

การตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติ
THAI TEXT DETECTION AND SEGMENTATION FROM NATURAL SCENES

ฉิทยา ฉิรุตติฉิเรณ
WITTAYA JIRATTITICHAREON

ฉิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
บัณฑิตวิทยาลัย
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2550

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติ

THAI TEXT DETECTION AND SEGMENTATION FROM NATURAL SCENES



วิทยา จิรัฎฐิติเจริญ

WITTAYA JIRATTITICHAREON

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2550

THAI TEXT DETECTION AND SEGMENTATION FROM NATURAL SCENES

WITTAYA JIRATTITICHAREON

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

2007

COPYRIGHT 2007

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

หัวข้อวิทยานิพนธ์	การตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติ
นักศึกษา	นาย วิทยา จิรจิตฺติเจริญ
รหัสนักศึกษา	47066450
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
พ.ศ.	2550
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผศ.ดร. ธนารัตน์ ชลิตาพงศ์

บทคัดย่อ

วิทยานิพนธ์ฉบับนี้ เสนอการตรวจจับและตัดแยกตัวอักษรภาษาไทยจากภาพในฉากธรรมชาติ วิธีการที่เสนอประกอบไปด้วยการเตรียมภาพ JPEG โดยแปลงภาพให้เป็นภาพระดับสีเทา เพื่อลดจำนวนช่องสัญญาณสีและพิจารณาเฉพาะค่าความเข้มแสง จากนั้นใช้วิธีการของเกาส์เซียนเบลตเพื่อกำจัดสัญญาณรบกวน และบล็อกสี่เหลี่ยมที่เกิดในภาพบีบอัด เมื่อผ่านขั้นตอนนี้แล้วสัญญาณรบกวนต่างๆ จะลดลง แต่ความคมชัดของเส้นขอบก็จะลดลงด้วย จึงต้องปรับปรุงภาพให้คมชัดขึ้น โดยใช้ฟิลเตอร์ความคมชัด จากนั้นทำการตรวจจับเส้นขอบโดยใช้ Laplacian of Gaussian (LoG) ซึ่งวิธีการนี้เป็นการหาเส้นขอบแบบอนุพันธ์อันดับสอง ที่ทนทานต่อสัญญาณรบกวน และการเปลี่ยนแปลงของความเข้มแสง ให้เส้นขอบที่บาง จากนั้นทำการตัดครองส่วนที่คาดว่าจะเป็นตัวอักษร โดยการเปรียบเทียบอัตราส่วนของแต่ละบริเวณปิดกับอัตราส่วนที่กำหนดไว้สำหรับตัวอักษรไทย จากนั้นทำการวิเคราะห์ลักษณะการวางตัวของข้อความภาษาไทย ซึ่งได้นำเสนอวิธีการใหม่ในการตรวจสอบกลุ่มตัวอักษรในข้อความได้แม้ว่าข้อความจะมีการเอียงและการถูกรบกวนแบบเปอร์สเปกทีฟ ซึ่งเกิดจากความไม่เหมาะสมของมุมมอง เมื่อได้แนวข้อความแล้วจะทำการกำหนดจุดที่มุมทั้งสี่ของกรอบข้อความ เพื่อเป็นข้อมูลในวิธีการแก้ไขข้อมูลภาพแบบสัมพรรค (Affine Rectification) ซึ่งวิธีการนี้จะแก้ไขระนาบข้อความที่บิดเบี้ยวให้กลับมาอยู่ในระนาบตรง ทำให้การตรวจจับตัวอักษรทำได้ถูกต้องยิ่งขึ้น นอกจากนี้ยังใช้ Gaussian Mixture Model (GMM) เพื่อวิเคราะห์สีในแบบจำลองต่างๆ และหาแบบจำลองที่เหมาะสมในการตัดแยกตัวอักษรออกจากพื้นหลัง สำหรับค่าพารามิเตอร์ของ GMM ได้แก่ค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐาน จะหาโดยใช้ Expectation Maximization (EM) จากการทดลองกับภาพถ่ายจำนวน 200 ภาพ ซึ่งมีจำนวนกลุ่มข้อความทั้งหมด 280 ข้อความและมีจำนวนตัวอักษรทั้งหมด 3,372 ตัว พบว่าระบบสามารถตรวจจับและตัดแยกกลุ่มข้อความได้ถูกต้องคิดเป็น 85.00 เปอร์เซ็นต์ สำหรับการตัดแยกตัวอักษรพบว่า ระบบสามารถการตรวจจับและตัดแยกตัวอักษรได้ถูกต้องคิดเป็น 89.86 เปอร์เซ็นต์

Thesis Title	Thai Text Detection and Segmentation from Natural Scenes
Student	Mr. Wittaya Jiratitichareon
Student ID.	47066450
Degree	Master of Science
Program	Information Technology
Year	2007
Thesis Advisor	Asst.Prof.Dr. Thanarat Chalidabhongse

ABSTRACT

This thesis proposes an approach to detect and segment Thai texts from natural scenes. First, the JPEG color image is converted to gray-scale image in order to reduce color channels and obtain only the intensity value. Then, Gaussian Blur is used to remove noise and artifacts that occur in compression image. Although noise in the image is removed, some image details are also reduced. To enhance image details, we use sharpening filter. In edge detection step, Laplacian of Gaussian (LoG) is used. LoG is the second order edge detector which is tolerant to noise and intensity changing. LoG also yields sharp edges. Each candidate text region is then extracted by comparing its connected component aspect ratio with defined Thai character aspect ratio. We proposed the new approach in Thai text layout analysis. The system can find groups of characters even in rotated or perspective layout caused by an inappropriate camera view angle. Four positions of text region corners are then defined and brought to Affine Rectification which is used to correct text rotation and recover deformation of text regions. The procedure can significantly improve text detection accuracy. Gaussian Mixture Model (GMM) is also employed for color model analysis in order to select an appropriate model to segment characters from background. We used Expectation Maximization (EM) to estimate GMM parameters such as mean and standard deviation. We tested the system on 200 images which have 280 text regions and 3,372 characters. The system can detect and segment text regions with 85.00% accuracy rate. For character segmentation, the system can detect and segment characters with 89.86% accuracy rate.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยดี ด้วยความกรุณาจากอาจารย์ผู้ควบคุมวิทยานิพนธ์ ผศ.ดร. ธนารัตน์ ชลิตาพงศ์ ที่ได้ให้คำปรึกษา และชี้แนะแนวทางในการทำวิจัย ตลอดจนการแก้ปัญหาของงานวิจัยนี้จนสำเร็จ

ขอขอบพระคุณคณาจารย์ ที่คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกๆ ท่าน ที่ได้ประสิทธิ์ประสาทวิชาความรู้ รวมถึงถ่ายทอดประสบการณ์ต่างๆ ให้แก่ข้าพเจ้า

คุณค่า และประโยชน์อันพึงเกิดจากวิทยานิพนธ์ฉบับนี้ ขอมอบแด่บิดา มารดา ผู้คอยดูแล และชี้แนะข้าพเจ้ามาโดยตลอด

วิทยา จิรัจิติเจริญ

สารบัญ (ต่อ)

	หน้า
2.2 งานวิจัยที่เกี่ยวข้อง.....	15
2.2.1 งานวิจัยการค้นหาคำศัพท์ด้วยคุณลักษณะของคำศัพท์.....	15
2.2.2 งานวิจัยการค้นหาคำศัพท์ด้วยคุณลักษณะของคำศัพท์.....	17
2.2.3 งานวิจัยการรู้จำคำศัพท์.....	17
บทที่ 3 การตรวจจับและตัดแยกคำศัพท์ไทยจากภาพในฉากธรรมชาติ.....	21
3.1 กระบวนการตรวจจับและตัดแยกคำศัพท์ไทยจากภาพป้ายในฉากธรรมชาติ.....	21
3.2 การตรวจจับข้อความ (Text Detection).....	22
3.3 การตัดแยกข้อความ (Text Extraction).....	30
บทที่ 4 การทดสอบการตรวจจับและตัดแยกคำศัพท์ไทยจากภาพในฉากธรรมชาติ.....	33
4.1 ชุดข้อมูลที่ใช้ในการทดสอบ.....	33
4.2 การทำงานของระบบการตรวจจับและตัดแยกคำศัพท์ไทยในฉากธรรมชาติ.....	33
4.3 ผลการทดสอบ.....	36
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....	49
5.1 ปัจจัยที่มีผลต่อคุณภาพของภาพป้าย.....	50
5.2 ข้อเสนอแนะในการทำวิจัยต่อไป.....	51
เอกสารอ้างอิง.....	55
ภาคผนวก.....	57
ภาคผนวก ก รูปที่ใช้ในการทดลอง.....	58
ภาคผนวก ข อัตราส่วนต่างๆ ที่ใช้ในการพิจารณารูปแบบของฟอนต์ภาษาไทย.....	63
ภาคผนวก ค ผลงานวิจัยที่ได้รับการตีพิมพ์เผยแพร่.....	71
ภาคผนวก ง โทรศัพท์มือถือรุ่นที่ใช้ถ่ายภาพได้ที่มีความละเอียดอย่างน้อย 640x480 พิกเซล.....	86
ประวัติผู้เขียน.....	87

สารบัญตาราง

ตารางที่	หน้า
3.1 แสดงการจัดกลุ่มตัวอักษร สระ และวรรณยุกต์ในภาษาไทย.....	23
4.1 ผลการทดลองระบบการตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติ กับภาพ ข้อความที่อยู่ในแนวนอนปกติ.....	37
4.2 ผลการทดลองระบบการตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติ กับภาพ ข้อความที่มีการหมุนไปจากแนวนอนปกติ.....	39
4.3 ผลการทดลองระบบการตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติ กับภาพ ข้อความที่มีการลู่เข้าแบบเปอร์สเปคทีฟ.....	41
4.4 เปรียบเทียบผลการทดสอบการตรวจจับตัวอักษร กับภาพถ่ายจำนวน 200 ภาพ.....	45
4.5 เปรียบเทียบผลการทดสอบการตัดแยกตัวอักษร ในแบบจำลองสี่ต่างๆ กับภาพถ่ายจำนวน 200 ภาพ.....	45
4.6 ผลการทดสอบระบบกับภาพถ่ายจำนวน 200 ภาพ.....	46
4.7 แสดงตัวอย่างของการตรวจจับและตัดแยกตัวอักษรที่ผิดพลาด.....	47

สารบัญรูป

รูปที่	หน้า
2.1 แสดงการกระจายแบบเกาส์เซียนในสองมิติ ที่มีค่าเฉลี่ย (Mean) ที่ $(0, 0)$ และ $\sigma = 1$	6
2.2 หน้ากากเกาส์เซียน (Gaussian Mask) ขนาด 3×3	7
2.3 ตัวกรองที่ทำให้ภาพคมชัด (Sharpening Filter).....	7
2.4 กราฟแสดงลักษณะการกระจายของข้อมูลแบบปกติ.....	8
2.5 ลักษณะของการบิดเบี้ยวภายใน (Internal Distortion).....	11
2.6 ลักษณะของการบิดเบี้ยวภายนอก (External Distortion).....	11
2.7 แสดงรูปอาคารที่มีการลู่อแบบเปอร์สเปกทีฟ (ก) เปอร์สเปกทีฟ 1 จุด (ข) เปอร์สเปกทีฟ 2 จุด (ค) เปอร์สเปกทีฟ 3 จุด.....	12
2.8 การสร้างเส้นแนวสายตาและเส้นที่ลากต่อไปจนบรรจบเพื่อหาจุดวานิชซึ่งในภาพอาคาร...	13
2.9 แสดงภาพป้ายข้อความ (ก) ป้ายภาษาอังกฤษ (ข) ป้ายภาษาจีน (ค) ป้ายภาษาไทย.....	13
2.10 การกำหนดค่าต่างๆ ในการวิเคราะห์เลขเอนท์ของข้อความในป้าย.....	14
3.1 แสดงกระบวนการตรวจจับและตัดแยกตัวอักษรภาษาไทยจากภาพในฉากธรรมชาติ.....	21
3.2 แสดงเส้นระดับของข้อความในภาษาไทย.....	24
3.3 แสดงภาพป้ายข้อความ (ก) ข้อความมีลักษณะการลู่อแบบเปอร์สเปกทีฟ (ข) ข้อความมีการหมุนไปจากแนวปกติ.....	25
3.4 การหาแนวข้อความโดยพิจารณาจากสมการเส้นตรง (ก) กรณีที่ข้อความมีการลู่อแบบเปอร์สเปกทีฟ (ข) กรณีที่ข้อความมีการหมุน.....	25
3.5 ตัวอย่างของการพิจารณาคู่ของกรอบตัวอักษร เพื่อหาสมการเส้นตรง.....	26
3.6 การหากรอบข้อความในกรณีที่ข้อความมีการหมุนและการลู่อแบบเปอร์สเปกทีฟ (ก) การกำหนดจุดต่างๆ รอบกลุ่มตัวอักษร (ข) การกำหนดจุด F_{mid} และ B_{mid} (ค) กรอบข้อความเมื่อแก้ไขแนวระย้าแล้ว.....	28
3.7 แสดงการสร้างกรอบล้อมรอบข้อความ (ก) ภาพต้นฉบับ (ข) ตรวจจับกลุ่มตัวอักษรและทำการล้อมกรอบข้อความ.....	29
3.8 ภาพข้อความที่ได้ทำการขยายกรอบแล้ว.....	30
3.9 แสดงค่าการกระจายของสีในแต่ละแบบจำลอง (R, G, B และ I) ของตัวอักษร “ร” “อ” และ “ช” จากป้ายที่ถ่ายจากสถานที่ และเวลาที่แตกต่างกัน.....	30
3.10 การแบ่งพื้นที่ในกรอบข้อความ.....	31
4.1 แสดงขั้นตอนการทำงานของระบบตรวจจับและตัดแยกตัวอักษรไทยในฉากธรรมชาติ.....	34

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.2 ผลลัพธ์ที่ตรวจจับและตัดแยกตัวอักษรภาษาไทยจากรูปที่ 4.1.....	36
5.1 ป้ายที่ระบบไม่สามารถตรวจจับข้อความได้.....	51
5.2 ตัวอย่างการรู้จำตัวอักษรไทยจากป้ายในฉากธรรมชาติ โดยใช้ซอฟต์แวร์ “อ่านไทย”.....	52
5.3 ตัวอย่างการรู้จำตัวอักษรไทยจากป้ายในฉากธรรมชาติ โดยใช้ซอฟต์แวร์ “อ่านไทย”.....	53
5.4 แสดงภาพที่มีขอบเขตป้ายชัดเจน.....	54

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันข้อมูลสารสนเทศมีความใกล้ชิดกับเรามาก ตัวอย่างหนึ่งของข้อมูลสารสนเทศที่เราพบเห็นได้ทั่วไปในชีวิตประจำวันคือ ข้อความในสถานที่ต่างๆ เช่น ข้อความท้องถนน ข้อความบอกสถานที่ ข้อความโฆษณาประชาสัมพันธ์ ข้อความเตือน ตลอดจนข้อความที่เป็นข้อมูลอื่นๆ ที่สามารถพบเห็นได้ในสิ่งแวดล้อมธรรมชาติ ข้อความที่ปรากฏจะให้ข้อมูลสารสนเทศอันเป็นประโยชน์แก่ผู้พบเห็น อย่างไรก็ตาม หากผู้พบเห็นข้อความไม่สามารถทราบข้อมูลจากข้อความที่มีการแจ้งข้อมูลสำคัญ ก็อาจเป็นอันตราย หรืออาจจะประสบอุบัติเหตุได้ ตัวอย่างเช่น นักท่องเที่ยวชาวต่างประเทศมาเที่ยวในเมืองไทย เขาเหล่านั้นไม่สามารถอ่านหรือทำความเข้าใจข้อความที่แจ้งเตือนถึงอันตราย เนื่องจากอ่านภาษาไทยไม่ได้ ซึ่งถ้ามีระบบการรู้จำตัวอักษรและแปลข้อความจากภาพ ก็จะสามารถช่วยชาวต่างชาติเหล่านั้นให้สามารถเข้าใจความหมายของข้อความนั้นๆ ได้

ในวิทยานิพนธ์ฉบับนี้ ได้วิจัยและพัฒนากระบวนการการตรวจจับและตัดแยกตัวอักษรจากภาพในฉากธรรมชาติ ฉากธรรมชาติในที่นี้หมายถึง สิ่งแวดล้อมจริงที่อยู่รอบๆ ข้อความ และข้อความอาจจะอยู่ในกรอบ เช่น ป้ายทางหลวง หรือเป็นข้อความที่ไม่อยู่ในกรอบ เช่น ข้อความที่ติดอยู่บนอาคาร เป็นต้น ซึ่งสิ่งแวดล้อมในธรรมชาติมักมีผลกระทบโดยตรงต่อคุณภาพของภาพถ่ายและความคมชัดของตัวอักษร ซึ่งเป็นการยากที่เราจะควบคุมสิ่งแวดล้อมให้เป็นไปดังที่เราต้องการ

การตรวจจับข้อความในภาพถ่ายจากสิ่งแวดล้อมธรรมชาตินั้น เป็นปัญหาที่ยาก และมีความท้าทายหลายประการ สิ่งแรกคือความหลากหลายของตัวอักษร ทั้งความแตกต่างของรูปแบบตัวอักษร ขนาด มุมมองและตำแหน่งของตัวอักษร นอกจากนี้ภาพข้อความตัวอักษรยังไม่อยู่ในลักษณะที่สมบูรณ์ อาจเนื่องจากการเคลื่อนที่ของอุปกรณ์ถ่ายภาพและทำการถ่ายภาพ หรือตัวอักษรถูกบดบังด้วยวัตถุอื่นๆ ที่อยู่ในฉากนั้น ในบางครั้งตัวอักษรที่ปรากฏมีลักษณะเป็นวัตถุสามมิติ ดังนั้นตัวอักษรก็มีการบิดเบี้ยวเนื่องจากวัสดุ หรือมีการลู่ตามลักษณะการวางของมุมมอง ความสกรปรกของบริเวณตัวอักษรก็เป็นอีกสิ่งหนึ่งที่ทำให้การตรวจจับข้อความทำได้ยาก นอกจากนี้ การสะท้อนของแสงแฟลช แสงจากหลอดไฟ และแสงสะท้อน (glare) จากดวงอาทิตย์ ต่างส่งผลต่อความคมชัดของข้อความบนภาพที่ถ่ายมาได้ทั้งสิ้น

การตรวจจับและตัดแยกตัวอักษรไทยในฉากธรรมชาติจะมีความแตกต่างกับการตรวจจับและตัดแยกตัวอักษรในระบบการรู้จำตัวอักษร (OCR) ก่อนข้างมาก โดยปกติแล้วโปรแกรมรู้จำตัวอักษรทั่วไปในห้องคลาดจะทำงานได้ดีมากกับภาพที่ได้จากการนำเอกสารมาสแกนด้วยเครื่องสแกนตั้งโต๊ะ เนื่องจากเราสามารถควบคุมการทำงานของระบบดังกล่าวได้เป็นอย่างดี ภาพสแกน

ของเอกสารที่ได้มักจะมีควมคมชัดมาก มีสัญญาณรบกวน (Noise) น้อย และสิ่งแวดล้อมไม่มีผลต่อคุณภาพของภาพสแกนเนื่องจากทำในระบบปิด แต่เมื่อนำระบบนี้มาใช้กับการรู้จำตัวอักษรในภาพถ่ายคุณภาพต่ำ ทำให้ผลการรู้จำผิดพลาดมาก เพราะคุณภาพของข้อมูลยังไม่เพียงพอที่จะทำให้ระบบสามารถรู้จำได้อย่างถูกต้องแม่นยำ จากการศึกษางานวิจัยของ Chen และคณะ [1] ซึ่งเป็นงานวิจัยเกี่ยวกับการตรวจจับและตัดแยกตัวอักษรออกจากภาพของฉากรวมชาติได้เสนอวิธีการรู้จำตัวอักษรภาษาจีนในภาพถ่ายฉากรวมชาติ ซึ่งผลการตรวจจับและตัดแยกให้ความถูกต้องแม่นยำสูง วิทยานิพนธ์นี้เสนอวิธีการตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากรวมชาติ ซึ่งคล้ายคลึงกับวิธีการของ Chen แต่อย่างไรก็ตาม เมื่อทดลองใช้วิธีการของ Chen กับตัวอักษรภาษาไทยแล้วได้ผลไม่ดีนัก เนื่องจากข้อความภาษาไทยมีรูปแบบตัวอักษรต่างจากภาษาจีน นอกจากนี้ภาษาไทยยังมีความซับซ้อนในการสร้างข้อความ เช่น การวางสระและวรรณยุกต์ในตำแหน่งบนและล่างข้อความ ดังนั้นจึงต้องมีการปรับปรุงวิธีการให้เหมาะสมกับตัวอักษรภาษาไทย ระบบจะต้องมีความคงทนต่อขนาดตัวอักษร สิ่งรบกวนต่างๆ ในภาพ รวมถึงการถูเข้าอันเนื่องมาจากมุมมองของการถ่ายภาพตัวอักษรในฉากรวมชาติ

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

วิทยานิพนธ์ฉบับนี้มีวัตถุประสงค์เพื่อเสนอวิธีตรวจจับและตัดแยกข้อความภาษาไทยจากภาพในฉากรวมชาติ ที่มีความทนทานต่อการเปลี่ยนแปลงลักษณะ ขนาด การบิดเบี้ยวเนื่องจากมุมมอง และการจัดกลุ่มของตัวอักษร เพื่อใช้ตรวจจับและตัดแยกตัวอักษรที่อยู่ในฉากรวมชาติได้อย่างถูกต้อง

1.3 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย

วิทยานิพนธ์ฉบับนี้ใช้วิธีการประมวลผลภาพ เพื่อตรวจจับและตัดแยกตัวอักษรจากภาพในฉากรวมชาติ โดยภาพตัวอักษรที่จะนำมาพิจารณาเพื่อหาขอบเขตข้อความและตัดแยกตัวอักษร จะต้องอยู่บนพื้นฐาน ดังต่อไปนี้

- 1) พื้นหลังของตัวอักษรเป็นสีพื้นเพียงสีเดียว ไม่มีลวดลาย
- 2) ข้อความเป็นแนวตรง ไม่ถูกออกแบบมาในลักษณะที่บิดเบี้ยวโค้งงอเพื่อความสวยงาม
- 3) ตัวอักษรถูกออกแบบให้มีความแตกต่างจากพื้นหลังทั้งสีและความเข้มแสง
- 4) ตัวอักษรแต่ละตัวถูกประกอบกันขึ้นเป็นข้อความที่มีขอบเขตชัดเจน
- 5) ตัวอักษรมีรูปแบบ (Font) ที่เป็นลักษณะของตัวพิมพ์ ไม่ได้ถูกเขียนขึ้นด้วยลายมือ
- 6) ตัวอักษรในบริบท (Context) ของข้อความเดียวกันจะมีขนาดเท่ากัน มีความเข้มแสงที่ใกล้เคียงกัน แต่หากเป็นคนละข้อความ สีของตัวอักษรอาจจะแตกต่างกันได้

สำหรับกระบวนการการตรวจจับและตัดแยกตัวอักษรจากภาพในฉากรวมชาติ มีขั้นตอนทั้งหมด 8 ขั้นตอนคือ

ขั้นตอนที่ 1 ทำการเตรียมภาพ โดยกำจัดสัญญาณรบกวน และปรับปรุงภาพเพื่อให้เส้นขอบมีความคมชัดขึ้น (Sharpening) จากนั้นทำแปลงภาพเป็นภาพระดับสีเทา (Gray-Scale Image)

ขั้นตอนที่ 2 ทำการตรวจจับเส้นขอบ (Edge Detection) ของวัตถุต่างๆ ในภาพ โดยใช้ LoG (Laplacian of Gaussian) เนื่องจากเป็นวิธีการวิเคราะห์เส้นขอบที่พิจารณาการเปลี่ยนแปลงแบบอนุพันธ์อันดับ 2 จึงทำให้ได้เส้นที่มีความน่าเชื่อถือ และทนทานต่อสัญญาณรบกวนที่อยู่ในภาพ

ขั้นตอนที่ 3 เมื่อได้เส้นขอบแล้ว จะทำการค้นหาส่วนประกอบที่เชื่อมติดกัน (Connected Component Labeling) เพื่อนำมาวิเคราะห์ว่าบริเวณใดบ้างที่มีโอกาสเป็นกลุ่มข้อความ และบริเวณใดไม่ใช่ข้อความ โดยนำส่วนประกอบนี้มาพิจารณาเปรียบเทียบกับค่าที่กำหนดไว้แล้วสำหรับตัวอักษรไทย ได้แก่ อัตราส่วนความกว้างต่อความสูง และอัตราส่วนพื้นที่ภายในเส้นขอบ (พื้นที่ตัวอักษร) ต่อพื้นที่สี่เหลี่ยมปิดล้อมเส้นขอบนั้น

ขั้นตอนที่ 4 ทำการวิเคราะห์รูปแบบการวางตัวของข้อความ หาแนวของข้อความว่าวางตัวแบบปกติ มีการหมุน หรือมีการถูเข้าแบบเปอร์สเปคทีฟ การหาแนวของข้อความจะพิจารณาจากจุดสูงสุดและต่ำสุดของแต่ละตัวอักษรที่อยู่ใกล้เคียงกัน เพื่อใช้ในการคำนวณหาค่าความชันเส้นขอบบรรทัดบนและล่าง เพื่อให้สามารถหากลุ่มตัวอักษรที่อยู่ในข้อความเดียวกัน และสร้างกรอบล้อมรอบข้อความได้อย่างถูกต้อง

ขั้นตอนที่ 5 ในขั้นตอนนี้จะทำการแก้ไขการถูเข้าแบบเปอร์สเปคทีฟและการหมุนของข้อความในกรอบที่หาได้ โดยใช้วิธีการแก้ไขแบบสัมพรรค (Affine Rectification) เพื่อแก้ไขภาพตัวอักษรที่ผิดเพี้ยนไปให้อยู่ในมุมมองแนวหน้าตรง ซึ่งจะช่วยให้หาขอบเขตของข้อความได้อย่างถูกต้องยิ่งขึ้น

ขั้นตอนที่ 6 เมื่อได้ภาพข้อความที่เป็นมุมมองหน้าตรงแล้ว จะใช้ขอบเขตของเส้นขอบที่เชื่อมติดกัน ที่ผ่านการกรองในขั้นตอนที่ 2 นำมาหาคุณสมบัติการกระจายตัวของสีและเปรียบเทียบค่าในแต่ละแบบจำลอง ได้แก่ R G B และ I เหตุที่เลือกใช้แบบจำลองเหล่านี้เพราะสามารถทำได้อย่างรวดเร็ว แยกความแตกต่างของพื้นหน้าและพื้นหลังได้ง่าย สะดวกต่อการตัดแยกตัวอักษรออกจากภาพฉากหลัง โดยเลือกใช้แบบจำลองผสมเกาส์เซียน (Gaussian Mixture Model - GMM) เพื่อเปรียบเทียบค่ากลางของการกระจายตัวขององค์ประกอบแต่ละตัว และใช้อัลกอริทึมของ Otsu [21] เพื่อกำหนดค่าเทรชโฮลด์ (Threshold) ที่เหมาะสมต่อการตัดแยกตัวอักษรออกจากฉากภาพ

ขั้นตอนที่ 7 เป็นขั้นตอนในการพิจารณาภาพโดยละเอียด เพื่อลดสัญญาณรบกวนและคัดเลือกรตัวอักษรให้ถูกต้องมากขึ้น โดยเริ่มจากการนำภาพขาวดำไปวิเคราะห์หาส่วนประกอบที่เชื่อมติดกัน โดยการพิจารณาให้ความสำคัญกับส่วนประกอบตามขอบเขตพื้นที่

ขั้นตอนสุดท้าย จะนำภาพขาวดำที่มีข้อมูลการวิเคราะห์ส่วนประกอบที่เชื่อมติดกันแล้วในขั้นตอนก่อนหน้านี้มาพิจารณาตัดแยกข้อความออกจากพื้นหลัง ซึ่งข้อมูลดังกล่าวประกอบด้วยข้อมูลพิกเซล ซึ่งเมื่อเสร็จขั้นตอนนี้ ผลลัพธ์ที่ได้จะเป็นส่วนของภาพพยัญชนะ สระ และวรรณยุกต์ไทยที่เป็นส่วนประกอบของข้อความถูกตัดแยกเป็นส่วนๆ

1.4 ขอบเขตการวิจัย

งานวิจัยนี้ เสนอวิธีการตรวจจับและตัดแยกข้อความภาษาไทยในฉากธรรมชาติ โดยมีขอบเขตการวิจัยดังนี้

- 1.4.1 ภาพข้อความที่ใช้ในการทดลองมีขนาด 640x480 พิกเซล
- 1.4.2 ภาพที่ถ่ายมานั้น มีระยะ และมุมมองการถ่ายภาพที่จำกัด ช่วงค่าหนึ่ง
- 1.4.3 พื้นหลังของตัวอักษรเป็นสีพื้นเพียงสีเดียว ไม่มีลวดลาย
- 1.4.4 ข้อความในภาพเป็นแนวตรง ไม่ถูกออกแบบมาในลักษณะที่บิดเบี้ยวโค้งงอเพื่อความสวยงาม
- 1.4.5 ตัวอักษรถูกออกแบบให้มีความแตกต่างจากพื้นหลังทั้งสีและความเข้มแสง
- 1.4.6 ตัวอักษรแต่ละตัวถูกประกอบกันขึ้นเป็นข้อความที่มีขอบเขตชัดเจน
- 1.4.7 ตัวอักษรมีรูปแบบ (Font) ที่เป็นลักษณะของตัวพิมพ์ ไม่ใช่ลายมือเขียน
- 1.4.8 ตัวอักษรในบริบท (Context) ของข้อความเดียวกันจะมีขนาดเท่ากัน มีความเข้มแสงที่ใกล้เคียงกัน แต่หากเป็นคนละข้อความ สีของตัวอักษรอาจจะแตกต่างกันได้
- 1.4.9 ผลลัพธ์ที่ได้จากระบบเป็นข้อมูลภาพตัวอักษร ที่ถูกตัดแยกไว้เป็นส่วนๆ
- 1.4.10 งานวิจัยนี้ทำเฉพาะส่วนของการตรวจจับและตัดแยกตัวอักษรภาษาไทยออกมาจากภาพถ่ายในฉากธรรมชาติ

1.5 ขั้นตอนการศึกษา

- 1.5.1 ศึกษางานวิจัยต่างๆ ที่เกี่ยวข้อง
- 1.5.2 ศึกษาลักษณะและรูปแบบการเขียนตัวอักษรและข้อความภาษาไทย
- 1.5.3 เก็บภาพตัวอักษรตัวอย่างที่จะนำมาใช้ในการทดลอง
- 1.5.4 ออกแบบอัลกอริทึมและเขียน โปรแกรมในการตรวจจับตัวอักษร
- 1.5.5 ออกแบบอัลกอริทึมและเขียน โปรแกรมในการจัดกลุ่มของตัวอักษร
- 1.5.6 ออกแบบอัลกอริทึมและเขียน โปรแกรมในการแก้ไขภาพเนื่องจากการลู่เข้า
- 1.5.7 บันทึกและสรุปผลการทดลอง
- 1.5.8 รวบรวมข้อมูลจัดทำเอกสารเพื่อนำเสนอเป็นวิทยานิพนธ์

1.6 ประโยชน์ที่คาดว่าจะได้รับ

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัยนี้ มีดังนี้

- 1.6.1 ทำให้ทราบถึงคุณลักษณะของตัวอักษรภาษาไทย เพื่อใช้เป็นพื้นฐานของการตรวจจับและจัดกลุ่มข้อความในภาษาไทย
- 1.6.2 จากการค้นคว้าจากงานวิจัยต่างๆ ทำให้ทราบวิธีการที่หลากหลาย ที่ใช้ในการค้นหาตัวอักษรในภาพ และสามารถปรับปรุงให้ได้วิธีการใหม่ที่มีความทนทานต่อสภาพแวดล้อมในฉากภาพ สัญญาณรบกวน รวมถึงการบิดเบี้ยวเนื่องจากมุมกล้อง
- 1.6.3 สามารถนำวิทยานิพนธ์นี้ประยุกต์ใช้ในระบบรู้จำตัวอักษรภาษาไทยในภาพถ่ายจากฉากรถรรมชาติ ที่ต้องการความรวดเร็วหรือมีข้อจำกัดในการประมวลผล เช่น ระบบอ่านข้อความบอกทางบนถนนที่ประมวลผลในเวลาจริง ระบบการรู้จำตัวอักษรไทยในอุปกรณ์มือถือ เช่น โทรศัพท์มือถือ คอมพิวเตอร์พกพา เป็นต้น

บทที่ 2

ทฤษฎีพื้นฐาน และงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีพื้นฐานต่างๆ และงานวิจัยที่เกี่ยวข้องในการทำวิจัยนี้

2.1 ทฤษฎีพื้นฐาน

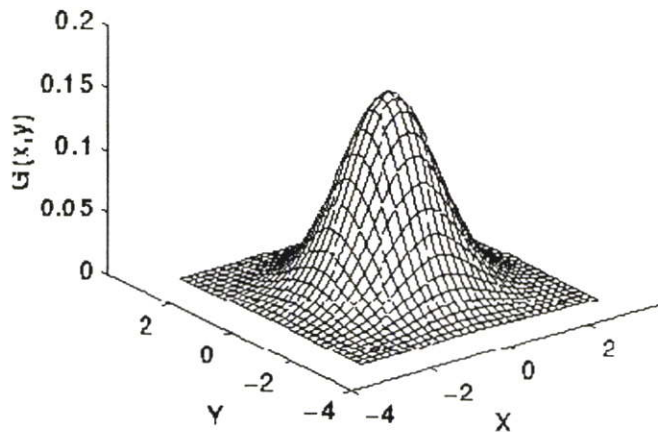
2.1.1 การเบลอแบบเกาส์เซียน (Gaussian Blur)

การเบลอแบบเกาส์เซียน (Gaussian Blur) หรือ ตัวกระทำทำให้เรียบแบบเกาส์เซียน (Gaussian Smoothing Operator) [25] เป็นตัวกระทำคอนโวลูชัน (Convolution) แบบสองมิติ ซึ่งถูกนำมาใช้เพื่อให้ภาพเบลอ โดยจุดประสงค์เพื่อลดสัญญาณรบกวนและรายละเอียดในภาพ โดยสามารถเขียนการกระจายเกาส์เซียนในสองมิติ ได้ดังนี้

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.1)$$

เมื่อ σ คือค่าความเบี่ยงเบนมาตรฐาน (Standard Deviation) ของการกระจาย ค่า x และ y คือค่าในแนวแกน x และ y ตามลำดับ

สามารถแสดงให้เห็นถึงการกระจายแบบสองมิติ ดังรูปที่ 2.1



รูปที่ 2.1 แสดงการกระจายแบบเกาส์เซียนในสองมิติ ที่มีค่าเฉลี่ย (Mean) ที่ $(0, 0)$ และ $\sigma = 1$

