

รหัสโครงการ SUT7-711-51-12-50



รายงานการวิจัย

เทคนิคการตรวจจับใบหน้าคนด้วยโครงข่าย ART แบบดัดแปลง

Human Face Detection Technique Based-on Modified

Adaptive Resonance Theory Network

ได้รับทุนอุดหนุนการวิจัยจาก

มหาวิทยาลัยเทคโนโลยีสุรนารี

ผลงานวิจัยเป็นความรับผิดชอบของหัวหน้าโครงการวิจัยแต่เพียงผู้เดียว



รายงานการวิจัย

เทคนิคการตรวจจับใบหน้าคนด้วยโครงข่าย ART แบบดัดแปลง

Human Face Detection Technique Based-on Modified

Adaptive Resonance Theory Network

คณะผู้วิจัย

หัวหน้าโครงการ

รองศาสตราจารย์ ดร.อาทิตย์ ศรีแก้ว

สาขาวิชาวิศวกรรมไฟฟ้า

สำนักวิชาวิศวกรรมศาสตร์

มหาวิทยาลัยเทคโนโลยีสุรนารี

ผู้ร่วมวิจัย

นางสาวอภิรดี อัมพะศิริ

ได้รับทุนอุดหนุนการวิจัยจากมหาวิทยาลัยเทคโนโลยีสุรนารี ปีงบประมาณ พ.ศ.2551

ผลงานวิจัยเป็นความรับผิดชอบของหัวหน้าโครงการวิจัยแต่เพียงผู้เดียว

กุมภาพันธ์ 2553

บทคัดย่อ

ระบบการตรวจจับหน้าคนเป็นส่วนหนึ่งของเทคโนโลยีทางคอมพิวเตอร์สำหรับการปฏิสัมพันธ์กันระหว่างคนและคอมพิวเตอร์ ซึ่งตลอดเวลาของการทำงาน คอมพิวเตอร์จำเป็นต้องทราบตำแหน่งของคนที่ต้องการปฏิสัมพันธ์ด้วย การศึกษาและพัฒนาระบบการตรวจจับและติดตามหน้าคนนั้นจึงยังคงเป็นหัวข้อที่ได้รับความสนใจอยู่ 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

Abstract

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

กิตติกรรมประกาศ

โครงการวิจัยนี้ได้รับความร่วมมือช่วยเหลือในด้านต่างๆ จากหลายๆ ฝ่าย จนสำเร็จไปได้ด้วยดี คณะผู้วิจัยขอขอบพระคุณศูนย์เครื่องมือวิทยาศาสตร์และเทคโนโลยี ที่ให้ความเอื้อเฟื้อทั้งทางด้านสถานที่ เครื่องมือและบุคลากร ขอขอบพระคุณสาขาวิชาวิศวกรรมไฟฟ้าและสำนักวิศวกรรมศาสตร์สำหรับการสนับสนุนในทุกๆ ด้าน การวิจัยครั้งนี้ได้รับทุนอุดหนุนการวิจัยจากมหาวิทยาลัยเทคโนโลยีสุรนารี ปีงบประมาณพ.ศ. 2551

คณะผู้วิจัย



สารบัญ

หน้า

บทคัดย่อ (ภาษาไทย).....	ก
บทคัดย่อ (ภาษาอังกฤษ).....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญรูป.....	ช
สารบัญตาราง.....	ฉ

บทที่

1 การตรวจจับใบหน้าคนด้วยการประมวลผลภาพ.....	1
1.1 กล่าวนำ.....	1
1.2 วัตถุประสงค์การวิจัย.....	1
1.3 ขอบเขตของงานวิจัย.....	1
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	2
2 ปรัชมนวัตกรรมกรรมและงานวิจัยที่เกี่ยวข้อง.....	3
2.1 การทบทวนวรรณกรรมที่เกี่ยวข้อง.....	3
2.1.1 วิธีเชิงความรู้.....	3
2.1.2 วิธีเชิงลักษณะ.....	3
2.1.3 วิธีเทียบเคียงแผ่นแบบ.....	4
2.1.4 วิธีเชิงลักษณะปรากฏ.....	5

สารบัญ (ต่อ)

	หน้า
2.1.4.1 วิธีหน้าลักษณะเฉพาะ.....	5
2.1.4.2 วิธีเชิงการกระจาย.....	6
2.1.4.3 โครงข่ายประสาทเทียม.....	6
2.1.4.4 วิธีเวกเตอร์เกี่ยวพัน.....	6
2.2 สรุป.....	7
3 ระบบการตรวจจับใบหน้าคนด้วยโครงข่าย ART.....	8
3.1 ระบบการคัดแยกสีผิวมนุษย์ออกจากพื้นหลัง.....	8
3.1.1 ค่าสีผิวมนุษย์.....	8
3.1.2 โครงข่ายประสาทเทียม.....	9
3.1.3 การคัดแยกสีผิวมนุษย์ด้วยโครงข่ายประสาทเทียมแบบแพร่กลับ.....	10
3.2 ระบบการแปลงข้อมูลภาพเป็นการวิเคราะห์ห้วงค์ประกอบหลัก.....	12
3.2.1 การหาค่าไอเกนและไอเกนเวกเตอร์ด้วยหลักการวิเคราะห์ห้วงค์ประกอบ.....	14
3.3 โครงข่าย ART.....	16
3.3.1 สถาปัตยกรรมของโครงข่าย ART.....	17
3.3.2 ประเภทของโครงข่าย ART.....	18
3.3.2.1 โครงข่าย ART1.....	18
3.3.2.2 โครงข่ายART2.....	18
3.3.2.3 โครงข่าย ARTMAP.....	19

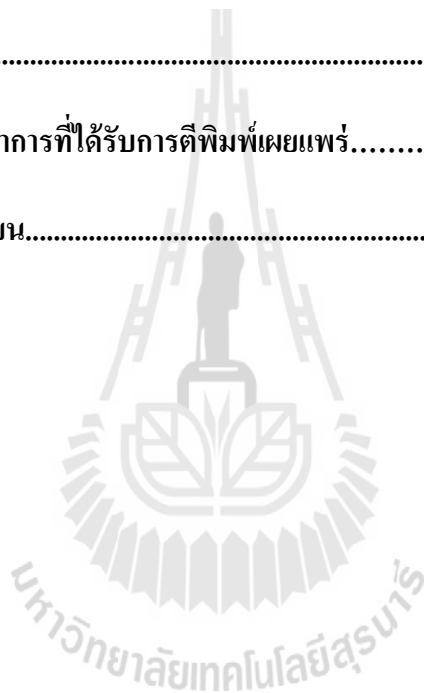
สารบัญ (ต่อ)

หน้า

3.2.2.4	โครงข่าย Fuzzy ART หรือ Fuzzy ARTMAP.....	20
3.4	ระบบตรวจจับหน้าคนจากภาพด้วยโครงข่าย ART.....	20
3.4.1	โครงสร้างของระบบ.....	20
3.4.2	หลักการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART.....	26
3.4.3	ตัวอย่างการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART.....	27
4	ผลการทดสอบระบบตรวจจับหน้าคนด้วยโครงข่าย ART.....	28
4.1	กล่าวนำ.....	28
4.2	ฐานข้อมูลภาพหน้ามนุษย์และภาพที่ใช้ในการทดสอบระบบ.....	28
4.2.1	ฐานข้อมูลภาพ CVSR.....	28
4.2.2	ฐานข้อมูลภาพ CMU VASC.....	29
4.2.3	ฐานข้อมูลภาพ AVTG.....	30
4.2.4	ฐานข้อมูลภาพ VALID และฐานข้อมูลภาพ CBCL.....	31
4.2.5	ภาพจากกล้องดิจิตอลฟูจิ F 601 และกล้องวิดีโอ.....	33
4.3	ทดสอบระบบตรวจจับหน้าคนด้วยโครงข่าย ART.....	34
4.3.1	ทดสอบระบบเมื่อมีการฝึกสอนโครงข่ายในลักษณะต่าง ๆ.....	34
4.3.2	ทดสอบระบบที่พารามิเตอร์ของโครงข่ายต่าง ๆ.....	46
4.3.3	ทดสอบระบบที่ขนาดของใบหน้าในภาพขนาดต่าง ๆ.....	48
4.3.4	ทดสอบระบบที่หน้าในลักษณะต่าง ๆ.....	52

สารบัญ (ต่อ)

	หน้า
4.3.5 ทดสอบระบบกับภาพที่มีหน้าคนในภาพมากกว่า 1 คน.....	56
4.4 สรุป.....	59
5 สรุปงานวิจัยและข้อเสนอแนะ.....	60
5.1 สรุปงานวิจัย.....	60
5.2 ข้อเสนอแนะ.....	60
รายการอ้างอิง.....	62
ภาคผนวก ก. บทความวิชาการที่ได้รับการตีพิมพ์เผยแพร่.....	65
ภาคผนวก ข. ประวัติผู้เขียน.....	67



สารบัญรูป

รูปที่	หน้า
2.1 ขอบเขตการพิจารณาของแต่ละลำดับชั้น.....	3
2.2 แบบจำลองไบน์ามนุษย์จากฐานข้อมูล MPEG-4.....	5
2.3 การกระจายตัวของข้อมูลตัวอย่าง.....	6
3.1 ตัวอย่างอินพุตของโครงข่ายประสาทเทียม.....	8
3.2 ค่าสีผิวมนุษย์ (ก) ปริภูมิ RGB (ข) ปริภูมิ HSV.....	9
3.3 ผลการคัดแยกสีผิวมนุษย์ออกจากสิ่งแวดล้อมด้วยโครงข่ายประสาทเทียมแบบแพร่กลับ (ก) ภาพต้นแบบ (ข) ผลการคัดแยกสีผิวมนุษย์ (ค) ขอบเขตของสีผิวมนุษย์ (ง) ผลการฝึกสอนโครงข่าย (จ) ขอบเขตการค้นหาน้ำคน.....	12
3.4 โครงข่ายประสาทเทียมแบบแพร่กลับในการคัดแยกสีผิวมนุษย์.....	12
3.5 ภาพตัวอย่างในการหาหน้าไอเคน (ก) ตัวอย่างภาพในปริภูมิสี RGB (ข) ตัวอย่างภาพในปริภูมิสี HSV.....	14
3.6 ภาพระดับเทาที่ทำการลดผลของความสว่าง (ก) สีผิวกลาง (ข) สีผิวเข้ม (ค) สีผิวอ่อน.....	15
3.7 ตัวอย่างการแปลงเมทริกซ์ภาพเป็นเวกเตอร์ภาพ.....	15
3.8 ภาพหน้าไอเคน แถวที่ 1 ภาพต้นแบบ แถวที่ 2 ภาพหน้าไอเคนจากไอเคนเวกเตอร์ค่าที่ 1 แถวที่ 3 ภาพหน้าไอเคนจากไอเคนเวกเตอร์ค่าที่ 10.....	16
3.9 แผนผังการทำงานของโครงข่าย ART.....	17
3.10 โครงข่ายประสาทเทียม ART.....	18
3.11 แผนผังการทำงานของโครงข่าย ART ในการหาค่าสอดคล้องที่เหมาะสมด้วย GA.....	22
3.12 แผนผังการค่าสอดคล้องที่เหมาะสมด้วย GA.....	23
3.13 โครงสร้างระบบการตรวจจับหน้าคนด้วยโครงข่าย ART.....	24

สารบัญรูป (ต่อ)

รูปที่	หน้า
3.14	แผนผังการโครงข่าย ART ในการคัดแยกหน้าคน..... 25
3.15	กล่องข้อมูลอินพุตของ โครงข่าย ART (ก) เส้นกำหนดขนาดกล่องข้อมูล (ข) กล่องข้อมูลอินพุตที่ ป้อนเข้าโครงข่าย ART (ค) กล่องข้อมูลที่ดีที่สุดที่แสดงตำแหน่งของหน้าคนในภาพ..... 26
3.16	ภาพตัวอย่างการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART (ก) ตัวอย่างภาพต้นแบบ (ข) ตัวอย่างภาพผลการทำงานของโครงข่าย ART (ค) ตัวอย่างภาพผลการตรวจจับหน้าคนด้วยโครงข่าย ART..... 27
4.1	ภาพตัวอย่างจากฐานข้อมูล CVSR.....29
4.2	ภาพตัวอย่างจากฐานข้อมูล CMU VASC.....30
4.3	ภาพตัวอย่างจากฐานข้อมูล AVTG.....31
4.4	ภาพตัวอย่างจากฐานข้อมูล VALID.....32
4.5	ภาพตัวอย่างจากฐานข้อมูล CBCL.....32
4.6	ภาพตัวอย่างจากกล้องฟูจิ F601.....33
4.7	ภาพตัวอย่างจากกล้องวีดีโอ.....34
4.8	ภาพตัวอย่างในการฝึกสอนโครงข่าย.....34
4.9	ตัวอย่างภาพหน้าไอเคน (ก) ภาพหน้าไอเคนที่จำนวนภาพในการหาหน้าไอเคน 3 ภาพ (ข) ภาพหน้า ไอเคนที่จำนวนภาพในการหาหน้าไอเคน 5 ภาพ (ค) ภาพหน้าไอเคนที่จำนวนภาพในการหาหน้า ไอเคน 10 ภาพ.....39
4.10	ตัวอย่างภาพผลการทดสอบระบบที่จำนวนตัวอย่างในการฝึกสอนต่าง ๆ.....37
4.11	ตัวอย่างภาพหน้าไอเคน.....42
4.12	ตัวอย่างภาพผลการทดสอบระบบ ของภาพตัวอย่างในการฝึกสอนที่คุณลักษณะต่าง ๆ.....44

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.13 ตัวอย่างภาพการทดสอบระบบของระบบที่กำหนดค่าสอดคล้อง ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802.....	48
4.14 ตัวอย่างภาพของหน้าคนในภาพขนาดต่าง ๆ (ก) หน้าคนขนาด 70% (ข) หน้าคนขนาด 55% (ข) หน้าคนขนาดน้อยกว่า 50%.....	49
4.15 ตัวอย่างภาพผลการทดสอบระบบ ของใบหน้าคนในภาพขนาดต่าง ๆ	50
4.16 ตัวอย่างภาพผลการทดสอบระบบ ของภาพหน้าในลักษณะต่าง ๆ	54
4.17 ตัวอย่างภาพผลการทดสอบระบบ ของภาพที่มีจำนวนหน้าคนในภาพมากกว่า 1 หน้า.....	57



สารบัญตาราง

ตารางที่	หน้า
3.1 ผลการฝึกสอน โครงข่ายประสาทเทียมเพื่อเลือกพารามิเตอร์ของโครงข่าย.....	11
4.1 ทดสอบระบบที่จำนวนภาพในการฝึกสอนโครงข่ายต่าง ๆ	36
4.2 ทดสอบระบบด้วยภาพที่มีคุณลักษณะต่าง ๆ	43
4.3 ความถูกต้องในการคัดแยกหน้าคนโดยค่าสอดส่องค่าต่าง ๆ	47
4.4 ความถูกต้องในการคัดแยกหน้าคนเมื่อขนาดของใบหน้าในภาพเปลี่ยนแปลง	49
4.5 ความถูกต้องในการคัดแยกหน้าคนในลักษณะต่าง ๆ	53





บทที่ 1

การตรวจจับใบหน้าคนด้วยการประมวลผลภาพ

1.1 กล่าวนำ

ในปัจจุบันเทคโนโลยีด้านหุ่นยนต์ได้รับความสนใจและพัฒนามากขึ้นอย่างรวดเร็ว มีการแข่งขันและส่งเสริมให้มีการคิดค้นเทคนิคใหม่ ๆ เพื่อให้หุ่นยนต์สามารถทำงานได้ใกล้เคียงกับมนุษย์มากที่สุด หนึ่งในการพัฒนาเพื่อให้หุ่นยนต์มีความใกล้เคียงกับมนุษย์คือ เพิ่มการมองเห็นให้หุ่นยนต์ เพื่อการเชื่อมต่อระหว่างคนและคอมพิวเตอร์ ซึ่งเป็นส่วนหนึ่งของการพัฒนาเทคโนโลยีการมองเห็นของคอมพิวเตอร์ (computer vision) การที่คอมพิวเตอร์จะสามารถสื่อสารกับมนุษย์ให้มีประสิทธิภาพดีคอมพิวเตอร์จำเป็นจะต้องทราบตำแหน่งของมนุษย์ที่ต้องการจะสื่อสารด้วย จึงมีการเพิ่มตาให้คอมพิวเตอร์ ซึ่งก็คือกล้องรับภาพ ดังนั้นเทคนิคการตรวจจับใบหน้าบุคคลจากภาพสี จึงเป็นเทคโนโลยีที่สำคัญในการพัฒนาเทคโนโลยีการมองเห็นของคอมพิวเตอร์

ระบบการตรวจจับใบหน้าคนจากภาพ ได้มีการค้นคว้าพัฒนาเทคนิคและวิธีการต่าง ๆ อยู่เสมอ โดยงานวิจัยการค้นหาหน้าคนจากภาพนั้น มีผู้ศึกษาและค้นคว้าวิจัยเรื่องนี้มากกว่า 20 ปีแล้ว และได้รับความสนใจมากในช่วง 5 ปีหลังที่ผ่านมา เหตุผลที่ทำให้งานนี้ได้รับความสนใจเนื่องจากความท้าทายในการค้นหาหน้าคน ที่มีความหลากหลายของใบหน้าที่ไม่ซ้ำกันในแต่ละบุคคล รวมถึงความเปลี่ยนแปลงไม่แน่นอนในการปรากฏของใบหน้า เช่น โครงสร้างทางกายภาพของใบหน้าที่แตกต่างกันไปตามเชื้อชาติ ตำแหน่งที่ตั้ง การวางท่า หรือการแสดงออกของสีหน้า เป็นต้น และปัจจัยที่เป็นสิ่งแวดล้อมของภาพที่มีใบหน้าปรากฏอยู่ก็มีผลต่อความถูกต้องในการตรวจจับใบหน้าด้วยเช่นกัน ซึ่งได้แก่ สีพื้นหลังใบหน้า ความสว่าง แสงและเงา เป็นต้น โดยในการวิจัยช่วงแรกจะพิจารณาเป็นภาพเดี่ยว และพัฒนาให้สามารถตรวจจับหน้าคนจากภาพได้มากกว่าหนึ่งคนขึ้นไป

1.2 วัตถุประสงค์การวิจัย

เพื่อศึกษาและพัฒนาวิธีการตรวจจับใบหน้าบุคคลจากภาพสีด้วยเครือข่าย ART แบบดัดแปลง

1.3 ขอบเขตของงานวิจัย

1.3.1 พัฒนาระบบการตรวจจับใบหน้าบุคคลจากภาพสี โดยใช้วิธีการวิเคราะห์องค์ประกอบหลัก (principal component analysis : PCA) ในการสร้างเวกเตอร์ลักษณะเฉพาะร่วมกับโครงข่ายประสาทเทียมแบบ Adaptive Resonance Theory (ART) ซึ่งจะมีการปรับแต่งรายละเอียดของโครงข่าย ให้มีความเหมาะสมในการคัดแยกความเป็นหน้าบุคคลออกจากสิ่งแวดล้อมให้ดียิ่งขึ้น

1.3.2 พัฒนาระบบการตรวจจับใบหน้าบุคคลจากภาพสีให้สามารถตรวจจับภาพที่มีหน้าคนมากกว่า 1 หน้าได้

1.3.3 ระบบที่ได้มีอัตราการตรวจจับใบหน้าบุคคลด้วยความถูกต้องมากกว่า 80% ด้วยความเร็วที่เหมาะสมในการนำไปใช้งานจริงได้

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1 เป็นองค์ความรู้ในการวิจัยและเผยแพร่ต่อไป – สำหรับการวิจัยต่อยอดและตีพิมพ์ของหน่วยวิจัยทางด้านการมองเห็นของหุ่นยนต์และคอมพิวเตอร์ (robot and computer vision)

1.4.2 นำไปสู่การผลิตเชิงพาณิชย์ – สามารถนำไปพัฒนาเป็นระบบรักษาความปลอดภัย สำหรับหน่วยงานต่างๆ ไปได้



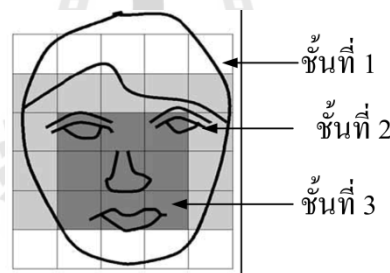
บทที่ 2

ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

2.1 การทบทวนวรรณกรรมที่เกี่ยวข้อง

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

2.1.1 วิธีเชิงความรู้ (Knowledge-based methods) เป็นวิธีการพิจารณาความสัมพันธ์ระหว่างลักษณะเด่นเฉพาะต่าง ๆ บนใบหน้า ที่มีตำแหน่งและองค์ประกอบหลักพื้นฐานตายตัวบนใบหน้า ตัวอย่างเช่น Yang and Huang (1994) ศึกษาวิธีเชิงความรู้แบบลำดับชั้น (hierarchical) ซึ่งพิจารณาแบ่งเป็น 3 ลำดับชั้น ดังแสดงในรูปที่ 2.1



รูปที่ 2.1 ขอบเขตการพิจารณาของแต่ละลำดับชั้น

ขั้นแรก พิจารณาหาตำแหน่งความน่าจะเป็นของใบหน้าโดยรวม ขั้นที่ 2 พิจารณาโดยค่า ฮิสโตแกรม (histogram) ร่วมกับการหาเส้นขอบ ขั้นที่ 3 พิจารณาลักษณะเด่นภายในของใบหน้า เช่น ความสมมาตรกันของตาทั้ง 2 ข้าง ตำแหน่งของจมูกและปากในแนวตั้งและแนวนอน

