สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การสืบค้นข้อมูลแบบเป็นลำดับชั้นในพื้นที่ที่สนใจและการทำดัชนีสำหรับระบบค้นคืนข้อมูลรูปภาพ

HIERACHICAL SEARCH BY REGION OF INTEREST FOR CONTENT-BASED IMAGE INDEXING AND RETRIEVAL

T 0 4 9 2 9 6

ทา บุนทัน

THA BOUNTHANH

เลขหมู่..............................
เลขทะเบียน 49296
วัน, เดือน, ปี 19 ก.พ. 2547

.b..............................
.i..............................

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์
บัณฑิตวิทยาลัย
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2546

ISBN 974-324-841-2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

HIERACHICAL SEARCH BY REGION OF INTEREST FOR CONTENT-BASED IMAGE INDEXING AND RETRIEVAL

THA BOUNTHANH

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR THE DEGREE OF

MASTER OF ENGINEERING IN COMPUTER ENGINEERING

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2003

ISBN 974-324-841-2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2003

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การสืบค้นข้อมูลแบบเป็นลำดับชั้นในพื้นที่ที่สนใจและการทำดัชนีสำหรับระบบค้นคืนข้อมูลรูปภาพ
นักศึกษา	นายทา บุนทัน
รหัสประจำตัว	44612803
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
พ.ศ.	2546
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รศ. สมศักดิ์ มิตะถา
อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม	ดร.อรฉัตร จิตต์โสภักตร์

บทคัดย่อ

วิทยานิพนธ์นี้นำเสนอวิธีการค้นคืนข้อมูลรูปภาพแบบเป็นลำดับชั้นของพื้นที่ที่สนใจและการดัชนีรูปภาพ ในการทำดัชนีรูปภาพ ทุกรูปภาพในฐานข้อมูลจะถูกแบ่งส่วนให้เป็นกลุ่มๆ โดยใช้เทคนิก Vector quantization (VQ). ลักษณะสำคัญของสี(color Histogram)และความสัมพันธ์เชิงพื้นที่ระหว่างคู่สี(Autocorrelogram)ในกลุ่มจะถูกนำมาวิเคราะห์แล้วบันทึกลงในฐานข้อมูลเพื่อเป็นคีย์ในการค้นหารูปภาพ ระบบค้นคืนข้อมูลรูปภาพที่นำเสนอนี้ ผู้ใช้สามารถเลือกพื้นที่ที่สนใจได้จากรูปภาพตัวอย่างและขบวนการสืบค้นจะนำเอาเฉพาะกลุ่มที่ถูกเลือกเท่านั้น ไปเปรียบเทียบความเหมือนกับกลุ่มรูปภาพที่อยู่ในฐานข้อมูล การค้นคืนข้อมูลรูปภาพแบบเป็นลำดับชั้นถูกนำมาใช้ในการค้นคืนรูปภาพที่มีความคล้ายคลึง การค้นหารูปภาพมีสองขั้นตอนการทำงาน ขั้นตอนแรก กลุ่มที่ใหญ่ที่สุดและอยู่บริเวณตรงกลางของพื้นที่ที่ผู้ใช้เลือกจะถูกนำมาใช้ในการค้นหารูปภาพก่อน ขั้นตอนที่สองคือนำเอากลุ่มที่เหลือจากผู้ใช้เลือกมาค้นหารูปภาพในชุดรูปภาพที่ได้มาจากขั้นตอนแรก ในการค้นคืนข้อมูลรูปภาพแบบเป็นลำดับชั้นของพื้นที่ที่สนใจทำให้ลดเวลาในการค้นหาและจากผลการทดลองเห็นว่าการสืบคืนข้อมูลรูปภาพตามพื้นที่ที่สนใจมีประสิทธิภาพในการค้นคืนรูปภาพที่มีสิ่งแวดล้อมแตกต่างกันและรูปภาพที่มีการเปลี่ยนแปลง เช่น การย่อขยายรูปภาพ การหมุนรูปภาพและ การปรากฏของพื้นที่ที่สนใจเพียงบางส่วน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	Hierachical Search by Region of Interest for Content-based Image Indexing and Retrieval
Student	Mr. Tha BOUNTHANH
Student ID.	44612803
Degree	Master Degree
Programme	Computer Engineering
Year	2003
Thesis Advisor	Assoc.Prof. Somsak Mitatha
Thesis Co-Advisor	Dr. Orachat Chitsobhuk

ABSTRACT

In this thesis, we present a Content-Based Image Retrieval (CBIR) system using hierarchical Region of Interest (ROI) query and indexing. During indexing process, all the images in the database are segmented using Vector Quantization (VQ) technique. Color features of color histogram and color correlogram analysis are then extracted from each segmented region and stored in the database as the keys to retrieve the relevant images. In the proposed retrieval system, users are allowed to select ROI directly over the sample or user's submission image and the query process then focuses on the content of the selected ROI in order to find those images containing similar regions from the database. The hierarchical region-of-interest query is performed to retrieve the similar images. Two-level search is exploited in this paper. In the first level, the most important regions, usually the large regions at the center of user's query, are used to retrieve images having similar regions using static search. This ensures that we can retrieve all the images having the most important regions. In the second level, all the remaining regions in user's query are used to search from all the retrieved images obtained from the first level. By using the proposed hierarchical ROI query, the searching time is reduced and the experimental results show good retrieval performance even in the case where the query ROI is in different scales, points of view, and background or partially appears in the indexing images.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ประสบความสำเร็จลุล่วงได้เป็นอย่างดี ด้วยคำแนะนำและคำปรึกษาจาก รศ.สมศักดิ์ มิตะถา และ ดร.อรณัตร จิตต์โสภักตร์ ซึ่งเป็นอาจารย์ผู้ควบคุมวิทยานิพนธ์และ อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม ที่ให้ความเอาใจใส่และ ให้ความช่วยเหลือเสมอมา ผู้วิจัยรู้สึกทราบซึ้งในความอนุเคราะห์จากท่านและ ขอขอบพระคุณเป็นอย่างสูง

ขอขอบคุณอาจารย์ทุกๆท่านและเพื่อนทุกคนที่ช่วยเหลือให้คำแนะนำต่างๆในการทำวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงด้วยดีและ สุดท้ายขอขอบคุณโครงการ AUN/SEED-Net และ JICA ที่ให้การสนับสนุนเสมอตลอดมา ขอขอบคุณเป็นอย่างยิ่ง

ทา บุนทัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

บทคัดย่อยภาษาไทย I
บทคัดย่อภาษาอังกฤษ II
กิตติกรรมประกาศ III
สารบัญ IV
สารบัญตาราง VI
สารบัญรูป VII
บทที่ 1 บทนำ 1
1.1 ความสำคัญและที่มา 1
1.2 วัตถุประสงค์ของงานวิจัย 2
1.3 ขอบเขตของงานวิจัย 2
1.4 เนื้อหาในวิทยานิพนธ์ 2

บทที่ 2 ข้อมูลภาพดิจิตอล (Digital Image Data) 4
2.1 บทนำ 4
2.2 การแสดงภาพดิจิตอล (Digital image representation) 4
2.3 โมเดลของสี (Color models) 5
2.4 โมเดลสีแบบอาร์จีบี (RGB color model) 5
2.5 การแสดงภาพสีบนจอมอนิเตอร์ของคอมพิวเตอร์ 6
2.6 ฮิสโตแกรมของสี (Color histogram) 8

บทที่ 3 การแบ่งกลุ่มข้อมูลรูปภาพ 10
3.1 บทนำ 10
3.2 การทำภาพให้ราบเรียบ 10
3.2.1 การทำภาพให้ราบเรียบด้วยค่าเฉลี่ย (Mean value smoothing) 10
3.2.2 การทำภาพให้ราบเรียบด้วยเทคนิค peer group filter 12
3.3 การจัดระดับสัญญาณหรือการควอนไตเซชั่น (Quantization) 15
3.3.1 การควอนไตซ์สัญญาณสเกลลาร์ (Scalar Quantization) 16
3.3.2 การควอนไตซ์สัญญาณเวคเตอร์ (Vector Quantization : VQ) 20

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

3.4 การนำข้อมูลจากการคลัสเตอร์แบบ general Lloyd algorithm (GLA) มาใช้ในการแบ่งกลุ่มรูปภาพด้วยวิธีเวคเตอร์ควอนไตเซชั่น 23

บทที่ 4 การสร้างดัชนีรูปภาพและการค้นคืนรูปภาพจากฐานข้อมูล 26
4.1 บทนำ 26
4.2 วิธีการสร้าง index สำหรับแต่ละรูปภาพในฐานข้อมูล 26
4.2.1 การแบ่งกลุ่มข้อมูลรูปภาพ 27
4.2.2 การคำนวณหาคุณลักษณะสำคัญของข้อมูลรูปภาพ 31
4.3 วิธีการค้นคืนข้อมูลภาพ 32
4.3.1 วิธีการวัดความแตกต่างของรูปภาพต้นฉบับและรูปภาพในฐานข้อมูล 33
4.3.2 วิธีการค้นหารูปภาพจากฐานข้อมูล 34

บทที่ 5 การทดลองและผลการทดลอง 36
5.1 บทนำ 36
5.2 การทดลองการค้นหาและผลการทดลอง 36
5.3 สรุปผลการทดลอง 51

บทที่ 6 สรุปและวิจารณ์ 52
6.1 บทนำ 52
6.2 สรุปและวิจารณ์ 52
6.2.1 การตีกรอบเลือกพื้นที่ที่สนใจ 52
6.2.2 ขนาดของฐานข้อมูล 53
6.2.3 ขนาดของภาพ 54
6.3 แนวทางในการพัฒนาต่อ 54

เอกสารอ้างอิง 55
ภาคผนวก ก ผลการทดลอง 56
ภาคผนวก ข บทความวิจัยที่ได้รับการตีพิมพ์ 68
ประวัติผู้เขียน 74

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ หน้า

3.1 The Generalize Lloyd Algorithm 23

3.2 แสดง Agglomerative Clustering Algorithm 25

4.1 แสดงผลการทดสอบโปรแกรมเพื่อหาค่าระดับสีที่ดีที่สุดสำหรับกำหนดค่าเริ่มต้น 29

5.1 แสดงผลการทดสอบโปรแกรมเพื่อหาค่าความต่างโดยรวมที่ดีที่สุด 37

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่		หน้า
2.1	แสดงภาพดิจิตอลแบบ Gray Scale	5
2.2	แสดงรูปแบบจำลองสีแบบอาร์จีบี (RGB color model)	6
2.3	แสดงไดอะแกรมของฮีสโตแกรมสามมิติของการแสดงภาพสีโดยการใช้รูปแบบจำลองสีแบบอาร์จีบีซึ่งมีขนาด 24 บิตต่อจุดภาพ	7
2.4	แสดงจุดภาพประกอบด้วยข้อมูลสีแดง สีเขียว และ สีน้ำเงิน	7
2.5	การแสดงฮีสโตแกรมของภาพของภาพสี	8
2.6	การแสดงกราฟฮีสโตแกรมของภาพที่คล้าย กัน แต่รูปภาพทั้งสองต่างกัน	9
3.1	แสดงการใช้เทมเพลทขนาด 3x3 ในการทำให้ภาพราบเรียบด้วยค่าเฉลี่ย	11
3.2	แสดงภาพผลลัพธ์จากการทำภาพให้ราบเรียบด้วยค่าเฉลี่ย	11
3.3	แสดงภาพผลลัพธ์จากการทำภาพให้ราบเรียบด้วย PGF	15
3.4	การแบ่งช่วงค่าของข้อมูลสำหรับการควอนไตซ์สัญญาณสเกลลาร์	16
3.5	ควอนไตเซอร์เชิงเส้นแบบเป็นขั้นบันได	17
3.6	ควอนไตเซอร์ขั้นบันไดแบบไม่เป็นเชิงเส้น	17
3.7	การแปลงสัญญาณอนาล็อกเป็นดิจิตอลด้วยวิธี PCM	18
3.8	ความผิดพลาดที่เกิดขึ้นจากกระบวนการควอนไตซิ่ง	19
3.9	แสดงการคำนวณค่าเวคเตอร์อ้างอิงจากการวิเคราะห์คลัสเตอร์	21
3.10	การเข้ารหัสและการถอดรหัสเวคเตอร์	22
4.1	ขบวนการสร้างดัชนี รูปภาพ	26
4.2.	ขบวนการแบ่งกลุ่มรูปภาพ	27
4.3	โฟล์วไดอะแกรม (Flow Diagram) ของ GLA	28
4.4	(a) (b) (c) คือรูปภาพที่กำหนดค่าเริ่มต้นด้วยการสุ่มตัวอย่าง และ (e) (f) (g) คือรูปภาพที่กำหนดค่าเริ่มต้นด้วยการกำหนดระดับสี	29
4.5	ตัวอย่างการแบ่งกลุ่มรูปภาพ	30
4.6	รูปภาพตัวอย่างและออร์โตโครรีโลแกรมของรูปภาพตัวอย่าง	32
4.7.	ขบวนการค้นคืนข้อมูลรูปภาพ	33
4.8	ผลลัพธ์สืบค้นด้วยการวิเคราะห์แบบเป็น static ในพื้นที่ที่สนใจ	34
4.9	ผลลัพธ์สืบค้นด้วยการวิเคราะห์แบบเป็นลำดับชั้นของพื้นที่ที่สนใจ	34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป(ต่อ)

4.10 การสืบค้นด้วยการวิเคราะห์ แบบผสมระหว่างการสืบค้นข้อมูลแบบเป็นลำดับชั้นและการสืบค้นข้อมูลแบบ static 35
5.1 ตัวอย่างการให้ข้อมูลสำหรับ Precision และ Recall 38
5.2 ผลการค้นคืนรูปภาพชุดที่ 1 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 40
5.3 ผลการค้นคืนรูปภาพชุดที่ 2 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 41
5.4 ผลการค้นคืนรูปภาพชุดที่ 3 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 42
5.5 ผลการค้นคืนรูปภาพชุดที่ 4 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 44
5.6 ผลการค้นคืนรูปภาพชุดที่ 5 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 45
5.7 ผลการค้นคืนรูปภาพชุดที่ 6 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 47
5.8 ผลการค้นคืนรูปภาพชุดที่ 7 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 48
5.9 ผลการเปรียบเทียบประสิทธิภาพการค้นคืนภาพทั้งหมด 49
5.10 ผลการเปรียบเทียบเวลาการค้นคืนภาพด้วยเทคนิคที่นำเสนอกับแบบ Static 50
6.1 ตัวอย่างในการคิวรี่รูปภาพที่มีการตีกรอบเลือกพื้นที่ที่สนใจในรูปภาพตัวอย่างที่แตกต่างกัน 53
6.2 ตัวอย่างผลการคิวรี่รูปภาพที่มีกลุ่มสีใกล้เคียงกัน 53
6.3 ตัวอย่างผลการแบ่งกลุ่มที่มีความแตกต่างของขนาดรูปภาพ (a) รูปภาพขนาดใหญ่ (b) รูปภาพขนาดเล็ก 54
ผก.1 ผลการค้นคืนรูปภาพชุดที่ 8 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R) (c) , Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 57
ผก.2 ผลการค้นคืนรูปภาพชุดที่ 9 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R) (c) , Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 58

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป(ต่อ)

ผก.3 ผลการค้นคืนรูปภาพชุดที่ 10 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R) (c) , Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 59
ผก.4 ผลการค้นคืนรูปภาพชุดที่ 11 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R) (c) , Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 60
ผก.5 ผลการค้นคืนรูปภาพชุดที่ 12 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R) (c) , Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 61
ผก.6 ผลการค้นคืนรูปภาพชุดที่ 13 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R) (c) , Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 62
ผก.7 ผลการค้นคืนรูปภาพชุดที่ 14 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R) (c) , Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 63
ผก.8 ผลการค้นคืนรูปภาพชุดที่ 15 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R) (c) , Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 64
ผก.9 ผลการค้นคืนรูปภาพชุดที่ 16 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R) (c) , Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 65
ผก.10 ผลการค้นคืนรูปภาพชุดที่ 17 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R) (c) , Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 66
ผก.11 ผลการค้นคืนรูปภาพชุดที่ 18 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R) (c) , Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3 67

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความสำคัญและที่มา

เนื่องจากปัจจุบันฐานข้อมูลรูปภาพดิจิตอลมีขนาดใหญ่และใช้กันอย่างกว้างขวางจึงจำเป็นต้องมีระบบค้นคืนข้อมูลรูปภาพที่มีประสิทธิภาพที่สามารถค้นคืนรูปภาพตามที่ผู้ใช้ต้องการได้ เทคนิคในระบบค้นคืนรูปภาพที่ใช้กันทั่วไปคือ Content Based Image Retrieval (CBIR) ซึ่งเป็นระบบค้นคืนข้อมูลรูปภาพจากลักษณะสำคัญต่างๆของรูปภาพได้แก่ สี (color) พื้นผิว (texture) และรูปร่าง (shape) เป็นต้น เทคนิคที่นิยมใช้ในการนำลักษณะสีของรูปภาพมาใช้สำหรับ CBIR คือ ฮิสโตแกรมของสี (color histogram) เนื่องจากเป็นลักษณะเด่นที่เห็นได้ชัดเจน ง่ายต่อการเข้าใจและเป็นวิธีที่คำนวณได้ง่าย เทคนิคของการใช้สีเป็นตัวดัชนีนี้แบ่งออกเป็น 2 วิธีคือ การใช้ global color เป็นดัชนีและการใช้ region color เป็นดัชนี ความแตกต่างของวิธีทั้งสองนี้คือ ในการใช้ดัชนีแบบ global นั้นจะใช้สีทั้งหมดที่ปรากฏในภาพในการเปรียบเทียบ ซึ่งจะมีประโยชน์สำหรับการค้นหาภาพที่ไม่เจาะจงตำแหน่งใดๆในภาพ เช่น ภาพป่าไม้ หรือ สนามฟุตบอล ในขณะที่การใช้ดัชนีแบบ region จะเป็นการเปรียบเทียบเฉพาะขอบเขตที่สนใจในภาพนั้นและใช้ประโยชน์สำหรับค้นหาส่วนของภาพที่สนใจ แต่วิธีการใช้ดัชนีแบบ region color นั้นมีความซับซ้อนกว่าแบบ global color แต่การใช้ฮิสโตแกรมของสี อย่างเดียวอาจจะทำให้ค้นคืนรูปภาพผิดพลาดได้ในกรณีที่ฐานข้อมูลมีรูปภาพมากอาจจะมีรูปภาพบางรูปที่แตกต่างกันแต่มีฮิสโตแกรมของสีคล้ายคลึงกันได้

เพื่อแก้ไขปัญหาดังกล่าวในวิทยานิพนธ์นี้ได้นำเอาวิธีการทำดัชนีแบบ region มาใช้ประโยชน์สำหรับค้นหาส่วนของภาพที่สนใจและภาพที่มีสิ่งแวดล้อมแตกต่าง โดยนำเอาลักษณะสีของภาพจากการทำฮิสโตแกรมของสี (color histogram) เฉพาะแต่ละขอบเขต เนื่องจากว่า เป็นวิธีที่เข้าใจง่ายและสามารถนำไปพัฒนาร่วมกับเทคนิคอื่นๆเพื่อปรับปรุงประสิทธิภาพให้ดีขึ้นได้ ฮิสโตแกรมของสีจะมีความใกล้เคียงกันแม้ว่าวัตถุในภาพจะมีการเปลี่ยนมุมมอง การหมุน การย่อขยายภาพ เราสามารถนำความสัมพันธ์เชิงพื้นที่ระหว่างคู่สี(color correlogram) มาใช้เพื่อแยกความแตกต่างระหว่างกลุ่มต่างๆ ที่มีฮิสโตแกรมของสีใกล้เคียงกันแต่มีการกระจายของตำแหน่งพิกเซลไม่เหมือนกัน เพื่อแก้ไขปัญหาในค้นคืนข้อมูลผิดพลาดได้เนื่องจากรูปภาพบางรูปที่แตกต่างกันอาจจะมีโทนสีคล้ายคลึงกันได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 วัตถุประสงค์ของงานวิจัย

1.2.1 ศึกษาถึงความสามารถและประสิทธิภาพการค้นหารูปภาพโดยใช้สีเป็นดัชนีให้กับภาพ

1.2.2 ศึกษาข้อจำกัดของแนวคิดและการแก้ไขเพื่อให้การกำหนดดัชนีโดยใช้สีของรูปภาพมีประสิทธิภาพมากที่สุด

1.2.3 ศึกษาแนวทางในการพัฒนาโดยใช้วิธี region color ร่วมกับออร์โตโครรีโลแกรม เพื่อลดข้อเสียเปรียบของวิธีใช้สีเพียงอย่างเดียว

1.2.4 ศึกษาข้อจำกัดและประยุกต์ใช้งานเวคเตอร์ควอนไตเซชั่นในการทำการแบ่งกลุ่มรูปภาพให้มีประสิทธิภาพมากที่สุด

1.2.5 ศึกษาแนวทางในการพัฒนาในการเพิ่มประสิทธิภาพการค้นหารูปภาพโดยใช้วิธีแบบสถิติและแบบเป็นลำดับชั้น

1.3 ขอบเขตของงานวิจัย

เป็นการนำเสนอวิธีการสร้างดัชนีรูปภาพและการค้นคืนข้อมูลรูปภาพแบบเป็นลำดับชั้นตามพื้นที่ที่สนใจ โดยการทำดัชนีรูปภาพที่นำเสนอนี้ รูปภาพแต่ละรูปถูกแบ่งส่วนให้เป็นจำนวน N กลุ่ม ก่อนขบวนการแบ่งสีรูปภาพให้เป็นจำนวน N กลุ่ม เราทำการกำจัด Noise และทำให้ภาพเรียบก่อน จากนั้นก็ดึงลักษณะสำคัญ (feature vector) ของกลุ่มย่อยแล้วนำไปเก็บไว้ในฐานข้อมูลเพื่อเป็น key ในการค้นหารูปภาพที่ต้องการ โดยการค้นหาข้อมูลรูปภาพที่นำเสนอในวิทยานิพนธ์นี้ ผู้ใช้สามารถกำหนดพื้นที่ที่ต้องการได้โดยตีกรอบสี่เหลี่ยมบนรูปภาพตัวอย่าง กลุ่มที่ผู้ใช้เลือกจะถูกนำมาเปรียบเทียบความเหมือน กับกลุ่มย่อยของรูปภาพที่เป็นเป้าหมายในฐานข้อมูล โดยใช้วิธีการค้นหาข้อมูลแบบผสมระหว่างการสืบ ค้นข้อมูลแบบเป็นลำดับชั้นและการสืบค้นข้อมูลแบบสถิติ

1.4 เนื้อหาในวิทยานิพนธ์

ในวิทยานิพนธ์นี้ ได้นำเสนอวิธีการสร้างดัชนีรูปภาพและการค้นคืนข้อมูลรูปภาพแบบเป็นลำดับชั้นตามพื้นที่ที่สนใจ โดยมีรายละเอียดทั้งหมดแบ่งเป็น 8 บท ดังต่อไปนี้

บทที่ 1 เป็นการกล่าวถึงความเป็นมาและความสำคัญของปัญหา แนวคิดที่ใช้ในการแก้ปัญหา วัตถุประสงค์และขอบเขตของการทำวิจัย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2 เป็นการกล่าวถึงลักษณะสำคัญของข้อมูลภาพดิจิตอล การแสดงภาพดิจิตอลรูปแบบจำลองของสี ฮิสโตแกรมของสีภาพ(color Histogram) และความสัมพันธ์เชิงพื้นที่ระหว่างคู่สี(color correlogram)

บทที่ 3 เป็นการกล่าวถึงการแบ่งกลุ่มข้อมูลรูปภาพ รวมถึงการทำภาพให้ราบเรียบ การจัดระดับสัญญาณหรือการควอนไตเซชั่น และการประยุกต์ general Lloyd algorithm (GLA) มาใช้ในการแบ่งกลุ่มรูปภาพด้ยวิธีเวคเตอร์ควอนไตเซชั่น

บทที่ 4 เป็นการกล่าวถึงวิธีการในการค้นหารูปภาพโดยใช้วิธีการค้นหาข้อมูลแบบผสมระหว่างการสืบค้นข้อมูลแบบเป็นลำดับชั้นและการสืบค้นข้อมูลแบบสถิติและกล่าวถึงวิธีการวัดความแตกต่างแบบ L1 ประยุกต์(modified L1 distance measure) สำหรับวัด ความแตกต่างของฮีสโตแกรมสี และออร์โตโครรีโลแกรม

บทที่ 5 เป็นการกล่าวถึงวิธีการทดลองการค้นหาข้อมูลรูปภาพและนำเสนอผลการทดลองพร้อมทั้งการวัดประสิทธิภาพของการค้นคืนข้อมูลรูปภาพ

บทที่ 6 เป็นการสรุปและวิจารณ์ และแนวทางในการพัฒนาต่อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ข้อมูลภาพดิจิตอล
(Digital Image Data)

2.1 คำนำ

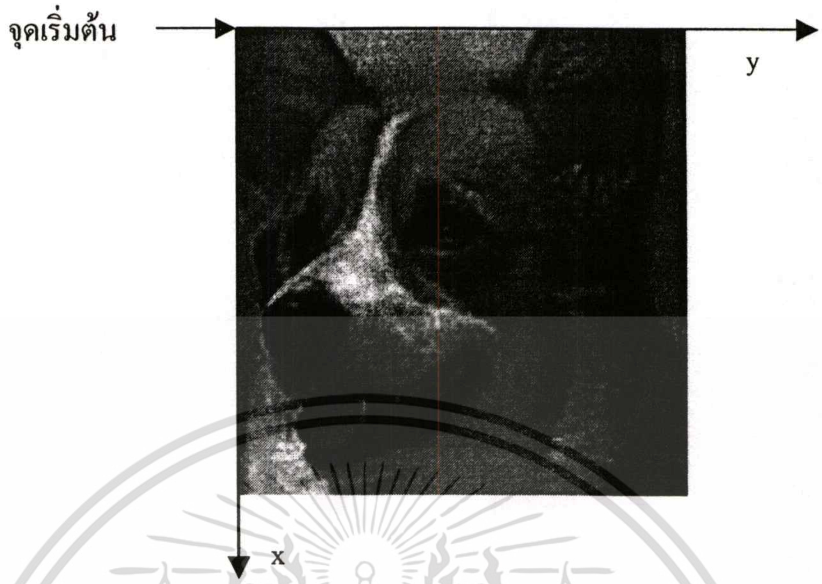
เนื้อหาภายในบทนี้จะเป็นการกล่าวถึงลักษณะของข้อมูลภาพดิจิตอล ซึ่งจะมีการแสดงภาพดิจิตอล ที่เป็นภาพขาวดำ ภาพสี รายละเอียดของภาพ (Resolution) จำนวนระดับสีเทาของภาพ ฮิสโตแกรมของภาพ โดยปกติการแปลความหมายจากภาพสีนั้นมีข้อได้เปรียบกว่าการแปลความหมายจากภาพขาวดำ เนื่องจากว่าในการวิเคราะห์ภาพที่ได้จากภาพสีนั้นสีจะเป็นตัวช่วยแยกแยะวัตถุ (Object Extraction) ที่สนใจออกมาจากฉากเบื้องหลังได้ดี ทำให้การวิเคราะห์หรือการแปลความหมายโดยการใช้ภาพสีมีประสิทธิภาพสูง

2.2 การแสดงภาพดิจิตอล (Digital image representation) [13]

ข้อมูลภาพดิจิตอลจะมีลักษณะเป็นแถวลำดับ (Array) แต่ละจุดภาพจะอยู่ในรูปสี่เหลี่ยมเล็กๆ แต่ละจุดภาพจะมีตำแหน่งบอกพิกัดสปาเชียล (Spatial coordinate) ในรูปคาร์ติเซียน (Cartesian) เป็น (x,y) เมื่อ x เป็นลำดับตำแหน่งในแนวนอน และ y เป็นลำดับตำแหน่งในแนวตั้ง ปกติแล้วตำแหน่งจุดภาพ (0,0) จะอยู่ทางด้านซ้ายมือสุดด้านบนของภาพ การจัดลำดับตำแหน่งของจุดภาพจะเรียงจากซ้ายไปขวาในแต่ละเส้นภาพและจัดลำดับของเส้นภาพจะเรียงจากบนลงล่าง ถ้าหากว่าเขียนอยู่ในรูปของฟังก์ชั่น f(x,y) จะทำให้ทราบว่าค่าของฟังก์ชั่นจะเป็นระดับความสว่างของจุดภาพดังกล่าว ซึ่งบางครั้งจะเรียกว่าค่าระดับสีเทา (Brightness or gray level) ในแต่ละภาพจะมีจำนวนระดับสีเทากี่ระดับขึ้นอยู่กับจำนวนบิตที่นำมาเข้ารหัสว่ามีกี่บิต (Bit) ในบางครั้งถ้ากำหนดให้ฟังก์ชั่นหรือค่าระดับสีเทาอยู่ในแกน z ก็จะสามารถพล็อต (Plot) ข้อมูลภาพดิจิตอลเป็นรูปสามมิติได้ แต่การมองภาพเป็นการมองในแถวตั้งฉาก ดังนั้นค่าความสูงของระดับสีเทาในแกน z จึงถูกปรับเปลี่ยนให้อยู่ในความขาวดำของจุดภาพในสองมิติแทนดังแสดงในรูปที่ 2.1 หากจุดภาพที่อยู่ประชิดติดกันมีค่าความแตกต่างของระดับสีเทาก็จะเกิดเป็นขอบของวัตถุในภาพขึ้น การเปลี่ยนแปลงค่าระดับสีเทาของจุดภาพต่างๆ นี้เองจะทำให้ผู้มองสามารถแยกแยะรายละเอียดของภาพหรือสามารถจำแนก (Classify) วัตถุต่างๆ ในภาพได้ เนื่องจากตำแหน่งโคออดิเนท (x,y) จะเป็นเลขจำนวนเต็มหน่วย (Integer) ดังนั้นแต่ละตำแหน่งโคออดิเนทจึงเป็นจุดภาพที่เรียกว่า พิกเซล (Pixel หรือ Picture element) โดยขนาดของจุดภาพขึ้นอยู่กับค่า x และ y ที่จะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบ่งไป ตัวอย่างเช่น รูปที่ 2.1 ซึ่งเป็นภาพขนาด 256x256 จุดภาพและค่าความแตกต่างของความสว่างหรือระดับสีเทาเป็น 256 ระดับ



รูปที่ 2.1 แสดงภาพดิจิตอลแบบ Gray Scale

2.3 โมเดลของสี (Color models)

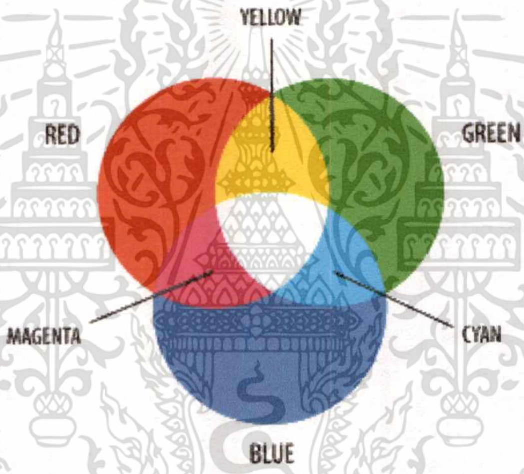
การแสดงภาพสีมีลักษณะเป็นรูประบบพิกัดแบบสามมิติ แต่ละสีเกิดจากการผสมระหว่างแกนทั้งสาม โดยโมเดลของสีในปัจจุบันมีการหันเหไปตามฮาร์แวร์ เช่น จอภาพสี เครื่องพิมพ์และการประยุกต์ใช้งานในด้านต่างๆ โมเดลของสีที่นิยมใช้กันมากสำหรับจอมอนิเตอร์สี และกล้องวีดีโอ RGB model (Red, Green, Blue) สำหรับโมเดลของสีแบบ CMY model (Cyan, Magenta, Yellow) นิยมใช้กับเครื่องพิมพ์สี ส่วนโมเดลของสีแบบ YIQ model เป็นมาตรา ฐานสำหรับการส่งกระจายภาพโทรทัศน์สี โดยที่ Y คือ luminance และ I กับ Q คือ ส่วนของ chromatic ที่เป็น inphase และ Quadrature ตามลำดับโมเดลของสีที่นิยมใช้กันมากสำหรับการประมวลผลภาพ (Image processing) นั่นคือ RGB model และ YIQ model สำหรับในบทนี้จะอธิบายเฉพาะโมเดลของสีแบบ RGB model เนื่องจากเป็นโมเดลของสีที่ใช้ในงานวิจัยที่เสนอนี้

2.4 โมเดลของสีแบบอาร์จีบี (RGB color model)

ลำของแสงสีขาวเมื่อผ่านแก้วปริซึมแล้วแสงที่ออกมาไม่ได้มีสีขาวเพียงสีเดียว แต่ประกอบไปด้วยสเปกตรัมที่ต่อเนื่องของสีมีย่านจากสีม่วง (Violet) ไปจนถึงสีแดง (Red) โดยพื้นฐานแล้วสีที่มนุษย์สามารถมองเห็นจากวัตถุนั้นเกิดจากธรรมชาติของแสงในการสะท้อนของวัตถุ ซึ่งแสงทั้งหมดทุกแสงจะถูกมองเห็นเป็นสีขาว วัตถุที่สะท้อนบางส่วนของสเปกตรัม (Spectrum) จะให้บาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เฉดของสี อย่างเช่น วัตถุที่มีสีเขียวจะสะท้อนช่วงคลื่นที่มีความยาว 500 ถึง 570 นาโนเมตร โดยการดูดซับเอาความยาวคลื่นที่เหลือนอกนั้นไว้หมดทำให้เราสามารถมองเห็นวัตถุชิ้นนั้นมีสีเขียวนั้นเอง ในย่านที่มนุษย์สามารถมองเห็นได้นั้นความยาวคลื่นของแสงจะเป็นตัวกำหนดสีที่อยู่ในหน่วยนาโนเมตร (nm) ในระบบโมเดลของแบบอาร์จีบี (RGB color model) นั้นมีสีหลักๆ อยู่ 3 สี คือ สีแดง (red color) สีเขียว(Green color) และสีน้ำเงิน(Blue color) จะถูกเรียกว่า สีปฐมภูมิ (Primary color) เมื่อมีการผสมของแสงสีแดง สีเขียว และน้ำเงินเข้าด้วยกันก็จะได้สีต่างๆ ดังแสดงในรูปที่ 2.2 ซึ่งจะสังเกตเห็นว่า สีน้ำเงินผสมสีแดงได้สีม่วง สีน้ำเงินผสมสีเขียวได้สีฟ้าแกมเขียว สีแดงผสมสีเขียวได้สีเหลือง วิธีการนี้เป็นวิธีที่มีประสิทธิภาพดีเพียงพอในการแสดงผลภาพและมีความนิยมใช้กันมาก เช่น การแสดงผลในมอนิเตอร์ (Monitor) สีสำหรับระบบคอมพิวเตอร์ทั่วไปเป็นต้น



