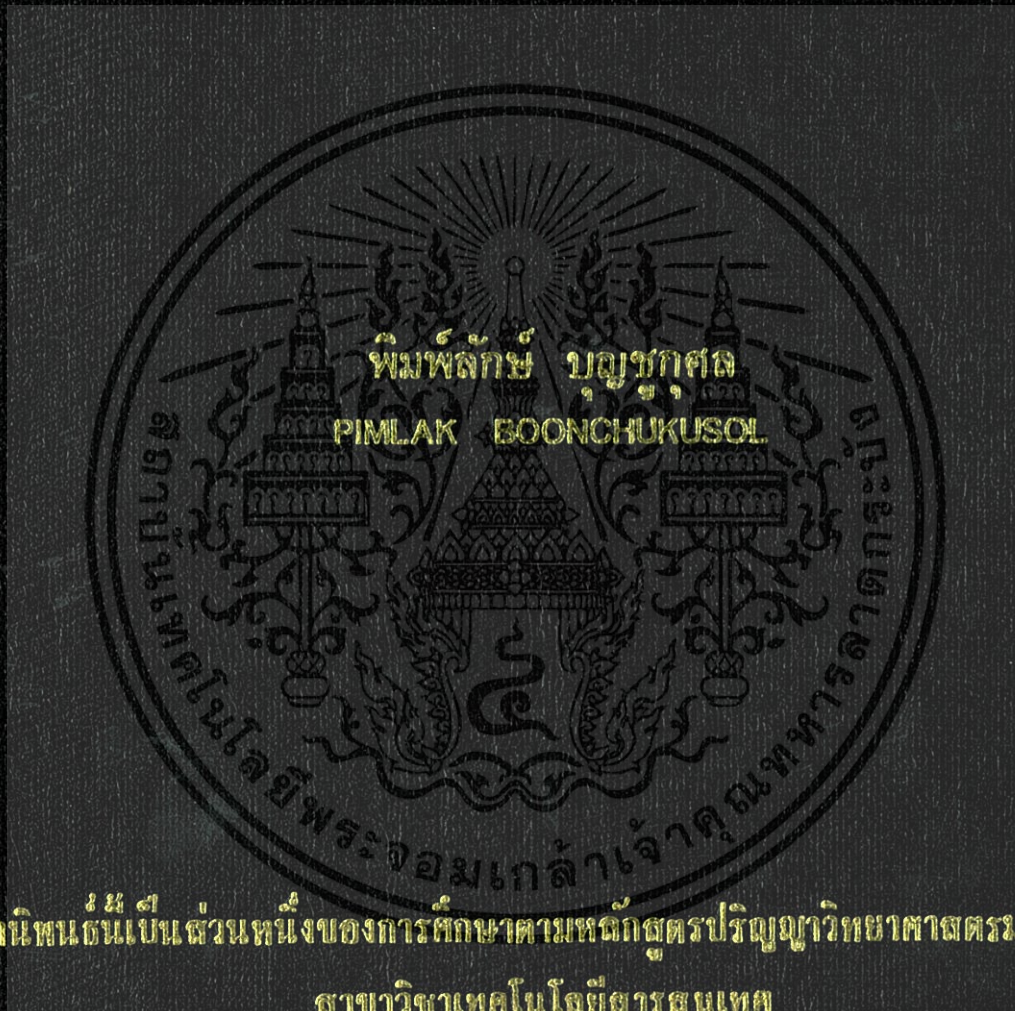


การตรวจจับข้อความภาษาไทยในภาพฉากธรรมชาติระยะปานกลาง
โดยการจำแนกกลุ่มเส้นขอบและการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน

THAI TEXT DETECTION FROM MEDIUM SHOT OF NATURAL SCENES
BY USING FAST BOUNDARY CLUSTERING AND MODIFIED
CONNECTED COMPONENT ANALYSIS



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาค้นคว้าตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2557

KMITL-2014-IT-M-001-004

การตรวจจับข้อความภาษาไทยในภาพฉากธรรมชาติระยะปานกลาง
โดยการจำแนกกลุ่มเส้นขอบและการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน

**THAI TEXT DETECTION FROM MEDIUM SHOT OF NATURAL SCENES
BY USING FAST BOUNDARY CLUSTERING AND MODIFIED
CONNECTED COMPONENT ANALYSIS**



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2557

KMITL-2014-IT-M-001-004

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**THAI TEXT DETECTION FROM MEDIUM SHOT OF NATURAL SCENES
BY USING FAST BOUNDARY CLUSTERING AND MODIFIED
CONNECTED COMPONENT ANALYSIS**



**A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENT FOR THE DEGREE OF
MASTER OF SCIENCE IN INFORMATION TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

2014

KMITL-2014-IT-M-001-004

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2014

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์หรือการป้องกันเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ การตรวจจับข้อความภาษาไทยในภาพจากธรรมชาติระยะปานกลางโดยการจำแนกกลุ่มเส้นขอบและการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน
Thai Text Detection from Medium Shot of Natural Scenes by Using Fast Boundary Clustering and Modified Connected Component Analysis

นักศึกษา นางสาวพิมพ์ลักษณ์ บุญชูกุล
รหัสประจำตัว ๕๑๖๖๐๔๐๓
ปริญญา วิทยาศาสตรมหาบัณฑิต
สาขาวิชา เทคโนโลยีสารสนเทศ
อาจารย์ที่ปรึกษาวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.กัณฑ์พงษ์ วรรณปัญญา

คณะกรรมการสอบวิทยานิพนธ์	ลายมือชื่อ
รองศาสตราจารย์ ดร.นพพร โชติกกำธร	
ผู้ช่วยศาสตราจารย์ ดร.ธนรัตน์ ชลิตาพงศ์	
ผู้ช่วยศาสตราจารย์ ดร.กัณฑ์พงษ์ วรรณปัญญา	
รองศาสตราจารย์ ดร.วรพจน์ กิริสุระเดช	
ดร.ณัฐพล พันธุ์วงศ์	

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

วัน/เดือน/ปี ที่สอบ วันพุธที่ ๑๔ พฤษภาคม ๒๕๕๗ เวลา ๐๙.๓๐ น.

สถานที่สอบ ณ ห้อง ๓๓๓ ชั้น ๓ คณะเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศรับรองแล้ว


(รองศาสตราจารย์ ดร.จันทร์บูรณ์ สลิตวิริยวงศ์)

คณบดีคณะเทคโนโลยีสารสนเทศ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่ควรนำออกนอกระบบโดยไม่ได้รับอนุญาต

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และทำซ้ำโดยไม่ได้รับอนุญาต

วันที่ ๒๖ เดือน พฤษภาคม พ.ศ. ๒๕๕๗

หัวข้อวิทยานิพนธ์

การตรวจจับข้อความภาษาไทยในภาพฉากธรรมชาติระยะปานกลางโดยการจำแนกกลุ่มเส้นขอบและการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน

นักศึกษา

นางสาวพิมพ์ลักษณ์ บุญชูกุล

รหัสนักศึกษา

53660403

ปริญญา

วิทยาศาสตรมหาบัณฑิต

สาขาวิชา

เทคโนโลยีสารสนเทศ

แขนงวิชา

เทคโนโลยีระบบสารสนเทศ

พ.ศ.

2557

อาจารย์ที่ปรึกษา

ผศ.ดร.กันต์พงษ์ วรรณปัญญา

บทคัดย่อ

การตรวจจับข้อความภาษาไทยในฉากธรรมชาติเป็นงานที่มีความท้าทายหลายประการเนื่องจากปัจจัยต่าง ๆ ที่ทำให้ฉากมีความซับซ้อน และงานวิจัยด้านการตรวจจับข้อความในฉากธรรมชาติต่าง ๆ ที่มีอยู่ยังไม่สามารถนำมาใช้ในการตรวจจับข้อความภาษาไทยได้อย่างสมบูรณ์เนื่องจากข้อความภาษาไทยประกอบด้วยสระและวรรณยุกต์ซึ่งมีลักษณะการวางแนวในข้อความต่างจากภาษาอื่น วิทยานิพนธ์ฉบับนี้ได้นำเสนอขั้นตอนวิธีตรวจจับข้อความภาษาไทยในฉากธรรมชาติซึ่งประกอบด้วย 2 ขั้นตอนหลักคือ 1) การจำแนกกลุ่มเส้นขอบโดยพิจารณาจากสีของวัตถุ สีของพื้นหลัง และตำแหน่งของวัตถุ เพื่อจำแนกตัวอักษรภาษาไทยออกจากพื้นหลังที่มีความซับซ้อน และ 2) การวิเคราะห์โครงสร้างและลักษณะการวางตัวของขอบเขตของวัตถุเพื่อตัดสินใจว่าวัตถุใดเป็นตัวอักษรภาษาไทย ซึ่งการจำแนกกลุ่มเส้นขอบเป็นวิธีจัดการข้อมูลให้อยู่ในรูปแบบที่ง่ายต่อการวิเคราะห์ขอบเขตของวัตถุ จากการทดลองกับภาพฉากธรรมชาติจำนวน 300 ภาพ ซึ่งประกอบด้วยขอบเขตตัวอักษรภาษาไทย 8,077 ขอบเขต พบว่าขั้นตอนวิธีนี้มีค่าความแม่นยำและค่าความครบถ้วนอยู่ที่ 0.87 และ 0.79

Thesis	Thai Text Detection from Medium Shot of Natural Scenes by Using Fast Boundary Clustering and Modified Connected Component Analysis
Student	Miss Pimlak Boonchukusol
Student ID.	53660403
Degree	Master of Science
Program	Information Technology
Major	Information System Technology
Year	2014
Thesis Advisor	Asst.Prof.Dr. Kuntpong Woraratpanya

ABSTRACT

Thai text detection from natural scenes is still a challenging task, since there are many unsolved issues. Furthermore, the existing related works cannot completely detect Thai text. The main reason is that Thai text layout has vowels and tonal marks that differ from other languages. This thesis proposes an approach to detect Thai text from natural scene images. This approach is composed of two main procedures. (i) Boundary clustering algorithm decomposes scene features into multilayer based on colors and positions. (ii) Connected component analysis by using a structure of text characters and character alignment in order to detect Thai text boundaries. Boundary clustering is applied to decompose character boundaries from other objects and improve the efficiency of Thai text detection from natural scenes. Based on 300 test images with 8,077 character boundaries, the experimental results demonstrate that the proposed approach achieves the high average precision and recall, 0.87 and 0.79.

กิตติกรรมประกาศ

ขอขอบคุณ ผศ.ดร.กนต์พงษ์ วรรณรัตน์ปัญญา อาจารย์ที่ปรึกษา ที่ได้กรุณาให้คำปรึกษา
ชี้แนะแนวทางในการทำวิจัย ตลอดจนตรวจแก้ไขวิทยานิพนธ์ฉบับนี้ให้สำเร็จลุล่วงได้ด้วยดี

ขอขอบคุณสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังที่ได้มอบโอกาสให้แก่
ข้าพเจ้าได้ไปเรียนรู้และแลกเปลี่ยนประสบการณ์ด้านการวิจัย ภาษา และวัฒนธรรมที่ประเทศญี่ปุ่น
เป็นระยะเวลา 4 เดือน

ขอขอบคุณคณาจารย์คณะเทคโนโลยีสารสนเทศ และเพื่อน ๆ สมาชิกห้องปฏิบัติการวิจัย
ด้านการรู้จำแบบและการประมวลผลภาพ (PRIP) ที่ได้ให้คำปรึกษา รวมทั้งถ่ายทอดความรู้ต่าง ๆ
ให้แก่ข้าพเจ้า

ท้ายที่สุด ขอขอบคุณบิดา มารดา และครอบครัวที่ทำให้กำลังใจสนับสนุนเป็นอย่างดี
จนกระทั่งสำเร็จการศึกษา

พิมพ์ลักษณ์ บุญชูกุล



สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VI
สารบัญภาพ.....	VII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา.....	3
1.3 กรอบแนวคิดในการทำวิจัย.....	3
1.4 ขอบเขตการวิจัย.....	5
1.5 ประโยชน์ที่ได้รับ.....	6
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	7
2.1 ทฤษฎีเกี่ยวกับตัวอักษรภาษาไทย.....	7
2.1.1 โครงสร้างพื้นฐาน.....	7
2.1.2 สัดส่วนต่าง ๆ ของตัวอักษร.....	8
2.2 การประมวลผลภาพดิจิทัล.....	9
2.2.1 ภาพระดับสีเทา.....	9
2.2.2 ภาพสี.....	10
2.2.3 การหาขอบภาพแบบแคนนี่.....	10
2.3 การจำแนกกลุ่มข้อมูล.....	12
2.4 สมการเส้นตรง.....	13
2.5 งานวิจัยที่เกี่ยวข้อง.....	13
2.5.1 การตรวจจับข้อความโดยพิจารณาจากพื้นที่.....	13
2.5.2 การตรวจจับข้อความโดยพิจารณาจากส่วนประกอบที่เชื่อมต่อกัน.....	14
2.5.3 การแยกตัวอักษรออกจากวัตถุอื่น ๆ ก่อนทำการตรวจจับข้อความ.....	15

สารบัญ (ต่อ)

	หน้า
บทที่ 3 วิธีดำเนินการวิจัย.....	16
3.1 กระบวนการเตรียมภาพ.....	16
3.2 กระบวนการจำแนกกลุ่มเส้นขอบ.....	17
3.3 การวิเคราะห์ขอบเขตของวัตถุ แบบ ก.....	20
3.3.1 การวิเคราะห์เพื่อตรวจจับตัวอักษรไทยที่วางแนวอยู่ในส่วนลำตัวหลัก.....	20
3.3.2 การวิเคราะห์เพื่อตรวจจับสระบนและสระล่าง.....	23
3.3.3 การวิเคราะห์เพื่อตรวจจับวรรณยุกต์.....	25
3.4 การวิเคราะห์ขอบเขตของวัตถุ แบบ ข.....	26
3.4.1 การวิเคราะห์เพื่อตรวจจับตัวอักษรไทยที่วางแนวอยู่ในส่วนลำตัวหลัก.....	26
3.4.2 การรวมกลุ่มข้อความ.....	29
3.4.3 การขยายเส้นกรอบเพื่อตรวจจับสระบน สระล่าง และวรรณยุกต์.....	30
บทที่ 4 ผลการทดลอง.....	33
4.1 การวัดประสิทธิภาพ.....	33
4.2 การออกแบบการทดลอง.....	34
4.3 การทดลองที่ 1.....	34
4.4 การทดลองที่ 2.....	37
4.5 การทดลองที่ 3.....	38
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....	42
เอกสารอ้างอิง.....	45
ภาคผนวก.....	48
ภาคผนวก ก ภาพที่ใช้ในการทดลอง.....	49
ภาคผนวก ข ผลงานวิจัยที่ได้รับการตีพิมพ์เผยแพร่.....	56
ประวัติผู้เขียน.....	75

สารบัญตาราง

ตารางที่	หน้า
1.1 ชุดข้อมูลที่ใช้ในการทดลอง.....	6
2.1 สมาชิกตัวอักษรภาษาไทย.....	7
3.1 สมาชิกตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนลำตัวหลัก จำแนกตามลักษณะของหาง.....	20
4.1 ผลการตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยใช้วิธี ACG.....	34
4.2 ผลการตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยใช้วิธี MCCA1.....	35
4.3 ผลการตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยใช้วิธี MCCA2.....	35
4.4 ผลการตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยใช้วิธี FBC ร่วมกับ MCCA2.....	37
4.5 ผลการตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยใช้วิธี BC ร่วมกับ MCCA2.....	38
4.6 ตัวอย่างผลลัพธ์ของการตรวจจับข้อความภาษาไทยที่ถูกต้องในกรณีต่าง ๆ.....	39
4.7 ตัวอย่างผลลัพธ์ของการตรวจจับตัวอักษรภาษาไทยที่ผิดพลาดในกรณีต่าง ๆ.....	40



สารบัญภาพ

ภาพที่	หน้า
1.1 กรอบแนวคิดในการทำวิจัย.....	4
2.1 ส่วนประกอบของตัวอักษรภาษาไทย.....	8
2.2 ตัวอย่างภาพระดับสีเทาขนาด 8 บิต.....	10
2.3 ตัวอย่างภาพสีและค่าความเข้มแสงของแต่ละพิกเซล.....	10
3.1 ผลลัพธ์ที่ได้จากกระบวนการเตรียมภาพในแต่ละขั้นตอน.....	17
3.2 การสกัดคุณลักษณะเด่นของวัตถุ.....	18
3.3 ผลลัพธ์ที่ได้จากการจำแนกกลุ่มเส้นขอบด้วยวิธี K-means.....	19
3.4 ตัวแปรที่ใช้ในการตัดสินใจว่าวัตถุใดเป็นตัวอักษรที่อยู่ระหว่างเส้นหลักกับเส้นฐาน.....	22
3.5 การคำนวณค่าความชันของ (ก) เส้นหลักและเส้นฐานของกลุ่มตัวอักษรที่ตรวจจับได้ (ข) เส้นหลักสมมติและเส้นฐานสมมติระหว่างกลุ่มตัวอักษรกับวัตถุใด ๆ.....	22
3.6 ตำแหน่งตัวแปรที่ใช้ในกฎสำหรับตัดสินใจว่าวัตถุใดเป็นสระบนหรือสระล่าง.....	25
3.7 ตัวอย่างผลลัพธ์ที่ได้จากการวิเคราะห์ขอบเขตของวัตถุแบบ ก.....	25
3.8 การคำนวณค่าความชันของเส้นหลักและเส้นฐานของตัวอักษรสองตัวที่ไม่มีหาง และเรียงต่อกัน.....	27
3.9 การฉายภาพเพื่อตรวจจับตัวอักษรที่อยู่ในแนวเดียวกัน (ก) ทิศทางในการฉายภาพ (ข) ตำแหน่งตัวแปรที่ใช้ในกฎเพื่อตัดสินใจว่าวัตถุที่พบจากการฉายภาพเป็นตัวอักษร หรือไม่.....	28
3.10 ตัวแปรที่ใช้ในการวิเคราะห์ว่าข้อความสองข้อความเป็นข้อความเดียวกันหรือไม่.....	30
3.11 การขยายเส้นกรอบเพื่อตรวจจับสระบน สระล่าง และวรรณยุกต์.....	30
3.12 ตัวอย่างผลลัพธ์ที่ได้จากการวิเคราะห์ขอบเขตของวัตถุแบบ ข.....	32
4.1 ตัวอย่างผลลัพธ์ที่ได้จากขั้นตอนวิธีตรวจจับข้อความภาษาไทยในวิทยานิพนธ์นี้ จำแนกตามชุดข้อมูล (ก) กลุ่มที่ 1 (ข) กลุ่มที่ 2 (ค) กลุ่มที่ 3.....	38
4.2 ตัวอย่างผลลัพธ์จากการจำแนกกลุ่มเส้นขอบที่ส่งผลให้ระบบไม่สามารถตรวจจับตัวอักษร ภาษาไทยได้.....	41

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในชีวิตประจำวันเราสามารถพบเห็นป้ายข้อความได้ทุกที่ทุกเวลา เช่น ป้ายบอกทาง ป้ายชื่อสถานที่ ป้ายแสดงคำเตือน และคำอธิบายต่าง ๆ ซึ่งข้อความเหล่านี้จัดเป็นสารสนเทศที่ช่วยให้การดำรงชีวิตของเราง่ายขึ้น ดังนั้น การตรวจจับข้อความภาษาไทยจึงเป็นขั้นตอนที่สำคัญในการนำข้อความเหล่านี้ไปประยุกต์ใช้ในงานด้านต่าง ๆ เช่น แอปพลิเคชันแปลภาษาบนสมาร์ตโฟนเพื่อช่วยเหลือชาวต่างชาติให้สามารถเข้าใจข้อความภาษาไทยที่ปรากฏบนป้าย ระบบตรวจจับและรู้จำป้ายทะเบียนรถ และระบบสังเคราะห์เสียงพูดเพื่อช่วยเหลือคนตาบอด เป็นต้น

การตรวจจับข้อความในฉากธรรมชาติเป็นงานที่มีความท้าทายหลายประการ เพราะนอกจากความหลากหลายของข้อความ ได้แก่ รูปแบบ ขนาด สี และทิศทางของตัวอักษรแล้ว ความไม่ชัดเจนของข้อความที่ปรากฏบนป้ายอันเนื่องมาจากสาเหตุต่าง ๆ เช่น ตัวอักษรซีดจาง สภาพแสงน้อย แสงสะท้อนจากดวงอาทิตย์ หลอดไฟ หรือแสงแฟลช ตลอดจนเงาของวัตถุอื่นที่พาดผ่านบริเวณป้าย สิ่งเหล่านี้ล้วนเป็นปัจจัยที่ส่งผลให้การตรวจจับข้อความในฉากธรรมชาติมีอัตราความผิดพลาดสูงกว่าการตรวจจับข้อความบนเอกสารที่ได้จากเครื่องสแกน ซึ่งมีการควบคุมปัจจัยแวดล้อมต่าง ๆ ที่จะมีผลต่อคุณภาพของข้อความ

วิธีการตรวจจับข้อความในฉากธรรมชาติที่มีอยู่ในปัจจุบันสามารถจำแนกเป็น 2 ประเภทหลัก คือ การตรวจจับข้อความโดยพิจารณาจากพื้นที่ (Region-based) และการตรวจจับข้อความโดยพิจารณาจากส่วนประกอบที่เชื่อมต่อกัน (Connected Component-based: CC-based) โดยวิธีแรกมีวัตถุประสงค์เพื่อตรวจจับและระบุตำแหน่งของพื้นที่ข้อความ โดยการวิเคราะห์พื้นผิว (Texture) เริ่มจากการสกัดเวกเตอร์คุณลักษณะเด่น (Feature Vector) ของพื้นที่ส่วนย่อย (Local Region) แล้วนำไปจำแนกด้วยตัวจำแนก (Classifier) เพื่อคาดคะเนว่าเป็นข้อความหรือไม่ สุดท้ายจึงนำพื้นที่ข้อความที่อยู่ใกล้กันรวมเข้าด้วยกันเป็นบล็อก [1], [2], [3], [4] เนื่องจากพื้นที่ข้อความมีลักษณะแตกต่างจากพื้นที่ที่ไม่ใช่ข้อความอย่างชัดเจน วิธีการนี้จึงสามารถตรวจจับข้อความได้อย่างมีประสิทธิภาพแม้ว่าจะมีสัญญาณรบกวน แต่ก็ยังมีข้อเสีย คือ ต้องใช้เวลาในการประมวลผลนานในทางกลับกัน วิธีตรวจจับข้อความโดยพิจารณาจากส่วนประกอบที่เชื่อมต่อกันเป็นการตัดแยก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนประกอบต่าง ๆ ในภาพโดยตรง ด้วยตัวดำเนินการหาขอบภาพหรือการจำแนกสี จากนั้นส่วนประกอบที่ไม่ใช่ข้อความจะถูกคัดกรองออกโดยการใช้อ็กรหรือตัวจำแนก [5], [6], [7] การตรวจจับข้อความด้วยวิธีนี้ใช้เวลาในการประมวลผลน้อย เนื่องจากส่วนประกอบที่ตัดแยกได้ต่อภาพมีจำนวนไม่มาก และส่วนประกอบของตัวอักษรที่ตรวจจับได้ก็สามารถนำไปใช้กับการรู้จำตัวอักษร (Character Recognition) ได้โดยตรง

Chucai Yi และคณะ [6] นำเสนอวิธี Adjacent Character Grouping เพื่อตรวจจับข้อความที่ปรากฏในภาพฉากธรรมชาติโดยวัดจากความสูง ความกว้าง จุดกึ่งกลาง ขนาดพื้นที่ตัวอักษร และลักษณะการเชื่อมต่อกันระหว่างวัตถุ วิธีการนี้สามารถตรวจจับข้อความภาษาอังกฤษในภาพฉากธรรมชาติได้อย่างมีประสิทธิภาพ แต่เมื่อนำมาใช้ในการตรวจจับข้อความภาษาไทยพบว่าไม่สามารถตรวจจับสระและวรรณยุกต์ภาษาไทยซึ่งมีลักษณะการวางแนวแตกต่างจากภาษาอังกฤษและผลลัพธ์ที่ได้มีวัตถุอื่น ๆ ที่ไม่ใช่ตัวอักษร แต่ถูกตรวจจับว่าเป็นตัวอักษรรวมอยู่ด้วยจำนวนมาก สำหรับงานวิจัยด้านการตรวจจับข้อความภาษาไทยที่ปรากฏในภาพฉากธรรมชาติ วิทยา จิรัฐติเจริญ [7] นำเสนอวิธีการตรวจจับและตัดแยกตัวอักษร โดยวัดจากความสูง ความกว้าง ขนาดพื้นที่ตัวอักษร และขนาดพื้นที่ที่รอบล้อม เพื่อตรวจจับตัวอักษรภาษาไทยส่วนใหญ่ที่วางอยู่ในแนวเดียวกัน จากนั้นจึงวิเคราะห์หาแนวของข้อความ กำหนดกรอบล้อมข้อความ แล้วขยายกรอบเพื่อครอบคลุมสระและวรรณยุกต์ภาษาไทย วิธีการนี้มีข้อดีคือ สามารถตรวจจับสระบน สระล่าง และวรรณยุกต์ได้ง่าย เนื่องจากทราบตำแหน่งของแนวข้อความ แต่การวิเคราะห์หาแนวข้อความทำได้ยาก เนื่องจากตัวอักษรส่วนใหญ่ที่ตรวจจับได้ในขั้นตอนแรกมีตัวอักษรที่มีหางด้านบนหรือด้านล่างรวมอยู่ด้วย จากข้อดีและข้อเสียของวิธีการข้างต้น วิทยานิพนธ์นี้จึงได้นำเสนอวิธีการตรวจจับข้อความในฉากธรรมชาติ โดยพัฒนาขั้นตอนวิธีของ Chucai Yi และคณะ ให้มีประสิทธิภาพมากยิ่งขึ้นเพื่อประยุกต์ใช้กับภาษาไทย กรอบแนวคิดหลักของวิธีการที่นำเสนอจะพิจารณาจากความสูง ความกว้าง จุดกึ่งกลาง และลักษณะการเชื่อมต่อกันระหว่างวัตถุไปพร้อม ๆ กับการวิเคราะห์หาแนวของข้อความเพื่อให้สามารถตรวจจับตัวอักษรภาษาไทยได้ถูกต้องและครบถ้วน โดยมีการกำหนดเกณฑ์ที่ใช้ในการตัดสินใจว่าวัตถุใดเป็นตัวอักษรให้สอดคล้องกับสัดส่วนและโครงสร้างของตัวอักษรภาษาไทย

อย่างไรก็ตาม เนื่องจากภาพฉากธรรมชาติมักมีพื้นหลังที่ซับซ้อนและมีวัตถุปรากฏอยู่บนภาพจำนวนมาก การวิเคราะห์คุณสมบัติและลักษณะการเชื่อมต่อกันของวัตถุเพียงอย่างเดียวจึงอาจกำจัดวัตถุที่ไม่ใช่ตัวอักษรแต่มีลักษณะคล้ายกับตัวอักษรได้ไม่ด้นัก ดังนั้น จึงมีนักวิจัยได้นำเสนอ

วิธีการแยกตัวอักษรออกจากวัตถุอื่น ๆ ก่อนทำการวิเคราะห์และตัดสินใจว่าวัตถุใดเป็นตัวอักษร โดยมีวัตถุประสงค์เพื่อช่วยเพิ่มความถูกต้องในการตรวจจับข้อความ [8], [9]

Chucui Yi ได้นำเสนอวิธีการ Boundary Clustering [9] ซึ่งสามารถจำแนกตัวอักษรออกจากวัตถุอื่น ๆ ได้ดี แต่ใช้เวลาในการประมวลผลนาน เนื่องจากในการสกัดคุณลักษณะเด่นจะต้องพิจารณาค่าสีของทุก ๆ พิกเซลที่ล้อมรอบพิกเซลขอบทั้งหมดแล้วหาค่าเฉลี่ยเพื่อให้ได้ค่าสีที่เป็นตัวแทนของแต่ละวัตถุ ดังนั้น วิทยานิพนธ์นี้จึงได้นำวิธีการจำแนกกลุ่มเส้นขอบมาใช้เพื่อแยกวัตถุที่น่าจะเป็นข้อความออกจากวัตถุอื่น ๆ ก่อนทำการวิเคราะห์และตัดสินใจว่าวัตถุใดเป็นตัวอักษรภาษาไทย และแก้ไขปัญหาด้านเวลาที่ใช้ในการประมวลผล โดยในการสกัดคุณลักษณะเด่นของวัตถุเพื่อนำไปใช้ในการจำแนกกลุ่มจะพิจารณาเฉพาะค่าสีของพิกเซลที่สนใจเท่านั้น

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

1.2.1 เพื่อพัฒนาขั้นตอนวิธีการตรวจจับข้อความภาษาไทยในฉากรวมชาติ โดยใช้วิธีการจำแนกกลุ่มเส้นขอบและวิธีวิเคราะห์ขอบเขตของวัตถุ

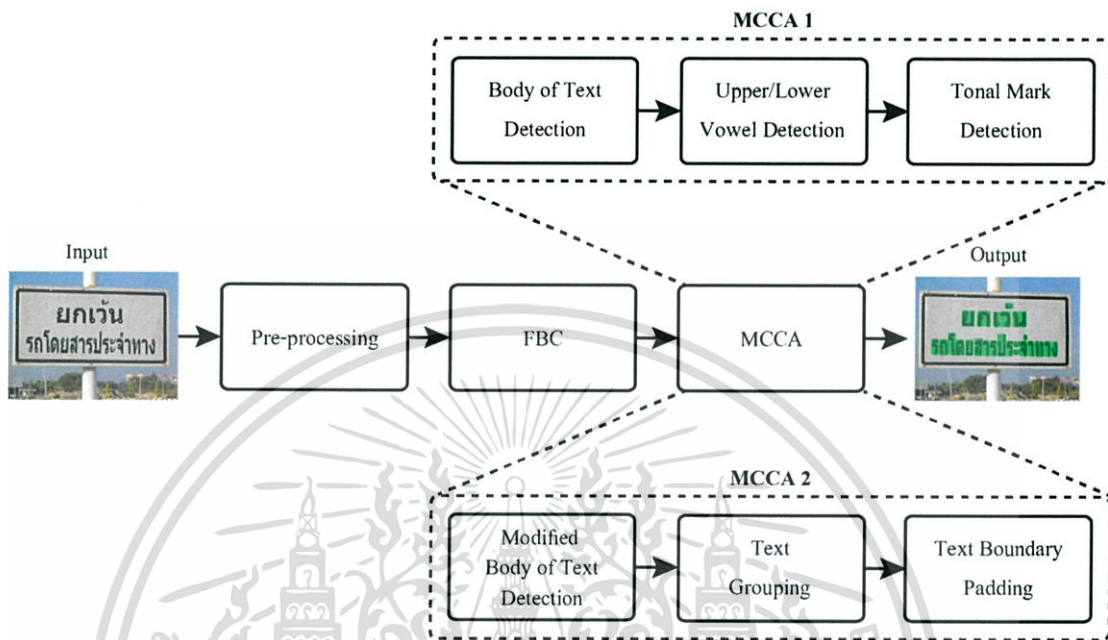
1.2.2 เพื่อพัฒนาขั้นตอนวิธีการจำแนกกลุ่มเส้นขอบให้สามารถประมวลผลได้รวดเร็วและมีประสิทธิภาพ

1.2.3 เพื่อพัฒนาขั้นตอนวิธีการวิเคราะห์ขอบเขตของวัตถุให้สามารถตรวจจับข้อความภาษาไทยในฉากรวมชาติได้อย่างมีประสิทธิภาพ