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

2.1.2 วิธีเชิงลักษณะ (Feature-based methods) เป็นการใช้อัลกอริทึมพิจารณาลักษณะเด่นและโครงสร้างของใบหน้า รวมทั้งความเปลี่ยนแปลงขององค์ประกอบต่าง ๆ ของภาพ เช่น แสงเงาตัวอย่างเช่น Chetverikov and Lerch (1993) ใช้ความแตกต่างของแสงเงาและเส้นแนวเป็นแบบจำลองในการค้นหา

ใบหน้า โดยแบบจำลองนั้นประกอบด้วยจุดความสว่างน้อย (จุดมืด) 2 จุด เพื่อแสดงดวงตา และจุดความสว่างมาก (จุดอ่อน) 3 จุด เพื่อแสดงโหนกแก้มและจมูก แล้วหาความสัมพันธ์ของระยะห่างและ ตำแหน่งของจุดต่าง ๆ เพื่อคัดเลือกแบบหน้าที่เหมาะสม ระบบนี้มีข้อจำกัดอยู่ที่เมื่อแสงเงาของสภาวะแวดล้อมเปลี่ยนแปลงจะทำให้ประสิทธิภาพในการค้นหาเปลี่ยนไป ต่อมา Graf, Chen, Petajan, and Cosatto (1995) พัฒนาทฤษฎีนี้โดยใช้งานในระบบภาพระดับเทา และนำตัวกรองช่วงผ่าน (band pass filter) ตัวกรองเกาส์เซียน (Gaussian filter) และค่าฮิสโทแกรม มาใช้หาค่าจุดสูงสุดและต่ำสุดเพื่อกำหนดขอบเขตของใบหน้า

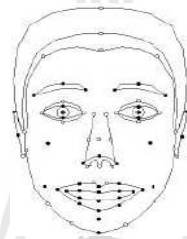
ลักษณะเฉพาะอีกประการหนึ่งที่นิยมนำมาพิจารณาหาใบหน้าบุคคล คือค่าสีผิวมนุษย์โดยมีการค้นหาในหลาย ๆ ปริภูมิสี โดย Crowley and Berard (1997) ทำการเก็บค่าเวกเตอร์ของสีผิวมนุษย์โดยแยกค่าสีออกเป็น R, G และ B เพื่อพิจารณาค่าฮิสโทแกรมในรูปของสมการ $h(r,g) \geq \alpha$ โดยค่า α คือค่าเริ่มเปลี่ยน (threshold) ของค่าสีผิวมนุษย์ที่ได้จากการสำรวจข้อมูลกลุ่มตัวอย่าง 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, Pramadhanto, and Yachida, (1996) ใช้ทฤษฎีคลุมเครือ (fuzzy theory) ในการคัดแยกสีผิวมนุษย์และสีผิวในปริภูมิสี CIE XYZ โดยทำการแปลงจากปริภูมิสี RGB นอกจากนี้ยังมีการนำเอาลักษณะเด่นของใบหน้าอื่น ๆ มาใช้ในการค้นหาตัวอย่างเช่น Yokoo and Hagiwara (1996) ใช้คุณสมบัติของโครงหน้าทั่วไปของมนุษย์ที่เป็นรูปไข่ สร้างแบบจำลองรูปวงรีเปรียบเทียบกับพารามิเตอร์ร่วมกับการค้นหาแบบจินเนติกอัลกอริทึม (genetic algorithm) จากการศึกษาพบว่าวิธีการนี้มีข้อจำกัดอยู่ที่รายละเอียดของใบหน้าทำการค้นหาต้องมีความชัดเจนและมีขนาดใหญ่เพียงพอต่อการค้นหา

2.1.3 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 เพื่อหาตำแหน่งความน่าจะเป็นของส่วนย่อยต่าง ๆ บนใบหน้า ที่สามารถเข้ากันได้ดีที่สุดกับแบบจำลองแผ่นแบบย่อยให้เป็นตำแหน่งของใบหน้าที่ต้องการต่อมา Tsukamoto, Lee, and Tsuji (1994) เสนอแบบจำลองคุณภาพสำหรับรูปแบบหน้า (qualitative model for face pattern : QMF) โดยใช้

พารามิเตอร์ของแสงสว่างและเส้นขอบใบหน้าเป็นแบบจำลองของใบหน้า และเทคนิคที่นิยมใช้อีกอันหนึ่งคือการวิเคราะห์องค์ประกอบหลัก (principal component analysis: PCA) โดย Jung, Lee C.W., and Lee Y.C (2002) ศึกษาการค้นหาคำแหน่งที่แน่นอนของใบหน้าโดยใช้หลักการพื้นฐานทางด้านเรขาคณิตวิเคราะห์องค์ประกอบของใบหน้าให้อยู่ในรูปของสมการทางคณิตศาสตร์

ต่อมาได้มีการพัฒนาแบบจำลองให้มีความสามารถปรับตัวได้เรียกว่า แผ่นแบบเปลี่ยนรูปได้ (deformable templates) โดยแผ่นแบบสามารถปรับตัวให้รองรับความยืดหยุ่นของโครงสร้างใบหน้าที่เปลี่ยนแปลงไปและลดปัญหาที่มีในแผ่นแบบคงตัว ตัวอย่างเช่น Malciu and Preteux (2002) นำแบบจำลองหน้าจากฐานข้อมูล MPEG-4 มาสร้างเป็นแบบจำลองแผ่นแบบพารามิเตอร์ (template parameterization) และพิจารณาความยืดหยุ่นของโครงสร้างใบหน้าเปรียบเทียบกับความยืดหยุ่นของสปริง เพื่อหาค่าพลังงานที่สปริงใช้น้อยที่สุดในการจะปรับตัวเพื่อให้ภาพที่หาเส้นขอบนั้นเข้ากับแบบจำลองมากที่สุด โดยทำการพิจารณาในส่วนของตาและปาก แสดงในรูปที่ 2.2

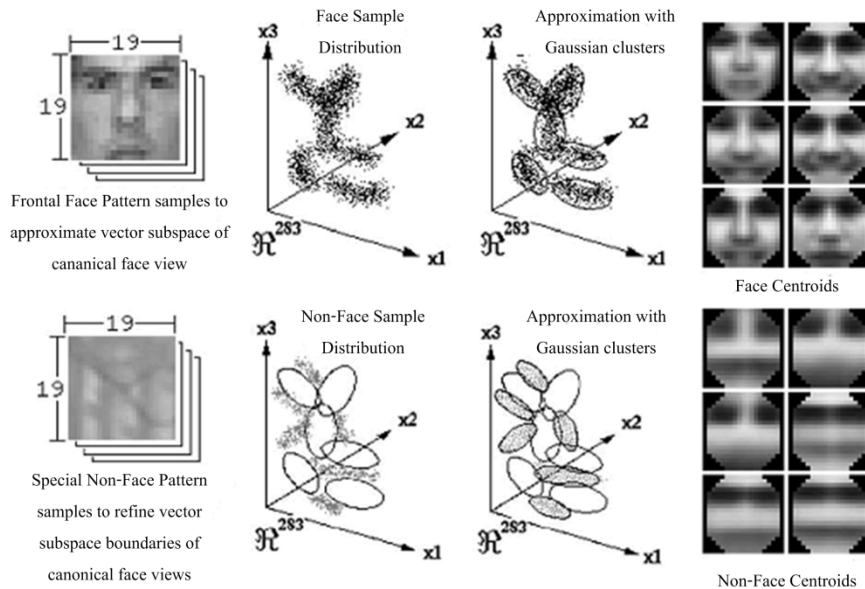


รูปที่ 2.2 แบบจำลองใบหน้ามนุษย์จากฐานข้อมูล MPEG-4

2.1.4 วิธีเชิงลักษณะปรากฏ (Appearance-based methods) เป็นการเปรียบเทียบภาพที่ต้องการค้นหาคำกับโครงสร้างแบบจำลองของใบหน้าทำการเรียนรู้และฝึกสอนให้ระบบจดจำ และนำความรู้ในฐานข้อมูลมาใช้ในการพิจารณา โดยสามารถแบ่งย่อยได้หลายวิธีด้วยกัน ในที่นี้จะทำการเสนอเพียงวิธีการที่ได้รับความนิยมใช้งานกันอย่างกว้างขวาง 4 วิธีดังนี้

2.1.4.1 วิธีหน้าลักษณะเฉพาะ (Eigenface Methods) เป็นการวิเคราะห์องค์ประกอบพื้นฐานของใบหน้าด้วยกระบวนการทางสถิติของใบหน้าที่หลากหลาย โดยหน้าลักษณะเฉพาะคือเซตของเวกเตอร์ลักษณะเฉพาะ (eigen vector) ที่สามารถหาได้จากเมทริกซ์ความแปรปรวนร่วมเกี่ยว (covariance matrix) สร้างเป็นแบบจำลองของใบหน้าที่รวมเอาลักษณะเด่นต่าง ๆ ของภาพใบหน้าตัวอย่างมารวมกันเพื่อหาค่าเฉพาะเจาะจงขององค์ประกอบบนใบหน้า ซึ่งในแต่ละบุคคลจะเป็นค่าเฉพาะของบุคคลนั้น ๆ ตัวอย่างเช่น Turk and Pentland (1991) นำภาพระดับเทามาแปลงเป็นเวกเตอร์เพื่อหาค่าลักษณะเฉพาะ และนำค่าลักษณะเฉพาะของตัวอย่างภาพหน้าบุคคลมาสร้างเป็นแบบจำลองหน้าลักษณะเฉพาะเพื่อค้นหาคำแหน่งของใบหน้า

2.1.4.2 วิธีเชิงการกระจาย (Distribution-Based Methods) เป็นการแสดงการกระจายตัวของรูปแบบข้อมูลตัวอย่างที่มีความเป็นหน้าและความไม่เป็นหน้าเพื่อใช้เป็นบรรทัดฐานในการตัดสินใจ ตัวอย่างเช่น Sung and Poggio (1998) นำฟังก์ชันเกาส์เซียน (Gaussian function) มาประมาณกลุ่มการกระจายของค่าเฉลี่ยกลุ่มตัวอย่าง แสดงในรูปที่ 2.3



รูปที่ 2.3 การกระจายตัวของข้อมูลตัวอย่าง

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

2.1.4.4 วิธีเวกเตอร์เกือหนุน (Support Vector Machines : SVMs) เป็นโครงข่ายแบบมีการฝึกสอนคล้ายคลึงกับโครงข่ายประสาทเทียม แต่มีหลักการที่แตกต่างกันตรงที่โครงข่ายประสาทเทียมนั้นจะทำการฝึกสอนเพื่อลดผลของค่าความผิดพลาด (empirical risk minimization) แต่ SVMs นั้นจะทำการฝึกสอนเพื่อลดของโครงสร้างความเสี่ยงให้น้อยที่สุด (structural risk minimization) และต้องการ ข้อมูล

ตัวอย่างที่มาก เพื่อลดความผิดพลาดของระบบ ตัวอย่างเช่น Osuna, Freund, and Girosi (1997) ทำการศึกษาเพื่อลดจำนวนในการฝึกสอน SVMs ให้มีปริมาณลดลงแต่ประสิทธิภาพในการทำงานไม่ลดลงไปด้วย

นอกจากนั้นยังมีการนำเอาวิธีการต่าง ๆ ข้างต้นมาใช้งานร่วมกัน โดยพัฒนาให้สามารถลดขอบเขตในการตรวจจับหรือเพิ่มประสิทธิภาพในการตรวจจับใบหน้าบุคคลให้ดียิ่งขึ้น ตัวอย่างเช่น Mitsukura, Fukumi and Akamatsu (2001) ใช้การค้นหาแบบจินเนติกอัลกอริทึมร่วมกับโครงข่ายประสาทเทียม LSDNN (Lip detection neural network and Skin Distinction Neural Network) โดยทำการค้นหาตำแหน่งของปากด้วย LSDNN และทำการค้นหาความสัมพันธ์ของระยะห่างระหว่างตำแหน่งของปากและสีผิวบริเวณโดยรอบด้วยการค้นหาแบบจินเนติกอัลกอริทึม และ Li, Gong, Sherrah, and Liddell (1999) นำเสนอวิธีเวกเตอร์เกือหนุ่ร่ว่กับปริภูมิลักษณะเฉพาะ (eigenspace) ในการตรวจจับใบหน้าบุคคล โดยใช้ปริภูมิลักษณะเฉพาะในการแบ่งขอบเขตความเป็นหน้าบุคคลออกจากสิ่งแวดล้อม และทำการพิจารณา ร่วมกับการตรวจจับรูปแบบใบหน้าด้วยวิธีเวกเตอร์เกือหนุ่

2.2 สรุป

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

บทที่ 3

ระบบการตรวจจับใบหน้าคนด้วยโครงข่าย ART

ระบบการตรวจจับใบหน้าคนด้วยโครงข่าย ART นี้ ประกอบไปด้วยระบบย่อย 3 ส่วนได้แก่ระบบคัดแยกสีผิวมนุษย์ออกจากพื้นหลัง ระบบแปลงข้อมูลภาพเป็นการวิเคราะห์ห้องค์ประกอบหลักและระบบตรวจจับใบหน้าคนด้วยโครงข่าย ART รายละเอียดต่างๆ มีดังต่อไปนี้

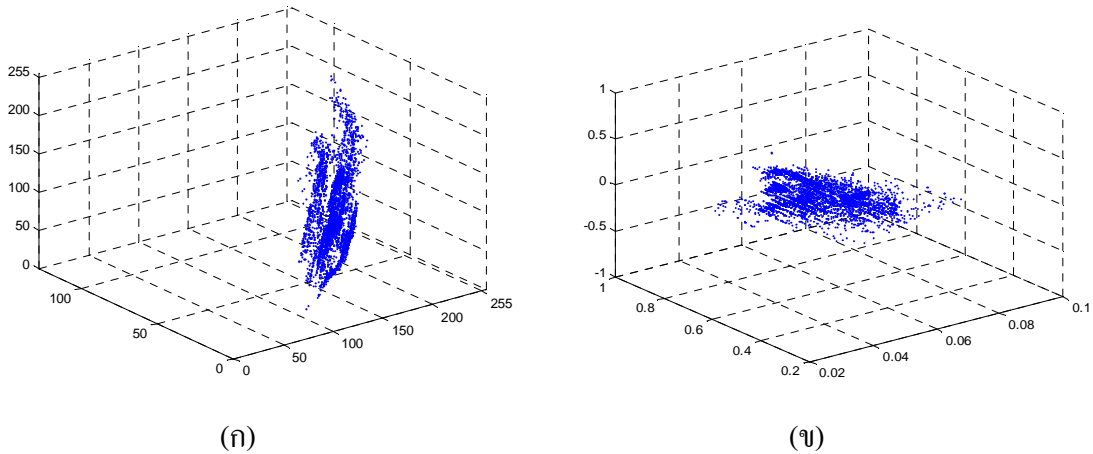
3.1 ระบบการคัดแยกสีผิวมนุษย์ออกจากพื้นหลัง

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

3.1.1 ค่าสีผิวมนุษย์ (Human Color) เป็นค่าที่สามารถคัดแยกออกจากสีของสิ่งแวดล้อมได้ โดยแปลงจากปริภูมิสี RGB เป็นปริภูมิสี HSV เพื่อแยกส่วนความเข้มแสงออกจากส่วนความเป็นสี เนื่องจากปริภูมิสี HSV สามารถกำจัดผลของความแตกต่างของค่าความเข้มแสงของสีผิวมนุษย์ และลดข้อจำกัดทางด้านแสงสว่างได้จากการตัดส่วนของค่า V ซึ่งเป็นค่าที่ใช้บอกระดับความสว่างของภาพออก จากนั้นตัดส่วนของภาพที่เป็นสีผิวมนุษย์มาเป็นอินพุตให้โครงข่าย โดยการเก็บตัวอย่างอินพุตจากภาพนั้นจะทำการเก็บตัวอย่างบริเวณใบหน้า และภาพใบหน้า 1 ภาพนั้นจะแบ่งบริเวณการเก็บตัวอย่างสีผิวมนุษย์ออกเป็น 5 ส่วนด้วยกัน เนื่องจากทั้ง 5 ส่วนนี้เป็นส่วนของสีผิวมนุษย์ที่ไม่มีสีอื่นที่ไม่ใช่สีผิวมนุษย์ผสมอยู่ด้วย และเป็นจุดหลักบนใบหน้าทำให้ง่ายในการเก็บข้อมูล โดยข้อมูลที่เก็บได้นั้นมีขนาดใหญ่เหมาะสม และในการเก็บข้อมูลถึง 5 จุดนั้น เพื่อให้ได้ตัวอย่างที่ครอบคลุมส่วนต่าง ๆ ของใบหน้ามากที่สุด ดังแสดงในรูปที่ 3.1 และค่าสีผิวมนุษย์ในปริภูมิสี RGB และ HSV แสดงในรูปที่ 3.2



รูปที่ 3.1 ตัวอย่างอินพุตของโครงข่ายประสาทเทียม



รูปที่ 3.2 ค่าสีผิวมนุษย์ (ก) ปริภูมิ RGB (ข) ปริภูมิ HSV

3.1.2 โครงข่ายประสาทเทียม (อาทิตย์ ศรีแก้ว, 2552) เป็นโครงข่ายที่สามารถปรับตัวเองต่อการตอบสนองของอินพุต ตามกฎการเรียนรู้ (learning rule) และเมื่อโครงข่ายได้เรียนรู้สิ่งที่ต้องการแล้ว โครงข่ายจะสามารถทำงานที่กำหนดไว้ได้ โดยโครงข่ายประสาทเทียมถูกพัฒนาคิดค้นมาจากหลักการทำงานของสมองมนุษย์ 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 และโครงข่ายประสาทเทียมมีความสามารถในการโต้ตอบ และตอบสนองต่อสถานะแวดล้อมได้ โดยเมื่อสถานะแวดล้อม

เปลี่ยนไป ตัวโครงข่ายจะสามารถตอบสนองกับการเปลี่ยนแปลงนั้น ๆ แล้วทำการฝึกฝนให้เข้ากับสภาวะแวดล้อมใหม่ได้

3.1.3 การคัดแยกสีผิวมนุษย์ด้วยโครงข่ายประสาทเทียมแบบแพร่กลับ โครงข่ายประสาทเทียมแบบแพร่กลับ เป็นโครงข่ายประสาทเทียมแบบหลายชั้น มีการเรียนรู้แบบมีผู้ฝึกสอน (supervised learning) คือ เรียนรู้ที่จะสร้างผลลัพธ์ที่ต้องการให้ได้ตามตัวอย่างที่ป้อนให้ และลดค่าความผิดพลาดของเอาต์พุตให้น้อยที่สุดโดยเทียบกับค่าน้ำหนักประสาท ซึ่งเป็นการวิเคราะห์หาค่าความผิดพลาดแบบกำลังสองเฉลี่ย (mean square error) ในอัลกอริทึมแบบแพร่กลับ มีการนำเสนอคู่อินพุตและเป้าหมายให้โครงข่ายเรียนรู้ แสดงในสมการที่ 2.9

$$\{p_1, t_1\}, \{p_2, t_2\}, \dots, \{p_Q, t_Q\} \quad (3.1)$$

ในการคัดแยกสีผิวมนุษย์ด้วยโครงข่ายประสาทเทียมแบบแพร่กลับจำเป็นต้องกำหนดคู่อินพุตและเป้าหมายที่ชัดเจน เพื่อให้โครงข่ายสามารถทำงานได้อย่างถูกต้องที่สุด โดยจะทำการเก็บตัวอย่างสีผิวมนุษย์จากที่ได้กล่าวไว้ในหัวข้อที่ 3.1.1 มาเป็นคู่อินพุตและเป้าหมายของโครงข่าย ซึ่งใช้ค่า H และค่า S ของภาพในช่วงที่เป็นสีผิวมนุษย์ระบุเป้าหมายว่าเป็นสีผิวมนุษย์ และค่า H และค่า S ของภาพ ในช่วงที่ไม่เป็นสีผิวมนุษย์ระบุเป้าหมายว่าไม่ใช่สีผิวมนุษย์มาฝึกสอนโครงข่าย ดังความสัมพันธ์ในสมการที่ 3.2

$$\begin{cases} P_{skin} = \begin{bmatrix} H_{skin} \\ S_{skin} \end{bmatrix}, t_{skin} = 1 \\ P_{nonskin} = \begin{bmatrix} H_{nonskin} \\ S_{nonskin} \end{bmatrix}, t_{nonskin} = 0 \end{cases} \quad (3.2)$$

โดยกำหนดตัวอย่างสีผิวมนุษย์ที่ใช้ฝึกสอนโครงข่ายจำนวน 100 ตัวอย่าง ขนาด 50x50 จุดภาพ จากจำนวนคน 20 คนซึ่งมีสีผิวแตกต่างกันออกไป และตัวอย่างภาพที่ไม่ใช่สีผิวมนุษย์จำนวน 100 ตัวอย่าง ขนาด 50x50 จุดภาพ ต่อมาทำการกำหนดตัวชี้วัดประสิทธิภาพ (performance index) เมื่อระบบมีการนำเสนอคู่อินพุตและเป้าหมายให้โครงข่ายเรียนรู้ ทำการป้อนแต่ละอินพุตให้กับโครงข่าย เอาต์พุตที่ได้จะถูกนำไปเปรียบเทียบกับเป้าหมาย เพื่อให้ค่าความผิดพลาดของเอาต์พุตและเป้าหมายมีค่าน้อยที่สุด ซึ่งค่าความผิดพลาดระหว่างเอาต์พุตและเป้าหมายนี้เองเป็นตัวชี้วัดประสิทธิภาพ โดยตัวชี้วัดประสิทธิภาพดังกล่าวนี้ คือค่าความผิดพลาดแบบกำลังสองเฉลี่ย (mean-square error) ของเอาต์พุตและเป้าหมาย

