การประมวลผลล่วงหน้าของภาพเอกสารสำหรับการรู้จำตัวอักษรภาษาไทย

นาย เกรียงศักดิ์ เหล็กดี

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

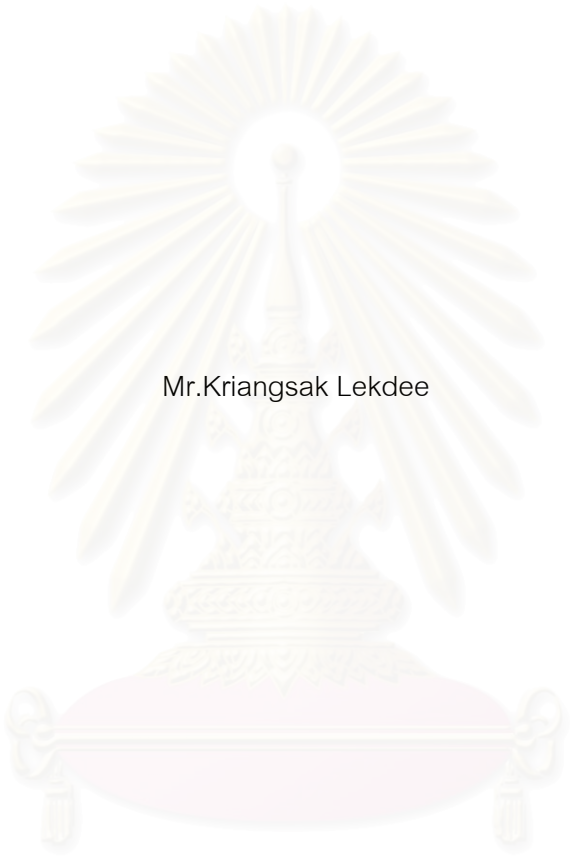
วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2546

ISBN 974-17-5780-8

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

DOCUMENT IMAGE PRE-PROCESSING FOR THAI OPTICAL CHARACTER RECOGNITION

Mr.Kriangsak Lekdee



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science

Department of Computer Engineering
Faculty of Engineering
Chulalongkorn University

Academic Year 2003
ISBN 974-17-5780-8

หัวข้อวิทยานิพนธ์	การประมวลผลล่วงหน้าของภาพเอกสารสำหรับการรู้จำตัวอักษรภาษาไทย
โดย	นาย เกรียงศักดิ์ เหล็กดี
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.ดิเรก ลาวัณย์ศิริ)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(รองศาสตราจารย์ ดร.สมชาย ประสิทธิ์จูตระกูล)

.....อาจารย์ที่ปรึกษา
(ผู้ช่วยศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล)

.....กรรมการ
(อาจารย์ ดร.ฐิต ศิริบูรณ์)

.....กรรมการ
(อาจารย์ ดร.อรรถสิทธิ์ สุรฤกษ์)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เกรียงศักดิ์ เหล็กดี : การประมวลผลล่วงหน้าของภาพเอกสารสำหรับการรู้จำตัวอักษรภาษาไทย. (DOCUMENT IMAGE PRE-PROCESSING FOR THAI OPTICAL CHARACTER RECOGNITION) อ.ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล, 63 หน้า. ISBN 974-17-5780-8.

วิทยานิพนธ์ฉบับนี้มีวัตถุประสงค์เพื่อพัฒนาโปรแกรมสำหรับการประมวลผลล่วงหน้าของภาพเอกสาร ซึ่งจะประกอบไปด้วยโปรแกรมในการสร้างภาพไบนารีจากภาพเอกสารที่ประกอบไปด้วยข้อความ พื้นหลัง หรือภาพถ่าย และโปรแกรมปรับมุมการเอียงของภาพเอกสารโดยอัตโนมัติ

งานวิจัยนี้ได้นำเสนอโปรแกรมในการสร้างภาพไบนารีด้วยอัลกอริทึม 2 อัลกอริทึมคือ อัลกอริทึมไนแบล็กหรืออัลกอริทึมของยาโนวิทซ์และบรัคสเตียน และโปรแกรมปรับมุมการเอียงของภาพเอกสารโดยอัตโนมัติอีก 2 อัลกอริทึมคืออัลกอริทึมฮัจทรานสฟอร์มและเนียเรสเนเบอร์คัทเตอริง ข้อมูลภาพเอกสารที่ใช้ทดสอบประกอบด้วยเอกสาร 2 ชุด เอกสารชุดแรกที่ทำการทดลองการไบไรเซชันจะด้วยอัลกอริทึมไนแบล็กหรืออัลกอริทึมของยาโนวิทซ์และบรัคสเตียน

ผลจากการทดลองในการทำไบนาไรเซชันโดยรวม อัลกอริทึมของไนแบล็กจะมีประสิทธิภาพมากกว่าอัลกอริทึมของยาโนวิทซ์และบรัคสเตียน จากภาพเอกสารจำนวน 10 ภาพ มีจำนวนตัวอักษรรวม 2,316 ตัว ผลการรู้จำตัวอักษรเมื่อนำภาพเอกสารที่ผ่านการทำไบนาไรเซชันด้วยอัลกอริทึมของอ็อดสึ จำนวนตัวอักษรที่ถูกต้องเท่ากับ 49.32% เอกสารภาพที่ผ่านอัลกอริทึมของไนแบล็กจำนวนตัวอักษรที่ถูกต้องเท่ากับ 63.16% เอกสารภาพที่ผ่านอัลกอริทึมของยาโนวิทซ์และบรัคสเตียน จำนวนตัวอักษรที่ถูกต้อง เท่ากับ 60.75%

ส่วนผลจากการทดลองในการปรับมุมการเอียงโดยอัตโนมัติ ด้วยอัลกอริทึมเนียเรสเนเบอร์คัทเตอริงจะเร็วกว่าและมีความละเอียดมากกว่าอัลกอริทึมฮัจทรานสฟอร์ม โดยที่อัลกอริทึมเนียเรสเนเบอร์คัทเตอริงมีค่าของมุมคลาดเคลื่อน ±0° ถึง ±2° และอัลกอริทึมฮัจทรานสฟอร์มมีค่าของมุมคลาดเคลื่อน คลาดเคลื่อน ±0° ถึง ±3°

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต.....

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่ออาจารย์ที่ปรึกษา.....

ปีการศึกษา 2546 ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

##4271407921: MAJOR COMPUTER SCIENCE
KEYWORD : LOCALLY ADAPTIVE BINARIZATION, THRESHOLD, DOUCMENT IMAGE, SKEW DETECTION

KRIENGSAK LEKDEE: DOCUMENT IMAGE PRE-PROCESSING FOR THAI OPTICAL CHARACTER RECOGNITION. THESIS ADVISOR: ASSIST. PROF DR. BOONSERM KIJSIRIKUL, 63 pp. ISBN 974-17-5780-8.

In this thesis, a set of computer programs for document image pre-processing for Thai optical character recognition are proposed. The binarization program generates binary images from document images composing of text, background, and/or image data. Another program performs automatic skew detection of document images.

Niblack's algorithm and Yanowitz and Bruckstein's algorithm were implemented for comparison of binarization. Hough Transform's algorithm and Nearest-Neighbor Clustering's algorithm were implemented for comparison of automatic skew detection.

The experimental result of binarization demonstrates that Niblack's algorithm is superior to Yanowitz and Bruckstein's algorithm. The accuracy of OCR after binarization with this algorithm is 63.75% whereas Yanowitz and Bruckstein's algorithm produces accuracy of only 60.75%.

For automatic skew detection, Nearest-Neighbor Clustering's algorithm provides higher speed and finer skew correction than Hough Transform's algorithm.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Department Computer Engineering	Student's signature…….
Field of study Computer Science	Advisor's signature……..
Academic year 2003	Co-advisor's signature……

กิตติกรรมประกาศ

การทำวิทยานิพนธ์ครั้งนี้สำเร็จลงได้อย่างสมบูรณ์ ผู้วิจัยต้องขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล เป็นอย่างยิ่งที่ได้กรุณาแนะนำหัวข้อวิทยานิพนธ์ ตลอดจนแนวทางการทำวิจัย สละเวลาให้คำปรึกษา ซึ่งเป็นผลผลักดันให้วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยดี ผู้วิจัยขอกราบขอบพระคุณ คุณพ่อ คุณแม่ และครอบครัวที่ได้ให้ชีวิต ให้การศึกษาและกำลังใจตลอดมาจนทำให้ผู้วิจัยได้มีวันนี้ ขอขอบคุณพี่ ๆ และ น้อง ๆ ที่ทำงานรวมทั้งเพื่อน ๆ ทุกคน ที่ให้กำลังใจและช่วยเหลือมาโดยตลอด และสุดท้ายนี้ ขอขอบพระคุณท่านคณะกรรมการการสอบวิทยานิพนธ์ที่ได้ช่วยพิจารณา ให้คำแนะนำตรวจทานแก้ไข และอนุมัติวิทยานิพนธ์ฉบับนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

บทคัดย่อภาษาไทย .....ง
บทคัดย่อภาษาอังกฤษ .....จ
กิตติกรรมประกาศ .....ฉ
สารบัญ .....ช
สารบัญตาราง .....ฌ
สารบัญภาพ .....ญ
บทที่
1. บทนำ.....1
1.1 ความเป็นมาและความสำคัญของปัญหา.....1
1.2 วัตถุประสงค์ของงานวิจัย.....2
1.3 ขอบเขตของวิทยานิพนธ์.....2
1.4 ขั้นตอนการดำเนินการวิจัย.....3
1.5 แนวคิดการดำเนินการวิจัย.....3
1.6 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย.....5

2. แนวความคิดและทฤษฎีที่เกี่ยวข้อง.....6
2.1 แนวคิดและทฤษฎี.....6
2.1.1 ประมวลผลล่วงหน้า.....6
2.1.2 การวิเคราะห์เอกสาร 19
2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง 21

3. วิธีการดำเนินงานวิจัย 23
3.1 ไบนาไรเซชัน 23
3.1.1 อัลกอริทึมของไนแบล็ก 23
3.1.2 อัลกอริทึมของยาโนวิทซ์และบรัคสเตียน 26
3.2 การปรับมุมเอียงของเอกสารแบบอัตโนมัติ 30
3.2.1 ฮัจทรานสฟอร์ม 30
3.2.2 เนียรีสเนเบอร์คัสเตอริง 32

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

4. การทดลองและผลการทดลอง……37
4.1 ไบนาไรเซชัน……37
4.1.1 ขั้นตอนการทดลอง……37
4.1.2 ผลการทดลอง……37
4.1.3 สรุปผลการทดลองของไบนาไรเซชัน……38
4.2 การทดลองการหมุนเอียงของเอกสาร ……39
4.2.1 ขั้นตอนการทดลอง ……39
4.2.2 ผลการทดลอง ……40
4.2.3 สรุปผลการทดลองของการปรับมุมการเอียงของเอกสารแบบอัตโนมัติ ……41

5. สรุปผลการวิจัยและข้อเสนอแนะ……42
5.1 สรุปผลการวิจัย……42
5.1.1 โปรแกรมไบนาไรเซชัน……42
5.1.2 โปรแกรมปรับมุมการเอียงของภาพเอกสารโดยอัตโนมัติ……42
5.2 ปัญหาและอุปสรรค……43

รายการอ้างอิง……44
ภาคผนวก……46
ภาคผนวก ก……47
ภาคผนวก ข……53
ภาคผนวก ค……57

ประวัติผู้เขียนวิทยานิพนธ์……63

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

ตารางที่

ตารางที่ 3.1 อัลกอริทึมพื้นฐานสำหรับฮัจทรานสฟอร์ม……..……..…30
ตารางที่ 3.2 อัลกอริทึมฮัจทรานสฟอร์มแบบใช้ Run-Length Encoding……..……31
ตารางที่ 3.3 อัลกอริทึม Nearest Neighbor Clustering…..……..…35
ตารางที่ 4.1 ผลการทดลองเปรียบเทียบประสิทธิภาพความถูกต้องหลังจากผ่านโปรแกรมรู้จำตัวอักษร……..……..38
ตารางที่ 4.2 ผลการทดลองเปรียบเทียบประสิทธิภาพของการหามุมเอียง…..……40

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

รูปที่
รูปที่ 2.1 ขั้นตอนการวิเคราะห์เอกสาร.....6
รูปที่ 2.2 จุดภาพรบกวน.....6
รูปที่ 2.3 การปรับค่าระดับความเข้มของจุด.....7
รูปที่ 2.4 ฮีสโตแกรมระดับความเข้มของภาพ $g(x, y)$ ที่มีองค์ประกอบภาพส่วนที่เป็นวัตถุค่อนข้างมืดบนส่วนพื้นหลังที่สว่าง.....8
รูปที่ 2.5 การจำแนกภาพด้วยวิธีการกำหนดค่าขีดแบ่งแบบครอบคลุม(ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์ (ค) ฮีสโตแกรมระดับความเข้มของภาพต้นฉบับ.....9
รูปที่ 2.6 ฮีสโตแกรมระดับความเข้มของภาพ g(x,y) ที่ความไม่สม่ำเสมอของระดับความเข้มในองค์ประกอบของภาพ.....10
รูปที่ 2.7 การจำแนกภาพด้วยวิธีการกำหนดค่าขีดแบ่งแบบปรับค่า (ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์ (ค) ฮีสโตแกรมระดับความเข้มของภาพต้นฉบับ.....11
รูปที่ 2.8 หน้าต่างขนาด 3 * 3 ที่ใช้สำหรับหาค่ากราเดียนท์โดยวิธีของโซเบิ้ล (ก) สำหรับหาค่า f_x (ข) สำหรับหาค่า f_y 13
รูปที่ 2.9 กราฟเส้นตรง.....14
รูปที่ 2.10 เส้นตรง $\rho - \theta$ 15
รูปที่ 2.11 วิธีการหามุมความเอียงของภาพเอกสารด้วยวิธีเนียรีสเนเบอร์ (ก). Text (ข). จุดศูนย์ถ่วงในตัวอักษรใน (ก) (ค). ส่วนเชื่อมระหว่างจุดที่ใกล้ๆ กัน (ง). กราฟแสดงมุมเอียงต่างๆ ที่ใกล้ 0 องศา.....16
รูปที่ 3.1 แสดงการเลื่อนของหน้าต่าง.....24
รูปที่ 3.2 ผังการทำงานของอัลกอริทึมของไนแบล็ก.....25
รูปที่ 3.3 ผังการทำงานของอัลกอริทึมของยาโนวิทซ์และบรัคสเตียน.....26
รูปที่ 3.4. แสดงการทำงานของการหาขีดแบ่งพื้นผิว.....28
รูปที่ 3.5 แสดงภาพผ่านการไบนาไรเซชัน ก) เป็นภาพต้นฉบับ ข) ภาพผ่านอัลกอริทึมของไนแบล็ก ค) ภาพผ่านอัลกอริทึมของยาโนวิทซ์ และบรัคสเตียน.....29
รูปที่ 3.6 อธิบายการเข้ารหัส Run-Length Encoding.....31
รูปที่ 3.7 การทำเนียเรสเนเบอร์คัสเตอร์ริง.....32

รูปที่ ก.1 แสดงแถบคำสั่ง.....47
รูปที่ ก.2 แสดงแถบกลุ่มคำสั่งของแถบคำสั่ง Pre-Processing.....48
รูปที่ ก.3 แสดงแถบกลุ่มคำสั่งของแถบคำสั่ง Filter.....48
รูปที่ ก.4 แสดงแถบกลุ่มคำสั่งของแถบคำสั่ง Edge Detection.....49
รูปที่ ก.5 แสดงแถบกลุ่มคำสั่งของแถบคำสั่ง Binarization.....49
รูปที่ ก.6 แสดงไดอะล๊อกของการเลือกสีพื้นหลัง.....50
รูปที่ ก.7 แสดงไดอะล๊อกของการเลือกพารามิเตอร์ในการประมวลผล ตามอัลกอริทึมของไนแบล็ก.....50
รูปที่ ก.8 แสดงไดอะล๊อกของการเลือกพารามิเตอร์ในการประมวลผล ตามอัลกอริทึมยาโนวิทซ์และบรัคสเตียน.....51
รูปที่ ก.9 แสดงแถบกลุ่มคำสั่งของแถบคำสั่ง Skew Detection.....51
รูปที่ ข.1.....53
รูปที่ ข.2.....53
รูปที่ ข.3.....53
รูปที่ ข.4.....54
รูปที่ ข.5.....54
รูปที่ ข.6.....54
รูปที่ ข.7.....55
รูปที่ ข.8.....55
รูปที่ ข.9.....55
รูปที่ ข.10.....56
รูปที่ ค.1.....57
รูปที่ ค.2.....58
รูปที่ ค.3.....59
รูปที่ ค.4.....60
รูปที่ ค.5.....61
รูปที่ ค.6.....62

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันการติดต่อกับหน่วยงานต่างๆ ไม่ว่าจะเป็นเอกชนหรือรัฐบาล มักจะใช้เอกสารในการติดต่อทั้งสิ้น เมื่อมีการใช้เอกสารมากขึ้นบางส่วนก็มีการทำลายไปบ้างบางส่วนก็มีการเก็บไว้บ้าง เอกสารบางอย่างที่ถูกเลือกให้เก็บไว้ในส่วนหนึ่งก็ต้องนำไปทำลายในช่วงเวลาต่อไป เพื่อช่วยลดภาระในการเก็บเอกสารและมีเอกสารบางส่วนที่ไม่สามารถทำลายได้ เช่น จดหมายเหตุ หนังสือต่างๆ เอกสารสำคัญ เป็นต้น การเก็บเอกสารเหล่านี้ในรูปของกระดาษนั้นจะมีอายุการใช้งานได้ไม่นาน ดังนั้นจึงมีความพยายามในการเก็บเอกสารเหล่านี้ให้มีอายุการใช้งานได้นานเช่น การเก็บในรูปไมโครฟิล์ม หรือการนำคอมพิวเตอร์เข้ามาช่วยในการจัดเก็บเอกสารลงบนสื่อเก็บข้อมูลทางคอมพิวเตอร์ (Computer Media) เช่น ซีดี-รอม (CD-ROM) ฮาร์ดดิสก์ (Hard disk) เทป (Tape) เป็นต้น

การนำคอมพิวเตอร์เข้ามาช่วยในการจัดเก็บเอกสาร มีขั้นตอนในการทำงานคือ ในขั้นแรกเราจะนำเอกสารที่ต้องการเก็บผ่านเครื่องกวาดภาพ (Scanner) ที่ทำหน้าที่รับข้อมูลจากเอกสารมาเก็บเป็นภาพเอกสาร (Document Image) จากนั้นก็นำภาพเอกสารไปเก็บบนสื่อเก็บข้อมูลทางคอมพิวเตอร์ต่อไป แต่ในบางครั้งเรามีความต้องการแก้ไขข้อมูลในภาพเอกสารนี้ หรือต้องการพิมพ์เป็นรายงานออกทางเครื่องพิมพ์ (Printer) การทำงานกับภาพเอกสารนั้นจะทำได้ยาก ดังนั้นจึงต้องมีการแปลงภาพเอกสารไปเป็นแฟ้มข้อมูลเอกสาร (Document Files) ในการแปลงนั้นเราจะใช้โปรแกรมที่เรียกว่าโปรแกรมการรู้จำตัวอักษร (Optical Character Recognition, OCR) เพื่อที่จะสามารถนำมาแก้ไขได้อย่างง่ายดายและสั่งพิมพ์เป็นรายงานตามต้องการได้ โปรแกรมการรู้จำตัวอักษรจะแปลงภาพเอกสารมาเป็นแฟ้มข้อมูลเอกสารได้ดีหรือมีข้อมูลเหมือนกับต้นฉบับเพียงใดนั้น จะขึ้นอยู่กับคุณภาพของภาพเอกสารนั้น

ภาพเอกสารที่ต้องการจะต้องเป็นภาพเอกสารที่มีคุณภาพดี แต่ในทางปฏิบัติพบว่าเอกสารที่มีคุณภาพดีนั้นหาได้ยาก ซึ่งปัญหาที่พบว่าภาพเอกสารนั้นขาดคุณภาพได้แก่ ภาพเอกสารที่มีลักษณะของการเน้นข้อความด้วยสีต่างๆ ภาพเอกสารที่มีรอยเลอะเทอะเปอะเปื้อน ภาพเอกสารที่มีการเขียนไม่ชัดเจนหรือการพิมพ์ที่ไม่มีคุณภาพ ภาพเอกสารที่ส่วนของพื้นหลัง (Background) มีรูปแบบที่ไม่แน่นอน และเงาที่เกิดขึ้นกับภาพขณะกวาดภาพเอกสารด้วยเครื่อง

กวาดภาพ เป็นต้น ซึ่งปัญหาต่างๆ เหล่านี้จะเป็นปัญหาหลักของระบบการวิเคราะห์เอกสาร (Document Analysis)

ดังนั้นในงานวิจัยนี้จะนำเสนอแนวความคิดที่ใช้สำหรับการแก้ปัญหาที่เกิดขึ้นดังกล่าว โดยนำภาพเอกสารที่เก็บในรูปของข้อมูลภาพระดับสีเทา (Gray Scale) มาใช้เทคนิคการทำไบนาไรเซชัน (Binarization Image) [6] [7] [8] เพื่อแยกภาพตัวอักษรและสัญลักษณ์กราฟิกเหล่านั้นออกมาเก็บในรูปของภาพไบนารี โดยที่จุดภาพสีดำจะนิยามถึงตัวอักษรและภาพกราฟิก และจุดภาพสีขาวนิยามถึงพื้นหลัง (Background) และนำภาพไบนารีที่ได้ไปหามุมการเอียง (Skew Angle) ของภาพว่าเอียงไปเท่าใด [16] [17] [18] [19] แล้วทำการปรับภาพไบนารีนั้นให้ตั้งตรง เพื่อจะนำภาพเอกสารที่สมบรูณ์นี้ไปใช้งานกับโปรแกรมการรู้จำตัวอักษร ให้มีความสามารถในการแปลงภาพเอกสารไปเป็นแฟ้มข้อมูลเอกสารได้เหมือนต้นฉบับมากขึ้น

1.2 วัตถุประสงค์ของงานวิจัย

เพื่อพัฒนาโปรแกรมสำหรับการประมวลผลล่วงหน้าของภาพเอกสาร ซึ่งจะประกอบไปด้วยโปรแกรมในการสร้างภาพไบนารีจากภาพเอกสารที่ประกอบไปด้วยข้อความ พื้นหลัง หรือภาพถ่าย และโปรแกรมปรับมุมการเอียงของภาพเอกสารโดยอัตโนมัติ

1.3 ขอบเขตของวิทยานิพนธ์

1.3.1. โปรแกรมจะทำงานภายใต้ระบบปฏิบัติการ Microsoft Windows

1.3.2. พัฒนาโปรแกรมในการสร้างภาพไบนารีจากภาพเอกสารที่ประกอบไปด้วยข้อความ สีพื้น ภาพถ่าย และโปรแกรมปรับมุมของการเอียงโดยอัตโนมัติ

1.3.3. แฟ้มข้อมูลรูปภาพเอกสารที่จะใช้ในการทดสอบจะต้องมีลักษณะดังนี้

1.3.3.1 แฟ้มข้อมูลรูปภาพเอกสารที่รับจากเครื่องกวาดภาพต้องจัดเก็บไฟล์แบบ BMP มีความละเอียด (resolution) 300 จุดต่อนิ้ว (dpi, dot per inch)

1.3.3.2 ภาพเอกสารที่ใช้ในการวิจัยจะต้องมีระดับความเข้มของสีเทา 256 ระดับ (Gray Scale) โดยลักษณะของภาพเอกสารที่ใช้ในการวิจัยส่วนที่เป็นตัวอักษรจะมีสีดำ และส่วนที่เป็นภาพถ่ายมีลักษณะของการไล่ระดับที่แตกต่างกันของสีเทา โดยที่จะมีระดับความเข้มอยู่ระหว่างสีขาวและสีดำ

1.3.3.3 ภาพเอกสารที่ใช้ในการวิจัยส่วนของตัวอักษรและภาพถ่ายนั้นจะแยกกันอย่างชันเจน กล่าวคือตัวอักษรและภาพถ่ายนั้นจะไม่อยู่ปะปนกัน เช่น ตัวอักษรจะไม่ปรากฏอยู่บนภาพถ่าย

1.3.3.4 ภาพเอกสารที่ใช้ในการวิจัยจะประกอบด้วยตัวอักษร พื้นหลัง หรือรูปภาพ

1.3.3.5 ภาพเอกสารที่ใช้ในการวิจัยมีการเอียงของภาพอยู่ระหว่าง 0-45 องศา

1.3.4. นำผลของการประมวลผลภาพเอกสารที่ได้ไปทดสอบกับโปรแกรมรู้จำตัวอักษรภาษาไทย

1.3.5. วิธีการวัดผลโปรแกรม

1.3.5.1 สำหรับโปรแกรมในส่วนการสร้างภาพไบนารี จะทำการทดสอบภาพเอกสารที่มีคุณลักษณะที่มีความแตกต่างกัน จำนวน 7 ภาพ และนำภาพที่ได้ไปผ่านโปรแกรมรู้จำตัวอักษรภาษาไทยเพื่อเปรียบเทียบระหว่างภาพต้นฉบับกับภาพที่ผ่านขบวนการไบนาไรเซชัน

1.3.5.2 สำหรับโปรแกรมในส่วนการหมุนภาพเอกสารโดยอัตโนมัติ จะทำการทดสอบภาพเอกสารที่มีมุมของความเอียงที่แตกต่างกัน จำนวน 5 ภาพ และจับเวลาเปรียบเทียบระหว่างการปรับมุมด้วยโปรแกรมปรับอัตโนมัติและการปรับด้วยมือ

1.4 ขั้นตอนการดำเนินการวิจัย

1.4.1. ศึกษางานวิจัยที่เกี่ยวกับการทำงานของวิธีการไบนาไรเซชันภาพเอกสาร และวิธีการหามุมการเอียงของภาพเอกสาร

1.4.2. ออกแบบขั้นตอนการทำงาน

1.4.3. พัฒนาโปรแกรมตามที่ได้ทำการออกแบบไว้

1.4.4. ทดสอบโปรแกรมกับเอกสารที่ใช้ในการทดลอง

1.4.5. สรุปผลการทำงาน และเขียนวิทยานิพนธ์

1.5 แนวคิดการดำเนินการวิจัย

ในวิทยานิพนธ์ฉบับนี้เน้นการทำไบนาไรเซชันและการปรับมุมการเอียงของภาพเอกสารที่ประกอบไปด้วย ตัวอักษร พื้นหลัง และภาพถ่าย เพื่อให้ได้ภาพเอกสารใหม่ในรูปแบบภาพไบนารี ประกอบด้วยขั้นตอนดังนี้

ขั้นตอนที่ 1 ลดจุดภาพรบกวนด้วยการใช้การกรองความถี่ต่ำ (Low pass filter) ซึ่งเป็นวิธีในการลดจุดภาพรบกวนได้อย่างมากและเป็นที่นิยมใช้กัน

ขั้นตอนที่ 2 สร้างภาพไบนารีของภาพเอกสารโดยการจำแนกส่วนของข้อความและภาพถ่าย ด้วยการทำไบนาไรเซชัน เราจะนำอัลกอริทึมของการทำไบนาไรเซชันอย่างน้อย 2 อัลกอริทึมด้วยกัน เช่น

