



รายงานการวิจัย

ระบบเฝ้าติดตามอาการง่วงนอนโดยใช้เทคนิคทางปัญญาประดิษฐ์ (Drowsiness Monitoring System Using Artificial Intelligent Technique)

ได้รับทุนอุดหนุนการวิจัยจาก
มหาวิทยาลัยเทคโนโลยีสุรนารี

ผลงานวิจัยเป็นความรับผิดชอบของหัวหน้าโครงการวิจัยแต่เพียงผู้เดียว



รายงานการวิจัย

ระบบเฝ้าติดตามอาการง่วงนอนโดยใช้เทคนิคทางปัญญาประดิษฐ์ (Drowsiness Monitoring System Using Artificial Intelligent Technique)

คณะผู้วิจัย

หัวหน้าโครงการ

ผู้ช่วยศาสตราจารย์ เรืออากาศเอก ดร. ประโยชน์ คำสวัสดิ์

สาขาวิชาวิศวกรรมโทรคมนาคม

สำนักวิชาวิศวกรรมศาสตร์

ได้รับทุนอุดหนุนการวิจัยจากมหาวิทยาลัยเทคโนโลยีสุรนารี ปีงบประมาณ พ.ศ. 2556

ผลงานวิจัยเป็นความรับผิดชอบของหัวหน้าโครงการวิจัยแต่เพียงผู้เดียว

กันยายน 2557

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณ มหาวิทยาลัยเทคโนโลยีสุรนารี ที่ได้ให้ทุนอุดหนุนการวิจัย ประจำปีงบประมาณ 2556

ผู้วิจัยขอขอบคุณ เจ้าหน้าที่ของสถานวิจัยสำนักวิชาวิศวกรรมศาสตร์และสถาบันวิจัยพัฒนา มหาวิทยาลัยเทคโนโลยีสุรนารี ที่ได้ให้ความสะดวกในด้านข้อมูล งานเอกสารแบบฟอร์มต่าง ๆ ที่ต้องใช้ในการดำเนินโครงการวิจัย จนทำให้งานเอกสารที่เกี่ยวข้องกับงานวิจัยนี้ดำเนินไปได้ด้วยดี

ผู้วิจัยขอขอบคุณ นายเกริกฤทธิ์ ศรีเคน นักศึกษาปริญญาโทที่ช่วยงานวิจัย ทำการทดลอง เก็บข้อมูล การวิจัย จนทำให้งานวิจัยนี้สำเร็จลุล่วง

ประโยชน์ คำสวัสดิ์

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อตรวจจับดวงตาสำหรับเฝ้าระวังอาการง่วงนอนของผู้ขับรถ โดยขั้นตอนแรกเป็นการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมและการคัดแยกสีผิว ต่อมาในขั้นตอนที่สองจะใช้เทคนิคและวิธีการ Integral projection และการตั้งค่าขีดเริ่มเปลี่ยนในการค้นหาตำแหน่งของดวงตา จากนั้นจะทำการประเมินสถานะของผู้ขับรถด้วยการคำนวณหาค่า PERCLOS ซึ่งเป็นการคำนวณหาอัตราส่วนของการปิดตาต่ำกว่า 20 เปอร์เซ็นต์ต่อการปิดตาระหว่าง 20-80 เปอร์เซ็นต์ของการเปิดตาสูงสุด ในขั้นสุดท้ายจะนำค่า PERCLOS ที่ได้มาตรวจจับอาการของผู้ขับรถ ผลการทดลองที่นำเสนอแสดงให้เห็นถึงประสิทธิผลของอัลกอริทึมและความเป็นไปได้ในการประยุกต์เพื่อการใช้งานได้จริง

คำหลัก : การเฝ้าระวังอาการง่วงนอน การตรวจจับใบหน้า การตรวจจับดวงตา

Abstract

The main objective of this research is to develop an eye detection technique for drowsiness monitoring system. Firstly, a face detector is used to locate face in the whole image with artificial neural networks and color skin segmentation. Secondly, the integral projection and threshold technique are applied to locate the exact position of the eyes. Then, the driver states are measured by calculating PERCLOS which is the ratio of the frame number of the eye between close and open 20%, and the frames of the eye between open 20% and open 80%. Finally, according to the PERCLOS parameter, it is easier to determine the drowsiness state of the driver. Experimental results show the feasibility and effectiveness of this algorithm.

Keywords: Drowsiness Detection, Face Detection, Eye Detection

สารบัญ

เรื่อง	หน้า
กิตติกรรมประกาศ	ก
บทคัดย่อ.....	ข
Abstract.....	ค
สารบัญ.....	ง
สารบัญตาราง	ฉ
สารบัญรูป.....	ช
บทที่ 1 บทนำ	1
1.1 ความสำคัญและที่มาของปัญหา	1
1.2 วัตถุประสงค์ของโครงการวิจัย	2
1.3 ขอบเขตของงานวิจัย.....	2
1.4 ขั้นตอนการดำเนินงาน.....	3
1.5 ประโยชน์ที่ได้รับ.....	3
บทที่ 2 ทฤษฎีและหลักการที่เกี่ยวข้อง.....	4
2.1 บทนำ	4
2.2 ทฤษฎีที่เกี่ยวข้องกับเทคนิคทางปัญญาประดิษฐ์.....	4
2.3 ทฤษฎีที่เกี่ยวข้องกับการประมวลผลภาพดิจิทัล.....	7
2.3.1 ปริภูมิสี (Color space) RGB	8
2.3.2 ปริภูมิสี HSV.....	8
2.3.3 การแปลงภาพจากปริภูมิสี RGB เป็นภาพระดับสีเทา (Gray scale)	9
2.3.4 ภาพขาว – ดำ (Binary image).....	10
2.3.5 การหาขอบภาพ (Edge detection).....	10
2.3.6 การขยายพิกเซล.....	12
2.3.7 การกร่อนพิกเซล.....	14

2.3.8	Retinex [2].....	15
2.4	การทบทวนวรรณกรรม/สารสนเทศ (Information) ที่เกี่ยวข้อง.....	19
บทที่ 3 การออกแบบระบบเฝ้าติดตามอาการง่วงนอน.....		24
3.1	บทนำ	24
3.2	การออกแบบระบบตรวจจับใบหน้า	24
3.2.1	การออกแบบการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม	24
3.2.2	การออกแบบการตรวจจับใบหน้าโดยใช้การคัดแยกสีผิว	30
3.2.3	การออกแบบการตรวจจับใบหน้าโดยใช้การคัดแยกสีผิวและเทคนิค MSRCR	33
3.3	การออกแบบระบบตรวจจับดวงตา	36
3.3.1	การค้นหาดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน	38
3.3.2	การค้นหาดวงตาด้วยวิธีการ Integral projection	39
3.4	การออกแบบระบบตรวจจับอาการง่วงนอน	44
3.5	สรุป	48
บทที่ 4 ผลการทดลอง		49
4.1	บทนำ	49
4.2	ผลการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม	49
4.3	ผลการตรวจจับใบหน้าโดยการคัดแยกสีผิว.....	54
4.4	ผลการตรวจจับใบหน้าโดยการคัดแยกสีผิวร่วมกับเทคนิค MSRCR.....	58
4.5	ผลการตรวจจับดวงตาและการตรวจจับอาการง่วงนอน	62
4.5.1	ผลการตรวจจับดวงตา	62
4.5.2	ผลการตรวจจับอาการง่วงนอน	67
บทที่ 5 บทสรุป		82
5.1	สรุป	82
5.2	แนวทางการวิจัยต่อไปในอนาคต.....	83
บรรณานุกรม.....		84
ภาคผนวก		88

สารบัญตาราง

หน้า

ตารางที่ 2.1: งานวิจัยที่เกี่ยวข้องด้านการตรวจจับอาการง่วงนอนขณะขับรถ	19
ตารางที่ 2.2: งานวิจัยที่เกี่ยวข้องด้านการตรวจจับใบหน้า.....	21
ตารางที่ 2.3: งานวิจัยที่เกี่ยวข้องด้านการตรวจจับดวงตา.....	21
ตารางที่ 2.4: งานวิจัยที่เกี่ยวข้องกับการตรวจจับอาการง่วงนอนโดยใช้เทคนิคทางปัญญาประดิษฐ์	23
ตารางที่ 3.1: พารามิเตอร์ของเทคนิค MSRCR ที่ใช้ในงานวิจัย	34
ตารางที่ 4.1: ผลการทดสอบการตรวจจับใบหน้าโดยใช้เทคนิคทางปัญญาประดิษฐ์	53
ตารางที่ 4.2: เปรียบเทียบร้อยละความถูกต้องระหว่างวิธีการของ P. P. Paul and M. Gavrilova กับวิธีการที่ นำเสนอ	55
ตารางที่ 4.3: เปรียบเทียบร้อยละความถูกต้องระหว่างวิธีตั้งการ Threshold และวิธีการ Integral projection.....	62
ตารางที่ 4.4: เปรียบเทียบการตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยนและ Integral projection....	64
ตารางที่ 4.5: ผลการตรวจจับอาการง่วงนอนในวิธีการที่ 1 ซึ่งใช้ระดับของ PERCLOS ที่กำหนดโดยงานวิจัย W. Qing et al. [30].....	70
ตารางที่ 4.6: ผลการตรวจจับอาการง่วงนอนในวิธีการที่ 2 ซึ่งใช้ระดับของ PERCLOS ที่กำหนดโดยงานวิจัย I. Garcia et al. [31]	75
ตารางที่ 4.7: ระดับค่า PERCLOS ที่มีการปรับระดับใหม่	77
ตารางที่ 4.8: ผลการตรวจจับอาการง่วงนอนกรณีใช้ระดับของ PERCLOS ที่มีการปรับระดับใหม่	78

สารบัญรูป

หน้า

รูปที่ 1.1 ภาพระบบเฝ้าระวังอาการง่วงนอนภายในรถยนต์.....	2
รูปที่ 2.1 โครงสร้างระบบประสาท.....	5
รูปที่ 2.2 โครงสร้างการทำงานของเครือข่ายประสาทเทียม.....	5
รูปที่ 2.3 การเรียนรู้แบบมีผู้ฝึกสอน.....	6
รูปที่ 2.4 การเรียนรู้แบบไม่มีผู้ฝึกสอน.....	7
รูปที่ 2.5 ลูกบาศก์แบบจำลองภาพสี RGB	8
รูปที่ 2.6 กรวยของแบบจำลองภาพสี HSV.....	9
รูปที่ 2.7 ค่าระดับสีเทา.....	9
รูปที่ 2.8 แสดงภาพในระดับสีเทาและภาพขาว-ดำ	10
รูปที่ 2.9 เทมเพลตสำหรับการหาขอบภาพในแนวนอน	11
รูปที่ 2.10 เทมเพลตสำหรับการหาขอบภาพในแนวตั้ง.....	11
รูปที่ 2.11 เทมเพลตการหาของภาพโดยวิธีโซเบล.....	12
รูปที่ 2.12 ภาพผลลัพธ์ที่ได้จากการหาขอบภาพ.....	12
รูปที่ 2.13 เทมเพลตการขยายพิกเซล	13
รูปที่ 2.14 ภาพจำลองการขยายพิกเซล	13
รูปที่ 2.15 ภาพจำลองการขยายพิกเซล	13
รูปที่ 2.16 ภาพจำลองการกร่อนพิกเซล.....	14
รูปที่ 2.17 ภาพจำลองการกร่อนพิกเซล.....	14
รูปที่ 2.18 การปรับปรุงโดยใช้เทคนิค Single-scale retinex.....	17
รูปที่ 2.19 การปรับปรุงโดยใช้เทคนิค Single-scale retinex.....	18
รูปที่ 2.20 การปรับปรุงโดยใช้เทคนิค MSR เทียบกับ MSRCR	19
รูปที่ 3.1 ระบบตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม	25
รูปที่ 3.2 ตัวอย่างภาพใบหน้าขนาด 200 x 200 พิกเซล.....	26
รูปที่ 3.3 ตัวอย่างภาพที่ไม่ใช่ภาพใบหน้าขนาด 200 x 200 พิกเซล	26
รูปที่ 3.4 การแปลงเวฟเล็ตแบบดีสครีท.....	27
รูปที่ 3.5 การแปลง DWT	27
รูปที่ 3.6 เมทริกซ์ A ที่ใช้ในการคำนวณสำหรับการวิเคราะห์องค์ประกอบ.....	28
รูปที่ 3.7 การทดลองเลือกค่าเวกเตอร์ไอเกนที่ค่าต่าง ๆ.....	29

รูปที่ 3.8 เลเยอร์ของเครือข่ายประสาทเทียม.....	30
รูปที่ 3.9 แผนภาพการตรวจจับใบหน้าโดยการคัดแยกสีผิว.....	30
รูปที่ 3.10 การแยกปริภูมิสี RGB.....	31
รูปที่ 3.11 การแปลงปริภูมิสี RGB เป็น HSV	32
รูปที่ 3.12 เอาต์พุตการคัดแยกสีผิว.....	32
รูปที่ 3.13 ภาพหลังการกำจัดบริเวณที่ไม่ใช่ใบหน้าออก.....	33
รูปที่ 3.14 ภาพหลังการตรวจจับ	33
รูปที่ 3.15 แผนภาพการตรวจจับใบหน้าด้วยวิธีการ	34
รูปที่ 3.16 เปรียบเทียบ Retinex แบบต่างๆ.....	35
รูปที่ 3.17 ตรวจจับใบหน้าหลังจากการทำ MSRCR.....	36
รูปที่ 3.18 แบ่งความสูงของใบหน้าเป็น 5 ส่วน.....	37
รูปที่ 3.19 บริเวณใบหน้า 2 ส่วน	37
รูปที่ 3.20 บริเวณใบหน้า 2 ส่วน.....	37
รูปที่ 3.21 หาขอบของวัตถุของบริเวณใบหน้า 2 ส่วน.....	37
รูปที่ 3.22 ขยายขอบของวัตถุของบริเวณใบหน้า 2 ส่วน.....	38
รูปที่ 3.23 ลักษณะของดวงตา	38
รูปที่ 3.24 ภาพระดับสีเทา.....	38
รูปที่ 3.25 ภาพหลังการตั้งค่าขีดเริ่มเปลี่ยน	39
รูปที่ 3.26 ลักษณะเฉพาะของตา	39
รูปที่ 3.27 การตรวจจับดวงตา.....	39
รูปที่ 3.28 บริเวณดวงตาที่เป็นภาพระดับสีเทา	40
รูปที่ 3.29 กราฟ Horizontal integral projection	41
รูปที่ 3.30 กราฟ Vertical integral projection.....	42
รูปที่ 3.31 ภาพม่านตาที่ได้จากวิธีการ Integral projection.....	44
รูปที่ 3.32 การติดตั้งกล้องวิดีโอในยานพาหนะ	44
รูปที่ 3.33 ขั้นตอนการตรวจจับอาการง่วงนอน.....	45
รูปที่ 3.34 เฟรมที่ใช้ในการคำนวณ PERCLOS [31]	46
รูปที่ 3.35 การกำหนดช่วงเวลาในการประมวลผล PERCLOS [25]	46
รูปที่ 3.36 กราฟ PERCLOS	47
รูปที่ 4.1 ภาพใบหน้าขนาด 200 x 200 พิกเซล	50
รูปที่ 4.2 ภาพที่ไม่ใช่ใบหน้าขนาด 200 x 200 พิกเซล.....	50
รูปที่ 4.3 การแปลง DWT	51
รูปที่ 4.4 เลเยอร์ของเครือข่ายประสาทเทียม.....	51

รูปที่ 4.5 กราฟการรู้จำของเครือข่ายประสาทเทียมในการทดลองที่ 1	52
รูปที่ 4.6 กราฟการรู้จำของเครือข่ายประสาทเทียมในการทดลองที่ 2	52
รูปที่ 4.7 กราฟการรู้จำของเครือข่ายประสาทเทียมในการทดลองที่ 3	53
รูปที่ 4.8 เปรียบเทียบความถูกต้อง.....	55
รูปที่ 4.9 แสดงเปอร์เซ็นต์การตรวจจับถูกต้องและผิดพลาด	56
รูปที่ 4.10 เปรียบเทียบร้อยละของการตรวจจับใบหน้าที่ถูกต้องและผิดพลาด	57
รูปที่ 4.11 ตัวอย่างการตรวจจับใบหน้าที่ถูกต้อง.....	57
รูปที่ 4.12 ตัวอย่างการตรวจจับใบหน้าที่ผิดพลาด.....	58
รูปที่ 4.13 การติดตั้งกล้องเว็บแคมภายในยานพาหนะ.....	58
รูปที่ 4.14 ภาพที่ใช้ในการทดสอบ	59
รูปที่ 4.15 แผนภาพการตรวจจับใบหนารวมกับเทคนิค MSRCR.....	60
รูปที่ 4.16 ตัวอย่างการปรับปรุงแสงสว่างด้วยเทคนิค MSRCR.....	61
รูปที่ 4.17 เปรียบเทียบการตรวจจับใบหน้า.....	61
รูปที่ 4.18 เปรียบเทียบการตรวจจับใบหน้าด้วยวิธีการคัดแยกสีผิว.....	62
รูปที่ 4.19 ภาพที่ใช้ในการทดสอบ	63
รูปที่ 4.20 เปรียบเทียบกราฟ Integral projection ของภาพดวงตา.....	66
รูปที่ 4.21 กราฟเปอร์เซ็นต์ของการปิดตาเทียบกับเฟรมเวลา	67
รูปที่ 4.22 กราฟ PERCLOS ของตาขวาในกรณีที่ผู้ทดสอบตื่นตัว.....	68
รูปที่ 4.23 กราฟ PERCLOS ของตาซ้ายในกรณีที่ผู้ทดสอบตื่นตัว.....	69
รูปที่ 4.24 กราฟ PERCLOS ของตาขวาในกรณีที่ผู้ทดสอบง่วงนอน	69
รูปที่ 4.25 กราฟ PERCLOS ของตาซ้ายในกรณีที่ผู้ทดสอบง่วงนอน.....	70
รูปที่ 4.26 การแจ้งเตือนผู้ทดสอบเมื่อมีอาการง่วงนอนในวิดีโอที่ผู้ทดสอบมีอาการง่วงนอน (PERCLOS > 0.4).....	71
รูปที่ 4.27 การแจ้งเตือนผู้ทดสอบเมื่อมีอาการง่วงนอนในวิดีโอที่ผู้ทดสอบมีอาการตื่นตัว (PERCLOS > 0.4).....	72
รูปที่ 4.28 กราฟ PERCLOS ของตาขวาในกรณีที่ผู้ทดสอบตื่นตัว.....	73
รูปที่ 4.29 กราฟ PERCLOS ของตาซ้ายในกรณีที่ผู้ทดสอบตื่นตัว.....	74
รูปที่ 4.30 กราฟ PERCLOS ของตาขวาในกรณีที่ผู้ทดสอบง่วงนอน	74
รูปที่ 4.31 กราฟ PERCLOS ของตาซ้ายในกรณีที่ผู้ทดสอบง่วงนอน.....	75
รูปที่ 4.32 การแจ้งเตือนผู้ทดสอบเมื่อมีอาการง่วงนอนในวิดีโอที่ผู้ทดสอบมีอาการง่วงนอน (PERCLOS > 0.23)	76
รูปที่ 4.33 การแจ้งเตือนผู้ทดสอบเมื่อมีอาการง่วงนอนในวิดีโอที่ผู้ทดสอบมีอาการตื่นตัว (PERCLOS > 0.23)	77

รูปที่ 4.34 การแจ้งเตือนผู้ทดสอบเมื่อมีอาการง่วงนอนในวิชิตส์น ที่ผู้ทดสอบมีอาการง่วงนอน (PERCLOS > 0.4).....	79
รูปที่ 4.35 การแจ้งเตือนผู้ทดสอบเมื่อมีอาการง่วงนอนในวิชิตส์น ที่ผู้ทดสอบมีอาการตื่นตัว (PERCLOS > 0.4).....	80

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของปัญหา

ในปัจจุบันนี้ การเกิดอุบัติเหตุเป็นสาเหตุการตายอันดับสองของคนไทยรองมาจากโรคมะเร็ง โดยอุบัติเหตุทางรถยนต์เป็นอุบัติเหตุที่เกิดขึ้นบ่อยเป็นอันดับต้น ๆ ซึ่งทำให้เกิดความสูญเสียทั้งชีวิตและทรัพย์สิน โดยสาเหตุหลักของการเกิดอุบัติเหตุทางรถยนต์คือ ความประมาทของผู้ใช้รถใช้ถนน ความไม่พร้อมของสภาพยานพาหนะที่ใช้และความไม่พร้อมของผู้ขับขี่ เช่น พักผ่อนไม่เพียงพอทำให้เกิดอาการง่วงนอน (Drowsiness) และเกิดการหลับในขณะขับรถ หรืออาจอยู่ในอาการมึนเมาจากการดื่มสุราหรือรับประทานยาที่อาจทำให้เกิดการง่วงซึม เป็นต้น นอกจากนี้การขับรถเป็นระยะเวลานานโดยไม่หยุดพักก็อาจทำให้เกิดความเมื่อยล้า (Fatigue) และการง่วงนอนซึ่งทำให้เกิดการหลับในขณะขับรถได้เช่นเดียวกัน

แนวทางหนึ่งที่จะสามารถลดการเกิดอุบัติเหตุในขณะขับรถอันเนื่องมาจากการหลับในคือ การแจ้งเตือนแก่ผู้ที่กำลังปฏิบัติงานในขณะที่กำลังเกิดอาการง่วง ด้วยการแจ้งเตือนหรือใช้ไฟสัญญาณเตือนเพื่อให้ผู้ปฏิบัติงานได้รับทราบ ทั้งนี้เพื่อที่จะได้แก้ไขอาการง่วงนั้นอย่างทันทั่วทั้งที่ วิธีการตรวจจับอาการง่วงนอนขณะขับรถนั้นสามารถทำได้โดยอาศัยการสังเกตอาการของผู้ปฏิบัติงานเช่น ลักษณะของเปลือกตาที่หริ่งมากกว่าปกติ และมีจำนวนครั้งของการกระพริบตานั้นน้อยลง มีการเคลื่อนไหวของศีรษะเอียงไปมา หรือการเกิดอาการนิ่งของเปลือกตาหรือศีรษะที่นานกว่าปกติ เป็นต้น [1]

วิธีการตรวจจับอาการง่วงนอนโดยอาศัยการสังเกตอาการของผู้ปฏิบัติงานนั้น สามารถทำได้โดยใช้หลักการของประมวลผลภาพดิจิทัล (Digital image processing) โดยหลักการดังกล่าวจะทำการตรวจจับดวงตาเพื่อค้นหาภาพดวงตาในข้อมูลภาพที่เรียงลำดับกัน (Image sequences) หรือที่เรียกว่าภาพวิดีโอ (Video) จากนั้นจะทำการเฝ้าสังเกตลักษณะของดวงตาและการกระพริบตาในข้อมูลภาพ เมื่อตรวจพบอาการง่วงนอนดังกล่าวก็จะทำการปลุกให้ตื่น หรือทำการแจ้งเตือนให้เปลี่ยนอิริยาบถหรือหยุดพักการทำงานชั่วคราว

เนื่องจากในปัจจุบัน เทคโนโลยีวงจรรวม (Integrated circuit technology) มีความเจริญรุดหน้าไปมาก ทำให้การประมวลผลของคอมพิวเตอร์ทำได้อย่างรวดเร็วและมีสมรรถนะสูงมากขึ้น ประกอบกับเทคนิคและวิธีการทางปัญญาประดิษฐ์ (Artificial intelligent techniques) ถูกนำมาประยุกต์ใช้งานอย่างกว้างขวางในด้านต่าง ๆ โดยเฉพาะในด้านการมองเห็นของคอมพิวเตอร์ (Computer vision) เนื่องจากเทคนิคและวิธีการดังกล่าวให้ผลลัพธ์ที่ดีในการค้นหาคำตอบที่เหมาะสมที่สุดในวงกว้าง (Global optimization) และสามารถค้นหาคำตอบ ทำการเรียนรู้และรู้จำได้อย่างถูกต้องและรวดเร็ว ซึ่งความสามารถเหล่านี้จะเพิ่มขึ้นตามสมรรถนะของคอมพิวเตอร์ที่เพิ่มขึ้น

ดังนั้นในงานวิจัยนี้ จึงมุ่งเน้นไปที่การนำเทคนิคและวิธีการทางปัญญาประดิษฐ์ เครือข่ายประสาทเทียม (Neural networks: NN) มาทำการประยุกต์ใช้เพื่อทำการค้นหาคำตอบที่เหมาะสมที่สุดของ

พารามิเตอร์ต่าง ๆ ที่ใช้ในการตรวจจับใบหน้าและดวงตา เพื่อให้ได้ผลลัพธ์ของการตรวจจับใบหน้าและดวงตา และการหาอัตราการกระพริบของเปลือกตาสำหรับระบบเฝ้าติดตามอาการง่วงนอนขณะขับรถที่ถูกต้องและแม่นยำ แผนภาพการประยุกต์ใช้ระบบเฝ้าติดตามอาการง่วงนอนโดยใช้เทคนิคการประมวลผลภาพดิจิทัลดังแสดงในรูปที่ 1.1



รูปที่ 1.1 ภาพระบบเฝ้าระวังอาการง่วงนอนภายในรถยนต์

1.2

วัตถุประสงค์ของโครงการวิจัย

1. เพื่อศึกษาวิธีการตรวจจับใบหน้าและดวงตาด้วยหลักการประมวลผลภาพดิจิทัล
2. เพื่อดำเนินการออกแบบตัวตรวจจับใบหน้าและดวงตาสำหรับระบบเฝ้าติดตามอาการง่วงนอนขณะขับรถ
3. เพื่อดำเนินการออกแบบโปรแกรมและสร้างชุดต้นแบบระบบเฝ้าติดตามอาการง่วงนอนขณะขับรถสำหรับแจ้งเตือนผู้ขับรถ เพื่อลดอัตราการเกิดอุบัติเหตุลง
4. เพื่อสร้างองค์ความรู้ใหม่ในการออกแบบระบบเฝ้าติดตามอาการง่วงนอนขณะขับรถด้วยวิธีการทางปัญญาประดิษฐ์

1.3

ขอบเขตของงานวิจัย

1. สัญญาณภาพวิดีโอที่ที่ใช้ในการทดสอบต้องมีความละเอียดอย่างต่ำในระดับ VGA 640×480 และมีความเร็วอย่างต่ำ 25 เฟรมต่อวินาที

2. ทำการออกแบบและพัฒนาระบบเฝ้าติดตามอาการร่งงนอนขณะขับรถโดยใช้โปรแกรม MATLAB C/C++ และโปรแกรม LabVIEW

1.4 ขั้นตอนการดำเนินงาน

วิธีการดำเนินการวิจัย มีขั้นตอนดังต่อไปนี้

1. ศึกษาวิธีการตรวจจับใบหน้าและดวงตาด้วยหลักการของการประมวลผลภาพดิจิทัล
2. ศึกษาวิธีการตรวจจับใบหน้าและดวงตาที่ประยุกต์ใช้เทคนิคทางปัญญาประดิษฐ์
3. ดำเนินการเก็บข้อมูลภาพวิดีโอของใบหน้าและดวงตาในขณะขับรถยนต์
4. ดำเนินการออกแบบตัวตรวจจับและคัดแยกภาพดวงตาจากภาพวิดีโอเพื่อนำไปใช้กับระบบตรวจจับอาการร่งงนอนขณะขับรถ โดยประยุกต์ใช้เทคนิคทางปัญญาประดิษฐ์
5. ดำเนินการออกแบบโปรแกรมและชุดต้นแบบระบบเฝ้าติดตามอาการร่งงนอนขณะขับรถ
6. ดำเนินการทดสอบและพัฒนาระบบเฝ้าติดตามอาการร่งงนอนขณะขับรถ

1.5 ประโยชน์ที่ได้รับ

1. เป็นองค์ความรู้ในการวิจัยในด้านการประมวลผลสัญญาณภาพดิจิทัลในมหาวิทยาลัยและหน่วยงานอื่น ๆ
2. นักวิจัยในโรงงานอุตสาหกรรมสามารถนำไปใช้ออกแบบและพัฒนาระบบเฝ้าติดตามอาการร่งงนอนเพื่อเพิ่มมูลค่าผลิตภัณฑ์
3. สามารถนำไปใช้ออกแบบและพัฒนาระบบเฝ้าติดตามอาการร่งงนอนสำหรับธุรกิจที่เกี่ยวข้องกับการขนส่งสินค้าหรือผู้โดยสาร

บทที่ 2

ทฤษฎีและหลักการที่เกี่ยวข้อง

2.1 บทนำ

ในบทนี้จะกล่าวถึงทฤษฎีที่ใช้ในงานวิจัย ได้แบ่งออกเป็น 2 ส่วนใหญ่ ๆ ได้แก่ ทฤษฎีที่เกี่ยวข้องกับเทคนิคทางปัญญาประดิษฐ์ และทฤษฎีที่เกี่ยวข้องกับการประมวลผลภาพดิจิทัล การปรับปรุงแสงสว่างด้วยเทคนิค Retinex และในตอนท้ายได้กล่าวถึงวรรณกรรมหรือบทความที่เกี่ยวข้องกับงานวิจัยที่มีมาแล้วในอดีต

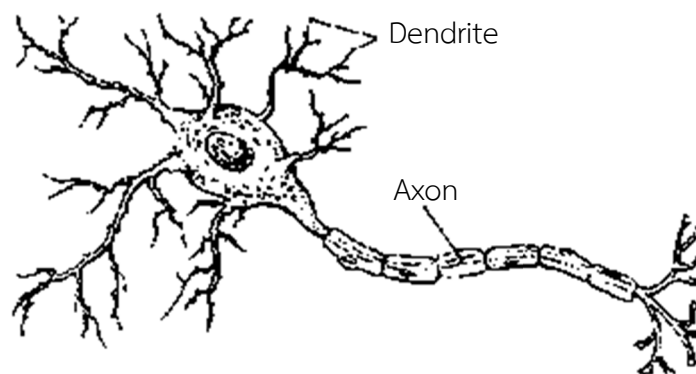
2.2 ทฤษฎีที่เกี่ยวข้องกับเทคนิคทางปัญญาประดิษฐ์

เครือข่ายประสาทเทียม [3] เป็นศาสตร์แขนงหนึ่งทางด้านปัญญาประดิษฐ์ที่สามารถนำไปประยุกต์ใช้กับงานหลายด้านได้อย่างมีประสิทธิภาพ เช่น การจำแนกรูปแบบ การทำนาย การควบคุม การหาความเหมาะสม และการจัดกลุ่ม เป็นต้น

หลักการสำคัญของเครือข่ายประสาทเทียม คือ ความพยายามที่จะลอกเลียนแบบการทำงานของเซลล์ประสาทในสมองมนุษย์เพื่อทำงานได้อย่างมีประสิทธิภาพ ลักษณะทั่วไปของเครือข่ายประสาทเทียมคือการที่โหนด (Node) ต่าง ๆ จำลองมาจากจุดประสานประสาท (Synapse) ของเซลล์ประสาทระหว่างปลายในการรับกระแสประสาท เรียกว่า เด็นไดรต์ (Dendrite) ซึ่งเป็นอินพุตและปลายในการส่งกระแสประสาท เรียกว่า แกนประสาท (Axon) ซึ่งเป็นเหมือนเอาต์พุตของเซลล์โดยมีฟังก์ชันเป็นตัวกำหนดสัญญาณส่งออก (Activation function or Transfer function) นั่นเอง

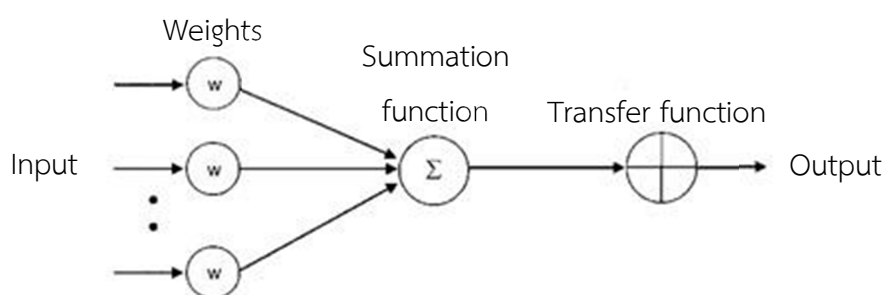
ลักษณะของเครือข่ายประสาทเทียมสามารถแบ่งได้ 2 แบบ คือ 1) เครือข่ายประสาทเทียมแบบชั้นเดียว (Single layer) ซึ่งจะมีเพียงชั้นสัญญาณประสาทขาเข้า และชั้นสัญญาณประสาทขาออกเท่านั้น เช่น เครือข่ายเพอร์เซ็ปตรอนอย่างง่าย (Simple perceptron) เป็นต้น และ 2) เครือข่ายประสาทเทียมแบบหลายชั้น (Multi-layer) ซึ่งมีลักษณะเช่นเดียวกับเครือข่ายประสาทเทียมแบบชั้นเดียว แต่จะมีชั้นแอบแฝง (Hidden) เพิ่มขึ้นโดยอยู่ส่วนกลางระหว่างชั้นนำข้อมูลป้อนเข้าและชั้นส่งข้อมูลออก ทั้งนี้ชั้นแอบแฝงอาจมี 1 ชั้น อย่งไรก็ตาม การแบ่งเครือข่ายประสาทเทียมตามประเภทการเรียนรู้ของเครือข่าย สามารถแบ่งได้ 2 ประเภท คือ การเรียนรู้แบบมีผู้สอน (Supervised learning) และการเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) โดยในปัจจุบันการพัฒนาเครือข่ายประสาทเทียมยังคงมีการดำเนินการอย่างต่อเนื่องและคาดว่าจะมีบทบาทอย่างมากในด้านการจำแนกรูปแบบ การพยากรณ์ การควบคุม การหาความเหมาะสมและการจัดกลุ่ม

ลักษณะโดยทั่วไปของโครงสร้างระบบประสาทแสดงดังแสดงในรูปที่ 2.1



รูปที่ 2.1 โครงสร้างระบบประสาท

เครือข่ายประสาทเทียมมีคุณลักษณะคล้ายกับการส่งผ่านสัญญาณประสาทในสมองของมนุษย์ กล่าวคือมีความสามารถในการรวบรวมความรู้ (Knowledge) โดยผ่านกระบวนการเรียนรู้ (Learning process) และความรู้เหล่านี้จะจัดเก็บอยู่ในรูปแบบค่าน้ำหนัก (Weight) ซึ่งสามารถปรับเปลี่ยนค่าได้เมื่อมีการเรียนรู้สิ่งใหม่ ๆ เข้าไปค่าน้ำหนักทำหน้าที่เปรียบเสมือนความรู้ที่รวบรวมไว้เพื่อใช้ในการแก้ไขปัญหาเฉพาะอย่างของมนุษย์การประมวลผลต่าง ๆ เกิดขึ้นในหน่วยประมวลผลย่อยเรียกว่าโหนด (node) ซึ่งโหนดเป็นการจำลองลักษณะการทำงานมาจากเซลล์การส่งสัญญาณ (Signal) ระหว่างโหนดที่เชื่อมต่อกัน (Connection) จำลองมาจากการเชื่อมต่อของเดินไดรต์และแกนประสาทในระบบประสาทของมนุษย์ภายในโหนดจะมีฟังก์ชันกำหนดสัญญาณส่งออกที่เรียกว่า ฟังก์ชันกระตุ้น (Activation function) หรือฟังก์ชันการแปลง (Transfer function) ซึ่งทำหน้าที่เปรียบเสมือนกระบวนการทำงานในเซลล์ ดังแสดงในรูปที่ 2.2



รูปที่ 2.2 โครงสร้างการทำงานของเครือข่ายประสาทเทียม

เครือข่ายประสาทเทียมประกอบด้วย 5 องค์ประกอบ ดังนี้

1. ข้อมูลป้อนเข้า (Input) คือข้อมูลที่เป็นตัวเลขหากเป็นข้อมูลเชิงคุณภาพต้องแปลงให้อยู่ในรูปเชิงปริมาณที่เครือข่ายประสาทเทียมยอมรับได้
2. ข้อมูลส่งออก (Output) คือผลลัพธ์ที่เกิดขึ้นจริง (Actual output) จากกระบวนการเรียนรู้ของเครือข่ายประสาทเทียม

3. ค่าน้ำหนัก (Weights) คือสิ่งที่ได้จากการเรียนรู้ของเครือข่ายประสาทเทียมหรือเรียกอีกอย่างหนึ่งว่า ค่าความรู้ (Knowledge) ค่านี้จะถูกเก็บเป็นทักษะเพื่อใช้ในการจดจำข้อมูล อื่น ๆ ที่อยู่ในรูปแบบเดียวกัน
4. ฟังก์ชันผลรวม (Summation function : S) เป็นผลรวมของข้อมูลป้อนเข้า (a_i) และค่าน้ำหนัก (w_i)
- 5.

