

การรู้จำภาพใบหน้าคน
(Face Recognition)



นางสาวกฤษณา บุญกร ไกรฤกษ์ รหัส 42010007
นางสาวสิริวรรณ ไชยแก้ว รหัส 42010383

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิศวกรรมอิเล็กทรอนิกส์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2545

เลขหมู่.....

เลขทะเบียน..... 50398

วันเดือนปี 13 พ.ค. 2547



เอกสารนี้เป็นเอกสารลิขสิทธิ์สงวนสำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้ไปใช้ในวัตถุประสงค์อื่น
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การรู้จำภาพใบหน้าคน
(Face Recognition)



ปฏิญานี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
สาขาวิศวกรรมอิเล็กทรอนิกส์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2545

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การรู้จำภาพใบหน้าคน

น.ส. กฤษณา บุญกรไกรฤกษ์ รหัส 42010007

น.ส. สิริวรรณ ไชยแก้ว รหัส 42010383

รศ. ดร. สุรพันธุ์ เอื้อไพบูลย์ (อาจารย์ที่ปรึกษา)

ปีการศึกษา 2545

บทคัดย่อ

ปริญญานิพนธ์ฉบับนี้ เป็นการออกแบบโปรแกรมเพื่อที่จะตรวจจับใบหน้าคน และทำการจดจำภาพใบหน้าคน จากภาพที่ถ่ายจากกล้องวิดีโอซึ่งต่อตรงเข้ากับคอมพิวเตอร์ โดยจะใช้โปรแกรม MATLAB ในการประมวลผลภาพเพื่อจะตรวจจับใบหน้าคนออกมาจากรูป ซึ่งการตรวจจับใบหน้าคนเป็นพื้นฐานของการรู้จำภาพใบหน้าคน และใช้ทฤษฎีโครงข่ายประสาทเทียมแบบย้อนกลับในการสอนให้คอมพิวเตอร์รู้จำใบหน้าคน เพื่อที่จะนำไปประยุกต์ใช้กับระบบรักษาความปลอดภัย ที่จะต้องใช้การพิสูจน์บุคคลก่อนให้บุคคลนั้นเข้าสู่บริเวณที่มีการรักษาความปลอดภัยอย่างเคร่งครัด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Face Recognition

Kritsana Boonkornkrirerk ID. 42010007

Siriwan Chaikaew ID. 42010383

Assoc. Prof. Dr. Surapan Aungpiboon(Advisor)

Year 2002

Abstract

This project is about programming for face recognition . By using MATLAB, we provide the source code for segment face region out of the picture. The picture was taken by video camera attached to the computer in order to process picture by computer. Face segmentation is the basis of face recognition. After that use Back-propagation Neural Network for training network. One of the applications of face recognition is to identify persons before letting them access to the restricted area.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อ	ก
Abstract	ข
สารบัญ	ก-จ
สารบัญรูปภาพ	ฉ-ณ
สารบัญตาราง	ญ
บทที่ 1 บทนำ	1
1.1 รายละเอียดโดยย่อของโครงการ	1
1.2 วัตถุประสงค์ของโครงการ	1
1.3 ขอบเขตของโครงการ	2
1.4 ประโยชน์หรือผลที่คาดว่าจะได้รับ	2
บทที่ 2 ทฤษฎี และ หลักการเบื้องต้น	4
2.1 ระบบการมองเห็นภาพ (Vision System)	4
2.1.1 การได้มาซึ่งภาพ (Image Acquisition)	4
2.1.2 กระบวนการประมวลผลภาพ (Image Processing)	5
2.1.3 ผลที่ได้และการแสดงผล (Output or Display)	5
2.2 หลักการเบื้องต้นของการประมวลผลภาพ (Image Processing Fundamentals)	6
2.2.1 พิกเซล (Pixel)	6
2.2.2 ตำแหน่งของพิกเซล (Pixel Location)	7
2.3 ระดับเกรย์ (Gray Scale)	9
2.4 ฮิสโตแกรม (Histogram)	11
2.5 การแปลงระดับเกรย์ (Gray-level Transformation)	12
2.6 พื้นฐานและระบบของสีโมเดล RGB (Red , Green , Blue)	13
2.7 บิตแมป (Bitmaps)	14
2.8 ความสว่าง (Brightness)	15

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ในการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.9 ชนิดของรูปภาพในกล่องเครื่องมือของโปรแกรมเมทเลป	15
2.9.1 คณิตรูปภาพ	15
2.9.2 ความหนาแน่นของรูปภาพ	17
2.9.3 ไบนารีของรูปภาพ	18
2.9.4 อาร์จีบีของรูปภาพ	18
2.10 รูปภาพที่เป็นมัลติเฟรมอาร์เรย์	21
2.10.1 ข้อจำกัดในการใช้มัลติเฟรม	22
2.11 สรุปชนิดของรูปภาพและคลาสต่างๆ	22
บทที่ 3 การออกแบบและโปรแกรมที่ใช้ในการประมวลผลภาพ	24
3.1 Skin Color Model	24
3.2 Skin Segmentation	26
3.3 Skin Regions	28
3.4 ทฤษฎีโครงข่ายประสาทเทียม (Artificial Neural Network)	29
3.5 โครงข่ายประสาทเทียมแบบชั้นเดียว (Single Layer Artificial Neural Network)	33
3.6 โครงข่ายประสาทเทียมแบบหลายชั้น (Multilayer Artificial Neural Network)	34
3.7 การฝึกสอนให้กับโครงข่าย (Training Algorithm)	34
3.8 ขั้นตอนการเรียนรู้แบบแพร่ย้อนกลับ (Back-propagation Learning)	35
3.9 สรุปขั้นตอนการประมวลผลรูป	37
บทที่ 4 การทดลองและผลการทดลอง	45
4.1 การทดลองหาประสิทธิภาพของการแยกสีผิว	45
4.2 การทดลองหาประสิทธิภาพของการตรวจจับหน้า	46
4.3 ผลการนำภาพต้นแบบมาทำการประมวลผลเพื่อตรวจจับหน้า	47
4.3.1 ผลการตรวจจับหน้าที่ได้ผลดี	47
4.3.1.1 ผลการตรวจจับหน้าตรงที่ได้ผลดี	49
4.3.1.2 ผลการตรวจจับหน้าที่เอียงที่ได้ผลดี	51

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ในการเรียนการสอนที่ไม่อนุญาตให้นำไปใช้ประโยชน์ด้วยการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.1.3 ผลการตรวจจับหน้าที่มีการสวมแว่นตาที่ได้ผลดี	53
4.3.1.4 ผลการตรวจจับหน้าที่ทำการหลับตาที่ได้ผลดี	55
4.3.2 ผลการตรวจจับหน้าที่ได้ผลไม่ดี	57
4.3.3 ผลการหาประสิทธิภาพของการตรวจจับหน้าจากภาพต้นแบบ	58
4.4 การหาประสิทธิภาพของภาพที่สอนให้รู้จำใบหน้า	58
4.5 การทดลองหาประสิทธิภาพของการจดจำภาพ	58
4.5.1 ผลการจดจำภาพใบหน้าที่ได้ผลดี	59
4.5.1.1 ผลการจดจำใบหน้าตรงที่ได้ผลดี	60
4.5.1.2 ผลการจดจำใบหน้าเอียงที่ได้ผลดี	61
4.5.1.3 ผลการจดจำใบหน้าที่มีการสวมแว่นตาที่ได้ผลดี	62
4.5.1.4 ผลการจดจำใบหน้าที่หลับตาที่ได้ผลดี	63
4.5.2 ผลการจดจำภาพใบหน้าที่ได้ผลไม่ดี	64
บทที่ 5 สรุปและวิจารณ์ผลการทดลอง	65
5.1 วิเคราะห์และสรุปผลการทดลอง	65
5.2 ปัญหาและข้อจำกัดของโปรแกรม	66
5.3 การพัฒนาโปรแกรมต่อไปในอนาคต	66
บรรณานุกรม (โปรแกรม MATLAB ที่ใช้ในการประมวลผลภาพเพื่อตรวจจับใบหน้า)	
กิตติกรรมประกาศ	
หนังสืออ้างอิง	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ

	หน้า
รูปที่ 1.1 แสดงภาพรวมของโครงการ	3
รูปที่ 1.2 บล็อกไดอะแกรม (Block Diagram) แสดงรายละเอียดโดยย่อของโครงการ	3
รูปที่ 2.1 พิกเซล a) แสดงลักษณะของภาพ	6
b) แสดงลักษณะของตำแหน่งพิกเซล	6
รูปที่ 2.2 คณิตศาสตร์แสดงพิกเซลในเมทริกซ์ภาพหนึ่ง	7
รูปที่ 2.3 a) ลักษณะของแสงที่ตกกระจายไม่เท่ากันบนพื้นผิวภาพ	8
b) ค่าพิกเซลของภาพบนพื้นผิวภาพ	8
รูปที่ 2.4 ตัวอย่างของระดับเกรย์ (Gray Scales)	10
รูปที่ 2.5 ฮิสโตแกรมของ 8 ระดับเกรย์ จากเมทริกซ์ 4×5	11
รูปที่ 2.6 ฮิสโตแกรมของระดับเกรย์ a) ใช้เทรชโฮลด์ 1 จุด (Single Threshold)	12
b) ใช้เทรชโฮลด์หลายจุด (Multiple Threshold)	12
รูปที่ 2.7 แสดงโมเดลสี RGB	14
รูปที่ 2.8 ค่าของพิกเซลที่ไปยังค่าที่เทียบเคียงสีในดัชนีรูปภาพ	16
รูปที่ 2.9 ค่าพิกเซลในความหนาแน่นของรูปภาพกำหนดระดับของสีเทา	17
รูปที่ 2.10 พิกเซลในไบนารีของรูปภาพมีค่าที่เป็นไปได้สองค่า คือ 0 หรือ 1	18
รูปที่ 2.11 ระบายสีของรูปภาพอาร์จีบี	19
รูปที่ 2.12 ระบายสีที่แยกจากกันของรูปภาพอาร์จีบี	21
รูปที่ 3.1 กราฟการกระจายสีผิวของบุคคลต่าง ๆ กัน	25
รูปที่ 3.2 กราฟการกระจายของสีผิว โดยวิธีการกระจายของเกาส์เซียน	25
รูปที่ 3.3 (รูปซ้าย) รูปต้นแบบ (รูปขวา) แสดงภาพ skin-lielihood	26
รูปที่ 3.4 (รูปซ้าย) ภาพ skin-lielihood	27
(รูปขวา) ภาพที่ผ่านการแยกสีผิว (skin-segmented image)	27
รูปที่ 3.5 (รูปซ้าย) บริเวณผิวที่ต้องการแยก (รูปขวา) หลังจากใช้การแยกแต่ละส่วนออกมา	28
รูปที่ 3.6 ส่วนของนิวรอล 1 หน่วย	30
รูปที่ 3.7 ฟังก์ชันเส้นตรง	31

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.8	ซิกมอยด์ฟังก์ชัน	32
รูปที่ 3.9	hyperbolic tangent function	33
รูปที่ 3.10	ลักษณะของนิวรอล 1 หน่วยที่ประกอบด้วยฟังก์ชันกระตุ้น	33
รูปที่ 3.11	โครงข่ายประสาทเทียมแบบหลายชั้น (Multilayer Artificial Neural Network)	34
รูปที่ 3.12	ไดอะแกรมของ Back-propagation Neural Network	35
รูปที่ 3.13	บล็อกไดอะแกรมแสดงขั้นตอนการประมวลผลรูป	38
รูปที่ 3.14	บล็อกไดอะแกรมแสดงขั้นตอนในการสอนให้เน็ตเวิร์กเรียนรู้	44
รูปที่ 4.1	(ซ้าย) ภาพค้นแบบ (ขวา) ภาพที่ได้จากการแยกสีผิวแบบสมบูรณ์	45
รูปที่ 4.2	(ซ้าย) ภาพที่ใช้ทดลอง (ขวา) ภาพที่ได้จากการแยกสีผิวแบบไม่สมบูรณ์	46
รูปที่ 4.3	ภาพที่ตรวจจับได้ถูกต้อง	47
รูปที่ 4.4	ภาพที่ตรวจจับไม่ถูกต้อง	47
รูปที่ 4.5	ภาพค้นแบบ	47
รูปที่ 4.6	ภาพสุ่มผิวหน้า	47
รูปที่ 4.7	ภาพ skin-likelihood	48
รูปที่ 4.8	ภาพที่ได้จากการแยกสีผิว	48
รูปที่ 4.9	ภาพที่ได้จากการหาจำนวน segment	48
รูปที่ 4.10	ภาพที่ได้จากการตรวจจับหน้าที่ได้ผลดี	48
รูปที่ 4.11	ภาพค้นแบบ	49
รูปที่ 4.12	ภาพ skin-likelihood	49
รูปที่ 4.13	ภาพที่ได้จากการแยกสีผิว	49
รูปที่ 4.14	ภาพที่ได้จากการหาจำนวน segment	50
รูปที่ 4.15	ภาพที่ได้จากการตรวจจับหน้าที่ได้ผลดี	50
รูปที่ 4.16	ภาพค้นแบบ	51
รูปที่ 4.17	ภาพ skin-likelihood	51
รูปที่ 4.18	ภาพที่ได้จากการแยกสีผิว	51
รูปที่ 4.19	ภาพที่ได้จากการหาจำนวน segment	52
รูปที่ 4.20	ภาพที่ได้จากการตรวจจับหน้าที่ได้ผลดี	52
รูปที่ 4.21	ภาพค้นแบบ	53
รูปที่ 4.22	ภาพ skin-likelihood	53
รูปที่ 4.23	ภาพที่ได้จากการแยกสีผิว	53

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับงานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.24	ภาพที่ได้จากการหาจำนวน segment	54
รูปที่ 4.25	ภาพที่ได้จากการตรวจจับหน้าที่ได้ผลดี	54
รูปที่ 4.26	ภาพต้นแบบ	55
รูปที่ 4.27	ภาพ skin-likelihood	55
รูปที่ 4.28	ภาพที่ได้จากการแยกสีผิว	55
รูปที่ 4.29	ภาพที่ได้จากการหาจำนวน segment	56
รูปที่ 4.30	ภาพที่ได้จากการตรวจจับหน้าที่ได้ผลดี	56
รูปที่ 4.31	ภาพต้นแบบที่ให้ผลตรวจจับหน้าไม่ดี	57
รูปที่ 4.32	ภาพที่ได้จากการแยกสีผิวที่ให้ผลตรวจจับหน้าไม่ดี	57
รูปที่ 4.33	ภาพที่ได้จากการตรวจจับหน้าที่ได้ผลไม่ดี	57
รูปที่ 4.34	รูปที่ตัดเฉพาะใบหน้า	59
รูปที่ 4.35	RGB profile	59
รูปที่ 4.36	ภาพ Output	59
รูปที่ 4.37	ภาพที่ทำการจดจำได้อย่างถูกต้อง	59
รูปที่ 4.38	รูปที่ตัดเฉพาะใบหน้า	60
รูปที่ 4.39	RGB profile	60
รูปที่ 4.40	ภาพ Output	60
รูปที่ 4.41	ภาพที่ทำการจดจำได้อย่างถูกต้อง	60
รูปที่ 4.42	รูปที่ตัดเฉพาะใบหน้า	61
รูปที่ 4.43	RGB profile	61
รูปที่ 4.44	ภาพ Output	61
รูปที่ 4.45	ภาพที่ทำการจดจำได้อย่างถูกต้อง	61
รูปที่ 4.46	รูปที่ตัดเฉพาะใบหน้า	62
รูปที่ 4.47	RGB profile	62
รูปที่ 4.48	ภาพ Output	62
รูปที่ 4.49	ภาพที่ทำการจดจำได้อย่างถูกต้อง	62
รูปที่ 4.50	รูปที่ตัดเฉพาะใบหน้า	63
รูปที่ 4.51	RGB profile	63
รูปที่ 4.52	ภาพ Output	63
รูปที่ 4.53	ภาพที่ทำการจดจำได้อย่างถูกต้อง	63

เอกสารนี้เป็นเอกสารที่ปราศจากลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ไม่อนุญาติให้นำไปใช้ประโยชน์ด้วยการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.54	รูปที่ตัดเฉพาะใบหน้า	64
รูปที่ 4.55	RGB profile	64
รูปที่ 4.56	ภาพ Output	64
รูปที่ 4.57	ภาพที่ไม่สามารถทำการจดจำได้อย่างถูกต้อง	64



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

	หน้า
ตารางที่ 2.1	
สรุปรชนิดของรูปภาพและคลาสต่างๆ	23
ตารางที่ 4.1	
ผลการทดลองการหาประสิทธิภาพของการแยกสีผิว	45
ตารางที่ 4.2	
ผลการทดลองหาประสิทธิภาพของการตรวจจับหน้า	46
ตารางที่ 4.3	
ผลการหาประสิทธิภาพของการตรวจจับหน้าจากภาพต้นแบบ	58
ตารางที่ 4.4	
ผลการหาประสิทธิภาพของภาพที่สอนให้รู้จำใบหน้า	58
ตารางที่ 4.4	
ผลการหาประสิทธิภาพของการจดจำภาพใบหน้า	58



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

ในปัจจุบันนี้วิวัฒนาการทางด้านคอมพิวเตอร์ได้ก้าวหน้าไปอย่างรวดเร็วมาก แต่การทำงานของคอมพิวเตอร์ก็ยังมีข้อจำกัดอยู่ คือคอมพิวเตอร์ไม่สามารถเรียนรู้และคิดเองได้เหมือนคน จากความต้องการที่จะให้คอมพิวเตอร์ทำงานได้เองเหมือนคน เพื่อที่จะนำคอมพิวเตอร์มาใช้งานแทนคนได้ ในงานที่มีความเสี่ยงสูง เพราะฉะนั้นจึงมีการคิดค้นหาวิธีการต่างๆที่จะสามารถทำให้คอมพิวเตอร์เรียนรู้และคิดเองได้ และวิธีการของไอเกนเฟซ (Eigenfaces) โดยจะนำรูปภาพใบหน้าคนที่ได้ มาหาโควาเรียนเมทริกซ์, ไอเกนเวกเตอร์ และไอเกนวาล्यू หลังจากนั้นนำภาพที่ต้องการทดสอบมาเปรียบเทียบกับไอเกน ซึ่งค่าไอเกนที่ต่างกันน้อยที่สุด คือภาพที่ตรงกันกับภาพที่อยู่ในฐานข้อมูล และด้วยวิธีการเช่นนี้เองทำให้เราสามารถให้คอมพิวเตอร์รู้จักและจดจำสิ่งที่เราต้องการให้จำได้

เมื่อคอมพิวเตอร์มีบทบาทกับชีวิตของมนุษย์มากขึ้น จึงได้มีการนำเทคโนโลยีทางด้านคอมพิวเตอร์นี้มาช่วยในงานต่างๆให้มนุษย์มีความสะดวกสบายมากขึ้น เช่น โครงการการรู้จำภาพใบหน้าคนนี้ได้นำระบบการมองเห็น (Vision System) มาประยุกต์ใช้กับ ไอเกนเฟซ เพื่อนำไปประยุกต์ใช้ในระบบรักษาความปลอดภัย หรือ ควบคุมการทำงานของระบบต่างๆได้

1.1 รายละเอียดโดยย่อของโครงการ

โครงการการรู้จำภาพใบหน้าคนนี้จะใช้วิธี ไอเกนเฟซ ในการสอนให้คอมพิวเตอร์รู้จักภาพใบหน้าคนต่างๆ ที่ได้รับมาจากกล้องวิดีโอที่ต่อกับคอมพิวเตอร์ และนำมาประยุกต์ใช้กับระบบการมองเห็น เพื่อนำภาพใบหน้าคนเข้าสู่ระบบไอเกนเฟซและให้คอมพิวเตอร์จดจำ ในการประมวลผลรูปภาพใบหน้าเพื่อใช้เป็นฐานข้อมูล และใช้โปรแกรมเมทแผลป กับวิธีการของไอเกนเฟซ เพื่อให้คอมพิวเตอร์รู้จักภาพใบหน้าคนได้

1.2 วัตถุประสงค์ของโครงการ

สร้างโปรแกรมใช้งาน (Application) โดยใช้ โปรแกรมเมทแผลปและวิธีการของไอเกนเฟซ เพื่อให้คอมพิวเตอร์สามารถรู้จำ บ่งชี้ภาพใบหน้าของแต่ละคนได้ ซึ่งจะรับใบหน้าคนมาจากกล้องวิดีโอที่ต่อกับคอมพิวเตอร์ โดยโปรแกรมจะมีการทำงานดังนี้ คือทำการรับภาพใบหน้าคนจากกล้องวิดีโอ แล้วทำการปรับแต่งภาพให้แต่ละภาพมีองค์ประกอบส่วนที่ไม่ใช่ใบหน้าเหมือนกัน

เพื่อให้เหลือแค่ตัวแปรเดียวคือใบหน้า จัดเก็บแต่ละภาพที่ได้ไว้ในหน่วยความจำของคอมพิวเตอร์ โดยสร้างฐานข้อมูลของใบหน้าแต่ละคนขึ้นมา แล้วดำเนินการตามวิธีของไอเคนเฟซ เพื่อคอมพิวเตอร์ให้รู้จำหน้าคน เมื่อคอมพิวเตอร์รู้จำภาพใบหน้าคนได้แล้วก็ทำการทดสอบการทำงานของโปรแกรมเพื่อให้ใช้งานได้จริง ซึ่งมีรูปแสดงภาพรวมของโครงงานดังรูป 1.1 และ Block diagram แสดงรายละเอียดโดยย่อ ดังรูป 1.2

1.3 ขอบเขตของโครงงาน

- ขอบเขตของการประมวลผลภาพ

เป็นการรับภาพหน้าคนจากกล้องวีดีโอที่ต่อกับคอมพิวเตอร์ และระยะห่างของกล้องวีดีโอกับคนถ่ายต้องอยู่ในระยะที่เหมาะสม ให้เห็นใบหน้าชัดเจน คือต้องอยู่ในลักษณะที่หน้าตรงกับกล้อง ไม่ก้ม หรือ ไม่เอียงหน้า

- ขอบเขตของการจดจำภาพใบหน้า

จำนวนฐานข้อมูลของภาพใบหน้าต้องมีจำนวน 10 คน และภาพที่นำมาทดสอบต้องเป็นภาพในลักษณะต่างๆของบุคคลในฐานข้อมูลนี้ ไม่สามารถบ่งชี้ภาพบุคคลที่ไม่อยู่ในฐานข้อมูลได้

1.4 ประโยชน์และผลที่คาดว่าจะได้รับ

โครงงานนี้ทำให้ได้โปรแกรมใช้งานที่สามารถระบุได้ว่าเป็นภาพใบหน้าใดในฐานข้อมูลนั้น และสามารถนำไปประยุกต์ใช้งานในด้านต่างๆได้ เช่น ระบบรักษาความปลอดภัย หรือใช้งานทางด้านฐานข้อมูลต่างๆ และสามารถนำไปพัฒนาใช้ร่วมกับนิวรอลเน็ตเวิร์ค (Neural Network) เพื่อให้สามารถจดจำได้ดียิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎี และ หลักการเบื้องต้น

2.1 ระบบการมองเห็นภาพ (Vision System)

กลไกระบบการมองเห็นภาพ (Vision System) นั้น จะหมายความรวมถึงทุกสิ่งทุกอย่างที่จำเป็นที่สามารถแปลงภาพนั้นๆ เป็นรหัสดิจิทัล เพื่อนำมาใช้ในระบบคอมพิวเตอร์ได้ , การปรับปรุงเปลี่ยนแปลงข้อมูล และการเสนอภาพที่ได้มาหลังจากการปรับปรุงเปลี่ยนแปลงข้อมูลแล้ว ระบบการมองเห็นภาพนี้ความยุ่งยากจะขึ้นอยู่กับการใช้งาน ซึ่งสามารถแบ่งเป็น 3 ขั้นตอนที่สำคัญ ดังนี้

- 2.1.1 การได้มาซึ่งภาพ (Image Acquisition)
- 2.1.2 กระบวนการประมวลผลภาพ (Image Processing)
- 2.1.3 ผลที่ได้ หรือ การแสดงผล (Output or Display)

ในปัจจุบันได้มีการประยุกต์ใช้งานระบบการมองเห็นอยู่มากมาย เช่น การใช้บาร์โค้ด , การพิมพ์สิ่งพิมพ์ต่างๆ , การประยุกต์ใช้งานในโรงงาน

2.1.1 การได้มาซึ่งภาพ (Image Acquisition)

การได้มาซึ่งภาพ หมายถึง การแปลงภาพในลักษณะทางกายภาพ ให้เป็นเซตของข้อมูลทางดิจิทัล ซึ่งเซตของข้อมูลนี้จะถูกส่งไปยังหน่วยประมวลผลต่อไป ฟังก์ชันของการได้มาซึ่งภาพนี้ แบ่งเป็น 4 เฟส คือ

1. การส่องสว่าง (Illumination)
2. รูปแบบของภาพ หรือ การทำให้ภาพชัดขึ้น (Image formation or Focusing)
3. การตรวจจับภาพ หรือ การรับภาพ (Image detection or Sensing)
4. รูปแบบผลของสัญญาณที่ได้จากกล้อง (Formatting camera output signal)