อัลกอริทึมที่ 1. คืออัลกอริทึมจำแนกภาพแบบขอบรอยต่อของภาพ ตามแนวอัลกอริทึมของ S. D. Yanowitz และ A. M. Bruckstein [8] ซึ่งเป็นอัลกอริทึมที่ใช้ในการจำแนกภาพจากขอบรอยต่อของภาพ

อัลกอริทึมที่ 2. คืออัลกอริทึมจำแนกภาพแบบกำหนดค่าขีดแบ่งเชิงบริเวณ (threshold surface) ตามแนวอัลกอริทึมของไนแบล็ก (Niblack's algorithm) [9] ที่ ∅ivind Due Trier และ Torfinn Taxt [7] ได้เสนอแนวทางในการปรับเปลี่ยนค่าพารามิเตอร์ของสมการและการเลือกขนาดของหน้าต่าง (window) ที่ใช้ในการคำนวณหาค่าขีดแบ่งแบบปรับค่าได้ และได้เสนอให้เพิ่มขั้นตอนหลังการประมวลผล (Postprocessing Step) มาใช้ประกอบกับอัลกอริทึมของไนแบล็ก เพื่อให้ประสิทธิภาพของการทำไบนาไรเซชันได้ผลที่ดีขึ้น

ขั้นตอนที่ 3 การหามุมในการเอียงของภาพเอกสาร และปรับมุมของเอกสารให้ตรงหากภาพเอกสารนั้นเอียงโดยอัตโนมัติ โดยการใช้แนวความคิดมาจากงานวิจัยการประมาณค่าความเอียงหน้าเอกสารสิ่งพิมพ์ โดยจะใช้อัลกอริทึมในการหามุมในการเอียงของภาพเอกสารอย่างน้อย 2 อัลกอริทึมด้วยกัน เช่น

อัลกอริทึมที่ 1. ใช้สมการฮัจทรานสฟอร์ม (Hough Transform) ในการหาค่ามุมความเอียงของภาพเอกสาร โดยนำค่าของจุดต่างๆ ที่เราสนใจมาผ่านสมการฮัจทรานสฟอร์มออกมาเป็นค่าของมุมที่เอียงอยู่ของภาพเอกสาร

อัลกอริทึมที่ 2. ใช้วิธีเนียเรสเนเบอร์คัสเตอร์ริง (Nearest-Neighbor Clustering) ในการหาค่ามุมความเอียงของภาพเอกสาร จะเป็นการกำหนดบริเวณของจุดภาพเป็นบล็อกก่อน หาจุดศูนย์ถ่วง (Centroid) ของแต่ละบล็อก แล้วพิจารณาจุดศูนย์ถ่วงที่อยู่ในระยะทางที่ใกล้ๆ กันมาคำนวณหาค่ามุมความเอียงของภาพเอกสาร

1.6 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย

1.6.1. ได้โปรแกรมในการสร้างภาพไบนารีจากภาพเอกสารที่ประกอบไปด้วยข้อความพื้นหลัง หรือภาพถ่าย

1.6.2. ได้โปรแกรมปรับมุมของการเอียงของภาพเอกสารโดยอัตโนมัติ

1.6.3. เพิ่มประสิทธิภาพในการรู้จำตัวอักษรให้แก่โปรแกรมการรู้จำตัวอักษร

1.6.4. ได้ภาพเอกสารที่พร้อมนำไปประมวลผลด้วยโปรแกรมการรู้จำตัวอักษรเช่น ภาพเอกสารไม่มีจุดภาพรบกวน ภาพเอกสารไม่มีการเอียงของภาพเอกสาร

1.6.5. เป็นแนวทางในการพัฒนางานทางด้านการรู้จำตัวอักษรในการส่วนของการวิเคราะห์เอกสาร

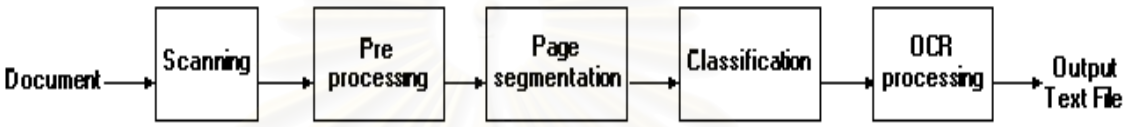
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1. แนวคิดและทฤษฎี

ระบบการวิเคราะห์เอกสาร (Document Analysis System) ประกอบด้วยขั้นตอน ดังรูปที่ 2.1



รูปที่ 2.1 ขั้นตอนการวิเคราะห์เอกสาร

2.1.1. ประมวลผลล่วงหน้า (Preprocessing)

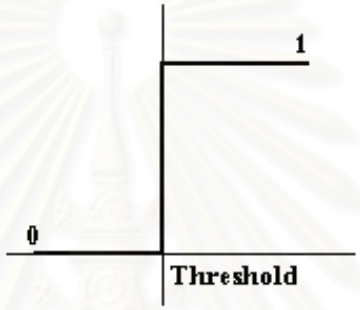
ขั้นตอนประมวลผลล่วงหน้าเป็นขั้นตอนแรกของโปรแกรมการรู้จำตัวอักษร นอกจากจะใช้ในการปรับลักษณะและคุณสมบัติต่างๆ ของรูปภาพเอกสารให้เหมาะสมกับขั้นตอนประมวลผลการรู้จำแล้ว ยังมีส่วนช่วยให้การตัดแยกตัวอักษรและการรู้จำตัวอักษรมีความถูกต้องมากยิ่งขึ้น โดยใช้วิธีการประมวลผลรูปภาพ ซึ่งบางวิธีสามารถทำได้เองด้วยเครื่องคอมพิวเตอร์แต่บางวิธีจำเป็นต้องใช้มนุษย์เป็นผู้พิจารณา ข้อมูลนำเข้าจะเป็นภาพเอกสาร (Document Image) ที่ได้จากการกวาดภาพ หรือแฟ้มข้อมูลรูปภาพที่ได้ทำการกวาดภาพเก็บไว้แล้ว ข้อมูลส่งออกเป็นภาพเอกสารที่เหมาะสมจะนำไปใช้ในขั้นตอนประมวลผลการรู้จำ ขั้นตอนก่อนประมวลผลโดยส่วนใหญ่จะประกอบด้วย

2.1.1.1 **การลดจุดภาพรบกวน** (Noise Reduction) จุดภาพรบกวนเป็นจุดภาพที่เราไม่ต้องการ ซึ่งเกิดขึ้นในขั้นตอนการกวาดภาพหรืออาจเกิดจากต้นฉบับที่เป็นเอกสารที่มีจุดภาพเกินปรากฏอยู่ ดังรูปที่ 2.2

รูปที่ 2.2 จุดภาพรบกวน

จุดภาพรบกวนมีผลทำให้รูปตัวอักษรเพี้ยนไปจากเดิม และจะทำให้ประมวลผลการรู้จำผิดพลาดด้วย จึงมีความจำเป็นจะต้องลดจุดภาพรบกวนเพื่อให้ได้รูปภาพที่ใกล้เคียงตามความเป็นจริง

2.1.1.2 **ไบนาไรเซชัน** (Binarization) เนื่องจากการกวาดภาพนั้น ผู้กวาดภาพสามารถกำหนดจำนวนระดับความเข้มสีของรูปภาพได้ เช่น 256 16 4 หรือ 2 ระดับ แต่การแปลงระดับความเข้มสีของรูปภาพจะขึ้นอยู่กับขั้นตอนประมวลผลการรู้จำนั้น ต้องการรูปภาพที่มีจำนวนระดับความเข้มเป็นเท่าไร โดยส่วนใหญ่จะใช้รูปที่มีจำนวนระดับความเข้มสี 2 ระดับ คือ ขาวและดำ และใช้วิธีการกำหนดค่าขีดแบ่ง (Threshold) เป็นตัวแบ่งระดับดังรูปที่ 2.3



รูปที่ 2.3 การปรับค่าระดับความเข้มของจุด

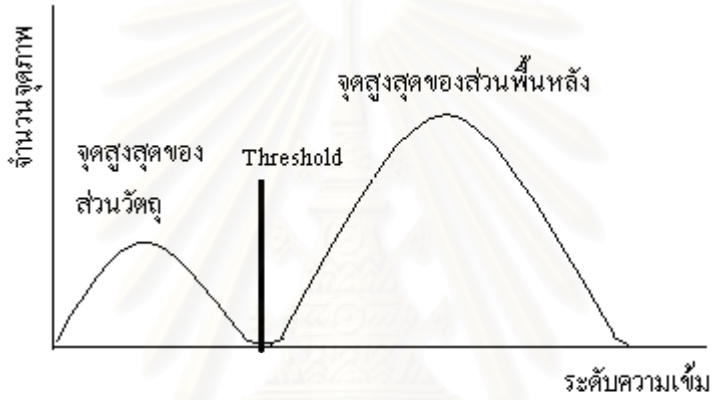
จะเห็นได้ว่าจุดที่มีระดับความเข้มมากกว่าหรือเท่ากับค่าขีดแบ่ง จะถูกปรับให้เป็น 1 และหากจุดที่มีระดับความเข้มน้อยกว่าค่าขีดแบ่ง จะถูกปรับให้เป็น 0 การกำหนดค่าขีดแบ่งถือว่าเป็นเทคนิคที่สำคัญในการประมวลผลภาพในส่วนของการจำแนกภาพ (Image Segmentation)

การจำแนกภาพหมายถึงการแยกองค์ประกอบหรือส่วนต่างๆ ที่อยู่ในภาพออกเป็นส่วนประกอบย่อยๆ ที่สัมพันธ์กับลักษณะทางกายภาพของรูปนั้น และส่วนประกอบที่ถูกแยกออกมาอาจถูกนำไปประมวลผลในส่วนอื่นต่อไป ซึ่งการจำแนกภาพจะมีหลักการทำงานในแนวเดียวกันกับสายตาคน คือสามารถแยกลักษณะเด่นออกมาจากภาพที่มองเห็นได้ และเทคนิคการกำหนดค่าขีดแบ่ง (Thresholding Technique) ถือว่าเป็นเทคนิคในการแบ่งแยกองค์ประกอบภาพที่ง่ายอีกเทคนิคหนึ่ง ซึ่งอาศัยหลักการที่ว่าจุดภาพที่มีคุณสมบัติของภาพอยู่ในช่วงหนึ่งจะถูกจัดเป็นกลุ่มๆ หนึ่ง โดยที่ระดับความเข้มหนึ่งสามารถที่จะแยกกลุ่มของจุดภาพเป็น 2 กลุ่มอย่างชัดเจน คือ กลุ่มของวัตถุ (Object) มีจุดภาพเป็นสีดำ กับกลุ่มของส่วนพื้นหลัง (Background) มีจุดภาพเป็นสีขาว ดังรูปที่ 2.4 ซึ่งแสดงฮีสโตแกรมระดับความเข้มของภาพ $g(x, y)$ ที่มีองค์ประกอบส่วนที่เป็นวัตถุค่อนข้างมืดบนส่วนพื้นหลังที่สว่าง ดังนั้นการแยกกลุ่มทั้งสองออกจากกันอย่างชัดเจนทำได้โดยเลือกค่าขีดแบ่ง T ที่มีค่าระดับความเข้มอยู่ระหว่างกลุ่มทั้งสองบนฮีสโตแกรมระดับความ

เข้มของภาพ และทำการตรวจสอบแต่ละจุดภาพ ถ้าค่า $g(x,y)$ น้อยกว่าค่าขีดแบ่ง $(g(x,y)<T)$ ถือว่าเป็นจุดภาพของส่วนที่เป็นวัตถุ และถ้ามากกว่าหรือเท่ากับค่าขีดแบ่ง $(g(x,y)\geq T)$ ถือว่าเป็นจุดภาพของส่วนที่เป็นพื้นหลังของภาพ ดังนั้นข้อมูลภาพ $g_T(x,y)$ ที่ผ่านการทำค่าขีดแบ่งสามารถนิยามด้วยสมการดังต่อไปนี้

$$g_T(x,y)=\begin{cases}0 & if \quad g(x,y)<T\\ 1 & if \quad g(x,y)\geq T\end{cases}$$

จุดภาพที่นิยามด้วย 0 คือ จุดภาพของส่วนที่เป็นวัตถุ และจุดภาพที่นิยามด้วย 1 คือจุดภาพของส่วนที่เป็นพื้นหลังของภาพ



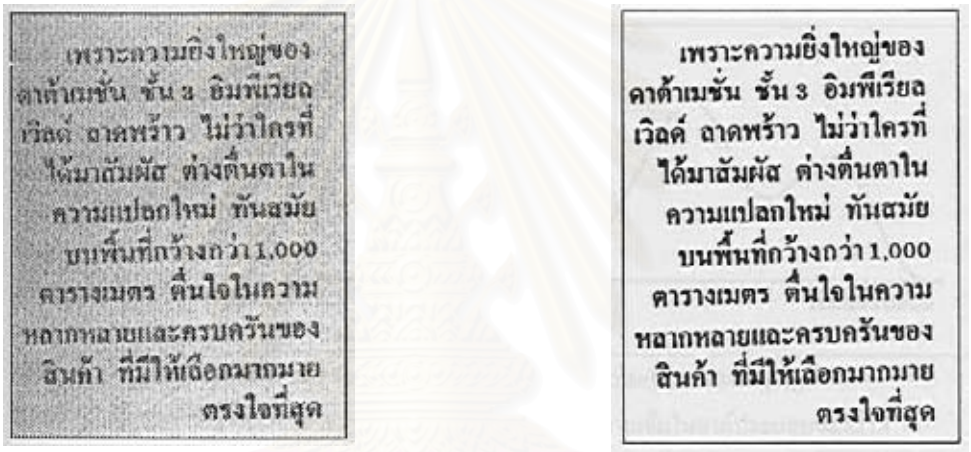
รูปที่ 2.4 ฮีสโตแกรมระดับความเข้มของภาพ $g(x,y)$ ที่มีองค์ประกอบภาพส่วนที่เป็นวัตถุค่อนข้างมืดบนส่วนพื้นหลังที่สว่าง

การจำแนกภาพโดยใช้เทคนิคการกำหนดค่าขีดแบ่ง เพื่อให้ได้ภาพผลลัพธ์ที่มีความถูกต้องเหมาะสมนั้นสิ่งสำคัญที่สุดคือ ค่าขีดแบ่งที่ใช้ เนื่องจากถ้าเลือกค่าขีดแบ่งที่ไม่มีความเหมาะสมภาพผลลัพธ์ที่ได้อาจไม่ถูกต้อง ดังนั้นปัญหาของการจำแนกภาพโดยวิธีการกำหนดค่าขีดแบ่งนี้คือทำอย่างไรจึงจะสามารถคำนวณหาค่าขีดแบ่งที่เหมาะสมสำหรับแต่ละภาพที่นำมาทำการจำแนกภาพ

ส่วนรูปแบบในการกำหนดค่ากับจุดภาพนั้น ภาพที่จุดภาพของส่วนที่เป็นวัตถุกับจุดภาพของที่เป็นส่วนพื้นหลังในแต่ละส่วนมีระดับความเข้มที่สม่ำเสมอ และระดับความเข้มระหว่างทั้ง 2 แตกต่างกันอย่างชัดเจนทั่วทั้งภาพ ซึ่งภาพลักษณะดังกล่าวนี้สามารถใช้ค่าขีดแบ่งเพียงค่าเดียวทำเป็นค่าขีดแบ่งกับแต่ละจุดภาพทั่วทั้งภาพได้ เรียกการกำหนดค่าขีดแบ่งในลักษณะนี้ว่า “การกำหนดค่าขีดแบ่งแบบครอบคลุม (Global Thresholding)” แต่ถ้าภาพที่มีระดับความเข้มไม่สม่ำเสมอเกิดขึ้นในส่วนของวัตถุหรือส่วนของพื้นหลังหรือทั้งสองส่วน ค่าขีดแบ่งแบบครอบคลุม

เพียงค่าเดียวไม่เหมาะสำหรับการทำค่าขีดแบ่งกับแต่ละจัดภาพทั่วทั้งภาพ ค่าขีดแบ่งที่ดีควรมีการปรับเปลี่ยนค่าตามตำแหน่งของจุดภาพนั้น คือ การใช้ค่าขีดแบ่งที่ต่างกันสำหรับจุดภาพที่ตำแหน่งต่างกันและเรียกการกำหนดค่าขีดแบ่งในลักษณะดังกล่าวว่า “การกำหนดค่าขีดแบ่งแบบปรับค่า (Adaptive Thresholding)”

2.1.1.2.1. การกำหนดค่าขีดแบ่งแบบครอบคลุม (Global Thresholding) สำหรับขั้นตอนการหาค่าขีดแบ่งที่ครอบคลุมภาพทั้งภาพโดยอัตโนมัติ โดยทั่วไปแล้วจะมีพื้นฐานของการดำเนินการอยู่บนฮีสโตแกรมระดับความเข้มของภาพ ซึ่งฮีสโตแกรมระดับความเข้มสามารถคำนวณจากจำนวนของจุดภาพในแต่ละระดับความเข้มของภาพ และทำการหาค่าขีดแบ่งในรูปแบบต่างๆ ที่สามารถแบ่งฮีสโตแกรมระดับของความเข้มของภาพออกเป็น 2 ส่วนได้อย่างถูกต้อง คือ ส่วนที่เป็นระดับความเข้มของวัตถุ กับส่วนที่เป็นระดับความเข้มของพื้นหลัง

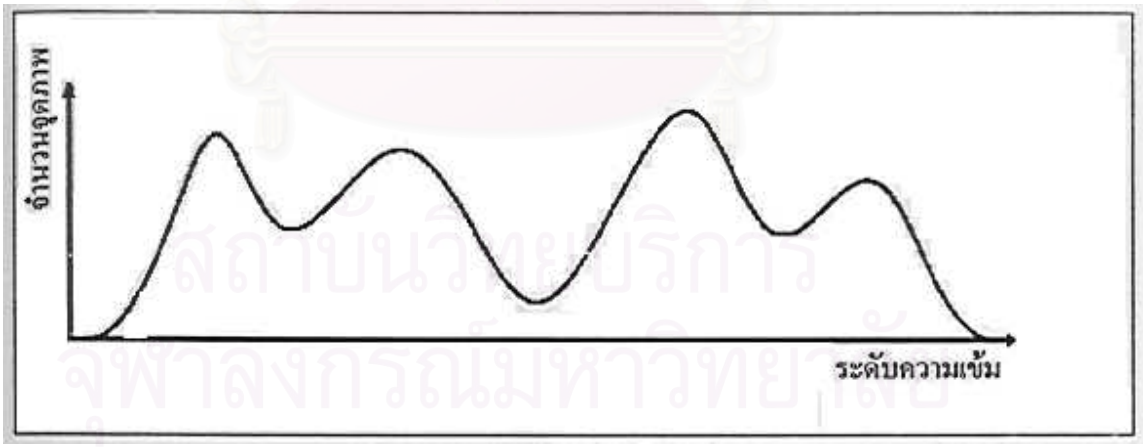


รูปที่ 2.5 การจำแนกภาพด้วยวิธีการกำหนดค่าขีดแบ่งแบบครอบคลุม
(ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์ (ค) ฮีสโตแกรมระดับความเข้มของภาพต้นฉบับ

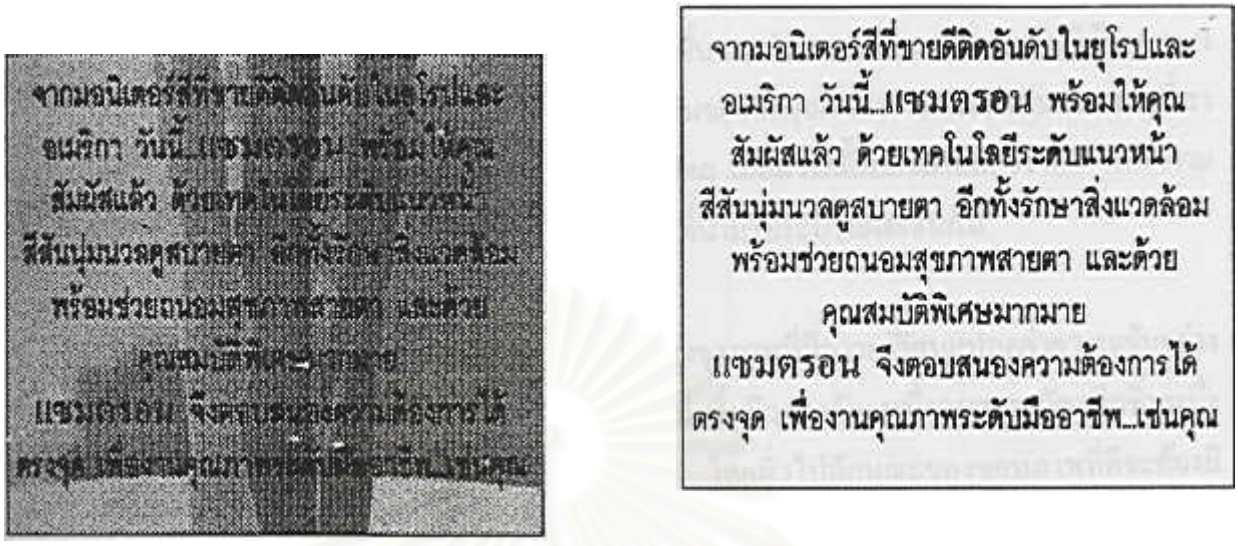
ตัวอย่างเช่น กรณีภาพที่มีอัตราความแตกต่างของระดับความเข้มระหว่างส่วนที่เป็นวัตถุกับส่วนของพื้นหลังสูง และระดับความเข้มที่เกิดขึ้นในแต่ละส่วนมีความสม่ำเสมอ กรณีนี้

สามารถใช้ระดับความเข้มที่มีจำนวนของจุดภาพที่ต่ำที่สุดและอยู่ระหว่างกลุ่มระดับความเข้ม (Peak) ทั้ง 2 กลุ่มฮีสโตแกรมจะเป็นค่าขีดแบ่งดังตัวอย่างในรูปที่ 2.5 หรือในกรณีทั่วๆ ไปค่าขีดแบ่งอาจจะพิจารณาจากค่าระดับความเข้มที่สามารถแบ่งฮีสโตแกรมออกเป็น 2 กลุ่มแล้วทำให้ความแปรปรวนที่เกิดขึ้นระหว่างกลุ่มที่มีค่ามากที่สุด และความแปรปรวนที่เกิดขึ้นภายในกลุ่มที่มีค่าต่ำสุด หลังจากนั้นนำค่าขีดแบ่งที่คำนวณหาได้ ไปกำหนดขีดแบ่งกับแต่ละจุดภาพทั่วทั้งภาพเพื่อให้ได้ภาพผลลัพธ์

2.1.1.2.2. การกำหนดค่าขีดแบ่งแบบปรับค่า (Adaptive Thresholding) กรณีที่ข้อมูลภาพมีความไม่สม่ำเสมอเกิดขึ้นในส่วนของวัตถุหรือส่วนของพื้นหลังหรือทั้งสองส่วน ซึ่งภาพในลักษณะเช่นนี้ฮีสโตแกรมระดับความเข้มของภาพจะมีลักษณะดังเช่นรูปที่ 2.6 การใช้ค่าขีดแบ่งแบบครอบคลุมเพียงค่าเดียวกับภาพอาจทำให้ภาพผลลัพธ์ที่ได้ไม่ถูกต้อง จากปัญหาที่เกิดขึ้นสามารถแก้ไขได้โดยคำนวณค่าขีดแบ่งของแต่ละจุดภาพเทียบกับข้อมูลต่างๆ จากจุดภาพข้างเคียง (Neighborhood of the pixel) ดังตัวอย่างในรูปที่ 2.7 เป็นการแสดงจำแนกภาพด้วยวิธีการกำหนดค่าขีดแบ่งแบบปรับค่า อัลกอริทึมของวิธีการกำหนดค่าขีดแบ่งแบบปรับค่ามีอยู่มาก [6] ดังเช่น อัลกอริทึมของไนแบล็ก (Niblack's Algorithm) [6] และอัลกอริทึมของยาโนวิทซ์และบรัคสเตียน (Yanowitz and Bruckstein's Algorithm) [7] ดังจะอธิบายตามลำดับต่อไป



รูปที่ 2.6 ฮีสโตแกรมระดับความเข้มของภาพ g(x,y) ที่ความไม่สม่ำเสมอของระดับความเข้มในองค์ประกอบของภาพ



รูปที่ 2.7 การจำแนกภาพด้วยวิธีการกำหนดค่าขีดแบ่งแบบปรับค่า
(ก) ภาพต้นฉบับ (ข) ภาพผลลัพธ์ (ค) ฮีสโตแกรมระดับความเข้มของภาพต้นฉบับ

- **อัลกอริทึมไนแบล็ก (Niblack's Algorithm) [6]** ในแนวความคิดของไนแบล็กจะกำหนดให้มีค่าขีดแบ่งหลายๆ จุดบนภาพ โดยค่าของขีดแบ่งนี้จะพิจารณาจากพื้นที่ที่กำลังพิจารณาในขณะนั้นหรือ T(x,y) ซึ่งคำนวณได้มาจากค่าเฉลี่ยท้องถิ่น (Local Mean) และค่าส่วนเบี่ยงเบนมาตรฐานท้องถิ่น (Local standard deviation) ซึ่งจะคำนวณได้จาก

$$T(x,y) = m(x,y) + k * s(x,y)$$

โดยที่ m(x,y) และ s(x,y) หมายถึงค่าเฉลี่ยท้องถิ่นและค่าส่วนเบี่ยงเบนมาตรฐานท้องถิ่น ตามลำดับในพื้นที่รอบข้างของจุดภาพ (x,y) ขนาดของพื้นที่รอบข้างของจุดภาพ (x,y) ควรจะมีขนาดเล็กเพื่อป้องกันการสูญหายของรายละเอียดของข้อมูลในพื้นที่ช่วงนั้น แต่ก็ควรให้มีพื้นที่

ขนาดใหญ่เพียงพอที่จะกำจัดสัญญาณรบกวนได้ จากงานวิจัยนี้พบว่าควรจะใช้พื้นที่รอบข้างเท่ากับ 15 * 15 จะเป็นตัวเลือกที่ดี และค่าของ k จะใช้เป็นเงื่อนไขที่สำคัญในการปรับแต่ง (adjust) เพื่อให้ได้ส่วนที่เป็นตัวอักษรออกมาให้มีประสิทธิภาพที่ดี เช่นเดียวกันจากงานวิจัยนี้ได้ใช้กำหนดค่าของ k = -0.2 ซึ่งจะได้ผลลัพธ์ที่ดีออกมา

ค่าเฉลี่ยท้องถิ่น (Local Mean) จะได้มาจาก

$$m(x,y) = \frac{1}{n}\sum_{x,y} g(x,y)$$

และค่าส่วนเบี่ยงเบนมาตรฐานท้องถิ่น (Local standard deviation) จะได้มาจาก

$$s(x,y) = \sqrt{\frac{1}{n}\sum_{x,y}((g(x,y) - m(x,y))^2}$$

โดยที่ n หมายถึงจำนวนของจุดภาพและ $g(x,y)$ หมายถึงค่าความความเข้มของจุดภาพที่ตำแหน่ง x,y

