University of Strathclyde

Department of Electronic and Electrical Engineering

Video Processing Analysis for Non-Invasive Fatigue Detection and Quantification

By

Masrullizam Mat Ibrahim

A thesis presented in fulfilment of the requirements for the degree of

Doctor of Philosophy

2014

Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has lead to the award of a degree.

Signed:
Name: Masrullizam Mat Ibrahim
Date:

Acknowledgements

It is a pleasure to thank the many people who made this thesis possible. I would like to begin by thanking my supervisor Prof. John Soraghan and Dr. Lykourgos Petropoulakis for giving me an excellent opportunity to carry out a very innovating and challenging research on video processing analysis for non-invasive fatigue detection and quantification. Prof. Soraghan and Dr. Petropoulakis provided me endless source of ideas and encouragement. Their enthusiasm in research, systematic organization at work and optimistic attitude towards life positively influenced my study and life. Their assistance and instructions during my time at the University of Strathclyde has been invaluable.

Not forget to my entire colleague in Centre for excellence in Signal and Image Processing (CeSIP) University of Strathclyde, which gave full support in term of technical, ideas and motivation. For the team project of sleep deprivation experiment which granted by Bridge the Gap (BTG) funding (Professor John Soraghan Dr Stephen Butler Dr L Petropoulakis Dr Simon Kyle, Mel McKendrick, and Jillian Hobson) and then further continued with the Scottish Sensor System Centre (SSSC) project (Gaetano DI Caterina, Carmine Clemente, George Mermiris, and Captain Lewis), the experience working with you all very valuable.

The most important, the closest vicinity of my PhD life, I owe my loving thanks to my wife Rainah Ismail and my son Muhammad Amir Najwan Masrullizam, and my daughters Aliyah Damia Masrullizam and Asma' Raudhah Masrullizam without their invaluable understanding, endless patience and encouragement it would have been impossible for me to finish this work.

My deepest and sincere thanks are always with my parents, parent-in-laws, also, to my brothers, brother-in-laws and sister-in-laws, also, the Malaysian Glaswegian friends for their constant support and prayers. I am so lucky and so proud to have such a wonderful family and friends.

Lastly, highly appreciation to The Ministry of Higher Education Malaysia and Universiti Teknikal Malaysia Melaka (UTeM), for the whole PhD study funding.

Abstract

Fatigue is a common symptom of weakness either physically or mentally. These symptoms may led to a drop in motivation, weakened sensitivity, slowing of responsiveness and inability to give full attention. All of these problems can cause adverse effects, such as accidents, especially those that require full attention as drivers of vehicles, and rail operators, the pilot of an aircraft or ship operators. This research investigates systems to detect and quantify the signs of fatigue using non-invasive facial analytics.

There are four main algorithms that represent the major contribution from the PhD research. These algorithms encompass facial fatigue detection and quantification system as a whole. Firstly, a new technique to detect the face is introduced. This face detection algorithm is an affiliation of colour skin segmentation technique, connected component of binary image usage, and learning machine algorithm. The introduced face detection algorithm is able to reduce the false positive detection rate by a very significant margin. For the facial fatigue detection and quantification, the major fatigue signs features are from the eye activity. A new algorithm called the , Interdependence and Adaptive Scale Mean Shift (IASMS) is presented. The IASMS is able to quantify the state of eye as well as to track non-rigid eye movement. IASMS integrates the mean shift tracking algorithm with an adaptive scale scheme, which is used to track the iris and quantify the iris size. The IASMS is associated with face detection algorithm, image enhanced scheme, eye open detection technique and iris detection method in the initialisation process. This proposed method is able to quantify the eye activities that represent the blink rate and the duration of eye closure.

The third contribution is yawning analysis algorithm. Commonly yawning is detected based on a wide mouth opening. Frequently however this approach is thwarted by the common human reaction to hand-cover the mouth during yawning. In this

research, a new approach to analyse yawning which takes into account the covered mouth is introduced. This algorithm combines with a new technique of mouth opening measurements, covered mouth detection, and facial distortion (wrinkles) detection. By using this proposed method, yawning is still able to detect even though the mouth is covered.

In order to have reliable results from the testing and evaluating of the developed fatigue detection algorithm, the real signs of fatigue are required. This research develops a recorded face activities database of the people that experience fatigue. This fatigue database is called as the Strathclyde Fatigue Facial (SFF). To induce the fatigue signs, ethically approved sleep deprivation experiments were carried out. In these experiments twenty participants, and four sessions were undertaken, which the participant has to deprive their sleep in 0, 3, 5, and 8 hours. The participants were subsequently requested to carry out 5 cognitive tasks that are related to the sleep loss.

The last contribution of this research is a technique to recognise the fatigue signs. The existing fatigue detection system is based on single classification. However, this work presents a new approach for fatigue recognition which the fatigue is classified into levels. The levels of fatigue are justified based on the sleep deprivation stages where the SFF database is fully used for training, testing and evaluation of the developed fatigue recognition algorithm. This fatigue recognition algorithm is then integrated into a Fatigue Monitoring Tool (FMT) platform. This FMT has been used to test the participant that carried out the tasks as ship crew in shipping bridge simulator.

Contents

List of Figures

List of Tables

Abbreviations

AAM	Active Appearance Model
AdaBoost	Adaptive Boosting
AECS	Average Eye Closure Speed
ARRB	Australian Road Research Board
ASM	Active Shape Model
ASTiD	Advisory System for Tired Driver
AU	Action Unit
AVMED	Institute of Aviation Medicine
BN	Bayesian Network
CAMSHIFT	Continuously Adaptive Mean Shift
CAS	Circadian Alertness Simulator
CAS-PEAL	Chinese Academic of Science - Pose, Expression, Accessories, and Lighting
CASIA	Chinese Academic of Science
CeSIP	Centre for excellence in Signal and Image Processing University of Strathclyde
CMU	Carnegie Mellon University
COPD	Chronic obstructive pulmonary disease
DFM	driver fatigue monitoring system
ECG	Electrocardiograph
ED	Distance between the centre of eyes
EDVTCS	Engine Driver Vigilance Telemetric Control System
EEG	Electroencephalography
EOG	Electrooculography

EMD	Distance between centre of mouth and the middle point between eyes
faceLAB™	Face Laboratory
FAID	Fatigue Audit Interdyne
FACS	Facial Action Coding System
FER	Focus of Eye Region
FERET	Facial Recognition Technology
FFS	Forward Features Selection
FMR	Focus Mouth Region
FMT	Fatigue Monitoring Tool
FRMS	Fatigue Risk Management System
FSI	Flag State Implementation
FWR	Focus Wrinkles Region
GUI	Graphical user interface
HIV	Human immunodeficiency virus
HOG	Histograms of Oriented Gradients
HMM	Hidden Markov Model
HSI	Hue Saturation Intensity
HSV	Hue Saturation Value
IASMS	Interdependence and Adaptive Scale Mean Shift
ICAO	International Civil Aviation Organization
ICE	Iris Challenge Evaluation
IMO	International Maritime Organisation
IR	Infrared
ISM	International Safety Management
Lab	Laboratory
LBP	Local Binary patterns
LDA	Linear Discrimination Analysis
LED	Light Emitting Diode
MIT	Massachusetts Institute of Technology

MLR	Multinomial Ridge Regression
MMU	Malaysia Multimedia University
mo	mount opening
NIR	Near Infrared
NN	Neural Network
PCA	Principal Component Analysis
PERCLOS	Prominent technique
PIE	CMU Pose, Illumination, and Expression
PsyKE	Psychology Knowledge Exchange & Enterprise Unit University of Strathclyde, and Glasgow Sleep Centre
PVT	Psychomotor vigilance task
RAAF	Royal Australian Air Force
RGB	Red Green Blue
RLBP	Regional Local Binary Pattern
SAD	Sum of Absolute Difference
SAFE	System for Air Crew Fatigue Evaluation
SART	Sustained Attention to Response Task
SFF	Strathclyde Facial Fatigue
SGLD	Second order statistical features
SMS	Safety Management System
SURF	Speed Up Robust Features
SVM	Support Vector Machine
3D	Three dimensional
TPMA	Three Process Model of Alertness
UBIRIS	Irises database from University of Beira Interior
UEC	University Ethic Committee of University of Strathclyde
YAWN	Yawning
YcbCr	Luminance and chroma component colour space
YR	Yawn rate
ZJU	Zhejiang University

List of Symbols

δ	Delta
θ	Theta
α	Alpha
f	PERCLOS curve over a certain period
t_1	The time the eye is closed for only 20%
t_2	The time when the eyes are 20% from completely closed
t_3	The times from eye open to eye 20% open (after being closed)
t_4	The times from eye open to eye 80% open (after being closed)
r	red
g	green
b	blue
c_b	Blue different chroma component
c_r	Red different chroma component
hue	Properties of colour
$y_{\min}$	Minimum y axis
$y_{\max}$	Maximum y axis
$x_{\min}$	Minimum x axis
$x_{\max}$	Maximum x axis
R	radius
ϵ	epsilon

$B(x, y)$	A bounding rectangle
sm	Small region of pixels
bg	Large region of pixels
$g'(x,y)$	Enhanced input image
$g(x,y)$	Input image
$Gl_{\min}$	The minimum input image intensities
$Gl_{\max}$	The maximum input image intensities
$Gl'_{\min}$	The transformed minimum input image intensities
$Gl'_{\max}$	The transformed minimum input image intensities
H_i	The cumulative histogram
X	The intensity value
aT	The adaptive threshold
$P_{\min}$	The minimum pixel value
$P_{\max}$	The maximum pixel value
sr	The aspect ratio of the bounding box shape
hbb	Height of then bounding box shape
wbb	Width of the bounding box shape
r	The iris radius
x_0	The coordinate of the iris centre in the x-direction
y_0	The coordinate of the iris centre in the y-direction
$I(x,y)$	The input iris image
$G_\sigma(r)$	The Gaussian function

$\hat{q} = \{\hat{q}_u\}_{u=1...m}$	The probability of the colour histogram of the iris
M	The number of histogram bins
x_i	Normalised pixel location from 1 to n with the target iris centred at 0
δ	Kronecker delta function
$b(x_i)$	The bin for pixel x_i
k	The Epanechnikov kernel function
C	Normalisation constant
$\hat{p}_u(y)$	The probability of colour histogram of the target iris candidate model
Y	Centre position of the current frame
H	Radius of weighting kernel
C_h	The normalisation constant
$d(y)$	The centre of the iris
$\hat{p}(y)$	Estimation of the Bhattacharyya coefficient
y_1	New location targeted iris
w_i	The weight
M_{00}	Zeroth moment of the region
M_{10}, M_{01}, M_{11}	First-order moments
$I(x, y)$	Probability pixel value within the object region in x and y range
x_c, y_c	The centroid point of the region
M_{20}, M_{02}	Second-order moments
μ_{20}, μ_{02}, μ_{11}	Rotation of ellipse
θ	Degree of orientation of the ellipse

a, b	The semi-major axis of the ellipse
A	Area of region computed from the zeroth moment
l_1	Length 1
l_2	Length 2
C_r	Centre point of the right eye
C_l	Centre point of the left eye
D_e	Distance between centre points of the irises
I_{FER}	Size of the FER
I_{SAD}	Sum of Absolute Difference (SAD) value of FER in between two frames
I_{SAD}	The normalise value of SAD
W	Width of FER
H	Height of FER
T_{AOi}	Threshold value of the iris area
Rem	The ratio of ED and EMD distances
x_1	Centre of right eye
x_2	Centre of left eye
ml	The measured length of mouth
mh	The height of the mouth
YR	The ratio of hl to height of FMR
$LBP_{P,R}(x_c, y_c)$	The result of Local Binary Pattern
i_c, i_p	Gray level values of the central pixel
	Surrounding pixels in the circle neighborhood

P	Surrounding pixels
R	Radius
ri	Rotational invariant
$s(x)$	Function binary LBP
$u2$	Uniform pattern
$riu2$	rotational invariant pattern with uniform pattern
$ROR(x,i)$	Circular bitwise right shift on the P-bit number x with i time
$LBP_{(P,R)}^{u2}$	LBP uniform pattern
$LBP_{(P,R)}^{riu2}$	Combination of rotational invariant pattern with uniform pattern
G_x	The gradient for the horizontal directions
G_y	The gradient for the vertical directions
FWR_{SAD}	Sum of the absolute values FWR
$Normalised FWR_{SAD}$	The normalise value of SAD
BR	Blink rate
T_{tec}	Total time of eye closed
A_{tec}	Average time of eye closed
Ac_{BR}	Accumulated BR
Ac_{NTtec}	Accumulated Normalised T_{tec}
Ac_{Atec}	Acculumated A_{tec}

Publication

Published and Presented:

1. N. A. Manap, G. Di Caterina, M.M. Ibrahim, and J.J. Soraghan.*"Co-operative surveillance cameras for high quality face acquisition in a real-time door monitoring system,"* 3rd European Workshop on Visual Information Processing (EUVIP), pp. 99-104, 2011.
2. G. Di Caterina, N.A. Manap, M. M. Ibrahim,and J.J. Soraghan*, 'Real time door access event detection and notification in a reactive smart surveillance system,"* Lecturer Notes in Computer Science, vol. 7340, pp. 459-467, Springer Berlin/Heidelberg, 2012.
3. M. M. Ibrahim, J. J. Soraghan, and L. Petropoulakis, "*Non-rigid eye movement tracking and eye state quantificatio*n," 19th International Conference on Systems, Signals and Image Processing (IWSSIP), pp. 280-283, 2012.
4. M. M. Ibrahim, J. J. Soraghan, and L. Petropoulakis, "*Covered Mouth Detection For Yawning,*" 3rd International Conference Signal and Image Processing Application (ICSIPA), Melaka Malaysia, Oct. 2013.
5. M. M. Ibrahim, J. J. Soraghan, L. Petropoulakis, "*Eye-state analysis using an interdependence and adaptive scale mean shift (IASMS) algorithm*", Elsevier, Biomedical Signal Processing and Control, Volume 11, Pages 53-62, 2014.

Submitted:

1. M. M. Ibrahim, J. J. Soraghan, and L. Petropoulakis' "*Yawning Analysis with Mouth Covered Detection*", Image and Vision Computing, Elsevier.

1. Introduction

1.1 Preface

Fatigue is synonymous with weakness symptoms that often occur either due to physical or mental health issues. Many people experience these symptoms either due to excessive physical activity or insufficient sleep. Fatigue is also a major symptom for those who are not healthy. Health care is a key to reducing the symptoms. This is done through a balanced diet, regular exercise to stay fit and sufficient rest and sleep. However, issues such as time constraints to exercise and high commitment to a career, make this symptom difficult to avoid [1, 2].

Fatigue effects cause a person to be in a non-normal state. Fatigue may lead to decreased level of concentration, slow the mental and physical reaction and may even cause a person to fall asleep. These effects which may lead to accidents happen easily, especially, to drivers of vehicles or those who work with heavy machinery. Accident statistics have shown that fatigue was the main factor amongst those who contributed the highest accident rate [3, 4]. Particularly, in industrial automotive and shipping, it is a major cause of accidents [5, 6]. Therefore, in order to reduce the negative impact of this symptom, a fatigue detection system which will be able to detect fatigue at an early stage and alert the user is desirable.

In general fatigue detection systems are divided into two categories, invasive and non-invasive systems. Invasive fatigue detection systems it is required the user to wear or attach sensory devices on part of the higher body. In general, the devices of this category measure the particular signals from the specific part of the body such as Electrocardiograph (ECG), Electroencephalography (EEG) and Electrooculography (EOG) [7-9]. These sensors must be worn correctly to avoid errors in measuring the signals. This approach, however, may lead to discomfort for the user, over and above the

possible difficulty of wearing the device, or it may interfere with the user movement. For non-invasive systems, the user is not required to wear any device, with the system operating automatically to detect fatigue signals. In this thesis, the research conducted focuses on a non-invasive system where fatigue signals are detected through face observation using a camera technology. Thus, the overall contents in this thesis discuss the algorithms for detection and quantification of the fatigue signs that were developed using image and video processing techniques.

1.2 Research Motivation

The human face is rich in information representing human behaviour and it provides a visible manifestation of a person's effective state, cognitive activity, intention personality and also includes spontaneous reactions due to fatigue. All this information can be read by detecting and measuring the reaction and movement of facial components such as mouth, eyes, nose, cheeks and forehead. For fatigue symptoms, the eye is the primary face component that represents fatigue signs. Dinges *et al*. [10] conducted an experiment to measure the alertness, as assessed by psychomotor vigilance, and found that the technique which measures the percentage of eyelid closure over time which reflects slow eyelid closures (droops rather than blinks) was valid to determine the alertness of drivers. This prominent technique is called PERCLOS and has been applied in a number of systems for detecting driver fatigue [11-14].

PERCLOS is a good technique to detect the level of vigilance of motorists. However, for some applications which also require fatigue detection systems and which involve, for example, airplane pilots, ship operators, or operators in control rooms, the PERCLOS technique is too sensitive. Therefore, we propose a new technique to detect, measure and classify fatigue according to different levels. The proposed measure of fatigue is based on the hours of sleep deprivation, which denotes the fatigue level equivalent to a person who did not sleep a given number of hours overnight. This measure of fatigue is chosen because many problems caused by fatigue are due to lack of sleep time as occurs, for example, in the shipping and aviation industries. This is

mainly due to the shift work which crews in shipping lines have to do while they are out at sea for a long time.

To classify fatigue according to hours of sleep deprivation, sleep deprivation data are needed for the development of algorithms, for testing and for evaluation. To obtain datasets, sleep deprivation experiments were conducted which involved 20 participants. Participants were sleep deprived for 0, 3, 5 and 8 hours. In each experiment participants were required to carry out certain cognitive tasks whilst their facial activities were recorded. These facial activities recording served as the main source for this research for training, testing and evaluating the developed algorithms.

1.3 Summary of Original Contributions

The main research contributions of this thesis are described below:

1. The first contribution is in the face acquisition operation, where a novel technique for detecting human faces is introduced. The first stage of the face detection algorithm is to locate the position of a face in the image before the next process is carried out. The introduced technique combines skin colour segmentation, connected components of binary image, and a machine learning classifier. The skin is segmented based on three colour spaces, RGB, HSV and YCbCr. The connected components are applied to eliminate unwanted segmented skin region. Lastly, the machine learning classifier is used to classify the remaining segmented colour skin region. In this research the Viola Jones classifier [15, 16] is applied to detect the face. With the use of the proposed technique the false positive error is significantly reduced.
2. The second contribution is to present a novel algorithm to quantify the eye state. The algorithm is able to detect and quantify the eye state in order to measure the activity of eye which represents any fatigue signs. The introduced novel algorithm is called Interdependence and Adaptive Scale Mean Shift (IASMS) algorithm. This algorithm applies a mean shift algorithm in combination with an

adaptive scale scheme in order to track non-rigid eye movement and measure the size of the iris in real time processing.

3. The third contribution is a novel technique to analyses yawning. Practically yawning is detected based upon the size of the mouth opening. In this novel technique, yawning may also be detected even though the mouth may be covered (by the palm of the hand usually). By using this technique the period of yawning can be measured.

4. The fourth contribution is the Strathclyde Facial Fatigue (SFF) videos footage database. This database is developed from the sleep deprivation experiments which involved twenty participants. There are four experimental sessions with participants being sleep deprived for 0, 3, 5 and 8 hours. The sleep deprivation experiment is setup and conducted by the Psychology Knowledge Exchange & Enterprise Unit (PsyKE), University of Strathclyde, and the Glasgow Sleep Centre, University of Glasgow From the experiments the participants' face activities are recorded whilst they carry out cognitive tasks. This database was used for training, testing and evaluation of the developed algorithms.

5. The last contribution is a new novel approach for fatigue recognition. The existing fatigue recognition systems are based on a single classification. In this research the fatigue signs are classed into four levels of fatigue. These levels are determined based on the sleep deprivation data from the SFF database. In this fatigue recognition system the features are extracted from the measurement of the eyes' activities using IASMS algorithm, and from yawning analysis results.

1.4 Thesis Organization

The reminder of this thesis is organised as follows:

Chapter 2 gives an overview of the recent field of fatigue detection, which covers the application of fatigue detection in several fields such as automation, aviation and

shipping industries. Then, the categories of fatigue detection based on measurement techniques are discussed. Lastly, the criteria of user acceptance to fatigue detection systems are reviewed.

Chapter 3 discusses the facial fatigue analysis systems that contain three operations, face acquisition, facial fatigue features extraction, and fatigue recognition. This chapter reviews the existing research for each of the operations. The related techniques that are applied in the next chapters are explained in detail and the most relevant techniques are critically reviewed.

Chapter 4 reviews the databases which are used in every stage of the facial fatigue system operation. This data is used for training, testing and evaluation of the developed algorithms. The main content of this chapter is the discussion on the novel SFF database. The data acquisition and experimental procedures are described, whereas, the experiment tasks can be referred in Appendix B.

Chapter 5 introduces two contributions of this research, face detection algorithm and Interdependence and Adaptive Scale Mean Shift (IASMS) algorithm for eye state analysis. The experimental results from these two novel algorithms are shown.

Chapter 6 introduces the novel algorithm to analyse yawning. There are three new techniques introduced in this algorithm, which involve the process to measure the mouth opening, covered mouth detection, and the wrinkle changes detection. The analysis results from the experiments using genuine yawning from the SFF database are discussed.

Chapter 7 presents the novel fatigue recognition technique which classifies the fatigue into various levels. The techniques to extract the feature vectors from IASMS and yawning analysis algorithms are explained. The developed algorithms that were tested in Fatigue Monitoring Tool (FMT) are described. The results for evaluating the performance of the developed algorithm are also shown.

Chapter 8 concludes the thesis along with suggestions for the future work.

2. Fatigue Detection

2.1 Introduction

Fatigue is a syndrome that is known as tiredness, exhaustion or lethargy, and it is a common health complaint. Generally, fatigue is defined as a feeling of lack of energy and can be caused by inadequate rest, long hours of physical or mental activity, sleep disturbance, excessive stress or a combination of these factors. In addition, the environment, place and time are also factors that can accelerate fatigue. Normally, a silent environment at night time can induce fatigue much easier compared to a noisy environment at daytime. In general, fatigue may result from several causes, which can be categorised into eight groups as shown in Table 2.1. Most of these causes are due to health conditions from which subjects may suffer. These causes may be difficult to manage or prevent from happening. However, some of these causes such as sleep deprivation, shift work and alcohol intake are manageable.

The symptoms of fatigue can be felt as well as having annoying effects such as impairment of hand-eye coordination, low motivation, poor concentration, slow reflexes and response, feeling overwhelmed and inability to pay attention. These symptoms increase the risk of errors in judgment and the tendency of risk taking. Furthermore, in more severe fatigue cases, the symptoms may result in failure to respond to changes in the surroundings and to information provided and also result in forgetting information. These effects can be truly dangerous, particularly for drivers or other transport operators as well as passengers.

This chapter gives an overview of the recent field of fatigue detection, which covers the application of fatigue detection in several fields such as automation, aviation and shipping industries. Then, the categories of fatigue detection based on measurement techniques are discussed. The criteria of user acceptance to fatigue detection systems are

subsequently reviewed, and the conclusions complete this chapter.

Table 2.1 Causes of fatigue

Categories	Causes of fatigue
Metabolic/ Endocrine	Anaemia; diabetes; Cushing's diseases; hypothyroidism; electrolyte abnormalities; kidney diseases; liver disease
Infectious	Cytomegalovirus; HIV infection; influenza (flu); malaria and many other infectious diseases
Cardiac (heart) and Pulmonary (lungs)	Congestive heart failure; coronary artery disease; valvular heart disease; chronic obstructive pulmonary disease (COPD); asthma; arrhythmias; pneumonia
Medications	Antidepressants; anti-anxiety medications; sedative medications; medication and drug withdrawal; antihistamines; steroids; some blood pressure medications; some antidepressants
Psychiatric (Mental Health)	Depression; anxiety; drug abuse; alcohol abuse; eating disorders (for example; bulimia; anorexia); grief and bereavement.
Sleep Problems	Sleep apnoea; sleep deprivation, reflux esophagitis; insomnia; narcolepsy; work shift work or work shift changes; pregnancy; extra night hours at "work"
Other	Cancer; rheumatology illnesses such as rheumatoid arthritis and systemic lupus; fibromyalgia; chronic fatigue syndrome; normal muscle exertion; obesity; chemotherapy and radiation therapy.

2.2 Fatigue in Industrial Applications

The consequences of fatigue such as decrease of vigilance, incorrect action and failure to respond to changes in the surroundings and information provided, may lead to serious injuries and, in the worst case, even death. For instance, in transportation

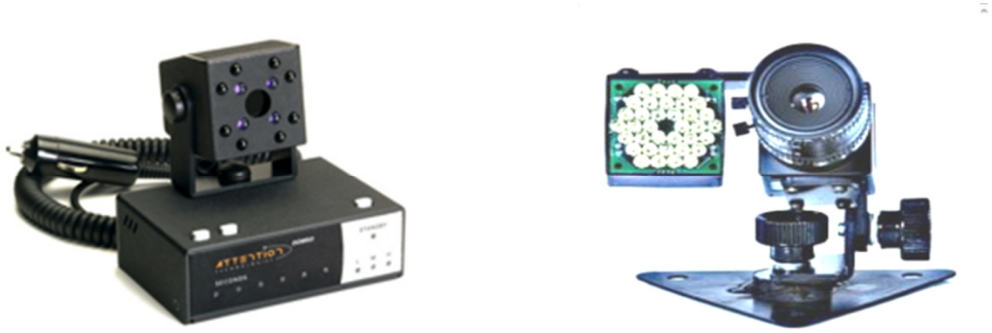
industry, an accident not only harms a driver, but also passengers and other road users may also be affected. The seriousness of these effects of fatigue, have resulted in the development of many fatigue detection systems as well as in the improvement of safety and management systems. Generally, the implementation of a fatigue prevention and detection system can be categorised into several groups of different applications, such as in automotive industry, aviation industry, and shipping industry.

2.2.1 Automation Industry

Significant development of fatigue monitoring and detection systems have emerged from the automotive industry. This is because fatigue contributes the largest number to road accidents.. According to the UK Department of Transport[17], in 2010, 1850 people were killed, and 22,660 were seriously injured. Furthermore, fatigued drivers have contributed to 20% of total road accidents [3]. In USA, the administration of National Highway Traffic Safety estimates that there are approximately 100,000 crashes each year caused by fatigue and drowsiness [4]. These percentages indicate the seriousness of fatigue impact on safety in road traffic, hence, many systems have been developed for fatigue detection and prevention.

There are several companies worldwide involved directly in the development of fatigue monitoring systems, particularly in ways for alerting the driver. For instance, Attention Technology Inc in San Diego US has introduced a Driver Fatigue Monitoring (DFM) system known as DD850 [18] (shown in Figure 2.1(a)) that uses infrared camera technology to monitor the eyes of the driver and measuring, in real-time, the eyes' position and eyelid closure. This device provides a visual gauge to represent the driver's drowsiness level from moderate to severe drowsiness. Two other companies that focus on eyes and face tracking are SensoMotoric Instruments GmbH [19] with their InSight^{TM} system, and Smart Eye AB [13] with the latest system named Smart Eye Pro 5.0 (shown in Figure 2.1(b)). Both systems offer non-invasive computer-vision-based operations which measure the head position, track the eyes and measure the behaviour of the eyes using a single Infra Red (IR) illumination camera to enable functionality in

bright lighting and also in dark conditions. Meanwhile, faceLAB^{TM}, which is a product from Seeing Machine, uses two cameras to allow a stereo system that provides wide field of view and to enable 3D function in the future. All the aforementioned products monitor the driver by keeping track of the eyes' behaviour such as eyelid openness, eye position, eye gaze, and blink frequency. In order to determine the fatigue alertness stage, most of the systems apply the prominent technique known as PERCLOS [10].



(a) (b)

Figure 2.1: (a) Driver Fatigue Monitoring (DFM) devices [18] and (b) Smart Eye Pro devices [13]. *(Images are permitted to be published)*

Currently, several automotive manufacturers have equipped their vehicles with fatigue monitoring technology. For example, Volkswagen Passat [20] has a system known as Driver Alert System. This system detects fatigue based on driver behaviour, steering wheel movement and lane deviation. When the system detects erratic steering wheel movement, the driver will be alerted through a visual display on the dashboard and a warning sound. Volvo [21] also has a system, almost similar to Volkswagen, which alerts the driver when the car crosses one of the road markings without an obvious reason. In addition, Mercedes Benz, not wanting to be left behind in this technology, introduced Attention Assist [22]. In this system, the driver behaviour is initially observed in order to create a unique driver profile in terms of speed, longitudinal and lateral acceleration, angle of the steering wheel, the way that the indicators and pedals are used. Then, the system will alert the driver when the driving

pattern has changed within a specific time interval. By contrast Toyota's [23] fatigue monitoring system is based on analysis of the driver face and head movement with near-infrared technology employed to monitor the eyelids and head position of the driver. Through integration with the Advanced Pre-Crash Safety system pre-crash warnings are issued. For example, when the Advanced Pre-Crash Safety system detects an obstacle ahead, and at the same time the Driver Monitoring System establishes that the driver's head has been turned away from the road for too long, the system automatically activates pre-crash warnings.

2.2.2 Aviation Industry

In the aviation industry, according to the report in [24], fatigue has contributed up to 10% of the aviation accidents. In order to reduce the percentage the accidents, action have been taken and plans have been made to ensure the crews are in the best possible condition whilst on duty. The International Civil Aviation Organization (ICAO) is a specialized agency that has set the standards and regulations necessary for safety and security. Regularity issues call for each aviation operator to operate a Fatigue Risk Management System (FRMS) [25]. FRMS is an advanced form of Safety Management System (SMS) [26] aiming to find a balance between safety, productivity and cost in an organisation. In an SMS system, the fatigue of crew members is managed through prescribed limits on maximum flight and duty hours based upon historical records and understanding of the fatigue symptoms through the normal work and rest period relationships. Additionally, in the FRMS, fatigue is taken into account by relating the effects of sleep and circadian rhythm as an added dimension of the management fatigue risk. Therefore, FRMS provides a means of allowing operators to work both safer and more efficiently.

However, FMRS is simply a management system, which is designed to manage and prevent crews from fatigue symptoms that can have adverse effects on them. The system appears to be incomplete without any element of monitoring and detection which is able to give an alertness and warning to crew pilot, especially during duty. The

Institute of Aviation Medicine (AVMED), which is one of the units of the Royal Australian Air Force (RAAF) involved with research in safety flying, has been working with Optalert Pty Ltd [27] to develop a device for a fatigue monitoring system for aviation environments. Optalert Pty Ltd is one of the leading industries in the development of fatigue detection technology and has produced a series of upgrades in fatigue detection systems since the company was established in 1994. In this collaborative research [28], Optalert Pty Ltd has produced a specific device to monitor the pilot on duty. Pilot behaviour is quite different compared to a vehicle driver. The main focus of a driver is on the road ahead whilst pilots spend less time looking ahead and most of the time they observe the controller panels. Optalert Pty Ltd has produced a device which employs a system of Infra Red (IR) reflectance oculography that has an IR transmitter and receiver bar attached to a spectacle's frame positioned below and in front of the eyes. The device measures the velocity of the eyelid closing and reopening during blink, and the duration of closure to quantify the level of fatigue. This device will alert the pilot when fatigue conditions have been reached.

2.2.3 **Shipping Industry**

Another industry that has similar considerations due to fatigue related situations is the shipping industry. In contrast to other two industries described in section 2.2.1 and 2.2.2, crew members in the shipping industry need to spend a very long time in the sea, which makes them suffer from lack of sleep and stress problems.

A conducted study in [29], where a sample of 79 crew and 6 patrol members were involved, showed that 44% of the participants worked more than 80 hours over a week, and 62% of them reported not getting enough sleep. Another study conducted for New-Zealand inter-island ferries [30] found that 61% officer crews felt they were always affected by fatigue when on duty. It was also found that 26% of the ferry samples were involved in a fatigue-related incident or accident in the last six months. Harma *et al*. [5] reported that as many as 40.6% of nautical officers falling sleep at work. Under International Maritime Organisation (IMO), a sub-committee called Flag

State Implementation (FSI), which deals with implementation of IMO instrument, port, flag and coastal state matters, have reported that 60% to 81% of all collisions and groundings were caused by human error factors [6].

In the Dutch shipping sector, as reported by Hotman *et al*. [6], a proper implementation of International Safety Management (ISM) with optimum organisation on board work patterns, lengthening of the rest periods and a reduction of administrative tasks reduced the fatigue levels among the crew members. Another system that has been implemented in shipping industry is the Watch System. This system provides the framework for the amount of rest individual shipping crew members can have. There a system measurement is introduced, a 2 shift-watch for crew working 12 hours and a 3 shift-watch for the crew with only eight hours on duty [31]. Based on study conducted in [32] a 3 shift-watch provides more rest than a 2 shift-watch. At the time of writing this thesis in author's knowledge, there were no commercially available systems for recognising fatigue amongst crew members in the shipping industry.

2.2.4 Other Industries Applications

Apart from the three industries described above, fatigue detection system also can be found in mining industry. ASTiD[33] is an example system, which monitor fatigue among mining worker. This system uses a steering sensor, with which their vehicle is equipped in order to measure the driver behaviour. Additionally, this system also provides a monitoring system for quantifying the duration of driving. The system measures whether the vehicle is being driven under monotonous conditions such as in a highway, mine pit, and long dull roads.

Further, the impact of fatigue also has led some researchers to study and attempt to resolve fatigue-based problems in the workplace. Most of these attempt to enhance the prevention and safety system as well as provide an effective work schedule [34]. The working schedule is important, especially for shift workers or those who work an extended duration. The statistics have shown an increasing number of accidents occur to workers after work in night shift and long extended work [35, 36].

2.3 Fatigue Detection Measurement Technique Categories

In general fatigue detection measurement techniques can be divided into the following five categories:Physiological measurement, Physical activity measurement, Behavioural measurement, Mathematical models and Hybrid techniques (a combination of two or more techniques). These techniques represent how potential fatigue signs are extracting and how these signs are examined in order to recognise a fatigue condition.

2.3.1 Physiological Measurement

Physiological measurement is a technique which quantifies the physiological element such as eye activities, brain activities, and heart rate. There are three physiological measures which are commonly applied: eye activities, electroencephalogram (EEG) which represents brain electrical activities, and electrodermal activities, associated with galvanic skin resistance.

2.3.1.1 Eye Activities

Most of the recent sleep detection devices apply the technique of measuring eye activities. The eye activities include blinking behaviour, eye closure, pupil size and eye movement pattern. The investigation of the effect of eye activities has been carried out extensively since the early 90s [2, 37]. Eye closure is an eye activity which closely related to fatigue symptoms. Dinges *et al.* were introduced a reliable technique called PERCLOS [10], which is based on eye closure. In this technique, the person is identified as fatigued when, over a certain period period of time, the eyes are closed for more than 80% of the time. This technique is used in many fatigue detection devices on the market nowadays.

In addition, there are also several systems which are based on the measurement of the eye blinks. Various blink characteristics such as rate and duration have been extensively investigated in relation to fatigue symptoms. The blink rate and blink duration are reported to increase proportionally with fatigue level [38, 39]. Therefore,

several researches focus to establish robust techniques for quantifying the behaviour of the eye blinks. For instance, Senaratne *et al*. [40] measure the duration of eye blink by applying an active flow technique and Hu *et al*. [7] used Electrooculography (EOG) signal to detect the blink and classified them by using Support Vector Machine (SVM) to determine fatigue. However, it should be noted that blinks are obviously sensitive to certain factors that may influence the rate of blinks such as the lighting condition. Hence, normally the blink rate is combined with eye closure measurement in order to result in more reliable detection such as products that have been commercialised by SensoMotoric Instrument Inc[19], Smart Eye AB [13], and Toyota Automotive [23].

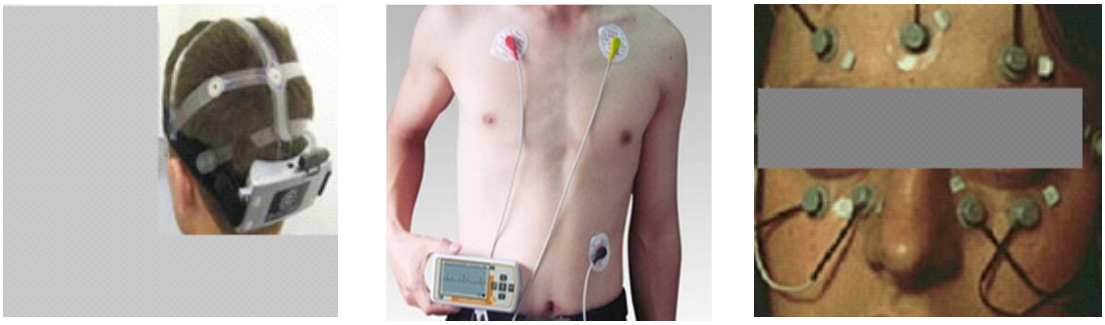
Another element of an eye that can represent fatigue symptoms is pupil size. It is found that fatigue is associated with a decrease in pupil size [41], and in a study [42] indicated that the pupil size has a strong relation with blink rate. In order to obtain the precise size of the pupil, Nishiyama *et al*. [43] used infra-red camera to measure pupil size in their research on fatigue for car drivers. Because the size is very small, a high specification camera is required that able to detect the pupil in certain distance. Since the change of pupil is very small in image, the illumination changes will affect the detection performance.

Instead of monitoring just eye activities some researchers have combined the eye activities with yawn detection. Yawn is absolutely a representative sign of fatigue, and which is described by the mouth opening much larger than normal. For instance, Omidyeganeh *et al*. [44] measured the mouth openness by applying a threshold value to YCbCr colour space, and yawn is determined based on the height of mouth opening. This technique depend only on single colour format, which lead the unstable performance because of illumination changes. Another approach is a utilization of a learning machine to train and classify the yawn, but the approach definitely requires a lot of yawning images. The crucial issue to be taken into consideration in developing a yawn detection algorithm is that the mouth is often covered by the hand during a yawn. There is should be a technique able to handle this situation.

