สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การศึกษาและเปรียบเทียบประสิทธิภาพในการค้นหาข้อมูลรูปภาพ

(A study and comparison of image retrieval performance)

T061462

นาย ธีร์ โรจจวัฒน์
นาย ธีระวุฒิ กิ่งไพบูลย์

วพ
ร 621 ก
2547
เลขหมู่.....
เลขทะเบียน.....61462
วัน,เดือน,ปี.17 ก.ค. 2549
b.11596454
i.....

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2547

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การศึกษาและเปรียบเทียบประสิทธิภาพในการค้นหาข้อมูลรูปภาพ

(A study and comparison of image retrieval performance)

โดย

นาย ธีร์ โรจจวัฒน์

นาย ธีระวุฒิ กิ่งไพบูลย์

อาจารย์ที่ปรึกษา

อาจารย์ ธนา หงษ์สุวรรณ อาจารย์ที่ปรึกษา

ผศ.ดร. อรฉัตร จิตต์โสภักตร์ อาจารย์ที่ปรึกษา

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2547

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์ปีการศึกษา 2547

ภาควิชา วิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การศึกษาและเปรียบเทียบประสิทธิภาพในการค้นหาข้อมูลรูปภาพ

A study and comparison of image retrieval performance

ผู้จัดทำ

1. นาย ธีร์ โรจจวัฒน์ รหัสประจำตัว 44010217

2. นาย ธีระวุฒิ กิ่งไพบูลย์ รหัสประจำตัว 44010230

_____ อาจารย์ที่ปรึกษา

(อาจารย์ ธนา หงษ์สุวรรณ)

_____ อาจารย์ที่ปรึกษา

(ผศ.ดร. อรฉัตร จิตต์โสภักตร์)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ศึกษาและเปรียบเทียบประสิทธิภาพในการค้นหาข้อมูลรูปภาพ

นาย ธีร์ โรจจวัฒน์
นาย ธีระวุฒิ กิ่งไพบูลย์
อาจารย์ ธนา หงษ์สุวรรณ อาจารย์ที่ปรึกษา
ผศ.ดร. อรฉัตร จิตต์โสภักตร์ อาจารย์ที่ปรึกษา
ปีการศึกษา 2548

บทคัดย่อ

ในปัจจุบัน มีการใช้ข้อมูลทางด้านมัลติมีเดียมากขึ้นไม่ว่าจะเป็นใน ด้านบันเทิง ด้านธุรกิจ หรือด้านความปลอดภัย มัลติมีเดียที่สำคัญประเภทหนึ่งก็คือรูปภาพ ซึ่งมีการใช้งานกันอย่างแพร่หลายในหลายๆด้าน เนื่องจากรูปภาพนั้นสามารถสร้างขึ้นได้ง่ายไม่ว่าจะด้วยกล้องถ่ายภาพ การสแกนรูปภาพจากสิ่งต่างๆ หรือจากโทรศัพท์เคลื่อนที่รุ่นใหม่ๆที่มีความสามารถในการถ่ายภาพด้วยกล้องที่ติดอยู่ได้ด้วย ผู้ที่สร้างรูปภาพกับผู้ใช้งานรูปภาพนั้นอาจจะไม่ใช่คนคนเดียวกัน ดังนั้นเมื่อผู้ใช้งานต้องการรูปภาพที่ถูกสร้างขึ้นมาโดยคนอื่นจึงต้องมีการค้นคืนรูปภาพที่ต้องการในฐานข้อมูล สำหรับงานบางอย่างต้องการความถูกต้องมากกว่าความรวดเร็ว งานบางอย่างต้องการความรวดเร็วมากกว่าความถูกต้อง หรือสำหรับงานบางงานก็ต้องการทั้งความถูกมากที่สุดและใช้เวลาในการค้นคืนรูปภาพที่น้อยที่สุด

ในโครงงานนี้จะทำการศึกษาและเปรียบเทียบประสิทธิภาพการค้นหาข้อมูลรูปภาพเพื่อให้ได้ความถูกต้องและรวดเร็วที่สุด โดยเลือกใช้ลักษณะเฉพาะที่จะสามารถเป็นการอ้างถึงรูปภาพได้ดีมา 3 อย่างคือ ฮีสโตแกรม (Histogram) , ฟูริเยร์ดิสสคริปเตอร์ (Fourier Descriptor) , เกเบอร์ (Gabor) ลักษณะเฉพาะที่เลือกมาใช้ทั้ง 3 อย่างนี้จะสามารถทำให้เราสามารถค้นหาภาพได้อย่างถูกต้องและใช้ความเร็วในการค้นหาต่างกันแล้วแต่ขนาดของฐานข้อมูลและลักษณะของภาพ และในโครงงานนี้ยังได้เสนอแนวทางใหม่ที่ทำให้สามารถแบ่งแยกรูปภาพที่คล้ายกันออกเป็นกลุ่มๆ ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

A study and comparison of image retrieval performance

Thee Rojjawat

Theerawut Kingpaibool

Thana Hongsuwan Advisor

Asst.Prof.Dr. Orachat Chitsobhuk Advisor

ABSTRACT

Nowadays , using multimedia data is becoming more popular among user all around the world such as for entertainment , business security etc. One of most important kind of multimedia is "still picture" which is vary widely being used for many occasion because it is vary easy to construct by camera , scanning or built-in camera in mobile phone. Sometime the user may not be the same person as the creator. So when the users want to search for picture that are created by other they have to retrieve picture from database. Some work requires accuracy more then speed and some requires both in order to have the exact results at minimum image retrieval performance.

This project research contains the information about efficiency of picture searching and comparing among them by using Histogram, Fourier Descriptor and Gabor features to measure its efficiency. Histogram has been used for determine picture color while Fourier Descriptor has been used for determine its line characteristic and Gabor has been used for determine surface characteristic. However, the problem is some features take a very long time to process in searching methods. To save times, and this project research will avoid using sequential search.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้คงไม่อาจเสร็จได้ด้วยดี หากไม่ได้รับความช่วยเหลือและความร่วมมือจากหลายๆ ฝ่ายด้วยกัน บุคคลแรกที่ต้องกล่าวถึงเพราะเป็นบุคคลสำคัญที่ทำให้วิทยานิพนธ์นี้เสร็จลงได้ดีก็คือ ผศ.ดร.อรฉัตร จิตต์โสภักตร์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ให้คำแนะนำ และความช่วยเหลือเสมอมา และ อาจารย์ ธนา หงษ์สุวรรณ ที่ให้คำปรึกษาอย่างดียิ่ง ซึ่งต้องขอขอบพระคุณเป็นอย่างมาก

และต้องขอขอบพระคุณบุคคลที่สำคัญที่สุดที่ทำให้ข้าพเจ้ามีวันนี้ได้คือ บิดา มารดา อันเป็นที่เคารพรักอย่างยิ่ง ซึ่งได้เลี้ยงดูข้าพเจ้าเป็นอย่างดี พร้อมทั้งให้ข้าพเจ้ามีโอกาสศึกษาอย่างเต็มที่ และยังให้กำลังใจ เอาใจใส่เสมอมาในทุกๆด้านอันหาที่เปรียบมิได้ ข้าพเจ้าขอระลึกในพระคุณอันสุดประมาณและขอกราบขอบพระคุณมา ณ ที่นี้

สุดท้ายต้องขอขอบคุณเพื่อนๆ ทุกคนที่คอยให้คำปรึกษา ความช่วยเหลือ และกำลังใจทำให้ข้าพเจ้าทำวิทยานิพนธ์นี้สำเร็จลงได้ด้วยดี

ธีร์ โรจจวัฒน์
ธีระวุฒิ กิ่งไพบูลย์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้าที่
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VII
สารบัญภาพ	VIII
บทที่ 1 บทนำ	1
1.1 ความสำคัญและที่มา	1
1.2 วัตถุประสงค์ของงานวิจัย	2
1.3 ขอบเขตของงานวิจัย	2
1.4 เนื้อหาโดยรวมของงานวิจัย	2
1.5 วิธีการดำเนินงาน	3
บทที่ 2 ข้อมูลภาพดิจิตอล	4
2.1 กลุ่มสีที่ใช้แสดงผล	4
2.1.1 คุณสมบัติของกลุ่มสีที่แสดงผล	4
2.1.1.1 Uniformity	5
2.1.1.2 Completeness	5
2.1.1.3 Compactness	5
2.1.1.4 Naturalness	5
2.1.2 กลุ่มสีแสดงผลแบบ RGB	5
2.2 Geometric Transformation	8
2.2.1 Enlarge or Reduce by scaling factor	8
2.2.2 การทำ Forward Mapping	9
2.2.3 การทำ Backward Mapping	9
2.2.3.1 Zero-order Interpolation (nearest neighbor)	10
2.2.3.2 First-order interpolation (bilinear interpolation)	11
2.2.3.3 High-order interpolation (cublic interpolation)	12
บทที่ 3 ลักษณะเฉพาะของภาพ	13

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

3.1 ฮีสโตแกรม (Histogram) 13
3.1.1 การ Quantization 14
3.1.2 การ Normalization 15
3.2 ฟูริเยร์ดิสสคริปเตอร์ (Fourier Descriptor) 16
3.2.1 Descrete Fourier Transform (DFT) 16
3.2.2 Fast Fourier Transform (FFT) 17
3.2.2.1 Radix-2 DIF-FFT 17
3.2.2.2 บัเตอร์ฟลาย (Butterfly) 18
3.2.3 เปรียบเทียบการทำ FFT กับ DFT 20
3.2.4 การทำภาพเป็นภาพขาวดำ (Gray Scale) 20
3.2.5 การตรวจหาขอบ (Edge detection) 22
3.2.6 การทำคอนโวลูชั่น (Convolution) 24
3.3 เกเบอร์ (Gabor) 26
3.3.1 ตัวกรองเกเบอร์(Gabor Filter) 26
3.3.2 การแสดงออกของพื้นผิว (Texture Representation) 27

บทที่ 4 การทำงานจองระบบ
4.1 ขั้นตอนการทำฮีสโตแกรม 29
4.2 ขั้นตอนการหาฟูริเยร์ดิสสคริปเตอร์ 31
4.3 การทำเกเบอร์โดยใช้ Fast Fourier Transform (FFT) 33

บทที่ 5 การจัดเก็บและการค้นหา 39
5.1 โครงสร้างฐานข้อมูล 39
5.1.1 Identifier 39
5.1.2 ImageFeature 40
5.2 การแยกกลุ่มรูปภาพด้วยลักษณะของสี 41
5.3 การจัดเก็บ ลักษณะเฉพาะของรูปภาพ ลงในฐานข้อมูล 43
5.4 วิธีการค้นหารูปภาพ 43
5.4.1 ดึงรูปภาพขึ้นมาเปรียบเทียบแบบปกติ(ทั้งหมด) 43
5.4.2 ดึงรูปภาพขึ้นมาเปรียบเทียบโดยใช้ Key ช่วย(เฉพาะที่คล้าย) 43

บทที่ 6 การทดลองและสรุปผล 47

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

6.1 ชุดการทดลอง	49
6.2 ผลการทดลอง	50
6.3 สรุปผลการทดลอง	68
บทที่ 7 วิจารณ์และสรุป	69
7.1 สรุปและวิจารณ์ผลการทดลอง	69
7.2 ข้อดี – ข้อเสียในงานวิจัย	69
7.2.1 การแบ่งรูปภาพใหญ่เป็นรูปภาพย่อย	69
7.2.2 การแบ่งกลุ่มรูปภาพด้วยแนวโน้มสี	70
7.3 การนำไปประยุกต์ใช้งาน	70
บรรณานุกรม	72

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้าที่
2.1 แสดงการกำหนดระดับสี	7
3.1 ระดับของสีที่ใช้ในการทำ Quantization	15
3.2 แสดงการเปรียบเทียบจำนวนการคูณเลขเชิงซ้อนของ DFT กับ FFT	20
5.1 ตัวอย่างตารางฐานข้องมูล Idebtifier ที่เก็บชื่อและ Key ของแต่ละภาพ	39
5.2 แสดงตัวอย่างฐานข้อมูล ImageFeature ที่เก็บลักษณะเฉพาะของรูปภาพต่างๆ	40

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่		หน้าที่
2.1	แสดงความสัมพันธ์ของสีในกลุ่มสีแสดงผลแบบ RGB	6
2.2	แสดงภาพ RGB และค่าในพิกเซลของภาพ RGB	7
2.3	แสดงการผสมสีในรูปแบบ RGB	8
2.4	แสดงภาพที่ทำการขยายโดยวิธี nearest neighbor	10
2.5	แสดงภาพที่ทำการขยายโดยวิธี bilinear interpolation	11
2.6	แสดงภาพที่ทำการขยายโดยวิธี cublic interpolation	12
3.1	แสดงภาพที่ทำฮีสโตแกรม	14
3.2	แสดงภาพที่มีกราฟฮีสโตแกรมเหมือนกันแต่ภาพต่างกัน	14
3.3	กราฟแสดงคู่ลำดับเชิงซ้อน	16
3.4	แสดงบัตเตอร์ฟลายของการคูณเลขเชิงซ้อน	18
3.5	แสดงการทำ บัตเตอร์ฟลาย N=4	19
3.6	แสดงค่าพิกเซลของภาพขาวดำ	21
3.7	แสดงภาพสีและภาพขาวดำ	21
3.8	การคำนวณ หาเส้นขอบโดยใช้ sobel mask	23
3.9	การคำนวณพิกเซลที่อยู่ขอบของภาพ	23
3.10	ภาพที่ก่อนนำมาตรวจสอบหาขอบและหลังนำมาตรวจสอบหาขอบ	24
3.11	ข้อมูลภาพต้นแบบและ kernel ที่ทำคอนโวลูชั่น	25
4.1	แสดงขั้นตอนการทำฮีสโตแกรม	30
4.2	แสดงขั้นตอนการทำ Fourier Transform	32
4.3	การขยายหน้ากากให้มีขนาดเท่ากับขนาดของรูปภาพย่อย โดยการเติม 0	33
4.4	แสดงการนำหน้ากากมาคำนวณ FFT ตามแนวแกน x และ y	34
4.5	แสดงการนำรูปภาพย่อยมาคำนวณ FFT ตามแนวแกน x และ y	35
4.6	แสดงการคูณระหว่างหน้ากากใหม่กับรูปภาพย่อยใหม่	36
4.7	แสดงการนำรูปภาพย่อยที่ได้จากการคูณมาทำ Inverse FFT	37
4.8	แสดงขั้นตอนการทำเกเบอร์โดยใช้วิธี FFT	38
5.1	ตำแหน่งของรูปภาพย่อยที่พิจารณา	41
5.2	รูปแบบการเปลี่ยนแปลงค่าเฉลี่ยสีในรูปภาพ	42
5.3	แสดงการแบ่งรูปภาพย่อยและหาลักษณะเฉพาะของรูปภาพย่อย	44
5.4	การจัดเก็บรูปภาพลงฐานข้อมูล	45
5.5	แสดงขั้นตอนการค้นหารูปภาพ	46
6.1	แบ่งภาพเป็น 9 รูปภาพย่อย	47

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

6.2 การแบ่งรูปภาพเป็นกลุ่มๆในฐานข้อมูล 47

6.3 กราฟแสดงเวลาที่ใช้ในการค้นหาของชุดรูปภาพ a โดยใช้ชุดการทดลองต่างๆ 52

6.4 กราฟแสดงเวลาที่ใช้ในการค้นหาของชุดรูปภาพ b โดยใช้ชุดการทดลองต่างๆ 55

6.5 กราฟแสดงเวลาที่ใช้ในการค้นหาของชุดรูปภาพ c โดยใช้ชุดการทดลองต่างๆ 58

6.6 กราฟแสดงเวลาที่ใช้ในการค้นหาของชุดรูปภาพ d โดยใช้ชุดการทดลองต่างๆ 61

6.7 กราฟแสดงเวลาที่ใช้ในการค้นหาของชุดรูปภาพ e โดยใช้ชุดการทดลองต่างๆ 64

6.8 กราฟแสดงเวลาที่ใช้ในการค้นหาของชุดรูปภาพ f โดยใช้ชุดการทดลองต่างๆ 67

6.9 กราฟเปรียบเทียบเวลาที่ในการค้นหาแต่ละชุดรูปภาพ 68

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความสำคัญและที่มา

ในปัจจุบันรูปภาพเป็นข้อมูลที่นับได้ว่ามีความสำคัญในด้านต่างๆ มากมาย ทั้งที่ใช้รูปภาพเพื่อเก็บเหตุการณ์เรื่องราวต่างๆที่เกิดขึ้น หรือใช้รูปภาพเป็นการอ้างถึงตัวตนเพื่อความปลอดภัย เช่น รูปติดบัตรต่างๆ เป็นต้น นอกจากนั้นรูปภาพในปัจจุบันยังสามารถถูกสร้างขึ้นได้อย่างง่ายกว่าอดีต เช่น กล้องถ่ายรูปดิจิตอล (Digital Camera) การจับภาพจากกล้องถ่ายวีดีโอ (Vedio Camera) การสแกนรูปภาพด้วยสแกนเนอร์ (Scanner) โทรศัพท์เคลื่อนที่ (Mobile Phone) บางรุ่นที่มีกล้องถ่ายภาพติดอยู่ เป็นต้น จึงทำให้รูปภาพกลายเป็นข้อมูลที่นิยมใช้งานมากขึ้น

หลังจากที่มีการสร้างภาพที่เป็นลักษณะของดิจิตอลขึ้นมาแล้ว ในขั้นตอนการค้นหานั้นถ้าหากเรามีรูปภาพที่เก็บไว้จำนวนไม่มาก เราก็สามารถค้นหารูปภาพที่ต้องการได้ด้วยตัวเราเอง แต่ถ้าหากเรามีการเก็บรูปภาพไว้จำนวนมากขึ้นและยังคงค้นหารูปภาพที่ต้องการนั้นด้วยตัวเองก็อาจจะสร้างความลำบากขึ้นได้ ฉะนั้นเราจึงมีการนำโปรแกรมคอมพิวเตอร์ (Computer Program) เข้ามาช่วยในการค้นหาเพื่อความถูกต้องและรวดเร็ว

เนื่องจากว่าในขั้นตอนการแยกแยะรูปภาพด้วยคนนั้น เราใช้สายตาในการมองความแตกต่างแต่ละรูปภาพ แต่คอมพิวเตอร์นั้นเราต้องหาวิธีในการที่จะให้คอมพิวเตอร์สามารถแยกแยะความแต่ต่างแต่ละรูปภาพได้ ดังนั้นเราจะต้องหาลักษณะเฉพาะ(Feature) ของภาพแต่ละภาพเพื่อให้คอมพิวเตอร์นั้นใช้ค่าเหล่านี้ในการเปรียบเทียบแต่ละรูปภาพ

ปัจจุบันมีการหาลักษณะเฉพาะของรูปภาพมากมายหลายวิธี สังเกตว่าวิธีต่างๆ ที่สร้างขึ้นมานั้นก็สร้างขึ้นมาจากวิธีในการแยกแยะรูปภาพของคนเรา เช่น การมองลักษณะของวัตถุในรูปภาพ การมองลักษณะของสี การมองลักษณะของลายเส้น เป็นต้น ซึ่งแต่ละวิธีก็มีข้อดี ข้อเสียแตกต่างกันไป บางวิธีอาจเหมาะกับภาพลักษณะหนึ่ง แต่ไม่เหมาะกับรูปภาพอีกลักษณะหนึ่ง บางวิธีอาจใช้เวลาน้อยในการแบ่งแยกรูปภาพแต่ได้ผลไม่ดี แต่บางวิธีอาจใช้เวลามากแต่สามารถแยกแยะรูปภาพได้ดี

ดังนั้นเราจึงต้องศึกษาลักษณะเฉพาะต่างๆ ของรูปภาพว่าอย่างไหนมีข้อดี ข้อเสียแตกต่างกันอย่างไร และถ้าเรานำวิธีหนึ่งมาใช้ควบคู่กับอีกวิธีหนึ่งจะได้ผลดีขึ้นมากน้อยเพียงใด โดยในงานวิจัยนี้เราสนใจเลือกลักษณะเฉพาะของรูปภาพมา 3 อย่างเพื่อทำการศึกษา คือ ฮีสโตรแกรม (Histogram) , ฟูริเยร์ดิสสคริปเตอร์ (Fourier Descriptor) , เกเบอร์ (Gabor) ซึ่งแต่ละอย่างก็จะมีจุดสนใจแตกต่างกันไป

นอกจากนั้นในงานวิจัยนี้เราได้เสนอแนวทางในการหาลักษณะเฉพาะของรูปภาพซึ่งจะมีผลทำให้การค้นหารูปภาพนั้นมีประสิทธิภาพในความตรงประเด็นมากขึ้น โดยมีแนวคิดมาจากวิธีการวาดภาพเหมือนซึ่งจะมีการแบ่งภาพออกเป็นส่วนๆ ก่อนที่จะคัดลอกทีละส่วนของภาพต้นแบบ ไปอีกภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หนึ่งที่ต้องการ ในงานวิจัยนี้ได้ประยุกต์ใช้วิธีในการวาดภาพโดยเราจะแบ่งรูปภาพออกเป็นรูปภาพย่อย (Subimage) ก่อนที่จะทำการลักษณะเฉพาะของแต่ละรูปภาพย่อย

ในการนำลักษณะเฉพาะของรูปภาพในฐานข้อมูลขึ้นมาเปรียบเทียบกับรูปภาพที่ต้องการนั้น เราต้องนำทุกภาพขึ้นมาเปรียบเทียบ แต่ในงานวิจัยนี้การเลือกรูปภาพขึ้นมาจากฐานข้อมูลเราจะเลือกรูปภาพย่อยขึ้นมาก่อน ถ้าเห็นว่าไม่เหมือนกันจะถือว่ารูปภาพนั้นไม่ตรงประเด็น แต่ยังไงเราก็ยังต้องเลือกรูปภาพย่อยของทุกรูปภาพทั้งหมดขึ้นมาเพื่อเปรียบเทียบอยู่ดี ดังนั้นเราจึงหาวิธีที่จะแบ่งแยกรูปภาพที่มีความคล้ายคลึงกันให้เป็นกลุ่มเดียวกันเสีย ทำให้ในขั้นตอนเราเลือกรูปภาพในฐานข้อมูลขึ้นมานั้นเราก็จะเลือกเฉพาะรูปภาพที่มีความคล้ายอยู่ในกลุ่มเดียวกันขึ้นมาเปรียบเทียบ ซึ่งจะช่วยให้เราสามารถค้นหารูปภาพที่ต้องการได้รวดเร็วยิ่งขึ้น

1.2 วัตถุประสงค์ของงานวิจัย

1.2.1 เพื่อเป็นการการศึกษาและทดลองเปรียบเทียบลักษณะเฉพาะของรูปภาพว่าแต่ละอย่างว่ามีประสิทธิภาพในการค้นหาแตกต่างกันอย่างไร

1.2.2 เพื่อเป็นการศึกษาการใช้ลักษณะเฉพาะของรูปภาพแต่ละอย่าง ให้ทำงานร่วมกันได้อย่างมีประสิทธิภาพ ในด้านที่มีความตรงประเด็นมากที่สุด และใช้เวลาการค้นหาน้อยที่สุด

1.2.3 เพื่อนำเสนอแนวคิดการหาลักษณะของภาพอย่างคราวๆ ให้ใช้งานร่วมกับการใช้ลักษณะเฉพาะในการค้นหารูปภาพให้มีความรวดเร็วยิ่งขึ้น

1.3 ขอบเขตของงานวิจัย

เนื่องจากงานวิจัยนี้เป็นงานวิจัยการทดลองซึ่งเกี่ยวข้องกับลักษณะเฉพาะของรูปภาพแค่ฮีสโตรแกรม , ฟูริเยร์ดิสสคริปเตอร์ , เกเบอร์ เท่านั้น ซึ่งลักษณะเฉพาะนี้ก็จะมีความสามารถที่แตกต่างกันไป ดังนั้นเราจึงมีต้องมีการกำหนดขอบเขตและข้อจำกัดเพื่อให้ผลที่ได้จากการทดลองนั้นมีความถูกต้องมากที่สุด ซึ่งเราจะแบ่งขอบเขตออกเป็น 3 จำพวก คือ

ขอบเขตเกี่ยวกับรูปภาพ รูปภาพที่ใช้ในการทดลองต้องมีขนาด 24 บิต(Bit) เท่านั้น และมีความชัดเจนของลายเส้นและสีระดับหนึ่ง ส่วนผลลัพธ์ที่ได้นั้นจะต้องเหมือนกับต้นแบบทุกประการ

ขอบเขตเกี่ยวกับฐานข้อมูล จำนวนรูปภาพที่อยู่ในฐานข้อมูลมีจำนวน 800 รูปภาพ โดยที่ไม่มีรูปภาพคู่ใดในฐานข้อมูลที่เหมือนกันทุกประการ

ขอบเขตเกี่ยวกับลักษณะเฉพาะของรูปภาพ เลือกมาใช้ในการทดลอง 3 อย่างคือ ฮีสโตแกรม , ฟูริเยร์ดิสสคริปเตอร์ , เกเบอร์ โดยค่าคงที่ต่างๆ ที่ใช้ในลักษณะต่างๆนั้นต้องไม่มีการเปลี่ยนแปลง

1.4 เนื้อหาโดยรวมของงานวิจัย

งานวิจัยนี้ได้กล่าวถึงการเลือกกลุ่มสีที่แสดงผล(Color Space) รวมทั้งรายละเอียดของกลุ่มสีที่แสดงผลที่ใช้ในศึกษาและการทดลอง และกล่าวถึงการเลือกลักษณะเฉพาะของรูปภาพที่ใช้ในการค้นหาภาพโดยเลือกมาศึกษา 3 อย่างคือ ฮีสโตแกรม , ฟูริเยร์ดิสสคริปเตอร์ และเกเบอร์ ซึ่งเป็นลักษณะเฉพาะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของรูปภาพที่มีประสิทธิภาพในการค้นหาภาพได้ดี และอธิบายวิธีการหาลักษณะเฉพาะของรูปภาพต่างๆ ในงานวิจัยนี้จึงกล่าวถึงวิธีการทำฮีสโตแกรมซึ่งเป็นวิธีที่ใช้บอกลักษณะสีของภาพ การทำฟูริเยร์ดิสสคริปเตอร์ซึ่งใช้บอกลักษณะของลายเส้นของภาพ การทำเกเบอร์ซึ่งใช้บอกลักษณะของพื้นผิวของภาพ ในการหาลักษณะเฉพาะของรูปภาพบางตัวยังมีความจำเป็นต้องทำภาพให้เป็นลักษณะขาว-ดำ (Gray Scale) ก่อน และมีการตรวจหาขอบ (Edge Detection) ด้วย ดังนั้นในงานวิจัยนี้จะกล่าวถึงวิธีการทำภาพให้เป็นขาว-ดำ กับ การทำการตรวจสอบขอบเพื่อหาเส้นขอบของวัตถุในรูปภาพ และในงานวิจัยยังเสนอวิธีในการค้นหารูปภาพ และยังได้ทำการทดลองเพื่อเปรียบเทียบความเหมาะสมของวิธีการค้นหาและลักษณะเฉพาะที่ใช้แต่ละแบบว่าเหมาะสมกับงานที่ต้องการหรือไม่อย่างไร เพราะว่าในงานบางงานไม่ได้คำนึงถึงความถูกต้องของการค้นหาเพียงอย่างเดียว แต่ยังคำนึงถึงเวลาที่ใช้ในการค้นหาด้วย

1.5 วิธีการดำเนินงาน

งานวิจัยชิ้นนี้เริ่มต้นด้วยการศึกษาข้อมูลรูปภาพดิจิตอล (Digitalimage Data) การประมวลผลของรูปภาพ (Image Processing) ซึ่งมีเรื่องหลักอยู่ 2 เรื่องคือ เรื่องเกี่ยวกับรูปภาพ และเรื่องเกี่ยวกับลักษณะเฉพาะของรูปภาพที่ใช้ในงานวิจัยชิ้นนี้ ซึ่งจะกล่าวถึงในบทที่ 2 และ 3 เพื่อเป็นการทำความเข้าใจข้อมูลต่างๆเกี่ยวกับรูปภาพและการใช้งานลักษณะเฉพาะนั้นๆ

หลังจากนั้นจะเป็นขั้นตอนในการออกแบบฐานข้อมูลเพื่อจัดเก็บรูปภาพ รวมถึงวิธีในการจัดเก็บรูปภาพและค้นหาว่ามีขั้นตอนการทำงานอย่างไร ดังจะกล่าวในบทที่ 4 เมื่อเราจัดเตรียมฐานข้อมูลเรียบร้อยแล้ว ก็มาถึงขั้นตอนการออกแบบการทดลอง,การทดลอง และสรุปผลการทดลอง ซึ่งในขั้นตอนของการออกแบบการทดลองนั้น เราใช้ชุดการทดลอง (Senario) ที่คาดว่าน่าจะมีประสิทธิภาพในการค้นหารูปภาพได้ดี ซึ่งมีรายละเอียดในบทที่ 5

