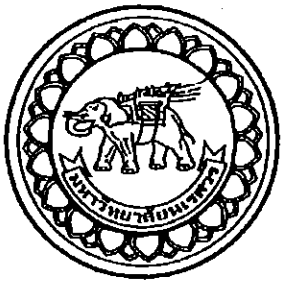


การตรวจวัดความเอียงและหมุนกลับของภาพเอกสารโดยอัตโนมัติ

Automatic Document Image Skew angle detection and correction

นายพงศกร ทองก่อ รหัส 51362053

นายพิศาล ห้าวแสน รหัส 51362091

ห้องสมุด คณะวิศวกรรมศาสตร์
วันที่รับ 25 มิ.ย. 2556
เลขทะเบียน 1627805
เลขเรียกหนังสือ ปร.
มหาวิทยาลัยนเรศวร W112ก 2554

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

ปีการศึกษา 2554

ใบรับรองปริญญานิพนธ์

หัวข้อโครงงาน การตรวจวัดความเอียงและการหมุนกลับของภาพเอกสารโดยอัตโนมัติ

ผู้ดำเนินโครงงาน นายพงศกร ทองก่อ รหัส 51362053

นายพิศาล ห้าวแสน รหัส 51362091

อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร.พนมขวัญ ริยะมงคล

สาขาวิชา วิศวกรรมคอมพิวเตอร์

ภาควิชา วิศวกรรมไฟฟ้าและคอมพิวเตอร์

ปีการศึกษา 2554

.....

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร อนุมัติให้โครงงานฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์

.....ประธานกรรมการ

(อาจารย์ รัฐภูมิ วรานุสาสน์)

.....กรรมการ

(อาจารย์ เศรษฐา ตั้งค้าวานิช)

.....กรรมการ

(อาจารย์ ภาณุพงศ์ สอนคม)

.....กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.พนมขวัญ ริยะมงคล)

หัวข้อโครงงาน	**การตรวจวัดความเอียงและหมุนกลับของภาพเอกสารโดยอัตโนมัติ**			
ผู้ดำเนินโครงงาน	**นายพงศกร**	**ทองก่อ**	**รหัส**	**51362053**
	นายพิศาล	**ห้าวแสน**	**รหัส**	**51362091**
อาจารย์ที่ปรึกษา	**ผู้ช่วยศาสตราจารย์ ดร. พนมขวัญ ริยะมงคล**			
สาขาวิชา	**วิศวกรรมคอมพิวเตอร์**			
ภาควิชา	**วิศวกรรมไฟฟ้าและคอมพิวเตอร์**			
ปีการศึกษา	2554			

.....

บทคัดย่อ

ความเอียงของภาพเอกสารเป็นข้อจำกัดหนึ่งที่ทำให้ประสิทธิภาพในการวิเคราะห์ข้อมูลในภาพเอกสารลดลง โครงงานนี้จึงได้พัฒนาขึ้นเพื่อที่จะทำให้ภาพเอกสารตรงโดยอัตโนมัติ ได้นำเอาเทคนิคการแบ่งภาพออกเป็นส่วนๆเข้ามาช่วยและ ใช้การหยดสีซึ่งช่วยให้สามารถแยกแยะส่วนของตัวอักษรและส่วนของรูปภาพได้ โปรแกรมมีประสิทธิภาพในการตรวจสอบความเอียงของภาพเอกสารได้ถูกต้องประมาณร้อยละ 80 สามารถใช้ได้กับภาพเอกสารที่มีรูปภาพและ ภาพเอกสารที่เอียงได้สูงสุด 90 องศา

Project title	Automatic Document Image Skew Angle Detection and Correction
Name	Mr. Pongsakorn Tongko ID. 51362053
	Mr. Pisan Hawsan ID. 51362091
Project advisor	Assistant Professor Panomkhawn Riyamongkol, Ph.D.
Major	Computer Engineering
Department	Electrical and Computer Engineering
Academic year	2011

.....

Abstract

The skew of the document image is one of the performance limitations of data analysis. This project has been developed to upright the skewed document image automatically. Dividing image into sections technique has been used. After that, Blob Coloring method has been applied to makes it possible to distinguish characters and parts of the picture. The efficiency of this program is about 80 %. This program can be used with the document that include picture and texts and rotated within ±90 degrees.

กิตติกรรมประกาศ

โครงงานนี้สำเร็จลุล่วงไปได้ด้วยดี อันเนื่องมาจากความอนุเคราะห์ในหลายๆสิ่ง ในหลายๆ ด้าน จากผู้มีพระคุณทุกท่าน ซึ่งก่อนอื่นใด ขอขอบพระคุณ พ่อแม่และครอบครัวที่ให้โอกาสทางการศึกษา และเป็นกำลังใจที่ดีเยี่ยมตลอดมา อีกทั้งยังเป็นกำลังทางทุนทรัพย์ที่สำคัญ ขอขอบพระคุณเพื่อนๆร่วมชั้นที่ให้คำปรึกษาในด้านต่างๆ ช่วยพูดสร้างแรงบันดาลใจ สร้างกำลังใจดีๆให้แก่กัน และที่สำคัญยิ่ง ที่เห็นจะขาดไม่ได้เลย ต้องขอขอบพระคุณ อาจารย์พนมขวัญ ริยะมงคล อาจารย์ที่ปรึกษาโครงงาน ที่เสียสละทั้งกำลังกายและกำลังใจในการทุ่มเทให้กับโครงงานนี้ ทั้งให้คำปรึกษา คำชี้แนะ ช่วยแก้ไขข้อผิดพลาดต่างๆ และคอยเป็นแรงกระตุ้นแรงผลักดันให้โครงงานนี้สำเร็จออกมาได้อย่างสมบูรณ์ตามเวลาที่กำหนด

ท้ายที่สุดนี้ ขอขอบพระคุณทุกๆท่านที่สละเวลาอันมีค่าให้ความสนใจในโครงงานนี้ ซึ่งผู้พัฒนาโครงงานหวังเป็นอย่างยิ่งว่า โครงงานนี้จะสร้างช่วยสร้างแนวทาง สร้างประโยชน์อย่างใดอย่างหนึ่งแก่ท่านไม่มากก็น้อย

นายพงศกร ทองก่อ

นายพิศาล ห้าวแสน

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....ก

บทคัดย่อภาษาอังกฤษ.....ข

กิตติกรรมประกาศ.....ค

สารบัญ.....ง

สารบัญรูป.....ฉ

สารบัญตาราง.....ช

บทที่ 1 บทนำ.....1

1.1 ความเป็นมาและความสำคัญของปัญหา..... 1

1.2 วัตถุประสงค์..... 2

1.3 ขอบเขตของโครงงาน..... 2

1.4 แผนการดำเนินงาน..... 3

1.5 ผลที่คาดว่าจะได้รับ..... 4

1.6 งบประมาณของโครงงาน..... 5

บทที่ 2 หลักการและทฤษฎีเบื้องต้น.....6

2.1 พื้นฐานเกี่ยวกับการประมวลผลภาพ..... 6

2.2 การแยกวัตถุออกจากภาพโดยการเทรชโฮลด์ (Threshold)..... 9

2.3 การแบ่งภาพออกเป็นส่วนๆ (block)..... 12

2.4 การหมุนภาพเอกสาร (Rotation)..... 13

2.5 การหาความเอียงของภาพเอกสารโดยการฉายแสง (Projection)..... 15

2.6 การหยดสี (Blob Coloring)..... 16

บทที่ 3 วิธีการดำเนินงาน.....17

3.1 แบ่งภาพเอกสารออกเป็นส่วนๆ(block)..... 18

3.2 ตรวจสอบแต่ละบล็อกว่าเป็นตัวอักษรหรือเปล่า..... 22

สารบัญ(ต่อ)

บทที่ 3 วิธีการดำเนินงาน.....16

3.1 แบ่งภาพเอกสารออกเป็นส่วนๆ(block)..... 17

3.2 ตรวจสอบแต่ละบล็อกว่าเป็นตัวอักษรหรือเปล่า..... 21

3.3 หมุนภาพและทำการฉายแสงในแนวนอน(HorizontalProjection Profile)..... 25

3.4 เลือกมุมที่จะนำไปใช้หมุนภาพเอกสารให้ตรง..... 31

3.5 หมุนภาพเอกสารให้ตรง..... 31

บทที่ 4 ผลการดำเนินงาน.....32

4.1 ผลการทดสอบโปรแกรมกับกลุ่มของภาพเอกสารที่มีเฉพาะตัวหนังสือ..... 34

4.2 ผลการทดสอบโปรแกรมกับกลุ่มของภาพเอกสารที่มีรูปภาพปะปนอยู่ด้วย..... 36

4.3 ตารางการเปรียบเทียบรูปภาพหลายๆ แบบ..... 38

บทที่ 5 บทสรุป.....46

5.1 สรุปและวิเคราะห์ผลการทดลอง..... 46

5.2 ปัญหาและแนวทางการแก้ไขปัญหา..... 46

5.3 แนวทางในการพัฒนาครั้งต่อไป..... 47

เอกสารอ้างอิง..... 48

ภาคผนวก..... 49

ประวัติผู้เขียนโครงงาน..... 59

สารบัญตาราง

ตารางที่ หน้า

3.1 แสดงขนาดบล็อกและอัตราส่วนความถูกต้องต่อเวลาของภาพเอกสารตัวหนังสือล้วน.....19

3.2 แสดงขนาดบล็อกและอัตราส่วนความถูกต้องต่อเวลาของภาพเอกสารที่มีรูปปน.....20

3.3 แสดงการหยดสี จำนวนพื้นที่ของการหยดสี.....22

3.4 การหาค่าเบี่ยงเบนมาตรฐานของตัวหนังสือของภาพเอกสารเป็นตัวหนังสือล้วน.....24

3.5 การหาค่าเบี่ยงเบนมาตรฐานของตัวหนังสือของภาพเอกสารที่มีรูปปน.....24

3.6 แสดงการหมุนภาพเอกสารที่ละ 5 องศา จากมุม -90 ถึง มุม 90 องศา.....30

3.7 แสดงการหมุนภาพเอกสารที่ละ 2 องศา จากมุม 6 ถึง มุม 14 องศา.....31

3.8 แสดงการหมุนภาพเอกสารที่ละ 1 องศา จากมุม 7 ถึง มุม 9 องศา.....32

3.9 แสดงค่าแอมพลิจูดสูงสุดของทุกบล็อกที่เป็นตัวอักษร.....33

4.1 ผลการทดสอบโปรแกรมกับเอกสารที่มีเฉพาะตัวหนังสือ.....36

4.2 ผลการทดสอบโปรแกรมกับภาพเอกสารที่มีรูปปนอยู่ด้วย.....38

4.3 แสดงการตรวจสอบความถูกต้องกับภาพเอกสารหลายๆแบบ.....41

สารบัญรูป

รูปที่ หน้า

2.1 แสดงการขยายภาพแบบบิตแมป....6
2.2 แสดงการขยายของภาพแบบเวกเตอร์....7
2.3 แสดงภาพโหมดสีแบบบิตแมป....7
2.4 แสดงภาพโหมดสีระดับเทา (Gray scale)....8
2.5 แสดงภาพโหมดสีแบบ 2 โทนสี (Duotone)....8
2.6 แสดงภาพโหมดสีแบบอาร์จีบี (RGB)....8
2.7 แสดงภาพโหมดสีแบบแล็บคัลเลอร์ (Lab Color)....9
2.8 แสดงภาพโหมดสีแบบ 8 บิตต่อ 1 จุดสี (Multichannel)....9
2.9 แสดงการแบ่งวัตถุออกจากพื้นหลัง....10
2.10 แสดงการเปลี่ยนแปลงของแต่ละบล็อก....12
2.11 แสดงการหมุนภาพรอบจุดหมุน (0,0)....13
2.12 แสดงการหมุนภาพเมื่อจุดหมุนไม่ได้อยู่ที่ (0,0)....14
2.13 การฉายแสงในแนวนอน (Horizontal Projection Profile)....15
2.14 แสดงการฉายแสงในแนวนอนของภาพเอกสารตรงและภาพเอกสารเอียง....15
2.15 แสดงการหยดสี (Blob Coloring) และจำนวนของพื้นวัตถุที่พบ....16
2.17 แสดงลักษณะการเชื่อมต่อกัน แบบ 8 ทิศทางและ 4 ทิศทาง....16
2.18 แสดงภาพตัวอย่างที่ได้จากการหยดสี (Blob Coloring) แบบ 8 ทิศทาง....16
3.1 แสดงแผนภาพขั้นตอนการดำเนินงาน....17
3.2 แสดงขนาดของรูปภาพและการแบ่งรูปภาพออกเป็นบล็อกๆ....18
3.3 แสดงอัตราส่วนความถูกต้องต่อขนาดของบล็อกที่เป็นภาพเอกสารตัวหนังสือล้วน....19
3.4 แสดงอัตราส่วนความถูกต้องต่อขนาดของบล็อกที่เป็นภาพเอกสารที่มีรูปปน....20
3.5 แสดงตัวอย่างการคำนวณหาจำนวนบล็อกของภาพเอกสารที่มีรูปภาพปน....21
3.6 แสดงบล็อกที่เป็นตัวอักษรจากการหยดสี (Blob Coloring)....26
3.7 แสดงภาพเริ่มต้นและภาพผลลัพธ์....33
4.1 แสดงหน้าจอหลักของโปรแกรม....34
4.2 แสดงคำสั่งของแมทแลบ (MATLAB) ใช้หมุนภาพเอกสารตามมุมที่ต้องการ....35

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญ

ที่มาและเหตุผลที่ทำให้เกิดโครงงานการตรวจสอบมุมเอียงและปรับภาพให้ตรงโดยอัตโนมัตินั้น มีเหตุผลหลักๆ อยู่ 2 ข้อด้วยกันคือ

1.1.1 เป็นกระบวนการประมวลผลเบื้องต้นเพื่อเตรียมภาพเอกสารก่อนเข้าสู่กระบวนการอื่นๆ

ในการวิเคราะห์ตัวอักษรในภาพเอกสารนั้นมักจะมีปัญหาเกี่ยวกับความชัดเจนของตัวอักษรในภาพ ไม่ว่าจะเป็นสัญญาณรบกวน (Noise) ต่างๆ, ความสว่างของแสง, ความเบลอความคมชัด และอีกหนึ่งปัญหาที่สำคัญของการวิเคราะห์ตัวอักษรในภาพเอกสารก็คือ "ภาพเอกสารมีความเอียง" ไม่ว่าจะเป็นกระบวนการรู้จำตัวอักษร (OCR), กระบวนการแปลภาษาจากภาพเอกสาร, การตรวจสอบฟอนต์จากภาพเอกสาร และกระบวนการอื่นๆที่เกี่ยวข้องกับการวิเคราะห์ตัวอักษรในภาพเอกสาร ความเอียงถือเป็นข้อจำกัดหนึ่งที่ทำให้กระบวนการต่างๆ มีประสิทธิภาพลดลง ผู้จัดทำจึงมีแนวคิดที่จะเขียนโปรแกรมตรวจหาค่าความเอียงและหมุนกลับภาพเอกสารนั้นให้ตรงโดยอัตโนมัติ

1.1.2 เพื่อให้ง่ายต่อการอ่านและการทำความเข้าใจของมนุษย์

ในบางเวลาเรามีความจำเป็นต้องเก็บข้อมูลในเอกสารที่เราสนใจไว้ โดยปกติทั่วไปเราก็จะนำไปถ่ายเอกสาร หรือบางคนอาจจดใส่กระดาษแต่ด้วยความก้าวหน้าของเทคโนโลยีในปัจจุบันไม่ว่าจะเป็นมือถือหรือคอมพิวเตอร์ต่างก็มีกล้องกันทั้งนั้น ซึ่งบางคนก็จะทำการเก็บข้อมูลในเอกสารด้วยการถ่ายรูป ซึ่งข้อดีคือสะดวก รวดเร็ว และง่าย ใครก็สามารถทำได้ แต่ข้อเสียก็มี คือภาพเอกสารที่ได้อาจไม่ชัดเจน อาจมีการบิดเบี้ยวไม่ตรง เวลาถ่ายรูปมือไม่นิ่งภาพเอกสารที่ออกมาเลยเอียงไม่เป็นไปตามที่ต้องการ ผู้จัดทำจึงมีแนวคิดที่จะเขียนโปรแกรมตรวจหาค่าความเอียงและหมุนกลับภาพเอกสารนั้นให้ตรงโดยอัตโนมัติเพื่อให้ง่ายต่อการอ่านข้อมูลในภาพเอกสาร

1.2 วัตถุประสงค์ของโครงงาน

1.2.1. เพื่อศึกษาหลักการและกระบวนการต่างๆของการปรับความเอียงของภาพเอกสารตั้งแต่การรับข้อมูล (input) จนกระทั่งได้ผลลัพธ์ตามที่ต้องการ (output)

1.2.2. เพื่อให้ได้โปรแกรมสำหรับปรับภาพเอกสารให้ตรงโดยอัตโนมัติ

1.3 ขอบข่ายโครงงาน

1.3.1 แสงสว่างในบริเวณที่ถ่ายภาพต้องสว่างเพียงพอที่จะเห็นภาพที่ถ่ายได้ชัดเจน

1.3.2 ความละเอียดของกล้องที่ถ่ายภาพเอกสารอยู่ในระดับ 2 ล้านพิกเซล (Megapixel) ขึ้นไป

1.3.3 ภาพเอกสารที่ใช้ในระบบประมวลผลต้องเอียงในแนวระนาบหรือภาพเอกสารต้องตั้งฉากกับกล้องที่ใช้ถ่ายภาพเอกสาร

1.3.4 ภาพเอกสารที่ต้องการปรับต้องเอียงในแนวระนาบไม่เกิน 90 องศา

1.3.5 ถ้าเป็นภาพเอกสารที่ประกอบไปด้วยตัวหนังสือและรูปภาพ บริเวณที่เป็นตัวหนังสือต้องมีมากกว่าบริเวณที่เป็นรูปภาพประมาณร้อยละ 70

1.4 ขั้นตอนการดำเนินงาน

1.4.1 ศึกษาค้นคว้าข้อมูลเกี่ยวกับการประมวลผลภาพ (Image Processing)

1.4.2 ศึกษาค้นคว้าข้อมูลเกี่ยวกับการเขียนโปรแกรมเชื่อมต่อกับกล้องเว็บแคม (Webcam) โดยใช้โปแกรมแมทแลบ (MATLAB)

1.4.3 เขียนโปรแกรมรับภาพเอกสารจากกล้องเว็บแคม (Webcam) และบันทึกภาพ

1.4.4 เขียนโปรแกรมตรวจหาความเอียงของภาพเอกสาร

1.4.5 เขียนโปรแกรมหมุนกลับภาพให้ตรงโดยอัตโนมัติ

1.4.6 ทดสอบการทำงานของระบบ

1.4.7 แก้ไขข้อผิดพลาดและเก็บรายละเอียดส่วนต่างๆของโครงงาน

1.4.8 สรุปผลการทดลองและจัดทำรูปเล่มโครงงาน

1.5 แผนการดำเนินงาน

ตารางที่ 1.1 แสดงขั้นตอนการดำเนินงาน

เดือน / แผนการปฏิบัติงาน	พ.ศ. 2554						พ.ศ. 2555	
	มิถุนายน	กรกฎาคม	สิงหาคม	กันยายน	พฤศจิกายน	ธันวาคม	มกราคม	กุมภาพันธ์
1.ศึกษาเกี่ยวกับประมวลผลภาพ	← →							
2.ศึกษาการเขียนโปรแกรมเชื่อมต่อเว็บแคม	← →	← →						
3.เขียนโปรแกรมรับภาพเอกสารจากเว็บแคม และบันทึกภาพ		← →						
4.เขียนโปรแกรมตรวจหาความเอียงของภาพเอกสาร			← →	← →				
5.เขียนโปรแกรมหมุนกลับภาพให้ตรงโดยอัตโนมัติ				← →	← →			
6.ทดสอบการทำงานของระบบ						← →		
7.แก้ไขข้อผิดพลาดและเก็บรายละเอียดส่วนต่างๆของโครงงาน							← →	← →
8.สรุปผลการทดลองและจัดทำรูปเล่มโครงงาน								← →