- **อัลกอริทึมของยาโนวิทซ์และบรัคสเตียน** (Yanowitz and Bruckstein's Algorithm) [7] เริ่มจากการนำภาพเอกสารต้นฉบับมาจำแนกหาขอบภาพ โดยใช้อนุพันธ์เชิงเส้นอันดับที่หนึ่งหรือกราเดียนท์ (Gradient, $\nabla f(x,y)$) โดยในการหาค่ากราเดียนท์บนข้อมูลภาพสามารถกำหนดอยู่ในรูปของเวคเตอร์ได้ดังสมการต่อไปนี้

$$\nabla f(x,y) = \begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

ขนาดของกราเดียนท์ (Magnitude of Gradient) จะเท่ากับอัตราการเปลี่ยนแปลงระดับความเข้มของจุดภาพในทิศทาง G ที่กำหนด สามารถหาได้จากสมการดังต่อไปนี้

$$|\nabla f(x,y)| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} = \sqrt{f_x^{\,2} + f_y^{\,2}}$$

สามารถเขียนให้อยู่ในรูปที่ง่ายโดยใช้การประมาณขนาดของกราเดียนท์ (Approximate Magnitude of Gradient) ได้ดังนี้

$$|\nabla f(x,y)| = |f_x| + |f_y|$$

เนื่องจากอัลกอริทึมของยาโนวิทซ์และบรัคสเตียนได้ใช้หน้าต่าง (Window) ขนาด 3 * 3 มากระทำกับโซเบิ้ลเอจโอเปอร์เรเตอร์ (Sobel's edge operator) หน้าต่างของโซเบิ้ลจะเป็นดังรูปที่ 2.8 (ก) และ (ข) สำหรับในการหาขนาดของกราเดียนท์

-1	0	1
-2	0	2
-1	0	1

(ก)

1	2	1
0	0	0
-1	-2	-1

(ข)

รูปที่ 2.8 หน้าต่างขนาด 3 * 3 ที่ใช้สำหรับหาค่ากราเดียนท์โดยวิธีของโซเบิ้ล
(ก) สำหรับหาค่า f_x (ข) สำหรับหาค่า f_y

เพราะฉะนั้นเราสามารถคำนวณหาขนาดของกราเดียนท์ได้ดังสมการดังต่อไปนี้

$$\left|\nabla f(x,y)\right| = \left(\begin{array}{l}\left|(f(x+1,y-1)+c*f(x+1,y)+f(x+1,y+1))-(f(x-1,y-1)+c*f(x-1,y)-f(x-1,y+1))\right| - \\ \left|(f(x-1,y+1)+c*f(x,y+1)+f(x+1,y+1))-(f(x-1,y-1)+c*f(x,y-1)-f(x+1,y-1))\right|\end{array}\right)$$

โดยที่ C เป็นค่าคงที่เท่ากับ 2 จากนี้เมื่อได้ค่าขนาดของกราเดียนท์จากภาพต้นฉบับ ซึ่งเราจะได้ขอบวัตถุต่างๆ ของภาพและนำไปทำให้ขอบของวัตถุให้เล็กลงเท่ากับ 1 จุดภาพด้วยอัลกอริทึมทินนิง (Thinning Algorithm) แล้วนำขอบภาพที่ได้ทาบลงไปในภาพต้นฉบับ เพื่อหาจุดกราเดียนท์สูงสุด (Point of maximal gradient) ซึ่งนั้นก็คือค่าขีดแบ่งของบริเวณ (Threshold Surface) นั้น จากที่กล่าวมาเป็นขั้นตอนอย่างคร่าวๆ ดังนั้นเราจะกำหนดค่าใหม่ของจุดภาพ P(x,y) ที่รอบการซ้ำที่ (n+1) ได้จากสมการดังต่อไปนี้

$$P_{n+1}(x,y) = P_n(x,y) + \frac{\beta * R_n(x,y)}{4}$$

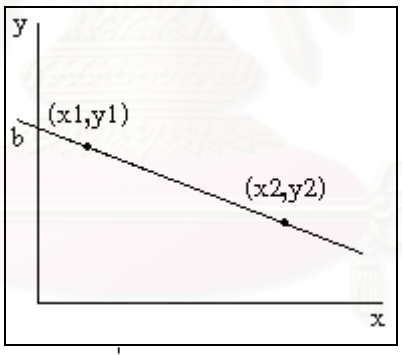
$$R_n(x,y) = P_n(x-1,y) + P_n(x+1,y) + P_n(x,y+1) + P_n(x,y-1) - 4P_n(x,y)$$

โดยที่ P(x,y) หรือโพเทนเชิลเซอเฟส (Potential Surface) ได้มาจากการทำซ้ำๆ (Iterative) ในเค้าร่างของการประมาณค่าในช่วง (Interpolation Scheme) [8] การประมาณค่าในช่วงของพื้นผิว (Interpolation Surface) คือการส่งค่าไปยังจุดภาพขอบภาพ และการประมาณค่าในช่วงตกค้าง (An Interpolation Residual) หรือ $R_n(x,y)$ คือค่าที่คำนวณได้มาจากจุดภาพในส่วนที่ไม่ใช่ขอบภาพ

ยาโนวิทซ์และบรัคสเตียนได้ให้คำแนะนำไว้ว่าค่าของ β ควรจะมีค่าอยู่ระหว่าง $1.0 < \beta < 2.0$ จะทำให้การลู่เข้า (Convergence) รวดเร็วขึ้น อย่างไรก็ตามจากงานวิจัย [7] ได้แนะว่าค่าของ $\beta > 1.0$ ผลลัพธ์เริ่มมีความผิดพลาดไป ดังนั้นควรเลือกค่าของ $\beta = 1.0$

2.1.1.3 **การหมุนภาพ** (Rotation) บางครั้งรูปภาพที่ได้จากการกวาดภาพเอกสารนั้นไม่ตรง อาจเป็นเพราะการกวาดภาพเอียงหรือต้นฉบับเอียง ซึ่งจะทำให้ขั้นตอนประมวลผลการรู้จำนั้นมีความผิดพลาดเกิดขึ้นได้ จึงต้องทำการหมุนภาพให้กลับมาตรงตามที่ต้องการ

- **ฮัจทรานสฟอร์ม** (Hough Transform) เป็นวิธีที่ได้รับความนิยมในการหาเส้นตรงในรูปภาพ โดยทฤษฎีนี้ใช้แนวคิดที่ว่า หากเรามีเส้นตรงในรูปภาพ ดังรูปที่ 2.9.



รูปที่ 2.9 กราฟเส้นตรง

ซึ่งมีสมการเส้นตรงคือ

$$y = mx + b$$

โดยที่ m คือ ความชันของเส้นตรงและ b คือจุดตัดแกน y จากสมการเส้นตรงจะได้ว่า หากเรานำจุด (x1,y1) ซึ่งเป็นจุดที่อยู่บนเส้นตรงไปแทนในสมการ จะได้

$$y1 = m \cdot x1 + b$$

จะเห็นว่าผลจากการแทนค่าจะให้ค่าความชัน m และจุดตัดแกน b และหากนำจุด (x2, y2) มาแทนค่าในสมการ จะได้

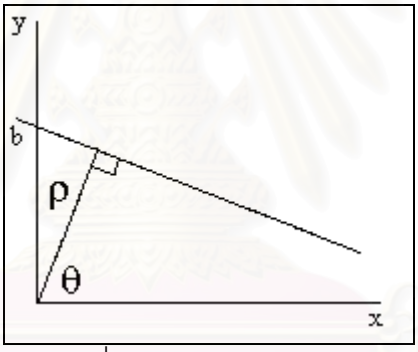
$$y2 = m \cdot x2 + b$$

ซึ่งจะให้ค่าความชัน m และจุดตัดแกนที่ b เช่นกัน เมื่อพิจารณาถึงตรงนี้จะเห็นว่าหากเรานำจุดทุกจุดในรูปภาพที่เราต้องการหาเส้นตรงมาแทนค่าลงในสมการเส้นตรง เพื่อหาค่าความชันและจุดตัดแกน จะได้ว่าความชันและจุดตัดแกนที่พบบ่อยๆ คือเส้นตรงที่อยู่บนรูปภาพนั่นเอง วิธีของฮัจทรานสฟอร์มจะสร้างแถวลำดับ (Array) ขนาด 2 มิติ ที่เรียกว่าแถวลำดับสะสม (Accumulator array) โดยมีตัวชี้ (Index) คือ ความชัน และ จุดตัดแกนสำหรับเก็บค่าจำนวนครั้งที่คำนวณได้ค่าของความชันและจุดตัดแกนนั้นๆ ทุกๆครั้งที่คำนวณได้ค่าความชันและจุดตัดแกนใดๆ ก็จะไปเพิ่มค่าในแถวลำดับสะสม [b , m]

ข้อบกพร่องของการใช้สมการเส้นตรง คือ จะเกิดปัญหาในกรณีที่สมการเส้นตรงในรูปภาพเป็นเส้นตรงที่ขนานกับแกน y ซึ่งจะให้ค่าความชันและจุดตัดแกนเป็นอนันตภาพ (Infinity) ดังนั้นฮัจทรานสฟอร์มเปลี่ยนสมการเส้นตรงที่ใช้เป็นสมการดังต่อไปนี้

$$\rho = x \cos\theta + y \sin\theta$$

โดยที่ ρ คือเส้นตรงที่ลากจากจุด (0,0) มาตั้งฉากกับเส้นตรงที่เราสนใจ และ θ คือมุมที่สมการเส้นตรงทำกับแกน x ดังรูปที่ 2.10



รูปที่ 2.10 เส้นตรง $\rho - \theta$

เนื่องจากสมการเส้นตรงถูกเปลี่ยนมาอยู่ในรูป $\rho - \theta$ ดังนั้นตัวชี้ของแถวลำดับสะสม จึงถูกเปลี่ยนจากความชัน กับจุดตัดแกน มาเป็น ρ กับ θ แทน โดยแถวลำดับสะสมจะมีจำนวนแถวเท่ากับ $\rho_{\max}$ ดังสมการ

$$\rho_{\max} = \sqrt{\text{ความกว้างของภาพเอกสาร}^2 + \text{ความสูงของภาพเอกสาร}^2}$$

โดยที่ค่าของ θ จะมีค่าตั้งแต่ –90 ถึง 90 องศา

- **เนียเรสเนเบอร์คัสเตอร์ริง (Nearest-Neighbor Clustering)** [13][14] การประมาณค่าความเอียงของภาพเอกสารด้วยวิธีเนียเรสเนเบอร์คัสเตอร์ริง ประกอบด้วย 4 ขั้นตอน คือ หาส่วนประกอบต่อเชื่อมกัน (Connected Component) หาเนียเรสเนเบอร์ (Nearest Neighbor) สำหรับแต่ละส่วนประกอบ (Component) หาความชันของแต่ละเนียเรสเนเบอร์และหาความชันของเอกสาร แต่ละขั้นตอนมีรายละเอียดดังนี้

ขั้นตอนที่ 1 หาส่วนประกอบต่อเชื่อมกัน (Connected Component)

การหาส่วนประกอบต่อเชื่อมกัน (Connected Component) เป็นการสร้างกล่องสี่เหลี่ยมล้อมรอบกลุ่มของจุดสีดำที่อยู่ติดกันเพื่อที่จะหาจุดศูนย์ถ่วง (Centroid) ของตัวอักษร และนำศูนย์ถ่วงที่ได้ไปหาเส้นบรรทัดต่อไป ผลจากการทำส่วนประกอบต่อเชื่อมกันจะได้กรอบสี่เหลี่ยมเป็นตัวแสดงขอบเขตของแต่ละตัวอักษรในภาพเอกสาร และจะเก็บกรอบสี่เหลี่ยมนี้ให้อยู่ในรูปของ

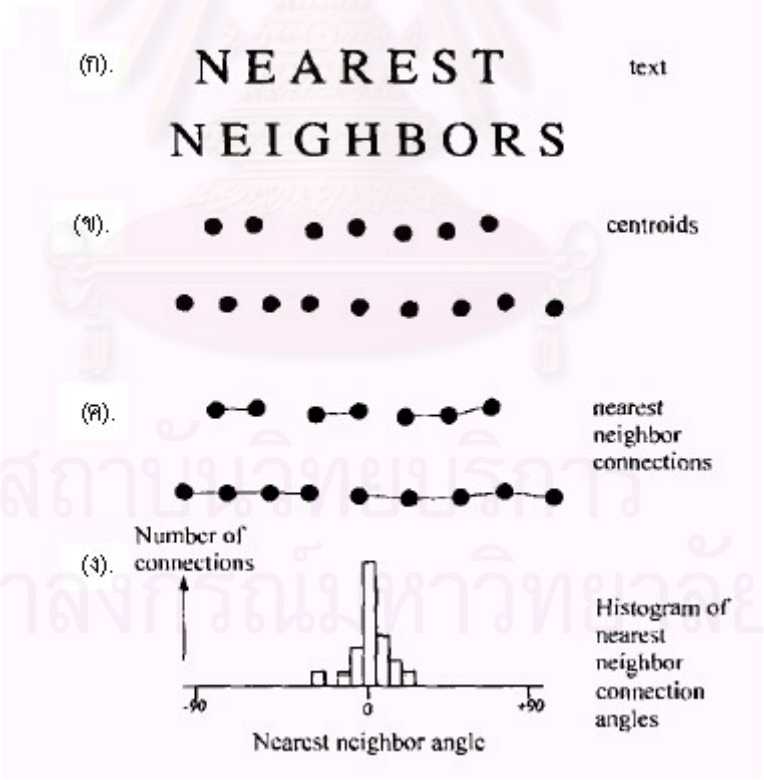
$D = \{C_1, C_2, \ldots, C_n\}$

$C_i = \{(X_{i1}, Y_{i1}), (X_{i2}, Y_{i2})\}$

โดยที่ C_i คือ กรอบสี่เหลี่ยมที่มีจุด (X_{i1}, Y_{i1}) เป็นพิกัดของมุมบนซ้ายมือของสี่เหลี่ยม และจุด (X_{i2}, Y_{i2}) เป็นพิกัดของมุมล่างขวามือของสี่เหลี่ยม

D คือเซตของกรอบสี่เหลี่ยมทั้งหมดในรูปภาพ

จากกรอบสี่เหลี่ยมที่หาได้ทำให้เราสามารถหาจุดศูนย์ถ่วงของตัวอักษรได้ดังรูปที่ 2.11



รูปที่ 2.11 วิธีการหามุมความเอียงของภาพเอกสารด้วยวิธีเนียเรสเนเบอร์

(ก). Text (ข). จุดศูนย์ถ่วงในตัวอักษรใน (ก) (ค). ส่วนเชื่อมระหว่างจุดที่ใกล้ๆ กัน

(ง). กราฟแสดงมุมเอียงต่างๆ ที่ใกล้ 0 องศา

ขั้นตอนที่ 2 หาเนียเรสเนเบอร์ (Nearest Neighbor) สำหรับแต่ละส่วนประกอบ (Component)

ก่อนอื่นเราจะทำการคัดสรรขนาดของส่วนประกอบ โดยจะหาขนาดของกรอบสี่เหลี่ยมเฉลี่ย เพื่อที่จะกำจัดรูปภาพที่จะให้กรอบสี่เหลี่ยมที่มีขนาดใหญ่และขนาดเล็กออก เช่น ตัวอักษรที่เป็นหัวข้อ ตัวอักษรที่มีขนาดเล็ก หรือเครื่องหมายต่างๆ เป็นต้น

$$\bar{c} = \left\{ \left(\frac{1}{n}\sum_{i=1}^{n} x_{i1}, \frac{1}{n}\sum_{i=1}^{n} y_{i1} \right), \left(\frac{1}{n}\sum_{i=1}^{n} x_{i2}, \frac{1}{n}\sum_{i=1}^{n} y_{i2} \right) \right\}$$

เป็นสมการที่ใช้ในการคำนวณหาขนาดของกรอบสี่เหลี่ยมเฉลี่ย ($\bar{c}$)

$$w(c_i) = |x_{i1} - x_{i2}|$$

เป็นสมการที่ใช้หาความกว้างของสี่เหลี่ยม

$$h(c_i) = |y_{i1} - y_{i2}|$$

เป็นสมการที่ใช้หาความสูงของสี่เหลี่ยม

สมมติให้ส่วนประกอบที่ผ่านการคัดสรรขนาดเป็นสมาชิกของเซต D' สมาชิกในเซตนี้จะต้องผ่านคุณสมบัติในสมการดังต่อไปนี้

$$D' = \{c_1, c_2, \ldots, c_m\}$$

$$\frac{1}{3}w(\bar{c}) < w(c_i) < 3w(\bar{c})$$

$$\frac{1}{3}h(\bar{c}) < h(c_i) < 3h(\bar{c})$$

กล่าวคือต้องมีความกว้างน้อยกว่าสามเท่าของความกว้างเฉลี่ย และมีความกว้างมากกว่าหนึ่งในสามเท่าของความกว้างเฉลี่ย รวมทั้งจะต้องมีความสูงน้อยกว่าสามเท่าของความสูงเฉลี่ย และมีความสูงมากกว่าหนึ่งในสามเท่าของความสูงเฉลี่ยอีกด้วย

เมื่อคัดสรรขนาดของส่วนประกอบเสร็จเรียบร้อยแล้ว เราจะนำส่วนประกอบที่เหลืออยู่มาหาเนียเรสเนเบอร์ คือการหาว่าส่วนประกอบใดอยู่บนเส้นบรรทัดเดียวกันบ้าง โดยเราจะนำจุดที่เป็นสมาชิกของ D' มา 1 จุด สมมติให้เป็นจุด Ci จากนั้นจะเลือกจุดที่มีระยะห่างที่สั้นที่สุดจากจุด Ci สมมติให้เป็นจุด Cj การคำนวณระยะห่างจะใช้สูตรเรขาคณิต ดังสมการ

$$e_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

หาก e_{ij} มีค่าน้อยกว่าห้าเท่าของ $w(\bar{c})$ จะถือว่าจุด Ci และ Cj อยู่บนเส้นบรรทัดเดียวกัน แต่ถ้าหาก e_{ij} มีค่ามากกว่าห้าเท่าของ $w(\bar{c})$ แปลว่าไม่มีจุดใดอยู่บนเส้นบรรทัดเดียวกันอีก และจะถือว่า Ci อยู่บนเส้นบรรทัดใหม่

ขั้นตอนที่ 3 หาความชันของแต่ละเนียเรสเนเบอร์

เมื่อได้จุดที่อยู่บนเส้นบรรทัดเดียวกัน (Nearest Neighbor) ก็นำจุดเหล่านั้นมาหาความชันโดยใช้สมการดังนี้

$$a_k = \frac{l_k \sum_{j=1}^{I_k} x_i y_i - \sum_{j=1} x_i \sum_{j=1}^{I_k} y_i}{l_k \sum_{j=1}^{I_k} x_i^2 - (\sum_{j=1}^{I_k} x_i)^2}$$

โดยที่ a_k คือ ความชันของเส้นบรรทัดที่ k

l_k คือ จำนวนจุดที่อยู่บนเส้นบรรทัดที่ k

(x_i, y_i) คือ จุดศูนย์ถ่วงของกรอบสี่เหลี่ยมของแต่ละส่วนประกอบที่อยู่บนเส้นบรรทัดที่ k

ขั้นตอนที่ 4 หาความชันของเอกสาร

จากความชันของแต่ละเนียเรสเนเบอร์ ที่หาได้นำมาหาความชันเฉลี่ย($\bar{a}$) ของภาพเอกสาร โดยใช้สมการดังนี้

$$\bar{a} = \frac{l_1 a_1 + l_2 a_2 + ... + l_k a_k}{l_1 + l_2 + ... + l_k}$$

แต่เนื่องจากจำนวนจุดของเนียเรสเนเบอร์มีผลต่อค่าที่ถูกต้องของความชัน อีกทั้งการใช้สมการหาความชันเฉลี่ยของภาพเอกสาร เพื่อหาค่าความชันเฉลี่ยจะให้ผลที่ไม่ถูกต้องหากจำนวนจุดที่อยู่บนเส้นบรรทัดเดียวกันมีค่าน้อยกว่าหรือเท่ากับหนึ่งในห้าเท่าของขนาดของไฟล์รูปภาพ ดังนั้นจึงสร้างข้อกำหนดเพิ่มเติมสำหรับการคำนวณ ดังนี้

ถ้าหากมีเส้นบรรทัดมากกว่า 1 เส้นที่มีจำนวนจุดที่อยู่บนเส้นบรรทัดเดียวกันมากกว่าหนึ่งในห้าของขนาดความกว้างของไฟล์รูปภาพ เราจะใช้สมการเพื่อหาค่าเฉลี่ยของความชันของเอกสาร

แต่ถ้าหากไม่มีเส้นบรรทัดใดเลยที่มีจำนวนจุดมากกว่าหนึ่งในห้าของขนาดความกว้างของไฟล์รูปภาพ เราจะใช้หลักของการหาฐานนิยม (Mode) คือ ความชันของเอกสารจะเป็นค่าความชันที่พบบ่อยที่สุด ยกตัวอย่างเช่น หากมีความชันแต่ละเนียเรสเนเบอร์ที่พบคือ a = {-15,-14,-10,-10,-10,-10,5,20,20} ค่าความชันของเอกสารคือ –10 องศา เพราะว่าเป็นความชันที่พบมากที่สุด

เมื่อได้ความชันของเอกสารแล้ว สามารถหามุมเอียงของเอกสารโดยใช้สมการดังนี้

$$\theta = \tan^{-1}(\bar{a})$$

2.1.1.4 **การแก้ไขมุมเอียง** (Skew Correction) การหมุนเอกสาร คือ การเปลี่ยนตำแหน่งของจุดที่อยู่บนรูปภาพที่เป็นเอกสารให้ไปอยู่ในตำแหน่งใหม่โดยตำแหน่งใหม่นี้ จะได้จากการคำนวณโดยใช้สมการทางคณิตศาสตร์

กำหนดให้จุด P(x,y) เป็นจุดในภาพเอกสารที่ยังเอียงเป็นมุม θ หลังจากที่ทำการหมุนจะได้จุดใหม่อยู่ที่ตำแหน่งจุด P'(x',y') โดยใช้สมการต่อไปนี้

$$(x',y') = (x,y)\begin{bmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{bmatrix}$$

$$= (x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta)$$

อย่างไรก็ตาม ตำแหน่งของจุดในระบบการประมวลผลภาพจะเป็นจำนวนเต็มเท่านั้น แต่เนื่องจากผลที่คำนวณได้จากสมการข้างต้น อาจทำให้ได้ค่าตำแหน่งของจุดที่ไม่ใช่จำนวนเต็ม และจะส่งผลให้ตำแหน่งของจุดที่ไม่ใช่จำนวนเต็ม และจะส่งผลให้ตำแหน่งดังกล่าวอยู่นอกขอบเขตของโดเมน (Domain) และพิสัย (Range) ในระบบการประมวลผลภาพ ดังนั้นจึงใช้ฟังก์ชันผกผัน (Inverse Function) ในการหาตำแหน่งใหม่ของจุดแทน สมการของฟังก์ชันผกผันมีดังนี้

$$(x,y) = (x',y')\begin{bmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{bmatrix}$$

$$= (x'\cos\theta + y'\sin\theta,\ -x'\sin\theta + y'\cos\theta)$$

2.1.2. การวิเคราะห์เอกสาร (Document analysis) [9]

การวิเคราะห์เอกสาร จะทำการวิเคราะห์ว่ามีตาราง รูปภาพ บรรทัดของตัวอักษร แบบของตัวอักษร ขนาดของตัวอักษร อยู่ที่ตำแหน่งใด แล้วทำการแยกตาราง รูปภาพ ออกแล้วนำส่วนที่เป็นภาพอักษรส่งไปให้ในส่วนรู้จำ การวิเคราะห์เอกสารมี ดังนี้

2.1.2.1 **การตัดแบ่งหน้าเอกสาร** (Page segmentation) [8][9] การตัดแบ่งหน้าเอกสาร เป็นวิธีการแบ่งส่วนรูปภาพของเอกสารออกเป็นบล็อกของข้อมูลที่มี คุณสมบัติเหมือนกัน วิธีการแบ่งหน้าเอกสาร โดยทั่วไปมี 3 วิธี คือ

2.1.2.1.1 วิธีการแบบบนลงล่าง (Top-Down Approach) วิธีการนี้เป็นเทคนิคของการแบ่งหน้าเอกสารออกเป็นส่วนหลักๆ ซึ่งแต่ละส่วนก็จะถูกแบ่งแยกออกเป็นส่วนย่อยๆ ลงไปอีก วิธีการนี้ต้องมีความรู้ในผังโครงสร้างของเอกสาร เพราะความรู้เกี่ยวกับเนื้องานในเอกสารจะช่วยในการวิเคราะห์โครงสร้างและรายละเอียดภายในซึ่งจะถูกแตกย่อยออกมาได้ ตัวอย่างเช่น หน้าเอกสารประกอบด้วย 3 คอลัมน์ และมีหัวเรื่องอยู่ด้านบน ระบบจะระบุลักษณะและวิเคราะห์แต่ละส่วน และทำการตัดแยกหน้าเอกสารออกเป็น ย่อหน้า บรรทัด คำ และสุดท้ายเป็นตัวอักษร

2.1.2.1.2 วิธีการแบบล่างขึ้นบน (Bottom-up Approach) โดยทั่วไปขั้นตอนการทำงานหลักๆ จะเริ่มตรวจสอบหาจุดดำในบรรทัดจากบรรทัดแรกไปบรรทัดสุดท้ายและจากซ้ายไปขวา เมื่อพบจุดดำที่ตำแหน่งใดก็จะค้นหาจุดดำที่อยู่ติดกัน (connected component) จนไม่สามารถหาจุดดำที่ติดกันได้อีกก็จะได้รูปของตัวอักษร ในแต่ละตัวอักษรจะหาจุดบริเวณใกล้เคียงในแนวแกน x เพื่อเชื่อมโยงตัวอักษรให้เป็นรูปแบบที่เป็นแถวหรือบรรทัด จากนั้นจะทำการคำนวณช่องว่างระหว่างตัวอักษรในแนวแกน x และอาศัยค่าที่คำนวณได้นี้ ใช้ในการแบ่งแยกบรรทัดไปเป็นคอลัมน์ สุดท้ายจะคำนวณภายในบรรทัดและช่องว่างระหว่างบรรทัด เพื่อรวมบรรทัดเข้าด้วยกันภายในคอลัมน์ไปเป็นรูปแบบของบล็อก

2.1.2.1.3 วิธีการแบบไฮบริด หรือ ลูกผสม (Hybrid Approach) จากการที่เอกสารบางประเภท ไม่สามารถทำการตัดแบ่งหน้าเอกสารโดยวิธีการแบบบนลงล่าง หรือ กลยุทธ์แบบล่างขึ้นบน แบบใดแบบหนึ่งได้ จึงเกิดวิธีการแบบไฮบริด เข้ามาช่วย ซึ่งเป็นการใช้วิธีการทำงานของทั้ง 2 วิธีข้างต้นเข้ามาช่วยร่วมกันทำงาน