การส่องสว่างเป็นตัวแปรสำคัญที่มีอิทธิพลต่อสัญญาณอินพุต (Input signal) ที่จะส่งต่อไปให้กับ ระบบการมองเห็นภาพ เพราะฉะนั้นเราจึงต้องออกแบบ ให้มีการส่องสว่างที่เหมาะสมกับการใช้งานที่แตกต่างกัน โดยที่ชนิดและวิธีการส่องสว่างของแหล่งกำเนิดแสง จะมีผลต่อกำลังงานของแสงที่ส่งออกมา ซึ่งจะมีผลต่อกระบวนการประมวลผลภาพ และผลของสัญญาณที่ได้รับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.2 กระบวนการประมวลผลภาพ (Image Processing)

การประมวลผลภาพ คือ การสร้างภาพใหม่โดยการแยกส่วนของข้อมูลที่เราสงสัย กับ สิ่งรบกวน (Noise) ออกจากกัน โดยการทำงานพื้นฐานของการประมวลผลคือ การกำจัดสิ่งรบกวนของภาพ (Noise elimination) , การปรับแต่งขอบภาพให้ดีขึ้น (Edge enhancement) , การกรองภาพ (Filtering) , การปรับปรุงเปลี่ยนแปลงค่าระดับเกรย์ (Grey scale modification) โดยทั่วไปแล้วจะใช้ฮาร์ดแวร์ในการประมวลผลภาพ แต่ถ้ามีการประมวลผลที่ซับซ้อนขึ้น ก็จะใช้ทั้ง ฮาร์ดแวร์ และซอฟต์แวร์ ซึ่งการซับซ้อนของการประมวลผลจะขึ้นอยู่กับลักษณะของการใช้งาน และจะมีวิธีการประมวลผล 3 อย่างคือ

1. ที่จุดเดียวกันในรูปภาพเดียวกัน (Point by point in one image)

คือการสร้างภาพใหม่โดยการเปลี่ยนค่าแบบจุดต่อจุด โดยจุดหนึ่งคือจุดในรูปภาพเดิม และอีกจุดคือจุดในรูปภาพใหม่ที่ได้รับการปรับปรุงเปลี่ยนแปลงจากรูปภาพเดิมแล้ว เช่น การแปลงภาพในระบบเลขฐาน 2 จากจุดที่มีค่าเป็น 0 ถูกเปลี่ยนเป็น 1 และจากค่า 1 ถูกเปลี่ยนกลับเป็น 0

2. ที่จุดเดียวกันในรูปภาพที่แตกต่างกัน (Using corresponding points on different images)

คือการสร้างภาพใหม่โดยการ จับคู่แต่ละจุดจากแหล่งภาพ 2 แหล่ง หรือมากกว่า จุดของรูปภาพที่แตกต่างกันสองจุดนี้ จะนำมาลบกันเพื่อหาค่าที่เปลี่ยนแปลงไป ข้อมูลของจุดในภาพใหม่จะเกิดจากการรวมกันของตัวเลขที่แตกต่างกันของภาพ 2 แหล่งนั้น

3. ที่บริเวณจุดนั้นในภาพนั้น (Using regional points in one image)

คือการสร้างภาพใหม่โดยการหาค่าเฉลี่ยรอบๆจุดนั้นในภาพนั้น . ค่าของจุดในภาพใหม่ คือ ค่าเฉลี่ยของจุดในภาพนั้นอันเดิม

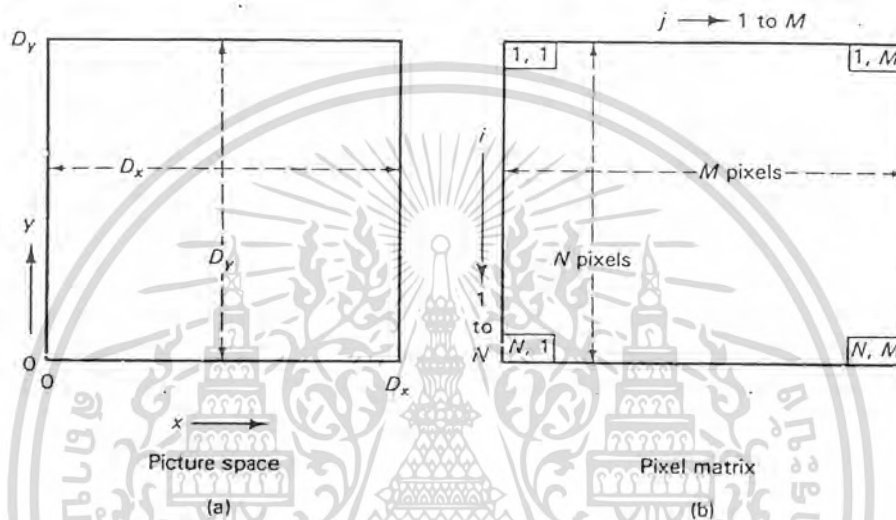
2.1.3 ผลที่ได้และการแสดงผล (Output or Display)

รูปแบบของการแสดงผลนั้นขึ้นอยู่กับว่า จะนำข้อมูลที่ได้ออกไปใช้ต่อในลักษณะใด หรือแสดงผลอย่างไรจึงจะเหมาะสมที่สุด เช่น การแสดงผลโดยการพิมพ์ , การแสดงผ่านหน้าจอ , การแสดงผลในรูปแบบของสัญญาณควบคุม เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 หลักการเบื้องต้นของการประมวลผลภาพ (Image Processing Fundamentals)

2.2.1 พิกเซล (Pixel)



รูปที่ 2.1 พิกเซล a) แสดงลักษณะของภาพ

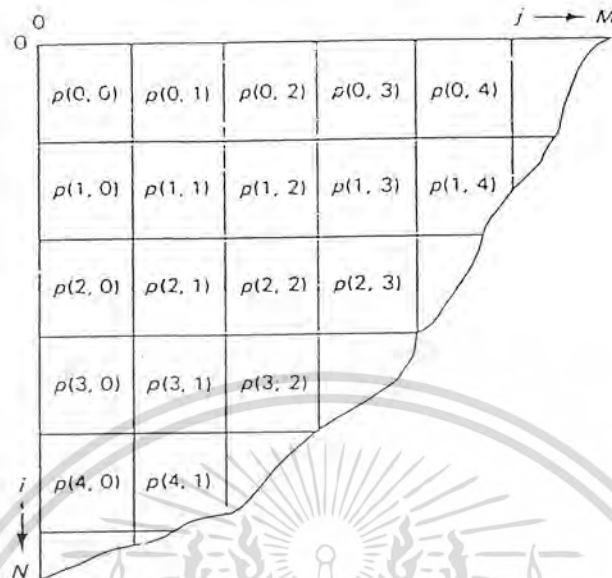
b) แสดงลักษณะของตำแหน่งพิกเซล

เมื่อ $x = D_x/N$ ที่เพิ่มขึ้น , $y = D_y/M$ ที่เพิ่มขึ้น

N = จำนวนพิกเซลที่มากที่สุดในแต่ละแถว , M = จำนวนพิกเซลที่มากที่สุดในแต่ละคอลัมน์

ในภาพหนึ่งๆ เราสามารถอธิบายได้ด้วยเมทริกซ์ของจุดพิกเซลขนาด $N \times M$ โดยใช้คู่ลำดับ $p(i,j)$ แทนค่าของจุดแต่ละจุด โดย i และ j ต้องไม่เป็นจำนวนสเกลาร์ลบ และ $p(i,j)$ นี้จะบ่งชี้ความเข้มของแสงที่จุดนั้นๆของภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.2 คณิตศาสตร์แสดงพิกเซลในเมทริกซ์ภาพหนึ่ง

ค่าที่กำกับแต่ละพิกเซลจะแสดงถึง ค่าเฉลี่ยของความเข้มของแสงในภาพ ที่จุดพิกเซลนั้น แทนอยู่ โดยที่ค่าของพิกเซลนี้จะเขียนแทนด้วย P_{ij} ซึ่งมีค่าตั้งแต่ 0 ถึง 1

2.2.2 ตำแหน่งของพิกเซล (Pixel Location)

ดังที่ได้กล่าวมาแล้วว่า ในภาพหนึ่งๆนั้น เราสามารถแทนอาร์เรย์ (Array) $N \times M$ และค่าในแต่ละจุดพิกเซล จะหมายถึงค่าเฉลี่ยของความเข้มของแสงที่ตกกระทบถึงภาพที่จุดพิกเซลนั้น

ตัวอย่างเช่น ในรูปที่ 2.3 เป็นภาพขนาด 10×10 นิ้ว ถ้าไม่มีแสงตกกระทบบริเวณด้านบนของภาพ แต่มีแสงตกกระทบมากที่บริเวณส่วนล่างของภาพ เราจะใช้ระบบเลขฐานสองแทนค่าความเข้มของการส่องสว่าง คือ 0 แทนบริเวณที่ไม่ถูกแสงเลย และ 1 แทนบริเวณที่มีความเข้มของแสงมากที่สุด

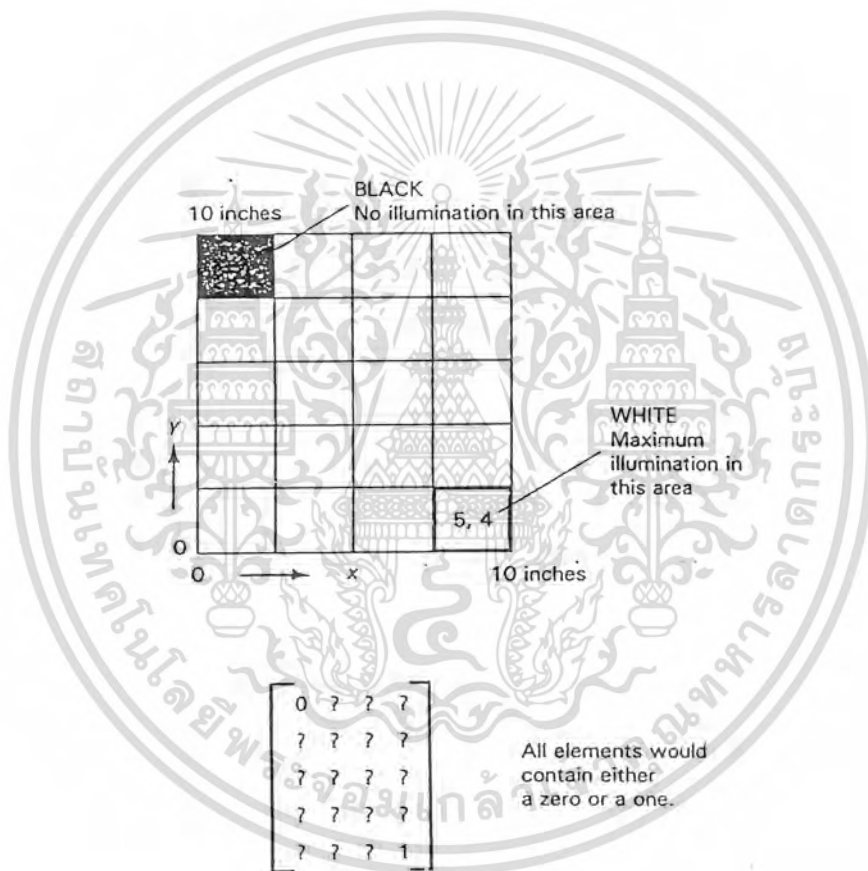
ในรูปที่ 2.3 นี้ เป็นภาพที่ถูกเขียนแทนด้วยเมทริกซ์ขนาด 5×4 (5 แถว, 4 หลัก) แต่แต่ละองค์ประกอบของภาพมีขนาด 2.5×2.0 นิ้ว และค่าในแต่ละองค์ประกอบของภาพนี้จะขึ้นอยู่กับแสงที่ตกกระทบเฉลี่ยบนพื้นที่นั้น

เอกสารนี้เป็นเอกสารของคณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
 บริเวณองค์ประกอบที่มีขนาด 2.5×2.0 นิ้ว ตรงส่วนมุมบนซ้ายของภาพ จะถูกแทนค่า
 ไม่ด้วยตำแหน่ง (1,1) ซึ่งในเมทริกซ์ 5×4 มีค่าเท่ากับ 0 แสดงว่าไม่มีแสงมาตกกระทบที่มีการนำไปใช้

บริเวณองค์ประกอบที่มีขนาด 2.5×2.0 นิ้ว ตรงส่วนมุมขวาล่างของภาพ จะถูกแทนด้วยตำแหน่ง (5,4) มีค่าเท่ากับ 1 แสดงว่ามีความเข้มของการส่องสว่างสูงสุด

ถ้าใช้ระบบ 16 ระดับเกรย์ (16 Gray level system) แทนระบบเลขฐานสอง จะได้ว่าที่จุดพิกเซล (1,1) จะมีค่าเท่ากับ 0 และ ที่จุดพิกเซล (5,4) จะมีค่าเท่ากับ 15

ผู้ออกแบบจะต้องกำหนดค่าเทรชโฮลด์ (Threshold value) ของความเข้มของการส่องสว่าง ที่จะใช้เป็นเกณฑ์ในการเปลี่ยนระดับจาก 0 เป็น 1



รูปที่ 2.3 a) ลักษณะของแสงที่ตกกระจายไม่เท่ากันบนพื้นผิวภาพ
b) ค่าพิกเซลของภาพบนพื้นผิวภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 ระดับเกรย์ (Gray Scale)

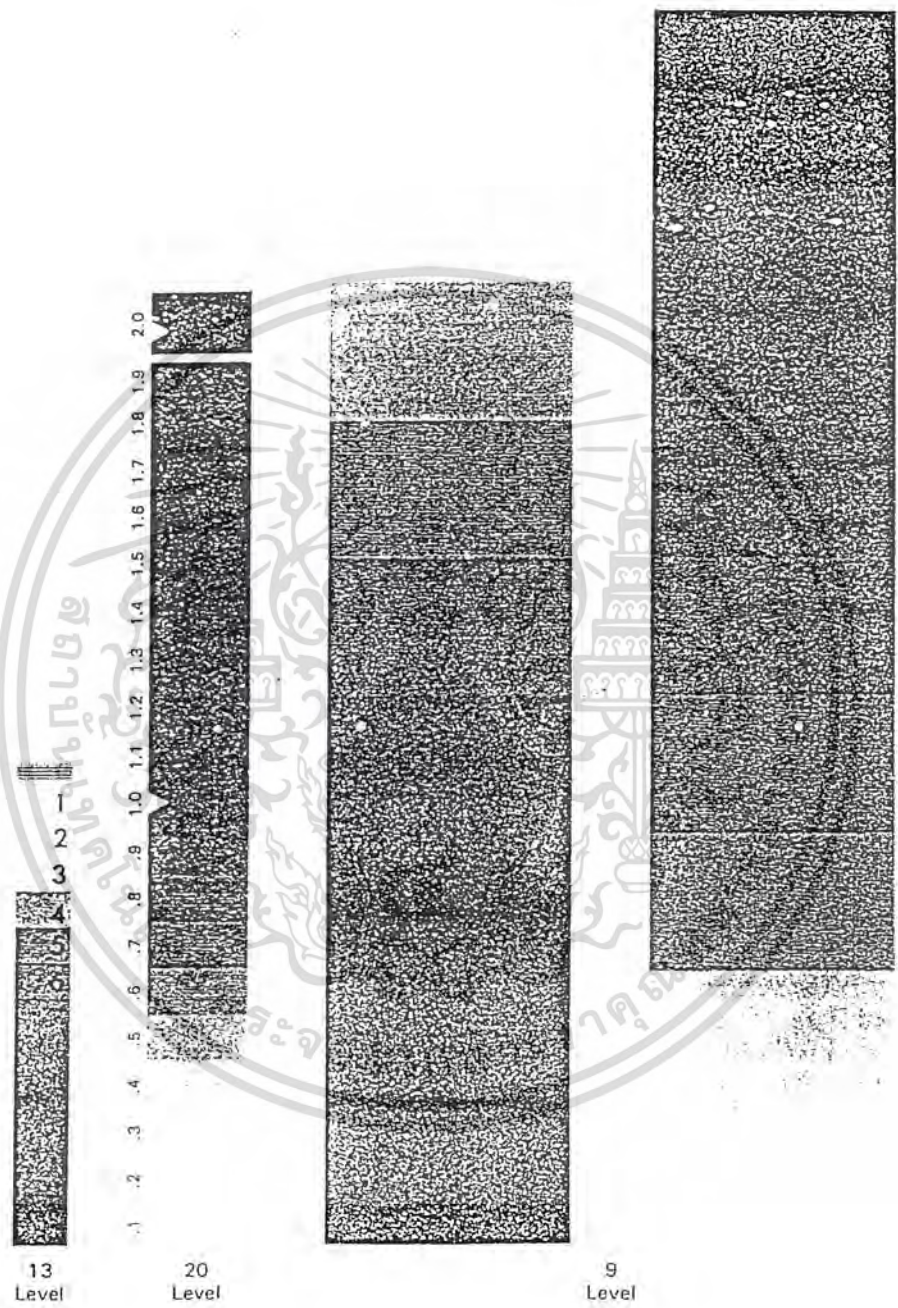
หากเราต้องการค่าของข้อมูลที่ละเอียดมากขึ้น จะต้องเพิ่มจำนวนบิตในการแสดงค่าของแต่ละพิกเซล เช่น ถ้าต้องการแบ่งความเข้มของการส่องสว่างให้มี 4 ระดับ ก็ต้องใช้เลขฐานสอง 2 บิต, 4 บิต สำหรับ 16 ระดับ และ 8 บิต สำหรับ 256 ระดับ ซึ่งจำนวนระดับที่ใช้ในระดับเกรย์นี้มักเป็นเลขยกกำลังของ 2 ส่วนค่าต่ำสุดคือ 0 ถูกกำหนดให้เป็นสีดำ และ 1 หรือตัวเลขที่น้อยกว่าค่าสูงสุดของระดับเกรย์อยู่ 1 (เช่น 15 สำหรับระดับเกรย์ 16 ระดับ) ถูกกำหนดให้เป็นสีขาว ค่าที่กำหนดให้ในแต่ละพิกเซลมักเป็นจำนวนเต็ม

Gray Scale	Gray Value Range
2^1 2 values	0, 1
2^3 8 values	0 to 7
2^4 16 values	0 to 15
2^8 256 values	0 to 255

ในยุคแรกๆของระบบการมองเห็นภาพ (Vision System) จะใช้ระบบเลขฐานสอง แต่ในปัจจุบัน เทคโนโลยีของไมโครโปรเซสเซอร์ (Microprocessors) เข้ามามีบทบาทมากขึ้น ซึ่งในไมโครโปรเซสเซอร์จะมีหน่วยความจำอย่างน้อย 8 บิต เพราะฉะนั้นการแบ่งระดับเป็น 16, 64 หรือ 256 จึงเป็นเรื่องธรรมดา และในตอนนี้อาจจะมีมากกว่า 256 ระดับ แต่ในการมองเห็นของมนุษย์ สามารถแยกความแตกต่างได้เพียง 10 ถึง 15 ระดับเท่านั้น ดังนั้นการแบ่งโดยละเอียดเป็น 64 หรือ 256 ระดับ ทำให้มนุษย์ไม่สามารถแยกความแตกต่างได้ จึงนำไปประยุกต์ใช้กับงานการประมวลผลภาพ หรือ ใช้ในขบวนการทางอุตสาหกรรมต่างๆ

จึงอาจกล่าวได้ว่าจำนวนระดับเกรย์เป็นตัวจำกัดรายละเอียดของภาพ ยิ่งแบ่งระดับเกรย์เป็นหลายระดับ ก็จะยิ่งเพิ่มคุณภาพของภาพนั้นด้วย และการเพิ่มจำนวนพิกเซล เช่น จาก 32×32 เป็น 250×250 ก็จะเป็นการเพิ่มความละเอียด (Resolution) และรายละเอียด (Detail) ของภาพ ซึ่งจะแตกต่างกับการขยายภาพ (Zoom) เพราะการขยายภาพ เป็นการเพิ่มขยายของแต่ละพิกเซลให้ใหญ่ขึ้น ไม่ได้เป็นการเพิ่มจำนวนพิกเซล แต่การแบ่งระดับเกรย์เป็นการเพิ่มจำนวนของพิกเซล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.4 ตัวอย่างของระดับเกรย์ (Gray Scales)
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4 ฮิสโตแกรม (Histogram)

ฮิสโตแกรมเป็นการแสดงให้เห็นถึงความถี่ของการนับจำนวนพิกเซล ที่มีค่าความเข้มแต่ละค่าหนึ่งในภาพระดับเกรย์ จากรูป 2.5 แกน x ในกราฟแสดงค่าระดับเกรย์ และ แกน y แสดงค่าจำนวนพิกเซลในระดับเกรย์นั้น โดยฮิสโตแกรมมีขั้นตอนการสร้างดังนี้

1. ทำการดิจิไทซ์ (Digitizing) ภาพ
2. นับจุดพิกเซลในแต่ละระดับเกรย์
3. พล็อตกราฟระหว่างจำนวนจุดกับระดับเกรย์



รูปที่ 2.5 ฮิสโตแกรมของ 8 ระดับเกรย์ จากเมทริกซ์ 4×5

รูปร่างของฮิสโตแกรมสามารถบอกลักษณะบางประการของภาพได้ เช่น ถ้าฮิสโตแกรมมีลักษณะที่แคบ จะหมายถึงขาดการคอนทราสต์ (contrast , การแยกแยะความผิดเพี้ยนของสี) ในภาพนั้น

ฮิสโตแกรมมีประโยชน์ในการกำหนดระดับเทรชโฮลด์ (Threshold) ซึ่งก็คือการเปลี่ยนภาพระดับเกรย์ให้เป็นภาพระบบเลขฐานสอง (Binary) หรือเพื่อปรับปรุงบางส่วนของสเปกตรัม (Spectrum) ของระดับเกรย์

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

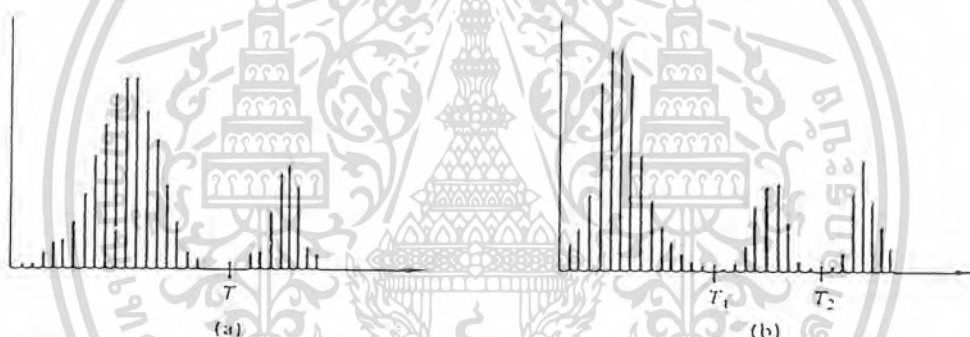
2.5 การแปลงระดับเกรย์ (Gray-level Transformation)

1. การเปลี่ยนแปลงความสว่าง (Global Alternative in brightness)

เป็นการใช้ค่าคงที่บวกหรือลบออกจากทุกพิกเซลของภาพ เพื่อเพิ่มหรือลดความสว่างของภาพ

2. การทำเทรชโฮล (Threshold)

เป็นการเปลี่ยนแปลงหรือหาแนวโน้มของค่าระดับเกรย์ในภาพเพื่อทำให้เป็นดิสครีท (Discreet) มากขึ้น โดยจะนำภาพมาทำฮิสโตแกรม แล้วกำหนดค่าระดับเกรย์ที่แน่นอนขึ้นมาเพื่อที่จะเป็นจุดตัดทำเทรชโฮล จากนั้นจะทำการตัดหรือปิดส่วนของระดับเกรย์ที่เราไม่ต้องการออกไป ดังรูปที่ 2.6



รูปที่ 2.6 ฮิสโตแกรมของระดับเกรย์ a) ใช้เทรชโฮล 1 จุด (Single Threshold)
b) ใช้เทรชโฮลหลายจุด (Multiple Threshold)

ประโยชน์ของเทรชโฮล ได้แก่ การทำให้เป็นภาพแบบไบนารี หากดูจากรูปที่ 2.5 a) จุดที่อยู่ในช่วงทางซ้ายมือทั้งหมดของจุดตัดเทรชโฮล T จะถูกทำเป็นสีขาว และจุดในช่วงทางขวามือทั้งหมดจะถูกทำให้เป็นสีดำ หรือ การช่วยให้หาขอบของภาพได้ง่ายขึ้น เป็นต้น แต่ทั้งนี้ก็เป็นกรยากที่จะกำหนดจุดตัดเทรชโฮลที่ดีที่สุดออกมาได้

3. บันช์ซิ่ง (Bunching) และการควอนไทซ์ (Quantize)