1.3 กรอบแนวคิดในการทำวิจัย

ขั้นตอนวิธีการตรวจจับข้อความภาษาไทยในฉากรวมชาติในวิทยานิพนธ์นี้มี 2 ขั้นตอนหลัก คือ ขั้นตอนการจำแนกกลุ่มเส้นขอบ (Fast Boundary Clustering: FBC) และขั้นตอนการวิเคราะห์ขอบเขตของวัตถุ (Modified Connected Component Analysis: MCCA) โดยเริ่มจากการเตรียมภาพ (Pre-processing) จากนั้นจึงสกัดคุณลักษณะเด่นของวัตถุในภาพแล้วนำคุณลักษณะที่สกัดได้ไปใช้ในการจำแนกกลุ่ม ซึ่งในขั้นตอนนี้วัตถุที่มีสีพื้นหลัง สีของวัตถุ และตำแหน่งใกล้เคียงกันจะถูกจำแนกให้อยู่กลุ่มเดียวกัน จากนั้นวัตถุทั้งหมดจะถูกนำมาวิเคราะห์เพื่อตัดสินใจว่าวัตถุใดเป็นตัวอักษรภาษาไทยโดยใช้กฎซึ่งสร้างขึ้นจากการวิเคราะห์สัดส่วนและโครงสร้างของตัวอักษรภาษาไทย โดยในวิทยานิพนธ์นี้มีวิธีการวิเคราะห์ขอบเขตของวัตถุ 2 แบบ คือ การ

วิเคราะห์ขอบเขตวัตถุแบบ ก (MCCA 1) [10] และการวิเคราะห์ขอบเขตวัตถุแบบ ข (MCCA 2) [11] ดังแสดงในภาพที่ 1.1 และกระบวนการดำเนินการวิจัยโดยรายละเอียดจะได้อธิบายในบทที่ 3



ภาพที่ 1.1 กรอบแนวคิดในการทำวิจัย

จากกรอบแนวคิดในภาพที่ 1.1 การวิเคราะห์ขอบเขตวัตถุแบบ ก ประกอบด้วย 3 ขั้นตอน ได้แก่ การตรวจจับตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนลำตัวหลัก (Body of Text Detection) การตรวจจับสระบนและสระล่าง (Upper/Lower Vowel Detection) และการตรวจจับวรรณยุกต์ (Tonal Mark Detection) โดยขั้นตอนการตรวจจับตัวอักษรที่วางแนวอยู่ในส่วนลำตัวหลักจะเริ่มจากการตรวจจับตัวอักษรที่เรียงต่อกันในแนวนอนให้ได้มากที่สุด จากนั้นวิเคราะห์หาแนวของข้อความโดยคำนวณจากตำแหน่งของตัวอักษรตัวแรกและตัวสุดท้าย แล้วตรวจจับตัวอักษรอื่น ๆ ที่อยู่ในแนวเดียวกันโดยยึดจากแนวของข้อความที่วิเคราะห์ได้ สุดท้ายทำการตรวจจับสระบน สระล่าง และวรรณยุกต์โดยใช้วิธีตรวจจับตัวอักษรที่เรียงต่อกันในแนวตั้ง การวิเคราะห์ขอบเขตวัตถุแบบ ก มีโอกาสมากที่แนวของข้อความที่วิเคราะห์ได้อาจไม่ใช่แนวข้อความที่แท้จริง ส่งผลให้การตรวจจับตัวอักษรภาษาไทยที่เหลือเกิดข้อผิดพลาดตามมา

เพื่อแก้ปัญหาดังกล่าว วิทยานิพนธ์นี้จึงได้นำเสนอการวิเคราะห์ขอบเขตวัตถุแบบ ข ซึ่งประกอบด้วย 3 ขั้นตอน ได้แก่ การตรวจจับตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนลำตัวหลัก

(Modified Body of Text Detection) การรวมกลุ่มข้อความที่อยู่ในแนวเดียวกัน (Text Grouping)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และการขยายกรอบล้อมรอบข้อความ (Text Boundary Padding) เพื่อตรวจจับสระบน สระล่าง และวรรณยุกต์ โดยขั้นตอนการตรวจจับตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนลำตัวหลักจะเริ่มจากการตรวจจับตัวอักษรสองตัวที่อยู่ติดกันแล้ววิเคราะห์หาแนวของข้อความจากตำแหน่งของตัวอักษรทั้งสองที่ตรวจจับได้ จากนั้นจึงทำการตรวจจับตัวอักษรอื่น ๆ ที่อยู่ในแนวเดียวกัน สุดท้ายทำการตรวจจับสระบน สระล่าง และวรรณยุกต์โดยยึดจากแนวของข้อความที่วิเคราะห์ได้ในขั้นตอนแรกเป็นหลัก

นอกจากกระบวนการวิเคราะห์ขอบเขตวัตถุแบบ ก และแบบ ข จะแตกต่างกันแล้ว เกณฑ์ที่กำหนดขึ้นเพื่อตัดสินใจว่าวัตถุใดเป็นตัวอักษรภาษาไทยก็แตกต่างกันด้วย โดยในการวิเคราะห์ขอบเขตวัตถุแบบ ก จะใช้เกณฑ์ที่สร้างขึ้นจากลักษณะความสัมพันธ์ระหว่างตัวอักษรภาษาไทยสองตัวที่เรียงต่อกันซึ่งได้จากการสังเกตข้อมูลบางส่วนที่ใช้ในการทดลอง ส่วนการวิเคราะห์ขอบเขตวัตถุแบบ ข จะใช้เกณฑ์จากแบบ ก บางส่วน ตัดสินใจร่วมกับเกณฑ์ที่สร้างขึ้นใหม่จากสัดส่วนมาตรฐานของตัวอักษรภาษาไทย [12]

1.4 ขอบเขตการวิจัย

1.4.1 การตรวจจับข้อความภาษาไทยเป็นการตรวจจับเฉพาะพยัญชนะ สระ และวรรณยุกต์

1.4.2 ชุดข้อมูลที่ใช้ในการทดสอบประสิทธิภาพเป็นชุดข้อมูลที่สร้างขึ้นใหม่ ประกอบด้วยภาพสี ขนาด 480 x 360 พิกเซล จำนวน 300 ภาพ แต่ละภาพมีคุณสมบัติดังนี้

1. ประกอบด้วยข้อความภาษาไทยจำนวน 1-5 บรรทัด และข้อความเป็นแนวตรง
2. ตัวอักษรในภาพเป็นตัวพิมพ์เท่านั้น
3. ตัวอักษรในข้อความเดียวกันมีขนาดเท่ากัน และมีค่าความเข้มแสงใกล้เคียงกัน
4. มุมมองภาพอยู่ในระดับเดียวกับสายตา (Normal Eye View) รูปร่างของวัตถุในภาพไม่ผิดเพี้ยนในลักษณะของทัศนมิติ (Perspective) หรือผิดเพี้ยนเพียงเล็กน้อย
5. พื้นหลังป้ายและฉากภาพมีลักษณะคงที่ ไม่เคลื่อนไหว

1.4.3 ชุดข้อมูลที่ใช้ในการทดลอง ถูกจำแนกเป็น 3 กลุ่ม กลุ่มละ 100 ภาพ โดยมีลักษณะและอัตราส่วนร้อยละของพื้นที่ป้ายต่อพื้นที่ทั้งหมดของข้อมูลแต่ละกลุ่มดังตารางที่ 1.1

ตารางที่ 1.1 ชุดข้อมูลที่ใช้ในการทดลอง

กลุ่มที่	ลักษณะของข้อมูล		ร้อยละของพื้นที่ป้ายต่อพื้นที่ทั้งหมด			
	พื้นหลังป้าย	ฉากภาพ	Min	Max	Mean	S.D.
1	เป็นวัสดุพื้นเรียบ	เป็นพื้นเรียบ เช่น ผนัง	11.22	100.00	50.71	27.15
2	มีสีเดียวและไม่มีลวดลาย	เป็นวิวทิวทัศน์ หรือมีวัตถุอื่น ๆ อยู่ในภาพ	5.08	89.12	36.98	17.49
3	ป้ายทำจากวัสดุพื้นไม่เรียบ มันวาว โปร่งใส มีลวดลาย หรือมีความซับซ้อนอันเนื่องมาจากแสงและเงา		13.51	100.00	54.32	27.71

1.4.4 การตรวจจับข้อความภาษาไทยในงานวิจัยนี้ไม่มีการรู้จำตัวอักษร

1.4.5 ผลลัพธ์ที่ได้เป็นภาพลักษณะชาติที่มีการระบุตำแหน่งของขอบเขตวัตถุที่ตรวจจับได้ว่าเป็นตัวอักษรภาษาไทย

1.4.6 งานวิจัยนี้ไม่ครอบคลุมการตรวจจับตัวอักษรในกรณีดังต่อไปนี้

1. ตัวอักษรที่มีขอบเขตไม่สมบูรณ์ อันเนื่องมาจากแสง เงา หรือภาพสะท้อนที่พาดทับบริเวณป้าย หรือตัวอักษรที่มีขอบเขตเชื่อมติดกับวัตถุหรือตัวอักษรอื่น
2. ตัวอักษรที่มีค่าความเข้มแสงของตัวอักษรและพื้นหลังใกล้เคียงกัน
3. ตัวอักษรที่ปรากฏบนป้ายที่ทำจากวัสดุโปร่งใสและมองเห็นวัตถุอื่น ๆ ที่อยู่ด้านหลังของป้าย

1.5 ประโยชน์ที่ได้รับ

1.5.1 ได้ขึ้นตอนวิธีตรวจจับข้อความภาษาไทยในลักษณะชาติแบบใหม่

1.5.2 สามารถนำขึ้นตอนวิธีตรวจจับข้อความภาษาไทยในลักษณะชาติที่พัฒนาขึ้นไปใช้งานได้อย่างมีประสิทธิภาพ

1.5.3 เป็นแนวทางในการพัฒนาขึ้นตอนวิธีตรวจจับข้อความภาษาไทยในลักษณะชาติแบบอื่น ๆ ในอนาคต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยทฤษฎีพื้นฐานต่าง ๆ ที่นำมาใช้ในงานวิจัย ได้แก่ ทฤษฎีที่เกี่ยวกับตัวอักษรภาษาไทย ทฤษฎีการประมวลผลภาพดิจิทัล (Digital Image Processing) เช่น ภาพระดับสีเทา (Intensity Image) ภาพสี (RGB Image) และการหาขอบภาพแบบแคนนี่ (Canny Edge Detection) ทฤษฎีการจำแนกกลุ่มข้อมูล (Data Clustering) ตลอดจนทฤษฎีทางด้านคณิตศาสตร์

2.1 ทฤษฎีเกี่ยวกับตัวอักษรภาษาไทย

ตัวอักษรภาษาไทยประกอบด้วยพยัญชนะ 44 ตัว สระ 16 ตัว วรรณยุกต์ 4 ตัว และอักขระพิเศษ 4 ตัว ดังแสดงในตารางที่ 2.1

ตารางที่ 2.1 สมาชิกตัวอักษรภาษาไทย

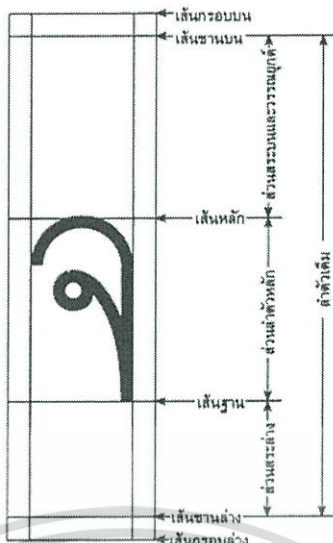
ประเภทตัวอักษร	สมาชิก
พยัญชนะ	ก ข ขล ค ฅ ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ศ ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศษ ส ห พ อ ฮ
สระ	อะ อา อี อึ อี้ อ้อ อุ อู โอ ไอ โอเอ แอ อุ ฤ ฦ
วรรณยุกต์	อํ ่ ้ ๊ ๋
อักขระพิเศษ	ั ็ ๋ ๑ ๒ ๓

2.1.1 โครงสร้างพื้นฐาน

ชุดตัวอักษรภาษาไทยในชุดเดียวกันมีโครงสร้างพื้นฐานของ สัด ส่วน ขนาดเส้น ช่องไฟ และรูปแบบตัวอักษร อยู่ในหลักเกณฑ์เดียวกัน ตัวอักษรภาษาไทยประกอบด้วย 3 ส่วน [12] คือ

- ส่วนสระบนและวรรณยุกต์ คือ ระยะระหว่างเส้นชานบนและเส้นหลัก
- ส่วนลำตัวหลัก คือ ระยะระหว่างเส้นหลักและเส้นฐาน
- ส่วนสระล่าง คือ ระยะระหว่างเส้นฐานกับเส้นชานล่าง ดังภาพที่ 2.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.1 ส่วนประกอบของตัวอักษรภาษาไทย

ในกระบวนการวิเคราะห์ขอบเขตวัตถุแบบ ก และแบบ ข จะมีการวิเคราะห์หาเส้นหลักและเส้นฐานของข้อความ เพื่อนำไปใช้ในการตรวจจับตัวอักษรภาษาไทยที่วางอยู่ในแนวเดียวกันให้ครบถ้วน และเพื่อให้ง่ายต่อการขยายกรอบล้อมรอบข้อความเพื่อตรวจจับสระบน สระล่าง และวรรณยุกต์ในกระบวนการวิเคราะห์ขอบเขตวัตถุแบบ ข

2.1.2 สัดส่วนต่าง ๆ ของตัวอักษร

1. ส่วนสูงของตัวอักษรภาษาไทย

- ส่วนสระบนและวรรณยุกต์ มีระยะ 7 ± 1 หน่วยอักษร
- ส่วนลำตัวหลัก มีระยะ 8 หน่วยอักษร
- ส่วนสระล่าง มีระยะ 4.5-5 หน่วยอักษร

2. ส่วนกว้างตัวอักษรภาษาไทย

- ตัวอักษรที่มีส่วนกว้างเท่ากับ 9 ± 1 หน่วยอักษร ได้แก่ ฉ
- ตัวอักษรที่มีส่วนกว้างเท่ากับ 8 ± 1 หน่วยอักษร ได้แก่ ฅ ฌ ฒ
- ตัวอักษรที่มีส่วนกว้างเท่ากับ 7 ± 1 หน่วยอักษร ได้แก่ จ
- ตัวอักษรที่มีส่วนกว้างเท่ากับ 6 ± 1 หน่วยอักษร ได้แก่ ฎ ฏ น บ ป พ ฟ ภ ษ พ สระอำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

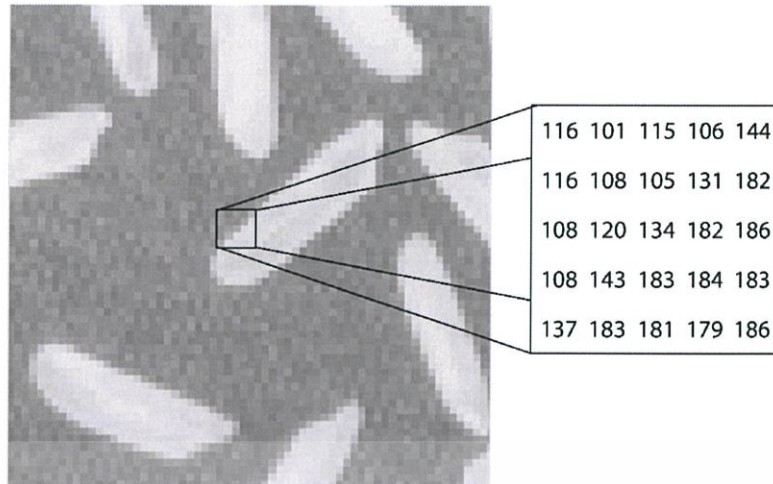
- ตัวอักษรที่มีส่วนกว้างเท่ากับ 5 ± 1 หน่วยอักษร ได้แก่ ก ค ฅ ฆ ท ต ถ ท ผ ฝ ม ย ฤ ล ศ ส ห อ ฮ โ ใ
- ตัวอักษรที่มีส่วนกว้างเท่ากับ 4 ± 1 หน่วยอักษร ได้แก่ ข ฃ ง จ ช ฌ ฐ ฌ ร ว ไม้หันอากาศหรือไม้ฝัด สระอา สระแอ สระอิ สระอี สระอือ สระอือ ไม้ยมก (๓) ไปยาลน้อย (๓)
- ตัวอักษรที่มีส่วนกว้างเท่ากับ 3 ± 1 หน่วยอักษร ได้แก่ สระอะ สระอุ ไม้โท ไม้ตรี ไม้ไต่คู้
- ตัวอักษรที่มีส่วนกว้างเท่ากับ 2 ± 1 หน่วยอักษร ได้แก่ ไม้จัตวา ทัณฑฆาต
- ตัวอักษรที่มีส่วนกว้างเท่ากับ 1-2 หน่วยอักษร ได้แก่ สระเอ สระอุ นิคหิต
- ตัวอักษรที่มีส่วนกว้างเท่ากับความหนาของเส้น-1 หน่วยอักษร ได้แก่ ไม้เอก

ในกระบวนการวิเคราะห์ขอบเขตวัตถุแบบ ข สัดส่วนต่าง ๆ เหล่านี้จะถูกนำไปสร้างเป็นกฎสำหรับตัดสินใจว่าวัตถุใดเป็นตัวอักษรภาษาไทย

2.2 การประมวลผลภาพดิจิทัล (Digital Image Processing)

2.2.1 ภาพระดับสีเทา (Intensity Image)

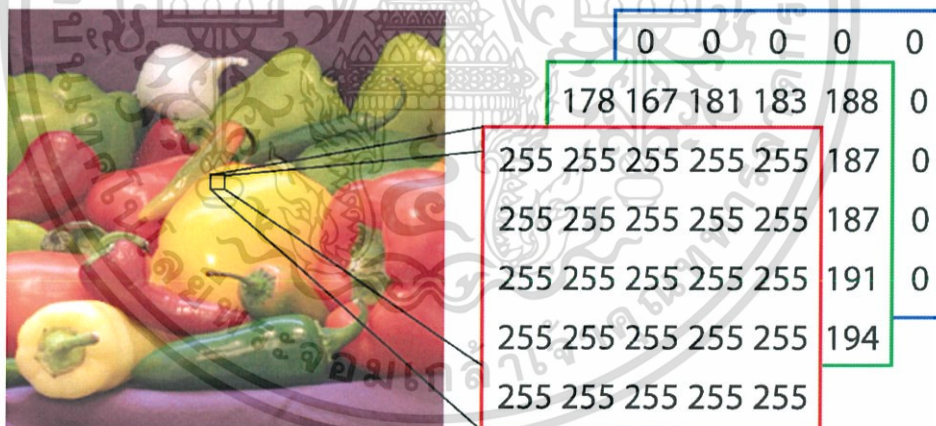
โดยปกติทั่ว ๆ ไปภาพระดับสีเทาจะมีขนาด 8 บิต สามารถแสดงค่าระดับสีเทาได้ 256 ระดับ และค่าที่เป็นไปได้คือ [0-255] เมื่อค่าความเข้มแสงมีค่าเป็น “0” จะหมายถึงพิกเซลนั้นมีความเข้มแสงต่ำที่สุดและจะทำให้พิกเซลนั้นเป็นสีดำ ในทางกลับกันหากค่าความเข้มแสงมีค่าเป็น “255” จะหมายถึงพิกเซลนั้นมีความเข้มแสงมากที่สุดและจะทำให้พิกเซลนั้นเป็นสีขาว อย่างไรก็ตาม ภาพระดับสีเทาอาจใช้จำนวนบิตมากกว่าหรือน้อยกว่า 8 บิตขึ้นอยู่กับความละเอียดของภาพที่ต้องการ [13], [14], [15] ภาพที่ 2.2 แสดงตัวอย่างภาพระดับสีเทาและค่าความเข้มแสงของแต่ละพิกเซล



ภาพที่ 2.2 ตัวอย่างภาพระดับสีเทาขนาด 8 บิต

2.2.2 ภาพสี (RGB Image)

กระบวนการจำแนกกลุ่มเส้นขอบจะพิจารณาจากค่าสีของพิกเซล โดยแต่ละพิกเซลของภาพสีจะประกอบด้วยระดับความเข้มแสง 3 ค่า คือ สีแดง สีเขียว และสีน้ำเงิน [13], [14], [15] ดังภาพที่ 2.3



ภาพที่ 2.3 ตัวอย่างภาพสีและค่าความเข้มแสงของแต่ละพิกเซล

2.2.3 การหาขอบภาพแบบแคนนี่ (Canny Edge Detection)

การหาขอบภาพแบบแคนนี่ [15] จะเริ่มจากการปรับภาพให้เรียบหรือกำจัดสัญญาณรบกวนในภาพโดยการคอนโวลูชันภาพต้นฉบับกับตัวกรองแบบเกาส์เซียน (Gaussian Filter) ดังสมการที่

2.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$S_{(i,j)} = G_{(i,j,\sigma)} * I_{(i,j)} \quad (2.1)$$

โดยที่ $S_{(i,j)}$ คือภาพที่ผ่านการกำจัดสัญญาณรบกวนแล้ว $G_{(i,j,\sigma)}$ คือตัวกรองแบบเกาส์เซียน และ $I_{(i,j)}$ คือภาพต้นฉบับ จากนั้นคำนวณหาการเปลี่ยนแปลงของค่าความเข้มแสงแต่ละพิกเซลในภาพ (Gradient) หรือค่าอนุพันธ์อันดับหนึ่งในทิศทางแกน x และ y ดังสมการที่ 2.2 และ 2.3 ตามลำดับ

$$P_{X(i,j)} \approx \frac{S_{(i,j+1)} - S_{(i,j)} + S_{(i+1,j+1)} - S_{(i+1,j)}}{2} \quad (2.2)$$

$$Q_{Y(i,j)} \approx \frac{S_{(i,j)} - S_{(i+1,j)} + S_{(i,j+1)} - S_{(i+1,j+1)}}{2} \quad (2.3)$$

นำค่าอนุพันธ์ที่ได้ไปคำนวณหาขนาด (Gradient Magnitude) และทิศทาง (Gradient Orientation) ของการเปลี่ยนแปลงค่าความเข้มแสงตามสมการที่ 2.4 และ 2.5 ตามลำดับ แล้วทำการค้นหาขอบภาพ โดยเลือกเฉพาะจุดที่ให้ค่าสูงสุดเฉพาะที่ (Local Maximum) เท่านั้น การหาขอบภาพโดยใช้อนุพันธ์อันดับหนึ่งนี้จะทำให้ได้ขอบที่บางเพียง 1 พิกเซล

$$M_{(i,j)} = \sqrt{P_{X(i,j)}^2 + Q_{Y(i,j)}^2} \quad (2.4)$$

$$\theta_{(i,j)} = \tan^{-1}\left(\frac{P_{X(i,j)}}{Q_{Y(i,j)}}\right) \quad (2.5)$$

อย่างไรก็ตาม ผลลัพธ์ที่ได้อาจมีเส้นขอบที่ไม่ใช่เส้นขอบที่แท้จริงปรากฏอยู่เนื่องมาจากสัญญาณรบกวนหรือลักษณะของวัตถุในภาพซึ่งพื้นผิวมีลวดลายหรือมีรายละเอียดภายในมาก ดังนั้น เพื่อลดปัญหาดังกล่าวจึงได้มีการกำหนดค่าอ้างอิง (Threshold: T) ขึ้นสองค่า คือ ค่าอ้างอิงสูง (High Threshold: T1) และ ค่าอ้างอิงต่ำ (Low Threshold: T2) หากพิกเซลใดมีค่ามากกว่า T1 แสดงว่าพิกเซลนั้นเป็นพิกเซลที่เป็นขอบ หากพิกเซลใดมีค่าน้อยกว่า T2 แสดงว่าพิกเซลนั้นไม่ใช่ขอบ และหากพิกเซลที่อยู่ติดกับพิกเซลที่เป็นข้อมีค่าอยู่ระหว่าง T1 และ T2 ก็จะถือว่าพิกเซลนั้นเป็นขอบด้วย กระบวนการนี้จะทำให้ขอบภาพมีความหนาหรือบาง

ในกระบวนการเตรียมภาพ ภาพฉากธรรมชาติซึ่งเป็นภาพสีจะถูกแปลงให้เป็นภาพระดับสีเทา แล้วทำการหาขอบเขตของวัตถุต่าง ๆ ที่ปรากฏอยู่ในภาพด้วยตัวดำเนินการหาขอบภาพ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบบแผนนี้ ซึ่งขอบเขตวัตถุที่ได้มานั้นจะถูกนำไปวิเคราะห์ว่าเป็นตัวอักษรภาษาไทยหรือไม่ในกระบวนการวิเคราะห์ขอบเขตวัตถุแบบ ก และ ข นอกจากนี้ ตำแหน่งของพิกเซลขอบ ค่าความเข้มแสงของพิกเซลในภาพระดับสีเทา และค่าความเข้มแสงของพิกเซลในภาพสี ก็จะถูกนำไปใช้ในขั้นตอนการสกัดคุณลักษณะเด่นในกระบวนการจำแนกกลุ่มเส้นขอบด้วย

2.3 การจำแนกกลุ่มข้อมูล (Data Clustering)

ขั้นตอนการจำแนกกลุ่มเส้นขอบเพื่อแยกตัวอักษรออกจากวัตถุอื่น ๆ ก่อนทำการวิเคราะห์และตัดสินใจว่าวัตถุใดเป็นตัวอักษรภาษาไทยจะใช้การจำแนกกลุ่มแบบ K-means [16] ซึ่งเป็นเทคนิคที่ใช้ในการจำแนกวัตถุจำนวน n วัตถุออกเป็น k กลุ่ม โดยที่ $k \leq n$ และมีเงื่อนไขดังนี้

1. แต่ละกลุ่มจะต้องประกอบด้วยวัตถุอย่างน้อย 1 วัตถุ
2. วัตถุแต่ละวัตถุจะต้องถูกจัดให้อยู่ในกลุ่มเพียงกลุ่มเดียวเท่านั้น

หลักการการทำงานของ K-means จะเริ่มจากการกำหนดค่า k หรือจำนวนกลุ่มที่ต้องการ จากนั้นสุ่มเลือกวัตถุเริ่มต้นจำนวน k ชุด ซึ่งแต่ละชุดที่ได้มาจะถูกกำหนดให้เป็นจุดศูนย์กลางเริ่มต้นของแต่ละกลุ่ม จากนั้นจำแนกวัตถุที่เหลือให้อยู่ในกลุ่มที่มีระยะห่างจากวัตถุถึงจุดศูนย์กลางของกลุ่มน้อยที่สุด แล้วคำนวณค่าเฉลี่ยเพื่อหาจุดศูนย์กลางใหม่ของกลุ่ม กระบวนการนี้จะถูกทำซ้ำจนกว่าค่าคลาดเคลื่อนกำลังสอง (Square Error) จะลู่เข้าหากันหรือมีค่าน้อยที่สุด ค่าคลาดเคลื่อนกำลังสองสามารถคำนวณได้จากสมการที่ 2.6

$$E = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2 \quad (2.6)$$

โดยที่ E คือ ผลรวมของค่าคลาดเคลื่อนกำลังสองของวัตถุทั้งหมดในชุดข้อมูล

p คือ ค่าข้อมูลของวัตถุ

m_i คือ ค่าเฉลี่ยของกลุ่ม C_i

เมื่อเสร็จสิ้นกระบวนการจำแนกกลุ่มเส้นขอบ วัตถุต่าง ๆ ในภาพจะถูกจำแนกออกเป็น k กลุ่ม โดยวัตถุที่มีสีพื้นหลัง สีของวัตถุ และตำแหน่งใกล้เคียงกันจะถูกจำแนกให้อยู่กลุ่มเดียวกัน

2.4 สมการเส้นตรง (Linear Equation)

ในกระบวนการวิเคราะห์ขอบเขตของวัตถุแบบ ก และ ข จะพิจารณาว่าวัตถุแต่ละวัตถุวางอยู่ในแนวเดียวกันหรือไม่จากค่าความชันระหว่างขอบเขตของวัตถุ 2 วัตถุ ซึ่งค่าความชันสามารถคำนวณได้จากสมการเส้นตรงแบบจุดและความชัน (Point-Slope Form) ดังสมการที่ 2.7

$$m = \frac{y-y_1}{x-x_1} \quad (2.7)$$

โดยที่ m คือ ค่าความชันของเส้นตรงที่ผ่านจุด (x,y) และ (x_1,y_1) [17]

2.5 งานวิจัยที่เกี่ยวข้อง

2.5.1 การตรวจจับข้อความโดยพิจารณาจากพื้นที่

Yu Zhong และคณะ [1] และ Julinda Gllavata และคณะ [2] นำการแปลงภาพ เช่น การแปลงโคไซน์แบบไม่ต่อเนื่อง (Discrete Cosine Transform: DCT) และการแปลงเวฟเล็ต (Wavelet Transform) มาใช้ในการสกัดคุณลักษณะเด่น แล้วใช้ค่าอ้างอิงเป็นเกณฑ์ในการคัดกรองพื้นที่ส่วนย่อยที่ไม่ใช่ข้อความออก สุดท้ายจึงนำพื้นที่ส่วนย่อยที่เหลืออยู่รวมเข้าด้วยกันโดยพิจารณาจากความสัมพันธ์เชิงพื้นที่ วิธีข้างต้นนี้สามารถตรวจจับเฉพาะข้อความที่มีขนาดคงที่เท่านั้น

Michael R. Lyu และคณะ [3] นำเสนอวิธีการตรวจจับข้อความภาษาจีนและภาษาอังกฤษที่ปรากฏในวิดีโอ เริ่มจากการสกัดคุณลักษณะเด่นโดยการหาเส้นขอบด้วยตัวดำเนินการหาขอบภาพโซเบล (Sobel Edge Detector) แล้วใช้ค่าอ้างอิงเป็นเกณฑ์ในการคัดกรองเส้นขอบของแต่ละพื้นที่ส่วนย่อยที่ไม่ใช่ข้อความออก ให้เหลือเฉพาะพื้นที่ส่วนย่อยที่ทนทานต่อการเปลี่ยนแปลงของแสงเท่านั้น สุดท้ายพื้นที่ดังกล่าวจะถูกรวมเข้าด้วยกันเป็นข้อความ โดยใช้การวิเคราะห์และฉายภาพ (Projection Profile Analysis: PPA) แบบวนซ้ำ วิธีการนี้สามารถตรวจจับได้เฉพาะข้อความที่มีการวางแนวในแนวนอนปกติเท่านั้น จึงไม่เหมาะที่จะนำมาใช้ในการตรวจจับข้อความในฉากธรรมชาติ ซึ่งยากที่จะควบคุมให้ของแนวข้อความในภาพไม่เอียงหรือหมุนไปจากแนวปกติเลย แม้ว่าขณะถ่ายภาพจะใช้เส้นตาราง (Grid) ที่แสดงบนหน้าจอช่วยก็ตาม

Manolis Delakis และคณะ [4] นำเสนอวิธีการตรวจจับข้อความ โดยใช้เครือข่ายประสาทเทียมแบบคอนโวลูชันนอล (Convolutional Neural Networks) ซึ่งเครือข่ายประสาทเทียมจะทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หน้าที่เรียนรู้และสกัดคุณลักษณะเด่นของชุดข้อมูลฝึกฝน โดยอัตโนมัติ นอกจากนี้ เครื่องข่ายประสาทเทียมยังทำหน้าที่ระบุว่าเป็นพื้นที่บริเวณใดในภาพเป็นข้อความและคัดกรองพื้นที่ของข้อความที่ไม่สมบูรณ์และพื้นที่ที่ไม่ใช่ข้อความออก