2.3.1.2 **EEG Based**

The brain normally produces tiny electrical signals when yawn is occurred this comes from the brain cells and nerves that send the signals to each other. The Electroencephalography (EEG) is utilised by recording this electrical signal activity as shown in Figure 2.2(a). The signals are significant in indicating the level of consciousness and alertness of the human brain. There are several algorithms that have been developed to measure the fatigue using EEGs for the all major EEG bands: delta is associated with slow movement of eye sleep, theta associates with inhibition of elicited responses, alpha associate with eye closure, and beta associates with focusing and alertness. Wilson and Bracewell [45] used wavelets to represent EEG in different scales, and applied neural networks to classify the signal of driver fatigue. Whilst, Zhou *et al*. [8] proposed a new method based on the bispectrum for feature extraction of EEG band and Budi *et al*. [9] examined the four electroencephalography (EEG) activities, (delta (δ), theta (θ), alpha (α) and for 52 subjects that show stable delta and theta activities over time, meanwhile, a slight decrease of alpha activity, and a significant decrease of beta activity.

In another approach EEG is also combined with other signals to increase the precision of fatigue detection. G. Yang *et al*. [46] combined EEG with Electrocardiograph (ECG) (shown in Figure 2.2(b)), and fatigue is classified using a dynamic Bayesian network. EEG can also be integrated with Electrooculography (EOG), as shown in Figure 2.2(c), which measures the resting potential of the retina eye. D. Sommer *et al*. [47] applied these two types of signals to determine a level of fatigue using Support Vector Machine (SVM) as a classifier. For using this EEG based technique, it is necessary to wear a device that is not convenient and it is likely to annoy users.



(a) (b) (c)

Figure 2.2: :(a) EEG measurement technique which the device apply on head(b) ECG measurement technique which device attached on body part and (c) EOG measurement technique which device attached on region of eyes.[45-47]. *(Images are permitted to be published)*

2.3.1.3 **Electrodermal Activities**

Electrodermalis, known as galvanic resistance, is a method of measuring electrical conductance of the skin. The electrical signal is varied based upon the moisture level and able to give an indication of physiological arousal, .such as fatigue Even though this technique is sensitive to many factors such as psychological stress, physical movement, and noise of humidity [48, 49], there are systems that have been established using this technique, such as an Engine Driver Vigilance Telemetric Control System (EDVTCS)[50].

EDVTCS is developed by company of NEUROCOM, which has been used within the Russian railway system. The galvanic skin resistance is designed as wristband as shown in Figure 2.3(a) so that it is simply to be worn by a railway operator. Another company uses the same measurement technique on its device, which is called 'SenseWear Armband'[51] (Figure 2.3(b)). However, BodyMediaInc. has developed this device for clinical use, purposely for scientific research and study for a patient program. Similar to the EEG based measurement technique, the user is required to wear a special device to measure the fatigue features.

(a)

(b)

Figure 2.3: (a) EDVTCS galvanic skin resistance sensor device (b) 'SenseWear Armband device, to measure electrical conductance of the skin. [50, 51]. *(Images are permitted to be published)*

2.3.2 Physical Activities Measurement

In this technique, the physical activities movement is measured, and it is commonly measured by actigraphy which is a method of monitoring the circle of human activity and rest. Actigraphy has frequently been used in research on sleep and circadian rhythms especially investigate the pattern of sleep and consciousness activities [52]. In order to determined fatigue state, the different amounts of body movement during sleep and wakefulness is identified [53].

This technique has also been applied to UK Civil Aviation industry by using a device called 'Actiwatch Alert' [54, 55]. This device is combined with the schedule duty system, which is used by the crew to minimise the unplanned sleep during on-duty days. Another device that uses the same technique is the 'SenseWear Armband' [51] that combines with the galvanic resistance technique. However, this technique is only able to detect precisely a person in sleep or nap condition over a relatively long period. Obviously, this technique is not suitable for situations that require full attention at all times like driving a car.

2.3.3 Behavioural Measurement

This technique is monitoring and measuring the behaviour of the subject based upon of the responses to the task that is being carried out. For example, the DAS200

Road Alert System introduced by Ellison Research Lab [56], which is a system monitoring lane deviation, inter vehicle distance, centre line and hard shoulder crossing during driving. The system will alert the driver when the tire crosses the specified lane. The same technique is also applied by Volkswagen [20] and Volvo [21], whilst the Mercedes Benz [22] has been combined with mathematical model techniques in order to detect the fatigue signs of a driver.

This measuring technique has also been applied in the mining industry, where ARRB Transport Research Ltd developed a device that is able to measure mining operator alertness level through the reaction of the operator to a visual and audible stimulus throughout the whole working shift [57]. The system is capable to identify reductions in alertness levels and allows intervention to counter the effects of fatigue. It is possible for this measuring technique to reach wrong judgments particularly when applied to fatigue detection system for vehicles. This is because poor road conditions, such as road damage, can cause drivers to swerve which will affect the decision of the system [39].

2.3.4 **Mathematical Model**

A mathematical model is developed to predict the fatigue and alert the user based on several factors used as inputs to the model such as a sleeping time, duration of current work and circadian influences. The outputs of the mathematical model formulation is a parameter of alertness, fatigue, sleepiness, performance measures and accident risk [58]. There are several models that have been developed for fatigue and alertness such as Fatigue Audit Interdyne (FAID)[59], Three Process Model of Alertness (TPMA) [60], System for Air Crew Fatigue Evaluation (SAFE) [61], and Circadian Alertness Simulator (CAS).

Fletcher and Dowson [59] have developed FAID which is a model based on sleep history and pattern of work and rest. The model uses a database that comprises 3500 days of wake, sleep and work, which is collected from 250 Australian shift workers. SAFE is a model that was developed for fatigue measurement on civil aviation crew members. In this model, an experiment was carried out for irregular pattern of

work and rest from 30 subjects in order to evaluate the output parameters of risk that associated with the duty schedule. In addition, there were 10,000 days of sleep alertness data that have been used by the CAS model. This database is collected from transportation workers. Essentially, using the mathematical model requires a huge amount of data set for training purposes. Models are also limited to the specific application based upon the trained data.

2.3.5 **Hybrid Technique**

Hybrid technique is a combination of one or more of the aforementioned measurement techniques. The primary aim of this combination is to obtain a more accurate assessment of fatigue from several fatigue related features. EEG signals were combined with eye activities by Rosario *et al*. [62], and a threshold is determined by prior experiment that conducted for fatigue people. Pritchett *et al*. [63] used EEG with the movement of the body including arm and head movement to detect fatigue. Furthermore, St John *et al*. [64] and Rodriguez *et al*. [65] incorporated the EEG, eye activities and behaviour measurement in order for more accurately detecting the fatigue. However this approach obviously the EEG device is required to be worn.

Behaviour measurement techniques are often implemented in automotive applications. For example, Tekade *et al* [66] have combined PERCLOS technique with driving information such as the lateral position, the steering wheel angle and the heading error. The fatigue is classified when the signs of fatigue are detected from one of the output measurement techniques. This approach is suitable only for monitoring the car driver. Furthermore, Fatigue Management International a company producing the Advisory System for Tired Driver (ASTiD) [33, 67], has combined the information from a steering wheel and a mathematical model to predict the sleepiness of the driver. The mathematical model is developed based on prior knowledge of the driver falling asleep pattern over a period of 24 hours. However, the sleep pattern of the people can be varied which difficult to standardize the detection system Meanwhile, the galvanic resistance has been applied together with user behaviour in the SenseWear product for predicting

the fatigue of users. The problem is the behavior or people are always different, that difficult to classify the fatigue signs. [51].

2.4 User Acceptance for Fatigue Detection Technologies

The development of a fatigue detection and monitoring system should consider several factors of user acceptance as discussed in [68-70], and a number of evaluation criteria of user acceptance should be taken into account for establishing fatigue detection systems. User acceptance can be divided into five elements: the ease of use, ease of learning, perceived value, advocacy, and user behaviour.

The element of ease of use encompasses the degree to which users find a system understandable, usable, and intuitive in its operation and maintenance. The system should not be designed to contribute stress and workload to users. Ease of learning should assess how much knowledge is required by the user to ensure smooth operation of the system. Furthermore, for perceived value, an assessment must be made of the degree to which the user perceives a safer and more alert condition as a function of applying the device. In the case of advocacy, the user considers endorsing or purchasing a system if a reliability indicator is proven. Considering, user behaviour, the system designer should anticipate possible user behaviour of user make the system robust against this.

Generally, the fatigue detection systems are grouped into intrusive and non-intrusive systems. A system is in the intrusive or invasive group when equipment used in the system give discomfort to the user, such as equipment that are worn and preventing the movement of the user. For example, EDVTCS galvanic skin resistance sensor and 'SenseWear Armband' require the user wear on the wrist and arm. Similarly, techniques based upon EEG signal, where there is specific device need to be attached on the body in order to read such brain and heart signals. In contrast, a non-invasive or non-intrusive system is one where no annoying devices need to be worn or are attached to the user. The common technique for non-invasive systems is to employ video and image processing technology. By using this approach, a fatigue detection system operates

automatically without the need of any advance preparation or setup. One or more cameras are the only required devices to act as sensors for establishing a fatigue detection system as reported in [13, 19, 22, 23, 27].

2.5 Conclusion

This chapter has reviewed fatigue detection in terms of application in industries, the categories of fatigue detection based upon measurement techniques and the criteria of user acceptance to fatigue detection systems in general. There are three main industries that have been considered, namely the automotive industry, the aviation industry, and the shipping industry. Automotive industry is the largest research and development effort for fatigue detection system. The aviation industry has established a process of fatigue prevention system aiming to ensure that crew members are in the best possible condition while working. The last industry, shipping, it aims to ensure the crew members who have inadequate rest understand their level of alertness. In author's knowledge, at present there are no commercially available fatigue detection systems available to the shipping industry. Apart from these three industries, fatigue detection is also implemented in the mining industry, as well as in the workplace, the research has paid attention to working schedule systems, especially for those workers who work extended hours and night shifts.

In general, the development of fatigue detection systems should also take in to account several criteria of user acceptance. Fatigue detection systems are commonly grouped into invasive and non-invasive systems, which based on how the device facilitates the users. Recently most of the development of fatigue detection system focuses on non-invasive methods, which mostly apply to image and video processing technology which measure the signs of fatigue in ways that are most acceptable by users.

3. Facial Fatigue Analysis

3.1 Introduction

The previous chapter reviewed several fatigue detection categories based on measurement techniques. This chapter discusses the facial fatigue analysis that is utilized in physiological measurement techniques. The face, in general, presents rich information that describes the behaviour of humans, which conveys the identity, emotive intent and disseminates. As discussed in Chapter 2, eyes and mouth are the facial components which most likely to provide the most distinctive signs of facial fatigue symptoms. In order to determine the level of fatigue, the facial feature component activities that represent the dominant signs of fatigue will be measured. The combination of the fatigue features of these components are classified.

A facial fatigue analysis system mainly is comprised of several operations as shown in Figure 3.1. This chapter reviews the most prominent techniques, which are employed in every operation of these systems. Firstly, the face acquisition part is discussed, and several categories of face detection algorithms and facial features component detection algorithms are reviewed. Then, techniques to extract the facial fatigue features from the dominant potential signs of fatigue such as eye activities and yawning detection are discussed. The last operation is fatigue recognition, and several techniques of single cue and multi-cue fatigue features are described. Conclusion of the chapter is provided.

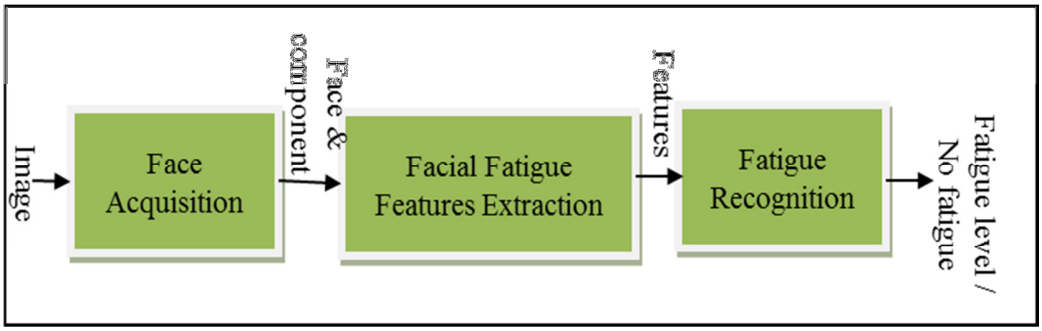


Figure 3.1: Facial Fatigue Analysis system that consists of three main operations where the input is image and the output is the fatigue level or no fatigue indication

3.2 Face Acquisition system

A face acquisition system is a crucial operation step in any facial analysis. This procedure is commonly performed before any specified operation for analysis of the localised face is carried out. During the face acquisition, there are two algorithms that are implemented sequentially the face detection algorithm and the facial features component detection algorithm.

3.2.1 Face Detection Algorithm

The task of a face detection algorithm is to locate the face in the image. The algorithm utilises several features that represent the face in order to precisely locate its position. Generally the face detection can be categorised into four methods based on how the facial features are represented and utilised for the face detection algorithm. These methods are: knowledge based method, features invariant method, template matching method and appearance based method [71, 72].

3.2.1.1 Knowledge Based Method

Knowledge based methods contain sets of rules that describe the sequence of the process to identify the face based upon the knowledge of the human face [71]. For instance, a face appears with two eyes which are symmetrically positioned, while a nose

location in between the eyes and the mouth and mouth size is wider than the eyes. A set of rules can be a combination of conditions for face detection algorithms such as the one implemented by Yang *et al*.[73]. After the possible face location is obtained, by using the sliding window approach scanning, the face is recovered by employing rules that define the characteristic of the eyes and mouth edges using the intensity of the image.

Kotropoulos *et al*. [74] applied the horizontal and vertical projection to find the face characteristics. The eyes, mouth and eyebrows are described by the darkest facial components, which project the highest values in horizontal and vertical projection. This straightforward face detection approach is easy to develop with simple rules to describe the features of the face and their relationships. For the symmetry frontal face with uncluttered background this method works effectively. However, using this method it is difficult to detect the face in different poses or multiple faces - especially in complex backgrounds. With this method it is also difficult to translate human knowledge into rules precisely.

3.2.1.2 **Features Invariant Based Method**

In features invariant based method the eyes, nose, mouth, ears and cheeks will be used as references to extract the facial features. The facial features could be from edge, intensity, shape, texture and colour of the face[71]. For example, Sirohey *et al*.[75] used edge as a feature to segment faces from a cluttered background. They used a Canny edge detector and heuristics to remove and group the edges so that only one face contour is preserved. Graf *et al*. [76] used greyscale image associated band pass filtering and morphological operations to enhance the regions with high intensity that indicate the shapes of eyes and mouth Then, by applying a threshold value and the connected component of the binary image, the essential facial components are identified. These features are then applied to the classifier to determine the face or non-face region.

Texture can be also used to extract the features of a face such as skin and hair. Augusteijn *et al*. [77]used this approach using second order statistical features Space

Gray Level Dependence (SGLD) to extract the feature vectors. Then, a neural network classifier was applied to train and classify every texture feature in the image. The feature most used and shown to be effective in face detection algorithms is skin colour. One of the colour spaces that can represent the human skin is YCbCr. The chrominance values Cb and Cr in the face region are narrowly distributed compared to the luminance value Y. Chai *et al*. [78] carried out an experiment using their dataset, and produced a range of chrominance values that denote the skin colour.

Singh *et al*. [79] have combined YCbCr with two other colour spaces RGB and HSI to segment. The segmentation of skin colour region is carried out separately for each colour space by employing threshold values specified by undertaking previous - experiments. Venn diagram theory has also been used to merge the segmented colour skin region. Since there are many colour spaces that can be used to identify skin colour in images Jose *et al*. [80] employed ten of the most commonly used colour skin approaches in order to find the best colour skin detection. This experiment consists of the detailed quantitative comparison for each colour spaces to decide which colour space model better aligns with skin colour. Based upon the experimental results, HSV colour space was found to be the colour space that best segments the skin region.

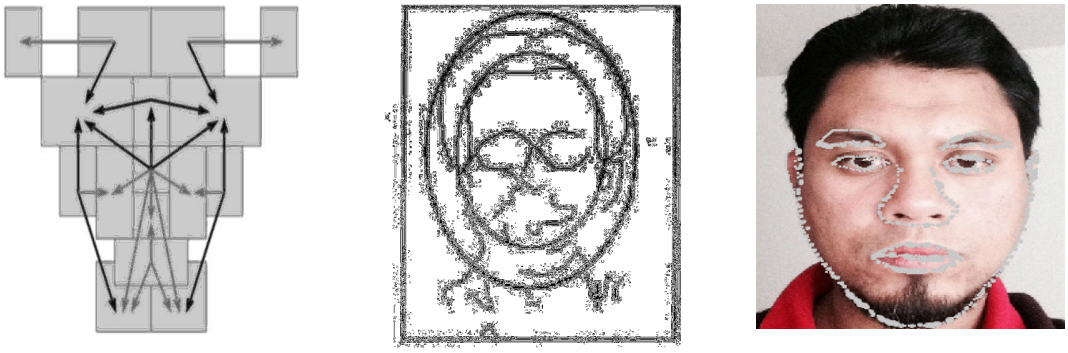
Feature invariant based method is very easy to implement especially for skin colour segmentation and intensity features. Features such as edges and shapes are invariant to pose and orientation changes. However, there are difficulties in term of extracting the features because of illumination, noise and occlusion in edge, shape and texture features. In complex backgrounds the features of the face can resemble the background features especially for skin colour.

3.2.1.3 **Template Matching**

In template matching method, a face pattern is manually predefined or parameterised by an algorithm. The face pattern is defined based on face components such as eyes, mouth and nose [71]. Generally, the template of the face can be

categorised into predefined template and deformable template. Figure 3.2(a) shows an example of a predefined template introduced by Scasscellati *et al*.[81]. The face is searched in the image based on this template, which comprises 16 regions that represent distinctive face components such as eyes and mouth, and also their 23 relations are defined. Miao *et al*.[82] applied a gravity centre template to detect the face. By using this template the algorithm is able to detect the face in multi-face rotations, but not suitable for complex background of image.

In deformable templates, the face characteristics are parametrically defined. For instance, Wang *et al*.[83] used the shape information obtained by applying a median filter and edge detector. The deformable template is designed by applying the elliptical ring as shown in Figure 3.2(b). The elliptical ring is able to represent the contour and almost all the edge points on the contour can be included. In addition, Edwards *et al*. [84] applied Active Shape Model (ASM) as shown in Figure 3.2(c), where is based on a statistical model template. By using ASM the face can be located, which the model estimate of the face location is based on the shape and intensity of the object. Generally, the template matching method is a simple algorithm and it is easy to detect the frontal face. But, it is inadequate for complex backgrounds and inefficient to deal with variations of pose, scale and shape of faces.



(a) (b) (c)

Figure 3.2: (a) Face template based on 16 face component regions with 12 relations as indicated by arrows. (b) Deformable template based on edges with elliptical ring around the face . (c) Active Shape Model (ASM) based template with the manually annotated is required . [83, 85, 86]. *(Images are permitted to be published)*

3.2.1.4 **Appearance Based Method**

In appearance based methods the image features are trained by algorithms unlike the template matching methods where the images are predefined by experts. In this method, statistical analysis, probability theory and machine learning algorithms are utilised [71, 87]. These algorithms associated with feature extraction approaches, such as the Harr-like feature extraction, Local Binary patterns (LBP), and Gabor texture feature extraction, in order to find discriminate functions for separating the hyperplanes, creating decision surfaces or setting the threshold function between face and non-face. Recently, the method's reliability has improved, so that it is implemented both in real-time applications and for experimental purposes [72].

There are many machine learning algorithms that have been applied and tested in order to find the best classifier which is robust enough particularly for real time implementation. For example, Shavers *et al.* [88] and Roohi *et al.* [89] have applied SVM to train the face features by dividing datasets in high dimension space based on kernel functions. Another machine learning algorithm is Principle Component Analysis (PCA) that transfers face images into a small set of characteristic image features called Eigen faces as implemented in [90-92]. Jing *et al.* [93] also combined PCA with SVM,

and used it to decrease the dimension of feature space for the purpose of training classifiers since PCA works well for analysing data in low and high dimensional subspace. Lastly, the SVM classifier is trained by using the face and non-face features represented by PCA, to verify the face candidate in the input image.

The prominent technique for face detection algorithm was introduced by Viola and Jones [16], which uses cascade classifiers to increase detection performance as well as to reduce computation time. This approach is proved to be efficient and robust [15, 94] and has been used as reference for many recent face detection algorithms [72, 87]. Figure 3.3 describes the cascade detection for a series of classifiers which were applied to every sub-window of input images. Each nod or stage represents a trained classifier using AdaBoost [95, 96]. Each stage has different complexity and the majority of sub-windows are rejected from the early stages until the classifiers become more complex to achieve low false positive rates. Viola and Jones applied Haar-like features to represent the face vector as shown in Figure 3.4. Haar-like features is computed by summing up the pixel intensities for every rectangular region, and find the difference of the values between them (between black and white regions). There are ten patterns rectangular that can be applied to produce the features vectors (Figure 3.4(a)). Figure 3.4(b) shows the location of Haar-like rectangular to extract features vectors from face. The details algorithm can be referred in Appendix A.

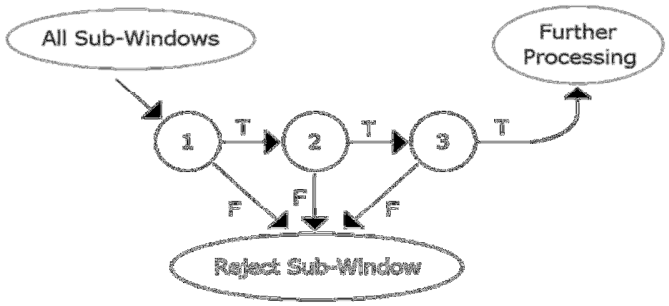


Figure 3.3: The cascade classifier which is a series of classifiers that applied to every sub-window of input images.

Based on the Viola Jones technique, several research outcomes resulted in improved ways to increase the robustness of face detection algorithms. For instance, Lee *et al*. [97] applied Multi-Block Local Binary Patterns (LBP) for feature extraction instead of Haar-like feature. Jianxin *et al*. [96] made improvements on the training side by introducing the Forward Features Selection (FFS) algorithm and fast pre-computing strategy for AdaBoost that can reduce training time by approximately 50 to 100 times. Minh-Tri *et al*. [98] improved the structure of the cascade classifier by creating a multiple exit structure. This technique trained a multi-exit boosted classifier by minimizing the number of weak classifiers which are required to obtain the targeted detection and false acceptance rates simultaneously. Each exit is associated with a weak classifier and a rejection decision at every exit is made if the intermediate boosted score is below a targeted threshold.

Appearance based methods utilise the capability of machine learning algorithms whose classification efficiency has been proved. In face detection this method is powerful, fast and fairly robust for real-time applications. The method demonstrates good empirical results and can detect the face for varying pose and orientation. However, this method needs a lot of positive and negative examples for training and testing. For instance, Viola and Jones [15] used 4916 face images and 9500 non-face images to train the classifier. This method also cannot escape from erroneous misclassifications, especially in the complex background images.

(a) (b)

Figure 3.4: (a) Example of Haar-like features pattern, (b) Example of Haar-like features extracted from the face [15].

3.2.2 Facial Features Component Detection Algorithm

A facial features component detection algorithm is a process to detect and find the location of the important facial components such as eyes, mouth, nose and eyebrows. These facial components describe the face activities which are crucial for the facial analysis operation. In order to accurately find every location of the components, the face should be located first. Then, the location of the components is obtained within the detected face region, which makes the algorithm not very complicated as face detection algorithm. Most of the developed algorithms have used eyes, mouth, and nose locations as reference points [76, 99, 100]. The other facial components can be located based on anthropometry of the face, which is calculated using the located reference components [101]. Generally in such cases, it is important to locate the eyes and mouth first. The techniques to detect these features can be categorised into; shape- based approaches, features based shape method and appearance based methods [71].

3.2.2.1 Shape Based Method

In shape based models, the eyes and mouth are detected based on the interior and exterior shape characteristics [102]. Basically in eyes tracking applications eyes are detected by the iris or the pupil shape. Basically the iris or pupils appear as elliptical shapes, Parez *et al*. [103] utilised threshold of image intensities to determine the centre of the pupil ellipse and edge detection is used to obtain the pupil boundaries. The problem of this technique is that several regions in the face may have intensities similar to the pupil or the iris region and the threshold values are difficult to set. In [104], Valenti *et al*. use isophote curvatures and edge orientation to detect the eyes. Isophote is a contour line of points of equal intensity obtained by slicing the intensity landscape with horizontal planes. However this technique may misclassify regions such as eyebrows that have a similar contour line.

For mouth detection, Eveno *et al*. [105] detected mouth region firstly by using horizontal intensities projection. Then the characteristic points of the mouth lips are

detected using a kind of active contour algorithm called 'jump snake'. This algorithm is easy to make adjustments to its parameter even when the initial point is far from the final edge. However, this technique is high computational for real-time application. Another technique that relies on the shape is the statistical shape modeling method introduced by Cootes *et al*. [106]. In this technique, the shape model is built from a set of shapes with corresponding landmarks. Cristinacce and Cootes [99] introduced the Shape Optimised Search (SOS) in order to improve the robustness and accuracy of detection. This shape model definitely requires a lot of image for modeling purposes.

3.2.2.2 **Features Shape Based Method**

In features based methods the common features such as edges, colour and texture are utilised to detect the facial component features [71]. For eye detection, the set of distinctive features around the eyes such as the edges, the border, the corner and the dark region of the pupil are normally used to represent eye features [107-109]. For example, Sirohey *et al*. [107] detect edges of eyes sclera with four Gabor wavelets. A nonlinear filter is formed to detect the left and right eye corners. The corners are used to locate the iris based on the edges. Another eye characteristic that can be used as a reference point is the region between the eyes as implemented by Kawato *et al*. [109].The brightness of the region between the eyes and the darkness of the region between forehead and the nose bridge are used as features to determine the location of the eyes.

Colour is one of the feature preferred by many researchers either for eyes or mouth detection since it is reliable and easy to implement [80]. Researches in [100, 108, 110] applied colour space to segment the lips colour, while, in [44, 74, 111] they projected horizontally and vertically the sum of colour value in order to find the location of the eyes and mouth. Even though the colour feature is easy to implement, there is a need to have algorithms that adapt to a variety of colour lips. This projection approach works well for symmetrical faces but not for non-symmetrical faces.

3.2.2.3 **Appearance Shape Based Method**

In this method, facial features component images must be used with statistical analysis, probability theory and the machine learning algorithm for training purposes [87, 112]. Referring to the success of the face detection technique introduced by Viola and Jones, a similar approach is applied to detect the facial features components as implemented in [100, 110, 113-116]. In addition, several others feature extractor techniques and classifiers have been introduced in the literature. For example, Vukadinovic *et al*. [108] detected facial feature components using a Gabor feature and boosted classifier. Kim *et al*. [117] used Speed Up Robust Features (SURF) to generate feature vectors for facial component features, and a two layer hierarchy of SVM is utilised for classification. This method can produce exceptionally good facial feature component detection with the detection rate is more that 95%, but a lot of images are required and the best features which optimise training parameters should be chosen.

3.3 Facial Fatigue Feature

As discussed in section 2.3, there are two face components activities which are distinctive in indicating the signs of fatigue: eyes' activity and yawning. From the eyes' activity the fatigue characteristics are detected from the blink behaviour, eye closure, pupil size, and eye movement. For yawning detection, the width of the mouth opening is measured with a wide open mouth typically recognized as a yawning. Moreover, other facial features components have also been considered by some researchers [118-120] for representing fatigue signs features.

3.3.1 Eye Activities

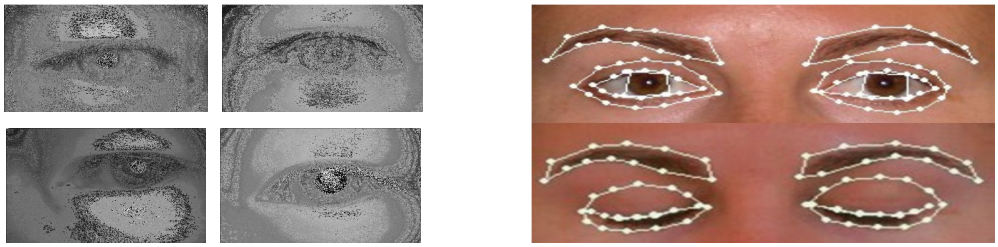
The eye activities represent the physical movement of the eye component that visually can be measured such as blinks and eye closure. In order to measure these eye activities, there are several techniques, which can be categorised in ways similar to face

detection and face components detection [71, 102], which are based on approaches that read the movement of the eye component. These categories are: template model based, features based, and appearance based.

3.3.1.1 **Template or Model Based**

The template based is an approach where the eye model needs to be designed first. In every image of a sequence of frames, the algorithm finds the high similarities ratio between the object in the image and the object that has been modeled. Wu *et al*. [121] and Chau *et al*. [122] designed the template model based for the open eye as shown in Figure 3.5(a). In the sequence of frames the correlation between the template model and the frame is computed. The eye is identified as being open or closed based upon the score of correlation, where a low score represents the eye being closed. This approach is easy to implement, but, it faces problems when dealing with multi-size of eye and multi-pose of face situations.

Another technique that is categorized in this group is a statistical active appearance shape model approach. In this technique, a set of manually annotated eye images, as shown in Figure 3.5(b), are required. These images with every annotated points are trained using statistical model algorithms, as applied in [123-125]. A large number of eye images are required for training purposes in order to ensure the algorithm performs robustly. With this technique, the problem of multiple-size seems to have been resolved, but the multi-orientation of face seems still to be limited on degree of orientation. Once again a lot of images for various faces and eyes shapes are required for training purposes.



(a) (b)

Figure 3.5: (a) Example of eye open used as a template [122], (b) Example of manually annotated eye image [123]. *(Images are permitted to be published)*

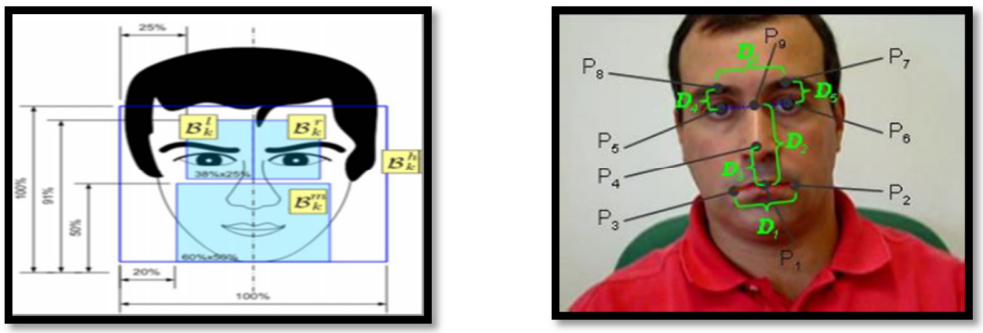
3.3.1.2 **Features Based Template Method**

The eye has several features, such as colour, texture, and edges that can be utilised to represent the eye activities. For example, Liying *et al.* [126] used the skin colour in order to distinguish between eye open and eye closed situations. Morris *et al.* [127] computed the variance of the colour map in the eye region to obtain the distinctive features of the eye state. Meanwhile, Azim *et al.* [128] measured the eye activities based upon the pupil size. The size was determined by detecting the bright part of the eye using an infrared camera. This approach of using colour features encounters the problem of illumination changes, which requires that algorithms must be able to deal with this challenge.

Instead of colour, distinctive points of the eye can be used to distinguish the eye‘s state as implemented by Jiménez-Pinto*et al.* [129]. They divided the face into right and left eye region, as shown in Figure 3.6 (a), and then by employing the Shi-Tomasi technique [130] the points of interest are obtained within these regions. The Shi-Tomasi salient points around the eye are tracked using the Lucas-Kanade algorithm. Based on these points, the eye pupil is localised and blink is detected from an intensity average of the eye pupil region.

In addition to the distinctive points, Batista *et al.* [101] used the eye corners to measure blinks. The eye corners are detected with reference to the eyebrow location based upon anthropometric face models as shown in Figure 3.6(b). In anthropometric

face model, the main facial features; eyes, nose and mouth location are detected. From these locations the distance between them are obtained, and will be used for the next process. In this case eyes are the main target to be used, which the blink is detected by applying a variance projection between corners. But, once again, the illumination changes, and the variation of the eye shape as well as the skin colour present challenges in detecting and tracking the essential eye points. In addition to the above, there are also approaches [40, 131, 132] that apply active flow algorithms to measure the eye's state. Active flow detects the changes of the eyelid in order to determine the eye open and closed states. By employing this approach, the speed of blink is efficiently measured. However, this technique is sensitive to the movement of the face.



(a) (b)

Figure 3.6: (a) The face is divided into interest regions in order to obtain salient points within the regions [130], (b) The position of the eyes, mouth, and nose based on anthropometric face models [101]. *(Images are permitted to be published)*

3.3.1.3 **Appearance Based Template method**

In the appearance based approach, the ingenuity of machine learning algorithms is fully utilised [87]. The algorithm trains the eye feature vectors extracted from the dataset of the eye activities in order to teach the machine learning algorithm. The eye features vectors are acquired from the features extractor, which can be denoted as a texture pattern, intensity value or edges. Subsequently, the learned machine algorithm

decides the status of the eye in every frame in a sequence based on the trained parameters. Wang *et al*. [133] extracted the eye's state features using the Gabor wavelet, and classified them by employing a Neural Network classifier. Furthermore, researches in [134-136], have applied an approach, similar to that introduced by Viola and Jones [15] for the face detection, which uses a cascade AdaBoost for training and classification, and Haar-like features for feature extraction. The same classifier has also been applied by Xu *et al*. [137], but they use an LBP for features extraction and the fatigue is determined by utilizing the PERCLOS technique.

The LBP has been proven to be a highly discriminative descriptor of the texture of image [138], thus, Senaratne *et al*. [40] have applied it to the eye's state features, where the fatigue is classified by SVM. The SVM classifier has also been chosen by Sanyuan *et al*. [139], where the eye's state features are extracted using the Histogram of Oriented Gradient (HOG) colour algorithm. Even though Senaratne *et al*. [40] have proved in their research that the classifier approach has better performance than the optical flow, but a lot of datasets of eye images are required for training purposes. Another challenge related to the learning machine approach is that the performance relies on the type of features used, which must be robust and consistent in many different images. The misclassification that commonly occurs during the process definitely degrades the performance of the algorithm.

3.3.2 **Yawning Detection**

Yawning is a symptom indicated visually by the mouth opening. Most research in yawning detection algorithms focuses on the measurement of the mouth opening. The algorithm must be able to differentiate between normal mouth opening and yawning. Similarly to the eye activities methods, the yawn detection can also be categorized into three approaches; features based, appearance based, and model based.

3.3.2.1 **Features Based Yawning Method**

In the features based method, the mouth opening can be measured by applying several features such as colour, edges and texture, which are able to describe the activities of the mouth. Commonly, there are two approaches to measure the mouth opening, mainly by tracking the lips movement and quantify the width of the mouth. Since the mouth is the wider facial component, colour is one of the prominent features able to distinguish the mouth region. Yao *et al*. [140] convert RGB color space to LaB colour space in order to use the 'a' component, which is found as acceptable for segmenting the lips region as shown in Figure 3.7(a). The lips represent the boundary of the mouth opening that is to be measured. YCbCr colour space was chosen by Omidyeganeh [44] for their yawning detection. They claim this colour space is able to indicate the mouth opening area which is the darkest colour region by setting a certain threshold. The yawning is detected by computing the ratio of the horizontal and vertical intensity projection of the mouth region. The mouth opening is detected as yawn when this ratio is above a set threshold. Implementation using the colour properties is easy but, to make the algorithm robust, several challenges must be addressed. The lips colour is different for everyone, and due to the different lighting conditions it is necessary to ensure that the threshold value is adapted to the changes.

Edges are another mouth features are able to represent the shape of the mouth opening. Alioua *et al*. [111] extracted the edges from the differences between the lips regions and the darkest region of the mouth (as shown in Figure 3.7(b)). The width of the mouth is measured in consecutive frames, and yawn is detected when the mouth is continuously opening widely more a set number of times. When using the edges to determine yawn, the challenge is that the changes of the edges are proportional to the illumination changes, which make it difficult to set the parameter of an edges detector. In order to locate accurately the mouth boundary, the active contour algorithm can be used as implemented in [141]. However, the computational cost of this approach is very high. The mouth corners have also been employed by Xiao *et al*. [142] to detect yawn. The mouth corners can be reference points in order to track the mouth region. Yawning

is classified based on the change of the texture in the corner of the mouth while the mouth is opened widely. Moreover, as in the case implemented for measuring eye activities, Jiménez-Pinto *et al*. [129] have detected the yawn based upon the silent point in the mouth region (as shown in Figure 3.7(a)). These salient points are used to describe the motion of the lips movement, and yawn is determined by examining the motion in the mouth region.

(a) (b)

Figure 3.7: (a) Lip segmentation using LaB colour space to highlight lips region [9], (b) Edges mouth detection method introduced in [111] to detect the boundary of uth. *(Images are permitted to be published)*

3.3.2.2 **Appearance Based Yawn Method**

The appearance based method, as reviewed in the eye activity approach, the statistical learning machine algorithm is applied and a distinctive feature needs to be extracted in order to train the algorithm. Lirong *et al*. [143] combined Haar-like features and variance value in order to train and classify the lips using SVM. With the additional integration of a Kalman filter their algorithm is also able to track the lip region. Lingling *et al*. [144] also applied the Haar-like features to detect the mouth region. The yawn is classified using colour properties to measure the breadth of the mouth opening.

Furthermore, Gabor wavelet features that represent texture was been used by Xiao *et al*. [142] for evaluating the degree of mouth opening. The texture features are extracted from the corners region of the mouth. The texture of the corner is different for mouth opening and closing. The yawn is then classified using Linear Discrimination

Analysis (LDA). Yang *et al*. [145] employed a back propagation neural network for yawn classification in three states: mouth closed, normal open, and wide open. The mouth region features are extracted based on the RGB colour properties. For this approach, as implemented in face detection algorithm and eye activities measurement technique, a huge numbers of datasets for images is required in order to obtain exceptional results.

3.3.2.3 **Model Based**

In the model based method it is necessary that the mouth or lips are modelled first. An Active Shape Model (ASM) for instance, requires a set of training images and, in every image, the important points that represent the structure of the mouth need to be marked. Then, the marked set of images is used for training using the statistical algorithm to establish the shape model as shown in Figure 3.8(a). Garcia *et al*. [146] and Anumas *et al*.[147] train the structure for the points of lips which are manually annotated (as shown in Figure 3.8(b)) using the ASM, and the yawn is measured based on the lips opening. Hachisuka *et al*. [148] employ Active Appearance Model (AAM), which annotates each essential facial point. For yawning detection three points are defined to represent the opening of the mouth. In this model based approach a lot of images required since every person has different facial structure.