ในบางครั้ง การบันช์ซิ่ง จะหมายรวมถึง การควอนไทซ์ ด้วย หรือ ใช้ลด

ระดับเกรย์ของภาพที่ไม่ต้องการลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
หมายเหตุ การควอนไทซ์ (Quantizing) และความผิดพลาด
ไม่ว่ากรณีใดๆ ทั้งสิ้น อภัยโทษให้แก่มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี และขออภัยต่อเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การควอนไทซ์เป็นการประมาณค่าระดับเกรย์ให้เป็นจำนวนเต็ม ซึ่งจะนำไปตามกฎ คือ ถ้าเป็นจุดทศนิยมให้ปัดขึ้นทั้งหมด และการปัดนี้ก็คือความผิดพลาดที่เกิดขึ้น

4. สปลิตติง (Splitting)

เป็นการเพิ่มความแตกต่างระหว่าง 2 กลุ่มของระดับเกรย์ เช่น ถ้าเรามีตัวอักษรเขียนอยู่บนฉาก โดยตัวเลขมีระดับเกรย์ที่ 98 , ฉากมีระดับเกรย์ที่ 99 ซึ่งตาของมนุษย์ไม่สามารถแยกแยะความแตกต่างเพียงเท่านี้ได้ และจะได้ฮิสโตแกรมที่มีลักษณะแบน เพราะฉะนั้นจึงแก้ปัญหาด้วยการสปลิตฮิสโตแกรม โดยทำการดึงค่า 99 ขึ้นเป็น 120 และดึงค่า 98 เป็น 80 ซึ่งจะทำให้ระหว่างตัวเลขและตัวอักษรมีความแตกต่างกันมากขึ้นจนสามารถสังเกตแยกแยะได้

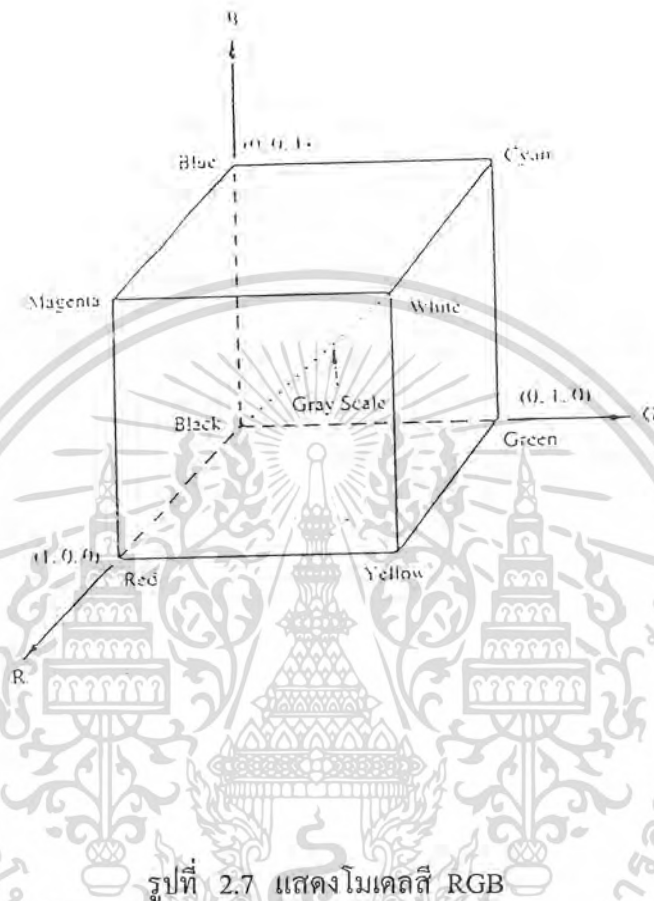
เทคนิคนี้จะมีประโยชน์มาก ในกรณีที่เราต้องการดึงเอาเฉพาะบางส่วนของภาพออกมา

2.6 พื้นฐานและระบบของสีโมเดล RGB (Red , Green , Blue)

ในโมเดลนี้ สีแต่ละสีจะอยู่ในรูปของสีปฐมภูมิ (แดง , เขียว , น้ำเงิน) โมเดลนี้มีโครงสร้างอยู่ในระบบของพิกัดคาร์ทีเซียน (Cartesian coordinate) มีลักษณะเป็นทรงลูกบาศก์ ดังรูปที่ 2.7 ค่าสี แดง , เขียว , น้ำเงิน จะอยู่ที่มุมทั้งสาม บนเส้นทแยงมุมตรงข้ามกัน และ ค่าสี คราม , ฟ้า , เหลือง จะอยู่ที่มุมทั้งสามในลักษณะเดียวกัน ส่วนสีดำจะอยู่ที่จุดกำเนิด สีขาวอยู่ที่มุมที่มีระยะทางไกลที่สุดจากแหล่งกำเนิด ค่าของระดับสีเทา (Gray Scale) จะอยู่บนเส้นที่เชื่อมระหว่างสีดำกับสีขาว ค่าสี คือ จุดที่อยู่บนผิวหรือในลูกบาศก์ถูกกำหนดค่าโดยเวกเตอร์ที่ชี้ออกจากจุดกำเนิด เพื่อความสะดวกเราจะสมมติให้ค่าสีถูกนอร์มอลไลซ์ (Normalize) ทั้งสามสี โดยให้มีค่าอยู่ในช่วง 0 ถึง 1 ลูกบาศก์ที่แสดงในรูปที่ 2.7 นี้ จึงเป็นลูกบาศก์หนึ่งหน่วย

ภาพในโมเดล RGB ประกอบด้วยภาพสามระนาบที่เป็นอิสระจากกัน สำหรับแต่ละสีปฐมภูมิเมื่อป้อนเข้าไปในมอนิเตอร์ (monitor) ที่เป็นแบบ RGB ภาพทั้งสามสีจะรวมตัวกันที่จอภาพกลายเป็นภาพสีผสม ดังนั้นการใช้โมเดล RGB ในการประมวลผลภาพนั้นจะสมเหตุสมผลเมื่อภาพถูกแยกออกโดยธรรมชาติให้อยู่ในเทอมของทั้งสามสี กล้องภาพสีส่วนใหญ่ที่ให้ภาพสีดิจิทัลจะอยู่ในรูปแบบของโมเดล RGB ดังนั้น โมเดลนี้จึงเป็น โมเดลที่สำคัญมากในการประมวลผลภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.7 แสดงโมเดลสี RGB

2.7 บิตแมป (Bitmaps)

บิตแมปเป็นอาร์เรย์ของบิตเพื่อแสดงลักษณะของจุดพิกเซล (Pixel) ในที่นี้จะกล่าวถึงรูปแบบข้อมูลบิตแมปบนวินโดวส์ (window) อย่างง่าย ซึ่งมีด้วยกัน 2 ลักษณะคือ GDI bitmaps และ DIBs bitmaps

ในส่วนของ GDI bitmaps นั้นถูกนำเสนอโดย Microsoft Foundation Class Library version 2.0 Cbitmap Class ตัวบิตแมปออปเจกจะเกี่ยวข้องกับโครงสร้างข้อมูลของวินโดวส์ ซึ่งถูกจัดการภายใต้โมดูลวินโดวส์ GDI นั้นแสดงให้เห็นว่ามันขึ้นอยู่กับอุปกรณ์ที่ใช้งาน ในโปรแกรมเราสามารถทำสำเนาข้อมูลบิตแมป แต่การจัดเรียงตัวของแต่ละบิตนั้น จะขึ้นอยู่กับฮาร์ดแวร์ที่ใช้งาน GDI bitmaps จะถูกรับส่งอย่างอิสระระหว่างโปรแกรมประยุกต์ภายในคอมพิวเตอร์เครื่องๆหนึ่งได้ แต่ด้วยลักษณะที่ขึ้นกับอุปกรณ์ที่ใช้งาน ดังนั้นการส่งผ่าน โมเด็มหรือดิสก์จึงไม่เหมาะสม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่ภายนอก
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

DIB bitmaps เป็นอีกรูปแบบหนึ่งของข้อมูลบิตแมปซึ่งสามารถแก้ไขปัญหาในส่วนของ การแลกเปลี่ยนข้อมูลระหว่างกัน คอมพิวเตอร์เครื่องใดก็ตามที่ใช้งานวินโดวส์สามารถที่จะ ประมวลผล DIB ได้ ซึ่งโดยทั่วไปมักเก็บอยู่ในไฟล์ฟอร์แมต BMP เมื่อไฟล์ BMP ถูกอ่านขึ้นมาจาก ดิสก์ มันจะทำการแปลงให้เป็น GDI bitmap แต่โปรแกรมสามารถทำงานโดยตรงกับรูปแบบ DIB ได้เลยถ้าจำเป็น

2.8 ความสว่าง (Brightness)

ความสว่างเป็นคุณสมบัติของแสงที่มนุษย์เป็นผู้กำหนดให้ระหว่างความสว่างมากที่สุดกับความมืด ความสว่างเป็นสิ่งที่รับรู้ได้แต่ไม่สามารถสัมผัสได้ ความสว่างสามารถรับรู้ได้เมื่อมีแสงตกกระทบกับเรตินารูปแท่งและรูปกรวยในตาของเรา ซึ่งจะก่อให้เกิดการตอบสนองที่ไม่เป็นเชิงเส้นและสลับซับซ้อน ความไวในการรับแสงของเราจะลดลงเมื่อแสงมีขนาดใหญ่มากขึ้น สำหรับความสว่างที่อยู่ในรูปภาพนั้น สามารถแสดงได้โดย Brightness Histogram ซึ่งเป็นกราฟการกระจายตัวของกลุ่มตัวเลข ในกราฟนี้จะแสดงถึงระดับสีเทาของแต่ละพิกเซลที่อยู่ในรูปภาพ หรืออาจจะกล่าวได้ว่าเป็นการแสดงให้เห็นถึงจำนวนของพิกเซลในรูปภาพที่อยู่ในแต่ละระดับของสีเทาว่ามีอยู่เป็นจำนวนเท่าไร

2.9 ชนิดของรูปภาพในกล่องเครื่องมือของโปรแกรมแมทเลบ

เครื่องมือในการประมวลผลรูปภาพรองรับรูปแบบพื้นฐาน 4 อย่างของรูปภาพคือ

- ดัชนีรูปภาพ (Index images)
- ความหนาแน่นของรูปภาพ (Intensity images)
- ไบนารีของรูปภาพ (Binary images)
- อาร์จีบีของรูปภาพ (RGB images)

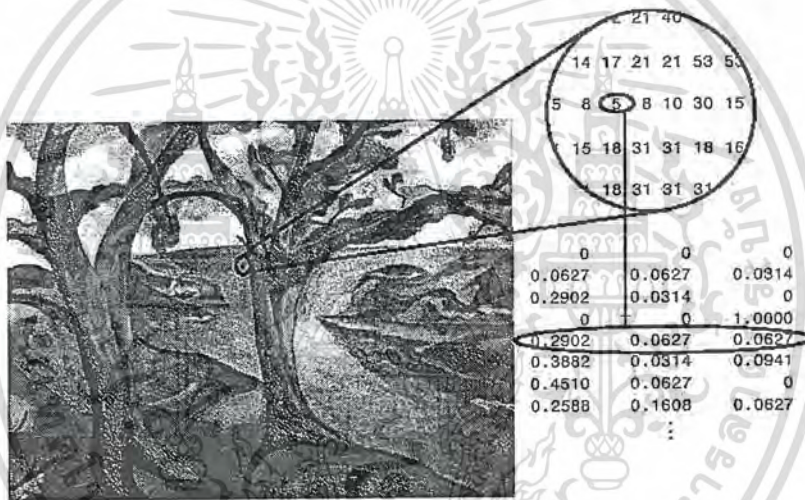
2.9.1 ดัชนีรูปภาพ

ดัชนีรูปภาพประกอบไปด้วยเมทริกซ์ของข้อมูล X และเมทริกซ์ของสีที่จะนำไปเทียบเคียง เมทริกซ์ของข้อมูลสามารถแบ่งออกเป็นคลาสของ uint8, uint16, หรือ double ส่วนเมทริกซ์ของสี นั้นจะเป็นอาร์เรย์ขนาด $m \times 3$ ของคลาส double ที่ประกอบด้วยเลขทศนิยมที่อยู่ในช่วง $[0,1]$ แต่ละแถวของการเทียบเคียงนั้นจะระบุคอมโพเนนต์ สีแดง, สีเขียว และสีน้ำเงินของแต่ละสีนั้น ดัชนีรูป

เอกสารนี้จะใช้การเทียบเคียงโดยตรงของแต่ละค่าในพิกเซลนั้น ไปยังค่าเทียบเคียงสีในพิกเซลนั้น ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เซตบนรูปภาพนั้นจะถูกกำหนดโดยใช้ค่าที่มีลักษณะเช่นเดียวกันกับ X เป็นดัชนีในการเทียบเคียง ค่าของจุดแรกแทนแถวแรกในการเทียบเคียง ค่าของจุดที่สองแทนแถวที่สอง และเป็นเช่นนี้เรื่อยไป

การเทียบเคียงสีมักจะถูกเก็บด้วยดัชนีรูปภาพและจะถูกนำมาใช้อย่างอัตโนมัติกับรูปภาพเมื่อเรียกใช้ฟังก์ชัน `imread` อย่างไรก็ตามในการใช้งานนั้น ไม่ได้ถูกจำกัดการใช้งานเฉพาะค่าพื้นฐานของการเทียบเคียงสีเท่านั้น แต่ยังสามารถเรียกใช้งานค่าใดๆในการเทียบเคียงสีได้ รูปภาพดังต่อไปนี้แสดงถึงโครงสร้างของดัชนีรูปภาพ พิกเซลที่อยู่ในรูปภาพจะถูกแสดงด้วยเลขจำนวนเต็ม ซึ่งจะชี้ไปยังค่าของสีที่เก็บอยู่ในการเทียบเคียงสีนั้น



รูปที่ 2.8 ค่าของพิกเซลชี้ไปยังค่าเทียบเคียงสีในดัชนีรูปภาพ

- คลาสและค่าออฟเซตของการเทียบเคียงสี

ความสัมพันธ์ระหว่างค่าที่อยู่ในเมทริกซ์รูปภาพกับค่าในการเทียบเคียงสีนั้นขึ้นอยู่กับคลาสของเมทริกซ์รูปภาพถ้าเมทริกซ์รูปภาพเป็นคลาสของ `double` ค่าของจุดแรกจะชี้ไปยังแถวแรกของการเทียบเคียง ค่าของจุดที่สองจะชี้ไปยังแถวที่สอง และเป็นเช่นนี้เรื่อยไป ถ้าเมทริกซ์รูปภาพเป็นคลาสของ `uint8` หรือ `uint16` มันจะเป็นค่าออฟเซต ค่าของจุด 0 จะชี้ไปยังแถวแรกของการเทียบเคียง ค่าของจุดที่หนึ่งจะชี้ไปยังแถวที่สอง และเป็นเช่นนี้เรื่อยไป

ค่าออฟเซตมักจะถูกใช้ในไฟล์ที่มีรูปแบบเป็นกราฟฟิกเพื่อจะทำให้จำนวนของสีที่สามารถรองรับได้มีมากที่สุด จากภาพข้างบนนี้ เมทริกซ์รูปภาพเป็นคลาส `double` เพราะมันไม่มีออฟเซต ดังนั้นค่าของจุดที่ห้าจะชี้ไปยังแถวที่ห้าของการเทียบเคียง

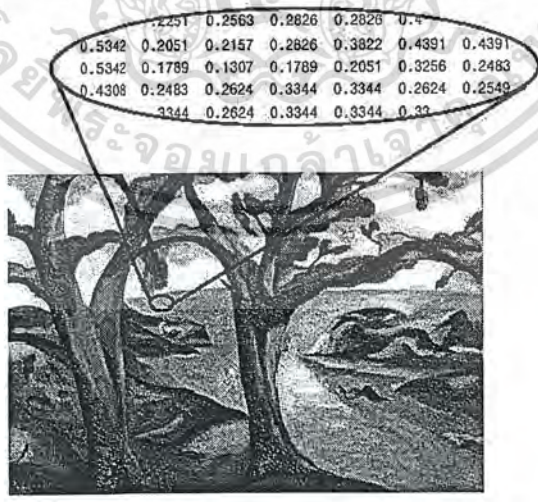
● ข้อจำกัดของการใช้ uint16

ในกล่องเครื่องมือนี้จะมีข้อจำกัดของการใช้ดัชนีรูปภาพที่เป็น คลาส uint16 เราสามารถที่จะอ่านรูปภาพนี้เข้าไปยังโปรแกรมเมทแลป และแสดงมันได้ แต่ก่อนที่จะทำกระบวนการของดัชนีรูปภาพแบบ uint16 จะต้องทำการเปลี่ยนให้เป็น double หรือ uint8 เสียก่อน ในการที่จะเปลี่ยนเป็น double ให้เรียกใช้ im2double เพื่อที่จะลดรูปภาพไปเป็น 256 สีหรือน้อยกว่านั้น (uint8) เรียก imapprox

2.9.2 ความหนาแน่นของรูปภาพ

ความหนาแน่นของรูปภาพคือเมทริกซ์ข้อมูล I ซึ่งค่าของมันนั้นจะแทนความหนาแน่นภายในระยะหนึ่งๆ โปรแกรมเมทแลปจะเก็บค่าความหนาแน่นของรูปภาพเป็นเมทริกซ์เดี่ยว โดยที่แต่ละค่าในเมทริกซ์ นั้นจะมีลักษณะเช่นเดียวกับหนึ่งพิกเซลในรูปภาพ เมทริกซ์สามารถเป็นคลาสของ double, uint8 หรือ uint16

ค่าที่อยู่ในเมทริกซ์ความหนาแน่นจะแทนความหนาแน่นที่หลากหลาย หรือระดับของสีเทา (Gray levels) โดยค่าความหนาแน่น 0 มักจะแทนด้วยสีดำ และค่าความหนาแน่น 1, 255 หรือ 65535 มักแทนด้วยสีขาว



รูปที่ 2.9 ค่าพิกเซลในความหนาแน่นของรูปภาพกำหนดระดับของสีเทา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.9.3 ไบนารีของรูปภาพ

ในไบนารีของรูปภาพ แต่ละพิกเซลจะสมมติให้เป็นหนึ่งในสองค่าที่ไม่ต่อเนื่อง ค่าสองค่านี้จะเป็นลักษณะของ เปิด หรือ ปิด ไบนารีของรูปภาพจะถูกเก็บในเมทริกซ์สองมิติของ 0 (ปิดพิกเซล) และ 1 (เปิดพิกเซล)

ไบนารีของรูปภาพถูกพิจารณาเป็นชนิดพิเศษของความหนาแน่นของรูปภาพ ซึ่งประกอบด้วยสีขาวและดำเท่านั้น อย่างไรก็ตาม เราสามารถคิดถึงไบนารีของรูปภาพเป็นลักษณะของดัชนีรูปภาพที่มีเพียงสองสีเท่านั้นก็ได้

ไบนารีของรูปภาพจะถูกเก็บเป็นอาร์เรย์ของคลาส double หรือ uint8 (ในกล่องเครื่องมือไม่รองรับไบนารีของรูปภาพที่เป็นคลาส uint16) อาร์เรย์ของคลาส uint8 โดยทั่วไปแล้วจะถูกนำมาใช้มากกว่าอาร์เรย์ของคลาส double เพราะ uint8 ใช้หน่วยความจำน้อยกว่ามาก ในกล่องเครื่องมือการประมวลผลรูปภาพ ฟังก์ชันใดๆที่คืนค่ากลับมาเป็นไบนารีของรูปภาพจะคืนค่ากลับมาเป็นอาร์เรย์ของ uint8 ถ้าแฟล็กเปิดใช้งานอยู่ ระยะเวลาของข้อมูลจะเป็น [0,1] ถ้าแฟล็กปิดใช้งาน ระยะเวลาข้อมูลเป็น [0,255]



รูปที่ 2.10 พิกเซลในไบนารีของรูปภาพมีค่าที่เป็นไปได้สองค่า คือ 0 หรือ 1

2.9.4 อาร์จีบีของรูปภาพ

อาร์จีบีของรูปภาพ หรืออาจเรียกเป็น สีแท้จริงของรูปภาพ (truecolor image) จะถูกเก็บในเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับครูใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่โดยไม่ขออนุญาต การค้า โปรแกรมเมทแลป เป็นอาร์เรย์ $m \times n \times 3$ ซึ่งจะกำหนดค่านี้นั้นที่สีแดง, สีเขียวและสีน้ำเงิน ไม่วากรณ์ใดๆ พงสน์ อีกทั้งห้ามมีเหตุดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับแต่ละพิกเซล อาร์จีบีของรูปภาพไม่ได้ใช้แพทเทจของสี สีในแต่ละพิกเซลจะถูกกำหนดโดยการรวมกันของความหนาแน่นของสีแดง, สีเขียว และสีน้ำเงินที่เก็บอยู่ในแต่ละระนาบของสีที่ตำแหน่งของแต่ละพิกเซล รูปแบบไฟล์กราฟฟิกจะเก็บอาร์จีบีของรูปภาพเป็น 24 บิต ซึ่งแต่ละคอมโพเนนท์ของสีแดง, สีเขียว และสีน้ำเงินจะมีขนาด 8 บิต ซึ่งจะทำให้เกิดจำนวนของสีที่เป็นไปได้ 16 ล้านสี เนื่องจากความละเอียดนี้เอง เราจึงสามารถแทนรูปภาพต่างๆที่พบเห็นได้ในชีวิตจริงด้วยรูปภาพแบบสีแท้จริง

อาร์เรย์อาร์จีบีใน MATLAB สามารถแบ่งเป็นคลาสของ double, uint8 หรือ uint16 ในอาร์เรย์อาร์จีบีที่เป็นคลาส double แต่ละคอมโพเนนท์ของสีจะมีค่าอยู่ระหว่าง 0 กับ 1 พิกเซลที่มีค่าคอมโพเนนท์ของสีเป็น (0,0,0) จะแสดงเป็นสีดำ และพิกเซลที่มีค่าคอมโพเนนท์ของสีเป็น (1,1,1) จะแสดงเป็นสีขาว คอมโพเนนท์ของสีทั้งสามสำหรับแต่ละพิกเซลนั้นจะถูกเก็บเป็นอาร์เรย์ขนาดสามมิติ ตัวอย่างเช่น คอมโพเนนท์ของสีแดง, เขียว และน้ำเงินของพิกเซล (10,5) จะถูกเก็บในอาร์จีบี (10,5,1), อาร์จีบี (10,5,2) และอาร์จีบี (10,5,3) ตามลำดับ

0.2235	0.1294	Blue	0.4190	0.2588	0.2588
0.5804	0.2902	0.0627	0.2902	0.2902	0.4824
0.1608	0.0627	0.0627	0.0627	0.2235	0.2588
0.5176	0.1922	0.0627	Green	0.1922	0.2588
0.5176	0.1294	0.1608	0.1294	0.1294	0.2588
0.5176	0.1608	0.0627	0.1608	0.1922	0.2588
0.5490	0.2235	0.5490	Red	0.7412	0.7765
0.5490	0.3882	0.5176	0.5804	0.5804	0.7765
0.4190	0.2588	0.2902	0.2588	0.2235	0.4824
0.2235	0.1608	0.2588	0.2588	0.1608	0.2588
0.2588	0.1608	0.2588	0.2588	0.2588	0.2588



รูปที่ 2.11 ระนาบสีของรูปภาพอาร์จีบี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการกำหนดสีของพิกเซลที่จุด (2,3) จะต้องดูที่อาร์เรย์ที่เก็บใน (2,3,1:3) สมมติว่า (2,3,1) ประกอบไปด้วยค่า 0.5176 (2,3,2) ประกอบไปด้วยค่า 0.1608 และ(2,3,3) ประกอบไปด้วยค่า 0.0627 ดังนั้นสีของพิกเซลที่จุด (2,3) คือ

0.5176 0.1608 0.0627

เพื่อที่จะแสดงหลักการของระนาบสีสามสีในอาร์เรย์ของรูปภาพนั้น โค้ดที่อยู่ข้างล่างนี้เป็นตัวอย่างของการสร้างอาร์เรย์ของรูปภาพที่ประกอบด้วยการไม่รบกวนกันของพื้นที่ในสีแดง, สีเขียว และสีน้ำเงิน จากนั้นจึงสร้างรูปภาพสำหรับแต่ละระนาบของสี (แดง,เขียวและน้ำเงิน) โดยจะแสดงแต่ละระนาบของสีที่แยกจากกัน และยังแสดงถึงรูปภาพต้นฉบับด้วย

```
RGB = reshape(ones(64,1) × reshape(jet(64),1,192),[64,64,3]);
```

```
R = RGB(:,:,1);
```

```
G = RGB(:,:,2);
```

```
B = RGB(:,:,3);
```