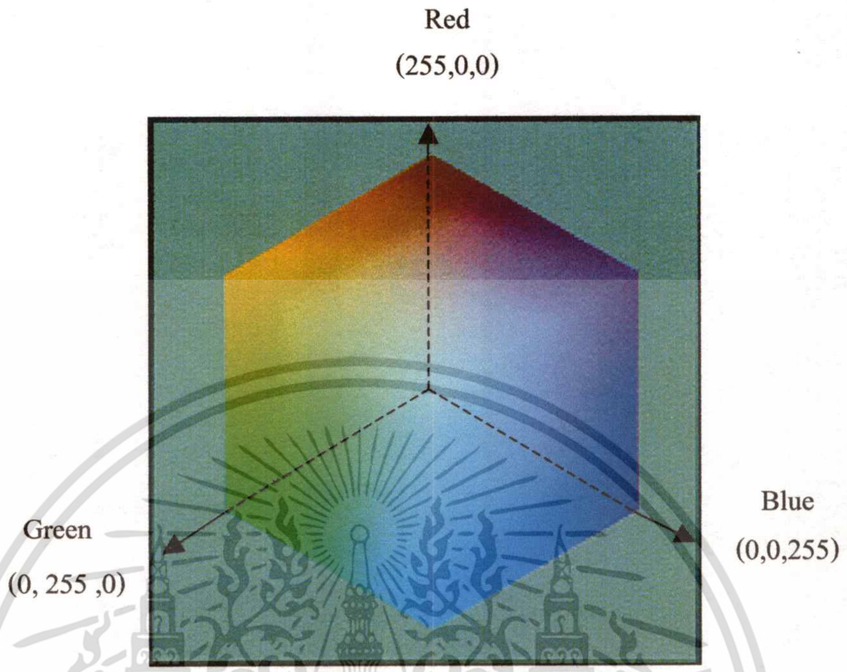
รูปที่ 2.2 แสดงโมเดลของสีแบบอาร์จีบี (RGB color model)

2.5 การแสดงภาพสีบนจอมอนิเตอร์คอมพิวเตอร์

ข้อมูลภาพสีอยู่ในรูปแบบของ RGB จะเป็นรูปแบบพื้นฐานของภาพที่มีการแสดงบนจอคอมพิวเตอร์โดยใช้ 1 พิกเซลของภาพประกอบด้วยองค์ประกอบสี 3 ค่า (channel) คือ สีแดง(red color) สีเขียว(Green color) และสีน้ำเงิน(Blue color) แต่ละ channel เป็นข้อมูลมีขนาด 8 บิต ดังนั้นข้อมูลสีหนึ่งจุดจะมีขนาด 24 บิต หรือเรียกว่าเป็นข้อมูล 24 บิตต่อหนึ่งจุดภาพ (Bit per pixel) นั่นหมายถึง แต่ละจุดภาพจะมีข้อมูลที่แตกต่างกันถึง 2^{24} ระดับหรือมีความแตกต่างกันเท่ากับ 16.7 ล้านระดับ ถ้ามีการแสดงผลภาพสีโดยการนำข้อมูลภาพทั้งสามแบนด์มาพล็อตเป็นฮีสโตแกรม (Histogram) สามมิติ ซึ่งแต่ละแกนมีค่าระดับสีเทาของข้อมูลภาพที่มีค่าได้ตั้งแต่ 0-255 ของแต่ละ channel ทำให้เกิดบล็อกสี่เหลี่ยมเล็กๆ ภายในฮีสโตแกรมดังกล่าวจำนวน 256x256x256 บล็อก

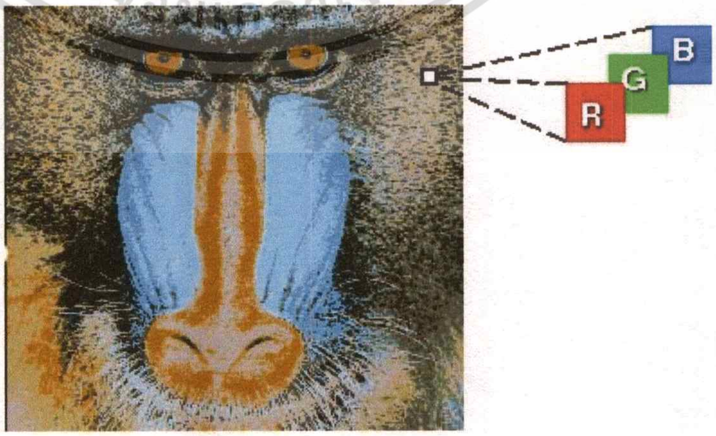
เอกสารนี้เป็นเอกสารที่เผยแพร่เพื่อการศึกษาเท่านั้น เมื่อนำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่แต่ละบล็อกก็คือข้อมูลของเฉดสีที่เกิดขึ้นในแต่ละจุดภาพนั่นเอง ฮีสโตแกรมภาพดังแสดงใน รูปที่ 2.3



รูปที่ 2.3 แสดงไดอะแกรมของฮีสโตแกรมสามมิติของการแสดงภาพสีโดยการใช้โมเดลของสีแบบ อาร์จีบีซึ่งมีขนาด 24 บิตต่อจุดภาพ

แต่ละจุดภาพจะประกอบด้วยข้อมูลสีแดง(red color) สีเขียว(Green color) และ สีน้ำเงิน (Blue color) เพื่อให้ได้ภาพสีที่สามารถแสดงรายละเอียดของทรัพยากรต่างๆ ดังแสดงในรูปที่ 2.4



รูปที่ 2.4 แสดงจุดภาพประกอบด้วยข้อมูลสีแดง สีเขียวและ สีน้ำเงิน

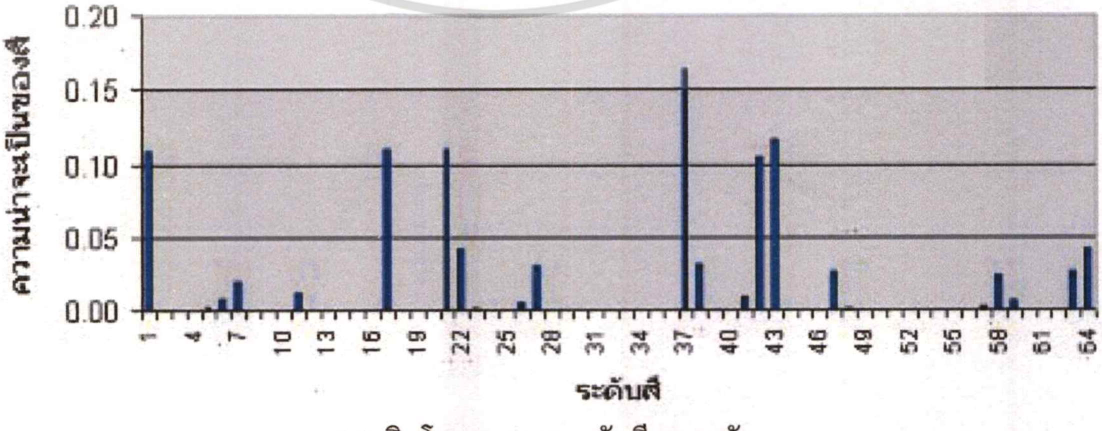
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6 ฮีสโตแกรมของภาพ (Image histogram)

ฮิสโตแกรมของสีประกอบไปด้วยระดับต่างๆ ของสี ซึ่งระดับต่างๆ นี้จะเป็นจำนวนของสีในแต่ละระดับที่ปรากฏในภาพ ฮิสโตแกรมของสีจะมีความทนทานต่อการเปลี่ยนมุมมองของภาพ เนื่องจากการเปลี่ยนมุมมองจะทำให้จำนวนพิกเซลรวมของสีไม่เกิดการเปลี่ยนแปลงมากนัก แต่ถ้ารูปร่างของภาพเมื่อมีการการเปลี่ยนมุมมองแล้วจะเกิดการเปลี่ยนแปลงไปมาก ฮิสโตแกรมจะเป็นเสมือนฟังก์ชั่นในการกำหนดความหน้าจะเป็นของสีที่ปรากฏในภาพ ซึ่งค่าของสีเดียวกันจะต้องอยู่ในระดับของสีเดียวกัน ฮิสโตแกรมของสีไม่เหมาะที่จะใช้ในการทำการจำ (recognition) เนื่องจากการจะระบุว่าภาพใดเหมือนกันนั้นจะต้องดูที่กลุ่มของสีที่ปรากฏในรูปว่าอยู่ในตำแหน่งเดียวกันหรือไม่ แต่ฮิสโตแกรมของสีไม่สามารถบอกได้ว่าระดับสีในแต่ละระดับอยู่ในตำแหน่งใดของภาพ อย่างไรก็ตาม ฮิสโตแกรมสีกลับได้รับความนิยมในการเป็นดัชนีของภาพ เพราะส่วนใหญ่แล้วพบว่า การที่สีที่ปรากฏในภาพมีความมัว เนื่องจากปัจจัยแวดล้อมต่างๆ ทำให้กลุ่มของสีที่ปรากฏในภาพเปลี่ยนตำแหน่งไป แต่ผลกระทบที่มีต่อฮิสโตแกรมของสีจะมีเพียงเล็กน้อยเท่านั้น ซึ่งทำให้การค้นหาภาพยังคงมีประสิทธิภาพที่ดีได้

ในวิทยานิพนธ์ที่ได้จัดทำนี้ การสร้างฮิสโตแกรมของสีโดยเก็บสีที่ปรากฏในภาพ โดยได้แบ่งค่า RGB ของรูปภาพออกเป็น 64 ระดับเพื่อลดความซับซ้อนและเวลาในการทำดัชนี [5]

(ก) ภาพสีที่อยู่ในแบบ RGB

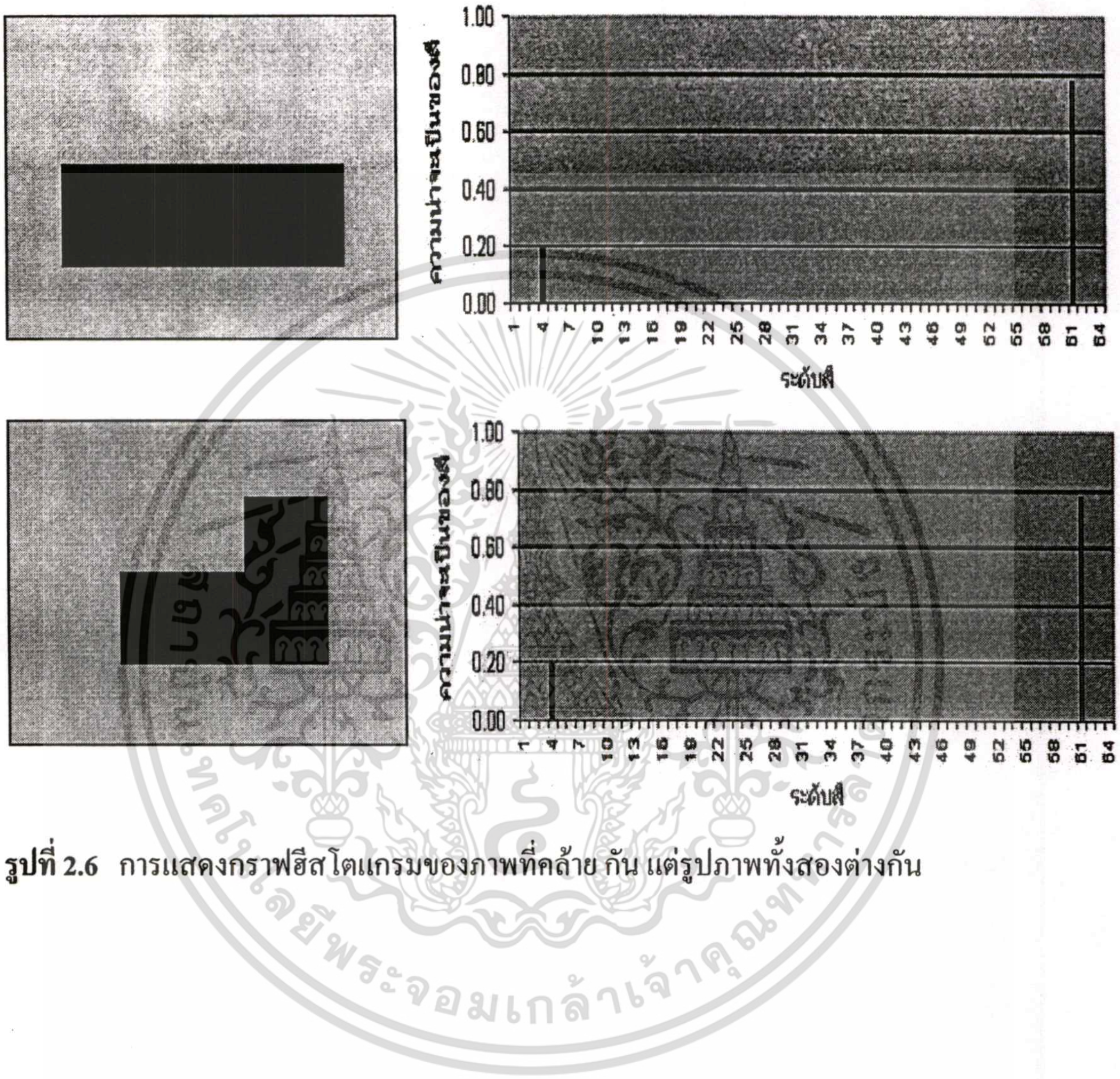


(ข) ฮิสโตแกรมของระดับสี 64 ระดับ

รูปที่ 2.5 การแสดงฮีสโตแกรมของภาพของภาพสี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับฮีสโตแกรมของภาพสองภาพที่มีรูปร่างกราฟคล้ายๆ กันไม่ได้หมายความว่าภาพทั้งสองจะคล้ายกัน ทั้งนี้เพราะฮีสโตแกรม ของภาพจะไม่เกี่ยวข้องกับตำแหน่งของจุดภาพ กล่าวคือ จุดภาพที่มีค่าระดับสีเทาเหมือนกันจากกราฟฮีสโตแกรมทั้งสองไม่ได้ปรากฏอยู่ที่ตำแหน่งเดียวกันภาพทั้งสองจึงไม่เหมือนกัน ดังแสดงในรูปที่ 2.6



รูปที่ 2.6 การแสดงกราฟฮีสโตแกรมของภาพที่คล้าย กัน แต่รูปภาพทั้งสองต่างกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การแบ่งกลุ่มข้อมูลรูปภาพ

3.1 บทนำ

การวิเคราะห์คลัสเตอร์ข้อมูลภาพ (clustering) นั้นเป็นกระบวนการหนึ่งที่สำคัญในการศึกษาด้านการประมวลผลภาพ โดยจุดมุ่งหมายของการวิเคราะห์คลัสเตอร์คือการจำแนกข้อมูลภาพอินพุทออกเป็นกลุ่มๆ โดยที่สมาชิกที่อยู่ในกลุ่มเดียวกันจะมีคุณลักษณะที่เหมือนหรือคล้ายคลึงกันข้อมูลที่ได้จากการจำแนกกลุ่มเหล่านี้สามารถนำไปใช้ประโยชน์ในการศึกษาด้านการประมวลผลภาพ เช่น สามารถนำไปใช้ช่วยในการหาขอบเขตของวัตถุภายในภาพสี วิทยานิพนธ์ที่ได้จัดทำนี้ได้นำข้อมูลจากการคลัสเตอร์ (cluster) แบบ generalized Lloyd algorithm (GLA))[7] มาใช้ในการแบ่งกลุ่มข้อมูลรูปภาพด้วยวิธีเวคเตอร์ควอนไตเซชั่น (Vector Quantization) ก่อนขบวนการแบ่งกลุ่มข้อมูลรูปภาพเราได้ใช้เทคนิค peer group filter (PGF)[6] ในการกำจัด Noise และทำให้ภาพเรียบ

3.2 การทำภาพให้ราบเรียบ (smoothing)

การทำภาพให้ราบเรียบคือการกำจัดสัญญาณรบกวนที่แปลกปลอมเข้ามาอยู่ในภาพ ซึ่งสัญญาณรบกวนดังกล่าวบางครั้งเกิดจากตัวจัดเก็บข้อมูลภาพหรือเกิดขึ้นระหว่างการส่งข้อมูลภาพผ่านระบบการสื่อสาร การกำจัดสัญญาณรบกวน และการเกลี่ยทำให้รายละเอียดภาพเรียบขึ้นนั้นเป็นขั้นตอนสำคัญสำหรับงานวิจัยด้านการประมวลผลภาพ ตัวอย่างเช่น ในขั้นตอนการควอนไตซ์สีการประมาณค่าของการเคลื่อนที่วัตถุ หรือการแยกส่วนของวัตถุในภาพ เทคนิคในการทำภาพให้ราบเรียบนั้นมีมากมายหลายวิธีเช่น การทำภาพให้ราบเรียบด้วยค่าเฉลี่ย (Mean value smoothing) และ การปรับเรียบด้วยค่ามัธยฐานหรืออาจเรียกว่าการกรองมัธยฐาน (Median Filter) แต่สิ่งที่เป็นปัญหาในการใช้วิธีการทำภาพให้ราบเรียบก็คือ การนำไปใช้อย่างขาดความระมัดระวัง โดยจะส่งผลก่อให้เกิดการเบลอ (blur) ของขอบภาพที่คมชัด

3.2.1 การทำภาพให้ราบเรียบด้วยค่าเฉลี่ย (Mean value smoothing)

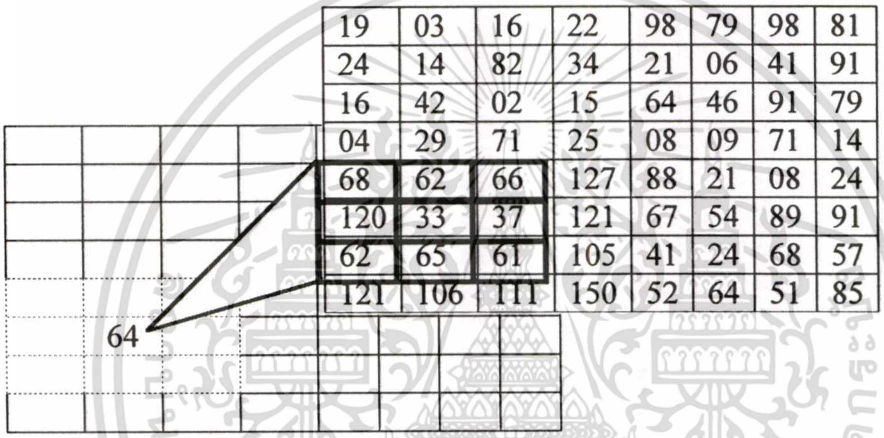
การกำจัดสัญญาณรบกวนจากการทำภาพให้ราบเรียบด้วยค่าเฉลี่ย จะเป็นการนำเอาระดับสีเทาของจุดภาพที่จุดข้างเคียงของจุดภาพที่ต้องการประมวลผลมาหาค่าเฉลี่ย เพื่อนำไปแทนค่าระดับสีเทาของจุดภาพนั้น ดังแสดงในรูปที่ 3.1 จากการนำค่าระดับสีเทาของจุดภาพข้างเคียงมาใช้ในการทำภาพให้ราบเรียบ จำนวนจุดภาพในละแวกข้างเคียงที่ใช้โดยมากมักจะเป็นเทมเพลทจัตุรัสขนาด 3x3 จุดภาพหรือ 5x5 จุดภาพ ปกติแล้วนิยมใช้เทมเพลท (Template) ขนาด 3x3 จุดภาพเพื่อลดเวลา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ห้ามมิให้นำไปใช้ประโยชน์ทางการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการคำนวณ การทำภาพให้ราบเรียบด้วยค่าเฉลี่ยจะเป็นการประมวลผลแบบกลุ่มจุดภาพ ซึ่งจะอยู่ในลักษณะการทำคอนโวลูชัน (Convolution) ระหว่างภาพกับเทมเพลทที่ใช้ เพื่อคำนวณค่าระดับสีเทาใหม่ของจุดภาพตำแหน่งเดิม ในการหาค่าเฉลี่ยนี้จะนำเทมเพลทที่ใช้ไปทาบบนภาพอินพุท หรือต้นฉบับ ซึ่งจุดกึ่งกลางของเทมเพลทจะถูกวางลงบนจุดภาพที่สนใจที่จะถูกทำการปรับเปลี่ยนค่าระดับสีเทา สมการที่ใช้สำหรับเทมเพลทขนาด 3x3 จุดภาพคือ

$$G(x,y)=\sum_{i=1}^{3}\sum_{j=1}^{3}W(i,j)I(x+2-i,y+2-j) \qquad (3.1)$$

เมื่อ W(i,j) เป็นเทมเพลทที่ใช้ I(x,y) เป็นภาพอินพุท G(x,y) เป็นภาพผลลัพธ์จากการทำภาพให้ราบเรียบด้วยค่าเฉลี่ย และ (x,y) เป็นพิกัดตำแหน่งของจุดภาพที่สนใจ



$$G(x,y)=(\frac{1}{9}\times 68)+(\frac{1}{9}\times 62)+(\frac{1}{9}\times 66)+(\frac{1}{9}\times 120)+(\frac{1}{9}\times 33)+$$
$$(\frac{1}{9}\times 37)+(\frac{1}{9}\times 62)+(\frac{1}{9}\times 65)+(\frac{1}{9}\times 61)=64$$

รูปที่ 3.1 แสดงการใช้เทมเพลทขนาด 3x3 ในการทำให้ภาพราบเรียบด้วยค่าเฉลี่ย

(a) ก่อนการทำภาพให้ราบเรียบ (b) หลังจากการทำภาพให้ราบเรียบ

รูปที่ 3.2 แสดงภาพผลลัพธ์จากการทำภาพให้ราบเรียบด้วยค่าเฉลี่ย

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำภาพให้ราบเรียบ โดยแทนค่าจุดภาพ (x,y) ด้วยค่าเฉลี่ยระดับสีเทาของกลุ่มจุดภาพข้างเคียงหรือที่เราเรียกว่าเทมเพลท โดยที่เทมเพลทที่เลือกนั้นเป็นเทมเพลทที่มีความกลมกลืนมากที่สุด แต่อย่างไรก็ตามเทมเพลทดังกล่าวก็ยังไม่สามารถให้ผลลพธ์ที่ดีได้ เมื่อนำไปใช้กับรูปภาพที่มีขอบเขตพื้นที่รูปทรงซับซ้อน เนื่องจากวิธีการทำภาพให้ราบเรียบนี้ใช้เทมเพลทรูปทรงสี่เหลียมจตุรัส ซึ่งขอบเขตพื้นที่ที่มีรูปทรงซับซ้อนเช่น รูปทรงตัววีและรูปทรงอื่นๆ เป็นต้น จะส่งผลก่อให้เกิดการรวมตัวของขอบเขตพื้นที่เข้ากันกับขอบเขตข้างเคียงที่อยู่ล้อมรอบหรือบางครั้งก็อาจจะแยกออกเป็นขอบเขตอิสระต่างหากด้วยระดับค่าสีเทาที่ผิดพลาด ภาพผลลัพธ์หลังจากการทำภาพให้ราบเรียบจะขาดความคมชัด (blur) ดังแสดงในรูปที่ 3.2 (b)

3.2.2 การทำภาพให้ราบเรียบด้วยเทคนิค peer group filtering [6]

สำหรับเทคนิคที่ง่ายๆ สำหรับการกำจัดสัญญาณรบกวน เช่น การทำภาพให้ราบเรียบด้วยค่าเฉลี่ย (Mean value smoothing) สามารถกำจัดสัญญาณรบกวนได้ แต่อย่างไรก็ตามค่าเฉลี่ยที่คำนวณได้จาก Local window จะถูกใช้ในการประมาณค่า ณ ตำแหน่งพิกเซลที่สนใจ ซึ่งอาจจะทำให้เป็นการ Blur ขอบของวัตถุในภาพได้

ในการจัดการกับข้อเสียดังกล่าวนี้ เราได้ใช้เทคนิคที่เรียกว่า Peer Group Filtering เพื่อใช้ในการกำจัด Noise ในภาพสี โดยถ้ากำหนดให้ x_0 (n) แทน Pixel vector ของภาพ ที่แสดงคุณสมบัติของข้อมูลสี ณ ตำแหน่ง n ที่เป็นจุดกึ่งกลาง (Center) ของ Window ขนาด w x w โดยทำการเรียงลำดับข้อมูลพิกเซลใน Window จากน้อยไปหามาก โดยเรียงตาม d_i(n) คือค่าผลต่าง เมื่อเทียบกับ x_0(n) โดยแทนด้วย x_i (n), $i = 0, \ldots, k = w^2 - 1$ โดยใช้วิธีการหาค่าผลต่างของ Euclidean distance measure ตัวอย่างเช่น

$$d_i(n) = \| x_0(n) - x_i(n) \|, i = 0,\ldots,k \qquad (3.2)$$

$$d_o(n) \le d_1(n) \le \ldots \le d_k(n) \qquad (3.3)$$

แล้วค่า Peer group P(n) ของ x_0 (n) ที่มีขนาด m(n) หาได้จาก

$$P(n) = \{x_i(n), i = 0,\ldots, m(n) - 1\} \qquad (3.4)$$

แนวความคิดของ Peer group filtering ถูกนำเสนอไว้ใน [6] ถูกใช้สำหรับกำจัด Noise และปรับปรุงภาพแบบ Grayscale ในงานวิจัยนี้ได้นำมาใช้ได้กับภาพสี และเพิ่มความสามารถให้ทำงานโดยอัตโนมัติในการเลือกขนาดของ m(n) สำหรับแต่ละพิกเซล โดยแนวความคิดพื้นฐานของ PGF สรุปได้เป็นสองขั้นตอนดังนี้

เอกสารนี้ใช้เพื่อการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. Classification : เป็นขั้นตอนในการแยก Peer group สำหรับแต่ละพิกเซล $x_0(n)$ ถ้าถูกตัดสินว่า $x_0(n)$ เป็นส่วนของสัญญาณรบกวนและไม่มี Peer group แล้ว Peer group ที่แท้จริงของพิกเซลตำแหน่งนั้นสามารถที่จะถูกประมาณด้วยค่าของพิกเซลอื่นๆ ที่เหลือใน window
2. Replacement : ทำการแทน $x_0(n)$ ด้วยค่าเฉลี่ยที่มีการ Weight แต่ละพิกเซลที่เป็นสมาชิกของ Peer group จะเห็นว่ามีลักษณะคล้ายกับการทำ Gaussian filtering ด้วยค่า binary mask ที่มีค่าเท่ากับ 1 เมื่อพิกเซลนั้นเป็นสมาชิกของ Peer group

จุดประสงค์ในการเฉลี่ยค่าที่เป็นสมาชิกของ Peer group แทนที่จะทำการเฉลี่ยค่าไปทั้ง Local window ก็เพื่อที่จะไม่ทำให้ขอบเกิดการ Blur มากขึ้น ในการหาค่าขนาดของ m(n) ที่เหมาะสมสำหรับแต่ละ Peer group ขึ้นอยู่กับค่าสถิติภายใน local window นั้นมีความสำคัญมากต่ออัลกอริทึมของ PGF ซึ่งวิธีการหนึ่งในการกำหนดขนาดดังกล่าวคือ ทำการกำหนดค่า threshold T(n) แล้ว m(n) หาได้จาก

$$d_{m(n)-1}(n) \le T(n) \quad and \quad d_{m(n)}(n) > T(n) \qquad (3.5)$$

อย่างไรก็ตาม เนื่องจากสถิติและ สัญญาณรบกวนสามารถที่จะทำการปรับให้เหมาะสมกับแต่ละภาพหรือภายในภาพเดียวกันก็ได้ ดังนั้นจึงเป็นการยากที่จะกำหนดค่า T(n) ที่เหมาะสมที่สุดสำหรับทุกๆ ภาพ ไว้ตายตัว

ถ้าภาพนั้นประกอบไปด้วยสีเพียงแค่ 2 กลุ่มใน window แล้วจะใช้ Fisher's linear discriminant [11] จะช่วยให้อัตราส่วนของการกระจายตัวภายใน class เดียวกัน ต่อการกระจายตัวแบบระหว่าง class มีค่าสูงสุด โดยสามารถที่จะถูกใช้เพื่อแยก 2 clusters นั้นได้ อย่างไรก็ตามถ้าหากมีสีมากกว่า 2 classes วิธีการดังกล่าวจะไม่สามารถที่จะใช้แยก cluster ที่มี $x_0(n)$ เป็น center pixel รวมถึงความซับซ้อนในการคำนวณ (computational complexity) ก็จะมีค่าสูงด้วย

วิธีการง่ายๆ สำหรับจัดการกับปัญหาดังกล่าวนี้คือใช้เฉพาะ ค่าผลต่าง 1 มิติ $d_i(n)$ สำหรับการประมาณค่าของ Fisher's discriminant โดยสามารถหาค่าได้ดังนี้

$$J(i) = \frac{|a_1(i) - a_2(i)|^2}{s_1^{\,2}(i) + s_2^{\,2}(i)}, i = 1,\ldots,k \qquad (3.6)$$

โดยที่

$$a_1(i) = \frac{1}{i}\sum_{j=0}^{i-1} d_j(n) \quad \text{and} \quad a_2(i) = \frac{1}{k+1-i}\sum_{j=0}^{k} d_j(n) \qquad (3.7)$$

$$s_1^{\,2}(i) = \sum_{j=0}^{i-1} |\, d_j(n) - a_1(i)\,|^2 \quad \text{and} \quad s_2^{\,2}(i) = \sum_{j=i}^{k} |\, d_j(n) - a_2(i)\,|^2 \qquad (3.8)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดย a_1 แทนค่าเฉลี่ยกลุ่มที่ 1 a_2 แทนค่าเฉลี่ยกลุ่มที่ 2 s_1^2 คือค่า Variance จากกลุ่มที่ 1 และ s_2^2 คือค่า Variance จากกลุ่มที่ 2

อัลกอริทึมจะทำการคำนวณหาค่า J(i) สำหรับแต่ละ i และหาค่าตำแหน่ง cut-off ที่ J(i) มีค่าสูงสุด ดังเช่น

$$m(n) = \arg\max_{i} J(i) \tag{3.9}$$

ค่า J(i) ที่ได้สามารถที่จะใช้เป็นตัวแยกสีทั้งสอง clusters ใน window ได้ และถ้าหากว่ามี cluster มากกว่า 2 clusters แล้วมีอันหนึ่งที่มี $x_0(n)$ อยู่ด้วยก็ยังสามารถที่จะถูกแยกออกจากสีอื่นๆ ที่เหลือ แต่ถ้าหากว่ามีสีเพียงแค่ cluster เดียวแล้ว Peer group จะไม่ประกอบไปด้วยทุกๆ จุดใน window แต่อย่างไรก็ตาม $x_0(n)$ จะยังคงให้ค่าที่เรียบ (Smooth) ด้วยการเฉลี่ยสมาชิกของ Peer group มันเอง

ขั้นตอนในการกำจัด Impulse noise นั้น ขั้นแรกจะทำการหาค่า First order differences ของค่าผลต่าง $d_i(n)$, $f_i(n)$ จะถูกคำนวณหาค่าก่อนที่จะทำการจัดกลุ่ม Peer group

$$f_i(n) = d_{i+1}(n) - d_i(n) \tag{3.10}$$

ขั้นตอนต่อไปคือการทดสอบเปรียบเทียบค่า $x_i(n)$ ตัวแรกและตัวสุดท้ายของ M points เพื่อใช้เช็คว่ามันเป็นส่วนของ Noise หรือไม่ ดังนี้

$$f_i(n) \le \alpha \tag{3.11}$$

โดยที่ M = w/2 หรือครึ่งหนึ่งของขนาด window และ α คือค่า threshold ของสัญญาณรบกวนในภาพ จะถูกกำหนดให้มีค่ามากสำหรับภาพที่มีสัญญาณรบกวนปริมาณมาก และค่าน้อยสำหรับภาพที่มีสัญญาณรบกวนปริมาณน้อย ในวิทยานิพนธ์ได้ใช้ α=2 ถ้า $f_i(n)$ ไม่เป็นไปตามเงื่อนไข แล้วจุดสิ้นสุดของค่า $x_j(n)$ สำหรับ $j \le i$ หรือ $j > i$ ใน Peer Groupจะถูกพิจารณาว่าเป็น impulse noise และจะถูกกำจัด ส่วน $d_j(n)$ ที่เหลือจะถูกใช้ในการประมาณหาค่า Peer group จริงๆ

วิธีการดังกล่าวจะทำงานภายใต้สมมติฐานว่าถ้า $x_0(n)$ เป็นส่วนหนึ่งของ impulse noise ซึ่งมักจะมีค่าที่แตกต่างจากสมาชิกอื่นๆ ภายใน window มาก ซึ่งสามารถที่จะถูกหาได้จากค่าของ $f_i(n)$ และในการกำจัด Noise อื่นๆ ที่รวมอยู่ใน window อาจจะมีผลต่อผลลัพธ์ที่ได้จากการจัดกลุ่ม peer group