2.5.2 การตรวจจับข้อความโดยพิจารณาจากส่วนประกอบที่เชื่อมต่อกัน

Rong-Chi Chang [5] นำเสนอวิธี Connected Component Labeling เพื่อคัดกรองวัตถุที่ไม่ต้องการออกโดยพิจารณาจากขนาดของภาพ ขนาดของวัตถุ และจำนวนพิกเซลเส้นขอบ วิธีการนี้สามารถตรวจจับตัวอักษรในภาพฉากธรรมชาติได้ครบถ้วน แต่ผลลัพธ์ที่ได้มีวัตถุอื่น ๆ ที่ไม่ใช่ตัวอักษรรวมอยู่ด้วยจำนวนมาก เนื่องจากวิธีนี้เป็นขั้นตอนการเตรียมภาพสำหรับกระบวนการรู้จำตัวอักษรเพื่อตัดสินใจว่าวัตถุที่ยังเหลืออยู่วัตถุใดเป็นตัวอักษรต่อไป ดังนั้น วิธี Connected Component Labeling จึงไม่เหมาะที่จะนำมาใช้ในการตรวจจับข้อความภาษาไทยในฉากธรรมชาติในวิทยานิพนธ์นี้ซึ่งไม่มีการรู้จำตัวอักษร

Chucui Yi และ Yingli Tian [6] นำเสนอวิธี Adjacent Character Grouping เพื่อตรวจจับข้อความที่ปรากฏในภาพฉากธรรมชาติโดยวัดจากความสูง ความกว้าง จุดศูนย์กลาง ขนาดพื้นที่ ตัวอักษร และลักษณะการเชื่อมต่อกันระหว่างวัตถุ หากมีวัตถุ 3 วัตถุอยู่ติดกันและมีคุณลักษณะตามเกณฑ์ที่กำหนดก็จะถูกตัดสินใจว่าเป็นกลุ่มของตัวอักษร จากนั้นจึงรวมกลุ่มของตัวอักษรเหล่านี้เข้าด้วยกันเป็นข้อความ วิธีการนี้สามารถตรวจจับข้อความภาษาอังกฤษในภาพฉากธรรมชาติได้อย่างมีประสิทธิภาพ แต่เมื่อนำมาใช้ในการตรวจจับข้อความภาษาไทยพบว่าไม่สามารถตรวจจับสระและวรรณยุกต์ภาษาไทยซึ่งมีลักษณะการวางแนวแตกต่างจากภาษาอังกฤษ และผลลัพธ์ที่ได้มีวัตถุอื่น ๆ ที่ไม่ใช่ตัวอักษรรวมอยู่ด้วยจำนวนมาก

วิทยา จิระจิตติเจริญ [7] นำเสนอวิธีการตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติ โดยเริ่มจากการทำให้ภาพเบลอลดสัญญาณรบกวน ปรับภาพให้คมชัด และตรวจจับเส้นขอบของวัตถุ จากนั้นตัดสินใจว่าวัตถุใดในภาพเป็นตัวอักษรด้วยกฎซึ่งขั้นตอนนี้ทำให้สามารถตรวจจับตัวอักษรภาษาไทยส่วนใหญ่ที่อยู่ระหว่างเส้นหลักและเส้นฐานได้ แล้วนำไปวิเคราะห์หาแนวข้อความ กรอบข้อความ และทำการขยายกรอบเพื่อให้ครอบคลุมสระและวรรณยุกต์ วิธีการนี้มีข้อดีคือ สามารถตรวจจับสระบน สระล่าง และวรรณยุกต์ได้ง่ายเนื่องจากทราบตำแหน่งของแนวข้อความ แต่การวิเคราะห์หาแนวข้อความทำได้ยากเนื่องจากตัวอักษรส่วนใหญ่ที่ตรวจจับได้ในขั้นตอนแรกมีตัวอักษรที่มีหางด้านบนหรือด้านล่างรวมอยู่ด้วย ดังนั้น ในวิทยานิพนธ์นี้จึงได้นำการวิเคราะห์หาแนวข้อความเข้ามาใช้เพื่อให้สามารถตรวจจับตัวอักษรภาษาไทยที่อยู่ในแนวเดียวกันได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ครบถ้วนมากยิ่งขึ้น และเพื่อให้ง่ายต่อการขยายกรอบล้อมรอบข้อความเพื่อตรวจจับสระบน สระล่าง และวรรณยุกต์ในกระบวนการวิเคราะห์ขอบเขตวัตถุแบบ x โดยในวิทยานิพนธ์นี้จะวิเคราะห์หาแนวของข้อความไปพร้อม ๆ กับการตัดสินใจโดยใช้กฎ

2.5.3 การแยกตัวอักษรออกจากวัตถุอื่น ๆ ก่อนทำการตรวจจับข้อความ

Manoj Kumar และคณะ [8] นำเสนอวิธีการ Multilayer Separation เพื่อแยกพิกเซลที่เป็นตัวอักษรออกจากพิกเซลพื้นหลังด้วยวิธีการควอนไทซ์ (Quantization) เพื่อลดจำนวนสีของภาพต้นฉบับ จากนั้นแปลงภาพที่ถูกลดจำนวนสีแล้วให้เป็นภาพระดับสีเทา แล้วจำแนกพิกเซลออกเป็นสามกลุ่มโดยพิจารณาจากค่าความเข้มแสงของแต่ละพิกเซล ข้อดีของการควอนไทซ์และแปลงภาพเป็นสีเทาคือ ช่วยลดปริมาณข้อมูลทำให้สามารถจำแนกได้ง่ายและรวดเร็ว แต่การลดจำนวนสีมีข้อเสีย คือ อาจทำให้ตัวอักษรและพื้นหลังมีค่าความเข้มแสงใกล้เคียงกันมากขึ้น และถูกจำแนกให้อยู่ในกลุ่มเดียวกัน ดังนั้น วิธีการ Multilayer Separation จึงไม่ถูกนำมาใช้ในวิทยานิพนธ์นี้

Chucui Yi และ Yingli Tian [9] นำเสนอวิธีการ Boundary Clustering เพื่อจำแนกกลุ่มวัตถุในภาพโดยพิจารณาเฉพาะบริเวณขอบของวัตถุ โดยผู้วิจัยเชื่อว่าข้อมูลช่วงที่มีความถี่สูงหรือขอบของวัตถุสามารถใช้ในการจำแนกกลุ่มได้ คุณลักษณะที่ใช้ในการจำแนกกลุ่ม ได้แก่ สีของวัตถุ สีของพื้นหลัง และค่าพิกัด y ของจุดกึ่งกลางของวัตถุ โดยการสกัดคุณลักษณะสีจะเริ่มจากการตรวจจับเส้นขอบของวัตถุต่าง ๆ ในภาพด้วยตัวดำเนินการหาขอบภาพแบบแคนนี่ จากนั้นเปรียบเทียบค่าความเข้มแสงของพิกเซลที่อยู่ล้อมรอบพิกเซลขอบแต่ละพิกเซลเพื่อหาคู่พิกเซลที่มีค่าความเข้มแสงต่างกันมากที่สุด ทำเช่นนี้ไปเรื่อย ๆ จนกว่าจะครบทุกพิกเซลขอบแล้วจึงคำนวณหาค่าความเข้มแสงสูงสุดเฉลี่ยและค่าความเข้มแสงต่ำสุดเฉลี่ย ได้เป็นคุณลักษณะสีของวัตถุ และสีของพื้นหลังที่จะนำไปใช้ในการจำแนกกลุ่ม วิธีการนี้สามารถแยกตัวอักษรออกจากวัตถุอื่น ๆ ได้ดี แม้ว่าสีของตัวอักษรจะคล้ายกับสีของพื้นหลัง แต่ต้องใช้เวลาในการประมวลผลนานเนื่องจากต้องพิจารณาค่าสีของทุก ๆ พิกเซลที่ล้อมรอบพิกเซลขอบทั้งหมดแล้วนำมาหาค่าเฉลี่ยเพื่อให้ได้ค่าสีที่เป็นตัวแทนของแต่ละวัตถุ ดังนั้น ในวิทยานิพนธ์นี้จะสกัดคุณลักษณะสีโดยพิจารณาเฉพาะพิกเซลที่สนใจเท่านั้นเพื่อลดระยะเวลาที่ใช้ในการประมวลผล

บทที่ 3

วิธีดำเนินการวิจัย

จากกรอบแนวคิดในการทำวิจัยซึ่งได้กล่าวไว้แล้วในบทที่ 1 ในบทนี้จะกล่าวถึงขั้นตอนวิธีการตรวจจับข้อความภาษาไทยในฉากธรรมชาติ ซึ่งประกอบด้วยกระบวนการเตรียมภาพ การจำแนกกลุ่มเส้นขอบเพื่อแยกวัตถุที่เป็นตัวอักษรออกจากวัตถุอื่น ๆ และการวิเคราะห์ขอบเขตของวัตถุเพื่อตัดสินใจว่าวัตถุใดเป็นข้อความภาษาไทย โดยในวิทยานิพนธ์นี้ได้นำเสนอวิธีการวิเคราะห์ขอบเขตของวัตถุ 2 วิธี คือ การวิเคราะห์ขอบเขตของวัตถุแบบ ก และการวิเคราะห์ขอบเขตของวัตถุแบบ ข

3.1 กระบวนการเตรียมภาพ

กระบวนการเตรียมภาพจะเริ่มจากการแปลงภาพสีให้เป็นภาพระดับสีเทา จากนั้นหาขอบเขตของวัตถุต่าง ๆ ที่ปรากฏอยู่ในภาพด้วยตัวดำเนินการหาขอบภาพแบบแคนนี่ ผลลัพธ์ที่ได้จากการหาขอบภาพจะเป็นภาพไบนารี ซึ่งพิกเซลขอบหรือพิกเซลสีขาวที่เชื่อมต่อกันในแต่ละชุดจะหมายถึงขอบเขตของแต่ละวัตถุ เมื่อได้วัตถุแล้วจะกำหนดเส้นกรอบสมมติ ซึ่งประกอบด้วย เส้นกรอบหน้า เส้นกรอบหลัง เส้นกรอบบน และเส้นกรอบล่าง ล้อมรอบแต่ละวัตถุโดยกำหนดให้ค่าพิกัด x ของเส้นกรอบหน้าและกรอบหลังมีค่าเท่ากับค่าพิกัด x ของพิกเซลขอบที่มีค่าน้อยที่สุดและมากที่สุดของวัตถุนั้นตามลำดับ และค่าพิกัด y ของเส้นกรอบบนและเส้นกรอบล่างมีค่าเท่ากับค่าพิกัด y ของพิกเซลขอบที่มีค่าน้อยที่สุดและมากที่สุดของวัตถุนั้นตามลำดับ ทั้งนี้ เส้นกรอบสมมติที่สร้างขึ้นจะถูกนำไปใช้ในขั้นตอนการวิเคราะห์ขอบเขตของวัตถุต่อไป กระบวนการเตรียมภาพแสดงได้ ดังภาพที่ 3.1



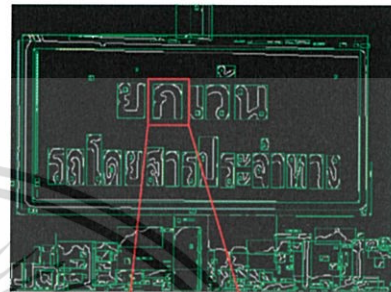
(ก)



(ข)



(ค)



(ง)



ภาพที่ 3.1 ผลลัพธ์ที่ได้จากระบวนการเตรียมภาพในแต่ละขั้นตอน (ก) ภาพต้นฉบับ (ข) ภาพระดับสีเทา (ค) ภาพไบนารีที่ได้จากการหาขอบภาพซึ่งประกอบด้วยชุดพิกเซลขอบที่เชื่อมต่อกัน (ง) เส้นกรอบสมมติที่สร้างขึ้นล้อมรอบพิกเซลขอบแต่ละชุด

3.2 กระบวนการจำแนกกลุ่มเส้นขอบ

ภาพฉากธรรมชาติส่วนใหญ่มักมีพื้นหลังที่ซับซ้อนเนื่องจากมีวัตถุที่ไม่ใช่ข้อความปรากฏอยู่ในรูปภาพจำนวนมากจึงทำให้การตรวจจับข้อความเป็นไปได้ยาก ดังนั้น การจำแนกวัตถุที่เป็นตัวอักษรออกจากวัตถุอื่นที่ไม่ใช่ตัวอักษร จึงเป็นวิธีที่จะช่วยลดความซับซ้อนของพื้นหลังของภาพฉากธรรมชาติและช่วยเพิ่มความถูกต้องในการวิเคราะห์ขอบเขตของวัตถุเพื่อตัดสินใจว่าวัตถุใดในภาพเป็นตัวอักษร เงื่อนไขที่ใช้ในการออกแบบกระบวนการจำแนกกลุ่มเส้นขอบมีดังนี้

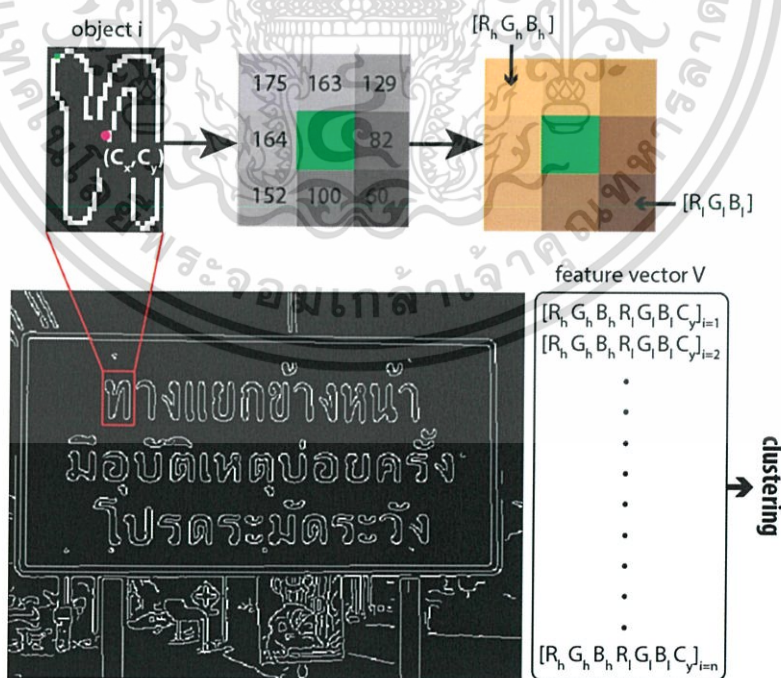
1. ถ้าเป็นตัวอักษรจะต้องมีสี่ของตัวอักษรเป็นสี่เหลี่ยมเพียงสี่เดียวและไม่มีลวดลาย
2. ถ้าเป็นตัวอักษรจะต้องมีสี่ของพื้นหลังเป็นสี่เหลี่ยมเพียงสี่เดียวและไม่มีลวดลาย
3. ถ้าเป็นข้อความภาษาไทยจะต้องมีการวางตัวในแนวนอน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การจำแนกกลุ่มเส้นขอบเป็นการสกัดคุณลักษณะเด่นของวัตถุในภาพแล้วนำคุณลักษณะเด่นที่สกัดได้ไปใช้ในการจำแนกกลุ่ม คุณลักษณะเด่นที่ต้องการสกัด ได้แก่ สีของวัตถุ สีของพื้นหลัง และตำแหน่งของวัตถุ ขั้นตอนการจำแนกกลุ่มเส้นขอบจะเริ่มจากสุ่มเลือกพิกเซลขอบมาจำนวน 1 พิกเซลต่อ 1 วัตถุ จากนั้นนำค่าความเข้มแสงในระดับภาพสีเทาของพิกเซลทั้งแปดที่อยู่ล้อมรอบพิกเซลที่เลือกไว้มาเปรียบเทียบกับเพื่อหาพิกเซลสองพิกเซลที่มีค่าความเข้มแสงแตกต่างกันมากที่สุด โดยจะเปรียบเทียบทีละคู่พิกเซลที่อยู่ตรงข้ามกันเท่านั้น พิกเซลที่ได้จะถูกนำไปหาค่าความเข้มแสงในระดับภาพสีเทากำหนดให้เป็นคุณลักษณะสีของวัตถุและสีของพื้นหลัง สมการที่ 3.1 แสดงเวกเตอร์คุณลักษณะเด่น V ของแต่ละวัตถุ

$$V_i = [R_h \ G_h \ B_h \ R_l \ G_l \ B_l \ C_y] \tag{3.1}$$

โดยที่ $[R_h, G_h, B_h]$ คือ ค่าสีแดง สีเขียว และสีน้ำเงินของพิกเซลที่มีค่าความเข้มแสงในระดับภาพสีเทาสูงที่สุด $[R_l, G_l, B_l]$ คือ ค่าสีแดง สีเขียว และสีน้ำเงินของพิกเซลที่มีค่าความเข้มแสงในระดับภาพสีเทาดำที่สุด และ $[C_y]$ คือ ค่าพิกัด y ของจุดศูนย์กลางของวัตถุในโดเมนตำแหน่ง กระบวนการสกัดคุณลักษณะแสดงได้ดังภาพที่ 3.2

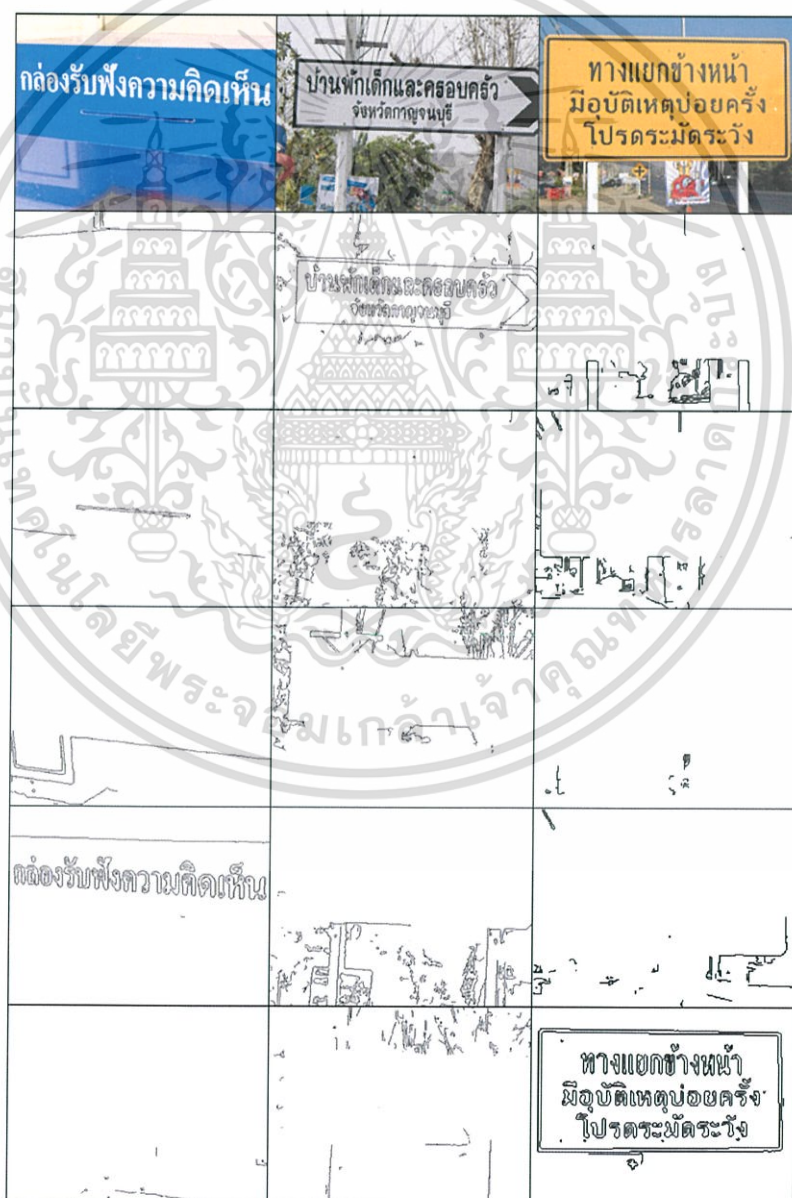


ภาพที่ 3.2 การสกัดคุณลักษณะเด่นของวัตถุ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เวกเตอร์คุณลักษณะเด่นที่สกัดได้ทั้งหมดจะถูกนำไปใช้ในการจำแนกกลุ่มด้วยอัลกอริทึม K-means ทั้งนี้ จำนวนกลุ่ม (k) เป็นปัจจัยที่สำคัญในการจำแนกกลุ่ม หากกำหนดค่า k น้อยเกินไป ก็จะทำให้ไม่สามารถจำแนกวัตถุที่เป็นตัวอักษรออกจากพื้นหลังที่ซับซ้อนได้ ในทางกลับกันหากกำหนดค่า k มากเกินไปก็อาจจะทำให้วัตถุที่เป็นตัวอักษรในข้อความเดียวกันถูกแยกไปอยู่คนละกลุ่ม ในวิทยานิพนธ์นี้กำหนดให้ $k = 5$ ซึ่งให้ผลลัพธ์ที่ดีที่สุดในการจำแนกกลุ่ม [9]

เมื่อเสร็จสิ้นกระบวนการจำแนกกลุ่มเส้นขอบ วัตถุต่าง ๆ ในภาพจะถูกจำแนกออกเป็น 5 กลุ่ม โดยวัตถุที่มีสีพื้นหลัง สีของวัตถุ และตำแหน่งใกล้เคียงกันจะถูกจำแนกให้อยู่กลุ่มเดียวกัน ตัวอย่างผลลัพธ์ที่ได้จากการจำแนกกลุ่มเส้นขอบแสดงได้ดังภาพที่ 3.3



ภาพที่ 3.3 ผลลัพธ์ที่ได้จากการจำแนกกลุ่มเส้นขอบด้วยวิธี K-means

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการศึกษาเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ไปเชิงพาณิชย์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 การวิเคราะห์ขอบเขตของวัตถุแบบ ก

เนื่องจากการวางแผนของข้อความภาษาไทยที่ปรากฏในฉากธรรมชาติส่วนใหญ่มักมีลักษณะคล้ายกัน กล่าวคือ ในแต่ละข้อความจะประกอบด้วยตัวอักษรตั้งแต่สองตัวขึ้นไปเรียงต่อกัน ในแนวนอน ดังนั้น ลักษณะเฉพาะของวัตถุต่าง ๆ ในภาพจึงสามารถนำมาใช้ตัดสินใจว่าวัตถุใด น่าจะเป็นตัวอักษรและวัตถุใดไม่ใช่ตัวอักษรได้

จากการวิเคราะห์ลักษณะของตัวอักษรภาษาไทยที่วางแผนอยู่ในส่วนลำตัวหลักพบว่า ตัวอักษรภาษาไทยสามารถจำแนกได้เป็น 3 กลุ่ม ตามลักษณะของหาง ได้แก่ ตัวอักษรที่ไม่มีหาง บนและหางล่าง ตัวอักษรที่มีเฉพาะหางบน และตัวอักษรที่มีเฉพาะหางล่าง ดังแสดงในตารางที่ 3.1 ซึ่งตัวอักษรส่วนใหญ่เป็นตัวอักษรที่ไม่มีทั้งหางบนและหางล่าง แสดงว่ากรณีที่ข้อความภาษาไทย จะประกอบด้วยตัวอักษรที่ไม่มีหางอย่างน้อยสองตัววางเรียงต่อกันมีโอกาสเกิดขึ้นได้มาก ดังนั้น หากสามารถตรวจจับตัวอักษรดังกล่าวได้ ก็จะทำให้ทราบตำแหน่งของเส้นหลักและเส้นฐานด้วย

ตารางที่ 3.1 สมาชิกตัวอักษรภาษาไทยที่วางแผนอยู่ในส่วนลำตัวหลัก จำแนกตามลักษณะของหาง

ลักษณะตัวอักษร	สมาชิก	จำนวน
ไม่มีหางบนและหางล่าง	ก ข ฃ ค ฅ ฆ ง จ ฉ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ผ พ ภ ม ย ร ล ว ย ห อ ะ อา เอ แอ ๆ	38
มีเฉพาะหางบน	ช ฌ ฬ ฟ ศ ส พ ฮ อ โ อ โ	12
มีเฉพาะหางล่าง	ฎ ฏ ฤ ฤ ๆ	5

กระบวนการวิเคราะห์ขอบเขตของวัตถุแบบ ก ประกอบด้วย 3 ขั้นตอนหลัก คือ การตรวจจับตัวอักษรภาษาไทยที่วางแผนอยู่ในส่วนลำตัวหลัก การตรวจจับสระบนและสระล่าง และการตรวจจับวรรณยุกต์

3.3.1 การวิเคราะห์เพื่อตรวจจับตัวอักษรภาษาไทยที่วางแผนอยู่ในส่วนลำตัวหลัก

การตรวจจับตัวอักษรภาษาไทยที่วางแผนอยู่ในส่วนลำตัวหลักจะเริ่มจากการหาวัตถุที่มีขนาดใกล้เคียงกันและเรียงต่อกันให้ได้มากที่สุด เพื่อระบุตำแหน่งเส้นหลักและเส้นฐานโดยพิจารณาจากตัวอักษรตัวแรกที่ตรวจจับได้ไปยังตัวอักษรตัวสุดท้าย จากนั้นวิเคราะห์วัตถุที่เหลือว่าเป็นตัวอักษรในแนวเดียวกันหรือไม่ โดยใช้เส้นหลักและเส้นฐานเป็นเกณฑ์ในการตัดสินใจ

จากการวิเคราะห์ลักษณะของตัวอักษรภาษาไทยที่ไม่มีหางและวางแนวอยู่ในส่วนลำตัวหลัก โดยพิจารณาจากตัวอย่างข้อมูลที่ใช้ในการทดลองพบว่า ตัวอักษรภาษาไทยที่เรียงต่อกันมีความสัมพันธ์กันดังนี้

1. ตัวอักษรในข้อความเดียวกันมีการวางแนวในแนวนอน โดยค่าพิกัด y ของจุดกึ่งกลางของตัวอักษรที่อยู่ด้านขวา (cy_j) จะต้องอยู่ระหว่างค่าพิกัด y ของกรอบบน (ty_i) กับกรอบล่าง (by_i) ของตัวอักษรด้านซ้าย ซึ่งสามารถสร้างเป็นกฎได้ดังสมการที่ 3.2

$$ty_i < cy_j < by_i \quad (3.2)$$

2. จุดกึ่งกลางของตัวอักษรที่อยู่ด้านซ้ายจะต้องไม่ซ้อนทับอยู่ในขอบเขตของตัวอักษรที่อยู่ด้านขวา โดยค่าพิกัด x ของจุดกึ่งกลางของตัวอักษรด้านซ้าย (cx_i) จะต้องมีย่านน้อยกว่าค่าพิกัด x ของเส้นกรอบหน้าของตัวอักษรที่อยู่ด้านขวา (lx_j) ดังสมการที่ 3.3

$$cx_i < lx_j \quad (3.3)$$

3. ตัวอักษรที่เรียงต่อกันจะต้องอยู่ใกล้กัน โดยในวิทยานิพนธ์นี้กำหนดให้ตัวอักษรที่เรียงต่อกันมีระยะห่างระหว่างตัวอักษรหรือช่องไฟ (d) น้อยกว่า 50% ของความกว้างของตัวอักษร ดังสมการที่ 3.4

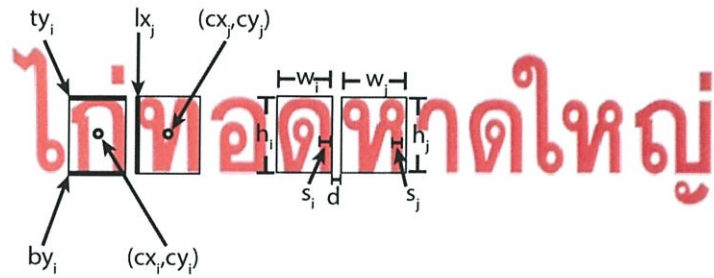
$$\frac{d}{w_i} < 0.5 \text{ และ } \frac{d}{w_j} < 0.5 \quad (3.4)$$

4. ตัวอักษรในข้อความเดียวกันมีขนาดใกล้เคียงกัน โดยในวิทยานิพนธ์นี้กำหนดให้ความกว้างของตัวอักษร (w) จะต้องมีค่าเท่ากันหรือมีค่ามากกว่าหรือเท่ากับ 80 % ของความกว้างของตัวอักษรที่เรียงต่อกัน และความสูงของตัวอักษร (h) จะต้องมีค่าเท่ากันหรือมีค่ามากกว่าหรือเท่ากับ 80% ของความสูงของตัวอักษรที่เรียงต่อกัน ดังสมการที่ 3.5 และ 3.6 ตามลำดับ

$$0.8 \leq \frac{w_i}{w_j} \leq 1 \quad (3.5)$$

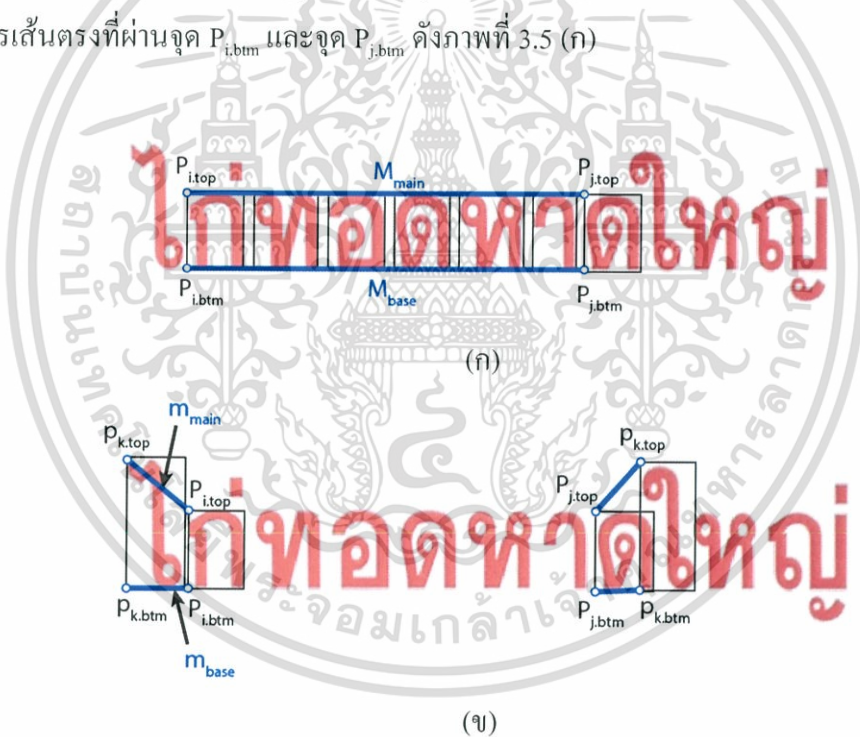
$$0.8 \leq \frac{h_i}{h_j} \leq 1 \quad (3.6)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.4 ตัวแปรที่ใช้ในการตัดสินใจว่าวัตถุใดเป็นตัวอักษรที่อยู่ระหว่างเส้นหลักกับเส้นฐาน