เมื่อมีข้อสรุปในการทดลองชิ้นนี้แล้ว เราก็จะนำผลการทดลองที่ได้ไปวิเคราะห์หาข้อดี ข้อเสียที่เกิดขึ้นในงานวิจัยชิ้นนี้ รวมทั้งเสนอแนวทางแก้ไขปรับปรุงซึ่งจะเป็นประโยชน์ในการนำงานวิจัยชิ้นไปประยุกต์ใช้งานต่อไปในอนาคต ในบทที่ 6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ข้อมูลภาพดิจิตอล (Digital Image Data)

การทำการประมวลผลรูปภาพต้องศึกษาพื้นฐานของรูปภาพก่อนว่ามีการสร้างขึ้นได้อย่างไร ซึ่งในรูปภาพแต่ละรูปจะมีองค์ประกอบต่างๆ มากมาย แต่องค์ประกอบที่จะนำมาประมวลผลนั้นเรียกว่า " พิกเซล(Pixel) " พิกเซลนั้นจะมีลักษณะเป็นจุดสี่เหลี่ยม ซึ่งเป็นองค์ประกอบเล็กๆ ที่มีอยู่มากมายบนภาพ การประมวลผลภาพต้องนำพิกเซลแต่ละรูปภาพมาทำการประมวลผล โดยที่แต่ละพิกเซลจะประกอบด้วยรายละเอียดสีต่าง ๆ มารวมกันเป็นพิกเซล ซึ่งเรียกกลุ่มของสีที่ใช้ว่า "Color Space" ซึ่งรูปแบบการบอกรายละเอียดของสีบนรูปภาพ โดยกลุ่มสีที่แสดงผลนั้นมีอยู่หลายแบบในที่นี้เราจะกล่าวถึงกลุ่มสีที่แสดงผลที่ใช้ในงานวิจัยนี้เท่านั้น

2.1 กลุ่มสีที่ใช้แสดงผล

การเกิดภาพต่างๆนบจอคอมพิวเตอร์ จะเกิดจากการเรียงตัวของจุดสี่เหลี่ยม สีต่างๆ ที่เราเรียกว่า " พิกเซล " ซึ่ง " พิกเซล " เป็นตัวกำหนดความคมชัดของภาพในจอคอมพิวเตอร์ ยิ่งเรากำหนดให้ภาพแต่ละภาพมีจำนวน " พิกเซล " มาก ภาพก็จะมีความละเอียดมากขึ้นตามไปด้วย แต่จะทำให้ไฟล์มีขนาดใหญ่ขึ้นด้วยตามความละเอียดของภาพ

พิกเซล (Pixel มาจากคำว่า "Picture Element") เป็นหน่วยพื้นฐานของสีที่สามารถโปรแกรมได้บนจอแสดงผลหรือบนภาพ ขนาดทางกายภาพของพิกเซลจะขึ้นอยู่กับการกำหนดความละเอียดของจอภาพ ถ้าตั้งการแสดงผลให้มีความละเอียดสูงสุด ขนาดทางกายภาพของพิกเซลก็จะเท่ากับขนาดทางกายภาพของจุด (Dot Pitch หรือ Dot Size) บนจอภาพ แต่ถ้าตั้งความละเอียดไว้ต่ำกว่าค่าสูงสุด พิกเซลจะมีขนาดใหญ่กว่าขนาดทางกายภาพของจุดบนจอภาพ นั่นคือแต่ละพิกเซลจะใช้มากกว่า 1 จุดในการแสดงผล

ในปัจจุบันนี้กลุ่มสีที่แสดงผลมีมากมายหลายแบบ เช่น กลุ่มสีที่แสดงผล RGB เป็นกลุ่มสีที่แสดงผลที่บอกแสดงลักษณะของสี 3 สี คือ สีแดง สีเขียว และสีน้ำเงิน , กลุ่มสีที่แสดงผล HSV เป็นกลุ่มสีที่แสดงผลที่บอกลักษณะของภาพโดยใช้ ระดับสี , ความเข้มของสี , ความสว่างของสี เป็นต้น แต่ในงานวิจัยนี้เราเลือกใช้กลุ่มสีที่แสดงผล RGB ในการแสดงรายละเอียดของภาพ

2.1.1 คุณสมบัติของกลุ่มสีที่แสดงผล

โดยทั่วไปแล้วภาพจะอยู่ในรูปแบบของกลุ่มสีที่แสดงผล RGB ซึ่งแทนสีทั้งหมดในกลุ่มสีที่แสดงผลด้วยเวกเตอร์ 3 มิติ เราจึงใช้กลุ่มสีที่แสดงผลนี้เป็นจุดเริ่มต้นของลักษณะเฉพาะของภาพได้โดยใช้ลักษณะเฉพาะของสี (Color Feature) จากนั้นค่อยทำการลดระดับของสีให้เหมาะสมต่อการทำฮีสโตแกรม คุณสมบัติดังกว่าลคือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.1.1 Uniformity

Uniformity เป็นคุณสมบัติที่ใช้ในการเปรียบเทียบความใกล้เคียงของเมตริกสีโดยกลุ่มสีที่แสดงจะมีคุณสมบัตินี้จะต้องคำนวณได้ง่าย อย่างเช่น การหาความใกล้เคียงของเมตริกสีโดยใช้ฟังก์ชั่นที่ไม่ขึ้นกับตำแหน่งในกลุ่มสีนั้น การทำการแปลงสีจะเป็นวิธีที่ทำให้กลุ่มสีมีคุณสมบัตินี้

2.1.1.2 Completeness

Completeness เป็นคุณสมบัติสำคัญที่กลุ่มสีที่แสดงผลจะต้องมีสีที่เราสามารถแยกแยะได้อย่างชัดเจน สามารถทำได้โดยการทำการลดระดับสี

2.1.1.3 Compactness

Compactness เป็นคุณสมบัติที่สีในกลุ่มสีที่แสดงผลจะต้องแตกต่างจากสีอื่นๆ ในกลุ่มสีที่แสดงผลเดียวกัน คุณสมบัตินี้จะเป็นตัวกำหนดมิติของภาพเพื่อให้แน่ใจว่า เมื่อผ่านการเปรียบเทียบจากกลุ่มสีที่แสดงผล RGB แล้วจะไม่มีสี 2 สีที่เหมือนกัน การลดระดับสีจะทำให้เกิดคุณสมบัตินี้

2.1.1.4 Naturalness

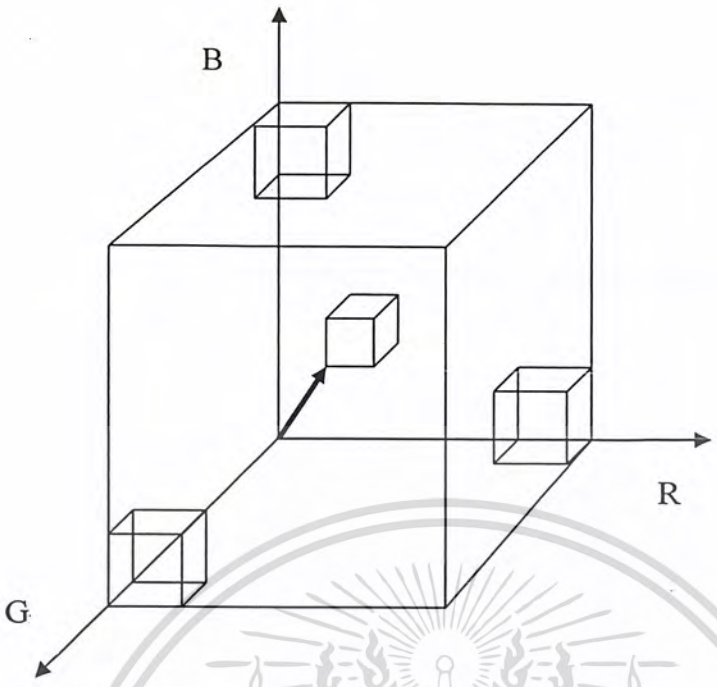
Naturalness เป็นคุณสมบัติของสีที่ผู้ใช้สามารถเข้าใจได้ง่าย ทำให้ผู้ใช้สามารถทำการดึงข้อมูลจากฐานข้อมูลโดยใช้สีเป็นดัชนีได้ ซึ่งสีดังกล่าวจะถูกแยกแยะออกเป็น 3 ส่วน คือ ความสว่าง (Brightness) , ระดับสี (Hue) , ความเข้มของสี (Saturation)

- ความสว่าง คือ คุณสมบัติที่มีผลกระทบโดยตรงต่อสีในด้านของแสง เช่น การเปลี่ยนแปลงระหว่างความสว่างถึงความมืด (Bright to Dim)
- ระดับสี คือ คุณสมบัติที่มีผลกระทบต่อสีที่จะทำให้สีเหมือนกับ สีแดง สีเหลือง สีเขียว สีน้ำเงิน สีฟ้าสีม่วง ฯลฯ
- ความเข้มของสี คือ คุณสมบัติที่มีผลกระทบต่อสีหนึ่งๆของระดับสี โดยพิจารณาถึงปริมาณของสี (Color Light) โดยไม่สนใจถึงความสว่าง

2.1.2 กลุ่มสีแสดงผลแบบ RGB

ภาพสีที่อยู่ในรูปแบบของ RGB จะเป็นรูปแบบพื้นฐานของภาพที่ใช้แสดงบนจอคอมพิวเตอร์ โดยแต่ละ พิกเซลของภาพจะประกอบด้วยองค์ประกอบของสี 3 ค่า คือสีแดง,สีเขียว,สีน้ำเงิน โดยแสดงองค์ประกอบสีสีละ n บิต ดังนั้นจะสามารถแสดงเฉดสีได้องค์ประกอบสีละ 2^n เฉดสี ที่แตกต่างกัน รูปที่ 2.1 แสดงความสัมพันธ์ของสีแดง สีเขียว และสีน้ำเงิน ใน 1 พิกเซล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.1 แสดงความสัมพันธ์ของสีในกลุ่มสีแสดงผลแบบ RGB

ในระบบแกน R จะแสดงปริมาณของสีแดงที่มีในพิกเซลนั้น แกน G จะแสดงปริมาณของสีเขียวที่มีในพิกเซลนั้น และแกน B จะแสดงปริมาณของสีน้ำเงินที่มีในพิกเซลนั้น เนื่องจากเก็บสีแยกเป็นคนละค่าของสีกันทำให้เมื่อมีการคำนวณต้องคำนวณแยกทีละค่าของสี

การผสมสีบนคอมพิวเตอร์นั้นอาศัยระดับสีหลัก 3 สีคือ สีแดง สีเขียว และ น้ำเงิน หรือที่เราเรียกว่า "RGB Colors" (Red-Green-Blue) นั้นเอง ความเหมือนจริงของสีคอมพิวเตอร์นั้น ขึ้นอยู่กับในหนึ่งจุดพิกเซลของการแสดงผลนั้นใช้ระดับของสี (Color depths) ที่มีค่าเท่าไหร่ เช่นถ้าสี RGB มีระดับของสีเป็น 8 planes นั่นคือ เราใช้ 8 บิตเก็บข้อมูลหนึ่งสี หมายความว่า เฉพาะระดับสีหลัก เช่น สีแดง สีเดียวก็จะมีความเป็นแดง อยู่ถึง $2^8 = 256$ ระดับ พูดง่ายๆ ก็คือว่า สมมติเราจะผสมสี หนึ่งสี จาก แดง-เขียว-น้ำเงิน (RGB) เราจะใช้สีแดง กี่ส่วนจาก 0 ถึง 255 ส่วน ใช้เขียว กี่ส่วนจาก 0 ถึง 255 และเช่นเดียวกัน สีน้ำเงินกี่ส่วน จาก 0 ถึง 255 สีที่เกิดขึ้นก็จะเกิดจากการผสมของสีทั้งสาม ในอัตราส่วนต่างๆ กัน เช่น สีเหลืองธรรมดา เกิดจากการผสมสี แดง 255 ส่วน สีเขียว 255 ส่วน และสีน้ำเงิน 0 ส่วน

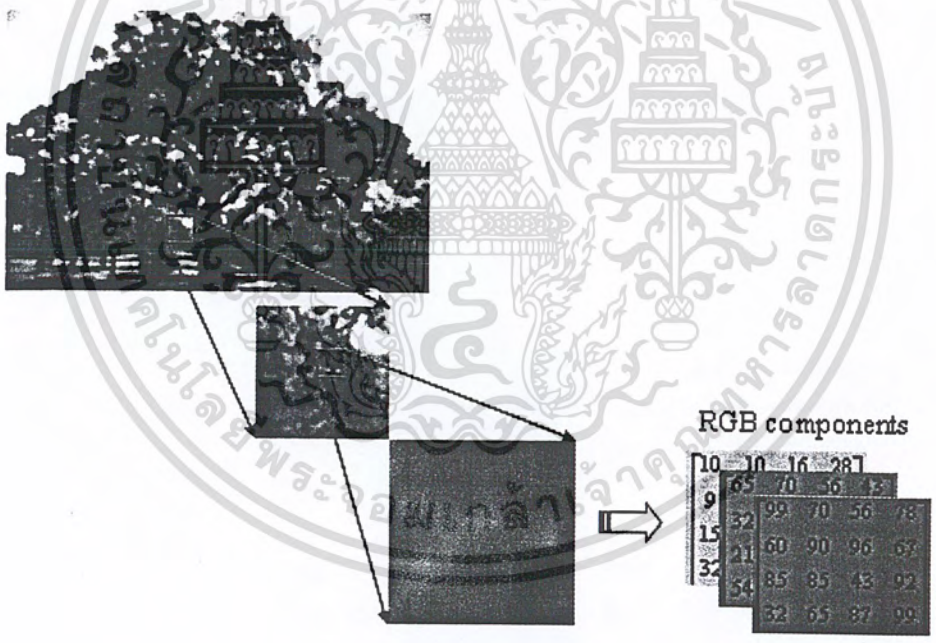
จากข้างต้น ระดับของสีแดง มีถึง 256 ระดับ สีเขียว 256 ระดับ สีน้ำเงิน 256 ระดับ ดังนั้น RGB ทั้งหมดใช้ 24 บิต (8+8+8) ในการแสดงสี RGB ของหนึ่งจุดพิกเซลซึ่งสามารถแสดงสีได้มากถึง 256 x 256 x 256 = 16.7 ล้านสี ตารางที่ 2.1 แสดงถึงจำนวนสีที่เป็นไปได้ ซึ่งขึ้นกับจำนวนบิตที่ใช้กำหนดระดับของสี และรูปที่ 2.2 แสดงการผสมสีของแม่สีในระบบ RGB

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จำนวนบิตที่ใช้เก็บสีต่อหนึ่งจุด	Color Mode Name	จำนวนสีที่แสดงได้
1	Black and White	2
4	16-Color(EGA)	16
8	Pseudo Color	256
16	Hi-Color	65,536
24	True-Color	16,777,216

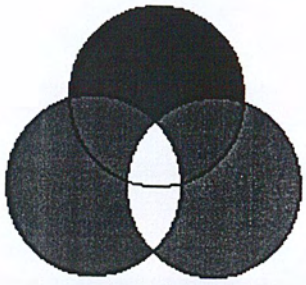
ตารางที่ 2.1 แสดงการกำหนดระดับสี

ค่าในแต่ละพิกเซลของสีของรูปภาพจะประกอบไปด้วยเวกเตอร์(Vector) ที่แสดงค่าของสีแดง สีเขียว และสีน้ำเงิน อย่างละ 8 บิต ดังนั้นรูปภาพ RGB 1 พิกเซล จะประกอบไปด้วยจำนวนบิตทั้งหมด 24 บิต ดังรูปที่ 2.2



รูปที่ 2.2 แสดงภาพ RGB และค่าในพิกเซลของภาพ RGB

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



RGB

รูปที่ 2.3 แสดงการผสมสีในรูปแบบ RGB

2.2 Geometric Transformtion

2.2.1 Enlarge or Reduce by scaling factor

การทำเปลี่ยนแปลงขนาดเป็นการขยายและย่อขนาดของรูปภาพซึ่งมีรูปแบบดังนี้

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} S & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

โดย

x' คือ ตำแหน่งของ x ใหม่ที่ทำการแปลงขนาดแล้ว

y' คือ ตำแหน่งของ y ใหม่ที่ทำการแปลงขนาดแล้ว

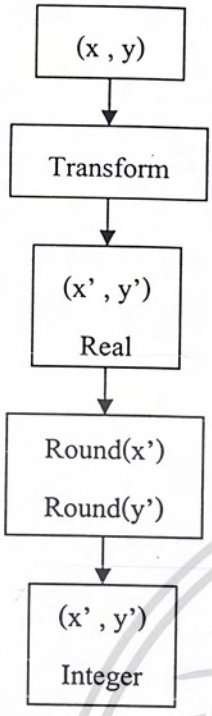
x คือ ตำแหน่งของพิกเซลสมแนวแกน x

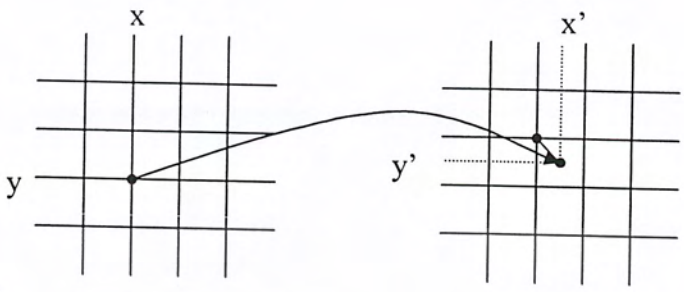
y คือ ตำแหน่งของพิกเซลสมแนวแกน y

S คือ ค่าคงที่ที่ใช้ทำการขยาย หรือ ย่อ รูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.2 การทำ Forward Mapping





f(x , y)

x , y real

g(x , y)

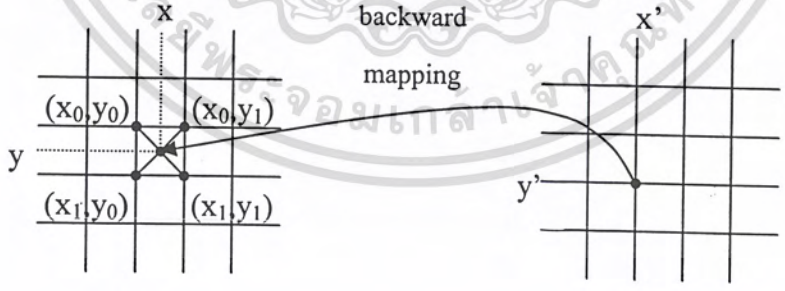
x' , y' integer

Transform

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_0 & a_1 \\ b_0 & b_1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$$

$$g(round(x'),(round(y')) = f(x,y)$$

2.2.3 การทำ Backward Mapping



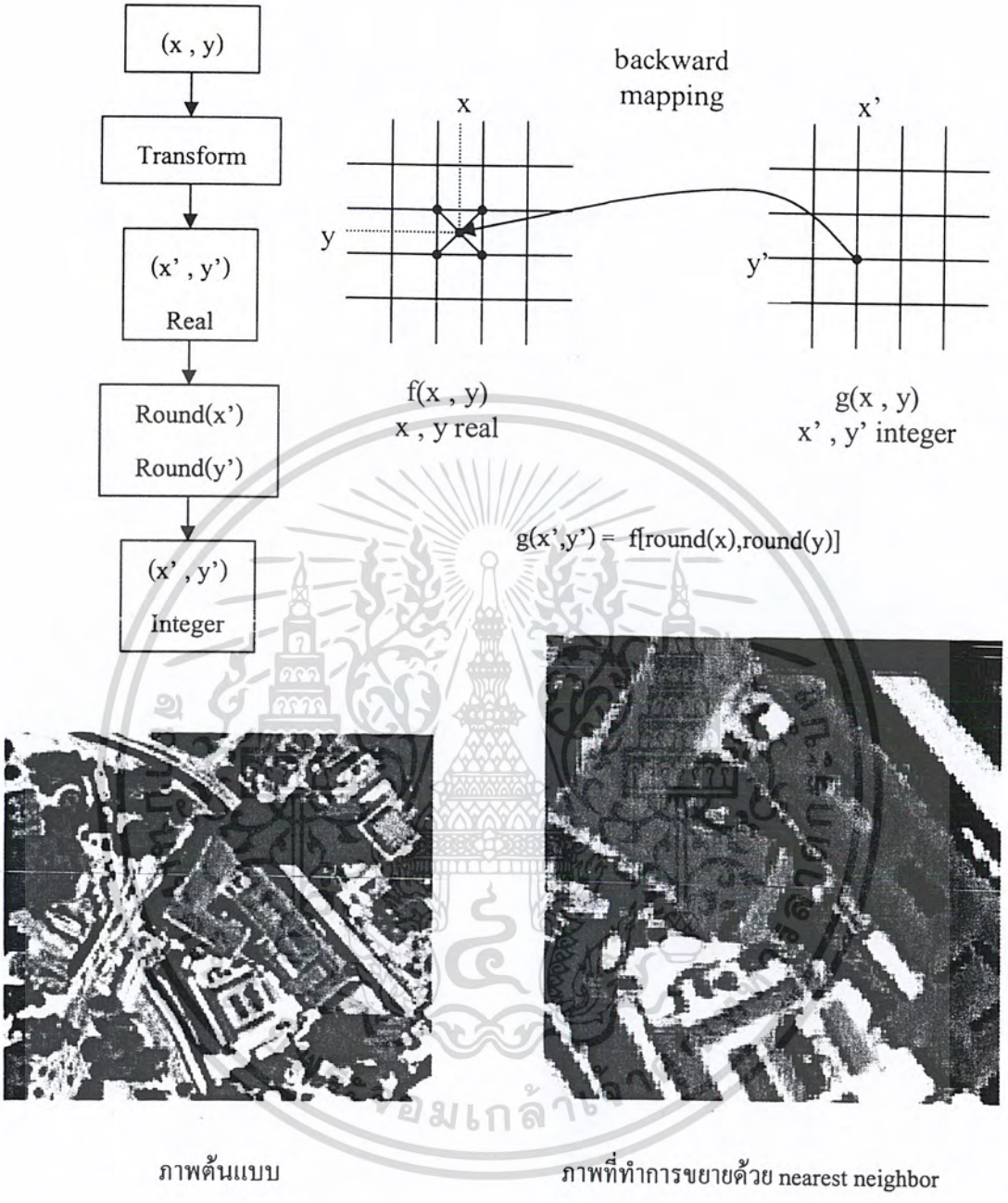
f(x , y)
x , y real

g(x , y)
x' , y' integer

$$g(x',y') = \text{interpolation}[f(x_0,y_0)]$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

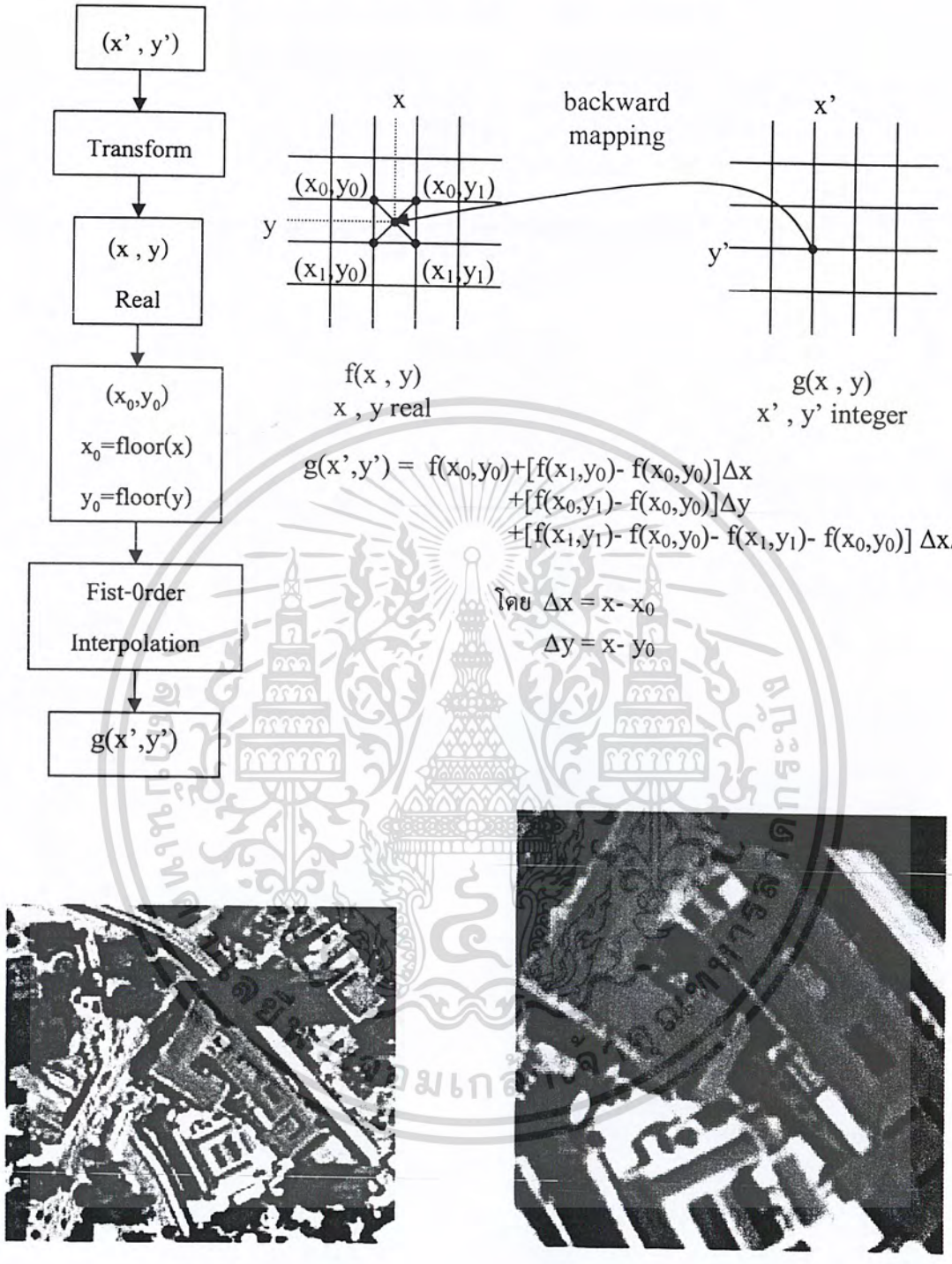
2.2.3.1 Zero-order Interpolation (nearest neighbor)



รูปที่ 2.4 แสดงภาพที่ทำการขยายโดยวิธี nearest neighbor

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.3.2 First-order interpolation (bilinear interpolation)



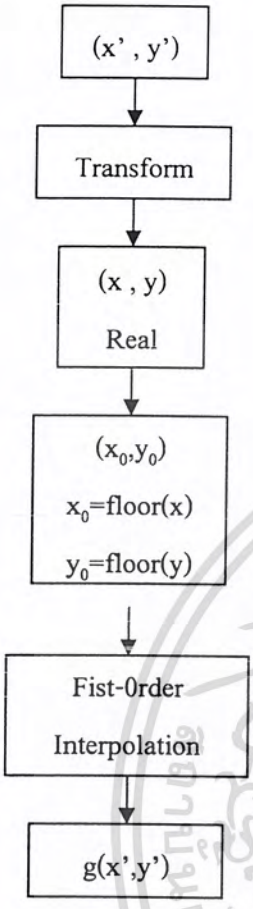
ภาพต้นแบบ

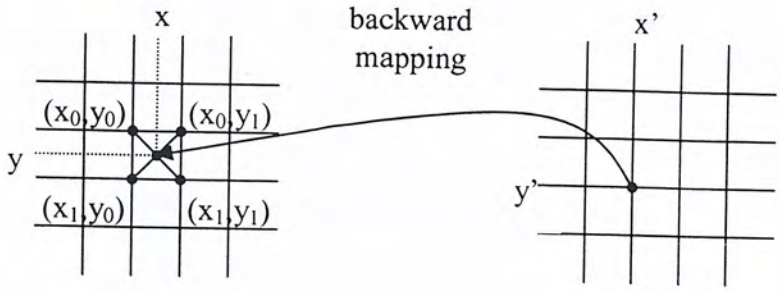
ภาพที่ทำการขยายด้วย bilinear interpolation