2.1.2.2 **การแบ่งแยกประเภทบล็อก (Block Classification)** บล็อกของข้อมูลที่ได้จากการตัดแบ่งหน้าเอกสาร จะต้องนำมาแบ่งแยกประเภทที่เราได้ กำหนดประเภทของเอกสารไว้ล่วงหน้าแล้ว โดยจะกำหนดฉลาก (Label) ให้กับบล็อกทุกบล็อกซึ่งอาจจะเป็น ชื่อเรื่อง ข้อความ รูปภาพ ตาราง เป็นต้น ถ้าเรารู้ผังโครงสร้างของเอกสารแล้วจะช่วยในการทำงานนี้ได้ ตัวอย่างเช่น ถ้าเรารู้ก่อนว่าเอกสารเป็นใบปะหน้าแฟกซ์ แล้วเราจะทราบว่าบล็อกที่อยู่ตรงกลางจะเป็นฉลากของที่อยู่ปลายทาง และบนสุดของเอกสารจะต้องเป็นฉลากของชื่อหน่วยงาน เป็นต้น

2.1.3 ขั้นตอนประมวลผลการรู้จำ (OCR-Processing)

ขั้นตอนประมวลผลการรู้จำตัวอักษรแบ่งออกเป็น 2 ขั้นตอน คือ

2.1.3.1 **การตัดแยกตัวอักษร (Character Segmentation)** เป็นขั้นตอนใช้แยกรูปของตัวอักษรออกจากรูปของเอกสาร โดยมีขั้นตอนย่อยอีก 2 ขั้นตอน เพื่อส่งให้กับขั้นตอนที่ 2.1.3.2 ต่อไป คือ

2.1.3.1.1. การวิเคราะห์โครงสร้างของเอกสาร (Format Analysis) เป็นขั้นตอนในการวิเคราะห์โครงสร้างของเอกสาร เช่น ส่วนใดควรจะเป็นบรรทัดเดียวกัน ส่วนใดเป็นย่อหน้า เป็นต้น แล้วส่งผลลัพธ์ที่เป็นรูปของบรรทัดไปยังขั้นต่อไป วิธีการที่ใช้ เช่น การตรวจสอบค่าฮีสโตแกรมของภาพในแนวนอน เพื่อหาพื้นที่ที่มีความน่าจะเป็นสูงที่จะเป็นช่องว่างระหว่างบรรทัด

2.1.3.1.2. การแยกรูปตัวอักษร (Character Segmentation) เป็นขั้นตอนแยกรูปภาพตัวอักษร (Character Image) ออกจากรูปภาพของบรรทัดที่ได้จากขั้นตอนการวิเคราะห์โครง

สร้างของเอกสาร มักใช้วิธีการล้อมกรอบตัวอักษร (Bounding Box Analysis) เพื่อกำหนดขอบเขตของตัวอักษรแต่ละตัว

2.1.3.2 การรู้จำตัวอักษร (Character Recognition) เป็นขั้นตอนที่รู้จำตัวอักษรโดยใช้รูปของตัวอักษรที่ได้รับจากขั้นตอนการตัดแยกตัวอักษร เพื่อให้ได้รหัสของตัวอักษร จะพบว่าข้อมูลนำเข้าคือรูปภาพของตัวอักษร (Character Image) ซึ่งได้รับมาจากการตัดแยกตัวอักษร และมีขั้นตอนย่อยอีก 2 ขั้นตอน คือ

2.1.3.2.1. การแยกลักษณะสำคัญของตัวอักษร (Feature Extraction) เป็นขั้นตอนย่อยที่ใช้ในการแยกลักษณะสำคัญของตัวอักษร เพื่อให้ได้คุณสมบัติเฉพาะของตัวอักษรนั้น

2.1.3.2.2. การจำแนกภาพ (Classification) เป็นขั้นตอนที่นำเอาคุณสมบัติเฉพาะของตัวอักษรมาพิจารณาและจำแนกตัวอักษรออกตามประเภทและได้ผลลัพธ์คือ รหัสของตัวอักษร (Character Code)

2.2. เอกสารและงานวิจัยที่เกี่ยวข้อง

2.2.1. Adaptive document image binarization โดย J. Sauvola และ M. Pietikäinen [5] ได้นำเสนออัลกอริทึมในการพิจารณาภาพเอกสารที่ประกอบไปด้วย ตัวอักษร พื้นหลังและภาพกราฟิก โดยใช้เทคนิคในการปรับค่าขีดแบ่งไปตามสถานการณ์ ในงานวิจัยนี้ได้ใช้อัลกอริทึมของไนแบล็ก (Niblack's Algorithm) ในการทำไบนาไรเซชันสำหรับภาพเอกสารที่เป็นส่วนของตัวอักษร และได้มีการปรับปรุงอัลอกริทึมให้มีประสิทธิภาพในการทำงานที่ดีขึ้น ในการแยกส่วนของพื้นหลังออกจากส่วนของตัวอักษร

2.2.2. Evaluation of binarization methods for document image โดย Øivind Due Trier และ Torfinn Taxt [6] ซึ่งได้นำเสนอการประเมินประสิทธิภาพการทำงานของอัลกอริทึมในการทำไบนาไรเซชันจำนวน 11 อัลกอริทึมว่ามีประสิทธิภาพเป็นอย่างไร จากผลลัพธ์ที่ได้ก็คืออัลกอริทึมที่ดีที่สุดคืออัลกอริทึมของไนแบล็ก ซึ่งเป็นอัลกอริทึมที่ได้มีการปรับปรุงให้มีประสิทธิภาพที่ดีขึ้น โดยใช้ขั้นตอนหลังการประมวลผล (Postprocessing Step) มาใช้ประกอบกับอัลกอริทึมของไนแบล็ก เพื่อให้ประสิทธิภาพของการทำไบนาไรเซชันได้ผลที่ดีขึ้น

2.2.3. A new method for image segmentation โดย S. D. Yanowitz และ A. M. Bruckstein [7] นำเสนออัลกอริทึมในการจำแนกภาพโดยใช้การจำแนกจากขอบรอยต่อของภาพ (Edge) จากค่าขีดแบ่งของบริเวณ (threshold surface) และค่าขีดแบ่งนี้จะได้มาจากการส่งค่าโพเทนเชิลเซอฟิส (Potential Surface) ไปที่ตำแหน่งของกราเดียนท์ที่มีค่าสูงสุด (Maxima of the

gradient) ของบริเวณนั้น ตำแหน่งของกราเดียนท์ที่มีค่าสูงสุดของบริเวณใดๆ จะคำนวณได้โดยใช้วิธีการของโซเบิ้ลหรือโซเบิ้ลเอจโอเปอร์เรเตอร์ (Sobel's edge operator) [14] เป็นต้น

2.2.4. การประมาณค่าความเอียงหน้าเอกสารสิ่งพิมพ์ โดย ยุทธพงษ์ รังสรรค์เสรี และกฤษณะ ชินสาร [3] ได้นำเสนอวิธีการในการหาค่ามุมความเอียงของภาพเอกสาร เป็นการหาค่าตำแหน่งของจุดภาพสีดำทางซ้ายสุดของแต่ละบรรทัดภาพเอกสารทั้งเอกสาร แล้วแบ่งบรรทัดภาพเอกสารออกเป็นบรรทัดตัวหนังสือ โดยพิจารณาจากความสัมพันธ์ของจุดภาพทางซ้ายสุดของบรรทัดที่อยู่ติดกัน ทำการรวมบรรทัดตัวหนังสือเป็นย่อหน้าของบรรทัดตัวหนังสือ เลือกย่อหน้าที่มีจำนวนบรรทัดภาพเอกสารมากที่สุด เป็นตัวแทนในการประมาณค่ามุมเอียงของหน้าเอกสาร

2.2.5. A document skew detection method using the Hough Transform โดย A. Amin และ S. Fischer [11] ได้นำเสนอวิธีการหาค่ามุมความเอียงของภาพเอกสาร โดยอาศัยหลักการของการวิเคราะห์ส่วนประกอบต่อเชื่อมกัน (Connected Component Analysis) ในการกำหนดบริเวณ (Region) ของจุดภาพสีดำที่ติดๆ กันให้เป็นบล็อกสี่เหลี่ยมแนวแบ่งเขต (Bounding Rectangle) แล้วนำบล็อกที่ได้มาพยายามรวมเป็นกลุ่ม จากนั้นหาจุดกึ่งกลางของกลุ่ม โดยจะมุ่งไปที่บล็อกที่ล่างสุดของกลุ่มแล้วนำมาคำนวณหามุมเอียงโดยใช้ฮัจทรานสฟอร์ม (Hough Transform) วิธีการนี้จะช่วยในการลดเวลาในการคำนวณได้ เพราะเนื่องจากจุดที่จะนำมาคำนวณนั้นน้อยลง ดังนั้นวิธีการนี้จะใช้เวลาในการประมวลผลน้อย

2.2.6. Improved document skew detection based on text line connected-component clustering โดย N. Liolios, N. Fakotokis และ G. Kokkinakis [14] ได้นำเสนอวิธีการหาค่ามุมความเอียงของภาพเอกสาร โดยอาศัยหลักการของการวิเคราะห์ส่วนประกอบต่อเชื่อมกันร่วมกับเนียเรสเนเบอร์คัสเตอร์ริง (Nearest-Neighbor Clustering) ในการความเอียงของภาพเอกสาร ผลของงานวิจัยได้บอกวิธีเนียเรสเนเบอร์คัสเตอร์ริง (Nearest-Neighbor Clustering) ในการหาค่าความเอียงนี้จะมีความถูกต้องสูงมากขึ้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

วิธีการดำเนินงานวิจัย

สำหรับงานวิจัยนี้ว่าด้วยเรื่องการเตรียมภาพเอกสารก่อนที่ไปขบวนการรู้จำตัวอักษร โดยมี 2 ส่วนคือ

1. ไบนาไรเซชัน (Binarization)
2. การปรับมุมเอียงของภาพเอกสารแบบอัตโนมัติ

3.1. ไบนาไรเซชัน

ไบนาไรเซชันเป็นการแปลงภาพเอกสารที่มีระดับความเข้มของสี 256 ระดับให้เหลือ 2 ระดับ คือสีขาวและสีดำเท่านั้น โดยโปรแกรมจะต้องหาค่าขีดแบ่ง (Threshold) เพื่อมาใช้ในการเปรียบเทียบกับระดับความเข้มของจุดภาพ จะมีค่ามากกว่าหรือน้อยกว่าค่าขีดแบ่งที่กำหนดขึ้นมา โดยหากค่าระดับความเข้มของจุดภาพมีค่าน้อยกว่าค่าขีดแบ่ง เช่นหากภาพมีจุดภาพสีขาวเป็นสีพื้นหลังของภาพเป็นส่วนใหญ่ และจุดภาพสีดำเป็นจุดภาพของวัตถุ เมื่อเปรียบเทียบระดับความเข้มของจุดภาพมีค่ามากกว่าค่าขีดแบ่งจะให้จุดภาพนั้นเป็นสีดำ และหากว่าน้อยกว่าค่าขีดแบ่งให้เป็นจุดภาพสีขาว การกำหนดค่าขีดแบ่งนี้มี 2 แบบคือแบบกำหนดค่าขีดแบ่งแบบครอบคลุม (Global Threshold) และแบบกำหนดค่าขีดแบ่งแบบปรับค่าได้ (Adaptive Threshold) ในงานวิจัยนี้จะศึกษาในกำหนดค่าขีดแบ่งแบบปรับค่าได้ ด้วยอัลกอริทึมของไนแบล็กและอัลกอริทึมของนาโนวิทซ์และบรัคสเตียน

3.1.1. อัลกอริทึมของไนแบล็ก

ในแนวคิดจะเป็นกำหนดให้มีค่าขีดแบ่งหลายๆ จุดบนภาพ โดยค่าของขีดแบ่งนี้จะพิจารณาจากพื้นที่ที่กำลังพิจารณาเป็นบริเวณย่อยๆ หรือเรียกว่าหน้าต่าง (Window) ขนาดของหน้าต่างเราสามารถกำหนดได้ว่าจะให้ขนาดของหน้าต่างมีขนาด 7*7, 9*9, 11*11, 13*13, 15*15, 21*21 หรือ 23*23 ขนาดใดก็ได้ โดยมีจุดที่สนใจจะอยู่ตรงกลางของหน้าต่างนำหน้าต่างที่ได้มาคำนวณหาค่าเฉลี่ยท้องถิ่น (Local Mean) และค่าส่วนเบี่ยงเบนมาตรฐานท้องถิ่น (Local standard deviation) เพื่อจะให้ได้ค่าขีดแบ่งหรือ T(x,y) คำนวณได้จากสูตร

$$T(x,y) = m(x,y) + k * s(x,y)$$

โดยที่ m(x,y) และ s(x,y) หมายถึงค่าเฉลี่ยท้องถิ่นและค่าส่วนเบี่ยงเบนมาตรฐานท้องถิ่น ตามลำดับในพื้นที่รอบข้างของจุดภาพ (x,y) ขนาดของพื้นที่รอบข้างของจุดภาพ (x,y)และค่าของ k จะใช้เป็นเงื่อนไขที่สำคัญในการปรับแต่ง (adjust) เพื่อให้ได้ส่วนที่เป็นตัวอักษรออกมาให้มีประสิทธิภาพที่ดี

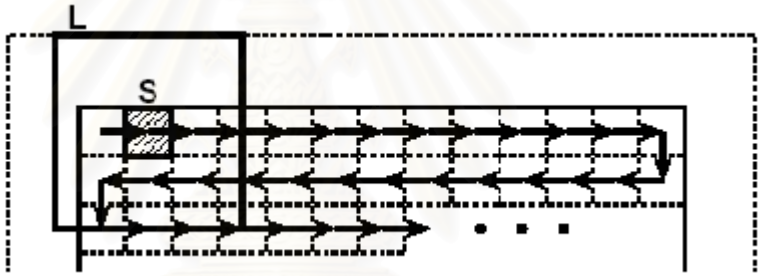
ค่าเฉลี่ยท้องถิ่นหาได้จากสูตร

$$m(x,y) = \frac{1}{n}\sum_{x,y} g(x,y)$$

และค่าส่วนเบี่ยงเบนมาตรฐานท้องถิ่น (Local standard deviation) จะได้มาจาก

$$s(x,y) = \sqrt{\frac{1}{n}\sum_{x,y}((g(x,y) - m(x,y))^2}$$

โดยที่ n หมายถึงจำนวนของจุดภาพและ $g(x,y)$ หมายถึงค่าความความเข้มของจุดภาพที่ตำแหน่ง x,y

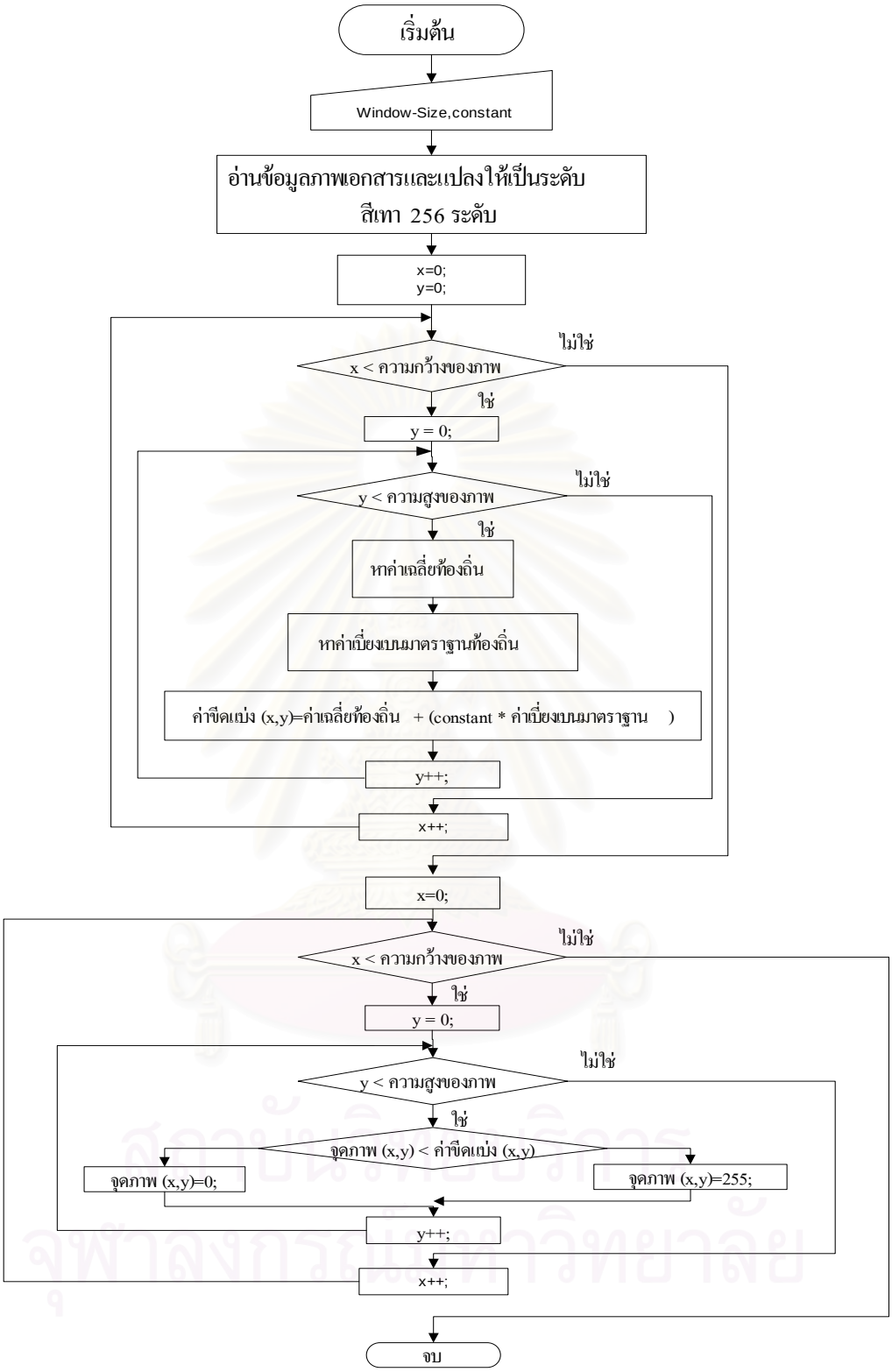


รูปที่ 3.1 การเลื่อนของหน้าต่าง

ในรูปที่ 3.1 โดยที่ s แทนจุดภาพที่กำลังพิจารณาและ L แทนหน้าต่างที่มาครอบจุดภาพ เมื่อทำการคำนวณแล้วหน้าต่างก็จะเลื่อนไปทางซ้ายจนสุด แล้วจึงเลื่อนลงเริ่มจากทางขวาสุดใหม่ จนกระทั่งเลื่อนลงไปจุดภาพแถวสุดท้ายและซ้ายสุด ครบทั้งภาพ เราก็จะได้ค่าขีดแบ่งที่คำนวณได้มาจากจุดภาพข้างเคียงที่อยู่ภายในหน้าต่าง มาทำการเปรียบเทียบกับจุดภาพต้นฉบับที่ละจุด เราสามารถเขียนเป็นผังการทำงานอัลกอริทึมของไนแบล็กได้ดังรูปที่ 3.2

จากสมการในการหาค่าขีดแบ่งของไนแบล็กตัวแปรที่สำคัญอีกตัวหนึ่งคือ k จะมีค่าระหว่าง 0.18 ถึง 0.2 ในกรณีที่ภาพเอกสารมีสีพื้นหลังเป็นสีดำ และ เป็น -0.18 ถึง -0.2 ในกรณีสีพื้นหลังเป็นสีขาว

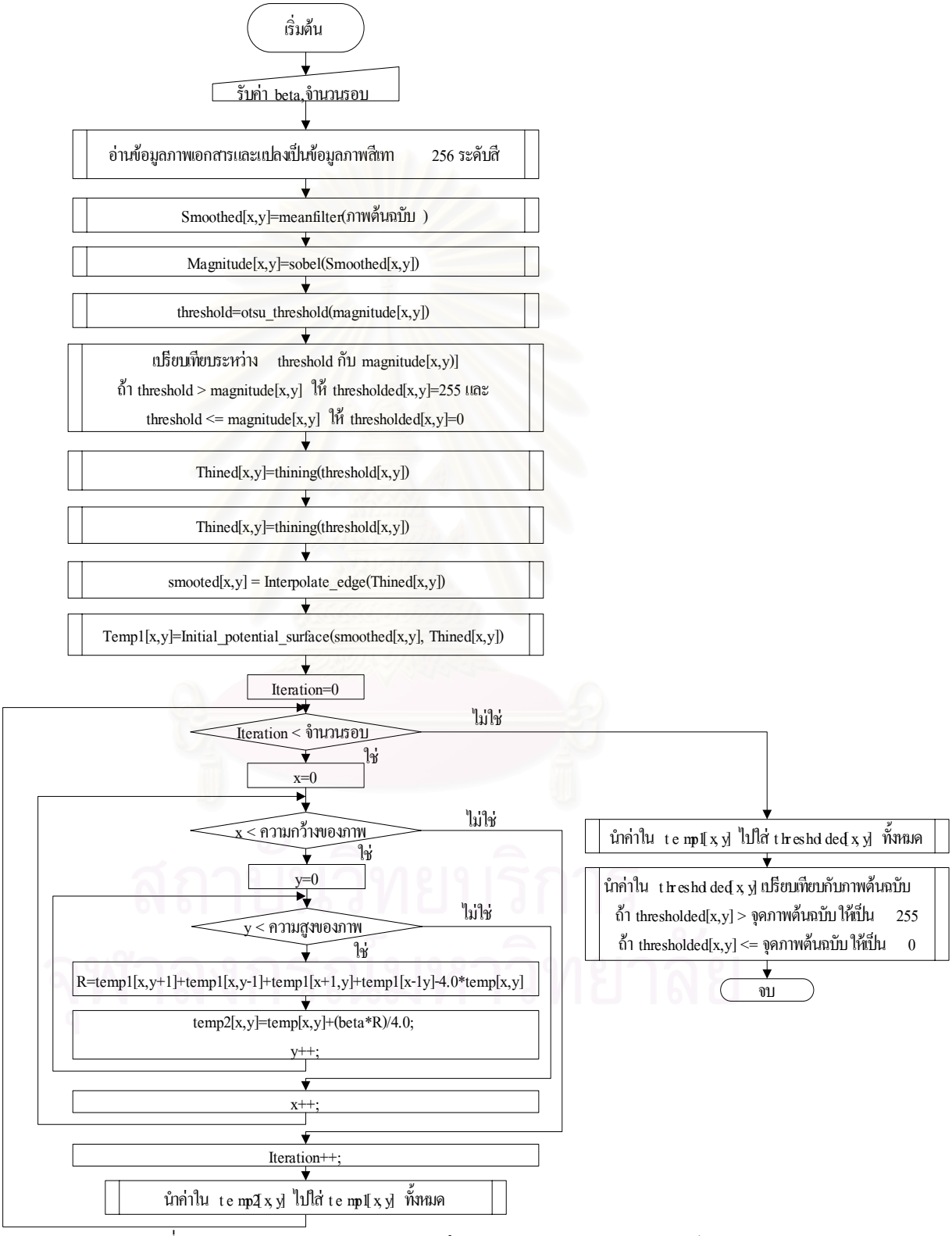
ในการวิจัยได้ทดลองกับภาพเอกสารที่มีพื้นหลังเป็นสีขาว โดยจะเปลี่ยนขนาดของหน้าต่างในการหาค่าขีดแบ่งจะได้ขนาดของหน้าต่างที่ดีที่สุด คือขนาด 15*15 และค่าของตัวแปรคงที่จะเท่ากับ -0.2



รูปที่ 3.2 ผังการทำงานของอัลกอริทึมของไนแบล็ก

3.1.2 อัลกอริทึมของยาโนวิทซ์และบรัคสเตียน

จากรูปที่ 3.3 เป็นการแสดงผังของการทำงานตามอัลกอริทึมของยาโนวิทซ์และบรัคสเตียน



รูปที่ 3.3 ผังการทำงานของอัลกอริทึมของยาโนวิทซ์และบรัคสเตียน

อัลกอริทึมของยาโนวิทซ์และบรัคสเตียนจะเป็นการหาขีดแบ่งพื้นผิว (Threshold Surface) โดยวิธีการนำภาพเอกสารมาจำแนกหาขอบภาพ โดยใช้อนุพันธ์เชิงเส้นอันดับที่หนึ่ง แล้วนำไปผ่านขบวนการทินนิง (Thining) เพื่อจะได้ค่าของบริเวณพื้นผิวที่มีค่าสูงสุด โดยจะอธิบายได้ ดังนี้

ขั้นตอนที่ 1 เตรียมภาพเอกสารต้นฉบับ โดยจะนำภาพเอกสารต้นฉบับไปผ่านขบวนการกรองความถี่แบบหาค่าเฉลี่ย (Mean Filter) เพื่อให้ภาพต้นฉบับมีความราบเรียบ แล้วนำผลที่ได้เก็บไว้ในแถวลำดับ (Array) ที่ชื่อว่า สมูท (Smoothed)

ขั้นตอนที่ 2 นำสมูทไปหาขอบภาพ ด้วยการใช้โซเบิลโอเปอเรเตอร์ (Sobel's Operator) ซึ่งเป็นการใช้อนุพันธ์เชิงเส้นอันดับที่ 1 ในการหาขอบภาพ แล้วนำผลที่ได้เก็บไว้ในแถวลำดับ ที่ชื่อว่า เม็กนิจูด (Magnitude) ดังรูปที่ 3.4. ข)

ขั้นตอนที่ 3 นำเม็กนิจูด ที่ได้ไปหาค่าขีดแบ่ง โดยจะใช้การหาค่าขีดแบ่งแบบครอบคลุมตามอัลกอริทึมของอ๊อดสึ (Otsu's Algorithm) เมื่อได้ค่าขีดแบ่งแล้วนำค่าที่ได้ไปเปรียบเทียบกับเม็กนิจูด นำผลที่ได้ไปเก็บไว้ในแถวลำดับ ชื่อว่า เทสโซ (Threshold)

ขั้นตอนที่ 4 นำเทสโซ มาทำทินนิงเพื่อหาโครงร่าง (Skeleton) ของภาพ นำผลที่ได้ไปเก็บไว้ในแถวลำดับ ชื่อว่า ทินนิง (Thining)

ขั้นตอนที่ 5 นำขอบภาพของทินนิงมาทำการแทรกเข้าไปในขอบภาพของสมูท ดังรูปที่ 3.4 ค) จะได้จุดที่ต้องแทรกจุดภาพ (Interpolation Point)

ขันตอนที่ 6 เป็นกระบวนการของการทำซ้ำ จำนวนครั้งของการทำซ้ำจะขึ้นอยู่กับผู้ใช้เป็นผู้กำหนด กระบวนการนี้จะเป็นการหาค่าขีดแบ่งพื้นผิว โดยการนำเอาสมูทและทินนิงไปหาค่าขีดแบ่งพื้นผิว จากสมการ