ภาพที่ 3.4 แสดงตัวอย่างข้อความภาษาไทยและตำแหน่งของตัวแปรต่าง ๆ ที่เกี่ยวข้อง ซึ่งหากวัตถุใดมีลักษณะเป็นไปตามกฎ (สมการที่ 3.2-3.6) แสดงว่าวัตถุนั้นน่าจะเป็นตัวอักษรภาษาไทยที่อยู่ระหว่างเส้นหลักกับเส้นฐาน จากนั้นจึงคำนวณหาค่าความชันของเส้นหลัก (M_{main}) จากสมการเส้นตรงที่ผ่านจุด $P_{i,top}$ และ $P_{j,top}$ และคำนวณหาค่าความชันของเส้นฐาน (M_{base}) จากสมการเส้นตรงที่ผ่านจุด $P_{i,btm}$ และจุด $P_{j,btm}$ ดังภาพที่ 3.5 (ก)



ภาพที่ 3.5 การคำนวณค่าความชันของ (ก) เส้นหลักและเส้นฐานของกลุ่มตัวอักษรที่ตรวจจับได้
(ข) เส้นหลักสมมติและเส้นฐานสมมติระหว่างกลุ่มตัวอักษรกับวัตถุใด ๆ

จากนั้น คำนวณหาค่าความชันระหว่างตัวอักษรภาษาไทยที่ตรวจจับได้กับวัตถุ k ใด ๆ โดยพิจารณาเงื่อนไข 2 ข้อ คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. หากวัตถุ k อยู่ด้านซ้ายของกลุ่มตัวอักษร จะคำนวณค่าความชันของเส้นหลักสมมติ (m_{main}) หรือเส้นตรงที่ผ่านจุด $p_{k,\text{top}}$ และ $P_{i,\text{top}}$ และคำนวณค่าความชันของเส้นฐานสมมติ (m_{base}) หรือเส้นตรงที่ผ่านจุด $p_{k,\text{btm}}$ และ $P_{i,\text{btm}}$

2. หากวัตถุ k อยู่ด้านขวาของกลุ่มตัวอักษร จะคำนวณค่าความชันของเส้นหลักสมมติ (m_{main}) หรือเส้นตรงที่ผ่านจุด $P_{j,\text{top}}$ และ $p_{k,\text{top}}$ และคำนวณค่าความชันของเส้นฐานสมมติ (m_{base}) หรือเส้นตรงที่ผ่านจุด $P_{j,\text{btm}}$ และ $p_{k,\text{btm}}$ ดังภาพที่ 3.5 (ข)

ค่าความชันของเส้นหลักสมมติและเส้นฐานสมมติที่คำนวณได้จะถูกนำไปเปรียบเทียบกับค่าความชันของเส้นหลักและเส้นฐานที่คำนวณไว้ก่อนหน้านี้ หากค่าความชันของเส้นหลักและเส้นหลักสมมติมีค่าใกล้เคียงกัน หรือค่าความชันของเส้นฐานและเส้นฐานสมมติมีค่าใกล้เคียงกัน ก็ถือว่าวัตถุนั้นเป็นตัวอักษรภาษาไทยที่อยู่ในแนวเดียวกัน โดยในวิทยานิพนธ์นี้กำหนดให้มีค่าผลต่างน้อยกว่า 0.05 ดังแสดงในสมการที่ 3.7 และ 3.8 ตามลำดับ ซึ่งขั้นตอนนี้จะทำให้สามารถตรวจจับตัวอักษรภาษาไทยบางส่วนที่มีลักษณะไม่เป็นไปตามกฎที่กำหนดไว้ในสมการที่ 3.2-3.6 ได้ครบถ้วน

$$|M_{\text{main}} - m_{\text{main}}| < 0.05 \quad (3.7)$$

$$|M_{\text{base}} - m_{\text{base}}| < 0.05 \quad (3.8)$$

3.3.2 การวิเคราะห์เพื่อตรวจจับสระบนและสระล่าง

เมื่อได้วัตถุที่คาดว่าเป็นตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนลำตัวหลักแล้ว วัตถุเหล่านี้จะถูกนำไปเปรียบเทียบกับวัตถุที่เหลือเพื่อตัดสินใจว่าวัตถุใดเป็นสระบนหรือสระล่าง จากการวิเคราะห์ลักษณะของตัวอักษรภาษาไทยพบว่า ตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนลำตัวหลักมีความสัมพันธ์กับสระบนและสระล่างดังนี้

1. ตำแหน่งของสระจะต้องอยู่ด้านบนหรือด้านล่างของตัวอักษรตัวใดตัวหนึ่ง โดยที่ค่าพิคัด y ของขอบด้านล่างของวัตถุที่เป็นสระบนจะต้องมีค่าน้อยกว่าค่าพิคัด y ของขอบด้านบนของตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนลำตัวหลัก ดังสมการที่ 3.9 และค่าพิคัด y ของขอบด้านบนของวัตถุที่เป็นสระล่างจะต้องมีค่ามากกว่าค่าพิคัด y ของขอบด้านล่างตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนลำตัวหลัก ดังสมการที่ 3.10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$ty_i > by_j \quad (3.9)$$

$$by_i < ty_j \quad (3.10)$$

นอกจากนี้ ค่าพิทัก x ของขอบด้านซ้ายหรือขวาของวัตถุที่เป็นสระจะต้องมีค่าอยู่ระหว่างขอบด้านซ้ายกับขอบด้านขวาของตัวอักษรที่ตรวจจับได้ ดังสมการที่ 3.11 และ 3.12

$$lx_i < lx_j < rx_i \quad (3.11)$$

$$lx_i < rx_j < rx_i \quad (3.12)$$

2. ระยะห่างระหว่างตัวอักษรที่ตรวจพบกับสระบนหรือสระล่าง (d) มีค่าน้อยกว่าหรือเท่ากับ 50% ของความสูงของตัวอักษรที่ตรวจจับได้ (h_i) ดังสมการที่ 3.13

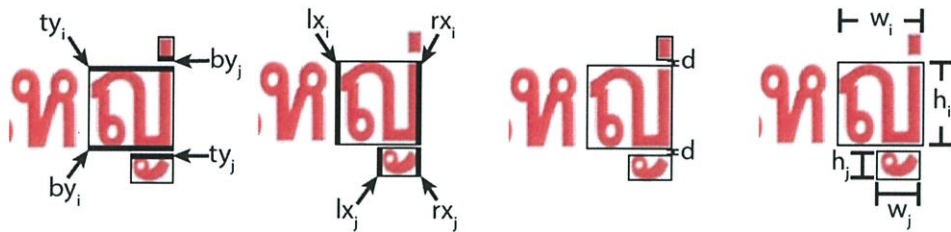
$$\frac{d}{h_i} \leq 0.5 \quad (3.13)$$

3. ความกว้างของสระ (w_j) มีค่าน้อยกว่าหรือเท่ากับสองเท่าของความกว้างของตัวอักษร (w_i) และความสูงของสระ (h_j) จะต้องมีค่าน้อยกว่าหรือเท่ากับความสูงของตัวอักษร (h_i) ดังสมการที่ 3.14 และ 3.15 ตามลำดับ

$$w_j \leq 2w_i \quad (3.14)$$

$$h_j \leq h_i \quad (3.15)$$

ภาพที่ 3.6 แสดงตัวอย่างข้อความภาษาไทยและตำแหน่งของตัวแปรต่าง ๆ ที่เกี่ยวข้อง โดยที่ i คือ ตัวอักษรที่ตรวจจับได้ในขั้นตอนการวิเคราะห์เพื่อตรวจจับตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนลำตัวหลัก และ j คือ วัตถุที่เป็นสระบนหรือสระล่าง



ภาพที่ 3.6 ตำแหน่งตัวแปรที่ใช้ในกฎสำหรับตัดสินใจว่าวัตถุใดเป็นสระบนหรือสระล่าง

3.3.3 การวิเคราะห์เพื่อตรวจจับวรรณยุกต์

สระบนที่ตรวจจับได้จะถูกนำไปเปรียบเทียบกับวัตถุที่เหลือเพื่อตัดสินใจว่าวัตถุใดเป็นวรรณยุกต์ ซึ่งจากการวิเคราะห์ลักษณะของตัวอักษรภาษาไทยพบว่า สระบนและวรรณยุกต์มีความสัมพันธ์กันในลักษณะเดียวกันกับตัวอักษรในส่วนลำตัวหลักและสระบน ดังนี้

1. ตำแหน่งของวรรณยุกต์จะต้องอยู่ด้านบนของสระบนตัวใดตัวหนึ่ง โดยที่ค่าพิคัด y ของขอบด้านล่างของวัตถุที่เป็นวรรณยุกต์ (by_j) จะต้องมีค่าน้อยกว่าค่าพิคัด y ของขอบด้านบนของสระบน (ty_i) ดังสมการที่ 3.9 และค่าพิคัด x ของขอบด้านซ้าย (lx_i) หรือขวา (rx_j) ของวัตถุที่เป็นวรรณยุกต์จะต้องอยู่ระหว่างขอบด้านซ้าย (lx_i) กับขอบด้านขวา (rx_j) ของวัตถุที่เป็นสระบน ดังสมการที่ 3.11 และ 3.12
2. ระยะห่างระหว่างวัตถุที่เป็นสระบนกับวัตถุที่เป็นวรรณยุกต์ (d) มีค่าน้อยกว่าหรือเท่ากับ 50% ของความสูงของวัตถุที่เป็นสระบน (h_i) ดังสมการที่ 3.13
3. ความกว้างของวรรณยุกต์ (w_j) มีค่าน้อยกว่าหรือเท่ากับสองเท่าของความกว้างของสระบน (w_i) ดังสมการที่ 3.14



ภาพที่ 3.7 ตัวอย่างผลลัพธ์ที่ได้จากการวิเคราะห์ขอบเขตของวัตถุแบบ ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ไปยังเว็บไซต์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการทดสอบประสิทธิภาพการตรวจจับข้อความภาษาไทยในฉากธรรมชาติด้วยวิธีวิเคราะห์ขอบเขตวัตถุแบบ ก พบว่าสามารถตรวจจับตัวอักษรภาษาไทยส่วนใหญ่ได้ครบถ้วน แต่ผลลัพธ์ที่ได้มีวัตถุที่ไม่ใช่ตัวอักษรรวมอยู่ด้วยเป็นจำนวนมาก ดังภาพที่ 3.7 เนื่องจากขั้นตอนการตรวจจับตัวอักษรที่วางแนวอยู่ในส่วนลำตัวหลักซึ่งเริ่มจากการหาวัตถุที่มีขนาดใกล้เคียงกันและวางอยู่ในแนวเดียวกันทีละคู่เรียงต่อกันไปเรื่อย ๆ โดยวัดจากตำแหน่งเส้นกรอบบน เส้นกรอบล่าง และจุดศูนย์กลางของวัตถุนั้น มีโอกาสมากที่จะตรวจจับได้วัตถุที่มีขนาดใกล้เคียงกันแต่ไม่ได้เรียงอยู่ในแนวเดียวกันทั้งหมด ทำให้ค่าความชันของเส้นหลักและเส้นฐานที่คำนวณจากตำแหน่งของตัวอักษรที่อยู่ด้านซ้ายสุดและตัวอักษรที่อยู่ด้านขวาสุดอาจไม่ใช่ค่าความชันที่แท้จริง ส่งผลให้การตรวจจับตัวอักษรภาษาไทยที่เหลือโดยเปรียบเทียบเฉพาะค่าความชันเกิดข้อผิดพลาดตามมา นอกจากนี้ค่าเกณฑ์ที่นำมาสร้างเป็นกฎอาจจะมีช่วงกว้างทำให้มีโอกาสตรวจจับได้วัตถุที่ไม่ใช่ตัวอักษรมากตามไปด้วย ปัญหาเหล่านี้สามารถแก้ไขได้โดยใช้การวิเคราะห์ขอบเขตของวัตถุแบบ ข ซึ่งจะกล่าวในหัวข้อต่อไป

3.4 การวิเคราะห์ขอบเขตของวัตถุแบบ ข

การวิเคราะห์ขอบเขตของวัตถุแบบ ข ประกอบด้วย 3 ขั้นตอนหลัก คือ การตรวจจับตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนลำตัวหลัก การรวมกลุ่มข้อความที่อยู่ในแนวเดียวกัน และการตรวจจับสระบน สระล่าง และวรรณยุกต์

3.4.1 การวิเคราะห์เพื่อตรวจจับตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนลำตัวหลัก

เริ่มจากการตรวจจับตัวอักษรภาษาไทยสองตัวที่ไม่มีหางและเรียงต่อกันเพื่อระบุตำแหน่งเส้นหลักและเส้นฐาน โดยใช้กฎที่สร้างขึ้นจากการสังเกตตัวอย่างข้อมูลที่ใช้ในการทดลองร่วมกับสัดส่วนมาตรฐานของโครงสร้างตัวอักษรภาษาไทย [12] จากนั้นวิเคราะห์วัตถุที่เหลือว่าเป็นตัวอักษรในแนวเดียวกันหรือไม่ โดยใช้เส้นหลักและเส้นฐานเป็นเกณฑ์ในการตัดสินใจโดยกฎที่นำมาใช้ในการตัดสินใจมีดังนี้

1. เมื่อพิจารณาสัดส่วนของตัวอักษรที่ไม่มีหางพบว่า ตัวอักษรที่มีส่วนกว้างน้อยที่สุดหรือ 1 หน่วยอักษร ได้แก่ สระเอ ตัวอักษรที่มีส่วนกว้างมากที่สุดหรือเท่ากับ 9 ± 1 หน่วยอักษร ได้แก่ ณ และสัดส่วนส่วนสูงของส่วนลำตัวหลักมีระยะ 8 หน่วยอักษร [12] ดังนั้น ช่วงค่าที่เป็นไปได้ของอัตราส่วนระหว่างส่วนกว้าง (w) กับส่วนสูง (h) ของตัวอักษรแต่ละตัวจึงแสดงได้ดังสมการที่ 3.16 และอัตราส่วนระหว่างส่วนกว้างของตัวอักษรสองตัวที่เรียงต่อกันแสดงได้ดังสมการที่ 3.17

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

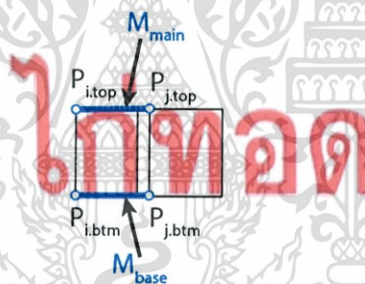
$$\frac{1}{8} \leq \frac{w_i}{h_i} \leq \frac{5}{4} \quad (3.16)$$

$$\frac{1}{10} \leq \frac{w_j}{w_j} \leq \frac{10}{1} \quad (3.17)$$

2. เส้นหลักและเส้นฐานระหว่างตัวอักษรสองตัวนี้จะต้องขนานกัน ดังนั้น ค่าความชันของเส้นหลัก (M_{main}) และเส้นฐาน (M_{base}) จะต้องมีค่าเท่ากันหรือใกล้เคียงกัน โดยในวิทยานิพนธ์นี้ กำหนดให้มีค่าผลต่างน้อยกว่า 0.05 ดังแสดงในสมการที่ 3.18

$$|M_{\text{main}} - M_{\text{base}}| < 0.05 \quad (3.18)$$

โดยค่าความชันของเส้นหลักคำนวณได้จากสมการเส้นตรงที่ผ่านจุด $P_{i,\text{top}}$ และ $P_{j,\text{top}}$ และค่าความชันของเส้นฐานคำนวณได้จากสมการเส้นตรงที่ผ่านจุด $P_{i,\text{btm}}$ และจุด $P_{j,\text{btm}}$ ดังภาพที่ 3.8

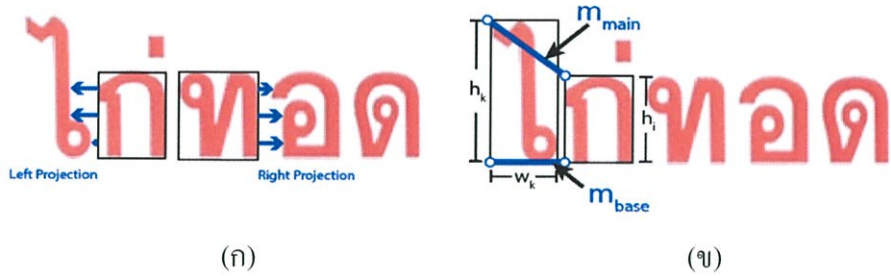


ภาพที่ 3.8 การคำนวณค่าความชันของเส้นหลักและเส้นฐานของตัวอักษรสองตัวที่ไม่มีหางและเรียงต่อกัน

นอกจากกฎที่กล่าวมาข้างต้นแล้ว สมการที่ 3.3 และ 3.4 เดิมที่ใช้ในขั้นตอนการวิเคราะห์ขอบเขตของวัตถุแบบ ก ก็จะถูกใช้เป็นเกณฑ์ในการตัดสินใจร่วมด้วย

จากสมมติฐานที่ว่าบริเวณข้อความและพื้นหลังมักจะมีค่าความเข้มแสงในภาพระดับสีเทาที่แตกต่างกันอย่างชัดเจน ดังนั้น การฉายภาพ (Projection) จึงถูกนำมาใช้เพื่อตัดสินใจว่าวัตถุอื่นที่เหลือเป็นตัวอักษรที่อยู่ในแนวเดียวกันหรือไม่ ถ้าตำแหน่งของจุดสิ้นสุดในการฉายภาพซ้อนทับอยู่ในเส้นกรอบของวัตถุใดก็ตาม วัตถุนั้นจะถูกนำมาพิจารณาว่าเป็นตัวอักษรหรือไม่โดยใช้กฎ และจะทำกระบวนการฉายภาพนี้ซ้ำจนกว่าจะไม่สามารถตรวจพบตัวอักษรตัวถัดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.9 การฉายภาพเพื่อตรวจจับตัวอักษรที่อยู่ในแนวเดียวกัน (ก) ทิศทางในการฉายภาพ (ข) ตำแหน่งตัวแปรที่ใช้ในกฎเพื่อตัดสินใจว่าวัตถุที่พบจากการฉายภาพเป็นตัวอักษรหรือไม่

ภาพที่ 3.9 (ก) แสดงทิศทางในการฉายภาพ โดยกำหนดให้กรอบหน้าของตัวอักษรที่ตรวจจับได้ที่อยู่ด้านซ้ายสุดเป็นจุดเริ่มต้นในการฉายภาพไปในทิศทางด้านซ้าย กรอบหลังของตัวอักษรที่ตรวจจับได้ที่อยู่ด้านขวาสุดเป็นจุดเริ่มต้นในการฉายภาพไปในทิศทางด้านขวา และกำหนดให้พิกเซลที่มีค่าความเข้มแสงแตกต่างจากพิกเซลก่อนหน้ามากกว่าเกณฑ์ที่กำหนดไว้เป็นจุดสิ้นสุดของการฉายภาพ วิทยานิพนธ์นี้กำหนดให้เกณฑ์มีค่าเท่ากับ 50

ภาพที่ 3.9 (ข) แสดงตัวแปรที่ใช้ในการตัดสินใจว่าวัตถุที่พบจากการฉายภาพเป็นตัวอักษรที่อยู่ในแนวเดียวกันหรือไม่ เกณฑ์ที่ใช้พิจารณามีดังนี้

1. เมื่อพิจารณาสัดส่วนของตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนลำตัวหลักพบว่า ตัวอักษรที่มีส่วนสูงน้อยที่สุดหรือ 2 หน่วยอักษร ได้แก่ สระอะ ตัวอักษรที่มีส่วนสูงมากที่สุดหรือเท่ากับ 13 หน่วยอักษร ได้แก่ โ โ โ [12] ดังนั้น ช่วงค่าที่เป็นไปได้ของอัตราส่วนระหว่างส่วนสูงของตัวอักษรที่ไม่มีหางที่ตรวจพบ (h_i) กับส่วนสูงของตัวอักษรที่กำลังสนใจ (h_k) จึงแสดงได้ดังสมการที่ 3.19

$$\frac{8}{13} \leq \frac{h_i}{h_k} \leq \frac{8}{2} \quad (3.19)$$

2. ช่วงค่าที่เป็นไปได้ของอัตราส่วนระหว่างส่วนกว้าง (w_k) กับส่วนสูง (h_k) ของตัวอักษรที่วางแนวอยู่ในส่วนลำตัวหลัก แสดงได้ดังสมการที่ 3.20

$$\frac{1}{13} \leq \frac{w_k}{h_k} \leq \frac{10}{2} \quad (3.20)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. หากตัวอักษรภาษาไทยอยู่ในแนวเดียวกัน ค่าความชันของเส้นหลัก (M_{main}) และเส้นหลักสมมติ (m_{main}) จะต้องมีค่าเท่ากันหรือใกล้เคียงกัน หรือค่าความชันของเส้นฐาน (M_{base}) และเส้นฐานสมมติ (m_{base}) จะต้องมีค่าเท่ากันหรือใกล้เคียงกัน โดยในวิทยานิพนธ์นี้กำหนดให้มีค่าผลต่างน้อยกว่า 0.05 ดังแสดงในสมการที่ 3.7 และ 3.8 ตามลำดับ

3.4.2 การรวมกลุ่มข้อความ

เนื่องจากกระบวนการจำแนกกลุ่มเส้นขอบมีโอกาสทำให้ตัวอักษรในข้อความเดียวกันถูกแยกออกไปอยู่คนละกลุ่ม โดยเฉพาะอย่างยิ่งในกรณีภาพฉากธรรมชาติที่มีพื้นหลังซับซ้อนน้อย ดังนั้น จึงมีการนำวัตถุในแต่ละกลุ่มกลับมารวมกันอีกครั้ง เพื่อวิเคราะห์ว่าข้อความที่ตรวจจับได้ในแต่ละกลุ่มเป็นข้อความเดียวกันหรือไม่ จากนั้นจึงทำการฉายภาพอีกครั้ง เพื่อให้สามารถตรวจจับตัวอักษรได้ครบถ้วน

จากการวิเคราะห์ลักษณะข้อความภาษาไทยพบว่ามีความสัมพันธ์กันดังนี้

1. หากข้อความทั้งสองเป็นข้อความเดียวกัน ค่าความชันของเส้นหลักของข้อความใดข้อความหนึ่ง (M_{main}) จะต้องมีค่าเท่ากันหรือใกล้เคียงกับค่าความชันของเส้นหลักสมมติระหว่างข้อความทั้งสอง (Mk_{main}) และค่าความชันของเส้นฐานของข้อความใดข้อความหนึ่ง (Mi_{base}) จะต้องมีค่าเท่ากับหรือใกล้เคียงกับค่าความชันของเส้นฐานสมมติระหว่างข้อความทั้งสอง (Mk_{base}) โดยในวิทยานิพนธ์นี้กำหนดให้มีค่าผลต่างน้อยกว่า 0.05 ดังแสดงในสมการที่ 3.21 และ 3.22 ตามลำดับ

$$|Mi_{\text{main}} - Mk_{\text{main}}| < 0.05 \quad (3.21)$$

$$|Mi_{\text{base}} - Mk_{\text{base}}| < 0.05 \quad (3.22)$$

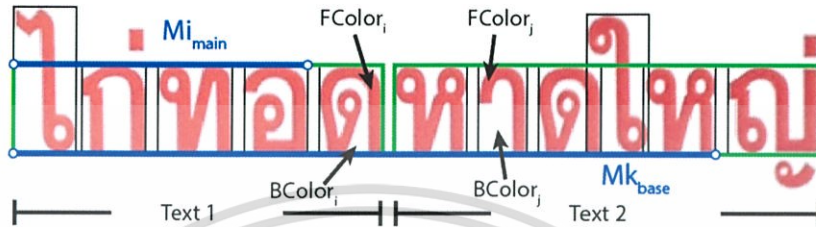
2. หากข้อความทั้งสองเป็นข้อความเดียวกัน ค่าเฉลี่ยของความเข้มแสงในระดับภาพสีของตัวอักษรของข้อความด้านซ้าย ($FColor_l$) จะต้องมีค่าใกล้เคียงกับข้อความด้านขวา ($FColor_r$) และค่าเฉลี่ยของความเข้มแสงในระดับภาพสีของพื้นหลังของข้อความด้านซ้าย ($BColor_l$) จะต้องมีค่าใกล้เคียงกับข้อความด้านขวา ($BColor_r$) โดยเปรียบเทียบค่าความเข้มแสงสีแดงกับสีแดง สีเขียวกับสีเขียว และสีน้ำเงินกับสีน้ำเงิน ทั้งนี้ กำหนดให้มีค่าผลต่างน้อยกว่าหรือเท่ากับ 50 ดังสมการที่ 3.23 และ 3.24 ตามลำดับ

$$|FColor_i - FColor_j| \leq 50 \quad (3.23)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาค้นคว้า ไม่นิพนธ์ให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$|BColor_i - BColor_j| \leq 50 \quad (3.24)$$

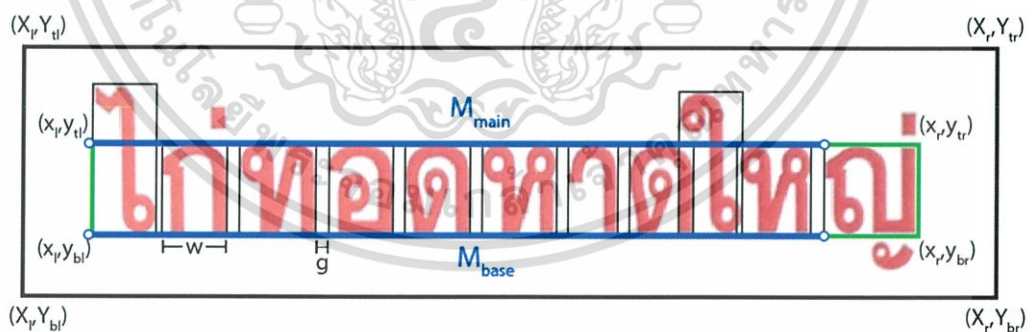
ตำแหน่งของตัวแปรที่ใช้ในกฎเพื่อตัดสินใจว่าข้อความสองข้อความเป็นข้อความเดียวกันหรือไม่ แสดงได้ดังภาพที่ 3.10



ภาพที่ 3.10 ตัวแปรที่ใช้ในการวิเคราะห์ว่าข้อความสองข้อความเป็นข้อความเดียวกันหรือไม่

3.4.3 การขยายเส้นกรอบเพื่อตรวจจับสระบน สระล่าง และวรรณยุกต์

ในขั้นตอนนี้ หากวัตถุที่เหลืออยู่วัตถุใดปรากฏอยู่ในบริเวณกรอบที่ขยายออกและมีสัดส่วนเป็นไปตามสมการที่ 3.25 วัตถุนั้นจะถูกตัดสินใจว่าเป็นตัวอักษร ภาพที่ 3.11 แสดงการขยายเส้นกรอบ โดยค่าพิสัยของมุมทั้งสี่ด้านกรอบข้อความที่ขยายออกสามารถคำนวณได้จากสมการที่ 3.26 - 3.31



ภาพที่ 3.11 การขยายเส้นกรอบเพื่อตรวจจับสระบน สระล่าง และวรรณยุกต์

$$\frac{1}{13} \leq \frac{w}{h} \leq \frac{5}{2} \quad (3.25)$$

$$X_l = x_l - (\bar{w} + \bar{g}) \quad (3.26)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$X_r = x_r + (\bar{w} + \bar{g}) \quad (3.27)$$

$$Y_{tl} = y_{tl} - M_{\text{main}}(x_l - X_l) - (y_{bl} - y_{tl}) \quad (3.28)$$

$$Y_{bl} = y_{bl} - M_{\text{base}}(x_l - X_l) + \frac{5}{8}(y_{bl} - y_{tl}) \quad (3.29)$$

$$Y_{tr} = M_{\text{main}}(X_r - x_r) + y_{tr} - (y_{br} - y_{tr}) \quad (3.30)$$

$$Y_{br} = M_{\text{base}}(X_r - x_r) + y_{br} + \frac{5}{8}(y_{br} - y_{tr}) \quad (3.31)$$

โดยที่	X_l	คือ ค่าพิกัด x ของเส้นรอบด้านซ้ายของกรอบที่ขยายออก
	X_r	คือ ค่าพิกัด x ของเส้นรอบด้านขวาของกรอบที่ขยายออก
	Y_{tl}	คือ ค่าพิกัด y ของมุมซ้ายบนของกรอบที่ขยายออก
	Y_{tr}	คือ ค่าพิกัด y ของมุมขวาบนของกรอบที่ขยายออก
	Y_{bl}	คือ ค่าพิกัด y ของมุมซ้ายล่างของกรอบที่ขยายออก
	Y_{br}	คือ ค่าพิกัด y ของมุมขวาล่างของกรอบที่ขยายออก
	x_l	คือ ค่าพิกัด x ของเส้นรอบด้านซ้ายของกรอบข้อความ
	x_r	คือ ค่าพิกัด x ของเส้นรอบด้านขวาของกรอบข้อความ
	y_{tl}	คือ ค่าพิกัด y ของมุมซ้ายบนของกรอบข้อความ
	y_{tr}	คือ ค่าพิกัด y ของมุมขวาบนของกรอบข้อความ
	y_{bl}	คือ ค่าพิกัด y ของมุมซ้ายล่างของกรอบข้อความ
	y_{br}	คือ ค่าพิกัด y ของมุมขวาล่างของกรอบข้อความ
	\bar{w}	คือ ส่วนกว้างเฉลี่ยของตัวอักษร
	\bar{g}	คือ ส่วนกว้างเฉลี่ยของระยะห่างระหว่างตัวอักษรแต่ละตัว
	M_{main}	คือ ค่าความชันของเส้นหลัก
	M_{base}	คือ ค่าความชันของเส้นฐาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 3.12 แสดงตัวอย่างผลลัพธ์ที่ได้จากการวิเคราะห์ขอบเขตวัตถุแบบ ข ซึ่งยังคงสามารถตรวจจับตัวอักษรภาษาไทยได้ครบถ้วน แต่วัตถุที่ไม่ใช่ตัวอักษรที่ตรวจจับได้มีจำนวนลดลง



ภาพที่ 3.12 ตัวอย่างผลลัพธ์ที่ได้จากการวิเคราะห์ขอบเขตของวัตถุแบบ ข

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการทดลอง

4.1 การวัดประสิทธิภาพ

การวัดประสิทธิภาพของกระบวนการตรวจจับข้อความภาษาไทยในวิทยานิพนธ์นี้จะวัดจากค่าความแม่นยำ (Precision: P) ค่าความครบถ้วน (Recall: R) และค่าความถูกต้องเฉลี่ย (F-measure: F) ซึ่งสามารถคำนวณได้จากสมการที่ 4.1-4.3 ตามลำดับ

$$P = \frac{TP}{TP + FP} \quad (4.1)$$

$$R = \frac{TP}{TP + FN} \quad (4.2)$$

$$F = \frac{2PR}{P + R} \quad (4.3)$$

โดยที่ TP คือ จำนวนขอบเขตวัตถุที่เป็นตัวอักษรและตรวจจับได้ว่าเป็นตัวอักษร
FP คือ จำนวนขอบเขตวัตถุที่ไม่ใช่ตัวอักษรแต่ตรวจจับได้ว่าเป็นตัวอักษร
FN คือ จำนวนขอบเขตวัตถุที่เป็นตัวอักษรแต่อัลกอริทึมไม่สามารถตรวจจับได้