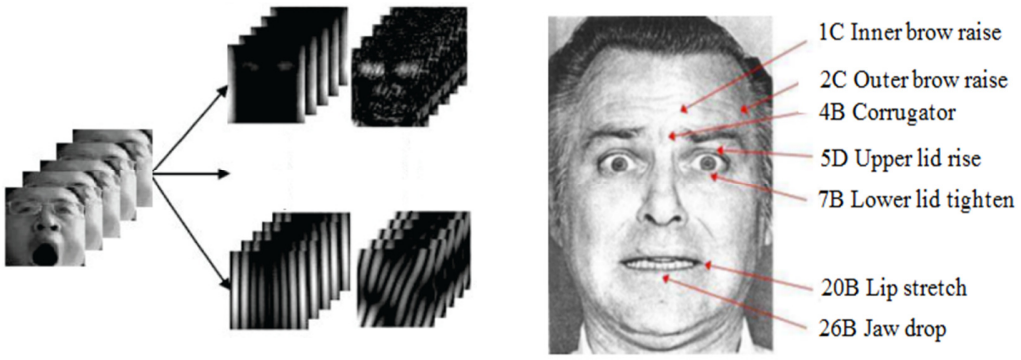
(a) (b)

Figure 3.8: (a) A marked mouth structure for Active Shape Model (ASM)[146], (b) Structure points of the lips are manually annotated [147].

3.3.3 Others Facial Component

Instead of the two major facial components, namely the eye activities and the yawning detection, which have been widely used for quantification of signs of fatigue, several researches use other facial features. Lu *et al*. [118] employed the whole facial features in order to recognize fatigue. Gabor wavelet features are also used to extract temporal features, which is a sequence of dynamic facial images that represent the signs of fatigue. Figure 3.9(a) shows the Gabor wavelet features that extracted from sequences of frames of yawning face. Several combinations of classifiers have been tested, such as Principle Component Analysis (PCA), Linear Discriminant Analysis (LDA), Hidden Markov Model (HMM), and AdaBoost for detecting fatigue. The AdaBoost classifier has produced the highest detection rates compared to the other classifiers. By using the same approach, Yang *et al*. [120] employed the Regional Local Binary Pattern (RLBP) to extract facial fatigue features, also using SVM as classifier.

In addition, the whole facial features have been used to measure fatigue by using The Facial Action Coding System (FACS). FAC is a prominent approach for facial expression recognition as outlined by Vural *et al*. [119]. FACS is widely used in facial expressions and has been adopted by Ekman and Friesen [149] to extract features of face. FACS has 46 Action Unit (AU) components that describe the facial muscles (as shown in Figure 3.9(b)). In Vural*et al*. [119] 45 AU are used, and an Adaboost classifier is applied to detect fatigue. Another noticeable sign of fatigue is a head nod. Usually this action can be detected through a head movement recognition process. However, by monitoring the movement of facial feature components, the nodding can be predicted as implemented in [12]. Xie *et al*. identify the centroid point of the face based on a structure of facial components, and analyse the movement of this point to determine nodding.



(a) (b)

Figure 3.9: (a) Dynamic facial images with Gabor wavelet features from sequences of frame for recognising fatigue [118], (b) Example of facial action decomposition from the Facial Action Coding System [119]. *(Images are permitted to be published)*

3.4 Facial Fatigue Recognition

Facial fatigue recognition is the last operation to be carried out in the facial fatigue detection system after completion of the face acquisition and of the facial fatigue features extraction operations. In this operation, the fatigue to be determined is based upon the selected features extracted from the face by using a specific technique. In general, the facial fatigue recognition can be grouped into a single cue recognition and multi-cue recognition. In the single cue recognition, only one indicator from the face is used for recognition process, and vice versa for multi-cue recognition.

3.4.1 Single Cue Recognition

Single cue recognition is mostly implemented for the eye activities features. The most prominent method that effectively detects fatigue from eye activities features is the PERCLOS (proportion/percentage of time in a minute that the eye is 80% closed) technique used in [12, 13, 18, 40, 62, 124, 126, 150]. PERCLOS, introduced by Dinges and Grace [10], measures the period of eye closure. The eye activity is classified as fatigue when 80% of a certain time period the eyes are closed. Figure 3.10 shows the

principle of the PERCLOS curve over a certain period. PERCLOS is computed as follows:

$$f = \frac{t_3 - t_2}{t_4 - t_1} \times 100\% \tag{3.1}$$

where f is normalized value of PERCLOS, t_1 is time requires for the eye to be 20% closed, while, t_2 is the time requires the eyes are to be 20% completely closed. t_3 and t_4 are the times requires eye open to eye to be 20% opened(after being closed) and to be 80% opened (after being closed), respectively. For normal blink, it take time in between 200 to 400 milliseconds (t_4+t_1). The t_1 is in range 10 to 40 millisecond, t_2 in the range 50 to 150 milliseconds and in the range 100 to 300 milliseconds. For the fatigue condition the time of eye closure is more than 700 milliseconds. The t_1 is less than 20 milliseconds. and t_2 is less than 100 milliseconds [8, 36, 37, 38]. The fatigue is detected when the computed f is 0.8 and more. This technique is reliable and successfully used in monitoring drivers whose full attention is required while driving.

The approach undertaken by Lu *et al* [118] and Yang *et al*[120]employed the whole face structure to examine signs of fatigue. The features of fatigue are extracted from the texture description of the whole face as shown in Figure 3.9(a). Then, a statistical learning machine classifier is used to detect the fatigue from the overall facial reactions. This approach definitely requires a huge number of face images, and it is difficult to classify fatigue signs from the tiny changes such as eye closure that implemented in the PERCLOS technique.

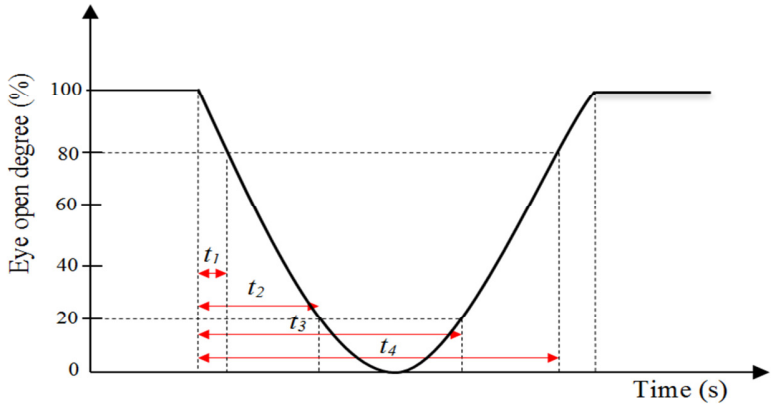


Figure 3.10: Principle of PERCLOS curve over a certain period. In normal condition t_1 is in range 10 to 40 millisecond, t_2 in the range 50 to 150 milliseconds, in the range 100 to 300 milliseconds. For fatigue condition, t_1+t_2 is more than 700 milliseconds, whilst, t_1 and t_2 less than 20 and 100 milliseconds respectively. [8, 36, 37, 38].

3.4.2 **Multiple Cues**

The combination of several facial features is an effective approach in order to obtain a more accurate assessment of fatigue signs. As discussed in section 3.3, there are two dominant features that successfully represent signs of fatigue: eye activities and yawning. Nevertheless, there are also others facial component features that can be applied as reported in [119, 151]. In a common operation, the fatigue recognition algorithm measures each of these facial fatigue features separately. Then all the features are combined for classification by using a state-of-art learning machine algorithm.

Two fatigue facial features that are often implemented in combination are the reliable eye activities measurement PERCLOS and the yawning [101, 129, 152]. In addition, Garcia *et al.* [146] have added another feature of eye activities called Average Eye Closure Speed (AECS), and combined it with PERCLOS and yawning to recognize fatigue. AECS is a technique which measures the average time of the eye closure speed, which is the time between the eye being fully open and eye being fully closed. The study carried in [153] shows that the AECS has a significant variation between persons in tired and normal state. The PERCLOS and AECS techniques are central to these fatigue

detection systems which emphasize lapse of concentration by the driver of the vehicle. There are also other features of eye activities which have been applied and integrated with yawn detection such as the one shown by Azim *et al*. [128]. They alternatively detected eye closure and half eye open separately. Subsequently these eye features are combined with yawning detection results and applied a SVM classifier to detect fatigue.

Besides the two essential components of facial fatigue features, the essential facial muscle features can also be utilised as implemented in facial expression recognition algorithms. FACS is one of the techniques for measuring the facial muscle movement. It possesses 46 action units and 45 of them have been selected initially by Vural *et al*. [119] for pre-fatigue detection experiments. Based on these experiments, only the 16 facial actions are selected as indicative for high prediction rates of fatigue signs. These 16 facial actions are then used for training and classification using Multinomial Ridge Regression (MRR). In research carried out by Qiang *et al*. [151] 21 features were used in their fatigue monitoring system. However, only five features are from the face, namely PERCLOS, AECS, eye gaze features, and yawn. Due the large number of fatigue features, Qiang *et al*. have used a Bayesian Network (BN) model to cope with the complexity of input features for determining fatigue symptoms. The technique that uses a lot of component of facial features is helpful for fatigue classification, but at the same time, it is very sensitive to common spontaneous reactions produced by the face.

3.5 Conclusion

This chapter provided a review of facial fatigue analysis which has three main components of operation. Firstly the face acquisition system is discussed which is divided into the face detection algorithm and the facial component detection algorithm. In the face detection algorithm four methods are reviewed, knowledge based, features based, template based and appearance based methods. Likewise, for the facial component detection, the shape based, features based, and appearance based methods were discussed. From these methods, the appearance based method, which employs

statistical and machine learning algorithms, is the most reliable and it is preferred by researchers.

After the face is detected and the components are located, the next operation is to extract the features from the facial component features that most prominently indicate the signs of fatigue, as discussed in Chapter 2. There are two components which are dominant in representing the signs of fatigue the eyes activities and the yawn detection. Similarly to the face acquisition operation, the measurement techniques for the activity are categorized into model based, features based and appearance based methods. In yawning detection algorithms, all the techniques applied are based on the degree of mouth opening. However there is no technique that can take into account yawn when the mouth is covered while yawning (a very common human gesture). Therefore, there is a technique should be developed to cope with this deficiency.

The last operation in a facial fatigue analysis system is the fatigue recognition algorithm. In general fatigue is detected based on single cue or multi-cue utilisation. In single cue fatigue recognition, the PERCLOS is the technique that has been mostly used. In multi-cue recognition, eye activities cues are commonly combined with yawning detection. Besides these two dominant fatigue features, there are also other facial features components that researchers have applied to measure the signs of fatigue.

4. Image Databases

4.1 Introduction

Databases of images and videos are required for training, testing and evaluating the performance of the developed algorithms. As discussed in chapter 3, the fatigue facial analysis system is divided into three sequential operations, and each of the operation uses the database repository. Firstly, for the face acquisition operations, several databases of face images are used in order to examine the performance of the new developed face detection algorithm. The next operation is a facial features extraction operation, which is a novel algorithm to measure the eye activities introduced. There are two categories of databases used here, the database of iris images and video footage from the eye blinks database. The iris images database is utilised for testing the developed iris localisation algorithm, while, the blink videos footage database is used for assessing the performance of the eye activity measurement algorithm.

One of the main contributions of this PhD research work is a development of facial fatigue video footage database. This database is named as Strathclyde Facial Fatigue (SFF) database. This database is developed which involved the Centre for excellence in Signal and Image Processing (CeSIP), University of Strathclyde, the Psychology Knowledge Exchange & Enterprise Unit (PsyKE), University of Strathclyde, and the Glasgow Sleep Centre, University of Glasgow. The SFF database contains results from a series of experiments that involved sleep deprived participants. This database has been fully used in the development of the fatigue recognition operation as well as in evaluating the facial features extraction operation. This chapter discusses the database with regards to the technical aspect and evaluation format of the database. For the novel facial fatigue video footage SFF database, the technical aspect of experiments is described and experimental tasks that were carried out are discussed.

4.2 Database for Face Acquisition

In face acquisition operations, face images databases are important, not only for testing and evaluating the developed algorithms, but also for using it for training purposes as, for example, implemented in [15, 16, 88-94]. For instance, Viola and Jones have used 5000 faces for their cascade classifier training [15, 94], with the face images cropped and scaled to 24 by 24 pixels as shown in Figure 4.1. Lee et al. [97] utilised 40,000 face images that are scaled to 19 by 19 pixels for training their face detector.

Figure 4.1 Example of face images used for training the features of face [15, 154]. *(Images are permitted to be published)*

There are several face image databases publicly available to use and these databases have been utilised by many researches - a brief description of some of the databases as shown in Table 4.1. The databases number (1) to (4) in Table 4.1 are among the most popular and they are used for the training of learning machine algorithms. Database (6) is utilised for testing the developed face detection algorithm in [71, 72]. More face image databases can be found in [155]. Figure 4.2 shows some of the images that contain faces that are used for testing.

Table 4.1 Some of the prominent face image databases

	Database Name	Features	Conditions
1	AR Face Database [156]	- Colour images - Size: 576 x 768 - 3288 images	All frontal views of neutral expression
2	CMU Pose, Illumination, and Expression (PIE) Database [157]	- Colour images - Size: 640 x 486 - 41,368 images	Variety of poses, illumination and expression
3	MIT Database [158]	- Greyscale images - Size: 120 x 128 - 433 images	Three different scale, lighting, and head tilt.
4	The Facial Recognition Technology (FERET) database [159]	- Colour images - Size: 256 x 384 - 14,051 images	Deference of poses, illumination, expression and time
5	CAS-PEAL Database [160]	- Colour images - Size: 360 x 480 - 30,900 images	Diversity of lighting, poses, distance, background, time and expression.
6	Combined MIT/CMU Test Set [161, 162]	- Greyscale/colour - Multi-size - 180 images	Multi sizes, pose, and distance.

Figure 4.2 Example of face images used for evaluating the performance of face detection algorithm [159-162]. *(Images are permitted to be published)*

4.3 Database for Eye Activities Measurement

In operations of fatigue features extraction the eye activity is one of the facial fatigue features to be measured. This thesis introduces a novel algorithm to measure the eye activity which is specifically discussed in the next chapter. The developed algorithm tracks and measures the iris size in order to quantify the state of the eye. Therefore in this algorithm, two types of databases are used; iris images databases and video footage of eye blinks database. The eye activity measurement algorithm is included in the operation of fatigue features extraction.

4.3.1 Iris Images Database

The iris image database is used for testing and evaluating the developed iris recognition system for biometric identification application. This is due to the fact that the structure of iris can represent a unique identity of a person. One of the prominent methods of iris recognition is introduced by Daugman [163, 164]. In iris recognition systems, the iris needs first to be located, and Daugman introduced an Integrodifferential operator in order to precisely detect the iris boundary. The iris region is extracted then and the structure of the iris pattern is encoded using the mathematical and statistical algorithms that represent the identification of the person.

In the developed eye measurement activity algorithms, which is discussed in next chapter, the iris databases are used only for testing the developed iris localisation algorithm. This is in order to ensure that the developed algorithm is able to work robustly. There are several iris databases which are publicly available as tabulated in Table 4.2. These databases contain a variety of features and image qualities as well as sizes and poses of irises, which are mostly applied to iris recognition research [165, 166]. Figure 4.3 shows same examples of iris images from various of databases.

Table 4.2 The iris databases

	Database Name	Numbers	Conditions/Features
1	CASIA 1,2,3 &43 [167]	- CASIA1: 756 iris images - CASIA2: 2,400 iris images - CASIA3: 22,034 iris images - CASIA4: 54,601 iris images	Variation of illumination condition for indoor and outdoor, and distance images. There are NIR images.
2	UBIRIS.v1&v2[165]	- UBIRIS.v1: 1,877 iris images - UBIRIS.v2: 11,102 iris images	Images captures in variable wavelength
3	MMU1&2 [168]	- MMU1: 450 iris images - MMU2: 995 iris images	Subject in variety of ages and nationality
4	ICE2005&2006	- ICE2005: 2953 iris images - ICE2006: 60,000 iris images	Broader range of quality
5	U of Bath [169]	- 16,000 iris images	High quality images

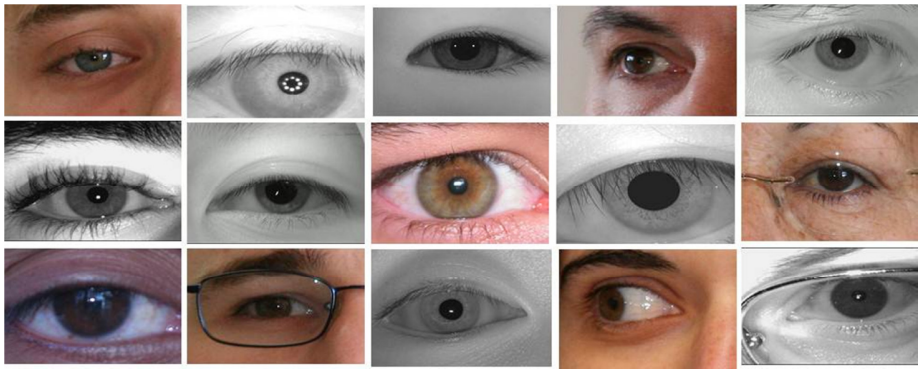


Figure 4.3 Example of iris images using for evaluating the performance of iris localisation algorithm [168-170]. *(Images are permitted to be published)*

4.3.2 Database for Blink Detection

In developing algorithms for measurement of eye activity, the detection of eye blink is very important as well as the capability to quantify the period of a blink. Therefore in this research, the Zhejiang University (ZJU) eye blink database [135] is used for evaluating the developed algorithm performance. The ZJU eye blink database contains 80 video clips and the numbers of blinks in the clips vary between one to six blinks in each case. There are a total of 255 blinks in this database. The video clips are recorded using a generic Pro5000 Logitech web camera where the frame size is 320 x 240 in 30fps. This database has also been used by researches in [131, 135, 171] for assessing

their developed algorithm performance. For further analysis, testing, and evaluating our developed algorithm, the Strathclyde Facial Fatigue Database that will be discussed in next section has been used. Figure 4.4 shows some examples of video images from the ZJU database.

Figure 4.4 Example video images of ZJU database that uses for evaluating performance of IASMS algorithm [135]. *(Images are permitted to be published)*

4.4 Strathclyde Facial Fatigue (SFF) Database

The Strathclyde Facial Fatigue database was developed in order to obtain genuine facial signs of fatigue. To author's knowledge there is currently no publicly available specific database of fatigue signs that can be used as reference. Therefore, our aim was to establish the facial fatigue database that can be utilised by any researcher. Most of the fatigue related researchers have used their own datasets for assessing their developed system. For example, Vural et al. [119] developed their fatigue database by playing a simulation video of driving. The participants had to drive over three hours and their faces were recorded during that time. Fan et. al. [172] recruited forty participants and recorded their faces for several hours in order to obtain the signs of fatigue from their faces.

Some of the researchers only test their algorithms based on video footage of persons who pretend to experience fatigue symptoms. For example, researchers in [44, 111, 141] detected yawn based on the width of the mouth opening and the algorithm was not tested on subject experiencing natural yawning. Hence, the SFF database provides genuine medium of facial fatigue signs from sleep deprivation experiments that involve twenty participants. As discussed in Chapter 2, sleep deprivation is one of the main

factors that contribute to fatigue symptoms and have significant importance to applications in the aviation, and shipping industries where crews suffer fatigue due to sleep loss and jetlag problems. Details of contribution and characteristics of the SFF are given in Appendix B

This SFF database consists of twenty participants with four experimental sessions and five tasks per session. The summary of the SFF video footage database is described in Table 4.3. Figure 4.5 shows examples of images from the video footage of the SFF database. The carried out tasks of sleep deprivation experiment are discussed in Appendix B.

Table 4.3 SFF database.

Participant	Session	Experiment Tasks	Database	Format
- 20 participants. - 10 male & female. - Age: 20 to 40 years.	- 4 session of experiment for each participant. - 0, 3, 5 and hour sleep deprivation	- 5 cognitive task for each session (face activities video recorded): 1. SART [187]. 2. PVT [188]. 3. Emotional face photo discrete categorisation task [188]. 4. Dynamic video social gesture dimensional categorisation task. 5. Boredom task	Every participant: - 5 video footages each session. - Total video footages are 20. - The duration video footage depends on the period task carried out. Total video footages in this database (20X20) are 80 videos.	Video footage: - avi format. - Xvid widescreen size - 25 frames per second.

Figure 4.5 Example images from video footage SFF *(Images are permitted to be published)*

4.5 Conclusion

This chapter presented and discussed the database of images and videos which are used in this research work for training, testing and evaluation of the developed algorithms. There are three operations in this facial fatigue detection system that uses the databases: face acquisition, eye activity measurement, and fatigue recognition. In the face acquisition operation, the face images databases are used to assess the performance of the developed face detection algorithm. For the iris localisation operation, iris images are used for testing the proposed algorithm. In fatigue recognition, a new video fatigue facial database that named as Strathclyde Facial Fatigue (SFF) is used. This database was developed through collaborative work between our research group CeSIP, PsyKE, and the Glasgow Sleep Centre.

This database is generated from the sleep deprivation experiments which involved twenty participants who, in four different sessions, were required to be sleep deprived for 0, 3, 5 and 8 hours. For each experimental session, participants were required to carry out five cognitive tasks that have sensitivity links to sleep loss. This database will be employed for training, testing and evaluation eye activity measurements,

yawning analysis and fatigue recognition algorithms the details of which are discussed in chapters 5, 6 and 7 respectively.

5. Face Acquisition and Eyes Activities Measurement

5.1 Introduction

As discussed in Chapter 2 and 3, eye activities represents a prominent feature in indicating the signs of fatigue. This chapter discusses the developed techniques and algorithms to measure eyes' activities. The general block diagram of the developed eyes' activities measurement system is shown in Figure 5.1. This system consists of four algorithms which operate sequentially. The two initial algorithms are concerned with face acquisition operations; face detection and facial component detection which are discussed in section 5.2. The face detection localises the region of the face and then the facial component detection locates the eyes and the mouth regions. In section 5.3 the next algorithm, iris localisation is discussed. The last algorithm, Interdependence and Adaptive Scale Means Shift (IASMS), measures the size of the iris as well as tracking the non-rigid iris movement in order to quantify the eye's activity. In this chapter two novel development are introduced, the face detection algorithm and the IASMS algorithm.

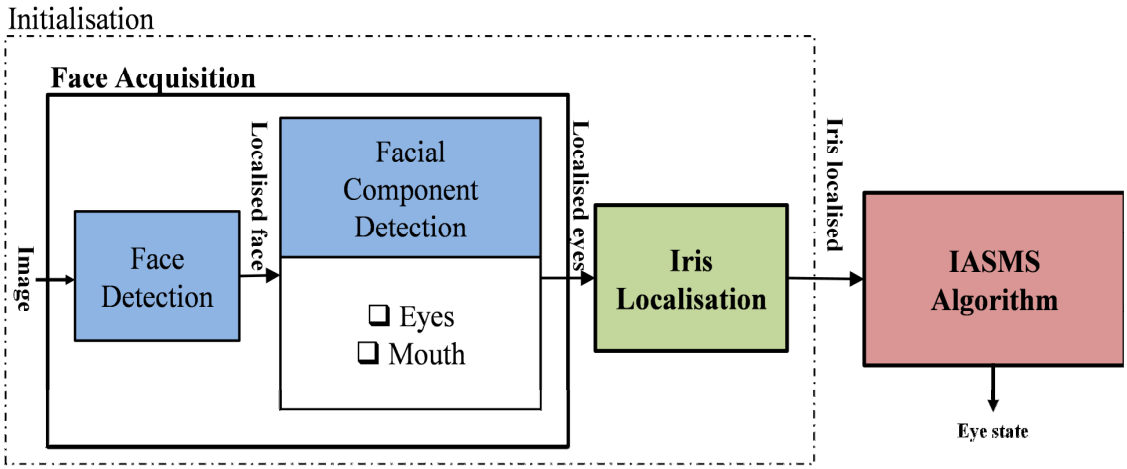


Figure 5.1 General Block diagram of eye activity measurement system which consists of face acquisition operation, iris localization operation and IASMS algorithm. The input of the system is image and the out is eye state

5.2 Face Acquisition

In face acquisition operation there are two algorithms that operate sequentially, face detection and facial component detection. The face region is firstly detected, and then the detected region is used for facial component extraction. The complexity of the facial component detection is less in compared to the face detection algorithm since searching region is already defined. There are only two facial components to be detected in this algorithm, eyes and mouth, this will be used to extract the signs of fatigue in this research work.

5.2.1 Face Detection

A new approach in face detection algorithm is introduced in this thesis. This algorithm applies a combination of colour skin segmentation, connected component of binary image, and a machine learning algorithm. Figure 5.2 shows the process flow of the algorithm where the colour segmentation is first implemented. Then, the connected components of the segmented colour skin region are used in order to form the bounding

rectangle. The size and shape of the bounding rectangle is then examined so as to select the rectangles most likely to contain a face. In the last process, each region of interest is examined to determine whether it contains a face or not by employing a learning machine algorithm.

5.2.1.1 **Colour Skin Segmentation**

As discussed in section 3.2.1.2, colour skin among the prominent features has been used to detect the face. Skin colour segmentation is a fast processing and highly robust to geometric variations of head pose and orientation of face[80]. In complex background scenes for example, a face detection algorithm needs to be more complex. However, using the skin segmentation the complexity can be reduced by eliminating the non-skin region during the initial process. In this algorithm three colour spaces are employed and the skin colour is detected based on classifying each pixel independently.

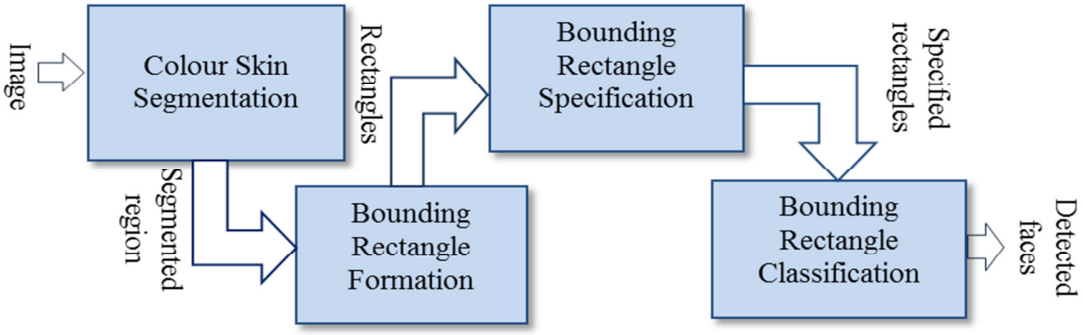


Figure 5.2 The flow of face detection algorithm with four sequential operations

These three colour spaces are widely used in skin detection research [80]. RGB is a basic colour space and the one that has mostly been applied in skin detection research, however it is not very robust to changes in lighting conditions. Therefore Kovac *et al.* [173] used gray world method as an adaptation technique in order to correct the images before applying skin detection. Another colour component which has given acceptable results is chrominance component [80] [174]. This component is computed by

subtracting the luminance component from B and R values of YCbCr color space. In order to adapt to the changing of lighting Yu-Ting *et al* [174] have modulated the range of YCbCr skin colour distribution. The last space colour used in this algorithm is HSV which represents colour in terms of colour-depth, colour-purity and colour brightness. It has rapidly gained ground in recent years and it has shown acceptable results as commented by Gonzalez *et al* [80] [175]. This combination colour space with adaptation technique is used for obtaining the best result to overcome sensitivity of image to illumination changes, ethnicity colour skin and the different cameras characteristics. The rule of skin colour segmentation of this algorithm is as in (5.1)

$$
\begin{aligned}
&If\ \ (r > 95\ \&\ g > 40\ \&\ 20)\ \&\\
&\quad ((\max(r,g,b) - \min(r,g,b)) > 15)\ \&\\
&\quad ((|r-g| > 15)\ \&\ (r-g)\ \&\ (r > b))\ \&\\
&\quad ((140 < c_b < 195)\ \&\ (140 < c_r < 165))\ \&\\
&\quad (0.01 < hue < 0.1)\\
&then\ \{region\ is\ a\ skin\ \ pixel\}
\end{aligned}
\tag{5.1}
$$

where *r*, g, and *b* represent the red, green and blue colour component respectively. C_b and C_r are the blue-difference and red-difference chroma components, and hue is another component from HSV (Hue/Saturation/Value) colour format.

5.2.1.2 **Rectangle Bounding Formation**

Result of the skin colour segmentation is a binary image as shown in Figure 5.3 that represents the skin and non-skin region. This binary image consists the connected components that uniquely labeled based on the region to which the pixels have same value are connected together. From the connected components the skin region bounding rectangle is formed. The regions which are not bounded by a rectangle are eliminated from the next process in the algorithm.

(a) (b)

Figure 5.3: Segmented skin colour region process result. (a) The input image,(b) the segmented region of skin detection. *(Images are permitted to be published)*

A bounding rectangle is formed based on eight extremal points; topmost left, topmost right, rightmost top, rightmost bottom, bottommost right, bottommost left, leftmost bottom and leftmost top as shown in Figure 5.4. For bounding rectangle formation there are four important points that must be defined. These are the maximum points (y_{max},x_{max}) and minimum points (y_{min},x_{min}) and are obtained according to the following rules:

$$\begin{aligned} y_{\min} &= \min \{y \mid (y, x) \in R\} \\ y_{\max} &= \max \{y \mid (y, x) \in R\} \\ x_{\min} &= \min \{x \mid (y, x) \in R\} \\ x_{\max} &= \max \{x \mid (y, x) \in R\} \end{aligned} \tag{5.2}$$

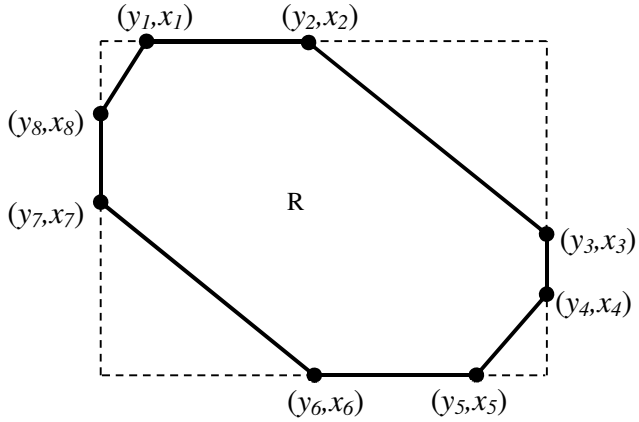


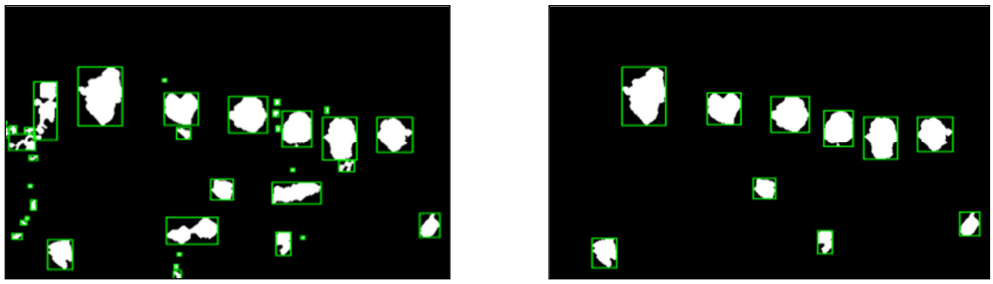
Figure 5.4: Eight extremal points in segmented region of binary image

5.2.1.3 **Bounding Rectangle Specification**

The results of all skin colour regions upon colour segmentation are bounded by rectangle $B(x, y)$ as shown in Figure 5.5(a). However there are some small regions which are segmented as skin colour regions but, obviously, cannot be presented as face regions. These are noise regions which need to be discarded. The following is an acceptable bounding rectangle $B(x, y)$ between the small region of pixels sm and large region of pixels bg :

$$B(x, y) = \begin{cases} 'retained' & if \ \ sm \le B(x, y) < bg \\ 'discarded' & otherwise \end{cases} \tag{5.3}$$

The value of sm is defined based on the minimum pixel that can represent a face. For instance, Viola *et al* [15] used a size 24x24 face image for extracting the Haar-like features to train the cascade classifier. In this algorithm, since the Viola Jane technique is applied in this work, the sm value is 24. The value of bg represents the largest possible size of a bounding rectangle, and it is always the smallest value of either the height or width of the image.



(a) (b)

Figure 5.5: (a) segmentation skin colour results are bounded by rectangle *B(x,y)*. (b) Output of bounding rectangle specification process.

Another aspect that needs to be checked on a bounding rectangle is its shape. The shape defines whether the rectangle bounds a face or a non-face object and it is measured by the ratio of *width* to *height* of the rectangle defined as follows:

$$0.83 \leq \frac{width}{height} \leq 1.27 \tag{5.4}$$

The ratio range is obtained from experimentation based on 98 images that contain 561 faces. Figure 5.6 shows examples of these images after skin colour segmentation and bounding rectangle formation processes were applied on them. In these images, the rectangles bound all the regions that were segmented as skin regions. However in this experiment, only the rectangles bounding ratio of face were measured. From the accumulated measurement values, the ratio range was concluded as shown in equation (5.4).

Figure 5.6 Example images that show the skin region bounding rectangles. *(Images are permitted to be published)*

5.2.1.4 **Bounding Rectangle Classification**

The final process is the bounding rectangle classification, where a classifier is used to evaluate whether the remaining rectangular boxes (as shown in Figure 5.7(b))

contain a face or not. In this research the Viola Jones [15] classifier is employed which fairly reliable and can opetrate in real-time [72]. This classifier, as discussed in section 3.2.1.4, applies the AdaBoost learning algorithms to train the cascade classifier. This technique extracts the features vectors, using a Haar-like features technique, from integral images [15].

(a) (b)

Figure 5.7: (a) The remain bounding rectangles after bounding rectangle specification operation, (b) the result of classification upon bounding rectangle which only faces detected.

5.2.2 Facial Component Detection

When the face is localised, the detected face region is further processed to find two important facial components which are: eyes and mouth. In this research the Viola-Jones technique [15] is used for eyes and mouth detection in the bounding rectangle classification as indicated in the previous section. Based on the technique discussed in section 3.2.1.4 and detailed in Appendix A, the images of the eyes and mouth are required for training purposes. In our algorithm the trained Viola Jones classification model of MATLAB R2012 is used. For eyes detection, the right and left eyes model is used with size of images of 12x18 pixels for training the classifier. For mouth detection, the model uses the size 15x25 pixels mouth images for training the classifier. The eyes and mouth images use Haar-like features and trained based on the Viola Jones technique as implemented in [136]. Figure 5.8 shows the region of the face, eyes and mouth detected using the Viola Jones technique.

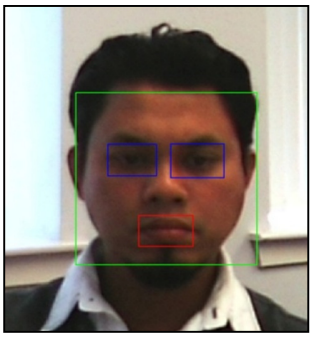


Figure 5.8 Face, eyes and mouth detected region. . *(Images are permitted to be published)*

5.3 Iris Localisation

In this research, the state of the eyes is determined based on the measurement of the iris size. Before the iris detection algorithm is carried out, the eyes are examined whether are closed or open. This stage is important to ensure the eyes are open initially to obtain the actual size of the iris. A new technique is introduced to determine the status of the eyes. This is based on an adaptive threshold value and a connected component binary image. Firstly, the input image is enhanced by transforming the histogram of the image to spread the level of colour values evenly but without changing the shape of the input histogram. A new histogram of image *g'(x,y)* is computed as follows [176]:

$$g'(x,y) = \frac{Gl'_{\max} - Gl'_{\min}}{Gl'_{\max} - Gl'_{\min}} [g(x,y) - Gl_{\min}] + Gl'_{\min} \tag{5.5}$$

where *g(x,y)* is the original histogram of image. This linear transform illustrated in Figure 5.9 stretches the histogram of the input image into the desired output intensity level of the histogram between $Gl'_{\min}$ and $Gl'_{\max}$, where $Gl_{\min}$ and $Gl_{\max}$ are the minimum and maximum input image intensities respectively. From this enhanced histogram, the cumulative histogram H_i is calculated (5.6) which generates the total number of index pixels in the histogram bins between 0 and 255 Figure 5.10(b)).

$$H_i = \sum_{j=1}^{i} h_j \tag{5.6}$$

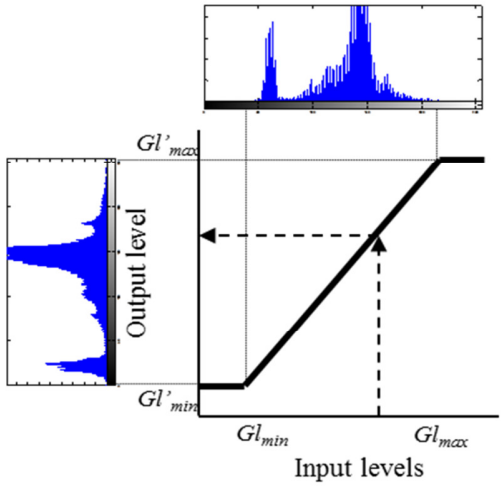
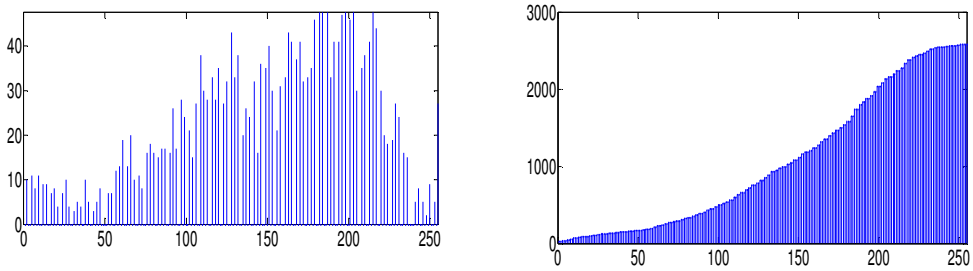


Figure 5.9 A linear transform that remaps intensity level minimum $Gl_{\min}$ and intensity level maximum $Gl_{\max}$ of input image into a new intensity level between $Gl'_{\min}$ (minimum) and $Gl'_{\max}$ (maximum).



