad (4.24)$$

$$\hat{X}'(k) = C(k)\hat{S}'(k) = C(k)A(k)\hat{S}(k - 1) \quad (4.25)$$

Since the difference of the observation and the estimation values represent a deviation, we define the difference as the measurement residual:

$$\tilde{X}(k) = X(k) - \hat{X}'(k) \quad (4.26)$$

If $w_1(k)$ is replaced by the $\tilde{X}(k)$ multiplying the correction matrix $H(k)$, the equation 4.22 can be written as:

$$\hat{S}(k) = A(k)\hat{S}(k - 1) + H(k)\tilde{X}(k)$$

$$= A(k)\hat{S}(k - 1) + H(k)\left[ X(k) - C(k)A(k)\hat{S}(k - 1) \right] \quad (4.27)$$

Therefore, we can build the one-step recursive Kalman model based on equation 4.24-4.27 in Figure 4-5. It should be noted that in the study, the input signal $X(k)$ is the PD signal after wavelet denoising and the output signal $\hat{S}(k)$ is the PD signal after Kalman filtering.
From equation 4.22, 4.23 and 4.27, we can get the following equation:

\[
\hat{S}(k) = A(k)\hat{S}(k-1) + H(K)[C(k)S(k) + w(k) - C(k)A(k)\hat{S}(k-1)]
\]

\[
= A(k)\hat{S}(k-1) + H(K)[A(k)\hat{S}(k-1) + w(k) - C(k)A(k)\hat{S}(k-1)] + C(k)A(k)S(k) + w(k) - C(k)A(k)S(k-1)
\]

\[
= A(k)S(k-1) + H(K)C(k)A(k)S(k-1) + w(k) - C(k)A(k)S(k-1)
\]

(4.28)

Assume the real value \(S(k)\) and its estimation \(\hat{S}(k)\) have the difference:

\[
\hat{S}(k) = S(k) - \hat{S}(k)
\]

\[
= \hat{S}(k-1) - \hat{S}(k) + w(k-1) - H(k)w(k)
\]

(4.29)

and its error covariance matrix is as follows:

\[
\tau(k) = E[S(k) - \hat{S}(k)\hat{S}(k)\hat{S}(k)\hat{S}(k)]
\]

(4.30)

In order to make the calculation easily, we define:

\[
\tau'(k) = E[S(k) - \hat{S}(k)(S(k)-\hat{S}(k))]
\]

(4.31)

Therefore, we can get

\[
\tau'(k) = E[(A(k)S(k) + w(k)) - A(k)\hat{S}(k)] + Q(k-1)
\]

(4.32)
\[ \varepsilon(k) = \left[ I - H(k)C(k) \right] \varepsilon(k)' \left[ I - H(k)C(k) \right]' + H(k)R(k)H(k)' \]
\[ = \varepsilon(k)' - H(k)C(k) \varepsilon(k)' + H(k) \left[ C(k)\varepsilon(k)'C(k)' + R(k) \right]H(k)' \]

(4.33)

If we define
\[ C(k)\varepsilon(k)'C(k)' + R(k) = SS' \]
(4.34)

Then we can get
\[ U = \varepsilon(k)'C(k)' \]
(4.35)

\[ \varepsilon(k) = \varepsilon(k)' - H(k)U - UH(k)' + H(k)SS'H(k)' \]
\[ = \varepsilon(k)' - U(SS')^{-1}U' + \left[ H(k)S - U(S')^{-1} \right]\left[ H(k)S - U(S')^{-1} \right]' \]

(4.36)

In order to minimize \( \varepsilon(k) \), \( H(k)S - U(S')^{-1} = 0 \), so

\[ H(k) = \varepsilon(k)'C(k)' \left[ C(k)\varepsilon(k)'C(k)' + R(k) \right]^{-1} \]
(4.37)

\[ \varepsilon(k) = \left[ I - H(k)C(k) \right] \varepsilon(k)' \]
(4.38)

\[ \hat{S}(k) = A(k)\hat{S}(k - 1) + H(K) \left[ X(k) - C(k)A(k)\hat{S}(k - 1) \right] \]
(4.39)

The total predictor-corrector algorithm for implementing this complete process, numerically, is represented in Figure 4-6.
4.1.2 Data Normalization and Feature Extraction

Before extracting the features to identify “stress” states in the Incongruent Stroop segments and “relaxation” states in the Congruent Stroop segments, the preprocessed PD and GSR signals were normalized to the range of [-1,1] since the baselines of these physiological signals collected from different subjects were variable. The normalization formula used is:

\[ u(n) = 2 \cdot \frac{u(n) - u_{\text{min}}}{u_{\text{max}} - u_{\text{min}}} - 1 \quad n=1,2,\ldots,M \]  \hspace{1cm} (4.40)
where \( u(n) \) and \( u'(n) \) are the original and normalized signals, respectively.

From observation of the signals collected, it was apparent that the PD signal tends to rise steeply when the subject transitions into each of the “stressed” affective states (incongruent Stroop segments), which can be taken as one of the characteristics to differentiate the “relaxation” vs. “stress” affective states of the computer user. To capture these characteristic transitions as an indicator of a “stress” segment, the Walsh transform of the initial portion of each segment was obtained. The Walsh transform represents a time series as a set of coefficients associated with a set of mutually orthogonal basis functions that take on two possible values, +1 or -1, only. The Walsh transform can be used to process time domain signals of any form, but is particularly useful for signals which contain rectangular or similar shapes, which display fast transitions [62]. The Walsh transform has been successfully applied to the detection of the interictal spikes in EEG data [63] and the upstrokes in the carotid pulse wave [64]. The 1D Walsh transform function implemented is defined as:

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

(4.41)

where \( y(m) \) is the one-dimensional sequence to be transformed and \( b[k](u) \) is the kth bit in the binary representation of the index u. \( W(u) \) is the u-th Walsh coefficient derived from the Walsh transforms. Application of the Walsh transform decomposes each original signal into a set of orthogonal functions, and encodes the decomposition in the resulting Walsh coefficients, which are ordered in a progression called “sequency”. The first and last few coefficients in this ordering are said to be the “low frequent” and “high frequent” components of the PD signal, respectively [65].
In this study, eight consecutive were analyzed, with 100 samples each, from the beginning of each segment (both C and IC). The eight mean values of these windows are formed into a sequence to represent the trend of the PD signal during the beginning of the C and IC segments.

In this study, only the “low frequent” components of the PD signal (which represent the overall trend of the signal, rather than the details) are of interest. Therefore the feature sought to represent the fast increases of PD in the IC segments is extracted from only the first two Walsh coefficients. Specifically, only the difference between the first and the second Walsh coefficient during the onset period of each Stroop segment is utilized as a feature, and denoted as “PDWalsh”.

While the main focus of this study is the use of features derived from the PD signal for affective assessment, the GSR signal was also recorded during the experiment and was considered as an additional source of information for the classification of segments. So, this study evaluated the affective assessment performance of three features derived from the pupil diameter signal, with respect to three features obtained from the GSR signal. To derive meaningful features from the GSR signals recorded, the three characteristic types of skin conductance variations were considered: Tonic, spontaneous and phasic. Tonic skin response is the basal activity, which represents the baseline level of skin conductance in the absence of any particular discrete environmental event. Spontaneous activity of GSR is in reality a change in physiological activity that occurs in the absence of any known stimuli. The typical frequency of spontaneous activity response is between one and three transient GSR increases or “responses” per minute. Phasic skin
conductance responses are generated as a reaction to stimuli and can be quantified by parameters such as amplitude, latency, rising time, and half recovery time [66-67]. In this context, it was decided that the mean amplitude value of the GSR signal in the segment (“GSRmean”), the number of discrete GSR responses found in the segment (“GSRnum”) and the average of their rising times (“GSRrisingTime”) would be extracted from each segment (C or IC) as three features from the preprocessed GSR signal. Therefore, there were a total of six features obtained from each segment for the purpose of classification of the segment as “relaxed” (C segment) or “stressed” (IC segment). Table 4-1 shows all the features extracted from each of the signals recorded during each Stroop segment interval.

<table>
<thead>
<tr>
<th>Biosignal</th>
<th>Features</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD (3 features)</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 the PD signal during the onset of each Stroop segment</td>
</tr>
<tr>
<td>GSR (3 features)</td>
<td>GSRmean</td>
<td>Mean value of the amplitude of each GSR samples</td>
</tr>
<tr>
<td></td>
<td>GSRrisingTime</td>
<td>(Average) Rising time of each GSR in a segment</td>
</tr>
<tr>
<td></td>
<td>GSRnum</td>
<td>Number of the GSR in a segment</td>
</tr>
</tbody>
</table>

4.1.3 Classification Algorithm

In order to assess the affective state (relaxation vs. stress) of a computer user through the features derived from the physiological signals recorded in each segment,
machine learning techniques were utilized to determine the class membership of each data segment. Specifically, in this work, these classifiers were implemented using the Waikato Environment for Knowledge Analysis (WEKA) software, which can be freely downloaded from http://www.cs.waikato.ac.nz/ml/weka/. This software provides tools for preprocessing, classification, regression and clustering [68-69]. We approached the classification task using five different algorithms from WEKA, to observe the classification effectiveness and robustness of the features we extracted (listed in Table 4-1). A brief summary of these algorithms is presented below.

