

การแบ่งส่วนประกอบเอกสาร  
โดยใช้เส้นขอบและการแบ่งในแนว XY

DOCUMENT SEGMENTATION  
USING CONTOUR AND XY CUT TECHNIQUE



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ.2548

ISBN 974-15-1978-8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**DOCUMENT SEGMENTATION**  
**USING CONTOUR AND XY CUT TECHNIQUE**



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

**2005**

**ISBN 974-15-1978-8**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



**COPYRIGHT 2005**

**SCHOOL OF GRADUATE STUDIES**

**KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์

การแบ่งส่วนประกอบเอกสาร โดยใช้เส้นขอบและการ  
แบ่งในแนว XY

นักศึกษา

นาย ณรงค์ชัย มุ่งแสงกลาง

รหัสประจำตัวนักศึกษา

44061631

ปริญญา

วิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชา

วิศวกรรมคอมพิวเตอร์

พ.ศ.

2548

อาจารย์ผู้ควบคุมวิทยานิพนธ์

รศ.ดร.บุญธีร์ เกียรติตราฐ

## บทคัดย่อ

การแบ่งส่วนประกอบของภาพเอกสาร โขนารี่เป็นขั้นตอนในกรทำงานเพื่อที่จะใช้ในการวิเคราะห์โครงสร้างของภาพเอกสาร ซึ่งโดยทั่วไปภาพเอกสารในหน้ากระดาษอาจจะประกอบไปด้วยคอลัมน์เดี่ยวหรือหลายคอลัมน์ ภายในแต่ละคอลัมน์อาจจะประกอบไปด้วยข้อความหรือรูปภาพ ดังนั้นงานในการแบ่งส่วนประกอบของภาพเอกสารจึงเป็นการกำหนดขอบเขตของภาพเอกสารแบบอัตโนมัติเพื่อที่จะนำไปใช้ในกระบวนการรู้จำ วิทยานิพนธ์ฉบับนี้ได้เสนอวิธีการแบ่งส่วนประกอบของภาพเอกสารที่มีความรวดเร็วและมีประสิทธิภาพ โดยอาศัยแนวความคิดแบบผสมผสานระหว่างการใช้การติดตามเส้นขอบและเทคนิคการแบ่งในแนว XY การทำงานจะเริ่มจากใช้เทคนิคในการติดตามเส้นขอบของภาพเพื่อให้ได้ขอบเขตโดยรวม จากนั้นใช้เทคนิคการแบ่งในแนว XY เข้ามาช่วยในการแก้ไขปัญหากรมเบ่งภาพเอกสารผิดพลาด ซึ่งในขั้นแรกของการแบ่งเอกสารในแนว XY จะเป็นการระบุตำแหน่งของจุดผิดพลาดบนเส้นขอบ ขั้นตอนต่อไปจะใช้ตำแหน่งของจุดผิดพลาดเพื่อที่จะกำหนดแนวทางการแบ่งส่วนประกอบครั้งใหม่ในแนวนอนหรือแนวตั้ง เพื่อที่จะตัดแบ่งจุดผิดพลาดเหล่านั้นให้มีความถูกต้องในส่วนของการคอลัมน์ยิ่งขึ้น วิธีการที่นำเสนอมีข้อดีคือไม่จำเป็นต้องเข้าถึงข้อมูลของจุดภาพพิกษ จุดภาพในการระบุจุดผิดพลาด ดังนั้นจึงมีความรวดเร็วกว่าวิธีการที่อาศัยการเข้าถึงข้อมูลของจุดภาพพิกษ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<b>Thesis Title</b>	Document Segmentation Using Contour and XY Cut Technique
<b>Student</b>	Narongchai Moongfangklang
<b>Student ID.</b>	44061631
<b>Degree</b>	Master of engineering
<b>Program</b>	Computer engineering
<b>Year</b>	2005
<b>Thesis Advisor</b>	Assoc.Prof. Boontee Kruaprachue

## ABSTRACT

Segmentation of binary document image is one of the document structure analysis processes. Normally image documents compose single-column or multi-columns, within each column consist text graphic and picture. Therefore, the task of image document segmentation is to identify the region of them automatically for using in optical character recognition. This thesis proposes the fast and efficient method for page segmentation of document containing nonrectangular block. The segmentation is based on edge following algorithm using small window of 16 by 32 pixels. This segmentation is very fast since only border pixels of paragraph are used without scanning the whole page. Still, the segmentation may contain error if the space between them is smaller than the window used in edge following. Consequently, proposed method reduces this error by first identifying the missed segmentation point using direction information in edge following, then, using XY cut at the missed segmentation point to separate the connected columns. The advantage of the proposed method is the fast identification of missed segmentation point. This methodology is faster with fewer overheads than other algorithms that need to access much more pixel of a document.

# สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VII
สารบัญรูป.....	IX
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการศึกษา.....	2
1.3 สมมุติฐานของการศึกษา.....	2
1.4 ทฤษฎีและแนวคิดที่ใช้ในการวิจัย.....	2
1.5 ขอบเขตของงานวิจัย.....	3
1.6 ขั้นตอนการศึกษา.....	3
1.7 รายละเอียดเครื่องมือและอุปกรณ์ที่ใช้ในการทำวิจัย.....	4
1.8 รายละเอียดของวิทยานิพนธ์.....	4
บทที่ 2 ทฤษฎีและแนวคิดที่เกี่ยวข้อง.....	5
2.1 การแสดงภาพดิจิทัล (Digital Image Representation).....	5
2.2 ขั้นตอนพื้นฐานในการประมวลผลภาพ.....	6
2.3 ความรู้พื้นฐานเกี่ยวกับข้อมูลภาพดิจิทัลพรอนิกส์.....	7
2.3.1 รูปแบบกลุ่มข้อมูลและระดับเฉดสี (Uniform Sampling and Quantization).....	7
2.3.2 ความสัมพันธ์ระหว่างจุดภาพ (Relationships between pixels).....	7
2.3.3 การเชื่อมต่อกันระหว่างจุดจะใช้ในการหาค่าขอบเขตพื้นที่ของวัตถุและ ส่วนประกอบ.....	8
2.4 รหัสแนวหรือรหัสลูกโซ่ (Chain Codes).....	8
2.5 คุณลักษณะพื้นฐานของภาพเอกสาร.....	9
2.6 ประเภทของเอกสารและการประยุกต์ใช้.....	10
2.6.1 เอกสารรูปแบบโครงสร้าง (structure articles).....	10
2.6.2 เอกสารที่มีรูปแบบไม่แน่นอน.....	11

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญ (ต่อ)

	หน้า
2.6.3 เอกสารกึ่งรูปแบบหรือกึ่งโครงสร้าง (Semi-structured layouts).....	13
2.6.4 แผนที่และภาพวาด.....	14
2.6.5 ภาพหน้าเว็บ และภาพจากวิดีโอ .....	15
2.7 ขั้นตอนการแบ่งหน้าเอกสารและวิธีการแบ่งประเภทขอบเขต .....	16
2.7.1 การแบ่งส่วนประกอบเอกสาร .....	17
2.7.2 การแบ่งประเภทขอบเขตเอกสาร .....	21
2.8 การแบ่งประเภทขอบเขตเอกสาร .....	23
2.8.1 วิธีการรีนแรงค์สมูทอัลกอริทึม (run/length smoothing).....	23
2.8.2 วิธีการแบ่งแนวรั้วแกน X-Y (recursive projection X-Y cut).....	23
2.8.3 วิธีการเดินวนรอบขอบ (contouring).....	24
2.8.4 วิธีการแบบไวท์ไทล์ (white tile) .....	24
2.9 เปรียบเทียบวิธีการทำงาน .....	25
2.10 วิเคราะห์วิธีการทำงาน .....	28
2.11 สรุป .....	31
บทที่ 3 การแบ่งส่วนประกอบภาพเอกสาร โดยให้เส้นขอบ และการแบ่งในแนว XY.....	34
3.1 รูปร่างของส่วนประกอบภาพเอกสาร .....	34
3.2 แนวคิดในการแบ่งส่วนประกอบภาพเอกสาร ในแนวนอน .....	35
3.3 วิธีการแบ่งภาพเอกสารเป็นส่วนประกอบใหญ่ .....	36
3.3.1 การหาขนาดหน้าต่างที่เหมาะสม .....	37
3.3.2 การหาตำแหน่งเริ่มต้นของหน้าต่าง .....	39
3.3.3 การเดินรอบขอบเขตส่วนประกอบใหญ่ .....	41
3.3.4 การหาขอบเขตส่วนประกอบใหญ่ (ขอบซ้ายขอบขวา).....	43
3.4 การระบุจุดผิดพลาดของแบ่งส่วนประกอบใหญ่ .....	45
3.5 การแบ่งส่วนประกอบซ้ำด้วยวิธีการแบ่งซ้ำในแนว XY .....	50
3.5.1 การแบ่งในแนวตั้ง (Vertically separation) .....	50
3.5.2 การแบ่งในแนวนอน (Horizontally separation).....	53
3.6 การปรับปรุงขอบเขตส่วนประกอบ .....	55

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญ (ต่อ)

	หน้า
บทที่ 4 วิธีการแบ่งจุดผิดพลาด.....	58
4.1 การติดตามเส้นขอบโดยแบ่งหน้าต่างย่อย.....	58
4.2 การแพร่โดยจำแนกประเภทของหน้าต่าง.....	61
4.4 การแพร่โดยอาศัยพิกัดตำแหน่งเป้าหมาย.....	66
4.5 การแบ่งซ้ำตามแนวแกน XY.....	69
บทที่ 5 ผลการทดลอง วิเคราะห์ และสรุปผลการทดลอง.....	71
5.1 ผลการทดลอง.....	71
5.2 เปรียบเทียบวิธีการอื่นๆ.....	72
5.2.1 การเปรียบเทียบเวลาและจำนวนจุดภาพที่ใช้ในการแบ่งส่วนประกอบ.....	72
5.2.2 การประเมินผลลัพธ์เชิงปริมาณในการแบ่งส่วนประกอบภาพเอกสาร (Quantitative evaluation of segmentation results).....	75
5.3 สรุปผลการทดลอง.....	79
5.4 วิเคราะห์ผลการทดลอง.....	83
5.5 ภาพตัวอย่างผลลัพธ์จากการประมวลผล.....	83
5.3.1 ภาพเอกสารขนาดเล็ก (รูปที่ 5.16-5.19).....	83
5.3.2 ภาพเอกสารขนาดกลาง (รูปที่ 5.20-5.23).....	85
5.3.3 ภาพเอกสารขนาดใหญ่ (รูปที่ 5.24-5.33).....	88
บทที่ 6 สรุปผลการวิจัยและข้อเสนอแนะ.....	98
6.1 สรุปผลการวิจัย.....	98
6.2 ข้อเสนอแนะแนวทางการทำวิจัยต่อ.....	99
เอกสารอ้างอิง.....	100
ภาคผนวก ก.....	106
ภาคผนวก ข.....	114
ภาคผนวก ค.....	141
ประวัติผู้เขียน.....	146

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญตาราง

ตารางที่	หน้า
2.1 แสดงคุณสมบัติที่สำคัญของวิธีการต่างๆ ที่ใช้ในการวิเคราะห์แผนผังของเอกสาร.....	26
2.2 แสดงเวลาที่ใช้ในการทำงานสำหรับวิธีการวิเคราะห์เอกสาร.....	27
2.3 แสดงผลงานวิจัยที่ผ่านมา.....	32
4.1 แสดงการจำแนกประเภทของหน้าต่าง.....	63
5.1 แสดงผลการทดลองด้วยวิธีการวิธีการการติดตามเส้นขอบและการแบ่งโนแนว XY.....	72
5.2 สรุปผลการทดลองพร้อมทั้งเปรียบเทียบผลการทดลองกับวิธีการอื่นๆ.....	80
5.3 แสดงผลสรุปของค่าเฉลี่ยการประเมินผลลัพธ์เชิงปริมาณเมื่อเปรียบเทียบกับวิธีการอื่น.....	81
5.4 แสดงผลสรุปของเปอร์เซ็นต์ค่าความถูกต้องเฉลี่ยที่ได้จากการประเมินผลลัพธ์เชิงปริมาณเมื่อเปรียบเทียบกับวิธีการอื่น.....	81
5.5 แสดงเวลาในการประมวลภาพเอกสารในแต่ละช่วง.....	82
ก.1 แสดงผลการทดลองแบ่งภาพเอกสาร โดยใช้หน้าต่างขนาดต่างๆ ดังนี้.....	108
ก.2 แสดงการเลือกค่าเรโซลูชันที่ใช้ในการทดลอง.....	110
ก.3 แสดงตัวอย่างการจำแนกประเภทข้อมูลหน้าต่าง โนแนวตั้ง.....	113
ข.1 แสดงผลการสำหรับการแบ่งขอบเขตภาพเอกสารขนาดเล็ก.....	115
ข.2 แสดงผลการสำหรับการแบ่งขอบเขตภาพเอกสารขนาดกลาง.....	117
ข.3 แสดงผลสำหรับการแบ่งขอบเขตภาพเอกสารขนาดใหญ่ที่เป็นตัวอักษรอย่างเดียว.....	119
ข.4 แสดงผลสำหรับการแบ่งขอบเขตภาพเอกสารขนาดใหญ่ที่ประกอบด้วยรูปภาพและตัวอักษร.....	122
ข.5 แสดงผลสำหรับการแบ่งขอบเขตภาพเอกสารขนาดใหญ่ที่ประกอบด้วยรูปภาพและตัวอักษรที่มีโครงสร้างติดกัน.....	125
ข.6 แสดงผลสำหรับการแบ่งขอบเขตภาพเอกสารขนาดใหญ่ที่มีโครงสร้างส่วนประกอบเป็นสี่เหลี่ยม.....	128
ข.7 เปรียบเทียบผลการแบ่งส่วนประกอบเชิงปริมาณ โดยเทียบกับจำนวนส่วนประกอบที่ประเมินด้วยสายตาสำหรับภาพเอกสารขนาดเล็ก.....	133
ข.8 เปรียบเทียบผลการแบ่งส่วนประกอบเชิงปริมาณ โดยเทียบกับจำนวนส่วนประกอบที่ประเมินด้วยสายตาสำหรับภาพเอกสารขนาดกลาง.....	134
ข.9 เปรียบเทียบผลการแบ่งส่วนประกอบเชิงปริมาณ โดยเทียบกับจำนวนส่วนประกอบที่ประเมินด้วยสายตาสำหรับภาพเอกสารขนาดใหญ่.....	135

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญตาราง (ต่อ)

ตารางที่

หน้า

ข.10	เปรียบเทียบผลการแบ่งส่วนประกอบเชิงปริมาณ โดยเทียบกับจำนวนส่วนประกอบที่ ประเมินด้วยสายตาสำหรับภาพเอกสารถามาคใหญ่ที่มีโครงสร้างสี่เหลี่ยม.....	139
------	---	-----



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญรูป

รูปที่	หน้า
2.1	แนวแกน X, Y ที่ใช้อ้างอิงที่กึ่ง ..... 5
2.2	ขั้นตอนพื้นฐานของกระบวนการประมวลผลภาพดิจิทัล ..... 7
2.3	ทิศทางของรหัสแนว..... 8
2.4	ตัวอย่างภาพเอกสารชนิดต่างๆ ..... 12
2.5	ชุดจุดภาพก่อนทำการรันเรจิสสมูท..... 23
2.6	ชุดจุดภาพหลังจากทำการรันเรจิสสมูทด้วยค่าที่กำหนดไว้คือ 3 ..... 23
2.7	การฉายเงาจุดดำในแนวตั้ง และช่วงที่จะแบ่ง ..... 24
2.8	การเดินวนรอบเส้นขอบของภาพเอกสาร ..... 24
2.9	กราฟที่ใช้แสดงการซัดเรียงไวก้าโทบ ..... 25
3.1	ตัวอย่างแผนผังแบบต่างๆ..... 35
3.2	ภาพรวมการทำงานของการที่นำเสนอ..... 36
3.3	ผลจากการเลือกขนาดหน้าต่างใหญ่ไป (32x32) ..... 38
3.4	ผลจากการเลือกขนาดหน้าต่างเล็กไป (16x16) ..... 38
3.5	ผลของการเลือกหน้าต่างที่เหมาะสม (16x32) ..... 38
3.6	การวางตำแหน่งหน้าต่างเริ่มต้นและลักษณะการเดินวนรอบส่วนประกอบใหญ่ของภาพเอกสารที่มีแผนผังสี่เหลี่ยม..... 40
3.7	การวางตำแหน่งหน้าต่างเริ่มต้นและลักษณะการเดินวนรอบส่วนประกอบใหญ่ของภาพเอกสารรูปแบบใดๆ ..... 40
3.8	ทิศทางการเดินหน้าต่าง..... 41
3.9	โครงสร้างข้อมูลที่เก็บส่วนประกอบใหญ่ ..... 44
3.10	แผนผังขั้นตอนการทำงานการแบ่งส่วนประกอบใหญ่..... 44
3.11	ลักษณะการติดตามเส้นขอบของภาพเอกสาร โบนารี ..... 45
3.12	แสดงการวางหน้าต่างรอบส่วนประกอบ..... 47
3.13	ความลึกของจุดผิดพลาด..... 47
3.14	ความกว้างของจุดผิดพลาด..... 47
3.15	แสดงจุดผิดพลาดที่ได้จากการแบ่งส่วนประกอบใหญ่ในครั้งแรก ..... 49
3.16	ลักษณะการเกิดจุดผิดพลาดในการแบ่งส่วนเอกสาร ..... 49
3.17	แนวทางในการแบ่งภาพเอกสาร ..... 51
3.18	แนวทางในการแบ่งภาพเอกสารสำหรับจุดผิดพลาดด้านบน ..... 52

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญรูป (ต่อ)

รูปที่	หน้า	
3.19	แนวทางในการแบ่งภาพเอกสารสำหรับจุดผิดพลาดด้านล่าง.....	52
3.20	แนวทางในการแบ่งภาพเอกสารสำหรับจุดผิดพลาดในแนวนอน .....	54
3.21	แผนผังขั้นตอนการทำงานการแบ่งส่วนประกอบใหญ่ซ้ำเพื่อแยกโครงสร้างส่วนประกอบที่ติดกัน.....	54
3.22	การแทรกถึงคีย์สลับของหน้าล่างในภาพเอกสาร.....	56
3.23	ขอบเขตส่วนประกอบใหญ่ที่ได้รับการปรับปรุงส่วนประกอบแล้ว.....	57
4.1	การแบ่งจุดผิดพลาด โดยใช้หน้าค่างขนาด 8x32 พิกเซลในการติดตามเส้นขอบ .....	59
4.2	การแบ่งส่วนประกอบตรงจุดผิดพลาด โดยใช้หน้าค่างขนาด 8x16 พิกเซลในการติดตามเส้นขอบ.....	60
4.3	แสดงการค่าตัวแปรต่างๆ ที่ใช้ในการจับเนคประเภทหน้าค่าง.....	62
4.4	แสดงตัวอย่างของหน้าค่างประเภทต่างๆ.....	63
4.5	การใช้วิธีกวรพรักับการจับเนคประเภทของหน้าค่าง.....	65
4.6	การแปร โดยกำหนดเป้าหมาย.....	67
4.7	ปัญหาของการแปรโดยกำหนดเป้าหมาย.....	68
4.8	การแบ่งจุดผิดพลาดด้วยวิธีการแบ่งจำนวนแถว XY.....	70
5.1	แสดงผลการเปรียบเทียบเวลาเฉลี่ยที่ใช้ในการประมวลผลของภาพเอกสารแต่ละชนิด .....	73
5.2	แสดงผลการเปรียบเทียบจำนวนจุดภาพที่เข้าถึงในการประมวลผลของภาพเอกสารแต่ละชนิด.....	73
5.3	แสดงผลการเปรียบเทียบเวลาเฉลี่ยที่ใช้ในการประมวลผลของภาพเอกสารแต่ละชนิด .....	73
5.4	แสดงผลการเปรียบเทียบจำนวนจุดภาพที่เข้าถึงในการประมวลผลของภาพเอกสารแต่ละชนิด.....	74
5.5	แสดงผลการเปรียบเทียบเวลาเฉลี่ยที่ใช้ในการประมวลผลของภาพเอกสารแต่ละชนิด .....	74
5.6	แสดงผลการเปรียบเทียบจำนวนจุดภาพที่เข้าถึงในการประมวลผลของภาพเอกสารแต่ละชนิด .....	74
5.7	แสดงแผนภูมิเปรียบเทียบความถูกต้องเฉลี่ย .....	76
5.8	แสดงแผนภูมิเปรียบเทียบค่าผิดพลาดเฉลี่ย .....	76
5.9	แสดงแผนภูมิเปรียบเทียบค่า Average Normalize.....	76
5.10	แสดงแผนภูมิเปรียบเทียบความถูกต้องเฉลี่ย .....	77
5.11	แสดงแผนภูมิเปรียบเทียบค่าผิดพลาดเฉลี่ย .....	77

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญรูป (ต่อ)

รูปที่	หน้า
5.12 แสดงแผนภูมิเปรียบเทียบค่า Average Normalize .....	77
5.13 แสดงแผนภูมิเปรียบเทียบความถูกต้องเฉลี่ย .....	78
5.14 แสดงแผนภูมิเปรียบเทียบค่าผิดพลาดเฉลี่ย .....	78
5.15 แสดงแผนภูมิเปรียบเทียบค่า Average Normalize .....	78
5.16 ผลการแบ่งภาพเอกสารที่ประกอบด้วยตัวอักษรอย่างเดียว .....	83
5.17 ผลการแบ่งภาพเอกสารที่ประกอบด้วยรูปภาพและตัวอักษร .....	84
5.18 ผลการแบ่งภาพเอกสารที่มีโครงสร้างติดกัน .....	84
5.19 ผลการแบ่งภาพเอกสารที่ไม่สามารถแบ่งซ้ำได้ .....	85
5.20 ผลการแบ่งภาพเอกสารที่ประกอบด้วยตัวอักษรอย่างเดียว .....	85
5.21 ผลการแบ่งภาพเอกสารที่ประกอบด้วยรูปภาพและตัวอักษร .....	86
5.22 ผลการแบ่งภาพเอกสารที่มีโครงสร้างติดกัน .....	86
5.23 ผลการแบ่งภาพเอกสารที่ไม่สามารถแบ่งซ้ำได้ .....	87
5.24 ผลการแบ่งภาพเอกสารที่ประกอบด้วยตัวอักษรอย่างเดียว .....	88
5.25 ผลการแบ่งภาพเอกสารที่ประกอบด้วยตัวอักษรอย่างเดียว .....	89
5.26 ผลการแบ่งภาพเอกสารที่ประกอบด้วยรูปภาพและตัวอักษร .....	90
5.27 ผลการแบ่งภาพเอกสารที่ประกอบด้วยรูปภาพและตัวอักษร .....	91
5.28 ผลการแบ่งภาพเอกสารที่มีโครงสร้างติดกัน .....	92
5.29 ผลการแบ่งภาพเอกสารที่มีโครงสร้างติดกัน .....	93
5.30 ผลการแบ่งภาพเอกสารที่มีโครงสร้างติดกัน .....	94
5.31 ผลการแบ่งภาพเอกสารที่ไม่สามารถแบ่งซ้ำได้ .....	95
5.32 ผลการแบ่งภาพเอกสารที่ไม่สามารถแบ่งซ้ำได้ .....	96
5.33 ผลการแบ่งภาพเอกสารที่ไม่สามารถแบ่งซ้ำได้ .....	97
ก.1 ภาพเอกสารที่ใช้ทดสอบการระบุประเภทของหน้าต่าง .....	112

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและความสำคัญของปัญหา

ในยุคโลกาภิวัตน์การติดต่อสื่อสารเป็นแบบโลกไร้พรมแดน การแลกเปลี่ยนข้อมูลสะดวกสบายและรวดเร็วขึ้น อัตราการสื่อสารแลกเปลี่ยนข้อมูลระหว่างองค์กรต่างๆ เพิ่มมากขึ้นในทุกๆ วัน การจัดเก็บข้อมูลและบันทึกการเปลี่ยนแปลงข้อมูลเป็นจำนวนมากในแต่ละวันนับว่ามีความจำเป็น และด้วยเหตุนี้จึงส่งผลให้เกิดการพัฒนาการจัดเก็บข้อมูลที่มีประสิทธิภาพ ถูกต้องและรวดเร็ว รวมทั้งสามารถที่จะทำงานแบบอัตโนมัติได้ ซึ่งหากพิจารณาขอบเขตของการจัดเก็บข้อมูลจะพบว่าการจัดเก็บให้ถูกต้องรวดเร็วและมีประสิทธิภาพได้นั้นต้องประกอบไปด้วยการทำงานในสองส่วนหลักๆ ส่วนแรกจะต้องมีการระบุขอบเขตของภาพเอกสารในแต่ละส่วนอย่างถูกต้องรวดเร็วและมีประสิทธิภาพ และส่วนที่สองเป็นขั้นตอนของกระบวนการรู้จำที่ถูกต้องรวดเร็ว และมีประสิทธิภาพเช่นกัน ส่วนงานในส่วนที่จะนำเสนอในวิทยานิพนธ์นี้จะอยู่ในส่วนแรก จากการสำรวจพบว่ามียุทธวิธีที่ถูกนำมาใช้ในการแก้ปัญหานี้ของกรมเอกสารอัตโนมัติ โดยทั่วไปจะมี 3 วิธีหลักๆ ที่ใช้ในการแบ่งภาพเอกสาร ได้แก่

1. วิธีการบนลงล่าง (top - down) เป็นวิธีการที่ขึ้นมาบนและมีการประมวลผลที่รวดเร็วถ้ารู้รายละเอียดของเนื้อหาแผนผังของเอกสาร การทำงานเริ่มจากหน้าเอกสารทั้งหมดในระดับสูงสุดจากนั้นจะนำมาแบ่งให้เป็นขอบเขตที่เล็กลง

2. วิธีการล่างขึ้นบน (bottom - up) เป็นวิธีการที่เริ่มจากรายละเอียดเล็กๆ รวมกันขึ้นเป็นองค์ประกอบของพื้นที่ในระดับที่ใหญ่ขึ้น โดยการทำงานจะเริ่มจากคร่าวๆ เข้าเป็นขอบเขตแล้วจากนั้นจะรวมขอบเขตที่ติดกันให้เป็นโครงสร้างที่ใหญ่ขึ้นตามลำดับ ดังนั้นจึงเป็นวิธีการที่ใช้เวลาในการประมวลผลนาน แต่ความถูกต้องแม่นยำสูงและยังสามารถใช้ได้กับภาพเอกสารที่มีความหลากหลาย และ

3. วิธีการแบบผสมผสาน (hybrid) เป็นวิธีการแบบผสมผสานที่อยู่ระหว่างสองวิธีแรก โดยจะนำข้อดีข้อด้อยของแต่ละวิธีมาพิจารณาและพยายามหาทางจำกัดข้อด้อยเหล่านั้น เพื่อให้ได้ผลลัพธ์ที่เหมาะสมตรงตามความต้องการ

แต่อย่างไรก็ตามแม้งานวิจัยในการแบ่งขอบเขตของภาพเอกสารจะมีมานาน และมีหลายวิธีที่ได้รับความนิยม แต่ยังมีวิธีที่เหมาะสมที่สุดในการแบ่งภาพเอกสารที่หลากหลายไม่ได้ จึงทำให้เกิดแนวความคิดในการประยุกต์วิธีการที่นำเสนอในวิทยานิพนธ์มาช่วยในการปรับปรุงการระบุขอบเขตของภาพเอกสาร โดยเฉพาะในส่วนของการล้นที่มีระยะห่างระหว่างล้นนั้นแคบเกินไป ให้มีความถูกต้องและได้รูปทรงใกล้เคียงกับขนาดของส่วนประกอบ (วัตถุ) จริงมากยิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 1.2 วัตถุประสงค์ของการศึกษา

สืบเนื่องจากงานวิจัยเรื่อง “A Fast and Efficient method for Document Segmentation for OCR” ของ B.Kruatrachue and P. Suthaphan [1] พบว่าเมื่อมีการแบ่งส่วนประกอบภาพเอกสารได้แล้วก็ยังไม่สามารถที่จะทราบได้ส่วนประกอบที่แบ่งได้นั้นมีความถูกต้องหรือไม่ จึงต้องมีการเดินวนรอบขอบของตัวอักษรในส่วนประกอบที่ได้ในครั้งแรกเพื่อหาจุดแบ่ง (คอถัมภ์) ระหว่างส่วนประกอบอีกครั้ง ซึ่งการทำงานดังกล่าวต้องมีการเข้าถึงจุดภาพที่อยู่ภายในส่วนประกอบโดยละเอียด ส่งผลให้ต้องใช้เวลาในการประมวลผลนาน ดังนั้นวัตถุประสงค์ในการศึกษาของวิทยานิพนธ์นี้ เพื่อพัฒนาวิธีการแบ่งส่วนประกอบของภาพเอกสารไบนารีแบบอัตโนมัติให้มีความถูกต้อง รวดเร็ว และมีประสิทธิภาพ เพื่อให้ได้ขอบเขตส่วนประกอบภาพเอกสารที่ใกล้เคียงกับรูปทรงของส่วนประกอบจริงยิ่งขึ้น สามารถนำไปประยุกต์ใช้กับระบบการกรู้อักษรตัวอักษร และจัดเก็บเอกสารได้

## 1.3 สมมุติฐานของการศึกษา

ในวิทยานิพนธ์นี้จะมีเสนอวิธีการแบ่งส่วนประกอบเอกสาร และผลการทดลองที่ได้ทดสอบกับภาพเอกสารหลายๆ รูปแบบ เช่น วารสารทางวิชาการ นิตยสาร ตีตราเรียน เป็นต้น โดยผลลัพธ์ที่ได้จะเป็นขอบเขตของส่วนประกอบต่างๆ ที่มีความถูกต้องและชัดเจนในส่วนของการแบ่งส่วนประกอบยิ่งขึ้น ใช้เวลาในการประมวลผลน้อย และสามารถนำไปประยุกต์ใช้กับระบบการค้นหาและจัดเก็บเอกสารได้

## 1.4 ทฤษฎีและแนวคิดที่ใช้ในการวิจัย

วิทยานิพนธ์นี้นำเสนอแนวคิดแบบผสมผสานในการแบ่งภาพเอกสารไบนารีที่มีการจัดเก็บในรูปแบบของไฟล์บิตแมปที่มีชื่อว่า “Fast Document Segmentation using Contour and XY cut Technique” ซึ่งประยุกต์แนวคิดเริ่มต้นมาจากงานวิจัยของ B.Kruatrachue and P. Suthaphan [1] โดยเนื้อหางานวิจัยเดิมจะกล่าวถึงการใช้เทคนิคในการติดตามเส้นขอบของภาพเอกสาร (Contour following technique) ด้วยหน้าต่างขนาด  $32 \times 32$  พิกเซล และใช้ค่าของจำนวนจุดสีดำภายในหน้าต่างที่สนใจเพื่อเป็นเงื่อนไขในการติดตามขอบของภาพเอกสาร ดังนั้นการเข้าถึงจุดสีดำในหน้าต่างใดๆ จึงไม่จำเป็นต้องมีการเข้าถึงทุกๆ จุดส่งผลให้การทำงานด้วยวิธีการดังกล่าวมีความรวดเร็วในการประมวลผล อีกทั้งสามารถติดตามขอบของภาพเอกสารที่มีรูปแบบใดๆ (ไม่จำกัดเฉพาะรูปร่างแบบแมนฮัตตัน) ได้เป็นอย่างดี แต่อย่างไรก็ตามในงานวิจัยดังกล่าวยังคงมีข้อด้อยอยู่หลายประการ ประการแรก คือ ในการแบ่งเสร็จแล้วไม่สามารถที่จะระบุได้ในทันทีว่าผลการแบ่งนั้นถูกต้องหรือไม่จึงต้องมีการโปรเจกต์ เพื่อตรวจสอบความถูกต้องของการแบ่งอีกครั้งทำให้เสียเวลาในส่วนของการประมวลผลเพิ่มขึ้น ประการที่สองเกิดขึ้นเมื่อระยะห่างระหว่างโครงสร้าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของส่วนประกอบเอกสารแคบมากๆ เป็นผลให้ไม่สามารถบ่งบอกจุดผิดพลาดในการแบ่งส่วนประกอบของภาพเอกสารระหว่างคอลัมน์แคบๆ ได้ทันทีต้องมีการโปรเจ็กอีกครั้ง

วิธีการแบ่งภาพเอกสารที่มีความรวดเร็วแนะนำเสนอในงานวิจัยของ G. Nagy and S. Seth [2] และ J. Ha and R.M. Haralick [3] ซึ่งจะนำเสนอการแบ่งเอกสารเข้าตามแนวแกนทั้งแนวตั้งและแนวนอน (X-Y cut technique) โดยมากเราจะรู้จักวิธีการนี้ในชื่อ “Projection X-Y cut” หรือ “X-Y tree decomposition” การทำงานจะอาศัยการโปรเจ็กทั้งสองแกนเพื่อหาแนวทางในการแบ่งขอบเขตของภาพเอกสาร จากนั้นนำขอบเขตของภาพเอกสารที่ได้มาเก็บไว้ในโครงสร้างข้อมูลแบบต้นไม้ แต่อย่างไรก็ตามถึงแม้วิธีการดังกล่าวจะมีความเร็วแต่ก็ยังให้ผลในการแบ่งส่วนประกอบที่มีรูปทรงในลักษณะแผ่นคี่เหลี่ยมเท่านั้น ทำให้ในการประมวลผลภาพเอกสารต้องเสียเนื้อที่หน่วยความจำ โดยไม่จำเป็นไปกับส่วนของพื้นที่ว่างๆ ที่ไม่ใช่เนื้อหาสาระที่สนใจ ซึ่งหน่วยความจำนับว่าเป็นทรัพยากรที่มีค่าในการประมวลผลข้อมูลจำนวนมาก

ดังนั้นแนวคิดหลักในวิทยานิพนธ์นี้ คือ หาวิธีการในการแบ่งขอบเขตส่วนประกอบของภาพเอกสารที่มีลักษณะของโครงสร้างรายละเอียดซับซ้อน ให้มีความถูกต้องในเรื่องของรูปทรง และเพิ่มประสิทธิภาพในการทำงานที่รวดเร็วยิ่งขึ้นด้วยการระบุจุดผิดพลาดของการแบ่งโดยอัตโนมัติ ซึ่งวิธีการที่ใช้ในวิทยานิพนธ์นี้ จะใช้วิธีการแบบผสมผสานระหว่างสองวิธีการ คือ วิธีการในการติดตามเส้นขอบ และ วิธีการการแบ่งเข้าตามแนวแกน X-Y โดยได้แสดงข้อมูลจุดผิดพลาดและระบุวิธีการแบ่งให้ถูกต้องยิ่งขึ้น

