

การตรวจจับอาการง่วงนอนโดยใช้วิธีการประมวลผลภาพ



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมโทรคมนาคม
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2558

**DROWSINESS DETECTION USING IMAGE
PROCESSING**

Kroekrit Sriken



**A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Master of Engineering in Telecommunication Engineering**

Suranaree University of Technology

Academic Year 2015

การตรวจจับอาการ่วงนอนโดยใช้วิธีการประมวลผลภาพ

มหาวิทยาลัยเทคโนโลยีสุรนารี อนุมัติให้นักวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

คณะกรรมการสอบวิทยานิพนธ์

(รศ. ดร.กิตติ อรรถกิจมงคล)

ประธานกรรมการ

(ผศ. ร.อ. ดร.ประโยชน์ คำสวัสดิ์)

กรรมการ (อาจารย์ที่ปรึกษาวิทยานิพนธ์)

(รศ. ดร.อาทิตย์ ศรีแก้ว)

กรรมการ

(ศ. ดร.ชูกิจ ลิ้มปีจันทร์)

รองอธิการบดีฝ่ายวิชาการและนวัตกรรม

(รศ. ร.อ. ดร.กนต์ธร ชำนิประศาสน์)

คณบดีสำนักวิชาวิศวกรรมศาสตร์

เกริกฤทธิ์ ศรีเคน : การตรวจจับอาการง่วงนอนโดยใช้วิธีการประมวลผลภาพ
(DROWSINESS DETECTION USING IMAGE PROCESSING) อาจารย์ที่ปรึกษา :
ผู้ช่วยศาสตราจารย์ เรืออากาศเอก ดร.ประโยชน์ คำสวัสดิ์, 109 หน้า.

งานวิจัยวิทยานิพนธ์นี้มีวัตถุประสงค์เพื่อพัฒนาเทคนิค และอัลกอริทึมการตรวจจับดวงตา สำหรับเฝ้าระวังอาการง่วงนอนโดยใช้วิธีการประมวลผลภาพดิจิทัลของผู้ขับขี่รถยนต์โดยขั้นตอนแรกเป็นการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมและการคัดแยกสีผิว ต่อมาในขั้นตอนที่สองจะใช้เทคนิคและวิธีการ อินทิกรัลโปรเจกชัน (Integral Projection) วิธีการตั้งค่าขีดเริ่มเปลี่ยน (Threshold) และวิธีการจับคู่แม่แบบ (Template Matching) ในการค้นหาตำแหน่งของดวงตา จากนั้นจะทำการประเมินสถานะของผู้ขับขี่รถยนต์ด้วยการคำนวณค่า PERCLOS ซึ่งเป็นการคำนวณหาอัตราส่วนของการปิดตาต่ำกว่าร้อยละ 20 ต่อการปิดตาระหว่างร้อยละ 20-80 ของการเปิดตาสูงสุด ในขั้นสุดท้ายจะนำค่า PERCLOS ที่ได้มาคาดเดาอาการของผู้ขับขี่รถยนต์ ผลการทดลองที่นำเสนอแสดงให้เห็นถึงประสิทธิภาพของอัลกอริทึมและความเป็นไปได้ในการประยุกต์เพื่อการใช้งานได้จริง



สาขาวิชา วิศวกรรมโทรคมนาคม
ปีการศึกษา 2558

ลายมือชื่อนักศึกษา _____
ลายมือชื่ออาจารย์ที่ปรึกษา _____

KROEKRIT SRIKEN : DROWSINESS DETECTION USING IMAGE
PROCESSING AND ARTIFICIAL INTELLIGENT TECHNIQUES. THESIS
ADVISOR : ASST. PROF. FLT. LT. PRAYOTH KUMSAWAT, Ph.D.,
109 PP.

DROWSINESS DETECTION/ FACE DETECTION/ EYE DETECTION

The main objective of this thesis 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 ,threshold and template matching 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.

School of Telecommunication Engineering

Academic Year 2015

Student's Signature _____

Advisor's Signature _____

กิตติกรรมประกาศ

วิทยานิพนธ์นี้ดำเนินการสำเร็จลุล่วงด้วยดี ผู้วิจัยขอขอบพระคุณบุคคลและกลุ่มบุคคลต่างๆ ที่ได้กรุณาให้คำปรึกษา แนะนำ รวมทั้งได้ให้ความช่วยเหลืออย่างดียิ่ง ทั้งด้านวิชาการและด้านการดำเนินงานวิจัย ซึ่งได้แก่

ผู้ช่วยศาสตราจารย์ เรืออากาศเอก ดร. ประโยชน์ คำสวัสดิ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้คำปรึกษา แนะนำ และแนะแนวทางอันเป็นประโยชน์ยิ่งต่องานวิจัย รวมถึงได้ช่วยตรวจทาน และแก้ไขรายงานวิทยานิพนธ์เล่มนี้จนทำให้มีความสมบูรณ์ยิ่งขึ้น รวมทั้งเป็นกำลังใจ และเป็นแบบอย่างที่ดีในการดำเนินชีวิตหลายๆ ด้านให้กับผู้วิจัยเสมอมา

ขอบคุณพี่ๆ เพื่อนๆ น้องๆ บัณฑิตศึกษาทุกท่านรวมถึงมิตรสหายทั้งในอดีตและปัจจุบันที่คอยถามไถ่และให้กำลังใจในการทำวิจัยมาโดยตลอด

สุดท้ายนี้ผู้วิจัยขอขอบคุณอาจารย์ผู้สอนทุกท่านที่ประสิทธิ์ประสาทความรู้ทางด้านต่างๆ ทั้งในอดีตและปัจจุบันและขอกราบขอบพระคุณบิดามารดา รวมถึงญาติพี่น้องของผู้วิจัยทุกท่านที่ได้ให้ความรักความอบอุ่นความห่วงใยการอบรมเลี้ยงดูและให้การสนับสนุนทางการศึกษาอย่างดียิ่งมาโดยตลอดรวมทั้งเป็นกำลังใจที่ยิ่งใหญ่ในยามที่ผู้วิจัยท้อและทุกข์ใจช่วยให้มีพลังเข้มแข็งพร้อมเผชิญกับปัญหาอุปสรรคต่างๆ จนทำให้ผู้วิจัยประสบความสำเร็จในชีวิตเรื่อยมา

เกริกฤทธิ์ ศรีเคน

สารบัญ

หน้า

บทคัดย่อ (ภาษาไทย).....	ก
บทคัดย่อ (ภาษาอังกฤษ).....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ช
สารบัญรูป.....	ซ
บทที่	
1 บทนำ	1
1.1 ความสำคัญและที่มาของปัญหา.....	1
1.2 วัตถุประสงค์ของโครงการวิจัย.....	2
1.3 ขอบเขตของงานวิจัย.....	3
1.4 ขั้นตอนการดำเนินงาน.....	3
1.5 ประโยชน์ที่ได้รับ.....	3
2 ทฤษฎีและหลักการที่เกี่ยวข้อง	4
2.1 บทนำ.....	4
2.2 ทฤษฎีที่เกี่ยวข้องกับเทคนิคทางปัญญาประดิษฐ์.....	4
2.3 ทฤษฎีที่เกี่ยวข้องกับการประมวลผลภาพดิจิทัล.....	8
2.3.1 ปริภูมิสี(Color Space) RGB.....	8
2.3.2 ปริภูมิสี HSV.....	9
2.3.3 การแปลงภาพจากปริภูมิสี RGB เป็นภาพระดับสีเทา (Gray Scale).....	10
2.3.4 ภาพขาว – ดำ (Binary Image).....	11
2.3.5 การหาขอบภาพ (Edge Detection).....	12
2.3.6 การขยายพิกเซล.....	14
2.3.7 การกร่อนพิกเซล.....	15
2.3.8 การจับคู่กับแม่แบบ (Template Matching).....	16

สารบัญ (ต่อ)

หน้า

2.3.9 Retinex [2].....	18
2.4 การทบทวนวรรณกรรม/สารสนเทศ (Information) ที่เกี่ยวข้อง	22
3 การออกแบบระบบเฝ้าติดตามอาการง่วงนอน.....	28
3.1 บทนำ.....	28
3.2 การออกแบบระบบตรวจจับใบหน้า.....	28
3.2.1 การออกแบบการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม.....	28
3.2.2 การออกแบบการตรวจจับใบหน้าโดยใช้การคัดแยกสีผิว	34
3.2.3 การออกแบบการตรวจจับใบหน้าโดยใช้การคัดแยกสีผิวและเทคนิค MSRCR.....	38
3.3 การออกแบบระบบตรวจจับดวงตา.....	41
3.3.1 การค้นหาดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน	43
3.3.2 การค้นหาดวงตาด้วยวิธีการ Integral Projection	45
3.3.3 การค้นหาดวงตาด้วยการจับคู่กับแม่แบบ.....	50
3.4 การออกแบบระบบตรวจจับอาการง่วงนอน	51
3.4.1 Percentage of Eye Closure Over Time (PERCLOS).....	53
3.5 สรุป.....	55
4 ผลการทดลอง.....	57
4.1 บทนำ.....	57
4.2 ผลการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม	57
4.3 ผลการตรวจจับใบหน้าโดยการคัดแยกสีผิว.....	62
4.4 ผลการตรวจจับใบหน้าโดยการคัดแยกสีผิวร่วมกับเทคนิค MSRCR.....	66
4.5 ผลการตรวจจับดวงตาและการตรวจจับอาการง่วงนอน	70
4.5.1 ผลการตรวจจับดวงตา.....	70
4.5.2 ผลการตรวจจับอาการง่วงนอน	77
4.6 บทสรุป	96

สารบัญ (ต่อ)

	หน้า
5 สรุป.....	98
5.1 สรุปผลงานวิจัย.....	98
5.2 แนวทางต่อไปในอนาคต.....	99
รายการอ้างอิง.....	100
ภาพผนวก	
ภาพผนวก ก.บทความวิชาการที่ได้รับการตีพิมพ์เผยแพร่ใน	
ระหว่างการศึกษา.....	104
ประวัติผู้เขียน.....	109



สารบัญตาราง

ตารางที่	หน้า
2.1	งานวิจัยที่เกี่ยวข้องด้านการตรวจจับอาการง่วงนอนขณะขับรถ23
2.2	งานวิจัยที่เกี่ยวข้องด้านการตรวจจับใบหน้า.....24
2.3	งานวิจัยที่เกี่ยวข้องด้านการตรวจจับดวงตา25
2.4	งานวิจัยที่เกี่ยวข้องกับการตรวจจับอาการง่วงนอนโดยใช้เทคนิคทางปัญญาประดิษฐ์27
3.1	พารามิเตอร์ของเทคนิค MSRCR ที่ใช้ในงานวิจัย.....39
4.1	ผลการทดสอบการตรวจจับใบหน้าโดยใช้เทคนิคทางปัญญาประดิษฐ์61
4.2	เปรียบเทียบร้อยละความถูกต้องระหว่างวิธีการของ P. P. Paul and M. Gavrilova กับวิธีการที่นำเสนอ63
4.3	เปรียบเทียบร้อยละความถูกต้องระหว่างวิธีตั้งการ Threshold วิธีการ Integral Projection และวิธี Template Matching70
4.4	ผลการทดลอง บุคคลที่ 171
4.5	ผลการทดลอง บุคคลที่ 272
4.6	ผลการทดลอง บุคคลที่ 373
4.7	ผลการเปรียบเทียบค่า PERCLOS กับเวลาในอิริยาบถที่ปกติ และง่วงนอนหรือเมื่อยล้าในบทความ [25].....77
4.8	เปรียบเทียบร้อยละการปิด – เปิดดวงตา ของผู้ทดสอบตื่นตัวกับผู้ทดสอบง่วงนอน94
4.9	ผลการจำแนกความผิดพลาดในกรณีต่างๆ ที่เกิดขึ้นตามตารางที่ 4.894

สารบัญรูป

รูปที่	หน้า
1.1	ภาพระบบเฟื่อะวังอาคารง้วนอนภายในรถยนต์.....2
2.1	โครงสร้างระบบประสาท.....5
2.2	โครงสร้างการทำงานของเครือข่ายประสาทเทียม.....6
2.3	การเรียนรู้แบบมีผู้ฝึกสอน.....7
2.4	การเรียนรู้แบบไม่มีผู้ฝึกสอน.....8
2.5	ลูกบาศก์แบบจำลองภาพสี RGB.....9
2.6	กรวยของแบบจำลองภาพสี HSV.....10
2.7	ค่าระดับสีเทา 0-255.....11
2.8	แสดงภาพในระดับสีเทาและภาพขาว-ดำ.....12
2.9	เทมเพลตสำหรับการหาขอบภาพในแนวนอน.....12
2.10	เทมเพลตสำหรับการหาขอบภาพในแนวนอน.....13
2.11	เทมเพลตการหาของภาพโดยวิธีโซเบล.....14
2.12	ภาพผลลัพธ์ที่ได้จากการหาขอบภาพ.....14
2.13	เทมเพลตการขยายพิกเซล.....14
2.14	ภาพจำลองการขยายพิกเซล.....15
2.15	ภาพจำลองการขยายพิกเซล.....15
2.16	ภาพจำลองการกร่อนพิกเซล.....16
2.17	ภาพจำลองการกร่อนพิกเซล.....16
2.18	การปรับปรุงโดยใช้เทคนิค Single-Scale Retinex.....21
2.19	การปรับปรุงโดยใช้เทคนิค Single-Scale Retinex.....21
2.20	การปรับปรุงโดยใช้เทคนิค MSR เทียบกับ MSRCR.....22
3.1	ระบบตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม.....29
3.2	ตัวอย่างภาพใบหน้าขนาด 200 × 200 พิกเซล.....30
3.3	ตัวอย่างภาพที่ไม่ใช่ภาพใบหน้าขนาด 200 × 200 พิกเซล.....30

สารบัญรูป (ต่อ)

รูปที่	หน้า
3.4	การแปลงเวฟเล็ดแบบดิสครีท31
3.5	การแปลง DWT31
3.6	เมทริกซ์ A ที่ใช้ในการคำนวณสำหรับการวิเคราะห์ห้วงค์ประกอบ32
3.7	การทดลองเลือกค่าเวกเตอร์ไอเกนที่ค่าต่าง ๆ33
3.8	เลขอร์ของเครือข่ายประสาทเทียม34
3.9	แผนภาพการตรวจจับใบหน้าโดยการคัดแยกสีผิว35
3.10	การแยกปริภูมิสี RGB36
3.11	การแปลงปริภูมิสี RGB เป็น HSV36
3.12	ภาพจากการคัดแยกสีผิว37
3.13	ภาพหลังจากการกำจัดบริเวณที่ไม่ใช่ใบหน้าออก38
3.14	ภาพจากหลังการตรวจจับใบหน้า38
3.15	แผนภาพการตรวจจับใบหน้าด้วยวิธีการ39
3.16	เปรียบเทียบ Retinex แบบต่างๆ40
3.17	ตรวจจับใบหน้าหลังจากการทำ MSRCR41
3.18	แบ่งความสูงของใบหน้าเป็น 5 ส่วน42
3.19	บริเวณใบหน้า 2 ส่วน42
3.20	บริเวณใบหน้า 2 ส่วน42
3.21	หาขอบของวัตถุของบริเวณใบหน้า 2 ส่วน43
3.22	ขยายขอบของวัตถุของบริเวณใบหน้า 2 ส่วน43
3.23	ลักษณะของดวงตา43
3.24	ภาพระดับสีเทา44
3.25	ภาพหลังการตั้งค่าขีดเริ่มเปลี่ยน44
3.26	ลักษณะเฉพาะของตา44
3.27	ผลการตรวจจับดวงตา44
3.28	บริเวณดวงตาที่เป็นภาพระดับสีเทา46
3.29	กราฟ Horizontal Integral Projection47

สารบัญรูป (ต่อ)

รูปที่	หน้า
3.30	กราฟ Vertical Integral Projection.....48
3.31	ภาพม่านตาที่ได้จากวิธีการ Integral Projection50
3.32	ภาพแม่แบบของตา50
3.33	การค้นหาจากซ้ายไปขวา บนลงล่าง50
3.34	การติดตั้งกล้องวิดีโอทัศนียภาพในยานพาหนะ52
3.35	ขั้นตอนการตรวจจับอาการง่วงนอน.....53
3.36	เฟรมที่ใช้ในการคำนวณ PERCLOS [26].....54
3.37	การกำหนดช่วงเวลาในการประมวลผล PERCLOS [26].....54
3.38	กราฟ PERCLOS55
4.1	ภาพใบหน้าขนาด 200 × 200 พิกเซล.....58
4.2	ภาพที่ไม่ใช่ใบหน้าขนาด 200 × 200 พิกเซล58
4.3	การแปลง DWT59
4.4	เลขอร์ของเครือข่ายประสาทเทียม59
4.5	กราฟการรู้จำของเครือข่ายประสาทเทียมในการทดลองที่ 1.....60
4.6	กราฟการรู้จำของเครือข่ายประสาทเทียมในการทดลองที่ 2.....60
4.7	กราฟการรู้จำของเครือข่ายประสาทเทียมในการทดลองที่ 3.....61
4.8	เปรียบเทียบความถูกต้อง63
4.9	แสดงร้อยละการตรวจจับถูกต้องและผิดพลาด.....64
4.10	เปรียบเทียบร้อยละของการตรวจจับใบหน้าที่ถูกต้องและผิดพลาด.....65
4.11	ตัวอย่างการตรวจจับใบหน้าที่ถูกต้อง65
4.12	ตัวอย่างการตรวจจับใบหน้าที่ผิดพลาด66
4.13	การติดตั้งกล้องเว็บแคมภายในยานพาหนะ66
4.14	ภาพที่ใช้ในการทดสอบ.....67
4.15	แผนภาพการตรวจจับใบหน้าที่รวมกับเทคนิค MSRCR68
4.16	ตัวอย่างการปรับปรุงแสงสว่างด้วยเทคนิค MSRCR.....69
4.17	ตัวอย่างการปรับปรุงแสงสว่างด้วยเทคนิค MSRCR.....69

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.18	เปรียบเทียบการตรวจจับใบหน้าด้วยวิธีการตัดแยกสีผิว.....70
4.19	เปรียบเทียบกราฟ Integral Projection ของภาพดวงตา.....76
4.20	กราฟร้อยละของการปิดตาเทียบกับเฟรมเวลา.....78
4.21	ฐานข้อมูลที่ใช้ในการทดสอบตื่นตัว.....79
4.22	ฐานข้อมูลที่ใช้ในการทดสอบขณะง่วงนอน.....80
4.23	หน้าต่างแสดงค่าต่างๆในการตรวจจับอาการง่วงนอน.....81
4.24	ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 1 สถานะตื่นตัว.....82
4.25	ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 1 สถานะตื่นตัว.....83
4.26	ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 1 สถานะง่วงนอน.....84
4.27	ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 1 สถานะง่วงนอน.....85
4.28	ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 2 สถานะตื่นตัว.....86
4.29	ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 2 สถานะตื่นตัว.....87
4.30	ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 2 สถานะง่วงนอน.....88
4.31	ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 2 สถานะง่วงนอน.....89
4.32	ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 3 สถานะตื่นตัว.....90
4.33	ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 3 สถานะตื่นตัว.....91
4.34	ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 3 สถานะง่วงนอน.....92
4.35	ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 3 สถานะง่วงนอน.....93
4.36	การแจ้งเตือนอาการง่วงนอนของบุคคลที่ 1.....95
4.37	การแจ้งเตือนอาการง่วงนอนของบุคคลที่ 2.....96
4.38	การแจ้งเตือนอาการง่วงนอนของบุคคลที่ 3.....96

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของปัญหา

ในปัจจุบันนี้ การเกิดอุบัติเหตุเป็นสาเหตุการตายอันดับสองของคนไทยรองมาจากโรคมะเร็ง โดยอุบัติเหตุทางรถยนต์เป็นอุบัติเหตุที่เกิดขึ้นบ่อยเป็นอันดับต้น ๆ ซึ่งทำให้เกิดความสูญเสียทั้งชีวิตและทรัพย์สิน โดยสาเหตุหลักของการเกิดอุบัติเหตุทางรถยนต์คือ ความประมาทของผู้ใช้รถใช้ถนน ความไม่พร้อมของสภาพยานพาหนะที่ใช้และความไม่พร้อมของผู้ขับขี่ เช่น พักผ่อนไม่เพียงพอทำให้เกิดอาการง่วงนอน (Drowsiness) และเกิดการหลับในขณะที่ขับรถ หรืออาจอยู่ในอาการมึนเมาจากการดื่มสุราหรือรับประทานยาที่อาจทำให้เกิดการง่วงซึม เป็นต้น นอกจากนี้ การขับรถเป็นระยะเวลานานโดยไม่หยุดพักก็อาจทำให้เกิดความเมื่อยล้า (Fatigue) และการง่วงนอนซึ่งทำให้เกิดการหลับในขณะที่ขับรถได้เช่นเดียวกัน

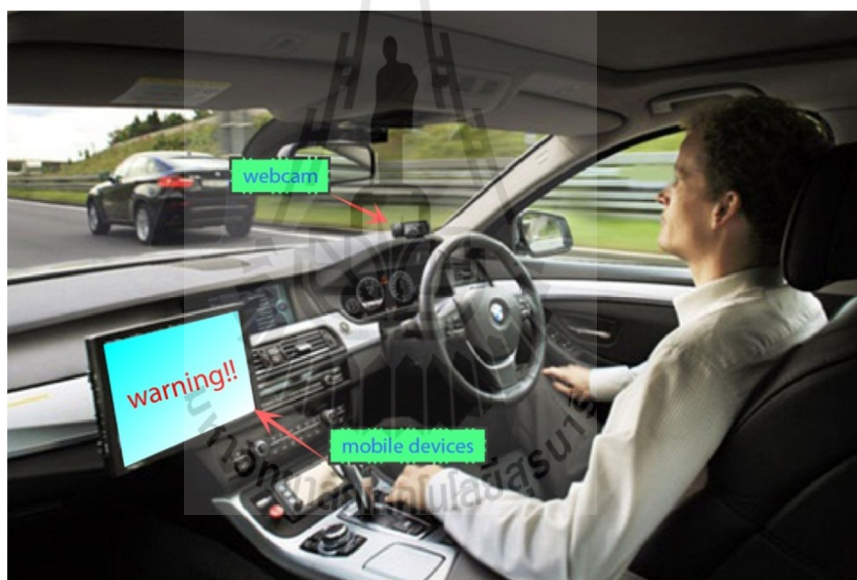
แนวทางหนึ่งที่จะสามารถลดการเกิดอุบัติเหตุในขณะที่ขับรถอันเนื่องมาจากการหลับในคือการแจ้งเตือนแก่ผู้ที่กำลังปฏิบัติงานในขณะที่กำลังเกิดการง่วง ด้วยการใส่เสียงเตือนหรือใช้ไฟสัญญาณเตือน เพื่อให้ผู้ปฏิบัติงานได้รับทราบ ทั้งนี้เพื่อที่จะได้แก้ไขอาการง่วงนั้นอย่างทันท่วงที วิธีการตรวจจับอาการง่วงนอนขณะขับรถนั้นสามารถทำได้โดยอาศัยการสังเกตอาการของผู้ปฏิบัติงานเช่น ลักษณะของเปลือกตาที่หริ่งมากกว่าปกติ และมีจำนวนครั้งของการกระพริบตานั้นน้อยลง มีการเคลื่อนไหวของศีรษะเอียงไปมา หรือการเกิดอาการนิ่งของเปลือกตาหรือศีรษะที่นานกว่าปกติ เป็นต้น [1]

วิธีการตรวจจับอาการง่วงนอนโดยอาศัยการสังเกตอาการของผู้ปฏิบัติงานนั้น สามารถทำได้โดยใช้หลักการของประมวลผลภาพดิจิทัล (Digital Image Processing) โดยหลักการดังกล่าวจะทำการตรวจจับดวงตาเพื่อค้นหาภาพดวงตาในข้อมูลภาพที่เรียงลำดับกัน (Image Sequences) หรือที่เรียกว่าภาพวิดีโอ (Video) จากนั้นจะทำการเฝ้าสังเกตลักษณะของดวงตาและการกระพริบตาในข้อมูลภาพ เมื่อตรวจพบอาการง่วงนอนดังกล่าวก็จะทำการปลุกให้ตื่น หรือทำการแจ้งเตือนให้เปลี่ยนอิริยาบถหรือหยุดพักการทำงานชั่วคราว

เนื่องจากในปัจจุบัน เทคโนโลยีวงจรรวม (Integrated Circuit Technology : IC) มีความเจริญรุดหน้าไปมาก ทำให้การประมวลผลของคอมพิวเตอร์ทำได้อย่างรวดเร็วและมีสมรรถนะสูงมากขึ้น ประกอบกับเทคนิคและวิธีการทางปัญญาประดิษฐ์ (Artificial Intelligent Techniques) ถูกนำมา

ประยุกต์ใช้งานอย่างกว้างขวางในด้านต่าง ๆ โดยเฉพาะในด้านการมองเห็นของคอมพิวเตอร์ (Computer Vision) เนื่องจากเทคนิคและวิธีการดังกล่าวให้ผลลัพธ์ที่ดีในการค้นหาคำตอบที่เหมาะสมที่สุดในวงกว้าง (Global Optimization) และสามารถค้นหาคำตอบ ทำการเรียนรู้และรู้จำได้อย่างถูกต้องและรวดเร็ว ซึ่งความสามารถเหล่านี้จะเพิ่มขึ้นตามสมรรถนะของคอมพิวเตอร์ที่เพิ่มขึ้น

ดังนั้นในงานวิจัยนี้ จึงมุ่งเน้นไปที่การนำเทคนิคและวิธีการทางปัญญาประดิษฐ์ เครือข่ายประสาทเทียม (Neural Networks : NN) มาทำการประยุกต์ใช้เพื่อทำการค้นหาค่าที่เหมาะสมที่สุดของพารามิเตอร์ต่าง ๆ ที่ใช้ในการตรวจจับใบหน้าและดวงตา เพื่อให้ได้ผลลัพธ์ของการตรวจจับใบหน้าและดวงตาและการหาอัตราการกระพริบของเปลือกตาสำหรับระบบเฝ้าติดตามอาการง่วงนอนขณะขับรถที่ถูกต้องและแม่นยำ แผนภาพการประยุกต์ใช้ระบบเฝ้าติดตามอาการง่วงนอน โดยใช้เทคนิคการประมวลผลภาพดิจิทัลแสดงในรูปที่ 1.1



รูปที่ 1.1 ภาพระบบเฝ้าระวังอาการง่วงนอนภายในรถยนต์

1.2 วัตถุประสงค์ของโครงการวิจัย

1.2.1 เพื่อศึกษาวิธีการตรวจจับใบหน้าและดวงตาด้วยหลักการประมวลผลภาพดิจิทัล

1.2.2 เพื่อดำเนินการออกแบบตัวตรวจจับใบหน้าและดวงตาสำหรับระบบเฝ้าติดตามอาการง่วงนอนขณะขับรถ

1.2.3 เพื่อดำเนินการออกแบบโปรแกรมและสร้างชุดต้นแบบระบบเฝ้าติดตามอาการง่วงนอนขณะขับรถสำหรับแจ้งเตือนผู้ขับขี่รถยนต์ เพื่อลดอัตราการเกิดอุบัติเหตุลง

1.2.4 เพื่อสร้างองค์ความรู้ใหม่ในการออกแบบระบบเฝ้าติดตามอาการร่งวงนอนขณะขับรถ ด้วยวิธีการทางปัญญาประดิษฐ์

1.3 ขอบเขตของงานวิจัย

1.3.1 สัญญาณภาพวิดีโอที่ใช้ในการทดสอบต้องมีความละเอียดอย่างต่ำในระดับ VGA 640×480 และมีความเร็วอย่างต่ำ 25 เฟรมต่อวินาที

1.3.2 ทำการออกแบบและพัฒนาระบบเฝ้าติดตามอาการร่งวงนอนขณะขับรถโดยใช้โปรแกรม MATLAB C/C++ และโปรแกรม LabVIEW

1.4 ขั้นตอนการดำเนินงาน

วิธีการดำเนินการวิจัย มีขั้นตอนดังต่อไปนี้

1.4.1 ศึกษาวิธีการตรวจจับใบหน้าและดวงตาด้วยหลักการของการประมวลผลภาพดิจิทัล

1.4.2 ศึกษาวิธีการตรวจจับใบหน้าและดวงตาที่ประยุกต์ใช้เทคนิคทางปัญญาประดิษฐ์

1.4.3 ดำเนินการเก็บข้อมูลภาพวิดีโอของใบหน้าและดวงตาในขณะขับรถยนต์

1.4.4 ดำเนินการออกแบบตัวตรวจจับและคัดแยกภาพดวงตาจากภาพวิดีโอเพื่อนำไปใช้กับระบบตรวจจับอาการร่งวงนอนขณะขับรถ โดยประยุกต์ใช้เทคนิคทางปัญญาประดิษฐ์

1.4.5 ดำเนินการออกแบบโปรแกรมและชุดต้นแบบระบบเฝ้าติดตามอาการร่งวงนอนขณะขับรถ

1.4.6 ดำเนินการทดสอบและพัฒนาระบบเฝ้าติดตามอาการร่งวงนอนขณะขับรถ

1.5 ประโยชน์ที่ได้รับ

1.5.1 เป็นองค์ความรู้ในการวิจัยในด้านการประมวลผลสัญญาณภาพดิจิทัลในมหาวิทยาลัยและหน่วยงานอื่น ๆ

1.5.2 นักวิจัยในโรงงานอุตสาหกรรมสามารถนำไปใช้ออกแบบและพัฒนาระบบเฝ้าติดตามอาการร่งวงนอนเพื่อเพิ่มมูลค่าผลิตภัณฑ์

1.5.3 สามารถนำไปใช้ออกแบบและพัฒนาระบบเฝ้าติดตามอาการร่งวงนอนสำหรับธุรกิจที่เกี่ยวข้องกับการขนส่งสินค้าหรือผู้โดยสาร

บทที่ 2

ทฤษฎีและหลักการที่เกี่ยวข้อง

2.1 บทนำ

ในบทนี้จะกล่าวถึงทฤษฎีที่ใช้ในงานวิจัย ได้แบ่งออกเป็น 2 ส่วนใหญ่ ๆ ได้แก่ทฤษฎีที่เกี่ยวข้องกับเทคนิคทางปัญญาประดิษฐ์ และทฤษฎีที่เกี่ยวข้องกับการประมวลผลภาพดิจิทัล การปรับปรุงแสงสว่างด้วยเทคนิค Retinex และในตอนท้ายได้กล่าวถึงวรรณกรรมหรือบทความที่เกี่ยวข้องกับงานวิจัยที่มีมาแล้วในอดีต

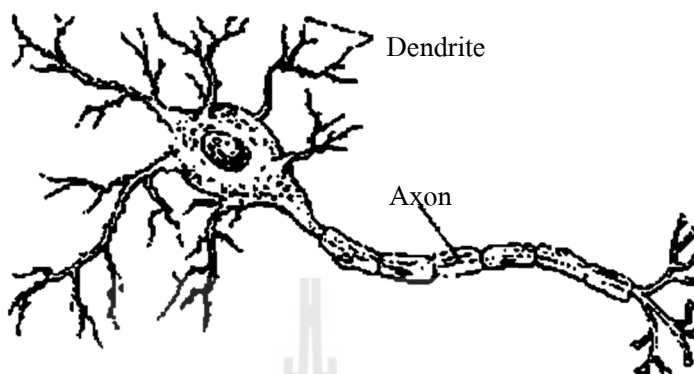
2.2 ทฤษฎีที่เกี่ยวข้องกับเทคนิคทางปัญญาประดิษฐ์

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

หลักการสำคัญของเครือข่ายประสาทเทียม คือ ความพยายามที่จะลอกเลียนแบบการทำงานของเซลล์ประสาทในสมองมนุษย์เพื่อทำงานได้อย่างมีประสิทธิภาพ ลักษณะทั่วไปของเครือข่ายประสาทเทียม คือการที่โหนด (Node) ต่าง ๆ จำลองมาจากจุดประสานประสาท (Synapse) ของเซลล์ประสาทระหว่างปลายในการรับกระแสประสาท เรียกว่า เด็นไดรต์ (Dendrite) ซึ่งเป็นอินพุต และปลายในการส่งกระแสประสาทเรียกว่า แกนประสาท (Axon) ซึ่งเป็นเหมือนเอาต์พุตของเซลล์ โดยมีฟังก์ชันเป็นตัวกำหนดสัญญาณส่งออก (Activation Function or Transfer Function) นั่นเอง