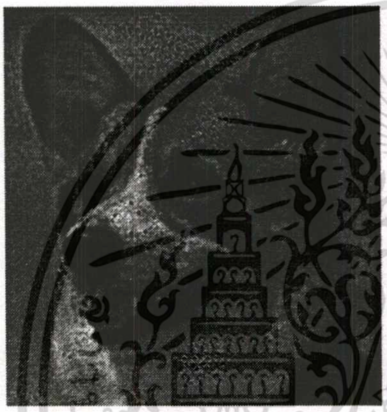
หลังจากการกำจัด impulse noise และทำการจัดกลุ่ม Peer group แล้ว พิกเซล $x_0(n)$ จะถูกแทนด้วยค่าเฉลี่ยของสมาชิก Peer group ที่มีการ weight ไว้ด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$x_{new}(n) = \frac{\sum_{i=0}^{m(n)-1} w_i p_i(n)}{\sum_{i=1}^{m(n)-1} w_i}, p_i(n) \in P(n) \qquad (3.12)$$

โดยที่ w_i คือค่า standard Gaussian weights ซึ่งขึ้นอยู่กับค่าตำแหน่งสัมพัทธ์ของ $p_i(n)$ โดย เทียบกับ $x_0(n)$

การปรับภาพให้เรียบ (smoothing) และการกำจัด Impulse Noise ในภาพสี ที่เรียกว่า Peer group filtering โดยอัลกอริทึมจะมีผลต่อการกำจัด Noise และการปรับสีของภาพให้เรียบขึ้น ซึ่งจะไม่ทำให้ขอบของวัตถุและรายละเอียดที่สำคัญในภาพนั้นเกิดการ Blur ดังแสดงในรูปที่ 3.3



(a) ก่อนการทำภาพให้ราบเรียบ

(b) หลังจากการทำภาพให้ราบเรียบ

รูปที่ 3.3 แสดงภาพผลลัพธ์จากการทำภาพให้ราบเรียบด้วย PGF

3.3 การจัดระดับสัญญาณหรือการควอนไตเซชั่น (Quantization)

เทคนิคหนึ่งที่ถูกนำมาใช้ในการเพิ่มความเร็วในการประมวลผลและการสื่อสารข้อมูลไม่ว่าจะเป็นข้อมูลเสียงหรือภาพก็คือการลดขนาดของข้อมูลต้นแบบให้มีขนาดเล็กลง การควอนไตเซชั่นเป็นเทคนิคหนึ่งซึ่งถูกนำมาใช้ในการลดขนาดข้อมูลต้นแบบ โดยการควอนไตเซชั่นนั้นจะใช้หลักการของการจัดระดับสัญญาณหรือข้อมูลอินพุทใหม่ เนื่องจากจำนวนระดับของข้อมูลอินพุทหรือข้อมูลต้นแบบนั้นประกอบด้วยระดับสัญญาณต่างๆ มากมาย ทำให้เราต้องใช้จำนวนบิตของสัญญาณดิจิตอลมากมาย การจัดระดับสัญญาณข้อมูลอินพุทใหม่หรือการควอนไตเซชั่นจึงถูกนำมาใช้ในการลดจำนวนบิตที่ใช้แทนข้อมูลแต่ละค่า โดยตัวควอนไตอร์จะทำการจัดแบ่งระดับสัญญาณของข้อมูลอินพุทใหม่ให้มีระดับสัญญาณจำกัดอยู่ 2^n ระดับ ทำให้เราสามารถแทนข้อมูลในแต่ละค่าของสัญญาณอินพุทด้วยสัญญาณดิจิตอลขนาด n บิต การควอนไตซ์สัญญาณนั้นสามารแบ่งออกได้เป็นสองแบบตามลักษณะของสัญญาณอินพุทคือ การควอนไตซ์สัญญาณสเกลลาร์ (Scalar Quantization) และการควอนไตซ์สัญญาณเวคเตอร์ (Vector Quantization)

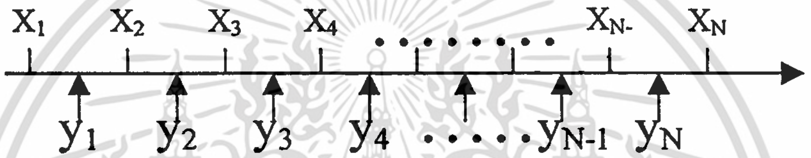
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.1 การควอนไตซ์สัญญาณสเกลลาร์ (Scalar Quantization)

การควอนไตซ์สัญญาณสเกลลาร์ถูกนำมาใช้ในการจัดระดับสัญญาณอินพุทแบบสเกลลาร์ โดยการควอนไตซ์สัญญาณสเกลลาร์Q เป็นการแม็ปค่าสัญญาณอินพุท x ให้เป็นสัญญาณ y ซึ่งอยู่ในช่วงของสัญญาณที่แน่นอน

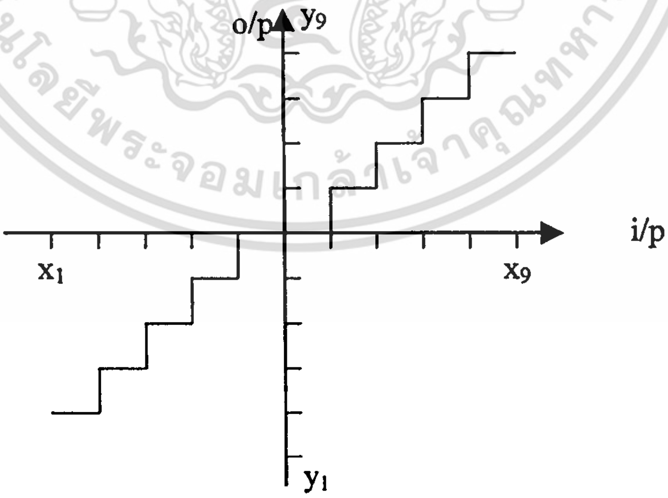
$$Q : x \longrightarrow y$$

การควอนไตซ์แบบสัญญาณสเกลลาร์แสดงได้โดยการแบ่งจำนวนบนแกนจำนวนจริงออกเป็นช่วงย่อยๆ ดังรูปที่ 3.1 กำแพงที่กั้นระหว่างช่วงจะเป็นตัวอ้างอิงในการตัดสินว่าระดับสัญญาณอินพุทควรถูกจัดอยู่ในช่วงใดเช่นหากระดับสัญญาณอินพุท x ตกอยู่ในช่วงที่ i เอาท์พุทที่ถูกเลือกใช้แทนสัญญาณ x ก็คือสัญญาณ yi



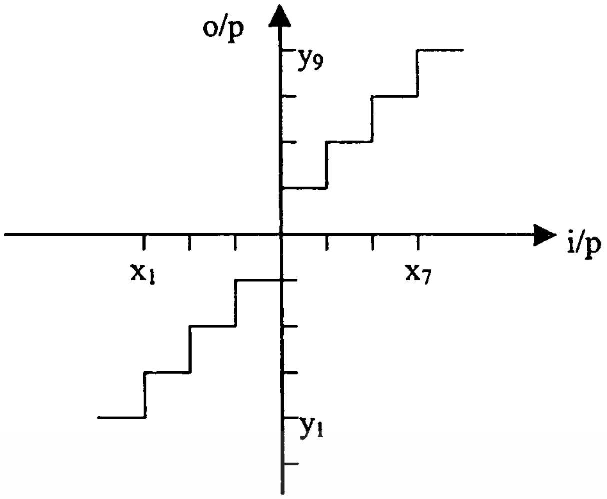
รูปที่ 3.4 การแบ่งช่วงค่าของข้อมูลสำหรับการควอนไตซ์สัญญาณสเกลลาร์

กราฟขั้นบันไดเชิงเส้นสองมิติสามารถนำมาใช้อธิบายการทำงานของตัวควอนไตเซอร์ดังรูปที่ 3.5



(a) Mid-tread Staircase Quantizer

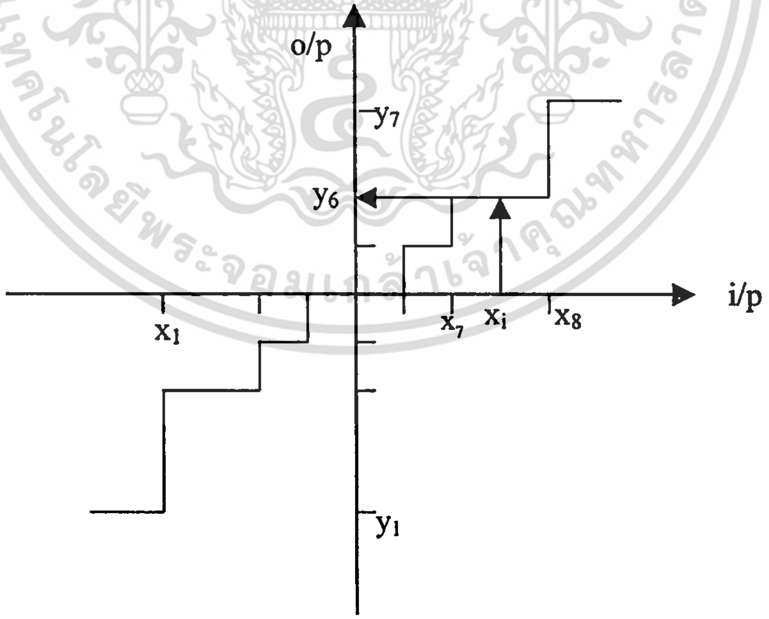
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(b) Mid-rise Staircase Quantizer

รูปที่ 3.5 ควอนไตเซอร์เชิงเส้นแบบเป็นขั้นบันได

จากรูปที่ 3.5 แกน X แทนระดับของสัญญาณอินพุทและแกน y แทนระดับสัญญาณเอาท์พุทเส้นแบ่งบนแกน x เป็นเส้นอ้างอิงที่ใช้ในการจัดระดับสัญญาณอินพุท ในขณะที่เส้นแบ่งบนแกน y เป็นระดับของสัญญาณเอาท์พุท โดยในรูปที่ 3.5 ทั้ง (a) และ (b) นั้นเป็นควอนไตเซอร์แบบเชิงเส้นทั้งคู่



รูปที่ 3.6 ควอนไตเซอร์ขั้นบันไดแบบไม่เป็นเชิงเส้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมมุติตามควอนไตเซอร์ที่แสดงในรูปที่ 3.6. ถ้ามีสัญญาณอินพุท x ซึ่งมีระดับสัญญาณตกอยู่ระหว่าง x_7 และ x_8 ดังนั้นเอาท์พุทจะถูกแม็ปเป็นระดับสัญญาณ y_6

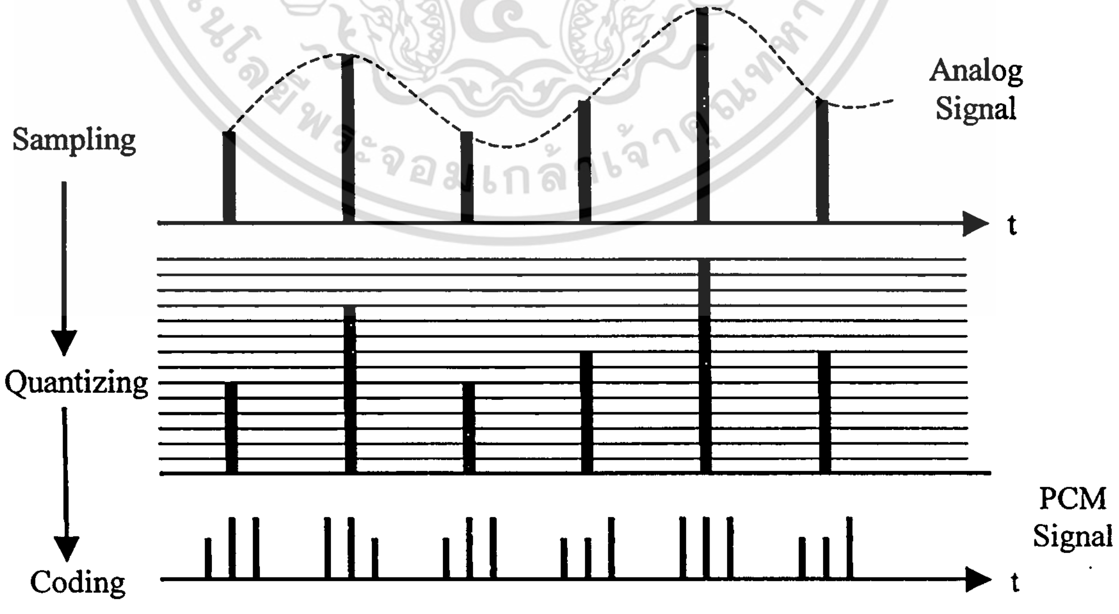
ในอุดมคติ สัญญาณเอาท์พุทในแต่ละช่วง (หรือที่เราเรียกว่า "โค๊ดเวิร์ด") จะเป็นค่าที่ดีที่สุดในช่วงนั้นๆ ซึ่งค่าที่ดีที่สุดในแต่ละช่วงก็คือค่ากึ่งกลางของแต่ละช่วงนั่นเอง ค่ากึ่งกลางของแต่ละช่วงนี้เราสามารถนิยามได้ว่าเป็นค่าเฉลี่ยของค่าอินพุททุกๆ ตัวที่ตกอยู่ในช่วงนั้นๆ ดังนั้นการที่จะคำนวณค่าเฉลี่ยในแต่ละช่วงได้จำเป็นต้องกำหนดขอบเขตที่แน่นอนของแต่ละช่วงก่อน ถ้าเรารู้ขอบเขตของแต่ละช่วงก็จะสามารถคำนวณค่ากึ่งกลางที่แน่นอนได้

วิธีการควอนไตเซอร์แบบสเกลลาร์มักจะถูกนำมาใช้ในการแปลงสัญญาณดิจิตอล โดยในการแปลงสัญญาณอนาล็อกเป็นดิจิตอลนั้น การสุ่มค่าตัวอย่างของสัญญาณอินพุทจะอยู่ในอัตราคงที่ และนำเอาค่าที่สุ่มของสัญญาณที่ได้นี้ไปทำการควอนไตซ์

เพื่อให้มีความเข้าใจและมองเห็นภาพการควอนไตซ์แบบสเกลลาร์ได้ชัดเจนมากยิ่งขึ้น ขอยกตัวอย่างการประยุกต์ใช้การควอนไตซ์สัญญาณแบบสเกลลาร์ในการแปลงสัญญาณอนาล็อกเป็นสัญญาณดิจิตอลด้วยวิธี PCM ซึ่งเป็นมาตราฐานของโทรศัพท์ดิจิตอลที่ใช้กันอยู่ในอเมริกาเหนือมาประกอบการทำความเข้าใจ

หลักการทำงานของพัลซ์โค๊ดโมดูเลชั่น (PCM) [7]

การทำพัลซ์โค๊ดโมดูเลชั่นนั้นถูกนำมาใช้ในการเปลี่ยนสัญญาณเสียงซึ่งเป็นสัญญาณอนาล็อกให้เป็นสัญญาณดิจิตอล PCM โดยสัญญาณเสียงอนาล็อกจะถูกเปลี่ยนไปเป็นสัญญาณดิจิตอล PCM โดยผ่านกระบวนการ การสุ่มตัวอย่าง ควอนไตซิ่ง และการเข้ารหัส ดังแสดงในรูปที่ 3.7



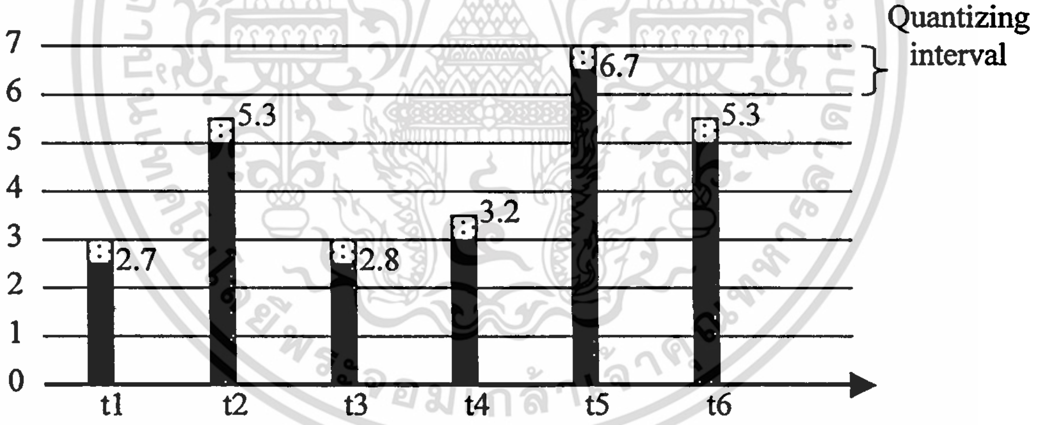
รูปที่ 3.7 การแปลงสัญญาณอนาล็อกเป็นดิจิตอลด้วยวิธี PCM

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับในภาคการสุ่มตัวอย่างนั้นก็คือการหยิบยกเอาแอมปลิจูดของสัญญาณอนาล็อกมาตามช่วงเวลาที่เหมาะสมเป็นช่วงๆ โดยอาศัยทฤษฎีการสุ่มตัวอย่างสัญญาณที่ได้จากการสุ่มตัวอย่างเหล่านี้ถูกเรียกว่าสัญญาณ PAM (Pulse Amplitude Modulation) ซึ่งเป็นสัญญาณที่ถูกจัดเรียงในสเกลของเวลา

หลังจากนั้นสัญญาณ PAM จะถูกจัดระดับโดยการควอนไตซ์ในกระบวนการควอนไตซ์นี้สเกลของแอมปลิจูดจะถูกแบ่งเป็นช่วงๆ เท่าๆ กัน ค่าควอนไตซ์แต่ละค่าเรียกว่าระดับควอนไตซ์และระยะระหว่างแต่ละระดับเรียกว่า " Quantizing Interval " ขนาดของแต่ละสุ่มตัวอย่างจะถูกจัดให้อยู่ในระดับการควอนไตซ์ที่อยู่ใกล้ที่สุด

สัญญาณ PAM ที่ได้ทำการควอนไตซ์เพียงแต่เป็นการประมาณค่าให้ใกล้เคียงกับสัญญาณอนาล็อกดั้งเดิม ดังนั้นจึงมีข้อผิดพลาดที่เกิดขึ้นบ้างระหว่างค่าระดับแอมปลิจูดของสัญญาณทั้งสอง ข้อผิดพลาดนี้เรียกว่าสัญญาณรบกวนที่เกิดจากการควอนไตซ์ดังแสดงในรูปที่ 3.6 ขนาดของสัญญาณรบกวนเหล่านี้กระจายอย่างมีระเบียบภายในช่วงห่างระหว่างระดับการควอนไตซ์โดยไม่คำนึงถึงแอมปลิจูดของสัญญาณเดิม หรือพูดอีกในหนึ่งก็คือ กำลังของสัญญาณรบกวนควอนไตซ์ซิ่งเกือบจะคงที่ ไม่ขึ้นอยู่กับระดับกำลังของสัญญาณ



รูปที่ 3.8 ความผิดพลาดที่เกิดขึ้นจากกระบวนการควอนไตซ์ซิ่ง

ถึงแม้สัญญาณรบกวนที่เกิดจากการควอนไตซิ่งไม่สามารถจะป้องกันไม่ให้เกิดขึ้นได้แต่ก็ควรจะให้มีน้อยที่สุด เพื่อให้คุณภาพของสัญญาณมีคุณภาพดี ถ้าเราดูจากรูปจะเห็นว่าสัญญาณรบกวนที่เกิดจากการควอนไตซ์มีความสัมพันธ์กับช่วงห่างระหว่างระดับการควอนไตซ์ เนื่องจากแอมปลิจูดของสัญญาณรบกวนที่เกิดจากการควอนไตซ์จะไม่เกินช่วงห่างระหว่างควอนไตซ์ ดังนั้นถ้าทำช่วงห่างระหว่างระดับการควอนไตซ์ให้แคบสัญญาณรบกวนก็จะสามารถลดลงได้

ขั้นตอนสุดท้ายในการแปลงสัญญาณอนาล็อกเป็นดิจิตอลด้วยวิธีพัลซ์โค๊ดโมดูเลชั่นก็คือการเข้ารหัสให้อยู่ในรูปไบนารีโค๊ด ไบนารีโค๊ดประกอบด้วยการรวมกันของยูนิตพัลซ์ซึ่งอยู่ในรูป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การมีพัลซ์หรือไม่มีพัลซ์คือ 1 และ 0 ในขั้นตอนนี้สัญญาณที่ผ่านการควอนไตซ์มาจะถูกแทนที่ด้วยโค๊ดไบนารีที่เรียกว่า " โค๊ดเวิร์ด " ซึ่งเป็นโค๊ดที่ใช้แทนในแต่ละระดับสัญญาณควอนไตซ์

3.3.2 การควอนไตซ์สัญญาณเวคเตอร์ (Vector Quantization)

การควอนไตซ์สัญญาณแบบเวคเตอร์นั้นเป็นเทคนิคที่พัฒนามาจากการควอนไตซ์ข้อมูลแบบสเกลลาร์ซึ่งเป็นปริมาณที่มีมิติเพียงมิติเดียว การเปลี่ยนแปลงจากข้อมูลใน 1 มิติมาเป็นข้อมูลในหลายมิตินั้นเป็นขั้นตอนหลักขั้นตอนหนึ่งในการพัฒนาไปสู่แนวคิดและเทคนิคใหม่ๆ ซึ่งสามารถนำมาเทียบเคียงกับแอพพลิเคชั่นที่ใช้การควอนไตซ์แบบสเกลลาร์ ในขณะที่การควอนไตซ์ปริมาณสเกลลาร์ถูกนำมาใช้ในการแปลงสัญญาณอนาล็อกเป็นสัญญาณดิจิตอล การควอนไตซ์ปริมาณแบบเวคเตอร์ก็ถูกนำมาใช้ในกระบวนการประมวลผลภาพ ซึ่งในเกือบทุกกรณีข้อมูลอินพุทจะถูกแทนอยู่ในรูปแบบของสัญญาณดิจิตอลอยู่แล้ว การวิเคราะห์และเทคนิคการออกแบบต่างๆ ของแอพพลิเคชั่นที่ใช้ในการควอนไตซ์แบบเวคเตอร์มักจะถูกนำมากล่าวไปพร้อมกันกับการควอนไตว์แบบสเกลลาร์

ปริมาณเวคเตอร์สามารถนำมาใช้ในการอธิบายข้อมูลในรูปแบบต่างๆ ได้มากมายเช่น ส่วนของสัญญาณคลื่นเสียงหรือภาพ ซึ่งการฟอร์มเวคเตอร์อย่างง่ายได้แก่การฟอร์มเวคเตอร์ขึ้นมาจากการสุ่มตัวอย่างของข้อมูลภาพหรือสัญญาณเสียง เวคเตอร์ควอนไตเซชั่นจึงมองได้ว่าเป็นส่วนหนึ่งของการรู้จำข้อมูล (Pattern Recognition) เมื่อข้อมูลอินพุทถูกประมาณและแทนด้วยข้อมูลพื้นฐานที่อยู่ในเซ็ตเซ็ตหนึ่งที่ถูกกำหนดขึ้น

เวคเตอร์ควอนไตเซชั่นเป็นวิธีการจัดระดับสัญญาณข้อมูลวิธีหนึ่งที่ถูกนำมาใช้กับข้อมูลในรูปแบบต่างๆ มากมายซึ่งรวมไปถึงข้อมูลภาพ เสียง การส่งโทรสาร หรือแม้แต่ข้อมูลทางดาวเทียมด้วยวิธีเวคเตอร์ควอนไตเซชั่นนี้ เวคเตอร์ที่ได้มาจากข้อมูลภาพจะถูกเข้ารหัสด้วยเวคเตอร์อ้างอิงกลุ่มหนึ่งหรือที่เราเรียกว่า "โค๊ดบุค(codebook) " จากทฤษฎีอัตราความผิดเพี้ยนของสัญญาณแสดงให้เห็นว่าวิธีการจัดระดับสัญญาณข้อมูลด้วยวิธีนี้สามารถให้คุณสมบัติการจัดระดับสัญญาณภาพที่ดีกว่าการเข้ารหัสด้วยเทคนิคอื่นๆ ซึ่งใช้หลักการเข้ารหัสโดยปริมาณสเกลลาร์ จุดมุ่งหมายของเวคเตอร์ควอนไตเซชั่นนั้นก็คือการคำนวณและออกแบบเซ็ตของเวคเตอร์อ้างอิงหรือโค๊ดบุคที่ทำให้ภาพเอาพุท์ที่ได้หลังการถอดรหัสมีอัตราความเพี้ยนเฉลี่ยต่ำที่สุด [15]

คำนิยามของเวคเตอร์ควอนไตเซชั่น

เวคเตอร์ควอนไตเซชั่น Q ใน k มิติซึ่งมีขนาดเป็น N เป็นการแม็ปปิ้งเวคเตอร์ใน k มิติกับเซ็ต C ที่แน่นอนซึ่งบรรจุด้วย N เอาท์พุท ซึ่งถูกเรียกว่าโค๊ดเวคเตอร์หรือโค๊ดเวิร์ด ดังนั้น

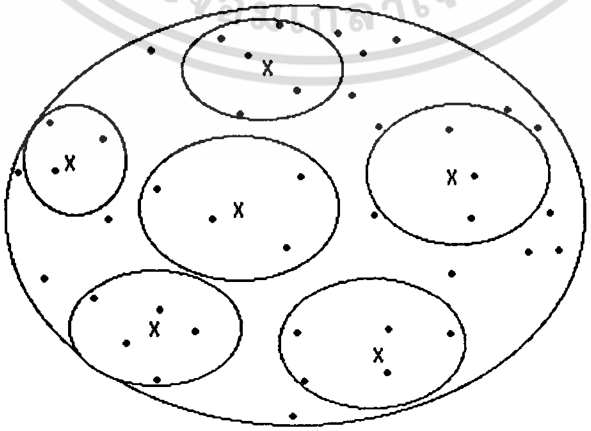
$$Q : R^k \rightarrow C$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ $C=\{Y_1,Y_2,Y_3,\ldots\ldots.Y_N\}$ และ$Y_i \in R_k$ สำหรับทุกๆ $i \in \{1,2,3,\ldots\ldots N\}$ เซ็ต C ถูกเรียกว่า " โค๊ดบุค " หรือ " โค๊ด " ซึ่งบรรจุด้วยเวคเตอร์ N เวคเตอร์แต่ละเวคเตอร์อยู่ใน R^k อัตราการเข้าโค๊ดหรือเราเรียกว่าอัตราของเวคเตอร์ควอนไตเซอร์ $r=(\log_2 N)/k$ ซึ่งเป็นอัตราที่บอกถึงจำนวนบิตที่ใช้ต่อหนึ่ง องค์ประกอบเวคเตอร์เพื่อใช้ในการแทนค่าอินพุทเวคเตอร์ และยังสามารถบอกถึงประสิทธิภาพของการออกแบบโค๊ดบุคนั้นด้วย โค๊ดบุคนั้นจึงเปรียบเสมือนเป็นตารางที่เก็บค่าของเวคเตอร์ต่างๆ เอาไว้ในหน่วย ความจำเพื่อใช้ในการแทนค่าเวคเตอร์อินพุท

เวคเตอร์ควอนไตเซชั่นเป็นรูปแบบใหม่ของการควอนไตซ์ที่พัฒนามาจากการควอนไตซ์ปริมาณสเกลลาร์ เวคเตอร์ควอนไตเซชั่นเป็นทางเลือกใหม่สำหรับการควอนไตซ์สัญญาณหรือข้อมูลแบบเวคเตอร์ ซึ่งปัจจุบันยังไม่มีเทคโนโลยีการเข้ารหัสใดที่มีประสิทธิภาพดีไปกว่าการเข้ารหัสแบบเวคเตอร์ควอนไตเซชั่น

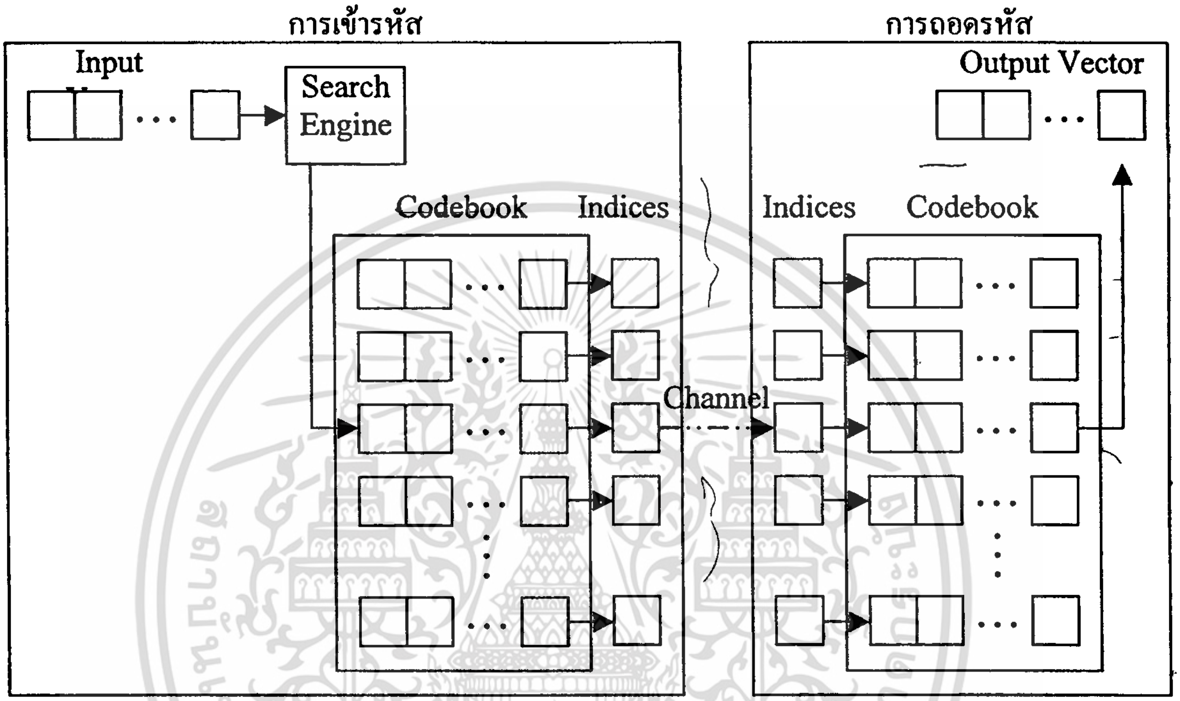
ขั้นตอนหลักที่สำคัญอีกขั้นตอนหนึ่งสำหรับแบ่งกลุ่มข้อมูลรูปภาพด้วยวิธีเวคเตอร์ควอนไตเซชั่นก็คือขั้นตอนการออกแบบโค๊ดบุค(Codebook Designing) หรือการคำนวณเซ็ตของเวคเตอร์อ้างอิง ในขั้นตอนการออกแบบโค๊ดบุคนี้จะเป็นขั้นตอนที่ทำหน้าที่คำนวณเซ็ตของค่าเวคเตอร์อ้างอิงที่เหมาะสมที่นำมาใช้ในการประมาณค่าเวคเตอร์อินพุท โดยค่าเวคเตอร์อ้างอิงที่ถูกคำนวณขึ้นมานี้เมื่อนำไปใช้ประ มาณค่าเวคเตอร์อินพุทแล้วจะต้องให้เวคเตอร์เอาท์พุทที่มีความผิดเพี้ยนไปจากสัญญาณอินพุทต่ำที่สุด สำหรับวิธีที่นิยมใช้ในการออกแบบโค๊ดบุคสำหรับแบ่งกลุ่มข้อมูลรูปภาพนั้นมักใช้วิธีการจำแนกกลุ่มข้อมูลหรือที่เราเรียกว่าการวิเคราะห์คลัสเตอร์นั่นเอง ในการคำนวณค่าเวคเตอร์อ้างอิงจากการวิเคราะห์ คลัสเตอร์ข้อมูลภาพอินพุทนี้ก็คือการจำแนกข้อมูลภาพออกเป็นกลุ่มหรือคลัสเตอร์แล้วทำการคำนวณค่าเวคเตอร์เฉลี่ยของเวคเตอร์แต่ละกลุ่มนี้ ค่าเวคเตอร์เฉลี่ยที่ได้ก็เป็นค่าเซ็นเตอร์ของทุกคลัสเตอร์ดังแสดงได้ในรูปที่ 3.9



รูปที่ 3.9 แสดงการคำนวณค่าเวคเตอร์อ้างอิงจากการวิเคราะห์คลัสเตอร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนสุดท้ายคือการเข้ารหัสข้อมูลภาพ เนื่องจากแต่ละเวคเตอร์ภายในเซ็ตของเวคเตอร์อ้างอิงจะมีรหัสหรือดัชนีประจำแต่ละเวคเตอร์ซึ่งเราเรียกว่า " โค๊ดเวิร์ด(Codeword) " ดังนั้นบางครั้งเราจึงเรียกเวคเตอร์ภายในเซ็ตของเวคเตอร์อ้างอิงว่า " โค๊ดเวคเตอร์ " ดังนั้นในฟังก์ชั่นบล็อคนี้เวคเตอร์เอาท์พุทที่ได้จากการประมาณค่าเวคเตอร์อินพุทด้วยเวคเตอร์อ้างอิงจะถูกนำมาเข้ารหัสด้วยโค๊ดเวิร์ดของแต่ละเวคเตอร์เอาท์พุท รูปที่ 3.10 แสดงระบบการเข้ารหัสและถอดรหัสเวคเตอร์



รูปที่ 3.10 การเข้ารหัสและการถอดรหัสเวคเตอร์

จากรูปที่ 3.10 ภายในฟังก์ชั่นบล็อคนี้เวคเตอร์ที่ได้จากภาพอินพุทแต่ละเวคเตอร์จะถูกนำมาประมาณค่าด้วยเวคเตอร์ค่าหนึ่งซึ่งอยู่ในเซ็ตของเวคเตอร์อ้างอิงหรือ " โค๊ดบุค " ซึ่งถูกสร้างขึ้นมาโดยกระบวนการเวคเตอร์ควอนไตเซชั่น โดยเวคเตอร์เอาพุทที่ได้จากการประมาณค่าแล้วนี้จะต้องมีความผิดพลาดของสัญญาณเมื่อเทียบกับสัญญาณอินพุทต่ำที่สุด เวคเตอร์แต่ละค่าภายในโค๊ดบุคนี้จะมี " โค๊ดเวิร์ด " หรือรหัสประจำตัวของแต่ละเวคเตอร์ เราจึงมักเรียกแต่ละเวคเตอร์อ้างอิงภายในโค๊ดบุคว่า " โค๊ดเวคเตอร์ "

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4 การวิเคราะห์คลัสเตอร์ข้อมูลภาพแบบ General Lloyd algorithm (GLA) ด้วยวิธีเวคเตอร์ควอนไตเซชั่นมาใช้ในการแบ่งกลุ่มรูปภาพ

การวิเคราะห์คลัสเตอร์แบบ General Lloyd Algorithm (GLA) ด้วยวิธีเวคเตอร์ควอนไตเซชั่น [11] เป็นการวิเคราะห์คลัสเตอร์วิธีหนึ่งที่มีการคำนวณแบบทำซ้ำ (iterative clustering)

ขั้นตอนแรกของกระบวนการ GLA ก็คือการสุ่มตัวอย่าง (random) ค่าเริ่มต้น จากนั้นก็ทำการแบ่งกลุ่มสีรูปภาพที่ได้จากการกำหนดค่าเริ่มต้นให้เป็นกลุ่มย่อยๆ จนกว่าค่า distortion ที่ได้ค่าใหม่จะเท่ากับหรือใกล้เคียงค่าเก่า (ก่อนที่จะมีการปรับปรุง) จึงจะเสร็จสิ้นการทำงาน ดังแสดงได้ในตารางที่ 3.1 โดย C คือ และ m คือเซ็ตของโค๊ดบุค

ตารางที่ 3.1 The Generalize Lloyd Algorithm

Table 3.1 The Generalize Lloyd Algorithm	
Step 1.	Begin with an initial codebook C_1. Set m = 1
Step 2 .	Give the codebook, C_m, perform the Lloyd Iteration to generate the improved codebook C_{m+1}.
Step 3.	Compute the average distortion for C_{m+1} . If it has changed by a small enough amount since the last iteration, stop. Otherwise set m+1 $\rightarrow$ m and go to Step 2.