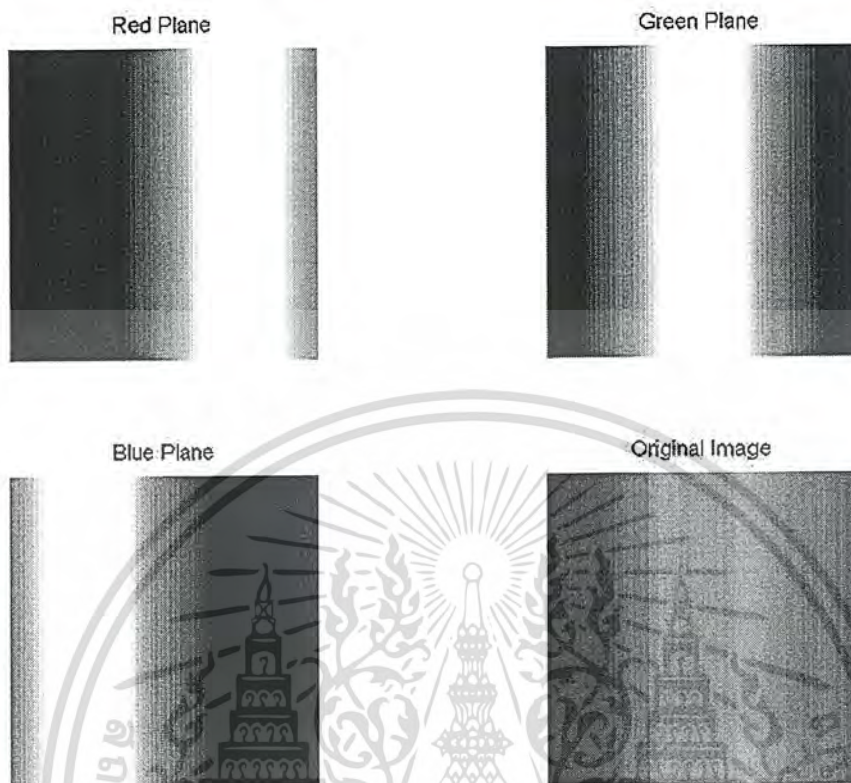
```
imshow(R)
```

```
figure, imshow(G)
```

```
figure, imshow(B)
```

```
figure, imshow(RGB)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.12 ระนาบสีที่แยกจากกันของรูปภาพอาร์จีบี

ในแต่ละสีระนาบสีที่แยกจากกันในรูปภาพประกอบไปด้วยพื้นที่ของสีขาว สีขาวนี้จะแสดงถึงค่าสูงสุดของแต่ละสีที่แยกจากกัน ตัวอย่างเช่น ในรูปภาพที่เป็นระนาบของสีแดง สีขาวจะแสดงถึงค่าสูงสุดของสีแดง ในขณะที่สีแดงถูกผสมด้วยสีเขียวหรือสีน้ำเงิน พิกเซลที่เป็นสีเทาจะปรากฏออกมา ส่วนพื้นที่ที่เป็นสีดำในรูปภาพนั้นจะแสดงค่าของพิกเซลที่ไม่มีค่าของสีแดงอยู่เลย เช่น $R = 0$

2.10 รูปภาพที่เป็นมัลติเฟรมอาร์เรย์

ในโปรแกรมประยุกต์บางตัวนั้น จำเป็นต้องทำงานร่วมกับรูปภาพจำนวนมากที่เกี่ยวข้องกับเวลาของการเปิดดู เช่น Magnetic Resonance Imaging (MRI) หรือ เฟรมของภาพยนตร์เป็นต้น

ในกล่องเครื่องมือการประมวลผลภาพนั้นรองรับการเก็บรูปภาพจำนวนมากในอาร์เรย์เดียวกัน แต่ละรูปภาพที่แยกจากกันจะเรียกว่าเฟรม ถ้าอาร์เรย์รองรับหลายๆเฟรม มันจะถูกนำมาต่อกัน

เป็นขนาด 4 มิติ เช่น อาร์เรย์ของภาพห้าภาพที่มีอาร์จีบีแบบ 400×300 จะเป็น $400 \times 300 \times 3 \times 5$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า สำหรับมัลติเฟรมที่มีความหนาแน่นหรือดัชนีรูปภาพที่เหมือนกันจะเป็น $400 \times 300 \times 1 \times 5$

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมีเหตุดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยการใช้คำสั่ง `cat` ในการเก็บรูปแต่ละรูปที่แยกจากกันในไฟล์ของมัลติเฟรม เช่น ถ้ามีกลุ่มของรูปภาพ A1, A2, A3, A4 และ A5 เราสามารถที่จะเก็บทั้งหมดนี้ภายในอาร์เรย์เดียวโดยใช้คำสั่ง

$$A = \text{cat}(4, A1, A2, A3, A4, A5)$$

เมื่อต้องการที่จะแยกแต่ละเฟรมออกจากรูปภาพที่เป็นมัลติเฟรม ตัวอย่างเช่น มีรูปภาพที่เป็นมัลติเฟรม `MULTI` คำสั่งต่อไปนี้จะทำการแยกเฟรมที่สามออกมา

$$\text{FRM3} = \text{MULTI}(:, :, 3)$$

ในรูปภาพที่เป็นมัลติเฟรมอาร์เรย์นั้น แต่ละรูปต้องมีขนาดที่เท่ากันและมีจำนวนของระนาบที่เท่ากันด้วย ในดัชนีรูปภาพที่เป็นมัลติเฟรม แต่ละรูปจะต้องใช้การเทียบเคียงแบบเดียวกัน

2.10.1 ข้อจำกัดในการใช้มัลติเฟรม

ฟังก์ชันจำนวนมากในกล่องเครื่องมือนี้จะทำงานเฉพาะสองหรือสามมิติแรกเท่านั้น เราสามารถใช้อาร์เรย์ขนาดสี่มิติด้วยฟังก์ชันต่อไปนี้ได้ แต่จำเป็นต้องทำงานแต่ละเฟรมโดยแยกจากกัน เช่น ตัวอย่างนี้เป็นการแสดงเฟรมที่เจ็ดในอาร์เรย์ `MULTI`

$$\text{imshow}(\text{MULTI}(:, :, :, 7))$$

ถ้าเราส่งอาร์เรย์ไปยังฟังก์ชันและอาร์เรย์นั้นมีมิติมากกว่าการทำงานที่ฟังก์ชันกำหนดไว้ ผลลัพธ์ที่เกิดขึ้นไม่อาจคาดเดาได้ ในบางกรณีฟังก์ชันจะทำงานในเฟรมแรกของอาร์เรย์อย่างถูกต้อง แต่ในกรณีอื่นๆผลลัพธ์ที่เกิดจากการทำงานนั้นอาจไม่มีความหมายอะไร

2.11 สรุปรูปชนิดของรูปภาพและคลาสต่างๆ

ตารางนี้เป็นการสรุปวิธีที่ โปรแกรมเมทแพล แปลข้อมูลเมทริกซ์เป็นสีของพิกเซล ซึ่งขึ้นอยู่กับชนิดของรูปภาพและคลาสที่ใช้ในการเก็บ

ชนิดรูปภาพ	คลาส <code>double</code>	คลาส <code>uint8</code> หรือ <code>uint16</code>
ไบนารี	รูปภาพเป็นอาร์เรย์ขนาด $m \times n$ ของหนึ่งกับศูนย์ ซึ่งแฟลทจะเปิดใช้งาน	รูปภาพเป็นอาร์เรย์ขนาด $m \times n$ ของหนึ่งกับศูนย์ ซึ่งแฟลทจะเปิดใช้งาน กล่องเครื่องมือไม่รองรับรูปภาพไบนารีที่เป็น <code>uint16</code>
ดัชนี	รูปภาพเป็นอาร์เรย์ขนาด $m \times n$ ของเลขจำนวนเต็มที่อยู่ในช่วง $[1, p]$	รูปภาพเป็นอาร์เรย์ขนาด $m \times n$ ของเลขจำนวนเต็มที่อยู่ในช่วง $[0, p-1]$

	การเทียบเคียงสีเป็นอาร์เรย์ขนาด $p*3$ ของเลขทศนิยมซึ่งค่าอยู่ในช่วง $[0,1]$	การเทียบเคียงสีเป็นอาร์เรย์ขนาด $p*3$ ของเลขทศนิยมซึ่งค่าอยู่ในช่วง $[0,1]$
ความหนาแน่น	รูปภาพเป็นอาร์เรย์ขนาด $m*n$ ของ	รูปภาพเป็นอาร์เรย์ขนาด $m*n$ ของเลข

ชนิดรูปภาพ	คลาส double	คลาส uint8 หรือ uint16
	เลขทศนิยมซึ่งค่าเป็นความสัมพัทธ์แบบเชิงเส้น โดยทั่วไปค่าอยู่ในช่วง $[0,1]$ การเทียบเคียงสีเป็นอาร์เรย์ขนาด $p*3$ ของเลขทศนิยม ซึ่งค่าอยู่ในช่วง $[0,1]$ และเป็นระดับของสีเทา	จำนวนเต็มซึ่งค่าเป็นความสัมพัทธ์แบบเชิงเส้น โดยทั่วไปค่าจะอยู่ในช่วง $[0,255]$ หรือ $[0,65535]$ การเทียบเคียงสีเป็นอาร์เรย์ขนาด $p*3$ ของเลขทศนิยม ซึ่งค่าอยู่ในช่วง $[0,1]$ และเป็นระดับของสีเทา
อาร์จีบี (สีแท้จริง)	รูปภาพเป็นอาร์เรย์ขนาด $m*n*3$ ของเลขทศนิยม ซึ่งค่าอยู่ในช่วง $[0,1]$	รูปภาพเป็นอาร์เรย์ขนาด $m*n*3$ ของเลขจำนวนเต็ม ซึ่งค่าอยู่ในช่วง $[0,255]$ หรือ $[0,65535]$

ตารางที่ 2.1 สรุปรูปชนิดของรูปภาพและคลาสต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การออกแบบและโปรแกรมที่ใช้ในการประมวลผลภาพ

3.1 Skin Color Model

ขั้นตอนในการแยกผิวของคนออกจากพื้นที่ที่ไม่ใช่ผิว ซึ่งจะใช้ความแตกต่างของสี เราจำเป็นต้องมีรูปแบบของสีผิวที่เชื่อถือได้ ที่สามารถจะปรับให้เหมาะกับสีผิวและแสงที่ส่องผ่านของแต่ละคนมีความแตกต่างกัน ต่อไปนี้จะอธิบายรูปแบบของสีผิวที่ใช้ในการแบ่งผิว

การแสดงผลภาพสีโดยทั่ว ๆ ไปจะแสดงในโหมด RGB ซึ่งไม่มีความเหมาะสมเพียงพอในการบอกคุณลักษณะของสีผิว ส่วนประกอบทั้ง 3 ส่วน (R,G,B) ไม่ได้แสดงถึงสีเท่านั้น แต่ยังสามารถแสดงถึงระดับความเข้มของแสงด้วย ค่าความเข้มของแสงอาจจะมีการเปลี่ยนแปลงตามแสงที่อยู่รอบ ๆ ตัวที่ส่องผ่านหน้าของคน ๆ หนึ่ง แต่ก็ไม่สามารถนำไปใช้แยกผิวออกจากพื้นที่ที่ไม่ใช่ผิว การแสดงสีแบบโครแมติก (Chromatic) ซึ่งรู้จักกันดีว่าเป็นสีบริสุทธิ์ซึ่งจะไม่มีองค์ประกอบของความเข้มแสงมาเกี่ยวข้อง โดยนิยามได้ดังนี้

$$r = R/(R+G+B)$$

$$b = B/(R+G+B)$$

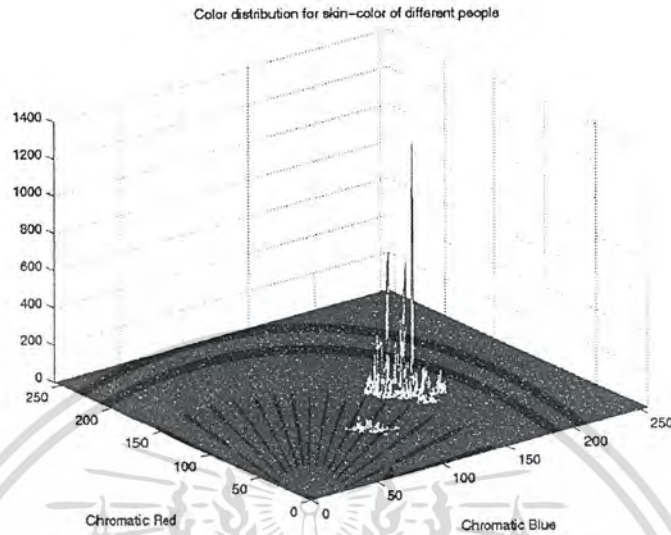
หมายเหตุ: สำหรับ g ไม่จำเป็นต้องคำนวณ เนื่องจากสามารถหาได้จาก $r+g+b = 1$

สีโครแมติกมีประโยชน์มากในการนำไปประยุกต์ใช้ โดยเฉพาะกรณีนี้ที่ได้นำมาประยุกต์ใช้ในการแยกผิวของคนออกจากพื้นที่ที่ไม่ใช่ผิว การกระจายของสีผิวของคนแต่ละคนจะแตกต่างกัน ถึงแม้ว่าจะมีความแตกต่างกันแต่ก็มีผลน้อยกว่าผลที่เกิดจากความแตกต่างความสว่างของแสงดังนั้นเราควรที่จะใช้สีโครแมติกในการแยกสีผิว

จากผิวตัวอย่างเราได้คัดแยกผิวโดยใช้วงจรกรองความถี่ต่ำผ่าน เพื่อลดผลของสัญญาณรบกวนจากสีผิวตัวอย่าง โดยสามารถแสดงผลตอบสนองอินพัลซ์ของวงจรกรองได้ดังนี้

$$1/9 \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



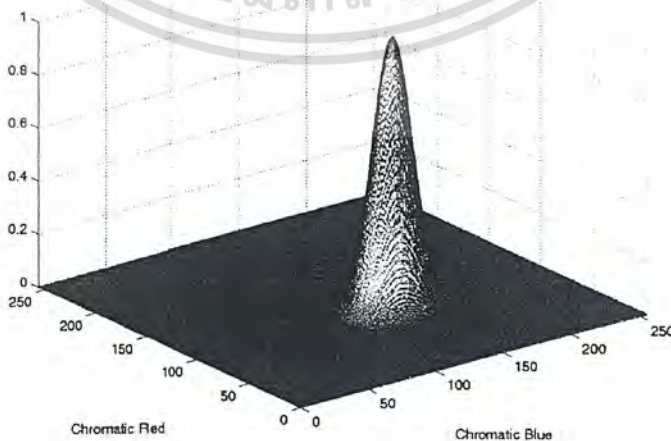
รูปที่ 3.1 กราฟการกระจายสีผิวของบุคคลต่าง ๆ กัน

จากรูปที่ 3.1 แสดงให้เห็นถึงการกระจายของสีผิว โดยวิธี โครแมติกของบุคคลต่าง ๆ ซึ่งจากฮิสโตแกรมจะเห็นได้ว่าการกระจายของสีจับกันเป็นกลุ่มก้อน สามารถแสดงการกระจายนี้ได้โดยใช้แบบจำลองเกาส์เซียน (Gaussian Model $N(m,C)$) ดังนี้

Mean: $m = E(x)$ โดย $x = (r \ b)^T$

Covariance: $C = E\{(x-m)(x-m)^T\}$

Fitting Skin Color into a Gaussian Distribution



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
รูปที่ 3.2 กราฟการกระจายของสีผิว โดยวิธีการกระจายของเกาส์เซียน
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การกระจายโดยวิธีของเกาส์เซียน (Gaussian Distribution) ทำให้เราสามารถทราบช่วงความเป็นไปได้ของแต่ละพิกเซลที่จะถือเป็นสีผิว นั่นคือถ้าแต่ละพิกเซลแปลงจาก สี RGB ไปเป็นสีโครแมติก ซึ่งจะให้ค่าของ (r,b) ความเป็นไปได้ของแต่ละพิกเซล (skin-likelihood) สามารถคำนวณได้จากสมการดังนี้

$$\text{Likelihood} = P(r,b) = \exp[-0.5(x-m)^T C^{-1}(x-m)]$$

$$\text{โดยที่ : } x = (r,b)^T$$

ดังนั้นแบบจำลองของสีผิวนี้สามารถที่จะแปลงจากภาพสีไปเป็นภาพเกรย์สเกล (grayscale) โดยที่แต่ละค่าของพิกเซลของเกรย์สเกล แสดงให้เห็นถึงค่าความเป็นไปได้ของแต่ละพิกเซลที่จะเป็นสีผิวจากภาพสี หากบริเวณที่มีสีเหมือนสีผิวมาก จะมีความเป็นไปได้มาก ภาพ skin-likelihood จะแสดงเป็นสีขาว และลดหลั่นลงมาเป็นระดับของสีเทา (ดังรูปที่ 3.3) ซึ่งหากใช้ค่าเทรชโวลต์ที่เหมาะสมจะสามารถที่จะแปลงภาพจากเกรย์สเกลไปเป็นภาพไบนารีสเกล (binary scale) ซึ่งสามารถที่จะแสดงบริเวณที่เป็นผิวกับบริเวณที่ไม่ใช่ผิวได้

3.2 Skin Segmentation

ในการที่จะตัดส่วนที่เป็นสีผิวออกมา จะต้องเริ่มจากการแปลงภาพต้นแบบที่เป็นภาพสีให้เป็นภาพเกรย์สเกลก่อน ซึ่งเรียกว่าการทำภาพ skin-likelihood โดยใช้ค่าที่ได้จากสมการในหัวข้อ 3.1 ดังตัวอย่างรูปที่ 3.3



รูปที่ 3.3 (รูปซ้าย) รูปต้นแบบ (รูปขวา) แสดงภาพ skin-likelihood

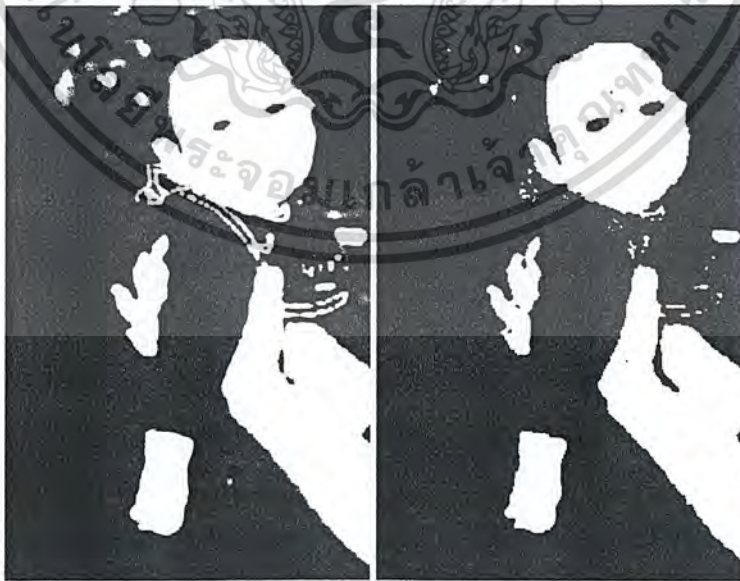
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.3 จะเห็นได้ว่าบริเวณที่เป็นผิว (เช่น หน้า มือ และแขน) จะมีความสว่างมากกว่าบริเวณที่ไม่ใช่ผิว และบริเวณที่ไม่ใช่ผิวจะไม่พิจารณาอีกต่อไป โดยจะแสดงเป็นสีดำ

เมื่อบริเวณที่เป็นผิวมีความสว่างมากกว่าบริเวณอื่น ๆ ทำให้สามารถแยกได้โดยการใช้ค่าเทรชโฮล เมื่อแต่ละคนมีสีผิวที่ต่างกันก็จะให้ค่าการกระจายของสีผิวต่างกัน ในขบวนการปรับค่าเทรชโฮลนี้จะต้องปรับค่าให้เหมาะสมตามสีผิวของแต่ละบุคคล ไม่สามารถใช้ค่าเทรชโฮลที่คงที่เพียงค่าเดียวกับทุกคนได้ วิธีที่ใช้หาค่าเทรชโฮลที่เปลี่ยนแปลงไปตามสีผิวของแต่ละคนได้จะเรียกว่า Adaptive Thresholding และค่าเทรชโฮลที่เหมาะสมกับแต่ละคนจะเรียกว่า optimal threshold

ในการปรับค่าเทรชโฮลนั้นหากปรับค่าให้ลดลง ก็จะเพิ่มบริเวณที่เลือก โดยบริเวณที่เลือกจะเพิ่มขึ้นน้อยลง (โดยกำหนดให้บริเวณผิวที่ต้องการเป็น 100เปอร์เซ็นต์) แต่หากค่าเทรชโฮลต่ำเกินไป บริเวณที่เลือกให้เป็นสีผิวก็จะมีมากเกินไปจริง คือจะรวมบริเวณที่ไม่ใช่สีผิวเข้าไปด้วย ดังนั้นบริเวณที่ถูกเลือกจะเพิ่มขึ้นอย่างรวดเร็ว ในโปรแกรมที่เราออกแบบจะใช้ค่าเทรชโฮลจาก 0.65 จนถึง 0.05 โดยจะลดลงทีละ 0.1 การเปลี่ยนแปลงของค่าเทรชโฮลที่ทำให้เกิดการเพิ่มขึ้นของบริเวณที่เลือกน้อยที่สุดจะถือว่าเป็นค่าเทรชโฮลที่เหมาะสมของคนนั้น หรือเรียกว่าค่า optimal threshold นั้นเอง

ในการใช้เทคนิค Adaptive Thresholding นี้จะเป็นการแยกบริเวณที่เป็นผิวและไม่ใช่ผิวได้เป็นอย่างดี



รูปที่ 3.4 (รูปซ้าย) ภาพ skin-likeness (รูปขวา) ภาพที่ผ่านการแยกสีผิว (skin-segmented image)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยวิธีนี้ บริเวณที่มีสีใกล้เคียงสีผิวแต่ไม่ใช่ผิวเนื้อคนก็จะถูกรวมเข้ามาด้วย เช่น ฉากหลัง หรือ เสื้อผ้า ดังนั้นจะต้องมีการแยกแต่ละบริเวณที่ถูกเลือกออกมาจากภาพ แล้วมาทำการประมวลผลต่อไป