แนวความคิดการใช้การเบลอแบบเกาส์เซียน ก็จะใช้ในการทำคอนโวลูชัน เมื่อพิจารณารูปภาพดิจิทัลจะพบว่าเป็นข้อมูลสองมิติที่ประกอบไปด้วยพิกเซลย่อยๆ ซึ่งเป็นข้อมูลดิสครีต ดังนั้นจึงต้องประมาณค่าฟังก์ชันเกาส์เซียนก่อนที่จะกระทำการคอนโวลูชัน ในทางทฤษฎี ค่าการ

กระจายเกาส์เซียนจะไม่เป็นศูนย์ในทุกๆ ที่ ทำให้ต้องใช้เคอร์เนลการคอนโวลูชันที่ใหญ่มาก แต่ในทางปฏิบัติจะใช้เคอร์เนลที่มีค่าจำนวนเต็ม ที่ประมาณค่าเกาส์เซียน ด้วย $\sigma = 1.0$ สำหรับขนาดหน้าต่างของเคอร์เนลอาจมีขนาดตั้งแต่ 3×3 5×5 หรือ 7×7 ในงานวิจัยนี้เลือกใช้ขนาด 3×3 เนื่องจากสามารถลดสัญญาณรบกวนได้ดี และใช้เวลาในการคำนวณคอนโวลูชันน้อยกว่าเมื่อเทียบกับการใช้หน้าต่างขนาดอื่นๆ

	1	2	1
1/16	2	4	2
	1	2	1

รูปที่ 2.2 หน้ากากเกาส์เซียน (Gaussian Mask) ขนาด 3×3

เมื่อได้หน้ากากเกาส์เซียน ดังรูปที่ 2.2 แล้ว จะนำไปทำการคอนโวลูชันกับภาพ ผลที่ได้คือภาพที่เบลอมากขึ้น และสัญญาณรบกวนต่างๆ ในภาพลดลง

2.1.2 การทำภาพให้คมชัด (Sharpening)

การลดสัญญาณรบกวนในภาพด้วยการทำให้ภาพเบลอนั้น มักจะทำให้รายละเอียดต่างๆ ในภาพเบลอลดหายไปด้วย ดังนั้นก่อนจะนำภาพไปประมวลผลในกระบวนการต่อไป จึงควรปรับปรุงให้ภาพมีความคมชัดก่อน ด้วยเทคนิคการทำภาพให้คมชัด [25]

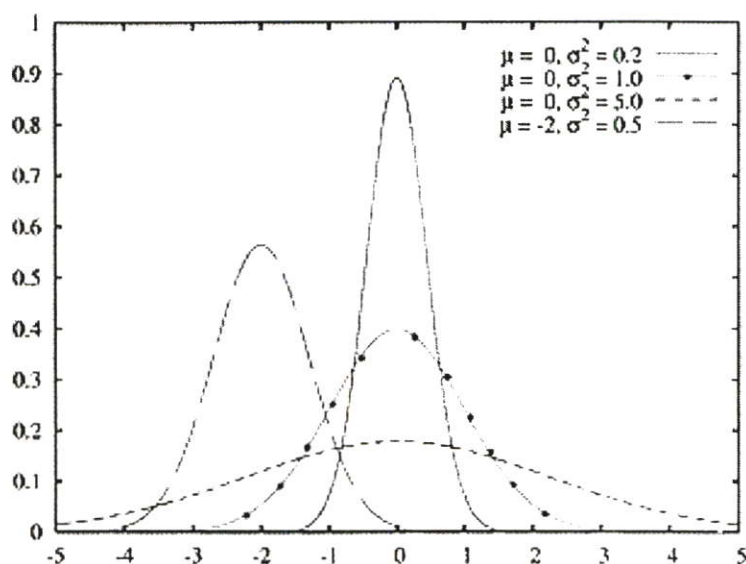
แนวคิดของการทำภาพให้คมชัดจะคล้ายกับวิธีการเบลอภาพแบบเกาส์เซียน กล่าวคือ จะใช้ตัวกรองให้คมชัด (Sharpening Filter) ซึ่งก็คือหน้ากากสำหรับทำคอนโวลูชันนั่นเอง

	-1	-2	-1
1/16	-2	12	-2
	-1	-2	-1

รูปที่ 2.3 ตัวกรองที่ทำให้ภาพคมชัด (Sharpening Filter)

เมื่อนำหน้ากากนี้ไปใช้ในการทำคอนโวลูชันกับภาพที่เบลอ ผลลัพธ์ที่ได้คือภาพที่มีความคมชัดขึ้นทั้งในรายละเอียดและเส้นขอบภาพ

2.1.3 แบบจำลองการกระจายแบบผสม (Mixture Distribution Model)



รูปที่ 2.4 กราฟแสดงลักษณะการกระจายของข้อมูลแบบปกติ

แบบจำลองการกระจายแบบผสม [2, 3, 4] ในทางสถิติแล้วเป็นการกระจายของความ เป็นไปได้ที่ประกอบขึ้นจากการกระจายของความเป็นไปได้อย่างหลายๆ การกระจาย ซึ่งสามารถ เขียนสมการพารามตริกได้ตามสมการที่ (2.2)

$$f_X(x) = \sum_{i=1}^n a_i f_Y(x; \theta_i) \quad (2.2)$$

โดยหากพิจารณาตัวแปรสุ่มแบบไม่ต่อเนื่อง ให้ x เป็นการผสมของตัวแปรสุ่มแบบไม่ ต่อเนื่องของตัวแปร Y ที่ประกอบด้วย การกระจายจำนวน n กลุ่ม ให้ $f_X(x)$ คือฟังก์ชันความ เป็นไปได้ของประชากร (Probability Mass Function) ของ X ที่เกิดจากผลรวมทั้งหมดของการ กระจายย่อย ให้ a_i เป็นสัดส่วนขององค์ประกอบ i และกำหนดให้ $0 < a_i < 1$ สำหรับ $i=1, \dots, n$ และ $a_1 + \dots + a_n = 1$ โดย θ_i เป็นพารามิเตอร์ที่ไม่ทราบค่า

2.1.4 การหาค่าพารามิเตอร์ในแบบจำลองการกระจายแบบผสมด้วยวิธี Expectation

Maximization

อัลกอริทึม Expectation Maximization (EM) [5, 6] ถูกนำมาใช้ในการหาค่าพารามิเตอร์ ของกลุ่มข้อมูลในแบบจำลองการกระจายแบบผสม เช่น ค่าเฉลี่ย (Mean) และค่าความแปรปรวน (Variance) จะทำโดยการสุ่มค่า การคำนวณค่าด้วยอัลกอริทึม EM นี้ประกอบด้วยสองขั้นตอนคือ ขั้นตอน Expectation และ Maximization

2.1.4.1 ขั้นตอน Expectation เริ่มจากการคำนวณความเป็นสมาชิกบางส่วน (Partial Membership) ของข้อมูลแต่ละตัวในการกระจายแต่ละส่วน โดยการคำนวณค่าความคาดหวัง (Expectation Values) สำหรับตัวแปรสมาชิกของข้อมูลแต่ละตัว เช่น สำหรับข้อมูล x_j และการกระจาย Y_i จะสามารถคำนวณค่าความเป็นสมาชิก $y_{i,j}$ ได้ดังสมการ (2.3)

$$y_{i,j} = \frac{a_i f_Y(x_j; \theta_i)}{f_X(x_j)} \quad (2.3)$$

2.1.4.2 ขั้นตอน Maximization เมื่อได้ค่า Expectation มาแล้ว ในขั้นต่อไปจะทำการคำนวณหาค่าสัมประสิทธิ์การผสม (Mixing Coefficient) a_i ซึ่งก็คือค่าเฉลี่ย (Mean) ของค่าความเป็นสมาชิกของข้อมูลทั้งหมด

$$a_i = \frac{1}{N} \sum_{j=1}^N y_{i,j} \quad (2.4)$$

ค่าพารามิเตอร์ θ_i สามารถคำนวณได้จากวิธีการ Expectation Maximization เช่นกัน โดยใช้ข้อมูล x_j ที่ถูกถ่วงน้ำหนักแล้วโดยใช้ค่าความเป็นสมาชิก ตัวอย่างเช่น ถ้าค่า θ คือค่า μ

$$\mu_i = \frac{\sum_j y_{i,j} x_j}{\sum_j y_{i,j}} \quad (2.5)$$

เพื่อคำนวณหาค่า a_i และ θ_i ใหม่อีกครั้ง เราจะกลับไปยังขั้นแรกคือขั้น Expectation แล้วคำนวณค่าความเป็นสมาชิกใหม่ กระบวนการจะถูกทำซ้ำอยู่อย่างนี้จนกว่าค่าพารามิเตอร์ในแบบจำลองผสมนี้จะลู่เข้าถึงค่าหนึ่งหรือไม่มีการเปลี่ยนแปลงอีก

2.1.5 อัลกอริทึมสำหรับหาค่าเทรชโฮลด์ของ Otsu (Otsu Thresholding Algorithm)

Otsu [21] ได้เสนออัลกอริทึมสำหรับการเลือกเทรชโฮลด์จากฮิสโตแกรมของภาพระดับสีเทา โดยเริ่มพิจารณาจากการทำอมัลไลซ์ฮิสโตแกรมให้อยู่ในรูปฟังก์ชันค่าการกระจายที่เป็นไปได้แบบคิสคริต (Discrete Probability Density Function)

$$p_q = \frac{n_q}{N} \quad q = 0, 1, 2, \dots, L-1 \quad (2.6)$$

เมื่อ N เป็นจำนวนรวมของพิกเซลทั้งหมดในภาพ n_q เป็นจำนวนพิกเซลที่มีระดับค่าความเข้มแสง ณ ตำแหน่ง q และ L เป็นจำนวนของระดับค่าความเข้มแสงทั้งหมดที่เป็นไปได้ในภาพ ถ้าเลือกค่า k มาใช้เป็นค่าเทรชโฮลด์เพื่อแบ่งกลุ่มฮิสโตแกรมออกเป็นสองกลุ่มคือ C_0 และ C_1 โดย C_0 เป็นเซตของพิกเซลที่อยู่ในช่วงค่า $[0, 1, \dots, k-1]$ และ C_1 เป็นเซตของพิกเซลที่อยู่ในช่วงค่า $[k, k+1, \dots, L-1]$ วิธีการของ Otsu จะเลือก k ที่ให้ค่าความแปรปรวนระหว่างกลุ่ม (between-class variance, σ_B^2) มีค่ามากที่สุด ซึ่ง σ_B^2 กำหนดโดยสมการ (2.7)

$$\sigma_B^2 = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 \quad (2.7)$$

เมื่อ

$$\omega_0 = \sum_{q=0}^{k-1} p_q \quad (2.8)$$

$$\omega_1 = \sum_{q=k}^{L-1} p_q \quad (2.9)$$

$$\mu_0 = \sum_{q=0}^{k-1} qp_q / \omega_0 \quad (2.10)$$

$$\mu_1 = \sum_{q=k}^{L-1} qp_q / \omega_1 \quad (2.11)$$

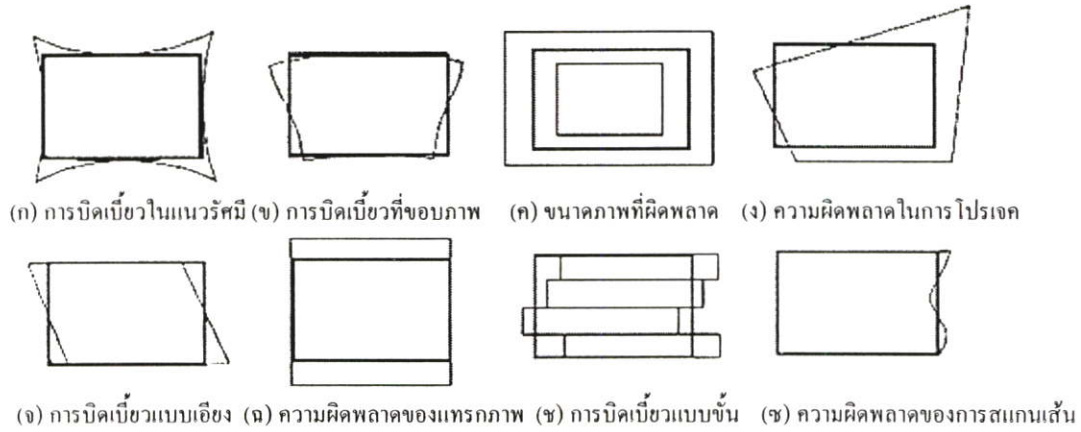
$$\mu_T = \sum_{q=0}^{L-1} qp_q \quad (2.12)$$

2.1.6 การบิดเบี้ยวทางเรขาคณิตของภาพ (Geometric Distortions of the Image)

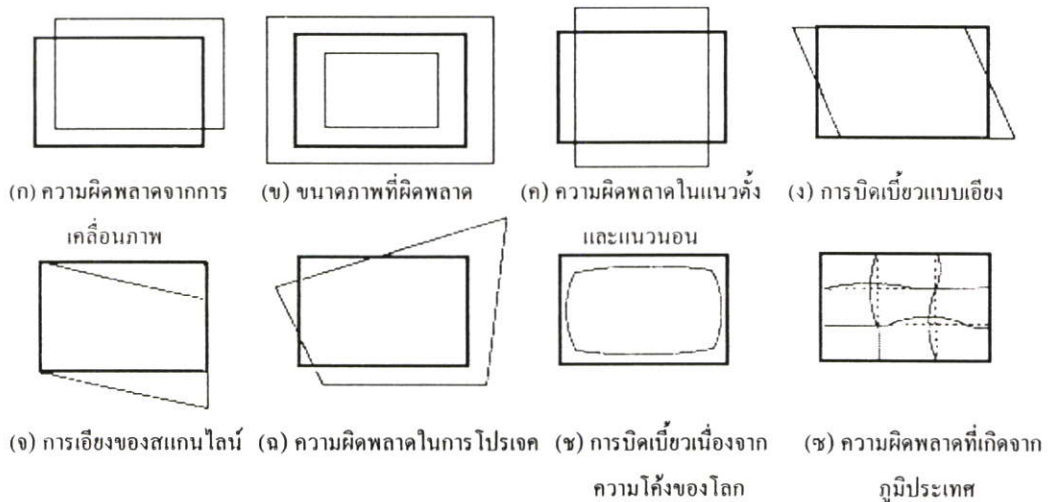
การบิดเบี้ยวทางเรขาคณิตของภาพ [24] เป็นความผิดพลาดที่อาจเกิดขึ้นได้ในขณะถ่ายภาพ ตามทฤษฎีแล้วหากอยู่ภายใต้เงื่อนไขในอุดมคติ พิกัดภาพจะถูกโปรเจกต์อย่างถูกต้อง แต่ในความเป็นจริงพบว่าพิกัดภาพที่ได้จริงกับอุดมคติมีความแตกต่างกัน ลักษณะของการบิดเบี้ยวทางเรขาคณิต แบ่งได้ออกเป็น 2 ประเภท คือ การบิดเบี้ยวภายใน (Internal Distortion) และการบิดเบี้ยวภายนอก (External Distortion)

การบิดเบี้ยวภายใน (Internal Distortion) มีผลมาจากเรขาคณิตของตัวรับภาพ (Sensor) ในทางทฤษฎี ตัวรับภาพในอุดมคติจะถูกโปรเจกต์ไปยังพิกัดภาพได้อย่างถูกต้องและแม่นยำ แต่ในความเป็นจริงนั้นตัวรับภาพและขบวนการรับภาพอาจมีข้อผิดพลาดที่ทำให้เกิดความบิดเบี้ยวได้

การบิดเบี้ยวภายนอก (External Distortion) มีผลมาจากทิศทางการวางตัวรับภาพ หรือเกิดจากรูปทรงของวัตถุเอง โดยอาจเกิดขึ้นจากตำแหน่งของการถ่ายภาพ มุมมองของการถ่ายภาพ และตัวการวางตัวของระนาบของพื้นผิวของวัตถุ



รูปที่ 2.5 ลักษณะของการบิดเบี้ยวภายใน (Internal Distortion)



รูปที่ 2.6 ลักษณะของการบิดเบี้ยวภายนอก (External Distortion)

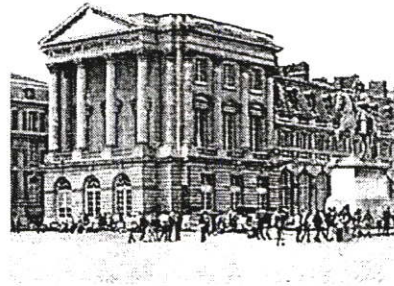
สำหรับรูปที่ 2.5 และ 2.6 เป็นตัวอย่างของภาพที่บิดเบี้ยวไป โดยได้แสดงขนาดภาพจริง (เส้นหนา) เปรียบเทียบกับภาพที่บิดเบี้ยว (เส้นบาง) จะเห็นได้ว่าในความผิดพลาดบางลักษณะ อาจเกิดได้จากการบิดเบี้ยวแบบภายในและแบบภายนอก ทั้งนี้เนื่องจากความบิดเบี้ยวนี้ อาจเกิดได้เนื่องจากคุณสมบัติของรับภาพเอง หรือเกิดจากลักษณะการวางตัว ของตัวรับภาพ ขณะถ่ายภาพก็ได้

2.1.7 การดูแบบเปอร์สเปกทีฟ (Perspective)

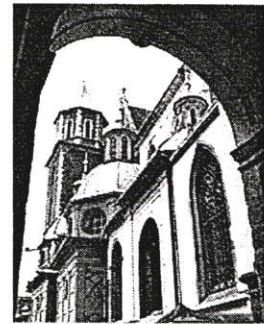
2.1.7.1 หลักการพื้นฐานของการดูแบบเปอร์สเปกทีฟ



(ก)



(ข)



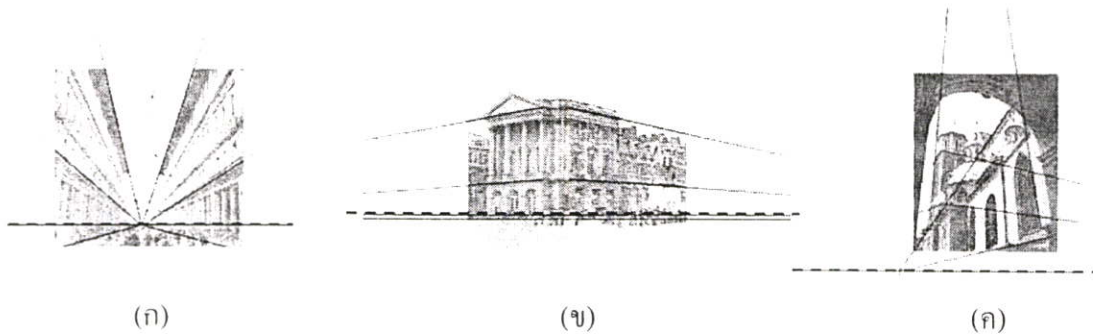
(ค)

รูปที่ 2.7 แสดงรูปอาคารที่มีการดูแบบเปอร์สเปกทีฟ (ก) เปอร์สเปกทีฟ 1 จุด (ข) เปอร์สเปกทีฟ 2 จุด (ค) เปอร์สเปกทีฟ 3 จุด

การดูแบบเปอร์สเปกทีฟ [23] ถูกนำมาใช้ในการวาดภาพและการเขียนแบบต่างๆ เนื่องจากเทคนิคนี้ทำให้เกิดการมองภาพเสมือนจริง วัตถุควมามีมิติไม่แบนราบ ในรูปที่ 2.7 แสดงความแตกต่างระหว่างเปอร์สเปกทีฟ 1 จุด 2 จุด และ 3 จุด ในรูป (ก) เป็นตัวอย่างของเปอร์สเปกทีฟ 1 จุด โดยเส้นแนววัตถุจะลู่เข้าที่จุดๆ หนึ่ง เรียกจุดที่เส้นแนววัตถุลู่เข้าไปจนบรรจบกันว่า จุดวานิชซึ่ง (Vanishing Point) ในรูป (ข) จะเห็นว่าภาพอาคารมีการลู่เข้าทั้งด้านซ้ายและขวา ทำให้เกิดจุดวานิชซึ่ง 2 จุดในแนวสายตา ในรูป (ค) จะเห็นเส้นองค์ประกอบของอาคารลู่เข้า 3 แนว คือ ด้านซ้ายและขวา และแนวคิงซึ่งมีจุดวานิชซึ่งอยู่บนท้องฟ้า

2.1.7.2 การสร้างเปอร์สเปกทีฟจากภาพถ่าย

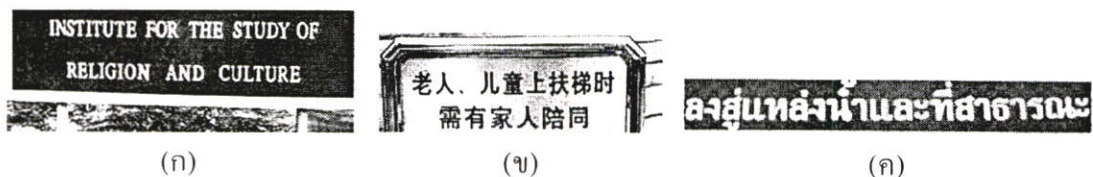
ในรูปที่ 2.8 เป็นการสร้างแนวเส้นสายตา (เส้นประ) และเส้นที่จะบรรจบกันเป็นจุดวานิชซึ่ง (เส้นทึบ) ให้แก่ภาพอาคารเดิม ในรูป (ก) และ (ข) เป็นตัวอย่างของเปอร์สเปกทีฟแบบปกติ ซึ่งผู้ถ่ายภาพไม่ต้องเอียงมุมกล้องขึ้นหรือลง สามารถถ่ายภาพในระดับสายตาปกติ ในทั้งสองรูปนี้ ผู้มองภาพจะเห็นเส้นแนวอาคารตรงดิ่ง สำหรับรูป (ค) เป็นตัวอย่างของเปอร์สเปกทีฟแบบหนอนมอง (Worm's Eye Perspective) ซึ่งผู้มองหรือผู้ถ่ายภาพจะต้องวางกล้องให้มีมุมต่ำกว่าเส้นสายตา เปอร์สเปกทีฟแบบนี้จะลู่เอียงขึ้นไปในทิศบน สามารถสร้างจุดวานิชซึ่งอยู่ข้างบนภาพได้ เรียกว่า จุดซันิธ (Zenith) ซึ่งทุกๆ เส้นแนวตั้งของอาคารจะลู่เข้าหาจุดนี้ ในทางกลับกัน หากสร้างเปอร์สเปกทีฟแบบบนมอง (Bird's Eye Perspective) จะได้จุดวานิชซึ่งที่ต่ำลงไป เรียกว่า จุดเนเดอร์ (Nadir)



รูปที่ 2.8 การสร้างเส้นแนวสายตาและเส้นที่ลากต่อไปจนบรรจบเพื่อหาจุดวานิชซึ่งในภาพอาคาร

2.1.8 การวิเคราะห์ความสัมพันธ์ของเลย์เอาต์ (Layout Analysis)

ความสัมพันธ์ของเลย์เอาต์ เป็นการพิจารณาการวางตัวของกลุ่มตัวอักษรในข้อความ ซึ่งแต่ละภาษาก็จะมีเลย์เอาต์แตกต่างกันไป ดังรูปที่ 2.9 แสดงให้เห็นเลย์เอาต์ของข้อความภาษาอังกฤษ ภาษาจีน และภาษาไทย



รูปที่ 2.9 แสดงภาพป้ายข้อความ (ก) ป้ายภาษาอังกฤษ (ข) ป้ายภาษาจีน (ค) ป้ายภาษาไทย

จะเห็นได้ว่า แต่ละภาษาจะมีเลย์เอาต์ข้อความที่แตกต่างกัน ขึ้นอยู่กับลักษณะของตัวอักษรและวิธีการนำตัวอักษรมาเรียงต่อกันให้เป็นข้อความ ภาษาอังกฤษและภาษาจีนมีเลย์เอาต์ใกล้เคียงกัน กล่าวคือ ข้อความจะสร้างขึ้นจากกลุ่มตัวอักษรที่เรียงต่อกันในบรรทัดเดียวกัน แตกต่างกับภาษาไทยที่มีการเรียงตัวของตัวอักษร วรรณยุกต์ และสระ อยู่ต่างบรรทัดกัน จากงานวิจัยของ Chen และคณะ [1] ซึ่งได้ศึกษาเลย์เอาต์ของภาษาจีนพบว่า ตัวอักษรจีนมีลักษณะเป็นบล็อกสี่เหลี่ยม ทำให้เมื่อประกอบกันเป็นข้อความจะได้เลย์เอาต์ที่เป็นกรอบ ซึ่งค่อนข้างมีขอบเขตชัดเจน แต่ละตัวอักษรในป้ายประกอบไปด้วยส่วนย่อยซึ่งเชื่อมต่อกัน ในบางครั้งส่วนย่อยจะวางตัวในทิศทางเดียวกันกับตัวอักษรในป้าย อย่างไรก็ตาม ปัญหาที่พบในการวิเคราะห์เลย์เอาต์ข้อความแบบอัตโนมัติสำหรับป้ายภาษาจีนคือ ระบบจะทราบได้อย่างไรว่าบริเวณนั้นเป็นหนึ่งตัวอักษร หรือเป็นเพียงส่วนย่อยของตัวอักษร โดยไม่ต้องวิเคราะห์ทั้งข้อความ

Chen และคณะ ได้ใช้เทคนิคการวิเคราะห์เลย์เอาต์ในการแก้ปัญหาดังกล่าว โดยมีเกณฑ์การวิเคราะห์ดังนี้

1. ส่วนประกอบย่อยในตัวอักษรหนึ่งตัว ควรจะอยู่ใกล้ๆ กัน นั่นคือ

$$\text{Dist}(i, j) / \max(\text{Width}(i), \text{Width}(j)) < C1 \quad (2.13)$$

2. เกณฑ์อัตราส่วน (Aspect Ratio-AP) ของแต่ละส่วนอยู่ในขอบเขตที่แน่นอน

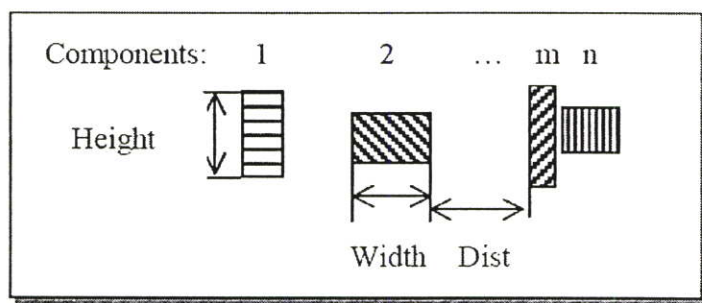
$$C2 < AP_i < C3 \quad (2.14)$$

3. หลังจากกำหนดขอบเขตตัวอักษรได้แล้ว จะพิจารณาตัวอักษรข้างเคียง ซึ่งในข้อความเดียวกันนั้น ตัวอักษรควรจะมีการเรียงตัวอยู่ใกล้กัน

$$\text{Align}(i, j) > C5 \quad (2.15)$$

$$\text{Align}(i, j) = C4 \cdot \frac{\text{Ol}(i, j)}{\max(\text{Height}_i, \text{Height}_j)} \quad (2.16)$$

ในสมการ (2.16) $\text{Ol}(i, j)$ คือความยาวของส่วนที่เหลื่อมกัน (Overlap) ของส่วนย่อย i และ j (สำหรับตัวอักษรจีนที่มีส่วนประกอบย่อยของตัวอักษรเหลื่อมกัน) สำหรับ $C1-C5$ เป็นค่าคงที่ที่ได้กำหนดไว้แล้ว ส่วนสัญลักษณ์อื่นๆ ได้ถูกกำหนดไว้ในรูปที่ 2.10



รูปที่ 2.10 การกำหนดค่าต่างๆ ในการวิเคราะห์เลย์เอาต์ของข้อความในป้าย

2.1.9 ซอฟต์แวร์รู้จำตัวอักษรไทยและอังกฤษ “อ่านไทย” (Thai & English OCR - ArnThai)

ซอฟต์แวร์รู้จำตัวอักษรไทย “อ่านไทย” [22] เป็นซอฟต์แวร์ที่พัฒนาโดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC) ซึ่งการพัฒนาซอฟต์แวร์นี้มีวัตถุประสงค์เพื่อวิจัยและพัฒนาทางด้านเทคโนโลยีภาพ โดยเน้นการวิเคราะห์ภาพและการรู้จำภาพเป็นหลัก

ซอฟต์แวร์ “อ่านไทย” เป็นหนึ่งในงานวิจัยด้านแปลงเอกสารภาษาไทยเก็บเป็นเพิ่มข้อความ ปัจจุบันได้พัฒนาถึงเวอร์ชัน 2.5 โดยมีคุณสมบัติของโปรแกรมดังนี้

- 1) ทำการแปลงเพิ่มภาพลักษณะเอกสารตัวพิมพ์ไทยและอังกฤษให้เป็นเพิ่มข้อความ
- 2) รับอินพุตเป็นเพิ่มภาพลักษณะเอกสาร ได้หลายรูปแบบทั้งภาพสี ภาพขาวดำ และภาพสีเทา
- 3) ส่งเอาต์พุตเป็นเพิ่มข้อความเอกสารแบบ “.txt”
- 4) อัตราความถูกต้องในการรู้จำอยู่ที่ประมาณ 95% ขึ้นไปกับเอกสารภาษาไทย
- 5) อัตราความถูกต้องในการรู้จำอยู่ที่ประมาณ 90% ขึ้นไปกับเอกสารภาษาไทยและเอกสารภาษาอังกฤษปนกัน
- 6) สามารถปรับแต่งภาพได้ตามต้องการ สามารถแปลงภาพจากภาพสีเป็นภาพสีเทา ภาพสีเทาเป็นภาพขาวดำหรือภาพขาวดำกลับเป็นภาพสีเทา รวมทั้งการสลับสีจากขาวเป็นดำ ดำเป็นขาว การย่อหรือขยายภาพ การหมุนภาพ 90 180 และ 270 องศา
- 7) มีการตรวจสอบตัวสะกดโดยจะแสดงไฮไลท์คำผิดให้อัตโนมติ และจะแสดงคำใกล้เคียงให้เลือกเพื่อแทนที่คำที่ผิด หรือจะแก้ไขคำผิดเองก็ได้
- 8) ฟอนต์ภาษาไทยสำหรับระบบปฏิบัติการที่เป็น MS- Windows ที่สามารถใช้ได้ดีคือ AngsanaUPC BrowalliaUPC CordiaUPC DilleniaUPC FreesiaUPC IrisUPC และ JasmineUPC
- 9) ฟอนต์ภาษาไทยสำหรับระบบปฏิบัติการที่เป็น Macintosh ที่สามารถใช้ได้ดีคือ AgfaTom DB75Narai DB95ThaiText DBFongNam DBNBarai DBSurawong DBThaiText EACChuanPim EACEact EACLanna EACPemai และ PSL-Text
- 10) ใช้งานได้ตั้งแต่ตัวอักษรขนาด 10 พอยท์ ขึ้นไป
- 11) เพิ่มภาพลักษณะเอกสารที่นำเข้ามาเป็นอินพุตนี้ควรจะสแกนเข้ามาที่ความละเอียด 200-400 dpi

2.2 งานวิจัยที่เกี่ยวข้อง

2.2.1 งานวิจัยการค้นหาตัวอักษรด้วยคุณลักษณะของตัวอักษร ใช้คุณสมบัติของตัวอักษรในการค้นหาและคัดแยกตัวอักษร เช่น สี ความกว้าง ความสูง อัตราส่วนของความกว้างต่อความสูง อัตราส่วนของพื้นที่ เป็นต้น

2.2.1.1 A Text Detection, Localization and Segmentation System for OCR in Images [7] Gllavata และคณะ เสนอการค้นหาตัวอักษรในฉากที่ซับซ้อนเพื่อใช้ในการทำดัชนีสำหรับการค้นหาภาพในฐานะข้อมูลภาพนิ่งและภาพเคลื่อนไหว โดยทำการค้นหาลักษณะของตัวอักษรโดยใช้การแปลงเวฟเล็ต ทำการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน (Connected Component-CC) เพื่อกำหนดขอบเขตของตัวอักษรแล้วทำการ

สกัดแต่ละตัวอักษรด้วยอัลกอริทึมค้นหาตัวอักษรแบบละเอียด สุดท้ายทำการแยกตัวอักษรและการแปลงเป็นภาพขาวดำด้วยการลดทอนสีด้วยการจัดกลุ่ม(Quantization) และการแปลงเวฟเล็ต

2.2.1.2 A New Approach for Video Text Detection [8] Cai และคณะ ได้เสนอการค้นหาตัวอักษรที่ใช้คุณสมบัติของตัวอักษรเช่น เส้นขอบตรง ความหนาแน่นของเส้นขอบ และการเรียงตัวในแนวระดับ ขั้นตอนแรกใช้อัลกอริทึมการค้นหาเส้นขอบภาพสีในสเปซของสีแบบ YUV และกรองเส้นขอบที่ไม่ใช่ตัวอักษรโดยเทียบกับค่าที่กำหนด (Threshold) จากนั้นใช้เทคนิคการเทียบค่าแบบเฉพาะ เพื่อรักษาตัวอักษรที่มีความแตกต่างน้อยและรวมถึงการทำให้ฉากหลังมีความซับซ้อนน้อยลง และกรองภาพโดยใช้คอนโวลูชัน 2 ขนาด สุดท้ายภาพของตัวอักษรถูกวิเคราะห์แบบเฉพาะที่ด้วยข้อมูลที่มีการจัดเตรียมไว้

2.2.1.3 Automatic Text Location in Images and Video Frames [9] Jain และ Yu เริ่มด้วยการลดปริมาณสีด้วยการทอนสีแบบจัดกลุ่มสี (Quantization) จากนั้นใช้อัลกอริทึมการจัดเรียงลำดับรูปภาพของแต่ละกลุ่มสีแบบหลายรูปแบบ เพื่อแยกสร้างภาพที่ประกอบด้วยพื้นหน้าและพื้นหลังหลาย ๆ รูปแบบ จากนั้นใช้การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน เพื่อพิจารณาภาพพื้นหน้าและพื้นหลังแต่ละภาพ เพื่อระบุบริเวณที่อาจจะเป็นข้อความ วิธีการนี้จะคัดแยกได้เฉพาะข้อความที่วางตัวในแนวนอนและมีขนาดใหญ่ ซึ่งสามารถใช้ค้นข้อความได้ดีกับภาพที่เป็นเอกสารที่มีการวางตัวของข้อความในแนวระดับ

2.2.1.4 Text Detection for Video Analysis [10] Agnihotri และ Dimitrova ได้เสนออัลกอริทึมที่พิจารณากรอบของสีแดง ในแบบจำลองสี RGB ด้วยลักษณะของสีของตัวอักษรที่พบบ่อยๆ โดยใช้คอนโวลูชันในการหาค่ากลางที่แตกต่างกันใน 2 กระบวนการ ซึ่งขั้นแรกทำการเพิ่มคุณภาพของภาพ และขั้นที่สองจะทำการให้เส้นขอบของตัวอักษร แล้วทำการจับกลุ่มของตัวอักษรจากการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน ในขั้นตอนสุดท้ายใช้กระบวนการแบบขาวดำ (binary) แบบปรกติ

