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 MyRio. 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 driver's 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, MyRio.

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 errors (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 fatigue 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

TABLE 1.1 LIST OF PREVIOUSLY DONE STUDIES

Past Study (Title, Author)	Brief Description	Advantages	Improvements of Our Study
Real-Time System for Driver Fatigue Detection by RGB-D Camera (Zhang, L., Liu, F. & Tang, J., 2015)	Driver fatigue —detector that makes use of an RGB-D sensor to measure eye state parameters and head pose in real-time	1.Real-Time Detection 2.Outputs are not based on one parameter only 3.Utilizes an RGB-D sensor for its depth data that serves as extra evidence for parameter detection 4.Using WLBP (Weber Local Binary Pattern), the accuracy of detection is at 82%	1.The system will be trained and calibrated using an existing database 2.Intensity images will not be used as inputs to the system in order to tolerate illumination changes
Eye Tracking for Everyone (Krafka, K., et.al, 2016)	A convolutional neural network, iTracker, was designed and implemented for eye tracking while utilizing GazeCapture which is a large-scale dataset	1.Utilizes a large-scale dataset. 2.Crowd sourcing was the method of data collection 3.Robust eye state tracking due to proper calibration	1.Reliability will be improved by adding head pose parameters 2.High variability will be encouraged to reduce the number of calibration that are necessary

Drowsy Driver Identification Using Eye Blink Detection (Ahmad, R. & Borole, J., 2015)	A non-intrusive machine vision-based concept was used to design and implement a drowsiness detection system based on the blink rate	1.All possible actions are considered and outputs are generated accordingly. 2.The parameters are not limited to the eyes. 3.Driver's attention is also considered 4.Drowsiness detection accuracy is currently at 95% with processing time of 2 sec.	Additional eye state parameter will be introduced in our study which is the blink duration
Real Time Implementation of Eye Tracking System Using Arduino Uno Based Hardware Interface (Venugopal, B. & D'souza, D., 2016)	Implemented a system using Arduino Uno and Zigbee wireless device for eye tracking and movement estimation	1.Has movement detection feature 2.Transmission of data is wireless	1.An additional device will not be required because the microcontroller to be used has built-in wireless capabilities.
Eye Gaze Tracking Based Driver Monitoring System (Mavely, A., Judith, J. & Kuruvilla, S., 2017)	Eye gaze tracking system for driver drowsiness was implemented using Raspberry Pi.	1.Notification using audio and vibration on the steering wheels. 2.Accuracy is high in any light conditions. 3.Less expensive	1.Different levels of drowsiness will be analyzed for different alarm notifications 2.Drowsiness detection will also be based on head pose parameters such as the tilting of the head.

Face and Head Detection for a Real-Time Surveillance System (Ishii, Y., Hongo, H., Yamamoto, K. & Niwa, Y., 2004)	Two classifiers were used in order to detect the face and head effectively which is then implemented using a PC and run at approximately 10fps for VGA input	1.Head detection is at 87.2% and its reliability is at 83.6% 2.Makes use of Four Directional Features for robust detection of patterns	1.False alarms will be greatly reduced because the distance between the sensor and the subject is ideal 2.Training procedure for the system will be improved because of the large-scale dataset

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 driver's 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 proposed thesis is to design a prototype that will utilize an RGB-D sensor that will be interfaced to a microcontroller for detection of blink rate, blink duration, and head pose estimation for evaluation of driver fatigue.

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 drowsiness of the driver
3. To create an alarm system that will notify the driver of their level of drowsiness.
   - To be able to light up the LED which will serve as an indicator of drowsiness level
   - To be able to alarm the sound of 25 dB for level of drowsiness and 35 dB for level of sleeping
   - To be able to automatically deactivate the alarm once the device detects that the driver is already on the alert level
4. To achieve at least 80% 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 consumer's 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 user's 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. There is an existing data for different blink rates, blink duration, and head pose estimation available for comparison.
2. The driver will stay awake after the alarm
3. Eyes are subjected under normal environment such that the normal blink rate and blink duration are not affected
4. There will be an error in fatigue estimation when driver continuously nods in situations such as jamming to music, passing by rumble strips, and the like.