\( K^* \) measures the distance between two feature vectors or instances using the entropic theory, based on the probability of transforming one instance into another by randomly choosing between all possible transformations [69-70]. With such a measure, given a test instance \( m \) whose class is unknown, its probability of being in category \( C \) is calculated by summing the probabilities from \( m \) to each instance that is a member of \( C \), which is shown below:

\[
P^*(C | m) = \sum_{n \in C} P^*(n | m)
\]  

(4.42)

The probabilities for each category are calculated and the category with the highest probability is considered as the classification of the new instance. This method is a refinement of the k-nearest-neighbor rule [69].

\textit{Multilayer Perceptron}, which is a feedforward artificial neural network model, consists of the input, hidden and output layers. The purpose of the training process is to make the output from the neural network to match the actual target values as closely as possible by adjusting the set of weight values [69], [71].
**Naïve Bayes** implements the probabilistic Naïve Bayes classifier, where the normal distribution is assumed for numerical features [72]. The form of the model assumes conditional independence among features and the parameters are estimated from the training instances [69].

**Random Forests** are composed of many classification trees. The classification of a new object is realized by propagating the input vector down each of the trees in the forest. Each tree depends on the values of a random vector, which are sampled independently and with the same distribution for all trees in the forest. Then each tree gives a classification and the classification having the most votes over all the trees in the forest is chosen [69], [73].

**JRip** implements the RIPPER (Repeated Incremental Pruning to Produce Error Reduction) algorithm, whose basic strategy is to find a rule set that models the data, by first using the IREP (Incremental Reduced Error Pruning) rule-learning algorithm to find an initial model, and then to iteratively improve that model, using an optimization procedure [69], [74-75].

In addition, in order to obtain a more accurate and realistic assessment of the classifiers, a k-fold cross validation method [76] was used. First, the given data sets are separated into two parts: the training sets and a test set, where the labels of the test set are considered unknown in the classifier training. Then the accuracy on these sets can more precisely reflect the performance that can be expected in classifying unknown data [69], [77]. In our study, the original data was first divided into several equal subsets, and one subset was tested using the classifier trained on the remaining subsets. This procedure
was repeated until every subset had been used once for testing. The overall accuracy for the classifier is based on the average performance over the all classification runs.

To evaluate the usefulness of the PD signal and compare with the performance of the GSR signal, the classification stage of this study was performed in three different phases, as illustrated in Table 4-2.

Table 4-2: Three different conditions studied in the classification phase of this study

<table>
<thead>
<tr>
<th>Phase Number</th>
<th>Condition</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>Only using the features extracted from the GSR signal (3 features from GSR are used for classification)</td>
</tr>
<tr>
<td>P3</td>
<td>Only using the features extracted from the PD signal (3 features from PD are used for classification)</td>
</tr>
</tbody>
</table>

4.1.4 Receiver Operating Characteristic (ROC) Curve

To compare the levels of affective state discrimination power of each of the six features listed in Table 4-2, their individual Receiver Operating Characteristic (ROC) curves were plotted and the area under each curve was calculated, based on the data segments used in the study. The ROC Curve is a graphical plot of the true positive rate (i.e., the fraction of all positive cases correctly classified) vs. false positive rate (i.e., the fraction of negative cases incorrectly classified as positive), for a binary classifier obtained by adjusting its threshold.

The origin of the coordinate axes in the ROC plot represents the situation in which the classifier produces no detections (and no false alarms) because the threshold is
set at the highest possible value of the detection signal. As the threshold is lowered, the sensitivity is increased (the classifier detects more positive examples) but the false alarm level may also increase (the classifier may detect more false positives). When the threshold is set at the lowest possible value of the detection signal, the classifier claims all the positive cases right (true positive rate = 1) but all the negative cases wrong (false positive rate = 1). All the ROC curves begin at the coordinate origin (0, 0) and end at the upper-right corner (1, 1). The area under the ROC curve can indicate the potential discriminant power of the classification signal. A ROC curve that extends closer to the upper-left corner of the ROC space encompasses a larger area under the curve and indicates that the detection signal used to construct that ROC curve has a greater potential for classification. The diagonal dividing the ROC space equally is the curve for the performance of a “random classifier”, which produces uniformly distributed random numbers without any relation to the input. The ROC analysis has been used for a long time in many fields [78], and it now used in machine learning and data mining [79].

In our study, the goal was to compare the discriminant power of each feature derived from the PD and GSR signals. For each feature, half of the segments correspond to “stressed” states induced by IC Stroop stimulation (ideal classifier output = 1), while the other half are known to be associated with “relaxed” (C Stroop) intervals (ideal classifier output = -1). Each point of an ROC curve represents the performance of a classifier based on a single feature, for a specific hypothetical threshold. The x coordinate of each point in a ROC curve represents the portion of the “positive classifier outputs” that are in disagreement to the ideal output (-1) (false positive rate) and the y coordinate represents the portion of the “positive classifier outputs” that match the ideal
(1) (true positive rate) for each threshold value tested. We created the ROC curve for each feature with the ROC MATLAB® scripts by Dr. G. C. Cawley (http://theoval.cmp.uea.ac.uk/matlab/#roc). These scripts not only sweep the complete range of normalized threshold values, [0, 1], and draw the ROC, but additionally estimate a “convex hull” (ROCCH) that fits the actual ROC points calculated [80]. In the result and discussion chapter, the ROCCH, is shown with dashed lines in the resulting plots. The area under the ROCCH, denoted “AUH”, is of interest because it is theoretically possible to operate at any point on the convex hull of the points in a ROC curve [81].

Figure 4-7 provides an overview of the research methodology employed for this off-line algorithm.

![Figure 4-7: The algorithm of off-line affective assessment based on PD signal](image)

### 4.2 On-line Affective Assessment

#### 4.2.1 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 PDs, to achieve on-line affective assessment. In the on-line processing algorithm, the hard threshold setting for eye blink removal, which is denoted as “Thresholdblink”, and the moving average window
method are applied in order to obtain the PD\textsubscript{r}. The moving average is the most common method for reducing noise in the digital signal processing (DSP) field, because it is not only simple but also optimal for retaining a sharp step response [82], which is suitable for the shape of the PD signal when it changes from a Congruent (C) to an Incongruent segment (IC). 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 a time span of 1 second, were selected as window length to achieve one PD\textsubscript{r} value per interval. However, because of the presence of eye blinks, a hard threshold must be set (“Threshold\textsubscript{blink}”). If the computer detects the value of current PD sample is smaller than the “Threshold\textsubscript{blink}”, the present PD sampling data is not going to be counted as the valid value in the current one second time interval. Otherwise, the present PD sampling data will contribute to achieving one PD\textsubscript{r} value in the current one second time interval. The equations used to implement this process are:

1) \[ m = 0 \] (4.43)

2) If \( PD(i) \geq \text{Threshold}_{\text{blink}} \)
   \[ X(i) = PD(i); \quad m = m + 1 \quad (60* n - 59 \leq i \leq 60* n \quad n = 1, 2, \ldots) \] (4.44)
   If \( PD(i) < \text{Threshold}_{\text{blink}} \)
   \[ X(i) = 0; \quad (60* n - 59 \leq i \leq 60* n \quad n = 1, 2, \ldots) \] (4.45)

3) \[ PD_r(n) = \left[ X(60* n - 59) + X(60* n - 58) + \ldots + X(60* n) \right] / m \] (4.46)

4.2.2 Development of On-line Affective Assessment Algorithm

After capturing the subject’s PD signal and averaging valid numbers of interval samples to define one PD\textsubscript{r} value per second, the next step is the identification of the affective state of the human subject. The most important parts of the waveforms in the
PDr signal are the large sudden increase and the large sudden decrease, which appear in the affective state transitions 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 [83], it is almost impossible to examine the difference between the frequency domain of “relaxation” and the “stress” affective states for the PDr signal. Therefore, the accurate detection of the large PDr signal increase or its large decrease, which implies the subject’s affective state transition, seems to be an appropriate and feasible way for the realization of on-line PD based affective recognition system.