(a) (b)

Figure 5.10: (a) Enhanced image histogram after linear transformation is implemented. (b) Cumulative histogram of the enhanced image as computed using equation (5.6).

The specific region that is extracted from the profile projection contains the eye and the eyebrow. The darkest area in this region represents the iris, the eyebrow and, also, the eyelashes. It has been experimentally observed that, on average, the darkest area constitutes approximately 10% of the entire obtained region of interest. Based on this percentage, the darkest area in the histogram is situated between 0 and X intensity levels. The intensity value X can thus be obtained by determining the area of the first 10% of the total number of pixels, contained in the enhanced cumulative histogram image (Figure 5.10(b)) and reading the corresponding intensity value from the x-axis.

$$0 < X < 255 \tag{5.7}$$

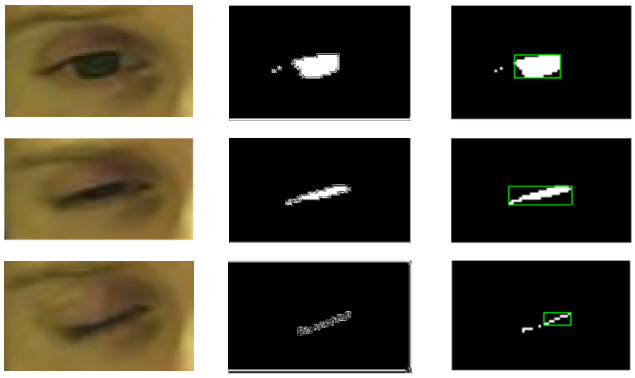
Then, the adaptive threshold value T_a is calculated as follows:

$$T_a = \frac{X - P_{\min}}{P_{\max} - P_{\min}} \tag{5.8}$$

where $P_{\max}$ and $P_{\min}$ are the maximum and minimum pixel value respectively.

The input eye region is then converted to a binary image using an adaptive threshold as shown in Figure 5.11(a) and (b). The connected components of the binary image are utilized to form a bounding box in the region of the interest as shown in Figure 5.11(c). The characteristic of the region of interest in this bounding box is examined in terms of the aspect ratio of the bounding box and the area of region of interest within it. Firstly, the aspect ratio of the bounding box shape sr of height (hbb) and width (wbb) is calculated (5.9). The experimental results have indicated that if this ratio is less than 0.4 the eye can be classified as closed. Otherwise the region of interest in the bounding box needs to be examined further.

$$sr = \begin{cases} 0 & \frac{hbb}{wbb} < 0.4 \\ 1 & otherwise \end{cases} \tag{5.9}$$



(a) (b) (c)

Figure 5.11: (a) Three state of the eye: eye fully opened, half closed, and fully closed. (b) Binary image when applied adaptive threshold on eye region. (c) Bounding box formed on segmented region for examining the darkest region.

Subsequently, the difference between eye open and closed is examined based on the area of the region of interest within the bounding box. A high percentage in the value of the region of interest, relative to the size of the bounding box, indicates that the eye is open, otherwise the eye is closed. Based on the experiments carried out if the area of region of interest is less than 70% of the bounding box area the eye is classified as closed.

Tracking the irises and quantifying their area are important tasks in the next algorithm. To ensure an accurate location of the irises and the best possible performance for the IASMS process a Daugman's iris localisation method [163, 164] is utilised. Daugman [163] introduced an Integro-Differential Operator to detect the boundary of the iris and the pupil. This technique performs circular edge detection by convolving the image with a Gaussian operator in order to smooth the image so that the edge information is highlighted. The operator equation (5.10) has been implemented in discrete form through the Matlab function *imfindcircles* .

$$\max(r, x_0, y_o)\left|G_\sigma(r) * \frac{\partial}{\partial r}\oint_{r,x_0,y_0}\frac{I(x,y)}{2\pi r}ds\right| \tag{5.10}$$

where x_0 and y_0 are the coordinates of the iris centre, and r is the iris radius. $I(x, y)$ is the input iris image smoothed by the then Gaussian function $G_\sigma(r)$. The smoothed image is scanned for searching a circular edge that has a maximum gradient change. This maximum gradient represents the boundary of the iris which is used, in turn, to find the centre of the iris. This final information is then delivered to the IASMS algorithm.

5.4 Interdependence And Adaptive Scale Mean Shift (IASMS) Algorithm

The Interdependence and Adaptive Scale Mean Shift (IASMS) algorithm is designed specifically for tracking and quantifying the iris size in order to analyse the state of the eyes. This algorithm is a hybrid of a conventional Mean Shift tracking algorithm, and an adaptive scale technique, which utilises moment features and interdependence tracking scheme based on eye's position.

5.4.1 Mean Shift Tracking Algorithm

The Mean Shift algorithm, introduced by Comaniciu *et al* [177], is a popular method for tracking due to its simplicity and efficiency. The Mean Shift algorithm tracks the object based on the probability of colour of the object. In this research, the Mean Shift tracking algorithm is employed to track the non-rigid iris movement wherein the iris is tracked based on the probability of the colour histogram of the iris $\hat{q} = \{\hat{q}_u\}_{u=1\ldots m}$, where *m* is the number of histogram bins. The probability colour histogram of the iris as a target mode $\hat{q}_u$ is computed as follows:

$$\hat{q}_u = c\sum_{i=1}^{n} k\left(\|x_i\|^2\right)\delta\left[b(x_i) - u\right] \tag{5.11}$$

where $\|x_i\|$ is an absolute value of normalised pixel which location from 1 to n with the target iris centred at 0. δ is a Kronecker delta function, while the function $b(x_i)$ determines the bin for pixel x_i. Mean shift tracking uses the Epanechnikov kernel [178] function k that represents the weight function of each pixel. C is a normalisation constant which is defined as follows

$$C = \frac{1}{\sum_{i=1}^{n} k(\|x_i\|^2)} \tag{5.12}$$

In order to track the iris, the probability of colour histogram of the target iris that called as a target candidate $\hat{p}_u(y)$ is computed as follows:

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n_k} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)\delta\left[b(x_i) - u\right] \tag{5.13}$$

where y is the centre position of the current frame and h is the radius of weighting kernel. The normalisation constant for this target model C_h is computed as follows:

$$C_h = \frac{1}{\sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)} \tag{5.14}$$

The primary task of the Mean Shift tracking algorithm is to find the target location in the current frame. The most probable location of the centre of the iris is obtained by

minimizing the distance between two distributions. The distance between target model and target candidate *d(y)* is defined as:

$$d(y) = \sqrt{1 - \hat{p}(y)} \tag{5.15}$$

where the $\hat{p}(y)$ is an estimation of the Bhattacharyya coefficient which can be computed as:

$$\hat{p}(y) = \sum_{u=1}^{m} \sqrt{\hat{p}_u(y)\hat{q}_u} \tag{5.16}$$

In order to find the acceptable centre point of the targeted iris in the next frame, the value $d(y)$ must be minimised while $\hat{p}(y)$ must be maximised. Comaniciu *et al* [177] derived the following method that can be used to compute a new location targeted iris y_1 as follows:

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i}{\sum_{i=1}^{n_h} w_i} \tag{5.17}$$

where the weight of the pixel w_i is computed as:

$$w_i = \sum_{u=1}^{m} \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}} \delta[b(x_i) - u] \tag{5.18}$$

5.4.2 **Adaptive Scale Scheme**

The adaptive scale scheme is an essential part in this algorithm in order to enable the algorithm to quantify the size of the iris. This scheme allows the iris scale to be quantified at the same time as the eye movement is tracked. The conventional Mean Shift tracking algorithm was designed for static colour distributions meaning that the target search region remains unchanged relative to variations in object size. However, Bradski [179] has extended the Mean Shift algorithm in order to deal with dynamic changes in colour probability distribution, and this technique is the Continuously Adaptive Mean Shift (CAMSHIFT) algorithm. The algorithm exploits the moment features, which represents a global description of the shape to obtain the area, height and width of the object. The area can be measured by finding the zeroth moment M_{00} of the region that is computed as follows:

$$M_{00} = \sum_{x,y} I(x, y) \tag{5.19}$$

When the area of the object is obtained, in this case the iris area, the centroid of the object can be recovered using first-order moment. The first order moment M_{10},M_{01}, and M_{11} are computed as:

$$\begin{aligned} M_{10} &= \sum_{x,y} xI(x, y) \\ M_{01} &= \sum_{x,y} yI(x, y) \\ M_{11} &= \sum_{x,y} xyI(x, y) \end{aligned} \tag{5.20}$$

where $I(x, y)$ is the probability pixel value within the object region in x and y range. The centroid point (x_c, y_c) of the region then is accomplished as:

$$x_c = \frac{M_{10}}{M_{00}}, \quad y_c = \frac{M_{01}}{M_{00}} \tag{5.21}$$

In the Mean Shift tacking algorithm essentially there are two forms of the search tracking region, rectangular and ellipsoidal. The form of the target search region can significantly impact on the performance of the Mean Shift Tracking algorithm. One of the major challenges faced when using the Mean Shift algorithm is the probability that the colour of the search region does not represent the exact probability of the tracked object [180]. Commonly, the object tracking is not precisely allocated in the search region form because of the shape of the object such as a car or a human where the background scene region is included in the search region form and which contributes noise in the probability colour of the search region. However, in this application, an ellipsoidal search region is employed. This search region form is appropriate with eye shape characteristic, thus, the aforementioned problem can be reduced.

In the CAMSHIFT algorithm, the target location of the object is estimated using a weighted image which is determined using the hue-based colour histogram. The hue colour is unstable at low saturations which in certain to below the threshold of pixels are not included in the colour histogram. In this algorithm, by referring to the results from the analysis carried out in [181], the weighted image from equation (5.18) is utilised to quantify the scale of the iris. Therefore, the value of the probability pixel $I(x, y)$ in (5.19) and (5.20) can be changed to the weight of the pixel w_i (5.18)

In order to obtain an accurate size of the iris, the width, height and also orientation must be acquired by computing the second order moment. The second order moment M_{20} and M_{02} are computed as follows:

$$M_{20} = \sum_{x,y} x^2 w(x, y), \quad M_{02} = \sum_{x,y} y^2 w(x, y) \tag{5.22}$$

From the second-order moment, the rotation of ellipse search target region can be described by calculating the second-order centre moment. The second-order centre

moment that able to represent that scale features are μ_{20}, μ_{02} and μ_{11} which computed as follows:

$$\mu_{20} = \frac{M_{20}}{M_{00}} - {x_c}^2, \mu_{02} = \frac{M_{02}}{M_{00}} - {y_c}^2, \mu_{11} = \frac{M_{11}}{M_{00}} - x_c y_c \tag{5.23}$$

The degree of orientation of the ellipse of the search target region can be derived by employing the second-order centre moment. The degree of the target region θ is computed as follows:

$$\theta = \frac{1}{2}\tan^{-1}\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \tag{5.24}$$

Ideally, this algorithm quantifies the eye's state based on the iris area. However, the actual region of the iris is difficult to compute in the searching region of the algorithm. Therefore, the iris area is approximated by the size of an elliptical search area. The ellipse of the search area can be calculated when the width and height of the ellipse are obtained. The semi-major axis of the ellipse a and the semi-minor axis of ellipse b is calculated as follows:

$$a = \sqrt{l_1 A/(\pi l_2)}, \quad b = \sqrt{l_2 A/(\pi l_1)} \tag{5.25}$$

where A is the area of region computed from the zeroth moment (5.19), while l_1 and l_2 that represent the length of region are calculated as follows:

$$l_1 = \sqrt{\frac{(\mu_{20} + \mu_{02}) + \sqrt{4{\mu_{11}}^2 + (\mu_{20} - \mu_{02})}}{2}} \tag{5.26}$$

$$l_2 = \sqrt{\frac{(\mu_{20} + \mu_{02}) - \sqrt{4\mu_{11}{}^2 + (\mu_{20} - \mu_{02})}}{2}} \tag{5.27}$$

5.4.3 Interdependence and Adaptive Scale Mean Shift Tracking

The common mean shift tracking is based on single tracking, which is independent of other objects. However, in this application, there are a pair of eyes and, therefore interrelated. By utilizing information from the pair of eyes, we will incorporate an interdependence technique that aims to be able to track the non-rigid movement of the eyes. There are two primary pieces of information employed in this algorithm, the centre locations of the irises and the distance between them. Initially the eye centres (5.18) are obtained from the pre-processing algorithms which will be discussed in the following section. Then the Mean shift algorithm updates these centre positions for every tracking iteration. The C_r and C_l in (5.28) denote the centre points of the right and left eye respectively.

$$C_r = (x_r, y_r) \qquad C_l = (x_l, y_l) \tag{5.28}$$

From the obtained centre points of the irises, the distance between centre point of eye left and right D_e is calculated as equation (5.29). The D_e is continually updated for every tracking iteration.

$$D_e(p,q) = \sqrt{(x_r - x_l)^2 + (y_r - y_l)^2} \tag{5.29}$$

The common drawback of the Mean Shift tracking algorithm is erroneous colour tracking. When dealing with colour, there is the possibility of multiple objects having similar colours. In this case, in order to avoid any erroneous results the following new approach is introduced. First a specific eye region is defined on the image known as the Focus of Eye Region (FER) with the size and location of the region updated in every

iteration. Then the Mean Shift algorithm is used to track the iris in the FER region of the image. The size of the FER depends on the centres of the irises C_r and C_l and the distance between them D_e . The size of FER for right eye $I_{FER}\ (x_R, y_R)$ and left eye $I_{FER}\ (x_L, y_L)$ are defined as follows:

$$\begin{aligned} I_{FER}\ (x_R, y_R) &= I(x_r, y_r) \qquad x_{Cr,l} - \frac{1}{3} D_e \le x_{r,l} \le x_{Cr,l} + \frac{1}{3} D_e \\ I_{FER}\ (x_L, y_L) &= I(x_l, y_l) \qquad y_{C,l} - \frac{1}{3} D_e \le y_{r,l} \le y_{Cr,l} + \frac{1}{3} D_e \end{aligned} \tag{5.30}$$

Occasionally the target object can be occluded; in this case the eye is usually occluded by the hand. To enable the system to evaluate when occlusion or wrong tracking occurs, the Sum of Absolute Difference (SAD) value of FER in between two frames I_{SAD} is obtained as indicated follows:

$$I_{SAD} = \sum_{i,j} \left| I_1(i,j) - I_2(x+i, y+j) \right| \tag{5.31}$$

The input image is converted to gray scale to obtain the normalised values. The SAD value is normalized that defined as I_{nSAD} which is computed as follows:

$$I_{nSAD} = \frac{I_{SAD}}{255 \times W \times H} \tag{5.32}$$

where *W and H* denote the width and height of FER respectively. The value of I_{nSAD} is examined in order to ensure that the eyes are always tracked accurately. Figure 5.11 shows a typical sequence of frames from a normal eye opening until the eye is occluded. As indicated in Fig. 1 the differences of SAD in FER of the eye opened and the eye closed does not indicate significant changes from frame 1 to frame 60 (Figure 5.12).

However, when the eye is totally occluded in frame 68 the normalized SAD value for the FER frame 68 increases rapidly (Figure 5.13). This significant change is used as an indicator in order to stop the tracking. The steps involved in the complete IASMS algorithm are given in Table 5.1.

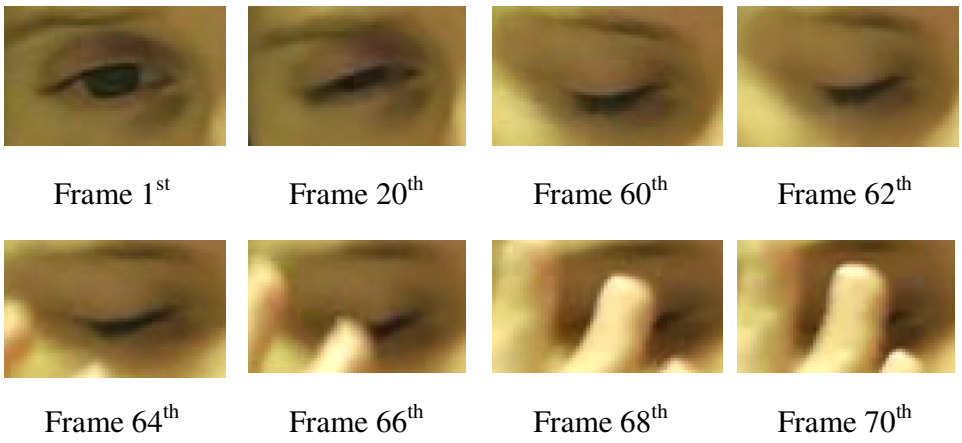


Figure 5.12: Sequences of frames from 1st frame until frame 70th that show eye begin with normally open until the eye occluded by hand

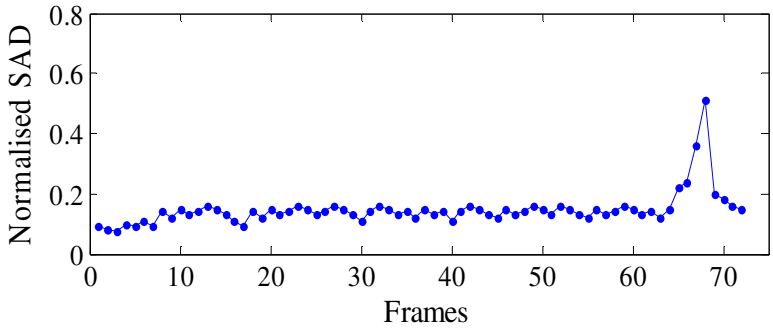


Figure 5.13 The normalised SAD values for a sequence of frames, which.is from 1st frame until frame 70th are shown. The normalised SAD increased when FER region is occluded as indicated in between frame 65 to 70.

Table 5.1 IASMS Algorithm

1. Three stages of the algorithm to obtain the iris region
   $\Rightarrow$ Face detection → Eyes detection → Iris localisation.
2. Calculate the distance D_e using (5.29).
3. Define the FER using (5.30).
4. Compute the probability colour distribution histogram $\hat{q}$ of the iris area using (5.11).
5. Compute the target candidate iris model $\hat{p}(y)$ using (5.16).
6. Compute the iris target candidate weight w_i using (5.18).
7. Compute the new iris target candidate y_1 using (5.17).
8. If $\|y_1 - y_0\| < \varepsilon$ then STOP (set default $\varepsilon = 0.1$). GO to the next step. Otherwise $y_0 = y_1$ and GO to step 2.
9. Estimate the width and height of the iris using (5.24) and (5.25). GO to step 2.
10. Calculate normalize value I_{nSAD} FER using (5.31) and (5.32).
11. Check I_{nSAD} if $I_{nSAD} > 0.4$ GO to step 1 otherwise GO to step 2.

5.5 Experiment Results

In order to evaluate the performance of the developed algorithms for face detection, blink detection and eye's state analysis, and eye tracking, extensive experiments have been performed. The databases of images and videos as discussed in Chapter 4 are utilised.

5.5.1 Face Detection

For the face detection algorithm evaluation performance, several images from the CMU [157], MIT [158], FERET [159],and CAS_PEAL [160] data bases were used. The results of the experiment can be viewed Table 5.2. The proposed algorithm, using a Viola-Jones classifier, is similar to the OpenCV-Viola Jones [15]. The same percentage

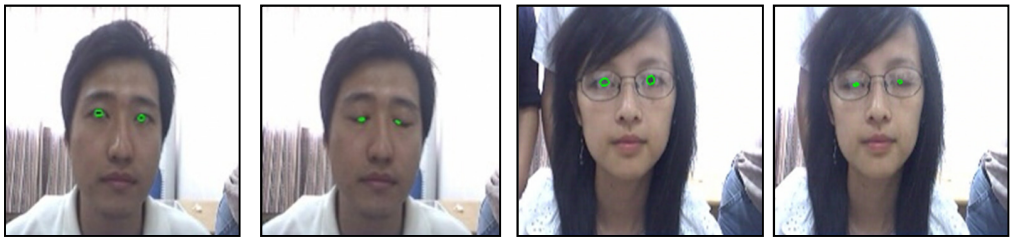
detection rate is obtained, however the false positive error is reduced significantly using the proposed algorithm.

Table 5.2 Face detection algorithm experiment result

Proposed Algorithm		Split Up SNoW Classifier Nielson [182]		Kienzle's Algorithm [183]		Tolga- Colour based detection [174]		OpenCV –Viola Jones [15]	
Detect rate (%)	False+ ve (%)	Detect rate (%)	False+ ve (%)	Detect rate (%)	False+ ve (%)	Detect rate (%)	False+ ve (%)	Detect rate (%)	False+ ve (%)
95.1	**1.8**	**95.1**	**7**	**91.2**	**11.9**	**74.5**	**13.8**	**95.1**	**19**

5.5.2 **Blink Detection**

For evaluating the performance of blink detection using the proposed algorithm, the ZJU eye blink database that contains 80 video clips is used. There are a total of 255 blinks in this database (example images from the ZJU database are shown in Figure 5.14) and the numbers of blinks in the clips varies between one to six blinks in each case. Since the IASMS tracking algorithm quantifies the size of the iris in sequences of frames, the blink of the eye is detected when the area of the iris becomes smaller as indicated in Figure 5.15. There are four video clips representing two, three, four and six blinks respectively and the results shown in Figure 5.15(a)-(d). The plot of the iris area drops substantially when the eye is closed and returns to its initial value when the eye is opened.



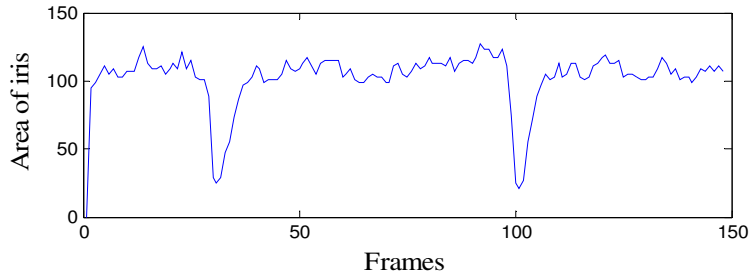
(a) (b) (c) (d)

Figure 5.14: Search area of iris, (a) and (c) are open eyes, (b) and (d) are closed eyes. *(Images are permitted to be published)*

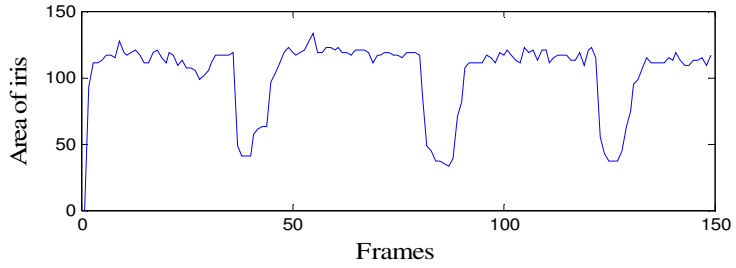
The performance of this algorithm in terms of blink detection rate is compared to the three methods that use the same database. These include Danisman's [36], Divjak's[15] and Gang's[19] methods. The IASMS algorithm successfully detected all 255 blinks in the video clip ZJU database (as shown in Table 5.3) compared to only 97% from the best of the other methods.

Table 5.3 Detection rate comparison for ZJU blink database.

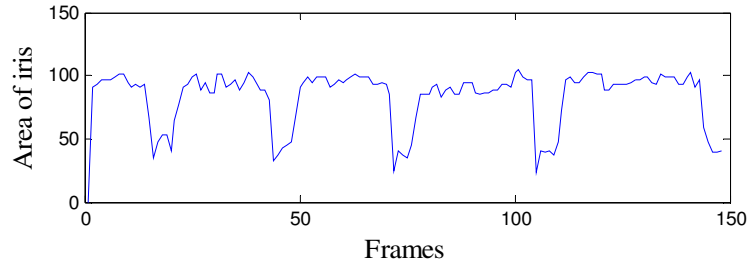
Method	Proposed method	Danisman's Method [171]	Divjak's Method [131]	Gang's Method [135]
Detection rate	100%	95%	97%	96%



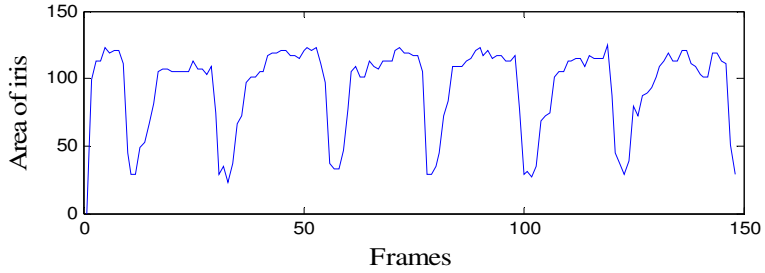
(a)



(b)



(c)



(d)

Figure 5.15: Results of eye state analysis for (a) 2 blinks; (b) 3 blinks; (c) 4 blinks; (d) 6 blinks.

5.5.3 **Eye's State Analysis**

In contemporary computer vision technology, not only a high accuracy of the eye blink detection rate is demanded, but also the ability to measure the period of the eye closure is required. To test the performance of the proposed algorithms, the Strathclyde Facial Fatigue (SFF) video footage database was used. The proposed algorithms track and evaluate the performance of the eye state analysis by examining the eye closure duration for every stage of sleep deprivation. Figure 5.16 shows the plots of the iris area over 30 seconds for each stage of sleep deprivation obtained from different experimental sessions in the SSF database. The aim is to find out the number of eye closures, total time of eye closure and the average time of eye closure during the 30 second period as shown in Table 5.4. The duration of eye closure is determined by the threshold value of the iris area T_{Aoi} which is calculated as follows:

$$T_{Aoi} = \frac{40}{100} \times Average\ of\ iris\ area \tag{5.33}$$

Table 5.4 Detection rate comparison for ZJU blink database.

Sleep Deprivation Stages	Average of area of iris	Number of eye closures	Total time of eye closure	Average time of eye closure
0 hour	59	4	0.73	0.18
3 hours	41	7	1.23	0.18
5 hours	130	10	6	0.85
8 hours	172	8	8.2	1

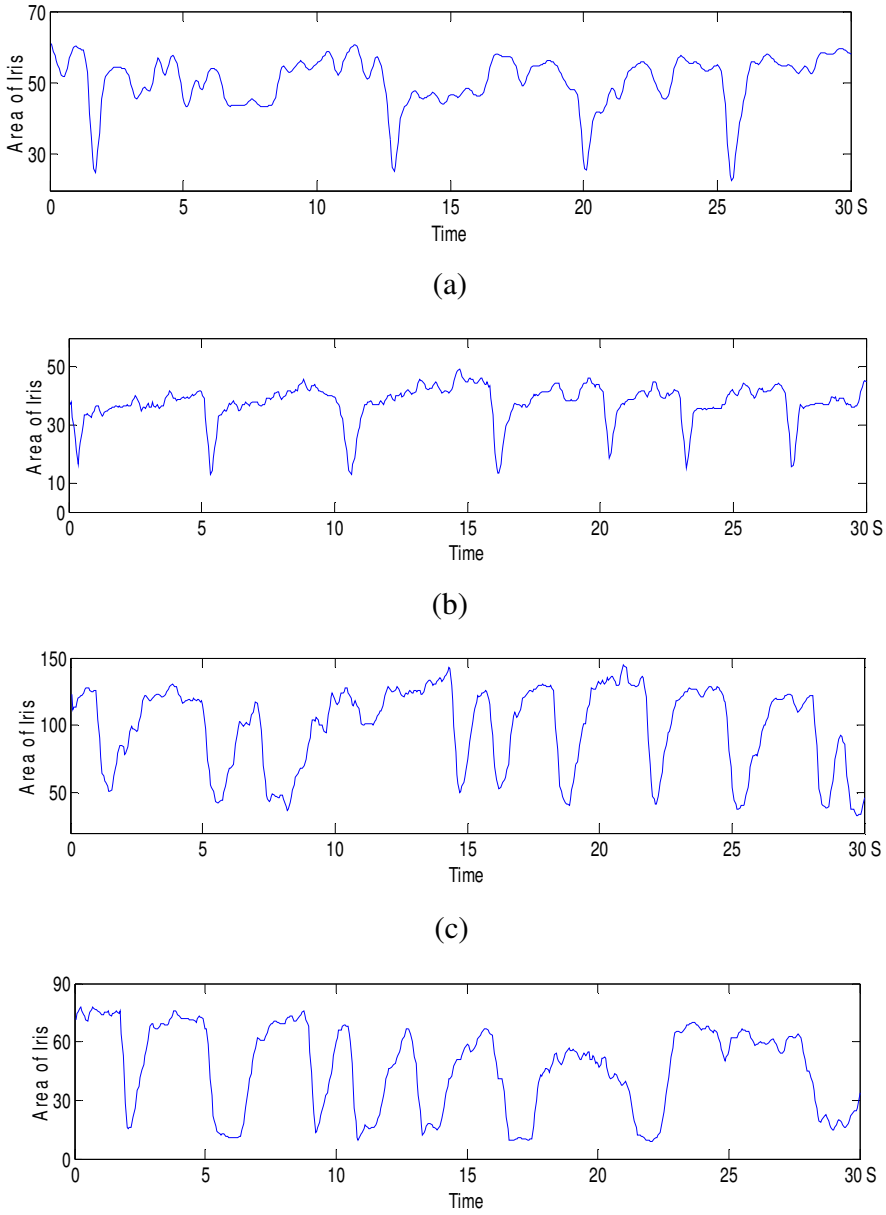


Figure 5.16: Plot of iris's area in 30 seconds: (a) 0 hour sleep deprivation; (b) 3 hours sleep deprivation; (c) 5 hours sleep deprivation; (d) 8 hours sleep deprivation

The 40% threshold value in (5.33) is determined from experiment that involved 420 frames of images from 12 videos footage (25fps) which are taken from JZU and (SFF) video footage database. In order to validate the accuracy of algorithm in

measuring the area of iris, the irises areas in each 420 frames are manually labelled, and then the pixel number of area are measured. Figure 5.17 show the plot of comparison between iris areas measured by the proposed algorithm over labelled iris area for two subjects. From the plot, the iris areas measured by developed algorithm are shown larger compared to labelled iris area. This is because the proposed algorithm also has been measured some background in tracking iris area. Table 5.5 shows the result of average iris area for eye open, eye closure and eyes half open from 420 frames tested. Based on this experiment, the best percentage threshold that can determine the eye closure by proposed algorithm is 40%. This 40 % represents the eyelashes area which is still recognised as a dark object in the search region even when the eye is closed.

For validating the accuracy in measuring period of eye closure, 1250 frames from SFF video footage for four different sleep deprivations have been used. Table 5.4 shows the result as evident that the IASMS tracking algorithm is capable of determining the duration of eye closure which is indicative of the difference between various stages of sleep deprivation. The average time of eye closure becomes progressively longer for longer periods of sleep deprivation.

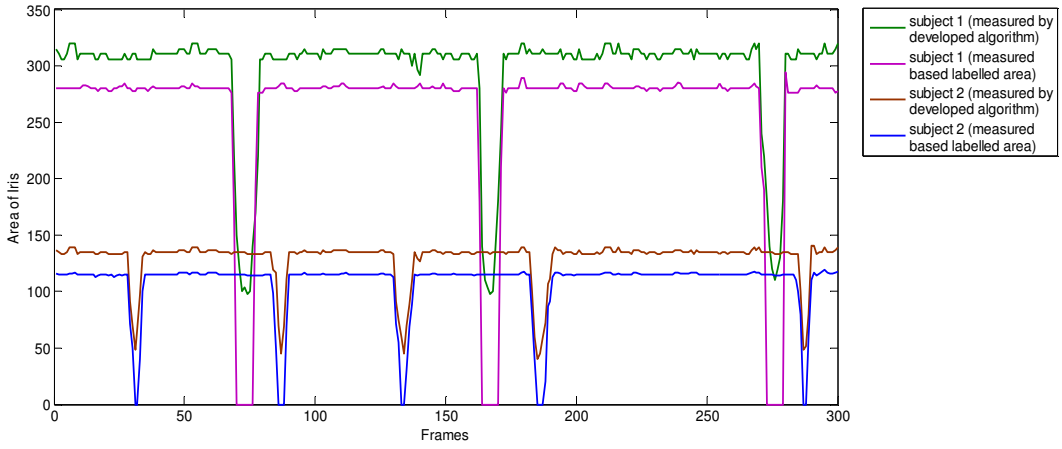


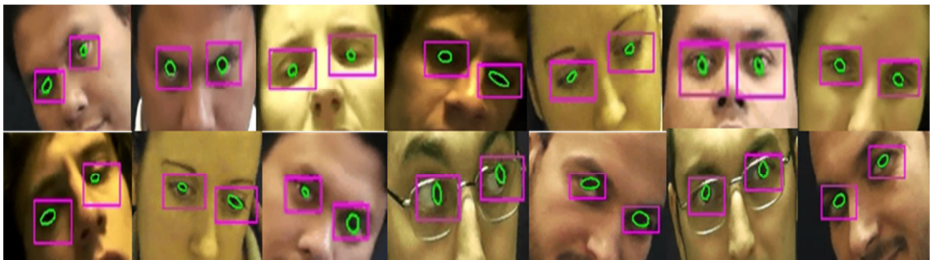
Figure 5.17: Plot of comparison between iris areas measured by the proposed algorithm over labelled iris area for two subjects

Table 5.5 . The average of iris area upon labelled measurement and algorithm measurement.

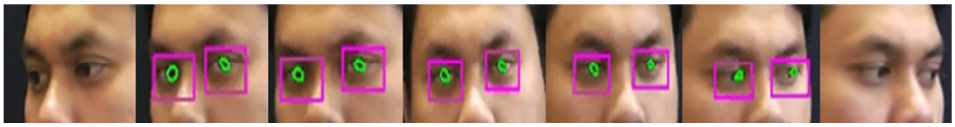
IRIS AREA MEASUREMENT	Eye opened	Eye closure	Half eye opened
Labelled	202	**0**	101
Proposed algorithm	224	86	155

5.5.4 Eyes Tracking

The IASMS tracking algorithm is intended to track the non-rigid eye movements in order to obtain the iris area plot (Figure 5.16). As discussed in section 5.4.3, the eye tracking algorithm can stop the iteration when the FER is occluded or when the FER is out of the tracked region. Similarly, the FER is also sensitive to rigid eye movement which may also cause the algorithm to stop the tracking iteration since, in such cases, facial information may suddenly move outside the camera view field. Figure 5.17(a) shows the tracked eyes images which indicate the algorithm's ability to track the eyes in multi-pose of face as long as the FER region is not occluded or out of the search region. In Figure 5.18(b) the FER is out of the search region when the face is turned too much to the right or left (first and last views). This eye tracking algorithm emphasizes that both eyes must be seen clearly in order to read the iris information, otherwise the algorithm stops and the initialisation scheme needs to be re-initiated to find out the new location of the eyes.



(a)



(b)

Figure 5.18: (a) Eyes tracking for multi-poses; (b) Eye tracking stops when FER is out of search region

5.6 Conclusion

This chapter has presented a system developed for measuring the activities of the eyes such as eye blinking and eye closure time in order to recognise symptoms of fatigue. The system consists of face acquisition operation algorithms, iris localisation algorithm, and IASMS algorithm. In face acquisition, a new face detection algorithm is introduced. This approach is a combination of colour skin segmentation, connected components binary image, and a machine learning classifier algorithm. Using a Viola-Jones classifier [15], the introduced approach is able to reduce the false positive detection rate by a very significant margin. The false positive error is a command error which happens particularly in complex background images.

In iris localisation algorithm, the Daugman's iris localisation method [163] is applied. This method is combined with a proposed technique to detect the eye open state in order to ensure that the actual size of iris is obtained. The last algorithm is the Interdependence and Adaptive Scale Mean Shift (IASMS) algorithm. This is a new proposed algorithm that has been designed for robust and accurate eye state tracking and

analysis. This algorithm is a combination of a tracking algorithm with an adaptive scale approach to track the eyes and measure the size of the iris. The iris area is measured over successive frames, and the eye closure is detected when the area of the iris is below a threshold value. From the experimental results, this algorithm is able to detect the eye closure as well as quantify the period of the eye closure. This is important for fatigue detection systems.

The algorithm also incorporated a new approach of eye tracking. This approach utilised the Mean Shift algorithm with eye location information to track non-rigid eye movement. In this algorithm, the tracking iteration is stopped when the eye is occluded or is out of the search region. The IASMS tracking algorithm was successfully combined with face acquisition algorithms, an eye and an iris localization algorithms, as well as an eye open detection approach in order to measure the eyes activities.

6. Yawning Analysis

6.1 Introduction

In chapter 5 novel techniques to measure the eyes activities which represent the dominant signs of fatigue were discussed. In this chapter, novel techniques to analyse yawning which is another common sign of fatigue are described. Yawning is a symptom visually indicated by a widely open mouth. Because of this, the entire yawning detection research focuses on measuring and classifying the opening of the mouth. Frequently however this approach is thwarted by the common human reaction to hand-cover the mouth during yawning. In this thesis, we introduce a new approach to analyse yawning by combining mouth opening measurements, covered mouth detection, and a facial distortion (wrinkles) detection as shown in Figure 6.1. In this approach the region interest; Focused Mouth Region ((FMR) and Focused Distortion Region (FDR) are initialised in beginning frame which is discussed in next section.

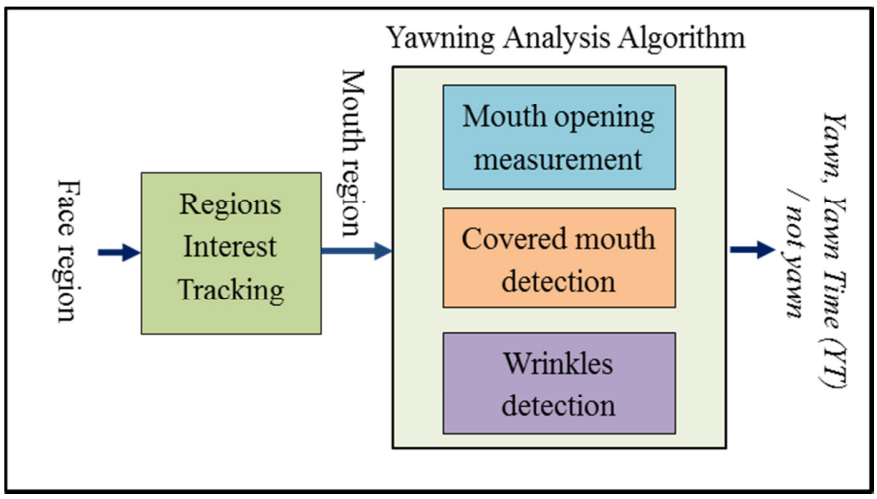


Figure 6.1 Yawning Analysis method.consists of region interest tracking operation and yawning analysis algorithm.