3.3 Skin Regions

การที่จะหาบริเวณผิวหน้าจากรูป skin-segmented image ตามที่เราต้องการ ซึ่งรูปนั้นจะเป็นภาพไบนารี แต่ละส่วน (segment) ที่เป็นสีขาว จะถูกแยกพิจารณาทีละส่วน โดยการสร้างรูปใหม่ที่มีแต่ส่วนนั้น คือจะมีเพียงส่วนเดียวที่เป็นสีขาว ส่วนอื่นจะถูกตัดออกไปก่อนโดยให้เป็นสีดำ เพื่อใช้ในการประมวลผล



รูปที่ 3.5 (รูปซ้าย) บริเวณผิวที่ต้องการแยก (รูปขวา) หลังจากใช้การแยกแต่ละส่วนออกมา

จากรูปที่ 3.5 แสดงให้เห็นถึงบริเวณผิวที่แยก โดยวิธี Skin Regions หลังจากการทดลองหลาย ๆ รูปแล้วทำให้ทราบได้ว่าบริเวณผิวหน้าจะต้องมีรู (hole) อย่างน้อย 1 รูในส่วนของผิวหน้านั้น ในการที่จะหาจำนวนรูภายในบริเวณนั้นสามารถคำนวณได้จาก Euler number

$$E = C - H$$

โดย E คือ Euler number

เอกสารนี้เป็น C คือ จำนวนของ connected components ปรึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณี H คือ จำนวนของรูที่ว่าง (holes) ในบริเวณที่ต้องการจนถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เราสามารถใช้โปรแกรม MATLAB ในการหาคำนวณหา Euler number ในกรณีนี้เรา กำหนดให้จำนวนของ connected components (ได้แก่ผิวหน้า) เป็น 1 ทำให้สามารถคำนวณค่า จำนวนของรูได้ดังนี้

$$H = 1 - E$$

ถ้าในบริเวณที่กำลังพิจารณามีมากกว่า 1 รู เราสามารถวิเคราะห์หาคคุณสมบัติในบริเวณ เฉพาะนั้น โดยเราจะต้องสร้างรูปใหม่โดยให้มีเฉพาะบริเวณนั้นเท่านั้น นอกนั้นกำหนดให้เป็นสีดำ แล้วทำการหาขนาดของบริเวณนั้น ในโครงการนี้จะกำหนดให้ผิวหน้าเป็นบริเวณสีผิวที่มีขนาด ใหญ่ที่สุดในภาพ ดังนั้นส่วนที่มีขนาดใหญ่ที่สุดจะถูกกำหนดให้เป็นส่วนที่มีผิวหน้าเลย

เมื่อได้ส่วนที่ประกอบด้วยผิวหน้าแล้ว เราจะต้องคำนวณหาความกว้างและความสูงของ ส่วนนั้นก่อน เพื่อคำนวณหาอัตราส่วนความสูงต่อความกว้าง (height to width ratio) ซึ่งบริเวณที่ เป็นหน้าคนนั้น จากการทดลอง จะมีอัตราส่วนมากกว่า 0.8 แต่ไม่น้อยกว่า 1.6 (ในโครงการนี้จะใช้ค่า สูงสุดของอัตราส่วนเป็น 1.3) หากมีค่าอัตราส่วนมากกว่า 1.3 เราจะกำหนดอัตราส่วนใหม่ให้เป็น 1.3 และหากอัตราส่วนน้อยกว่า 0.8 เราจะกำหนดอัตราส่วนใหม่ให้เป็น 1 เพื่อลดความผิดพลาดที่จะ เกิดจากบริเวณที่แยกออกมามีส่วนของฉากหลังหรือส่วนของคอและเสื้อผ้าติดมาด้วย โดยเฉพาะ ฉากหลังที่มีความกว้างและสูงมากเกินไป จะทำให้ผลการหาใบหน้าผิดพลาดได้มาก ดังนั้นเราจะ ต้องพิจารณาค่าจุดศูนย์กลางของบริเวณที่แยกออกมาใหม่ร่วมด้วย แล้วกำหนดจุด coordinate ที่ จะใช้เพื่อติดกรอบสี่เหลี่ยมเพื่อที่แสดงใบหน้าใหม่ โดยใช้ระยะทางที่น้อยที่สุดจากจุดศูนย์กลางมาถึง ขอบทั้งซ้าย ขวา บน และล่าง ซึ่งจะทำให้กรอบมีขนาดเล็กลงให้พอดีกับใบหน้า

3.4 ทฤษฎีโครงข่ายประสาทเทียม

นิเวศน์เน็ตเวิร์คเป็นความพยายามอีกอย่างของมนุษย์ในการที่จะลอกเลียนแบบธรรมชาติ โดยทั่วไปแล้วนิเวศน์เน็ตเวิร์คจะประกอบขึ้นจากหน่วยย่อยที่มีหน้าที่งานคล้ายคลึงกับการทำงาน ของมนุษย์ จากนั้นหน่วยย่อยเหล่านี้จะถูกจัดเชื่อมต่อกันเป็นโครงข่ายในลักษณะที่สัมพันธ์กับกาย วิภาคของสมองหรือบางครั้งอาจจะไม่สัมพันธ์ก็ได้ และ โดยการเชื่อมต่อกันในโครงข่ายนี้ทำให้ นิเวศน์เน็ตเวิร์คสามารถแสดงพฤติกรรมมากมายที่มีการตอบสนองคล้ายคลึงกับระบบประสาทใน สิ่งมีชีวิตยกตัวอย่างเช่น ความสามารถในการที่จะเรียนรู้จากประสบการณ์ การจัดกลุ่มของข่าวสาร จากตัวอย่างที่ให้ หรือจับลักษณะเด่นของอินพุทที่ประกอบจากข้อมูลที่ ไม่สมบูรณ์ เป็นต้น

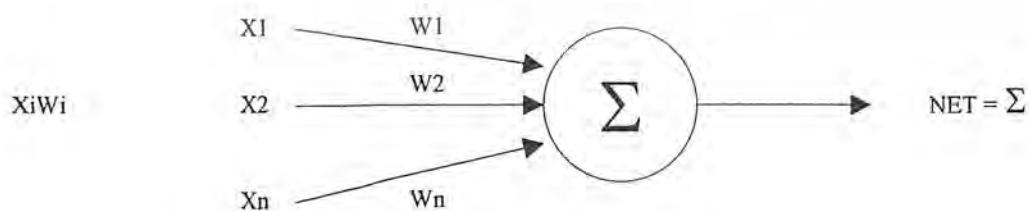
การศึกษาและสิ่งประดิษฐ์ทางด้านนิเวศน์เน็ตเวิร์ค ได้ถูกนำเสนอออกมาทั้งในรูปแบบของ ฮาร์ดแวร์และซอฟต์แวร์ โดยผลิตภัณฑ์ที่เป็นฮาร์ดแวร์มักเรียกว่า Neurocomputer ซึ่งจะทำการจำลอง การทำงานของเซลล์ประสาทบนวงจรทางไฟฟ้าหรือทางออปติคจากนั้นอาจถูกนำมาใช้ร่วมกับ ไมโครคอมพิวเตอร์ ฟังก์ชัน อีกฟังก์ชันหนึ่งคือแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มาใช้

คอมพิวเตอร์ธรรมดาได้ในลักษณะของ Co-processor ซึ่งสามารถเรียกใช้เหมือนกับเป็น โปรแกรมย่อยอันหนึ่งในระบบได้หรือในอีกแนวทางหนึ่งจะอยู่ในรูปของซอฟต์แวร์โดยสร้าง โปรแกรมเพื่อจำลองเซลล์ประสาทและจุดเชื่อมต่อของเซลล์ประสาทเหล่านั้นลงบนคอมพิวเตอร์ เพื่อศึกษาพฤติกรรมของมัน

โครงการยูนีเวอร์ซิตีเนตเวิร์คส์กำลังพิจารณาความสามารถพิเศษของยูนีเวอร์ซิตีเนตเวิร์คส์ ดังเช่นความสามารถในการเรียนรู้จากประสบการณ์หรืออีกหลายๆ อย่าง จะเห็นว่าความสามารถเหล่านั้นสามารถได้จากวิชาการทางด้านปัญญาประดิษฐ์ AI (Artificial Intelligence) ซึ่งจะสามารถเรียนรู้และจดจำ ซึ่งมีกระบวนการคล้ายคลึงกับกระบวนการทางสมองมนุษย์ โดยเฉพาะทางด้านระบบผู้เชี่ยวชาญ (Expert System) ซึ่งประสบความสำเร็จอย่างมากในหลายที่ แต่ก็ประสบปัญหาความยุ่งยากในการสร้างความสามารถสำหรับการเรียนรู้จากประสบการณ์ไม่น้อย จนหลายคนมองว่ายูนีเวอร์ซิตีเนตเวิร์คส์อาจเข้ามาแทนที่ AI แต่ที่จริงแล้วควรนำเทคนิคทั้งสองแบบมาใช้ร่วมกันได้เป็นระบบโดยที่แต่ละตัวทำงานในส่วนที่ตนเองถนัดมากกว่า

ความแตกต่างระหว่างคอมพิวเตอร์ธรรมดากับยูนีเวอร์ซิตีเนตเวิร์คส์ อีกอย่างหนึ่งคือในคอมพิวเตอร์ธรรมดาเมื่อต้องการให้กระทำงานอย่างใดสามารถทำได้โดยการสร้าง โปรแกรมแต่ในทางตรงกันข้ามยูนีเวอร์ซิตีเนตเวิร์คส์เมื่อต้องการให้การทำงานอย่างใดสามารถทำได้โดยการสอน (Training) โดยการฝึกสอนนี้จะทำให้นิวโรนเนตเวิร์คส์นำบทเรียนที่ได้จากการฝึกเข้าไปเก็บในโครงสร้างของยูนีเวอร์ซิตีเนตเวิร์คส์ และจากนั้นจะถูกเรียกกลับขึ้นมาใช้เมื่ออยู่ในระหว่างประมวลผลข้อมูล

การออกแบบสร้างประสาทเทียมมีสมมติฐานจาก โครงข่ายประสาทชีวภาพ ซึ่งมีความนำสัญญาณไฟฟ้าเคมีต่างกัน โมดูลโครงข่ายประสาทเทียมจึงต้องมีการถ่วงน้ำหนักก่อนนำไปใช้งาน เรียกว่า ไชแนปติกเวกท์ ปริมาณข้อมูลจะนำมารวมกัน และตัดสินใจด้วยระดับความสนใจของยูนีเวอร์ซิตีเนตเวิร์คส์ (Activation Level) แล้วส่งเอาที่พุดออกไปยังยูนีเวอร์ซิตีเนตเวิร์คส์อื่นๆ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ภายในขององค์กรเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
รูปที่ 3.6 ส่วนของนิวโรน 1 หน่วย
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.6 เป็นไดอะแกรมที่จำลองจากแนวความคิดของเซลล์สมองทางชีวภาพ สัญญาณอินพุต คือ X_1, X_2, \dots, X_n จะถูกป้อนเข้ามายังนิวรอล เปรียบได้กับสัญญาณไฟฟ้าเคมีเข้ามายังไซแนปส์ของเซลล์ประสาท ค่าอินพุตเหล่านี้จะถูกคูณด้วยค่าถ่วงน้ำหนัก (weight) ที่มีค่าตั้งแต่ 0.0-1.0 ผลรวมของสัญญาณทั้งหมดจะส่งออกมาโดย

$$NET = X_1W_1 + X_2W_2 + \dots + X_nW_n$$

หรืออาจเขียนเป็นรูปแบบของเวกเตอร์ได้ดังนี้

$$NET = \sum (X_i W_i)$$

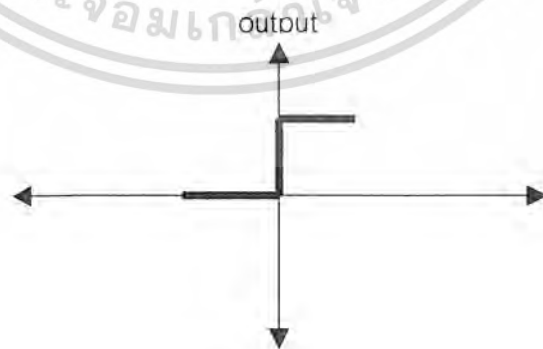
โดยที่

X = เวกเตอร์อินพุต

W = เวกเตอร์น้ำหนัก

NET = เวกเตอร์เอาต์พุต

จากนั้นจะถูกตัดสินใจโดยหน่วยเซลล์ด้วยฟังก์ชันการตัดสินใจ (Activation Function) และได้ค่าเอาต์พุตออกมา



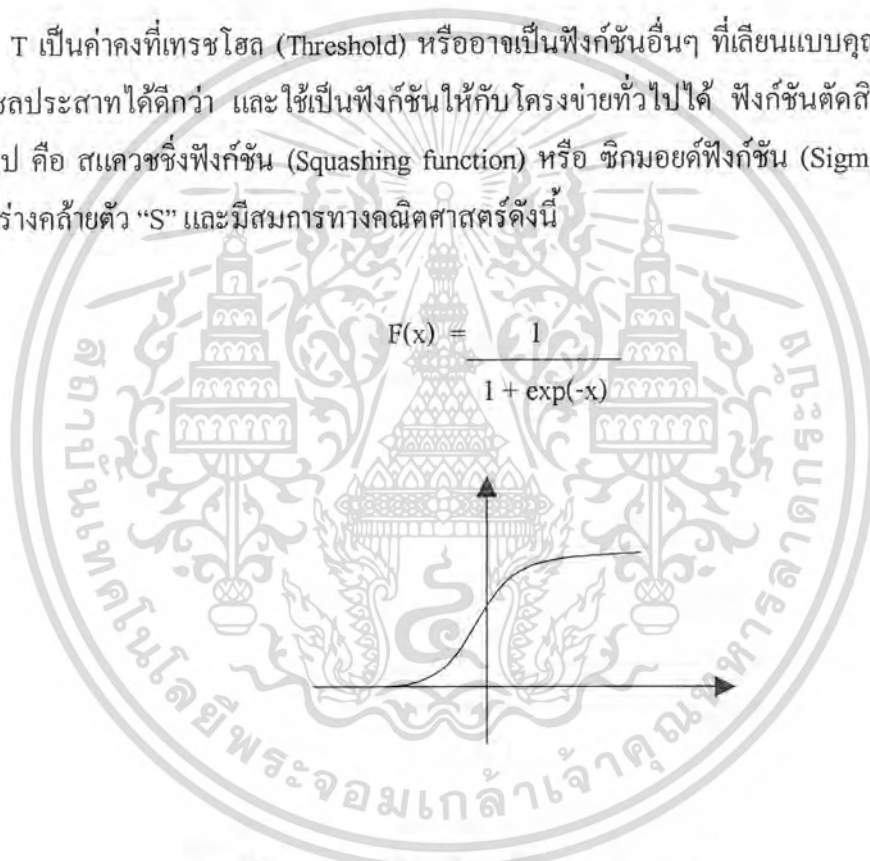
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้รูปที่ 3.7 ฟังก์ชันเส้นตรง
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$OUT = F(NET)$$

ฟังก์ชันตัดสินใจอาจเป็น hard linear function โดย

$$\begin{aligned} OUT &= 1 \text{ ถ้า } NET > T \\ &= 0 \text{ ถ้า } NET \text{ มีค่าเป็นกรณีอื่น} \end{aligned}$$

T เป็นค่าคงที่เทรชโฮลด์ (Threshold) หรืออาจเป็นฟังก์ชันอื่นๆ ที่เลียนแบบคุณสมบัติไม่คงที่ของเซลล์ประสาทได้ดีกว่า และใช้เป็นฟังก์ชันให้กับโครงข่ายทั่วไปได้ ฟังก์ชันตัดสินใจที่นิยมใช้กันทั่วไป คือ สแควชซิงฟังก์ชัน (Squashing function) หรือ ซิกมอยด์ฟังก์ชัน (Sigmoid function) ซึ่งมีรูปร่างคล้ายตัว “S” และมีสมการทางคณิตศาสตร์ดังนี้

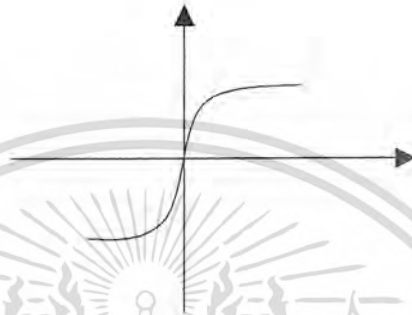


รูปที่ 3.8 ซิกมอยด์ฟังก์ชัน (Sigmoid function)

การใช้ซิกมอยด์ฟังก์ชัน จะทำให้เทรชโฮลด์ฟังก์ชันมีลักษณะ ฟังก์ชันไม่เป็นเชิงเส้นทำให้ได้ค่าอินพุตที่ไวต่อสัญญาณขนาดเล็กๆ และเฉลี่ยต่อสัญญาณแรงๆ คือสัญญาณไปทางบวกเพียงเล็กน้อยก็จะมีค่าเอาต์พุตใกล้เคียง “1” และสัญญาณเป็นลบค่าต่างๆ ใกล้เคียง “0” ขณะเดียวกันสัญญาณบวกขนาดใหญ่ๆ ก็ยังคงให้ค่าเอาต์พุตใกล้เคียง “1” และสัญญาณลบค่ามากๆ ก็ยังคงให้เอาต์พุตใกล้เคียงศูนย์ได้ คือมีคุณสมบัติ non linear gain นั่นเอง คลออสตรี (Grossberg, 1973) พบว่าสามารถแก้ปัญหา Noise saturation dilemma ได้ และทำให้นิวรอนสามารถทำงานได้กว้างขวางขึ้นไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นอกจากนี้ยังมีฟังก์ชันอื่นๆ อีกคือ ไฮเพอร์โบลิกฟังก์ชัน (Hyperbolic function) ซึ่งมีลักษณะคล้ายกับ logistic function คือ

$$OUT = \tanh(X)$$



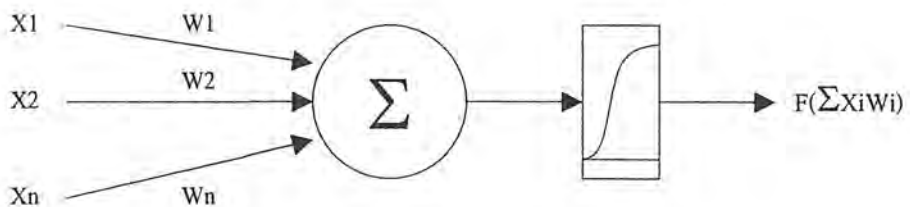
รูปที่ 3.9 hyperbolic tangent function

3.5 โครงข่ายประสาทเทียมแบบชั้นเดียว

(Single Layer Artificial Neural Network)

ถึงแม้ว่าลักษณะของนิวรอนเพียง 1 หน่วย ที่ประกอบด้วยฟังก์ชันบางอย่างได้ แต่ประสิทธิภาพที่เพิ่มขึ้นจากการทำงานมาจากการรวมนิวรอนหลายๆ หน่วยเข้าด้วยกัน

ลักษณะรูปแบบอย่างง่ายที่สุดของโครงข่ายนิวรอน จะจัดกลุ่มของนิวรอนมีลักษณะ 1 ชั้น ดังรูปที่ 3.10

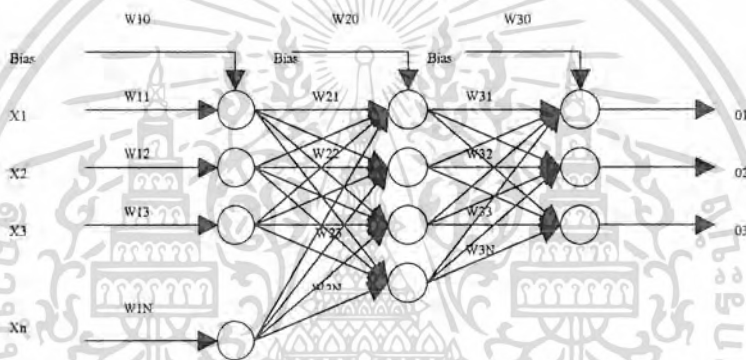


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น รูปที่ 3.10 ลักษณะของนิวรอน 1 หน่วย ที่ประกอบด้วยฟังก์ชันกระตุ้น

3.6 โครงข่ายประสาทเทียมแบบหลายชั้น

(Multilayer Artificial Neural Network)

ซึ่งประกอบด้วยอินพุตเลเยอร์ (Input Network) เอาต์พุตเลเยอร์ (Output Layer) และชั้นซ่อน (Hidden Layer) จำนวนตั้งแต่ 1 ชั้นขึ้นไป ทำให้ความสามารถของโครงข่ายดีขึ้น คือสามารถแก้ปัญหาที่ไม่เป็นเชิงเส้นได้ ปัญหาที่มักเกิดขึ้นจากการออกแบบโครงข่ายประสาทแบบหลายชั้นจะเป็นในแง่ของการกำหนดขนาดจำนวนหน่วยในแต่ละชั้น จำนวนชั้นซ่อนที่เหมาะสม และลักษณะการเชื่อมของแต่ละหน่วย แต่ละชั้น



รูปที่ 3.11 โครงข่ายประสาทเทียมแบบหลายชั้น (Multilayer Artificial Neural Network)

3.7 การฝึกสอนให้กับโครงข่าย

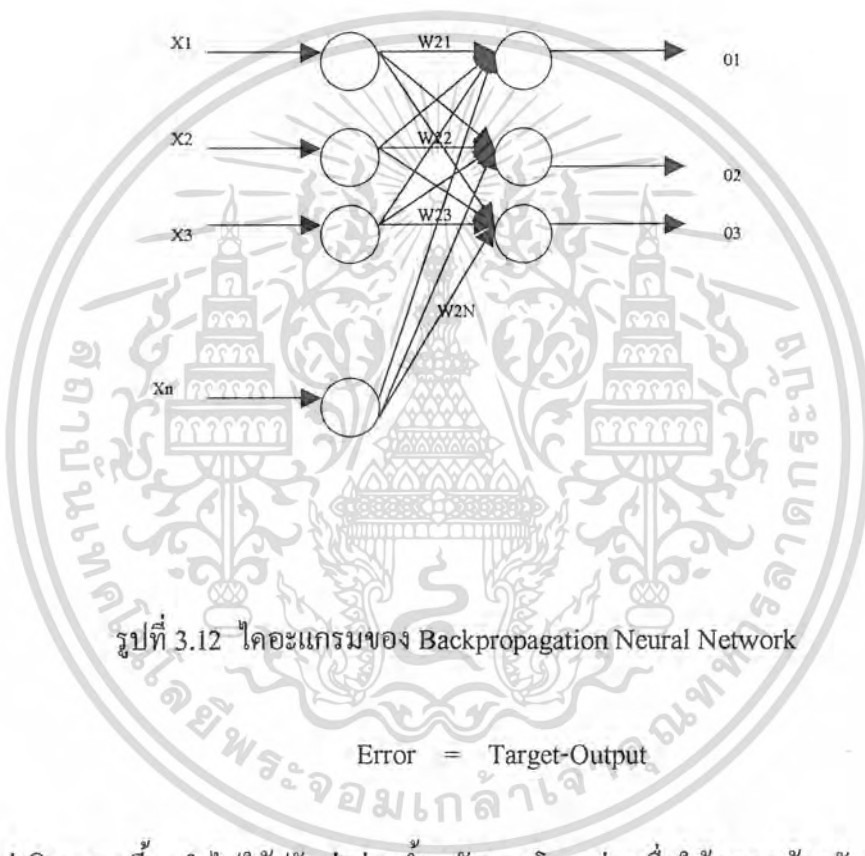
(Training Algorithm)