เทคนิคการวิเคราะห์คลัสเตอร์แบบ general Lloyd algorithm (GLA) ด้วยวิธีเวคเตอร์ควอนไตเซชั่นนี้จะใช้ local statistics ที่ได้หลังจากการทำ Peer group filtering เป็นค่าสำหรับการ weight ใน Vector quantization (VQ) โดยมีขั้นตอนดังต่อไปนี้

1. อัลกอริทึมที่เรียกว่า "peer group filtering" จะถูกนำมาใช้เพื่อที่ให้ภาพนั้นเรียบ (smooth) ขึ้น และเพื่อทำการกำจัด noise

2. หาค่าผลต่างสูงสุดของแต่ละ Peer group T(n), $T(n) = d_{m(n)-1}(n)$ ค่าของ T(n) จะเป็นตัวแสดงให้เห็นถึงความเรียบ (smoothness) ของ local region โดยที่ค่า weight ของแต่ละพิกเซลสามารถคำนวณหาได้จากสมการ (3.13)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$v(n) = \exp(-T(n)) \qquad (3.13)$$

พิกเซลที่เป็น Noise จะถูก weight ให้มีค่าน้อยกว่าพิกเซลที่อยู่ในพื้นที่ราบเรียบ (smooth region)

3. T_{avg} คือค่าเฉลี่ยของ T(n), จะแสดงถึงความเรียบของทั้งรูปภาพ โดยทั่วๆ ไปแล้วค่า T_{avg} ที่มีค่าสูง จะมีความราบเรียบน้อยในภาพ และ cluster ทั้งหลายจำเป็นที่จะต้องมีการควอนไตซ์สีของภาพ เพิ่มขึ้น และค่าเริ่มต้นของจำนวน cluster N เวคเตอร์ควอนไตเซชั่น สามารถที่จะประมาณค่าได้จาก

$$N = \beta T_{avg} \qquad (3.14)$$

โดยที่ β คืออัตราขยาย จะถูกกำหนดให้เท่ากับ 2 สำหรับในการทดลอง

4. ใช้อัลกอริทึมของ general Lloyd algorithm (GLA) จะถูกนำมาใช้สำหรับการควอนไตซ์พิกเซลสี เวคเตอร์ควอนไตเซชั่น โดยกฎสำหรับการ Update จะถูกปรับเปลี่ยนไปยังพิกเซลที่มีการ weightและ สำหรับ color cluster C_i คือ centroid

$$C_i = \frac{\sum v(n)x(n)}{\sum v(n)}, x(n) \in C_i \qquad (3.15)$$

ค่า centroid จะถูก shift ไปด้วยค่า weight ที่สูงกว่า

5. สำหรับค่า Initial cluster ของ GLA สามารถที่จะหาได้จากอัลกอริทึม Splitting initialization ซึ่งค่า weighted distortion หาได้จาก

$$D = \sum_i D_i = \sum_i \sum_n v(n) \| x(n) - c_i \|^2, x(n) \in C_i \qquad (3.16)$$

โดยที่ C_i คือ centroid ของ cluster C_i

x(n) และ v(n) คือ color vector และ perceptual weight สำหรับพิกเซล n จาก (3.13)

D_i คือ distortion รวมของ cluster C_i

ค่า D ที่หาได้จะถูกใช้เพื่อเป็นตัวบอกว่า cluster ไหนที่จะถูก Split

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. ขั้นตอนสุดท้ายของ VQ ค่า centriod ของ cluster จะถูกคำนวณโดยไม่มีการ weight เพื่อหาค่า center จริงๆ ของ cluster หลังจากใช้อัลกอริทึม GLA แล้ว บางครั้งมีพิกเซลจำนวนมากที่มีสีใกล้เคียงกันแต่มีมากกว่า 1 cluster เพราะเนื่องจาก GLA พยายามที่จะลดค่า global distortion ดังนั้นจึงต้องใช้อัลกอริทึม Agglomerative clustering [11] ซึ่งมีขั้นต้อนดังแสดงในตารางที่ 3.2

ตารางที่ 3.2 แสดง Agglomerative Clustering Algorithm

Table 3.2 Agglomerative Clustering Algorithm

<u>**begin**</u> <u>**initialize**</u> $c, \hat{c} \leftarrow n, D_i \leftarrow \{x_i\}, i=1,\ldots,n$

    <u>**do**</u> $\hat{c} \leftarrow \hat{c}-1$

        find nearest clusters, say, D_i and D_j

        merge D_i and D_j

    <u>**until**</u> $c = \hat{c}$

<u>**return**</u> c clusters

<u>**end**</u>

Agglomerative clustering จะทำงานบนค่า cluster centroids เพื่อทำการรวมกลุ่มสีที่มีค่าใกล้เคียงกัน ดังนั้นค่าผลต่างระหว่างสอง centroids ที่น้อยที่สุดจะถูกใช้เป็น threshold สำหรับการรวมพื้นที่ ซึ่งค่า threshold ดังกล่าวนี้ก็คือพารามิเตอร์สำหรับการควอนไตซ์ที่ได้เคยกล่าวไว้แล้วนั่นเอง ผลลัพธ์สุดท้ายของการควอนไตซ์ภาพจะได้ค่าพิกเซลซึ่งถูกแทนด้วยค่า cluster centroid ที่พิกเซลนั้นใกล้เคียงที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

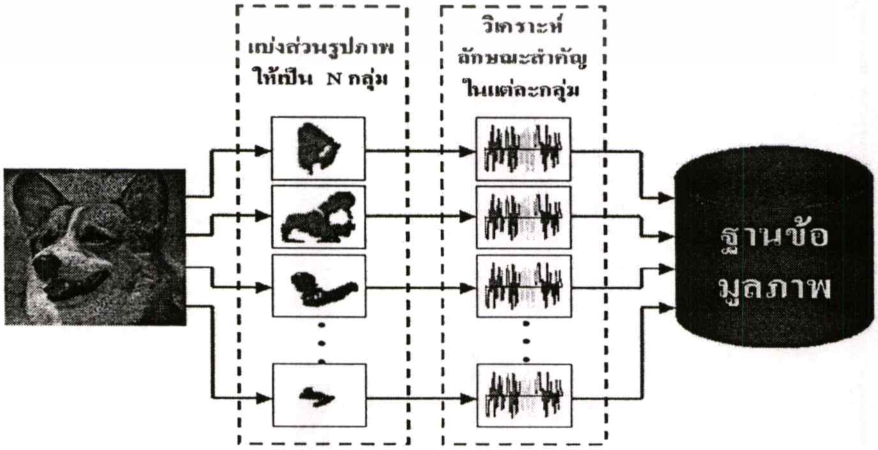
การสร้างดัชนีรูปภาพและการค้นคืนรูปภาพจากฐานข้อมูล

4.1. บทนำ

เทคนิคที่นิยมใช้ในการนำลักษณะสีของรูปภาพมาใช้สำหรับ CBIR คือฮิสโตแกรมของสี (color histogram) ซึ่งฮิสโตแกรมของสีเป็นวิธีหนึ่งที่ใช้ในการคำนวณหาค่าการกระจายของสีต่างๆทั้งหมด ซึ่งเป็นวิธีที่คำนวณได้ง่ายแต่วิธีนี้อาจจะทำให้ค้นคืนรูปภาพผิดพลาดได้ในกรณีที่ฐานข้อมูลมีรูปภาพมากอาจจะมีรูปภาพบางรูปที่แตกต่างกันแต่มีฮิสโตแกรมของสีคล้ายคลึงกันได้ เพื่อแก้ไขปัญหาดังกล่าวในวิทยานิพนธ์นี้ได้นำเอาวิธีการทำดัชนีแบบ region มาใช้ประโยชน์สำหรับค้นหาส่วนของภาพที่สนใจและภาพที่มีสิ่งแวดล้อมแตกต่าง เนื่องจาก ฮิสโตแกรมของสีจะมีความใกล้เคียงกันแม้ว่าวัตถุในภาพจะมีการเปลี่ยนมุมมอง การหมุน การย่อขยายภาพ นอกจากนั้นเราสามารถนำความสัมพันธ์เชิงพื้นที่ระหว่างคู่สี(color correlogram) มาใช้เพื่อแยกความแตกต่างระหว่างกลุ่มต่างๆ ที่มีฮิสโตแกรมของสีใกล้เคียงกันแต่มีการกระจายของตำแหน่งพิกเซลไม่เหมือนกัน เพื่อแก้ไขปัญหาในการค้นคืนข้อมูลผิดพลาด ได้เนื่องจากรูปภาพบางรูปที่แตกต่างกันอาจจะมีโทนสีคล้ายคลึงกันได้ ในวิทยานิพนธ์นี้เราใช้วิธีการค้นคืนข้อมูลแบบผสมระหว่างการสืบค้นข้อมูลแบบเป็นลำดับชั้นและการสืบค้นข้อมูลแบบstatic เพื่อลดเวลาในการค้นหาจากการสืบค้นข้อมูลแบบstatic และ ลดความผิดพลาดในการค้นคืนแบบเป็นลำดับชั้นหลายๆ ชั้น

4.2 การสร้างดัชนี (index) สำหรับแต่ละรูปภาพในฐานข้อมูล

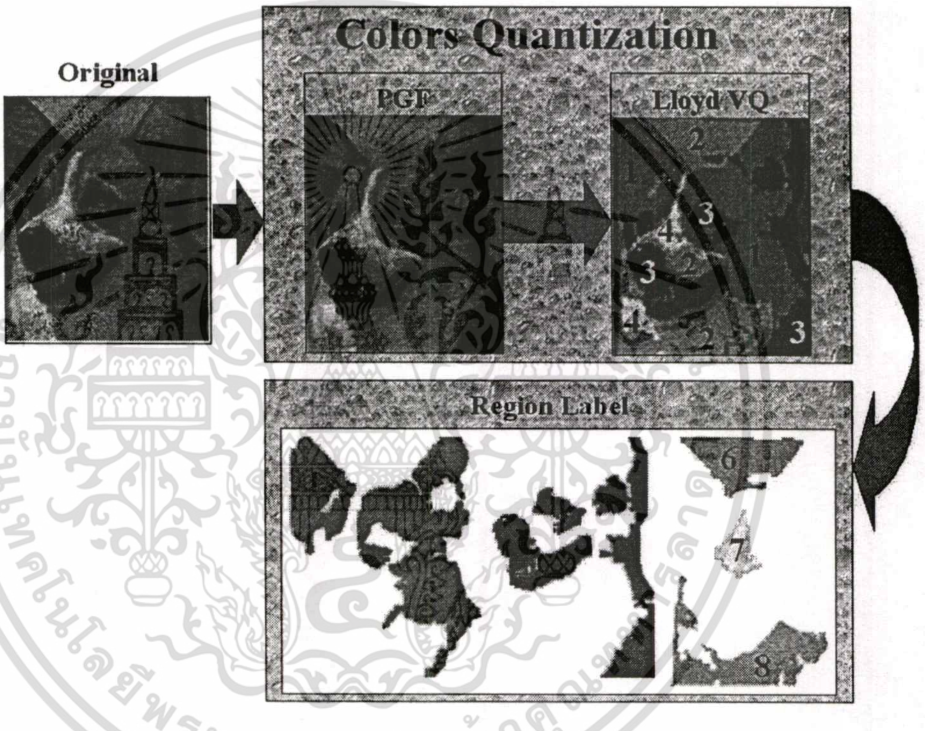
การทำดัชนีแบบ region รูปภาพแต่ละรูปจะถูกแบ่งส่วนให้เป็นกลุ่มๆ จากนั้นจะทำการคำนวณและนำลักษณะสีของภาพด้วยการทำฮิสโตแกรมของสี และ ความสัมพันธ์เชิงพื้นที่ระหว่างคู่สี เฉพาะแต่ละขอบเขตมาเก็บไว้ในฐานข้อมูลเพื่อเป็น key ในการค้นหา ดังรูปที่ 4.1



รูปที่ 4.1 ขบวนการสร้างดัชนี รูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีการทำดัชนีรูปภาพที่นำเสนอนี้ รูปภาพแต่ละรูปถูกแบ่งส่วนให้เป็นจำนวน N กลุ่มโดยใช้ เทคนิค generalized Lloyd algorithm (GLA) ด้วยวิธีเวคเตอร์ควอนไตเซชัน (VQ)[7] ในการแบ่งกลุ่มสีรูปภาพให้เป็นจำนวน N กลุ่ม ก่อนขบวนการแบ่งสีรูปภาพให้เป็นจำนวน N กลุ่มเราใช้เทคนิค peer group filter (PGF)[6] ในการกำจัด Noise และทำให้ภาพเรียบ เทคนิคดังกล่าวมีประสิทธิภาพในการกำจัดสัญญาณรบกวน (Noise) โดยไม่ทำให้ภาพผิดเพี้ยน และลดความชัดเจน(blur) ลง หลังจากการทำ PGF แล้วก็ใช้เทคนิค GLA ด้วยวิธี VQ ในการแบ่งสีรูปภาพให้เป็นจำนวน N กลุ่ม ดังแสดงในรูปที่ 4.2

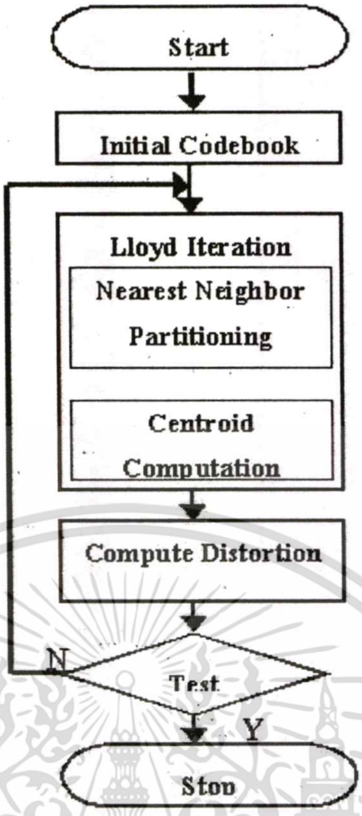


รูปที่ 4.2. ขบวนการแบ่งกลุ่มรูปภาพ

4.2.1 การแบ่งกลุ่มรูปภาพ

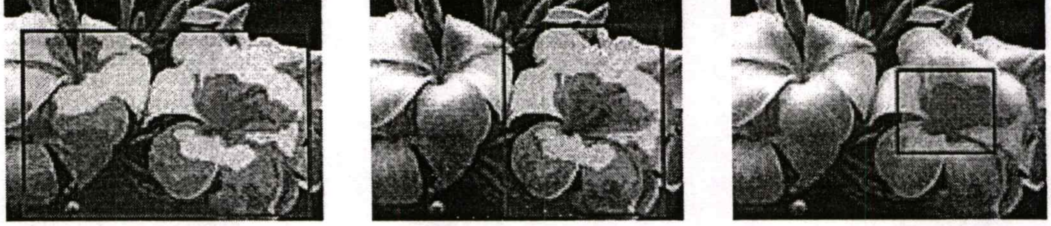
การแบ่งกลุ่มรูปภาพแบบ GLA ด้วยวิธีเวคเตอร์ควอนไตเซชัน (VQ) ขั้นตอนแรกของกระบวนการ GLA ก็คือการสุ่มตัวอย่างค่าเริ่มต้น จากนั้นก็ทำการแบ่งกลุ่มสีรูปภาพที่ได้จากการกำหนดค่าเริ่มต้นให้เป็นกลุ่มย่อยๆ จนกว่าค่า distortion ที่ได้ค่าใหม่จะเท่ากับหรือใกล้เคียงค่าเก่า (ก่อนที่จะมีการปรับปรุง) จึงจะเสร็จสิ้นการทำงาน ดังแสดงในรูปที่ 4.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



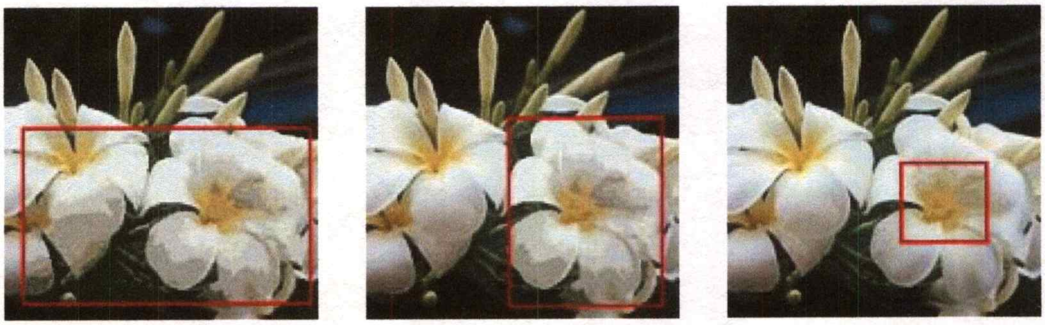
รูปที่ 4.3 โฟล์วไดอะแกรม (Flow Diagram) ของ GLA

วิธีหนึ่งที่ใช้ในการกำหนดค่าเริ่มต้นคือการสุ่มตัวอย่าง (random) หาค่าเริ่มต้น แต่การหาค่าเริ่มต้นแบบสุ่มตัวอย่างนี้ทำให้การแบ่งกลุ่มสีรูปภาพมีผลแตกต่างกันในกรณีวัตถุที่มีสิ่งแวดล้อมแตกต่างกันและรูปภาพที่มีการเปลี่ยนแปลง เช่น การย่อขยายรูปภาพ ขนาดที่ผู้ใช้เลือกพื้นที่ที่สนใจ ซึ่งทำให้มีปัญหาในการเปรียบเทียบความแตกต่างระหว่างกลุ่มที่ผู้ใช้เลือกและกลุ่มที่อยู่ในฐานข้อมูล พราะฉะนั้นในวิทยานิพนธ์นี้จึงใช้การแบ่งสีของรูปภาพให้เป็นไปตามระดับสีที่กำหนด เพื่อเป็นค่าเริ่มต้นของการแบ่งแยกประเภทสี ซึ่งสามารถแก้ปัญหาที่กล่าวมาข้างบนได้ ดังแสดงในรูปที่ 4.4



(a) (b) (c)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

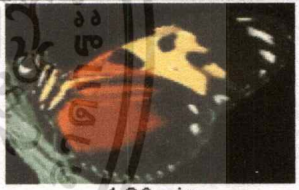
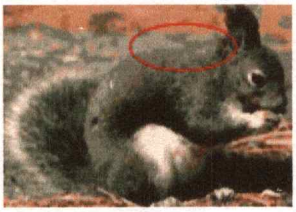


(e) (f) (g)

รูปที่ 4.4 (a) (b) (c) คือรูปภาพที่กำหนดค่าเริ่มต้นด้วยการสุ่มตัวอย่าง และ (e) (f) (g) คือรูปภาพที่กำหนดค่าเริ่มต้นด้วยการกำหนดระดับสี

โดยวิทยานิพนธ์นี้ได้แบ่งค่า RGB ของรูปภาพออกเป็นเท่ากับ 125 (5x5x5) ระดับมาใช้ในการกำหนดค่าเริ่มต้นคือการสุ่มตัวอย่าง โดยค่าดังกล่าวได้มาจากการทดลอง ดังแสดงในตาราง 4.1

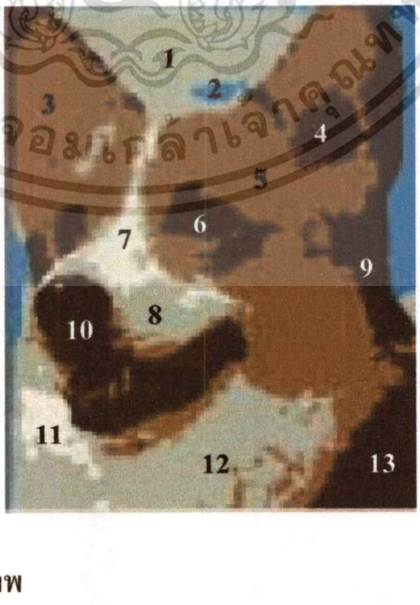
ตารางที่ 4.1 แสดงผลการทดสอบโปรแกรมเพื่อหาค่าระดับสีที่ดีที่สุดสำหรับกำหนดค่าเริ่มต้น

3x3x3 bin เท่า 27 ระดับ	5x5x5 binเท่า125 ระดับ	7x7x7 bin เท่า1343 ระดับ
1:35min	1:33min	 1:38min
1:35min	1:30min	1:36min
1:41min	1:39min	1:47min
 2:25min	2:14min	2:21min

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งาน เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากผลการทดลองเราเห็นได้ว่า ถ้าเราเลือกค่าจำนวน bin เท่ากับ (3x3x3) คือแบ่งค่า RGB ของรูปภาพออกเป็น 27 ระดับ ให้เป็นค่าเริ่มต้น เราจะเห็นว่าการแบ่งกลุ่มรูปภาพไม่สามารถแยกส่วนที่มีสีใกล้เคียงกันได้ และต้องเสียเวลาในการแบ่งกลุ่มดังกล่าวออกเป็นกลุ่มย่อยๆ จนกว่าค่า distortion ที่ได้ค่าใหม่จะเท่ากับหรือใกล้เคียงค่าเก่า ถ้าเราเลือกค่าจำนวน bin เท่ากับ (7x7x7) คือแบ่งค่า RGB ของรูปภาพออกเป็น 343 ระดับ จะเห็นว่าการแบ่งกลุ่มรูปภาพสามารถแยกระดับสีได้ละเอียดขึ้น ซึ่งทำให้ประสิทธิภาพในการแบ่งกลุ่มรูปภาพดีขึ้น แต่การที่เพิ่มระดับสีมากเกินความจำเป็นจะทำให้เราได้กลุ่มสีมาก ซึ่งจะทำให้เสียเวลาในการรวมกลุ่มที่มีค่าความแตกต่างระหว่างสองค่า centroids ตามค่าที่ตั้งไว้เพื่อให้ได้กลุ่มสีตามความต้องการ จากผลการทดลองเห็นได้ว่า ค่า bin เท่ากับ 5 คือแบ่งค่า RGB ของรูปภาพออกเป็น 125 ระดับ ให้เป็นค่าเริ่มต้นจะให้การประสิทธิภาพในการแบ่งกลุ่มรูปภาพใกล้เคียง กับค่า bin ที่ใหญ่กว่า แต่จะใช้เวลาน้อยกว่า

ภายหลังขบวนการกำหนดค่าเริ่มต้นของการแบ่งแยกประเภทสีแล้วก็ทำการแบ่งกลุ่มสีรูปภาพที่ได้จากการกำหนดค่าเริ่มต้นให้เป็นกลุ่มย่อยๆ โดยใช้ วิธีของ GLA ซึ่งจำนวนกลุ่มได้ถูกแบ่งออกเป็นกลุ่มเล็กๆ มากกว่าความต้องการจริง ความถูกต้องของการประมาณไม่สำคัญมากนักเพราะว่ากลุ่มเหล่านั้นจะถูกรวมกันในขั้นตอนต่อไป ภายหลังการทำ GLA จะมีกลุ่มสีที่มีค่าคล้ายกันมากกว่าหนึ่งกลุ่มเพราะว่าการทำ GLA มีจุดประสงค์อยู่ที่การลดค่า distortion ทั้งหมด ดังนั้นจึงใช้เทคนิค agglomerative clustering algorithm[7] ในการรวมกลุ่มที่มีค่าความแตกต่างระหว่างสองค่า centroids ตามค่าที่ตั้งไว้เพื่อให้ได้กลุ่มสีตามความต้องการ ภายหลังขบวนการการแบ่งสีรูปภาพแล้วจึงทำการกำหนดหมายเลขให้แต่ละกลุ่มสี (region labels) ดังแสดงในรูปที่ 4.5



รูปที่ 4.5 ตัวอย่างการแบ่งกลุ่มรูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2 การคำนวณหาคุณลักษณะสำคัญของข้อมูลรูปภาพ

หลังจากการแบ่งกลุ่มให้กับรูปภาพแล้วฮิสโตแกรมของสีและความสัมพันธ์เชิงพื้นที่ระหว่างคู่สี ในกลุ่มย่อยถูกนำมาใช้เป็นดัชนีรูปภาพ[4,8] ฮิสโตแกรมของสีคือการหาค่าความน่าจะเป็นของสีต่างๆ ในกลุ่มย่อย ถ้ากำหนดให้ $H_{ci}(I)$ แทนจำนวนพิกเซลในกลุ่มย่อย I ที่มีสีเป็น ci โดยที่ $h_{ci}(I)$ แทนฮิสโตแกรมของกลุ่มย่อย I ที่มีสี c_i และ n คือจำนวนพิกเซลในกลุ่มย่อย I โดยในวิทยานิพนธ์นี้ได้แบ่งค่า RGB ของรูปภาพออกเป็น 64 ระดับเพื่อลดความซับซ้อนและเวลาในการทำดัชนี [5]

$$h_{ci}(I) = \frac{H_{ci}(I)}{n} \qquad (4.1)$$

โครรีโลแกรมสี คือกราฟที่แสดงความสัมพันธ์เชิงพื้นที่ระหว่างคู่สี ตัวอย่างของการหาโครรีโลแกรมของสีกำหนดให้

I คือกลุ่มพื้นที่ย่อย

m คือจำนวนของสีในภาพโดยมีเซตของสีเป็น $\{c_1,\ldots,c_m\}$

p คือพิกเซลที่จุด (x,y)

I_c คือเซตของกลุ่มพื้นที่ย่อยที่มีสีเท่ากับ c

$\Gamma^{(k)}_{c_ic_j}(I)$ คือจำนวนของคู่สี C_i และ C_j ที่มีระยะห่าง k

$\gamma^{(k)}_{c_i,c_j}(I)$ คือโครรีโลแกรมของกลุ่มพื้นที่ย่อย I

$\alpha^{(k)}_c(I)$ คือออร์โตโครรีโลแกรมของกลุ่มพื้นที่ย่อย I

[n] คือเซตของระยะห่างระหว่างพิกเซล2 พิกเซลที่มีค่าเป็น {1,2…,n }

สามารถคำนวนได้ดังนี้ ถ้ากำหนดให้พิกเซล $p_1 = (x_1, y_1)$ และ $p_2 = (x_2, y_2)$ ดังนั้นระยะห่างระหว่างพิกเซล p_1 กับ p_2 คือ $|p_1-p_2|$ เท่ากับ max $\{|x_1-x_2|,|y_1-y_2|\}$ ตัวอย่าง ถ้าหยิบพิกเซลใดๆ ของสี ci ในรูปภาพขึ้นมาค่าโครรีโลแกรมของสีที่คำนวณคือค่าความน่าจะเป็นที่จะพบพิกเซลที่มีสี cj ห่างจากพิกเซลของสี ci ที่หยิบขึ้นมาเท่ากับ k หรืออาจจะกล่าวอีกอย่างหนึ่งว่าค่าโครรีโลแกรมของสีที่คำนวณได้คือค่าความน่าจะเป็นที่จะพบสี cj ที่ระยะห่าง k จากพิกเซลของสี ci

ถ้ากำหนดให้ i,j $\in$ [m] และ k $\in$ [n] ดังนั้นโครรีโลแกรมของกลุ่มพื้นที่ย่อย I คือ

$$\gamma^{(k)}_{c_i,c_j}(I) \cong \Pr_{p_1 \in I_{ci},\, p_2 \in I_{cj}} [p_1 \in I_{c_i}, p_2 \in I_{c_j}, | p_1 - p_2 | = \kappa]^{\gamma^{(k)}_{c_i,c_i}(I)} \qquad (4.2)$$

$$\gamma^{(k)}_{c_ic_j}(I) \cong \frac{\Gamma^{(k)}_{c_ic_j}(I)}{h_{C_i}(I).8k} \qquad (4.3)$$

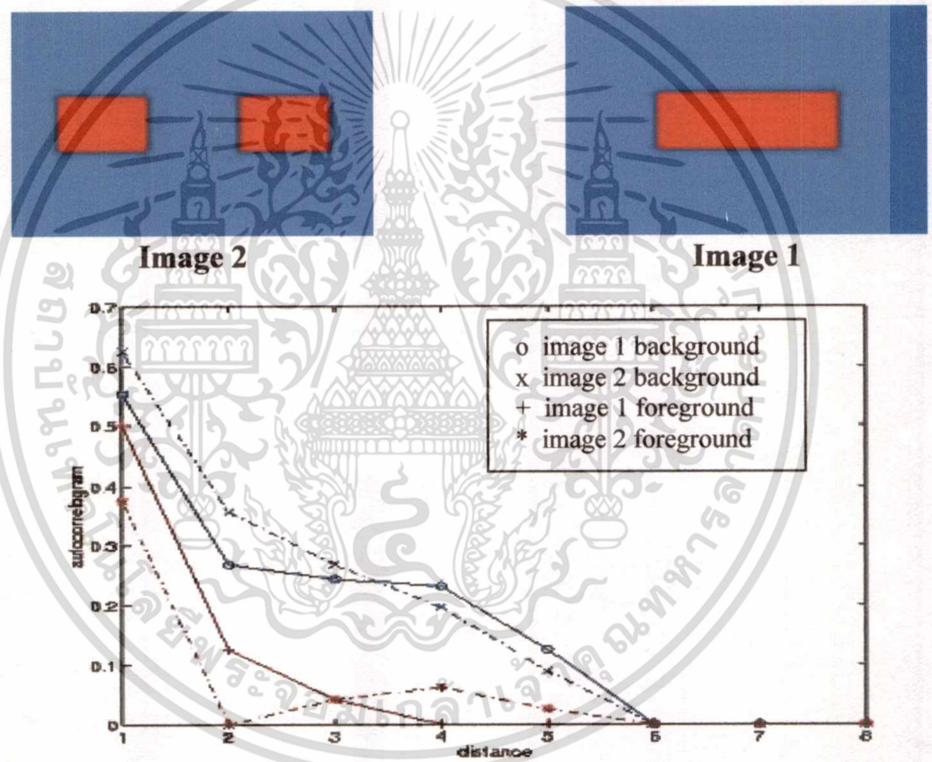
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในวิทยานิพนธ์นี้ได้นำ ความสัมพันธ์เชิงพื้นที่ระหว่างคู่สี มาใช้ร่วมกับ ฮีสโตแกรมสี เพื่อทำดัชนีรูปภาพให้แต่ละรูปภาพในฐานข้อมูล

ออร์โตโครรีโลแกรมเป็นความสัมพันธ์เชิงพื้นที่ระหว่างคู่สีที่เหมือนกัน โดยที่ออร์โตโครรีโลแกรมของกลุ่มพื้นที่ย่อย I ที่ $i \in [m]$ และ $k \in [n]$ คือ

$$\alpha_c^{(k)}(I) \cong \gamma_{c,c}^{(k)}(I) \tag{4.4}$$

ตัวอย่างออร์โตโครรีโลแกรมของรูปตัวอย่าง 2 รูปที่มีฮิสโตแกรมของสีเหมือนกัน ซึ่งมีเพียง 2 สีคือ ม่วงเป็นพื้นหลังกับสีเหลืองเป็นพื้นหน้า และ มี k ={1,2,3,4,5,6,7,8,9,10} แสดงดังรูปที่ 4.6