2.2.1.5 A Robust Algorithm for Text Detection in Images [11] Gllavata และคณะ [3] เสนอกระบวนการการกำหนดตัวอักษรแบบเฉพาะที่และสกัดอย่างอัตโนมัติในภาพสี โดยขั้นแรกแปลงจากภาพสีเป็นสีเฉดเทาและพิจารณาค่าเฉพาะช่องสัญญาณ Y แล้วหาบริเวณที่คาดว่าเป็นตัวอักษรจากภาพเส้นขอบ สุดท้ายเข้าสู่กระบวนการขาวดำด้วยการเทียบกับค่าที่กำหนด (Threshold) แบบทั่วไป

2.2.1.6 Textfinder: An Automatic System to Detect and Recognize Text in Images [12] Wu และคณะ เสนอการสกัดตัวอักษรแบบอัตโนมัติจากการใช้ข้อมูลของตัวอักษรที่มีเอกลักษณ์เฉพาะตัว โดยพิจารณาจากความกว้าง และความสูงของตัวอักษร

และทำการสกัดด้วยการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน เพิ่มประสิทธิภาพของการค้นหาตัวอักษรด้วยการใช้ฮิสโตแกรมแบบพื้นฐานสำหรับพื้นที่ของตัวอักษรแต่ละตัว

2.2.1.7 Localizing and Segmenting Text in Images and Videos [13] Lienhart

และ Wernicke เสนอวิธีการตรวจจับตัวอักษรจากการจัดตัวแบบแนวระดับ โดยตัวอักษรถูกระบุจากการใช้โครงข่ายประสาทเทียมแบบไปข้างหน้าหลายระดับชั้น ที่ผ่านการฝึกที่จะตรวจจับตัวอักษรที่ที่ขนาดตามที่ถูกกำหนดไว้ ภาพการ์เดิลของข้อมูลภาพสี่ให้คุณลักษณะสำหรับตัวอักษรแบบเฉพาะที่ การค้นหาแบบหลายขนาดเพื่อให้ทนทานต่อการเปลี่ยนแปลงขนาดของตัวอักษรที่พบในภาพ และเส้นของตัวอักษรเฉพาะที่ถูกปรับขนาดเพื่อจำกัดความสูงและทำการแยกส่วนเป็นภาพขาวดำ กระบวนการต่างๆ เริ่มด้วยการระบุสีของตัวอักษรและพื้นหลัง ต่อจากนั้นพื้นหลังถูกทำให้ลดลงจากการวิเคราะห์ส่วนที่ต่อเชื่อมกัน ขั้นสุดท้ายทำกระบวนการขาวดำเพื่อแยกตัวอักษรและพื้นหลังด้วยการเทียบกับค่าที่กำหนด

2.2.2 งานวิจัยการค้นหาตัวอักษรด้วยคุณลักษณะของตัวอักษร ใช้คุณลักษณะของตัวอักษรในการค้นหาและคัดแยกตัวอักษร เช่น การกระจายของข้อมูล คุณลักษณะเวฟเล็ต คุณลักษณะของข้อมูลในแต่ละบล็อกข้อมูล เป็นต้น

2.2.2.1 Automatic Text Detection In Video Frames Based on Bootstrap Artificial Neural Network and CED [14] Hao และคณะ ใช้ค่าคุณลักษณะของตัวกรองเกเบอร์ของเส้นขอบของภาพสี่ เป็นข้อมูลเข้าสำหรับกระบวนการหาบล็อกตัวอักษรด้วยโครงข่ายประสาทเทียม

2.2.2.2 Automatic Text Detection and Tracking in Digital Videos [15] Li และคณะ เสนอการพิจารณาตัวอักษรด้วยกรอบสี่เหลี่ยมขนาด 16x16 พิกเซล ว่าเป็นตัวอักษรหรือไม่โดยผ่านโครงข่ายประสาทเทียมแบบ 3 ลำดับชั้นบนพื้นฐานจากคุณสมบัติแบบเฉพาะที่จากของข้อมูลของค่าสัมประสิทธิ์เวฟเล็ตแบบความถี่สูงที่สุดทำการฉายโปรไฟล์เพื่อใช้วิเคราะห์การสกัดส่วนประกอบของตัวอักษรจากกล่องตัวอักษร

2.2.3 งานวิจัยการรู้จำตัวอักษร เนื่องจากระบบรู้จำตัวอักษรเป็นขั้นตอนที่นำข้อมูลที่ได้รับจากการค้นหาและคัดแยกตัวอักษรมาทำการหาคุณลักษณะเพื่อให้สามารถรู้จำว่าเป็นตัวอักษรใด ดังนั้นการศึกษาระบบดังกล่าวจึงมีความจำเป็นต่อการวิจัย เนื่องจากต้องเตรียมหรือวิเคราะห์ข้อมูลภาพให้เหมาะสมต่อการนำไปใช้ ซึ่งงานวิจัยที่เกี่ยวข้องเป็นงานวิจัยการรู้จำตัวอักษรภาษาจีน และภาษาไทย

2.2.3.1 Automatic Detection and Recognition of Signs from Natural Scenes

[1] Chen และคณะ ได้เสนอวิธีการในการตรวจจับป้ายภาษาจีนจากฉากธรรมชาติ โดยใช้วิธีการพิจารณาความละเอียดหลายระดับ (Multiresolution) และการแก้ไขการถ่วงเข้าของข้อความในภาพ (Affine Rectification Transformation) ซึ่งเกิดขึ้นจากการวางมุมมองกล้องที่ไม่เหมาะสม รวมถึงสร้างระบบการรู้จำตัวอักษรจากป้ายภาษาจีนในภาพ และแปลข้อความเป็นภาษาอังกฤษ

2.2.3.2 การรู้จำอักษรคัดลายมือภาษาไทยของคอมพิวเตอร์โดยวิธีวิเคราะห์

โครงสร้างแบบทันทีทันใด ผลงานวิจัยของนายชาญชัย คืออ่วม [16] เป็นวิธีการพิจารณาแบบ 2 ขั้นตอน คือ ขั้นตอนในการแยกแยะเส้น จะทำการวิเคราะห์ว่าเส้นที่ประกอบเป็นตัวอักษรเส้นใน โดยกรรมวิธีเปลี่ยนเป็นรหัสลูกโซ่แปดทิศ เพื่อให้ได้เส้นต่าง ๆ ได้แก่ หัวตัวอักษร วงกลมภายในตัวอักษรเส้นตรงที่เคลื่อนที่ขงทิศทั้ง 8 และเส้นหยักอยู่ในตำแหน่งใด ซึ่งตัวอักษรจะถูกแบ่งเป็น 6 ส่วน เท่า ๆ กันโดยอัตโนมัติ เพื่อให้ทราบตำแหน่งของเส้นและนำเส้นต่าง ๆ เข้าสู่ขั้นตอนที่ 2 เพื่อทำการวิเคราะห์ว่าเป็นตัวอักษรใด โดยพิจารณาหมายเลขช่องว่ามีเส้นที่กำหนดอยู่หรือไม่ มีการเริ่มต้นหรือการสิ้นสุดการเขียนที่ตำแหน่งช่องใดๆ ความยาวของเส้นบางเส้น และจำนวนเส้นในแนวตั้ง เป็นต้น

2.2.3.3 การจดจำรูปแบบตัวอักษรคัดลายมือภาษาไทยโดยวิธีแยกลักษณะเด่น

ผลงานวิจัยของนายประสาร ดังดีสานนท์ [17] เป็นวิธีการที่พิจารณาถึงโครงสร้างของรูปแบบตัวอักษร ทั้งลายเส้นและพื้นเบื้องหลังของลายเส้นควบคู่กันไปวิธีการพิจารณาลายเส้นและพื้นเบื้องหลังลายเส้น ผู้วิจัยได้แบ่งออกเป็นสามขั้นตอนด้วยกันคือ ขั้นตอนแรก เป็นการเปลี่ยนลักษณะลายเส้นและลักษณะของพื้นเบื้องหลังลายเส้น ไปเป็นรหัสเบื้องต้น (Initial Feature Extraction) ขั้นตอนที่สองการลดทอนรหัสเบื้องต้นของพื้นหลังลายเส้นที่ซ้ำซ้อนกันออกไป (Unification) ขั้นตอนที่สามเป็นการรวมรหัสลายเส้น และรหัสของพื้นหลังลายเส้นเข้าด้วยกันเพื่อให้เป็นคุณสมบัติของรูปแบบตัวอักษร (Concentration) เมื่อได้รหัสรวมของอักษรแต่ละตัวจะถูกจัดแยกกลุ่ม แล้วพิจารณาหาคุณสมบัติที่เป็นลักษณะเด่นเฉพาะของตัวอักษรนั้นๆ ที่ไม่เหมือนกันกับอักษรตัวอื่น ทำการแยกตัวอักษรออกจากกลุ่มออกไป รหัสที่เป็นคุณสมบัติเด่นของอักษรแต่ละตัวจะถูกพิจารณาจากบริเวณพื้นที่รอบๆ ตัวอักษร โดยผู้วิจัยจะพิจารณาค่าแล้วจึงกำหนดเป็นเงื่อนไข วิธีการนี้จึงใช้ได้ผลดีกับตัวอักษรต้นแบบ แต่หากลักษณะของตัวอักษรที่ต้องการวิเคราะห์เปลี่ยนรูปร่างแตกต่างจากต้นแบบแล้ว จะไม่สามารถวิเคราะห์ได้ อีกทั้งยังขาดต่อการเพิ่มเติมเงื่อนไขใหม่ๆ เพื่อแยกตัวอักษรออกจากกลุ่ม หากอักษรที่วิเคราะห์ไม่สามารถวินิจฉัยได้ แต่หากอักษรมีลักษณะเหมือนอักษรต้นแบบ ผลการรู้จำให้ความ

ถูกต้องสูงเกินกว่า 90 เปอร์เซ็นต์ แต่ว่าเวลาที่ใช้ในการรู้จำตัวอักษรแต่ละตัวต้องใช้เวลาประมาณ 3 นาทีซึ่งเป็นเวลานานเกินไป จึงไม่สามารถนำออกมาใช้งานได้จริง

2.2.3.4 การจดจำตัวอักษรลายมือเขียนภาษาไทยโดยการพิจารณาหัวตัวอักษร

ผลงานวิจัยของนายสุรพันธ์ เอื้อไพบูลย์ [18] ทำการเลือกพิจารณาส่วนหัวของตัวอักษรภาษาไทยที่มีลักษณะเป็นวงกลมเล็กๆ ซึ่งแตกต่างไปจากอักษรภาษาอื่นๆ ซึ่งมักไม่มีหัว ซึ่งผู้วิจัยได้ทำการตรวจสอบหัวของตัวอักษร แล้วจัดแบ่งกลุ่มตัวอักษรออกเป็นกลุ่มๆ คือ กลุ่มตัวอักษรที่ไม่มีหัว กลุ่มตัวอักษรที่มี 1 หัว กลุ่มอักษรที่มี 2 หัว และกลุ่มตัวอักษรที่มี 3 หัว ขั้นตอนการทำงานแบ่งออกเป็น 4 ขั้นตอน คือ ขั้นตอนที่หนึ่งเป็นการเตรียมภาพตัวอักษรแต่ละตัวไว้ 2 ชุด โดยชุดหนึ่งเป็นภาพปกติ และอีกชุดหนึ่งจะเป็นภาพตัวอักษรที่ผ่านการทำให้บางแล้ว ขั้นตอนที่สอง จะทำการศึกษาจากส่วนหัวตัวอักษรปกติ และคำนวณหาคุณสมบัติทางโทโพโลยีจากตัวอักษรที่ถูกทำให้บาง ในขั้นตอนสุดท้ายเป็นการวิเคราะห์หัวตัวอักษรด้วยเทคนิค ได้แก่ 1. Subhead Region 2. Feature Code 3. Head Style และ 4. Width Per Height Ratio การทำงานของระบบใช้กล้องเก็บภาพ CCD เป็นอุปกรณ์ป้อนข้อมูล (Input Device) ซึ่งการใช้งานคงยุ่งยากพอสมควร และต้องพึ่งผู้มีความชำนาญเป็นผู้ปรับแต่งความคมชัดในการถ่ายภาพ และในส่วนของอัลกอริทึมที่ใช้ เป็นการพิจารณาเฉพาะส่วนหัวของตัวอักษรซึ่งค่อนข้างจะเป็นการพิจารณาเฉพาะเจาะจงเกินไป หากข้อมูลเป็นภาพตัวอักษรที่เขียนแบบไม่มีหัว หรือลายเส้นของหัวขาดไม่สมบูรณ์หรือไม่เป็นวงปิด หรือไม่เป็นวงปิด หรือไม่ว่าหัวมีลักษณะที่บิดเบี้ยว ผลของการรู้จำย่อมผิดพลาดอย่างแน่นอน อย่างไรก็ตาม หากทุกอย่างเป็นไปได้ตามเงื่อนไขที่ผู้ทำวิจัยกำหนดไว้ในงานวิจัยนี้สามารถรู้จำตัวอักษรลายมือเขียนให้ผลของการรู้จำที่ตีพอสมควร

2.2.3.5 การรู้จำอักษรพิมพ์ภาษาไทยโดยวิธีการค้นหาลักษณะโครงสร้างของลายเส้น

ผลงานวิจัยของนายสุรสิทธิ์ ราตรี [19] ได้นำเสนอวิธีการค้นหาลักษณะโครงสร้างลายเส้นซึ่งเป็นลักษณะเด่นของตัวอักษร (Topological Feature Extraction) การพิจารณากระทำกับโครงสร้างของลายเส้นตัวอักษรที่ถูกแบ่งออกเป็น 8 ส่วนในแนวตั้ง แนวนอนและในแนวเส้นทแยงมุมทั้งสอง ในขั้นตอนต่อไปทำการแยกพิจารณาในแต่ละส่วน (Octant) หากคุณสมบัติทางโทโพโลยีของลายเส้น (Stroke) ของตัวอักษร ซึ่งจะได้รหัสแทนตัวอักษรที่เรียกว่า A-Code จากการรวมรหัสแสดงคุณลักษณะเด่นในแต่ละส่วนของตัวอักษรที่ถูกแบ่งทั้ง 8 ส่วน เนื่องจากการพิจารณากระทำกับลักษณะโครงสร้างลายเส้นของตัวอักษรอย่างละเอียดจึงสามารถเก็บคุณลักษณะเด่นของตัวอักษรได้ครบถ้วน ส่งผลให้ความถูกต้องของการรู้จำมีความถูกต้องสูงเกินกว่า 95 เปอร์เซ็นต์ แต่การแบ่งพื้นที่การพิจารณาออกเป็นส่วนๆ ที่มีพื้นที่เล็กเกินไป จะทำให้การพิจารณาโครงสร้างของลายเส้นในแต่ละส่วนอาจไม่ได้ลักษณะลายเส้นที่แท้จริงในกรณีที่ลายเส้น

ตัวอักษรที่มีความหนา แต่แนวทางแก้ไขสามารถทำได้โดยเพิ่มเติมส่วนของการทำตัวอักษรให้บาง ซึ่งจะได้โครงร่างของลายเส้นตัวอักษรที่แท้จริง แล้วจึงนำไปพิจารณาจะสามารถแก้ปัญหาส่วนนี้ได้ และคาดว่าจะได้ความถูกต้องแน่นอนแม้ว่ารูปแบบของตัวอักษร (Font) อาจจะแตกต่างกัน แต่ทั้งนี้ยังไม่สามารถรู้จำตัวอักษรแบบเอียง (Italic) เพราะตัวอักษรเอียงเมื่อแบ่งออกเป็น 8 ส่วน โดยใช้หลักการแบ่งเหมือนปกติ จะได้ผลที่แตกต่างกัน เมื่อนำไปพิจารณาเพื่อรู้จำอักษร โดยใช้หลักการเดียวกับตัวอักษรปกติ ก็จะทำให้เกิดความผิดพลาด

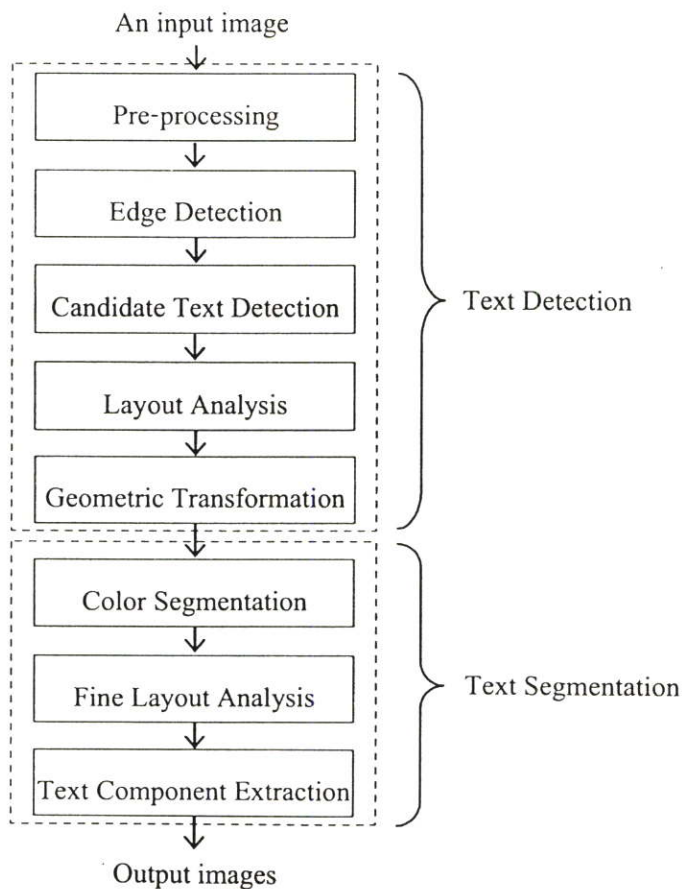
2.2.3.6 การรู้จำอักษรไทยโดยใช้ลักษณะเด่นของตัวอักษร ผลงานวิจัยของนาย กฤษณา วิไลลักษณ์ และคณะ[20] ซึ่งเสนอวิธีการรู้จำอักษรภาษาไทยโดยอาศัยลักษณะเด่นของตัวอักษร แนวคิดนี้ไม่ขึ้นอยู่กับขนาด ความหนา หรือการเอียงของตัวอักษร จึงช่วยลดเวลาในการจัดการเบื้องต้นในการใช้ลักษณะเด่นหลายๆ แบบแบ่งกลุ่มของตัวอักษร พบว่าตัวอักษรบางตัวอาจอยู่ได้หลายกลุ่ม เมื่อนำกลุ่มเหล่านี้มาหาตัวร่วมพบว่าจะได้กลุ่มที่เล็กลง ซึ่งมีสมาชิกไม่เกิน 7 ตัวอักษร จากนั้นเมื่อนำลักษณะพิเศษเฉพาะตัวอักษร มาแยกแยะอีกครั้ง ก็สามารถรู้จำอักษรได้อย่างมีประสิทธิภาพ ผลการทดสอบกับตัวพิมพ์ 5 รูปแบบ และลายมือเขียนภาษาไทยของ 10 คน เขียนข้อความทดสอบอักษรไทยมาตรฐานของสมาคมคอมพิวเตอร์แห่งประเทศไทย โดยมีพยัญชนะและสระทั้งหมด 3,870 ตัว พบว่าวิธีการดังกล่าวสามารถรู้จำได้ถูกต้องประมาณ 91 เปอร์เซ็นต์

บทที่ 3

การตรวจจับและตัดแยกตัวอักษรไทยจากภาพในฉากธรรมชาติ

3.1 กระบวนการตรวจจับและตัดแยกตัวอักษรไทยจากภาพป้ายในฉากธรรมชาติ

วิทยานิพนธ์ฉบับนี้ เสนอวิธีการตรวจจับและตัดแยกตัวอักษรไทยจากภาพในฉากธรรมชาติ วิธีการที่เสนอประกอบด้วย การเตรียมภาพก่อนประมวลผล การทำเป็นภาพระดับสีเทา การทำให้เบลอแบบเกาส์เซียน (Gaussian Blur) การทำภาพให้คมชัดขึ้น (Sharpening) การตรวจจับเส้นขอบโดยใช้ Laplacian of Gaussian (LoG) การเปรียบเทียบอัตราส่วนของแต่ละส่วนประกอบที่เชื่อมต่อกันกับอัตราส่วนที่กำหนดไว้สำหรับตัวอักษรไทย การวิเคราะห์แนวข้อความ การแก้ไขการเอียงและการถ่วง ซึ่งเกิดจากมุมมองที่ไม่เหมาะสมด้วยการแก้ไขแบบสัมพรรค (Affine Rectification) ซึ่งจะช่วยให้เพิ่มความถูกต้องในการตรวจจับตัวอักษร และใช้ Gaussian Mixture Model (GMM) ในการวิเคราะห์สีในแบบจำลองต่างๆ เพื่อหาแบบจำลองที่เหมาะสมในการตัดแยกตัวอักษรออกจากพื้นหลัง โดยมีโครงร่างกระบวนการที่นำเสนอ ดังแสดงในรูปที่ 3.1



รูปที่ 3.1 แสดงกระบวนการตรวจจับและตัดแยกตัวอักษรภาษาไทยจากภาพในฉากธรรมชาติ

3.2 การตรวจจับข้อความ (Text Detection)

ในกระบวนการตรวจจับตัวอักษร เริ่มจากการนำภาพอินพุต ผ่านเข้าสู่ขั้นตอนต่างๆ จำนวน 5 ขั้นตอน ดังนี้

- 1) **Pre-Processing** ทำการเตรียมภาพขั้นต้น เนื่องจากภาพชนิด JPEG มักจะมีบล็อกสีที่เหลี่ยม (Artifact) ปรากฏอยู่ทั่วไปในภาพ อันเป็นผลจากขั้นตอนของ Discrete Cosine Transform (DCT) ในกระบวนการบีบอัดภาพ หากไม่กำจัดบล็อกสีที่เหลี่ยมเหล่านี้ เมื่อทำการหาเส้นขอบวัตถุอาจทำให้ได้เส้นที่เหลี่ยมที่ไม่ต้องการจำนวนมาก ในการกำจัดบล็อกสีที่เหลี่ยมเหล่านี้จะใช้วิธีการเกาส์เซียนเบลอ (Gaussian Blur) โดยมีสมการดังนี้

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3.1)$$

เมื่อ σ คือค่าความเบี่ยงเบนมาตรฐาน (Standard Deviation) ค่า x, y คือค่าในแนวแกน x และ y ตามลำดับ

จากนั้นทำการปรับปรุงความคมชัดของเส้นขอบวัตถุในภาพ (Sharpening) ด้วยการทำคอนโวลูชันกับภาพด้วยตัวกรองที่กล่าวไปในหัวข้อ 2.1.2 จากนั้นทำแปลงภาพเป็นภาพระดับสีเทา (Gray-Scale Image) แล้วเข้าสู่การตรวจจับเส้นขอบ (Edge Detection)

- 2) **Edge Detection** เป็นขั้นตอนของการหาเส้นขอบ ซึ่งวิธีการที่เลือกใช้คือ LoG (Laplacian of Gaussian) เนื่องจากเป็นอัลกอริทึมสำหรับหาขอบของวัตถุในภาพ ที่มีความทนทานต่อสัญญาณรบกวน และให้เส้นขอบที่ชัดเจน สมการของ LoG แสดงได้ในสมการ (3.2)

$$\nabla^2 h(r) = -\left[\frac{r^2 - \sigma^2}{\sigma^4} \right] e^{-\frac{r^2}{2\sigma^2}} \quad (3.2)$$

เมื่อฟังก์ชันเกาส์เซียน $h(r) = -e^{-\frac{r^2}{2\sigma^2}}$ สำหรับ r^2 มีค่าเท่ากับ $x^2 + y^2$ เมื่อ x, y คือค่าพิกัดของพิกเซล และ σ คือ ค่า Standard Deviation

- 3) **Candidate Text Detection** เมื่อได้เส้นขอบแล้ว จะทำการค้นหาส่วนประกอบที่มีการเชื่อมติดกัน (Connected Component Labeling) เพื่อนำมาวิเคราะห์ต่อว่า บริเวณใดบ้างในภาพที่มีโอกาสเป็นกลุ่มข้อความ และบริเวณใดที่ไม่ใช่ข้อความ ซึ่งข้อความในภาษาไทยประกอบไปด้วยพยัญชนะ สระ และวรรณยุกต์ ซึ่งสามารถแยกออกเป็น

4 กลุ่ม ดังตารางที่ 3.1 เพื่อให้สามารถแยกบริเวณที่เป็นข้อความออกจากบริเวณที่ไม่ใช่ข้อความได้นั้น จะต้องนำแต่ละส่วนประกอบที่เชื่อมติดกัน มาหาอัตราส่วน ซึ่งกำหนดดังสมการ (3.3)

$$\text{Character_ratio} = \frac{\text{Character_Width}}{\text{Character_Height}} \quad (3.3)$$

อัตราส่วนที่กำลังจะพิจารณานี้ เป็นค่าโดยประมาณสำหรับตัวอักษรไทย ซึ่งมีค่าโดยเฉลี่ยอยู่ในช่วงความกว้างต่อความสูงค่าหนึ่ง ซึ่งจากการคำนวณพบว่ามีค่าอยู่ในช่วง 0.2-1.2 (อัตราส่วนต่างๆ ที่ใช้ในการพิจารณารูปแบบของฟอนต์ภาษาไทย อยู่ในภาคผนวก ข)

นอกจากนี้ จะใช้อัตราส่วนอีกค่าหนึ่งในการพิจารณาบริเวณที่คาดว่าจะ เป็นข้อความ นั่นคือ อัตราส่วนพื้นที่ตัวอักษรต่อพื้นที่กรอบล้อมรอบ ดังสมการ (3.4)

$$\text{Area_ratio} = \frac{\text{Character_Area}}{\text{Character_Boundary_Area}} \quad (3.4)$$

อัตราส่วนของพื้นที่ตัวอักษรต่อพื้นที่กรอบล้อมรอบ ที่น่าจะเป็นตัวอักษรคือมากกว่า 0.32 หากพิจารณาอัตราส่วนที่กำหนดในสมการ (3.3) และ (3.4) แล้ว ช่วงค่าดังกล่าวสามารถคัดกรองส่วนประกอบอื่นที่ไม่ใช่พยัญชนะรวมถึงส่วนประกอบที่ไม่ใช่ตัวอักษรได้ดี และส่วนประกอบใดที่ผ่านการคัดกรองตามที่กำหนดนั้น มีความเป็นไปได้สูงที่จะเป็นตัวอักษรไทยในกลุ่มของพยัญชนะ

ตารางที่ 3.1 แสดงการจัดกลุ่มตัวอักษร สระ และวรรณยุกต์ในภาษาไทย

Type	Members
Consonants	กขฌคจฉชฌญฎฐฑฒณดต ถทธนบปฝฝฝทฟกมยรฤลวศษส หฬอฮ
Vowels	อ ๑ ๒ ๓ ๔ ๕ ๖ ๗ ๘ ๙ ๑๐ ๑๑ ๑๒
Tones	— — — — — — — — — —
Speacial Symbols	๕๕ ยุก

การคัดกรองข้างต้นทำให้เราสามารถตรวจจับพยัญชนะส่วนใหญ่ในภาษาไทยได้ เพื่อใช้เป็นข้อมูลที่เหมาะสมสำหรับการหาแนวการวางตัวของข้อความในขั้นตอนต่อไป แต่หากพิจารณาตารางที่ 3.1 พบว่าบางพยัญชนะ สระ และวรรณยุกต์ ที่ไม่สามารถตรวจจับได้ในขั้นตอนนี้อาจมีจำนวนหนึ่ง ที่มีลักษณะก้ำกึ่งกับบริเวณที่ไม่ใช่ตัวอักษร ตัวอย่างเช่น สระ อี อี อี อี ไม้หันอากาศ อุ อุ และ ไม้เอก ซึ่งเป็นลักษณะตัวอักษรที่มีความซับซ้อนพิเศษ ที่ทำให้ระบบไม่สามารถตรวจจับข้อความได้ครบ จึงต้องพิจารณา ส่วนประกอบต่างๆ ของภาพอีกครั้งอย่างละเอียดเพื่อให้ได้ส่วนประกอบทั้งหมดของข้อความ ในขั้นตอนของการตัดแยกข้อความ และ Fine Layout Analysis

- 4) **Layout Analysis** เมื่อได้ตัวอักษรแล้ว ขั้นตอนต่อไปจะทำการวิเคราะห์รูปแบบการวางตัวของข้อความ เพื่อกำหนดขอบเขตของข้อความให้ชัดเจน ตัวอย่างของการพิจารณาข้อความ เป็นไปดังรูปที่ 3.2

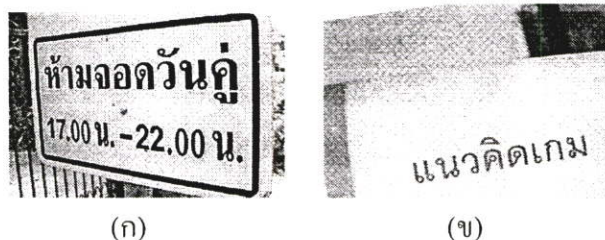


รูปที่ 3.2 แสดงเส้นระดับของข้อความในภาษาไทย

สมมติฐานเบื้องต้นของการหาเลย์เอาต์ข้อความ คือ ข้อความจะวางตัวอยู่ระหว่างเส้นที่ 1 และ 2 ทำให้สามารถพิจารณาแนวข้อความได้ดังนี้

- 1) หากความชันของเส้นที่ 1 และ 2 มีค่าใกล้เคียงกัน จะได้ว่าแนวข้อความวางอยู่ในแนวระดับปกติ หรือมีการหมุนเอียง
- 2) หากความชันของเส้นที่ 1 และ 2 มีค่าไม่เท่ากัน จะต้องลากเส้นตรงต่อออกไปจนตัดกันที่จุดจุดหนึ่ง ซึ่งเรียกว่า จุดวานิชซิง (Vanishing Point) จะได้ว่าข้อความมีการลู่เข้าแบบเปอร์สเปคทีฟ

ในการถ่ายภาพป้ายในฉากธรรมชาติ พบว่า ข้อความส่วนใหญ่ที่ปรากฏในภาพ มักมีการบิดเบี้ยวเนื่องจากมุมของการถ่ายภาพ ซึ่งทำให้เกิดลักษณะการลู่เข้า (Perspective Distortion) และการหมุน (Rotation) ดังรูปที่ 3.3 (ก) และ (ข) ตามลำดับ

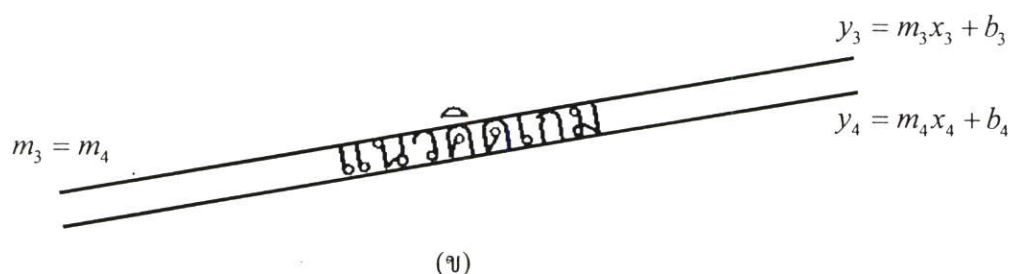
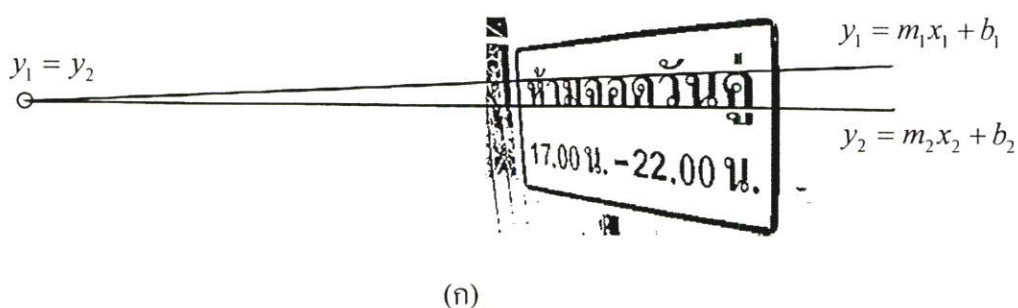


รูปที่ 3.3 แสดงภาพป้ายข้อความ (ก) ข้อความมีลักษณะการลู่อเข้าแบบเปอร์สเปคทีฟ (ข) ข้อความมีการหมุนไปจากแนวปกติ

ผู้วิจัยได้วิเคราะห์ปัญหาดังกล่าวข้างต้น พบว่า แม้ข้อความจะมีการลู่อเข้าหรือ หมุนไปจากแนวปกติเนื่องจากมุมมอง แต่แนวการวางตัวของตัวอักษรในข้อความยังคง ลักษณะของแนวเส้นตรง ดังนั้น เราสามารถพิจารณาคูสมบัติของข้อความในเชิง สมการเส้นตรง ดังนี้

$$y = mx + c \quad (3.5)$$

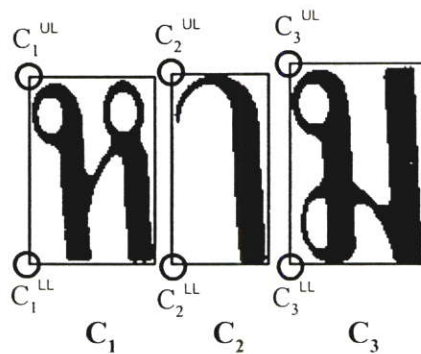
เมื่อ x , y เป็นค่าพิกัดในแนวแกน x และแกน y ตามลำดับ m เป็นค่าความชันของ เส้นตรงและ c เป็นค่าคงที่



รูปที่ 3.4 การหาแนวข้อความโดยพิจารณาจากสมการเส้นตรง (ก) กรณีที่ข้อความมีการลู่อเข้า แบบเปอร์สเปคทีฟ (ข) กรณีที่ข้อความมีการหมุน

แนวข้อความสามารถพิจารณาจากสมการเส้นตรง ดังแสดงในรูปที่ 3.4 โดยอยู่บนสมมติฐานว่า ถ้าข้อความมีลักษณะคู่เข้าแบบเปอร์สเปคทีฟ เส้นตรงที่สร้างขึ้นมาทั้งสองเส้นจะสามารถลากไปตัดกันได้ที่จุดๆ หนึ่ง ดังรูป 3.4 (ก) หากเส้นตรงทั้งสองไม่ตัดกัน แสดงว่าข้อความมีลักษณะปกติเพียงแต่มีการหมุน ดังรูป 3.4 (ข)

ดังนั้น ถ้าต้องการพิจารณาว่าตัวอักษร จะมีความสัมพันธ์กันในเชิงข้อความหรือไม่สามารถทำได้โดยหาความสัมพันธ์ระหว่างคู่ตัวอักษรใดๆ ว่ามีค่า m และ b เป็นเท่าใด ตัวอย่างของการพิจารณาคู่ของตัวอักษร (ต่อไปนี้จะเรียกว่า ส่วนประกอบย่อย) เป็นไปดังรูปที่ 3.5