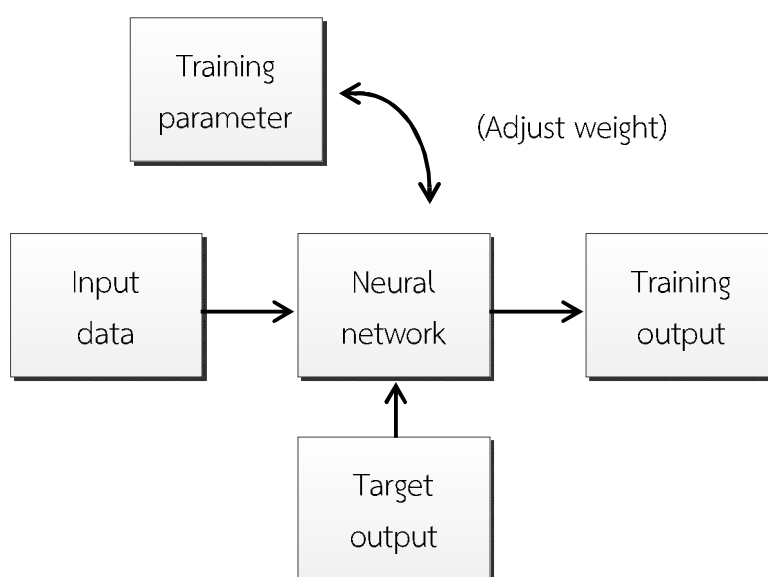
$$S = \sum_{i=1}^n a_i w_i \quad (2.1)$$

6. ฟังก์ชันการแปลง (Transfer function) เป็นการคำนวณการจำลองการทำงานของเครือข่ายประสาทเทียม เช่น ซิกมอยด์ฟังก์ชัน (Sigmoid function) ฟังก์ชันไฮเพอร์โบลิคแทนเจนต์ (Hyperbolic tangent function) เป็นต้น

ประเภทของการเรียนรู้ของเครือข่ายประสาทเทียม

1. การเรียนรู้แบบมีผู้สอน (Supervised learning)

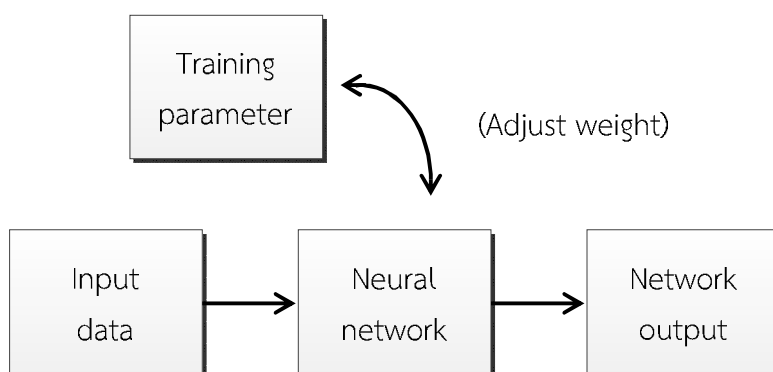
ข้อมูลจะประกอบด้วยตัวอย่างข้อมูลที่ต้องการสอน และผลลัพธ์ที่ต้องการให้เครือข่ายสร้าง เมื่อมีการนำข้อมูลในลักษณะเดียวกันมาเป็นข้อมูลป้อนเข้าเครือข่ายจะกำหนดผลลัพธ์ที่เป็นเป้าหมายให้กับข้อมูลป้อนเข้าแต่ละตัวเครือข่ายจะนำค่าผิดพลาดระหว่างค่าเป้าหมายกับค่าผลลัพธ์ที่ได้มาใช้ในการปรับค่าน้ำหนัก เพื่อให้ค่าผลลัพธ์ที่ใกล้เคียงกับเป้าหมายมากที่สุดดังแสดงในรูปที่ 2.3 ถ้าหากเปรียบเทียบกับมนุษย์จะเหมือนกับการสอนนักเรียนโดยมีครูผู้สอนคอยให้คำแนะนำ



รูปที่ 2.3 การเรียนรู้แบบมีผู้ฝึกสอน

2. การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning)

การเรียนรู้แบบนี้จะสอนเครือข่ายโดยการนำข้อมูลป้อนเข้าอย่างต่อเนื่องอย่างเดียวไม่มีการส่งค่าผลลัพธ์เป้าหมายให้กับข้อมูลป้อนเข้าแต่ละตัว การปรับน้ำหนักจะใช้ข้อมูลที่นำมาสอนเป็นตัวปรับค่า โดยค่าน้ำหนักจะปรับตามกลุ่มที่ข้อมูลป้อนเข้าที่มีรูปแบบคล้ายคลึงกันดังแสดงในรูปที่ 2.4 ถ้าหากเปรียบเทียบกับมนุษย์จะเหมือนกับการที่เราสามารถแยกแยะ พันธุ์พืช พันธุ์สัตว์ ตามลักษณะรูปร่างของมันได้ด้วยตัวเอง



รูปที่ 2.4 การเรียนรู้แบบไม่มีผู้ฝึกสอน

2.3 ทฤษฎีที่เกี่ยวข้องกับการประมวลผลภาพดิจิทัล

การตรวจจับใบหน้าและการเคลื่อนไหวของดวงตานั้นจะนำหลักการประมวลผลภาพดิจิทัลมาใช้ เช่น การแปลงค่าระหว่างปริภูมิสี การหาขอบวัตถุ การหาพื้นที่ของวัตถุ การหาจุดกึ่งกลางของวัตถุ เป็นต้น

วิดีโอ (Video) คือการเรียงต่อกันของภาพ แบ่งออกเป็น 2 ประเภทใหญ่ ๆ ได้แก่ วิดีทัศน์แบบอะนาล็อก (Analog video) และวิดีโอแบบดิจิทัล (Digital video) ซึ่งวิดีโอแบบอะนาล็อกนี้จะเก็บข้อมูลภาพและเสียงในรูปแบบของสัญญาณไฟฟ้า วิดีทัศน์แบบอะนาล็อกจะมีความไวต่อการผิดเพี้ยนสูง เมื่อมีการบันทึกต่อกันหลายๆ ครั้ง จะทำให้คุณภาพของภาพและเสียงลดลง สำหรับวิดีโอแบบดิจิทัลนั้นจะเก็บข้อมูลอยู่ในรูปของไฟล์คอมพิวเตอร์ที่สามารถนำกลับมาใช้ใหม่ได้ ทำสำเนาได้ และสามารถปรับแต่งแก้ไขได้

อัตราเฟรม (Frame rate) คืออัตราความเร็วในการแสดงภาพ อัตราในวิดีโอมีหน่วยเป็นเฟรมต่อวินาที (Frame per second: fps) เป็นหน่วยวัดปริมาณข้อมูลที่ใช้ในการเก็บบันทึกและแสดงวิดีโอ เช่น อัตราเฟรมของภาพยนตร์เท่ากับ 24 fps อัตราเฟรมระบบ PAL เท่ากับ 25 fps และอัตราเฟรมของ NTSC เท่ากับ 30 fps

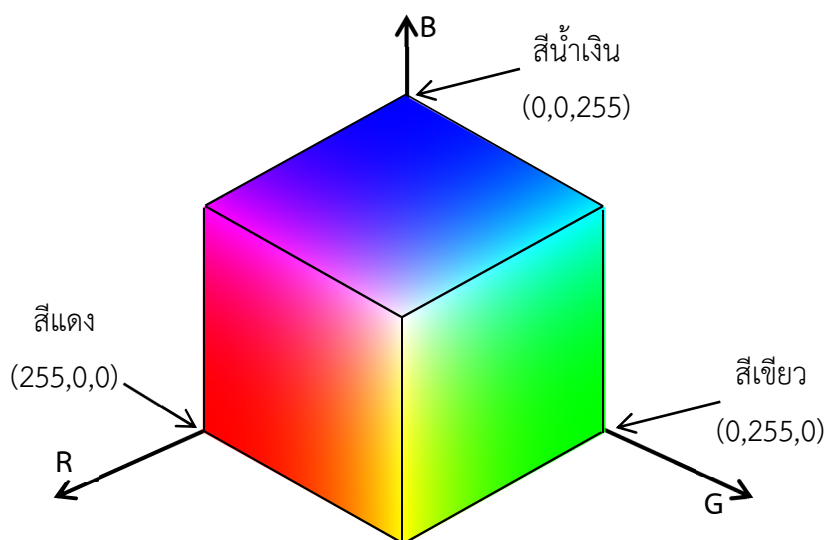
ความละเอียด (Resolution) คือความชัดของภาพ ความละเอียดของภาพขึ้นอยู่กับจำนวนพิกเซลทั้งหมด เช่น ความละเอียด 640 x 480 พิกเซล หมายถึง มีจำนวนพิกเซลแสดงผลเรียงกันในแนวนอน 640 พิกเซลและแนวตั้ง 480 พิกเซล

2.3.1 ปริภูมิสี (Color space) RGB

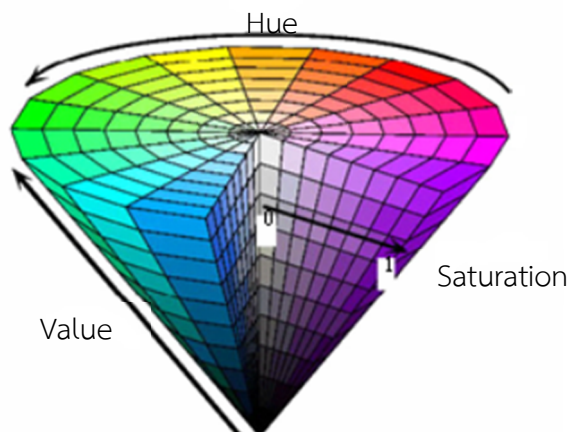
ปริภูมิสี RGB ประกอบไปด้วย 3 องค์ประกอบได้แก่ ค่าสีแดง (Red: R) ค่าสีเขียว (Green: G) และค่าสีน้ำเงิน (Blue: B) แต่ละสีจะเป็นอิสระต่อกันและมีค่าอยู่ระหว่าง 0-255 เมื่อนำสีทั้งสามมารวมกันที่ความเข้มสูงสุดจะได้สีขาว ส่วนใหญ่ปริภูมิสีนี้จะใช้ในอุปกรณ์ที่เกี่ยวข้องกับแสง เช่น จอภาพ กล้องดิจิทัล สแกนเนอร์ ดังแสดงในรูปที่ 2.5

2.3.2 ปริภูมิสี HSV

ปริภูมิสี HSV เป็นปริภูมิสีที่ประกอบด้วย 3 องค์ประกอบได้แก่ค่า H (Hue) คือค่าโทนสีหรือค่าสีส้มซึ่งมีค่าสีที่แตกต่างกันไปตามความถี่ของแสง ค่า S (Saturation) คือค่าความอิ่มตัวของสี ซึ่งเป็นค่าที่แสดงระดับสีเมื่อเทียบกับค่าโทนสี และค่า V (Value) คือค่าบอกระดับความสว่างของภาพ ซึ่งที่ระดับความสว่างต่ำสุดหมายถึงสีดำ ไม่ว่าจะมีย่านสี หรือค่าความอิ่มตัวของสีเท่าใด และระดับความสว่างสูงสุดหมายถึงสีขาว ซึ่งเป็นสีที่สว่างที่สุดของค่าโทนสี และค่าความอิ่มตัวของสีโดยภาพปริภูมิสี HSV สามารถหาได้จากการแปลงภาพปริภูมิสี RGB ด้วยสมการที่ (2.2) ถึงสมการที่ (2.5) และแสดงแบบจำลองภาพสี HSV ในรูปที่ 2.6



รูปที่ 2.5 ลูกบาศก์แบบจำลองภาพสี RGB



รูปที่ 2.6 กรวยของแบบจำลองภาพสี HSV

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360^\circ - \theta & \text{if } B > G \end{cases} \quad (2.2)$$

$$\theta = \cos^{-1} \left\{ \frac{0.5[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)^2}} \right\} \quad (2.3)$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (2.4)$$

$$V = \frac{\max(R, G, B)}{255} \quad (2.5)$$

2.3.3 การแปลงภาพจากปริภูมิสี RGB เป็นภาพระดับสีเทา (Gray scale)

การแปลงภาพปริภูมิสี RGB เป็นภาพระดับสีเทา เป็นการปรับให้ภาพแสดงถึงค่าความสว่าง (Brightness) ของภาพเพียงอย่างเดียวเท่านั้น ซึ่งค่าความสว่างของภาพ โดยทั่วไปในภาพขนาด 8 บิต ระดับสีเทาที่ประกอบด้วยค่าความสว่างที่แตกต่างกัน 256 ระดับ นั่นคือ จะมีค่าตั้งแต่ 0 ถึง 255 ดังแสดงในรูปที่ 2.7 โดยจะสังเกตเห็นได้ ระดับ 0 จะเป็นสีดำและระดับ 255 จะเป็นสีขาว



รูปที่ 2.7 ค่าระดับสีเทา

ในการแปลงภาพปริภูมิสี RGB ให้เป็นภาพระดับสีเทา สามารถคำนวณจากการแปลงค่าสีในภาพจากความสัมพันธ์ของการแปลงภาพสีเป็นภาพระดับสีเทาดังสมการที่ (2.6)

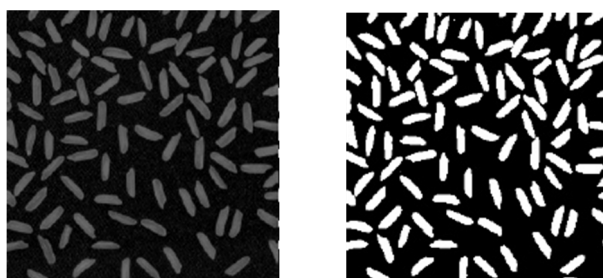
$$Grayscale_{(x,y)} = 0.299R_{(x,y)} + 0.587G_{(x,y)} + 0.114B_{(x,y)} \quad (2.6)$$

เมื่อ	<i>Grayscale</i>	คือ ค่าระดับสีเทาที่ได้จากการคำนวณของจุดภาพสี $RGB_{(x,y)}$
	<i>R</i>	คือ ค่าสีแดงในภาพระบบสี RGB ของจุดภาพของสี $R_{(x,y)}$
	<i>G</i>	คือ ค่าสีเขียวในภาพระบบสี RGB ของจุดภาพของสี $G_{(x,y)}$
	<i>B</i>	คือ ค่าสีน้ำเงินในภาพระบบสี RGB ของจุดภาพของสี $B_{(x,y)}$

2.3.4 ภาพขาว – ดำ (Binary image)

ภาพขาว – ดำ เป็นภาพที่มีเพียงสองระดับสีคือ สีขาวและสีดำ ซึ่งสีขาวจะแทนด้วยบิต 1 และสีดำจะแทนด้วยบิต 0 ภาพขาว – ดำได้จากการแปลงภาพจากภาพระดับสีเทาโดยใช้ค่า Threshold ถ้าพิกเซลในภาพระดับสีเทามีค่ามากกว่าค่า Threshold จะถูกแทนด้วยบิต 1 และในทางตรงกันข้ามถ้าพิกเซลในภาพระดับสีเทามีค่าน้อยกว่าค่า Threshold จะถูกแทนด้วยบิต 0 ดังสมการที่ (2.7) และรูปที่ 2.8

$$Binary\ image = \begin{cases} 1 & ; if\ grayscale \geq threshold \\ 0 & ; else \end{cases} \quad (2.7)$$



ก)

ข)

รูปที่ 2.8 แสดงภาพในระดับสีเทาและภาพขาว-ดำ

ก) ภาพในระดับสีเทา ข) ภาพขาว – ดำ

2.3.5 การหาขอบภาพ (Edge detection)

การหาขอบของภาพเป็นการหาเส้นรอบวัตถุที่อยู่ในภาพ เมื่อทราบเส้นรอบวัตถุ เราจะสามารถคำนวณหาพื้นที่ (ขนาด) หรือรู้จำนวนของวัตถุนั้นได้ อย่างไรก็ตาม การหาขอบภาพที่ถูกต้องสมบูรณ์ไม่ใช่เป็น

เรื่องที่ยาก โดยเฉพาะอย่างยิ่งการหาขอบของภาพที่มีคุณภาพต่ำ ขอบของภาพจะเห็นได้ชัด ถ้าความแตกต่างของความเข้มแสงจากจุดหนึ่งไปยังอีกจุดหนึ่งมีค่ามากขอบภาพก็จะเห็นได้ชัด ถ้าความแตกต่างมีค่าน้อยขอบภาพก็จะไม่ชัดเจน

เทมเพลต (Template) สำหรับหาขอบภาพ การหาขอบภาพในแนวนอนอย่างง่าย วิธีการก็คือหาผลต่างระหว่างจุดหนึ่งกับจุดที่อยู่ข้างล่าง (หรือข้างบน) ของจุดนั้น ดังนี้

$$Y_{diff}(x, y) = I(x, y) - I(x, y + 1) \quad (2.8)$$

โดยที่ Y_{diff} คือค่าความแตกต่างในแนวแกนตั้ง และ $I(x, y)$ คือค่าความเข้มแสงของจุดภาพที่ตำแหน่ง (x, y)

ผลของการใช้สมการ (2.8) มีค่าเทียบเท่ากับการคอลโวลูชันภาพด้วยเทมเพลต

1
-1

รูปที่ 2.9 เทมเพลตสำหรับการหาขอบภาพในแนวนอน

การหาขอบภาพในแนวตั้งก็สามารถหาได้เช่นเดียวกันคือ

$$X_{diff}(x, y) = I(x, y) - I(x - 1, y) \quad (2.9)$$

โดยที่ X_{diff} คือค่าความแตกต่างในแนวนอน และสมการที่ (2.9) มีให้ผลเทียบเท่ากับการคอนโวลูชันด้วยเทมเพลต

-1	1
----	---

รูปที่ 2.10 เทมเพลตสำหรับการหาขอบภาพในแนวตั้ง

บางครั้งเราต้องการรวมผลต่างของค่าความแตกต่างในแนวแกนแนวนอน และแกนตั้งเข้าด้วยกัน เพื่อที่จะได้มีตัววัดความแรงของขอบภาพ (Gradient magnitude) เพียงตัวเดียว เนื่องจากค่าความแตกต่างอาจมีค่าเป็นบวกหรือลบ ดังนั้น การบวกค่าความแตกต่างของทั้งสองแกนอาจทำให้ขอบภาพเกิดการหักล้างกันเอง ในทางปฏิบัติ เราจะต้องนำค่าสัมบูรณ์ (Absolute value) หรือค่ากำลังสอง (Squared value) ของค่าความแตกต่างของทั้งสองแกนมาบวกกันแทน

นอกจากหาความแรงของขอบภาพแล้ว การหาทิศทางของขอบภาพ (Gradient direction) ก็ มีประโยชน์เช่นกัน การหาทิศทางของขอบภาพสามารถทำได้โดยใช้สมการต่อไปนี้

$$GD(x, y) = \tan^{-1} \left\{ \frac{Y_{diff}(x, y)}{X_{diff}(x, y)} \right\} \quad (2.10)$$

เมื่อ $GD(x, y)$ คือ ทิศทางของขอบภาพที่ตำแหน่ง (x, y)

Y_{diff} คือ ค่าความแตกต่างในแนวแกนตั้ง

X_{diff} คือ ค่าความแตกต่างในแนวนอน

การหาขอบภาพโดยวิธีโซเบล การหาขอบภาพโดยวิธีโซเบล (Sobel edge detection) เป็นการหาขอบภาพโดยใช้เทมเพลตขนาด 3×3 สองเทมเพลต โดยเทมเพลตแรกจะใช้หาค่าความแตกต่างในแนวนอน (X_{diff}) และค่าความแตกต่างในแนวตั้ง (Y_{diff}) ดังแสดงในรูปที่ 2.11

$X_{diff} =$	<table><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>-2</td><td>0</td><td>2</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr></table>	-1	0	1	-2	0	2	-1	0	1	$Y_{diff} =$	<table><tr><td>1</td><td>2</td><td>3</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>-1</td><td>-2</td><td>-1</td></tr></table>	1	2	3	0	0	0	-1	-2	-1
-1	0	1																			
-2	0	2																			
-1	0	1																			
1	2	3																			
0	0	0																			
-1	-2	-1																			

รูปที่ 2.11 เทมเพลตการหาของภาพโดยวิธีโซเบล



ก)

ข)

รูปที่ 2.12 ภาพผลลัพธ์ที่ได้จากการหาขอบภาพ

ก) ภาพต้นฉบับ

ข) ภาพการหาขอบด้วยวิธี Sobel operator

2.3.6 การขยายพิกเซล

การขยายพิกเซลของภาพเป็นขั้นตอนการดำเนินการทางตรรกะ โดยทำการค้นหา (Scan) ตำแหน่งบนซ้ายไปยังตำแหน่งล่างขวา จะแทนพิกเซล 1 เมื่อมีค่าของพิกเซลใดๆ พิกเซลหนึ่งบน SE (Structuring element) ดังแสดงในรูปที่ 2.14 โดยมีสมการดังนี้

$$A \oplus B = \cup_{x \in B} (A_x) \quad (2.11)$$

B คือ Structuring Element

A คือ ภาพที่ต้องการขนาดพิกเซล

	1	
1	$\times$	1
	1	

รูปที่ 2.13 เทมเพลตการขยายฟิกเซล

The diagram illustrates the forward pass of a 2D convolution operation. It consists of two 8x8 grids. The left grid represents the input feature map, and the right grid represents the resulting feature map. A 3x3 kernel area is highlighted in red on the left grid, and the corresponding output is highlighted in yellow on the right grid. An arrow indicates the flow from the input to the output.

Input Feature Map (Left Grid):

0	1	0	1	0	0	0	0
1	x	1	1	0	0	0	0
0	1	1	1	0	0	0	0
0	0	1	1	0	0	0	0
0	0	0	1	0	0	0	0
0	0	0	1	0	0	0	0
0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0

Resulting Feature Map (Right Grid):

0	0	0	0	0	0	0	0
0	1	1	1	1	0	0	0
0	1	1	1	1	0	0	0
0	1	1	1	1	0	0	0
0	1	1	1	1	0	0	0
0	0	1	1	1	0	0	0
0	0	1	1	1	0	0	0
0	0	0	0	0	0	0	0

รูปที่ 2.14 ภาพจำลองการขยายพิกเซล

เมื่อทำการขยายพิกเซลโดยใช้การขยายพิกเซลจะได้ภาพดังแสดงในรูปที่ 2.15



ก)



๒)

รูปที่ 2.15 ภาพจำลองการขยายพิกเซล

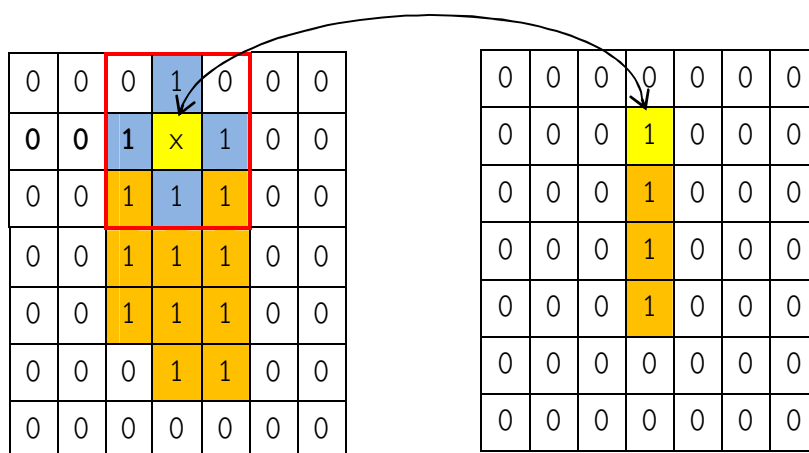
ก) ภาพต้นฉบับ

ข) ภาพหลังการขยายฟิสิกเซล

2.3.7 การกร่อนฟิกเซล

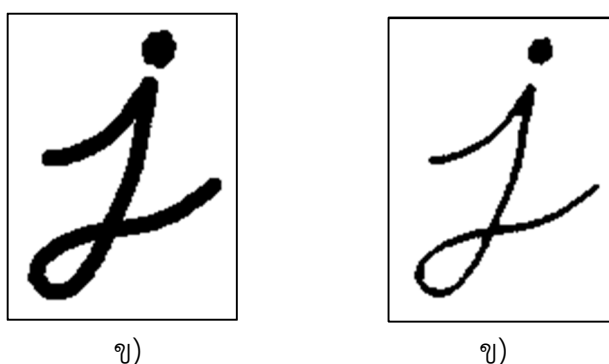
การกร่อนฟิกเซลเป็นวิธีการที่ตรงข้ามกับการขยายฟิกเซลคือการลดขนาดของฟิกเซล โดยทำการค้นหาจากตำแหน่งบนซ้ายไปยังตำแหน่งล่างขวา จะแทนฟิกเซล 0 เมื่อมีค่าของฟิกเซลบน SE ทั้งหมด ดังแสดงในรูปที่ 2.16 โดยมีสมการดังนี้

$$A \ominus B = \{w: B_w \subset A\} \quad (2.12)$$



รูปที่ 2.16 ภาพจำลองการกร่อนฟิกเซล

เมื่อทำการกร่อนฟิกเซลโดยใช้การกร่อนฟิกเซลจะได้ภาพดังแสดงในรูปที่ 2.17



รูปที่ 2.17 ภาพจำลองการกร่อนฟิกเซล

ก) ภาพต้นฉบับ

ข) ภาพหลังการกร่อนฟิกเซล

2.3.8 Retinex [2]

Retinex มาจากคำว่า Retina รวมกับ Cortex เป็นเทคนิคที่ช่วยในการปรับแสงสว่างในภาพ ในขณะที่ตาของเราสามารถมองเห็นสีได้อย่างถูกต้องถึงแม้มีแสงสว่างต่ำ แต่กล้องหรือกล้องวิดีโอที่คนไม่สามารถทำได้ Retinex สามารถคำนวณได้จากภาพขาเข้า

Single-scale retinex (SSR) เป็นการปรับความสว่างของภาพ ได้จากค่าแต่ละพิกเซลกับ พิกเซลข้างเคียงเรียกว่าฟังก์ชันเกาส์เซียน (Gaussian function) สามารถคำนวณได้จากสมการที่ (2.13)

$$R(x_1, x_2) = \log(I(x_1, x_2)) - \log(I(x_1, x_2) * F(x_1, x_2)) \quad (2.13)$$

เมื่อ I	คือ ภาพขาเข้า
R	คือ ภาพขาออกของ Single scale retinex
$\log$	คือ ฟังก์ชันลอการิทึมธรรมชาติ
(x_1, x_2)	คือ ตำแหน่งของพิกเซล
$*$	คือ ตัวดำเนินการคอนโวลูชัน (Convolution operator)
F	คือ ฟังก์ชันเซอร์ราวด์ (Surround function) โดยคำนวณได้ตามสมการที่ (2.14) ดังนี้

$$F(x_1, x_2) = K \cdot \exp[-(x_1^2 + x_2^2)/C^2] \quad (2.14)$$

เมื่อ σ	คือ ค่าเฉลี่ยมาตรฐานของเซอร์ราวด์ในฟังก์ชันแบบเกาส์เซียน (Gaussian surround)
x_1, x_2	คือ พิกัดของพิกเซล
K	คือ ค่านอร์มอลไลซ์ (Normalized) โดยคำนวณได้ตามสมการที่ (2.14) ดังนี้

$$K = \frac{1}{[\sum_{x_1} \sum_{x_2} F(x_1, x_2)]} \quad (2.15)$$

Mutis-scale retinex (MSR) เป็นการปรับความสว่างของภาพที่มีการพัฒนามาจาก Single scale retinex ที่มีจำนวนของช่องสี (Channel) เพิ่มขึ้น โดยคำนวณได้ตามสมการที่ (2.15) ดังนี้

$$R_{MSRi}(x_1, x_2) = \sum_{k=1}^K W_k \{ \log(x_1, x_2) - \log[F_k(x_1, x_2) * I_i(x_1, x_2)] \} \quad (2.16)$$

โดยที่ $i = 1, \dots, N$

เมื่อ I	คือ ภาพขาเข้า
R_{MSR}	คือ ภาพขาออกของ Multi-scale retinex
i	คือ ลำดับของช่องสี
(x_1, x_2)	คือ ตำแหน่งของพิกเซล
$*$	คือ ตัวดำเนินการคอนโวลูชัน (Convolution operator)
N	คือ ช่องสี (ภาพระดับสีเทา $N=1$, ภาพสี RGB $N=3$)
$\log$	คือ ฟังก์ชันลอการิทึมธรรมชาติ
F_k	คือ Surround function
W_k	คือ ค่าน้ำหนักที่เกี่ยวข้องกับ F_k
K	คือ สเกล
F_k	คือ Surround function โดยคำนวณได้ตามสมการ (2.17) ดังนี้

$$F_i(x_1, x_2) = K \cdot \exp[-(x_1^2 + x_2^2)/\sigma_k^2] \quad (2.17)$$

เมื่อ σ_k	คือ ค่าเฉลี่ยมาตรฐานของ Gaussian surrounds
K	คือ ค่านอร์มอลไลซ์ โดยคำนวณได้ตามสมการที่ (2.18) ดังนี้

$$K = \frac{1}{[\sum_{x_1} \sum_{x_2} F(x_1, x_2)]} \quad (2.18)$$

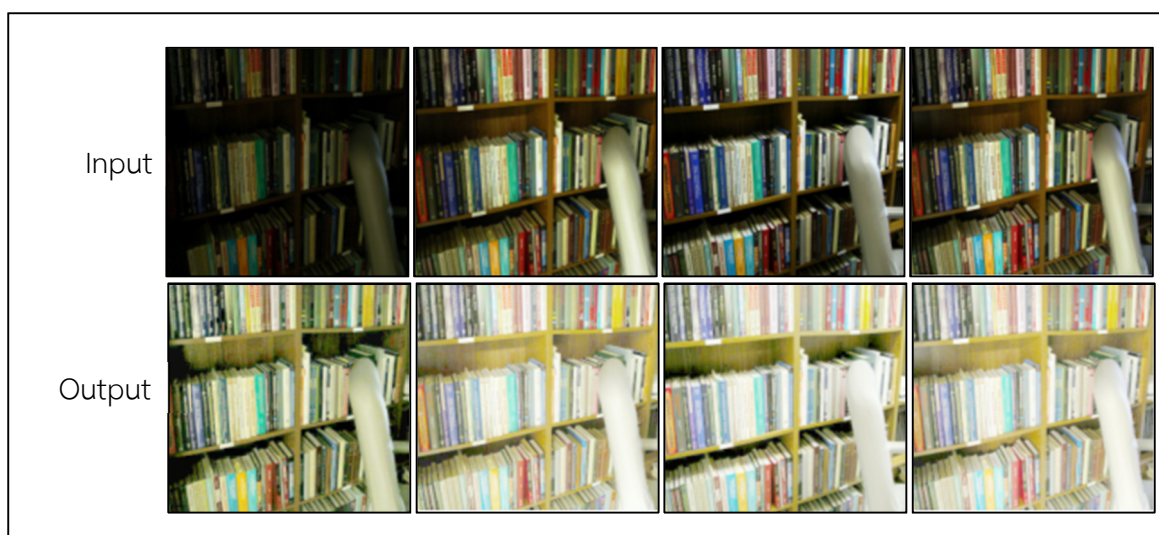
ในกรณีที่ เป็น Single scale retinex จะกำหนดให้ $K=1$ และ $W_1=1$

MSR with color restoration (MSRCR) เป็นการปรับความสว่างของภาพที่มีการพัฒนามาจาก Muti-scale retinex ที่มีพารามิเตอร์ (Parameter) ในการปรับปรุงสีในภาพเพิ่มเข้ามา โดยสามารถปรับปรุงสีให้เหมาะสมยิ่งขึ้น โดยคำนวณได้ตามสมการที่ (2.19) ถึง สมการ (2.20)

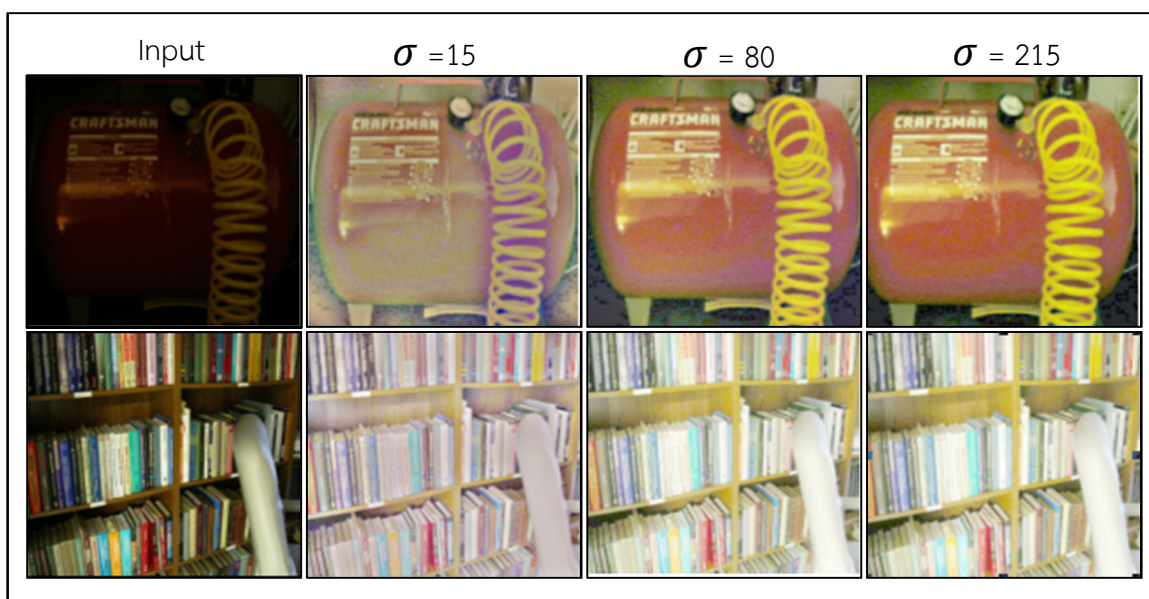
$$C_i(x_1, x_2) = \beta \left\{ \log[\alpha I_i(x_1, x_2)] - \log \left[\sum_{i=1}^N I_i(x_1, x_2) \right] \right\} \quad (2.19)$$

$$R_{MSRCRi}(x_1, x_2) = G[C_i(x_1, x_2) \{ \log I_i(x_1, x_2) - \log[F_k(x_1, x_2) * I_i(x_1, x_2)] \} + b] \quad (2.20)$$

เมื่อ I	คือ ภาพขาเข้า
R_{MSRCR}	คือ ภาพขาออกของ Multi-scale retinex
i	คือ ลำดับของช่องสี
(x_1, x_2)	คือ ตำแหน่งของพิกเซล
$*$	คือ ตัวดำเนินการคอนโวลูชัน (Convolution operator)
N	คือ ช่องสี (ภาพระดับสีเทา $N=1$, ภาพสี RGB $N=3$)
$\log$	คือ ฟังก์ชันลอการิทึมธรรมชาติ
F_k	คือ Surround function
K	คือ เลขของ Surround function หรือ สเกล
β, G	คือ ค่าคงที่อัตราขยาย (Gain constant)
α	คือ ค่าสัมประสิทธิ์ในการชดเชยค่าสี (Color restoration coefficient)
b	คือ Gain Offset value



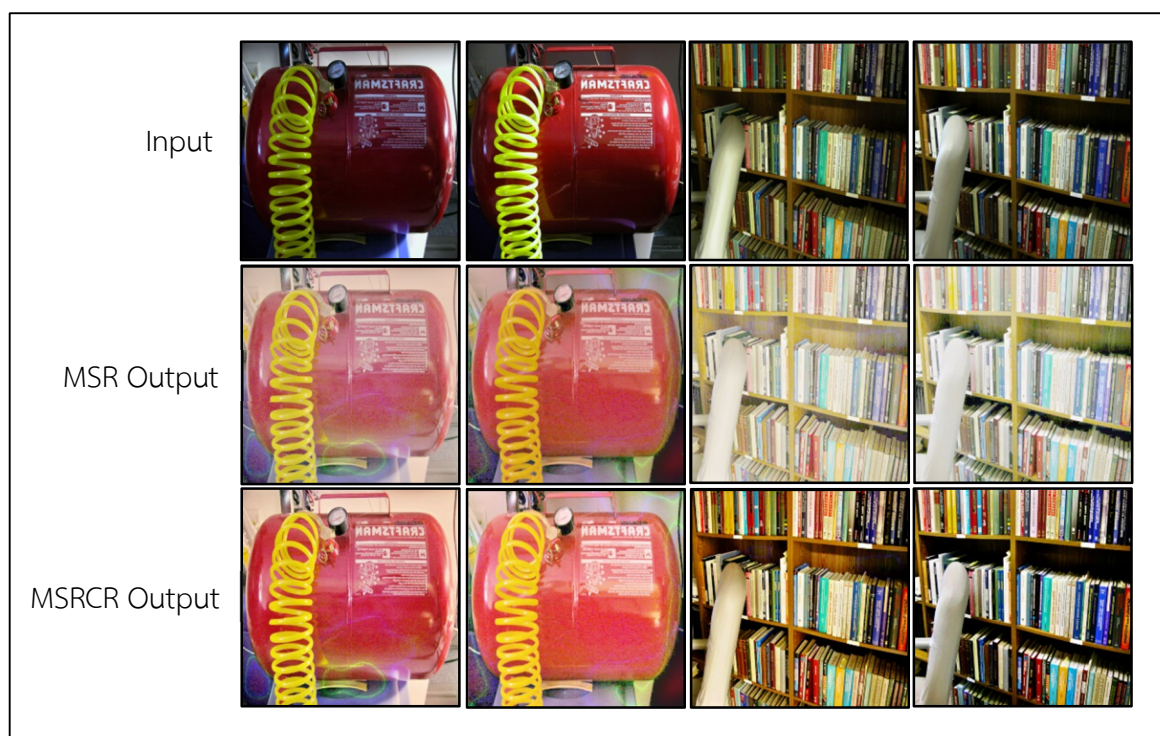
รูปที่ 2.18 การปรับปรุงโดยใช้เทคนิค Single-scale retinex



รูปที่ 2.19 การปรับปรุงโดยใช้เทคนิค Single-scale retinex

รูปที่ 2.18 เป็นการปรับปรุงโดยใช้เทคนิค Single-scale retinex โดยใช้พารามิเตอร์ $\sigma = 80$ โดยแถวบนเป็นภาพต้นแบบที่สภาพแสงต่าง ๆ และภาพแถวล่างเป็นภาพหลังการใช้เทคนิค Single-scale retinex ปรับปรุงแสงสว่างในภาพรูปที่ 2.19 เป็นการปรับปรุงโดยใช้เทคนิค Single-scale retinex โดยใช้พารามิเตอร์ $\sigma = 15, 80$ และ 215 และรูปที่ 2.20 เป็นการปรับปรุงโดยใช้เทคนิค MSR เทียบกับ MSRCR โดยใช้พารามิเตอร์ตามบทความวิจัยที่ได้นำเสนอโดย D. J. Jobson [2] ดังนี้

$$\begin{aligned}
 W_k &= 1/3 \\
 N &= 3 \\
 \sigma_1, \sigma_2, \sigma_3 &= 15, 80, 250 \\
 G &= 192 \\
 b &= 30 \\
 \alpha &= 125 \\
 \beta &= 46
 \end{aligned}$$



รูปที่ 2.20 การปรับปรุงโดยใช้เทคนิค MSR เทียบกับ MSRCR

2.4 การทบทวนวรรณกรรม/สารสนเทศ (Information) ที่เกี่ยวข้อง

การทบทวนวรรณกรรมและสารสนเทศที่เกี่ยวข้องในฐานข้อมูลต่าง ๆ สามารถสรุปเป็นตารางได้ดังต่อไปนี้

ตารางที่ 2.1: งานวิจัยที่เกี่ยวข้องด้านการตรวจจับอาการง่วงนอนขณะขับรถ

ปีที่พิมพ์ /ลำดับเอกสารอ้างอิง	คณะผู้วิจัย	องค์ความรู้ที่ได้จากบทความ
2004 [4]	Q. Ji Z. Zhu P. Lan	ในบทความนี้ได้นำเสนอการตรวจจับอาการง่วงนอนโดยใช้การสะท้อนของดวงตาจากหลอด LED ที่ติดอยู่กับกล้องเพื่อระบุตำแหน่งของดวงตาซึ่งสามารถใช้ได้ในเวลาที่ไม่มีการส่องสว่างหรือเวลากลางคืน
2011 [7]	R. C. Coetzer G. P. Hancke	ในบทความนี้ได้นำเสนอการตรวจจับอาการง่วงนอนโดยใช้ภาพจากการค้นหาดวงตา ว่ามีดวงตาดูอยู่หรือไม่โดยใช้ภาพของตา และไม่ใช่ตามารู้จำในเครือข่ายประสาทเทียมและ นำภาพทดสอบเป็นดวงตาหรือไม่