รูปที่ 2.5 แสดงภาพที่ทำการขยายโดยวิธี bilinear interpolation

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.3.3 High-order interpolation (cublic interpolation)





f(x , y)
x , y real

g(x , y)
x' , y' integer

$$g(x',y') = \sum_{m=-1}^{2}\sum_{n=-1}^{2} f(x_0+m,y_0+n)R(m-\Delta x)R(\Delta y-n)$$

$$R(k) = \frac{1}{6}[P(k+2)^3-4P(k-1)^3+6P(k)^3]$$

$$P(z) = z \quad \text{เมื่อ} \quad z > 0$$

$$0 \quad \text{เมื่อ} \quad z \leq 0$$

ภาพต้นแบบ

ภาพที่ทำการขยายด้วย cublic interpolation

รูปที่ 2.6 แสดงภาพที่ทำการขยายโดยวิธี cublic interpolation

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
ลักษณะเฉพาะของรูปภาพ (Feature)

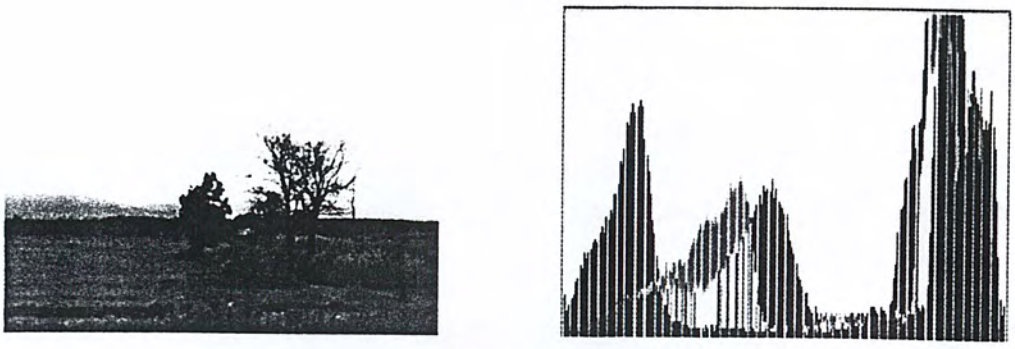
ลักษณะเฉพาะของรูปภาพเป็นการบอกลักษณะเฉพาะของรูปภาพแต่ละรูป ทำให้เรารูปว่าภาพแต่ละภาพแต่ต่างกันในการประมวลผลภาพ มีลักษณะเฉพาะของรูปภาพ อยู่หลายอย่างในการบอกลักษณะของรูปภาพ แต่ในงานวิจัยนี้เราได้เลือกที่จะใช้อยู่ 3 ลักษณะเฉพาะของรูปภาพที่เป็นลักษณะเฉพาะที่ใช้ระบุรูปภาพได้ดี คือ ฮีสโตแกรม , ฟูริเยร์ดิสสคริปเตอร์ , เกเบอร์ทั้ง 3 ลักษณะเฉพาะ นี้จะบอกลักษณะของภาพแตกต่างกันออกไป ฮีสโตแกรมจะบอกลักษณะสีของภาพ , ฟูริเยร์ดิสสคริปเตอร์จะบอกลักษณะลายเส้นของภาพ ส่วนเกเบอร์จะบอกลักษณะพื้นผิวของภาพ

3.1 ฮีสโตแกรม (Histogram)

ฮีสโตแกรมของสีประกอบด้วยระดับต่างๆของสี ซึ่งระดับต่างๆนี้จะมีจำนวนของสีในแต่ละระดับที่ปรากฏในภาพ ฮีสโตแกรมของภาพสามารถอธิบายความมืด ความสว่าง ความคมชัดสูงหรือต่ำ ฮีสโตแกรมของสีจะมีความทนทานต่อการเปลี่ยนมุมมองของภาพเนื่องจากการเปลี่ยนมุมมองจะทำให้จำนวนพิกเซลรวมของสีไม่เกิดการเปลี่ยนแปลงมากนัก แต่ถ้ารูปร่างของภาพเมื่อมีการเปลี่ยนแปลงมุมมองแล้วจะเกิดการเปลี่ยนแปลงไปมาก ฮีสโตแกรมจะเป็นเสมือนฟังก์ชั่นในการกำหนดความน่าจะเป็นของสีที่ปรากฏในภาพ ซึ่งค่าของสีเดียวกันจะต้องอยู่ในระดับของสีเดียวกัน แต่ฮีสโตแกรมของสีอย่างเดียวไม่สามารถที่ในการค้นหารูปภาพได้อย่างมีประสิทธิภาพ เนื่องจากว่าฮีสโตแกรมของสีไม่สามารถบอกได้ว่าระดับสีในแต่ละระดับอยู่ในตำแหน่งใดของภาพ อย่างไรก็ตาม ฮีสโตแกรมกลับได้รับความนิยมในการเป็นดัชนีของภาพ เพราะส่วนใหญ่แล้วพบว่าถ้าสีที่ปรากฏในรูปภาพมีความมัว เนื่องจากปัจจัยแวดล้อมต่างๆ ทำให้กลุ่มของสีที่ปรากฏในภาพเปลี่ยนตำแหน่งไป แต่ผลกระทบที่มีต่อ ฮีสโตแกรมของสีจะมีเพียงเล็กน้อยเท่านั้น ซึ่งทำให้การค้นหาภาพยังคงมีประสิทธิภาพที่ดีได้

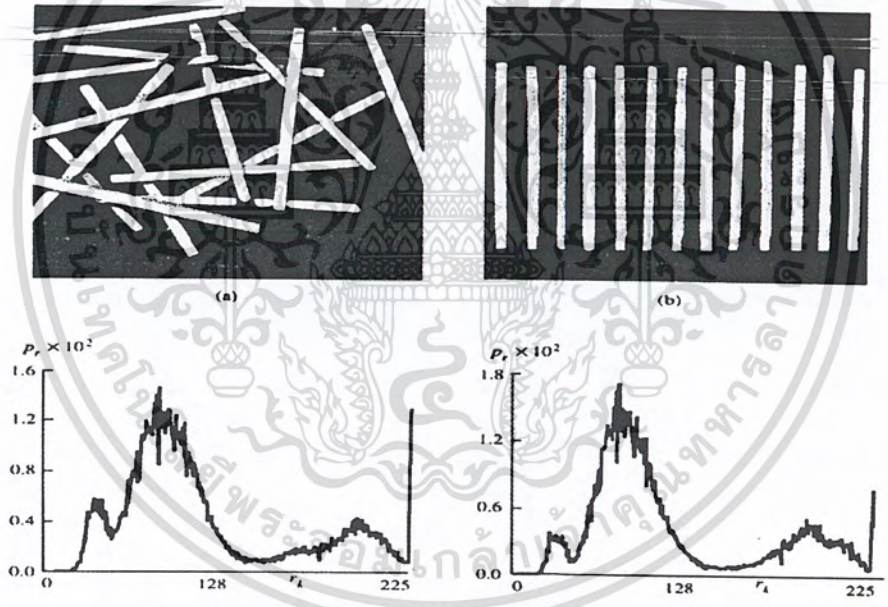
ฮีสโตแกรมจะถูกแสดงด้วยกราฟโดยค่าที่แสดงถึงความเข้มพิกเซลในแต่ละระดับ ซึ่งจะเริ่มตั้งแต่ระดับ 0 ถึง 255 ระดับของพิกเซลจะถูกวางอยู่บนแกน x ส่วนค่าที่แสดงถึงความถี่ที่เกิดขึ้นซ้ำๆกันของสีในระดับนั้นจะถูกแสดงบนแกน y รูปที่ 3.1 จะแสดงภาพที่ทำฮีสโตแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.1 แสดงภาพที่ทำฮีสโตแกรม

สำหรับฮีสโตแกรม (Global) ของภาพ 2 ภาพที่มีรูปร่างกราฟคล้ายๆ กัน ไม่ได้หมายความว่าภาพทั้ง 2 จะคล้ายกัน ทั้งนี้เพราะฮีสโตแกรมของภาพจะไม่เกี่ยวข้องกับตำแหน่งของจุดภาพกล่าวคือ จุดภาพที่มีค่าระดับสีเหมือนกัน จากกราฟฮีสโตแกรมทั้ง 2 ไม่ได้ปรากฏอยู่ที่ตำแหน่งเดียวกันของภาพ ดังนั้นภาพทั้ง 2 จึงไม่เหมือนกัน ดังแสดงในรูปที่ 3.2



รูปที่ 3.2 แสดงภาพที่มีกราฟฮีสโตแกรมเหมือนกันแต่ภาพต่างกัน

3.1.1 การ Quantization

ในงานวิจัยที่ได้จัดทำนี้ การสร้างฮีสโตแกรมของสีเพื่อเก็บสีที่ปรากฏอยู่ในภาพ โดยได้แบ่งค่า RGB ของรูปภาพถูกแบ่งออกเป็น 216 ระดับ (R = 6 ระดับ , G = 6 ระดับ , B = 6 ระดับ) เพื่อลดระดับความซับซ้อนของสีและเป็นการประหยัดเนื้อที่ในการจัดเก็บข้อมูลฮีสโตแกรมของแต่ละภาพ และทำให้ง่ายต่อการทำฮีสโตแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ช่วงของสี	ค่าที่ใช้แทน
0 – 42	0
43 – 85	1
86 – 128	2
129 – 171	3
172 – 213	4
214 - 255	5

ตารางที่ 3.1 ระดับของสีที่ใช้ในการทำ Quantization

เริ่มด้วยเราต้องหาว่าในพิกเซลที่เราสนใจนั้นนั้นประกอบไปด้วยค่าของสีอะไรบ้าง (RGB) เท่าไหร่ แล้วนำค่าที่หาได้ไปเปรียบเทียบกับตารางว่าค่าของพิกเซลที่เราสนใจนั้นอยู่ในช่วงสีไหน แล้วใช้ค่า 0-5 แทนการแสดงด้วยค่าระดับสีที่แท้จริง ในสีแต่ละสีจะแบ่งค่าออกเป็นทั้งหมด 6 ช่วงสี ซึ่งเดิมจะถูกแบ่งไว้ทั้งหมด 255 สี เพราะฉะนั้นเราสามารถแสดงระดับสีจริงๆที่ผสมสีทั้ง 3 แล้วได้ทั้งหมด 216 ระดับ เช่น ถ้า R = 192 , G = 64 , B = 255 ก็จะถูก quantized เป็น R = 4 , G = 1 , B = 5

3.1.2 การ Normalization

เนื่องจากขนาดของรูปภาพจะมีผลทำให้จำนวนพิกเซลในแต่ละรูปภาพไม่เท่ากัน ส่งผลให้ค่าฮีสโตแกรมของรูปภาพ 2 รูปภาพที่มีขนาดไม่เท่ากันนั้นไม่สามารถจะนำมาเปรียบเทียบกันได้ ดังนั้นเราจึงจำเป็นต้องทำการ Normalization เพื่อให้สามารถนำมาเปรียบเทียบกันได้ โดยเมื่อเราหาฮีสโตแกรมของรูปภาพได้แล้ว เราต้องนำความถี่ของแต่ละระดับสีมาหารด้วยจำนวนพิกเซลทั้งหมดที่มีในภาพ ผลที่ออกมาเราจะได้ค่าของความน่าจะเป็น (Probability) ของแต่ละระดับสีที่มีในรูปภาพ ซึ่งเมื่อรวมค่าความน่าจะเป็นของแต่ละระดับสีเข้าด้วยกันแล้วจะมีค่าเป็น 1

โดยสมการ คือ

$$p(n) = \frac{f(n)}{P \times Q}$$

โดย

p(n) - ค่าความน่าจะเป็นของจำนวนพิกเซลที่มีระดับสีที่ n

f(n) - ค่าความถี่ของพิกเซลที่มีในระดับสีที่ n

PxQ - จำนวนพิกเซลทั้งหมดที่มีในรูปภาพได้จากการนำจำนวนพิกเซลตามความกว้างคูณกับจำนวนพิกเซลตามความยาว

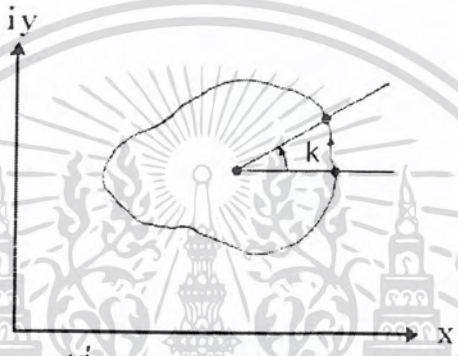
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 ฟูริเยร์ดิสสคริปเตอร์ (Fourier Descriptor)

เป็นการบรรยายลักษณะของเส้น (ที่ได้จากการหาขอบของรูปภาพ) เส้นขอบของสิ่งต่างๆในรูปภาพจะถูกแสดงด้วยจุดจำนวน K จุด บนระนาบ xy โดยเริ่มที่จุด (x_0,y_0) เพราะฉะนั้นจะได้คู่ลำดับของเส้นเป็น $(x_0,y_0),(x_1,y_1),(x_2,y_2),\ldots,(x_{k-1},y_{k-1})$ แต่ในรูปภาพ 1 รูปสามารถมีเส้นมากกว่า 1 เส้นได้ คู่ลำดับนี้สามารถแสดงด้วยรูปแบบ $x(k) = x_k$ และ $y(k) = y_k$ ในแต่ละเส้นขอบสามารถแสดงเป็นลำดับต่อเนื่องได้ คือ $s(k) = [x(k) , y(k)]$ โดย

$k = 0,1,2,\ldots,K-1$ แต่ละคู่ลำดับก็แทนได้ด้วยจำนวนเชิงซ้อน ดังรูปที่ 3.3

$$s(k) = x(k) + jy(k)$$



รูปที่ 3.3 กราฟแสดงคู่ลำดับเชิงซ้อน

โดยที่แกน x จะเป็นแกนของจำนวนจริง และแกน y จะเป็นแกนจินตภาพของจำนวนเชิงซ้อน ถึงแม้ว่าการอธิบายของเส้นจะถูกเปลี่ยนแปลงไปแต่โดยคุณสมบัติของเส้นขอบจะไม่ได้เปลี่ยนแปลงไป การเปลี่ยนแปลงให้เส้นถูกแทนด้วยชุดของจำนวนเชิงซ้อนเพื่อเป็นการลดปัญหาจาก 2 มิติเหลือ 1 มิติเดียว

3.2.1 Descrete Fourier Transform (DFT)

เมื่อเราได้เส้นที่ถูกแทนด้วยคู่ลำดับแล้ว เราต้องนำแปลงเพื่อให้เป็นจำนวนเชิงซ้อน โดยใช้วิธี Descrete Fourier Transform ดังแสดงในสมการ เป็นการแปลงชุดของคู่ลำดับ s(n) เป็นชุดของจำนวนเชิงซ้อน

$$a(k) = \frac{1}{N}\sum_{n=0}^{N-1} S(n)e^{-j2\pi(kn/N)}$$

โดย $k = 0,1,2,\ldots,N-1$ และ a(k) จะถูกเรียกว่า " Fourier descriptors " ของเส้นขอบ ในการแปลงกลับเป็น s(n) รูปแบบเดิมทำได้โดยใช้วิธี Inverse Fourier Transform ดังแสดงในสมการ

$$s(n) = \sum_{n=0}^{N-1} a(k)e^{j2\pi(kn/N)}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

โดย k = 0,1,2,…,N-1 อย่างไรก็ตามเราจะเลือกใช้เพียง P coefficients เท่านั้นจะไม่ใช้ทั้งหมดของ Fourier coefficients มันจะสมดุลโดยให้ a(k) = 0 สำหรับ k > P-1 (เมื่อเกินค่า P ให้กำหนดค่าที่เหลือใน a(u) ให้เป็น 0 ทั้งหมด)

$$\hat{s}(n) = \sum_{k=0}^{P-1} a(k) e^{j2\pi(kn/N)}$$

ถ้าต้องการค่าใกล้เคียงรูปภาพเดิมมากที่สุด จะต้องกำหนดค่า P ให้เท่ากับ N (ในสมการเพื่อจะได้ค่าที่ใกล้เคียง s(n) เดิมมากที่สุด)

โดย k = 0,1,2,…,N-1 ถ้าเรายิ่งกำหนด P ให้มีค่าใกล้เคียง N มากเท่าไหร่ รายละเอียดของภาพจะยิ่งมากขึ้น

ก่อนที่จะนำภาพมาหาฟูริเยร์ดิสสครีปเตอร์เราต้องหาเส้นในภาพให้ได้ก่อน โดยใช้วิธี "การหาขอบรูปภาพ (Edge Detector) " ซึ่งจะทำให้เราได้ขอบของสิ่งต่างๆ ที่อยู่ในภาพ แสดงออกมาในรูปแบบเส้น ซึ่งในภาพอาจจะมีเส้นแค่เส้นเดียวหรือหลายเส้นก็ได้ แล้วเราค่อยนำแต่ละพิกเซลที่ประกอบกันเป็นเส้นไปผ่าน Fourier Transform ผลที่ได้จะเป็นอาเรย์ของเลขทศนิยมที่มีขนาดอาเรย์เท่ากับจำนวนพิกเซลที่เรามี แต่ในงานวิจัยนี้เราไม่ได้ต้องการรายละเอียดเส้นขอบที่สูงมาก เราต้องการแค่บอกลักษณะของเส้นขอบอย่างคร่าวๆ จึงไม่จำเป็นที่จะต้องเอาค่าทั้งหมดในอาเรย์มาใช้ (กำหนดค่า P ให้พอบอกลักษณะของเส้น) สามารถตัดมาแค่บางส่วนได้

3.2.2 Fast Fourier Transform (FFT)

เนื่องจากการทำ Descrete Fourier Transform มีการคูณจำนวนเชิงซ้อนมากเลยทำให้สิ้นเปลืองทรัพยากรในการทำและใช้เวลาในการคำนวณมาก ทำให้การคำนวณยุ่งยากและช้ามาก

ดังนั้นจึงมีการทำฟูริเยร์ดิสสครีปเตอร์อีกวิธีหนึ่ง เรียกว่า " Fast Fourier Transform " เป็นการทำ Descrete Fourier Transform แบบเร็ว เพื่อช่วยลดทรัพยากรที่ใช้และเวลาในการคำนวณ โดยใช้วิธีการสลับตำแหน่งและการรวม (Recomposite) มีวิธีทำโดยการสลับตำแหน่งของพจน์มาไว้ด้วยกัน และรวมพจน์ให้อยู่ในรูปของการทำ Descrete Fourier Transform ที่มีจำนวน N ที่น้อยกว่าทำให้ลดรูปสมการให้คำนวณได้ง่ายขึ้น การทำ Fast Fourier Transform มีหลายแบบ แต่ในงานวิจัยนี้เราเลือกใช้ วิธี " Radix-2 DIF-FFT "

3.2.2.1 Radix-2 DIF-FFT

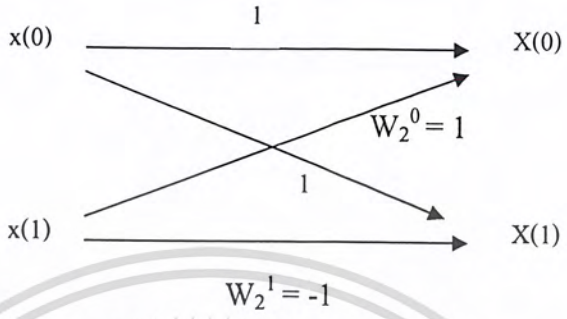
Radix-2 DIF-FFT เป็นการรวมการคำนวณ 2 แบบเข้าด้วยกัน คือ

1. วิธี Divide and conquer เป็นวิธีการลดจำนวนการคูณเลขเชิงซ้อนลง
2. ใช้วิธี Decimation in time กับ N สัญญาโดเมนเวลา โดยที่ N เป็นเลขยกกำลังของ 2 หรือเรียกว่า " Radix-2 " ดังนั้น จึงเรียกวิธีนี้ว่า " Radix-2 DIF-FFT "

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
61462

3.2.2.2 บัตเตอร์ฟลาย (Butterfly)

" บัตเตอร์ฟลาย " เป็นชื่อเรียกของกราฟการไหลของสัญญาณ (Signal flow graph) โดย 1 บัตเตอร์ ฟลาย จะมีการคูณเลขเชิงซ้อน 2 ครั้ง ดังรูปที่ 3.4



รูปที่ 3.4 แสดงบัตเตอร์ฟลายของการคูณเลขเชิงซ้อน

วิธีการดูบัตเตอร์ฟลาย คือ ตัวที่อยู่ปลายลูกศรจะเป็นผลลัพธ์ ส่วนตัวที่อยู่ต้นลูกศรจะนำไปคูณกับตัวที่อยู่ตรงกลางลูกศร แล้วนำแต่ละลูกศรที่ชี้ไปที่ตัวเดียวกันมาบวกกัน เช่น จากรูปที่ 3.3 จะได้

$$X(0) = x(0) + W_2^{\ 0}\, x(1)$$

$$X(1) = x(0) + W_2^{\ 1}\, x(1)$$

<u>ตัวอย่าง การทำ DIF-FFT ที่ N = 4</u>

จาก $X(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk}$ โดย k=0,1,...,N-1

และ $W_N = e^{-j2\pi/N}$

เมื่อ N = 4 → $X(k) = x(0)W_2^{\ 0.k} + x(1)W_4^{\ 1.k} + x(2)W_2^{\ 1.k} + x(3)W_4^{\ 3.k}$

$= x(0)W_2^{\ 0.k} + x(2)W_2^{\ 1.k} + x(1)W_4^{\ 1.k} + x(3)W_4^{\ 3.k}$

$= x(0)W_2^{\ 0.k} + x(2)W_2^{\ 1.k} + W_4^{\ 1.k}[x(1)W_2^{\ 0.k} + x(3)W_2^{\ 1.k}]$

$= X_1(k) + W_4^{\ k} X_2(k)$

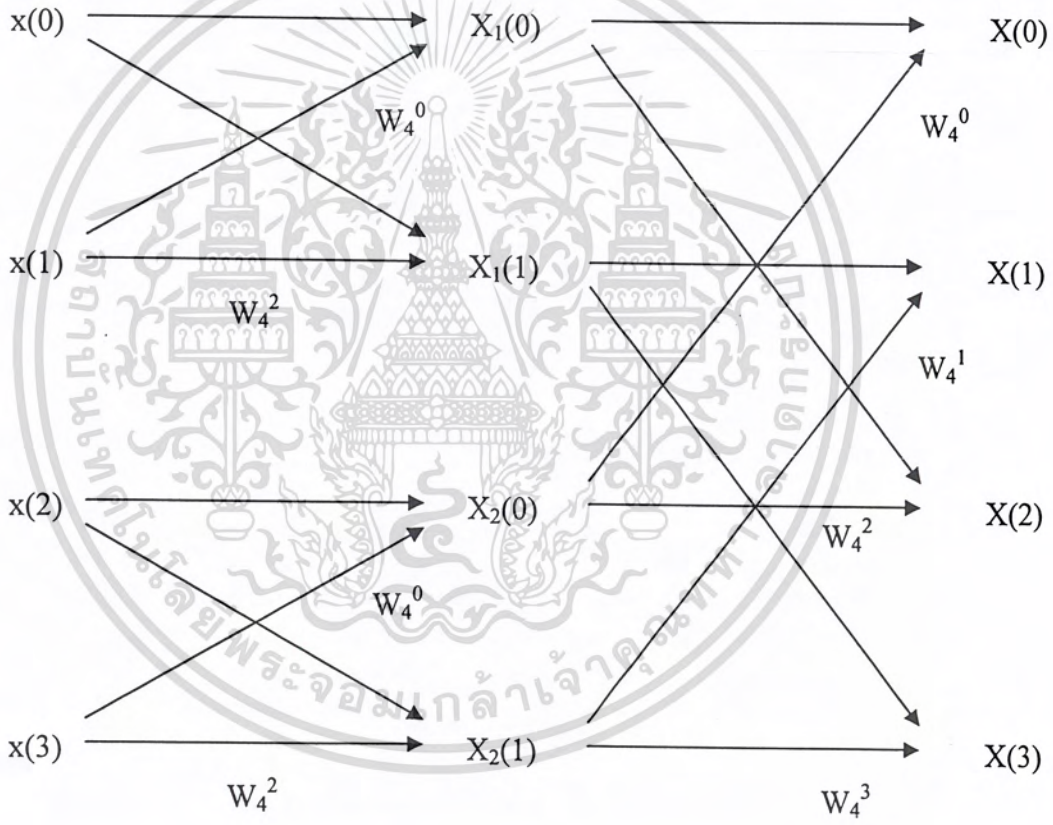
ดังนั้น DFT 4 จุด = DFT 2 จุด + $W_4^{\ k}$ * DFT 2 จุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เราสามารถลดรูปสมการลงได้ เป็น

$$X(0) = X_1(0) + W_4^{0}X_2(0)$$
$$X(1) = X_1(1) + W_4^{1}X_2(1)$$
$$X(2) = X_1(2) + W_4^{2}X_2(2)$$
$$= X_1(0) + W_4^{2}X_2(0)$$
$$X(3) = X_1(3) + W_4^{3}X_2(3)$$
$$= X_1(1) + W_4^{3}X_2(1)$$

ดังนั้นจะได้ บัตเตอร์ฟลาย ของ N = 4 ดังรูปที่ 3.5



หมายเหตุ : เส้นที่ไม่ได้กำหนดค่า ให้มีค่าเป็น 1

รูปที่ 3.5 แสดงการทำ บัตเตอร์ฟลาย N=4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.3 เปรียบเทียบการทำ FFT กับ DFT

การทำ Descrete Fourier Transform จะมีข้อเสีย คือ จะต้องมีการคูณเลขเชิงซ้อนเป็นจำนวนมาก ทำให้เสียเวลาและสิ้นเปลืองทรัพยากรในการคำนวณ ปัจจุบันจึงมีการทำ Fast Fourier Transform เพื่อจะลดจำนวนการคูณเลขเชิงซ้อน ซึ่งจะช่วยให้เราคำนวณได้เร็วขึ้นแล้ว ใช้ทรัพยากรในการคำนวณน้อยลง โดยปกติแล้ว Descrete Fourier Transform จะมีการคูณเลขเชิงซ้อนเป็นจำนวน N^2 ครั้ง แต่การทำ Fast Fourier Transform จะลดจำนวนการคูณเลขเชิงซ้อนลงจาก N^2 เหลือเพียง $N \log_2 N$ ครั้งเท่านั้น จะแสดงให้เห็นดังตารางที่ 3.2

N	DFT N^2	FFT $(N \log_2 N)$
2	4	2
4	16	8
8	64	24
...	...	...
256	65,536	2,048
512	262,144	4,608
1024	1,048,576	10,240

ตารางที่ 3.2 แสดงการเปรียบเทียบจำนวนการคูณเลขเชิงซ้อนของ DFT กับ FFT

3.2.4 การทำภาพเป็นภาพขาวดำ (Gray Scale)

รูปภาพปกติจะเป็นภาพสี RGB ซึ่งยากต่อการคำนวณ และการหาลักษณะเฉพาะของรูปภาพ บางลักษณะเฉพาะของรูปภาพ นั้นต้องทำภาพเป็นภาพขาวดำก่อน การทำภาพสี RGB ให้เป็นภาพขาวดำนั้นมีวิธีการดังนี้ โดยการนำ R,G,B ของพิกเซลแต่ละพิกเซลบนภาพมาคูณด้วยค่าคงที่แต่ละค่าแล้วนำค่าที่ได้ทั้งหมดมาบวกกัน เราจะต้องทำแบบนี้ไปทุกๆพิกเซลบนภาพ แล้วภาพที่ผ่านการทำเป็นภาพขาวดำจะออกมาเป็นภาพขวา-ดำ ดังสมการนี้

$$\text{Gray Scale} = (R*0.299) + (G*0.587) + (B*0.144)$$

<u>ตัวอย่างการทำภาพขาวดำ</u>

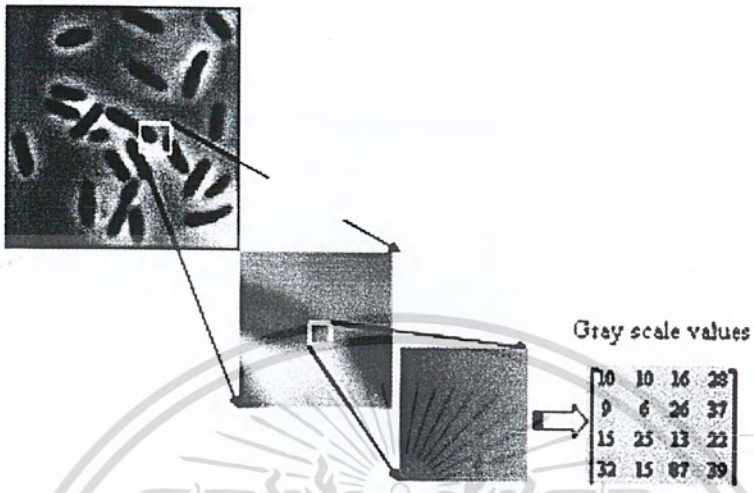
ให้ภาพๆหนึ่งมีพิกเซลที่มีค่า R=15 , G=35 , B=60

$$\text{Gray Scale} = (15*0.299) + (35*0.587) + (60*0.144)$$
$$= 4.485 + 20.545 + 8.64$$
$$= 33.67$$

