



ใบรับรองวิทยานิพนธ์

บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ

เรื่อง การตรวจจบและการแก้ไขการวางตัวของภาพ โดยใช้ซอฟต์แวร์แวกเตอร์แมชชีน

โดย นางสาวจิรา แก้วสุวรรณ

ได้รับอนุมัติให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ครุศาสตร์อุตสาหกรรมมหาบัณฑิต สาขาวิชาเทคโนโลยีคอมพิวเตอร์

คณาธิการบัณฑิตวิทยาลัย

(อาจารย์ ดร.มงคล หวังสถิตยวัฒน์)

21 พฤษภาคม 2550

คณะกรรมการสอบวิทยานิพนธ์

ประธานกรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.กัณฑ์พงษ์ วรรณปัญญา)

กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.สุพจน์ นิตย์สุวัฒน์)

กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.วินัย โพธิ์สุวรรณ)

กรรมการ

(ร้อยตำรวจตรี หญิง ดร.นิตาพรรณ สุวรรรัตน์)

การตรวจจับและแก้ไขการวางตัวของภาพโดยใช้ซอฟต์แวร์แมชชีน

นางสาวจิรา แก้วสุวรรณ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ครุศาสตร์อุตสาหกรรมมหาบัณฑิต
สาขาเทคโนโลยีคอมพิวเตอร์ ภาควิชาคอมพิวเตอร์ศึกษา
บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ปีการศึกษา 2549
ลิขสิทธิ์ของสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ชื่อ : นางสาวจิรา แก้วสุวรรณ
ชื่อวิทยานิพนธ์ : การตรวจจับและการแก้ไขการวางตัวของภาพโดยใช้ซอฟต์แวร์คอมพิวเตอร์
แมชชีน
สาขาวิชา : เทคโนโลยีคอมพิวเตอร์
สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ที่ปรึกษาวิทยานิพนธ์ : ผู้ช่วยศาสตราจารย์ ดร.กัณฑ์พงษ์ วรรณปัญญา
ผู้ช่วยศาสตราจารย์ ดร.สุพจน์ นิตย์สุวัฒน์
ปีการศึกษา : 2549

บทคัดย่อ

ในปัจจุบันการปรับทิศทางภาพเพื่อให้ได้ภาพในทิศทางที่ต้องการ มักใช้การปรับทิศทางภาพด้วยมือ แม้ว่าการปรับทิศทางภาพจะสามารถทำได้ง่ายและใช้กระบวนการทำเพียงไม่กี่ขั้นตอน แต่ในกรณีที่มีภาพจำนวนมาก เวลาที่เสียไปและกำลังแรงงานที่ใช้ไปจะเพิ่มมากขึ้น ดังนั้นในงานวิจัยนี้จึงได้นำเสนอเทคนิคการตรวจจับและแก้ไขทิศทางภาพให้ถูกต้อง โดยใช้วิธีการหาลักษณะเด่นของภาพจากค่าความสำคัญของสี และการหาค่าฮิสโตแกรมของขอบวัตถุภาพจากภาพย่อย จากนั้นทำการวิเคราะห์และรวมลักษณะเด่นของภาพจากข้อมูลภาพในทิศทางที่แตกต่างกันในแต่ละภาพ เพื่อแบ่งประเภทของภาพจากลักษณะเด่นที่ได้ ทำให้รู้ถึงทิศทางการวางตัวของภาพที่ป้อนเข้ามา สามารถตรวจจับและแก้ไขทิศทางการวางตัวของภาพได้ถูกต้องตามทิศทางที่เป็นจริงโดยใช้ซอฟต์แวร์คอมพิวเตอร์แมชชีนที่ได้ผ่านกระบวนการสอนให้ระบบเรียนรู้ เพื่อแยกแยะข้อมูลความแตกต่างของภาพจากข้อมูลภาพที่ใช้ในการสอนให้ระบบเรียนรู้ จำนวน 350 ภาพ และจำนวนภาพที่ใช้ในการทดสอบจำนวน 350 ภาพ สามารถให้ค่าความถูกต้องมากกว่าร้อยละ 88.00

(วิทยานิพนธ์มีจำนวนทั้งสิ้น 53 หน้า)

คำสำคัญ : การปรับทิศทางภาพ การตรวจจับภาพ ค่าความสำคัญของสี ค่าฮิสโตแกรมของขอบวัตถุภาพ กลุ่มทิศทาง

อาจารย์ที่ปรึกษาวิทยานิพนธ์

Name : Miss Chira Kaewsuwan
Thesis Title : Image Orientation Detection and Correction Using Support Vector Machine
Major Field : Computer Technology
King Mongkut's Institute of Technology North Bangkok
Thesis Advisors : Assistant Professor Dr. Kuntpong Woraratpanya
: Assistant Professor Dr. Supot Nitsuwat
Academic Year : 2006

Abstract

Currently, the operation of correcting image orientation is usually performed manually by technologists. Although the rotation operation is simple and takes a few moments, in case of a large digital photo archive, the accumulated time spent and cost required are considerable. Therefore, in this thesis, an automatic image orientation detection and correction is proposed. In this method, image features that are composed of color moments and edge direction histograms are extracted from sub-images. Then these image features are arranged and formed as a vector for each image. In order to create an automation system of detecting and correcting image orientation, the vectors of image features are modeled by using support vector machine. Based on 350 training images and 350 testing images, the experimental results show that the proposed method can perform accurately up to 88.00%.

(Total 53 pages)

Keywords : image orientation, image detection, color moment, edge direction histogram, bins

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จได้ด้วยความช่วยเหลืออย่างดียิ่งของ ผู้ช่วยศาสตราจารย์ ดร. กันต์ พงษ์ วรรณปัญญา และผู้ช่วยศาสตราจารย์ ดร. สุพจน์ นิตยสุวัฒน์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ผู้ให้ โอกาสและประสบการณ์ครั้งสำคัญ ตลอดจนคำแนะนำปรึกษาและแก้ไขข้อบกพร่อง ในการทำ วิทยานิพนธ์ครั้งนี้ ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอขอบพระคุณคุณพ่อ คุณแม่ ผู้เป็นกำลังใจที่ยิ่งใหญ่ที่เชื่อมั่นเสมอว่าลูกสามารถทำได้ ขอบพระคุณกำลังใจและความช่วยเหลือจากเพื่อนทุกคน คุณอริษกานต์ ไกรทอง เพื่อน จีบ รุต บอล อัน เตย บอย โย และโป้ง และขอกราบขอบพระคุณอาจารย์ภาควิชาคอมพิวเตอร์ศึกษาทุกท่าน

สุดท้ายนี้ประโยชน์และคุณความดีอันใด ที่เกิดจากวิทยานิพนธ์ฉบับนี้ ขอมอบให้แก่ คุณพ่อ คุณแม่ พี่และน้อง ครูอาจารย์และเพื่อนของข้าพเจ้าทุกท่านตลอดไป

จิรา แก้วสุวรรณ

สารบัญ

หน้า

บทคัดย่อภาษาไทย	ข
บทคัดย่อภาษาอังกฤษ	ค
กิตติกรรมประกาศ	ง
สารบัญตาราง	ช
สารบัญภาพ	ฅ
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์การวิจัย	2
1.3 ขอบเขตของการวิจัย	2
1.4 นิยามคำศัพท์	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ	3
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง	4
2.1 การค้นคืนภาพภาพโดยลักษณะเด่น	4
2.2 การหาค่าความสำคัญของสี	7
2.3 การหาค่าฮิสโตแกรมของวัตถุภาพ	8
2.4 ซัพพอร์ตเวกเตอร์แมชชีน	13
2.5 งานวิจัยที่เกี่ยวข้อง	21
บทที่ 3 การดำเนินงานวิจัย	22
3.1 ระบบการตรวจจับทิศทางการวางตัวของภาพโดยใช้ค่าความสำคัญของสี	22
3.2 ระบบการตรวจจับทิศทางการวางตัวของภาพด้วยใช้ค่าฮิสโตแกรม ของขอบวัตถุภาพ	25
3.3 ระบบการตรวจจับทิศทางการวางตัวของภาพด้วยการรวม 2 ลักษณะเด่น	26
3.4 วิธีการประเมินคุณภาพ	28
3.5 การเตรียมข้อมูลที่ใช้ในงานวิจัย	30
3.6 กระบวนการทำงานของซัพพอร์ตเวกเตอร์แมชชีน	34
บทที่ 4 ผลการทดลอง	36
4.1 ดำเนินการทดลอง	36
4.2 ดำเนินการทดลอง และผลการทดลอง	36

สารบัญ (ต่อ)

	หน้า
4.3 สรุปผลการทดลอง	47
บทที่ 5 สรุปและข้อเสนอแนะ	49
5.1 สรุปผลการทดลอง	49
5.2 ข้อเด่น และข้อด้อยของวิธีการที่พัฒนาขึ้นใหม่	49
5.3 ข้อเสนอแนะสำหรับงานวิจัยในอนาคต	50
5.4 ปัญหาของงานวิจัย	50
บรรณานุกรม	51
ประวัติผู้วิจัย	53

สารบัญตาราง

ตารางที่	หน้า
3-1 ข้อมูลที่ใช้ในงานวิจัย	30
3-2 ลักษณะการเก็บข้อมูลภาพด้วยค่าความสำคัญของสี	32
4-1 ผลการทดลองด้วยค่าความสำคัญของสีในรูปแบบที่ 1	37
4-2 ผลการทดลองด้วยค่าความสำคัญของสีในรูปแบบที่ 2	37
4-3 ผลการทดลองด้วยค่าความสำคัญของสีในรูปแบบที่ 3	38
4-4 ผลการทดลองด้วยค่าความสำคัญของสีในรูปแบบที่ 4	38
4-5 ผลการทดลองด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ในรูปแบบที่ 1	39
4-6 ผลการทดลองด้วยค่าฮิสโตแกรมของขอบเขตภาพ ในรูปแบบที่ 2	39
4-7 ผลการทดลองด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ในรูปแบบที่ 3	39
4-8 ผลการทดลองด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ในรูปแบบที่ 4	40
4-9 ผลการทดลองด้วยการรวม 2 ลักษณะเด่น ในรูปแบบที่ 1	40
4-10 ผลการทดลองด้วยการรวม 2 ลักษณะเด่น ในรูปแบบที่ 2	41
4-11 ผลการทดลองด้วยการรวม 2 ลักษณะเด่น ในรูปแบบที่ 3	41
4-12 ผลการทดลองด้วยการรวม 2 ลักษณะเด่น ในรูปแบบที่ 4	41
4-13 ผลการทดลองของระบบด้วยค่าความสำคัญของสี	42
4-14 ผลการทดลองของระบบด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ	43
4-15 ผลการทดลองของระบบด้วยการรวม 2 ลักษณะ	45
4-16 ผลการทดลองเปรียบเทียบด้านค่าความถูกต้องในการตรวจจับ	47
4-17 ผลการทดลองเปรียบเทียบเกี่ยวกับเวลาที่ใช้ในการประมวลผล	47
4-18 ผลการทดลองเปรียบเทียบด้วยขนาดของค่าลักษณะเด่น	47

สารบัญภาพ

ภาพที่	หน้า
2-1 แบบจำลองสีโหมค RGB	5
2-2 ตัวอย่างการหาของด้วยวิธี Gradient	9
2-3 การกำหนดค่า Threshold ในการแบ่งเส้นขอบด้วยวิธี Gradient และวิธี Laplacian	10
2-4 ตัวอย่างการหาของด้วยวิธี Laplacian	10
2-5 ขั้นตอนวิธีการหาขอบโดยวิธี Canny	11
2-6 ตำแหน่งข้อมูลสองกลุ่มในฟีเจอร์สเปซ (Feature Space)	14
2-7 การวางตัวของข้อมูลในลักษณะเชิงเส้น	16
2-8 การแบ่งกลุ่มข้อมูลที่มีข้อผิดพลาดโดยจุดข้อมูลที่อยู่ในแกนซัพพอร์ตทั้งสองข้าง	18
2-9 การเปลี่ยนมิติข้อมูลไปยังมิติที่สูงขึ้นด้วยฟังก์ชันของเคอร์เนล	19
3-1 ผังขั้นตอนระบบตรวจจับทิศทางการวางตัวของภาพด้วยค่าความสำคัญของสี	23
3-2 ขั้นตอนวิธีการดึงลักษณะเด่นด้วยค่าความสำคัญของสี	24
3-3 ผังขั้นตอนระบบตรวจจับทิศทางการวางตัวของภาพด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ	25
3-4 ขั้นตอนวิธีการดึงลักษณะเด่นด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ	26
3-5 ผังขั้นตอนระบบตรวจจับทิศทางการวางตัวของภาพด้วยการรวม 2 ลักษณะเด่น	27
3-6 ขั้นตอนวิธีการดึงลักษณะเด่นด้วยการรวม 2 ลักษณะเด่น	27
3-7 การจัดวางเวกเตอร์ข้อมูลโดยการรวม 2 ลักษณะเด่น	28
3-8 ตัวอย่างรูปภาพทั้ง 4 รูปแบบ	31
3-9 ลักษณะการจัดเก็บเวกเตอร์ข้อมูลด้วยค่าความสำคัญของสี	33
3-10 ลักษณะการจัดเก็บเวกเตอร์ข้อมูลด้วยฮิสโตแกรมของขอบวัตถุภาพ	33
3-11 ลักษณะการจัดเก็บเวกเตอร์ข้อมูลด้วยการรวม 2 ลักษณะเด่น	33
3-12 ทิศทางการหมุนภาพครั้งละ 90°	34
4-1 กราฟแสดงค่าความถูกต้องจากผลการทดลองของระบบด้วยค่าความสำคัญของสี	43
4-2 กราฟแสดงค่าความถูกต้องจากผลการทดลองของระบบด้วยค่าฮิสโตแกรม ของขอบวัตถุภาพ	44
4-3 กราฟแสดงค่าความถูกต้องจากผลการทดลองของระบบด้วยการรวม 2 ลักษณะเด่น	45
5-1 ลักษณะภาพที่เกิดปัญหาในการตรวจจับ	50

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันเทคโนโลยีของกล้องดิจิทัล (Digital Camera) มีการพัฒนาอย่างรวดเร็ว และเทคโนโลยีได้เข้ามาแทนที่กล้องถ่ายภาพที่ใช้ฟิล์มอย่างรวดเร็ว ด้วยรูปแบบการใช้งานที่ง่าย และสามารถถ่ายภาพได้ครั้งละจำนวนมาก จึงทำให้กล้องดิจิทัลกำลังได้รับความนิยมอย่างสูงจากผู้ใช้งานในขณะนี้

โดยปกติในการใช้งานกล้องดิจิทัล ก่อนที่มนุษย์จะทำการบันทึกภาพเข้าสู่กล้องดิจิทัล มนุษย์เป็นผู้กำหนดทิศทางการจัดวางของภาพเอง ซึ่งจากรูปแบบของการใช้งานกล้องดิจิทัล โดยทั่วไปสามารถเลือกทิศทางการถ่ายภาพด้วยกล้องดิจิทัลตามมุมมองในการจัดวางสำหรับการถ่ายภาพได้สองลักษณะด้วยกัน คือ ลักษณะการจัดวางในแนวตั้งและลักษณะการจัดวางในแนวนอน ดังนั้นเมื่อภาพที่ถ่ายจากกล้องดิจิทัลเหล่านี้ผ่านกระบวนการโอนถ่ายข้อมูล เข้าสู่เครื่องคอมพิวเตอร์ ภาพที่เกิดจากการจัดวางตามทิศทางที่ต้องการจากผู้ถ่ายภาพ ทำให้เกิดปัญหาทิศทางการวางตัวของภาพดิจิทัลที่แตกต่างกันออกไป ซึ่งปัญหาความแตกต่างของทิศทางของภาพเหล่านี้สามารถเป็นไปได้ 4 ทิศทางด้วยกันคือ 0° , 90° , 180° และ 270°

ปัญหาทิศทางการวางตัวของภาพดิจิทัลไม่ถูกต้องตามทิศทางที่เป็นจริงนี้ ในปัจจุบันผู้ใช้งานแก้ปัญหาการวางตัวของภาพดิจิทัลที่เกิดจากการวางตัวที่ไม่ถูกต้องตามทิศทางที่เป็นจริง ด้วยวิธีการเลือกปรับเปลี่ยนทิศทางการวางตัวของภาพดิจิทัล โดยใช้เครื่องมือสำหรับเปลี่ยนทิศทางการวางตัวของภาพในโปรแกรมสำหรับเปิดภาพโดยทั่วไป ซึ่งต้องทำการแก้ไขด้วยวิธีการเลือกภาพเพื่อทำการปรับเปลี่ยนทิศทาง ซึ่งหากจำนวนภาพดิจิทัลมีปริมาณมาก ผู้ใช้งานต้องทำการเลือกภาพที่มีปัญหาเองแล้ว ย่อมส่งผลให้ผู้ใช้งานสูญเสียเวลา และอาจทำให้เกิดความเบื่อหน่ายในการปรับเปลี่ยนทิศทางการวางตัวของภาพให้ถูกต้องตามทิศทางเป็นจริงตามไปด้วย

ในปัจจุบันยังไม่พบงานวิจัยที่เกี่ยวข้องกับการตรวจจับและแก้ไขทิศทางการวางตัวของภาพดิจิทัลในเมืองไทยแต่อย่างใด โดยมีแต่งานวิจัยของต่างประเทศโดย Vailaya [2] เป็นผู้บุกเบิกในยุคแรก ได้ทำการหาลักษณะเด่นของภาพจากค่าความสำคัญของสี (Color Moments) ต่อมา Wang และ Zhang [1] ได้พัฒนาต่อโดยเพิ่มการดึงลักษณะเด่นของภาพด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ (Edge Direction Histogram) และใช้วิธีการแบ่งกลุ่มข้อมูลด้วยซัพพอร์ตเวกเตอร์แมชชีน

(Support Vector Machine) ซึ่งมีประสิทธิภาพในด้านความเร็ว และความสามารถในการแยกแยะกลุ่มข้อมูลที่ดี ด้วยวิธีการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งข้อมูลออกจากกัน

จากวิธีการดึงลักษณะเด่นด้วยค่าความสำคัญของสี และค่าฮิสโตแกรมของขอบวัตถุภาพ และเทคนิคการแบ่งกลุ่มข้อมูลด้วยซัพพอร์ตเวกเตอร์แมชชีนที่มีประสิทธิภาพดี ผู้วิจัยจึงได้นำมาทำการพัฒนาต่อยอด เพื่อลดขนาดของข้อมูล จากวิธีการเดิมของ Wang และ Zhang [1] ในส่วนของขั้นตอนการหาค่าฮิสโตแกรมของขอบวัตถุภาพที่มีขนาด 925 ซึ่งถือว่ามีความใหญ่ เพื่อช่วยในการประมวลผล ให้มีประสิทธิภาพเร็วยิ่งขึ้น

1.2 วัตถุประสงค์การวิจัย

เพื่อพัฒนาขั้นตอนวิธีการตรวจจับและแก้ไขการวางตัวของภาพดิจิทัล โดยใช้ซัพพอร์ตเวกเตอร์แมชชีน

1.3 ขอบเขตของการวิจัย

1.3.1 ภาพที่นำมาทำการทดสอบเป็นภาพดิจิทัล ขนาด 192 x 128 พิกเซล (Pixels)

1.3.2 ภาพที่ใช้ในการทดสอบมีจำนวน กลุ่มละ 100 ภาพ แบ่งเป็นกลุ่มการสอนให้ระบบเรียนรู้ 50 ภาพ และกลุ่มทดสอบ 50 ภาพ

1.3.3 ใช้วิธีดึงลักษณะเด่นของภาพด้วยค่าความสำคัญของสี และค่าฮิสโตแกรมของขอบวัตถุภาพ

1.3.4 ใช้โปรแกรม MATLAB เวอร์ชัน 6.5 ในการวิเคราะห์และทำการทดลอง

1.4 นิยามคำศัพท์เฉพาะ

1.4.1 Color Moments หมายถึง การหาค่าความสำคัญของสี เพื่อดึงลักษณะเด่นของภาพ

1.4.2 Edge Direction Histogram หมายถึง การหาค่าฮิสโตแกรมของขอบวัตถุภาพ เพื่อดึงลักษณะเด่นของภาพ

1.4.3 Classification หมายถึง การจำแนกออกเป็นประเภท หรือการแบ่งกลุ่มข้อมูล

1.4.4 Feature Extraction หมายถึง การดึงลักษณะเด่น

1.4.5 Training หมายถึง การสอนให้ระบบเรียนรู้

1.4.6 Testing หมายถึง กระบวนการทดสอบการหาผลลัพธ์

1.4.7 Support Vector Machine หมายถึง วิธีการวิเคราะห์ข้อมูลเชิงเส้น โดยเน้นไปยังเส้นแบ่งแยกกลุ่มข้อมูลที่ดีที่สุด

1.4.8 Bins หมายถึง กลุ่มทิศทาง

1.4.9 Image orientation หมายถึง การปรับทิศทางภาพ

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 ได้วิธีการตรวจจับและแก้ไขทิศทางการวางตัวของภาพดิจิทัลให้ถูกต้องตามทิศทางที่เป็นจริง

1.5.2 เป็นแนวทางให้ผู้สนใจได้ค้นคว้าและศึกษาเกี่ยวกับการตรวจจับและแก้ไขทิศทางการวางตัวของภาพดิจิทัล