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

จากผลการทดลอง โครงข่ายที่ให้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยน้อยที่สุดคือ โครงข่าย 2-10-1 (ชั้นอินพุตมีเซลล์ประสาทเทียม 2 เซลล์ มีชั้นซ่อนเร้น 1 ชั้นมีเซลล์ประสาทเทียม 10 เซลล์ และชั้นเอาต์พุตมีเซลล์ประสาทเทียม 1 เซลล์) ใช้ฟังก์ชันถ่ายโอนแบบไฮเพอร์โบลิกแทนเจนต์ ซิกมอยด์ (Hyperbolic Tangent Sigmoid) และแบบลอค ซิกมอยด์ (Log-Sigmoid) ได้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยเท่ากับ 0.0408147 และโครงข่ายที่ให้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยน้อยรองลงมา คือ โครงข่าย 2-5-1 ใช้ฟังก์ชันถ่ายโอนแบบไฮเพอร์โบลิกแทนเจนต์ ซิกมอยด์ และแบบลอค ซิกมอยด์ ได้ค่าความผิดพลาดแบบกำลังสองเฉลี่ย 0.0430907 ดังนั้นจากผลการทดลอง จึงเลือกใช้โครงข่าย 2-5-1 ใช้ฟังก์ชันถ่ายโอนแบบไฮเพอร์โบลิกแทนเจนต์ ซิกมอยด์ เนื่องจากโครงข่าย 2-5-1 นั้นให้ค่าความผิดพลาดแบบกำลังสองเฉลี่ยน้อยกว่าไม่มากนักเมื่อเทียบกับโครงข่าย 2-10-1 แต่ใช้จำนวนเซลล์ประสาทเทียมน้อยกว่าโครงข่าย 2-10-1 ทำให้โครงข่ายมีขนาดเล็กกว่าโดยที่ประสิทธิภาพของโครงข่ายใกล้เคียงกัน โดยผลการคัดแยกสีผิวแสดงในรูปที่ 3.3 และสถาปัตยกรรมโครงข่ายประสาทเทียมแบบแพร่กลับในการแยกกลุ่มสีผิวมนุษย์แสดงในรูปที่ 3.4

ตารางที่ 3.1 ผลการฝึกสอน โครงข่ายประสาทเทียมเพื่อเลือกพารามิเตอร์ของโครงข่าย

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

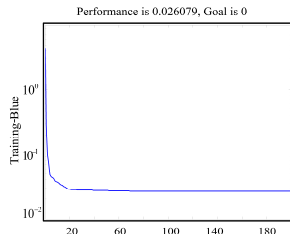


(ก)



(ข)

(ค)

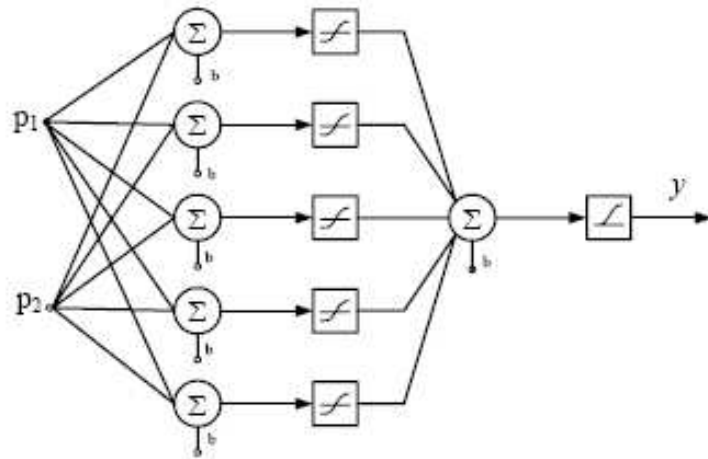


(ง)



(จ)

รูปที่ 3.3 ผลการคัดแยกสีผิวมนุษย์ออกจากสิ่งแวดล้อมด้วยโครงข่ายประสาทเทียมแบบแพร่กลับ (ก) ภาพต้นแบบ (ข) ผลการคัดแยกสีผิวมนุษย์ (ค) ขอบเขตของสีผิวมนุษย์ (ง) ผลการฝึกสอน โครงข่าย (จ) ขอบเขตการค้นหานำคน



รูปที่ 3.4 โครงสร้างของโครงข่ายประสาทเทียมแบบแพร่กลับในการแยกกลุ่มสีผิวมนุษย์

3.2 ระบบการแปลงข้อมูลภาพเป็นการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA)

วิธีการวิเคราะห์องค์ประกอบหลัก เป็นการแทนชุดข้อมูลใหม่ที่มีขนาดเล็กกว่าลงในชุดข้อมูลเดิม ซึ่งชุดข้อมูลใหม่นั้นจะมีความสัมพันธ์ระหว่างข้อมูลภายในชุดน้อยมาก โดยชุดข้อมูลเดิมที่มีความสัมพันธ์

ภายในข้อมูลขนาด 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 มาคำนวณหาค่าผลรวมความผิดพลาดน้อยที่สุด แสดงในสมการที่ 3.3

$$J_0(t_0) = \sum_{k=1}^M \|t_0 - t_k\|^2 \quad (3.3)$$

โดยค่า t_0 ที่ทำให้ J_0 มีค่าน้อยที่สุดหาได้จาก

$$\bar{t} = \frac{1}{M} \sum_{k=1}^M t_k \quad (3.4)$$

t คือค่าเฉลี่ยตัวอย่าง

$$\begin{aligned} J_{0(T_0)} &= \sum_{k=1}^M \|(T_0 - \bar{t}) - (t_k - \bar{t})\|^2 \\ &= \sum_{k=1}^M \|(T_0 - \bar{t})\|^2 - 2 \sum_{k=1}^M (T_0 - \bar{t})^T (t_k - \bar{t}) + \sum_{k=1}^M \|(t_k - \bar{t})\|^2 \\ &= \sum_{k=1}^M \|(T_0 - \bar{t})\|^2 - 2(T_0 - \bar{t})^T \sum_{k=1}^M (t_k - \bar{t}) + \sum_{k=1}^M \|(t_k - \bar{t})\|^2 \\ &= \sum_{k=1}^M \|(T_0 - \bar{t})\|^2 + \sum_{k=1}^M \|(t_k - \bar{t})\|^2 \end{aligned} \quad (3.5)$$

ทำการฉายข้อมูลลงบนเส้นค่าเฉลี่ยตัวอย่าง ได้สมการความสัมพันธ์ของตัวแปรใหม่ดังสมการที่ 3.6 โดยที่ \mathbf{u} คือเวกเตอร์ขนาด 1 หน่วย ที่มีทิศทางตามเส้นค่าเฉลี่ย และ y คือขนาดที่สัมพันธ์กันของระยะจาก t ถึง \bar{t}

$$\mathbf{t} = \bar{t} + y\mathbf{u} \quad (3.6)$$

พบว่าแกนฉายของข้อมูล 1 มิติ ก็คือไอเจนเวกเตอร์ที่สมนัยกับค่าไอเจนที่มากที่สุดตามลำดับ ดังนั้นในแนวทางเดียวกันเราสามารถขยายการแทนข้อมูลในรูปของ d มิติ โดยใช้การฉายข้อมูลลงเส้นข้อมูล d ดังแสดงในสมการที่ 3.7

$$\mathbf{t} = \bar{t} + y_d \mathbf{u}_d \quad (3.7)$$

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 3.8

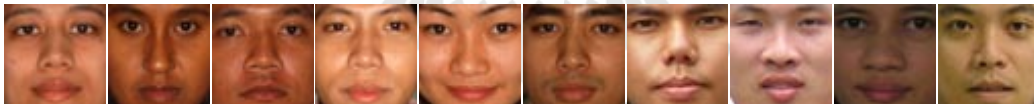
$$y_k = \mathbf{u}_1 t_1 + \mathbf{u}_2 t_2 + \dots + \mathbf{u}_{Dk} t_D = \mathbf{u}_k^T t \quad (3.8)$$

จากสมการที่ 3.8 สามารถเขียนสมการใหม่ได้โดยที่ y มีขนาด $d \times 1$ มิติ จากเวกเตอร์ t ขนาด $D \times 1$ มิติ และ U ขนาด $D \times d$ มิติ

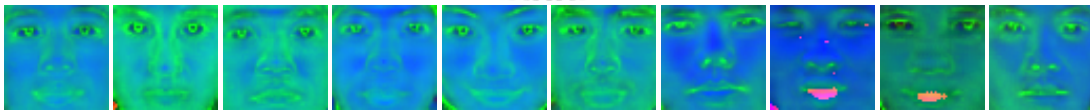
$$y = U^T t \quad (3.9)$$

จากสมการที่ 3.8 พบว่า u เป็นเมทริกซ์ที่ระบุแกนทิศทางขนาด 1 หน่วย และในสมการที่ 3.9 เมทริกซ์ U คือเมทริกซ์ความแปรปรวนร่วม และสามารถหาไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกนที่เรียงจากมากไปน้อยได้จากเมทริกซ์ความแปรปรวนร่วม U

3.2.1 การหาค่าไอเกนและไอเกนเวกเตอร์ ด้วยหลักการวิเคราะห์หองค์ประกอบ เป็นวิธีการหาหน้าไอเกน (Eigenface) ในปริภูมิหน้า (face Space) นำเสนอโดย 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



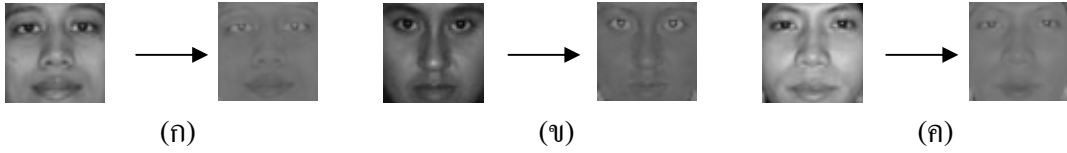
(ก)



(ข)

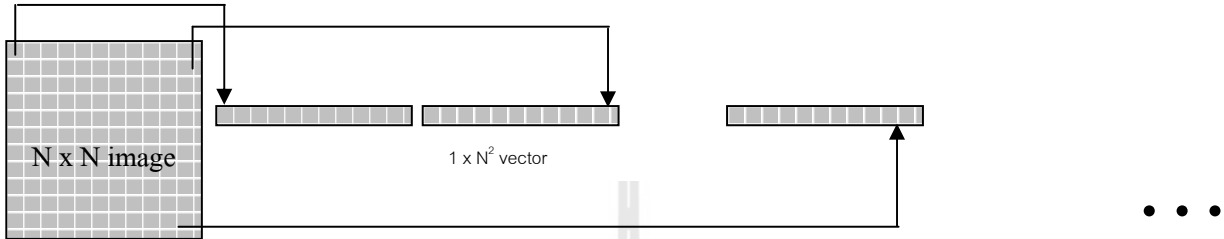
รูปที่ 3.5 ตัวอย่างภาพหน้าตัวอย่างในการหาหน้าไอเกน (ก) ตัวอย่างภาพในปริภูมิ RGB (ข) ตัวอย่างภาพในปริภูมิ HSV

โดยในรูปที่ 3.5 แสดงภาพตัวอย่างก่อนจะนำภาพมาทำการหาค่าไอเกนเวกเตอร์ซึ่ง จำเป็นต้องนำภาพในปริภูมิ RGB มาแปลงเป็นปริภูมิ HSV เพื่อลดผลของค่าความสว่าง ซึ่งมีผลต่อความถูกต้องของระบบการตรวจจับใบหน้า แล้วจึงแปลงภาพที่ทำการลดผลของค่าความสว่างแล้วนั้นเป็นปริภูมิ RGB อีกครั้ง และแปลงให้อยู่ในระดับเทาเพื่อให้เป็นภาพระนาบเดียว ดังแสดงในรูปที่ 3.6



รูปที่ 3.6 ภาพระดับเทาที่ได้จากการลดผลของความสว่าง (ก) สีผิวกลาง (ข) สีผิวเข้ม (ค) สีผิวอ่อน

แปลงเมทริกซ์ภาพตัวอย่างเป็นเวกเตอร์ขนาด $1 \times D$ ($D=mxn$) มิติ ดังแสดงในรูปที่ 3.7



รูปที่ 3.7 ตัวอย่างการแปลงเมทริกซ์ภาพเป็นเวกเตอร์ภาพ

จากนั้นนำเวกเตอร์ที่ได้มาจัดให้อยู่ในรูปของเมทริกซ์ ได้

$$\mathbf{I} = \{I_1, I_2, I_3, \dots, I_N\} \quad (3.10)$$

และนำเมทริกซ์ \mathbf{I} ที่ได้มาหาค่ากลางของชุดข้อมูลจากสมการที่ 3.11

$$\Phi_i = I_i - \frac{1}{N} \sum_{i=1}^N I_i \quad (3.11)$$

จะได้เวกเตอร์ค่ากลางของชุดข้อมูลในสมการที่ 3.12

$$\mathbf{A} = \{\Phi_1, \Phi_2, \Phi_3, \dots, \Phi_N\} \quad (3.12)$$

จากนั้นนำเมทริกซ์ \mathbf{A} ไปหาค่าเมทริกซ์ความแปรปรวน (covariance matrix: C) โดยเมทริกซ์ที่ได้มีขนาด $D \times D$ มิติ ดังแสดงในสมการที่ 3.13 และ 3.14

$$C = \frac{1}{N} \sum_{i=1}^N \Phi_i \Phi_i^T = \mathbf{A} \mathbf{A}^T \quad (3.13)$$

$$A v_i = u_i \quad (3.14)$$

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

จากค่าไอเกนเวกเตอร์ สามารถหาค่าหน้าไอเกน (Eigenface) ได้ โดยที่หน้าไอเกนค่า 1 นั้นจะแสดงลักษณะเฉพาะของหน้าได้ชัดเจนกว่าหน้าไอเกนค่า 2 ดังนั้น เราสามารถตัดหน้าไอเกนค่า 10 ออกได้โดยไม่

ทำให้เกิดความผิดพลาดกับระบบ เนื่องจากค่าไอเกนค่าที่ 10 นั้นแสดงคุณสมบัติของความเป็นหน้าไว้น้อยมากจนสามารถตัดค่านี้ทิ้งได้ โดยไม่ส่งผลกระทบต่อารแปลงเป็นหน้าไอเกน ดังนั้นเราจะสนใจหน้าไอเกน k ค่า โดยที่ $1 \leq k < N$ และแสดงการหาหน้าไอเกนในโดเมนหน้า (face space) จากสมการที่ 3.15

$$uu_k = (Av_k)^T \left(I_i - \frac{1}{N} \sum_{i=1}^N I_i \right) \quad (3.15)$$

ภาพต้นแบบจำนวน 10 ภาพสามารถหาภาพหน้าไอเกนได้ดังแสดงในรูปที่ 3.8



รูปที่ 3.8 ภาพหน้าไอเกน แถวที่ 1 ภาพต้นแบบในระดับเทา แถวที่ 2 ภาพต้นแบบในระดับเทาที่ทำการลดผลของความสว่างแล้ว แถวที่ 3 ภาพหน้าไอเกนจากไอเกนเวกเตอร์ค่าที่ 1 แถวที่ 4 ภาพหน้าไอเกนจากไอเกนเวกเตอร์ค่าที่ 10

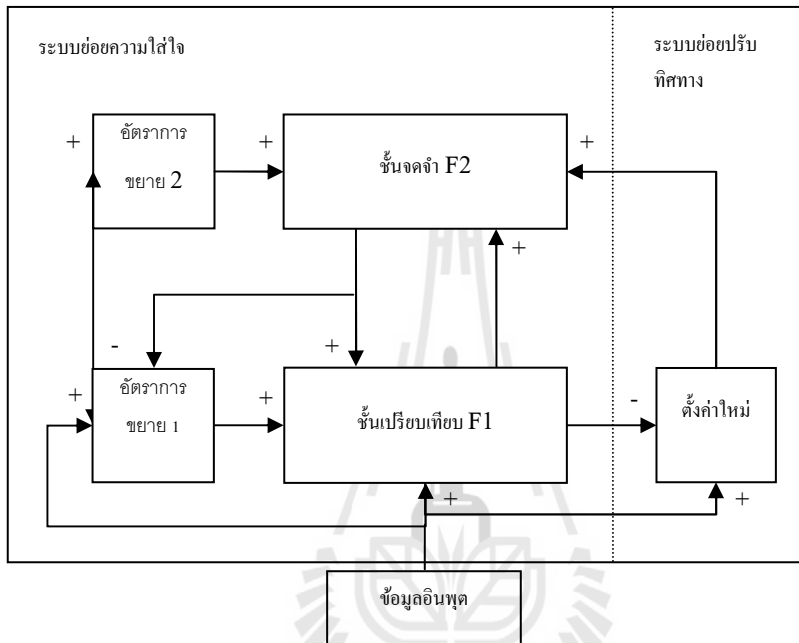
3.3 โครงข่าย ART (Adaptive Resonance Theory Network)

โครงข่าย ART เป็นโครงข่ายประสาทเทียมแบบไม่มีการฝึกสอน โดยเมื่อทำการป้อนค่าให้โครงข่ายโครงข่ายจะจดจำข้อมูลนั้นเป็นกลุ่มข้อมูลชุดที่ 1 เมื่อป้อนข้อมูลใหม่ที่มีค่าแตกต่างจากข้อมูลชุดแรก ระบบจะทำการเพิ่มกลุ่มข้อมูลใหม่เพื่อเก็บข้อมูลที่มีความแตกต่างกัน ซึ่งแตกต่างจากโครงข่ายประสาทเทียมแบบอื่นที่เมื่อทำการป้อนข้อมูลให้โครงข่ายด้วยข้อมูลชุดที่ 1 แล้ว ต้องการเพิ่มข้อมูลใหม่ให้โครงข่ายเรียนรู้เพิ่มเติม จำเป็นจะต้องทำการฝึกสอนโครงข่ายใหม่อีกครั้ง ด้วยข้อมูลชุดใหม่นั้น และทำให้โครงข่ายลืมความรู้เก่าที่ได้เรียนรู้มาจากข้อมูลชุดเดิมซึ่งปัญหานี้เรียกว่า ทวิปลเสถียรภาพสภาพพลาสติก (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)

3.3.1 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) ทำการปรับเทียบค่าน้ำหนักประสาทให้มีค่าสอดคล้องกับค่าอินพุตของระบบที่เข้ามา

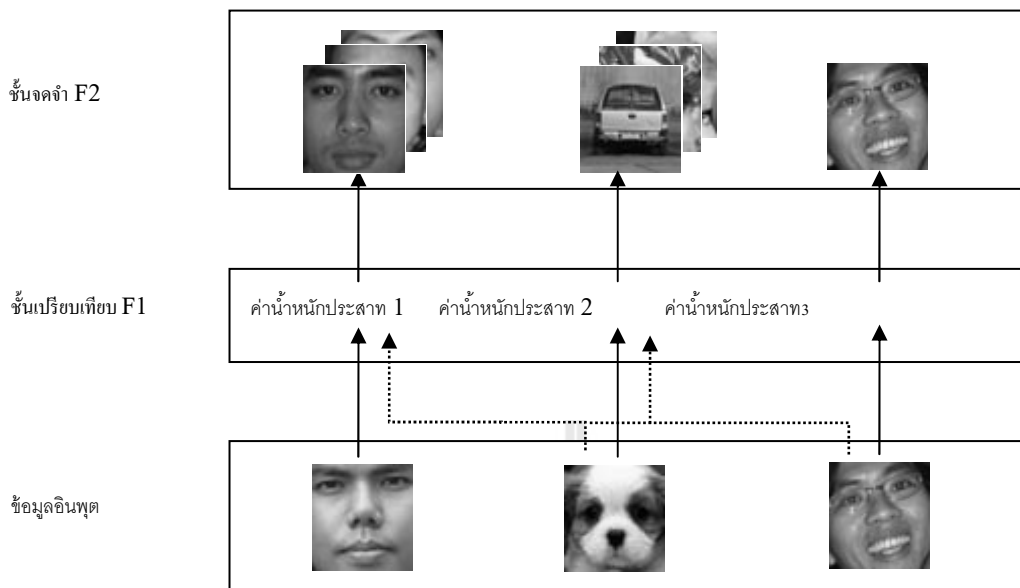
โครงสร้างของระบบทั้งสองแสดงในรูปที่ 3.9



รูปที่ 3.9 แผนผังการทำงานของโครงข่าย ART

โครงข่ายประสาทเทียม ART ที่ได้รับการฝึกสอนแล้วจะเก็บค่าข้อมูลอยู่ในรูปของค่าน้ำหนักประสาท (weight) โดยในชั้นเปรียบเทียบจะนำข้อมูลอินพุต (input data) ที่เข้ามานั้น มาพิจารณาเปรียบเทียบกับค่าน้ำหนักประสาทของกลุ่มข้อมูลที่อยู่ในชั้นจดจำ เรียกว่าค่าเทียบเคียง ถ้าข้อมูลอินพุตนั้นมีค่าสอดคล้องกับค่าน้ำหนักประสาทของกลุ่มใด ค่าเทียบเคียงที่ได้จากการเปรียบเทียบกับค่าน้ำหนักประสาทกลุ่มนั้นจะมีค่าน้อย และทำการคัดแยกกลุ่มข้อมูลอินพุตโดยการเปรียบเทียบค่าสอดคล้อง (vigilance: ρ) กับค่าเทียบเคียง ถ้าข้อมูลที่ได้มีค่าเทียบเคียงมากกว่าค่าสอดคล้อง ระบบจะทำการปรับค่าน้ำหนักประสาทของกลุ่มข้อมูลนั้น เพื่อทำการเก็บค่าข้อมูลใหม่ในกลุ่มที่มีค่าสอดคล้องกัน ถ้าข้อมูลที่ได้ มีค่าเทียบเคียงน้อยกว่าค่าสอดคล้อง แสดงว่าข้อมูลอินพุตนั้นไม่สอดคล้องกับค่าน้ำหนักประสาทในกลุ่มนั้น ระบบจะส่งค่าน้ำหนักประสาทของกลุ่มใหม่มาทำการเปรียบเทียบกับข้อมูลอินพุตต่อไป แต่ถ้าข้อมูลอินพุตนั้นไม่สอดคล้องกับ