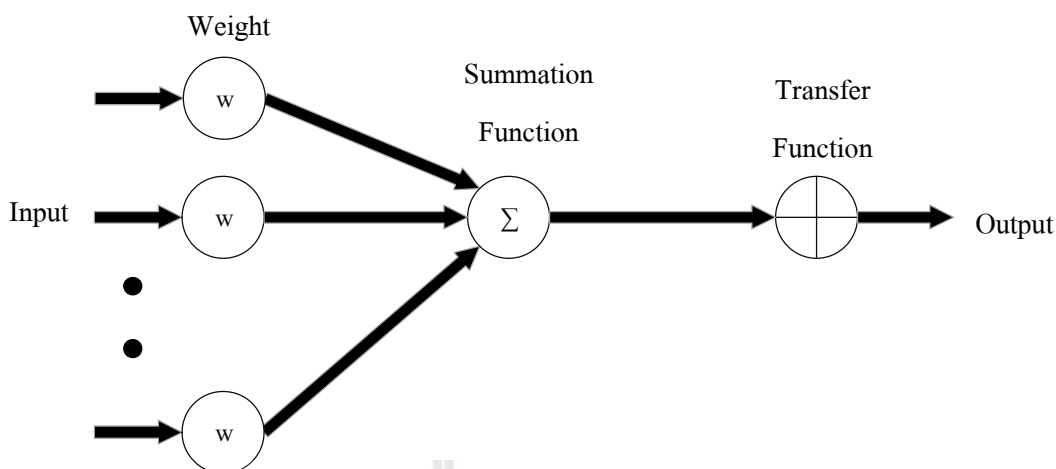
ลักษณะของเครือข่ายประสาทเทียมสามารถแบ่งได้ 2 แบบ คือ 1) เครือข่ายประสาทเทียมแบบชั้นเดียว (Single Layer) ซึ่งจะมีเพียงชั้นสัญญาณประสาทขาเข้า และชั้นสัญญาณประสาทขาออกเท่านั้น เช่น เครือข่ายเพอร์เซ็ปตรอนอย่างง่าย (Simple Perceptron) เป็นต้น และ 1) เครือข่ายประสาทเทียมแบบหลายชั้น (Multi-Layer) ซึ่งมีลักษณะเช่นเดียวกับเครือข่ายประสาทเทียมแบบชั้นเดียว แต่จะมีชั้นแอบแฝง (Hidden) เพิ่มขึ้น โดยอยู่ส่วนกลางระหว่างชั้นนำข้อมูลป้อนเข้าและชั้นส่งข้อมูลออก ทั้งนี้ชั้นแอบแฝงอาจมี 1 ชั้น อย่างไรก็ตาม การแบ่งเครือข่ายประสาทเทียมตามประเภทการเรียนรู้ของเครือข่าย สามารถแบ่งได้ 2 ประเภท คือ การเรียนรู้แบบมีผู้สอน (Supervised Learning) และการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) โดยในปัจจุบันการพัฒนาเครือข่ายประสาทเทียมยังคงมีการดำเนินการอย่างต่อเนื่องและคาดว่าจะมีบทบาทอย่างมากในด้าน

การจำแนกรูปแบบ การพยากรณ์ การควบคุม การหาความเหมาะสมและการจัดกลุ่มลักษณะ โดยทั่วไปของโครงสร้างระบบประสาทแสดงดังแสดงในรูปที่ 2.1



รูปที่ 2.1 โครงสร้างระบบประสาท

เครือข่ายประสาทเทียมมีคุณลักษณะคล้ายกับการส่งผ่านสัญญาณประสาทในสมองของมนุษย์ กล่าวคือมีความสามารถในการรวบรวมความรู้ (Knowledge) โดยผ่านกระบวนการเรียนรู้ (Learning Process) และความรู้เหล่านี้จะจัดเก็บอยู่ในรูปแบบค่าน้ำหนัก (Weight) ซึ่งสามารถปรับเปลี่ยนค่าได้เมื่อมีการเรียนรู้สิ่งใหม่ ๆ เข้าไป ค่าน้ำหนักทำหน้าที่เปรียบเสมือนความรู้ที่รวบรวมไว้เพื่อใช้ในการแก้ไขปัญหาเฉพาะอย่างของมนุษย์การประมวลผลต่าง ๆ เกิดขึ้นในหน่วยประมวลผลย่อยเรียกว่า โหนด (Node) ซึ่งโหนดเป็นการจำลองลักษณะการทำงานมาจากเซลล์การส่งสัญญาณ (Signal) ระหว่างโหนดที่เชื่อมต่อกัน (Connection) จำลองมาจากการเชื่อมต่อของเส้นใยประสาทและแกนประสาทในระบบประสาทของมนุษย์ภายในโหนดจะมีฟังก์ชันกำหนดสัญญาณส่งออกที่เรียกว่า ฟังก์ชันกระตุ้น (Activation Function) หรือฟังก์ชันการแปลง (Transfer Function) ซึ่งทำหน้าที่เปรียบเสมือนกระบวนการทำงานในเซลล์ดังแสดงในรูปที่ 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) ค่านี้จะถูกเก็บเป็นทักษะเพื่อใช้ในการจดจำข้อมูลอื่น ๆ ที่อยู่ในรูปแบบเดียวกัน
4. ฟังก์ชันผลรวม (Summation Function : S) เป็นผลรวมของข้อมูลป้อนเข้า (a_i) และค่าน้ำหนัก (w_i) ดังสมการที่ (2-1)

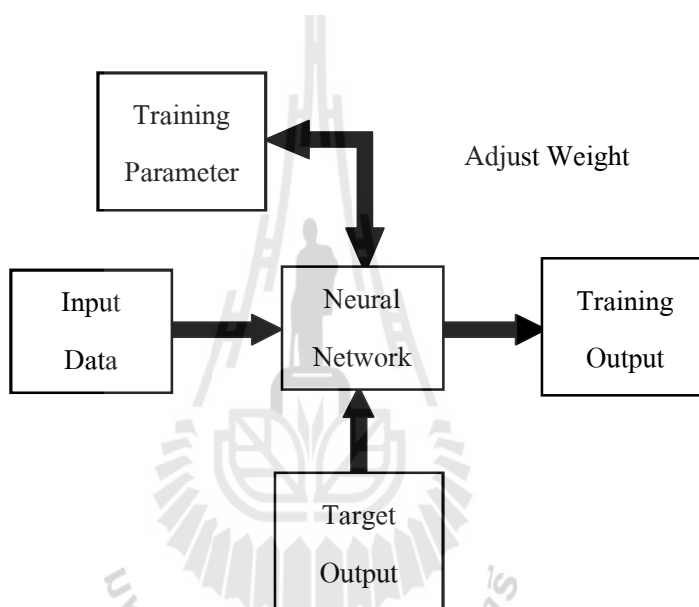
$$S = \sum_{i=1}^n a_i w_i \quad (2-1)$$

5. ฟังก์ชันการแปลง (Transfer Function) เป็นการคำนวณการจำลองการทำงานของเครือข่ายประสาทเทียม เช่น ซิกมอยด์ฟังก์ชัน (Sigmoid Function) ฟังก์ชันไฮเพอร์โบลิก-แทนเจนต์ (Hyperbolic Tangent Function) เป็นต้น

ประเภทของการเรียนรู้ของเครือข่ายประสาทเทียม

1. การเรียนรู้แบบมีผู้สอน (Supervised Learning)

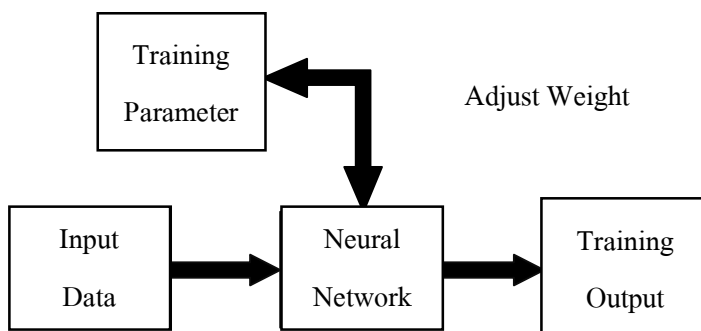
ข้อมูลจะประกอบด้วยตัวอย่างข้อมูลที่ต้องการสอน และผลลัพธ์ที่ต้องการให้เครือข่ายสร้าง เมื่อมีการนำข้อมูลในลักษณะเดียวกันมาเป็นข้อมูลป้อนเข้าเครือข่ายจะกำหนดผลลัพธ์ที่เป็นเป้าหมายให้กับข้อมูลป้อนเข้าแต่ละตัวเครือข่ายจะนำค่าผิดพลาดระหว่างค่าเป้าหมายกับค่าผลลัพธ์ที่ได้มาใช้ในการปรับค่าน้ำหนัก เพื่อให้ค่าผลลัพธ์ที่ใกล้เคียงกับเป้าหมายมากที่สุดดังแสดงในรูปที่ 2.3 ถ้าหากเปรียบเทียบกับมนุษย์จะเหมือนกับการสอนนักเรียน โดยมีครูผู้สอนคอยให้คำแนะนำ



รูปที่ 2.3 การเรียนรู้แบบมีผู้ฝึกสอน

2. การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning)

การเรียนรู้แบบนี้จะสอนเครือข่ายโดยการนำข้อมูลป้อนเข้าอย่างต่อเนื่องอย่างเดียว ไม่มีการส่งค่าผลลัพธ์เป้าหมายให้กับข้อมูลป้อนเข้าแต่ละตัว การปรับน้ำหนักจะใช้ข้อมูลที่นำมาสอนเป็นตัวปรับค่า โดยค่าน้ำหนักจะปรับตามกลุ่มข้อมูลอินพุตที่มีรูปแบบคล้ายคลึงกันดังแสดงในรูปที่ 2.4 ถ้าหากเปรียบเทียบกับมนุษย์จะเหมือนกับการที่เราสามารถแยกแยะ พันธุ์พืช พันธุ์สัตว์ ตามลักษณะรูปร่างของมันได้ด้วยตัวเอง



รูปที่ 2.4 การเรียนรู้แบบไม่มีผู้ฝึกสอน

2.3 ทฤษฎีที่เกี่ยวข้องกับการประมวลผลภาพดิจิทัล

การตรวจจับใบหน้าและการเคลื่อนไหวของดวงตานั้นจะนำหลักการประมวลผลภาพดิจิทัลมาใช้เช่นการแปลงค่าระหว่างปริภูมิสีการหาขอบวัตถุการหาพื้นที่ของวัตถุการหาจุดกึ่งกลางของวัตถุ เป็นต้น

วิดีโอ (Video) คือการเรียงต่อกันของภาพแบ่งออกเป็น 2 ประเภทใหญ่ๆ ได้แก่ วิดีโอแบบอะนาล็อก (Analog Video) และวิดีโอแบบดิจิทัล (Digital Video) ซึ่งวิดีโอแบบอะนาล็อกนี้จะเก็บข้อมูลภาพและเสียงในรูปแบบของสัญญาณไฟฟ้าวิดีโอแบบอะนาล็อกจะมีความไวต่อการผิดเพี้ยนสูงเมื่อมีการบันทึกต่อกันหลายๆ ครั้งจะทำให้คุณภาพของภาพและเสียงลดลง สำหรับวิดีโอแบบดิจิทัลนั้นจะเก็บข้อมูลอยู่ในรูปของไฟล์คอมพิวเตอร์ที่สามารถนำกลับมาใช้ใหม่ได้ทำสำเนาได้ และสามารถปรับแต่งแก้ไขได้

อัตราเฟรม (Frame Rate) คืออัตราความเร็วในการแสดงภาพอัตราในวิดีโอมีหน่วยเป็นเฟรมต่อวินาที (Frame Per Second : fps) เป็นหน่วยวัดปริมาณข้อมูลที่ใช้ในการเก็บบันทึกและแสดงวิดีโอ เช่น อัตราเฟรมของภาพยนตร์เท่ากับ 24 fps อัตราเฟรมระบบ PAL เท่ากับ 25 fps และอัตราเฟรมของ NTSC เท่ากับ 25 fps

ความละเอียด (Resolution) คือความชัดของภาพความละเอียดของภาพขึ้นอยู่กับจำนวนพิกเซลทั้งหมดเช่นความละเอียด 640×480 พิกเซลหมายถึงมีจำนวนพิกเซลแสดงผลเรียงกันเป็นแนวนอน 640 พิกเซลและแนวตั้ง 480 พิกเซล

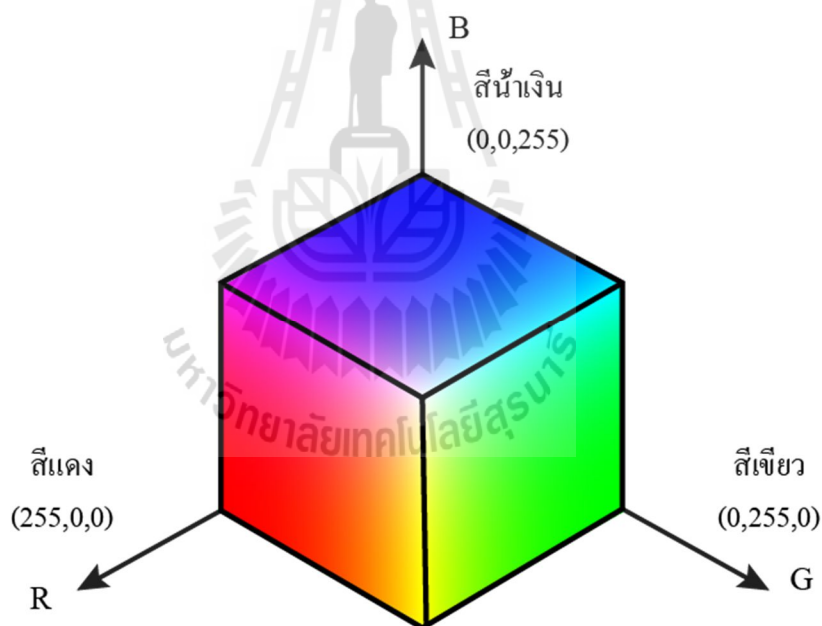
2.3.1 ปริภูมิสี (Color Space) RGB

ปริภูมิสี RGB ประกอบไปด้วย 3 องค์ประกอบได้แก่ ค่าสีแดง (Red : R) ค่าสีเขียว (Green : G) และค่าสีน้ำเงิน (Blue : B) แต่ละสีจะเป็นอิสระต่อกันและมีค่าอยู่ระหว่าง 0 – 255 เมื่อ

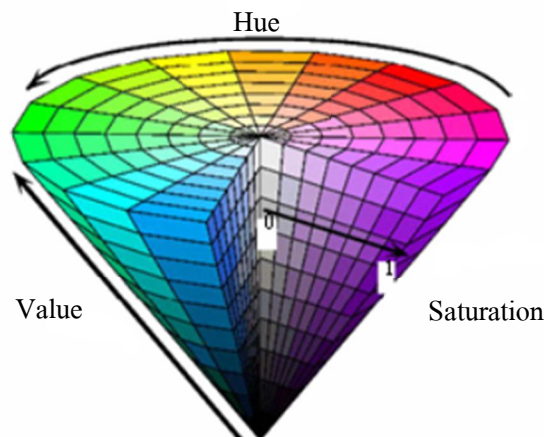
นำสีทั้งสามมารวมกันที่ความเข้มสูงสุดจะได้สีขาว ส่วนใหญ่ปริภูมิสีนี้จะใช้ในอุปกรณ์ที่เกี่ยวข้องกับแสง เช่น จอภาพ กล้องดิจิทัล สแกนเนอร์ ดังแสดงในรูปที่ 2.5

2.3.2 ปริภูมิสี HSV

ปริภูมิสี HSV เป็นปริภูมิสีที่ประกอบด้วย 3 องค์ประกอบได้แก่ค่า H (Hue) คือค่าโทนสีหรือค่าสีตันซึ่งมีค่าที่แตกต่างกันไปตามความถี่ของแสง ค่า S (Saturation) คือค่าความอิ่มตัวของสี ซึ่งเป็นค่าที่แสดงระดับสีเมื่อเทียบกับค่าโทนสี และค่า V (Value) คือค่าบอกระดับความสว่างของภาพ ซึ่งที่ระดับความสว่างต่ำสุดหมายถึงสีดำ ไม่ว่าจะมีย่านสีหรือค่าความอิ่มตัวสีเท่าใด และระดับความสว่างสูงสุดหมายถึงสีขาว ซึ่งเป็นสีที่สว่างที่สุดของค่าโทนสี และค่าความอิ่มตัวสีโดยภาพปริภูมิสี HSV สามารถหาได้จากการแปลงภาพปริภูมิสี RGB ด้วยสมการที่ (2-2) ถึงสมการที่ (2-5) และแสดงแบบจำลองภาพสี HSV ในรูปที่ 2.6



รูปที่ 2.5 ลูกบาศก์แบบจำลองภาพสี RGB



รูปที่ 2.6 กรวยของแบบจำลองภาพสี HSV

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases} \quad (2-2)$$

$$\theta = \cos^{-1} \left\{ \frac{0.5[(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)^2}} \right\} \quad (2-3)$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (2-4)$$

$$V = \frac{\max(R, G, B)}{255} \quad (2-5)$$

2.3.3 การแปลงภาพจากปริภูมิสี RGB เป็นภาพระดับสีเทา (Gray Scale)

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



รูปที่ 2.7 ค่าระดับสีเทา 0-255

ในการแปลงภาพปริภูมิสี RGB ให้เป็นภาพระดับสีเทานี้ สามารถคำนวณจากการแปลงค่าสีในภาพจากความสัมพันธ์ของการแปลงภาพสีเป็นภาพระดับสีเทาดังสมการที่ (2-6)

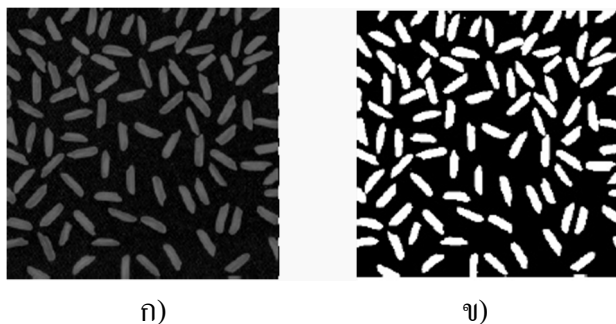
$$Grayscale_{(x,y)} = 0.299R_{(x,y)} + 0.587G_{(x,y)} + 0.114B_{(x,y)} \quad (2-6)$$

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

$$binaryimage = \begin{cases} 1 & ;if\ grayscale \geq threshold \\ 0 & else \end{cases} \quad (2-7)$$



รูปที่ 2.8 แสดงภาพในระดับสีเทาและภาพขาว-ดำ
ก) ภาพในระดับสีเทา ข) ภาพขาว – ดำ

2.3.5 การหาขอบภาพ (Edge Detection)

การหาขอบของภาพเป็นการหาเส้นรอบวัตถุที่อยู่ในภาพ เมื่อทราบเส้นรอบวัตถุ เราจะสามารถคำนวณหาพื้นที่ (ขนาด) หรือรูปร่างของวัตถุนั้นได้ อย่างไรก็ตาม การหาขอบภาพที่ถูกต้องสมบูรณ์ไม่ใช่เป็นเรื่องที่ง่าย โดยเฉพาะอย่างยิ่งการหาขอบของภาพที่มีคุณภาพต่ำ ขอบของภาพจะเห็นได้ชัด ถ้าความแตกต่างของความเข้มแสงจากจุดหนึ่งไปยังอีกจุดหนึ่งมีค่ามากขอบภาพก็จะเห็นได้ชัด ถ้าความแตกต่างมีค่าน้อย ขอบภาพก็จะไม่ชัดเจน

เทมเพลต (Template) สำหรับการหาขอบภาพ การหาขอบภาพในแนวนอนอย่างง่ายวิธีการก็คือหาผลต่างระหว่างจุดหนึ่งกับจุดที่อยู่ข้างล่าง (หรือข้างบน) ของจุดนั้น ดังนี้

$$Y_{diff}(x, y) = I(x, y) - I(x, y+1) \quad (2-8)$$

เมื่อ Y_{diff} คือ ค่าความแตกต่างในแนวแกนตั้ง และ $I(x, y)$ คือค่าความเข้มแสงของจุดภาพที่ตำแหน่ง (x, y)

ผลของการใช้สมการ (2-8) มีค่าเทียบเท่ากับการคอดโวลูชันภาพด้วยเทมเพลต

1
-1

รูปที่ 2.9 เทมเพลตสำหรับการหาขอบภาพในแนวนอน

การหาขอบภาพในแนวตั้งก็สามารถหาได้เช่นเดียวกันคือ

$$X_{diff}(x, y) = I(x, y) - I(x - 1, y) \quad (2-9)$$

เมื่อ X_{diff} คือ ค่าความแตกต่างในแนวนอน และสมการที่ (2-9) มีให้ผลเทียบเท่ากับการคอนโวลูชันด้วยเทมเพลต

-1	1
----	---

รูปที่ 2.10 เทมเพลตสำหรับการหาขอบภาพในแนวนอน

บางครั้งเราต้องการรวมผลต่างของค่าความแตกต่างในแนวแกนอนและแกนตั้งเข้าด้วยกัน เพื่อที่จะได้มีตัววัดความแรงของขอบภาพ (Gradient Magnitude) เพียงตัวเดียว เนื่องจากค่าความแตกต่างอาจมีค่าเป็นบวกหรือลบ ดังนั้น การบวกค่าความแตกต่างของทั้งสองแกนอาจทำให้ขอบภาพเกิดการหักล้างกันเอง ในทางปฏิบัติ เราจะต้องนำค่าสัมบูรณ์ (Absolute Value) หรือค่ากำลังสอง (Squared Value) ของค่าความแตกต่างของทั้งสองแกนมาบวกกันแทน

นอกจากหาความแรงของขอบภาพแล้ว การหาทิศทางของขอบภาพ (Gradient Direction) ก็มีประโยชน์เช่นกัน 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)

$Y_{diff}(x, y)$ คือ ค่าความแตกต่างในแนวแกนตั้ง (x, y)

$X_{diff}(x, y)$ คือ ค่าความแตกต่างในแนวนอน (x, y)

การหาขอบภาพโดยวิธีโซเบล การหาขอบภาพโดยวิธีโซเบล (Sobel Edge Detection) เป็นการหาขอบภาพโดยใช้เทมเพลตขนาด 3×3 สองเทมเพลต โดยเทมเพลตแรกจะใช้หาค่าความแตกต่างในแนวนอน (X_{diff}) และค่าความแตกต่างในแนวตั้ง (Y_{diff}) ดังแสดงในรูปที่

$$X_{diff} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad Y_{diff} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

รูปที่ 2.11 เเทมเพลตการหาของภาพโดยวิธีโซเบล



รูปที่ 2.12 ภาพผลลัพธ์ที่ได้จากการหาขอบภาพ

ก) ภาพต้นฉบับ

ข) ภาพการหาขอบด้วยวิธี Sobel Operator

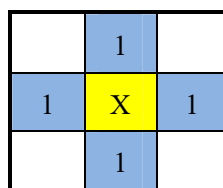
2.3.6 การขยายพิกเซล

การขยายพิกเซลของภาพเป็นขั้นตอนการดำเนินการทางตรรกะ โดยทำการค้นหา (Scan) ตำแหน่งบนซ้ายไปยังตำแหน่งล่างขวา จะแทนพิกเซล 1 เมื่อมีค่าของพิกเซลใดๆ พิกเซลหนึ่งบน SE (Structuring Element) ดังแสดงในรูปที่ 2.13 โดยมีสมการดังนี้

$$A \oplus B = \bigcup_{x \in B} (A_x) \quad (2-11)$$

เมื่อ B คือ Structuring Element

A คือ ภาพที่ต้องการขนาดพิกเซล



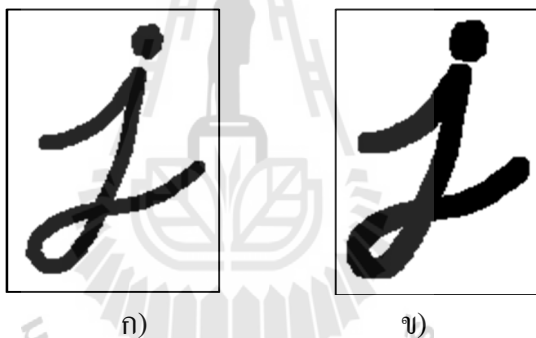
รูปที่ 2.13 เเทมเพลตการขยายพิกเซล

0	1	0	1	0	0	0			
1	X	1	1	0	0	0			
0	1	1	1	0	0	0			
0	0	1	1	0	0	0			
0	0	0	1	0	0	0			
0	0	0	1	0	0	0			
0	0	0	0	0	0	0			
0	0	0	0	0	0	0			

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

รูปที่ 2.14 ภาพจำลองการขยายพิกเซล

เมื่อทำการขยายพิกเซลโดยใช้การขยายพิกเซลจะได้ภาพดังแสดงในรูปที่ 2.15



รูปที่ 2.15 ภาพจำลองการขยายพิกเซล

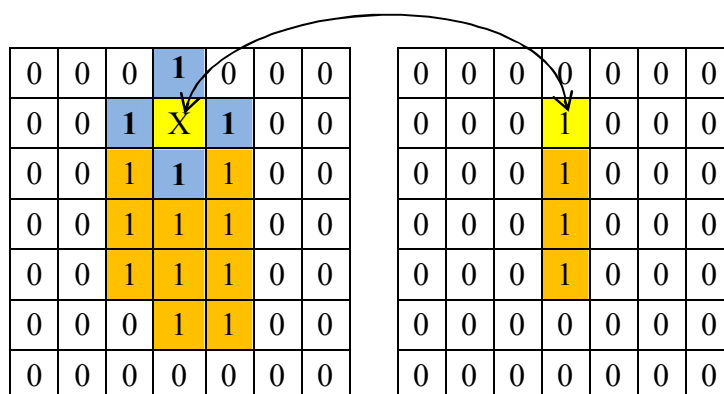
ก) ภาพต้นฉบับ

ข) ภาพหลังการขยายพิกเซล

2.3.7 การกร่อนพิกเซล

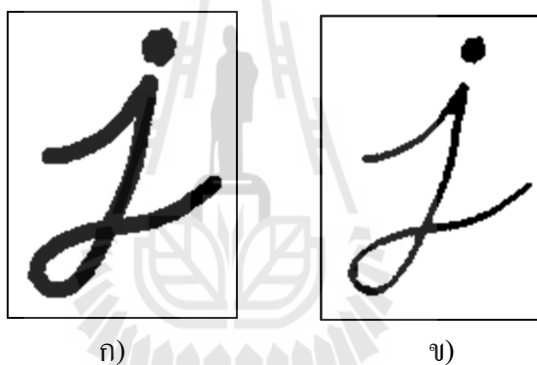
การกร่อนพิกเซลเป็นวิธีการที่ตรงข้ามกับการขยายพิกเซลคือการลดขนาดของพิกเซล โดยทำการค้นหาจากตำแหน่งบนซ้ายไปยังตำแหน่งล่างขวา จะแทนพิกเซล 0 เมื่อมีค่าของพิกเซลบน SE ทั้งหมด ดังแสดงในรูปที่ 2.16 โดยมีสมการดังนี้

$$A \ominus B = \{w: B_w \subset A\} \quad (2-12)$$



รูปที่ 2.16 ภาพจำลองการกร่อนฟิกเซล

เมื่อทำการกร่อนฟิกเซลโดยใช้การกร่อนฟิกเซลจะได้ภาพดังแสดงในรูปที่ 2.17



รูปที่ 2.17 ภาพจำลองการกร่อนฟิกเซล

ก) ภาพต้นฉบับ

ข) ภาพหลังการกร่อนฟิกเซล

2.3.8 การจับคู่กับแม่แบบ (Template Matching)

การจับคู่กับแม่แบบเป็นการนำภาพแม่แบบไปวางทับกับภาพที่พิจารณาไปยังจุดต่าง ๆ จากซ้ายไปขวา บนลงล่าง ทำการประมวลผลโดยใช้ค่าของฟิกเซลของภาพและแม่แบบที่อยู่ในตำแหน่งเดียวกัน เพื่อคำนวณหาค่าความเหมือนของแม่แบบกับบริเวณต่าง ๆ บนภาพซึ่งสำหรับใน Opencv 2.4.9.0 Library ฟังก์ชันในการคำนวณทั้งหมด 6 ฟังก์ชันดังสมการที่ (2-13) ถึง (2-18)

- CV_TM_SQDIFF

$$R(x, y) = \sum_{x', y'} (T(x', y') - I(x + x', y + y'))^2 \quad (2-13)$$

- CV_TM_SQDIFF_NORMED

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') - I(x + x', y + y'))^2}{\sqrt{\sum_{x', y'} T(x', y')^2 \cdot \sum_{x', y'} I(x + x', y + y')^2}} \quad (2-14)$$

- CV_TM_CCORR

$$R(x, y) = \sum_{x', y'} (T(x', y') \cdot I(x + x', y + y')) \quad (2-15)$$

- CV_TM_CCORR_NORMED

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') \cdot I(x + x', y + y'))}{\sqrt{\sum_{x', y'} T(x', y')^2 \cdot \sum_{x', y'} I(x + x', y + y')^2}} \quad (2-16)$$

- CV_TM_CCOEFF

$$R(x, y) = \sum_{x', y'} (T'(x', y') \cdot I'(x + x', y + y')) \quad (2-17)$$

- CV_TM_CCOEFF_NORMED

$$R(x, y) = \frac{\sum_{x', y'} (T'(x', y') \cdot I'(x + x', y + y'))}{\sqrt{\sum_{x', y'} T'(x', y')^2 \cdot \sum_{x', y'} I'(x + x', y + y')^2}} \quad (2-18)$$

โดยที่ T' และ I' หาได้จากสมการที่ (2-19) และ (2-20)

$$T'(x', y') = T(x', y') - 1/(w \cdot h) \cdot \sum_{x'', y''} T(x'', y'') \quad (2-19)$$

$$I'(x+x', y+y') = I(x+x', y+y') - 1/(w \cdot h) \cdot \sum_{x'', y''} I(x+x'', y+y'') \quad (2-20)$$

เมื่อ R	คือ ค่าความเหมือนของแม่แบบกับบริเวณต่าง ๆ บนภาพ
T	คือ ภาพแม่แบบ
I	คือ ภาพที่พิจารณา
w	คือ ความกว้างของแม่แบบ
h	คือ ความสูงของแม่แบบ
x', y'	คือ ตำแหน่งของพิกเซลบนภาพแม่แบบ
x, y	คือ ตำแหน่งของพิกเซลบนภาพที่พิจารณา

2.3.9 Retinex [2]

Retinex มาจากคำว่า Retina รวมกับ Cortex เป็นเทคนิคที่ช่วยในการปรับแสงสว่างในภาพ ในขณะที่ตาของเราสามารถมองเห็นสีได้อย่างถูกต้องถึงแม้มีแสงสว่างต่ำ แต่กล้องหรือกล้องวิดีโอที่ไม่สามารถทำได้ Retinex สามารถคำนวณได้จากภาพขาเข้า

Single-Scale Retinex (SSR) เป็นการปรับความสว่างของภาพ ได้จากค่าแต่ละพิกเซลกับพิกเซลข้างเคียงเรียกว่าฟังก์ชันเกาส์เซียน (Gaussian Function) สามารถคำนวณได้จากสมการที่ (2-21)

$$R(x, y) = \log(I(x, y)) - \log(I(x, y) * F(x, y)) \quad (2-21)$$

เมื่อ I	คือ ภาพขาเข้า
R	คือ ภาพขาออกของ Single Scale Retinex
\log	คือ ฟังก์ชันลอการิทึมธรรมชาติ
(x, y)	คือ ตำแหน่งของพิกเซล

- * คือ ตัวดำเนินการคอนโวลูชัน (Convolution Operator)
F คือ ฟังก์ชันเซอร์ราวด์ (Surround Function) โดยคำนวณได้ตามสมการที่ (2-22) ดังนี้

$$F(x, y) = K \exp[-(x^2 + y^2)/C^2] \quad (2-22)$$

- เมื่อ *C* คือ ค่าเฉลี่ยมาตรฐานของเซอร์ราวด์ในฟังก์ชันแบบเกาส์เซียน (Gaussian Surround)
(x, y) คือ พิกัดของพิกเซล
K คือ ค่านอร์มอลไลซ์ (Normalized) โดยคำนวณได้ตามสมการที่ (2-23) ดังนี้

$$K = \frac{1}{\sum_x \sum_y F(x, y)} \quad (2-23)$$

- Mutis-Scale Retinex (MSR) เป็นการปรับความสว่างของภาพที่มีการพัฒนาจาก Single Scale Retinex ที่มีจำนวนของช่องสี (Channel) เพิ่มขึ้น โดยคำนวณได้ตามสมการที่ (2-24) ดังนี้

$$R_{MSRi}(x, y) = \sum_{k=1}^K W_k \{ \log(x, y) - \log[F_k(x, y) * I_k(x, y)] \} \quad (2-24)$$