Then, mouth opening measurement with a new adaptive threshold for segmenting the mouth region is introduced. In section 6.4 the covered mouth detection that uses LBP features and a learning machine classifier are explained. Next, the distortion detection that apply the Sobel operator in region FDR is discussed. The yawning analysis algorithm which is a combination of mouth opening detection, covered mouth detection and distortion detection is described in the following section. Lastly, the experiment results over developed algorithms are explained before chapter ending with conclusion.

6.2 Region Interest Initialisation and Tracking

For region initialisation operation, the eyes and mouth location are necessary to be firstly detected. This can be obtained by applying the technique that discussed in section 5.2. When the eyes and mouth location are obtained, an anthropometric measurement is carried out to obtain the distinctive distances between these facial components. This information is then used in the formation of the two regions of interest.

6.2.1 Focused Mouth Region

The Focused Mouth Region (FMR) is formed based on the first detected location of the eyes and mouth. The distances of these two main face components are measured. A distance between the centre of eyes *ED*, and a distance between centre of mouth and the middle point between eyes *EMD* are obtained, as shown in Figure 6.2(a). Then, the ratio of these distances R_{EM} is computed as:

$$R_{EM} = \frac{ED}{EMD} \tag{6.1}$$

(a) (b)

Figure 6.2: (a) Anthropometric measurement on face. (b) Focused Mouth Region (FMR) is empirically defined using anthropometric measurement.

From these two distances the coordinates of the FMR coordinates (x_1,y_1) and (x_2,y_2) are empirically defined as:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_R \\ y_R + 0.75EMD \end{bmatrix} \tag{6.2}$$

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_L \\ y_L + 0.75EMD + 0.8ED \end{bmatrix} \tag{6.3}$$

where (x_R,y_R) and (x_L,y_L) are the centre points of the right eye iris and left eye iris respectively. The threshold values 0.75 and 0.8 which are used in equation (6.2) to (6.3) are based on the observation tests, undertaken on the SFF database. The FMR is dependent on the location and distance of the eyes. When the face moves forward the eye distance increases and so does the FMR. On the other hand, the FMR becomes smaller when the face moves backwards.

6.2.2 Focused Distortion Region

The Focused Distortion Region (FDR) is a region in the face that is used to measure changes of facial distortion. This region, which is most likely to undergo

changes (i.e. facial distortions) during yawning, was identified based on conducted experiments using the SFF database. The coordinates (x_1,y_1) and $(x_2,y_2)$of FDR, shown in Figure 6.3(b), are empirically defined as:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_R \\ y_R - 0.75EMD \end{bmatrix} \tag{6.4}$$

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_L \\ y_L + 0.75EMD \end{bmatrix} \tag{6.5}$$

(a) (b)

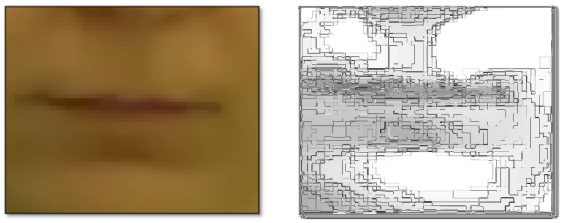
Figure 6.3: (a) Anthropometric measurement on face. (b) Focused Wrinkles Region (FWR) is empirically defined using anthropometric measurement.

6.3 Mouth Opening Measurement

Yawning is an involuntary action that causes us to open the mouth widely and breathe in deeply. In this paper a new technique to measure the width of mouth opening is introduced. This technique measures the mouth opening based on the darkest region between the lips. The yawning then is detected based upon the widest area of the darkest region. As discussed in Chapter 3 when applied techniques are based on intensities and colour values, the challenge is to obtain an accurate threshold value which will adapt to

multi-skin and lips colouration as well as to the lighting conditions. Therefore the proposed algorithm applies a new adaptive threshold in order to meet this challenge.

This adaptive threshold value is calculated during the initialisation operation after the FMR is formed. Due to the technique being depended on the intensity value of the region, it is necessary to improve the face image in order to heighten the darkest and the brightest facial regions. The enhancement is implemented by transforming the histogram of the image to spread the level of colour values evenly but without changing the shape of the input histogram as implemented in section 5.3. Figure 6.4 show example enhanced image of FMR. Subsequently, the initial adaptive threshold is calculated by refer the darkest area as shown in Figure 6.4. On average the darkest area constitutes approximately 5% or less of the entire obtained region of interest. Then, the adaptive threshold is calculated using the similar technique implemented in section 5.3 for eye open detection.



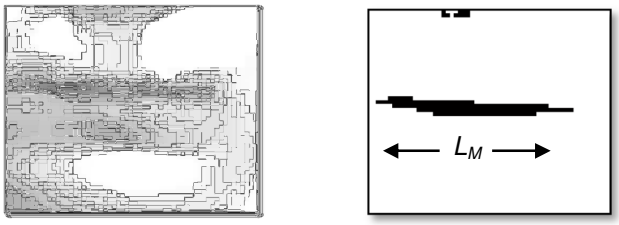
(a) (b)

Figure 6.4: (a) FMR image. (b) Enhanced FMR image

When the adaptive threshold T_a calculated, the FMR is segmented and the mouth length L_M is obtained as shown in Figure 6.5. The length of L_M must be within the range indicated in (6.6).

$$0.7ED < L_M < 0.9ED \tag{6.6}$$

If L_M is less than $0.7ED$, the value T_a is outside the limits given in (6.6). In this case T_a is increased or decreased accordingly in steps of 0.001, also repeating the segmentation process until the value of L_M is within the required range (6.6).



(a) (b)

Figure 6.5: (a)Enhanced FMR image (b)Segmented FMR image when apply adaptive threshold

After the adaptive threshold is finalised, the value is applied in the next sequence of frames for segmenting the darkest region between the lips. In order to determine the yawning, the height of the mouth $H_{M,}$ based on the darkest region in FMR, is measured as shown in Figure 6.6. Then, the Yawn Ratio (*YR)* of the height H_M over the height of FMR is computed as follows:

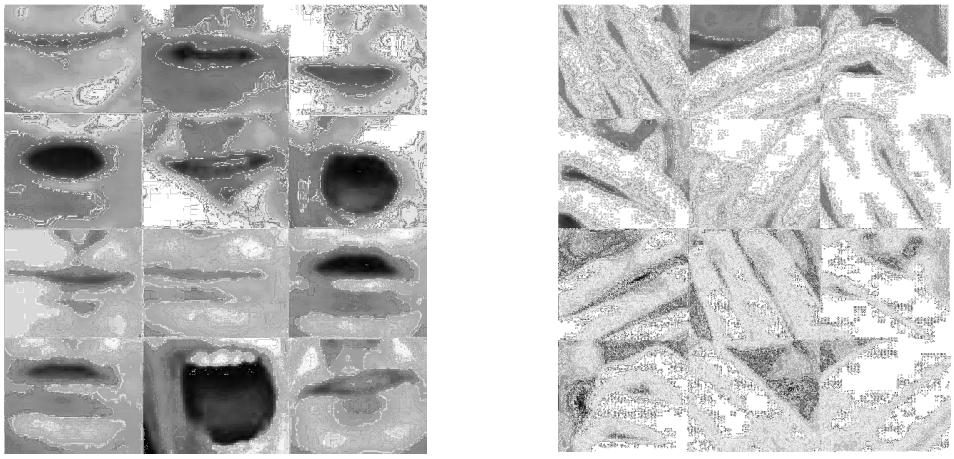
$$YR = \frac{H_M}{FMR_{height}} \tag{6.7}$$

(a) (b)

Figure 6.6: The measurement of high of mouth opening (a) input image (b) the segmented mouth opening region.

6.4 Mouth Covered Detection

It is not unusual to hand-cover the mouth during yawning. This means that the yawn can no longer be detected from the mouth opening. Therefore, this thesis introduces a new approach to detect yawn while the mouth is covered. In this approach the FMR is examined for whether the region is covered or not. Figure 6.7(b) shows image examples where the FMR is covered and Figure 6.7(a) when the mouth is not covered. From these images the texture difference with the FMR region covered and not covered is clearly visible. Hence, in order to differentiate between these two regions LBP features are chosen to extract the texture pattern from FMR. LBP features have been proven highly accurate descriptors of texture, robust to the monotonic gray scale changes as well as simple to implement computationally [138].



(a) (b)

Figure 6.7: (a) Example of uncovered mouth images. (b) Example of covered mouth images.

6.4.1 **Local Binary Pattern (LBP)**

Local Binary Pattern (LBP) originally applied on texture analysis to indicate the discriminative features of texture. With reference to Figure 6.8, the basic LBP code of the central pixel is obtained by subtracting its value form each of its neighbours and when a negative number is obtained it is substituted by a 0, else it is 1.

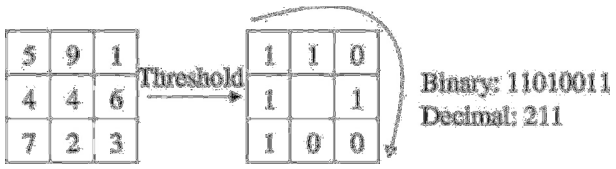
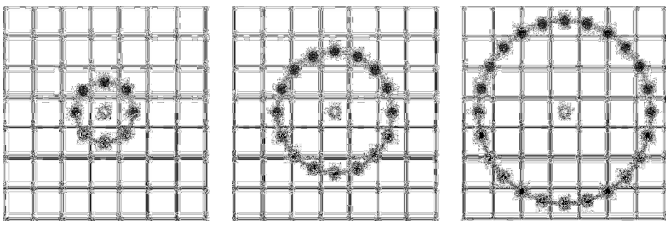


Figure 6.8: Example of the basic LBP operator for eight neighbors

The limitation of the basic LBP operator is that it represents a small scale feature structure which may be unable to capture large scale dominant features. In order to deal with the different scale of texture, Ojala *et al.* [184] presented an extended LBP operator where the radius and sampling points are increased. Figure 6.9 shows the extended operator where the notation (P,R) represent a neighborhood of P sampling points on a circle of radius R.



(a) (b) (c)

Figure 6.9: Example of extended LBP operator. (a) (8,1), (b) (16,2) and (c) (32,3)

The result of LBP can be formulated in decimal form $LBP_{P,R}(x_c, y_c)$ as follows:

$$LBP_{P,R}(x_c, y_c) = \sum_{P=0}^{P-1} s(i_P - i_c)2^P \tag{6.8}$$

where i_c and i_P denote gray level values of the central pixel, P represent surrounding pixels in the circle neighborhood with radius R and (xc, yc) is centre of region. *s(x)* is a function to differentiate between centre pixel value and surrounding pixel value which is defined as:

$$s(x) = \begin{cases} 1, & if\ x \geq 0 \\ 0, & if\ x < 0 \end{cases} \tag{6.9}$$

The operator LBP as formulated in (6.8) has a rotation effect if the image is rotated, then, the surrounding pixels in each neighborhood are moving accordingly along the perimeter of the circle. In order to remove this effect Ojala [184] proposed a rotational invariant (*ri*) LBP, and the operator rotational invariant $LBP^{ri}_{P,R}$ is computed as follows:

$$LBP^{ri}_{P,R} = \min\{ROR(LBP_{P,R}, i) \mid \quad i = 0,1,....., P-1\} \tag{6.10}$$

where *ROR(x,i)* performs a circular bitwise right shift on the *P*-bit number *x* with *i* time. Ojala *et al.* also proposed a LBP uniform pattern (*u2*), $LBP^{u2}_{(P,R)}$. The LBP is called uniform when it contains two bitwise transitions from 0 to 1 or vice versa. For example, 1111111 (0 transition) and 00111100 (2 transitions) are both uniform, for 100110001 (4 transitions) and 01010011(6 transitions) are not uniform.

Another LBP operator is a combination of rotational invariant pattern with uniform pattern, $LBP^{riu2}_{(P,R)}$. This operator is calculated by simply counting ones in uniform pattern codes and all non-uniform patterns are labeled in a single bin. The example of LBP histogram of the three types of LBP operators for covered and uncovered mouth region is shown in Figure 6.10.

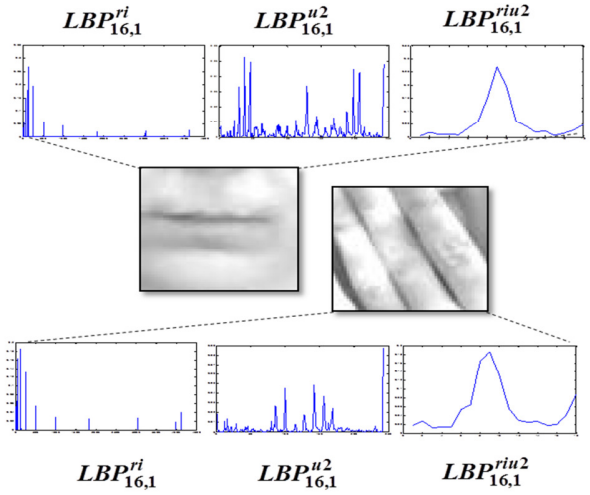


Figure 6.10: LBP histogram for not covered mouth and covered mouth regions for 3 operator with different radiuses

6.4.2 Distortions Detection

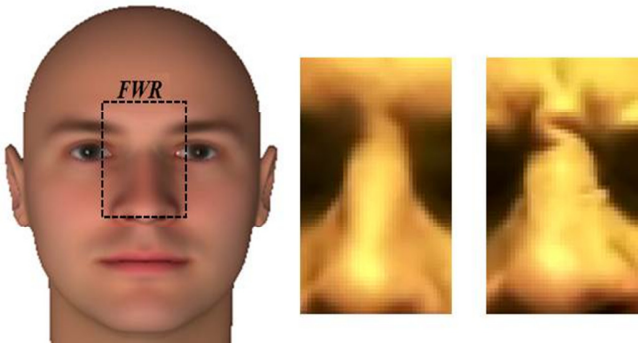
The ability to detect the covered mouth is not adequate to conclude that someone is yawning. The mouth is probably accidentally or intentionally covered - not necessarily due to yawning. Thus, this project introduces a technique to detect the distortions in a specific region of the face most likely to distort while yawning. This region is identified based on observations and the experiments conducted using the video footage from SFF database. The most affected region that shows distortion during yawning is labeled as Focused Distortion Region (FDR) as shown in Figure 6.11(a). Based upon the Figure 6.11(b) and (c), obviously, there are significant differences visually, to the area during yawning and normal situation.

In order to measure the changes in FDR, the work presented here uses edge detection to detect the distortions. Based on the experiments carried out, the Sobel operator [185] is chosen since it is able to detect most of the required edges. The Sobel operator first calculates the intensity gradient at each point in the region of interest. Then, it provides the direction of the largest possible increase from light to dark and the rate of change in the horizontal and vertical directions. The Sobel operator represents a partial derivative of f (x, y) and of the central point of a 3x3 area of pixels. The gradients for the

horizontal G_x and vertical G_y directions for the region of interest are then computed as follows [186] :

$$G_x = \{f(x+1, y-1) + sf(x+1, y) + f(x+1, y+1)\} - \{f(x-1, y-1) + 2f(x-1, y) + f(x-1, y-1)\} \quad (6.11)$$

$$G_y = \{f(x-1, y+1) + sf(x, y+1) + f(x+1, y+1)\} - \{f(x-1, y-1) + 2f(x, y-1) + f(x+1, y-1)\} \quad (6.12)$$



(a) (b) (c)

Figure 6.11: (a) Focused Wrinkles Region (FWR). (b) normal condition in FWR. (c) yawning condition in FWR

The distortions in FDR are detected based on the gradients in both directions and the result is shown in Figure 6.12. The numbers of edges detected in normal condition (Figure 6.12 (a)) is significantly different compared to yawning condition (Figure 6.12 (b)). In order to identify the distortion in FDR during yawning, the sum of the absolute values FDR_{SAD} is applied as in equation (6.13) to compute the number of wrinkles in the region. The normalized FDR_{SAD} is calculated as in (6.14) where *W and H* denote the width and height of FDR respectively. During yawning the numbers of the detected edges are increased as shown in Figure 6.13.

$$FDR_{SAD} = \sum_{i,j} \left| I_1(i,j) - I_2(x+i, y+j) \right| \tag{6.13}$$

$$Normalised\ FDR_{SAD} = \frac{FDR_{SAD}}{255 \times W \times H} \tag{6.14}$$

(a) (b)

Figure 6.12: FWR with input image and edges detected image. (a) normal condition in FWR. (b) yawning condition in FWR

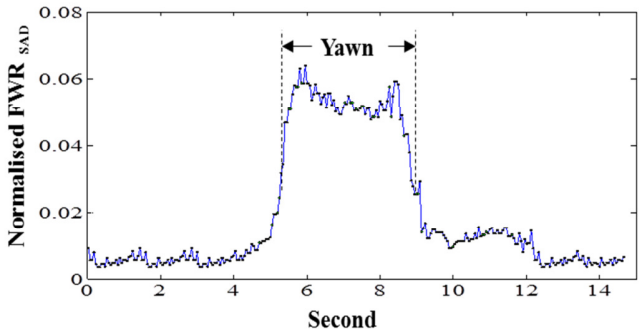


Figure 6.13: Normalised value of Focused Distortion Region (FDR) during yawning. The normalised value is dramatically increased over that 0.04 when yawn happened as shown in second of 5th to 9th.

6.5 Yawning Analysis

In general yawning can be categorised into three situations; a) mouth not-covered and wide open, b) mouth wide open for a short time before it gets covered and, c) mouth totally covered with invisible mouth opening. These situations are based on the observations made from the video footage of the SFF database. The yawn analysis

algorithm, as depicted in Figure 6.14, was therefore developed to deal with these three cases by introducing the new procedures as discussed earlier; i.e. mouth opening measurement, mouth covered detection and distortions detection.

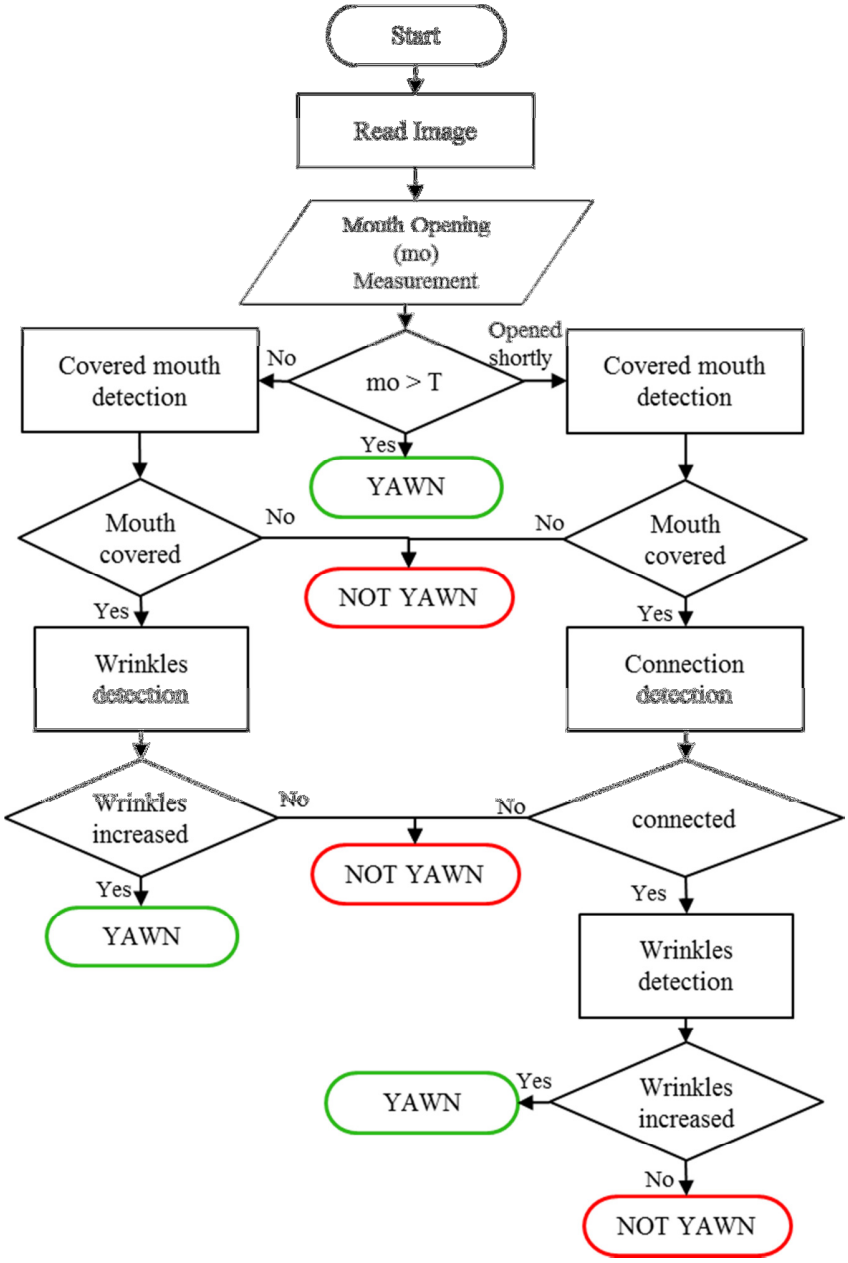


Figure 6.14: Flow chart of yawning analysis algorithm

In this algorithm, the status of the mouth opening, mouth covered, and distortion is examined periodically. The chosen time for measurements, based on the observations from the SFF database, is thirty seconds. This implies that yawning does not occur twice within this time period. In this algorithm the mouth opening is the first to be examined, and if the mouth opening measurement shows no yawning, then the mouth covered detection is triggered. The status of distortions detection is used when mouth covering is detected. In the situation where yawning occurs with the mouth wide open for a short time and then covered quickly, the algorithm checks the sequence of events, i.e. mouth open wide followed by mouth covered. As a result of this analysis, the yawn status is detected and the period of yawning is also obtained.

6.6 Experiment Results

Extensive experimentation has been carried out using the SFF database to evaluate the performance of the developed algorithms.

6.6.1 Mouth Opening Measurement

The mouth opening measurement algorithms are tested using the SFF database. In order to evaluate the accuracy of the mouth opening segmentation using the proposed adaptive threshold, the actual input image is compared with the segmented region of mouth opening Figure 6.15 shows examples of mouth opening images paired with the output image of mouth opening segmentation. 145 mouth opening images with varying lighting condition were used for testing. From these tests, it was shown that the algorithm was able to segment the darkest region as required for detecting yawning.

Figure 6.15: Mouth opening segmentation images

As discussed in section 4, yawning is determined by measuring the height of the segmented region of the mouth opening. In order to obtain the applicable threshold value that represents yawning, 25 of the real yawning scenes from the SFF database were used. From the yawning scenes it was found that the minimum value of Y*R* in equation (6.7) which denotes yawning is 0.5. This ratio value is applied in mouth opening measurements during yawn analysis.

6.6.2 **Mouth Covered Detection**

In order to detect the covered mouth, the FMR is examined in every video frame. The LBP features are extracted from the region, and then the learning machine algorithm decides the status. In this research, two classifiers were applied, Support Vector Machine (SVM) and Neural Network (NN), in order to determine their suitability for use in the analysis algorithm. A total of 500 covered mouth images and 100 non covered mouth images (as shown in Figure 6.7) of size 50X55 pixels have been used to train the NN and SVM classifiers. For the features extraction three operators of LBP; rotational invariant (*ri*), uniform pattern (*u2*), and rotational invariant pattern with uniform pattern (*riu2*), were applied with respective radius of 8, 16, and 24.

The results of the classification for 300 images for the SVM classifier in four different kernel functions and the NN in two different network layers are stated in Table 6.1. Table 6.1 describes the percentage of detection rate of true positive covered mouth

detection T_P, false positive F_P, true negative for uncovered mouth detection T_N and also the false negative F_N. The highest percentage of covered mouth detection is achieved by 5 layer network of Neural Network classifier which uses 16 sampling points of LBP uniform pattern. There are several classification results have obtained perfect detection for uncovered mouth detection. For false detection, 10 classification results show no errors for false positive covered mouth detection, and only 5 layer network of Neural Network classifier shows no error for false negative.

From the obtained classification results, the sensitivity and specificity of the detection for covered mouth and uncovered mouth can be calculated using equation (6.15 and (6.16). Sensitivity represents the ability of the developed algorithm to identify the covered mouth. The specificity relates to the ability of the algorithm to detect the non-covered mouth. Based on the classification results and the sensitivity and specificity calculation the performance of detection can be represented in Receiver Operating Characteristic (ROC) curve as shown in Figure 6.16 to Figure 6.18. These figures show the performance of six classification results for LBP features operator rotational invariant pattern, rotational invariant pattern with uniform pattern, and uniform pattern respectively. From these three curves, LBP uniform pattern show the best result classification particularly for Neural Network classifier.

$$Sensitivity = \frac{T_P}{T_P + F_P} \tag{6.15}$$

$$Specificity = \frac{T_N}{T_N + F_N} \tag{6.16}$$

Table 6.1 Mouth covered detection rate.

		ri		*u2*			*riu2*		
		8	16	8	16	24	8	16	24
SVM- rbf	T_P	59	50.8	51.6	54.2	55	98.3	84.1	66.7
	F_P	2.2	0	0	0	0	27.8	17.8	6.6
	T_N	97.8	100	100	100	100	72.2	82.2	93.4
	F_N	41	49.2	48.4	45.8	45	1.7	15.9	33.3
SVM-linear	T_P	97.5	90	95	98.3	95.8	97.5	99.2	94.2
	F_P	14.4	25	2.2	2.2	0	14.4	12.8	15.6
	T_N	85.6	75	97.8	97.8	100	85.6	87.2	84.4
	F_N	2.5	10	5	1.7	4.2	2.5	0.8	5.8
SVM-polynomial	T_P	86.7	63.3	57.8	50.8	55	95	82.5	91.7
	F_P	24.4	19.4	0	0	0	25	22.2	16.1
	T_N	75.6	80.6	100	100	100	75	77.8	83.9
	F_N	13.3	36.7	42.2	49.2	45	5	17.5	8.3
SVM-quadratic	T_P	80	63.3	83.3	78.3	69.2	95	85	85.8
	F_P	18.3	26.1	1.7	0	0	27.8	26.1	16.1
	T_N	81.7	73.9	98.3	100	100	72.2	73.9	83.9
	F_N	20	36.7	16.7	21.7	30.8	95	82.5	14.2
NN-(10)	T_P	92.5	90.8	83.3	97.5	96.7	98.3	92.5	95.8
	F_P	15	15	1.1	1.1	1.7	12.8	16.7	12.8
	T_N	85	85	98.9	98.9	98.3	87.2	83.3	87.2
	F_N	7.5	9.2	16.7	2.5	3.3	1.7	7.5	4.2
NN-(5)	T_P	91.2	95	90	100	98.3	99.2	92.5	95
	F_P	8.9	16.7	2.8	3.3	1.1	13.9	11.1	16.1
	T_N	92.1	83.3	97.2	96.7	98.9	86.1	88.9	83.9
	F_N	8.8	5	10	0	1.7	0.8	7.5	5

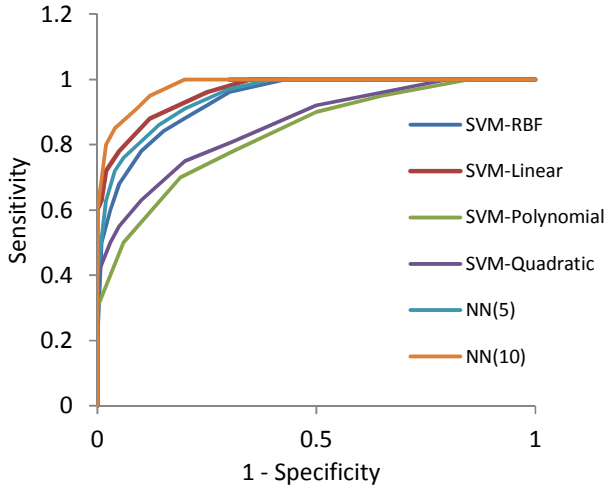


Figure 6.16: ROC curve for six classification results for LBP rotational invariant (*ri*) operator

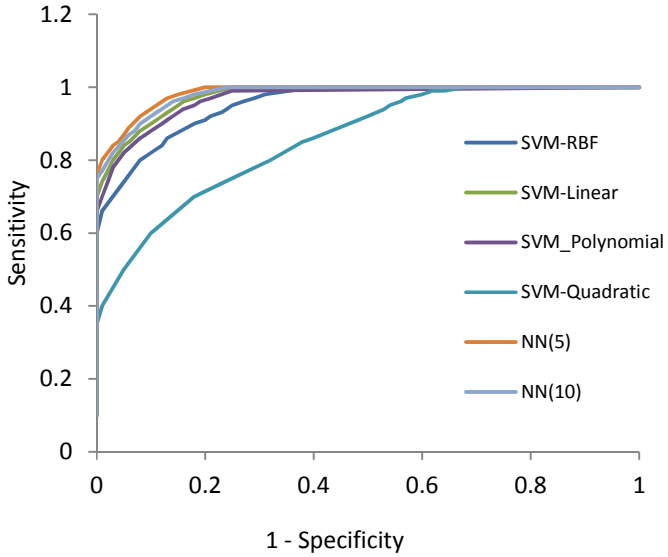


Figure 6.17: ROC curve for six classification results for LBP rotational invariant pattern with uniform pattern (*riu2*) operator.

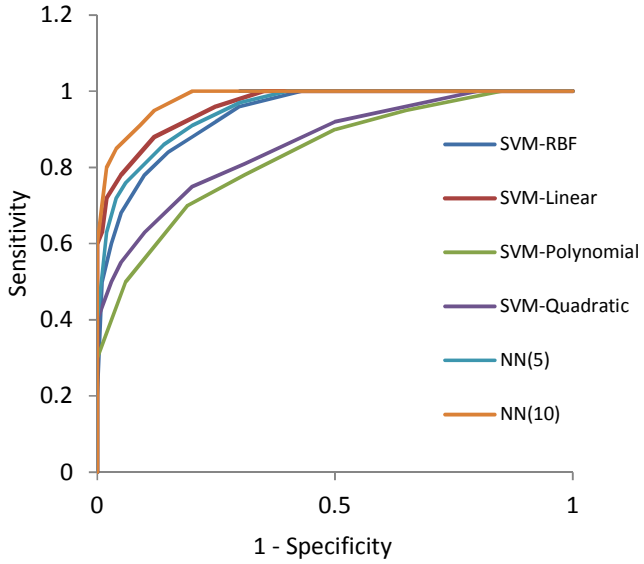


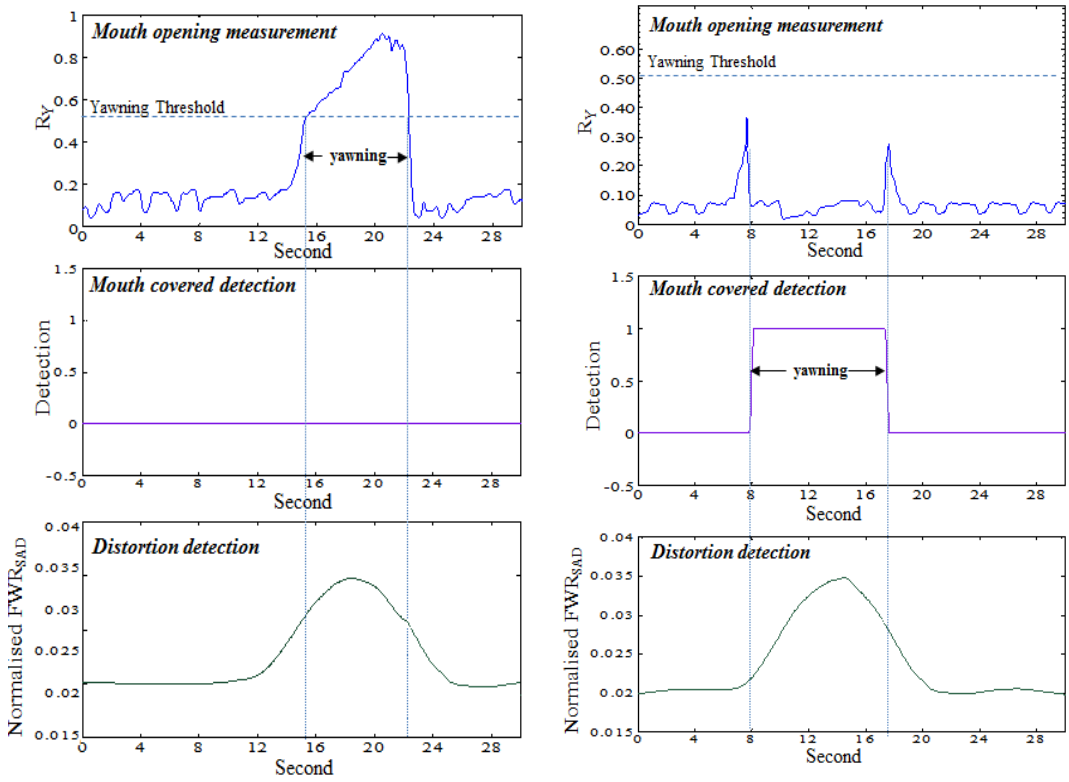
Figure 6.18: ROC curve for six classification results for LBP uniform pattern (*u2*) operator.

6.6.3 Yawning Analysis

The yawning analysis is a combination of mouth opening measurement, mouth covered detection, and distortions detection. These operations are examined every thirty seconds. For mouth opening measurement, the best threshold value that represents yawning is 0.5. Based on the detection results for mouth covered classification, LBP uniform pattern (*u2*) with radius 16 and a Neural Network as the classifier are chosen to determine the status of FMR. 30 video clips of genuine yawning images from the SFF database were used for evaluation of the algorithm's performance. These images were not used for training mouth covered detection. These videos contain 10 yawning images with non-covered mouth scenes, 14 with covered mouth scenes and 6 scenes with both situations present. The plots of the detection results in each situation of yawning are shown in Figure 6.19 and Figure 6.20.

In the situation of yawning with non-covered mouth, the plots of results are shown in Figure 6.19 (a). Yawning is detected when the height of the mouth opening is equal to

or greater than the threshold value. The yawning period is measured from the beginning of the intersection point between the threshold value and the ratio of mouth opening height to FMR height R_Y. In the situation where the mouth is covered during yawning, no R_Y value exceeds the threshold value as shown in Figure 6.19(b). In this case, the mouth covered detection is triggered. When mouth covered is detected the distortions within the mouth covered period are checked. Yawning is assumed to occur when the value of the wrinkles increases substantially within that period.



(a) (b)

Figure 6.19: Plots of detection result (a) yawning with mouth region not covered (b) yawning with covered mouth region

For the last situation of yawn analysis, where the mouth is opened wide over a short time before it is covered, the plots of the results are shown in Figure 6.20. The value of R_Y exceeds the threshold values for a short period before it drops quickly. The period of yawning is measured from the beginning of the intersection of R_Y with the threshold value until the end of the mouth covered detection. From the 30 video scenes of yawning, 28 of them can be detected successfully with the period measured accurately. However the algorithm failed to detect two yawning scenes where the mouth is completely covered. This is because the FDR did not indicate any distortion.

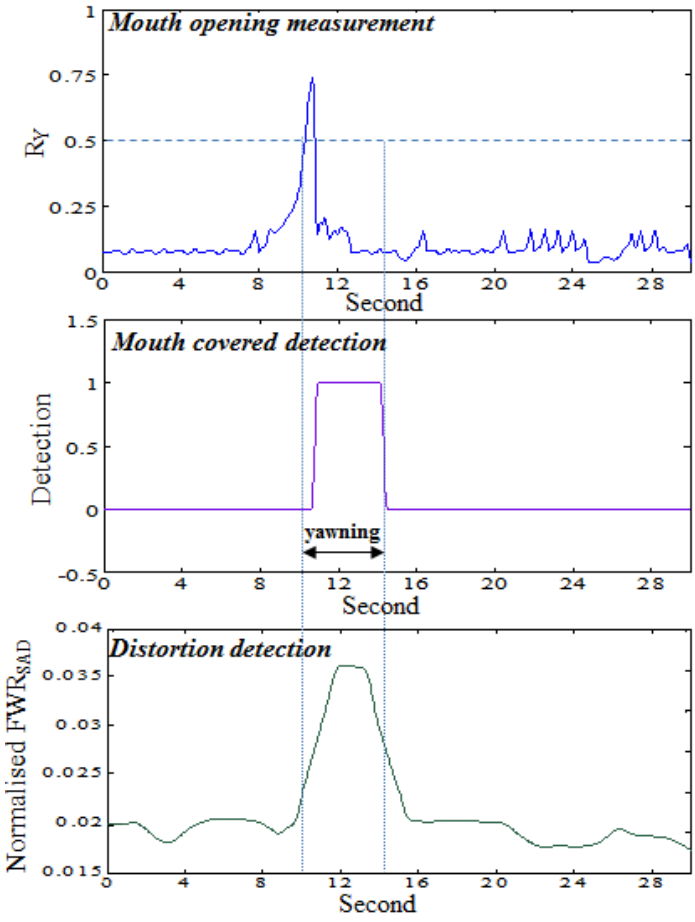


Figure 6.20: Plots of detection result for yawning with mouth not covered for a while before it is covered.

6.7 Conclusion

This chapter has presented a new approach to detect and analyse yawning which deals with both covered and uncovered mouth yawning situations. This approach is a combination of a new technique of mouth opening measurement, mouth covered detection, and distortion detection on specific region. The mouth opening is measured based on the darkest region between the lips. The darkest region is segmented using a proposed adaptive threshold. For detecting the covered mouth, an LBP uniform operator is applied to extract features of the mouth region, and then the mouth covered is classified using a learning machine classifier. Based on the carried out experiment, the multilayer perceptron (MLP) neural network with uniform pattern ($u2$) LBP features has shown the best result.