กลุ่มข้อมูลในชั้นจดจำกลุ่มใดเลยระบบจะทำการสร้างกลุ่มใหม่และเก็บข้อมูลอินพุตนั้นไว้ในกลุ่มใหม่ ดังแสดงในรูปที่ 3.10



โดย — แสดงว่าค่าเทียบเคียงมากกว่าค่าสอดคล้อง
 แสดงว่าค่าเทียบเคียงน้อยกว่าค่าสอดคล้อง

รูปที่ 3.10 โครงข่ายประสาทเทียม ART

3.3.2 ประเภทของโครงข่าย ART สามารถแบ่งออกได้เป็น 4 ประเภท โดยจะแตกต่างกันไปตามข้อมูลที่รับเข้ามาและรูปแบบของการเรียนรู้ ได้แก่

3.3.2.1 โครงข่าย ART1 (Carpenter and Grossberg, 1987) เป็นระบบที่มีความซับซ้อนน้อยที่สุด โดยจะรับข้อมูลอินพุตที่มีค่าเป็น 0 และ 1 เท่านั้น โดยข้อมูลอินพุตจะส่งให้ชั้นเปรียบเทียบ F1 ตัดสินใจว่าข้อมูลอินพุตนั้นควรเป็นข้อมูลกลุ่มใด ด้วยสมการที่ 3.16 และ 3.17

$$Y_j = \sum_{i=1}^n (b_{ij} I_i) \quad (3.16)$$

$$Y_j = \text{Max}(Y_j) \quad (3.17)$$

โดย I คือ อินพุตของระบบ

b คือ ไบอัส ของระบบ

Y_j คือ ค่ามากที่สุดของกลุ่มที่ J

3.3.2.2 โครงข่าย ART2 (Carpenter and Grossberg, 1987) เป็นระบบที่ทำการพัฒนาต่อจาก ART1 ให้สามารถรับข้อมูลเชิงอุปมาน (analog data) จึงมีโครงสร้างของระบบที่ซับซ้อนมากกว่า ART1 โดยที่ ART2 นั้นจะมีโครงสร้างเหมือนกับ ART1 ยกเว้นในชั้นเปรียบเทียบ F1 จะทำการเพิ่ม ชั้นย่อย

(sublayers) เพื่อให้สามารถรับข้อมูลแบบเชิงอุปมานได้ โดยข้อมูลอินพุตเชิงอุปมานนั้นมีขนาดใหญ่และมีความแตกต่างของข้อมูลมากกว่าข้อมูลอินพุตที่มีค่าเป็น 0 และ 1 เท่านั้น ที่ป้อนให้โครงข่าย ART1 ดังนั้นชั้นย่อยที่เพิ่มมาในชั้นเปรียบเทียบกับ F1 จึงทำหน้าที่ย่อยข้อมูลขนาดใหญ่ขึ้นให้เล็กลงตามสมการที่ 3.18

$$f(x) = \begin{cases} x & \text{if } x \geq \theta \\ 0 & \text{if } x < \theta \end{cases} \quad (3.18)$$

โดยค่า θ เป็นค่าที่ผู้ใช้โครงข่ายกำหนดเลือกค่าตามความเหมาะสมของข้อมูลอินพุต

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

$$Y_j(I) = \frac{|I \cap w_j|}{\alpha + |w_j|} \quad (3.19)$$

$$Y(I)_j = \text{Max}(Y_j) \quad (3.20)$$

เมื่อ \cap ในสมการที่ 3.19 คือตรรกศาสตร์ส่วนร่วมรวม (logical AND intersection) เป็นสัญลักษณ์ทางคณิตศาสตร์แสดงการรวมกันของข้อมูลอินพุตและค่าน้ำหนักประสาทที่มีค่าเหมือนกัน โดยกำหนดให้ค่า α มีค่าน้อยมากๆ และ Y_j คือกลุ่มที่มีความสอดคล้องกับอินพุตมากที่สุด

3.3.2.4 โครงข่าย Fuzzy ART หรือ Fuzzy ARTMAP (Carpenter, Grossberg, and Rosen, 1991) เป็นระบบที่ได้รับการพัฒนาจากโครงข่าย ART และโครงข่าย ARTMAP โดยนำตรรกศาสตร์คลุมเครือ (fuzzy logic) มารวมกับโครงข่าย ดังนั้นในการเปรียบเทียบข้อมูลอินพุตกับค่าน้ำหนักประสาทจึงใช้กฎตรรกศาสตร์มาพิจารณาด้วย ดังสมการที่ 3.21 และ 3.22

$$T_j(I) = \frac{|IAw_j|}{\alpha + |w_j|} \quad (3.21)$$

$$T(I)_j = \text{Max}(T_j) \quad (3.22)$$

เมื่อ A ในสมการที่ 3.21 คือตรรกศาสตร์คลุมเครือค่าต่ำสุดรวม (fuzzy AND minimum) เป็นสัญลักษณ์ทางคณิตศาสตร์แสดงการรวมกันของข้อมูลอินพุตและค่าน้ำหนักประสาทโดยพิจารณาค่าที่ต่ำที่สุดในการรวมกัน และถ้าค่า T_j มีค่ามากที่สุดมากกว่า 1 ค่า ระบบจะทำการเลือกเอาต์พุตของข้อมูลกลุ่มแรกเป็นกลุ่มที่เหมาะสมที่สุดของอินพุตนั้น

3.4 ระบบตรวจจับหน้าคนจากภาพด้วยโครงข่าย ART

ค่าไอเกนเวกเตอร์ที่คำนวณได้จากข้างต้น จะถูกนำมาเป็นอินพุตของโครงข่าย ART ซึ่งในที่นี้พิจารณาใช้โครงข่าย ART แบบ ARTMAP เนื่องจากเป็นโครงข่ายที่พัฒนามาจาก ART1 และ ART2 เพื่อให้สามารถรับข้อมูลอินพุตที่มีความซับซ้อนมากขึ้น ซึ่งหลักการในการคัดแยกกลุ่มข้อมูลของโครงข่ายทำให้สามารถแยกข้อมูลที่มีความซับซ้อนมากได้ชัดเจนมากกว่า ART1 และ ART2 ส่วนโครงข่ายแบบ Fuzzy ART หรือ Fuzzy ARTMAP นั้น ใช้หลักการทฤษฎีตรรกศาสตร์คลุมเครือมาช่วยในการคัดแยกกลุ่มข้อมูล ซึ่งมีโครงสร้างที่ซับซ้อนมากขึ้นไป ในการคัดแยกกลุ่มข้อมูลที่ต้องการแยกเพียงความเป็นหน้ามนุษย์และความไม่เป็นหน้ามนุษย์เท่านั้น

3.4.1 โครงสร้างของระบบ ป้อนชุดข้อมูลตัวอย่างเพื่อกำหนดกลุ่มให้โครงข่าย ART ด้วยค่าไอเกนเวกเตอร์ที่ได้จากใบหน้าตัวอย่าง โดยใช้ใบหน้าตัวอย่างจำนวน 20 หน้า ทำให้ได้ไอเกนเวกเตอร์ที่เป็นชุดข้อมูลตัวอย่างของหน้าคนเป็นกลุ่มข้อมูล กลุ่มที่ 1 ของโครงข่ายมีขนาด 20×20 มิติ และในทำนองเดียวกันทำการป้อนข้อมูลตัวอย่างที่ไม่ใช่ใบหน้าคนให้โครงข่าย ART ทำให้ชั้นเปรียบเทียบในโครงข่าย ART มีข้อมูลจดจำอยู่ 2 กลุ่ม คือ กลุ่มหน้าคนและกลุ่มที่ไม่ใช่หน้าคน โดยในการเปรียบเทียบขั้นแรกนี้เป็นการเปรียบเทียบเพื่อตรวจจับหน้าคน โดยเปรียบเทียบกับค่าน้ำหนักประสาทของกลุ่มข้อมูลที่เก็บค่าในชั้นจดจำ 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 $\rho_1 = \mu$

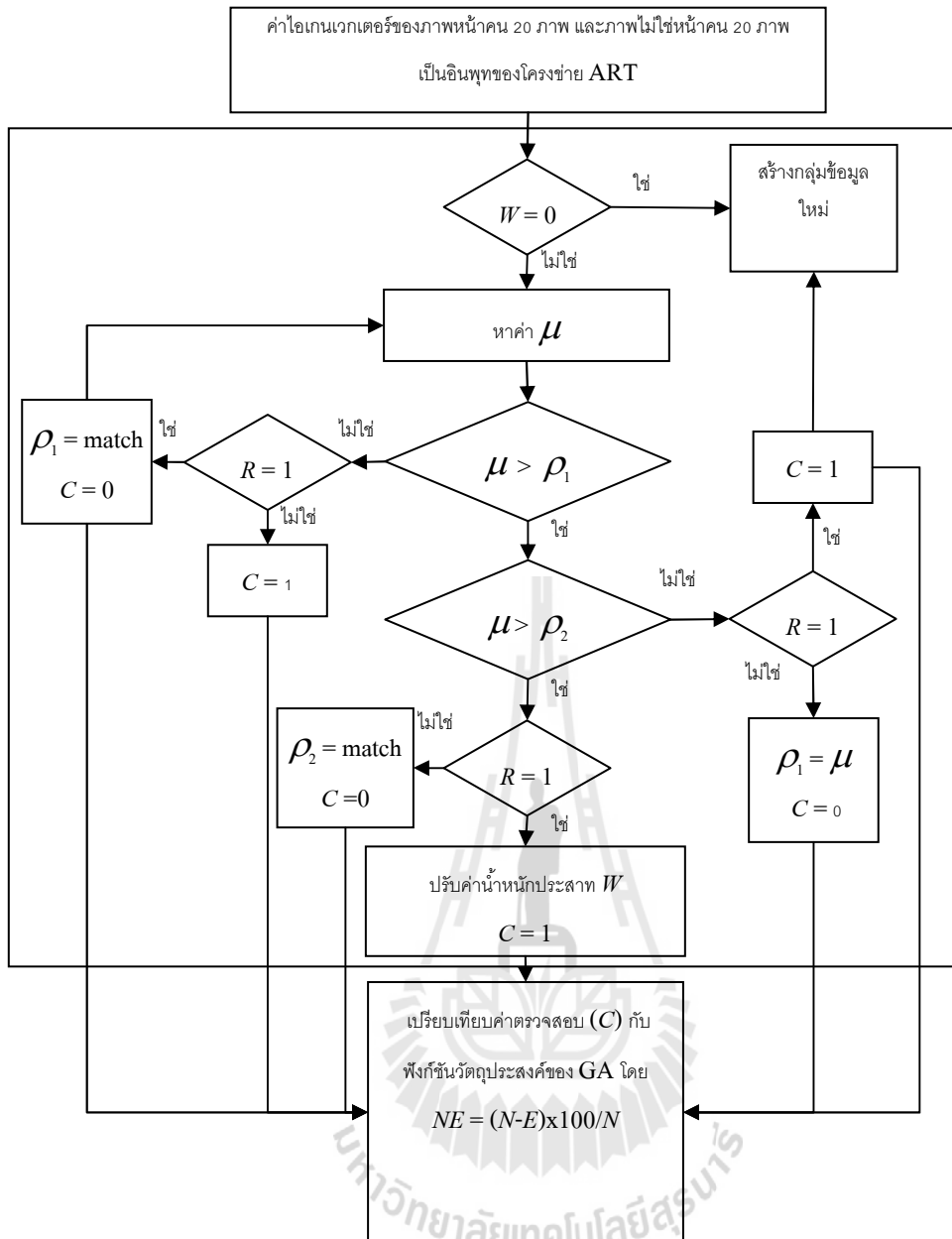
และถ้าข้อมูลตัวอย่างที่ป้อนให้โครงข่าย ART เป็นข้อมูลที่ไม่ใช่หน้าคน จะสามารถพิจารณาได้ 3 กรณีเช่นกันคือ

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

ในงานวิจัยนี้ ได้เพิ่มกลไกการกำหนดค่าพารามิเตอร์สอดคล้องที่เหมาะสม โดยใช้จินเนติกอัลกอริทึม (Genetic Algorithm) ทำการค้นหาค่าเหมาะที่สุด (optimization) แผนผังการทำงานของโครงข่าย ART ในการหาค่าสอดคล้องที่เหมาะสมด้วย GA มีแสดงในรูปที่ 3.11

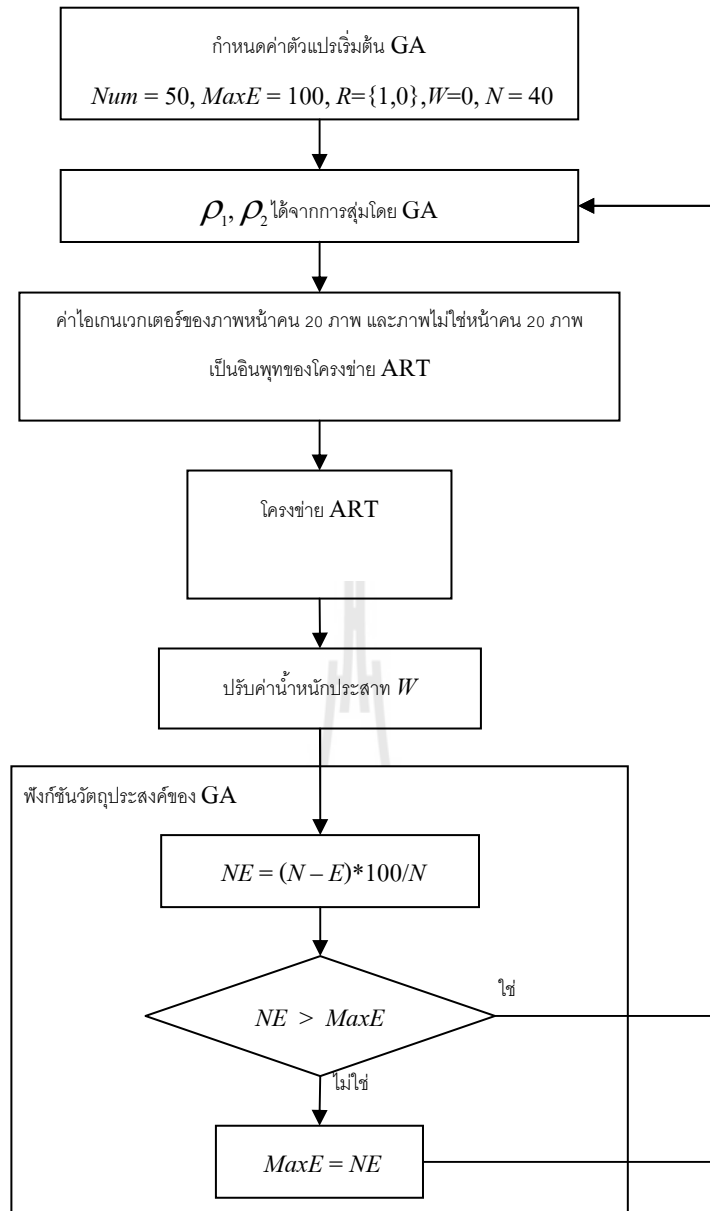
เนื่องจากชุดข้อมูลที่เป็นหน้าคนนั้น เป็นข้อมูลที่เกี่ยวข้องกับองค์ประกอบของหน้าในภาพที่มีความหลากหลายของหน้าแตกต่างกันตาม ลักษณะการวางท่า อารมณ์ของหน้า ขนาดของหน้า เป็นต้น ทำให้การหาค่าสอดคล้องที่เหมาะสมทำได้ยาก ดังนั้นจึงนำ GA มาช่วยในการหาค่าสอดคล้องที่เหมาะสมที่สุดในการแบ่งกลุ่มข้อมูลของโครงข่าย ART โดยทำการสร้างฟังก์ชันวัตถุประสงค์ให้ทำการแยกข้อมูลของหน้าคนจำนวน 20 หน้า และข้อมูลที่ไม่ใช่หน้าคนอีก 20 หน้าผ่านโครงข่าย ART โดยกำหนดค่าพารามิเตอร์เริ่มต้นของ GA ได้แก่จำนวนประชากร (population) จำนวน 50 ประชากร ($Num = 50$) ค่าความผิดพลาดเริ่มต้น เท่ากับ 100 ($MaxE = 100$) ค่าต้นแบบ (R) เป็นเซตแสดงสถานะของอินพุตว่าเป็นหน้าคนหรือไม่ใช่หน้าคน ในที่นี้ใช้ตัวอย่างที่เป็นหน้าคน 20 ตัวอย่าง และไม่ใช่หน้าคน 20 ตัวอย่าง ดังนั้นจะได้ค่า R เป็นเซตของตัวอย่างหน้าคน และตัวอย่างที่ไม่ใช่หน้าคน โดยที่ A มีค่าเป็น 1 และ B มีค่าเป็น 0 ค่าน้ำหนักประสาทเริ่มต้นเป็น 0 ($W = 0$) และจำนวนภาพตัวอย่างในการทดลอง 40 ตัวอย่าง ($N = 40$) จากนั้น ทำการหาค่าตัวแปรใน GA ค่าสอดคล้องเริ่มต้น ρ_1 และ ρ_2 โดยการสุ่มเลือกอิสระ และส่งค่าสอดคล้องที่ได้จากการสุ่มให้โครงข่าย ART





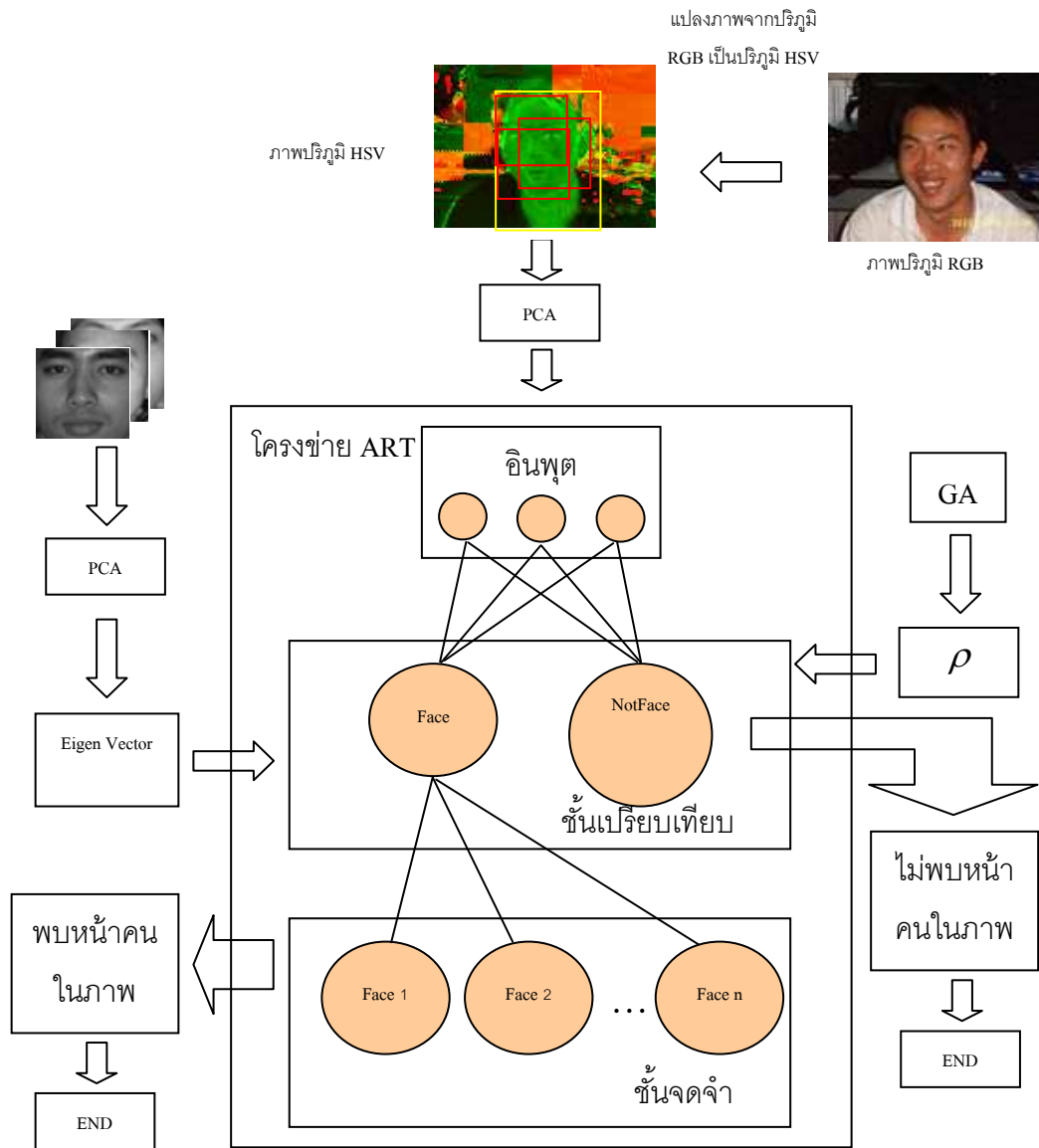
รูปที่ 3.11 แผนผังการทำงานของโครงข่าย ART ในการหาค่าสอดคล้องที่เหมาะสมด้วย GA

เมื่อได้ค่าตรวจสอบแล้วจะส่งค่าที่ได้ไปยังวัตถุประสงค์ของ GA เพื่อหาค่าความผิดพลาด (Net error: NE) โดย GA จะทำการปรับปรุงพันธุกรรมของประชากรเพื่อให้ได้ตัวแปรที่มีค่าความผิดพลาดที่น้อยที่สุด และสิ้นสุดการทำงานเมื่อ GA ครบจำนวนรอบสูงสุดที่กำหนดไว้ แสดงแผนผังการหาค่าสอดคล้องที่เหมาะสมโดย GA ในรูปที่ 3.12