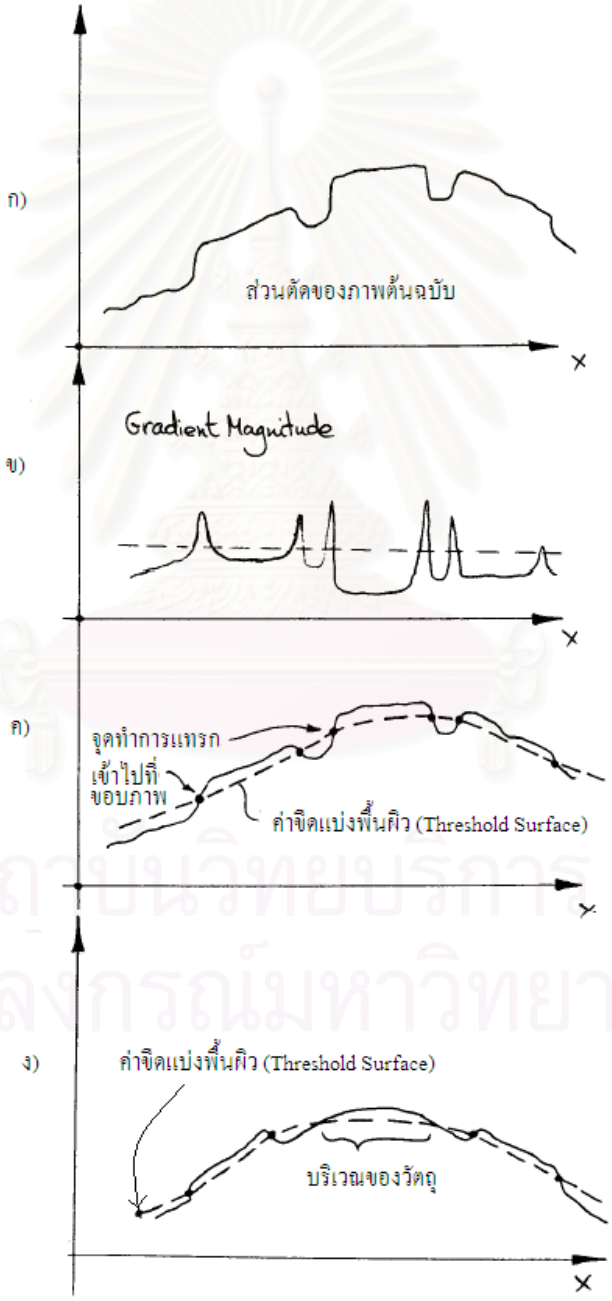
$$P_{n+1}(x,y) = P_n(x,y) + \frac{\beta * R_n(x,y)}{4}$$

$$R_n(x,y) = P_n(x-1,y) + P_n(x+1,y) + P_n(x,y+1) + P_n(x,y-1) - 4P_n(x,y)$$

โดยที่ P(x,y) หรือโพเทนเชิลเซอฟิส (Potential Surface) ได้มาจากการทำซ้ำ (Iterative) ในเค้าร่างของการประมาณค่าในช่วง (Interpolation Scheme) การประมาณค่าในช่วงของพื้นผิวคือการส่งค่าไปยังจุดภาพขอบภาพ และการประมาณค่าในช่วงตกค้าง (An Interpolation Residual) หรือ $R_n(x,y)$ คือค่าที่คำนวณได้มาจากจุดภาพในส่วนที่ไม่ใช่ขอบภาพ จะนำค่า P(x,y) ไปเก็บไว้ในแถวลำดับเทสโซเซอเฟส

จากยาโนวิทซ์และบรัคสเตียนได้ให้คำแนะนำไว้ว่าค่าของ β ควรจะมีค่าอยู่ระหว่าง $1.0 < \beta < 2.0$ จะทำให้การลู่เข้า (Convergence) รวดเร็วขึ้น อย่างไรก็ตามจากผลการทดลอง จะเห็นได้ว่าเมื่อค่าของ $\beta > 1.0$ ผลลัพธ์ที่ได้เริ่มมีความผิดพลาดไป ดังนั้นควรเลือกค่าของ $\beta = 1.0$

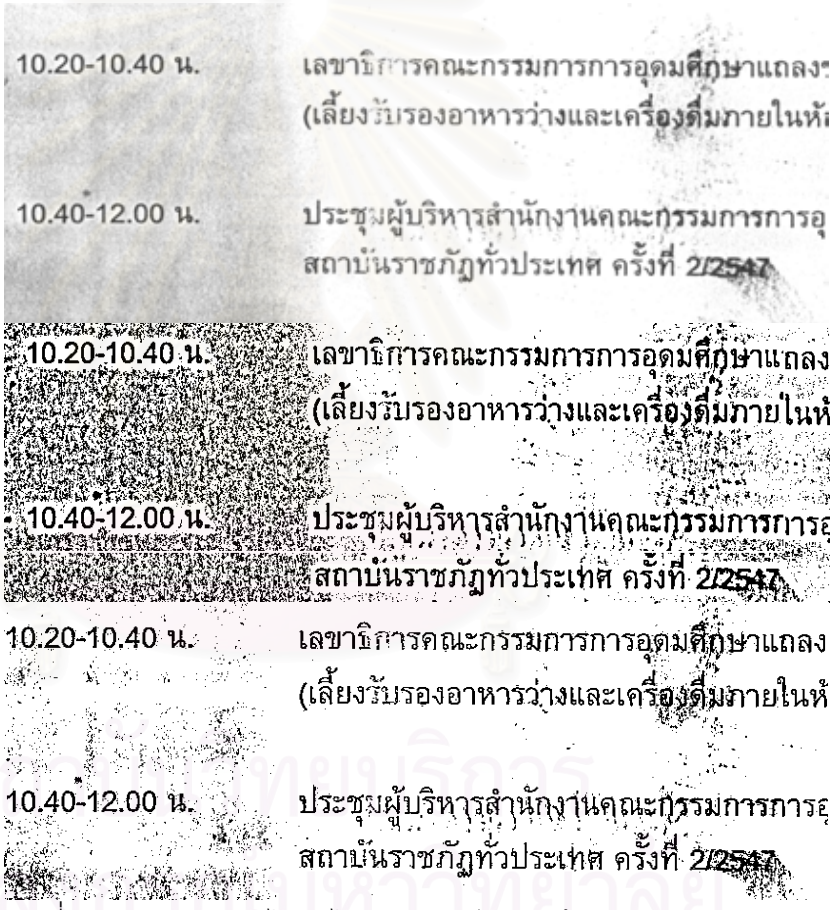
ขั้นตอนที่ 7 ค่าเทสโซเซอเฟสจะเป็นค่าขีดแบ่งที่ได้มา ที่จะนำมาเปรียบเทียบกับค่าจุดภาพต้นฉบับ หากค่าของจุดภาพต้นฉบับมีค่าน้อยกว่าค่าขีดแบ่ง จะทำให้จุดภาพต้นฉบับนั้นเป็นสีดำ ในทำนองเดียวกัน หากจุดภาพนั้นมีค่ามากกว่าจุดภาพต้นฉบับ จะทำให้จุดภาพต้นฉบับเป็นสีขาว



รูปที่ 3.4 การทำงานของการหาขีดแบ่งพื้นผิว

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 3.5 เป็นการแสดงภาพเอกสารที่ผ่านการทำไบนาไรเซชัน รูปที่ 3.5 ก) แสดงภาพเอกสารต้นฉบับ รูปที่ 3.5 ข) แสดงภาพเอกสารที่ผ่านการทำไบนาไรเซชันด้วยอัลกอริทึมของไนแบล็ก และรูปที่ 3.5 ค) ภาพเอกสารที่ผ่านการทำไบนาไรเซชันด้วยอัลกอริทึมของยาโนวิทซ์และบรัคสเตียน จะเห็นว่าภาพที่ผ่านอัลกอริทึมของยาโนวิทซ์และบรัคสเตียนจะมีจุดภาพรบกวนน้อยกว่าภาพของอัลกอริทึมของไนแบล็ก



ก)

ข)

ค)

รูปที่ 3.5 ภาพผ่านการไบนาไรเซชัน ก) เป็นภาพต้นฉบับ ข) ภาพผ่านอัลกอริทึมของไนแบล็ก ค) ภาพผ่านอัลกอริทึมของยาโนวิทซ์และบรัคสเตียน

3.2 การปรับมุมเอียงของภาพเอกสารแบบอัตโนมัติ

โปรแกรมการปรับมุมเอียงของภาพเอกสารแบบอัตโนมัติ ช่วยในการปรับมุมเอียงของเอกสารให้มีความถูกต้องมากขึ้น ในงานวิจัยนี้ใช้อัลกอริทึมในการหามุมเอียงของเอกสารหลักๆ เพียง 2 อัลกอริทึมได้แก่ฮัจทรานสฟอร์มและเนียเรสเนเบอร์คัสเตอร์ริง

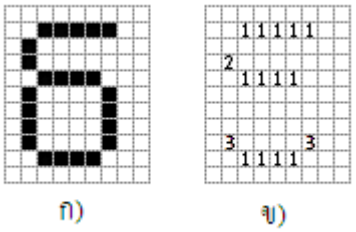
3.2.1. ฮัจทรานสฟอร์ม

นำเอกสารภาพที่ได้ซึ่งความเข้มของสีเป็น 256 ระดับสีเทา แปลงให้เป็นภาพเอกสารไบนาไรเซชันเพื่อหาค่าขีดแบ่ง โดยจะกำหนดค่าขีดแบ่งแบบครอบคลุมตามอัลกอริทึมของอ๊อดสึ นำภาพที่ผ่านการไบนาไรเซชัน โดยจะนำจุดภาพที่ผ่านเป็นสีดำมาคำนวณที่สมการเส้นตรง เพื่อจะหา ρ คือเส้นตรงที่ลากจากจุด (0,0) มาตั้งฉากกับเส้นตรงที่เราสนใจ และ θ คือมุมที่ทำกับแกน x โดยมีแถวลำดับตัวสะสม (Accumulator array) ซึ่งจะใช้เก็บสะสมค่าของ ρ และ θ ดังจะเห็นในอัลกอริทึมตารางที่ 3.1

ตารางที่ 3.1 อัลกอริทึมพื้นฐานสำหรับฮัจทรานสฟอร์ม

```
For (x)
    For (y)
    If (pixel is black)
    {     For (θ)
          {     Calculate ρ = x cos θ + y sin θ
                Increment Accumulator Array [ρ, θ]
          }
    }
```

หากทำการคำนวณตามอัลกอริทึมตารางที่ 3.1 จะต้องเสียเวลา เนื่องจากมีข้อมูลจำนวนมาก ดังนั้นจึงต้องทำการลดจำนวนข้อมูลลงโดยการใช้การเข้ารหัส Run-Length Encoding [12] เป็นวิธีที่ใช้ลดขนาดข้อมูลของไฟล์รูปภาพ โดยแทนที่จะเก็บข้อมูลตำแหน่งของจุดที่เป็นสีดำ ดังรูปที่ 3.6 ก) แต่จะเก็บจำนวนจุดสีดำที่อยู่ติดกัน โดยจะนับจุดสีดำที่อยู่ติดกันในแนวตั้งจากบนลงล่าง และจะใส่จำนวนจุดที่นับได้ลงไปในตำแหน่งสุดท้ายที่พบ ดังรูปที่ 3.6 ข) เป็นการลดความสำคัญของ noise และให้ความสำคัญกับจุดที่อยู่บนเส้นบรรทัดมากขึ้น แต่การทำ Run-Length Encoding แบบนี้ก็อาจเกิดข้อผิดพลาดได้ในกรณีที่มีรูปภาพอยู่ในเอกสาร เพราะรูปภาพนี้จะประกอบด้วยจุดสีดำที่อยู่ติดกันเป็นจำนวนมาก



รูปที่ 3.6 อธิบายการเข้ารหัส Run-Length Encoding

ดังนั้นลดข้อผิดพลาดให้คำนวณฮัจทรานสฟอร์มกับจุดซึ่งมีค่า Run-Length ระหว่าง 1 และ 25 (1/75 และ1/3 นิ้ว) ในการหามุมของภาพเอกสารจะนำแถวลำดับตัวสะสมที่มีค่ามากกว่าหนึ่งในสี่เท่าของความกว้างของไฟล์รูปภาพมาหาค่าฐานนิยม ซึ่งจะให้ค่ามุมเอียงของเอกสาร ทั้งนี้ค่าหนึ่งในสี่เท่าของความกว้างของไฟล์รูปภาพเกิดจากการที่เราต้องการกำจัดจุดภาพรบกวนออกไปมากที่สุด มุมเอียงของภาพเอกสารจะต้องเป็นมุมที่มีจุดมาตกที่มุมนั้นมากๆ จึงต้องมีการกำจัดออกไป ดังในตารางที่ 3.2

ตารางที่ 3.2 อัลกอริทึมฮัจทรานสฟอร์มแบบใช้ Run-Length Encoding

```
angle ={ } // Set of  possible angle initially empty
For x = 0 to bitmap width
{       For y = 0 to bitmap height
        {       Run_Length_Image = Find_Run_Length_Encoding()   }
}
For x = 0 to bitmap_width
{       For y = 0 to bitmap_height
        {       if  pixel in Run_Length_Image <>0 then
                {       For θ = -90 to 90
                        {       Calculate  ρ = x cos θ + y sin θ
                                Accumulator Array [ ρ , θ ] =
                                        Accumulator Array [ ρ , θ ] +
                                        Run_Length_Image [ x , y ]
                        }
                }
        }
}
```

For ρ = 1 to $\rho_{\max}$

{ For θ = -90 to 90

{ if $Accumulato\ rArray[\rho, \theta] \geq \frac{bitmap\ _width}{4}$ then

Add_Possible_Angle(θ, $angle$)

}

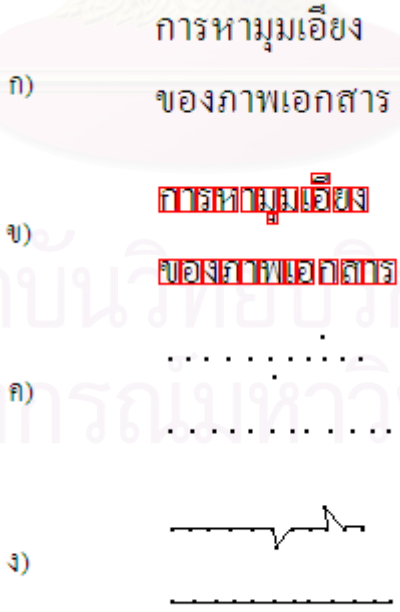
}

SkewAngle = Mode ($angle$)

Rotate_Document(SkewAngle)

3.2.2 เนียเรสเนเบอร์คัสเตอร์ริง

นำเอกสารภาพที่ได้ซึ่งความเข้มของสีเป็น 256 ระดับสีเทา แปลงให้เป็นภาพเอกสารไบนาไรเซชันเพื่อหาค่าขีดแบ่ง โดยจะให้การกำหนดค่าขีดแบ่งแบบครอบคลุมตามอัลกอริทึมของอ๊อดสึ นำภาพที่ผ่านการไบนาไรเซชัน หาส่วนประกอบต่อเชื่อมกัน (Connected Component) ของจุดภาพสีดำสร้างเป็นกล่องสี่เหลี่ยมล้อมรอบกลุ่มของจุดสีดำที่อยู่ติดกัน เพื่อที่จะหาจุดศูนย์ถ่วง (Centroid) ของตัวอักษร จากรูปที่ 3.7 ก) แสดงตัวอักษรที่เก็บอยู่ในภาพ หลังจากนั้นหาส่วนประกอบต่อเชื่อมกันของจุดภาพเพื่อสร้างกล่องสี่เหลี่ยม(c_i) ดังรูปที่ 3.7 ข)



รูปที่ 3.7 การทำเนียเรสเนเบอร์คัสเตอร์ริง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

กล่องสี่เหลี่ยมต่างๆ จะเป็นสมาชิกของเซต D จากนั้นนำกล่องสี่เหลี่ยมที่ได้มาหาจุดศูนย์ถ่วงได้ดังรูปที่ 3.7 ค) เขียนในรูปสมการดังนี้

$$D = \{c_1, c_2, ..., c_n\}$$

$$c_i = \{(x_{i1}, y_{i1}), (x_{i2}, y_{i2})\}$$

โดยที่ c_i คือ กรอบสี่เหลี่ยมที่มีจุด (x_{i1}, y_{i1}) เป็นพิกัดของมุมบนซ้ายมือของสี่เหลี่ยม และจุด (x_{i2}, y_{i2}) เป็นพิกัดของมุมล่างขวามือของสี่เหลี่ยม

D คือเซตของกรอบสี่เหลี่ยมทั้งหมดในรูปภาพ

ขั้นตอนต่อมาจะทำการคัดสรรขนาดของส่วนประกอบต่อเชื่อมกัน โดยจะหาขนาดของกรอบสี่เหลี่ยมเฉลี่ย เพื่อที่จะกำจัดรูปภาพที่จะให้กรอบสี่เหลี่ยมที่มีขนาดใหญ่, ตัวอักษรที่มีขนาดใหญ่ เช่น ตัวอักษรที่เป็นหัวข้อ, ตัวอักษรที่มีขนาดเล็ก จากสูตรที่ใช้ในการคำนวณหาขนาดของกรอบสี่เหลี่ยมเฉลี่ย ($\bar{c}$) ดังต่อไปนี้

$$\bar{c} = \left\{\left(\frac{1}{n}\sum_{i=1}^{n} x_{i1}, \frac{1}{n}\sum_{i=1}^{n} y_{i1}\right), \left(\frac{1}{n}\sum_{i=1}^{n} x_{i2}, \frac{1}{n}\sum_{i=1}^{n} y_{i2}\right)\right\}$$

นอกจากที่จะหาขนาดต้องหาความกว้างของสี่เหลี่ยมจากสมการ

$$w(c_i) = |x_{i1} - x_{i2}|$$

และความสูงเฉลี่ยจากสมการ

$$h(c_i) = |y_{i1} - y_{i2}|$$

เมื่อส่วนประกอบต่อเชื่อมกันที่ผ่านการคัดสรรขนาดเป็นสมาชิกของเซต D' สมาชิกในเซตนี้จะต้องผ่านคุณสมบัติ คือมีความกว้างน้อยกว่าสามเท่าของความกว้างเฉลี่ย และมีความกว้างมากกว่าหนึ่งในสามเท่าของความกว้างเฉลี่ย รวมทั้งจะต้องมีความสูงน้อยกว่าสองเท่าของความสูงเฉลี่ย และมีความสูงมากกว่าหนึ่งในสามเท่าของความสูงเฉลี่ย

$$D' = \{c_1, c_2, \ldots, c_m\}$$

$$\frac{1}{3}w(\bar{c}) < w(c_i) < 3w(\bar{c})$$

$$\frac{1}{3}h(\bar{c}) < h(c_i) < 2h(\bar{c})$$

เมื่อคัดสรรขนาดของส่วนประกอบเชื่อมต่อกันเสร็จเรียบร้อยแล้ว เราจะนำส่วนประกอบเชื่อมต่อกันที่เหลืออยู่มาหา เนียเรสเนเบอร์ นั่นก็คือการหาว่า ส่วนประกอบเชื่อมต่อกันอยู่บนเส้นบรรทัดเดียวกันบ้าง โดยเราจะนำจุดที่เป็นสมาชิกของเซต D' มา 1 จุด สมมติให้เป็นจุด Ci จากนั้นจะเลือกจุดที่มีระยะห่างที่สั้นที่สุดจากจุด Ci สมมติให้เป็นจุด Cj การคำนวณระยะห่างจะใช้สูตรเรขาคณิต ดังสมการต่อไปนี้

$$e_{ij} = \sqrt{(x_{i1} - x_{i2})^2 + (y_{i1} - y_{i2})^2}$$

หาก e_{ij} มีค่าน้อยกว่าสี่เท่าของ $w(\bar{c})$ จะถือว่าจุด Ci และ Cj อยู่บนเส้นบรรทัดเดียวกัน แต่ถ้าหาก e_{ij} มีค่ามากกว่าสี่เท่าของ $w(\bar{c})$ แปลว่าไม่มีจุดใดอยู่บนเส้นบรรทัดเดียวอีกแล้ว และจะถือว่า Ci อยู่บนเส้นบรรทัดใหม่ ที่ผ่านมาเป็นอัลกอริทึมของการหาเส้นหรือ Line Clustering Algorithm

ขั้นตอนต่อมาการหาความชันของแต่ละเนียเรสเนเบอร์ เมื่อได้จุดที่อยู่บนเส้นบรรทัดเดียวกัน นำจุดเหล่านั้นมาหาความชันโดยใช้สมการดังนี้

$$a_k = \frac{l_k \sum_{j=1}^{I_k} x_i y_i - \sum_{j=1} x_i \sum_{j=1}^{I_k} y_i}{l_k \sum_{j=1}^{I_k} x_i^2 - (\sum_{j=1}^{I_k} x_i)^2}$$

โดยที่ a_k คือ ความชันของเส้นบรรทัดที่ k

l_k คือ จำนวนจุดที่อยู่บนเส้นบรรทัดที่ k

(x_i, y_i) คือ จุดศูนย์ถ่วงของกรอบสี่เหลี่ยมของแต่ละส่วนประกอบที่อยู่บนเส้นบรรทัดที่ k

ขั้นตอนสุดท้ายหาความชันของเอกสาร นำความชันของแต่ละเนียเรสเนเบอร์มาหาความชันเฉลี่ย($\bar{a}$) ของภาพเอกสาร โดยใช้สมการดังนี้

$$\bar{a} = \frac{l_1 a_1 + l_2 a_2 + ... + l_k a_k}{l_1 + l_2 + ... + l_k}$$

แต่เนื่องจากจำนวนจุดของเนียเรสเนเบอร์มีผลต่อค่าที่ถูกต้องของความชัน อีกทั้งการใช้สมการหาความชันเฉลี่ยของภาพเอกสาร เพื่อหาค่าความชันเฉลี่ยจะให้ผลที่ไม่ถูกต้องหากจำนวนจุดที่อยู่บนเส้นบรรทัดเดียวกันมีค่าน้อยกว่าหรือเท่ากับหนึ่งในห้าเท่าของขนาดของไฟล์รูปภาพ ดังนั้นจึงสร้างข้อกำหนดเพิ่มเติมสำหรับการคำนวณ ดังนี้

ถ้าหากมีเส้นบรรทัดมากกว่า 1 เส้นที่มีจำนวนจุดที่อยู่บนเส้นบรรทัดเดียวกันมากกว่าหนึ่งในห้าของขนาดความกว้างของไฟล์รูปภาพ เราจะใช้สมการเพื่อหาค่าเฉลี่ยของความชันของเอกสาร

แต่ถ้าหากไม่มีเส้นบรรทัดใดเลย ที่มีจำนวนจุดมากกว่าหนึ่งในห้าของขนาดความกว้างของไฟล์รูปภาพ ใช้การเลือกความชันจากฐานนิยม

เมื่อได้ความชันของเอกสารแล้ว สามารถหามุมเอียงของเอกสารโดยใช้สมการดังนี้

$$\theta = \tan^{-1}(\bar{a})$$

ตารางที่ 3.3 อัลกอริทึม Nearest Neighbor Clustering [11]

Begin

D = Identify_Components()

D' = Remove_Small_And_Large_Components()

a = { } // Set of slope of Nearest Neighbor initially empty

l = { } // Set of number of Nearest Neighbor initially empty

While D' is not empty

Begin

Line = new Line() // Create an empty component line

Ci = Get_Top_Left_Component(D')

D' = D'-Ci

Do

Begin

Add_Component(Line, Ci)

Cj = Get_Shortest_Distance_Point(Ci,D')

$e_{ij} = \sqrt{(x_{i1} - x_{i2})^2 + (y_{i1} - y_{i2})^2}$

if $e_{ij} < 4 * w(\bar{c})$ then

Ci=Cj

D' = D' - Cj

End

While $e_{ij} < 4 * w(\bar{c})$

a_i = Calculate Slope(Line)

Add_Slope(a_i ,a)

Add_Number_of_Nearest_Neighbor(l_i ,l)

End

If every $l_i \le \frac{\text{Bitmap Width}}{5}$ then

Slope_of_Document = Mode(a_i)

Else Slope_of_Document = $\frac{l_1 a_1 + l_2 a_2 + \cdots + l_k a_k}{l_1 + l_2 + \cdots + l_k}$

SkewAngle = arctan(Slope_of_Document)

```
        Rotate_Document(SkewAngle)
End
```


สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4

การทดลองและผลการทดลอง

เนื่องด้วยงานวิจัยนี้มี 2 ส่วนคือส่วนการทำไบนาไรเซชัน และการหามุมการเอียงของเอกสาร ดังนั้นการทดลองจึงประกอบไปด้วย 2 ส่วนดังนี้

4.1 ไบนาไรเซชัน

4.1.1 ขั้นตอนการทดลอง

ทดลองโปรแกรมไบนาไรเซชัน ใน 3 อัลกอริทึมคือ อัลกอริทึมของอ็อดสึ, ไนแบล็กและอัลกอริทึมของยาโนวิทซ์และบรัคสเตียน โดยอัลกอริทึมของอ็อดสึเป็นการกำหนดค่าขีดแบ่งแบบครอบคลุม (Global Thresholding) อัลกอริทึมของไนแบล็กและอัลกอริทึมของยาโนวิทซ์และบรัคสเตียนเป็นการกำหนดค่าขีดแบ่งแบบปรับค่า (Adaptive Thresholding)

ในการทดลองจะทำการเปรียบเทียบระหว่าง 3 อัลกอริทึม โดยนำภาพที่ได้จาก 3 อัลกอริทึม ไปผ่านโปรแกรมรู้จำตัวอักษร แล้วเปรียบเทียบจำนวนคำที่ถูกต้องว่าอัลกอริทึมใดได้คำที่ถูกต้องมากกว่ากันโดยโปรแกรมรู้จำจะใช้โปรแกรมอ่านไทย เวอร์ชัน 2.5

โดยอัลกอริทึมของไนแบล็กโดยเลือกในขนาดของหน้าต่าง 15*15 และค่าของตัวแปรคงที่ที่พื้นหลังเป็นสีขาวจะให้มีค่าเท่ากับ –0.2 และหากพื้นหลังเป็นสีดำจะให้มีค่า 0.2

ส่วนอัลกอริทึมของยาโนวิทซ์และบรัคสเตียนจะตั้งค่าของการวนการทำซ้ำจำนวน 50 รอบ และให้ค่าเบต้า เท่ากับ 1.0

มาทำการทดลองบนเครื่องคอมพิวเตอร์ Pentium IV ความเร็ว 1.4 จิกะเฮิรตซ์ หน่วยความจำหลัก 128 เมกะไบต์ จำนวน 10 เอกสารมาทำการทดลอง

4.1.2 ผลการทดลอง

ผลการทดลองเป็นดังตารางที่ 4.1

ตารางที่ 4.1 ผลการทดลองเปรียบเทียบประสิทธิภาพความถูกต้องหลังจากผ่านโปรแกรมรู้จำตัวอักษร

	ภาพต้นฉบับ (ตัวอักษร)	ผ่านไบนาไรเซชันแบบครอบคลุม(ตัวอักษร)	ผ่านอัลกอริทึมของไนแบล็ก (ตัวอักษร)	ผ่านอัลกอริทึมของยาโนวิทซ์และบรัคสเตียน (ตัวอักษร)
เอกสารที่ 1	222	21	141	129
เอกสารที่ 2	553	346	336	273
เอกสารที่ 3	434	298	302	330
เอกสารที่ 4	442	245	371	383
เอกสารที่ 5	161	123	136	122
เอกสารที่ 6	145	103	74	37
เอกสารที่ 7	230	59	114	93
เอกสารที่ 8	193	127	137	180
เอกสารที่ 9	107	1	55	63
เอกสารที่ 10	244	24	59	49