4.2.2.1 Modified Backward Differential Method for Sudden PDr Signal Change Detection

A finite difference approximate to the value of some derivative of a scalar function \( z(x) \) at a point \( x_0 \) in its domain, say \( z'(x_0) \), relies on a suitable combination of sampled function values at nearby points [84]. There are three different types of forms for the finite difference, which are forward, backward and central difference respectively. A forward difference is an expression of the form:

\[
\Delta[z](x) = z(x + \lambda) - z(x)
\]

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

\[
\nabla[z](x) = z(x) - z(x - \lambda)
\]

Finally, the central difference is presented by:

\[
\delta[z](x) = z(x + \lambda) - z(x - \lambda)
\]
\( \lambda \), the step size, is assumed to be small: \( |\lambda| \ll 1 \). The finite difference methods are usually applied to numerically approximate the differential equations at the given equally spaced discrete points. The first order differential expression \( \frac{\partial z(x)}{\partial x} \) for a function \( z(x) = \rho \) can be approximately presented by the forward, the backward or the central finite difference equivalence, which are shown respectively as follows:

\[
\begin{align*}
\frac{\partial z(x)}{\partial x} & \approx \frac{\rho_{i+1} - \rho_i}{\Delta x} \\
\frac{\partial z(x)}{\partial x} & \approx \frac{\rho_i - \rho_{i-1}}{\Delta x} \\
\frac{\partial z(x)}{\partial x} & \approx \frac{(\rho_{i+1} - \rho_i) + (\rho_i - \rho_{i-1})}{2\Delta x}
\end{align*}
\]

In this study, the goal is to implement the algorithm for an on-line affective recognition system, so it is evident that only the present and the previous PDr signal data can be used. Therefore, only equation (4.51), which is backward differentiation, is feasible. In the 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 (4.51) can be written as

\[
\frac{\partial z(x)}{\partial x} = \rho_i - \rho_{i-1}
\]

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

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

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 \( \nabla z_i \) operator \( (\rho_i - \rho_{i-1}) \) takes a high positive value; whereas if \( \rho_{i-1} \) is high and \( \rho_i \) is low, the \( \nabla z_i \) 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_t signal according to the result calculated from (4.54), which can further indicate the affective state change of the human subject.

The differential operator has been successfully used in the digital signal and image processing areas with the aim of detecting significant changes. For example, N. Paivinen et al; utilized first and second derivative of EEG signal to extract time domain features for automatic seizure detection [85]. S. Suppappol et al; [86] and M. Paoletti et al; [87] 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 [88].

However, the traditional kernel for backward differentiation, equation 4.54, has limited effect on detecting the significant change of the PD_t signal at the beginning of (from “relaxation” to “stress”) and after (from “stress” to “relaxation”) the Incongruent segment (see Figure 5-11). Therefore, the modified backward differential method is developed and implemented, which is given by

\[
\frac{\partial f(x)}{\partial x}_{\text{Modified}} \approx \frac{(\rho_i - \rho_{i-1}) + (\rho_{i-1} - \rho_{i-2}) + (\rho_{i-2} - \rho_{i-3}) + (\rho_{i-3} - \rho_{i-4})}{4\Delta x}
\]

\[
\approx \frac{\rho_i - \rho_{i-4}}{4\Delta x}
\]

\[(4.55)\]
where $\rho_i$ is the current PD$_r$ signal value; $\rho_{i-1}$, $\rho_{i-2}$, $\rho_{i-3}$ and $\rho_{i-4}$ are the four neighboring previous PD$_r$ signal values. In a strict sense, equation (4.55) can not accurately present the mathematically defined derivative, but instead it estimates 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:

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

(4.56)

4.2.2.2 Shape Information Detection for PD$_r$ Signal

The shape information is a common used feature in the digital signal and digital image processing fields, whose most famous approach is the mathematical morphology [89-90]. Fei Zhang and Yong Lian used multiscale mathematical morphology for QRS detection in wearable ECG devices for a body area network [91]. Tomonari Yamaguchi et al; 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 [92]. Chee-Hung Henry Chu and Edward J. Delp provided a new approach to impulsive noise suppression of ECG signals by using mathematical morphological operators [93].

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 the logical transformation, which depends on the comparison of the point (pixel) neighborhoods in the input signal (image) with a pattern. By choosing the size and shape of the morphological operator, the researcher
can construct different patterns. 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 input value’s (pixel’s) neighborhood after transformation; whereas dilation is the operation that outputs the minimum value of all the points (pixels) in the input point’s (pixel’s) neighborhood after transformation.

The purpose of this study is to develop and implement an algorithm for on-line 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_r signal, is presented. The elementary single-scale mathematical morphology operations, which include modified dilation and modified erosion operators, for length N signal $\rho(n)$ are listed below:

Modified Dilation:

$$\rho \oplus o(n) = \max_i (\rho(n+1-i)+o(i)) \quad n = L, \ldots, N \ i = 1, \ldots, L \quad (4.57)$$

Modified Erosion:

$$\rho \ominus o(n) = \min_i (\rho(n+1-i)+o(i)) \quad n = L, \ldots, N \ i = 1, \ldots, L \quad (4.58)$$

where $i$ denotes the $i^{th}$ element in a length L structure element and $o(i)$ is a predefined morphological operator. The size and the value of the morphological operator can be adapted according to the specific application. In some common cases, a constant value is chosen for the structure element. For this study, in order to timely detect the shape variation of the PD_r signal of the subject without much delay, the length of the morphological operator of 5 and a constant value of 0 are implemented.
4.2.2.3 Algorithm for On-line Affective Assessment of the Computer User

In the 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$_t$ point, the current and the previous eight PD$_t$ points need to be considered.

**Step1: Preparation**

The pupil diameter can constrict to 1.5mm or dilate to about 8 to 9 mm. In addition, different individuals have 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 Figure 3-2), the mean value of twenty PD$_t$ points is calculated as the PD$_{Reference Value}$ for the definition of appropriate thresholds. In addition, two thresholds are needed for the affective state assessment, which are respectively regarded as the upper limit of the PD$_t$ signal amplitude fluctuation during the “relaxation” state and the lower limit of the PD$_t$ signal amplitude fluctuation during the “stress” state. These are denoted as “Threshold$_{relaxation}$” and “Threshold$_{stress}$” respectively. These two thresholds are calculated based on PD$_{Reference Value}$, 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:

\[
\text{Threshold}_{\text{relaxation}} = \text{PD}_{\text{Reference Value}} \times c1 \quad (c1>1) \quad (4.59)
\]

\[
\text{Threshold}_{\text{stress}} = \text{PD}_{\text{Reference Value}} \times c2 \quad (c2>1) \quad (4.60)
\]
Step2: Decision based-feature voting

Three features are extracted to identify the PD$_r$ signal change (from “relaxation” to “stress” or from “stress” to “relaxation”). Fulfillment of the criterion by each of the features is given a different level of importance in deciding the final output of the on-line system. If the current PD$_r$ signal value satisfies the criterion of one feature, then the corresponding weight score will be added.

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

\[
\text{If } \text{PD}_r(n) \geq \text{Threshold}_{\text{stress}} \quad (4.61) \\
\text{Weight}_{\text{stress}}(n) = \text{Weight}_{\text{stress}}(n) + 3 
\]

\[
\text{If } \text{PD}_r \leq \text{Threshold}_{\text{relaxation}} \quad (4.63) \\
\text{Weight}_{\text{relaxation}}(n) = \text{Weight}_{\text{relaxation}}(n) + 3 
\]

The second criterion for the second feature detection is based on the modified backward differential method, explained earlier. The derived values, which are obtained from the convolution of $K_{\text{Modified Backward}}$ with a period of the previous neighboring PD$_r$ signals, are denoted as DP$_{MB}$ (Differentiated PD$_r$ value after processing with modified backward differential method). In order to measure the amplitude of the DP$_{MB}$, a threshold has to be implemented, which is denoted as Threshold$_{MB}$. In addition, it is also necessary to count the number of DP$_{MB}$ greater than Threshold$_{MB}$ within a short time interval so as to detect whether DP$_{MB}$ values are consecutively large. The detailed approach for this criterion testing is as follows:
Calculate the $DP_{MB}$s  In order to determine affective state status for one current PD$_r$ value, the five $DP_{MB}$s for this current PD$_r$ point need to be calculated, which are $DP_{MB}(1)$, $DP_{MB}(2)$, $DP_{MB}(3)$, $DP_{MB}(4)$ and $DP_{MB}(5)$. These 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 interval to some extent.

\[ DP_{MB}(i) = 1 \times PD_r(n + 1 - i) - 1 \times PD_r(n - i - 3) \quad (i = 1, 2, 3, 4, 5) \]  

(4.65)

Calculate the number of $DP_{MB}$ greater than the absolute value of $Threshold_{MB}$

\[ Num1 = 0 \]  