In order to verify that the mouth covering is indeed due to yawning, the wrinkles of specific regions on the face are measured. Yawning is confirmed when wrinkles changes are detected. To detect the wrinkles Sobel operator edges detector is applied. In this yawning analysis algorithm, two regions of interest are introduced, FMR and FDR, which are monitored in every frame. FMR is for monitoring the mouth activities whereas the FDR is to measure the wrinkles. To analyse yawning the status of the mouth opening, the mouth covered detection, and the wrinkles changes detection are examined every thirty seconds. The analysis results produce the status of yawning as well as the period of yawning. In this research the genuine signs of yawning from SFF database video footage have applied for training, testing and validate the performance of the developed algorithms. Based on the conducted experiments, the developed algorithms shows very promising results.

7. Fatigue Recognition

7.1 Introduction

Chapter 5 introduced techniques to measure the eye state activities and Chapter 6 presented techniques to analyse yawn. In this chapter a new approach is introduced in which fatigue is classified based on sleep deprivation levels. Figure 7.1 shows the block diagram of the integration of the facial fatigue detection algorithms with the Fatigue Monitoring Tool (FMT). The FMT is a tool which is developed for monitoring facial characteristics, evaluating fatigue and classifying it into different levels (refer Appendix C for more details). In previous chapters, the algorithms for face image acquisition and the specific facial features extraction for fatigue signs were discussed. In this chapter, a fatigue recognition technique which associates the features vectors extraction techniques with the classification algorithm is explained.

In the next section, the fatigue recognition algorithm that comprises the features vectors extraction and the classification operation is discussed. During the features vector extraction operation, the features are extracted from eyes activities and yawn detection. The extracted feature vectors are then classified into levels using a machine learning algorithm. Subsequently in section 7.3, the conducted experimental results are presented and the performances of the developed fatigue recognition algorithms are discussed. Finally, section 7.4 contains the conclusion of the chapter.

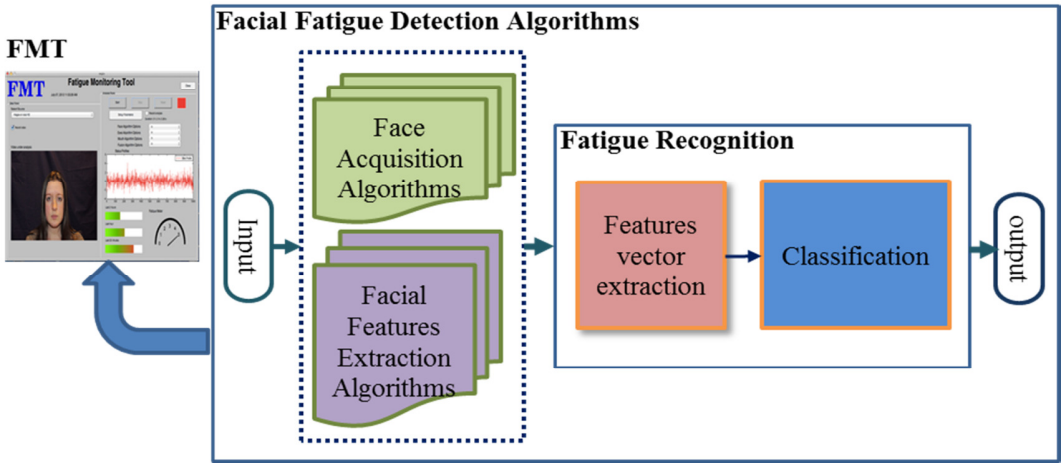


Figure 7.1: Block diagram of Fatigue Monitoring Tool system. The input of the system is image and the out is level of fatigue which based on sleep deprivation hours

7.2 Fatigue Recognition Algorithm

In this research, the signs of fatigue are classified into levels which are based on the sleep deprivation hours. There are four sleep deprivation hours used as references; 0, 3, 5 and 8 hours, which relate to the corresponding sleep deprivation hours of the SFF database. There are two operations involved in the fatigue recognition algorithms as shown in Figure 7.1; features vector extraction and classification.

7.2.1 Features Vectors Extraction

In order to classify the face activity into the fatigue levels, the features vectors need to be extracted from the specific face activities. Two main activities are applied in this research, eyes activities and yawn, which are specifically used to classify the fatigue signs.

7.2.1.1 Eyes Activities

As discussed in chapters 2 and 3, eyes activities are prominent features in fatigue representation. In chapter 5, a new technique is introduced to measure the eyes'

state in the sequence of frames. The eyes' state detection results are represented in a profile plot as shown in Figure 7.2. From this profile plot the distinctive feature vectors are extracted. The first feature is the Blink Rate (*BR*), which is counted from number of times the plot level is below a threshold value. For example in Figure 7.2, the *BR* is 3, which is detected based on the times the plot level is below the threshold value θ. The θ threshold value is calculated as follows:

$$\theta = 0.2 \max(ia) \tag{7.1}$$

where *ia* is the iris area in pixels. This equation (7.1) is defined based on experiments carried out based on 420 frames of images from 12 videos footage of SFF database and ZJU database.

The second features vector is extracted from the total of time of eye closed (*tec*). Each of blink has *tec,* and the total time of eye closed (T_{tec}) is computed by multiply with number of blink *n* and *m* is maximum number of blink as follows:

$$T_{tec} = \sum_{n=1}^{m} tec_n \tag{7.2}$$

The next feature vector is the average of T_{tec}, where the T_{tec} is divided by blink rate and the average of the total time of eye closed A_{tec} is computed as follows:

$$A_{tec} = \frac{\sum_{n=1}^{m} tec_n}{BR} \tag{7.3}$$

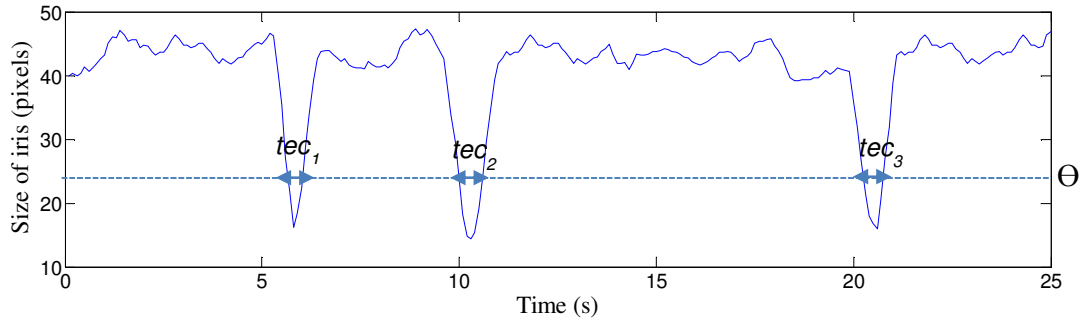


Figure 7.2: Eye state profile plot for three blinks. T_{ec} is time of eye closed

The detection and analysis of fatigue signs is a continuous process where the prior condition is essential to take into account. In order to increase the reliability of the fatigue features vectors, each of the three abovementioned feature vector of the accumulative values is computed. The accumulative value of blink rate $A_{BR,}$ total time of eye closed AT_{tec}, and average total time of eye closed AA_{tec} are calculated as follows:

$$A_{BR} = \frac{\sum_{k=1}^{l} BR_k}{l} \tag{7.4}$$

$$AT_{tec} = \frac{\sum_{k=1}^{l} T_{tec_k}}{l} \tag{7.5}$$

$$AA_{tec} = \frac{\sum_{k=1}^{l} A_{tec_k}}{l} \tag{7.6}$$

where k represent number of average computed, and, l denotes the cumulative time.

7.2.1.2 **Yawning Analysis**

The yawn analysis algorithm, discussed in chapter 6, is able to detect the yawn and also the yawn period. Figure 7.4 shows an example plot of yawn analysis that consists of the mouth opening measurement, covered mouth detection, and distortion detection. In this plot, yawn is detected in the situation where the mouth is covered, where facial distortion is detected. For the features vector, the Yawn Detection (*YD*) result is used to indicate '1' for yawn and '0' for not yawn. The second feature vector is the yawn time T_Y and it is computed as a normalized value by:

$$T_Y = \frac{yt}{tm} \tag{7.7}$$

where *yt* is the time yawn is detected as shown in Figure 7.3, and *tm* is duration of plot. Similar to the eye activities, the yawn features vector analysis also takes into account the history of yawn during monitoring. Therefore, cumulative values of yawn detection A_{YD} and yawn time AT_Y are computed as follows:

$$A_{YD} = \frac{\sum_{k=1}^{l} YD_k}{l} \tag{7.8}$$

$$AT_Y = \frac{\sum_{k=1}^{l} T_{Y_k}}{l} \tag{7.9}$$

where *k* represent number of average computed, and, *l* denotes the cumulative time.

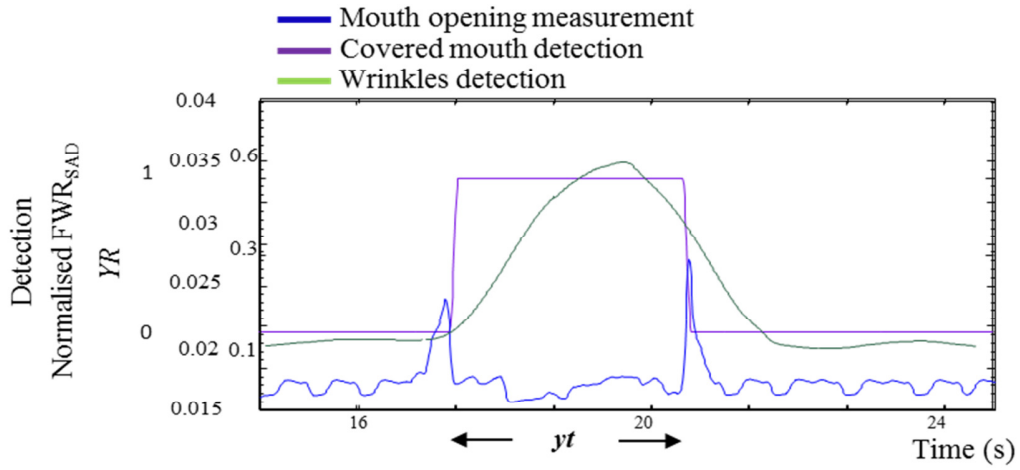


Figure 7.3: Results of three operations; mouth opening measurement, covered mouth detection and wrinkles detection. *yt* represents the time of yawning. This result is based on situation yawn is detected when mouth covered.

7.2.2 **Fatigue Classification**

In order to classify the face activities into levels of sleep deprivation, a set of features vectors need to be extracted from specific face activities. These features vectors are then trained using a machine learning classifier algorithm. Figure 7.4 shows a flowchart of the operation which is start with features vector extraction from plot of eye activity of mouth activity. The features extractor then generates the features vectors. These features vector are used to train the Neural Network (NN) classifier. The output of NN are the trained parameter that will be used in classification process. The video footages that were used for extracting the features vectors are obtained from the SFF database which represents the face activities for each case of sleep deprivation hours considered.

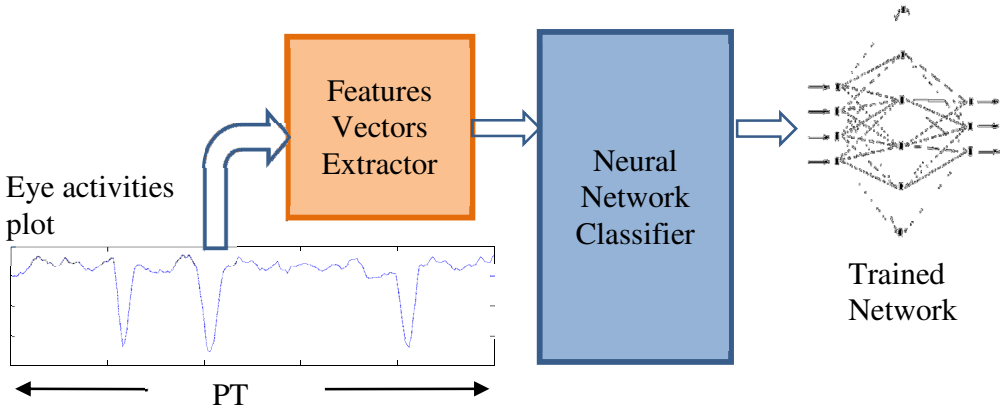


Figure 7.4: Training process where features vector are generated from extraction of eye activities. These features vectors are trained using Neural Network classifier, and the output are the trained parameters that to be used in classification operation.

7.2.2.1 **Features Extraction**

In order to train the classifier, we need to extract the best features vectors from the video footage SFF. Hence, the participants were selected based on eyes' status displaying significant differences between levels. Figure 7.5 shows the blink rate for the participants from the SFF database that was used for training the classifier. Based on the study in [187, 188] , the blink rate is supposed to change proportionally to the level of sleep loss. Some of the participants show significant strong signs of fatigue corresponding to the reading of 8 hours sleep loss where the blink rate drops dramatically because the eyes are closed for long periods of time (i.e. the person is asleep).

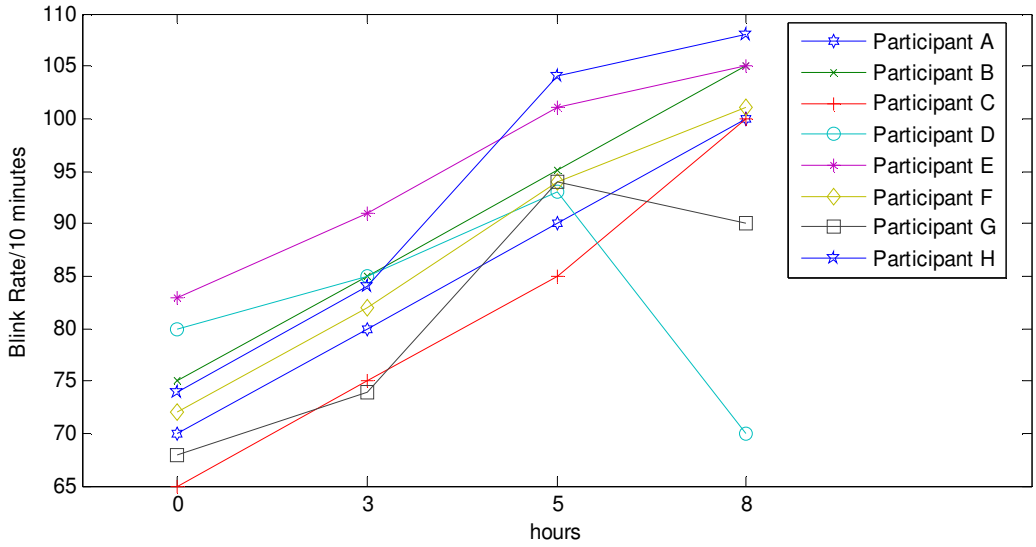


Figure 7.5: Blink rate of the eight participants from the SFF database that was used for training the classifier

7.2.2.2 **Training and Testing**

In this fatigue recognition system the classification is based on four stages of sleep deprivation hours. Since the classification requires making more than two decisions, a Multilayer perceptron Neural Network (MLPNN) was chosen as the classifier. This classifier has the following advantages: works well in learning and generalize the data, has smaller training set requirements, is fast, easy to implement and is highly reliable for multiple classifications [189-191].

The features vectors are extracted from certain time plots (*PT)* from eye activities and yawning analysis plots as shown in Figure 7.6. Initially, an experiment was conducted in order to find the best time of *PT* which involves 5 participants for training and testing. There are three *PT* applied to extract the features vectors, 15seconds, 30 seconds and 1 minute, and the results are shown in table 7.1. From the table with the small amount of dataset, 30 seconds and 1 minute intervals have produced the best results compared to 15 seconds. However, the 30-seconds PT allows more frequent updates of the fatigue status by the classification system.

Table 7.1 Training result for three PT

Plot Time (PT)	**15 seconds**	**30 seconds**	**1minute**
Classification rate	92%	98%	98%

For training the network, eight participants have been chosen for features extraction based on 30-seconds *PT*. A total of 3600 seconds of data were extracted from the participants and the evaluation matrix results for the training and testing process are shown in Figure 7.7. In this evaluation the data are divided into training, test, and validation process. Each data represent features vectors which are extracted from eyes state and yawning detection analysis for 30 seconds. There 318 data are used for testing, 70 for validation and 70 for test. For training, the detection rate is 94.6%, in which the most error is 2.4 %, which misclassification of 0 hour on 3 hours. In validation and test process the detection rate are more than 97%. Overall, the classification shows promising results with 95.4% detection rate with errors mostly occurring from misclassification between 0 and 3 hours and also between 5 and 8 hour

Figure 7.6: The confusion matrix for trained network. The trained parameters are evaluated based training confusion, validation confusion, test confusion and overall confusion performance.

7.2.2.3 **Fatigue levels**

As mentioned earlier the fatigue status level is based on the NN classification result. On the FMT, the status is represented using a gauge meter as shown in Figure C1in Appendix C. To ensure that the meter will not show inconsistent results,, the

fatigue level F_L is represented based on average results, rather than instant evaluations. The fatigue level F_L is calculated as follows:

$$F_L = \frac{PT}{t_a} \sum_{m=1}^{\frac{t_a}{PT}} C_m \tag{7.10}$$

where t_a (in seconds) is the time average to be used, C_m is the NN classification result, and PT is plot time. For example, if the time used for t_a is 3 minutes and PT is 30 seconds, the F_L result is based on averaging the classification result 6 times.

7.3 Experimental Results

Extensive experiments have been carried out to evaluate the performance of the developed fatigue detection algorithms using several video footages selected from the SFF database and video recorded on a shipping bridge simulator. In this research we aimed to apply the FMT to detect fatigue signs on shipping bridge crew members. The shipping bridge crew members are amongst the people that are always affected by the fatigue symptoms (as discussed in Chapter 2). Fortunately, we had an opportunity to carry out some experiments in the Glasgow Nautical College shipping bridge simulator.

Figure 7.7 shows the pictures of the shipping bridge simulator room taken during the experiments. In these experiments two participants were involved, where one of them had been sleep deprived for a whole night and the participant had no sleeping loss. In the simulator room the participants were required to perform standard tasks to manage the movement and direction of the ship for two hours.

As can be seen in Figure 7.7, the shipping simulator has two lightning conditions; switched off lighting room in which the source of lighting is only from the screen that projected the sea images, and the other is switched on lighting room condition where all the lights in the room are switched on. Both lighting condition have

been applied during the experiment in order to test the reliability of the developed algorithms.

Figure 7.7: The experiment was carried out in shipping bridge simulator room which involved two participants.

7.3.1 Validation of Measurement Accuracy

Two main algorithms were developed in order to detect and measure fatigue signs, namely an eye state analysis algorithm and a yawn detection algorithm. For validating the accuracy of these algorithms, there are several short video footages randomly selected from the SFF database. In these experiments, 7,550 frames were used to evaluate the performance of the eye activities and yawn detection. The detection results for each activity of the developed algorithms are compared to the activity observation results. The excellent

performance for the blink detection, eye closure time, yawn detection and yawning time of the developed algorithms are given in Table 7.2. All the experimented detection rates are over than 94%, with blink detection showing the highest accuracy.

Table 7.2 Performance of eyes activities and yawn

Activity	Blink	Eye Closure time	Yawn	Yawn time
Detection Rate	99.8%	95.8%	96%	94.6%

7.3.2 **Classification**

In this research, the fatigue signs are classified based upon sleep deprivation hours. For testing the performance of the developed fatigue monitoring system, the video footage SFF database and the recorded video of the participant involved in the shipping bridge simulator experiment are used. The results of the MLPNN classification for a 10-minute face activities monitoring, selected from the SFF database, are shown in Figure 7.8. This figure shows the classification results for four participants in different sleep deprivation states; 0 hour, 3 hours, 5 hours, and 8 hours respectively (Figure 7.8(a) to Figure 7.8(d)). For each sleep deprivation state, three different average times t_a (7.10): 30 seconds, 1 minute and 2 minutes is computed.

As above-mentioned the average time (7.10) is to avoid inconsistence classification result on gauge meter. Based on these results, two minutes average time produce the consistence classification result compared to the classification based 30 seconds time average. A consistence classification indication is important to prevent the fatigue level classification indicator fluctuation. Figure 7.9 shows the classification results based on a two minute average in ten minutes of monitoring. Figure 7.9a) shows the classification results for 0 and 3 hours sleep deprivation video footage, while the Figure 7.9(b) shows the results for 5 and 8 hours.

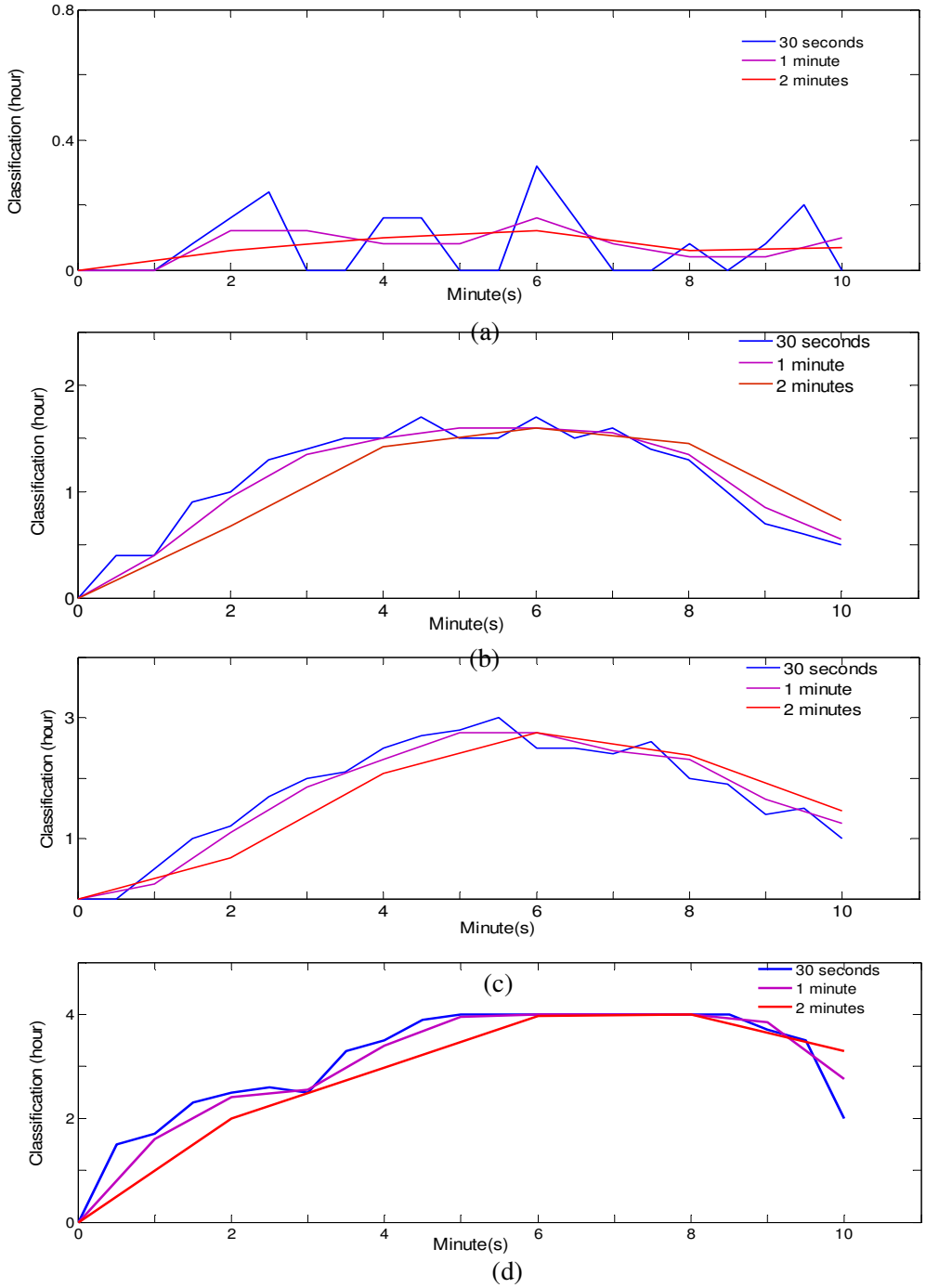
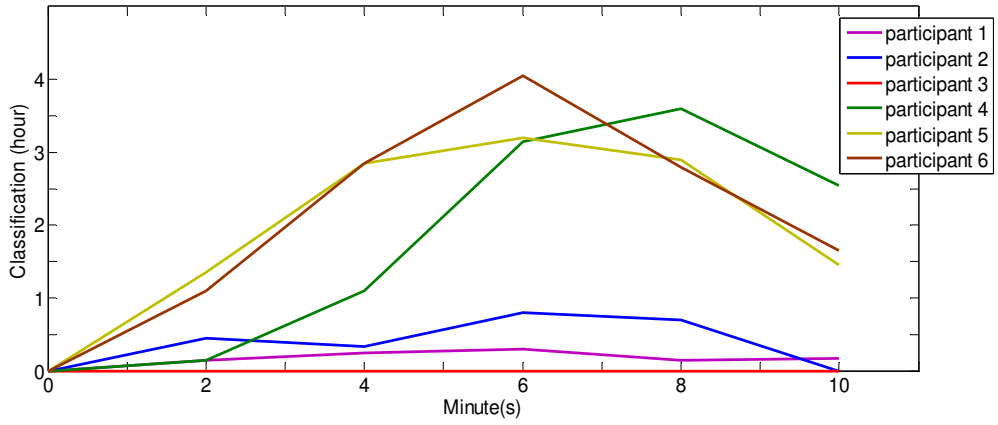
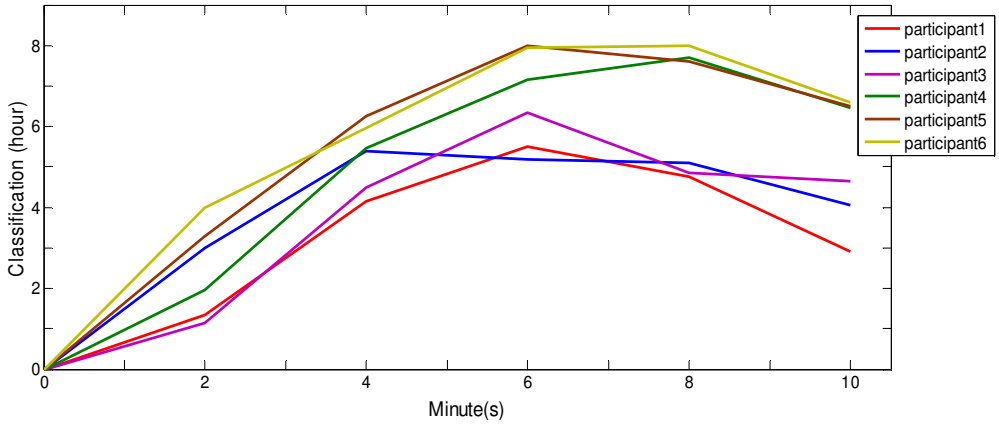


Figure 7.8: Fatigue classification results for 10 minutes from SFF video database; (a) 0hour, (b) 3 hours, (c) 5 hours, and (d) 8 hours sleep deprivation.



(a)



(b)

Figure 7.9: Classification results in 10 minutes monitoring; (a) for 0 hour and 3 hours, (b) 5 hours and 8 hours sleep deprivation video footages.

7.3.3 **Proposed Technique Vs PERCLOS**

PERCLOS is a prominent technique to detect fatigue. This technique, discussed in detail in Chapter 3, is based on detection of 80% of eye closure over a specific period

of time. In this experiment we compare the result of PERCLOS technique with our proposed technique using the video footage of the participants in the shipping bridge simulator experiment as shown in Figure 7.10. This result is based on two participants who conducted the shipping crew activities for 1hour and 20 minutes. One of the participants was sleep deprived for a whole night and the other was not sleep deprived. From Figure 7.10, for the proposed technique, the classification result for participant with 8 hours sleep deprived, it can be seen clearly climbing from 0 hour until a peak of fatigue at 8 hours classification. While for non-sleep deprived participant, the proposed technique indicate 1 hour as maximum of classification of fatigue.

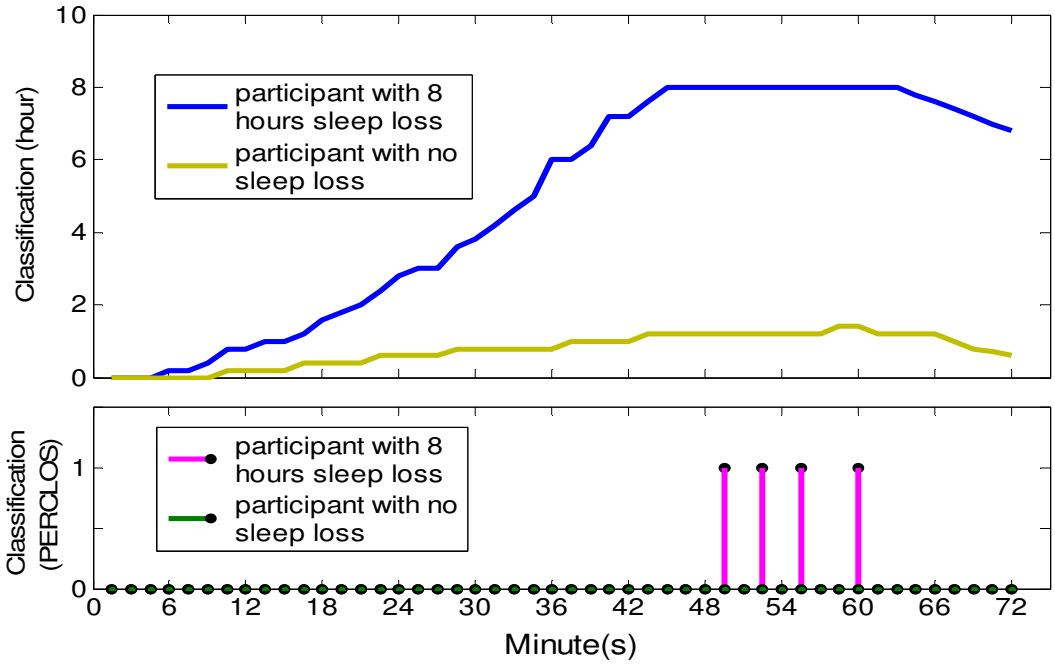


Figure 7.10: Proposed technique Vs PERCLOS results. The proposed technique produced the output based on level of fatigue signs, whilst PERCLOS based classification which result either detected of not.

For PERCLOS technique classification, no fatigue is detected for non-sleep deprived participant. However, four times fatigue is detected ('1' is represented detection in PERCLOS technique in Figure 7.10) for the sleep deprived participant.

Obviously, in the proposed technique, which is a combination of activities measurement of the eye and mouth, the fatigue signs are represented progressively in accordance with the progress of fatigue levels as they occur. By comparison, the PERCLOS technique, which only monitors the eye activities, the fatigue is detected only when the eye closure time reaches the specified percentage. The details analysis result of FMT that carried upon several SFF participants can be referred in Appendix D.

7.4 Conclusion

Chapter 7 has discussed a fatigue recognition technique using the Fatigue Monitoring Tool (FMT) to classify the facial fatigue signs. A new technique is introduced in which the fatigue signs are classified into certain levels based upon sleep deprivation hours. In this technique the features vectors are extracted from the eyes activities and the yawn detection. From these two components, ten features vectors are extracted for every thirty seconds of face activities. The extracted features vectors are then trained using a classifier. The MLPNN classifier is used in this research, and eight participants and 120 minutes of videos footage have been used for training the features vectors.

The fatigue recognition algorithm classifies the fatigue signs every thirty seconds, and the status of fatigue is displayed on the FMT based on the average calculated values. From the conducted experiments, the developed fatigue recognition indicates excellent classification results. Comparing the proposed technique and the prominent fatigue detection technique PERCLOS, the proposed technique is able to discern the flow of changes of fatigue signs compared to PERCLOS which is based on a single classification.

8. Conclusion and Future Work

8.1 Conclusion

This thesis has presented novel work of non-invasive fatigue detection and quantification algorithms. There are four novel algorithms which have been developed and discussed. One contribution for database development is also explained in this thesis.

Starting with the face acquisition operation, a novel approach of face detection algorithm was introduced, in which the algorithm is a combination of skin colour segmentation, connected component binary image formation and the Viola Jones classifier. Based on conducted experiments the proposed face detection algorithm has shown very good performance in which the false positive detection rate was reduced by a very significant margin.

Eye state is essentially measured eye activity which can be used to represent the dominant signs of fatigue. This thesis has introduced a new technique to measure eye activities. This is the Interdependence and Adaptive Scale Mean Shift (IASMS) algorithm, which is an association of mean shift tracking algorithm and an adaptive scale scheme. IASMS is integrated with a face detection algorithm, an image enhanced scheme, an eye open detection technique and an iris detection method in order to be able measure the eye state. Based on the conducted experimental results this proposed method successfully quantifies the eye states that denote blink rate and duration of eye closure.

Commonly, yawn is registered based on a wide mouth opening detection. This thesis has enhanced yawn detection techniques by introducing a new yawn analysis algorithm which takes into account yawning whilst the mouth is hand-covered; a very common and spontaneous human action. The algorithm incorporates a new mouth

opening measurement approach, a mouth covered detection, and a facial distortion (wrinkles) detection. Based on the experiments carried out, the proposed algorithm produces a very good yawn detection rate. This algorithm is able to analyse yawn in the three situations; when the mouth is widely open, when the mouth is wide open for a short period of time before is promptly covered, and when the mouth is fully covered.

Since a facial fatigue database was not available to be used for research purposes, this PhD research initiative has developed a video footage database that contains facial activities associated with signs of fatigue. In order to induce the genuine facial fatigue signs, sleep deprivation experiments were conducted, which involved twenty volunteering participants. This experiment was a collaborative work between CeSIP (University of Strathclyde), PsyKE (University of Strathclyde), and Glasgow Sleep Centre. In the experiments, the participants were required to go through four sessions of sleep deprivation of, 0, 3, 5 and 8 hours, in which the participant carried out different cognitive tasks. This database has been used for training, testing and evaluation for the yawning analysis and fatigue recognition algorithms.

Finally this thesis has presented a novel fatigue recognition algorithm that is able to detect and quantify the signs of fatigue into stages. This algorithm has been integrated with the Fatigue Monitoring Tool (FMT) platform in order to allow the implementation of the developed fatigue detection and quantification algorithms in a real-time monitoring system. The FMT classifies fatigue signs based on the sleep deprivation hours. The SFF database has been fully utilized for training, testing and also evaluation of the fatigue recognition algorithm. The FMT has also been tested on shipping crew members whilst carrying out routine shipping tasks by using a shipping bridge simulator. The conducted experiments have shown good results in terms of fatigue signs detection and classification.

8.2 Future Work

Based on the experiments carried out to develop a non-invasive facial fatigue detection and quantification algorithms, and the experience acquired from this work, there are some suggestions which could help to improve the performance and the robustness of the system. These are:

1. In this research the developed eye measurement algorithms were tested and evaluated based upon normal lighting conditions. In order to allow the developed system to work well under any lighting condition (including night time), the developed algorithm can be improved by testing and evaluating its performance using infrared video recordings.

2. Similarly, in this research the yawn is only analysed and tested under normal lighting conditions. For robustness, all the lighting conditions must really be considered.

3. In this research the eyes and mouth are the main components to be monitored. In future the algorithm development has to take in to account cases where monitored users wear sun glasses;. Other elements in human behaviour, such as head nodding, need to also be detected. In such cases though the challenge will be to distinguish between normal and fatigue induced behaviour. Other aspects that potentially can be measured are movements of the face and hand-face interactions at specific times. Commonly, people who begin to feel fatigued, start to touch and rub parts of their face. They also distort their facial characteristics in an attempt to alleviate fatigue.

4. In order to enable the fatigue system to monitor multiple of operators as, for example, in a ship bridge, additional elements, such as facial recognition, should be included in the system, so that unique identification results of facial activities can be obtained.

5. Instead of face component which able to indicate the signs of fatigue, hand gesture can be considered to be measured as fatigue sign. During fatigue people use hand to rub their face in order to restrain fatigue from getting worse. Therefore, another algorithm can be developed to measure and identify hand gesture that present the fatigue reaction

Appendices

Appendix A

Features

The Haar-like features use the intensity values of a pixel, and the change in contrast values between adjacent rectangular groups of pixels. The contrast variances between the pixel groups are used to determine relative light and dark areas. Two or three adjacent groups with a relative contrast variance form a Haar-like feature. Haar-like features, as shown in Figure A.1 are used to detect an image. Haar features can easily be scaled by increasing or decreasing the size of the pixel group being examined. This allows features to be used to detect objects of various sizes.

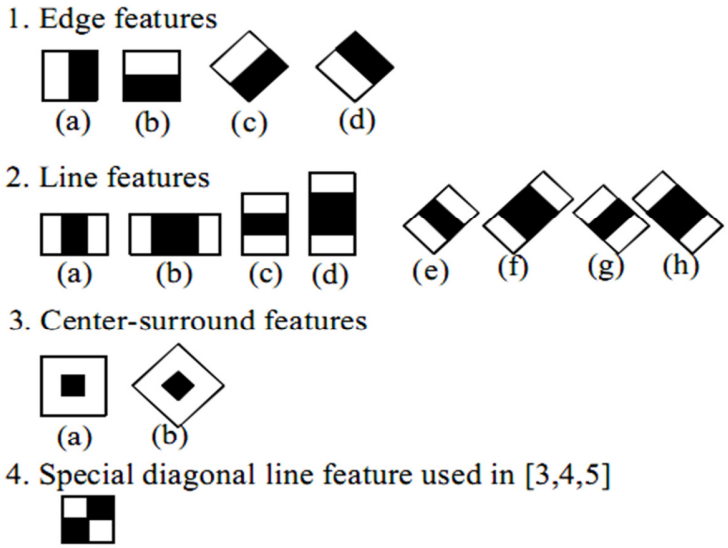


Figure A.1: Haar-like features

Integral Image

Rectangle features can be computed very rapidly using an intermediate representation for the image which we call the integral image. The integral image at location *x, y* as shown in Figure A.2 contains the sum of the pixels above and to the left of *x, y*, inclusive.

$$ii(x, y) = \sum_{x' \le x, y' \le y} i(x', y') \tag{A.1}$$

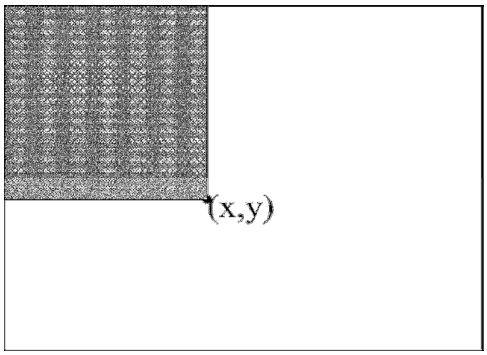


Figure A.2: The value of the integral image at point (*x, y*) is the sum of all the pixels above and to the left.

where *ii* (*x, y*) is the integral image and *i* (*x, y*) is the original image. Using the following pair of recurrences:

$$s(x, y) = s(x, y-1) + i(x, y) \tag{A.2}$$

$$ii(x, y) = ii(x-1, y) + s(x, y) \tag{A.3}$$

where *s*(*x, y*) is the cumulative row sum, *s*(*x*,−1) =0, and *ii* (−1, *y*) = 0. Using the integral image any rectangular sum can be computed in four array references as shown in Figure

A.3. The sum of the pixels within rectangle *D* can be computed with four array references. The value of the integral image at location 1 is the sum of the pixels in rectangle *A*. The value at location 2 is *A* + *B*, at location 3 is *A* + *C*, and at location 4 is *A* + *B* + *C* + *D*.