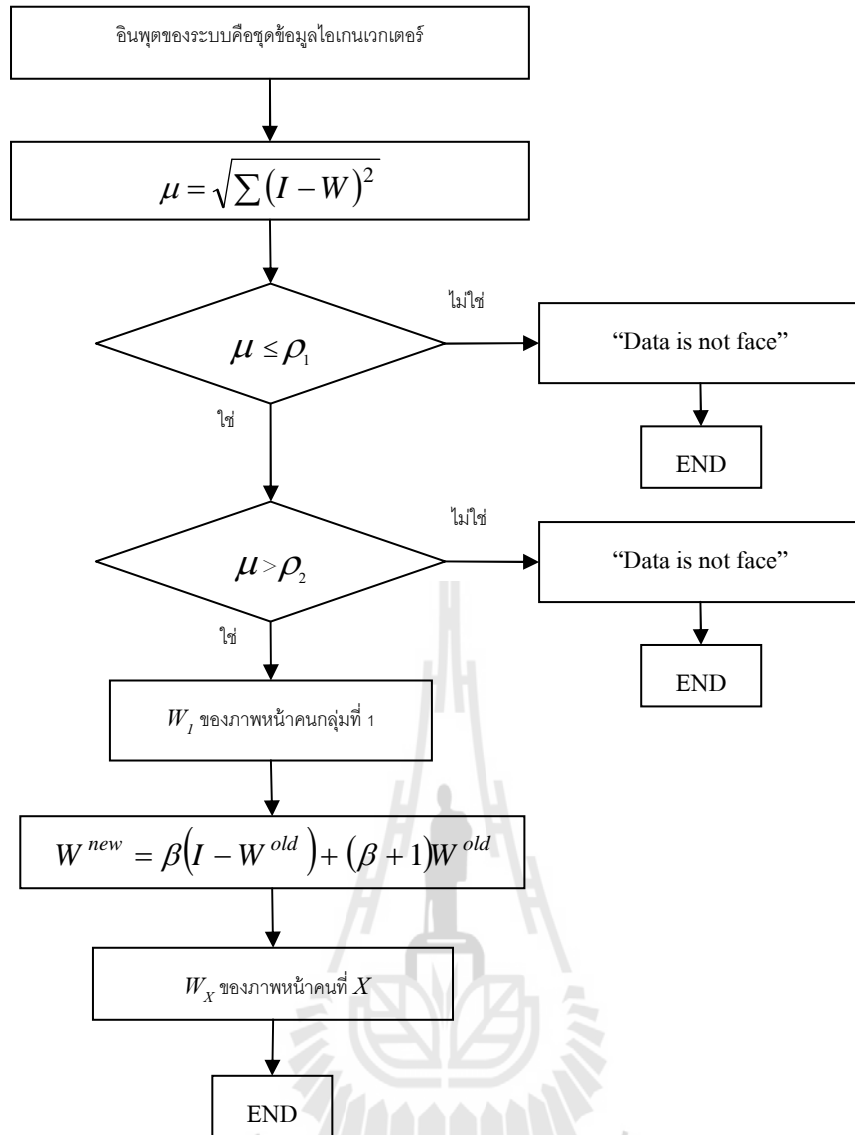
รูปที่ 3.12 แผนผังการหาค่าสอดคล้องที่เหมาะสมโดย GA

เมื่อได้ค่าสอดคล้องที่เหมาะสมกับระบบแล้ว นำค่าที่ได้นั้นมากำหนดเป็นค่าสอดคล้องของโครงข่าย ART โดยโครงสร้างของระบบการตรวจจับหน้าคนด้วยโครงข่าย ART มีโครงสร้างดังแสดงในรูปที่ 3.13



รูปที่ 3.13 โครงสร้างระบบการตรวจจับหน้าคนด้วยโครงข่าย ART

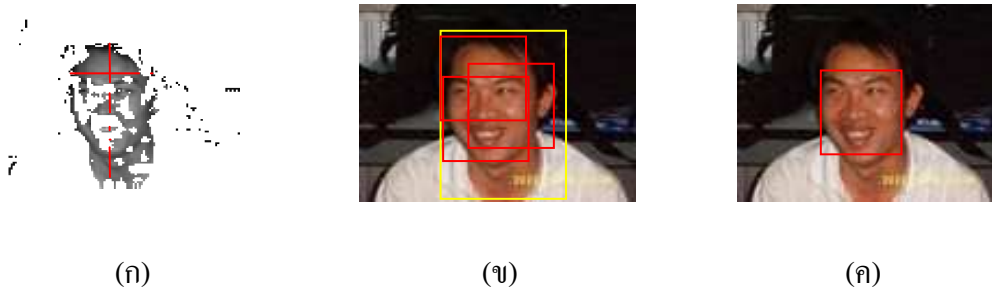
จากนั้นนำค่าสอดคล้องที่ได้การกระบวนการ GA มาใช้ในการเปรียบเทียบกับค่าเทียบเคียงที่ได้จากการคำนวณน้ำหนักประสาทของกลุ่มข้อมูลที่เหมาะสมที่สุด และในขั้นที่ 2 จะทำการเปรียบเทียบเพื่อแยกประเภทของหน้าที่มีลักษณะต่างกัน โดยแสดงโครงสร้างของโครงข่าย ART ในรูปที่ 3.14



รูปที่ 3.14 แผนผังโครงข่าย ART ในการคัดแยกหน้าคน

นำภาพที่ต้องการตรวจสอบมาแปลงจากปริภูมิสี RGB เป็นปริภูมิสี HSV จากนั้นมาทำการคัดแยกสีผิวมนุษย์ เพื่อกำหนดขอบเขตในการค้นหาให้แคบลง ได้ขอบเขตในการค้นหาที่คาดว่าน่าจะเป็นหน้าคน เนื่องจากมีค่าอยู่ในช่วงของสีผิวมนุษย์นั้น ต่อมาทำการหาขนาดของกล่องข้อมูลที่จะเข้าโครงข่าย ART ในรูปที่ 3.15 (ข) ในกรอบสีเหลืองคือขอบเขตทั้งหมดที่มีค่าสีผิวมนุษย์ที่ได้จากการคัดแยกสีผิวมนุษย์อยู่ และกรอบสีแดงคือกล่องข้อมูลที่จะป้อนเข้าโครงข่าย ART โดยสามารถหาได้จาก การหาค่าตำแหน่งของค่าจุดสีของค่าสีผิวมนุษย์ที่มีค่ามากที่สุดทั้งในแนวนตั้ง และแนวนอน เพื่อหาค่าที่น้อยที่สุดของทั้ง 2 แนวเป็นความยาวของกล่องข้อมูล แสดงในรูปที่ 3.15 (ก) เมื่อได้ขนาดของกล่องข้อมูลแล้ว ทำการหาค่าตำแหน่งของกล่องภาพที่จะเป็นอินพุตป้อนให้โครงข่าย ART โดยทำการเลือกกล่องข้อมูลที่มีค่าจุดสีที่เป็นสีผิวมนุษย์ในกล่องข้อมูลไม่น้อยกว่า 2 ใน 3 ของกล่องข้อมูลที่มีค่าจุดสีที่เป็นสีผิวมนุษย์ในกล่องมากที่สุด เป็นข้อมูลอินพุต

ป้อนเข้าโครงข่าย ART โดยทำการเปรียบเทียบกล่องข้อมูลอินพุตที่จะเข้าโครงข่ายกับข้อมูลหน้าในชั้นเปรียบเทียบชั้นแรกเพื่อหาตำแหน่งที่เหมาะสมที่สุดที่มีหน้าคนปรากฏอยู่ดังแสดงในรูปที่ 3.15 (ก)



รูปที่ 3.15 (ก) เส้นกำหนดขนาดกล่องข้อมูล (ข) กล่องข้อมูลอินพุตที่ป้อนเข้าโครงข่าย ART
(ค) กล่องข้อมูลที่ดีที่สุดที่เป็นตำแหน่งของหน้าคนในภาพ

3.4.2 หลักการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART ทำการสร้างโครงข่ายประสาทเทียม แปลงข้อมูลอินพุตของค่าไอเกนเวกเตอร์ในสมการที่ 3.23 ให้อยู่ในรูปของข้อมูลส่วนเติมเต็ม (complement data) แสดงในสมการที่ 3.24

$$v_1, v_2, \dots, v_M \quad (3.23)$$

$$I = (v, v^c) = (v_1, v_2, \dots, v_M, 1 - v_1, 1 - v_2, \dots, 1 - v_M) \quad (3.24)$$

โดย I คือข้อมูลส่วนเติมเต็ม

V คือไอเกนเวกเตอร์

ฝึกสอนโครงข่าย ด้วยกฎการเรียนรู้แบบแข่งขันแสดงในสมการที่ 3.25

$$W_i^{new} = \begin{cases} W_i^{old} & \text{if } \mu \geq \rho \\ \beta(I - W_i^{old}) + (1 + \beta)W_i^{old} & \text{if } \mu < \rho \end{cases} \quad (3.25)$$

โดย W^{new} คือค่าน้ำหนักประสาทใหม่

W^{old} คือค่าน้ำหนักประสาทเดิม

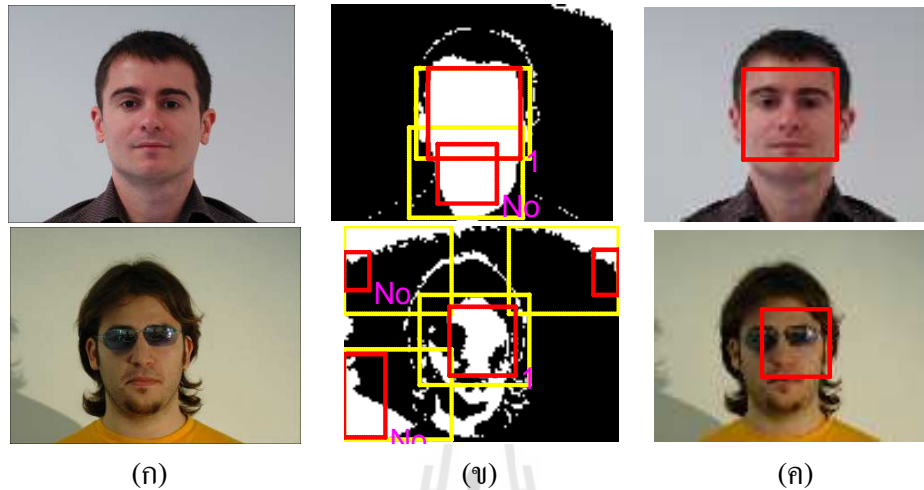
β คือค่าอัตราการเรียนรู้มีค่าอยู่ระหว่าง 0 ถึง 1

โครงข่าย ART จะทำการคัดแยกกลุ่มข้อมูลออกเป็น 2 กลุ่มข้อมูลคือ กลุ่มที่ 1 เป็นข้อมูลของภาพใบหน้าบุคคล และกลุ่มที่ 2 เป็นข้อมูลของภาพที่ไม่ใช่ใบหน้าบุคคล โดยพิจารณาคัดแยกด้วยค่าเทียบเคียง (μ) ดังแสดงในสมการที่ 3.26 เปรียบเทียบกับค่าสอดคล้อง (ρ)

$$\mu = \sqrt{(I - w_i^{old})^2} \quad (3.26)$$

ถ้าค่าเทียบเคียงที่ได้มีค่าน้อยกว่าค่าสอดคล้อง แสดงว่าข้อมูลอินพุตกับค่าน้ำหนักประสาทที่ใช้ไม่ได้อยู่กลุ่มเดียวกัน ระบบจะทำการเปรียบเทียบข้อมูลอินพุตกับค่าน้ำหนักประสาทของกลุ่มใหม่ จนกว่าจะได้กลุ่มที่ตรงกัน และถ้าข้อมูลอินพุตที่ได้นี้ ไม่ตรงกับกลุ่มข้อมูลของระบบกลุ่มใดเลย ระบบจะทำการจดจำข้อมูลอินพุตนี้ในกลุ่มที่สร้างขึ้นใหม่ต่อไป

3.4.3 ตัวอย่างการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART เมื่อกำหนดขอบเขตการค้นหาค้นหาโดยการพิจารณาบริเวณที่เป็นสีผิวมนุษย์เท่านั้น ระบบจะทำการสร้างกล่องข้อมูลอินพุต และนำกล่องข้อมูลนั้นไปแปลงเป็น ไอเกนเวกเตอร์เพื่อเป็นอินพุตให้โครงข่าย ART เพื่อคัดแยกความเป็นหน้าคน และไม่ใช่น้ำคนโดยเปรียบเทียบค่าเทียบเคียงของข้อมูลกับค่าสอดคล้องของระบบ แสดงได้ในรูปที่ 3.16



รูปที่ 3.16 ภาพตัวอย่างการทำงานของระบบตรวจจับหน้าคนด้วยโครงข่าย ART

(ก) ตัวอย่างภาพต้นแบบ (ข) ตัวอย่างภาพผลการทำงานของโครงข่าย ART

(ค) ตัวอย่างภาพผลการตรวจจับหน้าคนด้วยโครงข่าย ART

จากรูปที่ 3.16 (ข) พบว่าระบบจะทำการสร้างกล่องข้อมูลอินพุตที่มีค่าสีผิวมนุษย์มาก 1 ใน 4 ของพื้นที่กล่องทั้งหมดเป็นอินพุตของระบบ แสดงในกรอบภาพสี่เหลี่ยม และจุดภาพสีขาวแสดงถึงค่าจุดภาพนั้นมีความน่าจะเป็นที่จะเป็นสีผิวมนุษย์ โดยกล่องข้อมูลนี้อาจมีได้หลายอันขึ้นอยู่กับจำนวนจุดภาพที่เป็นสีผิวมนุษย์ในภาพนั้นมีค่าน้อยเพียงใด ต่อจากนั้นระบบจะทำการสร้างกล่องข้อมูลเพื่อเป็นอินพุตให้โครงข่าย แสดงในกรอบภาพสีแดงเพื่อหาจุดที่มีความน่าจะเป็นว่าเป็นหน้าคนภายในกล่องภาพแต่ละกล่องแล้วแสดงผลการคัดแยกหน้าคน โดยระบบจะแสดงตัวอักษร “No” บนกล่องภาพที่ไม่ใช่น้ำคน และแสดงตัวเลขบนกล่องภาพที่เป็นหน้าคน โดยถ้าในภาพนั้นระบบตรวจพบหน้าคนจำนวน 3 คน ระบบจะแสดงตัวเลขบนกล่องที่เป็นหน้าคนเรียงกันไปจาก 1 ถึง 3

บทที่ 4

ผลการทดสอบระบบตรวจจับหน้าคนด้วยโครงข่าย ART

4.1 กล่าวนำ

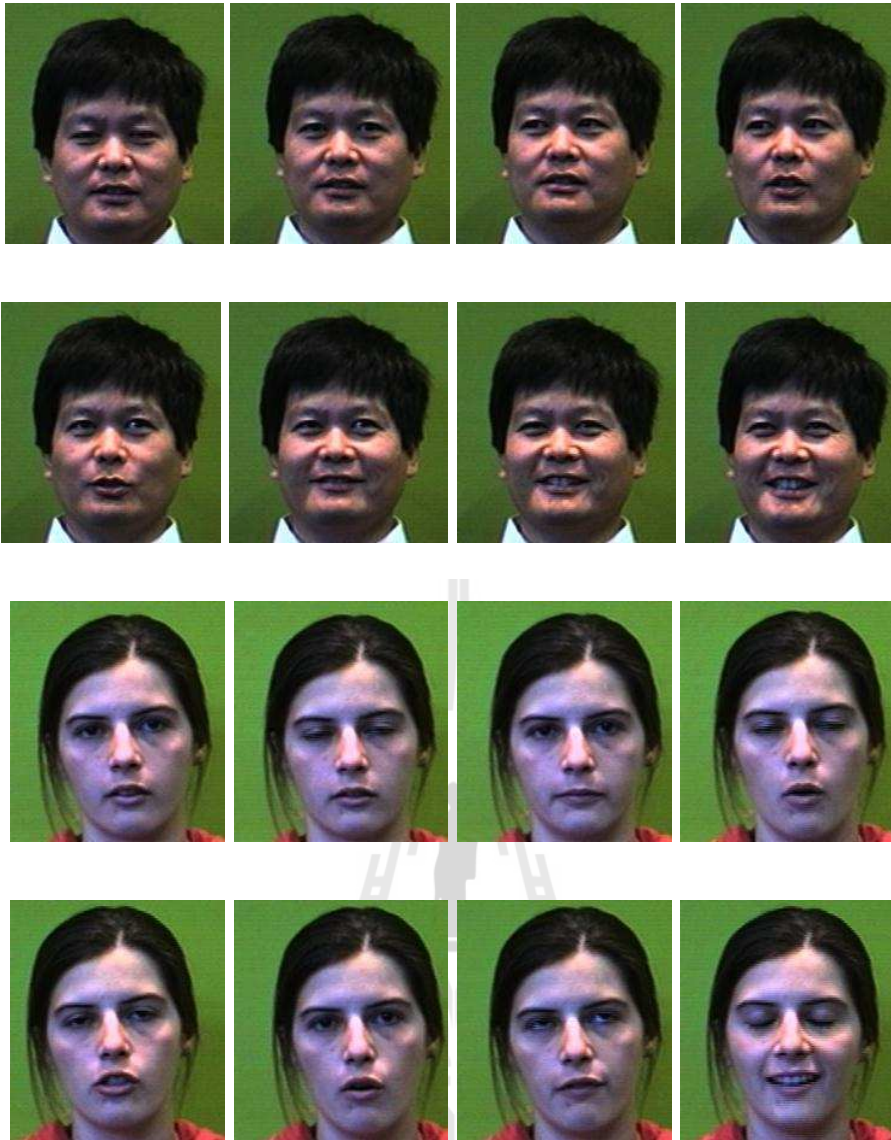
เมื่อทำการฝึกสอนโครงข่าย ART ด้วยภาพหน้าคน และภาพที่ไม่ใช่หน้าคนแล้ว ทำการทดสอบระบบด้วยภาพจากฐานข้อมูลต่าง ๆ เพื่อวัดประสิทธิภาพการทำงานของระบบว่ามีความสามารถในการตรวจจับหน้าคนจากภาพได้ดีเพียงไร โดยภาพที่ใช้ในการทดสอบนั้น ได้จากฐานข้อมูลภาพต่าง ๆ และภาพที่ผู้วิจัยได้ทำการบันทึกเองจากกล้องดิจิทัลยี่ห้อต่าง ๆ ในสภาวะแวดล้อมต่าง ๆ เพื่อทดสอบว่าระบบสามารถลดข้อจำกัดของสิ่งแวดล้อมที่มีผลต่อการตรวจจับหน้าคน ได้เป็นอย่างดี

4.2 ฐานข้อมูลภาพหน้ามนุษย์และภาพที่ใช้ในการทดสอบระบบ

ทำการรวบรวมข้อมูลภาพหน้ามนุษย์จากฐานข้อมูลต่าง ๆ ทางอินเทอร์เน็ตโดยฐานข้อมูลที่ใช้ในฉบับนี้ ถูกบันทึกภาพที่สภาวะแวดล้อมที่แตกต่างกัน ดังนั้นฐานข้อมูลแต่ละประเภทจึงนำมาทดสอบระบบที่ปัจจัยต่าง 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 กันและเป็นภาพเดี่ยว ซึ่งเป็นปัจจัยต่าง 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 ขนาดและที่ตำแหน่งของคนในภาพมีความไม่แน่นอน เพื่อแสดง

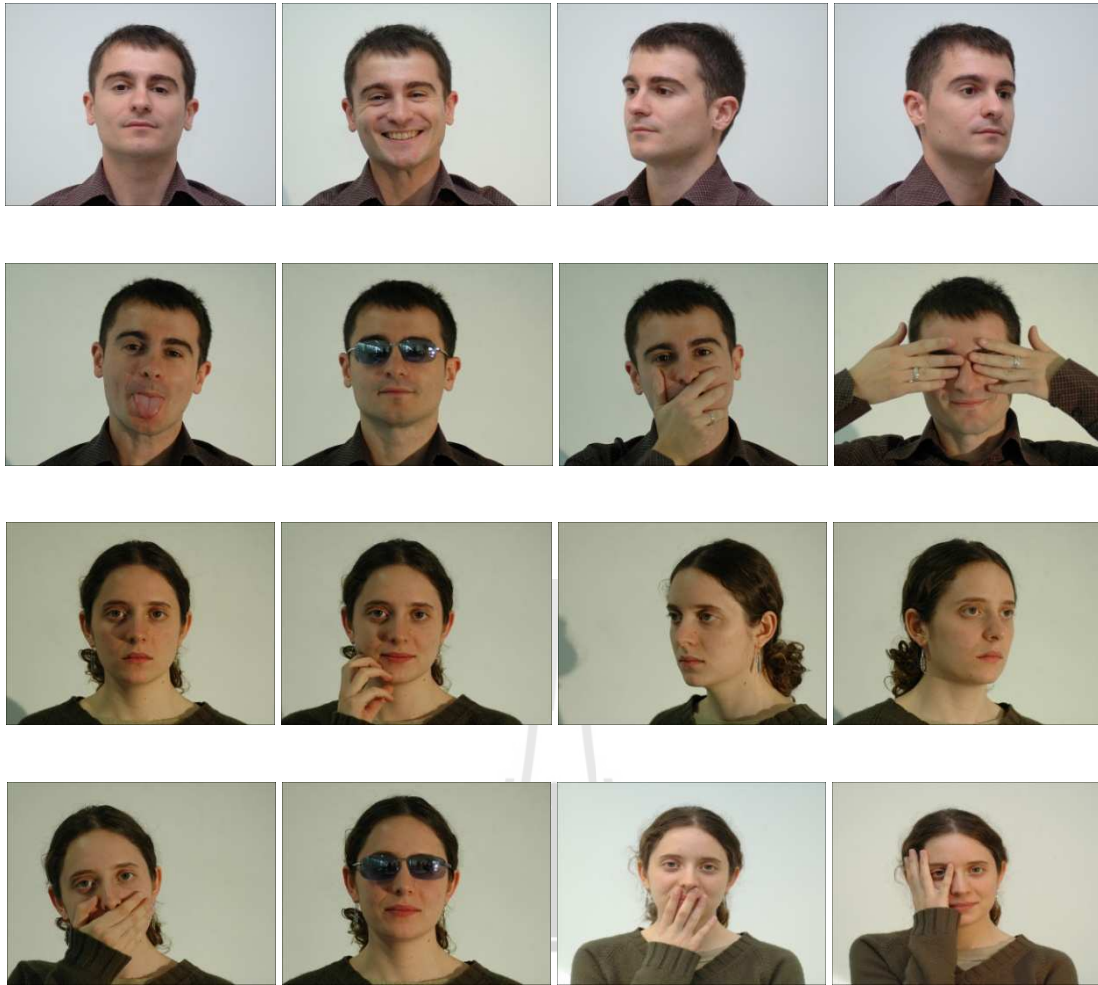
ให้เห็นว่า ระบบมีความสามารถในการตรวจจับหน้าคนในภาพโดยที่ปัจจัยต่าง ๆ ที่กล่าวมานั้น มีผลน้อยมากต่อประสิทธิภาพของระบบ