## 1.5 ขอบเขตของงานวิจัย

ขอบเขตของวิทยานิพนธ์ฉบับนี้เป็นการศึกษา ปรับปรุง และพัฒนาการประยุกต์ใช้แนวความคิดในการแบ่งส่วนประกอบภาพเอกสาร ให้มีความถูกต้องรวดเร็ว และมีประสิทธิภาพ โดยภาพเอกสารที่นำมาใช้ในการทดลองต้องเป็นเอกสารใบนารีที่ไม่มีกรอบ มีการสแกนด้วยความละเอียด 300 จุด/นิ้ว ไม่เอียงหรือได้รับการแก้ไขความเอียงแล้ว

## 1.6 ขั้นตอนการศึกษา

1. ศึกษาวิธีการแบ่งส่วนประกอบและแบ่งประเภทส่วนประกอบของภาพเอกสารในวิธีต่างๆ เพื่อใช้ในการกำหนดหัวข้อ เป้าหมาย วัตถุประสงค์ และขอบเขตของการทำวิทยานิพนธ์
2. จัดเตรียมข้อมูลทดสอบ ซึ่งภาพเอกสารที่ใช้ในการทดลองเป็นภาพใบนารีที่สแกนด้วยความละเอียด 300 จุด/นิ้ว ภาพเอกสารต้องไม่เอียง ไม่มีกรอบ ซึ่งภาพเอกสารที่ใช้เป็นตัวอย่างเป็นมาจาก วารสารเชิงวิชาการ ตำราเรียน นิตยสาร ฯลฯ โดยภาพจะประกอบด้วยตัวอักษรขนาดใหญ่ไม่เกิน 100 จุด โดยมีจำนวนภาพทดสอบทั้งหมด 212 ภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3 ทดลองหาหน้าตาที่เหมาะสม โดยเริ่มจากการหาขนาดของหน้าตาที่เหมาะสมในการติดตามขอบของภาพเอกสารที่หลากหลาย ซึ่งหลักในการพิจารณาถือจะไม่ทำให้เกิดความกำกวมในส่วนของคอลัมน์ของส่วนประกอบ

4 ทำการทดลองโดยนำวิธีการที่ได้ศึกษามาทดลอง จากนั้นทดลองตามแนวทางใหม่ที่ได้อ้างไว้เพื่อพิสูจน์แนวความคิดว่าได้ผลที่ถูกต้องและตรงตามแนวทางที่วางไว้หรือไม่

5 สรุปผลการทดลองพร้อมจัดทำความตีพิมพ์และพิมพ์วิทยานิพนธ์

### 1.7 รายละเอียดเครื่องมือและอุปกรณ์ที่ใช้ในการทำวิจัย

เครื่องคอมพิวเตอร์	=	อุปกรณ์สแกนภาพ (Scanner)
คุณสมบัติทางฮาร์ดแวร์		HP Scanjet
CPU :	Pentium IV 1.4 GHz	= เครื่องมือที่ใช้พัฒนา
RAM :	384 MB	Mficrosoft VC++ 6.0
HARD DISK :	80 GB	Photoshop 7.0
ระบบปฏิบัติการ (Operating System)		UltraEdit 3.2
Windows XP		

### 1.8 รายละเอียดของวิทยานิพนธ์

ในวิทยานิพนธ์ฉบับนี้แบ่งเนื้อหาการนำเสนอออกเป็น 6 บท ซึ่งแต่ละบทมีหัวข้อและเนื้อหา ดังต่อไปนี้

- บทที่ 1 กล่าวถึงความเป็นมาและความสำคัญของปัญหา วัตถุประสงค์ สมมุติฐานของการศึกษา รวมทั้งทฤษฎีแนวคิด และขอบเขตของงานวิจัย
- บทที่ 2 กล่าวถึงทฤษฎีและแนวคิดที่เกี่ยวข้องกับงานวิจัย หลักการเดิมที่เคยตีพิมพ์หรือวิจัยไว้ก่อน จะอธิบายและวิเคราะห์การทำงานทั้งข้อดีและข้อด้อย
- บทที่ 3 กล่าวถึงปัญหาเรื่องรูปแบบส่วนประกอบ การระบุจุดผิดพลาดในการแบ่งส่วนประกอบ โดยอัตโนมัติ แนวคิดที่นำมาประยุกต์ใช้ในการแบ่งส่วนประกอบของภาพเอกสารใบนารี แนวทางที่จะนำมาประยุกต์ใช้ในเรื่องการปรับปรุงการแบ่งส่วนประกอบของภาพเอกสารให้มีความถูกต้องและได้รูปทรงที่ใกล้เคียงกับส่วนประกอบที่ต้องการยิ่งขึ้น
- บทที่ 4 กล่าวถึงวิธีการแบ่งจุดผิดพลาดตลอดจนปัญหาที่พบในแต่ละวิธีการ
- บทที่ 5 กล่าวถึงผลการทดลอง สรุป และวิเคราะห์ผลการทดลองที่ได้
- บทที่ 6 กล่าวถึงข้อเสนอแนะ และแนวทางการทำวิจัยต่อ

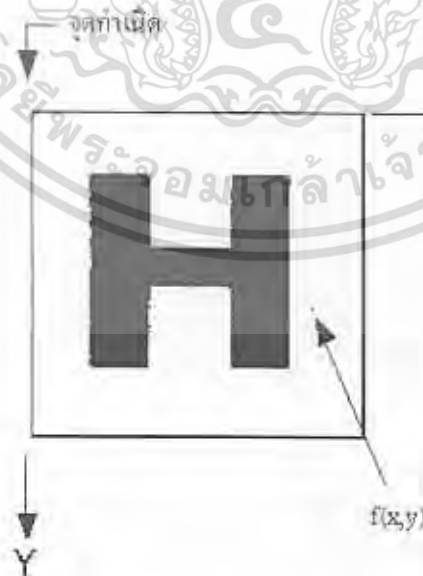
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ทฤษฎีและแนวคิดที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีและแนวคิดพื้นฐานที่เกี่ยวข้องกับงานวิจัย เริ่มต้นด้วยการทำความเข้าใจเกี่ยวกับการแสดงผลของภาพดิจิทัล ขั้นตอนพื้นฐานในการประมวลผลรูปภาพ ความรู้พื้นฐานเกี่ยวกับข้อมูลภาพอิเล็กทรอนิกส์ รหัสแนวหรือรหัสถูกใช้ซึ่งจะประยุกต์ใช้ในการติดตามขอบของภาพเอกสาร ความรู้พื้นฐานเกี่ยวกับคุณลักษณะพื้นฐานของภาพเอกสาร การจำแนกประเภทของเอกสารและการประยุกต์ใช้ ขั้นตอนการแบ่งหน้าเอกสารและวิธีการแบ่งประเภทขอบเขต การเปรียบเทียบวิธีการทำงาน รวมถึงการวิเคราะห์วิธีการทำงานของแต่ละวิธีการ และบทสรุป ตามลำดับ

### 2.1 การแสดงภาพดิจิทัล (Digital Image Representation)

โดยทั่วไปการอ้างสีในกรณีภาพสองมิติจะใช้ฟังก์ชัน  $f(x,y)$  โดย  $x$  และ  $y$  เป็นพิกัดตำแหน่งในแนวแกน  $X, Y$  ตามลำดับ ซึ่งค่าของฟังก์ชัน  $f$  ที่  $(x,y)$  จะเป็นระดับสี ดังรูปที่ 2.1 อธิบายแกนและพิกัดที่จะใช้ในวิทยานิพนธ์เล่มนี้ โดยภาพดิจิทัลสามารถนำไปพิจารณาให้อยู่ในรูปเมตริกซ์ที่ใช้แถว ( $Y$ ) และคอลัมน์ ( $X$ ) เป็นตัวบอกตำแหน่งจุดและระดับสีในภาพ ส่วนประกอบในแต่ละอาร์เรย์เรียกว่า “ส่วนประกอบภาพ (image elements)” หรือ “ส่วนประกอบรูปภาพ (picture elements)” หรือ “จุดภาพ (pixels หรือ pels)”



รูปที่ 2.1 แนวแกน  $X, Y$  ที่ใช้อ้างอิงพิกัด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.2 ขั้นตอนพื้นฐานในการประมวลผลภาพ

ขั้นตอนพื้นฐานในการประมวลผลภาพจะประกอบไปด้วย 5 ขั้นตอนหลักๆ ดังนี้

1. การทำภาพ (image acquisition) เป็นขั้นตอนนำภาพที่อยู่ในสื่ออื่นๆ ซึ่งไม่ใช่ดิจิทัล เช่น ภาพถ่าย ของจดหมาย มาแปลงให้เป็นดิจิทัลโดยผ่านอุปกรณ์ เช่น สแกนเนอร์ (scanner) กล้องดิจิทัล (digital camera) เป็นต้น

2. การปรับปรุงภาพ (preprocessing) เป็นขั้นตอนในการเพิ่มหรือแก้ไขรายละเอียดของเนื้อหาให้ชัดเจนยิ่งขึ้น เช่น การแก้ไขความเข้มของแสง (enhancing contrast) การปรับความเอียง (skew correction) และการลบสัญญาณรบกวน (removing noise) เป็นต้น

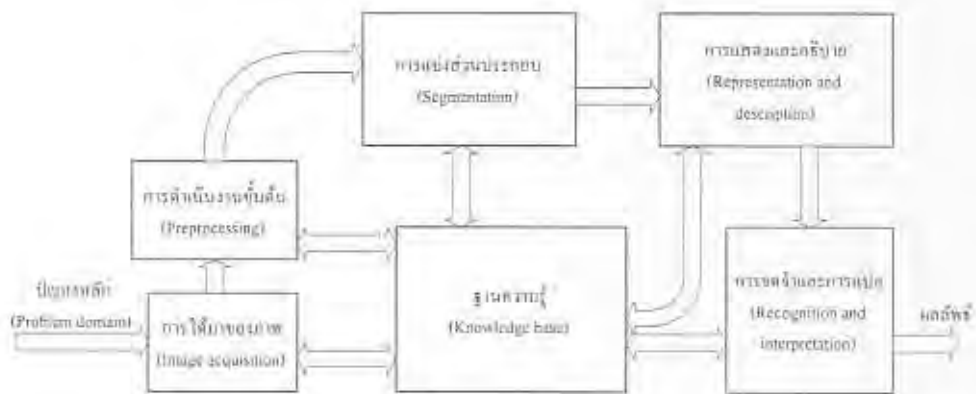
3. การแบ่งส่วนประกอบ (segmentation) เป็นขั้นตอนการแบ่งส่วนของข้อมูลภาพให้เป็นส่วนที่สัมพันธ์กัน โดยการแบ่งส่วนประกอบอัตโนมัติเป็นส่วนที่สำคัญที่สุดในงานประมวลผลภาพดิจิทัล ถ้าขั้นตอนการแบ่งภาพเอกสารสามารถทำได้ดีและเหมาะสมจะส่งผลให้ผลลัพธ์ที่ได้ดี แต่ถ้าอัลกอริทึมที่ใช้ในการแบ่งไม่ดี ก็จะทำให้ผลลัพธ์ที่ได้ไม่ดีตามไปด้วย ตัวอย่างเช่น ในกรณีการรู้จำตัวอักษร (character recognition) ต้องแบ่งแยกเฉพาะตัวอักษรและคำออกมาจากจากพื้น (background)

4. การแสดงและอธิบายผลลัพธ์ (representation and description) ผลลัพธ์ที่ได้จากขั้นตอนการแบ่ง โดยปกติจะเป็นข้อมูลรูปภาพ (raw pixel data) ส่วนประกอบ หรือเส้นขอบของพื้นที่ (constituting or boundary of region) หรือจุดทั้งหมดภายในพื้นที่ที่ถูกแบ่ง การแปลงข้อมูลให้ไปอยู่ในรูปแบบสำหรับการประมวลผลได้เป็นสิ่งจำเป็น ขั้นแรกที่ต้องดำเนินการคือ ต้องทำข้อมูลให้อยู่ในรูปขอบเขตหรือพื้นที่ที่สมบูรณ์ โดยขอบเขตแสดงได้โดยรูปร่างของรูปทรง เช่น มุม หรือเส้นขอบ พื้นที่ที่สามารถแสดงได้โดยพื้นผิว (texture) และรูปร่างของโครงสร้าง (skeletal shape) ในขั้นตอนนี้ควรจะอธิบายถึงลักษณะของข้อมูลและจุดเด่น ซึ่งจะสามารถแยกออกว่าวัตถุนี้แตกต่างจากวัตถุนั้นอย่างไร เช่น วัตถุที่เป็นชุดตัวอักษร ซึ่งจะมีการกระจายของรายละเอียดค่อนข้างสม่ำเสมอและจัดเรียงเป็นแนวตามแนวตั้งหรือแนวนอน ส่วนวัตถุที่เป็นภาพจะมีการกระจายของรายละเอียดไม่สม่ำเสมอ มีความหนาแน่นเป็นวงที่ และไม่จัดเรียงตามแนว เป็นต้น

5. การรู้จำและการแปล (recognition and interpretation) ขั้นตอนนี้เป็นการกำหนดค่า (label) ให้กับวัตถุซึ่งจะขึ้นอยู่กับรายละเอียด และการกำหนดความหมาย (meaning) ของวัตถุนั้นๆ ซึ่งได้กำหนดไว้ในฐานความรู้

ขั้นตอนทั้งหมดที่กล่าวมาข้างต้นมีความสัมพันธ์กับฐานความรู้ (knowledge base) ดังแสดงในรูปที่ 2.2 เริ่มตั้งแต่การนำปัญหาที่มีมาจัดอยู่ในฐานข้อมูลความรู้ (knowledge database) เพื่อให้ระบบประมวลผลภาพเข้าใจ ความรู้ในที่นี้หมายถึงรายละเอียดขอบเขตทั่วไปของภาพ ซึ่งความซับซ้อนของฐานความรู้ขึ้นอยู่กับเรื่องที่สนใจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.2 ขั้นตอนพื้นฐานของกระบวนการประมวลผลภาพดิจิทัล

## 2.3 ความรู้พื้นฐานเกี่ยวกับข้อมูลภาพอิเล็กทรอนิกส์

### 2.3.1 รูปแบบกลุ่มข้อมูลและระดับเฉดสี (Uniform Sampling and Quantization)

รูปแบบกลุ่มข้อมูลจะประมวลผลรูปภาพทางคอมพิวเตอร์ด้วยฟังก์ชันภาพ  $f(x,y)$  ซึ่งจะบ่งบอกถึงพิกัด (spatially) และความถี่หรือแอมพลิจูด (amplitude) โดยส่วนพิกัด  $(x,y)$  จะเรียกว่า "ข้อมูลภาพ (image sample)" ขณะที่ค่าของแอมพลิจูดเรียกว่า "ระดับเฉดสี (gray-level)" โดยข้อมูลภาพจะจัดเรียงอยู่ในรูปแบบอาร์เรย์ของเมตริกซ์ขนาด  $N \times M$  มิติ ซึ่ง  $N$  คือ จำนวนแถว  $M$  คือ จำนวนคอลัมน์ ดังแสดงในสมการ 2.1 ซึ่งแต่ละส่วนประกอบในอาร์เรย์จะถูกอ้างเป็นจุดภาพและสี

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,M-1) \\ f(1,0) & f(1,1) & \dots & f(1,M-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,M-1) \end{bmatrix} \dots (2.1)$$

### 2.3.2 ความสัมพันธ์ระหว่างจุดภาพ (Relationships between pixels)

สำหรับความสัมพันธ์ระหว่างจุดภาพที่อยู่ใกล้ชิดกันจะกำหนดโดยให้จุด  $p$  อยู่ที่พิกัด  $(x,y)$  จะมีจุดที่อยู่ใกล้  $p$  โนแนวตั้งและแนวนอนได้แก่  $(x+1,y)$ ,  $(x-1,y)$ ,  $(x,y+1)$ ,  $(x,y-1)$  ซึ่งจุดนี้เรียกว่า "ส่วนที่ติดกัน 4 ทิศของ  $p$  (4-neighbors)" หรือ  $N_4(p)$  ซึ่งแต่ละจุดจะมีระยะห่างจาก  $(x,y)$  แล้วยังมีส่วนที่ติดกันของ  $p$  ที่อยู่นอกเหนือจากนี้คือ  $(x+1,y+1)$ ,  $(x+1,y-1)$ ,  $(x-1,y+1)$ ,  $(x-1,y-1)$  ซึ่งเป็น  $N_8(p)$  และทั้งหมดเรียกว่า  $N_8(p)$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 2.3.3 การเชื่อมต่อกันระหว่างจุดจะใช้ในการหาค่าขอบเขตพื้นที่ของวัตถุและส่วนประกอบ

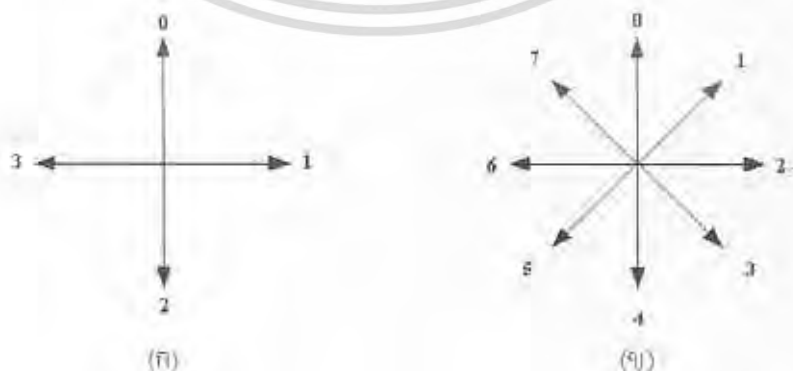
การเชื่อมต่อกันระหว่างจุดกำหนดได้ 3 ชนิดดังนี้

1. ส่วนติดกัน 4 ส่วน (4-connectivity) ของจุด  $p$  และ  $q$  กำหนดให้  $V$  เป็น 4 ส่วนที่ติดกัน ถ้า  $q$  อยู่ในเซตของ  $N_4(p)$
2. ส่วนติดกัน 8 ส่วน (8-connectivity) ของจุด  $p$  และ  $q$  กำหนดให้  $V$  เป็น 8 ส่วนที่ติดกัน ถ้า  $q$  อยู่ในเซตของ  $N_8(p)$
3. ส่วนติดกัน  $m$  ส่วน ( $m$ -connectivity) ของจุด  $p$  และ  $q$  กำหนด  $V$  เป็น  $m$  ส่วนที่ติดกัน ถ้า
  - $q$  อยู่ใน  $N_4(p)$  หรือ
  - $q$  อยู่ใน  $N_8(p)$  และ จุดของ  $N_4(p) \cap N_8(q)$  วางเปล่า

รูปภาพที่ใช้อ้างอิงในการเก็บลิ้มข้มอยู่กับสีของภาพ โดยทั่วไปแล้วภาพจะมีสีอย่างน้อย 2 สี นั่นคือการอ้างข้อมูล 1 บิตต่อ 1 พิกเซล ภาพที่ได้จะเป็นภาพขาวดำหรือไบนารี (Binary) ซึ่งมีค่า 0 และ 1 ความหลากหลายของสีจะเพิ่มขึ้นเมื่อมีการเพิ่มขนาดการเก็บข้อมูลต่อ 1 พิกเซล เช่น 8 บิตต่อ 1 พิกเซลจะมีการอ้างสีได้ถึง 256 สี คือ การอ้างถึงค่าสีจะอ้างเท่ากับ  $2^n$  เมื่อ  $n$  คือ ขนาดที่เก็บข้อมูลต่อ 1 พิกเซล

### 2.4 รหัสแนวหรือรหัสลูกโซ่ (Chain Codes)

รหัสแนวหรือรหัสลูกโซ่ถูกใช้ในการแสดงขอบเขตส่วนประกอบที่เรียงต่อกันไปในแนวเส้นตรงเดียวกัน ซึ่งจะระบุทั้งความยาวและทิศทาง โดยอยู่บนหลักเกณฑ์ 4 หรือ 8 ทิศทาง โดยที่ส่วนประกอบเรียงกันของข้อ (segment) จะระบุทิศทางแต่ละข้อเป็นรหัสตัวเลขตั้งแสดงในรูปที่ 2.3 ซึ่งขอบเขตส่วนประกอบจะถูกอธิบายในรูปแบบของพิกัด  $(x,y)$



รูปที่ 2.3 ทิศทางของรหัสแนว 4 ทิศทาง (ก) และ รหัสแนว 8 ทิศทาง (ข)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.5 คุณลักษณะพื้นฐานของภาพเอกสาร

ภาพเอกสารที่ถูกสแกนจะถูกนำมาแสดงในรูปแบบแบบไบนารี (Binary: 1 บิตต่อจุด) (เกรย์สเกล (gray-scale: 8 บิตต่อจุด) หรือสี (color: 8-24 บิตต่อจุด) ชนิดของภาพที่ใช้จะขึ้นอยู่กับวัตถุประสงค์ของการนำไปใช้และการแปลงระหว่างชนิด เช่น เปลี่ยนจากภาพสีไปเป็นเกรย์สเกลหรือไบนารี หรือเปลี่ยนจากภาพเกรย์สเกลเป็นไบนารี ซึ่งขึ้นกับความต้องการที่จะเพิ่มความเร็วในการคำนวณ ส่วนประกอบพื้นฐานที่แสดงในภาพ ได้แก่

1. ข้อความ (ตัวอักษร หรือตัวเลขในส่วนของข้อความ หัวเรื่อง ชื่อเรื่อง เซลล์ของตาราง คำอธิบายภาพ) ที่เกาะกลุ่มในภาพหรือกราฟิกเป็นคำอธิบาย
2. ตาราง (มีเส้นแบ่งและไม่มีเส้นแบ่ง)
3. สมการคณิตศาสตร์
4. รูปไบนารี รูปกราฟิก และรูปสี
5. กราฟิก (แผนผังวงวน ภาพลายเส้น กราฟ โคordinate โลกโก้ เป็นต้น)

ถ้ารับบางขอบเขตซึ่งเป็นข้อความบ่อยครั้งที่รายละเอียดของส่วนประกอบ ไม่ชัดเจนพอที่จะจำแนกประเภทไปเป็นข้อความหรือรูปภาพได้ ตัวอย่างเช่น ถ้าส่วนที่เป็นข้อความอยู่ใกล้ชิดกับรูปภาพหรือกราฟิกมากจนอาจจะถูกกำหนดให้เป็นขอบเขตรูปภาพ หรือในกรณีอื่นๆบางขอบเขตของส่วนประกอบอาจจะสามารถแบ่งให้เป็นประเภทข้อความได้ ตัวอย่างเช่น ภาพวาดลายเส้น ซึ่งหากเกิดกรณีแบบนี้มักจะจำเป็นต้องมีการประมวลผลพื้นที่เฉพาะ (Zone-specific processing) เพื่อใช้แก้ไขปัญหาขอบเขตส่วนประกอบที่ถูกแบ่งประเภทผิดพลาด

แผนผังทางภูมิศาสตร์ (Geometric document layout) และความเอียง(skew) เป็นลักษณะสองคุณลักษณะที่สำคัญของภาพเอกสาร แผนผังเอกสารเป็นแนวทางที่ใช้ในการจัดเรียงของส่วนประกอบ (ข้อความ รูปภาพ กราฟิกอื่นๆ) ของเอกสาร โดยทั่วไปสามารถที่จะแบ่งรูปแบบของแผนผังพื้นฐานได้สามรูปแบบคือ แผนผังแมนฮัตตัน (Manhattan layout) ซึ่งขอบเขตจะมีลักษณะเป็นรูปหลายเหลี่ยมที่คงที่ โดยมีขอบเขตนั้นเป็นเส้นตรง ความหนาแน่นและแนวตั้ง แผนผังสี่เหลี่ยม (rectangular layout) เป็นรูปแบบหนึ่งของแผนผังแมนฮัตตัน และรูปแบบสุดท้ายคือ แผนผังไร้หลักเกณฑ์ (arbitrary layout) จะมีขอบเขตเป็นรูปหลายเหลี่ยมที่ไม่คงที่ หรือมีพื้นที่ซ้อนทับกัน ความเอียงของเอกสารจะทำให้ภาพเอกสารที่เอียงเป็นปัญหาหลักของการกำหนดตำแหน่งหน้าเอกสาร ความเอียงจะทำให้เอกสารอยู่ไม่ถูกตำแหน่งที่ควรจะเป็น ซึ่งมักจะเกิดขึ้นแต่การสแกนภาพเอกสารเข้ามา ในหลายกรณีพบว่าจะทำให้ประสิทธิภาพในการแบ่งหน้าเอกสารและการแบ่งประเภทขอบเขตลดลง

วิธีวิเคราะห์แผนผังส่วนมากจะแบ่งขอบเขตและจัดประเภทภาพออกเป็น 3 ประเภท คือ ข้อความ กราฟิก และรูปภาพ โดยสมการคณิตศาสตร์จะถูกกำหนดเป็นประเภทข้อความ และตารางจะถูกกำหนดเป็นกราฟิก เมื่อจัดประเภทโดยใช้ข้อมูลที่จะเอียงขึ้นขั้นตอนนี้เรียกว่า “การทำเอกสารนี้เป็นเอกสารที่สแกนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความเข้าใจภาพเอกสาร” หรือ “document understanding” ซึ่งข้อความจะถูกจัดไปอยู่ในส่วนประกอบทางตรรกะ (logical layout)

## 2.6 ประเภทของเอกสารและการประยุกต์ใช้

เนื่องจากการใช้วิธีการใดวิธีการหนึ่งในการจัดประเภทของส่วนประกอบเอกสารที่มีส่วนประกอบหลากหลายและซับซ้อนจะมีความผิดพลาดค่อนข้างมาก จึงทำให้มีการเสนอวิธีการต่างๆ ที่ออกแบบมาเพื่อวิเคราะห์ภาพเอกสารที่มีส่วนประกอบภายในหลากหลายประเภท ซึ่งพบว่าส่วนมากมักจะเริ่มมาจากการแบ่งเอกสาร โดยใช้แผนผังของการจัดวางภาพเอกสารบนหน้ากระดาษ และเมื่อมีวิธีการใหม่ๆ เพิ่มขึ้นก็จะสามารถที่จะประเมินคุณภาพของวิธีการวิเคราะห์แผนผัง โดยอาศัยคุณสมบัติพื้นฐานของประเภทของส่วนประกอบเป้าหมายได้ แต่ถึงอย่างไรก็ตามยังไม่สามารถที่จะพิจารณาตามลำดับของใจความได้ หรือพิจารณาได้แต่ในประเภทที่จำกัด ในการจัดประเภทของชนิดเอกสารแบ่งตามลักษณะได้ดังนี้

1. เอกสารรูปแบบโครงสร้าง (วารสาร หนังสือพิมพ์ และจดหมายข่าว)
2. เอกสารที่มีรูปแบบไม่แน่นอน (โฆษณา ปกจีดี หนังสือและนิตยสาร)
3. เอกสารถึงรูปแบบ (ซองจดหมาย จดหมายและนามบัตร เช็กรถเช่า แบบฟอร์ม และเอกสารที่คล้ายตาราง)
4. แผนที่และภาพวาดเชิงวิศวกรรม
5. เอกสารที่มีรูปแบบเฉพาะตัว (หน้าโฮมเพจ และ รูปจากวิดีโอ)

ในส่วนนี้เป็นการสรุปภาพรวมของหลักการแบ่งส่วนประกอบ และการแบ่งประเภทส่วนประกอบที่ขึ้นอยู่กับตำแหน่งส่วนประกอบของตัวอย่างในแต่ละชนิด รวมถึงปัญหาที่เกิดขึ้นจากรูปที่ 2.4 แสดงตัวอย่างภาพเอกสารชนิดต่างๆ

### 2.6.1 เอกสารรูปแบบโครงสร้าง (structure articles)

วารสาร หนังสือพิมพ์ และจดหมายข่าว เป็นแหล่งข้อมูลข่าวสารที่หาง่ายและเป็นสิ่งพิมพ์ที่แพร่หลาย ภาพทั้งหมดที่เก็บมีความหลากหลาย (ข้อความที่มีขนาด รูปทรง และแบบที่แตกต่างกัน ความหลากหลายของกราฟิก ภาพไบนารี ภาพอาฟโทนและภาพสี) ซึ่งสามารถสแกนให้อยู่ในรูปแบบของภาพไบนารี เกรย์สเกล หรือภาพสีได้ โดยทั่วไปทางกายภาพแล้วขอบเขตจะถูกแบ่งแยกออกจากกัน แต่ถึงอย่างไรก็ตามถ้าช่องว่างระหว่างขอบเขตส่วนประกอบที่แตกต่างกันแคบมาก ซึ่งในบางกรณีอาจจะมีข้อความที่อยู่ในกราฟิกหรือรูป บางครั้งตัวหนังสือก็สีเข้ม (สิ่งพิมพ์ปกติ) หรือสีสว่าง (สีตรงข้ามกับปกติ) กว่าจากพื้น บทความในวารสาร หนังสือพิมพ์ และจดหมายข่าวเป็นตัวอย่างของเอกสารที่มีโครงสร้าง เพราะทุกส่วนประกอบบนหน้ามีการจัดเรียงลำดับตามชนิด เช่น หัวเรื่อง คำนำ และเนื้อหา โดยจะเชื่อมต่อเข้าด้วยกัน ด้วยกฎต่างๆ ไป ตัวอย่างของกฎ เช่น การอ่าน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตามลำดับของบล็อกข้อความจากบนลงล่าง และจากซ้ายไปขวา คำอธิบายภาพจะอยู่ใต้ภาพ แบบฟอร์มข้อความจะสัมพันธ์กับขนาดของพื้นที่ รูปหรือตารางจะอยู่หลังจากที่ข้อความอ้างถึง หนังสือพิมพ์จะมีความซับซ้อนมากกว่าวารสารหรือจดหมายข่าว แต่องค์ประกอบของโครงสร้างทางตรรกะจะน้อยกว่า ซึ่งเอกสารเหล่านี้สามารถแสดงในแผนผังแมนฮัตตัน และแผนผังไม่มีกฎเกณฑ์ได้ รวมถึงหน้าเอกสารของวารสารที่คล้ายตาราง แต่แผนผังนี้ใช้ได้ไม่ค่อยดีกับวารสารทั่วไป (non-technical magazines) ซึ่งงานหลักของสิ่งพิมพ์ประเภทนี้คือ การกำหนดตำแหน่งขอบเขตรวมถึงการประมาณและแก้ไขความเอียง (ซึ่งบางครั้งอาจจะเป็นตัวเลือก) การแยกเอกสารเฉพาะส่วนที่เหมือนกัน และนำไปจัดแบ่งประเภทขอบเขตเป็นฉากพื้น ข้อความ กราฟิก และรูปภาพ โดยทั่วไปบรรทัดข้อความที่ไม่มีความเอียงมีรูปแบบการจัดเรียงตามแนวตั้งและแนวนอน แต่สามารถเอียงได้ถ้า แกน X หรือแกน Y เอียง ปฏิบัติทุกขอบเขตบนภาพจะมีความเอียงเดียวกันทั้งหน้าเอกสาร แต่บางกรณีอาจมีขอบเขตที่มีทิศทาง แตกต่างจากขอบเขตอื่นๆ

ถึงแม้จะมีแผนผังโครงสร้างหลายรูปแบบแต่ปัญหาหลักที่ยังคงอยู่ คือ ความแตกต่างของชนิดภาพ เงื่อนไขการนำไปใช้ซึ่งการปรับปรุงให้ดี (ความเอียง รูปแบบแผนผังเอกสาร) และสัญญาณรบกวน

## 2.6.2 เอกสารที่มีรูปแบบไม่แน่นอน

นิยาม เอกสารที่มีรูปแบบไม่แน่นอนจะมีลักษณะคือขาดกฎทั่วไปที่จะนำมากำหนดส่วนประกอบภายในภาพเอกสาร ซึ่งโดยทั่วไปแล้วการกำหนดแผนผังของเอกสารจะขึ้นอยู่กับวัตถุประสงค์ของผู้ออกแบบ

### 2.6.2.1 โฆษณา

โฆษณาปกติจะตีพิมพ์อยู่ในนิตยสาร หรือหนังสือพิมพ์ และบนเว็บไซต์ซึ่งเป็นเอกสารอิเล็กทรอนิกส์ แผนผังของเอกสารดังกล่าวจะไม่คงที่เหมือนกับหัวเรื่องวารสารและหนังสือพิมพ์ พื้นที่ข้อความ ภาพ มักจะมีความเอียงที่แตกต่างกันได้ บรรทัดข้อความอาจจะเป็นรูปโค้ง ซึ่งตัวอักษรไม่เรียงลำดับกันเป็นแนวเส้นตรง และกลุ่มข้อความข้างในติดกับรูปภาพซึ่งมักพบได้บ่อยในการออกแบบ ในส่วนของข้อมูลสำคัญเพื่อที่จะดึงดูดสายตาของผู้อ่าน ดังนั้นโฆษณาจึงเป็นเอกสารแบบไม่เป็นโครงสร้าง [4, 5]

### 2.6.2.2 หน้าปก

ภาพหน้าปกเป็นภาพของเอกสารที่บ่อยครั้งจะใช้สีและตำแหน่งข้อความ [4] ที่หลากหลาย มีรูปแบบอักษรหลายขนาด รวมทั้งฉากพื้นมีหลายสีหรือหลายลวดลาย การวิเคราะห์แผนผังของภาพเพื่อจัดประเภทข้อความและรูปร่าง ซึ่งจะมีประโยชน์มากสำหรับการสืบค้นข้อมูล โดยปกติแล้วการสืบค้นข้อความง่ายกว่าค้นหาภาพ ซึ่งในการค้นหาภาพจะใช้สูตรอย่างใดอย่างหนึ่งในการค้นหา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เช่น รูปทรง พื้นผิว หรือสี ความท้าทายอยู่ที่แผนผังที่ไม่แน่นอนและพื้นจากมีหลายสี ซึ่งทำให้การตรวจหาตัวอักษรให้ถูกต้องทำได้ยาก



รูปที่ 2.4 ตัวอย่างภาพเอกสารชนิดต่างๆ\*

- |                   |                   |                 |             |
|-------------------|-------------------|-----------------|-------------|
| ก. สิ่งพิมพ์โฆษณา | ข. สิ่งพิมพ์โฆษณา | ค. หนังสือพิมพ์ | ง. บทความ   |
| จ. บทความ         | ฉ. หนังสือ        | ช. หน้าเว็บ     | ซ. หน้าเว็บ |

\* ที่มา Okun O, Doermann D & Pietikäinen M, "Page Segmentation and Zone Classification: the state of the art", Technical Report LAMP-TR-036(1999) , page 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 2.6.3 เอกสารกึ่งรูปแบบหรือกึ่งโครงสร้าง (Semi-structured layouts)

แผนผังกึ่งโครงสร้างมีส่วนประกอบโครงสร้างข้างในซึ่งผู้ใช้จะกรอกข้อมูลลงตรงฟิลด์ที่กำหนดให้ โดยแต่ละฟิลด์ไม่มีข้อกำหนดตายตัวว่าจะต้องวางอยู่ตรงส่วนไหนบนภาพเอกสาร หมายความว่าเอกสารสองชุดสามารถที่มีแผนผังที่แตกต่างกันทั้งๆ ที่มีส่วนประกอบข้างในเหมือนกัน