(4.66)

\[ \text{If} \quad DP_{MB}(i) > Threshold_{MB} \quad \text{Num1} = \text{Num1} + 1 \quad (i = 1, 2, 3, 4, 5) \]  

(4.67)

\[ Num2 = 0 \]  

(4.68)

\[ \text{If} \quad DP_{MB}(i) < -Threshold_{MB} \quad \text{Num2} = \text{Num2} + 1 \quad (i = 1, 2, 3, 4, 5) \]  

(4.69)

Modify the weight for “relaxation” or “stress” affective assessment

\[ \text{If} \quad \text{Num1} \geq 3 \quad \text{Weight}_{stress}(n) = \text{Weight}_{stress}(n) + 1 \]  

(4.70)

\[ \text{If} \quad \text{Num2} \geq 3 \quad \text{Weight}_{relaxation}(n) = \text{Weight}_{relaxation}(n) + 1 \]  

(4.71)

The third criterion is based on the shape information detection method which has also been discussed earlier. 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 short period of time to some extent. If the current PD$_r$
value is the maximum or the minimum of the latest five PD_r values, we call the point “morphological matched point”. The details for this criterion testing are proposed below:

(1) Calculate the number of “morphological matched points”

\[
\text{Num}_{\text{dilate}} = 0
\]  
(4.72)

If \( \text{PD}_r(n-i) = \max[\text{PD}_r(n-i), \text{PD}_r(n-i-1), \text{PD}_r(n-i-2), \text{PD}_r(n-i-3), \text{PD}_r(n-i-4)] \) 

\[
\text{Num}_{\text{dilate}} = \text{Num}_{\text{dilate}} + 1
\]  
(4.74)

\[
\text{Num}_{\text{erosion}} = 0
\]  
(4.75)

If \( \text{PD}_r(n-i) = \min[\text{PD}_r(n-i), \text{PD}_r(n-i-1), \text{PD}_r(n-i-2), \text{PD}_r(n-i-3), \text{PD}_r(n-i-4)] \) 

\[
\text{Num}_{\text{erosion}} = \text{Num}_{\text{erosion}} + 1
\]  
(4.77)

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

If \( \text{Num}_{\text{dilate}} \geq 3 \) \quad \text{Weight}_{\text{stress}}(n) = \text{Weight}_{\text{stress}}(n) + 1
\]  
(4.78)

If \( \text{Num}_{\text{erosion}} \geq 3 \) \quad \text{Weight}_{\text{relaxation}}(n) = \text{Weight}_{\text{relaxation}}(n) + 1
\]  
(4.79)

Step3: Affective state determination

The purpose of this step is to determine the final classification of the affective state of the subject. Analysis of the current PD_r value may yield three outcomes: 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

The current PD_r (n) is the key (upward) transition point, which indicates that the affective state of the human subject is changing from “relaxation” to “stress”.

66
If $\text{Weight}_{\text{relaxation}}(n) \geq 4$ and the previous PD, value in “stress” affective state

The current PD$_r(n)$ is the key (downward) transition point, which indicates that the affective state of the human subject is changing from “stress” to “relaxation”.

Otherwise

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

The figure 4-8 shows the algorithm of on-line affective assessment based on PD signal.

Figure 4-8: The algorithm of on-line affective assessment based on PD signal

4.3 Summary

This chapter proposes new affective assessment approaches (include both on-line and off-line algorithms) to classify the “stress” vs. “relaxation” states of computer users through the monitoring of the Pupil Diameter (PD) signal. In the off-line algorithm, the wavelet denoising and Kalman filtering methods were used to remove abruptly changing components of the raw PD signal first. Then three features, PDmean, PDmax and PDWalsh, were extracted from the filtered PD signal, to emphasize the most meaningful characteristics of PD, for classification purposes. Finally, five different classifiers were implemented in order to evaluate the performance and test the robustness of the PD signal for affective assessment. Similarly, ROC curves were developed to assess the
discriminating power of each individual feature. In the on-line algorithm, a hard threshold was set for eye blink removal and a moving average window was implemented for obtaining PD. Then three steps, which were preparation, feature-based decision voting and affective determination, were proposed to finally obtain the classification outcome of the on-line algorithm.
CHAPTER 5
RESULTS & DISCUSSION

In this study, the PD and the GSR signals measured during each Congruent (C) Stroop segment and each Incongruent (IC) Stroop segment were analyzed. The off-line and the on-line algorithms were both applied to the PD signal in order to achieve the affect assessment for each subject. For the off-line algorithm, the signal preprocessing, the feature extraction and the classification three steps were applied to the both PD and GSR signal (the GSR was used for comparison purposes). For the on-line algorithm, the preparation, feature-based decision voting and affective determination three steps were applied to the PD signal only. In this study, 42 subjects volunteered to participate in the experiment, with ages ranging from 20 to 50 (Mean: 27.5; Std. dev.: 5.14), from diverse professional and ethnic backgrounds. The results of each processing step are described in this chapter.

5.1 Off-line Affective Assessment Results

Generally, the off-line processing procedure for the PD signal-based affective assessment is composed of the following three steps. They are:

(i) Signal preprocessing: Remove the eye blinks and artifact influence on the recorded raw PD signal; Use the wavelet denoising and Kalman filtering method to remove the abrupt change of the PD signal;

(ii) Feature extraction: Data normalization is implemented first; then three features are extracted: PDmean, PDmax, PDWalsh;
(iii) Classification: Five different classifiers are utilized in order to evaluate the accuracy and robustness of PD signal for affective assessment.

5.1.1 Signal Preprocessing

In our research, the TOBII T60 eye tracker is utilized to monitor the pupil diameter at a frequency of 60Hz. An example of data from one of the 42 subjects who participated in the study is shown as follows:

Figure 5-1: A set of raw signals in the experiment. From top to bottom: the raw PD, the raw GSR and IL.
Figure 5-1 shows a set of signals (which 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. The three most important boundaries separate each congruent Stroop segment (C) from the incongruent Stroop segment (IC) that follows.

Because of the presence of eye blinks (signaled by a value of “4” in the validity code given by the TOBII system and identified as sudden transitions to a false PD value of zero), linear interpolation was implemented first to compensate the interruption in the PD data by each blink. This yields the PD signal after the removal of the eye blinks (Figure 5-2, the upper plot).

Although linear interpolation removes the blinking artifacts, it should be noted that there is still a substantial amount of fast variability that is not likely to represent pupil size changes due to affective variations. In fact, previous research has indicated that these fast variations in the PD signal could be due to quantization noise in the pupil diameter measurement.

On account of the noise included in the original PD signal, it is clear that it would be difficult to successfully classify the segments by using a threshold on the original PD data. Therefore, the goal in this study was to process the PD signal to retain primarily components that are derived from affective changes in the user, removing the noise components.
In this study, the main signal processing methods applied to the PD are wavelet denoising and Kalman filtering. The result of the PD signal after the wavelet denoising is illustrated in the middle plot of the Figure 5-2. Although the wavelet denoising method removes most of the abrupt changes of the original signal, there are still some abrupt transitions left. Therefore, Kalman filter is used for further processing. The result of the PD signal after Kalman filtering is illustrated in the lower plot of the Figure 5-2.

Figure 5-2: Original PD signal (upper), signal after wavelet denoising (middle) and signal after Kalman filtering (lower)
5.1.2 Normalization and Feature Extraction

In this step, the data is first normalized to the range of [-1, 1] and then three features are extracted from each of the PD and GSR signals, in order to perform further analysis.

Figure 5-3 shows the Walsh transform coefficients (bottom) obtained from sets of 8 mean values calculated from the beginning of the 6 segments in the PD data (top) from one experimental subject.

![The Mean Onset Value of the PD Signal for Six Stroop Segments of One Subject](image)

![Walsh Coefficients for Six Stroop Segments of One Subject](image)

Figure 5-3: The Walsh transform of the average onset value of the PD signal for six Stroop segments of one subject

The values of the “PDWalsh” feature for the three Congruent Stroop segments are 0.2730, 0.4233 and 0.2956, respectively, whereas for the three Incongruent Stroop segments are -0.5070, -0.3951 and -0.8192. These sample values exemplify the ability of
the “PDWalsh” feature to represent the presence or absence of the fast initial increase observed as a characteristic of the IC segments, in which the stimuli were expected to elicit stress in the subject. In addition to the rapid increase of PD observed in the IC segments, the maximum and average values of the preprocessed PD signal were also observed to be different for IC and C segments. As such, the maximum and average PD values in each segment under analysis were also used as features for the purpose of classification of segments.

As an example of the feature values obtained, the 6 features extracted from the PD and GSR data shown in Figure 5-2 are illustrated in the table below.