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

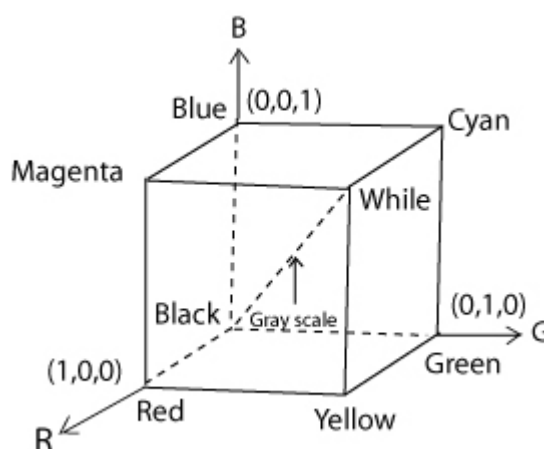
ในบทนี้เป็นการเสนอทฤษฎีที่จะนำมาประยุกต์ใช้ในการแก้ปัญหา การตรวจจับและแก้ไขทิศทางการวางตัวของภาพดิจิทัล โดยเสนอทฤษฎีที่เกี่ยวข้องสำหรับงานวิจัยนี้ เริ่มจากทฤษฎีการค้นคืนภาพโดยลักษณะเด่นของภาพ การหาค่าความสำคัญของสี การหาค่าฮิสโตแกรมของขอบวัตถุภาพ และทฤษฎีเกี่ยวกับซอฟต์แวร์เวกเตอร์แมชชีน

2.1 การค้นคืนภาพโดยลักษณะเด่น

เทคนิคการค้นคืนภาพในปัจจุบันได้ถูกเสนอขึ้นมาหลายวิธี เทคนิคเหล่านี้สามารถแบ่งตามลักษณะการค้นคืนภาพด้วยหลักเกณฑ์พื้นฐานที่ใช้โดยทั่วไปได้ 2 รูปแบบด้วยกัน ได้แก่ การค้นคืนภาพโดยใช้พื้นฐานทางตัวอักษร (Text-Based Image Retrieval) และการค้นคืนภาพโดยใช้พื้นฐานทางเนื้อหาของภาพ (Content-Based Image Retrieval) โดยแต่ละวิธีมีข้อดีและข้อเสียแตกต่างกัน ในการค้นคืนภาพจากข้อความ แต่ละภาพจะถูกอธิบายด้วยคำสำคัญและทำการค้นหาภาพโดยเปรียบเทียบหาคำสำคัญตรงกับข้อความที่สอบถาม สำหรับการค้นคืนภาพทางเนื้อหาแต่ละภาพจะถูกวิเคราะห์หาลักษณะเด่นที่สำคัญเพื่อหาตัวแทนของข้อมูลภาพสำหรับใช้แสดงถึงภาพนั้นๆ ซึ่งการหาลักษณะเด่นของภาพ ในงานวิจัยที่เกี่ยวข้องกับงานวิจัยชิ้นนี้ ส่วนใหญ่ได้นำเอาเทคนิคการค้นคืนรูปภาพโดยใช้ระบบค้นคืนเชิงเนื้อหาของภาพ หรือเรียกโดยย่อว่า CBIR นี้เข้ามาใช้เพื่อหาตัวแทนของข้อมูลภาพในการวิเคราะห์รูปภาพโดยอัตโนมัติ ซึ่งวิธีการแทนรูปภาพด้วยคุณลักษณะของภาพโดยทั่วไปที่นิยมใช้คือ คุณสมบัติของสี (Color) ลักษณะพื้นผิว (Texture) และรูปร่าง (Shape) เป็นต้น

2.1.1 คุณสมบัติของสี โดยทั่วไปในหาตัวแทนของข้อมูลภาพในการค้นคืนภาพ ได้นำเอาเทคนิคของสีมาใช้กันอย่างแพร่หลายมากที่สุด เนื่องจากสามารถทำได้ง่ายและมีความแน่นอนในระดับหนึ่ง โดยทั่วไปการแสดงผลภาพสีจะอยู่ในรูปแบบของสัญญาณสีแดง (R) สีเขียว (G) และสีน้ำเงิน (B) หรือเรียกว่า “RGB Color Space” นอกจากนี้ยังมีระบบสีอื่นๆ ที่นิยมใช้อย่างกว้างขวาง ได้แก่ แบบจำลองสีแบบ YUV แบบจำลองสีแบบ LAB แบบจำลองสีแบบ CMY และระบบสี HSV (หรือ HSL, HSB) ซึ่งจะกล่าวถึงแบบจำลองสีที่เกี่ยวข้องกับงานวิจัยนี้ได้แก่

2.1.1.1 แบบจำลองสี RGB [14] เป็นระบบที่ใช้กันโดยทั่วไป ใช้ในการแสดงผลบนจอมอนิเตอร์ เนื่องจากการกำหนดความเข้มของลำอเล็กตรอนที่ยิงไปกระทบบนจอภาพ แบบจำลองแสดงดังภาพที่ 2-1 ประกอบด้วยค่าของแม่สีหลัก 3 สี ได้แก่ ค่าของความหนาแน่นของสีแดง (Red) สีเขียว (Green) และสีน้ำเงิน (Blue) รูปภาพที่นำมาทำการประมวลผล เพื่อจัดเก็บลงฐานข้อมูลใช้ภาพดิจิทัลสีขนาด 24 บิตต่อพิกเซล นั่นคือ แต่ละสีมีข้อมูลได้ 8 บิต ดังนั้นแต่ละพิกเซลสามารถแทนค่าสีได้ถึง 16,777,216 สี



ภาพที่ 2-1 แบบจำลองสีโหมด RGB

2.1.1.2 แบบจำลองสี XYZ (XYZ Color Model) เป็นตัวกลางอ้างอิงการเปลี่ยนแปลงค่าระหว่างแบบจำลองสีทั้งหมด โดยหลักการคือ การให้แหล่งกำเนิดแสงที่ได้ทำการสมมุติขึ้นในอุดมคติเพื่อใช้เป็น Primary Colors สำหรับการผสมสีแบบ Additive ซึ่งเป็นการผสมสีของแสงสี และทำให้เกิดสีต่างๆ ขึ้น สีที่เกิดขึ้นจากการผสมนี้มนุษย์ไม่สามารถมองเห็นได้ทั้งหมด XYZ Model ประกอบด้วยสามส่วนคือ X Y และ Z ตามลำดับ โดย XYZ Model นี้จะคำนึงถึงระดับความสว่าง หรือ Illuminance ด้วย นอกเหนือจากข้อมูลที่บอกความแตกต่างของสี แต่ต้องการคำนึงถึงระดับความสว่างของสีเพียงอย่างเดียวก็สามารถที่จะแปลงให้อยู่ในรูป X และ Y โดยทางเลือกอีกวิธีหนึ่งคือ การใช้ระบบ xyY โดยที่ x และ y สองตัวแรกจะระบุตำแหน่งของสีบน CIE Diagram ส่วน Y ที่เพิ่มเข้ามาเพื่อบอกค่าความสว่างของสีนั้นๆ ทำให้สามารถที่จะเปลี่ยนกลับไปให้อยู่ในรูป XYZ Model ได้

เหตุที่ใช้ค่า Y เนื่องจากว่าในขั้นตอนการสมมุติแหล่งกำเนิดแสง Standard Primary Colors นั้น ได้มีการกำหนดให้ปริมาณของ Y ที่ใช้ในการกระตุ้น Cone Cell ทั้งสาม เพื่อที่จะทำให้เกิดเป็นสี

ต่างๆ ตลอดสเปกตรัมนั้นเท่ากับความสามารถในการตอบสนองต่อการถูกกระตุ้นของแสงแต่ละสี ตลอดทั้งสเปกตรัมของ Cone Cell ทั้งสามรวมกัน ความสามารถของ Cone Cell ทั้งสามดังกล่าว สามารถที่จะนำมาสร้างความสัมพันธ์ตลอดทั้งสเปกตรัมได้ เรียกว่า “Photopic Spectral Luminous Efficiency Curve” หรือ Curve มีลักษณะเป็นรูปโค้งคว่ำ ดังนั้นจะเห็นว่าเมื่อนำ Core Cell ทั้ง 3 รวมกัน จะมีความไวต่อแสงสีในช่วงกลางของสเปกตรัมซึ่งคือ บริเวณของสีเหลืองเขียวมากที่สุด และลดลงเมื่อค่อนข้างไปทางด้านสีแดงม่วงจนไม่สามารถมองเห็นได้ในที่สุด

จะเห็นได้ว่าส่วนที่เป็นข้อมูลที่บอกความแตกต่างของสี คือ xy และส่วนที่บอกความสว่างคือ Y โดยที่ xy นี้ จะถูกเรียกว่า Chromaticity Value ส่วน Y จะถูกเรียกว่า Luminance Value ถ้าหากมี ภาพที่ใช้ xyY Model โดยเราสามารถที่จะใช้เพียงค่า Y เพียงค่าเดียวเพื่อทำการแสดงภาพที่มีความ ถูกต้องด้านความสว่าง และ Contrast ของภาพแต่จะเป็นภาพที่มีสีขาวดำเหมือนกับการที่ดูภาพจาก โทรทัศน์ขาวดำ ซึ่งแตกต่างจาก RGB Model ที่จะไม่สามารถใช้ Channel หนึ่ง เพื่อทำการแสดง ภาพที่มีความสว่างและ Contrast ของภาพได้

2.1.1.3 แบบจำลองสี LUV [14] เป็นอีกรูปแบบหนึ่งที่สามารถสื่อความหมายและ สอดคล้องกับการแสดงความรู้สึกของมนุษย์ได้ การแปลงค่าจากแบบจำลองสี RGB เป็น แบบจำลองสี LUV จะต้องทำการแปลงค่า จาก RGB เป็น XYZ แล้วทำการแปลงค่า XYZ เข้าสู่ค่า ของแบบจำลอง LUV โดยกำหนดให้ XYZ เป็นค่าเริ่มต้นโดยค่าสีทั้งหมดจะถูกกำหนดเพียงค่าที่ เป็นค่าบวกเท่านั้น ที่ซึ่งค่า Y ตัวแรกจะถูกกำหนดแบบตายตัว เพียงรองรับคุณลักษณะของแสงและสี จากนั้นรวมค่า XYZ แสดงออกมาในลักษณะของกราฟ จะได้ Lowercase xyz นำมาทำนอมอนไลซ์ ได้ค่าของ xyz เป็น $x + y + z = 1$ ดังนั้น $x = X/(X+Y+Z)$, $y = Y/(X+Y+Z)$, $z = Z/(X+Y+Z)$ ด้วยเหตุนี้ xyz จึงอยู่บน $X+Y+Z = 1$ ของสีในรูปแบบ xyz มีลักษณะ Solid ในขั้นตอน Lowcase นี้ มีชื่อ เรียกว่า Chromaticity Value และแปรเปลี่ยนเป็น xy spces ของ $(X+Y+Z=1)$ spces ซึ่งถูกเรียกว่า Chromaticity Diagram ซึ่งในไดอะแกรมนี้โดยทั่วไปแล้วค่าสีพื้นฐานจะเหมือนกันแต่จะต่างกันที่ แสงสี (Brightness) เปรียบเทียบกับจุดที่เหมือนกัน ณ ตำแหน่งเดียวกัน ได้ไดอะแกรมพื้นฐานด้วย การกำหนด C-illuminant หรือ จุดสีขาว และนับจำนวนตัวเลขที่อยู่รอบนอกที่มีขนาดสอดคล้องกัน เกือบไว้ใน Wevelangnt ของนาโนมิเตอร์ ซึ่ง LUV space นั้น ใช้รูปแบบเดียวกันกับ CIE XYZ space

2.1.2. พื้นผิว (Texture) เป็นรูปแบบการเรียงตัวของสี เป็นวิธีที่ดีในการแยกประเภทของ รูปภาพ แต่ข้อเสียคือไม่สามารถอธิบายเป็นคำพูดที่ใช้อธิบายรูปภาพได้ เทคนิคต่างๆ ได้ออกเป็น 3 ประเภท คือ ความน่าจะเป็นกับสถิติ (Probabilistic/Statistic) แถบสี (Spectral) และ โครงสร้าง (Structurel)

2.1.3. รูปร่าง (Shape) แบ่งออกเป็น 2 ประเภทคือ เส้นขอบ (Boundary-based) และบริเวณ (Region-based) เส้นขอบจะพิจารณาเฉพาะเส้นล้อมรอบวัตถุ ส่วนบริเวณจะพิจารณาพื้นที่ทั้งหมดที่อยู่ในเส้นล้อมรอบวัตถุ คุณลักษณะของรูปร่างที่ต้องการจะต้องไม่ขึ้นกับการเปลี่ยนตำแหน่ง การหมุน และการเปลี่ยนขนาด วิธีหนึ่งที่มีคุณสมบัตินี้คือ Anglogram เป็นเทคนิคที่พัฒนามาจากวิธีการทำสามเหลี่ยมที่ชื่อว่า “Delauney Triangulation”

2.1.4. การทำดัชนีพื้นที่ของสี (Spatial Color Indexing) ในแต่ละวัตถุในรูปภาพจะมีบริเวณเป็นของตัวเอง วิธีนี้จะกำหนดดัชนีให้กับบริเวณของวัตถุต่างๆ ในรูปภาพ ในระยะแรกมีโปรแกรมที่ติดต่อผู้ใช้ที่สามารถกำหนดขอบเขตของวัตถุเพื่อใช้ในการค้นหาได้ ต่อมาได้มีการปรับปรุงโดยการรวมลักษณะทั้งสีและรูปร่างเข้าด้วยกัน

2.2 การหาค่าความสำคัญของสี

ค่าความสำคัญของสี [11] นี้ ถือเป็นเครื่องมือในการบอกความแตกต่างของลักษณะเด่นของสี ในอดีตใช้สำหรับการคำนวณหาค่าความคล้ายคลึงกันของสีบนภาพ ซึ่งค่าความคล้ายคลึงที่ได้มาสามารถใช้เป็นตัวชี้เพื่อเปรียบเทียบกับข้อมูลภาพในฐานข้อมูล สำหรับระบบการค้นคืนภาพเชิงเนื้อหา จากงานวิจัยที่ได้ทำการทดลองพบว่าสามารถให้ผลสำเร็จที่ดีในการหาการกระจายตัวที่คล้ายคลึงของพื้นที่สีบนภาพ

พื้นฐานของค่าความสำคัญของสีนี้ ถูกสร้างขึ้นมาจากสมมุติฐานการกระจายตัวของค่าสีบนภาพที่สามารถให้ความหมายความน่าจะเป็นของการกระจายของสีข้อมูลในแต่ละพื้นที่ โดยจะแสดงลักษณะพิเศษออกมาที่สามารถพิจารณาได้จากตัวเลขแต่ละค่าความสำคัญ เพราะฉะนั้นผลที่ได้มาจากค่าสีของภาพก็คือค่าเฉพาะในแต่ละค่าจากการกระจายตัวของข้อมูลสีนี้ ซึ่งค่าความสำคัญของสีนี้ที่ได้จากการดึงลักษณะเด่นนี้ จะเป็นตัวระบุค่าสีบนภาพ

Stricker และ Orengo [6] ใช้ 3 ค่าความสำคัญหลักในการหาค่าการกระจายตัวของสี นั่นก็คือ ค่าเฉลี่ย (Mean) ค่าความแปรปรวน (Variance) และค่าความบิดเบือน (Skewness) ค่าที่ได้จากทั้ง 3 ค่าความสำคัญหลักนี้ จะให้ผลลัพธ์ค่าสี 9 จำนวนด้วยกัน ซึ่งได้จากการคำนวณในแต่ละแชนแนลของภาพ ซึ่งสมการที่ใช้สำหรับการคำนวณค่าความสำคัญของสี สามารถหาได้จากสมการที่ (2-1), (2-2) และ (2-3) ตามลำดับ

$$\mu_i = \frac{1}{N} \sum_{j=1}^N p_{ij} \quad (2-1)$$

$$\sigma_i = \sqrt{\left(\frac{1}{N} \sum_{j=1}^N (p_{ij} - \mu_i)^2\right)} \quad (2-2)$$

$$S_i = \sqrt[3]{\left(\frac{1}{N} \sum_{j=1}^N (p_{ij} - \sigma_i)^3\right)} \quad (2-3)$$

เมื่อ μ_i คือ ค่าเฉลี่ย

σ_i คือ ค่าความแปรปรวน

S_i คือ ค่าความบิดเบือน

P_{ij} คือ ค่าของ i-th ที่เกิดในแต่ละ Channel ของสี จาก j-th ของพิกเซลภาพ

N คือ ผลรวมของพิกเซลทั้งหมดที่เกิดขึ้นบนภาพ

โดยปกติค่าความสำคัญของสีนี้ สามารถคำนวณได้ทั้งระบบสีแบบจำลองสีแบบ YCrCb แบบจำลองสีแบบ LAB แบบจำลองสีแบบ CMY และแบบจำลองสี HSV (หรือ HSL, HSB) เป็นต้น แต่ในงานวิจัยเกี่ยวข้อง [1], [2], [8] ได้ทำการทดลองและหาความสำคัญของสี โดยใช้สมการเพื่อหาค่าความสำคัญของสีเพียง 2 ค่า คือ ค่าเฉลี่ยและค่าความแปรปรวน ที่สามารถหาค่าความสำคัญของสีได้อย่างมีประสิทธิภาพและพบว่าสามารถให้ค่าความถูกต้องสูงจากความนิ่งของความอึมตัวของค่าสี เมื่อกระทำบนภาพในแบบจำลองสีแบบ LUV

2.3 การหาค่าฮิสโตแกรมของวัตถุภาพ (Edge Detection Histogram)

เทคนิคการหาขอบของวัตถุภาพเป็นคุณสมบัติหนึ่งที่สามารถเป็นตัวแทนของวัตถุในภาพได้อย่างชัดเจน ด้วยคุณสมบัติของขอบที่ไม่ขึ้นกับการเปลี่ยนตำแหน่ง การหมุน และการเปลี่ยนขนาด เมื่อนำมาสร้างฮิสโตแกรมแล้วสามารถได้ลักษณะเด่นสำหรับการเปรียบเทียบข้อมูลได้อย่างชัดเจน ซึ่งการหาค่าฮิสโตแกรมของขอบวัตถุภาพ [1] นี้ ได้นำเอาเทคนิคการหาขอบภาพโดยวิธีการ Canny มาประยุกต์ใช้ โดยการดัดแปลงขั้นตอนจากวิธีการหาขอบของ Canny หลังจากการคำนวณค่า Gradient แล้วจึงคำนวณพิกเซลภาพตามทิศทางค่ามุมของแต่ละกลุ่มทิศทาง ซึ่งสามารถแสดงทิศทางของจุดขอบบนพิกเซลภาพได้ ในงานวิจัยของ Wang และ Zhang [1] ได้ใช้วิธีการหาฮิสโตแกรม

ของวัตถุภาพจำนวนทั้งหมด 37 กลุ่มทิศทาง โดยกลุ่มทิศทางที่ 36 แรกมาจากการนับจำนวนของจุดขอบภาพที่สามารถเกิดทิศทางบนวัตถุภาพ เมื่อทำการควอนไทซ์ ครั้งละ 10 องศา และกลุ่มทิศทางสุดท้ายเป็นการนับจำนวนจุดขอบทั้งหมดที่ไม่มีทิศทางมุม คำนวณได้จากสมการที่ (2-4)

$$H(i) = h(i) / M \quad (2-4)$$

เมื่อ $H(i)$ คือ จำนวนจุดขอบในแต่ละทิศทางที่เกิดขึ้น

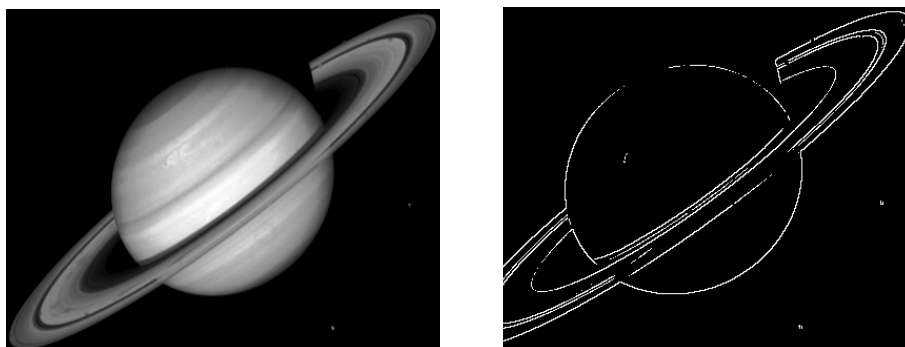
M คือ จำนวนพิกเซลทั้งหมดที่เกิดขึ้นบนวัตถุภาพ

2.3.1 วิธีการหาขอบภาพ (Edge Detection)

การหาขอบภาพ คือ การตรวจสอบว่าเส้นขอบลากผ่านหรือใกล้เคียงกับจุดใด โดยคำนวณจากการเปลี่ยนแปลงของความเข้มในตำแหน่งที่ใกล้เคียงกับจุดดังกล่าว [17] (ภาพที่ 2-2 (ก)) ซึ่งวิธีการหาขอบนั้นมีด้วยกันหลายวิธี แต่อย่างไรก็ตามสามารถแบ่งได้เป็น 2 กลุ่มหลัก คือวิธี Gradient และวิธี Laplacian โดยในแต่ละวิธีมีรายละเอียดดังต่อไปนี้