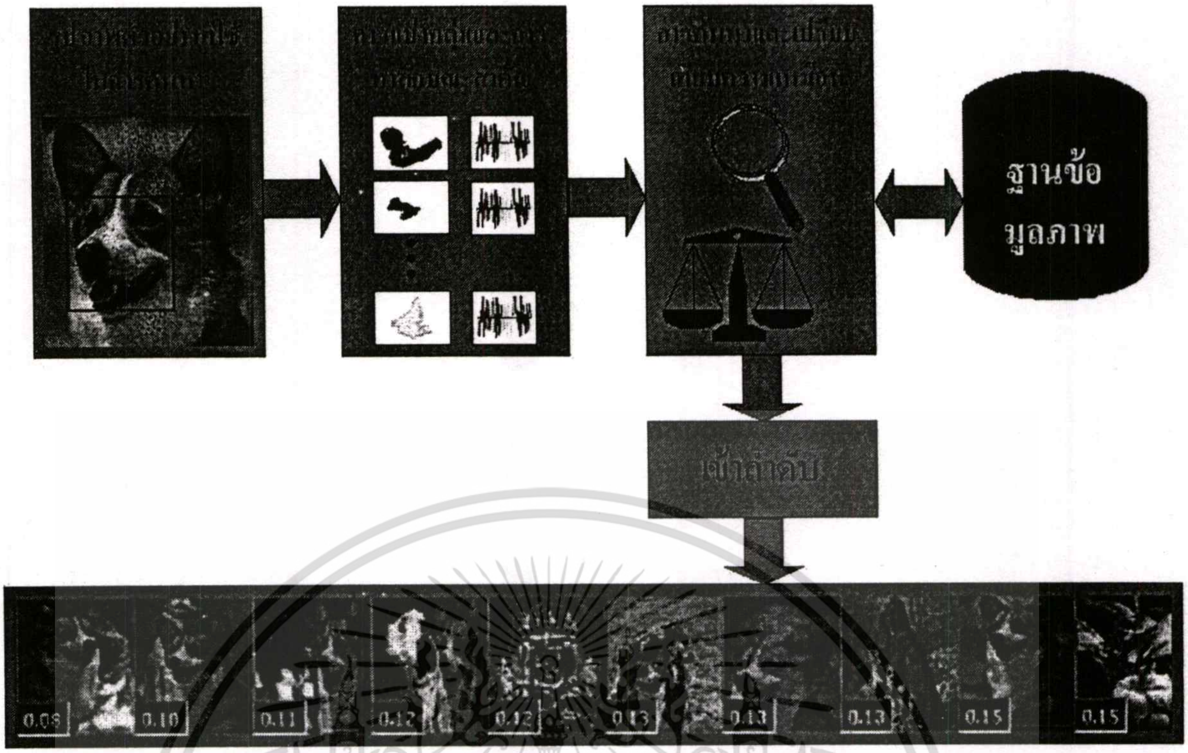


รูปที่ 4.6 รูปภาพตัวอย่างและออร์โตโครรีโลแกรมของรูปภาพตัวอย่าง

4.3 การเสนอแนวทางในการค้นคืนข้อมูลรูปภาพ

ระบบค้นคืนข้อมูลรูปภาพที่นำเสนอในวิทยานิพนธ์นี้ ผู้ใช้สามารถกำหนดพื้นที่ที่ต้องการได้โดยตีกรอบสี่เหลี่ยมบนรูปภาพตัวอย่าง พื้นที่ที่ถูกเลือกจะถูกแบ่งส่วนให้เป็นจำนวน N กลุ่ม โดยใช้ เทคนิคเดียวกันกับการทำดัชนีรูปภาพจากนั้นลักษณะสำคัญของสี และ ความสัมพันธ์เชิงพื้นที่ระหว่างคู่สีในกลุ่มที่ผู้ใช้เลือกจะถูกดึงมาเพื่อนำมาเปรียบเทียบความเหมือนกับกลุ่มย่อยของรูปภาพที่เป็นเป้าหมายในฐานข้อมูล ดังแสดงในรูปที่ 4.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.7. ขบวนการค้นคืนข้อมูลรูปภาพ

4.3.1 วิธีวัดความแตกต่างของรูปภาพต้นฉบับและรูปภาพในฐานข้อมูล

การวัดความแตกต่าง ในวิทยานิพนธ์นี้ได้ใช้เทคนิคประยุกต์การวัดความแตกต่างแบบ L1 (modified distance measure)[9] สำหรับการวัด ความแตกต่างของ ฮีสโตแกรมสี $|I-I'|_{h_1d_1}$ และออร์โตโครรีโลแกรม $|I-I'|_{\gamma_1d_1}$ ดังต่อไปนี้

$$|I-I'|_{h_1d_1} \underline{\Delta} \sum_{i\in[m]} \frac{|h_{c_i}(I)-h_{c_i}(I')|}{1+h_{c_i}(I)+h_{c_i}(I')} \qquad (4.5)$$

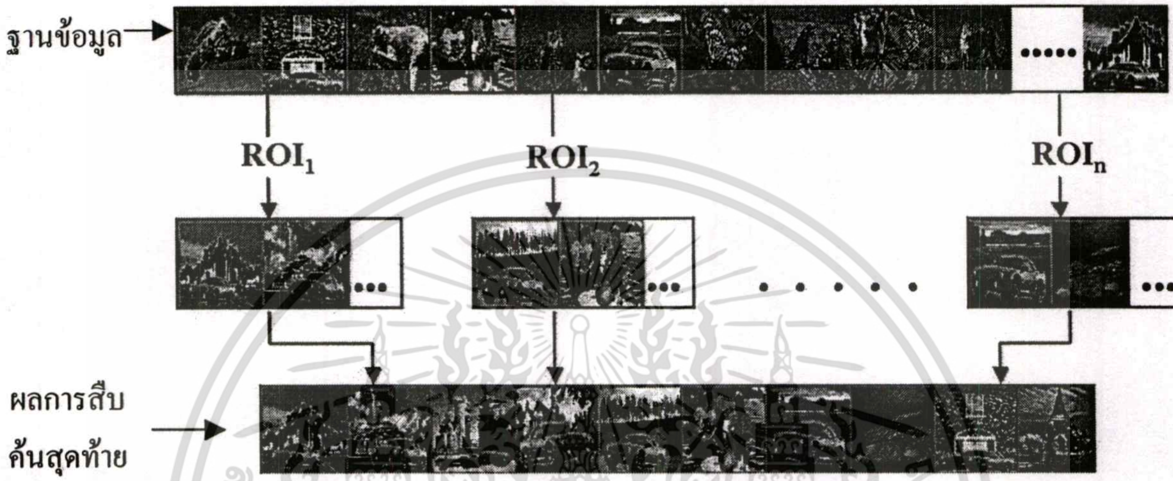
$$|I-I'|_{\gamma_1d_1} \underline{\Delta} \sum_{i\in[m]} \frac{|\gamma^{(k)}_{c_i,c_j}(I)-\gamma^{(k)}_{c_i,c_j}(I')|}{1+\gamma^{(k)}_{c_i,c_j}(I)+\gamma^{(k)}_{c_i,c_j}(I')} \qquad (4.6)$$

โดยที่ I และ I' คือรูปภาพตัวอย่างและรูปภาพในฐานข้อมูล $h_{ci}(I)$ และ $h_{ci}(I')$ คือ ฮีสโตแกรมสี และ $\gamma^{(k)}_{C_i,C_j}(I)$ และ $\gamma^{(k)}_{C_i,C_j}(I')$ คือความสัมพันธ์เชิงพื้นที่ระหว่างคู่สีของรูปภาพ I และ I' ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

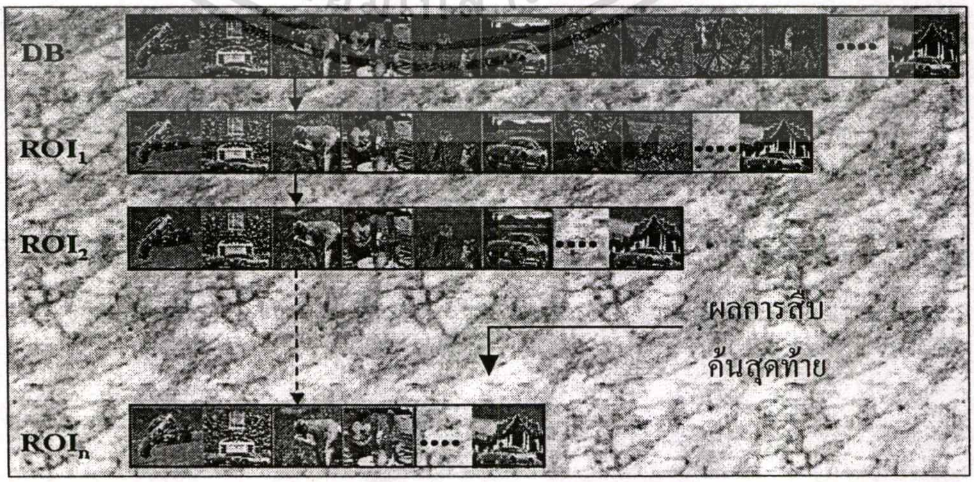
4.3.2 วิธีการค้นหารูปภาพจากฐานข้อมูล

เทคนิคในการค้นหารูปภาพจากฐานข้อมูลที่นิยมใช้คือ การสืบค้นข้อมูลแบบ static และแบบเป็นลำดับ โดยการสืบค้นข้อมูลแบบ static คือการค้นหาเพื่อให้ได้ผลลัพธ์ทุกๆ กลุ่มที่สนใจ แล้วนำมาทำเป็นผลตอบค่าเดียว ดังแสดงในรูปที่ 4.8 การสืบค้นข้อมูลแบบ static นี้จะเห็นว่าต้องการเวลาในการค้นหารูปภาพมาก



รูปที่ 4.8 ผลลัพธ์สืบค้นด้วยการวิเคราะห์แบบเป็น static ในพื้นที่ที่สนใจ

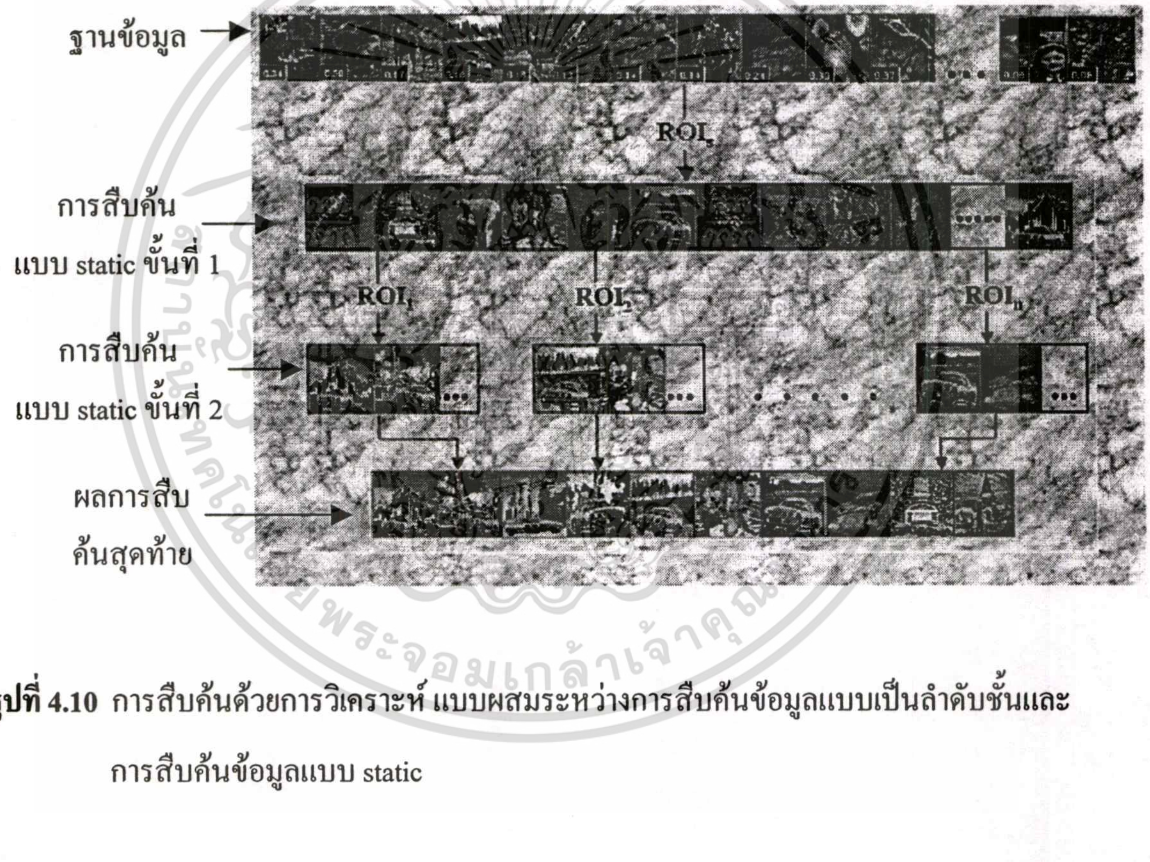
การสืบค้นข้อมูลแบบเป็นลำดับชั้น ระบบจะคืนรายการของรูปภาพที่จับคู่ได้จากการค้นหาในกลุ่มแรก จากนั้นกลุ่ม 2 ใช้เพื่อทำการค้นคืนข้อมูลจากผลที่ได้จากการค้นคืนก่อนหน้านี้และทำเช่นนี้ต่อไปจนกระทั่งครบทุกๆ กลุ่มที่ผู้ใช้เลือกทั้งหมด เมื่อทำการสืบค้นครบทุกกลุ่มพื้นที่สนใจแล้ว ผลลัพธ์สืบค้นสุดท้ายจะแสดงให้แก่ผู้ใช้ ดังแสดงในรูปที่ 4.9



รูปที่ 4.9 ผลลัพธ์สืบค้นด้วยการวิเคราะห์แบบเป็นลำดับชั้นของพื้นที่ที่สนใจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถึงแม้ว่าการสืบค้นข้อมูลแบบเป็นลำดับชั้นจะต้องการเวลาน้อยกว่าการสืบค้นข้อมูลแบบ static แต่เป็นไปได้ที่จะเกิดความผิดพลาดในลำดับชั้นต่างๆ ที่ทำให้เกิดมีข้อผิดพลาดในผลลัพธ์สุดท้ายได้ เพราะฉะนั้นในวิทยานิพนธ์นี้ได้นำเสนอการค้นคืนข้อมูลแบบผสมระหว่างการสืบค้นข้อมูลแบบเป็นลำดับชั้นและการสืบค้นข้อมูลแบบ static ทำให้ลดเวลาในการค้นหาเปรียบเทียบกับการสืบค้นข้อมูลแบบ static อย่างเดียวและลดความผิดพลาดการค้นหาข้อมูลแบบเป็นลำดับชั้นแบบอย่างเดียว การค้นหาข้อมูลแบบผสมนี้ ขั้นตอนแรก กลุ่มที่ใหญ่ที่สุดและอยู่บริเวณตรงกลางของพื้นที่ที่ผู้ใช้เลือกจะถูกนำมาใช้ในการค้นหารูปภาพแบบ static ก่อน ขั้นตอนที่สองคือนำเอากลุ่มที่เหลือทั้งหมดที่ผู้ใช้เลือกมาค้นหารูปภาพในชุดรูปภาพที่ได้ผลมาจากขั้นตอนแรกโดยใช้วิธีค้นหาแบบ static นำผลที่ได้ทั้งหมดจากการค้นหาแบบstaticในขั้นที่สองมาทำเป็นผลตอบค่าเดียว ดังแสดงในรูปที่ 4.10



รูปที่ 4.10 การสืบค้นด้วยการวิเคราะห์ แบบผสมระหว่างการสืบค้นข้อมูลแบบเป็นลำดับชั้นและการสืบค้นข้อมูลแบบ static

ROIs คือกลุ่มที่ใหญ่ที่สุดและอยู่บริเวณตรงกลางของพื้นที่ที่ผู้ใช้เลือก ROI_1 , ROI_2 , ROI_3 , ... ROI_n คือกลุ่มที่เหลือทั้งหมดที่ผู้ใช้เลือก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

การทดลองและผลการทดลอง

5.1 บทนำ

ในวิทยานิพนธ์นี้ได้ทำการทดลองค้นหารูปภาพโดยใช้วิธีการที่ได้นำเสนอมาคือทำดัชนีแบบ Region โดยใช้ลักษณะสีร่วมกับความสัมพันธ์เชิงพื้นที่ระหว่างคู่สี โดยทำการเปรียบเทียบผลการทดลองกับการค้นหารูปภาพโดยใช้ดัชนีแบบ global Histogram และ แบบ region Histogram โดยใช้โปรแกรมวิชวลเบสิก เวอร์ชั่น 6 (Visual Basic version 6) ในการประมวลผล และฐานข้อมูลของไมโครซอต์ฟเอ็กเซส 2000 (Microsoft Access 2000) ใช้เป็นฐานข้อมูล

5.2. การทดลองการค้นหารูปภาพ และ ผลการทดลอง

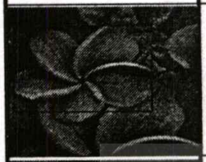
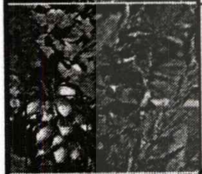
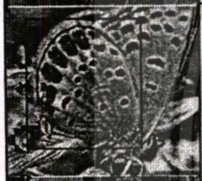
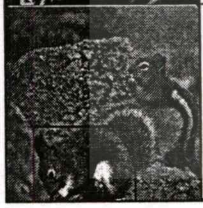
ในวิทยานิพนธ์นี้ได้พิจารณาสีของรูปภาพเป็น 64 สี ดังนั้นความแตกต่างของสีจะแบ่งออกเป็น 64 ระดับ และ เซตของค่า k={1,2,3,4,5,6,7,8,9,10} ซึ่งฐานข้อมูลรูปภาพมีจำนวนรูปทั้งหมด 500 รูป ประกอบด้วยรูปภาพ 50 ชุด เช่น กระรอก, ม้า, รถ, ปลา, ผีเสื้อ, เสือ, ดอกกุหลาบ, ดอกไม้ป่า, ฟักทอง, คน เป็นต้น โดยรูปภาพในแต่ละชุดมีสิ่งแวดล้อมหรือสถานที่ที่แตกต่างกัน รวมไปถึงรูปภาพที่มีการเปลี่ยนแปลงไปด้วยรูปแบบต่างๆ เช่น การย่อขยาย การหมุนรูปภาพ เป็นต้น

ในการทดลองนี้ได้ใช้รูปภาพจำนวน 250 รูปภาพ โดยสุ่มเลือกมาจากแต่ละชุด ชุดละ 5 รูปภาพสำหรับค้นคืนเพื่อค้นหารูปภาพใช้ค่าความต่างเท่ากับ 0.3

ค่าความต่างที่เหมาะสม คือค่าความต่างที่น้อยที่สุดของคู่ฮิสโตแกรมที่ยอมรับว่าภาพที่เป็นตัวอย่างและภาพที่เป็นเป้าหมายของคู่ฮิสโตแกรมนั้นคล้ายกัน การหาค่าความต่างที่เหมาะสมนี้ทำได้โดยการทดสอบโปรแกรมการค้นหารูปภาพหลายๆ ค่าความต่างเพื่อดูว่าค่าความต่างของฮิสโตแกรมใดที่ยังคงให้ภาพตัวอย่างที่คล้ายกับภาพเป้าหมาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 5.1 แสดงผลการทดสอบโปรแกรมเพื่อหาค่าความต่างโดยรวมที่ดีที่สุด

Query	ค่าความต่าง 0.5			ค่าความต่าง 0.3			ค่าความต่าง 0.1		
	จำนวนภาพที่ถูก	จำนวนภาพที่ผิด	เวลาที่ใช้ (min)	จำนวนภาพที่ถูก	จำนวนภาพที่ผิด	เวลาที่ใช้ (min)	จำนวนภาพที่ถูก	จำนวนภาพที่ผิด	เวลาที่ใช้ (min)
	10	0	3:04	10	0	2:39	9	1	1:37
	8	2	4:47	9	1	4:52	1	8	2:12
	9	1	3:42	9	1	2:09	0	10	1:24
	8	2	10:20	8	2	10:18	6	4	4:08
	9	1	2:55	9	1	2:42	3	7	1:18
	7	3	4:24	7	3	3:51	6	4	2:36
	5	5	7:54	4	6	5:38	4	6	2:17

จากผลการทดลองในตารางที่ 5.1 เห็นว่า การกำหนดค่าความต่างมากจะทำให้ผลลัพธ์มีความถูกต้อง แต่การที่กำหนดค่าความต่างมากจะทำให้ผลลัพธ์จากการค้นในลำดับแรก คือผลที่ได้จากการค้นด้วยกลุ่มที่ใหญ่ที่สุดและอยู่บริเวณตรงกลางของพื้นที่ที่ผู้ใช้เลือก ได้รูปภาพที่ความคล้ายคลึงมากมาย ซึ่งทำให้เสียเวลามากในการค้นหารูปภาพในระดับที่ 2 กรณีกำหนดค่าความต่างน้อยเกินไปจะทำให้ผลการค้นในลำดับแรกได้เฉพาะรูปภาพที่มีความเหมือนน้อยทำให้ไม่ได้รูปภาพบางรูปที่เป็นเป้าหมายที่ความแตกต่างค่อนข้างใกล้เคียง จากการทดลองเราเห็นว่าค่าความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต่างที่ให้ผลการค้นหารูปภาพมีประสิทธิภาพโดย รวมดีที่สุดคือ 0.3 สามารถค้นหารูปภาพได้ถูกต้อง และใช้เวลาน้อยกว่าค่าความต่างที่มีค่ามาก

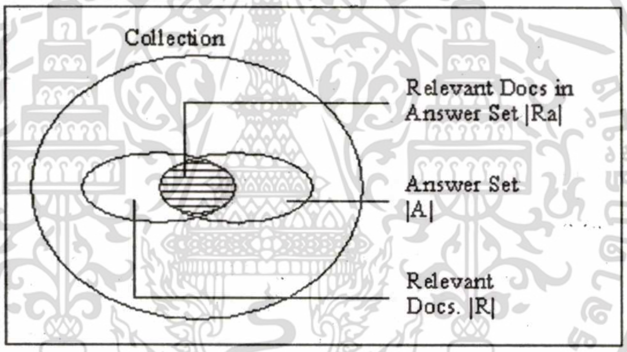
นอกจากนี้ได้นำผลความต่างที่หาได้มาทดลองค้นหารูปโดยใช้วิธีการที่ได้นำเสนอมาคือทำดัชนีแบบ Region โดยใช้ลักษณะสีร่วมกับความสัมพันธ์เชิงพื้นที่ระหว่างคู่สี โดยทำการเปรียบเทียบประสิทธิภาพกับการค้นหารูปภาพโดยใช้ดัชนีแบบ global Histogram และ แบบ region Histogram ในวิทยานิพนธ์นี้ได้ใช้วิธีการวัดประสิทธิภาพของการค้นคืนข้อมูลรูปภาพแบบ recall และ precision

Precision คือผลอัตราส่วนของจำนวนรูปภาพที่ถูกต้อง(Ra) ต่อจำนวนรูปภาพที่ต้องการ(A)

$$\text{Precision} = \frac{|Ra|}{|A|} \qquad (5.1)$$

Recall คือผลหานของจำนวนรูปภาพที่ถูกต้อง(Ra) กับ จำนวนรูปภาพที่มีในฐานข้อมูล

$$\text{Re}\,call = \frac{|Ra|}{|R|} \qquad (5.2)$$



รูปที่ 5.1 ตัวอย่างการให้ข้อมูลสำหรับ Precision และ Recall

การวัดประสิทธิภาพของการค้นคืนข้อมูลรูปภาพ โดยใช้การหาค่าเฉลี่ย precision ณ ระดับ recall r ซึ่งสามารถคำนวณได้ดังนี้

$$\overline{P}(r) = \sum_{i=1}^{N_q} \frac{P_i(r)}{N_q} \qquad (5.3)$$

P(r) คือค่าเฉลี่ย precision ที่ระดับ recall r N_q คือจำนวนของ query ที่ใช้ และ $P_i(r)$คือค่า precision ที่ระดับ recall r สำหรับ query i ใดๆ

ในการทดลองนี้จะแบ่งการค้นหารูปภาพออกเป็น 2 กรณี คือ การค้นหารูปภาพที่มีความเหมือน กัน และ การค้นหารูปภาพที่มีความคล้ายคลึงกัน

การทดลองแบบที่หนึ่งเป็นการทดลองการค้นหารูปภาพที่มีความเหมือนกัน ซึ่งเป็นการเปรียบ เทียบความสามารถในการค้นหารูปภาพที่มีสถานที่ที่แตกต่างกัน ความแตกต่างในแง่ของขนาด มุมมองของภาพและ การปรากฏของวัตถุเพียงบางส่วน ดังแสดงผลในรูปที่ 5.2-5.6

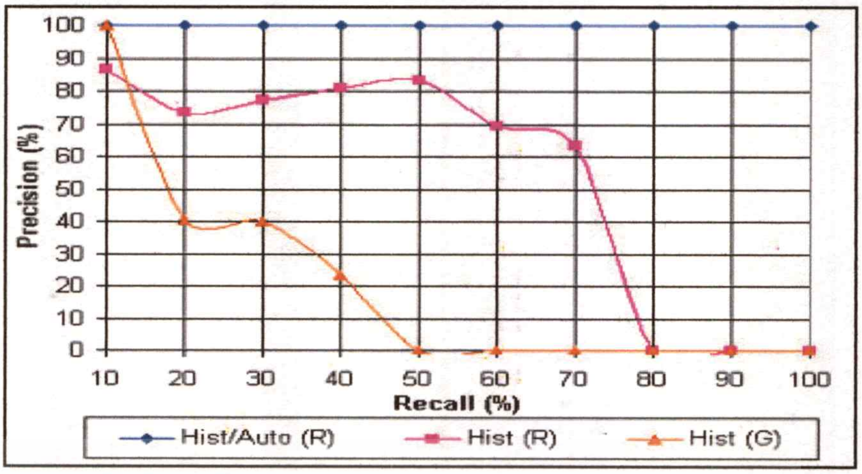
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(a)

(b)

(c)

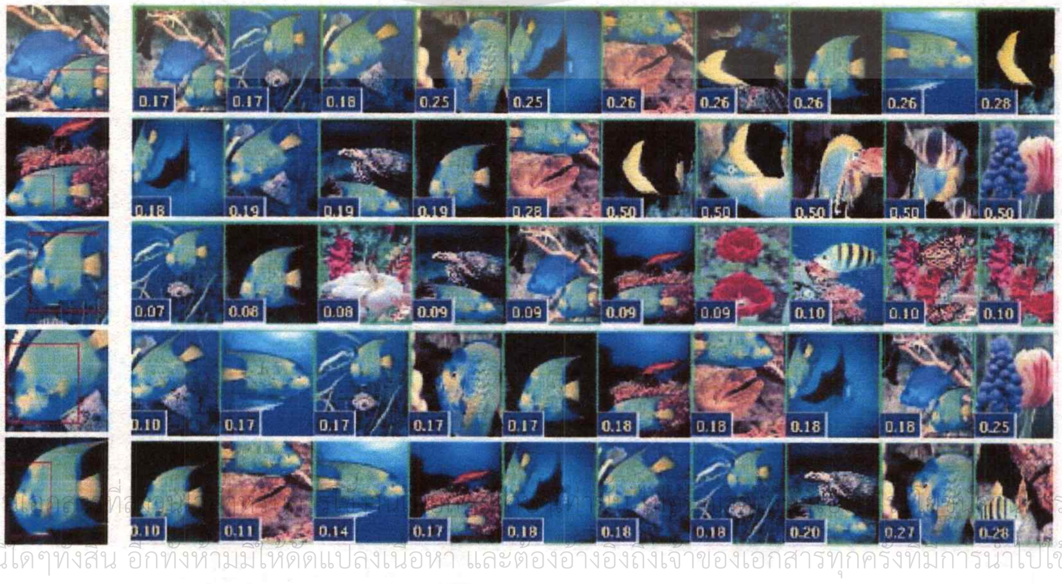
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



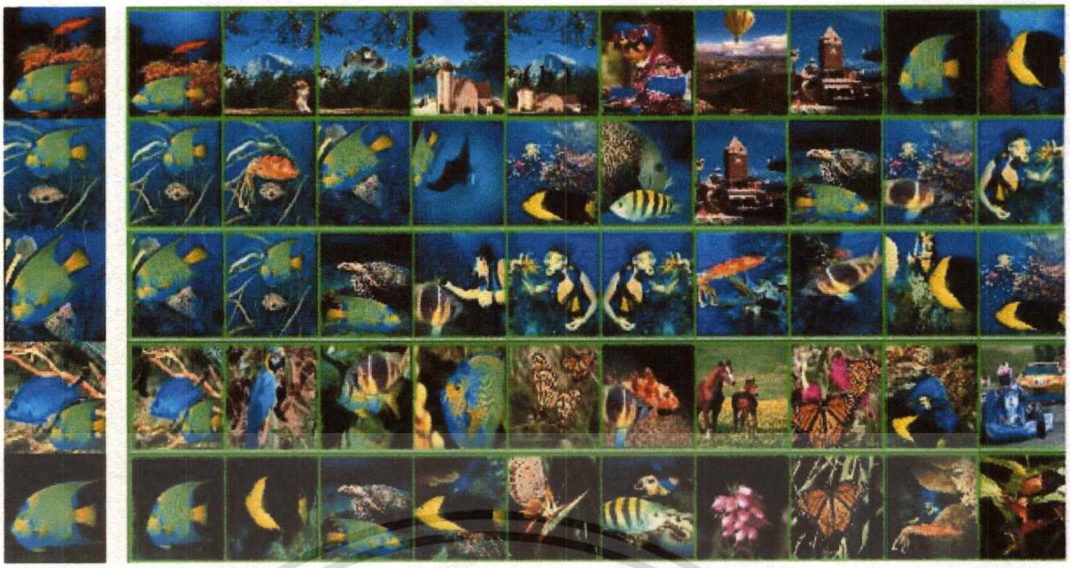
(d)

รูปที่ 5.2 ผลการค้นคืนรูปภาพที่มีความเหมือนกันชุดที่ 1 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

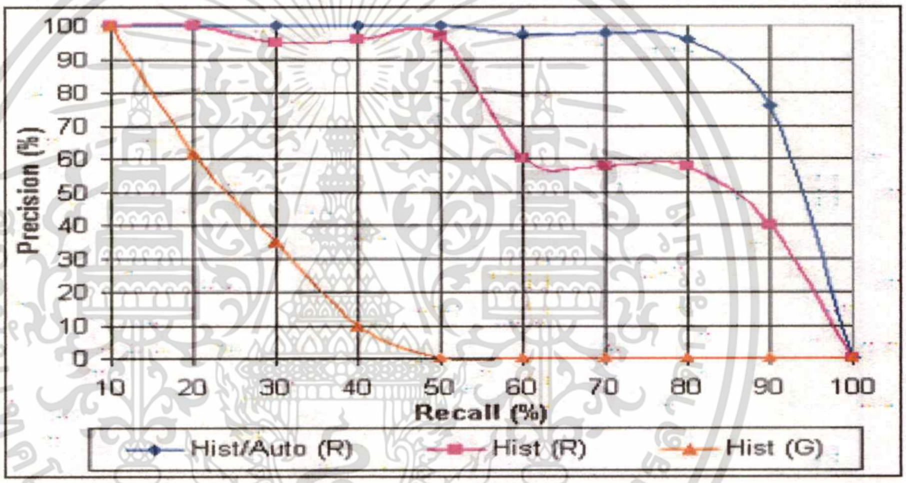
(a)



ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

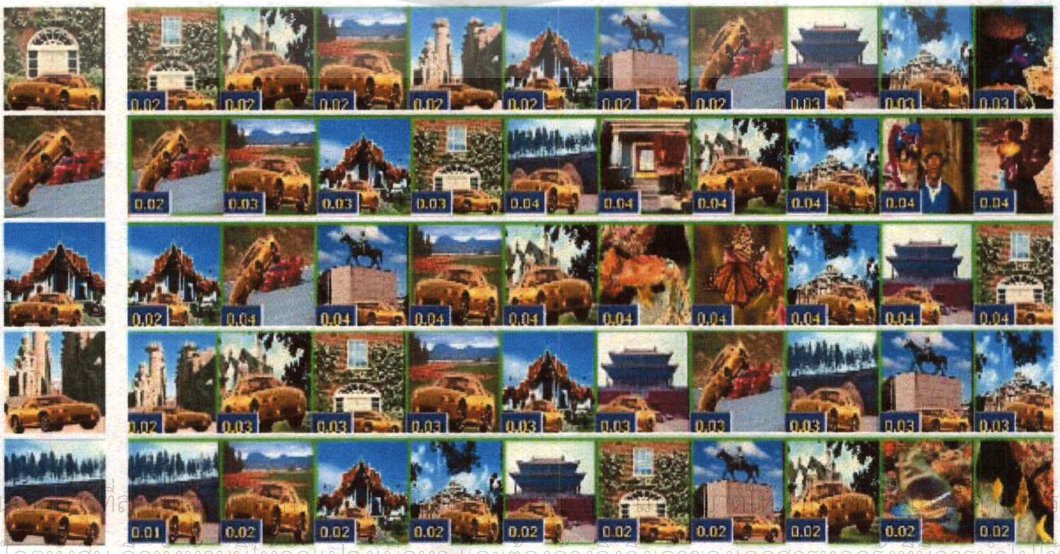


(c)

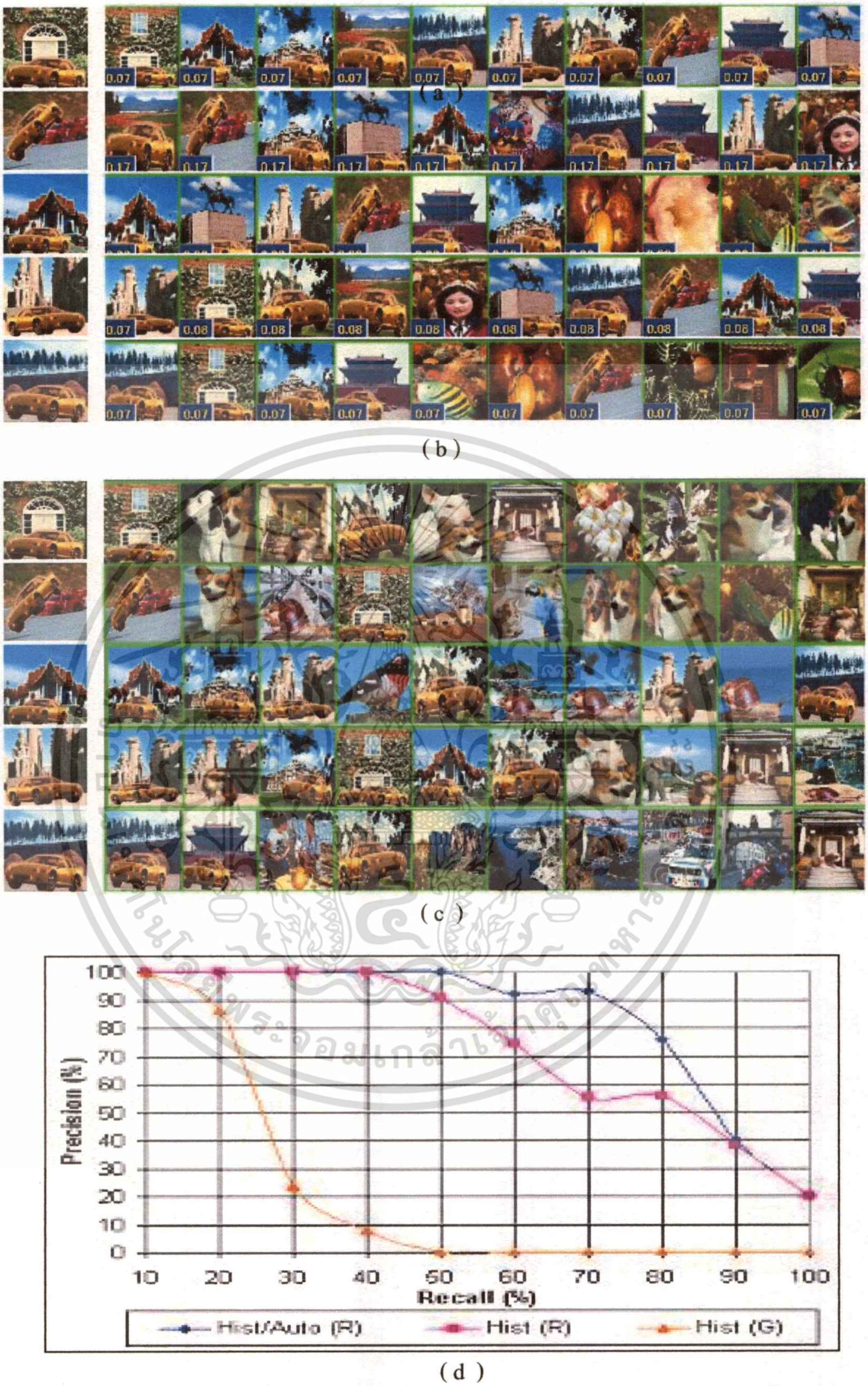


(d)