#### 2.6.3.1 ขอบจดหมาย เช็คนาการ และไปรษณียบัตร

ไปรษณียบัตรและเช็คนาการซึ่งเป็นตัวอย่างเอกสารแบบกึ่งโครงสร้าง ซึ่งเอกสารประเภทนี้สามารถกำหนดไว้ก่อนได้ว่าฟิลด์แต่ละฟิลด์เป็นบล็อกที่อยู่ รหัสไปรษณีย์ หรือผลรวมจำนวนเงิน ขอบเขตของฟิลด์นั้นๆ ที่อยู่ในเอกสารมีการจัดวางอย่างเป็นระเบียบแบบแผน นามบัตรก็ถูกรวมอยู่ในประเภทนี้ด้วยเพราะมีโครงสร้างทางตรรกะคล้ายคลึงกัน ภาพของเอกสารชนิดนี้จะเป็น ไบนารี เกรย์สเกล หรือสี และข้อความจะถูกพิมพ์บนพื้นผิวฉากที่ซับซ้อน

งานหลักที่ต้องดำเนินการของเอกสารประเภทนี้คือ การถ่วงน้ำหนักความแตกต่างของฟิลด์ตามลักษณะเฉพาะของแต่ละเอกสาร เช่น เช็คนาการ รายละเอียดนั้นจะรวมถึงฟิลด์ของลายเซ็น หมายเลขเช็ค วันที่ ผลรวมยอดเงินเป็นตัวเลข ผลรวมยอดเงินที่เป็นตัวอักษร เลขบัญชี ชื่อผู้ส่งจ่าย ที่อยู่ของสถาบันการเงิน และ/หรือตรา แต่สำหรับนามบัตรประกอบด้วย ชื่อผู้ถือ กิจการ และที่อยู่ ส่วนไปรษณียบัตรมีบล็อกที่อยู่นอกรวมที่ บาร์โค้ด (barcode) และสถานที่รับ ไปรษณีย์ ปัญหาหลักที่พบจากเอกสารประเภทนี้ ได้แก่

1. พื้นฉากที่ซับซ้อนทำให้ยากต่อการแปลงภาพเป็นภาพ ไบนารีหรือแบ่งส่วนประกอบของภาพ
2. อักษรตัวพิมพ์หรือตัวอักษรลายมือเขียนมีการสัมผัสหรือซ้อนทับกันบนเส้นบรรทัด
3. การเปลี่ยนแปลงของความสว่าง (ความเข้มของสี)
4. ตำแหน่งเอกสารที่ไม่แน่นอน มีการเคลื่อนย้ายตามการจัดวางรูปแบบของเอกสาร
5. ข้อจำกัดของเวลาที่ใช้ในการประมวลผล

#### 2.6.3.2 แบบฟอร์มและเอกสารที่เป็นตาราง

แบบฟอร์มและเอกสารที่เป็นตาราง เช่น กระดาษคำตอบ หรือใบส่งของ ที่ประกอบด้วยฟิลด์หรือเซลล์ข้างในซึ่งข้อมูลจะถูกเขียนหรือพิมพ์เข้าไป ในการวิเคราะห์ภาพเริ่มต้นปกติใช้ในรูปแบบภาพไบนารี จนถึงแบบฟอร์มสก็มี อย่างไรก็ตามถ้าจากพื้นของภาพเป็นรูปแบบเดียว ก็เป็นสิ่งที่ยากมากในการแบ่งข้อความออกมาจากแบบฟอร์ม นอกเหนือจากข้อความแล้วใบส่งของสามารถเก็บรูปเล็กๆ เช่น คราสัญลักษณ์ รูปแบบและเอกสารแบบตารางเป็นเอกสารกึ่งโครงสร้างที่มีตัวแบ่งเป็นตัวกำหนดขอบเขตข้อมูลตัวอักษรในตำแหน่งที่หลากหลาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

งานพื้นฐานที่จะใช้ในการแยกขอบเขตของฟิล์ม ขอบเขตบล็อก หรือการหาเส้นบรรทัดออก เพื่อที่จะสกัดเอาเฉพาะข้อความภายในแต่ละฟิล์ม ใบส่งของโดยปกติมีที่อยู่ของบริษัท และเลขที่บัญชีธนาคาร ซึ่งการดำเนินงานก็เหมือนกับการประมวลผลเชิงธนาคาร

การวิเคราะห์แบบฟอร์มมีอยู่สองประเภท ประเภทแรกคือรู้รูปแบบของแบบฟอร์มที่จะวิเคราะห์ กำหนดฟิล์มในฟอร์มให้อยู่รูปแบบใดรูปแบบหนึ่ง และรูปแบบ (model) ของฟิล์มที่วางหรือเซลล์ที่มียังไม่ได้ใส่ข้อมูล ในกรณีนี้งานแรกคือการกำหนดฟอร์มโดยใช้ค่าที่กำหนดหรือโครงสร้างที่สามารถถูกสกัดจากแบบซึ่งมีอยู่สองชนิดคือ ชนิดแรกเป็นแบบฟอร์มที่ถูกกำหนดเพียงครั้งเดียว เพื่อที่จะกำหนดพิกัดขอบเขตของฟิล์มข้อมูลที่ต้องการ ประเภทที่สองเป็นการวิเคราะห์แบบฟอร์มที่ไม่รู้เกี่ยวกับรูปแบบของแบบฟอร์มที่ใช้ เมื่อไม่รู้เกี่ยวกับรูปแบบของแบบฟอร์มที่ต้องใช้มาก่อน จึงต้องอาศัยอัลกอริทึมที่หาพิกัดที่สกัดแยกพื้นฐาน โครงสร้างแบบฟอร์มเป็นพื้นฐานของตัวแบ่งระหว่างเซลล์หรือฟิล์มของชนิดที่สองนี้จะคิดแปลงให้ยืดหยุ่น แต่จะใช้การคำนวณที่ซับซ้อนมาก และบางทีจึงได้ความถูกต้องน้อย

ปัญหาจะเพิ่มมากขึ้นเมื่อข้อมูลหรือข้อความถูกเขียนนอกขอบเขตฟิล์มที่กำหนดไว้ ซึ่งเป็นการยากที่จะกำหนดขอบเขตของข้อมูล ได้ถูกตั้ง การแปลงภาพ (การหมุน การย่อขยาย และการเปลี่ยนแปลงรูปร่าง) ดังนั้นจึงมีความยุ่งยากที่ต้องดำเนินการจับคู่รูปแบบที่ทราบล่วงหน้ากับภาพข้อมูลเข้า สำหรับใบส่งของที่ส่งทางเครื่องแฟกซ์จากผู้ขายถึงลูกค้า ลักษณะของภาพจะมีคุณภาพของรายละเอียดต่ำลงเพราะอาจเกิดการสูญเสียข้อมูลเนื่องจากการส่งแฟกซ์ และการประทับตราลงบนพื้นที่ที่เป็นข้อมูลเป็นอีกสาเหตุหนึ่งที่ทำให้เกิดผลกระทบที่ไม่ดีต่อการสกัดข้อมูลหรือข้อความของเอกสาร

## 2.6.4 แผนที่และภาพวาด

### 2.6.4.1 แผนที่

ภาพแผนที่ปกติจะเป็นภาพเกรย์สเกลหรือภาพสีที่ใช้บ่งบอกรายละเอียดเกี่ยวกับแผนที่ แผนที่นี้จะประกอบด้วยระดับสีหลายระดับ แต่ละระดับสีจะตรงกับประเภทของข้อมูลที่กำหนดไว้ ข้อมูลที่เป็นข้อความ เช่น ชื่อเมือง ประเทศ ทะเลสาบ และอื่นๆ หรือรายละเอียดที่กำหนดจะอยู่ในลักษณะกราฟิก เช่น ถนน ความแตกต่างของพื้นผิว และอื่นๆ

ข้อมูลที่เป็นข้อความบนแผนที่หรือภาพวาดที่มีขนาดเล็กและมีจำนวนน้อย อาจจะถูกรวมเข้าไปอยู่ในส่วนที่เป็นกราฟิก ดังนั้นจึงมีบางกฎที่ใช้สำหรับแยกส่วนที่ถูกรวมเข้าไป เพื่อที่จะสามารถแยกได้ถูกต้องว่าส่วนใดเป็นคิ้วหนังสือหรือขอบเขตที่แคบของแผนที่ เพื่อผลลัพธ์ที่ชัดเจนไม่คลุมเครือ และเพื่อกำหนดให้เป็นส่วนประกอบที่อยู่ใน โครงสร้างที่กำหนด ปกติข้อความในแผนที่จะสามารถปรากฏได้ทุกแห่งในแผนที่ และแนวบรรทัดของข้อความมีทิศทางไม่แน่นอน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถึงแม้แผนที่จะมีมากมายหลายประเภท เช่น เกี่ยวกับที่ดิน อุทกศาสตร์ ภูมิประเทศ อื่นๆ แต่ในการวิเคราะห์ภาพแผนที่สามารถแบ่งประเภทของงานได้ 3 ประเภทหลักๆ คือ

1. การแบ่งข้อความและกราฟิก
2. การกำหนดส่วนที่เป็นตัวอักษรและสัญลักษณ์พิเศษ
3. การกำหนดทิศทางของกราฟิก

สำหรับภาพสี่เหลี่ยมจะดำเนินการตั้งแต่การสแกนจากภาพที่ละสีโดยอาศัยสแกนเนอร์ที่ออกแบบมาเพื่อใช้ในงานนี้โดยเฉพาะ ตัวอักษรและสัญลักษณ์จะถูกแบ่งแยกออกจากกราฟิกที่ในแต่ละระดับชั้นของสีโดยใช้ขนาดของตัวอักษรและสัญลักษณ์ โดยข้อมูลทั้งหมดจะถูกส่งไปยังกระบวนการรู้จำ ขณะที่กราฟิกจะถูกทำให้บางลงและกำหนดขนาดและทิศทางของกราฟิก

เนื่องจากแผนที่มีรูปแบบที่หลากหลายจึงก่อให้เกิดวิธีการและระบบที่ใช้งานมากมาย โดยแต่ละวิธีอาจจะถูกนำไปประมวล โดยเฉพาะกับแผนที่ชนิดใดชนิดหนึ่งเท่านั้น ปัญหาโดยรวมคือตัวอักษรและสัญลักษณ์สัมผัสหรือซ้อนทับกราฟิก ข้อความเอียงหรือโค้ง ระยะห่างระหว่างตัวอักษรที่ใกล้เคียงและในข้อความเดียวกันมีระยะห่างที่ไม่แน่นอน ความแตกต่างของขนาดตัวอักษร (แม้ว่าจะล่วงรู้แบบอักษรที่ใช้ในการสร้างแผนที่) และสัญลักษณ์กราฟิกบางตัวที่คล้ายกับตัวอักษร

#### 2.6.4.2 ภาพวาดเชิงวิศวกรรม

ภาพวาดเชิงวิศวกรรมมีความท้าทายในการวิเคราะห์พอๆ กับแผนที่ ซึ่งในเอกสารของทั้งสองประเภทเก็บข้อมูลตัวอักษรขนาดเล็กและอยู่อย่างกระจัดกระจายไปทั่วพื้นที่ซึ่งกราฟิกมีขนาดใหญ่ อย่างไรก็ตามภาพวาดทางวิศวกรรมโดยทั่วไปเป็นภาพแบนราบหรือเกรย์สเกลที่มีรายละเอียดที่อยู่ห่างกัน และมีภูมิศาสตร์แบบปิด งานหลักของการวิเคราะห์แผนผังทางภูมิศาสตร์ของภาพวาดเชิงวิศวกรรม คือ ความต้องการที่จะแบ่งตัวอักษรหรือกราฟิก และแบ่งประเภทส่วนประกอบของกราฟิกเช่น เส้น (เส้นประ หรือเส้นจุด) เส้นโค้ง รูปทรง เป็นต้น

การแบ่งชนิดของภาพวาดเชิงวิศวกรรมส่วนมากจะมีวิธีการเฉพาะที่สามารถนำมาใช้ได้กับส่วนประกอบเพียงไม่กี่ชนิดที่แตกต่างกัน ปัญหาที่เป็นปัญหาหลักของเอกสารประเภทนี้คือข้อความที่ติดกับกราฟิก หรือ เอกสารต้นฉบับมีคุณภาพแย่

#### 2.6.5 ภาพหน้าเว็บ และภาพจากวิดีโอ

ภาพหน้าเว็บและภาพจากวิดีโอที่มีรายละเอียดเป็นอักษรถูกจัดให้เป็นเอกสารประเภทใหม่ คือ เอกสารอิเล็กทรอนิกส์ ซึ่งจะมีคุณลักษณะไม่เหมือนกับเอกสารอื่นๆ ที่อธิบายมาก่อนหน้านี้แล้ว โดยทั่วไปเอกสารจะเป็นหน้าสีและภาพจะมีความละเอียดต่ำ มีขนาดเล็ก ซึ่งจะทำให้การเรียกดูภาพทำได้รวดเร็ว โดยรูปภาพจะช่วยให้เกิดความเข้าใจที่ชัดเจนยิ่งขึ้น ส่วนประกอบหลักของเอกสารมีทั้งข้อความและรูปภาพที่มีขอบเขตและตำแหน่งไม่แน่นอน ตัวอักษรในหน้าเว็บที่เห็นเอกสารนี้เป็นเอกสารที่สแกนไวสำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อาจเป็นได้ทั้งตัวอักษรในรูปแบบตัวอักษรและตัวอักษรและรูปแบบรูปภาพ(Bitmap) อย่างใดอย่างหนึ่ง สำหรับภาพจากวีดิโอนั้นส่วนมากข้อความจะอยู่ในส่วนของภาพที่ค่อนข้างชัดชัด สิ่งที่น่าประหลาดอาจจะเป็นจากหรือข้อความกราฟิกอย่างใดอย่างหนึ่ง และความละเอียดของภาพที่น้อยกว่า 200-300 จุดต่อนิ้ว สำหรับเอกสารชนิดนี้จะทำให้การตรวจสอบและรู้จำส่วนประกอบทำได้ยาก

ภาพหน้าเว็บและภาพจากวีดิโอจัดอยู่ในประเภทเอกสารที่ไม่มีโครงสร้าง เพราะแผนผังของหน้าเว็บและภาพจากวีดิโอไม่สามารถจำกัดขอบเขตโดยการกำหนดคณไว้ก่อนได้ และในกรณีของภาพจากวีดิโอบ่อยครั้งที่ไม่สามารถกำหนดคณใดๆ ได้ ดังนั้นงานหลักของเอกสารประเภทนี้คือการกำหนดขอบเขตของข้อความที่อยู่ในภาพ ซึ่งมีวัตถุประสงค์เพื่อใช้ในการสืบค้นหาข้อมูลและทำดัชนีวีดิโอ ปัญหาหลักของเอกสารประเภทนี้ ได้แก่

1. สีของข้อความและฉากพื้นจะใกล้เคียงกันเมื่อความละเอียดของภาพต่ำ หรือข้อความที่มีความเข้มน้อยเมื่อเทียบกับฉากพื้น(ตัวอย่างข้อความที่อยู่บนหลอดฉายฉากพื้นที่ชัดชัด) ดังนั้นการทำเรโซ (Threshold) ไม่สามารถแบ่งข้อความและฉากพื้นได้
2. ส่วนประกอบข้อความอาจจะมีหลายสีหรือหลายเฉดสีหรือในกรณีที่ยากที่สุดคือไม่มีการจัดเรียงกันเป็นแนว (บรรทัดที่มีลักษณะเป็นคลื่น)
3. ส่วนประกอบข้อความที่ไม่ได้เป็นสีเดียวกัน และมีการขาดตอนไปเป็นส่วนย่อยหลายส่วน (ไม่ใช่เพราะสัญลักษณ์) ซึ่งมีสาเหตุอันเนื่องมาจากการออกแบบตัวอักษร
4. ผลกระทบที่เกิดจากการเคลื่อนไหวของภาพในวีดิโอ

## 2.7 ขั้นตอนการแบ่งหน้าเอกสารและวิธีการแบ่งประเภทขอบเขต

ในส่วนนี้แสดงขั้นตอนการแบ่งหน้าเอกสาร(หรือส่วนประกอบของเอกสาร)และวิธีการแบ่งประเภทขอบเขตของส่วนประกอบ โดยประเภทของเอกสารนั้นได้กล่าวมาแล้วในข้อที่ 2.6 ซึ่งได้แก่ เอกสารรูปแบบโครงสร้าง (วารสาร หนังสือพิมพ์ และจดหมายข่าว) เอกสารที่มีรูปแบบไม่แน่นอน (โฆษณา ปกซีดี หนังสือ และ วารสาร) และเอกสารที่ไม่มีรูปแบบใหม่ๆ เช่น ภาพหน้าเว็บ เป็นต้น โดยทั่วไปเอกสารจะมีการอธิบายรายละเอียดเป็นหน้า (ในรูปแบบกระดาษหรืออิเล็กทรอนิกส์) และประกอบด้วยขอบเขตรายละเอียดที่ต่างกัน ซึ่งต่างจากภาพวาดลายเส้น แผนที่ เชื้อธนาคร และแบบฟอร์ม เพราะโครงสร้างขอบเขตที่เหมือนกัน โดยทั้งหมดมีส่วนประกอบที่ผสมกันทั้งที่เป็นข้อความและกราฟิก และบ่อยครั้งที่ข้อความกับกราฟิกชิดกันหรือทับกัน

ในที่นี้การแบ่งส่วนประกอบหมายถึงการแบ่งภาพออกเป็นส่วนๆ โดยในแต่ละส่วนต้องเป็นชนิดเดียวกัน หรือพื้นที่แต่ละขอบเขตต้องเป็นข้อมูลเพียงชนิดเดียวเท่านั้น เช่น ข้อความ หรือกราฟิก และจะมีการจัดประเภทส่วนประกอบหลังจากแบ่งขอบเขตส่วนประกอบเอกสารแล้ว ดังนั้นการจัดแบ่งประเภทส่วนประกอบจะกำหนดประเภทได้จากพื้นฐานที่เป็นจุดเด่น ซึ่งหลายกรณีวิธีนั้นจะดำเนินการแบ่งส่วนประกอบและแบ่งประเภทของส่วนประกอบ โดยแต่ละวิธีการจะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กล่าวย่อๆ ไม่เน้นรายละเอียดของการนำไปใช้ จากนั้นจะยกตัวอย่างบางวิธีที่น่าสนใจในหัวข้อถัดไป

## 2.7.1 การแบ่งส่วนประกอบเอกสาร

การแบ่งหมวดหมู่ของวิธีการที่ใช้ในการแบ่งส่วนประกอบเอกสารสามารถจำแนกตามแนวคิดได้ 3 วิธีหลักๆ คือ วิธีการล่างขึ้นบน (bottom-up data-driven) วิธีการบนลงล่าง (top-down or model-driven) และ วิธีการผสมผสาน (hybrid) ซึ่งเป็นการผสมระหว่างวิธีการบนลงล่างและวิธีการล่างขึ้นบน) วิธีการบนลงล่าง ใช้มานานและมีการทำงานที่รวดเร็ว เมื่อรู้รายละเอียดของลำดับแผนผังเอกสาร ในกรณีนี้กระบวนการทำงานจะเริ่มทำจากหน้าเอกสารทั้งหมดที่อยู่ในระดับสูงสุดนำมาแบ่งทั้งหมดให้เป็นขอบเขตที่เล็กลงเช่น บล็อก บรรทัด คำ และตัวอักษร ส่วนวิธีการล่างขึ้นบนจะเริ่มจากจุดและรวมกลุ่มเข้าด้วยกันเป็นขอบเขต แล้วรวมขอบเขตที่ใกล้เคียงกันเข้าเป็นโครงสร้างที่ใหญ่ขึ้น ซึ่งวิธีการนี้มีความแม่นยำสูงสามารถใช้ได้กับเอกสารที่มีรูปแบบที่หลากหลาย แต่เวลาในการประมวลผลส่วนมากจะใช้เวลานาน วิธีการแบบผสมผสานเป็นวิธีที่อยู่ระหว่างกลางของวิธีการบนลงล่างและวิธีการล่างขึ้นบน โดยพยายามที่จะรวมความสามารถในการทำงานด้วยความรวดเร็วของวิธีการบนลงล่างและความถูกต้องแม่นยำในการทำงานของวิธีล่างขึ้นบน และจัดแบ่งประเภทของส่วนประกอบไว้เป็นข้อความหรือกราฟิก โดยวิธีการเหล่านี้สามารถแบ่งขอบเขตเอกสาร โดยอาศัยสวดลายพื้นที่แตกต่างกันเป็นข้อกำหนดในการแบ่งขอบเขต ซึ่งในงานวิจัยของ O. Okun และ M. Pietikäinen [6] ได้สรุปวิธีการใช้พื้นที่แบ่งพื้นที่โดยอาศัยสวดลาย

### 2.7.1.1 วิธีการล่างขึ้นบน

วิธีการจากล่างขึ้นบนจะใช้พื้นฐานของการวิเคราะห์ส่วนประกอบของภาพไบนารี นั่นคือรวมจำนวนจุดสีดำ ซึ่งเป็นข้อมูลของจุดภาพและรวมจุดภาพเหล่านั้นเข้าเป็นตัวอักษร รวมตัวอักษรเข้าเป็นบรรทัด และรวมบรรทัดย่อยๆ เข้าเป็นบล็อก (ย่อหน้า/ส่วนประกอบ) โดยใช้ความใกล้ชิดระหว่างส่วนประกอบที่ใกล้กันและขนาดของส่วนประกอบ การทำแบบลอเพื่อเชื่อมส่วนประกอบที่ใกล้ชิดกันเข้าด้วยกัน สำหรับวิธีการที่รู้จักกันโดยทั่วไป คือ โคอะแกรมไวโรนอย (Voronoi diagrams) ซึ่งเป็นวิธีการหลักของการจับกลุ่ม โดยระยะเวลาที่ใช้ในการประมวลผลในแต่ละภาพจะแตกต่างกันขึ้นอยู่กับวิธีการที่เลือกใช้ วิธีการในแบบล่างขึ้นบนนี้ส่วนมากจะสามารถรองรับการทำงานกับ (tolerant) เอกสารที่มีความเอียง (ซึ่งบางทีอาจจะมีหลายความเอียง คือ ส่วนประกอบในภาพแต่ละส่วนเอียงไปคนละทิศทาง) และเอกสารที่มีแผนผังไม่แน่นอนได้ โดยความเอียงจะถูกคำนวณจากผลรวมของบรรทัดข้อความที่แยกออกมา การเลือกตัวอย่างข้อมูลเป็นหลักสำคัญสำหรับวิธีการในกลุ่มนี้ ทุกขอบเขตจะถูกแสดงในโครงสร้างข้อมูลแบบต้นไม้แบบขึ้นลำดับ

สำหรับประมวลผลเอกสารที่มีความหลากหลายที่เสนอโดย S.Y. Wang และ T. Yagasaki [7]

เอกสารนี้เป็นเอกสารที่สแกนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สามารถใช้กับแบบฟอร์ม หรือ วารสาร ซึ่งตัวอักษรประกอบด้วยตัวภาษาอังกฤษ และตัวคันจิของญี่ปุ่น บล็อกที่ใกล้ชิดกันใช้เวลาในการประมวลผลที่รวดเร็วในการวิเคราะห์คุณสมบัติของบล็อก และการอ้างอิงโครงสร้างข้อมูลแบบต้นไม้

วิธีการในการวิเคราะห์ส่วนที่ติดกันของบล็อกสี่เหลี่ยมก่อนที่จะรวมส่วนประกอบประเภทเดียวกันเข้าด้วยกัน โดยใช้หน้าปกเล็กๆ จำนวนมากที่อยู่ในขอบเขตใหญ่ในการเชื่อมต่อบล็อก บล็อก สามารถลดเวลาที่ใช้ในการทำงาน ได้อย่างมากเมื่อเปรียบเทียบกับการเชื่อมต่อระหว่างจุด [8]

วิธีการที่นำเสนอโดย T. Pavlidis และ H. Zhou [9] คือ การสกัดแยกข้อมูลโดยการคำนวณค่าของจุดภายในหน้าต่าง  $5 \times 5$  พิกเซล ที่แต่ละจุดจะถูกสแกนภาพ โดยแถวต่อแถว วิธีการแบ่งส่วนที่อยู่ใกล้กัน โดยใช้บล็อกขนาดสี่เหลี่ยมรวมจุดเข้าเป็นพื้นที่เดียวกัน ประโยชน์ของวิธีการนี้สามารถทำได้ทั้งภาพไบนารีและเกรย์สเกล แต่ไม่รองรับกับเอกสารที่เอียงหรือมีรูปร่างแบบไม่แน่นอน

แนวคิดพื้นฐานของการวิเคราะห์ฉากพื้นของภาพไบนารีถูกนำมาเสนอใน [10-13] วิธีการที่ใช้ได้ประโยชน์ที่สุด คือวิธีการที่เรียกว่า กระเบื้องขาว (white tiles) [10, 11] ในการทำงานขั้นแรก จะสร้างโครงข่ายของสี่เหลี่ยม ซึ่งแต่ละอันจะกว้างเท่ากับความกว้างของช่องว่างสีขาว (ฉากพื้น) และเป็นสี่เหลี่ยมขนาดดังกล่าวที่อยู่ติดกันจนรอบเพื่อกำหนดขอบเขต วิธีการนี้เป็นการแสดงที่ชัดเจน ที่สามารถแสดงได้ทั้งแผนผังรูปร่างแบบแน่นอนชัดและรูปร่างแบบที่ไม่แน่นอน โดยไม่มีการแตกขอบเขตของรูปทรงที่ซับซ้อน วิธีนี้ไม่ต้องแก้ไขความเอียงของภาพเอกสารก่อน และสามารถที่จะกำหนดขอบเขตพื้นที่ข้อความที่เอียงในตำแหน่งที่แตกต่างกันได้ วิธีการนี้ทำงานได้รวดเร็ว เพราะเป็นการประมวลผลกับจุดภาพเพียงครั้งเดียว

การคำนวณงานเกี่ยวกับรูปร่างลักษณะแบบเปิดหรือปิด ถูกนำมาใช้ใน [14-16] เพื่อจัดกลุ่มจุดภาพรวมเข้าไบนารีในขอบเขตใหญ่ก่อนจัดประเภท การลบรอยภาพโดยใช้วิธีการรันเรจก์สมูท-อัลกอริทึม (Run Length Smoothing (RLS) ที่ใช้ใน [17-21] การคำนวณงานนี้คล้ายกับการขยายรูปร่างของภาพ บางวิธีในกลุ่มนี้ใช้ได้กับเอกสารที่มีความเอียงและแผนผังที่มีรูปร่างแบบไม่แน่นอน [15, 18, 22] ในขณะที่วิธีการอื่น [16, 17, 19, 20] ไม่สามารถที่จะใช้ได้

การสกัดข้อความจากภาพเอกสารสี่ที่ซับซ้อนใน [23-28] ซึ่งในหลายๆ วิธีกำหนดให้ตัวอักษรมีสีเดียวกันทั้งหมดและตัวอักษรจัดเรียงตามแนวนอน โดยที่ตัวอักษรและฉากพื้นต้องแยกออกจากกันได้ วิธีการที่เสนอโดย W.Y. Chen และ S.Y. Chen [23] ใช้การประมวลผลหน้าปกวารสารเชิงเทคนิค ซึ่งประกอบด้วยการลดคุณภาพสีโดยลดจำนวนสีลงตามการแบ่งขอบเขตและสีพื้น ส่วนวิธีการที่ A.K. Jain และ B. Yu [24] สามารถสกัดข้อความโดยการกำหนดค่าต่างๆ เพื่อนำมาทำการประมวลผลภาพ วิธีการนี้สามารถนำไปใช้กับภาพหน้าเว็บ หน้าปกซีดี และปกหนังสือ หรือ ภาพจากวีดีโอ ภาพหลายค่าซึ่งอาจจะเป็นภาพไบนารี (โฆษณา) ภาพเกรย์สเกล ภาพสี (ภาพเว็บในรูปแบบ GIF) หรือ ภาพสี่สี (วีดีโอ ปกหนังสือ หรือ ปกซีดี) ซึ่งจะถูกแบ่งย่อยไปเป็นส่วน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดที่แสดงด้านหน้าและฉากพื้นที่ถูกซับซ้อน ดังนั้นการวิเคราะห์ที่ส่วนประกอบที่ใกล้ชิดกันสามารถที่จะใช้กำหนดส่วนประกอบที่เป็นข้อความได้

วิธีการของ S. Messelodi และ C.M. Modena [25] มีข้อดีเหนือวิธีการอื่นในหลายประการด้วยกัน โดยวิธีการนี้ไม่ขึ้นกับความเอียงของบรรทัดข้อความ และสามารถที่จะดำเนินการรวมตัวอักษรเข้าด้วยกันอย่างถูกต้อง รวมถึงการรวมบรรทัดข้อความที่ใกล้ชิดกันเข้าด้วยกันได้อย่างถูกต้องด้วยเช่นกัน โดยขั้นตอนแรกจะตรวจสอบส่วนประกอบที่ใกล้ชิดกันในภาพไบนารีหนึ่งหรือหลายภาพ จำนวนของภาพขึ้นอยู่กับชนิดของภาพ (ไบนารี เกรย์สเกล หรือ สี) ตัวอย่าง สำหรับภาพไบนารีส่วนประกอบที่ใกล้ชิดกันจะถูกตรวจสอบบนสองภาพที่จะให้ผลบวกและผลลบของความเข้มของข้อความที่สอดคล้องกับฉากพื้น บรรทัดข้อความจะถูกสกัดโดยใช้แบ่งแยกความแตกต่างเป็นลำดับขั้น ในการแก้ปัญหาภาพเกรย์สเกลของปกหนังสือเป็นตัวอย่งที่แสดงการทำงานของวิธีการนี้ ดังนั้นจะเห็นว่าวิธีการนี้สามารถจัดการกับประเภทเอกสารอื่นได้ดีเช่นเดียวกัน

การวิเคราะห์ภาพหน้าเว็บของ T. Park, D. Kim และ K. Chung [26] ในตอนแรกทำการลดจำนวนของสี ปริมาณของสีจะตรวจสอบได้จากส่วนประกอบของภาพเอกสาร ในบริเวณที่ติดกันสำหรับแต่ละสีที่คงเหลือจากการลดจำนวนสี หลังจากนั้นนำตัวอักษรของส่วนประกอบที่สัมพันธ์กันในส่วนที่สนใจมาทำให้บังลง ตลอดจนคำนวณทั้งความกว้างและความสูง ตัวอักษรเป็นชุดแบบเส้นร่างที่ต่อกัน โดยมีความสัมพันธ์กับความกว้างที่คำนวณและอัตราส่วนระหว่างความกว้างและความสูงของอักษร เส้นร่างชุดข้อมูลจะพบในขอบเขตต่อเนื่องตัวอักษรแต่ละจะถูกใช้เป็นข้อมูลเข้าเพื่อที่จะรวมขอบเขตของกลุ่มตัวอักษรที่ใกล้ชิดกันเข้าเป็นบรรทัด ซึ่งเหมือนกันกับวิธีการของ S. Messelodi และ C.M. Modena [25] โดยวิธีการนี้สามารถที่จะตรวจหาทิศทางที่ไม่แน่นอนของข้อความตลอดจนบรรทัดข้อความที่เป็นแนวโค้งได้

### 2.7.1.2 วิธีบนลงล่าง

จุดเด่นของวิธีการบนลงล่างจะเป็นการดำเนินงานจากส่วนประกอบใหญ่ไปสู่ส่วนประกอบย่อย ซึ่งวิธีการในกลุ่มนี้ยังมีแนวความคิดที่หลากหลายดังนี้

- เริ่มต้นจากการประมวลภาพทั้งหมดโดยการแยกไปเป็นบล็อกเล็กๆ [29-30]
- การแปลงส่วนประกอบใหญ่โดยการใส่ภาพรวมทั้งหมดและอาศัยจุดภาพที่เปลี่ยนแปลงแล้วรวมจุดภาพเหล่านั้นเข้าด้วยกันเป็นกลุ่ม [31]
- การดำเนินงานสามารถเริ่มทำจากตัวแทนภาพอย่างหยาบหรือขอบเขตใหญ่ไปยังตัวแทนภาพแบบละเอียดหรือขอบเขตเล็กๆ [14, 32] โดยขอบเขตใหญ่จะถูกประมวลผลได้รวดเร็วในตัวแทนอย่างหยาบ และจะทำละเอียดอีกทีในตัวแทนที่ละเอียด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การวิเคราะห์โดยการฉายภาพเงาเป็นการนับจำนวนจุดไปตามแนวทิศทางใดทิศทางหนึ่งเป็นวิธีการแบบบนลงล่างที่รู้จักกันดีวิธีหนึ่ง ซึ่งใช้ได้กับภาพไบนารีที่ไม่มีความเอียง แล้วทำการแบ่งภาพซ้ำๆ จนมีขอบเขตที่เล็กลง โดยใช้พื้นฐานจุดขาว (ตำแหน่งที่มีจำนวนจุดสีดำน้อยที่สุด) ในการฉายภาพเงาในแนวตั้งและแนวนอน ซึ่งจะสอดคล้องกับตัวแบ่งระหว่างบล็อกสี่เหลี่ยม การฉายเงาจุดถูกนำมาใช้ประโยชน์โดย T.Akiyama และ N. Hagita [29] ขณะที่รูปร่างของบล็อกขอบเขตจะได้จากงานวิจัยของ J. Ha, R.M. Haralick และ I.T. Phillips [3] โดยมีจุดประสงค์หลักเพื่อเพิ่มความเร็วในการทำงาน

การเปลี่ยนแปลงภาพโดยใช้การรบกวนฟูวเคอร์ (Gabor filters) กับตำแหน่งที่หลากหลายในภาพนำเสนอโดย A.K. Jain และ S. Bhattacharjee [31] ซึ่งใช้ได้กับภาพเกรย์สเกลที่แบ่งจุดจากการเปลี่ยนแปลงภาพกับอัลกอริทึมสแควร์เออเรอร์ (squared-error) เพื่อตรวจจับขอบเขตของตัวอักษรและส่วนที่ไม่ใช่ตัวอักษร (ฉากพื้น รูป) วิธีการนี้ใช้ได้กับเอกสารที่มีความเอียงแต่ใช้เวลาในการทำงานมาก

การแบ่งเอกสารภาพเกรย์สเกลแบบพีระมิดสี่ชั้นของเกาส์เซียน (Gaussian) ทั้งสี่ระดับได้ถูกนำมาใช้โดย L. Cinque, L. Lombardi และ G. Mazini [14] ซึ่งจะใช้พีระมิดในการแสดงจุดเด่นทั้งสี่ของแผนที่ จุดเด่นเหล่านั้นประกอบไปด้วย ค่าเฉลี่ย ค่าความแปรปรวน ค่าคงที่ ค่ามัธยฐาน การทำงานจะทำจากภาพรายละเอียดระดับต่ำไปยังรายละเอียดที่สูงกว่า วิธีการนี้ไม่ขึ้นกับความเอียงและรูปแบบของเอกสาร