โดยที่ $i = 1, \dots, N$

- เมื่อ *I* คือ ภาพขาเข้า
R_{MSR} คือ ภาพขาออกของ Multi-Scale Retinex
i คือ ลำดับของช่องสี
(x, y) คือ ตำแหน่งของพิกเซล
 * คือ ตัวดำเนินการคอนโวลูชัน (Convolution Operator)
N คือ ช่องสี (ภาพระดับสีเทา $N=1$, ภาพสี RGB $N=3$)
 log คือ ฟังก์ชันลอการิทึมธรรมชาติ
K คือ ค่านอร์มอลไลซ์หรือสเกล
W_k คือ ค่าน้ำหนักที่เกี่ยวข้องกับ *F_k*
F_k คือ Surround Function โดยคำนวณได้ตามสมการ (2-25) ดังนี้

$$F_i(x, y) = K \cdot \exp\left[-(x^2 + y^2) / \sigma_k^2\right] \quad (2-25)$$

เมื่อ σ_k คือ ค่าเฉลี่ยมาตรฐานของ Gaussian Surrounds
 K คือ ค่านอร์มอลไลซ์ โดยคำนวณได้ตามสมการที่ (2-26) ดังนี้

$$K = \frac{1}{\sum_x \sum_y F(x, y)} \quad (2-26)$$

ในกรณีที่เป็น Single Scale Retinex จะกำหนดให้ $K = 1$ และ $W_1 = 1$

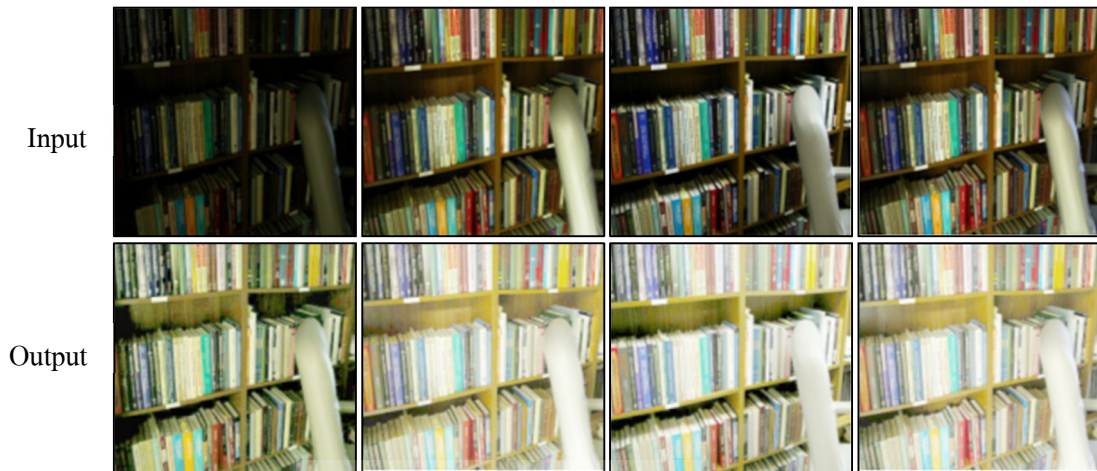
MSR with Color Restoration (MSRCR) เป็นการปรับความสว่างของภาพที่มีการพัฒนาจาก Multi-Scale Retinex ที่มีพารามิเตอร์ (Parameter) ในการปรับปรุงสีในภาพเพิ่มเข้ามา โดยสามารถปรับปรุงสีให้เหมาะสมยิ่งขึ้น โดยคำนวณได้ตามสมการที่ (2-27) ถึง สมการ (2-28)

$$C_i(x, y) = \beta \left\{ \log[\alpha I_i(x, y)] - \log \left[\sum_{i=1}^N I_i(x, y) \right] \right\} \quad (2-27)$$

$$R_{MSRCR}(x, y) = G \left[C_i(x, y) * \{ \log I_i(x, y) - \log [F_k(x, y) * I_i(x, y)] \} + b \right] \quad (2-28)$$

โดยที่ $i = 1, \dots, N$

เมื่อ I คือ ภาพขาเข้า
 R_{MSRCR} คือ ภาพขาออกของ Multi-Scale Retinex
 i คือ ลำดับของช่องสี
 (x, y) คือ ตำแหน่งของพิกเซล
 $*$ คือ ตัวดำเนินการคอนโวลูชัน (Convolution Operator)
 N คือ ช่องสี (ภาพระดับสีเทา $N=1$, ภาพสี RGB $N=3$)
 \log คือ ฟังก์ชันลอการิทึมธรรมชาติ
 F_k คือ Surround Function
 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

จากรูปที่ 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 โดยใช้พารามิเตอร์ตามบทความวิจัยที่ได้นำเสนอ โดย D. J. Jobson [2] ดังนี้

$$W_k = 1/3$$

$$N = 3$$

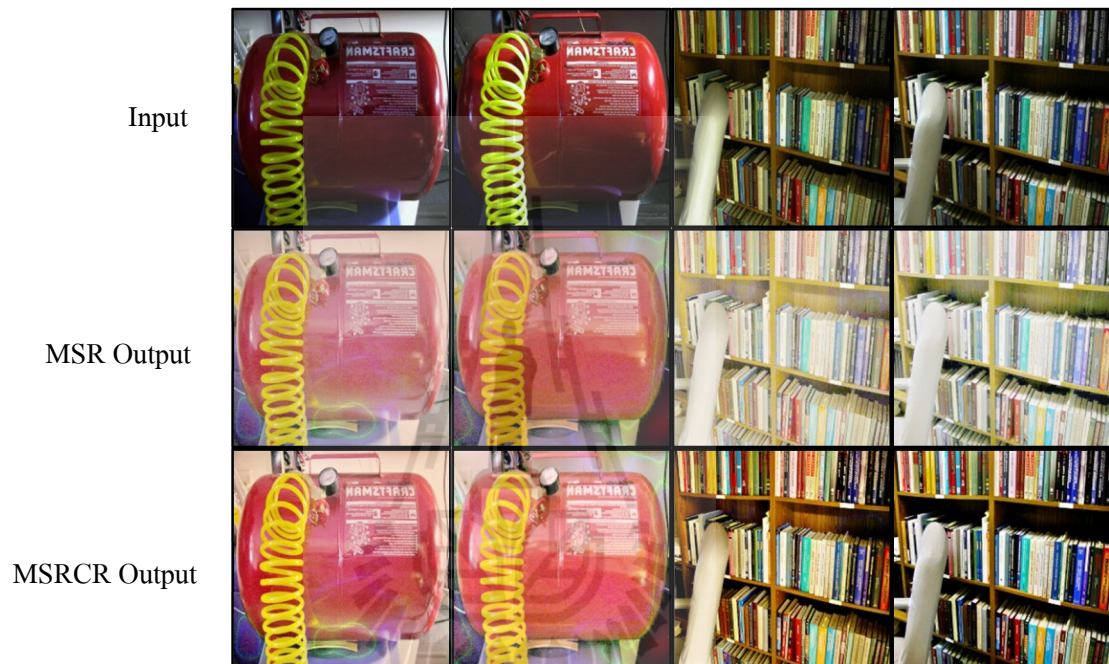
$$\sigma_1, \sigma_2, \sigma_3 = 15, 80, 250$$

$$G = 192$$

$$b = 30$$

$$\alpha = 125$$

$$\beta = 46$$



รูปที่ 2.20 การปรับปรุงโดยใช้เทคนิค MSR เทียบกับ MSRCR

2.4 การทบทวนวรรณกรรม/สารสนเทศ (Information) ที่เกี่ยวข้อง

การทบทวนวรรณกรรมและสารสนเทศที่เกี่ยวข้องในฐานข้อมูลต่าง ๆ สามารถสรุปเป็นตารางได้ดังต่อไปนี้

ตารางที่ 2.1 งานวิจัยที่เกี่ยวข้องด้านการตรวจจับอาการ่วงนอนขณะขับรถ

ปีที่พิมพ์ /ลำดับเอกสารอ้างอิง	คณะผู้วิจัย	องค์ความรู้ที่ได้จากบทความ
2004 [4]	Q. Ji Z. Zhu P. Lan	ในบทความนี้ได้นำเสนอการตรวจจับอาการ่วงนอนโดยใช้การสะท้อนของดวงตาจากหลอด LED ที่ติดอยู่กับกล้องเพื่อระบุตำแหน่งของดวงตาซึ่งสามารถใช้ได้ในเวลาที่ไม่มีแสงสว่างหรือเวลากลางคืน
2006 [1]	Q. Wang J. Yang M. Ren Y. Zheng	ในบทความนี้ได้นำเสนอการตรวจจับอาการ่วงนอนแบบต่างๆ เช่นการตรวจจับอาการ่วงนอนโดยใช้การสังเกตจากการเคลื่อนไหวของดวงตา อัตราการกระพริบตา ตำแหน่งของศีรษะ ลักษณะของปาก การตรวจจับโดยใช้ Active IR การตรวจจับโดยสังเกตการติดตามช่องทางเดินรถ
2009 [5]	P. R. Tabrizi R. A. Zoroofi	ในบทความนี้ได้นำเสนอการตรวจจับอาการ่วงนอนโดยการวิเคราะห์จากอัตราการกระพริบตาโดยหลังจากการตรวจจับใบหน้าแล้วจะทำการตรวจจับดวงตาในปริภูมิสี YCbCr และใช้ระยะห่างระหว่างขอบตากับขอบตาล่างบอกสถานะของตา
2010 [6]	S. Ribaric J. Lovrencic N. Pavesic	ในบทความนี้ได้นำเสนอการตรวจจับอาการ่วงนอนโดยสังเกตจากมุมการเอียงของศีรษะ โดยการนำเทคนิคทางปัญญาประดิษฐ์ นั่นคือเครือข่ายประสาทเทียมมาใช้ในการหามุมของศีรษะว่าอยู่ในระนาบ (In-Plane) หรือ ออกนอกระนาบ (Out-of-Plane)
2011 [7]	R. C. Coetzer G. P. Hancke	ในบทความนี้ได้นำเสนอการตรวจจับอาการ่วงนอนโดยการสังเกตจากการค้นหาดวงตา ว่ามีดวงตาอยู่หรือไม่โดยใช้ภาพของตา และไม่ใช้ตามาใช้ในการรู้จำของเครือข่ายประสาทเทียมและ นำภาพทดสอบเป็นดวงตาหรือไม่

ตารางที่ 2.1 งานวิจัยที่เกี่ยวข้องด้านการตรวจจับอาการง่วงนอนขณะขับรถ (ต่อ)

2012 [8]	X. Li E. Seignez P. Loonis	ในบทความนี้ได้นำเสนอการตรวจจับอาการง่วงนอนโดยนำวิธีการ 2 วิธีการมารวมกันได้แก่ การติดตาม (Tracking) ช่องทางเดินรถบนถนน และสถานะของใบหน้า ตัวอย่างเช่น มุมเอียงของใบหน้า ซึ่งเป็นเพิ่มความแม่นยำในการตรวจจับเพิ่มขึ้น
2012 [9]	I. Garcia S. Bronte L. M. Bergasa J. Almazan J. Yebes	ในบทความนี้ได้นำเสนอการตรวจจับอาการง่วงนอนจากการเปรียบเทียบอัตราการกระพริบตาโดยจะหาดวงตาจากการใช้อัลกอริทึม Integral Projection ทั้งแนวนอน (Horizontal) และแนวตั้ง (Vertical) เพื่อหาสถานะของตาโดยสังเกตจากรอยละของการลืมตา ในบทความนี้ยังได้ทดสอบกับสถานะการใช้งานจริงอีกด้วย

ตารางที่ 2.2 งานวิจัยที่เกี่ยวข้องด้านการตรวจจับใบหน้า

ปีที่พิมพ์ /ลำดับเอกสารอ้างอิง	คณะผู้วิจัย	องค์ความรู้ที่ได้จากบทความ
2011 [11]	P. P. Paul M. Gavilova	ในบทความนี้ได้นำเสนอการตรวจจับใบหน้าโดยใช้การวิเคราะห์ส่วนประกอบสำคัญการหาขอบวัตถุแบบ Canny และการคัดแยกสีผิวทำให้ประสิทธิภาพในการตรวจจับใบหน้าที่มีความถูกต้องสูง
2011 [10]	K. S. Park R. H. Park Y. G. Kim	ในบทความนี้ได้นำเสนอการตรวจจับใบหน้าโดยใช้บล็อก 3×3 เป็นเทมเพลตในการหาใบหน้าโดยเทมเพลตนี้จะหาความเข้มในแต่ละบล็อกเพื่อนำไปเปรียบเทียบกับภาพใบหน้าที่ต้องการทดสอบโดยลำดับความเข้มในบล็อก 3×3 ของภาพเทมเพลตและภาพใบหน้าที่นำมาทดสอบตรงกันจะถือว่าตำแหน่งนั้นเป็นใบหน้า

ตารางที่ 2.2 งานวิจัยที่เกี่ยวข้องด้านการตรวจจับใบหน้า (ต่อ)

2011 [12]	M. L. Chiang S. H. Lau	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยใช้การหาขอบของวัตถุและเปรียบเทียบกับ การตรวจจับใบหน้าโดยใช้การหาขอบของวัตถุแบบ Haar และ Canny ซึ่งในบทความนี้สามารถตรวจจับใบหน้าได้ ถึงแม้ว่าจะมีแสง เชื้อชาติ ขนาด และการเอียงของ ใบหน้า ที่แตกต่างกัน
--------------	---------------------------	---

ตารางที่ 2.3 งานวิจัยที่เกี่ยวข้องด้านการตรวจจับดวงตา

ปีที่พิมพ์ /ลำดับเอกสารอ้างอิง	คณะผู้วิจัย	องค์ความรู้ที่ได้จากบทความ
2007 [13]	C. W. Park J. M. Kwak H. Park Y. S. Moon	ในบทความนี้ได้นำเสนอการค้นหาดวงตาโดยใช้ตัวกรองกาบอร์ (Gabor Filter) เพื่อหาเส้นแนวนอนของตาและใช้ตัวกรอง ART (Angular Radial Transform) เพื่อหาลักษณะของตาที่เป็นวงกลมและรวมทั้งสองจะคำนวณหาตำแหน่งของตาออกมา
2011 [17]	S. Chen C. Liu	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยใช้ปริภูมิสี YUV แล้วใช้สมการ Projection Functions เพื่อค้นหาบริเวณของตา แล้วเลือกใช้การหาความเข้มในบริเวณนั้นเพื่อหาตำแหน่งของตาโดยแปลงภาพเป็นระดับสีเทาโดยที่ดวงตาจะมีค่าความเข้มที่ต่ำ วิธีการนี้เป็นวิธีการตรวจจับดวงตาที่มีความเร็ว สามารถตรวจจับดวงตาถูกต้องถึงร้อยละ 95.4
2012 [21]	Md. H. Rahman F. Jhumur Md. S. U. Yusuf T. Das M. Ahmsad	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยใช้ลักษณะของตาทั้งของข้างกับปากซึ่งเป็นรูปสามเหลี่ยม ทำการแปลงภาพให้อยู่ในปริภูมิสี YCbCr จากนั้นจะทำการหาใบหน้าโดยใช้การคัดแยกสีผิว จะพบว่าจะมีช่องว่างสามจุดของตาและปากมีลักษณะเป็นรูปสามเหลี่ยม

ตารางที่ 2.3 งานวิจัยที่เกี่ยวข้องด้านการตรวจจับดวงตา (ต่อ)

2010 [15]	N. Alioua A. Amine M. Rziza D. Aboutajdine	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยใช้ Circular Hough Transform (CHT) ซึ่งเป็นอัลกอริทึมที่มีความใหม่ที่ไม่ต้องใช้ภาพในการเรียนรู้ เมื่อนำอัลกอริทึมนี้มาเปรียบเทียบการหาขอบวัตถุพบว่ามี ความถูกต้องสูงกว่า
2011 [18]	M. A. A. Kashani M. M. Arani	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยอาศัยลักษณะของดวงตา เช่น คิ้ว หน้ตา เป็นต้น หลังจากตรวจจับใบหน้าโดยใช้สีผิวจะทำการหา ดวงตาโดยปกติแล้วตาจะมีลักษณะเป็นเส้นแนวนอน บทความนี้จะนำลักษณะนี้มาใช้หาดวงตา
2012 [20]	P. Goel S. Agarwal	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาที่นำ สองอัลกอริทึมรวมกัน ได้แก่ Haar Cascade Classifiers และคุณสมบัติต่าง ๆ บนใบหน้าเพื่อ ให้การตรวจจับมีประสิทธิภาพมากขึ้นเห็นได้จากผล การทดสอบที่มีการตรวจจับสูงถึงร้อยละ 98.75
2011 [16]	D. Y. Huang T. W Lin W. C. Hu M. S. Chen	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยการ ใช้สีในการคัดแยกสามารถตรวจจับได้ในแสงสว่างที่ เปลี่ยนแปลง สามารถหาได้จากการตรวจจับใบหน้า โดยการตรวจจับใบหน้าจะมีช่องว่างและระบุ ช่องว่างนี้เป็นดวงตา
2012 [19]	C. Phromsuthirak S. Umchid	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยใช้ รูปร่างลักษณะของตาและสีในการคัดแยก จะใช้ สมการ Eyemapc ที่ได้จากปริภูมิสี Cb และสมการ Eyemapl ที่ได้จากปริภูมิสี Cr และใช้ลักษณะของ ความกว้างและความสูงในการระบุตำแหน่งของตา
2009 [14]	N. V. Huan N. T. H. Binh H. Kim	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยตัว กรองวงกลม (Circular Filter) ปกติตาจะมีลักษณะ เป็น วงกลมหรือเป็นส่วนหนึ่งของวงกลม จึงได้นำ ลักษณะของตาคำนี้นำใช้ในการระบุตำแหน่งของตา

ตารางที่ 2.4 งานวิจัยที่เกี่ยวข้องกับการตรวจจับอาการง่วงนอนโดยใช้เทคนิคทางปัญญาประดิษฐ์

ปีที่พิมพ์ /ลำดับเอกสารอ้างอิง	คณะผู้วิจัย	องค์ความรู้ที่ได้จากบทความ
2005 [22]	F. H. C. Tivive A. Bouzerdoum	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาโดยการ ใช้เครือข่ายประสาทเทียมระบบสามารถแยกภาพที่มี ตา และไม่มีตา ในการใช้เครือข่ายประสาทเทียมใน บทความนี้สามารถบอกความถูกต้องในการคัดแยก ดวงตาได้ถึงร้อยละ 99
2007 [23]	T. Akashi Y. Wkasa K. Tanaka S. Karungaru M. Fukumi	ในบทความนี้ได้นำเสนอการตรวจจับดวงตาและ ติดตามดวงตาโดยใช้เทคนิคจินเนติกอัลกอริทึม (Genetic Algorithm) เป็นการติดตามดวงตาในภาพ วิดีโอมีความถูกต้องสูงถึงร้อยละ 97.9 และใช้เวลา ในการตรวจจับดวงตาโดยเฉลี่ย 28 มิลลิวินาทีต่อ เฟรม
2011 [7]	R. C. Coetzer G. P. Hancke	ในบทความนี้ได้นำเสนอการตรวจจับอาการง่วง นอนแบบเวลาจริงโดยใช้เทคนิคทางปัญญาประดิษฐ์ ได้แก่ เครือข่ายประสาทเทียม, เครื่องเวกเตอร์ เกือหนุน (Support Vector Machines: SVM) และ วิธีการค้นหาแบบตามูเชิงปรับตัว (Adaptive Boosting: AdaBoost)

บทที่ 3

การออกแบบระบบเฝ้าติดตามอาการง่วงนอน

3.1 บทนำ

เนื้อหาในบทนี้กล่าวถึงการออกแบบระบบเฝ้าติดตามอาการง่วงนอน โดยได้แบ่งเป็นหัวข้อ ดังนี้ การตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม การตรวจจับใบหน้าโดยใช้การคัดแยกสีผิว การตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน การตรวจจับดวงตาโดยใช้เทคนิค Integral Projection และในส่วนสุดท้ายจะเป็นบทสรุป

3.2 การออกแบบระบบตรวจจับใบหน้า

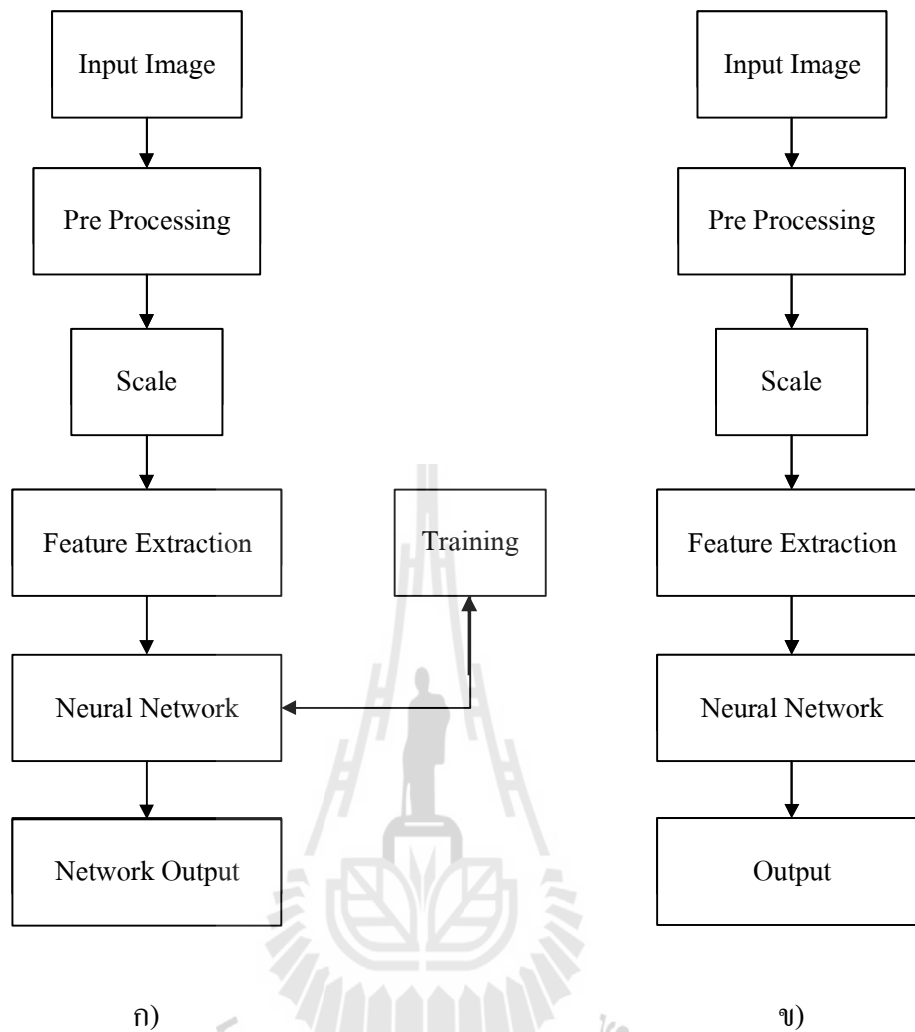
การตรวจจับอาการง่วงนอนโดยอาศัยการสังเกตอาการของผู้ขับขี่รถยนต์ ในงานวิจัยนี้ได้ใช้ดวงตาเป็นตัวบ่งชี้ถึงระดับอาการง่วงนอนของผู้ขับขี่รถยนต์ ในการค้นหาดวงตาในเฟรมภาพซึ่งมีขนาดใหญ่ขึ้นใช้เวลาในการประมวลผลมาก เพื่อลดเวลาในการค้นหาดวงตาลง ต้องค้นหาบริเวณของใบหน้าให้ได้เสียก่อน การตรวจจับใบหน้า ในบทนี้ได้ศึกษาวิธีการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม และการตรวจจับใบหน้าโดยใช้สีผิว

การตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมเป็นการนำภาพใบหน้าตัวอย่างมาให้เครือข่ายประสาทเทียมรู้จัก โดยภาพใบหน้าที่นำมาใช้ในการรู้จักมีการดึงลักษณะเด่นด้วยการวิเคราะห์ส่วนประกอบสำคัญ (Principal Components Analysis: PCA) และการแปลงเวฟเล็ตแบบดิสครีท (Discrete Wavelet Transform: DWT)

การตรวจจับใบหน้าโดยใช้การคัดแยกสีผิวเป็นการหาช่วงของสีผิวที่อยู่ในปริภูมิสี RGB และ HSV แล้วนำเทคนิคการประมวลผลภาพดิจิทัลมาใช้ประมวลผล เช่น การหาขอบภาพ (Edge Detection) การขยายพิกเซล (Dilation) การกร่อนพิกเซล (Erosion)

3.2.1 การออกแบบการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม

ในหัวข้อนี้เป็นการออกแบบการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม เป็นการนำภาพใบหน้ามาให้เครือข่ายประสาทเทียมรู้จัก โดยจะทำการดึงลักษณะเด่นจากภาพใบหน้าด้วยการวิเคราะห์ส่วนประกอบสำคัญ และการแปลงเวฟเล็ตแบบดิสครีท โดยมีแผนภาพการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมแสดงในรูปที่ 3.1



รูปที่ 3.1 ระบบตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม

ก) การรู้จำของเครือข่ายประสาทเทียม

ข) การทดสอบโดยการใช้เครือข่ายประสาทเทียม

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



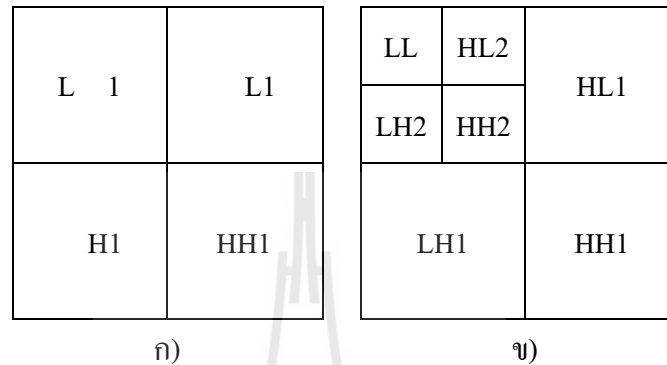
รูปที่ 3.2 ตัวอย่างภาพใบหน้าขนาด 200×200 พิกเซล



รูปที่ 3.3 ตัวอย่างภาพที่ไม่ใช่ภาพใบหน้าขนาด 200×200 พิกเซล

การดึงลักษณะเด่นโดยการแปลงเวฟเล็ตแบบดิสครีท

การแปลงเวฟเล็ตแบบดิสครีทเป็นการแปลงสัญญาณให้เป็นอนุกรมของเวฟเล็ต โดยเซตของผลลัพธ์จะเรียกว่าสัมประสิทธิ์โดยการแปลงเวฟเล็ตแบบดิสครีทระดับที่ 1 และระดับที่ 2 แสดงในรูปที่ 3.4

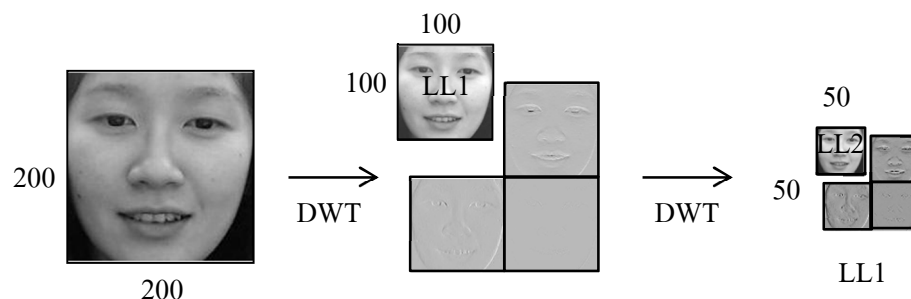


รูปที่ 3.4 การแปลงเวฟเล็ตแบบดิสครีท

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) เป็นวิธีการทางสถิติที่ใช้สร้างเมทริกซ์ของความแปรปรวนร่วม (Covariance Matrix) จากข้อมูลภาพ ถูกนำไปใช้ในการบีบอัดข้อมูล การวิเคราะห์องค์ประกอบหลักมาใช้ลดขนาดข้อมูล ข้อมูลของแพ้มกเกิดขึ้นอยู่ในรูปของเวกเตอร์ 1 มิติอยู่แล้ว นำเวกเตอร์ของทุกชุดข้อมูลมาจัดให้อยู่ในรูปแบบของเมทริกซ์ โดยเวกเตอร์ของชุดข้อมูลที่ 1 จะเป็นแถวที่ 1 ของเมทริกซ์ เวกเตอร์ของชุดข้อมูลที่ 2 จะเป็นแถวที่ 2 ของเมทริกซ์จนถึงเวกเตอร์ของชุดข้อมูลที่ n จะเป็นแถวที่ n ของเมทริกซ์ ดังนั้นจะได้เมทริกซ์ A มีมิติเป็น i และ j โดยที่ i หมายถึงข้อมูลชุดที่ i และ j หมายถึงมิติที่ j ดังแสดงในรูปแบบที่ 3.6

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

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)

$$\lambda = SS^T \quad (3-4)$$

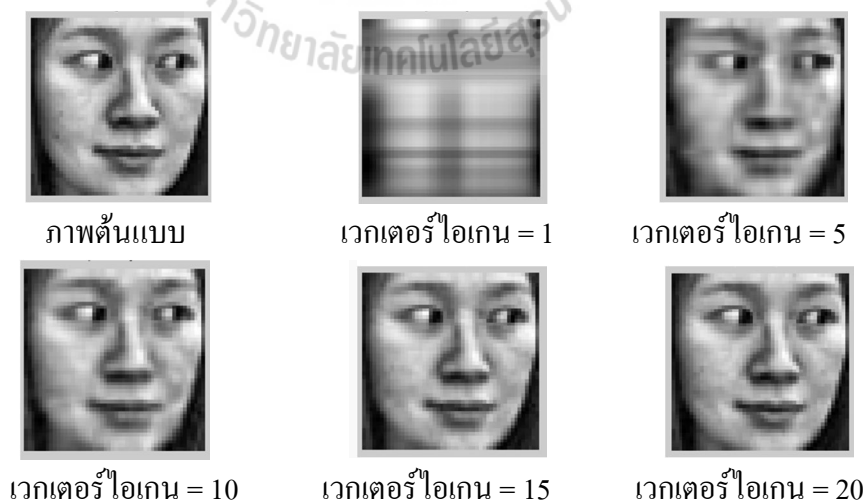
คำนวณเวกเตอร์ไอเกน (Eigen Vectors) ดังสมการที่ (3-5)

$$e = \frac{CS}{\sqrt{\lambda}} \quad (3-5)$$

ผลลัพธ์จะได้ค่าไอเกน และเวกเตอร์ไอเกน ข้อมูลทั้ง 2 มีความสัมพันธ์กัน (Correspondence) ซึ่งกันและกัน เวกเตอร์ไอเกนเป็นข้อมูลที่เปลี่ยนรูปแล้ว ไม่สามารถเห็นเป็นเวกเตอร์แพกเก็ตแบบเดิมได้ การนำข้อมูลเดิมกลับมาต้องนำเวกเตอร์ไอเกนมาคำนวณอีกครั้ง ดังสมการที่ (3-5) เมื่อไอเกนแพกเก็ตแรก (1st Eigen Package) คือการเลือกเวกเตอร์ไอเกนแรกของทุกเวกเตอร์ข้อมูลมาคำนวณ ส่วนไอเกนแพกเก็ตอื่น ๆ ก็ทำในรูปแบบเดียวกัน

การเลือกเวกเตอร์ไอเกน

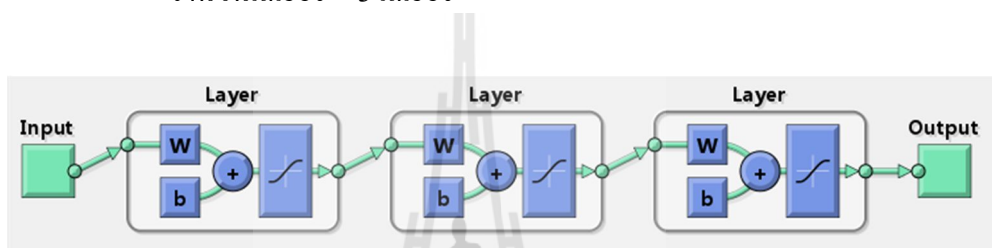
นำภาพหลังจากการแปลงเวฟเล็ดแบบดีสครีทที่มีขนาด 50×50 พิกเซลมาดึงลักษณะเด่นด้วยการวิเคราะห์ส่วนประกอบสำคัญ ในการเลือกค่าเวกเตอร์ไอเกนของการวิเคราะห์ส่วนประกอบสำคัญจะทดลองที่ค่าต่าง ๆ ตามรูปที่ 3.7