เทรนนิ่งอัลกอริทึมถูกจัดเป็น 2 ประเภท คือ แบบควบคุม (Supervise training) และแบบอิสระ (Unsupervised) โดยเทรนนิ่งแบบควบคุมต้องการข้อมูล 2 ชุดคือ อินพุต กับ ชุดเป้าหมาย เรียกเทรนนิ่งแพร์ (Training pairs) เช่น การเรียนรู้แบบแพร่ย้อนกลับ ส่วนการเทรนนิ่งแบบอิสระสร้างขึ้น โดยเปรียบเทียบการทำงานของสมองที่ไม่จำเป็นต้องมีผู้มาคอยคิดค่าเป้าหมายให้ก่อน โครงข่ายจะรับเพียงค่าอินพุตเข้าไปเท่านั้น เทรนนิ่งอัลกอริทึม จะเปลี่ยนแปลงค่าถ่วงน้ำหนักของโครงข่าย เพื่อสร้างเอาต์พุตที่มั่นคง และเมื่อค่าอินพุตมีการเปลี่ยนแปลงไปเพียงเล็กน้อย โครงข่ายจะยังสามารถชี้ได้ว่าเป็นข้อมูลเดิม เนื่องจากเอาต์พุตของโครงข่ายไม่มีการกำหนดมาก่อน ส่วนใหญ่จะถูกแปรรูปให้เข้าใจได้ จึงไม่สามารถใช้ตัดสินใจปัญหาที่มีความยากได้ มักใช้ในงานง่ายๆ ง่ายๆ เช่น การเปรียบเทียบเอกลักษณ์ รูปภาพ รูปแบบที่สัมพันธ์กันระหว่างอินพุต-เอาต์พุต ที่มีการนำไปใช้

3.8 ขั้นตอนการเรียนรู้แบบแพร่ย้อนกลับ

(Backpropagation Learning)

การฝึกสอนให้กับโครงข่ายประสาทแบบหลายชั้นใช้การเรียนรู้แบบแพร่ย้อนกลับ โดยโครงข่ายจะรับค่าอินพุต X และค่าเป้าหมาย (Target) ไว้ เมื่อคำนวณค่าเอาต์พุตในโครงข่ายแล้วจะเปรียบเทียบค่าผิดพลาดของเอาต์พุตที่คำนวณได้กับค่าเป้าหมาย



ค่าผิดพลาดนี้จะนำไปใช้ปรับค่าถ่วงน้ำหนักของโครงข่ายเพื่อให้สอดคล้องกับค่าเป้าหมาย การเรียนรู้จะซ้ำๆ ไปเรื่อยๆ จนโครงข่ายสามารถแยกแยะข้อมูลได้ ขั้นตอนการฝึกสอนสามารถสรุปเป็นขั้นๆ ได้ต่อไปนี้

1. สร้างโครงข่ายนิเวรอลที่มีจำนวนอินพุต และเอาต์พุตของปัญหาส่วนจำนวนชั้นซ่อนอาจเริ่มที่ค่ามากพอสมควร
2. ให้ค่าเริ่มต้นแก่ค่าถ่วงน้ำหนักในโครงข่าย โดยควรมีค่าในช่วง -1 ถึง 1
3. คำนวณค่า $\sum X_i W_i = U_i$ สำหรับแต่ละนิเวรอล j ของชั้นซ่อน
4. คำนวณค่ากระตุ้น (Activation) สำหรับแต่ละนิเวรอล j ของชั้นซ่อน $Y_j = U_i$

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการศึกษาเท่านั้น เมื่อผู้ใดเห็นนำไปเผยแพร่โดยไม่ได้รับอนุญาตไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งผู้ใดที่ประสงค์จะนำเอกสารนี้ไปเผยแพร่โดยไม่ได้รับอนุญาตให้มีการนำไปใช้

5. คำนวณค่า $\sum Y_i W_i = V_i$ สำหรับแต่ละนิวรอน k ของชั้นเอาต์พุท
6. คำนวณค่ากระตุ้น ของแต่ละนิวรอน k โดย $OUT = Z_k = f(V_i)$
7. นำค่า Z_k มาเปรียบเทียบกับเป้าหมายถ้าค่าผิดพลาดที่คำนวณได้น้อยกว่าระดับที่กำหนดไว้ ก็จบการเรียนรู้ มิฉะนั้นก็ทำต่อไป
8. คำนวณค่าปรับน้ำหนักสำหรับแต่ละค่าถ่วงน้ำหนัก
 - 8.1 สำหรับเส้นเชื่อมชั้นซ่อนกับชั้นเอาต์พุท

$$\Delta w_j = \alpha \delta_k y_j$$

โดยที่ $\delta_k = (Target\ k - Z_k) f'(U_j)$

- 8.2 สำหรับเส้นเชื่อมชั้นอินพุทกับชั้นซ่อน

$$\Delta w_{ij} = \alpha \delta_j x_j$$

โดยที่ $\delta_j = \sum \delta_k w_{jk} f'(U_j)$

ค่า α : อัตราการเรียนรู้ (Learning rate) มีค่าในช่วง 0 ถึง 1

9. ปรับค่าน้ำหนักเส้นเชื่อมจากนิวรอน r ไปนิวรอน s จะได้

$$W_{rs}(new) = W_{rs}(old) + \Delta W_{rs}$$

10. กลับไปทำข้อ 3

การปรับความเร็วในการเรียนรู้ อาจใช้การปรับค่า Gain ของ sigmoid function ในส่วนคิกริ ของ exponential การปรับอัตราการเรียนรู้ ขนาดจำนวนหน่วยในชั้นซ่อน และการเพิ่มโมเมนตัม แฟกเตอร์เพื่อคำนวณทิศทางการปรับค่าถ่วงน้ำหนัก นอกจากนี้ยังเลือกใช้อัตราการเรียนรู้เป็น ฟังก์ชันที่มีความเหมาะสมได้อีกด้วย

ฟังก์ชัน Sigmoid เป็นที่นิยมในการใช้เทรนนิ่งเนื่องจากมีความใกล้เคียงกับฟังก์ชันซัน บันได และสามารถคำนวณค่าอนุพันธ์ได้ง่ายอีกด้วย โดยอยู่ในรูปนิพจน์ของตัวฟังก์ชันเอง

$$f'(x) = f(x)[1-f(x)]$$

และจากข้อที่ 8 มีการใช้กฎของแคลคูลัสกฎนี้ใช้ปรับน้ำหนักโดยมีลักษณะดังนี้

$$\delta = T - A$$

โดยที่ T เป็นค่าเอาต์พุทที่ต้องการ

A เป็นค่าเอาต์พุทที่ออกมาจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ในเชิงวิชาการเท่านั้น มิใช่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดต่อเปลี่ยนแปลงหรือสังเคราะห์ของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากได้ค่าของเดลต้าแล้วค่านี้จะถูกนำไปคูณค่ากับค่าอินพุท แต่ละอินพุทเพื่อนำไปใช้ในการปรับค่าน้ำหนักให้เหมาะสม นอกจากนั้นยังต้องถูกนำไปคูณกับค่าอัตราการเรียนรู้เพื่อเพิ่มประสิทธิภาพในการฝึกสอนให้เร็วขึ้นหรือช้าลงซึ่งเขียนเป็นสูตรได้ดังนี้

$$\Delta w_{ij} = \alpha \delta_j x_j$$

โดยที่ Δw_{ij} หมายถึงค่าที่ต้องถูกนำมาใช้ในการปรับน้ำหนักของแต่ละค่าน้ำหนัก

x_i คือค่าอินพุทเวกเตอร์

δ_j คือค่าความผิดพลาดของเกร์-เดียน

ดังนั้นสูตรการปรับค่าน้ำหนักสามารถเขียนได้ดังนี้

$$w_{ij(n+1)} = w_{ij(n)} + \Delta w_{ij}$$

โดยที่ Δw_{ij} หมายถึงค่าที่ต้องถูกนำมาใช้ในการปรับน้ำหนักของแต่ละค่าน้ำหนัก

$w_{ij(n)}$ เป็นค่าน้ำหนักหลักการปรับ

$w_{ij(n+1)}$ เป็นค่าน้ำหนักก่อนการปรับ

3.9 สรุปขั้นตอนการประมวลผลรูป

การทำงานทั้งหมดจะใช้โปรแกรม MATLAB version 6.0.0.88 Release 12 ในการประมวลผลรูปเพื่อหาใบหน้า และในการเก็บรูปหน้าจะใช้กล้องดิจิทัล Creative Video Blaster Webcam Go เก็บรูปจากสถานที่ต่าง ๆ กัน โดยที่ฉากหลังจะเป็นแบบไหนก็ได้ (ไม่ต้องกำหนดให้เป็นสีพื้น) แต่ต้องถ่ายในที่ที่มีแสงเพียงพอ และในรูปจะต้องมีใบหน้าคนเพียงคนเดียว และจะบันทึกรูปทั้งหมดเป็น JPEG Format โดยจะใช้โปรแกรม Photoshop ช่วยในการลดขนาดของรูป และใช้ในการซูมตัวอย่างผิวหน้า จากนั้นใช้ทฤษฎีของโครงข่ายประสาทเทียมในการสอนเน็ตเวิร์คในการเรียนรู้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนการประมวลผลรูปจะแสดงได้เป็นบล็อกไดอะแกรมดังรูปที่ 3.13

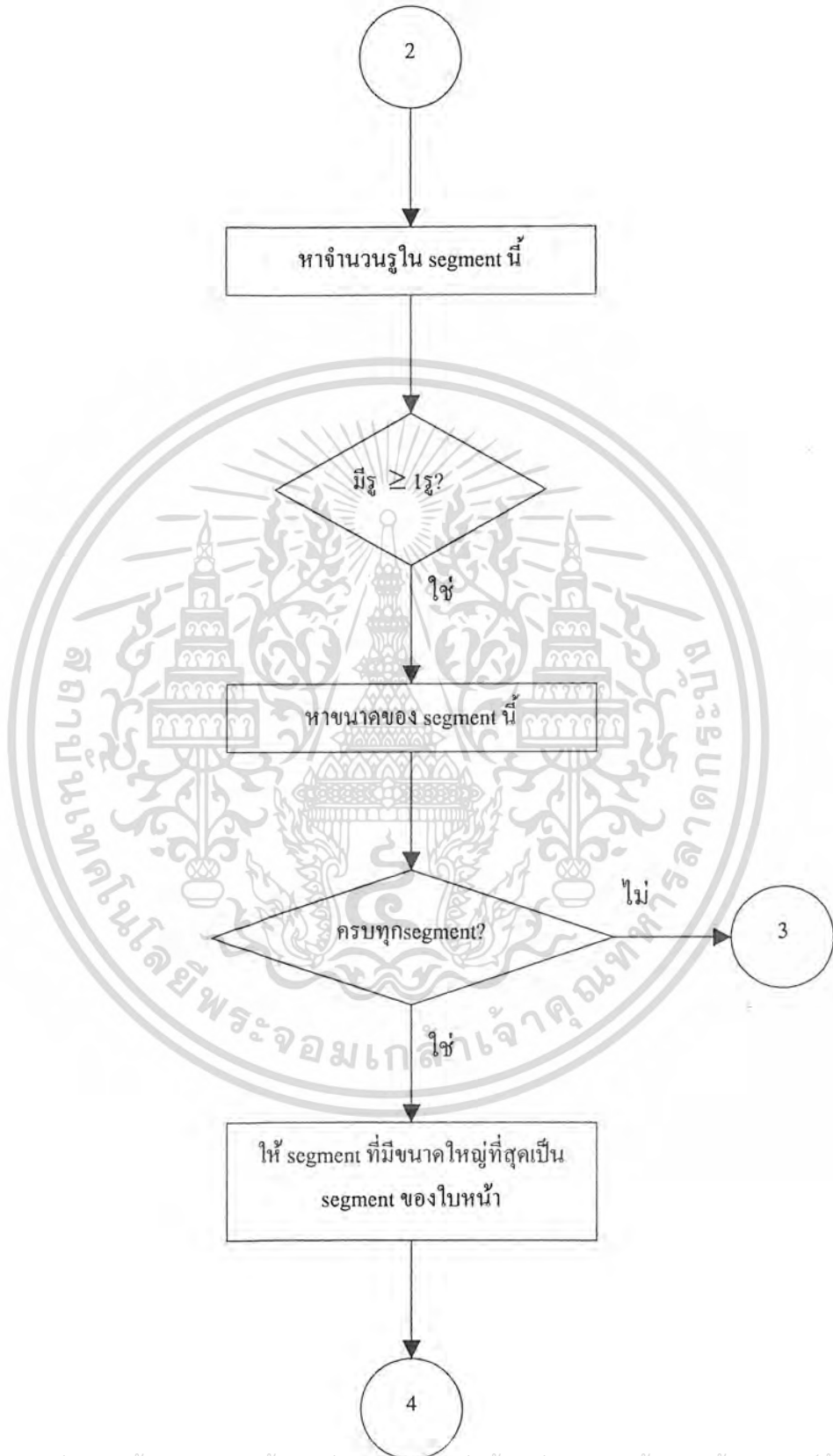
รูปที่ 3.13 บล็อกไดอะแกรมแสดงขั้นตอนการประมวลผลรูป



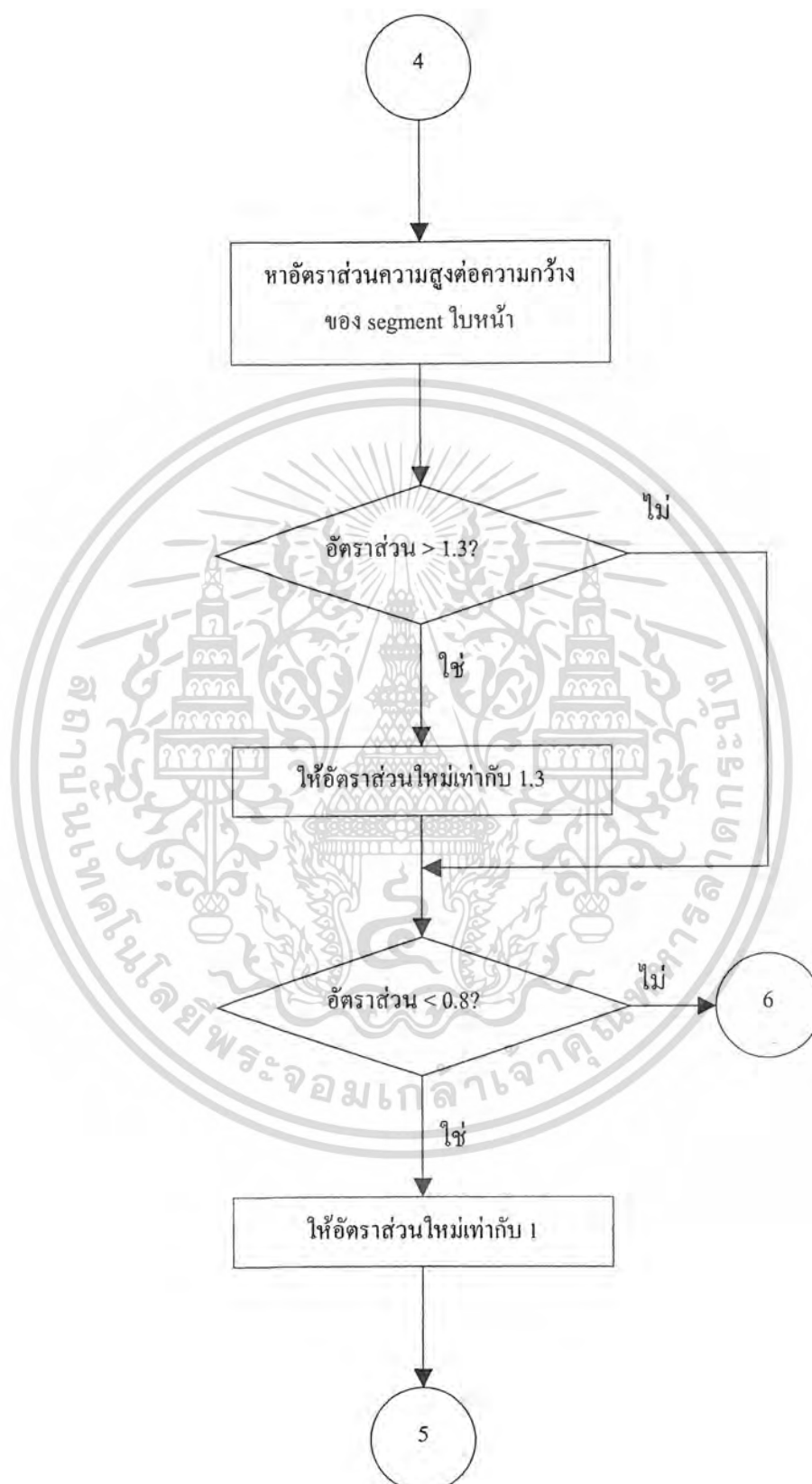
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



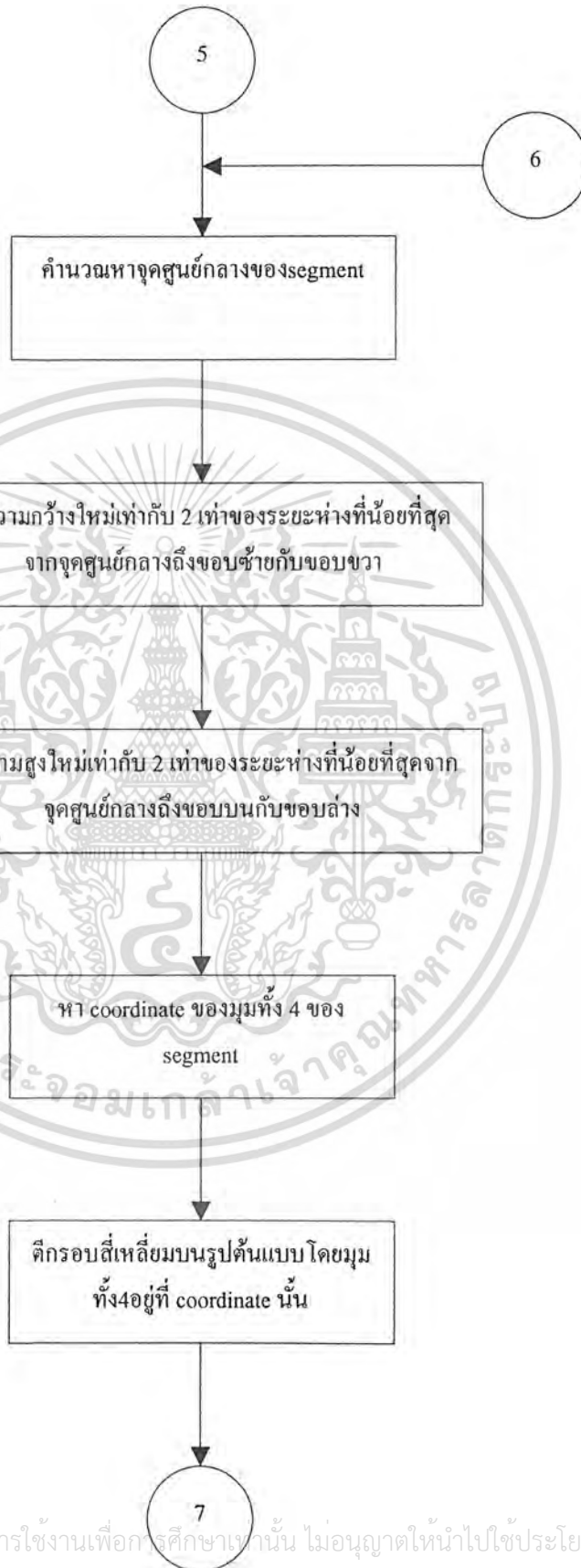
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

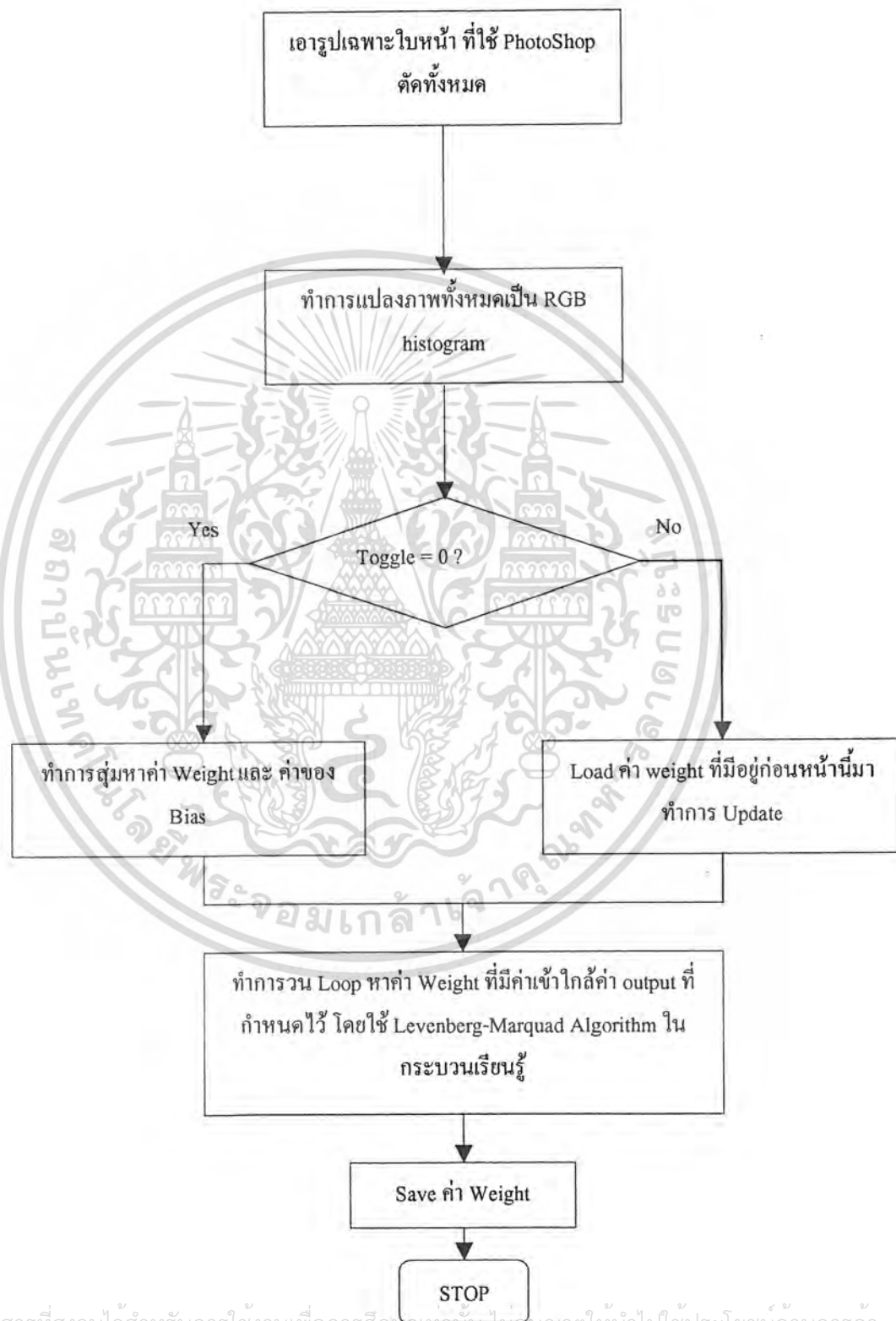


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.14 บล็อกโคอะแกรมแสดงขั้นตอนในการสอนให้เน็ตเวิร์กเรียนรู้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

การทดลองและผลการทดลอง

4.1 การทดลองหาประสิทธิภาพของการแยกสีผิว

วิธีการทดลอง

นำรูปตัวอย่างที่ใช้ทดลองมาทำการแยกสีผิว จากรูปที่เป็น RGB จะกลายเป็นรูปไบนารีที่ประกอบด้วย segment สีขาวที่จะแสดงถึงสีของภาพที่ใกล้เคียงสีผิวเฉลี่ยที่เราหาจากภาพต้นแบบ 10 คน โดยรูปตัวอย่างที่นำมาทดลองจะถูกถ่ายโดยกล้องตัวเดียวกัน แต่อาจจะต่างสถานที่

ผลการทดลอง

ตารางที่ 4.1 ผลการทดลองการหาประสิทธิภาพของการแยกสีผิว

ครั้งที่	จำนวนภาพที่ทดลอง	จำนวนภาพที่แยกได้สมบูรณ์	%ความถูกต้อง	%ความถูกต้องเฉลี่ย
1	15	10	66.67	75
2	12	10	83.33	

จากตารางที่ 4.1 ภาพที่แยกได้สมบูรณ์ บริเวณผิวหนังจะต้องเป็น segment เดียว หรือมีลักษณะติดกันเป็นส่วนเดียว ซึ่งจะมีลักษณะดังรูป



รูปที่ 4.1 (ซ้าย) ภาพต้นแบบ (ขวา) ภาพที่ได้จากการแยกสีผิวแบบสมบูรณ์

ภาพที่แยกได้ไม่สมบูรณ์ อาจเกิดจากภาพที่ถูกลำในบริเวณมีสีของแสงแตกต่างจากต้นแบบมาก หรือภาพที่มีขนาดของแสงไม่เหมาะสม ทำให้ภาพที่ได้จากการแยกสีผิวไม่ติดกันเป็น segment เดียว ดังรูปที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.2 (ซ้าย) ภาพที่ใช้ทดลอง (ขวา) ภาพที่ได้จากการแยกสีผิวแบบไม่สมบูรณ์

4.2 การทดลองหาประสิทธิภาพของการตรวจจับหน้า

วิธีการทดลอง

นำภาพตัวอย่างจากการทดลองในหัวข้อ 4.1 มาทำการตรวจจับหน้า โดยภาพ output จะเป็นภาพสีที่มีสีเหลืองล้อมรอบบริเวณใบหน้าที่ตรวจจับได้

ในการทดลองครั้งที่ 1 เรานำภาพที่ได้จากการแยกสีผิวของการทดลองหัวข้อ 4.1 มาทำการตรวจจับหน้า ซึ่งจะประกอบด้วยภาพที่แยกได้สมบูรณ์ 10 ภาพ และแยกไม่สมบูรณ์ 5 ภาพ

ในการทดลองครั้งที่ 2 เรานำเฉพาะภาพที่แยกสีผิวได้อย่างถูกต้อง (ภาพต้นแบบ) มาทำการตรวจจับหน้า

ผลการทดลอง

ตารางที่ 4.2 ผลการทดลองหาประสิทธิภาพของการตรวจจับหน้า

ครั้งที่	จำนวนภาพที่ทดลอง	จำนวนภาพที่ตรวจจับถูกต้อง	%ความถูกต้อง	%ความถูกต้องเฉลี่ย
1	15	10	66.67	83.33
2	10	10	100	

ภาพที่ตรวจจับถูกต้อง คือ ภาพที่ภายในกรอบสี่เหลี่ยมประกอบด้วยส่วนที่เป็น ตา จมูก ปาก คืออวัยวะทั้ง 3 อย่างจะต้องอยู่ภายในกรอบ ดังรูปที่ 4.3

ภาพที่ตรวจจับผิดพลาด คือ ภาพที่กรอบสี่เหลี่ยมไม่ล้อมรอบอวัยวะทั้ง 3 อย่าง (ตา จมูก ปาก) คือ อาจล้อมรอบเพียงจมูก กับปาก เป็นต้น ดังรูปที่ 4.4 โดยภาพที่ตรวจจับผิดพลาดจะเกิดจากภาพที่แยกไม่สมบูรณ์ทั้งหมด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.3 ภาพที่ตรวจจับได้ถูกต้อง

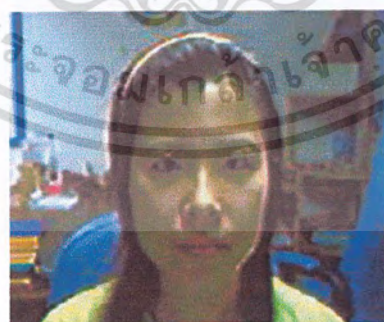


รูปที่ 4.4 ภาพที่ตรวจจับไม่ถูกต้อง

4.3 ผลการนำภาพต้นแบบมาทำการประมวลผลเพื่อตรวจจับหน้า

ผลการทดลองนี้จะเป็นผลที่ได้จากการใช้ภาพต้นแบบทั้งหมด 10 ภาพ จากคน 10 คน มาทำการประมวลผลตั้งแต่ต้นจนจบ โดยมีรูปประกอบการทดลองแต่ละขั้นตอนดังนี้

4.3.1 ผลการตรวจจับหน้าที่ได้ผลดี



รูปที่ 4.5 ภาพต้นแบบ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้รูปที่ 4.6 ภาพลุ่มผิวหนัง ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.7 ภาพ skin-likelihood



รูปที่ 4.8 ภาพที่ได้จากการแยกสีผิว



รูปที่ 4.9 ภาพที่ได้จากการหาจำนวน segment



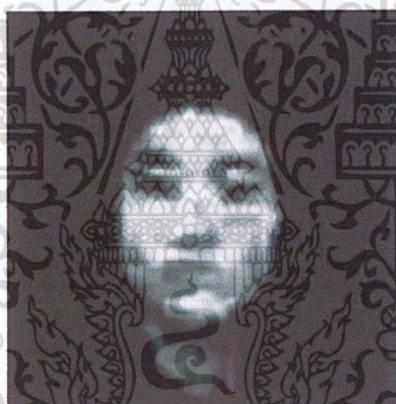
รูปที่ 4.10 ภาพที่ได้จากการตรวจจับหน้าที่ได้ผลดี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.1.1 ผลการตรวจจับหน้าตรงที่ได้ผลดี



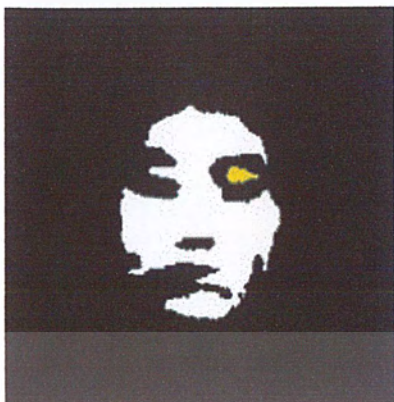
รูปที่ 4.11 ภาพต้นแบบ



รูปที่ 4.12 ภาพ skin-likeness



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับรูปที่ 4.13 ภาพที่ได้จากการแยกสีผิว อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.14 ภาพที่ได้จากการหาจำนวน segment



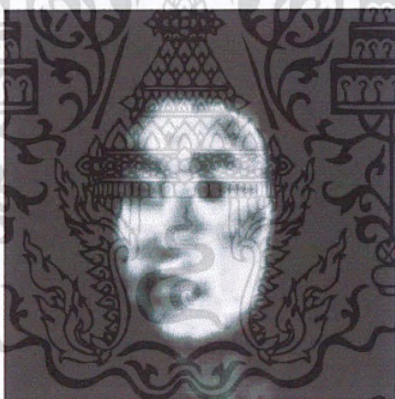
รูปที่ 4.15 ภาพที่ได้จากการตรวจจับใบหน้าที่ได้ผลดี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.1.2 ผลการตรวจจับหน้าที่เอียงที่ได้ผลดี



รูปที่ 4.16 ภาพต้นแบบ



รูปที่ 4.17 ภาพ skin-likelihood

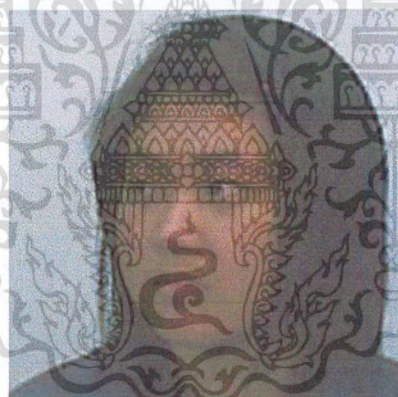


รูปที่ 4.18 ภาพที่ได้จากการแยกสีผิว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับก... ญาติให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้รูปที่ 4.18 นี้ถูกนำออกเผยแพร่ของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.19 ภาพที่ได้จากการหาจำนวน segment



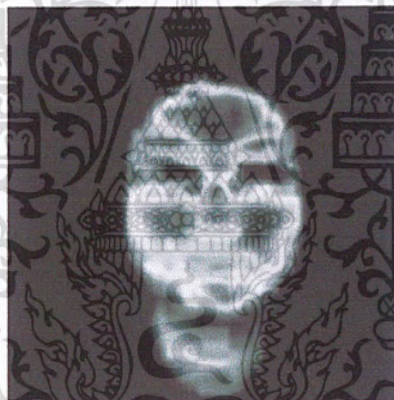
รูปที่ 4.20 ภาพที่ได้จากการตรวจจับหน้าที่ได้ผลดี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.1.3 ผลการตรวจจับหน้าที่มีการสวมแว่นตาที่ได้ผลดี