2006 [1]	Q. Wang J. Yang M. Ren Y. Zheng	ในบทความนี้ได้นำเสนอการตรวจจับอาการ่วงนอนแบบต่างๆ ที่มีเช่นการตรวจจับอาการ่วงนอนโดยใช้การสังเกตจากการเคลื่อนไหวของดวงตา อัตราการกระพริบตา ตำแหน่งของศีรษะ ลักษณะของปาก การตรวจจับโดยใช้ Active IR การตรวจจับโดยสังเกตการติดตามช่องทางเดินรถ
2009 [5]	P. R. Tabrizi R. A. Zoroofi	ในบทความนี้ได้นำเสนอการตรวจจับอาการ่วงนอนโดยการวิเคราะห์จากอัตราการกระพริบตาโดยหลังจากการตรวจจับใบหน้าแล้วจะทำการตรวจจับดวงตาในปริภูมิสี YCbCr และใช้ระยะห่างระหว่างขอบตาบนกับขอบตาล่างบอกสถานะของตา (หลับ/ไม่หลับตา)
2010 [6]	S. Ribaric J. Lovrencic N. Pavesic	ในบทความนี้ได้นำเสนอการตรวจจับอาการ่วงนอนโดยสังเกตจากมุมการเอียงของศีรษะ โดยการนำเทคนิคทางปัญญาประดิษฐ์ นั่นคือเครือข่ายประสาทเทียมมาใช้ในการหามุมของศีรษะว่าอยู่ในระนาบ (In-plane) หรือ นอกนอกระนาบ (Out-of-plane)
2012 [8]	X. Li E. Seignez P. Loonis	ในบทความนี้ได้นำเสนอการตรวจจับอาการ่วงนอนโดยนำวิธีการ 2 วิธีการมารวมกันได้แก่ การติดตาม (Tracking) ช่องทางเดินรถบนถนน และสถานะของใบหน้า ตัวอย่างเช่น มุมเอียงของใบหน้า ซึ่งเป็นเพิ่มความแม่นยำในการตรวจจับเพิ่มขึ้น
2012 [9]	I. Garcia S. Bronte L. M. Bergasa J. Almazan J. Yebes	ในบทความนี้ได้นำเสนอการตรวจจับอาการ่วงนอนจากการเปรียบเทียบอัตราการกระพริบตาโดยจะหาดวงตาจากการใช้อัลกอริทึม Integral projection ทั้งแนวนอน (Horizontal) และแนวตั้ง (Vertical) เพื่อหาสถานะของตาโดยสังเกตจากเปอร์เซ็นต์ของการลืมนตา ในบทความนี้ยังได้ทดสอบกับสถานะการใช้งานจริงอีกด้วย

ตารางที่ 2.2: งานวิจัยที่เกี่ยวข้องด้านการตรวจจับใบหน้า

ปีที่พิมพ์ /ลำดับเอกสารอ้างอิง	คณะผู้วิจัย	องค์ความรู้ที่ได้จากบทความ
2010 [10]	S. Kherchaoui A. Houacine	ในบทความนี้ได้นำเสนอการตรวจจับใบหน้าโดยใช้การคัดแยกสีผิวและลักษณะเฉพาะของใบหน้า โดยจะคัดแยกในปริภูมิสี CbCr และใช้ลักษณะเฉพาะของใบหน้าคือ ความกว้างและสูงของใบหน้า
2011 [12]	P. P. Paul M. Gavilova	ในบทความนี้ได้นำเสนอการตรวจจับใบหน้าโดยใช้ การวิเคราะห์ส่วนประกอบสำคัญการหาขอบวัตถุแบบ Canny และการคัดแยกสีผิวจึงทำให้ประสิทธิภาพในการตรวจจับใบหน้าที่มีความถูกต้องสูง
2011 [11]	K. S. Park R. H. Park Y. G. Kim	ในบทความนี้ได้นำเสนอการตรวจจับใบหน้าโดยใช้บล็อก 3x3 เป็นเทมเพลตในการหาใบหน้าโดยเทมเพลตนี้จะหาความเข้มในแต่ละบล็อกเพื่อนำไปเปรียบเทียบกับภาพใบหน้าที่ต้องการทดสอบโดยลำดับความเข้มในบล็อก 3x3 ของภาพเทมเพลตและภาพใบหน้าที่นำมาทดสอบตรงกันจะถือว่าตำแหน่งนั้นเป็นใบหน้า
2011 [13]	M. L. Chiang S. H. Lau	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยใช้การหาขอบของวัตถุและเปรียบเทียบกับารตรวจจับใบหน้าโดยใช้การหาขอบของวัตถุแบบ Haar และ Canny ซึ่งในบทความนี้สามารถตรวจจับใบหน้าได้ถึงแม้ว่าจะมีแสงเงา ขนตา และการเอียงของใบหน้าที่แตกต่างกัน

ตารางที่ 2.3: งานวิจัยที่เกี่ยวข้องด้านการตรวจจับดวงตา

ปีที่พิมพ์ /ลำดับเอกสารอ้างอิง	คณะผู้วิจัย	องค์ความรู้ที่ได้จากบทความ
2007 [14]	C. W. Park J. M. Kwak H. Park Y. S. Moon	ในบทความนี้ได้นำเสนอการค้นหาดวงตาโดยใช้ตัวกรองกาบอร์ (Gabor filter) เพื่อหาเส้นแนวนอนของตาและใช้ตัวกรอง ART (Angular radial transform) เพื่อหาลักษณะของตาที่เป็นวงกลมและรวมทั้งสองจะคำนวณหาตำแหน่งของตาออกมา

2009 [15]	N. V. Huan N. T. H. Binh H. Kim	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยตัวกรองวงกลม (Circular filter) ปกติตาจะมีลักษณะเป็นวงกลมหรือเป็นส่วนหนึ่งของวงกลม บทความนี้จึงได้นำลักษณะของตาตำานี้มาใช้ในการระบุตำแหน่งของตา
2011 [18]	S. Chen C. Liu	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยใช้ปริภูมิสี YUV แล้วใช้สมการ Projection functions เพื่อค้นหาบริเวณของตา แล้วเลือกใช้การหาความเข้มในบริเวณนั้นเพื่อหาดำแหน่งของตาโดยแปลงภาพเป็นระดับสีเทาโดยที่ดวงตาจะมีค่าความเข้มที่ต่ำ วิธีการนี้เป็นวิธีการตรวจจับดวงตาที่มีความเร็ว สามารถตรวจจับดวงตาถูกต้องถึง 95.4 เปอร์เซ็นต์
2010 [16]	N. Alioua A. Amine M. Rziza D. Aboutajdine	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยใช้ Circular Hough Transform (CHT) ซึ่งเป็นอัลกอริทึมที่มีความใหม่ที่ไม่ต้องใช้ภาพในการเรียนรู้ เมื่อนำอัลกอริทึมนี้มาเปรียบเทียบการหาขอบวัตถุพบว่ามีความถูกต้องสูงกว่า
2011 [19]	M. A. A. Kashani M. M. Arani	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยอาศัยลักษณะของดวงตา เช่น คิ้ว หน้ตา เป็นต้น หลังจากตรวจจับใบหน้าโดยใช้สีผิวจะทำการหาดวงตาโดยปกติแล้วตาจะมีลักษณะเป็นเส้นแนวนอนบทความนี้จะนำลักษณะนี้มาใช้หาดวงตา
2012 [20]	C. Phromsuthirak S. Umchid	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยใช้รูปร่างลักษณะของตาและสีในการคัดแยก จะใช้สมการ EyemapC ที่ได้จากปริภูมิสี Cb และสมการ EyemapL ที่ได้จากปริภูมิสี Cr และใช้ลักษณะของความกว้างและความสูงของตาในการระบุตำแหน่งของตา
2012 [21]	P. Goel S. Agarwal	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาที่นำสองอัลกอริทึมรวมกันได้แก่ Haar cascade classifiers และคุณสมบัติต่าง ๆ บนใบหน้าเพื่อให้การตรวจจับมีประสิทธิภาพมากขึ้นเห็นได้จากผลการทดสอบที่มีการตรวจจับสูงถึง 98.75 เปอร์เซ็นต์

2012 [22]	Md. H. Rahman F. Jhumur Md. S. U. Yusuf T. Das M. Ahmsad	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยใช้ลักษณะของตาทั้งของข้างกับปากซึ่งเป็นรูปสามเหลี่ยม ทำการแปลงภาพให้อยู่ในปริภูมิสี YCbCr จากนั้นจะทำการหาไบหน้าโดยใช้การคัดแยกสีผิว จะพบว่าจะมีช่องว่างสามจุดของตาและปากมีลักษณะเป็นรูปสามเหลี่ยม
2011 [17]	D. Y. Huang T. W Lin W. C. Hu M. S. Chen	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยใช้สีในการคัดแยกสามารถตรวจจับได้ในแสงสว่างที่เปลี่ยนแปลงสามารถหาได้จากการตรวจจับไบหน้าโดยการตรวจจับไบหน้าจะมีช่องว่างและระบุช่องว่างนี้เป็นดวงตา

ตารางที่ 2.4: งานวิจัยที่เกี่ยวข้องกับการตรวจจับอาการง่วงนอนโดยใช้เทคนิคทางปัญญาประดิษฐ์

ปีที่พิมพ์ /ลำดับเอกสารอ้างอิง	คณะผู้วิจัย	องค์ความรู้ที่ได้จากบทความ
2005 [23]	F. H. C. Tivive A. Bouzerdoun	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยการใช้เครือข่ายประสาทเทียมระบบสามารถแยกภาพที่มีตา และไม่มีตา ในการใช้เครือข่ายประสาทเทียมในบทความนี้สามารถบอกความถูกต้องในการคัดแยกดวงตาได้ถึง 99 เปอร์เซ็นต์
2007 [24]	T. Akashi Y. Wkasa K. Tanaka S. Karungaru M. Fukumi	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาและติดตามดวงตาโดยใช้เทคนิคจินเนติกอัลกอริทึม (Genetic algorithm) เป็นการติดตามดวงตาในภาพวิดีโอมีความถูกต้องสูงถึง 97.9 เปอร์เซ็นต์ และใช้เวลาในการตรวจจับดวงตาโดยเฉลี่ย 28 มิลลิวินาทีต่อเฟรม
2011 [7]	R. C. Coetzer G. P. Hancke	ในบทความนี้ได้นำเสนอการตรวจจับอาการง่วงนอนแบบเวลาจริงโดยใช้เทคนิคทางปัญญาประดิษฐ์ได้แก่ เครือข่ายประสาทเทียม, เครื่องเวกเตอร์เกื้อหนุน (Support vector machines: SVM) และวิธีการค้นหาแบบตามูเชิงปรับตัว (Adaptive boosting: AdaBoost)

บทที่ 3

การออกแบบระบบเฝ้าติดตามอาการง่วงนอน

3.1 บทนำ

เนื้อหาในบทนี้กล่าวถึงการออกแบบระบบเฝ้าติดตามอาการง่วงนอน โดยได้แบ่งเป็นหัวข้อดังนี้ การตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม การตรวจจับใบหน้าโดยใช้การคัดแยกสีผิว การตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน การตรวจจับดวงตาโดยใช้เทคนิค Integral projection และในส่วนสุดท้ายจะเป็นบทสรุป

3.2 การออกแบบระบบตรวจจับใบหน้า

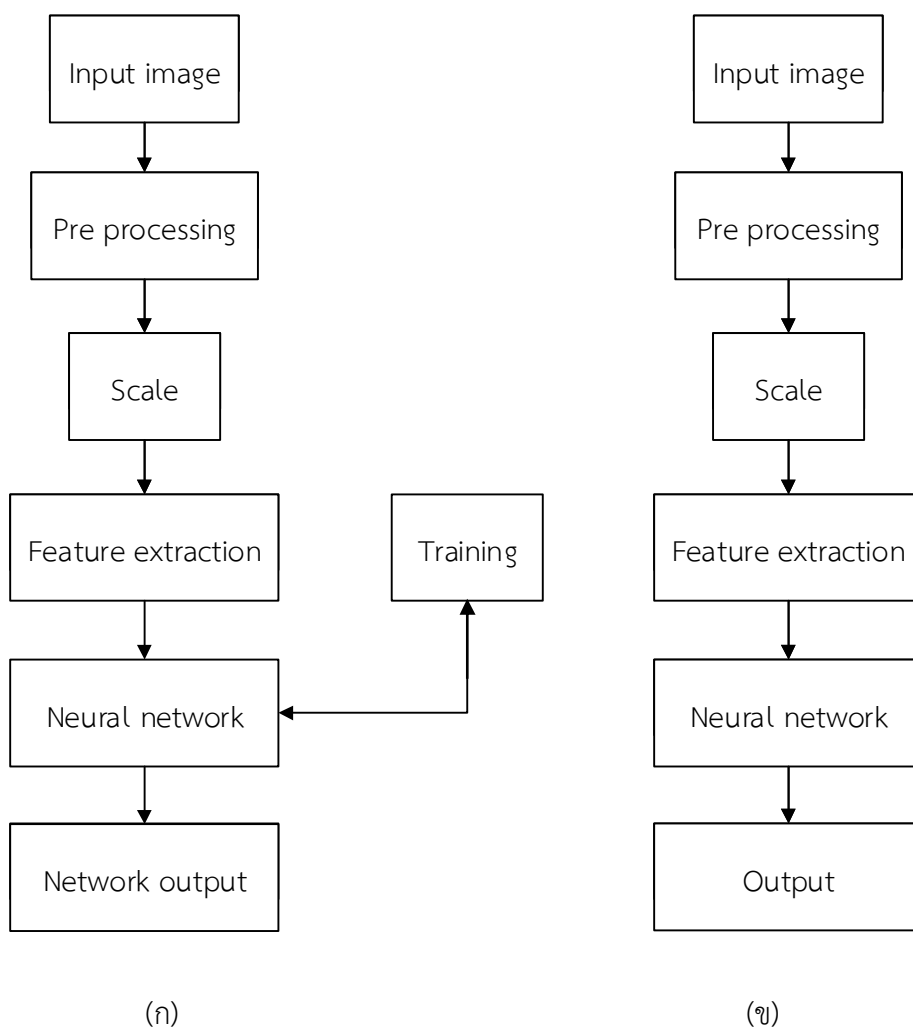
การตรวจจับอาการง่วงนอนโดยอาศัยการสังเกตอาการของผู้ขับรถ ในงานวิจัยนี้ได้ใช้ดวงตาเป็นตัวบ่งชี้ถึงระดับอาการง่วงนอนของผู้ขับรถ ในการค้นหาดวงตาในเฟรมภาพซึ่งมีขนาดใหญ่นั้นใช้เวลาในการประมวลผลมาก เพื่อลดเวลาในการค้นหาดวงตาลง ต้องค้นหาบริเวณของใบหน้าให้ได้เสียก่อน การตรวจจับใบหน้าในบทนี้ได้ศึกษาวิธีการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม และการตรวจจับใบหน้าโดยใช้สีผิว

การตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมเป็นการนำภาพใบหน้าตัวอย่างมาให้เครือข่ายประสาทเทียมรู้จำ โดยภาพใบหน้าที่นำมาใช้ในการรู้จำมีการดึงลักษณะเด่นด้วยการวิเคราะห์ส่วนประกอบสำคัญ (Principal components analysis: PCA) และการแปลงเวฟเล็ตแบบดิสครีท (Discrete wavelet transform: DWT)

การตรวจจับใบหน้าโดยใช้การคัดแยกสีผิวเป็นการหาช่วงของสีผิวที่อยู่ในปริภูมิสี RGB และ HSV แล้วนำเทคนิคการประมวลผลภาพดิจิทัลมาใช้ประมวลผลเช่น การหาขอบภาพ (Edge detection) การแปลงปริภูมิสี (Color transformations) การขยายพิกเซล (Dilation) การกร่อนพิกเซล (Erosion)

3.2.1 การออกแบบการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม

ในหัวข้อนี้เป็นการออกแบบการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม เป็นการนำภาพใบหน้ามาให้เครือข่ายประสาทเทียมรู้จำ โดยจะทำการดึงลักษณะเด่นด้วยการวิเคราะห์ส่วนประกอบสำคัญ และการแปลงเวฟเล็ตแบบดิสครีท โดยมีแผนภาพการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม แสดงในรูปที่ 3.1



รูปที่ 3.1 ระบบตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม

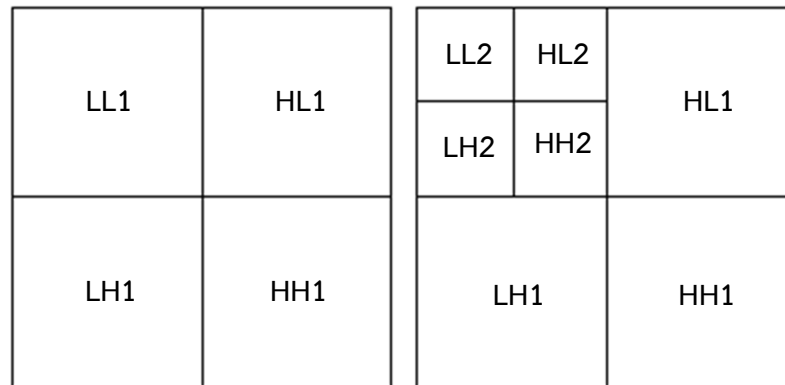
ก) การรู้จำของเครือข่ายประสาทเทียม

ข) การทดสอบโดยใช้เครือข่ายประสาทเทียม

การนำเครือข่ายประสาทเทียมมาใช้ในการค้นหาใบหน้าจะต้องมีการเตรียมฐานข้อมูลเพื่อการรู้จำของเครือข่ายประสาทเทียม ในงานวิจัยนี้ได้ใช้ฐานข้อมูลภาพใบหน้าและภาพที่ไม่ใช่ใบหน้าขนาด 200×200 พิกเซล อย่างละ 100 ภาพซึ่งเป็นฐานข้อมูลที่ผู้วิจัยได้สร้างขึ้นเองดังแสดงในรูปที่ 3.2 และรูปที่ 3.3 หลังจากนั้นจะนำภาพดังกล่าวมาปรับแสงสว่าง และทำการแปลงเวฟเล็ตแบบดีสครีทจำนวน 2 ระดับ และเลือกใช้ภาพแบนด์ย่อยความถี่ต่ำสุดที่มีขนาด 50×50 พิกเซล หลังจากนั้นจะนำภาพแบนด์ย่อยความถี่ต่ำดังกล่าวมาดึงลักษณะเด่นของภาพโดยใช้การวิเคราะห์ส่วนประกอบสำคัญและนำไปให้เครือข่ายประสาทเทียมรู้จำ

การดึงลักษณะเด่นโดยการแปลงเวฟเล็ตแบบดิสครีท

การแปลงเวฟเล็ตแบบดิสครีทเป็นการแปลงสัญญาณให้เป็นอนุกรมของเวฟเล็ต โดยเซตของผลลัพธ์จะเรียกว่าสัมประสิทธิ์โดยการแปลงเวฟเล็ตแบบดิสครีทระดับที่ 1 และระดับที่ 2 แสดงในรูปที่ 3.4



ก)

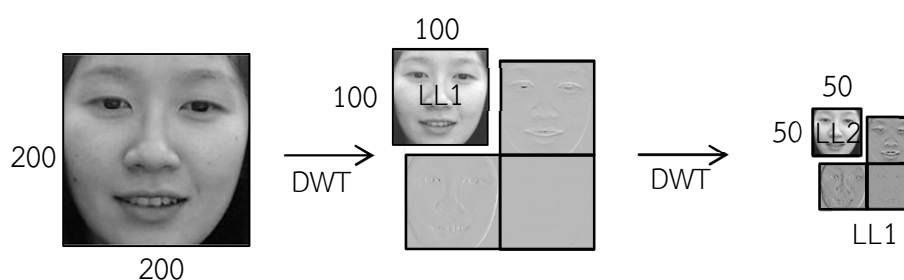
ข)

รูปที่ 3.4 การแปลงเวฟเล็ตแบบดิสครีท

ก) การแปลงเวฟเล็ตแบบดิสครีทระดับที่ 1

ข) การแปลงเวฟเล็ตแบบดิสครีทระดับที่ 2

เมื่อนำภาพใบหน้าขนาด 200×200 พิกเซลมาแปลงเวฟเล็ตแบบดิสครีทในระดับที่ 2 จะเหลือขนาดของภาพแบนด์ย่อย (Subband) ความถี่ต่ำสุด (LL2) ที่จะนำไปประมวลผลต่อเท่ากับ 50×50 พิกเซล จะได้ดังแสดงในรูปที่ 3.5



รูปที่ 3.5 การแปลง DWT

การดึงลักษณะเด่นด้วยการวิเคราะห์ส่วนประกอบสำคัญ

การวิเคราะห์ส่วนประกอบสำคัญ (Principal components analysis: PCA) เป็นวิธีการทางสถิติใช้สร้างเมทริกซ์ของความแปรปรวนร่วม (Covariance Matrix) จากข้อมูลภาพ ถูกนำไปใช้ในการบีบอัดข้อมูล การวิเคราะห์องค์ประกอบหลักมาใช้ลดขนาดข้อมูล ข้อมูลของแพคเกจนั้นอยู่ในรูปของเวกเตอร์ 1 มิติ

อยู่แล้ว นำเวกเตอร์ของทุกชุดข้อมูลมาจัดให้อยู่ในรูปแบบของเมทริกซ์ โดยเวกเตอร์ของชุดข้อมูลที่ 1 จะเป็นแถวที่ 1 ของเมทริกซ์ เวกเตอร์ของชุดข้อมูลที่ 2 จะเป็นแถวที่ 2 ของเมทริกซ์จนถึงเวกเตอร์ของชุดข้อมูลที่ n จะเป็นแถวที่ n ของเมทริกซ์ ดังนั้นจะได้เมทริกซ์ A มีมิติเป็น i และ j โดยที่ i หมายถึงข้อมูลชุดที่ และ j หมายถึงมิติที่ ดังแสดงในรูปที่ 3.6

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{bmatrix} \quad m = w \times h; n = \text{number of pattern}$$

รูปที่ 3.6 เมทริกซ์ A ที่ใช้ในการคำนวณสำหรับการวิเคราะห์องค์ประกอบ

คำนวณค่าเฉลี่ยของภาพในแต่ละหลักดังสมการที่ (3.1)

$$M_j = \frac{1}{n} \sum_{i=1}^n A_{ij} \quad 1 \leq j \leq m \quad (3.1)$$

คำนวณค่าเบี่ยงเบนมาตรฐานของภาพใบหน้าดังสมการที่ (3.2)

$$C_{ij} = A_{ij} - M_j \quad 1 \leq i \leq n, 1 \leq j \leq m \quad (3.2)$$

สร้างเมทริกซ์ของความแปรปรวนดังสมการที่ (3.3)

$$S = \frac{1}{n} \sum_{i=1}^n C_i C_i^T \quad (3.3)$$

คำนวณค่าไอเกนดังสมการที่ (3.4)

$$\lambda = SS^T \quad (3.4)$$

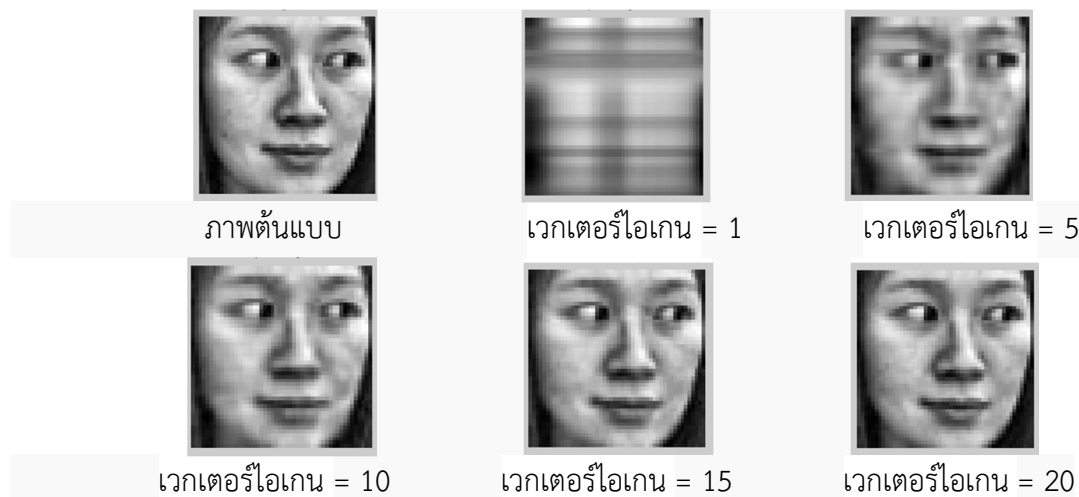
คำนวณเวกเตอร์ไอเกน (Eigen vectors) ดังสมการที่ (3.5)

$$e = \frac{CS}{\sqrt{\lambda}} \quad \text{เมื่อ } C = [c_{ij}] \quad (3.5)$$

ผลลัพธ์จะได้ค่าไอเกน และเวกเตอร์ไอเกน ข้อมูลทั้ง 2 มีความสมนัยกัน (Correspondence) ซึ่งกันและกัน เวกเตอร์ไอเกนเป็นข้อมูลที่เปลี่ยนรูปแล้ว ไม่สามารถเห็นเป็นเวกเตอร์แพกเก็ตแบบเดิมได้ การนำข้อมูลเดิมกลับมาต้องนำเวกเตอร์ไอเกนมาคำนวณอีกครั้ง ดังสมการ (3.5) เมื่อไอเกนแพกเก็ตแรก (1st Eigen Package) คือการเลือกเวกเตอร์ไอเกนแรกของทุกเวกเตอร์ข้อมูลมาคำนวณ ส่วนไอเกนแพกเก็ตอื่น ๆ ก็ทำในรูปแบบเดียวกัน

การเลือกเวกเตอร์ไอเกน

นำภาพหลังจากการแปลงเวฟเล็ตแบบดิสครีทที่มีขนาด 50×50 พิกเซลมาดึงลักษณะเด่นด้วยการวิเคราะห์ส่วนประกอบสำคัญ ในการเลือกค่าเวกเตอร์ไอเกนของการวิเคราะห์ส่วนประกอบสำคัญจะทดลองที่ค่าต่าง ๆ ตามรูปที่ 3.7

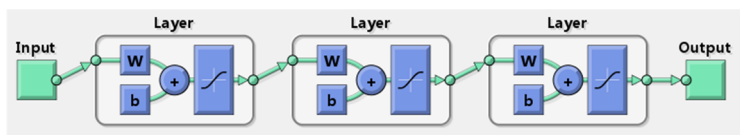


รูปที่ 3.7 การทดลองเลือกค่าเวกเตอร์ไอเกนที่ค่าต่าง ๆ

จากรูปที่ 3.7 จะเห็นว่าภาพเริ่มสามารถมองเห็นลักษณะของใบหน้าได้ที่ เวกเตอร์ไอเกน = 5 เพื่อลดจำนวนข้อมูลในการรู้จำของเครือข่ายประสาทเทียมลงจึงเลือกค่าเวกเตอร์ไอเกนที่มีค่าน้อย ๆ แต่สามารถคงความเป็นลักษณะเด่นของภาพเดิมอยู่

เมื่อนำข้อมูลที่ได้หลังการวิเคราะห์ส่วนประกอบสำคัญเข้าสู่การรู้จำของเครือข่ายประสาทเทียมจะเหลือข้อมูลที่ใช้รู้จำเป็น 250×200 เมื่อนำมาทดสอบกับภาพที่เป็นใบหน้า 30 ภาพและภาพที่ไม่ใช่ใบหน้า 30 ภาพ โดยกำหนดค่าในการรู้จำดังนี้

- 1) Error goal = 10×10^{-8}
- 2) จำนวนเลเยอร์ = 3 เลเยอร์

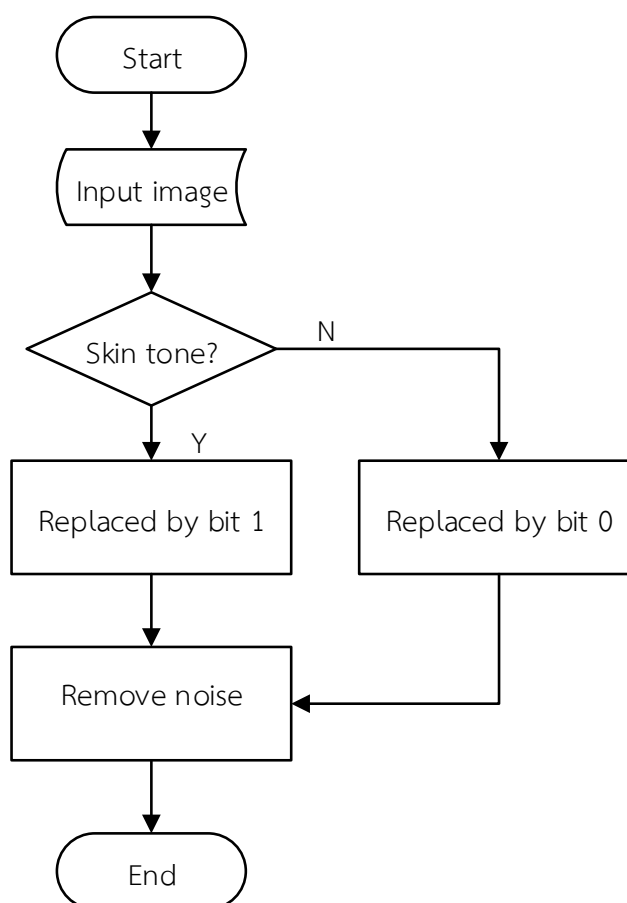


รูปที่ 3.8 เลเยอร์ของเครือข่ายประสาทเทียม

3) จำนวนรอบสูงสุด = 10000 รอบ

3.2.2 การออกแบบการตรวจจับใบหน้าโดยใช้การคัดแยกสีผิว

ในการการออกการตรวจจับใบหน้าโดยใช้การคัดแยกสีผิวผู้วิจัยได้ทำการเลือกใช้ปริภูมิสี RGB และ HSV มาใช้ จากหัวข้อ 3.2.1 การตรวจจับใบหน้าโดยใช้เทคนิคทางปัญญาประดิษฐ์สามารถตรวจจับได้เฉพาะใบหน้าที่เป็นหน้าตรงเนื่องจากภาพที่นำไปให้เครือข่ายประสาทเทียมรู้จำนั้นเป็นภาพใบหน้าตรงและมีขนาดคงที่จึงไม่เหมาะที่จะนำมาใช้กับการตรวจจับอาการง่วงนอนเนื่องจากขนาดของใบหน้าที่ได้จากกล้องเว็บแคมมีขนาดไม่คงที่เนื่องจากผู้ขับรถมีการหันซ้ายหรือหันขวา แนวทางหนึ่งที่สามารถแก้ปัญหานี้ได้คือทำการออกแบบการตรวจจับใบหน้าให้เหมาะสมโดยเลือกใช้การคัดแยกสีผิว โดยมีแผนภาพดังแสดงในรูปที่ 3.9



รูปที่ 3.9 แผนภาพการตรวจจับใบหน้าโดยการคัดแยกสีผิว

การหาช่วงของสีผิว สามารถทำได้โดยนำภาพใบหน้ามาแยกเป็นปริภูมิสี R, G, B และปริภูมิสี H, S, V โดยการแปลงปริภูมิสี RGB เป็น HSV ดังสมการที่ (3.6) ถึง (3.9)

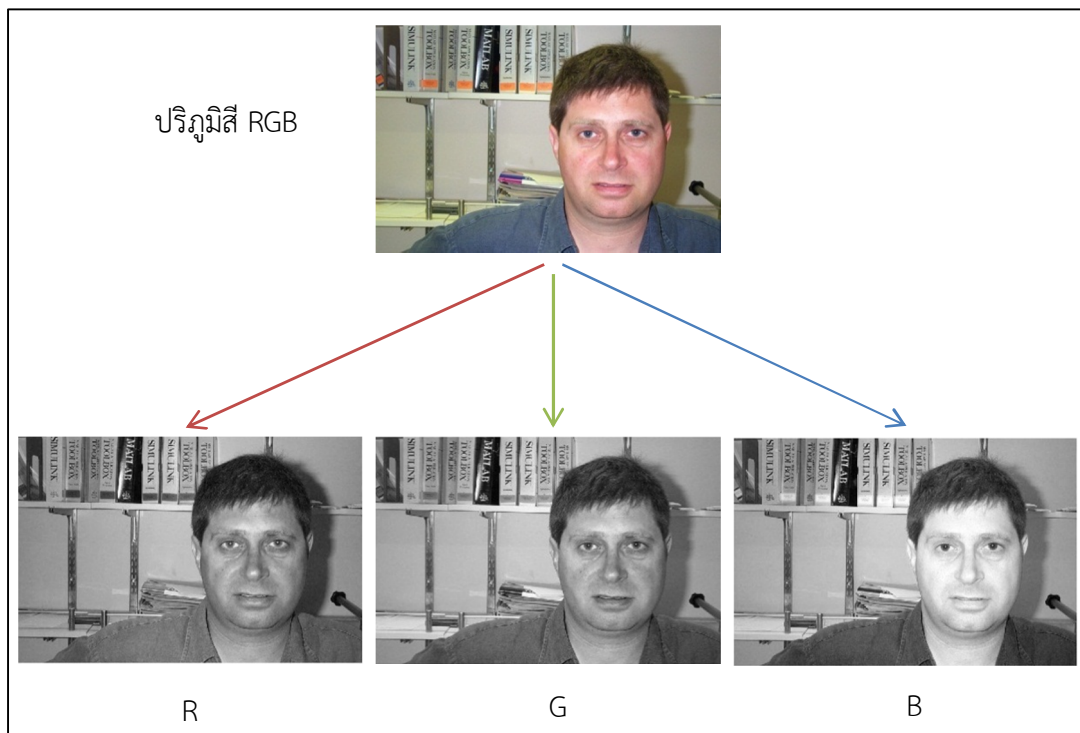
$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360^\circ - \theta & \text{if } B > G \end{cases} \quad (3.6)$$

$$\theta = \cos^{-1} \left\{ \frac{0.5[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)^2}} \right\} \quad (3.7)$$

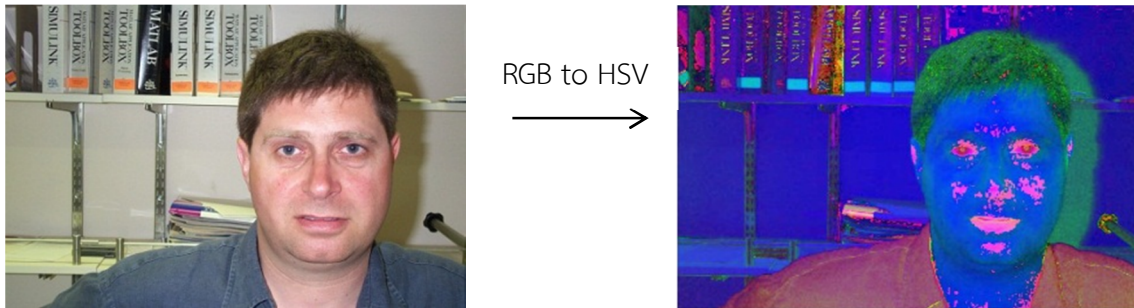
$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (3.8)$$

$$V = \frac{\max(R, G, B)}{255} \quad (3.9)$$

การตัดแยกสีผิว สีผิวเป็นช่วงสีกลุ่มหนึ่งในช่วงสีที่มีอยู่มากมาย เนื่องจากปริภูมิสี RGB เป็นปริภูมิสีที่แปรตามสภาพแสงได้ง่าย จึงนำปริภูมิสี HSV เข้ามาช่วยในการตัดแยกสีผิวโดยขั้นแรกนำภาพที่จะทดสอบมาแยกปริภูมิสี R G และ B ดังแสดงในรูปที่ 3.10 และแปลงปริภูมิสีจาก RGB เป็น HSV ดังแสดงในรูปที่ 3.11



รูปที่ 3.10 การแยกปริภูมิสี RGB



รูปที่ 3.11 การแปลงปริภูมิสี RGB เป็น HSV

นำภาพที่ได้ไปประมวลตามเงื่อนไขที่กำหนดตามสมการที่ (3.10) จะได้ดังแสดงในรูปที่ 3.12

$$\text{Detec}(x,y) = \begin{cases} 255 & \text{if } R(x,y) > G(x,y) \text{ and } |R(x,y) - G(x,y)| \geq 11 \\ & \text{and } 84 \leq R(x,y) \leq 153 \text{ and } 64 \leq G(x,y) \leq 94 \\ & \text{and } 30 \leq S(x,y) \leq 178 \text{ and } 76 \leq V(x,y) \leq 255 \\ 0 & \text{otherwise} \end{cases} \quad (3.10)$$

เมื่อ (x,y) คือ พิกัดของพิกเซลในภาพ

Detec คือ ภาพเอาต์พุตจากการประมวลผล



รูปที่ 3.12 เอาต์พุตการคัดแยกสีผิว

การตัดบริเวณที่ไม่ใช่ใบหน้า ทำได้โดยนับจำนวนพิกเซลในแต่ละวัตถุเพื่อตัดวัตถุที่มีจำนวนพิกเซลที่มากเกินไปหรือน้อยเกินไป การขยายพิกเซล และการกร่อนพิกเซล เพื่อกำจัดจุดเล็กๆ ที่ไม่ใช่ใบหน้า (Noise) โดยทั่วไปอัตราส่วนระหว่างความสูงต่อความกว้างของใบหน้าจะอยู่ระหว่าง 0.8 ถึง 2.6 และมีจำนวนของพิกเซลที่เป็นสีผิวมากกว่า 40 เพอร์เซ็นต์ หลังจากการกำจัดบริเวณที่ไม่ใช่ใบหน้าออก จะได้ดังแสดงในรูปที่ 3.13 และนำบริเวณดังกล่าวมา Crop จากภาพที่นำมาตรวจจับจะได้ดังแสดงในรูปที่ 3.14