ดังนั้นพิกเซลนี้มีค่าเท่ากับ 33.67

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าในแต่ละพิกเซลของภาพขาวดำ คือค่าความเข้มของแสง ณ แต่ละตำแหน่งของพิกเซลซึ่งจะอยู่ในรูปของภาพขาวดำ ดังรูปที่ 3.6



รูปที่ 3.6 แสดงค่าพิกเซลของภาพขาวดำ

ตัวอย่างภาพที่ทำเป็นภาพขาวดำ แสดงดังรูปที่ 3.7

รูปที่ 3.7 แสดงภาพสี และภาพขาวดำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.5 การตรวจหาขอบ (Edge detection)

ขอบรูปภาพ คือ ชุดของพิกเซลที่วางต่อๆกันอยู่บนขอบระหว่างพื้นที่สองส่วนของรูปภาพ ขอบของสิ่งของที่อยู่ในรูปภาพจะช่วยอธิบายรูปร่าง ลักษณะ ขนาด และอื่นๆ ที่อยู่ในรูปภาพ การตรวจหาขอบรูปภาพเป็นก้าวแรกในการแยกรูปภาพเป็นส่วน (Image Segment) แล้วนำมาประกอบรวมกันเป็นออบเจ็กต์หรือพื้นที่ใหม่ (Region) หรือสามารถนำขอบของส่วนต่างๆในภาพไปทำสิ่งอื่น ในที่นี้เราจะพิจารณาการตรวจหาขอบรูปภาพด้วยอนุพันธ์อันดับที่หนึ่งอนุพันธ์อันดับหนึ่ง (First Order Derivative) ได้แก่ โอเปอร์เรเตอร์ Sobel , Prewitt และ Frei-Chen

อนุพันธ์อันดับหนึ่ง ถ้านำมาหาขอบอิมเมจในทิศแนวนอน และแนวตั้ง ก็จะหาอนุพันธ์สมการที่ 1 เรา เรียกว่า " การหา Gradient " ซึ่งเป็นเวคเตอร์เมตริก แสดงได้ดังนี้

$$\text{Gradient f} = [H_r(x,y)] = \left[\frac{\partial}{\partial x} f(x,y)\right]$$

$$[H_c(x,y)] = \left[\frac{\partial}{\partial y} f(x,y)\right] \quad \text{.....(1)}$$

โดยที่ $\frac{\partial}{\partial x} f(x,y)$ จะเป็นการหาขอบอิมเมจในทางแนวนอน (x) และ $\frac{\partial}{\partial y} f(x,y)$ จะเป็นการหาขอบอิมเมจ

ในทิศทางแนวตั้ง (y) ถ้าหาทั้งสองทิศทาง จะเป็นขนาดของเวกเตอร์ (Magnitude Vector) ของสมการ (1) ซึ่งเขียนใหม่ได้ดังนี้

$$\text{Gradient f} = \left[H_r^2(x,y) + H_c^2(x,y)\right]^{1/2} \cong |H_r(x,y)| + |H_r(x,y)|$$

ส่วนทิศทางการตรวจหาขอบอิมเมจได้จากสมการ

$$\theta = \tan^{-1}\left[\frac{H_c(x,y)}{H_r(x,y)}\right]$$

โอเปอร์เรเตอร์ Sobel จะใช้ค่า Mask Coefficient ดังนี้

ในทิศทางแนวนอน H r (x , y)

$$\begin{bmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \text{หรือ} \quad \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

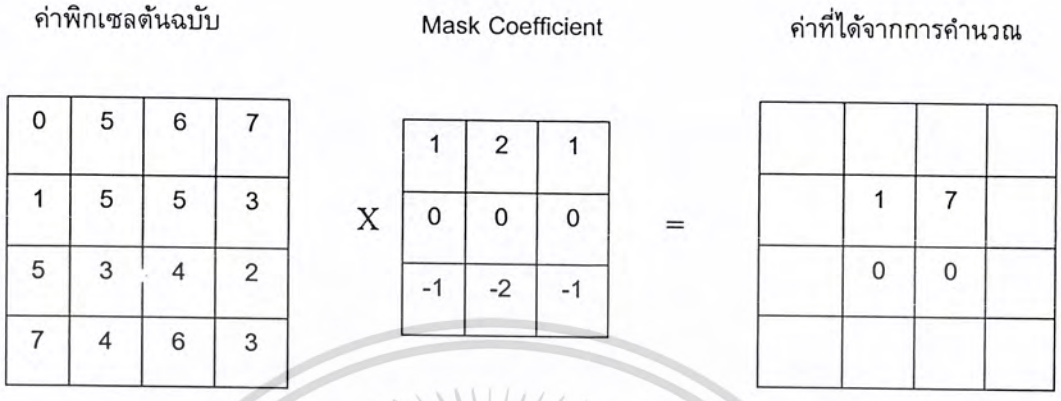
ในทิศทางแนวตั้ง H c (x , y)

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{หรือ} \quad \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างการคำนวณ

คำนวณหาค่าพิกเซลใหม่ เมื่อกำหนดค่า Mask Coefficient แบบ Sobel ในทิศทาง x อิมเมจมีระดับความสว่างหรือความเข้ม ตั้งแต่ 0 ถึง 7 และมีอิมเมจขนาด4 x 4พิกเซล ดังรูปที่ 3.8



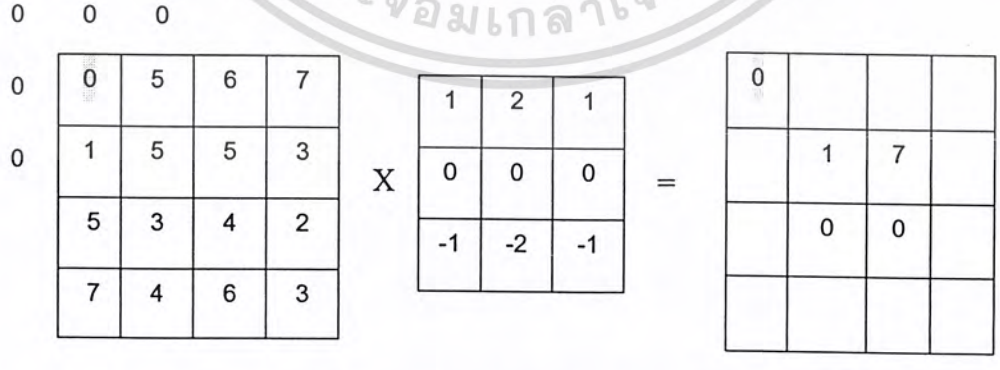
รูปที่ 3.8 การคำนวณหาเส้นขอบโดยใช้ Sobel Mask

เลือกพิกเซลที่ต้องการหาแล้วนำค่าของจุดนั้นและค่าที่อยู่รอบมาทำการคูณกับค่าที่อยู่ใน Mask Coefficient ตำแหน่งต่อตำแหน่ง เช่น จากรูป เราต้องการหาที่ตำแหน่ง (2,2) ซึ่งมีค่าเท่ากับ 5 ก็จะได้

$(0x1) + (5x2) + (6x1) + (1x0) + (5x0) + (5x0) + (5x(-1)) + (3x(-2)) + (4x(-1)) = 1$

แล้วนำค่าที่ได้ไปใส่ในผลลัพธ์ตำแหน่งนั้น ดังรูป จากตัวอย่างจะเห็นว่าอิมเมจมีระดับความสว่างหรือความเข้ม ตั้งแต่ 0 ถึง 7 ดังนั้นถ้าค่าที่คำนวณได้มีค่าน้อยกว่า 0 ก็จะแทนเป็นค่า 0 ส่วนค่าที่คำนวณได้มากกว่า 7 ก็แทนด้วย 7

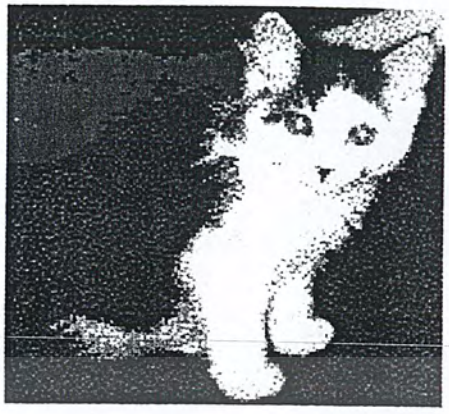
การคำนวณพิกเซลที่อยู่ขอบของหน้าต่างที่เราจะค้นหาในที่นี้เราใช้วิธีใส่ 0 ที่ตำแน่งที่ขาดไปจนมีขนาดเท่ากับ หน้าต่างที่กำหนดในที่นี้คือ 3x3 เพื่อให้สามารถคูณกันได้ ดังรูป 3.9



รูปที่ 3.9 การคำนวณพิกเซลที่อยู่ขอบของภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อผ่านขั้นตอนการตรวจสอบหาขอบแล้วจะได้ภาพดังรูปที่ 3.10



ภาพก่อนการหาขอบ

ภาพหลังการหาขอบ

รูปที่ 3.10 ภาพที่ก่อนนำมาตรวจสอบหาขอบ และหลังนำมาตรวจสอบหาขอบ

3.2.6 การทำคอนโวลูชั่น (Convolution)

การทำการ คอนโวลูชั่นภาพที่มีขนาด $M \times N$ โดยใช้ Kernel และฟังชันข้อมูลภาพ เขียนแทนด้วย $I(x, y)$ สำหรับขั้นตอนในการทำจะนำ $I(x, y)$ และหน้ากากขนาด 3×3 มาทำการซ้อนทับกับแฟ้มข้อมูลของภาพเอกสาร โดยใช้ตำแหน่งของหน้ากากคูณกับตำแหน่งเดียวกันของแฟ้มข้อมูล (คอนโวลูชั่น) โดย หน้ากาก ขนาด 3×3 มีลักษณะดังนี้

$$K(i,j) = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

วิธีการคำนวณดังสมการนี้

$$g(i,j) = \sum_{k=1}^{m}\sum_{l=1}^{n} I(i+k-1, j+l-1)K(k,l)$$

โดยที่ i เริ่มจาก 1 จนถึง $M-m+1$ และ j เริ่มจาก 1 จนถึง $N-n+1$

ในการทำคอนโวลูชั่นเป็นการทำการคำนวณล่วงหน้า(Preprocessing) ของข้อมูลอาเรย์ของรูปภาพต้นแบบ ระดับเท่า และมีแถวลำดับของตัวกรองซึ่งโดยทั่วไปแล้วจะต้องมีอย่างน้อยที่สุด 2 มิติ แถวลำดับ ตัวอย่างเช่น

$$g_{57} = I_{57}K_{11} + I_{58}K_{12} + I_{59}K_{13} + I_{67}K_{21} + I_{68}K_{22} + I_{69}K_{23}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังแสดงไว้ดังภาพ โดยมีภาพต้นแบบและ kernel ดังรูปที่ 3.11

I_{11}	I_{12}	I_{13}	I_{14}	I_{15}	I_{16}	I_{17}	I_{18}	I_{19}
I_{21}	I_{22}	I_{23}	I_{24}	I_{25}	I_{26}	I_{27}	I_{28}	I_{29}
I_{31}	I_{32}	I_{33}	I_{34}	I_{35}	I_{36}	I_{37}	I_{38}	I_{39}
I_{41}	I_{42}	I_{43}	I_{44}	I_{45}	I_{46}	I_{47}	I_{48}	I_{49}
I_{51}	I_{52}	I_{53}	I_{54}	I_{55}	I_{56}	I_{57}	I_{58}	I_{59}
I_{61}	I_{62}	I_{63}	I_{64}	I_{65}	I_{66}	I_{67}	I_{68}	I_{69}

K_{11}	K_{12}	K_{13}
K_{21}	K_{22}	K_{23}

รูปที่ 3.11 ข้อมูลภาพต้นแบบและ kernel ที่ทำคอนโวลูชั่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 เกเบอร์(Gabor)

การหาเกเบอร์ดิสสคริปเตอร์จะเป็น ลักษณะเฉพาะของรูปภาพ ที่ใช้บอกลักษณะพื้นผิวของภาพ โดยจะใช้หน้ากากที่มีขนาดตามที่กำหนดและเปลี่ยนแปลงตามค่า scale(การเปลี่ยนแปลงความถี่) และ oriantation(การเปลี่ยนแปลงมุม) ไปกรองภาพทุกพิกเซลทั่วรูปภาพเพื่อให้ได้ภาพใหม่ แล้วค่อยไปหาค่าเฉลี่ยของขนาด (μ_{mn}) และค่าเบี่ยงเบนมาตรฐานของภาพที่ได้จากการแปลงสัมประสิทธิ์ (σ_{mn}) มีรายละเอียดดังต่อไปนี้

3.3.1 ตัวกรองเกเบอร์ (Gabor Filter)

เกเบอร์เวฟเลท(Gabor Wavelet) ได้นำมาพิสูจน์แล้วว่าเป็นประโยชน์อย่างยิ่งในการวิเคราะห์พื้นผิวของภาพโดยรูปแบบการเรียงตัวของพิกเซล และได้นำมาใช้อย่างแพร่หลายในการประมวลผลภาพในการหาลักษณะเด่นของพื้นผิว(Texture Feature) ทำได้โดยการคำนวณหาค่าเฉลี่ย(μ) และค่าความแปรปรวน(σ)ของภาพที่ผ่านกระบวนการกรองเกเบอร์แล้ว โดยที่ทำการนอร์มอลไลซ์การหมุน (Rotation Normalization) ที่มีลักษณะการเคลื่อนตัวเป็นวงกลมของแต่ละอิลิเมนท์ขององค์ประกอบ (Feature Element) ทั่วทั้งภาพและจะมีทิศทางการเคลื่อนตัวที่เป็นลักษณะเด่นออกมาเหมือนๆกัน

ถ้ากำหนดให้ตำแหน่งของพิกเซลใดๆในภาพคือ I(x,y) และขนาดของภาพนั้นคือ PxQ โดยจะมีความสัมพันธ์ดังสมการคอนโวลูชั่นดังนี้

$$G_{mn}(x,y)=\sum_{s}\sum_{t} I(x-s,y-t)\Psi^{*}_{mn}(s,t)$$

จากสมการคอนโวลูชั่นนี้คือ รูปแบบการแปลงดิสครีตเกเบอร์เวฟเลท

โดยที่ s และ t คือ ตัวแปรขนาดของหน้ากาก ตัวกรอง (Filter Mask Size)

Ψ^{*}_{mn} คือ คอนจูเกทเชิงซ้อนของ Ψ_{mn}

Ψ_{mn} คือ ระดับของฟังชั่นความคล้ายกัน (Class of Self-similar Function) โดยได้จากการแผ่ขยายตัวและการหมุนตัวของเวฟเลทหลัก (Mother Wavelet) ดังนี้

$$\Psi(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp\left[--\frac{1}{2}\left(\frac{x^2}{\sigma_x{}^2}+\frac{y^2}{\sigma_y{}^2}\right)\right].\exp(j2\pi Wx)$$

โดย W คือ ความถี่ของการมอดูเลท

และจะได้ฟังชั่นความคล้ายกันของเกเบอร์เวฟเลท (Self-similar Gabor Wavelet) ดังนี้

$$\Psi_{mn}(x,y)=a^{-m}\Psi(\tilde{x},\tilde{y})$$

เมื่อ m และ n เจาะจงด้วยสเกล(Scale)และการปรับตัว(Orientation) ของเวฟเลทตามลำดับโดยที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

m = 0,1,2,3,4,…,M-1

n = 0,1,,2,3,4,…,N-1

$$\tilde{x} = a^{-m}\left(x\cos\theta + y\sin\theta\right)$$

$$\tilde{y} = a^{-m}\left(-x\cos\theta + y\sin\theta\right)$$

โดยที่ a > 1 และ $\theta = n\pi / N$

ตัวแปรในสมการข้างต้นนี้มีค่าดังต่อไปนี้

$$a = \left(u_h / u_l\right)^{\frac{1}{M-1}}$$

$$W_{m,n} = a^m U_l$$

$$\sigma_{x,m,n} = \frac{(a+1)\sqrt{2\ln 2}}{2\pi a^m (a-1)U_l}$$

$$\sigma_{y,m,n} = \frac{1}{2\pi \tan\left(\frac{\pi}{2N}\right)\sqrt{\frac{U_h^{\,2}}{2\ln 2} - \left(\frac{1}{2\pi\sigma_{x,m,n}}\right)^2}}$$

ในงานวิจัยของเราจะใช้ค่าคงที่ ซึ่งเป็นค่าคงที่ ที่ใช้ทั่วๆไปดังต่อไปนี้

$U_l = 0.05$, $U_h = 0.4$

s และ t มีช่วงตั้งแต่ 0 ถึง 60

หน้ากากตัวกรองขนาด 60 x 60

ซึ่งในงานวิจัยนี้เราจะเลือกค่า Scale เป็น 2 และ Orientation เป็น 2 ซึ่งหมายความว่าในรูปภาพที่เราจะหาเกเบอร์ดิสสคริปเตอร์นั้นเราจะสนใจที่ความถี่ 2 ค่าคือ ต่ำและสูง และสนใจลายเส้นที่มีมุม 2 มุม คือ แนวนอนและแนวตั้ง เพราะฉะนั้น เราจะได้กรณีทั้งหมดที่ทำ 4 กรณี คือ ที่ความถี่ต่ำสแกนหาเส้นในแนวนอน , ที่ความถี่ต่ำสแกนหาเส้นในแนวตั้ง , ที่ความถี่สูงสแกนหาเส้นในแนวนอน และที่ความถี่สูงสแกนหาเส้นในแนวตั้ง

3.3.2 การแสดงออกของพื้นผิว (Texture Representation)

ในส่วนนี้จะเป็นเรื่องของการแสดงออกพื้นผิวโดยใช้การแปลงเกเบอร์ (Gabor Transform) โดยกระบวนการคำนวณความคล้ายคลึงของพื้นผิว (Texture Similarity)และการนอร์มอลไลซ์การหมุน

หลังจากที่เราใช้ตัวกรองเกเบอร์บนรูปภาพแล้ว ผลลัพธ์ที่ได้คือความแตกต่างของการปรับตัวที่แตกต่างกันที่สเกล โดยเราจะได้อาเรย์ของขนาด (Array of Magnitudes) ดังสมการนี้

$$E_{(m,n)} = \sum_x \sum_y \left|G_{mn}(x,y)\right|$$

เมื่อ m = 0,1,…,M-1 และ n = 0,1,…,N-1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าของขนาด (Magniture) เหล่านี้ประกอบไปด้วยความแตกต่างของสเกลและการปรับตัวของรูปภาพ จุดประสงค์หลักของการกอบกู้พื้นผิวภาพพื้นฐาน (Texture-based Retrieval) ก็เพื่อที่จะค้นหารูปภาพที่มีองค์ประกอบคล้ายกัน หรือค้นหาบริเวณที่มีความคล้ายคลึงกันของพื้นผิว ซึ่งมันเป็นการคาดคะเนภาพที่เรากำลังสนใจ หรือบริเวณที่เรากำลังพิจารณาที่มีพื้นผิวเป็นเนื้อเดียวกัน (Homogenous Texture) โดยใช้ค่าเฉลี่ยของขนาด (μ_{mn}) และค่าเบี่ยงเบนมาตรฐานของขนาดที่ได้จากการแปลงสัมประสิทธิ์ (σ_{mn}) แสดงลักษณะของบริเวณที่พื้นผิวเป็นเนื้อเดียวกัน

$$\mu_{mn} = \frac{E(m,n)}{P \times Q}$$

$$\sigma_{mn} = \frac{\sqrt{\sum_x \sum_y \left(|G_{mn}(x,y)| - \mu_{mn}\right)^2}}{P \times Q}$$

ในองค์ประกอบของเวกเตอร์ f นี้แสดงถึงการแสดงออกของพื้นผิว (Texture Represition)ได้โดยสร้างมาจากค่าเฉลี่ยของขนาด μ_{mn} และค่าเบี่ยงเบียนมาตรฐานของขนาดที่ได้จากการเปลี่ยนแปลงสัมประสิทธิ์ σ_{mn} ให้เป็นส่วนประกอบของลักษณะเด่น ยกตัวอย่างเช่น มีขนาดสเกลคือ 5 และ 6 ค่าการปรับตัว เพราะฉะนั้นเราจะได้องค์ประกองของเวกเตอร์ f ดังนี้

$$f = (\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \ldots, \mu_{45}, \sigma_{45})$$

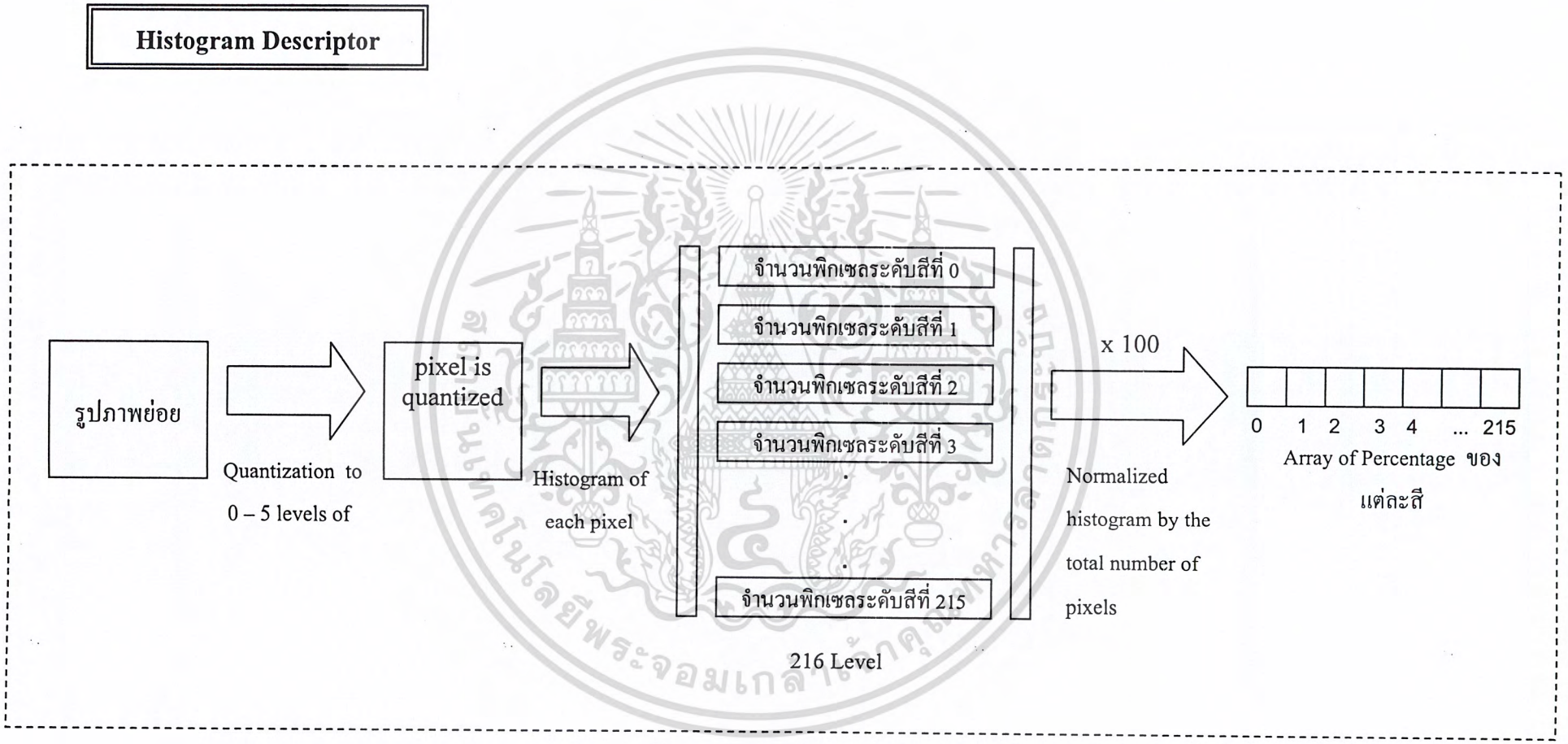
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่4
การทำงานของระบบ

4.1 ขั้นตอนการทำฮีสโตแกรม

หลังจากที่เราได้รูปภาพย่อยแล้ว ก็นำแต่ละรูปภาพย่อยไปทำการ Quantization ซึ่งเป็นการลดระดับของสีลงเพื่อให้ง่ายต่อการคำนวณ โดยในงานวิจัยนี้จะแบ่งเป็น 6 ระดับ ของแต่ละสี คือ R 6 ระดับ G 6 ระดับ และ B 6 ระดับรูปภาพย่อยที่ถูกทำ Quantization แล้ว ก็นำมาหาฮีสโตแกรมโดยนำแต่ละพิกเซลใน รูปภาพย่อย มาดูว่าอยู่ในระดับสีที่เท่าไหร่ แล้วเก็บเป็นความถี่สะสมในแต่ละระดับสี เมื่อสแกนดูทุกพิกเซลในรูปภาพย่อย เสร็จแล้วก็จะนำความถี่สะสมของแต่ละระดับสีที่มาทำการ Normalized โดยการนำความถี่สะสมของแต่ละระดับสีมาหารด้วยจำนวนพิกเซลทั้งหมด แล้วจะได้ผลลัพธ์ออกมาเป็นความน่าจะเป็นของแต่ละระดับสี นำค่าความน่าจะเป็นของแต่ละละดับสีคูณกับ 100 เพื่อให้ได้เป็นค่าเปอร์เซนต์ (%) หลังจากนั้นก็ปัดให้เป็นเลขจำนวนเต็มไปเก็บไว้ในอาเรย์ รูปที่ 4.1 แสดงขั้นตอนการทำฮีสโตแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