<table>
<thead>
<tr>
<th>Signals</th>
<th>Features</th>
<th>Stroop Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>PDmean</td>
<td>C1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.5330</td>
</tr>
<tr>
<td></td>
<td>PDmax</td>
<td>-0.2675</td>
</tr>
<tr>
<td></td>
<td>PDWalsh</td>
<td>0.2730</td>
</tr>
<tr>
<td>GSR</td>
<td>GSRmean</td>
<td>0.3120</td>
</tr>
<tr>
<td></td>
<td>GSRrising Time</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>GSRnum</td>
<td>0</td>
</tr>
</tbody>
</table>

5.1.3 Affective Assessment

While the SCWT was implemented in our experiment for the purpose of eliciting mild mental stress of the human subjects during controlled intervals, it is important to
verify for each subject and each pair of C and IC segments (denoted C/IC), that the stimuli were effective in producing a significant stress-related change, as hypothesized. Towards 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 second approach relied on the feedback provided by the subjects themselves, who were asked to describe their reactions in a questionnaire.

5.1.3.1 Selection of relevant C/IC segment pairs based on paired \( t \)-test

The GSR signal is one of the most commonly used physiological signals which can gauge some degree of stress affect of the human subject. 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 age: 27.52; std: 5.14) 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 the GSR signal with the affective state change, filtering and normalization preprocessing was applied 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 [94-95], with the significance values of 0.200 and 0.446 respectively. Therefore, it is possible to confirm that the values of “DiffGSRmean” were normally distributed. Then the paired \( t \)-test
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 verifies that the SCWT is, generally, an effective approach to elicit mild stress during Incongruent segments, allowing the subject to remain relaxed during Congruent segments. In addition, the 90% confidence interval of 126 “DiffGSRmean” values is [0.0453, 0.1414]. In order to remove the C/IC segments pairs in which stress elicitation was probably 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 segment (i.e., 140 segments) is as the data to be used in the rest of the off-line and on-line algorithm studies.

For the off-line system, five different classification algorithms were implemented through the features derived from the physiological signals recorded in each segment (see Table 4-1) in order to assess the affective state (relaxation vs. stress) of the human subjects. Specifically, in this 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 [68-69]. In this study, the five different classification algorithms are all from the WEKA for the purpose of observing the classification effectiveness and robustness of the features extracted (see Table 4-1). In addition, three classification phases were also performed, as illustrated in Table 4-2, in order to compare the classification efficiency of PD and GSR signal for affective assessment.
The data (pairs of C/IC segments) selected on the basis of the paired $t$-test, consisted of 70 congruent segments and 70 incongruent segments. These were considered as viable and 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 [96] was used. The accuracy rates from the experiments are shown in Table 5-2.

**Table 5-2**

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>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>
<td></td>
</tr>
<tr>
<td>P2:3 features from GSR (No PD)</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>
<td></td>
</tr>
<tr>
<td>P3:3 features from PD (No GSR)</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>
<td></td>
</tr>
</tbody>
</table>

5.1.3.2 Selection of relevant C/IC segment pairs based on 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 rating scales of affective state, to obtain a 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 (e.g., [97]), the collection of self-reports from subjects has been widely used in psychological and affective computing research filed [98-100].

In our questionnaire, two questions were posed: “How did you feel in Congruent segments?” and “How did you feel in Incongruent segments?”. An affective rating scale with 5 levels (from “1” to “5”) was listed under each question. “1” represents that the subject feels relaxed just like in a normal environment. “5” represents that the subject feels stressful just like he/she can not find the key of his/her house. “2”, “3” and “4” represent increasing, intermediate levels between “1” and “5”. The final results show that compared with the feeling in the Congruent segment, all the subjects felt more stressed 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 remove from consideration C/IC pairs of experiments with potentially insufficient stress elicitation, only data from subjects with difference evaluation score between Incongruent and Congruent segment greater than or equal to two was retained. In this way, only data from 31 subjects (i.e., 93 Congruent and 93 Incongruent segments) were kept for off-line and on-line affective processing, under this selection approach.

In addition, in order to obtain a more accurate and realistic assessment of the classifiers, the WEKA software and a 6-fold cross validation method were used. The accuracy rates from the experiments are shown in Table 5-3.
5.1.4 The Results of Receiver Operating Characteristic (ROC) Curve

The area under the ROC, denoted “AUROC”, is of interest because it is theoretically possible to operate at any point on the convex hull of the points in a ROC curve [101]. For the construction of the ROC curves for the different normalized features, the values of the feature for all the C and IC segments (for all subjects) were considered together.

The AUROC values of all the six features are presented in Table 5-4 and the ROC / ROCCH plots are shown for PDmean (Figure 5-4), PDmax (Figure 5-5), PDWalsh (Figure 5-6), GSRmean (Figure 5-7), GSRrisingTime (Figure 5-8) and GSRmean (Figure 5-9). It can be appreciated, in Table 5-4 and Figure 5-6, that the feature PDWalsh has the

<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>

Table 5-3
Results of affective assessment by five types of classification algorithms
(Selected data based on the questionnaire)
largest area under the standardized ROC curve. It is also noticeable that all three PD-derived features have significantly larger AUROC values than the features derived from the GSR signal.

Table 5-4
AUROC results for all 6 features derived from PD and GSR

<table>
<thead>
<tr>
<th>Physiological Signal</th>
<th>Features</th>
<th>AUROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD (3 features)</td>
<td>PDmean</td>
<td>0.9320</td>
</tr>
<tr>
<td></td>
<td>PDmax</td>
<td>0.9106</td>
</tr>
<tr>
<td></td>
<td>PDWalsh</td>
<td>0.9445</td>
</tr>
<tr>
<td>GSR (3 features)</td>
<td>GSRmean</td>
<td>0.5593</td>
</tr>
<tr>
<td></td>
<td>GSRrisingTime</td>
<td>0.5415</td>
</tr>
<tr>
<td></td>
<td>GSRnum</td>
<td>0.5579</td>
</tr>
</tbody>
</table>

Figure 5-4: The ROC/ROCCH of PDmean
Figure 5-5: The ROC/ROCCH of PDmax

Figure 5-6: The ROC/ROCCH of PDWalsh
Figure 5-7: The ROC/ROCCH of GSRmean

Figure 5-8: The ROC/ROCCH of GSRrisingTime
5.2 On-line Affective Assessment Results

5.2.1 The Results of the Physiological Signal Preprocessing

Figure 5-10 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 Figure 5-10 that while the length of the original PD signal in a typical record is about 28800 samples, the preprocessed signal is just about 480 data points (PD_r), due to the 60 samples contained in the sliding window used for averaging. In addition, it is apparent that the newly derived 480 data points (PD_r) still represent the variations of the original PD signal while simultaneously removing the noise to some extent.
5.2.2 The Results of On-line Affective Assessment

5.2.2.1 The Result of Modified Backward Differential Method for Sudden PD$_r$ Signal Change Detection

Figure 5-11 illustrates the results after the convolution of $K_{\text{Backward}}([-1 1])$ with a pair of segments (one Congruent segment and one Incongruent segment) of one subject’s PD$_r$ signal. The experimental protocol was designed to arouse “stress” in the human subject during the Incongruent segment (IC) while keeping the “relaxation” during the Congruent segment (C). It is conceivable that during the subject’s affective state variation period, there should be significant PD$_r$ signal value changes (increase or decrease). In Figure 5-11, the second vertical line denotes the expected affective state transition of the human subject was from “relaxation” to “stress”; whereas the third vertical line indicates the expected affective state transition of the human subject from “stress” to “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 Figure 5-11, 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 other time periods. In other words, the traditional first order backward differentiation operator can not effectively identify the sudden change of the PDₜ signal.

![Figure 5-11: The results after the convolution of K_{Backward} with a pair of segments of the PDₜ signal](image)

Figure 5-11: The results after the convolution of K_{Backward} with a pair of segments of the PDₜ signal

Figure 5-12 shows the results after the convolution of K_{Modified Backward} ([−1 0 0 0 1]) with the same pair of segments (one Congruent segment and the following Incongruent segment) of the PDₜ signal as in Figure 5-11. It is evident from the results that at the beginning of the Incongruent segment (a short time period after the second vertical line in Figure 5-12), there are four consecutive points with large positive values (slightly larger than 0.3); after the Incongruent segment (a short time period after the third vertical line in...
Figure 5-12), there are six consecutive points with large negative values (slightly less than -0.3). However, in other time durations, 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 convoluted by $K_{\text{Modified Backward}}$, it could be inferred that the PD$_r$ signal has the significant increase or decrease during those time intervals, 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.