4.1.3 สรุปผลการทดลองของไบนาไรเซชัน

สรุปผลการทดลองของทั้ง 3 อัลกอริทึมได้ดังนี้

อัลกอริทึมของอ็อดสึ เมื่อภาพเอกสารมีระดับสีพื้นหลัง 1 ระดับ การกำหนดค่าขีดแบ่งแบบครอบคลุมจะได้ผลดีที่สุดดังเช่นผลการทดลองของเอกสารที่ 6 จำนวนของการรู้จำตัวอักษรถูกต้องสูงที่สุด แต่ถ้าระดับสีของพื้นหลังมีมากกว่า 1 ระดับ จะทำให้ได้ภาพเอกสารที่ไม่ดี เมื่อไปผ่านโปรแกรมรู้จำตัวอักษร ทำให้การรู้จำตัวอักษรมีจำนวนตัวอักษรที่ถูกต่ำไปด้วย ถ้าหากภาพเอกสารมีระดับสีของพื้นหลังหลายระดับ และมีระดับของสีที่ใกล้เคียงกับสีของตัวอักษร ดังเช่น

เอกสารที่ 9 ภาพที่ได้มีความสมบูรณ์น้อยมากและมีข้อมูลบางอย่างหายไป เมื่อไปผ่านโปรแกรมรู้จำตัวอักษร การรู้ตัวอักษรที่ได้มีค่าความถูกต้องต่ำมาก อย่างไรก็ดีหากภาพเอกสารมีระดับสีของพื้นหลังที่แตกต่างกับระดับสีของตัวอักษรมาก และภาพเอกสารนั้นมีวรรณยุกต์อยู่ เมื่อผ่านการทำไบนาไรเซชันแล้ว ภาพที่ได้จะเห็นว่าวรรณยุกต์จะเห็นได้ชัดเจนขึ้น และส่งผลให้เมื่อเข้าโปรแกรมรู้จำตัวอักษร ผลที่ได้ก็จะมีวรรณยุกต์ออกมาด้วย

อัลกอริทึมของไนแบล็ก ผลของภาพเอกสารหลังจากผ่านไบนาไรเซชันเข้าสู่โปรแกรมรู้จำตัวอักษร ผลการรู้จำตัวอักษรมีความถูกต้องของตัวอักษรสูง จากผลการรู้จำภาษาไทยในส่วนของวรรณยุกต์นั้นจะสูญหายเป็นส่วนใหญ่ เนื่องจากภาพที่ได้จะมีจุดภาพรบกวนบริเวณวรรณยุกต์

ในส่วนการปรับพารามิเตอร์คือขนาดของหน้าต่างและค่าคงที่ ขนาดของหน้าต่างจะมีต่อความเร็วในการประมวลผล หากหน้าต่างขนาดใหญ่การประมวลผลจะช้า คุณภาพของภาพเอกสารจะมีคุณภาพดี มีจุดภาพรบกวนน้อยกว่าการใช้หน้าต่างขนาดเล็ก ส่วนการปรับค่าคงที่ระหว่าง -0.18 กับ -0.2 (ในกรณีที่ระดับสีของอักษรเป็นสีเข้ม) จากการทดลองได้ว่า ถ้าให้ค่าคงที่เท่ากับ -0.2 จะให้คุณภาพของภาพที่ได้ดีกว่าที่ค่าคงที่เท่ากับ –0.18

จากการทดลองอัลกอริทึมของของยาโนวิทซ์และบรัคสเตียน ผลของการรู้จำตัวอักษรถูกต้องสูงใกล้เคียงกับอัลกอริทึมของไนแบล็ก แต่อาจจะใช้เวลาในการประมวลผลมากกว่า ในบางภาพเอกสารเช่นเอกสารที่ 3 และ 4 จะเห็นว่าการรู้จำตัวอักษรของอัลกอริทึมของของยาโนวิทซ์และบรัคสเตียนมีความถูกต้องในการรู้จำสูงกว่าอัลกอริทึมของไนแบล็ก เพราะภาพที่ได้จากอัลกอริทึมของของยาโนวิทซ์และบรัคสเตียนมีจุดภาพรบกวนน้อยกว่าภาพของอัลกอริทึมของไนแบล็ก ทำให้การรู้จำตัวอักษรจึงได้ประสิทธิที่ดีกว่า

ในส่วนการทดลองการปรับค่าของพารามิเตอร์คือ เบต้าและจำนวนรอบ การปรับค่าของเบต้าเท่ากับ 1.0 จะเป็นค่าที่ดีที่สุด หากเป็น 2.0 จะทำให้ภาพมีจุดรบกวนเกิดขึ้นได้ ส่วนจำนวนรอบของการคำนวณไม่ควรน้อยกว่า 10 รอบ เพราะรอบยิ่งมากภาพก็จะมีความสะอาดขึ้น ส่งผลทำให้เมื่อนำภาพไปผ่านโปรแกรมรู้จำตัวอักษรจะได้ผลของการรู้จำตัวอักษรที่ถูกต้องมากขึ้น

4.2 การทดลองการหามุมการเอียงของภาพเอกสาร

4.2.1 ขั้นตอนการทดลอง

โปรแกรมสำหรับหามุมการเอียงของเอกสารที่สร้างขึ้น โดยใช้อัลกอริทึมของฮัจทรานสฟอร์มและเนียเรสเนเบอร์คัสเตอร์ริง มาทำการทดลองบนเครื่องคอมพิวเตอร์ Pentium IV ความเร็ว 1.4 จิกะเฮิรตซ์ หน่วยความจำหลัก 126 เมกะไบต์ จำนวน 6 เอกสารที่มีมุมการเอียงต่างๆกัน โดย

ภาพที่ 1 ถึง 5 จะมีความละเอียดเท่ากับ 200 จุดต่อนิ้ว ภาพที่ 6 จะมีความละเอียดเท่ากับ 75 จุดต่อนิ้ว ซึ่งเอกสารเหล่านี้รวบรวมมาจากนิตยสาร งานวิจัย และหนังสือต่างๆ ในการเตรียมภาพเอกสารที่ปรับมุมการเอียงของภาพไว้แล้ว จะใช้โปรแกรม Adobe Photoshop 7.0 ในการปรับมุม ในการทดลองจะทำการเปรียบเทียบประสิทธิภาพด้านเวลาและมุมการเอียงที่หาได้ ระหว่างการประมาณค่ามุมการเอียงด้วยสายตา และประมาณค่ามุมการเอียงโดยใช้โปรแกรมปรับมุมการเอียงของภาพเอกสารโดยอัตโนมัติ โดยทำการทดลองให้ผู้ทำการทดลอง 5 คน ทำการปรับมุมการเอียงของเอกสารที่ 1 ถึง เอกสารที่ 6 (ซึ่งมีตัวอย่างแสดงใน เอกสารอ้างอิง ค)

4.2.2 ผลการทดลอง

ผลการทดลองเป็นดังตารางที่ 4.2

ตารางที่ 4.2 ผลการทดลองเปรียบเทียบประสิทธิภาพของการหามุมการเอียง

		ฮัจทรานสฟอร์ม	เนียเรสเนเบอร์คัสเตอร์ริง	ผู้ทำการทดลองคนที่ 1	ผู้ทำการทดลองคนที่ 2	ผู้ทำการทดลองคนที่ 3	ผู้ทำการทดลองคนที่ 4	ผู้ทำการทดลองคนที่ 5
เอกสารที่ 1 (14^{o})	องศา	14	15.228	13.5	14	13.5	15	14
	เวลา (วินาที)	19	7					
เอกสารที่ 2 (40^{o})	องศา	39	39.2828	38	38	40	39	39
	เวลา (วินาที)	17	13					
เอกสารที่ 3 (18.3^{o})	องศา	16	18.476	17	18	18	17	19
	เวลา (วินาที)	18	17					
เอกสารที่ 4 (-10^{o})	องศา	-10	-9.099	-10	-9	-9	-11	-12
	เวลา (วินาที)	19	13					

เอกสารที่ 5 (-32°)	องศา	-30	-33.03	-33.5	-34	-33	-32.5	-33
	เวลา (วินาที)	23	17					
เอกสารที่ 6 (-30°)	องศา	-31	-28.19	-29	-31	-30	-29.5	-32
	เวลา (วินาที)	3	2					

4.2.3 สรุปผลการทดลองของการปรับมุมการเอียงของเอกสารแบบอัตโนมัติ

จากการทดลองจะเห็นว่าหากทำการปรับมุมการเอียง โดยใช้สายตาของคนในการประมาณค่าจะใช้เวลามากกว่า การปรับมุมการเอียงด้วยอัลกอริทึมของฮัจทรานสฟอร์มและใช้เนียเรสเนเบอร์คัสเตอร์ริง และนอกจากนั้นผู้ทำการทดลองจะต้องใช้จำนวนครั้งที่ปรับมุมการเอียงมากกว่า 1 ครั้ง ซึ่งการปรับแต่ละครั้งจะทำให้คุณภาพของเอกสารลดลง

ผลของอัลกอริทึมของฮัจทรานสฟอร์มในการหามุมการเอียงของเอกสาร มุมที่ได้จะมีค่าคลาดเคลื่อนเล็กน้อย อันเนื่องจากการแทนค่าของมุมการเอียงที่เป็นเลขจำนวนเต็ม ดังนั้นมุมผลลัพธ์ที่ได้จึงเป็นแค่เลขจำนวนเต็มเท่านั้น นอกจากนี้วิธีของฮัจทรานสฟอร์มยังสามารถหามุมการเอียงได้ดีสำหรับเอกสารที่มีจุดภาพรบกวนมากๆ ยกตัวอย่าง เช่นเอกสารที่มีรูปภาพ หรือเอกสารที่มีขอบสีดำซึ่งเกิดจากข้อผิดพลาดในขั้นตอนของการอ่านเอกสารจากเครื่องกวาดภาพ

ผลของอัลกอริทึมเนียเรสเนเบอร์คัสเตอร์ริงในการหามุมการเอียง จะให้ค่าของมุมการเอียงที่ถูกต้องและใกล้เคียงกับมุมการเอียงของเอกสารจริงเนื่องจากค่าของมุมการเอียงที่ได้สามารถเป็นค่าที่มีจุดทศนิยมได้ จึงทำให้มีความถูกต้องมากกว่าอัลกอริทึมของฮัจทรานสฟอร์ม นอกจากนี้หากภาพเอกสารมีจุดภาพรบกวนมาก อัลกอริทึมเนียเรสเนเบอร์คัสเตอร์ริงทำให้มุมการเอียงที่หาได้มีความผิดพลาดสูง แต่เมื่อทำการลดจุดภาพรบกวนลงแล้วอัลกอริทึมเนียเรสเนเบอร์คัสเตอร์ริง จะให้ค่าความผิดพลาดที่ลดลงและทำงานได้เร็วขึ้น

เมื่อทำการเปรียบเทียบระหว่างการอัลกอริทึมของฮัจทรานสฟอร์มและใช้เนียเรสเนเบอร์คัสเตอร์ริง ผลคืออัลกอริทึมเนียเรสเนเบอร์คัสเตอร์ริงจะใช้เวลาในการคำนวณน้อยกว่าฮัจทรานสฟอร์ม และองศาที่ได้ใกล้เคียงค่าที่ถูกต้องมากที่สุด

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

จากการทดลองและวิเคราะห์ผลที่ผ่านมา สามารถสรุปผลการวิจัยและข้อเสนอแนะเพื่อเป็นแนวทางในการพัฒนาต่อไปในอนาคต ดังนี้

5.1 สรุปผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างโปรแกรมสำหรับการประมวลล่วงหน้าของภาพเอกสาร ซึ่งประกอบด้วยโปรแกรม 2 ส่วนหลักคือ โปรแกรมไบนาไรเซชัน และโปรแกรมปรับมุมการเอียงของภาพเอกสารโดยอัตโนมัติ

5.1.1. โปรแกรมไบนาไรเซชัน

การทำไบนาไรเซชันเป็นการแปลงภาพเอกสารต่างๆ ที่ระดับของสี 256 สี ให้มีระดับสี 2 สี คือ สีขาวและสีดำ เพื่อให้ภาพเอกสารนั้นแยกส่วนของพื้นหลังกับตัวอักษรได้ จากการทดลองภาพเอกสารต้นฉบับจำนวน 10 ภาพ มีจำนวนตัวอักษรรวม 2,731 ตัว ผลการรู้จำตัวอักษรเมื่อนำภาพเอกสารที่ผ่านอัลกอริทึมของอ็อดสึ จำนวนตัวอักษรที่ถูกต้องเท่ากับ 1,347 ตัว คิดเป็น 49.32% เอกสารภาพที่ผ่านอัลกอริทึมของไนแบล็ก จำนวนตัวอักษรที่ถูกต้องเท่ากับ 1,725 ตัว คิดเป็น 63.16% เอกสารภาพที่ผ่านอัลกอริทึมของยาโนวิทซ์และบรัคสเตียน จำนวนตัวอักษรที่ถูกต้องเท่ากับ 1,659 ตัว คิดเป็น 60.75%

จาก 3 อัลกอริทึมในการทำไบนาไรเซชัน อัลกอริทึมของไนแบล็กให้จำนวนตัวอักษรของการรู้จำที่ถูกต้องมากที่สุด อัลกอริทึมของยาโนวิทซ์และบรัคสเตียนได้ผลรองลงมา และอัลกอริทึมของอ็อดสึให้ผลน้อยที่สุด

5.1.2. โปรแกรมปรับมุมการเอียงของภาพเอกสารโดยอัตโนมัติ

การปรับมุมการเอียงของภาพเอกสารด้วยการใช้โปรแกรมปรับนั้น สามารถทำได้ดีกว่าการปรับมุมเอียงด้วยสายตา เพราะค่าของมุมที่โปรแกรมปรับจะมีค่าของมุมที่ใกล้เคียงกับมุมที่ถูกปรับจริง โดยอัลกอริทึมฮัจทรานสฟอร์มจะมีค่าความคลาดเคลื่อนของมุมเอียง $\pm 0^{o}$ ถึง $\pm 3^{o}$ และอัลกอริทึมเนียเรสเนเบอร์คัตเตอริงจะมีค่าความคลาดเคลื่อนระหว่าง $\pm 0^{o}$ ถึง $\pm 2^{o}$ เฉพาะหากว่าใช้อัลกอริทึมเนียเรสเนเบอร์คัตเตอริง จะได้ค่าของมุมที่ใกล้เคียงมาก นอกจากนั้นการใช้โปรแกรม

ปรับมุมเอียงการเอียงของภาพเอกสารโดยอัตโนมัติยังช่วยรักษาคุณภาพของภาพได้ดีกว่าการปรับมุมเอียงของภาพเอกสารด้วยสายตา เพราะเป็นการปรับได้ในครั้งเดียว

5.2 ปัญหาและอุปสรรค

จากการศึกษาและทดลอง โปรแกรมแปลงข้อมูลภาพเอกสารที่มีระดับสี 256 ระดับสีเทาให้เป็นข้อมูลภาพที่มีระดับสีเพียง 2 ระดับและการปรับมุมเอียงของเอกสารแบบอัตโนมัติ สามารถสรุปปัญหา ที่มีผลต่อการทำงานของระบบ คือ

5.2.1 ผู้วิจัยขาดประสบการณ์ในการเขียนโปรแกรมแสดงผลและประมวลผลภาพภายใต้ภาษาวิชวลซีพลัสพลัสมาก่อน ทำให้เกิดความล่าช้าในการพัฒนาโปรแกรม

5.2.2 ผู้วิจัยมีประสบการณ์ทางด้านการประมวลผลภาพน้อยเกินไป จนบางครั้งในการศึกษาเพิ่มเติมด้วยตนเองเป็นไปด้วยความยากลำบาก

5.2.3 เนื่องจากต้องพัฒนาโปรแกรมต่อจากโครงงานรุ่นก่อนซึ่งไม่ได้มีเอกสารประกอบโปรแกรมไว้ให้ศึกษา จึงมีความยากลำบากในการสานงานต่อ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รายการอ้างอิง

1. ชาญฤทธิ์ สันตินานาเลิศ. การออกแบบและพัฒนาโปรแกรมโอซีอาร์ภาษาไทย. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย, 2542.
2. วิทยาศาสตร์ เทคโนโลยีและสิ่งแวดล้อม, กระทรวง, สำนักงานพัฒนาวิทยาศาสตร์ และเทคโนโลยีแห่งชาติ, ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ. อ่านไทย 1.0 [โปรแกรมคอมพิวเตอร์]. กรุงเทพฯ :ศูนย์เทคโนโลยี
3. ยุทธพงษ์ รังสรรค์เสรี และกฤษณะ ชินสาร. การประมาณค่าความเอียงหน้าเอกสารสิ่งพิมพ์., สารสนเทศลาดกระบัง, เล่มที่ 3, หน้า 14-21.
4. สิทธิชัย บุษหมั่น. การสร้างภาพไบนารีของภาพเอกสารโดยการจำแนกส่วนของข้อความและภาพถ่าย. วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้า เจ้าคุณทหารลาดกระบัง, 2540.
5. J. Sauvola and M. Pietikäinen. Adaptive document image binarization. Pattern Recognition. Vol. 33, 2000, 225-236.
6. D.T. Øivind and T. Torfinn. Evaluation of Binarization Methods for Document Images. IEEE Transaction on Pattern Analysis and Machine Intelligence. Vol. 17, NO. 3, 1995.
7. S. D. Yonowitz and A. M. Bruckstein. A new method for image segmentation. Comuter vision, Graphics and Image Processing, Vol. 46, No. 1, Apr. 1989, 82-95.
8. G.Nagy, S.C.Seth. and S.D. Stoddard. Document analysis with a expert system. Processing , Pattern Recognition in Practice II, Amsterdam, June 19-21, 1985.
9. S.N. Srihari. Document Image Understanding. IEEE Fall Joint Computer Conference, Dallas, Texas, 1986, 87-95.
10. R. Duda and P. Hart. Use of the Hough Transformation to detect lines and curves in picture. Communication of the ACM, Vol. 15, No. 1, 1972, 11-15.
11. Amin and S. Fischer. A Document Skew Detection Method Using the Hough Transform. Pattern Analysis & Applications, Vol. 3, Issue 3, 2000, 243-253.

12. S.C. Hinds, J.L. Fisher and D.P. D'Amato. A Document Skew Detection Method Using Run-length Encoding and the Hough Transform. Proceedings, 10th Internation Conference on Pattern Recognition, 1990, 464-468.

13. J. Xiaoyi, B. Horst and W. Dubravka. Skew Detection of Document Images by Focused Nearest- Neighbor Clustering. Document Analysis and Recognition, 1999. ICDAR '99. Proceedings of the Fifth International Conference on, 1999, 629-632.

14. N. Liolios, N. Fakotakis and G. Kokkinakis. Improved Document Skew Detection Based on Text Line Connected Component Clustering. Image Processing, 2001. Proceedings. 2001 International Conference on, Vol.1, 2001, 1098-1101.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก

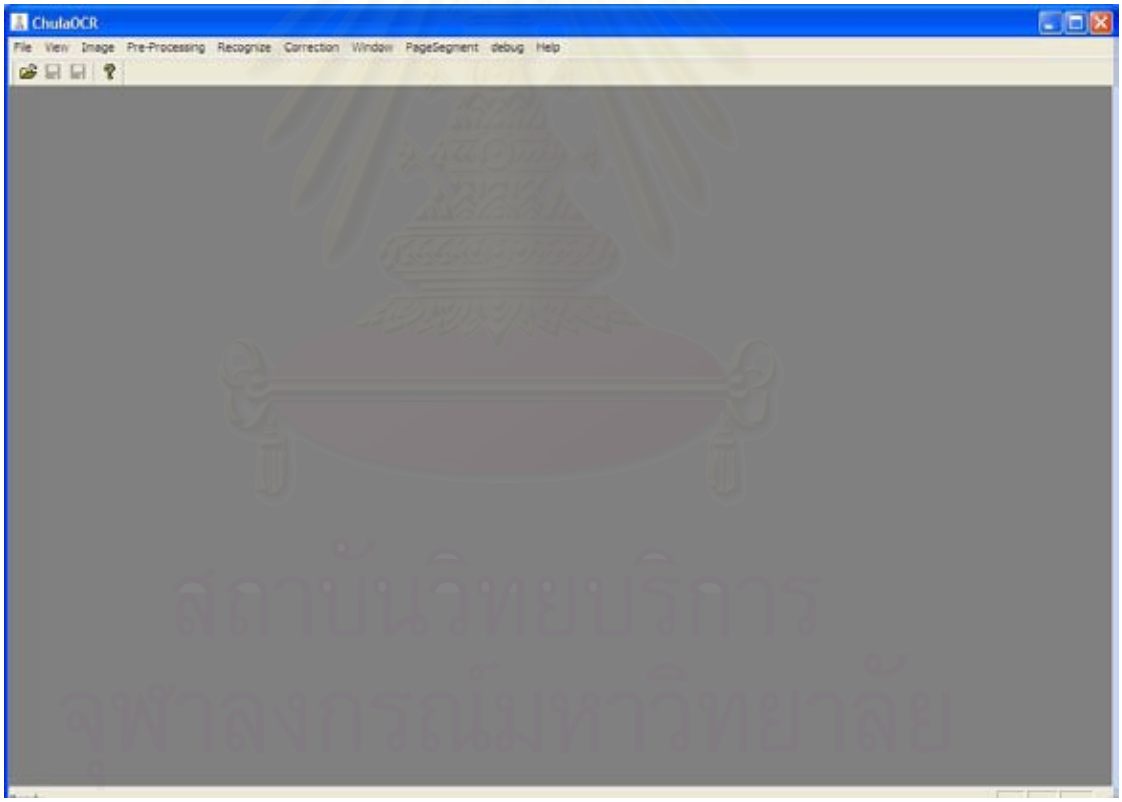
คู่มือการใช้งานโปรแกรมโอซีอาร์

ก.1 ความต้องการขั้นต่ำของโปรแกรม

เครื่องไอบีเอ็มพีซีคอมแพ็ตทิเบิล ที่มีหน่วยประมวลผลกลางเพนเทียมทูความเร็ว 350 เมกกะเฮิร์ตซ มีหน่วยความจำหลัก 128 เมกกะไบต์ ระบบปฏิบัติการ Windows 98 และเนื้อที่ในฮาร์ดดิสค์ จำนวน 30 เมกกะไบต์

ก.2 ส่วนประกอบของโปรแกรม

เมื่อเปิดโปรแกรมให้ทำงาน จะพอหน้าจอดังนี้



รูปที่ ก.1 แสดงแถบคำสั่ง

โปรแกรมจะประกอบด้วยแถบคำสั่งหลักดังนี้

1. File เป็นแถบคำสั่งรวบรวมชุดคำสั่งย่อยที่เกี่ยวกับการปิด/เปิด บันทึกข้อมูลภาพ หรืออ่านข้อมูลภาพจากสแกนเนอร์ โปรแกรมโอซีอาร์รุ่นนี้ยังคงชุดคำสั่งนี้จากรุ่นก่อน

2. View เป็นแถบคำสั่งที่รวมชุดคำสั่งการซ่อนแสดง แถบเครื่องมือ และแถบแสดงสถานะของโปรแกรม

3. Image เป็นแถบคำสั่งที่รวบรวมชุดคำสั่งการปรับแต่งข้อมูลภาพก่อนการทำการรู้จำเอกสาร

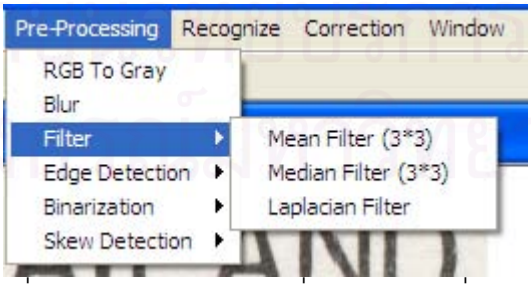
4. Pre-Processing เป็นแถบคำสั่งที่รวบรวมชุดคำสั่งการปรับแต่งข้อมูลภาพก่อนการทำการรู้จำเอกสาร เป็นส่วนหนึ่งของงานวิจัยนี้ ซึ่งจะประกอบไปด้วยแถบคำสั่งต่าง ๆ ดังต่อไปนี้

รูปที่ ก.2 แสดงแถบกลุ่มคำสั่งของแถบคำสั่ง Pre-Processing

4.1. RGB To Gray เป็นแถบคำสั่งสำหรับในการแปลงภาพเอกสารจากภาพสีเป็นภาพสีระดับสีเทา 256 ระดับสี

4.2. Blur เป็นแถบคำสั่งสำหรับภาพเอกสารให้เป็นภาพเบรอ เพื่อจะลดจุดรบกวนต่างภาพในภาพให้ลดลง

4.3. Filter เป็นแถบกลุ่มคำสั่งสำหรับกรองภาพเอกสาร ดังต่อไปนี้



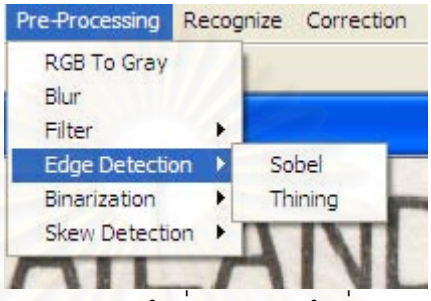
รูปที่ ก.3 แสดงแถบกลุ่มคำสั่งของแถบคำสั่ง Filter

4.3.1 Mean Filter (3*3) เป็นแถบคำสั่งสำหรับกรองภาพเอกสารแบบหาค่าเฉลี่ย จากจุดภาพรอบๆ ข้างจำนวน 8 จุด

4.3.2 Median Filter(3*3) เป็นแถบคำสั่งสำหรับกรองภาพเอกสารแบบหาค่ามัธยฐาน จากจุดภาพรอบๆ ข้างจำนวน 8 จุด

4.3.3 Laplacian Filter เป็นแถบคำสั่งสำหรับกรองภาพเอกสารแบบลาปาเซียน

4.4. Edge Detection เป็นแถบกลุ่มคำสั่งสำหรับหาขอบภาพของภาพเอกสาร โดยมีแถบคำสั่งย่อยๆ ดังต่อไปนี้

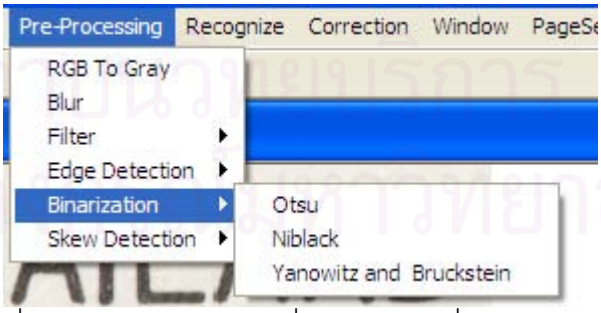


รูปที่ ก.4 แสดงแถบกลุ่มคำสั่งของแถบคำสั่ง Edge Detection

4.4.1 Sobel เป็นการหาขอบภาพ โดยใช้โซเบิลโอเปอร์เรเตอร์มากระทำคอนเวอลูชัน (Convolution)

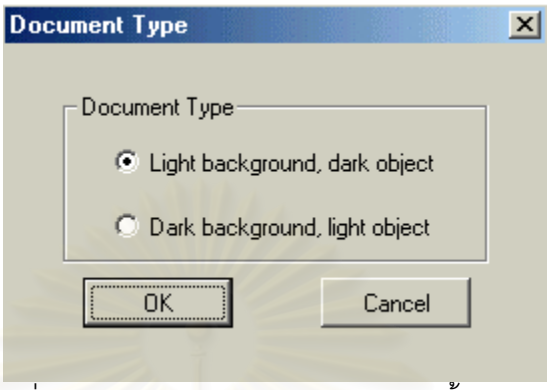
4.4.2 Thining เป็นการหากระดูก (Skeleton) ของภาพ โดยการพยายามลดจุดภาพข้างเคียง