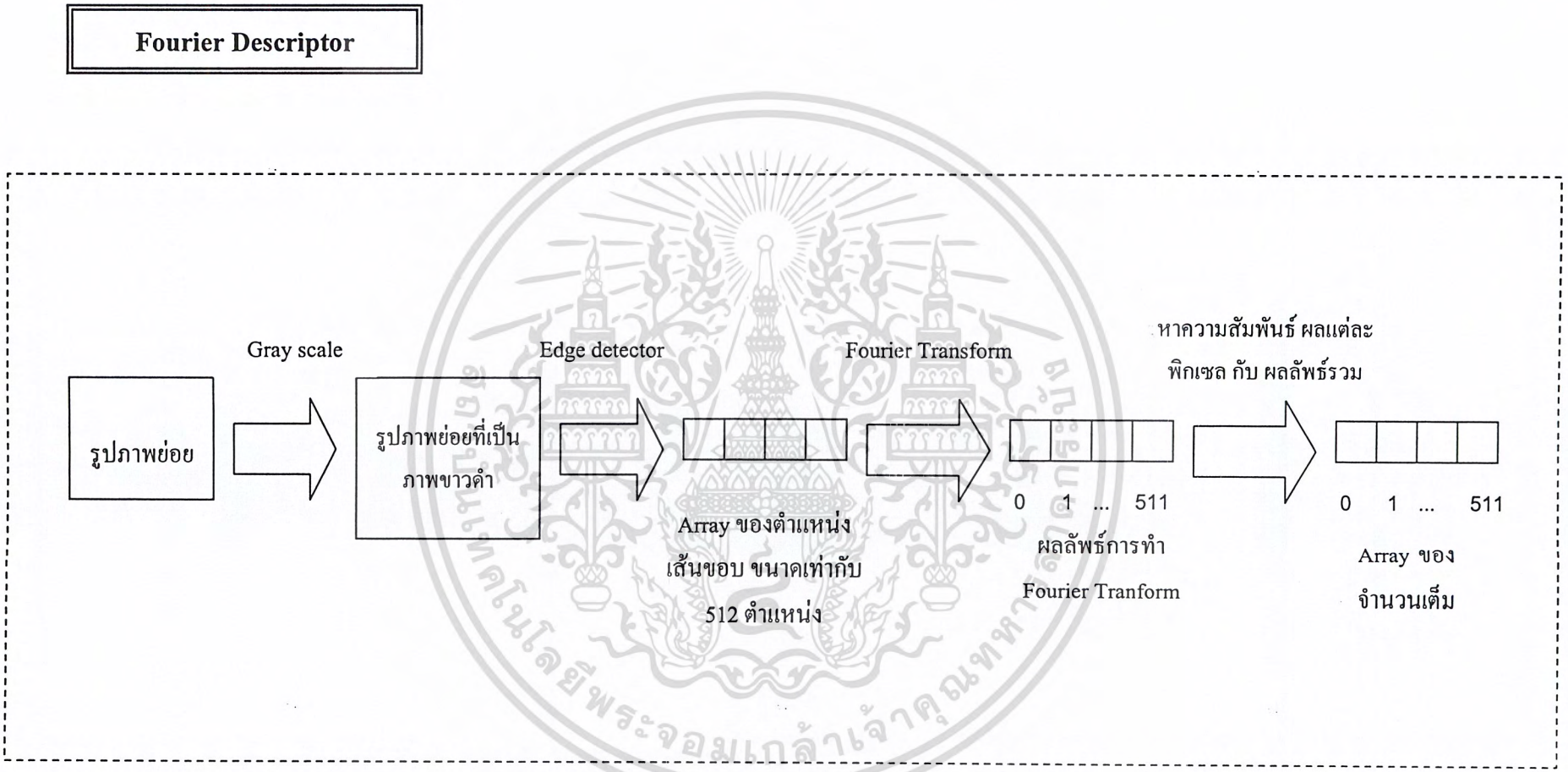


รูปที่ 4.1 แสดงขั้นตอนการทำฮีสโตแกรม

4.2 ขั้นตอนการหาฟูริเยร์ดิสสครีปเตอร์

หลังจากที่เราได้ รูปภาพย่อย แล้ว นำแต่ละ รูปภาพย่อย มาผ่านขั้นตอนการปรับรูปภาพที่มีส่วนประกอบของสี RGB มาทำเป็นรูปภาพขาวดำซึ่งก็ยังประกอบด้วยสี RGB เหมือนเดิม นำ รูปภาพย่อยที่เป็นรูปภาพขาว-ดำไปทำการตรวจสอบหาขอบ เพื่อเป็นการหาเส้นขอบ โดยการนำ sobel mask ไปคำนวณกับแต่ละพิกเซลในภาพ ถ้าพิกเซลใดให้ค่ามากกว่าค่าที่กำหนดไว้ให้เป็นขอบสีขาว นอกเหนือจากนั้นให้เป็นสีดำ จะได้ รูปภาพย่อย ที่เห็นเส้นขอบภายในของรูปภาพออกมา หลังจากนั้นทำการเลือกขอบของรูปภาพตั้งแต่พิกเซลแรกจนถึงพิกเซลที่ 512 ออกมาเพื่อไปทำ Fourier Transform โดยการนำสมการ Fourier ไปคำนวณกับตำแหน่งของพิกเซลที่เป็นขอบที่เลือกมา ผลที่ได้จะเป็นเลขทศนิยม นำผลที่ได้ทั้งหมด 512 ตัวมาหาว่าตัวเองนั้นเป็นเท่าไหร่ของผลรวมของทั้ง 512 ตัว แล้วนำทั้งหมดไปเก็บไว้ในอาเรย์ขั้นตอนการหา Fourier Transform ดังรูป 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.2 แสดงขั้นตอนการทำ Fourier Transform

4.3 การทำเกเบอร์โดยใช้ Fast Fourier Transform (FFT)

การทำเกเบอร์ปกติจะใช้เวลาในการคำนวณนานมาก เนื่องจากเราต้องทำคอนโวลูชั่น ของแต่ละหน้าการกับภาพทั้งภาพทุกพิกเซล ซึ่งเวลาในการทำจะเพิ่มขึ้นแบบเอกโปเนเชี่ยน(Exponential) ดังนั้นเราจึงเลือกจะใช้วิธีการทำเกเบอร์อีกวิธีหนึ่ง โดยการใช้วิธีการทำ Fast Fourier Transform (FFT) เข้ามาช่วยในการคำนวณ เพื่อลดการคำนวณเลขเชิงซ้อน ทำให้ใช้เวลาในการคำนวณน้อยลง และเวลาจะไม่เพิ่มขึ้นเป็นเอกโปเนเชี่ยน

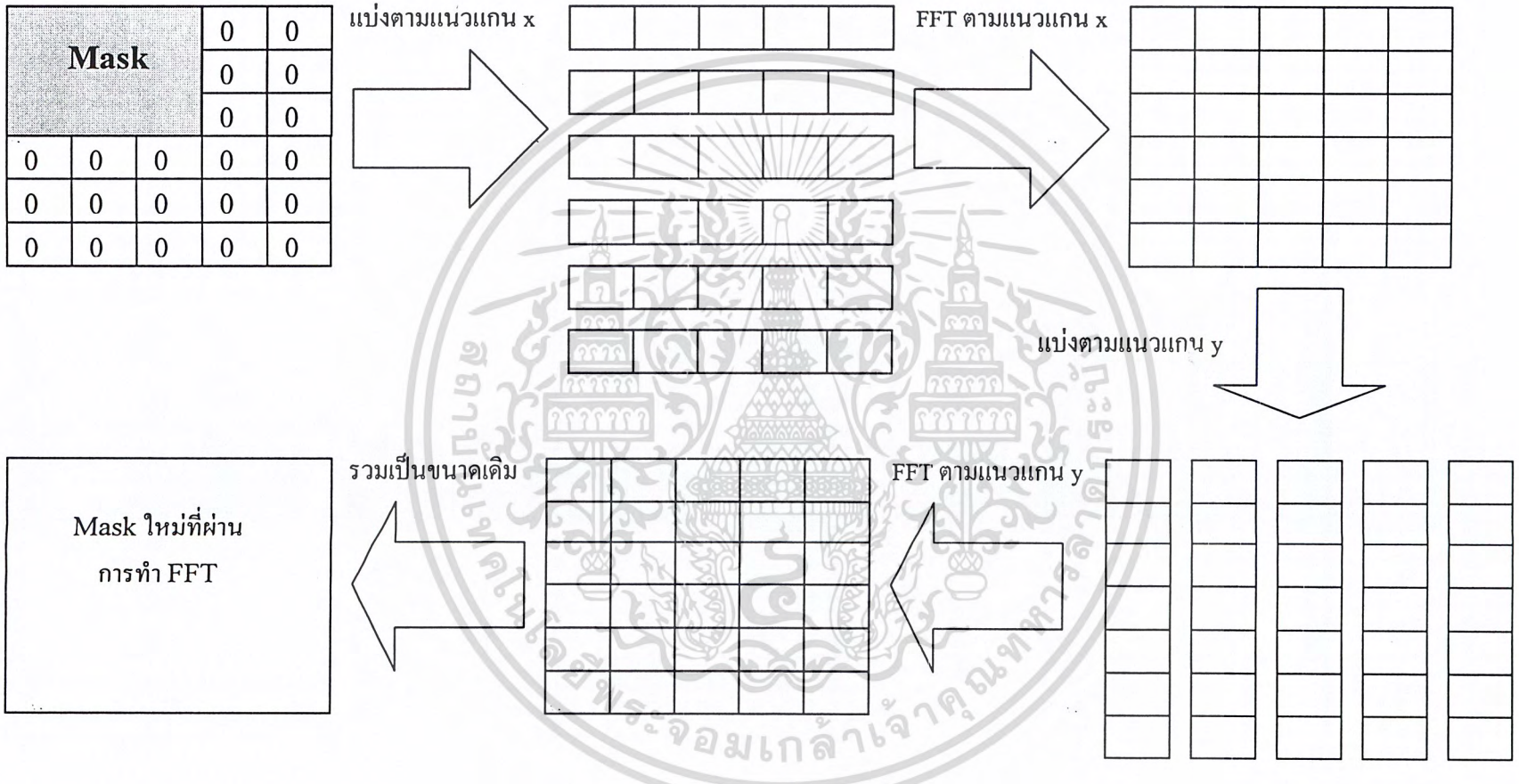
เริ่มต้นด้วยการนำ รูปภาพย่อย ที่ได้มาแปลงภาพให้เป็นขาว-ดำ เก็บไว้ก่อน แล้วก็หา mask มา 1 mask จากการใช้สมการคำนวณหา โดยหน้ากากจะมีขนาด 60 x 60 พิกเซล แล้วทำการขยายขนาดของหน้ากากคำนวณได้มาให้มีขนาดความกว้างและความยาวเท่ากับขนาดของรูปภาพย่อย โดยการเติม 0 เข้าไปในพิกเซลที่ขยายออกมา ดังรูปที่ 4.3

Mask			0	0
			0	0
			0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

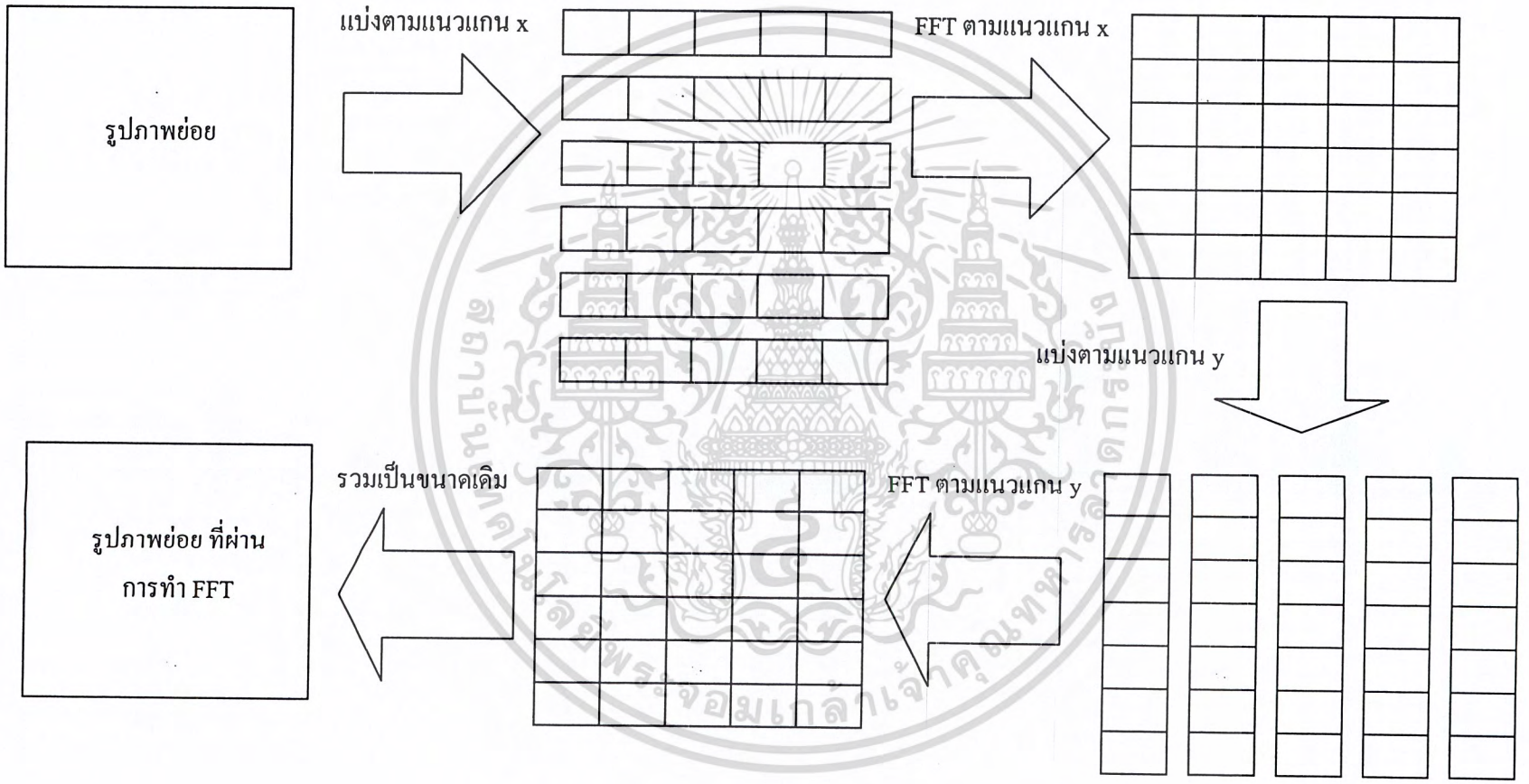
รูปที่ 4.3 การขยายหน้ากากให้มีขนาดเท่ากับขนาดของรูปภาพย่อย โดยการเติม 0

ขั้นตอนต่อมาเราจะนำหน้ากากที่ผ่านการขยายแล้วมาทำการคำนวณโดยใช้ FFT สแกนตามแนวแกน x และ y แล้วก็นำ รูปภาพย่อย มาทำแบบเดียวกัน คำนวณโดยใช้ FFT สแกนตามแนวแกน x และ y ดังรูปที่ 4.4 และ 4.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

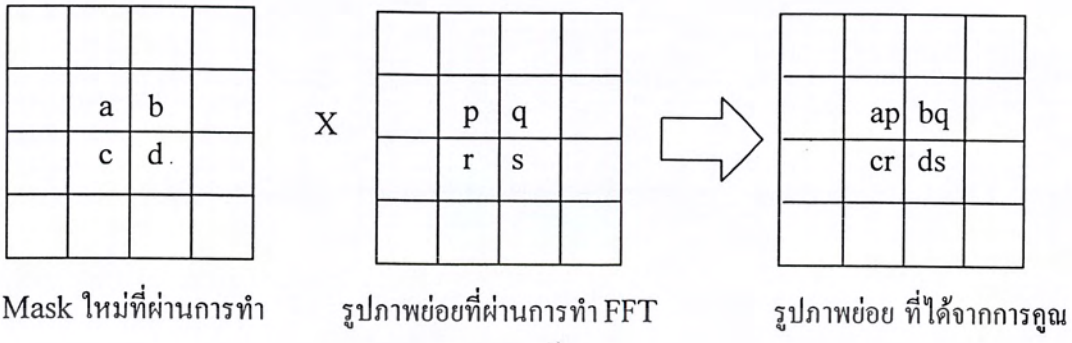


รูปที่ 4.4 แสดงการนำหน้ากากมาคำนวณ FFT ตามแนวแกน x และ y



รูปที่ 4.5 แสดงการนำรูปภาพย่อยมาคำนวณ FFT ตามแนวแกน x และ y

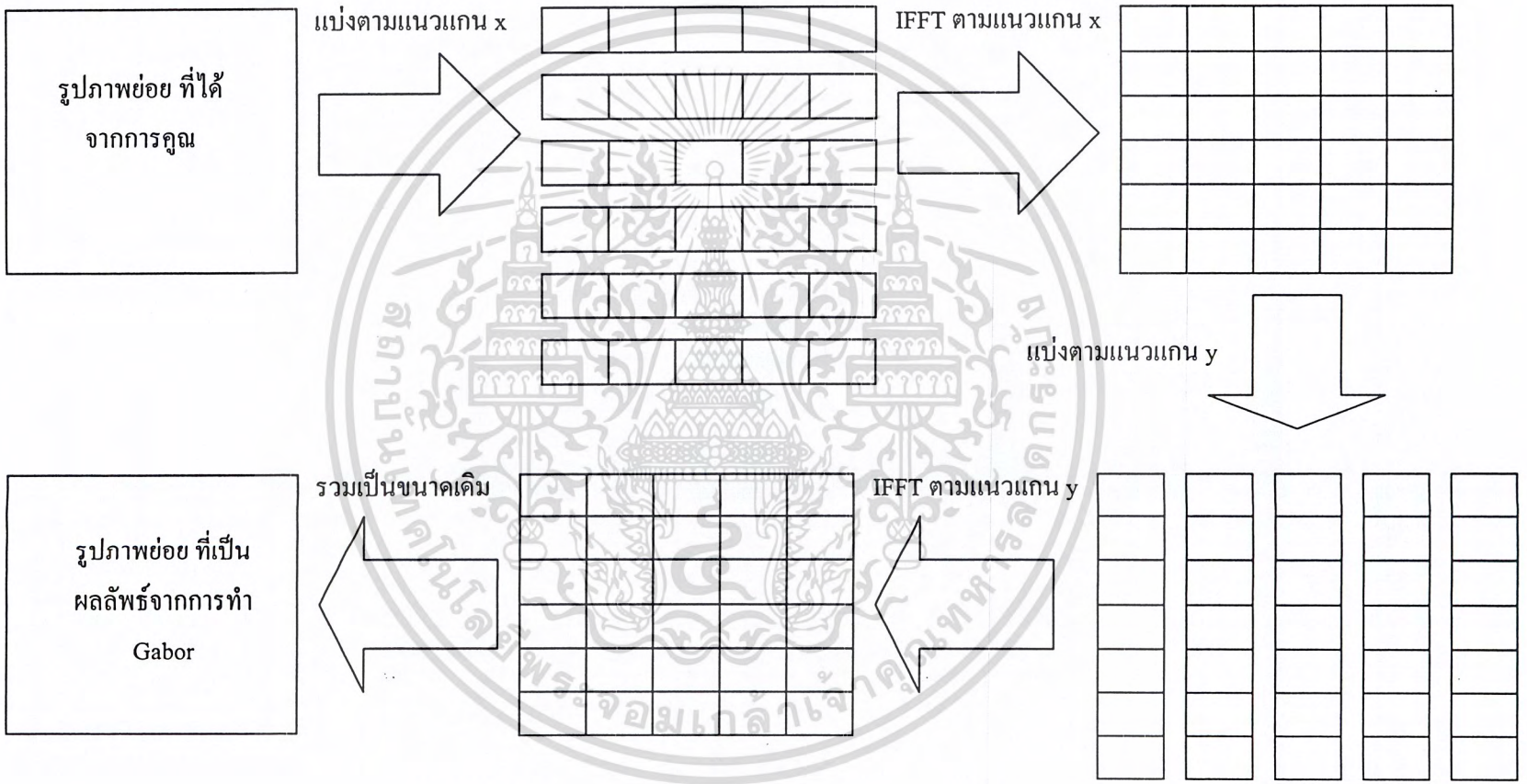
ขั้นตอนต่อไปในนำหน้ากากใหม่ที่ได้จากการทำ FFT และ รูปภาพย่อย ที่ได้จากการทำ FFT นำสองตัวมาคูณกันโดยจะนำค่าแต่ละพิกเซลมาคูณกันตำแหน่งต่อตำแหน่ง ดังรูปที่ 4.6



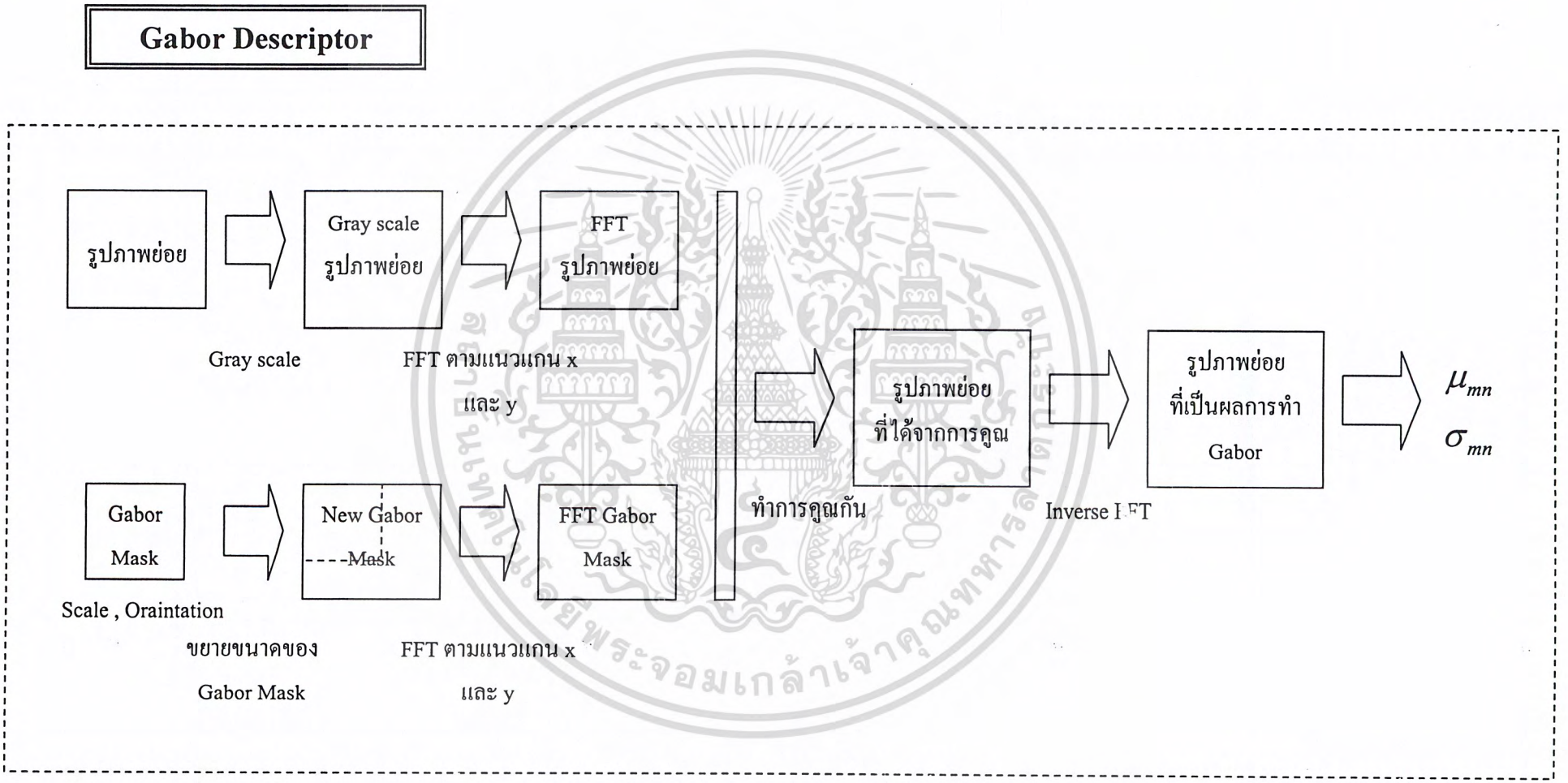
รูปที่ 4.6 แสดงการคูณระหว่างหน้ากากใหม่กับรูปภาพย่อยใหม่

หลังจากนั้นจะนำผลลัพธ์ที่ได้จากการคูณหน้ากากกับรูปภาพย่อย มาทำ Inverse FFT เราจะได้ภาพเกเบอร์ออกมาแล้วนำไปหาลักษณะเฉพาะของภาพนั้นซึ่งเป็น ค่าเฉลี่ยของขนาด (μ_{mn}) และค่าเบี่ยงเบนมาตรฐานของภาพที่ได้จากการแปลงสัมประสิทธิ์ (σ_{mn}) เก็บในอาเรย์แสดงดังรูป 4.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.7 แสดงการนำรูปภาพย่อยที่ได้จากการคูณมาทำ Inverse FFT



รูปที่ 4.8 แสดงขั้นตอนการทำเกเบอร์โดยใช้วิธี FFT

บทที่ 5
การเก็บข้อมูลและการค้นหา

5.1 โครงสร้างฐานข้อมูล

แบ่งเป็น 2 ตาราง(Table) คือ ImageFeature และ Identifier ดังนี้

5.1.1 Identifier

ImgNo	ImgName	KeyValue
1	animal1.bmp	1
2	animal2.bmp	2
3	animal3.bmp	3
4	animal4.bmp	4
5	view.1bmp	5
6	flower1.bmp	6
7	view2.bmp	7
8	flower2.bmp	8
9	flower3.bmp	9
10	texture1.bmp	10
...		...

ตารางที่ 5.1 ตัวอย่างตารางฐานข้อมูล Identifier ที่เก็บชื่อและ Key ของแต่ละภาพ

โดย

ImgNo - เป็นฟิล์ดตัวเลขที่เป็น auto increat เพื่อกำหนดหมายเลขของภาพ เช่น 1 ,2 ,3 ,... จะมีขนาดเท่ากับจำนานภาพที่มีอยู่ทั้งหมดในฐานข้อมูล

ImgName - เป็นฟิล์ดชื่อของรูปภาพ บอกว่าภาพนี้ชื่ออะไร ซึ่งจะมีขนาดเท่ากับจำนวนภาพที่มีอยู่ในฐานข้อมูล

KeyValue - เป็นฟิล์ดตัวเลขที่ใช้เก็บค่า Key ที่ได้จากรูปภาพเป็นรูปแบบการเปลี่ยนแปลงของสีในรูปภาพ ดังรูปที่ 4.2 ซึ่ง จะมีค่าตั้งแต่ 1 ถึง 13

* ใช้ ImgNo และ ImgName เป็น Primary Key ร่วมกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.1.2 ImageFeature

ImgNo	SubImg	Histogram	Fourier	Gabor
1	0	$x_0, x_1, x_2, \dots, x_{215}$	$y_0, y_1, y_2, \dots, y_{511}$	$\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{11}, \sigma_{11}$
1	1	$x_0, x_1, x_2, \dots, x_{215}$	$y_0, y_1, y_2, \dots, y_{511}$	$\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{11}, \sigma_{11}$
1	2	$x_0, x_1, x_2, \dots, x_{215}$	$y_0, y_1, y_2, \dots, y_{511}$	$\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{11}, \sigma_{11}$
1	3	$x_0, x_1, x_2, \dots, x_{215}$	$y_0, y_1, y_2, \dots, y_{511}$	$\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{11}, \sigma_{11}$
1	4	$x_0, x_1, x_2, \dots, x_{215}$	$y_0, y_1, y_2, \dots, y_{511}$	$\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{11}, \sigma_{11}$
1	5	$x_0, x_1, x_2, \dots, x_{215}$	$y_0, y_1, y_2, \dots, y_{511}$	$\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{11}, \sigma_{11}$
1	6	$x_0, x_1, x_2, \dots, x_{215}$	$y_0, y_1, y_2, \dots, y_{511}$	$\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{11}, \sigma_{11}$
1	7	$x_0, x_1, x_2, \dots, x_{215}$	$y_0, y_1, y_2, \dots, y_{511}$	$\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{11}, \sigma_{11}$
1	8	$x_0, x_1, x_2, \dots, x_{215}$	$y_0, y_1, y_2, \dots, y_{511}$	$\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{11}, \sigma_{11}$
2	1	$x_0, x_1, x_2, \dots, x_{215}$	$y_0, y_1, y_2, \dots, y_{511}$	$\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{11}, \sigma_{11}$
…	…	…	…	…
N	8	$x_0, x_1, x_2, \dots, x_{215}$	$y_0, y_1, y_2, \dots, y_{511}$	$\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{11}, \sigma_{11}$

ตารางที่ 5.2 แสดงตัวอย่างตารางฐานข้อมูล ImageFeature ที่เก็บลักษณะเฉพาะของรูปภาพต่างๆ

โดย

Histogram - $x_0, x_1, x_2, \dots, x_{215}$
ขนาดอาเรย์เท่ากับ 216 เป็นการเก็บลักษณะสีที่มี 216 ระดับ

Fourier - $y_0, y_1, y_2, \dots, y_{511}$
ขนาดอาเรย์เท่ากับ 512 เป็นการเก็บเส้นรูปภาพย่อยละ 512 พิกเซล

Gabor - $\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{11}, \sigma_{11}$
ขนาดอาเรย์เท่ากับ 60 เป็นการเก็บลักษณะสีตามแนวตั้งและนอนในความถี่ต่ำและสูง

* ใช้ ImgNo. กับ SubImg เป็น Primary Key ร่วมกัน

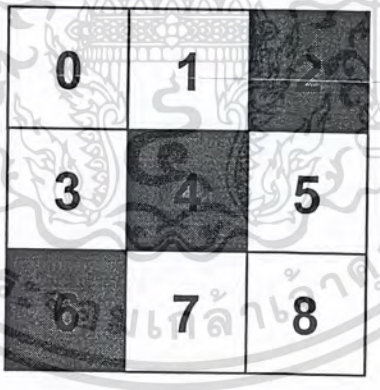
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 การแยกกลุ่มรูปภาพด้วยลักษณะของสี