รูปที่ 5.3 ผลการค้นคืนรูปภาพที่มีความเหมือนกันชุดที่ 2 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3



เอกสารนี้เป็น ... ค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.4 ผลการค้นคืนรูปภาพที่มีความเหมือนกันชุดที่ 3 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

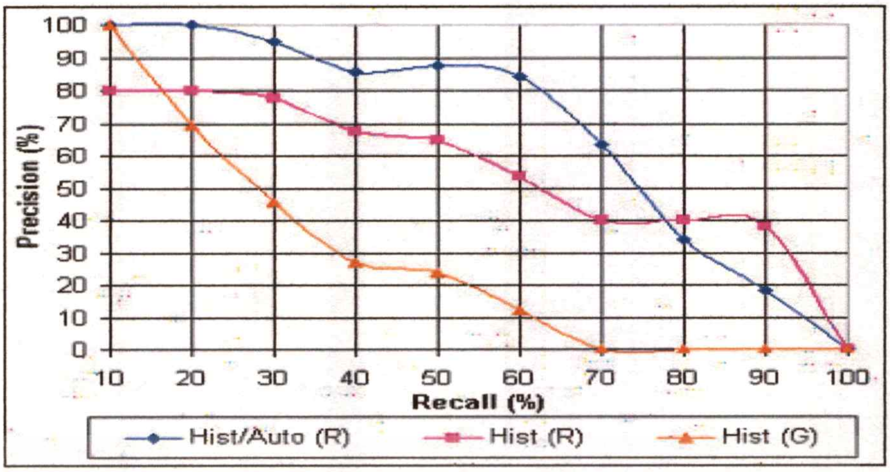
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(a)

(b)

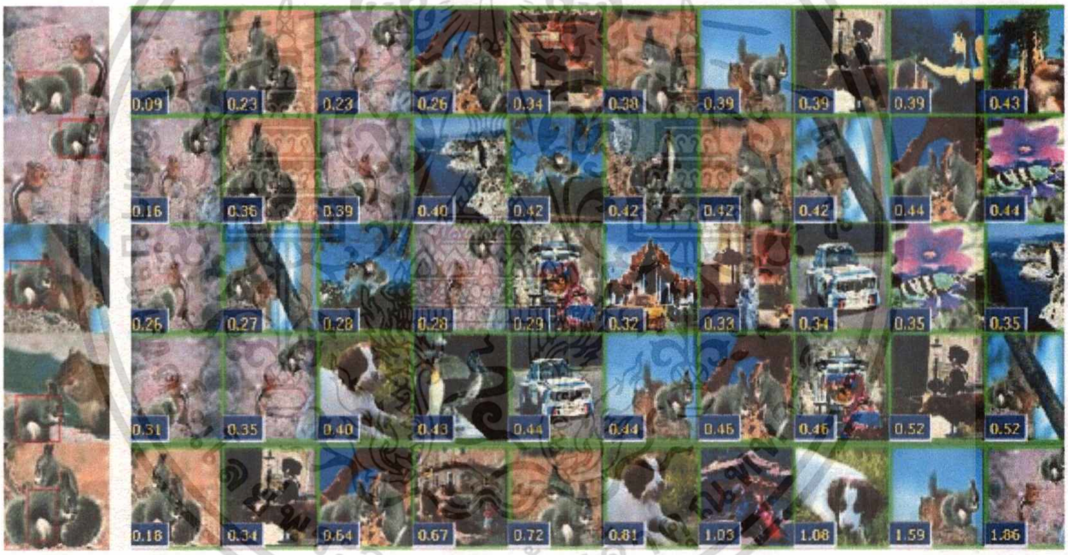
(c)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(d)

รูปที่ 5.5 ผลการค้นคืนรูปภาพที่มีความเหมือนกันชุดที่ 4 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

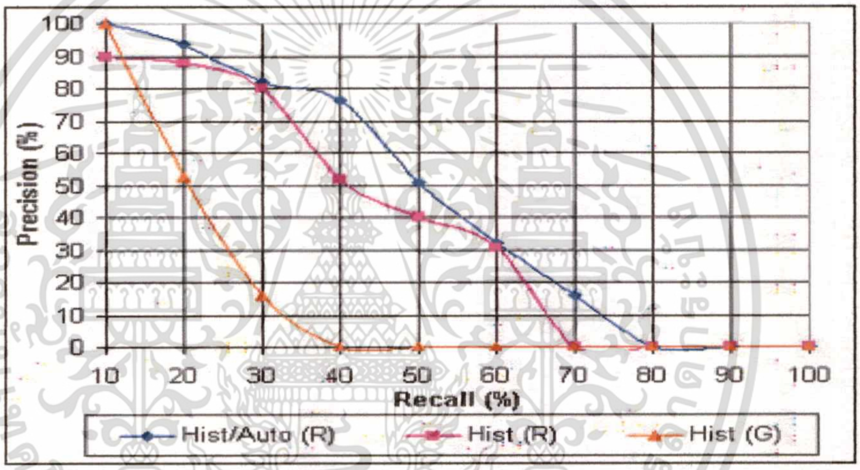


(a)

(b)

เอกสารนี้เป็น ... ค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(c)



(d)

รูปที่ 5.6 ผลการค้นคืนรูปภาพที่มีความเหมือนกันชุดที่ 5 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

จากผลการทดลองในกรณีค้นหารูปภาพที่มีความเหมือนกัน จะเห็นได้ว่าวิธีการที่ได้นำเสนอสามารถค้นรูปภาพได้ถูกต้องมากกว่าวิธีการทำดัชนีแบบ region Histogram และ แบบ global Histogram ดังรูปที่5.2-5.6 เนื่องจากว่า การใช้ฮิสโตแกรมร่วมกับออร์โตโครรีโลแกรมนั้นสามารถแยกความแตกต่างระหว่างกลุ่มๆที่มีฮิสโตแกรมของสีใกล้เคียงกันแต่มีการกระจายของตำแหน่งพิกเซลไม่เหมือนกันทำให้สามารถค้นหารูปภาพในกรณีรูปภาพที่มีสถานที่ที่แตกต่างกัน รูปภาพที่แตกต่างในแง่ของขนาด มุมมองของภาพและการปรากฏของวัตถุเพียงบางส่วนได้ สาเหตุที่ทำให้การค้นหารูปภาพแบบ region Histogram และ แบบ global คือการค้นหาข้อมูลรูปภาพโดยใช้สีอย่างเดียวมีประสิทธิภาพไม่ดี ทั้งนี้ขึ้นอยู่กับข้อจำกัดของตัว color histogram เองซึ่ง

เอกสารนี้เป็นเอกสารที่ลงในเวลาหรือการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไม้สามารถทำการค้นหารูปภาพเช่น ลักษณะของการกระจายสีในสภาวะแวดล้อมแตกต่างกันได้ ทำให้การค้นหารูปภาพยังคงพบรูปภาพที่ไม่เกี่ยวข้องมากและทำให้ผลลัพธ์ของภาพที่ต้องการหรือคาดหมายว่าจะได้ แต่ภาพที่ได้ออกมานั้นจะมีภาพส่วนหนึ่งที่ไม่ตรงกับประเภทของความต้องการออกมาด้วย แม้ว่าการทำดัชนีที่ได้นำเสนอให้ผลดีกว่าวิธีอื่นๆ แต่บางกรณีรูปภาพที่มีความเรียบน้อยทำให้การแบ่งกลุ่มมีความแตกต่างกันมากในกรณีมีการย่อขยายมาก ทำให้ความถูกต้องของการค้นคืนรูปภาพลดลง ดั่งรูปที 5.6

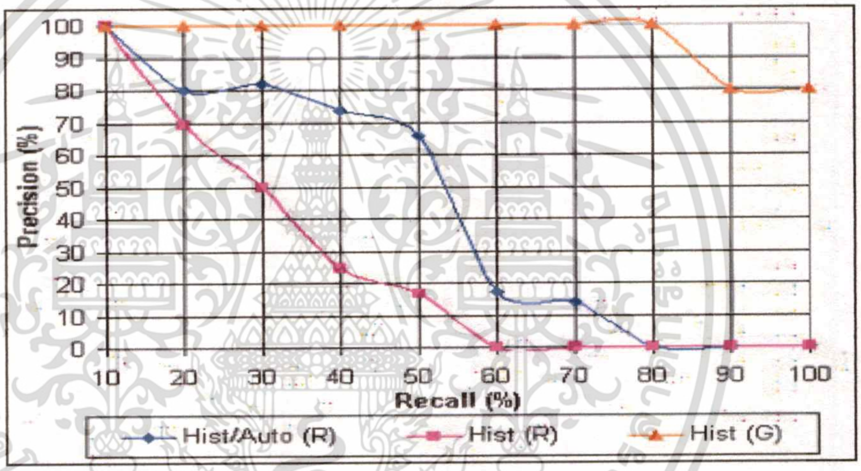
การทดลองแบบที่สองเป็นการทดลองการค้นหารูปภาพที่มีความคล้ายคลึงกัน ซึ่งเป็นการเปรียบ เทียบความสามารถในการค้นหารูปภาพที่มีสีและ การเกาะกลุ่มของสีใกล้เคียงกัน

(a)

(b)

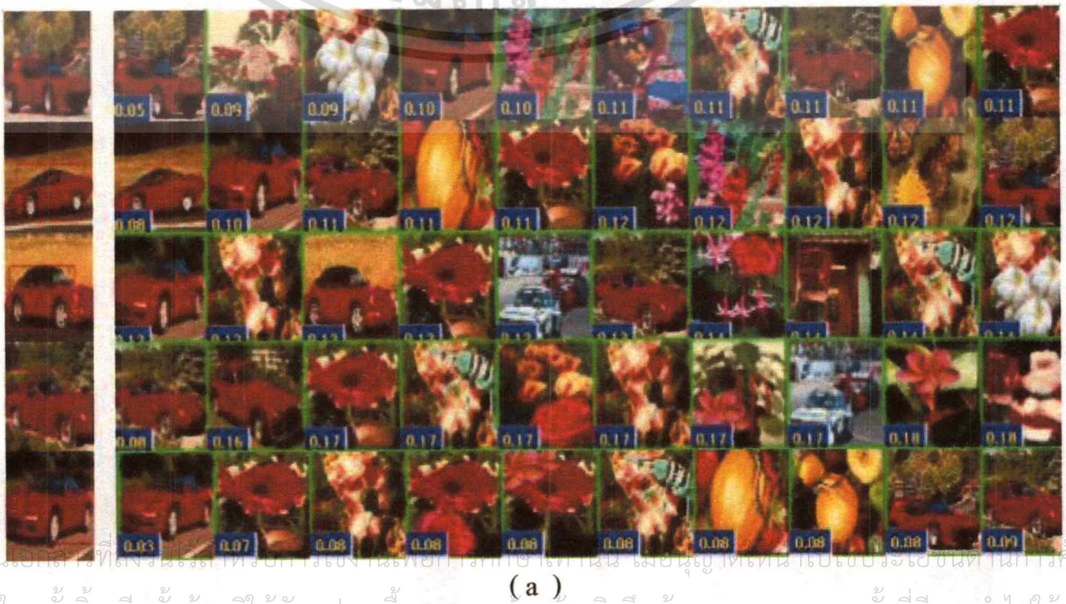
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(c)



(d)

รูปที่ 5.7 ผลการค้นคืนรูปภาพที่คล้ายกันชุดที่ 6 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

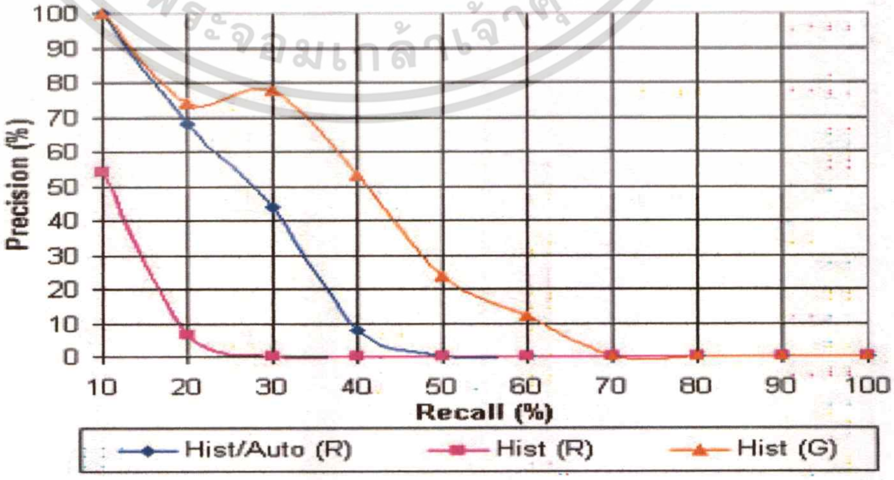


(a)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(b)

(c)



(d)

รูปที่ 5.8 ผลการค้นคืนรูปภาพที่คล้ายกันชุดที่ 7 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

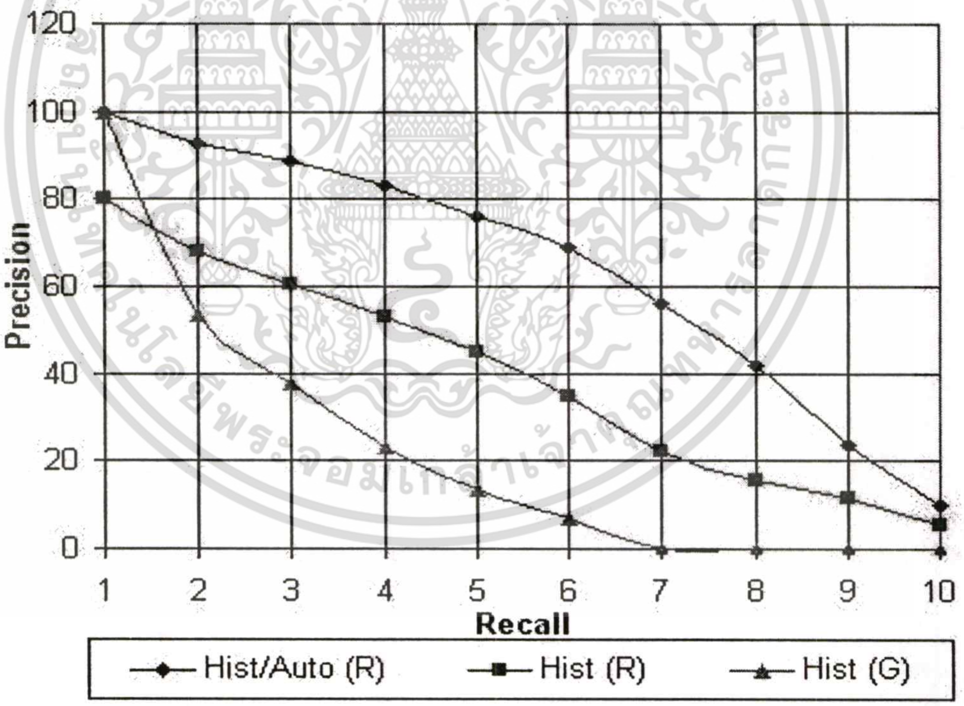
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากผลการทดลองในกรณีค้นหารูปภาพที่มีความคล้ายคลึงกัน จะเห็นได้ว่าวิธีการที่ได้นำเสนอยังมีประสิทธิภาพต่ำกว่าแบบ global Histogram แต่สูงกว่าแบบ region histogram เพียงอย่างเดียว เนื่องจากประสิทธิภาพของระบบที่นำเสนอขึ้นกับปัจจัยหลายประการ เช่น

(1) การแบ่งกลุ่มภาพของพื้นที่ที่ผู้ใช้เลือก มีความแตกต่างจากกลุ่มพื้นที่ที่แบ่งจากภาพในฐานข้อมูล ทำให้การเปรียบเทียบความแตกต่างต่างกันมาก เนื่องจากรูปร่างลักษณะของสิ่งที่ต้องการค้นหาค่อนข้างมีการจัดเรียงสีที่แตกต่างออกไป

(2) ลักษณะของภาพที่นำมาใช้ในฐานข้อมูลสำหรับค้นหาภาพที่คล้ายกันนั้นมีลักษณะของการกระจายสีในสภาวะแวดล้อมคล้ายกันทำให้ global Histogram ให้ผลดีกว่า ถ้าเราเพิ่มภาพในฐานข้อมูลให้มีวัตถุที่สนใจแต่อยู่ในสภาพแวดล้อมแตกต่างไปจากภาพที่นำมาค้นคืนก็จะเห็นได้ชัดว่า feature ที่เราเสนอจะให้ผลดีกว่า

ในวิทยานิพนธ์นี้ได้ทำการค้นหารูปภาพโดยรูปภาพตัวอย่าง 250 รูปภาพจากฐานข้อมูล และนำผลทั้งหมดมาเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยสุ่มเลือกใช้เทคนิคทั้ง 3 ดังแสดงรูปที่ 5.9



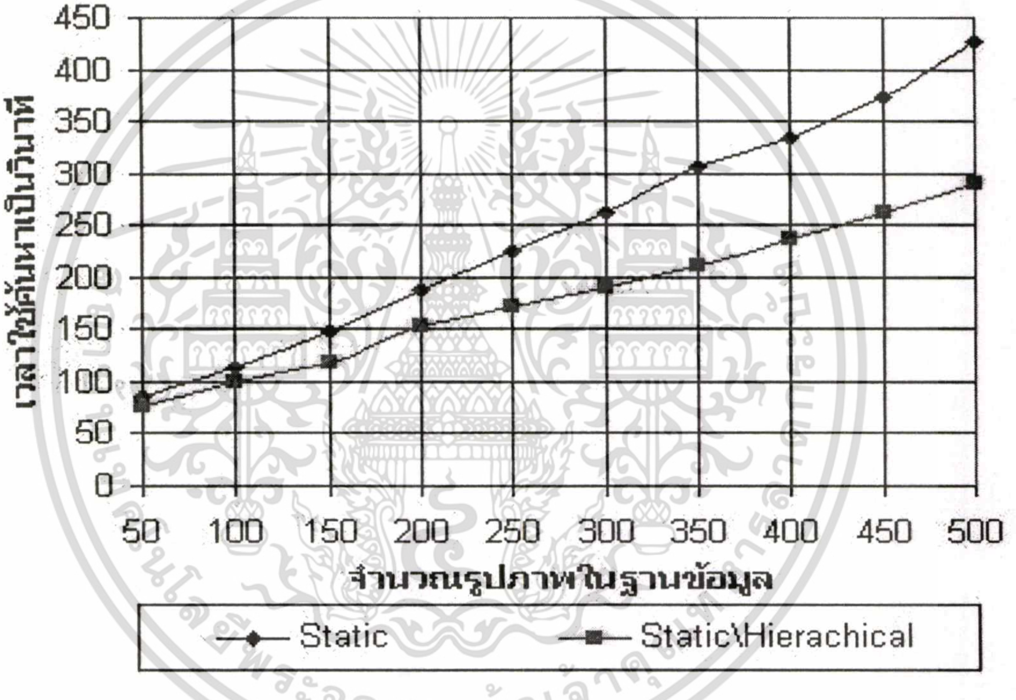
รูปที่5.9 ผลการเปรียบเทียบประสิทธิภาพการค้นคืนภาพทั้งหมด

จากผลการทดลองทั้งหมด ดังแสดงรูปที่5.9 ให้เห็นว่าวิธีที่นำเสนอสามารถค้นคืนข้อมูลรูปภาพได้ถูกต้องมากกว่าวิธี การทำดัชนีแบบ global Histogram และ แบบ region Histogram เนื่องจากว่า การใช้ฮิสโตแกรมร่วมกับออร์โตโครรีโลแกรมนั้นสามารถแยกความแตกต่างระหว่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มๆ ที่มีฮิสโตแกรมของสีใกล้เคียงกันแต่มีการกระจายของตำแหน่งพิกเซลไม่เหมือนกันทำให้สามารถค้นหารูปภาพในกรณีรูปภาพที่มีสถานที่ที่แตกต่างกัน รูปภาพที่แตกต่างในแง่ของขนาด มุมมองของภาพและการปรากฏของวัตถุเพียงบางส่วนได้

นอกจากนี้เรายังทำการเปรียบเทียบเวลาในการค้นหาข้อมูลด้วยวิธีที่นำเสนอคือการค้นหาข้อมูลรูปภาพแบบผสมระหว่างการสืบค้นข้อมูลแบบลำดับชั้นและการสืบค้นข้อมูลแบบ static กับวิธีการสืบค้นข้อมูลแบบ static อย่างเดียว ด้วยการเฉลี่ยเวลา ที่ระดับจำนวนข้อมูลรูปภาพในฐานข้อมูล ในวิทยานิพนธ์นี้ได้กำหนดฐานข้อมูลรูปภาพ เป็น 10 ระดับคือ 50, 100, 150, 200, 250, 300, 350, 400, 450 และ 500 รูปภาพ



รูปที่5.10 ผลการเปรียบเทียบเวลาการค้นคืนภาพด้วยเทคนิคที่นำเสนอกับแบบ Static

จากผลการทดลองดังแสดงรูปที่ 5.10 จะเห็นว่าการค้นหารูปภาพด้วยวิธีที่นำเสนอคือการค้นหาข้อมูลรูปภาพแบบผสมระหว่างการสืบค้นข้อมูลแบบลำดับชั้นและแบบ static ต้องการเวลาน้อยกว่าการสืบค้นข้อมูลแบบ static เพียงอย่างเดียวจะเห็นได้ว่า ถ้าข้อมูลรูปภาพในฐานข้อมูลยิ่งมีจำนวนมากขึ้นเท่าใด การสืบค้นข้อมูลแบบ static จะต้องการเวลามากในการสืบค้นมากขึ้นเท่านั้น เพราะทุกๆกลุ่มพื้นที่ที่ผู้ใช้เลือกจะต้องนำมาเปรียบเทียบกับกลุ่มย่อยของรูปภาพทั้งหมดในฐานข้อมูล สายเหตุที่ทำให้วิธีที่นำเสนอสามารถลดเวลาในการค้นหาได้เพราะสืบค้นที่นำเสนอจะเวลามากเฉพาะลำดับชั้นแรก คือต้องใช้กลุ่มที่ใหญ่ที่สุดและอยู่บริเวณตรงกลางของพื้นที่ที่ผู้ใช้เลือก ค้นหา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปภาพที่มีความเหมือนหรือคล้ายคลึงก่อน จากนั้นก็นำผลที่ได้จากขั้นตอนแรก มาค้นหากับกลุ่มที่เหลือที่ผู้ใช้เลือก ซึ่งทำให้การค้นหาของแต่ละกลุ่มใช้เวลาน้อยกว่าแบบ static

5.3 สรุปผลการทดลอง

จากผลการทดลองค้นหารูปภาพ สามารถสรุปผลการทดลองได้ว่าการค้นหารูปภาพด้วยการใช้ Color Histogram ร่วมกับ Autocorrelogram มีความสามารถในการค้นหารูปภาพที่ลักษณะเหมือนกันได้ตรงตามความต้องการสูง ดังรูปที่ 5.2-5.6 แต่ยังมีประสิทธิภาพต่ำในหาค้นรูปภาพที่มีความคล้ายกันน้อยดังรูปที่ 5.7 และ รูปที่ 5.8 เนื่องจากข้อจำกัดที่ได้กล่าวมา คือประสิทธิภาพในการแบ่งกลุ่มภาพและลักษณะของภาพที่นำมาใช้เป็นฐานข้อมูล นอกจากนี้รูปที่ 5.9 แสดงให้เห็นว่าวิธี region Histogram ร่วมกับ Autocorrelogram จะมีประสิทธิภาพในการค้นคืนรูปภาพโดยเฉลี่ยได้ดีกว่าวิธีอื่นๆ โดยเหตุที่ลักษณะสำคัญที่เลือกมาใช้นี้สามารถจัดการกับการกระจายของสีและความสัมพันธ์เชิงพื้นที่ระหว่างคู่สีในพื้นที่ที่สนใจ ซึ่งสามารถนำผลมาใช้ในการแยกความแตกต่างระหว่างกลุ่มต่างๆ และมีประสิทธิภาพในการค้นคืนรูปภาพในกรณีรูปภาพที่มีสิ่งแวดล้อมหรือสถานที่ที่แตกต่างกันและรูปภาพที่มีการเปลี่ยนแปลงไปได้ทั้งที่วัตถุที่สนใจมีความแตกต่างในแง่ของขนาด มุมมองของภาพและการปรากฏของวัตถุเพียงบางส่วนด้วย นอกจากนั้น การสืบค้นข้อมูลด้วยวิธีที่นำเสนอคือแบบผสมยังทำให้สามารถลดเวลาในการค้นหาได้อีกด้วย ดังแสดงในรูปที่ 5.10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6
สรุปและวิจารณ์

6.1 บทนำ

ในวิทยานิพนธ์นี้ได้นำเอาวิธีการทำดัชนีแบบ region มาใช้ประโยชน์สำหรับค้นหาส่วนของภาพที่สนใจและภาพที่มีสภาพแวดล้อมแตกต่าง โดยนำเอาลักษณะสีของภาพด้วยการทำฮิสโตแกรมของสี (color histogram) เฉพาะแต่ละขอบเขต เนื่องจากว่า เป็นวิธีที่เข้าใจง่ายและสามารถนำไปพัฒนาร่วมกับเทคนิคอื่นๆ เพื่อปรับปรุงประสิทธิภาพให้ดีขึ้นได้ ฮิสโตแกรมของสีจะมีความใกล้เคียงกันแม้ว่าวัตถุในภาพจะมีการเปลี่ยนมุมมอง การหมุน การย่อขยายภาพและความสัมพันธ์เชิงพื้นที่ระหว่างคู่สี(color correlogram) สามารถแยกความแตกต่างระหว่างกลุ่มต่างๆ ที่มีฮิสโตแกรมของสีใกล้เคียงกันแต่มีการกระจายของตำแหน่งพิกเซลไม่เหมือนกัน โดยใช้วิธีการสืบค้นข้อมูลแบบผสมระหว่างการสืบค้นข้อมูลแบบเป็นลำ ดับชั้นและการสืบค้นข้อมูลแบบ static

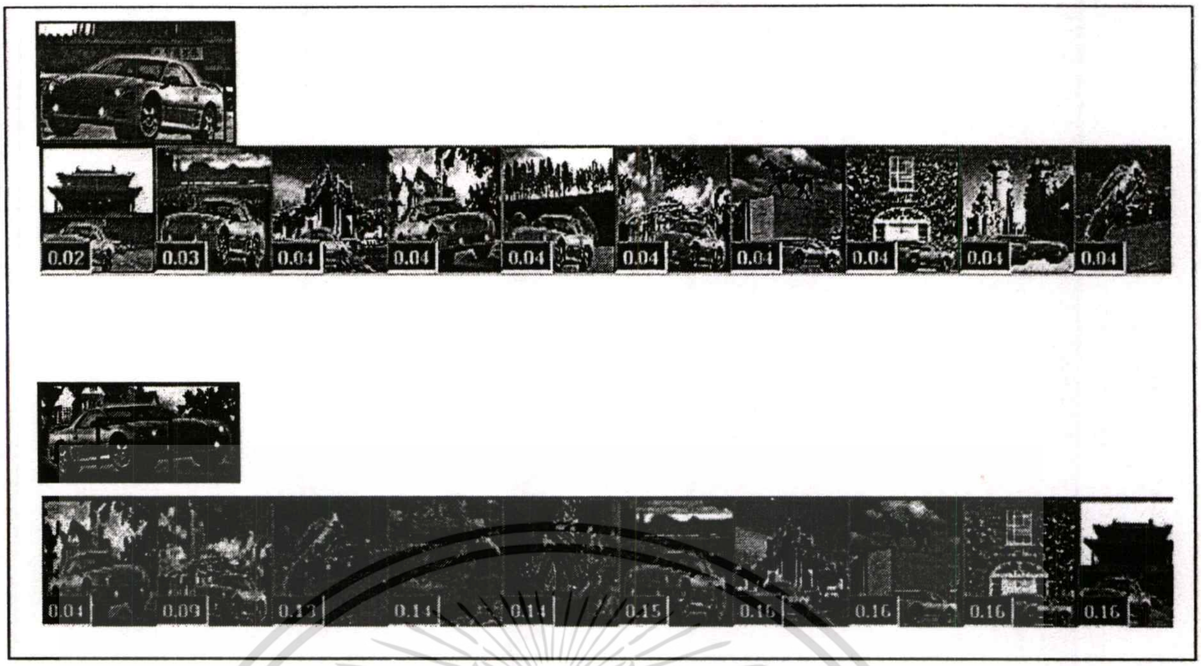
6.2 สรุปและวิจารณ์

วิทยานิพนธ์นี้ได้พัฒนาระบบการค้นคืนรูปภาพแบบผสมเป็นลำดับชั้นและแบบ static ของพื้นที่ที่สนใจและการทำดัชนีรูปภาพด้วยฮิสโตแกรมร่วมกับออร์โตโครรีโลแกรมเพื่อให้ระบบสามารถค้นหารูปภาพในกรณีรูปภาพที่มีสิ่งแวดล้อมหรือสถานที่ที่แตกต่างกันและรูปภาพที่มีการเปลี่ยนแปลงไปได้ทั้งที่วัตถุที่สนใจมีความแตกต่างในแง่ของขนาด มุมมองของภาพและการปรากฏของวัตถุเพียงบางส่วนด้วย แล้วทำการทดลองระบบที่ได้รับการพัฒนาขึ้นพร้อมทั้งเปรียบเทียบการทำงานกับการค้นหารูปภาพโดยใช้ดัชนีแบบ global Histogram และ region Histogram อย่างเดียว ในการค้นหารูปภาพในกรณีต่างๆ ดังต่อไปนี้ คือ รูปภาพที่มีสถานที่ที่แตกต่างกัน รูปภาพที่แตกต่างในแง่ของขนาด มุมมองของภาพและการปรากฏของวัตถุเพียงบางส่วน รูปภาพที่เหมือนและคล้ายกัน จากผลการทดลองปรากฏว่าระบบที่ได้รับการพัฒนานั้นสามารถค้นหารูปภาพที่มีความเหมือนใกล้เคียงกันในฐานมูลได้เกือบทั้งหมด แต่ยัง คงมีข้อจำกัดของระบบที่พัฒนาขึ้นมานี้ในบางกรณี ดังต่อไปนี้

6.2.1 การตีกรอบเลือกพื้นที่ที่สนใจ

การตีกรอบเลือกพื้นที่ที่สนใจในรูปภาพตัวอย่าง ในการทดลองการตีกรอบเลือกพื้นที่ที่สนใจในรูปภาพตัวอย่างที่ไม่เหมาะสมนั้นจะทำให้รูปภาพที่ค้นหาจะได้ภาพที่ไม่ตรงกับความต้องการ ดังนั้นการตีกรอบเลือกพื้นที่ที่สนใจในรูปภาพตัวอย่างที่เหมาะสมจะทำให้การค้นหารูปภาพมีความสามารถเพิ่มขึ้น ดังแสดงในรูปที่ 6.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.1 ตัวอย่างในการค้นหารูปภาพที่มีการตีกรอบเลือกพื้นที่ที่สนใจในรูปภาพตัวอย่างที่แตกต่างกัน

6.2.2 ขนาดของฐานข้อมูล

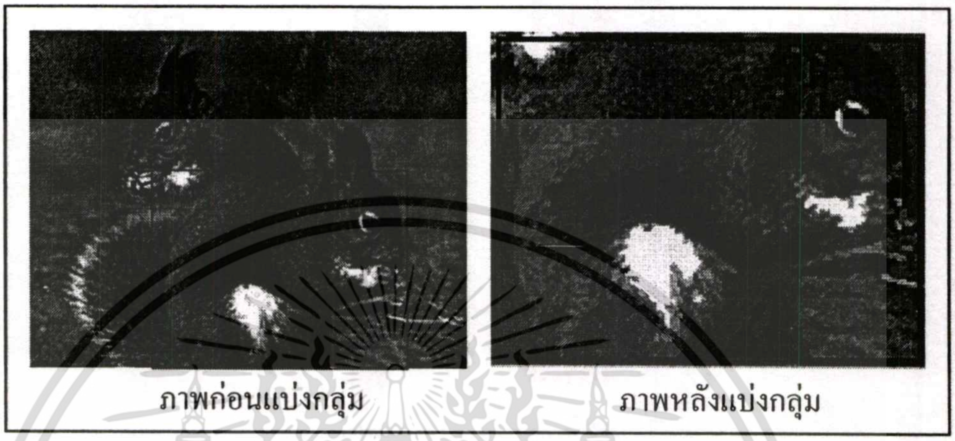
หากฐานข้อมูลมีขนาดใหญ่ขึ้นมาก ในการทดลองนั้นจะพบว่าการค้นหารูปภาพบางรูปภาพจากฐานข้อมูลนั้นให้ผลลัพธ์ของรูปภาพที่มีกลุ่มสีคล้ายคลึงกันซึ่งเป็นภาพที่ต้องการหรือคาดหมายว่าจะได้ แต่รูปภาพที่ได้ออกมานั้นจะมีโอกาสที่รูปภาพส่วนหนึ่งที่ไม่ตรงกับประเภทของความต้องการออกมาด้วย เช่น ต้องการรูปรถสีเหลือง แต่จะได้ภาพปลาที่มีสีเหลืองออกมาด้วย ดังแสดงในรูปที่ 6.2

รูปที่ 6.2 ตัวอย่างผลการค้นหารูปภาพที่มีกลุ่มสีใกล้เคียงกัน

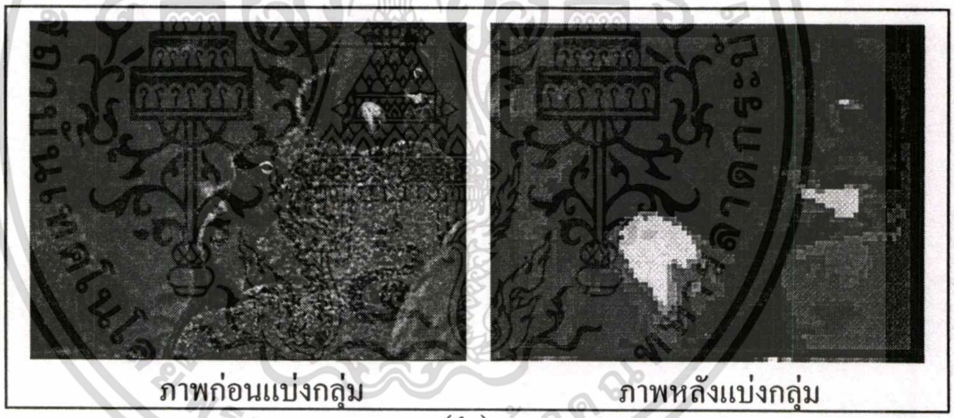
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.2.3 ขนาดของภาพ

รูปภาพที่มีการเปลี่ยนแปลงมาก เช่น การย่อขยาย ความเหมือน เป็นต้น จะทำให้การแบ่งกลุ่มของรูปภาพมีผลแตกต่างกันมาก ดังแสดงในรูปที่ 6.3 จะเห็นว่าภาพที่มีขนาดใหญ่จะแบ่งกลุ่มได้ละเอียดกลว่าภาพที่มีขนาดเล็ก ซึ่งทำให้มีปัญหาในการเปรียบเทียบความแตกต่างระหว่างกลุ่มที่ผู้ใช้เลือกและกลุ่มที่อยู่ในฐานข้อมูล



(a)



(b)

รูปที่ 6.3 ตัวอย่างผลการแบ่งกลุ่มที่มีความแตกต่างของขนาดรูปภาพ (a) รูปภาพขนาดใหญ่ (b) รูปภาพขนาดเล็ก

6.3 แนวทางในการพัฒนาต่อ

แม้ว่าการทำดัชนีรูปภาพโดยใช้ลักษณะสำคัญสีจะให้ผลดี แต่บางกรณี ลักษณะสำคัญสีอย่างเดียวไม่สามารถแยกความแตกต่างของกลุ่มที่มีสีคล้ายคลึงกันได้ แนวทางในการพัฒนาต่อเพื่อเพิ่มประสิทธิ ภาพในการค้นคืนรูปภาพ คือการใช้ลักษณะสำคัญพื้นผิว (Texture) หรือ รูปร่าง (shape) ร่วมกับลักษณะสำคัญสี นอกจากนั้น ผลจากขบวนการแบ่งกลุ่มรูปภาพก็มีผลต่อประสิทธิภาพในการค้นคืนรูปภาพ ถ้าวิธีการแบ่งกลุ่มรูปภาพที่ดีก็ยิ่งจะทำให้ประสิทธิภาพในการค้นคืนรูปภาพดีขึ้นอีกด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] Digital Libraries, Special issue of Communications of the ACM, vol. 38, April 1995.

[2] W. Y. Ma and B. S. Manjunath, "Netra: a toolbox for navigating large image databases", in Proc. IEEE ICIP, vol.1, pp. 568-571, Oct. 1997.

[3] M. Flickner et al., "Query by image and video content: The QBIC system", IEEE Computer, 28(9): 23-32, 1995.

[4] J. Huang et al., "Image Indexing using Color Correlograms", IEEE Comp. Soc. Conf. Vision and Pattern Recognition, pp. 762-768, 1997.

[5] O. Huseyin, T. Chen, and H. R. Wu, "Performance Evaluation of Multiple Regions-of-Interest Query for Accessing Image Databases", Proc. of Int. Symposium on Intelligent Multimedia, Video, and Speech Processing, pp. 300-303, May 2001.

[6] Y. Deng et al. "Peer Group Filtering and Perceptual Color Image Quantization", IEEE ISCAS, vol. 4, pp. 21-24, 1999.

[7] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic, Norwell, MA, 1992.

[8] G. Pass and R. Zabih, "Histogram refinement for content-based image retrieval", IEEE Workshop on Applications of Computer Vision, pp. 96-102, 1996.

[9] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*, Addison Wesley, 1999.

[10] K. Tang, J. Astola and Y. Neuvo, "Adaptive nonlinear multivariate image filtering for mixed noise removal", Proc. Of ISCAS, vol. 1, p. 427-30, 1993.

[11] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, 1970.

[12] R. M. Hararick, "Statistical and structural approaches to texture", Proc. Of IEEE 67(5):786-804, 1979

[13] R.C. Gonzalez and R.E. Wools, *Digital Image Processing*, Addison-Wesley, Reading, MA, 1992

[14] Linde, Y., Buzo, A. and Gray, R.M, "An algorithm for vector quantization design", IEEE Tran. Commu. 28, pp. 84-95, 1980.

[15] J.C., Ehrlich, R., and Full, W., "FCM: The fuzz C-means Clustering algorithm", Computer and Geosciences. 10, pp. 191-203, 1984.

[16] C. H., Lu. P.C. and Chung. J. C, "fast codebook generation algorithm for vector quantization of image", Pattern Regogn. Lett. 12, pp 605-609, 1991.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก ผลการทดลอง

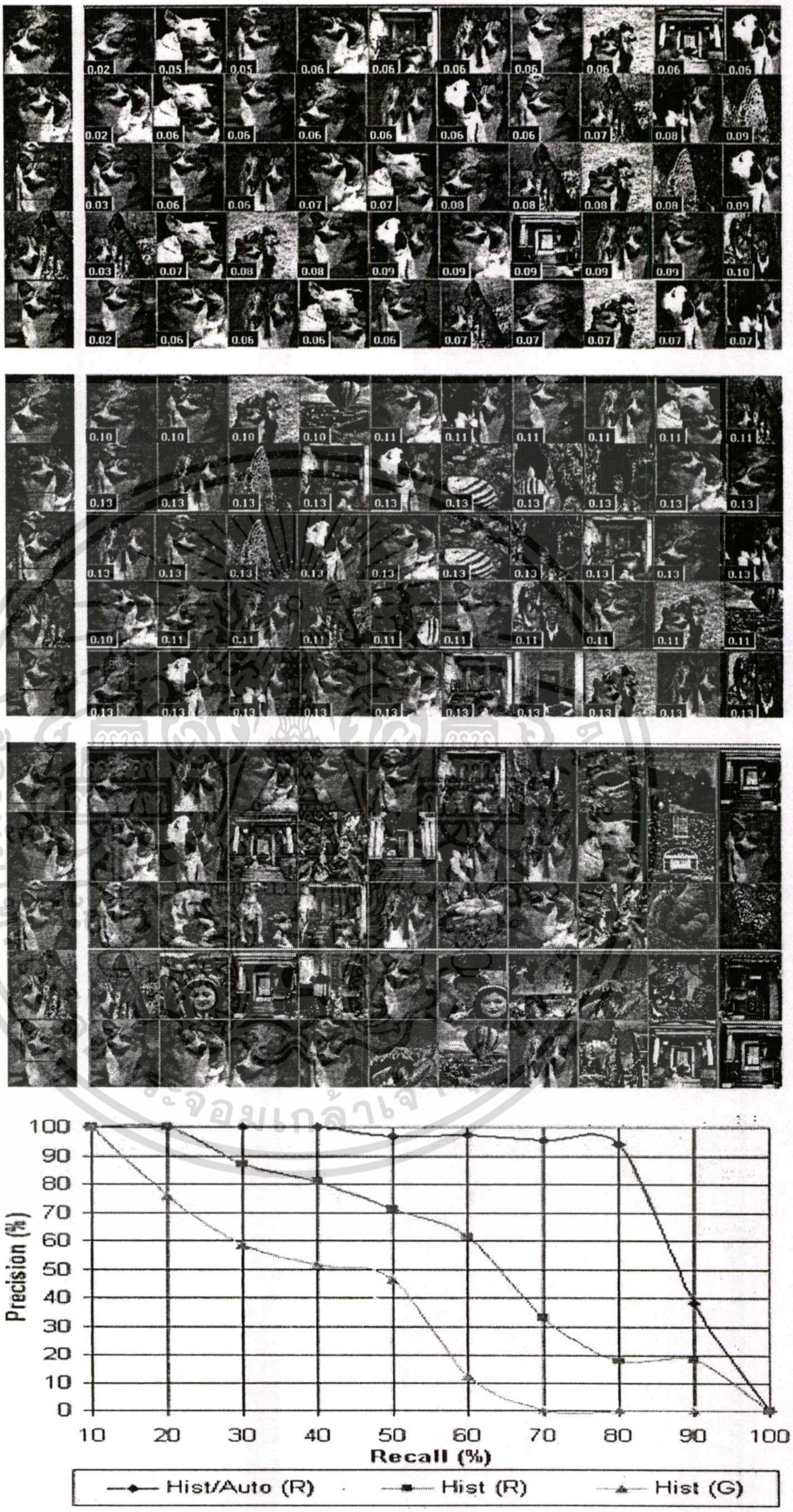
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(a)

(b)

(c)

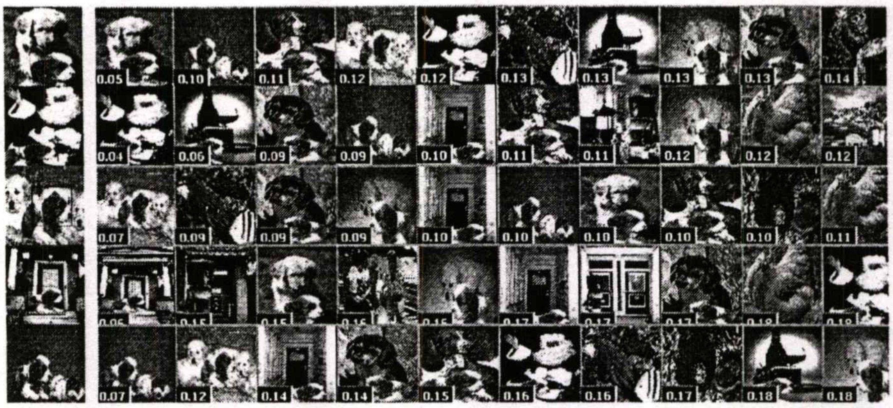
(d)



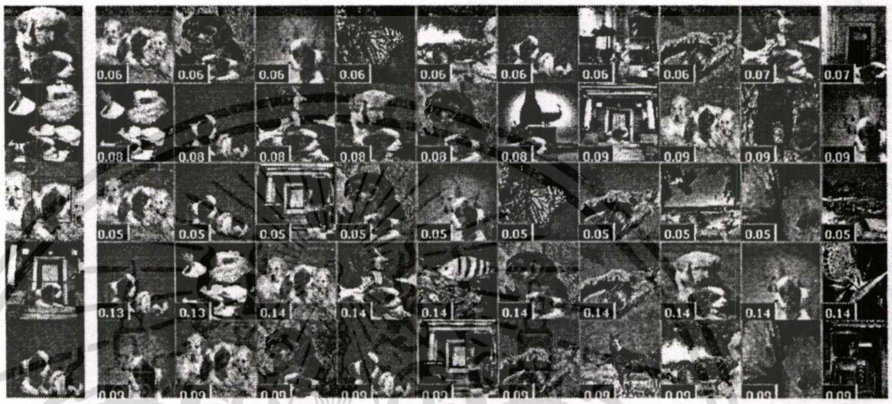
รูปที่ ผก.1 ผลการค้นคืนรูปภาพชุดที่ 8 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R) (c) , Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

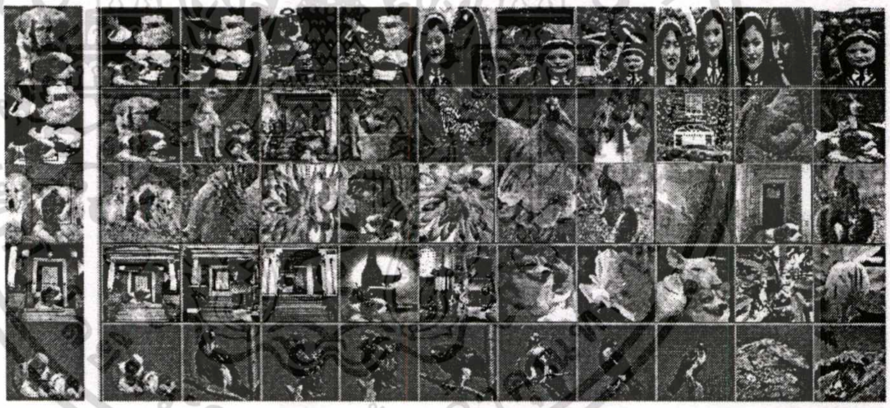
(a)



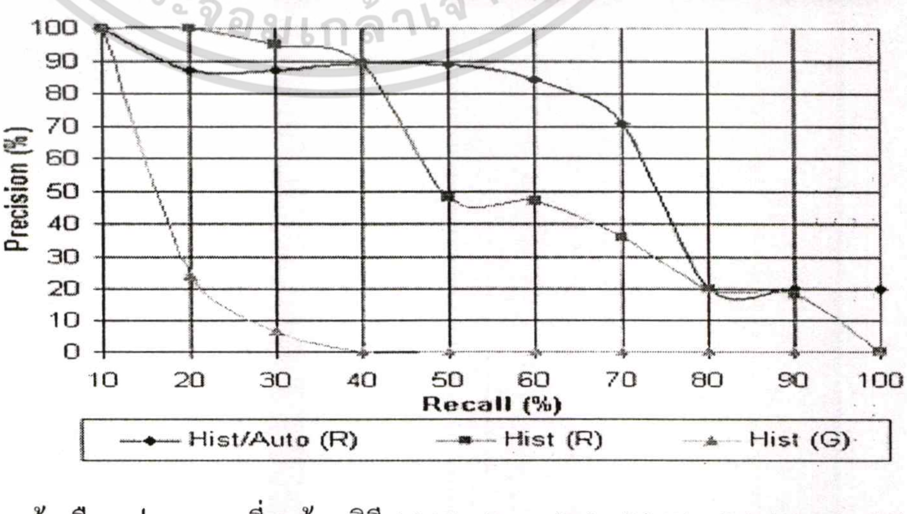
(b)



(c)



(d)



รูปที่ ผก.2 ผลการค้นคืนรูปภาพชุดที่ 9 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

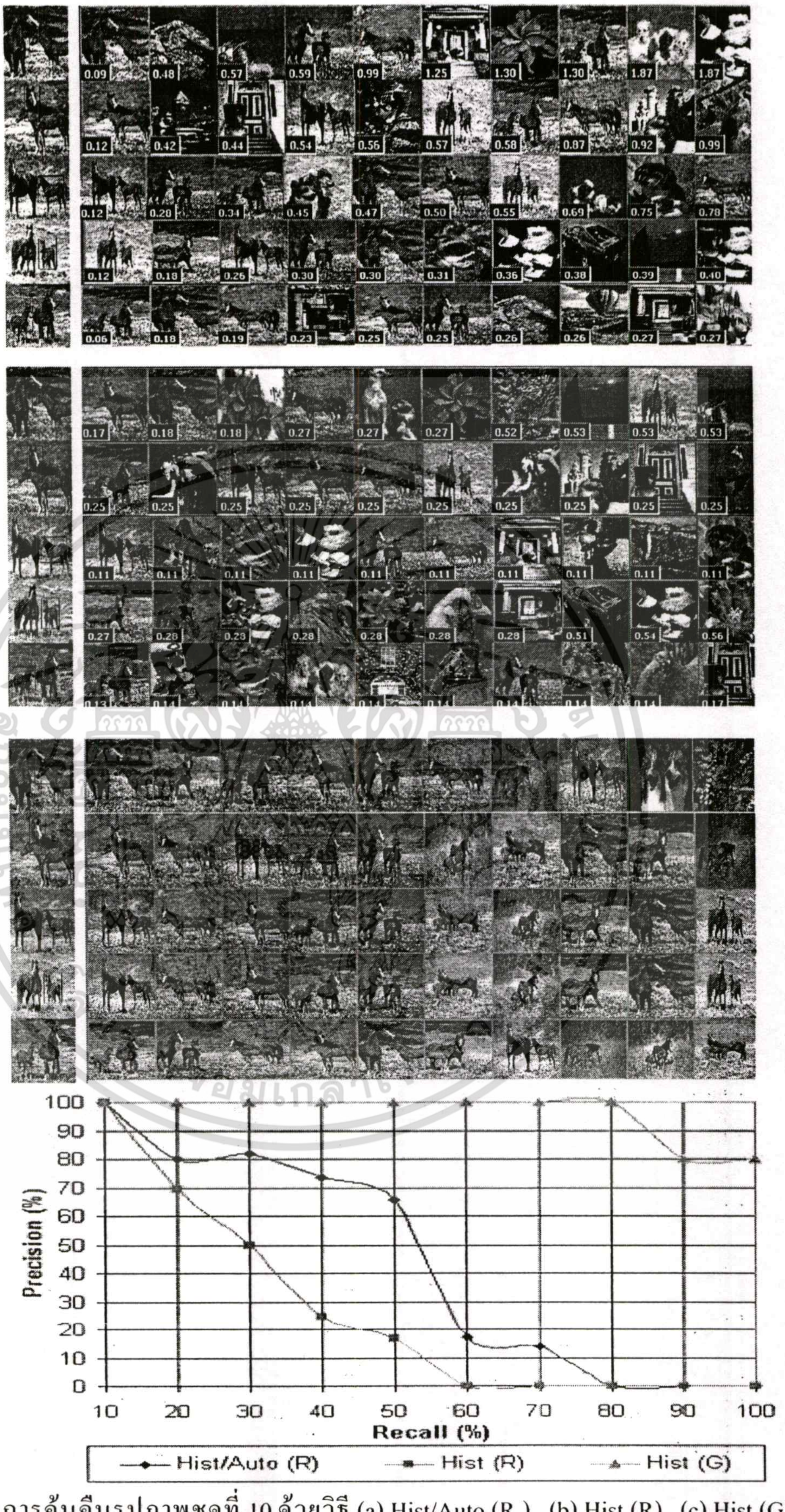
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(a)

(b)

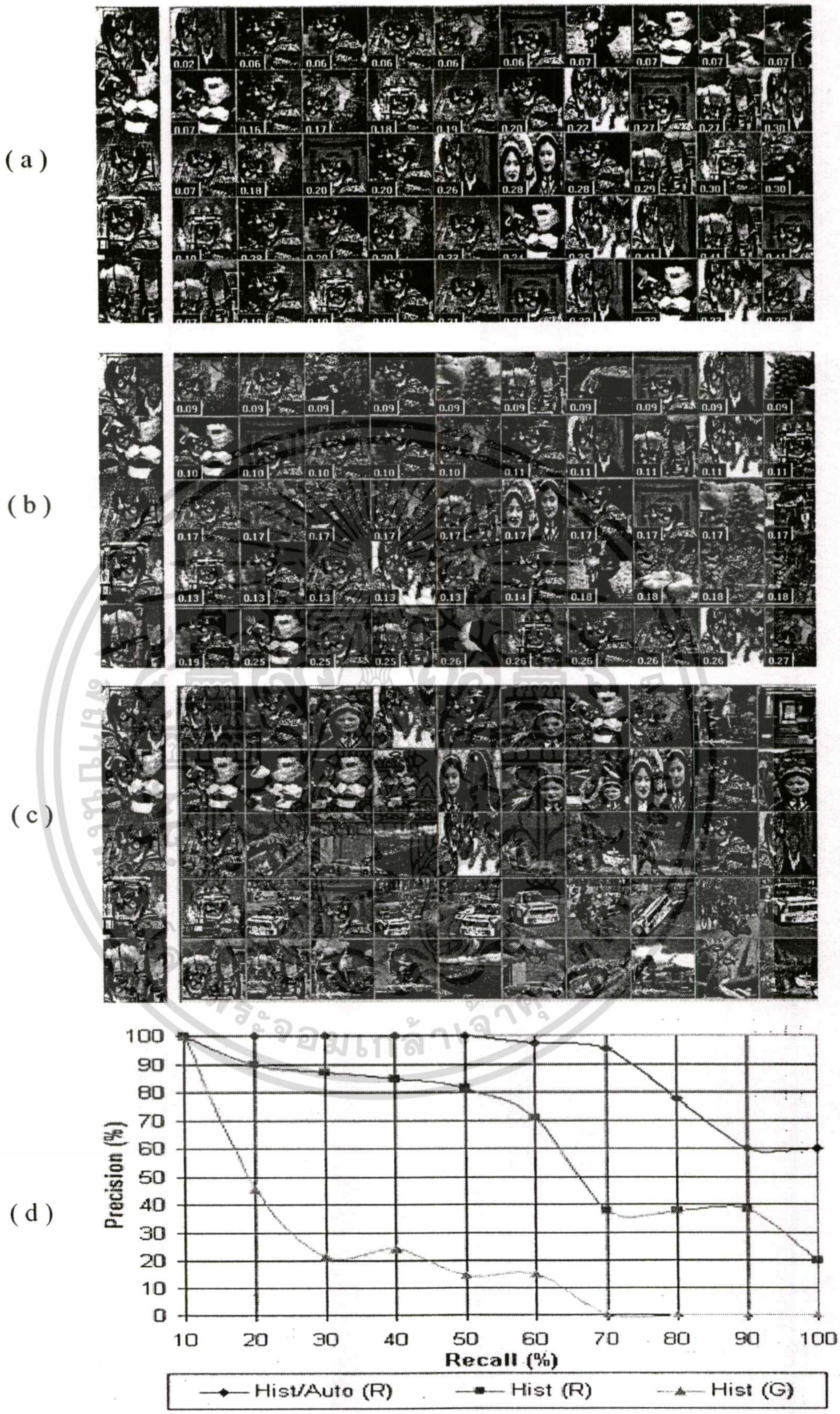
(c)

(d)



รูปที่ ผก.3 ผลการค้นคืนรูปภาพชุดที่ 10 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

เอกสารนี้เป็นเอกสารที่เผยแพร่เพื่อการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

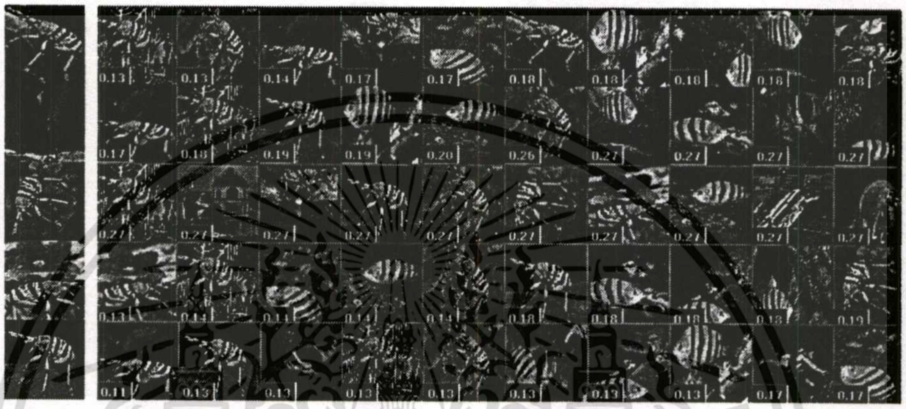


รูปที่ ผก.4 ผลการค้นคืนรูปภาพชุดที่ 11 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

เอกสารนี้เป็นเอกสาร ... ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

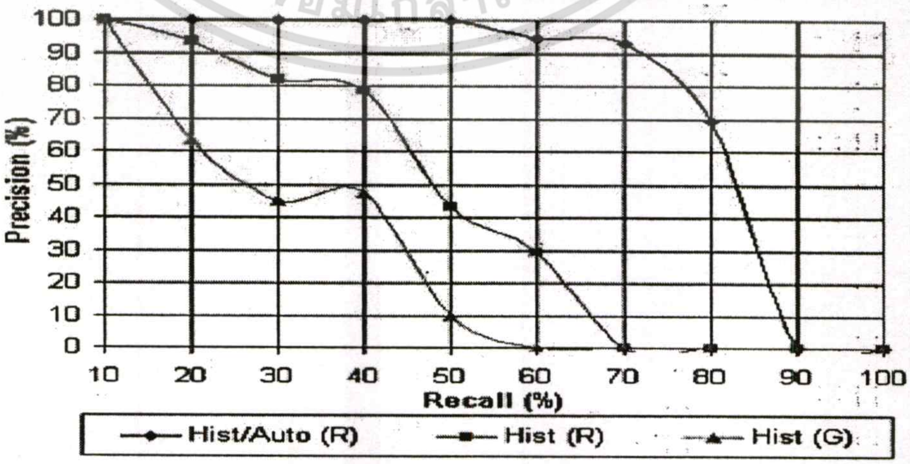
(a)

(b)



(c)

(d)



รูปที่ ผก.5 ผลการค้นคืนรูปภาพชุดที่ 12 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

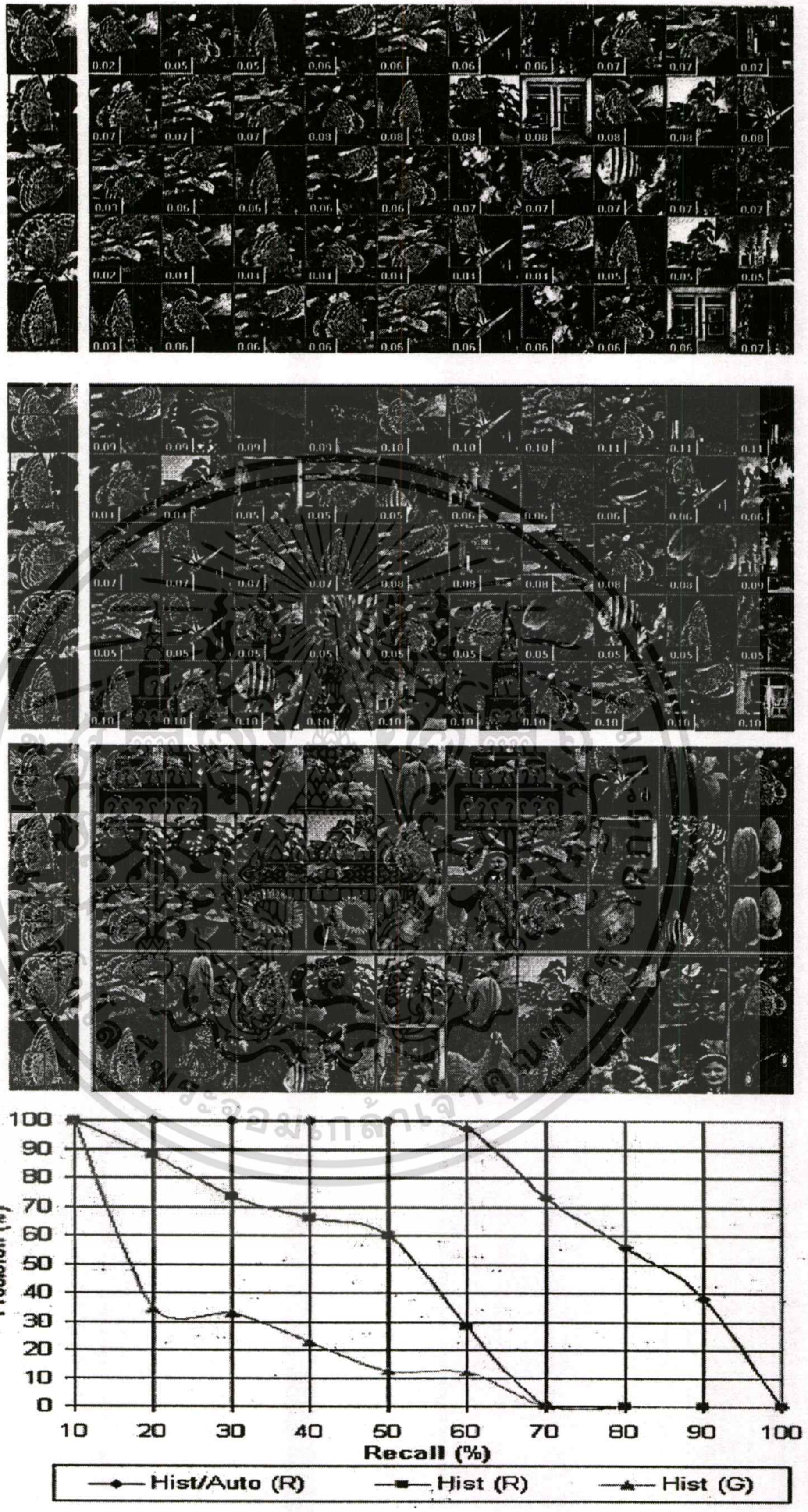
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(a)

(b)

(c)

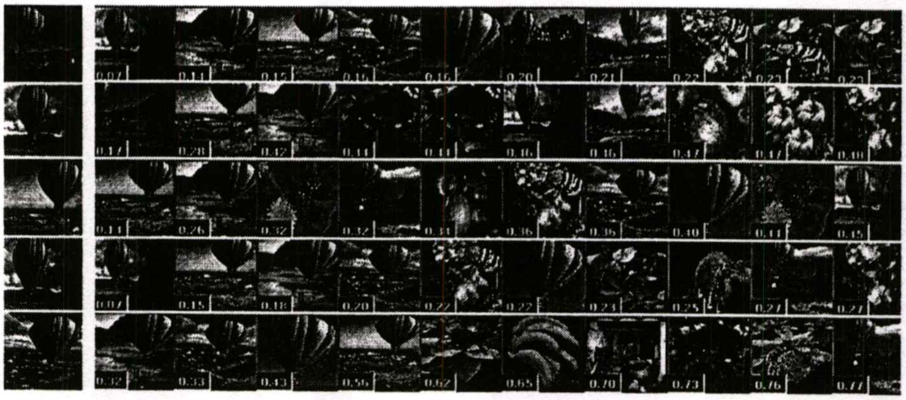
(d)



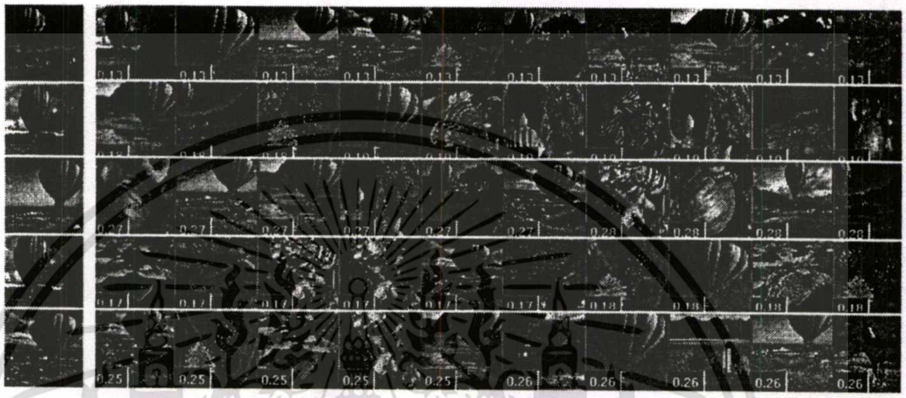
รูปที่ ผก.6 ผลการค้นคืนรูปภาพชุดที่ 13 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

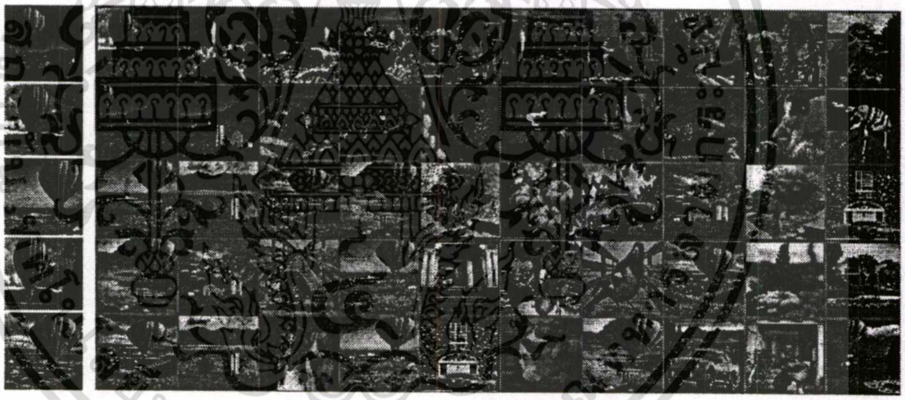
(a)