2.3.1.1 วิธี Gradient

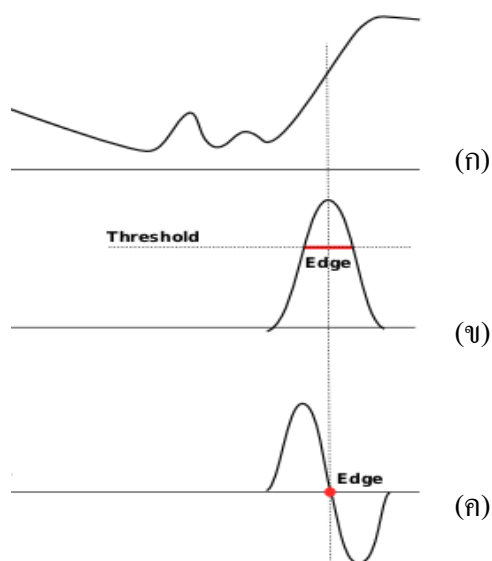
เป็นความคิดแรกที่ใช้วิเคราะห์หาขอบภาพ โดยใช้วิธีการวัดความเป็นแปลงความเข้ม (Gray level) ของจุดภาพที่กำลังพิจารณากับจุดภาพที่อยู่ข้างเคียง เพื่อการตัดสินใจต่อไปว่าเป็นขอบภาพหรือไม่ โดยวิธีนี้จะหาขอบด้วยการหาจุดต่ำสุดและจุดสูงสุดในรูปของอนุพันธ์อันดับหนึ่งของภาพ โดยจุดที่เป็นขอบจะอยู่ในส่วนที่เหนือค่า Threshold (ภาพที่ 2-2 (ข)) จึงอาจทำให้เส้นขอบที่ได้มีลักษณะหนา ตัวอย่างวิธีการหาขอบของกลุ่มนี้ เช่น Roberts, Prewitt และ Canny เป็นต้น ดังภาพตัวอย่างที่ 2-2



ภาพที่ 2-2 ตัวอย่างการหาขอบด้วยวิธี Gradient

2.3.1.2 วิธี Laplacian

วิธีนี้จะหาขอบโดยใช้อนุพันธ์อันดับ 2 โดยใช้จุดที่ค่า y เป็น “0” (ภาพที่ 2-2 (ค)) ซึ่งวิธีนี้จะใช้เวลาในการคำนวณมากกว่าวิธี Gradient สามารถใช้กรองความถี่ต่ำได้ขอบโครงภาพ ตัวอย่างวิธีการหาขอบของกลุ่มนี้ เช่น Laplacian of Gaussian และ Marrs-Hildreth เป็นต้น

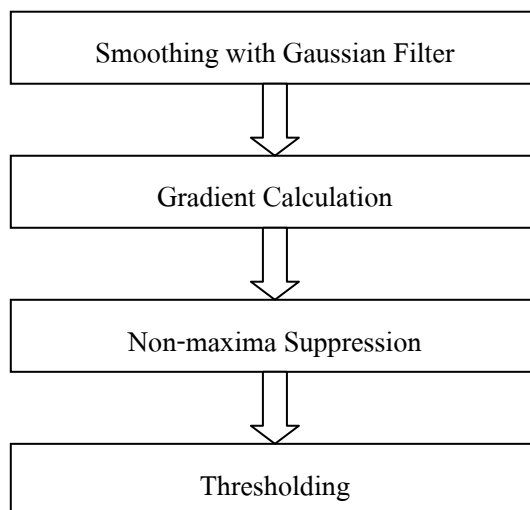


ภาพที่ 2-3 การกำหนดค่า Threshold ในการแบ่งเส้นขอบด้วยวิธี Gradient และวิธี Laplacian



ภาพที่ 2-4 ตัวอย่างการหาขอบด้วยวิธี Laplacian

ในงานวิจัยนี้เลือกใช้วิธีหาขอบ ด้วยวิธี Canny ซึ่งเป็นวิธีการหาขอบที่สามารถเลือกทิศทางการเกิดขอบภาพได้ และได้ผลดีในการหาขอบในภาพ ซึ่งอยู่ในประเภทวิธีการหาขอบแบบ Gradient ขั้นตอนวิธีการหาขอบโดยวิธีของ Canny แสดงดังภาพที่ 2-5



ภาพที่ 2-5 ขั้นตอนวิธีการหาขอบโดยวิธี Canny

การทำงานของ Canny Edge Detection นั้นเริ่มต้นจากการปรับภาพให้เรียบ (Smoothing) ด้วยตัวกรองเกาส์เซียน (Gaussian Filter) เพื่อกำจัดสัญญาณรบกวน หลังจากนั้น คำนวณค่าขนาด (Magnitude) และทิศทาง (Orientation) ของ Gradient โดยใช้การหาอนุพันธ์อันดับหนึ่ง ในขั้นตอนถัดมาจึงใช้ Non-maxima suppression กับ Gradient Magnitude เพื่อให้ได้ขอบที่บางลง และในขั้นตอนสุดท้ายใช้ Double Thresholding Algorithm เพื่อระบุพิกเซลที่เป็นขอบและช่วยเชื่อมต่อขอบ โดยในแต่ละขั้นตอนมีรายละเอียดดังต่อไปนี้

2.3.2.1 การขจัดสัญญาณรบกวน (Smoothing with Gaussian Filter)

ในขั้นตอนแรกของการหาขอบโดยอัลกอริทึมนี้จะต้องกำจัดสัญญาณรบกวนออกก่อนโดยใช้ตัวกรองเกาส์เซียน ซึ่งสามารถคำนวณได้จากการใช้หน้าต่าง (Mask) ขนาดเล็ก ขนาดของหน้าต่างเกาส์เซียนนี้ หากมีขนาดกว้างจะมีผลทำให้ลดสัญญาณรบกวนได้มาก แต่ถ้ากว้างมากเกินไปจะมีผลทำให้ขอบย่อยๆ ที่เป็นส่วนรายละเอียดนั้นหายไป สำหรับการคำนวณหาภาพที่ได้จากการใช้ตัวกรองเกาส์เซียน เป็นดังสมการที่ (2-5)

$$S_{ij} = I(i, j) * G(i, j; \sigma) \quad (2-5)$$

เมื่อ S_{ij} คือ ภาพภายหลังการกรอง
 $I(i, j)$ คือ ภาพที่ต้องการหาขอบ
 $G(i, j)$ คือ ตัวกรองแบบ Gaussian
 σ คือ ค่าเบี่ยงเบนมาตรฐานของการกระจาย

2.3.2.2 คำนวณหาทิศทางการเป็นขอบภาพ (Gradient Calculation)

ในขั้นแรกนำ Smoothing Image S_{ij} มาสร้าง x, y partial derivatives $P_{i,j}$ และ $Q_{i,j}$ ตามลำดับ ดังสมการที่ (2-6) และ (2-7)

$$P_{ij} = \{S_{i,j+1} - S_{i,j} + S_{i+1,j+1} - S_{i+1,j}\} / 2 \quad (2-6)$$

$$Q_{i,j} = \{S_{i,j} - S_{i+1,j} + S_{i,j+1} - S_{i+1,j+1}\} / 2 \quad (2-7)$$

เมื่อ $P_{i,j}$ คือ ค่าความแตกต่างในแกนแนวนอน
 $Q_{i,j}$ คือ ค่าความแตกต่างในแกนแนวตั้ง
 $S_{i,j}$ คือ ค่าความเข้มแสงของจุดภาพ

หลังจากนั้น ดำเนินการแปลงรูปแบบจาก Rectangular ไปเป็น Polar (Rectangular-To-Polar Conversion) เพื่อหาขนาดและทิศทางของ Gradient ตามสมการที่ (2-8)

$$M_{i,j} = \sqrt{P_{i,j}^2 + Q_{i,j}^2} \quad (2-8)$$

$$\theta = \arctan \frac{Q_{i,j}}{P_{i,j}} \quad (2-9)$$

จากสมการข้างต้นจะสามารถหาค่ามุม θ ออกมาได้ เมื่อแทนค่าตัวแปรในฟังก์ชัน $\arctan(x, y)$

2.3.2.3 การขจัดค่าที่ไม่มากที่สุด (Non-maxima Suppression)

สำหรับการหาขอบโดยวิธีการ Canny จุดที่ถือเป็นเส้นขอบได้นั้นต้องเป็นจุดที่ให้ค่าสูงสุดเฉพาะที่และเป็นทิศทางเดียวกับ Gradient ด้วย ซึ่งด้วยวิธีดังกล่าวนี้ทำให้ได้ขอบที่บางเพียง 1 พิกเซล ภาพที่ได้หลังการทำ Non-maxima Suppression จะให้ค่าเป็นศูนย์ในทุกจุดยกเว้นจุดที่เป็น Local Maxima Points ซึ่งจะยังคงค่าเดิมไว้

2.3.2.4 การกำหนดค่าขีดแบ่ง (Thresholding)

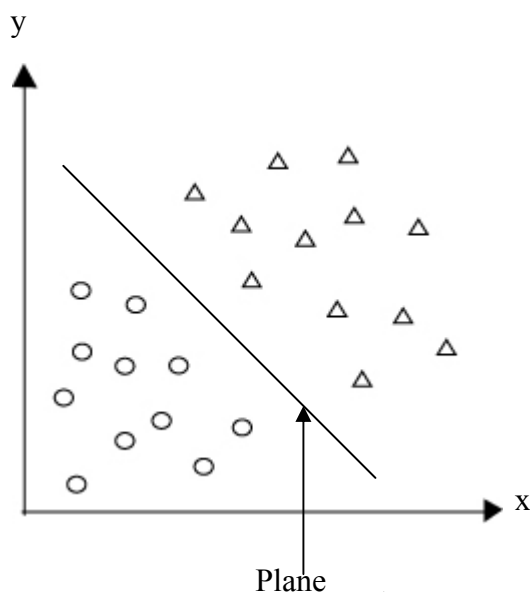
แม้ว่าภาพจะผ่านการขจัดสัญญาณรบกวนในขั้นตอนแรกแล้วก็ตาม ภาพที่ได้ก็ยังอาจมีเส้นขอบที่ไม่ใช่ขอบที่แท้จริงปรากฏอยู่อันเนื่องจากสัญญาณรบกวนหรือลักษณะของวัตถุในภาพเป็นพื้นผิวที่มีลวดลายหรือมีรายละเอียดภายในมาก ดังนั้น เพื่อลดปัญหาดังกล่าวจึงได้มีการกำหนดค่า Threshold ขึ้นมา 2 ค่า คือ High Threshold (T_1) และ Low Threshold (T_2) โดยพิกเซลที่มีค่ามากกว่า T_1 จะถูกปรับเป็น “1” (เป็นพิกเซลที่เป็นขอบ) แต่ถ้าน้อยกว่า T_2 จะถูกปรับเป็น 0 ส่วนค่าที่อยู่ระหว่างค่า Threshold ทั้งสอง การปรับเป็นค่า “0” หรือ “1” นั้นขึ้นอยู่กับพิกเซลที่อยู่รอบข้าง หากพบว่าพิกเซลที่อยู่รอบข้างของพิกเซลที่เป็นขอบ (ค่า $> T_1$) มีค่ามากกว่า T_2 แล้ว จะปรับค่าพิกเซลดังกล่าวให้มีค่าเป็น “1” และถือเป็นสมาชิกหนึ่งในภาพขอบด้วยเช่นกัน

2.4 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)

วิธีการของซัพพอร์ตเวกเตอร์แมชชีน [3], [4], [9] จัดเป็นเทคนิคที่ใช้ในการแก้ปัญหาทางด้านการรู้จำรูปแบบข้อมูล โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกป้อนเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นไปยังเส้นแบ่งแยกแยะกลุ่มข้อมูลได้ดีที่สุด (Optimal Separating Hyperplane) สำหรับรากฐานเดิมของซัพพอร์ตเวกเตอร์แมชชีน ถูกนำมาใช้กับข้อมูลที่เป็นเชิงเส้น แต่ในความเป็นจริงแล้วข้อมูลที่นำมาใช้ในระบบการสอนให้ระบบเรียนรู้ส่วนใหญ่มักเป็นข้อมูลแบบไม่เป็นเชิงเส้น ซึ่งสามารถแก้ปัญหาดังกล่าวด้วยการนำเคอร์เนลฟังก์ชันมาใช้โดยจะกล่าวต่อไปในเรื่องของเคอร์เนลฟังก์ชัน

2.4.1 แนวความคิดของซัพพอร์ตเวกเตอร์แมชชีน

แนวความคิดของซัพพอร์ตเวกเตอร์แมชชีน [16] นั้น เกิดจากการที่นำค่าของกลุ่มข้อมูลมาวางลงในฟีเจอร์สเปซ (Feature Space) จากนั้นจึงหาเส้นที่ใช้แบ่งข้อมูลทั้งสองออกจากกันโดยจะสร้างเส้นแบ่ง (Plane) ที่เป็นเส้นตรงขึ้นมา และเพื่อให้ทราบว่าเส้นตรงที่แบ่งสองกลุ่มออกจากกันนั้น เส้นตรงใดเป็นเส้นที่ดีที่สุด [15] ดังภาพที่ 2-6



ภาพที่ 2-6 ตำแหน่งข้อมูลสองกลุ่มในฟีเจอร์สเปซ (Feature Space)

สามารถนำมาเขียนเป็นสมการเพื่อใช้ในการแก้ปัญหา โดยข้อมูลที่นำมาวางลงในพื้นที่คุณลักษณะนั้นเป็นกลุ่มข้อมูลที่อยู่ในรูปของเวกเตอร์จะได้

$$X = ((x_1, y_1), \dots, (x_i, y_i)) \quad (2-10)$$

เมื่อ X คือ ชุดค่าลักษณะเด่น

จากนั้นทำการกำหนดสมการขึ้นมา ซึ่งสมการดังกล่าวมีลักษณะเป็นสมการเส้นตรง เพื่อนำมาสร้างเป็นเส้นตรงบนไฮเปอร์เพลนซึ่งแบ่งกลุ่มข้อมูลที่มีลักษณะเชิงเส้นสองกลุ่มออกจากกัน โดยมีการกำหนดกลุ่มของข้อมูลทั้งสองฝั่งเป็นเพียงสองค่าซึ่งแทนด้วยค่า y ซึ่งข้อมูลที่เป็นตัวกำหนดความชันและระนาบที่เกิดขึ้นบนไฮเปอร์เพลนเกิดจากคู่ข้อมูล (w, b) และกำหนดสมการที่เป็นตัวบ่งบอกข้อมูลแต่ละกลุ่มว่าอยู่ส่วนไหนของเส้นแบ่งไฮเปอร์เพลนแสดงได้ดังสมการที่ (2-11) และสมการที่ (2-12)

เมื่อนำสมการเงื่อนไขทั้งหมดมาวิเคราะห์เชิงเรขาคณิต โดยพิจารณาในกรณีที่ข้อมูลถูกแบ่งกลุ่มได้สมบูรณ์ตามเงื่อนไขของซัพพอร์ตเวกเตอร์แมชชีนในสมการที่ (2-13) และข้อมูลการสอนให้ระบบเรียนรู้ต้องอยู่ในรูปแบบของเชิงเส้น สามารถแสดงลักษณะการวางตัวของกลุ่มข้อมูลได้ดังภาพที่ 2-5 เวกเตอร์ของข้อมูลที่ถูกป้อนเข้าสู่ระบบการสอน เพื่อให้ระบบเรียนรู้แทนด้วยสมการ และข้อมูลทั้งสองด้านแบ่งเป็นบวกและลบ สถานะของข้อมูลจึงแทนด้วย y ซึ่งมีสองค่า

คือ $y=1$ และ $y=-1$ นั้น แต่ทั้งนี้ก็ยังตัดสินไม่ได้ว่าเส้นแบ่งนั้นควรจะเป็นเส้นใดจึงจะดีที่สุด วิธีการที่ใช้ในการหาเส้นแบ่งที่ดีที่สุดคือการเพิ่มเส้นขอบให้กับเส้นแบ่งทั้งสองข้าง ทำให้ได้เส้นใหม่ที่จะถือเป็นเส้นขอบของข้อมูลแต่ละฝั่งอีกด้วย เส้นขอบ (Margin) ที่เป็นเส้นขอบของเส้นแบ่งนั้นจะเป็นเส้นที่สัมผัสกับค่าข้อมูลในฟีเจอร์สเปซ (Feature Space) ที่ใกล้ที่สุด เส้นขอบของทั้งสองเส้นที่เพิ่มขึ้นมานี้ถูกแทนด้วยสมการ $w^T x + b \geq y \geq 1$ ถ้าอยู่ด้าน $y=1$ และ $w^T x + b \leq y \leq -1$ ถ้า $y=-1$ หากเส้นขอบของเส้นแบ่งใดๆ ที่มีความกว้างมากที่สุด แสดงให้เห็นว่าข้อมูลสองชุดมีการแยกกันชัดเจนมากที่สุด ดังนั้นเส้นแบ่งที่มีเส้นขอบกว้างที่สุดจึงเป็นเส้นแบ่งที่ดีที่สุด และเรียกตำแหน่งการสัมผัสข้อมูลที่ใกล้ที่สุดจากการเพิ่มขอบนี้ว่า “เวกเตอร์สนับสนุน” (Support Vector) โดยเรียกเส้นประที่แบ่งข้อมูลทั้งสองเส้นว่าเส้นขอบ ซึ่งสามารถเขียนเป็นสมการการคำนวณความกว้างของเส้นขอบต้องทำการคำนวณพจน์ให้อยู่ในรูปปกติ (Normalization) โดยคำนวณจากสมการที่ (2-11) และ (2-12) เมื่อแทนค่าลงไปแล้ว

$$w^T x + b \geq y \quad \text{เมื่อกำหนดให้ } y = 1 \quad (2-11)$$

$$w^T x + b \leq y \quad \text{เมื่อกำหนดให้ } y = -1 \quad (2-12)$$

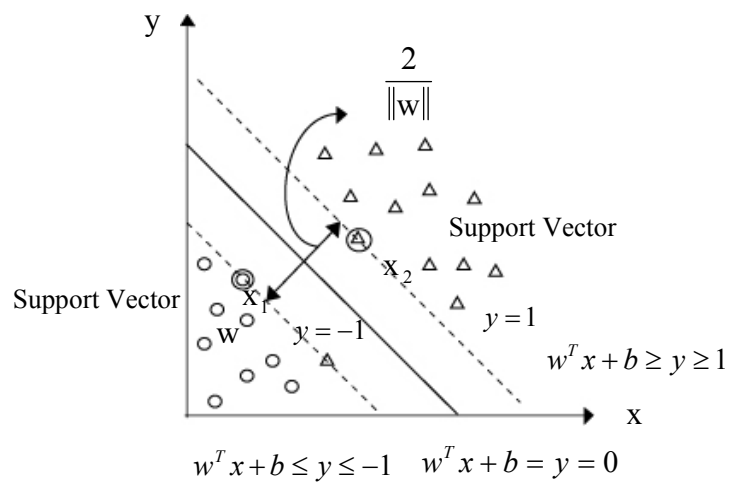
$$y(w^T x + b) - 1 \geq 0 \quad (2-13)$$

เมื่อ y คือ ค่ากลุ่มข้อมูล (1,-1)

w คือ ค่าความชัน

x คือ ค่าลักษณะเด่น

b คือ ค่าคงที่ (ค่าตัดแกน y)



ภาพที่ 2-7 การวางตัวของข้อมูลในลักษณะเชิงเส้น

$$\begin{aligned}
 w^T x^+ + b &= 1 \\
 w^T x^- + b &= -1 \\
 w^T (x^+ - x^-) &= 2 \\
 M &= \left(\frac{w}{\|w\|} \right)^T (x^+ - x^-) \\
 M &= \frac{2}{\|w\|}
 \end{aligned} \tag{2-14}$$

เมื่อ M คือ ความกว้างของเส้นขอบ

หลังจากที่ได้สมการที่ (2-13) และ (2-14) ของการหาเส้นแบ่งและค่าความกว้างตามลำดับแล้ว จึงทำการแก้สมการนั้นด้วย Lagrangian Dual Problem [15] ในสมการที่ (2-13)

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i [y_i ((w_i)^2 + b) - 1] \tag{2-15}$$

เมื่อ $\alpha_i \geq 0; \quad i = 1, 2, 3, \dots, N$

สมการที่ (2-15) ถูกนำมาหาอนุพันธ์ (Differential) เพื่อหาค่าน้ำหนักที่ดีที่สุด (w)

$$\frac{\partial L(w, b, \alpha)}{\partial w} = w - \sum_{i=1}^N y_i \alpha_i x_i = 0 \quad (2-16)$$

$$\frac{\partial L(w, b, \alpha)}{\partial b} = \sum_{i=1}^N y_i \alpha_i = 0 \quad (2-17)$$

หลังจากได้สมการที่ผ่านแก้ปัญหาคด้วย Dual Problem แล้วค่า w จะลดรูป และหาได้จากสมการที่ (2-18)

$$w = \sum_{i=1}^N \alpha_i y_i x_i \quad (2-18)$$

เมื่อ α คือ สัมประสิทธิ์คงที่

ค่า b นั้นไม่ปรากฏในการแก้ปัญหาคด้วย Dual Problem จึงหาได้จากสมการที่ (2-19)

$$b = -\frac{\max_{y_i=-1} (w^T x_i) + \min_{y_i=1} (w^T x_i)}{2} \quad (2-19)$$

เมื่อนำค่า w ที่แก้ปัญหาคไว้ ไปแทนในสมการที่ (2-20) ซึ่งเป็นสมการในการหาเส้นแบ่งจะได้