ถ้าเราไม่มีการใช้เกณฑ์เพื่อแบ่งแยกประเภทรูปภาพที่มีลักษณะคล้ายคลึงกันไว้ด้วยกัน จะทำให้ขั้นตอนในการดึงข้อมูลเพื่อนำมาเปรียบเทียบนั้น เราจะต้องดึงข้อมูลของรูปภาพทุกภาพที่มีอยู่ในฐานข้อมูลซึ่งอาจจะเป็นทั้งภาพที่คล้ายกันหรือไม่คล้ายคลึงกันเลย แล้วจึงค่อยนำรูปภาพมาเปรียบเทียบทีละรูปภาพว่าตรงกับความต้องการหรือไม่ ซึ่งเป็นการเสียเวลาในการพิจารณารูปภาพที่มีความแตกต่างกันอย่างชัดเจน ตามความเป็นจริงแล้วรูปภาพบางรูปที่ไม่มีส่วนคล้ายกับรูปภาพที่ต้องเลยก็ไม่น่าที่จะต้อง ถูกดึงขึ้นมาเปรียบเทียบด้วย ดังนั้นถ้าเรามีการกำหนดเกณฑ์สำหรับการแบ่งรูปภาพ ให้แบ่งแยกเป็นกลุ่มที่คล้ายกันอยู่ด้วยกัน จะทำให้ขั้นตอนในที่เราจะเลือกรูปภาพเพื่อมาเปรียบเทียบนั้นเราสามารถจะได้เฉพาะรูปภาพบางรูปที่มีลักษณะคล้ายกันมาพิจารณาเท่านั้นเพื่อเป็นการลดเวลาในการค้นหารูปภาพ

เนื่องจากว่ารูปภาพแต่ละรูปจะมีลักษณะของสีไม่สม่ำเสมอกันทั้งรูปภาพ ดังนั้นในงานวิจัยนี้เราจะแบ่งกลุ่มรูปภาพแต่ละรูปด้วยลักษณะของสีที่เปลี่ยนแปลงภายในรูปภาพ โดยเราสามารถจะพิจารณาสีโดยใช้วิธีการดูค่าสีเฉลี่ย(μ)ในแต่ละบริเวณภายในรูปภาพ เพราะสามารถคำนวณหาได้ง่ายไม่ยุ่งยากและสามารถบอกรายละเอียดของแต่ละรูปภาพได้ชัดเจนระดับหนึ่ง หลังจากนั้นจึงวิเคราะห์ค่าของสีเฉลี่ยที่เปลี่ยนแปลง แล้วจัดกลุ่มรูปภาพที่มีลักษณะการเปลี่ยนแปลงของสีเฉลี่ยรูปแบบเดียวกันไว้ด้วยกัน

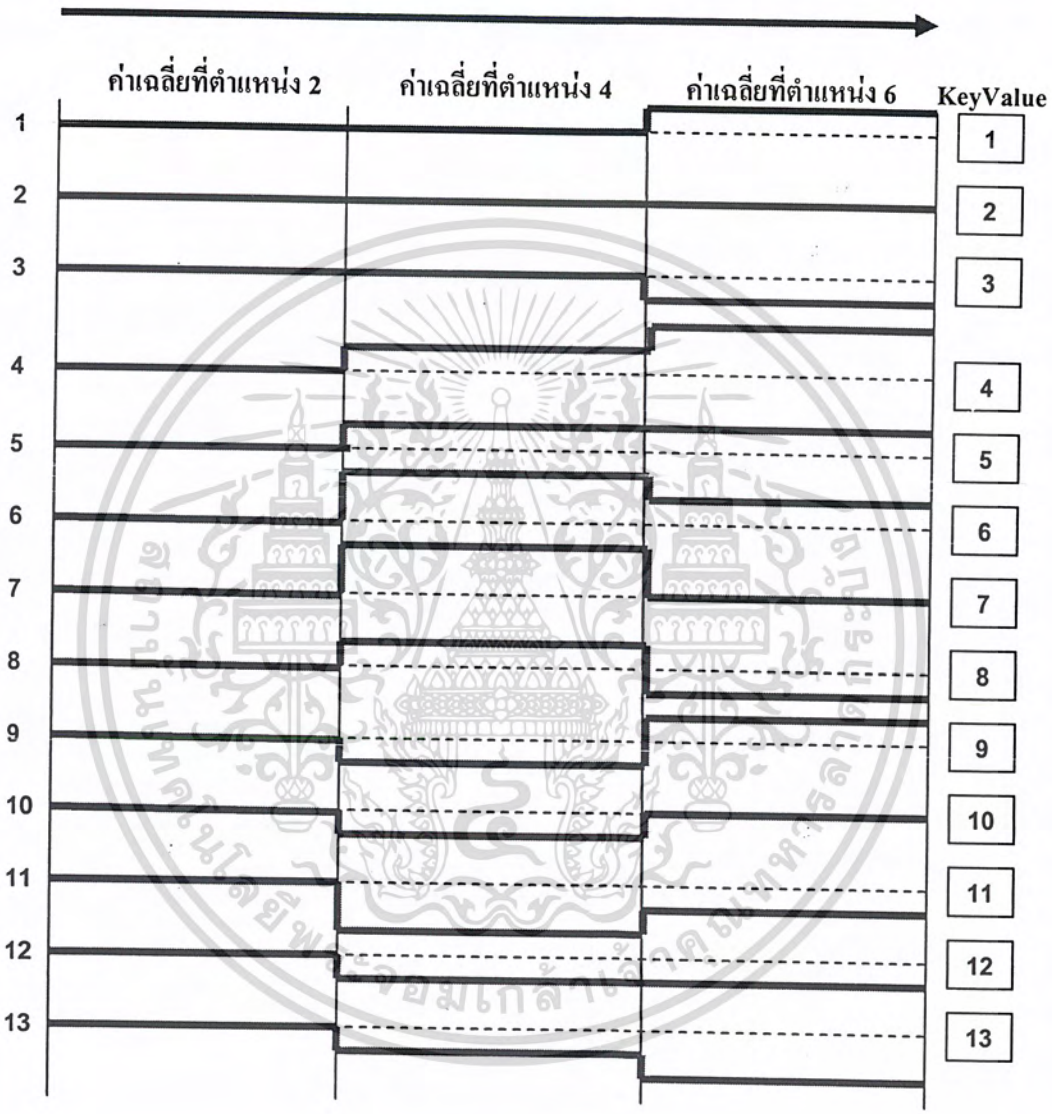
ค่าเฉลี่ยสีที่เราเลือกใช้เราจะไม่เลือกใช้ค่าเฉลี่ยสีของทั้งภาพ หลังจากที่เราได้แบ่งรูปภาพออกเป็น 9 ส่วนแล้ว เราจะเลือกหาค่าเฉลี่ยสีมาพิจารณาเพียงแค่ 3 รูปภาพย่อยจากทั้งหมด 9 รูปภาพย่อย คือในตำแหน่งที่ 2,4,6 เป็นแนวเส้นทแยงมุม เพื่อประหยัดเวลาการทำ ดังรูปที่ 4.1



รูปที่ 5.1 ตำแหน่งของรูปภาพย่อยที่พิจารณา

เมื่อเราหาค่าเฉลี่ยของสีที่ตำแหน่งดังกล่าวได้แล้ว เราก็นำมาพิจารณาหาแนวโน้มของค่าเฉลี่ยของสีที่เปลี่ยนแปลงจากมุมบนขวาไปมุมล่างซ้าย เพื่อจัดเก็บเป็นรูปแบบการเปลี่ยนแปลงของสีเฉลี่ยให้เป็นลักษณะเฉพาะอีกอย่างหนึ่งของแต่ละรูปภาพ โดยรูปแบบการเปลี่ยนแปลงของค่าเฉลี่ยสีสามารถแบ่งออกได้เป็น 13 แบบ ดังแสดงในรูปที่ 5.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.2 รูปแบบการเปลี่ยนแปลงค่าเฉลี่ยสีในรูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3 การจัดเก็บลักษณะเฉพาะของรูปภาพ ลงในฐานข้อมูล

เริ่มด้วยนำรูปภาพที่ต้องการจัดเก็บมาทำการปรับขนาดให้มีขนาดตามที่กำหนด แล้วแบ่งรูปภาพนั้นออกเป็น 9 ภาพย่อยๆ ที่มีขนาดเท่ากัน นำรูปภาพย่อยทั้ง 9 รูปมาหาลักษณะเฉพาะทั้งฮีสโตแกรม , ฟูริเยร์ดิสสคริปเตอร์ , เกเบอร์ดิสสคริปเตอร์ และรูปแบบการเปลี่ยนแปลงค่าเฉลี่ยสีของแต่ละรูปภาพย่อย หลังจากนั้นนำชื่อรูปภาพและรูปแบบการเปลี่ยนแปลงค่าเฉลี่ยสีไปเก็บลงในตารางชื่อ Identifier แล้วนำค่า ImgNo ไปใช้ในการเก็บในตาราง ImageFeature พร้อมลำดับรูปภาพย่อย , ฮีสโตแกรม , ฟูริเยร์ดิสสคริปเตอร์ และเกเบอร์ดิสสคริปเตอร์ ดังแสดงการจัดเก็บลักษณะเฉพาะและรูปแบบการเปลี่ยนแปลงค่าเฉลี่ยสีของรูปภาพ ดังรูปที่ 5.3 และ รูปที่ 5.4

5.4 วิธีการค้นหารูปภาพ

เริ่มจากนำภาพที่ต้องการค้นหามาปรับขนาดให้มีขนาดตามที่กำหนดก่อน แล้วแบ่งรูปภาพนั้นออกเป็น 9 รูปภาพย่อยๆ ที่มีขนาดเท่ากัน หลังจากนั้นก็หาลักษณะเฉพาะของรูปภาพที่ต้องการใช้ในการค้นหารูปภาพอย่างน้อย 1 อย่าง พร้อมทั้งค่าความตรงประเด็นที่ต้องการ(1% - 100%) แล้วจึงเตรียมลักษณะเฉพาะของรูปภาพที่ต้องการใช้ หลังจากได้ค่าลักษณะเฉพาะแล้วก็ดึงรูปภาพที่อยู่ในฐานข้อมูลมาเปรียบเทียบ โดยสามารถดึงรูปภาพขึ้นมาใช้เปรียบเทียบนั้นทำได้ 2 วิธี

5.4.1 ดึงรูปภาพขึ้นมาเปรียบเทียบแบบปกติ (ทั้งหมด)

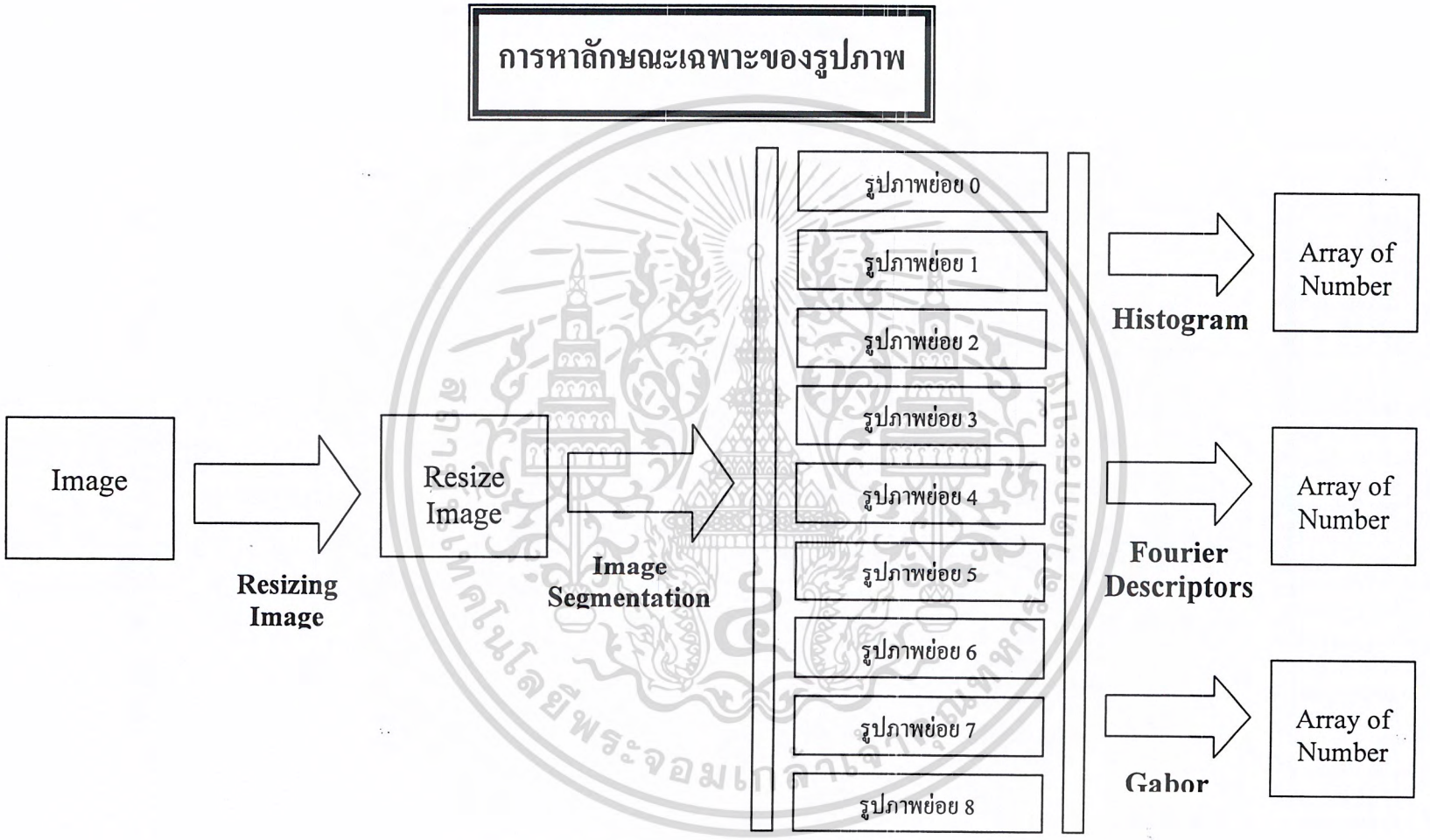
เลือกรูปภาพทั้งหมดที่อยู่ในฐานข้อมูล โดยไม่มีคัดเลือกก่อน ซึ่งรูปภาพบางรูปภาพที่อยู่ในฐานข้อมูลนั้นอาจไม่ได้มีความคล้ายคลึงกับรูปภาพที่ต้องการ

5.4.2 ดึงรูปภาพขึ้นมาเปรียบเทียบโดยใช้ Key ช่วย (เฉพาะที่คล้าย)

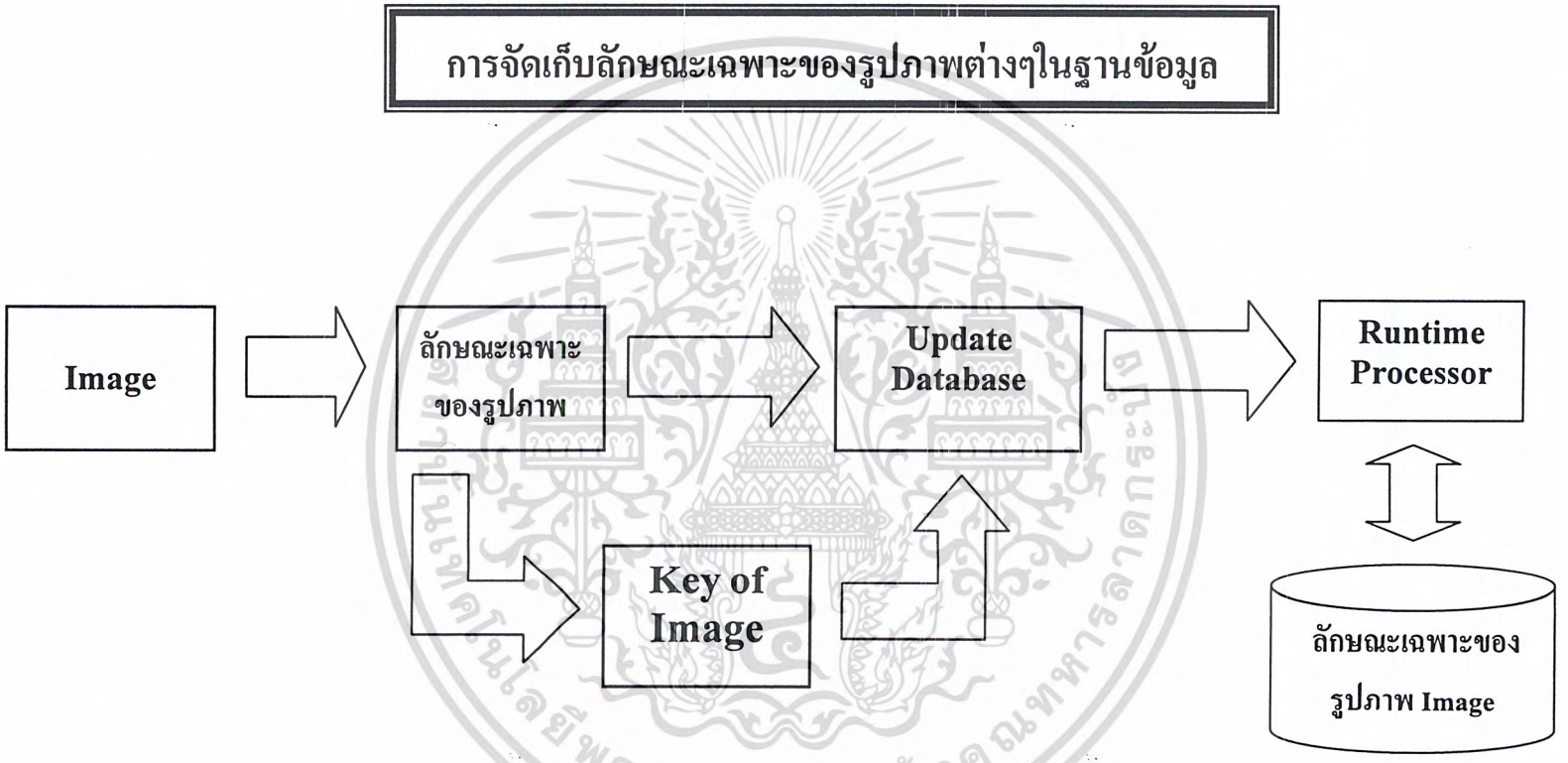
ก่อนที่จะทำการดึงรูปภาพขึ้นมาจากฐานข้อมูล จะต้องพิจารณาว่ารูปภาพที่ต้องการค้นหานั้นมีรูปแบบการเปรียบแปลงของค่าเฉลี่ยสีอย่างไร เพื่อหาคีย์ของรูปภาพที่ต้องการค้นหานั้นก่อน โดยพิจาณาเฉพาะรูปภาพย่อยที่ 2,4,6 เท่านั้นแล้วเปรียบเทียบรูปแบบการเปลี่ยนแปลงกับรูปแบบในรูปที่ 4.2 เมื่อได้คีย์ของรูปภาพแล้ว ก็เลือกเฉพาะรูปภาพในฐานข้อมูลที่มีคีย์เหมือนกันเท่านั้นขึ้นมาเปรียบเทียบกันกับรูปภาพที่ต้องการ

ขั้นตอนในการเปรียบเทียบลักษณะเฉพาะของรูปภาพที่หาได้จากรูปภาพย่อยที่ต้องการค้นหากับลักษณะเฉพาะของรูปภาพย่อยแต่ละรูปที่ดึงขึ้นมาจากฐานข้อมูลนั้น จะพิจารณาทุกค่าว่ามีค่าเท่ากันหรืออยู่ในช่วงที่เรารับได้หรือไม่ ซึ่งถ้าใช่ก็จะถือว่าภาพที่ดึงขึ้นมาจากฐานข้อมูลนั้นเป็นภาพเดียวกับรูปภาพที่ต้องการค้นหา ดังแสดงในรูปที่ 5.5

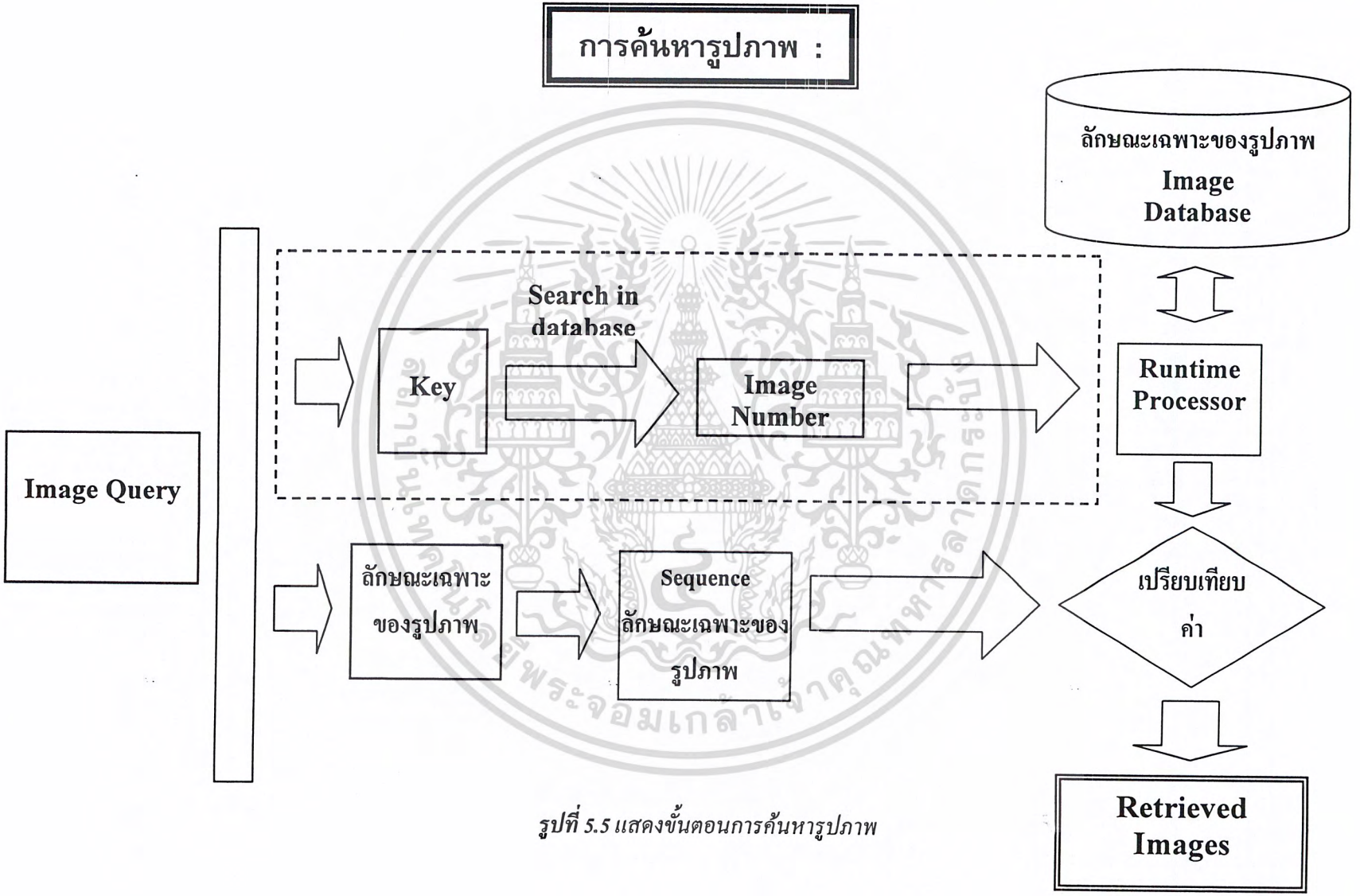
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.3 แสดงการแบ่งรูปภาพย่อยและหาลักษณะเฉพาะของรูปภาพย่อย



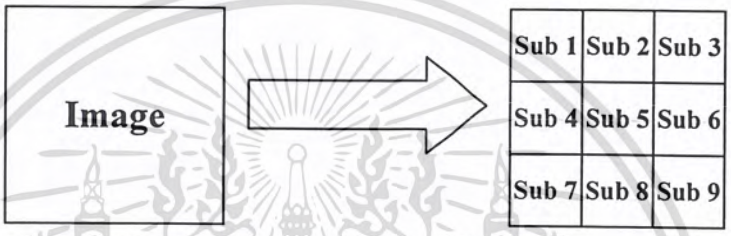
รูปที่ 5.4 การจัดเก็บรูปภาพลงฐานข้อมูล



รูปที่ 5.5 แสดงขั้นตอนการค้นหารูปภาพ

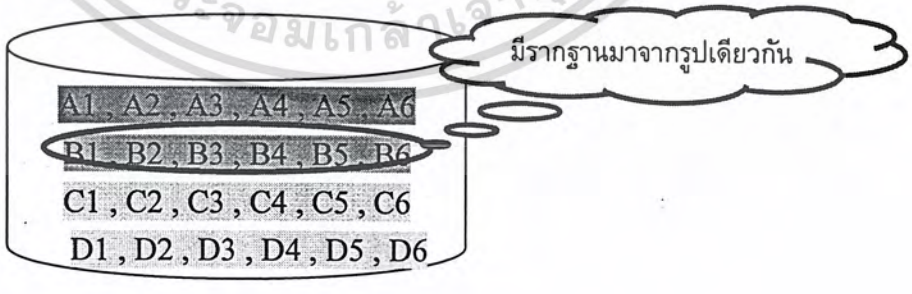
บทที่ 6
การทดลองและผลการทดลอง

ก่อนที่จะเริ่มการทดลองเราต้องมีฐานข้อมูลที่เก็บ ลักษณะเฉพาะของรูปภาพ ต่างๆของรูปภาพที่จะใช้ในการทดลอง รูปภาพต่างๆทุกรูปไม่ว่าที่อยู่ในฐานข้อมูลหรือภาพที่ใช้ในการค้นหาเราจะแบ่งแต่ละรูปออกเป็น 9 รูปย่อย (9 รูปภาพย่อย) คือแทนที่เราจะหาค่า ลักษณะเฉพาะของรูปภาพ ของทั้งรูปภาพเราก็หา ลักษณะเฉพาะของรูปภาพ ของแต่ละ รูปภาพย่อย ในแต่ละรูปภาพ เพราะจะทำให้การค้นหารูปภาพมีความแม่นยำประเด็นมากขึ้น ดังรูปที่ 6.1



รูปที่ 6.1 แบ่งภาพเป็น 9 รูปภาพย่อย

ในการทดลองนี้เราจะใช้กลุ่มของรูปภาพจำนวน 800 รูป ซึ่งน่าจะเป็นจำนวนที่สามารถทำให้เราเห็นความแตกต่างของความเร็วได้ และในกลุ่มรูปภาพเหล่านั้นก็จะมีกลุ่มรูปภาพย่อยๆที่เป็นรูปภาพเดียวกันแต่ถูกแก้ไขดัดแปลงให้มีความแตกต่างจากเดิม เช่น ทำการย่อ,ขยาย,เปลี่ยนแปลงสี,ลดความชัดเจน เป็นต้น เพื่อให้เห็นความแตกต่างของความตรงประเด็นได้ชัดเจนมากขึ้น



รูปที่ 6.2 การแบ่งรูปภาพเป็นกลุ่มๆในฐานข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การออกแบบการทดลองนั้น จะถูกแบ่งออกเป็นชุดการทดลองต่างๆทั้งหมด 6 Scenarios ซึ่งแต่ละชุดการทดลองก็จะมีวัตถุประสงค์ของการทดลองแตกต่างกันไป โดยการออกแบบการทดลองจะคำนึงถึงวัตถุประสงค์สำคัญหลักๆ 2 วัตถุประสงค์ คือ *ความถูกต้อง*(เปรียบเทียบที่ความตรงประเด็น) และ *ความรวดเร็ว*(เปรียบเทียบที่ response time) ที่มีผลต่อการค้นหารูปภาพ การทดลองที่ได้ออกแบบไว้ทั้งหมดอาจจะยังไม่ครอบคลุมทุกกรณี แต่จะเลือกมาเฉพาะบางกรณีที่เห็นว่าน่าสนใจและให้ผลที่แตกต่างกันอย่างชัดเจน โดยคำนึงถึงระยะเวลาในการทดลองด้วย การทดลองทุกชุดการทดลองจะใช้กับกลุ่มรูปภาพกลุ่มเดียวกัน

<u>ตัวแปรที่ต้องควบคุม</u>

1. ลักษณะเฉพาะของรูปภาพ ของรูปภาพที่ถูกจัดเก็บฐานข้อมูล
   รูปภาพที่จะใช้ในการค้นหาต้องเป็นชุดเดียวกัน ซึ่งจะใช้ชุดของรูปภาพที่มีรูปภาพลักษณะต่างๆกันไป เช่น รูปภาพที่เป็นสิ่งของเดียวๆ , รูปภาพที่เป็นพื้นผิว , รูปภาพที่เป็นลายเส้น , รูปภาพที่มีหลายๆองค์ประกอบ เป็นต้น จำนวนเท่าๆกัน
2. ค่าคงที่ต่างที่ใช้ในการหา ลักษณะเฉพาะของรูปภาพ
3. ต้องควบคุมความตรงประเด็นที่เราต้องการ เช่น ต้องการให้เหมือนกัน 100% เป็นต้น ซึ่งถ้าเรากำหนดความตรงประเด็นให้มีความตรงประเด็น 100% เราจะใช้เวลาในการค้นหาภาพน้อย ถ้าเราอยากได้รูปที่มีลักษณะคล้ายกันออกมาด้วยก็ไม่จำเป็นต้องกำหนดให้เหมือนกัน 100% แต่ก็จะต้องใช้เวลาในการค้นหานานขึ้น เนื่องจากวัตถุประสงค์เราต้องการภาพที่เหมือนกันจึงใช้ 100%
4. เครื่องที่ใช้ทดลองต้องเป็นเครื่องเดียวกัน และต้องไม่มีการเปลี่ยน config ต่างๆของเครื่อง รวมทั้งสภาพแวดล้อมต่างๆต้องเหมือนกันในทุกครั้งที่ทำการทดลอง