รูปที่ 3.5 ตัวอย่างของการพิจารณาคู่ของกรอบตัวอักษร เพื่อหาสมการเส้นตรง

จากรูปที่ 3.5 เป็นตัวอย่างของการพิจารณาข้อความ “ห้ามจอดวนคู่” เพื่อพยายามแทนแนวข้อความด้วยเส้นตรงสองเส้น เราจะเริ่มพิจารณาที่ส่วนประกอบย่อยในข้อความ นั่นคือ ตัวพยัญชนะทุกตัว และสระที่มีความสูงใกล้เคียงกับพยัญชนะ โดยเราจะกำหนดกรอบสี่เหลี่ยมที่แคบที่สุด ที่ล้อมรอบพยัญชนะนั้นได้ จากรูปที่ 3.5 จะได้กรอบ C_1 , C_2 และ C_3 แทนกรอบที่ล้อมรอบ “ห”, “า” และ “ม” ตามลำดับ กำหนดให้จุดบนซ้าย (C_i^{UL}) และล่างซ้าย (C_i^{LL}) ของแต่ละกรอบ แทนตำแหน่งบนสุดและล่างสุดของพยัญชนะนั้น จากนั้นทำการจับคู่ C_1 กับ C_2 , C_1 กับ C_3 และ C_2 กับ C_3 เพื่อนำค่าตำแหน่งพิกเซล x และ y ณ จุดนั้นๆ มาแทนในสมการเส้นตรง เมื่อหาสมการเส้นตรงได้ ก็จะทราบค่าความชันและค่าคงที่ของเส้นตรงบนและล่าง ของตัวอักษรแต่ละคู่ได้

จากวิธีวิเคราะห์ข้างต้น จะได้ว่า การจับคู่พิจารณาระหว่าง C_i กับ C_j ใดๆ จะสามารถคำนวณค่าได้ 4 ค่า ดังนี้

ค่าที่ 1 คือ $m_{UL_{ij}}$ เป็นค่าความชันของเส้นตรงที่ผ่านจุดบนซ้ายของ C_i และ C_j

ค่าที่ 2 คือ $b_{UL_{ij}}$ เป็นค่าคงที่ของเส้นตรงที่ผ่านจุดบนซ้ายของ C_i และ C_j

ค่าที่ 3 คือ m_{LL_j} เป็นค่าความชันของเส้นตรงที่ผ่านจุดล่างซ้ายของ C_i และ C_j

ค่าที่ 4 คือ b_{LL_j} เป็นค่าคงที่ของเส้นตรงที่ผ่านจุดล่างซ้ายของ C_i และ C_j

จากหลักการข้างต้น เราสามารถสร้างหลักเกณฑ์ในการวิเคราะห์แนวข้อความได้ดังนี้

- 1) ส่วนประกอบย่อยในข้อความหนึ่งๆ ควรจะวางตัวในแนวเส้นตรงเดียวกัน นั่นคือ

$$|m_{UL_{ij}} - m_{UL_{si}}| < \tau_1 \quad (3.6)$$

$$|m_{LL_{ij}} - m_{LL_{si}}| < \tau_1 \quad (3.7)$$

$$|b_{UL_{ij}} - b_{UL_{si}}| < \tau_2 \quad (3.8)$$

$$|b_{LL_{ij}} - b_{LL_{si}}| < \tau_2 \quad (3.9)$$

- 2) ค่าความชันอยู่ในขอบเขตที่แน่นอน นั่นคือ

$$\tau_3 < m_{UL_{ij}} < \tau_4 \quad (3.10)$$

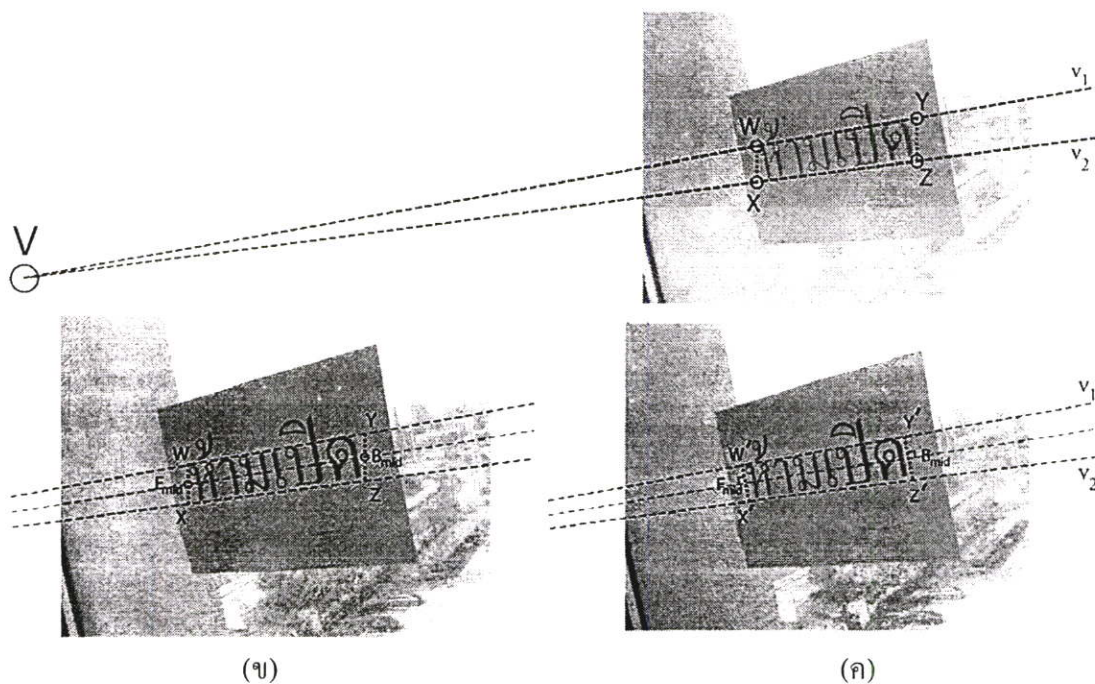
$$\tau_3 < m_{LL_{ij}} < \tau_4 \quad (3.11)$$

โดยค่า $\tau_1 - \tau_4$ เป็นค่าคงที่ที่ได้กำหนดไว้แล้ว

เพื่อให้การหาแนวข้อความทำได้ถูกต้องยิ่งขึ้น จำเป็นต้องพิจารณาให้ครอบคลุมถึงทิศการวางตัวของข้อความ ซึ่งนอกเหนือจากที่มีการลู่อแบบเปอร์สเปคทีฟและการหมุนข้างต้นแล้ว จากการวิจัยพบว่า ข้อความในภาพป้ายส่วนใหญ่จะมีการหมุนที่แฝงการลู่อด้วยซึ่งวิธีแก้ไข ทำได้โดย

- 1) กำหนดเส้นขอบเขตด้านหน้า (\overline{WX}) และหลังสุด (\overline{YZ}) ของกลุ่มข้อความ ดังแสดงในรูปที่ 3.6 (ก)
- 2) พิจารณาว่ามีการลู่อหรือไม่ โดยหาจุดความชันซึ่ง (V)

- 3) หาจุดกึ่งกลางของเส้นรอบด้านหน้า (F_{mid}) และด้านหลัง (B_{mid}) ดังรูป 3.6 (ข)
- 4) ลากเส้นตรงจากจุดวานิชซึ่ง ไปยังจุดกึ่งกลางของเส้นขอบที่ห่างจากจุดวานิชซึ่งมากกว่า เรียกเส้นตรงนี้ว่า $\overline{VB_{mid}}$
- 5) หาความชันของ $\overline{VB_{mid}}$ เพื่อสร้างสมการเส้นตรงที่ตั้งฉากกับ $\overline{VB_{mid}}$ 2 เส้น เส้นแรกผ่านจุด F_{mid} และเส้นที่สองผ่านจุด B_{mid}
- 6) สร้างเส้นรังสี \vec{v}_1 (ผ่านจุด W และ Y) และ \vec{v}_2 (ผ่านจุด X และ Z) แล้วหาจุดตัดกับสมการเส้นตรงในข้อ 5) เพื่อกำหนดจุด W', X', Y' และ Z' ซึ่งเป็นจุดมุมทั้งสี่มุมของกรอบตัวอักษร ดังรูป 3.6 (ค)

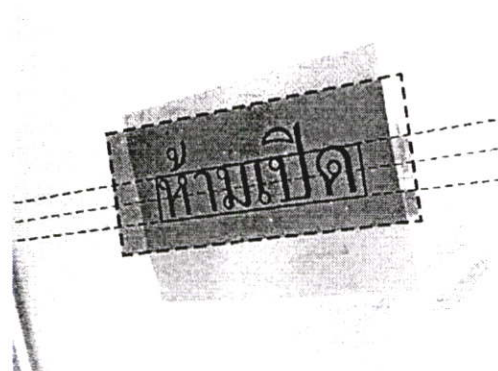


รูปที่ 3.6 การหากรอบข้อความในกรณีที่ข้อความมีการหมุนและการลู่อแบบเปอร์สเปกทีฟ
 (ก) การกำหนดจุดต่างๆ รอบกลุ่มตัวอักษร (ข) การกำหนดจุด F_{mid} และ B_{mid} (ค) กรอบข้อความเมื่อแก้ไขแนวระยะแล้ว

ผลจากการคำนวณเพื่อหาแนวเส้นของข้อความตามขั้นตอนข้างต้น ทำให้สามารถสร้างกรอบล้อมรอบข้อความในป้ายได้ ดังรูปที่ 3.7



(ก)



(ข)

รูปที่ 3.7 แสดงการสร้างกรอบล้อมรอบข้อความ (ก) ภาพต้นฉบับ (ข) ตรวจสอบกลุ่มตัวอักษรและทำการล้อมกรอบข้อความ

เมื่อได้กรอบล้อมรอบกลุ่มตัวอักษรที่อยู่ระหว่างเส้นที่ 1 และ 2 แล้ว ทำการขยายขอบเขตของกลุ่มตัวอักษรนั้นเพื่อให้ครอบคลุมสระ และวรรณยุกต์ ที่อยู่สูงขึ้นไปจากเส้นที่ 1 และต่ำลงมาจากเส้นที่ 2 นอกจากนี้ยังขยายขอบเขตด้านข้างเพื่อให้ครอบคลุมพยัญชนะและสระทั้งด้านหน้าและด้านหลัง โดยขยายตามสัดส่วนดังนี้

- 1) พิจารณาความสูงของพยัญชนะที่อยู่ระหว่างเส้นที่ 1 และ 2 ให้มีค่าเท่ากับ 1
- 2) เพิ่มขอบเขตบริเวณที่อยู่เหนือกว่าเส้นที่ 1 ไปอีกเป็นระยะ 1.2 เท่าของข้อ 1)
- 3) เพิ่มขอบเขตบริเวณที่อยู่ต่ำกว่าเส้นที่ 2 ไปอีกเป็นระยะ 0.7 เท่าของข้อ 1)
- 4) เพิ่มขอบเขตบริเวณด้านหน้าและด้านหลัง ไปอีกเป็นระยะ ด้านละ 1.2 เท่าของข้อ 1)

หลังจากขยายขอบเขตของกรอบตัวอักษรแล้ว จะได้จุดพิกัด 4 จุด คือ จุดบนซ้าย จุดล่างซ้าย จุดบนขวา และจุดล่างขวา เพื่อเป็นข้อมูลที่ใช้ในขั้นตอนต่อไป

- 5) **Geometric Transformation** ขั้นตอนนี้เป็นขั้นตอนของการปรับปรุงคุณภาพของภาพข้อความ ที่อาจมีการบิดเบี้ยวเนื่องจากการสแกนแบบเปอร์สเปกทีฟ และการหมุนเอียง โดยใช้การแก้ไขแบบสัมพรรค (Affine Transform) กับภาพที่มีกรอบล้อมรอบ และภาพที่มีการขยายกรอบล้อมรอบแล้ว

จากภาพป้ายภาษาไทยที่เป็นภาพอินพุต เมื่อผ่านขั้นตอนที่ 5) ของกระบวนการตรวจสอบข้อมูลแล้ว จะได้ภาพดังแสดงในรูปที่ 3.8 ซึ่งนำไปประมวลผลต่อในขั้นตอนการตัดแยกข้อความ (Text Extraction) ต่อไป



รูปที่ 3.8 ภาพข้อความที่ได้ทำการขยายกรอบแล้ว

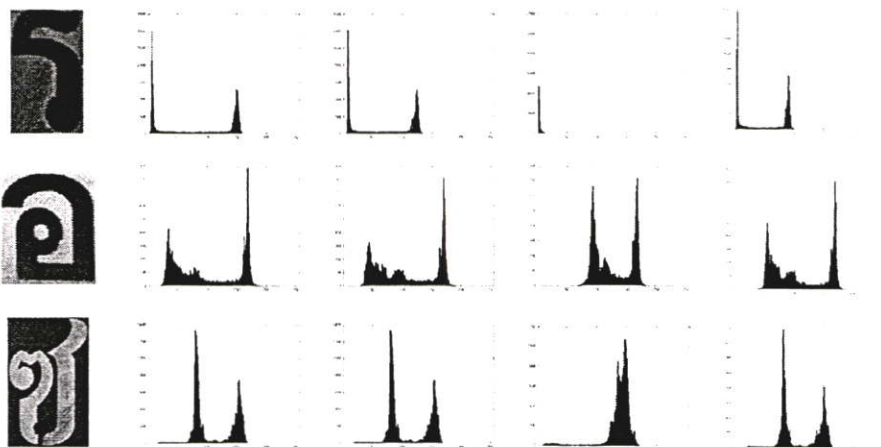
3.3 การตัดแยกข้อความ (Text Extraction)

กระบวนการการตัดแยกข้อความจากภาพ ประกอบด้วยขั้นตอน 3 ขั้นตอน ดังต่อไปนี้

- 1) **Color Segmentation** ในขั้นตอนนี้เป็นขั้นตอนของการตัดแยกตัวอักษรออกจากพื้นหลัง โดยใช้ความรู้พื้นฐานข้อ 2.1.3 แบบจำลองการกระจายแบบผสม เพื่อหาการกระจายของสีของตัวอักษรและพื้นหลัง เพื่อหาแบบจำลองสีที่เหมาะสมต่อการตัดพื้นหน้า (ข้อความ) ออกจากพื้นหลังของป้าย ซึ่งจะใช้แบบจำลองผสมเกาส์เซียนในการพิจารณาค่าการกระจายของสี ดังแสดงตัวอย่างในรูปที่ 3.9 โดยการพิจารณาจะใช้สมการ (3.12)

$$f(c) = (1 - \alpha)G_{Back}(\mu_b, \Sigma_b) + \alpha G_{Front}(\mu_f, \Sigma_f), (0 \leq \alpha \leq 1) \quad (3.12)$$

เมื่อ $G_{Back}(\cdot)$ เป็นการคำนวณการกระจายของสีของพื้นหลัง ส่วน $G_{Front}(\cdot)$ เป็นการคำนวณการกระจายของสีของพื้นหน้า



รูปที่ 3.9 แสดงค่าการกระจายของสีในแต่ละแบบจำลอง (R, G, B และ I) ของตัวอักษร “ร” “อ” และ “ช” จากป้ายที่ถ่ายจากสถานที่ และเวลาที่แตกต่างกัน

หลังจากการทดลองกับแบบจำลอง RGB และ HSI หลังจากทำการทดลองแล้วพบว่า แบบจำลองสีแดง สีเขียว สีน้ำเงิน และค่าความเข้มแสง (R, G, B, และ I ตามลำดับ) สามารถแยกตัวอักษรและพื้นหลังออกจากกันได้อย่างเหมาะสม โดยเลือกแบบจำลองทั้ง 4 เนื่องจากเป็นแบบจำลองที่ใช้การคำนวณในการแปลงแบบจำลองสีไม่มากนัก และเพียงพอต่อการตัดแยก ซึ่งหากเลือกจำนวนแบบจำลองสีที่น้อยลงจะทำให้ผลของการตัดแยกลดลง แต่หากเพิ่มจำนวนของแบบจำลองสีให้มากขึ้นจะเป็นการเพิ่มปริมาณการคำนวณทั้งการแปลงแบบจำลองสี และการคำนวณแบบจำลองการกระจาย ซึ่งจะทำให้เวลาในการตรวจจับและตัดแยกตัวอักษรช้าขึ้น และได้ผลลัพธ์ที่ดีใกล้เคียงกับจำนวนของแบบจำลองที่ได้เลือกไว้

จากนั้นจะพิจารณาคำนวณค่าในแต่ละแบบจำลองสี ๆ ละ 1 ครั้งต่อการพิจารณาส่วนของข้อความ เพื่อหาแบบจำลองสีที่เหมาะสมที่สุดในการตัดแยกข้อความ โดยจะเลือกแบบจำลองสีที่มีค่าความต่างของค่าเฉลี่ยที่มากที่สุด, $|\mu_b - \mu_r|$ มาใช้หลังจากนั้นจะทำการตัดแยกเป็นภาพขาวดำด้วยค่าเทรชโฮลด์ที่ได้จากการคำนวณด้วยอัลกอริทึม ของ Otsu ในหัวข้อ 2.1.5

- 2) **Fine Layout Analysis** เป็นขั้นตอนในการพิจารณาภาพโดยละเอียด เพื่อลดสัญญาณรบกวนและคัดเลือกตัวอักษรให้ถูกต้องมากขึ้น โดยเริ่มจากการนำภาพขาวดำไปวิเคราะห์หาส่วนประกอบที่เชื่อมติดกัน โดยการพิจารณาส่วนประกอบแต่ละตัวตามขอบเขตพื้นที่ที่ส่วนประกอบนั้นอยู่เพื่อให้สามารถคัดกรองสัญญาณรบกวน และตัดแยกตัวอักษรได้อย่างถูกต้องยิ่งขึ้น



รูปที่ 3.10 การแบ่งพื้นที่ในกรอบข้อความ

จากรูปที่ 3.10 พื้นที่ (Zone) ที่เป็นไปได้จะมีทั้งหมด 3 พื้นที่ ดังนี้

- 1) พื้นที่ของตัวพยัญชนะ ได้แก่ ตัวพยัญชนะไทยที่ไม่มีเชิง และสระ ได้แก่ อะ อา โอะ โอ แอะ แอ และ เอ พื้นที่ส่วนนี้พิจารณาจากตำแหน่งการวางตัวของตัวอักษรและอัตราส่วนความกว้างต่อสูงของสี่เหลี่ยมที่ล้อมรอบตัวอักษรให้ตรงกับสระ และพยัญชนะที่กำหนดไว้

2) พื้นที่ของสระและวรรณยุกต์ ได้แก่ อี อี้ อือ ไม้เอก ไม้โท ไม้ตรี ไม้จัตวา ตัวการ์ตูน และสระอำ พื้นที่ส่วนนี้พิจารณาจากอัตราส่วนความกว้างต่อสูงของสี่เหลี่ยมที่ล้อมรอบตัวอักษรให้ตรงกับสระและวรรณยุกต์ที่กำหนดไว้

3) พื้นที่ของพยัญชนะที่มีเชิง ได้แก่ ฐ ฎ ฏ ฎ ฤ และ ฃ และสระ ได้แก่ อุ อุ พื้นที่ส่วนนี้พิจารณาจากอัตราส่วนความกว้างต่อสูงของสี่เหลี่ยมที่ล้อมรอบตัวอักษรให้ตรงกับสระและส่วนของพยัญชนะที่กำหนดไว้ และรวมถึงการใช้ตำแหน่งการวางตัวร่วมด้วยเพื่อพิจารณาตัวอักษรที่มีเชิงของอักษรนั้นว่าเป็นตัวอักษรในพื้นที่นี้ด้วย

หลังจากผ่านขั้นตอนนี้แล้วจะได้ข้อมูลภาพขาวดำ ที่ผ่านพิจารณาหาตัวอักษรเพิ่มเติมจากขั้นตอนก่อนหน้านี้ ทั้งบริเวณด้านบน ด้านล่าง และด้านข้าง 2 ข้าง ตามลักษณะของการเขียนภาษาไทยเพื่อให้การตรวจจับและตัวแยกครอบคลุมตัวอักษรภาษาไทยและมีความสมบูรณ์ยิ่งขึ้น

- 3) **Text Component Extraction** เป็นขั้นตอนสุดท้ายของการตัดแยกข้อความออกจากพื้นหลัง โดยใช้ข้อมูลที่ได้จากการวิเคราะห์ส่วนที่เชื่อมต่อกันในขั้นตอนก่อนหน้านี้มาพิจารณาคัดแยก โดยนำภาพขาวดำที่มีข้อมูลการวิเคราะห์ส่วนประกอบที่เชื่อมติดกัน ที่ผ่านการกรองสัญญาณรบกวนออกไปแล้ว ซึ่งข้อมูลดังกล่าวประกอบด้วยข้อมูลพิกเซล ซึ่งเมื่อเสร็จขั้นตอนนี้ ผลลัพธ์ที่ได้จะเป็นส่วนของภาพ ที่เป็นส่วนประกอบย่อยของข้อความ ได้แก่ พยัญชนะ สระ และวรรณยุกต์ เป็นภาพค่าความเข้มแสง (Intensity Image) ของแบบจำลองที่เลือกไว้แล้วจากขั้นตอน Color Segmentation เพื่อเป็นข้อมูลนำเข้าสำหรับระบบรู้จำตัวอักษรที่ใช้ข้อมูลขาเข้าเป็นภาพค่าความเข้มแสง

บทที่ 4

การทดสอบการตรวจจับและตัดแยกตัวอักษรไทย จากภาพในฉากธรรมชาติ

4.1 ชุดข้อมูลที่ใช้ในการทดสอบ

ในการทดลองเพื่อตรวจจับและตัดแยกตัวอักษรไทยจากภาพที่ถ่ายในฉากธรรมชาติ งานวิจัยนี้ได้ทำการทดลองกับภาพดิจิทัลขนาด 640x480 พิกเซล ชนิดภาพ JPG จำนวน 200 ภาพ โดยภาพถูกถ่ายมาจากสถานที่ เวลา และมุมมองต่างๆ กัน

สำหรับภาพที่จะนำมาตรวจจับและตัดแยกข้อความภาษาไทย จะอยู่บนพื้นฐานต่อไปนี้

- 1) ภาพถูกถ่ายที่ระยะ และมุมมองกล้องที่จำกัดช่วงค่าหนึ่ง
- 2) พื้นหลังของข้อความเป็นสีพื้นเพียงสีเดียว ไม่มีลวดลายหรือหลายสีผสมกัน
- 3) ข้อความเป็นแนวตรง ไม่ถูกออกแบบมาในลักษณะที่บิดเบี้ยวโค้งงอ เพื่อความสวยงาม
- 4) ตัวอักษรถูกออกแบบให้มีความแตกต่างจากพื้นหลังทั้งสีและความเข้มแสง
- 5) ตัวอักษรแต่ละตัวถูกประกอบกันขึ้นเป็นข้อความที่มีขอบเขตชัดเจน
- 6) ตัวอักษรมีรูปแบบ (Font) ที่เป็นลักษณะของตัวพิมพ์ ไม่เป็นลายมือเขียน
- 7) ตัวอักษรในบริบท (Context) ของข้อความเดียวกันจะมีขนาดเท่ากัน มีความเข้มแสงที่ใกล้เคียงกัน แต่หากเป็นคนละข้อความ สีของตัวอักษรอาจจะแตกต่างกันได้

เมื่อพิจารณาภาพถ่ายแต่ละภาพแล้วพบว่า มีความหลากหลายของตัวอักษร ทั้งความแตกต่างของรูปแบบ ขนาด มุมมอง และตำแหน่งของตัวอักษร บางภาพตัวอักษรยังไม่อยู่ในลักษณะที่สมบูรณ์ อันเนื่องมาจากการเคลื่อนที่ของกล้องขณะถ่ายภาพ บางข้อความถูกบดบังด้วยเงา หรือวัตถุอื่นๆ ที่อยู่ในฉากภาพนั้น บางครั้งข้อความปรากฏอยู่บนวัตถุสามมิติ ดังนั้นตัวอักษรก็อาจมีการบิดเบี้ยวเนื่องจากวัสดุที่ใช้ทำพื้นหลังมีการบิดงอหรือผิววัสดุไม่เรียบ และพบว่าข้อความในภาพส่วนใหญ่มีการลู่ตามลักษณะการวางของมุมมอง นอกจากนี้ข้อความบางภาพมีความสกปรก มีจุดสะท้อนจากแสงแฟลชกล้องถ่ายรูป หลอดไฟ และดวงอาทิตย์ ปรากฏอยู่ในบริเวณข้อความ ซึ่งปัจจัยเหล่านี้ต่างส่งผลต่อความคมชัดของข้อความในภาพที่ถ่ายมาทั้งสิ้น

4.2 การทำงานของระบบการตรวจจับและตัดแยกตัวอักษรไทยในฉากธรรมชาติ

ในรูปที่ 4.1 แสดงผลที่ได้ในแต่ละขั้นตอน ของระบบการตรวจจับและตัดแยกตัวอักษรไทยในฉากธรรมชาติ



(ก)



(ข)



(ค)



(ง)



(จ)



(ฉ)



(ช)



(ซ)



(ฅ)



(ญ)



(ฎ)



(ฏ)

อนุบาลเด่นหล้า

(ฐ)

รูปที่ 4.1 แสดงขั้นตอนการทำงานของระบบตรวจจับและตัดแยกตัวอักษรไทยในฉากธรรมชาติ

จากรูปที่ 4.1 กระบวนการทำงานของระบบเริ่มจากนำภาพอินพุตรูป (ก) ซึ่งเป็นภาพขนาด 640x480 พิกเซล ชนิดภาพ JPG เข้าสู่ขั้นตอน Pre-processing เมื่อเสร็จแล้วจะได้ภาพระดับสีเทาดังรูป (ข) นำภาพที่ได้นี้เข้าสู่ขั้นตอน Edge Detection เพื่อหาเส้นขอบของวัตถุที่อยู่ในภาพ ผลที่ได้จะเป็นไปดังรูป (ค) จะเห็นได้ว่าข้อมูลภาพที่ได้ในขั้นนี้ เป็นเส้นขอบของบริเวณ

ต่างๆ ทั้งที่เป็นฉากภาพและตัวอักษรในกลุ่มข้อความปะปนกัน เพื่อที่จะตรวจจับบริเวณที่คาดว่า จะเป็นข้อความ และทำการแยกออกจากบริเวณที่ไม่ใช่ข้อความ จะต้องนำภาพเข้าสู่ขั้นตอนต่อไป คือ ขั้นตอน Candidate Text Detection ซึ่งจะตรวจจับตัวอักษรโดยพิจารณาอัตราส่วนของ ส่วนประกอบที่เชื่อมต่อกัน เส้นขอบของบริเวณปิดในภาพที่ไม่เป็นไปตามอัตราส่วนของ พยัญชนะไทยจะถูกคัดทิ้ง เมื่อเสร็จจากขั้นตอนนี้จะทำให้จำนวนเส้นขอบเหลือน้อยลงดังรูป (ง)-(จ) อย่างไรก็ตาม จะเห็นได้ว่ายังมีขอบภาพจำนวนหนึ่งที่ไม่ใช่ตัวอักษรติดอยู่ในภาพด้วย ขั้นต่อไปเป็นการวิเคราะห์หาเลขเอนท์ของกลุ่มตัวอักษรหรือ Layout Analysis โดยพิจารณาจาก คู่อักษรและแนวการวางตัว หากความชันของเส้นแนวข้อความและจุดความชันซึ่ง เพื่อค้นหา ข้อความในภาพวางตัวอยู่ในลักษณะใด ในตัวอย่างนี้เลขเอนท์ของข้อความมีลักษณะหมุนไปจาก แนวระดับปกติ และมีการลู่ออกแบบเปอร์สเปกทีฟ เมื่อผ่านการวิเคราะห์หาเลขเอนท์ของกลุ่มตัวอักษร แล้วจะได้ดังรูป (ข) จากนั้นจะกลับไปตัดส่วนของกลุ่มตัวอักษรในภาพต้นฉบับ จะได้ทั้งบริเวณ กลุ่มข้อความที่อยู่บนเส้นบรรทัดเดียวและขยายบริเวณให้ครอบคลุมสระวรรณยุกต์ที่อยู่สูงหรือ ต่ำกว่าเส้นบรรทัด ดังรูป (ข) กลุ่มข้อความนี้จะถูกนำไปแปลงในขั้นตอน Geometric Transformation เพื่อแก้ไขการเอียง หรือการลู่ออก ผลที่ได้จากขั้นนี้จะเป็นข้อความที่วางตัวในแนว ปกติ ดังรูป (ฉ) จากนั้นกลุ่มข้อความจะถูกนำไปเข้าสู่ขั้นตอน Color Segmentation ซึ่งตัวอย่างนี้ เลือกใช้ความเข้มแสงในแบบจำลองสีแดงในการตัดแยกตัวอักษรออกจากพื้นหลัง รูป (ญ) แสดง ภาพในความเข้มแสงสีแดง ขั้นต่อไปจะนำภาพไปทำเป็นภาพขาวดำและเข้าสู่ขั้นตอน Fine Layout Analysis เป็นการหาส่วนประกอบที่เชื่อมติดกันและจัดตั้งัญญาณรบกวน โดยการ พิจารณาให้ความสัมพันธ์กับส่วนประกอบตามขอบเขตพื้นที่ เพื่อให้การตรวจจับทำได้ถูกต้องยิ่งขึ้น ผลที่ได้เป็นไปดังรูป (ฉ) ขั้นตอนสุดท้ายเป็นขั้นตอนของ Text Component Extraction ซึ่งจะให้ ขอบเขตตัวอักษรในภาพขาวดำ ดังรูป (ฉ) สุดท้ายจะใช้ขอบเขตนี้ในการตัดแยกตัวอักษรจริงใน ภาพที่ได้เลือกแบบจำลองสีที่ถูกเลือกไว้แล้วก่อนหน้านี้ ผลลัพธ์สุดท้ายของระบบคือภาพตัว อักษรไทย ได้แก่ ภาพพยัญชนะ สระ และวรรณยุกต์ ถูกตัดแยกไว้ดังรูป (ฐ)

ในกระบวนการรู้จำ สามารถเลือกใช้ภาพตัวอักษรไทยที่เป็นภาพขาวดำ หรือภาพใน ความเข้มแสงในแบบจำลองสีที่เลือกไว้แล้ว รูปที่ 4.2 แสดงภาพขยายผลลัพธ์ที่ได้จากการ ตรวจจับและตัดแยกตัวอักษรไทยในฉากธรรมชาติ


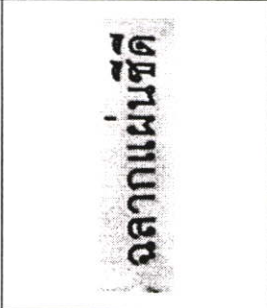
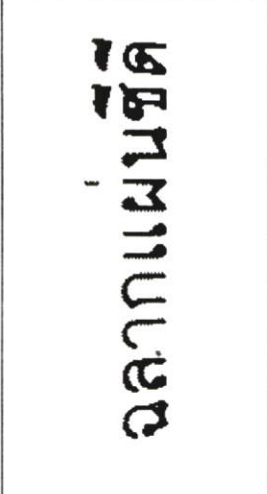



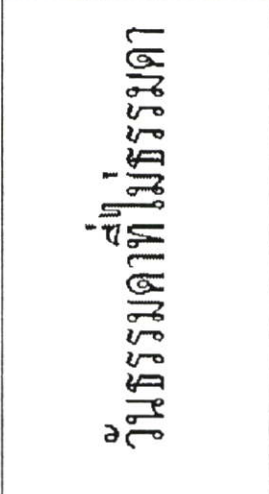
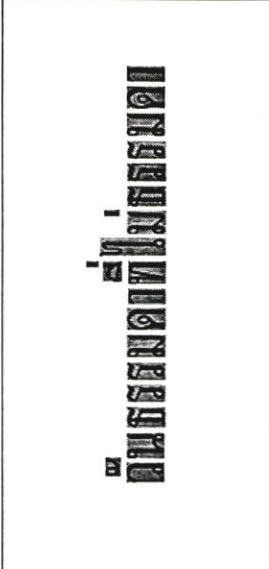


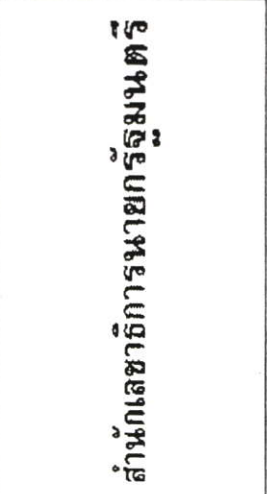



รูปที่ 4.2 ผลลัพธ์ที่ตรวจจับและตัดแยกตัวอักษรภาษาไทยจากรูปที่ 4.1

4.3 ผลการทดสอบ

ตัวอย่างผลการทดสอบระบบการตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติ เป็นไปดังตารางต่อไปนี้

ตารางที่ 4.1 ผลการทดลองระบบการตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติ กับภาพข้อความที่อยู่ในแนวนอนปกติ

ที่	ภาพ	กรอบข้อความ	ภาพขาวดำ	ภาพส่วนประกอบของข้อความที่ตัดแยกได้
1				
2				
3				

* 1. ภาพเบลอ 2. มีสัญลักษณ์อื่นที่ไม่ใช่ตัวอักษรปรากฏบนป้าย

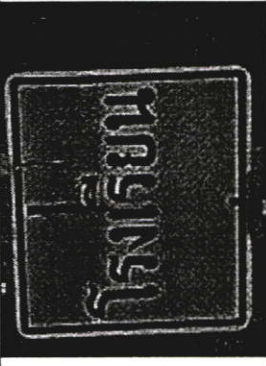


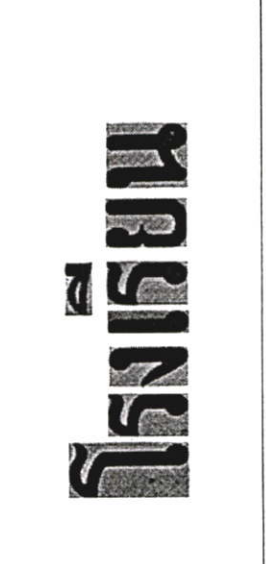

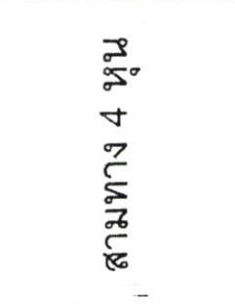

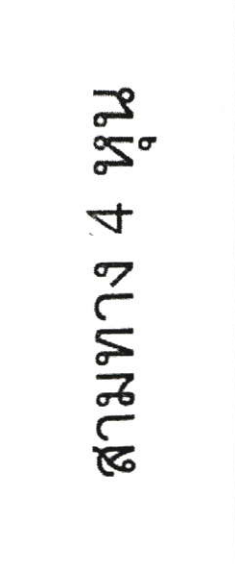




* 1. พื้นผิวป้ายไม่เรียบเนื่องจากวัสดุที่ใช้ทำป้าย 2. ป้ายมีเงาในตัวเอง 3. มีสัญลักษณ์อื่นที่ไม่ใช่ตัวอักษรปรากฏบนป้าย

4. ภาพค่อนข้างมืด ความคมชัดต่ำ (Low Contrast)