รูปที่ 3.13 ภาพหลังการกำจัดบริเวณที่ไม่ใช่ใบหน้าออก



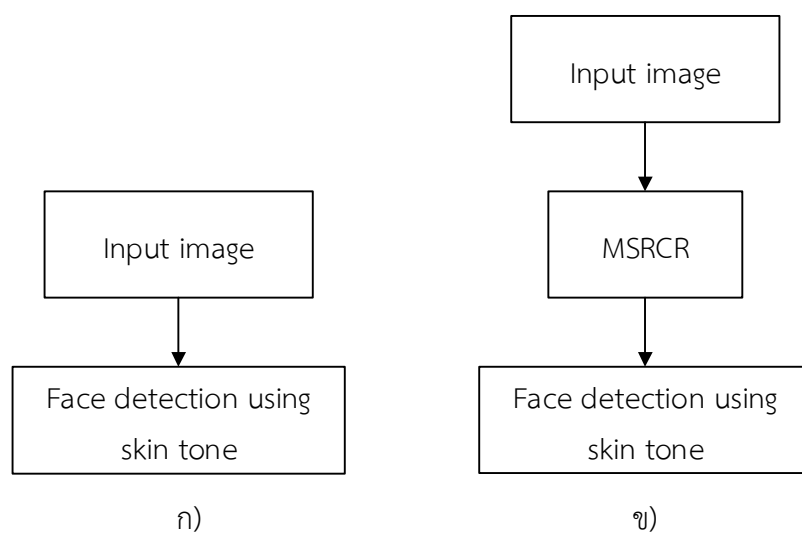
รูปที่ 3.14 ภาพหลังการตรวจจับ

3.2.3 การออกแบบการตรวจจับใบหน้าโดยใช้การตัดแยกสีผิวและเทคนิค MSRCR

การตรวจจับใบหน้าโดยใช้การตัดแยกสีผิวในหัวข้อ 3.2.2 เป็นการตรวจจับที่ใช้สีในการตัดแยกใบหน้าออกจากภาพพื้นหลัง ซึ่งการทดสอบในสภาวะใช้งานจริงที่มีแสงสว่างเปลี่ยนแปลงตลอดเวลา จึงต้องมีเทคนิคที่จะช่วยในการปรับปรุงภาพ (Image enhancement) ให้แสงสว่างภายในภาพเหมาะสมจนสามารถตรวจจับใบหน้าโดยการตัดแยกสีผิวได้ เทคนิคนี้คือ MSRCR (MSR with color restoration) ซึ่งเป็นเทคนิคที่ช่วยในการปรับแสงสว่างในภาพ โดยจะนำเทคนิค MSRCR มาเพื่อปรับแสงสว่างในภาพก่อนการตรวจจับใบหน้าโดยการตัดแยกสีผิวดังแสดงในรูปที่ 3.15 เทคนิค MSRCR สามารถช่วยปรับแสงในภาพที่มีแสงมืดเกินไป และสว่างเกินไปให้มีแสงสว่างที่เหมาะสมต่อการตรวจจับใบหน้า เนื่องจากในสภาวะใช้งานจริงปริมาณแสงสว่างในภาพจะมีการเปลี่ยนแปลงตลอดเวลา ในงานวิจัยนี้ได้ทำการเลือกใช้เลือกใช้พารามิเตอร์ของเทคนิค MSRCR ให้เหมาะสมกับรูปภาพขนาด 1920×1080 พิกเซล ซึ่งได้จากกล้องเว็บแคม รุ่น Logitech HD Pro Webcam C920 ดังนี้

ตารางที่ 3.1: พารามิเตอร์ของเทคนิค MSRCR ที่ใช้ในงานวิจัย

พารามิเตอร์	ค่าที่ใช้ในงานวิจัย
W_k	1/3
N	3
$\sigma_1, \sigma_2, \sigma_3$	500, 800, 1500
G	192
b	30
α	125
β	46



รูปที่ 3.15 แผนภาพการตรวจจับใบหน้าด้วยวิธีการ

ก) การตรวจจับใบหน้าทั่วไป

ข) การตรวจจับใบหน้าที่ใช้เทคนิค MSRCR



ก)



ข)



ค)



ง)



จ)



ฉ)

รูปที่ 3.16 เปรียบเทียบ Retinex แบบต่างๆ

ก) ภาพต้นแบบ

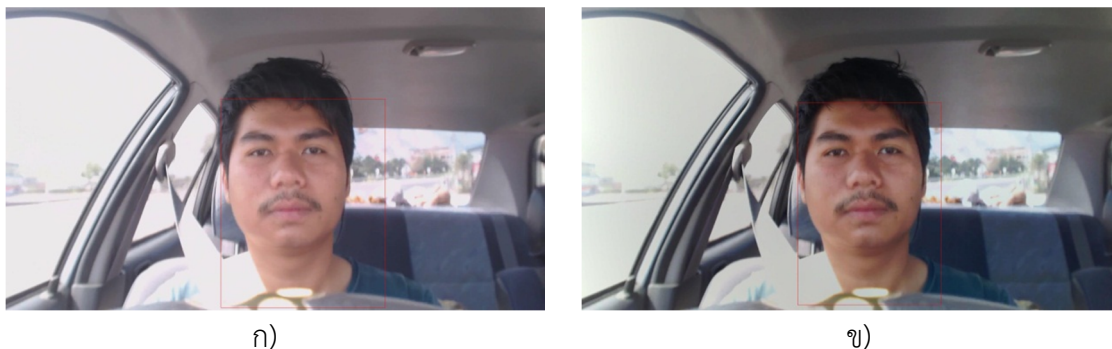
ข) ภาพที่ผ่าน SSR $\sigma = 500$

ค) ภาพที่ผ่าน SSR $\sigma = 800$

ง) ภาพที่ผ่าน SSR $\sigma = 1500$

จ) ภาพที่ผ่าน MSR $\sigma_1, \sigma_2, \sigma_3 = 500, 800, 1500$

ฉ) ภาพที่ผ่าน MSRCR $W_k = 1/3, N = 3, \sigma_1, \sigma_2, \sigma_3 = 500, 800, 1500, G = 192, b = 30, \alpha = 125, \beta = 46$



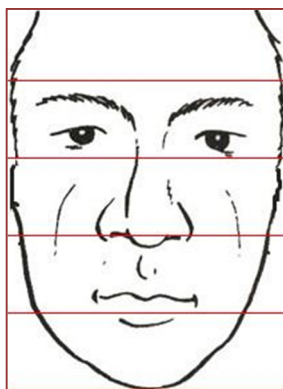
รูปที่ 3.17 ตรวจจับใบหน้าหลังจากการทำ MSRCR

- ก) รูปการตรวจจับใบหน้าโดยการคัดแยกสีผิวที่ไม่มีการปรับแสงสว่าง
 ข) รูปการตรวจจับใบหน้าโดยการคัดแยกสีผิวที่มีการปรับแสงสว่างด้วยเทคนิค MSRCR

นำภาพหลังการทำปรับปรุงแสงสว่างด้วยเทคนิค MSRCR มาตรวจจับใบหน้าโดยใช้เทคนิคการแยกสีผิวจะได้ดังแสดงในรูปที่ 3.17 โดยรูป ก) เป็นรูปการตรวจจับใบหน้าโดยการคัดแยกสีผิวที่ไม่มีการปรับแสงสว่าง และรูป ข) เป็นรูปการตรวจจับใบหน้าโดยการคัดแยกสีผิวที่มีการปรับแสงสว่างด้วยเทคนิค MSRCR จะเห็นว่าเมื่อมีการปรับแสงสว่างด้วยเทคนิค MSRCR แล้วทำให้การตรวจจับใบหน้าโดยการคัดแยกสีผิวถูกต้องมากขึ้น

3.3 การออกแบบระบบตรวจจับดวงตา

การตรวจจับดวงตาเป็นส่วนที่สำคัญมากสำหรับการตรวจจับอาการง่วงนอนโดยสามารถคาดเดาอาการง่วงนอนได้จากสถานะของดวงตา เช่น การหลับตาเป็นเวลานาน การกระพริบตาที่ช้าลง การหรี่ตามุมมองการมองของตา เป็นต้น การตรวจจับดวงตาโดยใช้การประมวลผลภาพดิจิทัลจะทำหลังจากการตรวจจับใบหน้าเพื่อลดพื้นที่ในค้นหาดวงตาลงทำให้เวลาในการค้นหาลดลงด้วยเช่นกัน หลังจากการตรวจจับใบหน้าจะได้ดังแสดงในรูปที่ 3.18 เมื่อแบ่งความสูงของใบหน้าเป็น 5 ส่วนจะเห็นว่าบริเวณที่มีดวงตาจะอยู่บนเส้นที่ 3 นับจากเส้นด้านล่างจึงสามารถตัดเอาเฉพาะบริเวณนี้มาประมวลผลในการค้นหาดวงตาต่อไปดังแสดงในรูปที่ 3.19



รูปที่ 3.18 แบ่งความสูงของใบหน้าเป็น 5 ส่วน



รูปที่ 3.19 บริเวณใบหน้า 2 ส่วน

นำบริเวณใบหน้า 2 ส่วนที่เหลือในรูปที่ 3.20 มาประมวลผลภาพดิจิทัลด้วยการหาขอบของวัตถุดังแสดงในรูปที่ 3.21 แล้วจากนั้นจะนำมาขยายขอบของวัตถุดังแสดงในรูปที่ 3.22 เพื่อให้แยกบริเวณของตาข่ายยิ่งขึ้นโดยนำภาพมากำจัดส่วนอื่น ๆ ที่ไม่ใช่บริเวณดวงตาเนื่องจากบริเวณของดวงตาจะมีขอบของภาพอยู่มากกว่าที่อื่น



รูปที่ 3.20 บริเวณใบหน้า 2 ส่วน



รูปที่ 3.21 หาขอบของวัตถุของบริเวณใบหน้า 2 ส่วน



รูปที่ 3.22 ขยายขอบของวัตถุของบริเวณใบหน้า 2 ส่วน



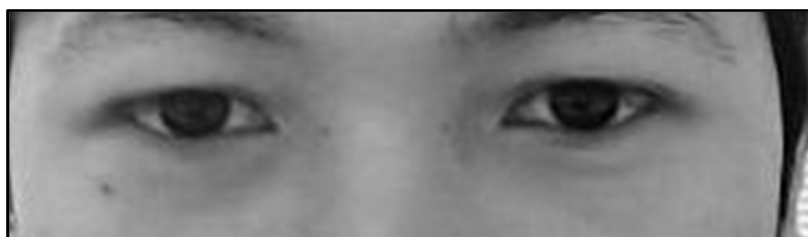
รูปที่ 3.23 ลักษณะของดวงตา

เมื่อได้บริเวณของดวงตาดังแสดงในรูปที่ 3.23 ต่อมาจะทำการค้นหาตานางานวิจัยนี้ได้นำเสนอวิธีการค้นหาตานา 2 วิธี ได้แก่ ค้นหาตานาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน และค้นหาตานาด้วยวิธี Integral projection ซึ่งวิธีการแรกจะเป็นการตั้งค่าขีดเริ่มเปลี่ยน เนื่องจากม่านตามีสีเข้มกว่าบริเวณอื่นจึงสามารถแยกม่านตาออกจากบริเวณอื่นได้ และวิธีการที่สองเป็นการรวมค่าของพิกเซลในแนวแกนตั้งและแนวแกนนอน

3.3.1 การค้นหาดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน

การค้นหาตานาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน (Threshold) สามารถทำได้โดยแปลงภาพในปริภูมิสี RGB เป็นภาพระดับสีเทา ดังแสดงในรูปที่ 3.24 ต่อมาจะทำการตั้งค่าขีดเริ่มเปลี่ยน เพื่อคัดแยกบริเวณที่สีเข้มกว่าบริเวณอื่นดังสมการที่ (3.11) จะได้ผลดังแสดงในรูปที่ 3.25 ซึ่งเป็นภาพขาว-ดำ เมื่อคัดแยกจากการคำนวณหาจำนวนพิกเซลในแต่ละวัตถุ ความกว้าง ความสูง และระยะห่างระหว่างตาทั้งสองข้างดังแสดงในรูปที่ 3.26 จะได้ตำแหน่งของม่านตาดังแสดงในรูปที่ 3.27

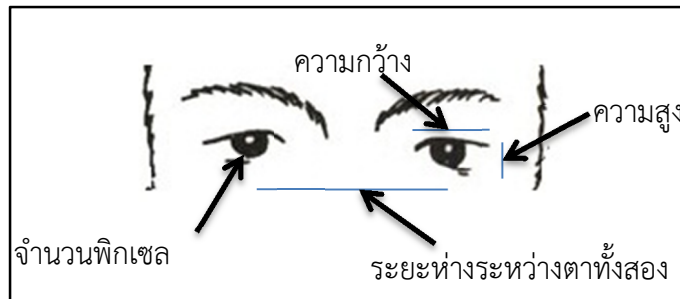
$$Binary\ image = \begin{cases} 1 & ; if\ grayscale \geq threshold \\ 0 & ; else \end{cases} \quad (3.11)$$



รูปที่ 3.24 ภาพระดับสีเทา



รูปที่ 3.25 ภาพหลังการตั้งค่าขีดเริ่มเปลี่ยน



รูปที่ 3.26 ลักษณะเฉพาะของตา



รูปที่ 3.27 การตรวจจับดวงตา

3.3.2 การค้นหาดวงตาด้วยวิธีการ Integral projection

Integral projection [9] เป็นการค้นหาค่าตำแหน่งของม่านตาโดยการหาผลรวมของค่าของพิกเซลของภาพระดับสีเทาในแนวแกนตั้งและแนวนอน เนื่องจากบริเวณของม่านตามีสีเข้มกว่าบริเวณอื่นจึงสามารถแยกตำแหน่งของม่านตาออกจากบริเวณอื่นได้

Horizontal integral projection เป็นการหาผลรวมของค่าพิกเซลในแนวแนวนอนดังสมการที่ (3.12) เมื่อนำรูปที่ 3.28 มาคำนวณหา Horizontal integral projection จะได้กราฟดังแสดงในรูปที่ 3.29

$$P_H(j) = \sum_{i=1}^m y_{(i,j)} \quad (3.12)$$

เมื่อ P_H คือ Horizontal integral projection

y คือ ภาพที่นำมาประมวลผล

- i คือ พิกัดของแถว
 j คือ พิกัดของหลัก
 m คือ จำนวนของแถว

Vertical integral projection เป็นการหาผลรวมของค่าพิกเซลในแนวแกนตั้งดังสมการที่ (3.13) เมื่อนำรูปที่ 3.28 มาคำนวณหา Vertical integral projection จะได้กราฟแสดงดังแสดงในรูปที่ 3.30

$$P_V(i) = \sum_{j=1}^n y_{(i,j)} \quad (3.13)$$

- เมื่อ P_V คือ Vertical integral projection
 y คือ ภาพที่นำมาประมวลผล
 i คือ พิกัดของแถว
 j คือ พิกัดของหลัก
 n คือ จำนวนของหลัก



ก)

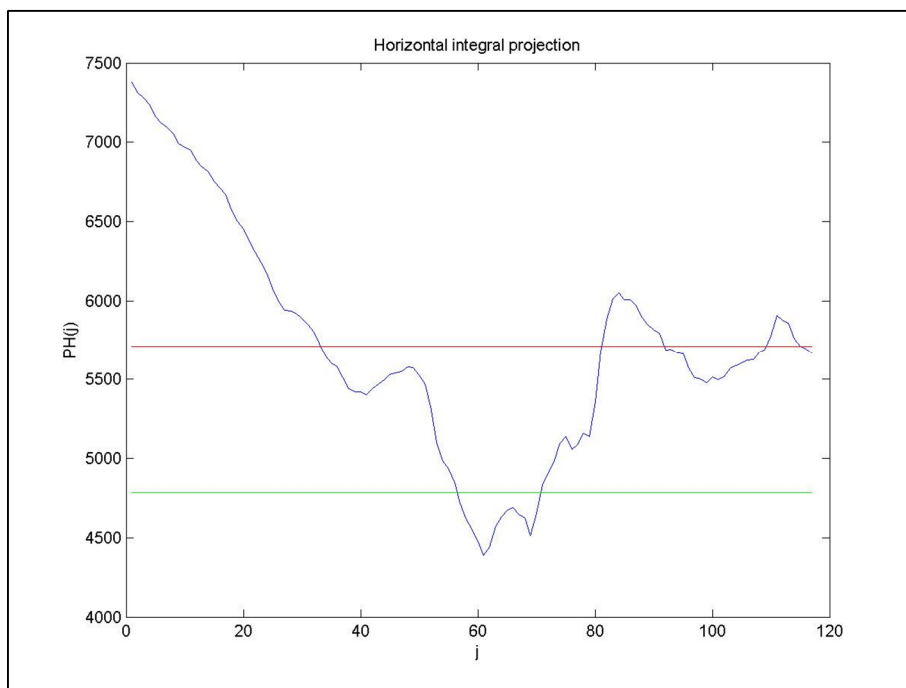


ข)

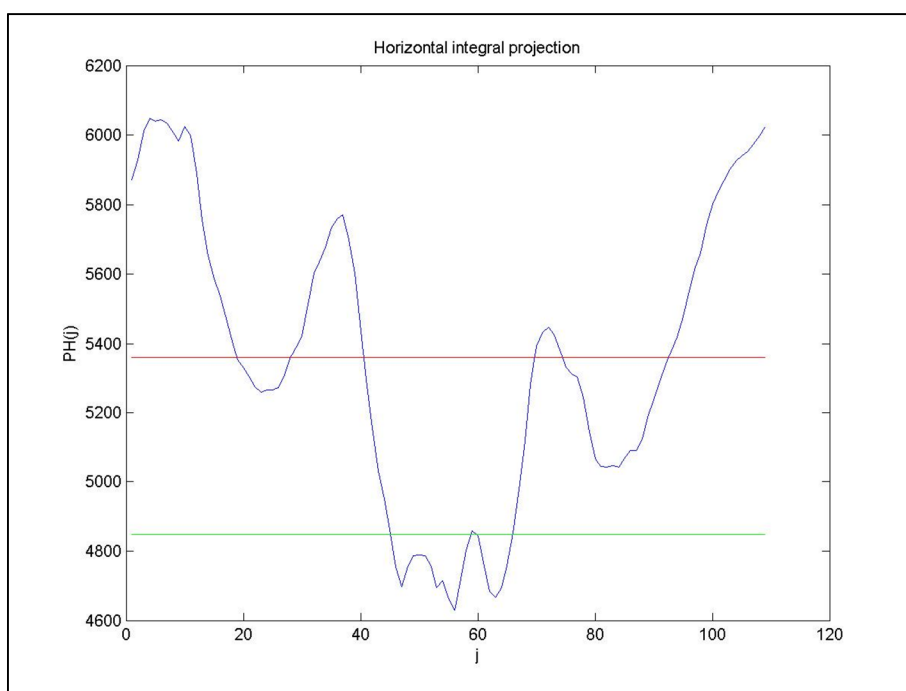
รูปที่ 3.28 บริเวณดวงตาที่เป็นภาพระดับสีเทา

ก) ตาขวา

ข) ตาซ้าย



ก)

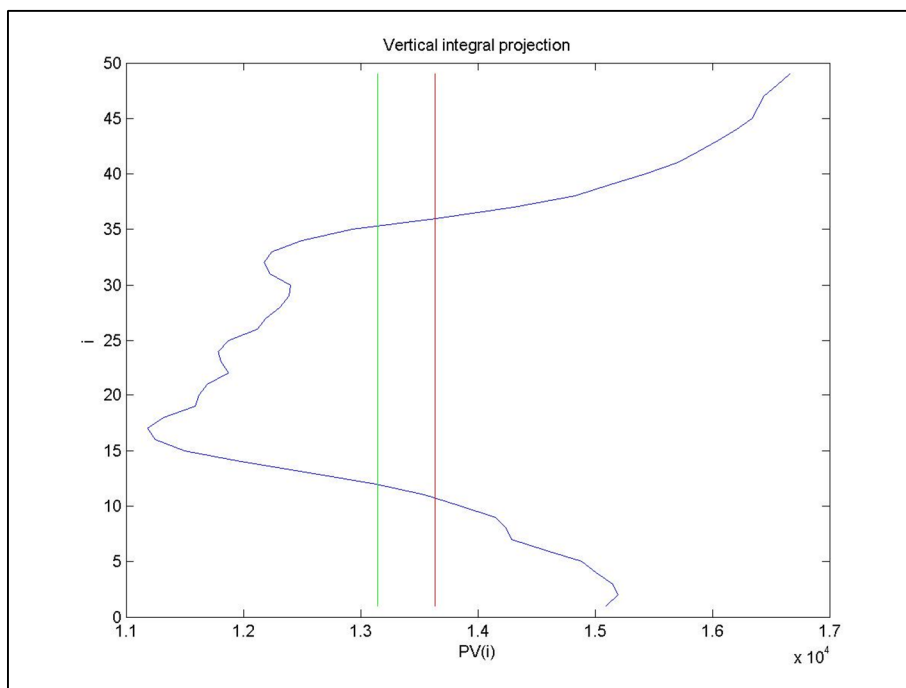


ข)

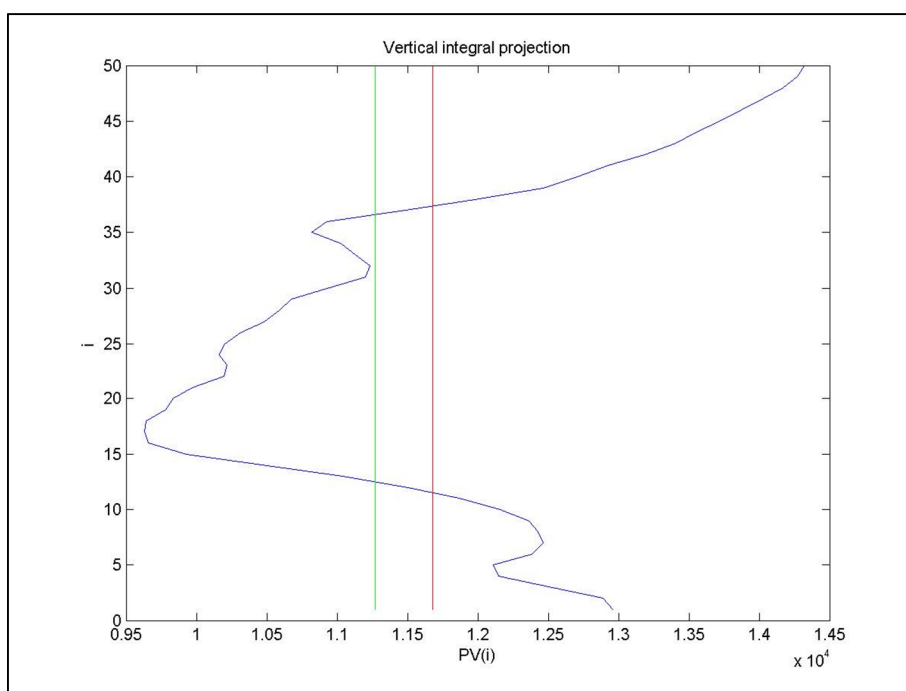
รูปที่ 3.29 กราฟ Horizontal integral projection

ก) กราฟ Horizontal integral projection ตาขวา

ข) กราฟ Horizontal integral projection ตาซ้าย



ก)



ข)

รูปที่ 3.30 กราฟ Vertical integral projection

ก) กราฟ Vertical integral projection ดาวหมา

ข) กราฟ Vertical integral projection ดาวซั้ว

รูปที่ 3.29 และรูปที่ 3.30 เป็นกราฟ Horizontal integral projection และ Vertical integral projection ตามลำดับ โดยเส้นกราฟสีแดงคือค่าเฉลี่ยของ P_H และ P_V สามารถคำนวณได้จากสมการที่ (3.14) และ (3.15) และเส้นกราฟสีเขียวคือค่าขีดเริ่มเปลี่ยนของม่านตาสามารถคำนวณได้จากสมการที่ (3.16) และ (3.17) เมื่อเลือกเฉพาะค่า P_H และ P_V ที่มีค่าต่ำกว่าค่าขีดเริ่มเปลี่ยน หรือเส้นกราฟสีเขียวจะได้บริเวณของม่านตาดังแสดงในรูปที่ 3.31

$$P_{H,arg} = \frac{\sum P_H}{m} \quad (3.14)$$

$$P_{V(arg)} = \frac{\sum P_V}{n} \quad (3.15)$$

เมื่อ $P_{H,arg}$	คือ ค่าเฉลี่ยของ P_H
m	คือ จำนวนของแถว
$P_{V,arg}$	คือ ค่าเฉลี่ยของ P_V
n	คือ จำนวนของหลัก

$$P_{H,Threshold} = P_{H,arg} - 0.7(P_{H,arg} - P_{H,min}) \quad (3.16)$$

$$P_{V,Threshold} = P_{V,arg} - 0.2(P_{V,arg} - P_{V,min}) \quad (3.17)$$

เมื่อ $P_{H,Threshold}$	คือ ค่าขีดเริ่มเปลี่ยนของม่านตาในแนวนอน
$P_{V,Threshold}$	คือ ค่าขีดเริ่มเปลี่ยนของม่านตาในแนวตั้ง
$P_{H,arg}$	คือ ค่าเฉลี่ยของ P_H
$P_{V,arg}$	คือ ค่าเฉลี่ยของ P_V
$P_{H,min}$	คือ ค่าต่ำสุดของ P_H
$P_{V,min}$	คือ ค่าต่ำสุดของ P_V



ก)

ข)

รูปที่ 3.31 ภาพม่านตาที่ได้จากวิธีการ Integral projection

ก) ตาขวา

ข) ตาซ้าย

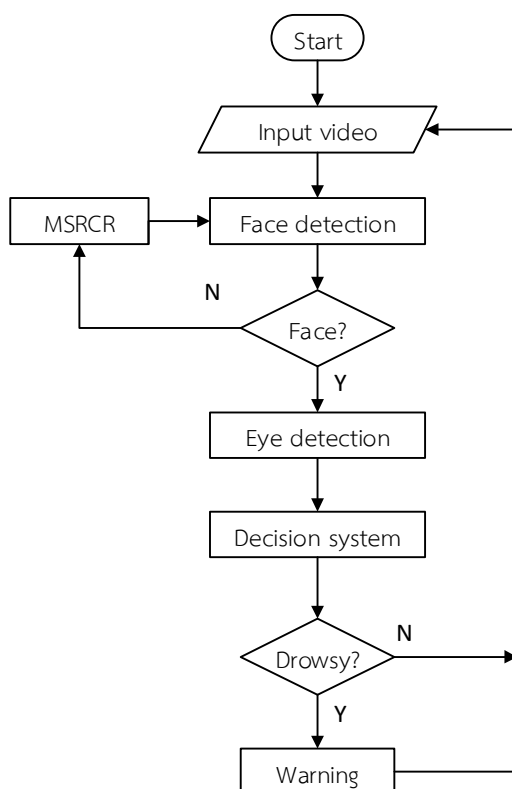
3.4 การออกแบบระบบตรวจจับอาการง่วงนอน

การตรวจจับอาการง่วงนอนที่ไม่รบกวนผู้ขับขี่วิธีหนึ่งคือ การตรวจจับอาการง่วงนอนด้วยการประมวลผลภาพดิจิทัลจากกล้องวิดีโอซึ่งวิธีการนี้เป็นการนำกล้องวิดีโอจับภาพใบหน้าของผู้ขับขี่ขณะขับขี่จากนั้นทำการประมวลผลด้วยอุปกรณ์เคลื่อนที่ และแจ้งเตือนด้วยเสียงเมื่อมีอาการง่วงนอนเกิดขึ้น ดังแสดงระบบตรวจจับอาการง่วงนอนในรูปที่ 3.32 เพื่อสังเกตลักษณะของใบหน้า ตรวจจับการเคลื่อนไหวของดวงตา และการเอียงของศีรษะ แล้วนำพารามิเตอร์อัตราการกระพริบตา มุมของการเอียงของศีรษะ มุมมองของคนขับ และอื่นๆ นำมาประมวลผลเพื่อคาดเดาถึงอาการง่วงนอนของผู้ขับขี่ซึ่งมีออกเป็น 3 ระดับ คือ ตื่นตัวหรือรู้สึกตัว เมื่อยล้า และง่วงนอน แล้วแจ้งเตือนผู้ขับขี่ให้ทราบถึงอาการง่วงนอนของตนเองโดยแจ้งเตือนเป็นข้อความบนหน้าจอของแท็บเล็ตพร้อมมีเสียงเตือนไปด้วยเมื่อผู้ขับขี่เริ่มมีอาการเมื่อยล้า หากผู้ขับขี่มีอาการเมื่อยล้าอยู่หรือเริ่มมีอาการง่วงนอนจะมีการแจ้งเตือนด้วยเสียงบ่อยครั้งมากขึ้นและเตือนให้หยุดการขับขี่ต่อไป



รูปที่ 3.32 การติดตั้งกล้องวิดีโอในยานพาหนะ

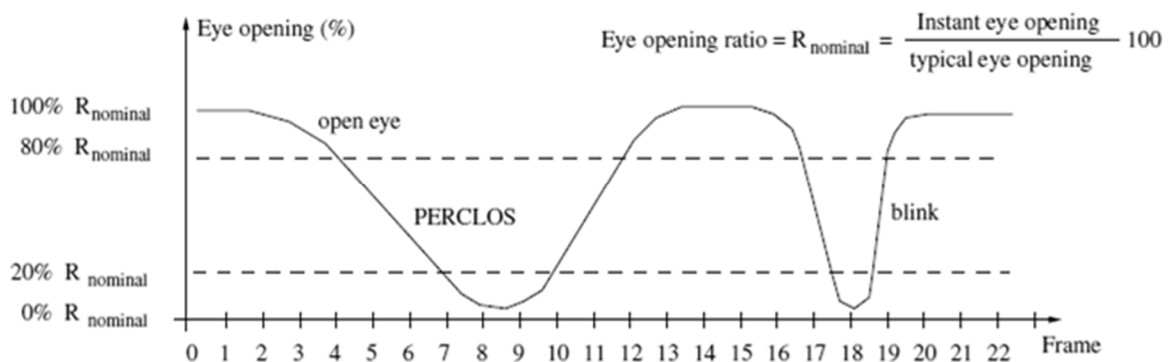
การตรวจจับอาการง่วงนอนด้วยการประมวลผลภาพดิจิทัลจากกล้องวิดีโอที่มีขั้นตอนดังแสดงในรูปที่ 3.33 ในขั้นแรกจะดึงเฟรมภาพจากกล้องวิดีโอซึ่งในงานวิจัยนี้จะติดตั้งกล้องเว็บแคมรุ่น Logitech HD Pro Webcam C920 ในรถยนต์ดังแสดงในรูปที่ 3.32 โดยเฟรมภาพที่ใช้มีขนาด 1920 x 1080 พิกเซลซึ่งมีความละเอียดสูงเพียงพอที่จะสามารถใช้ในการตรวจจับดวงตาของผู้ขับขี่รถยนต์ได้ จากนั้นจะนำเฟรมภาพมาตรวจจับใบหน้าโดยการคัดแยกสีผิว เนื่องจากวิธีการนี้สามารถตรวจจับได้ขณะหน้าของผู้ขับขี่เบี่ยงเล็กน้อยและสามารถตรวจจับได้กับใบหน้าของผู้ขับขี่รถทุกขนาด ส่วนการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมจะต้องใช้ภาพที่มาให้เครือข่ายประสาทเทียมรู้จำมากพอสมควรถึงจะสามารถตรวจจับใบหน้าที่มีอริยาบถแตกต่างจากใบหน้าตรงได้ ในขั้นตอนต่อมาจะตรวจสอบว่าในบริเวณที่ตรวจจับพบใบหน้าหรือไม่ ถ้าใช่จะทำการตรวจจับดวงตาต่อไป ถ้าไม่ใช่จะนำเฟรมภาพนี้มาปรับแสงในภาพด้วยวิธีการ MSRCR แล้วนำเฟรมภาพนี้กลับมาตรวจจับใบหน้าอีกครั้ง ในขั้นตอนต่อมาจะทำการตรวจจับดวงตาด้วยวิธีการ Integral projection เนื่องจากการตรวจจับดวงตาด้วยวิธีการ Integral projection นี้มีร้อยละการตรวจจับถูกต้องมากกว่าวิธีการตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยนเนื่องจากวิธีการ Integral projection สามารถตรวจจับดวงตาได้เมื่อภาพมีปริมาณของแสงสว่างที่เปลี่ยนแปลง เพื่อนำบริเวณของม่านตามาประมวลผลและตัดสินใจว่าผู้ขับขี่มีอาการง่วงนอนหรือไม่ ถ้าพบว่ามีอาการง่วงนอนระบบจะแจ้งเตือนให้ผู้ขับขี่ทราบเพื่อให้ผู้ขับขี่ตื่นหรือเปลี่ยนอริยาบถในการขับขี่ ในงานวิจัยนี้ได้ทำการจำลองการทำงานของระบบดังกล่าวด้วยโปรแกรม MATLAB โดยใช้เวลาในการประมวลผลประมาณ 11 วินาทีต่อเฟรม ส่วนในการประยุกต์เพื่อใช้งานจริงนั้น จะต้องทำการพัฒนาโปรแกรมให้สามารถทำงานได้แบบเวลาจริง



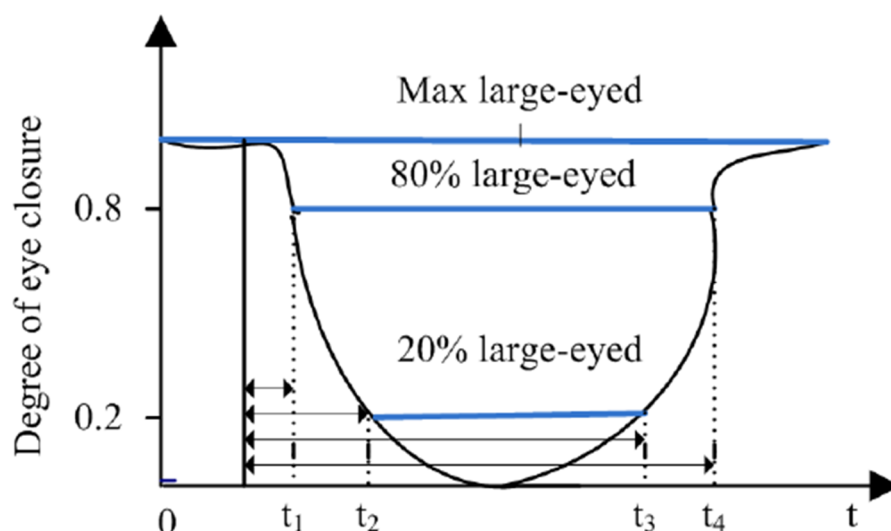
รูปที่ 3.33 ขั้นตอนการตรวจจับอาการง่วงนอน

ระบบตัดสินใจที่ใช้ในการตรวจจ้งอาการง่วงนอนจะใช้แนวทางการตรวจหาสถานะการปิด-เปิดของดวงตา โดยทำการคำนวณหาค่าร้อยละการปิดของดวงตา (Percentage of eye closure over time: PERCLOS) โดยปกติจะใช้เวลาในการกระพริบตา 0.3 วินาที [5] กล้องเว็บแคมรุ่น Logitech HD Pro Webcam C920 เฟรมภาพขนาด 1920 x 1080 พิกเซล มีอัตราเฟรม 30 เฟรมต่อวินาที แสดงว่าการกระพริบตา 1 ครั้งจะใช้ 7-9 เฟรมภาพ

PERCLOS [25] คือร้อยละของการปิดตาต่ำกว่า 20 เปอร์เซ็นต์ต่อการปิดตาระหว่าง 20-80 เปอร์เซ็นต์ของการเปิดตาที่สูงสุด จากรูปที่ 3.34 มีทั้งหมด 22 เฟรมภาพโดยช่วงที่นำมาประมวลผลหา PERCLOS คือช่วงเฟรมภาพที่ 4-12 สามารถคำนวณ PERCLOS ได้จากสมการที่ (3.18) โดยที่ $t_2 - t_3$ คือช่วงเวลาที่ตาปิดต่ำกว่า 20 เปอร์เซ็นต์ และ $t_1 - t_4$ คือ ช่วงเวลาที่ตาปิดระหว่าง 20-80 เปอร์เซ็นต์



รูปที่ 3.34 เฟรมที่ใช้ในการคำนวณ PERCLOS [31]



รูปที่ 3.35 การกำหนดช่วงเวลาในการประมวลผล PERCLOS [25]

$$PERCLOS = \frac{t_3 - t_2}{t_4 - t_1} \times 100\% \quad (3.18)$$

เมื่อ *PERCLOS* คือ ร้อยละการปิดของดวงตา

$t_2 - t_3$ คือ ช่วงเวลาที่ตาปิดต่ำกว่า 20 เปอร์เซนต์

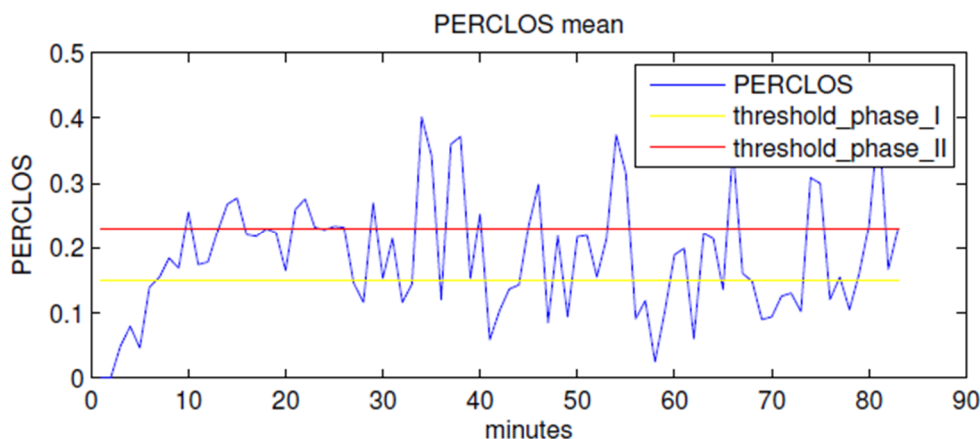
$t_1 - t_4$ คือ ช่วงเวลาที่ตาปิดระหว่าง 20-80 เปอร์เซนต์

ในงานวิจัยนี้ได้ใช้วิธีการคาดเดาอาการง่วงนอน 2 วิธีดังนี้

- วิธีการที่ 1 เป็นวิธีการที่นำเสนอในงานวิจัยโดย W. Qing et al. [30] ซึ่งมีการคาดการณ์ระดับของอาการง่วงนอนจากค่า PERCLOS โดยเมื่อมีค่า PERCLOS มากกว่า 0.4 หรือมีการปิดตานานเกิน 3 วินาทีจะสามารถคาดเดาได้ว่าผู้ทดสอบกำลังมีอาการง่วงนอนอยู่ในขณะนั้น ในการคำนวณหาค่า PERCLOS สามารถคำนวณได้จากสมการ (3.18)

- วิธีการที่ 2 ผู้วิจัยได้นำวิธีการคาดเดาอาการง่วงนอนที่นำเสนอโดย I. Garcia et al. [31] ซึ่งได้แบ่งระดับของอาการง่วงนอนออกเป็น 3 ระดับตามค่า PERCLOS ที่คำนวณได้ ดังแสดงในรูปที่ 3.36 ได้แก่

