Driver Fatigue Detection Using an RGB-D Sensor Based On Eye Tracking and Head Pose Estimation

A Thesis
Presented to the Faculty of the
Department of Electronics and Communications Engineering
Gokongwei College of Engineering
De La Salle University

In Partial Fulfillment of the
Requirements for the Degree of
Bachelor of Science in Electronics and Communications Engineering

by

GUEVARRA, Alnair M.
HERNANDEZ, Roy Stephen A.
LAGMAN, Maria Josefa M.
VILLANUEVA, Andre Micayle P.

November, 2018

De La Salle University

ORAL DEFENSE RECOMMENDATION SHEET

This thesis, entitled **Driver Fatigue Detection Using an RGB-D Sensor Based On Eye Tracking and Head Pose Estimation**, prepared and submitted by :

GUEVARRA, Alnair M.
HERNANDEZ, Roy Stephen A.
LAGMAN, Maria Josefa M.
VILLANUEVA, Andre Micayle P.

in partial fulfillment of the requirements for the degree of **Bachelor of Science in Electronics and Communications Engineering** (**BS-ECE**) has been examined and is recommended for acceptance and approval for **ORAL DEFENSE**.

____

Engr. Maria Antonette C. Roque
Adviser

November 12, 2018

ABSTRACT

Driver fatigue detection is very helpful in preventing vehicular accidents. These kinds of accidents happen all throughout the day especially in countries wherein public transport systems are under development and improvement. The two visual cues that will be analyzed carefully and integrated together in this study in order to accurately detect driver fatigue are the eye state and head pose. Eye tracking is a process of detecting, observing, and analyzing eye movements and its behavior; It utilizes an image capturing system that makes use of a developed algorithm for pupil detection and gaze estimation. The images will be captured using an RGB-D sensor which can work with a microcontroller. Nowadays, this technology is becoming ubiquitous due to its growing number of applications. One of which is being able to analyze driver fatigue and the possibility of preventing road accidents. A carefully designed prototype that detects several levels of drowsiness will be mounted on strategic locations that is optimal for data collection and analysis. The levels of drowsiness will be characterized by the time the drivers eyelids are closed, number of blinks, and head poses. Moreover, the said levels are based on psychological and scientific researches. In addition to eye tracking, analysis of head pose will also be included in order to utilize the depth values from the RGB-D sensor. The head pose refers to how tilted the head is to a specific direction. Thereafter, through head pose detection, it will also serve as a processing step for eye tracking. Having two visual cues, it will significantly help in reducing the number of false alarms caused by errors in the data gathered.

Index Terms—Eye Tracking, RGBD camera.

TABLE OF CONTENTS

De La Salle University

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

Contents

1.1 Background of the Study

Road accidents has become one of the most prevalent cause of death in the Philippines wherein 26% of it was caused by driver error (Tamayo, A., 2009). Driver error includes driver fatigue, the proposed thesis aims to create a device that would alert the driver in time to help in prevention of such accidents. One way of determining whether a driver is feeling drowsy is by observing eye behaviour. One of the significant eye metrics that determines whether a driver is feeling drowsy is the frequency of eye closure exceeding one second whereas shorter ones are considered as blinks (Verwey & Zaidel, 2000). The researchers intend to do a similar thing with regards to evaluating driver drowsiness by observing the eyes. In addition to the eye metrics to be used for driver fatigue detection, head pose estimation will also be an additional parameter using RGB-D camera. As compared to the conventional RGB camera, the advantage of an RGB-D camera is its capability of capturing depth which helps in 3-D representations and modelling. These cameras augment the usual images with its depth information. The augmentation happens in per-pixel basis. Along with the RGB image captured from the camera, the eye parameters, the depth data (head pose) will be used in order to develop an algorithm for driver fatigue detection.

The eye metrics to be used in this research are blink rate and blink duration. The blink rate refers to the number of blinks that has occurred within a given time period and the blink duration refers to the time where the eyes are in the close state before entering the open state. On a normal condition, the average of the blink duration of a human would range from 100-400ms and exceeding 1000ms is considered as a microsleep (Schiffman, H., 1990). Whereas the normal blink rate would range from 0.33blinks/s to 0.25blinks/s (Nakano, T., et al., 2013). It can be seen that as the number of blinks increases over a period of time, the higher the blink rate and the lower the number of blinks over a period of time, the lower the blink rate. One past study presents the use of Kinect as the RGB-D camera to track eye gazing by using the algorithm from the thesis Eye-Model-Based Gaze Estimation by RGB-D Camera (Jianfeng & Shigang, 2014). The Kinect sensor is able to acquire the pose and 3D position of the head. In past researches, RGB camera can be used for detection of driver drowsiness by utilization of image processing (Ballesteros, P., et al, 2012). In this past research in order to come up with a conclusion that the driver is drowsy and an alarm is needed to wake up the driver, the authors based it on the eye behavior and frequent nodding of the driver.

1.2 Prior Studies

The use of eye tracking in daily lives of people is becoming ubiquitous. There have been studies that make use of eye tracking for different applications such as personality tests, focus and attention analysis, and even used in the automotives industry. There have been studies about eye tracking and its general applications. The applications presented include scientific and academic research, market research, neuroscience and psychology, psychology research, medical research, usability research, packaging research, pc and gaming research, human factors and simulation, and ophthalmology (Punde et. al., 2017).

There were also studies that made use of eye tracking in order to study and analyze data from vehicle drivers. The data gathered and analyzed can be used in order to improve road safety and security as well as prevention of road accidents. Programming softwares was used in order to test the awareness of the driver and detect drowsiness in real-time. The addition of an alarm system for drivers will ensure that they can be alerted and reminded.

The data used in these researches came from a large database known as GazeCapture as well as manually captured with the use of different kinds of cameras such as RGB-D and even webcam. The use of webcam is essentially focused on simplicity because it is already integrated in some devices such as laptops. On the other hand, the use of RGB-D camera in some researches showed that the accuracy and quality of images have improved because in addition to the RGB image captured, there is also depth, meaning the head can be detected for head pose estimation which could then be used as a parameter for driver drowsiness detection. The microcontrollers used in past researches include Arduino and Raspberry Pi. These microcontrollers were used because of their simplicity, low cost, and low power consumption. The use of microcontrollers was utilized in past researches because algorithms can be programmed easily using applicable softwares.