การตรวจจับพื้นที่ข้อความบนพื้นผิวของฉากพื้นถูกนำเสนอโดย J.L. Chen [32] เป็นแนวคิดพื้นฐานที่ใช้การสกัดรายละเอียดของข้อความด้วยจุดเด่นของกฎหน้ากาก (Laws' masks) การแบ่งประเภทแบบหลายๆ โดยใช้บล็อกขนาด  $8 \times 8$  พิกเซล โดยแต่ละบล็อกมีการจัดเรียงที่ไม่ซ้อนทับกัน สามารถแบ่งประเภทของบล็อกได้เป็น ข้อความ ฉากพื้น และส่วนยังไม่ตัดสินใจ (fuzzy คือ ความน่าจะเป็นของขอบเขตระหว่างข้อความหรือฉากพื้น) และการแบ่งประเภทบล็อกของส่วนที่ยังไม่ตัดสินใจจะใช้การตรวจสอบในระดับจุดภาพ ทั้งการแบ่งแบบหลายและแบบละเอียดให้หลักการของ HMMs (Hidden Markov Models) ที่ใช้ 4 ถึง 8 สถานะ (states) ประโยชน์ของ HMMs ซึ่งข้อดีของ HMMs ที่เหนือกว่ากระบวนการเรียนรู้ของนิเวศอนเน็ตเวิร์กอื่นๆ คือแต่ละโมเดลสามารถเรียนรู้ได้อย่างอิสระจากโมเดลอื่นๆ ซึ่งเมื่อเพิ่มข้อมูลประเภทใหม่ๆ เข้าไป HMMs จะถูกสร้างขึ้นและเรียนรู้ เฉพาะตัวอย่างของประเภทนั้น (ส่วน HMMs ตัวอื่นๆ ไม่มีการเรียนรู้ใหม่) วิธีการนี้ไม่ไว (sensitive) ต่อข้อความและรูปแบบเอกสารชนิดตัวร่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 2.7.1.3 วิธีแบบผสมผสานและวิธีอื่นๆ

D. Wang และ S.N. Srihari [33] นำอัลกอริทึม RLS ของแนวคิดวิธีการล่างขึ้นบน มาใช้กับภาพไบนารีเพื่อตรวจหาข้อมูลบรรทัดข้อความและข้อมูลที่ไม่ใช่ข้อความ (รูปภาพ) และใช้วิธีบนลงล่างคือ วิธีการแบ่งซ้ำตามแนว X-Y (recursive X-Y cuts or RXYC) โดยจะมีการแยกบรรทัดข้อความที่ไม่ติดกันออกไปเป็นบล็อก วิธีการนี้ง่ายต่อการนำไปใช้และเวลาที่ใช้การทำงานค่อนข้างเร็ว แต่สามารถใช้วิเคราะห์ได้ในรูปแบบที่เป็นสี่เหลี่ยมเท่านั้นและต้องแปลงภาพเกรย์สเกลให้เป็นไบนารีก่อน

การตัดแปลงการแบ่งส่วนประกอบแบบแบ่งแล้วรวบรวม (split-and-merge) ของภาพเอกสารเกรย์สเกลไปเป็นขอบเขตที่มีเนื้อหาเดียวกัน เช่น การแสดงของโครงสร้างข้อมูลแบบต้นไม้ที่ถูกพัฒนาโดย J. Liu, Y.Y. Tang และ C.Y. Suen [34] ได้ทำการแบ่งและการรวมส่วนประกอบเข้าด้วยกัน โดยการคำนวณงานในเวลาเดียวกัน ถ้าบางขอบเขตไม่เป็นเนื้อหาเดียวกันก็จะแบ่งเป็นสี่เหลี่ยมย่อยสี่ส่วน โดยการใช้ค่าเรโซลูชันของการสลายเงา ถ้าสองพื้นที่ที่ติดกัน(อาจจะเป็นส่วนประกอบระดับเดียวกันหรือไม่เป็นก็ได้) มีรายละเอียดแบบเดียวกันจะถูกนำเข้ามารวมกัน ค่าเฉลี่ยและค่าความแปรปรวนของความหนาแน่นของจุดในแต่ละขอบเขตจะถูกใช้ในการตัดสินใจเกี่ยวกับการรวมและการแบ่งขอบเขต ซึ่งรูปแบบเอกสารจะมีรูปแบบที่แน่นอนไม่ได้ ในวิธีการนี้จำเป็นต้องอาศัยการประมาณและแก้ไขความเอียง ซึ่งสามารถนำไปประยุกต์ใช้กับภาพของวารสารเชิงวิชาการที่มีฉากในที่ซับซ้อน และเอกสารที่เป็นตาราง เหมือนกับวิธีการของ S.N. Srihari, T. Hong และ G. Srikanth [35]

วิธีการใน [36] เป็นการสกัดข้อความจากภาพสีที่ซับซ้อนของปกหนังสือและวารสาร โดยใช้การวิเคราะห์แบบผสมผสาน เทคนิคของวิธีการบนลงล่างจะเป็นการแบ่งภาพแบบต่างๆ ไปเป็นบล็อกสี่เหลี่ยมเล็กๆ และจะหยุดการแบ่งเมื่อจุดตัวอย่างมีสีน้อยกว่าสองสีที่แตกต่างกันภายในบล็อก บล็อกที่มีรูปข้อมูลเดียวกันถูกนำมาพิจารณาตลอดทั้งหมดของฉากพื้นชนิดเดียวกัน ขณะที่บล็อกที่ไม่มีรายละเอียดเหมือนกันมีสีที่แตกต่างกันอย่างน้อยสองสีที่สอดคล้องกันข้อความ วิธีการล่างขึ้นบนใช้ตรวจหาขอบเขตที่มีรายละเอียดเหมือนกันที่มีรูปร่างไม่แน่นอน โดยใช้คุณสมบัติของการเจริญเติบโตของขอบเขต (region growing) ผลลัพธ์ของทั้งสองวิธีเมื่อจะถูกนำมารวมกัน เพื่อระบุขอบเขตข้อความและส่วนที่ไม่ใช่ข้อความ โดยมีสมมุติฐานว่าข้อความจัดเรียงกันตามแนวอน

### 2.7.2 การแบ่งประเภทขอบเขตเอกสาร

วิธีการจำนวนมาก [7, 24, 25, 28, 29, 37-40] จะใช้จุดเด่นของส่วนประกอบที่ติดกันเพื่อแบ่งส่วนข้อความและรูปภาพในภาพไบนารี จุดเด่นนี้ประกอบด้วยขนาดของส่วนประกอบที่ติดกันและส่วนที่ใกล้ชิดกัน แนวการจัดเรียง ส่วนเริ่มต้นที่อยู่ใกล้กัน (proximity) และ สิ่งที่อยู่กันยาวขึ้น (elongation)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปร่างที่ได้จากการรันแรงค์จุดค่าและจุดขาวที่เห็นภายในแต่ละส่วนประกอบที่ถูกแบ่งใน [21], [33] การคำนวณค่ารันแรงค์จะดำเนินการในสี่ทิศทาง (แนวนอน แนวตั้ง เอียงซ้าย เอียงขวา) การแบ่งประเภทโดยใช้โครงสร้างข้อมูลแบบต้นไม้ (decision tree) ที่นำเสนอใน [41] และจุดเด่นของฉากพื้นกระเบื้องขาวที่ใช้ใน [42]

วิธีการ [43] เป็นการแบ่งประเภทโดยใช้วิธี ครอสคอรีเลชัน (cross-correlation) ระหว่างเส้นแนวสมแกนที่ติดกัน โดยมีการอ้างอิงจากระยะห่างระหว่างบรรทัด การรวมจุดเด่นของรันแรงค์และครอสคอรีเลชันระหว่างจุดที่นำเสนอใน [8]

วิธีการ [17-18] ใช้การฉายเงาทางแนวดิ่งเพื่อแบ่งข้อความและส่วนที่ไม่ใช่ข้อความออกจากกัน (การประมาณข้อความ) โดยการกระจายของจุดขาวและจุดดำเพื่อกำหนดถึงลักษณะว่าควรจะเป็นกราฟิกหรือรูปภาพ (ซึ่งตัวอักษรจะมีความหนาแน่นของจุดสีดำน้อยในเอกสารที่เป็นรูปแบบโครงสร้าง ส่วนเอกสารที่ไม่ใช่รูปแบบตัวหนังสือจะมีความหนาแน่นของจุดสีดำสูง) การคำนวณระดับสีเทาของภาพเกรย์สเกลทำให้รู้ถึงความแตกต่างระหว่าง ฉากพื้น ข้อความ กราฟิก และรูปภาพ ใน [14]

การจัดแบ่งประเภทโดยใช้วิธีการคำนวณแบบละมุน (soft computing) ที่อธิบายใน [11, 20, 22, 44] จะใช้นิวรอลเน็ตเวิร์ก (ในกรณีที่มีเงื่อนไขมากมายนี่จะมีขั้นตอนหลายระดับมากขึ้น) ขั้นตอนแรกเป็นการสอนให้นิวรอลเน็ตเวิร์กเรียนรู้ก่อน เริ่มจากตัวอย่างข้อความ กราฟิก และรูปภาพ ที่ต้องใช้ในการจัดแบ่งประเภท โดยขนาดบล็อกละจุดเด่นของรันแรงค์เป็นข้อมูลเข้าสำหรับนิวรอลเน็ตเวิร์ก [20] ขณะที่ในระดับค่าของส่วนประกอบชุดภาพที่เล็ก (wavelet packet components) จะถูกคำนวณจากหน้าตาเล็กๆ ที่จุดเด่น (feature) จะถูกนำมาใช้สำหรับการแบ่งประเภทโดยนิวรอล [15, 22] เพื่อที่จะทำการแบ่งประเภทของแต่ละหน้าตาได้อย่างถูกต้อง การรวมค่าสัดส่วนความน่าจะเป็นของกลุ่มข้อมูลที่ส่วนยังไม่ตัดสินใจ (fuzzy) เพื่อใช้ตัดสินใจจากข้อมูลที่ได้นั้นจากหน้าตาที่ใกล้เคียงกันที่เสร็จแล้ว การตัดสินใจเป็นการรวมการวัดที่มาจากภายในรายละเอียดของภาพที่อยู่ภายในมาตราวัดที่กำหนด แนวคิดที่ [16] นิวรอลเน็ตเวิร์กเรียนรู้จากชุดหน้าตาเล็กๆ ซึ่งจะแยกความแตกต่างระหว่างข้อความ ฉากพื้น ภาพวาดลายเส้นและรูปได้ดีที่สุด การรวมเข้าด้วยกัน (convolving) ของหน้าตาที่จุดเด่นของพื้นผิวที่ได้จากภาพหน้าเข้าจะถูกใช้ในการแบ่งประเภทของแต่ละจุดภาพโดยนิวรอลเน็ตเวิร์กไปเป็นหนึ่งในสามประเภท (ข้อความ ภาพวาดลายเส้น รูปภาพฮอฟโทนและฉากพื้น) ขอบเขตที่ถูกจัดไปเป็นประเภทแรกจะต้องดำเนินการเปรียบเทียบค่าที่ได้ของหน้าตาที่ค่ากลางซึ่งถูกกำหนด และถูกแบ่งโดยขนาดของส่วนประกอบที่เกี่ยวข้อง ซึ่งวิธีการนี้คงทนกับภาษาต่างๆ และสามารถแยกความแตกต่างของข้อความในแต่ละภาษา เช่น ภาษาอังกฤษ หรือภาษาจีน วิธีใน [15, 16, 22] จะแตกต่างจากวิธีอื่นๆ เพราะการแบ่งประเภทจุดค่อนหน้าตาจะทำก่อนขั้นตอนการแบ่งส่วนประกอบไปเป็นขอบเขต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.8 การแบ่งประเภทขอบเขตเอกสาร

ในหัวข้อนี้จะอธิบายถึงภาพรวมของการทำงานอย่างคร่าวๆ ของบางวิธีการที่ใช้แบ่งขอบเขตส่วนประกอบเอกสารที่กล่าวถึงในหัวข้อ 2.7 ดังต่อไปนี้

### 2.8.1 วิธีการรันเรงก์สุมูทลกอริทึม (run length smoothing)

เป็นวิธีการที่ใช้การเชื่อมต่อระหว่างจุดในแนวตั้งหรือแนวนอน เพื่อเชื่อมตัวจุดเข้าด้วยกัน หรือเรียกว่าวิธีการทำเบลอ โดยกำหนดค่าขึ้นมาค่าหนึ่งเพื่อใช้เป็นค่าระยะห่าง ซึ่งเป็นการกำหนดค่าระยะห่างของจุดค่าสองจุดทั้งในแนวตั้งหรือแนวนอน โดยมีเงื่อนไขว่าจะเพิ่มจุด (ดำ) เข้าระหว่างจุดค่าสองจุดที่มีระยะห่างกันน้อยกว่าค่าที่กำหนด ดังตัวอย่างในรูปที่ 2.5 เป็นจุดภาพก่อนการทำงาน รูปที่ 2.6 เป็นชุดของจุดภาพหลังจากการทำงานรันเรงก์สุมูทถูกกำหนดด้วยค่าที่ระยะห่างเท่ากับ 3 จุด



รูปที่ 2.5 ชุดจุดภาพก่อนทำการรันเรงก์สุมูท

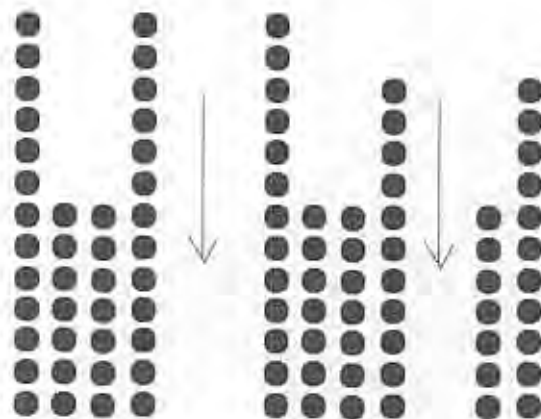


รูปที่ 2.6 ชุดจุดภาพหลังจากทำการรันเรงก์สุมูทด้วยค่าที่กำหนดไว้คือ 3

### 2.8.2 วิธีการแบ่งแนวข้ามแกน X-Y (recursive projection X-Y cut)

วิธีการแบ่งข้ามแนวแกน X-Y มีพื้นฐานมาจากรายงานภาพ ซึ่งลักษณะการทำงานจะเหมือนกับการพิจารณาจุดที่มีลูกแก้วอยู่อย่างกระจัดกระจาย เมื่อเอียงภาคนั้นไปข้างใดข้างหนึ่ง จะพบว่าลูกแก้วจะกองอยู่ที่ข้างที่เอียงลงไป นั่นคือหลักการทำงานของการฉายภาพเงา โดยลักษณะภาพเงาของเอกสาร โบนารีที่ได้ จะมีลักษณะเป็นช่วงความถี่ของจุดดำหรือการทำฮิสโตแกรมของจุดดำ ซึ่งในการแบ่งขอบเขตนั้นจะกำหนดค่าที่ยอมรับขึ้นมา เมื่อช่วงใดที่มีค่าน้อยกว่าที่กำหนดไว้ ช่วงนั้นจะถูกตัดแบ่งออกไป โดยการทำงานแบบตัดข้ามแกนนั้น อ้างขอบเขตในลักษณะแผนผังโครงสร้างข้อมูลแบบต้นไม้ โดยโหนดรากจะแสดงเป็นภาพเอกสารทั้งหน้า โดยโหนดลูกของแต่ละโหนดนั้นจะแสดงส่วนประกอบย่อยของโหนดเมื่อนั้นๆ ดังรูปที่ 2.7 แสดงการแบ่งตามแนวตั้ง โดยมีที่ลูกศรคือตำแหน่งที่ตัดแบ่งส่วนประกอบ

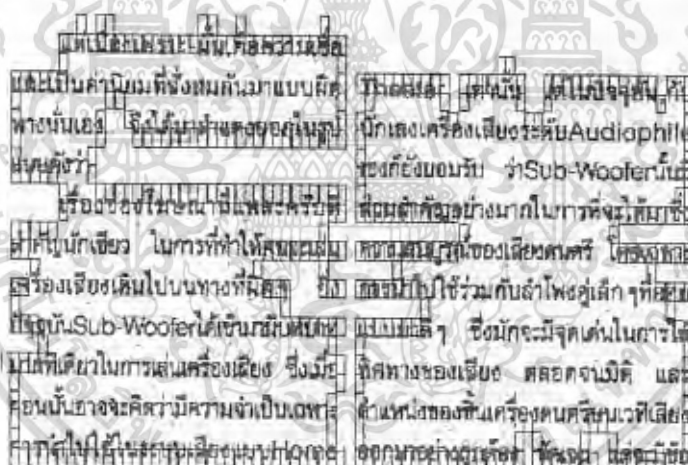
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.7 การฉายเงาจุดดำในแนวตั้ง และช่วงที่จะแบ่ง

### 2.8.3 วิธีการเดินวนรอบขอบ (contouring)

วิธีการนี้เป็นการเดินวนรอบขอบส่วนประกอบ โดยใช้จุดภาพที่ใกล้ชิดกันในทิศทาง 4 ทิศ หรือ 8 ทิศ ตามรหัสแนวหรือรหัสลูกโซ่ ดังแสดงในหัวข้อ 2.4 ซึ่งการเดินวนรอบขอบส่วนนี้ ทำการเดินรอบขอบส่วนประกอบจนวนกลับมายังจุดเริ่มต้น ผลลัพธ์ที่ได้คือขอบเขตของส่วนประกอบ



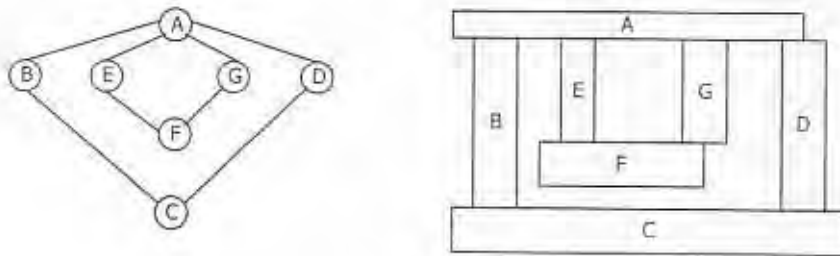
รูปที่ 2.8 การเดินวนรอบเส้นขอบของภาพเอกสารถ

### 2.8.4 วิธีการแบบไวท์ไทล์ (white tile)

วิธีการนี้จะนำส่วนที่เป็นฉากพื้นสีขาวรอบเนื้อหาสาระในภาพเอกสารมาเป็นตัวกำหนดขอบเขตของส่วนประกอบของภาพแบบไบนารี ซึ่งโดยปกติแล้วส่วนเนื้อหาสาระของภาพเอกสารจะถูกล้อมรอบด้วยช่องว่างสีขาว ลักษณะของช่องว่างหรือลักษณะของแผ่นกระเบื้องขาวนั้นจะขึ้นกับรูปร่างของเนื้อหาสาระของเอกสารบริเวณนั้นๆ อาจจะเป็นรูปหลายเหลี่ยม วงกลม สามเหลี่ยม ฯลฯ ดังนั้นเราจะแทนเนื้อหาในภาพเอกสารทั้งหมดด้วยโครงข่ายของช่องว่างสีขาว (Net of white tile) โดยการทำงานจะสร้างบล็อกสี่เหลี่ยมที่มีขนาดเท่ากับช่องว่างสีขาวหรือฉากพื้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ด้วยสมอร์วัตตอริทึมเพื่อที่จะหาเส้นรอบรูปของเนื้อหาที่สนใจ จากนั้นนำข้อมูลบล็อกสี่เหลี่ยม ช่องว่างสีขาวมาแสดงในลักษณะของโครงข่ายด้วยกราฟในรูปที่ 2.8 เพื่อใช้ในการระบุขอบเขต ส่วนประกอบในรูปแบบปิด



รูปที่ 2.9 กราฟที่ใช้แสดงการจัดเรียงไวท์โทน์

## 2.9 เปรียบเทียบวิธีการทำงาน

ตารางการเปรียบเทียบคุณสมบัติและประสิทธิภาพวิธีการในการแบ่งภาพเอกสาร และการระบุขอบเขตของส่วนประกอบที่ได้มาจากรายงานเชิงเทคนิคของ Okun O, Doermann D & Pietikäinen M [44] ซึ่งผลลัพธ์ได้แสดงไว้ในตาราง 2 ตาราง โดยในตารางที่ 2.1 “อ้างอิง” จะอ้างอิงวิธีการ “ชนิดเอกสาร” อ้างอิงประเภทของภาพเอกสารที่ถูกนำมาประมวลผลโดยวิธีการนั้น “ชนิดภาพ” ค่าที่เป็นไปได้มี B, G และ C คือแทน ภาพไบนารี ภาพเกรย์สเกล และ ภาพสี “ฉากพื้น” ค่าที่สามารถเป็น ได้คือ U (พื้นขาวหรือค่า) T (พื้นผิวที่เป็นลวดลาย) หรือ C (สี) “แผนผัง” ค่าที่เป็นไปได้มี R (แผนผังสี่เหลี่ยม) M (แผนผังแมนฮัตตัน) หรือ A (แผนผังที่ไม่มีรูปแบบ) “ความเอียง” นอกว่าวิธีการนี้ใช้งานกับเอกสารที่มีความเอียงบ้างได้ (Y) หรือไม่ได้ (N) “ชุดทดสอบ” และ “% ความถูกต้อง” อ้างอิงชุดขนาดที่นำมาทดสอบและความถูกต้องที่ได้รับจากวิธีการนั้นๆ “-” หมายความว่าจุดเด่นที่ให้มาไม่ได้ถูกอ้างอิงในบทความต้นฉบับ สัญลักษณ์ “+” แสดงถึง ผลลัพธ์ที่ไม่ถูกต้องอธิบายโดยรูปเดียว แต่มีหลายมาตรฐานที่นำมาใช้เช่น การแบ่งเป็นส่วน (fragmentation) และ อัตราการรวมเข้าด้วยกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.1 แสดงคุณสมบัติที่สำคัญของวิธีการต่างๆ ที่ใช้ในการวิเคราะห์แผนผังของเอกสาร

อ้างอิง	ชนิดเอกสาร	ชนิดภาพ	จากพื้น	เด่นชัด	ความเอียง	ชุดทดสอบ	%ความถูกต้อง
[2]	วารสาร	B	U	A	Y	-	-
[3]	วารสาร	B	U	R	N		99-100
[4]	วารสาร	B,G	U	M	N	-	-
[6]	วารสาร หนังสือพิมพ์ จดหมายข่าว โฆษณา	B	U	A	Y	40	+
[8]	วารสาร	B	U	A	Y	-	-
[9]	วารสาร	G	U	A	Y	100	-
[10]	วารสาร	G	U	A	Y		
[11]	วารสาร	G	U	M	N	-	-
[12]	วารสาร	B	U	R	N	-	-
[13]	วารสาร	B	U	R	Y	30	92-97
[14]	หนังสือพิมพ์	B	U	M	N	100	96
[15]	วารสาร	B	U	R	N	50	98.18-99.61
[16]	วารสาร หนังสือพิมพ์	B	U	R	N	100	-
[18]	ปกวารสาร	C	C	A	N	100	95.2-98
[19]	โฆษณา ภาพหน้าเว็บ ปกหนังสือ ภาพจากวิดีโอ	B, G, C	U, C	A	N	26 54 30 6,952	99.2 97.6 72.0 94.7
[20]	ปกหนังสือ	G	U	A	Y	100	91.2
[21]	หน้าเว็บ	C	C	A	Y	200	88.8-92
[23]	ภาพหน้าเว็บ	C	C	A	N	262	90
[25]	วารสาร	B	U	R	N	150	
[26]	หนังสือพิมพ์	G	U	A	Y	-	-
[28]	หนังสือพิมพ์	B	U	R	N	-	78-100
[29]	วารสาร แบบฟอร์ม เช็ทชนิดร	G	U, T	A	N	-	-
[31]	หนังสือและปกวารสาร	C	C	A	N	16	
[32]	วารสาร นมบัตร รายงานเชิงวิชาการ	B	U	A	N	30	93-100
[33]	โฆษณา จดหมาย ของจดหมาย	B	U	A	Y	150	-
[34]	วารสาร	B	U	M	N	150	19.1-99.4
[35]	ปกหนังสือและปกซีดี ภาพจากวิดีโอ	C	C	A	N		
[36]	วารสาร	B	U	R	N	979	97
[37]	ไม่ทราบ	G	T	A	Y		
[38]	วารสาร	B	U	R	Y		
[41]	วารสาร หนังสือพิมพ์	B	U	A	Y	114	-
[42]	วารสาร	G	U	A	N		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในตาราง 2.2 ประกอบด้วย “เวลา” นอกถึงเวลาที่ใช้ในการทำงานหน่วยเป็นวินาที ที่ทำงานบนเครื่องคอมพิวเตอร์รุ่นอะไร “ชนิด” “ขนาดภาพ” บอกขนาดของภาพ โดย “ความละเอียด” บอกถึงความละเอียดของภาพ โดยขนาดภาพมีหน่วยเป็นจุดต่อหน้า ส่วนความละเอียดมีหน่วยเป็นเป็นจุดต่อนิ้ว ส่วนรายละเอียดอื่นๆ จะเหมือนกับในตารางที่ 2.1

ตารางที่ 2.2 แสดงเวลาที่ใช้ในการทำงานสำหรับวิธีการวิเคราะห์เอกสาร

อ้างอิง	ชนิดเอกสาร	เวลา	ขนาดภาพ	ความละเอียด	ชนิด
[6]	วารสาร หนังสือพิมพ์ จดหมายข่าว โฆษณา	0.55	810x1151	100	HP 9000/735
[6]	วารสาร หนังสือพิมพ์ จดหมายข่าว โฆษณา	1.2	1215x1727	150	HP 9000/735
[6]	วารสาร หนังสือพิมพ์ จดหมายข่าว โฆษณา	6.5	2431x3455	300	HP 9000/735
[8]	วารสาร	20	1278x1746		Sun Spare2
[10]	วารสาร	22	A4	300	Sun Spare 20
[11]	วารสาร	60-85	264x332 ถึง 780x1080	100	Sun Spare 20
[14]	หนังสือพิมพ์	9	6592x9590		Pentium 350 II
[16]	วารสาร หนังสือพิมพ์	2.57	A4	300	Sun Spare
[18]	ปกวารสาร	95	2000x2679		Pentium 100
[18]	ปกวารสาร	180	1719x2476		Pentium 100
[19]	โฆษณา	0.15	548x769	150	Sun Ultra-Spare 1
[19]	ภาพหน้าเว็บ	0.11	385x234		Sun Ultra-Spare 1
[19]	ปกหนังสือ	0.4	763x537	50	Sun Ultra-Spare 1
[19]	ภาพจากวีดีโอ	0.09	160x120		Sun Ultra-Spare 1
[20]	ปกหนังสือ	0.01- 2.86	512x512		Pentium 200 Pro
[24]	วารสาร หนังสือพิมพ์	80	A4	400	-
[25]	วารสาร	~ 2	Letter-sized	300	Sun Spare 10
[26]	หนังสือพิมพ์	120	512x512	75	Sun Spare 2
[28]	หนังสือพิมพ์	2.6		100	Sun 3/60
[28]	หนังสือพิมพ์	9.5		200	Sun 3/60
[31]	ปกหนังสือและวารสาร	31.31	1600x2400	200	Sun Ultra Spare 5/10
[32]	วารสาร นามบัตร รายงานเชิงเทคนิค	4.8			PC 486
[33]	โฆษณา ซอจดหมาย จดหมาย	306.9- 563.6	< A4	300	Sun Spare IPX
[34]	วารสาร	1.3	A4	300	SG Indigo
[35]	ปกหนังสือและซีดี ภาพจากวีดีโอ	5.5-6	256x256		Sun Spare 20
[38]	วารสาร	0.9-1.9	A4	300	Sun Spare
[41]	วารสาร หนังสือพิมพ์	2.93	1053x1149	90	Pentium 200 Pro
[41]	วารสาร หนังสือพิมพ์	5.37- 7.03	2592x3300- 3114x3554	300	Pentium 200 Pro

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.10 วิเคราะห์วิธีการทำงาน

หัวข้อนี้จะอธิบายแนวคิดเกี่ยวกับเทคนิคการวิเคราะห์แผนผังเอกสาร [7, 24, 25, 34, 38, 40, 46, 47, 48] ซึ่งขั้นตอนที่วิธีการทั้งหมดนำไปใช้ในการแบ่งประเภทเอกสารต่างๆ และชนิดของภาพ แสดงให้เห็นว่า ไม่มีวิธีการแก้ปัญหาวิธีใดที่สมบูรณ์ที่จะสามารถที่ใช้ในปัญหาการแบ่งส่วนประกอบเอกสารและแบ่งประเภทส่วนประกอบที่หลากหลาย เพราะว่ามันมีขอบเขตกว้าง ยากต่อการจัดการ มีเงื่อนไขในการปรับแต่งมาก และทำนายผลลัพธ์ได้ยาก

สิ่งที่กล่าวมาทั้งหมดจึงจำเป็นต้องมีกระบวนการเฉพาะเพื่อพัฒนาวิธีการสำหรับแยกตัวอักษร จากภาพสีหรือภาพเกรย์สเกล วิธีการโดยทั่วไปจะใช้กฎที่ได้จากการเรียนรู้สำหรับการจัดแบ่งประเภทส่วนประกอบ จุดเด่นส่วนใหญ่ของวิธีการคือจำนวนเงื่อนไขที่พอที่จะทำให้เกิดความมั่นใจ เช่น

1. สีของตัวอักษรต้องเป็นสีเดียวกันเท่านั้น
2. สีจะเป็นตัวแบ่งความแตกต่างระหว่างข้อความและพื้นฉาก
3. ตัวอักษรรวมจะเป็นรูปแบบที่เรียงต่อกันเป็นบรรทัดในแนววนอน (เงื่อนไขนี้ไม่จำเป็นในวิธีการต่อไป [25, 26, 49, 50])

การแยกข้อความจากภาพสีไม่จำเป็นต้องใช้ OCR ในการตรวจสอบความถูกต้องของผลลัพธ์ที่ได้จากการสกัดแยก ซึ่งมีการแบ่งขอบเขตและประเภทของส่วนประกอบก่อนเพื่อให้การทำงานของ OCR ได้ลดจำนวนกฎเกณฑ์ที่ใช้และเป็นการเพิ่มความถูกต้องของการทำงาน

วิธีการวิเคราะห์ แผนผังกับภาพที่เป็นวารสาร หนังสือพิมพ์ หรือเอกสารที่เน้นข้อความอื่นๆ สามารถแบ่งได้เป็นสองกลุ่ม คือ กลุ่มแรกต้องทำการประมาณและแก้ไขความเอียงก่อนการวิเคราะห์ แผนผัง วิธีการชนิดนี้จะเป็นแผนผังชนิดสี่เหลี่ยมหรือแมนฮัตตัน เพราะการแก้ไขความเอียงก่อนจะทำให้การวิเคราะห์ แผนผัง ได้ง่ายขึ้น อย่างไรก็ตามความผิดพลาดที่เกิดจากการหาความเอียงจะลดความถูกต้องของวิธีการวิเคราะห์ แผนผัง ถ้าวิธีการนั้นไม่สามารถจัดการกับขอบเขตที่เอียงได้ ส่วนอีกประเภทหนึ่งจะจัดอยู่ในกลุ่มที่จะทำการแบ่งขอบเขตส่วนประกอบและจัดประเภทส่วนประกอบก่อน แล้วถึงมาประมาณความเอียงเฉพาะขอบเขตที่เป็นข้อความเท่านั้น ปกติวิธีการชนิดนี้จะใช้กับแผนผังที่มีความซับซ้อนและมีรูปแบบที่ไม่แน่นอน หรือภาพเอกสารที่มีหลากหลายความเอียง ในกรณีสุดท้ายการประมาณความเอียงจะถูกนำไปใช้กับบล็อกข้อความทุกบล็อก บางครั้งวิธีการนี้หลีกเลี่ยงใช้กระบวนการดำเนินงานที่พิเศษเพราะว่าการประมาณความเอียงจะไม่ทำในส่วนที่ไม่ใช่ข้อความ อย่างไรก็ตามการวิเคราะห์ แผนผัง ยังเป็นงานที่ยากกว่าการประมาณความเอียงและไม่ง่ายที่จะทำให้ถูกต้องเหมือนการประมาณความเอียง นั่นหมายความว่าความผิดพลาดที่เกิดจากการวิเคราะห์ แผนผัง ที่ซับซ้อนจะพบได้บ่อยมากกว่าการประมาณค่าความเอียง ซึ่งตัวบล็อกสุดท้ายขึ้นอยู่กับนำไปใช้และความซับซ้อนของปัญหา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การกำหนดรายละเอียดภาพจะแตกต่างกันสำหรับงานในแต่ละด้าน ตัวอย่างเช่น ภาพหน้าเว็บ[28] ต้องการรายละเอียดต่ำ แต่สำหรับภาพวารสารหรือหนังสือพิมพ์จะต้องการความละเอียดของภาพสูง การประมวลผลที่ความละเอียดต่ำจะใช้เวลาในคำนวณได้รวดเร็วกว่า แต่เอกสารที่ต้องการรายละเอียดที่แม่นยำจะทำให้เกิดความผิดพลาดและจะทำให้ขั้นตอนการทำงานต่อไปได้ผลลัพธ์ที่ไม่ดี การเลือกความละเอียดที่เหมาะสมไม่ใช่งานที่ง่าย เพราะความละเอียดที่ไม่เหมาะสมจะทำให้เกิดขั้นตอนประมวลผลเอกสารเพิ่มอีกมากมาย

โครงสร้างข้อมูลเป็นส่วนที่สำคัญที่สุดสำหรับวิธีการวิเคราะห์ที่แผนผังของภาพเอกสาร เพราะโครงสร้างข้อมูลที่ดีสามารถลดเวลาในการทำงานและกำหนดความยากง่ายในการเข้าถึงข้อมูลในการทำงานหลังจากการวิเคราะห์ที่แผนผัง ตัวอย่างโครงสร้างที่ยืดหยุ่นและมีประสิทธิภาพ [39, 51] ใช้วิธีการ BAG, [10, 11, 42] ใช้วิธีการ while tile, [34] ใช้วิธีการ quadtrees และ [8, 15, 22, 32] ใช้วิธีการ square block tessellation โดยวิธีทั้งหมดเป็นการแสดงข้อมูลที่ดี ไม่เพียงแต่เข้าถึงข้อมูลได้ง่ายแต่ยังเป็นอิสระต่อความเอียงหรือแผนผัง

สำหรับเทคนิคการคำนวณแบบละมุน (soft computing techniques) ถูกนำมาใช้ในการวิเคราะห์ที่แผนผังเอกสาร [5, 15, 16, 22] และในบทความ [20] อธิบายผลงานวิจัยที่ใช้นิเวศน์เน็ตเวิร์กสำหรับการแบ่งประเภทส่วนประกอบเอกสารอย่างครอบคลุม แต่อย่างไรก็ตามจำนวนข้อมูลที่ใช้ในการสอนให้รู้จำตัวอย่างมีขนาดใหญ่มาก (มากถึง 1,000,000 ใน [16]) และยังคงเชื่อมโยงกับจำนวนตัวอย่างที่ต้องการที่สอดคล้องความยากเคลื่อนที่ระหว่างกลาสมละภายในภคาสมนัก

ในบทความ [16, 32, 34] แสดงตัวอย่างประสิทธิภาพของวิธีการวิเคราะห์ที่แผนผังเอกสาร การสกัดข้อความจากฉากพื้นที่มีกมลสายซับซ้อน [32, 34] และการแบ่งภาษาที่แตกต่างกันภายในภาพเดียวกัน [16] ซึ่งงานวิจัยในด้านการวิเคราะห์ภาพเอกสารยังไม่ค่อยสนใจในด้านนี้

การใช้ฉากพื้นมาช่วยในการแบ่งจะทำให้แบ่งขอบเขตได้มากที่สุด เพราะถ้าปกติแล้วฉากพื้นจะเป็นตัวแบ่งขอบเขตระหว่างขอบเขตที่แตกต่างกัน ความละเอียดของภาพสามารถลดลงถึง 75-100 จุดต่อนิ้วก่อนทำการประมวลผล เป็นความคิดที่ดีถ้ารวมการแบ่งส่วนประกอบเอกสารกับการแบ่งประเภทส่วนประกอบเอกสารเข้าไว้ด้วยกัน จะทำให้ประหยัดเวลาในการทำงานลงเพราะข้อมูลจะถูกนำไปใช้เพียงหนึ่งครั้งเท่านั้น

เวลาที่ใช้ในการประมวลผลและความถูกต้องเป็นจุดเด่นที่สำคัญของวิธีการวิเคราะห์ที่แผนผังเอกสาร เวลาที่ใช้ในการทำงานจะมีช่วงเวลาที่แตกต่างกันมากเริ่มตั้งแต่ประมาณ 1 วินาทีไปจนถึงหลายนาทีต่อหนึ่งภาพ ซึ่งจะพบว่าวิธีการที่ทำงานได้เร็ว เช่น [11, 46] ซึ่งจะทำการประมวลผลความเอียงไปพร้อมกัน และไม่ขึ้นกับแผนผังเอกสาร ใช้เวลาในการประมวลผลภาพไปบนรีประมวล 5.5 – 7 วินาที โดยมีความละเอียด 300 จุดต่อนิ้ว และวิธีการที่เร็วกว่าคือ [39, 43, 51] ใช้เวลาทำงานเพียง 1-2 วินาทีในงานเดียวกัน แต่ในกรณีนี้ไม่มีทั้งการแก้ไขความเอียงของเอกสารซึ่งเป็นสิ่งจำเป็นหรือแม้แต่แผนผังที่เป็นรูปแบบ ไม่แน่นอนก็ทำไม่ได้ การคำนวณด้วยความเร็วเป็นสิ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำคัญมากสำหรับการเรียกค้นข้อมูลข่าวสาร เช่น ข้อความที่สกัดแยกจากภาพหน้าเว็บหรือจากวิดีโอ จากตัวอย่างดังกล่าวจะพบว่า เป็นงานที่ต้องการความเร็วมากกว่าความต้องการในการวิเคราะห์ภาพวารสารหรือหนังสือพิมพ์ ซึ่งขึ้นอยู่กับว่าต้องการการตอบสนองในทันทีหรือไม่ ในกรณีหลัง ผู้ใช้จะมีอิสระในการแก้ไขผลลัพธ์ได้นอกเหนือจากการกำหนดค่าตัวแปรใหม่ๆ ในกรณีที่เป็นรูปแบบการประมวลผลที่สามารถปรับเปลี่ยนค่าใหม่ได้ ซึ่งนับว่ามีความจำเป็นถ้าผลลัพธ์ก่อนหน้าไม่เป็นที่น่าพอใจ ดังนั้นการกำหนดตัวแปรที่เหมาะสมจะขึ้นอยู่กับรายละเอียดของงานในแต่ละด้าน

การเปรียบเทียบความถูกต้องของแต่ละวิธีการที่แตกต่างกันเป็นสิ่งที่ค่อนข้างยาก เพราะการทดสอบจะใช้ชุดข้อมูลที่แตกต่างกันและเงื่อนไขตั้งต้นที่แตกต่างกันในหลายกรณี ยกเว้นการวิเคราะห์ภาพที่เป็นข้อความสั้นๆ การประเมินคุณค่าของผลลัพธ์จากสายตา และประเมินคุณค่าจากประสิทธิภาพที่ตั้งไว้บางครั้งไม่สามารถที่จะกระทำได้นอกจากนั้นยังไม่มีกำหนดที่เป็นอันหนึ่งอันเดียวกันในการวัดความถูกต้อง สำหรับตัวอย่างการวัดอาจใช้อัตราส่วนของจำนวนขอบเขตที่แบ่งได้ถูกต้องต่อหน้า ในการแบ่งส่วนประกอบ และ จำนวนขอบเขตที่ถูกจัดประเภทต่อหน้า หรือถูกกำหนดโดยจำนวนของกรณีที่ขอบเขตถูกแบ่งหรือจำนวนของกรณีที่มีการรวมขอบเขตผิดพลาด การแบ่งขอบเขตที่มีรายละเอียดข้อมูลประเภทเดียวกันสองเป็นส่วนย่อยๆ หลายส่วน ซึ่งง่ายต่อการแก้ไขในขั้นต่อไปมากกว่าการที่รวมกับขอบเขตที่ไม่ใช่ประเภทเดียวกันมารวมไว้ด้วยกันซึ่งทำให้ยากต่อการตรวจสอบ ความถูกต้องจะมีค่าตั้งแต่ 70% ไปจนถึงเกือบ 100% สำหรับวิธีการนี้แตกต่างและประเภทของเอกสารที่มีหลากหลาย ความถูกต้องจะมีค่าสูงถ้าเป็นวารสารและหนังสือพิมพ์ และจะมีค่าต่ำเมื่อเป็นเว็บไซต์และภาพวิดีโอ

รูปแบบที่ถูกดึง (Ground-truthing) และการวิเคราะห์แผนผังเอกสารยังไม่มีทางออกที่สมบูรณ์ทุกวิธีการจะขึ้นต้นในความเอียงที่ไม่เปลี่ยนแปลง แต่เป็นการยากมากที่จะพิสูจน์ความจริงในประสิทธิภาพของวิธีทั้งหมดโดยใช้ชุดข้อมูลขนาดใหญ่ของภาพที่เอียง เพราะรูปแบบที่ถูกดึงปกติจะถูกสร้างด้วยภาพที่ตั้งตรงปราศจากความเอียง ระบบวัดประสิทธิภาพความถูกต้องจะประเมินคุณค่าจากผลลัพธ์ที่ได้จาก OCR [52] (ซึ่งตรงนี้จะ ไม่ชัดเจนว่าความผิดพลาดนั้นเกิดจากการแบ่งส่วนประกอบที่ไม่ดีหรือการรู้จำตัวอักษรไม่ถูกต้อง) หรือ ทั้งหมดจำเป็นที่จะต้องแบ่งรูปแบบให้ถูกต้องสำหรับแต่ละภาพที่เอียง [3, 53, 54, 55] ในกรณีหลังภาพที่ถูกสแกนมีความเป็นไปได้อย่างมีความเอียง และหลังจากที่สร้างรูปแบบที่ถูกดึงสำหรับพวกเอกสาร การทำงานแบบอัตโนมัติและการทำรูปแบบที่ถูกดึงให้ถูกต้องกินเวลาในการทำงานมาก (อาจจะ 5 – 10 นาทีต่อภาพที่ระดับพิกเซล เช่น ที่นำเสนอใน [3, 55] และ ขึ้นไปถึง 5 นาทีต่อหน้าในกรณีที่ทำระดับขอบเขตบล็อกคิงนำเสนอใน [54] ดังนั้นถ้ามีความจำเป็นต้องประเมินค่าประสิทธิภาพของวิธีการใดๆ ต้องใช้หนึ่งภาพที่สแกนเข้าด้วยความเอียงมุมต่างๆเริ่มตั้งแต่ 1 องศาถึง 100 องศา โดยมีการเพิ่มขึ้นทีละ 1 องศา ซึ่งจำเป็นต้องสแกนถึง 100 ครั้งในมุมที่แตกต่างกัน ในกรณีนี้การตรวจสอบรูปแบบที่ถูกดึงจะใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เวลาหลายชั่วโมง นั่นเป็นคำตอบว่าทำไมผู้วิจัยหลายท่านถึงใช้ชีวิตเก่าๆ ในการประเมินคุณค่า (ประเมินคุณค่าโดยใช้การมองเห็นด้วยสายตา) อย่างไรก็ตามก็อาจจะใช้ซอฟต์แวร์ที่ได้จากการประเมินคุณค่าของอัลกอริทึมวิเคราะห์แผนผังและการสร้างรูปแบบที่ถูกต้องเริ่มปรากฏใน [53] แนวคิดที่ [54] เป็นการการสร้างรูปแบบที่ถูกต้องโดยสามารถทำงานอย่างอัตโนมัติ แต่วิธีการนี้ยังทำได้เฉพาะเอกสารที่เป็นข้อความและไม่มีรูปภาพหรือกราฟิก

ในทางปฏิบัติจะดีมากหากมีการใช้อัลกอริทึมเฉพาะเจาะจงสำหรับชนิดเอกสารที่แตกต่างกัน และทำงานโดยมีประสิทธิภาพการทำงานอย่างเหมาะสมที่สุด ซึ่งลักษณะเด่นดังต่อไปนี้ควรจะถูกนำมาพิจารณาในกรณีต่างๆ

- ความคงทนต่อความเอียง (เอียงแบบเดียว หรือ เอียงหลายแบบ)
- ไม่ขึ้นกับรูปแบบแผนผังของเอกสาร
- สามารถสกัดข้อความได้ทั้งจากพื้นสีขาวและฉากพื้นสีตรงข้าม
- ง่ายต่อการเข้าถึงข้อมูล
- มีความเร็วและความถูกต้องสูง
- ไม่ขึ้นกับ ชนิด ขนาด และ ลักษณะของรูปแบบตัวอักษร

## 2.11 สรุป

ในบทนี้จะกล่าวถึงภาพรวมซึ่งเป็นศาสตร์ของการวิเคราะห์แผนผังเอกสาร และอธิบายถึงความก้าวหน้าของงานวิจัยด้านการแบ่งภาพเอกสารเริ่มแรกตั้งแต่ปี 1990 ซึ่งในตอนแรกกล่าวถึงงานด้านการวิเคราะห์ประเภทเอกสารที่หลากหลาย ซึ่งสามารถจัดได้เป็น 3 ประเภท คือ

1. กลุ่มที่มีโครงสร้าง
2. เอกสารที่มีแผนผังโครงสร้าง ไม่แน่นอน
3. เอกสารที่ไม่ใช่เวกเตอร์

ในกรณีเพิ่มเติมได้อธิบายวิธีการที่ใช้สำหรับกลุ่มนี้ อธิบายถึงจุดเด่นเช่นภาพ และ ชนิดของฉากพื้น ความละเอียดของภาพ เวลาที่ใช้ในการประมวลผล ความทนทานต่อความเอียง และแผนผังที่แตกต่างกัน

ถึงแม้จะมีงานวิจัยจำนวนมากมาในด้านนี้ แต่ก็ยังไม่มีวิธีที่เป็นพื้นฐานที่เหมาะสมจะใช้สำหรับประมวลผลภาพในประเภทเอกสารที่ต่างกัน ทั้งความถูกต้อง และระบบการทำงานที่เป็นแบบอัตโนมัติ จุดเด่นที่สำคัญในแต่ละวิธีการที่มีคือ ความเอียง รูปแบบของแผนผัง ความเร็วในการทำงาน ความถูกต้อง และยืดหยุ่น การแสดงตัวแทนภาพทำให้สามารถเข้าถึงข้อมูลได้ง่าย รูปแบบที่เหมาะสมของข้อความ กราฟิก และรูปภาพ การสร้างรูปแบบที่ถูกต้องแบบอัตโนมัติ การวัดประสิทธิภาพของวิธีการวิเคราะห์แผนผังสำหรับภาพไบนารีและเกรย์สเกล ตลอดจนระบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฐานความรู้อัตโนมัติหรือไม่อัตโนมัติ สำหรับงานวิจัยที่ผ่านมาจะแสดงในตารางที่ 2.3 ซึ่งจะแสดงถึง ชื่อผู้วิจัย วิธีการที่ใช้ และคุณลักษณะเด่นของแต่ละวิธีการ

ตารางที่ 2.3 แสดงผลงานวิจัยที่ผ่านมา

ลำดับ	ผู้แต่ง	ปี	วิธีการ	จุดเด่น
1	Wahl et al [56]	1982	Run length smoothing	ใช้เวลามากและไวต่อความเอียง
2	Nagy et al.[2]	1984	X-Y tree cut	ไวต่อความเอียง และเป็นบล็อกสี่เหลี่ยม
3	Wang et al.[33]	1989	Run Length smoothing and recursive X-Y cut	วิเคราะห์หนึ่งสัปดาห์ ไวต่อความเอียง
4	Fujisawa et al.[57]	1990	Top-down	เอกสารสิทธิบัตรประเทศญี่ปุ่น
5	Fisher et al.[58]	1990	Run-length smoothing and connected component extraction	กำหนดขอบเขตข้อความและรูปภาพ ไวต่อความเอียง
6	Pavlidis et al.[59]	1991	Column oriented projection	กำหนดขอบเขตข้อความและรูปภาพเหมาะสมกับความเอียงปานกลาง
7	Baird [60]	1992	Global-to-local strategy	เหมาะสมภาษาต่างๆ และแก้ไขความเอียง
8	Jain et al.[31]	1992	Gabor filtering	ใช้จุดค้นพื้นผิวของภาพเฉพาะ ใช้เวลาในการทำงานมาก
9	Labourgeots et al.[61]	1992	8x3 windows Filtering	เอกสาร ไม่เป็นรูปแบบ ความเอียงไม่มีผล
10	Pavlidis et al.[43]	1992	Horizontal smearing and bottom-up	กำหนดค่าคงที่ เหมาะกับความเอียง
11	Akindele et al.[62]	1993	White space tracing	บันทึกหลายเหลี่ยม สั้นเฉพาะขอบเขตข้อความเท่านั้น
12	Amamoto et al.[63]	1993	Morphological operation on white space	กำหนดตัวเขียนทั้งแนวตั้งและแนวนอน ความเอียงไม่มีผล

\* ที่มา Jain A.K., and Yu B. "Document Representation and Its Application to Page Decomposition", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 20, No. 3, March 1998.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.3 (ต่อ)

13	Ittner et al.[64]	1993	White space and minimum spanning tree	ภาษาและคำหนึ่งเป็นอิสระ ใช้เวลาในการคำนวณมาก
14	O’Gorman [65]	1993	k-nearest neighbor clustering	ใช้กับเอกสารที่มีโครงร่างไม่แน่นอน ให้ความถูกต้องสูง และใช้การคำนวณมาก
15	Antonacopoulos et al.[10], [42]	1994	Contours from white tiles	หาขอบเขตที่ไม่เป็นสี่เหลี่ยมและเอียง มีความผิดพลาดในการแบ่งตัวอักษรที่มีขนาดใหญ่
16	Zlatopolsky [66]	1994	Connected component extraction	ไวต่อค่าต่างๆ เอกสารมีหลายความเอียง
17	Doermann [67]	1995	Wavelet multiscale analysis	การคำนวณซับซ้อนมาก แบ่งเอกสารที่มีความซับซ้อน
18	Drivas et al.[37]	1995	Connected component grouping	แก้ไขความเอียง ใช้เวลาในการทำงานมาก
19	Ha et al.[30]	1995	Connected component-based projection profile	เร็วในการฉายภาพเงาของจุด และไวต่อความเอียง
20	Sylwester et al.[68]	1995	Trainable X-Y cut	มันคงมาก ความเอียงและสัญญาณรบกวน ไม่มีผล
21	Tang et al.[46]	1995	Modified fractal signature	ใช้กับเอกสารที่มีหลายภาษา ใช้การคำนวณมาก
22	Jain et al.[16], [69]	1996	Masks and neural network	ใช้จัดการกับเอกสารที่มีความซับซ้อน ใช้เวลานาน
23	Kise et al.[70]	1996	Background Thinning	ขอบเขตไม่เป็นสี่เหลี่ยม ความเอียงไม่มีผล
24	Liu et al.[71]	1996	Adaptive top-down and bottom-up	ขอบเขตไม่เป็นสี่เหลี่ยม ความเอียงไม่มีผล
25	Yamaahita et al.[72]	1996	Run length smearing and adaptive thresholding	ขนาดของอักษรและช่องว่างมีผลกระทบน้อย ความเอียงไม่มีผล

\* ที่มา Jain A.K., and Yu B. "Document Representation and Its Application to Page Decomposition", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 20, No. 3, March 1998.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## การแบ่งส่วนประกอบภาพเอกสารโดยใช้เส้นขอบ และการแบ่งในแนว XY

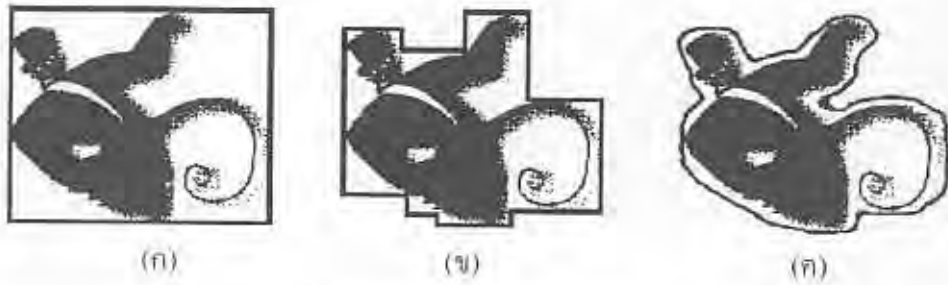
ในบทนี้จะกล่าวถึงวิธีการและอัลกอริทึมที่นำมาใช้ในวิทยานิพนธ์ที่มีชื่อว่า "การแบ่งส่วนประกอบภาพเอกสารโดยใช้เส้นขอบและการแบ่งในแนว XY" โดยมีแนวคิดมาจากการผสมผสานระหว่างสองวิธีการ คือ Contour following technique [1] และ Recursive X-Y cut technique [3] เนื้อหาเริ่มแรกของบทจะทำความเข้าใจเกี่ยวกับรูปร่างของส่วนประกอบในภาพเอกสาร รวมถึงแนวคิดหลักที่ใช้สำหรับแบ่งขอบเขตส่วนประกอบของภาพเอกสาร ไบนารีที่มีลักษณะของโครงสร้างรายละเอียดซับซ้อนให้มีความถูกต้อง การนำแนวคิดของการแบ่งส่วนประกอบที่ได้นำเสนอในวิทยานิพนธ์มาแบ่งส่วนประกอบของภาพเอกสาร ให้เป็นส่วนประกอบใหญ่ ภายหลังจากที่ได้ส่วนประกอบใหญ่แล้วจะทำการระบุจุดผิดพลาดระหว่างโครงสร้างที่ได้จากการแบ่งส่วนประกอบใหญ่ในรอบแรก โดยอัล โนมิด เพื่อที่จะเพิ่มประสิทธิภาพในการทำงานที่รวดเร็วยิ่งขึ้น จากนั้นจะทำการการแบ่งจุดผิดพลาดให้มีส่วนประกอบที่ถูกต้อง และสามารถแบ่งส่วนประกอบใหญ่ที่มี โครงสร้างส่วนประกอบที่ติดกันออกจากรัน และหัวข้อสุดท้ายเป็นการปรับปรุงและจัดเก็บขอบเขตของส่วนประกอบ

### 3.1 รูปร่างของส่วนประกอบภาพเอกสาร

สืบเนื่องจากการแบ่งรูปแบบของแผนผังโครงสร้างของเอกสารที่ได้กล่าว ไปแล้วในบทที่ 2 ซึ่งสามารถแบ่งแผนผังโครงสร้างของเอกสาร ได้ดังนี้ คือ แผนผังสี่เหลี่ยม แผนผังแมนฮัตตัน และ แผนผังที่มีรูปแบบไม่แน่นอน สำหรับการภาพเอกสารที่มีแผนผังแบบสี่เหลี่ยม นั้น เมื่อนำภาพมาแบ่งส่วนประกอบจะสามารถดำเนินการ ได้ง่ายกว่ารูปแบบอื่นๆ เพราะ โครงสร้างไม่ซับซ้อนการระบุขอบเขตจะสามารถกระทำ ได้อย่างชัดเจน แต่มีข้อด้อยในกรณีที่ภาพเอกสารมีพื้นที่ว่างเยอะ จะต้องเสียพื้นที่หน่วยความจำ ไปในการเก็บพื้นที่ว่างส่วนนั้น เป็นผลให้ ได้ ไฟล์ข้อมูลที่มีขนาดใหญ่เมื่อเปรียบเทียบกับแผนผัง โครงสร้างแบบอื่นๆ แผนผังแบบแมนฮัตตันเป็นแผนผังที่ค่อนข้าง จะให้รูปทรงที่ใกล้เคียงกับรูปทรงจริงของส่วนประกอบมากกว่าแผนผังแบบสี่เหลี่ยม แต่ก็ยังคงมีข้อจำกัดที่คล้ายๆ กัน แผนผังที่มีรูปแบบไม่แน่นอน กล่าวได้ว่าแผนผังรูปแบบสุดท้ายเป็นรูปแบบที่มีความยืดหยุ่นที่สุดเหมาะที่จะใช้ในการแสดงการแยกส่วนประกอบต่างๆ ของภาพเอกสารที่มีความหลากหลาย สามารถแสดงส่วนประกอบ ได้ใกล้เคียงความเป็นจริงมากที่สุด เพราะไม่จำกัดว่าขอบเขตของส่วนประกอบต้องเป็นสี่เหลี่ยมเท่านั้น ทำให้ไม่ต้องเสียเนื้อที่หน่วยความจำสำหรับพื้นที่ว่างโดยไม่จำเป็น ดังนั้นแนวทางในการศึกษาจึงเลือกแนวทางการที่จะแยกส่วนประกอบที่มี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แผนผังที่รูปร่างไม่แน่นอน เพราะเป็นปัญหาที่ยากและท้าทายประกอบด้วยเอกสาร โดยทั่วไปสามารถมีได้หลากหลายรูปแบบทั้ง วงกลม วงรี ตลอดจนรูปหลายเหลี่ยม ดังแสดงในรูปที่ 3.1



รูปที่ 3.1 ตัวอย่างแผนผังแบบต่างๆ

- (ก) แผนผังสีเหลี่ยม
- (ข) แผนผังแบบฮัตตัน
- (ค) แผนผังแบบรูปร่างไม่แน่นอน

จากรูปที่ 3.1 จะพบว่าแผนผังที่มีรูปร่างไม่แน่นอนสามารถที่จะแสดงขอบเขตได้ถูกต้องและใกล้เคียงรูปร่างและขนาดของภาพจริงมากที่สุด แต่ก็จะใช้ข้อมูลในการดึงอิงขอบเขตมากที่สุดเช่นกัน รองลงมาจะเป็นแผนผังแบบฮัตตัน และแผนผังแบบสีเหลี่ยมจะบอกขอบเขตได้หยาบที่สุด แต่จะใช้ข้อมูลในการระบุขอบเขตน้อยที่สุดด้วย

เมื่อมีความรู้ความเข้าใจส่วนประกอบของภาพเอกสารแล้ว งานถัดมาที่ต้องใส่ใจสำหรับการแบ่งภาพเอกสาร คือ ทำอย่างไรจึงจะสามารถระบุขอบเขตของส่วนประกอบของภาพเอกสารให้ใกล้เคียงกับความเห็นจริงมากที่สุด ซึ่งแนวคิดในการแบ่งส่วนประกอบภาพเอกสารในลักษณะนี้จะได้นำเสนอในหัวข้อถัดไป

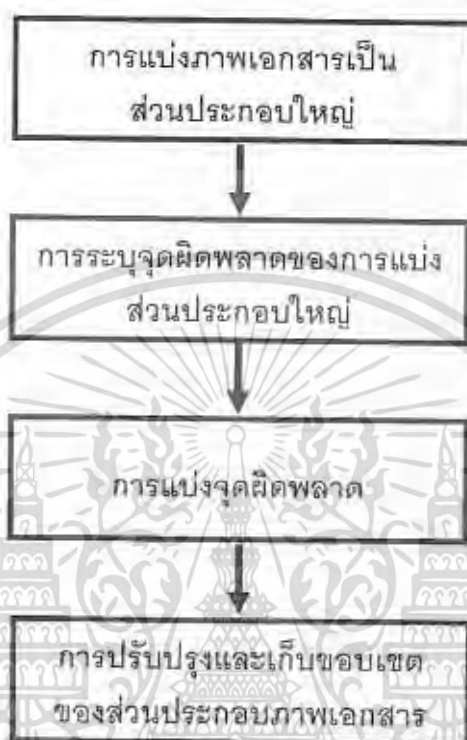
### 3.2 แนวคิดในการแบ่งส่วนประกอบภาพเอกสารใบนารี

ในหัวข้อนี้จะนำเสนอภาพรวมการทำงานของกระบวนการแบ่งส่วนประกอบเอกสาร โดยใช้เส้นขอบและการแบ่งในแนว XY โดยการทำงานจะแบ่งเป็น 4 ส่วนหลักๆ ด้วยกันคือ การแบ่งภาพเอกสารเป็นส่วนประกอบใหญ่ การระบุจุดผิดพลาดของการแบ่งส่วนประกอบใหญ่ การแบ่งจุดผิดพลาดและสุดท้ายเป็นการปรับปรุงเก็บขอบเขตของส่วนประกอบภาพเอกสาร

การทำงานในเบื้องต้นนั้นเริ่มจากการนำภาพเอกสารใบนารีมาทำการระบุขอบเขตของส่วนประกอบอย่างหยาบๆ ในครั้งแรกตามหลักการของการติดตามเส้นขอบ หรือ Contour Following Technique [1] จากนั้นนำขอบเขตที่ได้มาทำการวิเคราะห์เพื่อที่จะหาจุดที่มีการแบ่งผิดพลาดระหว่างโครงสร้างของเอกสารซึ่งเกิดจากระยะห่างในส่วนของคอสัมผัสน้อยเกินไป ถ้าค้นถัดมาเมื่อไรก็ตามที่ทราบตำแหน่งของจุดผิดพลาดก็จะใช้ประยุกต์ใช้หลักการแบ่งซ้ำแนวแกน XY

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หรือ Recursive XY Cut Technique เข้ามาช่วยตัดแบ่งจุดผิดพลาด ณ บริเวณดังกล่าวให้ถูกต้องยิ่งขึ้น และในที่สุดก็จะทำการจัดเก็บขอบเขตของส่วนประกอบใหม่ที่เกิดขึ้นภายหลังจากการแก้ไขจุดผิดพลาดครั้ง ภาพรวมของการทำงานแสดงในรูปที่ 3.2 ซึ่งหลักการทำงานของแต่ละหัวข้อจะนำเสนอในหัวข้อถัดไป



รูปที่ 3.2 ภาพรวมการทำงานของการทำงานที่นำเสนอ

### 3.3 วิธีการแบ่งภาพเอกสารเป็นส่วนประกอบใหญ่

แนวคิดในวิทยานิพนธ์มีพื้นฐานมากจากการติดตามเส้นขอบของภาพเอกสารในงานวิจัย [1] เป็นวิธีการวิเคราะห์ภาพเอกสารแบบผสมผสานระหว่างวิธีการจากล่างขึ้นบนและวิธีการจากบนลงล่าง โดยวิธีการเดิมจะใช้หน้าตาที่เหลี่ยมจัตุรัสขนาดเล็ก (32 x 32 พิกเซล) วางไปบนขอบของภาพเอกสาร ซึ่งนับว่าเป็นวิธีการทำงานแบบผสมผสานที่ดีวิธีการหนึ่ง เพราะจะทำให้รายละเอียดหรือเนื้อหาของภาพที่อยู่ใกล้กันถูกรวมเข้าด้วยกันด้วยความสูงและความกว้างของหน้าตา จึงส่งผลให้สามารถที่จะติดตามแนวเส้นขอบของภาพเอกสารและรวบรวมรายละเอียดเล็กๆ เป็นส่วนประกอบติดกันที่มีขนาดใหญ่ขึ้นได้ เช่น การรวมข้อความที่ถูกแบ่งด้วยช่องว่างของระยะห่างระหว่างบรรทัดหรือรวมตัวอักษรที่ถูกแบ่งด้วยระยะห่างระหว่างช่องไฟเข้าไว้ด้วยกันทำให้ได้ขอบเขตของกลุ่มข้อมูลที่มีขนาดใหญ่ขึ้นในระดับบล็อกที่เป็นขอบเขต โครงสร้างของภาพเอกสาร อีกทั้งรูปทรงหรือแผนผังของภาพเอกสารที่ได้จากการแบ่งของวิธีการนี้สามารถที่จะมีรูปแบบใดๆ ได้ตามขอบเขต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปร่างของส่วนประกอบของภาพเอกสาร โดยจะให้ผลลัพธ์ในการแบ่งภาพเอกสารที่มีรูปร่างใกล้เคียงกับวัตถุจริง และสามารถนำไปประยุกต์ใช้ได้กับเอกสารที่มีรูปแบบใดๆ (แผนผังที่มีรูปแบบไม่แน่นอน) ถ้าระยะห่างระหว่างโครงสร้างมากพอ ทั้งแผนผังโครงสร้างที่เป็นสี่เหลี่ยมและแผนผังแบบแมนฮัตตันได้ ถ้าพิจารณาในเรื่องของเวลาที่ใช้ในการประมวลผลจะพบว่า เป็นวิธีการที่ทำงานรวดเร็วเป็นที่น่าพอใจ เพราะไม่ได้เข้าถึงจุดภาพทุกจุดภายในหน้าต่าง ด้วยข้อดีที่ได้กล่าวมาทำให้เราสนใจและนำวิธีการดังกล่าวมาใช้ โดยมีการปรับปรุงการทำงานในบางส่วนให้เหมาะสมกับรูปแบบปัญหาที่สนใจ ซึ่งในหัวข้อนี้จะแบ่งการทำงานออกเป็น 4 ขั้นตอนดังนี้

### 3.3.1 การหาขนาดหน้าต่างที่เหมาะสม

การแบ่งส่วนประกอบใหญ่ที่นำมาประยุกต์ใช้จะเป็นการเดินทางรอบขอบเขตของภาพเอกสารด้วยหน้าต่างขนาด  $16 \times 32$  พิกเซล โดยจะได้ขอบเขตและขนาดของส่วนประกอบใหญ่ที่ค่อนข้างชัดเจน สำหรับขนาดของหน้าต่างที่เลือกใช้นั้นเป็นผลจากการทดลองกับภาพเอกสารที่มีความหลากหลาย จากการศึกษาโดยมีสมมติฐานว่าขนาดของหน้าต่างจะมีผลกับความถูกต้องในการแบ่งภาพเอกสาร ดังนั้นเราจึงทำการทดลองปรับเปลี่ยนขนาดของหน้าต่างเพื่อศึกษาผลลัพธ์ที่ได้จากการแบ่งของหน้าต่างแต่ละขนาด โดยเริ่มจากหน้าต่างขนาดที่ใหญ่ที่สุด คือ  $32 \times 32$  พิกเซล จนกระทั่งหน้าต่างเล็กที่สุดคือ  $8 \times 8$  พิกเซล แสดงดังตารางที่ ก.1 ในตารางภาคผนวก ก ซึ่งเงื่อนไขในการเลือกขนาดของหน้าต่างที่เหมาะสมจะพิจารณาจากเหตุผลหลายๆ ประการ ได้แก่ ประการแรกความสามารถในการแบ่งแล้วไม่ทำให้เกิดการแบ่งผิด อันเนื่องมาจากมีขีดจำกัดระหว่างบรรทัด หรือมีการแบ่งผิดจากการจัดวางระยะห่างระหว่างคอลัมน์น้อยเกินไป ประการที่สองแบ่งแล้วไม่ทำให้เกิดการรวมย่อหน้าที่ควรแยกออกจากกันเข้าด้วยกัน (under segment) ดังในรูปที่ 3.3 ประการที่สามท้ายแบ่งแล้วไม่ทำให้เกิดย่อหน้ายัดๆ (over segment) ในระดับบรรทัดมากขึ้นเกินไปดังแสดงในรูปที่ 3.4 ซึ่งจากการทดลองพบว่าหน้าต่างขนาด  $16 \times 32$  พิกเซลให้ผลในการแบ่งภาพเอกสารได้เหมาะสมที่สุดสำหรับเอกสารที่มีระยะห่างระหว่างโครงสร้างค่อนข้างแคบหรือเอกสารที่มีระยะห่างระหว่างโครงสร้างชัดเจนดังแสดงในรูปที่ 3.5 แสดงการใช้หน้าต่างขนาด  $16 \times 32$  พิกเซลในการติดตามขอบภาพเอกสาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แต่เมื่อมองหาระเบียบ, คิดควรวางเนื้อ	Theater เท่านั้น แต่ในปัจจุบันกับ
และเป็นคำนิยมที่สั่งสมกันมาแบบคิด	นักแต่งเครื่องเสียงระดับAudiophile
ทางนั้นเอง จึงได้มาสำแดงออกในรูป	เองก็ยังมียอมรับ ว่าSub-Woofernั้นมี
อนุพันธ์ว่า	ส่วนสำคัญอย่างมากในการที่จะได้มาซึ่ง
เรื่องของโฆษณานี้แหละครับที่	ความสมบูรณ์ของเสียงดนตรี โดยเหตุผล
สำคัญนักเขียน ในการที่ทำให้คนจะเล่น	การนำไปใช้ร่วมกับลำโพงคู่เล็ก ๆ ที่ยังมี
เครื่องเสียงเดินไปบนทางที่ผิดๆ ซึ่ง	แบบมาดี ๆ ซึ่งมักจะมึจุดเด่นในการให้
ปัจจุบันSub-Woofersได้เข้ามาบทบาท	ทิศทางของเสียง ตลอดจนมิติ และ
มีบทบาทเดียวในการเล่นเครื่องเสียง ซึ่งมี	ตำแหน่งของชิ้นเครื่องดนตรีบนเวทีเสียง
ก่อนนั้นอาจจะคิดว่ามีความจำเป็นเฉพาะ	ซอกกอกอย่างถูกต้อง ชัดเจน แต่จะมีข้อ
การนำไปใช้ในระบบเสียงแบบHofsky	

รูปที่ 3.3 ผลจากการเลือกขนาดหน้าต่างใหญ่ไป (32x32)

แต่เมื่อมองหาระเบียบ, คิดควรวางเนื้อ	Theater เท่านั้น แต่ในปัจจุบันกับ
และเป็นคำนิยมที่สั่งสมกันมาแบบคิด	นักแต่งเครื่องเสียงระดับAudiophile
ทางนั้นเอง จึงได้มาสำแดงออกในรูป	เองก็ยังมียอมรับ ว่าSub-Woofernั้นมี
อนุพันธ์ว่า	ส่วนสำคัญอย่างมากในการที่จะได้มาซึ่ง
เรื่องของโฆษณานี้แหละครับที่	ความสมบูรณ์ของเสียงดนตรี โดยเหตุผล
สำคัญนักเขียน ในการที่ทำให้คนจะเล่น	การนำไปใช้ร่วมกับลำโพงคู่เล็ก ๆ ที่ยังมี
เครื่องเสียงเดินไปบนทางที่ผิดๆ ซึ่ง	แบบมาดี ๆ ซึ่งมักจะมึจุดเด่นในการให้
ปัจจุบันSub-Woofersได้เข้ามาบทบาท	ทิศทางของเสียง ตลอดจนมิติ และ
มีบทบาทเดียวในการเล่นเครื่องเสียง ซึ่งมี	ตำแหน่งของชิ้นเครื่องดนตรีบนเวทีเสียง
ก่อนนั้นอาจจะคิดว่ามีความจำเป็นเฉพาะ	ซอกกอกอย่างถูกต้อง ชัดเจน แต่จะมีข้อ
การนำไปใช้ในระบบเสียงแบบHofsky	

รูปที่ 3.4 ผลจากการเลือกขนาดหน้าต่างเล็กไป (16x16)

แต่เมื่อมองหาระเบียบ, คิดควรวางเนื้อ	Theater เท่านั้น แต่ในปัจจุบันกับ
และเป็นคำนิยมที่สั่งสมกันมาแบบคิด	นักแต่งเครื่องเสียงระดับAudiophile
ทางนั้นเอง จึงได้มาสำแดงออกในรูป	เองก็ยังมียอมรับ ว่าSub-Woofernั้นมี
อนุพันธ์ว่า	ส่วนสำคัญอย่างมากในการที่จะได้มาซึ่ง
เรื่องของโฆษณานี้แหละครับที่	ความสมบูรณ์ของเสียงดนตรี โดยเหตุผล
สำคัญนักเขียน ในการที่ทำให้คนจะเล่น	การนำไปใช้ร่วมกับลำโพงคู่เล็ก ๆ ที่ยังมี
เครื่องเสียงเดินไปบนทางที่ผิดๆ ซึ่ง	แบบมาดี ๆ ซึ่งมักจะมึจุดเด่นในการให้
ปัจจุบันSub-Woofersได้เข้ามาบทบาท	ทิศทางของเสียง ตลอดจนมิติ และ
มีบทบาทเดียวในการเล่นเครื่องเสียง ซึ่งมี	ตำแหน่งของชิ้นเครื่องดนตรีบนเวทีเสียง
ก่อนนั้นอาจจะคิดว่ามีความจำเป็นเฉพาะ	ซอกกอกอย่างถูกต้อง ชัดเจน แต่จะมีข้อ
การนำไปใช้ในระบบเสียงแบบHofsky	

รูปที่ 3.5 ผลของการเลือกใช้หน้าต่างที่เหมาะสม (16x32)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.3.2 การหาตำแหน่งเริ่มต้นของหน้าต่าง

เมื่อเลือกขนาดของหน้าต่างที่เหมาะสมได้แล้ว ต่อไปจะเป็นการวางหน้าต่างคืนเพื่อเริ่มขั้นตอนการแบ่งภาพเอกสาร โดยจะเริ่มวางที่ตำแหน่งบนซ้ายสุดของรายละเอียดส่วนประกอบภาพเอกสาร แล้ววางหน้าต่างถัดไปเรียงตามขอบเขตของส่วนประกอบ ซึ่งผลลัพธ์ที่ได้จะเป็นขอบเขตของส่วนประกอบใหญ่ ซึ่งจะนำขอบเขตที่ได้นี้ไปหาจุดผิดพลาดของการแบ่งต่อไป กล่าวคือขอบเขตที่ได้จากขั้นตอนนี้ คือ ขอบเขตที่ถูกปิดล้อมด้วยหน้าต่างที่วางเรียงต่อกันจากจุดเริ่มต้นของหน้าต่างแรกจนกระทั่งวนกลับมาเจอหน้าต่างเริ่มต้นอีกครั้ง ดังนั้นขอบเขตที่ได้จึงเป็นขอบเขตส่วนประกอบใหญ่ที่ภายในมีรายละเอียดต่างๆ อยู่รวมกัน

การพิจารณาหน้าต่างด้านบนจะใช้การนับจำนวนจุดสีดำที่อยู่ในหน้าต่างบริเวณขอบของภาพเอกสาร ถ้าหน้าต่างใดมีจำนวนจุดสีดำภายในมากกว่าค่าที่กำหนดก็จะถือว่าเป็นหน้าต่างดำ โดยไม่จำเป็นที่จะต้องนับจำนวนจุดสีดำทั้งหมดในหน้าต่าง ดังนั้นจึงเป็นการลดการเข้าถึงจุดภาพทุกๆ จุดภาพในหน้าต่างลงได้ กล่าวคือเดินตามขอบภาพจะกระทำในระดับกลุ่มของจุด

จากวิธีการที่ได้กล่าวมาแล้วในช่วงขั้นตอนการพิจารณาหน้าต่างดำหรือหน้าต่างที่ใช้ในการติดตามเส้นขอบนั้นจะนำหลักการของรหัสฮูทโซมาใช้ นั่นคือจะวางหน้าต่างดำในทิศทางตามเข็มนาฬิกา เริ่มจากทิศ 0 ถึงทิศ 7 รวมทั้งหมดจะพิจารณาหน้าต่างใน 8 ทิศทางรอบหน้าต่างที่สนใจอยู่ในขณะนั้น ถ้าเจอหน้าต่างสีดำที่ทิศใดก่อนก็จะใช้หน้าต่างในทิศนั้นเป็นหน้าต่างเริ่มในการหาหน้าต่างถัดไป และดำเนินการซ้ำเช่นนี้ไปเรื่อยๆ จนจบทุกขอบหน้าต่างในภาพเอกสาร

จากการทดลองพบว่าขนาดหน้าต่างที่เหมาะสมในการติดตามขอบของภาพเอกสารคือขนาด  $16 \times 32$  พิกเซล การหาตำแหน่งเริ่มแรกของหน้าต่างขนาด  $16 \times 32$  พิกเซลนี้ จะสมมติหาจุดดำไปที่  $(x, y)$  โดยไปที่ละจุดในแนวคอลัมน์ ( $x$ ) เมื่อเจอจุดดำจุดแรก ให้กำหนด  $y$  เป็นตำแหน่งขอบบน และ  $x$  เป็นตำแหน่งขอบซ้าย ของหน้าต่าง ต่อจากนั้นจะทำการหาตำแหน่งของจุดดำทางซ้ายสุดที่อยู่ในช่วงแถว  $y-1$  ถึง  $y+32-1$  และตามแนวตั้ง  $x-1$  ถึง  $x+32-1$  โดยตำแหน่งทางซ้ายสุดที่พบจะกำหนดให้เป็นขอบซ้ายของหน้าต่างเริ่มต้น และทำการตรวจนับจุดดำในหน้าต่างเริ่มต้นว่ามีค่ามากกว่าหรือเท่ากับค่าจำนวนจุดสีดำที่กำหนด (Threshold) หรือไม่ โดยค่าของจำนวนจุดสีดำภายในหน้าต่างใดๆ หาได้จากสูตร

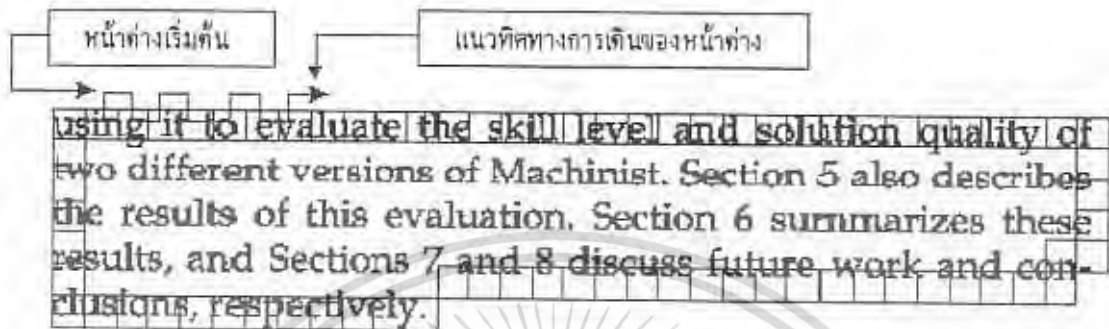
$$\text{Threshold} = 0.02 \times \text{Window Width} \times \text{Window Height}$$

$$\therefore \text{Threshold} = 0.02 \times 16 \times 32 = 10 \text{ Pixels}$$

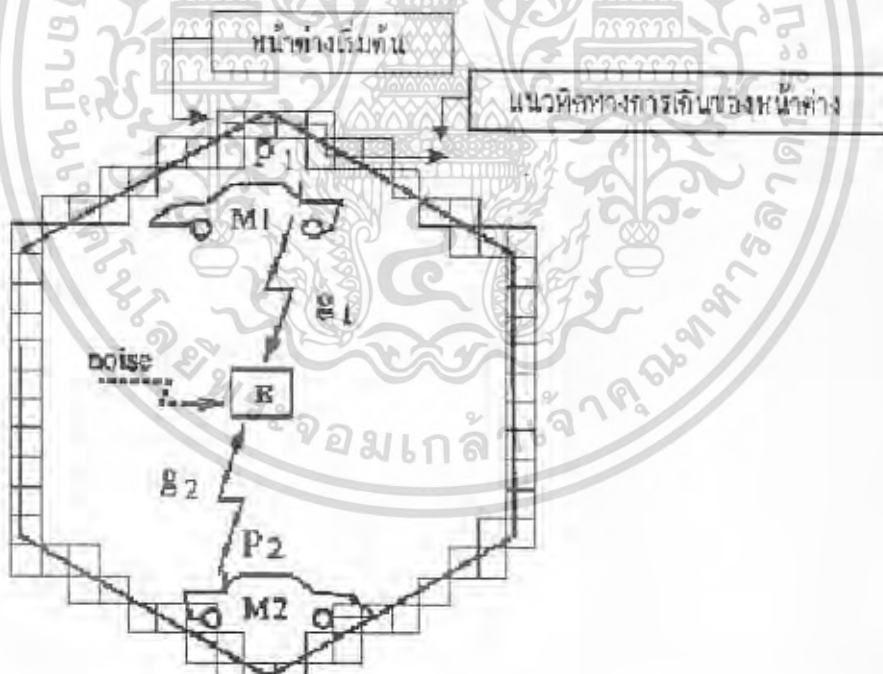
โดยค่า 0.02 เป็นสัมประสิทธิ์ที่ได้จากการทดลองปรับค่าเปอร์เซ็นต์ของจำนวนจุดสีดำที่ใช้ในหน้าต่างใดๆ ซึ่งผลโดยละเอียดแสดงในตารางที่ ก.2 ในตารางภาคผนวก ก ดังนั้นเพื่อตัดแบ่งภาพเอกสารที่หลากหลาย ถ้าพบว่ามีค่าจำนวนจุดสีดำเกินจะค่าที่กำหนดหน้าต่างนี้จะถือเป็นหน้าต่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เริ่มต้นในการวางรอบส่วนประกอบและเป็นหน้าต่างค่า ถ้าไม่ใช้ให้ถือว่ารายละเอียดภายในหน้าต่างนี้เป็นสัญญาณรบกวนและเป็นหน้าต่างขาว จากนั้นจะต้องทำการหาตำแหน่งหน้าต่างเริ่มต้นใหม่ต่อไป โดยเพื่อความชัดเจนจะแสดงแนวทางการจัดวางด้วยหน้าต่างขนาดใหญ่ตั้งในรูปที่ 3.6 แสดงการใช้หน้าต่างเดินวนรอบส่วนประกอบใหญ่ตั้งแต่หน้าต่างแรกจนถึงหน้าต่างสุดท้าย และรูปที่ 3.7 แสดงลักษณะส่วนประกอบใหญ่ที่มีแผนผัง โครงสร้างเอกสารแบบรูปร่างไม่แน่นอน



รูปที่ 3.6 การวางตำแหน่งหน้าต่างเริ่มต้นและลักษณะการเดินวนรอบส่วนประกอบใหญ่ของภาพเอกสารที่มีแผนผังสี่เหลี่ยม



รูปที่ 3.7 การวางตำแหน่งหน้าต่างเริ่มต้นและลักษณะการเดินวนรอบส่วนประกอบใหญ่ของภาพเอกสารรูปแบบใดๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.3.3 การเดินรอบขอบเขตส่วนประกอบใหญ่

การเดินรอบส่วนประกอบใหญ่ตามเข็มนาฬิกาเริ่มจากหน้าค่างบนสุดซ้ายสุด เดินวนรอบขอบเขตโดยทิศทาง 8 ทิศ โดยกำหนดทิศทางต่างๆ ดังต่อไปนี้

0 คือ บน	1 คือ บนขวา	2 คือ ขวา	3 คือ ล่างขวา
4 คือ ล่าง	5 คือ ล่างซ้าย	6 คือ ซ้าย	7 คือ บนซ้าย

$(x-1,y-1)$	$(x,y-1)$	$(x+1,y-1)$	7	0	1
$(x-1,y)$	$(x,y)$	$(x+1,y)$	6	ตำแหน่งปัจจุบัน	2
$(x-1,y+1)$	$(x,y+1)$	$(x+1,y+1)$	5	4	3

(ก)

(ข)

รูปที่ 3.8 ทิศทางการเดินหน้าค่าง (ก) ตำแหน่งต่างๆ ของหน้าค่าง (ข) การกำหนดทิศหน้าค่าง

ในรูปที่ 3.8 (ก) จะเป็นการกำหนดตำแหน่งของหน้าค่าง โดยที่ตำแหน่งที่พิจารณาคือตำแหน่ง  $(x,y)$  ตำแหน่งอื่นๆ จะถูกพิจารณาเทียบกับตำแหน่งนี้ใน 8 ทิศทางรูปที่ 3.8 (ข) เป็นการกำหนดทิศของหน้าค่างดังที่ได้กล่าวมาแล้วในตอนต้น ซึ่งลักษณะหน้าค่างที่กำหนดจะมี 2 ลักษณะคือ หน้าค่างดำและหน้าค่างขาว ซึ่งหน้าค่างดำ คือ หน้าค่างที่มีจุดดำที่อยู่ในหน้าค่างนั้นมากกว่า 10 จุด ถ้าไม่อยู่ในเงื่อนไขให้เป็นหน้าค่างขาว โดยขั้นตอนการทำงานทิศทางเดินของหน้าค่างเป็นดังนี้

1. การกำหนดทิศเริ่มต้นในการค้นหาและวางหน้าค่างดำ
  - 1.1 กรณีทิศทางปัจจุบันเป็นทิศ 0 : ทิศทางต่อไปที่เริ่มพิจารณาคือทิศ 5
  - 1.2 กรณีทิศทางปัจจุบันเป็นทิศ 1 : ทิศทางต่อไปที่เริ่มพิจารณาคือทิศ 6
  - 1.3 กรณีทิศทางปัจจุบันเป็นทิศ 2 : ทิศทางต่อไปที่เริ่มพิจารณาคือทิศ 7
  - 1.4 กรณีทิศทางปัจจุบันเป็นทิศ 3 : ทิศทางต่อไปที่เริ่มพิจารณาคือทิศ 0
  - 1.5 กรณีทิศทางปัจจุบันเป็นทิศ 4 : ทิศทางต่อไปที่เริ่มพิจารณาคือทิศ 1
  - 1.6 กรณีทิศทางปัจจุบันเป็นทิศ 5 : ทิศทางต่อไปที่เริ่มพิจารณาคือทิศ 2
  - 1.7 กรณีทิศทางปัจจุบันเป็นทิศ 6 : ทิศทางต่อไปที่เริ่มพิจารณาคือทิศ 3
  - 1.8 กรณีทิศทางปัจจุบันเป็นทิศ 7 : ทิศทางต่อไปที่เริ่มพิจารณาคือทิศ 4
  - 1.9 กรณีทิศทางปัจจุบันเป็นทิศ 8 : ทิศทางต่อไปที่เริ่มพิจารณาคือทิศ 5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2. การกำหนดทิศเริ่มต้นในการค้นหาและวางหน้าต่างขา

- 2.1 กรณีทิศทางปัจจุบันเป็นทิศ 0 : ทิศทางต่อไปที่จะพิจารณาคือทิศ 1
- 2.2 กรณีทิศทางปัจจุบันเป็นทิศ 1 : ทิศทางต่อไปที่จะพิจารณาคือทิศ 2
- 2.3 กรณีทิศทางปัจจุบันเป็นทิศ 2 : ทิศทางต่อไปที่จะพิจารณาคือทิศ 3
- 2.4 กรณีทิศทางปัจจุบันเป็นทิศ 3 : ทิศทางต่อไปที่จะพิจารณาคือทิศ 4
- 2.5 กรณีทิศทางปัจจุบันเป็นทิศ 4 : ทิศทางต่อไปที่จะพิจารณาคือทิศ 5
- 2.6 กรณีทิศทางปัจจุบันเป็นทิศ 5 : ทิศทางต่อไปที่จะพิจารณาคือทิศ 6
- 2.7 กรณีทิศทางปัจจุบันเป็นทิศ 6 : ทิศทางต่อไปที่จะพิจารณาคือทิศ 7
- 2.8 กรณีทิศทางปัจจุบันเป็นทิศ 7 : ทิศทางต่อไปที่จะพิจารณาคือทิศ 0

ข้อมูลพิกัดตำแหน่งของหน้าต่าง (ขอบบน ขอบซ้าย ขอบขวา และ ขอบล่าง) ทั้งหมดจะเก็บเป็นลิงค์ลิสต์เพื่อใช้ในหาขอบเขตส่วนประกอบด้านซ้ายและด้านขวา และสามารถขยายออกไปได้เรื่อยๆ

เมื่อทิศทางของหน้าต่างที่ค้นพบถูกใช้เป็นที่ทางที่เริ่มต้นในการค้นหาหน้าต่างที่อยู่ถัดไป จะกระทำได้อีกต่อเมื่อส่วนประกอบนั้นมีหน้าต่างอยู่อย่างน้อย 1 หน้าต่างเท่านั้น จึงจะทำให้เข้าสู่การสรุปการค้นหาส่วนประกอบเป็นกรณีเฉพาะ โดยอีกขบวนการย่อยต่างๆ มีความหมายดังนี้  
L = Block Left, R = Block Right, B = Block Bottom, T = Block Top

## 3. การระบุทิศทางในระหว่างการเดินทางของหน้าต่างรอบส่วนประกอบใหญ่ มีกฎดังนี้

- 3.1 กำหนดให้หน้าต่างแรกที่เจอเป็นหน้าต่างที่มีทิศทาง คือ 0 และเป็นหน้าต่างด้าน L
- 3.2 กรณีหน้าต่างปัจจุบันทิศทางคือ 0 : ทิศทางของหน้าต่างถัดไปจะเป็นหน้าต่างด้าน L
- 3.3 กรณีหน้าต่างปัจจุบันทิศทางคือ 1 : ทิศทางของหน้าต่างถัดไปจะเป็นหน้าต่างด้าน L
- 3.4 กรณีหน้าต่างปัจจุบันทิศทางคือ 2 : ทิศทางของหน้าต่างถัดไปจะเป็นหน้าต่างด้าน T
- 3.5 กรณีหน้าต่างปัจจุบันทิศทางคือ 3 : ทิศทางของหน้าต่างถัดไปจะเป็นหน้าต่างด้าน L
  - 3.5.1 ถ้าทิศก่อนหน้าไม่ใช่หน้าต่างด้าน L จะเป็นทิศทางหน้าต่างก่อนหน้าจะเป็นหน้าต่าง R
- 3.6 กรณีหน้าต่างปัจจุบันทิศทางคือ 4 ทิศทางของหน้าต่างถัดไปจะเป็นหน้าต่างด้าน R
- 3.7 กรณีหน้าต่างปัจจุบันทิศทางคือ 5 ทิศทางของหน้าต่างถัดไปจะเป็นหน้าต่างด้าน R
  - 3.7.1 ถ้าทิศก่อนหน้าไม่ใช่หน้าต่างด้าน L จะเป็นทิศทางหน้าต่างก่อนหน้าจะเป็นหน้าต่าง R
- 3.8 กรณีหน้าต่างปัจจุบันทิศทางคือ 6 ทิศทางของหน้าต่างถัดไปจะเป็นหน้าต่างด้าน B
- 3.9 กรณีหน้าต่างปัจจุบันทิศทางคือ 7 ทิศทางของหน้าต่างถัดไปจะเป็นหน้าต่างด้าน L

การเก็บหน้าต่างทั้งหมดในภาพเอกสารจะเก็บด้วยอาร์เรย์ (array) หนึ่งมิติ และ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

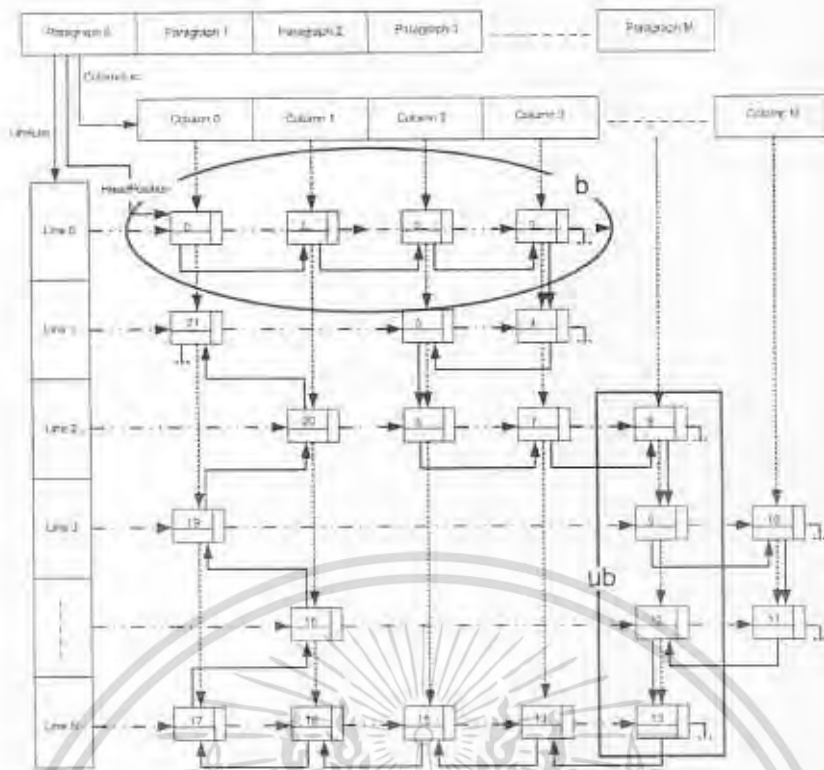
เนื่องจากการวางหน้าต่างมีทิศที่กำหนดแน่นอนตามที่กล่าว โดยจะไล่ไปจนเจอหน้าต่างค่า จากนั้นจะกำหนดทิศของหน้าต่างค่า โดยดูจากทิศหน้าต่างค่าครั้งที่ผ่านมา จากนั้นจะเริ่มวางหน้าต่างใหม่อีกครึ่งหนึ่งตามทิศที่กำหนด ซึ่งในการวางครั้งนี้จะเป็นหน้าต่างที่ซ้ำกับการวางครั้งที่แล้ว ดังนั้นเพื่อลดเวลาในการนับจำนวนจุดค่า จะใช้อาร์เรย์ (array) เก็บจำนวนจุดค่าที่นับได้ของภาพเอกสารทั้งภาพเฉพาะหน้าต่างที่เดินวนไปแล้ว ดังนั้นถ้าหน้าต่างใดเดินวนไปซ้ำก็จะไม่จำเป็นต้องนับจุดค่าใหม่ การใช้อาร์เรย์นี้จะลดเวลาลงได้มาก โดยเฉพาะหน้าต่างขาว เนื่องจากหน้าต่างขาวจะต้องมีการนับทุกจุด คือ 512 จุดสำหรับหน้าต่างขนาด 16x32 พิกเซล ในขณะที่หน้าต่างค่าอาจจะนับน้อยกว่า 512 จุด ขึ้นกับว่าเจอจุดค่ามากกว่า 2% เมื่อใด

### 3.3.4 การหาขอบเขตส่วนประกอบใหญ่ (ขอบซ้ายขอบขวา)

1. นำข้อมูลพิกัดหน้าต่างมาทำการจัดเรียงลำดับตามแนวแกน Y (ขอบบน) และ X (ขอบซ้าย) ด้วยวิธีการ merge sort จากน้อยไปมาก

2. จัดกลุ่มค่าแนวแกน y (แถว) ที่มีค่าเดียวกัน และมีค่า x (คอลัมน์) ของแต่ละพิกัดที่ต่อเนื่องกันเช่น ชุดแรก { (30,20,61,51), (30,52,61,83), (30,84,61,115), (30,116,61,147), (30,148,61,179) } ชุดที่สอง { (62,20,93,51) } ชุดที่สาม { (62,148,93,179), (62, 180, 93,211) } เป็นต้น

ข้อมูลขอบเขตส่วนประกอบใหญ่จะเก็บในโครงสร้างข้อมูลแบบลิงก์ลิสต์ ที่เก็บข้อมูลส่วนประกอบ คือ ขอบบน ขอบซ้าย ขอบขวา และขอบล่างของขอบหน้านั้นๆ นอกจากนั้นยังเก็บข้อมูลขอบของย่อหน้า โดยใช้กลุ่มหน้าต่างในแต่ละบรรทัดหน้าต่าง และกลุ่มของหน้าต่างที่เจอแสดงตามลำดับของลิงก์ลิสต์ดังแสดงในรูปที่ 3-9 และในรูปที่ 3-10 จะเป็นแผนผังขั้นตอนการทำงานในการหาขอบเขตใหญ่ ซึ่งจะเริ่มต้นการทำงานจากการนำภาพเอกสารที่ได้รับการแก้ไขความเอียงและแปลงให้เป็นไฟล์บิตแมปแล้วเข้ามาประมวลผลยังขั้นตอนแรก เพื่อที่จะหาตำแหน่งของหน้าต่างสี่ค่าหน้าต่างแรกในเอกสาร โดยหน้าต่างนี้จะเป็นหน้าต่างที่ใช้เริ่มต้นการเดินวนรอบส่วนประกอบใหญ่ของภาพเอกสาร จากกฎที่ได้กำหนดไว้หน้าต่างนี้มีทิศทางลือทิศ 0 และเป็นหน้าต่างด้านซ้าย เมื่อได้หน้าต่างแรกต่อมาก็จะดำเนินการวางหน้าต่างถัดไปตามกฎของทิศทางที่ได้กล่าวมาแล้วข้างต้น การตัดสินใจวางหน้าต่างเพื่อเก็บขอบเขตส่วนประกอบใหญ่จะวางไปจนกว่าจะพบหน้าต่างเริ่มต้น จากนั้นถ้าเดินรอบขอบเขตส่วนประกอบใหญ่หมดแล้ว ก็จะจัดเรียงตำแหน่งข้อมูลขอบเขตของหน้าต่างจากน้อยไปมากและหาขอบเขตซ้ายขวาของส่วนประกอบดังได้กล่าวไปแล้ว



รูปที่ 3.9 โครงสร้างข้อมูลที่เกี่ยวข้องประกอบใหญ่

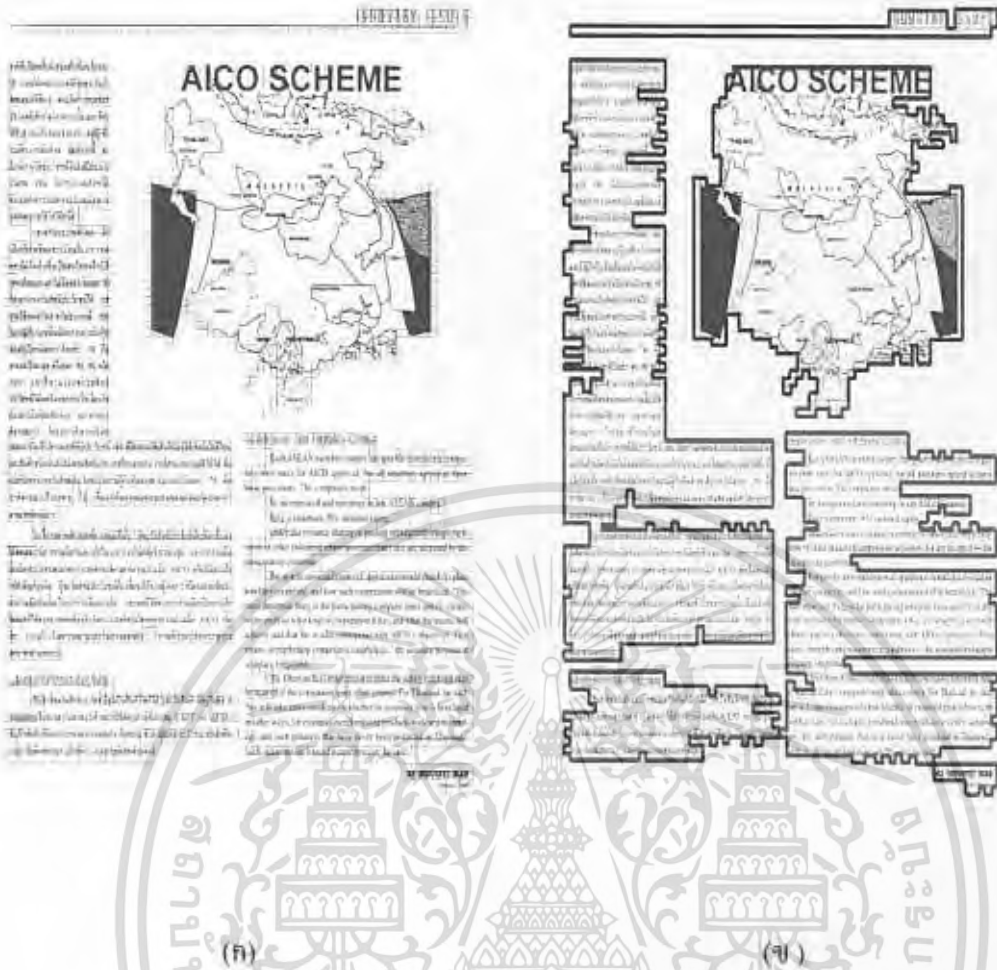


รูปที่ 3.10 แผนผังขั้นตอนการทำงานการแบ่งส่วนประกอบใหญ่

ซึ่งตัวอย่างผลลัพธ์ของการแบ่งและแนวเส้นขอบของภาพเอกสารตามงานวิจัย [2] ได้แสดงดัง

รูปที่ 3.11 (ก) และ 3.11 (ข) ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.11 ลักษณะการติดตามเส้นขอบของภาพเอกสาร โบราณ

ก.ตัวอย่างการติดตามขอบของภาพเอกสารด้วยหน้าต่างขนาด 32x32 พิกเซล  
ข.แนวเส้นขอบของภาพเอกสารที่ใช้หลักการติดตามเส้นขอบ

### 3.4 การระบุจุดผิดพลาดของแบ่งส่วนประกอบใหญ่

ภายหลังจากการเดินรอบส่วนประกอบใหญ่จะพบว่า การแบ่งโดยใช้วิธีการในการติดตามเส้นขอบ จะพบว่าในกรณีที่มีการวางรูปภาพติดกับส่วนของข้อความมากเกินไป หรือมีสัญลักษณ์เกิดขึ้นระหว่างโครงสร้างของส่วนประกอบ หรือมีการจัดวางระยะห่างระหว่างคอลัมน์ของเอกสารน้อยเกินไป ซึ่งบางครั้งอาจเป็นผลมาจากข้อจำกัดต่างที่มี เช่น กรณีของบทความจะมีข้อกำหนดในการตีพิมพ์ โดยจะกำหนดให้มีรูปแบบและจำนวนหน้ากระดาษไม่เกินจำนวนที่กำหนดไว้ ดังนั้นการบีบอัดเนื้อหาให้ได้มากที่สุดจึงเป็นสิ่งจำเป็นอย่างเลี่ยงไม่ได้จนกลายเป็นที่มาและสาเหตุในการแบ่งผิดพลาด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

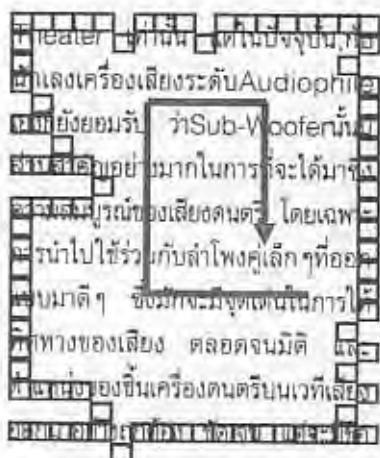
ความสามารถของอัลกอริทึมที่สามารถตัดแบ่งส่วนประกอบเอกสารให้ได้ใกล้เคียงกับรูปร่างจริงของส่วนประกอบนั้นๆ เป็นสิ่งที่ต้องการในการแบ่งภาพเอกสาร เพราะจะไม่ต้องเสียเนื้อที่หน่วยความจำในการเก็บส่วนที่ไม่ใช่เนื้อหาสาระของข้อมูล โดยเปล่าประโยชน์ ดังนั้นจึงเกิดแนวคิดในการพัฒนาและปรับปรุงการแบ่งส่วนประกอบใหญ่ ที่ได้จากการแบ่งในครั้งแรกให้สามารถแบ่งโครงสร้างส่วนประกอบที่ติดกันได้ถูกต้องยิ่งขึ้น ซึ่งก่อนที่จะทำการแบ่งอีกครั้งนั้น จำเป็นต้องรู้ตำแหน่งของจุดผิดพลาดที่จะต้องทำการแบ่งครั้งใหม่ก่อน ทั้งนี้เพื่อเพิ่มความรวดเร็วและเพิ่มประสิทธิภาพในการแบ่ง ลดการทำงานในระดับของจุดภาพหรือในระดับของพิกเซลลงไปได้มาก เพราะจะหาจุดผิดพลาดจากทิสหน้าต่างค่าที่วางตามขอบภาพ

จุดผิดพลาด (Missed Segmentation Point) คือ จุดที่มีการวางตำแหน่งของหน้าต่างค่าในทิสทวนเชื่อมนาฬิกาเมื่อเทียบกับหน้าต่างเดิม ซึ่งจะเป็นส่วนที่ติดกันระหว่างโครงสร้าง หรือระหว่างลอลัมน์ เพราะถ้าโครงสร้างไม่ติดกันจะทำการวางหน้าต่างในทิสทางตามเชื่อมนาฬิกา แต่ถ้ามีการเดินทวนเชื่อมนาฬิกาเมื่อไร แสดงว่ามีจุดแบ่งหรือมีการเว้นวรรค แต่เนื่องจากการเว้นวรรคมีหลากหลายระดับจึงทำให้เราได้จุดผิดพลาดขึ้น วนวนวน ดังนั้นจึงมีการกำหนดค่าความลึกที่เหมาะสมเข้ามาช่วยในการพิจารณาจุดผิดพลาด เพื่อให้ง่ายจะยกตัวอย่างรูปภาพที่เกิดจากการวางหน้าต่างขนาด 32x32 พิกเซล ดังในรูปที่ 3.12 (ก) ถูกสรแสดงการวางหน้าต่างตามเชื่อมนาฬิกา และรูป (ข) ในวงกลมสีแดงจุดที่มีการเว้นวรรคระหว่างลอลัมน์ และจุดผิดพลาดที่เป็นลอลัมน์ติดกัน

ในรูปที่ 3.12 (ข) ส่วนของวงกลมทั้งสองวงแสดงจุดที่มีการวางหน้าต่างทวนเชื่อมนาฬิกา ซึ่งพบว่าจุดที่เกิดการทวนเชื่อมมีความลึกหลายระดับ ถ้ารับในมุมมองคามวิสัยทัศน์ของมนุษย์นั้นจุดที่เกิดการวางหน้าต่างทวนเชื่อมนาฬิกาเป็นจุดเร้าตื้นๆ นั้น โดยมาจะไม่พิจารณาว่าเป็นจุดแบ่งระหว่างส่วนประกอบหรือระหว่างโครงสร้าง ส่วนบริเวณที่เหลี่ยมจะแสดงบริเวณที่มีจุดลอลัมน์ติดกันแต่ไม่สามารถระบุโดยใช้ทิสทวนเชื่อมของหน้าต่างได้ ซึ่งเป็นสาเหตุของการแบ่งส่วนประกอบผิดพลาด

ในรูปที่ 3.13 แสดงความลึก (H) และรูปที่ 3.14 แสดงความกว้าง (W) บริเวณจุดผิดพลาดถูกนำมาใช้พิจารณา เพื่อประกอบการตัดสินใจในการแบ่งภาพเอกสารให้มีความถูกต้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



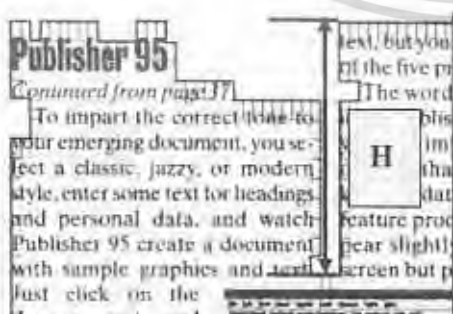
(ก)



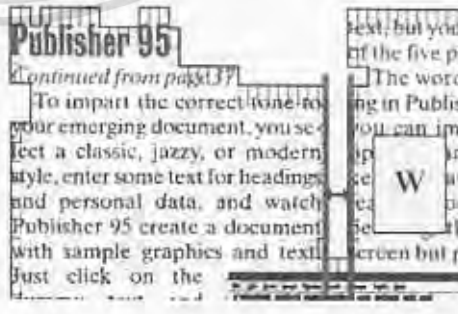
(ข)

รูปที่ 3.12 แสดงการวางหน้าต่างรอบส่วนประกอบ

- (ก) แสดงการเค้นมากรอบภาพเอกสาร โดยปกติจะมีการวางในทิศตามเข็มนาฬิกา
- (ข) แสดงจุดที่มีถ้าวางหน้าต่างทวนเข็มนาฬิกาและนำชะป็นจุดคิดพลาด



รูปที่ 3.13 ความลึกของจุดผิดพลาด



รูปที่ 3.14 ความกว้างของจุดผิดพลาด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นในวิทยานิพนธ์ฉบับนี้จะใช้คำนี้เป็นคำเริ่มต้นการทำงานในการพิจารณาจุดผิดพลาดในการแบ่งย่อหน้าใดๆ ในขั้นตอนถัดไปในรูปที่ 3.15 เป็นภาพตัวอย่างที่แสดงจุดผิดพลาดที่ได้จากการแบ่งส่วนประกอบใหญ่ในครั้งแรก ซึ่งจุดผิดพลาดจะถูกล้อมรอบด้วยวงกลม และมีการกำหนดตำแหน่งของจุดผิดพลาดไว้เพื่อใช้ในการตัดแบ่งต่อไป และแยกจุดผิดพลาดออกเป็นกรณีต่างๆ ได้อย่างชัดเจน โดยทั่วไปสามารถที่จะแบ่งประเภทของจุดผิดพลาดเป็นได้เป็น 2 ประเภท คือจุดผิดพลาดที่เกิดด้านบนและด้านล่างของภาพเอกสารดังรูป 3.16 โดยจะมีกฎในการกำหนดและระบุจุดผิดพลาด

นอกจากนี้แล้วเพื่อให้เกิดความมั่นใจว่าจุดผิดพลาดที่ได้เป็นจุดผิดพลาดที่ถูกต้อง ในกรณีที่ภาพเอกสารมีจุดผิดพลาดทั้งสองประเภทให้นำจุดผิดพลาดที่พบด้านบน (Top Missed Segmentation Point) คือ จุดผิดพลาดที่เกิดในลักษณะตัว U หงาย) และจุดผิดพลาดด้านล่าง (Bottom Missed Segmentation Point) คือ จุดผิดพลาดที่เกิดในลักษณะตัว U คว่ำ) มาพิจารณาประกอบกัน หากตำแหน่งของจุดผิดพลาดทั้งสองตรงกัน ก็จะสรุปได้ว่าจุดผิดพลาดตรงที่พบคือจุดผิดพลาดที่แท้จริง

สำหรับในกรณีที่ภายในส่วนประกอบใหญ่พบเพียงจุดผิดพลาดด้านบนหรือด้านล่างจุดเดียว โดยไม่สามารถที่จะระบุของจุดผิดพลาดที่ยืนยันตำแหน่งจุดผิดพลาดที่ถูกต้องได้ ก็จะพิจารณาความลึกจากขอบบนของย่อหน้า (top of paragraph) หรือความสูงจากขอบล่างของย่อหน้า (bottom of paragraph) เป็นหลัก

เมื่อพิจารณารูปที่ 3.12 หลังจากการให้นำหน้าตัวกินวนรอบขอบเขตส่วนประกอบในครั้งแรกจะพบจุดผิดพลาดทั้งหมด 4 จุด ประกอบด้วยจุดผิดพลาดด้านบนสองจุด คือ จุดที่ 1 และจุดที่ 2 และจุดผิดพลาดด้านล่างสองจุด คือ จุดที่ 3 และจุดที่ 4 เมื่อต้องการยืนยันจุดผิดพลาดที่ได้ว่าเป็นจุดที่มีความน่าเชื่อถือมากน้อยเพียงใด ก็ต้องนำจุดผิดพลาดที่เกิดด้านบนและด้านล่างมาเทียบตำแหน่งกัน โดยถ้ามีจุดผิดพลาดคู่ใดอยู่ในระยะของพิกัดค่า  $X$  ที่กำหนด (ค่าการทำงานของระยะ  $X$ ) ซึ่งอาจจะมีค่าเท่ากันหรือใกล้เคียงกันได้ ในที่นี้กำหนดให้มีความใกล้เคียงกันได้เท่ากับ 48 พิกเซลหรือมีความแตกต่างกันได้เท่ากับ 3 เท่าของขนาดหน้าด่าง โดยนับสองด้านทั้งด้านซ้ายและด้านขวาของตำแหน่งจุดผิดพลาด ซึ่งจากรูปจะพบว่าจุดผิดพลาดด้านบนที่ 1 และจุดผิดพลาดด้านล่างที่ 4 มีตำแหน่งพิกัดแกน  $X$  อยู่ในระยะค่าที่กำหนด จึงสามารถสรุปได้ว่าจุดดังกล่าวเป็นจุดผิดพลาดที่แท้จริง (จากการทดลองพบว่าพิกัดค่าข้อมูลแกน  $X$  มีจุดผิดพลาดด้านบน (1)  $X_1 = 44$  และจุดผิดพลาดด้านล่าง (4)  $X_4 = 41$ , ดังนั้น  $|X_1 - X_4| = 2$  หน้าด่างนั่นเอง) จุดผิดพลาดด้านบน (2) และจุดผิดพลาดด้านล่าง (3) ก็เช่นเดียวกันจากการทดลองพบว่าพิกัดค่าข้อมูลแกน  $X$  จุดผิดพลาดด้านบน (1)  $X_1 = 77$  และจุดผิดพลาดด้านล่าง (4)  $X_4 = 75$ , ดังนั้น  $|X_1 - X_4| = 2$  หน้าด่าง)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



### 3.5 การแบ่งส่วนประกอบซ้ำด้วยวิธีการแบ่งซ้ำในแนว XY

เมื่อทำการยืนยันจุดผิดพลาดที่ถูกต้องได้แล้ว ขั้นตอนถัดมาที่ต้องดำเนินการคือการแบ่งส่วนประกอบ ณ จุดผิดพลาดนั้นให้ถูกต้อง โดยพยายามหาวิธีการที่เหมาะสมมาช่วยเพิ่มประสิทธิภาพและความถูกต้องของการแบ่งภาพเอกสารที่ติดกันออกจากกัน สำหรับในขั้นตอนนี้จะนำการแบ่งซ้ำในแนว XY เข้ามาช่วยปรับปรุงภาพที่ได้จากการแบ่งส่วนประกอบใหญ่ให้มีความถูกต้อง ซึ่งจะเน้นในส่วนของคอลัมน์เพื่อที่จะแก้ปัญหาแบ่งคอลัมน์ที่ติดกันผิดพลาด โดยอาศัยการไล่พิกเซลในแนวแกน X และแนวแกน Y จากนั้นจะวางหน้าต่างลงบนแนวการไล่พิกเซลที่ยาวที่สุด ข้อดีหลักของวิธีการนี้คือใช้เวลาในการคำนวณค่อนข้างจะน้อยกว่าการเดินหน้าต่าง ประกอบกับวิธีการนี้สามารถที่จะทำงานได้ดีกับภาพเอกสารที่มีการจัดแบ่งคอลัมน์ชัดเจน โดยเฉพาะอย่างยิ่งภาพเอกสารที่มีรูปแบบโครงสร้างเป็นแบบสี่เหลี่ยม ในการแบ่งส่วนประกอบซ้ำจะถูกทำหลังจากเก็บขอบเขตของส่วนประกอบใหญ่ครั้งแรก

โดยขั้นตอนการทำงานจะเริ่มจากจุดผิดพลาดที่ได้และประเภทของจุดผิดพลาด ถ้าพบว่าจุดผิดพลาดที่พบในครั้งแรกไม่ให้จุดผิดพลาดที่แท้จริงก็จะไปพิจารณาจุดผิดพลาดถัดไปจนกว่าจะจบย่อหน้านั้นๆ ในทางตรงกันข้ามถ้าจุดนั้นได้รับการยืนยันว่าเป็นจุดผิดพลาดที่แท้จริงก็จะดำเนินการแบ่งตามลักษณะเฉพาะของจุดผิดพลาด ซึ่งโดยทั่วไปจะคิดแบ่งได้สองลักษณะทั้งแนวนอนและแนวตั้ง เงื่อนไขที่ใช้จบการทำงาน คือ ย่อหน้านั้นๆ แบ่งจุดผิดพลาดหมดทุกจุดแล้วและทุกย่อหน้าในภาพเอกสารถูกดำเนินการจนเสร็จ ซึ่งขั้นตอนของการทำงานแบ่งซ้ำแบ่งเป็น 2 แนวทางดังนี้

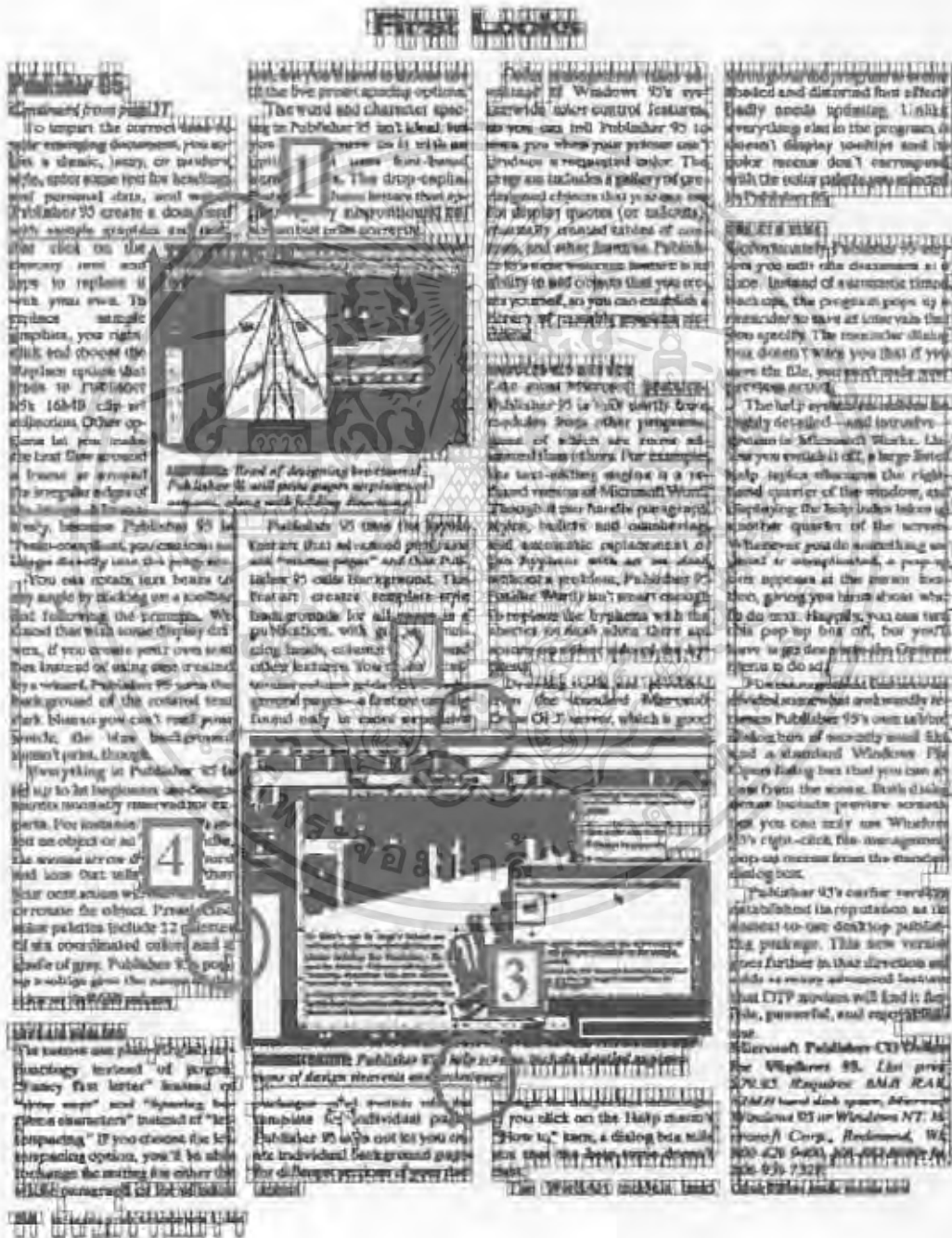
#### 3.5.1 การแบ่งในแนวตั้ง (Vertically separation)

ถ้ามีการแบ่งในแนวตั้งหรือแนวนอน X จะเกิดขึ้นกับจุดผิดพลาดประเภทใดนั้นจะพิจารณาถึงความกว้าง  $w$  บริเวณจุดผิดพลาดว่ามีความกว้างมากน้อยเพียงใด ถ้ามีความกว้างตั้งแต่ 2 หน้าต่างขึ้นไปจะถือว่าจอส่วนประกอบภาพที่มีความกว้าง เช่น รูปภาพ และต้องตัดในแนวนอน แต่ในทางกลับกันถ้าความกว้างของจุดผิดพลาดเท่ากับความกว้างขนาดหนึ่งหน้าต่างพอดี ก็แสดงว่าระยะห่างในการแบ่งแยกระหว่างคอลัมน์แคบมาก ดังนั้นเส้นแบ่งคอลัมน์จึงเป็นแนวตั้งหรือแนวตั้ง ซึ่งอาจจะเป็นผลมาจากข้อจำกัดในการจัดวางรูปแบบของเอกสาร หรือปัญหาจากสัญญาณรบกวนก็เป็นได้ ดังรูปที่ 3.17 จะพบว่าจุดผิดพลาดที่ 4 มีขนาดของจุดผิดพลาดเท่ากับความกว้างของหน้าขนาดหน้าต่างพอดี ดังนั้นจะใช้การแบ่งในแนวตั้ง นอกนั้นเป็นจุดผิดพลาดที่ต้องแบ่งในแนวนอน โดยการแบ่งจะวางหน้าต่างตามกฎดังนี้

- ถ้าจุดผิดพลาดที่พบอยู่ในหน้าต่างทิศทาง 3 และเป็นหน้าต่างด้านบน (R) จะถือว่าเป็นจุดผิดพลาดด้านบน และจะวางแนวการแบ่งของหน้าต่างลงล่างนับจากตำแหน่งจุดผิดพลาด ซึ่งการวางหน้าต่างจะวางไปจนกว่าจะพบตำแหน่งหน้าต่างเป้าหมายที่กำหนดไว้ (ตำแหน่งที่อยู่ด้านตรงข้ามของย่อหน้าย่อหน้านั้นๆ) หรือพบหน้าต่างที่มีค่าของจำนวนจุดสีดำมากกว่าค่าที่กำหนด และถ้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พบว่าผลของการวางหน้าค่างไม่สามารถวางหน้าค่างถึงหน้าค่างเป้าหมายได้ก็จะเปลี่ยนแนวการแบ่งไปเป็นแนวตอนแทน ซึ่งการดำเนินการในแนวนอนจะดำเนินไปจนกว่าจะพบหน้าค่างที่เป็นขอบของส่วนประกอบใหญ่ซึ่งยุติการทำงาน ในขณะที่การวางหน้าค่างในแนวนอนก็จะไล่พิทเซลในแนวตั้งทั้งบนและล่างไปจนเจอหน้าค่างส่วนประกอบใหญ่หรือไล่นในแนวตั้งไปจนเจอพิทเซลดำและเป็นระยะแนวตั้งที่มากที่สุดของแนวนอนนั้น และมีค่าแนวตั้งมากกว่าค่าที่กำหนดก็จะแบ่งในแนวตั้งนั้นต่อไป



รูปที่ 3.17 แนวทางในการแบ่งภาพเอกสาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.18 แนวทางในการแบ่งภาพเอกสารสำหรับจุดผลิตลาด้านบน

- ถ้าจุดผลิตสาดที่พบอยู่ในทิศทาง 7 และเป็นหน้าต่างด้านซ้าย (L) จะถือว่าเป็นจุดผลิตลาดด้านล่าง และจะวางแผนการแบ่งหรือวางหน้าต่างขึ้นบนนับจากตำแหน่งจุดผลิตลาด ซึ่งแนวทางในการแบ่งได้แสดงในรูปที่ 3.19 จากรูปมีตารางหน้าต่างในแนวตั้งขึ้น ไปจนชนเหนือของภาพไม่สามารถวางในทิศทางเดิมได้ ซึ่งในรูปจะแสดงด้วยลูกศรทิศชี้ขึ้น คือ ไปก็จะเปลี่ยนทิศทางการวางหน้าต่างไปเป็นแนวบนทั้งสองด้าน (ลูกศรสองทิศ) แทน และในขณะที่วาง ไปในแนวบนนั้นก็ทำการตรวจสอบด้วยว่ามีกรแบ่งในแนวตั้งอีกหรือไม่ ถ้ามีแนวแบ่งก็จะทำการแบ่งซ้ำตามเงื่อนไขจนไม่สามารถแบ่งได้



รูปที่ 3.19 แนวทางในการแบ่งภาพเอกสารสำหรับจุดผลิตลาดด้านล่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.5.2 การแบ่งในแนวนอน (Horizontally separation)

ในส่วนของการแบ่งซ้ำในแนวนอนหรือแนวแกน Y นั้นจะดำเนินการเมื่อขนาดของความกว้างบริเวณจุดผิดพลาดตั้งแต่ 2 เท่าของขนาดความกว้างของหน้าต่าง นั้นหมายความว่าภาพเอกสารมีการจัดวางส่วนของรูปภาพ และข้อความชิดติดกันในแนวนอนมากจนเทคนิคเดิมของการติดตามเส้นขอบไม่สามารถที่จะแบ่งแยกออกจากกันได้ ซึ่งแนวทางในการแบ่งได้แสดงในรูปที่ 3.17 ในส่วนของจุดผิดพลาดด้านบนที่ 1, 2 และจุดผิดพลาดด้านล่างที่ 3 จะกำหนดแนวทางในการแบ่งให้มี 2 ทิศทาง กือด้านซ้ายและขวา โดยใช้หน้าต่างตำแหน่งตรงมุมของจุดผิดพลาดแต่ละด้านในการเริ่มต้นขั้นตอนการแบ่ง

การแบ่งทั้งย่อหน้าย่อข้อด้านซ้ายและด้านขวาจะดำเนินการตามขั้นตอนที่คล้ายๆ กัน คือเมื่อทราบจุดผิดพลาดและได้รับการยืนยันว่าเป็นจุดผิดพลาดที่แท้จริงแล้ว จะทำการแบ่งระหว่างตำแหน่งของจุดผิดพลาดไปยังเป้าหมาย โดยตรง ซึ่งมีหลักการทำงานตามกฎดังนี้

- ถ้าจุดผิดพลาดที่พบอยู่ในทิศทาง 3, 4, 5, หรือ 6 และเป็นหน้าต่างด้านขวา (R) จะถือว่าเป็นจุดผิดพลาดด้านบน ดังนั้นจะกำหนดให้มีแนวทางการแบ่งของหน้าต่างไปทางซ้ายจนกว่าจะพบหน้าต่างที่เป็นขอบของส่วนประกอบใหญ่ที่เป็นเป้าหมายก็จะหยุดการทำงาน แต่ถ้าพบว่าผลการวางหน้าต่างไปแล้วพบหน้าต่างสีดำที่มีค่าของจำนวนจุดมากกว่าจำนวนจุดสีดำที่กำหนดก็จะเปลี่ยนการทำงานจากแนวเดิมไปกระทำในแนวตั้งแทน

- ถ้าจุดผิดพลาดที่พบอยู่ในทิศทาง 0, 1, 2 หรือ 7 และเป็นหน้าต่างด้านซ้าย (L) หรือด้านบน (T) ตามลำดับ จะถือว่าเป็นจุดผิดพลาดด้านบน ดังนั้นจะกำหนดให้มีแนวทางการแบ่งของหน้าต่างไปทางขวาจนกว่าจะพบหน้าต่างที่เป็นขอบของส่วนประกอบใหญ่ที่เป็นเป้าหมายก็จะหยุดการทำงาน แต่ถ้าพบว่าผลการวางหน้าต่างไปแล้วพบหน้าต่างสีดำที่มีค่าของจำนวนจุดมากกว่าจำนวนจุดสีดำที่กำหนดก็จะเปลี่ยนการทำงานจากแนวเดิมไปกระทำในแนวตั้งแทน

โดยทั้ง 2 กรณีที่กล่าวมาจะมีเงื่อนไขการถือกว่าจะวางหน้าต่างหรือไม่ขึ้นอยู่กับค่าจำนวนจุดสีดำในหน้าต่างนั้นๆ และผลการเดินระดับพิกเซลจะสามารถมีจุดสีขาวใดๆ ที่สามารถเชื่อมต่อกจากตำแหน่งของหน้าต่างจุดผิดพลาดถึงหน้าต่างที่เป็นเป้าหมายได้หรือไม่ ถ้ามีแม้แต่จุดเดียวที่ขาวตลอดก็จะทำการวางหน้าต่างเพื่อใช้แบ่งส่วนของภาพเอกสารนั้นๆ สำหรับแนวทางการแบ่งได้แสดงในรูปที่ 3.20



รูปที่ 3.20 แนวทางในการแบ่งภาพเอกสารสำหรับจุดผิดพลาดในแนวนอน

ภาพรวมของการทำงานการแบ่งส่วนประกอบซ้ำเพื่อแยกส่วนประกอบของโครงสร้างที่ติดกัน ออกจากกัน ได้แสดงไว้ในรูปที่ 3.21



รูปที่ 3.21 แผนผังขั้นตอนการทำงานการแบ่งส่วนประกอบใหญ่ซ้ำเพื่อแยกโครงสร้างส่วนประกอบที่ติดกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.6 การปรับปรุงขอบเขตส่วนประกอบ

เมื่อมีการแบ่งภาพเอกสารในส่วนของคอลัมน์ให้ถูกต้องแล้วขั้นตอนนี้คือสิ่งที่ต้องคำนึงถึงคือ ทำอย่างไรจะเก็บขอบเขตของส่วนประกอบใหม่ได้อย่างถูกต้อง นั่นก็คือการเก็บขอบเขตของหน้าต่างที่ใช้ในการติดตามของภาพใหม่ ต้องมีการแบ่งแยกย่อหน้าหรือส่วนของเอกสารที่ติดกันออกจากกัน ทุกเส้นแบ่งหนึ่งเส้นจะแบ่งส่วนประกอบออกเป็น 2 ส่วน และในส่วนประกอบที่ถูกแบ่งออกไปนั้น ถ้าสามารถแบ่งต่อได้อีกก็จะแบ่งส่วนนั้นออกอีกเป็น 2 ส่วน ทั้งนี้เนื่องจากวิธีการแบ่งจะแบ่งในแนวตั้งแล้วอาจจะต่อไปในแนวนอน แล้วก็อาจแบ่งในแนวตั้งต่อไปได้เรื่อยๆ

- ถ้ามีความกว้าง  $\varnothing$  จุดผิดพลาดตั้งแต่ 2 เท่าของขนาดหน้าต่าง จะกำหนดจุดเริ่มต้นในการแบ่งส่วนประกอบ 2 จุด คือ ตำแหน่งหน้าต่างที่มุมด้านซ้ายและขวาของจุดผิดพลาด จากนั้นทำการเก็บข้อมูลย่อหน้าด้านซ้ายและขวาของจุดผิดพลาดตามลำดับ พร้อมทั้งปรับปรุงขอบเขตของย่อหน้าที่เหลือ (ย่อหน้าด้านบนหรือล่าง)

ในรูปที่ 3.22 (ก) แสดงส่วนที่จะต้องทำการเก็บส่วนประกอบย่อยออกเป็น 3 ย่อหน้าด้วยกัน จุดผิดพลาดจะถูกล้อมด้วยวงกลม และส่วนประกอบที่ต้องทำการเก็บ คือ ย่อหน้าด้านซ้าย (หมายเลข 1) ของจุดผิดพลาด ย่อหน้าด้านขวา (หมายเลข 2) ของจุดผิดพลาดและต้องมีการปรับปรุงขอบเขตของย่อหน้าที่เหลือ (ย่อหน้าด้านล่างหมายเลข 3) โดยอาศัยการแก้ไขข้อมูลขอบเขตของส่วนประกอบใหญ่ที่ได้เดินเก็บขอบเขตไว้แล้ว ซึ่งทำการแก้ไขใหม่โดยการเพิ่มกลุ่ม (จุดถึงคลัสต์) ของหน้าต่างใหม่เข้าไป ณ ตำแหน่งเดียวกันสองจุด จุดแรกเป็นขอบเขตของส่วนประกอบใหม่ที่แบ่งโครงสร้างประกอบที่ติดกันแล้ว ส่วนจุดที่สองเป็นขอบเขตของส่วนประกอบที่เหลือที่ต้องทำการปรับปรุง

- ถ้ามีความกว้าง  $\varnothing$  จุดผิดพลาดเท่ากับหนึ่งหน้าต่างพอดี จะกำหนดจุดเริ่มต้นในการแบ่งส่วนประกอบ 2 จุด ซึ่งสองจุดนี้วางอยู่บนตำแหน่งเดียวกัน จากนั้นทำการเก็บข้อมูล 2 ย่อหน้า คือ ด้านซ้ายและขวาของจุดผิดพลาดตามลำดับ พร้อมทั้งปรับปรุงขอบเขตของย่อหน้าที่เหลือ แต่ในกรณีนี้จะเพิ่มเงื่อนไขพิเศษเข้าไปคือในกรณีที่พบหน้าต่างค่าก่อนหน้าต่างเป้าหมาย จึงจะต้องมีการเพิ่มตำแหน่งของจุดผิดพลาดและจำนวนย่อหน้าที่ต้องเก็บเพิ่มขึ้นมาด้วย

ในรูปที่ 3.22 (ข) แสดงส่วนประกอบใหม่ที่เกิดจากจุดผิดพลาดที่มีความกว้างเท่ากับขนาดของหน้าต่าง ซึ่งจุดผิดพลาดจะถูกล้อมรอบด้วยวงกลม โดยจะต้องทำการเก็บส่วนประกอบด้วยย่อหน้าด้านซ้าย (หมายเลข 1) ของจุดผิดพลาด และปรับปรุงขอบเขตย่อหน้าด้านขวา (หมายเลข 2) ของจุดผิดพลาด จากนั้นตามเงื่อนไขพบว่าการแบ่งเจอหน้าต่างค่าก่อนหน้าต่างเป้าหมาย จึงต้องเพิ่มจุดผิดพลาดในการแบ่งเข้าไปอีกจุดคือจุดที่ถูกล้อมด้วยกรอบสี่เหลี่ยม จากจุดผิดพลาดดังกล่าวทำให้เราเก็บย่อหน้าย่อยเพิ่มได้อีกคือย่อหน้าหมายเลข 3 และ 4 ตามลำดับ การแก้ไขข้อมูลขอบเขตของส่วนประกอบจะแก้ไขจากโครงสร้างเดิมที่ได้เดินเก็บขอบเขตไว้แล้ว และทำการแก้ไขใหม่โดยการเพิ่มถึงคลัสต์ของหน้าต่างใหม่เข้าไป ณ ตำแหน่งเดียวกันสองจุด จุดแรกเป็นขอบเขตของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนประกอบใหม่ที่แบ่งโครงสร้างที่คิดกันแล้ว ส่วนที่สองเป็นขอบเขตของส่วนประกอบที่เหลือที่  
ต้องทำการเก็บขอบเขตต่อไปจนกว่าครบเงื่อนไขที่กำหนด สำหรับลูกศรที่แสดงในภาพทั้งสองจะ  
แสดงทิศทางการแทรกกลุ่มของหน้าต่างในส่วนประกอบ

# First Looks

## Publisher 95

*(Continued from page 37)*

To impart the correct line to your entire document, you select a character or modern style, enter a font size, and press the **Enter** key. Publisher 95 creates a document with sample graphics and text. Just click on the dummy text and type to replace it with your own. To replace sample graphics, you right-click and choose the **Replace** option that leads to Publisher 95's 16MB clip-art



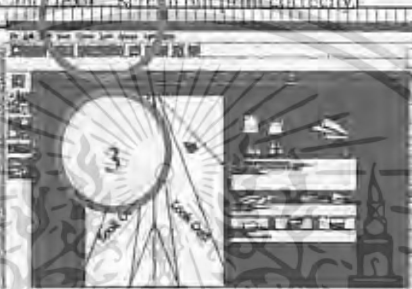
text, but you have to choose one of the five preset spacing options. The word and character spacing in Publisher 95 isn't ideal, but you can improve it with an option that adjusts the font-based kerning of all upper-case letters that appear slightly mispositioned on screen but print correctly.



Color management advantage of Windows-wide color-so you can tell warn you when produce a request program include designed object for display quickly manually created tents, and other or 95's most well- ability to add ob- ate yourself, so y library of reusable

Like most Microsoft Publisher 95 is

MODULES OLD XI



Publisher 95 create a document with sample graphics and text. Just click on the dummy text and type to replace it with your own. To replace sample graphics, you right-click and choose the **Replace** option that leads to Publisher 95's 16MB clip-art

near slightly mispositioned screen but print correctly.

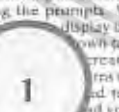
the display quotes (or callouts) manually created tables of contents, and other features. Publisher 95's most welcome feature is the ability to add objects that you create yourself, so you can establish a library of reusable objects.

Like most Microsoft Publisher 95 is built partly from modules from other programs, some of which are more advanced than others. For example, the first-edition engine is a revised version of Microsoft Word, though it can handle paragraphs, bullets and numbering, and automatic replacement, even hyphenation with an *em dash* (though a problem, Publisher 95, like Word, isn't smart enough to replace the hyphens with the correct *en dash* when there are no spaces between the letters.)

And the standard Microsoft Windows OLE server, which is good



Publisher 95 uses the **Master Pages** feature that advanced programs call "master pages" and that Publisher 95 calls **Background**. This feature creates template-style backgrounds for all pages in a publication, with graphics, running headlines, bullets, and other features. You can customize the background, usually found in the **Background** menu.



everything in Publisher 95 is up to let beginners use design secrets normally reserved for experts. For instance, when you click an object or an object handle, a mouse arrow displays a word that tells you whether your next action will move, resize, or rotate the object. Preselected color palettes include 12 letter-size coordinated colors and a shade of gray. Publisher 95's pop-up tooltips give the name of the



And the standard Microsoft Windows OLE server, which is good

(II)

รูปที่ 3.22 การแทรกถึงที่ลัดของหน้าต่างในภาพเอกสาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**Publisher 95**  
 Continued from page 37

To insert text, you select a character style, enter text for headings and personal data, and watch Publisher 95 format the text with sample graphics and text.