รูปที่ 3.7 การทดลองเลือกค่าเวกเตอร์ไอเกนที่ค่าต่าง ๆ

จากรูปที่ 3.7 ได้เลือกเวกเตอร์ไอเกิน = 5 เพื่อลดจำนวนข้อมูลในการรู้จำของเครือข่ายประสาทเทียมลงจึงเลือกค่าเวกเตอร์ไอเกินที่มีค่าน้อย ๆ แต่สามารถคงความเป็นลักษณะเด่นของภาพเดิมอยู่ เมื่อนำข้อมูลที่ได้หลังการวิเคราะห์ส่วนประกอบสำคัญเข้าสู่การรู้จำของเครือข่ายประสาทเทียมจะเหลือข้อมูลที่ใช้รู้จำเป็น 250×200 เมื่อนำมาทดสอบกับภาพที่เป็นใบหน้า 30 ภาพและภาพที่ไม่ใช่ใบหน้า 30 ภาพ โดยกำหนดค่าในการรู้จำดังนี้

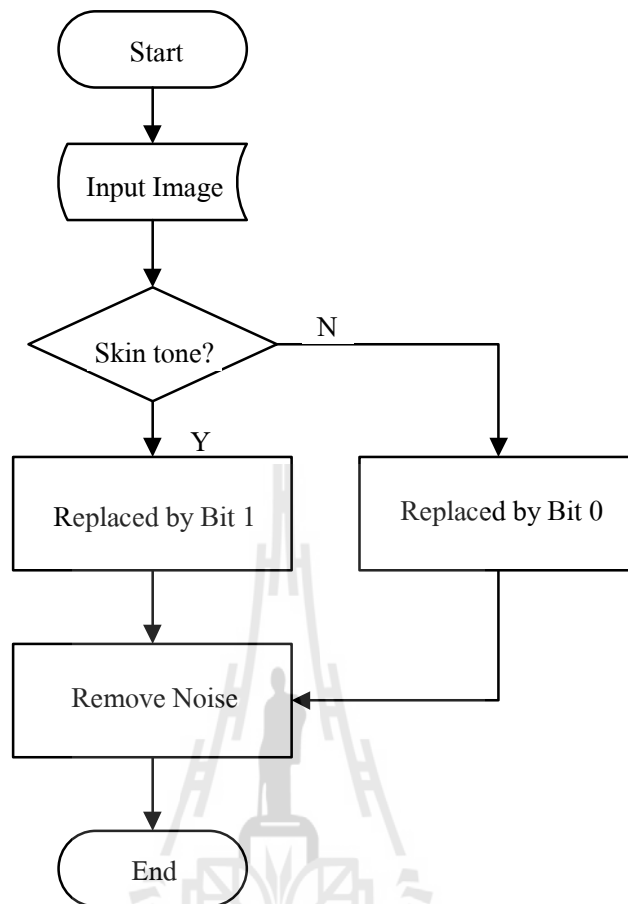
- Error Goal = 10×10^{-8}
- จำนวนรอบสูงสุด = 10000 รอบ
- จำนวนเลเยอร์ = 3 เลเยอร์



รูปที่ 3.8 เลเยอร์ของเครือข่ายประสาทเทียม

3.2.2 การออกแบบการตรวจจับใบหน้าโดยใช้การคัดแยกสีผิว

ในการการออกการตรวจจับใบหน้าโดยใช้การคัดแยกสีผิวผู้วิจัยได้ทำการเลือกใช้ปริภูมิสี RGB และ HSV มาใช้ จากหัวข้อ 3.2.1 การตรวจจับใบหน้าโดยใช้เทคนิคทางปัญญาประดิษฐ์สามารถตรวจจับได้เฉพาะใบหน้าที่เป็นหน้าตรงเนื่องจากภาพที่นำไปให้เครือข่ายประสาทเทียมรู้จำนั้นเป็นภาพใบหน้าตรงและมีขนาดคงที่จึงไม่เหมาะที่จะนำมาใช้กับการตรวจจับอาการง่วงนอนเนื่องจากขนาดของใบหน้าที่ได้จากกล้องเว็บแคมมีขนาดไม่คงที่เนื่องจากผู้จับขีรยนต์มีการหันซ้ายหรือหันขวา แนวทางหนึ่งที่สามารถแก้ปัญหานี้ได้คือทำการออกแบบการตรวจจับใบหน้าให้เหมาะสมโดยเลือกใช้การคัดแยกสีผิว โดยมีแผนภาพดังแสดงในรูปที่ 3.9



รูปที่ 3.9 แผนภาพการตรวจจับใบหน้าโดยการคัดแยกสีผิว

การหาช่วงของสีผิว สามารถทำได้โดยนำภาพใบหน้ามาแยกเป็นปริภูมิสี RGB และ ปริภูมิสี HSV โดยการแปลงปริภูมิสี RGB เป็น HSV ดังสมการที่ (3.6) ถึง (3.9)

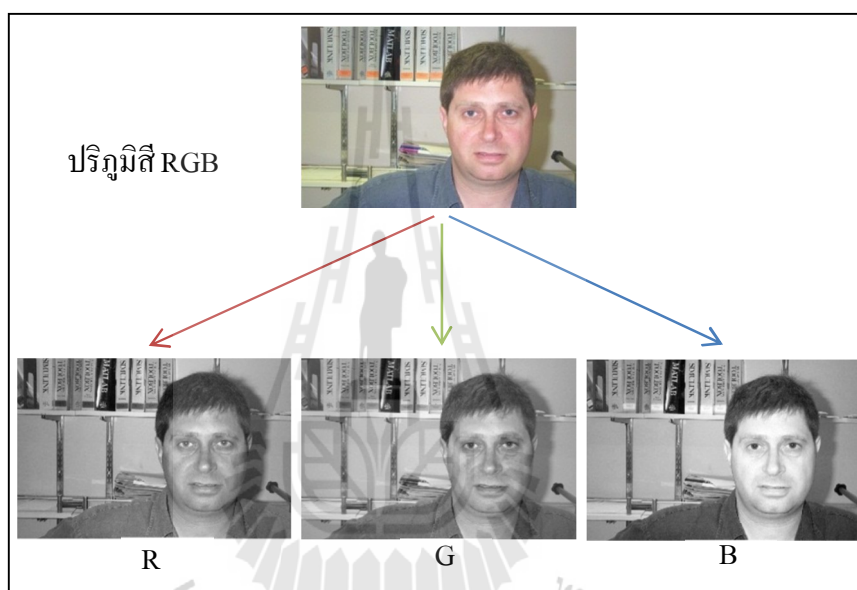
$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases} \quad (3-6)$$

$$\theta = \cos^{-1} \left\{ \frac{0.5[(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)^2}} \right\} \quad (3-7)$$

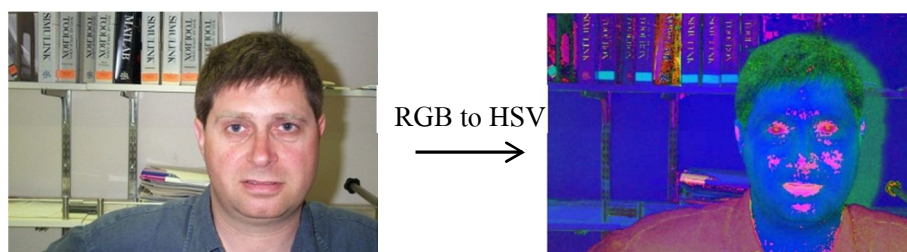
$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (3-8)$$

$$V = \frac{\max(R, G, B)}{255} \quad (3-9)$$

การคัดแยกสีผิว สีผิวเป็นช่วงสีกลุ่มหนึ่งในช่วงสีที่มีอยู่มากมาย เนื่องจากปริภูมิสี RGB เป็นปริภูมิสีที่แปรตามสภาพแสงได้ง่าย จึงนำปริภูมิสี HSV เข้ามาช่วยในการคัดแยกสีผิวโดยขั้นแรกนำภาพที่จะทดสอบมาแยกปริภูมิสี 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

จากนั้นนำภาพที่ได้ไปประมวลตามเงื่อนไขที่กำหนดตามสมการที่ (3.10) จะได้ดัง
แสดงในรูปที่ 3.12

$$Detec(x, y) = \begin{cases} 255 & \text{if } R(x, y) > G(x, y) \text{ and } |R(x, y) - G(x, y)| \geq 11 \\ & \text{and } 84 \leq R(x, y) \leq 153 \text{ and } 64 \leq G(x, y) \leq 95 \\ & \text{and } 30 \leq S(x, y) \leq 178 \text{ and } 76 \leq V(x, y) \leq 255 \\ 0 & \text{otherwise} \end{cases} \quad (3-10)$$

เมื่อ (x, y) คือ พิกัดของพิกเซลในภาพ

Detec คือ ภาพเอาต์พุตจากการประมวลผล



รูปที่ 3.12 ภาพจากการคัดแยกสีผิว

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



รูปที่ 3.13 ภาพหลังจากการกำจัดบริเวณที่ไม่ใช่ใบหน้าออก



รูปที่ 3.14 ภาพจากหลังการตรวจจับใบหน้า

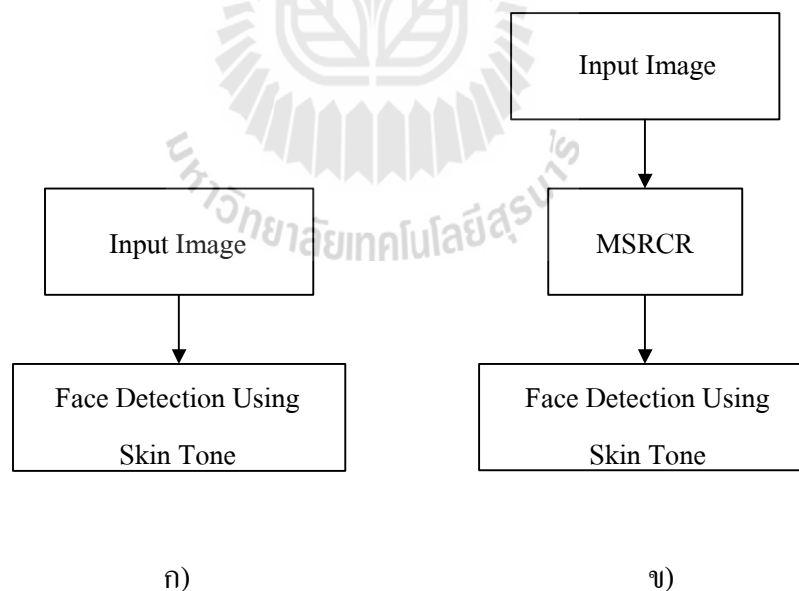
3.2.3 การออกแบบการตรวจจับใบหน้าโดยใช้การตัดแยกสีผิวและเทคนิค MSRCR

การตรวจจับใบหน้าโดยใช้การตัดแยกสีผิวในหัวข้อ 3.2.2 เป็นการตรวจจับที่ใช้ในการตัดแยกใบหน้าออกจากภาพพื้นหลัง ซึ่งการทดสอบในสภาวะใช้งานจริงที่มีแสงสว่างเปลี่ยนแปลงตลอดเวลา จึงต้องมีเทคนิคที่จะช่วยในการปรับปรุงภาพ (Image Enhancement) ให้แสงสว่างภายในภาพเหมาะสมจนสามารถตรวจจับใบหน้าโดยการตัดแยกสีผิวได้ เทคนิคนี้คือ MSRCR (MSR with Color Restoration) ซึ่งเป็นเทคนิคที่ช่วยในการปรับแสงสว่างในภาพ โดยจะนำเทคนิค MSRCR มาเพื่อปรับแสงสว่างในภาพก่อนการตรวจจับใบหน้าโดยการตัดแยกสีผิวดังแสดงในรูปที่ 3.15 เทคนิค MSRCR สามารถช่วยปรับแสงในภาพที่มีแสงมืดเกินไป และสว่างเกินไปให้มีแสงสว่างที่เหมาะสมต่อการตรวจจับใบหน้า เนื่องจากในสภาวะใช้งานจริงปริมาณแสง

สว่างในภาพจะมีการเปลี่ยนแปลงตลอดเวลา ในงานวิจัยนี้ได้ทำการเลือกใช้เลือกใช้พารามิเตอร์ของเทคนิค MSRCR ให้เหมาะสมกับรูปภาพขนาด 1920×1080 พิกเซล ซึ่งได้จากกล้องเว็บแคมรุ่น 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) การตรวจจับใบหน้าโดยใช้วิธีการทั่วไป
 ข) การตรวจจับใบหน้าที่เพิ่มเทคนิค MSRCR



รูปที่ 3.16 เปรียบเทียบ Retinex แบบต่างๆ

ก) ภาพต้นแบบ

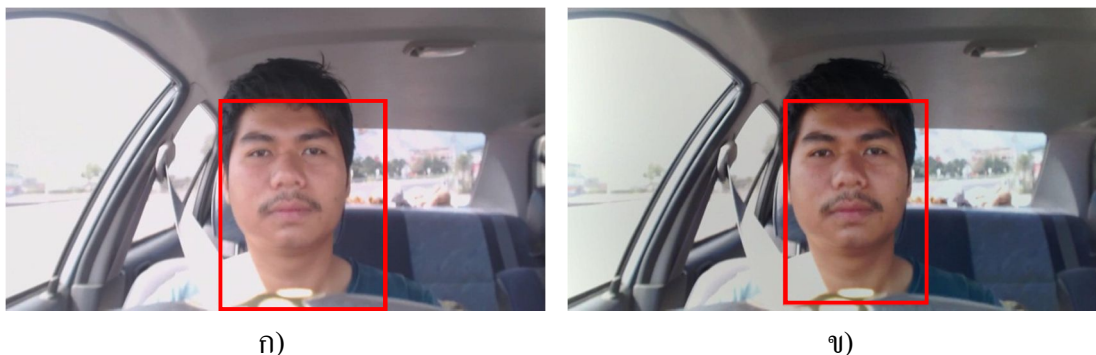
ข) ภาพที่ผ่าน SSR $\sigma = 500$

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



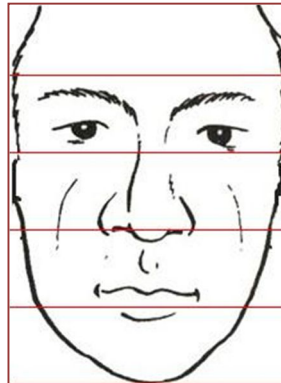
รูปที่ 3.17 ตรวจจับใบหน้าหลังจากการทำ MSRCR

- ก) รูปการตรวจจับใบหน้าโดยการคัดแยกสีผิวที่ไม่มีการปรับแสงสว่าง
- ข) รูปการตรวจจับใบหน้าโดยการคัดแยกสีผิวที่มีการปรับแสงสว่าง ด้วยเทคนิค MSRCR

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

3.3 การออกแบบระบบตรวจจับดวงตา

การตรวจจับดวงตาเป็นส่วนที่สำคัญมากสำหรับการตรวจจับอาการง่วงนอน โดยสามารถคาดเดาอาการง่วงนอนได้จากสถานะของดวงตา เช่น การหลับตาเป็นเวลานาน การกระพริบตาที่ช้าลง การหรี่ตา มุมมองการมองของตา เป็นต้น การตรวจจับดวงตาโดยใช้การประมวลผลภาพดิจิทัลจะทำหลังจากการตรวจจับใบหน้าเพื่อลดพื้นที่ในค้นหาดวงตาลงทำให้เวลาในการค้นหาลดลงด้วยเช่นกัน หลังจากการตรวจจับใบหน้าจะได้ดังแสดงในรูปที่ 3.18 เมื่อแบ่งความสูงของใบหน้าเป็น 5 ส่วนจะเห็นว่าบริเวณที่มีดวงตาจะอยู่บนเส้นที่ 3 นับจากเส้นด้านล่างจึงสามารถตัดเอาเฉพาะบริเวณนี้มาประมวลผลในการค้นหาดวงตาต่อไปดังแสดงในรูปที่ 3.19



รูปที่ 3.18 แบ่งความสูงของใบหน้าเป็น 5 ส่วน



รูปที่ 3.19 บริเวณใบหน้า 2 ส่วน

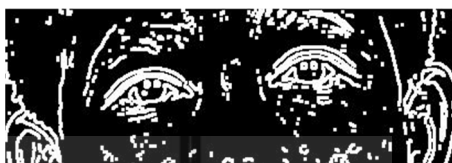
เมื่อนำบริเวณใบหน้า 2 ส่วนที่เหลือในรูปที่ 3.20 มาประมวลผลภาพดิจิทัลด้วยการหาขอบของวัตถุดังแสดงในรูปที่ 3.21 จากนั้นจะนำมาขยายขอบของวัตถุดังแสดงในรูปที่ 3.22 ทั้งนี้เพื่อให้ง่ายต่อการแยกบริเวณภาพดวงตาจึงนำภาพมากำจัดส่วนอื่น ๆ ที่ไม่ใช่บริเวณดวงตาเนื่องจากบริเวณของดวงตาจะมีขอบของภาพอยู่มากกว่าที่อื่น



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 ต่อมาจะทำการค้นหาตำแหน่งตาในงานวิจัยนี้ได้ นำเสนอวิธีการค้นหาตำแหน่งตา 2 วิธี ได้แก่ ค้นหาตำแหน่งตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน และค้นหาตำแหน่งตาด้วยวิธี Integral Projection ซึ่งวิธีการแรกจะเป็นการค้นหาตำแหน่งตาด้วยวิธี Integral Projection เนื่องจากตำแหน่งตาที่มีสีเข้มกว่าบริเวณอื่นจึงสามารถแยกตำแหน่งตาออกจากบริเวณอื่นได้ และวิธีการที่สองเป็นการรวมค่าของพิกเซลในแนวแกนตั้งและแนวแกนนอน

3.3.1 การค้นหาดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน

การค้นหาตำแหน่งตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน (Threshold) สามารถทำได้โดยแปลงภาพในปริภูมิสี RGB เป็นภาพระดับสีเทา ดังแสดงในรูปที่ 3.24 ต่อมาจะทำการตั้งค่าขีดเริ่มเปลี่ยน เพื่อคัดแยกบริเวณที่มีสีเข้มกว่าบริเวณอื่นดังสมการที่ (3-11) จะได้ผลดังแสดงในรูปที่ 3.25 ซึ่งเป็นภาพขาว-ดำ เมื่อคัดแยกจากการคำนวณหาจำนวนพิกเซลในแต่ละวัตถุ ความกว้าง ความสูง และระยะห่างระหว่างตาทั้งสองข้างดังแสดงในรูปที่ 3.26 จะได้ตำแหน่งของตำแหน่งตาดังแสดงในรูปที่ 3.27

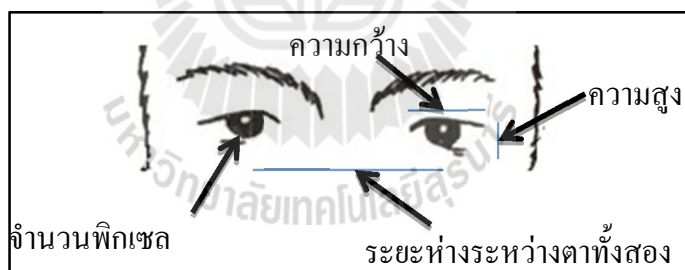
$$\text{Binary image} = \begin{cases} 1 & ; \text{if grayscale} \geq \text{threshold} \\ 0 & ; \text{else} \end{cases} \quad (3-11)$$



รูปที่ 3.24 ภาพระดับสีเทา



รูปที่ 3.25 ภาพหลังการตั้งค่าขีดเริ่มเปลี่ยน



รูปที่ 3.26 ลักษณะเฉพาะของตา



รูปที่ 3.27 ผลการตรวจจับดวงตา

3.3.2 การค้นหาดวงตาด้วยวิธีการ Integral Projection

Integral Projection [26] เป็นการค้นหาตำแหน่งของม่านตาโดยการหาผลรวมของค่าของพิกเซลของภาพระดับสีเทาในแนวแกนตั้งและแนวนอน 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
v	คือ ภาพที่นำมาประมวลผล
i	คือ พิกัดของแถว
j	คือ พิกัดของหลัก
n	คือ จำนวนของหลัก



ก)

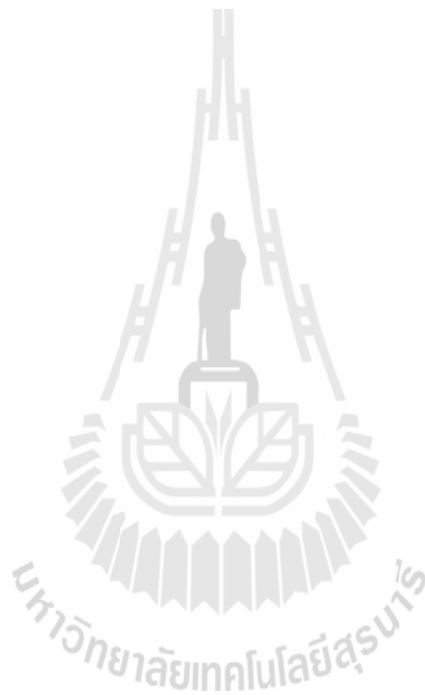


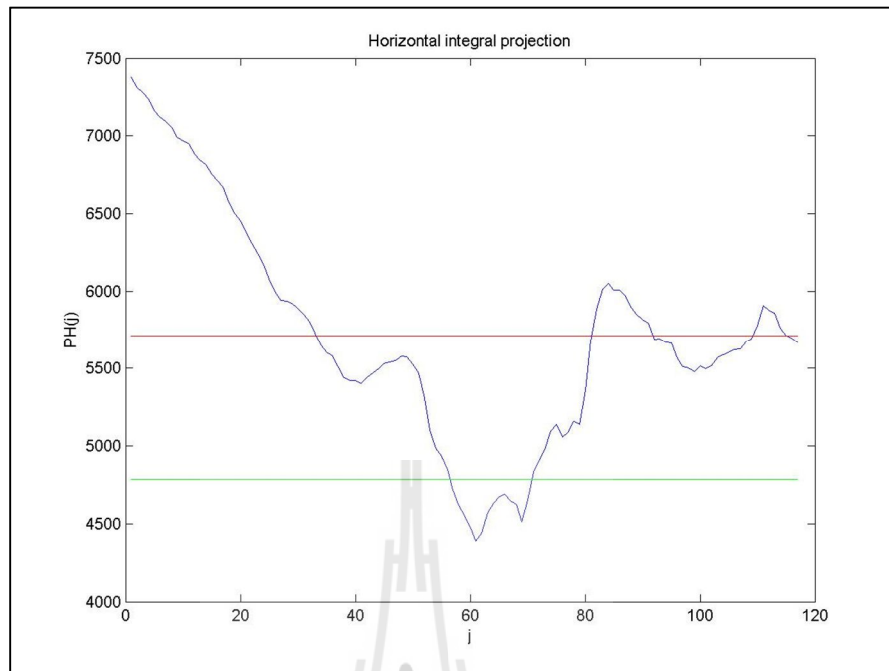
ข)

รูปที่ 3.28 บริเวณดวงตาที่เป็นภาพระดับสีเทา

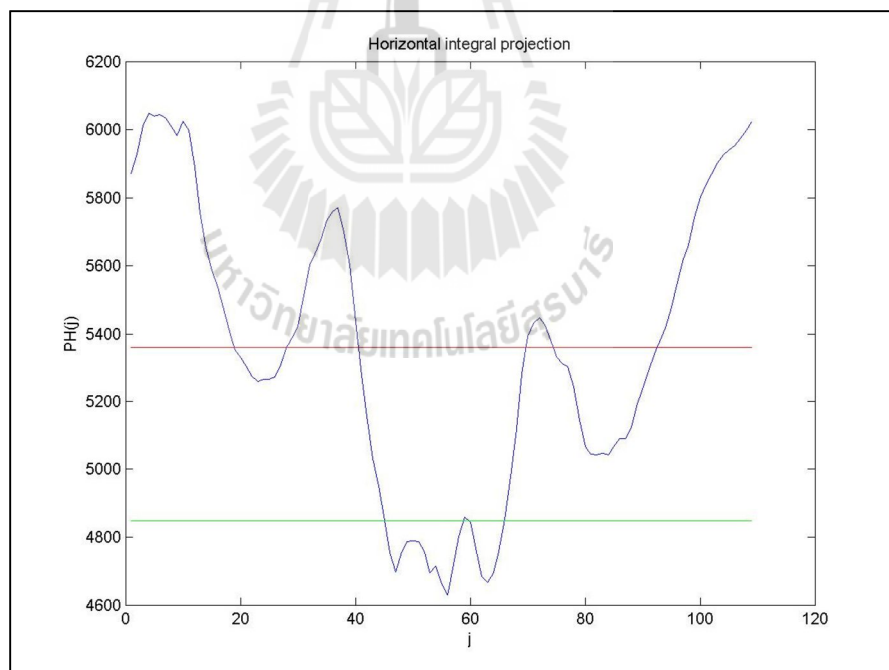
ก) ตาขว

ข) ตาช้ำ





ก)

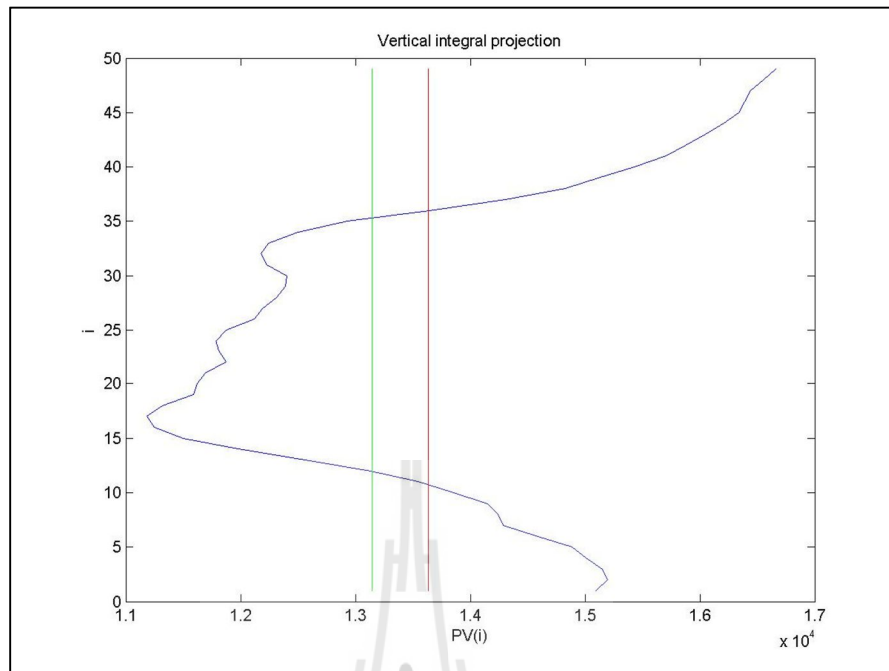


ข)

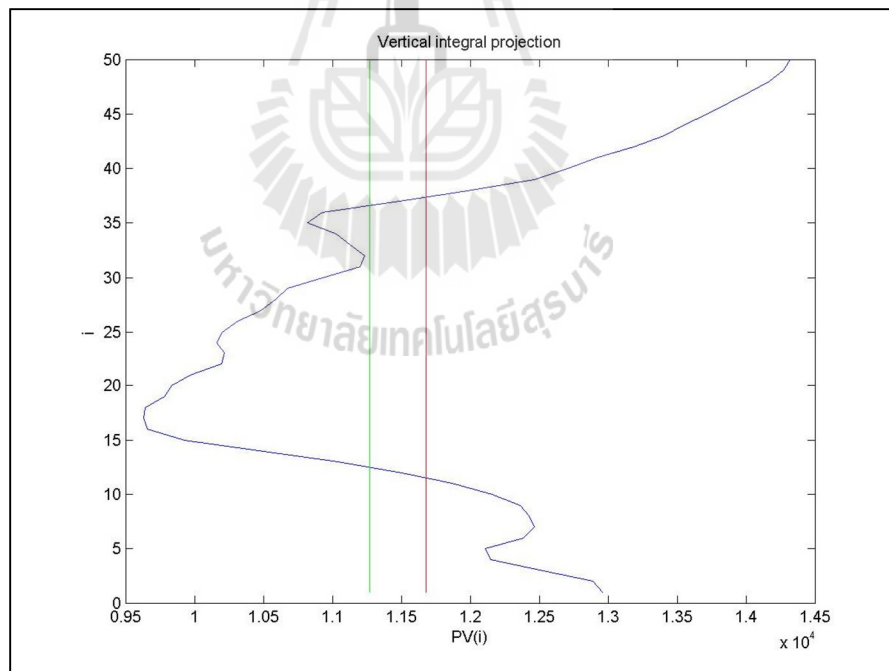
รูปที่ 3.29 กราฟ Horizontal Integral Projection

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 ดาขวา

ข) กราฟ Vertical Integral Projection ดาซ้าย

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

$$P_{H,arg} = \frac{\sum P_H}{m} \quad (3-14)$$

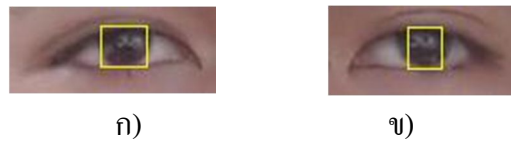
$$P_{V,arg} = \frac{\sum P_V}{n} \quad (3-15)$$

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

$$P_{H,Threshold} = P_{H,arg} - 0.7(P_{H,arg} - P_{H,min}) \quad (3-16)$$

$$P_{V,Threshold} = P_{V,arg} - 0.7(P_{V,arg} - P_{V,min}) \quad (3-17)$$

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



รูปที่ 3.31 ภาพม่านตาที่ได้จากวิธีการ Integral Projection

- ก) ตาขวา
ข) ตาซ้าย

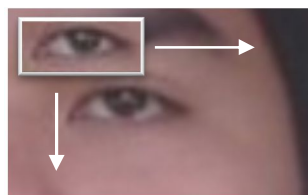
3.3.3 การค้นหาดวงตาด้วยการจับคู่กับแม่แบบ

การค้นหาดวงตาเป็นการนำภาพแม่แบบตาตั้งรูปที่ 3.32 ไปค้นหาในภาพใบหน้าโดยค้นหาเฉพาะบริเวณคาดว่าจะมีดวงตาตั้งรูปที่ 3.33 เพื่อลดกระบวนการในหาค้นหาดวงตาลงจะทำการค้นหาดวงตาโดยใช้แม่แบบตาซ้ายแก่ข้างเดียวเนื่องจากเวลาหลับตาหรือลืมตาปกปิดจะกระทำพร้อมกันทั้งสองข้าง โดยจะทำการค้นหาจากซ้ายไปขวา บนลงล่างตั้งรูปที่ 3.33 แล้วทำการประมวลผลโดยใช้ค่าของพิกเซลของภาพและแม่แบบที่อยู่ในตำแหน่งเดียวกัน เพื่อคำนวณหาค่าความเหมือนของแม่แบบกับภาพโดยได้ทำการทดลองกับฟังก์ชันต่าง ๆ ดังนี้ TM_SQDIFF , TM_SQDIFF_NORMED , TM_CCORR , TM_CCORR_NORMED , TM_CCOEFF , TM_CCOEFF_NORMED เพื่อเลือกที่เหมาะสมกับการค้นหาดวงตา



รูปที่ 3.32 ภาพแม่แบบของตา

- ก) ตาเปิด
ข) ตาปิด



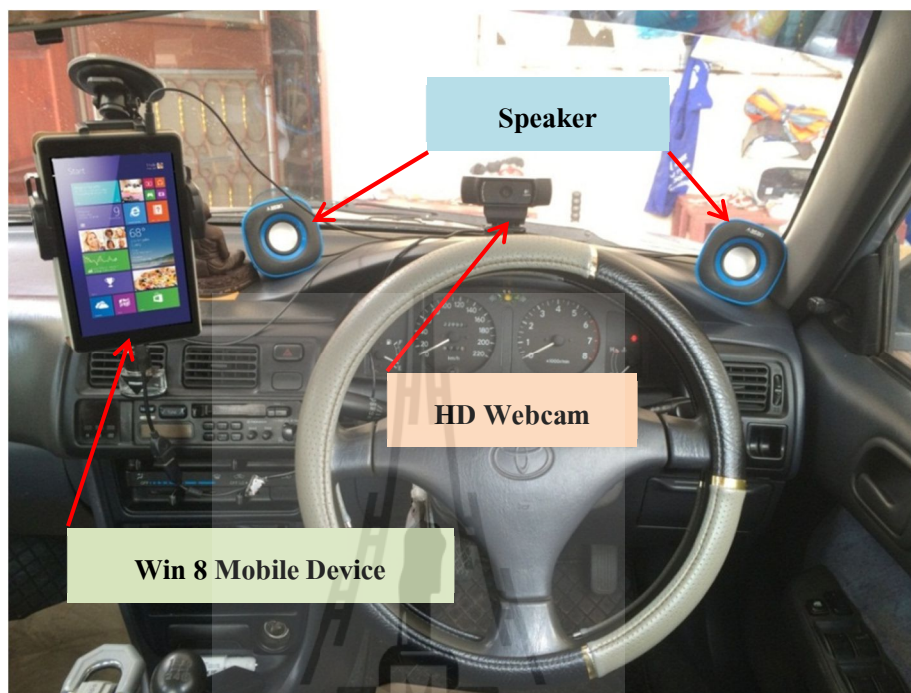
รูปที่ 3.33 การค้นหาจากซ้ายไปขวา บนลงล่าง