1.3 Problem Statement

One of the reasons for fatal traffic accidents is driver drowsiness. The researchers aim to detect driver drowsiness with three parameters: blink rate, blink duration, and head pose estimation. These three parameters will be integrated to be able to come up with the final evaluation which could then display the level of drowsiness of the driver through the use of indicator or alert the driver if the level of drowsiness reached the worst level. It will be based on the number of times the driver has nodded and if the drivers head deviates from its normal position for a long time. The problem to be resolved in this thesis is the accuracy of getting a correct evaluation of driver fatigue due to its dependency on acquiring the three parameters needed. The acquisition of the blink duration, blink rate, and head pose should also have a high accuracy so that false alarms could be avoided.

1.4 Objectives

1.4.1 General Objective(s)

The main objective of the thesis is to design and implement a system that will detect the level of fatigue of the driver and to alarm the driver once the level of fatigue reached a certain level not suitable for driving.

1.4.2 Specific Objectives

1. To develop an algorithm for detection of driver fatigue based on parameters such as blink duration, blink rate, and head pose estimation
2. To integrate eye state parameters and head pose estimation for determining the final evaluation of the level of driver fatigue
3. To utilize a microcontroller such that data acquired from the RGB-D sensor will be processed for driver fatigue evaluation
4. To create an alarm system, consisting of an indicator and a speaker, that will notify the driver of their level of drowsiness
5. To achieve at least 95% accuracy in determining the level of drowsiness: Alert, Low Alertness, Drowsy, and Sleeping

1.5 Significance of the Study

Eye tracking is very significant because almost every work and action require visual information; Eye tracking is able to track the behavior of the human eyes wherein it could observe how the human eye reacts in different situations. Some of the capabilities of eye trackers include detecting drowsiness, consciousness, and other mental states. The integration of these eye trackers to available electronic devices such as computers and mobile phones helps in analysis of consumers behavior and may lead to further technological advancements. Eye tracking technology also helps in researches and designs of fields including auto motives, medical, defense, and entertainment industries. These applications progress as time passes by wherein these progressions may root from the technology and research that this paper will provide.

In order to make the driver fatigue detection system become more accurate, it is also important to know the head position of the driver. Knowing the head position will enable the researchers to obtain more visual information regarding drowsiness level. The simplest of actions such as head swaying and nodding is already enough to know if the driver is exhausted. One example to know if the driver is exhausted includes the head not in the frontal position for an extended period of time. This means that the head is already tilted to one side or nodding off.

By understanding the users eye and head behavior, new safety and security measurements may be designed. In addition, there will be improvements on existing work based on data gathered from eye and head movements. The beneficiaries of this study include the user, the society and industries that make use of visual data for improvements, and future researchers as well. The user and the industries that continue to improve their products and devices help each other in such a way that the safety, security, and satisfaction of the users will be achieved given that their visual data and gaze patterns will be carefully analyzed by different industries. This study also helps in improving and testing of existing algorithms for eye tracking and head pose estimation. Therefore, this study will be of great help to future researchers that will engage and tackle this topic because eye tracking and head pose estimation is important and significant in this modern age.

1.6 Assumptions, Scope and Delimitations

1.6.1 Assumptions

1. The existing dataset to be used upon developing the system will come from a research study about eye tracking conducted by Krafka et. al. (2016) from Massachusetts Institute of Technology (MIT)

2. The RGB-D camera will have a difficulty in detecting the eyes if the person being evaluated has little eye openings

3. The device will treat head movements as nods or head tilts due to road humps or cracks which can affect the accuracy and reliability of the system

1.6.2 Scope

1. The device can measure the blink rate and blink duration of the eyes as additional parameters in determining whether the driver is drowsy or not.

2. It can also detect driver fatigue by using head pose estimation

3. The device should be able to detect and classify the level of drowsiness of the driver.

4. Evaluation of driver awareness using the device is still accurate even on low light conditions

5. The prototype will be tested in a laboratory setting and in a vehicle
6. The size of the RGB-D sensor to be placed at the dashboard of the vehicle will not exceed the 4 inches height allowance of the distracted driving law

1.6.3 Delimitations

1. The thesis will only be tested on standard cases, these does not include people with glasses, face or eye disorders, and eye irritation
2. The device will be placed inside closed/private vehicles to maximize the effectiveness of the alarm system
3. The system will not be interfaced with the computer box of the vehicle

1.7 Description and Methodology

Visual data gathered by the RGB-D camera will be used to study and analyze the eye movements and head pose of the driver. Moreover, there will be an alarm system when the camera detects driver fatigue which is not suited for driving a vehicle.

The head pose is estimated using the RGB-D camera and a head coordinate system is generated. If the system has detected that the driver has deviated from the reference or the normal position of the head for a certain time while the eye parameters are being monitored, then the device will evaluate if the driver has nodded off and an alarm will turn on to wake up the driver from sleeping. The head pose of the driver will serve as a basis for evaluation of driver fatigue along with the gathered blink rate and blink duration. The different head poses based from the estimated head pose results are given below in Figure 1.1.

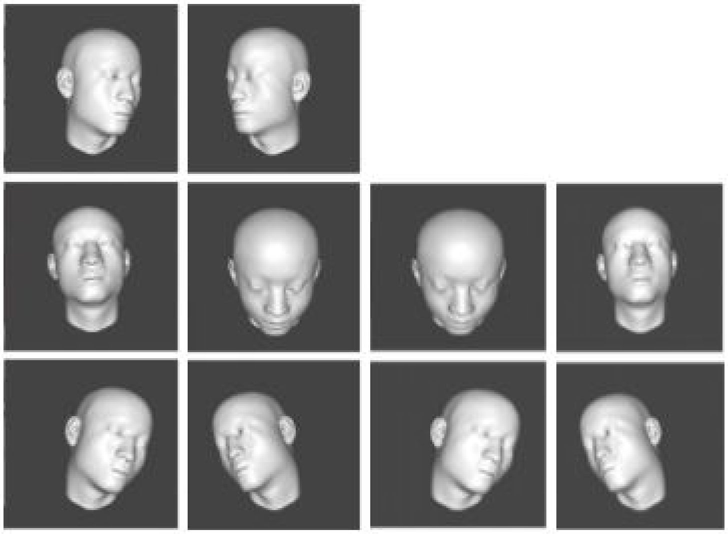

Fig. 1.1 Different head poses based from estimated head pose results (Liu, F. et. al., 2015)