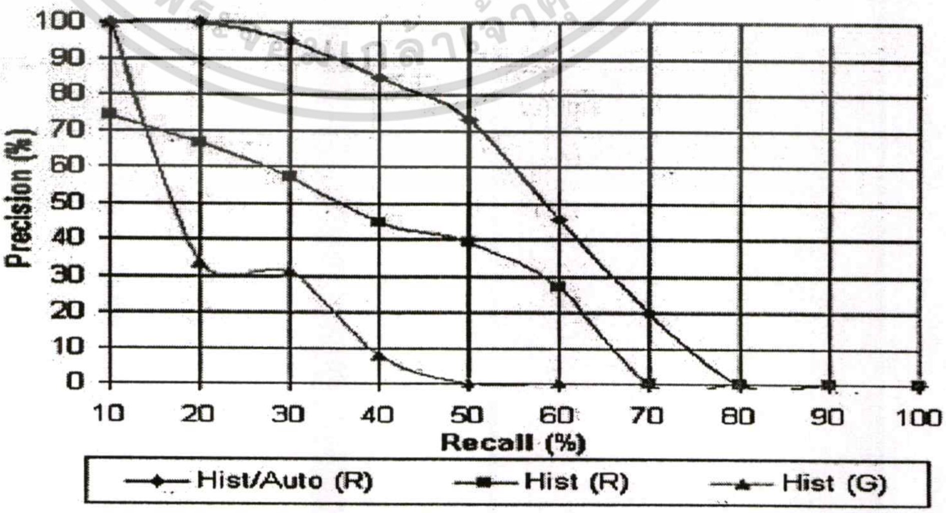
(b)



(c)



(d)



รูปที่ ผก.7 ผลการค้นคืนรูปภาพชุดที่ 14 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

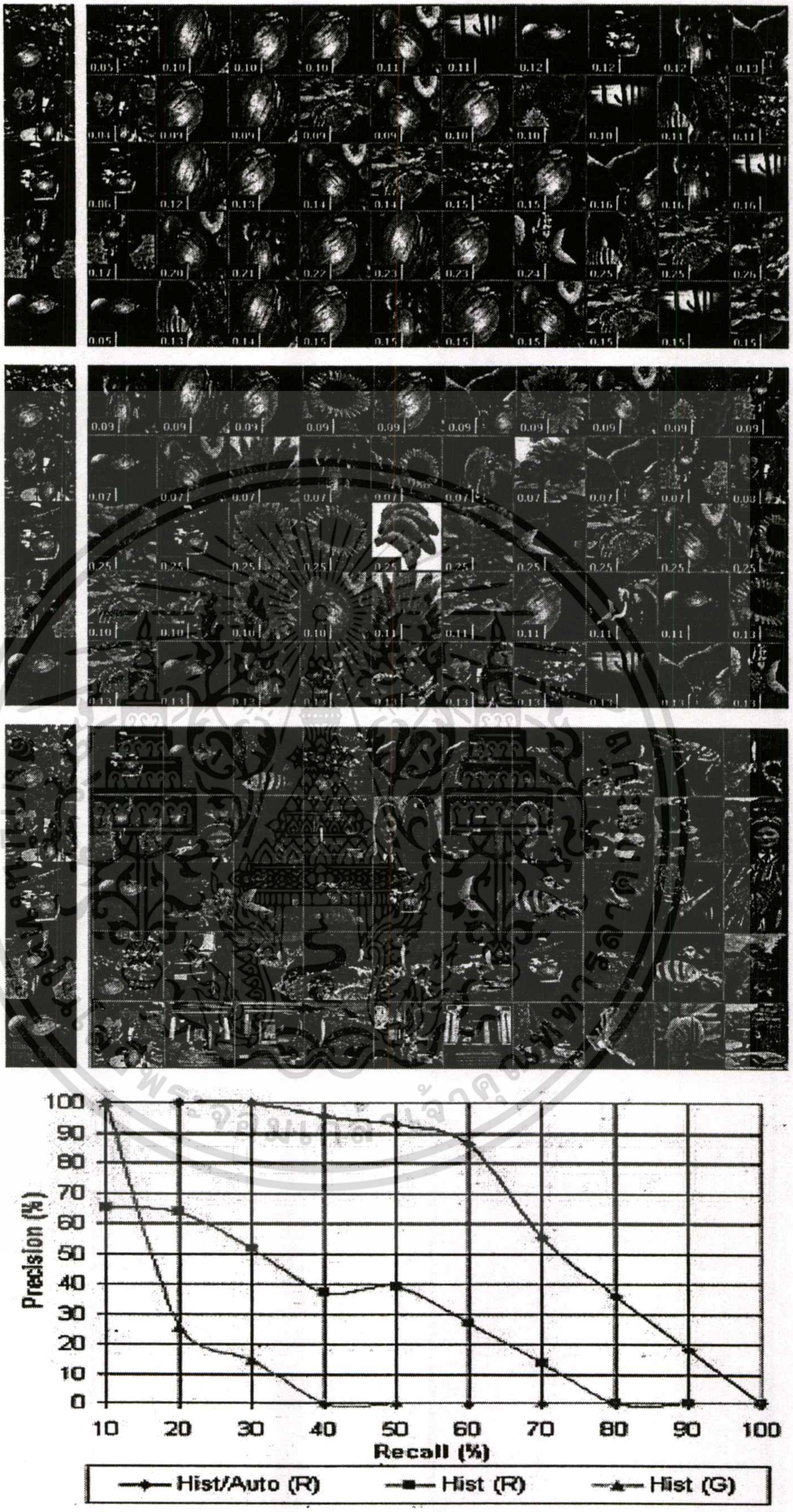
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(a)

(b)

(c)

(d)



รูปที่ ผก.8 ผลการค้นคืนรูปภาพชุดที่ 15 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

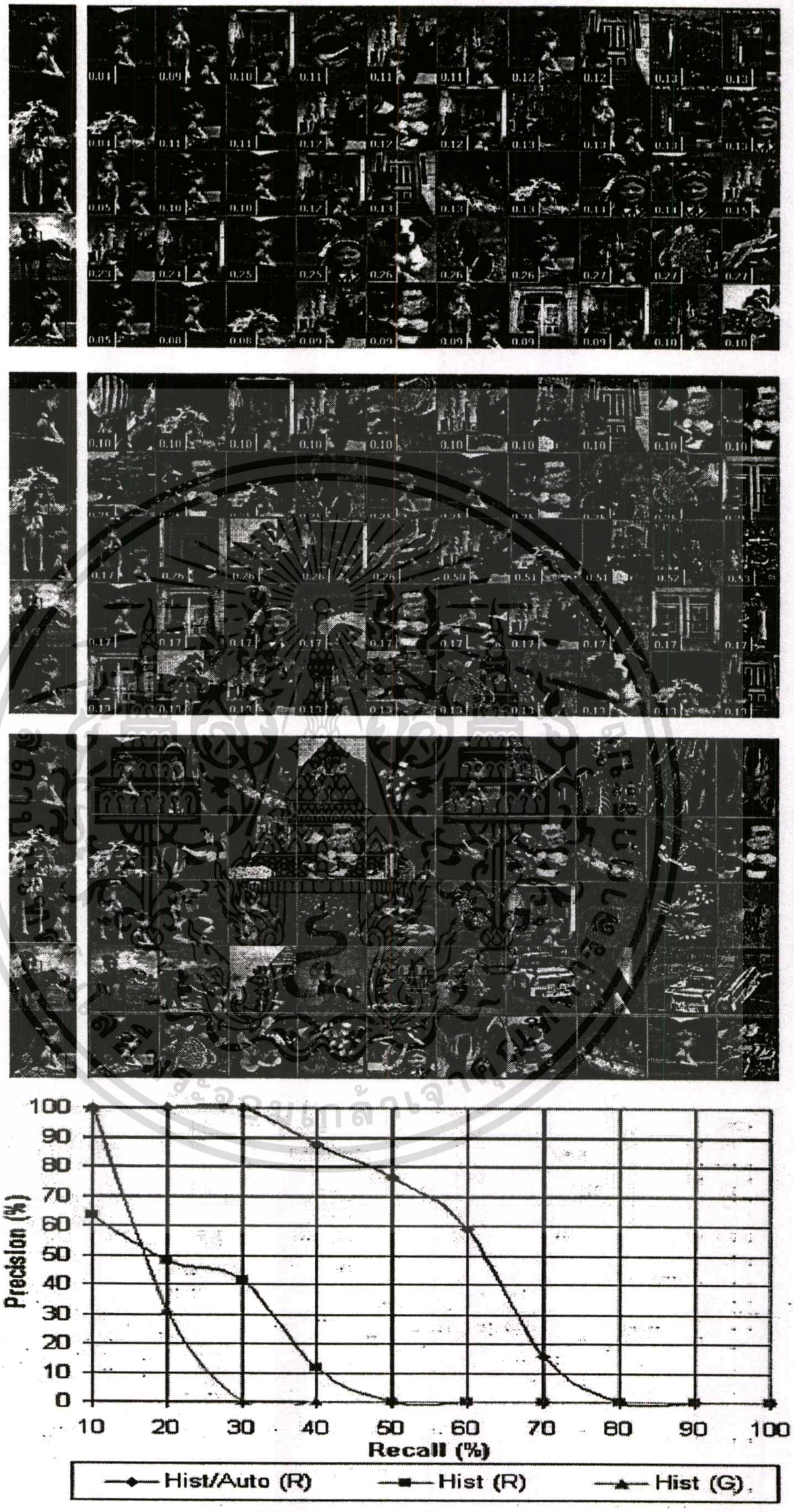
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(a)

(b)

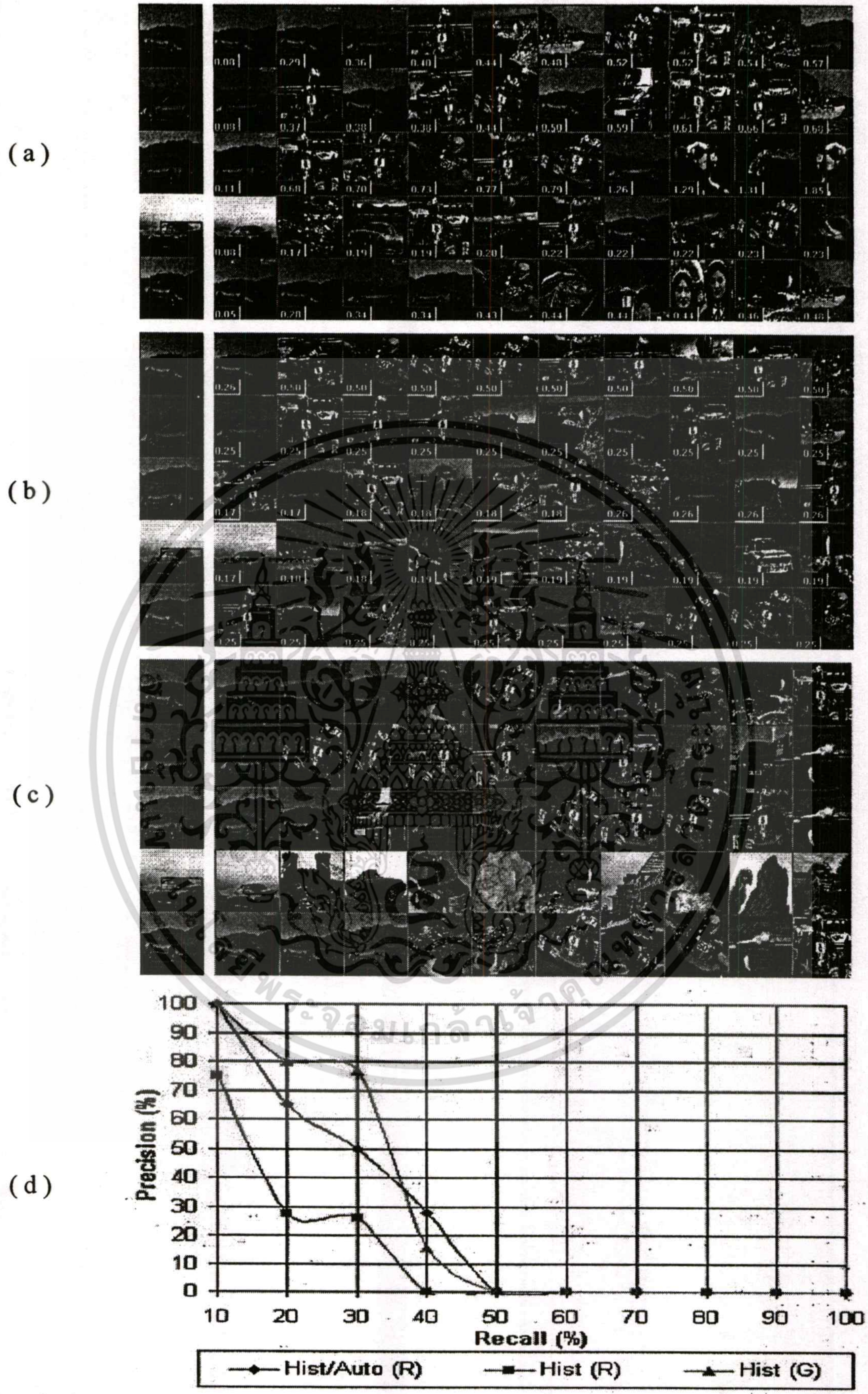
(c)

(d)



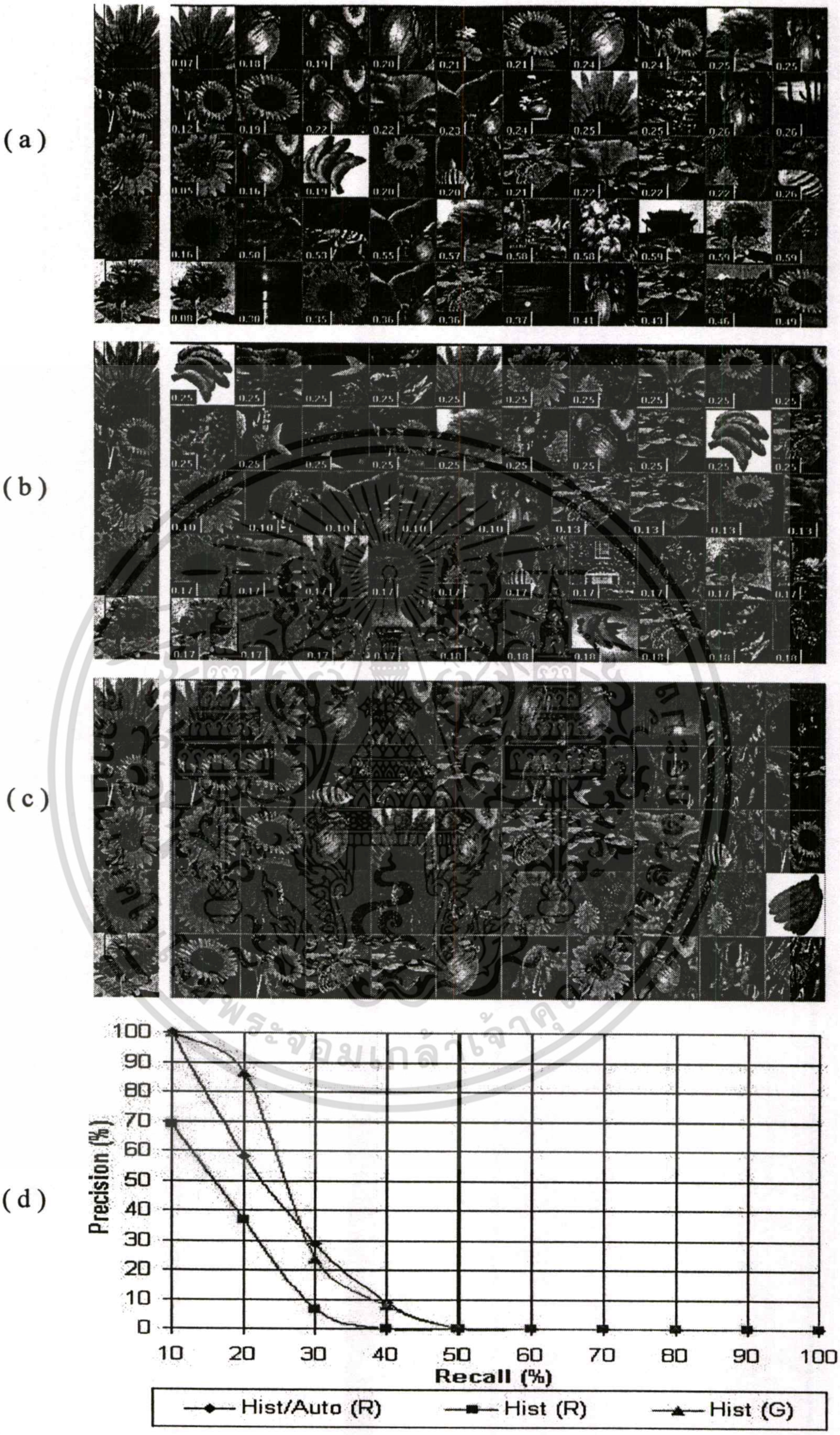
รูปที่ ผก.9 ผลการค้นคืนรูปภาพชุดที่ 16 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ผก.10 ผลการค้นคืนรูปภาพชุดที่ 17 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ผก.11 ผลการค้นคืนรูปภาพชุดที่ 18 ด้วยวิธี (a) Hist/Auto (R), (b) Hist (R), (c) Hist (G) และ (d) กราฟเปรียบเทียบประสิทธิภาพในการค้นคืนรูปภาพโดยใช้เทคนิคทั้ง 3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข บทความวิจัยที่ได้รับการตีพิมพ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5-я Международная конференция
и выставка

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ЕЕ ПРИМЕНЕНИЕ

12 – 14 марта 2003 г.,
Москва, Россия

ДОКЛАДЫ - 1

PROCEEDINGS - 1

МОСКВА 2003

IEEE-Russia

AUTEX Ltd

МЦНТИ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

HIERACHICAL REGION OF INTEREST QUERY AND INDEXING FOR CONTENT-BASED IMAGE RETRIEVAL

Bounthanh T., Chitsobhuk O., Mitatha S.

Department of Computer Engineering,
King Mongkut's Institute of Technology, Ladkrabang, 10520 Thailand

Abstract. In this paper, we present a Content-based Image Retrieval (CBIR) system using hierarchical Region of Interest (ROI) query and indexing. During indexing process, all the images in the database are segmented using Vector Quantization (VQ) technique. Color features of color histogram and color correlogram analysis are then extracted from each segmented region and stored in the database as the key to retrieve the relevant images. In the proposed retrieval system, users are allowed to select ROI directly over the sample or user's submission image and the query process then focuses on the content of the selected ROI in order to find those images containing similar regions from the database. The hierarchical region-of-interest query is performed to retrieve the similar images. Two-level search is exploited in this paper: In the first level, the most important regions, usually the large regions at the center of user's query, are used to retrieve images having similar regions using static search. This ensures that we can retrieve all the images having the most important regions. In the second level, all the remaining regions in user's query are used to search from all the retrieved images obtained from the first level. By using the proposed hierarchical ROI query, the searching time is reduced and the experimental results show good retrieval performance even in the case where the query ROI is in different scales, points of view, and background or partially appears in the indexing images.

1. Introduction

More and more digital images are captured and stored in the databases. One of the significant issues is how this information can be organized such that fast and efficient browsing, indexing and retrieval of the image content can be achieved [1]. A number of indexing and retrieval systems have been proposed to query the images in the database [1-3]. One approach is the content-based indexing and retrieval, which is widely adopted since all the meaningful attributes can be derived directly from the content containing in that image itself. The generated attributes are usually the visual features such as color, texture, or shape. These features are used as the key to retrieve the relevant images from the database according to user's query. Color histograms [3] are generally used as color features to describe the global color distribution in an image. Even though the histogram can be computed easily and is invariant to small changes in viewing position, it does not give any spatial or structure of colors in the image. This may cause false positives especially in large databases since different images may have similar global color distribution. To solve this problem, color correlogram [4, 8] was proposed in order to incorporate spatial correlation of colors with the global distribution of local spatial correlation of colors. This feature is easy to compute while the size of the feature is relatively small.

Retrieval systems are also important. Several retrieval systems [3,4] perform searches for relevant images from the database using features extracted from all regions of the user's query image. This type of query is good when user wants to find similar image to the example, which unlikely happens. Users usually try to search for such a specific object, such as "find a golden car", and sometimes do not pay attention on background or other objects in the sample image. Therefore, the retrieval system should allow human assistance to specify region-of-interest and then identify images that best describe the user's ROI query using the pre-calculated features.

2. The proposed indexing algorithm

In the proposed algorithm, each image is segmented into N homogeneous regions using vector quantization of color information. Before segmentation process, peer group filtering and perceptual color quantization [6] is applied to provide smoothness and noise reduction in the image. This technique was shown to be effectively removed noise without blurring the edges and details in the image. Then, the smoothed image is segmented using a modified general Lloyd (GLA) vector quantization technique. VQ algorithm requires a set of initial classification. One approach uses random initialization to generate the initial classification set [7]. However, using random initialization, the segmentation results may be different in the case that the same object or ROI is in different background images or the images contain the chosen ROI but in different sizes or points of view due to the randomness of pixel values. This leads to the difficulty in matching user's query ROI with the similar regions of the images in the database. Therefore, in this paper, color quantization is used to first generate the initial color sets. By using the proposed initialization technique, the segmentation process results in the similar segmented regions even in the problem cases discussed above. After initialization process, a modified GLA is performed to further nonlinearly classify the large initialized color sets into smaller groups using overall distortion of each group. The number of clusters is set to be slightly larger than the actual need. The correctness of the estimation is not critical because cluster will be merged in the next step. After GLA, a large number of pixels having approximately the same color will have more than one cluster because GLA is aimed to the minimize the global distortion. An agglomerative clustering algorithm [7] is performed on the cluster centroids to further merge close clusters such that the minimum distance between two centroids satisfied a preset threshold. After classification, the quantized color pixels are assigned region labels. Then, color features are extracted from each region such as histogram and color correlogram

เอกสารนี้ใช้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[4,8]. Color correlogram is a spatial extension of the histogram. Given any pixel of color C_i in the image color correlogram gives the probability that a pixel at distance k away from the given pixel is of color C_j. In this paper, the autocorrelogram is adopted in addition to color histogram. The autocorrelogram provides spatial correlation between identical colors. The vectors of histogram and autocorrelogram features are used as indexes for each image in the database.

3. The Retrieval System

In the proposed retrieval system, users are allowed to impose their desired requirements on the query image using a rectangular box. The ROI selection area is then segmented into N regions using the same technique as in the indexing process. For each region in the ROI, color features, color histogram and corresponding autocorrelogram, are extracted as described in previous section. Matching between user's selected ROI and that of each target image from the database is performed using similarity comparisons between the extracted features. The similarity measure used in the proposed system is a modified L_1 distance (relative measure of distance [9]) of histogram and autocorrelogram as:

$$|I - I'|_{h,d_1} \triangleq \sum_{i\in[m]} \frac{|h_{c_i}(I) - h_{c_i}(I')|}{1 + h_{c_i}(I) + h_{c_i}(I')}$$

$$|I - I'|_{\gamma,d_1} \triangleq \sum_{i\in[m]} \frac{|\gamma^{(k)}_{c_i,c_j}(I) - \gamma^{(k)}_{c_i,c_j}(I')|}{1 + \gamma^{(k)}_{c_i,c_j}(I) + \gamma^{(k)}_{c_i,c_j}(I')}$$

where I and I' are the query and the indexing image. $h_{C_i}(I)$ and $h_{C_i}(I')$ are the color histograms and $\gamma^{(k)}_{C_i,C_j}(I)$ and $\gamma^{(k)}_{C_i,C_j}(I')$ are color correlogram of I and I' images, respectively.

To retrieve relevant images, hierarchical query is performed instead of static query, which searches for N query regions to retrieve N query results and uses all the retrieved results to produce a single query result. Static query requires a large amount of search time due to a large number of regions containing in each indexing image. Even though hierarchical query requires less amount of search time, it is possible that a mistake from one level can propagate to next levels and cause retrieval errors at the end of query. Therefore, in this paper, two-level search is proposed, which not only can reduce the search time compared to the static search algorithm but also can reduce the effects of the mistakes propagating from one level to the next. The proposed hierarchical search is a combination of the static search and the traditional hierarchical search. In the first-level query, the most important regions from user's query, usually the large regions at the center of user's query, are matched with the regions in the indexing images using the static search technique. The indexing images, which contain all the most important regions and have matching errors less than predefined thresholds, are retrieved from the database. The refinement search is performed in the second-level query to look for the remaining regions in user's query from first-level retrieved images.

4. EXPERIMENTAL RESULTS

We have experimented the proposed hierarchical query and indexing approach using several color features such as global color histogram, regional color histogram, and a combination of regional color histogram and its corresponding autocorrelogram. These features are extracted from regions in the indexing images and the selected ROI of user's query. The number of histogram bins used is 64 bins while the distance set k = {1, 3, 5, 7} is used for computing autocorrelogram feature. The database consists of 100 images, including animals, cars, peoples, flowers, pumpkins, etc. Each set of images in the database contains an object with different scales, points of view, and backgrounds. Our query set consists of 50 queries, each chosen to represent various situations such as different views of the ROI, spatial translations, and backgrounds. The retrieval performance of all the queries is shown in Fig. 1.

The retrieval performance is evaluated using average precision at each recall level [x] as follow:

$$\overline{P}(r) = \sum_{i=1}^{N_q} \frac{P_i(r)}{N_q}$$

where P(r) is the average the precision at the recall level r, Nq is the number of queries used, and Pi(r) is the precision at recall r for the i-th query. From figure 1, the experimental results show that the retrieval performance of the proposed features outperforms the other two techniques.

In addition, we also show examples of the retrieval results of the butterfly query using each type of color features are presented in Fig. 1.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

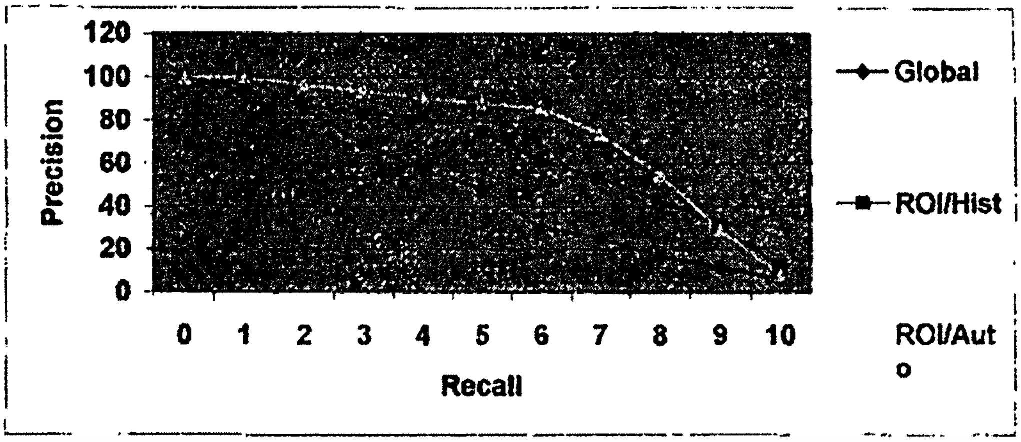


Fig. 1. The overall retrieval performance

(a)
(b)
(c)

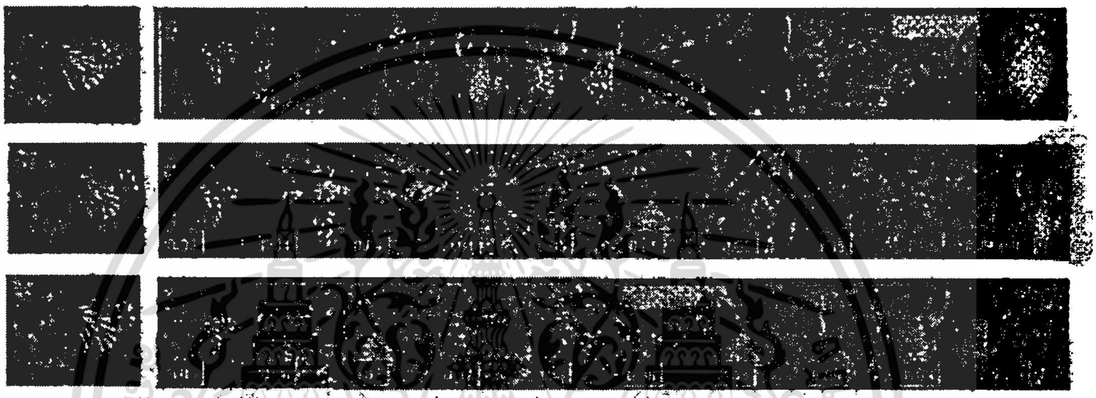


Figure 1. The retrieval results from different color features, (a) global histogram, (b) regional histogram, and (c) combined regional histogram and autocorrelogram.

As we can see from Fig. 1, the retrieval results of using combined regional color histogram and autocorrelogram are better than those of using global histogram and regional histogram only even in the case where the query ROI is in different scales, points of view, and background or partially appears in the indexing images. This is due to the color distribution and color structure information supported by the combined feature.

5. Conclusion

In this paper, we have investigated hierarchical query and indexing using regional color histogram and autocorrelogram analysis compared with the global color histogram and regional color histogram only. Since the selected features provide both color distribution and spatial correlation of colors in ROI, it can be used effectively to discriminate between different types of ROI. The retrieval performance is very efficient even in the case of different points of view, partial occlusion, and totally different backgrounds. Moreover, by using the proposed hierarchical search, we can reduce the searching time while still maintain good retrieval performance. Even though color indexing seems to give reasonable retrieval results, in some cases, only color features may not be enough to distinguish different types of ROI having similar colors. In order to further improve the retrieval performance, other features such as texture or shape should be incorporated with color information. Moreover, the results from the segmentation process are also effect the retrieval performance. The improvement of segmentation algorithm can gives better and thus more meaningful retrieval results

REFERENCES

[1] Digital Libraries, Special issue of Communications of the ACM, vol. 38, April 1995.

[2] W. Y. Ma and B. S. Manjunath, "Netra: a toolbox for navigating large image databases," in Proc. IEEE ICIP, vol.1, pp. 568-571, Oct. 1997.

[3] M. Flickner et al. "Query by image and video content: The QBIC system," IEEE Computer, 28(9): 23-32, 1995.

[4] J. Huang et al. "Image Indexing using Color Correlograms," IEEE Comp. Soc. Conf. Vision and Pattern Recognition, pp. 762-768, 1997.

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[5] O. Huseyin, T. Chen, and H. R. Wu, "Performance Evaluation of Multiple Regions-of-Interest Query for Accessing Image Databases," Proc. of Int. Symposium on Intelligent Multimedia, Video, and Speech Processing, pp 300-303, May 2001.

[6] Y. Deng et al. "Peer Group Filtering and Perceptual Color Image Quantization," IEEE ISCAS, vol. 4, pp. 21-24, 1999.

[7] A. Gersho and R. M. Gray, Vector Quantization and Signal Compression, Kluwer Academic, Norwell, MA, 1992.

[8] G. Pass and R. Zabih, "Histogram refinement for content-based image retrieval," IEEE Workshop on Applications of Computer Vision, pp. 96-102, 1996.

[9] R. Baeza-Yates and B. Ribeiro-Neto, Modern Information Retrieval, Addison Wesley, 1999.

DELIMITER SEARCH FOR LOCATING RULE PATTERN IN INTRUSION DETECTION SYSTEMS

Kruatrachue B., Threepak T., Hongsiwan T.

Department of Computer Engineering, Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang, Thailand

ABSTRACT. The goal of this research is to improve the run time performance of Network Intrusion Detection Systems (NIDS). The system performs intrusion detection by matching specific intrusion patterns with the incoming network packet data. Since the numbers of patterns is large (up to 2000) and keep increasing leading to huge search time. In this paper, we proposed a search time reduction by using delimiter search and implementing intrusion patterns into a dictionary of patterns with burst tries structure. The search time is compared between the snort Boyer-Moore algorithm and proposed method. The result of the proposed method on average take only 4.96 percent of time of original snort.

1. INTRODUCTION

Network Intrusion Detection System (NIDS) is the system which detect the intruder patterns from packet data capturing from network. The early NIDS is from the UC Davis Network Security Monitor [1, 2] using exact string matching. More recently, several NIDS tools have been developed and use broadly such as SNORT [3] and BRO [4]. Both systems still use exact string matching technique. The difficulties in implementing NIDS systems is to perform intrusion pattern matching in near-real-time in order to catch up with incoming network traffic. From these difficulties, the attacker can elude the NIDS by using insertion, evasion and Denial of Service techniques [5].

To overcome DOS, we try to improve intrusion pattern matching speed by finding delimiters in rules and in incoming data packet and use the delimiters and distance between delimiters for intruder patterns matching. Since the size of rules and packet is decrease, the search time is also decrease.

In order to verify the speed improvement of the proposed method we modified snort intruder patterns structure and matching methods and make comparisons on various type of network traffic.

2. SNORT INTRUDER PATTERN STRUCTURE AND SEARCHING ALGORITHM

Snort is a network intrusion detection system. Its function is to capture data from network and checking captured data with intrusion patterns. All the intrusion patterns are groups into a two dimensional list of rules as show in Figure 1. If content of packet data match any rule, it will alert to system administrator. This help system administrator know the intruder as fast as possible and secure the system.

The two dimensional list consist of two node type, Rule Tree Node (RTN) and Option Tree Node (OTN) is shown in Figure 1. The RTN contain the common properties of the rule, such as the source address, destination address, source port, destination port and protocol type (TCP, ICMP, UDP) of packet. The OTN contain the information for various options that can be added to each rule, such as TCP flags, ICMP codes and types, packet content.

If a content check is required, Snort uses a Boyer-Moore pattern matching algorithm to check the content string field in the OTN against the entire packet. If no string match exists, Snort will proceed to the next OTN in the list. This is a very time consuming process even though the Boyer-Moore perform some smart skipping without a shift and check a rule for the whole length of packet. Since the number of rule is large and keep increasing.

3. SNORT RULES

Snort rules are the set of sentence that describes about what snort do if it found some data in data packet matching with the content of rule. The grammar of rule is including snort response, rules of header and rules of content.

```
alert tcp any any -> any 80 (msg:"IIS-cmd?";flags:PA; content:".cmd?&"; nocase;)
```


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

นายทา บุนทัน เกิดเมื่อวันที่ 07 กันยายน 2508 ที่ จำปาสัก ประเทศลาว สำเร็จการศึกษาวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์ จากสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ปีการศึกษา 2543 และ ในปีการศึกษา 2544 ได้เข้าศึกษาต่อในระดับปริญญาโท หลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะบัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้