3.4 การออกแบบระบบตรวจจับอาการง่วงนอน

การตรวจจับอาการง่วงนอนที่ไม่รบกวนผู้ขับขี่รถยนต์วิธีหนึ่งคือ การตรวจจับอาการง่วงนอนด้วยการประมวลผลภาพดิจิทัลจากกล้องวิดีโอทัศนซึ่งวิธีการนี้เป็นการนำกล้องวิดีโอทัศนจับภาพใบหน้าของผู้ขับขี่รถยนต์ขณะขับรถจากนั้นทำการประมวลผลด้วยอุปกรณ์เคลื่อนที่ และแจ้งเตือนด้วยเสียงเมื่อมีอาการง่วงนอนเกิดขึ้น ดังแสดงระบบตรวจจับอาการง่วงนอนในรูปแบบที่ 3.34 เพื่อสังเกตลักษณะของใบหน้า ตรวจจับการเคลื่อนไหวของดวงตา และการเอียงของศีรษะ แล้วนำพารามิเตอร์อัตราการกระพริบตา มุมของการเอียงของศีรษะ มุมมองของคนขับ และอื่นๆ นำมาประมวลผลเพื่อคาดเดาถึงอาการง่วงนอนของผู้ขับขี่รถยนต์ซึ่งมีออกเป็น 3 ระดับคือ ตื่นตัวหรือรู้สึกตัว เมื่อยล้า และง่วงนอน แล้วแจ้งเตือนผู้ขับขี่รถยนต์ให้ทราบถึงอาการง่วงนอนของตนเอง โดยแจ้งเตือนเป็นข้อความบนหน้าจอของแท็บเล็ตพร้อมมีเสียงเตือนไปด้วยเมื่อผู้ขับขี่รถยนต์เริ่มมีอาการเมื่อยล้า หากผู้ขับขี่รถยนต์ยังมีอาการเมื่อยล้าอยู่หรือเริ่มมีอาการง่วงนอนจะมีการแจ้งเตือนด้วยเสียงบ่อยครั้งมากขึ้นและเตือนให้หยุดการขับรถต่อไป

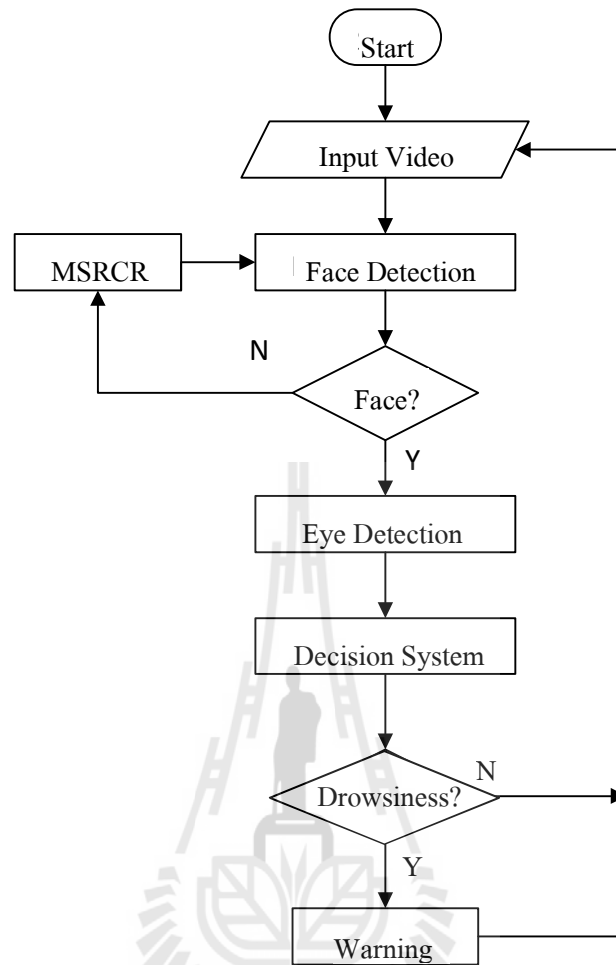
การตรวจจับอาการง่วงนอนด้วยการประมวลผลภาพดิจิทัลจากกล้องวิดีโอทัศนมีขั้นตอนดังแสดงในรูปแบบที่ 3.35 ในขั้นแรกจะดึงเฟรมภาพจากกล้องวิดีโอทัศนซึ่งในงานวิจัยนี้จะติดตั้งกล้องเว็บแคมรุ่น Logitech HD Pro Webcam C920 ในรถยนต์ดังแสดงในรูปแบบที่ 3.34 โดยเฟรมภาพที่ใช้มีขนาด 1920×1080 พิกเซลซึ่งมีความละเอียดสูงเพียงพอที่จะสามารถใช้ในการตรวจจับดวงตาของผู้ขับขี่รถยนต์ได้ จากนั้นจะนำเฟรมภาพมาตรวจจับใบหน้าโดยการคัดแยกสีผิว เนื่องจากวิธีการนี้สามารถตรวจจับได้ขณะหน้าของผู้ขับขี่รถยนต์เอียงเล็กน้อยและสามารถตรวจจับได้กับใบหน้าของผู้ขับขี่รถยนต์ทุกขนาด ส่วนการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมจะต้องใช้ภาพที่มาให้เครือข่ายประสาทเทียมรู้จำมากพอสมควรถึงจะสามารถตรวจจับใบหน้าที่มีอิริยาบถแตกต่างจากใบหน้าตรงได้ ในขั้นตอนต่อมาจะตรวจสอบว่าในบริเวณที่ตรวจจับพบใบหน้าหรือไม่ ถ้าใช่จะทำการตรวจจับดวงตาต่อไป ถ้าไม่ใช่จะนำเฟรมภาพนี้มาปรับแสงในภาพด้วยวิธีการ MSRCR แล้วนำเฟรมภาพนี้กลับมาตรวจจับใบหน้าอีกครั้งในขั้นตอนต่อมาจะทำการตรวจจับดวงตาด้วยวิธีการ Integral Projection เนื่องจากการตรวจจับดวงตาด้วยวิธีการ Integral Projection นี้มีร้อยละการตรวจจับถูกต้องมากกว่าวิธีการตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยนเนื่องจากวิธีการ Integral Projection สามารถตรวจจับดวงตาได้เมื่อภาพมีปริมาณของแสงสว่างที่เปลี่ยนแปลง เพื่อนำบริเวณของม่านตามาประมวลผลและตัดสินใจว่าผู้ขับขี่รถยนต์มีอาการง่วงนอนหรือไม่ ถ้าพบว่ามีอาการง่วงนอนระบบจะแจ้งเตือนให้ผู้ขับขี่รถยนต์ทราบเพื่อให้ผู้ขับตื่นหรือเปลี่ยนอิริยาบถในการขับรถ ในงานวิจัยนี้ได้ทำการจำลองการทำงานของระบบดังกล่าวด้วยโปรแกรม MATLAB

โดยใช้เวลาในการประมวลผลประมาณ 11 วินาทีต่อเฟรม ส่วนในการประยุกต์เพื่อใช้งานจริงนั้น จะต้องทำการพัฒนาโปรแกรมให้สามารถทำงานได้แบบเวลาจริง



รูปที่ 3.34 การติดตั้งกล้องวิดีโอทัศนียภาพในยานพาหนะ

มหาวิทยาลัยเทคโนโลยีสุรนารี



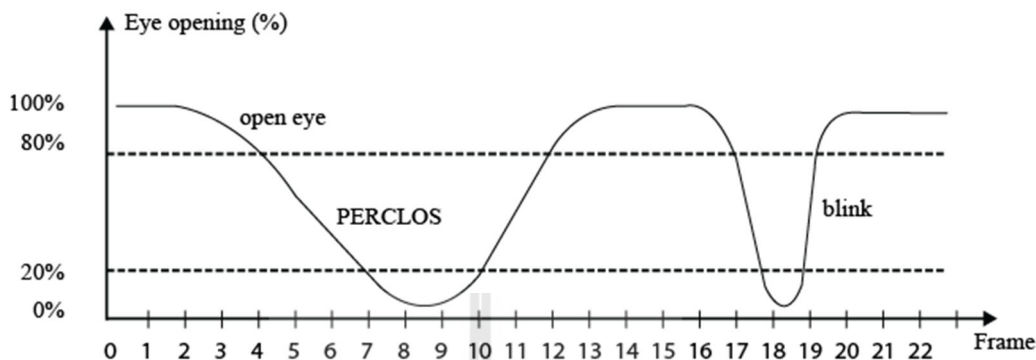
รูปที่ 3.35 ขั้นตอนการตรวจจับอาการง่วงนอน

3.4.1 Percentage of Eye Closure Over Time (PERCLOS)

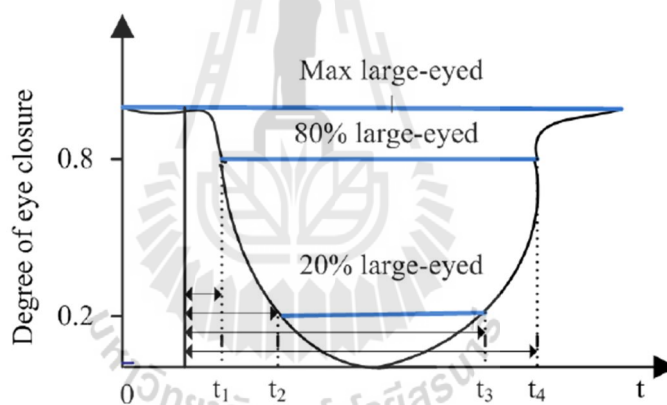
ระบบตัดสินใจที่ใช้ในการตรวจจับอาการง่วงนอนจะใช้แนวทางการตรวจหาสถานะการปิด-เปิดของดวงตา โดยทำการคำนวณหาค่าร้อยละการปิดของดวงตา (Percentage of Eye Closure Over Time: PERCLOS) โดยปกติจะใช้เวลาในการกระพริบตา 0.3 วินาที [25] ในงานวิจัยนี้ใช้กล้องเว็บแคมรุ่น Logitech HD Pro Webcam C920 เฟรมภาพขนาด 1920×1080 พิกเซล มีอัตราเฟรม 30 เฟรมต่อวินาที ในการเก็บข้อมูลภาพวิดีโอที่บันทึกแสดงให้เห็นว่าการกระพริบตา 1 ครั้งจะใช้เฟรมภาพประมาณ 7-9 เฟรมภาพ

PERCLOS [26] คือร้อยละของการปิดตาดำกว่าร้อยละ 20 ต่อการปิดตาระหว่างร้อยละ 20-80 ของการเปิดตาที่สูงสุด จากรูปที่ 3.36 มีทั้งหมด 22 เฟรมภาพโดยช่วงที่นำมาประมวลผลหา PERCLOS คือช่วงจากเฟรมภาพที่ 4 ถึงเฟรมที่ 12 สามารถคำนวณ PERCLOS ได้จากสมการที่

(3.18) โดยที่ $t_2 - t_3$ คือช่วงเวลาที่ตาปิดต่ำกว่าร้อยละ 20 และ $t_1 - t_4$ คือ ช่วงเวลาที่ตาปิดระหว่างร้อยละ 20 กับ 80



รูปที่ 3.36 เฟรมที่ใช้ในการคำนวณ PERCLOS [26]



รูปที่ 3.37 การกำหนดช่วงเวลาในการประมวลผล PERCLOS [26]

$$PERCLOS = \frac{t_3 - t_2}{t_4 - t_1} \times 100\% \tag{3-18}$$

เมื่อ *PERCLOS* คือ ร้อยละการปิดของดวงตา

t_2, t_3 คือ ช่วงเวลาที่ตาปิดต่ำกว่าร้อยละ 20

t_4, t_1 คือ ช่วงเวลาที่ตาปิดระหว่างร้อยละ 20-80

วิธีการคาดเดาอาการง่วงนอนโดยใช้ค่า PERCLOS มีดังนี้

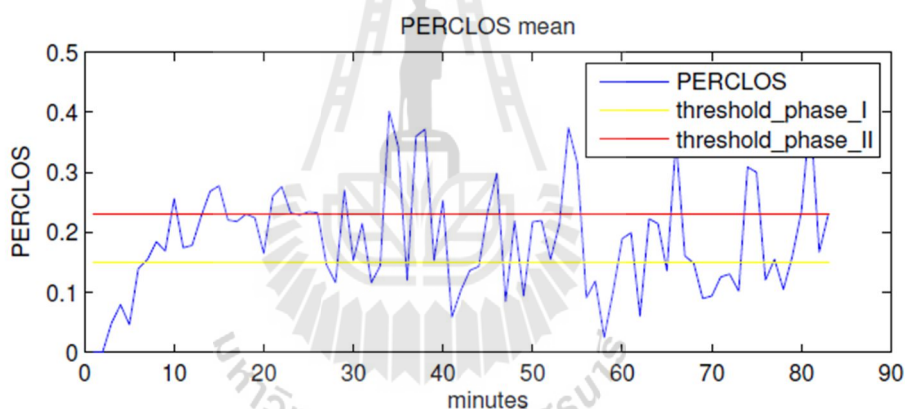
1) วิธีการที่ 1 เป็นวิธีการที่นำเสนอในงานวิจัยโดย W. Qing et al. [25] ซึ่งมีการคาดการณ์ระดับของอาการง่วงนอนจากค่า PERCLOS โดยเมื่อมีค่า PERCLOS มากกว่า 0.4 หรือมีการปิดตานานเกิน 3 วินาทีจะสามารถคาดเดาได้ว่าผู้ทดสอบกำลังมีอาการง่วงนอนอยู่ในขณะนั้น ในการคำนวณค่า PERCLOS สามารถคำนวณได้จากสมการ (3-18)

2) วิธีการที่ 2 เป็นวิธีการที่นำเสนอในงานวิจัยโดย I. Garcia et al. [26] ซึ่งได้แบ่งระดับของอาการง่วงนอนออกเป็น 3 ระดับตามค่า PERCLOS ที่คำนวณได้ ดังแสดงในรูปที่ 3.36 ได้แก่

ระดับที่ 1 ตื่นตัวหรือรู้สึกตัว (Wake) เป็นช่วงที่มี PERCLOS ต่ำกว่าร้อยละ 15

ระดับที่ 2 เมื่อยล้า (Fatigue) เป็นช่วงที่มี PERCLOS อยู่ระหว่างร้อยละ 15-23

ระดับที่ 3 ง่วงซึม (Sleepiness) เป็นช่วงที่มี PERCLOS สูงกว่าร้อยละ 23



รูปที่ 3.38 กราฟ PERCLOS

ในงานวิจัยนี้ได้นำเอาวิธีการคาดเดาอาการง่วงวิธีการที่ 1 มาใช้ในการตรวจจับอาการง่วงนอน โดยจะใช้การจับคู่แม่แบบมาใช้ในการตรวจจับดวงตา

3.5 สรุป

การออกแบบระบบเฝ้าติดตามอาการง่วงนอนที่ได้กล่าวมาในบทนี้ โดยได้กล่าวถึงการออกแบบการตรวจจับใบหน้า 2 วิธีการ วิธีการแรกคือการตรวจจับใบหน้าโดยใช้ใช้เครือข่ายประสาทเทียม ซึ่งเป็นการนำฐานข้อมูลภาพใบหน้าและภาพที่ไม่ใช่ใบหน้าขนาด 200×200 พิกเซล อย่างละ 100 ภาพมาดึงลักษณะเด่นด้วยวิธีการวิเคราะห์ส่วนประกอบสำคัญและการแปลงเวฟเล็ต

แบบสีสกรีนก่อนที่จะนำมาให้เครือข่ายเทียมรู้จำและได้ค่าน้ำหนักเพื่อนำไปใช้ตรวจจับใบหน้าต่อไป และวิธีการที่สองคือการตรวจจับใบหน้าโดยใช้การคัดแยกสีผิวเป็นการตรวจจับใบหน้าที่ใช้ช่วงสีผิวที่อยู่ในปริภูมิสี RGB และ HSV และยังได้นำเทคนิคการปรับปรุงแสงสว่างมาช่วยในการปรับแสงในภาพเพื่อให้การตรวจจับใบหน้าโดยการคัดแยกสีผิวมีประสิทธิภาพมากขึ้น ต่อมาได้กล่าวถึงการออกแบบการตรวจจับดวงตาด้วยวิธีการตั้งค่าขีดเริ่มเปลี่ยน และการตรวจจับดวงตาโดยใช้เทคนิค Integral Projection ซึ่งวิธีการแรกเป็นการค้นหาบริเวณของม่านตาจากลักษณะเฉพาะของดวงตา เนื่องจากความเข้มของม่านตาจะเข้มกว่าบริเวณอื่น และวิธีการที่สองเป็นการเป็นการค้นหาตำแหน่งของม่านตาโดยการหาผลรวมของค่าของพิกเซลของภาพระดับสีเทาในทั้งแกนตั้งและแกนนอน และในส่วนสุดท้ายเป็นการออกแบบการตรวจจับอาการง่วงนอนด้วยวิธีการคำนวณค่าร้อยละการปิดของดวงตาหรือ PERCLOS เพื่อเป็นแนวทางในการตรวจจับอาการง่วงนอนต่อไป



บทที่ 4

ผลการทดลอง

4.1 บทนำ

ในบทนี้จะนำเสนอผลการตรวจจับใบหน้าและดวงตาสำหรับระบบเฝ้าติดตามอาการง่วงนอน โดยได้แบ่งเป็นหัวข้อต่าง ๆ ประกอบด้วย ผลการการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมและการคัดแยกสีผิว ในส่วนของการคัดแยกสีผิวผู้วิจัยได้เพิ่มเทคนิค MSRCR เข้าไปด้วยเพื่อให้ผลการทดลองมีความถูกต้องมากยิ่งขึ้น จากนั้นจะนำเสนอเป็นผลการทดลองการตรวจจับดวงตาและส่วนสุดท้ายจะกล่าวถึงบทสรุปผลการทดลอง

4.2 ผลการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม

ในการนำเครือข่ายประสาทเทียมมาใช้ในการค้นหาใบหน้านั้นจะต้องมีการเตรียมฐานข้อมูลเพื่อใช้ในการรู้จำของเครือข่ายประสาทเทียม ในงานวิจัยนี้ได้ใช้ภาพใบหน้าและภาพที่ไม่ใช่ใบหน้าขนาด 200×200 พิกเซล อย่างละ 100 ภาพ ดังตัวอย่างแสดงในรูปที่ 4.1 และรูปที่ 4.2 จากนั้นจะนำภาพดังกล่าวมาทำการปรับความสว่าง และนำมาดึงลักษณะเด่นและลดขนาดโดยการแปลงเวฟเล็ดแบบดีสครีทให้มีขนาด 50×50 พิกเซล หลังจากนั้นจะนำภาพหลังการแปลงเวฟเล็ดแบบดีสครีทมาดึงลักษณะเด่นของภาพโดยใช้การวิเคราะห์ส่วนประกอบสำคัญและนำไปให้เครือข่ายประสาทเทียมรู้จำ โดยได้นำเสนอแผนภาพการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมดังแสดงในรูปที่ 3.1

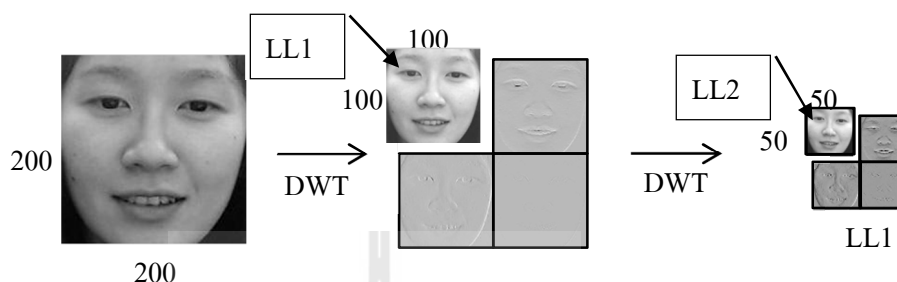


รูปที่ 4.1 ภาพใบหน้าขนาด 200×200 พิกเซล



รูปที่ 4.2 ภาพที่ไม่ใช่ใบหน้าขนาด 200×200 พิกเซล

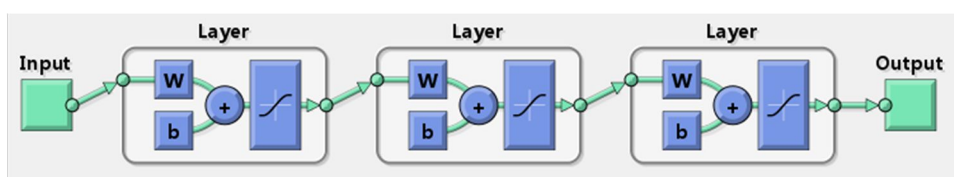
นำภาพใบหน้าขนาด 200×200 พิกเซลมาแปลงเวฟเล็ตแบบดีสครีทในระดับที่ 2 จะเหลือขนาดของภาพในแบนด์ย่อยความถี่ต่ำสุด (LL2) ที่จะนำไปประมวลผลต่อเท่ากับ 50×50 พิกเซล จะได้ดังแสดงในรูปที่ 4.3



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

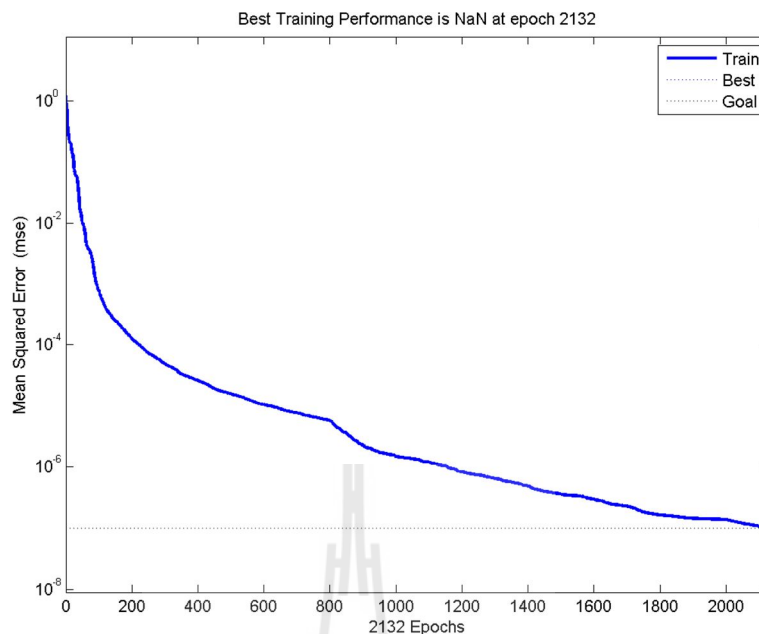
1. Error Goal = 10×10^{-8}
2. จำนวนเลเยอร์ = 3 เลเยอร์
3. จำนวนรอบสูงสุด = 10000 รอบ



รูปที่ 4.4 เลเยอร์ของเครือข่ายประสาทเทียม

การทดลองครั้งที่ 1

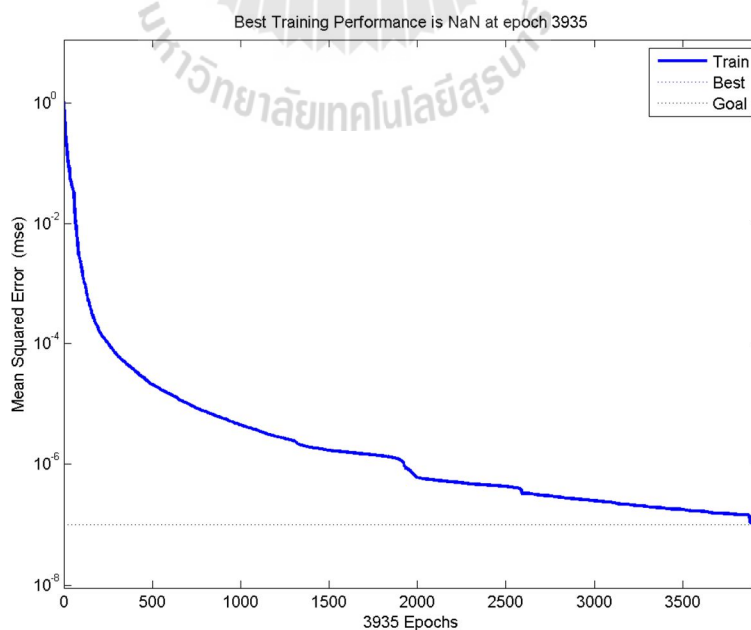
Eigenvectors = 5, Layer = 3-5-1('logsig' 'tansig' 'purelin') ใช้เวลาในการรู้จำ 107.229 วินาที จำนวน 2132 รอบ จะได้กราฟในการรู้จำดังแสดงในรูปที่ 4.5



รูปที่ 4.5 กราฟการรู้จำของเครือข่ายประสาทเทียมในการทดลองที่ 1

การทดลองครั้งที่ 2

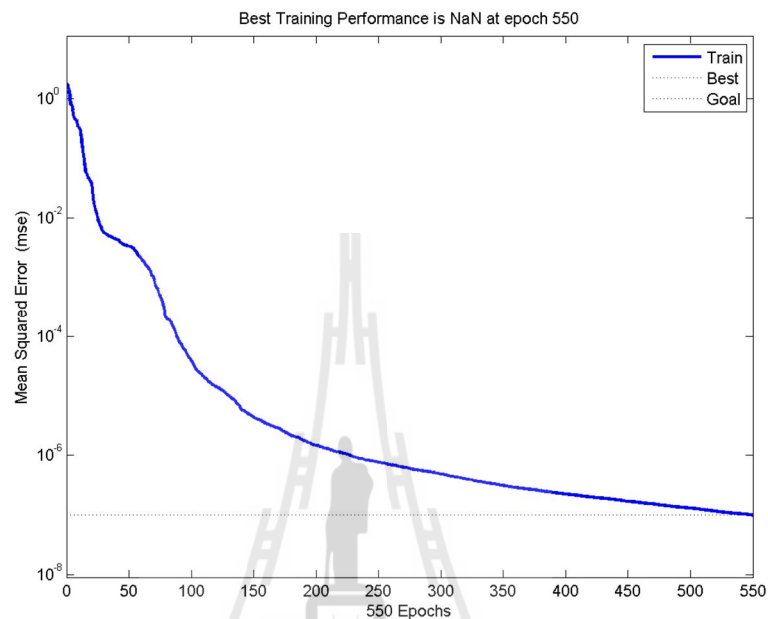
Eigenvectors = 5, Layer = 5-5-1('logsig' 'tansig' 'purelin') ใช้เวลาในการรู้จำ 98.574 วินาที จำนวน 3935 รอบ จะได้กราฟในการรู้จำดังแสดงในรูปที่ 4.6



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		ร้อยละความถูกต้องของการตรวจจับทั้งหมด
	ภาพใบหน้า	ภาพที่ไม่ใช่ใบหน้า	
ครั้งที่ 1	25	28	88.33
ครั้งที่ 2	22	28	83.22
ครั้งที่ 3	22	29	85.00
ค่าเฉลี่ย	23	28.33	85.55

ผลการตรวจจับใบหน้าโดยใช้เทคนิคเครือข่ายประสาทเทียมแสดงในตารางที่ 4.1 โดยทำการทดสอบภาพที่เป็นใบหน้า 30 ภาพและภาพที่ไม่ใช่ใบหน้า 30 ภาพจะใช้เวลาในการประมวลผล 0.075 วินาทีต่อภาพ พบว่ามีร้อยละการตรวจจับใบหน้าที่ถูกต้องประมาณร้อยละ 85.55 ซึ่งอาจเกิด

จากจำนวนภาพที่นำมาใช้ในการรู้จำของเครือข่ายประสาทเทียมยังไม่เพียงพอ การเพิ่มร้อยละความถูกต้องของการตรวจจับอาจทำได้โดยการเก็บรวบรวมข้อมูลใบหน้าที่ใช้ในการรู้จำให้มากขึ้น เนื่องจากผู้วิจัยเห็นว่า การตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมนั้นสามารถตรวจจับได้ตามลักษณะของใบหน้าที่มีการนำไปรู้จำเท่านั้นจึงไม่เหมาะที่จะนำมาตรวจจับใบหน้าของผู้ขับขี่รถยนต์จึงได้ออกแบบการตรวจจับใบหน้าโดยการคัดแยกสีผิวในหัวข้อถัดไป

4.3 ผลการตรวจจับใบหน้าโดยการคัดแยกสีผิว

ในการตรวจจับใบหน้าโดยการคัดแยกสีผิวจะนำอัลกอริทึมการตรวจจับใบหน้าโดยการคัดแยกสีผิวที่ออกแบบให้หัวข้อ 3.2.2 มาพัฒนาปรับปรุงให้ดีขึ้น โดยใช้โปรแกรม MATLAB และทดสอบกับภาพนิ่งที่ได้จากฐานข้อมูลจำนวน 5 ฐานข้อมูลดังนี้

- California Institute of Technology (CIT) [27] เป็นฐานข้อมูลใบหน้าทั้งหมด 450 ภาพจาก 27 คน มีขนาด 896×592 พิกเซล ซึ่งแต่ละภาพจะมีพื้นหลัง และแสงสว่างที่แตกต่างกัน
 - Bao Face [28] เป็นฐานข้อมูลใบหน้าที่มีขนาดแตกต่างกันประกอบด้วยภาพใบหน้าเดี่ยว 149 ภาพ และภาพใบหน้ากลุ่ม 221 ภาพ แต่เลือกเฉพาะส่วนที่เป็นใบหน้าเดี่ยวมาใช้ทดสอบเพื่อให้เหมาะสมกับงานวิจัย
 - Essex [29] เป็นฐานข้อมูลใบหน้าที่ได้จากคน 395 คน คนละ 20 ภาพ รวมทั้งหมด 7900 ภาพ
 - Georgia tech [30] 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 สามารถหาได้จากนำภาพหลังการตรวจจับใบหน้ามาตีตารางจากนั้นจะนับช่องที่มีการตรวจจับถูกต้องและผิดพลาดดังแสดงในรูปที่ 4.8 ก) และ ข) ซึ่งมีความถูกต้องเป็นร้อยละ 100 และ 84.5 ตามลำดับ



ก)



ข)

รูปที่ 4.8 เปรียบเทียบความถูกต้อง

ก) ถูกต้อง 100 %

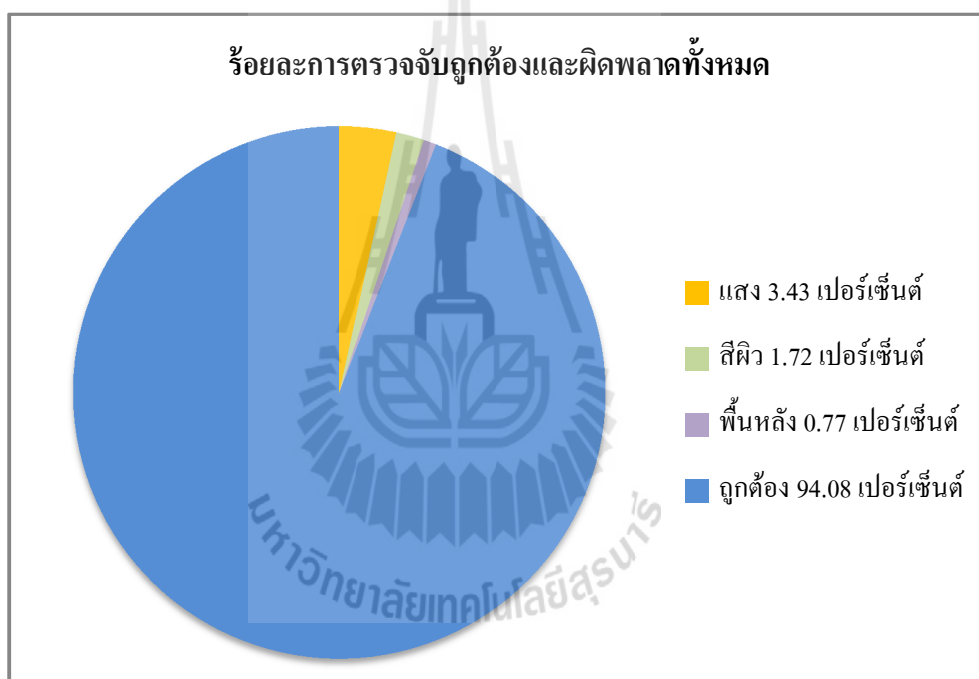
ข) ถูกต้อง 84.50 %

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

ฐานข้อมูลใบหน้า	วิธีการของ P. P. Paul and M. Gavrilova [11] (ร้อยละ)	วิธีการตรวจจับใบหน้าที่น่าสนใจ (ร้อยละ)
1. CIT	91.2	95.04
2. Bao Face	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