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

1.6.3 Delimitations

1. The RGB-D camera will have a difficulty in detecting the eyes if the person being evaluated has little eye openings
2. The thesis will only be tested on standard cases; it will not include people with eye disorders
3. The alarm placed inside a public vehicle will not be as effective as it is compared inside a private vehicle due to environmental noise

1.7 Description and Methodology

Visual data will be gathered using an RGB-D camera. The data gathered will be used by the RGB-D camera that follows a specific algorithm for each step of eye tracking. This eye tracker will be used to study and analyze the eye movements of the driver. Moreover, there will be an alarm system when the eye tracker detects driver fatigue which is not suited for driving a vehicle. In addition, the RGB-D sensor placement is at a position permitted by the law.

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 serves 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.

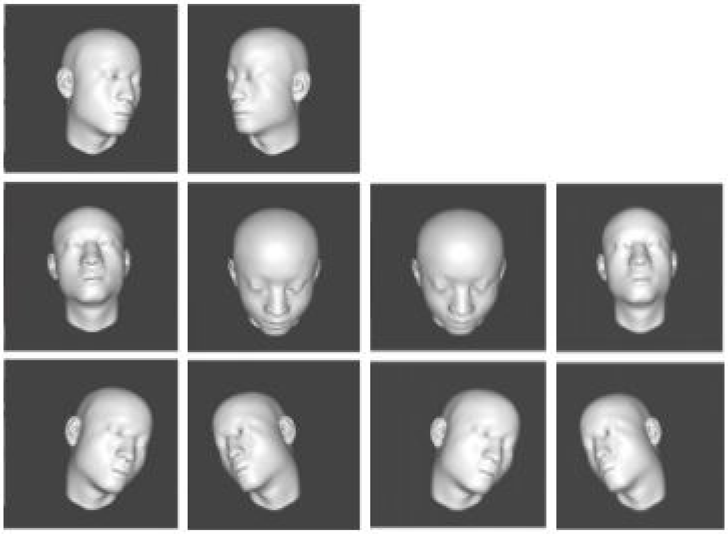

Fig. 1.1 Different head poses based from estimated head pose results (Liu, F. et. al., 2015)

TABLE 1.2 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(100-400ms)
Drowsy	High blink rate (greater than 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)

Using Labview, an algorithm will be made for the evaluation of driver fatigue. It is evaluated into four levels such as Awake, Low Alertness, Drowsy, and Sleeping. Each level of drowsiness will have an indicator and if the driver is evaluated as sleeping or drowsy an alarm will sound off to warn the driver.