- ระดับที่ 1 ตื่นตัวหรือรู้สึกตัว (Wake) เป็นช่วงที่มี PERCLOS ต่ำกว่า 15 เปอร์เซนต์
- ระดับที่ 2 เมื่อยล้า (Fatigue) เป็นช่วงที่มี PERCLOS อยู่ระหว่าง 15-23 เปอร์เซนต์
- ระดับที่ 3 ง่วงซึม (Sleepiness) เป็นช่วงที่มี PERCLOS สูงกว่า 23 เปอร์เซนต์



รูปที่ 3.36 กราฟ PERCLOS

3.5 สรุป

การออกแบบระบบเฝ้าติดตามอาการง่วงนอนที่ได้กล่าวมาในบทนี้ โดยได้กล่าวถึงการออกแบบ การตรวจจับใบหน้า 2 วิธีการ วิธีการแรกคือการตรวจจับใบหน้าโดยใช้ใช้เครือข่ายประสาทเทียม ซึ่งเป็นการ นำฐานข้อมูลภาพใบหน้าและภาพที่ไม่ใช่ใบหน้าขนาด 200 x 200 พิกเซล อย่างละ 100 ภาพมาดัดแปลง ลักษณะเด่นด้วยวิธีการวิเคราะห์ส่วนประกอบสำคัญและการแปลงเวฟเล็ตแบบดีสครีทก่อนที่จะนำมาให้เครือข่ายเทียม รู้จำและได้ค่าน้ำหนักเพื่อนำไปใช้ตรวจจับใบหน้าต่อไป และวิธีการที่สองคือการตรวจจับใบหน้าโดยใช้การคัด แยกสีผิวเป็นการตรวจจับใบหน้าที่ใช้ช่วงสีผิวที่อยู่ในปริภูมิสี RGB และ HSV และยังได้นำเทคนิคการปรับปรุง แสงสว่างมาช่วยในการปรับแสงในภาพเพื่อให้การตรวจจับใบหน้าโดยการคัดแยกสีผิวมีประสิทธิภาพมากขึ้น ต่อมาได้กล่าวถึงการออกแบบการตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน และการตรวจจับดวงตาโดยใช้ เทคนิค Integral projection ซึ่งวิธีการแรกเป็นการค้นหาบริเวณของม่านตาจากลักษณะเฉพาะของดวงตา เนื่องจากความเข้มของม่านตาจะเข้มกว่าบริเวณอื่น และวิธีการที่สองเป็นการเป็นการค้นหาตำแหน่งของม่าน ตาโดยการหาผลรวมของค่าของพิกเซลของภาพระดับสีเทาในทั้งแกนตั้งและแกนนอน และในที่สุดท้ายเป็น การออกแบบการตรวจจับอาการง่วงนอนด้วยวิธีการคำนวณหาค่าร้อยละการปิดของดวงตาหรือ PERCLOS เพื่อเป็นแนวทางในการตรวจจับอาการง่วงนอนต่อไป

บทที่ 4

ผลการทดลอง

4.1 บทนำ

ในบทนี้จะนำเสนอผลการตรวจจับใบหน้าและดวงตาสำหรับระบบเฝ้าติดตามอาการง่วงนอน โดยได้แบ่งเป็นหัวข้อต่าง ๆ ประกอบด้วย ผลการการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมและการคัดแยกสีผิว ในส่วนของการคัดแยกสีผิวผู้วิจัยได้เพิ่มเทคนิค MSRCR เข้าไปด้วยเพื่อให้ผลการทดลองมีความถูกต้องมากยิ่งขึ้น จากนั้นจะนำเสนอเป็นผลการทดลองการตรวจจับดวงตา และส่วนสุดท้ายจะกล่าวถึงบทสรุปผลการทดลอง

4.2 ผลการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม

ในการนำเครือข่ายประสาทเทียมมาใช้ในการค้นหาใบหน้านั้นจะต้องมีการเตรียมฐานข้อมูลเพื่อใช้ในการรู้จำของเครือข่ายประสาทเทียม ในงานวิจัยนี้ได้ใช้ภาพใบหน้าและภาพที่ไม่ใช่ใบหน้า ขนาด 200×200 พิกเซล อย่างละ 100 ภาพ ดังตัวอย่างแสดงในรูปที่ 4.1 และรูปที่ 4.2 จากนั้นจะนำภาพดังกล่าวมาทำการปรับความสว่าง และนำมาดึงลักษณะเด่นและลดขนาดโดยการแปลงเวฟเล็ตแบบดีสครีทให้มีขนาด 50×50 พิกเซล หลังจากนั้นจะนำภาพหลังการแปลงเวฟเล็ตแบบดีสครีท มาดึงลักษณะเด่นของภาพโดยใช้การวิเคราะห์ส่วนประกอบสำคัญและนำไปให้เครือข่ายประสาทเทียมรู้จำ โดยได้นำเสนอแผนภาพการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมดังแสดงในรูปที่ 3.1

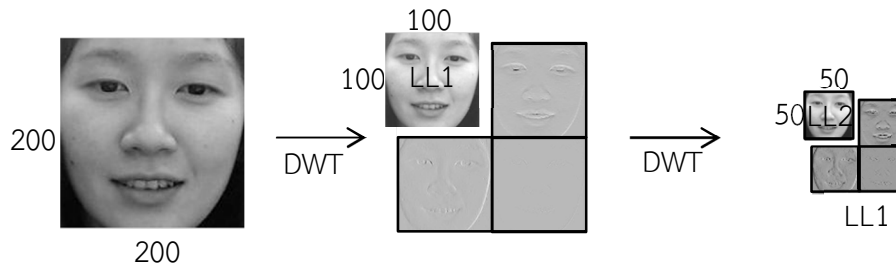


รูปที่ 4.1 ภาพใบหน้าขนาด 200 x 200 พิกเซล



รูปที่ 4.2 ภาพที่ไม่ใช่ใบหน้าขนาด 200 x 200 พิกเซล

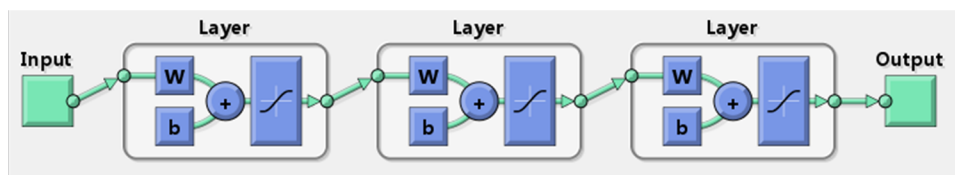
นำภาพใบหน้าขนาด 200 x 200 พิกเซลมาแปลงเวฟเล็ตแบบดิสครีทในระดับที่ 2 จะเหลือขนาดของภาพในแบนด์ย่อยความถี่ต่ำสุด (LL2) ที่จะนำไปประมวลผลต่อเท่ากับ 50 x 50 พิกเซล จะได้ดังแสดงในรูปที่ 4.3



รูปที่ 4.3 การแปลง DWT

จากนั้นนำภาพที่มีขนาด 50×50 พิกเซลมาดึงลักษณะเด่นโดยเลือกเวกเตอร์ไอเกน = 5 จะได้ข้อมูลขนาด 250×1 ต่อ 1 ภาพ เมื่อนำข้อมูลที่ได้หลังการวิเคราะห์ส่วนประกอบสำคัญเข้าสู่การรู้จำของเครือข่ายประสาทเทียมจะเหลือข้อมูลที่ใช้รู้จำเป็น 250×200 เมื่อนำมาทดสอบกับภาพที่เป็นใบหน้า 30 ภาพและภาพที่ไม่ใช่ใบหน้า 30 ภาพ เมื่อกำหนดค่าในการรู้จำดังนี้

1. Error goal = 10×10^{-8}
2. จำนวนเลเยอร์ = 3 เลเยอร์

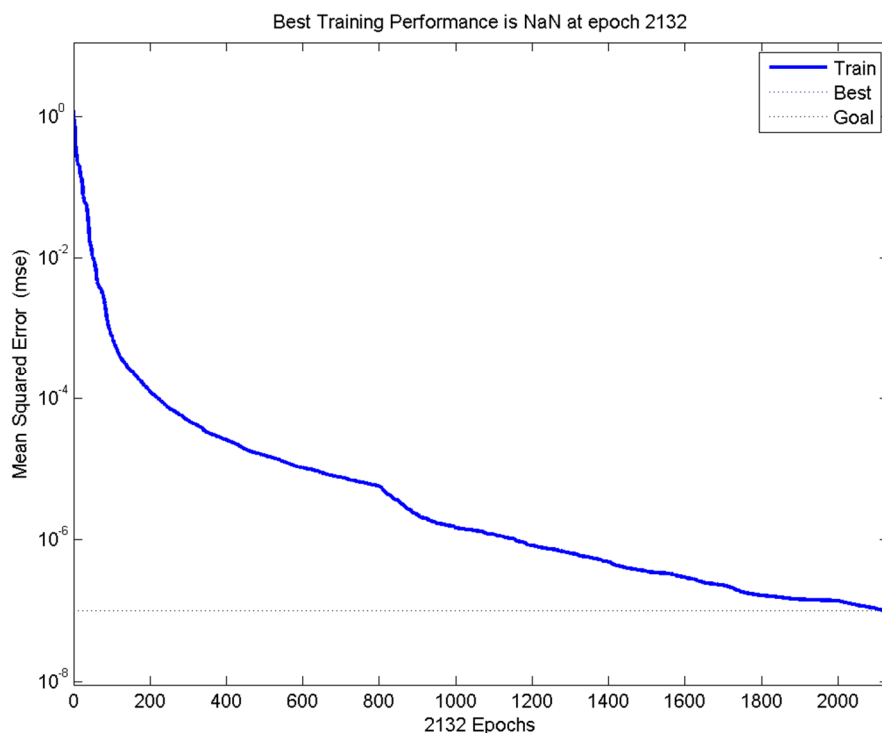


รูปที่ 4.4 เลเยอร์ของเครือข่ายประสาทเทียม

3. จำนวนรอบสูงสุด = 10000 รอบ

การทดลองครั้งที่ 1

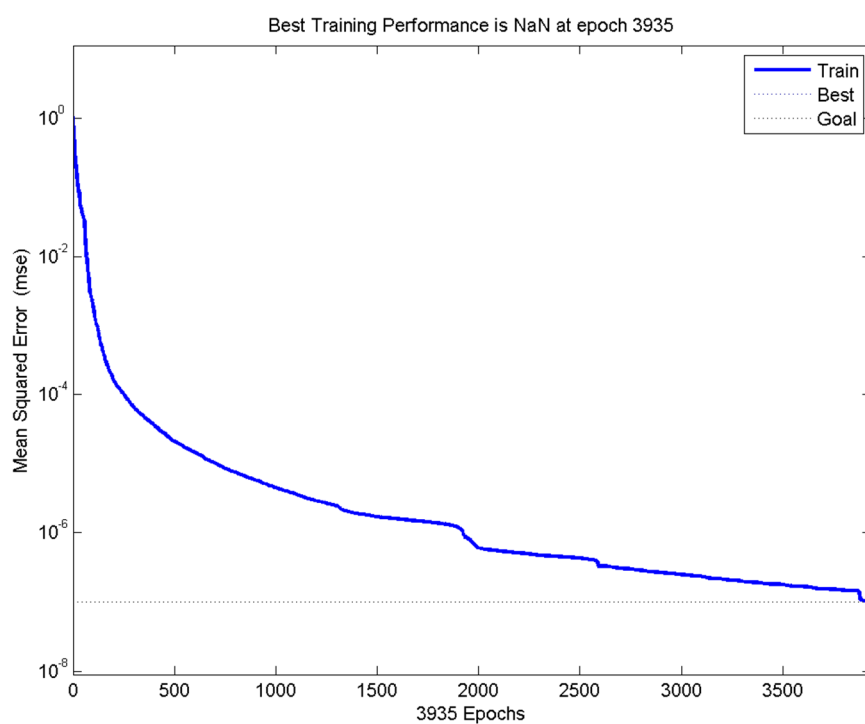
Eigenvectors = 5, Layer = 3-5-1('logsig' 'tansig' 'purelin') ใช้เวลาในการรู้จำ 107.229 วินาที จำนวน 2132 รอบ จะได้กราฟในการรู้จำดังแสดงในรูปที่ 4.5



รูปที่ 4.5 กราฟการรู้จำของเครือข่ายประสาทเทียมในการทดลองที่ 1

การทดลองครั้งที่ 2

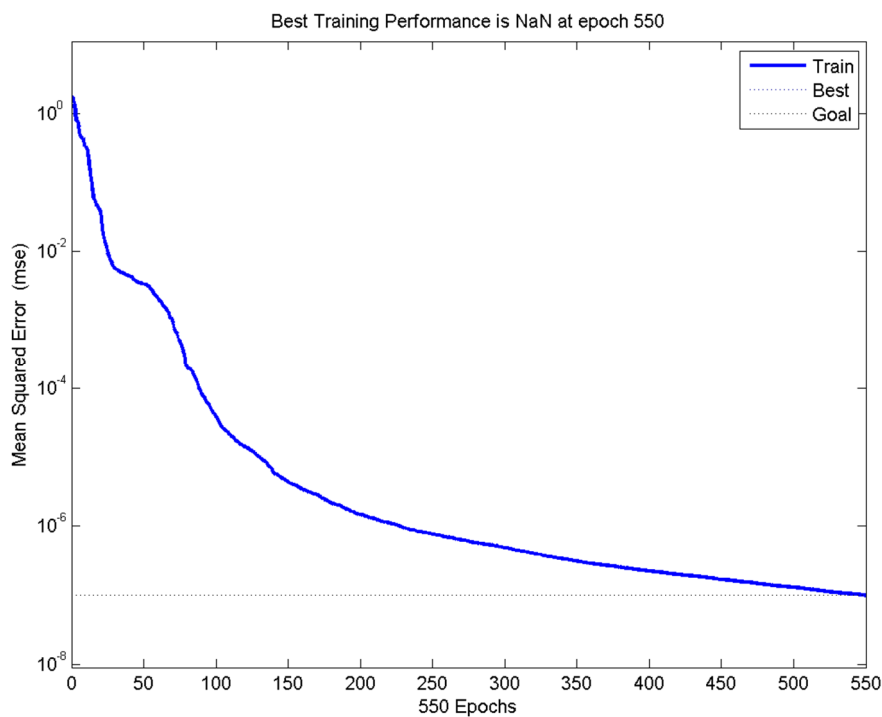
Eigenvectors = 5, Layer = 5-5-1('logsig' 'tansig' 'purelin') ใช้เวลาในการรู้จำ 98.574 วินาที จำนวน 3935 รอบ จะได้กราฟในการรู้จำดังแสดงในรูปที่ 4.6



รูปที่ 4.6 กราฟการรู้จำของเครือข่ายประสาทเทียมในการทดลองที่ 2

การทดลองครั้งที่ 3

Eigenvectors = 5, Layer = 5-5-1('tansig' 'tansig' 'tansig') ใช้เวลาในการรู้จำ 15.033 วินาที จำนวน 550 รอบ จะได้กราฟในการรู้จำดังแสดงในรูปที่ 4.7



รูปที่ 4.7 กราฟการรู้จำของเครือข่ายประสาทเทียมในการทดลองที่ 3

ตารางที่ 4.1: ผลการทดสอบการตรวจจับใบหน้าโดยใช้เทคนิคทางปัญญาประดิษฐ์

การทดลอง	ผลการตรวจจับที่ถูกต้อง		เปอร์เซ็นต์ความถูกต้องของการตรวจจับทั้งหมด (%)
	ภาพใบหน้า	ภาพที่ไม่ใช่ภาพใบหน้า	
ครั้งที่ 1	25	28	88.33
ครั้งที่ 2	22	28	83.33
ครั้งที่ 3	22	29	85.00
ค่าเฉลี่ย	23	28.33	85.55

ผลการตรวจจับใบหน้าโดยใช้เทคนิคเครือข่ายประสาทเทียมแสดงในตารางที่ 4.1 โดยทำการทดสอบภาพที่เป็นใบหน้า 30 ภาพและภาพที่ไม่ใช่ใบหน้า 30 ภาพจะใช้เวลาในการประมวลผล 0.075 วินาที ต่อภาพ พบว่ามีร้อยละการตรวจจับใบหน้าถูกต้องประมาณ 85.55 เปอร์เซ็นต์ ซึ่งอาจเกิดจากจำนวนภาพที่นำมาใช้ในการรู้จำของเครือข่ายประสาทเทียมยังไม่เพียงพอ การเพิ่มร้อยละความถูกต้องของการตรวจจับอาจทำได้โดยการเก็บรวบรวมข้อมูลใบหน้าที่ใช้ในการรู้จำให้มากขึ้น เนื่องจากผู้วิจัยเห็นว่าการตรวจจับใบหน้าโดย

ใช้เครือข่ายประสาทเทียมนั้นสามารถตรวจจับได้ตามลักษณะของใบหน้าที่มีการนำไปรู้จำเท่านั้นจึงไม่เหมาะที่จะนำมาตรวจจับใบหน้าของผู้ขับรถจึงได้ออกแบบการตรวจจับใบหน้าโดยการคัดแยกสีผิวในหัวข้อถัดไป

4.3 ผลการตรวจจับใบหน้าโดยการคัดแยกสีผิว

ในการตรวจจับใบหน้าโดยการคัดแยกสีผิวจะนำอัลกอริทึมการตรวจจับใบหน้าโดยการคัดแยกสีผิวที่ออกแบบให้หัวข้อ 3.23.2.2 มาพัฒนาปรับปรุงให้ดีขึ้นโดยใช้โปรแกรม Matlab และทดสอบกับภาพนิ่งที่ได้จากฐานข้อมูลจำนวน 5 ฐานข้อมูลดังนี้

- California Institute of Technology (CIT) [26] เป็นฐานข้อมูลใบหน้าทั้งหมด 450 ภาพจาก 27 คน มีขนาด 896×592 พิกเซล ซึ่งแต่ละภาพจะมีพื้นหลัง และแสงสว่างที่แตกต่างกัน
- Bao Face [27] เป็นฐานข้อมูลใบหน้าที่มีขนาดแตกต่างกันประกอบด้วยภาพใบหน้าเดี่ยว 149 ภาพ และภาพใบหน้ากลุ่ม 221 ภาพ แต่เลือกเฉพาะส่วนที่เป็นใบหน้าเดี่ยวมาใช้ทดสอบเพื่อให้เหมาะสมกับงานวิจัย
- Essex [28] เป็นฐานข้อมูลใบหน้าที่ได้จากคน 395 คน คนละ 20 ภาพ รวมทั้งหมด 7900 ภาพ
- Georgia Tech [29] เป็นฐานข้อมูลใบหน้าที่ได้จาก 50 คน คนละ 15 ภาพ แต่ละภาพจะมีขนาด 640×480 พิกเซล ซึ่งเป็นภาพที่มีแสงสว่างที่แตกต่างกัน
- Webcam เป็นฐานข้อมูลใบหน้าที่ได้จากกล้องเว็บแคมทั้งหมด 200 ภาพ จาก 10 คน มีขนาด 640×480 พิกเซล ซึ่งเป็นฐานข้อมูลที่ผู้วิจัยได้สร้างขึ้นเอง

ข้อมูลเชิงเทคนิคของกล้องเว็บแคม

- รุ่น Logitech HD Pro Webcam C920
- ความต้องการของระบบ Windows Vista[®], Windows[®] 7(32 บิต หรือ 64 บิต) หรือ Windows[®] 8
- โปรเซสเซอร์ 2.4 GHz Intel[®] Core 2 Duo
- หน่วยความจำ 2 GB ขึ้นไป
- พอร์ต USB 2.0 (USB 3.0)
- สามารถบันทึกภาพได้ในระดับ Full HD 1080p (สูงสุด 1920×1080 พิกเซล)

เมื่อนำไปทดสอบกับฐานข้อมูลจะได้ผลดังตารางที่ 4.2 ในการหาร้อยละของความถูกต้องของการตรวจจับใบหน้า สามารถหาได้จากนำภาพหลังการตรวจจับใบหน้ามาตีตารางจากนั้นจะนับช่องที่มีการ

ตรวจจับถูกต้องและผิดพลาดดังแสดงในรูปที่ 4.8 (ก) และ (ข) ซึ่งมีความถูกต้องเป็น 100 และ 84.5 เปอร์เซ็นต์ ตามลำดับ



(ก)



(ข)

รูปที่ 4.8 เปรียบเทียบความถูกต้อง

ก) ถูกต้อง 100 %

ข) ถูกต้อง 84.50 %

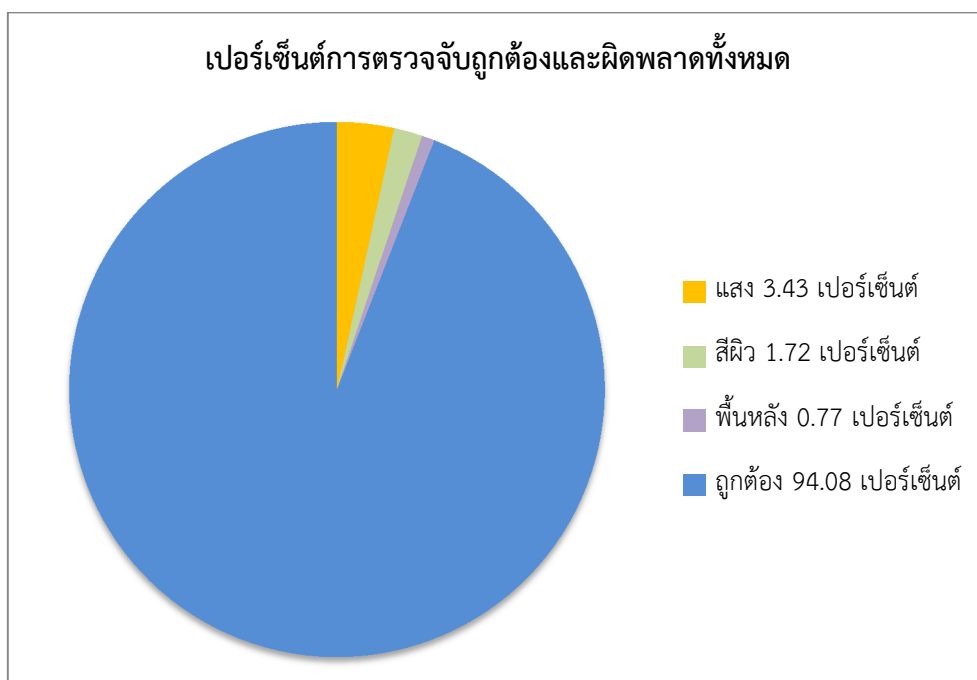
ตารางที่ 4.2: เปรียบเทียบร้อยละความถูกต้องระหว่างวิธีการของ P. P. Paul and M. Gavrilova กับวิธีการที่นำเสนอ

ฐานข้อมูลใบหน้า	วิธีการของ P. P. Paul and M. Gavrilova [12]	วิธีการตรวจจับใบหน้าที่นำเสนอ
1. CIT	91.2	95.04
2. BaoFace	92.5	92.68
3. Essex	93.7	93.79
4. Georgia Tech	85.2	94.82
5. Webcam	-	99.30
ค่าเฉลี่ย	90.65	95.13

ความผิดพลาดในการตรวจจับใบหน้าที่แสดงในรูปที่ 4.8 อาจมีสาเหตุมาจาก

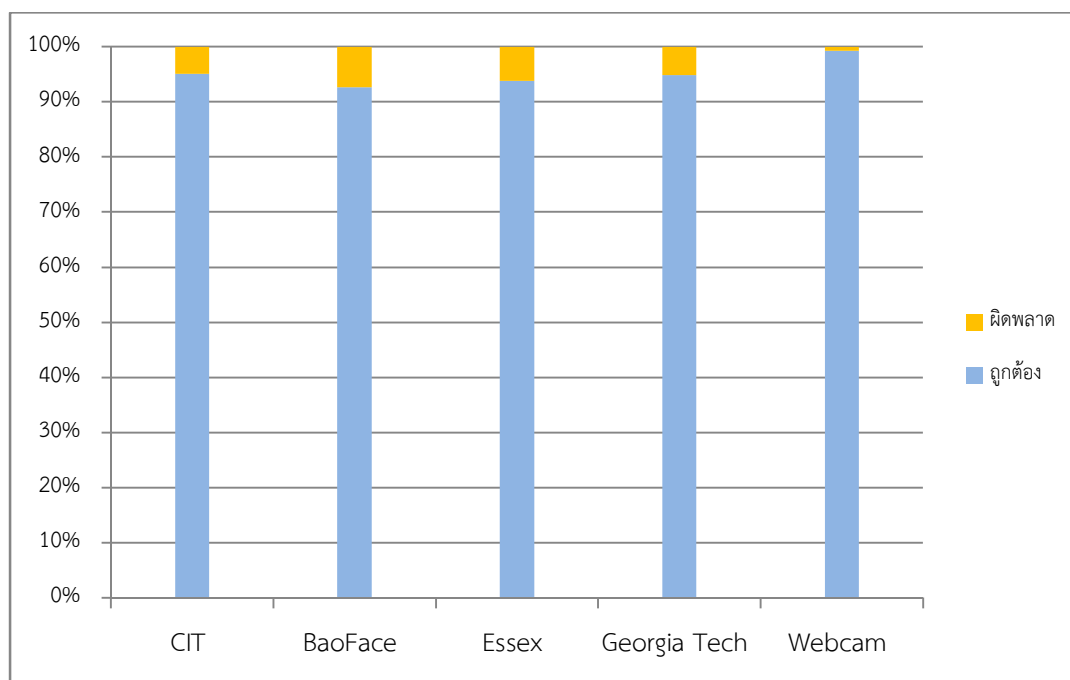
1. ภาพจากฐานข้อมูลเป็นภาพที่มีแสงสว่างไม่ปกติ ซึ่งทำให้การคัดแยกด้วยสีนั้นผิดพลาดเนื่องจากแสงสว่างมีผลต่อสีในภาพ ปัจจุบันนี้ทำให้การตรวจจับผิดพลาด 3.43 เปอร์เซ็นต์
2. ภาพจากฐานข้อมูลเป็นภาพใบหน้าของหลายเชื้อชาติ ซึ่งภาพตัวอย่างที่นำไปใช้หาขอบเขตของสีผิวโดยการคำนวณฮิสโตแกรมเป็นภาพใบหน้าของคนเอเชีย เมื่อนำไปทดสอบกับภาพใบหน้าของชาวสีอื่นจึงทำให้เกิดการตรวจจับผิดพลาด ปัจจุบันนี้ทำให้การตรวจจับผิดพลาด 1.72 เปอร์เซ็นต์

3. พื้นหลังของภาพในฐานข้อมูลมีสีที่คล้ายกับสีผิวจึงทำให้เกิดการตรวจจับผิดพลาดถึง 0.77 เปอร์เซ็นต์

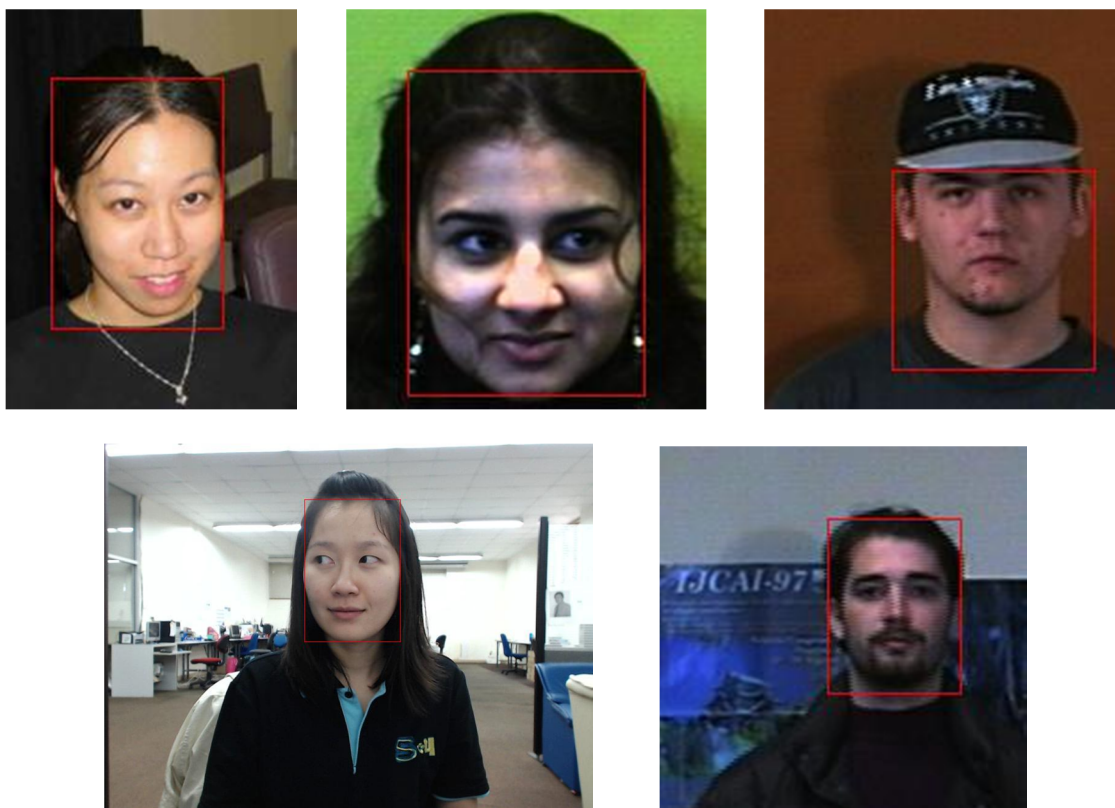


รูปที่ 4.9 แสดงเปอร์เซ็นต์การตรวจจับถูกต้องและผิดพลาด

จะเห็นได้ว่าแสงสว่างมีผลการต่ออัลกอริทึมการตรวจจับใบหน้ามากที่สุด ผู้วิจัยจึงได้สร้างฐานข้อมูลใบหน้าขึ้น โดยใช้กล้องที่มีการชดเชยแสงสว่าง เมื่อนำภาพในฐานข้อมูลนี้มาทดสอบจะได้ผลดังแสดงในรูปที่ 4.10 ซึ่งจะเห็นได้ว่าเมื่อมีการปัญหาของแสงสว่างในภาพโดยใช้กล้องที่มีการชดเชยแสง ทำให้การตรวจจับใบหน้ามีความถูกต้องมากขึ้นถึง 99.30 เปอร์เซ็นต์



รูปที่ 4.10 เปรียบเทียบร้อยละของการตรวจจับใบหน้าที่ถูกต้องและผิดพลาด



รูปที่ 4.11 ตัวอย่างการตรวจจับใบหน้าที่ถูกต้อง



รูปที่ 4.12 ตัวอย่างการตรวจจับใบหน้าที่ผิดพลาด

4.4 ผลการตรวจจับใบหน้าโดยการคัดแยกสีผิวร่วมกับเทคนิค MSRCR

ในการทดลองการตรวจจับใบหน้าโดยการคัดแยกสีผิวจะทำการติดตั้งกล้องเว็บแคมรุ่น Logitech HD Pro Webcam C920 ไว้ภายในยานพาหนะดังแสดงในรูปที่ 4.13 เพื่อเก็บภาพขณะขับรถ ภายใต้แสงสว่างที่เปลี่ยนแปลงตามธรรมชาติดังแสดงในรูปที่ 4.14



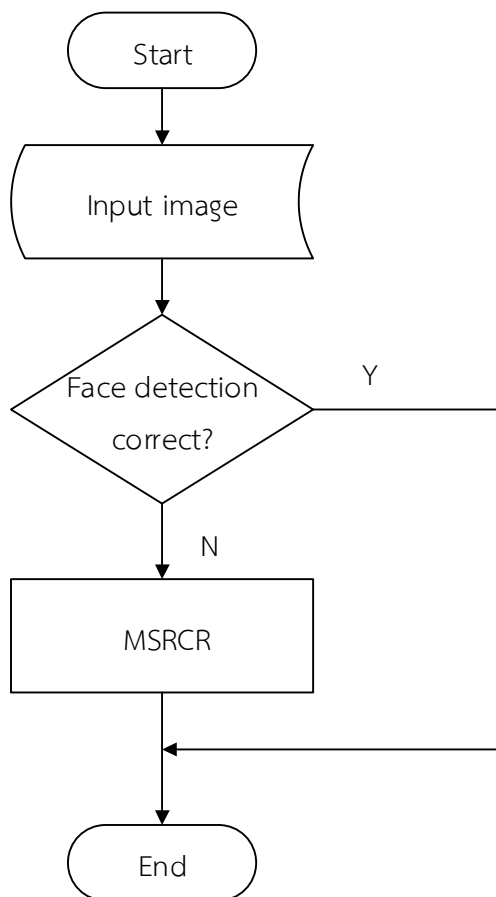
รูปที่ 4.13 การติดตั้งกล้องเว็บแคมภายในยานพาหนะ



รูปที่ 4.14 ภาพที่ใช้ในการทดสอบ

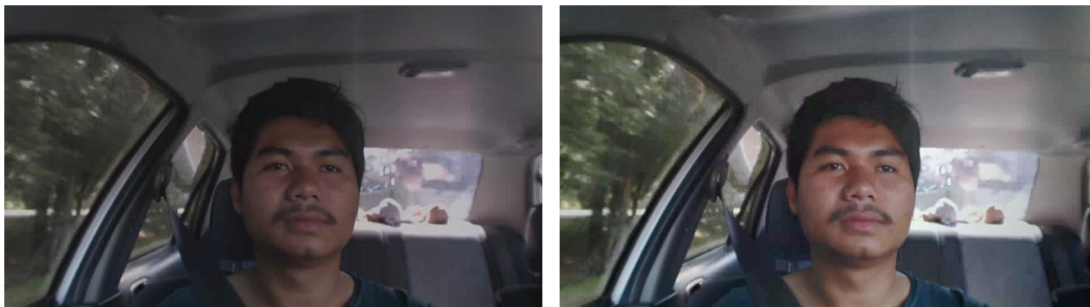
การทดสอบจะนำภาพขนาด 1920×1080 พิกเซลที่ได้จากกล้องเว็บแคมรุ่น Logitech HD Pro Webcam C920 มาทดสอบตรวจจับใบหน้าโดยใช้การคัดแยกสีผิวจะพบว่าทั้งผลการตรวจจับที่ถูกต้องและผลการจับที่ผิดพลาด จะนำภาพที่ตรวจจับผิดพลาดมาทำการปรับปรุงแสงสว่างภายในภาพด้วยเทคนิค MSRCR แล้วนำกลับมาตรวจจับใบหน้าโดยใช้การคัดแยกสีผิวดังแผนภาพดังแสดงในรูปที่ 4.15 ซึ่งเป็นแผนภาพการปรับปรุงแสงสว่างเมื่อโปรแกรมดึงเฟรมภาพจากวิดีโอทัศน์จะทำการตรวจจับใบหน้าโดยวิธีการคัดแยกสีผิว เมื่อพบว่ามีผลการตรวจจับผิดพลาดจะนำเฟรมภาพนั้นมาปรับปรุงแสงสว่างด้วยเทคนิค MSRCR จะได้ผลการปรับปรุงดังแสดงในรูปที่ 4.16 ซึ่งรูปที่ 4.16 ก) เป็นการปรับปรุงภาพที่มีแสงสว่างน้อยให้สว่างมากขึ้น และรูปที่ 4.16 ข) เป็นการปรับปรุงภาพที่มีแสงมากให้แสงสว่างลดน้อยลง โดยภาพซ้ายเป็นภาพดั้งเดิม

และภาพขาเป็นภาพหลังการปรับปรุงแสงสว่างโดยเทคนิค MSRCR

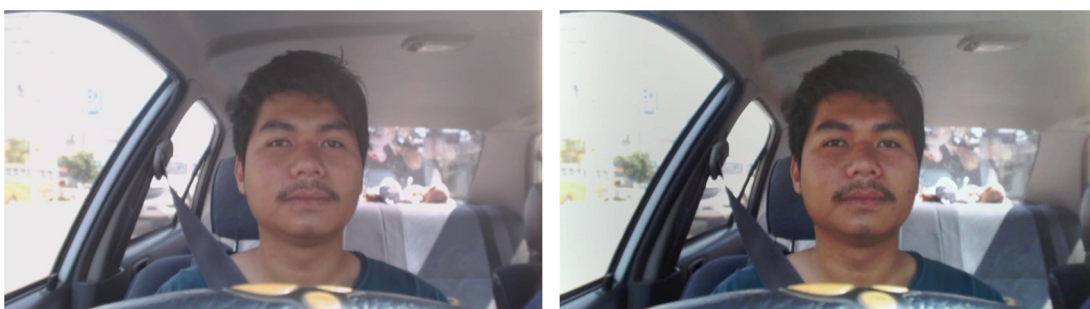


รูปที่ 4.15 แผนภาพการตรวจจับใบหน้ารวมกับเทคนิค MSRCR

นำเทคนิค MSRCR มาช่วยในการปรับปรุงแสงสว่างในภาพที่มีการตรวจจับใบหน้าโดยการคัดแยกสีผิวแล้วให้ผลการตรวจจับผิดพลาด โดยการทดสอบจะทดสอบกับภาพวิดีโอที่ได้ออกจากสภาพแสงสว่างจริงขณะขับรถ มีความยาว 24.11 นาที อัตราเฟรม 30 ภาพต่อวินาทีโดยจะดึงภาพมาประมวลผลทุก ๆ 10 เฟรมภาพ จะได้ภาพในการทดสอบทั้งหมด 4,341 ภาพ หลังจากนั้นนำภาพตรวจจับใบหน้าแล้วได้สุ่มเลือกภาพที่มีการตรวจจับใบหน้าผิดพลาดมา 500 ภาพเพื่อมาปรับปรุงแสงสว่างโดยใช้เทคนิค MSRCR แล้วนำไปตรวจจับใบหน้าโดยการคัดแยกสีผิวอีกครั้ง พบว่าเทคนิค MSRCR ช่วยปรับปรุงแสงสว่างในภาพให้มีผลการตรวจจับใบหน้าถึง 463 ภาพดังตัวอย่างแสดงรูปที่ 4.17 จะเห็นว่าการตรวจจับใบหน้าหลังการปรับแสงสว่างแล้วให้ความถูกต้องมากขึ้น และกราฟในรูปที่ 4.18 แสดงให้เห็นว่าการตรวจจับใบหน้าโดยการคัดแยกสีผิวรวมกับเทคนิค MSRCR จะสามารถปรับปรุงภาพจากเดิมที่มีร้อยละการตรวจจับถูกต้อง 77.45 เปอร์เซ็นต์ให้มีร้อยละการตรวจจับถูกต้อง 95.4 เปอร์เซ็นต์



ก)



ข)

รูปที่ 4.16 ตัวอย่างการปรับปรุงแสงสว่างด้วยเทคนิค MSRCR

- ก) ปรับปรุงภาพที่มีแสงสว่างน้อยให้สว่างมากขึ้น
- ข) ปรับปรุงภาพที่มีแสงมากให้แสงสว่างลดน้อยลง