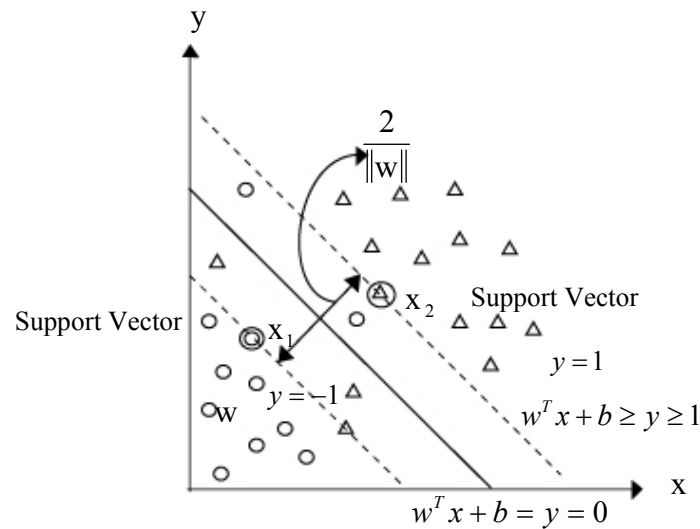
$$y_i \left(\sum_{i=1}^N \alpha_i y_i (x^T x_i) + b \right) - 1 \geq 0 \quad (2-20)$$

จากสมการที่ (2-20) คือการหาค่าสัมประสิทธิ์ α เพื่อนำมาใช้บอกการวางตัวของเส้นแบ่งข้อมูลต่อไป

2.4.2 ข้อผิดพลาดที่ไม่สามารถแบ่งกลุ่มข้อมูลได้

จากการหาซัพพอร์ตเวกเตอร์แมชชีนข้างต้น คิดในกรณีที่ข้อมูลไม่เกิดข้อผิดพลาด แต่ในความเป็นจริงผลลัพธ์ของการแบ่งข้อมูลด้วยซัพพอร์ตเวกเตอร์แมชชีน ย่อมมีข้อผิดพลาดเกิดขึ้น

จากค่าข้อมูลที่วางตัวในแกนซัพพอร์ตเวกเตอร์ การแก้ปัญหานี้คือการยอมให้มีการเกิดข้อผิดพลาดนี้ได้ในระยะหนึ่งที่ยอมรับได้ คำนี้นเรียกว่า “เวกเตอร์อนุโลม” (Slack Vector) ดังแสดงภาพที่ (2-6)



ภาพที่ 2-8 การแบ่งกลุ่มข้อมูลที่มีข้อผิดพลาดโดยจุดข้อมูลที่อยู่ในแกนซัพพอร์ตทั้งสองข้าง

เมื่อนำค่าเวกเตอร์อนุโลมเข้ามาในสมการที่ (2-15) ซึ่งเป็นขั้นตอนการแก้ปัญหา Dual Problem ก็เกิดตัวแปร C ขึ้นซึ่งก็เป็นค่าคงที่ค่าหนึ่ง ดังสมการที่ (2-21)

$$L(w, b, \xi, \alpha) = \frac{1}{2} \|w\|^2 + \frac{C}{2} \sum_{i=1}^N \xi_i^2 - \sum_{i=1}^N \alpha_i [y_i ((w_i \cdot w_i) + b) - 1 + \xi_i] \quad (2-21)$$

สมการที่ (2-21) ถูกนำมาหาอนุพันธ์ (Differential) เช่นเดียวกับขั้นตอนในวิธีการแก้สมการที่ (2-15)

$$\frac{\partial L(w, b, \xi, \alpha)}{\partial w} = w - \sum_{i=1}^N y_i \alpha_i x_i = 0 \quad (2-22)$$

$$\frac{\partial L(w, b, \xi, \alpha)}{\partial \xi} = C \xi - \alpha = 0 \quad (2-23)$$

$$\frac{\partial L(w, b, \xi, \alpha)}{\partial b} = \sum_{i=1}^N y_i \alpha_i = 0 \quad (2-24)$$

เมื่อ ξ คือ เวกเตอร์อนุโลม

หลังจากได้สมการที่ผ่านการแก้ปัญหาค่าด้วย Dual Problem แบบมีค่าเวกเตอร์อนุโลมแล้วค่า w ยังคงเป็นเช่นเดิมแต่จะได้ขอบเขตของ α ขึ้นมา

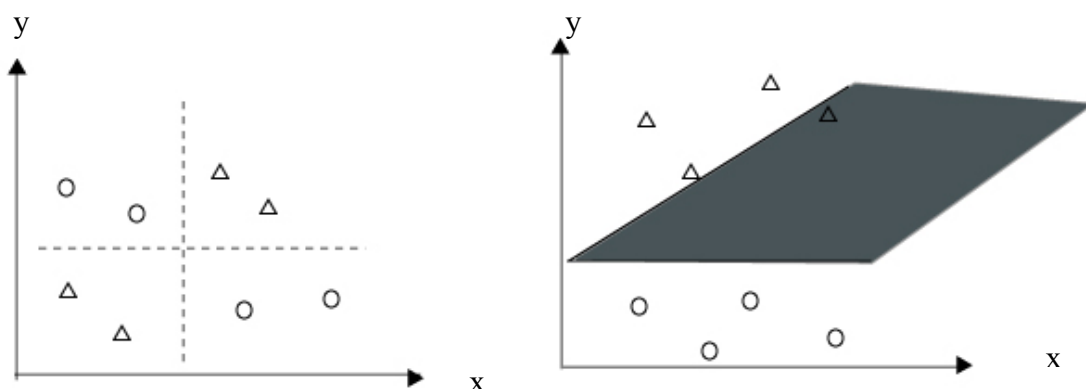
$$w = \sum_{i=1}^N \alpha_i y_i x_i \quad (2-25)$$

เมื่อ $C \geq \alpha \geq 0$

ค่า C จึงถือเป็นพารามิเตอร์อีกค่าหนึ่งที่จะต้องกำหนดขึ้นมาเพื่อใช้หาค่าสัมประสิทธิ์ α ซึ่งใช้สำหรับเพิ่มประสิทธิภาพของการแบ่งกลุ่มข้อมูล มีผลให้ระยะทางระหว่างเวกเตอร์มีความกว้างยิ่งขึ้นและยังทำหน้าที่ช่วยลดข้อผิดพลาดที่เกิดขึ้น

2.4.3 เคอร์เนล (Kernel)

เมื่อข้อมูลที่ใช้สำหรับการรู้จำมีการวางตัวกันในลักษณะตำแหน่งที่ไม่สามารถแบ่งกลุ่มได้โดยตรงตามวิธีการของซัพพอร์ตเวกเตอร์แมชชีนที่มีรากฐานมาจากข้อมูลเชิงเส้น จึงไม่สามารถนำมาใช้กับข้อมูลที่ไม่เป็นเชิงเส้นในลักษณะเช่นนี้ได้ โดยสามารถแก้ปัญหาลักษณะเช่นนี้ได้โดยการเปลี่ยนแปลงมิติของข้อมูลให้สูงขึ้นเพื่อช่วยในการเรียงตัวของข้อมูลเสียใหม่ที่เราเรียกว่า “พื้นที่มิติสูง” (Higher Dimensional Space)



ภาพที่ 2-9 การเปลี่ยนมิติข้อมูลไปยังมิติที่สูงขึ้นด้วยฟังก์ชันของเคอร์เนล

ในเคอร์เนลนั้นคือการคูณกันของชุดเวกเตอร์ของ x ใดๆ

$$K(x_i, x_j) = x_i^T x_j \quad (2-26)$$

เคอร์เนลที่นิยมใช้มีอยู่ 3 ชนิด [3] ด้วยกันคือ
โพลีโนเมียล (Polynomial) :

$$K(x_i, x_j) = (\lambda x_i^T x_j + r)^d, \gamma > 0 \quad (2-27)$$

เมื่อ d คือ ค่าของ Degree ในรูปเลขยกกำลัง

เรเดียลเบสฟังก์ชัน (Radial Basis Function-RBF) :

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (2-28)$$

เมื่อ γ คือ ค่าพารามิเตอร์

ซิกมอยด์ (Sigmoid) :

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r) \quad (2-29)$$

เมื่อ γ คือ ค่าพารามิเตอร์

ดังนั้นจากสมการของเคอร์เนลนั้นสามารถที่จะแทนลงไปในตำแหน่งของ $x_i^T x_j$ ในสมการที่ (2-20) ซึ่งสามารถจึงเขียนเป็นสมการได้ดังสมการที่ (2-30) ซึ่งเป็นสมการที่ใช้ในขั้นตอนที่จะเรียนรู้ว่าจะวางตำแหน่งเส้นแบ่งไว้ที่ตำแหน่งใดโดยทำงานร่วมกับเคอร์เนล เพื่อแปลงให้ข้อมูลที่ยากต่อการแบ่งแบบเชิงเส้นสามารถแบ่งได้เมื่อถูกทำให้เป็นข้อมูลแบบมิติสูง (Higher Dimension) ดังนั้นจึงมีอีกสมการหนึ่งที่ใช้ค่า w และ b เดิมมาจัดตำแหน่งของข้อมูลเพื่อที่ให้ทราบว่าข้อมูลนั้นเป็นกลุ่มใด กำหนดได้ดังสมการที่ (2-31)

$$y_i \left(\sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b \right) - 1 \geq 0 \quad (2-30)$$

$$f(x) = \text{sgn} \left(\sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b \right) \quad (2-31)$$

2.5 งานวิจัยที่เกี่ยวข้อง

ผลงานการวิจัยที่เกี่ยวข้องกับงานวิจัยนี้ได้แก่ งานวิจัยเกี่ยวกับการค้นหาทิศทางการวางตัวของภาพแบบอัตโนมัติของ Vailaya ในปี 1999 [2] ซึ่งใช้วิธีการหาค่าความสำคัญของสีของภาพแบบ LUV ฮิสโตแกรมของสี (Color histogram) ของภาพแบบ HSV ค่าฮิสโตแกรมของขอบวัตถุภาพ (Edge Direction Histogram) [7] และใช้คุณลักษณะเด่นของลายภาพแบบ MSAR ในกระบวนการแบ่งประเภทของภาพใช้ Bayesian Classifier ที่ซึ่งใช้ Learning Vector Quantization (LVQ) เป็นตัวแสดงลักษณะเด่นของกลุ่มข้อมูลแต่ละประเภทที่ได้จากการประเมินค่าความหนาแน่นของลักษณะเด่นที่สังเกตได้ ในการทดลอง Vailaya ได้นำเอาเทคนิคหลักการวิเคราะห์องค์ประกอบหลัก (PCA) และการวิเคราะห์ (LDA) มาทำการลดขนาดของข้อมูลด้วย ทำให้ผลการทดลองมีความถูกต้องดีขึ้น

ผลงานการวิจัยของ Wang และ Zhang ในปี 2001 [1] ซึ่งใช้วิธีการหาค่าความสำคัญของสีเช่นเดียวกันกับวิธีการของ Vailaya และได้เสนอเทคนิควิธีการหาค่าฮิสโตแกรมของขอบวัตถุภาพโดยการเพิ่ม กลุ่มทิศทาง พิเศษซึ่งได้ดัดแปลงมาจากวิธีการหาขอบภาพของ Vailaya [5] ให้กับค่าฮิสโตแกรมของขอบวัตถุภาพ Wang และ Zhang ใช้ Support Vector Machine ในการแบ่งประเภทของภาพ

ผลงานการวิจัยของ Zhang ในปี 2002 [11] ใช้วิธีการวิจัยโดยดึงลักษณะเด่นของภาพจากค่าความสำคัญของสี และค่าฮิสโตแกรมของขอบวัตถุภาพ และแบ่งกลุ่มข้อมูลด้วยซัพพอร์ตเวกเตอร์แมชชีน ซึ่งเป็นการพัฒนาต่อจาก Wang และ Zhang [2] โดยได้เพิ่มขั้นตอนแบ่งกลุ่มข้อมูลด้วยซัพพอร์ตเวกเตอร์แมชชีน โดยวิธีการ Boosting และได้ทดสอบกับภาพ 2 กลุ่ม คือ ภาพภายในอาคาร และภาพภายนอกอาคาร

ผลงานการวิจัยของ Jon Barker และ Zhou Su ในปี 2004 [8] ใช้วิธีดึงลักษณะเด่นด้วยหาค่าความสำคัญของสี และใช้วิธีการค่าฮิสโตแกรมของขอบวัตถุภาพ ใช้นิเวศน์เน็ตเวิร์กในการแบ่งกลุ่มข้อมูล

ผลงานการวิจัยต่อมาคือผลงานวิจัยของ V. Vapnik และคณะ [3] ซึ่งได้คิดค้นและพัฒนาวิธีการซัพพอร์ตเวกเตอร์แมชชีน โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกข้อมูล ซึ่งซัพพอร์ตเวกเตอร์แมชชีนนี้ จัดเป็นวิธีการที่ใหม่ มีจุดเด่นในการแยกแยะข้อมูลได้ดีกว่าวิธีการอื่น และในปัจจุบันมีผู้พัฒนาเครื่องมือสำหรับใช้งานซัพพอร์ตเวกเตอร์แมชชีนหลายรูปแบบด้วยกัน ช่วยให้งานวิจัยโดยใช้ซัพพอร์ตเวกเตอร์แมชชีนทำได้ความสะดวก และมีประสิทธิภาพมากยิ่งขึ้น

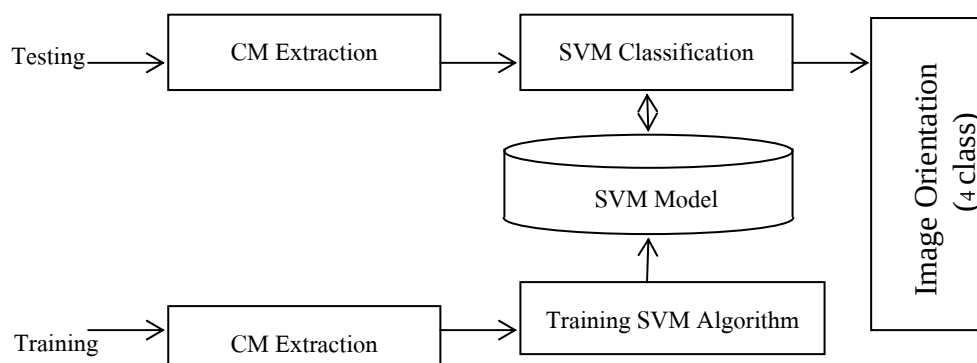
บทที่ 3

การดำเนินงานวิจัย

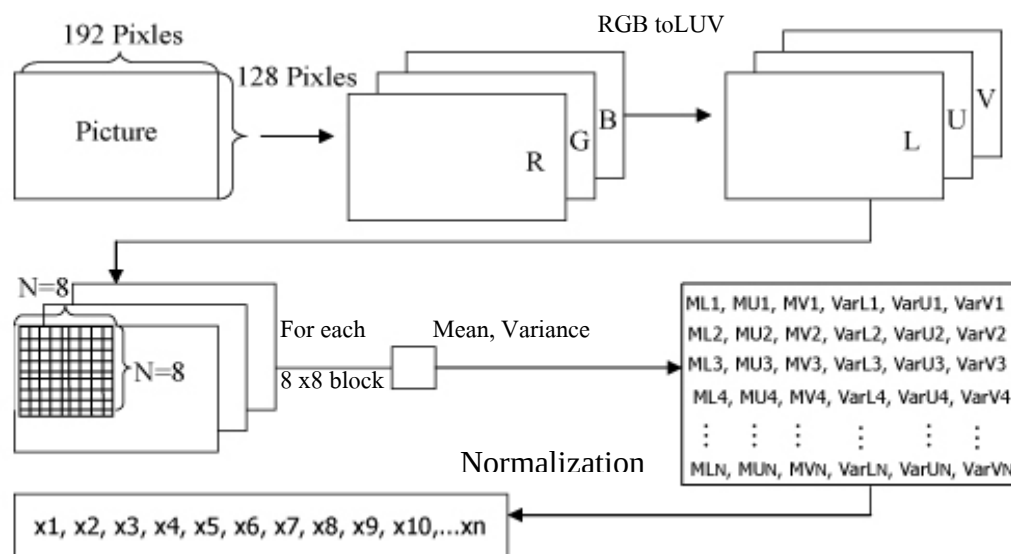
เนื้อหาของงานวิจัยในบทนี้จะกล่าวถึงโครงสร้างของระบบและขั้นตอนการวิจัยการตรวจจับและแก้ไขการวางตัวของภาพ โดยใช้ซอฟต์แวร์เวกเตอร์แมชชีน ซึ่งได้นำทฤษฎีและวิธีการที่กล่าวไว้ในบทที่ 2 มาใช้ในงานวิจัย ซึ่งมีขั้นตอนการดำเนินงานวิจัยดังนี้

3.1 ระบบการตรวจจับทิศทางการวางตัวของภาพโดยใช้ค่าความสำคัญของสี

ในขั้นตอนนี้จะเป็นการดึงลักษณะเด่นของภาพเพื่อแสดงลักษณะความคล้ายคลึงและความแตกต่างของข้อมูลสีในแต่ละพื้นที่ โดยเริ่มต้นจากการดึงลักษณะเด่นของภาพด้วยค่าความสำคัญของสี หลังจากนั้นเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ถึงความแตกต่างของฐานข้อมูลภาพทั้งหมด เพื่อให้ทราบถึงลักษณะเด่นของภาพในแต่ละประเภทเพื่อเก็บเป็นต้นแบบ (Model) ดังนั้นเมื่อมีข้อมูลภาพใหม่ๆ ที่ต้องการทดสอบเข้ามา ระบบจะทำการดึงลักษณะเด่นแล้วทำการแบ่งกลุ่มข้อมูลโดยเปรียบเทียบจากต้นแบบ ทำให้สามารถพิจารณาได้ว่าภาพที่เข้ามาเป็นลักษณะภาพประเภทใด โดยแบ่งเป็น 4 รูปแบบด้วยกันได้แก่ รูปแบบทิศทางการวางตัวของภาพ 0° , 90° , 180° และ 270° ซึ่งสามารถแสดงผังขั้นตอนวิธีการดำเนินการของระบบได้ดังภาพที่ 3-1



ภาพที่ 3-1 ผังขั้นตอนระบบตรวจจับทิศทางการวางตัวของภาพด้วยค่าความสำคัญของสี



ภาพที่ 3-2 ขั้นตอนวิธีการดึงลักษณะเด่นด้วยค่าความสำคัญของสี

จากภาพที่ 3-2 ในขั้นตอนแรกจะทำการปรับลดขนาดของภาพให้เป็น 192 x 128 พิกเซล สำหรับภาพในแนวนอน และขนาด 128 x 192 พิกเซล สำหรับภาพในแนวตั้ง แล้วทำการหาค่าความสำคัญของสี โดยทำการแปลงภาพจาก RGB เป็นภาพแบบ LUV เพื่อให้ได้ค่าความคงที่ของสี เมื่อทำการหมุนภาพที่ใช้ในการทดลอง จากการทดลอง [1], [2] เปรียบเทียบกับการแปลงค่าสีด้วยแบบจำลองสีอื่นๆ พบว่าวิธีการแปลงภาพจาก RGB เป็นภาพแบบ LUV ให้ค่าความถูกต้องดีที่สุด ดังนั้นในงานวิจัยนี้จึงได้ทำการทำทดลองโดยการกระทำบนภาพแบบ LUV เช่นเดียวกัน หลังจากนั้นทำการแบ่งภาพด้วยบล็อกย่อย ซึ่งก็คือ การแบ่งภาพออกเป็นภาพย่อยๆ (Subimage) ให้ได้ขนาด $N \times N$ ต่อบล็อก วัตถุประสงค์ของการแบ่งภาพนี้ก็เพื่อแบ่งอาณาเขตของบริเวณต่างๆ ในภาพ ซึ่งอาจจะสื่อความหมายที่แตกต่างกันออกไป เช่น บริเวณที่เป็นท้องฟ้า สนามหญ้า หรือพื้นน้ำ เป็นต้น ประโยชน์ของการแบ่งภาพด้วยบล็อกย่อยเพื่อช่วยในการเก็บค่าความสำคัญของสีได้อย่างละเอียดและสามารถเข้าใจข้อมูลที่แฝงอยู่ในภาพได้ดี ซึ่งขนาดของบล็อกที่ใช้ในการแบ่งภาพจะมีความสำคัญเป็นอย่างยิ่ง โดยทั่วไปความละเอียดของค่าที่ได้จะมีขนาดมากหรือน้อยขึ้นอยู่กับ การเลือกขนาดของ N สำหรับในงานวิจัยนี้ได้ งานวิจัยของ Zhang [11] ได้เสนอค่า N ที่เหมาะสมขนาด $N = 5 \sim 10$ และในงานวิจัยนี้ได้ทำการเลือก N ที่เหมาะสมโดยให้ $N = 8$

จากนั้นทำการหาค่าเฉลี่ย และค่าความแปรปรวน โดยกระทำในในแต่ละเซลล์แนลของภาพ ฉะนั้นผลรวมของค่าที่ได้ในแต่ละบล็อก จะได้เท่ากับ 6 ค่า (3 ค่าเฉลี่ย และ 3 ค่าความแปรปรวน) สมการสำหรับหาค่าเฉลี่ย และค่าความแปรปรวน หาได้จากสมการที่ (2-1) และ สมการที่ (2-2)