De La Salle University

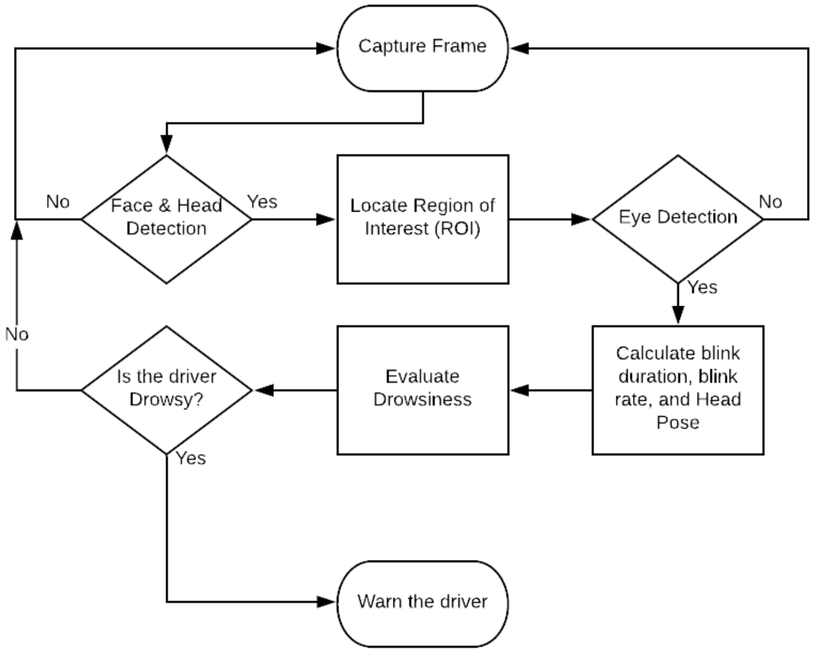

Fig. 1.2 Evaluation of level of Drowsiness based on blink rate and blink duration

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 drowsiness. The algorithm to be used for the evaluation of drowsiness will be classifiers. Some examples that can be used as classifiers are machine learning algorithms, fuzzy logic, etc.