The sum within *D* can be computed as 4 + 1 − (2 + 3).

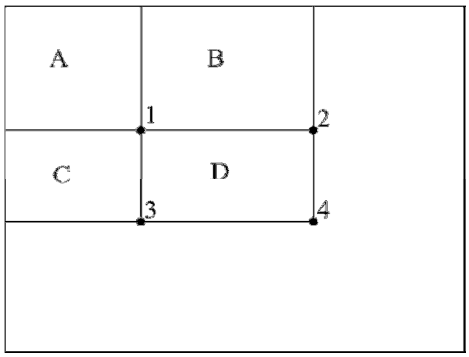


Figure A.3: Example four array reference points to compute the integral image.

Cascade classifier

The Cascade classifier technique is able to increase the detection performance while radically reducing computation time. Viola and Jones choose the Adaboost classifier because simple and more efficient. The classifier in this technique should be simpler classifiers, which are used to reject the majority of sub-windows before more complex classifiers are called upon to achieve low false positive rates (as in Figure A.4)

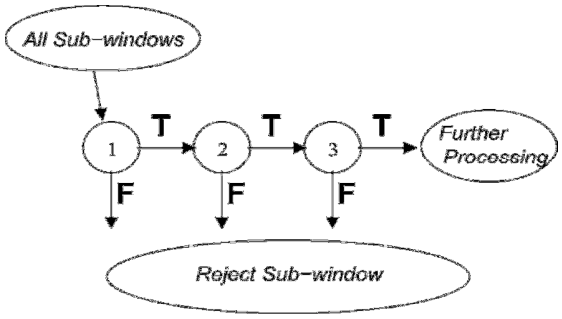


Figure A.4: Cascade classifier which series of classifier connected together.

Training Cascade Classifier

The cascade design process is driven from a set of detection and performance goals. For example in the face detection task, past systems have achieved good detection rates (between 85 and 95 percent) and extremely low false positive rates (on the order of 10−5 or 10−6). The number of cascade stages and the size of each stage must be sufficient to achieve similar detection performance while minimizing computation. Given a trained cascade of classifiers, the false positive rate of the cascade is:

$$F = \prod_{i=1}^{K} f_i \tag{A.4}$$

where F is the false positive rate of the cascaded classifier, K is the number of classifiers, and fi is the false positive rate of the ith classifier on the examples that get through to it. The detection rate is:

$$D = \prod_{i=1}^{K} d_i \tag{A.5}$$

where D is the detection rate of the cascaded classifier, K is the number of classifiers, and di is the detection rate of the ith classifier on the examples that get through to it.

Given concrete goals for overall false positive and detection rates, target rates can be determined for each stage in the cascade process. For example a detection rate of 0.9 can be achieved by a 10 stage classifier if each stage has a detection rate of 0.99 (since $0.9 \approx 0.9910$). While achieving this detection rate may sound like a daunting task, it is made significantly easier by the fact that each stage need only achieve a false positive rate of about 30% ($0.3010 \approx 6 \times 10-6$).

The number of features evaluated when scanning real images is necessarily a probabilistic process. Any given sub-window will progress down through the cascade,

one classifier at a time, until it is decided that the window is negative or, in rare circumstances, the window succeeds in each test and is labelled positive. The expected behavior of this process is determined by the distribution of image windows in a typical test set. The key measure of each classifier is its "positive rate", the proportion of windows which are labelled as potentially containing a face. The expected number of features which are evaluated is:

$$N = n_0 + \sum_{i=1}^{K} \left(n_i \prod_{j<i} P_j \right) \tag{A.6}$$

where N is the expected number of features evaluated, K is the number of classifiers, pi is the positive rate of the ith classifier, and ni are the number of features in the ith classifier. Interestingly, since faces are extremely rare, the "positive rate" is effectively equal to the false positive rate. The summary of cascade classifier training algorithm as described in Table A.1

Table A.1 The training algorithm for building a cascaded detector

User selects values for f, the maximum acceptable false positive rate per layer and d, the minimum acceptable detectionrate per layer.
• User selects target overall false positive rate, *Ftarget* .
• P = set of positive examples
• N = set of negative examples
• $F0 = 1.0$; $D0 = 1.0$
• $i = 0$
• while $Fi > Ftarget$
– $i \leftarrow i + 1$
– $ni = 0$; $Fi = Fi-1$
– while $Fi > f \times Fi-1$
 $ni \leftarrow ni + 1$
 Use P and N to train a classifier with ni features using AdaBoost
 Evaluate current cascaded classifier on validation set to determine Fi and Di .
 Decrease threshold for the ith classifier until the current cascaded classifier has a detection rate of at least
$d \times Di-1$ (this also affects Fi)
– $N \leftarrow$
– If $Fi > Ftarget$ then evaluate the current cascaded detector on the set of non-face images and put any false detections into the set N

Appendix B

Data acquisition

The sleep deprivation experiments were conducted by PsyKE and the Glasgow Sleep Centre, while CeSIP provided the technical facilities for image and video acquisition during the experiments. Due to the experiments involving the monitoring of human sleep activities the experimentswere designed to fulfil the specified requirements set out by the University Ethics Committee (UEC) of the University of Strathclyde. Twenty people, ten male and ten female with ages ranging between twenty to forty years old, participated in these ethically approved experiments. In order to ensure that the participants complied with the number of hours required to deprive their sleep an Actiwatch was used as shown in Figure 4.5. Actiwatch is an actigraphy- based device that measures the gross motion activity which allows quantification of physical activities and sleep.

Figure B1 Actiwatch device that used by participant for record the sleeping time. [55].

Experiments Procedure

During the sleep deprivation experiments each participant had to go through four experimental sessions. In each session the participants were sleep deprived for 0 hour (no sleep deprivation), 3, 5and 8 hours respectively. The participants carried out each experimental session in different weeks for four consecutively weeks and for different hours of sleep deprivation. During each experimental session the participants performed several cognitive tasks and the complete of schedule of each experimental session is

described in Table 4.3. The schedule represents the sequence of tasks that participants had to undergo.

Table B1 Schedule of sleep deprivation experiment.

	Task	Type of task	Time
1.	Sustained Attention to Response Task (SART)	Video recorded	(12 – 18) minutes
2.	Psychomotor vigilance task (PVT)	Video recorded	(12 – 18) minutes
3.	Pro & Anti saccade task with emotional faces (measures inhibition)	Eye tracking task	(30-40) minutes
4.	Emotional face photo discrete categorisation task	Video recorded	(5-10) minutes
5.	Dynamic video social gesture dimensional categorisation task	Video recorded	(10-15) minutes
6.	Boredom task	Video recorded	(10-15)minutes
7.	IAPS categorisation task	Eye tracking task	(30-40)minutes

Referring to **Error! Reference source not found.** the task types were divided into video recorded and eye tracking tasks. In video recorded tasks, the facial activities are recorded throughout the tasks, while in eye tracking tasks the facial activities are not recorded explicitly due to the fact that tasks involving the use of eye tracking equipment also capture parts of the face. Hence, the SFF database is developed based on five conducted cognitive tasks. These tasks are carried out in standard room lighting conditions with three video cameras used to record facial activities. The distance in between cameras is 120cm, and one of camera is located in middle to record front face activity. The high of cameras from the floor is 120cm and distance between camera and participant are approximately in range 110cm to 130cm. **Error! Reference source not**

found. shows the room used to carry out the five video recorded cognitive tasks. The cameras used were high definition cameras in the range 720 to 1080 pixels per frame, recording in PAL format.

Figure B2 The experiment room in normal lighting condition which has three cameras located 120cm in between each other. The high of camera is 120cm from the floor, and distance between camera and participant are approximately in range 110cm to 130cm. *(Images are permitted to be published)*

Experiment Tasks

The cognitive experimental tasks were conducted by PsyKE. The cognitive tasks were designed to test simple attention of the participants, sustained attention and their working memory. These cognitive tasks are associated with sensitivity to sleep loss and are able accelerate fatigue signs. The five tasks, during which recording of facial activities took place were:

A. Sustained Attention to Response Task (SART)

The Sustained Attention to Response Task is a computational based task designed by Robertson et al [187] to measure a person's ability to respond to infrequent and unpredictable stimuli during a period of rapid and rhythmic presentation of stimuli. In this task, the participants were asked to identify the characters that were presented rapidly on a computer screen (as shown in Figure B3). They had to press a key every time they saw any number presented rapidly on a computer screen but withhold a response to a specified low frequency digit.

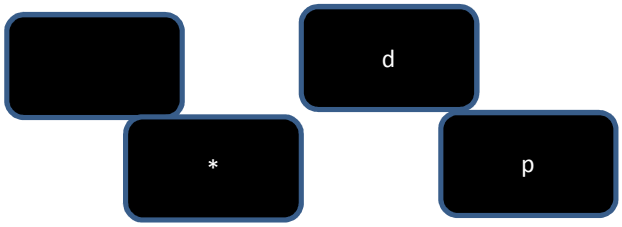


Figure B.3 Example of letters and symbols on screen that is measured in SART

B. Psychomotor Vigilance Task (PVT)

The Psychomotor Vigilance Task (PVT) was originally designed in 1985 [188] to measure sustained attention. This task has shown to be sensitive to sleepiness in clinical, experiment, and operational contexts, making it one of the most widely used in neurobehavioral tests, studies of sleep and in circadian rhythm research [187]. In this experiment, the participants required to identify the amount of LEDs displayed at a certain time. The accuracy of reading is highly correlated with sleep deficiency in terms of increased reaction time, reduced vigilance and performance variability.

C. Emotional face photo discrete categorisation task

In this task the participants performed face recognition tasks of increasing emotional intensity (as shown B.4), in which they had to evaluate several different affective face categories such as happy, sad, angry, and disgust. Based on study in [188], Helm et. al. found that sleep deprivation selectively impairs the accurate judgment of human facial emotion. This finding suggests that sleep loss discrete affective neural system which disrupts the identification of salient effective social cues. .

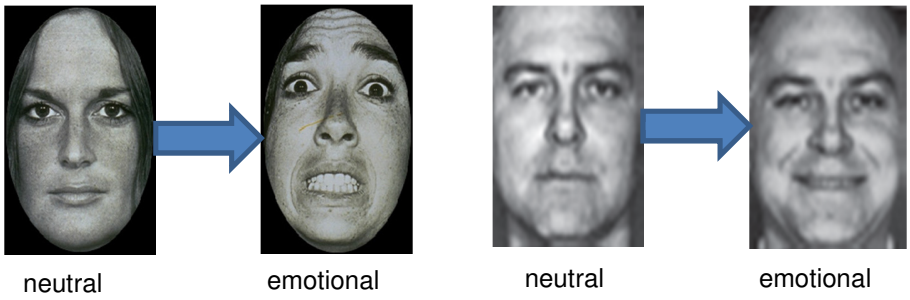


Figure B4 Increasing intensity of emotion

D. Dynamic video social gesture dimensional categorisation task

In this task the participant is required to categories the social gesture which is played on a monitor screen.

E. Boredom task

In this boredom task the participants were asked to track a moving light point and estimate the frequency of cycles across the screen as shown in Figure B5.

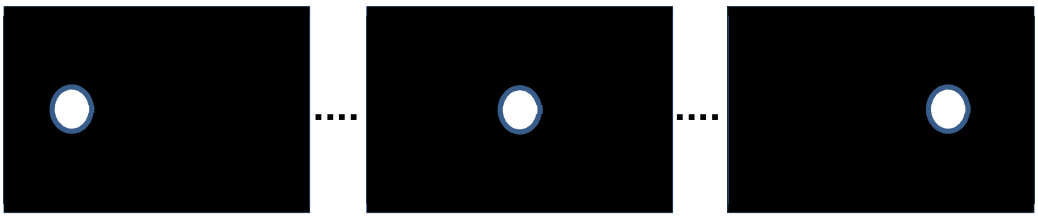


Figure B.5 Tracking moving light in boredom task

Characteristics of Database

The results of the cognitive tasks experiments, from the randomly selected participant, show that the numbers of errors occurring for attention tasks and psychomotor vigilance tasks are significantly increased when participant lost their sleep for five or eight hours (Figure B.6 and Figure B.7). These results indicate that the experiments successfully associated competence of task completion to sensitivity of sleep loss, and demonstrated the acceleration of fatigue signs related to this sleep loss.

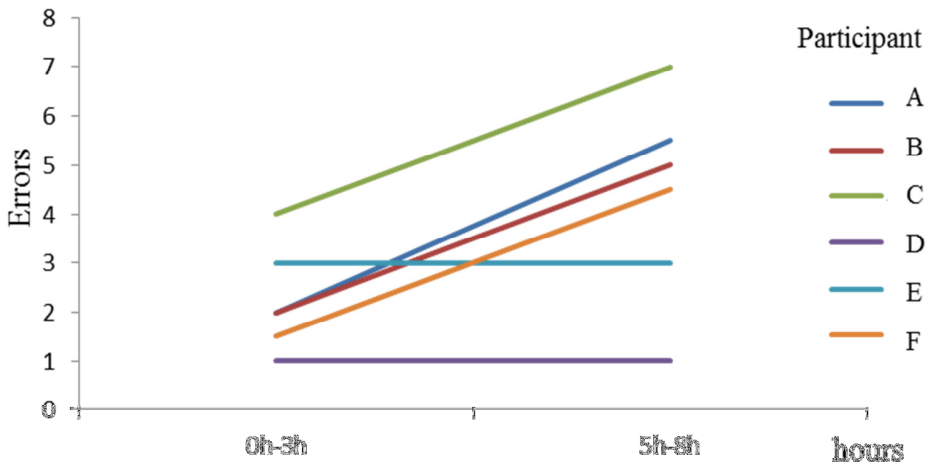


Figure B.6 Attention task errors

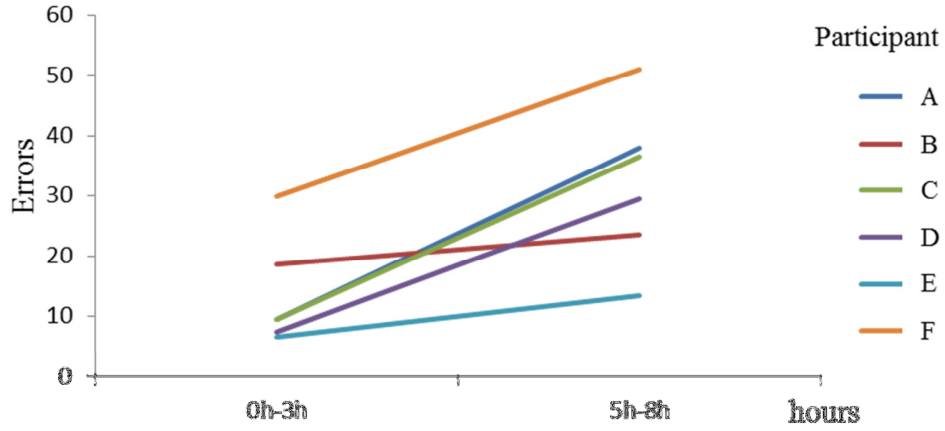


Figure B.7 PVT mean performance errors

Appendix C

Fatigue Monitoring Tool (FMT)

The Fatigue Monitoring Tool (FMT) platform, including the graphical user interface (GUI), was created by CeSIP group members who participated in the project. My part in this work was the algorithm development. The tool was designed with a GUI (as shown in Figure C1) which can be installed in any computer with a Windows XP or above operating system. The tool is able to monitor facial expressions and perform real-time processing by using any computer attached camera as well as a network camera. The tool also provides the storage that allows recording the monitoring activities including the result analysis. The monitored face is always analysed in predetermined time slots based on the developed algorithms. The current reading level of fatigue is indicated in a meter gauge as shown in Figure 7.2, and the average reading results of the last 5 hours can also be viewed in this tool.

The detail features of the FMT, with reference to the number labels, as shown in ,Figure C1 are:

1. *Source selector*: The FMT provides options to select the source input to be used in the analysis. This can be a network camera, a web camera or a stored pre-recorded video.

2. *Video recording option*: Users are allowed to select whether the video under analysis is to be recorded or not.
3. *Display the face monitoring*: Display the face which is being analysed.
4. *Control button*: Button to start, stop or reset the tool.
5. *Analysis recording option*: User is allowed to select whether the analysis results are to be recorded or not.
6. *Analysis parameter setup*: Certain analysis parameters such as eye activities plots and mouth open detection can be selected.
7. *Algorithm selection*: Users are allowed to select the algorithm to be used for the analysis.
8. *Out profile plot*: Show the current profile analysis plot.
9. *Fatigue level meter*: Show the current status of fatigue.
10. *Fatigue level history*: The history of fatigue levels.

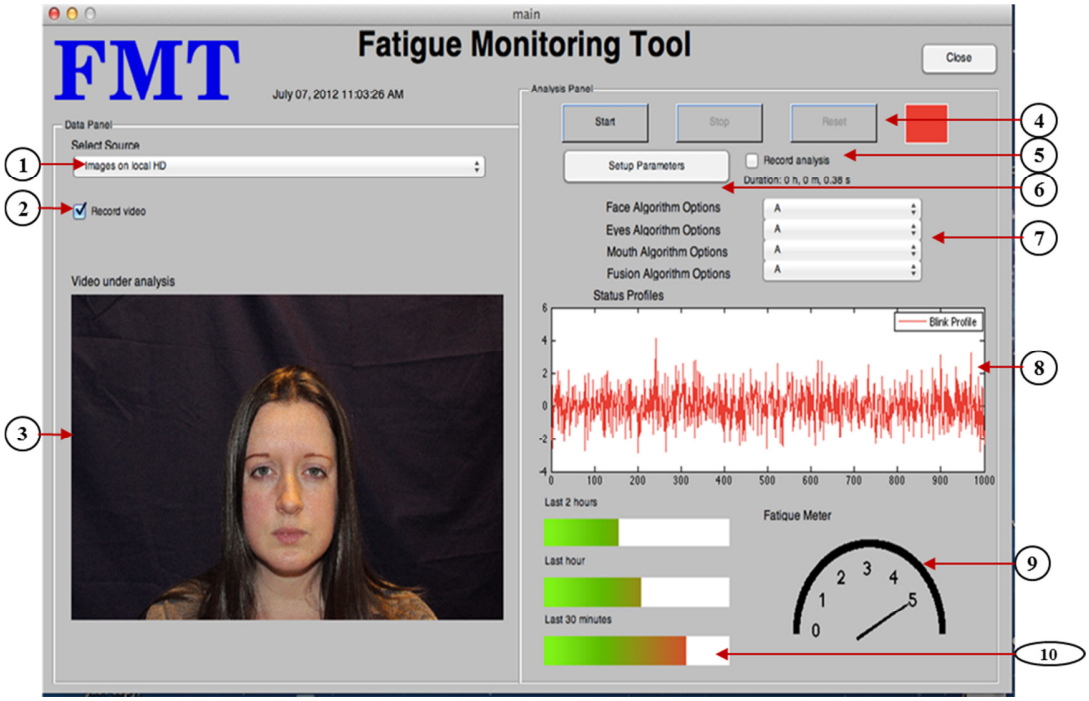
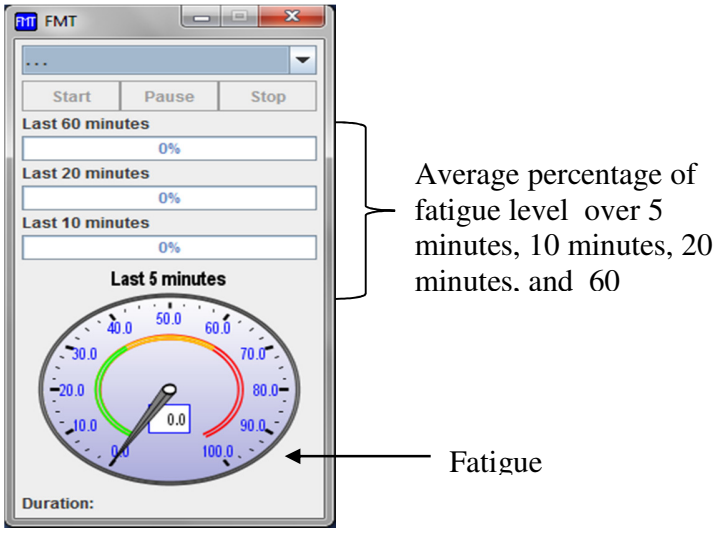


Figure C.1: Graphic user interface of FMT

Appendix D

Fatigue Monitoring Tool Analysis

Fatigue monitoring tool analysis indicates the result of average percentage of fatigue level in duration of 5 minutes, 10 minutes, 20 minutes, and 60 minutes. When the fatigue level reach read zone, which is 65% and above, isAlarm will indicate 'alarm' in analysis result.



Fatigue Monitoring Tool Analysis
Date: 22 Jun 2013, 13:39:26
Video file associated: analysis_1358861966499.mp4

time	60min	20min	10min	5min	isAlarm
0	0%	0%	0%	0%	
00:00:00.047	0%	0%	0%	0%	
00:00:15.055	0%	0%	0%	0%	
00:00:30.155	0%	0%	0%	0%	
00:00:45.022	0%	0%	0%	0%	
00:01:00.076	0%	0%	0%	0%	
00:01:15.052	0%	0%	0%	0%	
00:01:30.044	0%	0%	0%	0%	
00:01:45.020	0%	0%	0%	0%	
00:02:00.995	0.8%	2.5%	5%	10%	
00:02:15.223	0.8%	2.5%	5%	10%	
00:02:30.043	0.8%	2.5%	5%	10%	
00:02:45.518	1.7%	5%	10%	20%	
00:03:00.010	1.7%	5%	10%	20%	
00:03:15.064	2.5%	7.5%	15%	30%	
00:03:30.056	2.5%	7.5%	15%	30%	
00:03:45.126	3.3%	10%	20%	40%	
00:04:00.149	3.3%	10%	20%	40%	
00:04:15.593	3.6%	10.9%	21.9%	43.8%	
00:04:30.101	3.6%	10.9%	21.9%	43.8%	
00:04:45.108	4%	11.9%	23.8%	47.5%	
00:05:00.037	4%	11.9%	23.8%	47.5%	
00:05:15.122	4.3%	12.8%	25.6%	51.2%	
00:05:30.161	4.3%	12.8%	25.6%	51.2%	
00:05:45.121	4.6%	13.8%	27.5%	55%	
00:06:00.113	4.6%	13.8%	27.5%	55%	
00:06:15.136	4.9%	14.7%	29.4%	58.7%	
00:06:30.845	4.9%	14.7%	29.4%	58.7%	
00:06:45.571	5.2%	15.6%	31.2%	62.5%	
00:07:00.064	5.2%	15.6%	31.2%	62.5%	
00:07:30.031	5.5%	16.6%	33.1%	56.2%	
00:07:45.818	5.8%	17.5%	35%	50%	
00:08:00.061	5.8%	17.5%	35%	50%	
00:08:18.891	6.1%	18.4%	36.9%	43.8%	

Fatigue Monitoring Tool Analysis
Date: 21 Jun 2013, 10:52:14
Video file associated: analysis_1358765534048.mp4

time	60min	20min	10min	5min	isAlarm
0	0%	0%	0%	0%	
00:00:00.040	0%	0%	0%	0%	
00:00:05.000	0%	0%	0%	0%	
00:00:15.000	0%	0%	0%	0%	
00:00:30.000	0%	0%	0%	0%	
00:00:45.000	0%	0%	0%	0%	
00:00:45.040	0%	0%	0%	0%	
00:01:00.000	0%	0%	0%	0%	
00:01:15.000	0%	0%	0%	0%	
00:01:30.000	0%	0%	0%	0%	
00:01:45.400	0%	0%	0%	0%	
00:02:00.000	0%	0%	0%	0%	
00:02:15.000	0%	0%	0%	0%	
00:02:30.000	0%	0%	0%	0%	
00:02:45.200	0%	0%	0%	0%	
00:03:00.200	0%	0%	0%	0%	
00:03:15.040	0%	0%	0%	0%	
00:03:30.000	0%	0%	0%	0%	

00:03:45.040	0.8%	2.5%	5%	10%
00:04:00.000	0.8%	2.5%	5%	10%
00:04:15.720	0.8%	2.5%	5%	10%
00:04:30.320	0.8%	2.5%	5%	10%
00:04:45.000	0.8%	2.5%	5%	10%
00:05:00.280	0.8%	2.5%	5%	10%
00:05:15.200	0.8%	2.5%	5%	10%
00:05:30.640	0.8%	2.5%	5%	10%
00:05:45.120	0.8%	2.5%	5%	10%
00:06:00.600	0.8%	2.5%	5%	10%
00:06:15.800	0.8%	2.5%	5%	10%
00:06:30.200	0.8%	2.5%	5%	10%
00:06:45.520	0.8%	2.5%	5%	10%
00:07:00.360	0.8%	2.5%	5%	10%
00:07:15.640	0.8%	2.5%	5%	10%
00:07:30.120	0.8%	2.5%	5%	10%
00:07:45.320	0.8%	2.5%	5%	10%
00:08:00.000	0.8%	2.5%	5%	10%
00:08:15.200	0.8%	2.5%	5%	10%
00:08:30.040	0.8%	2.5%	5%	10%
00:08:45.080	0.8%	2.5%	5%	0%
00:09:00.080	0.8%	2.5%	5%	0%
00:09:15.600	0.8%	2.5%	5%	0%
00:09:30.480	0.8%	2.5%	5%	0%
00:09:45.320	0.8%	2.5%	5%	0%
00:09:58.560	0.8%	2.5%	5%	0%

Fatigue Monitoring Tool Analysis
Date: 22 Jun 2013, 10:39:11
Video file associated: analysis_1358851151104.mp4

time	60min	20min	10min	5min	isAlarm
0	0%	0%	0%	0%	
00:00:00.047	0%	0%	0%	0%	
00:00:15.741	0%	0%	0%	0%	
00:00:30.795	0%	0%	0%	0%	
00:00:45.740	0%	0%	0%	0%	
00:01:00.778	0%	0%	0%	0%	
00:01:15.645	0%	0%	0%	0%	
00:01:30.059	0%	0%	0%	0%	
00:01:45.192	0%	0%	0%	0%	
00:02:00.776	0.8%	2.5%	5%	10%	
00:02:15.503	0.8%	2.5%	5%	10%	
00:02:30.947	1.7%	5%	10%	20%	
00:02:45.470	1.7%	5%	10%	20%	
00:03:00.618	2.5%	7.5%	15%	30%	
00:03:15.766	2.5%	7.5%	15%	30%	
00:03:30.289	2.5%	7.5%	15%	30%	
00:03:45.936	2.5%	7.5%	15%	30%	
00:04:00.647	2.5%	7.5%	15%	30%	
00:04:15.249	2.5%	7.5%	15%	30%	
00:04:30.724	2.5%	7.5%	15%	30%	
00:04:45.294	2.5%	7.5%	15%	30%	
00:05:00.676	3.3%	10%	20%	40%	
00:05:15.215	3.3%	10%	20%	40%	
00:05:30.004	3.3%	10%	20%	40%	
00:05:45.292	4.2%	12.5%	25%	50%	
00:06:00.471	4.2%	12.5%	25%	50%	
00:06:15.306	5%	15%	30%	60%	
00:06:30.860	5.8%	17.5%	35%	70%	alarm
00:06:45.976	5.8%	17.5%	35%	70%	alarm
00:07:00.305	5.8%	17.5%	35%	70%	alarm
00:07:15.845	6.7%	20%	40%	70%	alarm
00:07:30.864	6.7%	20%	40%	70%	alarm
00:07:45.036	7.5%	22.5%	45%	70%	alarm
00:08:00.471	7.5%	22.5%	45%	70%	alarm

00:08:15.840	8.3%	25%	50%	70%	alarm
00:08:30.346	8.3%	25%	50%	70%	alarm
00:08:45.360	9.2%	27.5%	55%	80%	alarm
00:09:00.204	9.2%	27.5%	55%	80%	alarm
00:09:15.152	10%	30%	60%	90%	alarm
00:09:30.081	10%	30%	60%	90%	alarm
00:09:45.225	10.8%	32.5%	65%	100%	alarm
00:10:00.085	10.8%	32.5%	65%	100%	alarm
00:10:00.226	10.8%	32.5%	65%	100%	alarm
00:10:00.350	10.8%	32.5%	65%	100%	alarm
00:10:00.500	10.8%	32.5%	65%	100%	alarm
00:10:15.931	11.1%	33.4%	66.9%	93.8%	alarm
00:10:30.619	11.1%	33.4%	66.9%	93.8%	alarm
00:10:45.258	11.5%	34.4%	68.8%	87.5%	alarm
00:11:00.363	11.5%	34.4%	68.8%	87.5%	alarm
00:11:15.119	11.8%	35.3%	70.6%	81.2%	alarm
00:11:30.226	11.8%	35.3%	70.6%	81.2%	alarm
00:11:45.237	12.6%	37.8%	75.6%	81.2%	alarm
00:12:00.292	12.6%	37.8%	75.6%	81.2%	alarm
00:12:30.170	13.4%	40.3%	75.6%	81.2%	alarm
00:12:45.760	14.3%	42.8%	75.6%	81.2%	alarm
00:13:00.897	14.3%	42.8%	75.6%	81.2%	alarm
00:13:15.550	14.6%	43.8%	72.5%	75%	alarm
00:13:30.133	14.6%	43.8%	72.5%	75%	alarm
00:13:45.676	14.9%	44.7%	74.4%	68.8%	alarm
00:14:00.572	14.9%	44.7%	74.4%	68.8%	alarm
00:14:15.059	15.2%	45.6%	76.2%	62.5%	
00:14:45.433	15.2%	45.6%	76.2%	52.5%	
00:15:00.424	15.2%	45.6%	76.2%	52.5%	
00:15:15.260	15.2%	45.6%	71.2%	48.8%	
00:15:30.236	15.2%	45.6%	71.2%	48.8%	
00:15:45.352	15.2%	45.6%	66.2%	45%	
00:16:00.032	15.2%	45.6%	66.2%	45%	
00:16:15.694	15.2%	45.6%	61.2%	41.2%	
00:16:30.795	15.2%	45.6%	61.2%	41.2%	
00:16:45.319	15.2%	45.6%	56.2%	31.2%	
00:17:00.420	15.2%	45.6%	56.2%	31.2%	
00:17:15.380	15.2%	45.6%	51.2%	21.2%	
00:17:30.653	15.2%	45.6%	51.2%	21.2%	
00:17:46.284	15.2%	45.6%	46.2%	11.2%	
00:18:00.119	15.2%	45.6%	46.2%	11.2%	
00:18:15.596	15.2%	45.6%	41.2%	7.5%	
00:18:30.323	15.2%	45.6%	41.2%	7.5%	
00:18:45.455	15.2%	45.6%	36.2%	3.8%	
00:19:00.337	15.2%	45.6%	36.2%	3.8%	
00:19:15.048	15.2%	45.6%	31.2%	0%	
00:19:30.274	15.2%	45.6%	31.2%	0%	
00:19:45.624	16%	48.1%	31.2%	10%	
00:20:00.600	16%	48.1%	31.2%	10%	
00:20:15.139	16.9%	50.6%	34.4%	20%	
00:20:30.131	16.9%	50.6%	34.4%	20%	
00:20:45.653	17.2%	51.6%	34.4%	23.8%	
00:21:00.442	17.2%	51.6%	34.4%	23.8%	
00:21:15.137	17.5%	52.5%	34.4%	27.5%	
00:21:30.363	17.5%	52.5%	34.4%	27.5%	
00:21:45.167	17.8%	53.4%	31.2%	31.2%	
00:22:00.954	17.8%	53.4%	31.2%	31.2%	

Fatigue Monitoring Tool Analysis
Date: 23 Jun 2013, 15:14:26
Video file associated: analysis_1358867666522.mp4

time	60min	20min	10min	5min	isAlarm
0	0%	0%	0%	0%	
00:00:00.033	0%	0%	0%	0%	

00:00:15.081	0%	0%	0%	0%
00:00:30.393	0.1%	0.3%	0.6%	1.2%
00:00:45.210	0.1%	0.3%	0.6%	1.2%
00:01:00.555	0.2%	0.6%	1.2%	2.5%
00:01:15.240	0.2%	0.6%	1.2%	2.5%
00:01:30.057	0.2%	0.6%	1.2%	2.5%
00:01:45.171	0.3%	0.9%	1.9%	3.8%
00:02:00.351	0.4%	1.2%	2.5%	5%
00:02:30.744	0.5%	1.6%	3.1%	6.2%
00:02:30.777	0.5%	1.6%	3.1%	6.2%
00:03:00.939	0.6%	1.9%	3.8%	7.5%
00:03:30.870	0.7%	2.2%	4.4%	8.8%
00:03:45.786	0.7%	2.2%	4.4%	8.8%
00:04:00.537	0.8%	2.5%	5%	10%
00:04:30.600	0.9%	2.8%	5.6%	11.2%
00:05:00.993	1%	3.1%	6.2%	12.5%
00:05:15.711	1%	3.1%	6.2%	12.5%
00:05:30.594	1.1%	3.4%	6.9%	12.5%
00:06:00.789	1.2%	3.8%	7.5%	12.5%
00:06:30.753	1.4%	4.1%	8.1%	12.5%
00:06:45.306	1.4%	4.1%	8.1%	12.5%
00:07:00.024	1.4%	4.1%	8.1%	12.5%
00:07:30.582	1.6%	4.7%	9.4%	12.5%
00:08:00.777	1.7%	5%	10%	12.5%
00:08:30.543	1.8%	5.3%	10.6%	12.5%
00:08:45.096	1.8%	5.3%	10.6%	12.5%
00:09:00.870	1.9%	5.6%	11.2%	12.5%
00:09:15.159	1.9%	5.6%	11.2%	12.5%
00:09:30.867	2%	5.9%	11.9%	12.5%
00:10:00.402	2%	5.9%	11.9%	12.5%
00:10:30.234	2.1%	6.2%	12.5%	12.5%
00:11:00.759	2.3%	6.9%	12.5%	12.5%
00:11:30.294	2.3%	6.9%	12.5%	12.5%
00:11:45.837	2.4%	7.2%	12.5%	12.5%
00:12:00.126	2.4%	7.2%	12.5%	12.5%
00:12:30.552	2.5%	7.5%	12.5%	12.5%
00:13:00.450	2.6%	7.8%	12.5%	12.5%
00:13:30.843	2.8%	8.4%	12.5%	12.5%
00:14:00.114	2.8%	8.4%	12.5%	12.5%
00:14:30.243	2.9%	8.8%	12.5%	12.5%
00:14:45.720	3%	9.1%	12.5%	12.5%
00:15:00.999	3.1%	9.4%	12.5%	12.5%
00:15:15.057	3.1%	9.4%	12.5%	12.5%
00:15:30.930	3.2%	9.7%	12.5%	12.5%
00:15:30.963	3.2%	9.7%	12.5%	12.5%
00:15:30.996	3.2%	9.7%	12.5%	12.5%
00:15:45.648	3.2%	9.7%	12.5%	12.5%
00:15:45.681	3.2%	9.7%	12.5%	12.5%
00:15:45.714	3.2%	9.7%	12.5%	12.5%
00:15:45.747	3.2%	9.7%	12.5%	12.5%
00:15:45.780	3.2%	9.7%	12.5%	12.5%
00:15:45.813	3.2%	9.7%	12.5%	12.5%
00:15:45.846	3.2%	9.7%	12.5%	12.5%
00:15:45.879	3.2%	9.7%	12.5%	12.5%
00:15:45.912	3.2%	9.7%	12.5%	12.5%
00:15:45.945	3.2%	9.7%	12.5%	12.5%
00:15:45.978	3.2%	9.7%	12.5%	12.5%
00:15:46.011	3.2%	9.7%	12.5%	12.5%
00:15:46.044	3.2%	9.7%	12.5%	12.5%
00:15:46.077	3.2%	9.7%	12.5%	12.5%
00:15:46.110	3.2%	9.7%	12.5%	12.5%
00:15:46.143	3.2%	9.7%	12.5%	12.5%
00:15:46.176	3.2%	9.7%	12.5%	12.5%
00:15:46.209	3.2%	9.7%	12.5%	12.5%
00:15:46.242	3.2%	9.7%	12.5%	12.5%
00:15:46.275	3.2%	9.7%	12.5%	12.5%
00:15:46.308	3.2%	9.7%	12.5%	12.5%

00:15:46.341	3.2%	9.7%	12.5%	12.5%
00:15:46.374	3.2%	9.7%	12.5%	12.5%
00:15:46.407	3.2%	9.7%	12.5%	12.5%
00:15:46.440	3.2%	9.7%	12.5%	12.5%
00:15:46.473	3.2%	9.7%	12.5%	12.5%
00:15:46.506	3.2%	9.7%	12.5%	12.5%
00:15:46.539	3.2%	9.7%	12.5%	12.5%
00:15:46.572	3.2%	9.7%	12.5%	12.5%
00:15:46.605	3.2%	9.7%	12.5%	12.5%
00:15:46.638	3.2%	9.7%	12.5%	12.5%
00:15:46.671	3.2%	9.7%	12.5%	12.5%
00:15:46.704	3.2%	9.7%	12.5%	12.5%
00:15:46.737	3.2%	9.7%	12.5%	12.5%
00:15:46.770	3.2%	9.7%	12.5%	12.5%
00:15:46.803	3.2%	9.7%	12.5%	12.5%
00:16:00.729	3.2%	9.7%	12.5%	12.5%
00:16:30.891	3.3%	10%	12.5%	12.5%
00:16:45.675	3.4%	10.3%	12.5%	12.5%
00:17:00.327	3.4%	10.3%	12.5%	12.5%
00:17:15.210	3.5%	10.6%	12.5%	12.5%
00:17:30.258	3.5%	10.6%	12.5%	12.5%
00:18:00.420	3.6%	10.9%	12.5%	12.5%
00:18:15.204	3.8%	11.2%	12.5%	12.5%
00:18:30.912	3.8%	11.2%	12.5%	12.5%
00:19:00.711	3.9%	11.6%	12.5%	12.5%
00:19:30.444	4%	11.9%	12.5%	12.5%
00:19:45.096	4.1%	12.2%	12.5%	12.5%
00:20:00.903	4.1%	12.2%	12.5%	12.5%
00:20:30.537	4.2%	12.5%	12.5%	12.5%
00:20:45.090	4.3%	12.5%	12.5%	12.5%
00:21:00.798	4.3%	12.5%	12.5%	12.5%
00:21:30.333	4.4%	12.5%	12.5%	12.5%
00:22:00.957	4.5%	12.5%	12.5%	12.5%
00:22:30.426	4.6%	12.5%	12.5%	12.5%
00:22:45.903	4.7%	12.5%	12.5%	12.5%
00:23:00.159	4.7%	12.5%	12.5%	12.5%
00:23:30.420	4.8%	12.5%	12.5%	12.5%
00:23:45.897	4.9%	12.5%	12.5%	12.5%
00:24:00.153	4.9%	12.5%	12.5%	12.5%
00:24:15.795	5%	12.5%	12.5%	12.5%
00:24:30.008	5%	12.5%	12.5%	12.5%
00:25:00.873	5.1%	12.5%	12.5%	12.5%
00:25:30.870	5.2%	12.5%	12.5%	12.5%

References

[1] S. Bourgeois-Bougrine, P. Carbon, C. Gounelle, R. Mollard, and A. Coblentz, "Perceived Fatigue for Short- and Long-Haul Flights: A Survey of 739 Airline Pilots," *Aviation, Space, and Environmental Medicine,* vol. 74, pp. 1072-1077, 2003.