<u>ตัวแปรที่เปลี่ยนแปลงได้</u>

1. ลักษณะเฉพาะของรูปภาพ ที่จะใช้ในการค้นหารูปภาพ
2. วิธีในการค้นหารูปภาพ
3. รูปภาพที่ต้องการค้นหา

<u>เกณฑ์การวัดประสิทธิภาพ</u>

1. ความถูกต้อง(ตรงประเด็น)
2. ความรวดเร็ว(เวลาที่ใช้)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.1 ชุดการทดลอง

Part 1: Image Processing

Goal : หาลำดับก่อน-หลังของ ลักษณะเฉพาะของรูปภาพ ที่มีความถูกต้องและรวดเร็วในการค้นหารูปภาพ

Scenario 1 : Histogram

- เพื่อดูว่าลักษณะสีอย่างเดียวสามารถค้นหาภาพได้ถูกต้องและรวดเร็วแค่ไหน

Scenario 2 : Histogram + Fourier Descriptor

- เพื่อดูว่าลักษณะสีและลักษณะเส้นสามารถค้นหาภาพได้ถูกต้องและรวดเร็วแค่ไหน

Scenario 3 : Fourier Descriptor + Gabor

- เพื่อดูว่าลักษณะเส้นและลักษณะพื้นผิวสามารถค้นหาภาพได้ถูกต้องและรวดเร็วแค่ไหน

Part 2: Image Processing + Search Algorithm

Goal : พัฒนาลำดับของ ลักษณะเฉพาะของรูปภาพ ให้มีความรวดเร็วขึ้นโดยใช้วิธีค้นหาเข้ามาช่วย

Scenario 4 : Search Algorithm + Fourier Descriptor + Gabor

- เพื่อดูว่าลักษณะเส้นและลักษณะพื้นผิวสามารถค้นหาภาพได้ถูกต้องแค่ไหน และเมื่อเพิ่มวิธีค้นหาเพื่อเพิ่มความรวดเร็วในการหามากขึ้น

Scenario 5 : Search Algorithm + Histogram + Fourier Descriptor + Gabor

- เพื่อดูว่าลักษณะของสี,ลักษณะเส้น,ลักษณะพื้นผิวและใช้วิธีค้นหาจะค้นหาภาพได้ถูกต้องมากที่สุดแต่จะทำให้การทำงานช้าลงหรือไม่

ในการทดลองนั้นเราจะเลือกสุ่มรูปภาพต้นแบบมา 6 ชุด (a,b,c,d,e,f) ชุดละ 5 รูปที่ไม่ซ้ำกัน เพื่อใช้เป็น query ในการทดลองหารูปภาพที่เหมือนกัน โดยเลือกใช้วิธีที่การค้นหาต่างกันตาม senario ที่กำหนดไว้ หลังจากนั้นเราก็จะนำผลการค้นหาที่ตรงประเด็นและเวลาที่ใช้มาวัดผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.2 ผลการทดลอง

ชุดที่ a

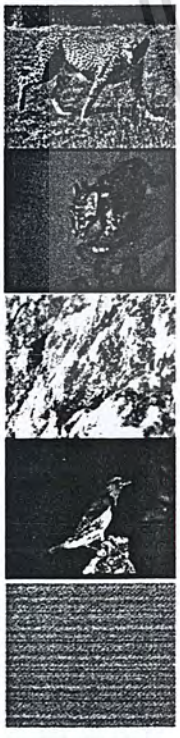
Scenario 1

รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 0 : 4	800
0 : 0 : 4	800
0 : 0 : 5	800
0 : 0 : 4	800
0 : 0 : 4	800

(a1)

Scenario 2



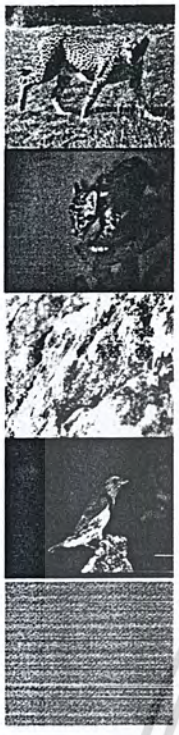
รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 0 : 4	800
0 : 0 : 5	800
0 : 0 : 5	800
0 : 0 : 4	800
0 : 0 : 5	800

(a2)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Scenario 3

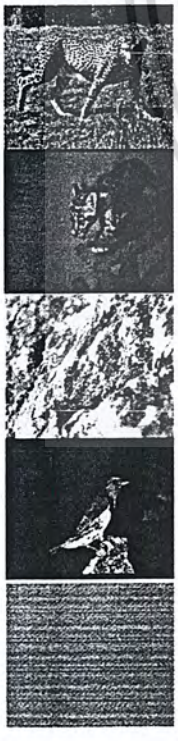


รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 17	800
0 : 1 : 16	800
0 : 1 : 16	800
0 : 1 : 17	800
0 : 1 : 17	800

(a3)

Scenario 4



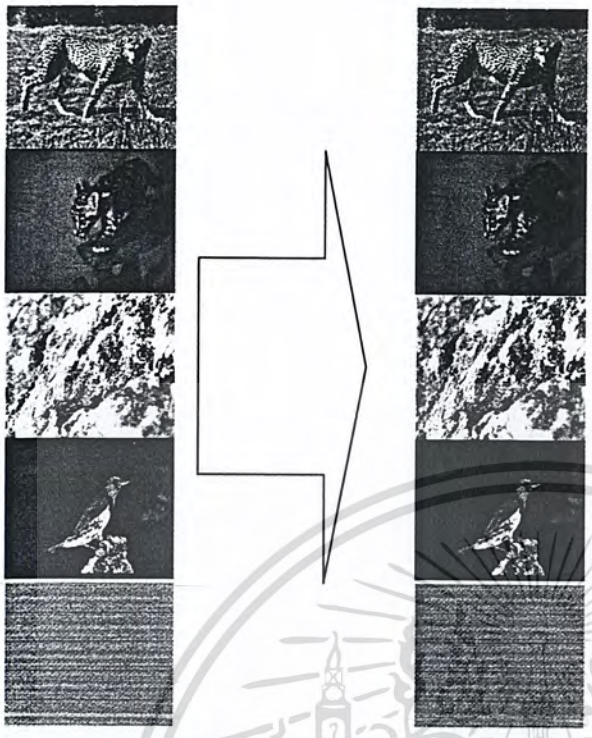
รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 13	148
0 : 1 : 13	148
0 : 1 : 12	72
0 : 1 : 13	183
0 : 1 : 12	3

(a4)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

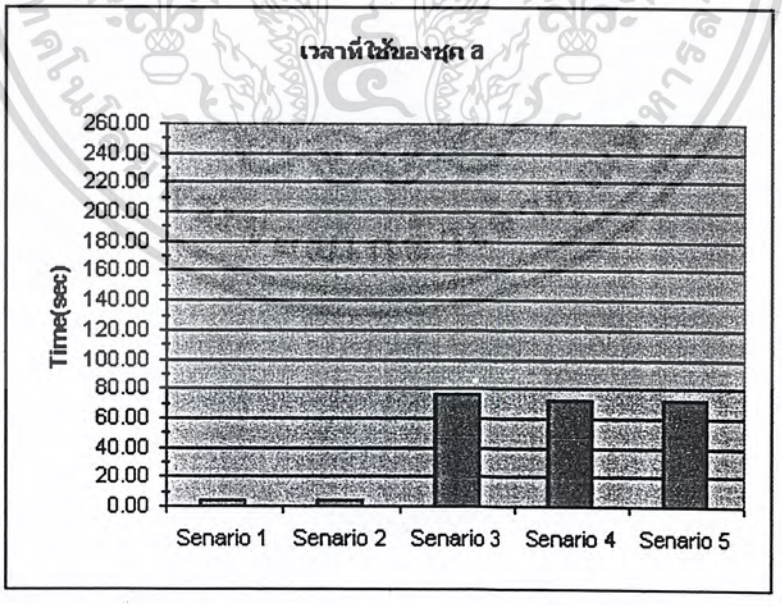
Scenario 5



รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 12	148
0 : 1 : 13	148
0 : 1 : 13	72
0 : 1 : 13	183
0 : 1 : 12	3

(a5)

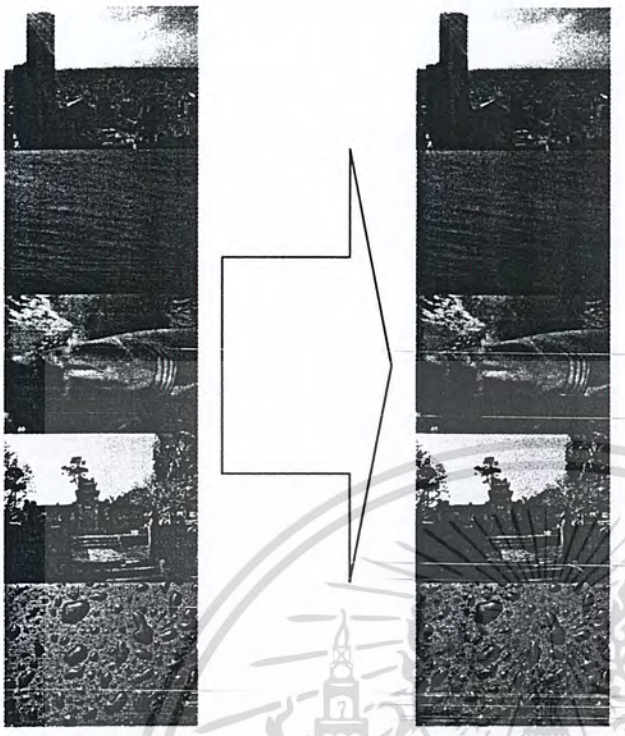


รูปที่ 6.3 กราฟแสดงเวลาที่ใช้ในการค้นหาของชุดรูปภาพ a โดยใช้ชุดการทดลองต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชุดที่ b

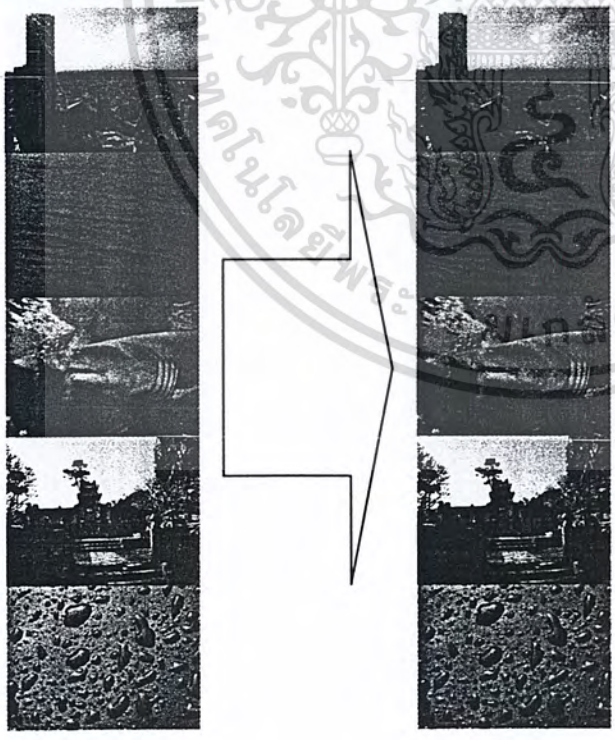
Scenario 1



เวลา	จำนวนภาพที่ค้นหา
0 : 0 : 5	800
0 : 0 : 4	800
0 : 0 : 5	800
0 : 0 : 5	800
0 : 0 : 5	800

(b1)

Scenario 2

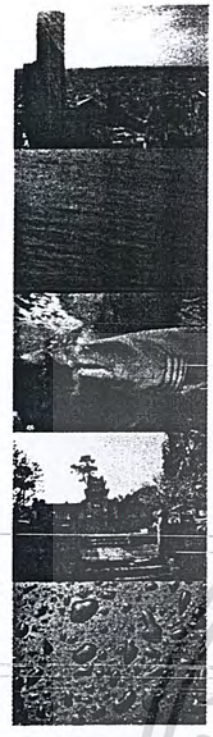


เวลา	จำนวนภาพที่ค้นหา
0 : 0 : 5	800
0 : 0 : 5	800
0 : 0 : 5	800
0 : 0 : 4	800
0 : 0 : 5	800

(b2)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Scenario 3



รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 16	800
0 : 1 : 16	800
0 : 1 : 17	800
0 : 1 : 16	800
0 : 1 : 17	800

(b3)

Scenario 4

รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 13	231
0 : 1 : 13	231
0 : 1 : 13	183
0 : 1 : 12	231
0 : 1 : 13	183

(b4)

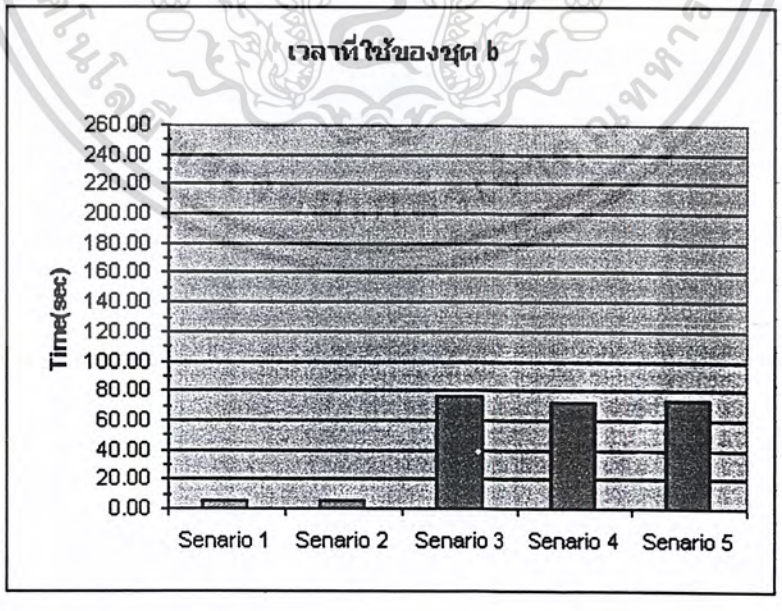
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Scenario 5

รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 13	231
0 : 1 : 14	231
0 : 1 : 13	183
0 : 1 : 13	231
0 : 1 : 12	183

(b5)

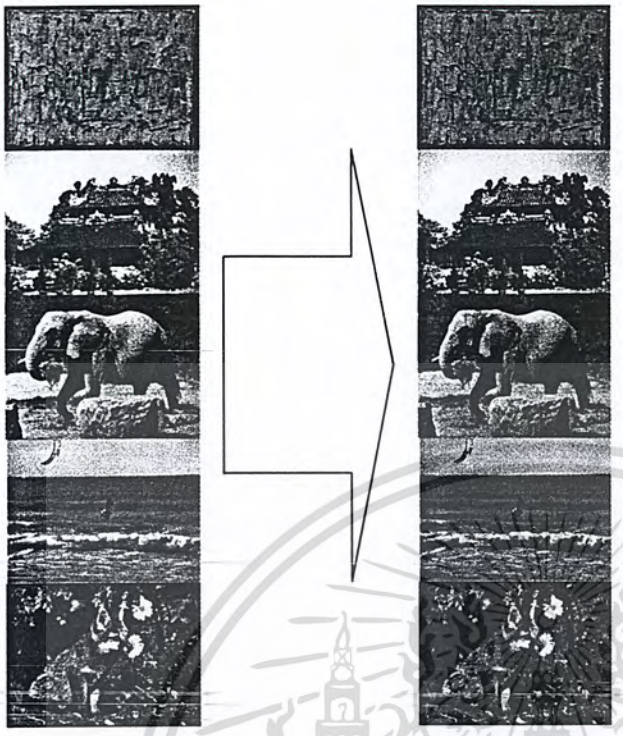


รูปที่ 6.4 กราฟแสดงเวลาที่ใช้ในการค้นหาของชุดรูปภาพ b โดยใช้ชุดการทดลองต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชุดที่ c

Scenario 1

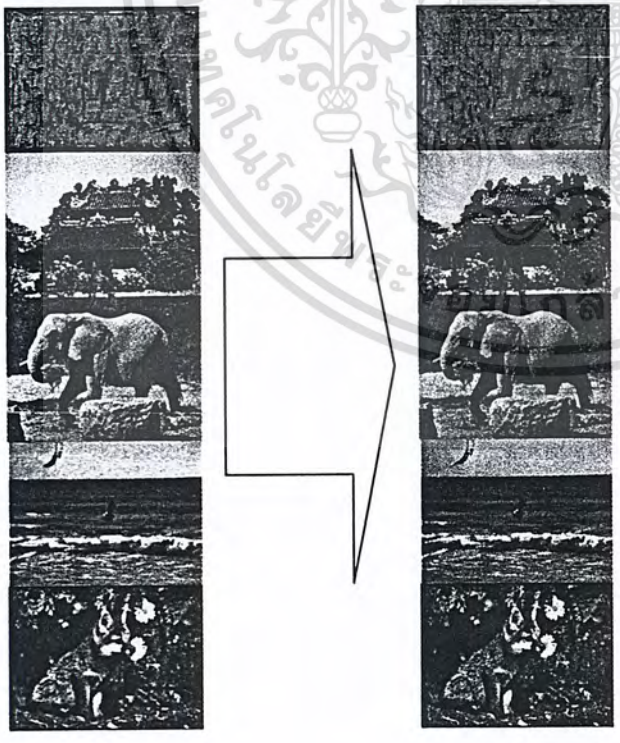


รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 0 : 4	800
0 : 0 : 4	800
0 : 0 : 4	800
0 : 0 : 4	800
0 : 0 : 4	800

(c1)

Scenario 2



รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 0 : 4	800
0 : 0 : 4	800
0 : 0 : 4	800
0 : 0 : 4	800
0 : 0 : 5	800

(c2)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Scenario 3

รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 16	800
0 : 1 : 16	800
0 : 1 : 17	800
0 : 1 : 16	800
0 : 1 : 16	800

(c3)

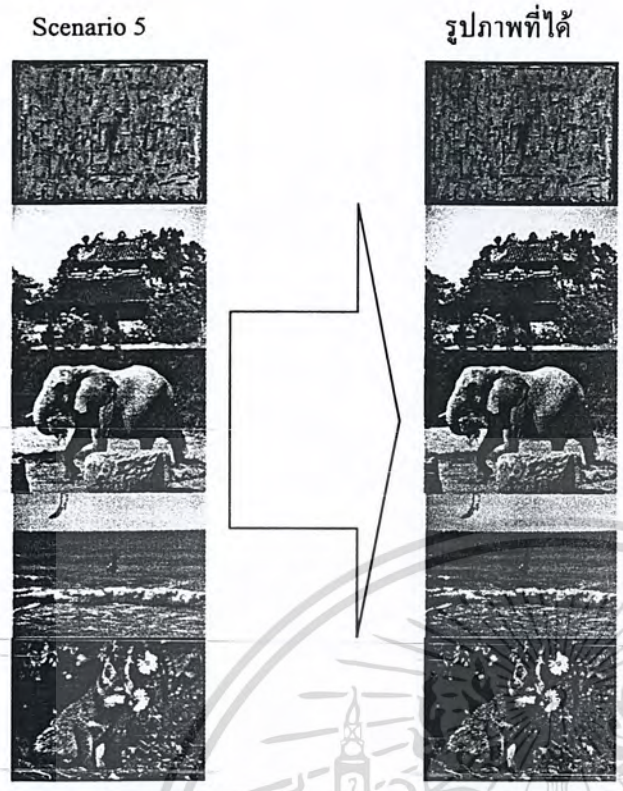
Scenario 4

รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 13	183
0 : 1 : 13	231
0 : 1 : 12	97
0 : 1 : 12	87
0 : 1 : 13	148

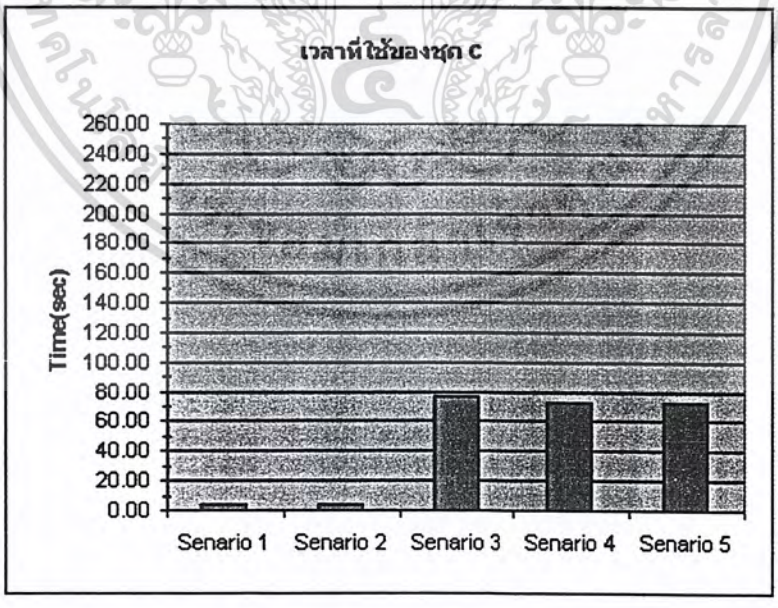
(c4)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 13	178
0 : 1 : 13	236
0 : 1 : 12	101
0 : 1 : 12	85
0 : 1 : 13	141

(c5)

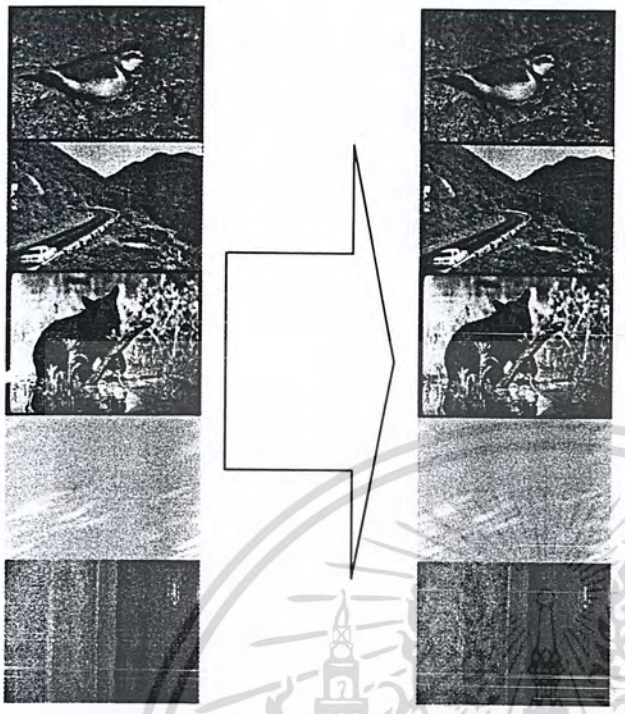


รูปที่ 6.5 กราฟแสดงเวลาที่ใช้ในการค้นหาของชุดรูปภาพ c โดยใช้ชุดการทดลองต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชุดที่ d

Scenario 1

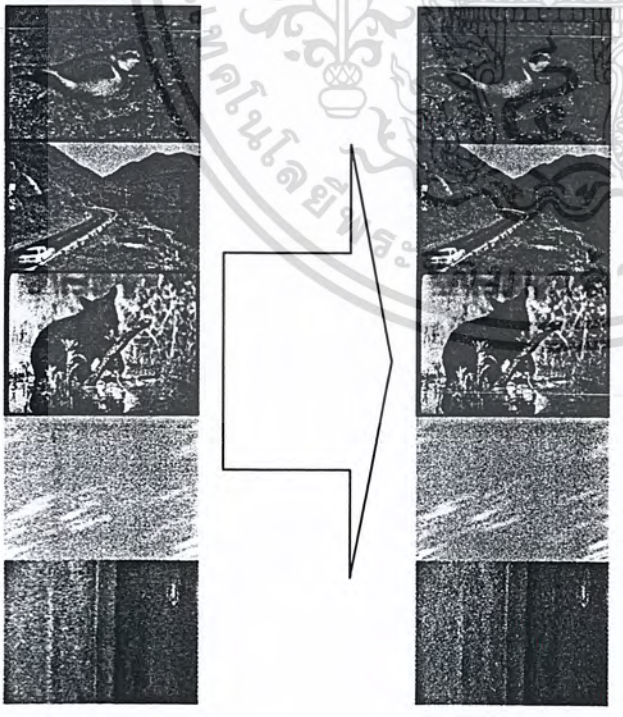


รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 0 : 5	800
0 : 0 : 4	800
0 : 0 : 4	800
0 : 0 : 4	800
0 : 0 : 4	800

(d1)

Scenario 2



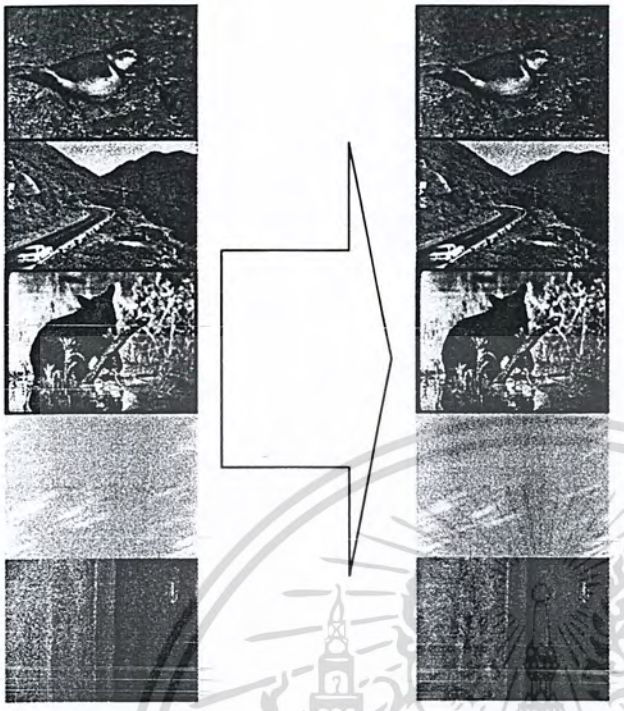
รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 0 : 4	800
0 : 0 : 5	800
0 : 0 : 5	800
0 : 0 : 5	800
0 : 0 : 5	800

(d2)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Scenario 3

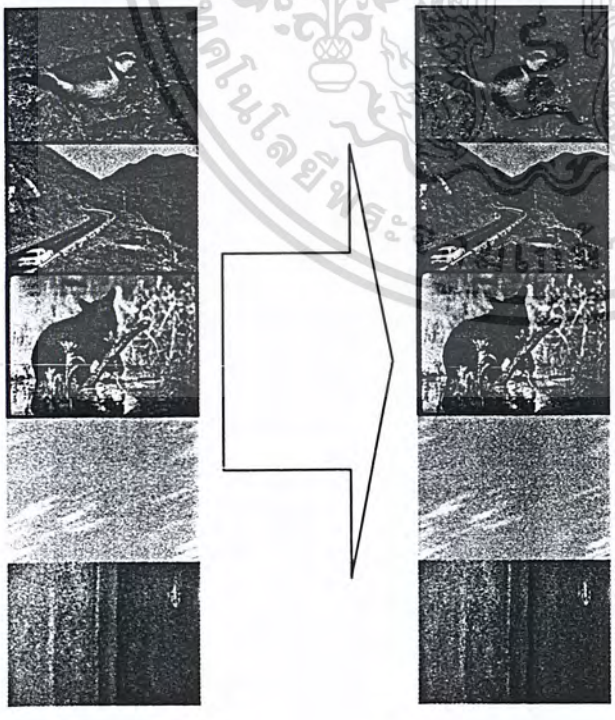


รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 16	800
0 : 1 : 17	800
0 : 1 : 16	800
0 : 1 : 16	800
0 : 1 : 17	800

(d3)