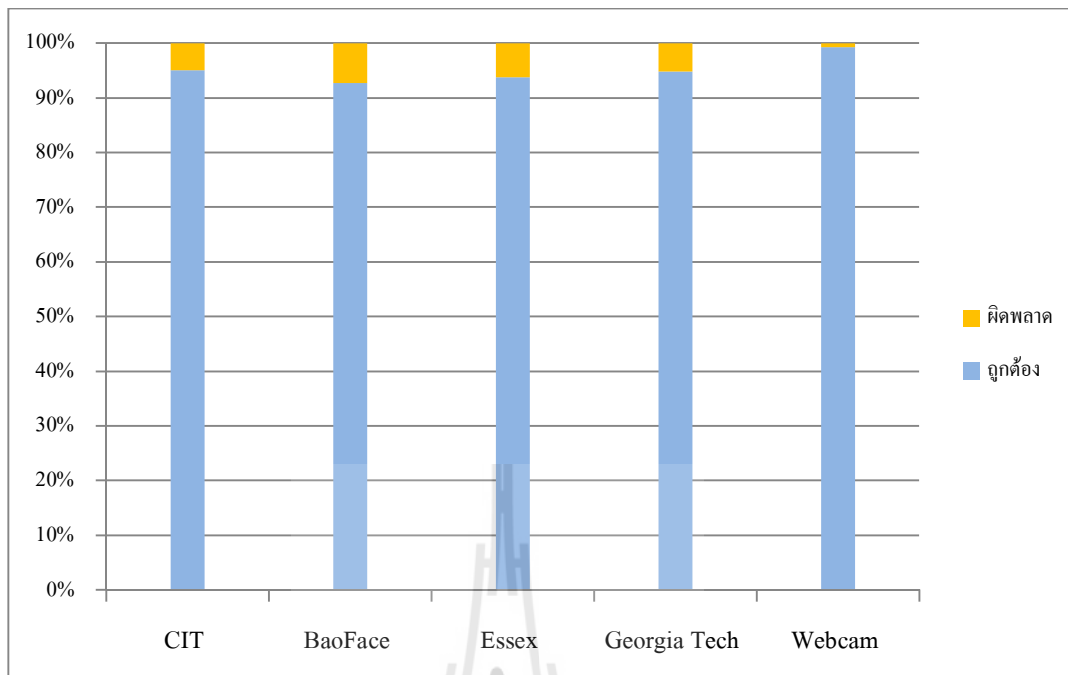
ความผิดพลาดในการตรวจจับใบหน้าที่แสดงในรูปที่ 4.8 อาจมีสาเหตุมาจาก

1. ภาพจากฐานข้อมูลเป็นภาพที่มีแสงสว่างไม่ปกติ ซึ่งทำให้การคัดแยกด้วยสีนั้นผิดพลาดเนื่องจากแสงสว่างมีผลต่อสีในภาพ ปัจจัยนี้ทำให้การตรวจจับผิดพลาดร้อยละ 3.43
2. ภาพจากฐานข้อมูลเป็นภาพใบหน้าของหลายเชื้อชาติ ซึ่งภาพตัวอย่างที่นำไปใช้หาขอบเขตของสีผิวโดยการคำนวณฮิตโตแกรมเป็นภาพใบหน้าของคนเอเชีย เมื่อนำไปทดสอบกับภาพใบหน้าของชาวสีอื่นจึงทำให้เกิดการตรวจจับผิดพลาด ปัจจัยนี้ทำให้การตรวจจับผิดพลาดร้อยละ 1.72
3. พื้นหลังของภาพในฐานข้อมูลมีสีที่คล้ายกับสีผิวจึงทำให้เกิดการตรวจจับผิดพลาดถึงร้อยละ 0.77

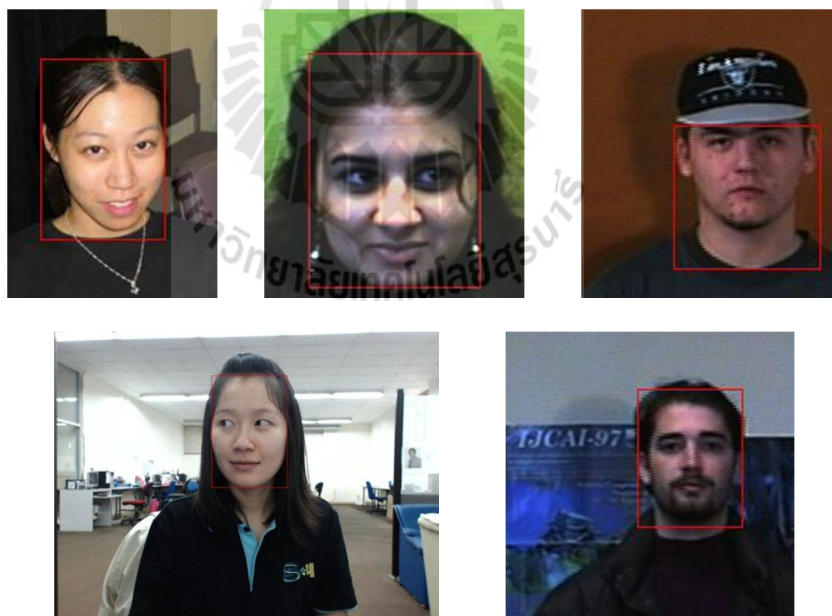


รูปที่ 4.9 แสดงร้อยละการตรวจจับถูกต้องและผิดพลาด

จะเห็นได้ว่าแสงสว่างมีผลการต่ออัลกอริทึมการตรวจจับใบหน้ามากที่สุด ผู้วิจัยจึงได้สร้างฐานข้อมูลใบหน้าขึ้น โดยใช้กล้องที่มีการชดเชยแสงสว่าง เมื่อนำภาพในฐานข้อมูลนี้มาทดสอบจะได้ผลดังแสดงในรูปที่ 4.10 ซึ่งจะเห็นได้ว่าเมื่อมีการปัญหาของแสงสว่างในภาพโดยใช้กล้องที่มีการชดเชยแสง ทำให้การตรวจจับใบหน้ามีความถูกต้องมากขึ้นถึงร้อยละ 99.30



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 เพื่อเก็บภาพขณะขับรถภายใต้แสงสว่างที่เปลี่ยนแปลงตามธรรมชาติดังแสดงในรูปที่ 4.14

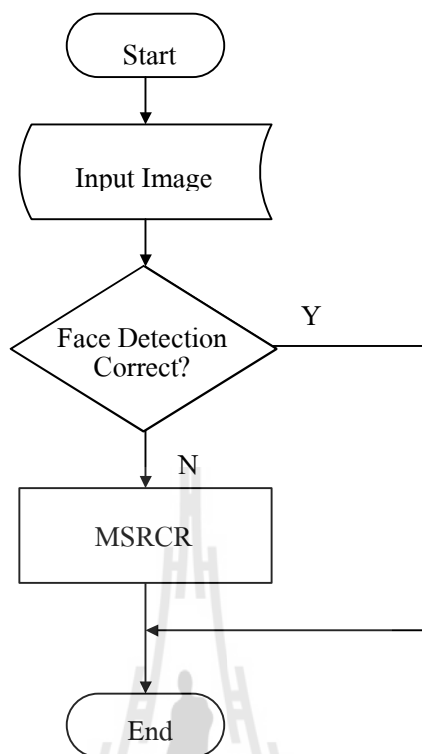


รูปที่ 4.13 การติดตั้งกล้องเว็บแคมภายในยานพาหนะ



รูปที่ 4.14 ภาพที่ใช้ในการทดสอบ

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

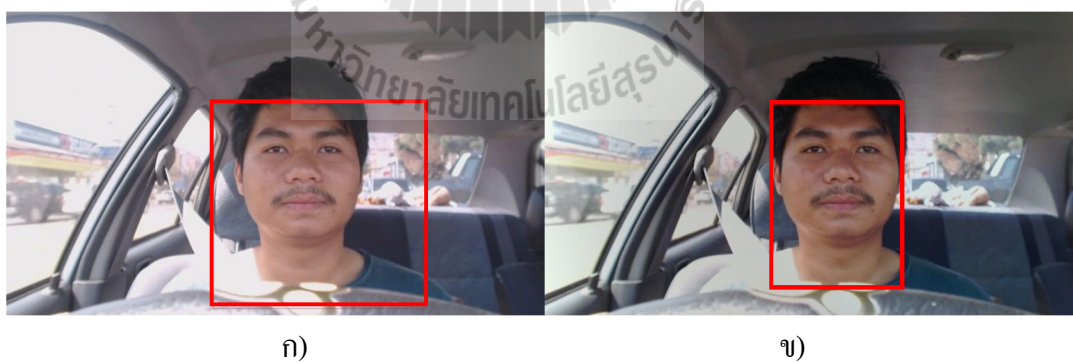


รูปที่ 4.15 แผนภาพการตรวจจับใบหน้าร่วมกับเทคนิค MSRCR

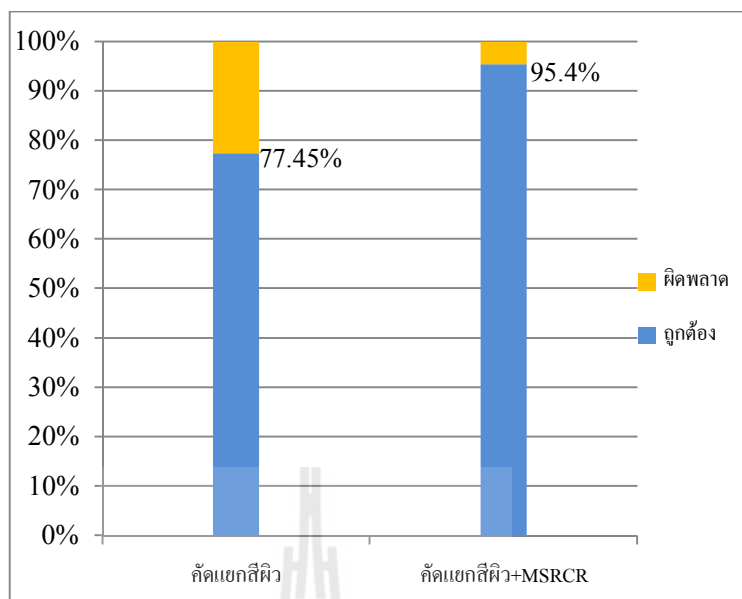
นำเทคนิค MSRCR มาช่วยในการปรับปรุงแสงสว่างในภาพที่มีการตรวจจับใบหน้าโดยการคัดแยกสีผิวแล้วให้ผลการตรวจจับผิดพลาด โดยการทดสอบจะทดสอบกับภาพวิดีโอที่ได้อาจสภาพแสงสว่างจริงขณะขับรถ มีความยาว 24.11 นาที อัตราเฟรม 30 ภาพต่อวินาทีโดยจะดึงภาพมาประมวลผลทุก ๆ 10 เฟรมภาพ จะได้ภาพในการทดสอบทั้งหมด 4,341 ภาพ หลังจากนำภาพตรวจจับใบหน้าแล้วได้สุ่มเลือกภาพที่มีการตรวจจับใบหน้าผิดพลาดมา 500 ภาพเพื่อมาปรับปรุงแสงสว่างโดยใช้เทคนิค MSRCR แล้วนำไปตรวจจับใบหน้าโดยการคัดแยกสีผิวอีกครั้ง พบว่าเทคนิค MSRCR ช่วยปรับปรุงแสงสว่างในภาพให้มีผลการตรวจจับใบหน้าถึง 463 ภาพดังตัวอย่างแสดงรูปที่ 4.17 จะเห็นว่าการตรวจจับใบหน้าหลังการปรับแสงสว่างแล้วให้ความถูกต้องมากขึ้นและกราฟในรูปที่ 4.18 แสดงให้เห็นว่าการตรวจจับใบหน้าโดยการคัดแยกสีผิวร่วมกับเทคนิค MSRCR จะสามารถปรับปรุงภาพจากเดิมที่มีร้อยละการตรวจจับถูกต้องร้อยละ 77.45 ให้มีร้อยละการตรวจจับถูกต้องร้อยละ 95.4



รูปที่ 4.16 ตัวอย่างการปรับปรุงแสงสว่างด้วยเทคนิค MSRCR
 ก) ปรับปรุงภาพที่มีแสงสว่างน้อยให้สว่างมากขึ้น
 ข) ปรับปรุงภาพที่มีแสงมากให้แสงสว่างลดน้อยลง



รูปที่ 4.17 ตัวอย่างการปรับปรุงแสงสว่างด้วยเทคนิค MSRCR
 ก) ปรับปรุงภาพที่มีแสงสว่างน้อยให้สว่างมากขึ้น
 ข) ปรับปรุงภาพที่มีแสงมากให้แสงสว่างลดน้อยลง



รูปที่ 4.18 เปรียบเทียบการตรวจจับใบหน้าด้วยวิธีการคัดแยกสีผิว และการประยุกต์ใช้ MSRCR ร่วมกับการคัดแยกสีผิว

4.5 ผลการตรวจจับดวงตาและการตรวจจับอาการง่วงนอน













4.5.1 ผลการตรวจจับดวงตา

ผลการตรวจจับดวงตาที่จะนำเสนอในหัวข้อนี้ ผู้วิจัยจะใช้ข้อมูลภาพที่ได้จากการติดตั้งกล้องเว็บแคม Logitech HD Pro Webcam C920 ทั้งหมด 9,000 ภาพแต่ละภาพมีขนาด 1920×1080 พิกเซล โดยเป็นการเก็บข้อมูลภาพจากการใช้งานจริงจากผู้ทดสอบ 3 คนคนละ 3,000 ภาพ จากนั้นจะนำภาพที่ได้มาทำการประมวลผลในห้องปฏิบัติการวิจัยเพื่อทำการตรวจจับใบหน้าโดยวิธีการคัดแยกสีผิว และทำการตรวจจับดวงตาด้วยวิธีการที่แตกต่างกัน 3 วิธีคือ วิธีการขีดเริ่มเปลี่ยน วิธีการ Integral Projection และวิธี Template Matching โดยมีผลการทดลองแสดงในตารางที่ 4.3

ตารางที่ 4.3 เปรียบเทียบร้อยละความถูกต้องระหว่างวิธีตั้งการ Threshold วิธีการ Integral Projection และวิธี Template Matching


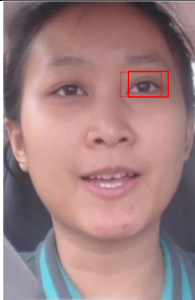
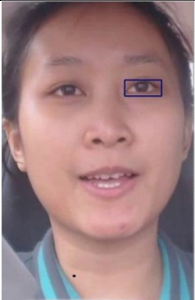

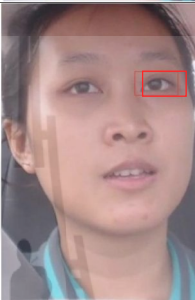
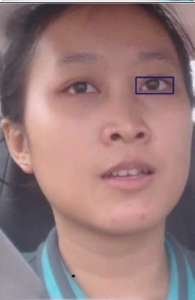

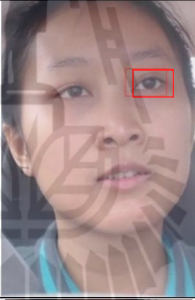
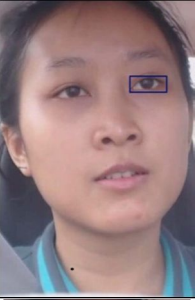
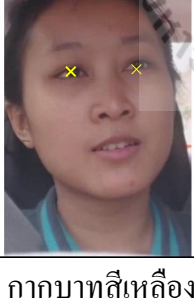


วิธีการ	ผู้ทดสอบที่ 1 3,000 ภาพ	ผู้ทดสอบที่ 2 3,000 ภาพ	ผู้ทดสอบที่ 3 3,000 ภาพ	ร้อยละความถูกต้อง
1. Threshold	93.73	97.57	93.07	94.79
2. Integral Projection	98.27	97.77	98.53	98.19
3. Template Matching	98.80	95.73	98.37	97.63

ตารางที่ 4.4 ผลการทดลอง บุคคลที่ 1

เฟรมที่	วิธีการ Threshold	วิธีการ Integral Projection	วิธีการ Template Matching
13			
401			
1204			
2324			



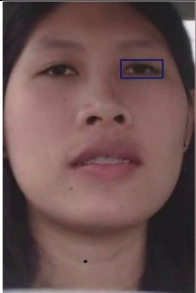
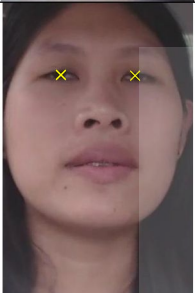


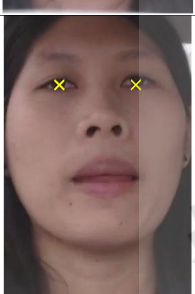

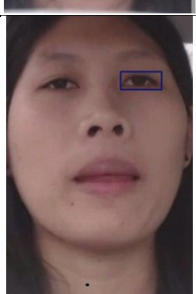

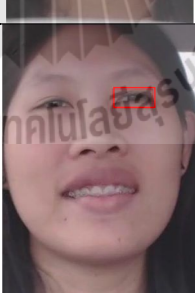

*หมายเหตุ กากบาทสีเหลืองแทนผลการตรวจจับวิธีการ Threshold
 สีเหลี่ยมสีแดงแทนผลการตรวจจับวิธีการ Integral Projection
 สีเหลี่ยมน้ำเงินแทนผลการตรวจจับวิธีการ Template Matching

ตารางที่ 4.5 ผลการทดลอง บุคคลที่ 2

เฟรมที่	วิธีการ Threshold	วิธีการ Integral Projection	วิธีการ Template Matching
446			
678			
1073			
2877			

*หมายเหตุ กากบาทสีเหลืองแทนผลการตรวจจับวิธีการ Threshold
 สีเหลี่ยมสีแดงแทนผลการตรวจจับวิธีการ Integral Projection
 สีเหลี่ยมน้ำเงินแทนผลการตรวจจับวิธีการ Template Matching

ตารางที่ 4.6 ผลการทดลอง บุคคลที่ 3

เฟรมที่	วิธีการ Threshold	วิธีการ Integral Projection	วิธีการ Template Matching
512			
847			
1212			
2162			

*หมายเหตุ กากบาทสีเหลืองแทนผลการตรวจจับวิธีการ Threshold
 สีเหลี่ยมสีแดงแทนผลการตรวจจับวิธีการ Integral Projection
 สีเหลี่ยมน้ำเงินแทนผลการตรวจจับวิธีการ Template Matching

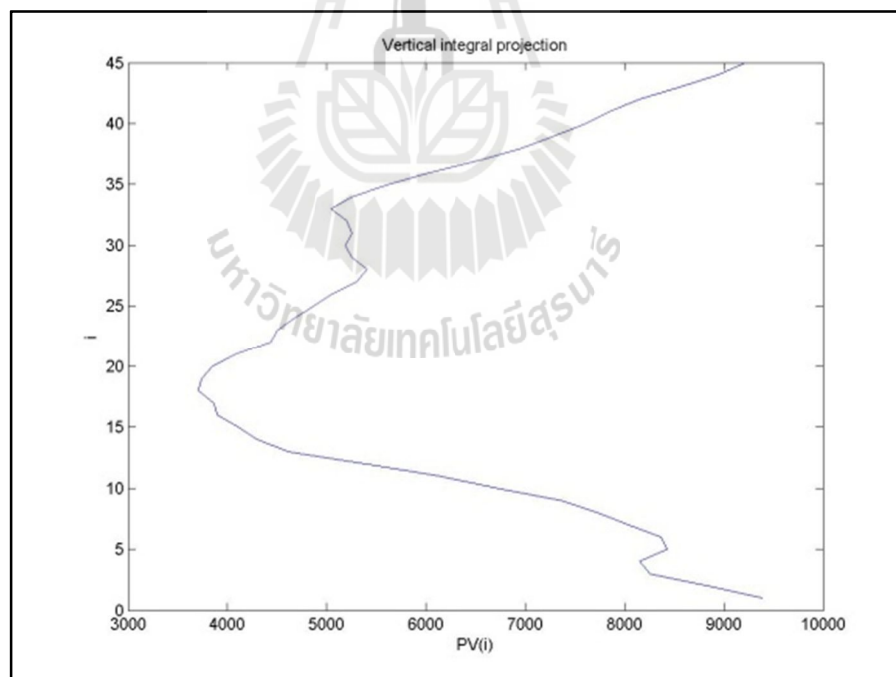
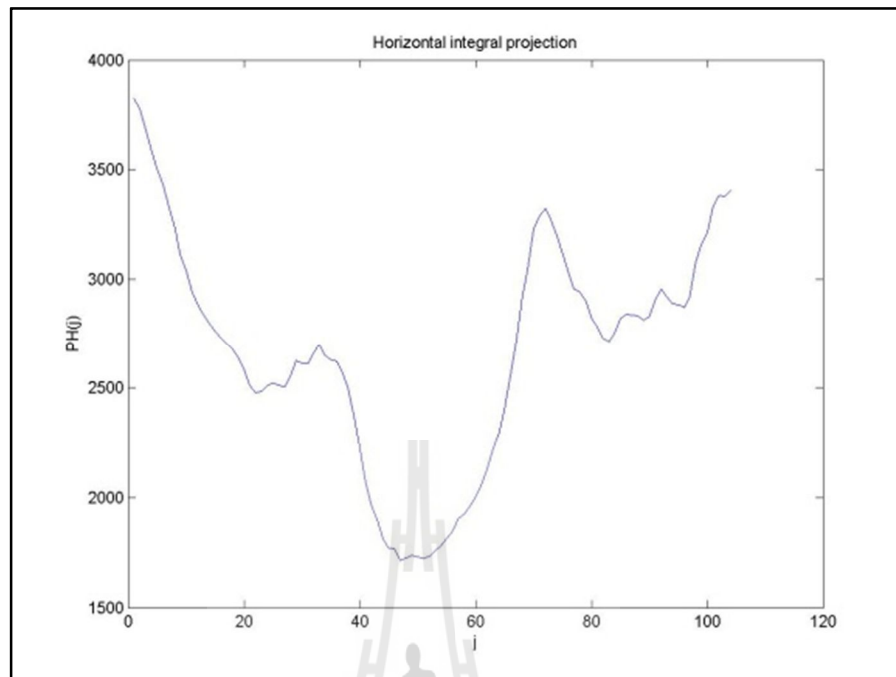
จากตารางที่ 4.3 จะเห็นว่าวิธีการ Threshold มีร้อยละการตรวจจับถูกต้องร้อยละ 94.79 วิธีการ Integral Projection มีร้อยละการตรวจจับถูกต้องร้อยละ 98.19 และวิธีการ Template Matching มีร้อยละการตรวจจับถูกต้องร้อยละ 97.63 โดยที่ใช้ฐานข้อมูลเดียวกันดังตารางที่ 4.4 ถึง

ตารางที่ 4.6 จะเห็นว่าการตรวจจับดวงตาโดยใช้วิธีการ Integral Projection และ Template Matching มีร้อยละการตรวจจับถูกต้องใกล้เคียงกันและ มากกว่าวิธีการ Threshold

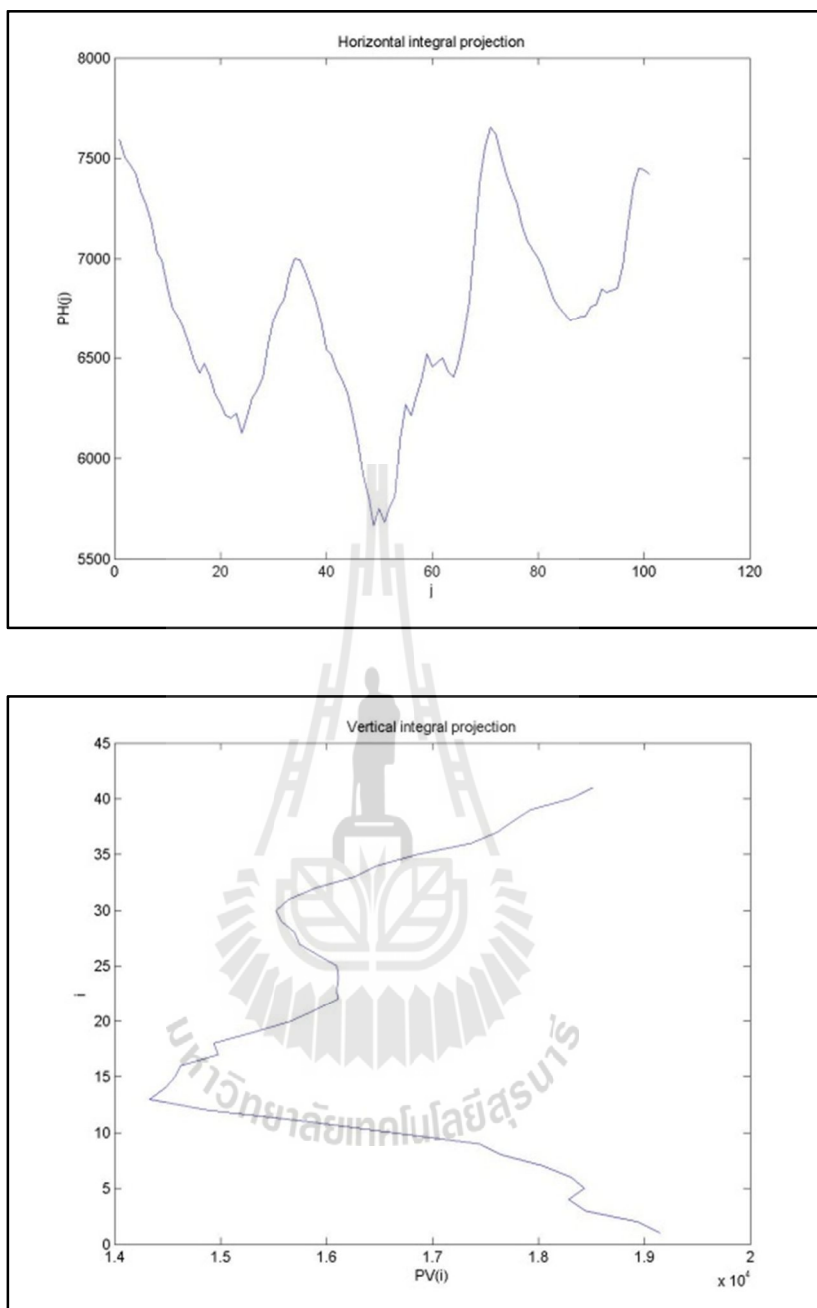
วิธีการ Threshold สามารถตรวจจับได้กับภาพที่มีแสงสว่างคงที่เนื่องจากการตั้ง Threshold ที่ค่าค่า ๆ หนึ่ง ซึ่งเมื่อภาพที่ได้จากสภาวะจริงจะมีแสงสว่างเปลี่ยนแปลงตลอดเวลา จะทำให้การตั้งค่า Threshold ไว้ที่ค่าเดียวส่งผลให้ผลการตรวจจับผิดพลาด และวิธีการนี้จะสามารถตรวจจับได้เฉพาะเวลาที่ผู้ทดสอบลืมตาหรือตาเปิด แต่ในสภาวะใช้งานจริงผู้ทดสอบมีทั้งตาเปิดและตาปิด วิธีการนี้จึงไม่เหมาะสำหรับการตรวจจับดวงตาในขณะที่ผู้ทดสอบขับรถ

วิธีการ Integral Projection เป็นการหาผลรวมของค่าความเข้มของพิกเซลในแนวแกนตั้ง และแนวนอนไม่ว่าแสงสว่างจะมากหรือน้อยบริเวณของม่านตาจะมีผลรวมของค่าความเข้มต่ำกว่าบริเวณอื่น ดังแสดงในรูปที่ 4.19 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 คือผลรวมค่าความเข้มของพิกเซลในแนวแนวนอน และแกน Y คือ พิกเซลในแนวแนวนอน) สังเกตได้ว่าบริเวณที่มีผลรวมของค่าความเข้มต่ำกว่าบริเวณอื่นของทุก ๆ กราฟจะเป็นบริเวณของม่านตา แต่เมื่อทดสอบในขณะที่ผู้ทดสอบหลับตาด้วยวิธีการนี้ จะมีการตรวจจับผิดพลาด

วิธีการ Template Matching เป็นการเทียบภาพที่นำมาทดสอบกับแม่แบบตาปิด และแม่แบบตาเปิด จะสามารถตรวจจับได้ทั้งผู้ทดสอบตาเปิด-ปิด และวิธีการนี้สามารถตรวจจับได้ภายใต้แสงสว่างที่เปลี่ยนแปลง จึงได้นำเอาวิธีการนี้มาใช้ในการตรวจจับดวงตา แล้วคาดเดาอาการง่วงนอนต่อไป



ก)



ข)

รูปที่ 4.19 เปรียบเทียบกราฟ Integral Projection ของภาพดวงตา

ก) กราฟของ Horizontal Integral Projection และ
Vertical Integral Projection ของภาพที่มีแสงน้อย

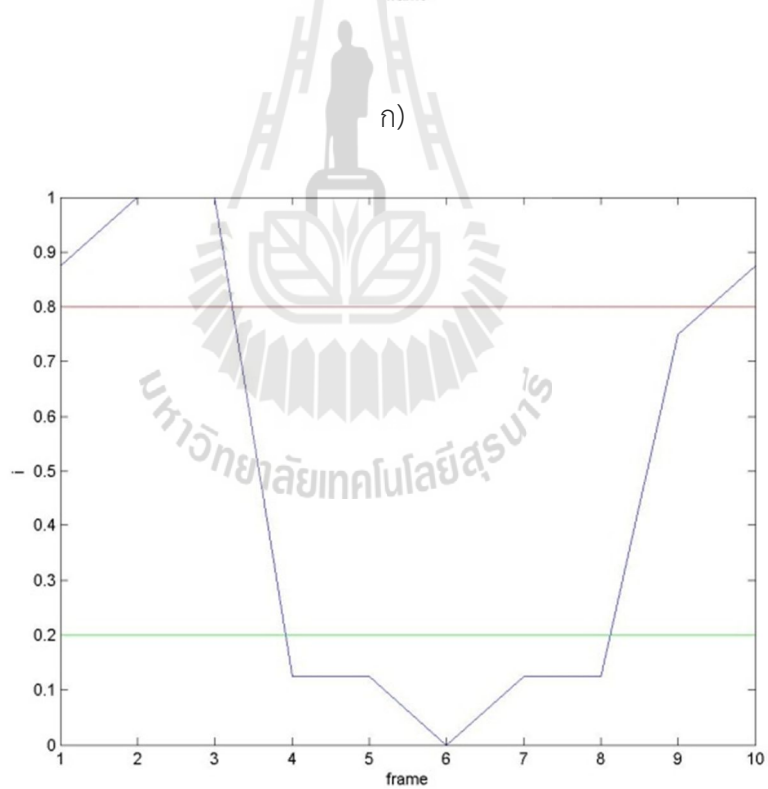
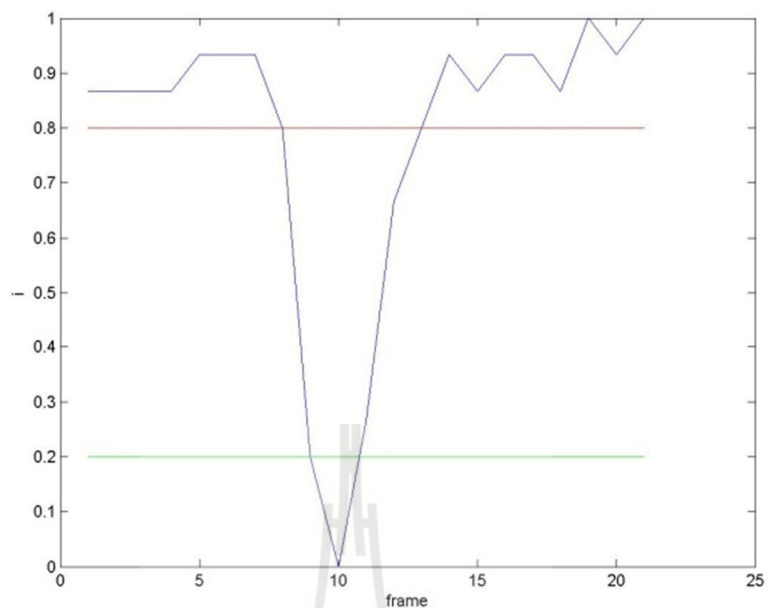
ข) กราฟของ Horizontal Integral Projection และ
Vertical Integral Projection ของภาพที่มีแสงมาก