จากนั้นทำการนอมอนไลเซชัน เพื่อช่วยลดขนาดของค่าลักษณะเด่นที่ได้ และเพื่อปรับระดับของค่าลักษณะเด่นทั้งหมดให้อยู่ในระดับเดียวกัน ซึ่งสมการในการนอมอนไลเซชัน ซึ่งสามารถหาได้จากสมการที่ (3-1) หลังทำการนอมอนไลเซชันเรียบร้อยแล้ว จะได้ค่าข้อมูลในลักษณะเวกเตอร์ $1 \times I$ มิติ

$$I = \frac{im - \min(im_N)}{\max(im) - \min(im_N)} \quad (3-1)$$

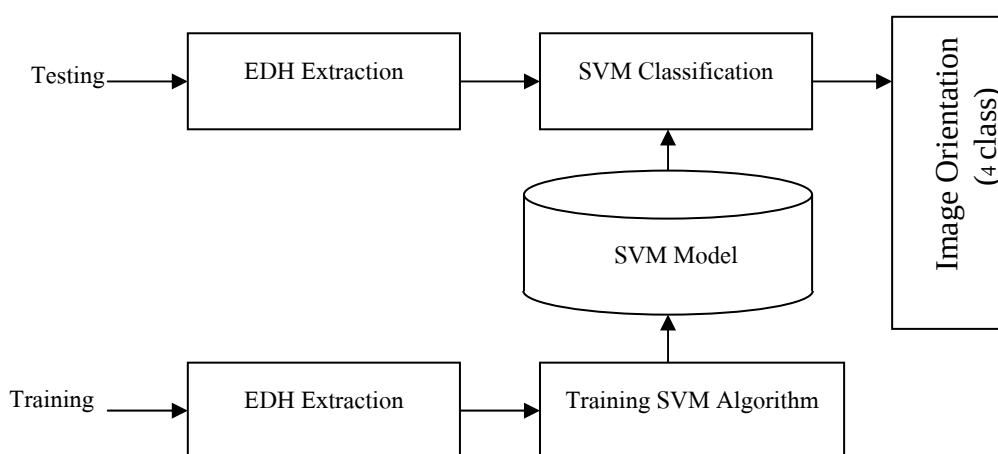
โดยที่ I คือ ผลลัพธ์ของข้อมูลหลังการแปลง

im คือ ข้อมูลเวกเตอร์ของภาพ

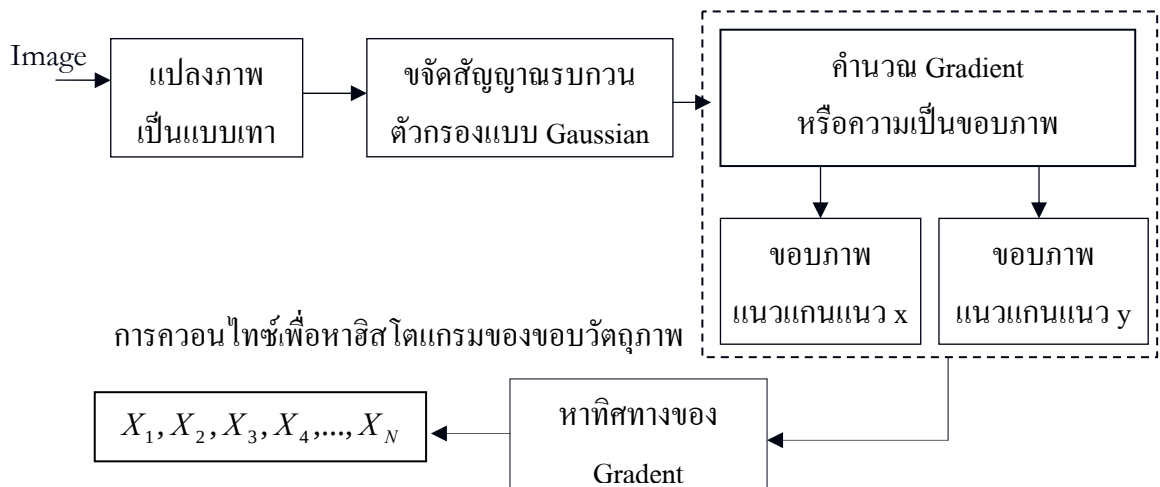
N คือ ผลรวมของพิกเซลทั้งหมดของภาพ

3.2 ระบบการตรวจจับทิศทางการวางตัวของภาพด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ

กระบวนการทำงานของการหาค่าฮิสโตแกรมของขอบวัตถุภาพนี้ เพื่อให้ได้ลักษณะของข้อมูลการกระจายตัวของจุดขอบในแต่ละมุม เพื่อนำมาแปลงเป็นฮิสโตแกรมสำหรับการเปรียบเทียบความแตกต่าง จากการหมุนภาพในทิศทางที่แตกต่างกัน โดยเริ่มจากการสอนให้ระบบเรียนรู้ถึงความแตกต่างของฐานข้อมูลภาพทั้งหมดสำหรับเก็บเป็นต้นแบบ (Model) เพื่อใช้ในการเปรียบเทียบกับข้อมูลภาพใหม่ๆ ที่ต้องการทดสอบ ทำให้ระบบสามารถพิจารณาได้ว่าภาพที่เข้ามาเป็นลักษณะภาพประเภทใดโดยแบ่งได้ 4 รูปแบบด้วยกันได้แก่ รูปแบบทิศทางการวางตัวของภาพ 0° , 90° , 180° และ 270° สามารถแสดงขั้นตอนวิธีการของระบบได้ดังภาพที่ 3-3



ภาพที่ 3-3 ฟังก์ชันตอนระบบตรวจจับทิศทางการวางตัวของภาพด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ



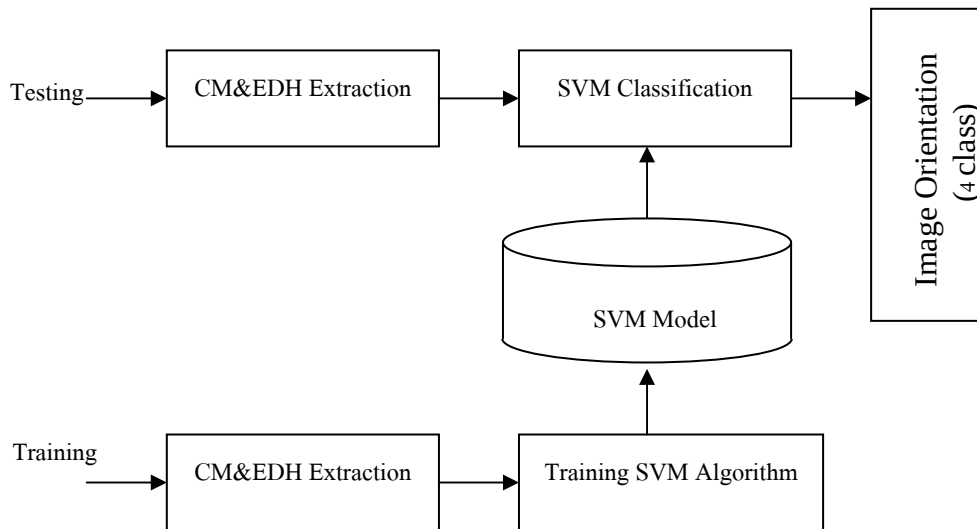
ภาพที่ 3-4 ขั้นตอนวิธีการดึงลักษณะเด่นด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ

จากภาพที่ 3-4 ในลำดับแรกจัดเตรียมข้อมูลภาพสำหรับการทดลอง ทำการปรับลดขนาดของภาพให้เป็น 192 x 128 พิกเซล สำหรับภาพในแนวนอน และขนาด 128 x 192 พิกเซล สำหรับภาพในแนวตั้งจัดสัญญาณรบกวนด้วยตัวกรองแบบ Gaussian เมื่อภาพนำเข้ามาจะทำการหาขอบของภาพ โดยต้องผ่านการจัดสัญญาณรบกวนก่อน ดังสมการที่ (2-5) หลังจากนั้นคำนวณ Gradient หรือความเป็นขอบภาพตามแนวแกน x (G_x) และแกน y (G_y) ซึ่งกระบวนการทำการ Gradient เป็นวิธีที่ใช้วิเคราะห์หาขอบภาพ โดยใช้วิธีการวัดความเป็นแปลงความเข้ม (Grey Level) ของจุดภาพที่กำลังพิจารณากับจุดภาพที่อยู่ข้างเคียง เพื่อการตัดสินใจต่อไปว่าเป็นขอบภาพหรือไม่ ทำการใส่ค่ามุม θ ให้แต่ละทิศทางของภาพ ดังสมการที่ (2-9) หลังจากนั้นทำการแปลงเป็นฮิสโตแกรมจากค่าของมุม θ ทำการควอนไทซ์เพื่อแบ่งเก็บค่ามุมที่เกิดขึ้นในแต่ละทิศทางภาพจากการแบ่งข้อมูลภาพออกเป็น $0^\circ - 360^\circ$ ซึ่งสามารถหาได้จากสมการที่ (2-4) หลังจากนั้นทำการแบ่งกลุ่มทิศทาง เพื่อบรรจุข้อมูล โดยในงานวิจัยนี้ได้ทำการเลือกค่า กลุ่มทิศทาง ที่เหมาะสมโดยให้กลุ่มทิศทาง เท่ากับ 37 ดังนั้นในขั้นตอนนี้ผลลัพธ์ของการหาค่าฮิสโตแกรมของวัตถุภาพจะให้ค่าลักษณะเด่นจำนวน 37 ค่า ในลักษณะ 1×37 มิติ

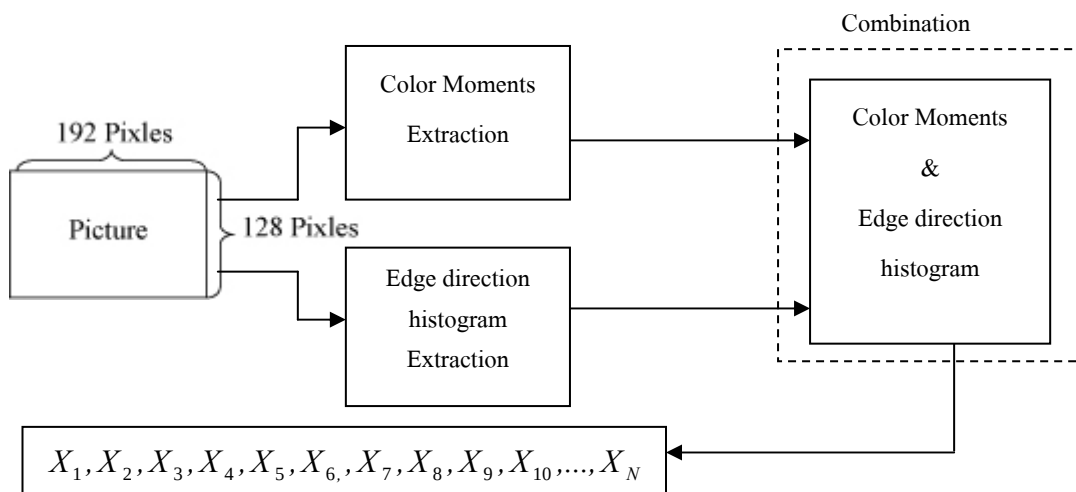
3.3 ระบบการตรวจจับทิศทางการวางตัวของภาพด้วยการรวม 2 ลักษณะเด่น

กระบวนการทำงานของการรวม 2 ลักษณะเด่นนี้ เกิดจากการรวมเอาลักษณะเด่นทั้ง 2 วิธีมารวม เข้าด้วยกัน คือ การหาค่าความสำคัญของสี และการหาค่าฮิสโตแกรมของขอบวัตถุภาพ เพื่อเก็บค่าลักษณะเด่นจากคุณสมบัติของสีและคุณสมบัติของรูปร่าง (เส้นขอบ) นำมาทำการสอนให้

ระบบเรียนรู้ เพื่อเก็บเป็นรูปแบบสำหรับภาพในแต่ละทิศทาง คือ 0° , 90° , 180° และ 270° เมื่อทำการอ่านไฟล์ภาพที่ต้องการทดสอบเข้ามาทำการดึงลักษณะเด่นแล้ว ทำให้ระบบสามารถเปรียบเทียบความแตกต่างของภาพในทิศทางที่แตกต่างกัน โดยแบ่งได้ 4 รูปแบบด้วยกันได้แก่ รูปแบบทิศทางการวางตัวของภาพ รูปแบบ 0° , 90° , 180° และ 270° สามารถแสดงขั้นตอนวิธีการของระบบ ได้ดังภาพที่ 3-5

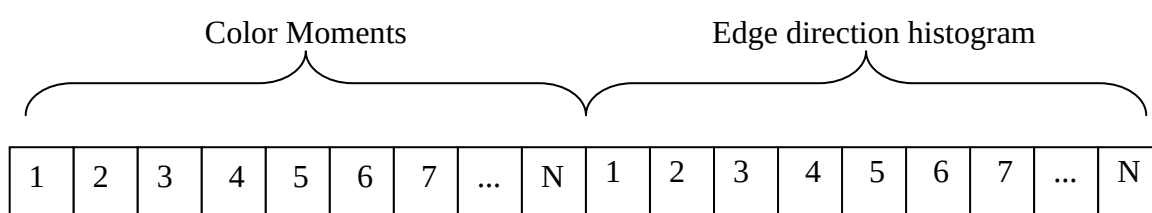


ภาพที่ 3-5 ผังขั้นตอนระบบตรวจจับทิศทางการวางตัวของภาพด้วยการรวม 2 ลักษณะเด่น



ภาพที่ 3-6 ขั้นตอนวิธีการดึงลักษณะเด่นด้วยการรวม 2 ลักษณะเด่น

จากภาพที่ 3-6 ในขั้นตอนแรกจัดเตรียมข้อมูลภาพสำหรับการทดลอง ทำการปรับลดขนาดของภาพให้เป็น 192 x 128 พิกเซล สำหรับภาพในแนวนอน และขนาด 128 x 192 พิกเซล สำหรับภาพในแนวตั้ง หลังจากนั้นทำการดึงลักษณะเด่นด้วยค่าความสำคัญของสี ดังรายละเอียดในขั้นตอนวิธีการดึงลักษณะเด่นด้วยค่าความสำคัญของสีในหัวข้อ 3.1 และทำการดึงลักษณะเด่นด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ดังรายละเอียดขั้นตอนวิธีการดึงลักษณะเด่นด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ในหัวข้อ 3.2 มาทำการคำนวณเพื่อหาค่าลักษณะเด่นทั้งสอง หลังจากนั้นนำค่าลักษณะเด่นทั้งสองที่ทำการรวมลักษณะเด่น มาจัดวางในรูปแบบ เวกเตอร์แบบ $1 \times I$ มิติ ดังแสดงในภาพที่ 3-7



ภาพที่ 3-7 การจัดวางเวกเตอร์ข้อมูลโดยการรวม 2 ลักษณะเด่น

3.4 วิธีการประเมินคุณภาพ

ระบบการตรวจจับและแก้ไขการวางตัวของภาพโดยใช้ซอฟต์แวร์เวกเตอร์แมชชีนจำเป็นต้องประเมินประสิทธิภาพของระบบ เพื่อใช้เป็นมาตรฐานชี้วัดคุณภาพของการสอนให้ระบบเรียนรู้ ซึ่งในงานวิจัยนี้ระบุใช้วิธีการประเมินประสิทธิภาพ ดังนี้

3.4.1 การคำนวณหาค่าความถูกต้อง (Accuracy Score) เป็นการเปรียบเทียบระหว่างข้อมูลลักษณะเด่นของภาพที่ต้องการทดสอบและข้อมูลลักษณะเด่นของภาพในต้นแบบ (Model) ซึ่งจะพิจารณาเฉพาะข้อมูลที่ตรงตามเกณฑ์ของจำนวนความถูกต้องของข้อมูลทั้งหมด ดังสมการที่ (3-2)

$$Asc = \frac{SF}{TF} \times 100 \quad (3-2)$$

เมื่อ Asc คือ ค่าความถูกต้อง

TF คือ จำนวนข้อมูลภาพทั้งหมดที่ใช้ในการทดลอง

SF คือ จำนวนข้อมูลภาพทดสอบที่ตรวจจับได้ถูกต้อง

3.4.2 การคำนวณน้ำหนักของภาพสำหรับการสอนให้ระบบเรียนรู้และการทดสอบ

3.4.2.1 รูปแบบที่ 1 เป็นการคำนวณน้ำหนักของภาพในกลุ่มการสอนให้ระบบเรียนรู้เพื่อสร้างเป็นต้นแบบหรือโมเดลสำหรับนำมาใช้ในการเปรียบเทียบกับค่าน้ำหนักของลักษณะเด่นของภาพที่นำมาทดสอบในการแบ่งกลุ่มข้อมูล โดยใช้วิธีการจัดรูปแบบข้อมูลภาพประเภทละ 100 ภาพ ทำการจัดรูปแบบโดยจัดให้เป็นกลุ่มภาพสำหรับการสอนให้ระบบเรียนรู้ 50 ภาพ และกลุ่มภาพสำหรับการทดสอบ 50 ภาพ

3.4.2.2 รูปแบบที่ 2 เป็นการคำนวณน้ำหนักของภาพในกลุ่มการสอนให้ระบบเรียนรู้เพื่อสร้างเป็นต้นแบบหรือโมเดลสำหรับนำมาใช้ในการเปรียบเทียบกับค่าน้ำหนักของลักษณะเด่นของภาพที่นำมาทดสอบในการแบ่งกลุ่มข้อมูล โดยใช้วิธีการจัดรูปแบบข้อมูลภาพประเภทละ 100 ภาพ ทำการจัดรูปแบบโดยจัดให้เป็นกลุ่มภาพสำหรับการสอนให้ระบบเรียนรู้ 40 ภาพ และกลุ่มภาพสำหรับการทดสอบ 60 ภาพ

3.4.2.3 รูปแบบที่ 3 เป็นการคำนวณน้ำหนักของภาพในกลุ่มการสอนให้ระบบเรียนรู้เพื่อสร้างเป็นต้นแบบหรือโมเดลสำหรับนำมาใช้ในการเปรียบเทียบกับค่าน้ำหนักของลักษณะเด่นของภาพที่นำมาทดสอบในการแบ่งกลุ่มข้อมูล โดยใช้วิธีการจัดรูปแบบข้อมูลภาพประเภทละ 100 ภาพ ทำการจัดรูปแบบโดยจัดให้เป็นกลุ่มภาพสำหรับการสอนให้ระบบเรียนรู้ 30 ภาพ และกลุ่มภาพสำหรับการทดสอบ 70 ภาพ

3.4.2.4 รูปแบบที่ 4 เป็นการคำนวณน้ำหนักของภาพในกลุ่มการสอนให้ระบบเรียนรู้เพื่อสร้างเป็นต้นแบบหรือโมเดลสำหรับนำมาใช้ในการเปรียบเทียบกับค่าน้ำหนักของลักษณะเด่นของภาพที่นำมาทดสอบในการแบ่งกลุ่มข้อมูล โดยใช้วิธีการจัดรูปแบบข้อมูลภาพประเภทละ 100 ภาพ ทำการจัดรูปแบบโดยจัดให้เป็นกลุ่มภาพสำหรับการสอนให้ระบบเรียนรู้ 20 ภาพ และกลุ่มภาพสำหรับการทดสอบ 80 ภาพ

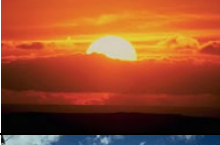
3.4.3 การวัดประสิทธิภาพกับงานวิจัยที่เกี่ยวข้อง

จากผลงานการวิจัยของ Wang และ Zhang [1] ซึ่งได้ทำการวิจัยโดยวิธีแรกทำการดึงลักษณะเด่นด้วยค่าความสำคัญของสี โดยให้ขนาดบล็อกย่อย เท่ากับ 8×8 ซึ่งได้ใช้พื้นที่ข้อมูลภาพ 3 ใน 4 ส่วน โดยเว้นข้อมูลส่วนตรงกลางของภาพไว้ วิธีที่สองทำการดึงลักษณะเด่นด้วยค่าฮิสโตแกรมของขอบวัตถุภาพด้วยการให้ค่ากลุ่มทิศทาง เท่ากับ 37 และให้ค่าบล็อกเท่ากับ 5 ซึ่งจะเท่ากับ $25 \times 37 = 925$ ค่า และวิธีที่สามการรวมลักษณะเด่นด้วยค่าความสำคัญของสีและค่าฮิสโตแกรมของขอบวัตถุภาพ นำมาทำการเปรียบเทียบกับงานวิจัยการตรวจจับและแก้ไขการวางตัวของภาพ โดยใช้ซอฟต์แวร์เคอร์แมมซินนี้ เพื่อเปรียบเทียบด้านค่าความถูกต้องในการตรวจจับ, ความเร็วที่ใช้ในการประมวลผล และด้านขนาดของค่าลักษณะเด่น