รูปที่ 4.21 ภาพต้นแบบ



รูปที่ 4.22 ภาพ skin-likelihood

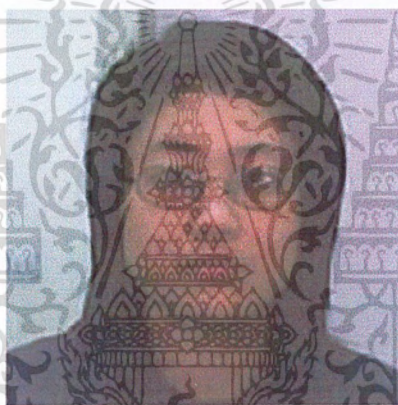


รูปที่ 4.23 ภาพที่ได้จากการแยกสีผิว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานที่ออกชื่อของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.24 ภาพที่ได้จากการหาจำนวน segment



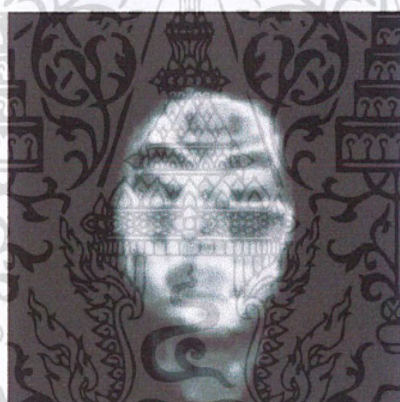
รูปที่ 4.25 ภาพที่ได้จากการตรวจจับหน้าที่ได้ผลดี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

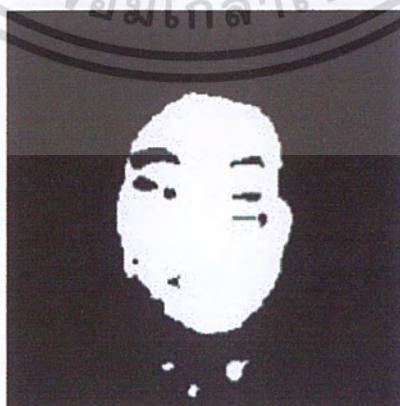
4.3.1.4 ผลการตรวจจับหน้าที่ทำการหลับตาที่ได้ผลดี



รูปที่ 4.26 ภาพต้นแบบ

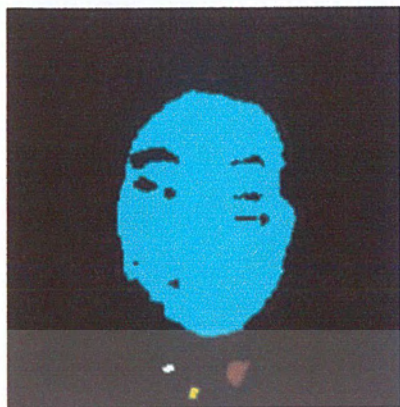


รูปที่ 4.27 ภาพ skin-likelihood

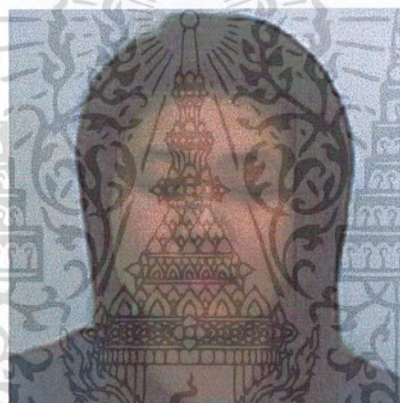


รูปที่ 4.28 ภาพที่ได้จากการแยกสีผิว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.29 ภาพที่ได้จากการหาจำนวน segment



รูปที่ 4.30 ภาพที่ได้จากการตรวจหน้าที่ได้ผลดี

จากรูปที่ 4.10, 4.15, 4.20, 4.25, 4.30 จะเห็นว่ากรอบสี่เหลี่ยมจะประกอบด้วย ตา จมูก ปาก ครบถ้วน และการที่เป็นภาพที่ตรวจจับหน้าที่ได้ผลดี คือ ขนาดของกรอบจะมีขนาดเท่ากับใบหน้าพอดี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.2 ผลการตรวจจับหน้าที่ได้ผลไม่ดี



รูปที่ 4.31 ภาพต้นแบบที่ให้ผลตรวจจับหน้าไม่ดี



รูปที่ 4.32 ภาพที่ได้จากการแยกสีผิวที่ให้ผลตรวจจับหน้าไม่ดี



รูปที่ 4.33 ภาพได้จากการตรวจจับหน้าที่ได้ผลไม่ดี

จากรูปที่ 4.33 จะเห็นว่าภาพที่ตรวจจับหน้าที่ได้ผลไม่ดี จะมีลักษณะของกรอบสี่เหลี่ยมล้อมรอบบริเวณกว้างกว่าใบหน้า ทั้งนี้ เนื่องจากการภาพต้นแบบที่มีฉากหลังมีสีคล้ายสีผิวอยู่ใกล้เคียงกับใบหน้า ทำให้รูปที่ได้จากการแยกสีผิว (รูปที่ 4.32) เกิด segment ใบหน้าขนาดใหญ่กว่าใบหน้าจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.3 ผลการหาประสิทธิภาพของการตรวจจับหน้าจากภาพต้นแบบ

ตารางที่ 4.3 ผลการหาประสิทธิภาพของการตรวจจับหน้าจากภาพต้นแบบ

จำนวนภาพต้นแบบที่ใช้ทดลอง	จำนวนภาพที่ตรวจจับได้ผลดี	จำนวนภาพที่ตรวจจับได้ผลไม่ดี	ประสิทธิภาพของการตรวจจับที่ได้ผลดี
10	7	3	70%

4.4 การหาประสิทธิภาพของภาพที่สอนให้รู้จำใบหน้า

วิธีทำการทดลอง

นำภาพที่ตัดเฉพาะใบหน้าที่มีสมบรูณ์ที่ใช้ในการสอนให้รู้จำใบหน้าทั้ง 8 คน กล่าวคือผลการตรวจจับหน้าที่ได้ผลดีในหัวข้อที่ 4.3 รูปที่ 4.10, 4.15, 4.20, 4.25, 4.30 นำมาทำเป็น RGB profile จากนั้นหาค่า output ของแต่ละรูปใบหน้าของแต่ละคน โดยผ่านสมการของ Neural Network แบบ Back-propagation เพื่อวัดประสิทธิภาพของภาพที่สอนให้รู้จำใบหน้า

ผลการทดลอง

ตารางที่ 4.4 ผลการหาประสิทธิภาพของภาพที่สอนให้รู้จำใบหน้า

ภาพที่ใช้ในการสอนให้รู้จำ	จำนวนภาพที่ใช้ในการสอนให้รู้จำที่ตรวจจับได้ผลดี	จำนวนภาพที่ไม่สามารถจดจำใบหน้าที่สอนได้	ประสิทธิภาพของการจดจำภาพที่สามารถจดจำได้ผลดี
8	8	0	100.00%

4.5 การทดลองหาประสิทธิภาพของการจดจำภาพ

วิธีทำการทดลอง

นำภาพที่ตัดเฉพาะใบหน้าที่มีสมบรูณ์กล่าวคือผลการตรวจจับหน้าที่ได้ผลดีในหัวข้อที่ 4.3 รูปที่ 4.10 นำมาทำเป็น RGB profile จากนั้นหาค่า output ของแต่ละรูปใบหน้าของแต่ละคน โดยผ่านสมการของ Neural Network แบบ Back-propagation และนำค่า output ที่ได้ ไปทำการเปรียบเทียบกับค่าใน Data base เพื่อทำการเปรียบเทียบว่าเป็นใบหน้าใคร

ผลการทดลอง

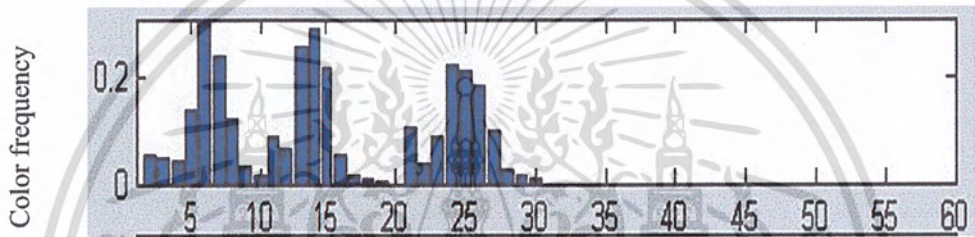
ตารางที่ 4.5 ผลการหาประสิทธิภาพของการจดจำภาพใบหน้า

ภาพที่ใช้ในการทดลอง	จำนวนภาพที่จดจำได้ผลดี	จำนวนภาพที่ไม่สามารถจดจำได้	ประสิทธิภาพของการจดจำที่ได้ผลดี
8	4	4	50.00%

4.5.1 ผลการจดจำภาพใบหน้าที่ได้ผลดี



รูปที่ 4.34 รูปที่ตัดเฉพาะใบหน้า

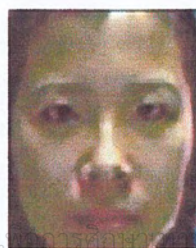


R[1-10], G[11-20], B[21-30]

รูปที่ 4.35 RGB profile



รูปที่ 4.36 ค่า Output



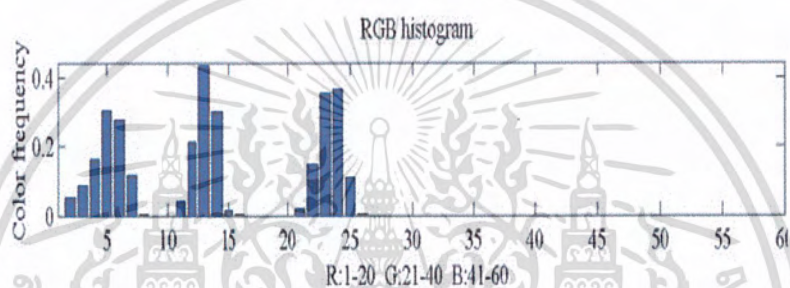
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานภายในเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งยังมีเหตุแต่เปลี่ยนแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.37 ภาพที่สามารถทำการจดจำได้อย่างถูกต้อง

4.5.1.1 ผลการจดจำใบหน้าตรงที่ได้ผลดี

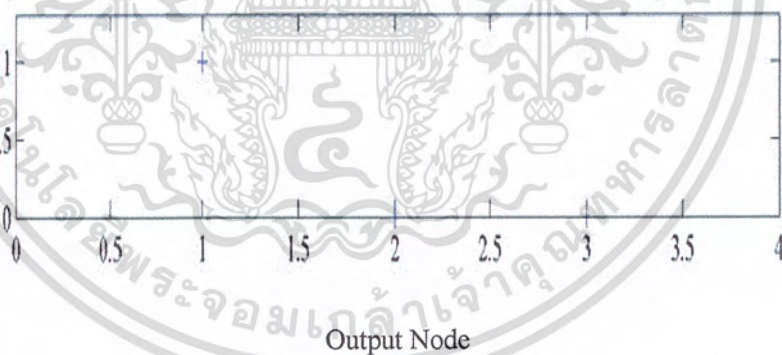


รูปที่ 4.38 รูปที่ตัดเฉพาะใบหน้า



รูปที่ 4.39 RGB profile

output

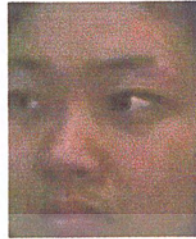


รูปที่ 4.40 ค่า output

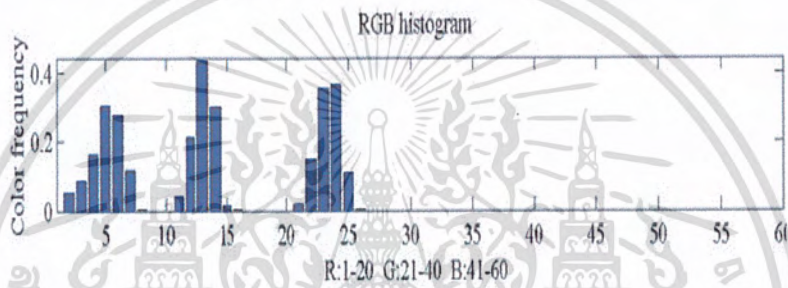


เอกสารนี้เป็นเอกสารที่สงวนไว้รูปที่ 4.41 ภาพที่สามารถทำการจดจำได้อย่างถูกต้องนำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5.1.2 ผลการจดจำใบหน้าเชิงที่ได้ผลดี

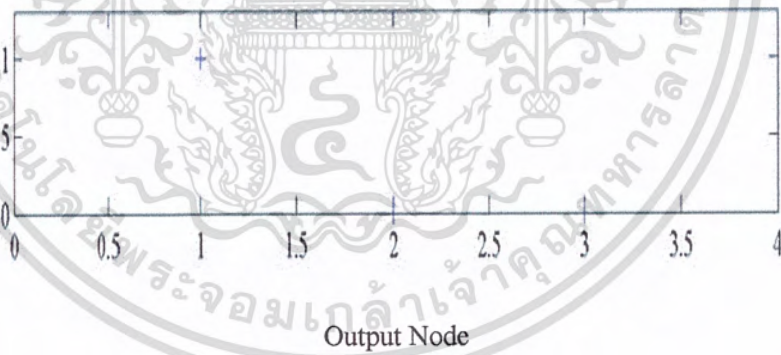


รูปที่ 4.42 รูปที่ตัดเฉพาะใบหน้า



รูปที่ 4.43 RGB profile

output



รูปที่ 4.44 ค่า output

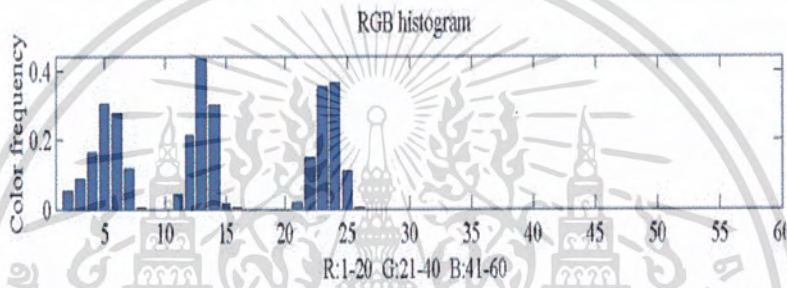


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ภายในองค์กรซึ่งห้ามการเผยแพร่หรือการนำข้อมูลไปใช้ประโยชน์ด้านการค้า
 รูปที่ 4.45 ภาพที่สามารถทำการจดจำได้อย่างถูกต้อง
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

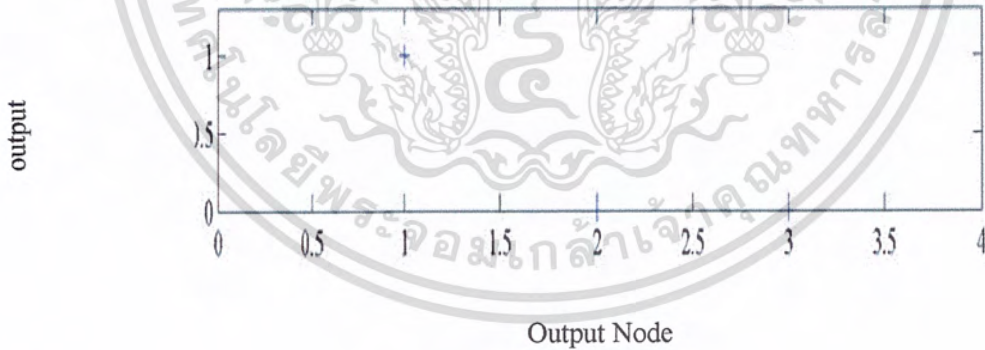
4.5.1.3 ผลการจดจำใบหน้าที่สวมแว่นตาที่ได้ผลดี



รูปที่ 4.46 รูปที่ตัดเฉพาะใบหน้า



รูปที่ 4.47 RGB profile

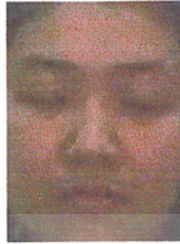


รูปที่ 4.48 ค่า output

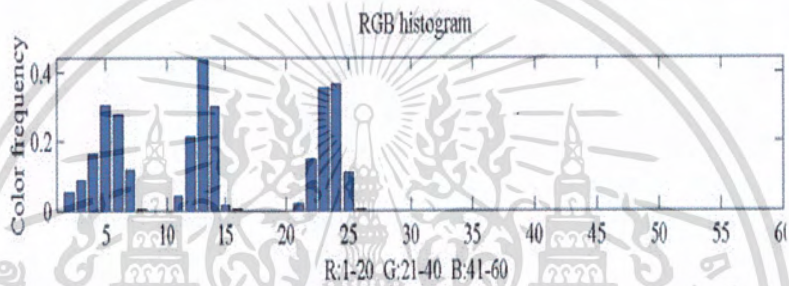


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ภายในองค์กรเท่านั้น การนำเอกสารนี้ไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

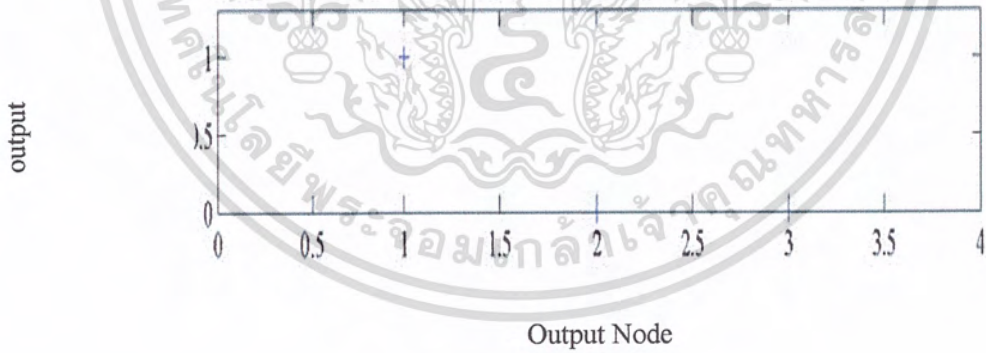
4.5.1.4 ผลการจดจำใบหน้าที่หลับตาที่ได้ผลดี



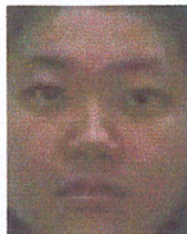
รูปที่ 4.50 รูปที่ตัดเฉพาะใบหน้า



รูปที่ 4.51 RGB profile



รูปที่ 4.52 ค่า output

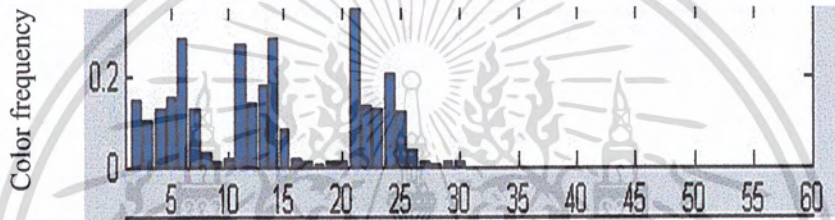


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเฉพาะเท่านั้น ไม่ควรเผยแพร่ไปใช้ประโยชน์ด้านการค้า
 รูปที่ 4.53 ภาพที่สามารถทำการจดจำได้อย่างถูกต้อง
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5.2 ผลการจดจำภาพใบหน้าที่ได้ผลไม่ดี

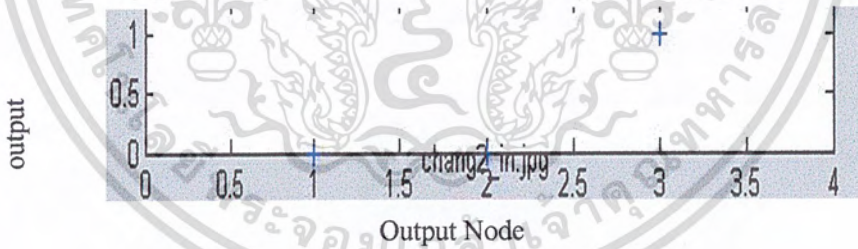


รูปที่ 4.54 ภาพที่ตัดเฉพาะใบหน้า

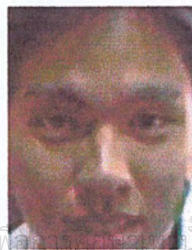


R[1-10], G[11-20], B[21-30]

รูปที่ 4.55 RGB profile



รูปที่ 4.56 ค่า output



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อ *research* เท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามเผยแพร่ต่อแหล่งอื่นและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.57 ภาพที่ไม่สามารถทำการจดจำได้อย่างถูกต้อง

บทที่ 5

สรุปและวิจารณ์ผลการทดลอง

5.1 วิเคราะห์และสรุปผลการทดลอง

1. จากการทดลองการแยกสีผิว ผลที่ได้จะดีหรือไม่ ขึ้นอยู่กับภาพที่เราเก็บไว้เป็นภาพต้นแบบ หากภาพต้นแบบมีแสงที่เพียงพอและมีแสงใกล้เคียงกับภาพสุ่มใบหน้า จะสามารถแยกสีผิวได้สมบูรณ์ นอกจากนี้ยังขึ้นอยู่กับภาพสุ่มสีผิวหน้าที่เก็บไว้เป็นฐานข้อมูลอีกด้วย ภาพเหล่านี้จะมีผลต่อค่าเทรซโฮลที่จะใช้ในการแยกสีผิวเป็นภาพใบหน้า
2. จากการทดลองการตรวจจับใบหน้าผลที่ได้จะขึ้นอยู่กับภาพแยกสีผิวที่นำมาทดลอง ภาพแยกสีผิวจะต้องมีลักษณะเป็น segment เดียวครอบคลุมทั่วใบหน้า (แยกได้สมบูรณ์) จะทำให้การตรวจจับใบหน้าได้ผลถูกต้อง 100%
3. ภาพที่ตรวจจับได้จะได้ผลดี จะต้องควบคุมฉากหลังไม่ให้มีสีผิวในบริเวณที่ใกล้เคียงกับใบหน้า หากฉากหลังมีสีผิวใกล้เคียงบริเวณใบหน้าจะทำให้การตรวจจับให้กรอบสีเหลี่ยมที่แสดงการตรวจจับมีขนาดใหญ่กว่าใบหน้าจริง
4. ภาพของคนๆ เดียวกันแต่ทำการถ่ายรูปที่สภาวะแวดล้อมที่ต่างกัน เช่น มีแสงที่ต่างกัน หรือ มีฉากหลังที่ต่างกัน หรือ ฉากหลังใกล้เคียงกับสีผิวหน้า ก็จะทำให้ภาพที่ตรวจจับได้ต่างกัน ทำให้เวลาที่มีการสอนให้เน็ตเวิร์กว่าใคร จะทำให้ผลในการทดสอบไม่ค่อยมีประสิทธิภาพเท่าที่ควร
5. ในการจดจำภาพใบหน้าแบบ Back-propagation Neural Network จะมีของเสีย คือค่า Weight ที่ทำการวน Loop แต่ละครั้งจะไม่ค่อยมีการ Update ค่า Weight ใหม่ ทำให้ค่าที่ใช้ในการประมวลผลไม่ค่อยมีประสิทธิภาพเท่าที่ควร
6. ถ้าหากจำนวน input , จำนวน hidden layer ที่ป้อนเข้าไปมีจำนวนมากๆ เพื่อเพิ่มประสิทธิภาพของการจดจำภาพใบหน้า หากมากเกินไปก็จะทำให้เสียเวลามากๆ ในการ Train

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ปัญหาและข้อจำกัดของโปรแกรม

โปรแกรมจะใช้ได้ผลดี ขึ้นกับภาพที่นำมาทดลอง จะต้องมิลักษณะที่มีแสงพอเหมาะ และมีแสงใกล้เคียงกันกับในทุก ๆ ภาพที่เก็บเป็นฐานข้อมูล จะทำให้แยกสีผิวได้สมบูรณ์ หากภาพมีลักษณะต่างจากนี้จะแยกสีผิวไม่ได้หรือได้ไม่สมบูรณ์ ส่วนข้อจำกัดอีกอย่างของการตรวจจับใบหน้าให้ได้ผลดี คือ ฉากหลังของภาพจะต้องไม่มีสีใกล้เคียงสีผิวขนาดใหญ่กว่าใบหน้า และจะต้องไม่มีสิ่งที่มีสีใกล้เคียงสีผิวใกล้ ๆ บริเวณใบหน้า

สำหรับส่วนในการสอนให้เน็ตเวิร์คเรียนรู้ใบหน้าแต่ละคนนั้น ข้อจำกัดก็คือ ในการใช้ Back-propagation Neural Network จะมีข้อจำกัดตรงที่ค่าของ Weight ที่ได้จะไม่มีการ Update ทำให้ค่า Weight ที่ออกมาไม่ค่อยจะถูกต้องเท่าที่ควร กล่าวคือ เวลาทำการ Test ประสิทธิภาพที่ได้จะไม่ 100 เปอร์เซ็นต์ ถ้าหากอยากเพิ่มประสิทธิภาพให้ดีขึ้นสามารถทำได้โดยการเพิ่มจำนวนของ input, และจำนวนของ hidden layer ให้มีค่าสูงๆ แต่ก็จะมีข้อเสียตรงที่จะต้องใช้เวลาในการประมวลผลที่จะได้ output ออกมา และปัญหาที่เจออีกก็คือ ถ้าหากภาพที่เราทำการตัดรูปใบหน้าไม่ได้รูปใบหน้าจริงๆ กล่าวคือมีส่วนอื่นถูกตัดมาด้วย เช่น ผม หรือ ฉากหลัง เวลามาภาพไปทำ RGB profile ก็จะได้ค่าที่มี error เกิดขึ้น

5.3 การพัฒนาโปรแกรมต่อไปในอนาคต

โปรแกรมควรจะต้องสามารถตรวจจับใบหน้าได้อย่างถูกต้องและได้ผลดีไม่ว่าฉากหลังจะมีลักษณะเช่นไร และอาจจะสามารถตรวจจับภาพที่มีใบหน้ามากกว่า 1 คนได้ ภาพใบหน้าเหล่านี้ อาจจะสามารถเอียงได้ ผลการตรวจจับจะต้องถูกต้องเช่นเดิม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรม MATLAB ที่ใช้ในการประมวลผลภาพเพื่อตรวจจับใบหน้า



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



โปรแกรม MATLAB ที่ใช้ในการตรวจจับใบหน้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

main.txt