TABLE 1.1 EVALUATION OF LEVEL OF DROWSINESS

Level of Drowsiness	**Description**
Awake	Normal blink rate (0.25 - 0.33 blink/s) Normal blink duration (100-400ms) Or Low blink rate (less than 0.25blink/s) Normal blink duration(100-400ms)
Low Alertness	Normal blink rate (0.33blink/s - 0.25blink/s) Long blink duration(400ms - 1s) Or Low blink rate (lower than 0.25blink/s) Long blink duration(400ms - 1s)
Drowsy	High blink rate (0.33blink/s) Long blink duration (400ms-1s) Or Normal blink rate Long blink duration (400ms - 1s)
Sleepy	Low Blink Rate (less than 0.25blink/s) Very long Blink Duration (greater than 1s)

It is evaluated into four levels such as Awake, Low Alertness, Drowsy, and Sleepy. Each level of drowsiness will have an indicator and if the driver is evaluated as sleepy or drowsy an alarm will sound off to warn the driver.

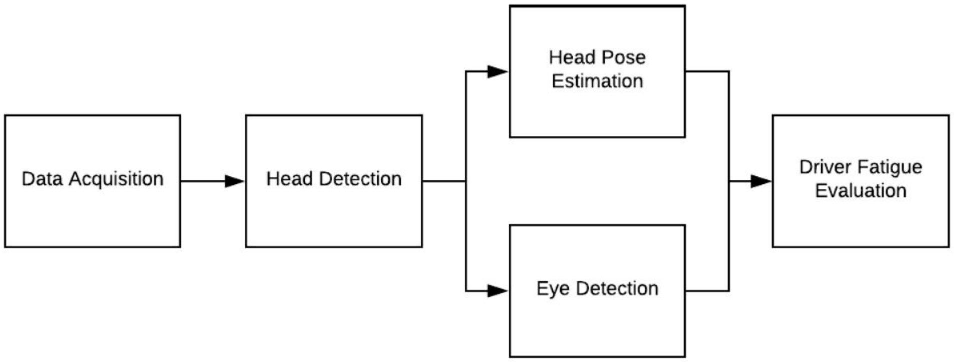

Fig. 1.2 General Framework of Driver Fatigue Evaluation System

The RGB-D camera is continuously taking a video of the face of the driver and frames will be captured. If the face and head is now detected, the edges of the faces will then be taken in order to find the location of the eyes. After eye detection, blink rate, blink duration, and head pose was calculated for evaluation of driver fatigue. The algorithm to be used for the evaluation of fatigue will be classifiers. Some examples that can be used as classifiers are machine learning algorithms, fuzzy logic, etc.

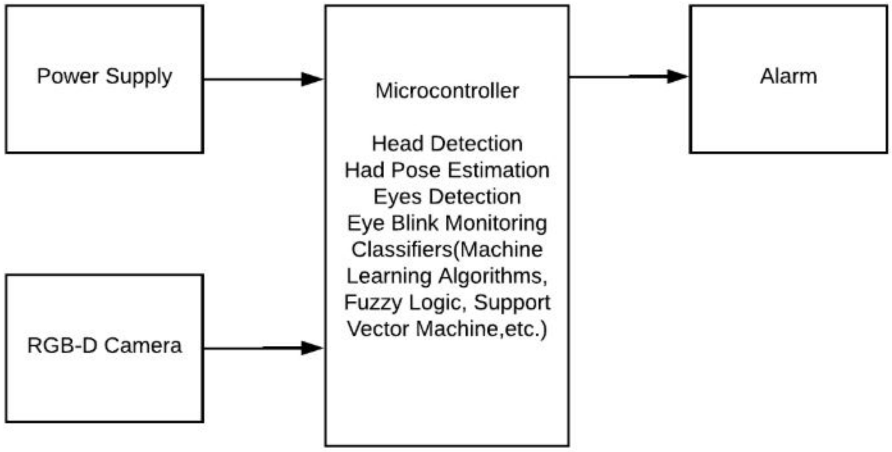

Fig. 1.3 Proposed Block Diagram of Driver Monitoring System

The RGB-D Camera will serve as the sensor for the eye movements and head movements. The microcontroller will then process the data acquired by the RGB-D. If the system has detected that the drivers fatigue level is high, the alarm system will turn on and notify the driver.

1.8 Overview

1.8.1 Gantt Chart

De La Salle University

The following figures shown below is the individual gantt chart of the researchers:

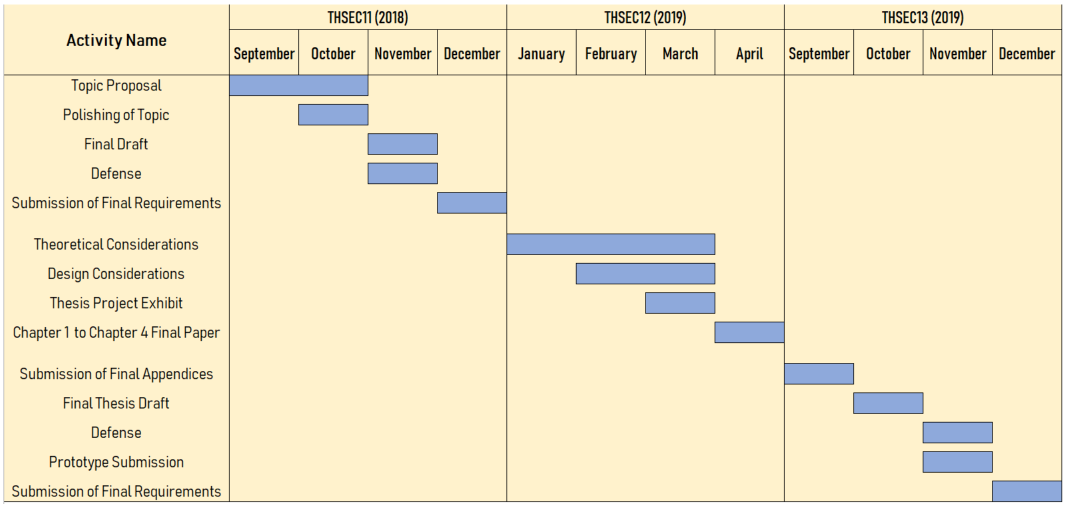

Fig. 1.4 Gantt Chart of Guevarra

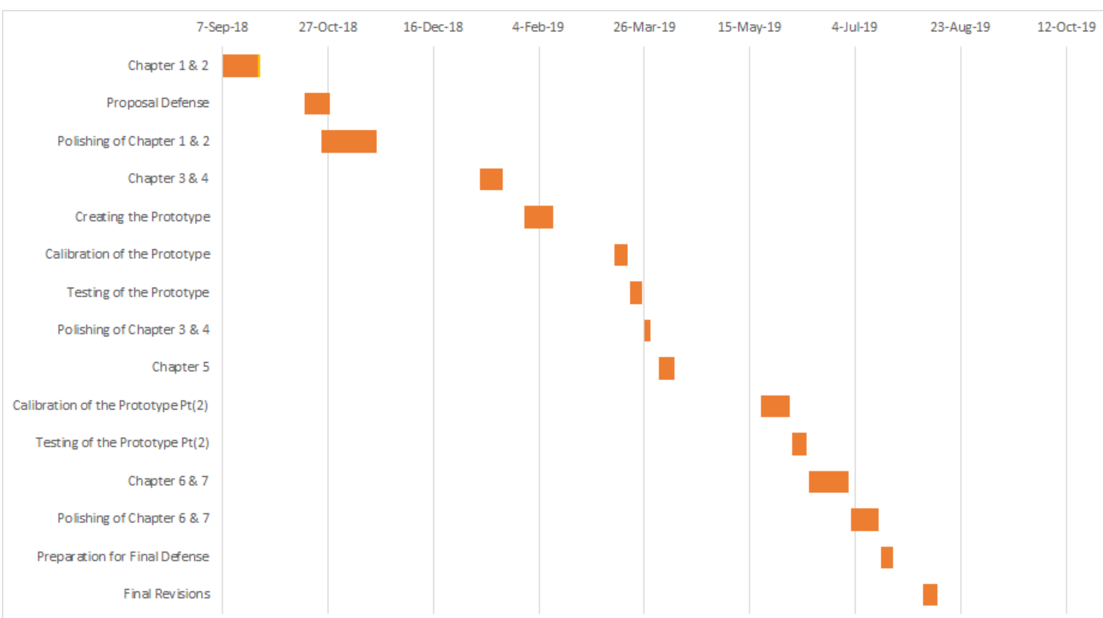

Fig. 1.5 Gantt Chart of Hernandez

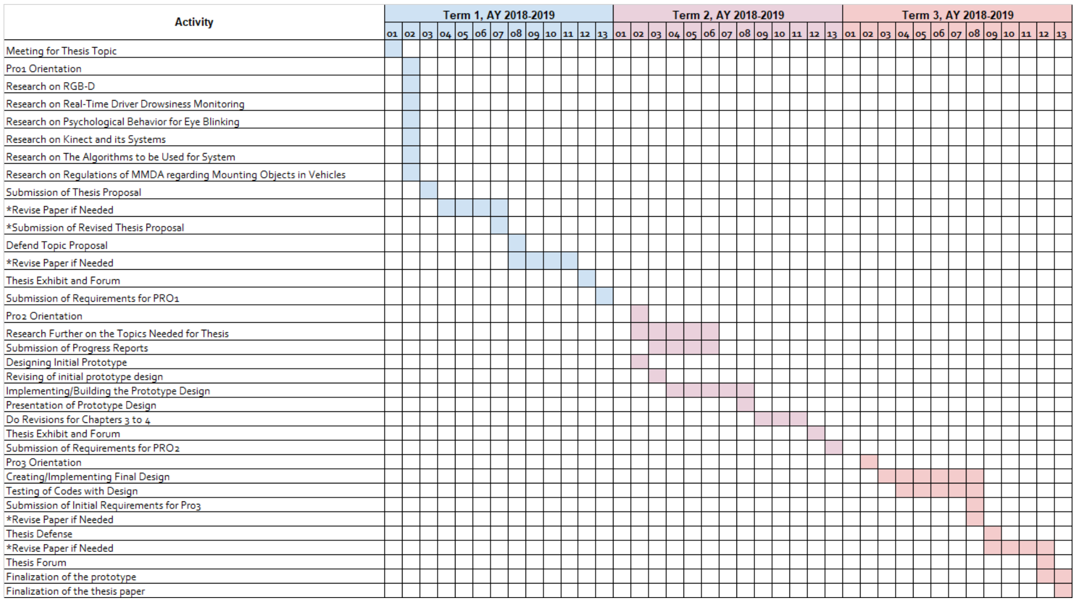

Fig. 1.6 Gantt Chart of Lagman

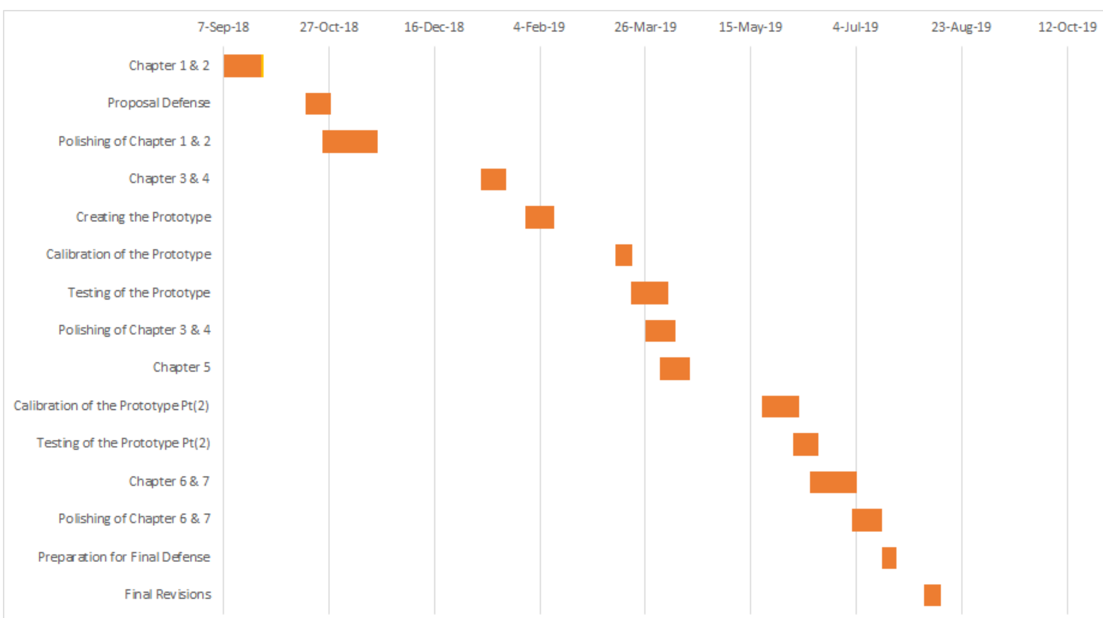

Fig. 1.7 Gantt Chart of Villanueva

1.8.2 Estimated Work Schedule and Budget

TABLE 1.2 PRICE LIST OF MATERIALS TO BE USED

Materials	Estimated Price(in PHP)
RGB-D Sensor	Php10,000-20,000
Microcontroller	Php450-25k
Laptop	Php15k-100k
Total Price	Php25k-140k

Chapter 2

LITERATURE REVIEW

Contents

2.1 Existing Work

2.1.1 Driver Awareness

Sleepiness can cause accidents in peoples daily lives (Cuentas & Gonzales, 2017). Sleepiness affects activities that require short visual and motor reaction times such as driving a car. The highest number of accidents in these kinds of activities are vehicular accidents according to World Health Organization (WHO). Eyelashes could be an indicator of sleepiness (Cuentas & Gonzales, 2017). Through image processing, a good eye tracker will be able to detect eyelashes and eye movements in order to determine if a person is sleepy or not. Driver awareness and fatigue also shows that a driver is powerless to stay awake (Ahmed et. al., 2015). Advancements in technology in order to prevent accidents are a challenge (Ahmed et. al., 2015). Eye trackers are considered to be parts of these advancements. Fatigue is another factor in these vehicular accidents since it affects the focus and attention of the driver. It also affects the movements and decision making of the driver which makes it a serious problem in driver awareness. By being able to prevent and monitor driver fatigue, there will be a huge reduction in the number of vehicular accidents (Cuentas & Gonzales, 2017).

2.1.2 Eye Tracking and its Applications

Eye tracking gives information on where do people look and what they ignore. The concept of eye tracking is very basic but the process and implementation can be very diverse and complex (Punde et. al., 2017). Vision is the most important sense (Punde et. al., 2017). In this study, detecting the behavior of drivers when driving especially at night is vital. By using eye tracking technology, the drivers focus, attention, and presence can be analyzed and studied. The eye trackers are important part of this study because it measure visual attention and has become an important research tool (Jadhav, 2017). The different eye tracker metrics are vital in studying the behavior of the drivers. These metrics include the gaze point, fixation, smooth pursuit and saccades. The gaze point is the basic unit of measuring eye movement. The fixation is a cluster of gaze points. The smooth pursuit is a slow eye movement that allows the eye to follow an object moving slowly. On the other hand, the saccade is a rapid eye movement between fixations. By knowing these metrics, the drivers eye movements can be studied more. The slow and rapid eye movements of the drivers eyes can lead to various inferences and conclusions.

Eye tracking is important in many aspects (Krafka et. al., 2016). The number of these aspects grow as technology advances. Even though there are so many aspects wherein eye tracking is applicable, it has yet to become an ubiquitous technology (Krafka et. al., 2016). Meaning, eye tracking technology is not yet widespread or being used globally unlike the technology of voice assistants and speech recognition. The application of eye tracking may range from human-computer interaction, medical diagnoses, psychological studies, and computer vision (Krafka et. al., 2016). It was strongly stated that the human gaze is the externally-observable indicator of human attention (Khosla, 2016). The use of eye tracking requires a device where it will be interfaced for it to reach its full capabilities and functionality. This technology does not require additional devices from users because it can be interfaced on devices that they already have such as mobile phones and computers (Krafka et. al., 2016). A convolutional neural network was interfaced to a mobile device. The results with and without calibration were compared. The results show that a prediction error of 1.34cm to 2.12cm can be obtained with calibration. Without calibration, the error goes up to 1.71cm and 2.53cm.

Non-intrusive ways of integrating eye tracking with other devices are better than intrusive ways (Nguyen, 2009). It is because non-intrusive ways do not require devices that cause distraction and annoyance to the user. Intrusive ways include using wearables such as headgears with cameras. On the other hand, non-intrusive ways of eye tracking include having a camera that is from a specific distance away from the user. There were three methods used in predicting eye gaze. These methods include detection using Haar-based features, Lucas Kanade method for tracking eye, and Gaussian process. The Haar-based method used thousands of samples of views and trains classifiers to detect object rapidly (Nguyen, 2009). On the other hand, the Lucas Kanade method uses set of pyramid representations of the original image and use the conventional tracking algorithm for each image pyramid (Nguyen, 2009). Lastly, the Gaussian process was used in eye tracking in such a way that given the gaze data of the user, the mean and covariance of these functions will be used to make predictions. (Nguyen, 2009).

2.1.3 Eye Tracking using Different Cameras (RGB-D and Web-cam)

Eye gaze tracking systems are usually based on the use of infrared lights. However, such systems may not work outdoor or may have a very limited head box for them to work (Xiong et, al., 2014). This research paper proposes a non-infrared based approach in order to track the eyes with the use of an RGBD camera using Kinect. Their method uses a personalized 3D face model. The system tracks the center of the iris and a set of 2D facial landmarks whose 3D locations are provided by the camera. (Xiong et, al., 2014). They used a non-IR based method because there are less-strict requirements for hardware and it is easier to integrate with laptops and tablets (Xiong et, al., 2014). Without the use of IR lighting, the visibility of the pupils were decreased, so they replace the task of pupil detection with iris detection. They compared the performance of using RGB from using RGBD and found out that there is more accurate tracking in RGBD. However, the results of their non-IR based method are not as accurate as the ones from the IR based approach.

The only RGB-D device used in this project is a Kinect sensor for gaze estimation for humans. The Kinect is set with a head coordinate system where the pupil center is detected from the image. The next step then involves calibration of the eyeball center where the system requests the human to gaze at a position creating a 3D eye model. The direction of the eye gaze is computed with the obtained calibrated eyeball center and pupil center. In their proposed method, it involves three main steps namely the face pose estimation, pupil center estimation, eyeball center calibration, and gaze estimation. The face pose estimation method uses the algorithm provided by Microsoft Kinect where additional devices are not needed to be able to estimate the head pose correctly. In pupil center estimation, an algorithm done by Febian Timm and Erhardt Barth is used that locates the pupil center by assuming that the pupil center P and the eyeball center C is a constant K. The eyeball center calibration is done by having two vectors: the target and the pupil center where the angle is obtained and these parameters are inputted to the given equation to calculate the eyeball center calibration. Lastly, the gaze estimation is obtained by the given parameters from the first three methods mentioned. Their project worked well with minimal error on the accuracy of the algorithm on detecting pupil center. (Jianfeng & Shigang, 2014).

In this research, they implemented an eye ball tracking system for patients that cannot perform any voluntary tasks related to daily life. A real time data stream of the eye movement is captured using a webcam that transfers data serially to MATLAB. (Mazhar et, al., 2015). After that, a sequential image processing scheme segments the iris of the eye and calculates the centroid. Their idea was to create a general eye tracking tool that can be interfaced with wheelchairs, moveable patient beds, alerting mechanisms for disabled ones etc. (Mazhar et, al., 2015). The hardware that they used was a webcam by Logitech which can be interfaced serially using a USB port of a computer. This method involves an intrusive system because the camera is mounted to a headband. Their eye detection algorithm involved image acquisition through the use of the webcam, then plane separation which is the RGB plane, then thresholding (segmentation), then open binary image, then structuring, then dilation, and finally centroid calculation.

2.1.4 Eye Tracking Applied on Drivers Awareness and Eye Behavior

Road accidents have been prevalent in India and has shown consistency in increasing each year. The highest percentage as to why road accidents occur are due to the drivers themselves. This project aims to determine driver drowsiness by using a webcam in determining the eye blinking action of the driver using an algorithm that is implemented using cascade object identifier from vision toolbox of MATLAB. Their system will take the recorded video intro frames where these frames are then processed one by one, which results to a more accurate eye detection as oppose to eye tracking. To calculate the eye blinking, they have used an equation to calculate average drowsiness, number of closed eye found over number of frames; A closed eye would equate to zero eye blinking and one if the eyes are partially or fully open. If the percentage of drowsiness is set to be more than the threshold value, an alert signal will be generated to the driver. Two systems were used to test the research: System I: Inspirion 15 (DELL), core i3 processor (64bit) with 4GB of RAM, VGA inbuilt web cam, MATLAB 13; and System II: Inspiron 15 (DELL), core 13 processor (64bit) with 8GB of RAM, 3 MP USB web cam, MATLAB 13. About 1000 frames are processed online and 1500 frames are processed offline; while both systems show efficient results in overall drowsiness detection, System II has showed a more promising result as compared to System I. (Ahman & Borole, 2015).

The objective of the project is to be able to simultaneously observe closed eyes and open mouth, when positive for yawning, the driver is alerted with a buzzer. A real-time video is captured by placing a camera in front of the driver using OpenCv, Viola Jones Algorithm, and Contour finding Algorithm are used in face detection. There are multiple detections considered in their system but the four main steps are as follows: face and eye detection, eye blinking detection, mouth detection, and warning system design. To detect face and eyes, they have used Viola Jones and Region of Interest (ROI) to locate the eyes. In eye blinking detection, the researchers only considered detecting the right eye in assumption that both eyes blink at the same time. The method used to get the eye blinking rate uses a ring buffer that calculates eye blinking rate every after first 100 first frames, writing 1 to the buffer when eye blinking is detected and 0 if otherwise. When the mouth is detected, decision of yawning is done by obtaining the difference between largest and smallest Y-coordinate values from the mouth; If the height surpasses the given threshold, the person is yawning. Their experiment lasted for 20 days that is divided into 6 sections: morning, afternoon, a given critical time, evening, night, and another given critical time. They also tested this on different test subjects based on sexuality, a person having or not having a mustaches and/or glasses, and age, giving a general best time result to use the system for each test subject.

2.1.5 Eye Tracking Implementation using Microcontrollers

The proposed system for eye tracking is implemented in real time using an arduino uno microcontroller and a zigbee wireless device. (Venugopal et, al., 2016). Two types of infrared eye following procedures are used which are enhanced pupil and dim pupil. Their system consists of three general modules. The first module consists of eye tracking and detection in real time using Viola-jones algorithm. The second module involves the detection of the pupils using Hough transformation. The third module comprises the detection and tracking of the pupils and setting a direction. The obtained direction is sent to the Arduino with a wireless zigbee device for transmission of data. (Venugopal et, al., 2016). Basically, the three modules consists of image pre-processing, pupil region detection and pupil detection with movement detection.

The main goal of developing the driver monitoring system using eye gaze tracking is to reduce road accidents and improve road safety. The algorithm detects the drowsiness of the driver and alerts the user using sound and vibration from steering. (Mavely et, al., 2016). The control of the whole framework operation is based on the Raspberry Pi 2. The project uses a Logitech c270 HD webcam and an LED IR illuminator for capturing an image. The camera is connected to the Pi using a USB port and it is placed above the steering wheel. An IR illumination is used to capture the users eyes during the night. For eye gaze tracking, CAMSHIFT (Continuously Adaptive Mean Shift) algorithm was used to extract the ROI (Region of Interest). The subjects drowsiness was detected if the eye gaze is not detected more than 2 secs. Their proposed method was compared with other types of approaches and found out that their method was very cost effective with a high rate of speed and accuracy (Mavely et, al., 2016). They compared it with the bioelectric signal and driver face monitoring methods.

2.1.6 Head Pose Estimation

This research focuses on a real-time driver fatigue detection system using an RGB-D camera where visual cues such as head pose and eye state were used to detect driver drowsiness. They did the 3D head pose estimation by making use of the RGB-D data caught by the camera. If the drivers head is not in its normal position for a long period of time or it happens frequently , then it will conclude that the driver is drowsy. An alarm is used to alert the driver after detecting driver fatigue. They also used eye states to further increase the accuracy of driver fatigue detection. In order to detect eye states, they used WLBP (Weber Local Binary Pattern) which is a local image descriptor. The performance of eye state detection was evaluated by collecting eye image samples from the RPI ISL Eye Training Database. After obtaining the data for the head pose and eye states, they used a support vector machine (SVM) in order to classify data for correct detection of driver fatigue. The combination of head pose and eye state has proven to increase the effectiveness of the system.

Another method for head pose estimation is to use Active Appearance Models (AAM) and Pose from Orthography and Scaling with Iterations (POSIT) (Dongare & Shah, 2017).

AAM is an object model that is composed of statistical data based on the shape and texture of the object. The AAM will produce a 2D model of the head pose based on its algorithm. However, 2D models are not totally accurate that is why POSIT will be used to generate 3D models (Dongare & Shah, 2017).

2.1.7 Head Detection

Head detection plays an important role in computer vision applications (Luo & Sang, 2009). Classifiers are proven to be very effective in object detection. However, these classifiers should have high detection rate and low false-positive rates in order to be reliable (Luo & Sang, 2009). This is where the problem lies. Online boosting can be applied in computer vision problems due to its strong adaptability (Luo & Sang, 2009). Online boosting can be done in such a way that the detector will initially be calibrated offline and self-adjust from online boosting when the environment changes. Weak classifiers will be improved through online boosting because there will be more an more training data that will be available. The object detected in online boosting can be verified through tracking. Tracking improves the detection rate (Luo & Sang, 2009). The tracking algorithm can be based from posterior probability which can be used to determine credibility coefficients of detected objects. Basically, the online boosting and tracking method work hand in hand to improve each other. Online boosting deals with the detection and tracking deals with the verification and increasing detection rate. Moreover, in order to use head detection with computer vision, the methods to be used must be effective during illuminant alterations and changes in its background (Ishii et al., 2004). Face detection methods that show good performance under various lighting conditions are already available. These different methods have their own edges over the others in terms of accuracy, efficiency, and cost. Notable methods include utilization of four directional features (FDF) and linear discriminant analysis (LDA). FDF is considered to be one of the most robust method to distinguish patterns (Ishii et al., 2004). The FDF represents the vertical, horizontal, and diagonal edge features of the image. FDF extraction follows a procedure. The procedure includes application of Prewitts operator in the four directions and the outputs will be fed to Gaussian filters in order to have low noise and deformation. Lastly, feature vectors will be extracted from the four edge images. On the other hand, the LDA will be used in order to create a discriminant space that will minimize the within-class variance and maximize the between-class variance of the images. The LDA method reduces the computational cost of head detection because the optimal class coefficient matrix will be obtained by solving the eigenvalue problem. This said matrix will then generate the optimal feature from primitive features. Therefore, by limiting the number of primitive feature vectors to a low number, the computational cost will also be lower (Ishii et al., 2004). In addition, the LDA finds a linear combination of features that characterizes two or more classes of objects which make it useful for pattern recognition and for object detection.

2.2 Lacking in the Approaches

2.2.1 Eye Tracking and its Applications

For the implementation done by Krafka et. al., eye tracking technology were only integrated to mobile devices. This means that the implementation used a small form factor camera that can only be integrated to small devices like mobile phones. Using GazeCapture as its database, the data is from crowdsourcing (Krafka et. al., 2016). Crowdsourcing was used in order to deviate from prior works which made use of original data from cameras included in their eye trackers. The advantage of having a large database is that models can be tested repetitively for robustness (Krafka et. al., 2016). The eye tracker used showed results before and after its calibration. The algorithm used in this specific eye tracker can only be used if the data is from a given database. However, Krafka (2016) strongly stated that this eye tracker outperformed state-of-the-art eye trackers by a large margin.

An introduction made by Punde et. al. (2017) shows that there are two types of eye trackers. These types include remote or screen based eye trackers and head-mounted or mobile eye tracker. The remote eye tracker requires the user to sit in front of the screen to interact with screen based content. On the other hand, head-mounted eye trackers does not distract users and lets them move freely. From their study, there are no direct applications per type of eye trackers and generalizations about both types of eye trackers were presented. This study was included in our research because of applications in automotives industry. Based on the eye trackers data, industries may improve their current products and devices. In addition, the visual data may be used for the industries analysis of eye movements and behaviors of users such as those of the psychological and medical applications (Punde et. al., 2017). This study is a helpful overview of eye tracking and its applications to different fields and practices. Even though there are no specific target field, the overall use, functionality, advantages, and disadvantages of eye trackers were presented clearly.

2.2.2 Head Pose Estimation

Through computer vision techniques, the detection of eye state, head pose, and facial features are easier. However, in the study conducted by Liu (2015), facial features were not part of their parameters for driver fatigue. Even though computer vision techniques nowadays are more advanced, the accuracy of detecting eye state and head pose is not guaranteed due to different illumination conditions and various facial features for different persons (Liu et. al, 2015). In addition, other facial actions were not taken into consideration in this study such as yawning or when the person being tested and monitored has face disorders that make it difficult for the RGB-D sensor to capture accurate data. Moreover, head pose estimation is very dependent on the RGB-D sensor since an RGB sensor cannot have depth values. These values are substantial and beneficial in driver fatigue detection since it allows another visual cue to be analyzed.

2.3 Summary

The proponents review of related literature were about eye-tracking technology, head pose estimation, and driver awareness. Driver awareness refers to the focus, attention, and functionality of the driver. Sleepiness and fatigue were considered to be major factors in vehicular accidents. The number of accidents can be minimized when these factors are being monitored using advancements in technology such as eye trackers and fatigue detectors. In the applications of eye tracking, it was discussed that eye tracking can be used to help persons with disabilities, to study the psychology of people, and to interact with computers easily. In the usage of RGB-D camera/webcam, their differences were seen by their quality of RGBD detection. In the utilization of drivers eye behavior in eye tracking, it was researched that there were studies to measure driver alertness and it can be based to make an effective eye tracking device for drivers. Lastly, the implementation of eye trackers using microcontrollers were dealt with because the proponents want to focus on having a lighter and compact device for eye detection rather than having laptops/computers as the brains of the eye tracking device. The literatures about these involved different types of microcontrollers like arduino/raspberry pi that were used to implement the eye tracking device. For head pose estimation, the use of an RGB-D sensor was emphasized to ensure that disadvantages of an RGB sensor will be compensated. By using an RGB-D sensor for head pose estimation, different illumination conditions will have the least effect on the captured images such as having depth values will make it easier detecting facial features. Having head pose estimation as another visual cue for driver fatigue is vital in having more accurate data and reduction of false alarms due to errors in data gathering. In addition by using Weber Local Binary Pattern (WLBP) for the images, the effectiveness and robustness were emphasized because the results showed that WLBP is better than other image descriptors. Moreover, the head pose helped in classifying the fatigue levels of the driver using a support vector machine (SVM).

De La Salle University

References

Ahmad, R., & Borole, J. (2015). Drowsy driver identification using eye blink detection. *IJISET-International Journal of Computer Science and Information Technologies*, *6*(1), 270–274.

Dongare, S. (2017). Eye tracking and head pose estimation for driver monitoring on raspberry pi. *International Journal of Advances in Electronics and Computer Science*, *4*(3), 4.

Ishii, Y., Hongo, H., Yamamoto, K., & Niwa, Y. (2004). Face and head detection for a real-time surveillance system. In *Pattern recognition, 2004. icpr 2004. proceedings of the 17th international conference on* (Vol. 3, pp. 298–301).

Jianfeng, L., & Shigang, L. (2014). Eye-model-based gaze estimation by rgb-d camera. In *Proceedings of the ieee conference on computer vision and pattern recognition workshops* (pp. 592–596).

Krafka, K., Khosla, A., Kellnhofer, P., Kannan, H., Bhandarkar, S., Matusik, W., & Torralba, A. (2016). Eye tracking for everyone. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 2176–2184).

Kumar, N., & Barwar, N. (2014). Detection of eye blinking and yawning for monitoring drivers drowsiness in real time. *Int. J. Application or Innovation in Engineering & Management*, *3*(11).

Luo, D., & Sang, N. (2009). Improving head detection from tracking. In *Information engineering and computer science, 2009. iciecs 2009. international conference on* (pp. 1–4).

Matouq, J., Al-Nabulsi, J., Al-Kazwini, A., Baniyassien, A., Al-Haj Issa, G., & Mohammad, H. (2014). Eye blinking-based method for detecting driver drowsiness. *Journal of medical engineering & technology*, *38*(8), 416–419.

Mavely, A. G., Judith, J., Sahal, P., & Kuruvilla, S. A. (2017). Eye gaze tracking based driver monitoring system. In *Circuits and systems (iccs), 2017 ieee international conference on* (pp. 364–367).

Mazhar, O., Shah, T. A., Khan, M. A., & Tehami, S. (2015). A real-time webcam based eye ball tracking system using matlab. In *Design and technology in electronic packaging (siitme), 2015 ieee 21st international symposium for* (pp. 139–142).

Nakano, T., Kato, M., Morito, Y., Itoi, S., & Kitazawa, S. (2013). Blink-related momentary activation of the default mode network while viewing videos. *Proceedings of the National Academy of Sciences*, *110*(2), 702–706.

De La Salle University

Nguyen, B. L. (2009). Eye gaze tracking. In *Computing and communication technologies, 2009. rivf'09. international conference on* (pp. 1–4).

Păsărică, A., Bozomitu, R., Cehan, V., Lupu, R., & Rotariu, C. (2015). Pupil detection algorithms for eye tracking applications. In *Design and technology in electronic packaging (siitme), 2015 ieee 21st international symposium for* (pp. 161–164).

Punde, P. A., Jadhav, M. E., & Manza, R. R. (2017). A study of eye tracking technology and its applications. In *2017 1st international conference on intelligent systems and information management (icisim)* (pp. 86–90).

Schiffman, H. R. (1990). *Sensation and perception: An integrated approach.* Oxford, England: John Wiley & Sons.

Singh, H., Bhatia, J. S., & Kaur, J. (2011). Eye tracking based driver fatigue monitoring and warning system. In *Power electronics (iicpe), 2010 india international conference on* (pp. 1–6).

Tamayo, A. (2009). Occurrence of traffic accidents in the philippines: An application of poisson regression analysis.

Vargas-Cuentas, N. I., & Roman-Gonzalez, A. (2017). Facial image processing for sleepiness estimation. In *Bio-engineering for smart technologies (biosmart), 2017 2nd international conference on* (pp. 1–3).

Venugopal, B., & Dsouza, D. (n.d.). Real time implementation of eye tracking system using arduino uno based hardware interface.

Verwey, W. B., & Zaidel, D. M. (2000). Predicting drowsiness accidents from personal attributes, eye blinks and ongoing driving behaviour. *Personality and individual differences*, *28*(1), 123–142.

Xiong, X., Liu, Z., Cai, Q., & Zhang, Z. (2014). Eye gaze tracking using an rgbd camera: a comparison with a rgb solution. In *Proceedings of the 2014 acm international joint conference on pervasive and ubiquitous computing: Adjunct publication* (pp. 1113–1121).

Zhang, L., Liu, F., & Tang, J. (2015). Real-time system for driver fatigue detection by rgb-d camera. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *6*(2), 22.

Zhang, Y., Xin, W., & Miao, D. (2015). Personality test based on eye tracking techniques. In *Affective computing and intelligent interaction (acii), 2015 international conference on* (pp. 832–837).

De La Salle University

Produced: November 12, 2018, 02:44