Figure 5-12: The results after the convolution of $K_{\text{Modified Backward}}$ with a pair of segments of the PD$_r$ signal
5.2.2.2 The Result of Shape Information Detection for PD$_r$ Signal Change

Figure 5-13 shows the results of a pair of segments (one Congruent segment and one Incongruent segment) of the PD$_r$ signal after processing with the modified dilation and erosion operations. Each stem in the figure is a PD$_r$ value, which represents the average value of the valid PD samples within one second. In order to better illustrate the effect of the morphological transformation, two additional line traces are used in the plot to show the results after the modified dilation and erosion operations. As shown in Figure 5-13, the solid line, the upper edge of the cluster of PD$_r$ stems, 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 results after modified erosion operation processing. These two lines can be used to describe the overall tendency of the PD$_r$ signal, especially during abrupt amplitude changes. It is apparent from Figure 5-13 that when the PD$_r$ signal has significant increases, the stems are almost consecutively on the solid line; whereas when the PD$_r$ signal has significant decreases, the stems are almost consecutively on the dashed 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.
5.2.2.3 The Result of Affective Assessment Based on the On-line Processing Algorithm

Figure 5-14 shows the affective assessment results of one subject from on-line PD signal processing. There are two traces in this figure. The upper trace is the preprocessed PD$_r$ signal. The lower trace is the on-line affective assessment result for the whole experiment. The negative outputs (stems) with value -1 denote that at that time, the system indicated that the affective state of the subject was relaxed; whereas the positive stems with value 1 indicate that the computer user is regarded as stressed at that time. It should be noted that only the Congruent and the Incongruent segments were processed for classification by the system, therefore, the system’s output at other time periods is shown as “0”. The accuracy rate for the whole example shown in Figure 5.14 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 the on-line processing
algorithm 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\%.

![Figure 5-14: The on-line affective assessment result for one subject](image)

5.3 Discussion of Off-line Affective Assessment Algorithm

5.3.1 Discussion of the Signal Preprocessing

As preprocessing, 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. The noise observed in the original PD signal contains components that change abruptly and quickly (similar to the “salt-and-pepper” noise in images). Accordingly, the use of a median filter, which runs through the signal sample by sample and replaces each sample with the
median of neighboring samples [102], could have been a possible solution. However, preliminary attempts using median filters, Finite Impulse Response (FIR) and Infinite Impulse Response (IIR) filters could not simultaneously remove the noise and provide a clear differentiation of the PD increases expected during the “IC” segments.

Wavelet denoising with the Daubechies wavelet and nine decomposition levels was applied to the raw PD signal and was able to remove the artifactual sudden changes from it. This was achieved by separating and discarding the small spikes or sudden changes which occur in a short time interval in the raw PD signal. Although the wavelet denoising method achieves a clear improvement in the signal, its output still includes some abrupt changes, which indicated the need to apply further processing. Therefore, Kalman filtering was implemented as the second signal preprocessing level to remove the remaining noise, obtaining more indicative output signals, such as the one shown in the lower plot of Figure 5-2. This figure exemplifies the ability of this method to significantly remove the remaining noise of the PD signal and make the increases in the incongruent segments of the processed signal more clearly distinguishable, providing a clearer indication of when the subject is stressed. While the two forms of filtering applied were not used to account for PD variations due to illumination changes, the encouraging levels of classification achieved even in the presence of illumination changes in segments IC2 and C3 illustrate the resilience of the approach used.

5.3.2 Discussion of Data Normalization and Feature Extraction

The feature extraction section of the study, sought to derive a few numerical values from each segment of preprocessed PD signal, which could efficiently capture the
differences expected between the pupil diameter in congruent and incongruent segments. Two logical features were the average and the maximum value of the processed PD in each segment. In addition, there was interest in deriving a feature that would highlight the sudden and sustained increase displayed by the pupil diameter when the subject becomes stressed. Accordingly, the difference value (“PDWalsh”) between the first and the second Walsh coefficients based on the PD signal during the onset of each Stroop segment was used as one of the features to differentiate the affective state of the computer users. An additional advantage of this selection of features is that the Walsh transform computation is simple, involving only addition and subtraction operations, resulting in a fast execution time. It should also be noted that the use of the difference between the first two Walsh coefficients is well matched to the sudden emergence of stress expected in our protocol, but may not be as useful in circumstances where stress builds up in a slow progression.

5.3.3 Discussion of Affective Assessment

In the study, five different classification algorithms were applied in order to investigate the effectiveness and robustness of the features derived from the physiological signals for differentiating between “relaxation” and “stress” user states, as evoked by congruent and incongruent Stroop stimulation, respectively. For the 70 valid C/IC pairs of data selected on the basis of the paired \( t \)-test, the outcomes of the classification, (Table 5-2), show that the monitoring and processing the PD and GSR signals (Phase 1) yield the average classification accuracy of 85.43%.

In the following two phases of stress classification the importance of each physiological signal for affective assessment is investigated. In phase 2 the PD signal is
excluded and only the GSR signal is used for classification analysis. The average classification rate drops dramatically, to 62.29%, which is obviously lower than the accuracy rate in the first phase (including the features derived from the PD and GSR signals). The results seem to hint that the discrimination contribution of the GSR features is moderate and the PD features might have been contributing strongly to the higher classification accuracy levels found when they were involved. This, in turn, may also suggest that the pupil diameter signal could be a very significant physiological signal to consider towards the development of an automated affective classification system. In order to verify our conjecture, in Phase 3, we use the three features extracted from the PD signal for classification, exclusively. We found that the average classification achieved with those features is 86.43% and the highest accuracy reached is 88.57%, achieved using the Naïve Bayes classifier approach. These observations suggest that the PD is, indeed, an important physiological signal for affective assessment, especially for “relaxation” vs. “stress” detection, as illustrated by our experiment. In addition, in comparison to the variance of the five classification rates in Phase 1 (where features from PD and GSR signals were used), which is 9.85%, and the variance in Phase 2 (where only GSR features were used), which is 3.16%, the variance in Phase 3 (involving only the PD features) is just 3.06% (This very small variance seems to indicate that the PD features contain strong discriminating information that can be exploited well by any of the classifiers).

The results of Table 5-3, which refers to the data chosen through the questionnaire (31 subjects’ data) are very similar to the results of Table 5-2 (for data selected through the t-test). In phase 1, an average classification accuracy of 85.27% was achieved. In phase 2, only the GSR signal is used for classification analysis (the PD signal is
excluded). The average classification rate drops significantly, to 51.94%, which is obviously lower than the accuracy rate in the first phase. In phase 3, the PD signal is only utilized and we get the average accuracy of 87.20%, which confirms that the PD signal is more effective than the GSR signal to make the affective assessment (“relaxation” vs. “stress”). In addition, the variances of results from the five classification algorithms are also investigated. The smallest one is 1.65%, which is in phase 3 (only the PD signal is considered), whereas the largest one is 2.25%, which is in phase 2 (only the GSR signal is considered). All these results indicate that the PD signal is not only effective but also robust for affective assessment.

5.3.4 Discussion of Receiver Operating Characteristic (ROC) Curve

We have also investigated the potential of the six types of features derived from the PD and GSR signals: PDmean, PDmax, PDWalsh, GSRmean, GSRrisingTime and GSRnum, to act as individual classification signals to differentiate between stress and relaxation states of the computer user. The results show that GSRmean, which is one of the physiological variables most commonly monitored for affective sensing, exhibited only moderate discriminating power with an AUROC value of 0.5593. Similarly, the GSRrisingTime and GSRnum features exhibited limited potential for this detection problem, with AUROC values of 0.5415 and 0.5579. In contrast, the three features derived from the PD signal (PDmean, PDmax and PDWalsh) displayed a strong potential for affective assessment, with AUROC values of 0.9320, 0.9106 and 0.9445, respectively. Representative ROC/ROCCH curves are included in Chapter 4 of this dissertation. It should be noted that the PDWalsh ROC/ROCCH curve exhibits a sharp slope from the coordinate origin and reaches into the high levels of the true positive rate immediately,
which represents a significant departure from the random classification diagonal. From this curve, it is easy to obtain a number of intermediate points for which the true positive rate is better than 0.8 while simultaneously having a false positive rate of less than 0.2.

5.4 Discussion of On-line Affective Assessment Algorithm

5.4.1 Discussion of Modified Backward Differential Method for Sudden \( \text{PD}_t \) Signal Change Detection

Traditional differential operators, which include forward, backward and central differential operators, have long been used in the digital signal and image processing fields with great success to detect significant changes. Because the purpose of this research is to develop an on-line \( \text{PD}_t \) signal based affective assessment (“relaxation” vs. “stress”) system for human subjects, it is evident that only the present and the previous \( \text{PD}_t \) data can be used. Therefore, the traditional backward differentiation operator was implemented first, with the results shown in Figure 5-11. It is widely known that according to the sign and the amplitude of the value derived from differentiation operators, the significant change of the \( \text{PD}_t \) signal can be automatically identified. As we can see from Figure 5-11, there are almost no values with large amplitude after the two boundaries of the Incongruent segment (IC), which could be interpreted as an indication that there is no significant change during the affective transition. However, in fact, it is evident from Figure 5-12 that the \( \text{PD}_t \) signal has important increase or decrease variation at the two boundaries of the Incongruent segment (IC). The reason why the traditional differentiation operator is not very effective in detecting them may be the small amplitude difference between two neighboring \( \text{PD}_t \) data points. Therefore, the modified differentiation operator \([-1 0 0 0 1]\) is proposed. This operator detects the difference
between the current valve and the valve 4 time steps back. The proper interval between the two detected points can not only represent a large difference between two points but also present the overall tendency of the PDₚ 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 PDₚ signal has undergone a significant change in this interval of time. In the algorithm, for one current PDₚ value classification, a set of five differences are to be calculated, which are displayed in equation 4.65. In other words, the current and the previous eight PDₚ data points are considered for voting, which can reliably outline the variation of the PDₚ signal, in that time interval.

5.4.2 Discussion of Shape Information Detection for PDₚ Signal Change

This extracted feature is to detect the shape information of the PDₚ 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 [103]. 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 PDₚ 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 achieve on-line detection of the sudden change (“increase” or “decrease”) of the PDₚ 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 PDₚ variations during the corresponding interval.

5.4.3 Discussion of Affective Determination

Two kinds of commonly used methods to select the more likely valid data for the affective assessment were used. One is the selection of relevant C/IC segment pairs based on paired t-test, which is a sound statistical approach. 70 out of 140 segment pairs were kept to perform further analysis, which resulted in the average classification rate of 72.30%. The other method is the selection of relevant individual data based on answers to a questionnaire, which is the commonly used approach in psychology. Data from 31 out of 42 subjects was kept and the average classification accuracy was 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 on-line 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 on-line PD signal based affective assessment are 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. The GSR signal is a type of widely used physiological signal to detect the emotions of human subjects. According to the physiological literature and the previous research on the GSR signal, “GSRmean”, “GSRnum” and “GSRrisingTime” features were extracted from the
all the Congruent ("C") and Incongruent ("IC") segments for off-line differentiation between “relaxation” and “stress” states of the computer user. However, GSR classification of both sets of selected data yields accuracy levels that are significantly lower than the accuracy of the on-line PD affective assessment algorithm (62.29% vs. 72.30%; 51.94% vs. 73.55%). These results and observation reinforce the belief 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 in classifying “relaxation” vs. “stress”.

5.5 Summary

This chapter first presented the results of the off-line algorithm for affective assessment. Examples of the PD signal after wavelet denoising and Kalman filtering were shown, illustrating the effectiveness of these methods to remove the abrupt variations of the raw PD signal. Then three features (PDmean, PDmax and PDWalsh) were extracted and processed by five different types of classification algorithm. The final classification results show that using the features derived from the PD signal exclusively can yield an average segment classification accuracy of 86.43% and 87.20%, respectively, for the 2 subsets of valid data kept for analysis. However, using the features from the GSR signal (excluding the features from the PD signal) could only yield average accuracies of 63.57% and 58.60%. Then this chapter presented the results of the on-line algorithm for affective assessment. There are mainly three steps for on-line affective assessment algorithm, which are preparation, decision based-feature voting and affective state determination. The final average accuracy results are 72.30% and 73.55% for the 2 subsets of valid data selected. These on-line results derived from the PD signal were
better (higher accuracy) than the accuracy for off-line GSR affective assessment. There are mainly three steps for on-line affective assessment algorithm, which are preparation, decision based-feature voting and affective state determination.
CHAPTER 6
CONCLUSIONS & FUTURE WORK

6.1 Conclusions

This dissertation proposed new approaches to recognize the “stress” affective state of the computer users through off-line and on-line processing of the pupil diameter (PD) signal, which is a non-invasive measurement that can be obtained by currently available eye gaze tracking systems. The proposed methodology includes both the off-line and the on-line methods. This dissertation also compared aim is to compare the classification efficiency of PD and GSR signals for stress detection.

In the off-line affective assessment algorithm, in order to emphasize the variations of the pupil diameter signal associated with affective changes of the computer user, wavelet denoising and Kalman filtering were applied on the original PD signal (after interpolation of eye blinks), sequentially, to remove the impulsive noise originally present in the signal. The signal obtained from this preprocessing displayed consistent increases that occur when the subject was stressed (i.e., during the incongruent Stroop segments). Therefore, the application of wavelet denoising and Kalman filtering seems to constitute an appropriate approach to obtain a processed PD signal that properly reflects the stressed states in the computer user.

In this study, a total of six features were extracted from the two monitored physiological signals (PD and GSR) in each congruent and incongruent Stroop segment of the protocol. Then, the features were processed by five different classification algorithms, in three different phases, in order to evaluate the overall usefulness of the
different physiological signals for affective assessment. To analyze only data for which stress elicitation appeared to be successful, two types of data selection methods were implemented. One was based on the paired $t$-test (statistics) and the other was based on self-evaluation (questionnaire) by the subjects. The final results show that using the features only derived from the PD signal can yield an average segment classification accuracy of 86.43% and 87.20% respectively for five different classifiers. However, using the features from the GSR signal (excluding the features from the PD signal) could only yield a maximum segment classification accuracy of 63.57% and 58.60%, respectively. 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 (one congruent and one incongruent Stroop segments), to emulate the illumination variations that can occur during ordinary computer use.

To further analyze the results from our experiment, the ROC/ROCCH plot for each of the six features derived from the PD and GSR signals was obtained. These graphs confirmed that the 3 individual features derived from the PD signal: PDmean, PDmax and PDWalsh, had normalized areas under the ROCCH curve of more than 0.9, whereas the 3 features derived from the GSR signal had areas under the ROCCH that did not surpass 0.6. These results show that the features derived from the PD signal exhibit a strong potential for discrimination between relaxed and stressed user states, as defined in the experiment performed, on the basis of congruent and incongruent Stroop stimulation.

In the on-line affective assessment algorithm, a hard threshold is first used to remove the eye blinks of the raw PD signal. Then a moving average window method
(length window=60 samples) is chosen to extract the representative data denoted as PD_r to reflect the overall subject state during each second. There are mainly three steps for on-line affective assessment algorithm, which are preparation, decision based-feature voting and affective state determination. In the preparation part, the reference value of the PD_r is the average value during one time interval in the introductory section 1 of the experiment. Then, two thresholds are inferred based on the previous calculated PD Reference Value, which are respectively regarded as the upper limit of the PD_r signal amplitude fluctuation during the “relaxation” state and the lower limit of the PD_r signal amplitude fluctuation during the “stress” state. In the decision based-feature voting part, three features are extracted to identify the affective state of the computer user. Different features were assigned weight scores. If the current PD_r signal value satisfies the criterion for one feature, the corresponding weight score will be modified accordingly. These three criteria are respectively based on the current PD_r value, the modified backward differential method and the shape information detection method. The last step is the affective state determination, which generates an affective status indication from the current PD_r (the possible outcomes are: from “relaxation” to “stress”, from “stress” to “relaxation” or no affect change). The final average accuracy levels achieved by the online processing method were 72.30% and 73.55%, for two different types of data selection methods (paired $t$-test and self-evaluation), which are clearly higher than the accuracy of the off-line GSR affective assessment.

In sum, the results of the 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.

6.2 Limitations of Affective Assessment Approaches Based on Pupil Diameter Signal

Although the proposed affective assessment approach based on off-line and on-line processing of the PD signal can obtain relatively promising results for differentiating “relaxation” vs. “stress” affective state, there are still some limitations.

The first limitation is that an accurate and robust eye tracking system is required (in our experiment, the TOBII T60 system is utilized), because almost all the physiological signals (including pupil diameter) are very subtle and easily interrupted by the environment, which may include improper setup or the operation of the system, excessive subject movement and so on. In addition, measuring the physiological signals (including pupil diameter) requires the subject to maintain a fixed position for a period of time, and fulfilling this requirement may increase the discomfort of the subjects.

The second limitation is that our approach to make the affective assessment (“relaxation” vs. “stress”) is not designed for a task in which mental stress would increase progressively. Indeed, the aim of this research was to target the detection of “emergence” of stress, because an opportune detection of the onset of stress would be the appropriate moment to initiate corrective or compensatory actions on the part of the computer. In addition, the design of the experiment used the “Stroop Color-Word Test” to elicit mild stress in subjects at very specific times. The sequence of segments that was
implemented provided 3 opportunities, per subject, to witness the congruent-to-incongruent transition, which is a rather sudden change.