รูปที่ 4.2 ภาพตัวอย่างจากฐานข้อมูลภาพ CMU VASC

4.2.3 ฐานข้อมูลภาพ AVTG

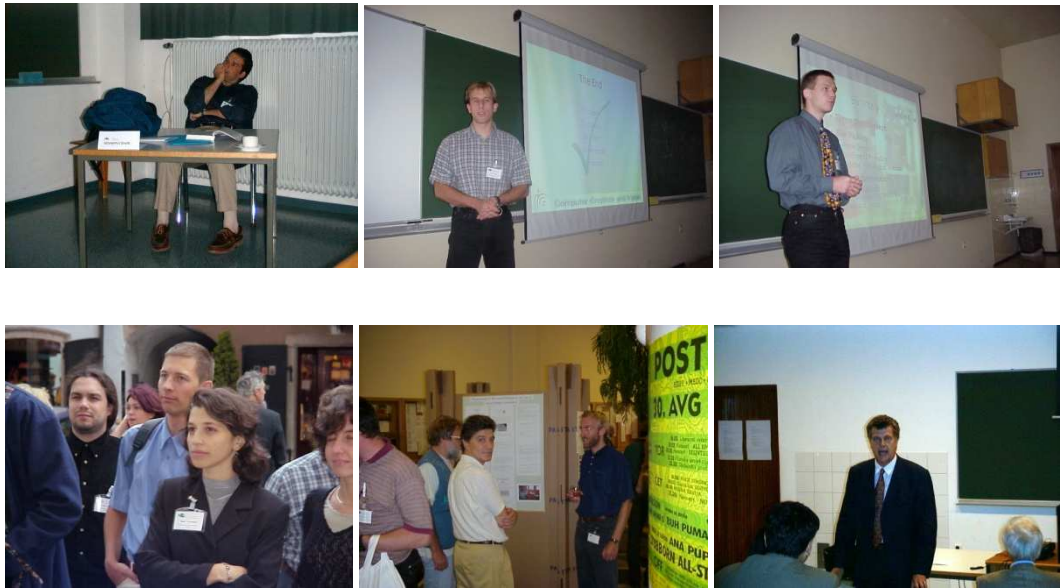
นำภาพจากฐานข้อมูลภาพ AVTG (Audio Visual Technologies Group of Technical University of Catalonia) แสดงในรูปที่ 4.3 มาใช้ในการทดสอบระบบการตรวจจับหน้าคน เมื่อลักษณะการวางท่าและการแสดงออกของหน้าคนในภาพมีการเปลี่ยนแปลง เนื่องจากภาพจากฐานข้อมูลภาพ AVTG นั้น เป็นภาพหน้าคน ที่มีความหลากหลายของลักษณะวางท่าและการแสดงออกของหน้าคน 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 ลักษณะวงท่าและการแสดงออกของหน้าคนในภาพ แสงและความสว่างในภาพ ขนาดและตำแหน่งของคนในภาพ พื้นหลังของภาพที่มีความซับซ้อน เป็นต้น เพื่อแสดงให้เห็นว่าระบบมีความสามารถในการตรวจจับหน้าคนในภาพ ภายใต้ปัจจัยต่าง ๆ ที่กล่าวมานั้น



รูปที่ 4.4 ภาพตัวอย่างจากฐานข้อมูลภาพ VALID



รูปที่ 4.5 ภาพตัวอย่างจากฐานข้อมูลภาพ CBCL

4.2.5 ภาพจากกล้องดิจิทัลฟูจิ F 601 และกล้องวิดีโอ

เป็นภาพที่ได้จากผู้วิจัยเป็นผู้บันทึกภาพเอง แสดงในรูปที่ 4.6 และ 4.7 ซึ่งขณะทำการบันทึกภาพนั้น ผู้วิจัยไม่ได้ทำการจำกัดสภาวะแวดล้อมของภาพ อันได้แก่ ความสว่าง ตำแหน่งของหน้าคน ในภาพ สีพื้นหลัง เพื่อให้ภาพที่ได้มีความหลากหลาย ในการทดสอบประสิทธิภาพของระบบ



รูปที่ 4.6 ภาพตัวอย่างจากกล้องฟูจิ F601



รูปที่ 4.7 ภาพตัวอย่างจากกล้องวิดีโอ

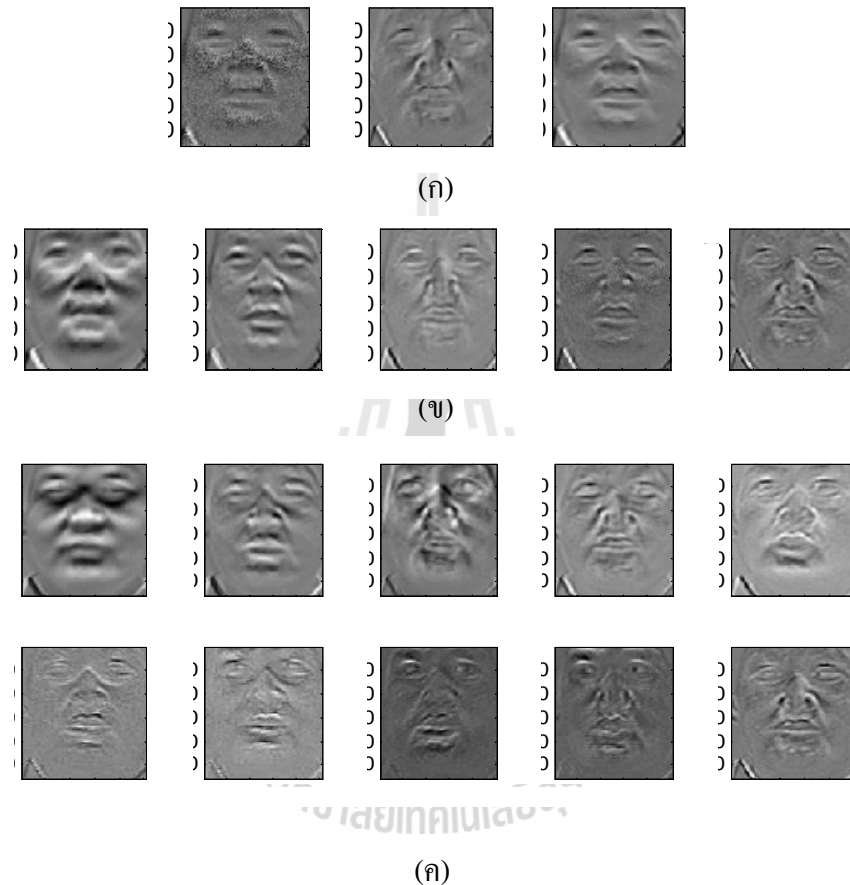
4.3 ทดสอบระบบตรวจจับหน้าคนด้วยโครงข่าย ART

4.3.1 ทดสอบระบบเมื่อมีการฝึกสอนโครงข่ายในลักษณะต่าง ๆ โดยจะทำการทดสอบระบบ เพื่อพิจารณาผลของการฝึกสอนระบบ ว่าจำนวนภาพในการฝึกสอนโครงข่ายมีผลต่อประสิทธิภาพของระบบอย่างไร โดยทำการทดสอบระบบด้วยภาพในการฝึกสอนระบบของคน ๆ เดียวกันที่มีลักษณะใบหน้าที่แตกต่างกันออกไป แสดงรูปตัวอย่างภาพที่ทำการทดลองในรูปที่ 4.8



รูปที่ 4.8 ภาพตัวอย่างในการฝึกสอนโครงข่าย

โดยจะทำการทดสอบระบบเมื่อไอเจนเวกเตอร์ที่เข้าสู่ระบบ ได้จากภาพตัวอย่างที่มีจำนวนภาพในการหาค่าไอเจนต่าง ๆ กันไป ดังแสดงค่าในตารางที่ 4.1 ซึ่งจำนวนภาพตัวอย่างที่ใช้หาค่าไอเจนนั้น มีผลต่อจำนวนมิติของข้อมูลไอเจนเวกเตอร์ที่เป็นข้อมูลในการฝึกสอน โครงข่าย โดยมิติของไอเจนเวกเตอร์ที่ใช้ในการฝึกสอน จะมีค่าเท่ากับจำนวนตัวอย่างในการฝึกสอน ดังนั้นเมื่อจำนวนภาพตัวอย่างในการฝึกสอนมากขึ้น จะเป็นผลให้ขนาดของข้อมูลไอเจนเวกเตอร์ในการฝึกสอนมากขึ้นด้วย โดยแสดงตัวอย่างภาพหน้าไอเจนของจำนวนภาพในการฝึกสอนในรูปที่ 4.9 และแสดงผลการทดลองในตารางที่ 4.1



รูปที่ 4.9 ตัวอย่างภาพหน้าไอเจน (ก) ภาพหน้าไอเจนที่จำนวนภาพในการหาหน้าไอเจน 3 ภาพ

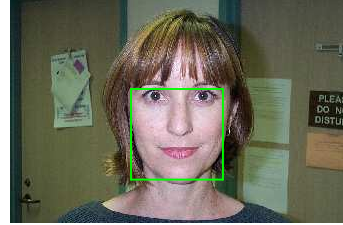
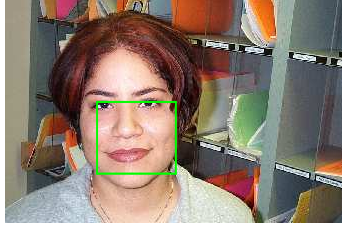
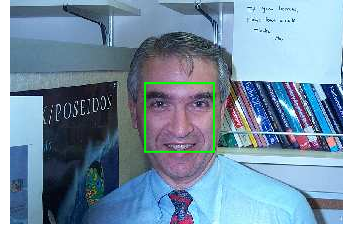
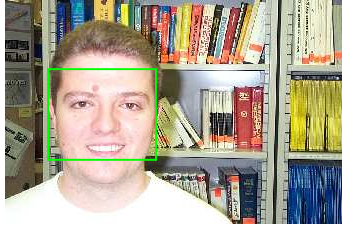
(ข) ภาพหน้าไอเจนที่จำนวนภาพในการหาหน้าไอเจน 5 ภาพ

(ค) ภาพหน้าไอเจนที่จำนวนภาพในการหาหน้าไอเจน 10 ภาพ

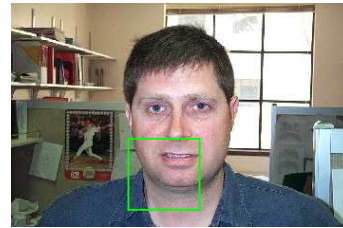
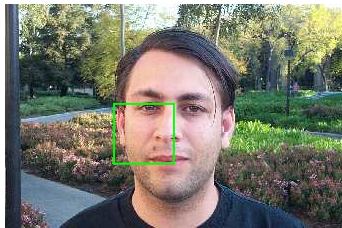
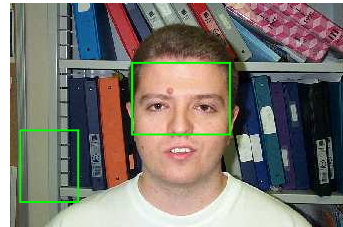
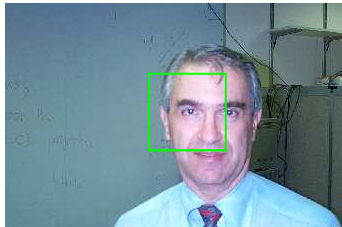
ตารางที่ 4.1 ทดสอบระบบที่จำนวนภาพในการฝึกสอน โครงข่ายต่าง ๆ

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 ภาพก็เพียงพอที่จะทำให้ระบบมีประสิทธิภาพในการตรวจจับหน้าคน ได้อยู่ในระดับที่ดี และเมื่อทำการเพิ่มจำนวนภาพในการฝึกสอนโครงข่ายมากขึ้น พบว่าโครงข่ายมีเปอร์เซ็นต์ความถูกต้องที่มากขึ้นจนถึงจำนวนภาพในการฝึกสอน 20 ภาพ เมื่อทำการเพิ่มจำนวนภาพในการฝึกสอนเป็น 30 ภาพ พบว่ามีเปอร์เซ็นต์ความถูกต้องน้อยลง ทั้งนี้อาจเนื่องมาจากการความซับซ้อนของมิติข้อมูลที่มากขึ้น อาจทำให้ระบบมีความสามารถในการคัดแยกความเป็นหน้าได้น้อยลง แต่อย่างไรก็ดีพบว่าผลการทดสอบระบบเมื่อจำนวนในการฝึกสอนโครงข่ายแตกต่างกันนั้น มีเปอร์เซ็นต์ความถูกต้องที่ไม่แตกต่างกันมากนัก ทั้งนี้เนื่องจากโครงข่าย ART นั้นเป็นโครงข่ายที่สามารถเรียนรู้ข้อมูลใหม่ได้โดยไม่ต้องทำการฝึกสอนซ้ำอีกโดยไม่ลืมข้อมูลเดิม แต่การฝึกสอนโครงข่ายด้วยจำนวนภาพตัวอย่างในปริมาณน้อยนั้น ก็อาจไม่เพียงพอที่จะทำให้ระบบมีประสิทธิภาพสูงได้ และผลจากค่าไอเกนเวกเตอร์ที่ดีที่จะเป็นข้อมูลอินพุตให้โครงข่าย ควรจะมีจำนวนมิติที่มากเพียงพอ จึงจะสามารถแสดงลักษณะของภาพได้ดีกว่าค่าไอเกนเวกเตอร์ที่มีขนาดมิติน้อย ๆ แสดงผลการทดสอบระบบในรูปแบบที่ 4.10

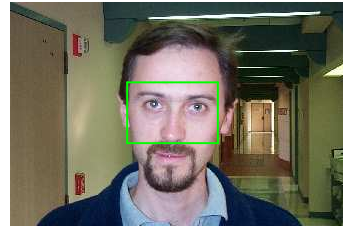
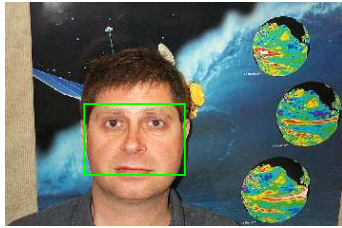
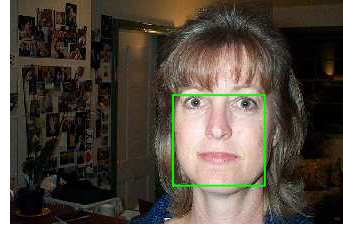
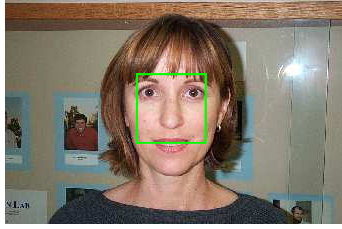


(ก) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 3 ภาพ

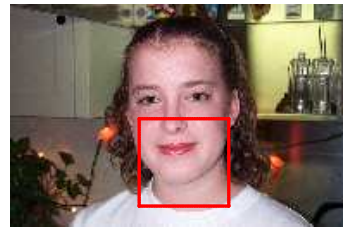


(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 3 ภาพ

รูปที่ 4.10 ตัวอย่างภาพผลการทดสอบระบบ ที่จำนวนตัวอย่างในการฝึกสอนต่าง ๆ

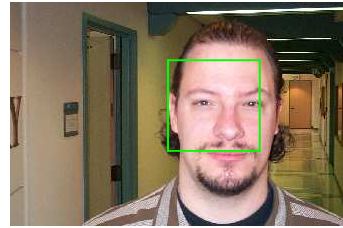
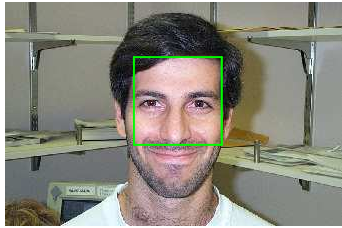
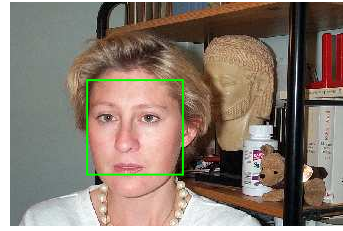
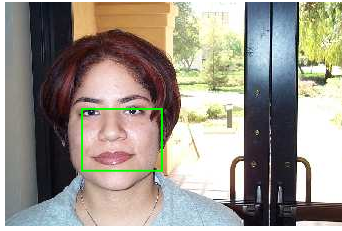


(ค) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 5 ภาพ



(ง) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 5 ภาพ

รูปที่ 4.10 ตัวอย่างภาพผลการทดสอบระบบ ที่จำนวนตัวอย่างในการฝึกสอนต่าง ๆ (ต่อ)

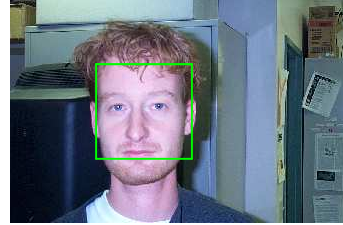
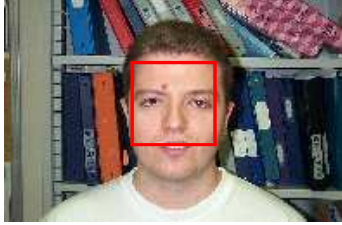
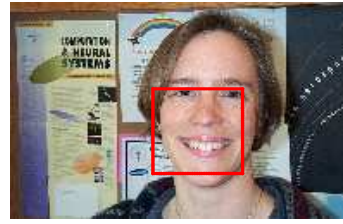
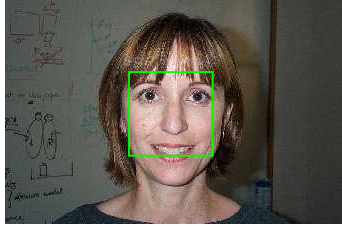


(จ) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 10 ภาพ

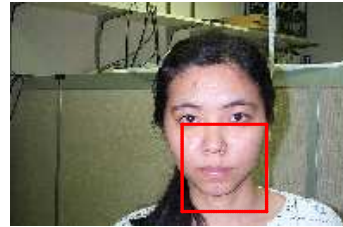


(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 10 ภาพ

รูปที่ 4.10 ตัวอย่างภาพผลการทดสอบระบบ ที่จำนวนตัวอย่างในการฝึกสอนต่าง ๆ (ต่อ)

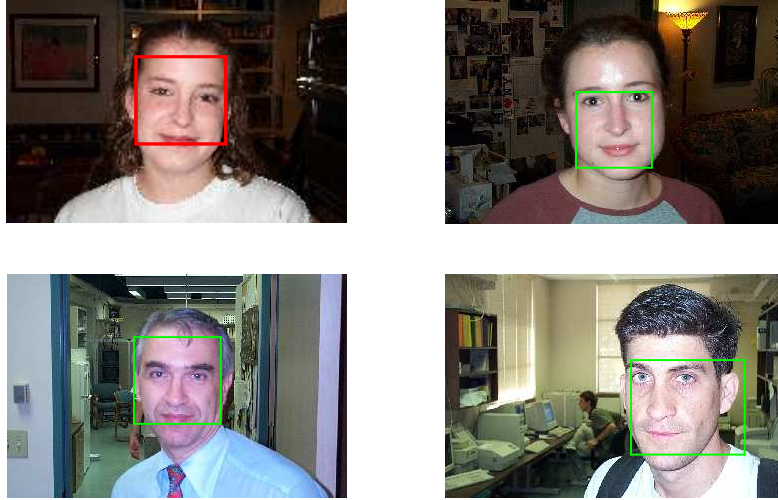


(ข) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 20 ภาพ

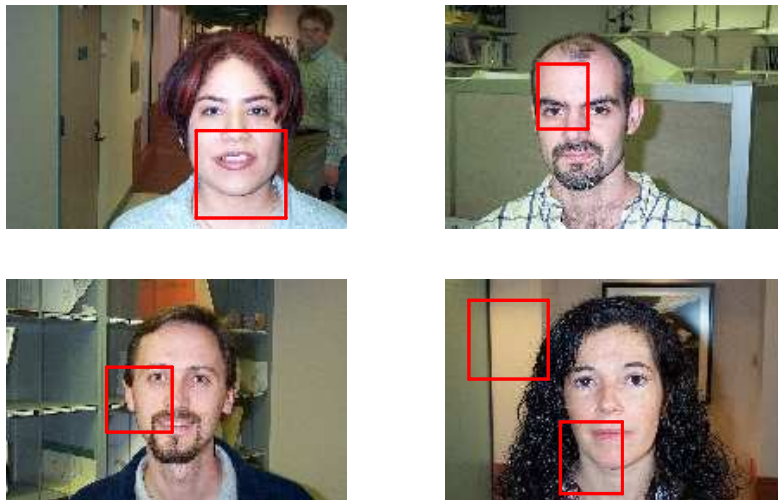


(ค) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 20 ภาพ

รูปที่ 4.10 ตัวอย่างภาพผลการทดสอบระบบ ที่จำนวนตัวอย่างในการฝึกสอนต่าง ๆ (ต่อ)