4.5.2 ผลการตรวจจับอาการง่วงนอน

หัวข้อนี้จะนำเสนอการตรวจจับดวงด้วยวิธี Template Matching สำหรับใช้ในการตรวจจับอาการง่วงนอน ร่วมกับวิธีการคำนวณค่า PERCLOS สามารถคำนวณได้จากกราฟของการปิดตาในการกระพริบตาแต่ละครั้งดังแสดงในรูปที่ 4.20 รูป ก) เป็นกราฟร้อยละของการปิดตาเทียบกับเฟรมภาพขณะตื่นตัว และ ข) เป็นกราฟร้อยละของการปิดตาเทียบกับเฟรมภาพขณะง่วงนอน โดยกราฟเส้นสีแดงคือร้อยละการปิดตาร้อยละ 80 และกราฟเส้นที่เขียวคือร้อยละการปิดตาร้อยละ 20 ค่า PERCLOS สามารถคำนวณได้ดังสมการที่ (3-17) ในการตรวจจับดวงตาด้วยวิธี Template Matching จะใช้การประมาณค่า PERCLOS ใน [26] ได้มีการประมาณค่า PERCLOS เทียบกับเวลา เมื่อค่า PERCLOS มากกว่า 0.4 เทียบได้กับการหลับตานานเป็นเวลา 3 วินาที ดังตารางที่ 4.7 โดยผู้วิจัยได้ใช้ฐานข้อมูลจากเว็บแคม Logitech HD Pro Webcam C920 ที่มีเฟรมภาพขนาด 1920x1080 พิกเซลซึ่งประกอบด้วยวิดีโอที่ผู้จับขี้ตื่นตัวหรือรู้สึกตัว (Wake) และที่ผู้จับขี้เมื่อยล้า (Fatigue) หรือง่วงนอน (Sleepiness) แต่ละวิดีโอที่มีความยาว 5 นาทีโดยมีอัตราเฟรม 30 เฟรมต่อวินาที จากผู้ทดสอบ 3 บุคคล ดังรูปที่ 4.21 และรูปที่ 4.22

ตารางที่ 4.7 ผลการเปรียบเทียบค่า PERCLOS กับเวลาในอิริยาบถที่ปกติ และง่วงนอนหรือเมื่อยล้า ในบทความ [25]

Test Number	PERCLOS (Conscious)	The Longest Continued Closure Time (Conscious)	PERCLOS (Fatigue)	The Longest Continued Closure Time (Fatigue)
1	8.49%	0.266 Sec	41.10%	3.361 Sec
2	6.25%	0.402 Sec	53.71%	3.733 Sec
3	6.75%	0.560 Sec	49.41%	3.162 Sec



ข)

รูปที่ 4.20 กราฟร้อยละของการปิดตาเทียบกับเฟรมเวลา

ก) กราฟร้อยละของการปิดตาเทียบกับเฟรมเวลาขณะตื่นตัว

ข) กราฟร้อยละของการปิดตาเทียบกับเฟรมเวลาขณะง่วงนอน



ก) บุคคลที่ 1



ข) บุคคลที่ 2



ค) บุคคลที่ 3

รูปที่ 4.21 ฐานข้อมูลที่ใช้ในการทดสอบต้นตัว



ก) บุคคลที่ 1



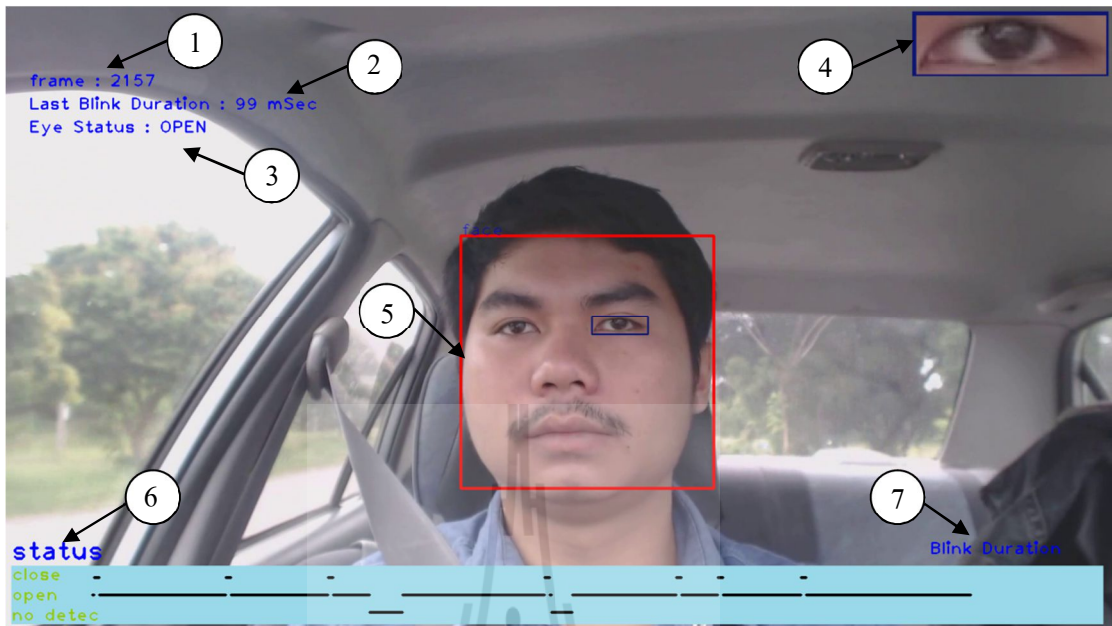
ข) บุคคลที่ 2



ค) บุคคลที่ 3

รูปที่ 4.22 ฐานข้อมูลที่ใช้ในการทดสอบขณะง่วงนอน

ในงานวิจัยนี้ได้ทดสอบระบบตรวจจับอาการง่วงนอนกับฐานข้อมูลที่ผู้ทดสอบตื่นตัวขณะขับรถ และผู้ทดสอบง่วงนอนที่ไม่ได้มีการขับรถ เนื่องจากในการขับรถในขณะที่ง่วงนอนอาจจะทำให้เกิดอันตรายต่อผู้ทดสอบได้ จึงได้จัดทำฐานข้อมูลในห้องทดลองแทนการทดสอบด้วยการขับขี่รถยนต์จริง และงานวิจัยนี้พัฒนาโปรแกรมด้วยโปรแกรม Microsoft Visual Studio 2010 และไลบรารี OpenCV 2.4.9 โดยมีหน้าต่างแสดงค่าต่างๆ ดังรูปที่ 4.23



รูปที่ 4.23 หน้าต่างแสดงค่าต่างๆในการตรวจจับอาการง่วงนอน

โดยมีรายละเอียดดังต่อไปนี้

หมายเลข 1 คือ ลำดับเฟรม

หมายเลข 2 คือ ระยะเวลาที่ใช้ในการกระพริบตาค่าสุด

หมายเลข 3 คือ สถานะของตา

หมายเลข 4 คือ ผลการตรวจจับดวงตา

หมายเลข 5 คือ บริเวณของใบหน้า

หมายเลข 6 คือ กราฟแสดงสถานะของตา

หมายเลข 7 คือ ระยะเวลาที่หลับตา

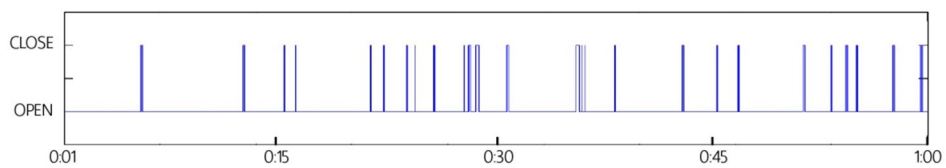
เมื่อฐานข้อมูลดังกล่าวมาทดสอบจะได้ผลการทดสอบดังนี้

บุคคลที่ 1

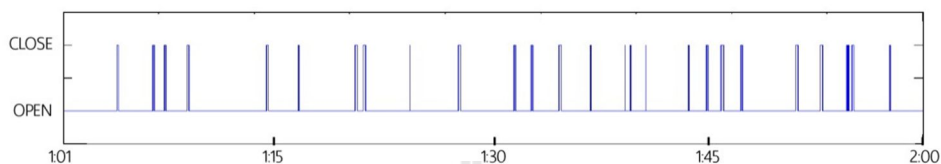
1. ผู้ทดสอบตื่นตัว



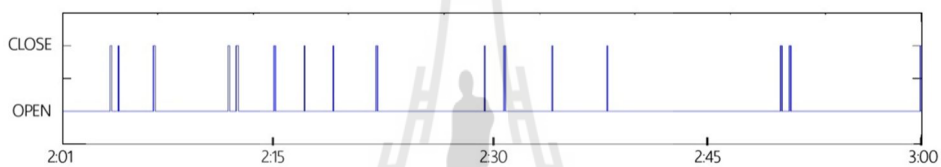
รูปที่ 4.24 ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 1 สถานะตื่นตัว



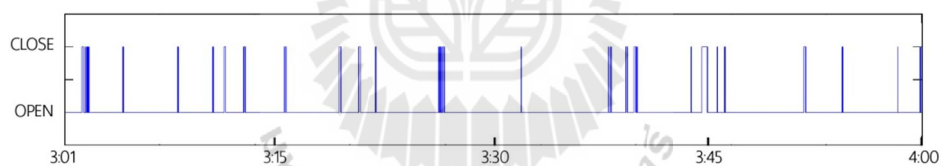
ก) นาทีที่ 1



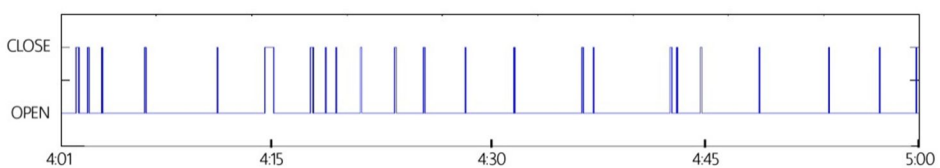
ข) นาทีที่ 2



ค) นาทีที่ 3



ง) นาทีที่ 4



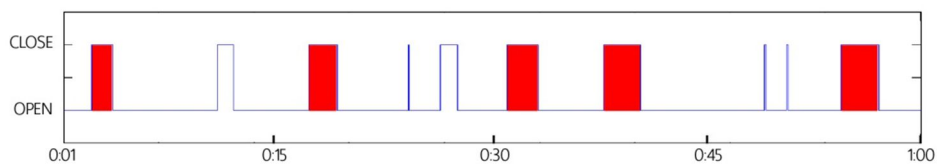
จ) นาทีที่ 5

รูปที่ 4.25 ผลการทดสอบการตรวจจับการง่วงนอนของบุคคลที่ 1 สถานะตื่นตัว

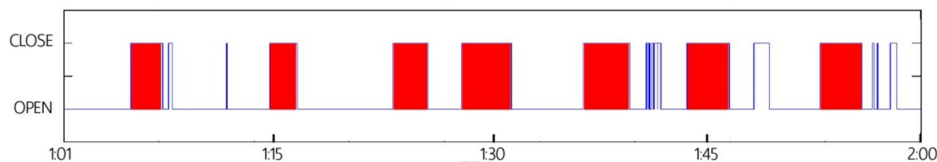
2. ผู้ทดสอบง่วงนอน



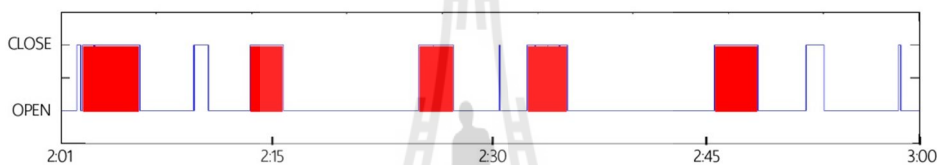
รูปที่ 4.26 ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 1 สถานะง่วงนอน



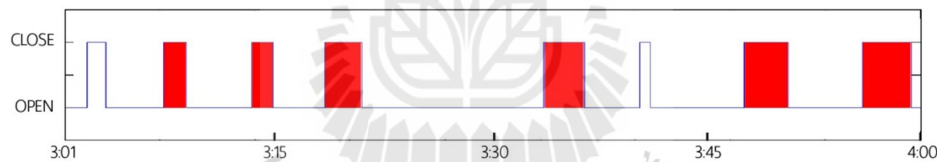
ก) นาทีที่ 1



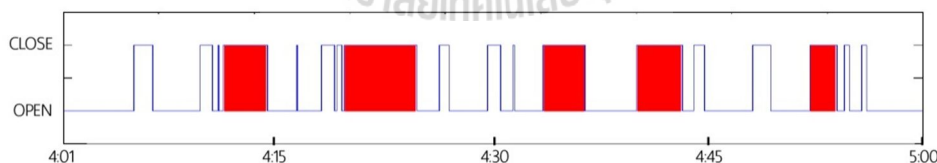
ข) นาทีที่ 2



ค) นาทีที่ 3



ง) นาทีที่ 4



จ) นาทีที่ 5

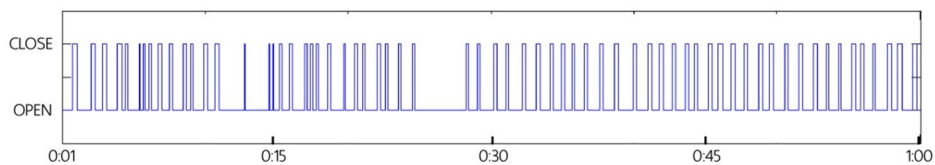
รูปที่ 4.27 ผลการทดสอบการตรวจจับการ่วงนอนของบุคคลที่ 1 สถานะ่วงนอน

บุคคลที่ 2

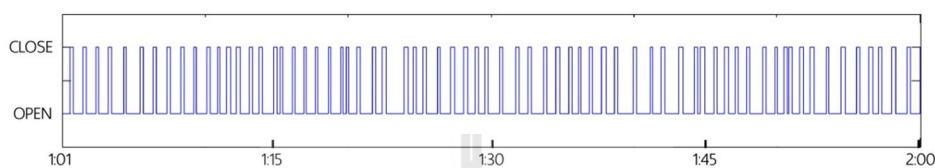
1. ผู้ทดสอบตื่นตัว



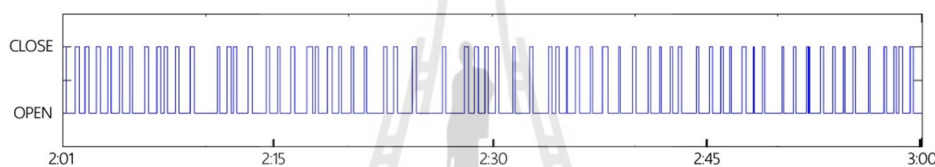
รูปที่ 4.28 ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 2 สถานะตื่นตัว



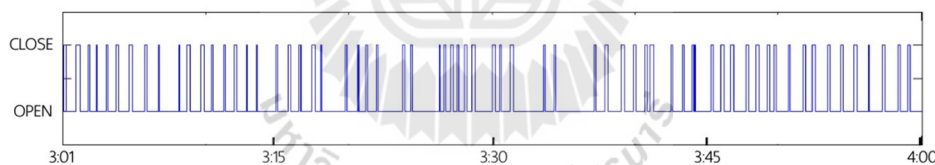
ก) นาทีที่ 1



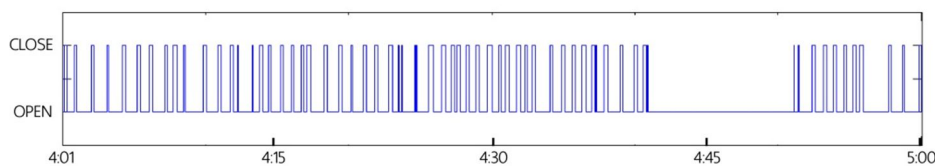
ข) นาทีที่ 2



ค) นาทีที่ 3



ง) นาทีที่ 4



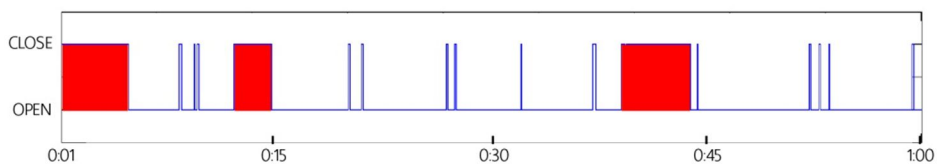
จ) นาทีที่ 5

รูปที่ 4.29 ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 2 สถานะตื่นตัว

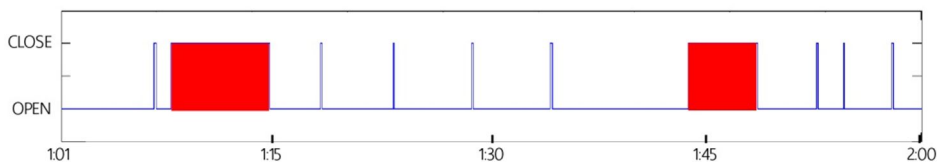
2. ผู้ทดสอบง่วงนอน



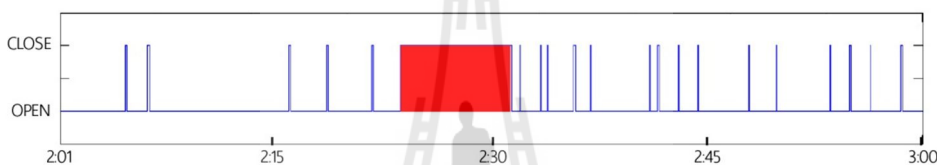
รูปที่ 4.30 ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 2 สถานะง่วงนอน



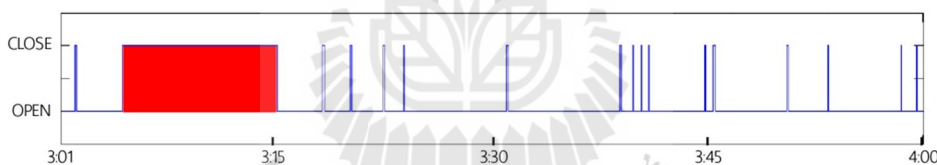
ก) นาทีที่ 1



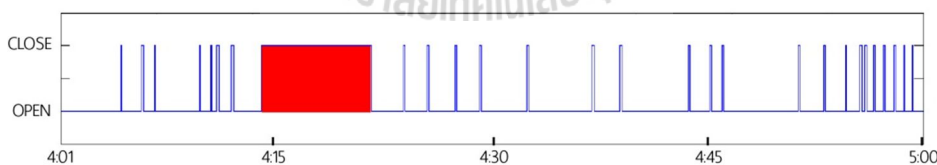
ข) นาทีที่ 2



ค) นาทีที่ 3



ง) นาทีที่ 4



จ) นาทีที่ 5

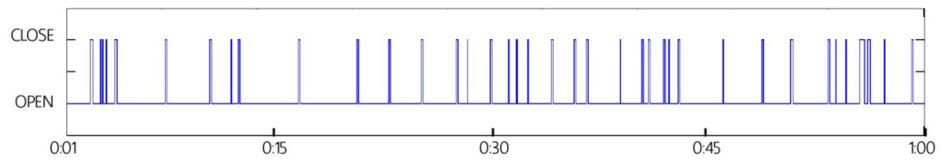
รูปที่ 4.31 ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 2 สถานะง่วงนอน

บุคคลที่ 3

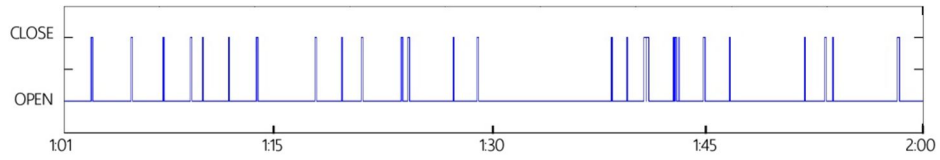
1. ผู้ทดสอบตื่นตัว



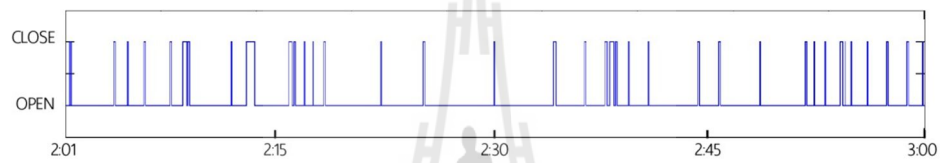
รูปที่ 4.32 ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 3 สถานะตื่นตัว



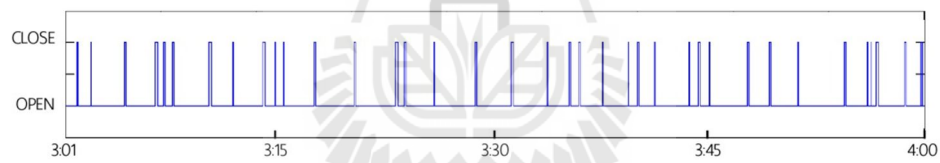
ก) นาทีที่ 1



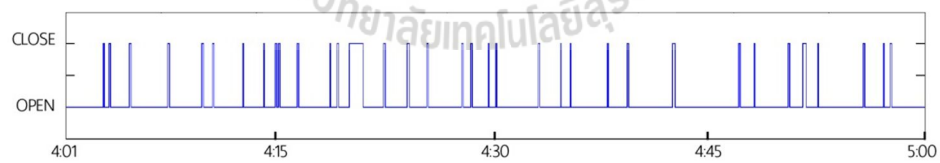
ข) นาทีที่ 2



ค) นาทีที่ 3



ง) นาทีที่ 4



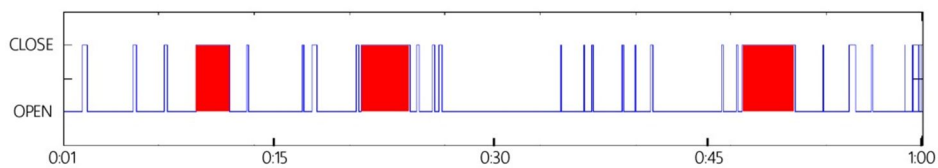
จ) นาทีที่ 5

รูปที่ 4.33 ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 3 สถานะตื่นตัว

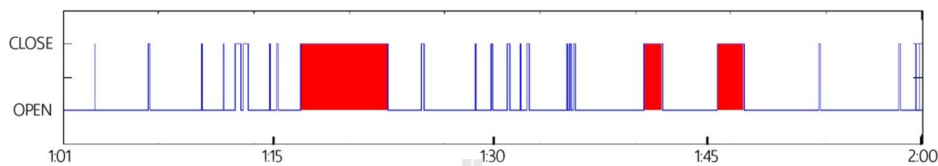
2. ผู้ทดสอบง่วงนอน



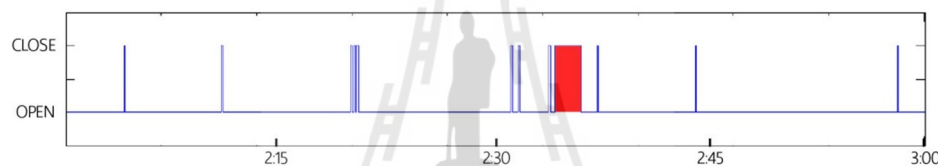
รูปที่ 4.34 ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 3 สถานะง่วงนอน



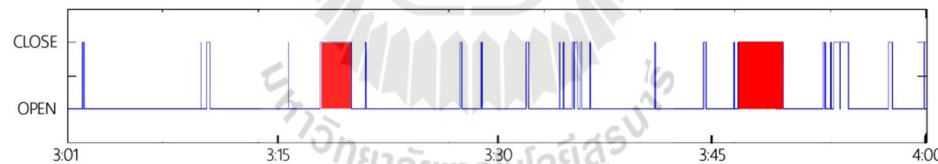
ก) นาทีที่ 1



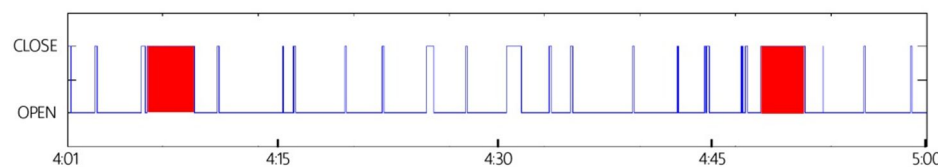
ข) นาทีที่ 2



ค) นาทีที่ 3



ง) นาทีที่ 4



จ) นาทีที่ 5

รูปที่ 4.35 ผลการทดสอบการตรวจจับอาการง่วงนอนของบุคคลที่ 3 สถานะง่วงนอน

ผลการทดลองพบว่า เมื่อทดสอบกับฐานข้อมูลที่ผู้ทดสอบตื่นตัวหรือไม่มีอาการง่วงนอนแสดงได้ดังรูปที่ 4.24 รูปที่ 4.28 และ รูปที่ 4.32 และกราฟสถานะของการเปิด-ปิดตา ดังรูปที่ 4.25 รูปที่ 4.29 และรูปที่ 4.33 โดยจะพบว่าผู้ทดสอบมีการปิดตาเป็นระยะเวลาสั้น ๆ (กระพริบตา) ระบบจึงไม่แจ้งเตือนและเมื่อทดสอบกับฐานข้อมูลที่ผู้ทดสอบง่วงนอนจะได้ผลการทดลองดังรูปที่ 4.26 รูปที่ 4.30 และ รูปที่ 4.34 และกราฟสถานะของการเปิด-ปิดตา ดังรูปที่ 4.27 รูปที่ 4.31 และ รูปที่ 4.35 จะเห็นว่ามี การปิดตาเป็นเวลานานและระบบมีการแจ้งเตือน (แถบสีแดง)

เมื่อนำผลการทดลองระหว่างผู้ทดสอบตื่นตัวกับผู้ทดสอบง่วงนอน มาเปรียบเทียบร้อยละการปิดตาและเปิดตาพบว่าฐานข้อมูลที่ผู้ทดสอบง่วงนอนมีการหลับตามากกว่าฐานข้อมูลที่ตื่นตัวดังแสดงในตารางที่ 4.8

ตารางที่ 4.8 เปรียบเทียบร้อยละการปิด – เปิดดวงตา ของผู้ทดสอบตื่นตัวกับผู้ทดสอบง่วงนอน

ฐานข้อมูล	ผู้ทดสอบตื่นตัว			ผู้ทดสอบง่วงนอน		
	ตาปิด (ร้อยละ)	ตาเปิด (ร้อยละ)	ผิดพลาด (ร้อยละ)	ตาปิด (ร้อยละ)	ตาเปิด (ร้อยละ)	ผิดพลาด (ร้อยละ)
บุคคลที่ 1	3.79	85.01	11.20	31.42	54.82	13.76
บุคคลที่ 2	19.07	48.50	32.43	19.34	80.55	0.11
บุคคลที่ 3	6.22	77.39	16.39	15.64	75.30	9.06

ตารางที่ 4.9 ผลการจำแนกความผิดพลาดในกรณีต่างๆ ที่เกิดขึ้นตามตารางที่ 4.8

ฐานข้อมูล	ผู้ทดสอบตื่นตัว			ผู้ทดสอบง่วงนอน		
	ถูกต้อง (ร้อยละ)	ผิดพลาด (ร้อยละ)		ถูกต้อง (ร้อยละ)	ผิดพลาด (ร้อยละ)	
		หน้า ✘	หน้า ✓ ตา ✘		หน้า ✘	หน้า ✓ ตา ✘
บุคคลที่ 1	88.8	9.36	1.84	86.24	13.74	0.02
บุคคลที่ 2	67.57	27.79	4.64	99.89	0.10	0.01
บุคคลที่ 3	83.61	14.30	2.09	90.94	9.03	0.03

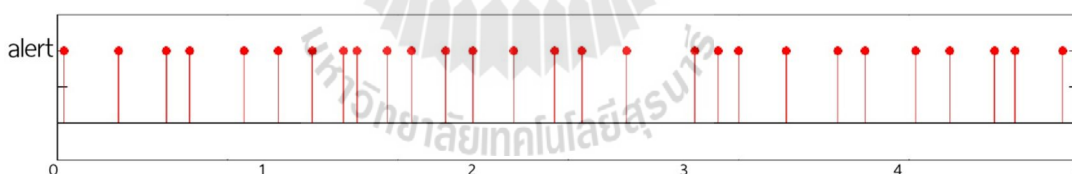
*หมายเหตุ หน้า ✘ แทนการตรวจจับใบหน้าผิดพลาด

หน้า ✓ ตา ✘ แทนการตรวจจับใบหน้าถูกต้องและตรวจจับดวงตาผิดพลาด

จากตารางที่ 4.8 พบว่าการทดสอบการคาดเดาอาการง่วงนอนของผู้ทดสอบทั้ง 3 บุคคลมีร้อยละการปิดตาเมื่อผู้ทดสอบง่วงนอนมีมากกว่าผู้ทดสอบตื่นตัว ซึ่งสอดคล้องกับความ เป็นจริงเมื่อมีอาการง่วงนอนจะมีการหลับตาเป็นเวลานาน ส่วนเวลาที่ผู้ทดสอบไม่มีอาการ ง่วงนอนจะเพียงแค่กระพริบตาเป็นช่วงเวลาสั้น ๆ เท่านั้น และความผิดพลาดที่เกิดขึ้นในการ ทดสอบเกิดจากการตรวจจับดวงตาผิดพลาด และการตรวจจับใบหน้าผิดพลาดส่งผลให้การ ตรวจจับดวงตาผิดพลาดด้วย เนื่องจากการทดสอบผู้ทดสอบมีการเคลื่อนไหว เช่นการหันหน้า ไปด้านข้างเพื่อมองกระจกหลัง และกระจกข้าง ซึ่งจะทำให้เกิดการตรวจจับผิดพลาดสามารถ จำแนกได้ดังแสดงในตารางที่ 4.9

ในการแจ้งเตือนผู้ทดสอบเมื่อพบว่าผู้ทดสอบมีอาการง่วงนอน ระบบจะทำการแจ้ง เตือนเมื่อผู้ทดสอบหลับตาช่วงเวลาหนึ่ง ในงานวิจัย [25] ได้เปรียบเทียบค่า PERCLOS ที่คาดเดาว่า ผู้ทดสอบง่วงนอนเท่ากับ 0.4 หรือร้อยละ 40 ซึ่งเทียบได้กับการหลับตาเป็นเวลา 3 วินาทีดังตารางที่ 4.7 ดังนั้นผู้วิจัยได้ออกแบบให้ระบบมีการแจ้งเตือนก่อนที่ผู้ทดสอบจะมีอาการง่วงนอน นั่นคือเมื่อ ผู้ทดสอบมีการหลับตาเป็นเวลา 1.5 วินาที เพื่อที่จะให้ผู้ทดสอบรู้สึกตัวหรือเปลี่ยนอิริยาบถก่อนที่ จะง่วงนอน และระบบได้มีการแจ้งเตือนดังนี้

- บุคคลที่ 1 มีการแจ้งเตือนทั้งหมด 28 ครั้ง ดังรูปที่ 4.36

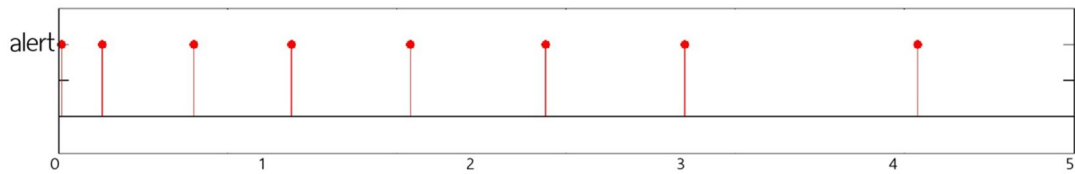


รูปที่ 4.36 การแจ้งเตือนอาการง่วงนอนของบุคคลที่ 1

จากการทดลองดังกล่าวพบว่ามีแจ้งเตือนที่เวลาดังนี้

0:02 0:18 0:32 0:39 0:55 1:05 1:15 1:24 1:28 1:37 1:44 1:54 2:02 2:14 2:26 2:34 2:47 3:07 3:14
3:20 3:34 3:49 3:57 4:12 4:22 4:35 4:41 และ 4:55 ตามลำดับ

- บุคคลที่ 2 มีการแจ้งเตือนทั้งหมด 8 ครั้ง ดังรูปที่ 4.37



รูปที่ 4.37 การแจ้งเตือนอาการง่วงนอนของบุคคลที่ 2

จากการทดลองดังกล่าวพบว่ามีแจ้งเตือนที่เวลาดังนี้

0:01 0:13 0:40 1:09 1:44 2:24 3:05 และ 4:14 ตามลำดับ

- บุคคลที่ 3 มีการแจ้งเตือนทั้งหมด 11 ครั้ง ดังรูปที่ 4.38



รูปที่ 4.38 การแจ้งเตือนอาการง่วงนอนของบุคคลที่ 3

จากการทดลองดังกล่าวพบว่ามีแจ้งเตือนที่เวลาดังนี้

0:10 0:22 0:47 1:17 1:41 1:46 2:35 3:18 3:47 4:06 และ 4:49 ตามลำดับ

จากผลการทดลองในตารางที่ 4.8 ทำการวิเคราะห์ความผิดพลาดที่เกิดขึ้นในการตรวจจับอาการง่วงนอนโดย

4.6 บทสรุป

บทนี้ได้กล่าวถึงผลการทดลองการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมและการตรวจจับใบหน้าโดยการคัดแยกสีผิว จากผลการทดลองจะเห็นว่า การตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมมีร้อยละความถูกต้องน้อยกว่าการตรวจจับใบหน้าโดยการคัดแยกสีผิวอาจเกิดจากจำนวนภาพที่นำมาใช้ในการรู้จำของเครือข่ายประสาทเทียมยังไม่เพียงพอ การเพิ่มร้อยละ

ความถูกต้องของการตรวจจับอาจทำได้โดยการเก็บรวบรวมข้อมูลใบหน้าที่ใช้ในการรู้จำให้มากขึ้น ส่วนการตรวจจับใบหน้าโดยการคัดแยกสีผิวเหมาะสำหรับการตรวจจับใบหน้าของผู้ขับขี่รถยนต์ เนื่องจากใบหน้าของผู้ขับขี่รถยนต์บางครั้งอาจจะไม่เป็นใบหน้าตรงซึ่งการใช้สีผิวในการตรวจจับใบหน้าสามารถที่จะตรวจจับได้ จากการทดลองพบว่าใบหน้าของผู้ขับขี่รถยนต์โดยทั่วไปนั้นจะมีการเปลี่ยนแปลงของแสงสว่างตลอดเวลาจึงต้องมีการนำเอาเทคนิค Retinex มาช่วยในการปรับแสงสว่าง ผลการทดลองได้แสดงให้เห็นว่า Retinex สามารถช่วยให้การตรวจจับใบหน้าโดยการคัดแยกสีผิวมีความถูกต้องมากยิ่งขึ้น

ผลการตรวจจับดวงตาด้วย วิธีการ 3 วิธี ได้แก่ วิธีการ Threshold วิธีการ Integral Projection และวิธีการ Template Matching ผลการทดลองพบว่าวิธีการ Integral Projection และวิธีการ Template Matching มีร้อยละของความถูกต้องในการตรวจจับสูงกว่าวิธีการตั้ง Threshold ทั้งนี้อาจเกิดจากฐานข้อมูลที่ใช้เป็นภาพวีดิทัศน์จากสถานะการใช้งานจริงขณะขับรถ ซึ่งมีแสงสว่างเปลี่ยนแปลงตลอดเวลา ทำให้วิธีการตั้ง Threshold และวิธีการ Integral Projection ไม่เหมาะกับการตรวจจับดวงตาในสถานการณ์ดังกล่าว ส่วนวิธีการ Template Matching ผู้วิจัยพบว่าวิธีการนี้เหมาะสมกับการตรวจจับดวงตาในภาพที่มีแสงสว่างที่เปลี่ยนแปลงตลอดเวลา รวมถึงกรณีที่ผู้ทดสอบมีการหลับตาหรือลืมตา เมื่อนำการตรวจจับดวงตาด้วยวิธีการนี้มาคาดเดาอาการง่วงนอน โดยใช้ผู้ทดสอบจำนวน 3 คนด้วยการคำนวณ PERCLOS โดยทดสอบในสถานะที่ผู้ทดสอบตื่นตัว และผู้ทดสอบง่วงนอน ผลการทดสอบโดยการคำนวณ PERCLOS ได้แสดงให้เห็นว่ามีความสอดคล้องกับสถานะของผู้ทดสอบในขณะนั้น โดยสามารถจะนำไปประยุกต์ใช้งานได้จริง

บทที่ 5

สรุป

5.1 สรุปผลงานวิจัย

งานวิจัยนี้ได้นำเสนอการตรวจจับอาการง่วงนอนของคนขับรถ ซึ่งผู้ขับขี่มีอาการง่วงนอน ซึ่งนับว่าเป็นสาเหตุอันดับต้น ๆ ของอุบัติเหตุทางรถยนต์ที่เกิดขึ้นในท้องถนนของประเทศไทย ระบบตรวจจับอาการง่วงนอนนี้ได้ออกแบบเป็นส่วน ๆ ดังนี้

1) การตรวจจับใบหน้า ในงานวิจัยนี้ได้นำเสนอการตรวจจับใบหน้า 2 วิธีได้แก่ การตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียม และการตรวจจับใบหน้าโดยการคัดแยกสีผิว การตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมจะนำฐานข้อมูลภาพใบหน้าและภาพที่ไม่ใช่ใบหน้า อย่างละ 100 ภาพมาดึงลักษณะเด่นด้วยวิธีการวิเคราะห์ส่วนประกอบสำคัญ (Principal Components Analysis: PCA) และการแปลงเวฟเล็ตแบบดิสครีต (Discrete Wavelet Transform: DWT) ก่อนที่จะนำมาให้เครือข่ายเทียมรู้จำและได้ค่าน้ำหนักเพื่อนำไปใช้ตรวจจับใบหน้าต่อไป และการตรวจจับใบหน้าโดยการคัดแยกสีผิวเป็นการตรวจจับใบหน้าที่ใช้ช่วงสีผิวที่อยู่ในปริภูมิสี RGB และ HSV โดยมีการนำการประมวลผลภาพดิจิทัล (Digital Image Processing: DIP) มาใช้ในการประมวลผล ซึ่งผลการทดลองแสดงให้เห็นว่าการตรวจจับใบหน้าโดยใช้การคัดแยกสีผิวเหมาะสมกับงานวิจัยนี้ เนื่องจากการตรวจจับใบหน้าโดยใช้สีผิวจะสามารถตรวจจับได้ใบหน้าได้ทุกขนาด ส่วนการตรวจจับใบหน้าโดยใช้เครือข่ายประสาทเทียมจะสามารถตรวจจับได้เฉพาะใบหน้าที่มีการรู้จำ ซึ่งมีข้อจำกัดในเรื่องของขนาด และการเอียงของใบหน้า ผู้วิจัยได้ทำการทดสอบกับฐานข้อมูล CIT BaoFace Essex Georgia Tech และฐานข้อมูลที่ผู้วิจัยจัดสร้างขึ้นเอง โดยพบว่าสามารถตรวจจับใบหน้าถูกต้องถึงร้อยละ 95.13 สาเหตุที่ทำให้การตรวจจับใบหน้าเกิดความผิดพลาดคือ แสงสว่างไม่ปกติ สีผิวจากหลายเชื้อชาติ พื้นหลังที่มีสีเหมือนสีผิว สาเหตุหลักที่ทำให้ตรวจจับผิดพลาดมากที่สุดคือ แสงสว่างไม่ปกติ ผู้วิจัยจึงได้นำเสนอการชดเชยแสงสว่างด้วยวิธี MSRCR เพื่อให้การตรวจจับใบหน้ามีความถูกต้องมากขึ้น ในการทดลองได้นำภาพใบหน้าที่มีการตรวจจับผิดพลาดมา 500 ภาพ สามารถปรับปรุงแสงสว่างทำให้การตรวจจับถูกต้องเพิ่มขึ้นจากค่าเฉลี่ยเดิมคือถูกต้องร้อยละ 77.45 เพิ่มขึ้นเป็นร้อยละ 95.40

2) การตรวจจับดวงตา หลังจากการตรวจจับใบหน้าโดยใช้การคัดแยกสีผิวจะได้บริเวณใบหน้า ในงานวิจัยนี้ได้นำเสนอการตรวจจับดวงตา 3 วิธีได้แก่ การตรวจจับดวงตาด้วยวิธีการตั้งค่าขีด

เริ่มเปลี่ยน (Threshold) การตรวจจับดวงตาด้วยวิธีอินทิกรัลโปรเจกชัน (Integral Projection) และการตรวจจับดวงตาด้วยวิธีการจับคู่แม่แบบ (Template Matching) เมื่อทดสอบกับฐานข้อมูลจริง ขณะขับรถพบว่า ผลการตรวจจับดวงตาโดยใช้วิธีการ Integral Projection วิธีการ Template Matching และวิธีการ Threshold มีร้อยละการตรวจจับถูกต้อง 98.19 97.63 และ 94.79 ตามลำดับ ซึ่งพบว่ามี 2 วิธีการที่ให้ผลการตรวจจับถูกต้องใกล้เคียงกัน ผู้วิจัยยังพบว่าวิธี Template Matching จะสามารถตรวจจับได้ในขณะที่ดวงตาเปิด และดวงตาปิด จึงได้เลือกวิธีการนี้ในการตรวจจับดวงตาเพื่อคาดเดาอาการง่วงนอนต่อไป

3) การตรวจจับอาการง่วงนอน ผู้วิจัยได้ทำการตรวจจับใบหน้าโดยใช้เทคนิคการคัดแยกสีผิว และทำการตรวจจับดวงตาด้วยวิธีการ Template Matching เพื่อคาดเดาอาการง่วงนอนจากระยะเวลาในเปิด - ปิดตาที่เปรียบเทียบกับค่า PERCLOS โดยทดสอบกับฐานข้อมูลที่ผู้วิจัยสร้างขึ้น ซึ่งประกอบด้วยวิดีโอที่ผู้ขับขี่ตื่นตัวหรือรู้สึกตัว (Wake) และวิดีโอที่ผู้ขับขี่เมื่อยล้า (Fatigue) หรือง่วงนอน (Sleepiness) แต่ละวิดีโอมีความยาว 5 นาทีโดยมีอัตราเฟรม 30 เฟรมต่อวินาที จากผู้ทดสอบ 3 คน เมื่อทดสอบระบบกับฐานข้อมูลดังกล่าวพบว่าสอดคล้องกับสถานะของผู้ทดสอบในขณะนั้น กล่าวคือ เมื่อทดสอบกับฐานข้อมูลของผู้ทดสอบง่วงนอนจะมีร้อยละการหลับตามากกว่าฐานข้อมูลของผู้ทดสอบตื่นตัว และเมื่อระบบตรวจจับพบว่าผู้ทดสอบมีอาการง่วงนอนระบบจะทำการแจ้งเตือนเพื่อให้ผู้ทดสอบรู้สึกตื่นตัวหรือเปลี่ยนอิริยาบถทำให้สามารถลดความเสี่ยงของการเกิดอุบัติเหตุบนท้องถนนจากการขับขีรถยนต์ได้

5.2 แนวทางต่อไปในอนาคต

- 1) การออกแบบและพัฒนาระบบเฝ้าติดตามอาการง่วงนอนสำหรับธุรกิจที่เกี่ยวข้องกับการขนส่งสินค้า หรือผู้โดยสาร
- 2) การพัฒนาอัลกอริทึมการตรวจจับอาการง่วงนอนนี้ด้วยอุปกรณ์เคลื่อนที่ (Mobile Devices) เช่น สมาร์ทโฟน (Smart Phone) และแท็บเล็ต (Tablet) ซึ่งเป็นอุปกรณ์เคลื่อนที่ที่ตัวประมวลผลมีประสิทธิภาพสูงใช้กันแพร่หลายและราคาไม่สูงมาก
- 3) การทดสอบการตรวจจับอาการง่วงนอนกับฐานข้อมูลกรณีของผู้ทดสอบมีอาการง่วงนอนในขณะที่ผู้ทดสอบขับขีรถยนต์จริง

รายการอ้างอิง

- Wang, Q., Yang, J., Ren, M., and Zheng, Y. (2006). **Driver fatigue detection :A survey**. 2006 6th World Congress on Intelligent Control and Automation. 2:8587-8591.
- Jobson, D. J., Rahman, Z., Woodell, and G. A., (1997). **A multiscale retinex for bridging the gap between color images and the human observation of scenes**. IEEE Transactions on Image Processing. 6(7): 965-976.
- ชนาวุฒิ ประกอบผล. (2552). **โครงข่ายประสาทเทียม Artificial Neural Networks**. วารสาร มฉก. วิชาการ 12 (24): 73-87.
- Ji, Q., Zhu, Z., and Lan, P., (2004). **Real-time nonintrusive monitoring and prediction of driver fatigue**. IEEE Transactions on Vehicular Technology. 53(4): 1052-1068.
- Tabrizi, P. R., and Zoroofi, R. A., (2009). **Drowsiness Detection Based on Brightness and Numeral Features of Eye Image**. Fifth International Conference on Intelligent Hiding and Multimedia Signal Processing. 1310-1313.
- Ribaric, S., Lovrencic, J., and Pavesic, N., (2010). **A neural-network-based system for monitoring driver fatigue**. 2010 15th IEEE Mediterranean Electrotechnical Conference. 1356-1361.
- Coetzer, R. C., and Hancke, G. P., (2011). **Eye detection for a real-time vehicle driver fatigue monitoring system**. IEEE Intelligent Vehicles Symposium(IV). 66-71
- Li, X., Seignez, E., and Loonis, P., (2012). **Reliability-based driver drowsiness detection using Dempster-Shafer theory**. 2012 12th International Conference on Control. 300-304.
- Garcia, I., Bronte, S., L. Bergasa, M., Almazan, J., and Yebes, J., (2012). **Vision-based drowsiness detector for real driving conditions**. 2012 Intelligent Vehicles Symposium Alcala de Henares. 618-623.
- Park, K. S., Park, R. H., and Kim, Y. G., (2011). **Face detection using the 3×3 block rank patterns of gradient magnitude images and a geometrical face model**. IEEE International Conference on Consumer Electronics (ICCE). 793-794.

- Paul, P. P., and Gavrilova, M., (2011). **PCA Based Geometric Modeling for Automatic Face Detection. International Conference on Computational Science and Its Applications.** 33-38.
- Chiang, M. L., and Lau, S. H., (2011). **Automatic multiple faces tracking and detection using improved edge detector algorithm.** International Conference on IT in Asia(CITA). 1-5.
- Park, C. W., Kwak, J. M., Park, H., and Moon, Y. S., (2007). **An Effective Method for Eye Detection Based on Texture Information.** International Conference on Convergence Information Technology. 586-589.
- Huan, N. V., N. Binh, T. H., and Kim, H., (2009). **Eye-verifier using ternary template for reliable eye detection in facial color images.** 2009 BTAS IEEE 3rd International Conference on. 1-6.
- Alioua, N., Amine, A., Rziza, M., and Aboutajdine, D., (2010). **Eye state analysis using iris detection based on Circular Hough Transform.** Multimedia Computing and Systems (ICMCS). 1-5.
- Huang, D. Y., Lin, T. W., Hu, W. C., and Chen, M. S., (2011). **Eye Detection Based on Skin Color Analysis with Different Poses under Varying Illumination Environment.** Fifth International Conference on Genetic and Evolutionary Computing. 252-255.
- Chen, S. and Liu, C., (2011). **Fast Eye Detection Using Different Color Spaces.** Systems Man and Cybernetics (SMC) IEEE International Conference on. 521-526.
- Kashani, M. A. A., Arani, M. M., and Fini, M. R. R., (2011). **Eye detection and tracking in images with using bag of pixels.** 2011 IEEE 3rd International Conference on. 64-68.
- Choopol Phromsuthirak and Sumet Umchid. **Development of a geometrical algorithm for eye detection in color images.** The 2012 Biomedical Engineering International Conference (BMEiCON-2012). 1-5.
- Goel, P., and Agarwal, S. (2012). **Hybrid Approach of Haar Cascade Classifiers and Geometrical Properties of Facial Features Applied to Illumination Invariant Gender Classification System.** International Conference on Computing Sciences. 132-136.

- Rahman, Md. H., Jhumur, F., Yusuf, Md. S. U., Das, T., and Ahmad, M. (2012). **An efficient face detection in color images using eye mouth triangular approach.** IEEE/OSA.IAPR International Conference on Informatics, Electronics & Vision. 530-535.
- Titive, F. H. C., and Bouzerdoum, A. (2005). **An eye feature detector based on convolutional neural network.** Proceedings of the Eighth International Symposium on. 90-93.
- Akashi, T., Wakasa, Y., Tanaka, K., Karungaru, S., and Fukumi, M. (2007). **Using Genetic Algorithm for Eye Detection and Tracking in Video Sequence.** Systemics, Cybernetics and Informatics. 5:72-78.
- Zhao, Y., Gao, Z., and Wu, W. (2010). **The detection algorithm of locomotive drivers' fatigue based on vision.** International Congress on Image and Signal Processing (CISP2010). 6:2686-2690.
- Qing, W., Xi, S. B., Bin, X., and Junjie, Z., (2010). **A PERCLOS-Based Driver Fatigue Recognition Application for Smart Vehicle Space.** Third International Symposium on Information Processing. 437-441.
- García, I., Bronte, S., Bergasa, L. M., Hernandez, N., Delgado, B., and Sevillano, M. (2010). **Vision-based drowsiness detector for a realistic driving simulator.** 2010 13th International IEEE Conference on. 887 – 894.
- Weber, M., Frontal Face Dataset, **California Institute of Technology**, available at <http://www.vision.caltech.edu/html-files/archive.html>
- R. Frischholz. **Baoface Database**, available at <http://www.Facedetection.com>
- Vision Group of Essex University Face Database**, available at <http://cswww.essex.ac.uk/mv/allfaces/index.html>
- Georgia Tech Face Database**, available at <ftp://ftp.ee.gatech.edu/pub/users/hayes/facedb/>



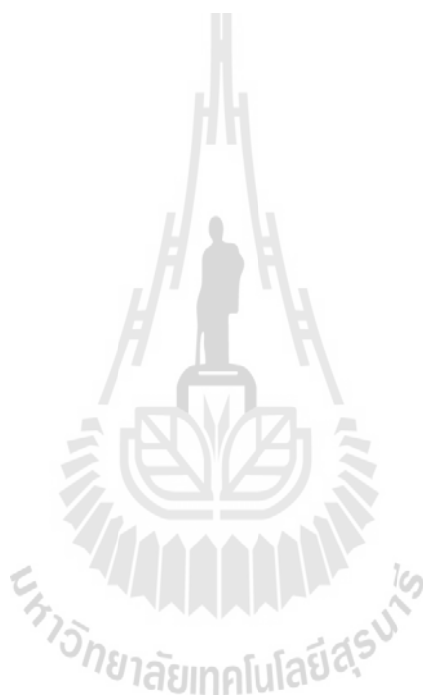
ภาคผนวก ก

บทความวิชาการที่ได้รับการตีพิมพ์เผยแพร่

มหาวิทยาลัยเทคโนโลยีสุรนารี

รายชื่อบทความวิชาการที่ได้รับการตีพิมพ์เผยแพร่ในระหว่างการศึกษา

เกริกฤทธิ์ ศรีเคน และ ประโยชน์ คำสวัสดิ์. (2556). การตรวจจับใบหน้าโดยใช้เทคนิคการคัดแยกสีผิว. การประชุมวิชาการงานวิจัยและพัฒนาเชิงประยุกต์ ครั้งที่ 5 (ECTI-CARD 2013), ณ โรงแรม สุรสัมมนาการ มหาวิทยาลัยเทคโนโลยีสุรนารี, วันที่ 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 พบว่าอัลกอริทึมที่ได้นำเสนอนี้มีการตรวจจับใบหน้าที่ถูกต้องถึง 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 ปัจจัยเหล่านี้ส่งผลให้ประสิทธิภาพในการขับรถเทียบเท่ากับกรณีแอลกอฮอล์ในเลือด 0.10 เปอร์เซ็นต์ ตามกฎหมายแล้วจะจำกัดปริมาณ

แอลกอฮอล์ในเลือดของผู้ขับขี่ไม่เกิน 0.05 เปอร์เซ็นต์ จะเห็นว่าปัญหาความเมื่อยล้าและอาการง่วงนอนส่งผลกระทบต่อผู้ขับขี่มากกว่าการมีปริมาณ แอลกอฮอล์ในเลือดตามที่กฎหมายกำหนดไว้ เพราะฉะนั้นจึงต้องออกแบบระบบที่เฝ้าระวังความเมื่อยล้าและอาการง่วงนอนของผู้ขับขี่เพื่อแจ้งเตือนให้ผู้ขับขี่ได้ทราบถึงระดับความเมื่อยล้าและอาการง่วงนอนของตนเอง

ระบบเฝ้าระวังความเมื่อยล้าและอาการง่วงนอนในอดีตมีหลายเทคนิค เช่น การวิเคราะห์คลื่นไฟฟ้าที่ได้จากสมอง (Electroencephalography: EEG) การสังเกตพฤติกรรมของยานพาหนะ การใช้กล้องวิดีโอที่ค้นตรวจการเคลื่อนไหวของดวงตา ซึ่งการใช้ EEG ในการตรวจจับความเมื่อยล้ามีข้อดีคือมีความแม่นยำสูงแต่จะมีข้อด้อยคือต้องใช้โพรบวัดติดกับศีรษะของผู้ขับขี่ทำให้เกิดการรบกวนผู้ขับขี่ในการขับยานพาหนะ การตรวจจับความเมื่อยล้าจากพฤติกรรมของยานพาหนะมีข้อดีคือไม่รบกวนผู้ขับขี่แต่จะมีข้อด้อยคือเมื่อรถเปลี่ยนช่องทางเดินรถ (Lane) และการชะลอความเร็วทำให้ระบบอาจสับสนและตัดสินใจผิดพลาดได้ว่าผู้ขับขี่กำลังมีความเมื่อยล้า และการใช้กล้องวิดีโอที่ค้นตรวจสอบการเคลื่อนไหวของดวงตาก็มีข้อดีคือไม่รบกวนผู้ขับขี่แต่จะมีข้อด้อยคือเมื่อผู้ขับขี่สวมแว่นตาและผลกระทบของแสงสว่าง ซึ่งปัญหานี้สามารถแก้ไขได้โดยการชดเชยแสงของกล้องวิดีโอ

งานวิจัยนี้ได้นำเสนอการตรวจจับใบหน้า ซึ่งเป็นขั้นตอนแรกของการตรวจจับความเมื่อยล้าและอาการง่วงนอนจากกล้องวิดีโอ เมื่อเปรียบเทียบกับข้อดีและข้อด้อยของระบบนี้พบว่าข้อดีของระบบนี้คือไม่มีการรบกวนผู้ขับขี่และไม่ก้าวก่ายในการขับรถของผู้ขับขี่ ส่วนข้อด้อยคือการนำกล้องวิดีโอมาใช้ในการตรวจจับใบหน้าจะมีผลกระทบจากแสงสว่าง แต่สามารถแก้ไขปัญหานี้ได้โดยใช้การชดเชยแสงสว่างจากกล้องวิดีโอ ซึ่งได้จากแหล่งกำเนิดแสง IR (Infrared) [2] ในส่วน

ของผู้ขับที่สวมแว่นตาที่ทำให้การตรวจจับผิดพลาด ซึ่งการตรวจจับผิดพลาดนี้เกิดขึ้นจากตัวผู้ขับเองไม่เกี่ยวกับระบบการตรวจจับความเมื่อยล้าและอาการง่วงนอนโดยการตรวจจับการเคลื่อนไหวของดวงตาให้ประสบผลสำเร็จนั้น สิ่งแรกที่ต้องทำคือการตรวจจับใบหน้าซึ่งในงานวิจัยนี้ผู้วิจัยได้นำเสนออัลกอริทึมที่ใช้ในการตรวจจับใบหน้าเพื่อนำไปสู่การตรวจจับการเคลื่อนไหวของดวงตาต่อไป

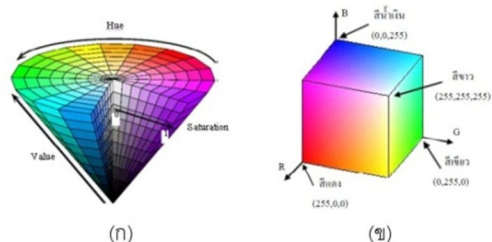
2. ความสำคัญที่มาของปัญหา

ระบบการตรวจจับความเมื่อยล้าและอาการง่วงนอนของผู้ขับจำเป็นมากในการลดอุบัติเหตุบนท้องถนนที่เกิดขึ้นจากความเมื่อยล้าและอาการง่วงนอนของผู้ขับ ระบบที่ดีจะต้องไม่รบกวนและก้าวก่ายผู้ขับ และต้องมีระบบเตือนที่สามารถเตือนให้ผู้ขับรู้ถึงระดับความเมื่อยล้าและอาการง่วงนอนที่เกิดขึ้นและยับยั้งการเกิดอุบัติเหตุที่อาจจะเกิดขึ้น โดยก่อนที่จะสามารถตรวจจับดวงตาได้จะต้องตรวจจับใบหน้าให้ได้เสียก่อน เนื่องจากความถูกต้องของการตรวจจับใบหน้าส่งผลต่อการตรวจจับดวงตา การตรวจจับใบหน้าที่มีในอดีตสามารถตรวจจับได้ถูกต้อง 93.7 เปอร์เซ็นต์ ผู้วิจัยจึงได้ออกแบบระบบที่สามารถตรวจจับใบหน้าที่มีความถูกต้องสูงมากกว่าเดิม

3. งานและทฤษฎีที่เกี่ยวข้อง

3.1 ปริภูมิสี HSV

ปริภูมิสี HSV เป็นปริภูมิสีที่ประกอบด้วย 3 องค์ประกอบได้แก่ค่า H (Hue) คือค่าโทนสีหรือค่าสีต้น ซึ่งมีค่าสีแตกต่างกันไปตามความถี่ของแสง 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)

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

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

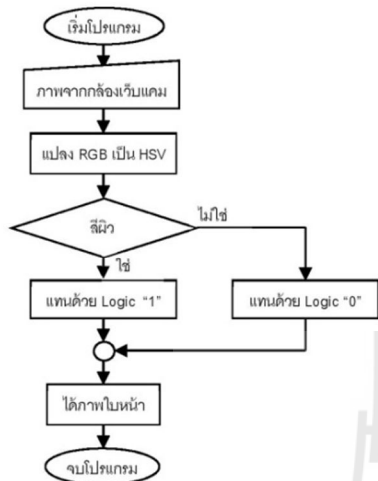
$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)} \quad (3)$$

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

4. รายละเอียดการพัฒนา

4.1 ภาพรวมของระบบ

ระบบตรวจจับใบหน้าดังแผนภาพในรูปที่ 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

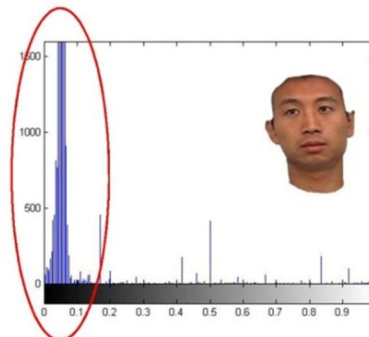
4.2.1 การเลือกใช้ปริภูมิสีที่บ่งบอกถึงใบหน้า

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) จากนั้นนำภาพใบหน้าหรือภาพจากกล้องวีดิทัศน์ที่ต้องการตรวจจับใบหน้ามาเข้าเงื่อนไข โดยจะแทนพิกเซลที่เป็นใบหน้าด้วยบิต '1' และพิกเซลที่ไม่ใช่ใบหน้าด้วยบิต '0'

ในงานวิจัยนี้ยังมีข้อจำกัดในเรื่องของแสงสว่างที่ทำให้การตรวจจับใบหน้าผิดพลาดและฐานข้อมูลที่นำมาทดสอบเป็นใบหน้าของคนจากหลายเชื้อชาติ 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	อัลกอริทึมการตรวจจับใบหน้า ที่นำเสนอ	อัลกอริทึมการตรวจจับ ใบหน้า [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 จะเห็นว่าการตรวจจับใบหน้าที่ได้นำเสนอมีความถูกต้องสูงสุดถึง 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)

หลังจากการทดสอบอัลกอริทึมการตรวจจับใบหน้าที่มีฐานข้อมูลใบหน้าพบว่าส่วนที่ผิดพลาดเกิดขึ้นจากหลายปัจจัย ปัจจัยแรกคือภาพมีแสงสว่างน้อย เพราะแสงสว่างเป็นปัจจัยที่ทำให้สีในภาพเกิดการผิดเพี้ยนไปซึ่งสามารถแก้ปัญหาได้โดยการใช้กล้องที่มีการชดเชยแสงสว่างหรือกล้องที่มีอินฟาเรด ซึ่งปัจจัยนี้ทำให้เกิดความผิดพลาด 3.43 เปอร์เซ็นต์ ปัจจัยที่สองเกิดจากสีของพื้นหลังที่มีสีคล้ายกับสีของใบหน้า ซึ่งเปอร์เซ็นต์การตรวจจับผิดพลาดในปัจจุบันนี้เท่ากับ 1.72 เปอร์เซ็นต์ และปัจจัยที่สามเกิดจากฐานข้อมูลใบหน้าที่เป็นภาพของใบหน้าจากหลายเชื้อชาติ ปัจจัยนี้ทำให้เกิดความผิดพลาด 0.77 เปอร์เซ็นต์

จะเห็นว่าการตรวจจับใบหน้าที่ผิดพลาดส่วนใหญ่เกิดจากภาพที่มีแสงสว่างน้อยซึ่งฐานข้อมูลที่ใช้ทดสอบไม่สามารถชดเชยแสงสว่างได้ แต่สำหรับการทดสอบกับกล้องเว็บแคมที่สามารถชดเชยแสงสว่างได้ จะเห็นว่าผลการทดลองกับกล้องเว็บแคมในตารางที่ 1 มีการตรวจจับใบหน้าถูกต้องถึง 99.30 เปอร์เซ็นต์ แสดงให้เห็นว่าเมื่อใช้กล้องที่มีการชดเชยแสงสว่าง จะทำให้การตรวจจับใบหน้ามีความถูกต้องมากขึ้น

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://cswww.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.

ประวัติผู้เขียน

นายเกริกฤทธิ์ ศรีเคน เกิดเมื่อวันที่ 2 ตุลาคม พ.ศ. 2532 เกิดที่อำเภอบัวใหญ่ จังหวัดนครราชสีมา สำเร็จการศึกษาระดับมัธยมศึกษาตอนปลายจากโรงเรียนบัวใหญ่ จังหวัดนครราชสีมา และสำเร็จการศึกษาระดับปริญญาตรี วิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมโทรคมนาคม จากมหาวิทยาลัยเทคโนโลยีสุรนารี จังหวัดนครราชสีมา เมื่อ พ.ศ. 2554 ขณะศึกษาระดับปริญญาโทเคยปฏิบัติงานเป็นผู้ช่วยสอนในรายวิชาของสาขาวิชาวิศวกรรมโทรคมนาคม และสาขาวิชาวิศวกรรมอิเล็กทรอนิกส์ สำนักวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี ได้แก่ ปฏิบัติการอิเล็กทรอนิกส์ขั้นสูง 1 ปฏิบัติการไมโครโปรเซสเซอร์ ปฏิบัติการอิเล็กทรอนิกส์ และปฏิบัติการวงจรอิเล็กทรอนิกส์ ในขณะที่ศึกษาอยู่ได้ทำงานเป็นผู้ช่วยวิจัยในโครงการวิจัย ดังนี้

1. โครงการวิจัยเรื่อง การออกแบบ และพัฒนาอัลกอริทึมการทำลายน้ำดิจิทัลสำหรับภาพสี
 2. โครงการวิจัยเรื่อง ระบบเฝ้าติดตามอาการง่วงนอน โดยใช้เทคนิคทางปัญญาประดิษฐ์
 3. โครงการวิจัยเรื่อง ระบบเครือข่ายเซ็นเซอร์ไร้สาย โดยใช้ระบบปฏิบัติการแอนดรอยด์
- ทางผู้วิจัยมีความสนใจทางด้าน การประมวลผลภาพดิจิทัล (Digital Image Processing) โครงการขยายประสาทเทียม การฝังลายน้ำดิจิทัล เครือข่ายเซ็นเซอร์ไร้สาย (Wireless Sensor Networks : WSN) การเขียนพัฒนาโปรแกรมระบบปฏิบัติการแอนดรอยด์ Internet of Things (IoT) และมีผลงานทางวิชาการที่ได้รับการตีพิมพ์เผยแพร่ในขณะที่ศึกษา 1 บทความ ดังรายชื่อที่ปรากฏในภาคผนวก ก.