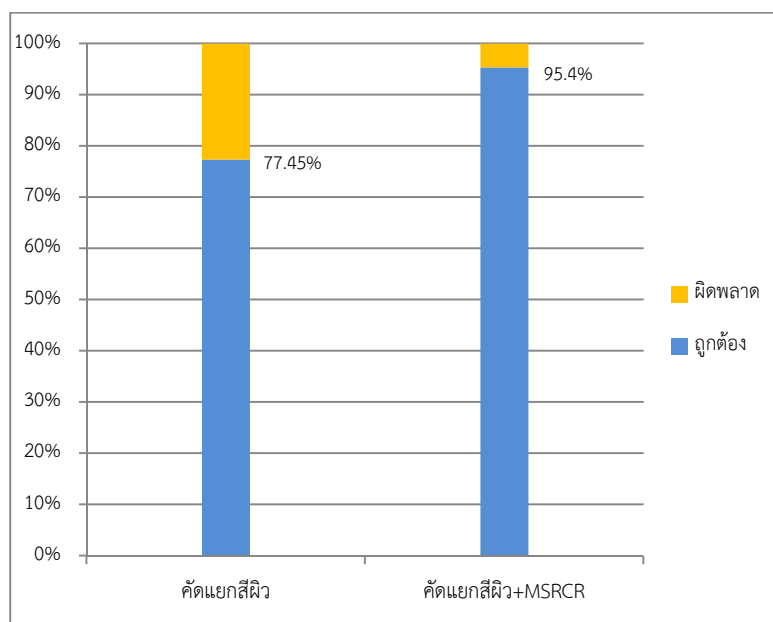


ก)

ข)

รูปที่ 4.17 เปรียบเทียบการตรวจจับใบหน้า

- ก) ภาพการตรวจจับใบหน้าที่ผิดพลาด
- ข) ภาพการตรวจจับใบหน้าหลังจากปรับปรุงแสงสว่างโดยใช้เทคนิค MSRCR



รูปที่ 4.18 เปรียบเทียบการตรวจจับใบหน้าด้วยวิธีการคัตแยกสีผิว

และการประยุกต์ใช้MSRCR ร่วมกับการคัตแยกสีผิว

4.5 ผลการตรวจจับดวงตาและการตรวจจับอาการง่วงนอน

4.5.1 ผลการตรวจจับดวงตา

ผลการตรวจจับดวงตาที่จะนำเสนอในหัวข้อนี้ ผู้วิจัยจะใช้ข้อมูลภาพที่ได้จากการติดตั้งกล้องเว็บแคม Logitech HD Pro Webcam C920 ทั้งหมด 4,617 ภาพแต่ละภาพมีขนาด 1920x1080 พิกเซล โดยเป็นการเก็บข้อมูลภาพจากการใช้งานจริง จากนั้นจะนำภาพที่ได้มาทำการประมวลผลในห้องปฏิบัติการวิจัยเพื่อทำการตรวจจับใบหน้าด้วยวิธีการคัตแยกสีผิว และทำการตรวจจับดวงตาด้วยวิธีการขีดเริ่มเปลี่ยน และวิธีการ Integral projection โดยมีผลการทดลองแสดงในตารางที่ 4.3

ตารางที่ 4.3: เปรียบเทียบร้อยละความถูกต้องระหว่างวิธีตั้งการ Threshold และวิธีการ Integral projection

วิธีการ	ตาขวา 4,617 ข้าง	ตาซ้าย 4,617 ข้าง	รวมทั้งสองตา 9,234 ข้าง	ร้อยละความถูกต้อง
Threshold	4,270	4,251	8,521	92.27
Integral projection	4,537	4523	9,060	98.11



รูปที่ 4.19 ภาพที่ใช้ในการทดสอบ

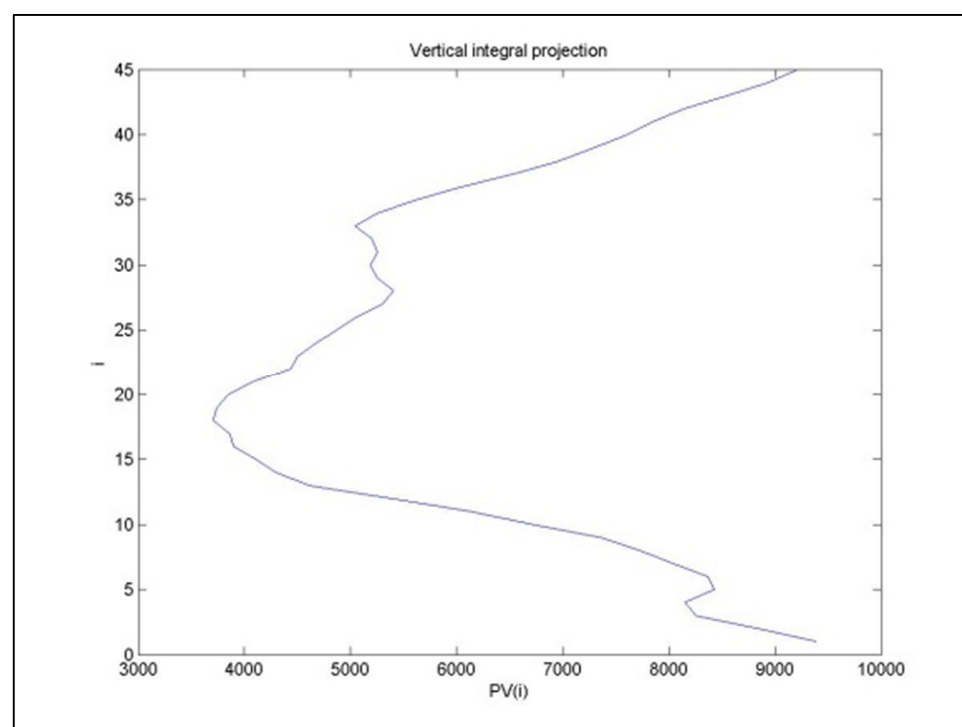
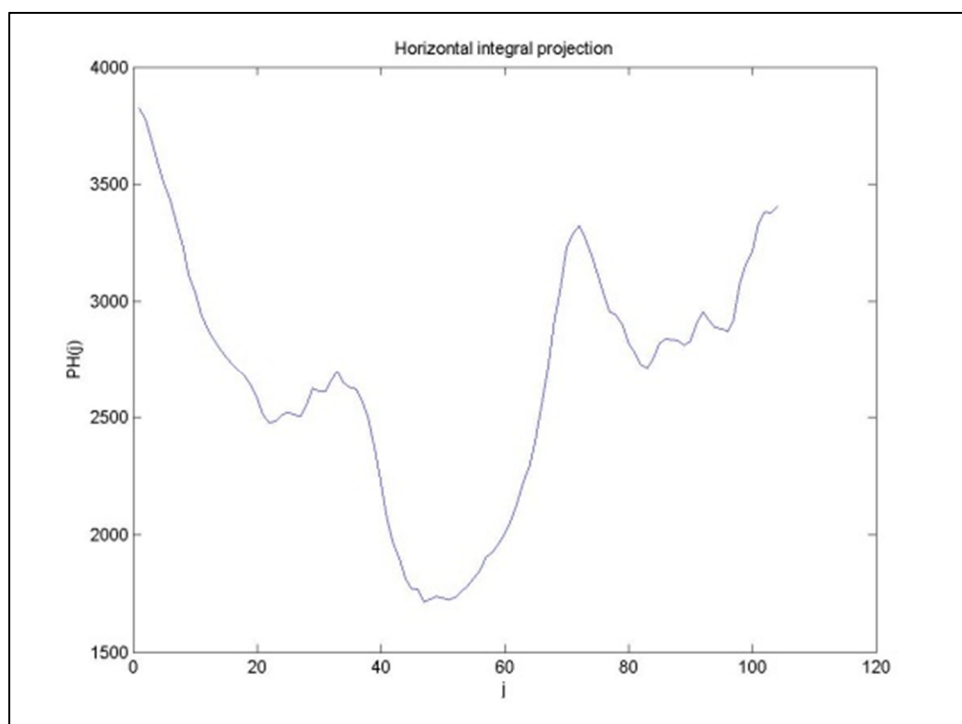
เมื่อนำภาพมาทดสอบการตรวจจับดวงตาด้วยวิธีการตั้ง Threshold และวิธีการ Integral projection ได้ผลการทดสอบดังตารางที่ 4.3

จากตารางที่ 4.3 จะเห็นว่าวิธีการ Integral projection มีร้อยละการตรวจจับถูกต้อง 98.11 เปอร์เซ็นต์ซึ่งมากกว่าวิธีการตั้ง Threshold ที่มีร้อยละการตรวจจับถูกต้อง 92.27 เปอร์เซ็นต์ โดยที่ใช้ฐานข้อมูลเดียวกันดังตารางที่ 4.4 ด้านซ้ายเป็นรูปจากการตรวจจับดวงตาด้วยวิธีการตั้ง Threshold และด้านขวาเป็นรูปจากการตรวจจับดวงตาด้วยวิธีการ Integral projection การตรวจจับดวงตาด้วยวิธีการตั้ง Threshold มีร้อยละความถูกต้องน้อยกว่าวิธีการ Integral projection อาจเกิดจากฐานข้อมูลที่ใช้มีแสงสว่างเปลี่ยนแปลงตลอดเวลา เนื่องจากภาพที่นำมาทดสอบได้จากสถานะใช้งานจริงขณะขับรถวิธีการตั้ง Threshold จึงไม่เหมาะสมกับการตรวจจับดวงตาในภาพที่มีแสงสว่างที่เปลี่ยนแปลงตลอดเวลา ส่วนวิธีการ Integral projection จะเหมาะสมกับการตรวจจับดวงตาในภาพที่มีแสงสว่างที่เปลี่ยนแปลงตลอดเวลา มากกว่า เนื่องจากวิธีการนี้เป็นการหาผลรวมของค่าความเข้มของพิกเซลในแนวแกนตั้งและแนวนอนไม่ว่าแสงสว่างจะมากหรือน้อยบริเวณของม่านตาจะมีผลรวมของค่าความเข้มต่ำกว่าบริเวณอื่นดังแสดงในรูปที่ 4.20 โดยรูป ก) เป็นกราฟของ Horizontal integral projection และ Vertical integral projection ของภาพที่มีแสงน้อย และ ข) เป็นกราฟของ Horizontal integral projection และ Vertical integral projection ของภาพที่มีแสงมาก (โดยกราฟของ Horizontal integral projection แกน X คือพิกเซลในแนวแกนตั้ง และแกน Y คือผลรวมค่าความเข้มของพิกเซลในแนวแกนตั้ง และกราฟของ Vertical integral projection แกน X คือ

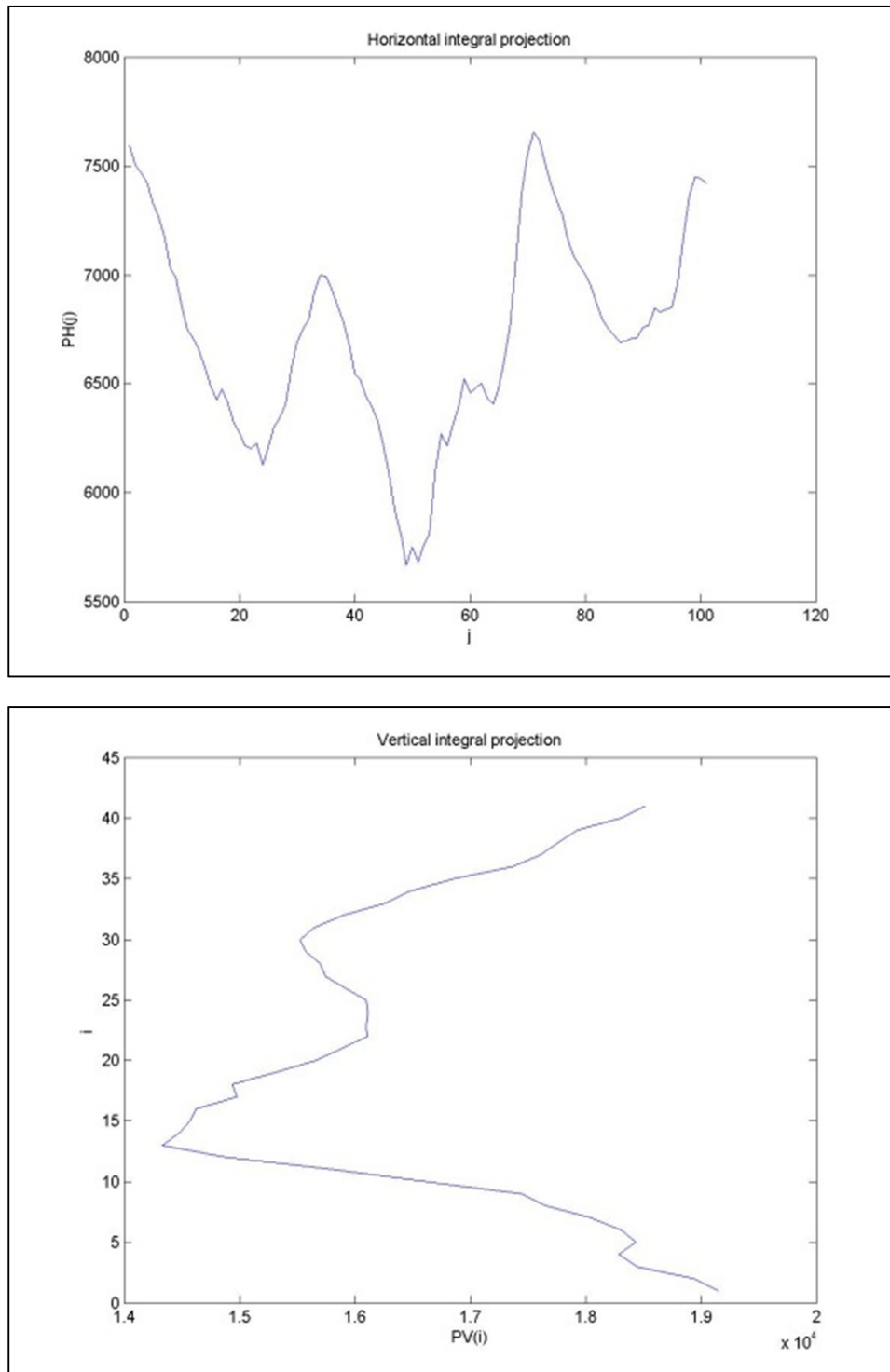
ผลรวมค่าความเข้มของพิกเซลในแนวแกนนอน และแกน Y คือ พิกเซลในแนวแกนนอน) สังเกตได้ว่าบริเวณที่มีผลรวมของค่าความเข้มต่ำกว่าบริเวณอื่นของทุก ๆ กราฟจะเป็นบริเวณของม่านตา

ตารางที่ 4.4: เปรียบเทียบการตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยนและ Integral projection

เฟรมที่	วิธีการตั้งค่าขีดเริ่มเปลี่ยน	วิธีการ Integral projection
198		
349		
1530		
2617		



n)



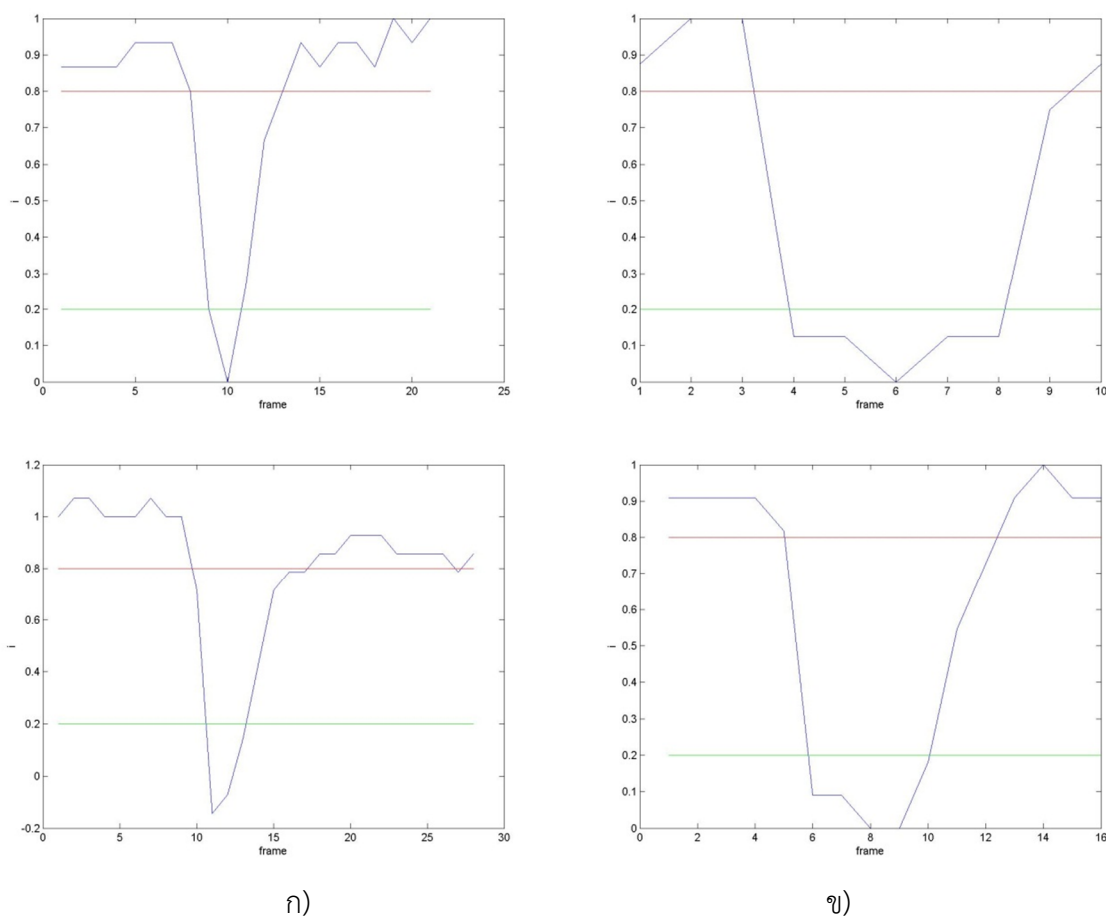
ข)

รูปที่ 4.20 เปรียบเทียบกราฟ Integral projection ของภาพดวงตา

- ก) กราฟของ Horizontal integral projection และ Vertical integral projection ของภาพที่มีแสงน้อย
- ข) กราฟของ Horizontal integral projection และ Vertical integral projection ของภาพที่มีแสงมาก

4.5.2 ผลการตรวจจับอาการง่วงนอน

ในการตรวจจับอาการง่วงนอนจะเริ่มจากการคำนวณค่า PERCLOS โดยสามารถคำนวณได้จาก กราฟเปอร์เซ็นต์ของการปิดตาในการกระพริบตาแต่ละครั้งดังแสดงในรูปที่ 4.21 รูป ก) เป็นกราฟเปอร์เซ็นต์ของการปิดตาเทียบกับเฟรมเวลาขณะตื่นตัว และ ข) เป็นกราฟเปอร์เซ็นต์ของการปิดตาเทียบกับเฟรมเวลาขณะง่วงนอน โดยกราฟเส้นสีแดงคือร้อยละการปิดตา 80 เปอร์เซ็นต์และกราฟเส้นที่เขียวคือร้อยละการปิดตา 20 เปอร์เซ็นต์ค่า PERCLOS คำนวณได้ดังสมการที่ (3.17) จากนั้นจะใช้วิธีการคาดเดาอาการง่วงนอน 2 วิธี ดังได้นำเสนอในหัวข้อที่ 3.4 เพื่อคาดเดา และแจ้งเตือนผู้ขับขี่รถยนต์ต่อไป

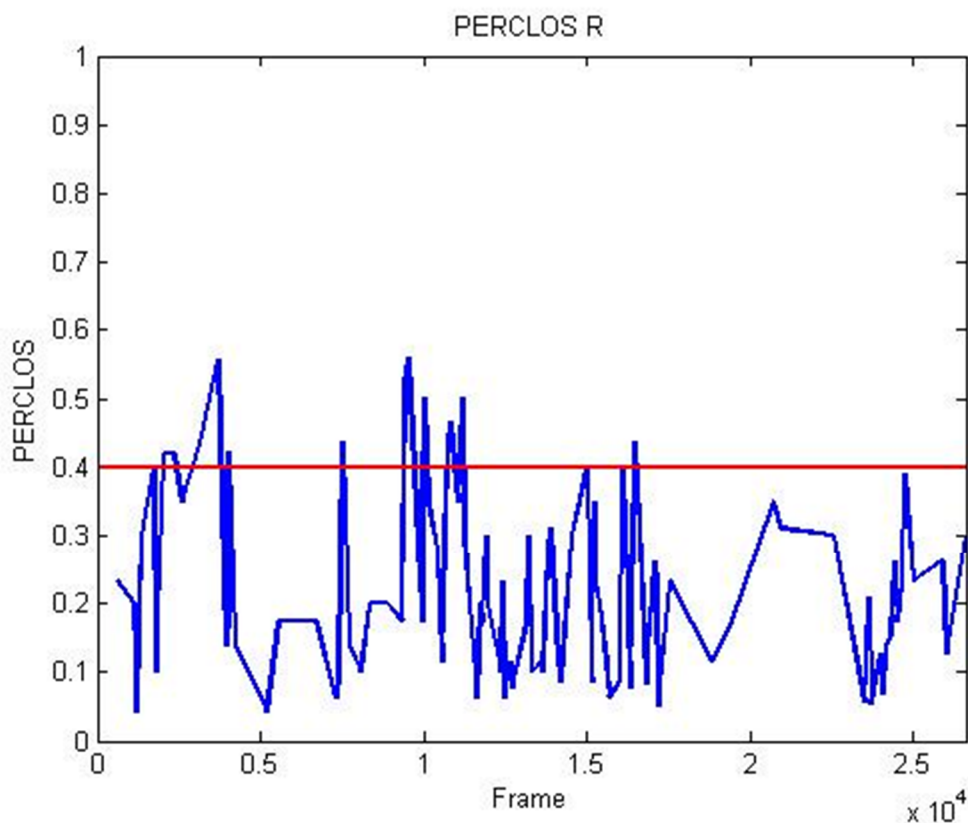


รูปที่ 4.21 กราฟเปอร์เซ็นต์ของการปิดตาเทียบกับเฟรมเวลา

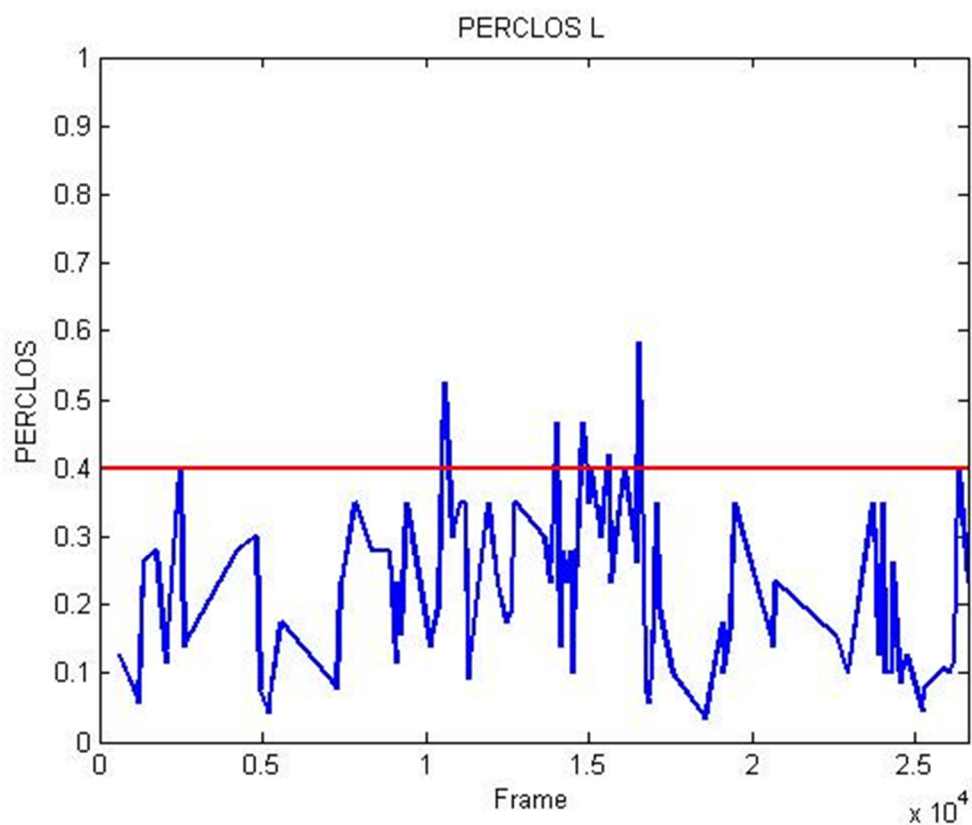
- ก) กราฟเปอร์เซ็นต์ของการปิดตาเทียบกับเฟรมเวลาขณะตื่นตัว
- ข) กราฟเปอร์เซ็นต์ของการปิดตาเทียบกับเฟรมเวลาขณะง่วงนอน

1) ผลการตรวจจับอาการง่วงนอนจากการคาดเดาอาการง่วงนอนวิธีการที่ 1

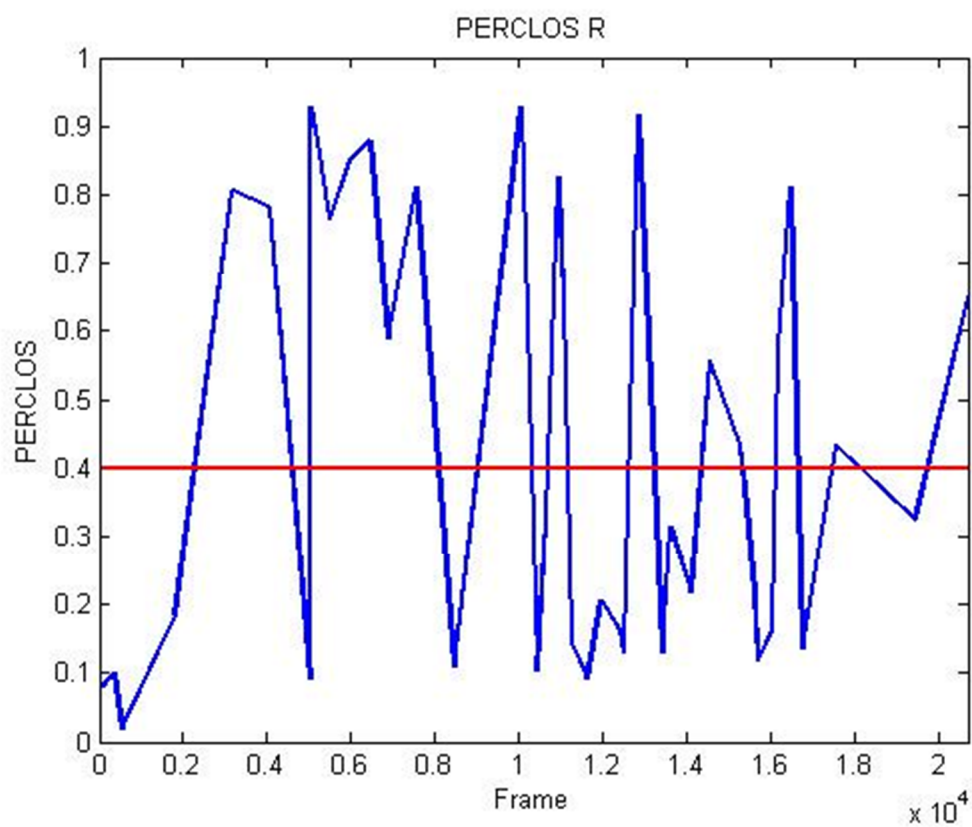
ในการคาดเดาอาการง่วงนอนวิธีการที่ 1 ผู้วิจัยได้นำวิธีการคำนวณ PERCLOS มาใช้คาดเดาอาการง่วงนอนของผู้ขับขี่รถยนต์ ตามรายละเอียดการออกแบบระบบตรวจจับอาการง่วงนอนที่นำเสนอไปแล้วในหัวข้อที่ 3.4 โดยเมื่อระบบพบว่ามีค่า PERCLOS มากกว่า 0.4 หรือมีการปิดตานานเกิน 3 วินาทีระบบจะคาดเดาได้ว่าผู้ทดสอบหรือผู้ขับขี่รถมีอาการง่วงนอน โดยผู้วิจัยได้ใช้ฐานข้อมูลจากเว็บแคม Logitech HD Pro Webcam C920 ที่มีเฟรมภาพขนาด 1920x1080 พิกเซลซึ่งประกอบด้วยวิดีโอที่ผู้ขับขี่ตื่นตัวหรือรู้สึกตัว (Wake) และที่ผู้ขับขี่เมื่อยล้า (Fatigue) และง่วงนอน (Sleepiness) แต่ละวิดีโอมีความยาว 20 นาทีมาใช้ในการทดสอบประสิทธิภาพของระบบ โดยมีผลการตรวจจับอาการง่วงนอนดังกราฟรูปที่ 4.22 ถึง รูปที่ 4.23 แสดงกราฟ PERCLOS ในขณะที่ผู้ขับขี่ตื่นตัวหรือรู้สึกตัวของชาวและตาซ้ายตามลำดับ โดยที่แกน X คือเฟรมภาพ และแกน Y คือค่า PERCLOS และรูปที่ 4.24 ถึง รูปที่ 4.25 แสดงกราฟ PERCLOS ในขณะที่ผู้ขับขี่เมื่อยล้าและง่วงนอนของชาวและตาซ้ายตามลำดับ โดยที่แกน X คือเฟรมภาพ และแกน Y คือค่า PERCLOS โดยกราฟเส้นสีน้ำเงินคือค่า PERCLOS และเส้นสีแดงคือ ค่า PERCLOS เท่ากับ 0.4



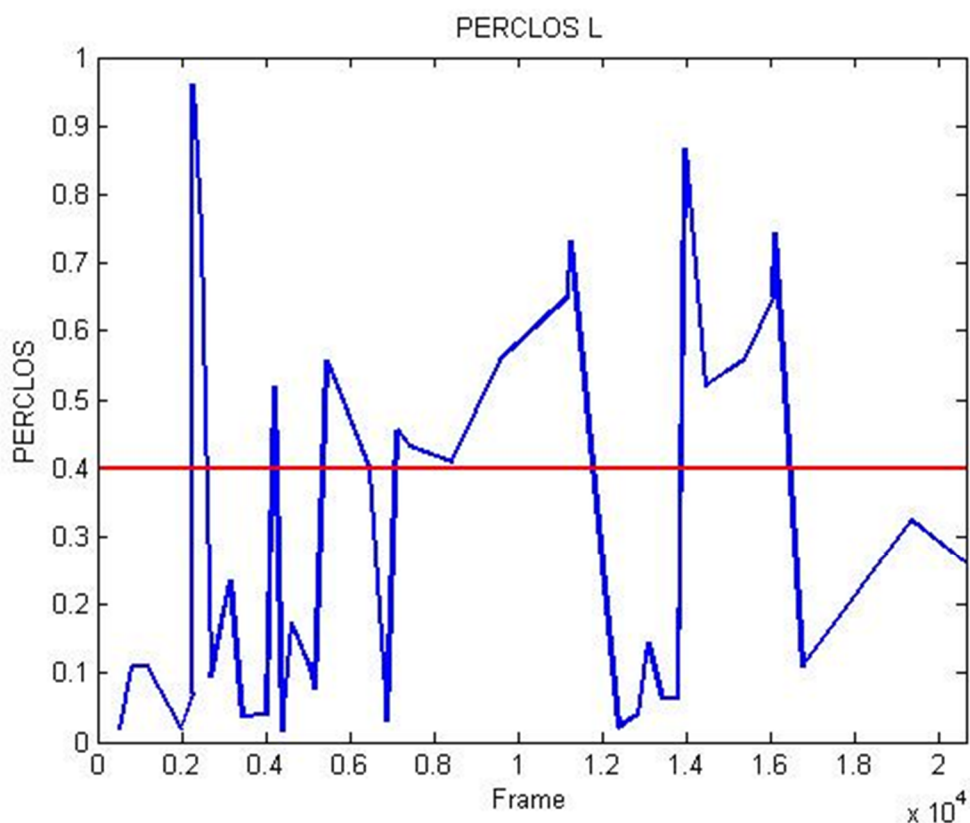
รูปที่ 4.22 กราฟ PERCLOS ของตาขวาในขณะที่ผู้ทดสอบตื่นตัว



รูปที่ 4.23 กราฟ PERCLOS ของตาซ้ายในขณะที่ผู้ทดสอบตื่นตัว



รูปที่ 4.24 กราฟ PERCLOS ของตาขวาในขณะที่ผู้ทดสอบง่วงนอน



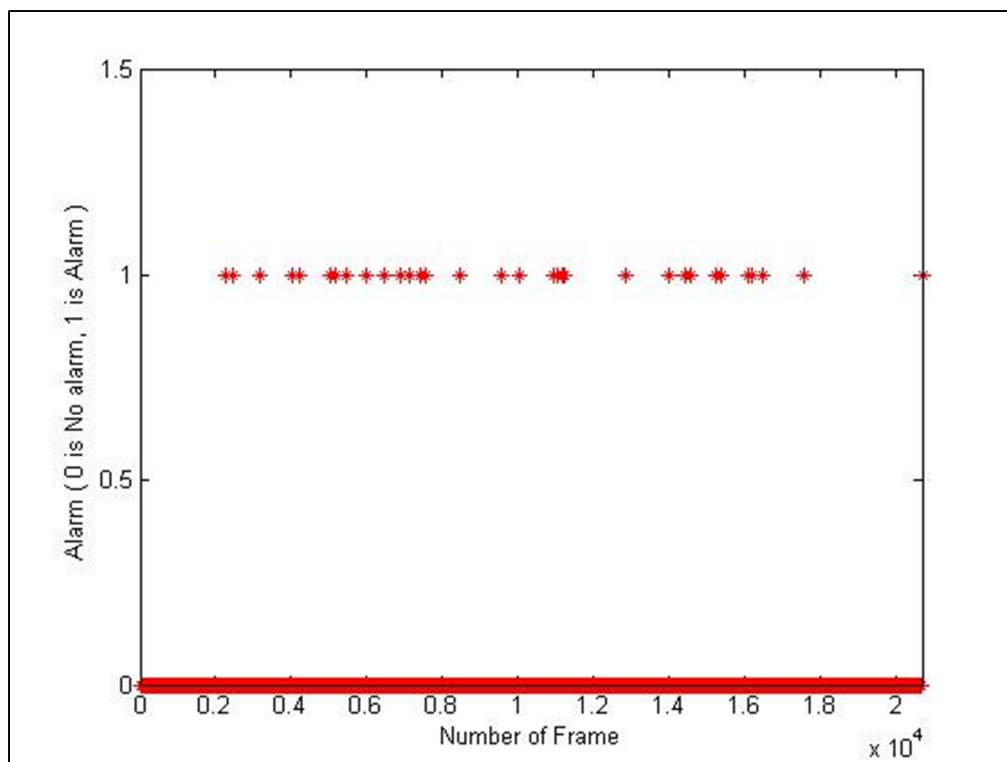
รูปที่ 4.25 กราฟ PERCLOS ของตาซ้ายในขณะที่ผู้ทดสอบง่วงนอน

จากผลการคำนวณ PERCLOS และการคาดเดาอาการง่วงนอนของวิธีการที่ 1 สามารถจำแนกระดับอาการง่วงนอนเป็นร้อยละดังตารางที่ 4.5

ตารางที่ 4.5: ผลการตรวจจับอาการง่วงนอนในวิธีการที่ 1 ซึ่งใช้ระดับของ PERCLOS ที่กำหนดโดยงานวิจัย W. Qing et al. [30]

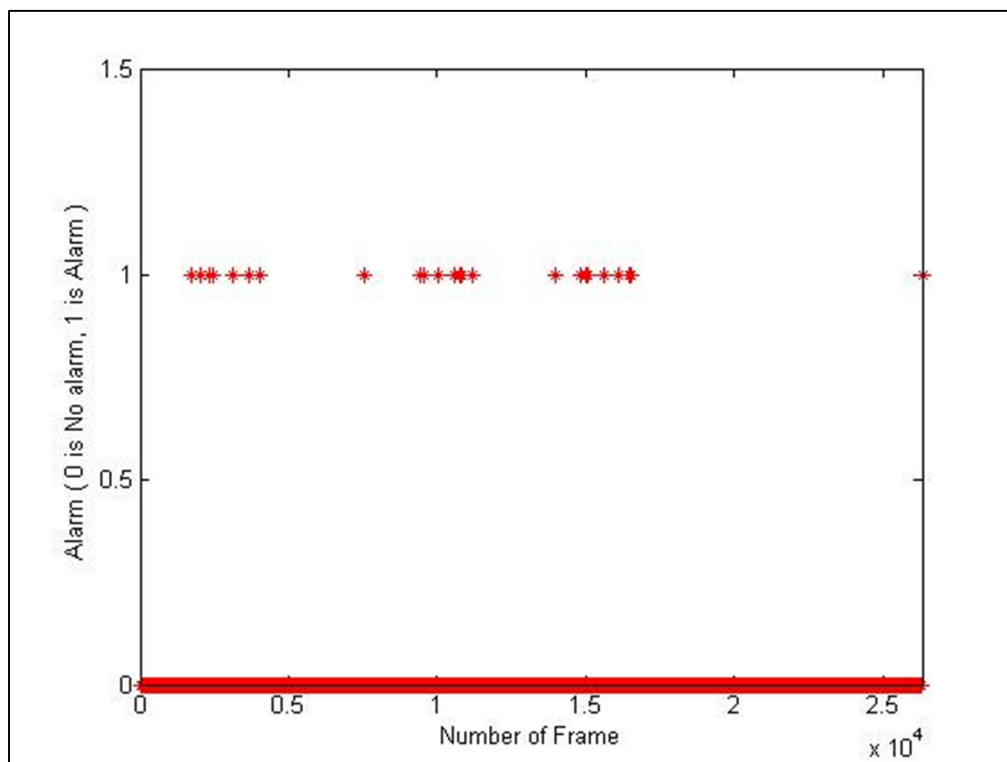
สถานะของผู้ทดสอบ	ฐานข้อมูลกรณีที่ผู้ทดสอบตื่นตัวหรือรู้สึกตัว			ฐานข้อมูลกรณีที่ผู้ทดสอบง่วงนอน		
	ตาขวา (เปอร์เซ็นต์)	ตาซ้าย (เปอร์เซ็นต์)	ค่าเฉลี่ย (เปอร์เซ็นต์)	ตาขวา (เปอร์เซ็นต์)	ตาซ้าย (เปอร์เซ็นต์)	ค่าเฉลี่ย (เปอร์เซ็นต์)
1.ตื่นตัวหรือรู้สึกตัว PERCLOS < 0.4	85.05	89.66	87.355	53.65	60.53	57.095
2.ง่วงนอน PERCLOS >= 0.4	14.95	10.34	12.645	46.34	39.47	42.905