* 1. ป้ายมีพื้นหลังลาย (Texture) 2. ฉากภาพมีความซับซ้อนมาก




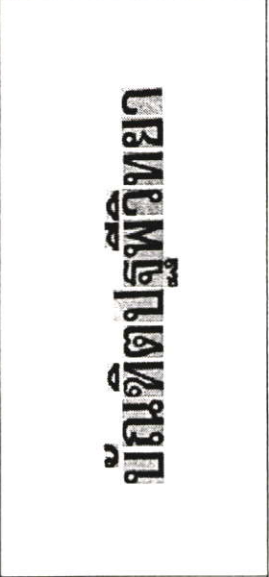



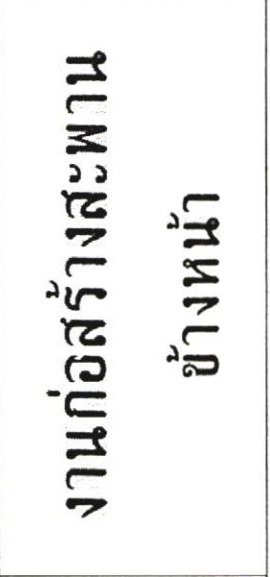




* สิ่งแวดล้อมที่มีผลกระทบต่อคุณภาพของภาพป้าย

ตารางที่ 4.1 (ต่อ)

ที่	ภาพ	กรอบข้อความ	ภาพขาวดำ	ภาพส่วนประกอบของข้อความที่ตัดแยกได้
4				
	* 1. ภาพถ่ายในสภาพแวดล้อมที่มีแสงสว่าง ความคมชัดต่ำ (Low Contrast) 2. ป้ายมีความสกปรก			
5				
	* 1. มีแสงสะท้อนจากพื้นผิวของป้าย 2. ฉากภาพมีความซับซ้อน 3. ข้อความในภาพมีการบิดเบือนเล็กน้อย			
6				
	* 1. ป้ายมีความสกปรก 2. มีสัญลักษณ์อื่นที่ไม่ใช่ตัวอักษรปรากฏบนป้าย 3. ฉากภาพมีความซับซ้อน			

* สิ่งแวดล้อมที่มีผลกระทบต่อคุณภาพของภาพป้าย

ตารางที่ 4.2 ผลการทดลองระบบการตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติ กับภาพข้อความที่มีการหมุนไปจากแนวอนุปกติ

ที่	ภาพ	กรอบข้อความ	ภาพขาวดำ	ภาพส่วนประกอบของข้อความที่ตัดแยกได้
1				
	* 1. ข้อความในภาพมีการหมุน 2. สัญลักษณ์บวกรวมที่เกิดขึ้นในกระบวนการรับภาพของกล้องดิจิทัล 3. เงาสะท้อนในป้าย			
2				
	* 1. ข้อความในภาพมีการหมุน 2. ป้ายมีรอยขีดข่วน			
3				
	* 1. ข้อความในภาพมีการหมุน 2. พื้นผิวป้ายไม่เรียบเนื่องจากวัสดุที่ใช้ทำป้าย 3. ป้ายมีความสกปรก			







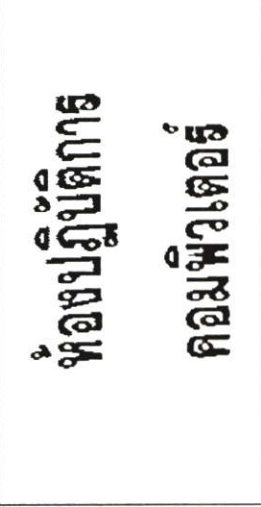

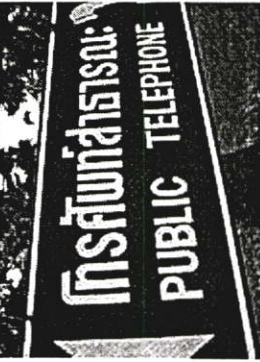



* ถึงแวดล้อมที่มีผลกระทบต่อคุณภาพของภาพป้าย

ตารางที่ 4.2 (ต่อ)

ที่	ภาพ	กรอบข้อความ	ภาพवाद้า	ภาพส่วนประกอบของข้อความที่ตัดแยกได้
4				
	* 1. ข้อความในภาพมีการหมุน 2. ป้ายมีรอยขีดข่วน 3. มีเส้นลวดปรากฏบนป้าย			
5				
	* 1. ข้อความในภาพมีการหมุน 2. ฉากภาพมีความซับซ้อน 3. มีสัญลักษณ์อื่นที่ไม่ใช่ตัวอักษรปรากฏบนป้าย 4. ตัวอักษรอยู่ชิดกันมาก			
6				
	* 1. ข้อความในภาพมีการหมุนเล็กน้อย 2. มีเงาจากสิ่งแวดล้อมสะท้อนบนพื้นป้าย			












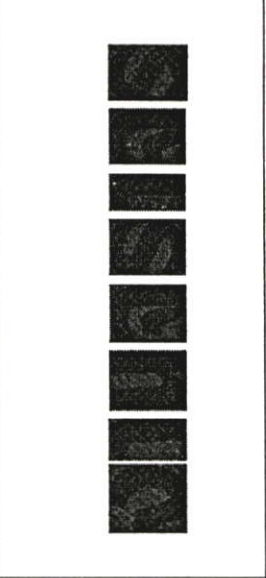
* ถึงแวดล้อมที่มีผลกระทบต่อดูแลคุณภาพของภาพป้าย

ตารางที่ 4.3 ผลการทดลองระบบการตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติ กับภาพข้อความที่มีการผู้เข้าแบบเปอร์สเปคทีฟ

ที่	ภาพ	กรอบข้อความ	ภาพขาวดำ	ภาพส่วนประกอบของข้อความที่ตัดแยกได้
1				
	* 1. ข้อความมีการผู้เข้าแบบเปอร์สเปคทีฟ 2. ฉากภาพมีความซับซ้อนมาก			
2				
	* 1. ข้อความมีการผู้เข้าแบบเปอร์สเปคทีฟ 2. ภาพเบลอ			
3				
	* 1. ข้อความมีการผู้เข้าแบบเปอร์สเปคทีฟ 2. ป้ายมีความสทกปรก 3. มีสัญลักษณ์อื่นที่ไม่ใช่ตัวอักษรปรากฏบนป้าย 4. ฉากภาพมีความซับซ้อน			











* สิ่งแวดล้อมที่มีผลกระทบต่อคุณภาพของภาพป้าย

ตารางที่ 4.3 (ต่อ)

ที่	ภาพ	กรอบข้อความ	ภาพขาวดำ	ภาพส่วนประกอบของข้อความที่ตัดแยกได้
4				
	<p>* 1. ข้อความมีการดูเข้าแบบเปอร์สเปคทีฟ 2. ภาพเบลอมาก 3. มี Artifact ปรากฏเนื่องจากเป็นภาพ JPG ที่ถูกบีบอัดสูง</p> <p>4. มีสัญลักษณ์ที่ไม่ใช่ตัวอักษรปรากฏบนป้าย 5. ฉากภาพมีความซับซ้อน</p>			
5				
	<p>* 1. ข้อความมีการดูเข้าแบบเปอร์สเปคทีฟ 2. ข้อความในภาพมีการหมุน 3. ฉากภาพมีความซับซ้อน 4. ป้ายมีลักษณะโปร่งแสง 5. มีเส้นลวดปรากฏบนป้าย</p>			
6				
	<p>* 1. ข้อความมีการดูเข้าแบบเปอร์สเปคทีฟ 2. ภาพถ่ายในสภาพแวดล้อมที่มีแสงสว่าง ความคมชัดต่ำ (Low Contrast) 3. ฉากภาพมีความซับซ้อน</p>			

* สิ่งแวดล้อมที่มีผลกระทบต่อคุณภาพของภาพป้าย

ตารางที่ 4.3 (ต่อ)

ที่	ภาพ	กรอบข้อความ	ภาพชาวตา	ภาพส่วนประกอบของข้อความที่ตัดแยกได้
7				
	* 1. ข้อความมีการดูเข้าแบบเปอร์สเปคทีฟ 2. ฉากภาพมีความซับซ้อน			
8				
	* 1. ข้อความมีการดูเข้าแบบเปอร์สเปคทีฟ 2. มีเงาจากสิ่งแวดล้อมสะท้อนบนพื้นป้าย			
9				
	* 1. ข้อความมีการดูเข้าแบบเปอร์สเปคทีฟเล็กน้อย 2. ป้ายมีความสกรปรก 3. ฉากภาพมีความซับซ้อน 4. ป้ายมีลักษณะโปร่งแสง			

* สิ่งแวดล้อมที่มีผลกระทบต่อคุณภาพของภาพป้าย

ตารางที่ 4.3 (ต่อ)

ที่	ภาพ	กรอบข้อความ	ภาพชาวตา	ภาพส่วนประกอบของข้อความที่ตัดแยกได้
10				
	* 1. ข้อความมีการดูเข้าแบบเปอร์สเปคทีฟ 2. พื้นผิวป้ายไม่เรียบเนื่องจากวัสดุที่ใช้ทำป้าย 3. ป้ายมีเงาในตัวเอง			
11				
	* 1. ข้อความมีการดูเข้าแบบเปอร์สเปคทีฟ 2. ป้ายมีความสกปรก			
12				
	* 1. ข้อความมีการดูเข้าแบบเปอร์สเปคทีฟ 2. ป้ายมีความสกปรก 3. มีสัญลักษณ์อื่นที่ไม่ใช่ตัวอักษรปรากฏบนป้าย			

* สิ่งแวดล้อมที่มีผลกระทบต่อคุณภาพของป้าย

ตารางที่ 4.4 เปรียบเทียบผลการทดสอบการตรวจจับตัวอักษร กับภาพถ่ายจำนวน 200 ภาพ

รูปแบบการหากลุ่มข้อความในการตรวจจับตัวอักษร		จำนวนที่ตรวจจับได้ จาก 280 ข้อความ	
		จำนวน	% ความถูกต้อง
1.	พิจารณาความสัมพันธ์ของตัวอักษรเฉพาะที่อยู่ในแนวนอนปกติ	161	57.50
2.	พิจารณาความสัมพันธ์ของตัวอักษรตามวิธีการที่นำเสนอ	238	85.00

การค้นหาข้อความ จำเป็นที่จะต้องหากลุ่มของตัวอักษรให้สมบูรณ์ทั้งข้อความ จึงจะสามารถตีความข้อความนั้นได้อย่างมีความหมาย ดังนั้น จึงพิจารณาการตรวจจับข้อความให้มีความสมบูรณ์ทั้งข้อความ จากผลการทดสอบเปรียบเทียบในตาราง 4.4 เมื่อพิจารณาความสัมพันธ์ของตัวอักษรเฉพาะที่อยู่ในแนวนอนปกติ ระบบให้ความถูกต้องในการตรวจจับข้อความอย่างครบถ้วน คิดเป็น 57.50 % โดยพบว่า 42.50 % ไม่สามารถค้นหาภาพข้อความที่มีการหมุนหรือการดูเข้าแบบเปอร์สเปคทีฟได้อย่างสมบูรณ์ ซึ่งต่างจากวิธีที่นำเสนอที่สามารถค้นหาข้อความได้ถูกต้องสมบูรณ์ถึง 85.00 % เนื่องจากการพิจารณาความสัมพันธ์ในรูปแบบของการวางตัวในแนวเส้นตรง และมีความทนทานต่อการหมุนและการดูเข้าแบบเปอร์สเปคทีฟ

ตารางที่ 4.5 เปรียบเทียบผลการทดสอบการตัดแยกตัวอักษรในแบบจำลองสีต่างๆ กับภาพถ่ายจำนวน 200 ภาพ

รูปแบบการเลือกแบบจำลอง สำหรับการตัดแยกตัวอักษร		จำนวนที่ตรวจจับและตัดแยกได้			
		จาก 3,372 ตัวอักษร		จาก 280 ข้อความ	
		จำนวน	% ความถูกต้อง	จำนวน	% ความถูกต้อง
1.	เลือกแบบจำลองสีเทา	2,597	77.01	238	85.00
2.	เลือกแบบจำลองสี R G B และ I ที่เหมาะสมด้วยอัลกอริทึม EM ตามวิธีการที่นำเสนอ	3,030	89.86	238	85.00

จากตารางที่ 4.5 พบว่า การใช้วิธีการคัดเลือกแบบจำลองสี R G B และ I ด้วยวิธี EM (Expectation Maximization) ให้ผลที่ดีกว่าการใช้เพียงแบบจำลองสีเทาของเพียงอย่างเดียว เนื่องจากภาพข้อความ ไม่ได้มีความแตกต่างเพียงความเข้มแสง แต่ยังมีหลากหลายเรื่องสีและแสงเงาในภาพ จึงทำให้แต่ละแบบจำลองสีมีความเหมาะสมกับภาพที่แตกต่างกัน

จากการทดสอบในขั้นตอนของการแก้ไขภาพแบบสัมพรรค (Affine Rectification) นั้น ไม่ได้ส่งผลต่อความถูกต้องของการตรวจจับตัวอักษร แต่ส่งผลต่อกระบวนการตัดแยกตัวอักษร เนื่องจากหากไม่ทำการแก้ไขภาพข้อความที่หมุนหรือลู่เข้าแบบเปอร์สเปกทีฟแล้ว จะทำให้ขอบเขตของภาพมีขนาดใหญ่และกินเนื้อที่ของภาพในส่วนที่ไม่เกี่ยวข้องกับกลุ่มของข้อความนั้น ซึ่งอาจจะเป็นกลุ่มของสัญญาณรบกวนอื่นๆ เช่น ใบไม้ หรือฉากธรรมชาติที่มีความคล้ายคลึงกับตัวอักษร ทำให้การตัดแยกเกิดผลที่ผิดพลาดขึ้น อีกทั้งความพิเศษของรูปแบบของการวางตัวอักษรในภาษาไทยที่มีการวางตัวในระดับของบรรทัดที่ต่างกันและขนาดตัวอักษรที่ต่างกัน ทำให้การพิจารณาแยกแยะตัวอักษรออกจากสัญญาณรบกวนเป็นไปได้ยาก

วิธีการคัดกรองสัญญาณรบกวนที่ไม่ใช่บริเวณข้อความภาษาไทยที่เสนอนั้น ประกอบด้วยการแก้ไขภาพแบบสัมพรรค (Affine Rectification) หลังจากนั้นจะได้ภาพข้อความที่คาดว่าจะอยู่ในแนวระดับเพื่อเข้าสู่กระบวนการตัดแยกตัวอักษรและขั้นตอน Fine Layout Analysis เพื่อลดสัญญาณรบกวนและได้กลุ่มตัวอักษรที่ถูกต้อง

จากการทดสอบระบบการตรวจจับและตัดแยกตัวอักษรภาษาไทยในฉากธรรมชาติทางขั้นตอนในบทที่ 3 กับภาพถ่ายจำนวน 200 ภาพ พบว่า ผลการทดสอบเป็นไปดังตารางที่ 4.6

ตารางที่ 4.6 ผลการทดสอบระบบกับภาพถ่ายจำนวน 200 ภาพ

	จำนวนทั้งหมด	จำนวนที่ตรวจจับ และตัดแยกได้	% ความถูกต้อง
ข้อความ	280	238	85.00
ตัวอักษร	3,372	3,030	89.86

ภาพที่นำมาทดสอบ บางภาพมีเพียงข้อความเดียว บางภาพมีมากกว่าหนึ่งข้อความ ผลการทดสอบในตารางที่ 4.6 หากผ่านกระบวนการที่ได้นำเสนอนั้นพบว่า ภาพจำนวน 200 ภาพ มีบริเวณข้อความรวมทั้งสิ้น 280 ข้อความ ซึ่งระบบสามารถตรวจจับและตัดแยกข้อความได้อย่างถูกต้องจำนวน 238 ข้อความ คิดเป็น 85.00 % หากพิจารณาเป็นตัวอักษร พบว่า จากจำนวนตัวอักษรรวมทุกภาพ 3,372 ตัว ระบบสามารถตรวจจับและตัดแยกตัวอักษรได้อย่างถูกต้องจำนวน 3,030 ตัว คิดเป็น 89.86 %

อย่างไรก็ตาม มีภาพส่วนหนึ่งที่ระบบไม่สามารถตรวจจับและตัดแยกข้อความได้อย่างถูกต้อง ในตารางที่ 4.7 แสดงตัวอย่างของการตรวจจับและตัดแยกตัวอักษรที่ผิดพลาด และสาเหตุที่เป็นไปได้ในแต่ละกรณี

ตารางที่ 4.7 แสดงตัวอย่างของการตรวจจับและตัดแยกตัวอักษรที่ผิดพลาด

ที่	ภาพ	กรอบข้อความ	ภาพขาวดำ	ภาพส่วนประกอบของข้อความที่ตัดแยกได้
1				
2				
3				




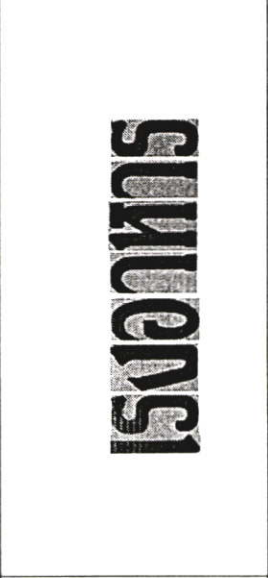




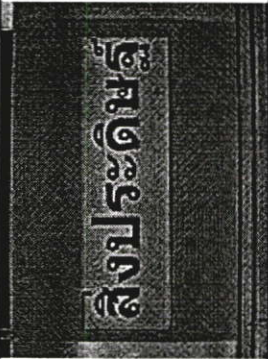
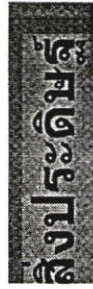


* ป้ายทำจากวัสดุที่มันวาว มีแสงสะท้อนของหลอดฟลูออโรสเซนส์ ทำให้การตรวจจับผิดพลาดเนื่องจากความไม่สม่ำเสมอของแสงบนป้าย

* ภาพนี้ถ่ายในเวลากลางวันมีแสงสว่างน้อย ภาพเบลอ ส่งผลต่อความคมชัดของตัวอักษร นอกจากนี้มีแสงไฟส่องกระทบป้ายเป็นบางบริเวณ ทำให้เกิดความไม่สม่ำเสมอของแสงบนป้าย ทำให้การตัดแยกทำได้ไม่สมบูรณ์

* ภาพเบลอเนื่องจากความสกรกของเลนส์ถ่ายภาพบริเวณพญูชนะ "ผ" ทำให้ระบบไม่สามารถตรวจจับและตัดแยกตัวอักษรนี้ได้ เนื่องจากมีระดับความเข้มแสงต่างกับตัวอักษรอื่นๆ ในข้อความเดียวกัน

* สาเหตุที่ทำให้การตรวจจับและตัดแยกตัวอักษรผิดพลาด

ตารางที่ 4.7 (ต่อ)

ที่	ภาพ	กรอบข้อความ	ภาพขาวดำ	ภาพส่วนประกอบของข้อความที่ตัดแยกได้
4				
	* มีแสงสะท้อนบริเวณส่วนบนของป้าย ทำให้ความเข้มแสงของสระ “โ” ไม่สม่ำเสมอ จึงไม่สามารถตรวจจับและตัดแยกได้อย่างถูกต้อง			
5				
	* ภาพถูกถ่ายภาพขณะรถวิ่ง ทำให้ภาพเบลอ ระบบสามารถตรวจจับและตัดแยกตัวอักษร ได้เฉพาะบรรทัดบน และจากระยะของการถ่ายภาพที่ค่อนข้างไกล ทำให้ตัวอักษรในบรรทัดล่างมีขนาดเล็ก ช่องไฟแคบ ไม่สามารถตรวจจับได้เนื่องจากตัวอักษรอยู่ติดกันเป็นกลุ่ม			
6				
	* สรวางติดกับพยัญชนะ ทำให้อัตราส่วนของส่วนที่เชื่อมติดกันไม่เกินไปตามที่กำหนด ระบบจึงไม่นำบริเวณ “สี” มาพิจารณา			

* สาเหตุที่ทำให้การตรวจจับและตัดแยกตัวอักษรผิดพลาด

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

งานวิจัยนี้ได้นำเสนอวิธีการตรวจจับและตัดแยกตัวอักษรไทยในฉากธรรมชาติ โดยมีวัตถุประสงค์เพื่อสร้างระบบที่สามารถตรวจจับและตัดแยกบริเวณข้อความในภาพถ่ายออกมาได้อย่างถูกต้อง แม้ว่าภาพจะถูกรบกวนจากปัจจัยต่างๆ ในสิ่งแวดล้อมธรรมชาติ โดยกระบวนการเริ่มจากการนำภาพจากกล้องถ่ายภาพดิจิทัล หรือ โทรศัพท์มือถือที่มีอุปกรณ์ถ่ายภาพชนิด JPEG ขนาด 640 x 480 พิกเซล เข้าสู่ขั้นตอนการเตรียมภาพ โดยแปลงจากภาพในแบบจำลอง RGB เป็นภาพระดับสีเทา เพื่อลดจำนวนช่องสัญญาณสี และพิจารณาเฉพาะค่าความเข้มแสง จากนั้นใช้ตัวกรองแบบเกาส์เซียนเพื่อลดสัญญาณรบกวนและบล็อกสี่เหลี่ยม (Artifact) โดยทำให้ภาพเบลอ แม้ว่าสัญญาณรบกวนจะลดลง แต่รายละเอียดเส้นขอบในภาพก็จะเบลอด้วย จึงต้องปรับปรุงภาพด้วยตัวกรองที่ทำให้ภาพคมชัด (Sharpening filter) จากนั้นหาเส้นขอบด้วยวิธีการ Laplacian of Gaussian (LoG) ซึ่งเป็นการหาเส้นขอบแบบอนุพันธ์อันดับสอง ที่ให้ผลลัพธ์เป็นเส้นขอบที่บางและคมชัด การหาเส้นขอบด้วยวิธีนี้เหมาะสมกับงานวิจัยนี้เนื่องจากตัวอักษรในภาพระยะชิดติดกัน หากใช้วิธีการอื่นจะได้เส้นขอบที่หนาทำให้ได้ผลลัพธ์ที่ยากต่อการตัดแยกตัวอักษร เมื่อได้เส้นขอบบริเวณต่างๆ แล้วจะเข้าสู่ขั้นตอนการคัดกรองเส้นขอบที่คาดว่าจะเป็นตัวอักษร โดยการหาส่วนประกอบที่เชื่อมติดกัน และพิจารณาค่าอัตราความกว้างต่อความสูง และอัตราพื้นที่ต่อพื้นที่สี่เหลี่ยมปิดล้อม เนื่องจากต้องการคัดกรองให้เหลือเพียงองค์ประกอบที่เหมาะสมต่อการพิจารณาการวางตัวของตัวอักษร ซึ่งวิธีการคัดกรองนี้ถูกออกแบบมาให้เหมาะสมกับระบบที่ใช้ในการตรวจจับตัวอักษรไทยโดยเฉพาะ ผู้วิจัยได้เสนอวิธีใหม่ในขั้นตอนการวิเคราะห์แนวการวางตัวของข้อความโดยพิจารณาจากความสัมพันธ์ของตัวอักษรแบบเชิงเส้น (สมการเส้นตรง) ให้สามารถตรวจจับข้อความได้แม้จะมีหมุนและการถูเข้าแบบเปอร์สเปคทีฟ เมื่อได้แนวข้อความแล้วจะทำการแก้ไขแบบสัมพรรค (Affine Rectification) ซึ่งเป็นการแปลงแบบเรขาคณิต โดยการกำหนดจุดสี่จุดที่มุมกรอบของข้อความ แล้วคำนวณหาเมตริกซ์ขนาด 4 x 4 ที่เป็นข้อมูลพื้นฐานสำหรับเข้าสู่กระบวนการแปลงพิกัดภาพโดยพิจารณาจุดของภาพใหม่แบบ Bilinear เพื่อได้ภาพของข้อความด้านที่ตั้งฉากกับมุมมองมีความเหมาะสมต่อการตัดแยกตัวอักษรในภาษาไทยที่มีรูปแบบการสร้างคำและประโยคที่ซับซ้อน

เมื่อได้เฉพาะบริเวณข้อความที่แน่นอนแล้ว จะเข้าสู่กระบวนการตัดแยกตัวอักษร โดยเลือกแบบจำลองสีที่เหมาะสมจากแบบจำลอง R G B และ I โดยอัลกอริทึม Expectation Maximization (EM) เพื่อหาค่าพารามิเตอร์สำหรับการกระจายแบบผสม การคัดเลือกแบบจำลองที่เหมาะสมกับภาพๆ หนึ่งนั้น จะพิจารณาจากค่าความแตกต่างของสองกลุ่มการกระจาย คือ สีตัวอักษรและพื้นหลัง หากแบบจำลองใดให้ค่าความแตกต่างมากที่สุดแสดงว่าควรเลือกใช้

แบบจำลองนั้นในการตัดแยกตัวอักษรออกจากพื้นหลัง จากนั้นหาค่าเทรชโฮลด์โดยใช้อัลกอริทึมของ Otsu ซึ่งเป็นอัลกอริทึมสำหรับหาค่าเทรชโฮลด์แบบปรับเปลี่ยนได้ (Adaptive Thresholding Algorithm) เพื่อให้ได้ภาพขาวดำสำหรับการตัดแยกตัวอักษรต่อไป เพื่อให้สามารถหาข้อความได้ครบถ้วนถูกต้องมากยิ่งขึ้น จะต้องใช้การพิจารณาข้อความแบบละเอียด เพื่อลดบริเวณที่ไม่ใช่ตัวอักษรไทย โดยพิจารณาจากรูปแบบประโยคและการวางตำแหน่งของตัวอักษรในขอบเขตพื้นที่และกลุ่มความเป็นไปได้ของตัวอักษรในแต่ละบริเวณ เนื่องจากภาพอาจมีบริเวณที่ไม่ใช่ตัวอักษรเพิ่มขึ้นจากขั้นตอนการขยายขอบเขตเพื่อให้ครอบคลุมสระ และวรรณยุกต์สำหรับภาษาไทย ซึ่งวิธีการนี้ออกแบบเพื่อให้เหมาะสมสำหรับการตัดแยกภาษาไทยจากภาพถ่าย ขั้นตอนสุดท้ายเป็นขั้นตอนการตัดแยกตัวอักษร เป็นการแยกส่วนของภาพออกมาเป็นส่วนๆ ได้แก่ พยัญชนะ สระ และวรรณยุกต์ เพื่อเตรียมให้กับระบบรู้จำตัวอักษร วิธีการในขั้นนี้จะทำโดยนำข้อมูลตัวอักษรจากส่วนประกอบเชื่อมติด มากำหนดพิกัด แล้วแยกส่วนประกอบนั้นออกจากภาพการส่องสว่างของแบบจำลองสีที่ได้เลือกไว้ในขั้นตอนก่อนหน้า ผลลัพธ์ที่ได้จากการตัดแยกภาพข้อความ จะได้ภาพของพยัญชนะ สระ และวรรณยุกต์ไทย ที่เป็นภาพในแบบจำลองความเข้มแสงที่เลือกไว้ เหตุที่ต้องตัดแยกตัวอักษรในภาพค่าความเข้มแสงนั้นก็เนื่องจากระบบรู้จำตัวอักษรในปัจจุบันจะใช้ทั้งข้อมูลภาพขาวดำ และภาพการส่องสว่างเป็นข้อมูลเข้าพื้นฐาน

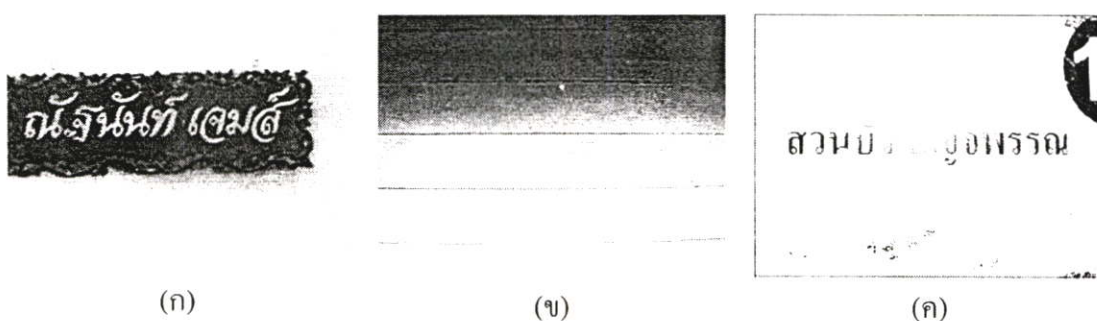
5.1 ปัจจัยที่มีผลต่อคุณภาพของภาพป้าย

จากการทดลองพบว่า ระบบจะสามารถตรวจจับและคัดแยกข้อความจากป้ายได้ดี ถ้าภาพอินพุตมีคุณภาพดี ดังนั้นขั้นตอนของการเตรียมภาพอินพุตหรือการถ่ายภาพ จึงเป็นขั้นตอนที่สำคัญมาก ทั้งนี้เนื่องจากการถ่ายภาพป้ายจากฉากธรรมชาติ มีความแตกต่างกับระบบการรับภาพทั่วไป เช่น การสแกนภาพเอกสาร โดยใช้เครื่องสแกน เนื่องจากการสแกนเอกสารนั้น เราสามารถกำหนดปัจจัยแวดล้อมต่างๆ อันจะมีผลต่อคุณภาพของภาพ เช่น ความเข้มแสง (Intensity) และความละเอียด (Resolution) เพื่อให้ได้ไฟล์ภาพสแกนที่มีคุณภาพดี แต่การถ่ายภาพป้ายจากฉากธรรมชาตินั้นพบว่า มีหลายปัจจัยที่มีผลกระทบโดยตรงต่อคุณภาพของภาพถ่าย ทำให้ภาพถ่ายมีคุณภาพไม่ดีเท่ากับภาพสแกน และปัจจัยเหล่านี้เป็นสิ่งที่ควบคุมได้ยาก

สรุปปัจจัยที่มีผลกระทบต่อคุณภาพของภาพป้ายที่ถ่ายในฉากธรรมชาติ มีดังนี้

- 1) ความชัดของฉากภาพ
- 2) สี และลวดลายของพื้นหลังป้าย
- 3) ข้อความในภาพไม่ชัดเจน เนื่องจากสีที่ซีดจาง หรือเกิดจากถ่ายภาพขณะที่มือสั่นหรือกล้องมีการเคลื่อนที่ ทำให้ข้อความที่ได้มีลักษณะมัว (Blur)
- 4) ป้ายเก่า มีรอยขีดข่วน มีสนิมหรือคราบสกปรก หรือมีวัสดุสำหรับยึดป้าย เช่น ตะปู น็อต ลวด ปรากฏอยู่ป้าย
- 5) ภาพที่ถูกถ่ายในสภาพแสงน้อย ทำให้ภาพมีความคมชัดต่ำ (Low Contrast)

- 6) แสงสะท้อนบนป้ายอันเนื่องมาจากแสงอาทิตย์ หลอดไฟ หรือแสงแฟลช
- 7) เงาของวัตถุอื่นที่พาดทับลงบนป้าย
- 8) การถูกวัตถุอื่นบังข้อความในป้าย เช่น ใบไม้ หรือสายไฟฟ้า
- 9) ความบิดเบี้ยวของภาพ (Image Distortion) ที่เกิดในขณะที่ถ่ายภาพ ซึ่งเกิดจากตัวรับภาพ หรือการลู่เข้าของข้อความในภาพอันเนื่องมาจากทิศทางของมุมมองกล้อง
- 10) วัสดุที่นำมาทำป้ายบิดเบี้ยว ยับ เช่น ป้ายผ้าหรือไวนิลที่มีรอยยับ ทำให้ข้อความเกิดการบิดเบี้ยวได้ หรือป้ายโลหะที่หักงอ ก็ทำให้การตรวจจับข้อความผิดพลาดเช่นเดียวกัน
- 11) การวางตัวของข้อความที่มีหลากหลายรูปแบบ



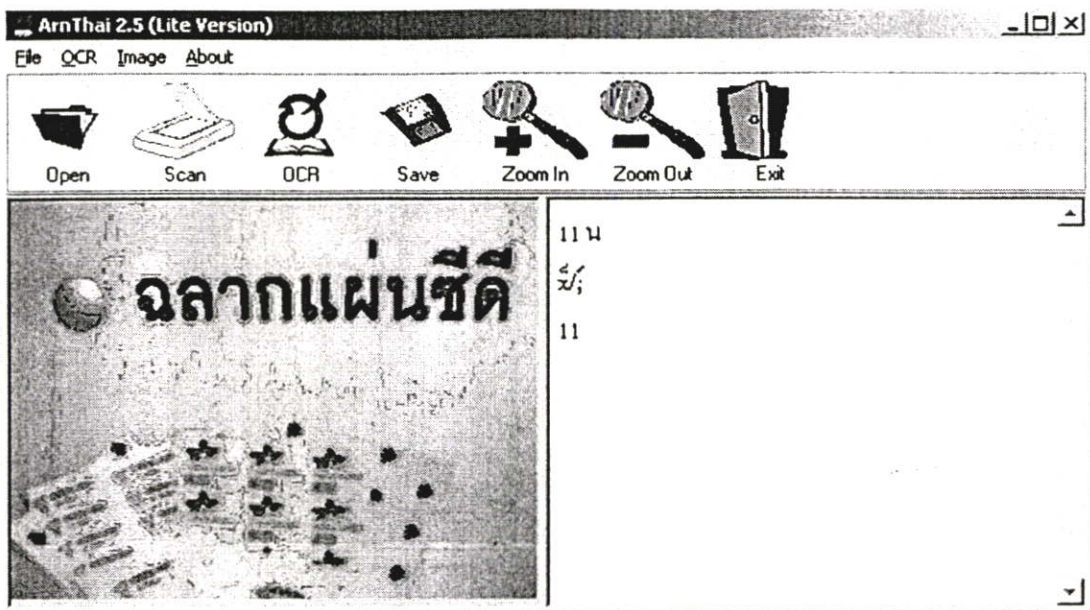
รูปที่ 5.1 ป้ายที่ระบบไม่สามารถตรวจจับข้อความได้

ในรูปที่ 5.1 เป็นตัวอย่างป้ายที่ระบบไม่สามารถตรวจจับข้อความได้ ในรูป (ก) เป็นป้ายที่ใช้ตัวอักษรแบบสวยงาม ไม่เป็นมาตรฐาน หากที่ระบบจะสามารถคัดกรองและตรวจจับได้ รูป (ข) เป็นป้ายที่กล้องทำการวัดแสงไม่เหมาะสม ทำให้ภาพป้ายที่ถ่ายมานั้นมีความแตกต่างของความเข้มแสงระหว่างพื้นหน้าและพื้นหลังค่อนข้างน้อย ไม่สามารถแยกตัวอักษรออกจากพื้นหลังได้ และในรูป (ค) เป็นป้ายที่มีแสงสะท้อนจากแฟลชกล้องถ่ายรูป ทำให้เกิดจุดสว่างจางกลางภาพ ทำให้ข้อความขาดหายไป