Scenario 4



รูปภาพที่ได้

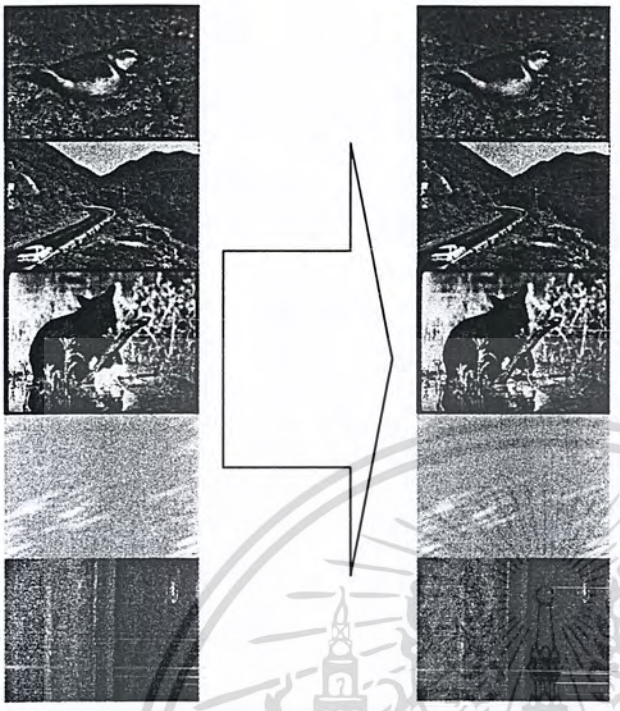
เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 13	183
0 : 1 : 12	72
0 : 1 : 12	87
0 : 1 : 12	72
0 : 1 : 12	72

(d4)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

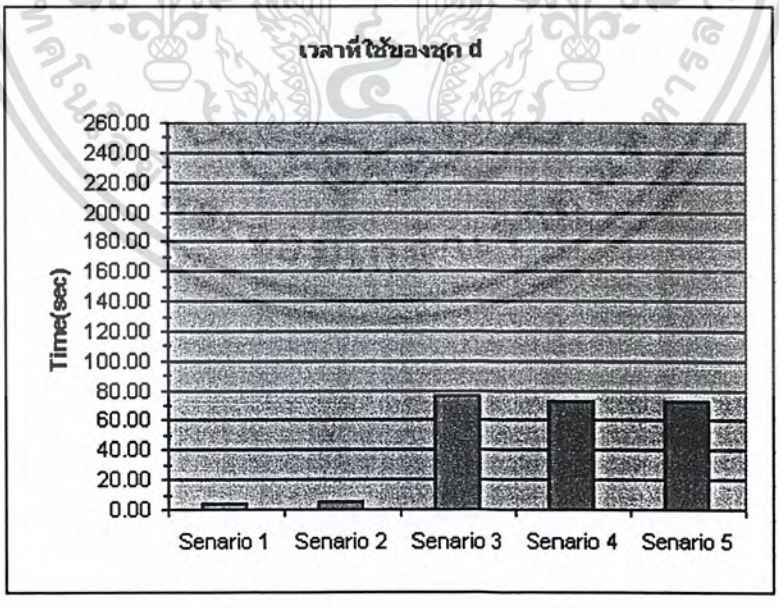
Scenario 5

รูปภาพที่ได้



เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 13	183
0 : 1 : 12	72
0 : 1 : 12	87
0 : 1 : 12	72
0 : 1 : 13	72

(d5)



***รูปที่ 6.6** กราฟแสดงเวลาที่ใช้ในการค้นหาของชุดรูปภาพ d โดยใช้ชุดการทดลองต่างๆ*

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชุดที่ e

Scenario 1

รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 0 : 4	800
0 : 0 : 5	800
0 : 0 : 5	800
0 : 0 : 5	800
0 : 0 : 5	800

(e1)

Scenario 2

รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 0 : 5	800
0 : 0 : 5	800
0 : 0 : 4	800
0 : 0 : 5	800
0 : 0 : 5	800

(e2)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Scenario 3

รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 16	800
0 : 1 : 16	800
0 : 1 : 16	800
0 : 1 : 16	800
0 : 1 : 16	800

(e3)

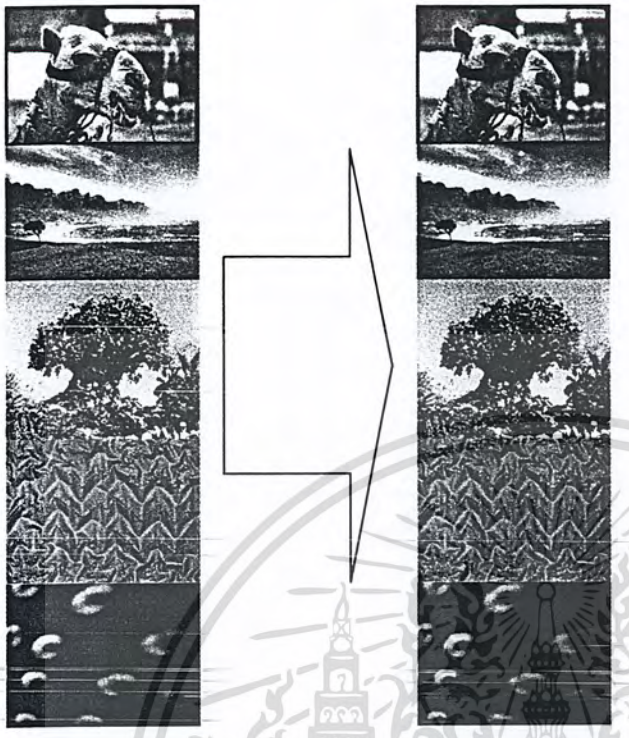
Scenario 4

รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 13	231
0 : 1 : 13	183
0 : 1 : 12	87
0 : 1 : 12	148
0 : 1 : 13	97

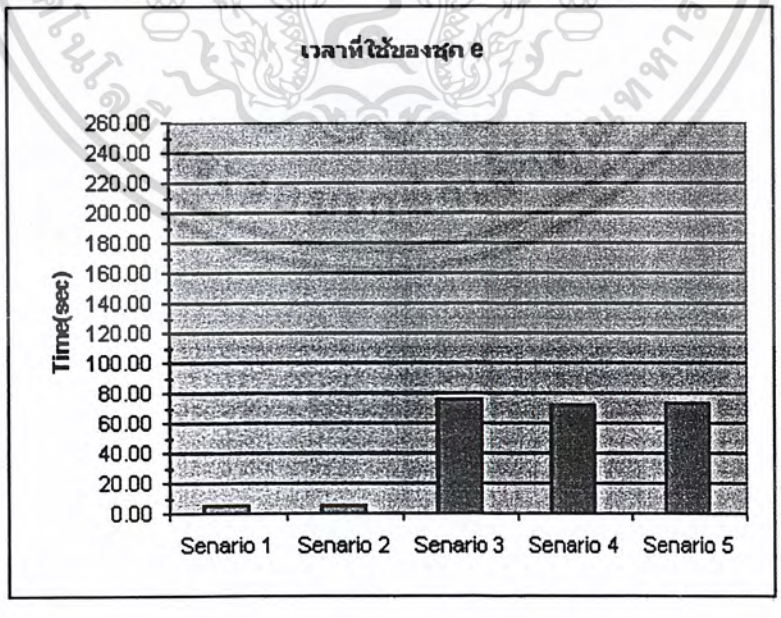
(e4)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 13	231
0 : 1 : 13	183
0 : 1 : 13	87
0 : 1 : 13	148
0 : 1 : 13	97

(e5)



รูปที่ 6.7 กราฟแสดงเวลาที่ใช้ในการค้นหาของชุดรูปภาพ e โดยใช้ชุดการทดลองต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชุดที่ f

Scenario 1

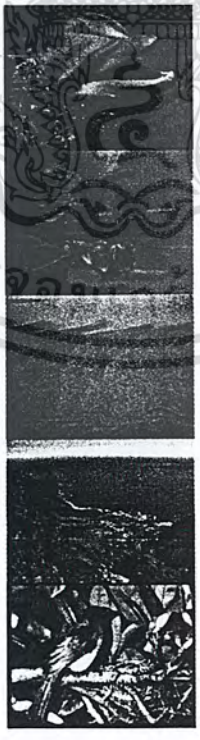
รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 0 : 5	800
0 : 0 : 4	800
0 : 0 : 5	800
0 : 0 : 5	800
0 : 0 : 4	800

(f1)

Scenario 2

รูปภาพที่ได้

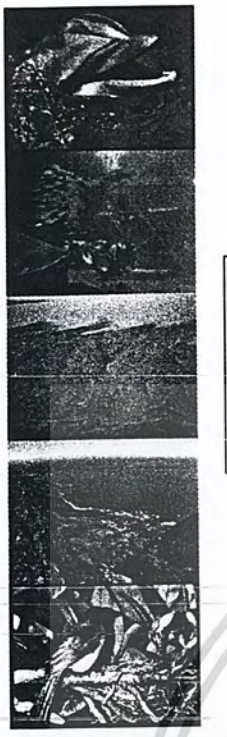


เวลา	จำนวนภาพที่ค้นหา
0 : 0 : 4	800
0 : 0 : 5	800
0 : 0 : 5	800
0 : 0 : 5	800
0 : 0 : 4	800

(f2)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Scenario 3



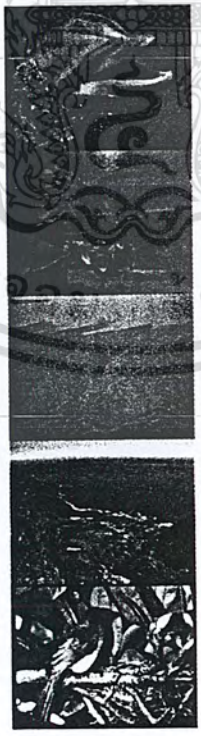
รูปภาพที่ได้

เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 16	800
0 : 1 : 16	800
0 : 1 : 16	800
0 : 1 : 16	800
0 : 1 : 16	800

(f3)

Scenario 4

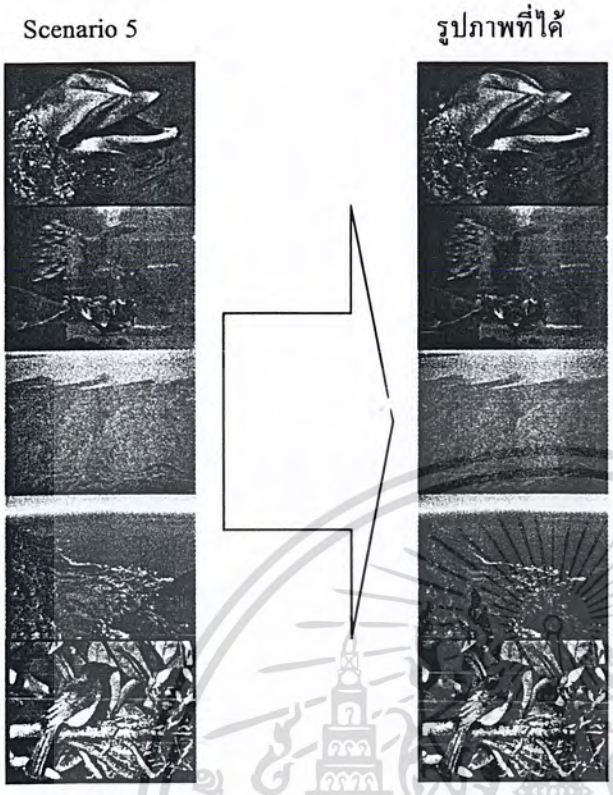
รูปภาพที่ได้



เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 12	148
0 : 1 : 13	231
0 : 1 : 13	183
0 : 1 : 13	87
0 : 1 : 12	183

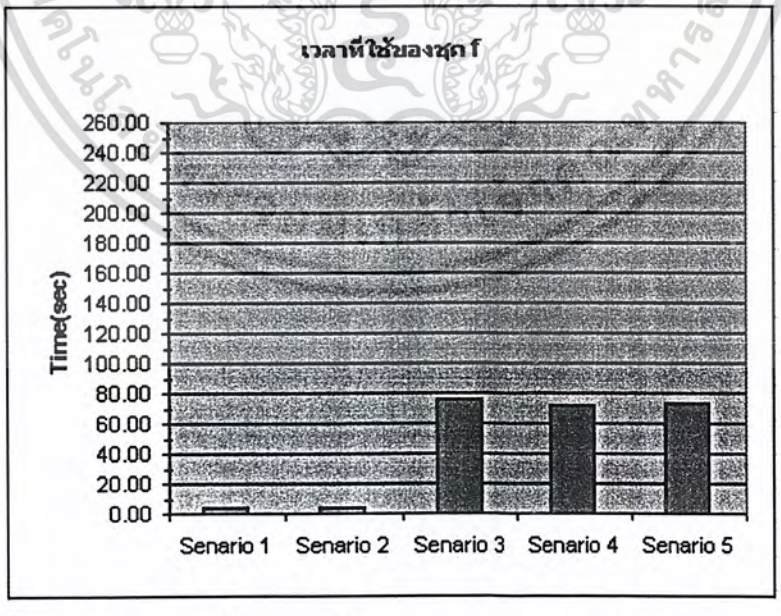
(f4)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เวลา	จำนวนภาพที่ค้นหา
0 : 1 : 13	148
0 : 1 : 13	231
0 : 1 : 13	183
0 : 1 : 12	87
0 : 1 : 13	183

(f5)

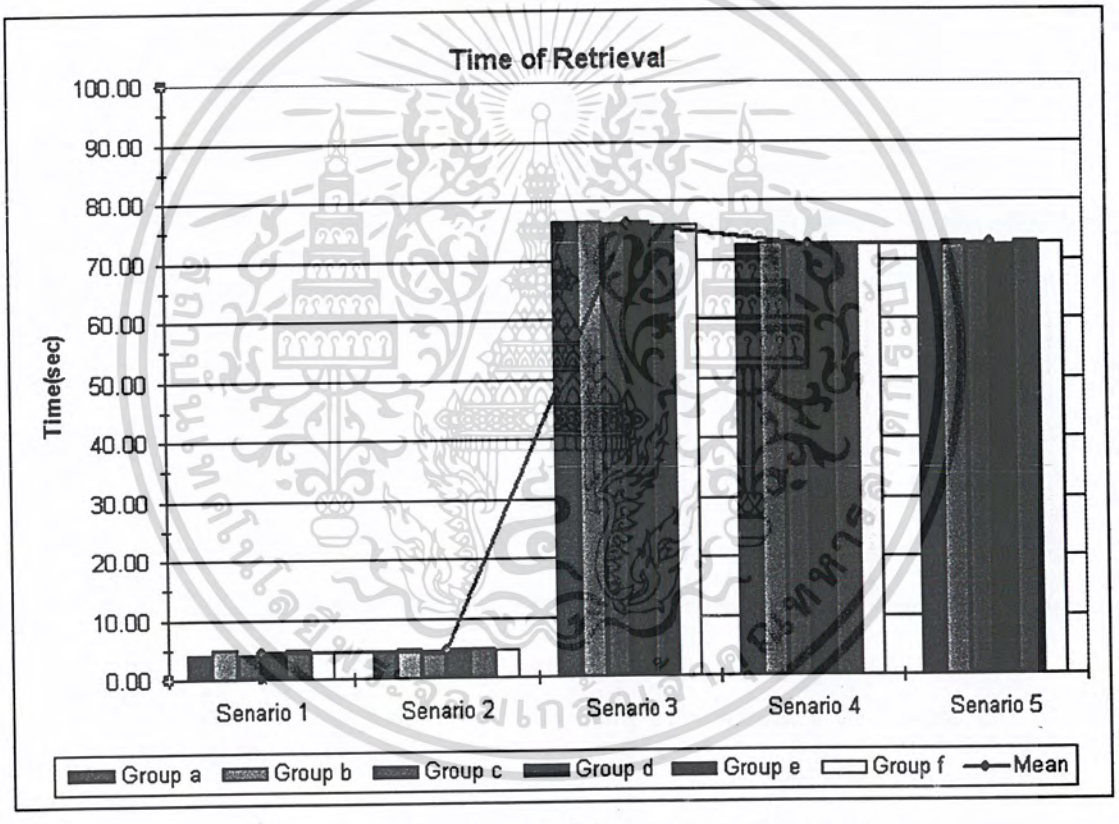


รูปที่ 6.8 กราฟแสดงเวลาที่ใช้ในการค้นหาของชุดรูปภาพ f โดยใช้ชุดการทดลองต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.3 สรุปผลการทดลอง

หลังที่เราได้ทำการทดลองค้นหารูปภาพตามการทดลองที่ได้ออกแบบไว้ โดยใช้ชุดรูปภาพทั้งหมด6 ชุดๆละ 5 รูปและค้นตามชุดการทดลอง ทั้ง 5 แบบ ผลลัพธ์ที่ได้คือสามารถค้นหารูปภาพที่ต้องการได้อย่างถูกต้องตามที่ต้องการเหมือนกันทุกชุดการทดลอง แต่เวลาที่ใช้ในการค้นหาแต่ละชุดการทดลองไม่เท่ากัน การค้นหาภาพโดยใช้ชุดการทดลองที่ 1 จะใช้เวลาในการค้นหาน้อยที่สุด ซึ่งเป็นการใช้ฮีสโตแกรมของภาพในการค้นหาอย่างเดียว ส่วนชุดการทดลองที่มีการใช้เกเบอร์เป็นส่วนหนึ่งในการค้นหาจะใช้เวลานานมากที่สุด และในชุดการทดลองที่ 4 และ 5 ที่มีการใช้คีย์ช่วยในการค้นหาภาพนั้นทำให้สามารถลดจำนวนรูปภาพในฐานข้อมูลที่นำมาเปรียบเทียบทำให้ใช้เวลาในการค้นหาน้อยลง



รูปที่ 6.9 กราฟเปรียบเทียบเวลาที่ใช้ในการค้นหา แต่ละชุดรูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 7
วิจารณ์และสรุป

7.1 สรุปและวิจารณ์ผลการทดลอง

การทดลองที่ได้ออกแบบนั้นเป็นการสุ่มตัวอย่างวิธีการค้นหาทั้งหมด ที่เกิดจากการผสมผสานกันระหว่างฮีสโตรแกรม , ฟูริเยร์ดิสสคริปเตอร์ และเกเบอร์ โดยเลือกจากความน่าสนใจที่แตกต่างกัน ทำให้บางชุดการทดลองนั้นอาจจะไม่มีประโยชน์ในการใช้งานจริง

โดยความเป็นจริงแล้วความหมายของความแม่นยำในงานบางงานไม่ได้หมายถึงค้นหารูปได้ตรงความต้องการเพียงอย่างเดียว แต่ยังอาจหมายถึงความต้องการรูปภาพที่มีความคล้ายคลึงกันอีกด้วย ซึ่งในงานวิจัยนี้ไม่ได้พูดถึง แต่ในการทดลองเราได้ทดลองใช้ความถูกต้องที่ไม่ถึง 100% ด้วย ซึ่งก็ให้ผลรูปภาพที่มีความคล้ายคลึงกันออกมาด้วย เช่น เราต้องการค้นหารูปภาพกระต่ายตัวเดียวที่อยู่ตรงกลางรูปภาพ โดยใช้วิธีการวิเคราะห์ฟูริเยร์ดิสสคริปเตอร์ ที่ตั้งค่าความเหมือนกันเพียง 90% ผลที่ได้ก็คือรูปภาพที่ออกมานั้นนอกจากจะได้รูปภาพกระต่ายที่ต้องการออกมาแล้ว ยังได้รูปภาพอื่นที่มีลักษณะลายเส้นคล้ายคลึงกัน เช่น กวางตัวเดียว ออกมาด้วย เป็นต้น ในการออกแบบการทดลองต่อไปเราอาจจะค้นหารูปภาพที่มีความคล้ายคลึงกันด้วย ซึ่งผลลัพธ์ที่ออกมาก็อาจจะมีทั้งถูกต้องและไม่ถูกต้อง แต่ผลลัพธ์ที่ดีที่สุดคือได้ผลลัพธ์ที่มีรูปภาพที่ไม่ตรงกับความต้องการเจือปนให้น้อยที่สุด

7.2 ข้อดี – ข้อเสียในงานวิจัย

7.2.1 การแบ่งรูปภาพใหญ่เป็นรูปภาพย่อย

การแบ่งรูปภาพใหญ่ให้เป็นรูปภาพย่อยๆ ก่อน ก่อนที่ทำการเก็บหรือเปรียบเทียบลักษณะเฉพาะของรูปภาพนั้นมีทั้งข้อดีและข้อเสีย ข้อดีของการแบ่งเป็นรูปภาพย่อยคือสามารถเปรียบเทียบภาพได้อย่างถูกต้องมากกว่า เพราะรายละเอียดที่เก็บนั้นเรามองรูปภาพย่อยเป็นเหมือนรูปภาพใหญ่จริงๆ

นอกจากนั้นในขั้นตอนการตรวจสอบขอบของวัตถุในรูปภาพจะให้ประสิทธิภาพที่ดีกว่า เนื่องจากว่าหลังจากที่เราตรวจหาเส้นขอบของทั้งรูปภาพแล้ว ในตอนที่เราต้องตัดเส้นบางส่วนทิ้งไปจะทำให้รายละเอียดบางอย่างหายไปได้ถ้าเรามองเป็นรูปภาพใหญ่ แต่ถ้าเรามองเป็นรูปภาพย่อยในบางรูปภาพย่อยที่ไม่มีเส้นมาก เราสามารถที่จะเก็บได้ทั้งหมด แต่ในกรณีที่รูปภาพย่อยนั้นมีเส้นมากเราก็จำเป็นต้องเสียรายละเอียดบางอย่างไปแค่ในรูปภาพย่อยนั้นๆ แต่รูปภาพย่อยที่เหลือยังคงมีรายละเอียดครบอยู่

ส่วนข้อเสียของการแบ่งเป็นรูปภาพย่อยก็คือรูปภาพที่เป็นผลลัพธ์จะไม่ยืดหยุ่น ถ้ารูปภาพมีการเปลี่ยนแปลงนิดหน่อย เช่น เส้นไม่ชัดเจนตรงส่วนหนึ่งของภาพ จะทำให้ไม่สามารถค้นหารูปภาพนั้นออกมาได้เลย วิธีแก้นั้นเราต้องปรับค่าความแตกต่างที่เรายอมรับได้ให้น้อยลง เช่น จากที่ต้องให้ถูกต้อง 100% ก็ปรับลดลงอาจจะเหลือกแค่ 90% เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อเสียอีกประการหนึ่งคือ เราต้องใช้พื้นที่ในการเก็บข้อมูลเพิ่มขึ้น แต่เดิมเราเก็บ 1 รูปใช้ 1 โรล เท่านั้น แต่ในงานวิจัยนี้เราต้องใช้ถึง 9 โรล เพื่อเก็บข้อมูลรูปภาพ 1 รูป แต่ถ้าข้อมูลในฐานข้อมูลนั้นมีขนาดไม่มากจนเกินไป และระบบจัดการฐานข้อมูลมีความสามารถดี ก็อาจจะทำให้ไม่มีปัญหามากนัก

7.2.2 การแบ่งกลุ่มรูปภาพด้วยแนวโน้มสี

การแบ่งแยกความแตกต่างของแต่ละรูปภาพโดยใช้ลักษณะการวิเคราะห์ค่าเฉลี่ยของสีที่เปลี่ยนแปลงไป มีข้อดีคือ เราสามารถที่จะแบ่งแยกรูปภาพออกเป็นกลุ่มๆได้อย่างดี ซึ่งมีผลช่วยลดเวลาในการค้นหารูปภาพให้น้อยลง แต่ก็มีข้อเสียเนื่องจากเราไม่ได้วิเคราะห์ลักษณะวัตถุภายในรูปภาพจะทำให้มีปัญหาเกิดขึ้นในกรณีที่เป็นภาพเดียวกันแต่ถูกถ่ายในช่วงเวลาที่ต่างกัน เช่น เช้า – เย็น ก็จะทำให้ค่าเฉลี่ยสีเปลี่ยนแปลงไป เป็นผลให้ไม่สามารถค้นหารูปภาพที่เหมือนกันแต่คนละสีไม่ได้ ต้องใช้วิธีค้นหาทั้งหมดแบบเดิม

7.3 การนำไปประยุกต์ใช้งาน

เนื่องจากการทดลองที่ออกแบบไว้นั้นเป็นการค้นหารูปภาพที่เหมือนกันทั้ง 100 % จึงทำให้อาจจะไม่สามารถนำไปใช้งานจริงๆได้ เพราะบางครั้งรูปต้นฉบับที่เราต้องการค้นหานั้นมีความแตกต่างกับรูปภาพในฐานข้อมูลนิดหน่อย เช่น รูปตึกในฐานข้อมูลเป็นรูปที่ถูกถ่ายขึ้นตอนกลางวันซึ่งมีแสงเงาชัดเจน แต่ว่ารูปต้นฉบับที่เราต้องการค้นหานั้นถูกถ่ายตอนเย็นก็จะเห็นแสงเงาเปลี่ยนแปลงไป ซึ่งจริงๆแล้วรูปภาพทั้ง 2 เป็นรูปภาพสิ่งเดียวกัน เป็นต้น

บางครั้งเราอาจจะต้องการค้นหารูปภาพที่เราสนใจมากกว่าเฉพาะรูปที่เป็นต้นฉบับในการค้นหา เช่น รูปภาพต้นฉบับที่เราต้องการค้นหาเป็นรูปดอกไม้ฉะนั้นรูปภาพผลลัพธ์ที่ควรจะออกมานั้นก็น่าจะเป็นรูปดอกไม้ทั้งหมดในฐานข้อมูลซึ่งอาจจะมีมากว่า 1 รูป เป็นต้น หรือในกรณีที่เราต้องการรูปภาพชนิดหนึ่งแต่ว่าไม่สามารถหารูปภาพที่จะมาเป็นต้นฉบับได้ เช่น บอกว่าต้องการค้นหารูปภาพดอกไม้ แต่ไม่มีรูปภาพต้นฉบับเป็นดอกไม้ที่ต้องการ เป็นต้น จึงมีความจำเป็นต้องบอกแค่ลักษณะของวัตถุที่อยู่ในภาพหรือสีสันในรูปของรูปภาพนั้น แล้วทำการค้นหาจากลักษณะดังกล่าว จากกรณีที่กล่าวมาข้างต้น เราจำเป็นต้องมีวิธีในการเรียนรู้ (Knowledge Base) และวิเคราะห์รูปภาพลักษณะต่างๆ ว่าเป็นรูปภาพอะไร โดยอาจจะดูลักษณะของวัตถุภายในรูปภาพ , ลักษณะลายเส้น หรือสีสันของรูปภาพนั้น

การวิเคราะห์รูปภาพแต่ละรูปนั้นจะมีหลายวิธี ซึ่งแต่ละวิธีก็จะมีประสิทธิภาพกับรูปภาพบางประเภท เช่น เราจะไม่สามารถใช้วิธีเดียวเดิมที่ให้ผลการค้นหารูปภาพพื้นผิวได้ดีไปใช้การการค้นหารูปภาพตึกต่างๆ เป็นต้น ดังนั้นเราควรจะเลือกวิธีที่เหมาะสมกับสิ่งที่เราสนใจ เช่น ในกรณีที่เราสนใจที่จะค้นหารูปตึกต่างๆ เราอาจใช้วิธีวิเคราะห์ลายเส้นที่มีลักษณะเป็นเส้นในแนวตั้ง กับแนวนอน ซึ่งเป็นลักษณะส่วนใหญ่ของตึก หรือในบางกรณีเราอาจจำเป็นต้องใช้หลายๆ วิธีในการช่วยกันค้นหาซึ่งอาจจะให้ผลที่ดีกว่า เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การตัดสินใจเลือกวิธีการค้นหารูปภาพที่ผิดอาจทำให้ผลการค้นหานั้นมีความผิดพลาด ผลจากการทดลองนี้จะช่วยให้สามารถตัดสินใจเลือกวิธีที่ใช้ให้เหมาะสมกับรูปภาพต่างๆ ที่ต้องการได้ดียิ่งขึ้น แต่นอกเหนือจากดูลักษณะของทั้งภาพแล้ว เราอาจจะดูลักษณะภายในรูปภาพที่เป็นวัตถุด้วย เช่น พิจาณาลักษณะของวัตถุ , สีสันของวัตถุ หรือจำนวนวัตถุในรูปภาพ เป็นต้น ซึ่งก็จะได้ผลลัพธ์ที่แตกต่างกันออกไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] Tha Bounthanh , "Hierachical search by region of interest for contentbased image indexing and retrieval" , p.4 - p.8 , 2003

[2] Gonzalez and Woods , Digital Image Processing 2nd Edition , Prentice Hall , 2002

[3] ยุทธนา ลีลาวัฒนกุล , Visual C++ .NET Episode One , ดวงกมลสมัย D.K.TODAY , p.151 – p.187 , 2003

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้