[2] J. A. Stern, D. Boyer, and D. Schroeder, "Blink Rate: A Possible Measure of Fatigue," *Human Factors: The Journal of the Human Factors and Ergonomics* vol. 36, pp. 285-297, 1994.

[3] P. Jackson, C. Hilditch, A. Holmes, N. Reed, N. Merat, and L. Smith, "Fatigue and Road Safety: A Critical Analysis of Recent Evidence," ed. London: Department for Transport: Department for Transport, 2011, p. 88.

[4] R. Resendes and K. H. martin, "Saving Lives Through Advanced Safety Technology," Faderal Highway Administration US, Washington DC, 2003.

[5] M. Harma, M. Partinen, R. Repo, M. Sorsa, and P. Siivonen, "Effects of 6/6 and 4/8 Watch Systems on Sleepiness among Bridge Officers," *Chronobiol Int,* vol. 25, pp. 413-23, 2008.

[6] I. Houtman, M. Miedema, K. Jettinghoff, A. Starren, J. Heinrich, and J. Gort, "Fatigue in the shipping industry," TNO-Netherlands Organisation for Applied Scientific Research, The Netherlands, 2005.

[7] S. Hu and G. Zheng, "Driver drowsiness detection with eyelid related parameters by Support Vector Machine," *Expert Systems with Applications,* vol. 36, pp. 7651-7658, 2009.

[8] S.-M. Zhou, J. Q. Gan, and F. Sepulveda, "Classifying mental tasks based on features of higher-order statistics from EEG signals in brain-computer interface," *Inf. Sci.,* vol. 178, pp. 1629-1640, 2008.

[9] B. T. Jap, S. Lal, P. Fischer, and E. Bekiaris, "Using EEG spectral components to assess algorithms for detecting fatigue," *Expert Systems with Applications,* vol. 36, pp. 2352-2359, 2009.

[10] D. F. Dinges and R. Grace, "PERCLOS: A Valid Psychophysiological Measure of Alertness As Assessed by Psychomotor Vigilance," Washington DC No. FHWA-MCRT-98-006, 1998.

[11] Q. Wu, B. Sun, B. Xie, and J. Zhao, "A PERCLOS-Based Driver Fatigue Recognition Application for Smart Vehicle Space," *Third International Symposium on Information Processing (ISIP),* pp. 437-441, 2010.

[12] J.-F. Xie, M. Xie, and W. Zhu, "Driver fatigue detection based on head gesture and PERCLOS," *International Conference on Wavelet Active Media Technology and Information Processing (ICWAMTIP),* 2012, pp. 128-131.

[13] A. S. Eye. *"Revolutionary Eye Tracking Technology"*. Available: http://www.smarteye.se/, [February 2012].
[14] S. Machines. *"faceLAB 5"*. Available: http://www.seeingmachines.com/product/facelab/, [Accessed: February 2012]
[15] P. Viola and M. J. JONES, "Robust Real-Time Face Detection," *International Journal of Computer Vision,* vol. 57(2), pp. 137–154, 2004.
[16] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* 2001, pp. I-511-I-518 vol.1.
[17] D. f. T. UK, "Reported Road Casualties in Great Britain: 2010 Annual Report " Department for Transport UK, United Kingdom,Statistical Release - 29 September 2011
[18] H. S. Group and *"Driver Fatigue, Lane Management & Warning Systems"*. Available: http://www.driverfatiguemonitor.com/dfm/dfm.html, Oct. 2010, [Accessed: February 2012].
[19] S. Instruments. *"SMI Gaze & Eye Tracking"*. Available: http://www.smivision.com/en/gaze-and-eye-tracking-systems/home.html, [Accessed: February 2012].
[20] Volkwagen. *"Driver Alert System"*. Available: http://www.volkswagen.co.uk/#/new/passat-vii/explore/experience/driver-assistance/driver-alert-system/, [Accessed: February 2012]
[21] ZerCustoms. *"Volvo Driver Alert Control and Lane Departure Warning"*. Available: http://www.zercustoms.com/news/Volvo-Driver-Alert-Control-and-Lane-Departure-Warning.html, 2011, [Accessed: February 2012].
[22] Daimler. *"Drowsiness-detection system warns drivers to prevent them falling asleep momentarily"*. Available: http://www.daimler.com/dccom/, [Accessed: February 2012].
[23] Toyota. *"Driver Monitoring System"*. Available: http://www.lexus.co.uk/range/ls/key-features/safety/safety-driver-monitoring-system.aspx, [Accessed: February 2012].
[24] J. A. Caldwell, "Fatigue in aviation," *Travel Medicine and Infectious Disease,* vol. 3, pp. 85-96, 2005.
[25] Q. S. Aviation. *"Fatigue Risk Management Systems (FRMS)"*. Available: www.aviation-quality-services.com/, [Acccessed: February 2012].
[26] S. R. Group, "Safety Management Systems – Guidance To Organisations," Civil Aviation Authority, 2008.
[27] Optalert. *"Alert & Alive"*. Available: http://www.optalert.com/, Jan. 2012 [Accessed February 2012].
[28] M. A. M. Corbett, "A drowsiness detection system for pilots: Optalert," *Aviat Space Environ Med,* vol. 80, pp. 149-149, 2009.
[29] M. R. Grech, T. J. Horberry, and M. S. Humphreys, "Fatigue and human error in the maritime domain," in *5th International Conference on Fatigue and Transportation*, Victoria, Australia, 2003.
[30] P. H. Gander and M. U. S. W. R. Centre, *A Review of Fatigue Management in the Maritime Sector*: Sleep Wake Research Centre, Massey University, 2005.

[31] M. Lützhöft, A. Dahlgren, A. Kircher, B. Thorslund, and M. Gillberg, "Fatigue at sea in Swedish shipping-a field study," *American journal of industrial medicine,* vol. 53, pp. 733-740, 2010.
[32] B. Thorslund, M. Lützhöft, M. Gillberg, and A. Kircher, "Fatigue at sea – A field study in Swedish shipping," WTI, Sweden2007.
[33] F. M. International.. "*ASTiD"* [online]. Available: http://www.fmig.org/astid.html, [Accessed: September 2012]
[34] D. Dawson, Y. Ian Noy, M. Härmä, T. Åkerstedt, and G. Belenky, "Modelling fatigue and the use of fatigue models in work settings," *Accident Analysis & Prevention,* vol. 43, pp. 549-564, 2011.
[35] L. K. Barger, B. E. Cade, N. T. Ayas, J. W. Cronin, B. Rosner, F. E. Speizer, and C. A. Czeisler, "Extended work shifts and the risk of motor vehicle crashes among interns," *New England Journal of Medicine,* vol. 352, pp. 125-134, 2005.
[36] A. E. Dembe, J. B. Erickson, R. G. Delbos, and S. M. Banks, "The impact of overtime and long work hours on occupational injuries and illnesses: new evidence from the United States," *Occupational and environmental medicine,* vol. 62, pp. 588-597, 2005.
[37] W. W. Wierwille, S. S. Wreggit, and R. R. Knipling, "DEVELOPMENT OF IMPROVED ALGORITHMS FOR ON-LINE DETECTION OF DRIVER DROWSINESS," presented at the International Congress on Transportation Electronics, Detroit, 1994.
[38] S. K. L. Lal and A. Craig, "Driver fatigue: Electroencephalography and psychological assessment," *Psychophysiology,* vol. 39, pp. 313-321, 2002.
[39] P. P. Caffier, U. Erdmann, and P. Ullsperger, "Experimental evaluation of eye-blink parameters as a drowsiness measure," *European Journal of Applied Physiology,* vol. 89, pp. 319-325, May 2003.
[40] R. Senaratne, B. Jap, S. Lal, A. Hsu, S. Halgamuge, and P. Fischer, "Comparing two video-based techniques for driver fatigue detection: classification versus optical flow approach," *Machine Vision and Applications,* vol. 22, pp. 597-618, 2011.
[41] Y. Morad, H. Lemberg, N. Yofe, and Y. Dagan, "Pupillography as an objective indicator of fatigue," *Current Eye Research,* vol. 21, pp. 535-542, 2000.
[42] M. Nakayama, "Influence of blink on pupillary indices," *IEEE Biomedical Circuits and Systems Conference (BioCAS),*. 2006, pp. 29-32.
[43] J. Nishiyama, K. Tanida, M. Kusumi, and Y. Hirata, "The pupil as a possible premonitor of drowsiness," *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.(EMBS)*, pp. 1586-1589, 2007.
[44] M. Omidyeganeh, A. Javadtalab, and S. Shirmohammadi, "Intelligent driver drowsiness detection through fusion of yawning and eye closure," *IEEE International Conference on Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS),* pp. 1-6, 2011.
[45] B. J. Wilson and T. D. Bracewell, "Alertness monitor using neural networks for EEG analysis," *IEEE Signal Processing Society Workshop Neural Networks for Signal Processing*, pp. 814-820 vol.2, 2000.

[46] G. Yang, Y. Lin, and P. Bhattacharya, "A driver fatigue recognition model based on information fusion and dynamic Bayesian network," *Information Sciences,* vol. 180, pp. 1942-1954, 2010.
[47] D. Sommer, M. Golz, U. Trutschel, and D. Edwards, "Biosignal Based Discrimination between Slight and Strong Driver Hypovigilance by Support-Vector Machines," in *Agents and Artificial Intelligence*. vol. 67, J. Filipe, A. Fred, and B. Sharp, Eds., ed: Springer Berlin Heidelberg, 2010, pp. 177-187.
[48] N. Wright and A. McGown, "Vigilance on the civil flight deck: incidence of sleepiness and sleep during long-haul flights and associated changes in physiological parameters," *Ergonomics,* vol. 44(1), pp. 82-106, 2001.
[49] M. Trimmel, M. Meixner-Pendleton, and S. Haring, "Stress response caused by system response time when searching for information on the Internet," *Human Factors: The Journal of the Human Factors and Ergonomics Society,* vol. 45, pp. 615-622, 2003.
[50] NEUROCOM., "*Engine Driver Vigilance Telemetric Control System (EDVTCS)"* Available:http://www.neurocom.ru/en2/product/edvtcs.html, [Accessed: June 2012].
[51] SenseWear.,"*Introducing the Enhanced BodyMedia SenseWear System"* [Available:http://sensewear.bodymedia.com/SW-Learn-More/Product-Overview, [Accessed June 2012].
[52] M. Michael Littner, C. A. Kushida, D. Dennis Bailey, R. B. Berry, D. G. Davila, and M. Hirshkowitz, "Practice parameters for the role of actigraphy in the study of sleep and circadian rhythms: an update for 2002," *Sleep,* vol. 26, p. 337, 2003.
[53] A. V. Dane, R. J. Schachar, and R. Tannock, "Does actigraphy differentiate ADHD subtypes in a clinical research setting?," *Journal of the American Academy of Child & Adolescent Psychiatry,* vol. 39, pp. 752-760, 2000.
[54] N. Wright and A. McGown, "Involuntary sleep during civil air operations: wrist activity and the prevention of sleep," *Aviation, Space, and Environmental Medicine,* vol. 75, pp. 37-45, 2004.
[55] N. Wright, D. Powell, A. Mcgown, E. Broadbent, and P. Loft, "Avoiding involuntary sleep during civil air operations: Validation of a wrist-worn alertness device," *Aviation, Space, and Environmental Medicine,* vol. 76, pp. 847-856, 2005.
[56] E. R. Lab, "*DS2000 Driver Alert System"*,Available: http://www.ellison-research.com/about.html, [Accessed: June 2012].
[57] N. Mabbott, "ARRB Pro-active fatigue management system," *Road and Transport Research,* vol. 12, 2003.
[58] A. Gundel, K. Marsalek, and C. ten Thoren, "A critical review of existing mathematical models for alertness," *Somnologie-Schlafforschung und Schlafmedizin,* vol. 11, pp. 148-156, 2007.
[59] G. D. Roach, A. Fletcher, and D. Dawson, "A model to predict work-related fatigue based on hours of work," *Aviation, space, and environmental medicine,* vol. 75, pp. A61-A69, 2004.

[60] T. Akerstedt, S. Folkard, and C. Portin, "Predictions from the three-process model of alertness," *Aviation, space, and environmental medicine,* vol. 75, pp. A75-A83, 2004.
[61] A. J. Belyavin and M. B. Spencer, "Modeling performance and alertness: the QinetiQ approach," *Aviation, space, and environmental medicine,* vol. 75, pp. A93-A103, 2004.
[62] H. De Rosario, J. S. Solaz, Rodri, x, N. guez, and L. M. Bergasa, "Controlled inducement and measurement of drowsiness in a driving simulator," *Intelligent Transport Systems, IET,* vol. 4, pp. 280-288, 2010.
[63] S. Pritchett, E. Zilberg, X. Zheng Ming, M. Karrar, S. Lal, and D. Burton, "Strengthening association between driver drowsiness and its physiological predictors by combining EEG with measures of body movement," *6th International Conference on Broadband and Biomedical Communications (IB2Com),* pp. 103-107, 2011.
[64] M. St John, M. R. Risser, and D. A. Kobus, "Toward a usable closed-loop attention management system: Predicting vigilance from minimal contact head, eye, and EEG measures," *Foundations of Augmented Cognition,* pp. 12-18, 2006.
[65] N. Rodriguez-Ibanez, M. A. Garcia-Gonzalez, M. Fernandez-Chimeno, and J. Ramos-Castro, "Drowsiness detection by thoracic effort signal analysis in real driving environments," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 6055-6058, 2011.
[66] P. M. Tekade and S. Gawali, "Investigation and New Method of No intrusive Detection of Driver Drowsiness."
[67] P. Philip, P. Sagaspe, N. Moore, J. Taillard, A. Charles, C. Guilleminault, and B. Bioulac, "Fatigue, sleep restriction and driving performance," *Accident Analysis & Prevention,* vol. 37, pp. 473-478, 2005.
[68] A. Whitlock and J. Pethick, "Driver vigilance devices: systems review," *Веб-узел Quintec Associates Limited: http://www. rssb. co. uk/Pages/Main. aspx,* 2002.
[69] E. Bekiaris, S. Nikolaou, and A. Mousadakou, "Design guidelines for driver drowsiness detection and avoidance," *System for effective Assessment of driver vigilance and Warning According to traffic risK Estimation (AWAKE) Deliverable,* vol. 9. 2002.
[70] L. Barr, S. Popkin, and H. Howarth, "An evaluation of emerging driver fatigue detection measures and technologies," 2009.
[71] Y. Ming-Hsuan, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 24, pp. 34-58, 2002.
[72] M.-H. Yang, "Face Detection," University of California, Merced, CA 95344, 2009.
[73] G. Yang and T. S. Huang, "Human face detection in a complex background," *Pattern Recognition,* vol. 27, pp. 53-63, 1994.
[74] C. Kotropoulos and I. Pitas, "Rule-based face detection in frontal views," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* vol. 4, pp. 2537-2540 vol.4, 1997.

[75] S. A. Sirohey, "Human Face Segmentation and Identification," Center of Automation Research University of Maryland CS-TR-3176, 1993.
[76] H. P. Graf, E. Cosatto, D. Gibbon, M. Kocheisen, and E. Petajan, "Multi-modal system for locating heads and faces," *Second International Conference on Automatic Face and Gesture Recognition,* pp. 88-93, 1996.
[77] M. F. Augusteijn and T. L. Skufca, "Identification of human faces through texture-based feature recognition and neural network technology," *IEEE International Conference on Neural Networks,* pp. 392-398 vol.1, 1993.
[78] D. Chai and K. N. Ngan, "Face segmentation using skin-color map in videophone applications," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 9, pp. 551-564, 1999.
[79] S. K. Singh, D. S. Chauhan, M. Vatsa, and R. Singh, "A Robust Skin Color Based Face Detection Algorithm," *Tamkang Journal of Science and Engineering,* vol. 6, pp. 227-234, 2003.
[80] J. M. Chaves-González, M. A. Vega-Rodríguez, J. A. Gómez-Pulido, and J. M. Sánchez-Pérez, "Detecting skin in face recognition systems: A colour spaces study," *Digital Signal Processing,* vol. 20, pp. 806-823, 2010.
[81] B. Scassellati, "Eye finding via face detection for a foveated active vision system," *National Conference on Artificial Intelligence,* pp. 969-976, 1998.
[82] J. Miao, B. Yin, K. Wang, L. Shen, and X. Chen, "A hierarchical multiscale and multiangle system for human face detection in a complex background using gravity-center template," *Pattern Recognition,* vol. 32, pp. 1237-1248, 1999.
[83] J. Wang and T. Tan, "A new face detection method based on shape information," *Pattern Recognition Letters,* vol. 21, pp. 463-471, 2000.
[84] G. J. Edwards, C. J. Taylor, and T. F. Cootes, "Learning to identify and track faces in image sequences," *Third IEEE International Conference on Automatic Face and Gesture Recognition,* pp. 260-265, 1998.
[85] B. Scassellati, "Eye finding via face detection for a foveated active vision system," in *Proceedings of the National Conference on Artificial Intelligence*, 1998, pp. 969-976.
[86] G. J. Edwards, C. J. Taylor, and T. F. Cootes, "Learning to identify and track faces in image sequences," *Third IEEE International Conference on Automatic Face and Gesture Recognition,* pp. 260-265, 1998.
[87] C. Zhang and Z. Zhang, "A Survey of Recent Advances in Face Detection," Microsoft Corporation MSR-TR-2010-66, 2010.
[88] C. Shavers, R. Li, and G. Lebby, "An SVM-based approach to face detection," *Thirty-Eighth Southeastern Symposium on System Theory (SSST)* pp. 362-366, 2006.
[89] M. Roohi, G. Mirjalily, and M. T. Sadeghi, "Face Detection Using a Modified SVM-Based Classifier," *International Conference on Computational Intelligence and Multimedia Applications,* vol. 2, pp. 354-360, 2007.
[90] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 12, pp. 103-108, 1990.

[91] H. M. El-Bakry and Z. Qiangfu, "Fast Neural Implementation of PCA for Face Detection," *International Joint Conference on Neural Networks (IJCNN)* pp. 806-811, 2006.
[92] L. Liying and G. Weiwei, "The Face Detection Algorithm Combined Skin Color Segmentation and PCA," *International Conference on Information Engineering and Computer Science (ICIECS)* 2009.
[93] Z. Jing, Z. Xue-dong, and H. Seok-wun, "A Novel Approach Using PCA and SVM for Face Detection," *Fourth International Conference on Natural Computation (ICNC)* vol. 3, pp. 29-33, 2008.
[94] P. Viola and M. Jones, "Robust real-time face detection," *Eighth IEEE International Conference on Computer Vision (ICCV),* vol. 2, 2001.
[95] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. I-511-I-518 vol. 1, 2001.
[96] W. Jianxin, S. C. Brubaker, M. D. Mullin, and J. M. Rehg, "Fast Asymmetric Learning for Cascade Face Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 30, pp. 369-382, 2008.
[97] S.-W. Lee, S. Li, L. Zhang, R. Chu, S. Xiang, S. Liao, and S. Li, "Face Detection Based on Multi-Block LBP Representation," in *Advances in Biometrics* vol. 4642, ed: Springer Berlin / Heidelberg, 2007.
[98] P. Minh-Tri, V. D. D. Hoang, and C. Tat-Jen, "Detection with multi-exit asymmetric boosting," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 1-8, 2008.
[99] D. Cristinacce and T. Cootes, "A comparison of shape constrained facial feature detectors," *Sixth IEEE International Conference on Automatic Face and Gesture Recognition,* pp. 375-380, 2004.
[100] P. I. Wilson and J. Fernandez, "Facial feature detection using Haar classifiers," *Journal of Computing Sciences in Colleges,* vol. 21, pp. 127-133, 2006.
[101] J. Batista, "A Drowsiness and Point of Attention Monitoring System for Driver Vigilance," *IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 702-708, 2007.
[102] E. Hjelmås and B. K. Low, "Face Detection: A Survey," *Computer Vision and Image Understanding,* vol. 83, pp. 236-274, 2001.
[103] A.Pérez, M.L.Córdoba, A.García, and R.Méndez, "A Precise Eye-Gaze Detection and Tracking System," presented at the WSCG, 2003.
[104] R. Valenti and T. Gevers, "Accurate eye center location and tracking using isophote curvature," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 1-8, 2008.
[105] N. Eveno, A. Caplier, and P. Y. Coulon, "Accurate and quasi-automatic lip tracking," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 14, pp. 706-715, 2004.
[106] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer vision and image understanding,* vol. 61, pp. 38-59, 1995.

[107] S. Sirohey, A. Rosenfeld, and Z. Duric, "A method of detecting and tracking irises and eyelids in video," *Pattern Recognition,* vol. 35, pp. 1389-1401, 2002.
[108] D. Vukadinovic and M. Pantic, "Fully automatic facial feature point detection using Gabor feature based boosted classifiers," *IEEE International Conference on Systems, Man and Cybernetics,* pp. 1692-1698, 2005.
[109] S. Kawato and J. Ohya, "Real-time detection of nodding and head-shaking by directly detecting and tracking the between-eye," *Fourth IEEE International Conference on Automatic Face and Gesture Recognition.,* pp. 40-45, 2000.
[110] T. Vatahska, M. Bennewitz, and S. Behnke, "Feature-based head pose estimation from images," *7th IEEE-RAS International Conference on Humanoid Robots,* pp. 330-335, 2007.
[111] N. Alioua, A. Amine, M. Rziza, and D. Aboutajdine, "Driver's Fatigue and Drowsiness Detection to Reduce Traffic Accidents on Road Computer Analysis of Images and Patterns." vol. 6855, P. Real, D. Diaz-Pernil, H. Molina-Abril, A. Berciano, and W. Kropatsch, Eds., ed: Springer Berlin / Heidelberg, 2011, pp. 397-404.
[112] B. Fasel and J. Luettin, "Automatic facial expression analysis: a survey," *Pattern Recognition,* vol. 36, pp. 259-275, 2003.
[113] D. W. Hansen and J. P. Hansen, "Robustifying Eye Interaction," *Conference on Computer Vision and Pattern Recognition Workshop (CVPRW),* pp. 152-152, 2006.
[114] I. Fasel, B. Fortenberry, and J. Movellan, "A generative framework for real time object detection and classification," *Computer Vision and Image Understanding,* vol. 98, pp. 182-210, 2005.
[115] Z. Ping, S. Xiaoyou, and L. Feng, "A novel eye detecting technology Based on Adaboost and gray-scale information," *International Conference on Computer Design and Applications (ICCDA),* vol. 1, pp. V1-563-V1-566, 2010.
[116] S. Mingli, T. Dacheng, S. Zhuo, and L. Xuelong, "Visual-Context Boosting for Eye Detection," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, ,* vol. 40, pp. 1460-1467, 2010.
[117] K. Donghoon and R. Dahyot, "Face Components Detection Using SURF Descriptors and SVMs," *International Machine Vision and Image Processing Conference, (IMVIP)*, pp. 51-56, 2008.
[118] B. Lu, Y. Sun, and L. Wang, "Dynamic Face Fatigue Detection Based on Feature-Lever Fusion," in *Multimedia Communications (Mediacom), 2010 International Conference on*, 2010, pp. 123-126.
[119] E. Vural, M. Cetin, A. Ercil, G. Littlewort, M. Bartlett, and J. Movellan, "Drowsy driver detection through facial movement analysis," *Human–Computer Interaction,* pp. 6-18, 2007.
[120] H. Yang, X. Jiang, Y. Zhang, and L. Wang, "Fatigue Detection Based on Regional Local Binary Patterns Histogram and Support Vector Machine," *International Conference on Computer Science and Electronics Engineering (ICCSEE),* pp. 628-631, 2012.

[121] J.-D. Wu and T.-R. Chen, "Development of a drowsiness warning system based on the fuzzy logic images analysis," *Expert Systems with Applications,* vol. 34, pp. 1556-1561, 2008.
[122] M. Chau and M. Betke, "Real time eye tracking and blink detection with USB cameras," Boston University, Boston, USA., Technical Report 2005-01,2005.
[123] I. Bacivarov, M. Ionita, and P. Corcoran, "Statistical models of appearance for eye tracking and eye-blink detection and measurement," *IEEE Transactions on Consumer Electronics,* vol. 54, pp. 1312-1320, 2008.
[124] F. Sukno, S.-K. Pavani, C. Butakoff, and A. Frangi, "Automatic Assessment of Eye Blinking Patterns through Statistical Shape Models Computer Vision Systems." vol. 5815, M. Fritz, B. Schiele, and J. Piater, Eds., ed: Springer Berlin / Heidelberg, 2009, pp. 33-42.
[125] H.-Z. Dong and M. Xie, "Real-time driver fatigue detection based on simplified landmarks of AAM," *International Conference on Apperceiving Computing and Intelligence Analysis (ICACIA)*, pp. 363-366, 2010.
[126] L. Liying and Q. Haoxiang, "The Study of Driver Fatigue Monitor Algorithm Based on Skin Color Segmentation," *International Symposium on Intelligent Information Technology Application Workshop (IITAW)* , pp. 463-466, 2008.
[127] T. Morris, P. Blenkhorn, and F. Zaidi, "Blink detection for real-time eye tracking," *J. Netw. Comput. Appl.,* vol. 25, pp. 129-143, 2002.
[128] T. Azim and M. Jaffar, "Automatic Fatigue Detection of Drivers through Pupil Detection and Yawning," 2009.
[129] J. Jiménez-Pinto and M. Torres-Torriti, "Driver alert state and fatigue detection by salient points analysis," *IEEE International Conference on Systems, Man and Cybernetics,(SMC),* pp. 455-461, 2009.
[130] J. Shi and C. Tomasi, "Good features to track," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 593-600, 2004.
[131] M. Divjak and H. Bischof, "Eye blink based fatigue detection for prevention of Computer Vision Syndrome," presented at the IAPR Conference on Machine Vision Applications, Yokohama, Japan., 2009.
[132] M. Lalonde, D. Byrns, L. Gagnon, N. Teasdale, and D. Laurendeau, "Real-time eye blink detection with GPU-based SIFT tracking," *Fourth Canadian Conference on Computer and Robot Vision (CRV)*, pp. 481-487, 2007.
[133] R.-b. Wang, K.-y. Guo, S.-m. Shi, and J.-w. Chu, "A monitoring method of driver fatigue behavior based on machine vision," *IEEE Intelligent Vehicles Symposium,* pp. 110-113, 2003.
[134] I. Fasel, B. Fortenberry, and J. Movellan, "A generative framework for real time object detection and classification," *Comput. Vis. Image Underst.,* vol. 98, pp. 182-210, 2005.
[135] P. Gang, S. Lin, W. Zhaohui, and L. Shihong, "Eyeblink-based Anti-Spoofing in Face Recognition from a Generic Webcamera," *IEEE 11th International Conference on Computer Vision (ICCV)*, pp. 1-8, 2007.

[136] M. Castrillón, O. Déniz, C. Guerra, and M. Hernández, "ENCARA2: Real-time detection of multiple faces at different resolutions in video streams," *Journal of Visual Communication and Image Representation,* vol. 18, pp. 130-140, 2007.
[137] C. Xu, Y. Zheng, and Z. Wang, "Efficient eye states detection in real-time for drowsy driving monitoring system," in *Information and Automation, 2008. ICIA 2008. International Conference on*, 2008, pp. 170-174.
[138] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 29, pp. 915-928, 2007.
[139] Z. Sanyuan and S. Tingzhi, "Driver Fatigue Detection Based on Eye Status," *International Conference on Multimedia Technology (ICMT),* , pp. 1-4, 2004.
[140] W. Yao, Y. Liang, and M. Du, "A real-time lip localization and tacking for lip reading," *3rd International Conference on Advanced Computer Theory and Engineering (ICACTE),* vol. 6, pp. V6-363, 2010.
[141] M. Mohanty, A. Mishra, and A. Routray, "A non-rigid motion estimation algorithm for yawn detection in human drivers," *International Journal of Computational Vision and Robotics,* vol. 1, pp. 89-109, 2009.
[142] F. Xiao, Y. Bao-Cai, and S. Yan-Feng, "Yawning Detection for Monitoring Driver Fatigue," *International Conference on Machine Learning and Cybernetics.,* vol. 2, pp. 664-668, 2007.
[143] W. Lirong, W. Xiaoli, and X. Jing, "Lip Detection and Tracking Using Variance Based Haar-Like Features and Kalman filter," *Fifth International Conference on Frontier of Computer Science and Technology (FCST),* pp. 608-612, 2010.
[144] L. Lingling, C. Yangzhou, and L. Zhenlong, "Yawning detection for monitoring driver fatigue based on two cameras," *12th International IEEE Conference on Intelligent Transportation Systems (ITSC),* pp. 1-6, 2009.
[145] Y. Yang, J. Sheng, and W. Zhou, "The Monitoring Method of Driver's Fatigue Based on Neural Network," *International Conference on Mechatronics and Automation (ICMA),* pp. 3555-3559, 2007.
[146] H. García, A. Salazar, D. Alvarez, and Á. Orozco, "Driving fatigue detection using active shape models," *Advances in Visual Computing,* pp. 171-180, 2010.
[147] S. Anumas and S.-c. Kim, "Driver fatigue monitoring system using video face images & physiological information," *Biomedical Engineering International Conference (BMEiCON),* pp. 125-130, 2012.
[148] S. Hachisuka, K. Ishida, T. Enya, and M. Kamijo, "Facial expression measurement for detecting driver drowsiness," *Engineering Psychology and Cognitive Ergonomics,* pp. 135-144, 2011.
[149] P. Ekman and W. V. Friesen, "Facial action coding system: A technique for the measurement of facial movement. Palo Alto," *CA: Consulting Psychologists Press. Ellsworth, PC, & Smith, CA (1988). From appraisal to emotion: Differences among unpleasant feelings. Motivation and Emotion,* vol. 12, pp. 271-302, 1978.
[150] Q. Wu, B. Sun, B. Xie, and J. Zhao, "A PERCLOS-Based Driver Fatigue Recognition Application for Smart Vehicle Space," *Third International Symposium on Information Processing (ISIP),* pp. 437-441, 2010.

[151] J. Qiang, Z. Zhiwei, and P. Lan, "Real-time nonintrusive monitoring and prediction of driver fatigue," *IEEE Transactions on Vehicular Technology,* vol. 53, pp. 1052-1068, 2004.
[152] L. Weiwei, S. Haixin, and S. Weijie, "Driver fatigue detection through pupil detection and yawing analysis," *International Conference on Bioinformatics and Biomedical Technology (ICBBT),* pp. 404-407, 2010.
[153] Q. Ji and X. Yang, "Real time visual cues extraction for monitoring driver vigilance," *Computer Vision Systems,* pp. 107-124, 2001.
[154] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. I-511-I-518 vol.1, 2001.
[155] M. Grgic and K. Delac. *"Face Recognition Homepage"*. Available: http://www.face-rec.org/databases/, [Accessed: Jan. 2013].
[156] A. Martinez and R. Benavente, "The AR face database, 1998," *Computer Vision Center, Technical Report,* vol. 3, 2007.
[157] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 25, pp. 1615-1618, 2003.
[158] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (CVPR)*, pp. 586-591, 1991.
[159] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 22, pp. 1090-1104, 2000.
[160] G. Wen, C. Bo, S. Shiguang, C. Xilin, Z. Delong, Z. Xiaohua, and Z. Debin, "The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, ,* vol. 38, pp. 149-161, 2008.
[161] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 20, pp. 23-38, 1998.
[162] K.-K. Sung and T. Poggio, "Example-based learning for view-based human face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 20, pp. 39-51, 1998.
[163] J. Daugman, "How iris recognition works," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 14, pp. 21-30, 2004.
[164] J. Daugman, "New Methods in Iris Recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics,* vol. 37, pp. 1167-1175, 2007.
[165] H. Proenca, S. Filipe, R. Santos, J. Oliveira, and L. A. Alexandre, "The UBIRIS.v2: A Database of Visible Wavelength Iris Images Captured On-the-Move and At-a-Distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 32, pp. 1529-1535, 2010.
[166] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, "Image understanding for iris biometrics: A survey," *Computer Vision and Image Understanding,* vol. 110, pp. 281-307, 2008.

[167] CASIA, "*Biometrics Idea Test",* Available:http://www.idealtest.org/findTotalDbByMode.do?mode=Iris, [Accessed: May 2013]
[168] MMU, "*Iris Recognition Homepage*" Available: http://pesona.mmu.edu.my/~ccteo/, [Accessed: May 2013].
[169] Bath, "Bath Iris Image Database" Available: http://www.smartsensors.co.uk/information/bath-iris-image-database/, [Accessed: May 2013].
[170] CASIA, *"Biometrics Idea Test"*. Available: http://www.idealtest.org/findTotalDbByMode.do?mode=Iris, 2013 [Accessed: May 2013].
[171] T. Danisman, I. M. Bilasco, C. Djeraba, and N. Ihaddadene, "Drowsy driver detection system using eye blink patterns," *International Conference on Machine and Web Intelligence (ICMWI),* pp. 230-233, 2010.
[172] X. Fan, B.-C. Yin, and Y.-F. Sun, "Dynamic Human Fatigue Detection Using Feature-Level Fusion," in *Image and Signal Processing*, ed: Springer, 2008, pp. 94-102.
[173] J. Kovac, P. Peer, and F. Solina, "Human skin color clustering for face detection," *IEEE Region 8 Computer as a Tool EUROCON,* vol. 2, pp. 144-148 vol.2, 2003.
[174] P. Yu-Ting, R. Shang-Jang, S. Mon-Chau, and L. Yi-Chi, "A Simple and Accurate Color Face Detection Algorithm in Complex Background," *IEEE International Conference on Multimedia* pp. 1545-1548, 2006.
[175] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods," *Pattern Recognition,* vol. 40, pp. 1106-1122, 2007.
[176] E. R. Davies, *Machine vision: theory, algorithms, practicalities*: Morgan Kaufmann, 2004.
[177] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence, ,* vol. 25, pp. 564-577, 2003.
[178] B. W. Silverman, *Density estimation for statistics and data analysis* vol. 26: Chapman & Hall/CRC, 1986.
[179] G. R. Bradski, "Computer Vision Face Tracking For Use in a Perceptual User Interface," *Intel Technology Journal,* vol. 2, pp. 214--219, 1998.
[180] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Comput. Surv.,* vol. 38, p. 13, 2006.
[181] N. Jifeng, Y. Shuqin, and Y. Fuzeng, "An Algorithm of Adaptive Deformation Estimation of Moving Object in the Mean Shift Algorithm," *International Symposium on Computer Network and Multimedia Technology (CNMT),* pp. 1-4, 2004.
[182] M. Nilsson, J. Nordberg, and I. Claesson, "Face Detection using Local SMQT Features and Split up Snow Classifier," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* vol. 2, pp. 589-592, 2007.

[183] P. Breuer, K. Kwang-In, W. Kienzle, B. Scholkopf, and V. Blanz, "Automatic 3D face reconstruction from single images or video," *8th IEEE International Conference on Automatic Face & Gesture Recognition,* pp. 1-8, 2008.
[184] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 24, pp. 971-987, 2002.
[185] R. O. Duda and P. E. Hart, *Pattern classification and scene analysis* vol. 3: Wiley New York, 1973.
[186] J.-Y. Zhang, Y. Chen, and X.-x. Huang, "Edge detection of images based on improved Sobel operator and genetic algorithms," *International Conference on Image Analysis and Signal Processing (IASP)*, pp. 31-35, 2009.
[187] T. ÅKerstedt, M. Ingre, G. Kecklund, A. Anund, D. Sandberg, M. Wahde, P. Philip, and P. Kronberg, "Reaction of sleepiness indicators to partial sleep deprivation, time of day and time on task in a driving simulator – the DROWSI project," *Journal of Sleep Research,* vol. 19, pp. 298-309, 2010.
[188] C. Anderson, A. W. Wales, and J. A. Horne, "PVT lapses differ according to eyes open, closed, or looking away," *Sleep,* vol. 33, p. 197, 2010.
[189] M. M. Ibrahim, J. S. Soroghan, and L. Petropoulakis, "Mouth covered detection for yawn," *IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, pp. 89-94, 2013.
[190] M. M. Ibrahim, J. J. Soraghan, and L. Petropoulakis, "Non-rigid eye movement tracking and eye state quantification," *19th International Conference on Systems, Signals and Image Processing (IWSSIP),* pp. 280-283, 2012.
[191] M. M. Ibrahim, J. J. Soraghan, and L. Petropoulakis, "Eye-state analysis using an interdependence and adaptive scale mean shift (IASMS) algorithm," *Biomedical Signal Processing and Control,* vol. 11, pp. 53-62, 2014.