3.5 การเตรียมข้อมูลที่ใช้ในงานวิจัย

ภาพที่ใช้ในงานวิจัยนี้แบ่งเป็น 3 กลุ่มด้วยกัน คือ กลุ่มภาพภายในอาคาร กลุ่มภาพนอกอาคารและกลุ่มภาพถ่ายหน้าตรง ขนาดภาพที่ใช้ในการทดลองขนาด 192 x 128 พิกเซล ภาพส่วนใหญ่ได้มาจากฐานข้อมูลงานวิจัยเกี่ยวกับการค้นคืนภาพของ Wang [10] และข้อมูลจากการถ่ายภาพสี่ด้วยกล้องดิจิทัล ซึ่งภาพที่นำมาใช้ในการทดลองอาจมีลักษณะการวางตัวในลักษณะแนวตั้ง และลักษณะแนวนอน ทำการแบ่งข้อมูลตามประเภทของภาพ เลือกข้อมูลที่จะใช้ทำการทดสอบและตั้งชื่อภาพสำหรับการวิจัยครั้งนี้ เพื่อให้ได้ลักษณะเด่นที่ครอบคลุมกับการทดลอง ดังแสดงในตารางที่ 3-1

ตารางที่ 3-1 ข้อมูลที่ใช้ในงานวิจัย

ลำดับที่	กลุ่มข้อมูลภาพ	ประเภทของภาพ	ภาพตัวอย่าง	จำนวนทั้งหมด
1	ภาพภายในอาคาร	ห้องเรียน		100
2	ภาพนอกอาคาร	ทะเล		100
		พื้นที่หญ้า		100
		พระอาทิตย์ตก		100
		ภูเขาน้ำแข็ง		100
		ทะเลทราย		100

ตารางที่ 3-1 (ต่อ)

ลำดับที่	กลุ่มข้อมูลภาพ	ประเภทของภาพ	ภาพตัวอย่าง	จำนวนทั้งหมด
3	ภาพถ่ายหน้าตรง	ใบหน้ามนุษย์		100

3.5.1 ขั้นเตรียมข้อมูลสำหรับการสอนให้ระบบเรียนรู้

การเตรียมข้อมูล สำหรับการสอนการให้ระบบเรียนรู้ถือเป็นจุดเน้นของงานวิจัยนี้ โดยขั้นตอนการสอนให้ระบบเรียนรู้ได้เน้นที่ตัวอย่างข้อมูลที่นำมาใช้จากข้อมูลไฟล์ภาพประเภทละ 100 ภาพ ทำการแบ่งเป็นชุดข้อมูลสำหรับการสอนให้ระบบเรียนรู้ 50 ภาพ ทำการหมุนภาพ ครั้งละ 90° เพื่อใช้เป็นชุดข้อมูลสำหรับการสอนให้ระบบเรียนรู้ ดังนั้นจะได้รูปแบบการเตรียมข้อมูลภาพ สำหรับกลุ่มข้อมูลในการสอนให้ระบบเรียนรู้ 4 รูปแบบด้วยกัน ดังแสดงในภาพที่ 3-5



(ก)



(ข)



(ค)



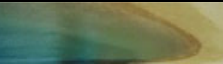
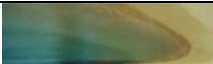
(ง)

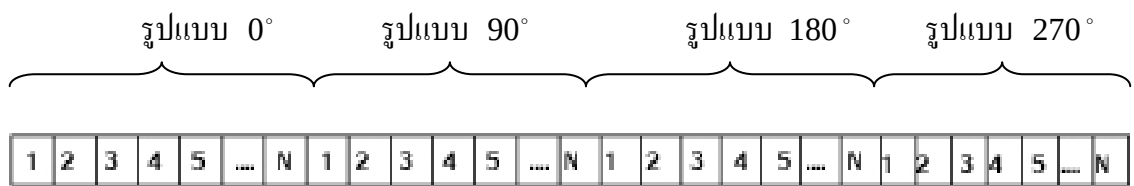
ภาพที่ 3-8 ตัวอย่างรูปภาพทั้ง 4 รูปแบบ (ก) รูปแบบ 0° (ข) รูปแบบ 90° (ค) รูปแบบ 180° และ (ง) รูปแบบ 270°

3.5.1.1 ขั้นตอนการเตรียมลักษณะเด่นของภาพสำหรับการสอนให้ระบบเรียนรู้ด้วยค่าความสำคัญของสี ซึ่งสามารถหาได้จากหัวข้อ 3.1 นำข้อมูลภาพมาทำการหมุนภาพ ครั้งละ 90° คำนวณหาค่าความสำคัญของสี โดยเก็บค่าเพียงจำนวน 1 ใน 4 ส่วนของวัตถุภาพโดยตัดเก็บเฉพาะส่วนบนสุดของภาพเพื่อเปรียบเทียบหารูปแบบการจัดวางที่แตกต่างกัน ซึ่งจากลักษณะของภาพ โดยทั่วไปหากใช้วิธีการสังเกตตามข้อมูลที่เป็นจริงของภาพ พบว่าวัตถุภาพที่เกิดขึ้นโดยทั่วไป เมื่อ

ทำการหมุนเปลี่ยนภาพในทิศทางตรงกันข้ามกัน ลักษณะของสีที่เกิดบนวัตถุภาพจะมีความแตกต่างกันอย่างชัดเจน ตัวอย่างเช่น กลุ่มภาพที่สามารถสังเกตได้เด่นชัดในงานวิจัยนี้ ได้แก่ กลุ่มภาพแบบมีท้องฟ้า ซึ่งหากภาพมีลักษณะทิศทางการวางตัวที่ถูกต้อง ส่วนบนสุดของภาพจะเป็นท้องฟ้า และส่วนล่างสุดจะเป็นพื้นดิน ดังนั้นหากพิจารณาแล้วเลือกเก็บภาพจำนวน 1 ใน 4 ส่วนของภาพแล้วทำการจัดเรียงข้อมูลดังแสดงในตารางที่ 3-2 จะเห็นได้ว่าภาพที่จัดวางในลักษณะที่ถูกต้อง (รูปแบบ 0°) ประกอบด้วยข้อมูลภาพ 4 ส่วนด้วยกัน ข้อมูลของภาพที่สามารถสังเกตได้เด่นชัดก็คือ ส่วน (ก) ซึ่งเป็นข้อมูลภาพในลำดับแรกซึ่งเกือบทั้งหมดจะเป็นท้องฟ้า ส่วน (ข) ซึ่งเป็นข้อมูลในลำดับที่ 2 จะแสดงข้อมูลที่เด่นชัดในลักษณะข้อมูลพื้นดิน-ท้องฟ้า ข้อมูลในลำดับที่ 3 ส่วน (ค) จะเป็นข้อมูลที่เด่นชัดในส่วนของพื้นดิน และข้อมูลในลำดับที่ 4 ส่วน (ง) จะแสดงข้อมูลที่เด่นชัดในลักษณะข้อมูลท้องฟ้า-พื้นดิน เพราะฉะนั้นเมื่อทำลักษณะเช่นเดียวกันนี้ โดยการหมุนครั้งละ 90° จำนวน 4 ครั้ง ให้กับภาพ จะสามารถเก็บค่าลักษณะข้อมูลที่มีลักษณะที่แตกต่างกัน 4 รูปแบบ ได้แก่ รูปแบบ $0^\circ, 90^\circ, 180^\circ$ และ 270° ดังนั้นในขั้นตอนสุดท้ายของการหาค่าความสำคัญของสีเมื่อทำการทำอนโมไนเซชันเรียบร้อยแล้ว ทำการเก็บข้อมูลลักษณะเด่นเพียง 1 ใน 4 ส่วน รวมลักษณะเด่นที่ได้ดังลักษณะภาพที่แสดงในตารางที่ 3-2 จากภาพขนาด 192×128 พิกเซล เวกเตอร์จะมีขนาด 1×576 ฉะนั้นเมื่อรวมลักษณะเด่นทั้ง 4 ขนาดของเวกเตอร์ทั้งหมดจะเท่ากับ $1 \times 2,304$ ดังแสดงในภาพที่ 3-7 เป็นข้อมูลเวกเตอร์ของภาพสำหรับการใช้ในการสอนการให้ระบบเรียนรู้ด้วยซัพพอร์ตเวกเตอร์แมชชีน ในขั้นตอนต่อไป

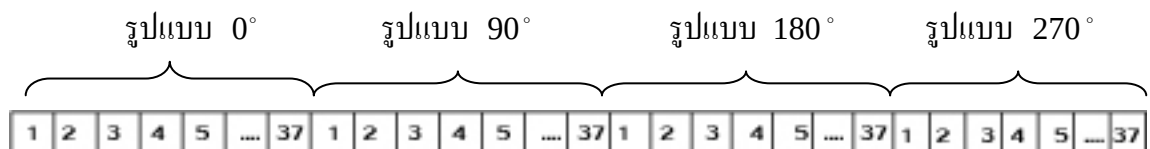
ตารางที่ 3-2 ลักษณะการเก็บข้อมูลภาพด้วยค่าความสำคัญของสี

รูปแบบ	ลักษณะรูปแบบการเก็บข้อมูลภาพ			
	(ก)	(ข)	(ค)	(ง)
0°				
90°				
180°				
270°				



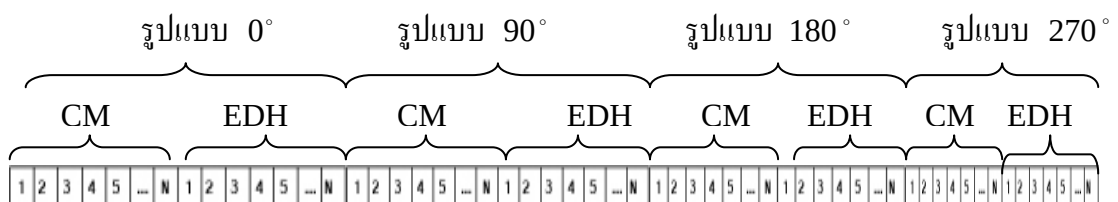
ภาพที่ 3-9 ลักษณะการจัดเก็บเวกเตอร์ข้อมูลด้วยค่าความสำคัญของสี

3.5.1.2 ขั้นตอนการเตรียมลักษณะเด่นของภาพสำหรับการสอนให้ระบบเรียนรู้ด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ สามารถหาได้จากหัวข้อ 3.2 โดยทำการหมุนภาพครั้งละ 90° เช่นเดียวกัน แต่จะต่างกันโดยการหาค่าฮิสโตแกรมจะทำการเก็บค่าข้อมูลจากพื้นที่ทั้งหมดของภาพ ดังนั้นเมื่อให้จำนวน กลุ่มทิศทาง เท่ากับ 37 ค่า ทำการหมุนเพื่อเปลี่ยนทิศทาง 4 ครั้ง ดังนั้นข้อมูลสำหรับการสอนให้ระบบเรียนรู้เวกเตอร์จะมีขนาดเท่ากับ 1×148 สามารถแสดงได้ดังภาพที่ 3-8



ภาพที่ 3-10 ลักษณะการจัดเก็บเวกเตอร์ข้อมูลด้วยฮิสโตแกรมของขอบวัตถุภาพ

3.5.1.3 ขั้นตอนการเตรียมข้อมูลการรวม 2 ลักษณะเด่นด้วยค่าความสำคัญของสีและค่าฮิสโตแกรมของขอบวัตถุภาพ จากลักษณะเด่นที่ได้มาทั้ง 2 รูปแบบ จากหัวข้อ 3.3 ทำการรวมลักษณะเด่นที่ได้ ในการหมุนครั้งละ 90° โดยทำการเก็บค่าความสำคัญของสีก่อน และตามด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ดังนั้นเมื่อทำการรวมลักษณะเด่นที่ได้ เวกเตอร์จะมีขนาด $1 \times 2,452$ ดังภาพที่ 3-9



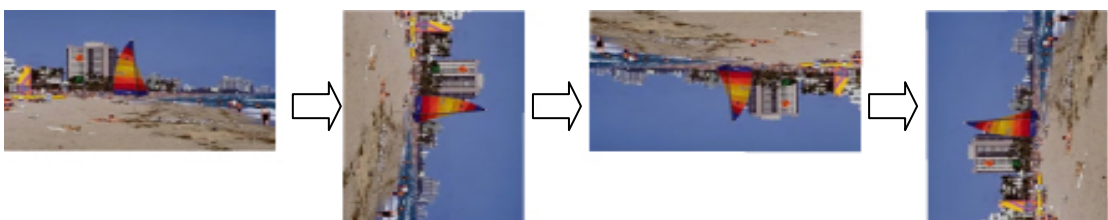
ภาพที่ 3-11 ลักษณะการจัดเก็บเวกเตอร์ข้อมูลด้วยการรวม 2 ลักษณะเด่น

3.6 กระบวนการทำงานของซัพพอร์ทเวกเตอร์แมชชีน

3.6.1 ขั้นตอนการสอนให้ระบบเรียนรู้ด้วยซัพพอร์ทเวกเตอร์แมชชีน จากขั้นตอนการดึงลักษณะเด่นของภาพเข้าสู่ขั้นตอนการสอนให้ระบบเรียนรู้ โดยในขั้นตอนนี้ได้นำเอาค่าเวกเตอร์ของข้อมูลที่ได้จากการดึงลักษณะเด่นจากหัวข้อที่ 3.5.1.1 , 3.5.1.2 และ 3.5.1.3 นำมาทำการสอนให้ระบบเรียนรู้ ซึ่งข้อมูลที่ได้นี้ก็คือนำค่า x_i จากนั้นทำการกำหนดค่า y_i ให้กับข้อมูลซึ่งจะมีทั้งหมด 4 ค่าด้วยกันคือ 1, 2, 3 และ 4 ซึ่งเป็นการกำหนดประเภทของข้อมูลลักษณะเด่นให้กับการสอนให้ระบบเรียนรู้ ซึ่งการแบ่งกลุ่มข้อมูลลักษณะเด่นนี้เป็นการแบ่งกลุ่มข้อมูลแบบหลายกลุ่ม (Multi - Class Support Vector Classifier) แบบหนึ่งรายการ- เปรียบเทียบกับ- หนึ่งรายการ (one-against-one) [3] ดังนั้นจากข้อมูลชุดการสอนให้ระบบเรียนรู้จำนวน 50 รายการ ค่าลักษณะเด่นที่ได้จากกลุ่มภาพ 4 กลุ่ม จึงเท่ากับ 200 รายการ แบ่งเป็นกลุ่มภาพแบบ 0° จำนวน 50 รายการ กำหนดค่า y_i เท่ากับ 1 กลุ่มภาพแบบ 90° จำนวน 50 รายการ กำหนด y_i เท่ากับ 2 กลุ่มภาพแบบ 180° จำนวน 50 รายการ กำหนดค่า y_i เท่ากับ 3 และ กลุ่มภาพแบบ 270° จำนวน 50 รายการ โดยกำหนดค่า y_i เท่ากับ 4 สมการที่ใช้สำหรับการสอนให้ระบบเรียนรู้ ก็คือสมการทางคณิตศาสตร์ในสมการที่ (2-30) ซึ่งเป็นสมการที่ใช้ในการสอนให้ระบบเรียนรู้ บางครั้งค่าลักษณะเด่นที่ได้นี้ไม่สามารถแบ่งแยกได้ด้วยสมการแบบซัพพอร์ทเวกเตอร์แมชชีนแบบเชิงเส้น จึงจะต้องใช้ เคอร์เนลเข้ามาช่วย ในการทำให้ข้อมูลสองกลุ่มวางตัวจัดเรียงใหม่ เคอร์เนลที่ใช้มี 4 ชนิดด้วยกันได้แก่ เชิงเส้น โพลีโนเมียล เรเดียลเบสิสฟังก์ชัน และซิกมอยด์

หลังจากที่ผ่านกระบวนการนี้แล้ว ผลลัพธ์ที่ได้ออกมาคือชุดต้นแบบ (Model) เพื่อใช้สำหรับการแบ่งกลุ่มของข้อมูลในขั้นทดสอบต่อไป

3.6.2 ขั้นตอนการทดสอบด้วยซัพพอร์ทเวกเตอร์แมชชีน โดยในขั้นตอนนี้โปรแกรมจะทำการอ่านไฟล์ภาพที่ต้องการทดสอบเข้ามาทีละภาพ ซึ่งระบบการทำงานของโปรแกรมจะทำการหมุนเปลี่ยนทิศทางให้กับภาพจำนวน 4 ครั้ง ครั้งละ 90° ในทิศทางทวนเข็มนาฬิกา ทำการเก็บค่าลักษณะเด่นที่ได้ ดังภาพที่ 3-10



ภาพที่ 3-12 ทิศทางการหมุนภาพครั้งละ 90°

หลังจากนั้นทำการดึงลักษณะเด่นของภาพสำหรับการทดสอบ เมื่อโปรแกรมทำการหมุนภาพและเก็บค่าลักษณะเด่นในแต่ละทิศทางพร้อมทั้งรวมลักษณะเด่นเรียบร้อยแล้ว เข้าสู่กระบวนการทดสอบโดยระบบจะทำการเปรียบเทียบกับชุดต้นแบบ จะแสดงผลที่ได้ โดยโปรแกรมสามารถที่จะรายงานได้เป็นใน 4 รูปแบบด้วยกัน คือ หากภาพที่ใช้ในการทดสอบมีการวางตัวในรูปแบบ 0° จะแสดงผลเป็น 1 หากภาพที่ใช้ในการทดสอบมีการวางตัวในรูปแบบ 90° จะแสดงผลเป็น 2 หากภาพที่ใช้ในการทดสอบมีการวางตัวในรูปแบบ 180° จะแสดงผลเป็น 3 และหากภาพที่ใช้ในการทดสอบมีการวางตัวรูปแบบ 270° จะแสดงผลเป็น 4 เมื่อซอฟต์แวร์คอมพิวเตอร์แมชชีนตรวจจับทิศทางเรียบร้อยแล้ว จะทำการหมุนเพื่อเปลี่ยนทิศทางการวางตัวของภาพเพื่อแก้ไขทิศทางการวางตัวของภาพให้ถูกต้องตามทิศทางที่เป็นจริง ซึ่งสมการที่ใช้ในการแบ่งกลุ่มข้อมูลด้วยซอฟต์แวร์คอมพิวเตอร์แมชชีนนี้ ก็คือสมการที่ (2-31)

บทที่ 4

ผลการวิจัย

ในบทนี้กล่าวถึงวิธีการทดลอง และผลการทดลองที่ได้จากการการตรวจจับและการแก้ไขการวางตัวของภาพโดยใช้ซอฟต์แวร์คอมพิวเตอร์แมชชีน โดยใช้ลักษณะเด่นที่กล่าวไว้ในบทที่ 3

4.1 ดำเนินการทดลอง

ภาพที่นำมาใช้ในการทดลองนี้ข้อมูลภาพที่ใช้โดยส่วนใหญ่เป็นข้อมูลภาพที่ได้จากฐานข้อมูลภาพจากอินเทอร์เน็ต เพื่อให้เป็นไปตามขอบเขตของการวิจัยจึงได้ทำการปรับลดขนาดของภาพให้เป็น 192 x 128 พิกเซล สำหรับภาพในแนวนอน และขนาด 128 x 192 พิกเซล สำหรับภาพในแนวตั้ง ภาพที่ใช้ในฐานข้อมูลทั้งหมดเป็นภาพแบบ 2 มิติมีจำนวน 3 กลุ่มด้วยกันคือ กลุ่มภาพภายในอาคาร กลุ่มภาพนอกอาคาร และกลุ่มภาพถ่ายหน้าตรง และได้ทำการแบ่งออกตามประเภทของภาพออกเป็น 7 ประเภทได้แก่ ห้องเรียน ทะเล พื้นหญ้า พระอาทิตย์ตก ภูเขาน้ำแข็ง ทะเลทราย และภาพถ่ายใบหน้ามนุษย์ โดยแต่ละประเภทมีจำนวน 100 ภาพ

หลังจากทำการเตรียมภาพแล้ว ขั้นตอนต่อมาเป็นขั้นตอนการทดลองซึ่งทำการประมวลผลด้วยเครื่องคอมพิวเตอร์รุ่น Intel Celeron M มีความเร็วในการประมวลผล 1.6 GHz หน่วยความจำ 256 MB บนระบบปฏิบัติการ Microsoft Windows XP และใช้โปรแกรม MATLAB เวอร์ชัน 6.5 ในการคำนวณเพื่อการประมวลผลสำหรับการตรวจจับและแก้ไขภาพดิจิทัล

4.2 ดำเนินการทดลอง และผลการทดลอง

ในบทนี้ได้จัดรูปแบบการทดลองมีขั้นตอนการหาผลของการวิจัยที่สำคัญ ประกอบด้วย 3 ขั้นตอนดังนี้

4.2.1 ผลการทดลองการคำนวณน้ำหนักของภาพสำหรับการสอนให้ระบบเรียนรู้และการทดสอบ

การทดลองการคำนวณน้ำหนักของภาพสำหรับการสอนให้ระบบเรียนรู้นี้ เพื่อใช้สำหรับคำนวณหาน้ำหนักที่เหมาะสมสำหรับการจัดกลุ่มข้อมูลให้กับการสอนให้ระบบเรียนรู้และการจัดกลุ่มข้อมูลสำหรับการทดสอบ โดยได้แบ่งวิธีการทดลองออกเป็น 4 รูปแบบด้วยกัน ตามหัวข้อ 3.4.2 ซึ่งจะได้รูปแบบสำหรับการคำนวณน้ำหนักภาพ ได้แก่ รูปแบบที่ 1 (50:50) รูปแบบที่