The third limitation is that our algorithm can not recognize many other types of affect of the computer users may experience. Russell and Feldman Barrett [104-105] proposed the theoretical assumption that the emotional state can be defined by two or three dimensions (1. valence, 2. arousal, 3. dominance) induced in their subjects. Herbon et al. [106] proposed that pupil diameter may be a rather strong indicator of the valence of an emotion. Therefore, it is possible that pupil diameter alone may not effectively differentiate between affective states which have the same levels of valence but different arousal types, such as “fear” vs. “sadness”. Some other types of observation such as facial expression and speech may be required for multiple affect recognition.

6.3 Future Work

This dissertation developed the approach for off-line and on-line differentiation of the affective state of the computer user (“relaxation” vs. “stress”) through the non-invasive monitoring of the PD signal. However, in the future work, the proposed approach could be enhanced by following additional research avenues.

The first suggestion is to integrate other affective sensing modalities to recognize some more types of the emotions of the computer user. Apart from the physiological signals, including pupil diameter, there are other subject characteristics which can reflect the emotion of the computer user to some extent, such as body gesture, facial expression and speech. Although these affective sensing modalities have their own advantages and
disadvantages, integration of many other recognition methods may provide a promising opportunity to reach a better performance in multi-affective sensing.

A possible second enhancement is to apply the developed off-line and on-line algorithms to data from whose stress affect is not elicited by the Stroop Color-Word Test, in order to test the general effectiveness and robustness of the approach. For example, in the future work, the researcher can use “joyful” vs. “terrible music” or “comedy” vs. “horror” films to elicit the “relaxation” and the “stress” emotions in the computer user.

The last recommendation is to evaluate the system with larger subject pools. In our experiment, 42 subjects volunteered to participate. In order to make further verify that each pair of C and IC segments (denoted C/IC) really involved a significant stress-related change, in our study, two types of data selection approaches were applied. The first is the selection of relevant C/IC segments pairs based on paired \( t \)-test analysis. The other method is to apply a traditionally used questionnaire to assess emotional state from self-reports by the subjects. After the processing of the relevant data selection, 70 out of 126 pairs of C/IC segments and 31 out of 42 subjects’ data were retained as the final valid data. These amounts could expanded by collecting data from a larger pool of subjects.
REFERENCES


[61] M. St-Pierre and D. Gingras, “Comparison between the unscented Kalman filter and the extended Kalman filter for the position estimation module of an integrated


APPENDICES
MEMORANDUM

To: Dr. Armando Barreto
CC: File
From: Jada Dixon, MJ, MPH, IRB Coordinator
Date: February 9, 2012
Proposal Title: "Online Affective Recognition of a Computer User a Biosignal Processing Approach
IRB Approval # 020311-01

The Health Sciences Institutional Review Board of Florida International University has re-approved your study for the use of human subjects via the Expedited Review process. Your study was found to be in compliance with this institution's Federal Wide Assurance (00000060).

Your IRB approval date is February 2, 2012 and it will expire on February 3, 2013. As a requirement of IRB approval you are required to:

1) Provide immediate written notification to the IRB of:
   • Any additions to, or changes in the procedures involving human subjects.
   • Every serious or unusual or unanticipated adverse event as well as problems with the rights or welfare of the human subjects. Confirmation of receipt of serious AE reports must be made with the IRB office.
2) Utilize copies of the date stamped consent document(s) for the recruitment of subjects and receive annual renewal of consent documents.
3) Receive annual review and re-approval prior to your expiration date. Projects should be submitted for renewal at least 30 days in advance of the expiration date.
4) Close out your project with the IRB office when the study is finished or discontinued.

Special Conditions: N/A

Please note that your IRB approval number is indicated above. For further information, you may visit the IRB website at http://research.fiu.edu/compliance/compliance.html.
ADULT CONSENT TO PARTICIPATE IN A RESEARCH STUDY
On-line Affective Recognition of a Computer User through A Biosignal Processing Approach

PURPOSE OF THE STUDY
You are being asked to be in a research study. The purpose of this study is to design and develop a Digital Signal Processing (DSP) approach for affective sensing through pupil diameter monitoring. We are trying to collect and process data from real-time pupil diameter measurements.

NUMBER OF STUDY PARTICIPANTS
If you decide to be in this study, you will be one of about 30 people in this research study.

DURATION OF THE STUDY
Your participation will require about 40 minutes of your time, in a single visit to the FIU DSP Laboratory.

PROCEDURES
If you agree to be in the study, we will ask you to do the following things:
1. Sit in front of an eye gaze tracking system (TOBII T60), which looks just like a standard LCD computer monitor, with your head facing the computer screen throughout the experiment. You will manipulate a standard computer mouse as you normally do when you use any computer. You will place the middle and ring fingers of your other hand on a small device (Galvanic Skin Response sensor) that also looks like a computer mouse. The Galvanic Skin Response sensor where you will place two fingers is a battery-driven device, which will measure the moisture in your skin. The desk-mounted TOBII T60 eye tracking system that will be capturing the image of your eyes also uses very low power infrared light to illuminate your eyes, so that the location and the size of your pupil can be continuously measured. None of the sensors used in the experiment are invasive, and they will not cause any form of sensation in your fingers.
2. During the performance of the experiment, you will be instructed automatically by the computer to participate in a Stroop Color-Word Interference Test, which is implemented with Macromedia Flash. In this test, you will be shown words that designate color names. Regardless of the meaning of the word shown, you need to click one of the five buttons shown on the screen, to indicate the font color used in the word shown.

RISKS AND/OR DISCOMFORTS
No significant risks are expected to be associated with your participation in this study. It is not foreseen that the experiment may cause you any harm or severe discomfort. However, if you get upset or feel discomfort during the experiment, you may ask to take a break.

BENEFITS
You will not get any direct benefit from participating in the study. However, your help will give us valuable data and information in order to verify our system for the proposed study.
ALTERNATIVES
There are no known alternatives available to you other than not taking part in this study. However, any significant new findings developed during the course of the research which may relate to your willingness to continue participation will be provided to you.

CONFIDENTIALITY
The records of this study will be kept private and will be protected to the fullest extent provided by law. In any sort of report we might publish, we will not include any information that will make it possible to identify a subject. Research records will be stored securely and only the researcher team will have access to the records. However, your records may be reviewed for audit purposes by authorized University or other agents who will be bound by the same provisions of confidentiality.

COMPENSATION & COSTS
You will not receive payment for your participation in this study. On the other hand, you will not be responsible for any costs to participate in this study.

RIGHT TO DECLINE OR WITHDRAW
Your participation in this study is voluntary. You are free to participate in the study or withdraw your consent at any time during the study. Your withdrawal or lack of participation will not affect any benefits to which you are otherwise entitled. The investigator reserves the right to remove you without your consent at such time that they feel it is in the best interest.

RESEARCHER CONTACT INFORMATION
If you have any questions about the purpose, procedures, or any other issues relating to this research study you may contact Dr. Armando Barreto at the FIU Engineering Center (10555 West Flagler Street), Room EC-3981, Miami, FL 33114, (305) 348-3711, or by email at barreto@fiu.edu.

IRB CONTACT INFORMATION
If you would like to talk with someone about your rights of being a subject in this research study or about ethical issues with this research study, you may contact the FIU Office of Research Integrity by phone at 305-348-2494 or by email at ori@fiu.edu.

PARTICIPANT AGREEMENT
I have read the information in this consent form and agree to participate in this study. I have had a chance to ask any questions I have about this study, and they have been answered for me. I understand that I am entitled to a copy of this form after it has been read and signed.

Signature of Participant

Supriya Srinivasan

Printed Name of Participant


Signature of Person Obtaining Consent


05/18/2012

Date

05/18/2012

Date
Experiment Evaluation Form for Stroop Color-Word Test

This experiment evaluation form is just used for the statistical purpose.

How did you feel when you were doing the experiment?
1 means that you felt relaxed, just like you were in the environment of normal life.
5 means that you felt stressed, just like you found that you could not find the key of your house.
2, 3 and 4 represent the intermediate affective level from 1 to 5 respectively. (stress affect increases gradually with the number)

In the Congruent segment (the color of the font matches its meaning), I felt
1 2 3 4 5

In the Incongruent segment (the color of the font does not match its meaning), I felt
1 2 3 4 5

Name: Ameya Nayak
Date: 06/11/2012
VITA
PENG REN

Education

Ph.D. in Biomedical Engineering
Florida International University
Advisor: Armando Barreto
Aug. 2008 - Present
Miami, FL (U.S.A)

M.S. in Biomedical Engineering
University of Electronic Science and Technology of China
Chengdu, Sichuan (P.R.China)

B.S. in Biomedical Engineering
University of Electronic Science and Technology of China
Chengdu, Sichuan (P.R.China)

Publications