(ฉ) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 30 ภาพ

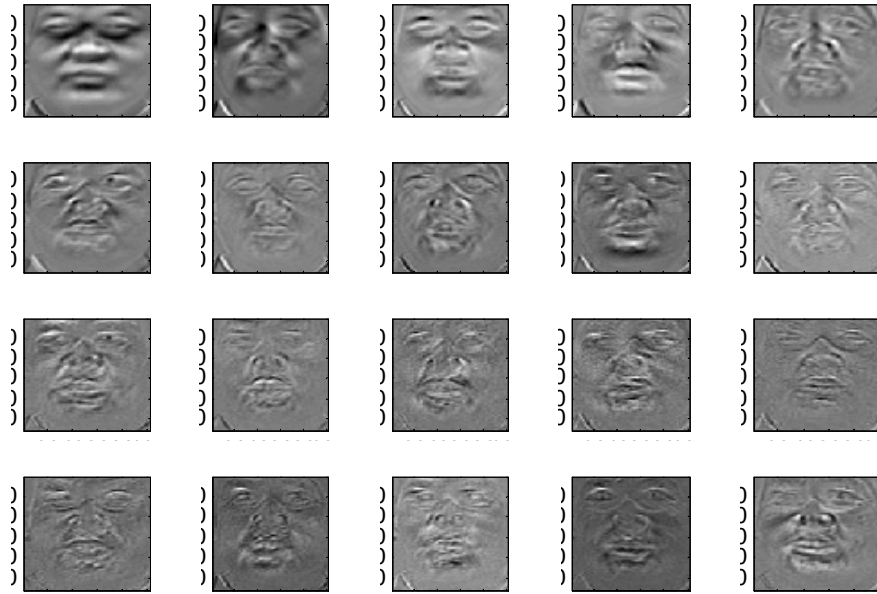


(ฅ) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของจำนวนตัวอย่างในการฝึกสอน 30 ภาพ

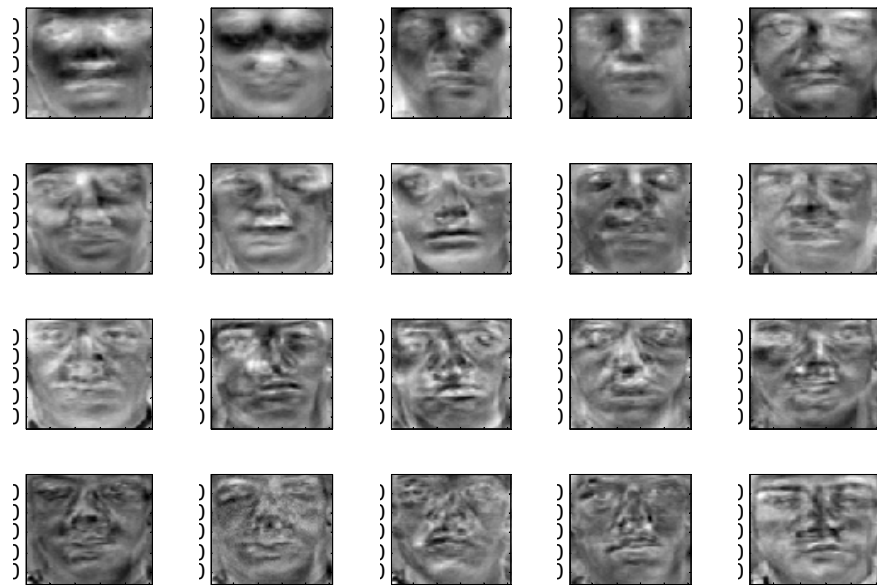
รูปที่ 4.10 ตัวอย่างภาพผลการทดสอบระบบ ที่จำนวนตัวอย่างในการฝึกสอนต่าง ๆ (ต่อ)

เมื่อทราบผลของการทดสอบจำนวนภาพในการฝึกสอน โคร่งข่ายแล้วว่ามีผลต่อประสิทธิภาพของระบบ ดังนั้นในการทดลองต่อมาเราจึงทำการเลือกใช้จำนวนภาพในการฝึกสอน 20 ภาพ เนื่องจากการทดลองข้างต้นได้แสดงให้เห็นแล้วว่า ทำให้ระบบมีประสิทธิภาพดีที่สุด โดยจะพิจารณาคูณสมบัติของภาพที่ใช้ในการฝึกสอนว่ามีผลต่อระบบอย่างไร ด้วยการแบ่งภาพในการฝึกสอนออกเป็น 3 กลุ่มหลัก ๆ คือ กลุ่มแรกทำการฝึกสอนด้วยภาพของบุคคลคนเดียวกันในลักษณะใบหน้า และการวางท่าที่แตกต่างกัน กลุ่มที่ 2 ทำการฝึกสอนด้วยภาพของคนหลาย ๆ คนในลักษณะหน้าตรงเหมือนกันหมด และกลุ่มที่ 3 ทำการฝึกสอนด้วยภาพของคนหลาย ๆ คนในลักษณะใบหน้าและการวางท่าที่แตกต่างกัน โดยแสดงตัวอย่างภาพที่

ใช้ในการฝึกสอนโครงข่ายในตารางที่ 4.2 และแสดงภาพหน้าไอเกนของตัวอย่างแต่ละแบบในรูปที่ 4.11
ได้ผลการทดลองแสดงในตารางที่ 4.2

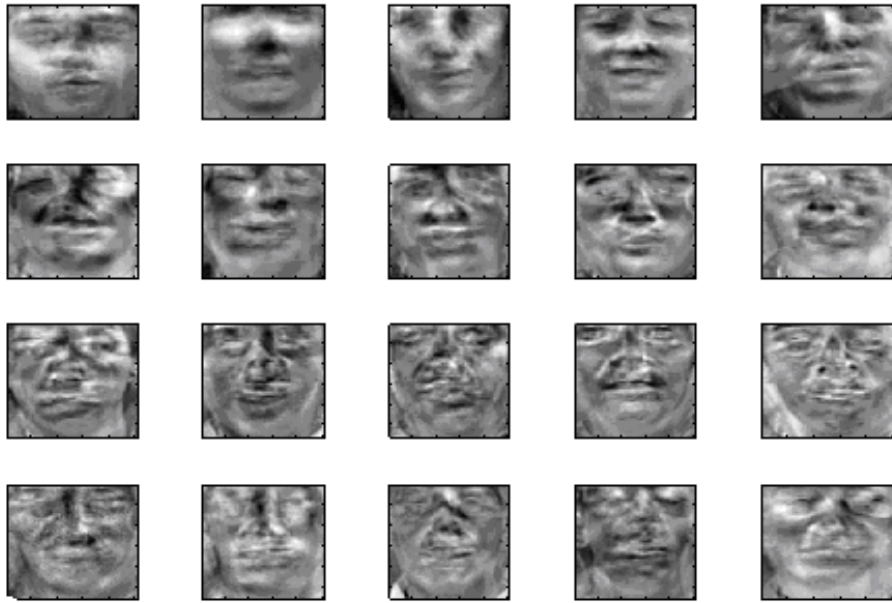


(ก) ภาพหน้าไอเกนของภาพบุคคลเดียวกัน ที่ลักษณะใบหน้าต่างกัน



(ข) ภาพหน้าไอเกนของภาพบุคคลต่างกัน ที่ลักษณะใบหน้าตรง




รูปที่ 4.11 ตัวอย่างภาพหน้าไอเกน



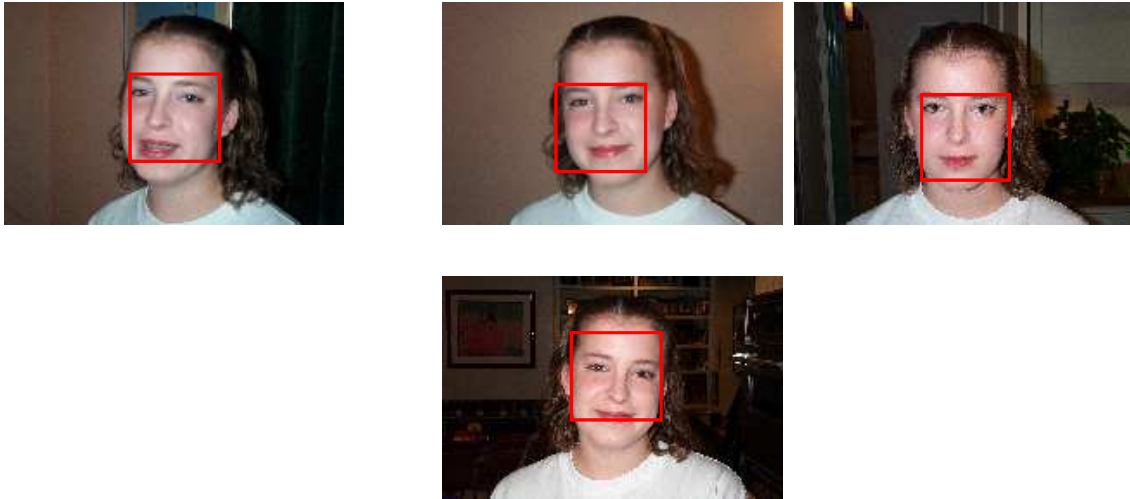
(ค) ภาพหน้าไอเคนของภาพบุคคลต่างกัน ที่ลักษณะใบหน้าต่างกัน

รูปที่ 4.11 ตัวอย่างภาพหน้าไอเคน (ต่อ)

ตารางที่ 4.2 ทดสอบระบบด้วยภาพที่มีคุณลักษณะต่าง ๆ

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

จากการทดลองพบว่า การฝึกสอนด้วยภาพบุคคลคนเดียวกันในลักษณะต่าง ๆ ให้เปอร์เซ็นต์ความถูกต้องที่ดีที่สุด และเมื่อพิจารณาจากภาพหน้าไอเคนที่ได้พบว่า ภาพหน้าไอเคนของข้อมูลที่เป็นบุคคลเดียวกันนั้นจะแสดงลักษณะเค้าโครงของหน้าได้ชัดเจนที่สุดเช่นเดียวกัน โดยค่าไอเคนที่ชัดเจนที่สุดนั้นจะเป็นค่าไอเคนที่เกิดจากค่าไอเคนค่าแรก ในรูปที่ 4.12 แสดงภาพตัวอย่างในการทดสอบระบบด้วยภาพที่มีคุณลักษณะต่าง ๆ

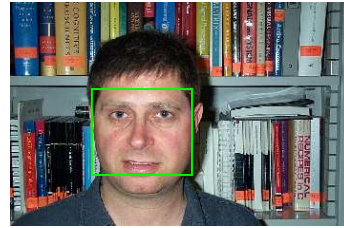
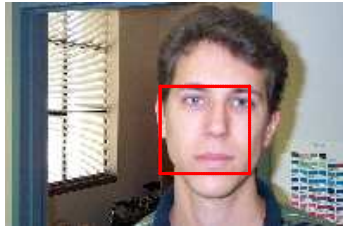
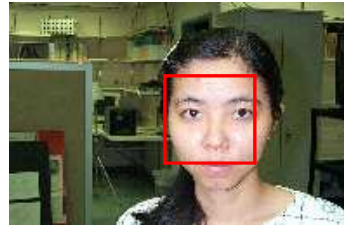
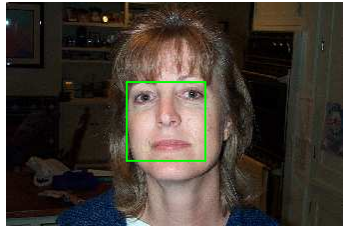


(ก) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้องของตัวอย่างในการฝึกสอน โครงข่าย
ของบุคคลเดียวกันที่ลักษณะใบหน้าต่าง ๆ

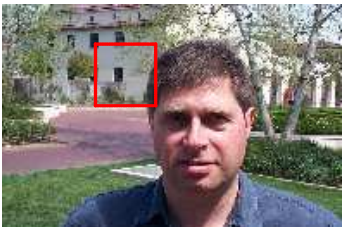
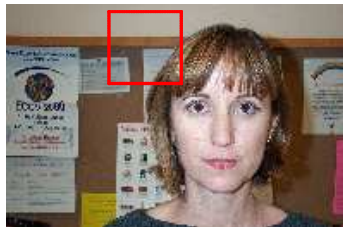


(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้องของตัวอย่างในการฝึกสอน โครงข่าย
ของบุคคลเดียวกันที่ลักษณะใบหน้าต่าง ๆ

รูปที่ 4.12 ตัวอย่างภาพผลการทดสอบระบบ ของภาพตัวอย่างในการฝึกสอนที่คุณลักษณะต่าง ๆ



(ค) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้องของตัวอย่างในการฝึกสอน โครงข่าย
ของบุคคลต่างกัน ที่ลักษณะหน้าตรง



(ง) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้องของตัวอย่างในการฝึกสอน โครงข่าย
ของบุคคลต่างกัน ที่ลักษณะหน้าตรง

รูปที่ 4.12 ตัวอย่างภาพผลการทดสอบระบบ ของภาพตัวอย่างในการฝึกสอนที่คุณลักษณะต่าง ๆ (ต่อ)



(จ) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้องของตัวอย่างในการฝึกสอน โครงข่ายของบุคคลต่างกัน ที่ลักษณะใบหน้าต่าง ๆ



(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้องของตัวอย่างในการฝึกสอน โครงข่ายของบุคคลต่างกัน ที่ลักษณะใบหน้าต่าง ๆ

รูปที่ 4.12 ตัวอย่างภาพผลการทดสอบระบบ ของภาพตัวอย่างในการฝึกสอนที่คุณลักษณะต่าง ๆ (ต่อ)

อย่างไรก็ดีจากผลการทดลองแสดงให้เห็นว่า คุณลักษณะของภาพที่ทำการฝึกสอนต่าง ๆ กันนั้น ให้ผลการตรวจจับหน้าคนได้มีประสิทธิภาพดีไม่แตกต่างกันมากนัก ดังนั้นปัจจัยในการเลือกใช้ข้อมูลในการฝึกสอนโครงข่าย จึงอาจไม่จำเป็นต้องพิจารณาถึงความหลากหลายของข้อมูลฝึกสอน ทั้งนี้เนื่องจากโครงข่าย ART นั้นมีการพัฒนาปรับตัวให้สามารถรับข้อมูลที่หลากหลายได้ในอนาคต

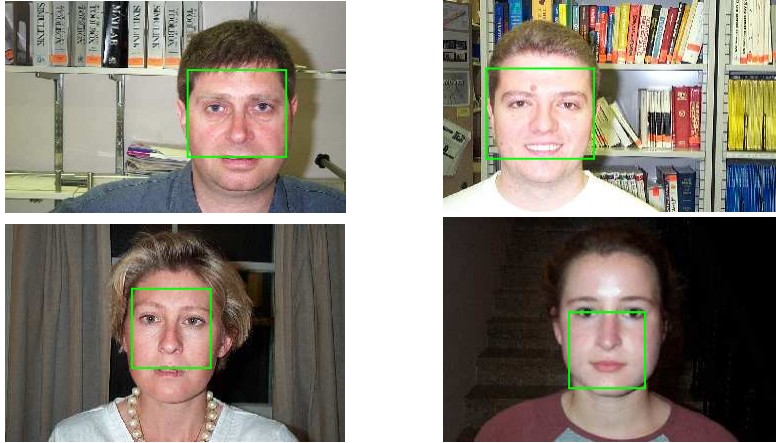
4.3.2 ทดสอบระบบที่พารามิเตอร์ของโครงข่ายต่าง ๆ โดยทำการทดสอบโครงข่ายเมื่อทำการเปลี่ยนแปลงค่าสอดคล้อง (ρ) ซึ่งเป็นพารามิเตอร์ที่กำหนดขีดความสามารถในการแบ่งกลุ่มข้อมูลของ

โครงข่าย ART โดย ρ_1 เป็นค่าขีดความสามารถในการแบ่งกลุ่มของความเป็นหน้า และความไม่เป็นหน้าออก จากกัน ส่วน ρ_2 เป็นค่าขีดความสามารถในการแบ่งกลุ่มความเป็นหน้าออกเป็นกลุ่มต่าง ๆ โดยเมื่อทำการเปลี่ยน ค่าพารามิเตอร์แล้วได้ผลการทดลองดังแสดงในตารางที่ 4.3

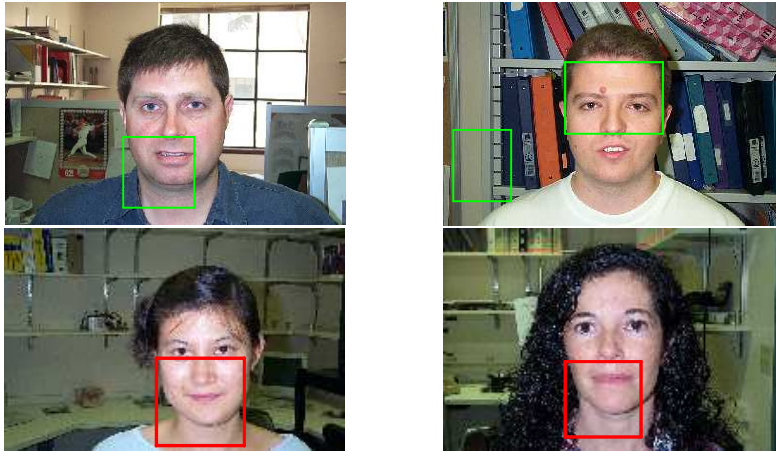
ตารางที่ 4.3 ความถูกต้องในการคัดแยกหน้าคนโดยค่าสอดคล้องค่าต่าง ๆ

ค่าสอดคล้อง		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 ระบบที่กำหนดค่าสอดคล้อง ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802 ให้ผลการทดสอบระบบที่มี ประสิทธิภาพดีที่สุด คือ ให้เปอร์เซ็นต์ความถูกต้องในการตรวจจับ ได้ดีทั้งในส่วนภาพที่เป็นหน้าคน และ ภาพที่ไม่ใช่หน้าคน โดยแสดงผลการทดสอบระบบในรูปแบบที่ 4.13



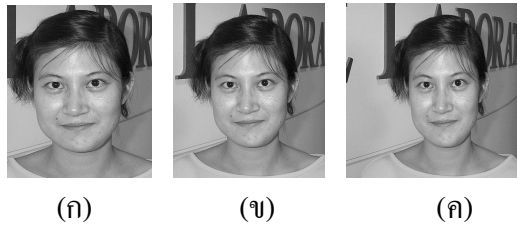
(ก) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้องของระบบที่กำหนดค่าสอดคล้อง ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802



(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้องของระบบที่กำหนดค่าสอดคล้อง ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802


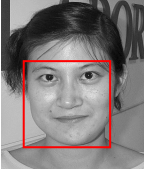

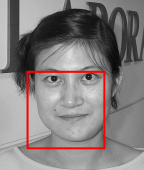

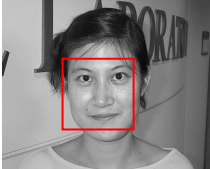
รูปที่ 4.13 ตัวอย่างภาพการทดสอบระบบของระบบที่กำหนดค่าสอดคล้อง ρ_1 มีค่า 0.851 และ ρ_2 มีค่า 0.802

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

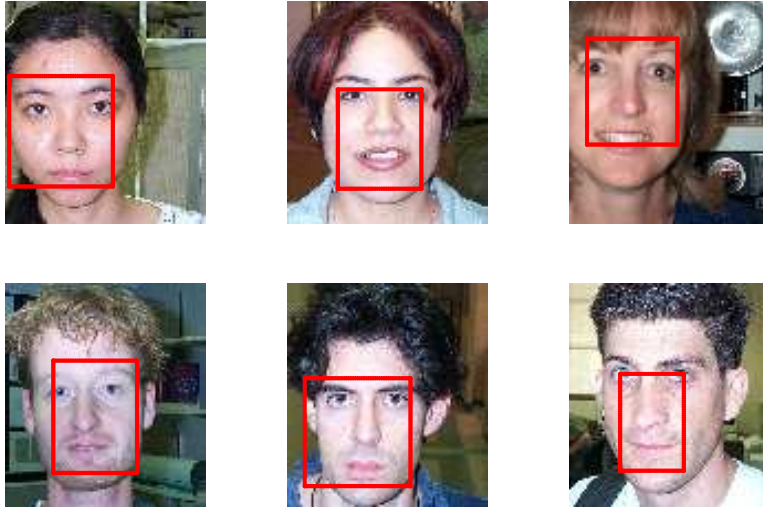


รูปที่ 4.14 ตัวอย่างภาพที่ขนาดหน้าคนในภาพต่าง ๆ (ก) หน้าคนขนาด 70%
(ข) หน้าคนขนาด 55% (ค) หน้าคนขนาดน้อยกว่า 50%

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

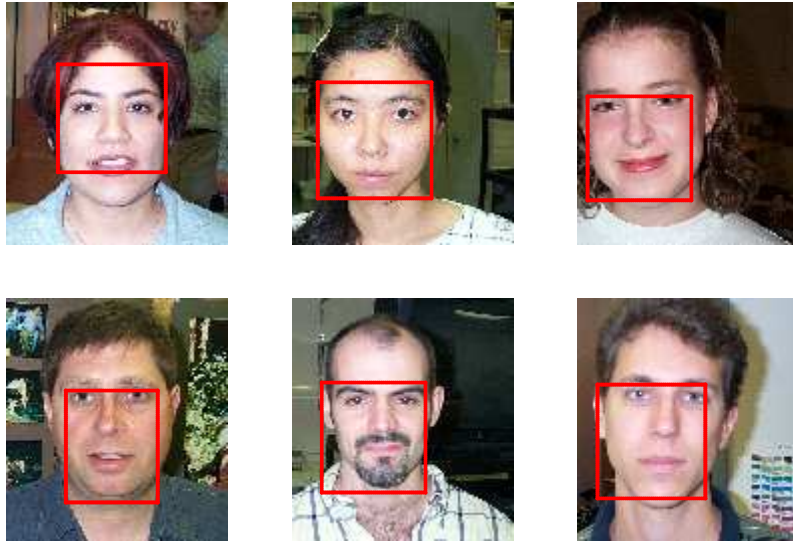


(ก) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพที่ขนาดใบหน้าคนในภาพ มีขนาด 70% ของขนาดภาพทั้งหมด

