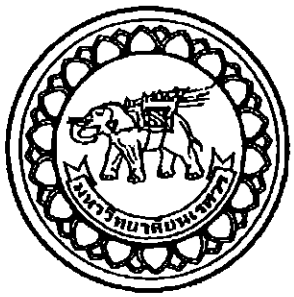


ระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความ

Converting Image of Thai Document to Text

นายทศพร ใจแก้ว รหัส 51361995

นายพสธร นามสง่า รหัส 51352060

ห้องสมุดคณะวิศวกรรมศาสตร์
วันที่รับ - 2 ก.ค. 2558
เลขทะเบียน 16280109
เลขเรียกหนังสือ ว.
มหาวิทยาลัยนเรศวร

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

ปีการศึกษา 2554

ใบรับรองปริญญานิพนธ์

หัวข้อโครงงาน ระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความ

ผู้ดำเนินโครงงาน นายทศพร ใจแก้ว รหัส 51361995

นายพสธร นามสง่า รหัส 51362060

อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. พนมขวัญ ริยะมงคล

สาขาวิชา วิศวกรรมคอมพิวเตอร์

ภาควิชา วิศวกรรมไฟฟ้าและคอมพิวเตอร์

ปีการศึกษา 2554

.....

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร อนุมัติให้โครงงานฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์

คณะกรรมการสอบโครงงานวิศวกรรม

.....ประธานกรรมการ

(อาจารย์รัฐภูมิ วรานุสาสน์)

.....กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร. พนมขวัญ ริยะมงคล)

.....กรรมการ

(ดร. พงศ์พันธ์ กิจสนาโยธิน)

.....กรรมการ

(ดร. สุรเดช จิตประไพกุลศาล)

หัวข้อโครงงาน	ระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความ
ผู้ดำเนินโครงงาน	นายทศพร ใจแก้ว รหัสนิสิต 51361995
	นายพสธร นามสง่า รหัสนิสิต 51362060
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร. พนมชวัญ ริยะมงคล
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
ภาควิชา	วิศวกรรมไฟฟ้าและคอมพิวเตอร์
ปีการศึกษา	2554

.....

บทคัดย่อ

ระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความที่พัฒนาขึ้น สามารถประมวลผลภาพของเอกสารและแปลงเป็นไฟล์ข้อความโดยถ่ายภาพจากกล้องเว็บแคมเป็นไฟล์ข้อความ จุดประสงค์เพื่อแก้ไขปัญหาเกี่ยวกับการรับเอกสารจากที่อื่นหรือไฟล์เอกสารได้สูญหายไป โดยมีขั้นตอนการประมวลผลขั้นต้นเพื่อกำจัดสัญญาณรบกวน การปรับปรุงภาพ การวิเคราะห์เอกสาร และการสกัดลักษณะเด่น โดยใช้ทฤษฎีของเฮาซดอรฟดิสแทนซ์มาใช้ในการรู้จำ จากนั้นตัวอักษรที่ได้จะถูกบันทึกเป็นไฟล์ข้อความ

ความถูกต้องของระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความมีค่าความถูกต้องของตัวอักษรมากกว่าร้อยละ 90 ข้อดีของระบบนี้คือสามารถแปลงภาพเอกสารเป็นไฟล์ข้อความและมีขั้นตอนการประมวลผลขั้นต้นที่ดีที่สามารถปรับปรุงภาพเอกสารได้ในสภาวะที่ไม่ดี

Project Title	Converting Image of Thai Document to Text
Name	Mr. Thossaporn Jaikaew ID. 51361995
	Mr. Possatorn Namsanga ID. 51362060
Project Advisor	Assistant Professor Panomkhawn Riyamongkol, Ph.D.
Major	Computer Engineering
Department	Electrical and Computer Engineering
Academic Year	2011

.....

Abstract

Converting image of Thai document to text file system has been developed. This system can process and convert image document which captured from webcam into text file. The purpose is to resolve typing problem when users receive a document from other sources or lose document files. The pre-processing process includes eliminating the interference, enhancing the image, analyzing character, and extraction features. Hausdorff distance has been used in the recognition process. Then, the output characters will be save to text file.

The accuracy of correct characters is more than 90 percent. The advantage of this system is the ability to convert the image to text file, and good pre-processing process that can enhance the image in bad condition.

กิตติกรรมประกาศ

โครงงานวิศวกรรมฉบับนี้เรียบร้อยสมบูรณ์ได้ เนื่องจากได้รับความอนุเคราะห์และสนับสนุนจาก ผู้ช่วยศาสตราจารย์ ดร.พนมขวัญ ริยะมงคล ซึ่งสละเวลาในการช่วยให้คำปรึกษา ให้คำแนะนำไต่ถามความก้าวหน้าของวิทยานิพนธ์มาโดยตลอด แนะนำขั้นตอนต่างๆในการทำการทดลอง เพื่อให้การทดลองเป็นไปอย่างรอบคอบและถูกต้อง

ขอขอบพระคุณคณะอาจารย์ทุกท่าน บิดามารดาและเพื่อนวิศวกรรมคอมพิวเตอร์ที่ให้การสนับสนุนช่วยเหลือในการทำการทดลอง และเป็นกำลังใจแก่คณะผู้ดำเนินโครงงาน ขอขอบพระคุณทุกๆท่านที่มิได้เอ่ยนามถึง ที่มีส่วนร่วมช่วยให้โครงงานนี้สำเร็จลุล่วงไปได้ด้วยดี

คณะผู้ดำเนินโครงงานวิศวกรรม

ทศพร ใจแก้ว

พสธร นามสง่า

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ก

บทคัดย่อภาษาอังกฤษ ..... ข

กิตติกรรมประกาศ ..... ค

สารบัญ ..... ง

สารบัญตาราง ..... ฉ

สารบัญภาพ ..... ช

บทที่ 1 บทนำ 1

1.1 ที่มาและความสำคัญของโครงงาน 1

1.2 วัตถุประสงค์ของโครงงาน 1

1.3 ขอบเขตของโครงงาน 2

1.4 ขั้นตอนการดำเนินงานของโครงงาน 2

1.5 แผนการดำเนินงานของโครงงาน 2

1.6 ผลที่คาดว่าจะได้รับจากโครงงาน 3

1.7 งบประมาณของโครงงาน 3

บทที่ 2 ทฤษฎีและหลักการที่เกี่ยวข้อง 4

2.1 การทำงานเกี่ยวกับการรับภาพและประมวลผลภาพ 4

2.1.1 วิธีการจัดเก็บข้อมูลภาพ 4

2.1.2 การนำเข้าข้อมูลภาพ 5

2.1.3 กระบวนการประมวลผลขั้นต้น 6

2.1.4 การวิเคราะห์เอกสาร 20

2.1.5 การสกัดลักษณะเด่น 21

2.1.6 ขั้นตอนการรู้จำ 23

2.2 ทฤษฎีเฮาซดอรฟดิสแทนซ์ 25

สารบัญ(ต่อ)

หน้า

บทที่ 3 วิธีการดำเนินงานโครงงาน 27
3.1 ขั้นตอนการรับภาพ 29
3.2 ขั้นตอนการประมวลผลขั้นต้น 31
3.2.1 การแปลงภาพในระดับเทา 31
3.2.2 การแปลงภาพเป็นขาวดำสองระดับ 31
3.2.3 การกรองสัญญาณรบกวน 32
3.3 ขั้นตอนการวิเคราะห์เอกสาร 34
3.3.1 Thinning Morphological image processing 34
3.3.2 การตัดแยกตัวอักษร สระ และวรรณยุกต์ 34
3.4 ขั้นตอนการสกัดลักษณะเด่น 37
3.4.1 การหาจำนวนหัวตัวอักษร 37
3.5 ขั้นตอนการรู้จำ 38

บทที่ 4 ผลการทดลองและการวิเคราะห์ 40
4.1 ผลการทดลองส่วนของขั้นตอนการรับภาพ 40
4.2 ผลการทดลองส่วนของการประมวลผลขั้นต้น 41
4.3 ผลการทดลองส่วนของขั้นตอนการวิเคราะห์เอกสาร 43
4.4 ผลการทดลองในส่วนของขั้นตอนการสกัดลักษณะเด่น 44
4.5 ผลการทดลองในส่วนของขั้นตอนการรู้จำ 45
4.5.1 ผลการทดลองในส่วนการรู้จำปกติภายใต้สภาวะควบคุม 45
4.5.1.1 ผลการทดลองการรู้จำตัวอักษรภาษาไทย 45
4.5.1.2 ผลการทดลองส่วนการรู้จำตัวอักษรภาพเอกสาร 46
4.5.1.3 ผลการทดลองส่วนการรู้จำตัวอักษรที่ได้รับสัญญาณรบกวน 47
4.5.2 ผลการทดลองการรู้จำเมื่อเทียบกับโปรแกรมอ่านไทย 48
4.5.3 ผลการทดลองการรู้จำที่เกี่ยวกับความทนทานของระบบ 51
4.5.3.1 ความทนทานของระบบกับภาพถ่ายที่องศาไม่ปกติ 51
4.5.3.2 ความทนทานของระบบกับภาพเอกสารในแนวเอียง 52
4.5.3.3 ความทนทานของระบบกับตัวอักษรภาษาไทยรูปแบบอื่น 54

สารบัญ(ต่อ)

หน้า

บทที่ 5 สรุปผลการดำเนินงานของโครงงาน 56

5.1 สรุปผลการทดลองของโครงงาน 57

5.2 ปัญหาและอุปสรรคของโครงงาน 58

5.3 ข้อเสนอแนะและแนวทางการพัฒนา 58

เอกสารอ้างอิง 59

ภาคผนวก ก. การใช้โปรแกรม 60

ประวัติผู้ดำเนินโครงงาน 74

สารบัญตาราง

ตารางที่ หน้า

4.1 แสดงผลการรู้จำตัวอักษรภาษาไทยพยัญชนะ สระ และตัวเลข 45

4.2 ร้อยละของความถูกต้องของการแปลงเป็นไฟล์เอกสาร 46

4.3 แสดงภาพเอกสารที่ถูกสัญญาณรบกวน 47

4.4 แสดงการเปรียบเทียบไฟล์ข้อความจากโปรแกรมอ่านไทยกับระบบ 49

4.5 แสดงร้อยละของความถูกต้องของการเปรียบเทียบกับโปรแกรมอ่านไทย 50

4.6 แสดงการผลค่าความถูกต้องเมื่อภาพเอกสารหมุนในมุมก้มและมุมเงย 51

4.7 แสดงผลการทดลองความถูกต้องของภาพเอกสารในแนวเอียงที่องศาต่างๆ 53

4.8 แสดงการเปรียบเทียบความถูกต้องเมื่อใช้ตัวอักษรชนิด Angsana New 54

4.9 แสดงการเปรียบเทียบความถูกต้องเมื่อใช้ตัวอักษรชนิด Cordia New 54

4.10 แสดงการเปรียบเทียบความถูกต้องเมื่อใช้ตัวอักษรชนิด Tahoma 55

5.1 แสดงผลการทดลองของระบบภายใต้สภาวะควบคุมในกรณีต่างๆ 57

5.2 แสดงผลการเทียบกับโปรแกรมอ่านไทยภายใต้สภาวะควบคุมในกรณีต่างๆ 57

5.3 แสดงปัญหาและอุปสรรคและแนวทางการแก้ไข 58

สารบัญรูป

รูปที่		หน้า
2.1	ภาพแบบบิตแมพ	4
2.2	ภาพแบบบิตแมพ(ต่อ)	5
2.3	ภาพที่เกิดจากการได้รับสัญญาณรบกวน	6
2.4	รูปแบบ 8- neighborhood และแบบ 4-neighborhood ของการคัดกรองโดยพื้นที่	7
2.5	แสดงการสแกนภาพของวัตถุ	7
2.6	แสดงการตรวจสอบความถูกต้องของพิกเซล	8
2.7	แสดงตัวอย่างการให้ป้ายชื่อในแต่ละเงื่อนไข	8
2.8	แสดงภาพหลังการสแกนครั้งแรก และการยุบรวมเซตค่าสมมูล	9
2.9	แสดงภาพหลังการปรับค่าป้ายชื่อ	9
2.10	แสดงองค์ประกอบของโครงสร้างและภาพตั้งต้นของการขยายขนาด	10
2.11	แสดงขั้นตอนการยูเนียนแม่แบบของการขยายขนาด	11
2.12	แสดงองค์ประกอบของโครงสร้างและภาพตั้งต้นของการกร่อน	12
2.13	แสดงเหตุการณ์ที่แม่แบบกับข้อมูลภาพไม่เหมือนกัน	12
2.14	แสดงเหตุการณ์ที่แม่แบบกับข้อมูลภาพเหมือนกัน	12
2.15	แสดงเหตุการณ์ที่แม่แบบข้อมูลภาพมีความเหมือนกันอีกเหตุการณ์หนึ่ง	13
2.16	แสดงภาพผลลัพธ์สุดท้ายที่ผ่านการกร่อน	13
2.17	แสดงการทำโอเพนนิ่งเพื่อกำจัดสัญญารบกวนและเส้นเชื่อม	13
2.18	ภาพการทำให้บาง	14
2.19	แม่แบบขนาด 3x3 เพื่อใช้ในการพิจารณาค่าของพิกเซล	14
2.20	กฎในการตัดสินใจสำหรับการลบพิกเซลที่พิจารณา	15
2.21	การหมุนแม่แบบ 180 องศา	16
2.22	การเทรสโฮลแบบปกติกับการเทรสโฮลแบบแบ่งส่วนเทรสโฮล	18
2.23	ภาพแสดงการทำกรองภาพโดยใช้ค่ากลาง	19
2.24	แสดงภาพผลลัพธ์จากการลบกัน และปรับค่าระดับเทา	19
2.25	แสดงภาพผลลัพธ์จากการเทรสโฮล	20
2.26	แสดงการโปรเจคชันในแนวนอน (Horizontal Projection)	20
2.27	แสดงการโปรเจคชันในแนวตั้ง (Vertical Projection)	20
2.28	แสดง 4-Connected	21

สารบัญรูป(ต่อ)

รูปที่		หน้า
2.29	แสดง 8-Connected	21
2.30	แสดงการเทสีแบบ 4-Connected	22
2.31	แสดงการเทสีแบบ 8-Connected	22
2.32	แสดงภาพอินพุตขนาด 3×3 และภาพเอาต์พุตขนาด 2×2	24
2.33	แสดงจุดกึ่งกลางของภาพเอสต์พุตบนภาพอินพุต	24
2.34	การให้ค่าแก่ภาพเอาต์พุตจากการเลือกค่าจากตำแหน่งที่ใกล้ที่สุด	24
3.1	โครงสร้างของระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความ	27
3.2	ขั้นตอนการทำงานของระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความ	28
3.3	ภาพเอกสารภาษาไทยแบบ TH Sarabun PSK	29
3.4	แม่แบบที่ใช้ในการรู้จำ	29
3.5	การแปลงภาพเอกสารสี 24 บิตเป็นภาพระดับเทา	31
3.6	การแปลงภาพเอกสารระดับเทาเป็นภาพขาวดำสองระดับ	31
3.7	การกำจัดสัญญาณรบกวนโดยวิธีการคัดกรองโดยพื้นที่	32
3.8	การทำโอเพนนิ่งเพื่อลดขอบที่ไม่เรียบของภาพ	32
3.9	ผังขั้นตอนการเตรียมภาพและขั้นตอนการประมวลผลขั้นต้น	33
3.10	การทำตัวอักษรบางลง	34
3.11	ภาพแสดงการโปรเจคชันแนวนอนเฟส 1	34
3.12	ภาพแสดงการโปรเจคชันแนวตั้งเฟส 1	35
3.13	ภาพแสดงการโปรเจคชันแนวนอนเฟส 2	35
3.14	ภาพแสดงการโปรเจคชันแนวตั้งเฟส 2	35
3.15	ผังแสดงขั้นตอนการวิเคราะห์เอกสาร	36
3.16	การเทสี (Flood Fill) เพื่อหาตำแหน่งตัวอักษร	37
3.17	การทำการคัดกรองโดยพื้นที่เพื่อนับจำนวนหัวตัวอักษร	37
3.18	ภาพแสดงการรู้จำตัวอักษรโดยทฤษฎีเฮาซดอรฟดิสแทนซ์	38
3.19	ภาพแสดงการรู้จำตัวอักษรโดยทฤษฎีเฮาซดอรฟดิสแทนซ์(ต่อ)	39
4.1	ภาพที่ผ่านการเทรสโฮลจากกล้องที่ไม่มีโฟกัส	40
4.2	ภาพที่ผ่านการเทรสโฮลจากกล้องที่มีโฟกัส	40
4.3	ภาพแสดงการเทรสโฮลด้วยวิธีต่างๆ	41

สารบัญรูป(ต่อ)

รูปที่ หน้า

4.4 ภาพที่ไม่ได้ผ่านการทำการคัดกรองโดยพื้นที่มาทำให้บาง 42

4.5 ภาพที่ผ่านการทำการคัดกรองโดยพื้นที่มาทำให้บาง 42

4.6 ภาพตัวอักษรที่ยังไม่ได้ทำโอเพนนิ่งแล้วนำไปทำให้บาง 42

4.7 ภาพตัวอักษรที่ผ่านการทำโอเพนนิ่งแล้วนำไปทำให้บาง 42

4.8 ภาพตัวอักษรที่ไม่ได้ผ่านขั้นตอนการทำให้บางมาทำการโปรเจคชัน 43

4.9 ภาพตัวอักษรที่ผ่านขั้นตอนการทำให้บางมาทำการโปรเจคชัน 43

4.10 การสกัดลักษณะเด่นที่ผลตรงตามความเป็นจริงและแม่แบบที่เหลือ 44

4.11 การสกัดลักษณะเด่นที่ผลไม่ตรงตามความเป็นจริงและแม่แบบที่เหลือ 44

4.12 แสดงผลการรู้จำพยัญชนะ สระ ตัวเลข ภาษาไทย 45

4.13 แสดงภาพที่ถ่ายในแนว 0 องศา 51

4.14 กราฟแสดงผลการทดลองความถูกต้องของภาพเอกสารในมุมก้มและมุมเงย 52

4.15 แสดงภาพที่ถ่ายในแนวปกติและความถูกต้องในการแปลงร้อยละ 100 52

4.16 กราฟแสดงผลการทดลองความถูกต้องของภาพเอกสารในแนวเอียงที่องศาต่างๆ 54

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของโครงงาน

ปัจจุบันหน่วยงานราชการ องค์กรเอกชน และรัฐวิสาหกิจ ได้มีการนำเอารูปแบบตัวอักษรมาตรฐานแห่งชาติ 13 รูปแบบมาใช้ในงานด้านเอกสาร โดยทั่วไปในงานด้านเอกสารนั้นจะมีรูปแบบการใช้ประโยคหรือข้อความที่เหมือนกัน โดยในกรณีที่มีการรับเอกสารจากที่อื่นหรือไฟล์เอกสารต้นฉบับหายไปซึ่งจะต้องมีการพิมพ์ใหม่ โครงงานนี้จึงจัดทำขึ้นเพื่อเป็นเครื่องมือที่ช่วยอำนวยความสะดวกในงานด้านเอกสาร โดยอาศัยการรับภาพจากกล้องเว็บแคมและนำมาแปลงเป็นไฟล์ตัวอักษร ซึ่งสามารถตัดอักษรหรือประโยคที่ต้องการมาใช้โดยไม่ต้องเสียเวลาในการพิมพ์

การรู้จำ (Recognize) คืองานวิจัยแนวหนึ่งที่มีความสำคัญ และเป็นแขนงหนึ่งของระบบปัญญาประดิษฐ์ (Artificial Intelligence) การรู้จำมีหลายชนิด เช่น การรู้จำวัตถุ การรู้จำเสียง การรู้จำภาพ และการรู้จำตัวอักษรเป็นต้น

การรู้จำตัวอักษรเป็นการนำเอารูปตัวอักษรที่ได้มาทำการวิเคราะห์ว่าเป็นตัวอักษรใด การวิเคราะห์สามารถใช้ได้หลายแบบ เช่นเฮาซดรอฟดิสแทนซ์ (Hausdroff Distance) ลำดับการเปลี่ยนแปลงของเส้น (Stroke Changing Sequence) การวิเคราะห์ด้วยเพนนาไลซ์ ดิสคริมิแนนท์ (Penalised Discriminant Analysis) การใช้ฟัซซี่ลอจิค (Fuzzy Logic) และการใช้โครงข่ายประสาทเทียม (Neural Network) เป็นต้น

จากคุณสมบัติของเฮาซดรอฟดิสแทนซ์ (Hausdroff Distance) ผู้ทำโครงงานจึงสนใจที่จะนำมาทำการใช้ในงานด้านนี้ เพื่อให้ได้ผลการรู้จำที่มีประสิทธิภาพมากที่สุด โดยอ้างอิงการใช้รูปแบบตัวอักษรชนิด TH Saraban PSK มาใช้ในการศึกษา ซึ่ง รูปแบบนี้เป็นรูปแบบทั่วไปในงานด้านเอกสาร

1.2 วัตถุประสงค์ของโครงงาน

เพื่อพัฒนาโปรแกรมสำหรับการแปลงข้อความภาษาไทยจากเอกสารเป็นไฟล์ข้อความ

1.3 ขอบเขตของโครงงาน

1.3.1 ตัวอักษรที่ใช้ศึกษาเป็นตัวอักษรภาษาไทยชนิด TH Saraban PSK

1.3.2 ตัวอักษรที่ใช้ศึกษาเป็นตัวอักษรตัวตรงขนาดปกติ ไม่หนา ไม่เอียง

1.3.3 ขนาดตัวอักษรที่รู้จำคือตั้งแต่ 16 หน่วยขึ้นไป

1.3.4 กล้องเว็บแคมสามารถปรับโฟกัสได้

1.3.5 ความละเอียดของกล้องเว็บแคมที่ใช้ต้องมีความละเอียด 640 x 480 ขึ้นไป

1.3.6 อัตราบรรทัดต่อรูปภาพอยู่ที่ 4-5 บรรทัดต่อ 1 รูปที่ความละเอียด 640 x 480 ขึ้นไป

1.4 ขั้นตอนการดำเนินงานของโครงงาน

1.4.1 ศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับโครงงาน

1.4.2 ศึกษาทฤษฎีการทำงานแบบเฮาซดอรฟดิสแทนซ์

1.4.3 พัฒนาระบบการรับภาพโดยใช้กล้องเว็บแคม

1.4.4 ออกแบบโปรแกรมและส่วนประกอบ

1.4.5 พัฒนาโปรแกรมแปลงข้อความภาษาไทยจากเอกสารเป็นไฟล์ข้อความ

1.4.6 ทดสอบการใช้งานและแก้ไขข้อบกพร่อง

1.4.7 ทดสอบการทำงานของโปรแกรมและสรุปผลการทดสอบโปรแกรม

1.5 แผนการดำเนินงานของโครงงาน

กิจกรรม	ระยะเวลา										
	ปี 2554							ปี 2555			
	มิ.ย.	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.	เม.ย.
1. ศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับโครงงาน	■	■									
2. ศึกษาทฤษฎีเฮาซดอรฟดิสแทนซ์		■	■								
3. พัฒนาระบบการรับภาพโดยใช้กล้องเว็บแคม			■	■	■						
4. ออกแบบโปรแกรมและส่วนประกอบ					■	■	■				
5. เขียนโปรแกรมและพัฒนาโปรแกรม						■	■	■			
6. ทดสอบการใช้งานและแก้ไขข้อบกพร่อง								■	■		
7. สรุปผลการทำโครงงานและจัดทำรายงาน										■	■

1.6 ผลที่คาดว่าจะได้รับจากโครงงาน

1.6.1 ได้โปรแกรมสำหรับการแปลงข้อความภาษาไทยจากเอกสารเป็นไฟล์ข้อความ

1.6.2 สามารถประยุกต์ใช้กับข้อความภาษาอังกฤษ

1.7 งบประมาณของโครงงาน

1.7.1 ค่าใช้จ่ายในการทำรายงาน	เป็นจำนวนเงิน	1000	บาท
1.7.2 ค่าอุปกรณ์ในการทำโครงงาน	เป็นจำนวนเงิน	1000	บาท
	รวมเป็นเงินทั้งสิ้น	2000	บาท

หมายเหตุ ขออนุมัติถัวเฉลี่ยทุกรายการ

มหาวิทยาลัยนเรศวร

บทที่ 2
ทฤษฎีและหลักการที่เกี่ยวข้อง

ทฤษฎีและหลักการต่าง ๆ ที่จะกล่าวต่อไปนี้เป็นส่วนประกอบของระบบการแปลงข้อความจากเอกสารเป็นไฟล์ข้อความโดยใช้ทฤษฎีเฮาซดอรฟดิสแทนซ์ (Hausdorff Distance) โดยการใช้กล้องเว็บแคมที่มีความละเอียด 640 x 480 ในการรับภาพและแปลงออกมาให้เป็นไฟล์ข้อความ

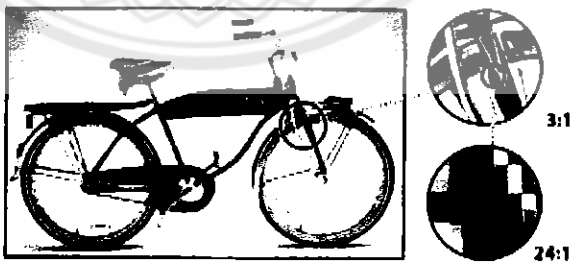
2.1 การทำงานเกี่ยวกับการรับภาพและประมวลผล [1]

2.1.1 วิธีการจัดเก็บข้อมูลภาพ

2.1.1.1 ภาพบิตแมพ (Bitmap) [5]

บิตแมพเป็นภาพที่ประกอบขึ้นด้วยจุดสีต่างๆ ที่มีจำนวนคงที่ตายตัวตามการสร้างภาพที่มีความละเอียดของภาพต่างกันไป หากขยายภาพบิตแมพจะเห็นว่ามีลักษณะเป็นตารางเล็กๆ ซึ่งแต่ละบิตคือ ส่วนหนึ่งของข้อมูลคอมพิวเตอร์

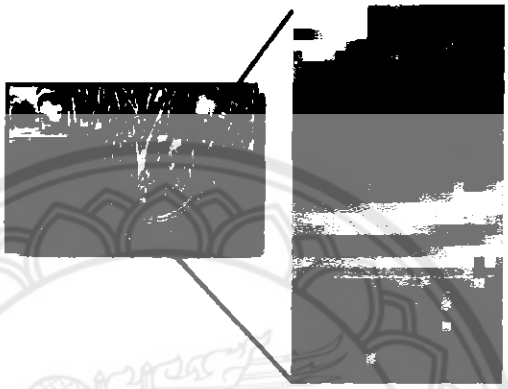
เนื่องจากบิตแมพมีค่าพิกเซลจำนวนคงที่จึงทำให้มีข้อจำกัดในเรื่องการขยายขนาดภาพ การเปลี่ยนขนาดภาพทำโดยเพิ่มหรือลดพิกเซลจากที่มีอยู่เดิม เมื่อขยายภาพให้ใหญ่ขึ้นความละเอียดของภาพจึงลดลง และถ้าเพิ่มค่าความละเอียดมากขึ้นก็จะทำให้ไฟล์มีขนาดใหญ่และเปลืองเนื้อที่หน่วยความจำมากขึ้นตามไปด้วย ภาพที่ขยายใหญ่ขึ้นจะมองเห็นเป็นตารางสี่เหลี่ยมเรียงต่อกันทำให้ขาดความสวยงาม



รูปที่ 2.1 ภาพแบบบิตแมพ

ภาพแบบบิตแมพจึงเหมาะสำหรับงานกราฟิกในแบบที่ต้องการให้แสงเงาในรายละเอียดเป็นไฟล์ที่เหมาะกับการทำงานกับภาพเหมือนจริงประเภทภาพถ่าย เพราะบิตแมพมีช่องสัญญาณพิเศษ เรียกว่า ช่องสัญญาณแอลฟา ซึ่งเป็น 32 บิตหรือสีสมจริง เช่น ภาพที่นำมาใช้กับ Photoshop จะเป็นภาพเหมือน ภาพถ่าย เพราะไฟล์ที่ได้จาก Photoshop เป็นบิตแมพในขณะที่ไฟล์ที่

สร้างจาก Illustrator จะเหมือนการ์ตูนหรือภาพเขียน เพราะเป็นไฟล์แบบเวกเตอร์นอกจากนี้ยังเหมาะสำหรับภาพที่ต้องการระบายสี สร้างสี หรือกำหนดสีที่ต้องการความละเอียดและสวยงาม ไฟล์ภาพแบบบิตแมพในระบบวินโดวส์คือ ไฟล์ที่มีนามสกุล . BMP, .PCX. , .TIF, .GIF, .JPG, .MSP, .PCD เป็นต้น สำหรับโปรแกรมที่ใช้สร้างกราฟิกแบบนี้คือ โปรแกรม Paint ต่างๆ เช่น Paintbrush, Photoshop, Photostyler เป็นต้น



รูปที่ 2.2 ภาพแบบบิตแมพ (ต่อ)

ชนิดภาพบิตแมพ

1. ภาพไบนารีหรือภาพขาวดำเป็นภาพที่ใช้สีเพียง 2 สีคือดำและขาวในการแสดงผล
2. ภาพระดับเทาเป็นภาพที่แสดงผลด้วยสีเทาระดับความเข้ม 0 – 255 ระดับสี
3. สีเป็นภาพสี รองรับการใช้สีได้สูงสุด 24 บิต

2.1.2 การนำเข้าข้อมูลภาพ (Image Acquisition) [4]

ส่วนมากจะคิดว่าส่วนมากภาพที่อยู่ในคอมพิวเตอร์จะเป็นรูปที่ได้มาจากกล้องวีดีโอ ดังเช่นรูปภาพที่เห็นอยู่ในโทรทัศน์ ซึ่งในความเป็นจริงนั้นรูปภาพที่ได้มานั้นมาจากอุปกรณ์หลายๆ อย่าง มีวิธีการจัดเก็บที่แตกต่างกัน

2.1.2.1 เว็บแคม (Webcam)

เว็บแคม (Webcam ย่อมาจาก Web Camera) คือ กล้องวีดีโอที่ถ่ายทอดภาพนิ่งหรือภาพวิดีโอผ่านระบบเครือข่าย เว็บไซต์ โปรแกรม ถือเป็นอุปกรณ์นำข้อมูล (อินพุต) ที่สามารถจับภาพเคลื่อนไหวให้ไปปรากฏในจอภาพ และสามารถส่งภาพเคลื่อนไหวหรือภาพนิ่งนี้ไปให้คนอีกฟากหนึ่งเห็นตัวเราเคลื่อนไหวได้เหมือนอยู่ต่อหน้า ปัจจุบันมีทั้งแบบที่เชื่อมต่อกับคอมพิวเตอร์ผ่านสายยูเอสบี และเชื่อมต่อแบบไร้สาย

2.1.2.2 รูปแบบภาพสี (Color Images)

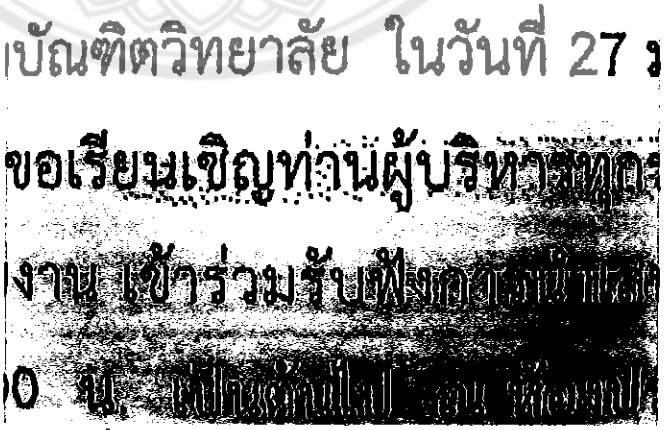
การนำภาพแบบมีสีเข้าสู่คอมพิวเตอร์จะเปรียบเสมือนการนำภาพเข้าสู่คอมพิวเตอร์ 3 รูปในเวลาเดียวกัน ซึ่งรายละเอียดขึ้นอยู่กับระบบชนิดของสี ในระบบ RGB จะใช้ความเข้มเป็น สีแดง สีเขียว และสีน้ำเงิน ในกล้องถ่ายรูปจะทำการแยกแถบสีดังกล่าวออกจากกันจึงทำให้ได้รูปภาพ 3 รูปเป็นของสีแดง สีเขียว และสีน้ำเงินตามลำดับ

2.1.3 กระบวนการประมวลผลขั้นต้น (Pre-Processing) [2]

ก่อนที่โปรแกรมจะสามารถบอกได้ว่ารูปภาพที่ส่งเข้าไปประกอบด้วยตัวอักษรอะไรบ้าง จำเป็นจะต้องผ่านขั้นตอนที่สำคัญหลายขั้นตอน ขั้นตอนนี้มักถูกเรียกรวมกันว่ากระบวนการประมวลผลขั้นต้น (Pre-Processing) ซึ่งเป็นขั้นตอนในการปรับแต่งและจัดเตรียมข้อมูลให้เหมาะสมกับขั้นตอนการรู้จำต่อไป ขั้นตอนเหล่านี้มีความสำคัญต่อประสิทธิภาพโดยรวมของระบบ เพราะหากมีความผิดพลาดเกิดขึ้นในส่วนนี้ ก็จะส่งผลกระทบไปยังส่วนถัดไปของระบบด้วยขั้นตอนการประมวลผลเบื้องต้นที่สำคัญได้แก่

2.1.3.1 การกรองสัญญาณรบกวน (Noise Filtering)

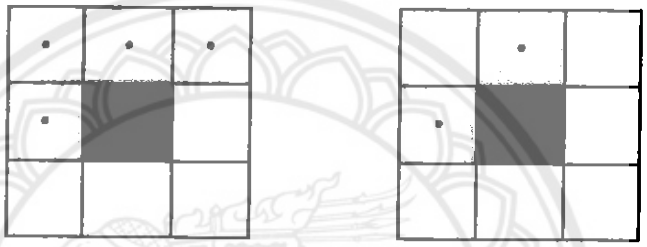
การกรองสัญญาณรบกวนมีจุดประสงค์เพื่อปรับปรุงส่วนของรูปภาพที่เป็นสิ่งแปลกปลอมอันไม่พึงประสงค์ออกไป โดยสัญญาณรบกวนที่เกิดขึ้นส่วนใหญ่มักจะมาจากคุณภาพของเอกสารต้นฉบับที่นำมาทำการอ่าน ซึ่งต้นเหตุสำคัญที่ทำให้ความถูกต้องของโปรแกรมลดลง จึงจำเป็นที่จะต้องจัดการกับส่วนเกินเหล่านี้ออกไปให้ได้มากที่สุดเท่าที่จะเป็นไปได้ แต่ยังไม่มีวิธีการใดที่รับรองได้ว่าสามารถจัดการกับสัญญาณรบกวนได้ โดยสมบูรณ์ ดังนั้นส่วนการรู้จำจะต้องมีความทนทานต่อการรบกวนเหล่านี้ได้พอสมควร



รูปที่ 2.3 ภาพที่เกิดจากการได้รับสัญญาณรบกวน

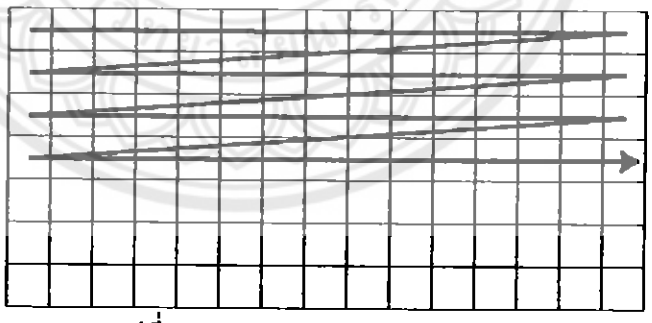
2.1.3.1.1 การให้ป้ายชื่อกับวัตถุ (Region Labeling)

การให้ป้ายชื่อกับวัตถุที่เชื่อมต่อกันเป็นวิธีการหนึ่งในการนับจำนวนวัตถุ หรือในการให้ชื่อแก่วัตถุในภาพ ซึ่งภาพที่ว่านี้จะเป็นภาพขาวดำสองระดับ การคัดกรองวัตถุโดยพื้นที่นั้นเป็นวิธีการที่นิยมใช้กันมากในการประมวลผลภาพ (Digital Image Processing) เพราะหลังจากการทำการคัดกรองวัตถุโดยพื้นที่แล้วนั้นจะได้คุณลักษณะต่างๆ ของวัตถุนั้นมาด้วย อาทิเช่น พื้นที่ของวัตถุนั้น และตำแหน่ง เป็นต้น เพื่อนำคุณลักษณะต่างๆนั้นมาใช้ประโยชน์ต่อไป การให้ชื่อแก่วัตถุนั้นจะมีอยู่ 2 ลักษณะ คือ แบบ 8-neighborhood และ แบบ 4-neighborhood ซึ่งทั้งสองแบบจะมีรูปแบบในการตรวจสอบป้ายชื่อที่แตกต่างกันดังรูป

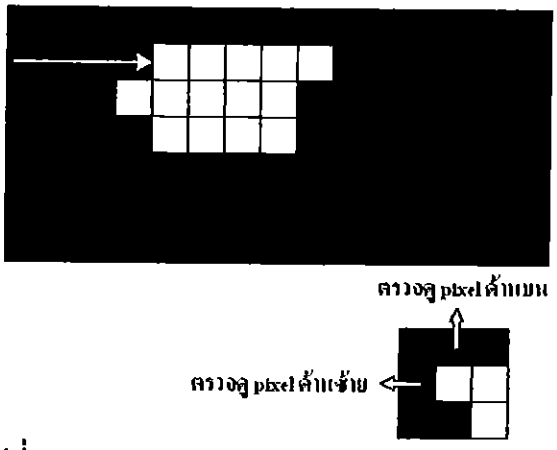


รูปที่ 2.4 รูปแบบ 8- neighborhood และแบบ 4- neighborhood สำหรับการคัดกรองโดยพื้นที่

หลักการของการให้ชื่อแก่วัตถุนั้น จะทำการสแกนภาพ 2 รอบ ในการสแกนครั้งแรกนั้นจะเป็นการให้ป้ายชื่อแก่วัตถุ ซึ่งจะทำการสแกนจากซ้ายไปขวา และบนลงล่าง ดังรูปที่ 2.5



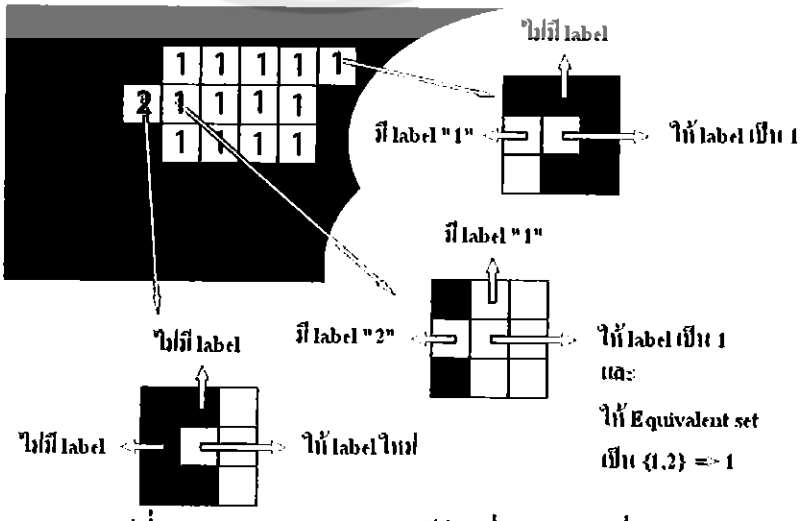
รูปที่ 2.5 แสดงการสแกนภาพของวัตถุ



รูปที่ 2.6 แสดงการตรวจสอบความถูกต้องของพิกเซล

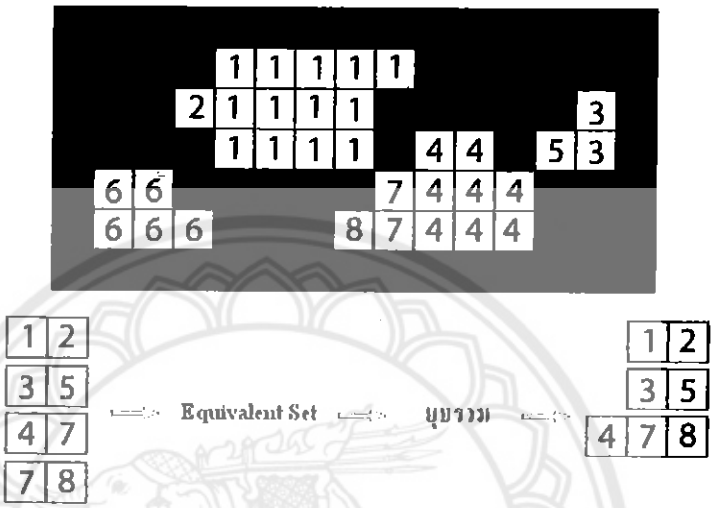
ในการสแกนนั้นจะพิจารณาเฉพาะพิกเซลที่มีค่าเป็น 1 (สีขาว) เท่านั้น จากนั้นจะทำการตรวจดูพิกเซลด้านบน และพิกเซลด้านซ้าย(สำหรับแบบ 4-neighborhood) ดังรูปที่ 2.6 ซึ่งเป็น พิกเซลที่ทำการสแกนผ่านมาแล้วว่า

1. ถ้าพิกเซลด้านบน และพิกเซล ด้านซ้าย ไม่มีป้ายชื่อจะทำการให้ป้ายชื่อแก่พิกเซลที่กำลังพิจารณาอยู่ด้วยป้ายชื่อใหม่
2. ถ้าพิกเซลด้านบน หรือพิกเซลด้านซ้าย พิกเซลใดพิกเซลหนึ่ง มีป้ายชื่อแล้ว จะทำการให้ป้ายชื่อแก่พิกเซลที่กำลังพิจารณาอยู่ ด้วยป้ายชื่อนั้น
3. ถ้าพิกเซลด้านบน และพิกเซลด้านซ้าย ทั้งสองพิกเซลมีป้ายชื่อแล้ว จะทำการให้ ป้ายชื่อแก่พิกเซลที่กำลังพิจารณาอยู่ด้วยป้ายชื่อที่มีค่าน้อยสุดของทั้งสองพิกเซลแล้วทำการเก็บค่าไว้ในเซตค่าสมมูล (เซตค่าสมมูลจะใช้ในการบอกว่าป้ายชื่อต่างกันนี้ คือวัตถุเดียวกัน)



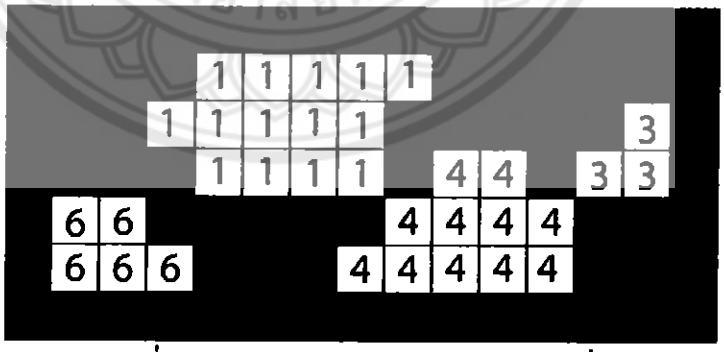
รูปที่ 2.7 แสดงตัวอย่างการให้ป้ายชื่อในแต่ละเงื่อนไข

หลังจากการสแกนครั้งแรกในภาพขาวดำสองระดับ ก็จะได้วัตถุที่มีป้ายชื่อซึ่งในแต่ละวัตถุอาจจะมีป้ายชื่ออยู่หลายค่า ดังรูปที่ 2.8 ดังนั้นในการสแกนครั้งที่สอง จึงเป็นการสแกนเพื่อปรับค่าป้ายชื่อในวัตถุ แต่ก่อนทำการปรับค่านั้น ต้องทำการยุบรวมเซตค่าสมมูลก่อน เพื่อที่จะได้รู้ว่าป้ายชื่อใดคือวัตถุเดียวกันบ้าง



รูปที่ 2.8 แสดงภาพหลังการสแกนครั้งแรก และการยุบรวมเซตค่าสมมูล

หลังจากการสแกนครั้งที่สองนั้นจะได้ ภาพขาวดำสองระดับที่แต่ละวัตถุมีเพียงป้ายชื่อเดียวดังรูปที่ 2.9



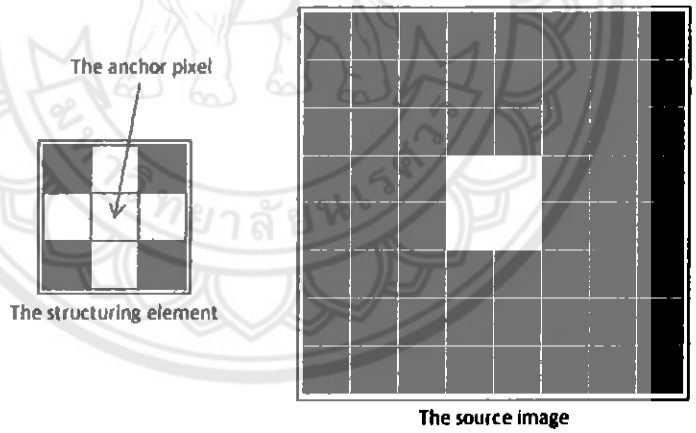
รูปที่ 2.9 แสดงภาพหลังการปรับค่าในป้ายชื่อ

2.1.3.2 การเปลี่ยนแปลงโครงสร้างของภาพ (Morphological image processing)

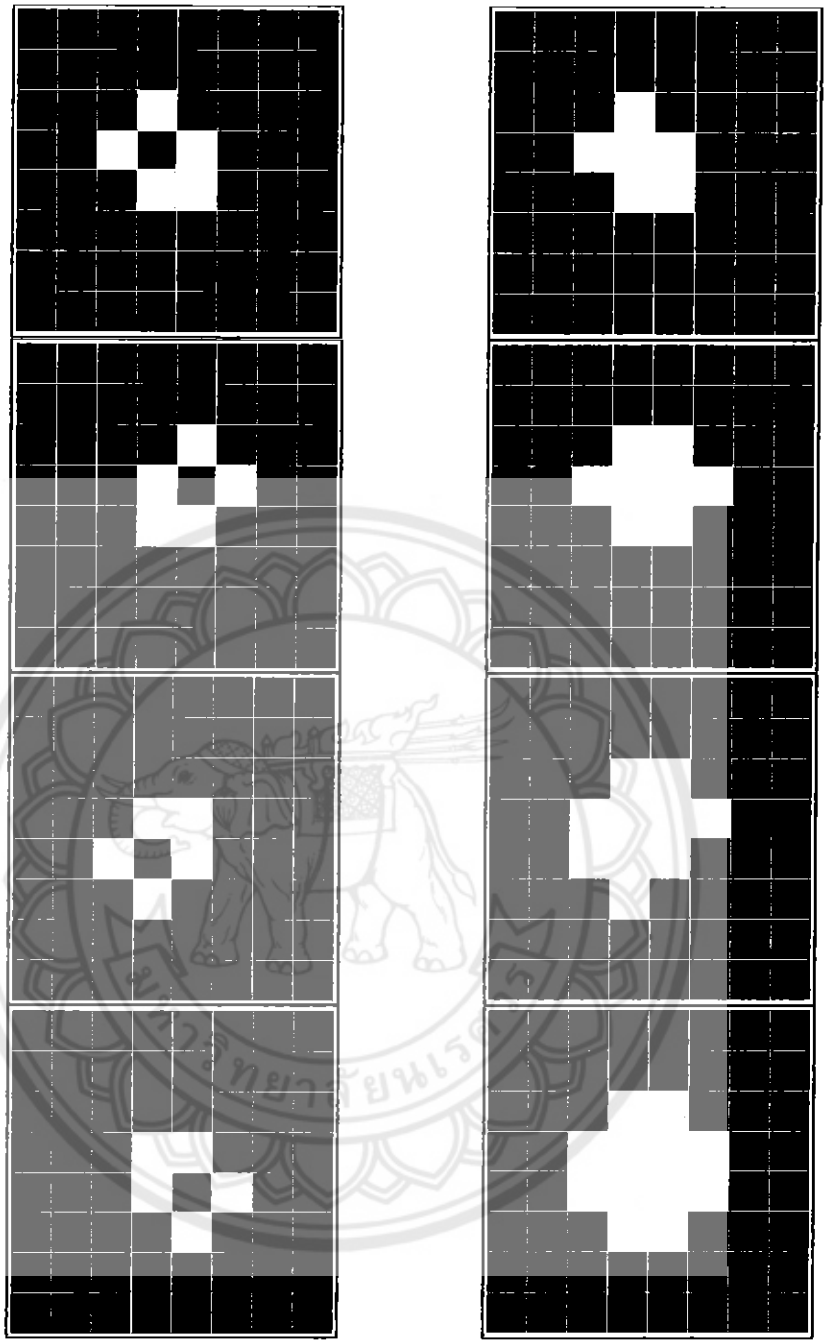
การประมวลผลภาพโดยการเปลี่ยนแปลงลักษณะรูปร่างหรือโครงสร้างของภาพ เป็นอีกหนึ่งวิธีการที่มักนิยมใช้กันในขั้นตอนการประมวลผลขั้นต้น ภาพที่ใช้ในการประมวลผลภาพด้านรูปร่างหรือโครงสร้างนั้นส่วนใหญ่จะเป็นภาพขาวดำสองระดับซึ่งมีโอเปอเรชันที่ได้ใช้ในโครงงานนี้ได้แก่

- Dilation (การขยายขนาด)
- Erosion (การย่อขนาด)
- Opening (การทำย่อขนาดแล้วตามด้วยการขยายขนาด)
- Thinning (การทำให้บาง)

การขยายขนาด (Dilation) คือ กระบวนการที่ทำให้วัตถุในภาพขยายออก ด้วยการใช้แม่แบบขนาดและรูปร่างต่างๆ แล้วแต่จะกำหนด จากนั้นนำเอาแม่แบบไล่เทียบไปบนภาพเฉพาะพิกเซลที่มีค่าเป็น 1 ทั่วทั้งภาพ โดยใช้จุดกึ่งกลางของแม่แบบเป็นจุดอ้างอิง และภาพผลลัพธ์จะมีค่าเท่ากับการทำการยูเนียนแม่แบบกับข้อมูลภาพดังรูปที่ 2.11



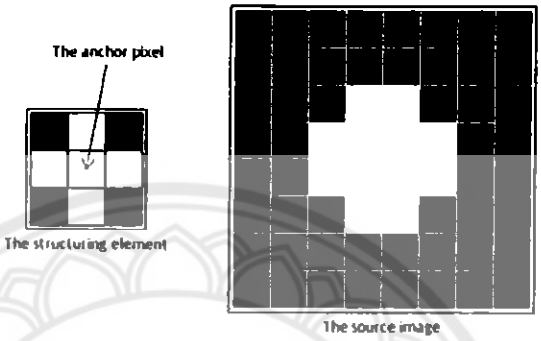
รูปที่ 2.10 แสดงองค์ประกอบของโครงสร้าง(ซ้าย) และภาพตั้งต้น(ขวา) ของการขยายขนาด



รูปที่ 2.11 แสดงขั้นตอนการยูเนียนแม่แบบของการขยายขนาด

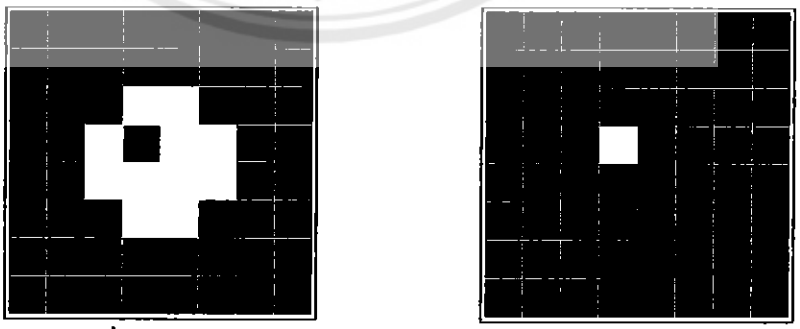
จากรูปที่ 2.11 แสดงวิธีการขยายขนาด โดยฝั่งซ้ายจะแสดงตำแหน่งของแม่แบบที่ไล่เทียบไปบนภาพตั้งต้น และฝั่งขวาแสดงผลลัพธ์จากการยูเนียนแม่แบบกับภาพตั้งต้นที่ละพิกเซลจนได้ผลลัพธ์สุดท้าย

การย่อขนาด คือ กระบวนการที่ทำให้วัตถุในภาพย่อลง ด้วยการใช้แม่แบบขนาดและรูปร่างต่างๆ แล้วแต่จะกำหนด จากนั้นนำเอาแม่แบบไล่เทียบไปบนภาพเฉพาะพิกเซลที่มีค่าเป็น 1 ทั่วทั้งภาพ โดยใช้จุดกึ่งกลางของแม่แบบเป็นจุดอ้างอิง จากนั้นจะทำการเปรียบเทียบข้อมูลภาพกับแม่แบบถ้าข้อมูลภาพมีความเหมือนกันกับแม่แบบจะให้ พิกเซล บนภาพผลลัพธ์ที่ตรงกับจุดอ้างอิงนั้น มีค่าเป็น 1 ดังรูปภาพตัวอย่าง

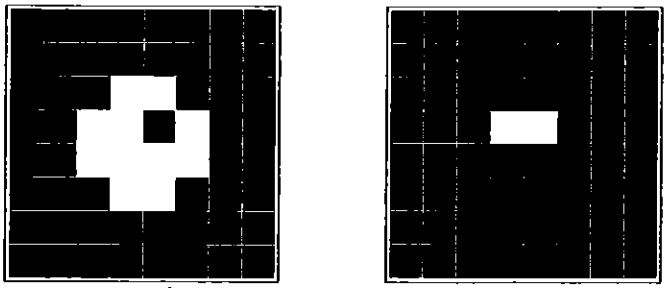


รูปที่ 2.12 แสดงองค์ประกอบของโครงสร้าง (ซ้าย) และภาพตั้งต้น (ขวา) ของการย่อขนาด

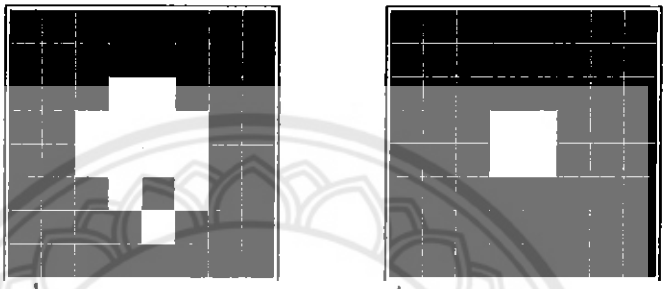
รูปที่ 2.13 แสดงเหตุการณ์ที่แม่แบบกับข้อมูลภาพไม่เหมือนกัน



รูปที่ 2.14 แสดงเหตุการณ์ที่แม่แบบกับข้อมูลภาพเหมือนกัน



รูปที่ 2.15 แสดงเหตุการณ์ที่แม่แบบกับข้อมูลภาพมีความเหมือนกันอีกเหตุการณ์หนึ่ง



รูปที่ 2.16 แสดงภาพผลลัพธ์สุดท้ายที่ทำการย่อขนาดเสร็จแล้ว

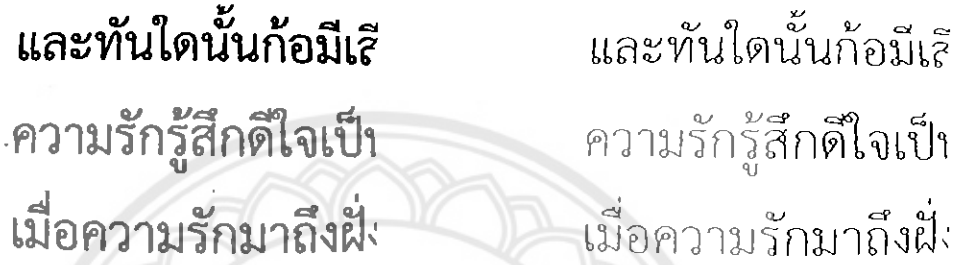
โอเพนนิ่ง (Opening) คือ กระบวนการที่ทำย่อขนาดก่อนแล้วตามด้วยการขยายขนาดโดยใช้แม่แบบชุดเดียวกัน การโอเพนนิ่งนั้นมีประโยชน์ในการกำจัดเส้นเชื่อมระหว่างวัตถุออกไปได้ และยังสามารถนำไปประยุกต์ใช้กับการกำจัดสัญญาณรบกวนกับภาพที่เป็นภาพตัวอักษรได้อีกด้วย ดังแสดงในรูป

รูปที่ 2.17 แสดงการทำโอเพนนิ่งเพื่อกำจัดสัญญารบกวนและเส้นเชื่อม
(ซ้าย) ภาพที่ได้รับสัญญาณรบกวนและมีเส้นเชื่อม
(ขวา) ภาพที่ถูกกำจัดสัญญาณรบกวน และเส้นเชื่อมของตัวอักษรออกด้วยแม่แบบขนาด 2×2

การที่จะลบเส้นเชื่อมหรือสัญญาณรบกวนออกได้นั้น แม่แบบที่ใช้นั้นจะต้องมีความกว้างมากกว่าเส้นเชื่อมหรือสัญญาณรบกวนนั้นๆ

2.1.3.2.1 การทำให้บาง (Thinning)

การทำให้บางเป็นเทคนิคหนึ่งที่นิยมใช้กันในขั้นตอนการประมวลผลขั้นต้นของระบบการรู้จำตัวอักษร ซึ่งเป็นการทำให้ตัวอักษรนั้นบางลงเรื่อยๆ โดยการลบขอบของตัวอักษรนั้นๆออก จนตัวอักษรนั้นเหลือความหนาเพียง 1 พิกเซล ในขณะที่สัดส่วนของตัวอักษร และรูปร่างของตัวอักษรนั้นยังต้องคงเดิม นอกจากนี้ยังเป็นการช่วยลดจำนวนของ พิกเซล ที่จะต้องนำมาใช้ในขั้นตอนการประมวลผล หรือการรู้จำอีกต่อไป



รูปที่ 2.18 การทำภาพเอกสารให้บาง

ขั้นตอนในการทำภาพให้บางนั้นจะทำกันเป็นรอบๆ (Iteration) โดยในแต่ละรอบจะแบ่งออกเป็น 2 ขั้นตอนย่อย ดังนี้

1. ในขั้นตอนย่อยที่ 1 นั้น จะพิจารณาลบพิกเซลนั้นๆ ออกเมื่อ เป็นจริงตามเงื่อนไข G1,G2 และ G3 ทั้งหมด
2. ในขั้นตอนย่อยที่ 2 นั้น จะพิจารณาลบพิกเซลนั้นๆ ออกเมื่อ เป็นจริงตามเงื่อนไข G1,G2 และ G3' ทั้งหมด

เงื่อนไข G1 : $X_H(p) = 1$ เมื่อ

$$X_H(p) = \sum_{i=1}^{4} b_i \tag{2-1}$$

$$b_i = \begin{cases} 1, if\ x_{2i-1} = 0\ and\ (x_{2i} = 1\ or\ x_{2i+1} = 1) \\ 0, \text{otherwise} \end{cases} \tag{2-2}$$

X_4	X_3	X_2
X_5	p	X_1
X_6	X_7	X_8

รูปที่ 2.19 แม่แบบขนาด 3x3 เพื่อใช้ในการพิจารณาค่าของพิกเซล

เมื่อ $X_1, X_2, X_3, \ldots, X_8$ คือ ค่าของ พิกเซลข้างเคียง p , p คือพิกเซลที่พิจารณา

เงื่อนไข G2 :

$$2 \leq \min[n_1(p), n_2(p) \leq 3 \tag{2-3}$$

เมื่อ

$$n_1(p) = \sum_{k=1}^{4} x_{2k-1} \vee x_{2k} \tag{2-4}$$

$$n_2(p) = \sum_{k=1}^{4} x_{2k} \vee x_{2k+1} \tag{2-5}$$

เงื่อนไข G3 : $(x_2 \vee x_3 \vee \bar{x}_8) \wedge x_1 = 0$ (2-6)

เงื่อนไข G3' : $(x_6 \vee x_7 \vee \bar{x}_4) \wedge x_5 = 0$ (2-7)

เมื่อทำครบ 2 ขั้นตอนย่อยแล้วจะถือว่าจบ 1 รอบ (Iteration) จากนั้นจะทำวนซ้ำในลักษณะเดิมจากขั้นตอนย่อยที่หนึ่ง และสอง ไปเรื่อยๆ เป็นรอบๆ ทำไปจนกระทั่งรูปภาพนั้นไม่มีการเปลี่ยนแปลงจากเดิม และภาพนั้นก็จะเป็นผลลัพธ์ของการทำให้บาง

หรือเมื่อพิจารณาจากเงื่อนไขในขั้นตอนย่อยที่ 1 และขั้นตอนย่อยที่ 2 แล้วสามารถที่จะสรุปได้ดังนี้

ขั้นตอนย่อยที่หนึ่งนั้นจะใช้แม่แบบทั้งหมด 8 แบบดังแสดงในรูปที่ 2.16 เป็นกฎในการตัดสินใจสำหรับการลบพิกเซลที่พิจารณาอยู่นั้นๆออก

x	1	x
1	1	0
x	x	0

x	0	0
1	1	0
x	1	x

x	1	0
1	1	0
x	1	1

1	0	0
1	1	1
x	1	1

0	0	0
x	1	1
x	1	1

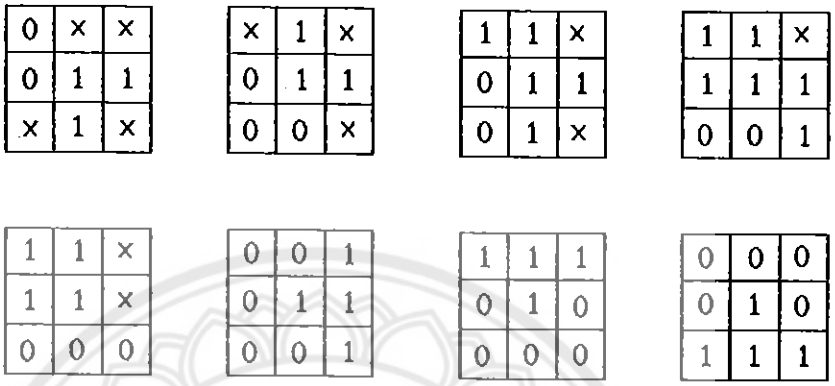
1	0	0
1	1	0
1	0	0

0	0	0
0	1	0
1	1	1

1	1	1
0	1	0
0	0	0

รูปที่ 2.20 กฎในการตัดสินใจสำหรับการลบพิกเซลที่พิจารณา

ขั้นตอนย่อยที่สองนั้นภาพอินพุต คือ ภาพที่ผ่านการลบพิกเซลที่เป็นขอบจากขั้นตอนย่อยที่หนึ่งมาแล้ว ในขั้นตอนนี้จะใช้แม่แบบที่คล้ายกับแม่แบบที่ใช้ในขั้นตอนย่อยแรก เพียงแต่ต้องทำการหมุนแม่แบบไป 180 องศา ก่อน ดังแสดงในรูปที่ 2.17



รูปที่ 2.21 การหมุนแม่แบบ 180 องศา

เมื่อ 1 คือ พิกเซลที่เป็นสีขาว

0 คือ พิกเซลที่เป็นสีดำ

X คือ พิกเซลที่เป็นสีขาว หรือสีดำ ก็ได้ (Don't care)

โดยแต่ละขั้นตอนย่อยนั้นจะใช้ แม่แบบขนาด 3×3 รูปแบบต่างๆ มาใช้ในการเปรียบเทียบค่าของพิกเซลข้างเคียง และถ้าหากค่าของพิกเซลข้างเคียงนั้นมีค่าเป็นไปตามแม่แบบที่กำหนดนั้นแล้ว แสดงว่าพิกเซลที่กำลังพิจารณาอยู่เป็นขอบของตัวอักษรแล้วต้องพิจารณาลบออก

2.1.3.3 การปรับเป็นภาพขาวดำสองระดับ (Binarization) [6]

การทำภาพขาวดำสองระดับ เป็นการแปลงข้อมูลภาพระดับเทาให้เป็นภาพขาวดำ ซึ่งมีค่าระดับความสว่างเป็น 2 ค่าคือ 1 กับ 0 และเรียกภาพแบบนี้ว่า ภาพไบนารี (Binary) โดยอาศัยจุดเทรสโฮล (Threshold) ในการกำหนดระดับขาวหรือดำ ดังสมการที่ (2-8) ภาพขาวดำสองระดับช่วยให้การวิเคราะห์ภาพ และการคำนวณเป็นไปได้ง่าย และสะดวก

$$g(x,y) = \begin{cases} 1 => f(x,y) \geq T \\ 0 => f(x,y) < T \end{cases} \quad (2\text{-}8)$$

โดยที่ f(x,y) คือ ภาพระดับเทาของตำแหน่ง x,y

T คือ ค่าเทรสโฮล

g(x,y) คือ ภาพขาวดำสองระดับที่ได้ของตำแหน่ง x,y

ลักษณะของภาพขาวดำสองระดับที่ได้ขึ้นกับค่าเทรสโฮล ดังนั้นถ้าค่าเทรสโฮลมีค่าสูงภาพที่ได้จะเป็นสีดำมาก ในทางกลับกัน ถ้าค่าเทรสโฮลมีค่าต่ำมาก ภาพที่ได้จะเป็นสีขาวมาก ด้วยเหตุนี้ การกำหนดค่าเทรสโฮลที่เหมาะสมจะช่วยให้ภาพขาวดำสองระดับที่ได้มีความถูกต้องใกล้เคียงกับภาพจริงมากที่สุด

2.1.3.3.1 การเทรสโฮลแบบอัตโนมัติ (Otsu's Threshold)

เป็นวิธีการเลือกค่าเทรสโฮลแบบอัตโนมัติวิธีการหนึ่ง โดยการแบ่งภาพที่ถูกเทรสโฮลแล้วออกเป็นกลุ่มของพิกเซล 2 กลุ่มคือ กลุ่มพื้นหน้าและกลุ่มพื้นหลัง จากนั้นจะทำการคำนวณหาค่าเทรสโฮลที่เหมาะสมที่สุด ที่เมื่อแบ่งภาพออกเป็น 2 กลุ่มแล้วมีค่าความแปรปรวนที่ต่ำที่สุด

$$\sigma_w^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t) \quad (2\text{-}9)$$

เมื่อ $\sigma_w^2(t)$ คือ ค่าความแปรปรวน

ω_i คือ ค่าน่าจะเป็นของแต่ละกลุ่มพิกเซลที่ถูกแบ่งออกโดยค่าเทรสโฮล t

σ_i^2 คือ ค่าความแปรปรวนของแต่ละกลุ่ม

โดยค่า $\omega_1(t)$ และ $\omega_2(t)$ หาได้จาก

$$\omega_1(t) = \sum_{i=0}^{t} p_i \quad (2\text{-}10)$$

$$\omega_2(t) = \sum_{i=t+1}^{L} p_i \quad (2\text{-}11)$$

เมื่อ p_i คือค่าความน่าจะเป็นของค่าระดับเทาที่ i

t คือค่าเทรสโฮล

L คือ ระดับเทาสุดท้าย

และค่า $\sigma_1^2(t)$ และ $\sigma_2^2(t)$ หาได้จาก

$$\sigma_1^2(t) = \sum_{i=0}^{t} p_i(x_i - \mu_1)^2 \quad (2\text{-}12)$$

$$\sigma_2^2(t) = \sum_{i=t+1}^{L} p_i(x_i - \mu_2)^2 \quad (2\text{-}13)$$

เมื่อ μ_i คือ ค่า *mean*

x_i คือ ค่าระดับเทาที่ i

โดยที่ค่า μ_1 และ μ_2 ค่าได้จาก

$$\mu_1 = \sum_{i=0}^{t} p_i \cdot x_i \quad (2\text{-}14)$$

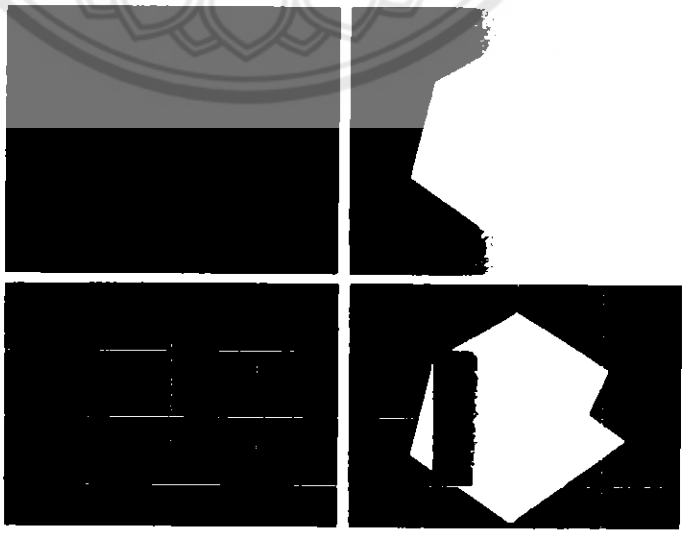
$$\mu_2 = \sum_{i=t+1}^{L} p_i \cdot x_i \quad (2\text{-}15)$$

2.1.3.3.2 การเทรสโฮลแบบปรับค่า (Adaptive Threshold)

การเทรสโฮลแบบปรับค่าเป็นการทำเทรสโฮลที่เหมาะกับภาพที่มีสภาพแสงที่ไม่เรียบหรือมีเงาบนภาพ การทำเทรสโฮลแบบปรับค่านั้นมีหลายวิธี แต่ในที่นี้จะกล่าวถึง 2 วิธีเท่านั้น คือการปรับค่าเทรสโฮลโดยการแบ่งภาพ และการปรับค่าเทรสโฮลโดยใช้ค่ากลาง

การปรับค่าเทรสโฮลโดยการแบ่งภาพ (Adaptive thresholding by segmentation)

เป็นการทำเทรสโฮลกับโดยการแบ่งภาพออกเป็นหลายส่วนย่อยๆ แล้วทำการเทรสโฮลกับภาพที่ละส่วน โดยที่แต่ละส่วนนั้นอาจจะใช่ค่าเทรสโฮลที่แตกต่างกัน



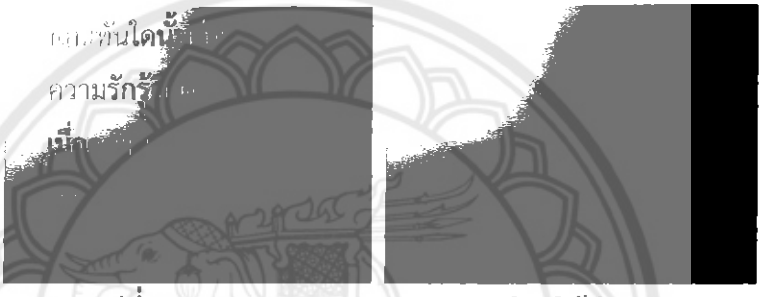
รูปที่ 2.22 การเทรสโฮลแบบปกติการทำเทรสโฮลกับโดยการแบ่งภาพ

จากรูปที่ 2.22 แสดงการเปรียบเทียบให้เห็นว่าการทำเทรสโฮลแบบปกติทีเดียวทั้งภาพกับภาพที่มีสภาพแสงไม่สม่ำเสมอนั้น จะทำให้รายละเอียดของวัตถุในภาพนั้นหายไป แต่เมื่อทำการแบ่งภาพที่มีสภาพแสงไม่สม่ำเสมอออกเป็นส่วนๆ แล้วทำการเทรสโฮลโดยใช้ค่าเทรสโฮลต่างๆ นั้น จะทำให้ได้รายละเอียดของวัตถุในภาพเพิ่มขึ้นจากเดิม

การปรับค่าเทรสโฮลโดยใช้ค่ากลาง (Adaptive thresholding by median)

การทำเทรสโฮลแบบนี้นั้นเป็นการทำเทรสโฮลที่เหมาะกับภาพที่เป็นตัวหนังสือที่มีสภาพแสงไม่สม่ำเสมอ โดยหลักการจะอาศัยการทำการกรองภาพโดยใช้ค่ากลางซึ่งรูปแบบที่ใช้นั้นควรจะมีขนาดใหญ่กว่าความหนาของตัวอักษร โดยมีขั้นตอนดังนี้

1. ทำการกรองภาพโดยใช้ค่ากลางกับภาพตั้งต้น

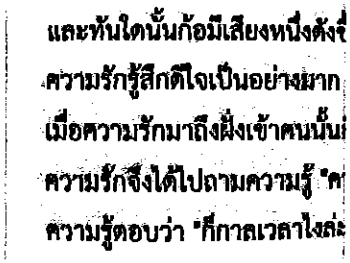


รูปที่ 2.23 ภาพแสดงการทำกรองภาพโดยใช้ค่ากลาง
(ซ้าย) ภาพตั้งต้น (ขวา) ภาพที่ผ่านการทำการกรองภาพโดยใช้ค่ากลาง

จากรูปที่ 2.23 ขวา สังเกตได้ว่าหลังจากการทำการกรองภาพโดยใช้ค่ากลางด้วยรูปแบบขนาด 51×51 นั้น จะทำให้ได้ภาพที่มีลักษณะเหมือนเป็นภาพพื้นหลังของตัวอักษร

2. นำภาพมาลบกัน

โดยใช้ภาพตั้งต้นเป็นตัวตั้ง แล้วลบออกด้วยภาพที่ผ่านการทำการกรองภาพโดยใช้ค่ากลาง จากนั้นทำการปรับค่าระดับเทาให้อยู่ระหว่าง 0-255



รูปที่ 2.24 แสดงภาพผลลัพธ์จากการลบกัน และปรับค่าระดับเทา

จากรูปที่ 2.24 สังเกตได้ว่าผลลัพธ์ที่ได้จากการลบกัน และผ่านการปรับค่าระดับเทาแล้วนั้นจะทำให้ได้ภาพที่เป็นตัวหนังสือที่ตัดสภาพของแสงที่ไม่สม่ำเสมอออกไป

3. การทำเทรสโฮล

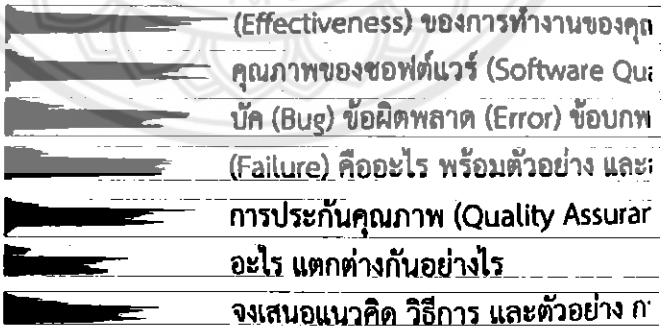
และทันใดนั้นก้อมีเสียงหนึ่งดังขึ้
ความรักรู้สึกดีใจเป็นอย่างมาก
เมื่อความรักมาถึงฝั่งเข้าคนนั้นแ
ความรักจึงได้ไปถามความรู้ "ค
ความรู้ตอบว่า "ก็กาลเวลาไงล่ะ

รูปที่ 2.25 แสดงภาพผลลัพธ์จากการเทรสโฮล

ในการทำเทรสโฮลนั้นจะทำการเทรสโฮลด้วยค่า 0 แต่ก่อนการทำเทรสโฮลนั้นอาจจะนำภาพมีลบกับค่าคงที่ใดๆ ก่อนก็ได้ตามที่เหมาะสม แล้วจึงทำการเทรสโฮลด้วยค่า 0 ซึ่งผลลัพธ์สุดท้ายที่ได้นั้นจะเป็นภาพขาวดำโดยมีพื้นหลังเป็นสีขาว และตัวอักษรเป็นสีดำ

2.1.4 การวิเคราะห์เอกสาร (Document Analysis) [1]

ขั้นตอนการวิเคราะห์เอกสารเป็นขั้นตอนในการตัดแยกบรรทัด และตัดแยกออกเป็นตัวอักษรกับวรรณยุกต์และสระ ในการตัดแยกนั้นจะใช้วิธีการที่เรียกว่า การหาค่าโปรเจคชันตามแนวนอน (Horizontal Projection) โดยการนับค่าจุดสีดำ ในภาพตามแนวนอน เพื่อให้รู้ว่าแต่ละบรรทัดนั้นอยู่ตำแหน่งใด และการหาค่าโปรเจคชันในแนวตั้ง (Vertical Projection) โดยทำการนับค่าจุดสีดำ ในภาพตามแนวตั้ง เพื่อให้รู้ว่าแต่ละตัวอักษรนั้นอยู่ตำแหน่งใดบ้าง



รูปที่ 2.26 แสดงการโปรเจคชันในแนวนอน (Horizontal Projection)

รูปที่ 2.27 แสดงการโปรเจคชันในแนวตั้ง (Vertical Projection)

2.1.5 การสกัดลักษณะเด่น (Feature Extraction) [5][6]

การสกัดลักษณะสำคัญเป็นการดึงเอาโครงสร้างพื้นฐานที่สำคัญของตัวอักษรนั้นออกมา โดยโครงสร้างพื้นฐานที่ว่าจะต้องมีการกำหนดไว้ก่อนว่าจะมีอะไรบ้าง มีการนิยามอย่างไร ตัวอย่างเช่น สำหรับภาษาไทยอาจกำหนดว่าตัวอักษรภาษาไทยทั้งหมดประกอบด้วยโครงสร้างพื้นฐานคือ เส้นตรง(แนวตั้ง/นอน)เส้นเอียง หัว(วงกลม) ส่วนโค้ง ส่วนเว้า จุดแตกกิ่ง จุดตัด เป็นต้น

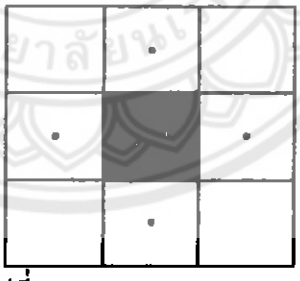
2.1.5.1 การเทสี (Flood Fill)

การเทสีเป็นวิธีการที่ใช้ในการเทสีลงบนพื้นที่ พื้นที่หนึ่งที่มีค่าสีเดียวกัน ตัวอย่างที่เห็นได้ชัดคือ เครื่องมือถังสีในโปรแกรม Paint เมื่อเลือกสีแล้วกดเทลงจุดๆ หนึ่ง จะเห็นได้ว่าพื้นที่ที่ติดกันและมีสีเดียวกันกับจุดเริ่มจะกลายเป็นสีใหม่ที่ถูกเทลงไป หรืออย่างเช่นในเกม Minesweeper เมื่อกดลงบนพื้นที่ที่ไม่มีระเบิดและไม่มีตัวเลข เกมจะทำการเปิดพื้นที่ที่ไม่มีระเบิดและตัวเลขที่อยู่ติดกันออกมา

การทำเทสีนั้นจะต้องการอินพุตอยู่ 3 ตัว ได้แก่

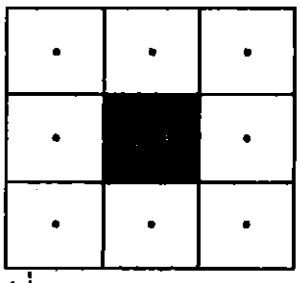
1. จุดเริ่มต้น (x, y)
2. สีที่ต้องการระบาย
3. สีของพื้นที่ที่ต้องการระบาย

การทำการเทสีนั้นจะมีการพิจารณาจุดถัดไปที่จะทำการให้สีอยู่ 2 แบบ คือ แบบ 4-Connected และแบบ 8-Connected โดยที่การพิจารณาแบบ 4-Connected นั้นจะมีการพิจารณา 4 จุดรอบข้าง ดังรูปที่ 2.24



รูปที่ 2.28 แสดง 4-Connected

ในขณะที่แบบ 8-Connected นั้นจะมีการพิจารณา 8 จุดรอบข้าง ดังรูปที่ 2.29



รูปที่ 2.29 แสดง 8-Connected

เมื่อทำการเทสีแบบ 4-Connected แล้ว จะมีบางจุดที่จะไม่สามารถเข้าถึงได้ด้วยลักษณะการพิจารณาของแบบ 4-Connected เอง ดังรูปที่ 2.30

ก. ข.

รูปที่ 2.30 แสดงการเทสีแบบ 4-Connected

เมื่อ พิกเซล ตรงกลางที่มีจุดสีแดง (รูป ก.) คือจุดเริ่มต้น แล้วสีแดง คือ สีที่ต้องการระบาย และสีขาว คือ สีของพื้นที่ที่ต้องการระบาย และเมื่อทำการเทสีแบบ 4-Connected แล้วจะได้ผลลัพธ์ดังรูป ข. แต่เมื่อทำการเทสีแบบ 8-Connected โดยที่มีจุดเริ่มต้นจุดเดียวกัน และสีที่ต้องการระบายกับสีของพื้นที่ที่ต้องการระบายเหมือนกัน ผลลัพธ์จะได้ดังรูปที่ 2.31

รูปที่ 2.31 แสดงการเทสีแบบ 8-Connected

2.1.6 ขั้นตอนการรู้จำ (Recognition) [2]

ขั้นตอนนี้ถือเป็นหัวใจของระบบ เพราะเป็นส่วนที่จะตัดสินว่ารูปตัวอักษรที่ส่งเข้าไปเป็นรหัสตัวอักษรอะไร เช่นเดียวกับส่วนอื่นๆ ที่มีวิธีการหลากหลายซึ่งนำมาใช้เพื่อให้ได้ผลการทำงานที่ดีที่สุด เทคนิคใหม่ๆ ได้ถูกพัฒนาขึ้นมาอย่างต่อเนื่องเพื่อรองรับกับปัญหาที่เกิดจากเอกสารที่มีความสลับซับซ้อนมากขึ้น

2.1.6.1 วิธีทางการเข้าคู่รูปแบบ (Template Matching)

วิธีการเข้าคู่รูปแบบเป็นวิธีการแรกๆ ที่มาใช้ในการรู้จำตัวอักษร หลักการโดยทั่วไปคือ จะต้องมีรูปแบบ (Template) ที่สร้างขึ้นมาสำหรับอ่านตัวอักษร โดยมีการกำหนดตำแหน่งสำคัญที่สามารถใช้แยกแยะความแตกต่างระหว่างตัวอักษรแต่ละตัว เวลาทำงานก็ให้นำรูปภาพที่ต้องการอ่านไปทาบบนแบบเพื่อวัดความคล้ายคลึงกันของภาพกับต้นแบบ จากนั้นก็ระบุว่าเป็นรหัสตัวอักษรอะไร โดยใช้ค่าผ่านระดับหรือวิธีการบางอย่างในการตัดสิน วิธีการนี้จะค่อนข้างอ่อนไหวต่อข้อมูลแทรกซ้อน ขนาด และการเอียงของตัวอักษร จึงจำเป็นต้องมีขั้นตอนการปรับแต่งข้อมูลที่ดี นอกจากนั้นขั้นตอนการเปรียบเทียบก็ไม่ใช่ว่าสามารถเทียบกันแบบจุดต่อจุดได้ เพราะในทางปฏิบัติตัวอักษรที่ส่งเข้าสามารถมีความแปรปรวนได้หลายรูปแบบ ดังนั้นวิธีการเทียบก็ต้องมีประสิทธิภาพเพียงพอที่จะรองรับกับปัญหาดังกล่าวได้

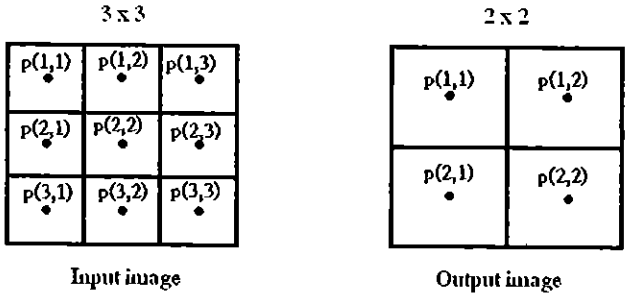
2.1.6.2 การปรับขนาดของภาพ (Image scaling)

การปรับขนาดของภาพเป็นกระบวนการในการปรับเปลี่ยนขนาดของภาพดิจิตอล ไม่ว่าจะเป็นการย่อภาพ หรือขยายภาพ ในการปรับเปลี่ยนขนาดของภาพดิจิตอลนั้น มีวิธีการอยู่หลายวิธี เช่น Nearest-neighbor interpolation, Bilinear interpolation หรือ Bicubic interpolation เป็นต้น ในแต่ละวิธีการก็จะมีข้อดีข้อเสียที่แตกต่างกันออกไป แต่ในโครงงานนี้จะกล่าวถึงการปรับเปลี่ยนขนาดของภาพด้วยวิธี Nearest-neighbor interpolation เท่านั้น เนื่องจากเป็นวิธีการที่ง่ายที่สุด Nearest neighbor interpolationเป็นวิธีการหนึ่งในการให้ค่าความเข้มแสงแก่พิกเซลใหม่กับภาพใหม่ในการ ย่อ หรือขยายภาพ โดยหลักการคือ จะเอาค่าความเข้มแสงของพิกเซลที่อยู่ใกล้ที่สุดมาใช้ โดยการคำนวณจากจุดกึ่งกลางของพิกเซล นั้นๆ

ในการย่อหรือขยายภาพจะประกอบไปด้วย 3 ขั้นตอน

1. สร้างภาพใหม่
2. คำนวณพิกัด
3. ให้ค่าสีแก่ภาพใหม่

จากรูปที่ 2.28 เป็นตัวอย่างที่จะแสดงให้ดูว่าถ้าสมมติมีภาพขนาด 3×3 พิกเซล ต้องการย่อให้เหลือขนาด 2×2 พิกเซล นั้นต้องทำอย่างไร

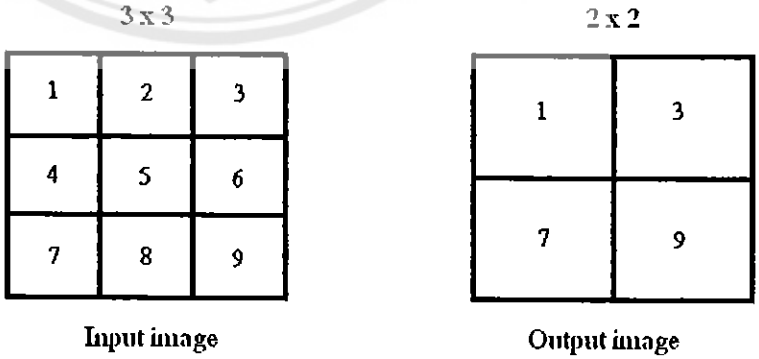


รูปที่ 2.32 แสดงภาพอินพุตขนาด 3×3 และภาพเอาต์พุตขนาด 2×2

ในขั้นตอนแรกนั้นจะต้องทำการสร้างภาพใหม่ตามขนาดที่ต้องการดังรูปที่ 2.33

รูปที่ 2.33 แสดงจุดกึ่งกลางของภาพเอาต์พุตบนภาพอินพุต

ในขั้นตอนที่ 2 นั้นจะเป็นการหาว่าจุดกึ่งกลางของแต่ละพิกเซลในภาพใหม่นั้นอยู่ที่ตำแหน่งใดในภาพเก่า จากรูปที่ 2.33 จะแสดงให้เห็นว่าจุดกึ่งกลางของพิกเซลในภาพเอาต์พุต ขนาด 2×2 แต่ละพิกเซลอยู่ที่ตำแหน่งใดในภาพอินพุตขนาด 3×3 เมื่อเอาภาพมาเทียบกัน สังเกตได้จากจุดสีแดงที่อยู่ในภาพอินพุตขนาด 3×3



รูปที่ 2.34 การให้ค่าแก่ภาพเอาต์พุตจากการเลือกค่าจากตำแหน่งที่ใกล้ที่สุด

ในขั้นตอนสุดท้าย ก็จะเป็นการให้ค่าแก่ภาพเอาต์พุตขนาด 2×2 พิกเซล จากการเลือกเอาค่าจากตำแหน่งพิกเซลที่อยู่ใกล้ที่สุดมาใส่

ห้องสมดคณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

2.2 ทฤษฎีเฮาซดอรฟดิสแทนซ์ (Hausdorff Distance) [3][4]

ในปี 1868-1942 นักคณิตศาสตร์ชาวเยอรมันชื่อ Felix Hausdorff ได้ค้นพบวิธีเปรียบเทียบภาพ 2 ภาพ ว่าคล้ายหรือแตกต่างกัน ด้วยวิธีการคำนวณระยะทางจากตำแหน่งใด ๆ บนภาพทุก ๆ จุด ตำแหน่งขึ้นใหม่โดยค่าของจุดภาพเหล่านั้นจะบอกถึงระยะห่างของจุดภาพนั่นเอง ซึ่งต่อมาเรียกว่า เทคนิคเฮาซดอรฟดิสแทนซ์ (Hausdorff Distance Technique) วิธีการเฮาซดอรฟดิสแทนซ์ดำเนินการโดยกำหนดให้ภาพที่ต้องการทดสอบการจำแนก (Test) เป็นภาพ 2 ระดับ ตัวอักษรให้มีค่าเป็น 0 พื้นหลังกำหนดให้มีค่าเป็น 1 กรณีจำแนกภาพตัวอักษร ตำแหน่งบนภาพที่นำมาทดสอบที่มีค่าเป็น 0 คือสมาชิกของเซต A และตำแหน่งบนภาพแม่แบบ (Model) ที่มีค่าเป็น 0 คือสมาชิกของเซต B สมการเฮาซดอรฟดิสแทนซ์นิยามดังสมการที่ (2-16) และสมการที่ (2-17)

$$H(A,B) = \max(h(A,B), h(B,A)) \tag{2-16}$$

เมื่อ

$$h(A,B) = \underset{a\in A}{max}\ \underset{b\in B}{min} ||a-b|| \tag{2-17}$$

$$h(B,A) = \underset{b\in B}{max}\ \underset{a\in A}{min} ||b-a|| \tag{2-18}$$

โดยที่

A	คือ	เซตของจุดภาพที่มีค่าเป็น 1 บนภาพที่ต้องการทดสอบการรู้จำ
B	คือ	เซตของจุดภาพที่มีค่าเป็น 1 บนภาพแม่แบบ
$\|\|a-b\|\|$	คือ	ระยะห่างระหว่างภาพ A ไปยังภาพ B โดยใช้การวัดระยะห่างแบบยุคลิด (Euclidean Distance)

จากสมการข้างต้นกำหนดให้จุดภาพที่มีค่าเป็น 0 ของภาพที่ทดสอบ คือสมาชิกของเซตข้อมูล A ซึ่งสามารถเขียนในรูปของเซตคือ $A=\{a_1,a_2,a_3,\ldots,a_p\}$, $a_i = (S_i, t_i)$, $1 \le i \le m$ และกำหนดให้จุดภาพที่มีค่าเป็น 0 ของภาพแม่แบบคือสมาชิกของเซตข้อมูล B เขียนในรูปของเซตคือ $B = \{b_1,b_2,b_3,\ldots,b_q\}$, $b_j = (x_j, y_j)$, $1 \le i \le n$ เมื่อ m และ n คือจำนวนจุดบนภาพที่นำมาทดสอบและจำนวนจุดบนภาพแม่แบบตามลำดับแล้วพิจารณาเปรียบเทียบระยะทางโดยสมการเฮาซดอรฟดิสแทนซ์ นิยามแทนด้วย *H(A,B)* จากสมการที่ 2-2 นิยาม h(A,B) คือค่าความห่างจากเซตของข้อมูล A ไปยังเซตของข้อมูล B โดยการนำสมาชิกของเซต A ซึ่งนิยามแทนด้วย $a \in A$ หนึ่งตำแหน่งไปวัดระยะทาง (Distance) กับสมาชิกของเซต B ทุก ๆ ตำแหน่ง แล้วเลือกระยะทางที่มีค่า

น้อยที่น้อยที่สุดไว้ จากนั้นเลื่อนตำแหน่งสมาชิกเซต a แล้วนำไปวัดระยะทางกับสมาชิกของเซต B อีกครั้งทำซ้ำจนกระทั่งสิ้นสุดตำแหน่งสมาชิกของเซต A เลือกค่าที่วัดระยะทางระหว่าง A กับ B ที่มีค่ามากที่สุดเพียงค่าเดียวออกมาซึ่งเป็นตำแหน่งที่ภาพทดสอบไกลจากภาพแม่แบบมากที่สุด สำหรับสมการที่ 2-1 เป็นการวัดระยะทางจากสมาชิกของเซต A ไปยังสมาชิกของเซต B และจากสมาชิกของเซต B ไปยังสมาชิกของเซต A แล้วเลือกระยะทางที่ไกลที่สุด ซึ่งก็คือเป็นการวัดจุดที่มีความผิดพลาด (Mismatch) ระหว่าง A กับ B มากที่สุดนั่นเอง

จะเห็นได้ว่าค่าของ H นี้จะมีค่าอยู่ในช่วง [0,∞] โดยค่า H ที่มีค่ามากขึ้นจะแสดงถึงความถูกต้องที่ลดลง ฉะนั้นค่าความถูกต้องจึงแปรผกผันกับค่า H ข้อดีของวิธีการเฮาซดอรฟดิส-แทนซ์ คือ ทนทานต่อการเปลี่ยนแปลงของข้อมูล เช่น แสง สัญญาณการรบกวน ภาพถูกบัง และภาพในมุมเอียง เป็นต้น

บทที่ 3

วิธีการดำเนินงานโครงงาน

เมื่อได้ทำการวิเคราะห์ออกแบบขั้นตอนในการดำเนินงานของระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความ จะมีขั้นตอนในการดำเนินงานทั้งหมดดังนี้

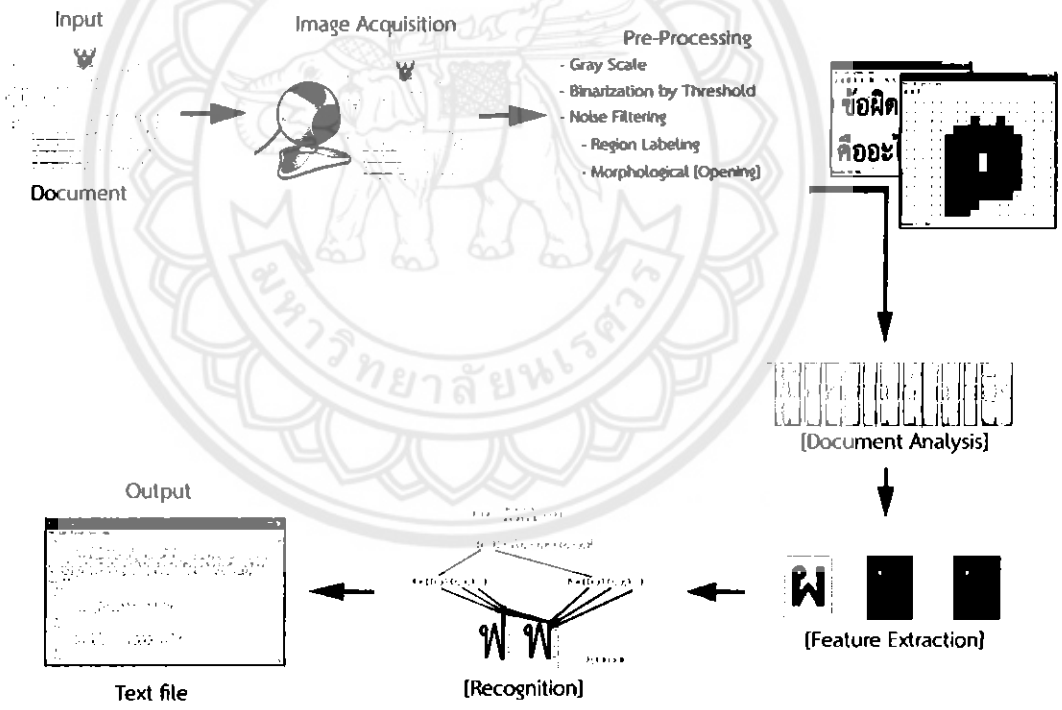
3.1 ขั้นตอนการรับภาพ

3.2 ขั้นตอนการประมวลผลขั้นต้น

3.3 ขั้นตอนการวิเคราะห์เอกสาร

3.4 ขั้นตอนการสกัดลักษณะเด่น

3.5 ขั้นตอนการรู้จำ



รูปที่ 3.1 โครงสร้างของระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความ

ในส่วนของระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความมีขั้นตอนการทำงานทั้งหมดดังรูปที่ 3.2 โดยมีขั้นตอนการทำงานดังนี้คือ ขั้นแรกจะเป็นขั้นตอนการรับภาพจากกล้องเว็บแคม โดยกล้องเว็บแคมจะทำการรับภาพเข้ามาเก็บเป็นไฟล์ภาพสี 24 บิต จากนั้นจะถูกแปลงเป็นภาพระดับเทา 256 ระดับ แล้วภาพจะถูกแปลงให้เป็นภาพขาวดำสองระดับด้วยวิธีการเทรสโฮล ซึ่งจะเหลือข้อมูลภาพเพียง 2 ระดับเท่านั้นคือ 0 และ 1 จากนั้นจะทำการแยกบรรทัด แยกตัวอักษร สระ และวรรณยุกต์ ออกจากกัน โดยอาศัยคุณสมบัติการโปรเจคชัน จากนั้นจะทำการแบ่งชนิดตัวอักษร สระ และวรรณยุกต์ โดยอาศัยการคัดแยกจากลักษณะเด่น เพื่อนำไปสู่ขั้นตอนการรู้จำและทำการแปลงเป็นไฟล์เอกสารต่อไป

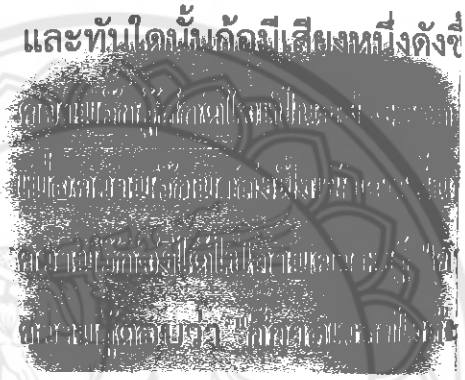
รูปที่ 3.2 ขั้นตอนการทำงานของระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความ

3.1 ขั้นตอนการรับภาพ

3.1.1 การเตรียมข้อมูลภาพ

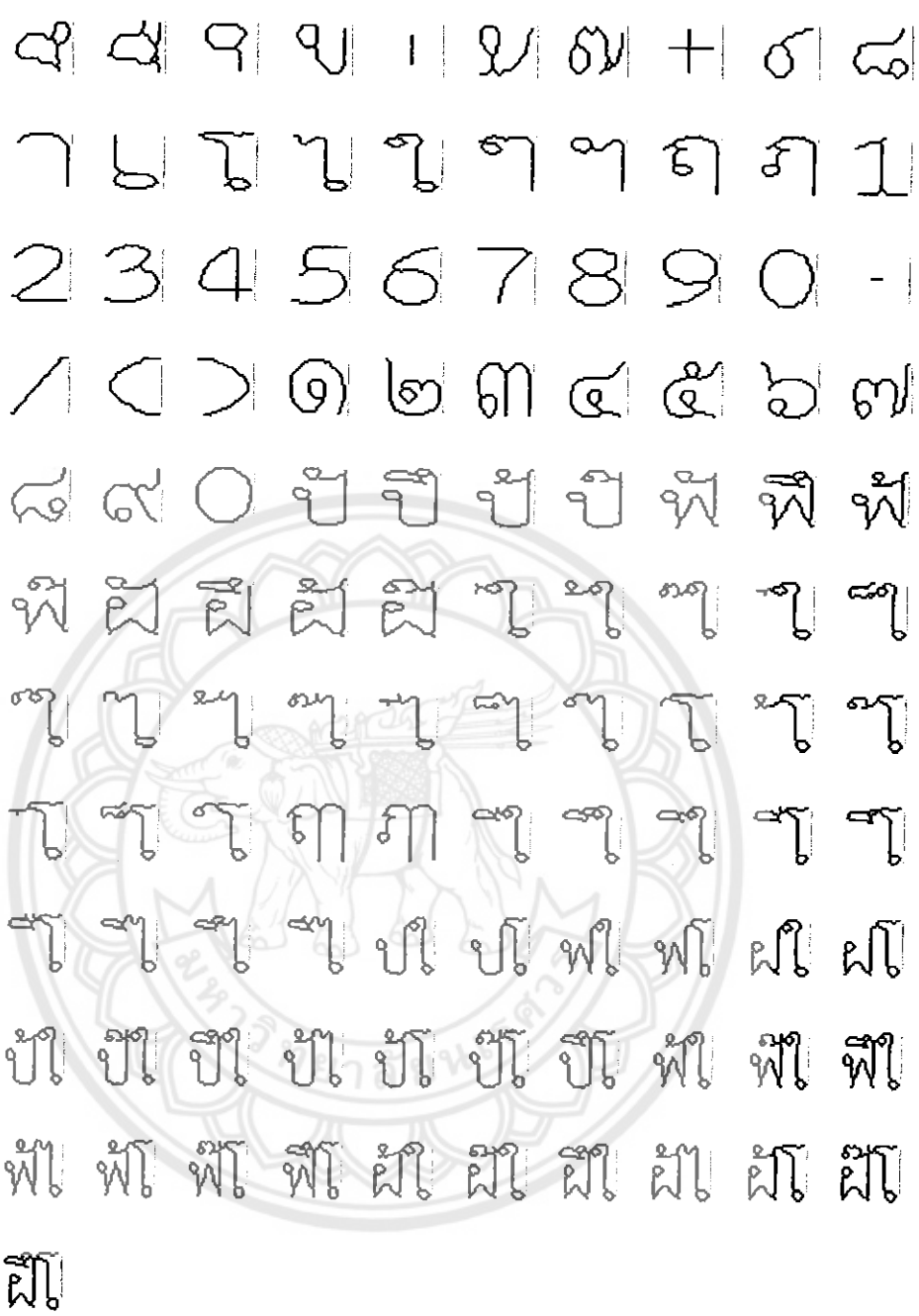
การหาข้อมูลเกี่ยวกับรูปภาพจะต้องมีการเตรียมข้อมูลของรูปภาพเพื่อนำมาใช้เป็นชุดที่ใช้ในการประมวลผลการรู้จำ ซึ่งที่มาของรูปเกิดขึ้นจากการถ่ายภาพเอกสารด้วยกล้องเว็บแคม โดยกระบวนการเตรียมข้อมูลมีดังนี้

1. นำเอกสารซึ่งประกอบด้วยตัวอักษรภาษาไทยและเลขอารบิก แบบ TH Sarabun PSK ขนาด 16 หน่วยขึ้นไป นำไปถ่ายภาพโดยใช้กล้องเว็บแคมที่ความละเอียด 640 x 480 พิกเซล



รูปที่ 3.3 ภาพเอกสารภาษาไทยแบบ TH Sarabun PSK

2. นำรูปภาพต้นแบบของตัวอักษรภาษาไทยและตัวเลขอารบิกแบบ TH Surabun PSK จำนวน 1 ภาพต่อ 1 ตัวอักษรในขนาด 30 x 30 พิกเซล รูปภาพที่ใช้เป็นต้นแบบจะประกอบด้วยชุดของรูปแบบตัวอักษรทั้งหมด 161 รูปแบบ ดังแสดงในรูปที่ 3.4



รูปที่ 3.4 ต้นแบบที่ใช้ในการรู้จำ

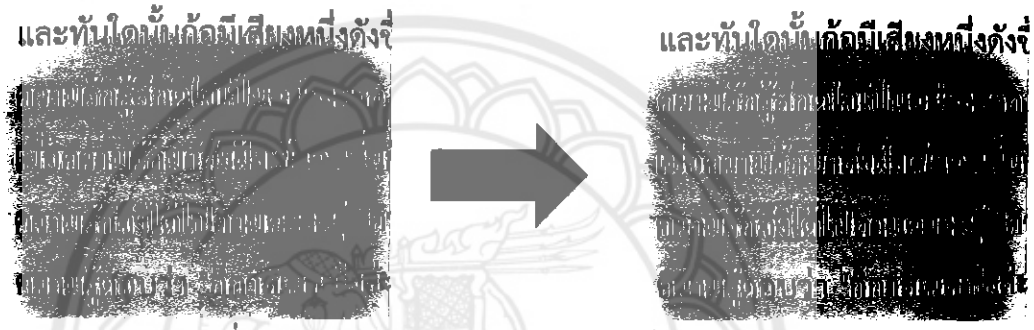
จากรูปที่ 3.4 จะสังเกตได้ว่าจะมีรูปต้นแบบสำหรับตัวอักษร กับสระ หรือวรรณยุกต์ที่ติดกัน เพื่อช่วยเพิ่มความถูกต้องในการรู้จำ

3.2 ขั้นตอนการประมวลผลขั้นต้น

เป็นขั้นตอนในการปรับปรุงและจัดเตรียมข้อมูลภาพในเหมาะสมกับขั้นตอนถัดไป ขั้นตอนเหล่านี้มีความสำคัญต่อประสิทธิภาพโดยรวมของระบบ ซึ่งประกอบด้วยส่วนสำคัญดังนี้

3.2.1 การแปลงภาพในระดับเทา

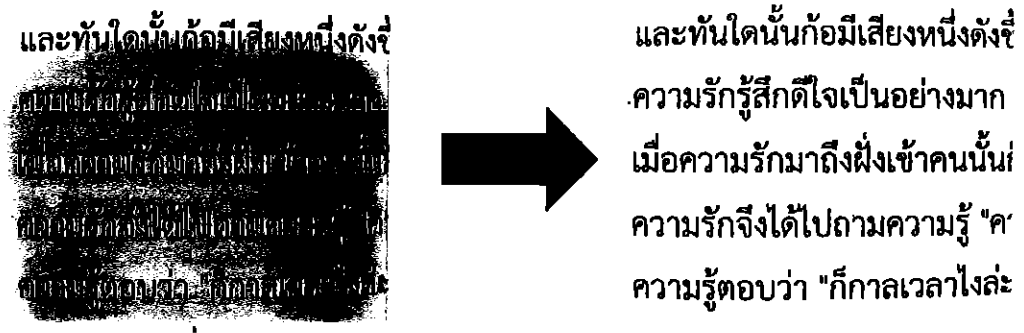
หลังจากได้ทำการเก็บข้อมูลภาพสี 24 บิตจากกล้องเว็บแคม และทำการกำจัดสัญญาณรบกวนออกไป ในขั้นตอนนี้ข้อมูลภาพในระดับสี จะนำมาทำการแปลงภาพให้เป็นภาพระดับเทา 256 ระดับ เพื่อความสะดวกในการนำไปใช้งานในขั้นตอนถัดไป



รูปที่ 3.5 การแปลงภาพเอกสารสี 24 บิตเป็นภาพระดับเทา

3.2.2 การแปลงภาพเป็นขาวดำสองระดับ

เมื่อได้รับข้อมูลภาพในระดับเทาแล้วจะแปลงภาพให้เป็นขาวดำสองระดับ ซึ่งจะเหลือข้อมูลภาพเพียง 2 ระดับคือ 0 และ 1 โดย 0 จะเป็นสีดำและ 1 จะเป็นสีขาว ทั้งนี้เพื่อแยกตัวหนังสือกับพื้นหลังออกจากกัน วิธีการแยกจะใช้การเทรสโฮลโดยให้เลือกวิธีการดังกล่าวซึ่งจะประกอบด้วย การเทรสโฮลด้วยวิธีของโอซึ การปรับค่าเทรสโฮลโดยใช้ค่ากลาง และการปรับค่าเทรสโฮลโดยการแบ่งภาพ



รูปที่ 3.6 การแปลงภาพเอกสารระดับเทาเป็นภาพขาวดำสองระดับ

3.2.3 การกรองสัญญาณรบกวน (Noise Filtering)

ในขั้นตอนนี้จะเป็นการกำจัดสัญญาณรบกวนที่ไม่ต้องการออกไปเพื่อที่จะลดความผิดพลาดในการทำงานในขั้นตอนการวิเคราะห์เอกสารและขั้นตอนต่อไปโดยจะมีวิธีการดังนี้

3.2.3.1 การคัดกรองวัตถุโดยพื้นที่ (Filtering by Area)

ในขั้นตอนนี้จะเป็นการกำจัดสัญญาณรบกวนที่ได้หลังจากการแปลงภาพขาวดำสองระดับเพื่อลดความผิดพลาดในการนำไปใช้ในขั้นตอนการวิเคราะห์เอกสาร โดยการคัดกรองนั้นจะทำการคัดกรองพื้นที่ของวัตถุที่มีพื้นที่น้อยกว่าที่กำหนด

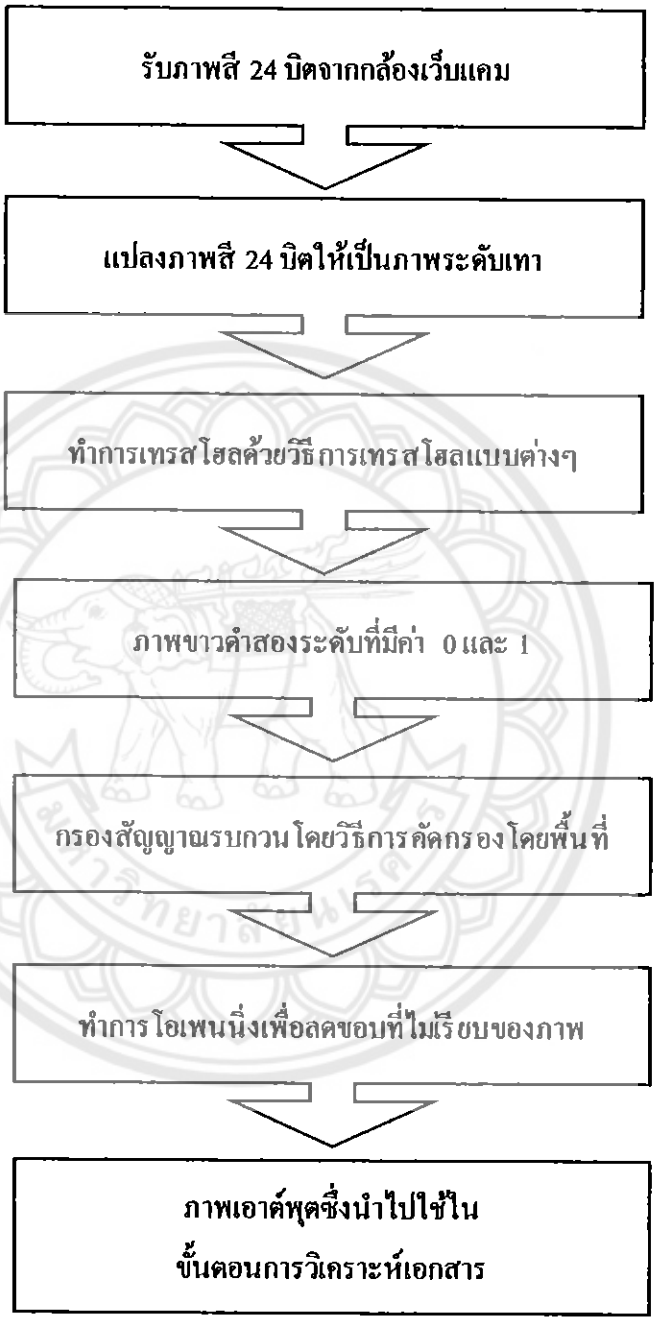
รูปที่ 3.7 การกำจัดสัญญาณรบกวนโดยวิธีการคัดกรองโดยพื้นที่

3.2.3.2 การทำโอเพนนิ่ง

เมื่อทำการคัดกรองโดยพื้นที่ของวัตถุเพื่อทำการกำจัดสัญญาณรบกวนแล้ว จะเข้าสู่ขั้นตอนการโอเพนนิ่งเพื่อทำให้ภาพตัวอักษรมีความเหมาะสมกับการทำให้บาง โดยการลดขอบที่ไม่เรียบของภาพตัวอักษร

รูปที่ 3.8 การทำโอเพนนิ่งเพื่อลดขอบที่ไม่เรียบของภาพ

ขั้นตอนในส่วนของขั้นตอนการรับภาพและขั้นตอนการประมวลผลขั้นต้น ได้ทำการแสดงแผนผังลำดับการทำงานดังรูปที่ 3.9 คือ



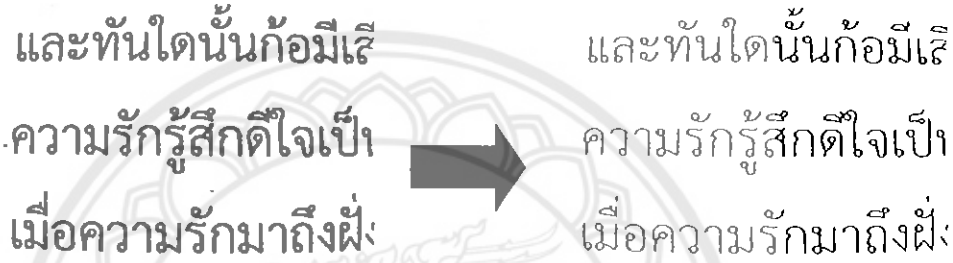
รูปที่ 3.9 ผังขั้นตอนการเตรียมภาพและขั้นตอนการประมวลผลขั้นต้น

3.3 ขั้นตอนการวิเคราะห์เอกสาร

หลังจากทำการแปลงภาพเป็นขาวดำสองระดับแล้ว ผลที่ได้พื้นหลังของภาพเอกสารจะเป็นสีขาวและสีดำจะเป็นส่วนของตัวหนังสือ ในขั้นตอนนี้จะทำการตัดแยกตัวอักษร สระ และวรรณยุกต์ โดยมีขั้นตอนการทำงานดังนี้

3.3.1 Thinning Morphological Image Processing

ขั้นตอนนี้จะนำตัวอักษรที่ได้จากการแปลงภาพเป็นขาวดำสองระดับให้บางลง เพื่อทำการแยกตัวอักษร สระ หรือวรรณยุกต์ที่ติดกันและเพื่อลดจำนวนจุดที่ต้องใช้ในการรู้จำ



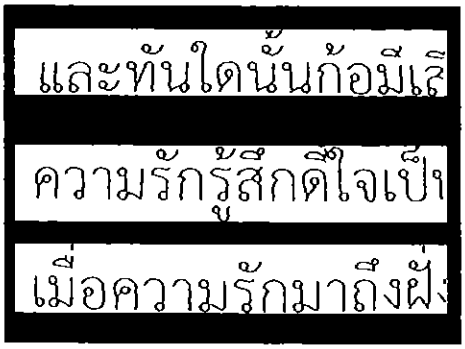
รูปที่ 3.10 การทำให้ตัวอักษรบางลง

3.3.2 การตัดแยกตัวอักษร สระ และวรรณยุกต์ (Projection)

ขั้นตอนการตัดแยกตัวอักษร สระ และวรรณยุกต์นั้นจะใช้วิธีการโปรเจคชัน ตามแนวตั้งและแนวนอน

3.3.2.1 การโปรเจคชันแนวนอน (เฟส 1)

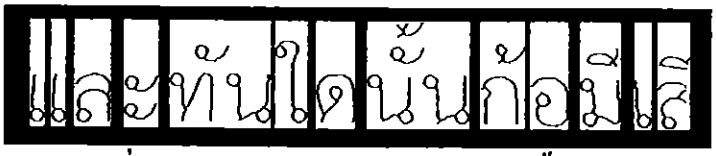
ในเฟสนี้จะเป็นการทำโปรเจคชันแนวนอนกับภาพ ผลที่ได้คือภาพตัวอักษรที่ได้ทำการตัดแยกเป็นแต่ละบรรทัด



รูปที่ 3.11 ภาพแสดงการโปรเจคชันแนวนอนเฟส 1

3.3.2.2 การโปรเจคชันแนวตั้ง (เฟส 1)

ในเฟสนี้จะเป็นการโปรเจคชันในแนวตั้งเพื่อตัดแยกคำหรือตัวอักษร ผลที่ได้จะได้ภาพตัวอักษรที่ถูกตัดแยกเป็นคำหรือตัวอักษรเดี่ยวๆ



รูปที่ 3.12 ภาพแสดงการโปรเจคชันแนวตั้งเฟส 1

3.3.2.3 การโปรเจคชันแนวนอน (เฟส 2)

ในเฟสนี้จะเป็นการทำการโปรเจคชันแนวนอนในครั้งที่ 2 กับภาพที่ผ่านกระบวนการโปรเจคชันในขั้นตอนที่ 3.3.2.1 และ 3.3.2.2 เพื่อเป็นการตัดสระกับวรรณยุกต์ที่อยู่บนตัวอักษร และสระที่อยู่ข้างล่างตัวอักษร

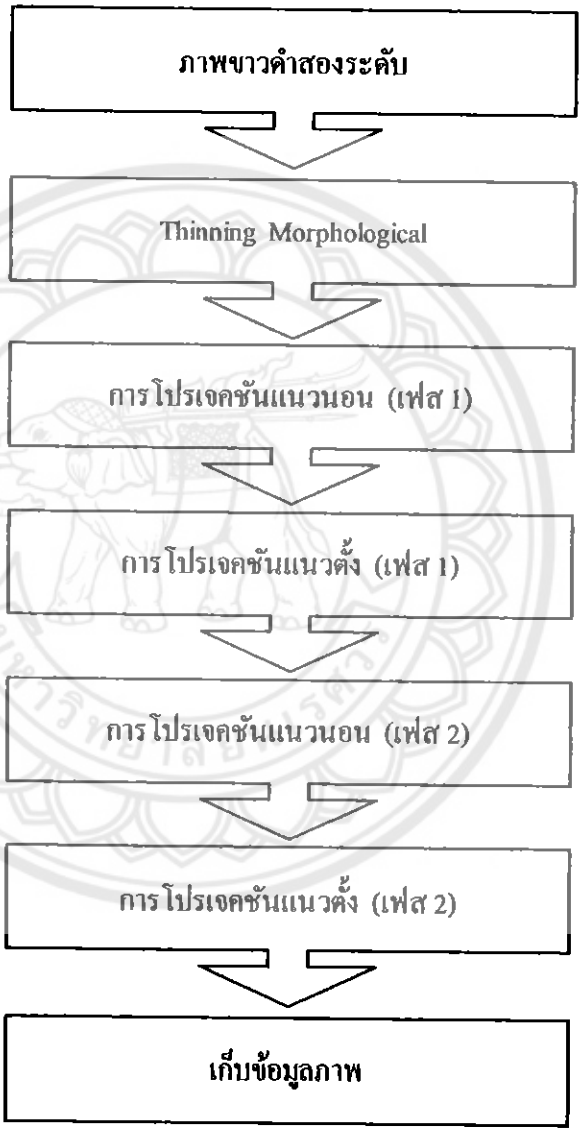
รูปที่ 3.13 ภาพแสดงการโปรเจคชันแนวนอนเฟส 2

3.3.2.4 การโปรเจคชันแนวตั้ง (เฟส 2)

เมื่อภาพที่ผ่านขั้นตอนที่ 3.3.2.3 ในการทำการตัดสระกับวรรณยุกต์บนตัวอักษร และสระที่อยู่ต่ำกว่าตัวอักษร แต่ในบางภาพตัวอักษรยังมีภาพที่มีตัวอักษรอยู่มากกว่าหนึ่งตัว จึงจำเป็นต้องทำการโปรเจคชันแนวตั้งอีกครั้ง

รูปที่ 3.14 ภาพแสดงการโปรเจคชันแนวตั้งเฟส 2

ทั้งนี้เมื่อทำการตัดแยกตัวอักษรแล้วก็ตาม แต่ก็จะยังมีตัวอักษรกับสระ หรือวรรณยุกต์ ที่ไม่สามารถตัดแยกออกเป็นตัวเดี่ยวๆได้ โดยเฉพาะตัวอักษรหางยาวที่มีสระหรือวรรณยุกต์ติดอยู่เช่น ปี ป่ ป้ ป๊ ฟ ฟ้ ฝ ฝ้ เป็นต้น ดังนั้นเพื่อให้ตัวอักษรกับสระ หรือวรรณยุกต์แยกออกจากกันทั้งหมดจึงต้องใช้วิธีการให้ป้ายชื่อแก่วัตถุ แล้วหยิบเอาแต่ละวัตถุไปสร้างเป็นรูปใหม่ หรือเพิ่มรูปภาพต้นแบบสำหรับตัวอักษรกับสระ หรือวรรณยุกต์ที่ติดกัน



รูปที่ 3.15 ผังแสดงขั้นตอนการวิเคราะห์เอกสาร

3.4 ขั้นตอนการสกัดลักษณะเด่น

การสกัดลักษณะสำคัญของตัวอักษร สระ และวรรณยุกต์นั้น สามารถทำได้หลายวิธีตามที่ได้ศึกษามาได้แก่ การหาจำนวนหัว

3.4.1 การหาจำนวนหัวของตัวอักษร

ตัวอักษรภาษาไทยแทบทุกตัวประกอบด้วยหัว โดยที่หัวของแต่ละตัวอักษรมีตำแหน่งที่แตกต่างกันบางตัวเหมือนกันบางตัวไม่เหมือนกัน การหาจำนวนหัวในที่นี้ไม่ได้หมายถึงหัวตัวอักษรเพียงอย่างเดียว แต่ยังรวมถึงช่องว่างในตัวอักษรด้วย เช่น ห พ จะถือว่ามี 2 หัว เป็นต้น

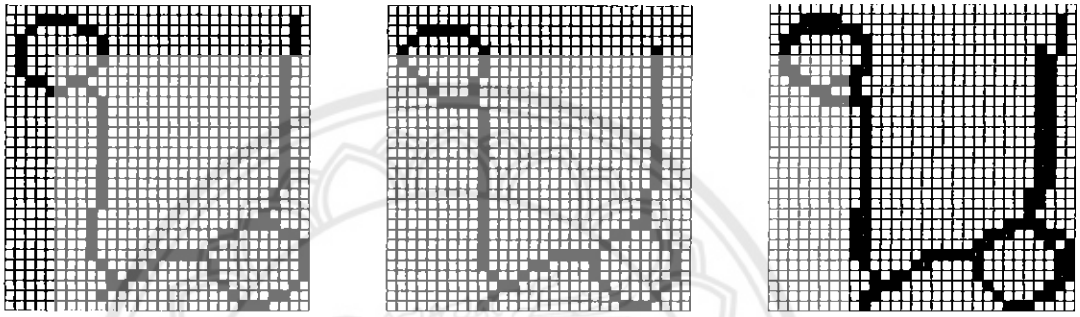
รูปที่ 3.16 การเทสี (Flood Fill) เพื่อหาตำแหน่งตัวอักษร

โดยขั้นตอนการทำงานจะเริ่มจากการทำการเทสี (Flood Fill) ที่เป็นสีเดียวกับตัวอักษรที่จะทำการหาหัวลงไปในภาพ จากนั้นจะทำการคัดกรองวัตถุโดยพื้นที่เพื่อทำการนับจำนวนของหัวที่ได้จากขั้นตอนการเทสีในการหาค่าหัวของตัวอักษรว่ามีจำนวนหัวเท่าไร ในขั้นตอนนี้จะทำให้ทราบจำนวนหัวของแต่ละตัวอักษรได้ ทั้งนี้เพื่อเพิ่มประสิทธิภาพของความรวดเร็วของข้อมูลที่ใช้ในขั้นตอนการรู้จำ

รูปที่ 3.17 การทำการคัดกรองโดยพื้นที่เพื่อนับจำนวนหัวตัวอักษร

3.5 ขั้นตอนการรู้จำ

หลังจากที่ได้ภาพที่ผ่านการตัดแยกบรรทัดและเป็นตัวอักษรเดียวมาแล้ว ขั้นตอนต่อมาคือการรู้จำ ในขั้นตอนการรู้จำนี้จะใช้ทฤษฎีเฮาซดอรฟดิสแทนซ์ในการรู้จำ โดยจะนำภาพแต่ละภาพที่ได้จากการตัดแยกไปทำการปรับขนาดของภาพเพื่อให้มีขนาดเท่ากันกับภาพแม่แบบ จากนั้นทำการเปรียบเทียบกับภาพแม่แบบที่มีอยู่แล้วเลือกภาพที่มีความใกล้เคียงที่สุดออกมาภาพหนึ่ง ด้วยวิธีการเฮาซดอรฟดิสแทนซ์เพื่อที่จะนำดัชนีของภาพแม่แบบนั้นไปแปลงเป็นตัวอักษรต่อไป



A B C

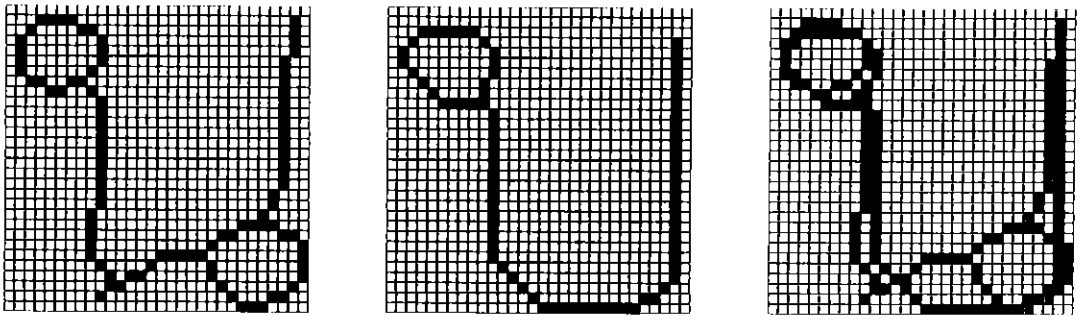
รูปที่ 3.18 ภาพแสดงการรู้จำตัวอักษรโดยทฤษฎีเฮาซดอรฟดิสแทนซ์
(A) ภาพที่จะนำมาทำการรู้จำ
(B) ภาพแม่แบบ
(C) แสดงการเปรียบเทียบ

$h(A,B) = \text{MaxMin } |A-B|$ $h(A,B) = 3.6056$

$h(B,A) = \text{MaxMin } |B-A|$ $h(B,A) = 1.4142$

$H = \text{Max}(h(A,B), h(B,A))$ $H = 3.6056$

ค่า H ที่ได้นั้นจะเรียกว่าค่าเฮาซดอรฟเป็นค่าที่แสดงถึงระยะทางที่สั้นที่สุดระหว่าง A-B กับ B-A ที่มีค่ามากที่สุด ดังนั้นจึงกล่าวได้ว่า ภาพที่มีความใกล้เคียงกันจะมีค่า H ที่ต่ำ เพราะฉะนั้นภาพที่เหมือนกันแล้ว ทางอุดมคติเมื่อนำมาทำการรู้จำจะได้ ค่า $H = 0$



A B C

รูปที่ 3.19 ภาพแสดงการรู้จำตัวอักษรโดยทฤษฎีเฮาซดอรฟดิสแทนซ์ (ต่อ)

(A) ภาพที่จะนำมาทำการรู้จำ

(B) ภาพแม่แบบ

(C) แสดงการเปรียบเทียบ

h(A,B) = MaxMin |A-B|	h(A,B) = 7
h(B,A) = MaxMin |B-A|	h(B,A) = 4.2426
H = Max(h(A,B), h(B,A))	H = 7

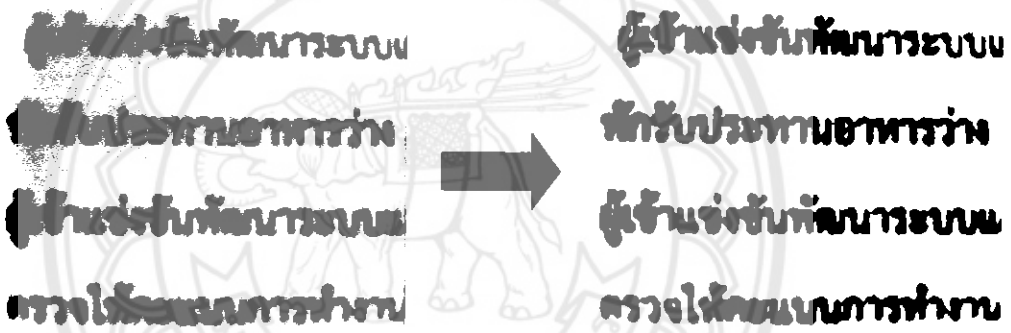
สังเกตได้ว่าเมื่อนำรูปภาพที่ไม่ถูกต้องมาเปรียบเทียบกันแล้ว ค่า H ที่ได้นั้นจะมีค่าที่มากกว่ารูปภาพที่ถูกต้อง ดังนั้นจึงสามารถที่จะเลือกภาพที่มีความคล้ายคลึงกันมากที่สุดได้ด้วยการหาภาพที่มีค่า H ที่ต่ำที่สุด

บทที่4

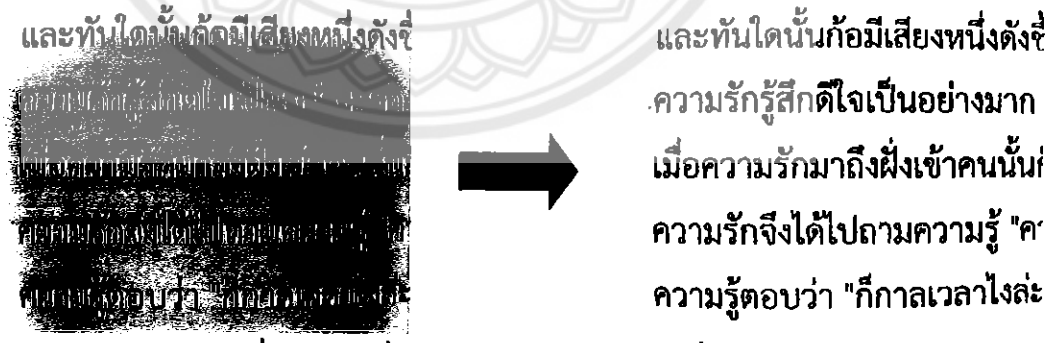
ผลการทดลองและการวิเคราะห์

4.1 ผลการทดลองส่วนของขั้นตอนการรับภาพ

การทดลองการรับภาพจากกล้องเว็บแคมนั้น ผลการทดลองพบว่าในการรับภาพจากกล้องเว็บแคมนั้น ภาพต้นแบบที่ถ่ายโดยกล้องเว็บแคมที่ไม่สามารถปรับโฟกัสได้ ไม่สามารถนำเอามาใช้ในการทดลองได้เนื่องจากภาพที่ได้ไม่มีความคมชัดและความละเอียดต่ำ เมื่อผ่านขั้นตอนการเทรสโฮลแล้วจะทำให้สูญเสียรายละเอียดของตัวอักษร ซึ่งเมื่อใช้กล้องเว็บแคมที่สามรถปรับโฟกัสได้นั้น จะได้ภาพที่มีความคมชัดและความละเอียดสูง และจะได้รายละเอียดในขั้นตอนการเทรสโฮล



รูปที่ 4.1 ภาพที่ผ่านการเทรสโฮลจากกล้องที่ไม่มีโฟกัส



รูปที่ 4.2 ภาพที่ผ่านการเทรสโฮลจากกล้องที่มีโฟกัส

4.2 ผลการทดลองส่วนของขั้นตอนการประมวลผลขั้นต้น

ในการทดลองขั้นแรกเป็นส่วนของการเตรียมภาพ เพื่อรองรับภาพท่าว่างหรือมืดเกินไป ดังรูปที่ 4.3 (ก) เป็นภาพที่มีสภาพแสงที่ไม่เรียบคือมีทั้งส่วนที่มืดและส่วนที่สว่างอยู่ในรูป รูปที่ 4.3 (ข) จะทำการเทรสโฮลโดยใช้วิธีการปรับค่าเทรสโฮลโดยใช้ค่ากลาง รูปที่ 4.3 (ค) เป็นการเทรสโฮลโดยใช้วิธีการปรับค่าเทรสโฮลแบบอัตโนมัติ และรูปที่ 4.3 (ง) เป็นการเทรสโฮลภาพโดยใช้วิธี การปรับค่าเทรสโฮลโดยการแบ่งภาพ โดยใช้รูปต้นแบบคือ รูป 4.3 (ก) เนื่องจากในสภาพแวดล้อมมีแสงที่แตกต่างกันจึงไม่สามารถกำหนดวิธีการเทรสโฮลที่แน่นอนได้ จึงเลือกวิธีการเทรสโฮลได้หลายแบบเพื่อให้เหมาะสมกับการทำงานในขั้นตอนอื่นๆ

รูปที่ 4.3 ภาพแสดงการเทรสโฮลด้วยวิธีต่างๆ

(ก.) ภาพที่มีทั้งส่วนที่มืดและส่วนที่สว่าง

(ข.) เทรสโฮลด้วยวิธีการปรับค่าเทรสโฮลโดยใช้ค่ากลาง

(ค.) เทรสโฮลด้วยวิธีการเทรสโฮลแบบอัตโนมัติ

(ง.) เทรสโฮลด้วยวิธการปรับค่าเทรสโฮลโดยการแบ่งภาพ

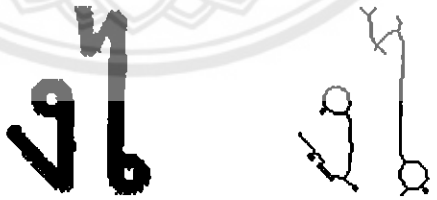
จากการที่ทำการแปลงภาพเป็นขาวดำสองระดับโดยใช้วิธีการเทรสโฮล และพบว่ายังมีส่วนของสัญญาณรบกวนที่ทำให้การทำให้บางของภาพมีปัญหา ดังนั้นจึงจำเป็นต้องทำการกำจัดสัญญาณรบกวนก่อน โดยใช้วิธีการคัดกรองโดยพื้นที่

รูปที่ 4.4 ภาพที่ไม่ได้ผ่านการทำการคัดกรองโดยพื้นที่มาทำให้บาง

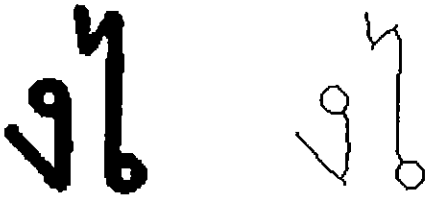
รูปที่ 4.5 ภาพที่ผ่านการการคัดกรองโดยพื้นที่มาทำให้บาง

จะสังเกตได้ว่าก่อนการทำการคัดกรองโดยพื้นที่นั้นจะมีสัญญาณรบกวนอยู่ในภาพ ส่งผลให้เมื่อทำให้บางจะยังคงมีจุดที่ไม่ต้องการอยู่ ซึ่งจะทำให้เกิดความผิดพลาดในขั้นตอนการวิเคราะห์เอกสาร แต่หลังจากการทำการคัดกรองโดยพื้นที่แล้ว สัญญาณรบกวนจะลดลงจะเหลือเพียงตัวอักษรที่ต้องการ เพื่อทำให้ภาพเหมาะสมกับการทำให้บางต่อไป

นอกจากนี้เมื่อทำการแปลงภาพเป็นขาวดำสองระดับแล้วยังพบอีกว่าตัวอักษรที่ได้มีขอบที่ไม่เรียบ ทำให้ภาพนั้นไม่เหมาะต่อการทำให้บางจึงจำเป็นต้องทำการกำจัดลดขอบที่ไม่เรียบของตัวอักษรออกไป โดยการทำการโอเพนนิ่ง



รูปที่ 4.6 ภาพตัวอักษรที่ยังไม่ได้ทำโอเพนนิ่งแล้วนำไปทำให้บาง



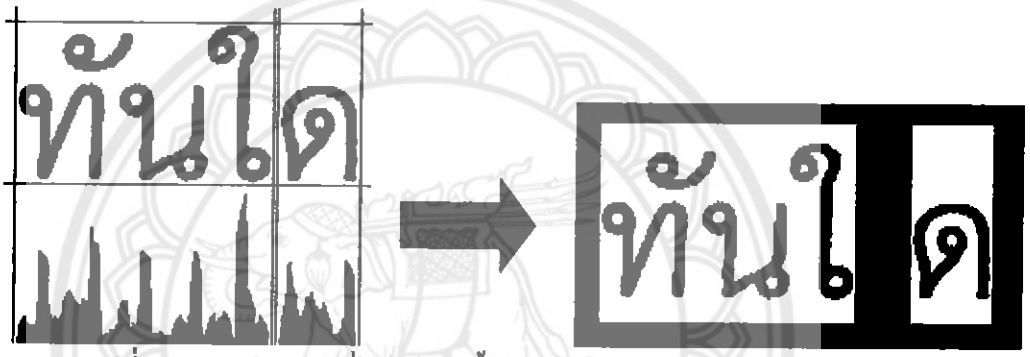
รูปที่ 4.7 ภาพตัวอักษรที่ผ่านการทำโอเพนนิ่งแล้วนำไปทำให้บาง

4.3 ผลการทดลองส่วนของขั้นตอนการวิเคราะห์เอกสาร

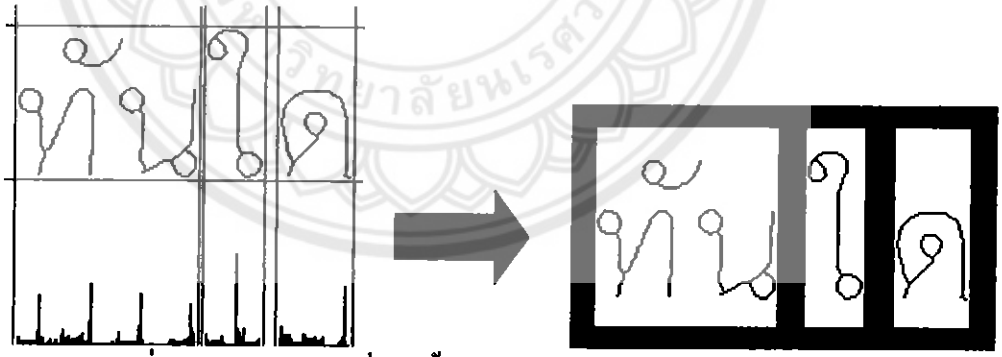
การทดลองการวิเคราะห์เอกสารจะใช้ภาพที่ผ่านขั้นตอนการประมวลผลขั้นต้นมาใช้ซึ่งผ่านการแปลงภาพเป็นขาวดำสองระดับและการทำการคัดกรองโดยพื้นที่ และการทำ โอเพนนิ่งกับการทำให้บางแล้ว

ในการทดลองในส่วนนี้พบว่าก่อนหน้าที่ไม่มีขั้นตอนการทำให้บางนั้นตัวอักษรที่ซ้อนกันการโปรเจคชันจะไม่สามารถใช้ได้เนื่องจากตัวอักษรที่ได้ออกมาจะติดกันอยู่เป็นภาพเดียวกัน

แต่เมื่อมีการทำให้บางกับภาพแล้ว การโปรเจคชันจะสามารถแยกตัวอักษรที่ซ้อนกันนั้นออกจากกันได้



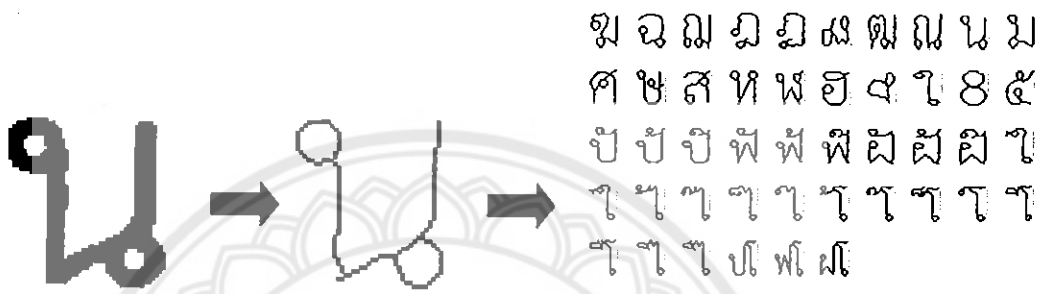
รูปที่ 4.8 ภาพตัวอักษรที่ไม่ได้ผ่านขั้นตอนการทำให้บางมาทำการโปรเจคชัน



รูปที่ 4.9 ภาพตัวอักษรที่ผ่านขั้นตอนการทำให้บางมาทำการโปรเจคชัน

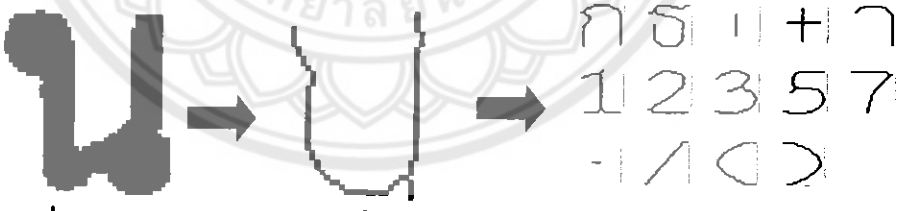
4.4 ผลการทดลองในส่วนของขั้นตอนการสกัดลักษณะเด่น

เมื่อทำการการสกัดลักษณะเด่นแล้ว ผลการทดลองที่ได้อาจจะมีความคลาดเคลื่อนเนื่องจากในขั้นตอนนี้ ทำเพื่อลดจำนวนภาพที่ใช้ทดสอบมาเปรียบเทียบกับภาพต้นแบบน้อยลง ตัวอย่างเช่น จากที่ต้องเปรียบเทียบภาพที่ทดสอบกับภาพต้นแบบ 20 รูปแบบ แต่เมื่อภาพทดสอบผ่านขั้นตอนการสกัดลักษณะเด่นแล้ว จะมีการเปรียบเทียบกับภาพต้นแบบเพียง 5 รูปแบบ เป็นต้น



รูปที่ 4.10 การสกัดลักษณะเด่นที่ผลตรงตามความเป็นจริง และแม่แบบที่เหลือ

ทั้งนี้ผลการทดลองของการสกัดลักษณะเด่นนี้ อาจจะเกิดความผิดพลาดขึ้นได้ โดยลักษณะเด่นที่สกัดได้ อาจไม่ตรงตามภาพต้นแบบที่เป็นจริงเพราะอาจจะเกิดจากการรับภาพที่ผิดพลาด และอาจจะส่งผลให้การรู้จำที่อาศัยการสกัดลักษณะเด่นมีความผิดพลาดไปด้วย แต่ถ้ามีการสกัดลักษณะเด่นที่ถูกต้องก็จะช่วยทำให้ขั้นตอนการรู้จำมีประสิทธิภาพมากยิ่งขึ้น



รูปที่ 4.11 การสกัดลักษณะเด่นที่ผลไม่ตรงตามความเป็นจริง และแม่แบบที่เหลือ

จากรูปที่ 4.11 เนื่องจากการรับภาพเข้ามาที่ผิดพลาด ส่งผลให้เมื่อทำการสกัดลักษณะเด่นแล้วนั้นผลที่ได้คือเป็นตัวอักษรที่ไม่มีหัว ซึ่งทำให้รูปแม่แบบที่เหลือนั้นไม่มีรูปที่ถูกต้อง

4.5 ผลการทดลองในส่วนของขั้นตอนการรู้จำ

ในส่วนของผลการทดลองในขั้นตอนการรู้จำนั้น จะทำการรับภาพจากกล้องเว็บแคมเพื่อทำการประมวลผลในส่วนของการประมวลผลขั้นต้น การวิเคราะห์เอกสาร การสกัดลักษณะเด่น จนมาถึงขั้นตอนนี้คือขั้นตอนการรู้จำ โดยจะนำตัวอักษรที่ได้จากขั้นตอนข้างต้นมาทำการเปรียบเทียบกับตัวอักษรต้นแบบที่ได้ทำการเก็บไว้ ในผลการทดลองในส่วนการรู้จำจะทำการแบ่งผลการทดลองออกเป็น 3 ส่วนดังนี้

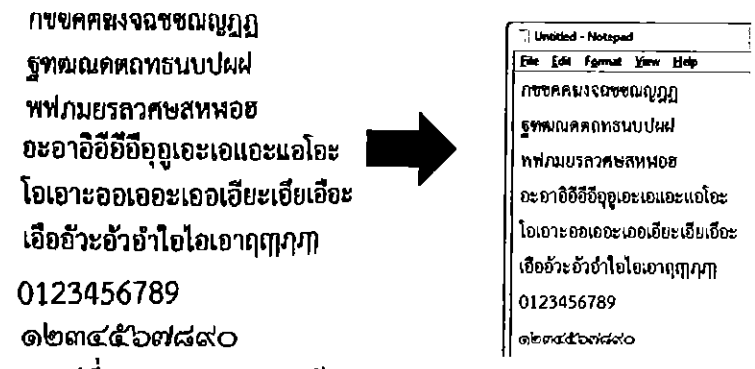
4.5.1 ผลการทดลองในส่วนการรู้จำปกติภายใต้สภาวะควบคุม

ในผลการทดลองส่วนนี้จะทำการถ่ายภาพเอกสารจากกล้องเว็บแคมโดยให้อยู่ภายใต้สภาวะควบคุม คือจะทำการควบคุมปริมาณความสว่างของสภาพแสง สัญญาณรบกวน และทำการวัดค่าความถูกต้องของผลการทดลองจากจำนวนตัวอักษรที่ถูกต้องหลังจากที่แปลงเป็นไฟล์ข้อความแล้ว

4.5.1.1 ผลการทดลองการรู้จำตัวอักษรภาษาไทย

ในผลการทดลองนี้จะรับภาพตัวอักษรภาษาไทยจำนวน 44 ตัวอักษร สระ 32 ตัว ตัวเลขอารบิกและตัวเลขไทย 10 ตัว ภายใต้สภาวะควบคุม โดยจะควบคุมสภาพแสงที่มีความสม่ำเสมอและไม่มีสัญญาณรบกวน และทำการวัดค่าความถูกต้องของผลการทดลองส่วนการรู้จำได้เป็นร้อยละ 98.83 ทั้งนี้เพื่อทดสอบแม่แบบของระบบว่าสามารถอ่านได้ทุกตัวอักษรหรือไม่

ตัวอักษรภาษาไทย	ตัวอักษรทั้งหมด	✓	✗	ร้อยละความถูกต้อง	ตัวอักษรที่ผิด
พยัญชนะไทย 44 ตัว	44	44	44	100	-
สระไทย 32 ตัว	32	31	1	96.87	ี
ตัวเลขอารบิก	10	10	10	100	-
ตัวเลขไทย	10	10	10	100	-



รูปที่ 4.12 แสดงผลการรู้จำพยัญชนะ สระ ตัวเลข ภาษาไทย

4.5.1.2 ผลการทดลองส่วนการรู้จำตัวอักษรภาพเอกสาร

ในผลการทดลองส่วนการรู้จำตัวอักษรภาพเอกสารจะเป็นภาพเอกสารที่ถ่ายจากกล้องเว็บแคมภายใต้สภาวะควบคุมซึ่งจะควบคุมให้มีความแตกต่างกันทางด้านสภาพแสง มีสัญญาณรบกวน สิ่งแปลกปลอมภายในภาพ จำนวน 10 ชุดการทดลองมาทำการรู้จำ จากนั้นจะทำการนับจำนวนตัวอักษรที่สมบูรณ์ จำนวนตัวอักษรที่นับได้ จำนวนตัวอักษรที่ผิดพลาด และทำการวิเคราะห์เป็นความถูกต้องร้อยละเท่าใด

เมื่อทำการแปลงภาพที่รับเข้ามาจากกล้องเว็บแคมแล้วหลังจากที่ผ่านขั้นตอนการรู้จำซึ่งเป็นขั้นตอนสุดท้ายเพื่อให้ได้ออกมาซึ่งไฟล์เอกสาร โดยทำการทดลองจากภาพเอกสารจำนวน 10 ชุดการทดลองได้ผลการทดลองดังตารางที่ 4.2

ตารางที่ 4.2 ร้อยละของความถูกต้องของการแปลงเป็นไฟล์เอกสาร

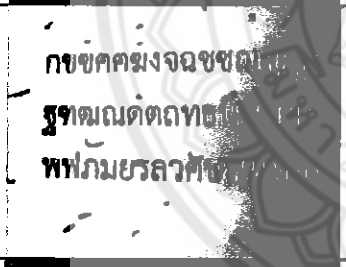
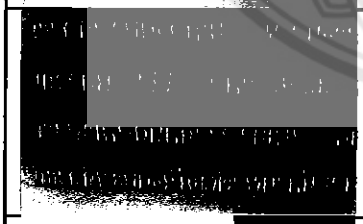
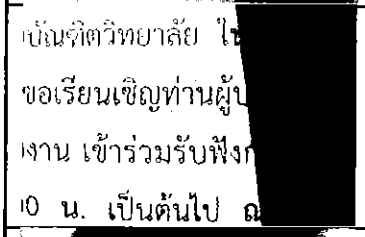
ชุดทดลอง	ตัวอักษรทั้งหมด	✓	✗	ร้อยละความถูกต้อง	ตัวอักษรที่รู้จำผิด
1	94	88	6	93.61	ต(2) ช ี ่ ้
2	93	86	7	92.47	ต ใ น ิ ี ่ ้
3	62	60	2	96.77	ี ่
4	84	82	2	97.61	ช ิ
5	93	87	6	93.54	ช(2) ี(2) ่(2)
6	74	74	0	100.00	
7	61	59	2	96.72	ช ี
8	77	76	1	98.70	ี
9	75	66	9	88.00	ี(2) ี(2) ่(2) ค ู ้
10	97	95	2	97.93	ณ ่

ผลการทดลองในส่วนนี้เมื่อทำการนำชุดการทดลองจำนวน 10 ชุดที่ได้ผ่านการประมวลผลภาพขั้นต้น และทำการรู้จำจากนั้นจะทำการแปลงเป็นไฟล์ข้อความ ระบบสามารถแปลงเป็นไฟล์ข้อความได้ค่าความถูกต้องเฉลี่ยโดยนับจำนวนตัวอักษรที่สมบูรณ์ จำนวนตัวอักษรที่นับได้ จำนวนตัวอักษรที่ผิดพลาด และทำการวิเคราะห์เป็นความถูกต้องร้อยละ 95.53 ตัวอักษรที่ผิดพลาดส่วนใหญ่ได้แก่สระจำพวกสระบน ี ี ี ิ ี ทั้งนี้ขึ้นอยู่กับมุมในการถ่ายภาพเอกสารจากกล้องเว็บแคม

4.5.1.3 ผลการทดลองส่วนการรู้จำตัวอักษรที่ได้รับสัญญาณรบกวน

ผลการทดลองนี้จะทำการรับภาพเอกสารภายใต้สภาวะควบคุม คือทำการเพิ่มสัญญาณรบกวนชนิดต่างๆ เข้ามาที่เป็นสัญญาณรบกวนทางด้านงานเอกสาร เช่น การใช้ปากกาเน้นข้อความ การขีดเขียน สัญญาณรบกวนทั่วไป ภาพที่ถูกถ่ายในสภาพแสงที่มีความเข้มแสงที่แตกต่างกัน เข้าไปในภาพเอกสารเพื่อวัดประสิทธิภาพการประมวลผลขั้นต้น ตัวอักษรที่ผิดพลาดส่วนใหญ่กระจายออกไป ทั้งนี้ขึ้นอยู่กับความสามารถในการถ่ายภาพเอกสารจากกล้องเว็บแคม วัดค่าความถูกต้องได้ร้อยละ 97.56

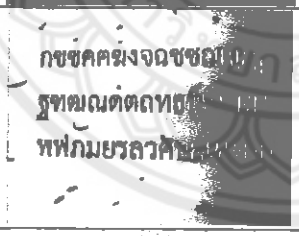
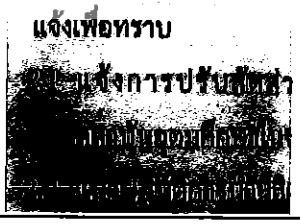
ตารางที่ 4.3 แสดงภาพเอกสารที่ถูกสัญญาณรบกวน

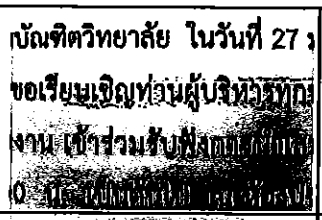
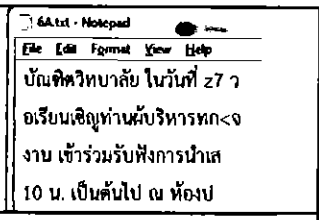
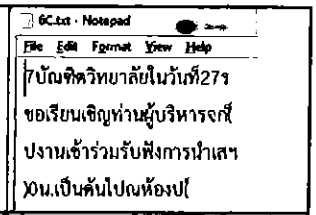
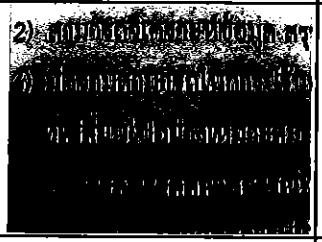
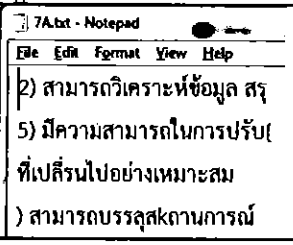
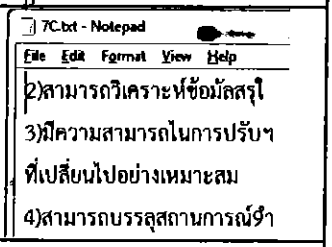
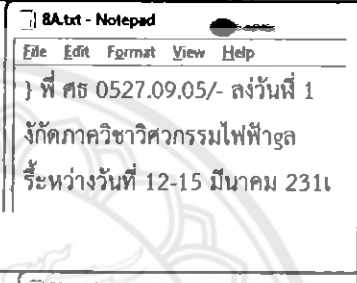
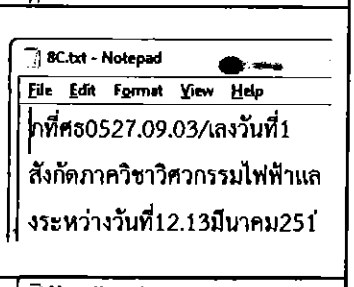
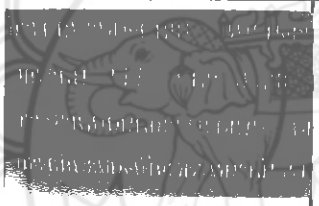
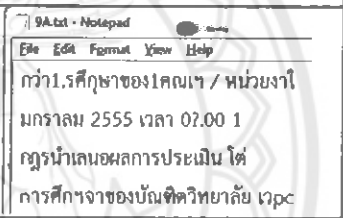
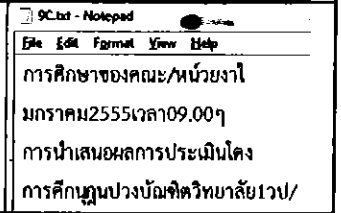
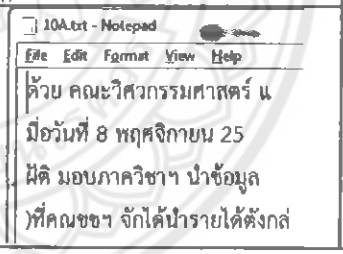
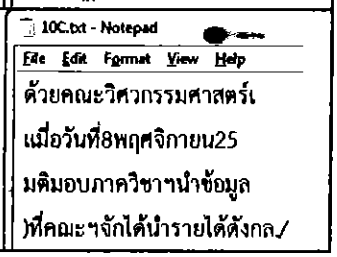
ภาพที่มีสัญญาณรบกวน	ตัวอักษรทั้งหมด	✓	×	ร้อยละความถูกต้อง	ตัวอักษรที่ผิด
	44	29	15	65.90	ก ค ค ซ ฎ ฐ ด ต บ ฝ พ ฟ ม ย ศ ห
	44	41	3	93.18	ฝ ย ร
	74	66	8	89.18	ี ่ ศ ษ ย อ ฌ เ
	94	87	7	92.55	น ี ่ ท ุ ต ฌ
	80	78	2	97.50	-(2)

4.5.2 ผลการทดลองการรู้จำเมื่อเทียบกับโปรแกรมอ่านไทย

ผลการทดลองในส่วนนี้จะทำการเปรียบเทียบความถูกต้องของโปรแกรมอ่านไทยกับ ระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความ โดยจะทำการนำภาพเอกสารชนิดเดียวกันจำนวน 10 ชุดการทดลองมาทำการรู้จำและวัดค่าความถูกต้องระหว่างโปรแกรมอ่านไทยและระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความโดยภาพเอกสารที่ใช้ในการทดลองนั้นจะอยู่ภายใต้สภาวะควบคุม ซึ่งจะประกอบด้วย การได้รับสัญญาณรบกวน สภาพแสงที่ไม่สม่ำเสมอ ภาพที่ถูกถ่ายในแนวเอียง ก้ม และ เงยกล้องเว็บแคม

ตารางที่ 4.4 แสดงการเปรียบเทียบไฟล์ข้อความจากโปรแกรมอ่านไทยกับระบบ

การทดลองที่	ภาพต้นแบบ	โปรแกรมอ่านไทย	ระบบการแปลงข้อความ
1		1A.txt - Notepad File Edit Format View Help สอนประจำภาคการศึกษา ห์ให้ภาควิชารวบรวม และ 1รศึกษาปลาย ปีการศึกฯา	1C.txt - Notepad File Edit Format View Help สอนประจำภาคการศึกษา ห์ให้ภาควิชารวบรวมและ)รศึกษาปลายปีการศึกษา
2		2A.txt - Notepad File Edit Format View Help ?)ชญคคมงจฉฑฌช9 ฐฑ)T1ณ1 ?+ดฉฑธนบปผk พพ 1 ภ@jยรลวศษสหพออ	2C.txt - Notepad File Edit Format View Help กขชศศมงจฉชฌญฎฎอ ฐฑฝ์ณคคฉฑธนนปผฝ์ ศฟ์ภฉงมรลวคษสฑพออ
3		3A.txt - Notepad File Edit Format View Help กขข์คคมงจฉชฌญฎฎ ฐฑฑฌณดตฉฑธนิ้บปผฝ พฟ ภพ มmลวคษสหพออ	3C.txt - Notepad File Edit Format View He กขชคคมงจฉชฌญฎฎ ฐฑฑฌณดตฉฑธนนปผม พฟ์ภมปีโลวศษสหพออ
4		4A.txt - Notepad File Edit Format View Help แจ้ง เพื่อ ทราบ 2.1 แจ้งการปรับสัดส่ว สถาบันอุดมศึกษาในส 2.Z แถวปฏิบัติการนำนิสิเ	4C.txt - Notepad File Edit Format View Help แจ้งเพื่อทราบ 2.1แจ้งการปรับสัดส่ว สถาบันอุดมศึกษาในf 2.2แนวปฏิบัติการนำนิสิโ
5	เป็นศิตวิทยาลัย ไ ขอเรียนเชิญท่านผู้บ เงาน เข้าร่วมรับฟังก 10 น. เป็นต้นไป ณ	5A.txt - Notepad File Edit Format View Help 1บัณฑิตวิทยาลัยฎีวันที่ 27 ว ขอเรีย1 พนจ.เจช1ญห่+ านผู้)ริหารทุกส1 1งาน เข้าร่วมรบฟศจ\ารนำเสๆ 0 น. เป็นต้นไป ณ หองปส(	5C.txt - Notepad File Edit Format View Help 7บัณฑิตวิทยาลัยไป.งวันที่27ร ขอเรียนเชิญท่านผู้บริหารฤกโ(ปงานเข้าร่วมรับฟังการนำเสๆ)0น.เป็นต้นไป71งห้องป(

6			
7			
8			
9			
10			

จากตารางที่ 4.4 แสดงการเปรียบเทียบภาพเอกสารที่ผ่านการแปลงเป็นไฟล์ข้อความระหว่างโปรแกรมอ่านไทยและระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความเพื่อให้เห็นว่าภายใต้สภาวะควบคุมที่เหมือนกัน เช่น การเพิ่มสัญญาณรบกวน สภาพแสงที่ไม่สม่ำเสมอ ว่ามีประสิทธิภาพในการแปลงภาพเอกสารเป็นไฟล์ข้อความนั้นมีความแตกต่างกันอย่างไร

จากการทดลองการเปรียบเทียบความถูกต้องของโปรแกรมอ่านไทยกับ ระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความโดยทำการเปรียบเทียบภาพเอกสารที่ถ่ายได้มาจากกล้องเว็บแคมจำนวน 10 ชุดโดยจะประกอบด้วยภาพเอกสารปกติ ภาพที่มีสัญญาณรบกวน ภาพที่มีสภาพแสงที่แตกต่างกัน จากนั้นจะทำการเปรียบเทียบจำนวนตัวอักษรที่สามารถอ่านได้แล้วนำผลการทดลองมาเปรียบเทียบซึ่งจะได้ผลการทดลองดังตาราง 4.5

ตารางที่ 4.5 แสดงร้อยละของความถูกต้องของการเปรียบเทียบกับโปรแกรมอ่านไทย

ชุดการทดลอง	ตัวอักษรทั้งหมด	โปรแกรมอ่านไทย			ระบบการแปลงฯ			ความถูกต้องเมื่อเทียบกับโปรแกรมอ่านไทย
		✓	✗	ร้อยละ	✓	✗	ร้อยละ	
1	61	60	1	98.36	60	1	98.36	= 0
2	44	12	32	27.27	29	15	65.90	> 11.90
3	44	34	7	77.27	41	3	93.18	> 20.91
4	75	73	2	97.33	74	1	98.66	> 1.33
5	94	80	14	85.10	87	7	92.55	> 7.45
6	91	84	7	92.30	86	5	94.50	> 2.2
7	97	93	4	95.87	95	2	97.93	>2.06
8	80	74	6	92.50	78	2	97.50	> 5.00
9	74	66	8	89.18	66	8	89.18	= 0
10	96	93	3	95.87	93	3	96.87	= 0

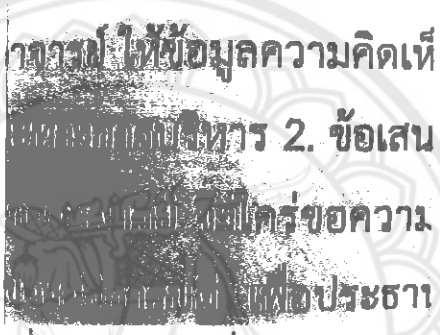
จากชุดทดลอง 10 ชุด จะทำการรับภาพที่ได้รับสัญญาณรบกวนชนิดต่างๆ เข้ามาที่เป็นสัญญาณรบกวนทางด้านงานเอกสาร เช่น การใช้ปากกาเน้นข้อความ รอยหมึก สัญญาณรบกวนทั่วไป ภาพที่ถูกถ่ายในสภาพแสงที่มีความเข้มแสงที่แตกต่างกัน พบว่า โปรแกรมอ่านไทยมีความถูกต้องในการรู้จำเฉลี่ยที่ร้อยละ 85.10 และในระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความมีความถูกต้องเฉลี่ยที่ร้อยละ 92.46 เนื่องจากในโปรแกรมอ่านไทยระบบการประมวลผลขั้นต้น (Pre-Processing) ยังไม่มีประสิทธิภาพเท่าที่ควรเมื่อรับภาพถ่ายเอกสารในสภาวะที่สภาพแสงไม่สม่ำเสมอและมีสัญญาณรบกวนเป็นจำนวนมาก

4.5.3 ผลการทดลองการรู้จำที่เกี่ยวกับความทนทานของระบบ

ผลการทดลองส่วนนี้เป็นการทดลองเพิ่มเติมเพื่อพิสูจน์ข้อจำกัดของระบบเกี่ยวกับการถ่ายภาพโดยกล้องเว็บแคมและภาพเอกสารตามได้ที่จำกัดไว้ในขอบเขต ในผลการทดลองส่วนนี้จะทำการทดลองเกี่ยวกับความทนทานของระบบว่าสามารถรองรับข้อจำกัดอื่นๆนอกเหนือจากขอบเขตได้หรือไม่โดยแบ่งเป็น 3 แบบการทดลองดังนี้

4.5.3.1 ความทนทานของระบบกับภาพเอกสารที่ถูกหมุนในมุมก้มและมุมเงย

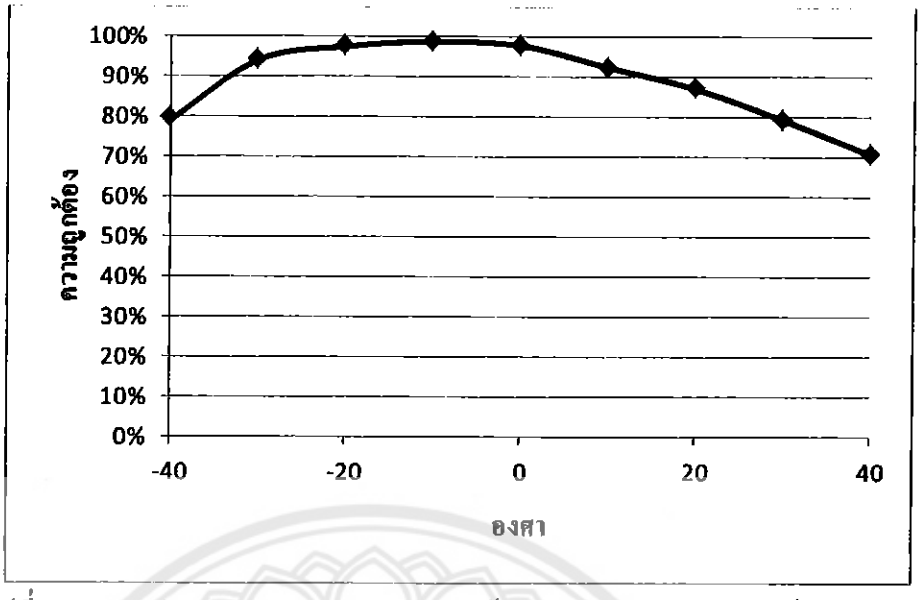
ในผลการทดลองส่วนนี้จะทำการนำภาพเอกสาร มาทำการทดลองเพื่อที่จะทำให้ทราบถึงความทนทานของระบบว่ารองรับได้กับภาพเอกสารที่ถ่ายโดยกล้องเว็บแคมจากมุมก้มและมุมเงยได้หรือไม่



รูปที่ 4.13 แสดงภาพที่ถ่ายในแนว 0 องศา

ตารางที่ 4.6 แสดงการผลค่าความถูกต้องเมื่อภาพเอกสารหมุนในมุมก้มและมุมเงย

การทดลองที่	องศาเปลี่ยน	ตัวอักษรทั้งหมด	✓	×	ร้อยละความถูกต้อง	ตัวอักษรที่ผิด
1	-40	107	85	22	79.43	า(3) ข(2) ค(2) ต(2) ย ไ ห ้ เ ภ ่ ค ธ ร ม ิ น
2	-30	87	82	5	94.25	า(2) ข เ ื
3	-20	83	81	2	97.59	า(2)
4	-10	88	87	1	98.86	์
5	0	91	89	2	97.80	์ ์
6	+10	92	85	7	92.39	า ์ ไ ส ค ข ื
7	+20	85	74	11	87.05	า(5) ข(2) ์ ไ ้ ส
8	+30	87	69	18	79.31	า(5) ร(2) อ(2) ท ห ข ึ ง ิ บ ะ
9	+40	92	65	27	70.65	า(8) ค(2) ห(2) อ(2) ้(2) ิ(2) ร ข ค ์ น ท ส ง บ

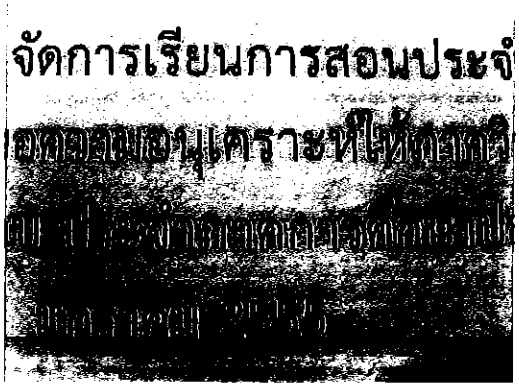


รูปที่ 4.14 กราฟแสดงผลการทดลองความถูกต้องของภาพเอกสารในมุมก้มและมุมเงย

จากตารางที่ 4.6 และรูปที่ 4.14 พบว่าเมื่อในภาพถ่ายเอกสารถูกหมุนในมุมก้มและมุมเงย ในกรณีนี้คือการถ่ายภาพเอกสารจากกล้องเว็บแคมโดยมีการถ่ายในมุมก้มหรือเงยกล้องขึ้นจะทำให้ประสิทธิภาพความทนทานของระบบมีความถูกต้องลดลงเรื่อยๆ จนไม่สามารถที่จะรู้จำได้ เนื่องจากเมื่อมีการก้มหรือเงยกล้องเว็บแคมจะทำให้ไม่สามารถถ่ายภาพตัวอักษรได้โฟกัสทั้งรูปภาพ

4.5.3.2 ความทนทานของระบบกับภาพเอกสารแนวเอียง

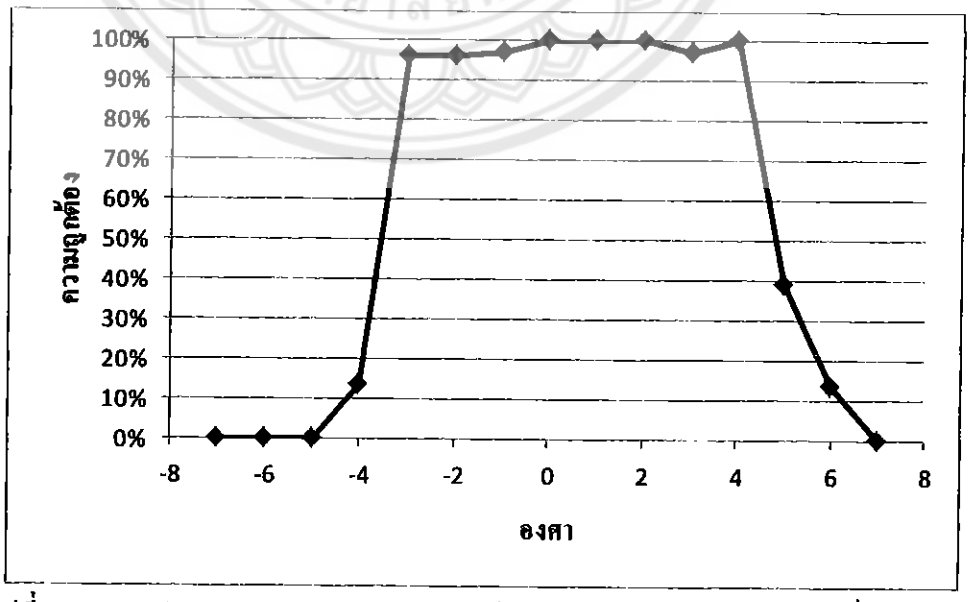
ในผลการทดลองส่วนนี้จะเป็นการแปลงภาพเอกสารที่ระบบสามารถแปลงเป็นไฟล์ข้อความได้ถูกต้องร้อยละ 100 ดังรูปที่ 4.15 มาทำการทดลอง โดยจะทำการหมุนภาพไปเรื่อยๆ ตั้งแต่ -1 องศาถึง -10 องศาและ 1 องศาถึง 10 องศาจากนั้นทำการวัดความถูกต้องในขณะที่ภาพเอกสารเปลี่ยนแปลงไปเรื่อยๆ ว่ามีประสิทธิภาพในการแปลงภาพเอกสารเป็นเท่าใด



รูปที่ 4.15 แสดงภาพที่ถ่ายในแนวปกติและความถูกต้องในการแปลงร้อยละ 100

ตารางที่ 4.7 แสดงผลการทดลองความถูกต้องของภาพเอกสารในแนวเอียงที่องศาต่างๆ

การทดลองที่	องศาที่เปลี่ยน	ตัวอักษรทั้งหมด	✓	✗	ร้อยละความถูกต้อง
1	-7	74	0	74	0
2	-6	74	0	74	0
3	-5	74	0	74	0
4	-4	74	10	64	13.51
5	-3	74	71	3	95.94
6	-2	74	71	3	95.94
7	-1	74	72	2	97.29
8	0	74	74	0	100
9	+1	74	74	0	100
10	+2	74	74	0	100
11	+3	74	72	2	97.29
12	+4	74	74	0	100
13	+5	74	29	45	39.18
14	+6	74	10	64	13.51
15	+7	74	0	74	0



รูปที่ 4.16 กราฟแสดงผลการทดลองความถูกต้องของภาพเอกสารในแนวเอียงที่องศาต่างๆ

จากตารางที่ 4.7 จะทำการหมุนภาพเอกสาร(Rotate) โดยใช้ภาพเอกสารที่ระบบสามารถแปลงเป็นไฟล์เอกสารโดยมีความถูกต้องร้อยละ 100 พบว่าเมื่อมีการหมุนภาพเอกสารตามเข็มนาฬิกา หรือ ทวนเข็มนาฬิกาแล้ว ประสิทธิภาพในรู้จำยังคงถูกต้องมากกว่าร้อยละ 90 ที่ -3 ถึง 4 องศา และประสิทธิภาพในการรู้จำลดลงเมื่อน้อยกว่า -3 องศา และมากกว่า 4 องศา สาเหตุเกิดจากการโปรเจคชันในขั้นตอนการตัดแยกบรรทัด เมื่อมีการหมุนภาพจะทำให้บรรทัดของตัวอักษรซ้อนกัน ทำให้การแปลงเป็นไฟล์ภาพผิดพลาดเพิ่มขึ้นเรื่อยๆตามองศาของภาพเอกสารที่เปลี่ยนไป

4.5.3.3 ความทนทานของระบบกับตัวอักษรภาษาไทยรูปแบบชนิดอื่น

ในผลการทดลองในส่วนนี้จะใช้ภาพต้นแบบของระบบที่เป็นรูปแบบตัวอักษรชนิด TH Sarabun PSK มาทำการเปรียบเทียบกับรูปแบบตัวอักษรภาษาไทยชนิด Angnasa New, Cordia New และ Tahoma เนื่องจากเป็นรูปแบบตัวอักษรพื้นฐานที่ถูกฝังมากับไมโครซอฟท์ออฟฟิศ และเพื่อทดสอบความทนทานของระบบว่าสามารถที่จะทำการรู้จำตัวอักษรที่ไม่มีในภาพต้นแบบได้หรือไม่

ตารางที่ 4.8 แสดงการเปรียบเทียบความถูกต้องเมื่อใช้ตัวอักษรชนิด Angsana New

ชุดที่	อักษรทั้งหมด	✓	×	ร้อยละความถูกต้อง
1	64	52	12	81.25
2	83	62	21	74.70
3	76	59	17	77.63
4	97	79	18	81.44
5	104	80	24	76.92

ตารางที่ 4.9 แสดงการเปรียบเทียบความถูกต้องเมื่อใช้ตัวอักษรชนิด Cordia New

ชุดที่	อักษรทั้งหมด	✓	×	ร้อยละความถูกต้อง
1	57	46	11	80.70
2	87	52	35	59.77
3	88	69	19	78.41
4	91	72	19	79.12
5	85	58	27	68.24

ตารางที่ 4.10 แสดงการเปรียบเทียบความถูกต้องเมื่อใช้ตัวอักษรชนิด Tahoma

ชุดที่	อักษรทั้งหมด	✓	x	ร้อยละความถูกต้อง
1	70	32	38	45.71
2	46	31	15	67.39
3	63	43	20	68.25
4	50	31	19	62.00
5	69	47	22	68.12

จากตารางที่ 4.8 - 4.10 เมื่อใช้รูปแบบตัวอักษรชนิดอื่นที่ไม่ได้มีในรูปต้นแบบ พบว่าประสิทธิภาพในการรู้จำสามารถที่จะรู้จำตัวอักษรได้ แต่เนื่องจากชนิดของรูปแบบตัวอักษรนั้นมีลักษณะเด่นของตัวอักษรที่แตกต่างกันทำให้ประสิทธิภาพในการรู้จำลดน้อยลง โดยเมื่อใช้รูปแบบ Angsana New มีความถูกต้องในการรู้จำร้อยละ 78.39 รูปแบบ Cordia New มีความถูกต้องในการรู้จำร้อยละ 73.25 และรูปแบบ Tahoma มีความถูกต้องในการรู้จำร้อยละ 62.30 เมื่อทำการเปรียบเทียบกับรูปแบบตัวอักษรชนิด TH Sarabun PSK ที่เป็นรูปต้นแบบของการรู้จำของระบบ เพราะรูปแบบตัวอักษรแต่ละชนิดมีความแตกต่างและลักษณะเฉพาะของพยัญชนะ และสระในรูปแบบตัวอักษรแต่ละตัวไม่เหมือนกัน

บทที่ 5

สรุปผลการดำเนินงานโครงงาน

ระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความถูกพัฒนาขึ้นเพื่อตอบสนองการใช้งานด้านเอกสารในกรณีที่มีการรับเอกสารจากที่อื่นหรือไฟล์เอกสารต้นฉบับสูญหาย ซึ่งต้องทำการพิมพ์ใหม่ เพื่ออำนวยความสะดวกในการใช้งานด้านเอกสารได้

หลักการพื้นฐานที่เกี่ยวข้องจะเป็นการรับภาพเอกสารโดยใช้กล้องเว็บแคมในการถ่ายภาพ การประมวลผลขั้นต้น การวิเคราะห์เอกสาร การค้นหาลักษณะเด่นของตัวอักษร และทฤษฎีของเฮาซ์ดอรฟดิสแทนซ์เพื่อใช้งานในการรู้จำตัวอักษร

โดยขั้นตอนการดำเนินงานจะเริ่มจากการรับภาพโดยใช้กล้องเว็บแคม จากนั้นทำการประมวลผลขั้นต้น ซึ่งประกอบด้วย การแปลงภาพในระดับเทา การแปลงภาพเป็นขาวดำสองระดับ และการกรองสัญญาณรบกวน ในขั้นตอนต่อมาจะเป็นการวิเคราะห์เอกสาร โดยจะทำการตัดแยกบรรทัด และตัวอักษรออกจากกัน ก่อนเข้าสู่ขั้นตอนการสกัดลักษณะเด่น ซึ่งใช้วิธีการหาจำนวนหัวตัวอักษร จากนั้นในขั้นตอนสุดท้ายจะเป็นการรู้จำโดยใช้ทฤษฎีของเฮาซ์ดอรฟดิสแทนซ์

ในผลการทดลองการถ่ายภาพเอกสารจากกล้องเว็บแคมจะต้องใช้กล้องเว็บแคมที่มีโฟกัสเพื่อให้ตัวอักษรในภาพที่มีความคมชัด ซึ่งนำไปสู่ขั้นตอนการประมวลผลขั้นต้นได้แก่ การทำเทรสโฮลภาพเพื่อแยกตัวอักษรออกจากพื้นหลัง โดยในการทำเทรสโฮลนั้นไม่สามารถที่จะกำหนดวิธีการเทรสโฮลที่แน่นอนได้ เนื่องจากในสภาพแวดล้อมมีแสงที่แตกต่างกัน การกรองสัญญาณรบกวน การลบขอบภาพที่ไม่เรียบเพื่อให้ง่ายต่อการทำให้ตัวอักษรบาง และจะนำไปทำการวิเคราะห์เอกสารเพื่อทำการตัดแยกพยัญชนะ สระ วรรณยุกต์ โดยการทำตัวอักษรให้บางนั้นจะมีความสำคัญต่อขั้นตอนการวิเคราะห์เอกสาร ซึ่งจะทำให้สามารถที่จะตัดแยกพยัญชนะ สระ วรรณยุกต์ที่ซ้อนกันอยู่ได้ เมื่อได้ตัวอักษรที่ตัดแยกแล้ว จะนำไปค้นหาลักษณะเด่นของตัวอักษรโดยใช้การนับจำนวนหัวตัวอักษร ซึ่งจะช่วยลดจำนวนภาพต้นแบบที่จะนำมาใช้ในการเปรียบเทียบได้ สุดท้ายจะนำไปทำการทดลองการรู้จำโดยจะทำการรู้จำกับตัวอักษรภาษาไทย สระ วรรณยุกต์ ตัวเลขภาษาไทยและตัวเลขอารบิก ภาพเอกสารที่ได้รับสัญญาณรบกวน ภาพเอกสารเมื่อเทียบกับโปรแกรมอ่านไทย และการทดลองเกี่ยวกับความทนทานของระบบซึ่งได้ทำการหาข้อสรุปดังต่อไปนี้

5.1 สรุปผลการทดลองของโครงงาน

5.2 สรุปปัญหาและอุปสรรคของโครงงาน

5.3 ข้อเสนอแนะและแนวทางการพัฒนา

5.1 สรุปผลการทดลองของโครงงาน

การแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความนั้น จากผลการทดลองที่ได้มาทั้งหมด จะเป็นการวัดประสิทธิภาพ โดยจะทำการนับจำนวนตัวอักษรที่ผิดกับจำนวนตัวอักษรที่สมบูรณ์ทั้งหมดในชุดการทดลอง แล้วทำการคำนวณหาค่าความถูกต้อง โดยทำการทดลองจำนวน 10 ชุด พบว่ามีความถูกต้องในการรู้จำถึงร้อยละ 90 และทำการเปรียบเทียบความถูกต้องกับโปรแกรมอ่านไทยพบว่ามีประสิทธิภาพในการรู้จำมากกว่าโปรแกรมอ่านไทยร้อยละ 5 เนื่องจากมีกระบวนการประมวลผลขั้นต้นที่ดีกว่า ทั้งนี้ระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความจะใช้ระยะเวลาในการประมวลผลในการแปลงเป็นไฟล์ข้อความที่นานเนื่องจากระบบ ถูกพัฒนาขึ้นโดยโปรแกรมแมทแลปซึ่งเป็นโปรแกรมที่ใช้คำนวณทางด้านวิศวกรรมศาสตร์และวิทยาศาสตร์ นอกจากนั้นยังได้แสดงผลการทดลองเพื่อให้เห็นข้อจำกัดของการถ่ายภาพเอกสารว่าต้องถ่ายในมุมปกติ ไม่เอียง ก้ม หรือ เงย กล้องเว็บแคม เพราะจะทำให้ประสิทธิภาพในการรู้จำลดลงตามการถ่ายภาพเอกสารที่ผิดพลาด

ผลทดลองของระบบภายใต้สภาวะควบคุมในกรณีต่างๆ

ตารางที่ 5.1 แสดงผลการทดลองของระบบภายใต้สภาวะควบคุมในกรณีต่างๆ

ผลการทดลองของระบบการแปลงข้อความจากภาพเอกสาร		ร้อยละความถูกต้อง
การรู้จำตัวอักษรภาษาไทย สระ วรรณยุกต์ตัวเลขอารบิกและตัวเลขไทย		99.21
การรู้จำภาพเอกสารภายใต้สภาวะควบคุม		95.53
การรู้จำภาพเอกสารที่ได้รับสัญญาณรบกวนภายใต้สภาวะควบคุม		97.56
การรู้จำภาพเอกสารที่ถูกหมุนในมุมก้มและมุมเงย		--> 0
การรู้จำภาพเอกสารในแนวเอียง		--> 0
การรู้จำตัวอักษรไทยภาษาไทยของรูปแบบชนิดอื่น	Angsana New	78.38
	Cordia New	73.24
	Tahoma	62.29

ผลการทดลองของระบบเมื่อเทียบกับโปรแกรมอ่านไทย

ตารางที่ 5.2 แสดงผลการเทียบกับโปรแกรมอ่านไทยภายใต้สภาวะควบคุมในกรณีต่างๆ

ผลการทดลองของระบบการแปลงข้อความจากภาพเอกสาร	ร้อยละความถูกต้อง
การรู้จำภาพเอกสารเมื่อเทียบกับโปรแกรมอ่านไทยภายในสภาวะควบคุม	> 5.09

5.2 ปัญหาและอุปสรรคของโครงงาน

จากการทดลองเพื่อหาปัจจัยที่ส่งผลกระทบต่อการทำงานของระบบการแปลงภาพเอกสารเป็นไฟล์ข้อความ ทำให้ทราบถึงสาเหตุต่างๆที่ทำให้เกิดข้อผิดพลาด โดยส่งผลให้ไม่สามารถทำการแปลงภาพเอกสารเป็นไฟล์ข้อความได้อย่างสมบูรณ์ และทำให้ระบบเกิดข้อจำกัดในการใช้งานดังต่อไปนี้

ตารางที่ 5.3 แสดงปัญหาและอุปสรรคและแนวทางการแก้ไข

ปัญหาและอุปสรรคของโครงงาน	แนวทางการแก้ไข
1.การถ่ายภาพเอกสารด้วยกล้องเว็บแคมจึงไม่สามารถถ่ายภาพทั้งเอกสารได้	1.อาจจะต้องมีการพัฒนาให้มีการถ่ายภาพเอกสารหลายๆภาพแล้วนำภาพมาเชื่อมต่อกัน
2.เนื่องจากเลนส์กล้องมีความโค้งเว้า อาจส่งผลต่อการถ่ายภาพเอกสารได้	2.มีการสร้างขอบเขตเพื่อตัดขอบภาพที่เกิดจากเลนส์โค้งหรืออาจถ่ายภาพเอกสารจากมุมสูง
3.แม่แบบของตัวอักษรเพียงแม่แบบเดียวคือ TH Sarabun PSK เมื่อมีการรู้จำกับรูปแบบตัวอักษรชนิดอื่นจะมีความถูกต้องในการรู้จำน้อยลง	3.เพิ่มแม่แบบเพื่อสามารถทำการรู้จำกับรูปแบบตัวอักษรชนิดอื่นให้มีความถูกต้องให้การรู้จำและมีความหลากหลายในการใช้งาน

5.3 ข้อเสนอแนะและแนวทางการพัฒนา

5.3.1 ความรู้ที่จำเป็นต่อการพัฒนาต่อ

1. ศึกษาเกี่ยวการกับทำงานของกล้องเว็บแคมเพื่อแก้ปัญหาการถ่ายภาพทั้งเอกสารและความโค้งเว้าของเลนส์
2. การศึกษาอัลกอริทึมที่ช่วยในการประมวลผลและเพิ่มประสิทธิภาพของความถูกต้องในการรู้จำ
3. การศึกษารูปแบบของตัวอักษรเพื่อใช้พัฒนาให้ผู้ใช้งานสามารถเพิ่มแม่แบบด้วยตนเองได้

5.3.2 แนวทางการพัฒนาต่อ

1. เชื่อมต่อกล้องเว็บแคมเพื่อทำการรับภาพให้ใช้งานได้แบบเวลาจริง
2. เพิ่มรูปแบบของการรู้จำ เช่น ชนิดรูปแบบตัวอักษรอื่นๆ รูปแบบตัวหนาและตัวเอียง
3. ปรับปรุงในส่วนของการประมวลผลให้รวดเร็วขึ้น
4. ระบบสามารถใช้งานกับภาพที่มีลักษณะเป็นมุมเอียงได้
5. พัฒนาการตรวจสอบและแก้ไขความถูกต้องของการแปลงข้อความโดยอัตโนมัติ

เอกสารอ้างอิง

[1] "Morphological Image Processing." [online]. Available : http://fivedots.coe.psu.ac.th/~montri/Teaching/image/morph.DOC. 2554.

[2] "Thai Optical Character Recognition." [online]. Available : http://thaiocr.phaisarn.com/. 2554.

[3] สนั่น ศรีสุข. (2542). "การรู้จำอักษรตัวพิมพ์ภาษาไทยโดยใช้เฮาซดอรฟดิสแทนซ์." ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีมหานคร, กรุงเทพมหานคร

[4] ราตรี จันทนะทรัพย์ และ นิดาพรรณ สุรีรัตนันท์. (2548) "การรู้จำป้ายทะเบียนรถยนต์ไทยโดยใช้เทคนิคเฮาซดอรฟดิสแทนซ์." ภาควิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์อุตสาหกรรม สถาบันเทคโนโลยีพระจอมเจ้าพระนครเหนือ, กรุงเทพมหานคร

[5] ศุภกิจ จันทรา และ อรรฌพัฒน์ วงศ์ผัน. (2550) "โปรแกรมวิเคราะห์ตัวอักษรไทยโดยใช้โครงข่ายประสาทเทียม." วิทยานิพนธ์ปริญญาวิศวกรรมศาสตร์ ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร, พิษณุโลก

[6] ภาณุเดช วศินวรรธนะ และ วัชรพงษ์ เจนพนัส.(2550) "โปรแกรมตรวจจับหมายเลขทะเบียนรถยนต์." วิทยานิพนธ์ปริญญาวิศวกรรมศาสตร์ ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร, พิษณุโลก

ภาคผนวก ก.

คู่มือการใช้งานระบบ

1. ระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความ คืออะไร

ระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความเป็นโครงงานนิสิตสาขาวิศวกรรมคอมพิวเตอร์ เป็นโปรแกรมที่ถูกพัฒนาขึ้นโดยการเขียนโปรแกรมด้วยภาษา MATLAB ซึ่งใช้งานโดยทดสอบผ่านรูปแบบตัวอักษรภาษาไทยชนิด TH Sarabun PSK เพื่อตอบสนองชีวิตในปัจจุบันที่ต้องแข่งขันกับเวลาในเรื่องการทำงานด้านเอกสาร ระบบนี้มีประสิทธิภาพในการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความโดยการถ่ายภาพเอกสารจากกล้องเว็บแคม ใช้งานง่าย ลักษณะการใช้งานเป็นแบบ Graphic User Interface ซึ่งสะดวกต่อการใช้งาน

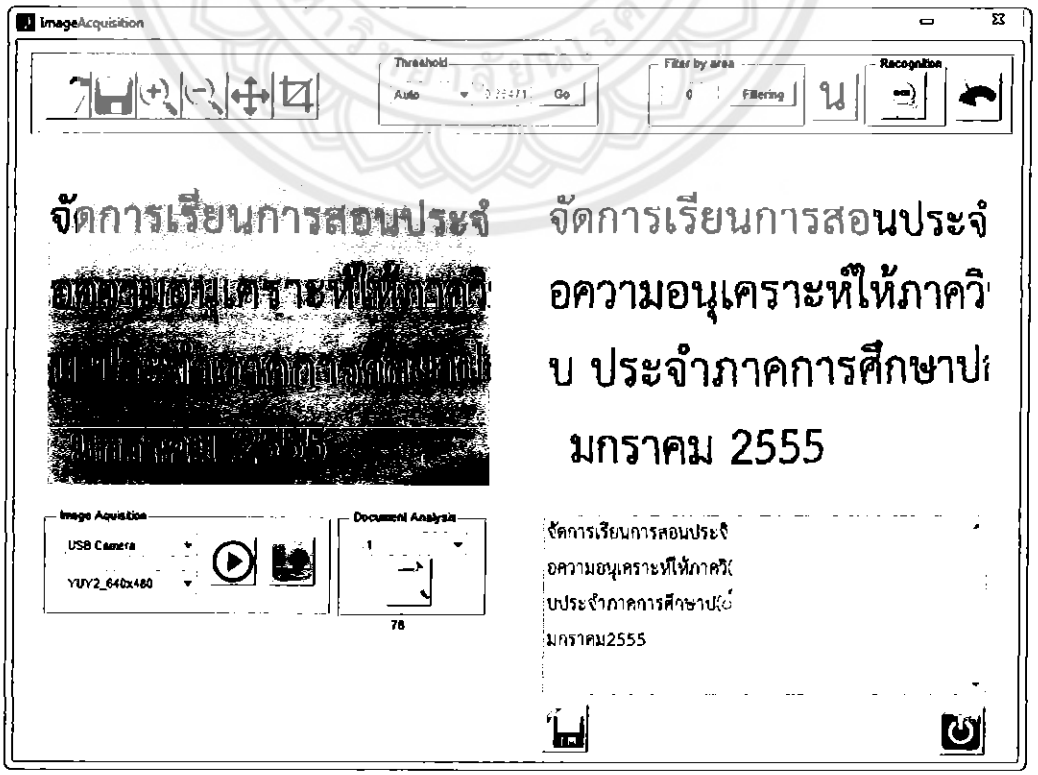
คุณสมบัติของระบบการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความ

1.สามารถแปลงภาพเอกสารเป็นไฟล์ข้อความ

2.มีประสิทธิภาพสูงสำหรับการปรับปรุงภาพที่มีสัญญาณรบกวน

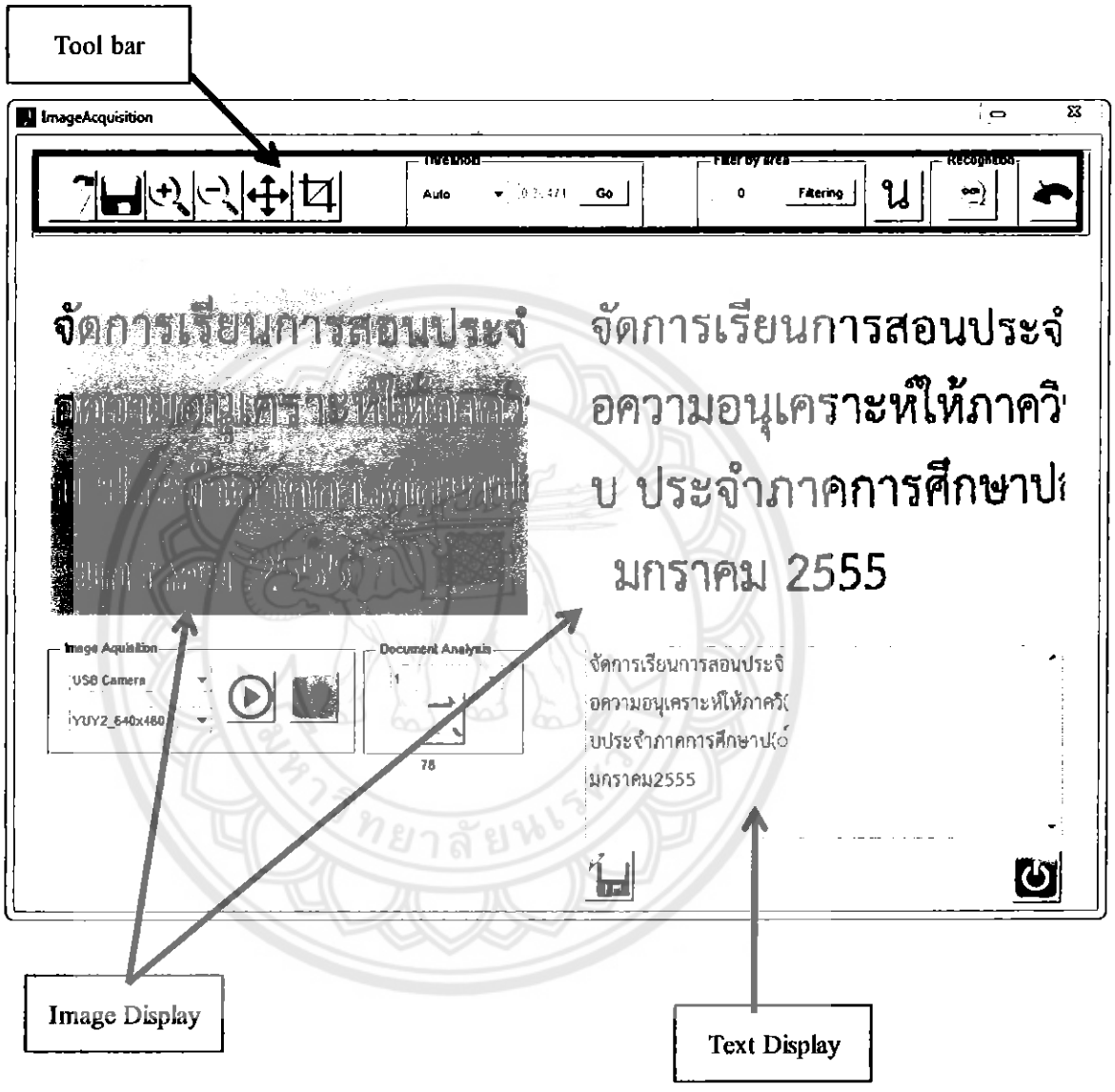
3.มีความสะดวกมากกว่าการใช้งานผ่านเครื่องสแกน

4.มีความถูกต้องของการแปลงเป็นไฟล์ข้อความสูงถึงร้อยละ 90



2. Graphic User Interface

GUI ประกอบด้วย Tool bar, Image Display, Text Display ซึ่งทุกรายการได้ออกมามาเพื่อความเหมาะสมในการใช้งานและเพื่ออำนวยความสะดวกแก่ผู้ใช้



อธิบายการทำงานของตัวโปรแกรม

: (Open file) ใช้เพื่อทำการ browse หรือเปิดรูปภาพเอกสารในแฟ้มที่มีการถ่ายเก็บไว้

: (Save Image) ใช้เพื่อทำการบันทึกรูปภาพเอกสารที่ทำการถ่ายภาพเอาไว้

: (Zoom in) ใช้เพื่อทำการขยายขนาดรูปภาพเพื่อดูรูปภาพหรือทำการตัดรูปภาพ

: (Zoom out) ใช้เพื่อทำการลดขนาดรูปเพื่อดูรูปภาพหรือทำการตัดรูปภาพ

: (Move) ใช้เพื่อทำการย้ายรูปเพื่อดูภาพรวมของภาพที่ถ่าย

: (Crop) ใช้เพื่อการทำตัดภาพเพื่อตัดภาพเอกสารเฉพาะส่วนที่ต้องการแปลงไฟล์ข้อความ

: (Start) ใช้เพื่อเริ่มการทำงานของกล้องเว็บแคม

: (Stop) ใช้เพื่อหยุดการทำงานของกล้องเว็บแคม

: (Capture) ใช้เพื่อทำการถ่ายภาพเอกสารที่ต้องการถ่ายปุ่มนี้จะเปิดขึ้นเมื่อกดปุ่ม Start

: (Enable/Disable find Hole) ใช้เพื่อทำการลดจำนวนการค้นหาตัวอักษรที่เหมือนกัน

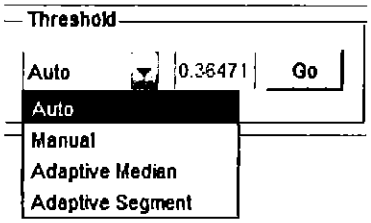
: (Separate Character) ใช้เพื่อคัดแยกบรรทัดของพยัญชนะ สระ วรรณยุกต์

: (Recognize) ใช้เพื่อทำการรู้จำและแปลงภาพเอกสารเป็นไฟล์ข้อความ

: (Save file) ใช้เพื่อทำการบันทึกตัวอักษรที่แปลงแล้วเป็นไฟล์ข้อความ (Text file)

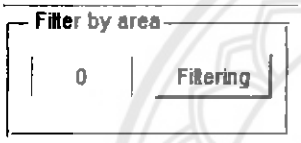
: (Reset) ใช้เพื่อทำการยกเลิกการกระทำทั้งหมดที่ได้ทำมาแล้วก่อนหน้า

: (Exit) ใช้เพื่อออกจากโปรแกรมเมื่อการทำงานสิ้นสุดแล้ว



: (Threshold) ใช้เพื่อปรับภาพเป็นขาวดำสองระดับเพื่อทำการแปลงเป็นไฟล์ข้อความ โดยมีให้เลือก 4 รูปแบบการ Threshold ดังนี้

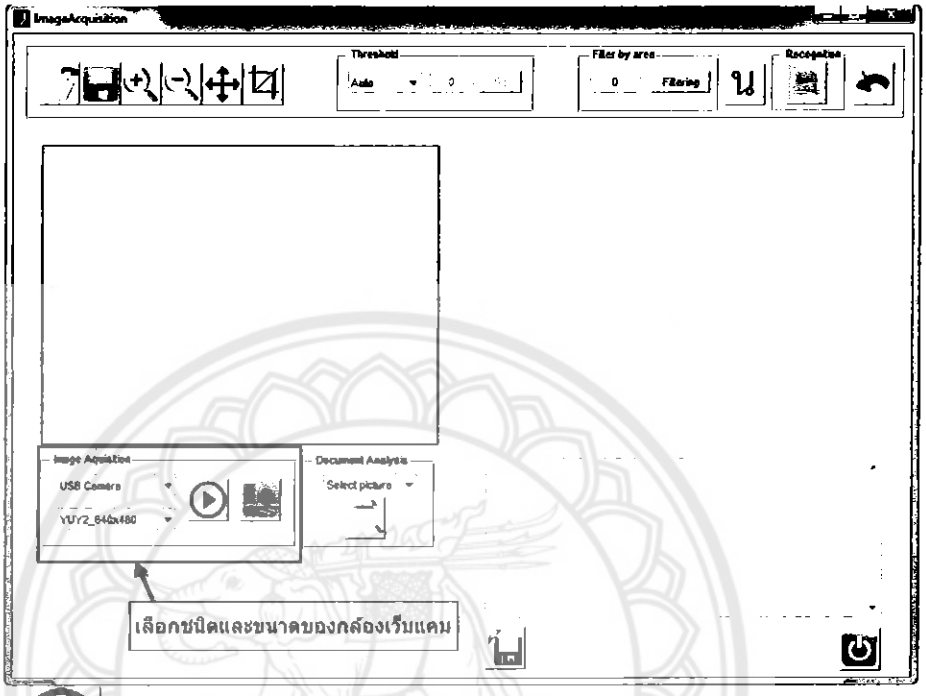
- Auto Threshold เป็นการปรับภาพเอกสารที่มีลักษณะปกติ
- Manual คือการปรับภาพเอกสารที่มีลักษณะปกติแต่ต้องการปรับค่าเอง
- Adaptive Median คือการปรับภาพเอกสารที่มีลักษณะแสดงสว่างของภาพไม่เท่ากัน
- Adaptive Segment คือการปรับภาพเอกสารที่มีลักษณะแสดงสว่างของภาพไม่เท่ากัน



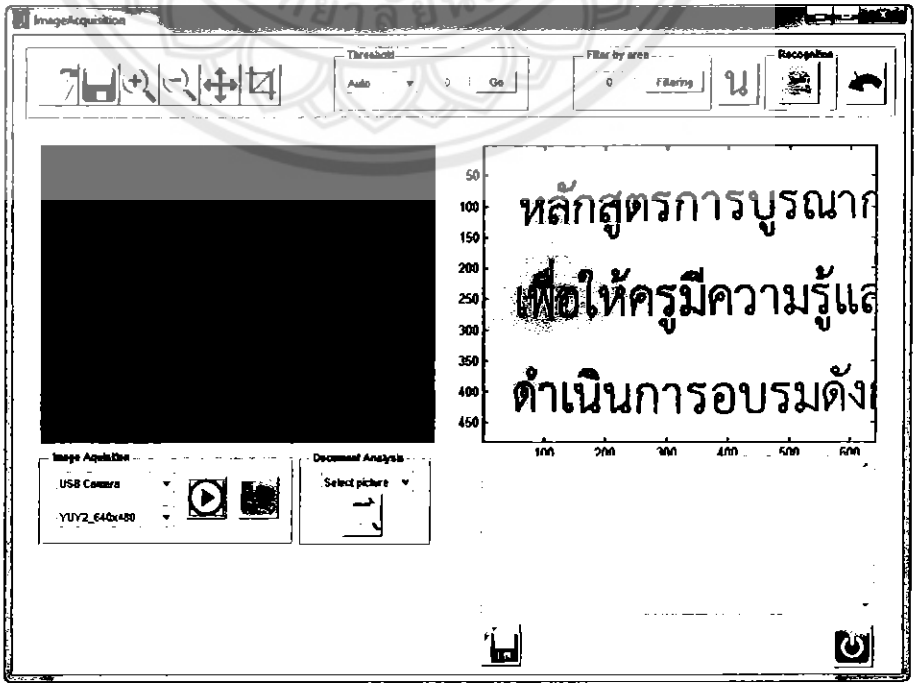
: (Filter) ใช้เพื่อทำการกำจัดสัญญาณรบกวนหรือสิ่งแปลกปลอมที่เข้ามาในภาพถ่ายเอกสารเพื่อป้องกันการรู้จำที่ผิดพลาด

3. ขั้นตอนการถ่ายภาพเอกสารด้วยกล้องเว็บแคม

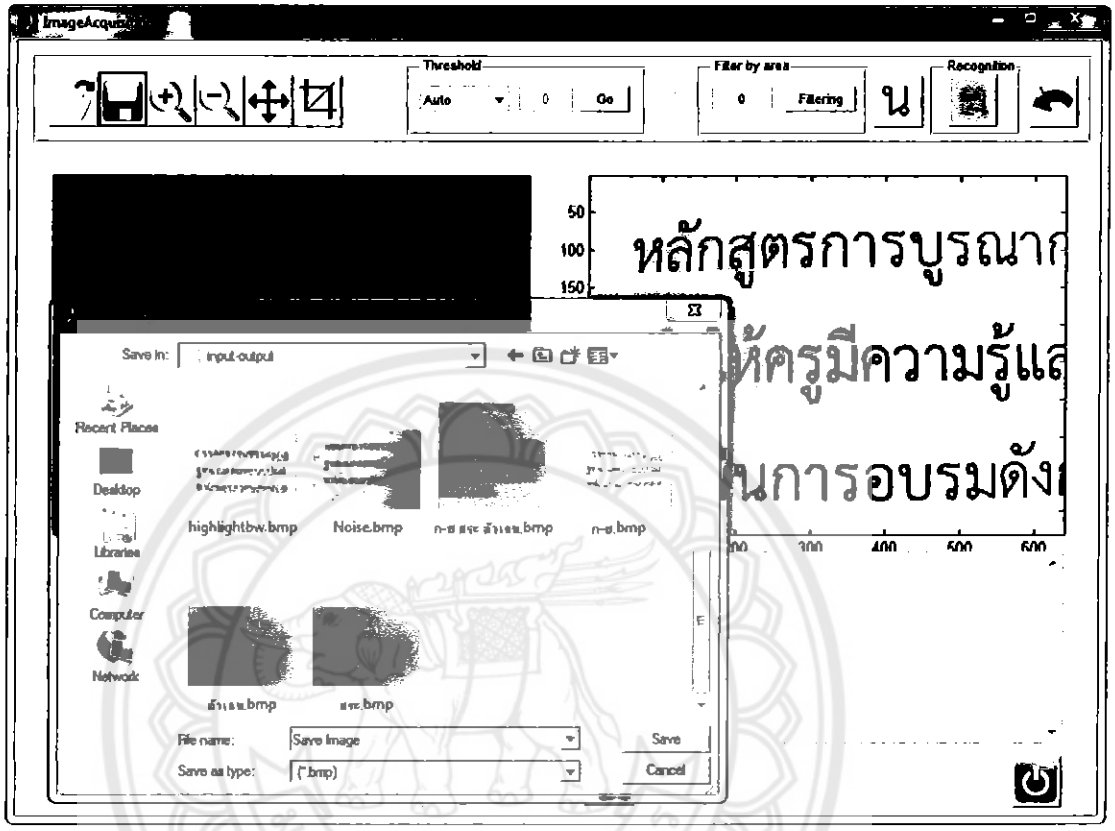
1. เมื่อเปิดโปรแกรมขึ้นมาแล้วให้ทำการเลือกชนิดและความละเอียดของกล้องเว็บแคมที่ต้องการใช้งาน



2. กดปุ่ม เพื่อเริ่มการทำงานของกล้องเว็บแคมจากนั้นทำการถ่ายภาพเอกสารโดยที่ปุ่ม เพื่อทำการถ่ายภาพหลังจากนั้นให้กดปุ่ม เพื่อหยุดการทำงานของกล้องเว็บแคม

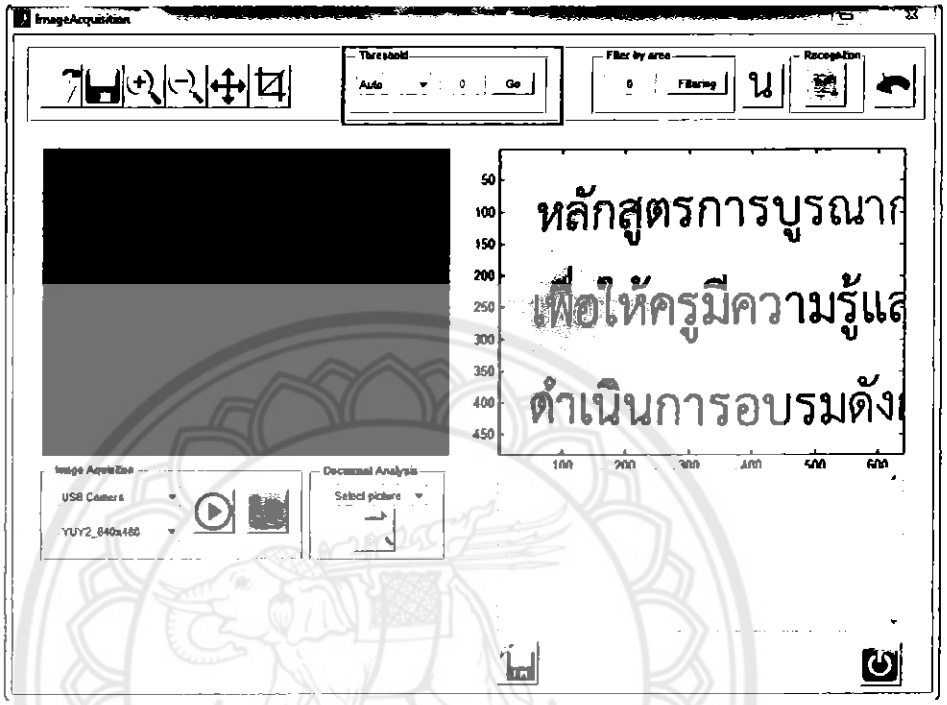


3. จากนั้นเมื่อถ่ายภาพเอกสารเรียบร้อยแล้วจะปรากฏภาพเอกสารที่ถ่ายในช่อง Image Display ขึ้น จากนั้นทำการ Save Image โดยกดปุ่ม ตั้งชื่อภาพแล้วกด save เพื่อนำไปใช้งานในครั้งถัดไป



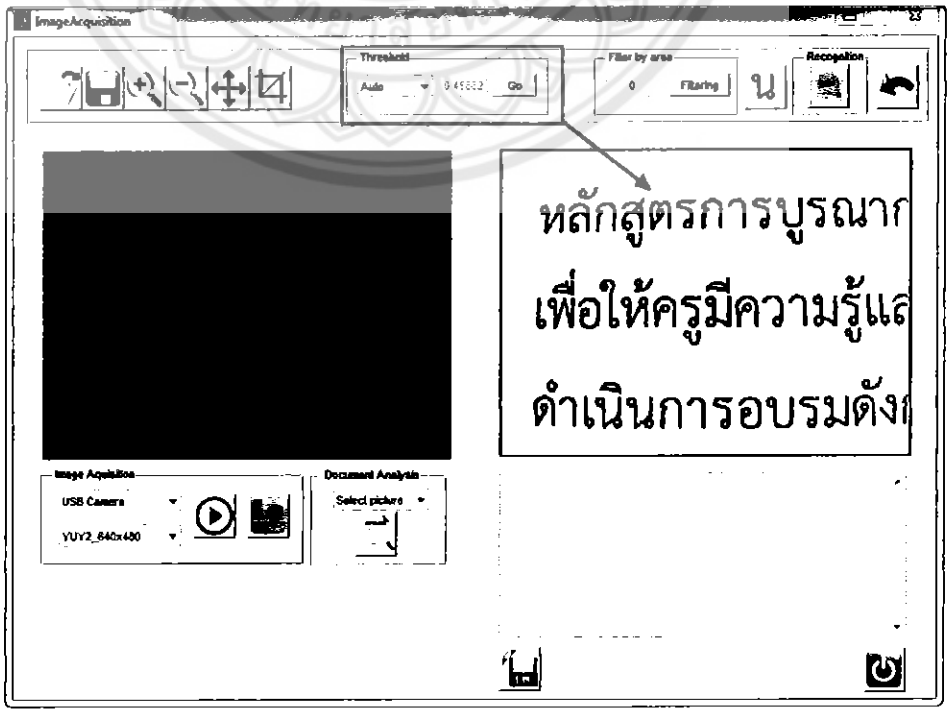
4. ขั้นตอนการแปลงภาพเป็นขาวดำสองระดับ

1. เมื่อทำการถ่ายภาพเอกสารหรือทำการเปิดภาพเอกสารจากแฟ้มจะต้องทำการปรับภาพเป็นขาวดำสองระดับก่อน



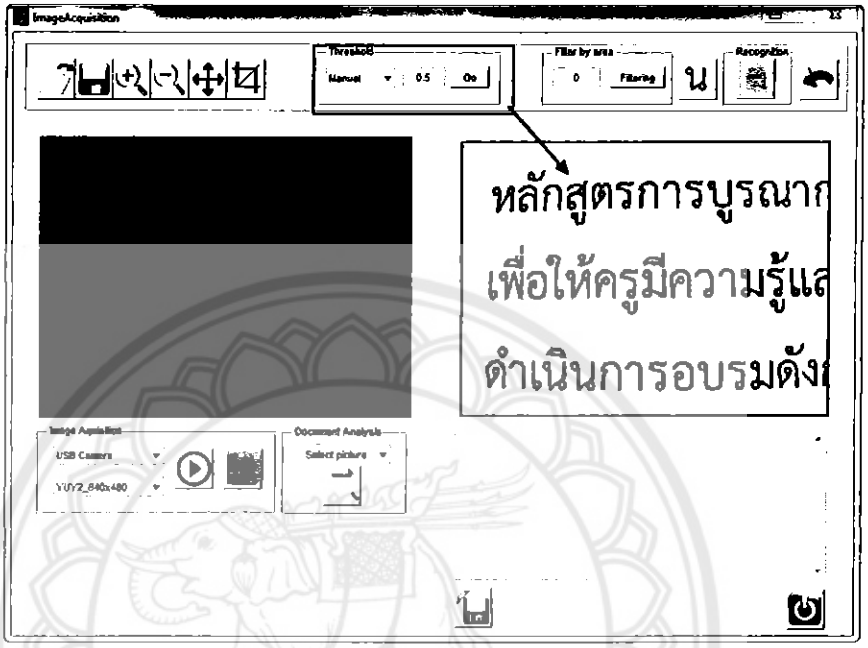
การเลือกปรับค่าขาวดำสองระดับแบบ Auto

2. การเลือกการปรับภาพแบบ Auto จะใช้งานกับเอกสารปกติทั่วไปที่ไม่มีสัญญาณรบกวนมากและสภาพแสดงมีความเป็นปกติ



การเลือกปรับค่าขาวดำสองระดับแบบ Manual

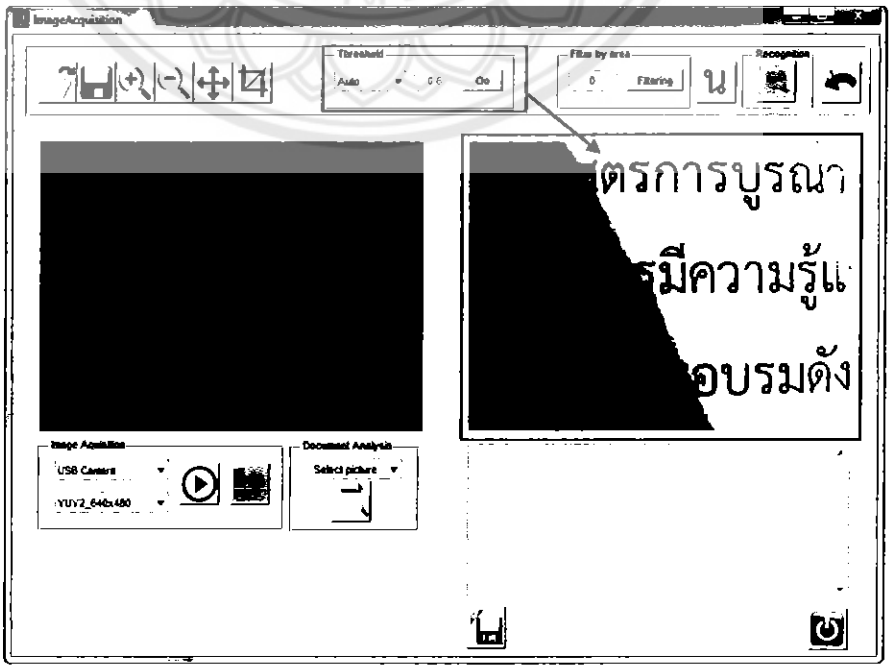
3. การเลือกปรับภาพแบบ Manual จะใช้งานกับเอกสารปกติทั่วไปที่ไม่มีสัญญาณรบกวนมากและสภาพแสดงมีความเป็นปกติแต่มีความแตกต่างกันกรณีที่แบบ Auto อาจทำให้ตัวหนังสือในภาพขาดจากกันจึงต้องเปลี่ยนมาใช้แบบ Manual



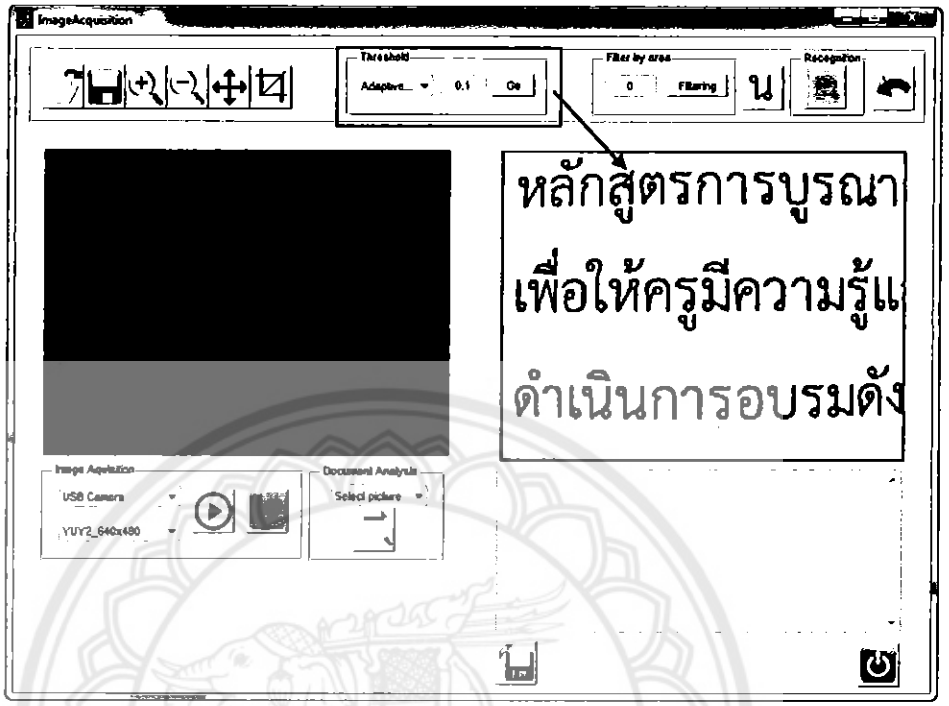
การเลือกปรับค่าขาวดำสองระดับแบบ Adaptive Median

4. การเลือกปรับภาพแบบ Adaptive Median จะใช้ในกรณีที่การรับภาพเอกสารมีความสว่างของสภาพแสงที่ไม่สม่ำเสมอกัน

ตัวอย่างภาพที่มีปัญหาเรื่องความสว่างของสภาพแสดงที่ไม่สม่ำเสมอกันหากใช้แบบ Auto

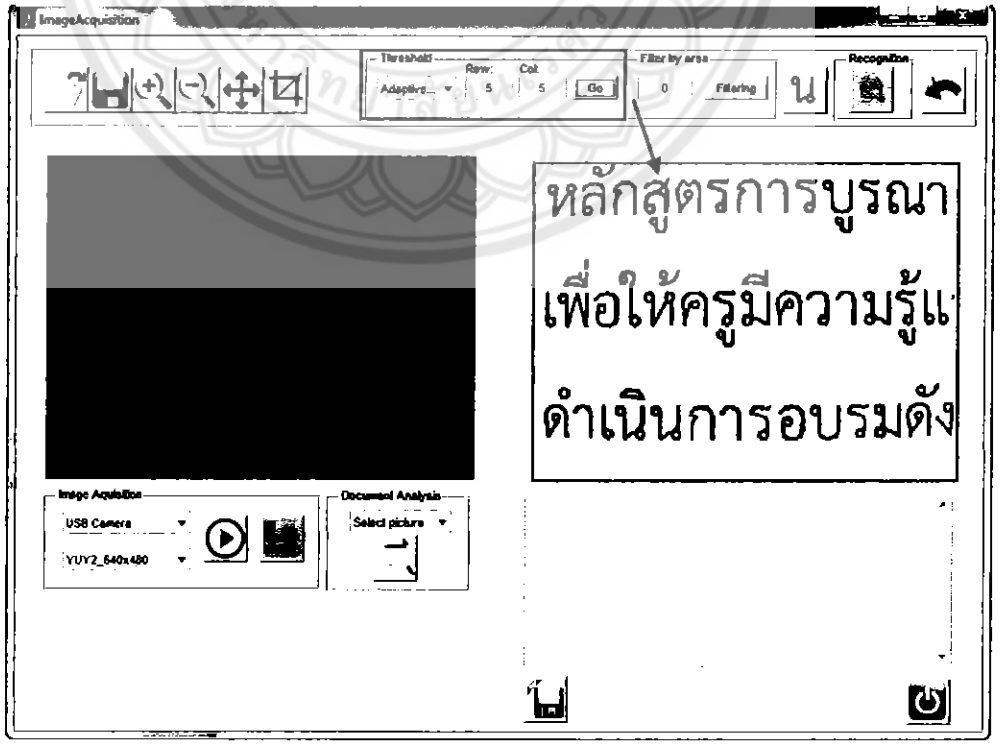


ดังนั้นหากการรับภาพมีความสว่างของสภาพแสงที่ไม่สม่ำเสมอกันจะต้องใช้การแปลงเป็นภาพขาวดำสองระดับแบบ Adaptive Median แทน



การเลือกปรับค่าขาวดำสองระดับแบบ Adaptive Segment

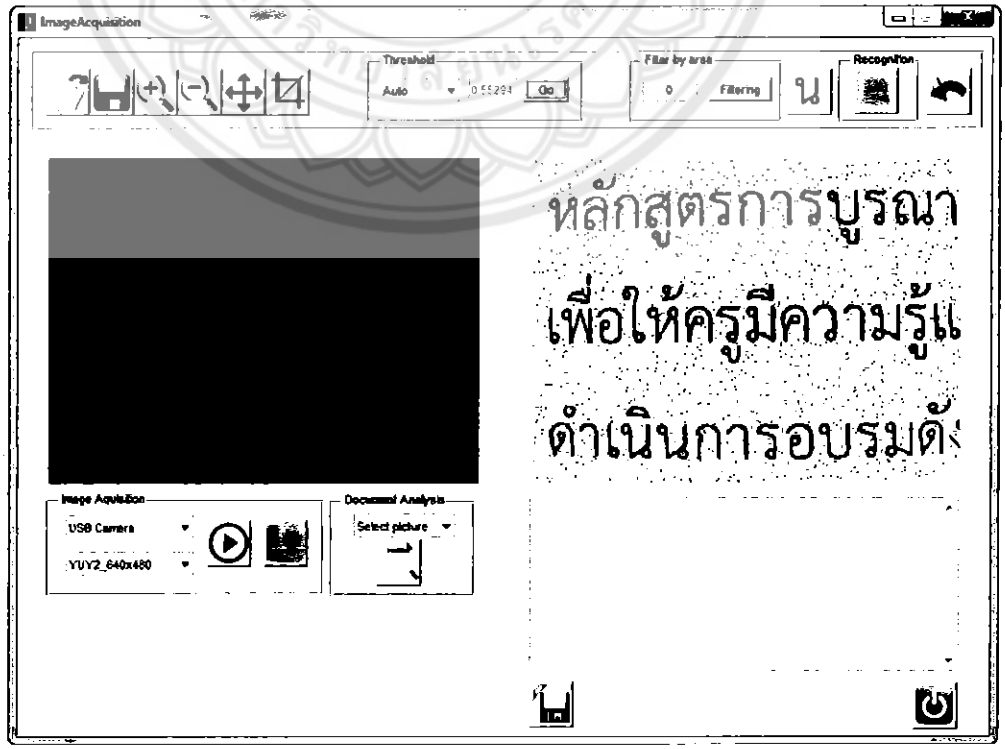
5. การเลือกปรับค่าขาวดำสองระดับแบบ Adaptive Segment จะมีลักษณะที่คล้ายกับ Adaptive Median แต่แทนที่จะใส่ค่าเทรสโฮลเข้าไป แต่ใส่เป็นค่าของแถวและคอลัมภ์ที่จะทำการแบ่งภาพ



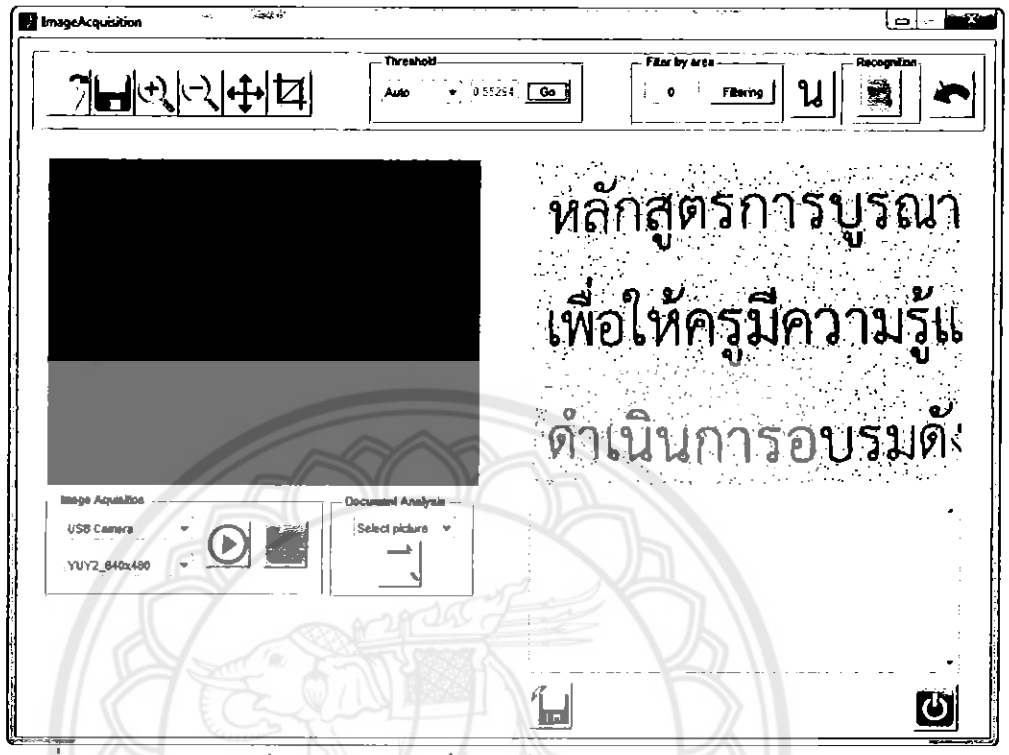
5. ขั้นตอนการทำงานการแปลงข้อความจากภาพเอกสารเป็นไฟล์ข้อความ

5.1 ทำการรับภาพเอกสารจากกล้องเว็บแคมโดยการถ่ายภาพเอกสารในส่วนที่ต้องการแปลงเป็นไฟล์ข้อความ

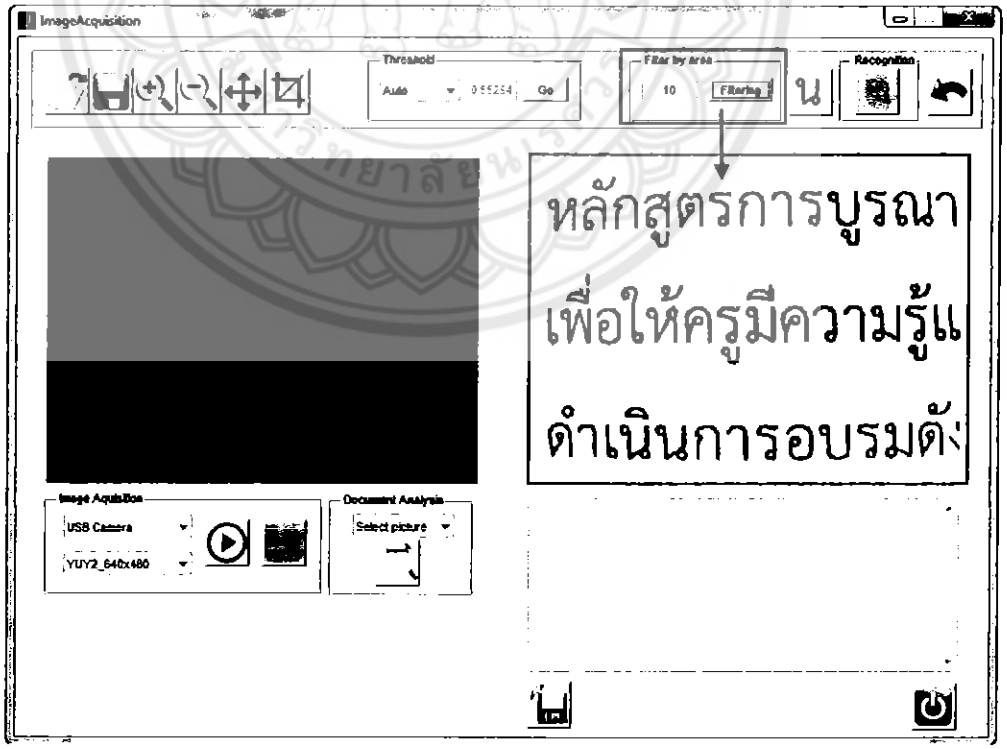
5.2 ทำการแปลงภาพเอกสารเป็นภาพขาวดำสองระดับในที่นี่จะใช้แบบ Auto



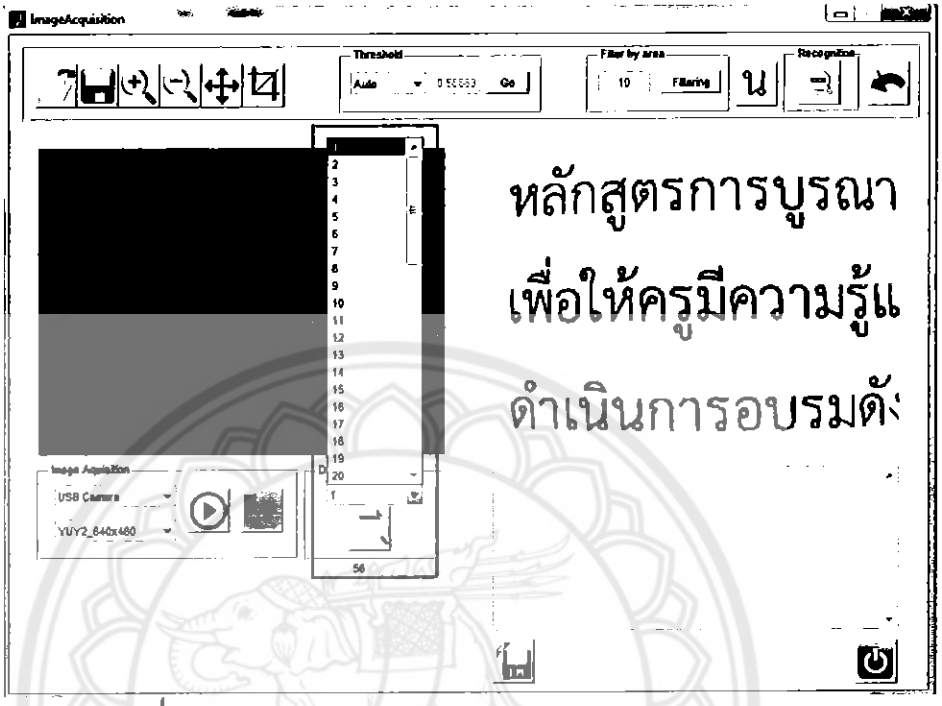
5.3 เมื่อทำการแปลงภาพเป็นขาวดำสองระดับแล้วหากมีสัญญาณรบกวนหรือสิ่งแปลกปลอมภายในภาพเอกสารให้กดปุ่ม "Filter" และใส่ค่าที่ต้องการลงใช้ช่องว่าง



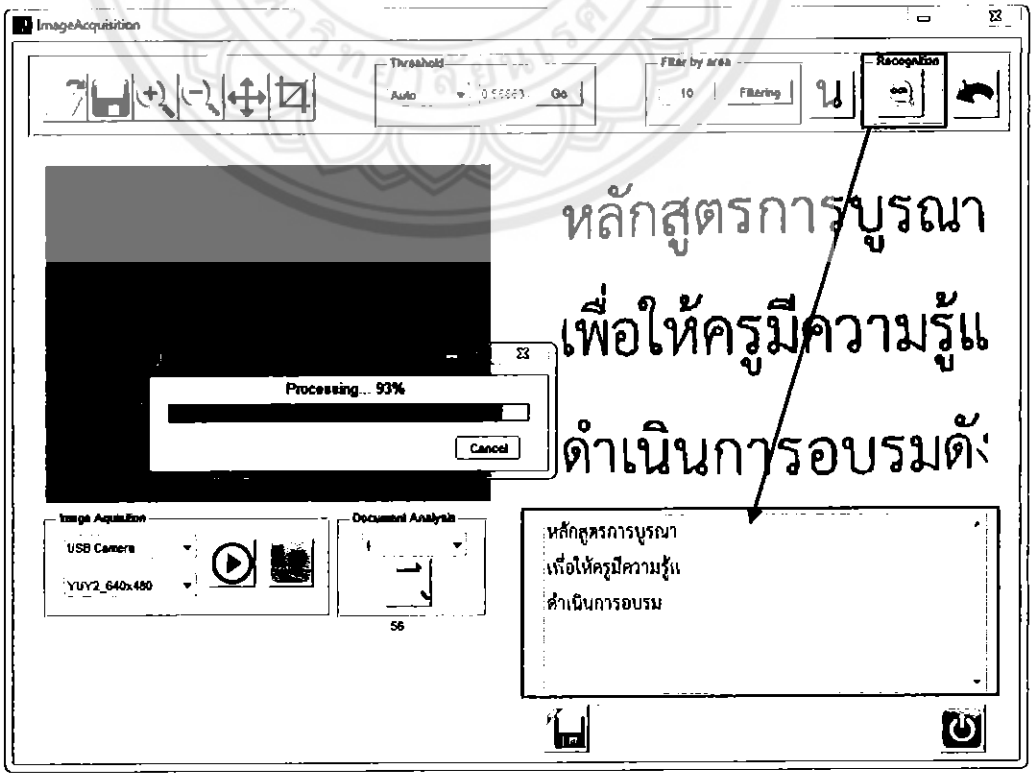
หลังจากที่กดปุ่ม "Filter" แล้วสัญญารบกวนก็จะหายไป



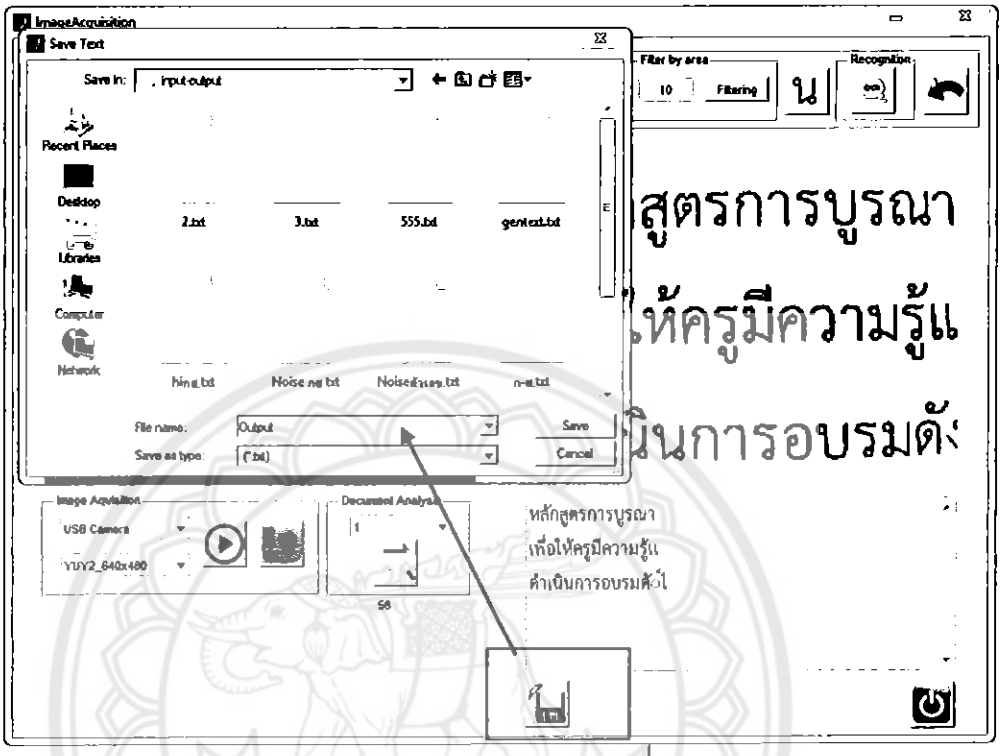
5.4 กดปุ่ม เพื่อทำการวิเคราะห์ตัวอักษรว่าเป็นตัวอักษรชนิดใด เมื่อกดปุ่ม จะปรากฏจำนวนตัวอักษรที่อ่านได้ทั้งหมดว่ามีจำนวนกี่ตัว



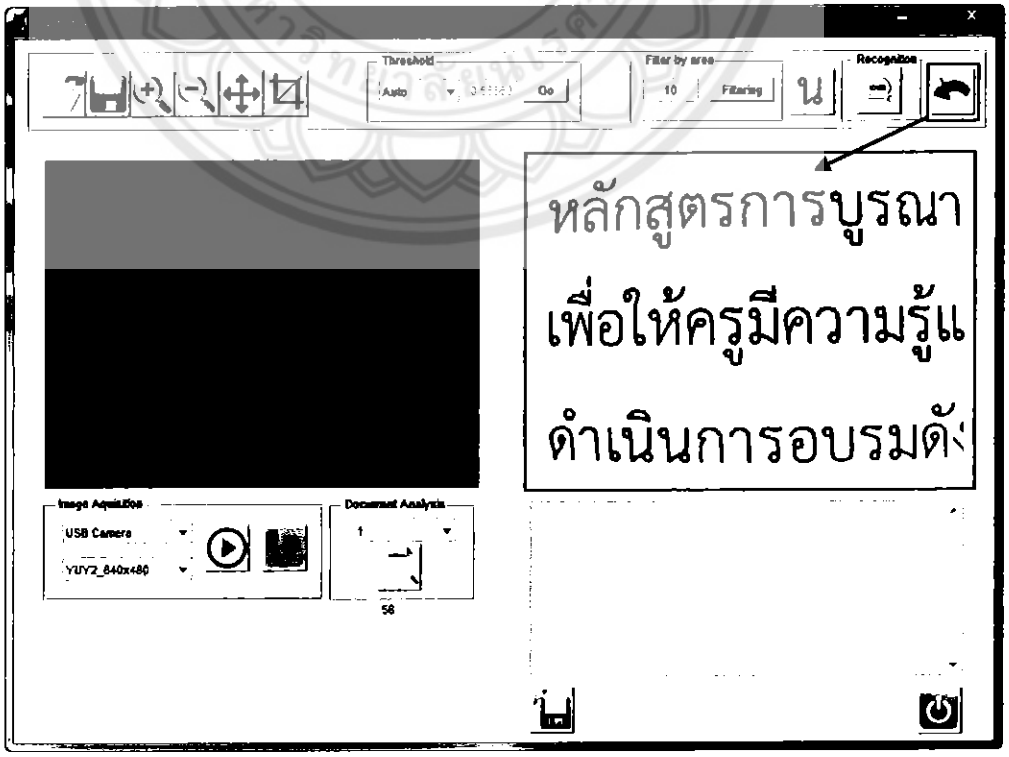
5.5 กดปุ่ม เพื่อทำการแปลงภาพเอกสารที่ได้ทำการปรับปรุงเรียบร้อยแล้วมาเป็นไฟล์ข้อความเพื่อนำไปใช้งานในขั้นตอนต่อไป



5.6 กดปุ่ม เพื่อทำการบันทึกข้อความที่แปลงออกมาแล้วให้เป็นไฟล์ข้อความโดย ตั้งชื่อของไฟล์ แล้วก็กดตรงคำว่า Save โปรแกรมก็จะทำการบันทึก



5.7 หากต้องการยกเลิกกระทำที่ทำไปทั้งหมดให้กดปุ่ม เพื่อกลับไปเริ่มทำงานใหม่



5.8 หากต้องการออกจากโปรแกรมให้กดปุ่ม เพื่อทำการออกจากโปรแกรม