4.5. Binarization เป็นแถบกลุ่มคำสั่งสำหรับการปรับภาพเอกสารที่มีระดับสี 256 ระดับสีเทาให้เป็นข้อมูลภาพที่มีระดับสีเพียง 2 ระดับ คือสีขาวและสีดำ โดยมีแถบคำสั่งย่อยๆ ดังต่อไปนี้



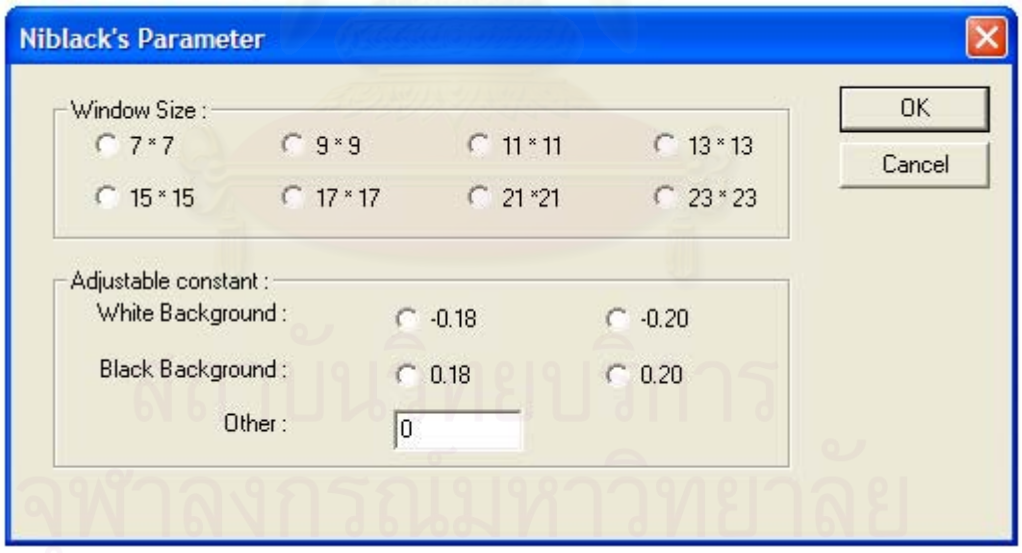
รูปที่ ก.5 แสดงแถบกลุ่มคำสั่งของแถบคำสั่ง Binarization

4.5.1 Otsu คำสั่งสำหรับการปรับภาพเอกสารตามวิธีการของอ็อดสึ โดยจะปรากฏหน้าต่างให้เลือกประเภทของภาพว่าเป็นภาพสีพื้นเข้มหรืออ่อน ดังรูป



รูปที่ ก.6 แสดงไดอะล๊อกของการเลือกสีพื้นหลัง

4.5.2 Niblack คำสั่งสำหรับการปรับภาพเอกสารตามวิธีการของไนแบล็ก โดยจะปรากฏหน้าต่างให้เลือกขนาดของหน้าต่างที่ใช้ในการคำนวณ นอกจากนั้นยังมีให้เลือกค่าคงที่ต่างๆ ดังรูป



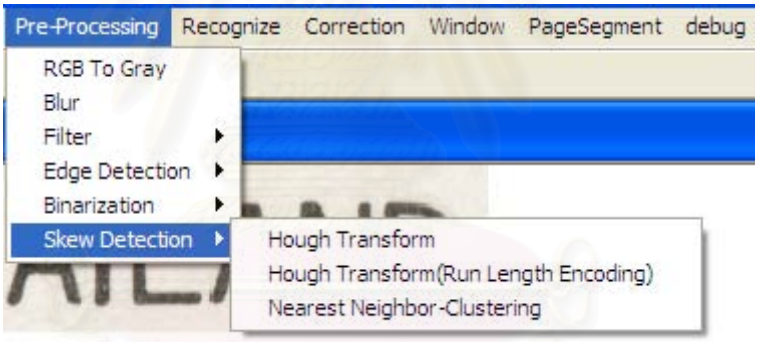
รูปที่ ก.7 แสดงไดอะล๊อกของการเลือกพารามิเตอร์ในการประมวลผล ตามอัลกอริทึมของไนแบล็ก

4.5.3 Yanowitz and Bruckstein คำสั่งสำหรับการปรับภาพเอกสารตามวิธีการของยาโนวิทซ์และบรัคสเตียน โดยจะปรากฏหน้าต่างให้ใส่ค่าคงที่ตรงช่อง Beta Constant: โดยค่าที่ใส่นั้นต้องมีค่าระหว่าง 1.0 ถึง 2.0 เท่านั้น และอีกช่อง Max Iterations: จะ

เป็นการกำหนดจำนวนการทำซ้ำ ควรมีค่ามากกว่า 1 รอบ แต่ไม่เกิน 500 รอบ รอบมากจะทำให้ในการคำนวณนานขึ้น โปรแกรมช้า แต่ภาพก็จะมีค่าขีดแบ่งที่ดีขึ้นด้วย นอกจากนั้นยังมีให้เลือกค่าคงที่ต่างๆ ดังรูป

รูปที่ ก.8 แสดงไดอะล๊อกของการเลือกพารามิเตอร์ในการประมวลผลตามอัลกอริทึมยาโนวิทซ์และบรัคสเตียน

4.6. Skew Detection เป็นแถบกลุ่มคำสั่งสำหรับหามุมเอียงของเอกสาร โดยมีแถบคำสั่งย่อยๆ ดังต่อไปนี้



รูปที่ ก.9 แสดงแถบกลุ่มคำสั่งของแถบคำสั่ง Skew Detection

4.6.1 Hough Transform เป็นการปรับมุมเอียงของภาพเอกสาร โดยใช้จุดภาพทั้งภาพ มาหามุมเอียง

4.6.2 Hough Transform (Run Length Encoding) เป็นการปรับมุมเอียงของภาพเอกสาร โดยใช้การเข้ารหัส เพื่อให้ใช้จุดภาพในการคำนวณน้อยลงและเร็วขึ้น มาหามุมเอียง

4.6.3 Nearest Neighbor-Clustering เป็นการปรับมุมเอียงของภาพเอกสารด้วยการ พยายามจัดจุดภาพเป็นบล็อกแล้วนำไปหามุมเอียง

5. Recognize เป็นแถบคำสั่งที่รวบรวบคำสั่งสำหรับการรู้จำ นอกจากนี้จะมีคำสั่งการเลือกโซนที่ต้องการทำการรู้จำแบบเลือกด้วยตนเอง นอกจากนี้ในแถบคำสั่งยังมีคำสั่งสำหรับการเลือกหน่วยรู้จำตัวอักษร(Select recognition unit) โดยสามารถเลือกใช้หน่วยรู้จำได้ทั้งแบบใช้ นิวรอลเน็ตเวอร์ค และ แบบใช้ ซัพพอร์ตเวคเตอร์แมชชีน

6. Correction เป็นแถบคำสั่งสำหรับการแก้ไขคำผิดที่หลังจากการรู้จำ

7. Window เป็นแถบคำสั่งที่รวบรวมคำสั่งการวางรูปแบบการแสดงผลของหน้าต่าง

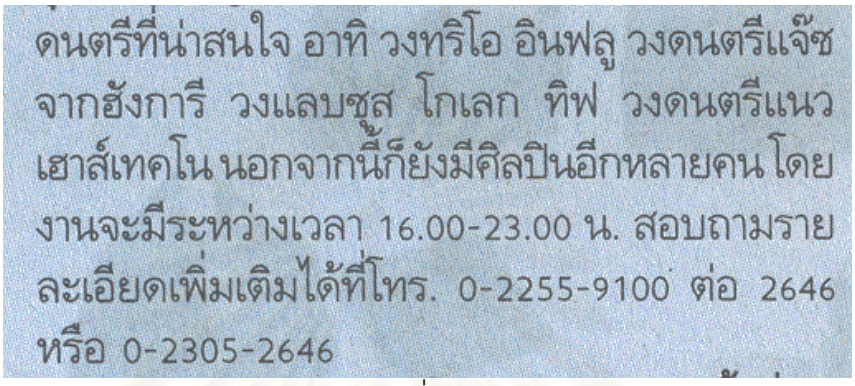
8. Page Segment เป็นแถบคำสั่งสำหรับการแบ่งส่วนหน้ากระดาษแบบอัตโนมัติ Debug เป็นแถบคำสั่งพิเศษ สำหรับผู้พัฒนาโปรแกรม ภายในจะประกอบด้วยชุดคำสั่งเพื่อดู การตัดตัวอักษร ลำดับของการตัด และ ลำดับของการรู้จำตัวอักษร

9. Help เป็นแถบคำสั่งอธิบายตัวโปรแกรม

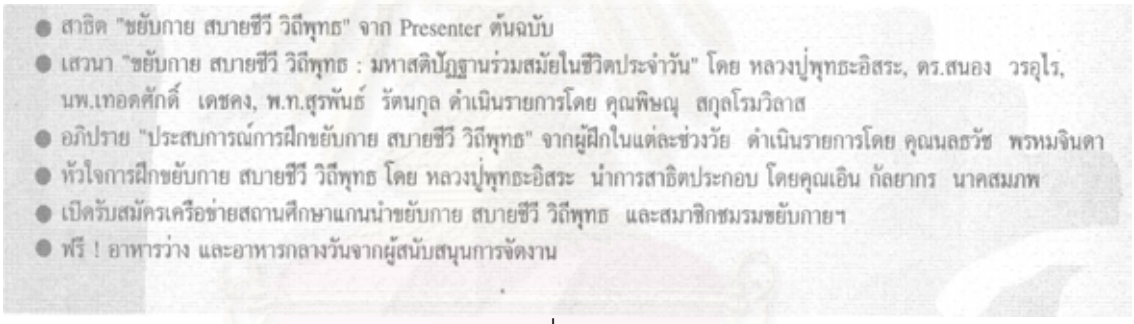
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ข

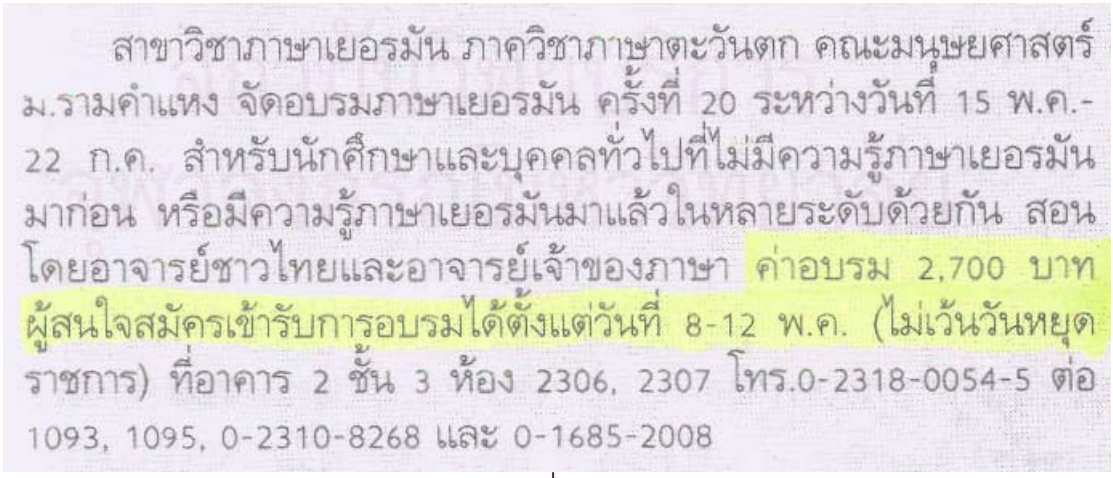
เอกสารที่ใช้ในการทดสอบการทำไบนาไรเซชัน



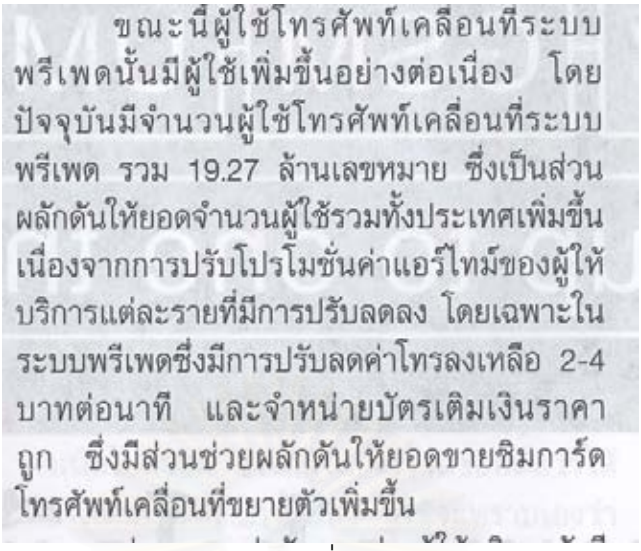
รูปที่ ข.1



รูปที่ ข.2

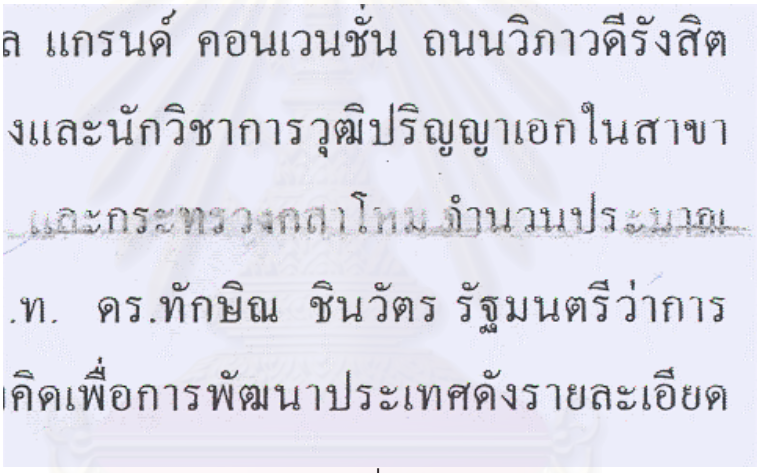


รูปที่ ข.3



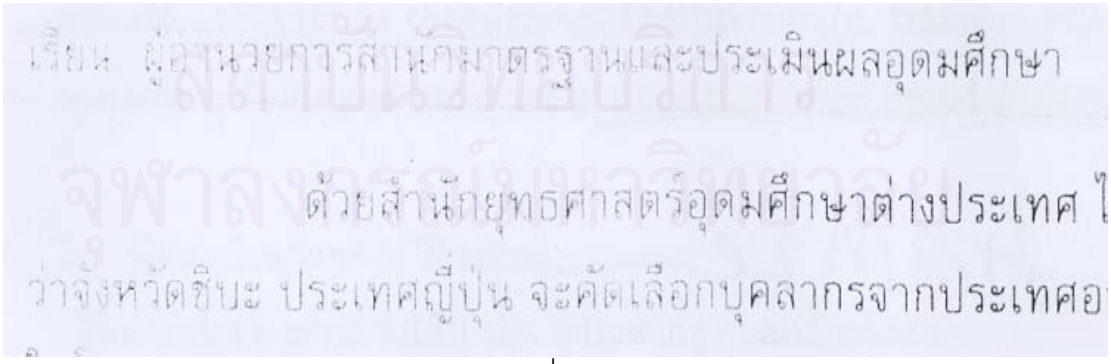
ขณะนี้ผู้ใช้โทรศัพท์เคลื่อนที่ระบบพรีเพดนั้นมีผู้ใช้เพิ่มขึ้นอย่างต่อเนื่อง โดยปัจจุบันมีจำนวนผู้ใช้โทรศัพท์เคลื่อนที่ระบบพรีเพด รวม 19.27 ล้านเลขหมาย ซึ่งเป็นส่วนผลักดันให้ยอดจำนวนผู้ใช้รวมทั้งประเทศเพิ่มขึ้น เนื่องจากการปรับโปรโมชั่นค่าแอร์ไทม์ของผู้ให้บริการแต่ละรายที่มีการปรับลดลง โดยเฉพาะในระบบพรีเพดซึ่งมีการปรับลดค่าโทรลงเหลือ 2-4 บาทต่อนาที และจำหน่ายบัตรเติมเงินราคาถูก ซึ่งมีส่วนช่วยผลักดันให้ยอดขายซิมการ์ดโทรศัพท์เคลื่อนที่ขยายตัวเพิ่มขึ้น

รูปที่ ข.4



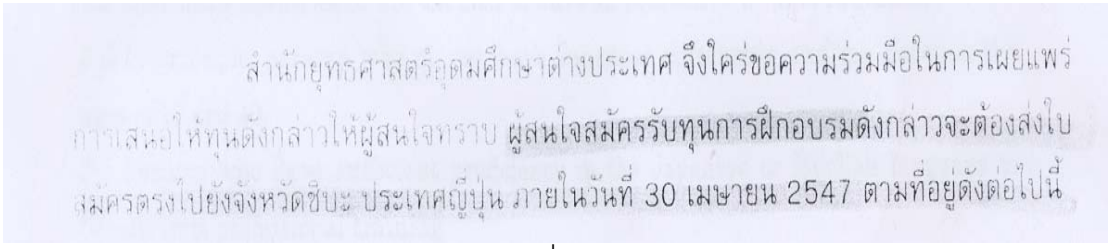
ล แกรนด์ คอนเวนชั่น ถนนวิภาวดีรังสิต
งและนักวิชาการวุฒิปริญญาเอกในสาขา
และกระทรวงคมนาคม จำนวนประมาณ
.ท. ดร.ทักษิณ ชินวัตร รัฐมนตรีว่าการ
คิดเพื่อการพัฒนาประเทศดังรายละเอียด

รูปที่ ข.5

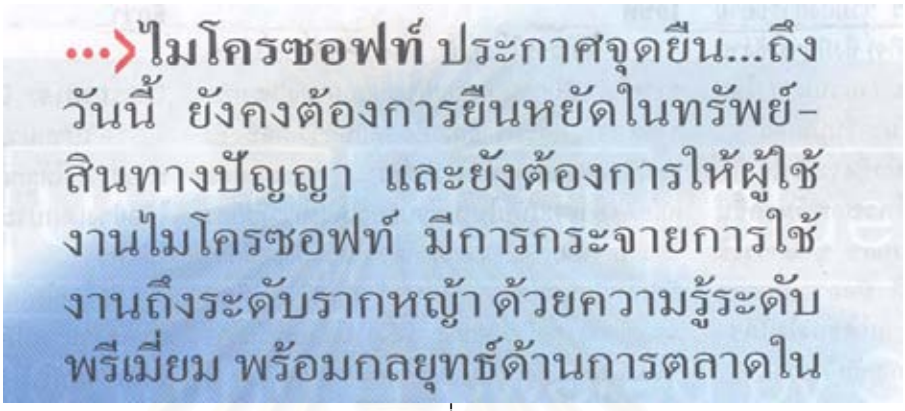


เรียน ผู้อำนวยการสำนักมาตรฐานและประเมินผลอุดมศึกษา
ด้วยสำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ ไ
ว่าจังหวัดชิบะ ประเทศญี่ปุ่น จะคัดเลือกบุคลากรจากประเทศอ

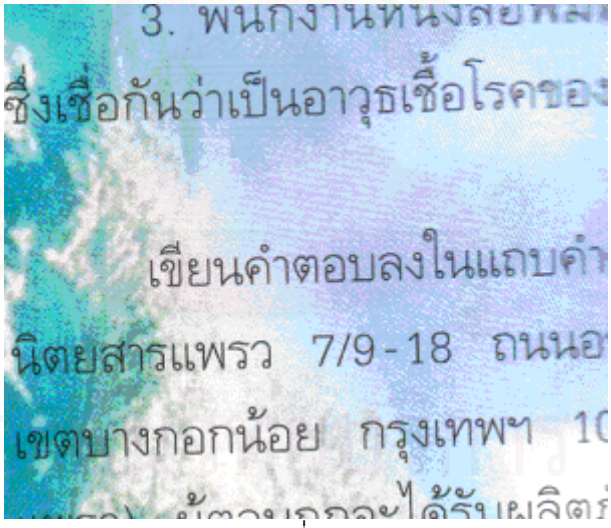
รูปที่ ข.6



รูปที่ ข.7

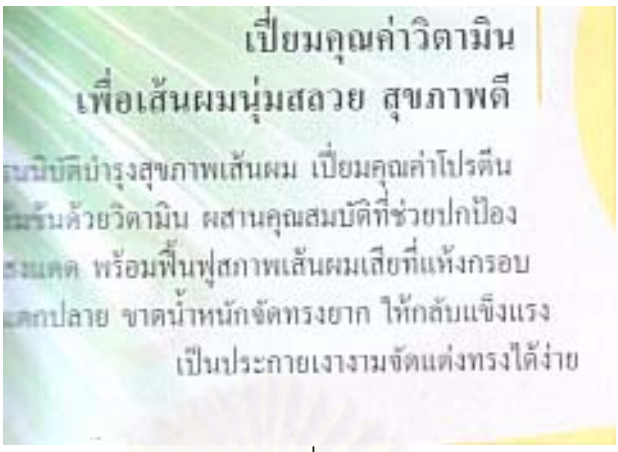


รูปที่ ข.8



รูปที่ ข.9

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

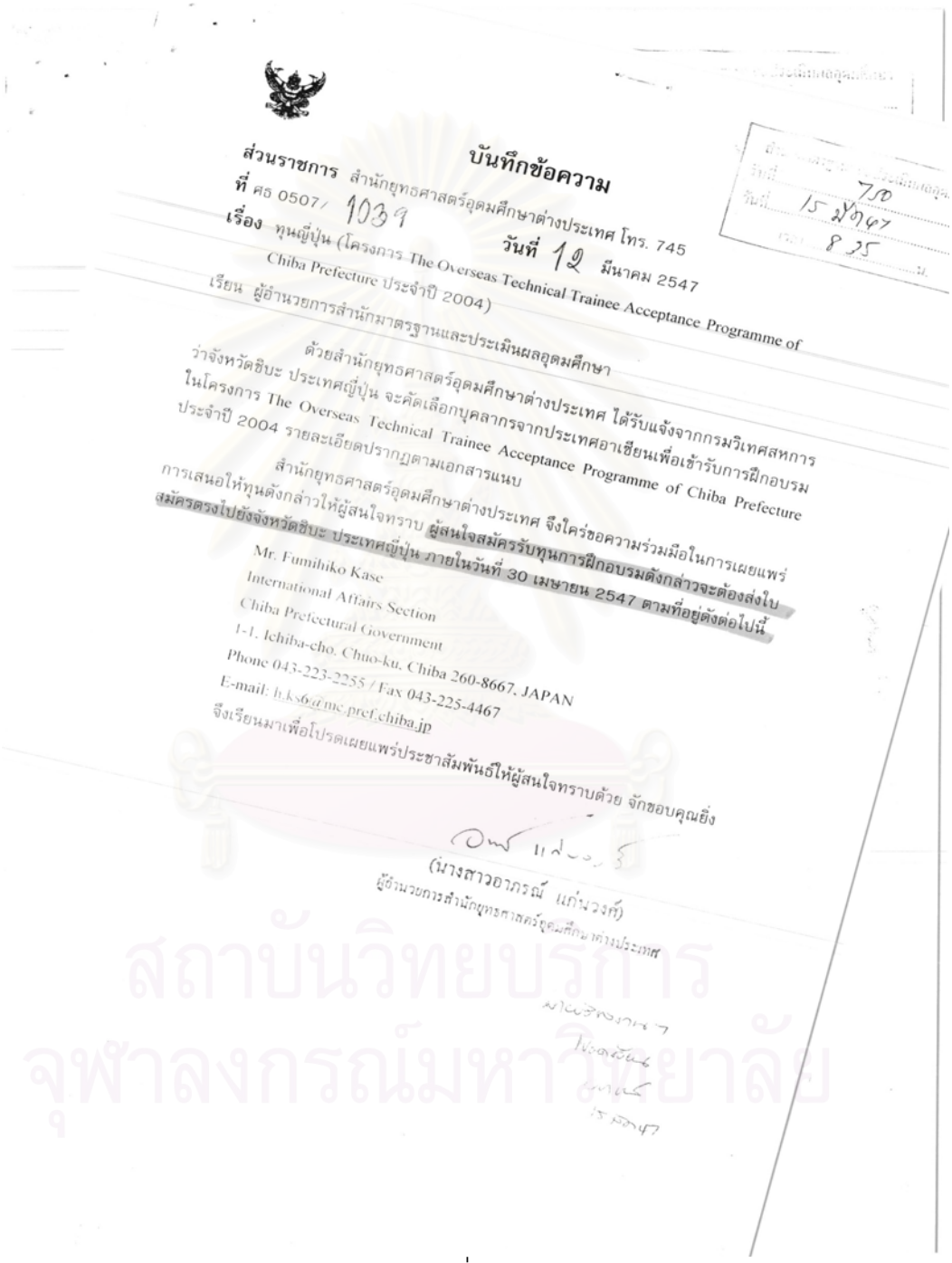


รูปที่ ข.10

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ค

เอกสารที่ใช้ในการทดสอบอัลกอริทึมปรับมุมเอียงโดยอัตโนมัติ



บันทึกข้อความ

ส่วนราชการ สำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ โทร. 745
ที่ ศธ 0507/ 1039 วันที่ 12 มีนาคม 2547
เรื่อง ทุนญี่ปุ่น (โครงการ The Overseas Technical Trainee Acceptance Programme of Chiba Prefecture ประจำปี 2004)
เรียน ผู้อำนวยการสำนักมาตรฐานและประเมินผลอุดมศึกษา

ด้วยสำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ ได้รับแจ้งจากกรมวิเทศสหการ ว่าจังหวัดชิบะ ประเทศญี่ปุ่น จะคัดเลือกบุคลากรจากประเทศอาเซียนเพื่อเข้ารับการฝึกอบรม ในโครงการ The Overseas Technical Trainee Acceptance Programme of Chiba Prefecture ประจำปี 2004 รายละเอียดปรากฏตามเอกสารแนบ

สำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ จึงใคร่ขอความร่วมมือในการเผยแพร่การเสนอให้ทุนดังกล่าวให้ผู้สนใจทราบ ผู้สนใจสมัครรับทุนการฝึกอบรมดังกล่าวจะต้องส่งใบสมัครตรงไปยังจังหวัดชิบะ ประเทศญี่ปุ่น ภายในวันที่ 30 เมษายน 2547 ตามที่อยู่ดังต่อไปนี้