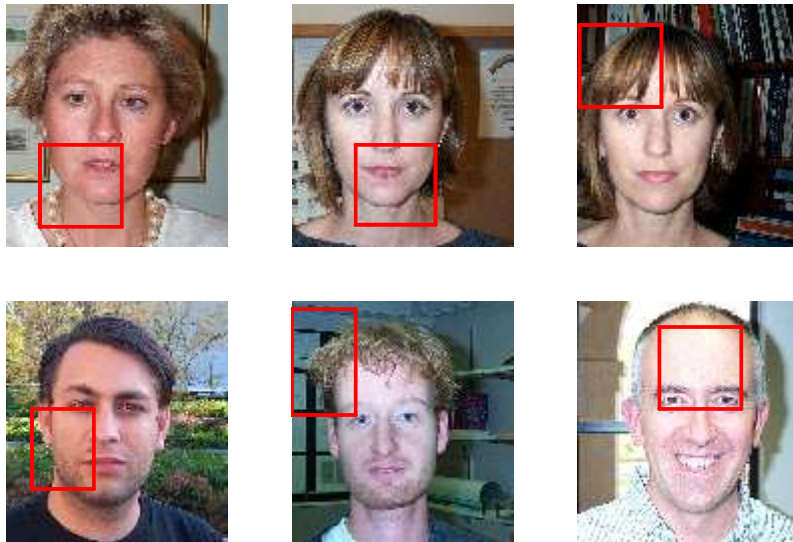


(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพที่ขนาดใบหน้าคนในภาพ มีขนาด 70% ของขนาดภาพทั้งหมด

รูปที่ 4.15 ตัวอย่างภาพผลการทดสอบระบบ ของใบหน้าคนในภาพขนาดต่าง ๆ

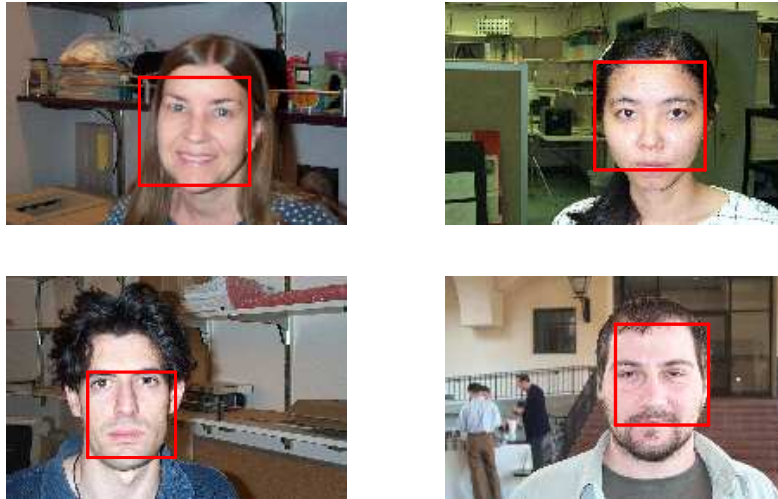


(ค) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพที่ขนาดใบหน้าคนในภาพ มีขนาด 55% ของขนาดภาพทั้งหมด

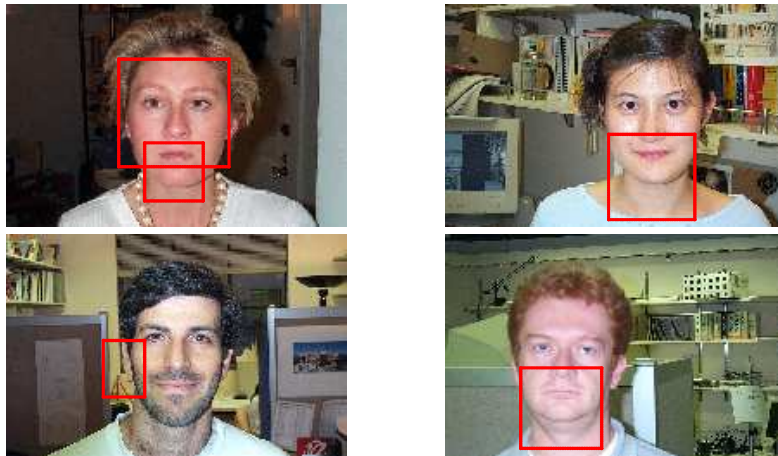


(ง) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพที่ขนาดใบหน้าคนในภาพ มีขนาด 55% ของขนาดภาพทั้งหมด

รูปที่ 4.15 ตัวอย่างภาพผลการทดสอบระบบ ของใบหน้าคนในภาพขนาดต่าง ๆ (ต่อ)



(จ) ตัวอย่างภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพที่ขนาดใบหน้าคนในภาพ มีขนาดน้อยกว่า 50% ของขนาดภาพทั้งหมด



(ข) ตัวอย่างภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพที่ขนาดใบหน้าคนในภาพ มีขนาดน้อยกว่า 50% ของขนาดภาพทั้งหมด

รูปที่ 4.15 ตัวอย่างภาพผลการทดสอบระบบ ของใบหน้าคนในภาพขนาดต่าง ๆ (ต่อ)

4.3.4 ทดสอบระบบที่หน้าในลักษณะต่าง ๆ การวางท่าของหน้าคนในภาพมีผลทำให้ระบบ ทำงานได้ต่างกัน กล่าวคือทำให้ความชัดเจนของหน้าในภาพจะต่างกัน โดยภาพหน้าตรงเป็นภาพที่มีความชัดเจนของลักษณะเฉพาะของหน้าที่ดีที่สุด และเมื่อน้ำมีสิ่งมาปิดบังหรือขวางกั้นส่วนใดส่วนหนึ่งบนใบหน้า ความชัดเจนของหน้านั้นจะน้อยลง เช่น ใบหน้าเอียงอาจจะให้เห็นลักษณะของดวงตา และ โครงหน้าไม่ชัดเจน หรือหน้าที่มีมือมาบังส่วนใดส่วนหนึ่งไว้ ทำให้ลักษณะเฉพาะของหน้าในส่วนนั้นหายไป ซึ่งจะ ได้ผลการทดสอบระบบกับภาพหน้าคนในลักษณะต่าง 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 ทั้งนี้เนื่องมาจากมีสิ่งขีดขวางมาปกปิดส่วนของใบหน้า เป็นผลให้ลักษณะในส่วนนั้นหายไป และมีดที่ปิดบังส่วนของใบหน้ายังเป็นสีผิวมนุษย์ ทำให้ขอบเขตในการค้นหากว้างมากขึ้น เป็นผลให้มีพื้นที่ความผิดพลาดในการค้นหาเพิ่มขึ้นด้วย และในส่วนของหน้าที่ปิดปาก ก็เช่นเดียวกัน ส่วนภาพคนใส่แว่นตาดำนั้น ได้ถูกปิดบังใบหน้าเช่นเดียวกัน แต่ส่งผลต่อความถูกต้องของระบบน้อยกว่าการปิดบังด้วยมือ ทั้งนี้อาจเนื่องมาจากสีของแว่นตานั้นไม่ได้มีสีเดียวกับสีผิวมนุษย์ และ โดยเป็นการปิดบังในบริเวณที่ไม่กว้าง ส่วนภาพที่ให้ผลการทดสอบระบบมีเปอร์เซ็นต์ความถูกต้องที่ดีที่สุดนั้น คือภาพหน้าตรง รองลงมาคือภาพหน้าเอียงซ้ายและหน้าเอียงขวา ทั้งนี้ก็เนื่องจากภาพหน้าตรงนั้นแสดงรายละเอียดของลักษณะเฉพาะชัดเจนที่สุด และเหมือนภาพตัวอย่างในการฝึกสอนมากที่สุด ส่วนภาพหน้าเอียงซ้าย และภาพหน้าเอียงขวา ให้เปอร์เซ็นต์ความถูกต้องที่ใกล้เคียงกัน เนื่องจากมีลักษณะใบหน้าที่ใกล้เคียงกัน แต่กลับด้านซ้าย ขวากัน

เท่านั้น แต่โดยภาพรวมนั้น ระบบสามารถคัดแยกความเป็นหน้าคน ได้ถูกต้องอยู่ในเกณฑ์ที่ดี ไม่ว่ากับหน้า
ในลักษณะใด ๆ ก็ตาม โดยรูปที่ 4.16 แสดงผลการทดสอบระบบของหน้าคนในลักษณะต่าง ๆ



(ก) ภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพหน้าตรง



(ข) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าตรง



(ค) ภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพหน้าเอียงซ้าย



(ง) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าเอียงซ้าย



(จ) ภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพหน้าเอียงขวา

รูปที่ 4.16 ตัวอย่างภาพผลการทดสอบระบบ ของภาพตัวอย่างหน้าในลักษณะต่าง ๆ



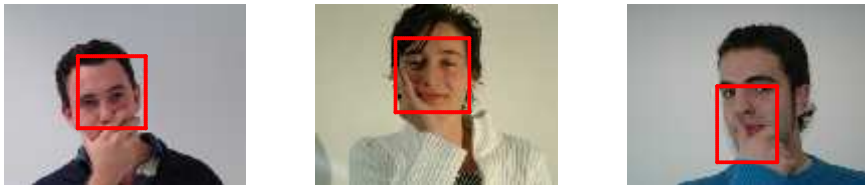
(ก) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าเอียงขวา



(ข) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าใส่แว่นดำ



(ค) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าใส่แว่นดำ



(ง) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าปิดปาก



(ฉ) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าปิดปาก

รูปที่ 4.16 ตัวอย่างภาพผลการทดสอบระบบ ของภาพตัวอย่างหน้าในลักษณะต่าง ๆ (ต่อ)



(ก) ภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพหน้าปิดตา



(ข) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าปิดตา

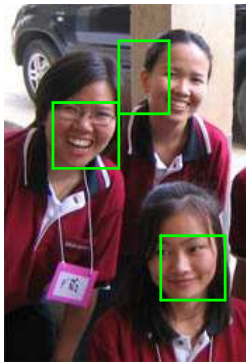
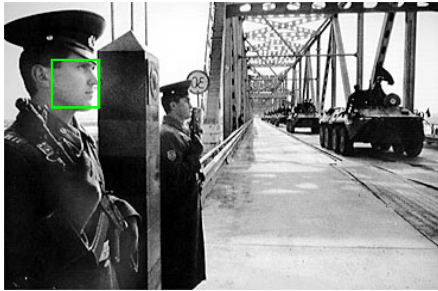
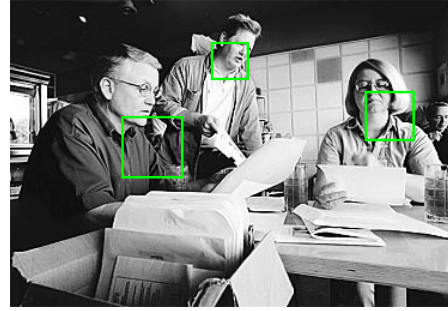
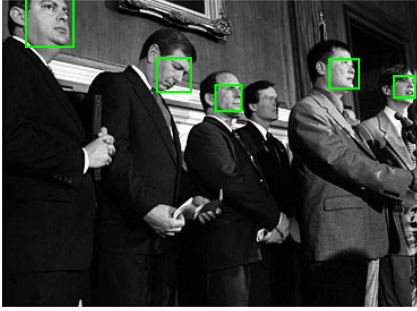
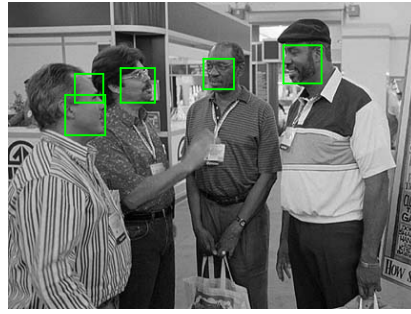
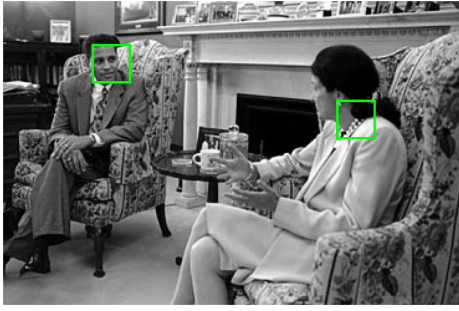
รูปที่ 4.16 ตัวอย่างภาพผลการทดสอบระบบ ของภาพตัวอย่างหน้าในลักษณะต่าง ๆ (ต่อ)

4.3.5 ทดสอบระบบกับภาพที่มีหน้าคนในภาพมากกว่า 1 คน โดยบุคคลในภาพนั้นอาจมีลักษณะการวางท่า ลักษณะใบหน้า ขนาดของหน้าในภาพแตกต่างกันออกไป หรือถูกปิดบังลักษณะของหน้าในบางส่วนได้เช่นกัน ทั้งนี้เนื่องจากธรรมชาติของมนุษย์นั้นไม่ได้มีลักษณะการวางท่า หรือการแสดงออกทางสีหน้าที่เป็นรูปแบบตายตัว ดังนั้นระบบการคัดแยกหน้าคนที่จึงต้องสามารถตรวจจับหน้าคนได้ในทุกลักษณะ และต้องสามารถหาหน้าคนในภาพได้อย่างครบถ้วนตามจำนวนหน้าจริงในภาพ โดยแสดงผลการทดลองในรูปที่ 4.17



(ก) ภาพผลการทดสอบระบบที่ถูกต้อง ของตัวอย่างภาพหน้าคนมากกว่า 1 หน้าในภาพ

รูปที่ 4.17 ตัวอย่างภาพผลการทดสอบระบบของภาพที่มีจำนวนหน้าคนในภาพมากกว่า 1 หน้าในภาพ



(ข) ภาพผลการทดสอบระบบที่ไม่ถูกต้อง ของตัวอย่างภาพหน้าคนมากกว่า 1 หน้าในภาพ

รูปที่ 4.17 ตัวอย่างภาพผลการทดสอบระบบของภาพที่มีจำนวนหน้าคนในภาพมากกว่า 1 หน้าในภาพ (ต่อ)

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 และมีลักษณะหน้าคนที่มีแน่นอน ซึ่งจากการทดลองพบว่าที่จำนวนภาพตัวอย่างในการหาค่าไอเกนเวกเตอร์ที่เหมาะสม คือ 20 ภาพ ของภาพบุคคลคนเดียวกัน ในลักษณะการแสดงออกของสีหน้าต่าง ๆ กัน

2. ค่าพารามิเตอร์ของระบบ โดยพบว่าเมื่อเราทำการปรับค่าสอดคล้อง ซึ่งเป็นค่าที่ใช้เป็นตัวกำหนดกลุ่มของค่าน้ำหนักประสาทของระบบ โดยทำการสุ่มค่าด้วย GA ทำให้ได้ค่าสอดคล้องของระบบที่เหมาะสม คือมี 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

งานวิจัยนี้ดำเนินการศึกษาและพัฒนา โครงสร้างระบบการตรวจจับหน้าคนด้วยโครงข่าย ART เพื่อให้สามารถนำไปประยุกต์ใช้ให้เป็นประโยชน์กับงานด้านอื่น ๆ ได้ต่อไป ซึ่งระบบที่พัฒนาขึ้นประกอบไปด้วยการคัดแยกสีผิวมนุษย์ออกจากสิ่งแวดล้อมโดยโครงข่ายประสาทเทียมแบบแพร่กลับ และการตรวจจับตำแหน่งของหน้าคนในภาพโดยโครงข่าย ART โดยใช้ไอเกนเวกเตอร์ที่ได้จากหลักการวิเคราะห์องค์ประกอบ มาเป็นอินพุตของโครงข่าย โดยใช้การค้นหาแบบ GA ในการหาค่าที่เหมาะสมที่สุดของค่าพารามิเตอร์ของระบบ และการดำเนินงานวิจัยดังกล่าวสำเร็จคล่องตามวัตถุประสงค์ โดยสามารถสรุปผลการศึกษาวิจัยและพัฒนาทางวิศวกรรมเป็นข้อสรุปได้ดังต่อไปนี้

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

รายละเอียดของระบบการตรวจจับใบหน้าคน พร้อมกับโครงสร้างของโครงข่าย ART ซึ่งเป็นกลไกหลัก ได้กล่าวไว้ในบทที่ 3 โดยโครงสร้างของระบบประกอบด้วย การคัดแยกความเป็นหน้าคนและความไม่เป็นหน้าคนออกโดยโครงข่าย ART ที่ทำการประยุกต์ให้เสมือนมีโครงสร้างในชั้นเปรียบเทียบ 2 ชั้น

ในบทที่ 4 แสดงรายละเอียดของการทดสอบระบบในหลายปัจจัย เพื่อแสดงให้เห็นว่าระบบมีความหลากหลายในการตรวจจับหน้าคน สามารถจับหน้าคนได้ในภาพที่มีพื้นหลังที่ซับซ้อน ในสถานะแสงและสิ่งแวดล้อมที่สามารถมองเห็นองค์ประกอบของหน้าคนได้ชัดเจน และสามารถตรวจจับหน้าคนในภาพได้มากกว่า 1 หน้า

5.2 ข้อเสนอแนะ

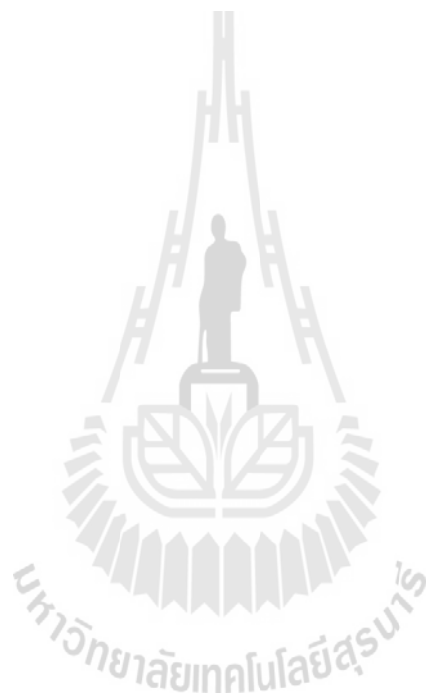
จากการดำเนินงานวิจัยที่ผ่านมาและผลที่ได้ทำให้เกิดแนวคิดและข้อเสนอแนะในการดำเนินงานวิจัยต่อไปในอนาคต ดังต่อไปนี้

1. ระบบของงานวิจัยนี้ยังไม่สามารถตรวจจับหน้าคนเมื่อมีสิ่งกีดขวางได้ดึ้นัก เนื่องจากระบบใช้ลักษณะเฉพาะของหน้าในการหาค่าไอเกนเวกเตอร์เป็นอินพุตของระบบ ดังนั้นเพื่อพัฒนาระบบให้สามารถ

ตรวจจับหน้าคนแม้มีสิ่งกีดขวางซ้อนทับกับหน้าได้ ควรจะต้องเพิ่มเติมเทคนิคในการตรวจจับหน้าคน เช่น ให้มีการตรวจจับดวงตาร่วมกับการตรวจจับหน้าคนด้วยโครงข่าย ART เป็นต้น

2. เพื่อให้ระบบสามารถใช้งานได้อย่างหลากหลาย จำเป็นต้องมีภาพตัวอย่างที่ใช้ในการหาค่าไอเกนเวกเตอร์ที่หลากหลายที่จะใช้แทนกลุ่มข้อมูลได้ทุกกลุ่ม

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



รายการอ้างอิง

- อาทิตย์ ศรีแก้ว. (2552). **ปัญญาเชิงคำนวณ**. สำนักวิชาวิศวกรรมศาสตร์: มหาวิทยาลัยเทคโนโลยีสุรนารี.
- ชมพู่ ทรัพย์ปทุมสิน. (2548). **วิธีการใหม่แบบพันทางในการแยกส่วนมือจากภาพสี**. วิทยานิพนธ์ปริญญาวิศวกรรมศาสตรมหาบัณฑิต มหาวิทยาลัยเทคโนโลยีสุรนารี.
- Bellhumeur V., Hespanha J. and Kriegeman 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**, Jhon 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

บทความวิชาการที่ได้รับการตีพิมพ์



รายชื่อบทความที่ได้รับการตีพิมพ์

อภิรดี อัมพะศิริและอาทิตย์ ศรีแก้ว, เทคนิคการตรวจจับหน้าคนด้วยโครงข่าย ART, the 11th National Computer Science Engineering Conferences (NCSEC 2007), มหาวิทยาลัยศรีปทุม, พฤศจิกายน, 2551.





ประวัติผู้วิจัย

ดร.อาทิตย์ ศรีแก้ว เกิดเมื่อวันที่ 19 พฤศจิกายน พ.ศ. 2515 สำเร็จการศึกษาระดับปริญญาตรีใน สาขาวิชาวิศวกรรมอิเล็กทรอนิกส์จาก สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง เมื่อ พ.ศ. 2537 สำเร็จการศึกษาระดับ ปริญญาโทและปริญญาเอกสาขาวิศวกรรมไฟฟ้าจาก Vanderbilt University ประเทศสหรัฐอเมริกา เมื่อ พ.ศ. 2540 และ 2543 ตามลำดับ ปัจจุบันดำรงตำแหน่งผู้ช่วยศาสตราจารย์ สาขาวิชาวิศวกรรม ไฟฟ้า สำนักวิชาวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี มีความสนใจงานวิจัยทางด้านการมองเห็นของคอมพิวเตอร์และหุ่นยนต์ การประมวลผลภาพ และระบบทางปัญญาประดิษฐ์

นางสาวอภิรดี อัมพะลิวรี 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 ก. มีความสนใจทางด้านการประมวลผลสัญญาณภาพ การมองเห็นของคอมพิวเตอร์