ค่าความแม่นยำ คือ อัตราส่วนระหว่างจำนวนของขอบเขตตัวอักษรที่ตรวจจับได้สำเร็จต่อจำนวนขอบเขตวัตถุที่ตรวจจับได้ทั้งหมด ค่าความครบถ้วน คือ อัตราส่วนระหว่างจำนวนของขอบเขตตัวอักษรที่ตรวจจับได้สำเร็จต่อจำนวนขอบเขตตัวอักษรทั้งหมด และค่าความถูกต้องเฉลี่ย คือ ค่าเฉลี่ยของค่าความแม่นยำและค่าความครบถ้วน หากค่าความแม่นยำมีค่าน้อยแสดงว่าอัลกอริทึมตรวจจับได้วัตถุที่ไม่ใช่ตัวอักษรจำนวนมาก หากค่าความครบถ้วนมีค่าน้อยแสดงว่าอัลกอริทึมไม่สามารถตรวจจับตัวอักษรส่วนใหญ่ได้ และหากค่าความถูกต้องเฉลี่ยมีค่ามากแสดงว่าอัลกอริทึมมีประสิทธิภาพมาก

4.2 การออกแบบการทดลอง

เพื่อทดสอบประสิทธิภาพของอัลกอริทึมที่คิดค้นขึ้น ผู้วิจัยได้แบ่งการทดลองออกเป็น 3 การทดลอง คือ การทดลองที่ 1 มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพในการตรวจจับข้อความภาษาไทยในฉากธรรมชาติด้วยวิธี Adjacent Character Grouping [6] วิธีการวิเคราะห์ขอบเขตของวัตถุแบบ ก และวิธีการวิเคราะห์ขอบเขตของวัตถุแบบ ข โดยวัดจากค่าความแม่นยำและค่าความครบถ้วน การทดลองที่ 2 มีวัตถุประสงค์เพื่อทดสอบประสิทธิภาพในการตรวจจับข้อความภาษาไทยในฉากธรรมชาติด้วยวิธีการจำแนกกลุ่มเส้นขอบร่วมกับวิธีการวิเคราะห์ขอบเขตของวัตถุแบบ ข โดยวัดจากค่าความแม่นยำและค่าความครบถ้วน และการทดลองที่ 3 มีวัตถุประสงค์เพื่อทดสอบประสิทธิภาพของวิธีการจำแนกกลุ่มเส้นขอบโดยเปรียบเทียบระหว่างวิธี Boundary Clustering [9] และวิธีจำแนกกลุ่มเส้นขอบในวิทยานิพนธ์นี้ในด้านเวลาที่ใช้ในการประมวลผลและผลลัพธ์ที่ได้จากการตรวจจับข้อความภาษาไทย

4.3 การทดลองที่ 1

การทดลองที่ 1 เป็นการทดสอบประสิทธิภาพการตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยใช้วิธี Adjacent Character Grouping (ACG) [6] วิธีการวิเคราะห์ขอบเขตวัตถุแบบ ก (MCCA1) และวิธีการวิเคราะห์ขอบเขตวัตถุแบบ ข (MCCA2) ได้ผลการทดลองดังตารางที่ 4.1-4.3

ตารางที่ 4.1 ผลการตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยใช้วิธี ACG

กลุ่มที่	จำนวนตัวอักษรทั้งหมด	จำนวนขอบเขตวัตถุที่ตรวจจับได้		เวลาที่ใช้ในการประมวลผล (วินาที/ภาพ)	ประสิทธิภาพ		
		ตัวอักษร	ไม่ใช่ตัวอักษร		P	R	F
1	2,325	1,655	536	0.0158227	0.85	0.72	0.75
2	2,587	1,818	4,572	2.4205654	0.40	0.71	0.45
3	3,165	1,955	1,849	1.5088525	0.58	0.64	0.58

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 ผลการตรวจจับข้อความภาษาไทยในฉากรรรมชาติโดยใช้วิธี MCCA1

กลุ่ม ที่	จำนวน ตัวอักษร ทั้งหมด	จำนวนขอบเขตวัตถุ ที่ตรวจจับได้		เวลาที่ใช้ ในการประมวลผล (วินาที/ภาพ)	ประสิทธิภาพ		
		ตัวอักษร	ไม่ใช่ตัวอักษร		P	R	F
1	2,325	2,070	709	0.0766852	0.82	0.90	0.84
2	2,587	2,181	6,085	0.5319829	0.39	0.84	0.49
3	3,165	2,312	2,844	0.2135527	0.52	0.76	0.58

ตารางที่ 4.3 ผลการตรวจจับข้อความภาษาไทยในฉากรรรมชาติโดยใช้วิธี MCCA2

กลุ่ม ที่	จำนวน ตัวอักษร ทั้งหมด	จำนวนขอบเขตวัตถุ ที่ตรวจจับได้		เวลาที่ใช้ในการ ประมวลผล (วินาที/ภาพ)	ประสิทธิภาพ		
		ตัวอักษร	ไม่ใช่ตัวอักษร		P	R	F
1	2,325	2,029	114	0.3681617	0.95	0.90	0.92
2	2,587	2,100	439	1.2197364	0.82	0.83	0.80
3	3,165	2,107	660	0.6249197	0.78	0.70	0.71

เมื่อพิจารณาผลการทดสอบประสิทธิภาพในการตรวจจับข้อความภาษาไทยในฉากรรรมชาติด้วยวิธี ACG เปรียบเทียบกับวิธี MCCA1 พบว่า วิธี MCCA1 ได้ค่าความครบถ้วนเฉลี่ยสูงกว่าวิธี ACG อย่างชัดเจน โดยที่ค่าความแม่นยำเฉลี่ยของวิธี MCCA1 มีค่าลดลงเล็กน้อย ดังนั้น วิธี MCCA1 จึงสามารถแก้ปัญหากรณีที่ผลลัพธ์จากการตรวจจับข้อความภาษาไทยในฉากรรรมชาติด้วยวิธี ACG ไม่สามารถตรวจจับตัวอักษรภาษาไทยบางส่วน โดยเฉพาะอย่างยิ่งสระและวรรณยุกต์ได้ แต่ไม่สามารถแก้ปัญหากรณีที่ผลลัพธ์จากการตรวจจับข้อความมีวัตถุที่ไม่ใช่ตัวอักษรรวมอยู่ด้วยจำนวนมาก

เมื่อพิจารณาผลการทดสอบประสิทธิภาพในการตรวจจับข้อความภาษาไทยในฉากรรรมชาติด้วยวิธี ACG เปรียบเทียบกับวิธี MCCA2 พบว่า วิธี MCCA2 ได้ค่าความแม่นยำและค่าความครบถ้วนเฉลี่ยสูงกว่าวิธี ACG อย่างชัดเจน ดังนั้น วิธี MCCA2 จึงสามารถแก้ปัญหากรณีที่วิธี ACG ไม่สามารถตรวจจับสระและวรรณยุกต์ภาษาไทยได้ และสามารถ

แก้ปัญหากรณีที่ผลลัพธ์จากการตรวจจับข้อความมีวัตถุที่ไม่ใช่ตัวอักษรรวมอยู่ด้วยจำนวนมากเมื่อตรวจจับข้อความด้วยวิธี ACG ได้ด้วย

เมื่อพิจารณาผลการทดสอบประสิทธิภาพในการตรวจจับข้อความภาษาไทยในฉากรวมชาติด้วยวิธี MCCA1 เปรียบเทียบกับวิธี MCCA2 พบว่า วิธี MCCA2 ได้ค่าความแม่นยำเฉลี่ยสูงกว่าวิธี MCCA1 อย่างชัดเจน โดยที่ค่าความครบถ้วนเฉลี่ยของวิธี MCCA2 มีค่าลดลงเล็กน้อย ดังนั้น วิธี MCCA2 จึงสามารถแก้ปัญหากรณีที่ผลลัพธ์จากการตรวจจับข้อความมีวัตถุที่ไม่ใช่ตัวอักษรรวมอยู่ด้วยจำนวนมากที่เกิดขึ้นเมื่อตรวจจับข้อความด้วยวิธี MCCA1 ได้อย่างมีประสิทธิภาพ แต่เกณฑ์ที่ใช้ในการตัดสินใจว่าวัตถุใดเป็นตัวอักษรภาษาไทยของวิธี MCCA2 อาจไม่ครอบคลุมสัดส่วนของตัวอักษรภาษาไทยที่ปรากฏในฉากรวมชาติเท่ากับเกณฑ์ของวิธี MCCA1

ความสัมพันธ์แบบผกผันของค่าความแม่นยำและค่าความครบถ้วนในลักษณะนี้ถือเป็นธรรมชาติของข้อมูลอย่างหนึ่ง [18], [19] กล่าวคือ ความต้องการที่จะตรวจจับตัวอักษรให้ครบถ้วนมากที่สุดอาจส่งผลให้ผลลัพธ์ที่ได้มีวัตถุที่ไม่ใช่ตัวอักษรรวมอยู่ด้วยมากขึ้น และความต้องการที่จะกำจัดวัตถุที่ไม่ใช่ตัวอักษรให้ได้มากที่สุดก็อาจส่งผลให้ตรวจจับตัวอักษรได้ครบถ้วนน้อยลงเช่นกัน ดังนั้น การตัดสินใจว่าจะมุ่งเน้นที่ค่าความแม่นยำหรือค่าความครบถ้วนจึงขึ้นกับลักษณะของการนำผลลัพธ์ที่ได้ไปใช้งานต่อ หากต้องการให้การเรียนรู้จำตัวอักษรมีความถูกต้องสูงก็จะต้องเน้นที่ค่าความแม่นยำเป็นหลัก เพราะยังผลลัพธ์ที่ได้จากการตรวจจับข้อความมีวัตถุที่ไม่ใช่ตัวอักษรรวมอยู่ด้วยน้อยเท่าไร ก็ยังช่วยลดภาระการทำงานและเพิ่มอัตราความถูกต้องของการรู้จำตัวอักษรมากขึ้นเท่านั้น เช่นเดียวกันกับงานด้านการสืบค้นข้อมูล (Data Retrieval) และการจำแนกกลุ่มข้อความ (Text Classification) ที่ให้ความสำคัญกับค่าความแม่นยำมากกว่าค่าความครบถ้วน [20], [21]

ทั้งนี้ เนื่องจากค่าความถูกต้องเฉลี่ยของวิธี MCCA2 มีค่าสูงกว่าวิธี MCCA1 ในการทดลองที่ 2 และ 3 จึงใช้วิธี MCCA2 ในการตรวจจับข้อความภาษาไทยในฉากรวมชาติร่วมกับวิธีการจำแนกกลุ่มเส้นขอบและวิธี Boundary Clustering [9]

4.4 การทดลองที่ 2

การทดลองที่ 2 เป็นการทดสอบประสิทธิภาพของขั้นตอนวิธีตรวจจับข้อความภาษาไทยในฉากรวมชาติโดยใช้วิธีการจำแนกกลุ่มเส้นขอบ (FBC) ร่วมกับวิธีการวิเคราะห์ขอบเขตของวัตถุแบบ ข (MCCA2) ได้ผลการทดลองดังตารางที่ 4.4

ตารางที่ 4.4 ผลการตรวจจับข้อความภาษาไทยในฉากรวมชาติโดยใช้วิธี FBC ร่วมกับ MCCA2

กลุ่มที่	จำนวนตัวอักษรทั้งหมด	จำนวนขอบเขตวัตถุที่ตรวจจับได้		เวลาที่ใช้ในการประมวลผล (วินาที/ภาพ)		ประสิทธิภาพ		
		ตัวอักษร	ไม่ใช่ตัวอักษร	FBC	MCCA2	P	R	F
1	2,325	2,000	108	0.2915313	0.3372842	0.96	0.89	0.91
2	2,587	2,044	348	1.1191281	0.8867380	0.85	0.81	0.80
3	3,165	2,002	452	0.6369254	0.4949015	0.81	0.66	0.71

จากตารางที่ 4.3 และ 4.4 แสดงให้เห็นว่า การตรวจจับข้อความภาษาไทยด้วยวิธี FBC ร่วมกับ MCCA2 ได้ค่าความแม่นยำเฉลี่ยสูงกว่าการตรวจจับข้อความภาษาไทยด้วยวิธี MCCA2 เล็กน้อย แต่ก็ส่งผลให้ค่าความครบถ้วนเฉลี่ยที่ได้มีค่าลดลงเล็กน้อยเช่นกัน อย่างไรก็ตาม เมื่อพิจารณาจำนวนขอบเขตวัตถุที่ตรวจจับได้พบว่า จำนวนตัวอักษรที่ตรวจจับได้ที่ลดลงมีค่าน้อยกว่าจำนวนวัตถุที่ไม่ใช่ตัวอักษรที่สามารถกำจัดได้เมื่อทดสอบกับข้อมูลกลุ่มที่ 2 และ 3 ดังนั้น วิธี FBC ที่ $k = 5$ จึงเหมาะที่จะนำมาใช้ในกรณีที่พื้นหลังป้ายหรือฉากภาพมีความซับซ้อนเพื่อช่วยลดจำนวนวัตถุที่ไม่ใช่ตัวอักษรที่ปรากฏในผลลัพธ์ แต่ไม่เหมาะที่จะนำมาใช้ในกรณีที่พื้นหลังป้ายและฉากภาพไม่ซับซ้อน นอกจากนี้ วิธี FBC ยังช่วยให้การวิเคราะห์ขอบเขตวัตถุเพื่อตัดสินใจว่าวัตถุใดเป็นตัวอักษรภาษาไทยง่ายขึ้นดังจะเห็นได้จากเวลาประมวลผลเฉลี่ยของกระบวนการ MCCA2 ที่มีค่าลดลงหลังผ่านกระบวนการ FBC แล้ว ภาพที่ 4.1 แสดงตัวอย่างผลลัพธ์ที่จากการตรวจจับข้อความภาษาไทยด้วยวิธี FBC ร่วมกับ MCCA2



ภาพที่ 4.1 ตัวอย่างผลลัพธ์ที่ได้จากขั้นตอนวิธีตรวจจับข้อความภาษาไทยในวิทยานิพนธ์นี้
จำแนกตามชุดข้อมูล (ก) กลุ่มที่ 1 (ข) กลุ่มที่ 2 (ค) กลุ่มที่ 3

4.5 การทดลองที่ 3

ในการทดลองสุดท้าย วิธี Boundary Clustering (BC) [9] จะถูกนำมาเปรียบเทียบกับวิธีการจำแนกกลุ่มเส้นขอบ (FBC) ในวิทยานิพนธ์นี้ จากการทดสอบประสิทธิภาพของขั้นตอนวิธีตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยใช้วิธี BC ร่วมกับวิธีการวิเคราะห์ขอบเขตของวัตถุแบบ ข (MCCA2) ได้ผลการทดลองดังตารางที่ 4.5

ตารางที่ 4.5 ผลการตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยใช้วิธี BC ร่วมกับ MCCA2




กลุ่มที่	จำนวนตัวอักษรทั้งหมด	จำนวนขอบเขตวัตถุที่ตรวจจับได้		เวลาที่ใช้ในการประมวลผล (วินาที/ภาพ)		ประสิทธิภาพ		
		ตัวอักษร	ไม่ใช่ตัวอักษร	BC	MCCA2	P	R	F
1	2,325	2,041	123	31.0728250	0.3384369	0.95	0.90	0.92
2	2,587	2,087	415	59.6057200	0.9038761	0.83	0.82	0.80
3	3,165	2,087	644	46.4637300	0.5196165	0.78	0.69	0.71

จากตารางที่ 4.4 และ 4.5 พบว่า การตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยใช้วิธี FBC ร่วมกับ MCCA2 และวิธี BC ร่วมกับ MCCA2 ได้ค่าความแม่นยำเฉลี่ยและค่าความครบถ้วนเฉลี่ยใกล้เคียงกัน แสดงว่า วิธี FBC สามารถสกัดคุณลักษณะสีเพื่อใช้ในการจำแนกกลุ่มเส้นขอบได้ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มีประสิทธิภาพใกล้เคียงกับวิธี BC แต่วิธี FBC สามารถประมวลผลได้เร็วกว่าโดยใช้เวลาในการประมวลผลเฉลี่ย 0.68 วินาที/ภาพ ในขณะที่วิธี BC ใช้เวลาในการประมวลผลเฉลี่ย 45.71 วินาที/ภาพ

การตรวจจับข้อความในฉากธรรมชาตินั้นเป็นงานที่มีความท้าทายหลายประการอันเนื่องมาจากปัจจัยต่าง ๆ ที่อาจส่งผลให้การตรวจจับข้อความในฉากธรรมชาติเกิดความผิดพลาด กรณีที่ขั้นตอนวิธีในวิทยานิพนธ์นี้สามารถตรวจจับข้อความภาษาไทยได้อย่างถูกต้องและกรณีที่ไม่สามารถตรวจจับตัวอักษรภาษาไทยได้ แสดงได้ดังตารางที่ 4.6 และ 4.7 ตามลำดับ

ตารางที่ 4.6 ตัวอย่างผลลัพธ์ของการตรวจจับข้อความภาษาไทยที่ถูกต้องในกรณีต่าง ๆ

ที่	ภาพ	คำอธิบาย
1		พื้นหลังป้ายและฉากภาพเป็นพื้นเรียบ ไม่มีลวดลาย และมีค่าความเข้มแสงสม่ำเสมอ ตัวอักษรมีค่าความเข้มแสงแตกต่างจากพื้นหลังอย่างชัดเจน กระบวนการเตรียมภาพสามารถตัดแยกขอบเขตของตัวอักษรได้ถูกต้อง และตัวอักษรที่ตัดแยกได้มีสัดส่วนเป็นไปตามเกณฑ์ที่กำหนด
2		พื้นหลังป้ายเป็นพื้นเรียบ ไม่มีลวดลาย และมีค่าความเข้มแสงสม่ำเสมอ ตัวอักษรมีค่าความเข้มแสงแตกต่างจากพื้นหลังอย่างชัดเจน สามารถตัดแยกขอบเขตของตัวอักษรได้ถูกต้อง ตัวอักษรที่ตัดแยกได้มีสัดส่วนเป็นไปตามเกณฑ์ที่กำหนด และสามารถกำจัดวัตถุที่ไม่ใช่ตัวอักษรที่ปรากฏบริเวณฉากภาพได้อย่างมีประสิทธิภาพ
3		พื้นหลังป้ายมีความเข้มแสงไม่สม่ำเสมอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.6 (ต่อ)

ที่	ภาพ	คำอธิบาย
4		ป้ายที่ได้รับผลกระทบจากเงาและแสงสะท้อน

ตารางที่ 4.7 ตัวอย่างผลลัพธ์ของการตรวจจับตัวอักษรภาษาไทยที่ผิดพลาดในกรณีต่าง ๆ

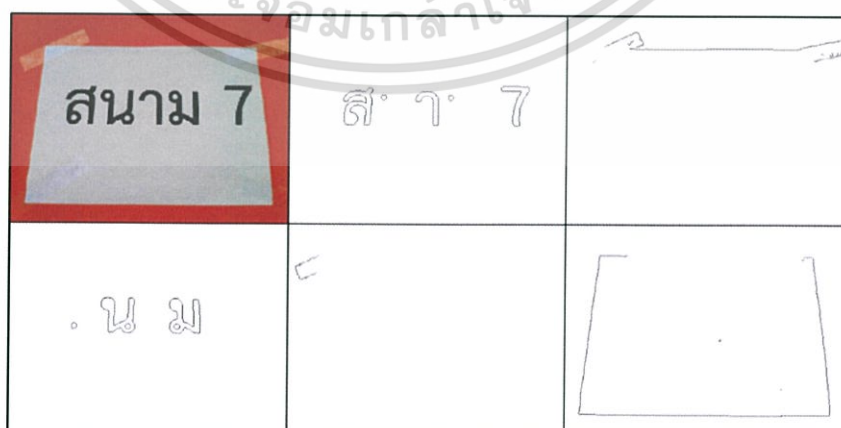
ที่	ภาพ	คำอธิบาย
1		ป้ายมีลักษณะ โปรงใสและมีวัตถุอื่น ๆ ปრაกฏเป็นฉากหลังของตัวอักษร ทำให้ไม่สามารถระบุขอบเขตวัตถุได้อย่างชัดเจนในขั้นตอนการเตรียมภาพ
2		ภาพสะท้อนที่พาดทับบริเวณป้ายซึ่งทำจากวัสดุมันวาว ทำให้ตัดแยกขอบเขตของวัตถุได้ไม่สมบูรณ์ และมีสัดส่วนไม่เป็นไปตามกฎที่กำหนดไว้
3		ตัวอักษร “ฟ” ถูกตัดแยกออกเป็น 2 วัตถุเนื่องจากบริเวณเส้นทแยงขึ้นลงของตัวอักษรนี้มีลักษณะบางกว่าบริเวณอื่น นอกจากนี้ ตัวอักษร “ฟ” และไม้หันอากาศอยู่ติดกัน ทำให้ขอบเขตวัตถุที่ตัดแยกได้ในขั้นตอนการเตรียมภาพมีสัดส่วนไม่เป็นไปตามกฎที่กำหนดไว้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.7 (ต่อ)

ที่	ภาพ	คำอธิบาย
4		เงาที่พาดผ่านบริเวณตัวอักษรทำให้ขอบเขตของตัวอักษรที่ตัดแยกได้ไม่สมบูรณ์ และมีสัดส่วนไม่เป็นไปตามกฎที่กำหนดไว้
5		ขอบเขตของ “สระอา” อยู่ติดกับขอบเขตของวัตถุอื่น ทำให้ขอบเขตวัตถุที่ตัดแยกได้ไม่สมบูรณ์ และมีสัดส่วนไม่เป็นไปตามกฎที่กำหนดไว้

นอกจากนี้ ยังมีกรณีที่การตรวจจับตัวอักษรภาษาไทยเกิดข้อผิดพลาดเนื่องจากวิธีการจำแนกกลุ่มเส้นขอบมีโอกาทำให้วัตถุที่เป็นตัวอักษรในข้อความเดียวกันถูกแยกไปอยู่คนละกลุ่มเมื่อกำหนดค่า k ไม่เหมาะสมกับลักษณะของข้อมูล ภาพที่ 4.2 แสดงกรณีที่ผลลัพธ์จากการจำแนกกลุ่มเส้นขอบที่ส่งผลให้ในขั้นตอนการวิเคราะห์ขอบเขตวัตถุแบบ x ไม่สามารถตรวจจับตัวอักษรภาษาไทยที่วางแนวอยู่ในส่วนคำตัวหลักได้ เพราะในแต่ละกลุ่มไม่มีวัตถุสองวัตถุที่เรียงต่อกันและมีลักษณะเป็นไปตามเกณฑ์ที่กำหนดไว้



ภาพที่ 4.2 ตัวอย่างผลลัพธ์จากการจำแนกกลุ่มเส้นขอบ

ที่ส่งผลให้ระบบไม่สามารถตรวจจับตัวอักษรภาษาไทยได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

วิทยานิพนธ์นี้มีวัตถุประสงค์เพื่อพัฒนาขั้นตอนวิธีการตรวจจับข้อความภาษาไทยในฉากธรรมชาติด้วยวิธีจำแนกกลุ่มเส้นขอบและวิธีวิเคราะห์ขอบเขตของวัตถุ โดยมุ่งเน้นที่การปรับปรุงขั้นตอนวิธี Boundary Clustering (BC) [9] ด้านเวลาที่ใช้ในการประมวลผล และมุ่งเน้นที่การพัฒนาขั้นตอนวิธี Adjacent Character Grouping (ACG) [6] ให้มีประสิทธิภาพมากยิ่งขึ้นเพื่อประยุกต์ใช้กับภาษาไทย โดยกระบวนการจำแนกกลุ่มเส้นขอบที่นำเสนอในวิทยานิพนธ์นี้จะสกัดคุณลักษณะเด่นของวัตถุเพื่อนำไปใช้ในการจำแนกกลุ่ม โดยพิจารณาเฉพาะค่าสีของพิกเซลที่สนใจเท่านั้น เพื่อให้สามารถประมวลผลได้อย่างรวดเร็ว และในกระบวนการวิเคราะห์ขอบเขตของวัตถุเพื่อตัดสินใจว่าวัตถุใดเป็นตัวอักษรภาษาไทย จะพิจารณาจากความสูง ความกว้าง จุดกึ่งกลาง และลักษณะการเชื่อมต่อกันระหว่างวัตถุไปพร้อม ๆ กับการวิเคราะห์หาแนวของข้อความเพื่อให้สามารถตรวจจับตัวอักษรภาษาไทยได้ถูกต้องและครบถ้วน โดยมีการกำหนดเกณฑ์ที่ใช้ในการตัดสินใจว่า วัตถุใดเป็นตัวอักษรให้สอดคล้องกับสัดส่วนและโครงสร้างของตัวอักษรภาษาไทย ซึ่งวิทยานิพนธ์นี้ได้นำเสนอวิธีการวิเคราะห์ขอบเขตของวัตถุ 2 แบบ คือ การวิเคราะห์ขอบเขตวัตถุแบบ ก และการวิเคราะห์ขอบเขตวัตถุแบบ ข

5.1.1 ประสิทธิภาพของวิธีการวิเคราะห์ขอบเขตของวัตถุแบบต่าง ๆ

ในการทดลองที่ 1 ผู้วิจัยได้ทำการเปรียบเทียบประสิทธิภาพในการตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยพิจารณาจากส่วนประกอบที่เชื่อมต่อกัน 3 วิธี ได้แก่ วิธี ACG วิธีการวิเคราะห์ขอบเขตของวัตถุแบบ ก และวิธีการวิเคราะห์ขอบเขตของวัตถุแบบ ข จากผลการทดสอบแสดงให้เห็นว่า วิธีการวิเคราะห์ขอบเขตของวัตถุแบบ ก และแบบ ข สามารถตรวจจับตัวอักษรภาษาไทยได้ครบถ้วนดีกว่าวิธี ACG โดยวิธีการวิเคราะห์ขอบเขตวัตถุแบบ ก และแบบ ข ได้ค่าครบถ้วนเฉลี่ยไม่แตกต่างกันมากนัก แต่เมื่อพิจารณาค่าความแม่นยำเฉลี่ยและค่าความถูกต้องเฉลี่ยพบว่า วิธีการวิเคราะห์ขอบเขตวัตถุแบบ ข ได้ค่าความแม่นยำเฉลี่ยและค่าความถูกต้องเฉลี่ยสูงกว่าวิธี ACG และวิธีการวิเคราะห์ขอบเขตวัตถุแบบ ก อย่างชัดเจน ดังนั้น วิธีการวิเคราะห์ขอบเขตวัตถุแบบ ข จึงเป็นวิธีการตรวจจับข้อความภาษาไทยที่มีประสิทธิภาพดีที่สุดใน และสามารถแก้ไข

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัญหาที่เกิดขึ้นเมื่อตรวจจับข้อความภาษาไทยในฉากรวมชาติด้วยวิธี ACG ได้ทั้ง 2 กรณี คือ 1) ไม่สามารถตรวจจับสระและวรรณยุกต์ภาษาไทยได้ และ 2) ผลลัพธ์จากการตรวจจับข้อความมีวัตถุที่ไม่ใช่ตัวอักษรรวมอยู่ด้วยจำนวนมาก

5.1.2 ประสิทธิภาพของการตรวจจับข้อความภาษาไทยในฉากรวมชาติด้วยวิธีการจำแนกกลุ่มเส้นขอบร่วมกับวิธีการวิเคราะห์ขอบเขตของวัตถุ

ในการทดลองที่ 2 ผู้วิจัยได้ทำการทดสอบประสิทธิภาพในการตรวจจับข้อความภาษาไทยในฉากรวมชาติด้วยวิธีการจำแนกกลุ่มเส้นขอบร่วมกับวิธีการวิเคราะห์ขอบเขตของวัตถุแบบ ข เพื่อเปรียบเทียบกับประสิทธิภาพในการตรวจจับข้อความภาษาไทยในฉากรวมชาติด้วยวิธีการวิเคราะห์ขอบเขตวัตถุแบบ ข เพียงอย่างเดียว จากการทดสอบแสดงให้เห็นว่า วิธีการจำแนกกลุ่มเส้นขอบช่วยจัดการข้อมูลให้อยู่ในรูปแบบที่ง่ายต่อการวิเคราะห์ว่าวัตถุใดเป็นตัวอักษรภาษาไทย และช่วยลดจำนวนวัตถุที่ไม่ใช่ตัวอักษรแต่ตรวจจับได้ว่าเป็นตัวอักษรได้ค้เมื่อมีการกำหนดจำนวนกลุ่ม (k) ให้เหมาะสมกับลักษณะของข้อมูล จากผลการทดสอบพบว่าได้ค่าความแม่นยำ ค่าความครบถ้วน และค่าความถูกต้องเฉลี่ยอยู่ที่ 0.87, 0.79 และ 0.81 ตามลำดับ ซึ่งการที่ค่าความแม่นยำสูงกว่าค่าความครบถ้วนนี้ มีลักษณะเหมือนกันกับผลการแข่งขัน ICDAR Robust Reading 2013 ด้านการตรวจจับข้อความในฉากรวมชาติ โดยผู้ชนะในการแข่งขันนี้ได้ค่าความแม่นยำ ค่าความครบถ้วน และค่าความถูกต้องเฉลี่ยอยู่ที่ 0.88, 0.66 และ 0.76 ตามลำดับ [22]

5.1.3 การเปรียบเทียบประสิทธิภาพของวิธีการจำแนกกลุ่มเส้นขอบกับวิธี BC

ในการทดลองที่ 3 ผู้วิจัยได้ทำการทดสอบประสิทธิภาพในการตรวจจับข้อความภาษาไทยในฉากรวมชาติด้วยวิธี BC ร่วมกับวิธีการวิเคราะห์ขอบเขตของวัตถุแบบ ข เพื่อเปรียบเทียบกับประสิทธิภาพในการตรวจจับข้อความภาษาไทยในฉากรวมชาติด้วยวิธีการจำแนกกลุ่มเส้นขอบร่วมกับวิธีการวิเคราะห์ขอบเขตของวัตถุแบบ ข จากการทดสอบแสดงให้เห็นว่า วิธีการจำแนกกลุ่มเส้นขอบสามารถแก้ไขปัญหาด้านเวลาที่ใช้ในการประมวลผลของวิธี BC ได้ โดยที่ประสิทธิภาพในการสกัดสกัดคุณลักษณะเด่นของวัตถุเพื่อนำไปใช้ในการจำแนกกลุ่มไม่ลดลง

จากการวิเคราะห์ผลลัพธ์ที่ได้จากการทดลองสามารถสรุปได้ว่า ขั้นตอนวิธีการตรวจจับข้อความภาษาไทยในฉากรวมชาติที่นำเสนอนี้จะสามารถตรวจจับข้อความภาษาไทยได้อย่างมีประสิทธิภาพก็ต่อเมื่อ กระบวนการเตรียมภาพสามารถระบุขอบเขตของตัวอักษรแต่ละตัวได้อย่างสมบูรณ์ ซึ่งปัจจัยที่อาจส่งผลให้ขอบเขตของตัวอักษรที่ตัดแยกได้ไม่สมบูรณ์ สามารถสรุปได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. แสงเงา ภาพสะท้อน หรือสิ่งสกปรกที่พาดทับบริเวณข้อความซึ่งทำให้ค่าความเข้มแสงของตัวอักษรไม่สม่ำเสมอหรือมีค่าแตกต่างกันมากกว่าค่าใดค่าหนึ่ง
2. ป้ายและข้อความที่มีสีซีดจาง ตลอดจนแสงเงา หรือภาพสะท้อนที่พาดทับบริเวณป้ายซึ่งทำให้ตัวอักษรและพื้นหลังมีค่าความเข้มแสงแตกต่างกันน้อยกว่าค่าใดค่าหนึ่ง
3. วัตถุที่ใช้ทำป้ายมีลักษณะโปร่งใสและมองเห็นวัตถุอื่น ๆ ที่อยู่ด้านหลังของป้าย
4. ข้อความที่มีระยะช่องไฟระหว่างตัวอักษรน้อยกว่าค่าใดค่าหนึ่ง
5. วัตถุใด ๆ ในภาพที่บดบังบริเวณของข้อความ
6. รูปแบบของตัวอักษรที่บริเวณเส้นทแยงขึ้นหรือลงมีลักษณะบางกว่าบริเวณอื่น