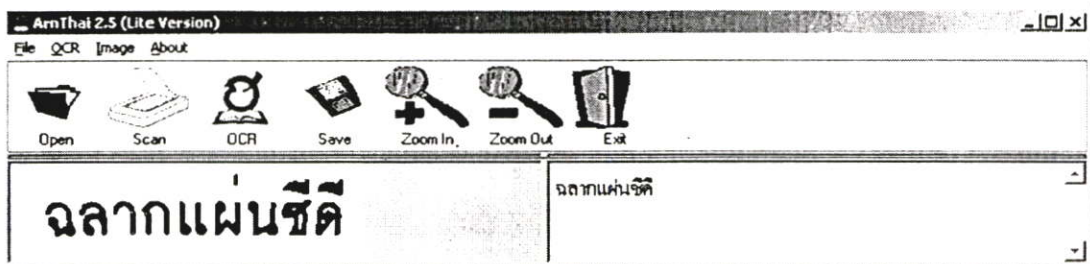
5.2 ข้อเสนอแนะในการทำวิจัยต่อไป

แนวทางการทำวิจัยต่อไปในอนาคต สำหรับการพัฒนาระบบรู้จำตัวอักษร (OCR) โดยใช้ภาพเอาต์พุตที่ได้จากงานวิจัยนี้ไปใช้ในการรู้จำตัวอักษรภาษาไทย ซึ่งจะทำให้ได้ระบบที่สามารถอ่านข้อความจากป้ายภาษาไทยได้

ผู้วิจัยได้ทำการทดสอบการรู้จำตัวอักษรไทยจากภาพในฉากธรรมชาติ โดยใช้ซอฟต์แวร์ “อ่านไทย” เวอร์ชัน 2.5 เปรียบเทียบผลการรู้จำตัวอักษรที่ได้จากภาพต้นฉบับที่ไม่ได้ทำการตรวจจับตัวอักษรและภาพที่ผ่านการตรวจจับและตัดแยกตัวอักษรแล้ว ได้ผลดังรูปที่ 5.2 และ 5.3

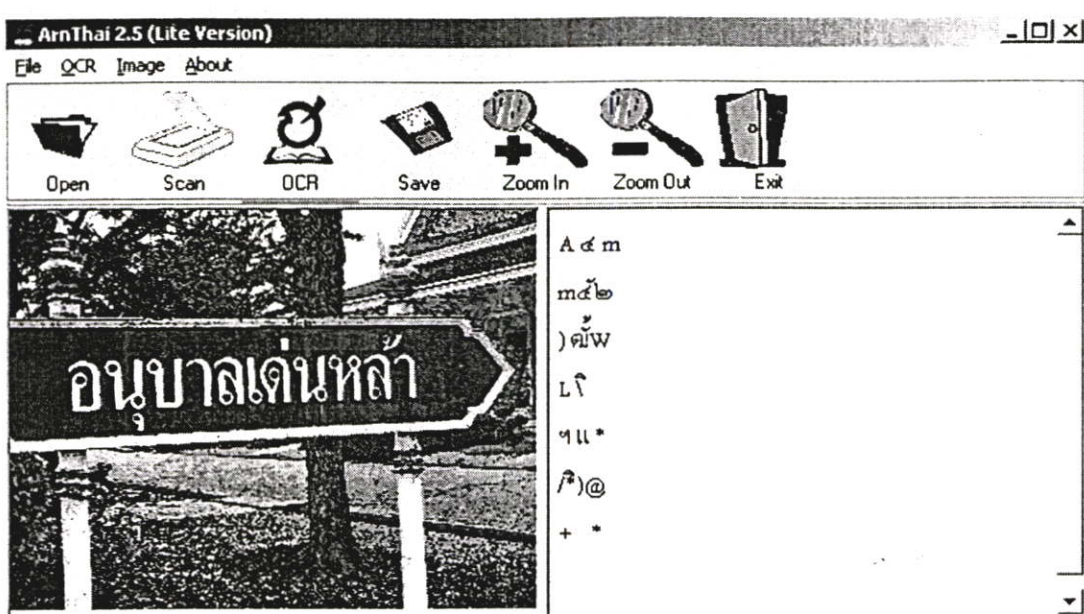


(ก)

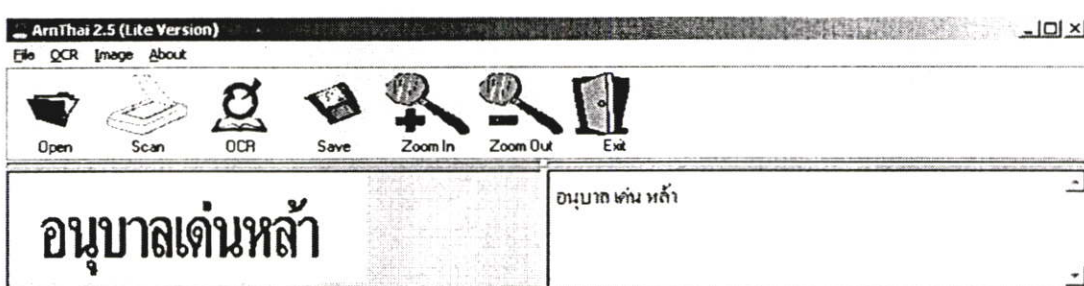


(ข)

- รูปที่ 5.2 ตัวอย่างการรู้จำตัวอักษรไทยจากป้ายในฉากธรรมชาติ โดยใช้ซอฟต์แวร์ “อ่านไทย”
- (ก) ช่องซ้ายมือเป็นภาพต้นฉบับที่ไม่ได้ผ่านระบบการตรวจจับและตัดแยกตัวอักษรไทย สำหรับช่องขวามือเป็นผลการรู้จำตัวอักษร
- (ข) ช่องซ้ายมือเป็นภาพข้อความที่ระบบตรวจจับและตัดแยกได้ สำหรับช่องขวามือเป็นผลการรู้จำตัวอักษร



(ก)



(ข)

รูปที่ 5.3 ตัวอย่างการรู้จำตัวอักษรไทยจากป้ายในฉากธรรมชาติ โดยใช้ซอฟต์แวร์ “อ่านไทย”

(ก) ช่องซ้ายมือเป็นภาพต้นฉบับที่ไม่ได้ผ่านระบบการตรวจจับและตัดแยกตัวอักษรไทย สำหรับช่องขวามือเป็นผลการรู้จำตัวอักษร

(ข) ช่องซ้ายมือเป็นภาพข้อความที่ระบบตรวจจับและตัดแยกได้ สำหรับช่องขวามือเป็นผลการรู้จำตัวอักษร

จะเห็นได้ว่า ซอฟต์แวร์ “อ่านไทย” ไม่สามารถให้ผลการรู้จำตัวอักษรที่ดีได้กับภาพต้นฉบับ เนื่องจากซอฟต์แวร์ไม่มีวิธีการหาบริเวณข้อความแบบอัตโนมัติ เมื่อนำภาพต้นฉบับผ่านการตรวจจับและตัดแยกตัวอักษรก่อน แล้วนำผลของส่วนภาพที่ตัดแยกได้ไปเข้าซอฟต์แวร์นี้อีกครั้ง ปรากฏว่าซอฟต์แวร์สามารถรู้จำตัวอักษรได้ดี สำหรับตัวอย่างอื่นๆ ของการรู้จำตัวอักษรจากภาพถ่ายป้ายในฉากธรรมชาติ แสดงในตารางที่ 5.1

อย่างไรก็ตาม ซอฟต์แวร์ “อ่านไทย” จะให้ผลการรู้จำที่ดีกับบางรูปแบบ (Font) ของตัวอักษรเท่านั้น ดังนั้นในการพัฒนาระบบรู้จำตัวอักษรภาษาไทย ควรจะคำนึงถึงรูปแบบตัวอักษรที่หลากหลายบนป้ายทั่วไปด้วย

ระบบที่เสนอในวิทยานิพนธ์นี้สามารถตรวจจับและตัดแยกตัวอักษรไทยที่อยู่ในฉากธรรมชาติสำหรับภาพที่มีขนาด 640x480 พิกเซล ซึ่งมีขนาดเล็ก อย่างไรก็ตาม ในปัจจุบันอุปกรณ์รับภาพได้รับการพัฒนาไปอย่างรวดเร็ว ทำให้สามารถถ่ายภาพได้ที่ขนาดใหญ่ขึ้น ซึ่งหากต้องการตรวจจับตัวอักษรในภาพให้ได้ความแม่นยำสูงโดยใช้ภาพขนาดใหญ่ จำเป็นต้องพัฒนาระบบให้มีความเหมาะสมกับภาพที่มีข้อมูลปริมาณมาก โดยอาจพัฒนาให้ระบบตรวจจับกรอบป้ายก่อน ดังตัวอย่างในรูปที่ 5.4 เพื่อลดข้อมูลที่ไม่ใช่บริเวณข้อความในภาพลง หรือพัฒนาให้เป็นระบบตรวจจับตัวอักษรสำหรับภาพถ่ายขนาดใหญ่ เป็นต้น



รูปที่ 5.4 แสดงภาพที่มีขอบเขตป้ายชัดเจน

เอกสารอ้างอิง

- [1] Xilin Chen, Jie Yang, Jing Zhang, and Alex Waibel, **Automatic Detection and Recognition of Signs from Natural Scenes**, IEEE Transactions on Image Processing, Vol. 13, No. 1, Jan 2004.
- [2] **Mixture Model** [Online]. Available: http://en.wikipedia.org/wiki/Mixture_model
- [3] G. McLachlan, D. Peel Finite Mixture Models, , Wiley, 2000.
- [4] Marin, J.M., Mengersen, K. and Robert, C.P. **Bayesian modelling and inference on mixtures of distributions**. Handbook of Statistics 25, D. Dey and C.R. Rao (eds). Elsevier-Sciences .
- [5] **Expectation-maximization algorithm** [Online]. Available: http://en.wikipedia.org/wiki/Expectation-maximization_algorithm
- [6] Arthur Dempster, Nan Laird, and Donald Rubin. **Maximum likelihood from incomplete data via the EM algorithm**. Journal of the Royal Statistical Society, Series B, 39(1):1–38, 1977.
- [7] J. Gllavata, R. Ewerth and B. Freisleben, **A Text Detection, Localization and Segmentation System for OCR in Images**, Proc. Of IEEE Sixth International Symposium on Multimedia Software Engineering (ISMSE'04), 2004.
- [8] M. Cai, J. Song, and M. R. Lyu, **A New Approach for Video Text Detection**, Proc. Of IEEE Int'l Conference on Image Processing, Rochester, New York, USA , 2002, pp. 117-120.
- [9] A. K. Jain, and B. Yu, **Automatic Text Location in Images and Video Frames**, Pattern Recognition 31(12), 1998, pp. 2055-2076.
- [10] L. Agnihotri, and N. Dimitrova, **Text Detection for Video Analysis**, Proc. of Int'l Conf. on Multimedia Computing and Systems, Florence, 1999, pp. 109-113.
- [11] J. Gllavata, R. Ewerth, and B. Freisleben, **A Robust Algorithm for Text Detection in Images**, Proc. of 3rd Int'l Symposium on Image and Signal Processing and Analysis, Rome, 2003, pp. 611-616.
- [12] V. Wu, R. Manmatha, and E.M. Riseman, **Textfinder: An Automatic System to Detect and Recognize Text in Images**, IEEE Transact. on Pattern Analysis and Machine Intelligence, Vol. 21, Issue 11, 1999, pp. 1224-1229.

- [13] R. Lienhart, and A. Wernicke, **Localizing and Segmenting Text in Images and Videos**, IEEE Transact. On Circuits and Systems for Video Technology, Vol. 12, Nr. 4, 2002, pp. 256-268.
- [14] Y. Hao, Z. Yi, H. Zengguang, and T. Min, **Automatic Text Detection In Video Frames Based on Bootstrap Artificial Neural Network and CED**, Journal of WSCG Vol.11, No.1, Plzen, Czech Republic, 2003, ISSN 1213-6972.
- [15] H. Li, D. Doermann, and O. Kia, **Automatic Text Detection and Tracking in Digital Videos**, IEEE Transact. on Image Processing, Vol. 9, Nr. 1, 2000, pp. 147-156.
- [16] ชามุขชัย คีอ่วม, การรู้จำอักษรคัดลายมือภาษาไทยของคอมพิวเตอร์โดยวิธีวิเคราะห์โครงสร้างแบบทันทีทันใด, วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2542.
- [17] ประสาร ตั้งศิษานนท์, การจดจำรูปแบบตัวอักษรคัดลายมือภาษาไทย โดยใช้วิธีแยกลักษณะเด่น, วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2529.
- [18] สุรพันธ์ เอื้อไพบูลย์, การจดจำตัวอักษรลายมือเขียนภาษาไทยโดยการพิจารณาหัวตัวอักษร, วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2531.
- [19] สุรสิทธิ์ ราศรี, การรู้จำอักษรพิมพ์ภาษาไทยโดยวิธีการค้นหาลักษณะโครงสร้างของลายเส้น, วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2532.
- [20] กฤษณา วิไลลักษณ์, วิเชียร เปรมชัยสวัสดิ์ และนุชรี เปรมชัยสวัสดิ์, การรู้จำอักษรไทยโดยใช้ลักษณะเด่นของตัวอักษร, การประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 21, กรุงเทพมหานครฯ, 2541, หน้า 90-93.
- [21] Otsu, N., **A Threshold Selection Method from Gray-Level Histograms**, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No. 1, 1979, pp. 62-66.
- [22] **ArnThai** [Online]. Available: http://arnthai.nectec.or.th/arnthai_thai.htm
- [23] **Basic Principles of Perspective Drawing for the Technical Illustrator** [Online]. Available: http://www.khulsey.com/perspective_basics.html
- [24] Geometric Distortions of the Image [Online]. Available: <http://www.profc.udec.cl/~gabriel/tutoriales/rsnote/cp9/cp9-3.htm>
- [25] R. Gonzalez and R. Woods, **Digital Image Processing**, Addison-Wesley Publishing Company, 2003.

ภาคผนวก

ภาคผนวก ก
รูปที่ใช้ในการทดลอง

บัณฑิตปรุพิวิทยา

สุสานแม่ใจ
เทศบาลตำบลแม่ใจ

คูไฟฟ้า

บ้านแม่ใจ ขอย 1
เทศบาลตำบลแม่ใจ

แม่ใจการเกษตร

หยุด

อ.สามพราน

ตัดผมชาย

ปลอดภัย

สถานีวิทยุ

สวนหลวง ร.๙

ทต. 0+603
ลาดยาว จตุจักร
กรุงเทพมหานคร

ห้ามเปิด

ห้องปฏิบัติการ
คอมพิวเตอร์

ห้องบรรยาย 3

ฉลากแผ่นซีดี

ทางออก

โทรศัพท์สาธารณะ
PUBLIC TELEPHONE

โทรศัพท์สาธารณะ
PUBLIC TELEPHONE

บริการอาหาร-เครื่องดื่ม

อนุบาลเด่นหล้า

สถานที่ปลอดบุหรี่

ห้ามเข้าก่อนได้รับอนุญาต

โรงเรียนศึกษาพิเศษ

เภสัชวิทยาและพิษวิทยา
Pharmacology and Toxicology

เขตปลอดบุหรี่

อบรมฟรี

วิทยาลัยราชภัฏ 5.5
วิศกรรมย์กาวาสารวิจิตร

อาทิตย์

← ทางเข้า
คณะวิทยาศาสตร์

เหรียญทอง
สร้างพิเศษ 40°

ห้ามเปิด

เวลาให้บริการ
08.00น-21.00น

ห้ามเปิด

ห้ามเปิด

ข้างหน้า
ทางเอก

ทางโค้ง

เสาร์ - อาทิตย์

วงจรถ่ายรูป

ภาพบางภาพไปจำหน่ายต้องใช้เวลา

Text - มานพ ทัศนธร
Editor - Ananlaya

บริการซ่อมรถ
Tel: 0-2217-5655

ไมพริ่งไบโ คอนวูด
ติดตั้งง่าย

ที่ตั้งแพทย์แผนก +

โรงเรียน

ปลอดภัยไว้ก่อน

www.lbtop.com
ขายด่วน !
Tel: 02-2374568-70

ตักแตนทอด

สำนักเลขาธิการนายกรัฐมนตรี

ห้องอดรถ

ผู้ใช้บริการสนามเทนนิส
กรุณาปฏิบัติตามกฎระเบียบ
อย่างเคร่งครัด

แบนเนอร์
Quality of Life

↑↓
รถวิ่ง
สวนทาง

รถวิ่ง
สวนทาง

รถวิ่ง
สวนทาง

รถเดิน
ทางเดียว

รถเดิน
ทางเดียว

ระวังรถ
ทางขวา

ระวังรถ
ทางขวา

ระวังรถ
ทางขวา

ระวังรถ
ทางขวา

ขบได้เร็วที่สุด
กับการขบโดย
© 0 2273 5001

รามอินทรา
Ram Inthra
บางนา
Bang Na

ระวัง
ตรวจจับ
ความเร็ว

แจ้งเหตุเบี่ยงตัว
โทร. 1543

หยุดให้ทาง/เลนผ่าน
←
เมื่อช่วยผ่านตลอด

ห้ามแซง

ห้ามแซง

เริ่มต้นจุดพ่อน้ำมัน

คืนเดือน
ห้ามหยุด
กีดขวาง
ทางเข้า-ออก

↑
ขึ้นครีวินทร์

↑
ขึ้นครีวินทร์

ยินดีต้อนรับ
02-322-1082-3

↑
สุวรรณภูมิ
←

สุวรรณภูมิ
ชิดซ้าย

ช่องซ้าย
ผ่านตลอด

ช่องซ้าย
ผ่านตลอด

←
ถนนถึงแล้ว

←
ห้ามจอด

←
ทางออก

←
ทางออก

เขตห้ามหยุด

เขตห้ามหยุด

เขตห้ามหยุด

เขตห้ามหยุด

ห้ามจอดตลอดแนว
ฝาผนังล้อคล้อ

ห้ามจอดตลอดแนว
ฝาผนังล้อคล้อ

ห้ามเข้าไปในตึกคณะ

สถาปัตยกรรม

←
ห้ามจอดตลอดแนว

สำหรับสนนโดย

โรงอาหาร 3
คณะกรรมการศาสตร์

←
ห้ามจอดตลอดแนว

บุคคลภายนอก
ห้ามจอด

ห้ามจอด
ตลอดแนว

วัดคุมสัมพันธ

ห้ามนำสัตว์เลี้ยงเข้ามาในสนามเทนนิส

สวนสัตว์

สนามนี้ใช้เฉพาะนักศึกษา

สนามนี้ใช้เฉพาะนักศึกษา

สนามนี้ใช้เฉพาะนักศึกษา

สนามนี้ใช้เฉพาะนักศึกษา

สนามเทนนิส 4

ห้ามรถทุกชนิดเข้าบริเวณนี้

อาคารยิมเนยเจียม 2

โรงอาหาร

ห้องพักอาจารย์

ห้องพักอาจารย์

ห้องพักอาจารย์

อาคารบรรยาย

← โรงอาหาร

← โรงอาหาร

ห้องบรรยาย

เขตปลอดบุหรี่

เขตปลอดบุหรี่

สอนขับรถยนต์

หยุดในทางก่อนผ่าน
เลี้ยวซ้ายผ่านตลอด

ระวังรถจักรยานยนต์

รถทางขวา
ไปก่อน

ทึกลับรถ
100 ม.

หม้อน้ำ

BMU
เขียนตาไฟ 25.
30.

โค้งอันตราย
โปรดลดความเร็ว

ทางแยก
ข้างหน้า

นครราชสีมา
บุรีรัมย์

นครราชสีมา

คุ้มบลิสง
9 9-๗

เจาะบ่อบาดาล
02-316-3155
081-940-8988

โรงเรียนวัดประชาบำรุง
ตำบลหนองนก อำเภอหนองบัวระเหว จังหวัดขอนแก่น

โรงเรียน

สิ่งประดิษฐ์

ผลิตจากไม้
ต่างประเทศ

บล็อกแก้ว
47

สมาร์ทบอร์ด
140

รับของเอง

ดีศรีม

ดีเทา

ดีสี

ดีเทา

หน้าต่าง

วงกบ

วงกบ

สามทาง 4 หุน

สามทาง 4 หุน

ตรงเกลียวใน 4 หุน

ตรงเกลียวนอก 4 หุน

กลอนประตู

มือจับประตู

หยุดอันอาทิตย์

เนื้อหนียว สวยนาน

สวยเหมือนไม้จริง

เบื้อเคมีชว ทน...นาน

ของจังหวัด

ของจังหวัด

ชนิดซ้าย

รถจักรยาน
จักรยานยนต์

ลดความเร็ว
เข้าเขตโรงเรียน

ลดความเร็ว
เข้าเขตโรงเรียน

ถูกจับแน่

จำคุกไม่เกิน
1 ปี 2,000 - 10

สัญญาณไฟ
ข้างหน้า

ห้ามจอดวนตี
17.00 น. - 22.00 น.

ห้ามจอดวนตี
17.00 น. - 22.00 น.

ห้ามจอดวนตี
17.00 น. - 22.00 น.

ห้ามจอดวนตี
17.00 น. - 22.00 น.

รถวิ่ง
สวนทาง

รถวิ่ง
สวนทาง

ถนนก่อสร้างสะพาน
ข้างหน้า

ถนนก่อสร้างสะพาน
ข้างหน้า

ถนนก่อสร้างสะพาน
300 ม.

เครื่องจักร
กำลังทำงาน

ลดความเร็ว

หากพบเห็น
ถนนชำรุด

เมืองไม้ดอกไม้ประดับ
เฉลิมพระเกียรติฯ
ระยะทาง 3 กม. →

สายด่วน
1784

แจ้งสาธารณภัย

ลดความเร็ว

ที่จอดรถ

FOOD COURT

← ↗

ซุ้มของรางวัล

องค์พระ

TEPPANYAKI A LA CARTE

กล้วยน้ำไทเสวยอบ

หอพักศาสตร์

ทางขึ้น

ทางเข้า

ห้ามเข้าก่อนได้รับอนุญาต

สถานที่นี้ปลอดบุหรี่

รถจักรยาน และ แหนบตาเวียน

กรุงเทพมหานคร

สโมสรนักศึกษาหอพัก

สำหรับเด็ก
อายุ 4-6 ปี

สนามเด็กเล่น

8
ใต้ราว-ยึดตัว

กทท.
ขยะเปียก

สะพานข้ามลำน้ำ

หอรัชมงคล

สวนรมณี

11
เรือนเพาะชำ

ภาคผนวก ข

อัตราส่วนต่างๆ ที่ใช้ในการพิจารณารูปแบบของฟอนต์ภาษาไทย

พอนต์	Angsana New				AngsanaUPC				Browallia New				BrowalliaUPC			
ชนิด	ตัวปกติ (plain)		ตัวหนา (bold)		ตัวปกติ (plain)		ตัวหนา (bold)		ตัวปกติ (plain)		ตัวหนา (bold)		ตัวปกติ (plain)		ตัวหนา (bold)	
ตัวอักษร	อัตรส่วน		อัตรส่วน		อัตรส่วน		อัตรส่วน		อัตรส่วน		อัตรส่วน		อัตรส่วน		อัตรส่วน	
	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ
ก	0.763	0.341	0.775	0.478	0.763	0.341	0.775	0.478	0.825	0.305	0.875	0.458	0.825	0.305	0.875	0.458
ข	0.763	0.372	0.863	0.486	0.763	0.372	0.863	0.486	0.838	0.394	0.913	0.523	0.838	0.394	0.913	0.523
ฃ	0.913	0.364	0.938	0.482	0.913	0.364	0.938	0.482	0.888	0.386	0.963	0.531	0.888	0.386	0.963	0.531
ค	0.763	0.365	0.825	0.504	0.763	0.365	0.825	0.504	0.788	0.376	0.913	0.524	0.788	0.376	0.913	0.524
ค	0.785	0.390	0.863	0.477	0.785	0.390	0.863	0.477	0.775	0.396	0.913	0.528	0.775	0.396	0.913	0.528
ฅ	0.975	0.370	1.049	0.477	0.975	0.370	1.049	0.477	1.038	0.386	1.100	0.517	1.038	0.386	1.100	0.517
ง	0.563	0.353	0.625	0.414	0.563	0.353	0.625	0.414	0.588	0.338	0.663	0.458	0.588	0.338	0.663	0.458
จ	0.638	0.408	0.775	0.480	0.638	0.408	0.775	0.480	0.713	0.367	0.838	0.451	0.713	0.367	0.838	0.451
ฉ	0.825	0.362	0.888	0.432	0.825	0.362	0.888	0.432	0.950	0.347	1.113	0.435	0.950	0.347	1.113	0.435
ช	0.888	0.347	0.825	0.485	0.888	0.347	0.825	0.485	0.871	0.350	0.897	0.455	0.871	0.350	0.897	0.455
ช	0.963	0.354	1.000	0.484	0.963	0.354	1.000	0.484	0.952	0.349	0.966	0.445	0.952	0.349	0.966	0.445
ง	1.038	0.390	1.125	0.505	1.038	0.390	1.125	0.505	1.113	0.413	1.225	0.519	1.113	0.413	1.225	0.519
ญ	1.075	0.359	1.175	0.498	1.075	0.359	1.175	0.498	1.150	0.373	1.238	0.504	1.150	0.373	1.238	0.504
ฎ	0.550	0.336	0.632	0.437	0.550	0.336	0.632	0.437	0.574	0.342	0.598	0.462	0.574	0.342	0.598	0.462
ฎ	0.550	0.342	0.627	0.438	0.550	0.342	0.627	0.438	0.566	0.350	0.620	0.470	0.566	0.350	0.620	0.470
ฐ	1.733	0.379	0.763	0.495	1.733	0.379	0.763	0.495	1.647	0.446	1.895	0.548	1.647	0.446	1.895	0.548
ฑ	1.075	0.351	1.138	0.451	1.075	0.351	1.138	0.451	1.138	0.347	1.213	0.466	1.138	0.347	1.213	0.466
ฒ	1.063	0.413	1.225	0.522	1.063	0.413	1.225	0.522	1.150	0.420	1.284	0.549	1.150	0.420	1.284	0.549
ณ	1.100	0.386	1.198	0.479	1.100	0.386	1.198	0.479	1.350	0.336	1.475	0.471	1.350	0.336	1.475	0.471
ด	0.763	0.390	0.850	0.485	0.763	0.390	0.850	0.485	0.800	0.388	0.938	0.543	0.800	0.388	0.938	0.543
ด	0.772	0.388	0.863	0.514	0.772	0.388	0.863	0.514	0.800	0.393	0.938	0.542	0.800	0.393	0.938	0.542
ถ	0.763	0.370	0.800	0.489	0.763	0.370	0.800	0.489	0.813	0.346	0.875	0.481	0.813	0.346	0.875	0.481
ท	0.888	0.369	0.938	0.488	0.888	0.369	0.938	0.488	0.925	0.338	1.000	0.467	0.925	0.338	1.000	0.467
น	0.725	0.424	0.763	0.547	0.725	0.424	0.763	0.547	0.848	0.387	0.875	0.486	0.848	0.387	0.875	0.486
ธ	0.925	0.326	0.963	0.399	0.925	0.326	0.963	0.399	1.000	0.320	1.138	0.453	1.000	0.320	1.138	0.453
บ	0.900	0.358	0.925	0.462	0.900	0.358	0.925	0.462	0.900	0.339	0.975	0.469	0.900	0.339	0.975	0.469
ป	0.563	0.278	0.557	0.357	0.563	0.278	0.557	0.357	0.605	0.268	0.655	0.378	0.605	0.268	0.655	0.378
ผ	0.738	0.422	0.850	0.529	0.738	0.422	0.850	0.529	0.763	0.404	0.825	0.592	0.763	0.404	0.825	0.592
ฝ	0.454	0.321	0.531	0.405	0.454	0.321	0.531	0.405	0.513	0.314	0.555	0.473	0.513	0.314	0.555	0.473
พ	0.925	0.391	1.050	0.472	0.925	0.391	1.050	0.472	1.000	0.386	1.050	0.546	1.000	0.386	1.050	0.546
ฟ	0.591	0.304	0.648	0.365	0.591	0.304	0.648	0.365	0.672	0.296	0.667	0.414	0.672	0.296	0.667	0.414
ภ	0.813	0.348	0.888	0.425	0.813	0.348	0.888	0.425	0.913	0.309	0.913	0.453	0.913	0.309	0.913	0.453
ม	0.813	0.353	0.900	0.490	0.813	0.353	0.900	0.490	0.813	0.411	0.963	0.546	0.813	0.411	0.963	0.546
ย	0.638	0.408	0.750	0.566	0.638	0.408	0.750	0.566	0.700	0.423	0.800	0.570	0.700	0.423	0.800	0.570
ร	0.613	0.296	0.700	0.363	0.613	0.296	0.700	0.363	0.738	0.272	0.800	0.361	0.738	0.272	0.800	0.361
ฤ	0.550	0.316	0.512	0.419	0.550	0.316	0.512	0.419	0.596	0.286	0.642	0.411	0.596	0.286	0.642	0.411
ล	0.750	0.386	0.763	0.519	0.750	0.386	0.763	0.519	0.713	0.431	0.850	0.537	0.713	0.431	0.850	0.537
ฬ	0.586	0.301	0.603	0.355	0.586	0.301	0.603	0.355	0.679	0.261	0.664	0.385	0.679	0.261	0.664	0.385

พอนด์	Angsana New				AngsanaUPC				Browallia New				BrowalliaUPC			
ชนิด	ตีวปกติ (plain)		ตีวหนา (bold)		ตีวปกติ (plain)		ตีวหนา (bold)		ตีวปกติ (plain)		ตีวหนา (bold)		ตีวปกติ (plain)		ตีวหนา (bold)	
ตีวอักษร	อัตราส่วน	กว้าง / สูง	พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ		กว้าง / สูง	พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ		กว้าง / สูง	พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ		กว้าง / สูง	พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ		กว้าง / สูง	พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ	
			พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ	พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ		พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ	พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ		พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ	พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ		พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ	พื้นที่ตีวอักษร / พื้นที่กรอบล้อมรอบ			
ว	0.675	0.258	0.725	0.336	0.675	0.258	0.725	0.336	0.700	0.261	0.813	0.335	0.700	0.261	0.813	0.335
ศ	0.814	0.368	0.863	0.499	0.814	0.368	0.863	0.499	0.831	0.372	0.943	0.468	0.831	0.372	0.943	0.468
ษ	0.863	0.374	0.925	0.505	0.863	0.374	0.925	0.505	1.125	0.353	1.213	0.468	1.125	0.353	1.213	0.468
ส	0.872	0.324	0.938	0.454	0.872	0.324	0.938	0.454	0.838	0.399	0.841	0.493	0.838	0.399	0.841	0.493
ห	0.949	0.315	0.975	0.410	0.949	0.315	0.975	0.410	0.925	0.355	0.988	0.495	0.925	0.355	0.988	0.495
พ	1.012	0.321	1.225	0.439	1.012	0.321	1.225	0.439	1.163	0.357	1.055	0.463	1.163	0.357	1.055	0.463
อ	0.759	0.362	0.772	0.484	0.759	0.362	0.772	0.484	0.700	0.421	0.850	0.511	0.700	0.421	0.850	0.511
ช	0.688	0.405	0.772	0.508	0.688	0.405	0.772	0.508	0.850	0.406	0.833	0.518	0.850	0.406	0.833	0.518
ฉ	1.793	0.386	1.941	0.403	1.793	0.386	1.941	0.403	2.103	0.415	1.970	0.429	2.103	0.415	1.970	0.429
อา	0.506	0.305	0.519	0.394	0.506	0.305	0.519	0.394	0.675	0.228	0.738	0.330	0.675	0.228	0.738	0.330
อ	1.906	0.331	1.806	0.414	1.906	0.331	1.806	0.414	2.069	0.452	1.833	0.426	2.069	0.452	1.833	0.426
อ	1.632	0.336	1.512	0.429	1.632	0.336	1.512	0.429	1.538	0.417	1.610	0.475	1.538	0.417	1.610	0.475
อ	1.500	0.319	1.444	0.389	1.500	0.319	1.444	0.389	1.419	0.404	1.533	0.409	1.419	0.404	1.533	0.409
อ	1.386	0.382	1.354	0.446	1.386	0.382	1.354	0.446	1.538	0.445	1.610	0.504	1.538	0.445	1.610	0.504
อุ	0.813	0.507	0.944	0.558	0.813	0.507	0.944	0.558	0.605	0.488	0.738	0.562	0.605	0.488	0.738	0.562
อ	1.235	0.458	1.270	0.572	1.235	0.458	1.270	0.572	1.116	0.501	1.385	0.617	1.116	0.501	1.385	0.617
เอ	0.288	0.399	0.313	0.526	0.288	0.399	0.313	0.526	0.325	0.405	0.388	0.546	0.325	0.405	0.388	0.546
ไอ	0.405	0.217	0.404	0.294	0.405	0.217	0.404	0.294	0.388	0.227	0.490	0.277	0.388	0.227	0.490	0.277
ไอ	0.366	0.250	0.370	0.348	0.366	0.250	0.370	0.348	0.367	0.237	0.438	0.321	0.367	0.237	0.438	0.321
ไอ	0.391	0.232	0.449	0.323	0.391	0.232	0.449	0.323	0.432	0.232	0.431	0.316	0.432	0.232	0.431	0.316
อ	1.000	0.358	1.200	0.434	1.000	0.358	1.200	0.434	1.459	0.332	1.179	0.437	1.459	0.332	1.179	0.437
อ	0.387	0.871	0.487	0.690	0.387	0.871	0.487	0.690	0.360	1.000	0.441	1.000	0.360	1.000	0.441	1.000
อ	0.981	0.338	1.149	0.447	0.981	0.338	1.149	0.447	1.529	0.373	1.442	0.455	1.529	0.373	1.442	0.455
อ	1.313	0.332	1.463	0.411	1.313	0.332	1.463	0.411	1.897	0.379	1.875	0.462	1.897	0.379	1.875	0.462
อ	1.000	0.411	1.075	0.424	1.000	0.411	1.075	0.424	0.969	0.444	1.152	0.470	0.969	0.444	1.152	0.470
อ	1.052	0.410	1.069	0.342	1.052	0.410	1.069	0.342	1.107	0.357	1.224	0.475	1.107	0.357	1.224	0.475
๑	1.000	0.435	1.000	0.473	1.000	0.435	1.000	0.473	1.000	0.412	1.143	0.482	1.000	0.412	1.143	0.482
๒	1.033	0.333	0.989	0.382	1.033	0.333	0.989	0.382	1.000	0.349	1.061	0.426	1.000	0.349	1.061	0.426
๓	1.448	0.418	1.466	0.444	1.448	0.418	1.466	0.444	1.339	0.443	1.298	0.589	1.339	0.443	1.298	0.589
๔	0.932	0.266	0.956	0.312	0.932	0.266	0.956	0.312	1.011	0.261	1.038	0.357	1.011	0.261	1.038	0.357
๕	0.989	0.321	0.978	0.344	0.989	0.321	0.978	0.344	0.891	0.325	1.062	0.382	0.891	0.325	1.062	0.382
๖	0.714	0.257	0.711	0.298	0.714	0.257	0.711	0.298	0.759	0.278	1.012	0.309	0.759	0.278	1.012	0.309
๗	0.978	0.283	0.989	0.360	0.978	0.283	0.989	0.360	1.148	0.279	1.181	0.351	1.148	0.279	1.181	0.351
๘	0.921	0.279	0.935	0.339	0.921	0.279	0.935	0.339	0.951	0.305	1.000	0.356	0.951	0.305	1.000	0.356
๙	0.685	0.302	0.700	0.398	0.685	0.302	0.700	0.398	0.843	0.341	1.095	0.326	0.843	0.341	1.095	0.326
ค่าเฉลี่ย (ก-ฮ)	0.804	0.358	0.868	0.464	0.804	0.358	0.868	0.464	0.861	0.358	0.936	0.481	0.861	0.358	0.936	0.481
ค่าสูงสุด (ก-ฮ)	1.100	0.424	1.225	0.566	1.100	0.424	1.225	0.566	1.350	0.431	1.475	0.592	1.350	0.431	1.475	0.592
ค่าต่ำสุด (ก-ฮ)	0.454	0.258	0.512	0.336	0.454	0.258	0.512	0.336	0.513	0.261	0.555	0.335	0.513	0.261	0.555	0.335