1.6 ผลที่คาดว่าจะได้รับ

1.6.1 สามารถหมุนกลับภาพเอกสารที่เป็นตัวหนังสือล้วนให้ตรงได้ 80% ขึ้นไป และ60%ขึ้นไปสำหรับภาพเอกสารที่มีรูปภาพปะปน โดยใช้จำนวนภาพเอกสารที่สามารถหมุนได้ตรงต่อจำนวนภาพเอกสารทั้งหมด เป็นตัวชี้วัด

1.6.2 สามารถนำความรู้เกี่ยวกับการวิเคราะห์รูปภาพไปประยุกต์ใช้กับงานอื่นได้

1.6.3 ได้พัฒนากระบวนการคิดวิเคราะห์อัลกอริทึม (Algorithm) ที่เหมาะสมกับปัญหา

1.7 งบประมาณของโครงงาน

1.7.1 ค่าถ่ายเอกสารและค่าเข้าเล่มโครงงาน	1000.00	บาท
1.7.2 ค่าหมึกพิมพ์	800.00	บาท
1.7.3 ค่าวัสดุอุปกรณ์ในการทำโครงงาน	200.00	บาท
รวมเป็นเงินทั้งสิ้น	2000.00	บาท

มหาวิทยาลัยนเรศวร

บทที่2
หลักการและทฤษฎี

2.1 พื้นฐานเกี่ยวกับการประมวลผลภาพ

2.1.1 ชนิดของรูปภาพ โดยทั่วไปแบ่งออกเป็น 2 ชนิดคือ ภาพแบบบิตแมปและภาพแบบเวกเตอร์

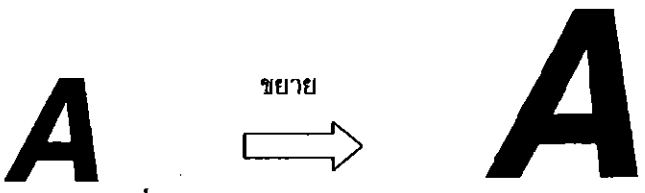
2.1.1.1 ภาพบิตแมป (Bitmap)

เป็นภาพที่ประกอบด้วยจุดสีต่างๆที่มีความเข้มและสีแตกต่างกัน มีจำนวนคงที่ตายตัว มีลักษณะเป็นสี่เหลี่ยมเล็กๆเป็นช่องๆเหมือนตารางรวมๆกันทำให้เกิดภาพ มีข้อดี คือเหมาะกับภาพที่ต้องการระบายสี สร้างสี กำหนดสีที่ต้องการความละเอียดได้ง่าย ข้อเสีย คือ การขยายขนาดของภาพจะทำให้จุดโตขึ้นและเห็นเป็นรอยหยักๆ ถ้าเป็นเส้นโค้งก็จะยิ่งเห็นได้ชัด ไฟล์ภาพพวกนี้มีหลายรูปแบบ อาทิ เช่น BMP, TIF, JPG, PCT ฯลฯ โปรแกรมที่ใช้ เช่น Photoshop, Paint ฯลฯ [1]

รูปที่ 2.1 แสดงการขยายของภาพบิตแมป

2.1.1.2 ภาพเวกเตอร์ (Vector)

เป็นภาพที่สร้างด้วยส่วนประกอบของเส้นลักษณะต่างๆ เป็นภาพที่เกิดจากการกำหนดพิกัดและการคำนวณค่าบนระนาบสองมิติ รวมทั้งมุมและระยะทาง ตามทฤษฎีเวกเตอร์ในทางคณิตศาสตร์ ในการก่อให้เกิดเป็นเส้น หรือรูปภาพ ข้อดีคือ สามารถย่อขยายได้ โดยคุณภาพไม่เปลี่ยนแปลง ข้อเสียคือภาพไม่เหมือนภาพจริงเป็นได้เพียงภาพวาด หรือใกล้เคียงภาพถ่ายเท่านั้น ข้อมูลภาพพวกนี้ได้แก่ไฟล์สกุล svg, ps, eps, ai โปรแกรมที่ใช้สร้างข้อมูลภาพประเภทนี้เช่น Adobe Illustrator, InkScape



รูปที่ 2.2 แสดงการขยายของภาพเวกเตอร์

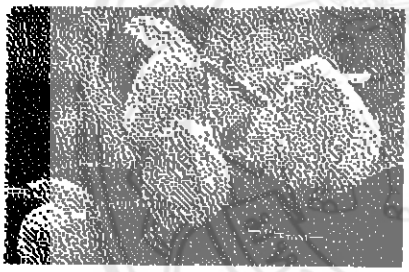
2.1.2ความละเอียดของภาพ

ความละเอียดเป็นสิ่งที่บอกถึงคุณภาพของภาพ หน่วยที่นิยมใช้บอกถึงความละเอียดของภาพคือ พิกเซลต่อนิ้ว(Pixel/Inch) ค่านี้บอกให้เราทราบว่าภาพมีจำนวนพิกเซลกี่พิกเซลในหนึ่งนิ้ว และยังทำให้คำนวณจำนวนจุดทั้งหมดของภาพได้อีกด้วย ตัวอย่างเช่นภาพขนาด 1x1นิ้ว มีความละเอียดเท่ากับ 5 พิกเซลต่อนิ้ว ภาพนี้จะมีพิกเซลทั้งหมดเท่ากับ 25 พิกเซล การเลือกใช้ภาพที่มีความละเอียดสูงๆนั้นเป็นสิ่งที่ดีแต่การใช้ความละเอียดมากกว่าอุปกรณ์แสดงผล เราจะไม่สามารถใช้ประโยชน์จากความละเอียดที่เพิ่มมากขึ้นมานั้น อีกทั้งยังทำให้คอมพิวเตอร์ทำงานช้าลงอีกด้วย การเลือกความละเอียดที่ถูกต้องคือ เลือกความละเอียดตามอุปกรณ์แสดงผลที่คุณใช้งาน [2]

2.1.3 โหมดสีต่างๆ

โหมดสีคือ รูปแบบการผสมสีด้วยเทคนิควิธีการต่างๆ ให้ได้มาซึ่งสีที่แสดงออกบนจอภาพหรือเครื่องพิมพ์ โดยโหมดสีที่นิยมใช้กันมี [3]

2.1.3.1 โหมดสีแบบบิตแมป(Bitmap)



โหมดสีที่มีการเก็บข้อมูลสีเดียว 1 บิตต่อ1 พิกเซลซึ่งจะทำให้รูปโหมดนี้มีเพียงสีขาวหรือดำและไม่สามารถไล่เฉดสีได้ทำให้ภาพหยาบมากและไม่สามารถตกแต่งใดๆได้ข้อดีของโหมดภาพแบบนี้คือ ภาพที่ได้จากโหมดนี้จะมีขนาดที่เล็กมาก สามารถใช้สร้างภาพลายเส้นหรือโลโก้ที่ไม่ต้องการสีสันได้

รูปที่ 2.3แสดงภาพโหมดสีแบบบิตแมป

2.1.3.2 โหมดสีระดับเทา(Grayscale)

โหมดสำหรับภาพขาวดำ สามารถไล่เฉดสีได้ทำให้ภาพมีความคมชัดกว่าบิตแมปมาก

รูปที่ 2.4แสดงภาพโหมดสีระดับเทา (Grayscale)

2.1.3.3 โหมดสีแบบ2 โทนสี(Duotone)

โหมดสีที่สามารถปรับความคมชัดและเฉดสีของภาพระดับเทาให้มีความหลากหลายมากขึ้น สามารถใช้สีอื่น ๆ เข้ามาเสริมในสีดำ ทำให้ภาพมีความน่าสนใจขึ้น

รูปที่ 2.5แสดงภาพโหมดสีแบบ 2 โทนสี (Duotone)

2.1.3.4 โหมดสีแบบอาร์จีบี (RGB Color)

โหมดสีที่ถอดคุณสมบัติของภาพแบบอาร์จีบี มาสร้างเห็นโหมดภาพโดยมีสี แดง เขียวและน้ำเงิน โดยแต่ละสีจะไล่ได้256 ระดับโดยใช้หลักการการรวมแสงสีซึ่งสามารถสร้างสีได้สูงสุด 16.7 ล้านสีหลักการแสดงสีของ จอคอมพิวเตอร์นั้นจะแสดงเป็นอาร์จีบีอยู่แล้วฉะนั้นไม่ว่าจะเลือกโหมดการทำงานใดก็ตาม การแสงดผลบนจอภาพก็จะใช้เป็น RGB อยู่เช่นเดิม

รูปที่ 2.6แสดงภาพโหมดสีแบบอาร์จีบี

2.1.3.5 โหมดสีแบบแลบคัลเลอร์ (Lab Color)

เป็นโหมดสีที่มีความเหมือนจริงที่สุด สำหรับงานพิมพ์ที่เป็น Photo CD หรือภาพที่ต้องการใช้งานระหว่างระบบคอมพิวเตอร์ที่แตกต่างกัน

รูปที่ 2.7แสดงภาพโหมดสีแบบแลบคัลเลอร์

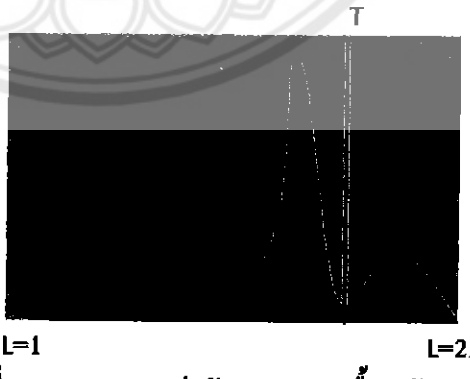
2.1.3.6 โหมดสีแบบ8 บิตต่อ 1 พิกเซล (Multichannel)

การเก็บสีจำนวน 8 บิตต่อ 1 พิกเซลทำให้มีความจำกัดของจำนวนสี ซึ่งใช้กับการพิมพ์แบบพิเศษที่ไม่ต้องการความละเอียด และสีที่ฉูดฉาดมากนัก

รูปที่ 2.8แสดงภาพโหมดสีแบบ8บิตต่อ 1 พิกเซล

2.2 การแยกวัตถุออกจากภาพโดยการเทรชโฮลด์(Threshold)

เทคนิคการกระทำแบบเทรชโฮลด์เป็นเทคนิคการประมวลผลภาพอย่างง่ายๆ เพื่อที่จะแบ่งแยกส่วนพื้นหน้า (Foreground) หรือวัตถุออกจากพื้นหลัง (Background) โดยใช้ค่าระดับความเข้มคงที่ค่าหนึ่งเป็นตัวกำหนดในการแยกแยะ เพื่อให้ภาพผลลัพธ์ที่ได้เป็นภาพ (Binary) ที่มีค่าระดับความเข้มเทา เพียง 2 ระดับเท่านั้นคือขาวและดำ ซึ่งการกำหนดค่าเทรชโฮลด์นั้น ถ้าค่าเทรชโฮลด์ที่เรากำหนดไว้มีค่าไม่เหมาะสม เช่น ค่าเทรชโฮลด์ที่มากเกินไปอาจทำให้รายระเอียดบางส่วนของภาพขาดหายไป หรือค่าเทรชโฮลด์ที่น้อยเกินไปก็อาจมีบางสิ่งที่ไม่พึงประสงค์ปะปนมาด้วย เช่น สัญญาณรบกวน (Noise) ดังนั้นเราจะต้องมีการกำหนดค่าเทรชโฮลด์ที่เหมาะสม ซึ่งในปัจจุบันมีผู้เสนอวิธีในการหาค่าเทรชโฮลด์หลายวิธี ซึ่งแต่ละวิธีจะถูกนำไปใช้ในงานที่มีลักษณะที่แตกต่างกันออกไป



รูปที่ 2.9แสดงการแบ่งวัตถุออกจากพื้นหลัง

เมื่อได้ค่า T ที่ต้องการก็นำไปหาค่าของแต่ละพิกเซลได้ตามสมการ 2.1

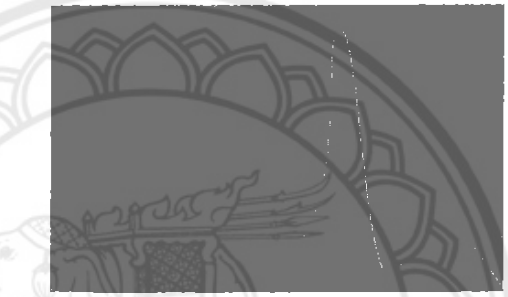
$$f(i,j) = \begin{cases} 1 & if \quad f(i,j) < T \\ 0 & if \quad f(i,j) \geq T \end{cases} \quad (2.1)$$

2.2.1 การเทรชโฮลด์(Threshold) โดย Otsu's Thresholding Method

ที่เลือกอธิบายเทคนิคการเทรชโฮลด์แบบOtsu เพราะในโครงงานนี้ใช้การเทรชโฮลด์แบบ Otsu's Thresholdingเป็นหลัก อัลกอลิทึมOtsu เป็นอัลกอลิทึมที่ฉลาดมากในการวิเคราะห์แผนภาพฮิสโตแกรมและตัดสินใจเลือกค่าเทรชโฮลด์การเลือกค่าเทรชโฮลด์ของ Otsu จะเลือกค่าเทรชโฮลด์โดยคำนวณค่าความแปรปรวนจากฮิสโตแกรม[4]

การหาค่าความแปรปรวนระหว่างกลุ่ม โดยใช้ k เป็นค่าในการแบ่งกลุ่ม เริ่มจาก k =1 ถึง k=255

$$\sigma_B^2 = \omega_0\omega_1(\mu_0 - \mu_1)^2 \quad (2.2)$$



$$P_i = n_i / N,\ P_i \geq 0,\ \sum_{i=1}^{L} P_i = 1$$

P_i คือการ Normalization ค่าให้อยู่ระหว่าง 0 ถึง 1

n_i คือจำนวนพิกเซล ณ ระดับที่ i

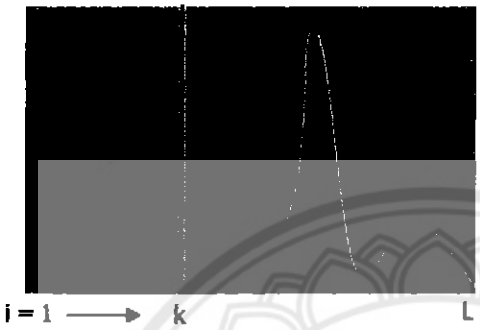
N คือผลรวมของจำนวนพิกเซลในแต่ละระดับ $N = n_1 + n_2 + ... + n_L$

L คือพิกเซลของภาพสีเทา $\{1,2,...,L\}$

$\omega_0 = \sum_{i=1}^{K} P_i = \omega(k)$ (2.3)

ω_0 คือค่าความน่าจะเป็นของกลุ่มที่ 1

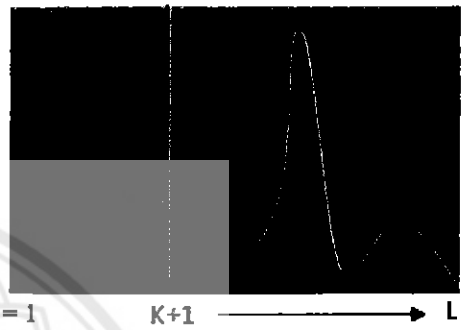
$\sum_{i=1}^{K} P_i$ คือผลรวมของ P_i ตั้งแต่ $i = 1$ ถึง k



$\omega_1 = \sum_{i=k+1}^{L} P_i$ (2.4)

ω_1 คือค่าความน่าจะเป็นของกลุ่มที่ 2

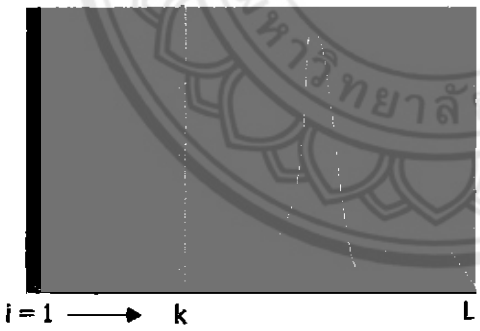
$\sum_{i=k+1}^{L} P_i$ คือผลรวมของ P_i ตั้งแต่ $i = k+1$ ถึง L



$\mu_0 = \mu(k)/\omega_0$ (2.5)

μ_0 คือค่าเฉลี่ยของกลุ่มที่ 1

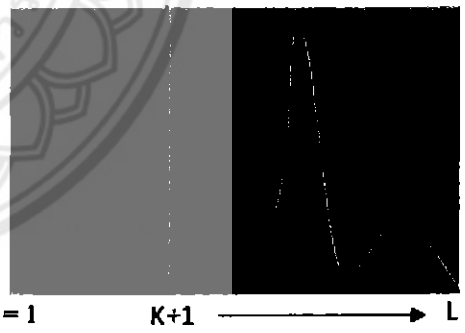
$\mu_k = \sum_{i=1}^{K} ip_i$ คือผลรวมของ $i * P_i$ ตั้งแต่ $i = 1$ ถึง k



$\mu_1 = \sum_{i=K+1}^{L} ip_i / \omega_1$ (2.6)

μ_1 คือค่าเฉลี่ยของกลุ่มที่ 2

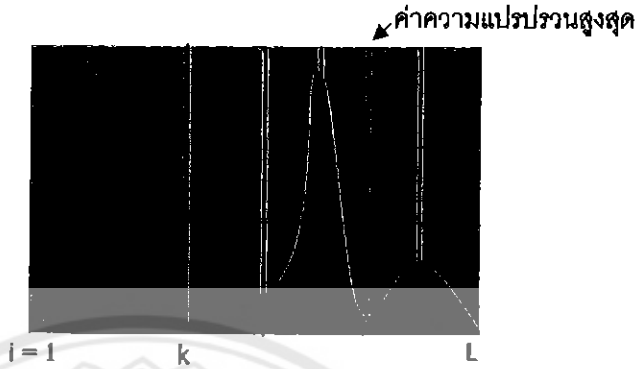
$\sum_{i=K+1}^{L} ip_i / \omega_1$ คือผลรวมของ $(i * p_i)/\omega_1$ ตั้งแต่ $i = k+1$ ถึง L



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

σ_B^2 คือค่าความแปรปรวนของกลุ่มตั้งแต่รอบที่ 1 ถึง 255

โดย $k = 1-255$

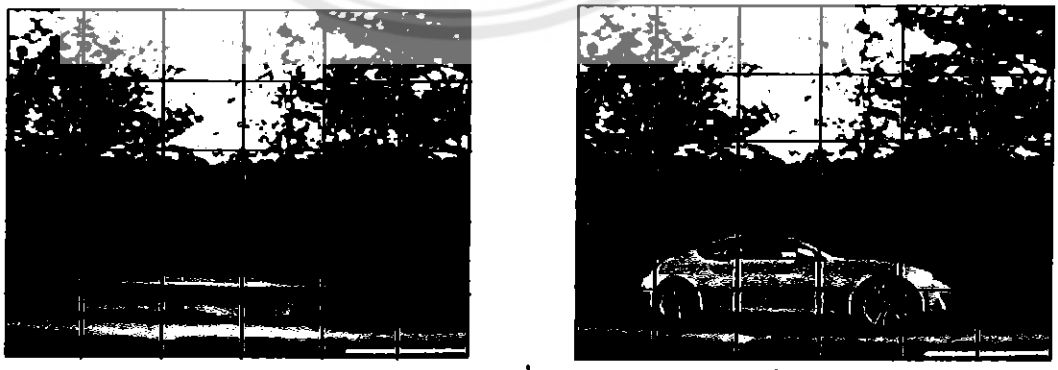


ค่าความแปรปรวนสูงสุดของกลุ่ม ตั้งแต่รอบที่ $k = 1-255$ คือค่าเทรชโฮลด์ที่ได้มา

2.3 การแบ่งภาพออกเป็นส่วนๆ(block)

ภาพหนึ่งภาพเราสามารถตัดออกมาเป็นส่วนๆได้ ซึ่งในที่นี้เราจะเรียกส่วนที่ตัดออกมาว่าบล็อก(Block) [5]เราแบ่งภาพออกเป็นส่วนๆ เพื่อที่จะได้นำแต่ละส่วนไปวิเคราะห์วิจัยเปรียบเทียบได้ง่ายขึ้น เพราะว่าแต่ละส่วนของภาพจะให้ข้อมูลที่ตรงจุดมากกว่าภาพทั้งภาพ เช่น ตั้งกล้องจับภาพเอาไว้ ดูการเปลี่ยนแปลงบนท้องถนน โดยที่ด้านบนเป็นท้องฟ้า เราไม่ต้องนำมาคิดก็ได้ เราหยิบแค่บล็อกข้างล่างมาวิเคราะห์พอ

รูปข้างล่างนี้ก็เป็นตัวอย่างหนึ่งที่เราอาศัยการตรวจจับการเปลี่ยนแปลงของแต่ละบล็อก

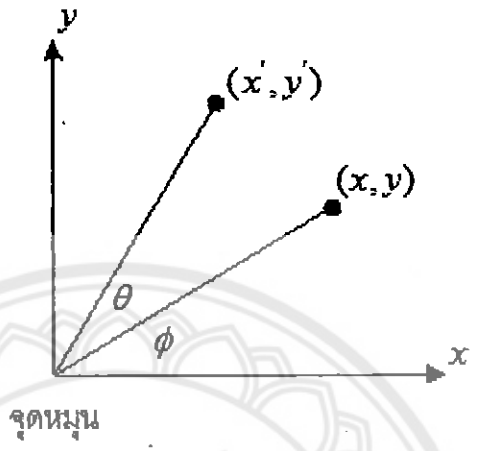


รูป 2.10แสดงการเปลี่ยนแปลงของแต่ละบล็อก

หลักของการแบ่งบล็อกนั้น ก็ขึ้นอยู่กับความต้องการของผู้เขียนโปรแกรมว่าต้องการกี่บล็อก ขนาดเล็ก ใหญ่เท่าไหร่ จากตัวอย่างในรูป 2.10 แบ่งภาพออกเป็น 5*6 = 30 บล็อก

2.4 การหมุนภาพเอกสาร (Rotation)

การหมุนภาพก็คือ หมุนตำแหน่งของภาพในระนาบ x, y จากจุด (x, y) ไปยังจุด (x', y') [6]



รูป 2.11 แสดงการหมุนภาพรอบจุดหมุน (0,0)

จากรูปที่ 2.11 จะได้ว่า

$$x = r\cos(\phi)$$
$$y = r\sin(\phi) \tag{2.8}$$

และ

$$x' = r\cos(\phi+\theta) = r(\cos\phi\cos\theta + \sin\phi\sin\theta)$$
$$y' = r\sin(\phi+\theta) = r(\sin\phi\cos\theta - \cos\phi\sin\theta) \tag{2.9}$$

เพราะฉะนั้นจากสมการที่ (2.7) และ (2.8) จะได้สมการของการหมุนรอบจุดหมุนดังนี้คือ

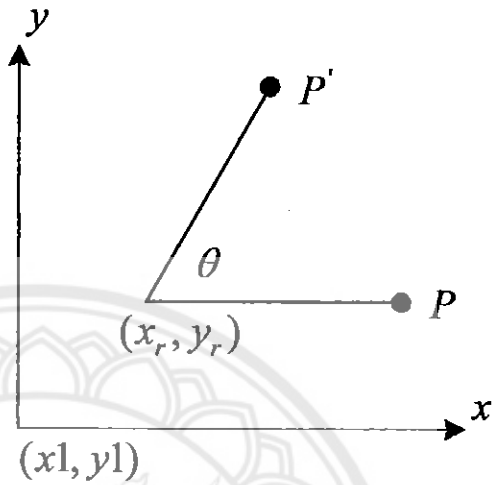
$$x' = x\cos(\theta) + y\sin(\theta)$$
$$y' = y\cos(\theta) - x\sin(\theta) \tag{2.10}$$

ซึ่งสามารถเขียนให้อยู่ในรูปแบบของเมตริกได้ มีลักษณะดังนี้คือ $P' = R.P$ เมื่อ

$$P' = \begin{bmatrix} x' \\ y' \end{bmatrix} \qquad P = \begin{bmatrix} x \\ y \end{bmatrix} \text{ และ } R = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{2.11}$$

2.4.1 การหมุนภาพเมื่อจุดหมุนไม่ได้อยู่ที่จุดกำเนิด (Origin)

พิจารณาเมื่อจุดหมุนไม่ได้อยู่ในตำแหน่ง (0,0) ย้ายไปอยู่ที่ตำแหน่ง (x_r, y_r)



รูป 2.12 แสดงการหมุนภาพเมื่อจุดหมุนไม่ได้อยู่ที่ (0,0)

สมการการหมุนรอบจุดหมุนใด ๆ ที่ไม่ใช้จุด Origin มีลักษณะดังนี้คือ

$$x' = (x - x_r)\cos\theta - (y - y_r)\sin\theta + x_r$$
$$y' = (x - x_r)\sin\theta + (y - y_r)\cos\theta + y_r \qquad (2.12)$$

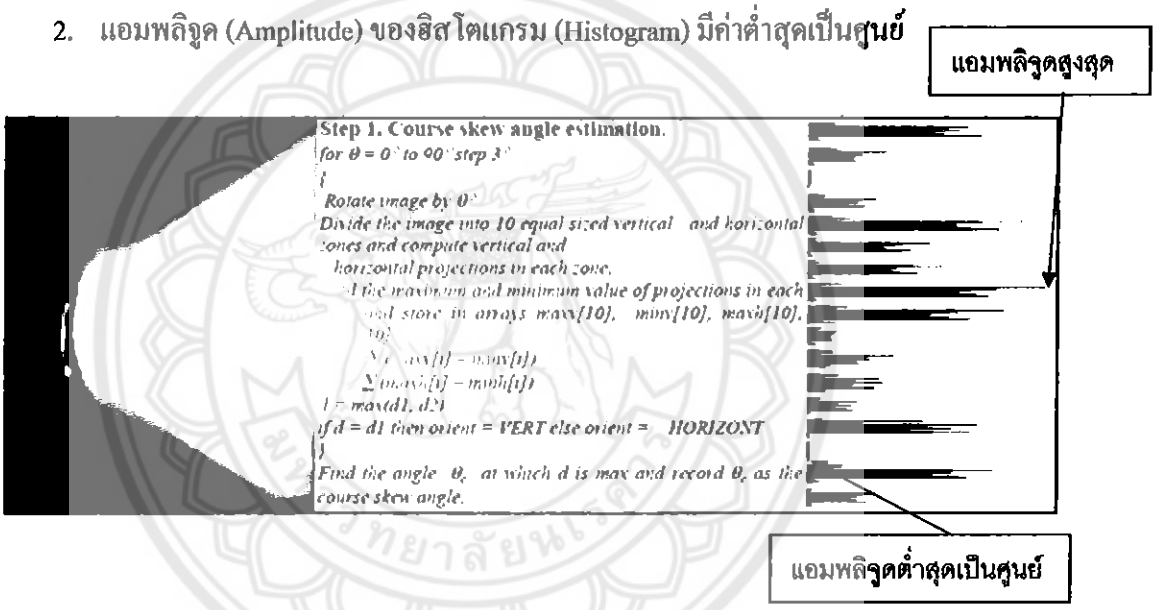
สำหรับการหมุนในการทดลองของโครงงานนี้จะมีจุดหมุนอยู่ที่จุดศูนย์กลางของภาพ

2.5การหาค่าความเอียงของภาพเอกสารโดยการฉายแสงในแนวนอน(Horizontal Projection Profile)

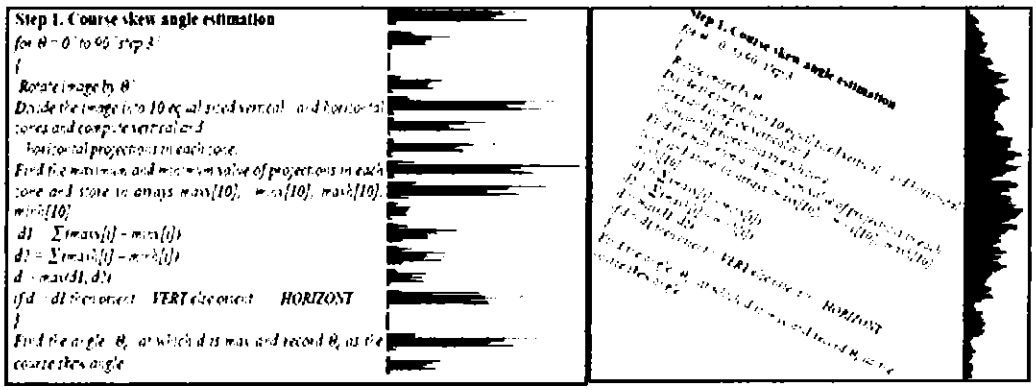
การปรับความเอียงของภาพเอกสารนั้นมีวิธีการต่างๆมากมาย การฉายแสงในแนวนอน (Horizontal Projection Profile) เป็นหนึ่งในวิธีที่ได้รับความนิยมเป็นอย่างมาก แนวคิดของอัลกอลิทึมนี้ก็คือการฉายแสงในแนวนอนแล้วดูการเกิดเงาจากวัตถุ ดังตัวอย่างในรูป 2.2 วัตถุในที่นี้ก็คือตัวอักษร ใช้กระทำกับภาพเอกสารที่เป็นภาพขาวดำ (Binary Image) มีผลลัพธ์ออกมาเป็นแผนภาพฮิสโตแกรม (histogram)ในแนวนอนของแต่ละแถวตามจำนวนพิกเซลที่มีสีดำ [7]

การฉายแสงในแนวนอน(Horizontal Projection Profile) เมื่อภาพเอกสารตรง

1. แอมพลิจูด (Amplitude) ของฮิสโตแกรม (Histogram) มีค่าสูงสุด
2. แอมพลิจูด (Amplitude) ของฮิสโตแกรม (Histogram) มีค่าต่ำสุดเป็นศูนย์



รูป 2.13 การฉายแสงในแนวนอน(Horizontal Projection Profile)



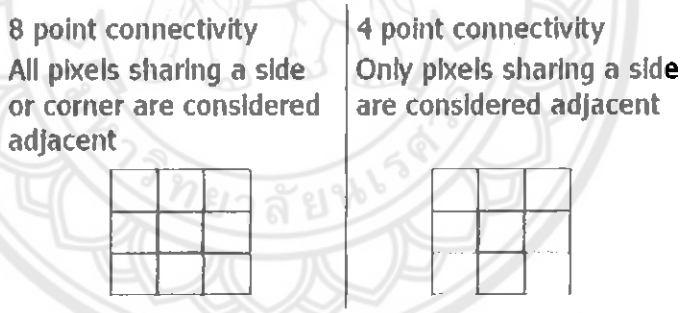
รูป 2.14 แสดงการฉายแสงในแนวนอนของภาพเอกสารตรงและภาพเอกสารเอียง

2.6การหยดสี (Blob Coloring)

การหยดสี (Blob Coloring)คือ กระบวนการในการระบุพื้นที่แต่ละพื้นที่ที่เชื่อมต่อกันใช้แนวคิดของการหยดสีลงบนพื้นที่ สีที่หยดลงไปก็จะกระจายออกไปพื้นที่รอบๆจนเต็มทั่วทั้งพื้นที่นั้น [8] อย่างเช่นตัวอย่างในรูป

รูป 2.16 แสดงการหยดสีและจำนวนของพื้นวัตถุที่พบ

เมื่อภาพถูกแบ่งออกเป็นหลายๆส่วน การหยดสีเป็นเทคนิคที่ช่วยในการหาบริเวณที่เชื่อมต่อกัน Blob Coloring มีหลายวิธีด้วยกัน เป็นต้นว่า การหยดสีแบบ 4 ทิศทางใช้ในการหาบริเวณที่เป็นเส้น หรือการหยดสีแบบ 8 ทิศทางใช้ในการหาพื้นที่ที่เชื่อมต่อกันตามภาพที่2.17



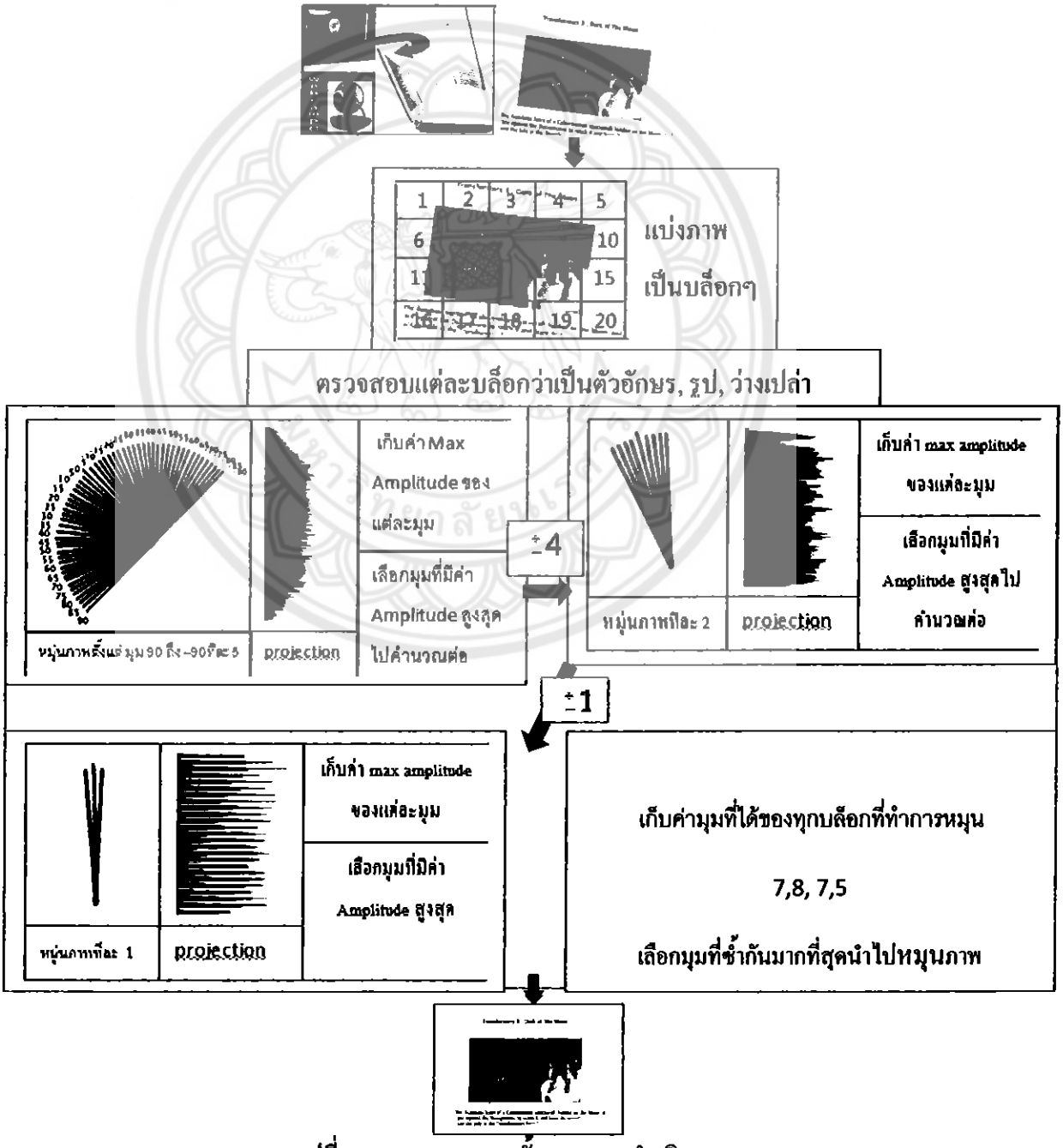
รูปที่ 2.17แสดงลักษณะการเชื่อมต่อกัน แบบ 8 และ 4 ทิศทาง

รูปที่ 2.18แสดงภาพตัวอย่างที่ได้จากการหยดสีแบบ 8 ทิศทาง

จากรูปที่ 2.18 พื้นที่ที่เชื่อมต่อกันจะมีสีเดียวกัน สีที่เทลงบนพื้นที่จะเป็นการสุ่มและจะให้พื้นที่ใกล้ๆกันมีสีไม่ซ้ำกัน

บทที่ 3
วิธีการดำเนินงาน

ขั้นตอนและวิธีการดำเนินงานนั้น มีขั้นตอนหลักๆดังนี้ คือแบ่งภาพออกเป็นส่วนๆ, ตรวจสอบแต่ละบล็อกว่าเป็นบล็อกของตัวหนังสือหรือเปล่า, ทำการหมุนและฉายแสงในแนวนอน เพื่อหามุมที่ภาพเอกสารเอียง และขั้นตอนสุดท้ายก็คือหมุนภาพเอกสารกลับให้ตรง

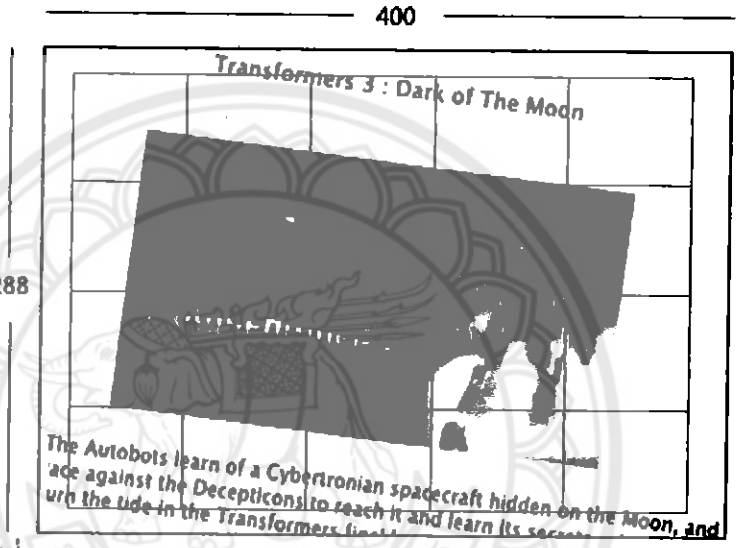


รูปที่ 3.1แสดงแผนภาพขั้นตอนการดำเนินงาน

3.1 แบ่งภาพเอกสารออกเป็นส่วนๆ (Block)

ขั้นตอนนี้เป็นการแบ่งภาพเป็นบล็อกเล็กๆ เพื่อแยกกันประมวลผล ไม่ประมวลผลทีเดียวทั้งภาพ บางบล็อกที่ไม่เข้าเงื่อนไขตามที่เราต้องการ ก็ไม่ต้องทำการประมวลผล ละเลยบล็อกนั้นไป การแบ่งภาพเป็นบล็อกๆจะเป็นการช่วยลดข้อมูลที่ไม่จำเป็นลงได้

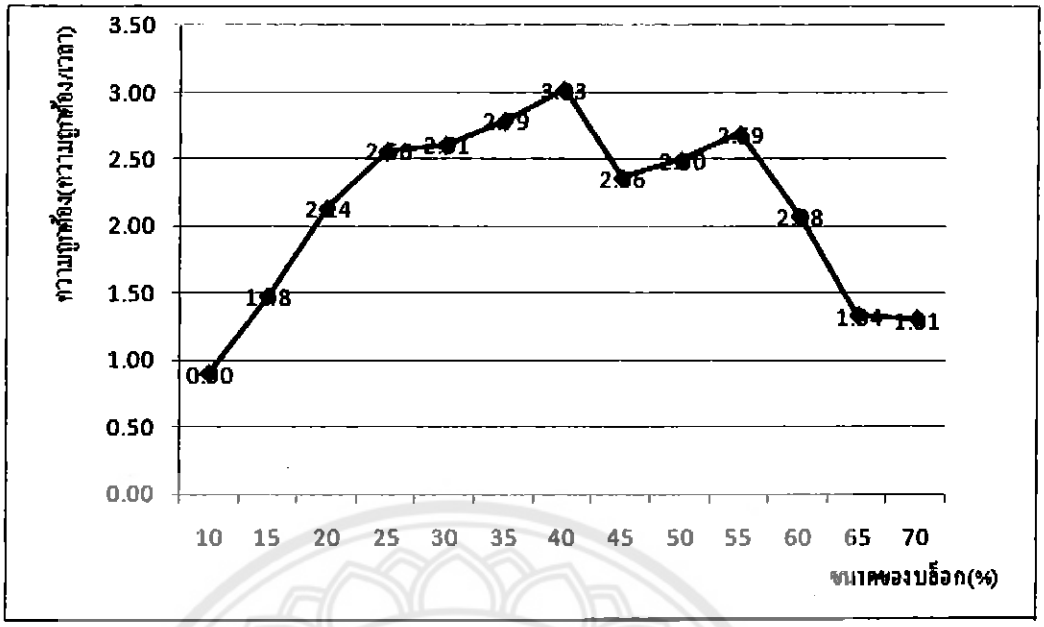
หลักของการแบ่งคือ แต่ละบล็อกต้องเป็นสี่เหลี่ยมจัตุรัสเพราะว่ามันจะมีผลต่อกระบวนการฉายแสงในแนวนอน (Horizontal projection profile) ในขั้นตอนต่อไป



รูปที่ 3.2 แสดงขนาดของรูปภาพและการแบ่งรูปภาพออกเป็นบล็อกๆ

การหาขนาดและจำนวนบล็อก

- ให้ค่าเริ่มต้น(default)ของขนาดบล็อก = x% (ในโปรแกรมสามารถปรับค่านี้ตรงส่วนของการตั้งค่า(Setting)ได้ตามต้องการ)
- ขนาดบล็อก = ค่า default $* \left(\frac{\text{ความสูง}+\text{ความยาว}}{2}\right)$
- นำขนาดบล็อกหารความยาวเพื่อหาจำนวนบล็อกในแนวนอน และนำขนาดบล็อกหารความสูงเพื่อหาจำนวนบล็อกในแนวตั้ง ถ้าหารแล้วเหลือเศษให้ปัดทิ้ง
- ความยาวภาพลบความยาวของบล็อกทั้งหมดในแนวนอนแล้วหาร 2 จะเท่ากับกรอบด้านข้างซ้าย-ขวา
- ความสูงภาพลบความสูงของบล็อกทั้งหมดในแนวตั้งแล้วหาร 2เท่ากับกรอบด้านบน-ล่าง

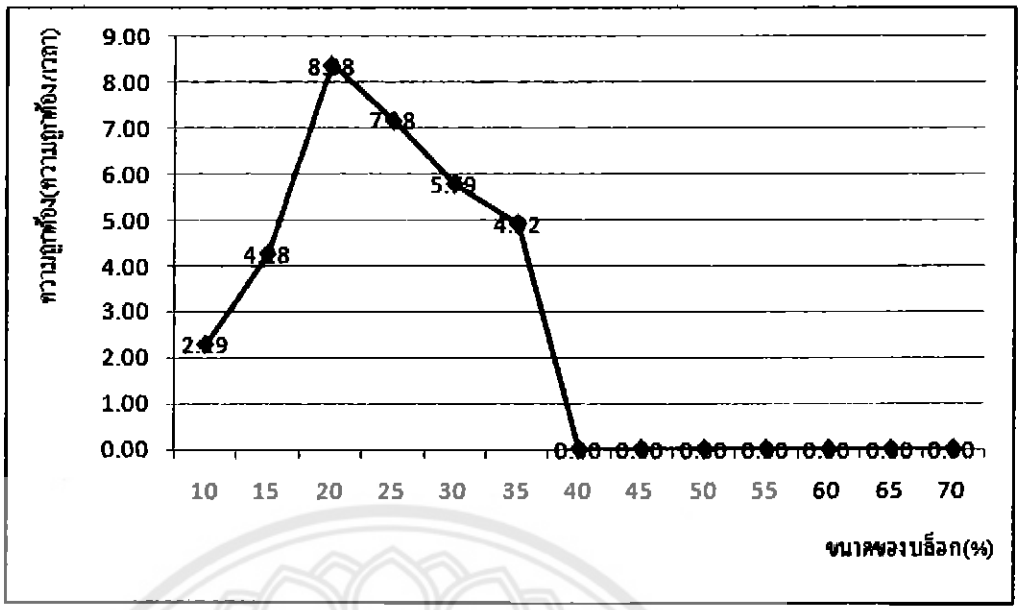


รูปที่ 3.3แสดงอัตราส่วนความถูกต้องต่อขนาดของบล็อกที่เป็นภาพเอกสาร**ตัวหนังสือล้วน**

ตารางที่3.1 แสดงขนาดบล็อกและอัตราส่วนความถูกต้องต่อเวลาของภาพเอกสาร**ตัวหนังสือล้วน**

ขนาดบล็อก(ร้อยละ)	ความถูกต้อง	เวลา	อัตราส่วนความถูกต้องต่อเวลา
10	100.0	111.1	0.90
15	100.0	67.7	1.48
20	100.0	46.8	2.14
25	83.3	32.5	2.56
30	66.7	25.5	2.61
35	66.7	23.9	2.79
40	58.3	19.3	3.03
45	41.7	17.6	2.36
50	33.3	13.4	2.50
55	33.3	12.4	2.69
60	25.0	12.0	2.08
65	16.7	12.5	1.34
70	16.7	12.7	1.31

จากกราฟจะเห็นได้ว่าค่าความถูกต้องต่อเวลามีค่าสูงสุดที่ขนาดของบล็อกมีค่าร้อยละ 40



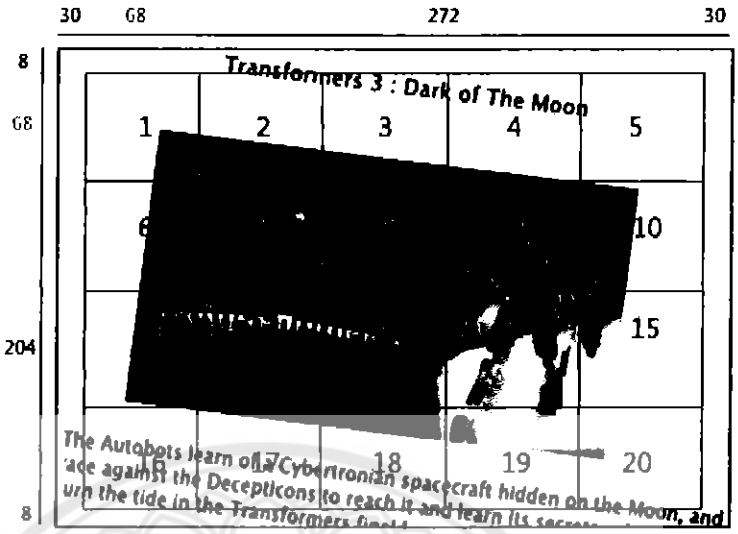
รูปที่ 3.3แสดงอัตราส่วนความถูกต้องต่อขนาดของบล็อกที่เป็นภาพเอกสารที่มีรูปประกอบ

ตารางที่3.2 แสดงขนาดบล็อกและอัตราส่วนความถูกต้องต่อเวลาของภาพเอกสารที่มีรูปประกอบ

ขนาดบล็อก(ร้อยละ)	ค่าความถูกต้อง	เวลา	ค่าความถูกต้อง/เวลา
10	100	43.61395	2.29
15	100	23.37155	4.28
20	100	11.9276	8.38
25	75	10.439275	7.18
30	25	4.31575	5.79
35	25	5.078	4.92
40	0	2.3004	0.00
45	0	1.996075	0.00
50	0	1.956875	0.00
55	0	2.2084	0.00
60	0	2.1022	0.00
65	0	1.84745	0.00
70	0	2.858	0.00

จากกราฟจะเห็นได้ว่าค่าความถูกต้องต่อเวลามีค่าสูงสุดที่ขนาดของบล็อกมีค่าเป็นร้อยละ20

ตัวอย่างการคำนวณค่าต่างๆ



รูปที่ 3.5แสดงตัวอย่างการคำนวณหาจำนวนบล็อกของภาพเอกสาร**ที่มีรูปภาพประกอบ**

Default =20%, ความยาว = 400, ความสูง = 288

$$\text{ขนาดบล็อก} = \frac{20}{100} * \frac{(400+288)}{2} = 68.8$$

จำนวนบล็อกในแนวตั้ง= 288/68 = 4.235 -> 4

จำนวนบล็อกในแนวนอน= 400/68 = 5.88 ->5

$$\text{กรอบซ้าย- ขวา} = \frac{400-(5*68)}{2} = 30$$

$$\text{กรอบบน- ล่าง} = \frac{288-(4*68)}{2} = 8$$

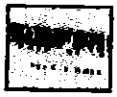
3.2 ตรวจสอบแต่ละบล็อกว่าเป็นตัวอักษรหรือเปล่า

ในการตรวจสอบจะใช้หลักการหยดสี (Blob coloring) เพื่อนับจำนวนการหยดสีและหาขนาดของแต่ละพื้นที่ที่ทำการหยดสีในทุกๆ บล็อก (block) ของภาพเอกสาร เพื่อใช้เป็นข้อมูลในการตัดสินใจว่า บล็อกนั้นเป็นตัวอักษรหรือเปล่า

ตารางที่ 3.3 แสดงการหยดสี จำนวนพื้นที่ของการหยดสี ขนาดของแต่ละพื้นที่ ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) ของแต่ละบล็อก

No.	block	Blob	จำนวน (blob)	ขนาดทุกblob (Pixel)	sd
1			1	580	0
2			11	1210, 52, 1, 37 , 60, 64, 42, 43, 76, 1, 93	351.8000
3			13	39, 28, 63, 107, 73, 1, 1, 10, 8, 103, 68, 43, 64	36.4695
4			7	65,64,71,78,66,160,140	41.1333
5			1	73	0
6			1	4599	0
7			1	12385	0
8			1	12148	0
9			1	10671	0
10			1	4232	0

ตารางที่ 3.3 (ต่อ)

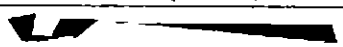
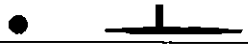
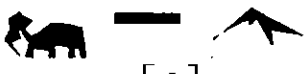
11			1	5918	0
12			7	10553,71,18,20,1,28,5	3.9797e +003
13			3	11229,14,8	6.4767e +003
14			9	2002,65,2,4,3329,4,194,2,27	1.2048e +003
15			1	2502	0
16			38	89,41,31,26,54,58,52,63,77,34,60,69,84,72,71,478,50,40,56,39,44,57,18,3,156,76,44,59,44, 41,54,35,9,6,58,51,40,8	74.5714
17			34	2145,52,30,14,33,62,68,92,27,56,53,54,32,31, 66,45,32,53,45,78,49,40,85,23,36,71, 67,59,128,50,13,34,26,42	360.0824
18			28	3429,3,39,57,56,29,53,33,56,54,57,3,19,48,41,57, 60,40,21,2,33,62,55,55,43,29,20,34	640.8649
19			26	74,4,36,537,117,98,99,98,16,49,2,77,27,12,3, 92,17,3,194,10,80,307,10,58,65,52	154.343
20			9	273,8,66,46,74,56,253,53,27	97.3453

การทดลองการหยดสี (Blob coloring) เพื่อหาค่าเบี่ยงเบนมาตรฐานของพื้นที่ตัวหนังสือ เพื่อใช้เป็นค่าแบ่งตัวหนังสือกับรูปภาพ

ตารางที่ 3.4 การหาค่าเบี่ยงเบนมาตรฐานของพื้นที่ตัวหนังสือเมื่อภาพเอกสารเป็นตัวหนังสือล้วน

ภาพเอกสารขาว-ดำ	ค่าเบี่ยงเบนมาตรฐานของพื้นที่
English ภาษาอังกฤษ . .	75.0855
การสร้างภาพตัดขวางทางการแพทย์	17.1208
$p(t,\theta) = -ln\left[\frac{I}{I_0}\right] = \int_{t,\theta} \mu(x,y)\,ds$	23.7737
1. Parallel-beam Projection	12.8250
Parallel-beam เป็นเรย์โปรเจ็คชั่นแบบง่าย	21.7401
$p_\theta(t) = \int_{-\infty}^{\infty} f(t,s)ds$ สมการ	27.3417
(Point-Source) ดังนั้นข้อมูลเรย์โปรเจ็คชั่น	13.1363
กระจายออกเป็นรูปกรวยดังรูปที่ 6-	13.0130
Equally Spaced Collinear Detectors	24.3541
สมการเรย์โปรเจ็คชั่นแบบ Fan-beam....	15.0762

ตารางที่ 3.5 แสดงการหาค่าเบี่ยงเบนมาตรฐานของตัวหนังสือเมื่อภาพเอกสารมีรูปประกอบ

ภาพเอกสารขาว-ดำ	ค่าเบี่ยงเบนมาตรฐานของพื้นที่
English ภาษาอังกฤษ . .	123.5526
 การสร้างภาพตัดขวางทางการแพทย์ 	97.8999
 $p(t,\theta) = -ln\left[\frac{I}{I_0}\right] = \int_{t,\theta} \mu(x,y)\,ds$	272.3834
1. Parallel-beam Projection	102.9360

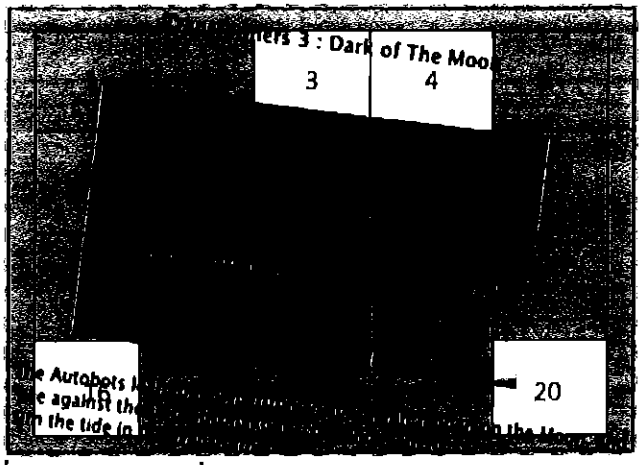
ตารางที่ 3.5 (ต่อ)

Parallel-beam เป็นเรย์โปรเจ็คชั่นแบบง่าย	151.8133
$p_\theta(t) = \int_{-\infty}^{\infty} f(t,s)ds$ สมการเรย์โปรเจ็คชั่น	134.2838
1. Parallel-beam Projection	119.5005
กระจายออกเป็นรูปกรวยดังรูปที่ 6-	186.4432
Equally Spaced Collinear Detectors	114.2912
สมการเรย์โปรเจ็คชั่นแบบ Fan-beam....	147.0073

จากตารางที่ 3.4 จะเห็นได้ว่าหาค่าเบี่ยงเบนมาตรฐานของพื้นที่ตัวหนังสือจะอยู่ที่ประมาณ 20 และจากตารางที่ 3.5 เมื่อมีรูปภาพมาปนกับตัวหนังสือทำให้ค่าเบี่ยงเบนมาตรฐานมีค่าเพิ่มมากขึ้น เนื่องมาจากรูปภาพมีพื้นที่เล็กใหญ่ไม่เท่ากัน บางพื้นที่มีขนาดใหญ่มาก บางพื้นที่มีขนาดเล็ก ซึ่งทำให้ค่าเบี่ยงเบนมาตรฐานมีค่ามาก

จากตารางทั้ง 2 ตาราง จะใช้ค่าเบี่ยงเบนมาตรฐานที่ 150 เป็นตัวแบ่งว่าบล็อกแต่ละบล็อกนั้นเป็นตัวหนังสือหรือเป็นภาพ

เมื่อพิจารณาตารางที่ 3.3 บล็อกที่เป็นตัวอักษรจากการหยดสีคือ บล็อกที่ 3, 4, 16, 20



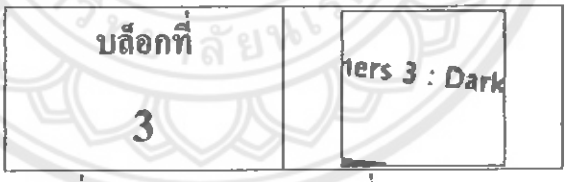
รูปที่ 3.6แสดงบล็อกที่เป็นตัวอักษรจากการหยดสี(Blob coloring)

3.3 หมุนภาพและทำการฉายแสงในแนวนอน(HorizontalProjection Profile)

ในขั้นตอนนี้มี 2 กระบวนการย่อยที่ทำควบคู่กันไป 2 กระบวนการ **คือการหมุนภาพและการทำ**การฉายแสงในแนวนอน โดยเมื่อหมุนภาพเสร็จ 1 ครั้งก็จะทำการฉายแสงในแนวนอน1 ครั้ง และเก็บค่าแอมพลิจูดสูงสุด (Max amplitude)ที่ได้

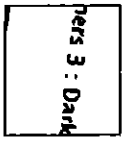
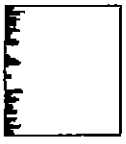
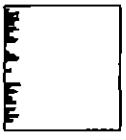
3.3.1 หมุนภาพเอกสารที่ละ 5 องศา และทำการฉายแสงในแนวนอน เก็บค่าแอมพลิจูดสูงสุด

ตัวอย่างการประมวลผลของบล็อกที่ 3

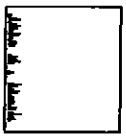
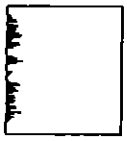


รูปที่ 3.7 แสดงบล็อกตัวอย่างที่ใช้ในการหมุน

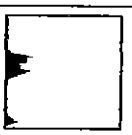
ตารางที่ 3.6แสดงการหมุนภาพเอกสารที่ละ 5 องศา จากมุม -90 ถึง มุม 90 องศา

มุม	Image rotated	projection	Max amplitude
-90			10
-85			9

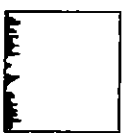
ตารางที่ 3.6 (ต่อ)

-80			9
-75			10
-70			8
-65			8
-60			7
-55			10
-50			8
-45			40
-40			9
-35			11
-30			11

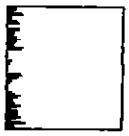
ตารางที่ 3.6 (ต่อ)

-25			10
-20			21
-15			27
-10			31
-5			33
0			37
5			56
10			57
15			36
20			34
25			32
30			26

ตารางที่ 3.6 (ต่อ)

35			22
40			22
45			20
50			18
55			17
60			17
65			17
70			16
75			16
80			16
85			17

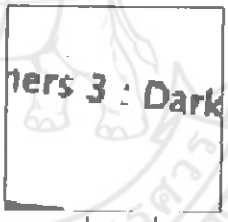
ตารางที่ 3.6 (ต่อ)

90			18

หลังจากที่หมุน -90 ถึง 90 ครบทุกมุมแล้ว ก็จะทำการเลือกมุมที่มีค่าแอมพลิจูดสูงสุดมาทำการหมุนต่อ โดยนำมุมที่ได้มาบวก 4 และลบ 4 จากตัวอย่างในตารางข้างบนมุมที่มีค่าแอมพลิจูดสูงสุดคือ 10

ดังนั้นช่วงของในการหมุนของขั้นตอนถัดไป คือ (10-4)= 6 ถึง (10+4)= 14

3.3.2หมุนภาพเอกสารที่ละ2 องศา ตั้งแต่ 6 องศาถึง 14 องศา และทำ projection เก็บค่า Max amplitude



รูปที่ 3.8แสดงรูปเดิมรูปที่ 3.7 ที่เราต้องการทำการหมุน

ตารางที่ 3.7แสดงการหมุนภาพเอกสารที่ละ 2 องศา จากมุม 6 ถึง มุม 14 องศา

มุม	Image rotated	projection	Max amplitude
6			58
8			62
10			57

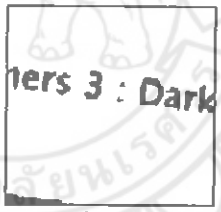
ตารางที่ 3.7 (ต่อ)

12			51
14			39

หลังจากที่หมุน 6 ถึง 14 ครบทุกมุมแล้ว ก็จะทำการเลือกมุมที่มีค่าแอมพลิจูดสูงสุดมาทำการหมุนต่อ โดยนำมุมที่ได้มาบวก 1 และลบ 1 จากตัวอย่างในตารางข้างบนมุมที่มีค่าแอมพลิจูดสูงสุดคือ 8

ดังนั้นช่วงของในการหมุนของขั้นตอนสุดท้าย คือ (8-1)= 7 ถึง (8+1)= 9

3.3.3 หมุนภาพที่ละ1 องศา และ projection เก็บค่า max amplitude



รูปที่ 3.9แสดงรูปเดิมรูปที่ 3.7 ที่เราต้องการทำการหมุน

ตารางที่ 3.8แสดงการหมุนภาพเอกสารที่ละ 1 องศา จากมุม 7 ถึง มุม 9 องศา

มุม	Image rotated	projection	Max amplitude
7			64
8			62
9			53

ค่าแอมพลิจูดสูงสุดที่ได้คือ 64 ซึ่งตรงกับมุม 7 องศา ดังนั้นจึงสรุปได้ว่ามุมที่จะทำให้บล็อกนี้ตรงคือ 7 องศา หลังจากนั้นก็ทำวิธีการเดิมกับทุกบล็อกที่เป็นบล็อกของตัวหนังสือ ผลลัพธ์จากการประมวลผลของทุกๆบล็อก

ตารางที่3.9แสดงค่าแอมพลิจูดสูงสุดของทุกบล็อกที่เป็นตัวอักษร

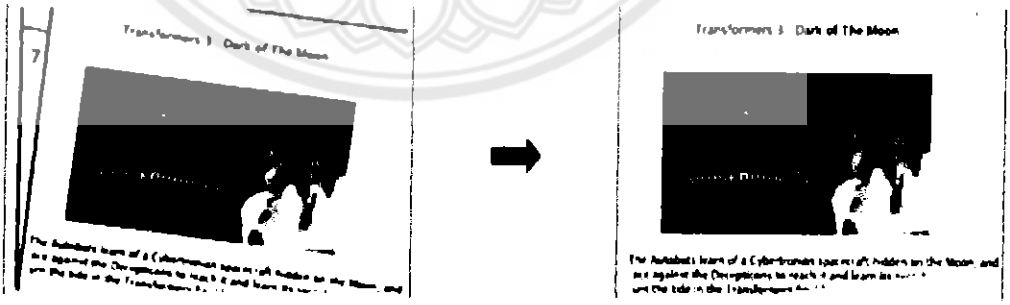
บล็อกที่	ผลลัพธ์
3	7
4	8
16	7
20	5

3.4 เลือกมุมที่จะนำไปใช้หมุนภาพเอกสารให้ตรง

เลือกมุมที่ซ้ำกันมากที่สุดที่จะนำไปใช้หมุนภาพให้ตรง จากตารางที่ มุมที่ซ้ำกันมากที่สุด คือ 7 องศา

3.5 หมุนภาพเอกสารให้ตรง

การหมุนภาพให้ตรงซึ่งเป็นการหมุนครั้งสุดท้าย จะใช้การหมุนแบบ bilinear เพราะว่า bilinear จะสูญเสียรายละเอียดของภาพเอกสารน้อยกว่าการหมุนแบบธรรมดา

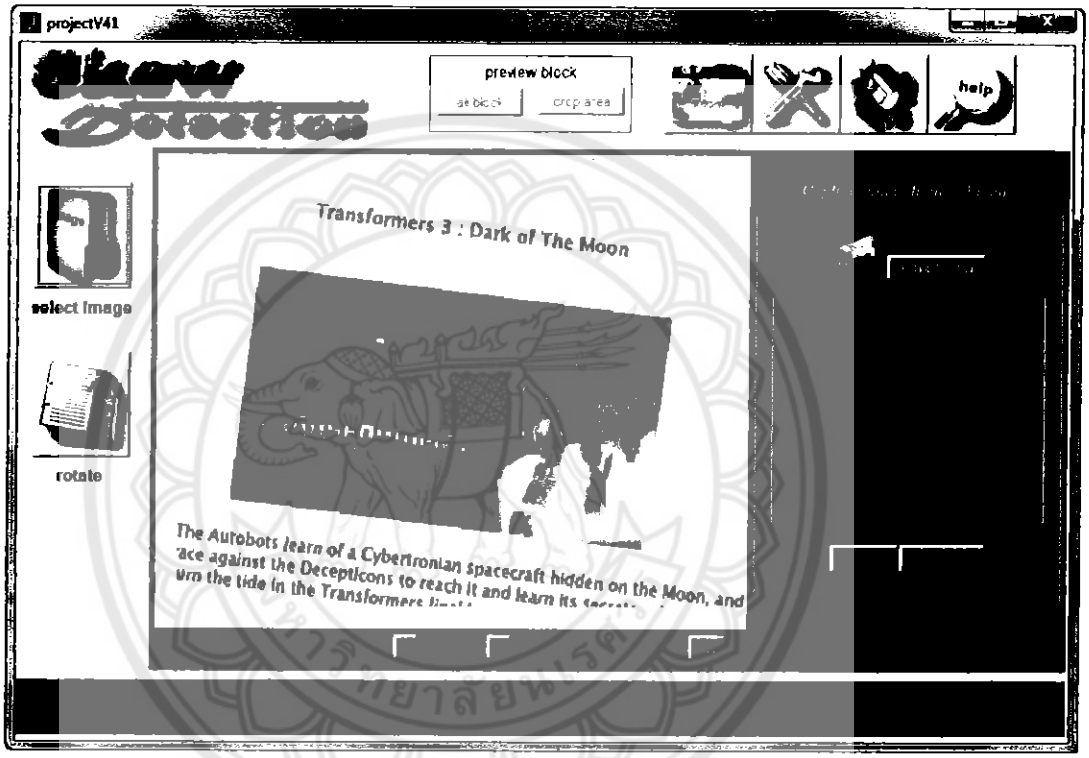


รูปที่ 3.10แสดงภาพเริ่มต้นและภาพผลลัพธ์

บทที่ 4

ผลการดำเนินงาน

โปรแกรมในภาพข้างล่างเป็นโปรแกรมที่เขียนจากแมทแลบ(MATLAB) เป็นผลจากการนำอัลกอริทึมและเทคนิคต่างๆที่กล่าวในข้างต้น มาประยุกต์ใช้เข้าด้วยกัน



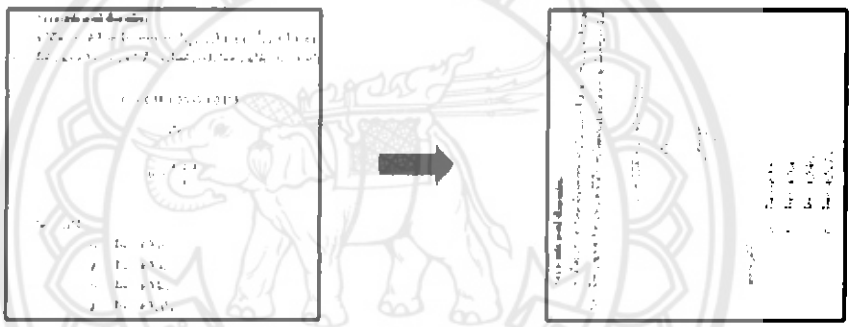
รูปที่ 4.1 แสดงหน้าจอหลักของโปรแกรม

ภาพเอกสารที่ใช้ในการทดสอบมีทั้งเอกสารในคอมพิวเตอร์, อินเตอร์เน็ตและภาพถ่ายจากกล้องเว็บแคม ซึ่งจะมีมุมเอียงแตกต่างกันไป ภาพเอกสารจะถูกทำให้เอียงโดยโปรแกรม MATLAB ตามตัวอย่างในรูปที่ 4.2 เพื่อที่จะได้ทราบว่าภาพนั้นเอียงไปกี่องศา เอาไว้เปรียบเทียบกับมุมที่โปรแกรมวิเคราะห์ได้ ทำให้สามารถวัดความคลาดเคลื่อนออกมาเป็นเปอร์เซ็นต์ได้ ซึ่งจะแน่นอนกว่าการมองด้วยตาเปล่า

skew image = imrotate(straight image,angle,'bilinear')

skew image=ภาพเอกสารที่ถูกทำให้เอียง, straight image=ภาพเอกสารตรง

angle=มุมที่ต้องการให้เอียง, bilinear=หมุนแบบbilinear



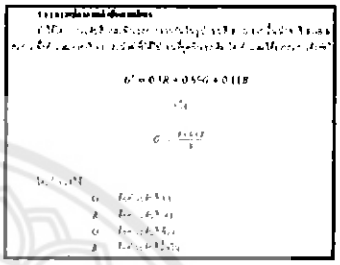
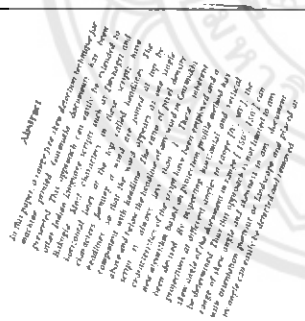
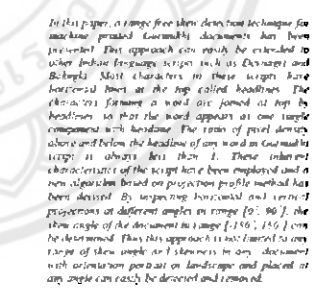
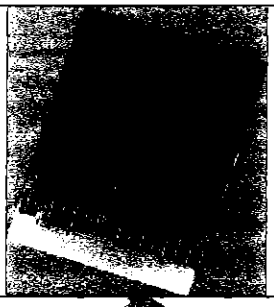
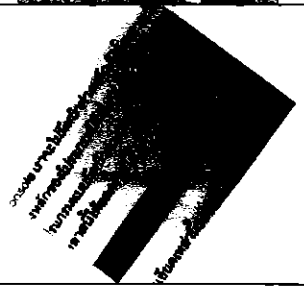
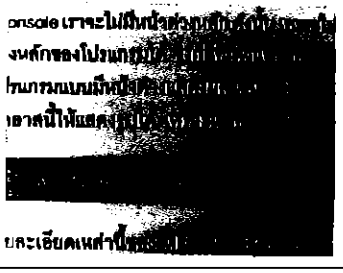
straight image skew image

skew image = imrotate(straight image, 89,'bilinear')

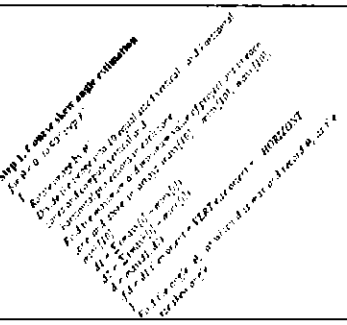
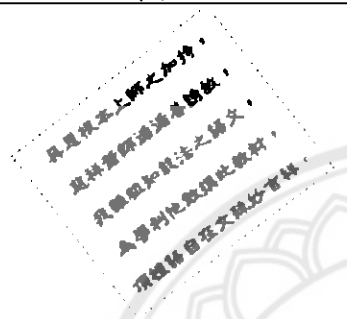
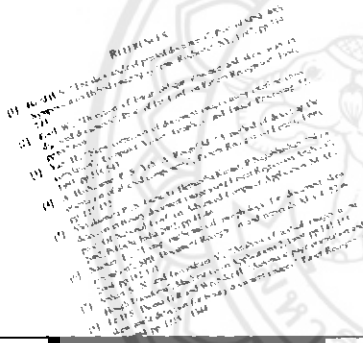
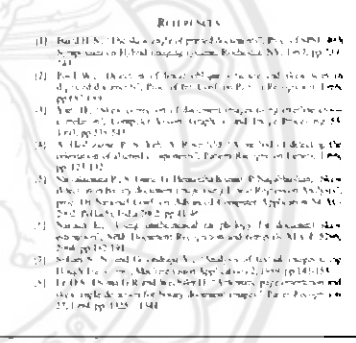
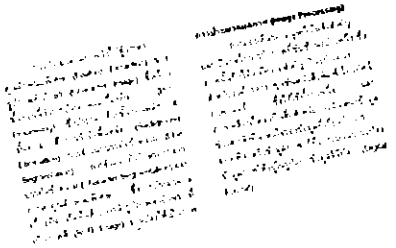
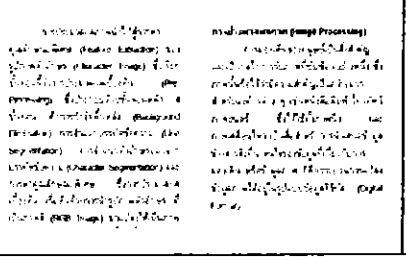
การทดสอบโปรแกรมจะแบ่งภาพเอกสารออกเป็น 2 กลุ่มคือ กลุ่มของภาพเอกสารที่มีตัวหนังสือ, ตาราง แต่จะไม่มีรูปภาพปะปนอยู่ และกลุ่มของภาพเอกสารที่มีรูปภาพปะปนอยู่ด้วย แต่ต้องมีบริเวณที่มีตัวหนังสือประกอบอยู่ด้วยอย่างน้อย 70% และบริเวณที่มีตัวหนังสือนั้นจะต้องเป็นตัวหนังสือสีเข้มและมีพื้นหลังที่มีสีอ่อนกว่า

4.1 ผลการทดสอบโปรแกรมกับกลุ่มของภาพเอกสารที่มีเฉพาะตัวหนังสือ

ตารางที่ 4.1ผลการทดสอบโปรแกรมกับกลุ่มของภาพเอกสารที่มีเฉพาะตัวหนังสือ

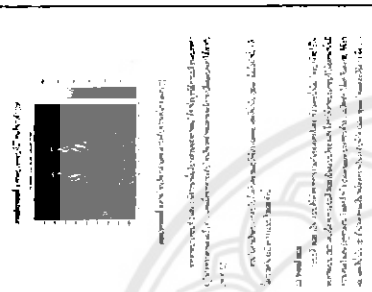
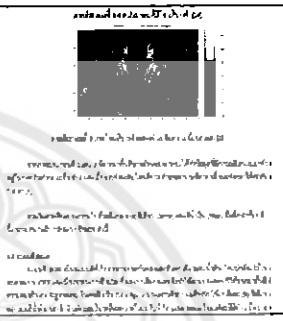
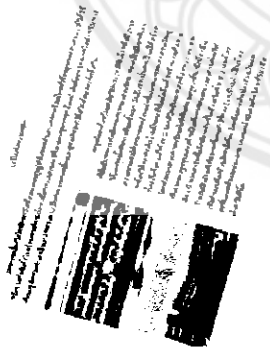
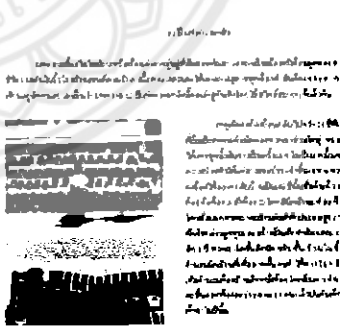
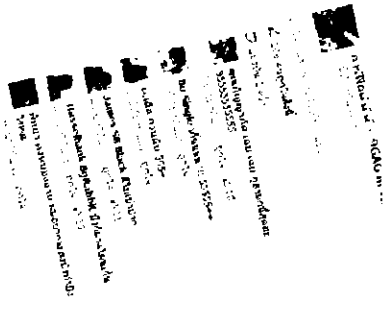
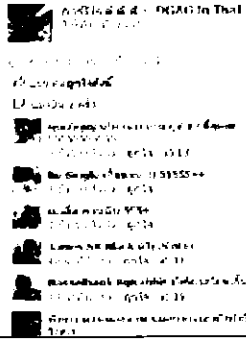
ลำดับ	ก่อนประมวลผล		หลังประมวลผล		
	ภาพ	มุมเอียง (องศา)	ภาพ	มุมที่ detect ได้	คลาดเคลื่อน
1		89		-89	0
2		86		-86	1
3		69		-69	0
4		72		-71	0
5		54		-54	0

ตารางที่ 4.1 (ต่อ)

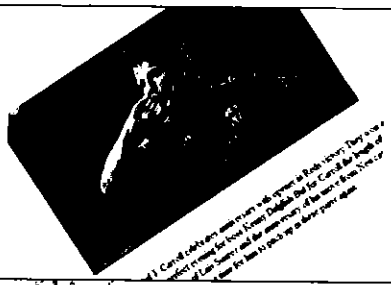
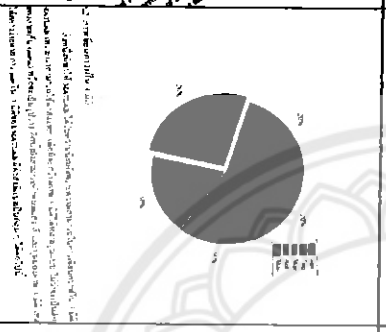
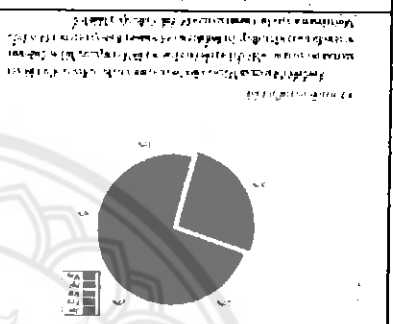
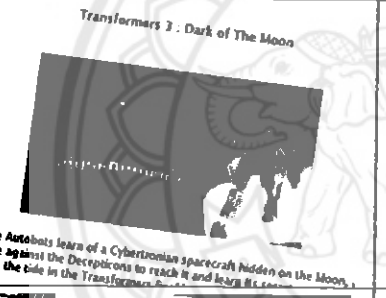
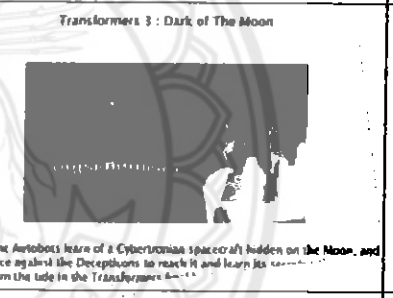
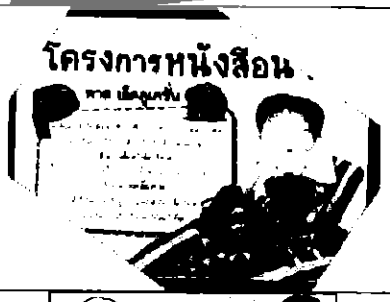
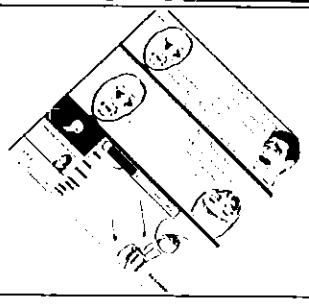
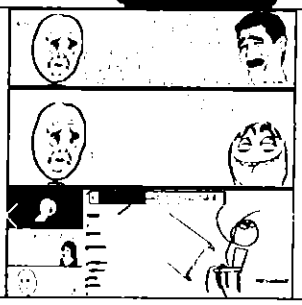
6		46	**Step 1. Course skew angle estimation.** *for θ = 0° to 90° step 3°* *{* *Rotate image by θ°* *Divide the image into 10 equal sized vertical and horizontal zones and compute vertical and horizontal projections in each zone.* *Find the maximum and minimum value of projections in each zone and store in arrays maxv[10], minv[10], maxh[10], minh[10]* *d1 = ∑(maxv[i] – minv[i])* *d2 = ∑(maxh[i] – minh[i])* *d = max(d1, d2)* *if d = d1 then orient = VERT else orient = HORIZONT* *}* *Find the angle θ, at which d is max and record θ, as the course skew angle.*	-46	0
7		38	具慶根本上師之加持， 堪祥喜師諸通善說故， 我雖粗知戲法之語文， 為學利他敢撰此教材， 頂禮諸自在文殊妙吉祥。	-38	0
8		21		-21	0
9		16		-16	0
10		11		-11	0

4.2 ผลการทดสอบโปรแกรมกับกลุ่มของภาพเอกสารที่มีรูปภาพปะปนอยู่ด้วย

ตารางที่ 4.2ผลการทดสอบโปรแกรมกับภาพเอกสารที่มีรูปปนอยู่ด้วย

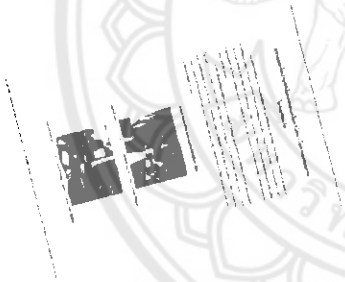
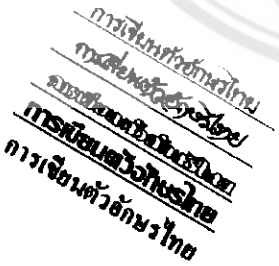
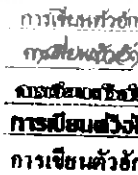
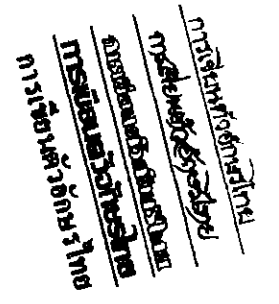
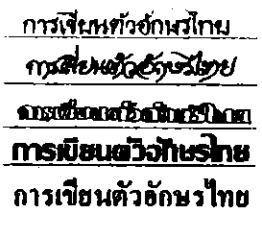
ลำดับ	ก่อนประมวลผล		หลังประมวลผล		
	ภาพ	มุมเอียง (องศา)	ภาพ	มุมที่ detect ได้	คลาดเคลื่อน
1		90		-90	0
2		-88		-92	102
3		74		-74	0
4		-78		78	0

ตารางที่ 4.2 (ต่อ)

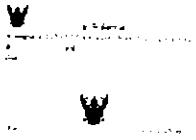
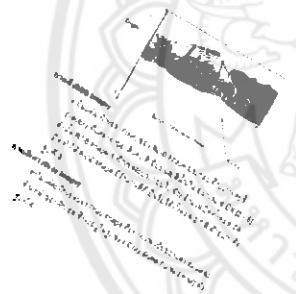
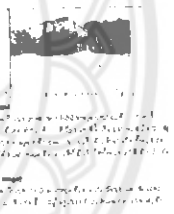
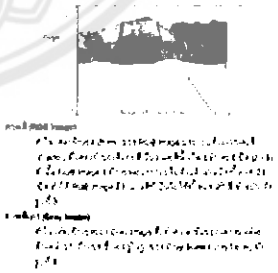
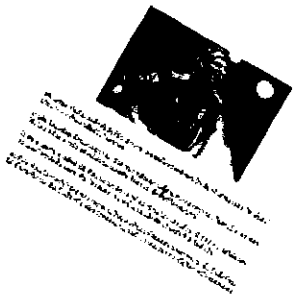
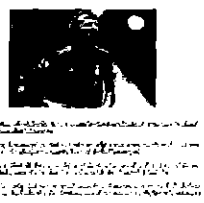
5		34		-34	0
6		-88		-92	102
7		-7		7	0
8		-6		7	1
9		33		-33	0
10		-45		45	0

4.3 การเปรียบเทียบรูปภาพหลายๆ แบบโดยการเลือกมุมสองมุมคือ มุม -30 และ -75 ในการนำมาทดสอบ เพื่อตรวจความถูกต้องของโปรแกรม

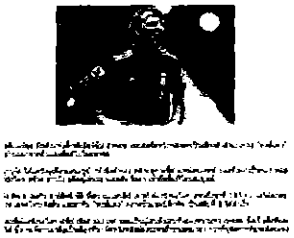
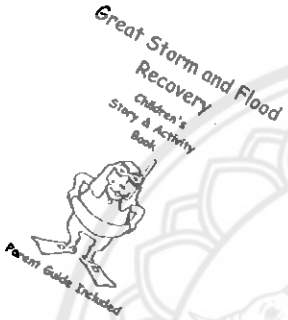
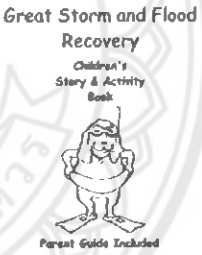
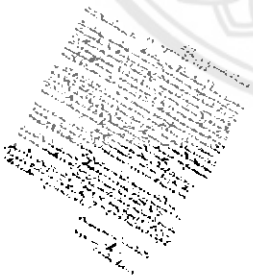
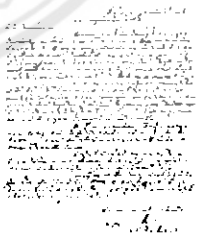
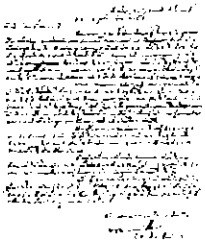
ตารางที่ 4.3แสดงการตรวจสอบความถูกต้องกับภาพเอกสารหลายๆ แบบ

ภาพก่อนการประมวลผล			ภาพหลังการประมวลผล		
No.	ภาพที่มีความเอียง	มุมเอียง (องศา)	ภาพที่ผ่านการปรับแล้ว	มุมที่หมุนกลับ (องศา)	คลาดเคลื่อน
1		-30		30	0
2		-75		75	0
3		-30		29	0
4		-75		75	0

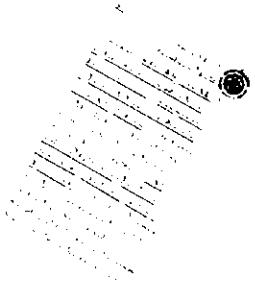
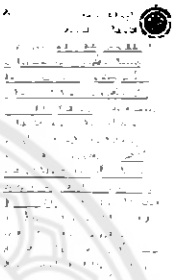
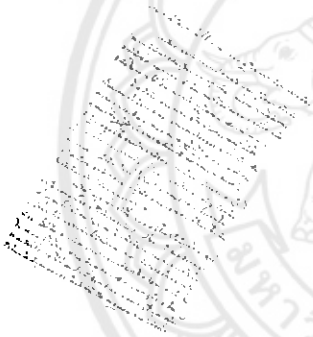
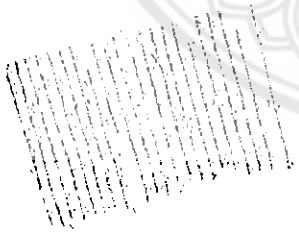
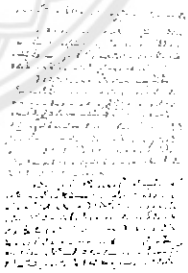
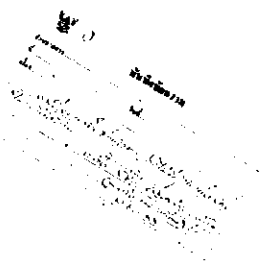
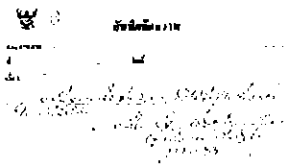
ตารางที่ 4.3 (ต่อ)

5		-30		29	1
6		-75		75	0
7		-30		31	1
8		-75		75	0
9		-30		30	0

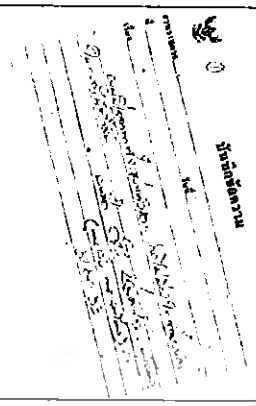
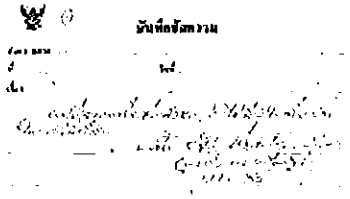
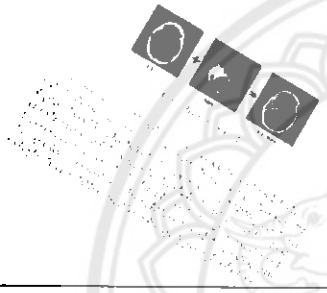
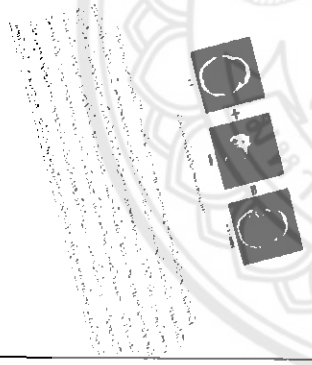
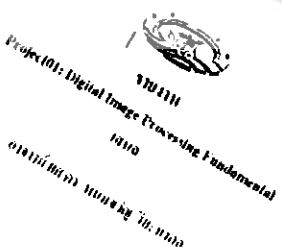
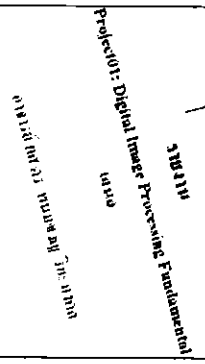
ตารางที่ 4.3 (ต่อ)

10		-75		75	0
11		-30		30	0
12		-75		75	0
13		-30		30	0
14		-75		75	0

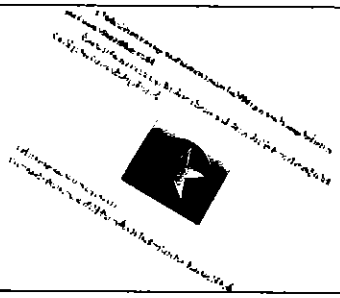
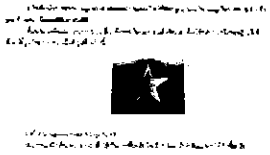
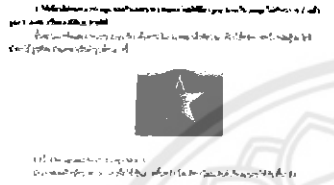
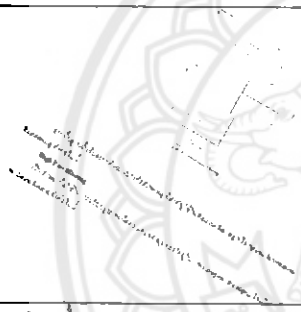
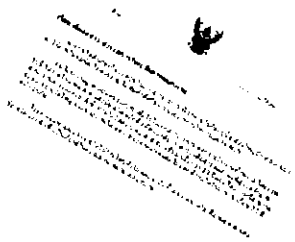
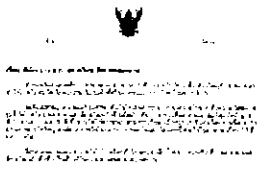
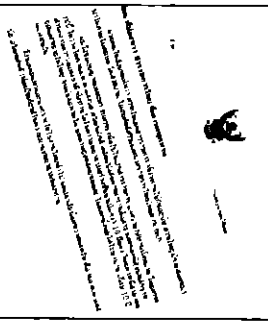
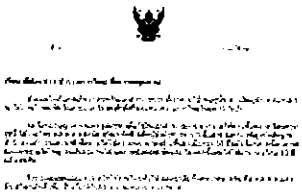
ตารางที่ 4.3 (ต่อ)

15		-30		30	0
16		-75		74	1
17		-30		30	0
18		-75		74	1
19		-30		31	1

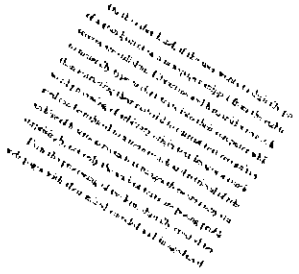
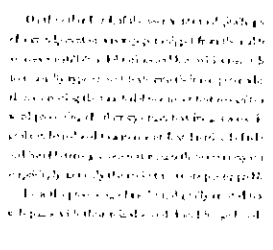
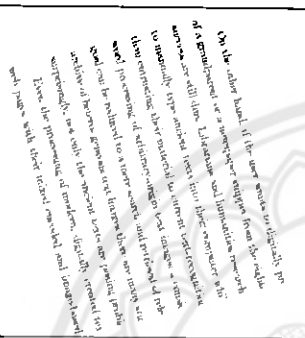
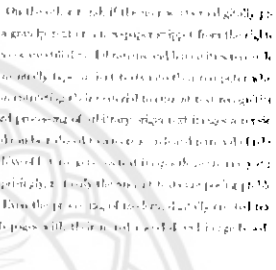
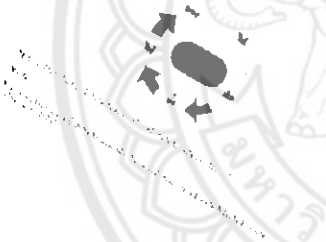
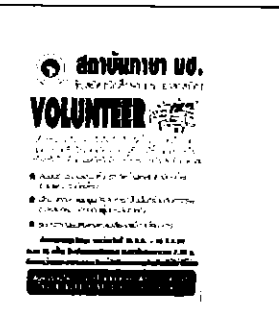
ตารางที่ 4.3 (ต่อ)

20		-75		76	1
21		-30		30	0
22		-75		75	0
23		-30		31	1
24		-75		75	0

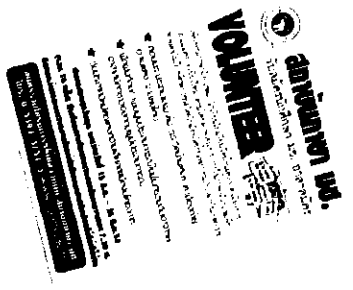
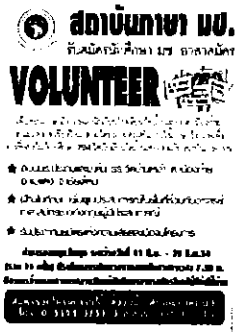
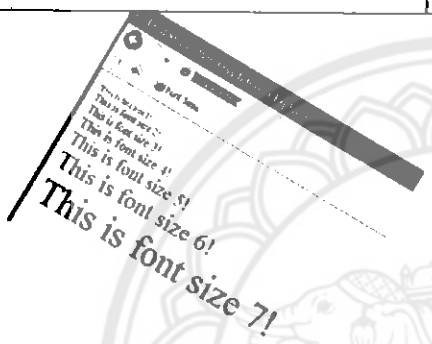
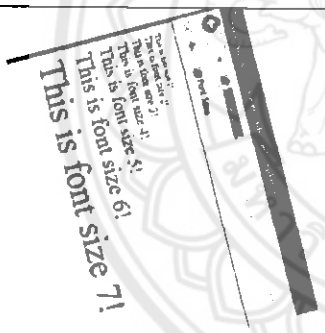
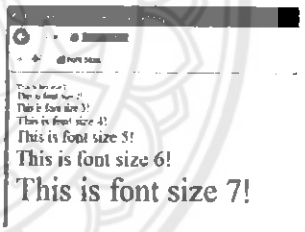
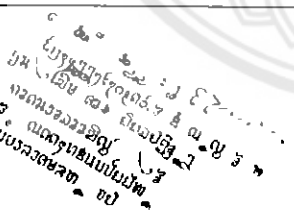
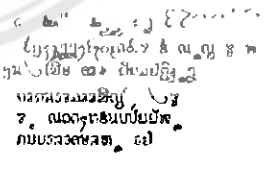
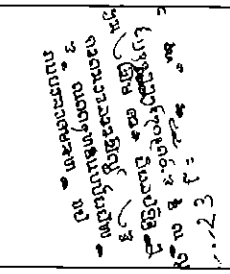
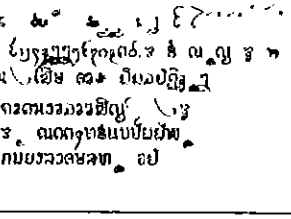
ตารางที่ 4.3 (ต่อ)

25		-30		30	0
26		-75		75	0
27		-30		31	1
28		-75		75	0
29		-30		30	0
30		-75		75	0

ตารางที่ 4.3 (ต่อ)

31		-30		30	0
32		-75		75	0
33		-30		30	0
34		-75		75	0
35		-30		30	0

ตารางที่ 4.3 (ต่อ)

36		-75		75	0
37		-30		30	0
38		-75		75	0
39		-30		30	0
40		-75		75	0

ผลจากการหมุนในตารางด้านบนทั้งหมด 40 ภาพภาพที่ทำการหมุนได้ตรงมี 32 ภาพ และภาพที่ผิดพลาดไป 1 องศา มี 8 ภาพ

บทที่ 5
บทสรุป

5.1 สรุปและวิเคราะห์ผลการทดลอง

โครงงานนี้ได้นำเสนอวิธีการโดยใช้หลักการประมวลผลภาพและคอมพิวเตอร์วิชันมาช่วยวิเคราะห์ความเอียงของเอกสารและปรับให้ตรงโดยอัตโนมัติ โดยใช้อัลกอริทึมการฉายแสงในแนวนอน (Horizontal Projection Profile) เป็นอัลกอริทึมหลัก และใช้อัลกอริทึมการเทสี (Blob Coloring) กับ การแบ่งบล็อกมาช่วยเสริมให้การวิเคราะห์ความเอียงมีประสิทธ์ภาพมากขึ้น ให้สามารถประมวลผลได้กับภาพเอกสารหลายๆแบบ ถ้าใช้เพียงอัลกอริทึมการฉายแสงในแนวนอนจะมีข้อจำกัดหลายอย่าง เช่น ใช้กับภาพเอกสารที่มีรูปภาพปะปนไม่ได้ ความถูกต้องน้อยลงเมื่อภาพเอกสารมีความเอียงมากขึ้น

ความถูกต้องของการวิเคราะห์ความเอียงนั้นขึ้นอยู่กับองค์ประกอบหลักๆดังต่อไปนี้

1. ความชัดเจนของภาพเอกสาร ไม่ว่าจะเป็นเรื่องของแสงเงา ความคมชัดของตัวหนังสือหรือสัญญาณ
   รบกวน(Noise) ต่างๆ ยิ่งมีความชัดเจนของภาพเอกสารมากโอกาสของความถูกต้องก็มีสูง
2. ปริมาณของตัวหนังสือก็มีผลเช่นเดียวกัน โดยเฉพาะภาพเอกสารที่มีรูปภาพปะปนอยู่ด้วย ถ้ามีรูปมากเกินโอกาสผิดพลาดมีสูง จากการทดลองของโปรแกรม ถ้ามีรูปกับตัวหนังสืออย่างละประมาณร้อยละ 50 โอกาสที่จะหมุนภาพให้ตรงมีประมาณร้อยละ 60 ถ้ามีรูปประมาณร้อยละ 30 ตัวหนังสือประมาณร้อยละ 70 โอกาสที่จะหมุนภาพให้ตรงมีประมาณร้อยละ 80 ส่วนภาพเอกสารที่เป็นตัวหนังสือล้วนไม่ค่อยมีปัญหาเท่าไหร่ โอกาสหมุนได้ถูกอยู่ที่ประมาณร้อยละ 90
3. ความเอียงของภาพเอกสาร ยิ่งเอียงมากโอกาสหมุนผิดก็มีมาก โดยฉะเพราะมุมที่เอียงตั้งแต่ -85 ขึ้นไปจนถึง -90 และมุมที่มีโอกาสผิดได้บ่อยๆอีกมุมหนึ่งก็คือ

5.2 ปัญหาและแนวทางการแก้ไขปัญหา

5.2.1 อัตราความเร็วในการวิเคราะห์ความเอียงของภาพเอกสารและหมุนกลับให้ตรง

เนื่องจากขอบเขตของความเอียงมีมาก คือมุมตั้งแต่ -90 ถึง 90 ซึ่งรวมๆแล้วก็ 180 มุม กว่าจะตรวจพบมุมที่ถูกต้อง กว่าจะพบมุมที่คิดว่าตรง โปรแกรมใช้เวลาค่อนข้างนานมาก จึงทำให้อัตราความเร็วในการทำงานอยู่ในขั้นต่ำ

การแก้ไขปัญหา ใช้การหมุนแบบบล็อกเดียว การหมุนแบบบล็อกเดียวคือใช้พื้นที่เพียงส่วนเดียวในการวิเคราะห์ความเอียงของภาพเอกสาร สมมติให้ขนาดของบล็อกแต่ละบล็อก

อยู่ที่ร้อยละ 20 ของภาพเอกสาร เพราะฉะนั้นพื้นที่ที่ใช้วิเคราะห์ก็จะเล็กลงถึง 5 เท่าทำให้อัตราความเร็วเพิ่มขึ้น 5 เท่า แต่ก็มีความเสี่ยงที่จะหมุนออกมาไม่ตรงมากกว่าการหมุนแบบทุกบล็อก

5.2.2 การกำหนดค่าคงที่

ในโปรแกรมได้มีการกำหนดค่าหลายค่าขึ้นมาเพื่อใช้เป็นค่าคงที่ในการตัดสินใจทำอย่างใดอย่างหนึ่งเช่น ค่าเบี่ยงเบนมาตรฐาน(Standard Deviation) ซึ่งกำหนดไว้ว่า ถ้าบล็อกใดเป็นบล็อกของตัวหนังสือจะมีค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) ไม่เกิน 100 ซึ่งเป็นค่าที่เกิดจากการทดลอง ถ้าเกิน 150 ถือว่าเป็นรูปภาพ ซึ่งการกำหนดในลักษณะนี้มันก็ไม่แน่นอนเสมอไป และยังมีอีกหลายๆค่าที่มีลักษณะนี้

การแก้ไขปัญหา สร้างหน้าจอGUI ส่วนของตั้งค่าขึ้นมาให้ผู้ใช้งานสามารถปรับค่าต่างๆเองได้ และสร้างปุ่มกำหนดค่าเริ่มต้น (Default) ไว้เป็นค่าที่เกิดจากการทดลอง ผู้ใช้งานสามารถเลือกค่าตามต้องการหรือจะใช้ค่าเริ่มต้น (Default) ที่กำหนดไว้ให้ก็ได้ แต่การใช้งานโปรแกรมดูเหมือนจะเป็นอัตโนมัติ เพราะต้องมีการกำหนดค่าก่อน แนวคิดที่จะแก้ไขปัญหานี้ในครั้งต่อไปก็คือใช้การเรียนรู้ของคอมพิวเตอร์ (Machine Learning)

5.3 แนวทางในการพัฒนาครั้งต่อไป

แนวทางในการพัฒนาครั้งต่อไป ได้มีแนวคิดว่าอยากจะเอาการเรียนรู้ของเครื่องคอมพิวเตอร์(Machine Learning) มาช่วยวิเคราะห์ความเอียงของภาพเอกสารและช่วยในการแยกแยะว่าแต่ละบล็อกในภาพ บล็อกไหนเป็นบล็อกของตัวหนังสือ บล็อกไหนเป็นบล็อกของรูปภาพ และพัฒนาเรื่องของอัตราความเร็วให้มีมากขึ้น

อีกเรื่องที่อยากต่อยอดคือ การปรับความเอียงของวัตถุ, ตรวจสอบว่าวัตถุตั้งตรงหรือเปล่า, เอียงกี่องศาหรือ การรักษาสมดุลให้วัตถุตั้งตรงอยู่ตลอดเวลา เมื่อมีปัจจัยภายนอก เช่น ลมมากระทำก็ยังสามารถรักษาสมดุลไว้ได้

เอกสารอ้างอิง

[1] ยุทธนา แสงสุวรรณ. (16 พฤษจิกายน 2552). กราฟิกที่ใช้ในงานคอมพิวเตอร์. สืบค้นเมื่อ 18 มิถุนายน 2554, จาก http://images.lekyutthana.multiply.multiplycontent.com/journal

[2] เกียรติพงษ์ กตติสุพัฒน์. (2553). รู้จักกับความละเอียดของภาพ. สืบค้นเมื่อ 18 มิถุนายน 2554, จาก http://www.108award.com/index.php?lay=show&ac=article&Id=538696525&Ntype=1

[3] edu-mine. (2552). โหมดสีต่างๆ. สืบค้นเมื่อ 18 มิถุนายน 2554, จาก http://www.edu-mine.com/photoshop/lesson5_colormode.html

[4] ยุทธพงษ์ รังสรรค์เสรี และ พรพรรณ์ คุลยกาญจน์. (2543). การจำแนกข้อมูลภาพโดยการทำเทรชโฮลด์หลายระดับของค่าทางสถิติอันดับที่สองของระดับสีเทา. วิทยานิพนธ์, มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ. สืบค้นเมื่อ 28 กุมภาพันธ์ 2554,จาก http://kucon.lib.ku.ac.th/Fulltext/KC3811020.pdf

[5] นิรุธ อำนวยศิลป์. (2554). การประมวลผลภาพและวีดีโอด้วย. กรุงเทพฯ: บริษัทด่านสุทธาการพิมพ์จำกัด

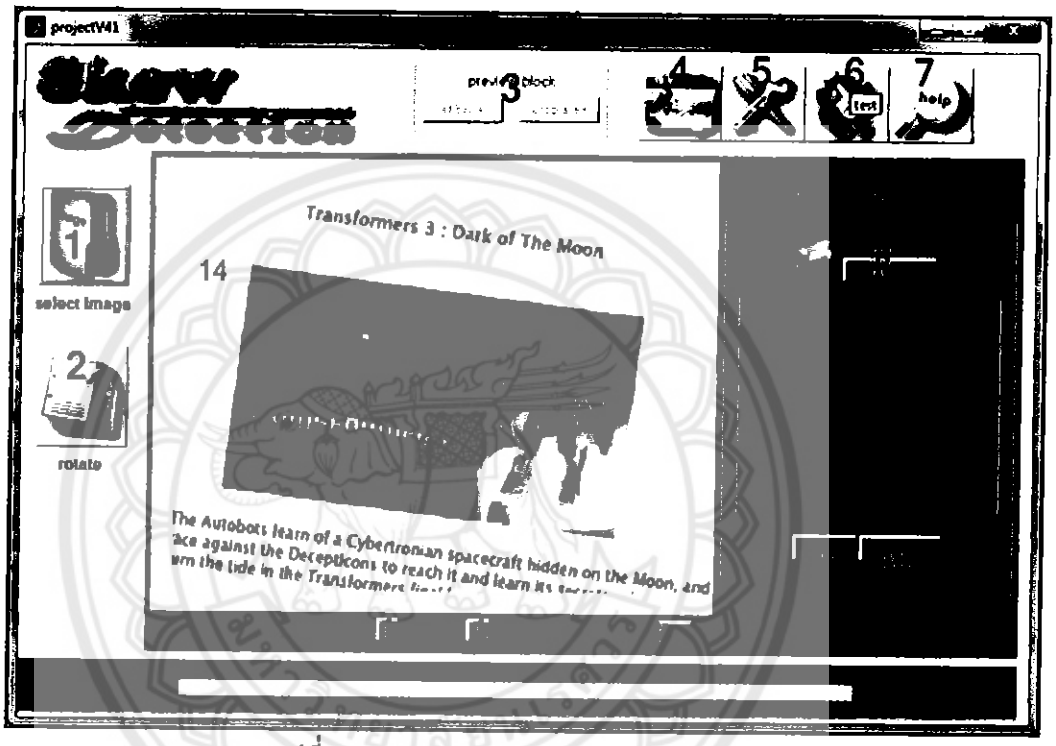
[6] มนตรี กาญจนะเดชา. (2554). การแปลงข้อมูลภาพสองมิติ. สืบค้นเมื่อ 10 กรกฎาคม 2554, จาก http://fivedots.coe.psu.ac.th/~montri/Teaching/image/image.html

[7] JONATHAN J. HULL. "Document Image Skew Detection: Survey and Annotated Bibliography", in document Analysis Systems II, J.J.Hull, S.L.Taylor (Eds.), World Scientific, pp. 40-64,1998.

[8] ศิวาพร ศรเจริญ. (2548). การค้นหาและติดตามวัตถุในระนาบ 2 มิติ. วิทยานิพนธ์, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี. กรุงเทพฯ. สืบค้นเมื่อ 7 กุมภาพันธ์ 2555, จาก http://cpe.kmutt.ac.th/previousproject/2005 /2/index.htm

ภาคผนวก

1. ส่วนประกอบต่างๆของโปรแกรมปรับภาพเอกสารให้ตรงโดยอัตโนมัติ

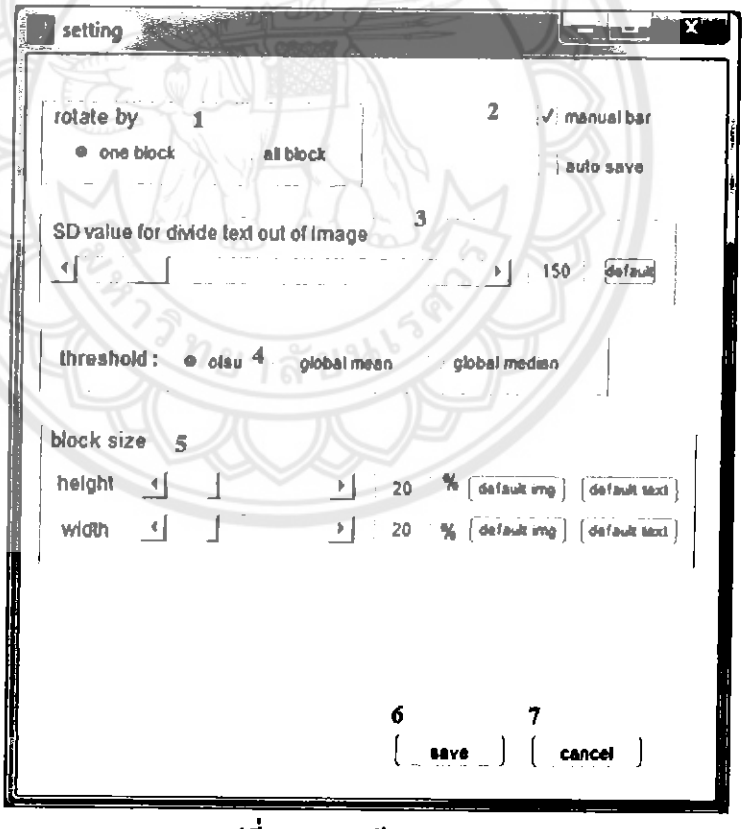


รูปที่ 1แสดงส่วนประกอบต่างๆของโปรแกรม

1. ปุ่ม select image=เลือกภาพเอกสารที่ต้องการปรับให้ตรง สามารถเลือกภาพเอกสารที่ต้องการจะปรับได้มากกว่า 1 ภาพ
2. ปุ่ม rotate=หมุนภาพเอกสารให้ตรง
3. ปุ่ม preview block=แสดงการแบ่งบล็อกทั้งหมดของภาพเอกสารและบล็อกที่เป็นตัวหนังสือ
4. ปุ่ม directory=ภาพเอกสารที่บันทึกอัตโนมัติ (Auto Save) จะถูกบันทึกไว้ที่นี่
5. ปุ่ม setting=เซ็ตค่าต่างๆในโปรแกรม
6. ปุ่ม tester= ตรวจเช็คมุมของภาพเอกสารที่ดีเท็คได้
7. ปุ่ม help= อธิบายวิธีใช้โปรแกรม
8. ปุ่ม start camera = เปิดกล้องเว็บแคม
9. ปุ่ม capture = ถ่ายภาพจากกล้องเว็บแคม

10. ปุ่ม capture&rotate = ถ่ายภาพและหมุนภาพ เมื่อถ่ายภาพเอกสารเสร็จ ภาพจะหมุนตรงโดยอัตโนมัติ
11. ปุ่มปรับภาพเอกสารให้หมุนทีละ +90 องศา
12. ปุ่มปรับภาพเอกสารให้หมุนทีละ-90 องศา
13. บันทึก (Save) ภาพที่แสดงบนหน้าต่างแสดงภาพเอกสาร
14. หน้าต่างแสดงภาพเอกสาร
15. หน้าจอแสดงภาพจากกล้องเว็บแคม
16. Progress bar = แถบแสดงความคืบหน้าขณะกำลังรันโปรแกรม

2. ส่วนประกอบต่างๆและการใช้งานในหน้าต่าง setting

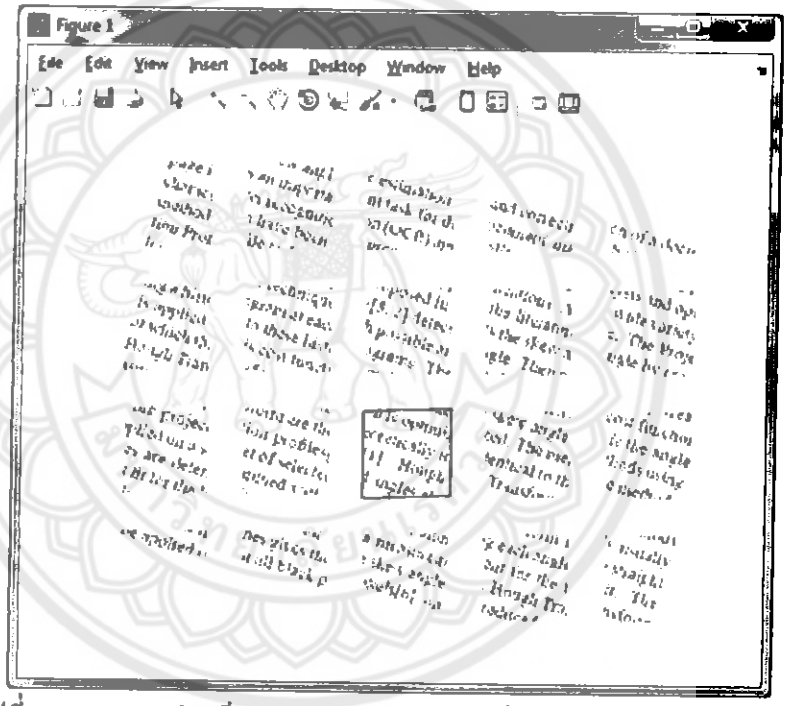


รูปที่ 2แสดงหน้าต่าง setting

1 rotate by

1.1 one block คือ จะใช้นำบล็อก (block) เดียวที่อยู่ตรงกลางของภาพเอกสาร ตัวอย่างดังรูปที่ 3 บล็อกที่อยู่ตรงกลางคือ บล็อกที่มีกรอบสีแดง จะนำบล็อกนี้บล็อกเดียวไปประมวณผล ซึ่งวิธีการนี้จะทำให้โปรแกรมประมวลผลได้เร็วมาก แต่ความถูกต้องก็จะลดลง วิธีการนี้เหมาะสำหรับภาพเอกสารที่เป็นตัวหนังสือทั้งหมด ไม่มีรูปภาพประกอบ

1.2 all block คือ นำทุกบล็อกไปประมวลผล วิธีการนี้จะทำให้โปรแกรมทำงานช้า แต่มีความถูกต้องสูง ใช้ได้ทั้งภาพเอกสารที่มีรูปภาพประกอบและ ภาพเอกสารที่เป็นตัวหนังสือทั้งหมด



รูปที่ 3แสดงการแบ่งบล็อก (block) และการหาบล็อกตรงกลางของโปรแกรม

2. เป็นส่วนของ check box

- manual barถ้าใส่เครื่องหมายถูกคือให้แสดงแถบ manual bar ในหน้าต่างโปรแกรมหลัก
- auto save เมื่อปรับภาพเอกสารเสร็จ ให้บันทึก (save) ภาพเอกสารนั้นโดยอัตโนมัติ

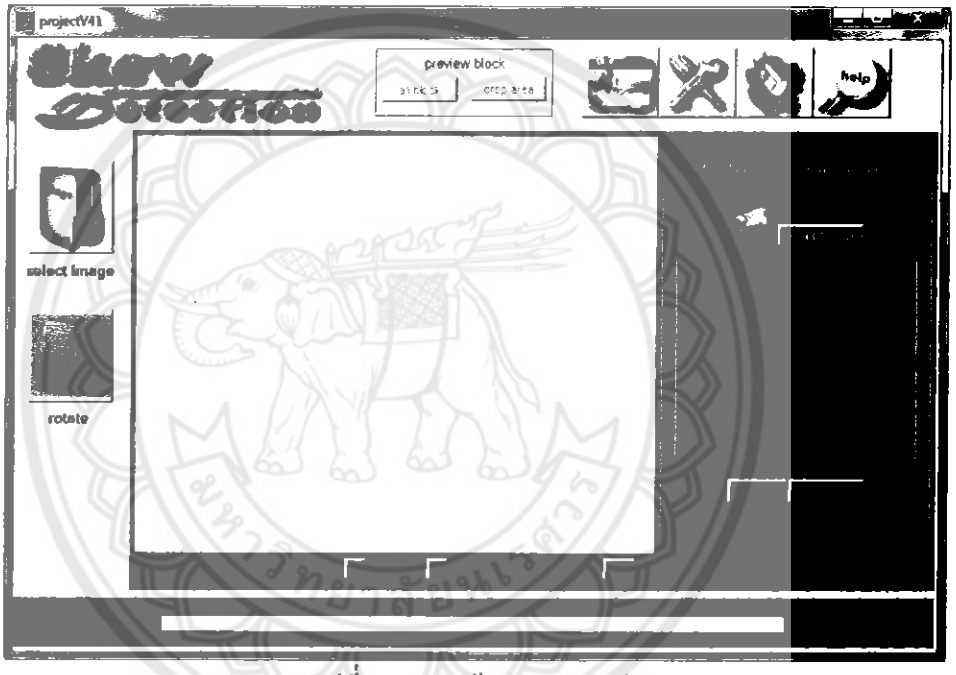
3. กำหนดค่าเบี่ยงเบนมาตรฐาน (standard deviation)

ค่าเริ่มต้น(default) ของโปรแกรมจะตั้งไว้ที่ 150 ผู้ใช้(user) สามารถเปลี่ยนแปลงค่าตามที่ต้องการได้

4. กำหนดวิธีการเทรชโฮลด์ มี 3 แบบให้เลือก คือ Otsu, Global Mean และ Global Median ค่าเริ่มต้น (default) ของโปรแกรมจะเป็นวิธีการของ Otsu's Threshold

5. กำหนดขนาดของบล็อก (block) แบ่งเป็น 2 ส่วนคือขนาดความกว้างของบล็อกและขนาดความยาวของบล็อกค่าเริ่มต้น(default) ของโปรแกรม ถ้าเป็นภาพเอกสารที่เป็นตัวหนังสือทั้งหมดจะตั้งค่าไว้ที่ 40 ถ้าภาพเอกสารมีรูปภาพประกอบด้วยจะตั้งค่าไว้ที่ 20

2. วิธีการใช้งานโปรแกรม

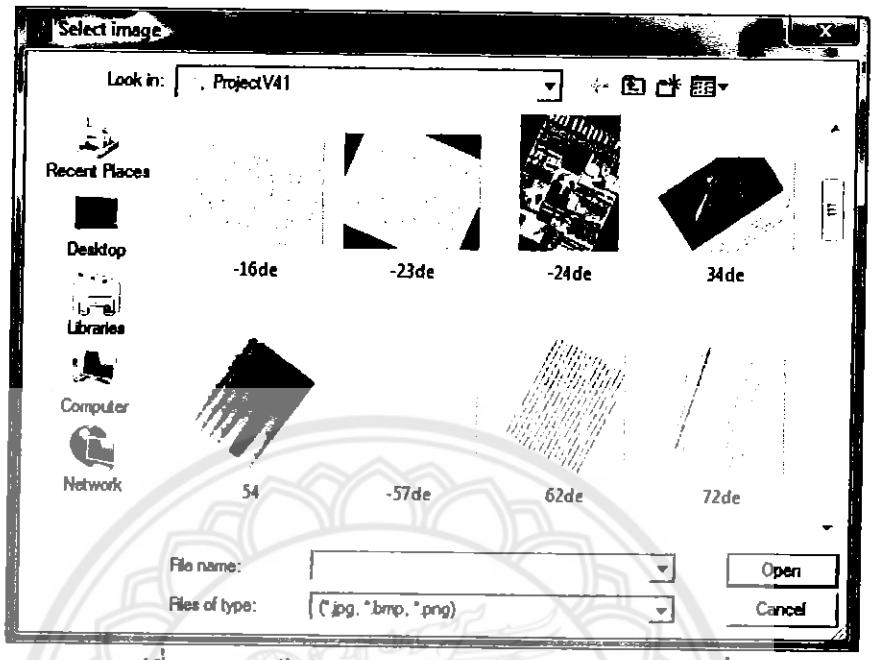


รูปที่ 4แสดงหน้าจอหลักของโปรแกรม

2.1 ปรับภาพเอกสารจากไฟล์ภาพที่มีอยู่แล้ว

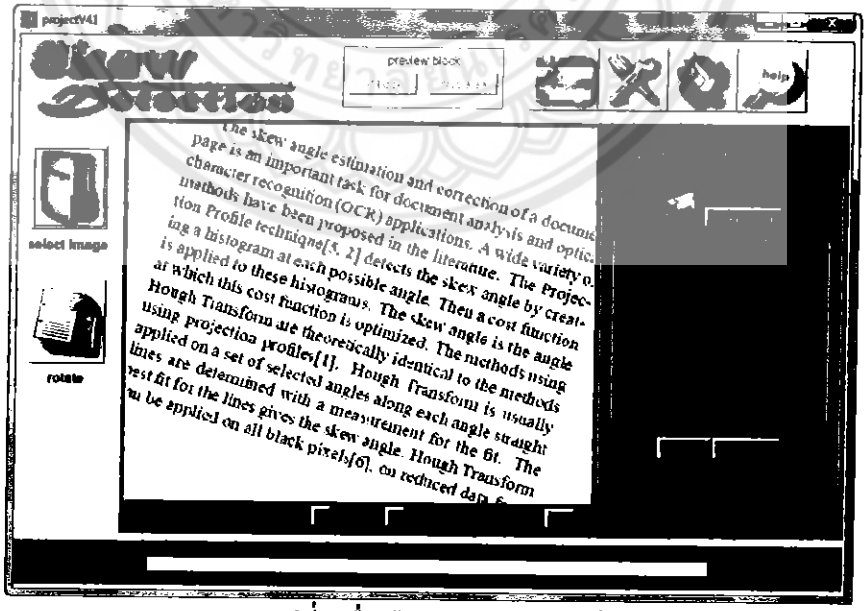
เมื่อเปิดโปรแกรมปรับภาพเอกสารให้ตรงโดยอัตโนมัติ จะมีลักษณะของหน้าจอหลักดังรูปที่ 2

1. คลิ๊กที่ปุ่ม select imageจะปรากฏหน้าต่าง Select image ดังในรูป



รูปที่ 5แสดงหน้าต่าง Select image ให้เลือกภาพเอกสารที่ต้องการ

2. เลือกภาพเอกสารที่ต้องการ แล้วคลิ๊ก Open
3. ภาพเอกสารที่เลือกจะปรากฏที่หน้าต่างแสดงภาพเอกสารดังรูปที่ 4



รูปที่ 6เมื่อเลือกภาพเอกสารเสร็จแล้ว

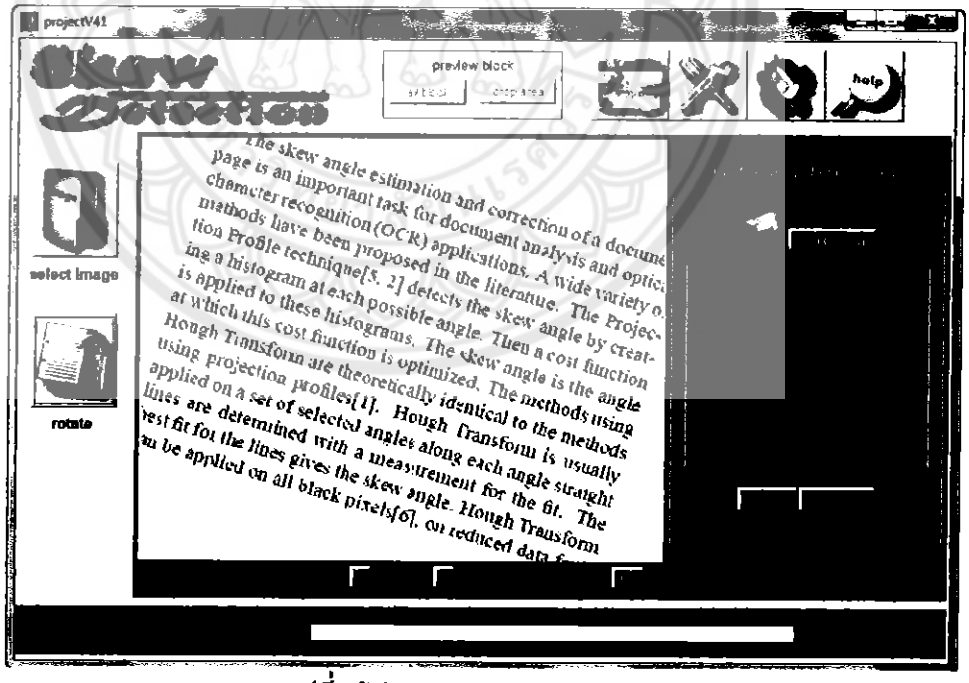
4. คลิ๊กที่ปุ่ม rotate เพื่อปรับภาพเอกสารให้ตรง

ขณะโปรแกรมกำลังทำงาน จะมีแถบแสดงความคืบหน้าของการทำงานอยู่ด้านส่วนล่างสุดของโปรแกรม โดยจะแบ่งออกเป็น 3 ส่วนย่อยๆดังนี้

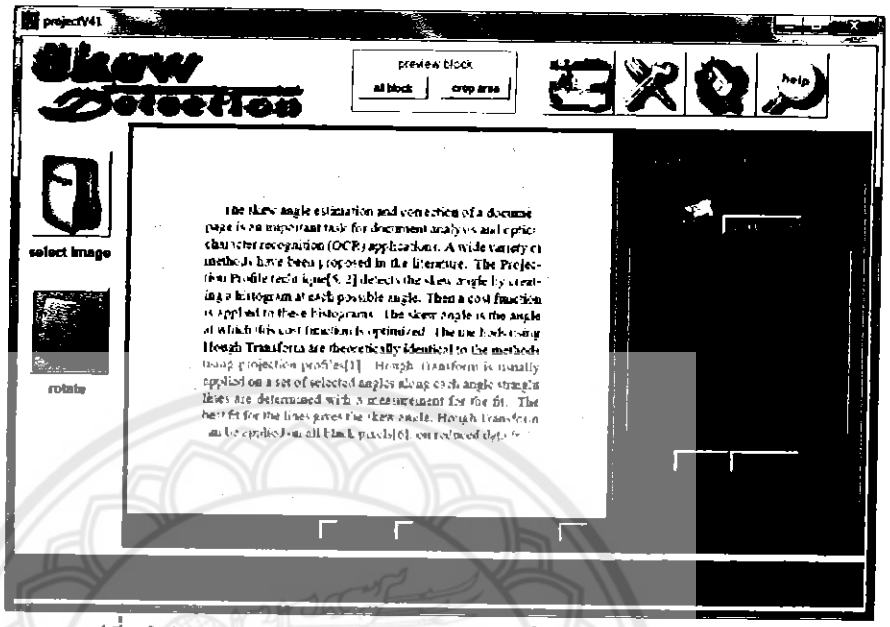
รูปที่ 7แถบแสดงความคืบหน้าการทำงานของโปรแกรม

1. จำนวนของภาพเอกสาร เลขตัวหน้าแสดงภาพเอกสารที่กำลังทำอยู่ เลขตัวหลังคือจำนวนภาพเอกสารทั้งหมด
2. แถบแสดงความคืบหน้าของการประมวลผลของแต่ละภาพ
3. ร้อยละ (%) การประมวลผลของแต่ละภาพ

จากตัวอย่างในรูปที่ 5 คือ กำลังประมวลผลภาพเอกสารที่ 1 จากทั้งหมด 1 ภาพ และโปรแกรมทำงานได้ 23 %



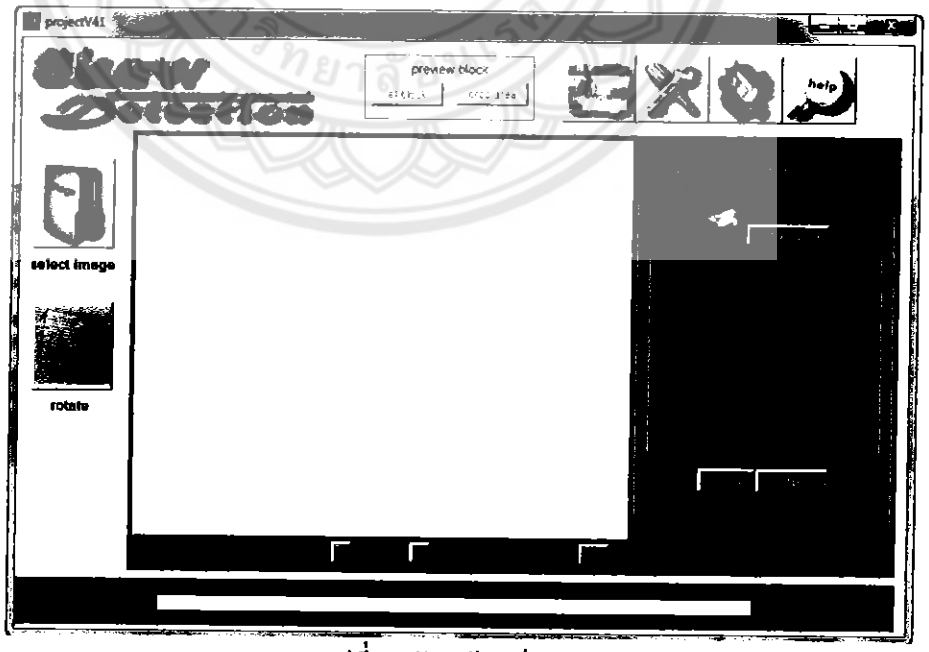
รูปที่ 8 โปรแกรมกำลังประมวลผล



รูปที่ 9โปรแกรมปรับภาพเอกสารให้ตรงเสร็จเรียบร้อยแล้ว

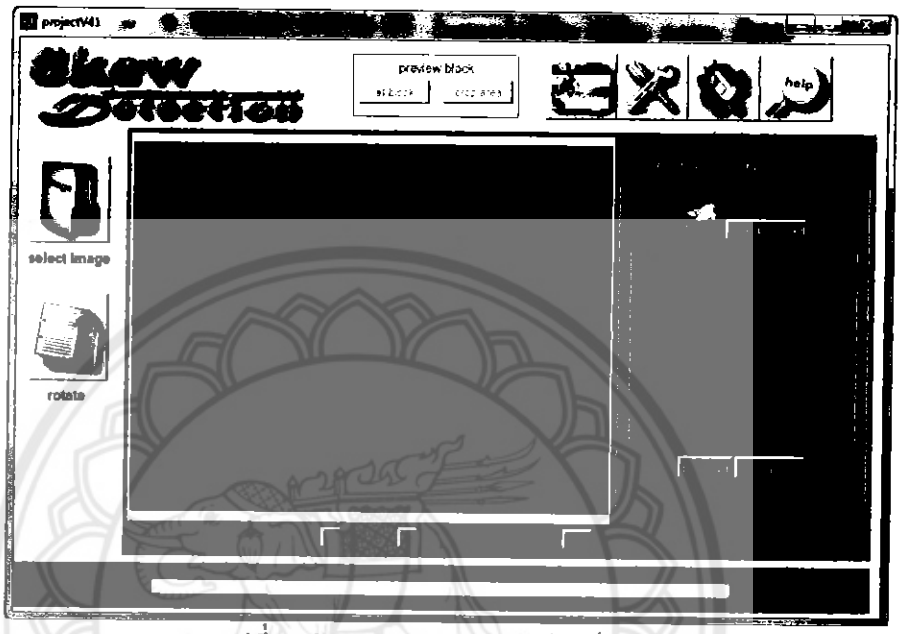
2.2 ปรับภาพเอกสารโดยการถ่ายภาพจากกล้องเว็บแคม

1. คลิ๊กที่ปุ่ม start cameraเพื่อเปิดกล้องเว็บแคม



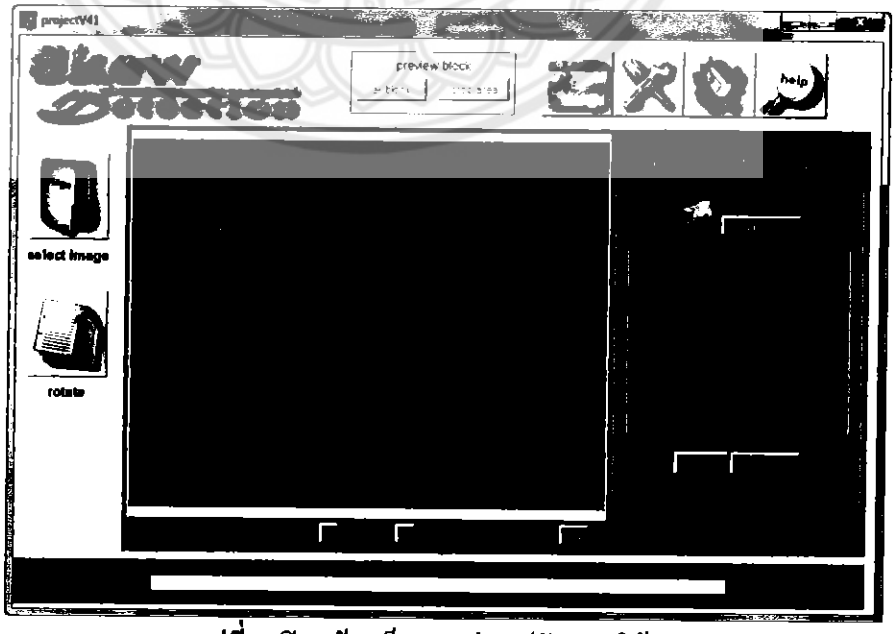
รูปที่ 10เปิดกล้องเว็บแคม

2. ถ่ายภาพที่ต้องการ ตอนถ่ายภาพนั้นควรมีแสงสว่างเพียงพอ ไม่มืดจนเกินไป



รูปที่ 11ถ่ายภาพเอกสารจากกล้องเว็บแคม

3. เมื่อถ่ายภาพที่ต้องการเสร็จแล้ว ให้คลิ๊กที่ปุ่ม stop camera เพื่อปิดกล้องเว็บแคม

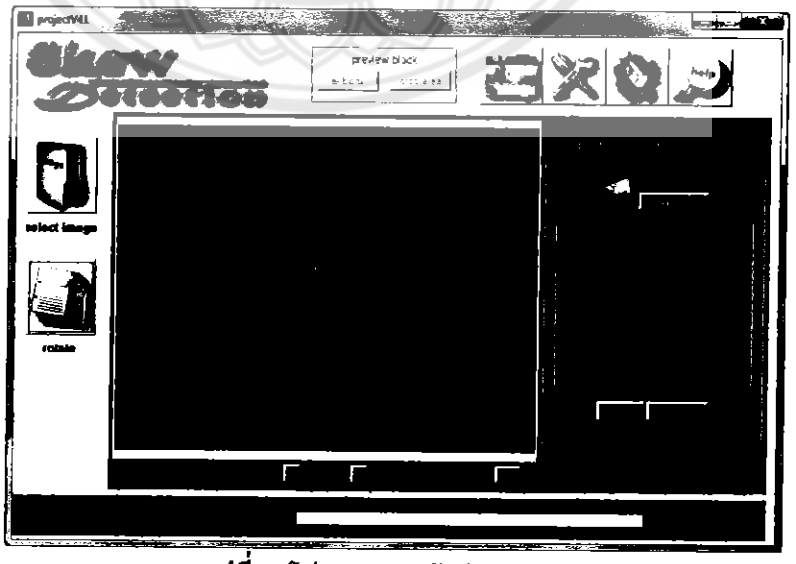


รูปที่ 12ปิดกล้องเว็บแคมก่อนปรับภาพให้ตรง

4.ถ้าต้องการบันทึกภาพที่ถ่ายเพื่อเก็บไว้ ให้คลิ๊กที่ปุ่ม save จะมีหน้าต่าง Save Image As ปรากฏขึ้นมา ให้ตั้งชื่อและเลือกที่อยู่ตามที่ต้องการ เสร็จแล้วให้คลิ๊กปุ่ม save

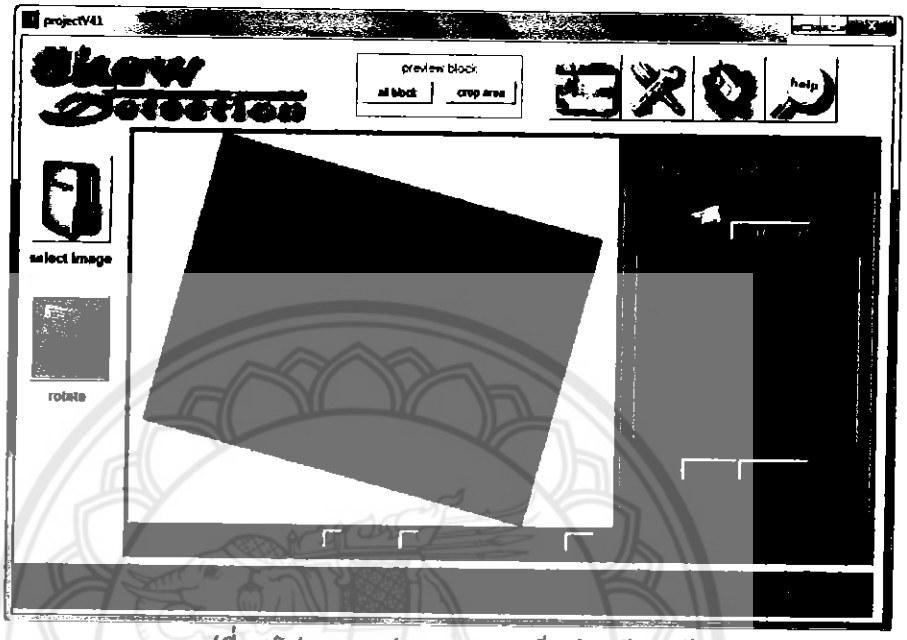
รูปที่ 12แสดงหน้าต่าง Save Image As

5. คลิ๊กปุ่ม rotate เพื่อปรับภาพเอกสารให้ตรง



รูปที่ 13โปรแกรมกำลังประมวลผล

6. โปรแกรมประมวลผลเสร็จเรียบร้อยแล้ว



รูปที่ 13 โปรแกรมประมวลผลเสร็จเรียบร้อยแล้ว