```
%-----Main1-----
```

```
%plot skin-color distribution
```

```
[cb1,cr1] = chromadist('ex_bom1.jpg');
```

```
[cb2,cr2] = chromadist('ex_chang.jpg');
```

```
[cb3,cr3] = chromadist('ex_chang1.jpg');
```

```
[cb4,cr4] = chromadist('ex_chang2.jpg');
```

```
[cb5,cr5] = chromadist('ex_day.jpg');
```

```
[cb6,cr6] = chromadist('ex_day1.jpg');
```

```
[cb7,cr7] = chromadist('ex_desh.jpg');
```

```
[cb8,cr8] = chromadist('ex_desh2.jpg');
```

```
[cb9,cr9] = chromadist('ex_koi.jpg');
```

```
[cb10,cr10] = chromadist('ex_koi1.jpg');
```

```
cb = [cb1 cb2 cb3 cb4 cb5 cb6 cb7 cb8 cb9 cb10];
```

```
cr = [cr1 cr2 cr3 cr4 cr5 cr6 cr7 cr8 cr9 cr10];
```

```
figure(1), colordistplot(cb,cr);
```

```
bmean = mean(cb);
```

```
rmean = mean(cr);
```

```
brcov = cov(cb,cr);
```

```
colorchart = zeros(256);
```

```
for b = 0:255
```

```
for r = 0:255
```

```
    x = [(b-bmean); (r-rmean)];
```

```
    colorchart(b+1,r+1) = exp(-0.5*x'*inv(brcov)*x);
```

```
end
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

main.txt

```
end
```

```
figure(2),surf(colorchart)
```

```
segmentskin('koi1.jpg', bmean, rmean, brcov);
```

```
final('koi1.jpg','koi1_seg.jpg');
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลง ผลิตซ้ำ และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

center.txt

% Computes the centroid of a skin region, using its area. See reference for more% details.

```
function [xmean, ymean] = center(bw)
```

```
bw=bwfill(bw,'holes');
```

```
area = bwarea(bw);
```

```
[m n] =size(bw);
```

```
bw=double(bw);
```

```
xmean =0; ymean = 0;
```

```
for i=1:m,
```

```
    for j=1:n,
```

```
        xmean = xmean + j*bw(i,j);
```

```
        ymean = ymean + i*bw(i,j);
```

```
    end;
```

```
end;
```

```
xmean = xmean/area;
```

```
ymean = ymean/area;
```

```
xmean = round(xmean);
```

```
ymean = round(ymean);
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

chromadist.txt

```
function [cb, cr] = chromadist(filename)
% return the chromatic components of the image
% low pass filtering is carried out to remove noise

im = imread(filename);
imycb = rgb2ycbcr(im);
lpf = 1/9 * ones(3);
cb = imycb(:,:,2);
cb = filter2(lpf, cb);
cb = reshape(cb, 1, prod(size(cb)));
cr = imycb(:,:,3);
cr = filter2(lpf, cr);
cr = reshape(cr, 1, prod(size(cr)));
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

colordistplot.txt

```
function plot = colordistplot(cb,cr)
```

```
chroma = zeros(256);
```

```
cb = round(cb);
```

```
cr = round(cr);
```

```
for i = 1:length(cb)
```

```
    chroma(cb(i),cr(i)) = chroma(cb(i),cr(i)) + 1;
```

```
end
```

```
surf(chroma)
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

final.txt

```
function [finalpic] = final(imoriginal,imsegmented)
```

```
imsource = imread(imoriginal);
```

```
%find holes in each region
```

```
skincolor = imread(imsegmented);
```

```
BW = im2bw(skincolor);
```

```
[L,numobj] = bwlabel(BW,8);
```

```
map = [0 0 0; jet(numobj)];
```

```
figure(1),imshow(L+1,map,'notruesize');
```

```
for i=1:numobj,
```

```
    [x,y] = find(bwlabel(BW) == i);
```

```
    bwsegment = bwselect(BW,y,x,8);
```

```
    [L,numbojs] = bwlabel(bwsegment,4);
```

```
    numfeatures = bweuler(bwsegment,4);
```

```
    numholes = 1 - numfeatures;
```

```
    %if there is at least one hole in the region, compute the area of region.
```

```
    %then compare with other regions.
```

```
    if (numholes >= 1)
```

```
        area(i) = bwarea(bwsegment);
```

```
    end
```

```
end;
```

```
[y i] = max(area);
```

```
%maximum value is considered face region
```

```
[x,y] = find(bwlabel(BW) == i);
```

```
bwface = bwselect(BW,y,x,8);
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลง Page 1 และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

final.txt

```
[m n] = size(bwface);
```

```
%compute the width & height of region
```

```
[l,r,u,d] = resize(bwface);
```

```
wx = (r - l + 1); % width
```

```
ly = (d - u + 1); % height
```

```
% Get the ration between the height (ly) and width(wx) of the region
```

```
hwratio = ly/wx
```

```
% To be used if we find regions that are very tall
```

```
% Adjust the ratio by reducing the height of the region
```

```
if (hwratio > 1.3)
```

```
    ly = floor(1.3 * wx); % approx. computation of the new height
```

```
% This is to eliminate the parts of the image that are cut
```

```
% after computing the new height of the region
```

```
[l,r,u,d] = resize(bwface);
```

```
for i = (u+ly):m,
```

```
    for j = 1:n,
```

```
        bwface(i,j) = 0;
```

```
    end
```

```
end
```

```
% Get the new ratio
```

```
hwratio = ly/wx
```

```
end;
```

```
% To be used if we find regions that are very wide
```

```
% Adjust the ratio by reducing the width of the region
```

```
if (hwratio < 0.8)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลง Page 2 และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

final.txt

```
wx = ly; % approx. computation of the new width

% This is to eliminate the parts of the image that are cut
% after computing the new width of the region
[l,r,u,d] = reysize(bwface);
for j = (l+wx):n,
    for i = 1:m,
        bwface(i,j) = 0;
    end
end
% Get the new ratio
hwratio = ly/wx
end;

%compute center of the region
[cx cy] = center(bwface);

[l,r,u,d] = reysize(bwface);

%assume that face region is centered on cx to avoid error from background
x1 = cx - l;
x2 = r - cx;
x = min(x1,x2);
l = cx - x;
r = cx + x;

%assume that face region is centered on cy to avoid error from background
y1 = cy - u;
y2 = d - cy;
y = min(y1,y2);
u = cy - y;
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลง **Page 3** และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

final.txt

```
d = cy + y;
```

```
%display the original image with the region that is a face by drawing
```

```
%a rectangular
```

```
rectcoor = [l u (r-l) (d-u)];
```

```
figure(2),imshow(imsource);
```

```
hd = rectangle('Position',rectcoor);
```

```
set(hd,'edgecolor','w');
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

resize.txt

```
function [left, right, up, down] = reccsize(A)
```

```
[m n] = size(A);
```

```
left = -1;
```

```
right = -1;
```

```
up = -1;
```

```
down = -1;
```

```
for j=1:n,
```

```
for i=1:m,
```

```
if (A(i,j) ~= 0)
```

```
left = j;
```

```
break;
```

```
end;
```

```
end;
```

```
if (left ~= -1) break; end;
```

```
end;
```

```
for j=n:-1:1,
```

```
for i=1:m,
```

```
if (A(i,j) ~= 0)
```

```
right = j;
```

```
break;
```

```
end;
```

```
end;
```

```
if (right ~= -1) break; end;
```

```
end;
```

```
for i=1:m,
```

```
for j=1:n,
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

resize.txt

```
if (A(i,j) ~= 0)
    up = i;
    break;
end;
end;
if (up ~= -1) break; end;
end;

for i= m:-1:1,
    for j=1:n,
        if (A(i,j) ~= 0)
            down = i;
            break;
        end;
    end;
    if (down ~= -1) break; end;
end;

if (left == -1) left = 1; end;
if (right == -1) right = n; end;
if (up == -1) up=1; end;
if (down == -1) down = m; end;
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

rgb_trainall.txt

```
function [x, outputs] = rgb_trainall(N,p, toggle)

% Trains all images in directory images.
% We give file names of all jpg images.
% Plots final residual of each image at end.
% Given # bins N and # hidden nodes p.
% If toggle = 0, then we start with random initial weights.
% Otherwise assumes inital weights w given in file rgb_weights unless toggle=0.

warning off;

%Training rate lambda.
lam = 0.1;

% # iterations.
itMAX = 1000;

% Human faces: y=1
x(:,1) = image2rgbhist('bom2_in.jpg', N);
y(1,1) = 0;y(2,1)=0;y(3,1)=0;

x(:,2) = image2rgbhist('chang2_in.jpg', N);
y(1,2) = 0;y(2,2)=0;y(3,2)=1;

x(:,3) = image2rgbhist('koi2_in.jpg', N);
y(1,3) = 0;y(2,3)=1;y(3,3)=0;

x(:,4) = image2rgbhist('leek2_in.jpg', N);
y(1,4) = 0;y(2,4)=1;y(3,4)=1;
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

rgb_trainall.txt

```
x(:,5) = image2rgbhist ('ron2_in.jpg', N);  
y(1,5) = 1;y(2,5)=0;y(3,5)=0;
```

```
x(:,6) = image2rgbhist ('mai2_in.jpg', N);  
y(1,6) = 1;y(2,6)=0;y(3,6)=1;
```

```
x(:,7) = image2rgbhist ('num2_in.jpg', N);  
y(1,7) = 1;y(2,7)=1;y(3,7)=0;
```

```
x(:,8) = image2rgbhist ('may2_in.jpg', N);  
y(1,8) = 1;y(2,8)=1;y(3,8)=1;
```

```
% Run through Levenberg-Marquadt algorithm.
```

```
if toggle == 0
```

```
    w=0;
```

```
else
```

```
    load rgb_weights w;
```

```
end;
```

```
[w,res] = levmar (x,y, p, w, lam, itMAX);
```

```
save rgb_weights w;
```

```
% Output overall residual on final weights w.
```

```
% Residual = Network Output - Desired Output.
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

rgb_trainall.txt

```
for i=1:8
    outputs(:,i) = forward(x(:,i),3,p,w);
    %overall_res(i) = outputs(i) - y(1,i);
end;

subplot(3,1,1);
plot(outputs(1,:));
subplot(3,1,2);
plot(outputs(2,:));
subplot(3,1,3);
plot(outputs(3,:));
%title('Network outputs');
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

rgb_forward.txt

```
function [y] = rgb_forward (file_name)
```

```
% rgb_forward.m
```

```
% Uses RGB histogram approach for identifying image.
```

```
% Given an image file (i.e. jpeg), returns value y
```

```
% y = 1 for face, y = 0 for not face
```

```
% Assumes weights are stored as w in rgb_weights.mat
```

```
% Plots results
```

```
x = image2rgbhist (file_name, 10);
```

```
load rgb_weights w;
```

```
y = forward (x, 3, 20 , w);
```

```
subplot (3,1,1);
```

```
imshow(file_name);
```

```
subplot(3,1,2);
```

```
bar(x);
```

```
xlabel('R:1-20 G:21-40 B:41-60');
```

```
axis([1,60,0,max(x)+0.01]);
```

```
ylabel('Color frequency')
```

```
title('RGB histogram')
```

```
subplot(3,1,3);
```

```
plot(y, '+');
```

```
axis([0.5, 1.5, 0, 1.3]);
```

```
xlabel(y);
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

image2rgbhist1.txt

```
function [x] = image2rgbhist1 (file_name, N)

% Writes an image at file_name (i.e. jpg) into 3 histograms.
% Each histogram has N bins.
% The vector is the appended R, G, then B histograms.
% We scale x at the end so all entries are between 0 and 1.
% This is a necessary condition for feeding into the neural network.
% The resize_matrix operation assumes each image is at least 50x50 pixels.

% Get bin centers.
for i=1:N
    Bins(i,1) = i/N - 1/(2*N);
end;

% Interpolation factor I
I = 50;

A = double(file_name);
B = resize_matrix(A,I,I);
B = B/260;
[total_rows total_columns three] = size(B);

for row = 1:total_rows
    for column = 1:total_columns
        v(column+(row-1)*total_columns,1) = B(row,column,1);    % R
        v(column+(row-1)*total_columns,2) = B(row,column,2);    % G
        v(column+(row-1)*total_columns,3) = B(row,column,3);    % B
    end;
end;
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

image2rgbhist1.txt

```
total_hist = hist(v,Bins);
```

```
hist_R = total_hist(:,1);
```

```
hist_B = total_hist(:,2);
```

```
hist_G = total_hist(:,3);
```

```
x=[hist_R;hist_G;hist_B];
```

```
x = x / (I^2);
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

forward.txt

```
function [y] = forward (x, n, p ,w)

% forward.m

% Calculates forward output of given input x with weights w.
% x(1:m) vector corresponding to given input
% n is # of nodes in output (T)
% p number of nodes in hidden layer
% w(1:P) vector corresponding to weights -- P=p*(m+1)+n*(p+1)
% y(1:n) vector corresponding to network output

m=length(x);
xp=[x;1];
tem=w(1:p*(m+1));
W1=reshape(tem,p,m+1);
tem=w(p*(m+1)+1:p*(m+1)+n*(p+1));
W2=reshape(tem,n,p+1);
a=W1*xp;
z=1./(1+exp(-a));
zp=[z;1];
y=W2*zp;
%y=1./(1+exp(-y));
```

resize_matrix.txt

```
function [B] = resize_matrix (A, r, c)

% To use type "[B] = resize_matrix (A, #, #)".
% Resizes image matrix A by picking out pixels.
% Resulting matrix B will have r rows by c columns.
% Each pixel has three values: RGB.

[rows columns three] = size(A);

row_sample = floor (rows/r);
column_sample = floor (columns/c);

for i=1:r
    for j=1:c
        B(i,j,:) = A(i*row_sample, j*column_sample, :);
    end;
end;
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

test1.txt

```
function [y] = Joy1(filename,image_seg)
finalpic=final(filename,image_seg);
figure(1),subplot(5,1,1);
imshow(filename);
y=rgb_forward1(finalpic);
figure(1),subplot(5,1,5);
if y == [0;0;0]
    imshow('bom2_in.jpg');
elseif y == [0;0;1]
    imshow('chang2_in.jpg');
elseif y == [0;1;0]
    imshow('koi2_in.jpg');
elseif y == [0;1;1]
    imshow('lek2_in.jpg');
elseif y == [1;0;0]
    imshow('ron2_in.jpg');
elseif y == [1;0;1]
    imshow('mai2_in.jpg');
elseif y == [1;1;0]
    imshow('num2_in.jpg');
elseif y == [1;1;1]
    imshow('may2_in.jpg');
end;
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

levmar.txt

```
function [w, res] = levmar (X,T, p, w0, lam, itMAX)
```

```
% levmar -- Levenberg-Marquard Algorithm
```

```
% Input
```

```
% (X,T) are nodes to go through
```

```
% p is number of nodes in hidden layer
```

```
% w0 is initialization for weights
```

```
% lam is initialization for learning rate lambda
```

```
% itMAX is number of iterations
```

```
% Output
```

```
% w is the final weights matrix
```

```
% res is the list of error at each iteration
```

```
% Compute number of unknowns: nunk.
```

```
[m, nex] = size(X);
```

```
[n, nex] = size(T);
```

```
nunk = p*(m+1) + n*(p+1);
```

```
if w0 == 0
```

```
    w = 3*rand(nunk,1); % Random weights.
```

```
else
```

```
    w = w0;
```

```
end;
```

```
% Levenberg-Marquard Algorithm
```

```
for it = 1:itMAX
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อ Page 1 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

levmar.txt

```
[Etot, Gtot, Jtot, R] = etot(X,T,p,w,1);  
An = Jtot' * Jtot + lam * eye(nunk, nunk);  
  
dw = An \ (Jtot' * reshape(R, n*nex, 1) );  
wtst = w - dw;  
[Etot, Gtot, Junk, R] = etot(X,T,p,wtst,0);  
  
if Etst <= Etot  
    w = wtst; lam = lam / 10;  
else  
    lam = lam * 10;  
end  
  
res(it) = Etot;  
  
end
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหามาใช้ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

eone.txt

```
function [E,G,y,J] = eone(x,t,p,w,toggle)
%
% program to generate misfit and gradient of misfit given an
% input-output pair
%
% x(1:m) vector corresponding to given input
% t(1:n) vector corresponding to target output
% p number of nodes in hidden layer
% w(1:P) vector corresponding to weights -- P=p*(m+1)+n*(p+1)
% y(1:n) vector corresponding to network output
% E scalar for norm of (t-y)
% G(1:P) gradient of E with respect to w
% J(1:n,1:P) Jacobian of y with respect to w
% toggle 1 -> compute Jacobian, 0 otherwise
%
m=length(x);
n=length(t);
xp=[x;1];
tem=w(1:p*(m+1));
W1=reshape(tem,p,m+1);
tem=w(p*(m+1)+1:p*(m+1)+n*(p+1));
W2=reshape(tem,n,p+1);
a=W1*xp;
z=sigmoid(a);
zp=[z;1];
y=W2*zp;
E=0.5*norm(t-y)^2;
```

```
% gradient with respect to W2
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

cone.txt

```
del=y-t;
tem=del*zp';
g2=reshape(tem,n*(p+1),1);

% gradient with respect to W1
q=W2(:,1:p)'\*del;
tem=(sigmoid(a).\*q)*xp';
g1=reshape(tem,p*(m+1),1);
G=[g1;g2];
J=[];
if toggle ==1

% jacobian with respect to W2
J2=[];
for l=1:p+1
    J2=[J2 zp(l)*eye(n,n)];
end

% jacobian with respect to W1
J1=[];
for l=1:m+1
    J1=[J1 W2(:,l:p)].*(ones(n,1)*sigmoid(a'))\*xp(l)];
end
J=[J1 J2];
end
```



etot.txt

```
function [Etot,Gtot,Jtot,R]=etot(X,T,p,w,toggle)
% compute total misfit
% Etot total misfit
% Gtot gradient of total misfit wrt weight w
% R residual matrix, each column is residual vector
% Jtot jacobian of R wrt w
% toggle == 1 compute jacobian
%
[m,nex]=size(X);
[n,nex]=size(T);
% choose number of nodes in hidden layer
Etot=0; Gtot=zeros(size(w)); Jtot=[];
for k=1:nex
    x=X(:,k); t=T(:,k);
    [E,G,y,J]=eone(x,t,p,w,toggle);
    Etot=Etot+E;
    Gtot=Gtot+G;
    Jtot=[Jtot;J];
    R(:,k)=y-t;
end
```



sigmoid.txt

```
function s=sigmoid(t)
s=1./(1+exp(-t));

% Alternative
% s=(exp(t)-exp(-t))./(exp(t)+exp(-t));
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ปริญญานิพนธ์ฉบับนี้จะไม่สำเร็จล่วงลงไปได้ ถ้าไม่ได้รับความร่วมมือจากหลาย ๆ ฝ่ายที่สำคัญ โดยเฉพาะความอุปการะของ รศ. ดร. สุรพันธุ์ เอื้อไพบูลย์ อาจารย์ที่ปรึกษาโครงการนี้ ที่ได้ให้ความเอื้อเฟื้อทางด้านข้อมูล, อุปกรณ์ต่าง ๆ รวมทั้งการให้คำปรึกษาและการให้ความรู้ที่เป็นประโยชน์เกี่ยวกับโครงการเป็นอย่างดี รวมทั้งความช่วยเหลือของเพื่อน ๆ , รุ่นพี่ และรุ่นน้อง ที่ช่วยเป็นแบบภาพรูปภาพใบหน้าทีนำมาใช้เป็นฐานข้อมูล รวมทั้งให้คำปรึกษาและคอยเป็นกำลังใจให้ผ่านพ้นปัญหาต่าง ๆ ได้ด้วยดี จึงขอฝากคำขอบคุณทุก ๆ ท่าน มา ณ ที่นี้ด้วย

ถ้าหากปริญญานิพนธ์ฉบับนี้มีข้อบกพร่องหรือความผิดพลาดประการใด คณะผู้จัดทำขอน้อมรับ และจะทำการปรับปรุงให้ถูกต้อง

คณะผู้จัดทำ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หนังสืออ้างอิง

1. Louis J. Galbiati , Jr. , “Machine Vision and Digital Image Processing Fundamentals” ,
Prentice Hall , Inc. , 1990
2. Gregory A. Baxes , “Digital Image Processing Principles and Applications” ,
John Wiley & Sons , Inc. , 1994
3. Rafael C. Gonzalez and Richard E. Woods , “Digital Image Processing” ,
Addison – Wesley Publishing , 1992
4. รศ. มนต์ สัจวรศิลป์ , วรรัตน์ ภัทรอมรกุล , “คู่มือการใช้งาน MATLAB ฉบับสมบูรณ์” ,
อินโฟเพรส , 2000
5. Jie Yang and Alex Waibel , “A Real-Time Face Tracker”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้