2 (40:60) รูปแบบที่ 3 (30:70) และรูปแบบที่ 4 (20:80) นำมาทำการทดลองถึงลักษณะเด่นของภาพด้วยค่าความสำคัญของสี ค่าฮิสโตแกรมของขอบวัตถุภาพ และการรวม 2 ลักษณะเด่นเปรียบเทียบกับเคอร์เนลที่ใช้ 4 วิธี คือ เชิงเส้น โพลีโนเมียล เรเดียลเบสิสฟังก์ชัน และซิกมอยด์ เพื่อหาค่าเคอร์เนลที่เหมาะสมและดีที่สุด โดยกำหนดพารามิเตอร์ $C = 1000$ ทำการทดสอบกับภาพกลุ่มตัวอย่าง จำนวน 100 ภาพ และใช้วิธีการคำนวณหาค่าความถูกต้องดังสมการที่ 3-2

4.2.1.1 ผลการทดลองการคำนวณน้ำหนักของภาพสำหรับการสอนให้ระบบเรียนรู้ และการทดสอบด้วยค่าความสำคัญของสี

ตารางที่ 4-1 ตารางผลการทดลองด้วยค่าความสำคัญของสีในรูปแบบที่ 1

รูปแบบการทดลอง		เคอร์เนลที่เหมาะสม	ค่าความถูกต้อง
กลุ่มการสอนให้ระบบเรียนรู้	ทดสอบ		
50	50	Linear	90%
		Polynomia	70%
		RBF	70 %
		Sigmoid	70 %

ตารางที่ 4-2 ตารางผลการทดลองด้วยค่าความสำคัญของสีในรูปแบบที่ 2

รูปแบบการทดลอง		เคอร์เนลที่เหมาะสม	ค่าความถูกต้อง
กลุ่มการสอนให้ระบบเรียนรู้	ทดสอบ		
40	60	Linear	83%
		Polynomia	68%
		RBF	68%
		Sigmoid	68%

ตารางที่ 4-3 ตารางผลการทดลองด้วยค่าความสำคัญของสีในรูปแบบที่ 3

รูปแบบการทดลอง		เคอร์เนล ที่เหมาะสม	ค่าความถูกต้อง
กลุ่มการสอนให้ ระบบเรียนรู้	ทดสอบ		
30	70	Linear	70%
30	70	Polynomia	65%
		RBF	65%
		Sigmoid	64%

ตารางที่ 4-4 ตารางผลการทดลองด้วยค่าความสำคัญของสี ในรูปแบบที่ 4

รูปแบบการทดลอง		เคอร์เนล ที่เหมาะสม	ค่าความถูกต้อง
กลุ่มการสอนให้ ระบบเรียนรู้	ชุดทดสอบ		
20	80	Linear	82%
		Polynomia	62%
		RBF	61%
		Sigmoid	60%

จากผลการทดลองหารูปแบบที่เหมาะสมทั้ง 4 รูปแบบ ได้แก่ รูปแบบที่ 1 (50:50) รูปแบบที่ 2 (40:60) รูปแบบที่ 3 (30:70) และรูปแบบที่ 4 (20:80) เพื่อกำหนดน้ำหนักสำหรับการจัดรูปแบบให้กับกลุ่มภาพสำหรับการสอนให้ระบบเรียนรู้ และการทดสอบ สามารถสรุปได้ว่า ผลการทดลองด้วยค่าความสำคัญของสีในรูปแบบที่ 1 (50:50) มีประสิทธิภาพดีที่สุดและเหมาะสมที่สุด จากผลการทดลองที่ให้ค่าความถูกต้องสูงสุดร้อยละ 90 เคอร์เนลที่เหมาะสมสำหรับการทดลองด้วยค่าความสำคัญของสี ก็คือ เคอร์เนลแบบเชิงเส้น

4.2.1.2 ผลการทดลองการค่าน้ำหนักของภาพสำหรับการสอนให้ระบบเรียนรู้ และการทดสอบด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ

ตารางที่ 4-5 ตารางผลการทดลองด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ในรูปแบบที่ 1

รูปแบบการทดลอง		เคอร์เนล ที่เหมาะสม	ค่าความถูกต้อง
กลุ่มการสอนให้ ระบบเรียนรู้	ทดสอบ		
50	50	Linear	84 %
		Polynomia	90 %
		RBF	92 %
		Sigmoid	90 %

ตารางที่ 4-6 ตารางผลการทดลองด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ในรูปแบบที่ 2

รูปแบบการทดลอง		เคอร์เนล ที่เหมาะสม	ค่าความถูกต้อง
กลุ่มการสอนให้ ระบบเรียนรู้	ทดสอบ		
40	60	Linear	76 %
		Polynomia	78 %
		RBF	83 %
		Sigmoid	83 %

ตารางที่ 4-7 ตารางผลการทดลองด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ในรูปแบบที่ 3

รูปแบบการทดลอง		เคอร์เนล ที่เหมาะสม	ค่าความถูกต้อง
กลุ่มการสอนให้ ระบบเรียนรู้	ทดสอบ		
30	70	Linear	71 %
		Polynomia	77 %
		RBF	78 %
		Sigmoid	77 %

ตารางที่ 4-8 ตารางผลการทดลองด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ในรูปแบบที่ 4

รูปแบบการทดลอง		เคอร์เนล ที่เหมาะสม	ค่าความถูกต้อง
กลุ่มการสอนให้ ระบบเรียนรู้	ชุดทดสอบ		
20	80	Linear	72 %
		Polynomia	81 %
		RBF	81 %
		Sigmoid	81 %

จากผลการทดลองหารูปแบบที่เหมาะสม ทั้ง 4 รูปแบบ ได้แก่ รูปแบบที่ 1 (50:50) รูปแบบที่ 2 (40:60) รูปแบบที่ 3 (30:70) และรูปแบบที่ 4 (20:80) เพื่อกำหนดน้ำหนักสำหรับการจัดรูปแบบให้กับกลุ่มภาพสำหรับการสอนให้ระบบเรียนรู้ และการทดสอบ สามารถสรุปได้ว่าด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ในรูปแบบที่ 1 (50:50) มีประสิทธิภาพดีที่สุด จากค่าความถูกต้องร้อยละ 92 และเคอร์เนลที่เหมาะสมสำหรับการทดลองด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ก็คือ เคอร์เนลแบบเรเดียลเบสิสฟังก์ชัน

4.2.1.3 ผลการทดลองการค่าน้ำหนักของภาพสำหรับการสอนให้ระบบเรียนรู้ และการทดสอบด้วยการรวม 2 ลักษณะเด่น

ตารางที่ 4-9 ตารางผลการทดลองด้วยการรวม 2 ลักษณะเด่น ในรูปแบบที่ 1

รูปแบบการทดลอง		เคอร์เนล ที่เหมาะสม	ค่าความถูกต้อง
กลุ่มการสอนให้ ระบบเรียนรู้	ทดสอบ		
50	50	Linear	84 %
		Polynomia	90 %
		RBF	92 %
		Sigmoid	88 %

ตารางที่ 4-10 ตารางผลการทดลองด้วยการรวม 2 ลักษณะเด่น ในรูปแบบที่ 2

รูปแบบการทดลอง		เคอร์เนล ที่เหมาะสม	ค่าความถูกต้อง
กลุ่มการสอนให้ ระบบเรียนรู้	ทดสอบ		
40	60	Linear	72 %
		Polynomia	74 %
		RBF	61 %
		Sigmoid	71 %

ตารางที่ 4-11 ตารางผลการทดลองด้วยการรวม 2 ลักษณะเด่น ในรูปแบบที่ 3

รูปแบบการทดลอง		เคอร์เนล ที่เหมาะสม	ค่าความถูกต้อง
กลุ่มการสอนให้ ระบบเรียนรู้	ทดสอบ		
30	70	Linear	75 %
		Polynomia	80 %
		RBF	72 %
		Sigmoid	80 %

ตารางที่ 4-12 ตารางผลการทดลองด้วยการรวม 2 ลักษณะเด่น ในรูปแบบที่ 4

รูปแบบการทดลอง		เคอร์เนล ที่เหมาะสม	ค่าความถูกต้อง
กลุ่มการสอนให้ ระบบเรียนรู้	ชุดทดสอบ		
20	80	Linear	81 %
		Polynomia	68 %
		RBF	71 %
		Sigmoid	70 %

จากผลการทดลองหารูปแบบที่เหมาะสม ทั้ง 4 รูปแบบ ได้แก่ รูปแบบที่ 1 (50:50) รูปแบบที่ 2 (40:60) รูปแบบที่ 3 (30:70) และรูปแบบที่ 4 (20:80) เพื่อกำหนดหานักสำหรับ

การจัดรูปแบบให้กับกลุ่มภาพสำหรับการสอนให้ระบบเรียนรู้ และการทดสอบ สามารถสรุปได้ด้วยการรวม 2 ลักษณะเด่นในรูปแบบที่ 1 (50:50) มีประสิทธิภาพดีที่สุด จากค่าความถูกต้องร้อยละ 92 และเคอร์เนลที่เหมาะสมสำหรับการทดลองด้วยการรวม 2 ลักษณะเด่น ก็คือ เคอร์เนลแบบเรเดียลเบสิสฟังก์ชัน

4.2.2 ผลการทดลองเพื่อประเมินหาค่าความถูกต้องสูงสุดให้กับระบบ จากหัวข้อ 4.2.1 ในช่วงต้นซึ่งได้ดำเนินการทดลองเพื่อคำนวณน้ำหนักของภาพสำหรับการสอนให้ระบบเรียนรู้และการทดสอบ จากค่าความสำคัญของสี ค่าฮิสโตแกรมของขอบวัตถุภาพ และการรวม 2 ลักษณะเด่น จึงทำให้สามารถจัดน้ำหนักของภาพสำหรับการสอนให้ระบบเรียนรู้และการทดสอบให้กับข้อมูลภาพที่ดีที่สุดในแต่ละระบบได้ ดังนั้นในขั้นตอนนี้จึงได้นำผลการทดลองที่ได้ดังกล่าวมาจัดน้ำหนักของภาพสำหรับการใช้ในการทดสอบกับข้อมูลภาพในฐานข้อมูลทั้งหมดที่จัดเตรียมไว้ในหัวข้อที่ 3-5 ทำการทดสอบหาค่าความถูกต้องสูงสุดของภาพในแต่ละประเภท สามารถแสดงผลการทดลองได้ดังต่อไปนี้

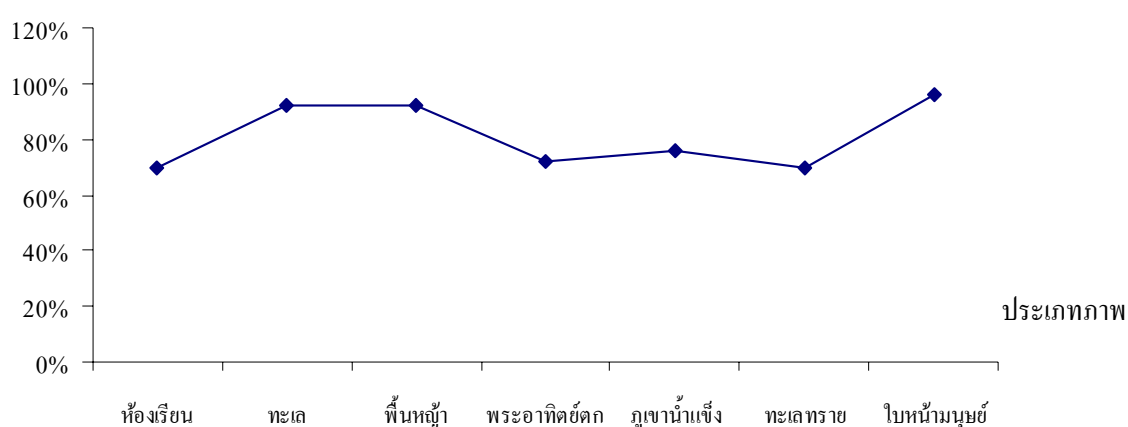
4.2.2.1 ผลการทดลองประเมินหาค่าความถูกต้องของระบบด้วยค่าความสำคัญของสี จากผลการทดลองเพื่อการคำนวณน้ำหนักของภาพสำหรับการสอนให้ระบบเรียนรู้และการทดสอบด้วยค่าความสำคัญของสี ของระบบจากหัวข้อ 4.2.1.1 ทำให้สามารถหารูปแบบที่เหมาะสมสำหรับการทดลองได้ รูปแบบที่เหมาะสมที่สุด ก็คือรูปแบบที่ 1 (50:50) และเคอร์เนลที่เหมาะสมที่สุด ก็คือ เคอร์เนลแบบเชิงเส้น ดังนั้นจึงนำมาทำการทดลองเพื่อประเมินหาค่าความถูกต้องของระบบด้วยค่าความสำคัญของสีนี้ สามารถแสดงผลการทดลองทั้งหมดจากข้อมูลได้ดังตารางที่ 4-13

ตารางที่ 4-13 ตารางผลการทดลองของระบบด้วยค่าความสำคัญของสี

ลำดับที่	กลุ่มข้อมูลภาพ	ประเภทภาพ	ค่าความถูกต้อง
1	ภาพภายในอาคาร	ห้องเรียน	70 %
2	ภาพภายนอกอาคาร	ทะเล	92 %
		พื้นหญ้า	92 %
		พระอาทิตย์ตก	72 %
		ภูเขาน้ำแข็ง	76 %
		ทะเลทราย	70 %
3	ภาพถ่ายหน้าตรง	ใบหน้านมนุษย์	96%

จากผลการทดลองที่ได้ ผลจากการทดลองของระบบด้วยค่าความสำคัญของสี พบว่าประเภทภาพที่ระบบสามารถตรวจจับได้ดีที่สุดคือ ประเภทภาพถ่ายใบหน้ามนุษย์สามารถให้ค่าความถูกต้องสูงสุดร้อยละ 96 และประเภทภาพที่สามารถจับได้ต่ำได้แก่ ประเภทภาพชั้นเรียน ประเภทภาพทะเลทราย และประเภทภาพทะเลทราย ซึ่งผลการดิ่งลักษณะเด่นด้วยค่าความสำคัญของสีของระบบ ให้ค่าความถูกต้องเฉลี่ยอยู่ที่ร้อยละ 81

ค่าความถูกต้อง (%)



ภาพที่ 4-1 กราฟแสดงค่าความถูกต้องจากการทดลองของระบบด้วยค่าความสำคัญของสี

4.2.2.2 ผลการทดลองหาค่าความถูกต้องของระบบด้วยค่าฮิสโตแกรมของขอบวัตถุ ภาพ จากผลการทดลองเพื่อการคำนวณน้ำหนักของภาพสำหรับการสอนให้ระบบเรียนรู้ ในตารางผลการทดลองด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ ในหัวข้อ 4.2.1.2 ทำให้ได้ผลการทดลองที่เหมาะสมที่สุด คือ รูปแบบที่ 1 (50:50) ด้วยเคอร์เนลแบบเรเดียลเบสิสฟังก์ชัน นำมาใช้ทำการทดลองหาความถูกต้องของระบบในกระบวนการดิ่งลักษณะเด่นของภาพด้วยค่าฮิสโตแกรมของขอบวัตถุภาพสามารถแสดงผลการทดลองได้ดังตารางที่ 4-14

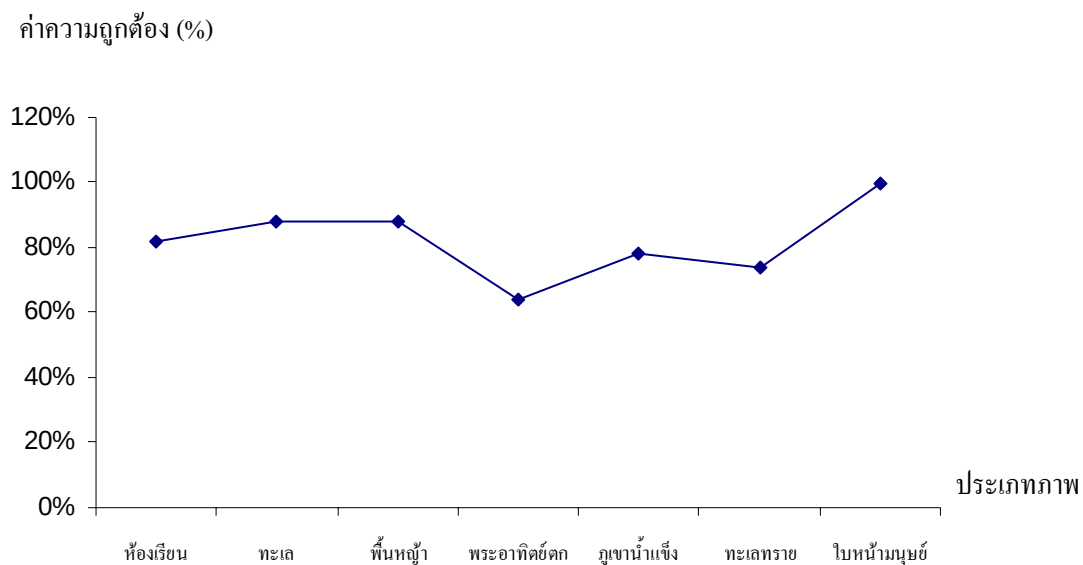
ตารางที่ 4-14 ตารางผลการทดลองของระบบด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ

ลำดับที่	กลุ่มข้อมูลภาพ	ประเภทภาพ	ค่าความถูกต้อง
1	ภาพภายในอาคาร	ห้องเรียน	82 %
2	ภาพภายนอกอาคาร	ทะเล	88 %
		พื้นหญ้า	88 %

ตารางที่ 4-14 (ต่อ)

ลำดับที่	กลุ่มข้อมูลภาพ	ประเภทภาพ	ค่าความถูกต้อง
2	ภาพภายนอกอาคาร	พระอาทิตย์ตก	65 %
		ภูเขา น้ำแข็ง	78 %
		ทะเลทราย	74 %
3	ภาพถ่ายหน้าตรง	ใบหน้านามนุษย์	100 %

จากผลการทดลองของระบบที่ได้จากค่าฮิสโตแกรมของขอบวัตถุภาพ พบว่าประเภทภาพที่ระบบสามารถตรวจจับได้ดีที่สุดคือ ประเภทภาพใบหน้านามนุษย์สามารถให้ค่าความถูกต้องสูงสุดร้อยละ 100 และประเภทภาพที่สามารถจับได้ต่ำสุดได้แก่ ประเภทภาพพระอาทิตย์ตก ซึ่งจับได้ร้อยละ 65 ซึ่งผลการดิ่งลักษณะเด่นของระบบด้วยค่าความฮิสโตแกรมของขอบวัตถุภาพ ให้ค่าความถูกต้องเฉลี่ยอยู่ที่ร้อยละ 82 และสามารถเขียนกราฟแสดงผลการทดลองได้ดังภาพที่ 4-2



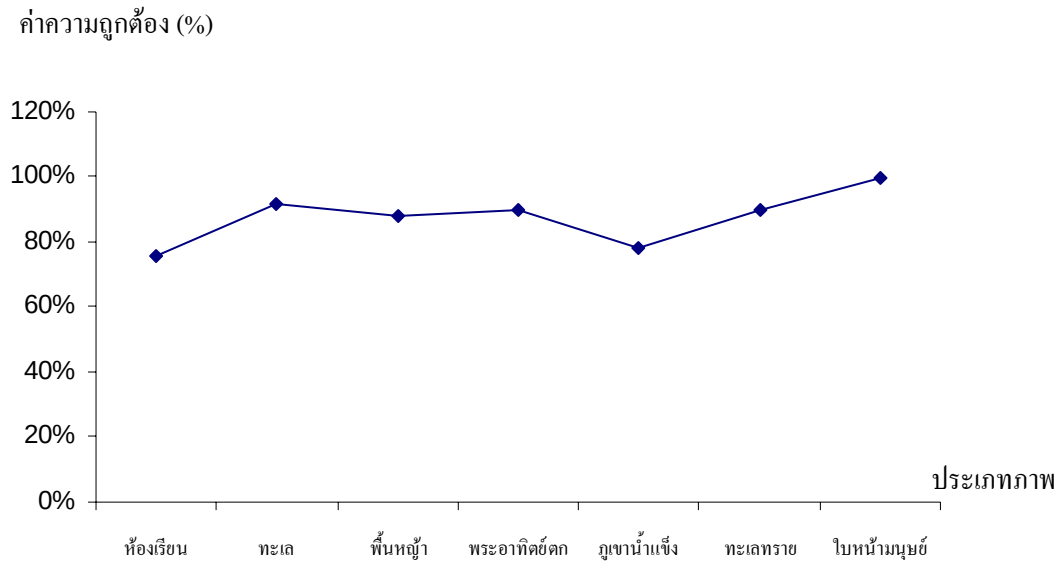
ภาพที่ 4-2 กราฟแสดงค่าความถูกต้องจากผลการทดลองของระบบด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ

4.2.2.3 ผลการทดลองหาค่าความถูกต้องของระบบด้วยการรวม 2 ลักษณะเด่น จากผลการทดลองเพื่อการคำนวณน้ำหนักของภาพสำหรับการทดสอบให้ระบบเรียนรู้และการทดสอบด้วยการรวม 2 ลักษณะเด่น และการหารูปแบบเคอร์เนลที่เหมาะสม ของระบบจากหัวข้อ 4.2.1.3 ทำให้ทราบว่า รูปแบบที่ 1 (50:50) เป็นวิธีการที่ดีที่สุด ด้วยเคอร์เนลแบบเรเดียลเบสิสฟังก์ชันนำมาทำการทดลองหาค่าความถูกต้องของระบบสามารถแสดงผลการทดลองได้ดังตารางที่ 4-15