จากการทดสอบพบว่าระบบมีการแจ้งเตือนเกิดขึ้นเมื่อทดสอบกับวิดีโอที่ผู้ทดสอบมีอาการง่วงนอน พบว่าระบบมีการแจ้งเตือนผู้ขับขี่รถยนต์จำนวน 33 ครั้งจากการคำนวณค่า PERCLOS ทั้งหมด 79 ค่าคิดเป็น 41.77 เปอร์เซ็นต์ โดยเกณฑ์คือการแจ้งเตือนที่เกิดขึ้น โดย 0 ไม่มีการแจ้งเตือน (0 is No alarm) และ 1 มีการแจ้งเตือน (1 is Alarm) และเกณฑ์คือ เฟรมที่นำมาประมวลผลแจ้งเตือนจากค่า PERCLOS ซึ่งจากผลการแจ้งเตือนในเฟรมต่าง ๆ ดังรูปที่ 4.26 พบว่าสอดคล้องกับสถานการณ์ง่วงนอนจริง ๆ ของผู้ขับรถในวิดีโอที่กล่าวคือ เมื่อผู้ขับขี่รถยนต์มีอาการง่วงนอนก็ควรจะถูกระบบทำการแจ้งเตือน



รูปที่ 4.26 การแจ้งเตือนผู้ทดสอบเมื่อมีอาการง่วงนอนในวิดีโอที่ผู้ทดสอบมีอาการง่วงนอน (PERCLOS > 0.4)

ในกรณีการทดสอบพบว่าระบบมีการแจ้งเตือนเกิดขึ้นเมื่อทดสอบกับวิดีโอที่ผู้ทดสอบมีอาการตื่นตัวหรือรู้สึกตัว ระบบมีการแจ้งเตือนผู้ขับขี่รถยนต์จำนวน 24 ครั้งจากการคำนวณค่า PERCLOS ทั้งหมด 194 ค่าคิดเป็น 12.37 เปอร์เซ็นต์ในเฟรมต่าง ๆ รูปที่ 4.27 ซึ่งสอดคล้องกับสถานการณ์ตื่นตัวจริง ๆ ของผู้ขับรถในวิดีโอที่กล่าว คือระบบมีการแจ้งเตือนน้อยเนื่องจากผู้ขับขี่มีอาการตื่นตัวหรือรู้สึกตัวในขณะที่ขับรถ



รูปที่ 4.27 การแจ้งเตือนผู้ทดสอบเมื่อมีอาการง่วงนอนในวิดีโอทัศน์
ที่ผู้ทดสอบมีอาการตื่นตัว (PERCLOS > 0.4)

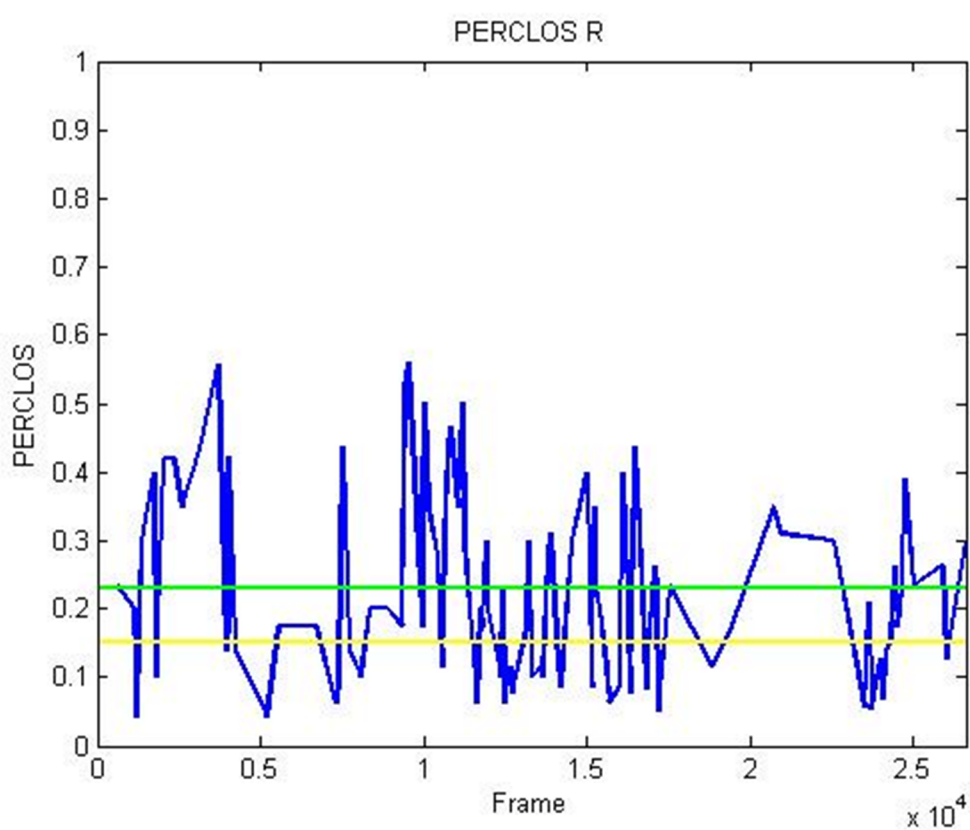
จากกราฟในรูปที่ 4.22 และ รูปที่ 4.23 จะเห็นว่าค่า PERCLOS ที่ได้จากการกระพริบตาแต่ละครั้งส่วนมากจะมีค่าต่ำกว่า 0.4 หรือต่ำกว่า 40 เปอร์เซ็นต์ และกราฟในรูปที่ 4.24 และรูปที่ 4.25 จะมีค่า PERCLOS ที่ได้จากการกระพริบตาแต่ละครั้งส่วนมากจะมีค่าสูงกว่า 0.4 หรือสูงกว่า 40 เปอร์เซ็นต์ ซึ่งสอดคล้องกับผลการทดลองในตารางที่ 4.5 โดยพบว่าร้อยละการง่วงนอนเมื่อทดสอบกับวิดีโอทัศน์ที่ผู้ทดสอบมีอาการง่วงนอนจะมากกว่าการทดสอบกับวิดีโอทัศน์ที่ผู้ทดสอบมีอาการตื่นตัวหรือรู้สึกตัว ซึ่งพบว่ามีร้อยละการง่วงนอนสูงถึง 42.905 เปอร์เซ็นต์ในขณะที่ผลการทดสอบอาการง่วงนอนในวิดีโอทัศน์ที่ผู้ทดสอบมีอาการตื่นตัวหรือรู้สึกตัว จะมีอาการง่วงนอนเพียง 12.645 เปอร์เซ็นต์

2) ผลการตรวจจับอาการง่วงนอนจากการคาดเดาอาการง่วงนอนวิธีการที่ 2

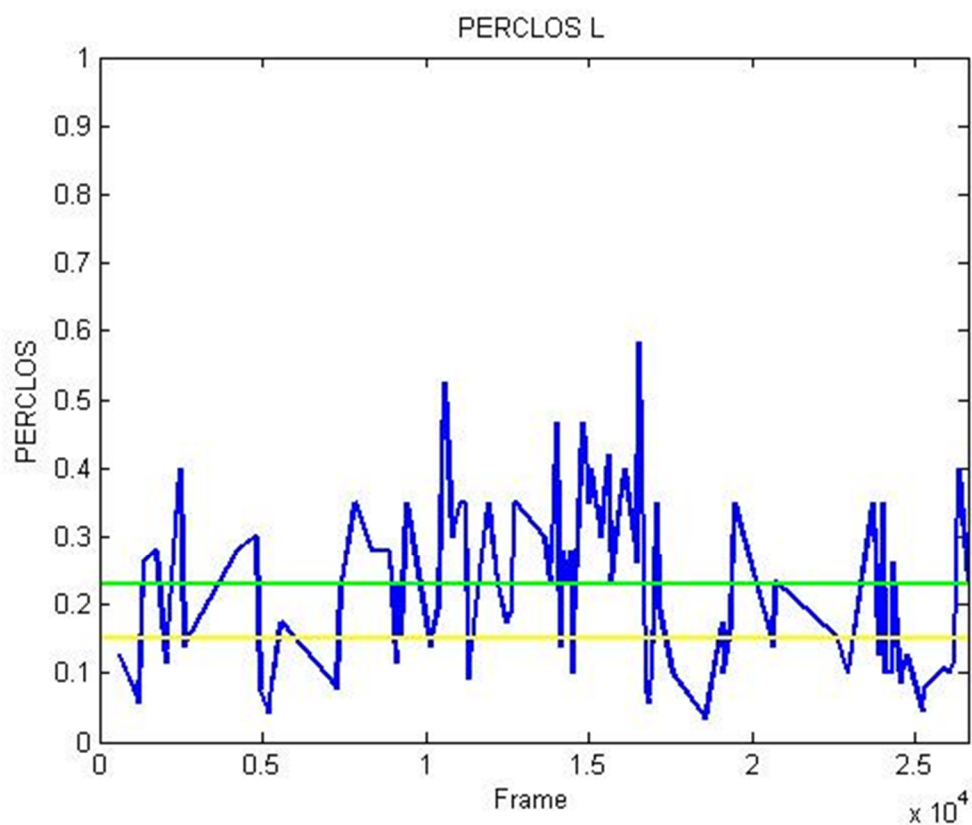
ในการคาดเดาอาการง่วงนอนวิธีการที่ 2 ผู้วิจัยได้นำวิธีการคำนวณ PERCLOS มาใช้คาดเดาอาการง่วงนอนของผู้ขับซีรี่ย์ยนต์ 3 ระดับ ตามรายละเอียดการออกแบบระบบตรวจจับอาการง่วงนอนที่นำเสนอไปแล้วในหัวข้อที่ 3.4 ดังนี้คือ

- ระดับที่ 1 ตื่นตัวหรือรู้สึกตัว (Wake) เป็นช่วงที่มี PERCLOS ต่ำกว่า 15 เปอร์เซ็นต์
- ระดับที่ 2 เมื่อยล้า (Fatigue) เป็นช่วงที่มี PERCLOS อยู่ระหว่าง 15-23 เปอร์เซ็นต์
- ระดับที่ 3 ง่วงนอน (Sleepiness) เป็นช่วงที่มี PERCLOS สูงกว่า 23 เปอร์เซ็นต์

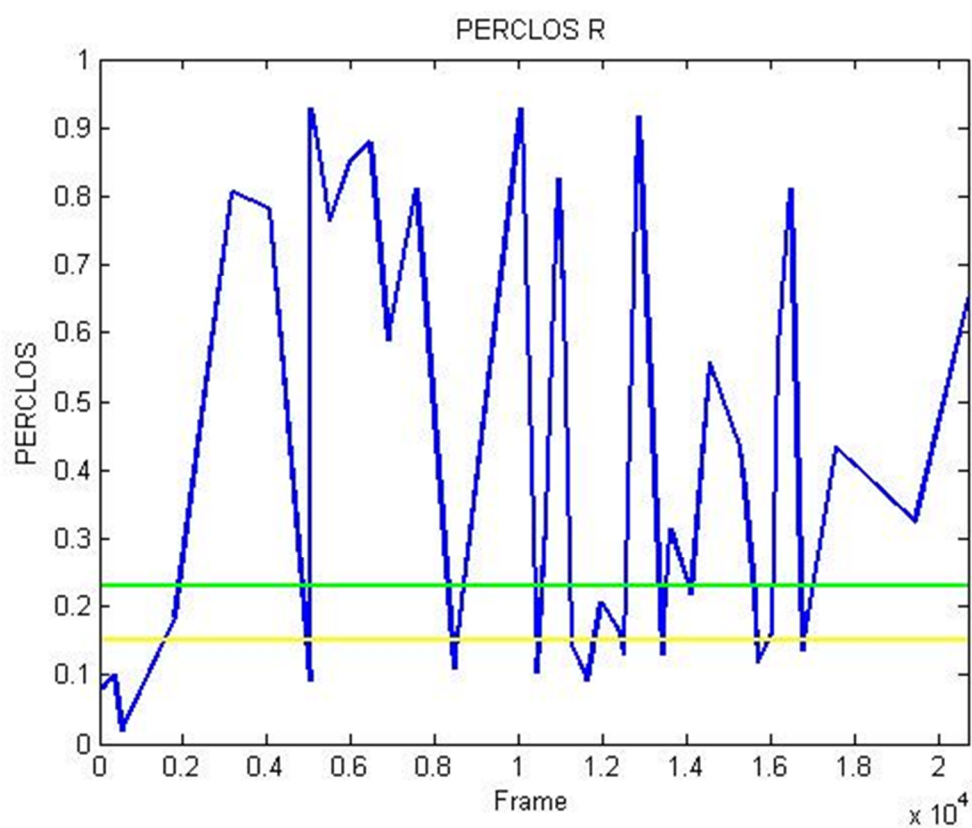
โดยผู้วิจัยได้ใช้ฐานข้อมูลจากเว็บแคม Logitech HD Pro Webcam C920 ที่มีเฟรมภาพขนาด 1920x1080 พิกเซลซึ่งประกอบด้วยวิดีโอที่ผู้ขับขี่ตื่นตัวหรือรู้สึกตัว (Wake) และที่ผู้ขับขี่เมื่อยล้า (Fatigue) แต่ละวิดีโอที่ตีความมีความยาว 20 นาทีที่ใช้ในการทดสอบประสิทธิภาพของระบบซึ่งมีผลการตรวจจับอาการง่วงนอนดังกราฟในรูปที่ 4.28 ถึง รูปที่ 4.29 โดยนำเสนอกราฟ PERCLOS ในขณะที่ผู้ขับรถตื่นตัวหรือรู้สึกตัวของชาวและตาซ้ายตามลำดับ โดยที่แกน X คือเฟรมภาพ และแกน Y คือค่า PERCLOS และรูปที่ 4.30 ถึง รูปที่ 4.31 แสดงกราฟ PERCLOS ในขณะที่ผู้ขับรถเมื่อยล้าและง่วงนอนของชาวและตาซ้ายตามลำดับ โดยแกน X คือเฟรมภาพ และแกน Y คือค่า PERCLOS โดยเส้นกราฟสีน้ำเงินคือค่า PERCLOS เส้นกราฟสีเขียวคือ ค่า PERCLOS เท่ากับ 0.23 และเส้นกราฟสีเหลืองคือ ค่า PERCLOS เท่ากับ 0.15



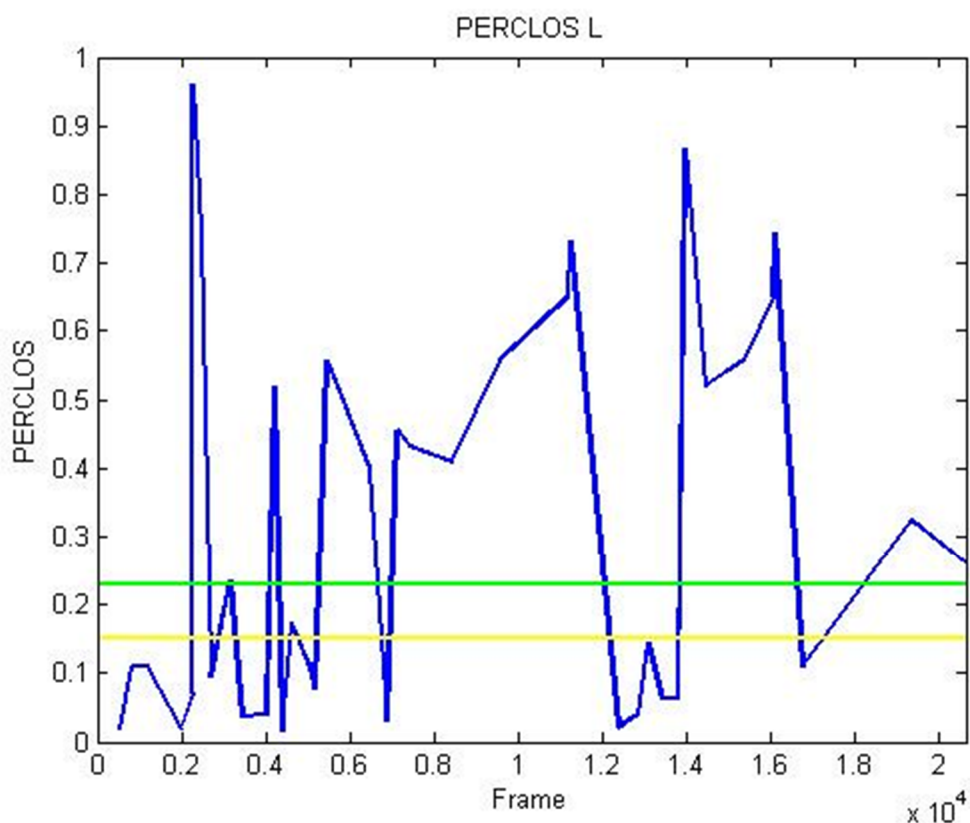
รูปที่ 4.28 กราฟ PERCLOS ของตาขวาในขณะที่ผู้ทดสอบตื่นตัว



รูปที่ 4.29 กราฟ PERCLOS ของตาซ้ายในขณะที่ผู้ทดสอบตื่นตัว



รูปที่ 4.30 กราฟ PERCLOS ของตาขวาในขณะที่ผู้ทดสอบง่วงนอน



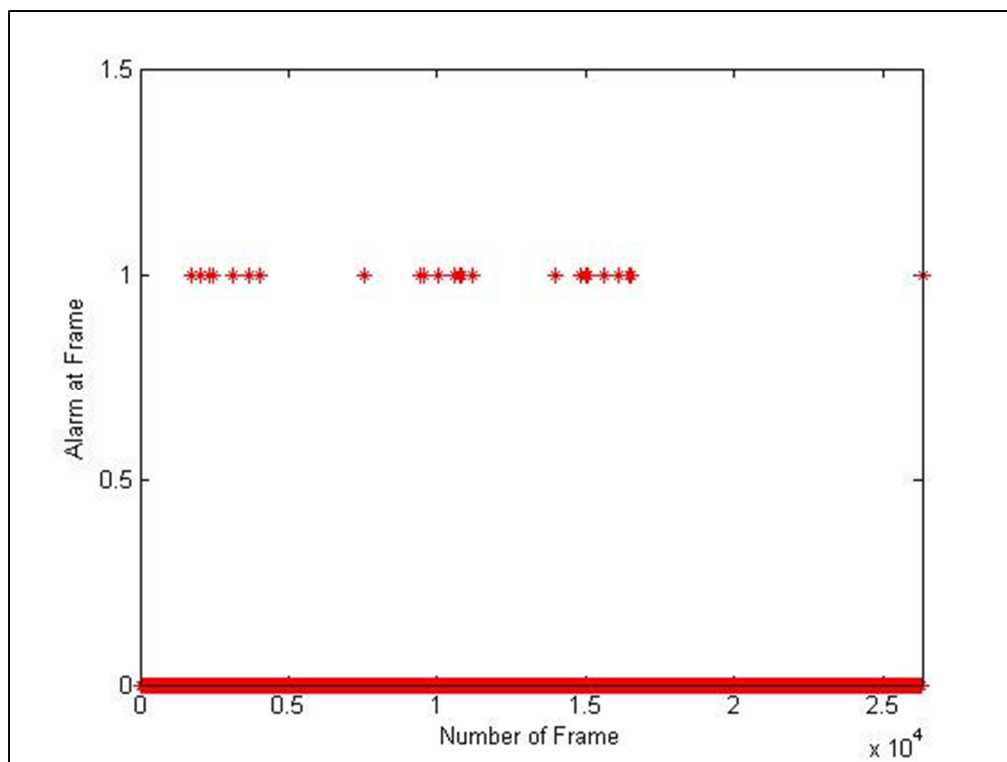
รูปที่ 4.31 กราฟ PERCLOS ของตาซ้ายในขณะที่ผู้ทดสอบง่วงนอน

ผลการคำนวณ PERCLOS และคาดเดาอาการง่วงนอนในวิธีการที่ 2 สามารถจำแนกระดับอาการง่วงนอนเป็นร้อยละดังตารางที่ 4.6 ดังนี้

ตารางที่ 4.6: ผลการตรวจจับอาการง่วงนอนในวิธีการที่ 2 ซึ่งใช้ระดับของ PERCLOS ที่กำหนดโดยงานวิจัย I. Garcia et al. [31]

สถานะของผู้ทดสอบ	ฐานข้อมูลกรณีที่ผู้ทดสอบตื่นตัวหรือรู้สึกตัว			ฐานข้อมูลกรณีที่ผู้ทดสอบง่วงนอน		
	ตาขวา (เปอร์เซ็นต์)	ตาซ้าย (เปอร์เซ็นต์)	ค่าเฉลี่ย (เปอร์เซ็นต์)	ตาขวา (เปอร์เซ็นต์)	ตาซ้าย (เปอร์เซ็นต์)	ค่าเฉลี่ย (เปอร์เซ็นต์)
1.ตื่นตัวหรือรู้สึกตัว PERCLOS < 0.15	34.58	28.97	31.775	29.27	30.36	29.815
2.เมื่อยล้า 0.15 < PERCLOS < 0.23	19.63	29.91	24.77	14.63	33.93	48.56
3.ง่วงนอน PERCLOS > 0.23	45.79	41.12	43.455	56.10	35.71	45.905

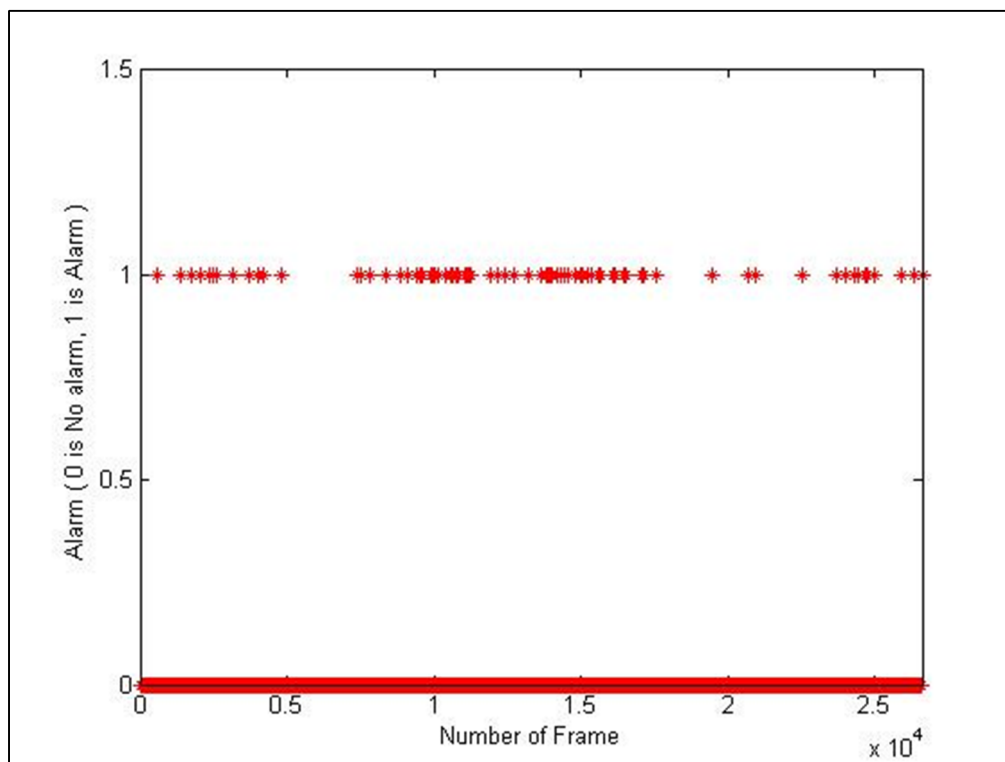
จากการทดสอบพบว่าระบบมีการแจ้งเตือนเกิดขึ้นเมื่อทดสอบกับวิดีโอที่ผู้ทดสอบมีอาการง่วงนอนจำนวน 38 ครั้งจากการคำนวณค่า PERCLOS ทั้งหมด 79 ค่าคิดเป็น 48.10 เปอร์เซ็นต์ในเฟรมต่าง ๆ ดังรูปที่ 4.32 โดยแกนตั้งคือการแจ้งเตือนที่เกิดขึ้น โดย 0 ไม่มีการแจ้งเตือน (0 is No alarm) และ 1 มีการแจ้งเตือน (1 is Alarm) และแกนนอนคือ เฟรมที่นำมาประมวลผลแจ้งเตือนจากค่า PERCLOS ซึ่งสอดคล้องกับสถานการณ์จริง ๆ ของผู้ขับรถในวิดีโอที่กล่าวคือ เมื่อผู้ขับขีรถยนต์มีอาการง่วงนอนก็ควรจะถูกกระตุ้นการแจ้งเตือนบ่อย ๆ



รูปที่ 4.32 การแจ้งเตือนผู้ทดสอบเมื่อมีอาการง่วงนอนในวิดีโอ

ที่ผู้ทดสอบมีอาการง่วงนอน (PERCLOS > 0.23)

และผลการทดสอบในกรณีที่ผู้ทดสอบมีอาการตื่นตัวหรือรู้สึกตัว พบว่ามีการแจ้งเตือนผู้ขับรถจำนวน 77 ครั้งจากการคำนวณค่า PERCLOS ทั้งหมด 194 ค่าคิดเป็น 39.69 เปอร์เซ็นต์ในเฟรมต่าง ๆ ดังรูปที่ 4.33 ซึ่งพบว่าผลการทดสอบไม่สอดคล้องกับสถานการณ์จริง ๆ ของผู้ขับรถในวิดีโอที่กล่าวคือ ระบบไม่ควรจะทำการแจ้งเตือนบ่อย ๆ เนื่องจากผู้ขับขีรถยนต์มีอาการตื่นตัวหรือรู้สึกตัวในขณะที่ขับรถ จึงจำเป็นต้องปรับตั้งการทำงานของระบบแจ้งเตือนใหม่



รูปที่ 4.33 การแจ้งเตือนผู้ทดสอบเมื่อมีอาการง่วงนอนในวิดีโอที่
ที่ผู้ทดสอบมีอาการตื่นตัว (PERCLOS > 0.23)

อย่างไรก็ตามการตั้งค่าระดับในการคาดเดาอาการง่วงนอนทั้ง 2 วิธีดังกล่าว สามารถปรับเปลี่ยนได้โดยทำการทดลองหลาย ๆ ครั้งและเก็บรวบรวมข้อมูลจนสามารถตั้งค่าระดับที่เหมาะสมจนสามารถใช้ในการคาดเดาได้ว่าผู้ทดสอบหรือผู้ขับรถมีอาการง่วงนอนจริง

จากการเก็บข้อมูล และวิเคราะห์ผลข้อมูล นักวิจัยได้ทำการประยุกต์วิธีการคาดเดาอาการง่วงนอนโดยนำระดับการตั้งค่า PERCLOS ของวิธีการที่ 1 และวิธีการที่ 2 มาปรับตั้งระดับใหม่ให้เหมาะสมโดยนำเสนอตารางที่ 4.7

ตารางที่ 4.7: ระดับค่า PERCLOS ที่มีการปรับระดับใหม่

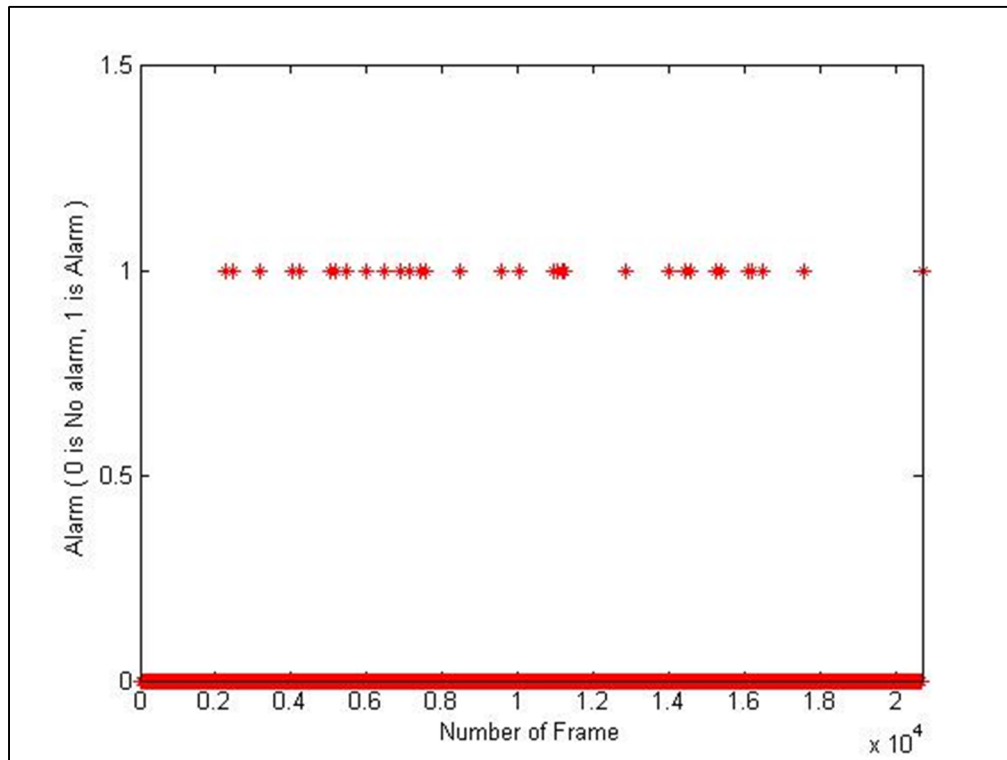
ระดับที่	ค่า PERCLOS ในวิธีการที่ 2	ค่า PERCLOS ที่มีการปรับระดับใหม่
1 ตื่นตัวหรือรู้สึกตัว (Wake)	ต่ำกว่า 15 เปอร์เซ็นต์	ต่ำกว่า 15 เปอร์เซ็นต์
2 เหนื่อยล้า (Fatigue)	อยู่ระหว่าง 15-23 เปอร์เซ็นต์	อยู่ระหว่าง 15-40 เปอร์เซ็นต์
3 ง่วงนอน (Sleepiness)	สูงกว่า 23 เปอร์เซ็นต์	สูงกว่า 40 เปอร์เซ็นต์

เมื่อนำค่า PERCLOS ที่มีการปรับระดับใหม่ตามตารางที่ 4.7 มาใช้ในการคาดเดาอาการง่วงนอนโดยใช้วิธีทัศนชุดเดิมจะได้ผลการทดลองดังตารางที่ 4.8

ตารางที่ 4.8: ผลการตรวจจับอาการง่วงนอนกรณีใช้ระดับของ PERCLOS ที่มีการปรับระดับใหม่

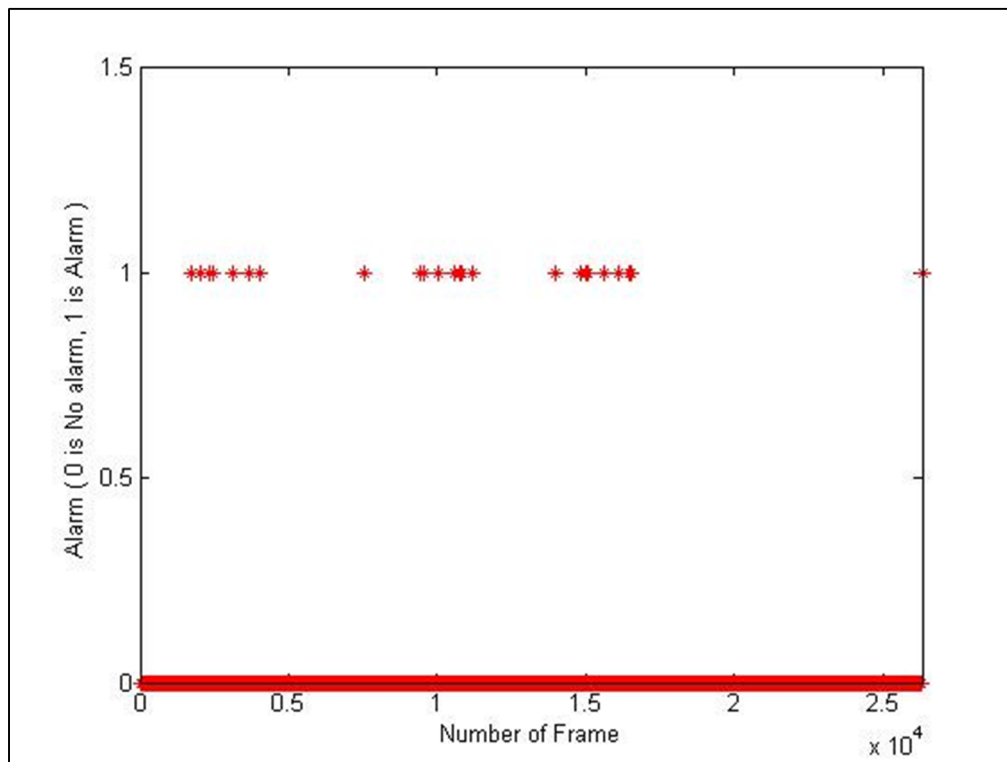
สถานะของผู้ทดสอบ	ฐานข้อมูลกรณีที่ผู้ทดสอบตื่นตัวหรือรู้สึกตัว			ฐานข้อมูลกรณีที่ผู้ทดสอบง่วงนอน		
	ตาขวา (เปอร์เซ็นต์)	ตาซ้าย (เปอร์เซ็นต์)	ตาขวา (เปอร์เซ็นต์)	ตาซ้าย (เปอร์เซ็นต์)	ตาขวา (เปอร์เซ็นต์)	ตาซ้าย (เปอร์เซ็นต์)
1.ตื่นตัวหรือรู้สึกตัว PERCLOS < 0.15	53.27	49.43	51.35	36.59	50.00	43.295
2.เมื่อยล้า 0.15 < PERCLOS < 0.4	31.78	40.23	36.005	17.07	10.53	13.800
3.ง่วงนอน PERCLOS > 0.4	14.95	10.34	12.645	46.34	39.47	42.905

จากการทดสอบพบว่าระบบมีการแจ้งเตือนเกิดขึ้นเมื่อทดสอบกับวิธีทัศนชุดเดิมที่ผู้ทดสอบมีอาการง่วงนอน พบว่าระบบมีการแจ้งเตือนผู้ขับขี่รถยนต์จำนวน 33 ครั้งจากการคำนวณหาค่า PERCLOS ทั้งหมด 79 ค่าคิดเป็น 41.77 เปอร์เซ็นต์ โดยเกณฑ์คือการแจ้งเตือนที่เกิดขึ้น โดย 0 ไม่มีการแจ้งเตือน (0 is No alarm) และ 1 มีการแจ้งเตือน (1 is Alarm) และเกณฑ์นอนคือ เฟรมที่นำมาประมวลผลแจ้งเตือนจากค่า PERCLOS ซึ่งจากผลการแจ้งเตือนในเฟรมต่าง ๆ ดังรูปที่ 4.34 พบว่าสอดคล้องกับสถานการณ์จริง ๆ ของผู้ขับขี่ในวิธีทัศนชุดังกล่าวคือ เมื่อผู้ขับขี่รถยนต์มีอาการง่วงนอนก็ควรจะถูกระบบทำการแจ้งเตือน



รูปที่ 4.34 การแจ้งเตือนผู้ทดสอบเมื่อมีอาการง่วงนอนในวิดีโอที่
ที่ผู้ทดสอบมีอาการง่วงนอน (PERCLOS > 0.4)

และผลการทดสอบในกรณีที่ผู้ทดสอบมีอาการตื่นตัวหรือรู้สึกตัว พบว่ามีการแจ้งเตือนผู้ขับรถยนต์จำนวน 24 ครั้งจากการคำนวณค่า PERCLOS ทั้งหมด 194 ค่าคิดเป็น 12.37 เปอร์เซ็นต์ในเฟรมต่าง ๆ ดังรูปที่ 4.35 ซึ่งพบว่าผลการทดสอบสอดคล้องกับสถานการณ์จริง ๆ ของผู้ขับรถในวิดีโอที่กล่าวคือระบบไม่ควรจะทำการแจ้งเตือนบ่อย ๆ เนื่องจากผู้ขับที่มีอาการตื่นตัวหรือรู้สึกตัวในขณะที่ขับรถ



รูปที่ 4.35 การแจ้งเตือนผู้ทดสอบเมื่อมีอาการง่วงนอนในวิดีโอที่
ที่ผู้ทดสอบมีอาการตื่นตัว (PERCLOS > 0.4)

4.6 สรุป

บทนี้ได้กล่าวถึงผลการทดลองการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมและการตรวจจับใบหน้าโดยการคัดแยกสีผิว จากผลการทดลองจะเห็นว่า การตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมมีร้อยละความถูกต้องน้อยกว่าการตรวจจับใบหน้าโดยการคัดแยกสีผิวอาจเกิดจากจำนวนภาพที่นำมาใช้ในการรู้จำของเครือข่ายประสาทเทียมยังไม่เพียงพอ การเพิ่มร้อยละความถูกต้องของการตรวจจับอาจทำได้โดยการเก็บรวบรวมข้อมูลใบหน้าที่ใช้ในการรู้จำให้มากขึ้น ส่วนการตรวจจับใบหน้าโดยการคัดแยกสีผิวให้ร้อยละความถูกต้องมากกว่าการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมแล้วยังเหมาะสำหรับการตรวจจับใบหน้าของผู้ขับรถเนื่องจากใบหน้าของผู้ขับรถบางครั้งอาจไม่เป็นใบหน้าตรงซึ่งการใช้สีผิวในการตรวจจับใบหน้าสามารถที่จะตรวจจับได้ ในการตรวจจับใบหน้าของผู้ขับรถโดยทั่วไปนั้นพบว่าการเปลี่ยนแปลงของแสงสว่างตลอดเวลาจึงต้องมีการนำเอาเทคนิค Retinex มาช่วยในการปรับแสงสว่าง จากผลการทดลองจะเห็นว่า Retinex สามารถช่วยให้ร้อยละการตรวจจับใบหน้าโดยการคัดแยกสีผิวเพิ่มขึ้น และในหัวข้อต่อมาจะเป็นผลการตรวจจับดวงตาด้วยวิธีการตั้ง Threshold และวิธีการ Integral projection จากผลการทดลองพบว่าวิธีการ Integral projection มีร้อยละของความถูกต้องในการตรวจจับสูงกว่าวิธีการตั้ง Threshold ทั้งนี้อาจเกิดจากฐานข้อมูลที่ใช้มีแสงสว่างเปลี่ยนแปลงตลอดเวลา เนื่องจากภาพวิดีโอที่นำมาทดสอบเป็นภาพวิดีโอจากสถานะการใช้งานจริงขณะขับรถดังนั้นวิธีการตั้ง Threshold จึงไม่เหมาะสมกับการตรวจจับ