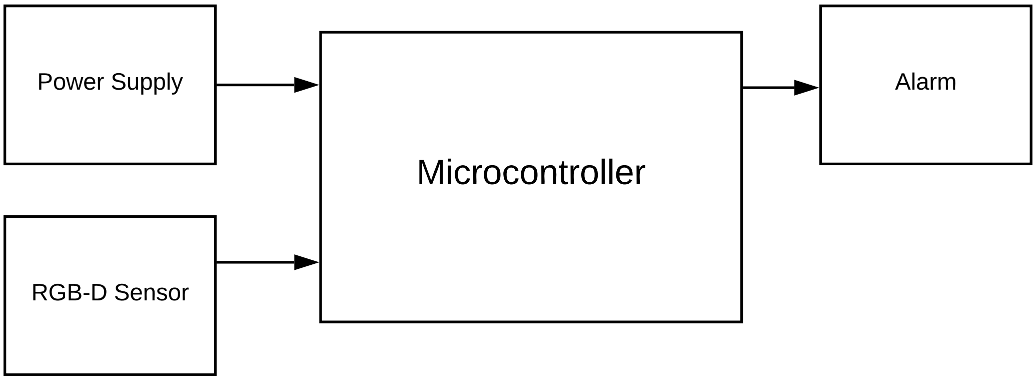

Fig. 1.3 Proposed Block Diagram of Driver Monitoring System

1.8 Overview

1.8.1 Gantt Chart

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

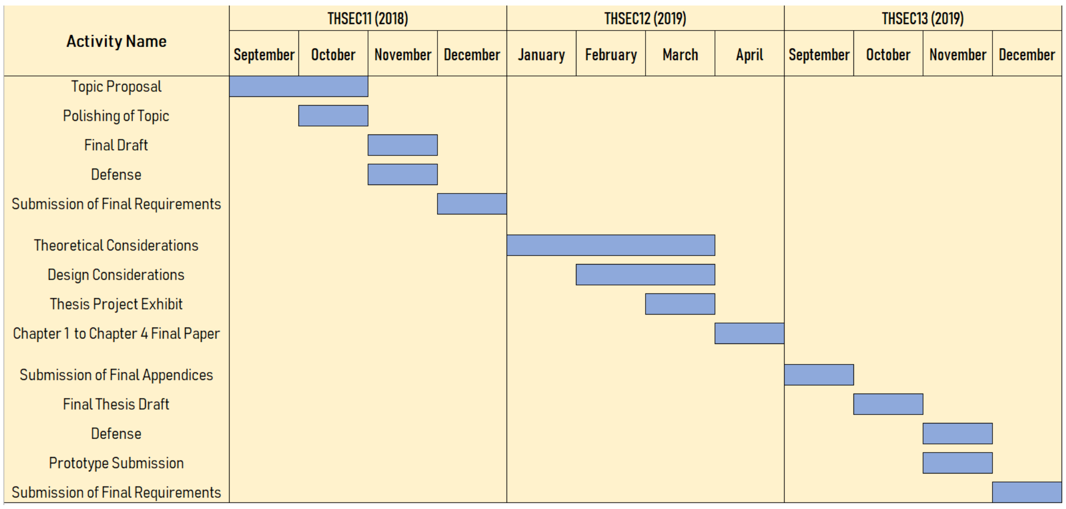

Fig. 1.4 Gantt Chart of Guevarra

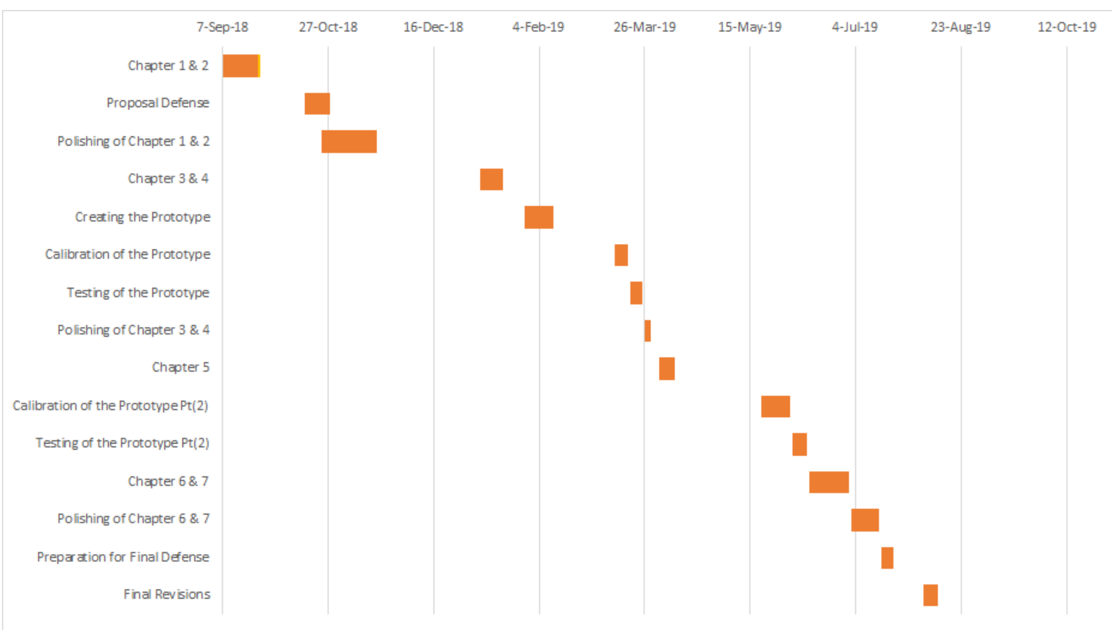

Fig. 1.5 Gantt Chart of Hernandez

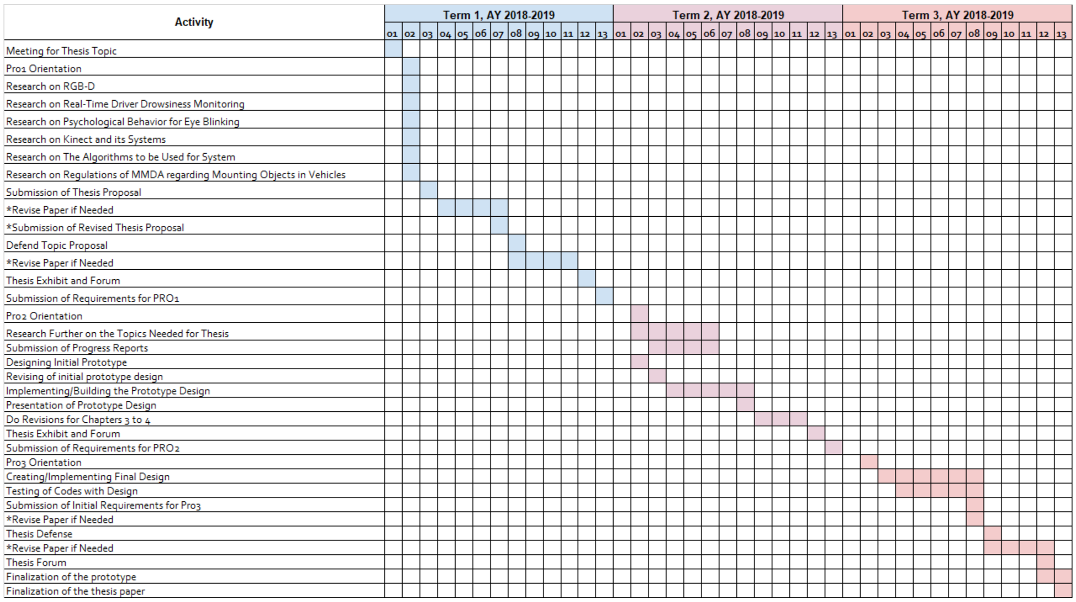

Fig. 1.6 Gantt Chart of Lagman

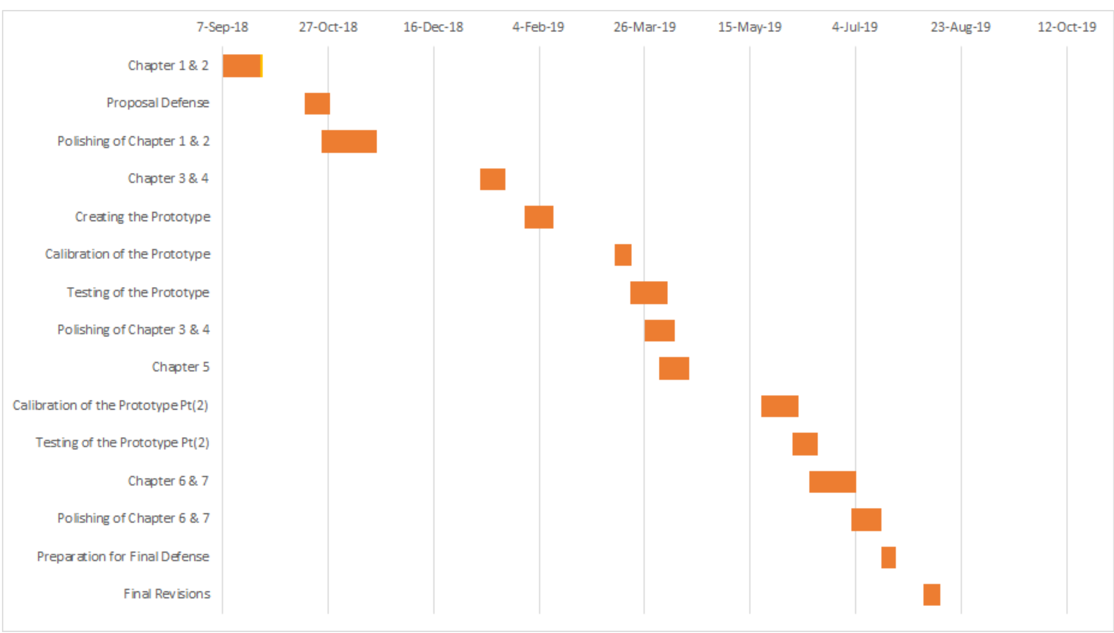

Fig. 1.7 Gantt Chart of Villanueva

1.8.2 Estimated Work Schedule and Budget

TABLE 1.3 PRICE LIST OF MATERIALS TO BE USED

Materials	Estimated Price(in PHP)
RGB-D Sensor	10,000-20,000
Microcontroller	450-2,000
Laptop	Already Available
Total Price	10,450-28,300

Chapter 2

LITERATURE REVIEW

Contents

2.1 Existing Work

2.1.1 Driver Awareness

Sleepiness can cause accidents in people's 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 the driver's 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 driver's 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 driver's 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 Webcam)

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 Driver's 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 user's eyes during the night. For eye gaze tracking, CAMSHIFT (Continuously Adaptive Mean Shift) algorithm was used to extract the ROI (Region of Interest). The subject's 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 driver's 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).

2.2 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 driver's 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 WLBP for the images, the effectiveness and robustness were emphasized because the

De La Salle University

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).

References