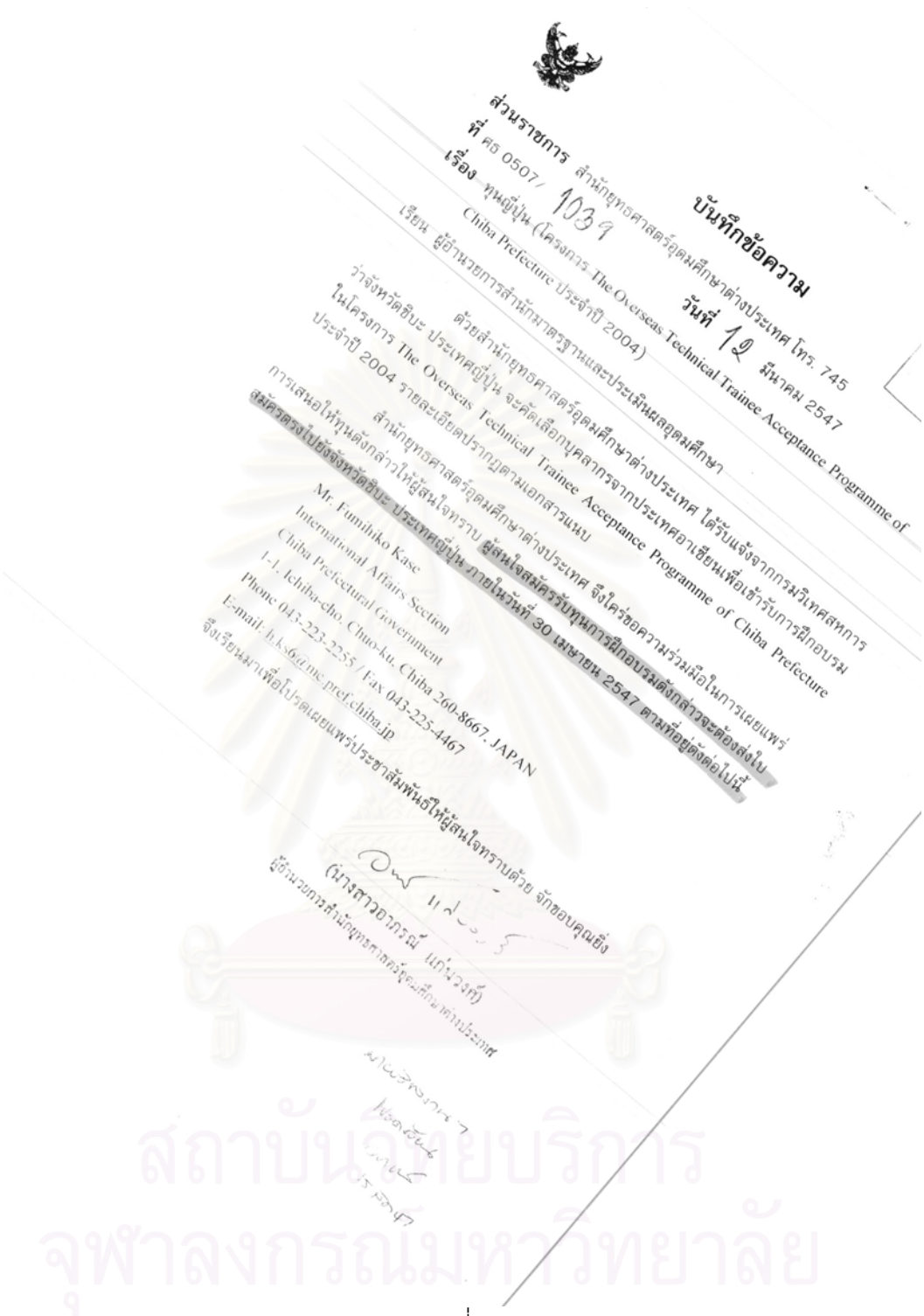
Mr. Fumihiko Kase
International Affairs Section
Chiba Prefectural Government
1-1, Ichiba-cho, Chuo-ku, Chiba 260-8667, JAPAN
Phone 043-223-2255 / Fax 043-225-4467
E-mail: h.ks6@mc.pref.chiba.jp

จึงเรียนมาเพื่อโปรดเผยแพร่ประชาสัมพันธ์ให้ผู้สนใจทราบด้วย จักขอบคุณยิ่ง

(นางสาวอาภรณ์ แก่นวงศ์)
ผู้อำนวยการสำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ

รูปที่ ค.1.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



บันทึกข้อความ

ส่วนราชการ สำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ โทร. 745

ที่ ศธ 0507/ 1039 วันที่ 19 มีนาคม 2547

เรื่อง ทุนญี่ปุ่น (โครงการ The Overseas Technical Trainee Acceptance Programme of Chiba Prefecture ประจำปี 2004)

เรียน ผู้อำนวยการสำนักมาตรฐานและประเมินผลอุดมศึกษา

ด้วยสำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ ได้รับแจ้งจากกรมวิเทศสหการ ว่าจังหวัดชิบะ ประเทศญี่ปุ่น จะคัดเลือกบุคลากรจากประเทศอาเซียนเพื่อเข้ารับการฝึกอบรม ในโครงการ The Overseas Technical Trainee Acceptance Programme of Chiba Prefecture ประจำปี 2004 รายละเอียดปรากฏตามเอกสารแนบ

สำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ จึงใคร่ขอความร่วมมือในการเผยแพร่ การเสนอให้ทุนดังกล่าวให้ผู้สนใจทราบ ผู้สนใจสมัครรับทุนการฝึกอบรมดังกล่าวจะต้องส่งใบ สมัครตรงไปยังจังหวัดชิบะ ประเทศญี่ปุ่น ภายในวันที่ 30 เมษายน 2547 ตามที่อยู่ดังต่อไปนี้

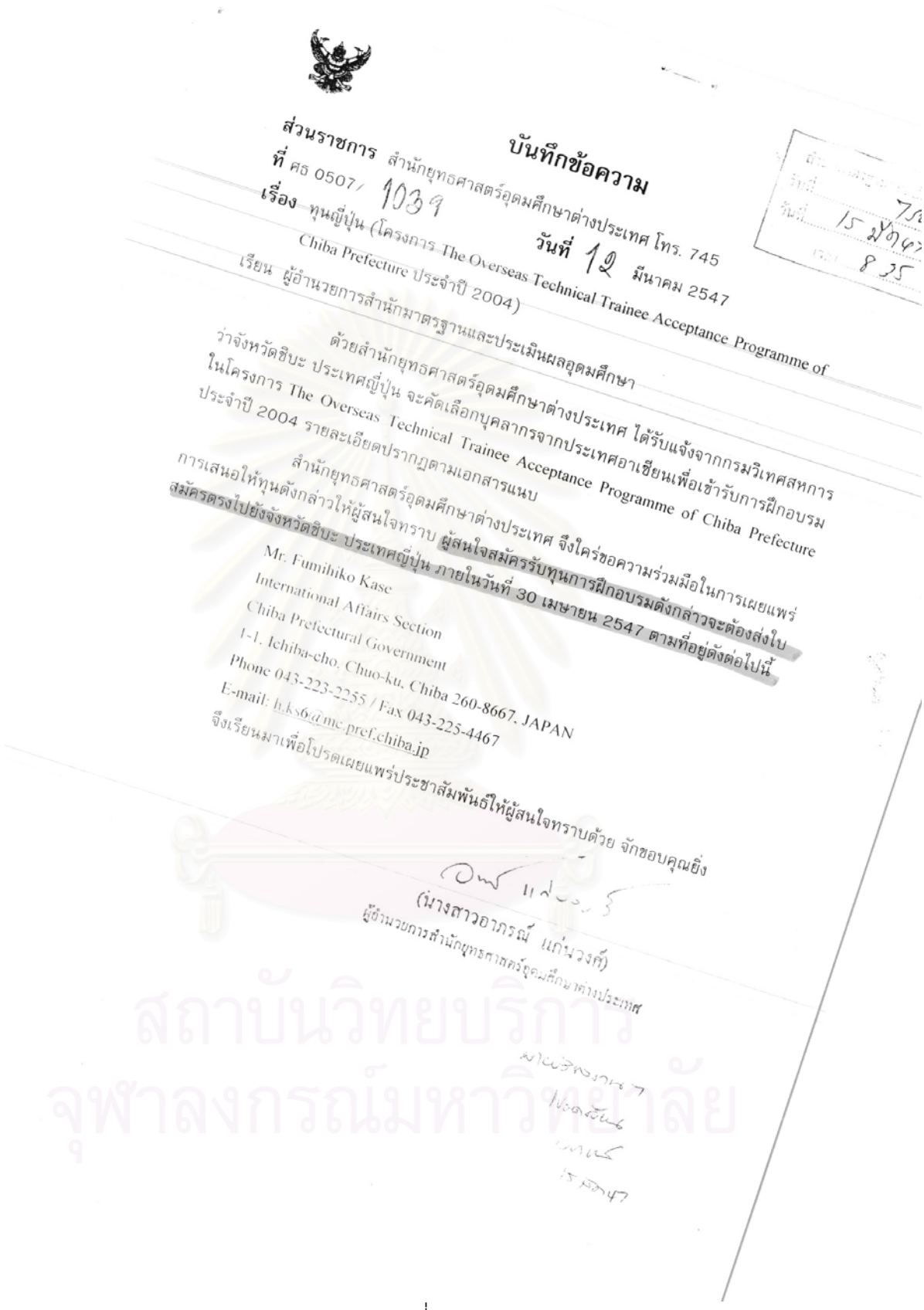
Mr. Fumihiko Kase
International Affairs Section
Chiba Prefectural Government
1-1, Ichiba-cho, Chuo-ku, Chiba 260-8667, JAPAN
Phone 043-223-2255 / Fax 043-225-4467
E-mail: h.kase@mc.pref.chiba.jp

จึงเรียนมาเพื่อโปรดเผยแพร่ประชาสัมพันธ์ให้ผู้สนใจทราบด้วย จักขอบคุณยิ่ง

(นางสาวอาภรณี นภโหวกที)
ผู้อำนวยการสำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ ค.2.



บันทึกข้อความ

ส่วนราชการ สำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ โทร. 745
ที่ ศธ 0507/ 1039 **วันที่** 12 มีนาคม 2547
เรื่อง ทุนญี่ปุ่น (โครงการ The Overseas Technical Trainee Acceptance Programme of Chiba Prefecture ประจำปี 2004)
เรียน ผู้อำนวยการสำนักมาตรฐานและประเมินผลอุดมศึกษา

ด้วยสำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ ได้รับแจ้งจากกรมวิเทศสหการ ว่าจังหวัดชิบะ ประเทศญี่ปุ่น จะคัดเลือกบุคลากรจากประเทศอาเซียนเพื่อเข้ารับการฝึกอบรม ในโครงการ The Overseas Technical Trainee Acceptance Programme of Chiba Prefecture ประจำปี 2004 รายละเอียดปรากฏตามเอกสารแนบ

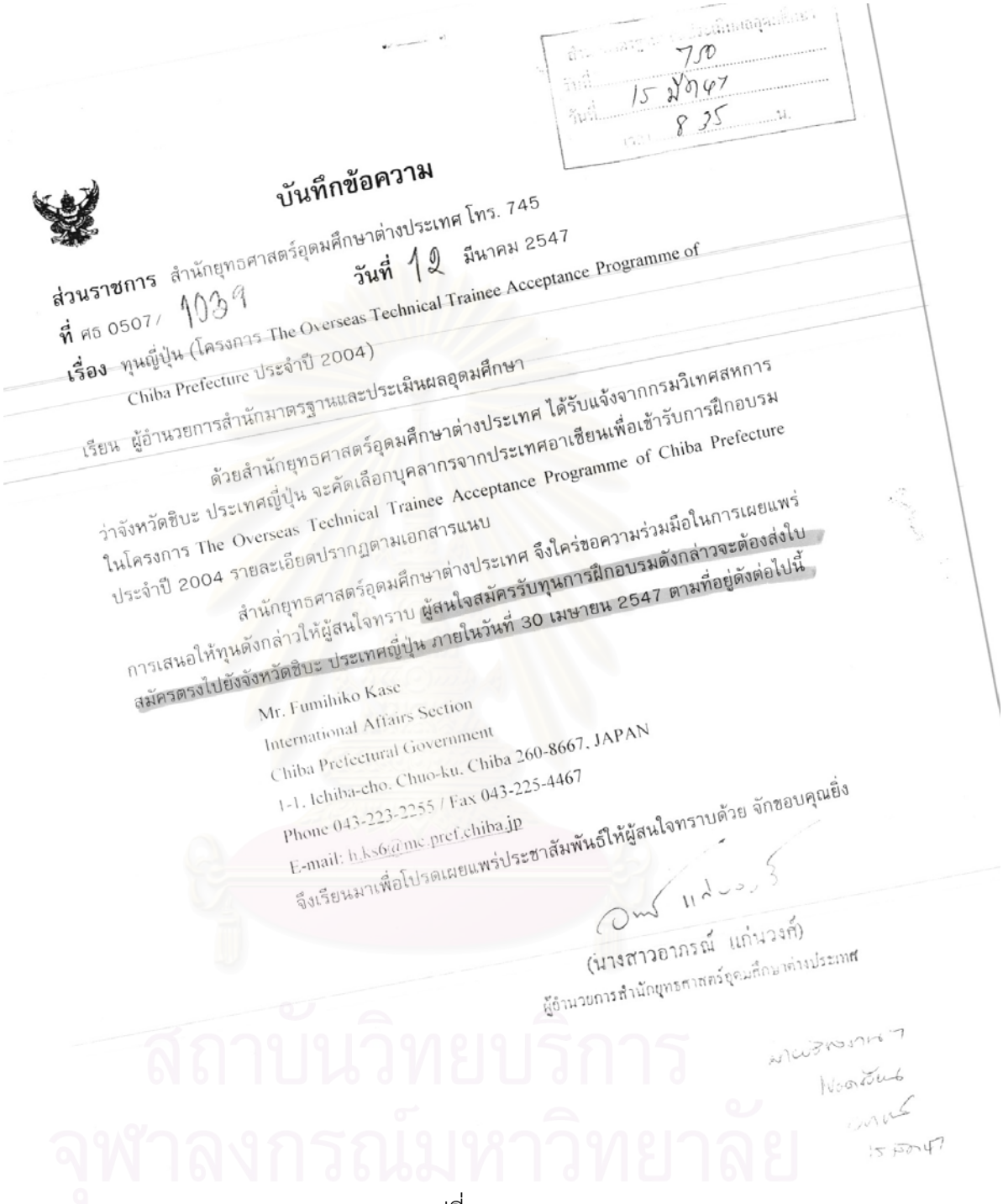
สำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ จึงใคร่ขอความร่วมมือในการเผยแพร่การเสนอให้ทุนดังกล่าวให้ผู้สนใจทราบ ผู้สนใจสมัครรับทุนการฝึกอบรมดังกล่าวจะต้องส่งใบสมัครตรงไปยังจังหวัดชิบะ ประเทศญี่ปุ่น ภายในวันที่ 30 เมษายน 2547 ตามที่อยู่ดังต่อไปนี้

Mr. Fumihiko Kase
International Affairs Section
Chiba Prefectural Government
1-1, Ichiba-cho, Chuo-ku, Chiba 260-8667, JAPAN
Phone 043-223-2255 / Fax 043-225-4467
E-mail: b.ks6@mc.pref.chiba.jp

จึงเรียนมาเพื่อโปรดเผยแพร่ประชาสัมพันธ์ให้ผู้สนใจทราบด้วย จักขอบคุณยิ่ง

(นางสาวอาภรณ์ แก่นวงศ์)
ผู้อำนวยการสำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ

รูปที่ ค.3.



บันทึกข้อความ

ส่วนราชการ สำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ โทร. 745
ที่ ศธ 0507/ 1039 **วันที่** 12 มีนาคม 2547
เรื่อง ทุนญี่ปุ่น (โครงการ The Overseas Technical Trainee Acceptance Programme of Chiba Prefecture ประจำปี 2004)
เรียน ผู้อำนวยการสำนักมาตรฐานและประเมินผลอุดมศึกษา

ด้วยสำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ ได้รับแจ้งจากกรมวิเทศสหการ ว่าจังหวัดชิบะ ประเทศญี่ปุ่น จะคัดเลือกบุคลากรจากประเทศอาเซียนเพื่อเข้ารับการฝึกอบรม ในโครงการ The Overseas Technical Trainee Acceptance Programme of Chiba Prefecture ประจำปี 2004 รายละเอียดปรากฏตามเอกสารแนบ

สำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ จึงใคร่ขอความร่วมมือในการเผยแพร่การเสนอให้ทุนดังกล่าวให้ผู้สนใจทราบ ผู้สนใจสมัครรับทุนการฝึกอบรมดังกล่าวจะต้องส่งใบสมัครตรงไปยังจังหวัดชิบะ ประเทศญี่ปุ่น ภายในวันที่ 30 เมษายน 2547 ตามที่อยู่ดังต่อไปนี้

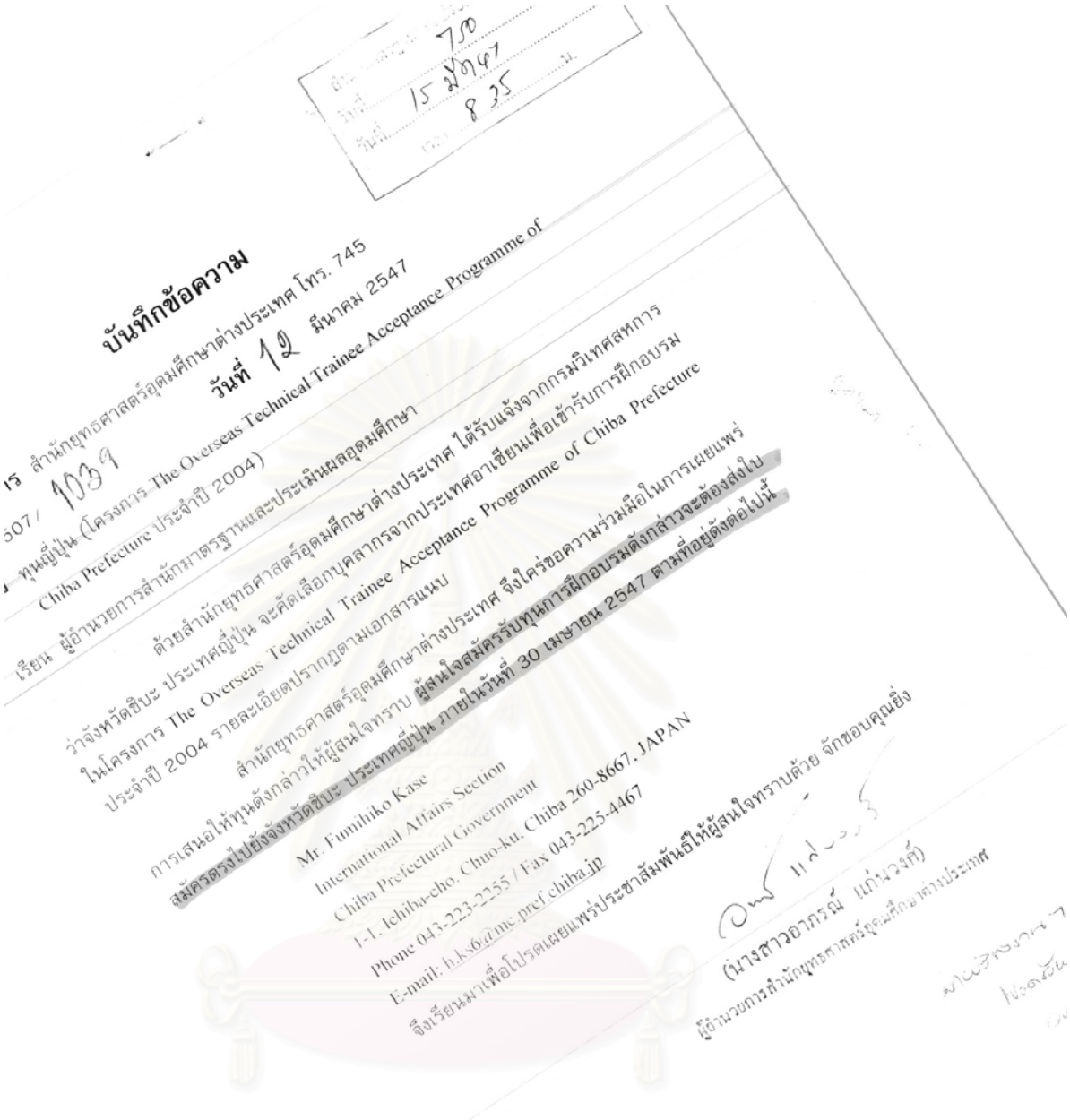
Mr. Fumihiko Kase
International Affairs Section
Chiba Prefectural Government
1-1, Ichiba-cho, Chuo-ku, Chiba 260-8667, JAPAN
Phone 043-223-2255 / Fax 043-225-4467
E-mail: h.ks6@mc.pref.chiba.jp

จึงเรียนมาเพื่อโปรดเผยแพร่ประชาสัมพันธ์ให้ผู้สนใจทราบด้วย จักขอบคุณยิ่ง

(นางสาวอาภรณ์ แก่นวงศ์)
ผู้อำนวยการสำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ ค.4.



บันทึกข้อความ

ที่ 507/ 1094 สำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ โทร. 745

วันที่ 12 มีนาคม 2547

เรื่อง ทุนญี่ปุ่น (โครงการ The Overseas Technical Trainee Acceptance Programme of Chiba Prefecture ประจำปี 2004)

เรียน ผู้อำนวยการสำนักมาตรฐานและประเมินผลอุดมศึกษา

ด้วยสำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ ได้รับแจ้งจากกรมวิเทศสหการ ว่าจังหวัดชิบะ ประเทศญี่ปุ่น จะคัดเลือกบุคลากรจากประเทศอาเซียนเพื่อเข้ารับการฝึกอบรม ในโครงการ The Overseas Technical Trainee Acceptance Programme of Chiba Prefecture ประจำปี 2004 รายละเอียดปรากฏตามเอกสารแนบ

สำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ จึงใคร่ขอความร่วมมือในการเผยแพร่การเสนอให้ทุนดังกล่าวให้ผู้สนใจทราบ ผู้สนใจสมัครรับทุนการฝึกอบรมดังกล่าวจะต้องส่งใบสมัครไปยังจังหวัดชิบะ ประเทศญี่ปุ่น ภายในวันที่ 30 เมษายน 2547 ตามที่อยู่ดังต่อไปนี้

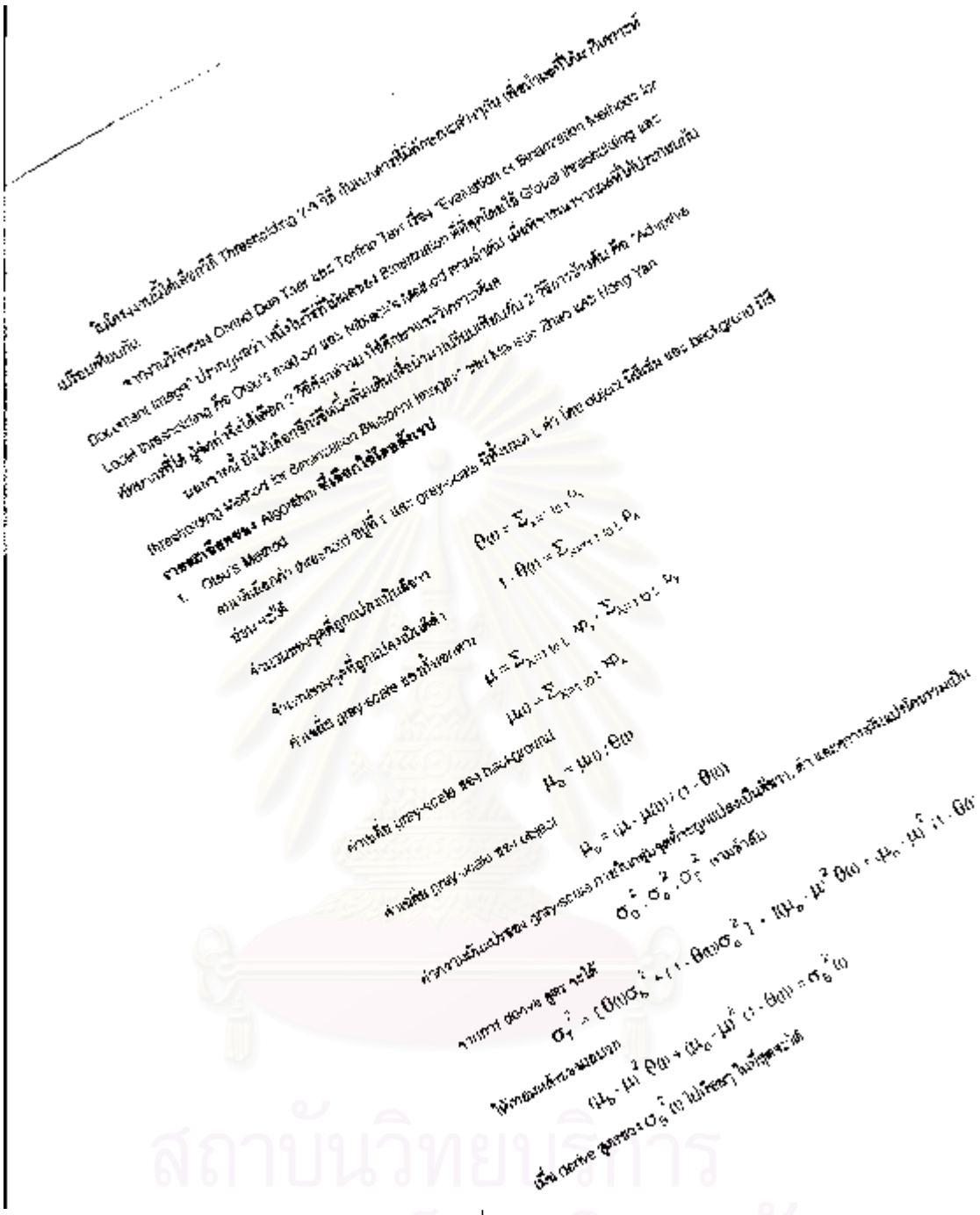
Mr. Fumihiko Kase
International Affairs Section
Chiba Prefectural Government
1-1, Ichiba-cho, Chuo-ku, Chiba 260-8667, JAPAN
Phone 043-223-2255 / Fax 043-225-4467
E-mail: h.k-6@mc.pref.chiba.jp

จึงเรียนมาเพื่อโปรดเผยแพร่ประชาสัมพันธ์ให้ผู้สนใจทราบด้วย จักขอบคุณยิ่ง

(นางสาวอากรณ์ แก่นวงศ์)
ผู้อำนวยการสำนักยุทธศาสตร์อุดมศึกษาต่างประเทศ

รูปที่ ค.5.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ ค.6.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นายเกรียงศักดิ์ เหล็กดี เกิดวันที่ 1 กรกฎาคม พ.ศ.2516 ที่จังหวัดพิษณุโลก เป็นบุตรนางมุ เหล็กดีและนายเหว่า เหล็กดี จบการศึกษาระดับประถมศึกษาที่โรงเรียนพิณพลราษฎร์ ตรงตั้งจิต 12 ระดับมัธยมศึกษาตอนต้นโรงเรียนพิษณุโลกพิทยาคม จังหวัดพิษณุโลก จบระดับปริญญาตรีจากสถาบันเทคโนโลยีพระจอมเกล้า เจ้าคุณทหารลาดกระบัง คณะครุศาสตร์อุตสาหกรรม สาขาอิเล็กทรอนิกส์และคอมพิวเตอร์ วิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาระดับปริญญาโท สาขาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย