

เทคนิคการตรวจจับหน้าคนด้วยโครงข่ายทฤษฎีเรโซแนนซ์แบบปรับตัว

นางสาวอภิรดี อัมพะลิริ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2551

**HUMAN FACE DETECTION TECHNIQUE BASED ON
ADAPTIVE RESONANCE THEORY NETWORK**

Apiradee Ampawasiri

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

Suranaree University of Technology

Academic Year 2008

เทคนิคการตรวจจับหน้าคนด้วยโครงข่ายทฤษฎีเรโซแนนซ์แบบปรับตัว

มหาวิทยาลัยเทคโนโลยีสุรนารี อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตาม
หลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

คณะกรรมการสอบวิทยานิพนธ์

(รศ. ดร.กิตติ อัครกิจมงคล)

ประธานกรรมการ

(ผศ. ดร.อาทิตย์ ศรีแก้ว)

กรรมการ (อาจารย์ที่ปรึกษาวิทยานิพนธ์)

(ผศ. ร.อ. ดร.ประโยชน์ คำสวัสดิ์)

กรรมการ

(ศ. ดร.ไพโรจน์ สัตยธรรม)

รองอธิการบดีฝ่ายวิชาการ

(รศ. น.อ. ดร.วรพจน์ ขำพิศ)

คณบดีสำนักวิชาวิศวกรรมศาสตร์

อภिरดี อัมพะละสิริ : เทคนิคการตรวจจับหน้าคนด้วยโครงข่ายทฤษฎีเรโซแนนซ์แบบปรับตัว
(HUMAN FACE DETECTION TECHNIQUE BASED ON ADAPTIVE RESONANCE
THEORY NETWORK) อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร.อาทิตย์ ศรีแก้ว, 92 หน้า

ระบบการตรวจจับหน้าคนเป็นส่วนหนึ่งของเทคโนโลยีทางคอมพิวเตอร์สำหรับการปฏิสัมพันธ์กันระหว่างคนและคอมพิวเตอร์ ซึ่งตลอดเวลาของการทำงาน คอมพิวเตอร์จำเป็นต้องทราบตำแหน่งของคนที่ต้องการปฏิสัมพันธ์ด้วย การศึกษาและพัฒนาระบบการตรวจจับและติดตามหน้านั้นจึงยังคงเป็นหัวข้อที่ได้รับความสนใจอยู่ เนื่องจากเป็นระบบที่มีความเปลี่ยนแปลงไม่แน่นอนของหน้าคนในหลากหลายปัจจัยอันประกอบด้วย โครงสร้างทางกายภาพ เชื้อชาติ ตำแหน่งที่ตั้ง การวางท่าและความซับซ้อนของสภาพแวดล้อม กระทั่งปัจจุบัน ได้มีการนำเสนอการค้นคว้าและวิจัยด้านการค้นหาหน้าคนแล้วเป็นอย่างมากมาย ที่ซึ่งหลายๆ งานนั้นสามารถนำไปใช้งานได้จริง งานวิจัยนี้ได้นำเสนอระบบการตรวจจับหน้าคนโดยใช้โครงข่ายประสาทเทียมแบบทฤษฎีเรโซแนนซ์แบบปรับตัว เป็นกลไกหลักสำหรับคัดแยกความเป็นหน้าคน โดยก่อนทำการคัดแยก ระบบจะทำการแปลงข้อมูลภาพด้วยการวิเคราะห์องค์ประกอบหลัก (principal component analysis : PCA) เพื่อเป็นการเตรียมข้อมูลที่เป็นตัวแทนของข้อมูลภาพนั้น ข้อมูลภาพที่ผ่านการแปลงนี้จะมีขนาดของข้อมูลที่เล็กลง และยังคงข้อมูลที่บ่งบอกความเป็นหน้าคน โดยที่ประสิทธิภาพในการคัดแยกหน้าคนของโครงข่ายสามารถยอมรับได้ นอกจากนั้นโครงข่ายยังมีลักษณะพิเศษที่สามารถเรียนรู้รูปแบบใหม่ได้ โดยยังคงจดจำรูปแบบก่อนหน้านั้น เพื่อให้ระบบการตรวจจับหน้าคนมีการเรียนรู้ที่ดีขึ้นเรื่อยๆ เหมาะสมกับการใช้งานการจดจำหน้าคนได้เป็นอย่างดี นอกจากนั้นแล้ว ในขั้นตอนการเรียนรู้ของโครงข่าย จินเนติกอัลกอริทึมได้ถูกนำมาช่วยในการหาค่าพารามิเตอร์ของโครงข่ายที่เหมาะสมที่สุด เพื่อที่จะมั่นใจว่าระบบมีประสิทธิภาพในการคัดแยกมากที่สุด ระบบที่นำเสนอนี้สามารถเพิ่มขีดความสามารถของกระบวนการเรียนรู้ ทำให้ได้มาซึ่งการจดจำหน้าคนที่ดีขึ้น

APIRADEE AMPAWASIRI : HUMAN FACE DETECTION TECHNIQUE

BASE-ON ADAPTIVE RESONANCE THEORY NETWORK. THESIS

ADVISOR : ASST. PROF. ARTHIT SRIKAEW, Ph.D., 92 PP.

FACE DETECTION/EIGENFACE/EIGEN VECTOR/ART

A human face detection system is one of computer technology for interaction between human and computers. It is necessary for computers to, most of the time, be capable of knowing human location. Human detection and tracking has always been an active field of study and development. Face detection is one of the complicated area of research due to wide variations of human faces including physical structure, human nationality, location, poses and typical complex surroundings. Nowadays, many face detection researches have been proposed in which many of them have already been available in real-world applications. This research presents a human face detection using adaptive resonance theory neural network for classification. The system transforms face images into a principal component analysis which provides an appropriate representation of the face image. This transformed image has smaller size and gives a compact form of human faces. This allows the network to efficiently classify face images. Moreover the network's unique ability of learning new pattern while still maintaining the previous ones makes the face detection system more preferable. In learning phase, the network parameters has been optimized using genetic algorithm in order to ensure maximum efficiency of classification. The proposed system shows desirable results which can be improved to achieve better performance through the process of learning

School of Electrical Engineering

Academic Year 2008

Student's Signature _____

Advisor's Signature _____

กิตติกรรมประกาศ

วิทยานิพนธ์นี้ดำเนินการสำเร็จลุล่วงด้วยดี เนื่องจากได้รับความช่วยเหลืออย่างดียิ่ง ทั้งด้านวิชาการและด้านดำเนินงานวิจัย จากบุคคล และกลุ่มบุคคลต่างๆ ผู้วิจัยขอขอบพระคุณบุคคลทั้งหลาย ซึ่งได้แก่

ผู้ช่วยศาสตราจารย์ ดร.อาทิตย์ ศรีแก้ว อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้คำปรึกษาแนะนำ และแนะแนวทางอันเป็นประโยชน์ยิ่งต่องานวิจัย รวมถึงได้ช่วยตรวจทาน และแก้ไขรายงานวิทยานิพนธ์เล่มนี้จนทำให้มีความสมบูรณ์ยิ่งขึ้น รวมทั้งเป็นกำลังใจให้กับผู้วิจัยเสมอมา

รองศาสตราจารย์ นาวาอากาศโท ดร.สรายุทธ สุจิตจร ผู้ช่วยศาสตราจารย์ ดร.กิตติ อรรถกิจมงคล และอาจารย์ประจำสาขาวิชาวิศวกรรมไฟฟ้า มหาวิทยาลัยเทคโนโลยีสุรนารีทุกท่าน ที่ได้กรุณาให้คำปรึกษา แนะนำ ความรู้ทางด้านวิชาการอย่างดียิ่งมาโดยตลอด

คุณอัญชลี รักด่านกลาง คุณภัทราวรรณ สิทธิวินกุล และบุคลากรศูนย์เครื่องมือวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีสุรนารีทุกท่าน ที่อำนวยความสะดวกในการทำงานของข้าพเจ้า

ตลอดจนพี่ ๆ เพื่อน ๆ น้อง ๆ บัณฑิตศึกษาทุกท่าน ที่คอยถามไถ่และให้กำลังใจในการทำวิจัยมาโดยตลอด

ท้ายนี้ผู้วิจัยขอกราบขอบพระคุณอาจารย์ผู้สอนทุกท่าน ที่ประสิทธิ์ประสาทความรู้ทางด้านต่าง ๆ ทั้งในอดีตและปัจจุบัน ขอกราบขอบพระคุณบิดา มารดา รวมถึงญาติพี่น้องของผู้วิจัยทุกท่านที่ได้ให้ความรัก ความอบอุ่น ความหวังใจ การอบรมเลี้ยงดู และให้การสนับสนุนทางด้านการศึกษาอย่างดียิ่งมาโดยตลอด

อภิรดี อัมพะะศิริ

สารบัญ

หน้า

บทคัดย่อ (ภาษาไทย).....	ก
บทคัดย่อ (ภาษาอังกฤษ).....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ช
สารบัญรูป.....	ซ
คำอธิบายสัญลักษณ์และคำย่อ.....	ฅ
บทที่	
1 บทนำ.....	1
1.1 ความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์การวิจัย.....	1
1.3 ขอบเขตของงานวิจัย.....	2
1.4 ขอบเขตของงานวิจัย.....	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	2
1.6 รูปแบบวิทยานิพนธ์.....	2
2 ปรัชญาวรรณกรรมและงานวิจัยที่เกี่ยวข้อง.....	4
2.1 กล่าวนำ.....	4
2.2 แบบจำลองภาพในปริภูมิสีต่าง ๆ.....	4
2.2.1 ภาพเชิงดิจิทัล.....	4
2.2.2 ภาพระดับเทา.....	5
2.2.3 ปริภูมิสี RGB.....	5
2.2.4 ปริภูมิ HSV.....	6
2.3 การคัดแยกสีผิวมนุษย์ออกจากพื้นหลัง.....	7
2.3.1 ค่าสีผิวมนุษย์.....	8
2.3.2 โครงข่ายประสาทเทียม.....	9
2.3.3 สถาปัตยกรรมของโครงข่ายประสาทเทียม.....	9
2.3.4 ฟังก์ชันถ่ายโอนของโครงข่ายประสาทเทียม.....	10

สารบัญ (ต่อ)

หน้า

2.3.5	โครงข่ายประสาทเทียมแบบแพร่กลับ.....	13
2.3.6	การคัดแยกสีผิวมนุษย์ด้วยโครงข่ายประสาทเทียมแบบแพร่กลับ.....	13
2.4	วิธีการวิเคราะห์ห้วงค้ำประกอบหลัก.....	16
2.4.1	การหาค่าไอเกนและไอเกนเวกเตอร์ ด้วยหลักการวิเคราะห์ห้วงค้ำประกอบ.....	18
2.5	จินเนติกอัลกอริทึม.....	21
2.6	งานวิจัยที่เกี่ยวข้อง.....	22
2.6.1	วิธีเชิงความรู้.....	22
2.6.2	วิธีเชิงลักษณะ.....	22
2.6.3	วิธีเทียบเคียงแผ่นแบบ.....	23
2.6.4	วิธีเชิงลักษณะปรากฏ.....	24
2.6.4.1	วิธีหน้าลักษณะเฉพาะ.....	24
2.6.4.2	วิธีเชิงการกระจาย.....	24
2.6.4.3	โครงข่ายประสาทเทียม.....	25
2.6.4.4	วิธีเวกเตอร์เกี่ยวพัน.....	25
2.7	สรุป.....	26
3	การตรวจจับหน้าคนด้วยโครงข่าย ART.....	27
3.1	กล่าวนำ.....	27
3.2	โครงข่าย ART.....	27
3.2.1	สถาปัตยกรรมของโครงข่าย ART.....	27
3.2.2	ประเภทของโครงข่าย ART.....	29
3.2.2.1	โครงข่าย ART1.....	29
3.2.2.2	โครงข่ายART2.....	30
3.2.2.3	โครงข่าย ARTMAP.....	30
3.2.2.4	โครงข่าย Fuzzy ART หรือ Fuzzy ARTMAP.....	30
3.3	ระบบตรวจจับหน้าคนจากภาพด้วยโครงข่าย ART.....	31
3.3.1	โครงสร้างของระบบ.....	31
3.3.2	หลักการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART.....	37
3.3.3	ตัวอย่างการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART.....	38
3.4	สรุป.....	39

สารบัญ (ต่อ)

หน้า

4 ผลการทดสอบระบบตรวจจับหน้าคนด้วยโครงข่าย ART.....	40
4.1 กล่าวนำ.....	40
4.2 ฐานข้อมูลภาพหน้ามนุษย์และภาพที่ใช้ในการทดสอบระบบ.....	40
4.2.1 ฐานข้อมูลภาพ CVSR.....	40
4.2.2 ฐานข้อมูลภาพ CMU VASC.....	41
4.2.3 ฐานข้อมูลภาพ AVTG.....	42
4.2.4 ฐานข้อมูลภาพ VALID และฐานข้อมูลภาพ CBCL.....	43
4.2.5 ภาพจากกล้องดิจิตอลฟูจิ F 601 และกล้องวิดีโอ.....	44
4.3 ทดสอบระบบตรวจจับหน้าคนด้วยโครงข่าย ART.....	46
4.3.1 ทดสอบระบบเมื่อมีการฝึกสอนโครงข่ายในลักษณะต่าง ๆ.....	46
4.3.2 ทดสอบระบบที่พารามิเตอร์ของโครงข่ายต่าง ๆ.....	59
4.3.3 ทดสอบระบบที่ขนาดของใบหน้าในภาพขนาดต่าง ๆ.....	60
4.3.4 ทดสอบระบบที่หน้าในลักษณะต่าง ๆ.....	64
4.3.5 ทดสอบระบบกับภาพที่มีหน้าคนในภาพมากกว่า 1 คน.....	68
4.4 สรุป.....	71
5 สรุปงานวิจัยและข้อเสนอแนะ.....	72
5.1 สรุปงานวิจัย.....	72
5.2 ข้อเสนอแนะ.....	72
รายการอ้างอิง.....	74
ภาคผนวก	
ภาคผนวก ก. โครงข่ายประสาทเทียมแบบแพร่กลับ.....	77
ภาคผนวก ข. ภาพผลการทดสอบระบบ.....	80
ภาคผนวก ค. บทความวิชาการที่ได้รับการตีพิมพ์เผยแพร่ในระหว่างศึกษา.....	90
ประวัติผู้เขียน.....	92

สารบัญตาราง

ตารางที่	หน้า
2.1 ผลการฝึกสอนโครงข่ายประสาทเทียมเพื่อเลือกพารามิเตอร์ของโครงข่าย.....	14
4.1 ทดสอบระบบที่จำนวนภาพในการฝึกสอนโครงข่ายต่าง ๆ.....	48
4.2 ทดสอบระบบด้วยภาพที่มีคุณลักษณะต่าง ๆ.....	55
4.3 ความถูกต้องในการคัดแยกหน้าคนโดยค่าสอดคล้องค่าต่าง ๆ.....	59
4.4 ความถูกต้องในการคัดแยกหน้าคนเมื่อขนาดของใบหน้าในภาพเปลี่ยนแปลง.....	61
4.5 ความถูกต้องในการคัดแยกหน้าคนในลักษณะต่าง ๆ.....	65

สารบัญรูป

รูปที่	หน้า
2.1 ช่วงจุดภาพระดับเทา.....	5
2.2 ลูกบาศก์จำลองของปริภูมิ RGB.....	6
2.3 แบบจำลองกรวยภาพสีของปริภูมิ HSV.....	7
2.4 ภาพในปริภูมิสี RGB ภาพระดับเทา และปริภูมิสี HSV.....	7
(ก) ภาพในปริภูมิสี RGB.....	7
(ข) ภาพระดับสีเทา.....	7
(ค) ภาพในปริภูมิสี HSV.....	7
2.5 ตัวอย่างอินพุตของโครงข่ายประสาทเทียม.....	8
2.6 ค่าสีผิวมนุษย์	
(ก) ปริภูมิ RGB.....	8
(ข) ปริภูมิ HSV.....	8
2.7 แบบจำลองเซลล์ประสาทเทียมหลายอินพุต.....	10
2.8 รูปแบบสัญลักษณ์เซลล์ประสาทเทียม.....	10
2.9 โครงข่ายประสาทเทียมแบบหลายชั้น.....	12
2.10 การกำหนดพารามิเตอร์ของโครงข่ายประสาทเทียมแบบหลายชั้น.....	12
2.11 ผลการคัดแยกสีผิวมนุษย์ออกจากสิ่งแวดล้อมด้วยโครงข่ายประสาทเทียมแบบแพร่กลับ	
(ก) ภาพต้นแบบ.....	15
(ข) ผลการคัดแยกสีผิวมนุษย์.....	15
(ค) ขอบเขตของสีผิวมนุษย์.....	15
(ง) ผลการฝึกสอนโครงข่าย.....	15
(จ) ขอบเขตการค้นหาน้ำคน.....	15
2.12 โครงข่ายประสาทเทียมแบบแพร่กลับในการคัดแยกสีผิวมนุษย์.....	15
2.13 ภาพตัวอย่างในการหาน้ำไอเคน.....	18
(ก) ตัวอย่างภาพในปริภูมิสี RGB.....	18
(ข) ตัวอย่างภาพในปริภูมิสี HSV.....	18
2.14 ภาพระดับเทาที่ทำการลดผลของความสว่าง.....	18
(ก) สีผิวกลาง.....	18

สารบัญรูป (ต่อ)

รูปที่	หน้า
(ข) สีส้มเข้ม.....	18
(ค) สีส้มอ่อน.....	18
2.15 ตัวอย่างการแปลงเมทริกซ์ภาพเป็นเวกเตอร์ภาพ.....	19
2.16 ภาพหน้าไอเคน แถวที่ 1 ภาพต้นแบบ แถวที่ 2 ภาพหน้าไอเคนจากไอเคนเวกเตอร์ค่า ที่ 1 แถวที่ 3 ภาพหน้าไอเคนจากไอเคนเวกเตอร์ค่าที่ 10.....	20
2.17 ขอบเขตการพิจารณาของแต่ละลำดับชั้น.....	22
2.18 แบบจำลองใบหน้ามนุษย์จากฐานข้อมูล MPEG-4.....	24
2.19 การกระจายตัวของข้อมูลตัวอย่าง.....	25
3.1 แผนผังการทำงานของโครงข่าย ART.....	28
3.2 โครงข่ายประสาทเทียม ART.....	29
3.3 แผนผังการทำงานของโครงข่าย ART ในการหาค่าสอดคล้องที่เหมาะสมด้วย GA.....	33
3.4 แผนผังการค่าสอดคล้องที่เหมาะสมด้วย GA.....	34
3.5 โครงสร้างระบบการตรวจจับหน้าคนด้วยโครงข่าย ART.....	35
3.6 แผนผังการโครงข่าย ART ในการคัดแยกหน้าคน.....	36
3.7 ก่อสร้างข้อมูลอินพุตของโครงข่าย ART	
(ก) เส้นกำหนดขนาดกล่องข้อมูล.....	37
(ข) กล่องข้อมูลอินพุตที่ป้อนเข้าโครงข่าย ART.....	37
(ค) กล่องข้อมูลที่ดีที่สุดที่แสดงตำแหน่งของหน้าคนในภาพ.....	37
3.8 ภาพตัวอย่างการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART.....	38
(ก) ตัวอย่างภาพต้นแบบ.....	38
(ข) ตัวอย่างภาพผลการทำงานของโครงข่าย ART.....	38
(ค) ตัวอย่างภาพผลการตรวจจับหน้าคนด้วยโครงข่าย ART.....	38
4.1 ภาพตัวอย่างจากฐานข้อมูล CVSR.....	41
4.2 ภาพตัวอย่างจากฐานข้อมูล CMU VASC.....	42
4.3 ภาพตัวอย่างจากฐานข้อมูล AVTG.....	43
4.4 ภาพตัวอย่างจากฐานข้อมูล VALID.....	44
4.5 ภาพตัวอย่างจากฐานข้อมูล CBCL.....	44
4.6 ภาพตัวอย่างจากกล้องฟูจิ F601.....	45

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.7 ภาพตัวอย่างจากกล้องวิดีโอ.....	46
4.8 ภาพตัวอย่างในการฝึกสอนโครงข่าย.....	46
4.9 ตัวอย่างภาพหน้าไอเคน	
(ก) ภาพหน้าไอเคนที่จำนวนภาพในการหาหน้าไอเคน 3 ภาพ.....	47
(ข) ภาพหน้าไอเคนที่จำนวนภาพในการหาหน้าไอเคน 5 ภาพ.....	47
(ค) ภาพหน้าไอเคนที่จำนวนภาพในการหาหน้าไอเคน 10 ภาพ.....	47
4.10 ตัวอย่างภาพผลการทดสอบระบบ ที่จำนวนตัวอย่างในการฝึกสอนต่าง ๆ	
(ก) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวน ตัวอย่างในการฝึกสอน 3 ภาพ.....	49
(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวน ตัวอย่างในการฝึกสอน 3 ภาพ.....	49
(ค) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวน ตัวอย่างในการฝึกสอน 5 ภาพ.....	50
(ง) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวน ตัวอย่างในการฝึกสอน 5 ภาพ.....	50
(จ) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวน ตัวอย่างในการฝึกสอน 10 ภาพ.....	51
(ฉ) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวน ตัวอย่างในการฝึกสอน 10 ภาพ.....	51
(ช) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวน ตัวอย่างในการฝึกสอน 20 ภาพ.....	52
(ซ) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวน ตัวอย่างในการฝึกสอน 20 ภาพ.....	52
(ณ) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวน ตัวอย่างในการฝึกสอน 30 ภาพ.....	53
(ญ) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวน ตัวอย่างในการฝึกสอน 30 ภาพ.....	53

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.11 ตัวอย่างภาพหน้าไอเคน	
(ก) ภาพหน้าไอเคนของบุคคลเดียวกัน ที่ลักษณะใบหน้าต่างกัน.....	54
(ข) ภาพหน้าไอเคนของบุคคลต่างกัน ที่ลักษณะใบหน้าตรง.....	54
(ค) ภาพหน้าไอเคนของบุคคลต่างกัน ที่ลักษณะใบหน้าต่างกัน.....	55
4.12 ตัวอย่างภาพผลการทดสอบระบบ ของภาพตัวอย่างในการฝึกสอนที่คุณลักษณะต่าง ๆ	
(ก) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างในการฝึกสอน โครงข่ายของบุคคลเดียวกันที่ลักษณะใบหน้าต่าง ๆ.....	56
(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างในการฝึกสอน โครงข่ายของบุคคลเดียวกันที่ลักษณะใบหน้าต่าง ๆ.....	56
(ค) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างในการฝึกสอน โครงข่ายของบุคคลต่างกัน ที่ลักษณะใบหน้าตรง.....	57
(ง) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างในการฝึกสอน โครงข่ายของบุคคลต่างกัน ที่ลักษณะใบหน้าตรง.....	57
(จ) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างในการฝึกสอน โครงข่ายของบุคคลต่างกันที่ลักษณะใบหน้าต่าง ๆ.....	58
(ฉ) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างในการฝึกสอน โครงข่ายของบุคคลต่างกันที่ลักษณะใบหน้าต่าง ๆ.....	58
4.13 ตัวอย่างภาพการทดสอบระบบของระบบที่กำหนดค่าสอดคล้อง ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802	
(ก) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของระบบที่กำหนดค่าสอดคล้อง ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802.....	60
(ข) ตัวอย่างภาพผลการทดสอบระบบไม่ถูกต้อง ของระบบที่กำหนดค่าสอดคล้อง ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802.....	60
4.14 ตัวอย่างภาพของหน้าคนในภาพขนาดต่าง ๆ	
(ก) หน้าคนขนาด 70%.....	61
(ข) หน้าคนขนาด 55%.....	61
(ข) หน้าคนขนาดน้อยกว่า 50%.....	61

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.17 ตัวอย่างภาพผลการทดสอบระบบ ของภาพที่มีจำนวนหน้าคนในภาพมากกว่า 1 หน้า	
(ก) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของภาพที่มี จำนวนหน้าคนมากกว่า 1 หน้า ในภาพ.....	69
(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของภาพที่มี จำนวนหน้าคนมากกว่า 1 หน้า ในภาพ.....	70
ข.1 ภาพผลการทดสอบระบบ.....	81

บทที่ 1

บทนำ

1.1 ความสำคัญของปัญหา

ในปัจจุบันเทคโนโลยีด้านหุ่นยนต์ได้รับความสนใจและพัฒนาขึ้นอย่างรวดเร็ว มีการแข่งขันและส่งเสริมให้มีการคิดค้นเทคนิคใหม่ ๆ เพื่อให้หุ่นยนต์สามารถทำงานได้ใกล้เคียงกับมนุษย์มากที่สุด หนึ่งในการพัฒนาเพื่อให้หุ่นยนต์มีความใกล้เคียงกับมนุษย์คือ เพิ่มการมองเห็นให้หุ่นยนต์ เพื่อการเชื่อมต่อระหว่างคนและคอมพิวเตอร์ ซึ่งเป็นส่วนหนึ่งของการพัฒนาเทคโนโลยีการมองเห็นของคอมพิวเตอร์ (computer vision) การที่คอมพิวเตอร์จะสามารถสื่อสารกับมนุษย์ให้มีประสิทธิภาพดี คอมพิวเตอร์จำเป็นจะต้องทราบตำแหน่งของมนุษย์ที่ต้องการจะสื่อสารด้วย จึงมีการเพิ่มตาให้คอมพิวเตอร์ ซึ่งคือกล้องรับภาพ ดังนั้นเทคนิคการตรวจจับใบหน้าบุคคลจากภาพสี จึงเป็นเทคโนโลยีที่สำคัญในการพัฒนาเทคโนโลยีการมองเห็นของคอมพิวเตอร์

ระบบการตรวจจับหน้าคนจากภาพ ได้มีการค้นคว้าพัฒนาเทคนิคและวิธีการต่าง ๆ อยู่เสมอ โดยงานวิจัยการค้นหาน้ำคนจากภาพนั้น มีผู้ศึกษาและค้นคว้าวิจัยเรื่องนี้มากกว่า 20 ปีแล้ว และได้รับความสนใจมากในช่วง 5 ปีหลังจากที่ผ่านมานี้ เหตุผลที่ทำให้งานนี้ได้รับความสนใจเนื่องมาจากความท้าทายในการค้นหาน้ำคน ที่มีความหลากหลายของใบหน้าที่ไม่ซ้ำกันในแต่ละบุคคล รวมถึงความเปลี่ยนแปลงไม่แน่นอนในการปรากฏของใบหน้า เช่น โครงสร้างทางกายภาพของใบหน้าที่แตกต่างกันไปตามเชื้อชาติ ตำแหน่งที่ตั้ง การวางท่า หรือการแสดงออกของสีหน้าเป็นต้น และปัจจัยที่เป็นสิ่งแวดล้อมของภาพที่มีใบหน้าปรากฏอยู่ก็มีผลต่อความถูกต้องในการตรวจจับใบหน้าด้วยเช่นกัน ซึ่งได้แก่ สีพื้นหลังใบหน้า ความสว่าง แสงและเงา เป็นต้น โดยในการวิจัยช่วงแรกจะพิจารณาเป็นภาพเดี่ยว และพัฒนาให้สามารถตรวจจับหน้าคนจากภาพได้มากกว่าหนึ่งคนขึ้นไป

1.2 วัตถุประสงค์การวิจัย

1.2.1 เพื่อศึกษาและพัฒนาวิธีการตรวจจับใบหน้าบุคคลจากภาพสี

1.2.2 เพื่อศึกษาและพัฒนาวิธีการจดจำใบหน้าบุคคลจากภาพสี

1.3 ข้อตกลงเบื้องต้น

ลักษณะการวางท่าของใบหน้าต้องอยู่ในตำแหน่งที่มีลักษณะเด่นในพิจารณาชัดเจน คือต้องเป็นภาพหน้าตรง โดยมีการก้มหน้าหรือเงยหน้าในระดับที่กำหนดไว้ล่วงหน้า ไม่ถูกปิดบังหรือซ่อนทับจากวัตถุหรือร่างกายของบุคคลอื่น

1.4 ขอบเขตของงานวิจัย

1.4.1 พัฒนาระบบการตรวจจับหน้าคนจากภาพสี โดยใช้วิธีการวิเคราะห์องค์ประกอบหลัก (principal component analysis : PCA) ในการสร้างเวกเตอร์ลักษณะเฉพาะร่วมกับโครงข่ายประสาทเทียม แบบ Adaptive Resonance Theory (ART) ในการคัดแยกความเป็นหน้าคนออกจากสิ่งแวดล้อม

1.4.2 พัฒนาระบบการตรวจจับใบหน้าบุคคลจากภาพสีให้สามารถตรวจจับภาพที่มีหน้าคนมากกว่า 1 หน้าได้

1.4.3 พัฒนาระบบการจดจำใบหน้าบุคคลด้วยหน้าลักษณะเฉพาะ (eigenface)

1.4.4 ระบบที่ได้มีอัตราการตรวจจับใบหน้าบุคคลด้วยความถูกต้องมากกว่า 80%

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 เป็นแนวทางในการพัฒนาระบบการมองเห็นของคอมพิวเตอร์

1.5.2 พัฒนาการวิจัยเกี่ยวกับการค้นหาใบหน้าบุคคล

1.5.3 ประยุกต์ใช้งานในด้านการพัฒนาหุ่นยนต์ ระบบอุตสาหกรรมอัตโนมัติ ด้านความปลอดภัย เป็นต้น

1.6 รูปเล่มวิทยานิพนธ์

วิทยานิพนธ์ฉบับนี้ประกอบด้วยเนื้อหาทั้งหมด 5 บท และภาคผนวกโดยมีรายละเอียดของเนื้อหาในแต่ละบทดังนี้

บทที่ 1 กล่าวถึงความสำคัญของปัญหา วัตถุประสงค์การวิจัย ข้อตกลงเบื้องต้น ขอบเขตของการวิจัย และประโยชน์ที่คาดว่าจะได้รับจากงานวิจัยนี้

บทที่ 2 กล่าวถึง การประมวลผลภาพ การกำหนดขอบเขตในการตรวจจับหน้าคนเบื้องต้นด้วยการคัดแยกสีผิวมนุษย์ออกจากสิ่งแวดล้อม โครงข่ายประสาทเทียม และงานวิจัยที่เกี่ยวข้อง

บทที่ 3 อธิบายทฤษฎีที่เกี่ยวข้อง ได้แก่ หลักการวิธีการวิเคราะห์องค์ประกอบหลัก การหาค่าไอเกน ไอเกนเวกเตอร์ โครงข่าย ART โครงสร้างของระบบตรวจจับหน้าคน และการจดจำหน้าคนจากค่าไอเกน

บทที่ 4 นำเสนอผลการทดสอบระบบตรวจจับหน้าคนด้วยโครงข่าย ART โดยทำการทดสอบระบบกับฐานข้อมูลภาพต่าง ๆ เช่น ฐานข้อมูลภาพ CVSR ฐานข้อมูล CMU VASC เป็นต้นในสภาวะสิ่งแวดล้อมในภาพต่าง ๆ กัน ภาพพื้นหลังที่ซับซ้อน และภาพพื้นหลังที่ไม่ซับซ้อนทดสอบระบบด้วยภาพหน้าคนที่มีการวางท่าต่าง ๆ กัน เช่น หน้าตรง หน้าเอียง หน้าที่มีการปิดบังบางส่วนของใบหน้า ทดสอบระบบด้วยภาพหน้าคนที่แสดงอารมณ์ต่าง ๆ กัน เช่น หน้านิ่ง หน้ายิ้ม หน้าหัวเราะ หน้าโกรธหรือเคียด ทดสอบระบบด้วยภาพหน้าคนที่มีหน้าคนปรากฏอยู่ในภาพ ต่าง ๆ กัน เช่น ภาพเดี่ยว ภาพที่มีหน้าคนปรากฏในภาพ 2 คน ภาพที่มีหน้าคนปรากฏในภาพ 3 คน และภาพที่มีหน้าคนปรากฏในภาพ 4 คนขึ้นไป เป็นต้น

บทที่ 5 สรุปผลการทดลองของงานวิจัยนี้ และข้อเสนอแนะสำหรับงานวิจัยและพัฒนาในอนาคต

ภาคผนวก ประกอบไปด้วยเนื้อหา 3 ส่วน ได้แก่ ภาคผนวก ก. แสดงรายละเอียดของโครงข่ายประสาทเทียมแบบแพร่กลับ ภาคผนวก ข. แสดงภาพที่ใช้ในการทดสอบระบบและภาพผลการทดสอบระบบที่เสนอในงานวิจัยนี้ ภาคผนวก ค. บทความทางวิชาการที่ได้รับการตีพิมพ์เผยแพร่ในระหว่างศึกษา

บทที่ 2

ปรีทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

2.1 กล่าวนำ

การค้นหาน้ำคนจากภาพสีนั้น จำเป็นที่จะต้องรู้จักปริภูมิสีต่าง ๆ ซึ่งมีการพิจารณาค่าสีที่แตกต่างกันออกไป โดยภาพสีจากปริภูมิหนึ่งสามารถแปลงค่าไปสู่อีกปริภูมิหนึ่งได้ และมักใช้ประโยชน์จากความแตกต่างของค่าสีในการคัดแยกสีผิวมนุษย์ออกจากสีอื่น ๆ ซึ่งการคัดแยกสีผิวมนุษย์ออกจากสิ่งแวดล้อมนี้ จะช่วยให้สามารถลดขอบเขตในการค้นหาน้ำคนให้แคบลงได้ กล่าวคือ บริเวณที่มีสีผิวมนุษย์ปรากฏอยู่จะมีความเป็นไปได้สูงที่จะมีน้ำคนปรากฏอยู่ ดังนั้นถ้าเราเลือกพิจารณาเฉพาะบริเวณที่มีสีผิวมนุษย์ปรากฏอยู่ จะช่วยให้ระบบทำงานได้เร็วขึ้นและลดการผิดพลาดในการค้นหาให้น้อยลง

บทที่ 2 นี้นำเสนอ แบบจำลองภาพในปริภูมิสีต่าง ๆ รวมไปถึงการคัดแยกสีผิวมนุษย์ออกจากสิ่งแวดล้อม วิธีการวิเคราะห์องค์ประกอบหลัก และทฤษฎีที่เกี่ยวข้องในการค้นหาน้ำคนจากภาพสี ตลอดจนนำเสนอผลงานวิจัยที่เกี่ยวข้องกับการค้นหาน้ำคนจากภาพสีที่น่าสนใจ ดังรายละเอียดต่อไปนี้

2.2 แบบจำลองภาพในปริภูมิสีต่าง ๆ

ปริภูมิสีของภาพดิจิทัลนั้นมีอยู่ด้วยกันหลากหลายปริภูมิสี ได้แก่ RGB, HSV (HSI), YCrCb, YUV, YIQ, YES, CIE XYZ และ CIE LUV ซึ่งภาพสีต่าง ๆ ที่เราเห็นนั้น โดยทั่วไปแล้วจะเป็นภาพสีในปริภูมิสี RGB และภาพสีในปริภูมิ RGB นั้นยังสามารถแปลงไปเป็นภาพระดับเทา ซึ่งเป็นภาพระนาบเดียวที่ง่ายต่อการนำไปวิเคราะห์และเป็นอินพุตของระบบ

2.2.1 ภาพเชิงดิจิทัล (Digital Image) คือฟังก์ชัน 2 มิติ $f(x,y)$ ของค่าความเข้มแสง (Intensity) โดยที่ x และ y เป็นค่าแสดงตำแหน่งในระบบพิกัดฉาก และขนาดของ f ที่ตำแหน่ง (x,y) คือค่าความเข้มแสงของภาพที่จุดนั้น ๆ โดยภาพสีเกิดจากค่าความเข้มแสงที่เป็นปริมาณเวกเตอร์ประกอบด้วย 3 องค์ประกอบ คือ เวกเตอร์สีแดง เวกเตอร์สีเขียว และเวกเตอร์สีน้ำเงิน ซึ่งประกอบกันด้วยอัตราส่วนต่าง ๆ กันเป็นภาพสี 1 ภาพ (ภาพปริภูมิ RGB)

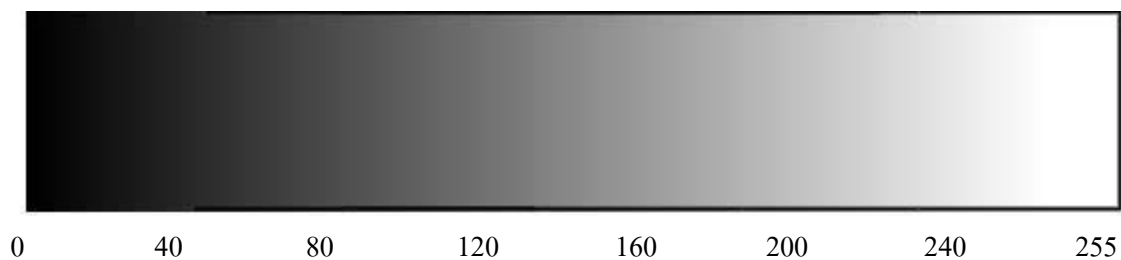
2.2.2 ภาพระดับเทา (Gray Scale Image) มีค่าความเข้มแสงเป็นปริมาณสเกลาร์หรือเรียกว่า ภาพสีเดียว (Monochrome Image) โดยค่า $f(x,y)$ ต้องไม่เป็นศูนย์และมีค่าจำกัด สามารถแยกองค์ประกอบเป็น 2 ส่วน คือ องค์ประกอบความสว่าง (illumination component : $i(x,y)$) และ องค์ประกอบสะท้อนของแสง (reflectance component : $r(x,y)$) โดยที่ $f(x,y)$ เกิดจากการคูณกันขององค์ประกอบทั้งสอง แสดงในสมการที่ 2.1

$$f(x,y) = i(x,y) \cdot r(x,y) \quad (2.1)$$

โดยที่ $i(x,y)$ ต้องมีค่ามากกว่าศูนย์และมีค่าจำกัด ส่วน $r(x,y)$ ต้องมีค่ามากกว่าศูนย์และน้อยกว่าหนึ่ง ดังนั้นค่าจุดภาพระดับเทาจะมีค่าเป็นช่วง แสดงในสมการที่ 2.2

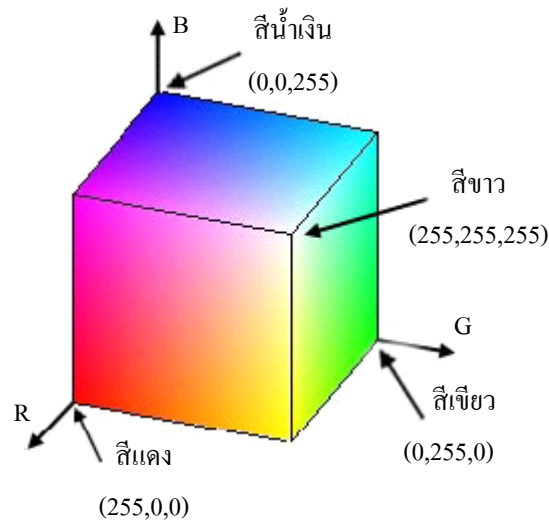
$$L_{min} \leq l \leq L_{max} \quad (2.2)$$

เรียกช่วง $[L_{min}, L_{max}]$ นี้ว่าระดับเทา (gray scale) ซึ่งโดยปกติแล้วนิยามเลื่อนช่วงนี้ให้มีค่าเป็น $[0, 255]$ โดยที่ L_{min} เท่ากับศูนย์ คือจุดภาพสีดำ และ L_{max} เท่ากับ 255 คือจุดภาพสีขาว ส่วนจุดภาพที่มีค่าในช่วงระหว่างนี้จะแสดงระดับของสีเทา ดังรูปที่ 2.1



รูปที่ 2.1 ช่วงจุดภาพระดับเทา

2.2.3 ปริภูมิสี RGB เป็นปริภูมิที่ประกอบด้วยองค์ประกอบของสเปกตรัม 3 แม่สีปฐมภูมิ (primary spectral components) ได้แก่ สีแดง (red: R) สีเขียว (green: G) และสีน้ำเงิน (blue: B) ซึ่งแต่ละสีนั้นอยู่บนระนาบภาพ (image planes) ที่เป็นอิสระต่อกัน บนพื้นฐานระบบพิกัดคาร์ทีเซียนดังแสดงในรูปที่ 2.2 และในกล้องสิดิจิตอลส่วนมากนิยมใช้รูปแบบการรับภาพแบบภาพสี RGB ดังนั้น ปริภูมิสี RGB จึงเป็นปริภูมิที่สำคัญ



รูปที่ 2.2 ลูกบาศก์จำลองของปริภูมิ RGB

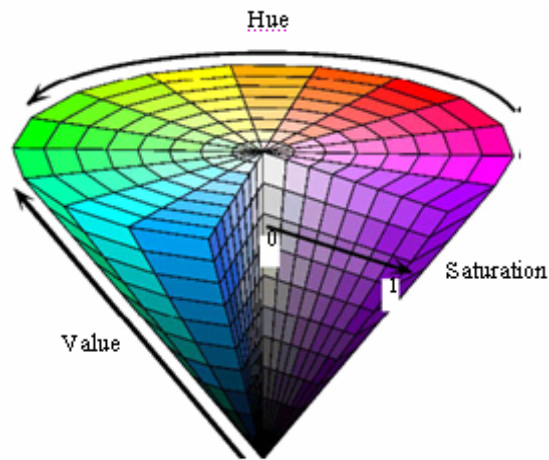
2.2.4 ปริภูมิสี HSV เป็นปริภูมิที่ประกอบด้วย 3 องค์ประกอบได้แก่ ค่า H (Hue) คือค่าโทนสี หรือค่าสีต้น ซึ่งมีค่าสีแตกต่างกันไปตามความถี่ของแสง ค่า S (Saturation) คือค่าความอิ่มตัวสี ซึ่งเป็นค่าที่แสดงระดับสีเมื่อเทียบกับค่าโทนสี และค่า V (Value) คือค่าบอกระดับความสว่างของภาพ ซึ่งที่ระดับความสว่างต่ำสุดหมายถึงสีดำ ไม่ว่าค่าโทนสี หรือค่าความอิ่มตัวสีเท่าใด และระดับความสว่างสูงสุดหมายถึงสีขาว ซึ่งเป็นสีที่สว่างที่สุดของค่าโทนสี และค่าความอิ่มตัวสี โดยภาพปริภูมิสี HSV สามารถหาได้จากการแปลงภาพปริภูมิสี RGB ด้วยสมการที่ 2.3 ถึงสมการที่ 2.6 และแสดงแบบจำลองภาพสี HSV ในรูปที่ 2.3

$$H = \begin{cases} H_1 & \text{if } B \leq G \\ 360^\circ - H_1 & \text{if } B > G \end{cases} \quad (2.3)$$

$$H_1 = \cos^{-1} \left\{ \frac{0.5[(R-G) + (R-B)]}{\sqrt{(R-G)(R-G) + (R-B)(R-B)}} \right\} \quad (2.4)$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (2.5)$$

$$V = \frac{\max(R, G, B)}{255} \quad (2.6)$$



รูปที่ 2.3 แบบจำลองกรวยภาพสีของปริภูมิ HSV

และจากสมการที่ได้กล่าวไปแล้วนั้น จะสามารถแปลงภาพในปริภูมิ RGB เป็นภาพระดับเทา และภาพในปริภูมิ HSV ได้ดังรูปที่ 2.4



(ก)



(ข)



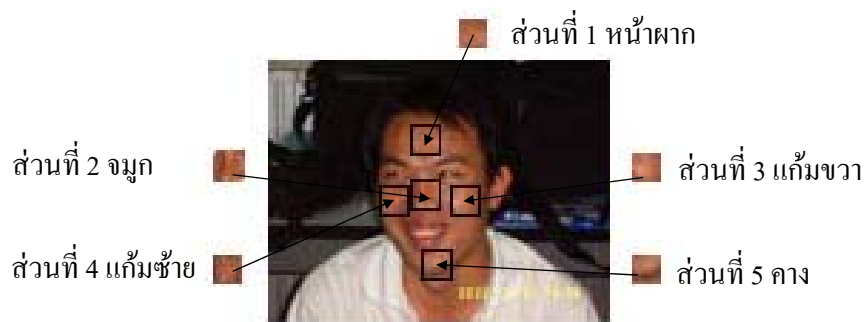
(ค)

รูปที่ 2.4 ภาพในปริภูมิสี RGB ภาพระดับเทา และปริภูมิสี HSV (ก) ภาพในปริภูมิสี RGB
(ข) ภาพระดับเทา (ค) ภาพในปริภูมิสี HSV

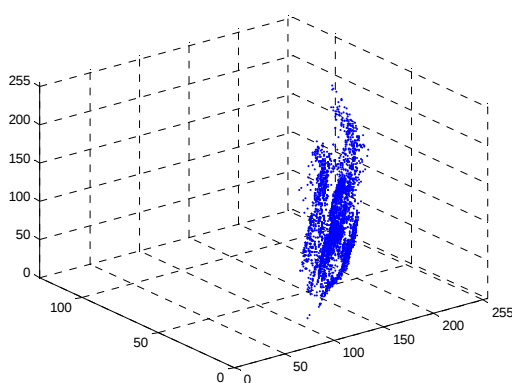
2.3 การคัดแยกสีผิวมนุษย์ออกจากพื้นหลัง

สีผิวมนุษย์นั้น มีความหลากหลายแตกต่างกันไปขึ้นอยู่กับลักษณะพันธุกรรมและเชื้อชาติ แต่จะพบว่าความแตกต่างของสีผิวมนุษย์ที่เกิดขึ้นนั้น มาจากความแตกต่างทางด้านความเข้มแสง (intensity) มากกว่าทางด้านความเป็นสี (chrominance) ดังนั้นการแยกสีผิวมนุษย์ทำได้โดยการตัดส่วนที่เป็นความเข้มแสงออก พิจารณาเฉพาะส่วนของความเป็นสี ทำให้ได้ค่าสีผิวมนุษย์ที่แตกต่างจากค่าสีอื่น ๆ จากนั้นใช้ช่วงของค่าสีผิวมนุษย์นี้เป็นอินพุตให้โครงข่ายประสาทเทียมแบบแพร่กลับทำการคัดแยกสีผิวมนุษย์ออกจากสีอื่น ๆ เพื่อกำหนดขอบเขตในการค้นหาหน้ามนุษย์ให้แคบลง (ชมพู่ ทรัพย์ปทุมสิน, 2548)

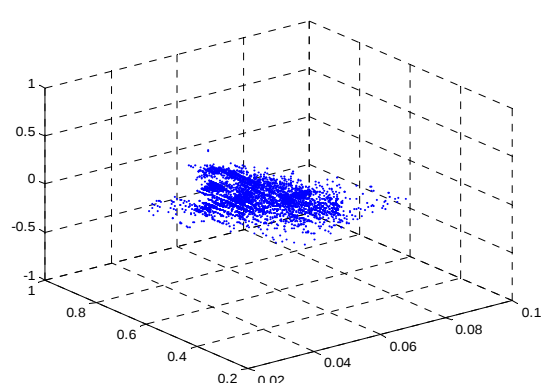
2.3.1 ค่าสีผิวมนุษย์ (Human Color) เป็นค่าที่สามารถคัดแยกออกจากสีของสิ่งแวดล้อมได้ โดยแปลงจากปริภูมิสี RGB เป็นปริภูมิสี HSV เพื่อแยกส่วนความเข้มแสงออกจากส่วนความเป็นสี เนื่องจากปริภูมิสี HSV สามารถกำจัดผลของความแตกต่างของค่าความเข้มแสงของสีผิวมนุษย์ และลดข้อจำกัดทางด้านแสงสว่างได้จากการตัดส่วนของค่า V ซึ่งเป็นค่าที่ใช้บอกระดับความสว่างของภาพออก จากนั้นตัดส่วนของภาพที่เป็นสีผิวมนุษย์มาเป็นอินพุตให้โครงข่าย โดยการเก็บตัวอย่างอินพุตจากภาพนั้น จะทำการเก็บตัวอย่างบริเวณใบหน้า และภาพใบหน้า 1 ภาพนั้นจะแบ่งบริเวณการเก็บตัวอย่างสีผิวมนุษย์ ออกเป็น 5 ส่วนด้วยกัน เนื่องจากทั้ง 5 ส่วนนี้เป็นส่วนของสีผิวมนุษย์ที่ไม่มีสีอื่นที่ไม่ใช่สีผิวมนุษย์ผสมอยู่ด้วย และเป็นจุดหลักบนใบหน้าทำให้ง่ายในการเก็บข้อมูล โดยข้อมูลที่เก็บได้นั้นมีขนาดใหญ่เหมาะสม และในการเก็บข้อมูลถึง 5 จุดนั้น เพื่อให้ได้ตัวอย่างที่ครอบคลุมส่วนต่าง ๆ ของใบหน้ามากที่สุด ดังแสดงในรูปที่ 2.5 และค่าสีผิวมนุษย์ในปริภูมิสี RGB และ HSV แสดงในรูปที่ 2.6



รูปที่ 2.5 ตัวอย่างอินพุตของโครงข่ายประสาทเทียม



(ก)



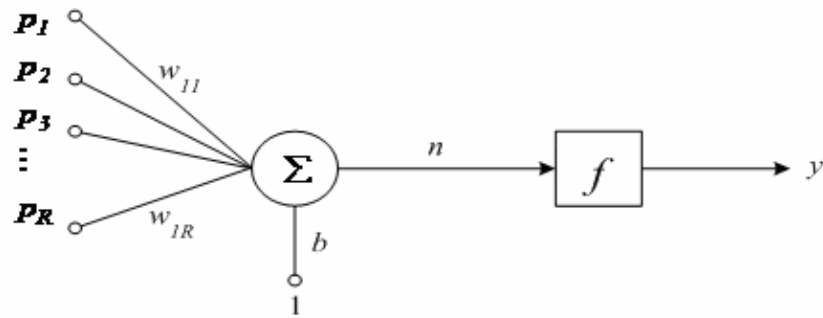
(ข)

รูปที่ 2.6 ค่าสีผิวมนุษย์ (ก) ปริภูมิ RGB (ข) ปริภูมิ HSV

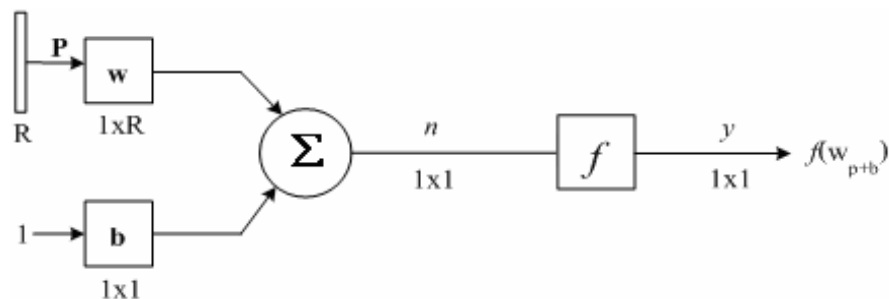
2.3.2 โครงข่ายประสาทเทียม (อาทิตย์ ศรีแก้ว, 2552) เป็นโครงข่ายที่สามารถปรับตัวเองต่อการตอบสนองของอินพุต ตามกฎการเรียนรู้ (learning rule) และเมื่อโครงข่ายได้เรียนรู้สิ่งที่ต้องการแล้ว โครงข่ายจะสามารถทำงานที่กำหนดไว้ได้ โดยโครงข่ายประสาทเทียมถูกพัฒนาคิดค้นมาจากหลักการการทำงานของสมองมนุษย์ ซึ่งประกอบไปด้วยหน่วยประมวลผลเรียกว่าเซลล์ประสาทเทียม (neuron) โดยจำนวนเซลล์ประสาทในสมองมนุษย์นั้นมีอยู่ประมาณ 10^{11} เซลล์ และมีการเชื่อมต่อกันอย่างมากมาย สมองมนุษย์จึงเสมือนเป็นคอมพิวเตอร์ที่สามารถปรับตัวเองได้ ไม่เป็นเชิงเส้น และทำงานแบบขนานในการดูแลจัดการการทำงานร่วมกันของเซลล์ประสาทในสมอง ดังนั้นโครงข่ายประสาทเทียมที่เลียนแบบมาจากการทำงานของสมองมนุษย์นี้ จึงมีความสามารถในการเรียนรู้จากตัวอย่าง และการทำให้เป็นกรณีทั่วไป (generalize) ซึ่งถือว่าเป็นคุณลักษณะสำคัญของโครงข่ายประสาทเทียม โดยโครงข่ายจะถูกฝึกสอนด้วยรูปแบบต่าง ๆ ที่ต้องการให้โครงข่ายเรียนรู้ผ่านกฎการเรียนรู้ ขบวนการเรียนรู้ได้ของโครงข่ายนั่นเองที่ทำให้โครงข่ายมีความแตกต่างจากการทำงานของโปรแกรมคอมพิวเตอร์อื่น ๆ และการที่โครงข่ายถูกทำให้เป็นกรณีทั่วไปทำให้โครงข่ายสามารถจำแนกแยกแยะรูปแบบของอินพุตแบบใหม่ ๆ ที่โครงข่ายไม่รู้จักมาก่อนได้ โครงข่ายจะทำการเก็บข้อมูลความรู้ในระหว่างขั้นตอนของการเรียนรู้ โดยเก็บไว้ที่น้ำหนักประสาท (synaptic weights) โครงสร้างของตัวเซลล์ประสาทเทียมภายในโครงข่ายมีอยู่หลายชนิด ซึ่งเป็นองค์ประกอบสำคัญที่ทำให้คุณลักษณะต่าง ๆ ของโครงข่ายแตกต่างกันออกไป ไม่ว่าจะเป็นการจัดวางเรียงตัวของเซลล์ประสาทเทียม กฎการเรียนรู้ที่ทำให้เกิดการปรับเปลี่ยนค่าของจุดประสาท และเงื่อนไขในการฝึกฝนของ นอกจากนี้โครงข่ายประสาทเทียมยังมีข้อดีอีกมากมาย เช่น มีความทนทานต่อความบกพร่อง เพราะข้อมูลภายในโครงข่ายได้ถูกกระจายไปทั้งโครงข่ายตามเซลล์ประสาทเทียมต่าง ๆ การจะทำให้ทั้งระบบไม่สามารถทำงานได้นั้น จะต้องทำให้เกิดความเสียหายอย่างหนักเท่านั้น และโครงข่ายประสาทเทียมมีความสามารถในการได้ตอบ และตอบสนองต่อสภาวะแวดล้อมได้ โดยเมื่อสภาวะแวดล้อมเปลี่ยนไป ตัวโครงข่ายจะสามารถตอบสนองกับการเปลี่ยนแปลงนั้น ๆ แล้วทำการฝึกฝนให้เข้ากับสภาวะแวดล้อมใหม่ได้

2.3.3 สถาปัตยกรรมของโครงข่ายประสาทเทียม (Neural network architecture) โครงสร้างของโครงข่ายประสาทเทียมประกอบด้วย 2 ส่วนหลัก ๆ คือ แบบจำลองเซลล์ประสาทเทียม และสถาปัตยกรรมการเชื่อมต่อกันเป็นโครงข่ายของเซลล์ประสาทเทียม โดยแบบจำลองของเซลล์ประสาทเทียมแสดงในรูปที่ 2.7 และรูปแบบสัญลักษณ์เซลล์ประสาทเทียมแสดงในรูปที่ 2.8

เซลล์ประสาทเทียมมี R อินพุต แต่ละอินพุตย่อย $p_1, p_2, p_3, \dots, p_R$ ถูกคูณด้วยค่าน้ำหนักประสาท (weight) ของแต่ละตัว $w_{11}, w_{12}, w_{13}, \dots, w_{1R}$ และ ไบอัส b (bias) เป็นอีกหนึ่งอินพุตที่มีค่าน้ำหนักประสาทที่เท่ากับ 1 โดยอินพุตทั้งสองถูกรวม (sum) ได้เอาต์พุตเป็น n เรียกว่าเน็ตอินพุต (net input) ซึ่งเป็นอินพุตให้กับฟังก์ชันถ่ายโอน f (transfer function) และได้เอาต์พุตของเซลล์ประสาทเทียมคือ y โดยเอาต์พุตของเซลล์ประสาทเทียมสามารถคำนวณได้จากสมการที่ 2.7



รูปที่ 2.7 แบบจำลองเซลล์ประสาทเทียมหลายอินพุต



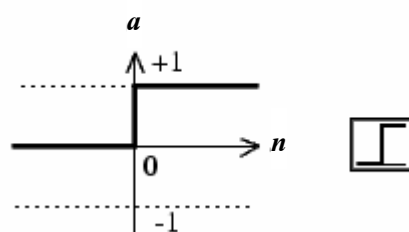
รูปที่ 2.8 รูปแบบสัญลักษณ์เซลล์ประสาทเทียม

เห็นได้ว่าเอาต์พุตของเซลล์ประสาทเทียมขึ้นอยู่กับน้ำหนักประสาท $\mathbf{W}$ และไบอัส $\mathbf{b}$ ซึ่งถูกปรับค่าจากกฎการเรียนรู้

$$y = f(\mathbf{Wp} + \mathbf{b}) \quad (2.7)$$

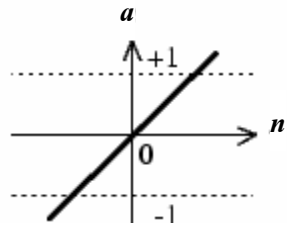
2.3.4 ฟังก์ชันถ่ายโอนของโครงข่ายประสาทเทียม (Transfer function) เป็นฟังก์ชันแบบเชิงเส้น หรือไม่เป็นเชิงเส้นก็ได้ ซึ่งมีอยู่ด้วยกันหลายรูปแบบ แต่ที่นิยมใช้กันแพร่หลายได้แก่

2.3.4.1 ฟังก์ชันถ่ายแบบฮาร์ดลิมิต (Hard Limit Transfer Function) ฟังก์ชันถ่ายโอนชนิดนี้จะให้ค่าเอาต์พุต 2 ค่า ตามค่าอินพุต กล่าวคือถ้าอินพุตมีค่าต่ำกว่า 0 จะให้ค่าเอาต์พุตเป็น 0 และถ้าอินพุตมีค่ามากกว่าหรือเท่ากับ 0 จะให้ค่าเอาต์พุตเป็น 1



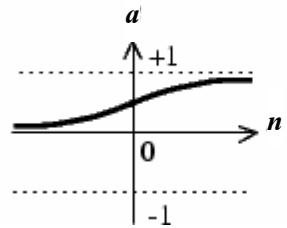
$$a = \text{hard lim}(n) \quad , \quad a = \begin{cases} 0 & n < 0 \\ 1 & n \geq 0 \end{cases}$$

2.3.4.2 ฟังก์ชันถ่ายแบบเส้นตรง (Linear Transfer Function) ฟังก์ชันถ่ายโอนชนิดนี้จะให้ค่าเอาต์พุตเท่ากับค่าอินพุตที่ป้อนเข้ามา



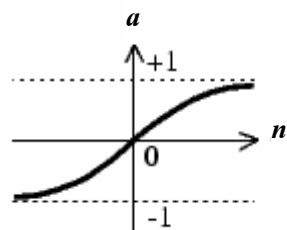
$$a = \text{purelin}(n) = n$$

2.3.4.3 ฟังก์ชันถ่ายแบบลอจ ซิกมอยด์ (Log-sigmoid Transfer Function) ฟังก์ชันถ่ายโอนชนิดนี้จะให้ค่าเอาต์พุตอยู่ระหว่างช่วง 0 ถึง 1



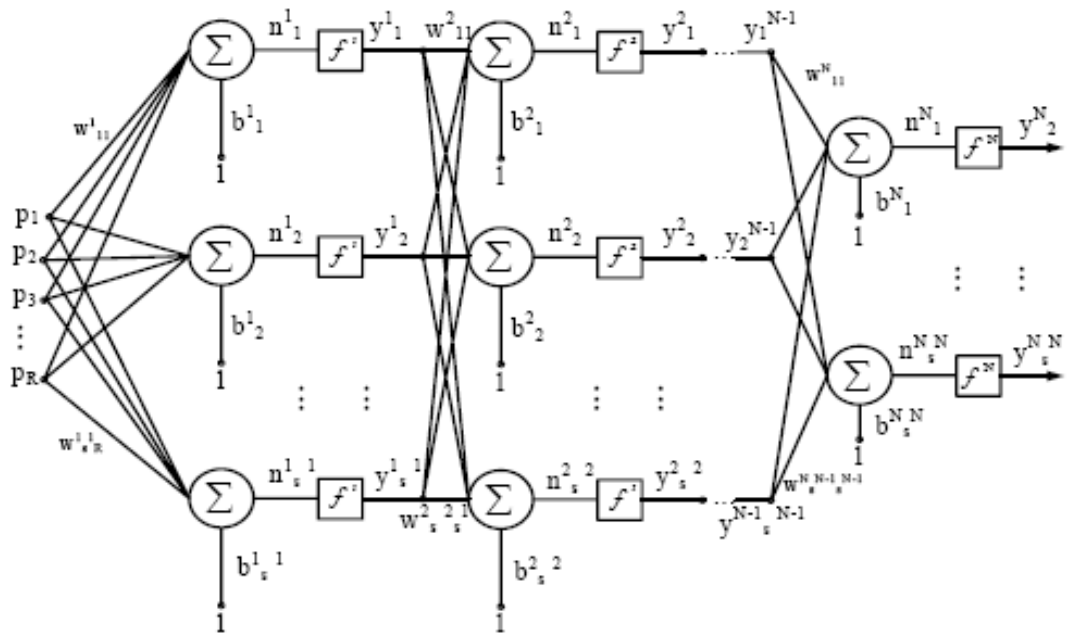
$$a = \text{log sig}(n) = \frac{1}{1 + e^{-n}}$$

2.3.4.4 ฟังก์ชันถ่ายโอนแบบไฮเพอร์โบลิกแทนเจนต์ ซิกมอยด์ (Hyperbolic Tangent Sigmoid) ฟังก์ชันถ่ายโอนชนิดนี้จะให้ค่าเอาต์พุตอยู่ระหว่างช่วง 1 ถึง -1



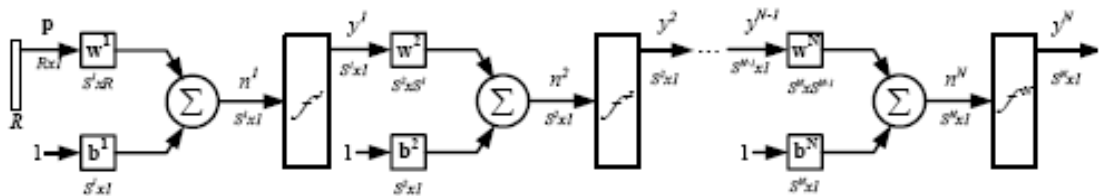
$$a = \text{tan sig}(n) = \frac{2}{(1 + e^{-2n}) - 1}$$

โดยมากในทางปฏิบัติ โครงข่ายประสาทเทียมมีโครงสร้างหลายชั้น (multiple layer) แต่ละชั้นมีเมทริกซ์น้ำหนักประสาท $\mathbf{W}$ เวกเตอร์ $\mathbf{b}$ เนตเอาต์พุต n และเอาต์พุต y ของชั้นนั้น แสดงในรูปที่ 2.9



รูปที่ 2.9 โครงข่ายประสาทเทียมแบบหลายชั้น

จากโครงสร้างโครงข่ายประสาทเทียมแบบหลายชั้น จำนวนเซลล์ประสาทเทียมของแต่ละชั้นคือ $R-S^1-S^2-\dots-S^N$ ซึ่งโดยปกติแล้วชั้นแรกจะเป็นชั้นอินพุต (input layer) ทำหน้าที่รับอินพุตจากภายนอกโครงข่าย ในชั้นสุดท้ายจะเป็นชั้นเอาต์พุตสำหรับส่งค่าเอาต์พุตออกจากโครงข่าย ส่วนชั้นระหว่างอินพุตและเอาต์พุตเรียกว่าชั้นซ่อนเร้น (hidden layer) และแต่ละชั้นสามารถมีจำนวนเซลล์ประสาทเทียมแตกต่างกันได้ โดยที่เอาต์พุตของแต่ละชั้นจะเป็นอินพุตให้กับชั้นถัดไป แสดงในรูปที่ 2.10



รูปที่ 2.10 การกำหนดพารามิเตอร์ของโครงข่ายประสาทเทียมแบบหลายชั้น

และความสัมพันธ์ของค่าเอาต์พุตแต่ละชั้นแสดงในสมการที่ 2.8

$$y^{l+1} = f^{l+1}(W^{l+1}y^l + b^{l+1}) \quad (2.8)$$

โดยที่ $l=0, 1, 2, \dots, N-1$ ชั้นแรกเป็นชั้นอินพุตที่รับอินพุตโดยตรงจากภายนอกโครงข่าย คือ $y^0 = p$ และชั้นสุดท้ายเป็นชั้นเอาต์พุต คือ y^N ซึ่งในโครงข่ายหลายชั้นจะมีพารามิเตอร์ค่อนข้างมาก ดังนั้นการจะนำโครงข่ายประสาทเทียมไปประยุกต์ใช้งานจึงต้องคำนึงถึงการออกแบบพารามิเตอร์ต่าง ๆ เช่น จำนวนชั้น จำนวนเซลล์ประสาทเทียมในแต่ละชั้น จำนวนอินพุต จำนวนเอาต์พุต ชนิดของฟังก์ชันถ่ายโอน ฯลฯ ให้เหมาะสมกับการใช้งาน

2.3.5 โครงข่ายประสาทเทียมแบบแพร่กลับ (Backpropagation Neuron Network) โครงข่ายประสาทเทียมแบบแพร่กลับ เป็นโครงข่ายประสาทเทียมแบบหลายชั้น มีการเรียนรู้แบบมีผู้ฝึกสอน (supervised learning) คือ เรียนรู้ที่จะสร้างผลลัพธ์ที่ต้องการให้ได้ตามตัวอย่างที่ป้อนให้ และลดค่าความผิดพลาดของเอาต์พุตให้น้อยที่สุดโดยเทียบกับค่าน้ำหนักประสาท ซึ่งเป็นการวิเคราะห์หาค่าความผิดพลาดแบบกำลังสองเฉลี่ย (mean square error) ในอัลกอริทึมแบบแพร่กลับ มีการนำเสนอคู่อินพุตและเป้าหมายให้โครงข่ายเรียนรู้แสดงในสมการที่ 2.9

$$\{p_1, t_1\}, \{p_2, t_2\}, \dots, \{p_Q, t_Q\} \quad (2.9)$$

2.3.6 การคัดแยกสีผิวมนุษย์ด้วยโครงข่ายประสาทเทียมแบบแพร่กลับ จำเป็นจะต้องกำหนดคู่อินพุตและเป้าหมายที่ชัดเจน เพื่อให้โครงข่ายสามารถทำงานได้อย่างถูกต้องที่สุด โดยจะทำการเก็บตัวอย่างสีผิวมนุษย์จากที่ได้กล่าวไว้ในหัวข้อที่ 2.3.1 มาเป็นคู่อินพุตและเป้าหมายของโครงข่าย ซึ่งใช้ค่า H และค่า S ของภาพในช่วงที่เป็นสีผิวมนุษย์ระบุเป้าหมายว่าเป็นสีผิวมนุษย์ และค่า H และค่า S ของภาพในช่วงที่ไม่เป็นสีผิวมนุษย์ระบุเป้าหมายว่าไม่ใช่สีผิวมนุษย์มาฝึกสอนโครงข่าย ดังความสัมพันธ์ในสมการที่ 2.10

$$\begin{cases} P_{skin} = \begin{bmatrix} H_{skin} \\ S_{skin} \end{bmatrix}, t_{skin} = 1 \\ P_{nonskin} = \begin{bmatrix} H_{nonskin} \\ S_{nonskin} \end{bmatrix}, t_{nonskin} = 0 \end{cases} \quad (2.10)$$

โดยกำหนดตัวอย่างสีผิวมนุษย์ที่ใช้ฝึกสอนโครงข่ายจำนวน 100 ตัวอย่าง ขนาด 50x50 จุดภาพ จากจำนวนคน 20 คนซึ่งมีสีผิวแตกต่างกันออกไป และตัวอย่างภาพที่ไม่ใช่สีผิวมนุษย์จำนวน 100 ตัวอย่าง ขนาด 50x50 จุดภาพ ต่อมาทำการกำหนดตัวชี้วัดประสิทธิภาพ (performance index) เมื่อระบบมีการนำเสนอคู่อินพุตและเป้าหมายให้โครงข่ายเรียนรู้ ทำการป้อนแต่ละอินพุตให้กับโครงข่าย เอาต์พุตที่ได้จะถูกนำไปเปรียบเทียบกับเป้าหมาย เพื่อให้ค่าความผิดพลาดของเอาต์พุตและเป้าหมายมีค่าน้อยที่สุด ซึ่งค่า

ความผิดพลาดระหว่างเอาต์พุตและเป้าหมายนี้เองเป็นตัวชี้วัดประสิทธิภาพ โดยตัวชี้วัดประสิทธิภาพดังกล่าวนี้ คือค่าความผิดพลาดแบบกำลังสองเฉลี่ย (mean-square error) ของเอาต์พุตและเป้าหมาย

จากนั้นเพื่อให้ระบบมีประสิทธิภาพมากที่สุด จำเป็นจะต้องออกแบบพารามิเตอร์ของโครงข่ายประสาทเทียมแบบแพร่กลับซึ่งได้แก่ จำนวนชั้น จำนวนเซลล์ประสาทเทียม และถ่ายโอน ให้เหมาะสมกับงานที่ทำ ซึ่งในงานวิจัยเรื่องวิธีการใหม่แบบพันทางในการแยกส่วนมือจากภาพสี (ชมพู่ ทรัพย์ปทุมสิน, 2548) ได้ทำการทดลองหาพารามิเตอร์ที่เหมาะสมของโครงข่ายประสาทเทียมแบบแพร่กลับในการคัดแยกสีผิวมนุษย์ออกจากพื้นหลัง โดยใช้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยเป็นตัวชี้วัดประสิทธิภาพ ได้ผลการทดลองดังตารางที่ 3.1

ตารางที่ 2.1 ผลการฝึกสอนโครงข่ายประสาทเทียมเพื่อเลือกพารามิเตอร์ของโครงข่าย

โครงข่าย	ฟังก์ชันถ่ายโอน	ค่าความผิดพลาดแบบกำลังสองเฉลี่ย
2-2-1	Tansig — Tansig	0.0823830
2-2-1	Tansig — Purelin	0.0781679
2-2-1	Purelin — Purelin	0.1501080
2-2-1	Tansig — Logsig	0.0498176
2-5-1	Tansig — Purelin	0.0746313
2-5-1	Tansig — Logsig	0.0430907
2-10-1	Tansig — Purelin	0.0689867
2-10-1	Tansig — Logsig	0.0408147
2-5-3-1	Tansig — Logsig — Tansig	0.5086721
2-5-3-1	Tansig — Logsig — Logsig	0.0682925
2-5-3-1	Tansig — Logsig — Purelin	0.0678671
2-5-3-2-1	Tansig — Logsig — Logsig — Purelin	0.2496315

จากผลการทดลอง โครงข่ายที่ให้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยน้อยที่สุดคือ โครงข่าย 2-10-1 (ชั้นอินพุตมีเซลล์ประสาทเทียม 2 เซลล์ มีชั้นซ่อนเร้น 1 ชั้นมีเซลล์ประสาทเทียม 10 เซลล์ และชั้นเอาต์พุตมีเซลล์ประสาทเทียม 1 เซลล์) ใช้ฟังก์ชันถ่ายโอนแบบไฮเพอร์โบลิคแทนเจนต์ ซิกมอยด์ (Hyperbolic Tangent Sigmoid) และแบบลอจ ซิกมอยด์ (Log-Sigmoid) ได้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยเท่ากับ 0.0408147 และโครงข่ายที่ให้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยน้อยรองลงมา คือ โครงข่าย 2-5-1 ใช้ฟังก์ชันถ่ายโอนแบบไฮเพอร์โบลิคแทนเจนต์ ซิกมอยด์ และแบบลอจ ซิกมอยด์ ได้ค่าความผิดพลาดแบบกำลังสองเฉลี่ย 0.0430907 ดังนั้นจากผลการทดลอง จึงเลือกใช้โครงข่าย 2-5-1 ใช้ฟังก์ชันถ่ายโอนแบบไฮเพอร์โบลิคแทนเจนต์ ซิกมอยด์ เนื่องจากโครงข่าย 2-5-1 นั้นให้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยน้อยกว่าไม่มากนักเมื่อเทียบกับ โครงข่าย 2-10-1 แต่ใช้จำนวนเซลล์ประสาทเทียมน้อย

กว่าโครงข่าย 2-10-1 ทำให้โครงข่ายมีขนาดเล็กกว่าโดยที่ประสิทธิภาพของโครงข่ายใกล้เคียงกัน โดยผลการตัดแยกสีผิวแสดงในรูปที่ 2.11 และสถาปัตยกรรมโครงข่ายประสาทเทียมแบบแพร่กลับในการแยกกลุ่มสีผิวมนุษย์แสดงในรูปที่ 2.12



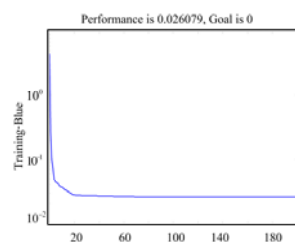
(ก)



(ข)



(ค)



(ง)

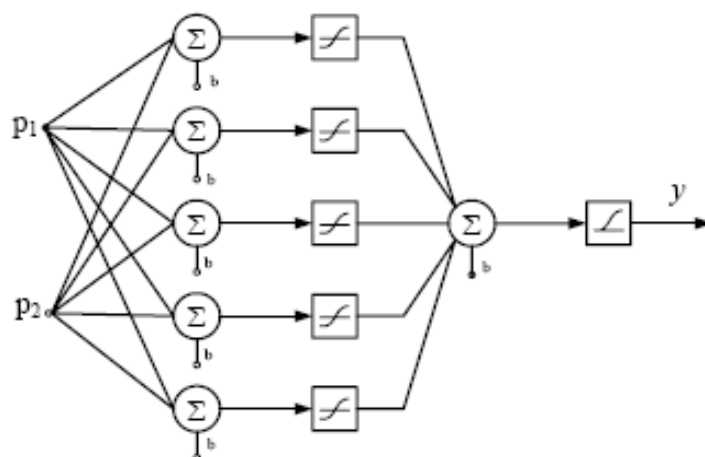


(จ)

รูปที่ 2.11 ผลการตัดแยกสีผิวมนุษย์ออกจากสิ่งแวดล้อมด้วยโครงข่ายประสาทเทียมแบบแพร่กลับ

(ก) ภาพต้นแบบ (ข) ผลการตัดแยกสีผิวมนุษย์ (ค) ขอบเขตของสีผิวมนุษย์

(ง) ผลการฝึกสอนโครงข่าย (จ) ขอบเขตการค้นหาหน้าคน



รูปที่ 2.12 โครงข่ายประสาทเทียมแบบแพร่กลับในการแยกกลุ่มสีผิวมนุษย์

2.4 วิธีการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA)

วิธีการวิเคราะห์องค์ประกอบหลัก เป็นการแทนชุดข้อมูลใหม่ที่มีขนาดเล็กกว่าลงในชุดข้อมูลเดิม ซึ่งชุดข้อมูลใหม่นั้นจะมีความสัมพันธ์ระหว่างข้อมูลภายในชุดน้อยมาก โดยชุดข้อมูลเดิมที่มีความสัมพันธ์ภายในข้อมูลขนาด D มิติ สามารถใช้การรวมเชิงเส้นในการลดข้อมูลเป็น d มิติ โดยที่ $d < D$ และเรียกชุดข้อมูลใหม่นี้ว่า องค์ประกอบหลัก (Principal Components)

วิธีการวิเคราะห์องค์ประกอบหลักสามารถทำได้ 2 วิธี คือ ทฤษฎีผลรวมของระยะทางที่น้อยที่สุด (Least Squared Distance) นำเสนอโดย Duda (2001) และ วิธีการเปลี่ยนตัวแปร นำเสนอโดย Hotelling (2001)

1. ทฤษฎีผลรวมของระยะทางที่น้อยที่สุด (Least Squared Distance) เป็นวิธีที่มีแนวคิดมาจากการหาจุดและเส้นที่มีความเหมาะสมที่สุดของชุดข้อมูลตัวอย่างขนาด M ค่า โดยแต่ละค่ามีขนาด D มิติ ทำให้ได้ข้อมูลตัวอย่างทั้งหมด คือ $t_1, t_2, \dots, t_M$ แล้วแทนข้อมูลตัวอย่างนั้นด้วยข้อมูลใหม่ ขนาด D มิติ และ 1 มิติ โดยที่ t_0 คือข้อมูลที่ถูแทนเป็น 0 มิติ (0-dimensional representation) เพื่อหาความสัมพันธ์ของข้อมูลตัวอย่างกับข้อมูลใหม่ โดยใช้ t_0 มาคำนวณหาผลรวมความผิดพลาดน้อยที่สุด แสดงในสมการที่ 2.11

$$J_0(t_0) = \sum_{k=1}^M \|t_0 - t_k\|^2 \quad (2.11)$$

โดยค่า t_0 ที่ทำให้ J_0 มีค่าน้อยที่สุดหาได้จาก

$$\bar{t} = \frac{1}{M} \sum_{k=1}^M t_k \quad (2.12)$$

t คือค่าเฉลี่ยตัวอย่าง

$$\begin{aligned} J_0(t_0) &= \sum_{k=1}^M \|(T_0 - \bar{t}) - (t_k - \bar{t})\|^2 \\ &= \sum_{k=1}^M \|(T_0 - \bar{t})\|^2 - 2 \sum_{k=1}^M (T_0 - \bar{t})^T (t_k - \bar{t}) + \sum_{k=1}^M \|(t_k - \bar{t})\|^2 \\ &= \sum_{k=1}^M \|(T_0 - \bar{t})\|^2 - 2(T_0 - \bar{t})^T \sum_{k=1}^M (t_k - \bar{t}) + \sum_{k=1}^M \|(t_k - \bar{t})\|^2 \\ &= \sum_{k=1}^M \|(T_0 - \bar{t})\|^2 + \sum_{k=1}^M \|(t_k - \bar{t})\|^2 \end{aligned} \quad (2.13)$$

ทำการฉายข้อมูลลงบนเส้นค่าเฉลี่ยตัวอย่าง ได้สมการความสัมพันธ์ของตัวแปรใหม่ดังสมการที่ 2.14 โดยที่ $\mathbf{u}$ คือเวกเตอร์ขนาด 1 หน่วย ที่มีทิศทางตามเส้นค่าเฉลี่ย และ y คือขนาดที่สัมพันธ์กันของระยะจาก t ถึง $\bar{t}$

$$\mathbf{t} = \bar{t} + y\mathbf{u} \quad (2.14)$$

พบว่าแกนฉายของข้อมูล 1 มิติ ก็คือไอเจนเวกเตอร์ที่สมนัยกับค่าไอเจนที่มากที่สุดตามลำดับ ดังนั้นในแนวทางเดียวกันเราสามารถขยายการแทนข้อมูลในรูปของ d มิติ โดยใช้การฉายข้อมูลลงเส้นข้อมูล d ดังแสดงในสมการที่ 2.15

$$\mathbf{t} = \bar{t} + y_d\mathbf{u}_d \quad (2.15)$$

และจะเรียกค่า y_d ว่าองค์ประกอบหลัก (Principal component) ของข้อมูลขนาด d มิติ

2. วิธีการเปลี่ยนตัวแปร มีแนวคิดมาจากการแปลงเชิงเส้นของตัวแปรที่มีความสัมพันธ์กันกับตัวแปรใหม่ที่มีความสัมพันธ์กันน้อยตามลำดับการลดลงของความแปรปรวน โดยกำหนดให้เวกเตอร์ t ใด ๆ ขนาด D มิติ มีค่าเฉลี่ย $\bar{t}$ และเมทริกซ์ความแปรปรวนร่วม ขนาด $D \times D$ มิติ สามารถคำนวณหาเวกเตอร์ใหม่ y ที่มีความสัมพันธ์กันน้อยตามลำดับการลดลงของค่าความแปรปรวน จากหลักการรวมกันเชิงเส้นแสดงความสัมพันธ์ดังสมการที่ 2.16

$$y_k = \mathbf{u}_1 t_1 + \mathbf{u}_2 t_2 + \dots + \mathbf{u}_{Dk} t_D = \mathbf{u}_k^T \mathbf{t} \quad (2.16)$$

จากสมการที่ 2.16 สามารถเขียนสมการใหม่ได้โดยที่ y มีขนาด $d \times 1$ มิติ จากเวกเตอร์ t ขนาด $D \times 1$ มิติ และ $\mathbf{U}$ ขนาด $D \times d$ มิติ

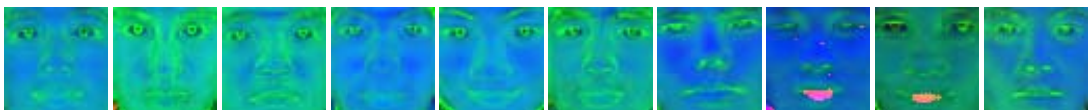
$$y = \mathbf{U}^T \mathbf{t} \quad (2.17)$$

จากสมการที่ 2.16 พบว่า $\mathbf{u}$ เป็นเมทริกซ์ที่ระบุแกนทิศทางขนาด 1 หน่วย และในสมการที่ 2.17 เมทริกซ์ $\mathbf{U}$ คือเมทริกซ์ความแปรปรวนร่วม และสามารถหาไอเจนเวกเตอร์ที่สมนัยกับค่าไอเจนที่เรียงจากมากไปน้อยได้จากเมทริกซ์ความแปรปรวนร่วม $\mathbf{U}$

2.4.1 การหาค่าไอเกนและไอเกนเวกเตอร์ ด้วยหลักการวิเคราะห์องค์ประกอบ เป็นวิธีการหาหน้าไอเกน (Eigenface) ในปริภูมิหน้า (face Space) นำเสนอ โดย Turk และ Pentland โดยทำการคำนวณหาไอเกนเวกเตอร์ของเมทริกซ์ความแปรปรวนของภาพหน้าตัวอย่าง ที่สมนัยกับค่าไอเกนที่มีค่าความแปรปรวนจากมากไปน้อย โดยนำภาพหน้าตัวอย่างดังกล่าวไปฉายด้วยไอเกนเวกเตอร์ จะได้ไอเกนเวกเตอร์ในปริภูมิหน้าไอเกน ซึ่งสามารถคำนวณได้โดยนำภาพหน้าตัวอย่าง จำนวน N ภาพ คือ $A_1, A_2, A_3, \dots, A_N$ ที่มีมิติขนาดเท่ากับคือ $m \times n$ มิติ



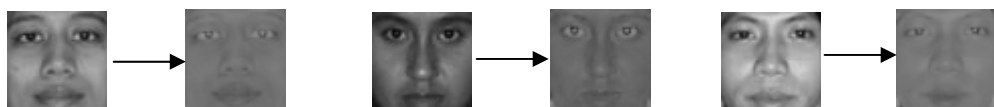
(ก)



(ข)

รูปที่ 2.13 ตัวอย่างภาพหน้าตัวอย่างในการหาหน้าไอเกน (ก) ตัวอย่างภาพในปริภูมิ RGB
(ข) ตัวอย่างภาพในปริภูมิ HSV

โดยในรูปที่ 2.13 แสดงภาพตัวอย่างก่อนจะนำภาพมาทำการหาค่าไอเกนเวกเตอร์ซึ่ง จำเป็นต้องนำภาพในปริภูมิ RGB มาแปลงเป็นปริภูมิ HSV เพื่อลดผลของค่าความสว่าง ซึ่งมีผลต่อความถูกต้องของระบบการตรวจจับหน้าคน แล้วจึงแปลงภาพที่ทำการลดผลของความสว่างแล้วนั้นเป็นปริภูมิ RGB อีกครั้ง และแปลงให้อยู่ในระดับเทาเพื่อให้เป็นภาพระนาบเดียว ดังแสดงในรูปที่ 2.14



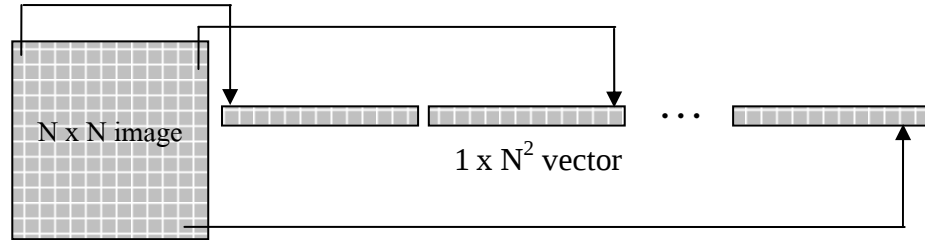
(ก)

(ข)

(ค)

รูปที่ 2.14 ภาพระดับเทาที่ได้จากการลดผลของความสว่าง (ก) สีผิวกลาง (ข) สีผิวเข้ม (ค) สีผิวอ่อน

แปลงเมทริกซ์ภาพตัวอย่างเป็นเวกเตอร์ขนาด $1 \times D$ ($D=m \times n$) มิติ ดังแสดงในรูปที่ 2.15



รูปที่ 2.15 ตัวอย่างการแปลงเมทริกซ์ภาพเป็นเวกเตอร์ภาพ

จากนั้นนำเวกเตอร์ที่ได้มามาจัดให้อยู่ในรูปของเมทริกซ์ ได้

$$\mathbf{I} = \{I_1, I_2, I_3, \dots, I_N\} \quad (2.18)$$

และนำเมทริกซ์ $\mathbf{I}$ ที่ได้มาหาค่ากลางของชุดข้อมูลจากสมการที่ 2.19

$$\phi_i = I_i - \frac{1}{N} \sum_{i=1}^N I_i \quad (2.19)$$

จะได้เวกเตอร์ค่ากลางของชุดข้อมูลในสมการที่ 2.20

$$\mathbf{A} = \{\phi_1, \phi_2, \phi_3, \dots, \phi_N\} \quad (2.20)$$

จากนั้นนำเมทริกซ์ $\mathbf{A}$ ไปหาค่าเมทริกซ์ความแปรปรวน (covariance matrix: $\mathbf{C}$) โดยเมทริกซ์ที่ได้มีขนาด $D \times D$ มิติ ดังแสดงในสมการที่ 2.21 และ 2.22

$$\mathbf{C} = \frac{1}{N} \sum_{i=1}^N \phi_i \phi_i^T = \mathbf{A} \mathbf{A}^T \quad (2.21)$$

$$\mathbf{A} \mathbf{v}_i = \mathbf{u}_i \quad (2.22)$$

จะได้ค่าไอเกน u_i จำนวนเท่ากับจำนวนภาพต้นแบบคือ N ค่า และไอเกนเวกเตอร์ v_i ขนาด $D \times D$ จุดภาพ ซึ่งเป็นข้อมูลขนาดใหญ่ พิจารณาเห็นว่าสามารถลดขนาดของค่าทั้งสองลงได้ โดยค่าไอเกนและไอเกนเวกเตอร์ของ AA^T และ $A^T A$ นั้นจะเป็นค่าที่มีความสอดคล้องกัน ดังนั้นเราจะหาค่าไอเกนและไอเกนเวกเตอร์จาก $L = A^T A$ จะได้ค่าไอเกน N ค่าและไอเกนเวกเตอร์มีขนาด $I \times N$ มิติ

จากค่าไอเกนเวกเตอร์ สามารถหาค่าหน้าไอเกน (Eigenface) ได้ โดยที่หน้าไอเกนค่า 1 นั้นจะแสดงลักษณะเฉพาะของหน้าได้ชัดเจนกว่าหน้าไอเกนค่า 2 ดังนั้น เราสามารถตัดหน้าไอเกนค่า 10 ออกได้โดยไม่ทำให้เกิดความผิดพลาดกับระบบ เนื่องจากค่าไอเกนค่าที่ 10 นั้นแสดงคุณสมบัติของความเป็นหน้าไว้น้อยมากจนสามารถตัดค่านี้ทิ้งได้ โดยไม่ส่งผลกระทบต่อการเล่นเป็นหน้าไอเกน ดังนั้นเราจะสนใจหน้าไอเกน k ค่า โดยที่ $1 \leq k < N$ และแสดงการหาหน้าไอเกนในโดเมนหน้า (face space) จากสมการที่ 2.23

$$uu_k = (Av_k)^T \left(I_i - \frac{1}{N} \sum_{i=1}^N I_i \right) \quad (2.23)$$

ภาพต้นแบบจำนวน 10 ภาพสามารถหาภาพหน้าไอเกนได้ดังแสดงในรูปที่ 2.16



รูปที่ 2.16 ภาพหน้าไอเกน แถวที่ 1 ภาพต้นแบบในระดับเทา แถวที่ 2 ภาพต้นแบบในระดับเทาที่ทำการลดผลของความสว่างแล้ว แถวที่ 3 ภาพหน้าไอเกนจากไอเกนเวกเตอร์ค่าที่ 1 แถวที่ 4 ภาพหน้าไอเกนจากไอเกนเวกเตอร์ค่าที่ 10

2.5 จีแนติกอัลกอริทึม (Genetic algorithm)

จีแนติกอัลกอริทึม หรือ GA (อาทิตย์ ศรีแก้ว, 2552) เป็นวิธีการค้นหาคำตอบที่ดีที่สุดโดยใช้หลักการคัดเลือกแบบธรรมชาติและหลักการทางสายพันธุ์ ประกอบไปด้วย 3 กระบวนการที่สำคัญได้แก่ การคัดเลือกสายพันธุ์ (selection) ซึ่งเป็นขั้นตอนในการคัดเลือกประชากรที่ดีในระบบไปเป็นต้นกำเนิดสายพันธุ์เพื่อให้กำเนิดลูกหลานในรุ่นถัดไป ปฏิบัติการทางสายพันธุ์ (genetic operation) ซึ่งเป็นกรรมวิธีการเปลี่ยนแปลงโครโมโซมด้วยวิธีการทางสายพันธุ์ และการแทนที่ (replacement) ซึ่งเป็นขั้นตอนการนำเอาลูกหลานกำเนิดใหม่ไปแทนที่ประชากรเก่าในรุ่นก่อน กระบวนการภายในของ GA ทำให้คำตอบของระบบที่มีอยู่เกิดวิวัฒนาการในตัวเองอันจะนำไปสู่การปรับตัวให้กลายเป็นคำตอบที่ดีกว่าและดีที่สุดได้ รายละเอียดขององค์ประกอบในวัฏจักร GA มีดังนี้

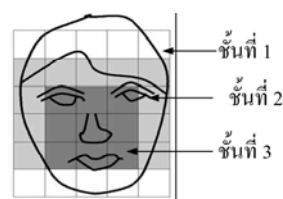
1. ประชากร (population) ประกอบไปด้วยกลุ่มของโครโมโซม (chromosome) ซึ่งเป็นตัวแทนของคำตอบในระบบที่ต้องการค้นหา
2. ต้นกำเนิดสายพันธุ์ (parents) กลุ่มประชากรที่ถูกคัดเลือกเพื่อเป็นตัวแทนในการให้กำเนิดสายพันธุ์ใหม่ในรุ่นถัดไป
3. สายพันธุ์ใหม่ (offspring) เป็นประชากรกลุ่มใหม่ที่ได้รับการถ่ายทอดสายพันธุ์มาจากต้นกำเนิดสายพันธุ์ โดยคาดหวังที่จะได้รับสายพันธุ์ที่ดีที่สุดเพื่อถ่ายทอดต่อ ๆ กันในประชากรรุ่นถัดไป (next generation)
4. ฟังก์ชันวัตถุประสงค์ (objective function) เป็นกระบวนการสำหรับการประเมินผลคำตอบของระบบว่าดีหรือไม่ดีแค่ไหน โครโมโซมที่นำไปประเมินค่าด้วยฟังก์ชันวัตถุประสงค์จะต้องอยู่ในรูปแบบที่ระบบเข้าใจ ซึ่งเรียกว่า รูปแบบฟีโนไทป์ (phenotype) หลังจากการประเมินผลคำตอบของระบบแล้วระบบจะส่งค่าความเหมาะสม (fitness value) เพื่อนำไปใช้เป็นเครื่องมือสำหรับตัดสินใจคัดเลือกโครโมโซมที่ดีเพื่อใช้สำหรับการสืบสายพันธุ์เป็นลูกหลานต่อไป โดยการประเมินค่าคำตอบของระบบนั้น ๆ ขึ้นอยู่กับฟังก์ชันวัตถุประสงค์ของงานแต่ละงาน และค่าความเหมาะสมที่ให้กับโครโมโซมแต่ละตัวจะมีการเปรียบเทียบกันเองในกลุ่มประชากร

GA มีขั้นตอนการทำงานโดยเชื่อมโยงเข้ากับระบบในโลกจริง เพื่อทำการค้นหาคำตอบที่ต้องการคำตอบของระบบที่ต้องการให้ GA ค้นหาจะอยู่ในรูปของโครโมโซมในกลุ่มของประชากร ดังนั้นระบบจะสามารถรู้ได้ว่าคำตอบที่มีอยู่ใน GA ณ เวลาหนึ่ง ๆ นั้นดีหรือไม่ดีอย่างไรด้วยการประเมินค่าของโครโมโซม โดยระบบจะมีการเชื่อมต่อกับ GA ผ่านฟังก์ชันวัตถุประสงค์สำหรับการใช้ในการประเมินค่าของโครโมโซม

2.6 งานวิจัยที่เกี่ยวข้อง

การค้นหาน้ำคนจากภาพนั้น เป็นงานวิจัยที่มีได้รับความสนใจศึกษากันอย่างกว้างขวาง เหตุผลที่ทำให้งานนี้ได้รับความสนใจเนื่องมาจากความท้าทายในการค้นหาใบหน้าบุคคล ที่มีความหลากหลายของใบหน้าที่ไม่ซ้ำกันในแต่ละบุคคล รวมถึงความเปลี่ยนแปลงไม่แน่นอนในการปรากฏของใบหน้า เช่น โครงสร้างทางกายภาพของใบหน้าที่แตกต่างกันไปตามเชื้อชาติ ตำแหน่งที่ตั้ง การวางท่า หรือการแสดงออกของสีหน้า เป็นต้น โดยในการวิจัยช่วงแรกจะพิจารณาเป็นภาพเดี่ยว สามารถแบ่งประเภทการตรวจจับออกได้เป็น 4 ประเภทหลัก ๆ คือ

2.6.1 วิธีเชิงความรู้ (Knowledge-based methods) เป็นวิธีการพิจารณาความสัมพันธ์ระหว่างลักษณะเด่นเฉพาะต่าง ๆ บนใบหน้า ที่มีตำแหน่งและองค์ประกอบหลักพื้นฐานตายตัวบนใบหน้า ตัวอย่างเช่น Yang and Huang (1994) ศึกษาวิธีเชิงความรู้แบบลำดับชั้น (hierarchical) ซึ่งพิจารณาแบ่งเป็น 3 ลำดับชั้น ดังแสดงในรูปที่ 2.17



รูปที่ 2.17 ขอบเขตการพิจารณาของแต่ละลำดับชั้น

ชั้นแรก พิจารณาคำถามตำแหน่งความน่าจะเป็นของใบหน้าโดยรวม ชั้นที่ 2 พิจารณาโดยค่าฮิสโตแกรม (histogram) ร่วมกับการหาเส้นขอบ ชั้นที่ 3 พิจารณาลักษณะเด่นภายในของใบหน้า เช่น ความสมมาตรกันของตาทั้ง 2 ข้าง ตำแหน่งของจมูกและปากในแนวตั้งและแนวนอน

ต่อมา Kotropoulos and Pitas (1997) ได้ศึกษาวิธีฐานกฎแบบกำหนดเขต (rule-based localization method) ซึ่งเป็นวิธีการที่คล้ายคลึงกับวิธีของ Yang and Huang แตกต่างกันที่พิจารณาค่าฮิสโตแกรมทั้งในแนวตั้งและแนวนอน แต่ยังไม่สามารถขจัดปัญหาของภาพที่มีพื้นหลังซับซ้อนและปัญหาความหลากหลายของใบหน้าบุคคล รวมทั้งการวางท่าที่แตกต่างกันได้

2.6.2 วิธีเชิงลักษณะ (Feature-based methods) เป็นการใช้อัลกอริทึมพิจารณาลักษณะเด่นและโครงสร้างของใบหน้า รวมทั้งความเปลี่ยนแปลงขององค์ประกอบต่าง ๆ ของภาพ เช่น แสงเงา ตัวอย่างเช่น Chetverikov and Lerch (1993) ใช้ความแตกต่างของแสงเงาและเส้นแนวเป็นแบบจำลองในการค้นหาใบหน้า โดยแบบจำลองนั้นประกอบด้วยจุดความสว่างน้อย (จุดมืด) 2 จุด เพื่อแสดงดวงตา และจุดความสว่างมาก (จุดอ่อน) 3 จุด เพื่อแสดงโหนกแก้มและจมูก แล้วหาความสัมพันธ์ของระยะห่างและตำแหน่งของจุดต่าง ๆ เพื่อคัดเลือกแบบหน้าที่เหมาะสม ระบบนี้มีข้อจำกัดอยู่ที่เมื่อแสงเงาของสภาวะแวดล้อมเปลี่ยนแปลงจะทำให้ประสิทธิภาพในการค้นหาเปลี่ยนไป ต่อมา Graf, Chen, Petajan, and

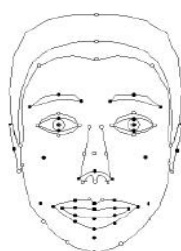
Cosatto (1995) พัฒนาทฤษฎีนี้โดยใช้งานในระบบภาพระดับเทา และนำตัวกรองช่วงผ่าน (band pass filter) ตัวกรองเกาส์เซียน (Gaussian filter) และค่าฮิสโทแกรม มาใช้หาค่าจุดสูงสุดและต่ำสุดเพื่อกำหนดขอบเขตของใบหน้า

ลักษณะเฉพาะอีกประการหนึ่งที่นิยมนำมาพิจารณาหาใบหน้าบุคคล คือค่าสีผิวมนุษย์โดยมีการค้นหาในหลาย ๆ ปริภูมิสี โดย Crowley and Berard (1997) ทำการเก็บค่าเวกเตอร์ของสีผิวมนุษย์โดยแยกค่าสีออกเป็น R, G และ B เพื่อพิจารณาค่าฮิสโทแกรมในรูปของสมการ $h(r,g) \geq \tau$ โดยค่า τ คือค่าเริ่มเปลี่ยน (threshold) ของค่าสีผิวมนุษย์ที่ได้จากการสำรวจข้อมูลกลุ่มตัวอย่าง ต่อมา Saxe and Fould (1996) ประยุกต์ใช้วิธีการดังกล่าวกับปริภูมิสี HSV และใช้แนวคิดค่าฮิสโทแกรมตัดผ่าน (histogram intersection) เปรียบเทียบค่าฮิสโทแกรมควบคุมกับค่าฮิสโทแกรมปัจจุบัน ซึ่งทำให้ประสิทธิภาพดีกว่าการใช้ค่าเริ่มเปลี่ยน

Chai and Ngan (1998) พิจารณาค่าสีผิวในปริภูมิสี YCrCb โดยพิจารณาในรูปของสมการ $Cr1 \leq Cr \leq Cr2$ และ $Cb1 \leq Cb \leq Cb2$ ซึ่งค่า $[Cr1, Cr2]$ และ $[Cb1, Cb2]$ เป็นค่าเริ่มเปลี่ยนของค่าสีผิวมนุษย์ที่ได้จากการสำรวจข้อมูลกลุ่มตัวอย่าง ซึ่งต้องระมัดระวังในการเลือกใช้ให้เหมาะสมกับภาพที่ทำการทดลอง และเป็นข้อเสียของการคัดแยกสีผิวมนุษย์ด้วยค่าเริ่มเปลี่ยน ต่อมามีการนำการคัดแยกสีผิวมาใช้ร่วมกับวิธีการอื่น ๆ ในการค้นหาใบหน้า ตัวอย่างเช่น Wu, Yokoyama, Pramadihanto, and Yachida, (1996) ใช้ทฤษฎีคลุมเครือ (fuzzy theory) ในการคัดแยกสีผิวมนุษย์และสีผิวในปริภูมิสี CIE XYZ โดยทำการแปลงจากปริภูมิสี RGB นอกจากนี้ยังมีการนำเอาลักษณะเด่นของใบหน้าอื่น ๆ มาใช้ในการค้นหาตัวอย่างเช่น Yokoo and Hagiwara (1996) ใช้คุณสมบัติของโครงหน้าทั่วไปของมนุษย์ที่เป็นรูปไข่ สร้างแบบจำลองรูปวงรีเปรียบเทียบพารามิเตอร์ร่วมกับการค้นหาแบบจินเนติกอัลกอริทึม (genetic algorithm) จากการศึกษาพบว่าวิธีการนี้มีข้อจำกัดอยู่ที่รายละเอียดของใบหน้าที่ต้องการค้นหาต้องมีความชัดเจนและมีขนาดใหญ่เพียงพอต่อการค้นหา

2.6.3 วิธีเทียบเคียงแผ่นแบบ (Template matching methods) เป็นการเปรียบเทียบภาพที่ต้องการค้นหากับโครงสร้างแบบจำลองของใบหน้ามาตรฐาน โดยเก็บข้อมูลความสัมพันธ์อย่างอิสระของส่วนต่าง ๆ บนใบหน้า ได้แก่ โครงสร้างใบหน้า ตา จมูก และปาก ในทำหน้าตรง ตัวอย่างเช่น Sakai, Nagao and Fujibayashi, (1969) เสนอโครงสร้างแบบจำลองแผ่นแบบย่อย (subtemplates model) โดยใช้ตัวกรองโซเบล (sobel filter) หาเส้นขอบ เพื่อหาดำแหน่งความน่าจะเป็นของส่วนย่อยต่าง ๆ บนใบหน้า ที่สามารถเข้ากันได้ดีที่สุดกับแบบจำลองแผ่นแบบย่อยให้เป็นตำแหน่งของใบหน้าที่ต้องการ ต่อมา Tsukamoto, Lee, and Tsuji (1994) เสนอแบบจำลองคุณภาพสำหรับรูปแบบหน้า (qualitative model for face pattern : QMF) โดยใช้พารามิเตอร์ของแสงสว่างและเส้นขอบใบหน้าเป็นแบบจำลองของใบหน้า และเทคนิคที่นิยมใช้อีกอันหนึ่งคือการวิเคราะห์องค์ประกอบหลัก (principal component analysis: PCA) โดย Jung, Lee C.W., and Lee Y.C (2002) ศึกษาการค้นหาดำแหน่งที่แน่นอนของใบหน้าโดยใช้หลักการพื้นฐานทางด้านเรขาคณิต วิเคราะห์องค์ประกอบของใบหน้าให้อยู่ในรูปของสมการทางคณิตศาสตร์

ต่อมาได้มีการพัฒนาแบบจำลองให้มีความสามารถปรับตัวได้เรียกว่า แผ่นแบบเปลี่ยนรูปได้ (deformable templates) โดยแผ่นแบบสามารถปรับตัวให้รองรับความยืดหยุ่นของโครงสร้างใบหน้าที่เปลี่ยนแปลงไปและลดปัญหาที่มีในแผ่นแบบคงตัว ตัวอย่างเช่น Malciu and Preteux (2002) นำแบบจำลองหน้าจากฐานข้อมูล MPEG-4 มาสร้างเป็นแบบจำลองแผ่นแบบพารามิเตอร์ (template parameterization) และพิจารณาความยืดหยุ่นของโครงสร้างใบหน้าเปรียบเทียบกับความยืดหยุ่นของสปริง เพื่อหาค่าพลังงานที่สปริงใช้น้อยที่สุดในการจะปรับตัวเพื่อให้ภาพที่หาเส้นขอบนั้นเข้ากับแบบจำลองมากที่สุด โดยทำการพิจารณาในส่วนของตาและปาก แสดงในรูปที่ 2.18

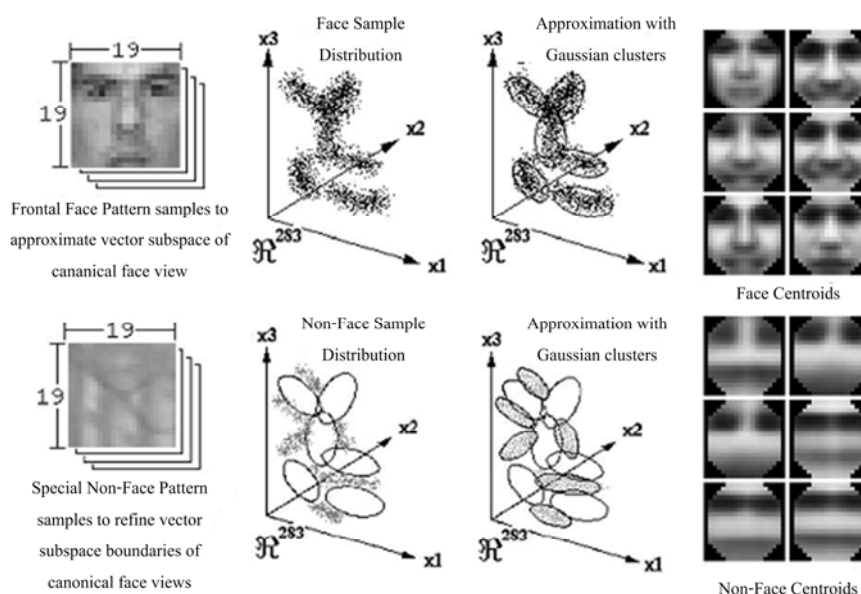


รูปที่ 2.18 แบบจำลองใบหน้ามนุษย์จากฐานข้อมูล MPEG-4

2.6.4 วิธีเชิงลักษณะปรากฏ (Appearance-based methods) เป็นการเปรียบเทียบภาพที่ต้องการค้นหากับโครงสร้างแบบจำลองของใบหน้าที่ทำกรเรียนรู้และฝึกสอนให้ระบบจดจำ และนำความรู้ในฐานข้อมูลมาใช้ในการพิจารณา โดยสามารถแบ่งย่อยได้หลายวิธีด้วยกัน ในที่นี้จะทำการเสนอเพียงวิธีการที่ได้รับความนิยมใช้งานกันอย่างกว้างขวาง 4 วิธีดังนี้

2.6.4.1 วิธีหน้าลักษณะเฉพาะ (Eigenface Methods) เป็นการวิเคราะห์องค์ประกอบพื้นฐานของใบหน้าด้วยกระบวนการทางสถิติของใบหน้าที่หลากหลาย โดยหน้าลักษณะเฉพาะคือเซตของเวกเตอร์ลักษณะเฉพาะ (eigen vector) ที่สามารถหาได้จากเมทริกซ์ความแปรปรวนร่วมเกี่ยว (covariance matrix) สร้างเป็นแบบจำลองของใบหน้าโดยรวมเอาลักษณะเด่นต่าง ๆ ของภาพใบหน้าตัวอย่างมารวมกันเพื่อหาค่าเฉพาะเจาะจงขององค์ประกอบบนใบหน้า ซึ่งในแต่ละบุคคลจะเป็นค่าเฉพาะของบุคคลนั้น ๆ ตัวอย่างเช่น Turk and Pentland (1991) นำภาพระดับเทาามาแปลงเป็นเวกเตอร์เพื่อหาค่าลักษณะเฉพาะ และนำค่าลักษณะเฉพาะของตัวอย่างภาพหน้าบุคคลมาสร้างเป็นแบบจำลองหน้าลักษณะเฉพาะเพื่อค้นหาคำแหน่งของใบหน้า

2.6.4.2 วิธีเชิงการกระจาย (Distribution-Base Methods) เป็นการแสดงการกระจายตัวของรูปแบบข้อมูลตัวอย่างที่มีความเป็นหน้าและความเป็นหน้าเพื่อใช้เป็นบรรทัดฐานในการตัดสินใจ ตัวอย่างเช่น Sung and Poggio (1998) นำฟังก์ชันเกาส์เซียน (Gaussian function) มาประมาณกลุ่มการกระจายของค่าเฉลี่ยกลุ่มตัวอย่าง แสดงในรูปที่ 2.19



รูปที่ 2.19 การกระจายตัวของข้อมูลตัวอย่าง

2.6.4.3 โครงข่ายประสาทเทียม (Artificial Neural Networks) เป็นการฝึกสอนโครงข่ายด้วยโครงสร้างใบหน้าที่มีความซับซ้อน โดยทำการปรับค่าน้ำหนักประสาทเพื่อให้ค่าความผิดพลาดกำลังสองเฉลี่ยมีค่าลดลง ความถูกต้องของโครงข่ายประสาทเทียมขึ้นกับตัวอย่างที่ทำการฝึกสอน จำนวนชั้นนิวรอน และจำนวนนิวรอนที่เหมาะสม ตัวอย่างเช่น Rowley, Baluja, and Kanade (1998) ใช้โครงข่ายประสาทเทียมแบบแพร่กลับในการฝึกสอนข้อมูลตัวอย่าง โดยทำการแบ่งข้อมูลออกเป็นภาพตัวอย่างใบหน้าที่บุคคลและภาพที่ไม่ใช่ใบหน้าที่บุคคลอย่างละ 1,048 ภาพ และแบ่งชุดข้อมูลออกเป็น 15 ชุดข้อมูล รวมข้อมูลตัวอย่างทั้งหมดได้ 15,720 ภาพ ซึ่งการฝึกสอนที่มีจำนวนข้อมูลตัวอย่างที่ครอบคลุมและมากเพียงพอจะช่วยลดความผิดพลาดของระบบ

2.6.4.4 วิธีเวกเตอร์เกือหนุน (Support Vector Machines : SVMs) เป็นโครงข่ายแบบมีการฝึกสอนคล้ายคลึงกับโครงข่ายประสาทเทียม แต่มีหลักการที่แตกต่างกันตรงที่โครงข่ายประสาทเทียมนั้นจะทำการฝึกสอนเพื่อลดผลของค่าความผิดพลาด (empirical risk minimization) แต่ SVMs นั้นจะทำการฝึกสอนเพื่อลดของโครงสร้างความเสี่ยงให้น้อยที่สุด (structural risk minimization) และต้องการข้อมูลตัวอย่างที่มาก เพื่อลดความผิดพลาดของระบบ ตัวอย่างเช่น Osuna, Freund, and Girosi (1997) ทำการศึกษาเพื่อลดจำนวนในการฝึกสอน SVMs ให้มีปริมาณลดลงแต่ประสิทธิภาพในการทำงานไม่ลดลงไปด้วย

นอกจากนั้นยังมีการนำเอาวิธีการต่าง ๆ ข้างต้นมาใช้งานร่วมกัน โดยพัฒนาให้สามารถลดขอบเขตในการตรวจจับหรือเพิ่มประสิทธิภาพในการตรวจจับใบหน้าที่บุคคลให้ดียิ่งขึ้น ตัวอย่างเช่น Mitsukura, Fukumi and Akamatsu (2001) ใช้การค้นหาแบบจีนเนติกอัลกอริทึมร่วมกับโครงข่ายประสาทเทียม LSDNN (Lip detection neural network and Skin Distinction Neural Network) โดยทำการ

ค้นหาตำแหน่งของปากด้วย LSDNN และทำการค้นหาความสัมพันธ์ของระยะห่างระหว่างตำแหน่งของปากและสี่ผิวนบริเวณโดยรอบด้วยการค้นหาแบบจีนเนติกอัลกอริทึม และ Li, Gong, Sherrah, and Liddell (1999) นำเสนอวิธีเวกเตอร์เกือหนุนร่วมกับปริภูมิลักษณะเฉพาะ (eigenspace) ในการตรวจจับใบหน้าบุคคล โดยใช้ปริภูมิลักษณะเฉพาะในการแบ่งขอบเขตความเป็นหน้าบุคคลออกจากสิ่งแวดล้อม และทำการพิจารณาร่วมกับการตรวจจับรูปแบบใบหน้าด้วยวิธีเวกเตอร์เกือหนุน

2.7 สรุป

บทที่ 2 นำเสนอรายงานผลการสืบค้นวรรณกรรมวิจัยย้อนหลังที่เกี่ยวข้องกับงานวิจัยที่จะดำเนินการ จากฐานข้อมูล IEEE IEE Science Direct และอื่น ๆ ซึ่งทำให้ทราบถึงแนวทางการวิจัยที่เกี่ยวข้อง ระเบียบวิธีที่ผู้วิจัยอื่น ๆ ได้นำมาใช้ ผลการดำเนินงาน ข้อเสนอแนะต่าง ๆ จากคณະนักวิจัยตั้งแต่อดีตจนถึงปัจจุบัน จากการสืบค้นปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง พบว่าเทคนิคการตรวจจับหน้าคนจากภาพสีส่วนใหญ่เป็นการวิจัยด้วยเทคนิคการประมวลผลภาพและคอมพิวเตอร์วิทัศน์ จากข้อมูลภาพทั้งจากการเก็บตัวอย่างภาพด้วยตัวเองและจากฐานข้อมูลรูปต่าง ๆ ที่ถ่ายภาพภายใต้การควบคุมปัจจัยต่าง ๆ ได้แก่ แสงสว่าง ความซับซ้อนของภาพพื้นหลัง และ ลักษณะการวางท่า เป็นต้น ซึ่งเทคนิคในการตรวจจับหน้าคนแต่ละแบบนั้น ล้วนมีข้อจำกัดต่าง ๆ กันไป จากแนวคิดดังกล่าวงานวิจัยนี้จึงได้ทำการพัฒนาเพื่อลดข้อจำกัดของขนาดข้อมูลภาพอินพุต เพื่อให้ระบบมีขนาดเล็กลง และสามารถทำงานได้ดีขึ้น โดยที่ระบบมีการพัฒนาตัวเองเพื่อให้ระบบเกิดความเป็นปัจจุบันอยู่เสมอ ซึ่งเป็นข้อจำกัดในงานวิจัยที่ใช้โครงข่ายประสาทเทียมในการตรวจจับหน้าคน

บทที่ 3

การตรวจจับหน้าคนด้วยโครงข่าย ART

3.1 กล่าวนำ

Adaptive Resonance Theory Network หรือ โครงข่าย ART นั้นมีการใช้งานในหลากหลายรูปแบบ ตัวอย่างเช่น Edelman, Intrator, and Poggio, (1998) ได้นำโครงข่ายมาช่วยในการคัดแยกภาพวัตถุทรงกระบอกที่ต้องการออกจากวัตถุอื่น ๆ และในการประยุกต์ใช้งานโครงข่ายประสาทเทียม ART เพื่อคัดแยกความเป็นหน้าบุคคลออกจากสิ่งแวดล้อนั้น จะทำการพิจารณาองค์ประกอบของใบหน้าบุคคลเฉพาะส่วนที่เป็นจุดเด่นของใบหน้า โดยภาพใบหน้าตัวอย่างที่มีความละเอียดของข้อมูลภาพที่ชัดเจนนั้นเป็นส่วนสำคัญในการลดข้อจำกัดของการตรวจจับและเพิ่มประสิทธิภาพในการตรวจจับ แต่จะพบว่าข้อมูลที่มีความละเอียดของภาพสูงนั้นจะมีขนาดของภาพที่ใหญ่ตามไปด้วย ดังนั้นในงานวิจัยนี้จึงได้อาศัยหลักการของวิธีการวิเคราะห์องค์ประกอบหลัก ในการลดขนาดของข้อมูลตัวอย่าง โดยวิธีการวิเคราะห์องค์ประกอบหลักนี้ จะแปลงองค์ประกอบหลักของข้อมูลภาพตัวอย่างให้อยู่ในรูปของไอเกนเวกเตอร์ ซึ่งมีขนาดของเมทริกซ์เท่ากับจำนวนข้อมูลตัวอย่าง ซึ่งแสดงรายละเอียดในหัวข้อที่ผ่านมา เพื่อนำผลลัพธ์ที่ได้เป็นอินพุตของโครงข่ายประสาทเทียม ART ในการคัดแยกความเป็นหน้าออกจากสิ่งแวดล้อน

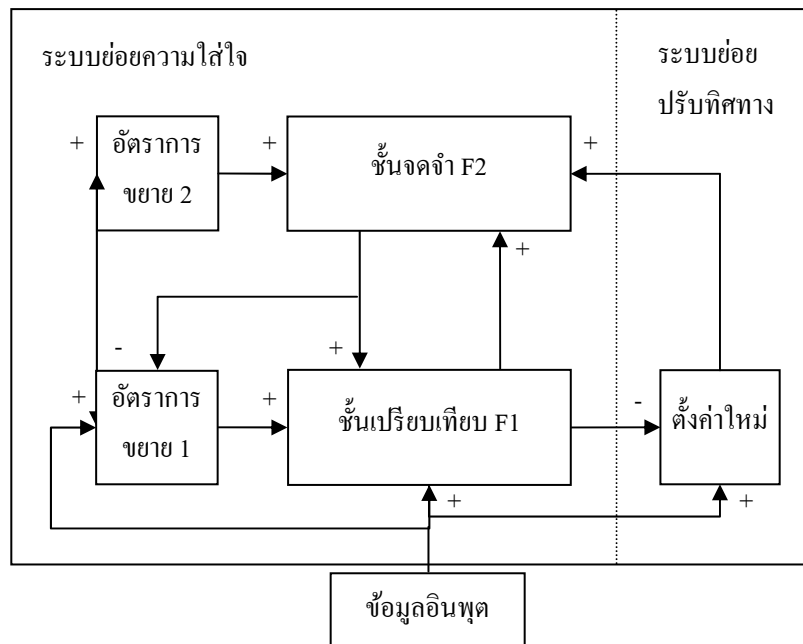
3.2 โครงข่าย ART (Adaptive Resonance Theory Network)

โครงข่าย ART เป็นโครงข่ายประสาทเทียมแบบไม่มีการฝึกสอน โดยเมื่อทำการป้อนค่าให้โครงข่าย โครงข่ายจะจดจำข้อมูลนั้นเป็นกลุ่มข้อมูลชุดที่ 1 เมื่อป้อนข้อมูลใหม่ที่มีค่าแตกต่างจากข้อมูลชุดแรก ระบบจะทำการเพิ่มกลุ่มข้อมูลใหม่เพื่อเก็บข้อมูลที่มีความแตกต่างกัน ซึ่งแตกต่างจากโครงข่ายประสาทเทียมแบบอื่นที่เมื่อทำการป้อนข้อมูลให้โครงข่ายด้วยข้อมูลชุดที่ 1 แล้ว ต้องการเพิ่มข้อมูลใหม่ให้โครงข่ายเรียนรู้เพิ่มเติม จำเป็นจะต้องทำการฝึกสอนโครงข่ายใหม่อีกครั้ง ด้วยข้อมูลชุดใหม่นั้น และทำให้โครงข่ายลืมความรู้เก่าที่ได้เรียนรู้มาจากข้อมูลชุดเดิมซึ่งปัญหานี้เรียกว่า ทวิบทเสถียรภาพสภาพพลาสติก (stability-plasticity dilemma) (Guo, Li, and Chan, 1998) ดังนั้น Carpenter and Grossberg (1991) จึงได้พัฒนาโครงข่าย ART เพื่อแก้ปัญหาดังกล่าว กล่าวคือโครงข่าย ART มีคุณสมบัติสำคัญ 2 ประการ คือสามารถคงความรู้ที่เคยเรียนรู้หรือรูปแบบที่เรียนรู้มาแล้วได้ (stability) และสามารถเรียนรู้ความรู้หรือรูปแบบใหม่เพิ่มเติมได้ไปพร้อม ๆ กัน (plasticity)

3.2.1 สถาปัตยกรรมของโครงข่าย ART โครงข่ายประสาทเทียม ART เป็นโครงข่ายประสาทเทียมแบบมีการฝึกสอนโดยใช้กฎการเรียนรู้แบบแข่งขัน แบ่งการทำงานออกเป็น 2 ระบบย่อย คือ

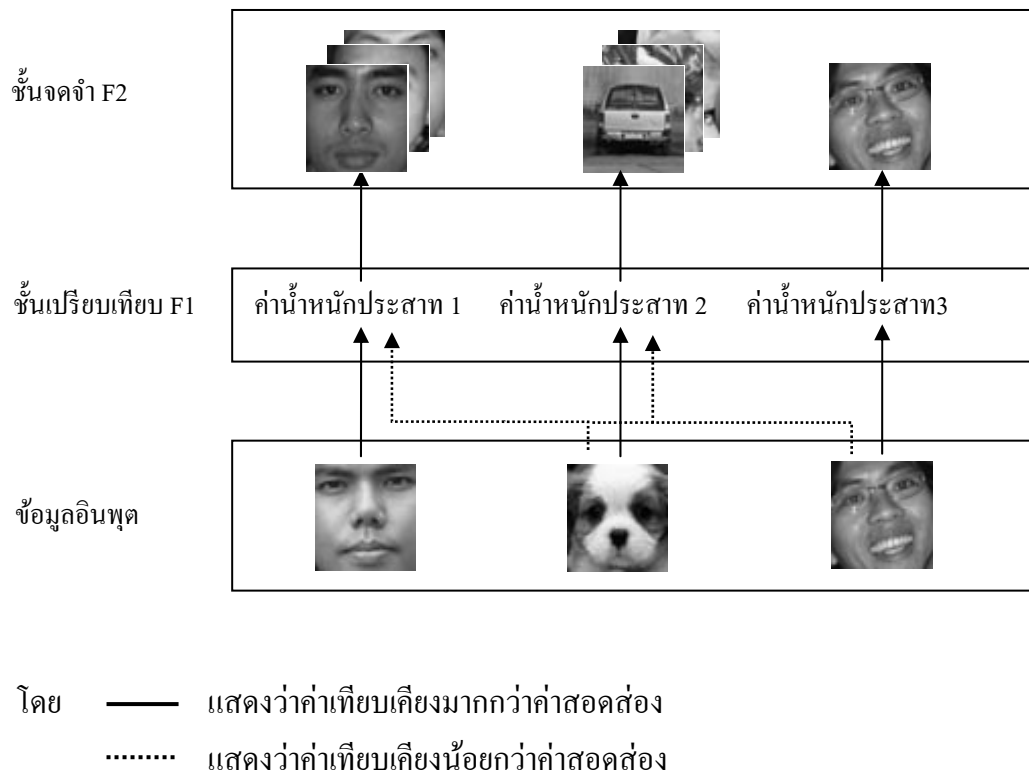
1. ระบบย่อยความใส่ใจ (attentional subsystem) มีโครงสร้างแบบ 2 ชั้น คือชั้นเปรียบเทียบ (comparison layer: F1) และชั้นจดจำ (recognition layer: F2)

2. ระบบย่อยปรับทิศทาง (orienting subsystem) ทำการปรับเทียบค่าน้ำหนักประสาทให้หมาะสมสอดคล้องกับค่าอินพุตของระบบที่เข้ามา
โครงสร้างของระบบทั้งสอง แสดงในรูปที่ 3.1



รูปที่ 3.1 แผนผังการทำงานของโครงข่าย ART

โครงข่ายประสาทเทียม ART ที่ได้รับการฝึกสอนแล้วจะเก็บค่าข้อมูลอยู่ในรูปของค่าน้ำหนักประสาท (weight) โดยในชั้นเปรียบเทียบจะนำข้อมูลอินพุต (input data) ที่เข้ามานั้น มาพิจารณาเปรียบเทียบกับค่าน้ำหนักประสาทของกลุ่มข้อมูลที่อยู่ในชั้นจดจำ เรียกค่านี้อีกว่าค่าเทียบเคียง ถ้าข้อมูลอินพุตนั้นมีค่าสอดคล้องกับค่าน้ำหนักประสาทกลุ่มใด ค่าเทียบเคียงที่ได้จากการเปรียบเทียบกับค่าน้ำหนักประสาทกลุ่มนั้นจะมีค่าน้อย และทำการคัดแยกกลุ่มข้อมูลอินพุตโดยการเปรียบเทียบค่าสอดคล้อง (vigilance: ρ) กับค่าเทียบเคียง ถ้าข้อมูลที่ได้มีค่าเทียบเคียงมากกว่าค่าสอดคล้อง ระบบจะทำการปรับค่าน้ำหนักประสาทของกลุ่มข้อมูลนั้น เพื่อทำการเก็บค่าข้อมูลใหม่ในกลุ่มที่มีค่าสอดคล้องกัน ถ้าข้อมูลที่ได้มีค่าเทียบเคียงน้อยกว่าค่าสอดคล้อง แสดงว่าข้อมูลอินพุตนั้นไม่สอดคล้องกับค่าน้ำหนักประสาทในกลุ่มนั้น ระบบจะส่งค่าน้ำหนักประสาทของกลุ่มใหม่มาทำการเปรียบเทียบกับข้อมูลอินพุตต่อไป แต่ถ้าข้อมูลอินพุตนั้นไม่สอดคล้องกับกลุ่มข้อมูลในชั้นจดจำกลุ่มใดเลยระบบจะทำการสร้างกลุ่มใหม่และเก็บข้อมูลอินพุตนั้นไว้ในกลุ่มใหม่ ดังแสดงในรูปที่ 3.2



รูปที่ 3.2 โครงข่ายประสาทเทียม ART

3.2.2 ประเภทของโครงข่าย ART สามารถแบ่งออกได้เป็น 4 ประเภท โดยจะแตกต่างกันไปตามข้อมูลที่รับเข้ามาและรูปแบบของการเรียนรู้ ได้แก่

3.2.2.1 โครงข่าย ART1 (Carpenter and Grossberg, 1987) เป็นระบบที่มีความซับซ้อนน้อยที่สุด โดยจะรับข้อมูลอินพุตที่มีค่าเป็น 0 และ 1 เท่านั้น โดยข้อมูลอินพุตจะส่งให้ชั้นเปรียบเทียบ F1 ตัดสินใจว่าข้อมูลอินพุตนั้นควรเป็นข้อมูลกลุ่มใด ด้วยสมการที่ 3.1 และ 3.2

$$Y_j = \sum_{i=1}^n (b_{ij} I_i) \quad (3.1)$$

$$Y_J = \text{Max}(Y_j) \quad (3.2)$$

โดย I คือ อินพุตของระบบ
 b คือ ไซนัส ของระบบ
 Y_J คือ ค่ามากที่สุดของกลุ่มที่ J

3.2.2.2 โครงข่าย ART2 (Carpenter and Grossberg, 1987) เป็นระบบที่ทำการพัฒนาต่อจาก ART1 ให้สามารถรับข้อมูลเชิงอุปมาน (analog data) จึงมีโครงสร้างของระบบที่ซับซ้อนมากกว่า ART1 โดยที่ ART2 นั้นจะมีโครงสร้างเหมือนกับ ART1 ยกเว้นในชั้นเปรียบเทียบ F1 จะทำการเพิ่มชั้นย่อย (sublayers) เพื่อให้สามารถรับข้อมูลแบบเชิงอุปมานได้ โดยข้อมูลอินพุตเชิงอุปมานนั้นมีขนาดใหญ่และมีความแตกต่างของข้อมูลมากกว่าข้อมูลอินพุตที่มีค่าเป็น 0 และ 1 เท่านั้น ที่ป้อนให้โครงข่าย ART1 ดังนั้นชั้นย่อยที่เพิ่มมาในชั้นเปรียบเทียบ F1 จึงทำหน้าที่ย่อยข้อมูลขนาดใหญ่ขึ้นให้เล็กลงตามสมการที่ 3.3

$$f(x) = \begin{cases} x & \text{if } x \geq \theta \\ 0 & \text{if } x < \theta \end{cases} \quad (3.3)$$

โดยค่า θ เป็นค่าที่ผู้ใช้โครงข่ายกำหนดเลือกค่าตามความเหมาะสมของข้อมูลอินพุต

3.2.2.3 โครงข่าย ARTMAP (Carpenter, Grossberg, and Reynolds, 1991) เป็นระบบที่มีผู้ฝึกสอน (supervised) และเสมือนมีโครงข่าย ART 2 ชุด โดยชุดแรกจะทำหน้าที่รับข้อมูลอินพุต และชุดที่ 2 รับข้อมูลเอาต์พุตที่มีความสัมพันธ์กับข้อมูลอินพุต ระบบจะทำการเรียนรู้จากความสัมพันธ์ของข้อมูลอินพุตกับเอาต์พุตนั้น โดยคำว่า ARTMAP นั้นมีความหมายว่าระบบจะดำเนินการเรียนรู้เพื่อจับคู่หรือแมพ (match) ข้อมูลอินพุตและเอาต์พุตที่ต้องการ โดยอินพุตก่อนเข้าโครงข่ายต้องผ่านการทำตัวเข้ารหัสแบบเติมเต็ม (Complement Coder) เพื่อปรับข้อมูลอินพุตให้มีขนาดเป็นสองเท่า ทำให้โครงข่ายสามารถพิจารณาขอบเขตของรูปแบบข้อมูลได้ครอบคลุมขึ้น จากนั้นส่งข้อมูลไปทำการเปรียบเทียบเพื่อหากลุ่มข้อมูลที่เหมาะสมที่สุด โดยทำการเปรียบเทียบด้วยสมการที่ 3.4 และ 3.5

$$Y_j(I) = \frac{|I \cap w_j|}{\alpha + |w_j|} \quad (3.4)$$

$$Y(I)_J = \text{Max}(Y_j) \quad (3.5)$$

เมื่อ $\cap$ ในสมการที่ 3.4 คือตรรกศาสตร์ส่วนร่วมรวม (logical AND intersection) เป็นสัญลักษณ์ทางคณิตศาสตร์แสดงการรวมกันของข้อมูลอินพุตและค่าน้ำหนักประสาทที่มีค่าเหมือนกัน โดยกำหนดให้ค่า α มีค่าน้อยมากๆ และ Y_j คือกลุ่มที่มีความสอดคล้องกับอินพุตมากที่สุด

3.2.2.4 โครงข่าย Fuzzy ART หรือ Fuzzy ARTMAP (Carpenter, Grossberg, and Rosen, 1991) เป็นระบบที่ได้รับการพัฒนามาจากโครงข่าย ART และโครงข่าย ARTMAP โดยนำ

ตรรกศาสตร์คลุมเครือ (fuzzy logic) มารวมกับโครงข่าย ดังนั้นในการเปรียบเทียบข้อมูลอินพุตกับค่าน้ำหนักประสาทจึงใช้กฎตรรกศาสตร์มาพิจารณาด้วย ดังสมการที่ 3.6 และ 3.7

$$T_j(I) = \frac{|I \wedge w_j|}{\alpha + |w_j|} \quad (3.6)$$

$$T(I)_J = \text{Max}(T_j) \quad (3.7)$$

เมื่อ $\wedge$ ในสมการที่ 3.6 คือตรรกศาสตร์คลุมเครือค่าต่ำสุดรวม (fuzzy AND minimum) เป็นสัญลักษณ์ทางคณิตศาสตร์แสดงการรวมกันของข้อมูลอินพุตและค่าน้ำหนักประสาทโดยพิจารณาค่าที่ต่ำที่สุดในรวมกัน และถ้าค่า T_j มีค่ามากที่สุดมากกว่า 1 ค่า ระบบจะทำการเลือกเอาต์พุตของข้อมูลกลุ่มแรกเป็นกลุ่มที่เหมาะสมที่สุดของอินพุตนั้น

3.3 ระบบตรวจจับหน้าคนจากภาพด้วยโครงข่าย ART

ค่าไอเกนเวกเตอร์ที่คำนวณได้ในบทที่ 2 ถูกนำมาเป็นอินพุตของโครงข่าย ART ซึ่งในที่นี้พิจารณาใช้โครงข่าย ART แบบ ARTMAP เนื่องจากเป็นโครงข่ายที่พัฒนามาจาก ART1 และ ART2 เพื่อให้สามารถรับข้อมูลอินพุตที่มีความซับซ้อนมากขึ้น ซึ่งหลักการในการคัดแยกกลุ่มข้อมูลของโครงข่ายทำให้สามารถแยกข้อมูลที่มีความซับซ้อนมากได้ชัดเจนมากกว่า ART1 และ ART2 ส่วนโครงข่ายแบบ Fuzzy ART หรือ Fuzzy ARTMAP นั้น ใช้หลักตรรกศาสตร์คลุมเครือมาช่วยในการคัดแยกกลุ่มข้อมูล ซึ่งมีโครงสร้างที่ซับซ้อนมากขึ้นไป ในการคัดแยกกลุ่มข้อมูลที่ต้องการแยกเพียงความเป็นหน้ามนุษย์และความไม่เป็นหน้ามนุษย์เท่านั้น

3.3.1 โครงสร้างของระบบ ป้อนชุดข้อมูลตัวอย่างเพื่อกำหนดกลุ่มให้โครงข่าย ART ด้วยค่าไอเกนเวกเตอร์ที่ได้จากใบหน้าตัวอย่าง โดยใช้ใบหน้าตัวอย่างจำนวน 20 หน้า ทำให้ได้ไอเกนเวกเตอร์ที่เป็นชุดข้อมูลตัวอย่างของหน้าคนเป็นกลุ่มข้อมูล กลุ่มที่ 1 ของโครงข่ายมีขนาด 20x20 มิติ และในทำนองเดียวกันทำการป้อนข้อมูลตัวอย่างที่ไม่ใช่ใบหน้าคนให้โครงข่าย ART ทำให้ชั้นเปรียบเทียบในโครงข่าย ART มีข้อมูลจดจำอยู่ 2 กลุ่ม คือ กลุ่มหน้าคนและกลุ่มที่ไม่ใช่หน้าคน โดยในการเปรียบเทียบขั้นแรกนี้เป็นการเปรียบเทียบเพื่อตรวจจับหน้าคน โดยเปรียบเทียบกับค่าน้ำหนักประสาทของกลุ่มข้อมูลที่เก็บค่าไว้ชั้นจดจำ และโครงข่ายจะทำการจำแนกกลุ่มข้อมูลที่เข้าโครงข่าย โดยเมื่อมีชุดข้อมูลใหม่เข้าสู่โครงข่ายโครงข่ายจะทำการตรวจสอบชุดข้อมูลนั้นโดยการเปรียบเทียบค่าเทียบเคียง (match: μ) กับค่าสอดคล้อง (ρ) เพื่อหาค่าตรวจสอบ (check: C) ในพิจารณาว่า โครงข่ายทำการคัดแยกกลุ่มข้อมูลได้ถูกต้องเพียงใด ถ้าข้อมูลตัวอย่างที่ป้อนให้โครงข่าย ART เป็นข้อมูลหน้าคน จะสามารถพิจารณาได้ 3 กรณีคือ

1. ข้อมูลตัวอย่างมีค่า μ มากกว่า ρ_1 และ ρ_2 จะให้ค่าตรวจสอบมีค่าเป็น 1 และโครงข่าย ART จะทำการปรับค่าน้ำหนักประสาทของกลุ่มข้อมูล

2. ข้อมูลตัวอย่างมีค่า μ มากกว่าค่า ρ_1 แต่น้อยกว่า ρ_2 จะให้ค่าตรวจสอบมีค่าเป็น 1 และโครงข่าย ART จะทำการสร้างกลุ่มข้อมูลใหม่

3. ข้อมูลตัวอย่างมีค่า μ น้อยกว่า ρ_1 จะให้ค่าตรวจสอบมีค่าเป็น 0 และโครงข่าย ART จะทำการปรับค่าสอดคล้อง $\rho_1 = \mu$

และถ้าข้อมูลตัวอย่างที่ป้อนให้โครงข่าย ART เป็นข้อมูลที่ไม่ใช่หน้าคน จะสามารถพิจารณาได้ 3 กรณีเช่นกันคือ

1. ข้อมูลตัวอย่างมีค่า μ มากกว่า ρ_1 และ ρ_2 จะให้ค่าตรวจสอบมีค่าเป็น 0 และโครงข่าย ART จะทำการปรับค่าสอดคล้อง $\rho_2 = \mu$

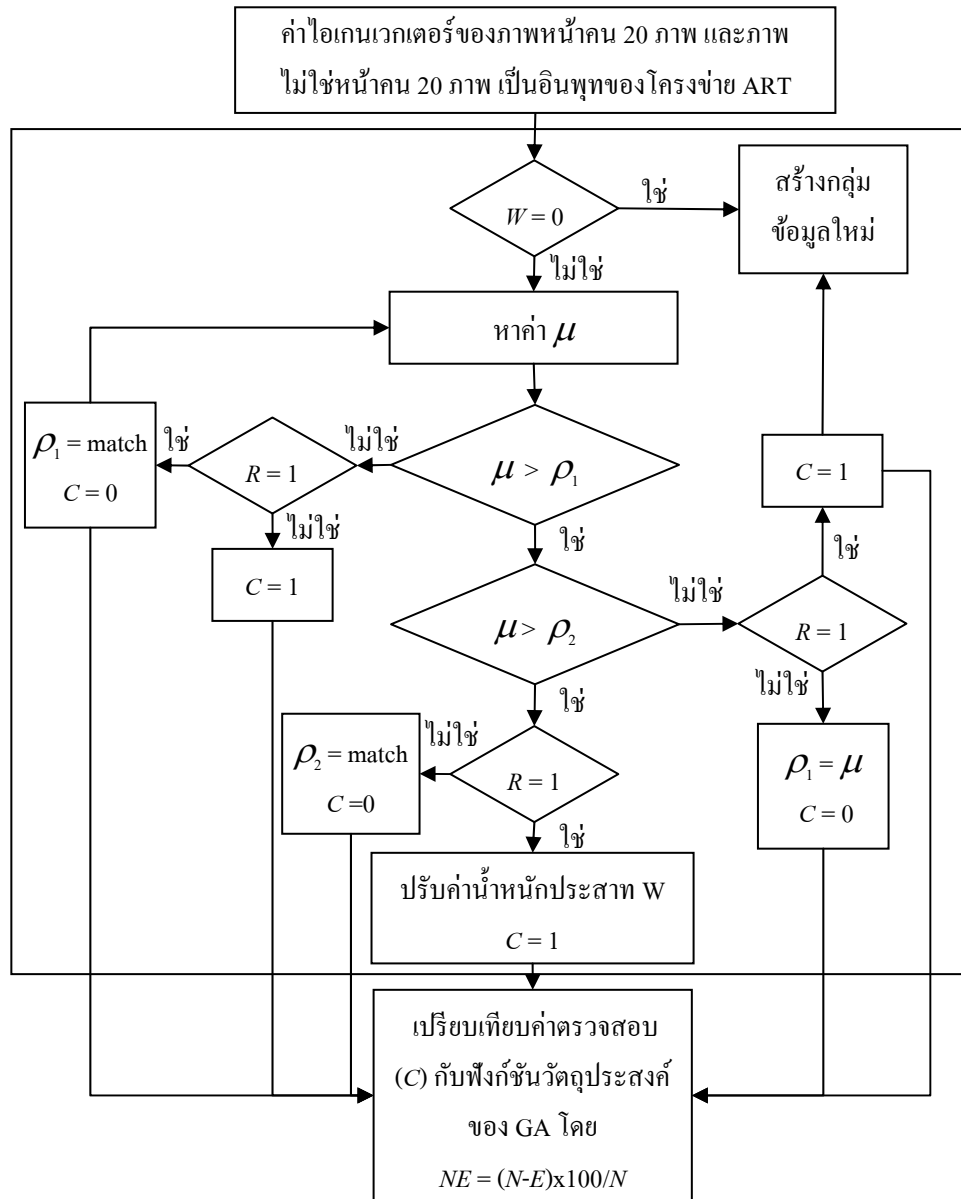
2. ข้อมูลตัวอย่างมีค่า μ มากกว่าค่า ρ_1 แต่น้อยกว่า ρ_2 จะให้ค่าตรวจสอบมีค่าเป็น 0 และโครงข่าย ART จะทำการปรับค่าสอดคล้อง $\rho_1 = \mu$

3. ข้อมูลตัวอย่างมีค่า μ น้อยกว่า ρ_1 จะให้ค่าตรวจสอบมีค่าเป็น 1

โดยแสดงผลแผนผังการทำงานของโครงข่าย ART ในการหาค่าสอดคล้องที่เหมาะสมด้วย GA ในรูปที่

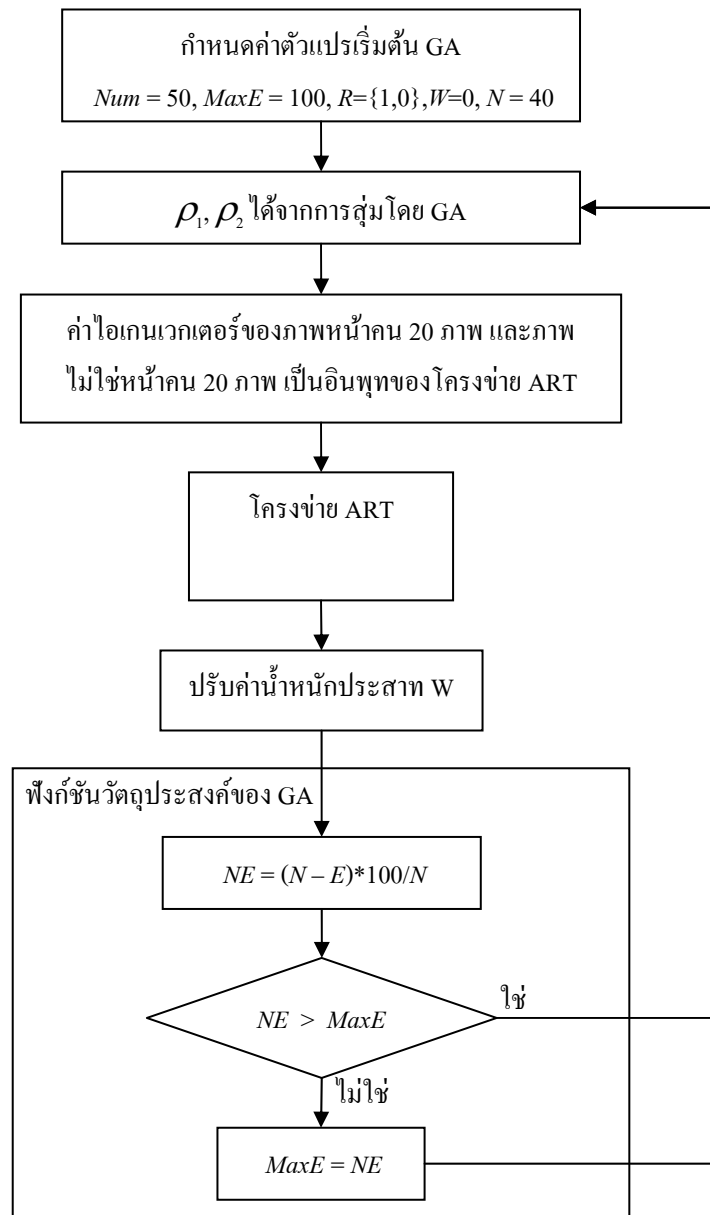
3.3

เนื่องจากชุดข้อมูลที่เป็นหน้าคนนั้น เป็นข้อมูลที่เกี่ยวข้องกับองค์ประกอบของหน้าในภาพที่มีความหลากหลายของหน้าแตกต่างกันตาม ลักษณะการวางท่า อารมณ์ของหน้า ขนาดของหน้า เป็นต้น ทำให้การกำหนดค่าสอดคล้องที่เหมาะสมทำได้ยาก ดังนั้นจึงนำ GA มาช่วยในการหาค่าสอดคล้องที่เหมาะสมที่สุดในการแบ่งกลุ่มข้อมูลของโครงข่าย ART โดยทำการสร้างฟังก์ชันวัตถุประสงค์ให้ทำการแยกข้อมูลของหน้าคนจำนวน 20 หน้า และข้อมูลที่ไม่ใช่หน้าคนอีก 20 หน้าผ่านโครงข่าย ART โดยกำหนดค่าพารามิเตอร์เริ่มต้นของ GA ได้แก่จำนวนประชากรในการเจริญพันธุ์จำนวน 50 ประชากร ($Num = 50$) ค่าความผิดพลาดเริ่มต้น เท่ากับ 100 ($MaxE = 100$) ค่าต้นแบบ (R) เป็นเซตแสดงสถานะของอินพุตว่าเป็นหน้าคนหรือไม่ใช่หน้าคน ในที่นี้ใช้ตัวอย่างที่เป็นหน้าคน 20 ตัวอย่าง และไม่ใช่นหน้าคน 20 ตัวอย่าง ดังนั้นจะได้ค่า R เป็นเซตของตัวอย่างหน้าคน และตัวอย่างที่ไม่ใช่หน้าคน โดยที่ A มีค่าเป็น 1 และ B มีค่าเป็น 0 ค่าน้ำหนักประสาทเริ่มต้นเป็น 0 ($W = 0$) และจำนวนภาพตัวอย่างในการทดลอง 40 ตัวอย่าง ($N = 40$) จากนั้น ทำการกำหนดค่าตัวแปรใน GA ค่าสอดคล้องเริ่มต้น ρ_1 และ ρ_2 โดยการสุ่มเลือกอิสระและส่งค่าสอดคล้องที่ได้จากการสุ่มให้โครงข่าย ART



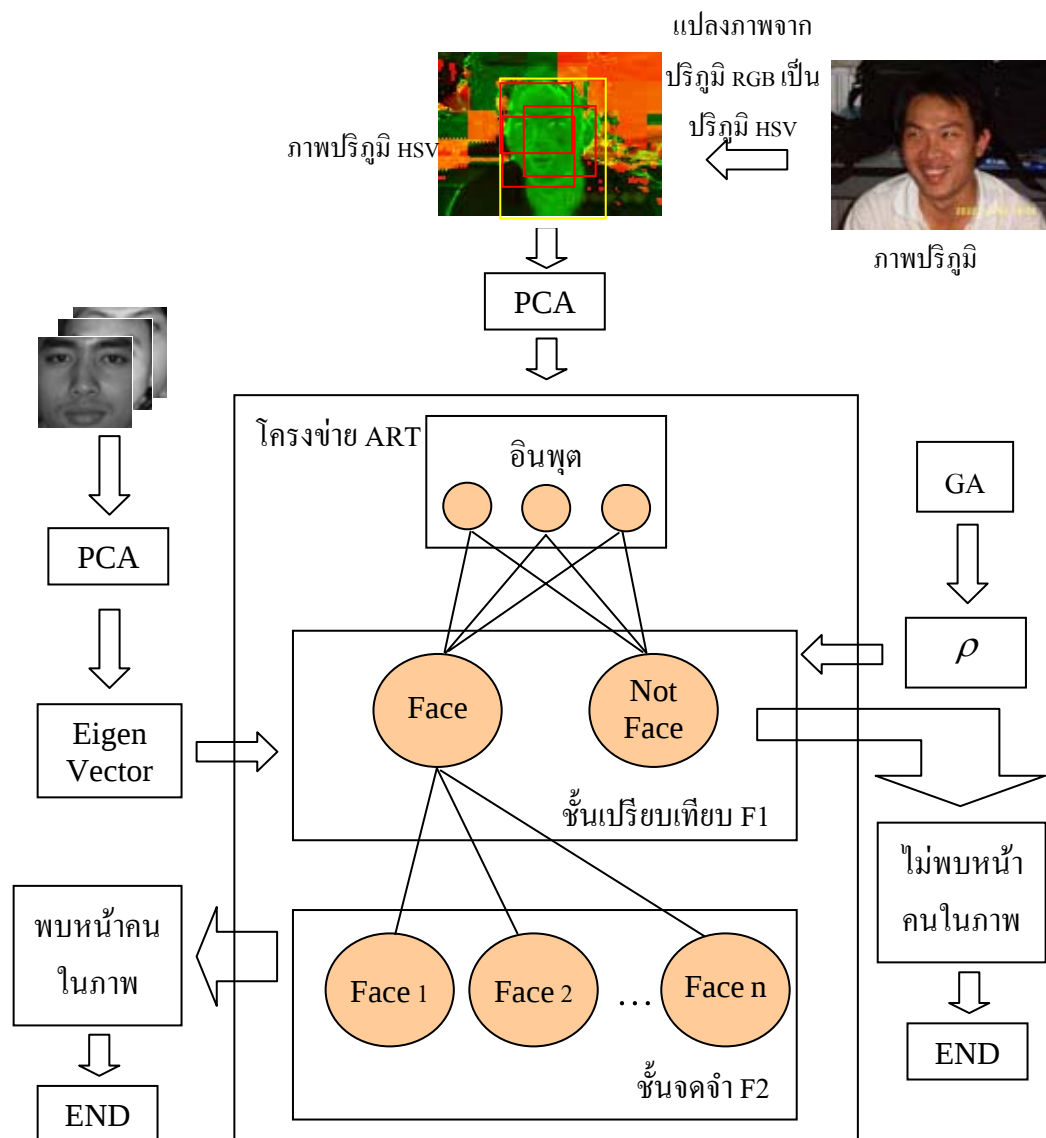
รูปที่ 3.3 แผนผังการทำงานของโครงข่าย ART ในการหาค่าสอดคล้องที่เหมาะสมด้วย GA

เมื่อได้ค่าตรวจสอบแล้วจะส่งค่าที่ได้ไปยังวัตถุประสงค์ของ GA เพื่อหาค่าความผิดพลาด (Net error: NE) โดย GA จะทำการปรับปรุงพันธุกรรมของประชากรเพื่อให้ได้ตัวแปรที่มีค่าความผิดพลาดที่น้อยที่สุด และสิ้นสุดการทำงานเมื่อ GA ครบจำนวนรอบสูงสุดที่กำหนดไว้ แสดงแผนผังการหาค่าสอดคล้องที่เหมาะสมโดย GA ในรูปที่ 3.4



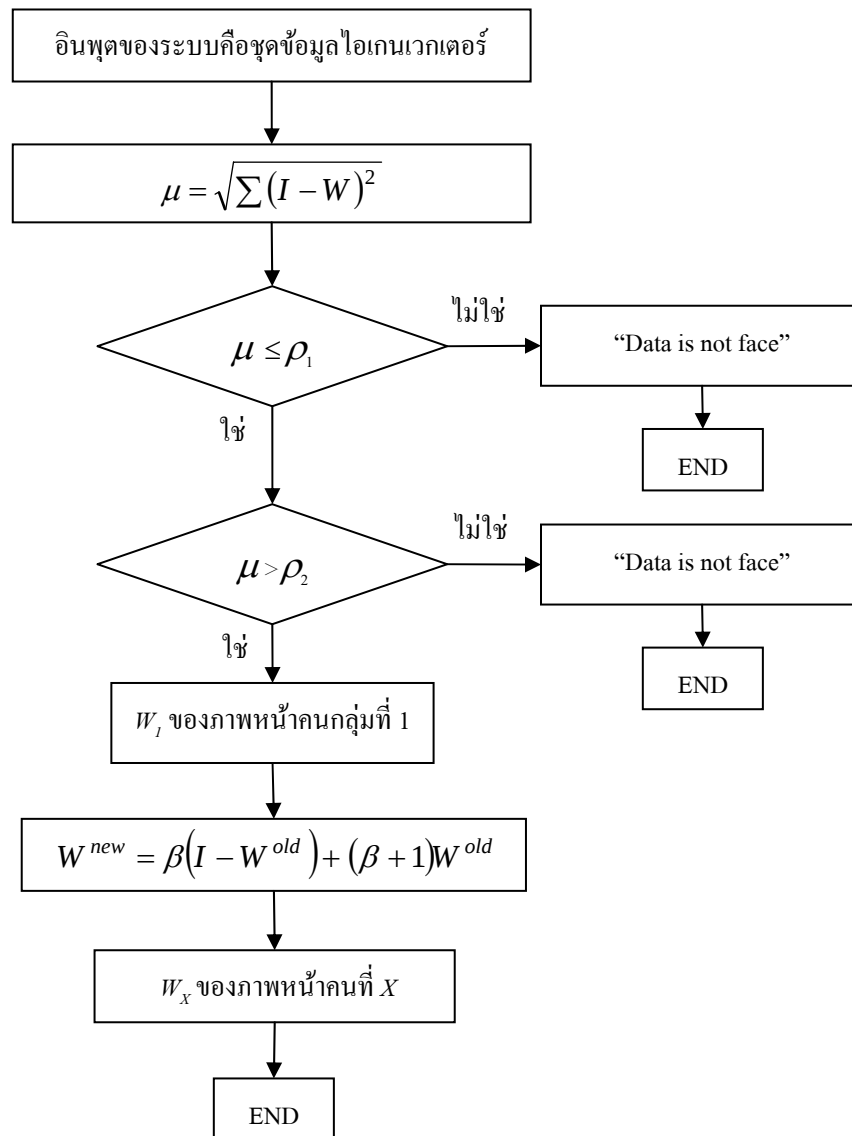
รูปที่ 3.4 แผนผังการหาค่าสอดคล้องที่เหมาะสมโดย GA

เมื่อได้ค่าสอดคล้องที่เหมาะสมกับระบบแล้ว นำค่าที่ได้นั้นมากำหนดเป็นค่าสอดคล้องของโครงข่าย ART โดยโครงสร้างของระบบการตรวจจับหน้าคนด้วยโครงข่าย ART มีโครงสร้างดังแสดงในรูปที่ 3.5



รูปที่ 3.5 โครงสร้างระบบการตรวจจับหน้าคนด้วยโครงข่าย ART

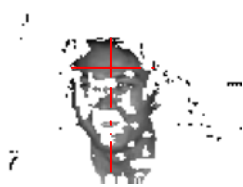
จากนั้นนำค่าสอดคล้องที่ได้รับการกระบวนการ GA มาใช้ในการเปรียบเทียบกับค่าเทียบเคียงที่ได้จากการคำนวณน้ำหนักประสาทของกลุ่มข้อมูลที่เหมาะสมที่สุด และในชั้นที่ 2 จะทำการเปรียบเทียบเพื่อแยกประเภทของหน้าที่มีลักษณะต่างกัน โดยแสดงโครงสร้างของโครงข่าย ART ในรูปที่ 3.6



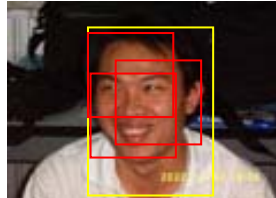
รูปที่ 3.6 แผนผังโครงข่าย ART ในการคัดแยกหน้าคน

นำภาพที่ต้องการตรวจสอบมาแปลงจากปริภูมิสี RGB เป็นปริภูมิสี HSV จากนั้นมาทำการคัดแยกสีผิวมนุษย์เพื่อกำหนดขอบเขตในการค้นหาให้แคบลง ได้ขอบเขตในการค้นหาที่คาดว่าน่าจะเป็นหน้าคน เนื่องจากมีค่าอยู่ในช่วงของสีผิวมนุษย์นั้น ต่อมาทำการหาขนาดของกล่องข้อมูลที่จะเข้าโครงข่าย ART ในรูปที่ 3.7 (ข) ในกรอบสี่เหลี่ยมคือขอบเขตทั้งหมดที่มีค่าสีผิวมนุษย์ที่ได้จากการคัดแยกสีผิวมนุษย์อยู่ และกรอบสีแดงคือกล่องข้อมูลที่จะป้อนเข้าโครงข่าย ART โดยสามารถหาได้จาก การหาดำแหน่งของค่าจุดสีของค่าสีผิวมนุษย์ที่มีค่ามากที่สุดทั้งในแนวตั้ง และแนวนอน เพื่อหาค่าที่น้อยที่สุดของทั้ง 2 แนวเป็นความยาวของกล่องข้อมูล แสดงในรูปที่ 3.7 (ก) เมื่อได้ขนาดของกล่องข้อมูลแล้ว ทำการหาดำแหน่งของกล่องภาพที่จะเป็นอินพุตป้อนให้โครงข่าย ART โดยทำการเลือกกล่องข้อมูลที่มีค่าจุดสีที่เป็นสีผิวมนุษย์ในกล่องข้อมูลไม่น้อยกว่า 2 ใน 3 ของกล่องข้อมูลที่มีค่าจุดสีที่เป็นสีผิวมนุษย์ในกล่องมากที่สุด เป็นข้อมูล

อินพุตป้อนเข้าโครงข่าย ART โดยทำการเปรียบเทียบกล่องข้อมูลอินพุตที่จะเข้าโครงข่ายกับข้อมูลหน้าในชั้นเปรียบเทียบชั้นแรกเพื่อหาตำแหน่งที่เหมาะสมที่สุดที่มีหน้าคนปรากฏอยู่ดังแสดงในรูปที่ 3.7 (ก)



(ก)



(ข)



(ค)

รูปที่ 3.7 (ก) เส้นกำหนดขนาดกล่องข้อมูล (ข) กล่องข้อมูลอินพุตที่ป้อนเข้าโครงข่าย ART
(ค) กล่องข้อมูลที่ดีที่สุดที่เป็นตำแหน่งของหน้าคนในภาพ

3.3.2 หลักการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART ทำการสร้างโครงข่ายประสาทเทียม แปลงข้อมูลอินพุตของค่าไอเกนเวกเตอร์ในสมการที่ 3.8 ให้อยู่ในรูปของข้อมูลส่วนเติมเต็ม (complement data) แสดงในสมการที่ 3.9

$$v_1, v_2, \dots, v_M \quad (3.8)$$

$$I = (v, v^c) = (v_1, v_2, \dots, v_M, 1 - v_1, 1 - v_2, \dots, 1 - v_M) \quad (3.9)$$

โดย I คือข้อมูลส่วนเติมเต็ม

V คือไอเกนเวกเตอร์

ฝึกสอนโครงข่าย ด้วยกฎการเรียนรู้แบบแข่งขันแสดงในสมการที่ 3.10

$$W_i^{new} = \begin{cases} W_i^{old} & \text{if } \mu \geq \rho \\ \beta(I - W_i^{old}) + (1 + \beta)W_i^{old} & \text{if } \mu < \rho \end{cases} \quad (3.10)$$

โดย W^{new} คือค่าน้ำหนักประสาทใหม่

W^{old} คือค่าน้ำหนักประสาทเดิม

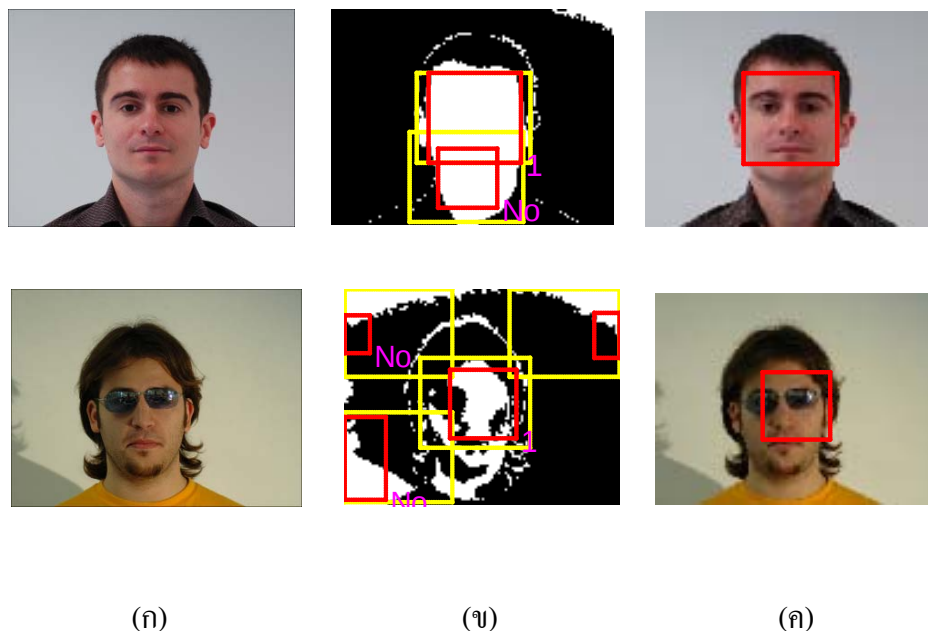
β คือค่าอัตราการเรียนรู้มีค่าอยู่ระหว่าง 0 ถึง 1

โครงข่าย ART จะทำการคัดแยกกลุ่มข้อมูลออกเป็น 2 กลุ่มข้อมูลคือ กลุ่มที่ 1 เป็นข้อมูลของภาพใบหน้าบุคคล และกลุ่มที่ 2 เป็นข้อมูลของภาพที่ไม่ใช่ใบหน้าบุคคล โดยพิจารณาคัดแยกด้วยค่าเทียบเคียง (μ) ดังแสดงในสมการที่ 3.11 เปรียบเทียบกับค่าสอดคล้อง (ρ)

$$\mu = \sqrt{(I - w_i^{old})^2} \quad (3.11)$$

ถ้าค่าเทียบเคียงที่ได้มีค่าน้อยกว่าค่าสอดคล้อง แสดงว่าข้อมูลอินพุตกับค่าน้ำหนักประสาทที่ใช้ไม่ได้ของกลุ่มเดียวกัน ระบบจะทำการเปรียบเทียบข้อมูลอินพุตกับค่าน้ำหนักประสาทของกลุ่มใหม่ จนกว่าจะได้กลุ่มที่ตรงกัน และถ้าข้อมูลอินพุตที่ได้นี้ ไม่ตรงกับกลุ่มข้อมูลของระบบกลุ่มใดเลย ระบบจะทำการจดจำข้อมูลอินพุตนี้ในกลุ่มที่สร้างขึ้นใหม่ต่อไป

3.3.3 ตัวอย่างการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART เมื่อกำหนดขอบเขตการค้นหาโดยการพิจารณาบริเวณที่เป็นสี่ผิวนุญช์เท่านั้น ระบบจะทำการสร้างกล่องข้อมูลอินพุต และนำกล่องข้อมูลนั้นไปแปลงเป็นไอเกนเวกเตอร์เพื่อเป็นอินพุตให้โครงข่าย ART เพื่อคัดแยกความเป็นหน้าคน และไม่ใช่หน้าคนโดยเปรียบเทียบค่าเทียบเคียงของข้อมูลกับค่าสอดคล้องของระบบ แสดงได้ในรูปที่ 3.8



รูปที่ 3.8 ภาพตัวอย่างการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART

(ก) ตัวอย่างภาพต้นแบบ (ข) ตัวอย่างภาพผลการทำงานของโครงข่าย ART

(ค) ตัวอย่างภาพผลการตรวจจับหน้าคนด้วยโครงข่าย ART

จากรูปที่ 3.8 (ข) พบว่าระบบจะทำการสร้างกล่องข้อมูลอินพุตที่มีค่าสีผิวมนุษย์มาก 1 ใน 4 ของพื้นที่กล่องทั้งหมดเป็นอินพุตของระบบ แสดงในกรอบภาพสีเหลือง และจุดภาพสีขาวแสดงถึงค่าจุดภาพนั้นมีความน่าจะเป็นที่จะเป็นสีผิวมนุษย์ โดยกล่องข้อมูลนี้อาจมีได้หลายอันขึ้นอยู่กับจำนวนจุดภาพที่เป็นสีผิวมนุษย์ในภาพนั้นมีค่ามากน้อยเพียงใด ต่อจากนั้นระบบจะทำการสร้างกล่องข้อมูลเพื่อเป็นอินพุตให้โครงข่าย แสดงในกรอบภาพสีแดงเพื่อหาจุดที่มีความน่าจะเป็นว่าเป็นหน้าคนภายในกล่องภาพแต่ละกล่อง แล้วแสดงผลการคัดแยกหน้าคน โดยระบบจะแสดงตัวอักษร “No” บนกล่องภาพที่ไม่ใช่หน้าคน และแสดงตัวเลขบนกล่องภาพที่เป็นหน้าคน โดยถ้าในภาพนั้นระบบตรวจพบหน้าคนจำนวน 3 คน ระบบจะแสดงตัวเลขบนกล่องที่เป็นหน้าคนเรียงกันไปจาก 1 ถึง 3

3.4 สรุป

โครงสร้างของโครงข่าย ART ให้ในการตรวจจับหน้าคนจากภาพสีนั้น ประกอบด้วยชั้นจดจำและชั้นเปรียบเทียบ โดยในชั้นจดจำนั้นได้ใช้ค่าสอดคล้องในการเปรียบเทียบ 2 ค่า โดยค่าที่ได้นั้นได้จากการสุ่มเลือกค่าที่เหมาะสมด้วย GA โดยกำหนดพารามิเตอร์เริ่มต้นของ GA ดังนี้ คือ กำหนดจำนวนรอบสูงสุดในการค้นหา 50 รอบ ค่าความผิดพลาดสูงสุดเป็น 100 ช่วงของข้อมูลที่ต้องการเลือกเป็นค่าเริ่มต้นอยู่ในช่วง 0 ถึง 1 ค่าน้ำหนักประสาทเริ่มต้นเป็นค่า 0 และ จำนวนประชากรมีค่า 40 โดยมีอินพุตของระบบเป็นไอเกนเวกเตอร์ที่ได้จากการวิเคราะห์องค์ประกอบ ของภาพตัวอย่างหน้าคนจำนวน 20 ภาพ และภาพตัวอย่างที่ไม่ใช่หน้าคนอีก 20 ภาพ ซึ่งการทดสอบระบบจะได้กล่าวถึงในบทต่อไป

บทที่ 4

ผลการทดสอบระบบตรวจจับหน้าคนด้วยโครงข่าย ART

4.1 กล่าวนำ

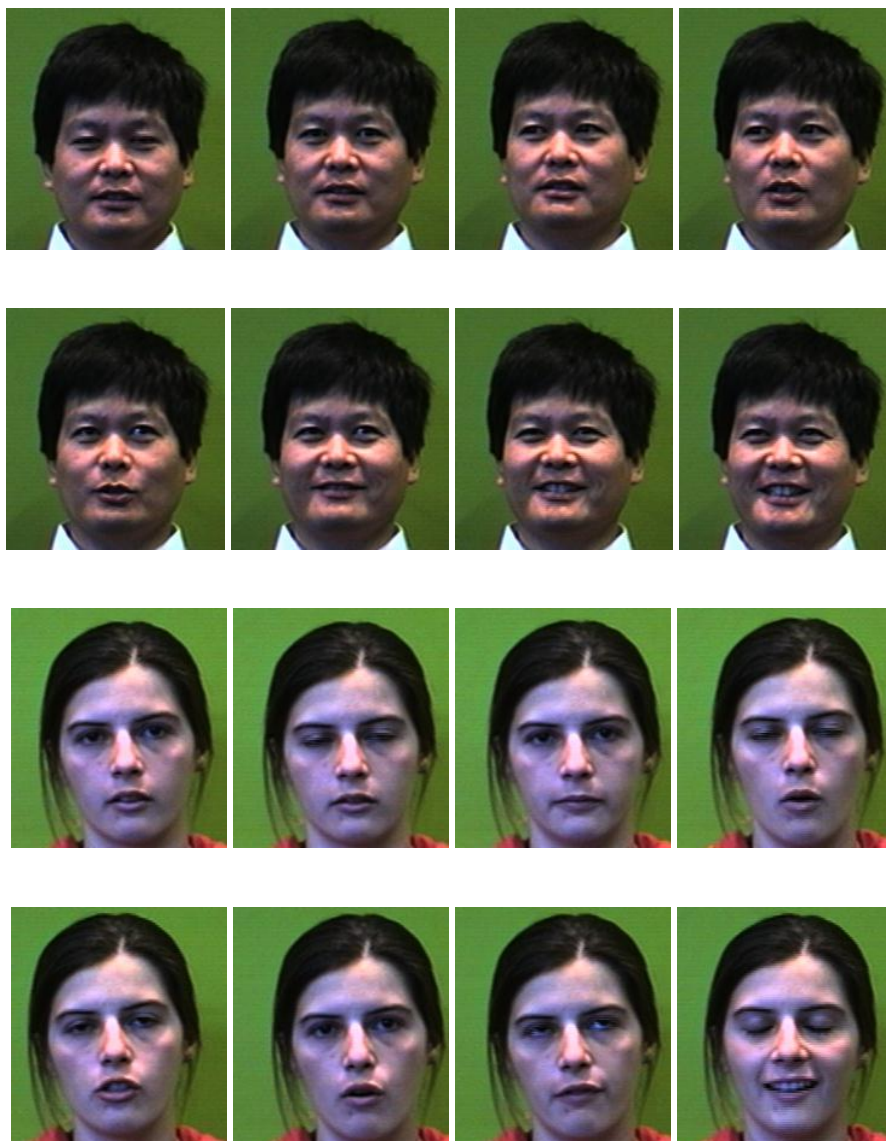
เมื่อทำการฝึกสอนโครงข่าย ART ด้วยภาพหน้าคน และภาพที่ไม่ใช่หน้าคนแล้ว ทำการทดสอบระบบด้วยภาพจากฐานข้อมูลต่าง ๆ เพื่อวัดประสิทธิภาพการทำงานของระบบว่ามีความสามารถในการตรวจจับหน้าคนจากภาพได้ดีเพียงไร โดยภาพที่ใช้ในการทดสอบนั้นได้จากฐานข้อมูลภาพต่าง ๆ และภาพที่ผู้วิจัยได้ทำการบันทึกเองจากกล้องดิจิทัลยี่ห้อต่าง ๆ ในสภาวะแวดล้อมต่าง ๆ เพื่อทดสอบว่าระบบสามารถลดข้อจำกัดของสิ่งแวดล้อมที่มีผลต่อการตรวจจับหน้าคนได้เป็นอย่างดี

4.2 ฐานข้อมูลภาพหน้ามนุษย์และภาพที่ใช้ในการทดสอบระบบ

ทำการรวบรวมข้อมูลภาพหน้ามนุษย์จากฐานข้อมูลต่าง ๆ ทางอินเทอร์เน็ตโดยฐานข้อมูลที่ใช้ นั้นถูกบันทึกภาพที่สภาวะแวดล้อมที่แตกต่างกัน ดังนั้นฐานข้อมูลแต่ละประเภทจึงใช้นำมาทดสอบระบบที่ปัจจัยต่าง ๆ แตกต่างกันไป ดังนี้

4.2.1 ฐานข้อมูลภาพ CVSR

นำภาพจากฐานข้อมูลภาพ CVSR (Computer Vision Science Research Face database of University of Essex, UK) แสดงในรูปที่ 4.1 มาใช้ในการทดสอบระบบเมื่อมีการฝึกสอนโครงข่ายในลักษณะต่าง ๆ เนื่องจากภาพจากฐานข้อมูลภาพ CVSR นั้นเป็นภาพที่มีการเปลี่ยนแปลงลักษณะการแสดงออกของหน้าที่มีความหลากหลาย และมีการจำกัดสีพื้นหลังของภาพให้มีสีเดียวกันและไม่มีความซับซ้อน ความสว่างในภาพมีความใกล้เคียงกัน ขนาดของหน้าคนในภาพมีขนาดเท่า ๆ กันและเป็นภาพเดี่ยวซึ่งเป็นปัจจัยต่าง ๆ ที่มีผลต่อประสิทธิภาพของระบบ ทำการทดลองนี้นั้น สามารถจำกัดอยู่ที่เพียงการทดสอบประสิทธิภาพในการฝึกสอนระบบเท่านั้น

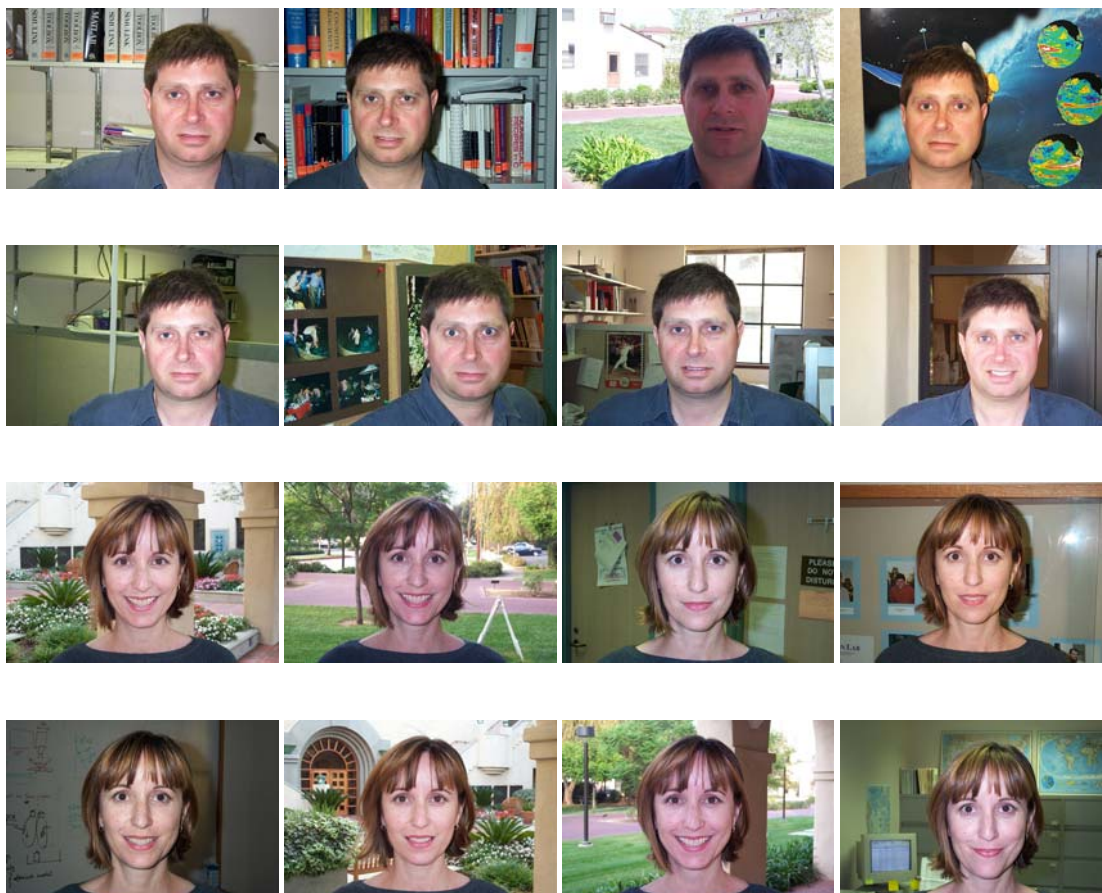


รูปที่ 4.1 ภาพตัวอย่างจากฐานข้อมูลภาพ CVSR

4.2.2 ฐานข้อมูลภาพ CMU VASC

นำภาพจากฐานข้อมูลภาพ CMU VASC (Carnegie Mellon University Vision and Autonomous System Center's Image Database) แสดงในรูปที่ 4.2 มาใช้ในการทดสอบระบบการตรวจจับหน้าคน เมื่อระบบมีการฝึกสอนโครงข่ายในลักษณะต่าง ๆ ทดสอบระบบเมื่อมีการเปลี่ยนแปลงค่าพารามิเตอร์ของโครงข่าย และทดสอบระบบเมื่อมีการเปลี่ยนแปลงขนาดของใบหน้าคนในภาพ โดยเหตุผลที่นำฐานข้อมูล CMU VASC มาใช้ในการทดสอบระบบเนื่องจากภาพจากฐานข้อมูลภาพ CMU VASC นั้นเป็นภาพที่มีปัจจัยภายในภาพ ที่มีความซับซ้อนและมีผลต่อประสิทธิภาพในการตรวจจับหน้าคน ไม่ว่าจะเป็นการเปลี่ยนแปลงลักษณะการแสดงออกของหน้าที่มีความหลากหลาย พื้นหลังของภาพมีความซับซ้อน ความสว่างในภาพที่ไม่สม่ำเสมอ ขนาดและตำแหน่งของคนในภาพมีความไม่แน่นอน เพื่

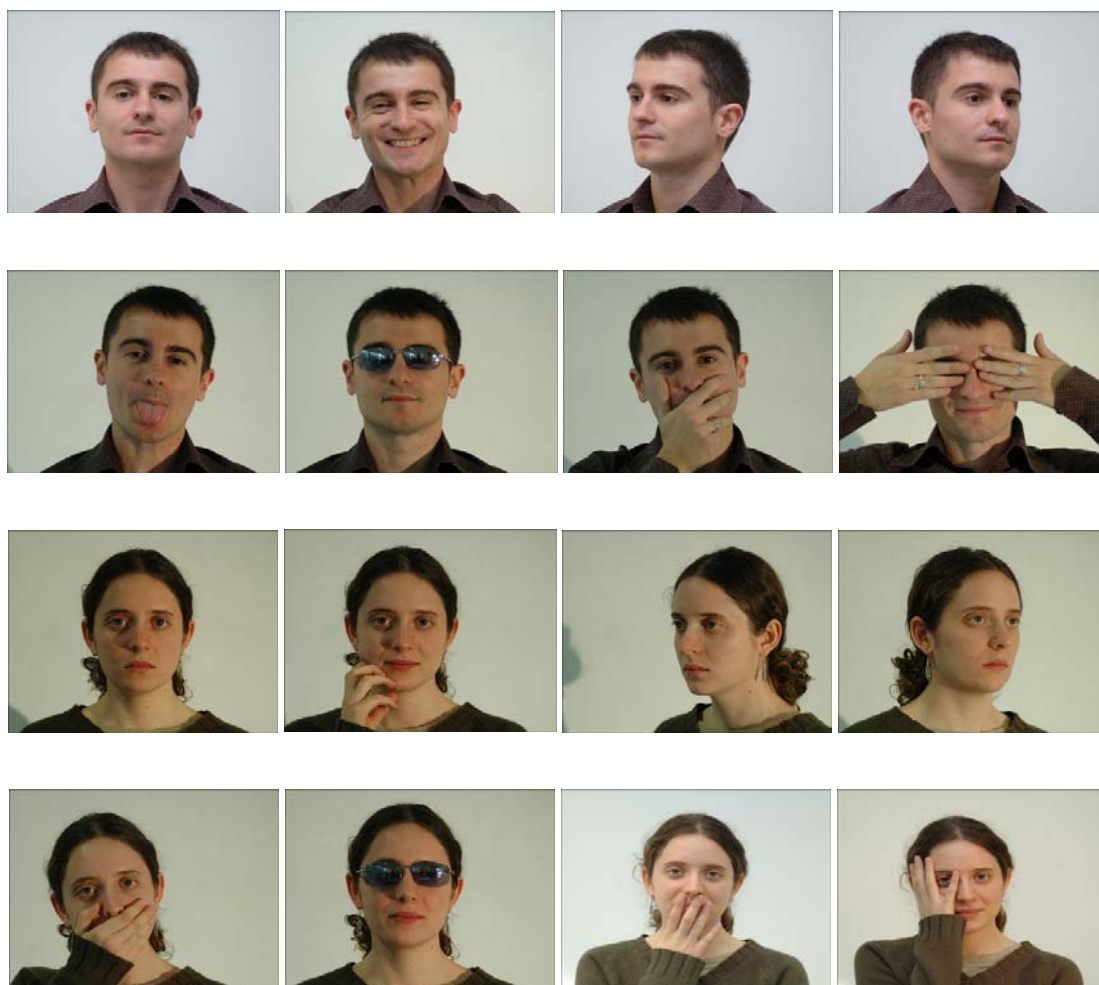
แสดงให้เห็นว่า ระบบมีความสามารถในการตรวจจับหน้าคนในภาพโดยที่ปัจจัยต่าง ๆ ที่กล่าวมานั้น มีผลน้อยมากต่อประสิทธิภาพของระบบ



รูปที่ 4.2 ภาพตัวอย่างจากฐานข้อมูลภาพ CMU VASC

4.2.3 ฐานข้อมูลภาพ AVTG

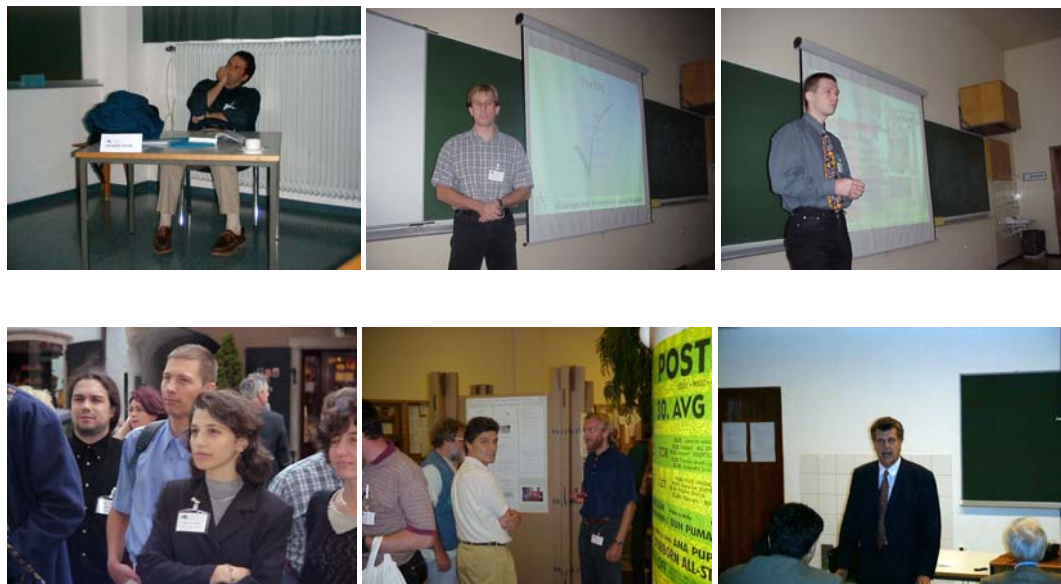
นำภาพจากฐานข้อมูลภาพ AVTG (Audio Visual Technologies Group of Technical University of Catalonia) แสดงในรูปที่ 4.3 มาใช้ในการทดสอบระบบการตรวจจับหน้าคน เมื่อลักษณะการวางท่าและการแสดงออกของหน้าคนในภาพมีการเปลี่ยนแปลง เนื่องจากภาพจากฐานข้อมูลภาพ AVTG นั้น เป็นภาพหน้าคน ที่มีความหลากหลายของลักษณะวางท่าและการแสดงออกของหน้าคน และมีการจำกัดสีพื้นหลังของภาพให้มีสีเดียวกันและไม่มีความซับซ้อน ความสว่างในภาพมีความใกล้เคียงกัน ขนาดของหน้าคนในภาพมีขนาดเท่า ๆ กันและเป็นภาพเดี่ยวซึ่งเป็นปัจจัยต่าง ๆ ที่มีผลต่อประสิทธิภาพของระบบการทดลองนี้นั้น สามารถจำกัดอยู่ที่เพียงการทดสอบประสิทธิภาพของระบบเมื่อลักษณะการวางท่าและการแสดงออกของหน้าคนในภาพมีการเปลี่ยนแปลงไปเท่านั้น



รูปที่ 4.3 ภาพตัวอย่างจากฐานข้อมูลภาพ AVTG

4.2.4 ฐานข้อมูลภาพ VALID และฐานข้อมูลภาพ CBCL

นำภาพจากฐานข้อมูลภาพ VALID (VALID Database of University college Dublin) แสดงในรูปที่ 4.4 และฐานข้อมูลภาพ CBCL (Center for Biological and Computational Learning Face database of Massachusetts Institute of Technology) แสดงในรูปที่ 4.5 มาใช้ในการทดสอบระบบการตรวจจับหน้าคน เมื่อจำนวนคนในภาพมีการเปลี่ยนแปลง เนื่องจากภาพจากฐานข้อมูลภาพ VALID และฐานข้อมูลภาพ CBCL นั้น เป็นภาพที่มีความหลากหลายและความไม่แน่นอนของปัจจัยที่มีผลต่อประสิทธิภาพของระบบตรวจจับหน้าคนในภาพ ได้แก่ จำนวนหน้าคนในภาพ ลักษณะวางท่าและการแสดงออกของหน้าคนในภาพ แสงและความสว่างในภาพ ขนาดและตำแหน่งของคนในภาพ พื้นหลังของภาพที่มีความซับซ้อน เป็นต้น เพื่อแสดงให้เห็นว่าระบบมีความสามารถในการตรวจจับหน้าคนในภาพภายใต้ปัจจัยต่าง ๆ ที่กล่าวมานั้น



รูปที่ 4.4 ภาพตัวอย่างจากฐานข้อมูลภาพ VALID



รูปที่ 4.5 ภาพตัวอย่างจากฐานข้อมูลภาพ CBCL

4.2.5 ภาพจากกล้องดิจิตอลฟูจิ F 601 และกล้องวิดีโอ

เป็นภาพที่ได้จากผู้วิจัยเป็นผู้บันทึกภาพเอง แสดงในรูปที่ 4.6 และ 4.7 ซึ่งขณะทำการบันทึกภาพนั้น ผู้วิจัยไม่ได้ทำการจำกัดสถานะแวดล้อมของภาพ อันได้แก่ ความสว่าง ตำแหน่งของหน้าคนในภาพ สีพื้นหลัง เพื่อให้ภาพที่ได้มีความหลากหลาย ในการทดสอบประสิทธิภาพของระบบ



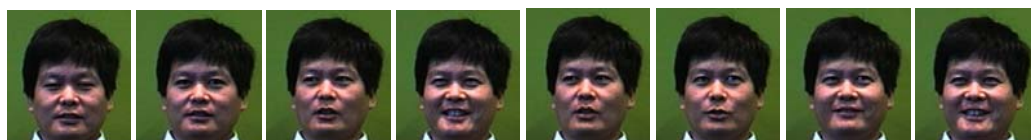
รูปที่ 4.6 ภาพตัวอย่างจากกล้องฟูจิ F601



รูปที่ 4.7 ภาพตัวอย่างจากกล้องวิดีโอ

4.3 ทดสอบระบบตรวจจับหน้าคนด้วยโครงข่าย ART

4.3.1 ทดสอบระบบเมื่อมีการฝึกสอนโครงข่ายในลักษณะต่าง ๆ โดยจะทำการทดสอบระบบเพื่อพิจารณาผลของการฝึกสอนระบบ ว่าจำนวนภาพในการฝึกสอนโครงข่ายมีผลต่อประสิทธิภาพของระบบอย่างไร โดยทำการทดสอบระบบด้วยภาพในการฝึกสอนระบบของคน ๆ เดียวกันที่มีลักษณะใบหน้าที่แตกต่างกันออกไป แสดงรูปตัวอย่างภาพที่ทำการทดลองในรูปที่ 4.8



รูปที่ 4.8 ภาพตัวอย่างในการฝึกสอนโครงข่าย

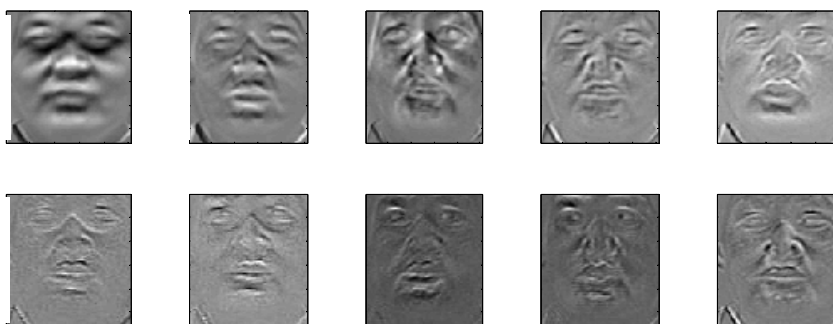
โดยจะทำการทดสอบระบบเมื่อไอเจนเวกเตอร์ที่เข้าสู่ระบบ ได้จากภาพตัวอย่างมีจำนวนภาพในการหาค่าไอเจนต่าง ๆ กันไป ดังแสดงค่าในตารางที่ 1 ซึ่งจำนวนภาพตัวอย่างที่ใช้หาค่าไอเจนนั้น มีผลต่อจำนวนมิติของข้อมูลไอเจนเวกเตอร์ที่เป็นข้อมูลในการฝึกสอนโครงข่าย โดยมิติของไอเจนเวกเตอร์ที่ใช้ในการฝึกสอน จะมีค่าเท่ากับจำนวนตัวอย่างในการฝึกสอน ดังนั้นเมื่อจำนวนภาพตัวอย่างในการฝึกสอนมากขึ้น จะเป็นผลให้ขนาดของข้อมูลไอเจนเวกเตอร์ในการฝึกสอนมากขึ้นด้วย โดยแสดงตัวอย่างภาพหน้าไอเจนของจำนวนภาพในการฝึกสอนในรูปที่ 4.9 และแสดงผลการทดลองในตารางที่ 1



(ก)



(ข)



(ค)

รูปที่ 4.9 ตัวอย่างภาพหน้าไอเจน (ก) ภาพหน้าไอเจนที่จำนวนภาพในการหาหน้าไอเจน 3 ภาพ

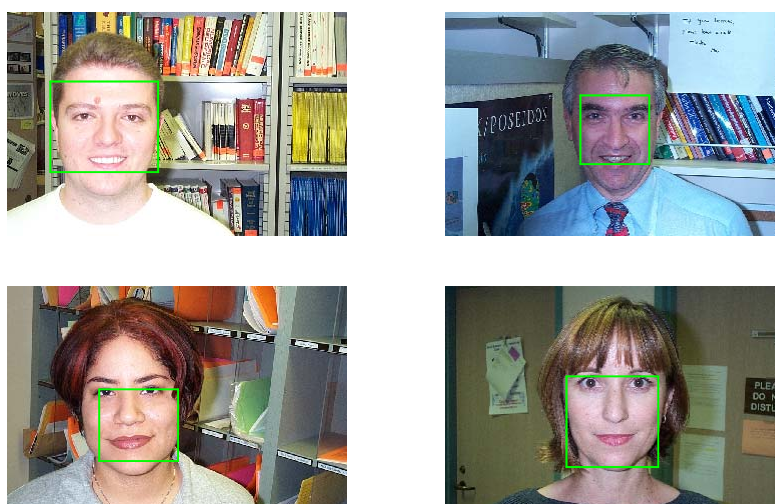
(ข) ภาพหน้าไอเกนที่จำนวนภาพในการหาหน้าไอเกน 5 ภาพ

(ค) ภาพหน้าไอเกนที่จำนวนภาพในการหาหน้าไอเกน 10 ภาพ

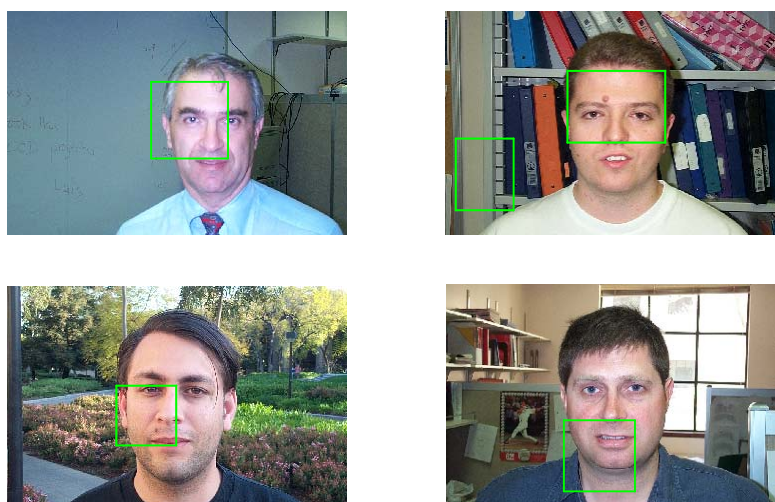
ตารางที่ 4.1 ทดสอบระบบที่จำนวนภาพในการฝึกสอนโครงข่ายต่าง ๆ

จำนวนภาพในการฝึกสอนโครงข่าย (ภาพ)	ความถูกต้อง (%)
3	95.45
5	97.25
10	97.17
15	97.48
20	98.86
30	98.42

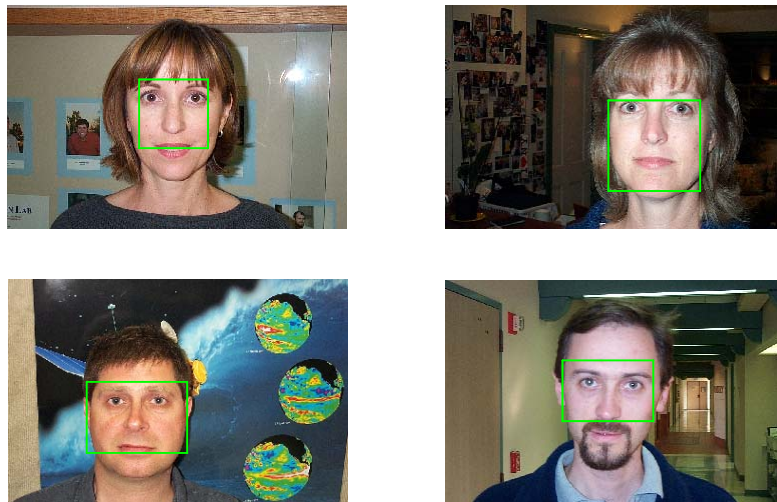
โดยทำการทดสอบความถูกต้องของระบบกับภาพในฐานข้อมูลภาพ CMU VASC จำนวน 100 ภาพ ได้ผลการทดลองดังตารางที่ 4.1 จากการผลการทดลองพบว่าเมื่อจำนวนภาพตัวอย่างในการฝึกสอนเพียง 3 ภาพก็เพียงพอที่จะทำให้ระบบมีประสิทธิภาพในการตรวจจับหน้าคนได้อยู่ในระดับที่ดี และเมื่อทำการเพิ่มจำนวนภาพในการฝึกสอนโครงข่ายมากขึ้น พบว่าโครงข่ายมีเปอร์เซ็นต์ความถูกต้องที่มากขึ้น จนถึงจำนวนภาพในการฝึกสอน 20 ภาพ เมื่อทำการเพิ่มจำนวนภาพในการฝึกสอนเป็น 30 ภาพ พบว่ามีเปอร์เซ็นต์ความถูกต้องน้อยลง ทั้งนี้อาจเนื่องมาจากการความซับซ้อนของมิติข้อมูลที่มากขึ้น อาจทำให้ระบบมีความสามารถในการคัดแยกความเป็นหน้าได้น้อยลง แต่อย่างไรก็ดีพบว่าผลการทดสอบระบบเมื่อจำนวนในการฝึกสอนโครงข่ายแตกต่างกันนั้น มีเปอร์เซ็นต์ความถูกต้องที่ไม่แตกต่างกันมากนัก ทั้งนี้เนื่องจากโครงข่าย ART นั้นเป็นโครงข่ายที่สามารถเรียนรู้ข้อมูลใหม่ได้โดยไม่ต้องทำการฝึกสอนซ้ำอีกโดยไม่ลืมข้อมูลเดิม แต่การฝึกสอนโครงข่ายด้วยจำนวนภาพตัวอย่างในปริมาณน้อยนั้น ก็อาจไม่เพียงพอที่จะทำให้ระบบมีประสิทธิภาพสูงได้ และผลจากค่าไอเกนเวกเตอร์ที่ดีที่จะเป็นข้อมูลอินพุตให้โครงข่าย ควรมีจำนวนมิติที่มากเพียงพอ จึงจะสามารถแสดงลักษณะของภาพได้ดีกว่าค่าไอเกนเวกเตอร์ที่มีขนาดมิติน้อย ๆ แสดงผลการทดสอบระบบในรูปแบบที่ 4.10



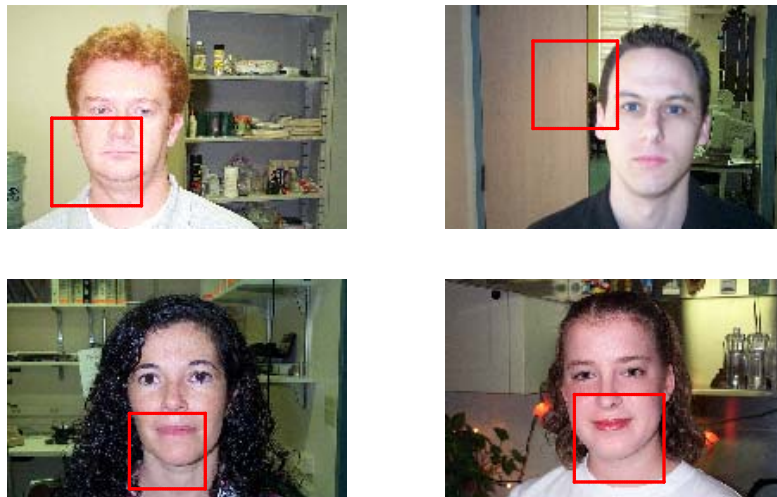
(ก) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 3 ภาพ



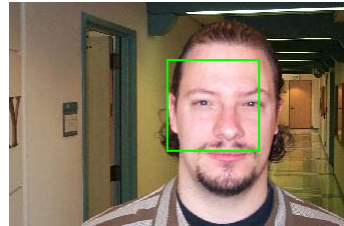
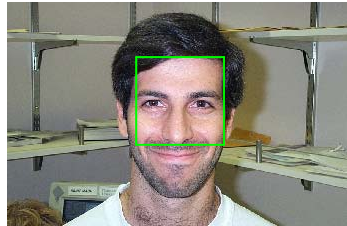
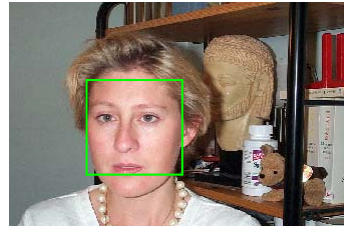
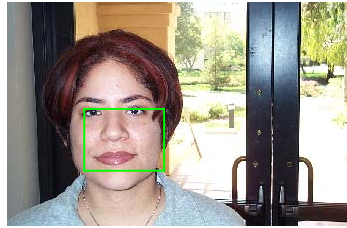
(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 3 ภาพ



(ค) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 5 ภาพ



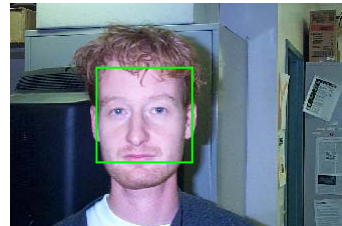
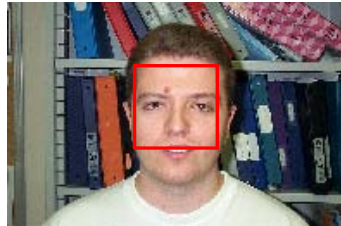
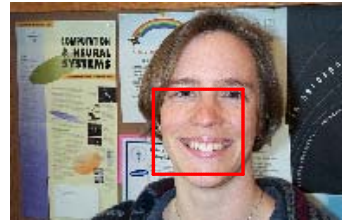
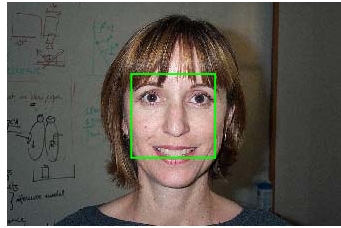
(ง) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 5 ภาพ



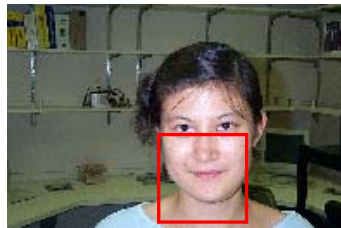
(จ) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 10 ภาพ



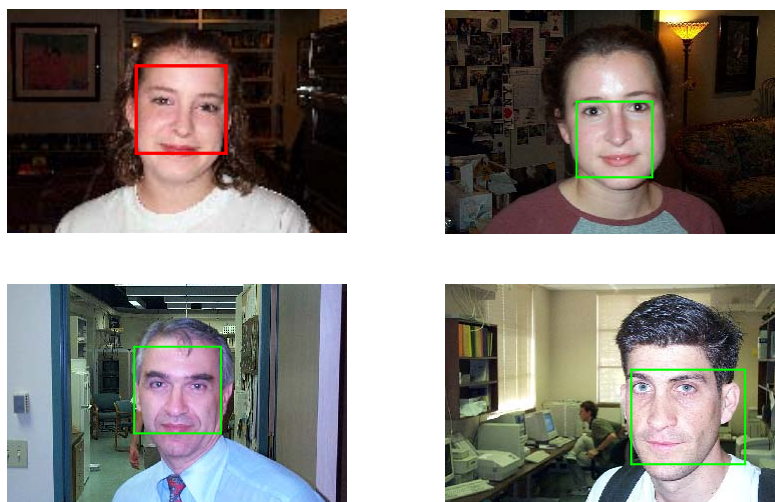
(ฉ) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 10 ภาพ



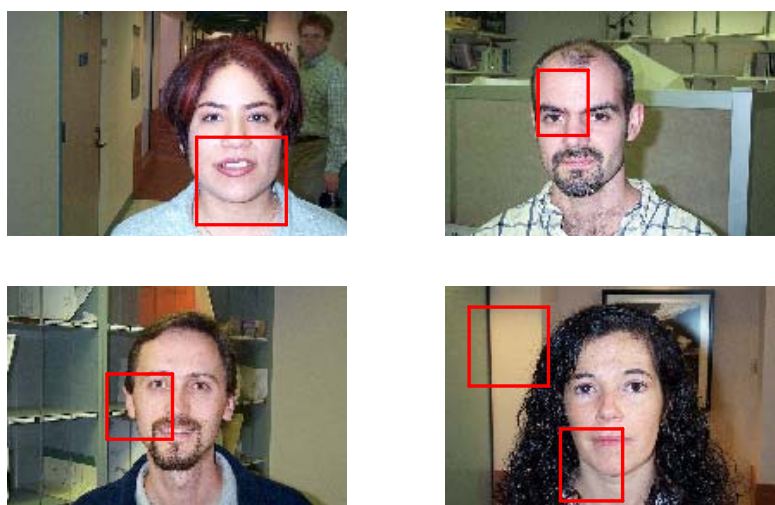
(ข) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 20 ภาพ



(ค) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 20 ภาพ



(ณ) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 30 ภาพ



(ญ) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 30 ภาพ

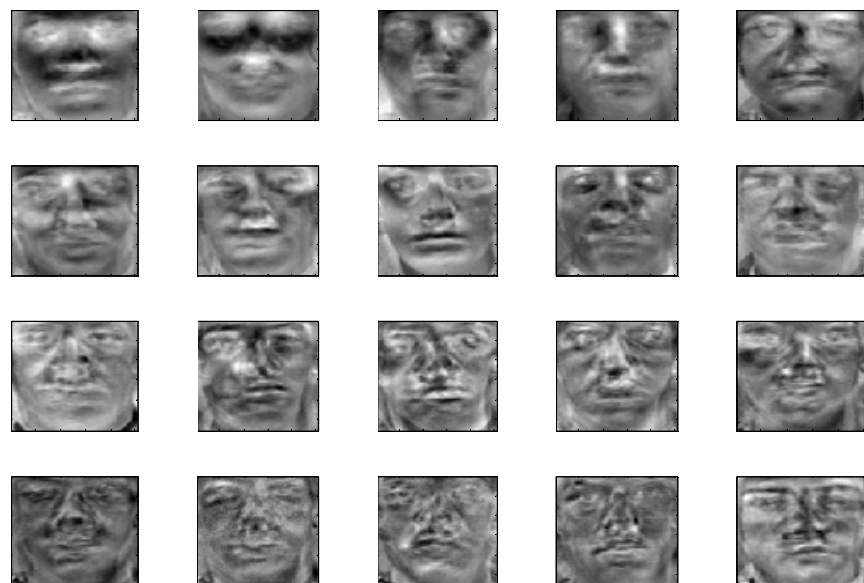
รูปที่ 4.10 ตัวอย่างภาพผลการทดสอบระบบ ที่จำนวนตัวอย่างในการฝึกสอนต่าง ๆ

เมื่อทราบผลของการทดสอบจำนวนภาพในการฝึกสอนโครงข่ายแล้วว่ามีผลต่อประสิทธิภาพของระบบ ดังนั้นในการทดลองต่อมาเราจึงทำการเลือกใช้จำนวนภาพในการฝึกสอน 20 ภาพ เนื่องจากการทดลองข้างต้นได้แสดงให้เห็นแล้วว่า ทำให้ระบบมีประสิทธิภาพดีที่สุด โดยจะพิจารณาคุณสมบัติของภาพที่ใช้ในการฝึกสอนว่ามีผลต่อระบบอย่างไร ด้วยการแบ่งภาพในการฝึกสอนออกเป็น 3 กลุ่มหลัก ๆ คือ กลุ่มแรกทำการฝึกสอนด้วยภาพของบุคคลคนเดียวกันในลักษณะใบหน้า และการวางท่าที่แตกต่างกัน กลุ่มที่ 2 ทำการฝึกสอนด้วยภาพของคนหลาย ๆ คนในลักษณะหน้าตรงเหมือนกันหมด และกลุ่มที่ 3

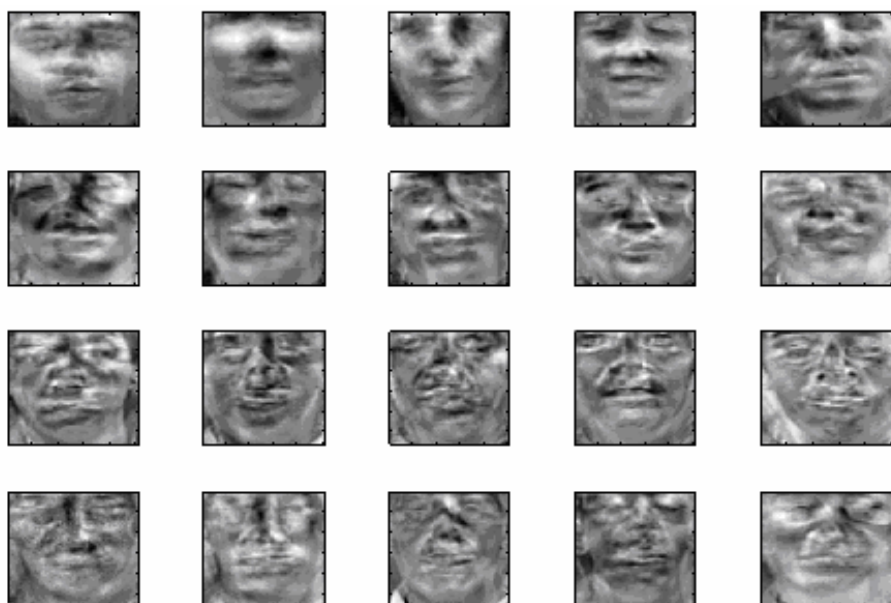
ทำการฝึกสอนด้วยภาพของคนหลาย ๆ คนในลักษณะใบหน้าและการวางท่าที่แตกต่างกัน โดยแสดงตัวอย่างภาพที่ใช้ในการฝึกสอนโครงข่ายในตารางที่ 4.2 และแสดงภาพหน้าไอเกนของตัวอย่างแต่ละแบบในรูปที่ 4.11 ได้ผลการทดลองแสดงในตารางที่ 4.2



(ก) ภาพหน้าไอเกนของภาพบุคคลเดียวกัน ที่ลักษณะใบหน้าต่างกัน



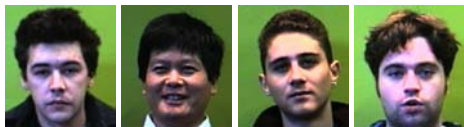
(ข) ภาพหน้าไอเกนของภาพบุคคลต่างกัน ที่ลักษณะใบหน้าตรง



(ค) ภาพหน้าไอเจนของภาพบุคคลต่างกัน ที่ลักษณะใบหน้าต่างกัน

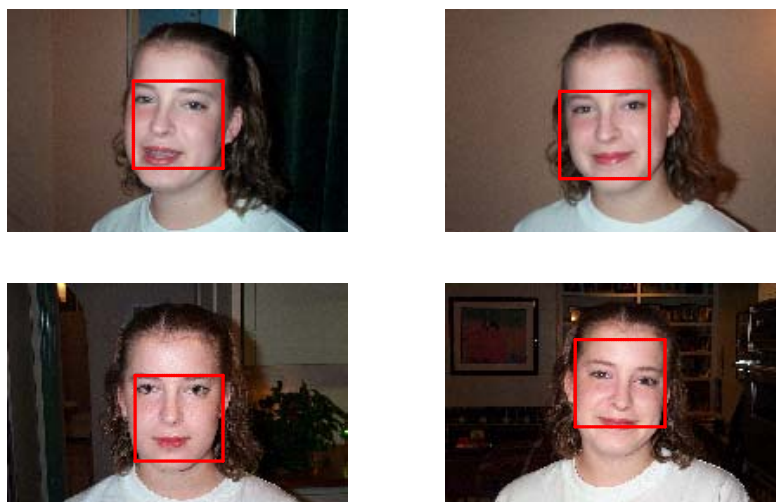
รูปที่ 4.11 ตัวอย่างภาพหน้าไอเจน

ตารางที่ 4.2 ทดสอบระบบด้วยภาพที่มีคุณลักษณะต่าง ๆ

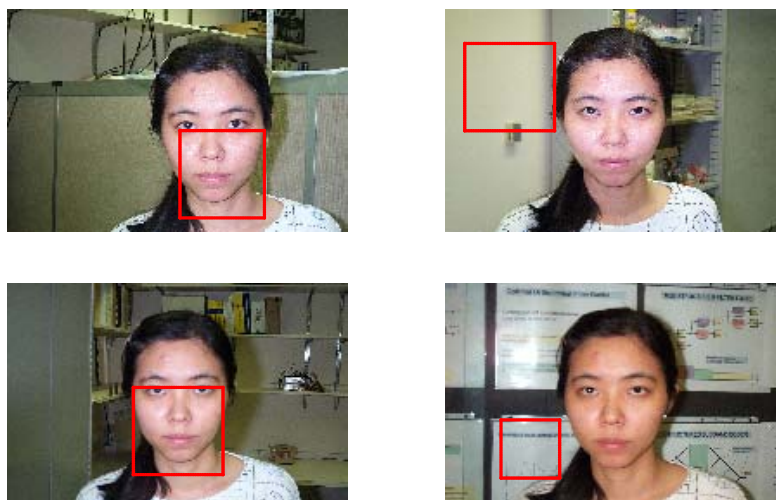
ภาพในการฝึกสอนโครงข่าย	ความถูกต้อง (%)
ภาพบุคคลเดียวกัน ที่ลักษณะใบหน้าต่าง ๆ 	98.86
ภาพบุคคลต่างกัน ที่ลักษณะใบหน้าตรง 	97.48
ภาพบุคคลต่างกัน ที่ลักษณะใบหน้าต่าง ๆ 	96.68

จากการทดลองพบว่าการฝึกสอนด้วยภาพบุคคลคนเดียวกันในลักษณะต่าง ๆ ให้เปอร์เซ็นต์ความถูกต้องที่ดีที่สุด และเมื่อพิจารณาจากภาพหน้าไอเจนที่ได้พบว่า ภาพหน้าไอเจนของข้อมูลที่เป็นบุคคลเดียวกันนั้นจะแสดงลักษณะเค้าโครงของหน้าได้ชัดเจนที่สุดเช่นเดียวกัน โดยค่าไอเจนที่ชัดเจนที่สุดนั้น

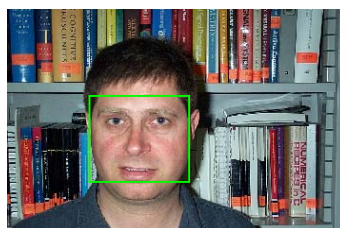
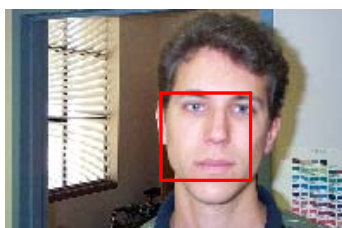
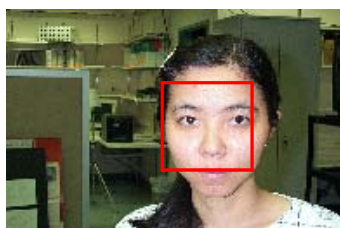
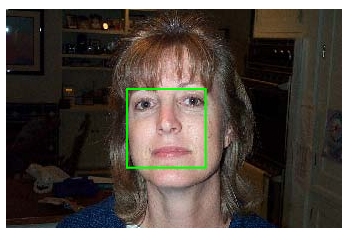
จะเป็นค่าไอเจนที่เกิดจากค่าไอเจนค่าแรก ในรูปที่ 4.12 แสดงภาพตัวอย่างในการทดสอบระบบด้วยภาพที่มีคุณลักษณะต่าง ๆ



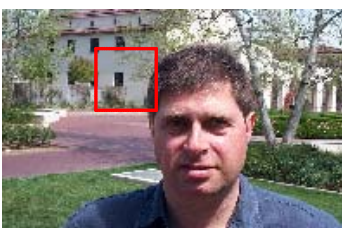
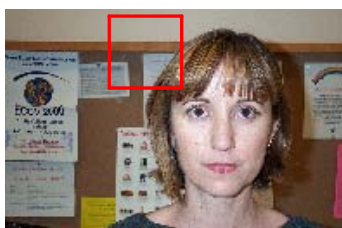
(ก) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้องของตัวอย่างในการฝึกสอนโครงข่ายของบุคคลเดียวกันที่ลักษณะใบหน้าต่าง ๆ



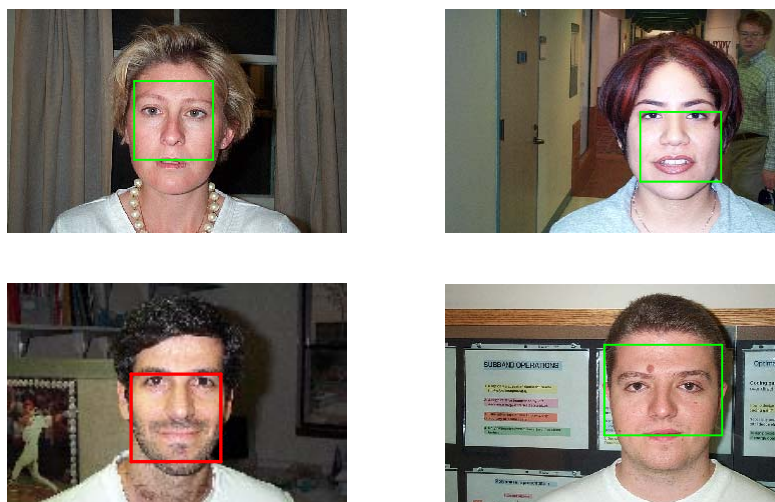
(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้องของตัวอย่างในการฝึกสอนโครงข่ายของบุคคลเดียวกันที่ลักษณะใบหน้าต่าง ๆ



(ค) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้องของตัวอย่างในการฝึกสอนโครงข่าย
ของบุคคลต่างกัน ที่ลักษณะหน้าตรง



(ง) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้องของตัวอย่างในการฝึกสอนโครงข่าย
ของบุคคลต่างกัน ที่ลักษณะหน้าตรง



(จ) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้องของตัวอย่างในการฝึกสอนโครงข่ายของบุคคลต่างกัน ที่ลักษณะใบหน้าต่าง ๆ



(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้องของตัวอย่างในการฝึกสอนโครงข่ายของบุคคลต่างกัน ที่ลักษณะใบหน้าต่าง ๆ

รูปที่ 4.12 ตัวอย่างภาพผลการทดสอบระบบ ของภาพตัวอย่างในการฝึกสอนที่คุณลักษณะต่าง ๆ

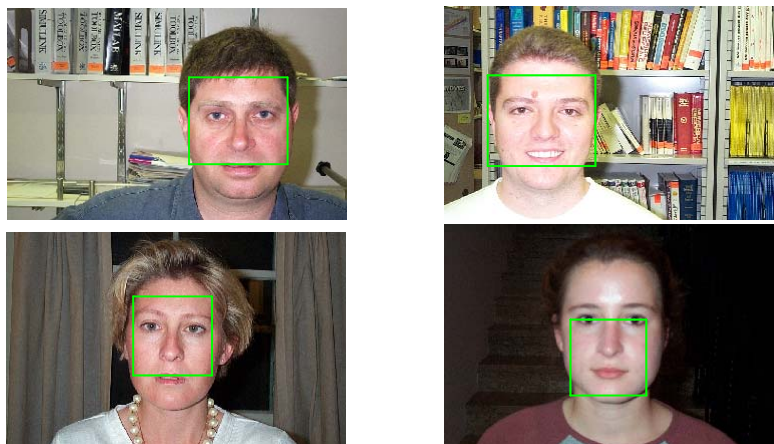
อย่างไรก็ดีจากผลการทดลองแสดงให้เห็นว่า คุณลักษณะของภาพที่ทำการฝึกสอนต่าง ๆ กันนั้น ให้ผลการตรวจจับหน้าคนได้มีประสิทธิภาพดีไม่แตกต่างกันมากนัก ดังนั้นปัจจัยในการเลือกใช้ข้อมูลในการฝึกสอนโครงข่าย จึงอาจไม่จำเป็นต้องพิจารณาถึงความหลากหลายของข้อมูลฝึกสอน ทั้งนี้ เนื่องจากโครงข่าย ART นั้นมีการพัฒนาปรับตัวให้สามารถรับข้อมูลที่หลากหลายได้ในอนาคต

4.3.2 ทดสอบระบบที่พารามิเตอร์ของโครงข่ายต่าง ๆ โดยทำการทดสอบโครงข่ายเมื่อทำการเปลี่ยนแปลงค่าสอดส่อง (ρ) ซึ่งเป็นพารามิเตอร์ที่กำหนดขีดความสามารถในการแบ่งกลุ่มข้อมูลของโครงข่าย ART โดย ρ_1 เป็นค่าขีดความสามารถในการแบ่งกลุ่มของความเป็นหน้า และความเป็นหน้าออกจากกัน ส่วน ρ_2 เป็นค่าขีดความสามารถในการแบ่งกลุ่มความเป็นหน้าออกเป็นกลุ่มต่าง ๆ โดยเมื่อทำการเปลี่ยนค่าพารามิเตอร์แล้วได้ผลการทดลองดังแสดงในตารางที่ 4.3

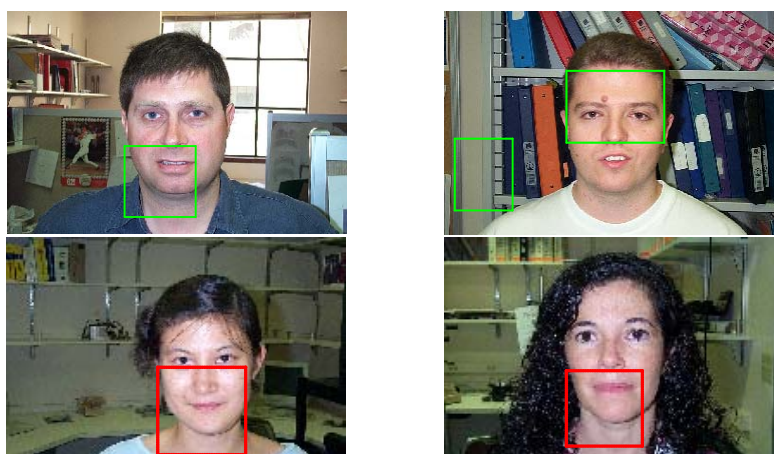
ตารางที่ 4 ความถูกต้องในการคัดแยกหน้าคนโดยค่าสอดส่องค่าต่าง ๆ

ค่าสอดส่อง		ความถูกต้องในการคัดแยก (%)	
ρ_1	ρ_2	ภาพใบหน้าบุคคล	ไม่ใช่ภาพใบหน้าบุคคล
0.554	0.567	80.06	72.56
	0.802	82.74	71.98
0.644	0.678	85.56	86.67
	0.802	88.59	96.57
0.786	0.802	91.76	66.75
	0.946	82.88	87.73
0.851	0.802	98.86	94.67
	0.946	83.21	86.65

พบว่า ระบบที่กำหนดค่าสอดส่อง ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802 ให้ผลการทดสอบระบบที่มีประสิทธิภาพดีที่สุด คือให้เปอร์เซ็นต์ความถูกต้องในการตรวจจับได้ดีทั้งในส่วนภาพที่เป็นหน้าคน และภาพที่ไม่ใช่หน้าคน โดยแสดงผลการทดสอบระบบในรูปที่ 4.13



(ก) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้องของระบบที่กำหนดค่าสอดคล้อง
 ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802



(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้องของระบบที่กำหนดค่าสอดคล้อง
 ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802

รูปที่ 4.13 ตัวอย่างภาพการทดสอบระบบของระบบที่กำหนดค่าสอดคล้อง
 ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802

4.3.3 ทดสอบระบบที่ขนาดของใบหน้าในภาพขนาดต่าง ๆ โดยจะทำการทดสอบโครงข่าย ART เมื่อทำการทดสอบกับภาพเดี่ยวที่มีหน้าคนในภาพขนาดประมาณ 70%, 55% และ น้อยกว่า 50% ของขนาดภาพทั้งหมดแสดงตัวอย่างภาพที่ใช้ทำการทดสอบในรูปที่ 4.14 และแสดงผลการทดสอบระบบในตารางที่ 4.4



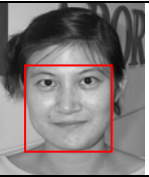
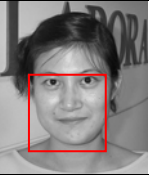
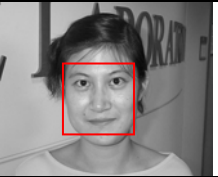
(ก)

(ข)

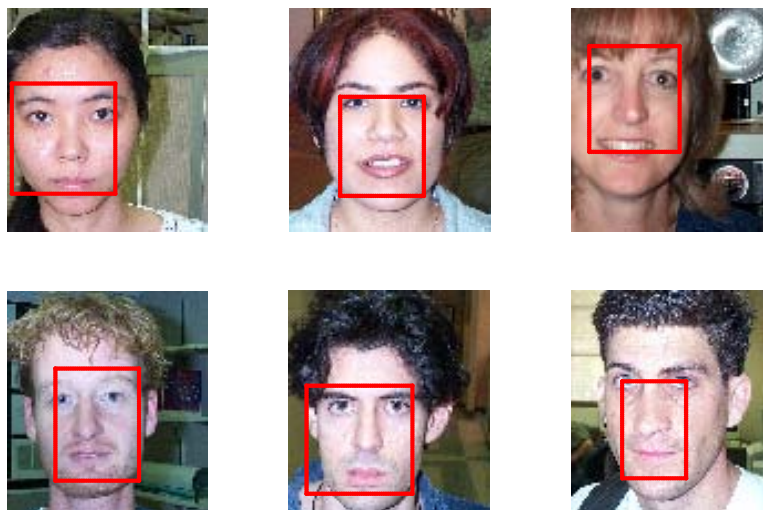
(ค)

รูปที่ 4.14 ตัวอย่างภาพที่ขนาดหน้าคนในภาพต่าง ๆ (ก) หน้าคนขนาด 70%
(ข) หน้าคนขนาด 55% (ค) หน้าคนขนาดน้อยกว่า 50%

ตารางที่ 4.4 ความถูกต้องในการคัดแยกหน้าคนเมื่อขนาดของใบหน้าในภาพเปลี่ยนแปลง

ภาพตัวอย่างในการทดสอบ	ขนาดของหน้าคนในภาพ (%)	ความถูกต้อง (%)	ภาพตัวอย่างผลการทดสอบ
	70	98.78	
	55	97.46	
	น้อยกว่า 50	88.72	

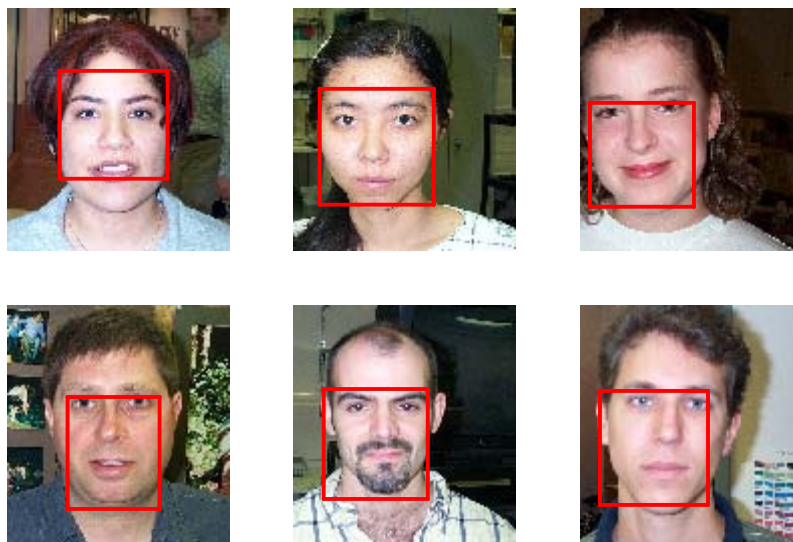
พบว่าระบบสามารถแยกภาพที่มีหน้าคนในภาพที่มีขนาด 70% ของขนาดภาพทั้งหมดได้ถูกต้องมากกว่า ภาพที่มีหน้าคนในภาพขนาดอื่น ๆ เนื่องจากภาพที่มีขนาดของใบหน้าใหญ่ จะทำให้เห็นลักษณะและรายละเอียดของภาพได้ชัดเจน ทำให้ระบบสามารถจดจำลักษณะเฉพาะได้ดี การตรวจจับจึงทำได้ถูกต้องมากกว่าภาพใบหน้าคนขนาดเล็ก โดยเมื่อทำการทดลองต่อไปพบว่าภาพที่มีขนาดเล็กมาก ๆ ระบบจะมองว่าไม่ใช่หน้าคน เนื่องจากลักษณะเฉพาะของใบหน้าไม่เด่นชัดและมีขนาดของสีผิวมนุษย์น้อยมากเมื่อเทียบกับภาพใบหน้าขนาดใหญ่ โดยภาพตัวอย่างผลการทดสอบระบบ แสดงในรูปที่ 4.15



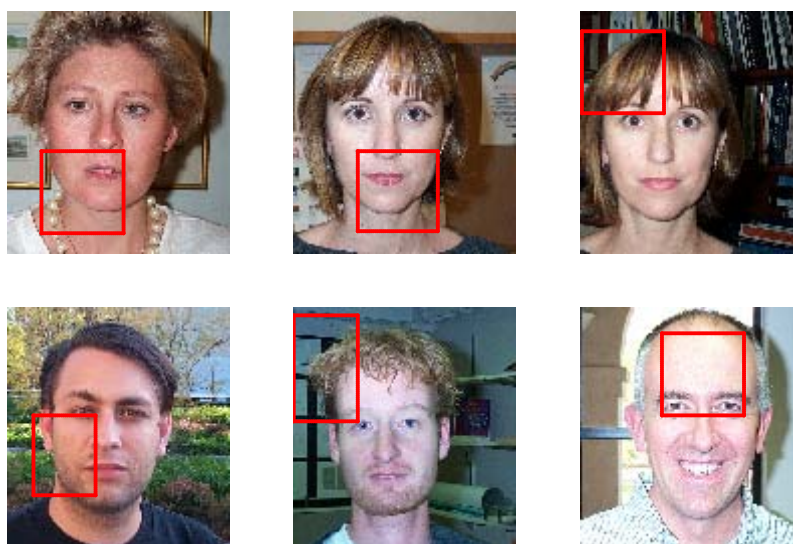
(ก) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพที่ขนาดใบหน้าคนในภาพ มีขนาด 70% ของขนาดภาพทั้งหมด



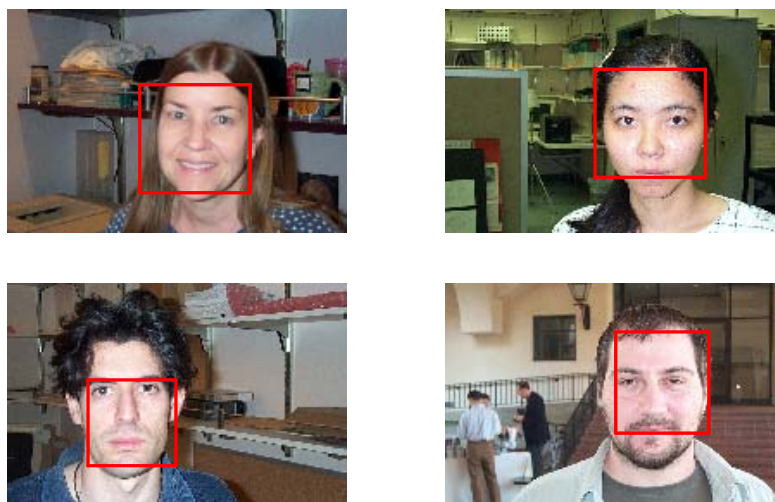
(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพที่ขนาดใบหน้าคนในภาพ มีขนาด 70% ของขนาดภาพทั้งหมด



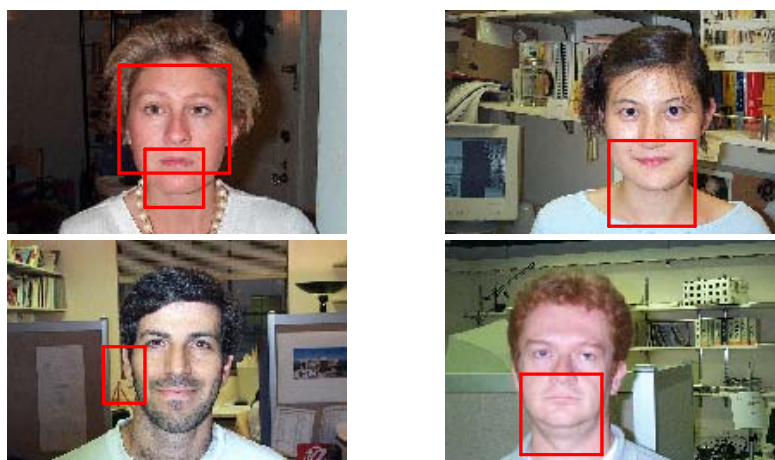
(ค) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพที่ขนาดใบหน้าคนในภาพ มีขนาด 55% ของขนาดภาพทั้งหมด



(ง) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพที่ขนาดใบหน้าคนในภาพ มีขนาด 55% ของขนาดภาพทั้งหมด



(จ) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพที่ขนาดใบหน้าคนในภาพ มีขนาดน้อยกว่า 50% ของขนาดภาพทั้งหมด



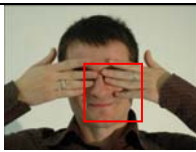
(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพที่ขนาดใบหน้าคนในภาพ มีขนาดน้อยกว่า 50% ของขนาดภาพทั้งหมด

รูปที่ 4.15 ตัวอย่างภาพผลการทดสอบระบบ ของใบหน้าคนในภาพขนาดต่าง ๆ

4.3.4 ทดสอบระบบที่หน้าในลักษณะต่าง ๆ การวางท่าของหน้าคนในภาพมีผลทำให้ระบบทำงานได้ต่างกัน กล่าวคือทำให้ความชัดเจนของหน้าในภาพจะต่างกัน โดยภาพหน้าตรงเป็นภาพที่มีความชัดเจนของลักษณะเฉพาะของหน้าที่ดีที่สุด และเมื่อหน้ามีสิ่งมาปิดบังหรือขวางกั้นส่วนใดส่วนหนึ่งบนใบหน้า ความชัดเจนของหน้านั้นจะน้อยลง เช่น ใบหน้าเอียงอาจจะให้เห็นลักษณะของดวงตา และโครง

หน้าไม่ชัดเจน หรือหน้าที่มีมือมาบังส่วนใดส่วนหนึ่งไว้ ทำให้ลักษณะเฉพาะของหน้าในส่วนนั้นหายไป ซึ่งจะได้ผลการทดสอบระบบกับภาพหน้าคนในลักษณะต่าง ๆ แสดงในตารางการทดสอบที่ 4.5

ตารางที่ 4.5 ความถูกต้องในการคัดแยกหน้าคนในลักษณะต่าง ๆ

ภาพตัวอย่างในการทดสอบ	ลักษณะการวางท่าของหน้าในภาพ	ความถูกต้องในการคัดแยก (%)	ภาพตัวอย่างผลการทดสอบ
	หน้าตรง	97.66	
	หน้าเอียงซ้าย	95.48	
	หน้าเอียงขวา	93.51	
	ใส่แว่นดำ	83.41	
	ปิดปาก	77.78	
	ปิดตา	70.46	

จากการทดลองพบว่าในภาพหน้าคนที่นำมือปิดตานั้น มีเปอร์เซ็นต์ความถูกต้องที่น้อยที่สุด ทั้งนี้เนื่องมาจากมีสิ่งขีดขวางมาปกปิดส่วนตาของใบหน้า เป็นผลให้ลักษณะในส่วนนั้นหายไป และมีดที่ปิดบังส่วนของใบหน้านั้นยังเป็นสีผิวมนุษย์ ทำให้ขอบเขตในการค้นหากว้างมากขึ้น เป็นผลให้มีพื้นที่ความผิดพลาดในการค้นหามากขึ้นด้วย และในส่วนของหน้าที่ปิดปาก ก็เช่นเดียวกัน ส่วนภาพคนใส่แว่นตานั้น ได้ถูกปิดบังใบหน้าเช่นเดียวกัน แต่ส่งผลต่อความถูกต้องของระบบน้อยกว่าการปิดบังด้วยมือ ทั้งนี้อาจเนื่องมาจากสีของแว่นตานั้นไม่ได้มีสีเดียวกับสีผิวมนุษย์ และโดยเป็นการปิดบังในบริเวณที่ไม่กว้าง ส่วนภาพที่ให้ผลการทดสอบระบบมีเปอร์เซ็นต์ความถูกต้องที่ดีที่สุดนั้น คือภาพหน้าตรง รองลงมาคือภาพหน้าเอียงซ้ายและหน้าเอียงขวา ทั้งนี้ก็เนื่องจากภาพหน้าตรงนั้นแสดงรายละเอียดของ

ลักษณะเฉพาะชัดเจนที่สุด และเหมือนภาพตัวอย่างในการฝึกสอนมากที่สุด ส่วนภาพหน้าเอียงซ้าย และภาพหน้าเอียงขวา ให้เปอร์เซ็นต์ความถูกต้องที่ใกล้เคียงกัน เนื่องจากมีลักษณะใบหน้าใกล้เคียงกัน แต่กลับด้านซ้าย ขวากันเท่านั้น แต่โดยภาพรวมนั้น ระบบสามารถคัดแยกความเป็นหน้าคนได้ถูกต้องอยู่ในเกณฑ์ที่ดี ไม่เท่ากับหน้าในลักษณะใด ๆ ก็ตาม โดยรูปที่ 4.16 แสดงผลการทดสอบระบบของหน้าคนในลักษณะต่าง ๆ



(ก) ภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพหน้าตรง



(ข) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าตรง



(ค) ภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพหน้าเอียงซ้าย



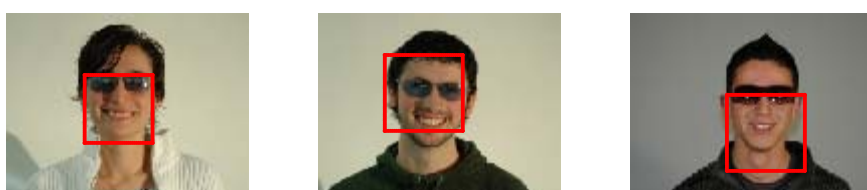
(ง) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าเอียงซ้าย



(จ) ภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพหน้าเอียงขวา



(ฉ) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าเอียงขวา



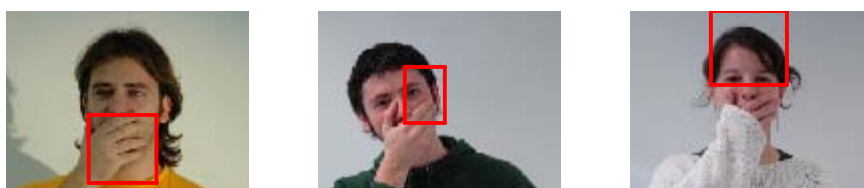
(ช) ภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพหน้าใส่แว่นดำ



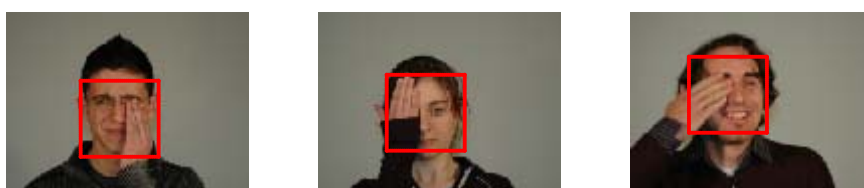
(ซ) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าใส่แว่นดำ



(ณ) ภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพหน้าปิดปาก



(ญ) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าปิดปาก



(ฎ) ภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพหน้าปิดตา

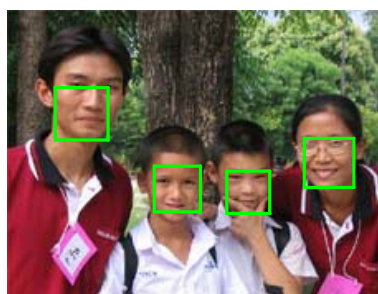
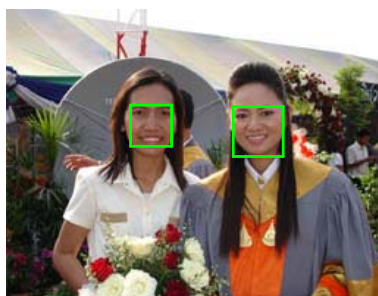
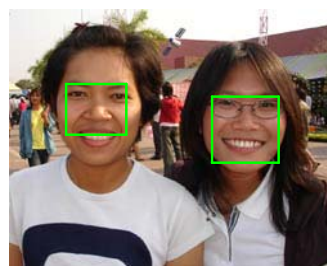
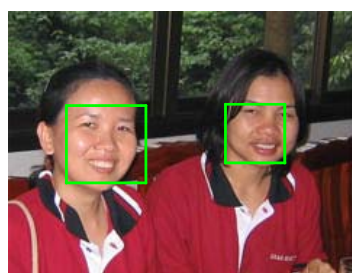
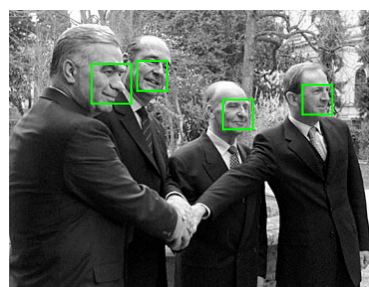
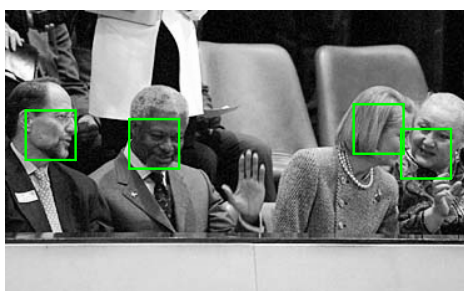
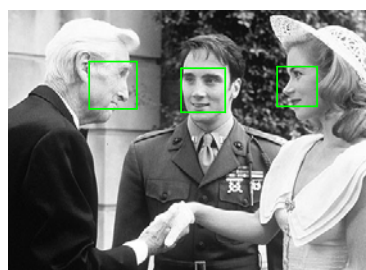
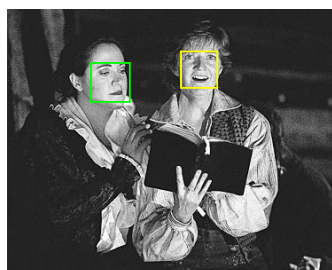
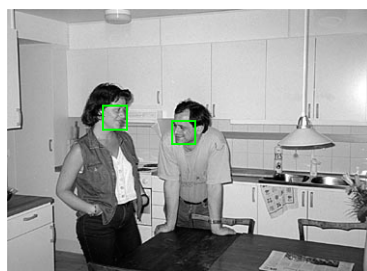
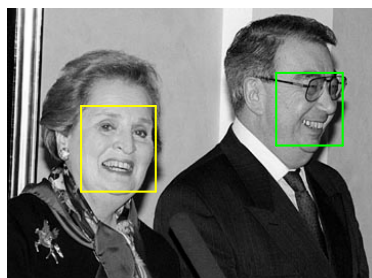


(ฏ) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าปิดตา

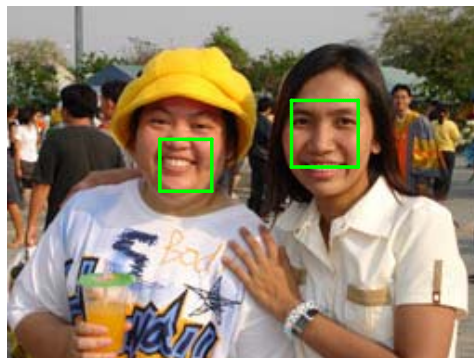
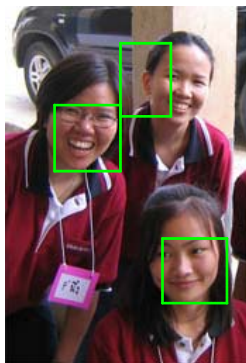
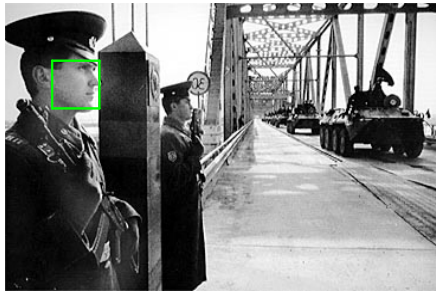
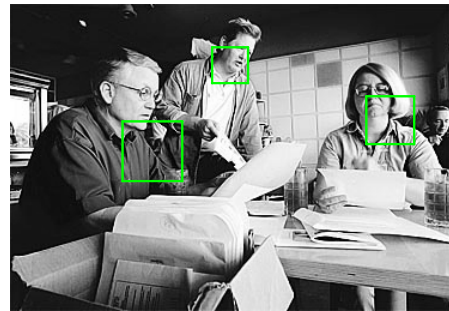
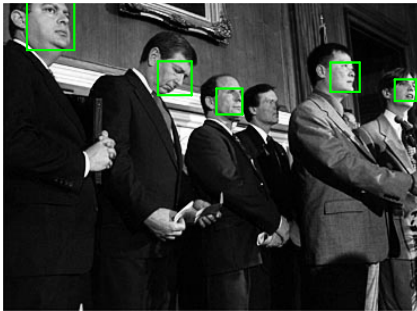
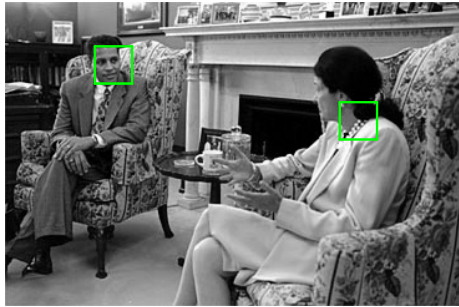
รูปที่ 4.16 ตัวอย่างภาพผลการทดสอบระบบ ของภาพตัวอย่างหน้าในลักษณะต่าง ๆ

4.3.5 ทดสอบระบบกับภาพที่มีหน้าคนในภาพมากกว่า 1 คน โดยบุคคลในภาพนั้นอาจมีลักษณะการวางท่า ลักษณะใบหน้า ขนาดของหน้าในภาพแตกต่างกันออกไป หรือถูกปิดบังลักษณะของหน้าในบางส่วนได้เช่นกัน ทั้งนี้เนื่องจากธรรมชาติของมนุษย์นั้นไม่ได้มีลักษณะการวางท่า หรือการ

แสดงออกทางสีหน้าที่เป็นรูปแบบตายตัว ดังนั้นระบบการคัดแยกหน้าที่ดีจึงต้องสามารถตรวจจับหน้าคนได้ในทุกลักษณะ และต้องสามารถหาหน้าคนในภาพได้อย่างครบถ้วนตามจำนวนหน้าจริงในภาพ โดยแสดงผลการทดลองในรูปแบบที่ 4.17



(ก) ภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพหน้าคนมากกว่า 1 หน้าในภาพ



(ข) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าคนมากกว่า 1 หน้าในภาพ

รูปที่ 4.17 ตัวอย่างภาพผลการทดสอบระบบ ของภาพที่มีจำนวนหน้าคนในภาพมากกว่า 1 หน้าในภาพ

จากผลการทดลองพบว่าระบบสามารถตรวจจับหน้าคนในภาพได้ถูกต้อง 97.84% และยังสามารถตรวจจับหน้าคนในภาพได้มากกว่า 1 หน้าในภาพ ไม่ว่าหน้าคนในภาพนั้นจะมีลักษณะใบหน้า หรือการวางท่าของคนในภาพที่แตกต่างกันออกไป หรือมีส่วนใดส่วนหนึ่งของใบหน้าที่ถูกปิดบังก็ตาม แต่ก็พบว่าระบบยังคงมีข้อผิดพลาดในการตรวจจับ กล่าวคือเกิดการตรวจจับไม่ครบถ้วน คือมีใบหน้าบางใบหน้าที่ระบบไม่สามารถระบุความเป็นหน้าได้ และเกิดการตรวจจับเกิน คือทำการระบุความเป็นหน้าคนในส่วนที่ไม่มีหน้าคนปรากฏอยู่ ซึ่งความผิดพลาดนี้อาจเนื่องมาจากหน้าคนในภาพนั้นมีขนาดเล็กเกินกว่าจะแสดงลักษณะของหน้าอย่างชัดเจน หรือส่วนมือที่มีความเป็นสีผิวมนุษย์อาจมีค่าไอเกนเวกเตอร์ ซึ่งเป็นข้อมูลอินพุตของระบบที่ใกล้เคียงกับความเป็นหน้าคนในบางลักษณะก็เป็นได้

4.4 สรุป

จากผลการทดสอบระบบที่ปัจจัยต่าง ๆ พบว่าระบบมีเปอร์เซ็นต์ความถูกต้องในการตรวจจับหน้าคนที่ดี ขึ้นอยู่กับปัจจัยดังต่อไปนี้

1. จำนวนภาพตัวอย่างและลักษณะของภาพตัวอย่างในการหาค่าไอเกนเวกเตอร์ ต้องมีจำนวนมากเพียงพอที่จะแสดงลักษณะเฉพาะของความเป็นหน้าคน และมีลักษณะหน้าคนที่มีความแน่นอน ซึ่งจากการทดลองพบว่าจำนวนภาพตัวอย่างในการหาค่าไอเกนเวกเตอร์ที่เหมาะสม คือ 20 ภาพ ของภาพบุคคลคนเดียวกันในลักษณะการแสดงออกของสีหน้าต่าง ๆ กัน
2. ค่าพารามิเตอร์ของระบบ โดยพบว่าเมื่อเราทำการปรับค่าสอดคล้อง ซึ่งเป็นค่าที่ใช้เป็นตัวกำหนดกลุ่มของค่าน้ำหนักประสาทของระบบโดยทำการสุ่มค่าด้วย GA ทำให้ได้ค่าสอดคล้องของระบบที่เหมาะสมคือมี ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802
3. ขนาดของหน้าคนในภาพ โดยหน้าคนที่มีขนาดใหญ่จะให้เปอร์เซ็นต์ความถูกต้องในการตรวจจับหน้าคนมากที่สุด กล่าวคือยิ่งหน้าคนในภาพมีขนาดใหญ่เท่าใด ก็จะทำให้รายละเอียดของลักษณะเฉพาะของหน้าคนเด่นชัดยิ่งขึ้นตามไปด้วย
4. สิ่งแวดล้อมและการวางท่าของบุคคลในภาพ พบว่าภาพที่สีพื้นหลังที่ไม่ซับซ้อนจะทำให้ระบบมีประสิทธิภาพในการตรวจจับหน้าคนมากกว่าภาพที่มีสีพื้นหลังที่ซับซ้อน และลักษณะการวางท่า เช่น มุมการหันของหน้า มือ แวนตามีผลให้เปอร์เซ็นต์ความถูกต้องในการตรวจจับหน้าคนลดลง

บทที่ 5

สรุปงานวิจัยและข้อเสนอแนะ

5.1 สรุปงานวิจัย

งานวิจัยวิทยานิพนธ์นี้ดำเนินการศึกษาและพัฒนา โครงสร้างระบบการตรวจจับหน้าคนด้วยโครงข่าย ART เพื่อให้สามารถนำไปประยุกต์ใช้ให้เป็นประโยชน์กับงานด้านอื่น ๆ ได้ต่อไป ซึ่งระบบที่พัฒนาขึ้นประกอบไปด้วยการคัดแยกสีผิวมนุษย์ออกจากสิ่งแวดล้อมโดยโครงข่ายประสาทเทียมแบบแพร่กลับ และการตรวจจับตำแหน่งของหน้าคนในภาพโดยโครงข่าย ART โดยใช้ไอคอนเวกเตอร์ที่ได้จากหลักการวิเคราะห์องค์ประกอบ มาเป็นอินพุตของโครงข่าย โดยใช้การค้นหาแบบ GA ในการหาค่าที่เหมาะสมที่สุดของค่าพารามิเตอร์ของระบบ และการดำเนินงานวิจัยวิทยานิพนธ์ดังกล่าวสำเร็จลุล่วงตามวัตถุประสงค์ โดยสามารถสรุปผลการศึกษาวิจัยและพัฒนาทางวิศวกรรมเป็นข้อสรุปได้ดังต่อไปนี้

การปฏิวัติวรรณกรรมและงานวิจัยที่เกี่ยวข้อง ได้รับการรายงานไว้ในบทที่ 2 ซึ่งพบว่าผู้สนใจทำการศึกษาวิจัยในเรื่องนี้อย่างกว้างขวาง และได้ใช้เทคนิควิธีการต่าง ๆ กันออกไปในการทดลอง โดยในการตรวจจับหน้าคนมีปัจจัยที่เป็นข้อจำกัดในการตรวจจับหน้าคนให้มีประสิทธิภาพดีเป็นจำนวนมาก ได้แก่ ความหลากหลายของลักษณะ การวางท่า สีผิวมนุษย์ที่แตกต่างกันไปตามเชื้อชาติ ความสว่างของแสง ตำแหน่งที่หน้าคนจะปรากฏ สีพื้นหลัง และการเปลี่ยนแปลงต่าง ๆ ที่สามารถเกิดขึ้นได้อยู่เสมอ ดังนั้นงานวิจัยนี้จึงดำเนินการพัฒนาวิธีการตรวจจับหน้าคนโดยโครงข่าย ART เพื่อให้สามารถตรวจจับหน้าคนในภาพได้ภายใต้ข้อจำกัดที่ได้กล่าวมาแล้วนั้น

โครงสร้างของโครงข่าย ART ได้กล่าวไว้ในบทที่ 3 โดยโครงสร้างของระบบประกอบด้วย การคัดแยกความเป็นหน้าคนและความไม่เป็นหน้าคนออกโดยโครงข่าย ART ที่ทำการประยุกต์ให้เสมือนมีโครงสร้างในชั้นเปรียบเทียบกับ 2 ชั้น