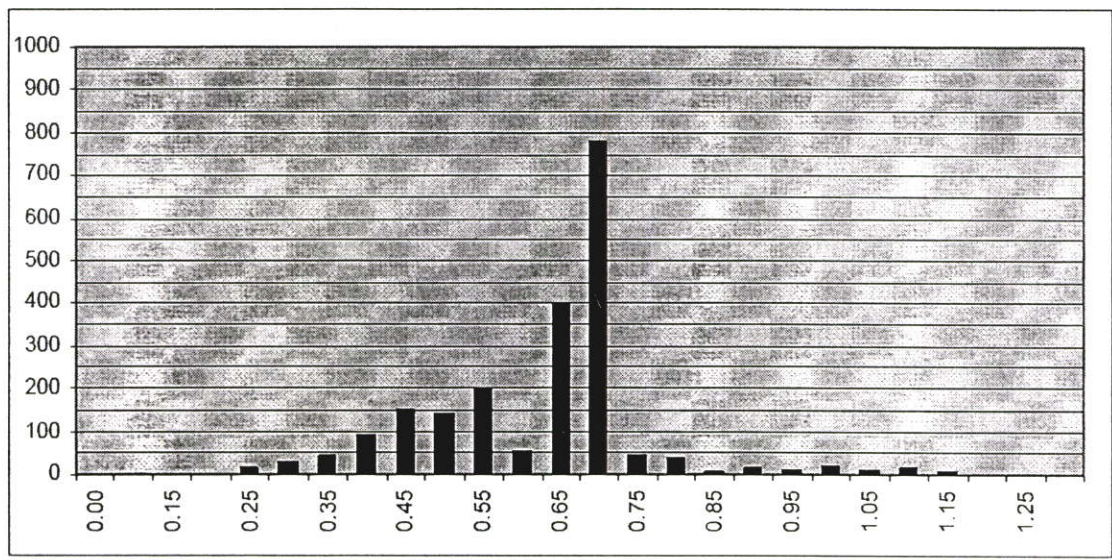
พอนด์	Cordia New				CordiaUPC				DilleniaUPC				FreesiaUPC			
ชนิด	ตัวปกติ (plain)		ตัวหนา (bold)		ตัวปกติ (plain)		ตัวหนา (bold)		ตัวปกติ (plain)		ตัวหนา (bold)		ตัวปกติ (plain)		ตัวหนา (bold)	
ตัวอักษร	อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน	
	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ
ก	0.753	0.254	0.850	0.411	0.753	0.254	0.850	0.411	0.750	0.216	0.738	0.448	0.713	0.337	0.813	0.527
ข	0.728	0.359	0.877	0.514	0.728	0.359	0.877	0.514	0.696	0.302	0.763	0.554	0.675	0.455	0.888	0.588
ฃ	0.788	0.372	0.900	0.524	0.788	0.372	0.900	0.524	0.747	0.298	0.848	0.535	0.725	0.449	0.888	0.569
ค	0.728	0.350	0.813	0.549	0.728	0.350	0.813	0.549	0.775	0.272	0.788	0.520	0.750	0.460	0.888	0.611
ค	0.750	0.387	0.863	0.556	0.750	0.387	0.863	0.556	0.785	0.288	0.810	0.536	0.750	0.467	0.888	0.623
ฅ	0.963	0.325	1.062	0.515	0.963	0.325	1.062	0.515	0.911	0.302	0.925	0.543	0.938	0.418	1.025	0.596
ง	0.593	0.288	0.675	0.418	0.593	0.288	0.675	0.418	0.550	0.244	0.575	0.476	0.538	0.376	0.663	0.515
จ	0.738	0.285	0.838	0.443	0.738	0.285	0.838	0.443	0.688	0.256	0.675	0.494	0.688	0.392	0.775	0.544
ฉ	0.877	0.308	1.037	0.424	0.877	0.308	1.037	0.424	0.925	0.252	1.000	0.464	0.877	0.366	0.963	0.505
ช	0.753	0.324	0.835	0.447	0.753	0.324	0.835	0.447	0.744	0.260	0.767	0.505	0.689	0.386	0.895	0.520
ช	0.741	0.339	0.857	0.445	0.741	0.339	0.857	0.445	0.791	0.254	0.837	0.485	0.722	0.374	0.895	0.511
ง	1.207	0.296	1.296	0.480	1.207	0.296	1.296	0.480	1.125	0.282	1.113	0.559	1.049	0.416	1.198	0.572
จ	1.256	0.270	1.325	0.461	1.256	0.270	1.325	0.461	1.125	0.250	1.113	0.515	1.088	0.373	1.225	0.548
ฎ	0.566	0.261	0.648	0.425	0.566	0.261	0.648	0.425	0.598	0.254	0.576	0.500	0.640	0.342	0.702	0.504
ฎ	0.566	0.265	0.680	0.422	0.566	0.265	0.680	0.422	0.598	0.255	0.576	0.496	0.640	0.348	0.684	0.513
ฐ	1.500	0.396	1.644	0.530	1.500	0.396	1.644	0.530	1.273	0.345	1.244	0.547	1.889	0.461	2.162	0.560
ฑ	1.050	0.315	1.163	0.466	1.050	0.315	1.163	0.466	1.013	0.259	0.975	0.462	0.975	0.384	1.125	0.526
ฒ	1.222	0.360	1.244	0.554	1.222	0.360	1.244	0.554	1.139	0.309	1.215	0.572	1.150	0.453	1.293	0.599
ณ	1.354	0.263	1.383	0.451	1.354	0.263	1.383	0.451	1.250	0.243	1.363	0.474	1.163	0.386	1.272	0.538
ด	0.728	0.366	0.863	0.558	0.728	0.366	0.863	0.558	0.775	0.284	0.788	0.542	0.750	0.459	0.888	0.629
ด	0.738	0.365	0.863	0.569	0.738	0.365	0.863	0.569	0.785	0.294	0.810	0.553	0.750	0.460	0.888	0.639
ถ	0.744	0.291	0.850	0.466	0.744	0.291	0.850	0.466	0.750	0.258	0.738	0.534	0.713	0.384	0.813	0.567
ท	0.864	0.311	1.000	0.478	0.864	0.311	1.000	0.478	0.825	0.248	0.850	0.473	0.900	0.343	1.013	0.526
น	0.704	0.335	0.850	0.514	0.704	0.335	0.850	0.514	0.747	0.274	0.684	0.555	0.725	0.388	0.827	0.548
ธ	0.963	0.255	1.025	0.423	0.963	0.255	1.025	0.423	0.963	0.230	1.050	0.432	1.037	0.337	1.025	0.479
บ	0.901	0.269	1.000	0.439	0.901	0.269	1.000	0.439	0.825	0.237	0.838	0.478	0.900	0.316	1.013	0.467
ป	0.598	0.207	0.645	0.341	0.598	0.207	0.645	0.341	0.579	0.190	0.583	0.384	0.621	0.240	0.698	0.384
ผ	0.704	0.373	0.800	0.556	0.704	0.373	0.800	0.556	0.738	0.306	0.788	0.566	0.688	0.450	0.813	0.647
ฝ	0.467	0.290	0.516	0.426	0.467	0.290	0.516	0.426	0.518	0.241	0.548	0.448	0.474	0.356	0.560	0.518
พ	0.951	0.329	1.088	0.526	0.951	0.329	1.088	0.526	0.925	0.275	1.050	0.504	0.963	0.432	1.100	0.594
ฟ	0.631	0.250	0.702	0.393	0.631	0.250	0.702	0.393	0.649	0.214	0.730	0.392	0.664	0.334	0.759	0.463
ภ	0.890	0.249	0.975	0.397	0.890	0.249	0.975	0.397	0.838	0.227	0.850	0.462	0.925	0.305	0.988	0.460
ม	0.829	0.318	0.914	0.505	0.829	0.318	0.914	0.505	0.800	0.284	0.813	0.556	0.825	0.419	0.938	0.588
ย	0.753	0.363	0.838	0.565	0.753	0.363	0.838	0.565	0.675	0.310	0.713	0.590	0.700	0.436	0.825	0.606
ร	0.598	0.282	0.768	0.396	0.598	0.282	0.768	0.396	0.709	0.220	0.696	0.456	0.675	0.310	0.790	0.418
ฤ	0.473	0.229	0.630	0.391	0.473	0.229	0.630	0.391	0.522	0.205	0.513	0.428	0.509	0.318	0.571	0.477
ล	0.744	0.356	0.889	0.537	0.744	0.356	0.889	0.537	0.725	0.296	0.725	0.560	0.713	0.444	0.788	0.623
ฬ	0.566	0.194	0.731	0.337	0.566	0.194	0.731	0.337	0.583	0.180	0.591	0.370	0.652	0.248	0.705	0.391

พอนด์	Cordia New				CordiaUPC				DilleniaUPC				FreesiaUPC			
ชนิด	ตัวปกติ (plain)		ตัวหนา (bold)		ตัวปกติ (plain)		ตัวหนา (bold)		ตัวปกติ (plain)		ตัวหนา (bold)		ตัวปกติ (plain)		ตัวหนา (bold)	
ตัวอักษร	อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน	
	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ
ว	0.622	0.258	0.728	0.379	0.622	0.258	0.728	0.379	0.638	0.203	0.638	0.406	0.638	0.295	0.728	0.413
ศ	0.656	0.345	0.809	0.506	0.656	0.345	0.809	0.506	0.767	0.237	0.778	0.458	0.730	0.407	0.907	0.553
ษ	1.123	0.285	1.150	0.466	1.123	0.285	1.150	0.466	1.025	0.255	1.075	0.461	1.100	0.336	1.213	0.474
ศ	0.670	0.343	0.811	0.525	0.670	0.343	0.811	0.525	0.722	0.254	0.722	0.487	0.700	0.375	0.828	0.537
ห	0.901	0.300	1.025	0.459	0.901	0.300	1.025	0.459	0.825	0.272	0.913	0.497	0.925	0.347	1.038	0.508
พ	0.813	0.269	1.053	0.412	0.813	0.269	1.053	0.412	0.730	0.229	0.802	0.401	0.926	0.348	1.054	0.497
อ	0.753	0.362	0.938	0.497	0.753	0.362	0.938	0.497	0.688	0.296	0.688	0.584	0.713	0.421	0.800	0.596
ช	0.690	0.381	0.837	0.525	0.690	0.381	0.837	0.525	0.659	0.328	0.706	0.583	0.682	0.433	0.814	0.580
ฉ	2.071	0.352	2.250	0.455	2.071	0.352	2.250	0.455	1.793	0.346	1.636	0.542	2.333	0.437	2.226	0.540
อ	0.630	0.185	0.725	0.311	0.630	0.185	0.725	0.311	0.475	0.176	0.550	0.337	0.625	0.235	0.700	0.374
อ	2.172	0.389	2.057	0.490	2.172	0.389	2.057	0.490	1.964	0.384	1.667	0.612	2.000	0.537	2.172	0.692
อ	1.500	0.333	1.946	0.565	1.500	0.333	1.946	0.565	1.410	0.322	1.310	0.576	1.500	0.494	1.676	0.648
อ	1.465	0.342	2.139	0.515	1.465	0.342	2.139	0.515	1.250	0.326	1.191	0.546	1.639	0.476	1.692	0.593
อ	1.500	0.368	1.973	0.591	1.500	0.368	1.973	0.591	1.410	0.345	1.310	0.611	1.500	0.527	1.703	0.664
อ	0.639	0.444	0.811	0.590	0.639	0.444	0.811	0.590	0.583	0.454	0.625	0.586	0.659	0.513	0.762	0.579
อ	1.111	0.474	1.459	0.608	1.111	0.474	1.459	0.608	1.027	0.442	0.976	0.609	1.500	0.579	1.450	0.619
อ	0.321	0.362	0.400	0.496	0.321	0.362	0.400	0.496	0.291	0.357	0.380	0.515	0.338	0.425	0.425	0.547
อ	0.364	0.214	0.439	0.317	0.364	0.214	0.439	0.317	0.358	0.163	0.353	0.327	0.386	0.232	0.490	0.291
อ	0.423	0.194	0.523	0.287	0.423	0.194	0.523	0.287	0.349	0.172	0.383	0.323	0.381	0.265	0.458	0.331
อ	0.365	0.216	0.445	0.307	0.365	0.216	0.445	0.307	0.369	0.161	0.396	0.305	0.386	0.244	0.451	0.317
อ	1.389	0.386	1.375	0.504	1.389	0.386	1.375	0.504	1.093	0.338	1.000	0.486	1.349	0.423	1.422	0.502
อ	0.259	1.000	0.429	1.000	0.259	1.000	0.429	1.000	0.207	1.000	0.273	1.000	0.265	1.000	0.441	1.000
อ	1.275	0.335	1.526	0.436	1.275	0.335	1.526	0.436	1.289	0.350	1.341	0.541	1.500	0.417	1.541	0.534
อ	1.844	0.425	2.030	0.469	1.844	0.425	2.030	0.469	1.513	0.336	1.419	0.511	1.359	0.398	1.452	0.493
อ	0.971	0.351	1.077	0.469	0.971	0.351	1.077	0.469	0.893	0.343	0.879	0.521	1.135	0.344	1.324	0.512
อ	1.155	0.276	1.161	0.466	1.155	0.276	1.161	0.466	1.052	0.239	0.984	0.436	1.100	0.348	1.224	0.478
๑	0.985	0.314	1.046	0.465	0.985	0.314	1.046	0.465	1.000	0.286	0.941	0.493	1.000	0.400	1.143	0.481
๒	0.988	0.294	0.875	0.434	0.988	0.294	0.875	0.434	0.963	0.257	0.894	0.461	0.874	0.339	1.085	0.420
๓	1.345	0.372	1.179	0.550	1.345	0.372	1.179	0.550	1.211	0.308	0.937	0.548	1.103	0.438	1.246	0.568
๔	1.313	0.273	0.964	0.350	1.313	0.273	0.964	0.350	1.027	0.241	0.949	0.414	1.012	0.264	1.038	0.360
๕	1.063	0.264	0.964	0.417	1.063	0.264	0.964	0.417	0.787	0.251	0.875	0.413	1.012	0.308	1.062	0.385
๖	0.840	0.238	0.897	0.350	0.840	0.238	0.897	0.350	0.808	0.212	0.750	0.366	0.919	0.259	1.012	0.311
๗	1.000	0.260	1.079	0.370	1.000	0.260	1.079	0.370	0.988	0.219	0.940	0.429	1.000	0.303	1.181	0.357
๘	1.273	0.243	0.943	0.351	1.273	0.243	0.943	0.351	1.055	0.240	0.987	0.406	0.952	0.307	1.000	0.357
๙	1.086	0.225	1.071	0.345	1.086	0.225	1.071	0.345	1.086	0.246	1.051	0.419	0.988	0.296	1.120	0.325
ค่าเฉลี่ย (ก-ฮ)	0.804	0.309	0.914	0.471	0.804	0.309	0.914	0.471	0.791	0.259	0.818	0.495	0.793	0.382	0.909	0.537
ค่าสูงสุด (ก-ฮ)	1.354	0.387	1.383	0.569	1.354	0.387	1.383	0.569	1.250	0.328	1.363	0.590	1.163	0.467	1.293	0.647
ค่าต่ำสุด (ก-ฮ)	0.467	0.194	0.516	0.337	0.467	0.194	0.516	0.337	0.518	0.180	0.513	0.370	0.474	0.240	0.560	0.384

พอนด์	IrisUPC				JasmineUPC				KodchiangUPC				LilyUPC			
ชนิด	คิ้วปกติ (plain)		คิ้วหนา (bold)		คิ้วปกติ (plain)		คิ้วหนา (bold)		คิ้วปกติ (plain)		คิ้วหนา (bold)		คิ้วปกติ (plain)		คิ้วหนา (bold)	
อัตราส่วน คิ้วอักษร	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ		กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ		กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ		กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ		กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ		
		พื้นที่ตัวอักษร	พื้นที่กรอบล้อมรอบ		พื้นที่ตัวอักษร	พื้นที่กรอบล้อมรอบ		พื้นที่ตัวอักษร	พื้นที่กรอบล้อมรอบ		พื้นที่ตัวอักษร	พื้นที่กรอบล้อมรอบ		พื้นที่ตัวอักษร	พื้นที่กรอบล้อมรอบ	
ก	0.725	0.370	0.850	0.504	0.925	0.364	0.988	0.517	0.778	0.341	0.840	0.488	0.775	0.546	0.900	0.657
ข	0.775	0.429	0.925	0.561	0.825	0.433	0.988	0.538	0.630	0.436	0.778	0.538	0.638	0.610	0.800	0.702
ฅ	0.813	0.408	0.963	0.556	0.838	0.429	1.075	0.524	0.667	0.395	0.790	0.532	0.763	0.515	0.963	0.591
ค	0.675	0.469	0.825	0.601	0.913	0.457	1.049	0.588	0.827	0.387	0.901	0.522	0.950	0.645	1.098	0.724
ค	0.688	0.494	0.825	0.614	0.914	0.460	1.063	0.622	0.838	0.383	0.914	0.513	0.950	0.652	1.123	0.724
ฅ	0.838	0.431	1.013	0.569	0.988	0.417	1.160	0.517	0.840	0.368	0.926	0.486	0.913	0.473	1.049	0.566
ง	0.488	0.410	0.613	0.493	0.763	0.366	0.850	0.494	0.543	0.335	0.605	0.443	0.738	0.481	0.888	0.550
จ	0.725	0.396	0.863	0.510	0.800	0.384	0.901	0.514	0.622	0.354	0.695	0.453	0.763	0.489	0.913	0.607
ฉ	0.900	0.354	1.138	0.443	0.825	0.418	0.963	0.547	0.732	0.347	0.805	0.475	0.825	0.540	1.000	0.637
ช	0.788	0.398	0.976	0.490	0.739	0.404	0.929	0.518	0.567	0.409	0.708	0.509	0.763	0.589	0.913	0.680
ช	0.825	0.400	1.012	0.483	0.773	0.394	1.024	0.497	0.600	0.382	0.719	0.499	0.900	0.502	1.063	0.590
ฅ	1.238	0.382	1.333	0.515	1.350	0.389	1.469	0.533	1.146	0.359	1.232	0.507	1.213	0.552	1.400	0.648
ญ	1.138	0.382	1.288	0.530	1.388	0.356	1.469	0.505	1.146	0.341	1.244	0.487	1.213	0.542	1.388	0.660
ฎ	0.508	0.375	0.592	0.505	0.705	0.367	0.719	0.498	0.619	0.313	0.676	0.464	0.680	0.500	0.761	0.614
ฎ	0.508	0.380	0.592	0.494	0.705	0.375	0.717	0.505	0.629	0.313	0.686	0.453	0.680	0.498	0.761	0.617
ฐ	1.513	0.421	2.088	0.538	0.825	0.449	0.925	0.555	2.875	0.454	3.313	0.577	0.705	0.491	0.830	0.617
ฑ	0.900	0.403	1.088	0.513	1.050	0.368	1.200	0.493	0.888	0.332	0.950	0.476	0.950	0.486	1.063	0.593
ฒ	1.075	0.517	1.235	0.618	1.400	0.449	1.575	0.604	1.235	0.386	1.341	0.509	1.383	0.585	1.654	0.657
ณ	1.188	0.373	1.321	0.482	1.463	0.356	1.568	0.501	1.146	0.357	1.256	0.499	1.213	0.547	1.400	0.644
ด	0.675	0.508	0.825	0.628	0.925	0.471	1.063	0.612	0.827	0.387	0.901	0.529	0.950	0.630	1.098	0.717
ด	0.750	0.490	0.825	0.636	0.913	0.477	1.063	0.625	0.825	0.402	0.914	0.520	0.950	0.642	1.111	0.722
ถ	0.775	0.373	0.850	0.530	0.875	0.381	0.951	0.529	0.753	0.359	0.815	0.516	0.788	0.586	0.900	0.688
ท	0.700	0.463	0.913	0.571	0.951	0.388	1.050	0.555	0.888	0.331	0.975	0.479	0.788	0.576	0.900	0.669
น	0.613	0.466	0.813	0.535	0.838	0.421	0.913	0.569	0.646	0.350	0.720	0.488	0.813	0.538	0.963	0.623
ธ	0.900	0.387	1.075	0.489	0.938	0.375	1.025	0.514	0.852	0.348	0.963	0.467	0.788	0.570	0.900	0.667
บ	0.775	0.395	0.900	0.521	0.950	0.353	1.038	0.495	0.877	0.317	0.963	0.440	0.788	0.537	0.900	0.638
ป	0.492	0.317	0.570	0.422	0.600	0.300	0.650	0.428	0.617	0.261	0.678	0.364	0.620	0.463	0.720	0.566
ผ	0.663	0.495	0.750	0.661	0.925	0.464	1.075	0.568	0.838	0.393	0.950	0.536	1.200	0.579	1.400	0.653
ฝ	0.411	0.386	0.477	0.527	0.583	0.381	0.695	0.476	0.588	0.325	0.667	0.443	0.943	0.455	1.114	0.510
พ	0.800	0.462	0.975	0.596	1.050	0.462	1.263	0.587	0.950	0.393	1.088	0.518	1.200	0.567	1.400	0.638
ฟ	0.508	0.340	0.609	0.454	0.672	0.376	0.822	0.475	0.667	0.320	0.763	0.422	0.943	0.446	1.114	0.498
ภ	0.738	0.392	0.888	0.493	0.975	0.342	1.062	0.476	0.864	0.310	0.938	0.443	0.875	0.498	1.038	0.611
ม	0.759	0.470	0.913	0.582	0.963	0.384	1.075	0.520	0.864	0.355	0.938	0.486	0.788	0.573	0.889	0.668
ย	0.700	0.458	0.775	0.635	0.838	0.437	0.900	0.591	0.728	0.381	0.802	0.520	0.825	0.627	0.926	0.745
ร	0.563	0.328	0.738	0.396	0.813	0.367	0.900	0.491	0.578	0.313	0.683	0.443	0.750	0.493	0.875	0.620
ฤ	0.517	0.308	0.630	0.443	0.648	0.327	0.700	0.462	0.560	0.297	0.606	0.445	0.630	0.526	0.720	0.617
ล	0.759	0.416	0.838	0.541	0.888	0.396	0.975	0.553	0.732	0.322	0.780	0.473	0.788	0.569	0.913	0.673
ฤ	0.492	0.323	0.657	0.427	0.722	0.293	0.782	0.406	0.651	0.262	0.697	0.391	0.710	0.462	0.830	0.549

พอนต์	IrisUPC				JasmineUPC				KodchiangUPC				LilyUPC			
ชนิด	ตัวปกติ (plain)		ตัวหนา (bold)		ตัวปกติ (plain)		ตัวหนา (bold)		ตัวปกติ (plain)		ตัวหนา (bold)		ตัวปกติ (plain)		ตัวหนา (bold)	
ตัวอักษร	อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน		อัตราส่วน	
	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ	กว้าง / สูง	พื้นที่ตัวอักษร / พื้นที่กรอบล้อมรอบ
ว	0.557	0.324	0.675	0.395	0.813	0.296	0.889	0.445	0.598	0.269	0.659	0.378	0.713	0.403	0.800	0.529
ศ	0.663	0.474	0.778	0.547	0.849	0.452	0.967	0.561	0.787	0.349	0.876	0.474	0.864	0.606	1.011	0.699
ษ	0.900	0.385	1.063	0.506	1.000	0.386	1.125	0.518	0.926	0.323	1.025	0.430	0.873	0.541	0.988	0.643
ส	0.873	0.400	0.871	0.523	0.826	0.396	0.878	0.529	0.700	0.303	0.780	0.426	0.716	0.567	0.839	0.668
ห	0.825	0.393	0.938	0.521	0.975	0.431	1.050	0.573	0.877	0.374	0.963	0.495	0.813	0.571	0.925	0.668
พ	0.796	0.385	1.069	0.472	0.922	0.413	1.104	0.481	0.844	0.352	0.956	0.479	0.943	0.444	1.105	0.503
อ	0.738	0.392	0.838	0.525	0.838	0.407	0.913	0.543	0.707	0.334	0.744	0.458	0.838	0.482	0.900	0.620
ช	0.700	0.463	0.779	0.542	0.791	0.399	0.856	0.506	0.678	0.315	0.736	0.416	0.761	0.494	0.818	0.624
ฉ	2.000	0.355	2.111	0.411	2.800	0.438	2.760	0.545	3.188	0.483	3.000	0.571	2.944	0.673	3.429	0.720
อา	0.575	0.276	0.675	0.360	0.713	0.276	0.877	0.370	0.568	0.258	0.642	0.372	0.625	0.375	0.750	0.482
อื	2.286	0.422	2.321	0.496	4.571	0.633	4.688	0.702	5.200	0.485	5.083	0.567	6.200	0.968	5.308	0.957
อึ	1.853	0.430	1.829	0.573	2.560	0.454	2.500	0.501	2.348	0.321	2.667	0.410	2.696	0.558	3.136	0.679
อฺ	1.658	0.383	1.914	0.488	2.625	0.501	2.621	0.525	2.261	0.299	2.625	0.350	2.818	0.567	2.958	0.616
อฌ	1.882	0.466	1.778	0.573	2.560	0.504	2.500	0.574	1.565	0.353	1.750	0.442	1.955	0.598	2.091	0.694
อุ	0.686	0.555	0.968	0.703	0.759	0.624	1.000	0.697	0.923	0.492	1.107	0.582	0.900	0.867	1.083	0.885
อู	1.286	0.545	1.813	0.682	1.741	0.573	2.036	0.653	1.808	0.466	2.107	0.579	2.048	0.641	2.231	0.717
เอ	0.300	0.542	0.400	0.618	0.300	0.545	0.413	0.617	0.363	0.412	0.450	0.484	0.213	1.000	0.300	1.000
โอ	0.421	0.271	0.515	0.311	0.323	0.340	0.547	0.349	0.361	0.278	0.400	0.383	0.449	0.326	0.507	0.400
ออ	0.349	0.289	0.447	0.366	0.339	0.334	0.461	0.412	0.336	0.305	0.372	0.402	0.433	0.353	0.475	0.411
ไอ	0.433	0.243	0.500	0.322	0.347	0.325	0.469	0.396	0.387	0.270	0.438	0.349	0.449	0.326	0.511	0.402
อเ	1.220	0.315	1.385	0.394	1.733	0.504	1.628	0.508	1.429	0.333	1.528	0.438	2.240	0.698	2.250	0.767
อึ	0.250	1.000	0.586	1.000	0.480	0.950	0.731	0.955	0.789	0.789	1.211	0.794	0.789	1.000	0.731	1.000
อช	1.256	0.307	1.390	0.370	2.316	0.543	2.500	0.533	2.556	0.413	2.737	0.518	2.353	0.679	2.304	0.718
อฌ	1.371	0.382	1.457	0.382	2.714	0.519	2.625	0.566	3.471	0.495	3.632	0.588	3.222	0.741	3.947	0.808
อ๓	1.000	0.356	1.467	0.626	1.440	0.478	1.586	0.558	1.684	0.493	1.947	0.602	1.714	0.593	1.808	0.633
อ๔	1.241	0.454	1.283	0.608	1.145	0.407	1.210	0.510	1.125	0.376	1.228	0.452	1.116	0.535	1.116	0.645
อ๕	0.857	0.423	0.986	0.539	0.944	0.403	1.041	0.507	0.970	0.323	1.121	0.406	0.917	0.509	0.976	0.568
อ๖	0.925	0.398	1.075	0.482	0.903	0.359	1.042	0.445	0.975	0.324	1.150	0.418	0.832	0.508	0.905	0.570
อ๗	1.111	0.470	1.259	0.599	1.000	0.423	1.100	0.534	1.241	0.481	1.400	0.562	1.143	0.611	1.257	0.681
อ๘	1.013	0.243	1.025	0.319	0.710	0.313	0.859	0.392	0.850	0.241	0.963	0.333	0.904	0.414	0.957	0.478
อ๙	1.013	0.288	1.025	0.372	0.710	0.361	0.859	0.443	0.863	0.300	1.264	0.437	0.905	0.449	0.947	0.547
อ๑	0.900	0.286	0.963	0.400	0.755	0.255	0.830	0.341	0.829	0.273	0.915	0.355	0.874	0.396	0.958	0.458
อ๒	1.078	0.308	1.253	0.389	1.000	0.313	1.105	0.403	1.075	0.308	1.313	0.385	0.990	0.462	1.063	0.533
อ๓	1.051	0.230	1.038	0.327	0.768	0.299	0.830	0.390	0.852	0.245	0.951	0.332	0.853	0.422	0.926	0.499
อ๔	1.025	0.295	1.138	0.337	0.811	0.304	0.854	0.372	1.049	0.277	1.263	0.345	0.947	0.391	1.042	0.441
ค่าเฉลี่ย (ก-ข)	0.754	0.409	0.890	0.526	0.908	0.396	1.024	0.526	0.782	0.348	0.867	0.475	0.874	0.539	1.016	0.634
ค่าสูงสุด (ก-ข)	1.238	0.517	1.333	0.661	1.463	0.477	1.575	0.625	1.235	0.436	1.341	0.538	1.383	0.652	1.654	0.745
ค่าต่ำสุด (ก-ข)	0.411	0.308	0.477	0.395	0.583	0.293	0.650	0.406	0.543	0.261	0.605	0.364	0.620	0.403	0.720	0.498

จากตารางแสดงอัตราส่วนต่างๆ ที่ใช้ในการพิจารณารูปแบบของพจนต์ภาษาไทย โดยใช้รูปแบบตัวอักษรจำนวน 12 รูปแบบ มีตัวอย่างของตัวอักษรไทยทั้งหมด 2,044 ตัว ได้แก่ พยัญชนะ สระ วรรณยุกต์ และตัวเลขไทย เมื่อสร้างกราฟฮิสโตแกรมของค่าอัตราส่วนความกว้างต่อความสูงจากข้อมูลในตารางก่อนหน้านี้ สามารถแสดงได้ดังกราฟในรูปที่ 1



รูปที่ 1 กราฟฮิสโตแกรมของค่าอัตราส่วนความกว้างต่อความสูงจากข้อมูลตัวอักษรไทย 12 รูปแบบ

จากกราฟ จะได้ค่าเฉลี่ยเท่ากับ 0.647 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.134 สำหรับช่วงค่าที่สามารถครอบคลุมความเป็นไปได้ของตัวอักษรจะอยู่ในช่วง 0.20 ถึง 1.20

ภาคผนวก ก

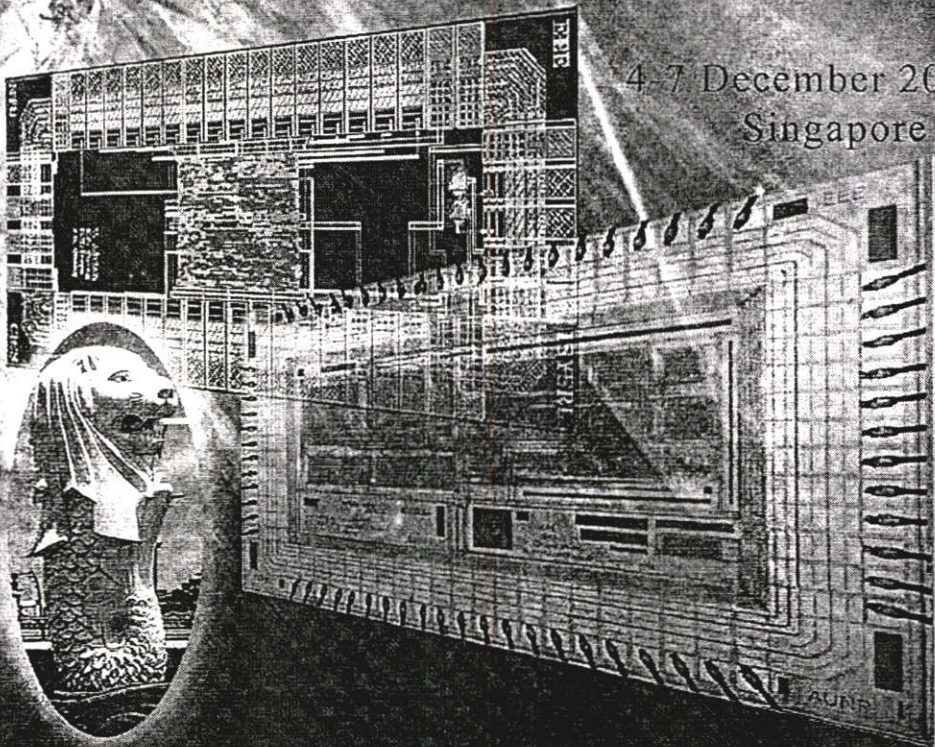
ผลงานวิจัยที่ได้รับการตีพิมพ์เผยแพร่

- [1] W. Jirattitichareon and T. H. Chalidabhongse, **Automatic Detection and Segmentation of Text in Low Quality Thai Sign Images**, Proc. of 2006 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS'06), Singapore, December 4-7, 2006.
- [2] W. Jirattitichareon and T. H. Chalidabhongse, **Text Detection and Segmentation from Thai Sign Images**, Proc. of the International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC'06), Thailand, July 10-13, 2006.

APCCAS 2006

2006 IEEE Asia Pacific Conference
on Circuits and Systems

4-7 December 2006
Singapore



Final Program and Abstracts



D2-PM1-RM6.4**A Real-Time Digital Baseband Channel Emulation System for OFDM Communications**

C C Fu, T P Wang, K C Chang, C H Liao, T D Chiueh
National Taiwan University, TAIWAN

This paper presents the design and implementation of a digital baseband channel emulator. The proposed emulator consists of several wireless channel effects and RF front-end impairments. The channel effects include an additive white Gaussian noise source and a multi-path fading channel with up to 12 faders that can have programmable delays, magnitude, and a Doppler spectrum or a flat spectrum. Other RF electronics effects such as power amplifier non-linearity, carrier frequency offset, oscillator pulling, phase noise, I/Q imbalance, DC offset, and sampling clock offset are also modeled and designed in digital logic. The whole emulator is implemented on FPGA and can run up to 40 MHz, making it suitable for most OFDM wireless systems.