5.2 ข้อเสนอแนะ

1. เนื่องจากวิธีการจำแนกกลุ่มเส้นขอบมีโอกาสทำให้วัตถุที่เป็นตัวอักษรในข้อความเดียวกันถูกแยกไปอยู่คนละกลุ่มเมื่อกำหนดค่า k ไม่เหมาะสมกับข้อมูล ดังนั้น การนำวิธีจำแนกกลุ่มข้อมูลแบบอื่นที่มีความยืดหยุ่นกว่า K-means Clustering มาใช้ จึงอาจจะแก้ปัญหานี้ได้ เช่น วิธี Agglomerative Hierarchical Clustering [16] ซึ่งไม่จำเป็นต้องกำหนดจำนวนกลุ่มก่อน โดยการจำแนกกลุ่มด้วยวิธีนี้จะเริ่มจากการกำหนดจำนวนกลุ่มให้เท่ากับจำนวนข้อมูล จากนั้นรวมกลุ่มข้อมูล 2 กลุ่มเข้าด้วยกัน โดยวัดจากระยะห่างหรือความคล้าย แล้วพิจารณาข้อมูลตัวถัดไปว่าควรรวมอยู่ในกลุ่มที่มีอยู่แล้วหรือควรรวมกับข้อมูลอื่นได้เป็นกลุ่มใหม่ ทำเช่นนี้เรื่อย ๆ จนกระทั่งข้อมูลทุกตัวไม่มีการเปลี่ยนแปลงกลุ่มอีก
2. ตัวอักษรที่ปรากฏบนป้ายในฉากรถรรมชาติอาจมีสัดส่วนและโครงสร้างไม่เป็นไปตามกฎที่สร้างขึ้นอันเนื่องมาจากปัจจัยต่าง ๆ เช่น ความหลากหลายของรูปแบบตัวอักษร และการบีบย่อ หรือขยายตัวอักษรเพื่อให้สามารถแสดงผลบนป้ายได้ครบถ้วน ปัญหานี้อาจแก้ไขได้โดยนำการเรียนรู้ของเครื่อง (Machine Learning) มาช่วยสร้างกฎที่ใช้ในการตัดสินใจว่าวัตถุใดเป็นตัวอักษรให้มีความยืดหยุ่นและสามารถตรวจจับข้อความได้หลากหลายรูปแบบมากยิ่งขึ้น
3. กรณีตรวจจับตัวอักษรในข้อความได้ไม่ครบถ้วน อาจนำการรู้จำตัวอักษรและระบบช่วยสะกดคำเข้ามาประยุกต์ใช้เพื่อคาดคะเนตัวอักษรที่ขาดหายไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

- [1] Zhong, Y., Zhang H. J. and Jain, A. K. “Automatic Caption Localization in Compressed Video” **IEEE Trans. on Pattern Analysis and Machine Intelligence.** vol. 22, no. 4, 2000. pp. 385-392
- [2] Gllavata, J., Ewerth, R. and Freisleben, B. “Text Detection in Images Based on Unsupervised Classification of High-frequency Wavelet Coefficients” **Proc. 17th International Conference on Pattern Recognition.** 2004. pp. 425-428
- [3] Lyu, M. R., Song, J. and Cai, M. “A Comprehensive Method for Multilingual Video Text Detection, Localization and Extraction” **IEEE Trans. on Circuits and Systems for Video Technology.** vol. 15, no. 2, 2005. pp. 243-255
- [4] Delakis, M. and Garcia, C., “Text Detection with Convolutional Neural Networks” **Proc. International Conference on Computer Vision Theory and Applications.** 2008. pp. 290-294
- [5] Chang, R. “Intelligent Text Detection and Extraction from Natural Scene Images” **Proc. 15th North-East Asia Symposium on Nano, Information Technolgy and Reliability.** 2011. pp. 23-28
- [6] Yi, C. and Tian, Y. “Text String Detection From Natural Scenes by Structure-Based Partition and Grouping” **IEEE Trans. on Image Process.** vol. 20, no. 9, September 2011. pp. 2594-2605
- [7] วิทยา จิรัฎฐิติเจริญ. “การตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากรมชาติ” วิทยานิพนธ์ วิทยาศาสตร์มหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง. 2550.
- [8] Kumar, M., Kim, Y. C. and Lee, G. S. “Text Detection using Multilayer Separation in Real Scene Images” **Proc. 10th IEEE International Conference on Computer and Information Technology.** 2010. pp. 1413-1417

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง (ต่อ)

- [9] Yi, C. and Tian, Y. “Localizing Text in Scene Images by Boundary Clustering, Stroke Segmentation and String Fragment Classification” **IEEE Trans. on Image Process.** vol. 21, no. 9, 2012. pp. 4256-4268
- [10] Woraratpanya, K. and Boonchukusol, P. “Thai Text Detection from Natural Scenes by using Boundary Clustering and Connected Component Analysis” **Proc. 9th National Conference on Computing and Information Technology.** 2013. pp. 665-672
- [11] Woraratpanya, K., Boonchukusol, P., Kuroki, Y. and Kato, Y. “Improved Thai Text Detection from Natural Scenes” **Proc. 2013 International Conference on Information Technology and Electrical Engineering.** 2013. pp. 137-142
- [12] ราชบัณฑิตยสถาน. 2540. **มาตรฐานโครงสร้างตัวอักษรไทย ฉบับราชบัณฑิตยสถาน.** กรุงเทพฯ : อรุณการพิมพ์.
- [13] สมเกียรติ อุดมหารธนากุล. 2554. **การประมวลผลภาพดิจิทัลเบื้องต้น.** กรุงเทพฯ : สำนักพิมพ์ที่อป.
- [14] อรฉัตร จิตต์โสภักตร์. 2552. **ทฤษฎีการประมวลผลภาพดิจิทัล.** กรุงเทพฯ : สวงนกิจพริ้นท์ แอนด์ มีเดีย.
- [15] Gonzalez, R. C. and Woods, R. E. 2002. **Digital Image Processing (2nd Edition).** New Jersey : Prentice Hall.
- [16] Han, J. and Kamber, M. 2006. **Data Mining: Concepts and Techniques (2nd Edition).** San Francisco : Morgan Kaufmann Publishers.
- [17] “**Linear Equation.**” [Online]. Available : http://en.wikipedia.org/wiki/linear_equation. 2013.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง (ต่อ)

- [18] Alvarez, S. A. **“An Exact Analytical Relation Among Recall, Precision, and Classification Accuracy in Information Retrieval”** Department of Computer Science, Boston College. 2002.
- [19] Jizba, R. **“Measuring Search Effectiveness.”** [Online]. Available :
http://www.creighton.edu/fileadmin/user/HSL/docs/ref/Searching_-recall_Precision.pdf
 2014.
- [20] Riloff, E. and Lehnert, W. **“Information Extraction as a Basis for High-Precision Text Classification”** *ACM Trans. on Information System.* 1994. pp. 296-333
- [21] Takamura, H. **“Critical Issues and Information Security and Managing Risk”** **Speech presented at the 10th International Conference on Computing and Information Technology**, Phuket, Thailand. May 2014.
- [22] Karatzas, D., Shafait, F., Uchida, S., Iwamura, M., Gomez i Bigorda, L., Robles Mestre, S., Mas, J., Fernandez Mota, D., Almazan Almazan, J., de las Heras, L.-P. **“ICDAR 2013 Robust Reading Competition”** *Proc. 2013 12th International Conference on Document Analysis and Recognition.* 2013. pp. 1484-1493

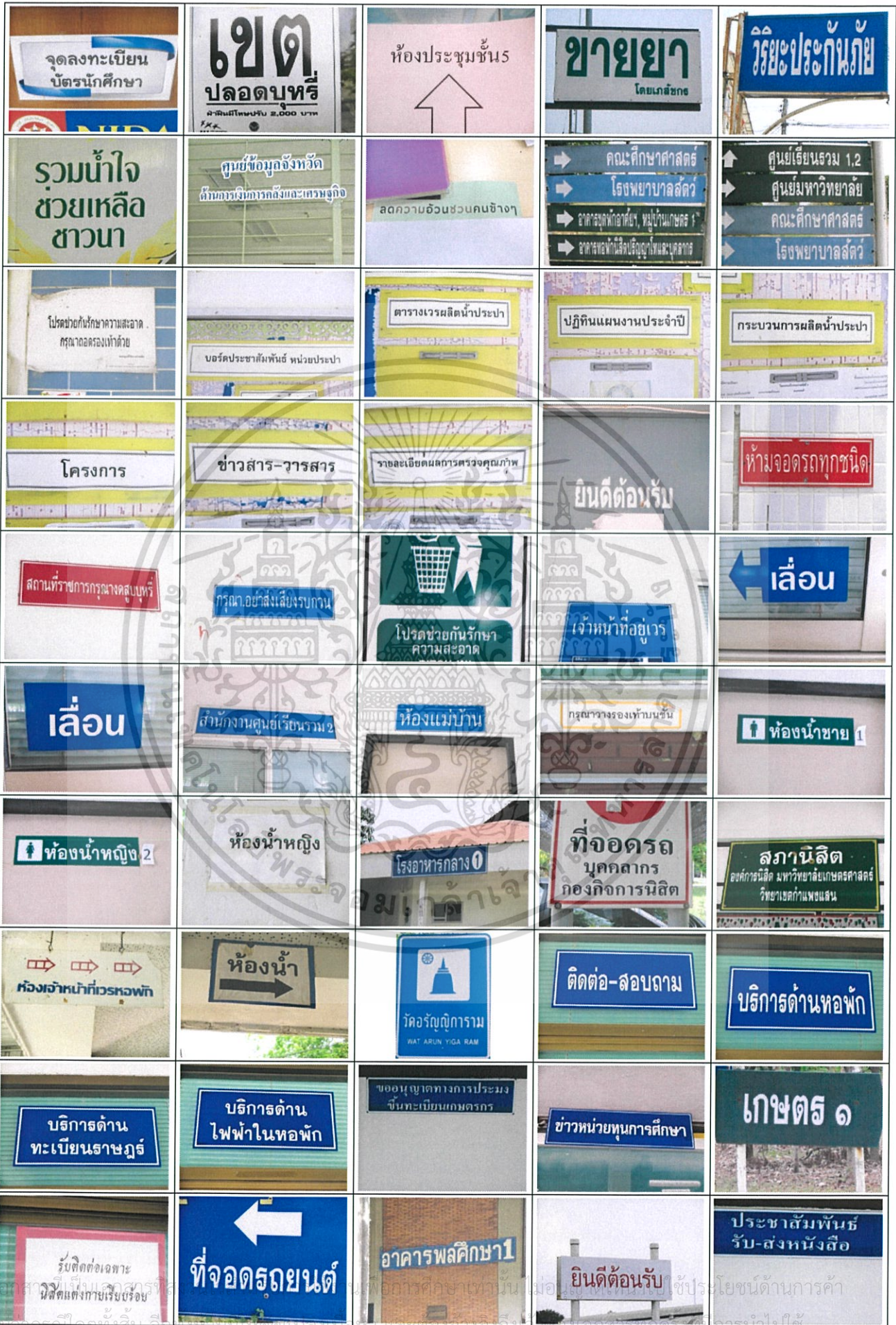


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

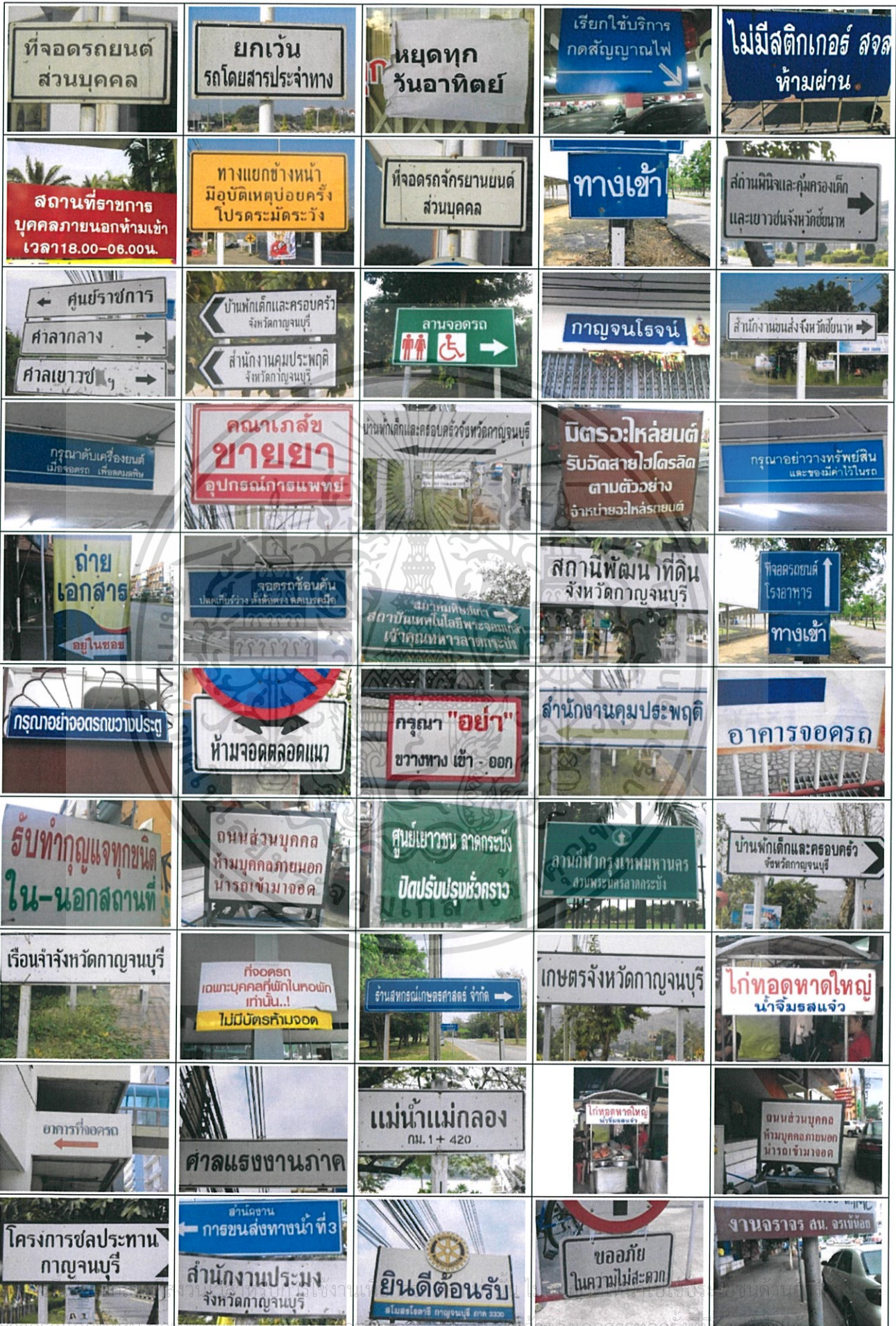


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้





สงวนลิขสิทธิ์ © 2015 โดย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี. การนำภาพไปใช้โดยไม่ได้รับอนุญาตถือว่าผิดกฎหมาย. การนำภาพไปใช้โดยไม่ได้รับอนุญาตถือว่าผิดกฎหมาย. การนำภาพไปใช้โดยไม่ได้รับอนุญาตถือว่าผิดกฎหมาย.





สำรวจการจราจรที่ถนนวิภาวดีรังสิต กรุงเทพมหานคร เมื่อวันที่ 15 กรกฎาคม 2564





เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



THE 9TH NATIONAL CONFERENCE ON
COMPUTING AND INFORMATION
TECHNOLOGY

PROCEEDINGS OF NCCIT 2013
THE 9TH NATIONAL CONFERENCE ON COMPUTING AND INFORMATION TECHNOLOGY

9TH-10TH MAY 2013

KING MONKUT'S UNIVERSITY OF TECHNOLOGY NORTH BANGKOK

WWW.NCCIT.NET

FACULTY OF INFORMATION TECHNOLOGY
KING MONKUT'S UNIVERSITY OF TECHNOLOGY NORTH BANGKOK

บทความวิจัย

การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ
ครั้งที่ 9

9-10 พฤษภาคม 2556

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ



FACULTY OF INFORMATION TECHNOLOGY

KING MONKUT'S UNIVERSITY OF TECHNOLOGY NORTH BANGKOK

WWW.IT.KMUTNB.AC.TH

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thursday May 9th, 2013		
NCCIT 2013 Room V		
Image Processing, Computer Vision, and Pattern Recognition		
Time	Title/Author	Page
13.00-13.20 NCCIT2013-183	Buddha Amulet Information Retrieval using Digital Images Combined with Feature Extraction and K-nearest Neighbor Techniques <i>Patompon Fugthong and Phayung Meesad</i>	621
13.20-13.40 NCCIT2013-165	Utilizing MHI for Human's Gesture with Repeating-path Trajectory <i>Pawinee Jaroonphan and Nongluck Covavisaruch</i>	627
13.40-14.00 NCCIT2013-191	Performance Comparison of Eye Tracking by Mean-Shift IPv4 and IPv6 and Lacus-Kanade Algorithms <i>Veerawan Janthanasub, Maleerat Sodanil, Phayung Meesad and Metha Sirigool</i>	633
14.00-14.20 NCCIT2013-204	Human Detection using Census Transform-Histogram of Oriented Gradients based feature and a Modified Version of the Extreme Learning Machine for Generalized Radial Basis Function Neural Networks <i>Chumpol Senapan and Khamron Sunat</i>	639
14.20-14.40 NCCIT2013-235	The Music Expression Recognition System using Case Based Reasoning <i>Watchariya Phonchareon and Maleerat Sodanil</i>	645
14.40-15.00 NCCIT2013-45	A Framework of Automatic Screening for Diabetic Retinopathy using Image Processing <i>Krisana Sriboran, Rapeepron Lordkeaw and Jantima Polpinij</i>	651
15.00-15.30	Coffee Break	
15.30-15.50 NCCIT2013-221	The Face Detection Optimization with Image Filtering Technique <i>Witsanukorn Thodthong, Pattanaphong Pothipassa and Khamron Sunat</i>	658
15.50-16.10 NCCIT2013-109	Thai Text Detection from Natural Scenes by Using Boundary Clustering and Connected Component Analysis <i>Kunpong Woraratpanya and Pimlak Boonchukusol</i>	665
16.10-16.30 NCCIT2013-158	Fast Searching and Counting Object for Coulter Plate Using Cross Scanner Technique <i>Riditchai Kanvijit and Jakkree Srinonchat</i>	673
16.30-16.50 NCCIT2013-79	Reversible Watermarking Algorithm Using Adaptive Embedding On Prediction And Sorting <i>Kitsada Somjit and Sirichai Rujipatthanapong</i>	678
16.50-17.10 NCCIT2013-34	Development of Computer Program based on Graphical User Interface (GUI) of SCILAB in Digital Image Processing for Under Graduate Student <i>Kitti Surpare and Meechai Lohakan</i>	684
17.10-17.30 NCCIT2013-200	Handwritten Character Segmentation on Buddhadasa Indapanno Documents <i>Wanwisa Chiwakulmongkhon and Tanasanee Phientrakul</i>	691
17.30-22.00	Welcome Dinner and Best Paper Award Presentation	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยการจำแนกกลุ่มเส้นขอบ และการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน Thai Text Detection from Natural Scenes by Using Boundary Clustering and Connected Component Analysis

กัณฑ์พงษ์ วรรัตน์ปัญญา (Kuntpong Woraratpanya)¹ และพิมพ์ลักษณ์ บุญชูกุลศุล (Pimlak Boonchukusol)²

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

¹kuntpong@it.kmitl.ac.th, ²pimlak.b@gmail.com

บทคัดย่อ

การตรวจจับข้อความในฉากธรรมชาติเป็นงานที่มีความท้าทายหลายประการ โดยเฉพาะอย่างยิ่งจากที่พื้นหลังมีความซับซ้อนจึงทำให้มีการพัฒนาวิธีการตรวจจับข้อความเกิดขึ้นมากมาย อย่างไรก็ตาม วิธีการเหล่านี้ไม่สามารถนำมาใช้กับการตรวจจับข้อความภาษาไทยได้ทั้งหมดเนื่องจากภาษาไทยมีลักษณะแตกต่างจากภาษาอื่น ๆ ดังนั้นงานวิจัยนี้จึงนำเสนอวิธีการตรวจจับข้อความภาษาไทยในฉากธรรมชาติซึ่งประกอบด้วย 2 ขั้นตอนหลักคือ 1) การจำแนกกลุ่มเส้นขอบโดยการวิเคราะห์ค่าสีและตำแหน่งของพิกเซลขอบเพื่อลดความซับซ้อนของพื้นหลัง 2) การวิเคราะห์ลักษณะการวางตัวของเส้นขอบหรือส่วนประกอบในแต่ละกลุ่มเพื่อตัดสินใจว่าส่วนประกอบใดเป็นตัวอักษรภาษาไทย ผลจากการทดสอบโดยใช้รูปภาพทั้งสิ้น 150 ภาพ ซึ่งประกอบด้วยตัวอักษรภาษาไทยจำนวน 5,555 ตัว พบว่าวิธีที่นำเสนอมีความแม่นยำและค่าเรียกกลับคืนเฉลี่ยเป็น 0.69 และ 0.88 ตามลำดับ

คำสำคัญ: การตรวจจับข้อความภาษาไทย ฉากธรรมชาติ การจำแนกกลุ่มเส้นขอบ การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน

Abstract

Text detection from natural scenes is a challenging task. Furthermore, the existing related works cannot detect Thai text completely, since Thai text layout has vowels and tonal marks that differ from other languages. This paper proposes an approach to detect Thai text from natural scenes. This approach consists of two main steps:

(i) clustering edge pixels based on color pairs and spatial positions in order to reduce the background complexity and (ii) connected component analysis to detect Thai characters. Based on 150 test images with 5,555 characters, the experimental results demonstrate that the proposed approach achieves the average of precision and recall 0.69 and 0.88, respectively.

Keyword: Thai text detection, natural scene, boundary clustering, connected component analysis

1. บทนำ

ปัจจุบันกล้องถ่ายภาพโดยเฉพาะอย่างยิ่งกล้องที่มาพร้อมกับอุปกรณ์พกพาต่าง ๆ นั้นได้ถูกนำมาใช้งานอย่างแพร่หลาย จึงส่งผลให้จำนวนรูปภาพที่ถ่ายจากฉากธรรมชาติดีปริมาณเพิ่มมากขึ้น ข้อความที่ปรากฏในรูปภาพจัดเป็นข้อมูลอย่างหนึ่ง เช่น คำอธิบาย คำเตือน ป้ายบอกทาง ป้ายชื่อสถานที่ ซึ่งหากมีระบบที่สามารถตรวจจับและประมวลผลข้อความที่ปรากฏในภาพฉากธรรมชาติได้ก็ทำให้ทราบข้อมูลที่อยู่ในรูปภาพเหล่านั้น

การตรวจจับข้อความในภาพฉากธรรมชาติเป็นปัญหาที่มีความท้าทายหลายประการเนื่องจากปัจจัยต่าง ๆ เช่น ความซับซ้อนของฉากภาพ ขนาดและรูปแบบของตัวอักษร สีและลวดลายของตัวอักษรและพื้นหลังป้าย ข้อความในภาพไม่ชัดเจนเนื่องจากตัวอักษรสีจางหรือภาพที่ถูกถ่ายในสภาพแสงน้อย แสงสะท้อนจากดวงอาทิตย์ หลอดไฟ หรือแสงแฟลชเงาของวัตถุอื่นที่พาดผ่านบริเวณป้าย ลักษณะการวางตัวของ

ข้อความที่มีหลากหลายรูปแบบ ตลอดจนข้อความที่มีคีย์เวิร์ด เนื่องจากป้ายไม้เรียบหรือมุมกล้องขณะถ่ายภาพ ปัจจัยเหล่านี้ ทำให้การตรวจจับข้อความในฉากธรรมชาติมีอัตราความผิดพลาดสูงกว่าการตรวจจับข้อความที่ได้จากการสแกนเอกสารด้วยเครื่องสแกนซึ่งมีการควบคุมปัจจัยแวดล้อมต่าง ๆ ที่จะมีผลต่อคุณภาพของภาพ [1]

อย่างไรก็ตาม งานวิจัยต่าง ๆ ที่เกี่ยวข้องกับการตรวจจับข้อความในภาพฉากธรรมชาติ นั้น ไม่สามารถนำมาใช้กับการตรวจจับตัวข้อความภาษาไทยได้ทั้งหมด เพราะภาษาไทยมีลักษณะแตกต่างจากภาษาอื่น เช่น การวางสระและวรรณยุกต์ ในตำแหน่งบนและล่างของตัวอักษร ดังนั้นจึงต้องมีการปรับปรุงกระบวนการให้เหมาะสมที่จะนำมาใช้ในการตรวจจับข้อความภาษาไทย

2. วรรณกรรมที่เกี่ยวข้อง

กระบวนการจำแนกกลุ่มเส้นขอบจะเริ่มจากการเตรียมภาพ โดยการแปลงเป็นภาพสีให้เป็นภาพระดับสีเทา (Intensity Image) โดยปกติทั่ว ๆ ไปภาพระดับสีเทามีขนาด 8 บิต สามารถแสดงค่าระดับสีเทาได้ทั้งหมด 256 ระดับ และค่าที่เป็นไปได้คือ 0-255 เมื่อค่าความเข้มแสงมีค่าเป็น 0 จะหมายถึงพิกเซลนั้นมีค่าความเข้มแสงต่ำและจะทำให้พิกเซลนั้นเป็นสีขาว ในทางกลับกัน หากค่าความเข้มแสงมีค่าเป็น 255 จะหมายถึงพิกเซลนั้นมีค่าความเข้มแสงมากและจะทำให้พิกเซลนั้นเป็นสีดำ [2], [3], [4] ซึ่งในการหาค่าความหนาของตัวอักษรเพื่อนำไปใช้วิเคราะห์ส่วนประกอบที่เชื่อมต่อกันในงานวิจัยนี้ก็จะพิจารณาจากค่าความเข้มแสงของพิกเซลที่เปลี่ยนแปลงไปในภาพระดับสีเทาเช่นกัน

2.1 การหาขอบภาพ

การหาขอบของวัตถุต่าง ๆ ในงานวิจัยนี้จะใช้ตัวดำเนินการหาขอบภาพแบบแคเนนนี่ (Canny Edge Detector) โดยจะเริ่มจากการขจัดสัญญาณรบกวนออกจากภาพด้วยตัวกรองแบบเกาส์เซียน (Gaussian Filter) จากนั้นจะคำนวณหาค่าอนุพันธ์หรือค่าการไล่ระดับหรือสี แล้วนำไปเปรียบเทียบกับค่า Threshold 2 ค่าที่ได้กำหนดไว้ คือ T1 และ T2 โดยที่ T1 > T2 หากพิกเซลใดมีค่ามากกว่า T1 แสดงว่าพิกเซลนั้นเป็นพิกเซลของเส้นขอบ นอกจากนี้ พิกเซลที่อยู่ติดกับพิกเซลขอบ ที่มีค่า

มากกว่า T2 ก็จะถือว่าเป็นพิกเซลขอบเช่นเดียวกัน [2] ตัวดำเนินการหาขอบภาพแบบแคเนนนี่มีข้อดีคือสามารถตรวจจับเส้นขอบของตัวอักษรที่มีสีใกล้เคียงกับพื้นหลังได้

2.2 การจำแนกกลุ่ม

ในการจำแนกกลุ่มเส้นขอบจะใช้ค่าสีของพิกเซลที่อยู่รอบพิกเซลขอบคู่ที่มีค่าสีแตกต่างกันมากที่สุดเป็นตัวแปรในการจำแนกกลุ่ม แต่ละพิกเซลของภาพสี (RGB Image) ประกอบด้วยค่าระดับความเข้มแสง 3 ค่า คือ สีแดง สีเขียว และสีน้ำเงิน [2], [3], [4] ดังนั้น ในการวิเคราะห์หาคู่พิกเซลใดมีค่าสีแตกต่างกันมากที่สุดจะต้องทำการแปลงค่าระดับความเข้มแสงทั้งสามค่าของแต่ละพิกเซลให้เป็นมาตรฐานเดียวกันก่อนจึงจะสามารถนำค่าสีของแต่ละคู่พิกเซลมาเปรียบเทียบกันได้

คะแนนมาตรฐาน (Standard Score or Z-Score) เป็นตัวเลขที่ไม่มีหน่วย ได้จากการวัดค่าของข้อมูลว่าห่างจากค่าเฉลี่ยเป็นกี่เท่าของค่าเบี่ยงเบนมาตรฐาน ดังสมการที่ (1)

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

โดยที่ z คือค่าคะแนนมาตรฐาน x คือค่าของข้อมูลที่ต้องการแปลงให้เป็นคะแนนมาตรฐาน μ คือค่าเฉลี่ยของข้อมูล และ σ คือค่าเบี่ยงเบนมาตรฐาน [5]

การจำแนกกลุ่มเส้นขอบในงานวิจัยนี้จะใช้การจำแนกกลุ่มแบบ k-means ซึ่งเป็นเทคนิคที่ใช้ในการแบ่งข้อมูลจำนวน n ตัวออกเป็น k กลุ่มโดยที่ $k \leq n$ และมีเงื่อนไขคือ 1) แต่ละกลุ่มจะต้องมีข้อมูลอย่างน้อย 1 ตัว และ 2) ข้อมูลแต่ละตัวจะต้องถูกจัดให้อยู่ในกลุ่มเพียงกลุ่มเดียวเท่านั้น

หลักการทำงานของ k-means จะเริ่มจากการสุ่มเลือกข้อมูลจำนวน k ค่าและกำหนดให้เป็นค่าเฉลี่ยหรือจุดศูนย์กลางเริ่มต้นของแต่ละกลุ่ม ข้อมูลที่เหลือจะถูกจัดเข้าให้อยู่ในกลุ่มที่มีค่าใกล้เคียงกันมากที่สุดโดยคำนวณจากระยะห่างระหว่างข้อมูลและค่าเฉลี่ยของกลุ่ม จากนั้นจะคำนวณหาค่าเฉลี่ยหรือจุดศูนย์กลางของแต่ละกลุ่มใหม่ กระบวนการนี้จะถูกทำซ้ำจนกว่าค่าเฉลี่ยของแต่ละกลุ่มจะไม่มีเปลี่ยนแปลง [6]

2.3 งานวิจัยที่เกี่ยวข้อง

ในปี ค.ศ. 2012 Chucai Yi และ Yingli Tian ได้ออกแบบวิธีจำแนกกลุ่มคู่สีบนพื้นฐานของ Gaussian Mixture Model (GMM) และ Expectation-Maximization (EM) โดย

พิจารณาจากค่าเฉลี่ยของสีและตำแหน่งของพิกเซลทุกพิกเซลที่เรียงต่อกันเป็นเส้นขอบของวัตถุ ค่าเฉลี่ยของสีได้มาจากค่าสีของพิกเซล 2 พิกเซลที่มีค่าแตกต่างกันมากที่สุดจากทั้งหมด 8 พิกเซลที่อยู่ล้อมรอบพิกเซลขอบแต่ละพิกเซล [7] จุดค้อยของวิธีการนี้คือต้องใช้เวลาในการประมวลผลค่อนข้างนานกว่าจะได้ค่าเฉลี่ยของแต่ละวัตถุที่จะนำไปใช้ในการจำแนกกลุ่ม

ในปี ค.ศ. 2011 Chucui Yi และ Yingli Tian ได้นำเสนอวิธีการจัดกลุ่มและวิเคราะห์โครงสร้างของข้อความเพื่อตรวจจับข้อความที่ปรากฏในภาพฉากธรรมชาติโดยวัดจากความสูง ความกว้าง จุดศูนย์กลาง ขนาดพื้นที่ และลักษณะการเชื่อมต่อระหว่างวัตถุโดยวัตถุที่เป็นตัวอักษรจะต้องเชื่อมต่อกับวัตถุที่อยู่ทางซ้ายและวัตถุที่อยู่ทางขวา หากวัตถุใดมีโครงสร้างและลักษณะการเชื่อมต่อที่แตกต่างไปจากเกณฑ์ที่กำหนดไว้ก็จะถือว่าวัตถุชิ้นนั้นไม่ใช่ตัวอักษร [8] จุดค้อยของวิธีการนี้คือไม่สามารถตรวจจับข้อความที่ประกอบด้วยตัวอักษรน้อยกว่า 3 ตัวและไม่สามารถตรวจจับสระและวรรณยุกต์ภาษาไทยที่อยู่ด้านบนและด้านล่างของพยัญชนะ

ในปี พ.ศ. 2550 วิทชา จิรัฐติเจริญ ได้นำเสนอวิธีการคัดกรองส่วนประกอบที่คาดว่าจะเป็นตัวอักษรภาษาไทยโดยการเปรียบเทียบอัตราส่วนของแต่ละบริเวณปิดกับอัตราส่วนที่กำหนดไว้ โดยอัตราส่วนความกว้างต่อความสูงของพยัญชนะภาษาไทยจะอยู่ในช่วง 0.2-1.2 และอัตราส่วนของพื้นที่ของพยัญชนะภาษาไทยต่อพื้นที่กรอบล้อมจะมีค่ามากกว่า 0.32 อัตราส่วนนี้ได้มาจากการพิจารณารูปแบบตัวอักษรภาษาไทยจำนวน 12 รูปแบบ ได้แก่ Angsana_New, AngsanaUPC, Browallia_New, BrowalliaUPC, Cordia_New, CordiaUPC, DilleniaUPC, FreesiaUPC, IrisUPC, JasmineUPC, KodchiangUPC และ LilyUPC [9] จุดค้อยของวิธีการนี้คืออัตราส่วนที่นำมาใช้ในการคัดกรองส่วนประกอบที่คาดว่าจะเป็นตัวอักษรภาษาไทยนั้นเฉพาะเจาะจงกับรูปแบบตัวอักษรภาษาไทยที่กำหนดไว้เท่านั้น

3. วิธีการดำเนินการวิจัย

กระบวนการตรวจจับข้อความภาษาไทยในฉากธรรมชาติในงานวิจัยนี้แบ่งออกเป็น 2 ส่วน คือ การจำแนกกลุ่มเส้นขอบและการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน

3.1 การจำแนกกลุ่มเส้นขอบ

ภาพฉากธรรมชาติส่วนใหญ่มักมีพื้นหลังที่ซับซ้อนเนื่องจากมีวัตถุที่ไม่ใช่ข้อความ หลากสี หลากขนาดและรูปร่างปรากฏอยู่ในรูปภาพทำให้การตรวจจับข้อความเป็นไปได้ยาก ดังนั้นการจำแนกข้อความออกจากวัตถุอื่นที่ไม่ใช่ตัวอักษรจึงเป็นวิธีที่จะช่วยลดความซับซ้อนของพื้นหลังของภาพฉากธรรมชาติและช่วยลดระยะเวลาในการประมวลผลและวิเคราะห์ส่วนประกอบที่เชื่อมต่อกันในขั้นตอนต่อไป

สมมติฐานที่ใช้ในการออกแบบกระบวนการจำแนกกลุ่มเส้นขอบมีดังนี้

- สีตัวอักษรเป็นสีพื้นเพียงสีเดียวและไม่มีสวดลาย
- พื้นหลังของตัวอักษรเป็นสีพื้นเพียงสีเดียวและไม่มีสวดลาย
- ตัวอักษรของข้อความเดียวกันจะมีสีตัวอักษรและสีพื้นหลังใกล้เคียงกัน
- ข้อความภาษาไทยมีการวางตัวในแนวนอน

กระบวนการจำแนกกลุ่มเส้นขอบจะเริ่มจากการเตรียมภาพโดยการแปลงเป็นภาพสีให้เป็นภาพระดับสีเทา จากนั้นจะทำการตรวจจับเส้นขอบของวัตถุต่างๆ ที่ปรากฏอยู่ในรูปภาพโดยใช้ตัวดำเนินการหาขอบภาพแบบแคนนี่

เมื่อได้เส้นขอบแล้ว จะทำการหาสีของวัตถุและสีของพื้นหลังโดยการนำค่าสีแดง สีเขียว และสีน้ำเงินของพิกเซลทั้งแปดที่อยู่ล้อมรอบพิกเซลขอบของวัตถุพิกเซลใดก็ได้ไปคำนวณหาพิกเซล 2 พิกเซลที่มีค่าสีแตกต่างกันมากที่สุดโดยการเปรียบเทียบทีละคู่ คือ พิกเซลตำแหน่งที่ 1 กับ 8, 2 กับ 7, 3 กับ 6 และ 4 กับ 5 ตำแหน่งของพิกเซลที่อยู่ล้อมรอบพิกเซลขอบแสดงได้ดังภาพที่ 1

1	2	3
4	พิกเซล ขอบ	5
6	7	8

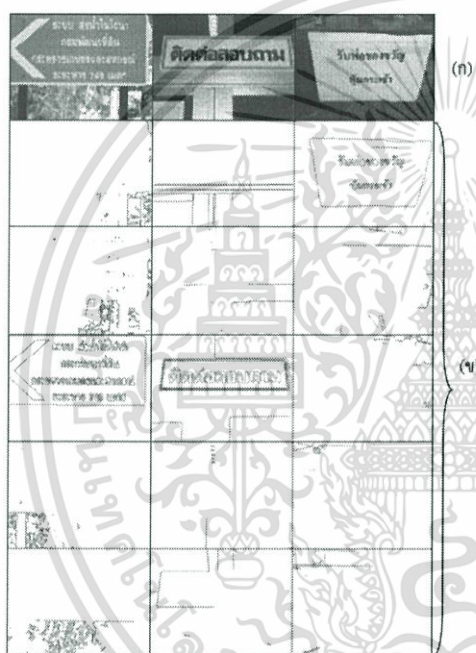
ภาพที่ 1: แสดงตำแหน่งของพิกเซลขอบและพิกเซลที่อยู่ล้อมรอบ

นอกจากสีของวัตถุและสีพื้นหลัง ตำแหน่งของวัตถุก็จะถูกนำมาพิจารณาในการจำแนกกลุ่มเส้นขอบด้วยเช่นกัน

$$x = [R_h, G_h, B_h, R_l, G_l, B_l, C_y]$$

โดยที่ x คือวัตถุหรือเส้นขอบที่ต้องการจำแนกกลุ่ม R_h, G_h, B_h คือค่าสีแดง สีเขียว และสีน้ำเงินของพิกเซลที่มีค่าสูงที่สุด R_l, G_l, B_l คือค่าสีแดง สีเขียว และสีน้ำเงินของ

พิกเซลที่มีค่าต่ำที่สุด และ C_y คือค่าพิกัดในแนวแกน y ของจุดศูนย์กลางของวัตถุ ตัวแปรเหล่านี้จะถูกใช้ในการจำแนกกลุ่มเส้นขอบด้วยเทคนิค k -means โดยกำหนดให้ $k=5$ ซึ่งให้ผลลัพธ์ในการจำแนกกลุ่มที่ดีที่สุด [7] หากกำหนดค่า k น้อยเกินไปก็จะทำให้ไม่สามารถจำแนกวัตถุที่เป็นตัวอักษรออกจากพื้นหลังที่ซับซ้อน ในทางกลับกัน หากกำหนดค่า k มากเกินไปก็อาจจะทำให้วัตถุที่เป็นตัวอักษรในข้อความเดียวกันถูกแยกไปอยู่คนละกลุ่ม ผลลัพธ์ที่ได้จากการจำแนกกลุ่มเส้นขอบซึ่งได้ทดลองกับภาพตัวอย่างจำนวน 3 ภาพ แสดงได้ดังภาพที่ 2



ภาพที่ 2: (ก) ตัวอย่างภาพต้นฉบับ และ
(ข) ผลลัพธ์ที่ได้จากการจำแนกกลุ่มเส้นขอบเป็น 5 กลุ่ม

อย่างไรก็ตาม เนื่องจากค่าเฉลี่ยหรือจุดศูนย์กลางเริ่มต้นของแต่ละกลุ่มได้มาจากการสุ่มเลือกข้อมูล ดังนั้น ผลลัพธ์ที่ได้จากการจำแนกกลุ่มเส้นขอบเป็น 5 กลุ่ม จึงไม่เหมือนกันในแต่ละครั้งที่ประมวลผลภาพภาพเดียวกัน

3.2 การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน

ข้อความที่ปรากฏในฉากธรรมชาติส่วนใหญ่มักจะมีลักษณะการวางแนวที่คล้ายกัน กล่าวคือ ข้อความจะประกอบด้วยตัวอักษรตั้งแต่ 2 ตัวขึ้นไปที่เกี่ยวข้องกันใน

แนวนอน ดังนั้น ลักษณะการเชื่อมต่อกันระหว่างวัตถุ 2 วัตถุที่ปรากฏในฉากธรรมชาติจึงสามารถนำมาใช้ในการตัดสินใจว่าวัตถุใดเป็นตัวอักษรและวัตถุใดไม่ใช่ตัวอักษร ดังนั้น งานวิจัยนี้จึงได้นำวิธีการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกันของ Chucai Yi และ Yingli Tian [8] มาปรับปรุงและพัฒนาต่อเพื่อให้สามารถตรวจจับข้อความภาษาไทยได้ถูกต้องมากยิ่งขึ้น

3.2.1 การวิเคราะห์เพื่อตรวจจับตัวอักษรภาษาไทยที่อยู่ในแนวเดียวกัน

เกณฑ์ที่กำหนดขึ้นเพื่อใช้ในการตัดสินใจว่าวัตถุใดเป็นพยัญชนะภาษาไทย มีดังนี้

- ตัวอักษรภาษาไทยมีการวางแนวในแนวนอนโดยเรียงจากซ้ายไปขวา ดังนั้น ค่าพิกัด y ของจุดกึ่งกลางของตัวอักษรที่อยู่ทางขวาจะต้องอยู่ระหว่างส่วนที่สูงที่สุดและต่ำที่สุดของตัวอักษรทางซ้าย
- ตัวอักษรภาษาไทยที่อยู่ในกลุ่มเดียวกันย่อมมีขนาดใกล้เคียงกัน ดังนั้น ความกว้างของตัวอักษรจะต้องมีค่ามากกว่าหรือเท่ากับ 80 % ของความกว้างของตัวอักษรที่อยู่ติดกัน และความสูงของตัวอักษรจะต้องมีค่ามากกว่าหรือเท่ากับ 80% ของความสูงของตัวอักษรที่อยู่ติดกัน
- ตัวอักษรที่อยู่ติดกันควรมีระยะห่างระหว่างตัวอักษรไม่มากเกินไป ดังนั้น ระยะระหว่างตัวอักษรที่อยู่ติดกันจะต้องมีค่าน้อยกว่า 50% ของความกว้างของตัวอักษรที่อยู่ทางซ้าย
- ความหนาของตัวอักษรที่อยู่ติดกันจะต้องมีค่าใกล้เคียงกัน ดังนั้น ผลต่างระหว่างความหนาของตัวอักษรที่อยู่ติดกันจะต้องมีค่าน้อยกว่า 50% ของความหนาของตัวอักษรที่อยู่ทางซ้าย
- พื้นที่ของตัวอักษรที่อยู่ติดกันจะต้องไม่มีส่วนที่ซ้อนทับกันมากเกินไป ดังนั้น จุดกึ่งกลางของตัวอักษรที่อยู่ทางซ้ายจะต้องไม่ซ้อนทับอยู่ในขอบเขตของตัวอักษรที่อยู่ทางขวา ทั้งนี้ การหาค่าความหนาของตัวอักษรจะพิจารณาจากค่าความเข้มแสงของพิกเซลที่เปลี่ยนแปลงไปในภาพระดับสีเทา โดยจะเริ่มวัดจากพิกเซลขอบที่อยู่บริเวณกึ่งกลางด้านขวาของวัตถุไล่ไปทางซ้ายที่ละพิกเซล หากค่าความเข้มแสงระหว่างพิกเซลมีค่ามากกว่าหรือเท่ากับ 50 จะถือว่าพิกเซลนั้นเป็นขอบอีกด้านของตัวอักษร และให้ระยะห่างระหว่างพิกเซลทั้งสองเป็นค่าความหนาของตัวอักษร

วัตถุที่ตรวจจับได้ว่าเป็นพยัญชนะภาษาไทยที่เรียงต่อกัน จะถูกนำไปคำนวณหาเส้นระดับข้อความ โดยการหาค่าความชันของเส้นที่ 3 และ 4 แล้วนำไปเปรียบเทียบกับค่าความชันระหว่างวัตถุที่ตรวจจับได้ว่าเป็นพยัญชนะภาษาไทยและวัตถุอื่น ๆ ที่เหลือ หากค่าความชันของเส้นที่ 3 หรือ 4 มีค่าใกล้เคียงกันจะถือว่าวัตถุนั้นเป็นพยัญชนะภาษาไทยที่อยู่ในบรรทัดเดียวกัน กระบวนการนี้จะทำให้สามารถตรวจจับพยัญชนะและสระภาษาไทยบางส่วนที่ไม่สามารถตรวจจับได้ในขั้นตอนแรกเนื่องจากมีลักษณะ โครงสร้างแตกต่างจากเกณฑ์ที่กำหนดไว้ได้ ถูกต้องมากขึ้น เส้นระดับของข้อความภาษาไทยแสดง ได้ดังภาพที่ 3



ภาพที่ 3: แสดงเส้นระดับของข้อความภาษาไทย

3.2.2 การวิเคราะห์เพื่อตรวจจับสระและวรรณยุกต์ภาษาไทย

เมื่อได้วัตถุที่คาดว่าเป็นตัวอักษรภาษาไทยแต่ละบรรทัดแล้ว จะนำวัตถุเหล่านี้ไปเปรียบเทียบกับวัตถุอื่น ๆ ที่เหลือเพื่อตัดสินใจว่าวัตถุใดเป็นสระหรือวรรณยุกต์ที่อยู่ด้านบนและด้านล่างของพยัญชนะที่ตรวจพบ เกณฑ์ที่กำหนดขึ้นเพื่อใช้ในการตัดสินใจมีดังนี้

- ตำแหน่งของสระจะต้องอยู่ด้านบนหรือด้านล่างของพยัญชนะตัวใดตัวหนึ่ง ดังนั้น ค่าพิทัก x ของขอบด้านซ้ายหรือขวาของวัตถุที่เป็นสระจะต้องมีค่าอยู่ระหว่างขอบด้านซ้ายและขอบด้านขวาของวัตถุที่เป็นพยัญชนะ
- ระยะห่างระหว่างพยัญชนะและสระที่อยู่ด้านบนหรือด้านล่างควรมีระยะไม่มากจนเกินไป ดังนั้น ระยะห่างระหว่างวัตถุที่เป็นพยัญชนะและวัตถุที่เป็นสระจะต้องมีค่าน้อยกว่าหรือเท่ากับ 50% ของความสูงของวัตถุที่เป็นพยัญชนะ
- ความกว้างของสระจะต้องมีค่าน้อยกว่าหรือเท่ากับสองเท่าของความกว้างของพยัญชนะ
- ความสูงของสระจะต้องมีค่าน้อยกว่าหรือเท่ากับครึ่งหนึ่งของความสูงของพยัญชนะ
- ตำแหน่งของวรรณยุกต์จะต้องอยู่ด้านบนของสระที่อยู่ระหว่างเส้นที่ 2 และ 3 ตัวใดตัวหนึ่ง ดังนั้น ค่าพิทัก x ของขอบด้านซ้ายหรือขวาของวัตถุที่เป็นวรรณยุกต์จะต้องอยู่ระหว่างขอบด้านซ้ายและขอบด้านขวาของวัตถุที่เป็นสระ

- ระยะห่างระหว่างสระและวรรณยุกต์ควรมีระยะไม่มากจนเกินไป ดังนั้น ระยะห่างระหว่างวัตถุที่เป็นสระและวัตถุที่เป็นวรรณยุกต์จะต้องมีค่าน้อยกว่าหรือเท่ากับ 50% ของความสูงของวัตถุที่เป็นสระ
- ความกว้างของวรรณยุกต์จะต้องมีค่าน้อยกว่าหรือเท่ากับสองเท่าของความกว้างของสระ

อัลกอริทึมสำหรับวิเคราะห์ส่วนประกอบที่เชื่อมต่อกันเพื่อตัดสินใจว่าวัตถุใดเป็นตัวอักษรภาษาไทยแสดง ได้ดังภาพที่ 4

```

FOR every clusters (5 clusters)
  FOR every connected components C and C'
    IF topBorder(y)C ≤ centroid(y)C ≤ bottomBorder(y)C
      AND (0.8 × widthC) ≥ widthC
      AND (0.8 × heightC) ≥ heightC
      AND (leftBorder(x)C - rightBorder(x)C) ≤ (0.5 × widthC)
      AND |thicknessC - thicknessC'| ≤ (0.5 × widthC)
      AND centroid(x)C < leftBorder(x)C
      RETURN CC (Connected Characters)
    END IF
  END FOR

FOR every connected characters CC and left components LC
  IF |topSlopeCC - topSlopeLC| < 0.05
    OR |bottomSlopeCC - bottomSlopeLC| < 0.05
    RETURN LC = CC
  END IF
END FOR

FOR every connected characters CC and left components LC
  IF (leftBorder(x)CC < leftBorder(x)LC < rightBorder(x)CC OR
    leftBorder(x)CC < rightBorder(x)LC < rightBorder(x)CC)
    AND verticalGapDistance(CC, LC) ≤ (0.5 × heightCC)
    AND widthLC ≤ (2 × widthCC) AND heightLC ≤ heightCC
    RETURN LC = UV, LV (Upper Vowels and Lower Vowels)
  END IF
END FOR

FOR every upper vowels UV and left components LC
  IF (leftBorder(x)UV < leftBorder(x)LC < rightBorder(x)UV OR
    leftBorder(x)UV < rightBorder(x)LC < rightBorder(x)UV)
    AND verticalGapDistance(UV, LC) ≤ (0.5 × heightUV)
    AND widthLC ≤ (2 × widthUV)
    RETURN LC = TM (Tonal Marks)
  END IF
END FOR
END FOR
  
```

ภาพที่ 4: แสดงอัลกอริทึมสำหรับวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน

เมื่อเปรียบเทียบผลลัพธ์ที่ได้จากการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกันด้วยวิธีของ Chucai [8] กับวิธีที่พัฒนาขึ้นในงานวิจัยนี้จะพบว่า วิธีที่พัฒนาขึ้นในงานวิจัยนี้สามารถตรวจจับวรรณยุกต์และสระที่อยู่ระหว่างเส้นระดับข้อความเส้นที่ 1 กับ 3 และเส้นที่ 4 กับ 5 ได้ถูกต้องมากกว่าวิธีของ Chucai ตัวอย่างผลลัพธ์แสดงได้ดังภาพที่ 5

อาการเก็บรหัส (ก)	ตัวอักษร (ข)	ตัวอักษรเก็บรหัส (ค)
-------------------	--------------	----------------------

ภาพที่ 5: (ก) ตัวอย่างภาพต้นฉบับ และผลลัพธ์ที่ได้จากการวิเคราะห์ ส่วนประกอบที่เชื่อมต่อกันด้วย (ข) วิธีของ Chucai เปรียบเทียบกับ (ค) วิธีที่พัฒนาขึ้นในงานวิจัยนี้

3.3 การเตรียมการทดลอง

สร้างฐานข้อมูลภาพขึ้นใหม่ โดยเป็นภาพสี RGB ที่ถ่ายด้วยกล้องดิจิทัล เพิ่มข้อมูลภาพเป็นชนิด JPEG มีขนาดภาพเท่ากับ 480x360 พิกเซล จำนวน 150 ภาพ ประกอบด้วยภาพที่มีพื้นหลังไม่ซับซ้อนจำนวน 50 ภาพ ภาพที่มีพื้นหลังซับซ้อนปานกลางจำนวน 50 ภาพ และภาพที่มีพื้นหลังซับซ้อนมากจำนวน 50 ภาพ เกณฑ์ที่ใช้ในการแบ่งกลุ่มแสดงดังตารางที่ 1

ตารางที่ 1: แสดงเกณฑ์ที่ใช้ในการแบ่งกลุ่มพื้นหลังออกเป็นประเภทไม่ซับซ้อน ซับซ้อนปานกลาง และซับซ้อนมาก

พื้นหลัง	เกณฑ์ที่ใช้ในการแบ่งกลุ่ม	
	จากภาพ	พื้นหลังป้าย
ไม่ซับซ้อน		ไม่ซับซ้อน
ซับซ้อนปานกลาง	ซับซ้อน อย่างใดอย่างหนึ่ง	
ซับซ้อนมาก	ซับซ้อน	

จากตารางที่ 1 จากภาพซับซ้อน หมายถึง ภาพที่มีวัตถุต่าง ๆ เช่น อาคาร ต้นไม้ ฯลฯ ปรากฏอยู่ในฉากภาพ และพื้นหลังป้ายซับซ้อน หมายถึง พื้นหลังป้ายมีลวดลาย หรือข้อความบนป้ายไม่ชัดเจน ซึ่งอาจเกิดจากตัวอักษรสีซีดจาง ภาพที่ถ่ายในสภาพแสงน้อย ภาพที่มีแสงสะท้อนหรือเงาพาดผ่านบริเวณป้าย

งานวิจัยนี้ใช้โปรแกรม MATLAB [10] ในการพัฒนาอัลกอริทึมสำหรับประมวลผลภาพฉากธรรมชาติเพื่อตรวจจับข้อความภาษาไทย โดยการจำแนกกลุ่มเส้นขอบและวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน

3.4 การวัดประสิทธิภาพ

การวัดประสิทธิภาพของกระบวนการตรวจจับข้อความภาษาไทยในงานวิจัยนี้จะวัดจากค่าความแม่นยำ (Precision) หรืออัตราส่วนของการตรวจจับตัวอักษรที่ถูกต้องจากจำนวนวัตถุทั้งหมดที่ตรวจจับได้ว่าเป็นตัวอักษร และค่าการเรียกกลับคืน (Recall) หรืออัตราส่วนของการตรวจจับตัวอักษรที่ถูกต้องจากจำนวนตัวอักษรทั้งหมด สามารถคำนวณได้ดังสมการที่ (2) และ (3)

$$\text{precision} = \frac{tp}{tp + fp} \quad (2)$$

$$\text{recall} = \frac{tp}{tp + fn} \quad (3)$$

โดยที่ tp คือวัตถุที่เป็นตัวอักษรและตรวจจับได้ว่าเป็นตัวอักษร fp คือวัตถุที่ไม่ใช่ตัวอักษรแต่ตรวจจับได้ว่าเป็นตัวอักษร และ fn คือวัตถุที่เป็นตัวอักษรแต่อัลกอริทึมไม่สามารถตรวจจับได้

4. ผลการดำเนินงาน

จากการทดสอบประสิทธิภาพของกระบวนการตรวจจับข้อความภาษาไทยในงานวิจัยนี้กับภาพฉากธรรมชาติในฐานข้อมูลที่สร้างขึ้นมาจำนวน 150 ภาพ ซึ่งจำแนกเป็น 3 กลุ่มคือ ภาพที่มีพื้นหลังไม่ซับซ้อน ภาพที่มีพื้นหลังซับซ้อนปานกลาง และภาพที่มีพื้นหลังซับซ้อนมาก ได้ผลดังตารางที่ 2

ตารางที่ 2: แสดงประสิทธิภาพของกระบวนการตรวจจับข้อความภาษาไทยในฉากธรรมชาติโดยการจำแนกกลุ่มเส้นขอบและวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน (BC+CC) เปรียบเทียบกับกรณีที่ไม่จำแนกกลุ่มเส้นขอบ (CC)

พื้นหลัง	จำนวนตัวอักษร	ประสิทธิภาพ			
		BC + CC		CC	
		Precision	Recall	Precision	Recall
ไม่ซับซ้อน	1,537	0.84	0.90	0.70	0.94
ซับซ้อนปานกลาง	2,038	0.66	0.85	0.42	0.87
ซับซ้อนมาก	1,980	0.57	0.89	0.37	0.93

จากตารางที่ 2 เมื่อเปรียบเทียบค่าความแม่นยำของทั้ง 3 กลุ่มจะพบว่า ยิ่งพื้นหลังของภาพฉากธรรมชาติที่ใช้ในการทดสอบประสิทธิภาพมีความซับซ้อนมาก ค่าความแม่นยำที่ได้ก็จะยิ่งมีค่าลดลง เนื่องจากพื้นหลังที่ซับซ้อนจะมีวัตถุที่ไม่ใช่ตัวอักษรปรากฏอยู่ในภาพเป็นจำนวนมาก หากลักษณะการเชื่อมต่อกันของวัตถุเหล่านี้ใกล้เคียงกับลักษณะการเชื่อมต่อกันของตัวอักษรก็จะทำให้วัตถุเหล่านี้ถูกตรวจจับได้ว่าเป็นตัวอักษร ทั้งนี้ เมื่อเปรียบเทียบค่าการเรียกกลับคืนของทั้ง 3 กลุ่มจะพบว่ามีการเรียกกลับคืนใกล้เคียงกัน

กรณีที่ไม่จำแนกกลุ่มเส้นขอบ จากการทดสอบประสิทธิภาพพบว่ามีความแม่นยำน้อยกว่ากรณีที่จำแนกกลุ่มเส้นขอบก่อนทำการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน

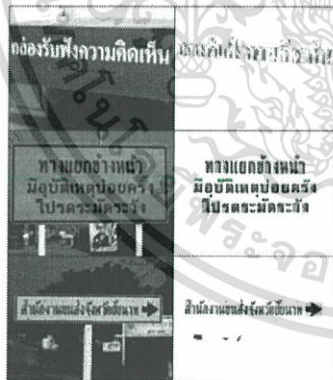
เนื่องจากวัตถุที่ไม่ใช่ตัวอักษรที่นำมาวิเคราะห์และเปรียบเทียบกันมีจำนวนมากขึ้น ทำให้โอกาสที่วัตถุเหล่านี้จะมีลักษณะการเชื่อมต่อกันใกล้เคียงกับตัวอักษรเพิ่มขึ้น สำหรับค่าการเรียกกลับคืนที่ได้ในกรณีนี้พบว่ามีความมากกว่าเล็กน้อย เนื่องจากไม่มีโอกาสที่วัตถุที่เป็นตัวอักษร ในข้อความเดียวกันถูกแยกไปอยู่คนละกลุ่ม

ตารางที่ 3: แสดงประสิทธิภาพด้านเวลาที่ใช้ในประมวลผลในขั้นตอนวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน กรณีที่จำแนกกลุ่มเส้นขอบ (BC+CC) เปรียบเทียบกับกรณีที่ไม่จำแนกกลุ่มเส้นขอบ (CC)

พื้นที่หลัง	เวลาที่ใช้ในการประมวลผลเฉลี่ย (วินาที/ภาพ)	
	BC + CC	CC
ไม่ซับซ้อน	0.052776	0.057434
ซับซ้อนปานกลาง	0.060448	0.077608
ซับซ้อนมาก	0.072102	0.120136

จากตารางที่ 3 เมื่อเปรียบเทียบเวลาที่ใช้ในการประมวลผลของทั้งสองกรณีพบว่า กรณีที่จำแนกกลุ่มเส้นขอบใช้เวลาในการประมวลผลเฉลี่ยต่อหนึ่งภาพน้อยกว่ากรณีที่จำแนกกลุ่มเส้นขอบ และยิ่งพื้นที่หลังของภาพมีความซับซ้อนมาก การจำแนกกลุ่มเส้นขอบก็จะช่วยลดเวลาที่ต้องใช้ในการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกันได้มากขึ้น

ตัวอย่างผลลัพธ์ที่ได้จากการตรวจจับข้อความภาษาไทยในฉากธรรมชาติในงานวิจัยนี้แสดงได้ดังภาพที่ 6



ภาพที่ 6: (ซ้าย) ตัวอย่างภาพต้นฉบับ และ (ขวา) ผลลัพธ์ที่ได้จากการตรวจจับข้อความภาษาไทยในฉากธรรมชาติ จำแนกตามความซับซ้อนของพื้นที่หลัง (ก) ไม่ซับซ้อน (ข) ซับซ้อนปานกลาง (ค) ซับซ้อนมาก

5. สรุป

งานวิจัยนี้ได้นำเสนอวิธีการตรวจจับข้อความภาษาไทยในฉากธรรมชาติ เริ่มจากการเตรียมภาพโดยการแปลงเป็นภาพสีให้เป็นภาพระดับสีเทา จากนั้นจะทำการตรวจจับเส้นขอบของวัตถุต่าง ๆ ที่ปรากฏอยู่ในรูปภาพและนำเส้นขอบที่ได้ไปจำแนกกลุ่มโดยพิจารณาจากสีของวัตถุ สีของพื้นหลัง และตำแหน่งของวัตถุ เส้นขอบที่ผ่านการจำแนกกลุ่มแล้วจะถูกนำไปวิเคราะห์โครงสร้างและลักษณะการเชื่อมต่อกันระหว่างวัตถุเพื่อตัดสินใจว่าวัตถุใดเป็นตัวอักษรและวัตถุใดไม่ใช่ตัวอักษร ระบบตรวจจับข้อความภาษาไทยในฉากธรรมชาตินี้สามารถนำไปพัฒนาต่อยอดร่วมกับระบบรู้จำตัวอักษรเพื่อนำไปประยุกต์ใช้ชีวิตประจำวันด้านต่าง ๆ อาทิ ระบบตรวจจับข้อความผ่านกล้องและแปลภาษาโดยอัตโนมัติบนสมาร์ตโฟน ระบบตรวจจับและอ่านป้ายทะเบียนรถยนต์สำหรับงานด้านจราจร เป็นต้น

อย่างไรก็ตาม เพิ่มข้อมูลภาพขนาด 480x360 พิกเซลที่ใช้ในการทดลองนี้เป็นภาพขนาดเล็ก ดังนั้น หากต้องการตรวจจับข้อความในภาพที่มีขนาดใหญ่ จำเป็นต้องปรับปรุงวิธีการที่ใช้ในงานวิจัยนี้ให้มีความเหมาะสมกับข้อมูลที่มีปริมาณมากขึ้น

เอกสารอ้างอิง

- [1] X. Chen, J. Yang, J. Zhang, and A. Waibel, "Automatic Detection and Recognition of Signs From Natural Scenes" *IEEE Trans. on Image Process.*, vol. 13, no. 1, pp. 87-99, 2004.
- [2] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing (2nd Edition)*, Prentice Hall, New Jersey, 2002.
- [3] อรรถกร จิตต์โสภักดิ์, *ทฤษฎีการประมวลผลภาพดิจิทัล*, สงวนลิขสิทธิ์ แอนด บิเคย์, กรุงเทพฯ, 2552.
- [4] สมเกียรติ อุดมหาระพาส, *การประมวลผลภาพดิจิทัลเบื้องต้น*, สำนักพิมพ์ท็อป, กรุงเทพฯ, 2554.
- [5] Kenneth J. Meier, Jeffrey L. Brudney, and John Bohte, *Applied Statistics for Public and Nonprofit Administration (7th Edition)*, Thomsan Wadsworth, Belmont, 2009.
- [6] Jiawei Han, and Micheline Kamber, *Data Mining: Concepts and Techniques (2nd Edition)*, Morgan Kaufmann Publishers, San Francisco, 2006.
- [7] C. Yi and Y. Tian, "Localizing Text in Scene Images by Boundary Clustering, Stroke Segmentation, and String Fragment Classification" *IEEE Trans. on Image Process.*, vol. 21, no. 9, pp. 4256-4268, 2012.
- [8] C. Yi and Y. Tian, "Text String Detection From Natural Scenes by Structure-Based Partition and Grouping" *IEEE Trans. on Image Process.*, vol. 20, no. 9, pp. 2594-2605, 2011.

- [9] วิชา จิรจิตเจริญ, *การตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติ*, วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย, ศงส.ลาดกระบัง, 2550.
- [10] Rafael C. Gonzalez, Richard E. Woods, and Steven L. Eddins, *Digital Image Processing Using MATLAB*, Pearson Prentice Hall, New Jersey, 2004.

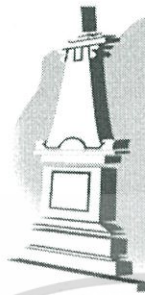


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ISSN: 2088-6578
ISBN: 978-1-4799-0423-5



50th ANNIVERSARY
DEPARTMENT OF ELECTRICAL
ENGINEERING AND INFORMATION
TECHNOLOGY UGM



ICITEE 2013

Proceedings of

2013 International Conference
on Information Technology and
Electrical Engineering

7-8 October 2013

The Sahid Rich Jogja Hotel

Yogyakarta, Indonesia

**“Intelligent and Green Technologies for
Sustainable Development”**



Organized by

Department of Electrical Engineering
and Information Technology
Universitas Gadjah Mada, Indonesia

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปเผยแพร่โดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



of Information Technology, Universitas Ma Chung; Monika N.U. Prihastyanti, Ma Chung Research Center for Photosynthetic Pigments, Universitas Ma Chung; Leenawaty Limantara, Ma Chung Research Center for Photosynthetic Pigments, Universitas Ma Chung; Tatas H.P. Brotosudarmo, Ma Chung Research Center for Photosynthetic Pigments, Universitas Ma Chung

Abstract-Nitrogen is one of the important nutrients elements for the growth of soybean plants. In this paper we propose mobile application that can be used nondestructively to estimate the nitrogen content of soybean leaves. We named this software "Mata Daun". The primary concept of this software is to relate the RGB (Red, Green, Blue) value of the captured soybean image with its nitrogen content. Furthermore, the captured image is processed into Enhanced Color Visibility (ECV) index using digital image processing method for the ease of software algorithm process. Calibration process and field trial were conducted to found the relation between ECV index and soybean leaves nitrogen content. The calibration result showed that the nitrogen readings by this application had a fairly strong relationship ($R2 = 0.70$) with the soybean leaves nitrogen content (Agriexpert GCN-6000 readings). The field test result also gave the same strong positive relationship between predicted and real soybean leaves nitrogen content ($R2 = 0.93$).

6 Femicidae Swarm System

Thammarat Taengtang, King Mongkut's Institute of Technology Ladkrabang; Witthaya Siththivet, King Mongkut's Institute of Technology Ladkrabang; Kitti Paithoonwattanakij, King Mongkut's Institute of Technology Ladkrabang

Abstract-Improved an ant colony by firefly algorithm, in this paper is proposed the method that is interwoven between ant colony optimization and firefly algorithm to increase efficiency of solving the traveling salesman problem it is called that Femicidae swarm system (FSS). It uses relationship between pheromone and distance which is attractiveness and absorption coefficient. This method is based on ant colony optimization which state transition rule of ant colony is improved by adding detection, which is a condition of distances. The performance of FSS is divided into two parts: the speed and tour length of a result. A speedy result of FSS is faster than ACS and tour length of FSS is near the best result.

7 Semantic Interrelation in Distributed System Through Green Computing Ontology

Herlina Jayadianti, Universidade do Minho; Lukito Edi Nugroho, Gadjah Mada University; Carlos B. Sousa Pinto, University of Minho; Paulus Insap Santosa, Gadjah Mada University; Wahyu Widayat, Gadjah Mada University, Faculty of Economic and Development

Abstract-Green computing refers to the system that provides minimal impact on the environment. When we are talking about green computing we discuss about how much energy is used by the system, such as energy used by the system, time used for the search process, and how effective the system is. Related to that issue, through this paper we want to propose a new effort to achieve Green Computing in heterogeneous data in distributed system. The technology chosen to deal with them is Ontology. We try to generate a common ontology including a common set of terms, based on the several ontologies available, in order

to make possible to share the common terminology (set of terms) that it implements, between different communities. If a very large amount of distributed data is not managed and distributed properly, user will need more time to do a search process. The longer the search is done, the more energy is used.

TS-1_3C: Software Engineering, Services, and Information Technology 6

Monday, 07 October 15.30-17:50 Orchid 2

1 Improved Thai Text Detection from Natural Scenes

Kuntpong Woraratpanya, Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang; Pimlak Boonchukusol, Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang; Yoshimitsu Kuroki, Kurume National College of Technology; Yasushi Kato, Tsuruoka National College of Technology

Abstract-Thai text detection from natural scenes is still a challenging task for language translation applications, since there are many unsolved issues. Furthermore, the existing related works cannot completely detect Thai text. The main reason is that Thai text layout has vowels and tonal marks that differ from other languages. This paper proposes an approach to detect Thai text from natural scenes. The approach consists of two main procedures. (i) Fast boundary clustering algorithm decomposes scene features into multilayers, so that it is faster and easier to analyze Thai text characters. (ii) Modified connected component analysis method is applied to such scene features in order to detect Thai text boundaries. Based on 150 test images with 4,920 characters, the experimental results demonstrate that the proposed approach achieves the high average precision and recall, 0.80 and 0.90.

2 Likelihood Calculation Classification for Indonesian Language News Documents

Aini Fuddoly, Universiti Teknologi Petronas; Jafreezal Jaafar, Universiti Teknologi PETRONAS; Norshuhani Zamin, Mrs

Abstract-Text categorization has been an important research area that seeks to classify textual documents into a group of predetermined categories. Unfortunately, the interest towards Indonesian news classification has been very little. In this paper, we propose a text categorization algorithm based on Bracewell method that uses the likelihood calculation between the article and the category's keywords. Through experiments, the algorithm succeeded in classifying Indonesian news corpus with accuracy as high as 93,84% in offline environment, 93,82% in online environment, and 80% benchmarking against human evaluation.

3 News Recommendation in Indonesian Language Based on User Click Behavior

Diandra Desyaputri, Swiss German University; Alva Erwin, Swiss German University; Maulahikmah Galinium, Swiss German University; Didi Nugrahadi, Beritagar

Abstract-Recommendation system has been

Improved Thai Text Detection from Natural Scenes

Kuntpong Woratpanya¹

Faculty of Information Technology,
King Mongkut's Institute of Technology Ladkrabang,
Bangkok, Thailand 10520
Email: kuntpong@it.kmitl.ac.th¹

Pimlak Boonchukusol²

Faculty of Information Technology,
King Mongkut's Institute of Technology Ladkrabang,
Bangkok, Thailand 10520
Email: pimlak.b@gmail.com²

Yoshimitsu Kuroki³

Kurume National College of Technology,
Fukuoka, Japan 835-8555
Email: kuroki@kurume-nct.ac.jp³

Yasushi Kato⁴

Tsuruoka National College of Technology,
Tsuruoka, Yamagata, Japan 997-8511
Email: president@tsuruoka-nct.ac.jp⁴

Abstract—Thai text detection from natural scenes is still a challenging task for language translation applications, since there are many unsolved issues. Furthermore, the existing related works cannot completely detect Thai text. The main reason is that Thai text layout has vowels and tonal marks that differ from other languages. This paper proposes an approach to detect Thai text from natural scenes. The approach consists of two main procedures. (i) Fast boundary clustering algorithm decomposes scene features into multilayers, so that it is faster and easier to analyze Thai text characters. (ii) Modified connected component analysis method is applied to such scene features in order to detect Thai text boundaries. Based on 150 test images with 4,920 characters, the experimental results demonstrate that the proposed approach achieves the high average precision and recall, 0.80 and 0.90.

Keywords—Thai text detection; natural scene; fast boundary clustering; modified connected component analysis

I. INTRODUCTION

Text labels are everywhere in our lives, such as traffic signs, place name signs, warning signs, and description signs. These texts in natural scenes are information that makes our live easier; therefore, Thai text detection is a very important procedure for applying to language translation applications for smart phone to help foreign tourists understand Thai language signs.

Automatic text detection from natural scenes is still a challenging task, since there are many unsolved issues, which are not only variety of text: font style, size, color, and orientation but also text intensity affected by lighting conditions: low light, highlight, and shadow. Thus, error rate of text detection from natural scenes is higher than that of OCR systems which work on scanned documents under controlled environment [1].

In natural scene images, text is usually printed as a string which contains more than one character. The string most likely appears similar in size, character color, background color, and character alignment. By nature of these features, the text detection without a character recognition can be performed by using adjacent character grouping and connected

component labeling proposed by C. Yi et al. [2], and R. Chang [3], respectively. However, these approaches usually fail to remove other objects which resemble text characters. To solve this problem, the multilayer separation method introduced by M. Kumar et al. [4] was employed to split text characters from other objects. In the same way, C. Yi et al. [5] proposed boundary clustering to group the boundary pixels with bigram color uniformity on the border of text and attachment surface. These algorithms take more computational time in data preparation for clustering, since all neighborhood pixels of all edge pixels of each boundary are computed.

In this paper, we propose an approach to detect Thai text with multiple sizes and colors from natural scene images. This approach is composed of two main procedures: (i) fast boundary clustering and (ii) modified connected component analysis. In the former procedure, the features based on colors and positions are extracted. Then k-means algorithm is applied to such features in order to decompose character boundaries from other objects. In the latter procedure, the decomposed features are analyzed by using a structure of text characters, such as position, size, stroke width, distance between two components, and character alignment, in order to detect Thai text boundary. The proposed method improves the efficiency of Thai text detection from natural scene images.

The remainder of this paper is organized as follows. Fast boundary clustering and modified connected component analysis, which are two principal procedures of the proposed approach, are described in section II and section III, respectively. Section IV presents and discusses results of experiments. Finally, the conclusion is given in section V.

II. FAST BOUNDARY CLUSTERING

Boundary clustering is a procedure to decompose boundary objects into multilayers, which is very useful to accurately detect the target objects. C. Yi et al. [5] proposed boundary clustering to group the boundary pixels with bigram color uniformity on the border of text and attachment surface. This method effectively decomposes boundary objects, but it consumes a lot of computing time, especially complex back-

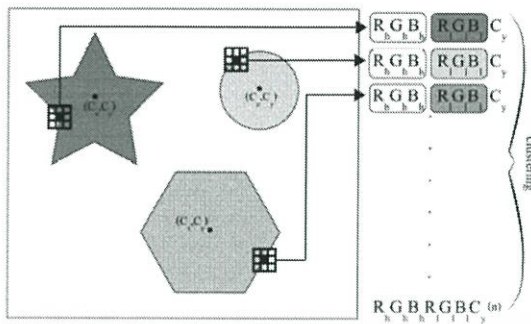


Fig. 1. Feature vector extraction based on the proposed assumptions.

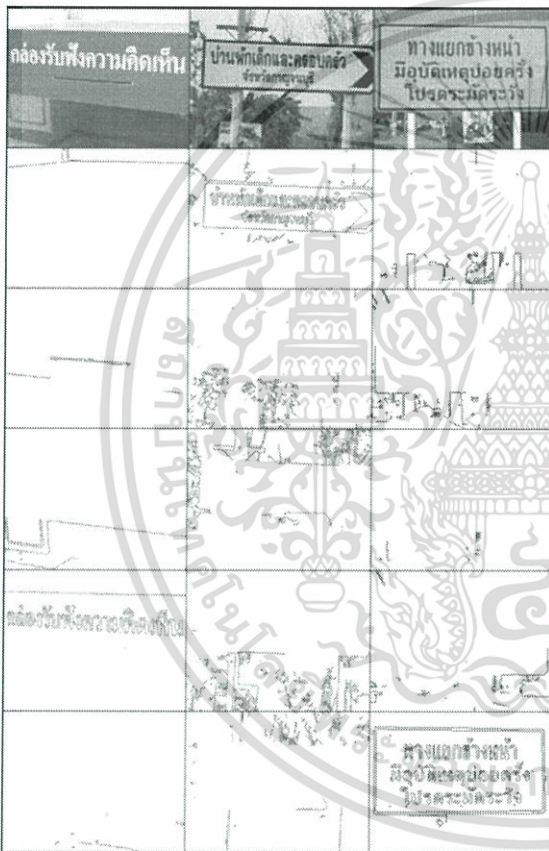


Fig. 2. Boundary clustering, the first row is three original images and other rows are boundary layers after clustering.

grounds. Therefore, in order to effectively separate Thai text strings from the complex backgrounds, this paper proposes the fast boundary clustering algorithm based on the essential assumptions described as follows:

- Characters in a same scene have almost the same foreground and background colors.
- Text foreground and background colors are normally a uniform color.

Based on these assumptions, the proposed approach can speed up by using the following two steps:

In the first step, Canny edge detector is used to obtain object edge sets. As mentioned by the above assumptions, the feature of each object edge is extracted by determining the maximum color difference between two opposite pixels on a 3×3 mask as demonstrated in Fig. 1. In this way, the feature vector, V_i , of each object edge is formed from three components, the highest color intensity, the lowest color intensity, and the spatial position of boundary, i.e., y-coordinate of the centroid of boundary. The vector can be defined as (1). The highest and lowest color intensities are the sum of z-scores of RGB values. The z-score equation is defined as (2).

$$V_i = [R_h, G_h, B_h, R_l, G_l, B_l, C_y]_i, \quad (1)$$

where $[R_h, G_h, B_h]$, $[R_l, G_l, B_l]$, and $[C_y]$ imply pixel color values of the highest intensity, pixel color values of the lowest intensity, and the spatial position, respectively.

$$z = \frac{x - \mu}{\sigma}, \quad (2)$$

where z , x , μ , and σ denotes standardized z-score, values of R , G , or B , mean of each color value, and standard deviation of each color value, respectively.

In the second step, the feature vectors are decomposed into multilayers by means of k-means clustering algorithm. The number of clusters, K , is an important factor. If K is too small, text boundaries are not effectively extracted from complex backgrounds. On the other hand, if K is too large, the text boundaries are probably broken into several fragments and assigned into different boundary layers. Fig. 2 illustrates three examples of boundary layers after clustering. Edge pixels with similar color pairs and spatial positions are perfectly grouped into the same layer. This shows the best results of applying the clustering algorithm proposed in [5], when the number of clusters is equal to 5.

III. MODIFIED CONNECTED COMPONENT ANALYSIS

Connected component analysis is a procedure to detect text characters and filter unwanted objects out. C. Yi et al. [2] proposed adjacent character grouping to decide whether three connected components are character sets and merge the character sets together. This method effectively detects characters, but its accuracy probably reduces after boundary clustering, when the connected characters are less than three characters on each boundary layers. R. Chang [3] proposed connected component labeling to filter unwanted objects based on image sizes, component sizes, and the number of edge pixels. This method is a preprocessing step for a recognition

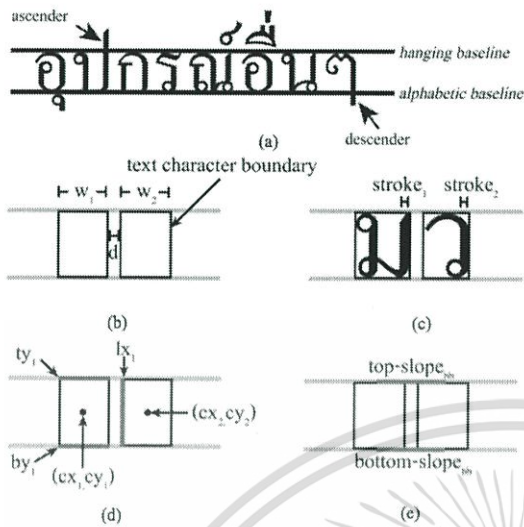


Fig. 3. Two-connected component detection model, (a) Thai word typography line, (b) width of text character boundary and distance between two boundaries, (c) stroke width of character, (d) borders and centroids of boundary, and (e) slopes between two boundaries.

phase to decide whether detected connected components is a character, but it cannot remove unwanted connected objects similar to connected characters. Moreover, these methods were designed to detect English characters only, thus they cannot completely detect Thai vowels and tonal marks.

Therefore, in order to effectively detect Thai characters on multilayers, this paper proposes the modified connected component analysis algorithm. Thai text strings in natural scene images usually appear in alignment, namely, each text character in a text string have to connect to another text character. The structure features among all components can be used to determine whether the connected components belong to text characters or unexpected noises. In the same way, most Thai typography characters, including 44 consonants, 19 vowels, and 4 tonal marks, are printed as a text string, which contains two or more characters. Therefore, to identify two connected components and locate its alphabetic and hanging baselines as shown in Fig. 3(a), the essential constraints are defined and described below; the constraint parameters are described in Table I; and a threshold parameter T_n with the suitable value is defined in Table II.

- Thai characters commonly align on the alphabetic baseline, such that $1/T_1 \leq w_1/w_2 \leq T_1$ as shown in Fig. 3(b).
- A gap of two connected characters is not far from each other character, such that $d/w_1 < 1/T_2$ and $d/w_2 < 1/T_2$ as shown in Fig. 3(b).
- Two connected characters usually appear in the same font size, such that $(|stroke_1 - stroke_2|/stroke_1) < 1/T_2$ and $(|stroke_1 - stroke_2|/stroke_2) < 1/T_2$ as shown in Fig. 3(c).

TABLE I. TERMINOLOGIES OF TWO-CONNECTED DETECTION MODEL

Parameter	Description
alphabetic baseline	A line which most characters align on
hanging baseline	A line is above alphabetic baseline with character height
ascender	A part of characters that extend above the hanging baseline
descender	A part of characters that extend below the alphabetic baseline
ty	y-coordinate of top-border of boundary
by	y-coordinate of bottom-border of boundary
lx	x-coordinate of left-border of boundary
Ty	y-coordinate of top-border of text string boundary
By	y-coordinate of bottom-border of text string boundary
(cx,cy)	centroid of boundary
(Cx,Cy)	centroid of text string boundary
w	width of boundary
h	height of boundary
W	width of two connected boundaries
H	height of two connected boundaries
stroke	stroke width of a character
d	distance between two connected boundaries
FColor ₁	average RGB values of foreground of text string boundary
FColor ₂	average RGB values of background of text string boundary
slope _{ab}	slope between two boundaries
slope _{cb}	slope between two-connected boundary and another boundary
slope _a	slope of text string boundary
slope _{bb}	slope between two text string boundaries

- Two-connected character area can be slightly overlapped, such that $cx_1 < lx_2$ as shown in Fig. 3(d).
- Two connected characters are usually aligned in a horizontal direction, such that $ty_1 < cy_2 < by_1$ as shown in Fig. 3(d).
- Two connected characters without ascender or descender usually align between the alphabetic baseline and the hanging baseline, such that $|\text{top-slope}_{cb} - \text{bottom-slope}_{bb}| < T_3$ as shown in Fig. 3(e).

After alphabetic baseline and hanging baseline of each text string are located, left and right projections are used to determine whether the remaining boundaries in the same layer belong to characters in the same text string as depicted in Fig. 4(a). The starting point is the minimum x-coordinate of the detected text string boundary for the left projection and the maximum x-coordinate of the detected text string boundary for the right projection. Text characters and its background usually have distinct intensity values, thus the end point is the x-coordinate of the pixel whose the intensity value differs

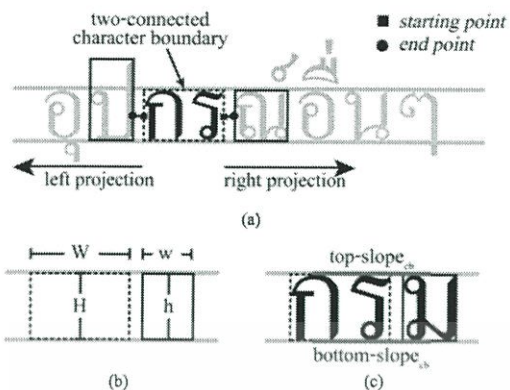


Fig. 4. Thai character boundary detection using projection model, (a) direction of projection and position of starting and end points, (b) size of two-connected boundary and other boundaries, and (c) slopes between two-connected boundary and another boundary.

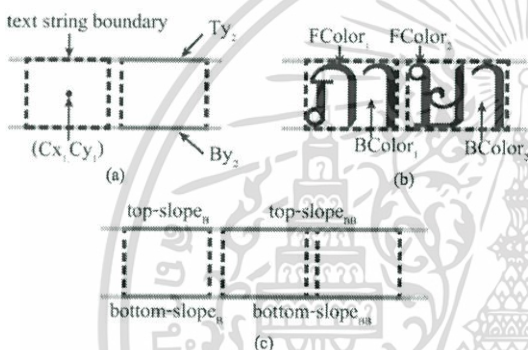


Fig. 5. Text string boundary grouping model, (a) borders and centroid of text string boundary, (b) foreground and background colors of text string boundary, and (c) slopes of text string boundary and slopes between two text string boundaries.

from previous pixel greater than T_4 . The starting point and end point of the left projection are y -coordinate of the center between the hanging baseline and the alphabetic baseline. If the end point is overlapped in a boundary, this boundary is determined whether it is a character or an unwanted object by using the following constraints.

- $1/T_1 \leq w/W \leq T_1$ as shown in Fig. 4 (b).
- $1/T_5 \leq h/H \leq T_2$ as shown in Fig. 4 (b).
- Either top border or bottom border of each Thai character boundary must be connected to the alphabetic baseline or the hanging baseline, such that $|\text{top-slope}_{cb} \square \text{top-slope}_{bb}| < T_3$ or $|\text{bottom-slope}_{cb} \square \text{bottom-slope}_{bb}| < T_3$ as shown in Fig. 4(c).

The projection procedure continuously performs until no more characters in the same text string is detected. However, each text string is probably assigned into different boundary layers, especially natural scene images with low background

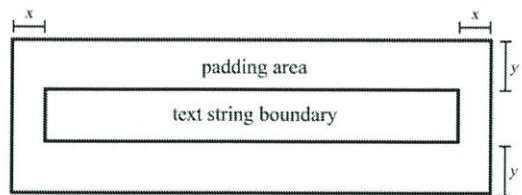


Fig. 6. The padding area used to detect Thai vowels and tonal marks.

complexity. Therefore, the constraints to group these text strings together are defined as follows:

- A text string is usually aligned in horizontal direction, such that $Ty_2 < Cy_1 < By_2$ as shown in Fig. 5(a).
- A text string usually appears in the same foreground and background colors, such that $|FCColor_1 \square FCColor_2| \leq T_4$ and $|BColor_1 \square BColor_2| \leq T_4$ as shown in Fig. 5(b).
- A text string has only one alphabetic baseline and one hanging line, such that $|\text{top-slope}_B \square \text{top-slope}_{BB}| < T_3$ and $|\text{bottom-slope}_B \square \text{bottom-slope}_{BB}| < T_3$ as shown in Fig. 5(c).

In order to increase the accuracy of detecting characters, which align on the alphabetic baseline, the left and right projections are used again. Furthermore, remaining boundaries that appear between left and right borders of each text string boundary are also detected.

However, in some cases, Thai characters do not align on the alphabetic baseline, since some vowels are above the hanging baseline or below the alphabetic baseline. Furthermore, all tonal marks always are above the hanging baseline. We determine whether the remaining boundaries belong to Thai characters that align above or lower the detected text string by using text string boundary surrounding padding. If a component is within the padding area, this component should be a character. The padding area size is specified as shown in Fig. 6, where x represents the sum of average boundary width and average distance between two connected boundaries and y represents average boundary height.

IV. EXPERIMENTS

In order to evaluate the efficiency of the proposed method in terms of precision, recall, and computing time, the experimental methods are setup as follows.

A. Dataset

There are two experimental datasets. The first dataset consists of 100 test images which contain 1 Thai character per each image. The pixel resolutions of test images are 64×64 . The second dataset comprises of 150 scene images with 4,920 character boundaries. Each image consists of one to four text strings and all of text strings are horizontal. This dataset is classified into three groups based on background complexity. The pixel resolution of test images is 480×360 .

TABLE II. PARAMETER VALUES FOR EXPERIMENTS

Parameter	T ₁	T ₂	T ₃	T ₄	T ₅
Threshold Value	7	2	0.05	50	4

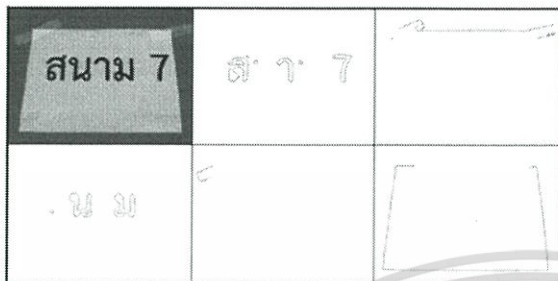


Fig. 7. A failure result from FBC method affected to accuracy of Thai text detection.

B. Performance Evaluation

The performance is evaluated by two metrics, *precision* and *recall*, defined as (3) and (4), respectively.

$$\text{precision} = \frac{tp}{tp + fp} \quad (3)$$

and

$$\text{recall} = \frac{tp}{tp + fn} \quad (4)$$

where *tp*, *fp*, and *fn* represent a number of successful detected characters, a number of detected noises, and a number of unsuccessful detected characters, respectively. In other words, precision is the ratio of the number of successful detected text characters to the total number of detected characters, whereas recall is the ratio of the number of successful detected characters to the total number of characters. The low precision implies overestimation while the low recall implies underestimation. The significant parameters of experiments are set up in Table II.

C. Results and Discussions

In order to evaluate the efficiency of the proposed method, three experiments are setup. The first experiments aim to test the fast boundary clustering (FBC) algorithm in terms of computing time and an error rate of the feature extraction, respectively. The second experiment aims to test Thai text detection. This experiment compares the results of a connected component analysis (CCA) with a modified connected component analysis (MCCA) method in terms of precision and recall. Finally, the last experiment aims to test Thai text detection without character recognition by using a combination of FBC and MCCA methods in terms of precision and recall. The first dataset is used for testing an error rate of the feature extraction and the second dataset is used for the others.

TABLE III. COMPARISON OF COMPUTING TIME OF BC AND FBC ALGORITHMS

Background Complexity	Average Processing Time (Sec/Image)	
	BC	FBC
Low	98.09	0.98
Medium	184.13	2.46
High	241.76	3.57

TABLE IV. COMPARISON OF EVALUATION RESULTS CCA AND MCCA ALGORITHMS

Background Complexity	Performance			
	CCA		MCCA	
	Precision	Recall	Precision	Recall
Low	0.77	0.93	0.89	0.94
Medium	0.42	0.90	0.77	0.92
High	0.33	0.81	0.66	0.85

TABLE V. EVALUATION RESULTS OF A COMBINATION OF FBC AND MCCA ALGORITHMS

Background Complexity	Performance	
	Precision	Recall
Low	0.90	0.93
Medium	0.81	0.92
High	0.70	0.84

In the first experiments, the boundary clustering (BC) algorithm proposed in [5] is implemented as a baseline. This algorithm extracts color features by determining all neighborhood pixels of all edges. In comparison of average processing time, the experimental results show that the BC algorithm takes an average 174.66 seconds/image whereas the FBC takes an average 2.34 seconds/image as presented in Table III. Furthermore, the error rate of color feature extraction is 2.42% when compared to the baseline algorithm. This shows that the FBC algorithm is not only faster but also more effective to extract color features.

In the second experiment, CCA algorithm proposed in [6] is implemented as a baseline. Table IV presents a comparison of Thai text detection with CCA and MCCA algorithms. It is clearly that MCCA achieves better precision and recall than CCA algorithm. Particularly, in medium and high background complexities, the precisions of the MCCA algorithm are greatly improved.

In the last experiment, a combination of FBC and MCCA methods is evaluated. Table V shows the experimental results. It is evident that the proposed method achieves a better precision. An important reason to support this achievement is that the FBC algorithm effectively decomposes character boundaries from other unwanted boundaries. This makes it easier to detect two connected character boundaries. Moreover, the FBC algorithm help improve the precision affected



Fig. 8. A part of experimental results of Thai text detection marked in cyan.

by background complexity, since scene images with high background complexity have more objects that resemble Thai text characters. In terms of recall, both MCCA algorithm and MCCA with FBC algorithm yields almost the same results. However, as text boundaries are probably broken into several fragments and assigned into different boundary layers, it is easy to fail to detect text characters. Fig. 7 shows an example of a failure result from FBC algorithm. In this case, the objects are decomposed into five layers and each layer does not contain two-connected component under the defined constraints in section III, thus MCCA algorithm fails to detect Thai text. Fig. 8 shows a part of Thai text detection results.

V. CONCLUSION

In this paper, we have proposed an approach to detect Thai text from natural scenes. This approach consists of two main procedures: fast boundary clustering (FBC) and modified connected component analysis (MCCA). The FBC procedure uses for separating text characters from other objects and the MCCA procedure uses for determining whether which object is Thai text character. The combination of FBC and MCCA procedures improves the precision of medium and high background complexities. Based on experimental results, the proposed approach achieves the high average precision and recall.

VI. REFERENCES

- [1] X. Chen, J. Yang, J. Zhang, and A. Waibel, "Automatic Detection and Recognition of Signs From Natural Scenes," *IEEE Trans. on Image Process.*, vol. 13, no. 1, pp. 87-99, January 2004.
- [2] C. Yi and Y. Tian, "Text String Detection From Natural Scenes by Structure-Based Partition and Grouping," *IEEE Trans. on Image Process.*, vol. 20, no. 9, pp. 2594-2605, September 2011.
- [3] R. Chang, "Intelligent Text Detection and Extraction from Natural Scene Images," *Proc. 15th North-East Asia Symposium on Nano, Information Technology and Reliability*, pp. 23-28, 2011.
- [4] M. Kumar, Y. C. Kim, and G. S. Lee, "Text Detection using Multilayer Separation in Real Scene Images," *Proc. 10th IEEE International Conference on Computer and Information Technology*, pp. 1413-1417, 2010.
- [5] C. Yi and Y. Tian, "Localizing Text in Scene Images by Boundary Clustering, Stroke Segmentation, and String Fragment Classification," *IEEE Trans. on Image Process.*, vol. 21, no. 9, pp. 4256-4268, September 2012.
- [6] K. Woratpanya and P. Boonchukusol, "Thai Text Detection from Natural Scenes by using Boundary Clustering and Connected Component Analysis," *Proc. 9th National Conference on Computing and Information Technology*, pp. 665-672, May 2013.

ประวัติผู้เขียน

ชื่อ-นามสกุล	นางสาวพิมพ์ลักษณ์ บุญชูกุล
วัน เดือน ปีเกิด	31 สิงหาคม 2531
ที่อยู่	77/217 ถนนนวมินทร์ 75 แขวงนวมินทร์ เขตบึงกุ่ม กรุงเทพฯ 10240
ประวัติการศึกษา	2553 วิทยาศาสตรบัณฑิต สาขาเทคโนโลยีสารสนเทศ (เกียรตินิยมอันดับ 1) มหาวิทยาลัยเกษตรศาสตร์



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้