ในบทที่ 4 แสดงรายละเอียดของการทดสอบระบบในหลายปัจจัย เพื่อแสดงให้เห็นว่าระบบมีความหลากหลายในการตรวจจับหน้าคน สามารถจับหน้าคนได้ในภาพที่มีพื้นหลังที่ซับซ้อน ในสภาวะแสงและสิ่งแวดล้อมที่สามารถมองเห็นองค์ประกอบของหน้าคนได้ชัดเจน และสามารถตรวจจับหน้าคนในภาพได้มากกว่า 1 หน้า

5.2 ข้อเสนอแนะ

จากการดำเนินงานวิจัยวิทยานิพนธ์ที่ผ่านมาและผลที่ได้ทำให้เกิดแนวคิดและข้อเสนอแนะในการดำเนินงานวิจัยต่อไปในอนาคต ดังต่อไปนี้

1. ระบบของงานวิจัยนี้ยังไม่สามารถตรวจจับหน้าคนเมื่อมีสิ่งกีดขวางได้คั่น เนื่องจากระบบใช้ลักษณะเฉพาะของหน้าในการหาค่าไอเจนเวกเตอร์เป็นอินพุตของระบบ ดังนั้นเพื่อพัฒนาระบบให้สามารถตรวจจับหน้าคนแม้มีสิ่งกีดขวางซ้อนทับกับหน้าได้ ควรจะต้องเพิ่มเติมเทคนิคในการตรวจจับหน้าคน เช่นให้มีการตรวจจับดวงตา ร่วมกับการตรวจจับหน้าคนด้วยโครงข่าย ART เป็นต้น

2. เพื่อให้ระบบสามารถใช้งานได้อย่างหลากหลาย จำเป็นต้องมีภาพตัวอย่างที่ใช้ในการหาค่าไอเจนเวกเตอร์ที่หลากหลายที่จะใช้แทนกลุ่มข้อมูลได้ทุกกลุ่ม

3. เพื่อให้ระบบสามารถใช้งานได้ตามเวลาจริง (real time) ควรนำวิธีการที่ได้นำเสนอในงานวิจัยนี้ ซึ่งใช้การประมวลผลค่อนข้างสูง นำไปพัฒนาด้วยโปรแกรมที่มีการประมวลผลที่รวดเร็ว เช่น โปรแกรมด้วยภาษาซี เป็นต้น เพื่อให้สามารถคำนวณและแสดงผลการตรวจจับหน้าคนได้ตามเวลาจริง

รายการอ้างอิง

- อาทิตย์ ศรีแก้ว. (2552). **ปัญญาเชิงคำนวณ**. สำนักวิชาวิศวกรรมศาสตร์: มหาวิทยาลัยเทคโนโลยีสุรนารี.
- ชมพู ทรัพย์ปทุมสิน. (2548). **วิธีการใหม่แบบพันทางในการแยกส่วนมือจากภาพสี**. วิทยานิพนธ์ปริญญาวิศวกรรมศาสตรมหาบัณฑิต มหาวิทยาลัยเทคโนโลยีสุรนารี.
- Bellhumeur V., Hespanha J. and Kriegsmann D., (1997) **Eigenface vs. fisherface: recognition using class specific linear projection**, IEEE Trans. On PAMI, v. 19, pp. 711-720.
- Carpenter G. A. and Grossberg S., (1987) **ART2 Self-organizing of stable category recognition codes for analog input pattern**, In Applied Optics, vol. 26, pp. 4919-4930.
- Carpenter G. A., Grossberg S. and Reynolds J.H., (1991) **ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network**, Neural Network, vol. 4, pp. 565-588.
- Carpenter G. A., Grossberg S. and Rosen D.B., (1991) **Fuzzey ART : Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System**, Neural Network, vol. 4, pp. 759-771.
- Chai D. and Ngan K.N., (1998) **Locating Facial Region of a Head-and-Shoulders Color Image**, Proc. Third Int'l Conf. Automatic Face and Gesture Recognition, pp. 124-129.
- Chetverikov D. and Lerch A., (1993) **Multiresolution Face Detection**, Theoretical Foundations of Computer Vision, vol. 69, pp. 131-140.
- Crowley J.L. and Berard F., (1997) **Multi-Modal Tracking of Faces for Video Communications**, Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 640-645.
- Duda R.O., Hart P.E. and Stork D.G., (2001) **Pattern Classification**, John Wiley & Sons.
- Edelman S., Intrator N. and Poggio T., (1998) **Complex cells and object recognition**, Vision Res., vol. 38, pp. 2657-2684.
- Hotelling, (2001) **A primer of Multivariate Statistics**, Lawrence Erlbaum Associates.
- Gong Y.Li, Sherrah S.J. and Liddell H., (2000) **Multi-view Face Detection Using Support Vector Machines and Eigenspace Modelling**, Proc. IEEE 2000

- Graf H.P., Chen T., Petajan E., and Cosatto E., (1995) **Locating Faces and Facial Parts**, Proc. First Int'l Workshop Automatic Face and Gesture Recognition, pp. 41-46.
- Guo G., Li S.Z. and Chan K., (1998) **Face Recognition by Support Vector Machines**.
- Jung D. J., Lee C.W. and Lee Y.C., (2002) **PCA-Base Real-Time Face Detection and Tracking**.
- Kelly M.D., (1970) **Visual identification of people by computer**, Tech. rep. AI-130, Stanford AI Project, Stanford.
- Kotropoulos C. and Pitas I., (1997) **Rule-Based Face Detection in Frontal Views**, Proc. Int'l Conf. Acoustics, Speech and Signal Processing, vol. 4, pp. 2537-2540.
- Lawrence S., Giles C.L., Tsois A.C. and Back A.D., (1997) **Face recognition: A convolutional neural-network approach**, IEEE Trans. Neural Network 8, pp. 98-113.
- Li S.Z. and Lu J., (1999) **Face recognition using the nearest feature line method**, IEEE Trans. Neural Netw. 10, pp. 439-443.
- Malciu M. and Preteux F., (2002) **Tracking facial features in video sequences using a deformable model-based approach**, ARTEMIS Project Unit.
- Mitsukura Y., Fukumi M. and Akamatsu N., (2001) **A detection Method of Face Regions in Color Image by Using Evolutionally Computation**, Proc. IEEE 2001.
- Nefian A.V. et al., (1998) **Hidden markov models for face recognition**, In Proceedings, International Conference on Acoustics, Speech and Signal Proceeding, pp. 2721-2724.
- Osuna E., Freund R., Girosi and F., (1997) **Training Support Vector Machines: An Application to Face Detection**, Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 130-136.
- Penev P. and Atick J., (1996) **Local feature analysis: A general statistical theory for object representation**, Neural Syst. 7, pp. 477-500.
- Phillips P.J., Wechsler H., Huang J. and Rauss P., (1998) **The FERET database and evaluation procedure for face-recognition algorithms**, Image Vis. Comput. 16, pp. 295-306.
- Rowley H., Baluja S., and Kanade T., (1998) **Rotation Invariant Neural Network-Based Face Detection**, Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 38-44.
- Sakai T., Nagao M., and Fujibayashi S., (1969) **Line Extraction and Pattern Detection in a Photograph**, Pattern Recognition, vol. 1, pp. 233-248.
- Saxe D. and Foulds R., (1996) **Toward Robust Skin Identification in Video Images**, Proc. Second Int'l Conf. Automatic Face and Gesture Recognition, pp. 379-384.
- Sung K.-K. and Poggio T., (1998) **Example-Based Learning for View-Based Human Face Detection**, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 39-51, Jan.

- Tanaka T. and Weitzenfeld A., (1997) **Adaptive Resonance Theory**, pp.157-165.
- Tsukamoto A., Lee C.-W., and Tsuji S., (1994) **Detection and Pose Estimation of Human Face with Synthesized Image Models**, Proc. Int'l Conf. Pattern Recognition, pp. 754-757.
- Turk M. and Pentland A., (1991) **Eigenfaces for Recognition**, J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86.
- Umeki H. and Mizutani H., (1996) **Dynamic Link Matching for Multiple Object Recognition**, Proc. IEEE 1996.
- Wu C.J., Sung A.H. and Soliman H.S., (1994) **A Fuzzey ART Network with Fuzzey Control for Image Data Compression**, Proc. of IASTED Int. Conf. on Modeling, Simulation and Control in the Process Industry, pp. 95-98.
- Wu H., Yokoyama T., Pramadihanto D., and Yachida M., (1996) **Face and Facial Feature Extraction from Color Image**, Proc. Second Int'l Conf. Automatic Face and Gesture Recognition, pp. 345-350.
- Yang G. and Huang T.S., (1994) **Human Face Detection in Complex Background**, Pattern Recognition, vol. 27, no. 1, pp. 53-63.
- Yokoo Y. and M Hagiwara., (1996) **Human Face Detection Method using Genetic Algorithm**, Proc. IEEE 1996.
- Zhao W., Chellappa R., Phillips P.J. and Rosenfeld A., (2003) **Face recognition : A literature survey**, ACM Computing surveys, v. 35, pp. 399-458.

ภาคผนวก ก

โครงข่ายประสาทเทียมแบบแพร่กลับ

เมื่อป้อนแต่ละอินพุตให้กับโครงข่าย เอาต์พุตที่ได้จะถูกนำไปเปรียบเทียบกับเป้าหมาย อัลกอริทึมจะทำการปรับพารามิเตอร์ของโครงข่าย ซึ่งได้แก่ค่าน้ำหนักประสาทและค่าไบอัส เพื่อให้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยของเอาต์พุตและเป้าหมายมีค่าน้อยที่สุด และมีตัวชี้วัดประสิทธิภาพ (performance index) คือ

$$F(\mathbf{x}) = E[(\mathbf{t}-\mathbf{y})^T(\mathbf{t}-\mathbf{y})] \quad (\text{ก.1})$$

โดยที่ $\mathbf{x}$ คือเมทริกซ์ของค่าน้ำหนักประสาทและค่าไบอัส ซึ่งอัลกอริทึมแบบแพร่กลับสำหรับปรับค่าน้ำหนักประสาทและค่าไบอัส ณ รอบที่ $k+1$ ด้วยการประมาณค่าความผิดพลาดกำลังสองเฉลี่ยและค่าคงที่การเรียนรู้ α แสดงในสมการที่ ก.2 และ สมการที่ ก.3

$$w_{ij}^m(k+1) = w_{ij}^m(k) - \alpha \frac{\partial \hat{F}}{\partial w_{ij}^m} \quad (\text{ก.2})$$

$$b_i^m(k+1) = b_i^m(k) - \alpha \frac{\partial \hat{F}}{\partial b_i^m} \quad (\text{ก.3})$$

ได้เอาต์พุตของโครงข่ายชั้นที่ m เป็นฟังก์ชันของค่าน้ำหนักประสาทและค่าไบอัส แสดงในสมการที่ ก.4

$$n_i^m = \sum_{j=1}^{S^{m-1}} w_{ij}^m y_j^{m-1} + b_i^m \quad (\text{ก.4})$$

โดยกำหนดให้ความไวของค่าความผิดพลาด (error sensitivity) คือ

$$\delta_i^m = \frac{\partial \hat{F}}{\partial n_i^m} \quad (\text{ก.5})$$

ซึ่งสามารถแยกพิจารณาได้เป็น 2 กรณีคือ กรณีชั้นที่ m เป็นชั้นเอาต์พุต ($m=N$) โดยที่ N คือจำนวนชั้นของโครงข่าย จะได้

$$\delta_i^N = -2(t_i - y_i^N) f'(n_i^N) \quad (\text{ก.6})$$

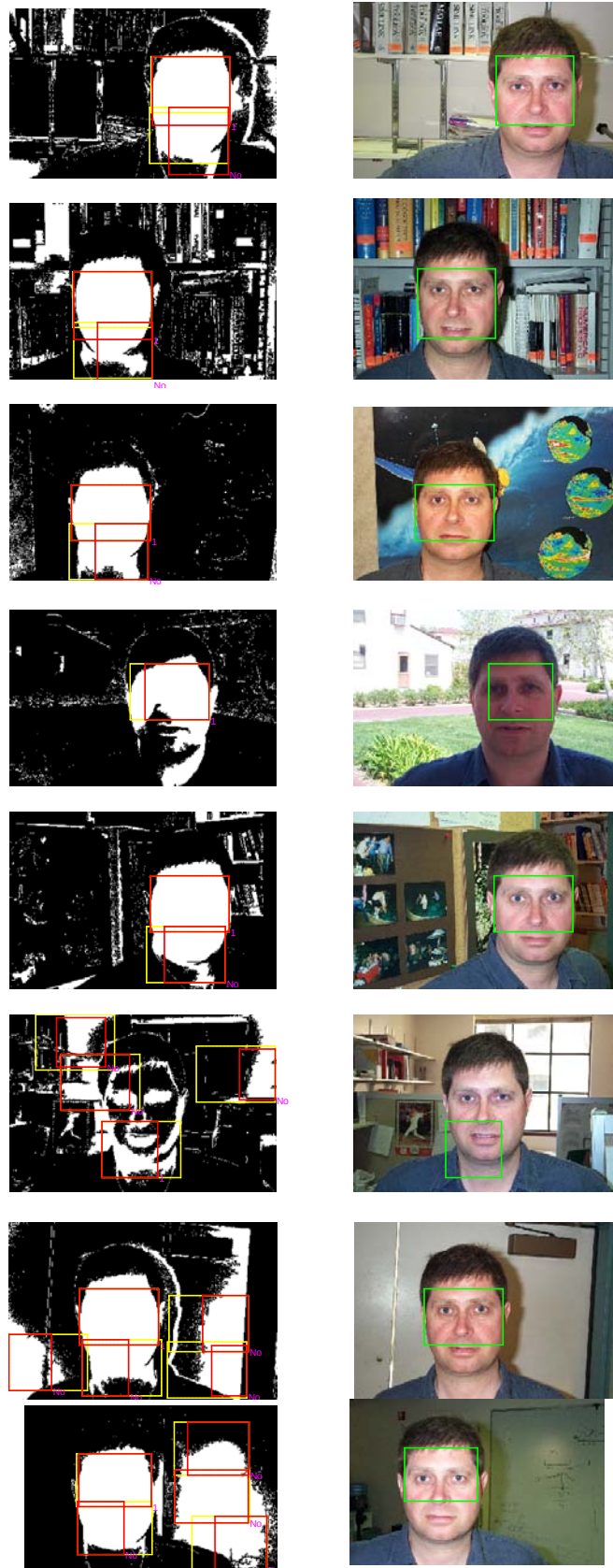
และในกรณีชั้นที่ m เป็นชั้นซ่อนเร้น จะได้

$$\delta_i^m = f'(n_i^m) \sum_{l=1}^{S^{m+1}} \delta_l^{m+1} w_{il}^{m+1} \quad (\text{ก.7})$$

จากสมการข้างต้นแสดงให้เห็นว่า ค่าความไวของค่าความผิดพลาดที่ได้ในชั้นที่พิจารณาสามารถคำนวณได้จากองค์ประกอบในชั้นถัดไป ซึ่งองค์ประกอบค่าความไว เป็นส่วนที่ทำให้อัลกอริทึมแบบแพร่กลับแตกต่างไปจากอัลกอริทึม LMS การปรับค่าน้ำหนักประสาทและไบอัสยังเป็นไปในแบบลงชันสุด (steepest descent) ซึ่งจะนำโครงข่ายไปในทิศทางที่ทำให้ค่าความผิดพลาดกำลังสองเฉลี่ยมีค่าน้อยที่สุด

ภาคผนวก ข

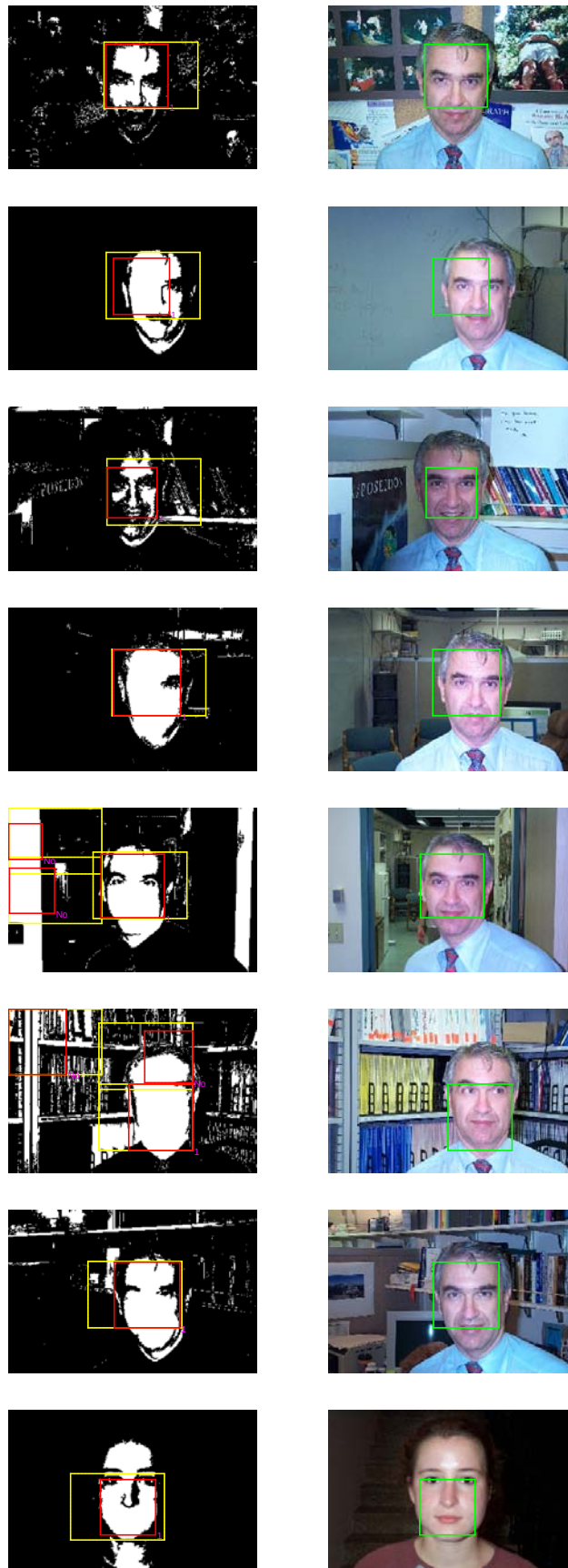
ภาพผลการทดสอบระบบ



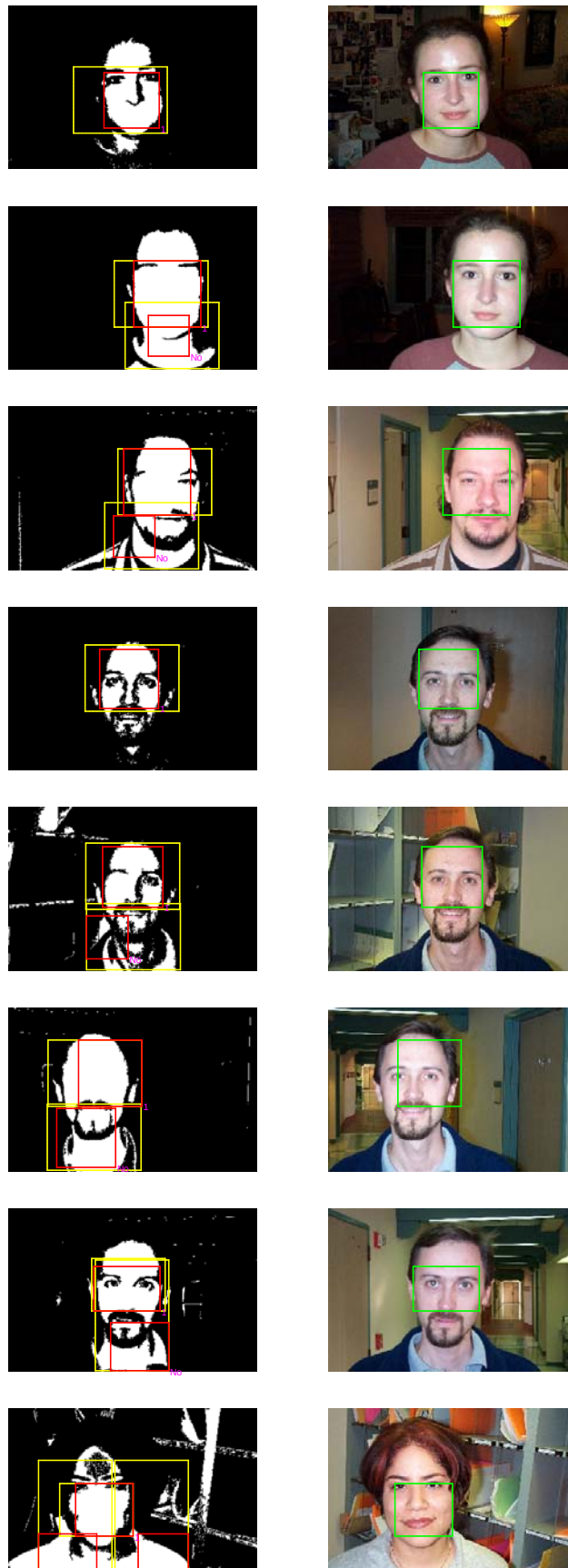
รูปที่ ข.1 ภาพผลการทดสอบระบบ



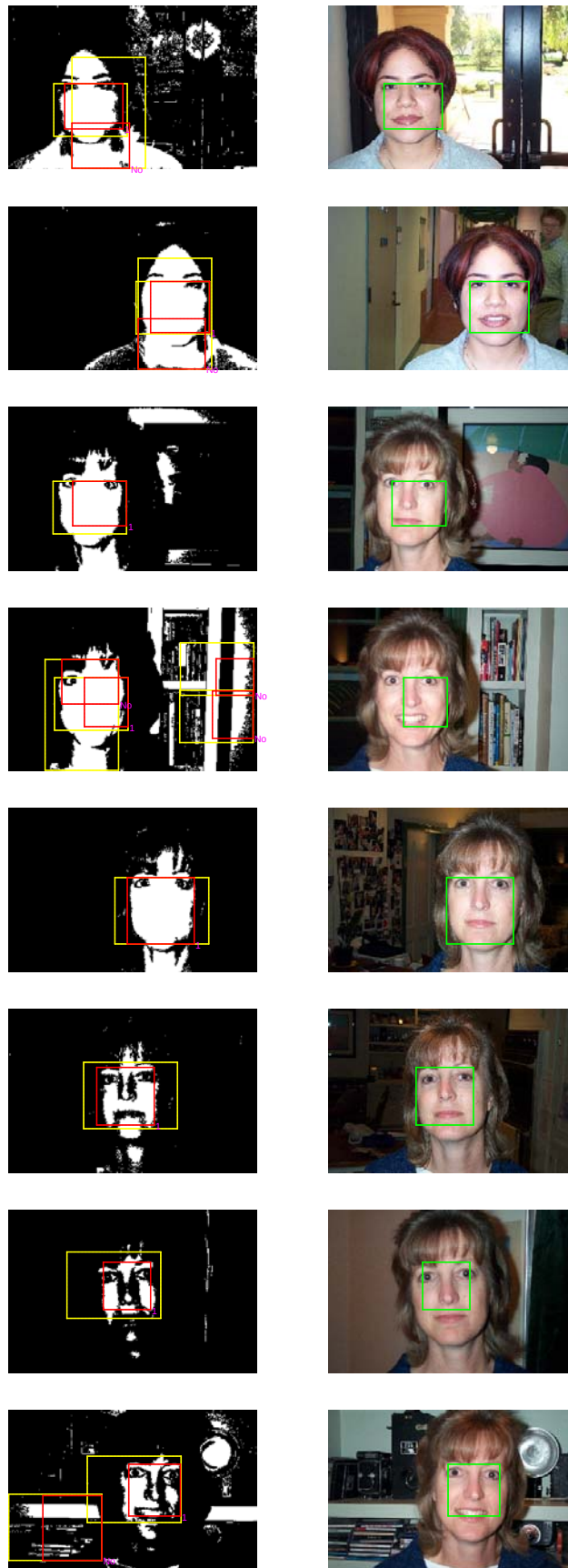
รูปที่ ข.1 ภาพผลการทดสอบระบบ (ต่อ)



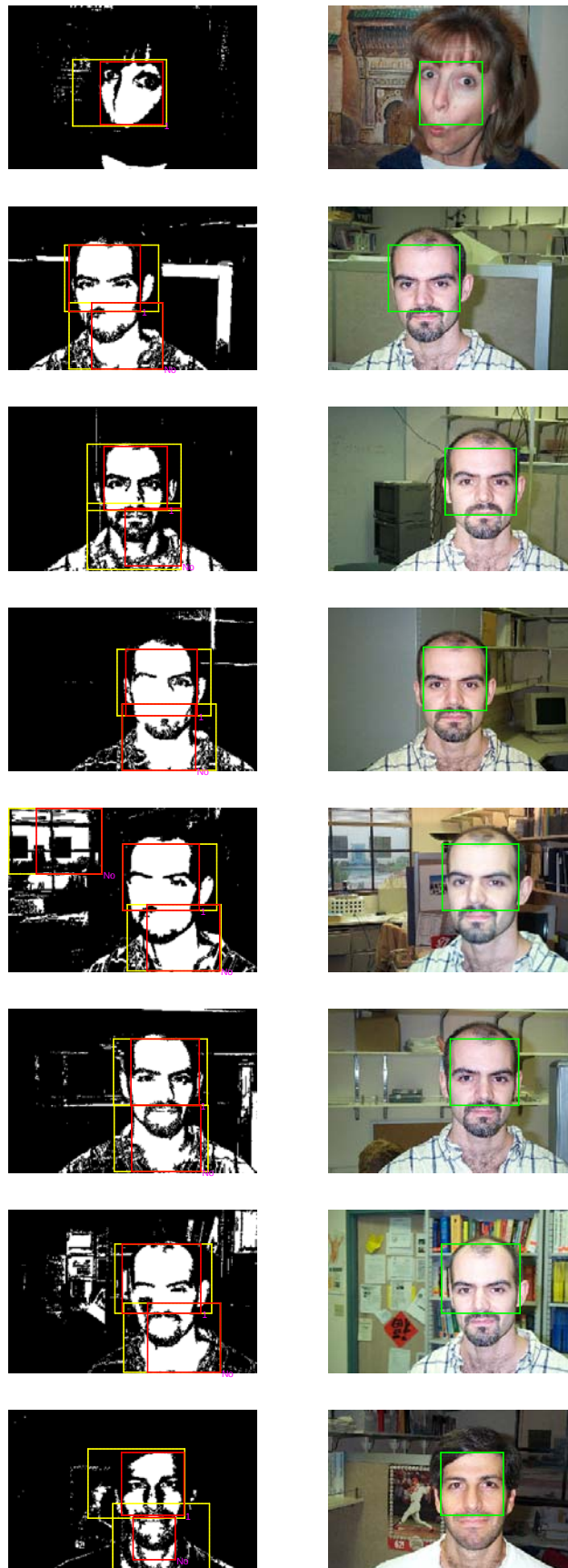
รูปที่ ข.1 ภาพผลการทดสอบระบบ (ต่อ)



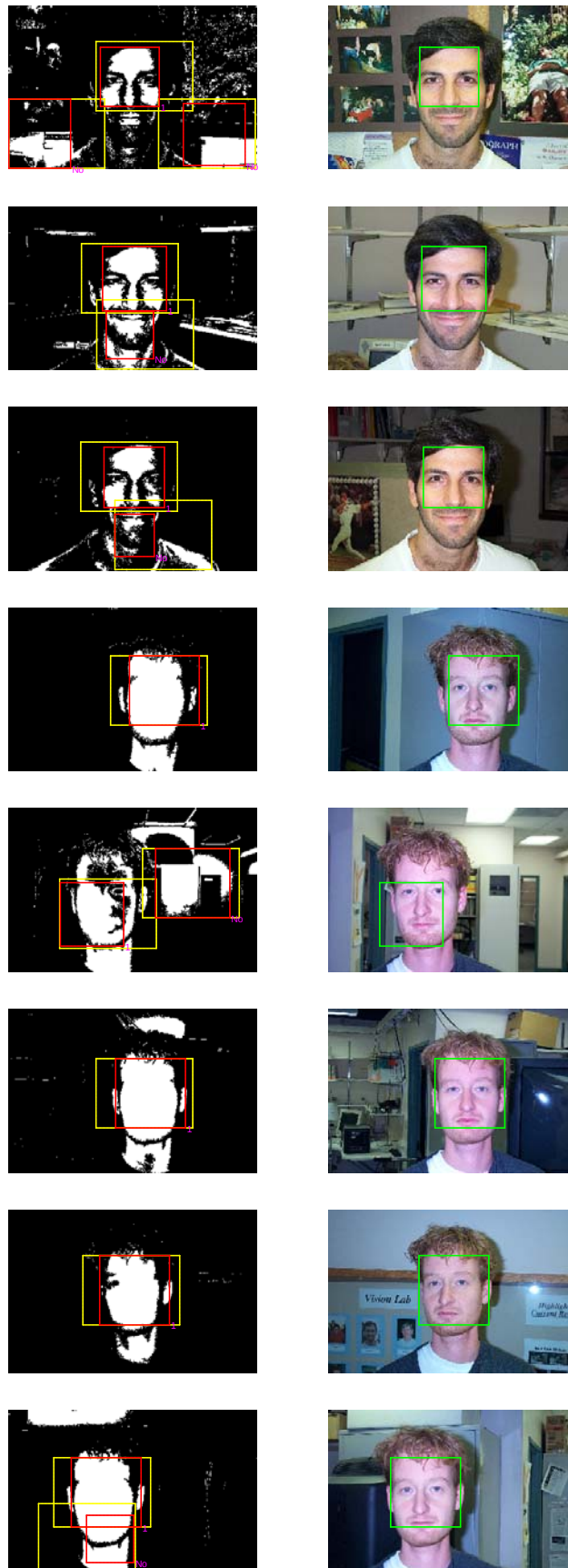
รูปที่ ข.1 ภาพผลการทดสอบระบบ (ต่อ)



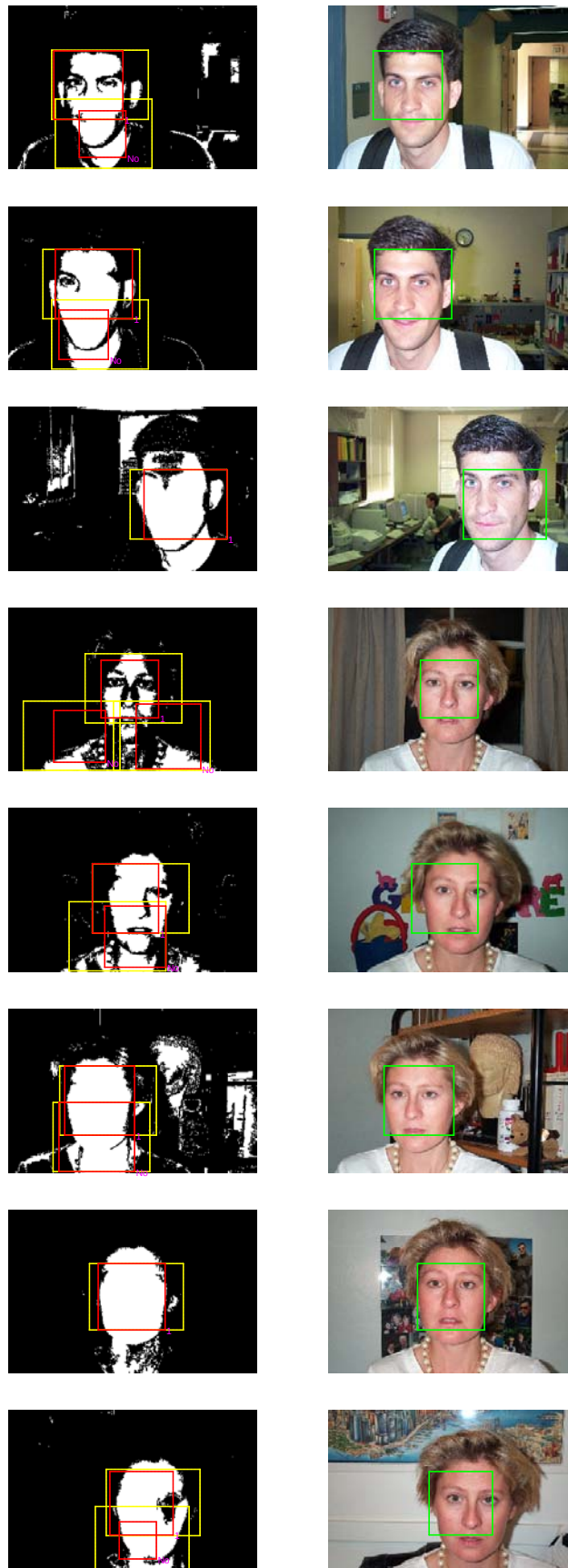
รูปที่ ข.1 ภาพผลการทดสอบระบบ (ต่อ)



รูปที่ ข.1 ภาพผลการทดสอบระบบ (ต่อ)



รูปที่ ข.1 ภาพผลการทดสอบระบบ (ต่อ)



รูปที่ ข.1 ภาพผลการทดสอบระบบ (ต่อ)



รูปที่ ข.1 ภาพผลการทดสอบระบบ (ต่อ)

ภาคผนวก ค

บทความทางวิชาการที่ได้รับการตีพิมพ์เผยแพร่ในระหว่างศึกษา

บทความทางวิชาการที่ได้รับการตีพิมพ์เผยแพร่ในระหว่างศึกษา

บทความวิจัยที่ได้รับการตีพิมพ์ในการประชุมวิชาการระดับชาติ

อภิรดี อัมพะละสิริ (2550). เทคนิคการตรวจจับหน้าคนด้วยโครงข่าย ART, การประชุมทางวิชาการ
วิทยาการคอมพิวเตอร์ และวิศวกรรมคอมพิวเตอร์แห่งชาติ ครั้งที่ 11 (NCSEC 2007), กรุงเทพฯ,
หน้า 447-452.

การเผยแพร่ผลงานในการประชุมวิชาการระดับชาติ

อภิรดี อัมพะละสิริ (2551). เทคนิคการตรวจจับหน้าคนด้วยโครงข่าย ART, การประชุมวิชาการ
บัณฑิตศึกษา มทส. ครั้งที่ 1 ณ มหาวิทยาลัยเทคโนโลยีสุรนารี จังหวัดนครราชสีมา.

ประวัติผู้เขียน

นางสาวอภิรดี อัมพะสวัสดิ์ เกิดเมื่อวันที่ 18 ตุลาคม พ.ศ. 2523 ที่อำเภอเมือง จังหวัดสุโขทัย สำเร็จการศึกษาระดับมัธยมศึกษาตอนปลายจากโรงเรียนสุโขทัยวิทยาคม จังหวัดสุโขทัย และ สำเร็จการศึกษาระดับปริญญาตรี วิศวกรรมศาสตรบัณฑิต (วิศวกรรมไฟฟ้า) จากมหาวิทยาลัยเทคโนโลยีสุรนารี จังหวัดนครราชสีมา เมื่อ พ.ศ. 2545 ขณะศึกษาระดับปริญญาโท เคยสอนวิชาปฏิบัติการของสาขาวิชาวิศวกรรมไฟฟ้า สำนักวิชาวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี ได้แก่ ปฏิบัติการวงจรและอุปกรณ์ ปฏิบัติการวงจรอิเล็กทรอนิกส์ ปฏิบัติการวิศวกรรมไฟฟ้า 1 ปฏิบัติการเครื่องจักรกลไฟฟ้า และมีผลงานทางวิชาการที่ได้รับการตีพิมพ์เผยแพร่ในขณะศึกษา ดังรายชื่อที่ปรากฏในภาคผนวก ค. มีความสนใจทางด้านการประมวลผลสัญญาณภาพ การมองเห็นของคอมพิวเตอร์