ตารางที่ 4-15 ตารางผลการทดลองของระบบด้วยการรวม 2 ลักษณะเด่น

ลำดับที่	กลุ่มข้อมูลภาพ	ประเภทภาพ	ค่าความถูกต้อง
1	ภาพภายในอาคาร	ห้องเรียน	76%
2	ภาพภายนอกอาคาร	ทะเล	92%
		พื้นที่หญ้า	90%
		พระอาทิตย์ตก	90%
		ภูเขาน้ำแข็ง	78%
		ทะเลทราย	90%
3	ภาพถ่ายหน้าตรง	ใบหน้านมนุษย์	100%

จากผลการทดลองของระบบในการตรวจจับและการแก้ไขการวางตัวของภาพโดยใช้ซอฟต์แวร์เมชชีนนี้ พบว่าผลจากการรวม 2 ลักษณะเด่น สามารถให้ค่าความถูกต้องสูงกับทุกประเภทภาพ โดยประเภทภาพห้องเรียนให้ค่าความถูกต้องร้อยละ 76 ทะเลให้ค่าความถูกต้องร้อยละ 92, พื้นที่หญ้าให้ค่าความถูกต้องร้อยละ 90, พระอาทิตย์ตกให้ค่าความถูกต้องร้อยละ 90, ภูเขา น้ำแข็งให้ค่าความถูกต้องร้อยละ 78, ทะเลทรายให้ค่าความถูกต้องร้อยละ 90 และใบหน้านมนุษย์ให้ค่าความถูกต้องร้อยละ 100 ซึ่งผลการทดลองหาค่าความถูกต้องของระบบจากการรวม 2 ลักษณะเด่นให้ค่าความถูกต้องเฉลี่ยอยู่ที่ร้อยละ 88



ภาพที่ 4-3 กราฟแสดงค่าความถูกต้องจากผลการทดลองของระบบด้วยการรวม 2 ลักษณะเด่น

4.2.3 การทดลองเปรียบเทียบกับงานวิจัยที่เกี่ยวข้อง จากผลงานการวิจัยของ Wang และ Zhang [1] ได้ใช้วิธีการดึงลักษณะเด่นของภาพด้วยค่าความสำคัญของสี ค่าฮิสโตแกรมของขอบวัตถุภาพ และได้ทำการรวมลักษณะเด่นด้วยค่าความสำคัญของสีและค่าฮิสโตแกรมของขอบวัตถุภาพ (รวม 2 ลักษณะเด่น) ทำการเปรียบเทียบกับงานวิจัยหัวข้อการตรวจจับและแก้ไขทิศทางภาพด้วยซัพพอร์ตเวกเตอร์แมชชีน (Image Orientation Detection and Correction Using Support Vector Machine :ICSVM) ซึ่งนำมาพัฒนาต่อจากลักษณะเด่นทั้ง 3 รูปแบบ โดยแก้ปัญหาในส่วนข้อด้อยของวิธีการของ Wang และ Zhang [1] ซึ่งจากวิธีการเดิมของ Wang และ Zhang [1] ได้ใช้วิธีการหาค่าความสำคัญของสีโดยทำการเก็บค่าจากจำนวน 3 ใน 4 ส่วนของภาพ โดยไม่พิจารณาส่วนข้อมูลพื้นที่ตรงกลางของภาพและใช้วิธีการหาค่าฮิสโตแกรมของขอบวัตถุภาพโดยการใช้บล็อกขนาด $5 \times 5 = 25$ บล็อก สำหรับเก็บค่าฮิสโตแกรมของขอบวัตถุภาพ ดังนั้นผู้วิจัยจึงได้นำมาทำการทดลองเปรียบเทียบในวิธีการที่ได้พัฒนาขึ้นตามขั้นตอนวิธีการที่กล่าวไว้ในบทที่ 3 โดยทำการเปรียบเทียบข้อมูลทางด้านความถูกต้องในการตรวจจับ ด้านเวลาที่ใช้ในการประมวลผล และด้านขนาดของข้อมูลลักษณะเด่น โดยเลือกฐานข้อมูลภาพตัวอย่างขึ้นมา 100 ภาพ แบ่งเป็นกลุ่มการสอนให้ระบบเรียนรู้ 50 ภาพ และกลุ่มทดสอบ 50 ภาพ ทำการปรับลดขนาดของข้อมูลภาพเป็น 64×64 พิกเซล สำหรับแนวนอนและแนวตั้ง ทำการทดลองเปรียบเทียบด้วยค่าลักษณะเด่นทั้ง 3 รูปแบบ คือ ค่าความสำคัญของสี ค่าฮิสโตแกรมของขอบวัตถุภาพ และการรวม 2 ลักษณะเด่น ทำการทดลองด้วยเคอร์เนลแบบเชิงเส้น ค่าพารามิเตอร์ $C=1000$ สามารถแสดงผลการทดลองได้ดังรายละเอียด

ตารางที่ 4-16 ผลการทดลองเปรียบเทียบด้านค่าความถูกต้องในการตรวจจับ

ผลงานการวิจัย	ค่าความถูกต้อง		
	ค่าความสำคัญของสี	ค่าฮิสโตแกรม ของขอบวัตถุภาพ	รวม 2 ลักษณะเด่น
งานวิจัยที่เกี่ยวข้อง	92 %	92 %	96 %
ICSVM	96 %	100 %	100 %

ตารางที่ 4-17 ผลการทดลองเปรียบเทียบเกี่ยวกับเวลาที่ใช้ในการประมวลผล

ผลงานการวิจัย	ความเร็วในการประมวลผล		
	ค่าความสำคัญของสี	ค่าฮิสโตแกรม ของขอบวัตถุภาพ	รวม 2 ลักษณะเด่น
งานวิจัยที่เกี่ยวข้อง	0.5320 วินาที	0.6250 วินาที	0.7970 วินาที
ICSVM	0.7340 วินาที	0.0470 วินาที	0.5470 วินาที

ตารางที่ 4-18 ผลการทดลองเปรียบเทียบด้วยขนาดของค่าลักษณะเด่น

ผลงานการวิจัย	ขนาดของค่าลักษณะเด่น		
	ค่าความสำคัญของสี	ค่าฮิสโตแกรม ของขอบวัตถุภาพ	รวม 2 ลักษณะเด่น
งานวิจัยที่เกี่ยวข้อง	288	925	1213
ICSVM	384	148	532

4.3 สรุปผลการทดลอง

ผลการทดลองคำนวณน้ำหนักของภาพสำหรับการสอนให้ระบบเรียนรู้ และการทดสอบจากหัวข้อ 4.2.1 สามารถให้รูปแบบการจัดชุดข้อมูลสำหรับการสอนให้ระบบเรียนรู้ และการทดสอบที่เหมาะสม ได้แก่ ระบบการตรวจจับทิศทางการวางตัวของภาพโดยใช้ค่าความสำคัญของสี รูปแบบที่เหมาะสมสำหรับการทดลอง คือ รูปแบบที่ 1 (50 :50) โดยเคอร์เนลที่เหมาะสม คือ เคอร์เนลแบบเชิงเส้น ระบบการตรวจจับทิศทางการวางตัวของภาพโดยใช้ค่าฮิสโตแกรมของขอบวัตถุภาพ รูปแบบที่เหมาะสมสำหรับการทดลอง คือ รูปแบบที่ 1 (50 :50) โดยเคอร์เนลที่เหมาะสม คือ เคอร์เนลแบบเรเดียลเบสิสฟังก์ชัน และระบบการตรวจจับการวางตัวของภาพโดยใช้วิธีการ

รวม 2 ลักษณะเด่น รูปแบบที่เหมาะสมสำหรับการทดลอง คือ รูปแบบที่ 1 (50 :50) โดยเคอร์เนลที่เหมาะสม คือ เคอร์เนลแบบเรเดียลเบสิสฟังก์ชัน

จากการทดลองที่ผ่านมา พบว่าผลการทดลองของระบบด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ สามารถให้ผลการทดลองสูงกว่าผลการทดลองด้วยค่าความสำคัญของสี ซึ่งผลการทดลองด้วยค่าฮิสโตแกรมของขอบวัตถุภาพให้ค่าความถูกต้องเฉลี่ยร้อยละ 82 และผลการทดลองของระบบด้วยค่าความสำคัญของสีให้ค่าความถูกต้องเฉลี่ยร้อยละ 81 โดยประเภทภาพที่สามารถตรวจจับได้ดีซึ่งให้ค่าความถูกต้องสูงทั้ง 2 รูปแบบ คือ กลุ่มภาพถ่ายหน้าตรง ประเภทภาพใบหน้ามนุษย์ และประเภทภาพที่ผลการทดลองให้ค่าความถูกต้องน้อยสุด คือ กลุ่มภาพนอกอาคาร ประเภทภาพพระอาทิตย์ตก

จากผลการทดลองของระบบด้วยวิธีการรวม 2 ลักษณะเด่น พบว่าให้ผลการทดลองที่มีประสิทธิภาพสูงกับภาพทุกประเภท และที่สำคัญสามารถแก้ปัญหาภาพที่ตรวจจับได้ยากในกรณีของประเภทภาพพระอาทิตย์ตก จากการทดลอง ในรูปแบบที่ 1 และรูปแบบที่ 2 ได้ ซึ่งผลการทดลองพบว่าสามารถให้ค่าความถูกต้องสูงร้อยละ 90

จากผลการทดลองที่ได้ในตารางที่ 4-16 ตารางที่ 4-17 และตารางที่ 4-18 ซึ่งเป็นผลการทดลองเปรียบเทียบกับงานวิจัยที่เกี่ยวข้อง พบว่าผลการเปรียบเทียบด้านความถูกต้องในงานวิจัยนี้สามารถให้ผลการทดลองสูงกว่าและให้ข้อเด่นในด้านสามารถลดขนาดของเวกเตอร์ของข้อมูล ส่งผลให้ความเร็วของระบบมีประสิทธิภาพสูงกว่าวิธีการของ Wang และ Zhang [1] โดยเฉพาะผลการทดลองด้วยค่าฮิสโตแกรมของขอบวัตถุภาพ สามารถให้ประสิทธิภาพในการตรวจจับที่ดีและความสามารถในการประมวลผลที่รวดเร็ว จึงสามารถแก้ปัญหางานวิจัยในด้านความเร็วและขนาดของเวกเตอร์ข้อมูลที่มีขนาดใหญ่ให้มีขนาดเล็กลงได้

บทที่ 5

สรุปและข้อเสนอแนะ

5.1 สรุปผลการทดลอง

งานวิจัยนี้ได้นำเสนอวิธีการตรวจจับและแก้ไขการวางตัวของภาพโดยใช้ซอฟต์แวร์แมชชีน โดยผู้วิจัยได้ศึกษาทฤษฎีของงานวิจัยต่างๆ และได้พัฒนาขั้นตอนการรวมลักษณะเด่นจากค่าความสำคัญของสีและค่าฮิสโตแกรมของขอบวัตถุภาพโดยการหมุนครั้งละ 90° เพื่อเก็บค่าข้อมูลลักษณะเด่นในแต่ละทิศทาง ซึ่งจากผลการทดลองที่ได้ในบทที่ 4 พบว่า

ผลการทดลองเปรียบเทียบเพื่อหารูปแบบที่ดีที่สุด สำหรับการให้น้ำหนักการสอนให้ระบบเรียนรู้ และการทดสอบ พบว่ารูปแบบที่ 1(50:50) วิธีการจัดรูปแบบข้อมูลภาพประเภทละ 100 ภาพ ทำการจัดรูปแบบโดยจัดให้เป็นกลุ่มภาพสำหรับการสอนให้ระบบเรียนรู้ 50 ภาพ และกลุ่มภาพสำหรับการทดสอบ 50 ภาพ สามารถให้ผลการทดลองที่มีประสิทธิภาพดีที่สุด

ผลการทดลองในด้านค่าความถูกต้องในการตรวจจับของระบบโดยวิธีการรวม 2 ลักษณะเด่น สามารถให้ค่าความถูกต้องดีที่สุด โดยให้ผลการทดลองเฉลี่ยร้อยละ 88 กลุ่มภาพที่สามารถตรวจจับได้ดีที่สุด คือ ประเภทภาพถ่ายใบหน้ามนุษย์ สามารถให้ค่าความถูกต้องสูงร้อยละ 100

ผลการทดลองด้วยค่าฮิสโตแกรมของขอบวัตถุภาพสามารถให้ค่าความถูกต้องที่ดีและให้ค่าความเร็วในการประมวลผลที่สูงแก่ระบบ ซึ่งส่งผลให้งานวิจัยครั้งนี้สามารถพัฒนาและแก้ปัญหาในเรื่องขนาดของवेक्टरที่มีขนาดใหญ่ให้มีขนาดเล็กลง โดยจากวิธีการของ Wang และ Zhang [1] ที่ใช้บล็อกขนาด 5×5 และให้ค่า bins ที่ 37 ดังนั้นवेक्टरที่ได้จะมีขนาด 925 แต่ในงานวิจัยนี้สามารถลดขนาดของवेक्टरโดยให้มีขนาดเพียง 148

5.2 ข้อเด่น และข้อด้อยของวิธีการที่พัฒนาขึ้นใหม่

ข้อเด่นของงานวิจัยนี้คือ สามารถลดขนาดของवेक्टरที่มีขนาดใหญ่ให้เล็กลง ด้านความถูกต้องของการตรวจจับก็มีค่าความถูกต้องสูง แต่ข้อด้อยของงานวิจัยก็คือ เมื่อทำการ 2 รวมลักษณะเด่นเข้าด้วยกัน ส่งผลให้ขนาดของवेक्टरมีลักษณะเป็นมิติเดียวที่ต่อเนื่องกัน อาจมีผลในการทดสอบกับภาพที่มีขนาดใหญ่

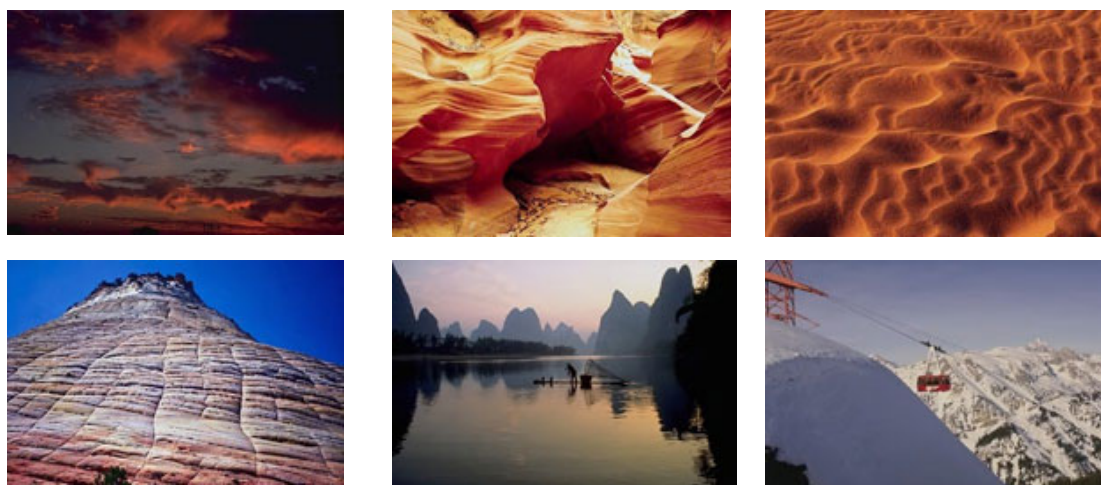
5.3 ข้อเสนอแนะสำหรับงานวิจัยในอนาคต

งานวิจัยนี้ ทำการดึงลักษณะเด่นด้วยวิธีเก็บค่าทิศทางจากการหมุนภาพ 4 ทิศทาง และทำการเก็บเวกเตอร์ในลักษณะมิติเดียว ส่งผลให้เวกเตอร์ที่ได้มีขนาดใหญ่ ซึ่งสามารถแก้ปัญหาโดยใช้วิธีการแบ่งกลุ่มข้อมูลด้วยซัพพอร์ตเวกเตอร์แมชชีนในลักษณะรูปแบบของ Multi – class แบบคู่ขนานเพื่อช่วยเพิ่มความเร็วในการประมวลผล

ในส่วนของการดึงลักษณะเด่นด้วยค่าความสำคัญของสี ซึ่งหากทำการเก็บลักษณะเด่นครั้งหนึ่งของภาพ จากการเก็บในแต่ละทิศทางแล้ว จะสามารถช่วยเพิ่มความถูกต้องในการตรวจจับได้สูงขึ้น แต่เวลาในการประมวลผลจะช้า ซึ่งหากผู้สนใจทำการพัฒนาต่อในส่วนของการค่าความสำคัญของสีนี้ ควรเพิ่มขั้นตอนของการลดทอนข้อมูลหรือการบีบอัดภาพ อาจจะใช้วิธีการ PCA , LDA หรือ 2DPCA เป็นต้น

5.4 ปัญหาของงานวิจัย

ปัญหาที่พบในงานวิจัยนี้คือ ลักษณะเด่นที่ใช้ในงานวิจัยนี้สามารถตรวจจับภาพได้ดีในกลุ่มภาพที่เกิดค่าสีที่แตกต่างกัน และจากภาพที่เกิดลดทอนบนพื้นผิวจำนวนมาก หากเป็นภาพในลักษณะแบบพื้นผิวที่ราบเรียบ ภาพที่มีลักษณะสีคล้ายคลึงกันหรือภาพที่มนุษย์ได้ทำการตกแต่งหรือสร้างขึ้นจากโปรแกรมทางด้านกราฟิกแล้ว การตรวจจับทิศทางการวางตัวของภาพจะทำได้ยาก ซึ่งสามารถแสดงตัวอย่างภาพที่เกิดปัญหาในการตรวจจับได้ดังภาพที่ 5-1



ภาพที่ 5-1 ลักษณะภาพที่เกิดปัญหาในการตรวจจับ

บรรณานุกรม

1. Y. Wang and H. Zhang. "Content-based image orientation detection with support vector machines." IEEE Workshop on Content-based Access of Image and Video Libraries. 2001.
2. A. Vailaya, H. Zhang, and A.K. Jain. "Automatic image orientation detection." in: Proc. Sixth IEEE Internat. Conf. on Image Processing. 1999.
3. Chih-Chung Chang and Chih-Jen Lin. "LIBSVM." A Library for Support Vector Machines. September, 2006.
4. Chapelle O. Chapelle, P. Haffner and V. Vapnik. "SVMs for Histogram-Based Image Classification." Submitted to IEEE Transactions on Neural Networks. 1999 .
5. Vailaya A. Vailaya, A. K. Jain and H. J. Zhang. "On image classification: City images vs. landscapes." Pattern Recognition. 1998.
6. M. Stricker and M. Orengo. "Similarity of color images." In SPIE Conference on Storage and Retrieval for Image and Video Databases III. Feb, 1995.
7. A.K.Jain and A. Vailaya. "Image retrieval using color and shape". Pattern Recognition. August, 1996.
8. Jon Barker and Zhou Su. "Automatic Image Orientation Detection." Department of Computer Science The University Of Sheffield . August, 2004.
9. E. Mayoraz and E. Alpaydm. "Support vector machines for multi-class classification." in: IDIAP Research Report. May, 1998.
10. James Z. Wang, Jia Li and Gio Wiederhold. "SIMPLIcity: Semantics-sensitive Integrated Matching for Picture Libraries." IEEE Trans. on Pattern Analysis and Machine Intelligence. 2001.
11. Lei Zhang, Mingjing Li and Hong-Jiang Zhang. Boosting image orientation detection with indoor vs. outdoor classification. Microsoft Research Asia. 2002.

12. J.F. Canny. "A computational approach to edge detection." IEEE Trans. Pattern Anal. Mach., 1986.
13. Tai Sing Lee. Canny Edge Detection. CMU 15-385 Computer Vision. 2002.
14. <http://www.hike.te.chiba-u.ac.jp/ikeda/CIE>. [15 พฤษภาคม 2550]
15. Nello Cristianini, John Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University press. 2000.
16. อานนท์ นามสนิท. การจำแนกกลุ่มเพลงไทยโดยใช้ซัพพอร์ตเวกเตอร์แมชชีนส์. บัณฑิตวิทยาลัย สถาบันเทคโนโลยี พระจอมเกล้าพระนครเหนือ. 2006 .
17. <http://www.cs.su.ac.th/~prawim/doc/>. [12 เมษายน 2550]

ประวัติผู้วิจัย

ชื่อ : นางสาวจิรา แก้วสุวรรณ
ชื่อวิทยานิพนธ์ : การตรวจจับและการแก้ไขการวางตัวของภาพโดยใช้ซอฟต์แวร์คอมพิวเตอร์
สาขาวิชา : เทคโนโลยีคอมพิวเตอร์

ประวัติ

ประวัติส่วนตัว เกิดวันจันทร์ที่ 8 มกราคม พ.ศ. 2522 ที่จังหวัดพัทลุง
ประวัติการศึกษา สำเร็จการศึกษาระดับปริญญาตรี สาขาวิชาเอก วิศวกรรมคอมพิวเตอร์ จากสถาบันเทคโนโลยีราชมงคล วิทยาเขตพระนครเหนือ
ประวัติการทำงาน พ.ศ. 2548 – 2549 พนักงานในตำแหน่งเว็บไซต์ บริษัท อินโฟรมาซอทพี (ประเทศไทย) จำกัด
สถานที่ติดต่อ บ้านเลขที่ 120 ม.4 ต.ร่มเมือง อ.เมือง จ.พัทลุง 93000
Email : Chira_phat@hotmail.com