D2-PM1-RM6.5**Synchronization Issues in OFDM Systems**

H Zhou, *A V Malipatil, Y F Huang
University of Notre Dame, UNITED STATES
*LSI Logic, Inc., UNITED STATES

Orthogonal frequency division multiplexing (OFDM) is a bandwidth efficient modulation technique suitable for broadband wireless communications in frequency selective fading channels, yet, its performance is sensitive to timing and carrier synchronization errors. This paper provides an overview on the causes and the effects of these synchronization errors. We survey various symbol timing and carrier acquisition schemes available in the literature including coarse and fine synchronization techniques, using cyclic prefix or training symbols. Blind estimators, which also do not require training symbols, are discussed as viable alternatives when data throughput is very critical.

Regular Session D2-PM1-RM7

Pattern Recognition I

1400 - 1515 12/6/2006 Wednesday

Room : Falcon

Chair(s) : Pei-Yung Hsiao

TAIWAN

D2-PM1-RM7.1**Footstep Recognition with Psycho-acoustics Parameter**

A Itai, H Yasukawa
Aichi Prefectural University, JAPAN

The characteristics of a footstep are determined by the gait, the footwear and the floor. Accurate footstep analysis would be useful in various applications, home security service, surveillance and understanding of human action since the gait expresses personality, age and gender. The feasibility of personal identification has been confirmed by using the feature parameter of footsteps, however, the recognition rate of this method decreases as the number of subjects increases. This paper applies psycho-acoustics parameter to feature extraction. Results show that the parameter proposed herein yields effective and practical personal identification.

D2-PM1-RM7.2**A Robust Anti-collusion Coding in Digital Fingerprinting System**

J Yang, X Xu
Wuhan University of Technology, CHINA

A novel robust anti-collusion encoding scheme which combines the Balanced Incomplete Block Design (BIBD) code with Turbo code in digital fingerprinting system is proposed. The fingerprints are coded by Turbo code before embedding and added to the wavelet domain which is suitable for the visual characteristics of human. The experimental results indicate that the proposed algorithm can successfully resist not only the collusion attacks by colluders, but also the other attacks such as noise when the digital products are transmitted in the channel. Meanwhile it can decrease the bit error rate (BER) of the fingerprints and enhance the overall resistance of the system.

D2-PM1-RM7.3**Automatic Detection and Segmentation of Text in Low Quality Thai Sign Images**

W Jirattichareon, T H Chalidabhongse
King Mongkut's Institute of Technology,
Ladkrabang, THAILAND

A system for automatic detection and segmentation of text in low quality Thai sign images is presented in this paper. The method is designed as a part of a real-time Thai sign translator system which can be used in many applications. First, an input image is pre-processed to enhance its quality. Secondly, we apply LoG (Laplacian of Gaussian) to the image

for edge detection. After edge detection, in the third step, we perform connected component labeling and morphological operations for contour filling. To detect an element in a Thai sentence, we have to set some appropriate ratios and compare them with each closed region to find consonants, vowels, tones and special symbols. Next, we employ 4-line Thai character criteria for layout analysis. Finally, we use GMM (Gaussian Mixture Model) to represent foreground and background, and perform color segmentation in selected color model. Finally, a method for perspective distortion correction is also performed to prepare the segmented texts be ready for further recognition process. We tested the system on 192 Thai sign images which contain total of 4681 characters. The images were captured from several environments in various lighting conditions. The detection accuracy is 90.22%.

D2-PM1-RM7.4

Real-Time Implementation of a Particle Filter with Integrated Voice Activity Detector for Acoustic Speaker Tracking

A M Johansson, E A Lehmann, S Nordholm
Western Australian Telecommunications
Research Institute, AUSTRALIA

In noisy and reverberant environments, the problem of acoustic source localisation and tracking (ASLT) using an array of microphones presents a number of challenging difficulties. One of the main issues when considering real-world situations involving human speakers is the temporally discontinuous nature of speech signals: the presence of silence gaps in the speech can easily misguide the tracking algorithm, even in practical environments with low to moderate noise and reverberation levels. This work focuses on a real-time implementation of the ASLT algorithm proposed in [1], which circumvents this problem by integrating measurements from a voice activity detector (VAD) within the tracking algorithm framework. The algorithm is here optimized for low computational complexity, and is implemented on a PC based real-time system. The resulting computational load is calculated and is presented along with real measurements of the true execution speed for the considered algorithm implementation. The results show that the algorithm is suitable for implementation in currently existing low-power embedded systems.

D2-PM1-RM7.5

Early Detection on the Condition of Pancreas Organ as the Cause of Diabetes Mellitus by Real Time Iris Image Processing

A D Wibawa, *M H Purnomo
Hang Tuah University, INDONESIA
*Institut Teknologi Sepuluh Nopember,
INDONESIA

Iridology is an alternative method in evaluating the condition of our internal organ by looking at the image of iris. Evaluating the iris is done by detecting the presence of some broken tissues in iris. In this paper, input image of the iris will be taken by using a video camera in real time. The presence of broken tissues in iris in a certain area will represent the condition of certain organ according to the chart of iris. The chart of iris that will be used in this research is the international chart of iris Dr. Bernard Jensen. The organ that will be observed is Pancreas. Pancreas's position in iris is at 07.15 - 07.45 when a circle of iris is divided by 120 points. Some image processing methods will be used to enhance the quality of image of iris so that the broken tissues in area of pancreas can be detected clearly. Finally, the result of this detecting method will be compared with two kind of testing. The first is the result will be compared with the data from Iridology Practitioner in Surabaya and the second is the result will be compared with Insulin normality test.

Regular Session D2-PM2-RM1

Graph Theory and Computing

1545 - 1700 12/6/2006 Wednesday

Room : Swallow

Chair(s) : Toshinori Yamada

D2-PM2-RM1.1

On Finding a Solution in the Core of a Multicommodity Flow Game on a Spider

T Yamada, K Karasawa
Saitama University, JAPAN

Motivated by the development of an efficient and stable routing scheme for the Internet, Papadimitriou introduced a multicommodity flow game and raised the problem of whether the core of a multicommodity flow game is always nonempty. Markakis and Saberi settled the problem affirmatively. However, their proof is not constructive, and it is not known how to find a solution in the core of the game, to the best of my knowledge. This paper presents a

Automatic Detection and Segmentation of Text in Low Quality Thai Sign Images

Wittaya Jirattitichareon
Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
s7066450@it.kmitl.ac.th

Thanarat H. Chalidabhongse
Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
thanarat@it.kmitl.ac.th

Abstract—A system for automatic detection and segmentation of text in low quality Thai sign images is presented in this paper. The method is designed as a part of a real-time Thai sign translator system which can be used in many applications. First, an input image is pre-processed to enhance its quality. Secondly, we apply LoG (Laplacian of Gaussian) to the image for edge detection. After edge detection, in the third step, we perform connected component labeling and morphological operations for contour filling. To detect an element in a Thai sentence, we have to set some appropriate ratios and compare them with each closed region to find consonants, vowels, tones and special symbols. Next, we employ 4-line Thai character criteria for layout analysis. Finally, we use GMM (Gaussian Mixture Model) to represent foreground and background, and perform color segmentation in selected color model. Finally, a method for perspective distortion correction is also performed to prepare the segmented texts be ready for further recognition process. We tested the system on 192 Thai sign images which contain total of 4681 characters. The images were captured from several environments in various lighting conditions. The detection accuracy is 90.22%.

Keywords—text detection, text segmentation, thai sign translation

1. INTRODUCTION

A sign is a one-way communication form that we are familiar in our everyday lives. Signs are used in many purposes such as notification, advertising, and warning. We encounter signs in everywhere. Signs make our lives easier if we can read and understand their meanings, but we may face even dangers if we do not understand the message embedded in signs. For example, foreign tourists might get into trouble if they do not understand the Thai warning signs on roads, the national parks or other attractive places.

Automatic detection and segmentation of text in signs taken from natural scenes is very difficult task. Sign Images have the variety of text, background and natural factors. Text can vary in font, size, orientation position and color.

Text can be distorted by shape of the sign it embedded [1] or by camera view angle. It can also be blurred or overlaid by shadow of the other objects. Signs can have variety in their background texture because of the materials they are made from. Signs made from canvas may have wrinkle and hard to detect the exact text region. Unlike canvas signs, wood or metal signs are rigid. They can not be wrinkled but its texture may also cause error text detection. Natural factors, such as the reflection of light and camera flash, can affect the quality of images.

Although Thai optical character recognition (Thai OCR) systems have been developed for a decade [2], [3], [4], but they are not suitable for sign recognition. The OCR systems work well on high quality scanned documents under a controlled environment. They may have a low accuracy rate for sign recognition tasks because of low quality images we mentioned above.

Xilin Chen et al. [5] developed an approach for detection and recognition of signs. The system shows high accuracy rate for Chinese text detecting in sign images. However, for Thai sign detection and recognition task. We found many differences between Chinese and Thai language. Chinese has square-like characters. All characters will be arranged in the same level for composing one sentence. For Thai, the sentence structure is more complicated. A sentence is built up from several elements such as consonants, vowels, tonal marks and other special symbols and these elements may not be aligned in the same line level. The Fig. 1 shows Chinese and Thai sign images.

路段請勿通行

(a)

ลงสู่แหล่งน้ำและที่สาธารณะ

(b)

Figure 1. Sign images (a) A Chinese sign. (b) A Thai sign

There are two challenges for text detection and segmentation of Thai signs: where is the text position and how to segment text elements which are in many levels in one word.

In this paper, we presented a system for automatic detection and segmentation of text in low quality Thai sign images which is a part of real-time Thai sign recognition system. The aim of our development is helping the international tourists overcome the language barriers wherever they are in Thailand. The rest of this paper is organized as follows: section II, we describe the problems of detecting and segmenting of text in low quality Thai sign images. Section III, we present the overview of the proposed system. Section IV we report experiment results. Section V concludes the paper and outlines the future work.

II. PROBLEM DESCRIPTION

In this research, we focus on text detection and segmentation of text in low quality images. Some examples of low quality images captured from natural scenes are shown in Fig. 2.

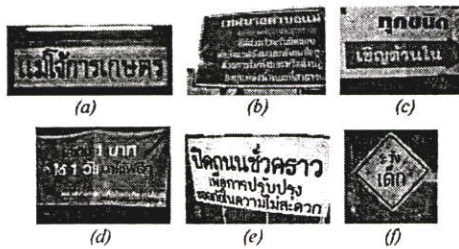


Figure 2. Some examples of signs captured from several conditions and places (a) A sign with reflected light. (b) A sign with a glare from sunlight (c) A sign casted by shadows of surrounding objects. (d) A canvas sign with wrinkles. (e) A perspective sign image. (f) A sign with the metal oxides.

The environment conditions, which affect on text in an image, are summarized and shown in table I.

TABLE I EFFECTS ON TEXT IN SIGN IMAGES CAUSED BY SEVERAL ENVIRONMENT CONDITIONS

Conditions	Effects to text in sign images			
	Size	Contrast	Color	Highlight
Lighting	x	✓	✓	✓
Shadows	x	✓	✓	x
Distance	✓	✓	x	x
Camera view angle	✓	✓	x	x

Compared with the scanned documents in OCR tasks, sign images used in the proposed system are much lower in quality. In OCR, high quality scanned documents are used.

Text in documents usually comes from computer-generated text. But for our experiment system, text can vary in font, size, color and layout. Text can be distorted because of sign shape or camera angle. It can be blurred or overlaid by surrounding object shadows. The variety of signs in their background texture because of the materials they are made from. Furthermore, several conditions such as deformations, highlights, shadows, and specularly can also decrease the accuracy of detection rate.

III. THE SYSTEM OVERVIEW

The schema for an automatic detection and segmentation of text in low quality Thai sign images is shown in Fig. 3.

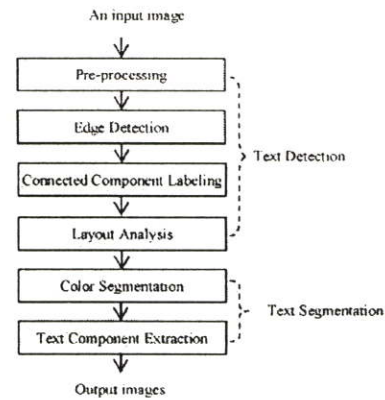


Figure 3. The schema of the proposed system

There are two main processes in the proposed system. One is text detection and another is text segmentation.

A. Text Detection

The characteristics of signs which can be used in our experiment are based on the following assumptions:

- 1) The text is designed with high contrast to its background in both of color and intensity.
- 2) Each character is composed of one or several connected regions.
- 3) The characters in the same context have almost the same in both of size and intensity but may be different in color.
- 4) The characters in the same context have almost the same background patterns.

For text detection, an input image (JPEG) must be processed in pre-processing step to decrease artifacts from DCT procedure. Without this process, it is very difficult task to find an edge set because of noise in original image. Next,

IV. EXPERIMENT RESULTS

We tested the system on 192 low quality images captured from natural scenes. The resolution of each image is 640x480 pixels. The images were taken from several places under uncontrolled conditions. There are many factors that can affect the contrast of text in sign, for instances, fluorescent light reflection, sunlight glare, casting shadows, blurred text, and perspective due to camera angles. The complex of text layout in sign is another factor that may decrease the quality of the images.

Fig. 6 shows examples of our text detection and segmentation results. The red rectangle indicates the detected text regions. We tested the system on a set of 192 various low quality Thai sign images that contain the total of 4681 characters. The system yielded a recognition rate of 90.22%. The ground-truth used in this experiment was from human detection and segmentation.



Figure 6. Examples of correct detection and segmentation results.

However, there were some fault detections due to some problems as shown in Fig. 7. The detection problems arise when signs have complex layout. Another problem we faced is the distortion of the text due to perspective.



Figure 7. Examples of fault detection and segmentation results.

One of the common problems for camera taken images is the perspective distortion. To solve this problem, we employed geometric transformation [6] to restore geometric distortion. The results after the distortion restoration are shown in Fig. 8. This step is done for two purposes, one is to improve the text segmentation accuracy and the other is to prepare the images for further recognition process.



Figure 9. Distortion recovery results.

V. CONCLUSION

In this paper, we presented a system for automatic detection and segmentation of text in low quality Thai sign images. This research is designed as a part of a real-time Thai sign translator system which can be used in many applications such as driving and tourist assistant systems.

REFERENCES

- [1] J. Ohya, A. Shio, and A. Akamatsu, "Recognition of characters in scene images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, no. 2, pp.214-220, 1994.
- [2] S. Watcharabutsarakham, "Using projection and loop for segmentation of touching Thai typewritten," *IEEE International Symposium on Communications and Information Technology*, 2004 Vol. 1, Oct 2004, pp.504-508.
- [3] Watcharabutsarakham, "Segmentation for touching Thai typewritten," *IEEE Region 10 Conference, TENCON 2004 Vol. A*, Nov 2004, pp.199-202.
- [4] N. Premchaiswadi, W. Premchaiswadi and S. Narita, "Segmentation of Horizontal and vertical Touching Thai Characters," *IEICE TRANS.*, pp.987-994, 2000.
- [5] Xilin Chen, Jie Yang, Jing Zhang, and Alex Waibel, "Automatic Detection and Recognition of Signs From Natural Scenes," *IEEE Transactions on Image Processing*, Vol. 13, No. 1, Jan 2004.
- [6] M. S. Brown and W. B. Seales, "Document restoration using 3D shape: A general deskewing algorithm for arbitrarily warped documents," *Proc. ICCV*, vol. 2, 2001, pp. 367-374.



Message from General Chairs

Welcome to the 21 st International Technical Conference on Circuit/Systems, Computers and Communications (ITC-CSCC2006). This is the 2 nd ITC-CSCC held in Thailand. ITC-CSCC2006 is held during July 10-13 2006 in Chiang Mai, the rose of the north, which is a beautiful city with a long history and old traditions.

The goal of the conference is to promote research and development activities. The conference has become a respected forum for the substantive interchange of ideas in all areas of electronics, computers and communications. To realize this objective, we provide an environment in which researchers can exchange and discuss their research ideas, techniques and applications. Around 500 participants from more than 10 different countries are participated in this conference. Over 540 papers are accepted and it is considered to be a big academic conference held in Thailand.

This conference will not become possible if we did not have a strong collaborative work among the colleagues from Korea and Japan. I would like to express my cordial regards to all of them.

I also would like to express my sincere gratitude to all of my colleagues on organizing the conference and their time over the past 6 months. I am grateful to all the authors and speakers for their contribution on research works and participations for making ITC-CSCC2006 a memorable event.

General Chair: Booncharoen Surinavakul
General Co-Chair: Hong Tae Jeon
General Co-Chair: Mitsunori Makino

Foreward of TPCs

The ITC-CSCC is the international conference in circuit, system, computer and communication that has been hosted by Korea and Japan for 20 years. The steering committee of ITC-CSCC has accepted Thailand as the host of ITC-CSCC 2006. This is the great honor for Thailand, and counted as the second time for Thailand to host the conference next to the first one held in Phuket in 2003. The organizing committee of ITC-CSCC 2006 proposed Chiangmai as the venue, and the proposal has been accepted at the steering committee meeting in Jeju Island, Korea in 2005.

There are more than 600 papers in the ITC-CSCC 2006 submitted from Korea, Japan, Thailand and others. This caused technical program co-chairs to work hard for managing the reviewing process. The technical program committee consisting of well-known professors from Korea, Japan and Thailand also worked very hard for reviewing the papers in order to keep quality of the papers presented in the ITC-CSCC.

The technical program starts in the morning of July 11 by the honorable keynote speeches addressed by Prof. Jin Koo Rhee (Dongguk University, Korea), Dr. Osamu Yamauchi (Nippon Telegraph and Telephone Corporation, Japan), and Prof. Sethaporn Cusripituck (The National Telecommunications Commission of Thailand, Thailand). The regular technical program is organized in seven parallel oral sessions and one poster session starting from the afternoon of July 11 through 13, 2006. Hopefully, the keynote speeches and research results presented in the technical program will benefit all participants of ITC-CSCC 2006.

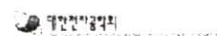
On behalf of technical program chair, I would like to express my appreciations to keynote speakers, authors, technical program committee members, reviewers, and session chairs for their great contributions. Hopefully, all participants of ITC-CSCC 2006 will enjoy presentation, discussion, and proud history & culture of Chiangmai. We look forward to seeing all of you again and again in the next ITC-CSCC.

Technical Program Chair: Kosm Chamnongthai
Technical Program Co-Chair: Kukjun Chun
Technical Program Co-Chair: Takao Onoye

Sponsors



National Electronics and Computer
Technology Center, Thailand



The Institute of Electronics Engineers
of Korea (IEEK), Korea



The Institute of Electronics,
Information and Communication
Engineers (EICE), Japan



The Electrical Engineering/Electronics,
Computer, Telecommunications
and Information Association (ECTD),
Thailand



In association with
IEEE Thailand Section



Computer Vision

Visual Tracking and Recognition based on Segmentation, Simplified SIFT and Evolutionary Particle Filters

Hyung-Soo Kim, Yong-Gun Jo, Hoon Kang (Chung-Ang Univ.)

Development of High-speed Tunnel Fire Detection Algorithm Using the Global and Local Features

Byoungmoo Lee, Dongil Han (Sejong Univ. Seoul)

Model-Based Object Recognition by Maximization of Combined Mutual Information of Intensity and Orientation

Yong Sun Kim, Jae Hak Lee, Jong Beom Ra (KAIST)

Text Detection and Segmentation from Thai Sign Images

Witaya Jirattichareon and Thanarat H. Chalidabhongse (King Mongkut's Inst. of Tech Ladkrabang)

<<< Back >>>

Sponsors



National Electronics and Computer
Technology Center, Thailand



The Institute of Electronics Engineers
of Korea (IEEK, Korea)



The Institute of Electronics,
Information and Communication
Engineers (IEICE, Japan)



The Electrical Engineering/Electronics,
Computer, Telecommunications
and Information Association (ECTI),
Thailand



In association with
IEEE Thailand Section

- Improved Channel Estimation Based on LMMSE Method for OFDM Systems III-189
 *Yeon-Seok Kang, Tae-Wook Hwang, Young Soo Kim, Jin-Sang Kim,
 Doug-Young Suh (Kyunghee Univ.)

Session [ThAM2-5] Ubiquitous Network 1

Time: Thursday July 13, 2006, 10:30-11:50

Location: HuaySomPoi1

Chairperson(s): Daranee Hormdee (Khon Kaen Univ.)

- A Study on Rendering Engine for EPG Program III-193
 *Sang-Yep Nam (Kyung Moon College), Sang-Won Lee (KETI), Chun-Woo Lee
 (Woosong Information College), Deuk-Jea Park, In-Jung Park (Dankook Univ.)
 Investigation of Symmetric Key Cryptographic Algorithms for Security of Wireless
 Sensor Networks III-197
 *Kyung Jun Choi, Jong-In Song (Gwangju Inst. of Science and Tech.)
 Home Station. Novel Architecture of Home Gateway and Its Implementation with
 Switch Chip and Common Communication Protocol III-201
 *Kwang-Soon Choi, Yang-Keun Ahn, Young-Choong Park,
 Kwang-Mo Jung (Korea Electronics Tech Inst.)
 Optimization of Water Period for Best Effort Service in wireless MAN III-205
 *Jin Kyung Park, Woo Cheol Shin, Jun Ha, Cheon Won Choi (Dankook Univ.)

Session [ThAM2-6] Computer Vision

Time: Thursday July 13, 2006, 10:30-11:50

Location: HuaySomPoi2

Chairperson(s): Gen Fujita (Osaka Univ. of Electro-Communications)

- Visual Tracking and Recognition based on Segmentation, Simplified SIFT and
 Evolutionary Particle Filters III-209
 *Hyoung-Soo Kim, Yong-Gun Jo, Hoon Kang (Chung-Ang Univ.)
 Development of High-speed Tunnel Fire Detection Algorithm Using the Global and
 Local Features III-213
 *Byoungmoo Lee, Dongil Han (Sejong Univ. Seoul)
 Model-Based Object Recognition by Maximization of Combined Mutual Information of
 Intensity and Orientation III-217
 *Yong Sun Kim, Jae Hak Lee, Jong Beom Ra (KAIST)
 Text Detection and Segmentation from Thai Sign Images III-221
 *Wittaya Jirattichareon and Thanarat H. Chalidabhongse
 (King Mongkut's Inst. of Tech. Ladkrabang)

Session [ThAM2-7] Linear Nonlinear Systems 2

Time: Thursday July 13, 2006, 10:30-11:50

Location: HuayNamDang1

Chairperson(s): Sang-Yep Nam (Kyung Moon College)

- Design and Implementation of Feedback-Limited Nonlinear Shift Registers III-225
 *Kunihiko Kudou, Akio Tsuneda, Takahiro Inoue (Kumamoto Univ.)
 Measurement of Reflected Wave of an MOSFET for Crosstalk Analysis III-229
 *Apiruk Sangkrasin, Takemi Watanabe, Fujihiko Matsumoto,
 Yasuaki Noguchi (National Defense Academy, Japan)
 Directional Algebraic Reconstruction Technique for Electrical Impedance Tomography III-233
 *J.H. Kim, B.C. Kang, B.Y. Choi (Kyungpook National Univ.),
 K.Y. Kim (Cheju National Univ.)
 Performance of Asynchronous DS-CDMA Systems on Multipath Propagation Channels
 with Digitalized Chaotic Sequences III-237
 *Daisaburo Yoshioka (Sojo Univ.), Akio Tsuneda (Kumamoto Univ.)

Session [ThPM1-1] Signal Processing 4

Time: Thursday July 13, 2006, 13:00-14:40

Location: HuayNamKow1

Chairperson(s): Nobuhiko Sugino (Tokyo Inst. of Tech.)

TEXT DETECTION AND SEGMENTATION FROM THAI SIGN IMAGES

Wittaya Jirattitichareon and Thanarat H. Chalidabhongse

Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang
Bangkok, THAILAND
{s7066450@it.kmitl.ac.th and thanarat@it.kmitl.ac.th}

ABSTRACT

This paper describes a new approach for text detection and segmentation from Thai sign images. The approach consists of five main steps. We first detect the edges in the image using LoG edge detector. Next, connected components are segmented and analyzed to locate character components such as consonants, vowels, tone, and special notations. Layout analysis is then performed. Gaussian mixture model of color is used to model the color distribution of characters, and then the system uses it in separating the characters from the background color. The final step is to extract text components using information from previous steps. The experimental results show our system can detect and segment Thai texts from sign images with 87.14% accuracy rate.

Keywords: Pattern Recognition, Text Detection, Text Segmentation

1. INTRODUCTION

Signs appear everywhere and are used for information communication, such as news, advertisement, warning, etc. They are ubiquitously seen at the street, highway, tourist places, and most common places around us. However, the only way people can exploit signs is to understand the languages, words, symbols, images or other protocols shown on the signs.

This is our underlying motivation to develop a vision system that can recognize and translate Thai signs from images in real-time. This report describes our work on detecting and segmenting texts from Thai sign images which is a part of the whole system. Although there have been quite a number of researches in Thai OCR such as the works appeared in [1,2,3], those approaches do not work well for Thai sign recognition due to many limitations. Most of the previous works in Thai OCR work on document images that are scanned with fine resolution under controlled environment. However, in sign recognition, the images are captured from natural scenes that are usually out of control. The taken images might contain noises from many sources such as illumination reflection from sun or camera flash, and casting shadow from other objects. Moreover, the image

of the text and sign can be affected from perspective transformation which causes the texts in the same sign having varying sizes. All of these affect the performance of text detection, segmentation and sign recognition.

Chen, Yang, Zhang, and Waibel [4] developed a system that detects, recognizes, and translates Chinese signs from images. The system is capable to translate Chinese sign to English. Applying their techniques directly to Thai sign recognition is not simple due to differences of languages. Chinese characters have blocking characteristics while Thai characters are not. Thai characters arrange in four levels and each of them has different size. Moreover, some words contain connected characters. Fig.1 shows samples of Chinese and Thai texts appear on signs.



(a)



(b)

Fig. 1: Sign images: (a) A Chinese sign (b) A Thai sign

This report is organized as follows: the next section describes characteristics of sign images. Then, the third section overviews our system. Experiments and results are shown in Section 4, and the Section 5 lists the conclusion and future works.

2. THAI SIGN IMAGES

Fig. 2 shows samples of Thai sign images taken from several places under different environments. As we can see from these samples, there are quite a numbers of factors that can affect the quality of sign images which will determine the quality of text detection and segmentation. These factors are summarized in Table 1. Comparing the quality of images that are used in typical OCR and in this research, there are many differences. Those images for OCR are scanned with high resolution while the sign images are usually taken from the natural scene that can be interfered by many uncontrolled factors. There are lots of variations in fonts, sizes, and text

Based on our experiments, if the ratio is between 0.5-1.8, the connected component is likely to be text. Otherwise, it will be classified as non-text region. However, as shown in Table 2, there are some vowel characters that this ratio criterion does not work. Thus, some extra works must be done to include these characters in text regions. The Gaussian mixture model of grey level is used to separate the background and foreground pixels of sign images. From pixel classification, we can recover those vowel characters that were missed from using the connected component ratio criterion.

After obtaining the text regions, the layout analysis is then performed to locate the text bounding box. Figure 4 shows a Thai word that composes of several characters which are distributed over four levels. The main characters are located in the normal level which is between line 1 and line 2. The other characters such as vowels, tones, and special symbols usually locate in other levels. We will use this sample word in explanation of our layout analysis.



Fig. 4: A sample of Thai word that characters distribute into all four levels.

The layout analysis is done as follows:

1. Locate the top and bottom boundaries of normal level which contains main text regions (line 1 and line 2 in Fig. 4). Then, we compare these top and bottom boundaries with the boundaries of the neighbor text regions. If they are about the same, we can merge these neighbor regions into one text region.

2. For vowels and tones that locate above and below the normal level, we will apply the same algorithm in step 1 to locate and estimate the distance between line 1 and line 4 and the distance between line 2 and line 5.

3. Use Gaussian Mixture Model of grey level to classify the pixels and recover the missing symbols from previous step. The GMM equation is shown in Eq. 3.

$$f(c) = (1-\alpha)G_{\text{back}}(\mu_b, \sum_b) + \alpha G_{\text{fwd}}(\mu_f, \sum_f), (0 \leq \alpha \leq 1) \quad (3)$$

where $G_{\text{back}}(\cdot)$ is the distribution of background and $G_{\text{fwd}}(\cdot)$ is the distribution of foreground.

However, signs can contain several lines of texts. We perform the same algorithm explained above throughout the entire sign regions. Thus a sign might contain several text regions. Each text region is then passed to next process which is text extraction.

3.2 Text Extraction

From previous steps, we obtain several text regions with their boundaries. Next step is to extract each of the text

characters. The Gaussian Mixture Models are used to model the distribution of colors and brightness (R, G, B, H and I) of background and foreground pixels. Figure 5 shows sample of color models. By using result from connected component layout analysis and color classification based on GMM of background/foreground color models, we then can extract the text characters from images.

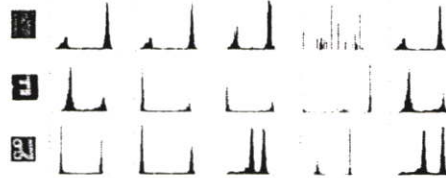


Fig. 5: Distribution of background and foreground color and brightness (R, G, B, H and I) of sign images.

4. EXPERIMENTAL RESULTS

We tested our system on 70 low resolution (640x480) sign images taken from natural scenes in several places. Each image contains some challenges such as reflections, glare, shadows, blurs, dirty background, etc. Most of the signs contain several lines of texts. Some same line texts have different colors. Some images were taken at night, so the texts are low contrast and brightness.

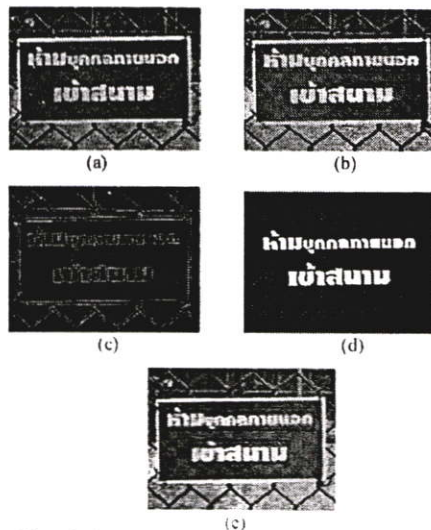


Fig. 6: Detection and segmentation result on a low resolution image. (a) input color image (b) grey-scale image (c) edge image from LoG edge detector (d) connected component result (e) result after layout analysis.

Figure 6 shows the detection and segmentation result on a sign image captured from a tennis court. Although the input image is quite blur and casted by shadows from tennis court fence, but the output shows a reasonably correct result.

Figure 7 shows results of our system applied on other sample images. Over the 70 tested images, our Thai text detection and segmentation from sign images yields 87.14% accuracy rate. The ground-truth used in this test was from human detection and segmentation.



Fig. 7: Results on some sample sign images. The segmented text regions are bounded by red bounding boxes.

However, there were some detection errors found which are shown in Figure 8. Problems include complex layout, having other pictures or symbols that have similar compact ratio as the characters, strong perspective effect, low background/foreground contrast, etc.

5. CONCLUSION AND FUTURE WORKS

This report describes a novel approach to detect and segment Thai text from sign images which is a part of a

real-time Thai sign recognition and translation. The current detection and segmentation results yield a reasonable accuracy rate. However, there are some problems that we plan to cope with in the enhanced system such as the errors from perspective effects, low contrast images, etc. The research will be also extended in the OCR and translation areas for the future works.



Fig. 8: Results of some text segmentation errors.

6. REFERENCES

- [1] S. Watcharabutsarakham, "Using projection and loop for segmentation of touching Thai typewritten," *IEEE International Symposium on Communications and Information Technology*, Vol. 1, pp.504 – 508, 2004.
- [2] Watcharabutsarakham, "Segmentation for touching Thai typewritten," *IEEE Region 10 Conference TENCN*, Vol. A, pp.199 – 202, 2004.
- [3] N. Premchaiswadi, W. Premchaiswadi and S. Narita, "Segmentation of Horizontal and vertical Touching Thai Characters," *IEICE Transaction*, pp.987-994, 2000.
- [4] Xilin Chen, Jie Yang, Jing Zhang, and Alex Waibel, "Automatic Detection and Recognition of Signs From Natural Scenes," *IEEE Transactions on Image Processing*, Vol. 13, No. 1, January 2004.

ภาคผนวก ง

โทรศัพท์มือถือรุ่นที่ใช้ถ่ายภาพได้ที่มีความละเอียดอย่างน้อย 640x480 พิกเซล

ตารางแสดงตัวอย่างโทรศัพท์มือถือรุ่นต่างๆ ที่สามารถถ่ายภาพได้ที่มีความละเอียดอย่างน้อย 640x480 พิกเซล ซึ่งเป็นขนาดภาพที่ใช้ทดสอบในงานวิจัยนี้

ที่	ยี่ห้อ	รุ่น
1	โนเกีย	3110, 3230, 3250, 3250, 3250, 3650, 3660 5070, 5140i, 5200, 5300, 5500, 5500, 5700 6080, 6085, 6086, 6101, 6103, 6110, 6111, 6120, 6125, 6131, 6136, 6151, 6170, 6230, 6230i, 6233, 6234, 6255, 6260, 6270, 6280, 6288, 6290, 6300, 6600, 6620, 6630, 6650, 6670, 6680, 6708, 6820, 6822 7088, 7200, 7260, 7270, 7280, 7360, 7370, 7373, 7380, 7390, 7600, 7610, 7710 8800, 9500 E50, E61i, E65, E70, E90 N70, N71, N72, N73, N76, N77, N80, N90, N91, N92, N93, N93i, N95
2	โซนี่ อิริคสัน	D750i K200i, K220i, K300i, K310i, K320i, K500i, K510i, K550i, K600i, K608i, K610i, K610im, K618i, K700, K750i, K790i, K800i, K810i P800, P900, P910, P990i S600i, S700 T230, T630 V600i, V630i, V800 W200i, W300i, W550i, W580i, W610i, W660i, W700i, W710i, W800i, W810i, W830i, W850i, W880i, W900i Z1010, Z310i, Z500, Z520i, Z530i, Z550i, Z558i, Z610i, Z710i, Z750i, Z800i
3	O ₂	O ₂ Frame, O ₂ Atom Life, O ₂ Zinc, O ₂ Graphite, O ₂ X2i O ₂ XDA Stealth, O ₂ XDA Orion
4	Dopod	U1000 , P800w, D810
5	HP	iPAQ hw6965, iPAQ hw6365

ประวัติผู้เขียน

ชื่อ	นายวิทยา จิริจิติเจริญ
วัน เดือน ปีเกิด	1 สิงหาคม 2523
ที่อยู่	37 หมู่ 1 ต.หนองแสง อ.ประจันตคาม จ.ปราจีนบุรี 25130
ประวัติการศึกษา	2546 สถาปัตยกรรมศาสตรบัณฑิต สาขาสถาปัตยกรรม สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ประวัติการทำงาน	2546 – 2547 ผู้ช่วยนักวิจัย คณะสถาปัตยกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง 2549 – 2550 อาจารย์ประจำสาขาแอนิเมชัน วิทยาลัยศิลปะ สื่อ และเทคโนโลยี มหาวิทยาลัยเชียงใหม่