ดวงตาในสถานการณ์ดังกล่าว ส่วนวิธีการ Integral projection จะเหมาะสมกับการตรวจจับดวงตาในภาพที่มีแสงสว่างที่เปลี่ยนแปลงตลอดเวลามากกว่า เนื่องจากวิธีการนี้เป็นการหาผลรวมของค่าความเข้มของพิกเซลในแนวแกนตั้งและแกนนอนไม่ว่าแสงสว่างจะมากหรือน้อยบริเวณของม่านตาจะมีผลรวมของค่าความเข้มต่ำกว่าบริเวณอื่น เมื่อนำการตรวจจับดวงตาดังกล่าวมาคำนวณหาค่า PERCLOS จะสามารถคาดเดาอาการง่วงนอนของผู้ขับรถได้

บทที่ 5

บทสรุป

5.1 สรุป

งานวิจัยนี้นำเสนอระบบเฝ้าติดตามอาการง่วงนอน ซึ่งได้ออกแบบการตรวจจับใบหน้า การตรวจจับดวงตา เพื่อคำนวณร้อยละการหลับตาที่สามารถคาดเดาถึงอาการง่วงนอนของผู้ขับขี่รถยนต์ สามารถสรุปงานวิจัยทั้งหมดได้ดังนี้

ในงานวิจัยนี้ได้ออกแบบการตรวจจับใบหน้า 2 วิธี โดยวิธีการแรกเป็นการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมซึ่งเป็นการนำฐานข้อมูลภาพใบหน้าและภาพที่ไม่ใช่ใบหน้าขนาด 200×200 พิกเซล อย่างละ 100 ภาพมาทำการสกัดลักษณะเด่นด้วยวิธีการวิเคราะห์ส่วนประกอบสำคัญและการแปลงเวฟเลตแบบดิสครีทก่อนที่จะนำมาให้เครือข่ายประสาทเทียมรู้จำและได้ค่าน้ำหนักเพื่อนำไปใช้ตรวจจับใบหน้าต่อไป และวิธีการที่สองเป็นการตรวจจับใบหน้าโดยการตัดแยกสีผิวซึ่งเป็นการตัดแยกสีผิวเป็นการตรวจจับใบหน้าที่ใช้ช่วงสีผิวที่อยู่ในปริภูมิสี RGB และ HSV โดยมีการนำวิธีการประมวลผลภาพดิจิทัล (Digital Image Processing: DIP) มาใช้ในการประมวลผล การตรวจจับใบหน้าโดยการตัดแยกสีผิวมีร้อยละการตรวจจับที่ถูกต้องมากกว่าการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมโดยการตรวจจับใบหน้าโดยการตัดแยกสีผิวมีผลการตรวจจับถูกต้องเฉลี่ย 95.13 เปอร์เซ็นต์และการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมมีผลการตรวจจับถูกต้องเฉลี่ย 85.55 เปอร์เซ็นต์ ซึ่งแสดงให้เห็นว่าการตรวจจับใบหน้าโดยการตัดแยกสีผิวเหมาะสำหรับการตรวจจับใบหน้าของผู้ขับรถเนื่องจากวิธีการนี้สามารถตรวจจับใบหน้าได้ถูกต้องในขณะที่ใบหน้าของผู้ขับเรียงเล็กน้อยและสามารถตรวจจับใบหน้าของผู้ขับได้ทุกขนาด ส่วนการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมจะต้องใช้ข้อมูลภาพใบหน้าสำหรับให้เครือข่ายประสาทเทียมรู้จำเป็นจำนวนมากพอสมควรถึงจะสามารถตรวจจับใบหน้าที่มีอริยาบถแตกต่างจากใบหน้าตรงได้ ในงานวิจัยนี้ยังได้มีการประมวลผลภาพก่อน (Pre-processing) นั่นคือการปรับแสงสว่างในภาพก่อนการตรวจจับใบหน้าโดยใช้เทคนิค Multi-scale retinex with color restoration (MSRCR) มาใช้ในการปรับค่าแสงสว่างในภาพให้เหมาะสมก่อนที่จะทำการตรวจจับใบหน้า

ในงานวิจัยนี้ได้ออกแบบการตรวจจับดวงตา 2 วิธี โดยวิธีแรกคือ การตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยนซึ่งเป็นการค้นหาบริเวณของม่านตาจากลักษณะเฉพาะของดวงตาที่มีความเข้มของม่านตาจะเข้มกว่าบริเวณอื่น และวิธีการที่สองคือ การตรวจจับดวงตาด้วยวิธีการ Integral projection ซึ่งเป็นการเป็นการค้นหาตำแหน่งของม่านตาโดยการหาผลรวมของค่าของพิกเซลของภาพระดับสีเทาในทั้งแกนตั้งและแกนนอน การตรวจจับดวงตาโดยเทคนิค Integral projection มีร้อยละการตรวจจับที่ถูกต้องมากกว่าการตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน โดยการตรวจจับดวงตาโดยเทคนิค Integral projection มีผลการตรวจจับถูกต้องเฉลี่ย 98.11 เปอร์เซ็นต์และการตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยนมีผลการ

ตรวจจับถูกต้องเฉลี่ย 92.27เปอร์เซ็นต์ แสดงให้เห็นว่าการตรวจจับดวงตาโดยเทคนิค Integral projection เหมาะสำหรับนำมาตรวจจับดวงตามากกว่าการตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน เนื่องจากวิธีการเป็นการหาผลรวมของค่าความเข้มของพิกเซลในแนวแกนตั้งและแนวนอนไม่ว่าแสงสว่างจะมากหรือน้อย บริเวณของม่านตาจะมีผลรวมของค่าความเข้มต่ำกว่าบริเวณอื่น จึงสามารถตรวจจับดวงตาได้ถึงแม้ว่าภาพมีแสงสว่างเปลี่ยนแปลงตลอดเวลา ส่วนการตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยนจะมีการตรวจจับผิดพลาดขณะที่แสงสว่างในภาพมีการเปลี่ยนแปลงเนื่องจากวิธีการนี้เป็นการตั้งค่าขีดเริ่มเปลี่ยนที่คงที่ไม่สามารถปรับเปลี่ยนค่าขีดเริ่มเปลี่ยนได้เองเมื่อแสงสว่างในภาพมีการเปลี่ยนแปลงไป

การคาดเดาอาการง่วงนอนของผู้ขับรถสามารถคาดเดาได้จากการคำนวณหาเปอร์เซ็นต์การปิดตา หรือ PERCLOS โดยมีเงื่อนไขในการตัดสินใจได้ 2 วิธี ได้แก่ วิธีการแรกเป็นการคาดการณ์ระดับของอาการง่วงนอนจากค่า PERCLOS เมื่อมีค่า PERCLOS มากกว่า 0.4 หรือมีการปิดตานานเกิน 3 วินาที วิธีการที่สองจะเป็นการแบ่งระดับของอาการง่วงนอน 3 ระดับ คือ ระดับที่ 1 ตื่นตัวหรือรู้สึกตัว (Wake) เป็นช่วงที่มี PERCLOS ต่ำกว่า 15 เปอร์เซ็นต์ ระดับที่ 2 เมื่อยล้า (Fatigue) เป็นช่วงที่มี PERCLOS อยู่ระหว่าง 15-23 เปอร์เซ็นต์ ระดับที่ 3 ง่วงซึม (Sleepiness) เป็นช่วงที่มี PERCLOS สูงกว่า 23 เปอร์เซ็นต์ และเมื่อผู้วิจัยได้ทดสอบและเก็บรวบรวมข้อมูลจนได้ปรับระดับของค่า PERCLOS ใหม่โดยประยุกต์ใช้แนวทางของวิธีการที่ 1 และวิธีการที่ 2 มาทำการปรับตั้งระดับการแจ้งเตือนใหม่เป็นระดับที่ 1 ตื่นตัวหรือรู้สึกตัว (Wake) เป็นช่วงที่มี PERCLOS ต่ำกว่า 15 เปอร์เซ็นต์ ระดับที่ 2 เมื่อยล้า (Fatigue) เป็นช่วงที่มี PERCLOS อยู่ระหว่าง 15-40 เปอร์เซ็นต์ ระดับที่ 3 ง่วงซึม (Sleepiness) เป็นช่วงที่มี PERCLOS สูงกว่า 40 เปอร์เซ็นต์ทำให้ผลการตรวจจับอาการง่วงนอนของผู้ขับรถสอดคล้องกับสถานการณ์จริง ๆ ของผู้ขับรถ และผลการทดลองได้แสดงให้เห็นการคาดเดาอาการง่วงนอนตามวิธีการที่นำเสนอสามารถคาดเดาอาการง่วงนอน และแจ้งเตือนผู้ขับขี่รถยนต์ได้จริง

5.2 แนวทางการวิจัยต่อไปในอนาคต

1. พัฒนาอัลกอริทึมการตรวจจับอาการง่วงนอนนี้กับอุปกรณ์เคลื่อนที่ (Mobile devices) เช่น สมาร์ทโฟน (Smart phone) และแท็บเล็ต (Tablet) ซึ่งเป็นอุปกรณ์เคลื่อนที่ที่นิยมใช้กันแพร่หลายและราคาไม่สูงมาก
2. พัฒนาอัลกอริทึมการตรวจจับใบหน้า และดวงตาให้มีร้อยละความถูกต้องมากขึ้น ซึ่งจะให้การคาดเดาอาการง่วงนอนมีความถูกต้องมากขึ้น

บรรณานุกรม

- [1] Q. Wang, J. Yang, M. Ren and Y. Zheng, "Driver Fatigue Detection :A Survey," In Proc. Proceedings of the 6th World Congress on Intelligent Control and Automation, pp. 8587-8591, 2006
- [2] D. J. Jobson, Z. Rahman, G. A. Woodell, "A Multiscale Retinex for Bridging the Gap Between Color Images and the Human Observation of Scenes," In Proc. IEEE Transactions on Image Processing, Vol. 6, No. 7, pp. 965-976, 7 July 1997
- [3] ธนาวุฒิ ประกอบผล, "โครงข่ายประสาทเทียม Artificial Neural Networks," วารสาร มฉก.วิชาการ 73, พศ. 2552
- [4] Q. Ji, Z. Zhu and P. Lan, "Real-Time Nonintrusive Monitoring and Prediction of Driver Fatigue," In Proc. IEEE Transactions on Vehicular Technology, Vol. 53, No. 4, pp. 1052-1068, 2004
- [5] P. R. Tabrizi and R. A. Zoroofi, "Drowsiness Detection Based on Brightness and Numeral Features of Eye Image," In Proc. Fifth International Conference on Intelligent Hiding and Multimedia Signal Processing, pp. 1310, 2009
- [6] S. Ribaric, J. Lovrencic and N. Pavesic, "A Neural-Network-Based System for Monitoring Driver Fatigue," In Proc. 2010 15th IEEE Mediterranean Electrotechnical Conference, pp. 1356-1361, 2010
- [7] R. C. Coetzer and G. P. Hancke, "Eye detection for a real-time vehicle driver fatigue monitoring system," In Proc. IEEE Intelligent Vehicles Symposium(IV) , pp. 66-71, June 5-9, 2011
- [8] X. Li, E. Seignez and P. Loonis, "Reliability-based Driver Drowsiness Detection Using Dempster-Shafer Theory," In Proc. 2012 12th International Conference on Control, pp. 300-304, December 2012
- [9] I. Garcia, S. Bronte, L. M. Bergasa, J. Almazan and J. Yebes, "Vision-based drowsiness detector for Real Driving Conditions," In Proc. 2012 Intelligent Vehicles Symposium Alcala de Henares, pp. 618-623, 2012

- [10] S. Kherchaoui and A. Houacine, "Face Detection Based On A Model Of The Skin Color With Constraints And Template Matching," In Proc. International Conference on Machine and Web Intelligence (ICMWI 2010), pp. 469-472, 2010
- [11] K. S. Park, R. H. Park and Y. G. Kim, "Face Detection Using the 3x3 Block Rank Patterns of Gradient Magnitude Image and a Geometrical Face Model," In Proc. IEEE International Conference on Consumer Electronics (ICCE), pp. 793-794, 2011
- [12] P. P. Paul and M. Gavrilova, "PCA Based Geometric Modeling for Automatic Face Detection," In Proc. International Conference on Computational Science and Its Applications, pp. 33-38, 2011
- [13] M. L. Chiang and S. H. Lau, "Automatic Multiple Faces Tracking and Detection using Improved Edge Detector Algorithm," In Proc. International Conference on IT in Asia(CITA), pp. 1-5, 2011
- [14] C. W. Park, J. M. Kwak, H. Park and Y. S. Moon, "An Effective Method for Eye Detection Based on Texture Information," In Proc. International Conference on Convergence Information Technology, pp. 586-589, 2007
- [15] N. V. Huan, N. T. H. Binh and H. Kim, "Eye-verifier Using Ternary Template for Reliable Eye Detection in Facial Color Image," In Proc. 2009 BTAS IEEE 3rd International Conference on, pp. 1-6, 2009
- [16] N. Alioua, A. Amine, M. Rziza and D. Aboutajdine, "Eye state analysis using iris detection based on Circular Hough Transform," In Proc. Multimedia Computing and Systems (ICMCS), pp. 1-5, 2010
- [17] D. Y. Huang, T. W. Lin, W. C. Hu and M. S. Chen, "Eye Detection Based on Skin Color Analysis with Different Poses under Varying Illumination Environment," In Proc. Fifth International Conference on Genetic and Evolutionary Computing, pp. 252-255, 2011
- [18] S. Chen and C. Liu, "Fast Eye Detection Using Different Color Spaces," In Proc. Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on, pp. 521-526, 2011

- [19] M. A. A. Kashani, M. M. Arani, M. R. R. Fini, "Eye detection and tracking in images with using bag of pixels," In Proc. 2011 IEEE 3rd International Conference on, pp. 64-68, 2011
- [20] C. Phromsuthirak and S. Umchid, "Development of a Geometrical Algorithm for Eye Detection in Color Images," In Proc. The 2012 Biomedical Engineering International Conference (BMEICON-2012), 2012
- [21] P. Goel and S. Agarwal, "Hybrid approach of Haar Cascade Classifiers and Geometrical Properties of Facial Features Applied to Illumination Invariant Gender Classification System," In Proc. International Conference on Computing Sciences, pp. 132-136, 2012
- [22] Md. H. Rahman, F. Jhumur, Md. S. U. Yusuf, T. Das and M. Ahmad, "An Efficient Face Detection in Color Images Using Eye Mouth Triangular Approach," In Proc. IEEE/OSA.IAPR International Conference on Informatics, Electronics & Vision, pp. 530-535, 2012
- [23] F. H. C. Tivive and A. Bouzerdoum, "An Eye Feature Detector Base on Convolutional Neural Network," In Proc. Proceedings of the Eighth International Symposium on, pp. 90-93, 2005
- [24] T. Akashi, Y. Wakasa, K. Tanaka, S. Karungaru and M. Fukumi, "Using Genetic Algorithm for Eye Detection and Tracking in Video Sequence," In Proc. Systemics, Cybernetics and Informatics, Vol.5 , pp. 72-78, 2007
- [25] Y. Zhao. Z. Gao, W. Wu, "The Detection Algorithm of Locomotive Drivers' Fatigue Based on Vision," In Proc. International Congress on Image and Signal Processing (CISP2010), pp. 2686-2690, 2010
- [26] Weber, M., Frontal Face Dataset, California Institute of Technology, available at <http://www.vision.caltech.edu/html-files/archive.html>, 1999
- [27] R. Frischholz. Baoface Database, available at <http://www.Facedetection.com>
- [28] Vision Group of Essex University Face Database, available at <http://cswww.essex.ac.uk/mv/allfaces/index.html>

- [29] Georgia Tech Face Database, available at <ftp://ftp.ee.gatech.edu/pub/users/hayes/facedb/>
- [30] W. Qing, S. B. Xi, X. Bin and Z. Junjie, "A PERCLOS-based Driver Fatigue Recognition Application for Smart Vehicle Space," In Proc. Third International Symposium on Information Processing, pp. 437-441, 2010
- [31] I. Garcia, S. Bronte, L. M. Bergasa, B. Delgado and M. Sevillano, "Vision-based drowsiness detector for a Realistic Driving Simulator," In Proc. 2010 13th International IEEE Annual Conference on Intelligent Transportation Systems, pp. 887-894, September 19-22, 2010

ภาคผนวก

ผลงานวิจัยที่ตีพิมพ์เผยแพร่

เกริกฤทธิ์ ศรีเคน และประโยชน์ คำสวัสดิ์, “การตรวจจับใบหน้าโดยการคัดแยกสีผิว,” การประชุมวิชาการ งานวิจัย และพัฒนาเชิงประยุกต์ ครั้งที่ 5, นครราชสีมา, 8-10 พฤษภาคม 2556, หน้า 389 – 392.

การตรวจจับใบหน้าโดยใช้เทคนิคการคัดแยกสีผิว

เกริกฤทธิ์ ศรีเคน และ ประโยชน์ คำสวัสดิ์

สาขาวิชาวิศวกรรมโทรคมนาคม สำนักวิชาวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี

E-mail: b.kroekrit@hotmail.co.uk

บทคัดย่อ

บทความฉบับนี้ได้นำเสนอการตรวจจับใบหน้าโดยใช้เทคนิคการคัดแยกสีผิว (Skin Tone) ซึ่งจะคัดแยกในระดับพิกเซล (Pixel) ว่าพิกเซลนั้นเป็นสีผิวหรือไม่เป็นสีผิวเพื่อคัดแยกใบหน้าออกจากพื้นหลัง อัลกอริทึมนี้สามารถใช้ได้กับภาพที่มีแสงสว่างน้อย ภาพที่มีการแสดงสีหน้าต่างๆ และสามารถใช้กับภาพที่มีขนาดที่แตกต่างกัน เมื่อนำเทคนิคการตรวจจับใบหน้าไปทดสอบกับฐานข้อมูลใบหน้า CIT, BaoFace, Essex และ Georgia Tech พบว่าอัลกอริทึมที่ได้นำเสนอนี้มีการตรวจจับใบหน้าถูกต้องถึง 95.04 เปอร์เซ็นต์

Abstract

In this paper, we proposed an efficient face detection algorithm based on the skin tone technique which can classify each pixel as skin or non-skin individually, independently from its neighbors. This algorithm is robust against illumination, background, facial expression changes and also works for images of different sizes. The face database used in our experiments consisted of pictures of male and female obtained from the CIT, BaoFace, Essex and Georgia Tech face databases. Experimental results show that the proposed algorithm is good enough to localize a human face in an image with an accuracy of 95.04%.

คำสำคัญ

การคัดแยกสีผิว, การตรวจจับใบหน้า, ความเมื่อยล้า

1. บทนำ

ในปัจจุบันปัญหาอุบัติเหตุที่เกิดขึ้นบนท้องถนนส่วนใหญ่เกิดจากการที่ผู้ขับขี่เกิดความเมื่อยล้า (Fatigue) หรืออาการง่วงนอน (Drowsy) ถึง 57 เปอร์เซ็นต์ของอุบัติเหตุบนท้องถนนทั้งหมด [1] และเสียชีวิตจากความเมื่อยล้าและอาการง่วงนอนถึง 20 เปอร์เซ็นต์ ปัจจัยที่ทำให้เกิดความเมื่อยล้าและอาการง่วงนอน ได้แก่ การพักผ่อนไม่เพียงพอ การอดนอนเป็นระยะเวลา 24 ชั่วโมงติดต่อกัน ปัจจัยเหล่านี้ส่งผลให้ประสิทธิภาพในการขับรถเทียบเท่ากับการมีแอลกอฮอล์ในเลือด 0.10 เปอร์เซ็นต์ ตามกฎหมายแล้วจะจำกัดปริมาณ

แอลกอฮอล์ในเลือดของผู้ขับไม่เกิน 0.05 เปอร์เซ็นต์ จะเห็นว่าปัญหาความเมื่อยล้าและอาการง่วงนอนส่งผลกระทบต่อผู้ขับมากกว่าการมีปริมาณ แอลกอฮอล์ในเลือดตามที่กฎหมายกำหนดไว้ เพราะฉะนั้นจึงต้องออกแบบระบบที่เฝ้าระวังความเมื่อยล้าและอาการง่วงนอนของผู้ขับเพื่อแจ้งเตือนให้ผู้ขับได้ทราบถึงระดับความเมื่อยล้าและอาการง่วงนอนของตนเอง

ระบบเฝ้าระวังความเมื่อยล้าและอาการง่วงนอนในอดีตมีหลายเทคนิค เช่น การวิเคราะห์คลื่นไฟฟ้าที่ได้จากสมอง (Electroencephalography: EEG) การสังเกตพฤติกรรมของยานพาหนะ การใช้กล้องวิดีโอทัศน์ตรวจจับการเคลื่อนไหวของดวงตา ซึ่งการใช้ EEG ในการตรวจจับความเมื่อยล้ามีข้อดีคือมีความแม่นยำสูงแต่จะมีข้อด้อยคือต้องใช้โพรมิเตอร์ติดกับศีรษะของผู้ขับทำให้เกิดการรบกวนผู้ขับในการขับยานพาหนะ การตรวจจับความเมื่อยล้าจากพฤติกรรมของยานพาหนะมีข้อดีคือไม่รบกวนผู้ขับแต่จะมีข้อด้อยคือเมื่อรถเปลี่ยนช่องทางเดินรถ (Lane) และการชะลอความเร็วทำให้ระบบอาจสับสนและตัดสินใจผิดพลาดได้ว่าผู้ขับกำลังมีความเมื่อยล้า และการใช้กล้องวิดีโอทัศน์ตรวจสอบการเคลื่อนไหวของดวงตามีข้อดีคือไม่รบกวนผู้ขับแต่จะมีข้อด้อยคือเมื่อผู้ขับสวมแว่นตาและผลกระทบของแสงสว่าง ซึ่งปัญหานี้สามารถแก้ไขได้โดยการชดเชยแสงของกล้องวิดีโอทัศน์

งานวิจัยนี้ได้นำเสนอการตรวจจับใบหน้า ซึ่งเป็นขั้นตอนแรกของการตรวจจับความเมื่อยล้าและอาการง่วงนอนจากกล้องวิดีโอทัศน์ เมื่อเปรียบเทียบข้อดีและข้อด้อยของระบบนี้พบว่าข้อดีของระบบนี้คือไม่มีการรบกวนผู้ขับและไม่ก่อกวนในการขับรถของผู้ขับ ส่วนข้อด้อยคือการนำกล้องวิดีโอทัศน์มาใช้ในการตรวจจับใบหน้าจะมีผลกระทบจากแสงสว่าง แต่สามารถแก้ไขปัญหานี้ได้โดยใช้การชดเชยแสงสว่างจากกล้องวิดีโอทัศน์ ซึ่งได้จากแหล่งกำเนิดแสง IR (Infrared) [2] ในส่วน

ของผู้ขับที่สวมแว่นตาที่ทำให้การตรวจจับผิดพลาด ซึ่งการตรวจจับผิดพลาดนี้เกิดขึ้นจากตัวผู้ขับเองไม่เกี่ยวกับระบบการตรวจจับความเมื่อยล้าและอาการง่วงนอนโดยการตรวจจับการเคลื่อนไหวของดวงตาให้ประสบผลสำเร็จนั้น สิ่งแรกที่ต้องทำคือการตรวจจับใบหน้าซึ่งในงานวิจัยนี้ผู้วิจัยได้นำเสนออัลกอริทึมที่ใช้ในการตรวจจับใบหน้าเพื่อนำไปสู่การตรวจจับการเคลื่อนไหวของดวงตาต่อไป

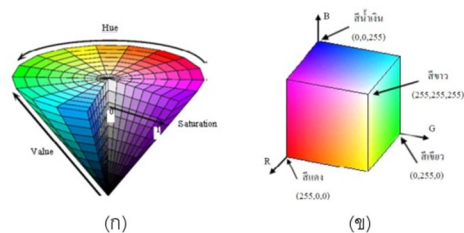
2. ความสำคัญที่มาของปัญหา

ระบบการตรวจจับความเมื่อยล้าและอาการง่วงนอนของผู้ขับจำเป็นมากในการลดอุบัติเหตุบนท้องถนนที่เกิดขึ้นจากความเมื่อยล้าและอาการง่วงนอนของผู้ขับ ระบบที่ดีจะต้องไม่รบกวนและก้าวก่ายผู้ขับ และต้องมีระบบเตือนที่สามารถเตือนให้ผู้ขับรู้ถึงระดับความเมื่อยล้าและอาการง่วงนอนที่เกิดขึ้นและยับยั้งการเกิดอุบัติเหตุที่อาจจะเกิดขึ้น โดยก่อนที่จะสามารถตรวจจับดวงตาได้จะต้องตรวจจับใบหน้าให้ได้เสียก่อน เนื่องจากความถูกต้องของการตรวจจับใบหน้าส่งผลต่อการตรวจจับดวงตา การตรวจจับใบหน้าที่มีในอดีตสามารถตรวจจับได้ถูกต้อง 93.7 เปอร์เซ็นต์ ผู้วิจัยจึงได้ออกแบบระบบที่สามารถตรวจจับใบหน้าที่มีความถูกต้องสูงมากกว่าเดิม

3. งานและทฤษฎีที่เกี่ยวข้อง

3.1 ปริภูมิสี HSV

ปริภูมิสี HSV เป็นปริภูมิสีที่ประกอบด้วย 3 องค์ประกอบได้แก่ค่า H (Hue) คือค่าโทนสีหรือค่าสีัน ซึ่งมีค่าที่แตกต่างกันไปตามความถี่ของแสง ค่า S (Saturation) คือค่าความอิ่มตัวของสี ซึ่งเป็นค่าที่แสดงระดับสีเมื่อเทียบกับค่าโทนสี และค่า V (Value) คือค่าบอกระดับความสว่างของภาพ ซึ่งที่ระดับความสว่างต่ำสุดหมายถึงสีดำ ไม่ว่าค่าโทนสี หรือค่าความอิ่มตัวของสีเท่าใด และระดับความสว่างสูงสุดหมายถึงสีขาว ซึ่งเป็นสีที่สว่างที่สุดของค่าโทนสี และค่าความอิ่มตัวของสีโดยภาพปริภูมิสี HSV สามารถหาได้จากการแปลงภาพปริภูมิสี RGB ด้วยสมการที่ 1 ถึงสมการที่ 4 และแสดงกรวยของแบบจำลองภาพสี HSV ในรูปที่ 1 ก)



รูปที่ 1 ก) กรวยของแบบจำลองภาพสี HSV

ข) ลูกบาศก์แบบจำลองของภาพสี RGB

3.2 ปริภูมิสี RGB

ปริภูมิสี RGB เป็นปริภูมิสีที่ประกอบไปด้วย 3 องค์ประกอบได้แก่ ค่าสีแดง (Red: R) ค่าสีเขียว (Green: G) และค่าสีน้ำเงิน (Blue: B) แต่ละสีจะเป็นอิสระต่อกัน และมีค่าอยู่ระหว่าง 0-255 เมื่อนำสีทั้งสามมารวมกันที่ความเข้มสูงสุดจะได้สีขาว ส่วนใหญ่ปริภูมิสีนี้จะใช้ในอุปกรณ์ที่เกี่ยวข้องกับแสง เช่น จอภาพ กล้องดิจิทัล สแกนเนอร์ ดังรูปที่ 1 ข)

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360^\circ - \theta & \text{if } B > G \end{cases} \quad (1)$$

$$\theta = \cos^{-1} \left\{ \frac{0.5[(R-G)+(R-B)]}{\sqrt{(R-G)^2 + (R-B)^2}} \right\} \quad (2)$$

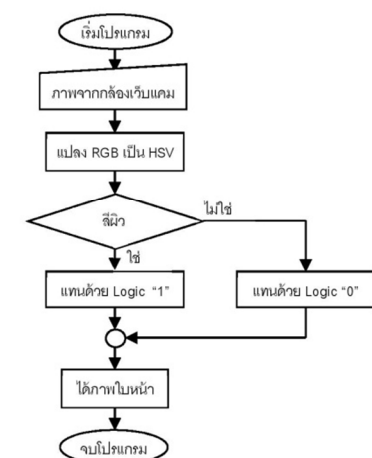
$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)} \quad (3)$$

$$V = \frac{\max(R,G,B)}{255} \quad (4)$$

4. รายละเอียดการพัฒนา

4.1 ภาพรวมของระบบ

ระบบตรวจจับใบหนาดังแผนภาพในรูปที่ 2 มีขั้นตอนดังนี้คือเมื่อได้รับภาพจากกล้องเว็บแคม (Webcam) ที่เป็นปริภูมิสี RGB แล้วจะแปลงภาพจากปริภูมิสี RGB เป็นปริภูมิสี HSV จากนั้นจะเลือกปริภูมิสี R G B และ H มาเข้าเงื่อนไขเพื่อสกัดเอาเฉพาะส่วนที่เป็นใบหน้าออกจากพื้นหลังโดยแทนค่าพิกเซล (Pixel) ที่เป็นใบหน้าด้วยบิต '1' และแทนค่าพิกเซลที่ไม่ใช่ใบหน้าด้วยบิต '0'



รูปที่ 2 แผนภาพระบบตรวจจับใบหน้า

4.2 การออกแบบและพัฒนาระบบ

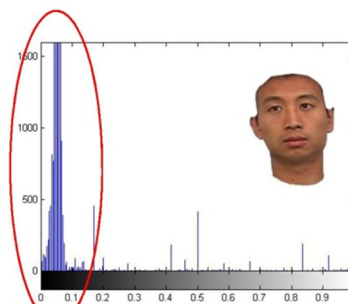
4.2.1 การเลือกใช้ปริภูมิสีที่บ่งบอกถึงใบหน้า

ผู้วิจัยได้นำภาพใบหน้าของคนเอเชียมาคำนวณหาฮิสโตแกรมเพื่อหาค่าขอบเขตของสีผิว ดังแสดงในรูปที่ 3 ซึ่งจะเห็นว่ากราฟฮิสโตแกรมของปริภูมิสี H ที่ได้จากภาพใบหน้า จะมีช่วงขอบเขตที่เป็นสีผิวอยู่ระหว่าง 0.02 - 0.08 เพื่อให้การตรวจจับใบหน้ามีความถูกต้องมากขึ้น จะทำการคำนวณหาฮิสโตแกรมในปริภูมิสี R G และ B ด้วยเช่นกัน

4.2.2 การหาบริเวณใบหน้า

สร้างเงื่อนไขขอบเขตจากทุกปริภูมิสีที่ได้จากหัวข้อที่แล้ว โดยนำเงื่อนไขทั้งหมดมารวมกัน (And) จากนั้นนำภาพใบหน้าหรือภาพจากกล้องวิดีโอคนที่ต้องการตรวจจับใบหน้ามาเข้าเงื่อนไข โดยจะแทนพิกเซลที่เป็นใบหน้าด้วยบิต '1' และพิกเซลที่ไม่ใช่ใบหน้าด้วยบิต '0'

ในงานวิจัยนี้ยังมีข้อจำกัดในเรื่องของแสงสว่างที่ทำให้การตรวจจับใบหน้าผิดพลาดและฐานข้อมูลที่น่ามาทดสอบเป็นใบหน้าของคนจากหลายเชื้อชาติ ซึ่งผู้วิจัยได้ออกแบบระบบนี้ให้เหมาะสมกับคนเอเชียเท่านั้น



รูปที่ 3 ฮิสโตแกรมจากภาพใบหน้า

5. ผลการทดลอง

5.1 ฐานข้อมูลที่ใช้ในการทดลอง

ในการทดลองจะนำอัลกอริทึมที่ออกแบบ มาทดสอบกับฐานข้อมูล California Institute of Technology (CIT) [3] จำนวน 450 ภาพ, Baoface dataset (BaoFace) [4] จำนวน 149 ภาพ, Vision group of Essex University Face Database (Essex) [5] จำนวน 7,900 ภาพ, Georgia Tech Face Database (Georgia Tech) [6] จำนวน 750 ภาพ และภาพใบหน้าจากกล้องเว็บแคม จำนวน 200 ภาพ

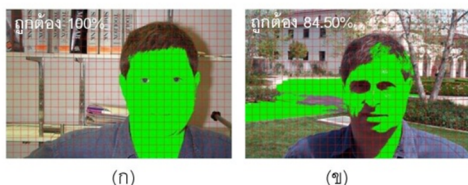
5.2 ผลการทดลองและการวิจารณ์ผล

หลังจากทดสอบอัลกอริทึมการตรวจจับใบหน้าที่ได้ออกแบบกับฐานข้อมูลใบหน้าข้างต้นได้ผลการทดลองดังตารางที่ 1

ตารางที่ 1 ผลการทดลองกับฐานข้อมูลใบหน้า

Face database	อัลกอริทึมการตรวจจับใบหน้าที่น่าเสนอ	อัลกอริทึมการตรวจจับใบหน้า [7]
	ถูกต้อง (%)	ถูกต้อง (%)
CIT [3]	95.04	91.2
BaoFace [4]	92.68	92.5
Essex [5]	93.79	93.7
Georgia Tech [6]	94.82	85.2
Webcam	99.30	-

จากผลการทดลองในตารางที่ 1 จะเห็นว่า การตรวจจับใบหน้าที่ได้นำเสนอมีความถูกต้องสูงถึง 95.04 เปอร์เซ็นต์ ซึ่ง



รูปที่ 4 ก) ตรวจจับถูกต้อง ข) ตรวจจับผิดพลาด

มากกว่าการตรวจจับใบหน้าที่น่าเสนอใน [7] โดยขั้นตอนที่ใช้ในการหาความถูกต้องของการตรวจจับใบหน้า คือนำภาพหลังการตรวจจับใบหน้ามาตีเป็นตารางดังรูปที่ 4 เพื่อนับช่องของตารางที่ตรวจจับถูกต้องและตรวจจับผิดพลาด ผลที่ได้จากการนับจะมี 2 ลักษณะดังนี้คือ

1) ตรวจจับถูกต้องคือ พื้นที่ที่เป็นใบหน้าแล้วระบบตัดสินใจว่าเป็นใบหน้า และพื้นที่ที่ไม่ใช่ใบหน้าแล้วระบบตัดสินใจว่าไม่ใช่ใบหน้า ดังรูปที่ 4 ก)

2) ตรวจจับผิดพลาดคือ พื้นที่ที่ไม่ใช่ใบหน้าแล้วระบบตัดสินใจว่าเป็นใบหน้าและพื้นที่ที่เป็นใบหน้าแล้วระบบตัดสินใจว่าไม่ใช่ใบหน้า ดังรูปที่ 4 ข)

หลังจากการทดสอบอัลกอริทึมการตรวจจับใบหน้าที่ฐานข้อมูลใบหน้าพบว่าส่วนที่ผิดพลาดเกิดขึ้นจากหลายปัจจัย ปัจจัยแรกคือภาพมีแสงสว่างน้อย เพราะแสงสว่างเป็นปัจจัยที่ทำให้สีในภาพเกิดการผิดเพี้ยนไปซึ่งสามารถแก้ปัญหาได้โดยการใช้อีกกล้องที่มีการชดเชยแสงสว่างหรือกล้องที่มีอินฟราเรด ซึ่งปัจจัยนี้ทำให้เกิดความผิดพลาด 3.43 เปอร์เซ็นต์ ปัจจัยที่สองเกิดจากสีของพื้นหลังที่มีสีคล้ายกับสีของใบหน้า ซึ่งเปอร์เซ็นต์การตรวจจับผิดพลาดในปัจจัยนี้เท่ากับ 1.72 เปอร์เซ็นต์ และปัจจัยที่สามเกิดจากฐานข้อมูลใบหน้าที่เป็นภาพของใบหน้าจากหลายเชื้อชาติ ปัจจัยนี้ทำให้เกิดความผิดพลาด 0.77 เปอร์เซ็นต์

จะเห็นว่าการตรวจจับใบหน้าที่ผิดพลาดส่วนใหญ่เกิดจากภาพที่มีแสงสว่างน้อยซึ่งฐานข้อมูลที่ใช้ทดสอบไม่สามารถชดเชยแสงสว่างได้ แต่สำหรับการทดสอบกับกล้องเว็บแคมที่สามารถชดเชยแสงสว่างได้ จะเห็นว่าผลการทดลองกับกล้องเว็บแคมในตารางที่ 1 มีการตรวจจับใบหน้าถูกต้องถึง 99.30 เปอร์เซ็นต์ แสดงให้เห็นว่าเมื่อใช้กล้องที่มีการชดเชยแสงสว่างจะทำให้การตรวจจับใบหน้ามีความถูกต้องมากขึ้น

6. บทสรุป

งานวิจัยนี้เป็นการตรวจจับใบหน้าโดยใช้เทคนิคการคัดแยกสีผิวโดยใช้สีในการแยกเอาใบหน้าออกจากพื้นหลัง ปริภูมิสีที่นำมาใช้คือ R, G, B (จากปริภูมิสี RGB) และ H (จากปริภูมิสี HSV) ซึ่งเหมาะสำหรับภาพใบหน้าที่มีแสงสว่างปกติและเหมาะสำหรับชาวเอเชีย อัลกอริทึมการตรวจจับใบหน้าที่ผู้วิจัยได้นำเสนอนี้มีความถูกต้องของการตรวจจับใบหน้ามากกว่าอัลกอริทึมการตรวจจับใบหน้าเดิมที่มีในอดีต เมื่อนำไปทดสอบกับฐานข้อมูลเดียวกัน ซึ่งมีความถูกต้องมากถึง 95.04 เปอร์เซ็นต์

ในการพัฒนางานวิจัย ผู้วิจัยจะนำอัลกอริทึมดังกล่าวไปพัฒนาโปรแกรมเป็นแอปพลิเคชันสำหรับแท็บเล็ต (Tablet) เพื่อตรวจจับอาการง่วงนอนจากใบหน้าและดวงตาของผู้ขับและแจ้งเตือนต่อไป

7. กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณมหาวิทยาลัยเทคโนโลยีสุรนารีที่ได้ให้ทุนอุดหนุนการวิจัย

8. เอกสารอ้างอิง

- [1] R.Coetzer and G.Hancke, "Driver Fatigue Detection: A Survey," In Proc. International AFRICAN Conference, Vol. 1, pp. 1-6, 2009.
- [2] Guoqing Xu, Yangsheng Wang, Jituo Li, Xiaoxu Zhou, "Real Time Detection of Eye Corners and Iris Center from Images Acquired by Usual Camera," In Proc. Second International Conference on Intelligent Networks and Intelligent Systems, 2009.
- [3] Weber, M., Frontal Face Dataset, California Institute of Technology, <http://www.vision.caltech.edu/html-files/archive.html>, 1999
- [4] R. Frischholz. Baoface Database. Available: <http://www.facedetection.com>
- [5] Vision Group of Essex University Face Database, available at <http://cswwww.essex.ac.uk/mvallfaces/index.html>
- [6] Georgia Tech Face Database, <ftp://ftp.ee.gatech.edu/pub/users/hayes/facedb/>.
- [7] Padma Polash Paul, Marina Gavrilova, "PCA Based Geometric Modeling for Automatic Face Detection," International Conference on Computational Science and Its Applications, pp. 33-38, 2011.