1

2

3

**AIRBORNE:** Tired of designing brochures? Publisher 95 will print paper airplanes on your documents.

4

**AIRBORNE:** Tired of designing brochures? Publisher 95 will print paper airplanes on your documents.

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

(ก)

รูปที่ 3.23 ขอบเขตส่วนประกอบใหญ่ที่ได้รับการปรับปรุงส่วนประกอบแล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## วิธีการแบ่งจุดผิดพลาด

ในบทนี้จะกล่าวถึงวิธีการแบ่งส่วนประกอบตรงจุดผิดพลาด เพื่อหาวิธีการที่เหมาะสมในการแบ่งส่วนประกอบย่อยของโครงสร้างที่ติดกัน ซึ่งจะเน้นในส่วนที่ติดกันระหว่างคอลัมน์ที่ได้จากการแบ่งภาพเอกสารในรอบแรก ให้มีความถูกต้องชัดเจนยิ่งขึ้น ซึ่งแต่ละแนวทางจะกล่าวถึงวิธีการของการศึกษา ผลจากการทดลอง และปัญหาที่พบในแต่ละวิธี โดยสองวิธีแรกที่มีแนวคิดมาจาก Contour Following Technique [1] โดยอาศัยการแบ่งขนาดหน้าต่างให้เล็กลงเพื่อประยุกต์ใช้ในการติดตามเส้นขอบของภาพที่แคบขึ้น ณ บริเวณจุดผิดพลาดที่พบ การจำแนกประเภทของหน้าต่างร่วมกับหลักการแพร่โดยระบุประเภทของหน้าต่างที่ยินยอมให้ทำการแพร่ได้ การแพร่โดยใช้ตำแหน่งพิกัดเป้าหมายที่กำหนด และท้ายสุดจะเป็นวิธีการในการแบ่งข้ามแนวแกน X-Y ซึ่งเป็นวิธีการที่เลือกใช้ในวิทยานิพนธ์ฉบับนี้ ซึ่งรายละเอียดได้กล่าวไว้แล้วในบทที่ 3

### 4.1 การติดตามเส้นขอบโดยแบ่งหน้าต่างย่อย

วิธีการนี้มีแนวคิดเริ่มต้นมาจากเทคนิคการติดตามเส้นขอบหรือ Contour Following Technique ซึ่งจะถูกนำมาประยุกต์ใช้ภายหลังจากที่เดินรอบขอบของส่วนประกอบใหญ่และสามารถที่จะระบุตำแหน่งของจุดผิดพลาดได้แล้ว โดยมีแนวคิดในการทดลองว่า ถ้าหากลดความกว้างของหน้าต่างให้เล็กลงและคงความสูงของหน้าต่างให้เท่าเดิมเพื่อให้สามารถรวมกลุ่มของข้อความระหว่างบรรทัดเข้าด้วยกัน หลังจากนั้นจะสามารถชอกซอนเข้าไปในระหว่างโครงสร้างที่แคบๆ ได้ จึงน่าจะเหมาะสำหรับแบ่งจุดผิดพลาดระหว่างโครงสร้างหรือคอลัมน์ที่แคบๆ ได้ดี

จากแนวคิดดังกล่าว ในการทำการทดลองจึงปรับลดขนาดความกว้างของหน้าต่างจากเดิมขนาด 16x32 พิกเซล เหลือเพียง 8x32 พิกเซล (1 ใน 2 ของ 16x32 พิกเซล) โดยในการทดลองจะเริ่มใช้หน้าต่างขนาด 16x32 พิกเซลเดินเพื่อเก็บขอบเขตของส่วนประกอบใหญ่ ต่อมาระบุจุดผิดพลาดจากการเดินรอบส่วนประกอบใหญ่ จากนั้นจะทำการวางหน้าต่างขนาด 8x32 พิกเซล ซึ่งเริ่มต้นติดตามเส้นขอบ ณ จุดผิดพลาดเป็นต้นไป หน้าต่างจะวางบนส่วนที่เป็นสีขาวของหน้ากระดาษ (ส่วนแบ่งคอลัมน์) ไปจนเจอขอบของหน้าต่างเดิม (หน้าต่างขนาด 61x32 พิกเซล) โดยมีแนวคิดว่า หน้าต่างจะจะมีจำนวนจุดสีดำไม่เกินค่าที่กำหนด (ค่า Threshold)

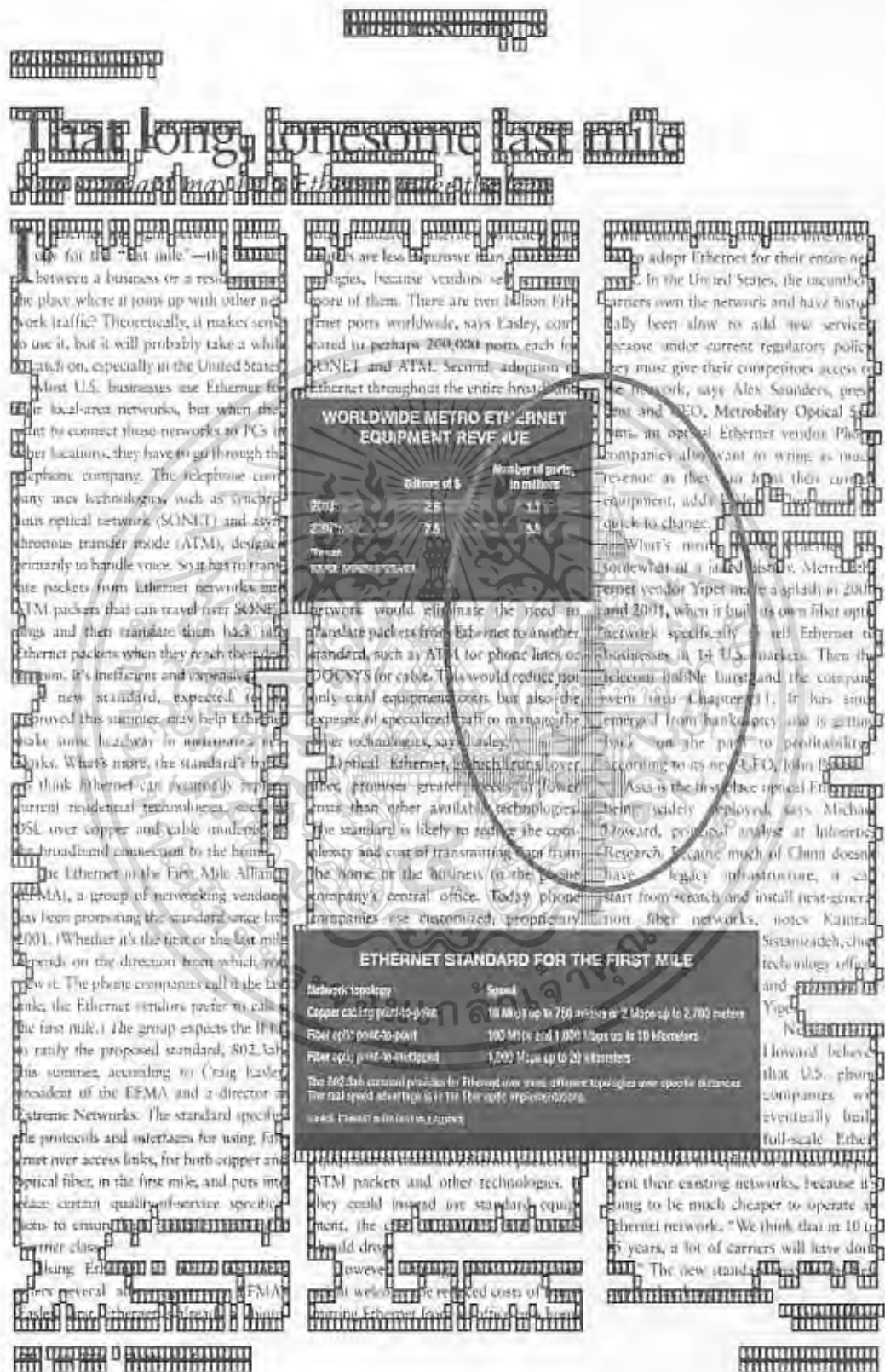
$$\text{Threshold} = 0.02 \times \text{Window Width} \times \text{Window Height}$$

$$\therefore \text{Threshold} = 0.02 \times 8 \times 32 = 5 \text{ Pixels}$$

ดังนั้นหน้าต่างขาวที่จะวางลงไปต้องมีจำนวนจุดสีดำในหน้าต่างน้อยกว่า 5 พิกเซล และถ้า

จำนวนจุดมากกว่านี้จะถือว่าเป็นหน้าต่างดำ จากผลการทดลองพบว่าหน้าต่างขนาดดังกล่าว เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

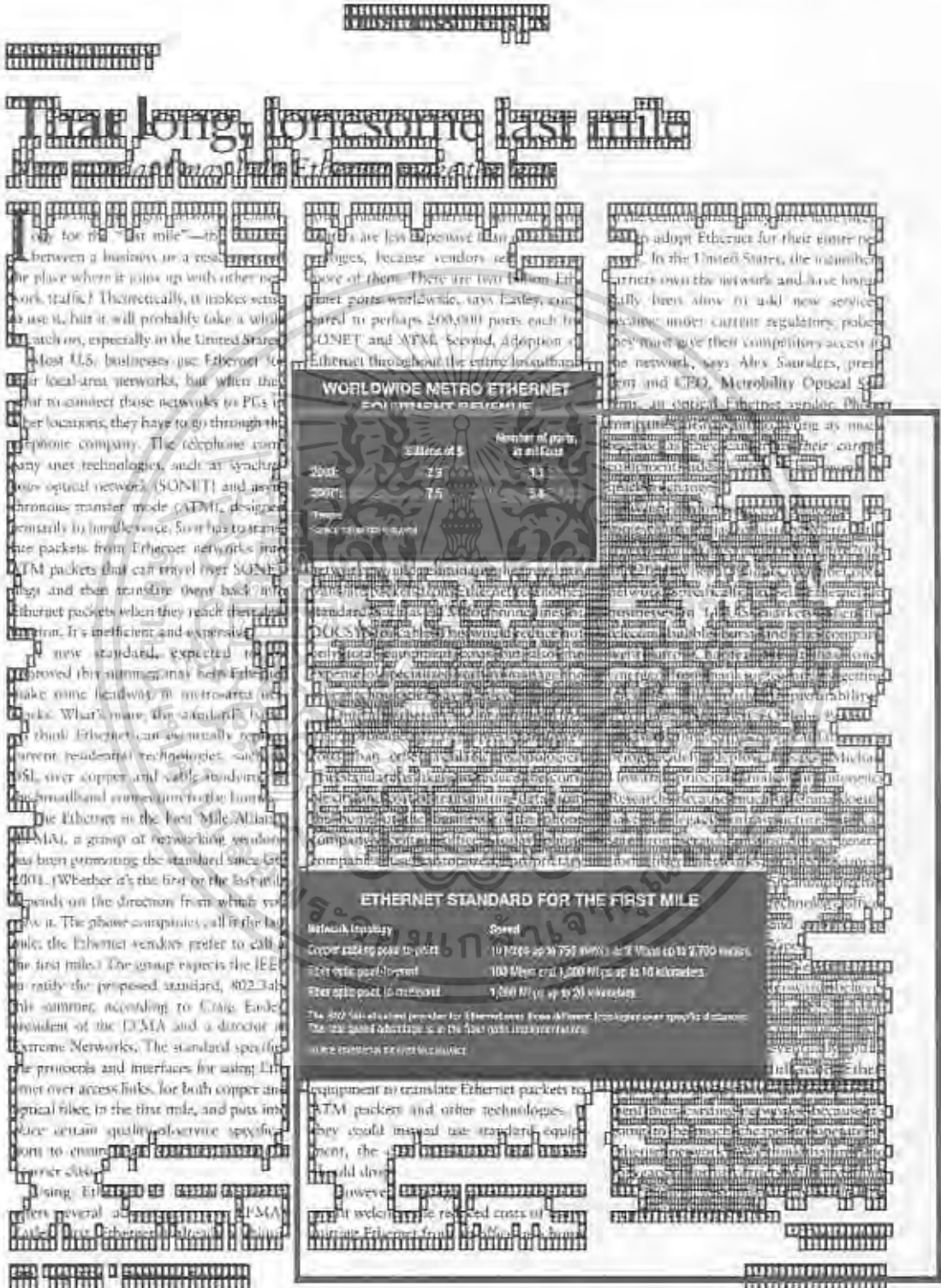
สามารถที่จะแทรกในแนวดิ่งได้แต่ไม่สามารถแทรกในแนวอนค่อไปได้ ดังแสดงในรูปที่ 4.1



รูปที่ 4.1 การแบ่งจุดศัพทภาคโดยใช้หน้าตาขนาด 8x32 พิกเซลในการติดตามเส้นขอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากข้อค้อยของหน้าต่างขนาดเต็มจึงเกิดแนวคิดในการทำการทดลองปรับลดขนาดความกว้างของหน้าต่างจากเต็มขนาด 8x32 พิกเซล เหลือเพียง 8x16 พิกเซล โดยมีขั้นตอนการทำงานเช่นเดียวกับที่กล่าวมาแล้ว ซึ่งค่าจำนวนจุดสีดำที่กำหนดว่าหน้าต่างดำหรือไม่ดำได้จาก



รูปที่ 4.2 การแบ่งส่วนประกอบตรงจุดผลิตผลโดยใช้น้ำต่างขนาด 8x16 พิกเซลในการติดตาม

เส้นขอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถ้าหากวิธีการนี้สามารถทำงานได้รวดเร็วเช่นกัน แต่ปัญหาที่พบของแนวทางนี้ คือ เมื่อหน้าดำมีขนาดเล็กทำให้เกิดการแทรกซ้อนไปทุกๆ ตำแหน่งที่เป็นช่องว่าง จากภาพจะพบว่าผลลัพธ์ของการแบ่งระหว่างโครงสร้างที่ติดกันจะได้หน้าดำกระจายเกือบทั่วทั้งส่วนประกอบที่ต้องการแบ่ง ซึ่งเป็นสิ่งที่ไม่พึงประสงค์ ดังนั้นวิธีการนี้จึงไม่เหมาะที่จะใช้ในการแบ่งจุดผิดพลาดระหว่างโครงสร้าง

## 4.2 การแพร่โดยจำแนกประเภทของหน้าดำ

ภายหลังจากความคิดในการใช้หน้าดำขนาดเล็กไม่เป็นผล จึงมีแนวความคิดริเริ่มการแพร่หน้าดำจากจุดที่ผิดพลาดขึ้น โดยหลักการจะใช้การแพร่หน้าดำสีขาวขนาด  $16 \times 32$  พิกเซลใน 8 ทิศทางตามรหัสถูกใจ ซึ่งเกณฑ์การตัดสินใจในการวางหน้าดำขึ้นอยู่กับประเภทของหน้าดำและตำแหน่งของหน้าดำ การจำแนกประเภทของหน้าดำจะแบ่งออกเป็น 4 ประเภท โดยการแบ่งหน้าดำออกเป็นประเภทต่างๆ โดยมีแนวคิดที่ว่า หน้าดำที่วางอยู่กลางคอลัมน์ที่มีตัวอักษรอยู่เต็มหน้าดำกับหน้าดำบริเวณขอบของคอลัมน์ที่มีส่วนของตัวอักษรหรือสัญลักษณ์อยู่เล็กน้อย ถึงแม้หน้าดำทั้งสองชนิดจะถือว่าเป็นหน้าดำคล้ายกัน แต่หน้าดำชนิดหลังอาจจะถือว่าเป็นหน้าดำขาวได้ ทั้งนี้เพื่อประโยชน์ในการแพร่ที่สามารถซอกซอนเข้าไปในช่องว่างแคบๆ ได้ หรือแพร่เพื่อแบ่งส่วนประกอบที่ติดกันออกจากกันได้ ซึ่งตัวแปรที่ใช้ในการจำแนกความแตกต่างของหน้าดำ คือ ค่าของการเปลี่ยนแปลงระหว่างจุดสีขาวและสีดำ (Transition value) และค่าของผลรวมของจุดสีดำใน 4 ทิศทาง สำหรับตัวอย่างของการแยกประเภทของหน้าดำ เช่น การแบ่งแยกระหว่างกลุ่มของตัวอักษรและรูปภาพ เป็นต้น โดยมีสมมุติฐานในการแบ่งประเภทของหน้าดำดังต่อไปนี้

1. หน้าดำขาว คือ หน้าดำที่มีการเปลี่ยนแปลงระหว่างจุดดำและขาวน้อยมาก หรือแทบไม่มีเลย และมีจำนวนจุดสีดำในหน้าดำน้อยกว่าครึ่งจำนวนจุดสีดำที่กำหนด
2. หน้าดำบริเวณขอบของคอลัมน์ตัวอักษร คือ หน้าดำที่มีค่าการเปลี่ยนแปลงระหว่างจุดดำและขาวน้อย และมีจำนวนจุดสีดำในหน้าดำมากกว่าครึ่งจำนวนจุดสีดำที่กำหนดเล็กน้อย
3. หน้าดำภายในคอลัมน์ตัวอักษร คือ หน้าดำที่มีค่าการเปลี่ยนแปลงระหว่างจุดดำและขาวเยอะ และมีจำนวนจุดสีดำในหน้าดำมากกว่าครึ่งจำนวนจุดสีดำที่กำหนด
4. หน้าดำที่เนื้อหาภายในเป็นรูปภาพ คือ หน้าดำที่มีค่าการเปลี่ยนแปลงระหว่างจุดดำและขาวน้อย เพราะเนื้อหาภายในส่วนใหญ่เป็นสีดำ และมีจำนวนจุดสีดำในหน้าดำมากกว่าครึ่งจำนวนจุดสีดำที่กำหนดเยอะมากๆ

ขั้นตอนแรกก่อนทำการแพร่เราต้องทราบประเภทของหน้าดำก่อนว่าตำแหน่งที่สนใจเป็นหน้าดำประเภทไหน การแยกประเภทหน้าดำจะใช้หลักการสแกนเพื่อเช็คจำนวนจุดสีดำ (Summation of black pixel) และค่าที่ได้จากการเปลี่ยนแปลงระหว่างจุดดำและจุดขาว (Transfer

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

value) ในหน้าต่างทั้ง 4 ทิศทาง ประกอบด้วย แนวตั้ง แนวนอน และแนวทแยงอีกสองทิศทาง โดยการกำหนดตัวแปรต่างๆ ดังนี้

- ผลรวมของจุดสีดำในแนวตั้ง (SV: Summation of black pixel in vertical direction)
- ผลรวมของจุดสีดำในแนวนอน (SH: Summation of black pixel in horizontal direction)
- ผลรวมของจุดสีดำในแนวตั้ง (OL: Summation of black pixel in oblique left direction)
- ผลรวมของจุดสีดำในแนวตั้ง (OR: Summation of black pixel in oblique right direction)
- ผลของการเปลี่ยนแปลงระหว่างจุดขาวและดำในแนวตั้ง (TV: Transition value between black and white in vertical direction)
- ผลของการเปลี่ยนแปลงระหว่างจุดขาวและดำในแนวตั้ง (TH: Transition value between black and white in horizontal direction)
- ผลของการเปลี่ยนแปลงระหว่างจุดขาวและดำในแนวตั้ง (TL: Transition value between black and white in oblique left direction)
- ผลของการเปลี่ยนแปลงระหว่างจุดขาวและดำในแนวตั้ง (TR: Transition value between black and white in oblique right direction)

ในรูปที่ 4.3 แสดงตัวอย่างการหาค่าตัวแปรต่างๆ ที่นำมาใช้ในการจำแนกประเภทหน้าต่าง โดยจะมีการสแกนทั้ง 4 ทิศทาง คือ ทิศแนวตั้ง ทิศแนวนอน และทิศทางทแยงอีก 2 ทิศ ส่วนในรูปที่ 4.4 (ก)-(ง) ในวงกลมและวงรีจะแสดงตัวอย่างการแยกประเภทหน้าต่าง

### The Classification Type of Window (16x32 pixels)

Vertical Direction



Summation of Black Pixel (SV) = 96  
Transition Value Pixel (TV) = 30

Oblique Left Direction



Summation of Black Pixel (OL) = 7  
Transition Value Pixel (TL) = 4

Horizontal Direction



Summation of Black Pixel (SH) = 98  
Transition Value Pixel (TH) = 42

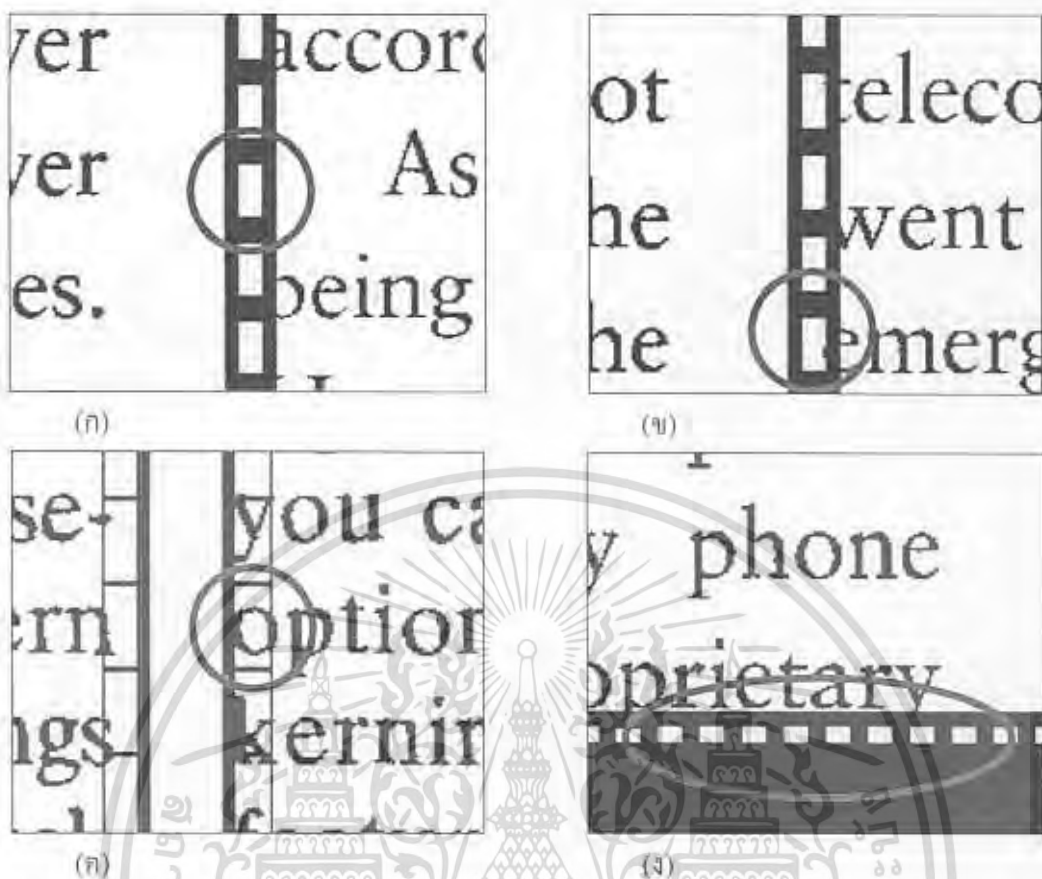
Oblique Right Direction



Summation of Black Pixel (OR) = 7  
Transition Value Pixel (TR) = 4

รูปที่ 4.3 แสดงการค่าตัวแปรต่างๆ ที่ใช้ในการจำแนกประเภทหน้าต่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4 แสดงตัวอย่างของหน้าต่างประเภทต่างๆ

- (ก) แสดงหน้าต่างขาว
- (ข) หน้าต่างขอบของคอลัมน์
- (ค) หน้าต่างตัวอักษร
- (ง) หน้าต่างรูปภาพ

ก่อนการทดลองได้ทำการวัดค่าต่างๆ โดยอ้างอิงจากขนาดของหน้าต่าง เพื่อใช้เป็นข้อมูลในการแบ่งประเภทของหน้าต่าง ซึ่งพอสรุปค่าได้ดังต่อไปนี้

ตารางที่ 4.1 แสดงการจำแนกประเภทของหน้าต่าง

SV	SH	OL	OR	TV	TH	TL	TR	Type of window
$\geq 50$	$\geq 50$	$> 8$	$> 8$	$\leq 16$	$\leq 32$	$\geq 8$	$\geq 8$	Picture
$\geq 50$	$\geq 50$	$\leq 8$	$\leq 8$	$\geq 32$	$\geq 32$	$< 8$	$< 8$	Character
$< 20$	$< 20$	$\leq 2$	$\leq 2$	$< 8$	$< 16$	$\leq 2$	$\leq 2$	White
$< 50$	$< 50$	$\leq 4$	$\leq 4$	$< 16$	$< 32$	$\leq 4$	$\leq 4$	Part of character or Part of symbol

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการทดลองวัดค่าต่างๆ ในภาพเอกสารพบว่าค่าของ OL, OR, TL และ TR ไม่สามารถที่จะนำมาจำแนกความแตกต่างระหว่างประเภทของหน้าต่างไว้ ดังนั้นจึงไม่ได้นำค่าทั้งสี่ค่ามาใช้ ในพื้นที่กรอบสี่เหลี่ยมในรูปที่ 4.5 แสดงผลลัพธ์ที่ได้จากการทดลองวางหน้าต่างโดยการจำแนกประเภทของหน้าต่างก่อนวาง จากผลที่ได้พบว่าค่าที่กำหนดในตารางที่ 4.1 นั้นไม่เพียงพอในการจำแนกประเภทของหน้าต่าง เพราะยังมีข้อมูลจริงบางส่วนมีค่าไม่ตรงตามตารางทำให้จำแนกหน้าต่างผิดประเภทและส่งผลให้การแบ่งไม่ตีความที่คาดหวัง และส่งผลให้การแบ่งมีการกระจายของหน้าต่างทั่วทั้งโครงสร้างที่ต้องการแบ่ง ซึ่งข้อมูลของผลลัพธ์ในการจำแนกประเภทของหน้าต่างบางส่วนของผลการทดลองได้นำเสนอในตารางที่ ก.3 ในภาคผนวก ก

ข้อดีของวิธีการ คือ ต้องมีการเข้าถึงจุดภาพทุกจุดภายในหน้าต่างที่สนใจ และเข้าถึงข้อมูลในระดับพิกเซลของทุกๆ หน้าต่างรอบหน้าต่างที่สนใจใน 8 ทิศทางตามรหัสลูกโซ่ เพื่อวัดระดับของความสัมพันธ์ของหน้าต่างในกลุ่มที่สนใจ ทำให้สามารถเลือกหน้าต่างที่เหมาะสมได้สำหรับการวางหน้าต่างในแต่ละครั้ง ซึ่งถ้าพิจารณาให้ละเอียดจะพบว่าหากข้อมูลมีขนาดใหญ่และมีจุดผิดพลาดเป็นจำนวนมาก การคำนวณอาจต้องการเวลายาวนานพอควรในการเข้าถึงจุดภาพเหล่านั้น

การทดสอบข้อมูลในการจำแนกประเภทของหน้าต่างต้องอาศัยประเภทของข้อมูลภาพเอกสารที่มีความหลากหลาย และต้องใช้ฐานข้อมูลขนาดใหญ่ เพื่อที่จะทดสอบและกึ่งหนดค่าที่เหมาะสมสำหรับค่าเริ่มต้นการทำงานของแต่ละตัวแปรในตาราง ซึ่งการจำแนกประเภทของหน้าต่างให้ถูกต้องและเหมาะสมกับภาพเอกสารทุกๆ ประเภทนั้นอาจจะมีคามผิดพลาด เพราะการจัดรูปแบบของเอกสารแต่ละประเภทนั้นมีความหลากหลาย และเอกสารบางประเภท เช่น สิ่งพิมพ์โฆษณา แผ่นพับ จุลสาร ฯลฯ จะถูกจัดความวัตถุประสงค์ของผู้ออกแบบเป็นหลักคั้งที่ได้กล่าวไว้แล้วในบทที่ 2

จากการทดลองได้ข้อสรุปว่าผลของการจำแนกประเภทหน้าต่างไม่ถูกต้องเนื่องจากข้อมูลจริงส่วนมากมีการซ้อนทับกันหรือปนกันอยู่ โดยเฉพาะอย่างยิ่งหน้าต่างที่เป็นส่วนของตัวอักษรหรือสัญลักษณ์กับหน้าต่างตัวอักษร ซึ่งมักจะให้ค่าข้อมูลของ SV, SH, TV, และ TH สำหรับบางหน้าต่างอาจไม่ตรงตามความเป็นจริง ดังนั้นการระบุประเภทของหน้าต่างนั้นๆ จะต้องมึวิธีการจัดกลุ่มของข้อมูลอย่างมีประสิทธิภาพและลดการเข้าถึงจุดภาพลง เพื่อไม่ให้เกิดความกำกวมหรือซ้อนทับกันของกลุ่มข้อมูลจนยากที่จะแยกประเภทได้และให้เวลาในการประมวลผลที่รวดเร็วขึ้น อันจะส่งผลให้กระบวนการแปรเป็นไปอย่างถูกต้องและมีประสิทธิภาพ



#### 4.4 การแพร่โดยอาศัยพิกัดตำแหน่งเป้าหมาย

จากวิธีที่ผ่านมาพบว่า การแพร่โดยใช้การจำแนกประเภทของหน้าต่างไม่มีประสิทธิภาพ จึงมีแนวคิดที่จะนำวิธีการแพร่ของหน้าต่างขนาด  $16 \times 32$  พิกเซล กลับมาประยุกต์ใช้โดยกำหนดทิศทางและตำแหน่งเป้าหมายที่เหมาะสมให้กับการแบ่งเพิ่มเข้าไป โดยมีแนวคิดว่า ถ้าเราสามารถเลือกตำแหน่งของหน้าต่างที่เหมาะสมประกอบกับการแพร่อย่างมีทิศทางและเป้าหมายปลายทางของหน้าต่างที่ต้องแบ่งลงจะให้ผลลัพธ์ในการแบ่งจุดผิดพลาดดีขึ้น

สำหรับการทดลองยังคงยึดหลักการแพร่ด้วยหน้าต่างขนาด  $16 \times 32$  พิกเซลจากจุดผิดพลาดเช่นเดิม เพราะจากการวิธีการสองวิธีแรก ได้พิสูจน์แล้วว่าหน้าต่างเล็กๆ ไม่ได้เหมาะสมกับการแบ่งส่วนเอกสาร ณ จุดใดๆ เสมอไป การวางหน้าต่างจะเลือกที่วางหน้าต่างในตำแหน่งใดและทิศทางใดนั้นขึ้นอยู่กับค่าความสำคัญ (Priority) ของหน้าต่างเมื่อเทียบกับพิกัดตำแหน่งหน้าต่างที่เป็นเป้าหมาย กล่าวคือให้พิจารณาค่าพิกัดในแนวแกน Y ก่อน ถ้าพิกัดแกน Y เท่ากับจุดเป้าหมายก็จะมาพิจารณาพิกัดแกน X กับจุดเป้าหมาย และจะทำการเลือกพิกัดตำแหน่งของหน้าต่างที่ใกล้เป้าหมายที่สุดในขณะนั้นๆ นอกจากนี้พิกัดตำแหน่งแล้วยังให้หลักการสแกนหาแนวเส้นสีขาวที่มีความหนา 4 พิกเซลในแนวตั้งหรือแนวนอนตามแต่เงื่อนไขกำหนด ซึ่งคาดว่าจะจะเป็นเส้นกั้นแบ่งระหว่างกลุ่มส่วนประกอบในหน้าต่างร่วมด้วย

ผลลัพธ์ที่ได้จากการทดลองบางตัวอย่างแสดงในรูปที่ 4.6 กรอบสี่เหลี่ยมแสดงพื้นที่ของโครงสร้างที่สนใจจะทำการแบ่ง จากพื้นที่ดังกล่าวพบว่าสามารถแบ่งภาพเอกสารได้ถูกต้อง แนวทางของวิธีการนี้ใช้เวลาในการประมวลผลที่ค่อนข้างเร็ว ความถูกต้องค่อนข้างสูง ทั้งยังได้จำนวนย่อหน้าในการแบ่งที่ใกล้เคียงกับความเป็นจริงมากที่สุด อย่างไรก็ตามยังมีข้อผิดพลาดบางประการ คือ ถ้าหากว่าการแบ่งแล้วไม่สามารถไปถึงหน้าต่างเป้าหมายได้ แต่กลับไปเจอกับหน้าต่างที่เป็นส่วนของตัวอักษร แล้วจะให้ผลลัพธ์ในการแบ่งที่ผิดพลาด นั่นคือจะตัดแบ่งเอาส่วนของกลุ่มตัวอักษรที่ควรจะอยู่ย่อหน้าเดียวกันไปไว้ผิดส่วน ดังแสดงในรูปที่ 4.7 ในกรอบสี่เหลี่ยมจะพบว่าปัญหาสำหรับการแพร่คือควรหยุดการวางหน้าต่างเมื่อไร ค่าจำนวนจุดสีดำเท่าไรจึงจะเหมาะสมทำให้มีวางหน้าต่างเข้าไปในเนื้อหาขงภาพ และขณะเดียวกันก็ต้องเป็นค่าที่เหมาะสมที่พอจะแพร่เข้าไประหว่างช่องว่างแคบๆ ของโครงสร้างส่วนประกอบที่มีสัญญาณรบกวนรวมอยู่ด้วยเช่นกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# WORLDWIDE METRO ETHERNET EQUIPMENT REVENUE

## ETHERNET STANDARD FOR THE FIRST MILE

# That long, lonesome last mile

...ly for the "last mile"—the ...  
 ...between a business or a resi...  
 ...the place where it joins up with other ne...  
 ...work traffic? Theoretically, it makes sense...  
 ...to use it, but it will probably take a while...  
 ...to catch on, especially in the United States...  
 ...Most U.S. businesses use Ethernet netw...  
 ...for local-area networks, but when they...  
 ...want to connect those networks to PC...  
 ...in other locations, they have to go through...  
 ...the telephone company. The telephone com...  
 ...pany uses technologies, such as synchro...  
 ...nous optical network (SONET) and asyn...  
 ...chronous transfer mode (ATM), designe...  
 ...d primarily to handle voice. So it has to tr...  
 ...anslate packets from Ethernet networks to...  
 ...ATM packets that can travel over SON...  
 ...nets and then translate them back to E...  
 ...thernet packets when they reach their de...  
 ...stination. It's inefficient and expensive...  
 ...The new standard, expected to be...  
 ...approved this summer, may help Ethere...  
 ...t make some headway in metro-area ne...  
 ...tworks. What's more, the standard's bu...  
 ...s think Ethernet can eventually replac...  
 ...e current residential technologies, such...  
 ...as DSL over copper and cable modems...  
 ...for a broadband connection to the home...  
 ...The Ethernet in the First Mile Allian...  
 ...ce (EFMA), a group of networking vendor...  
 ...has been promoting the standard since...  
 ...2001. (Whether it's the first or the last...  
 ...depends on the direction from which you...  
 ...view it. The phone companies call it the...  
 ...last mile; the Ethernet vendors prefer to...  
 ...call it the first mile.) The group expects the...  
 ...standard to ratify the proposed standard, 802...  
 ...this summer, according to Craig Easley...  
 ...resident of the EFMA and a director at...  
 ...Extreme Networks. The standard spec...  
 ...ifies the protocols and interfaces for us...  
 ...ing Ethernet over access links, for both...  
 ...copper and optical fiber, in the first mile, and puts...  
 ...in place certain quality-of-service spec...  
 ...ifications to ensure that the standard...  
 ...can be used by all carrier class...  
 ...Using Eth...  
 ...ers several ad...  
 ...EFMA...  
 ...Easley, first Ethernet...  
 ...already...  
 ...China

...s are less expensive than...  
 ...ologies, because vendors see...  
 ...more of them. There are two billion Eth...  
 ...ernet ports worldwide, says Easley, com...  
 ...pared to 1 billion ATM ports.

...SONET and ATM. Second, adoption of...  
 ...Ethernet throughout the entire broadband



...network would eliminate the need to...  
 ...translate packets from Ethernet to another...  
 ...standard, such as ATM for phone lines or...  
 ...DOCSIS for cable. This would reduce not...  
 ...only total equipment costs but also the...  
 ...expense of specialized staff to manage the...  
 ...other technologies, says Easley.

...Optical Ethernet, which runs over...  
 ...fiber, promises greater speeds at lower...  
 ...costs than other available technologies...  
 ...The standard is likely to reduce the com...  
 ...plexity and cost of transmitting data fr...  
 ...om the home or the business to the pho...  
 ...ne company's central office. Today phone...  
 ...companies use customized, proprietary

Network topology	Speed
Copper coaxial point-to-point	10 Mbps up to 750 meters or 2 Mbps up to 2,700 meters
Fiber optic point-to-point	100 Mbps and 1,000 Mbps up to 10 kilometers
Fiber optic point-to-multipoint	1,000 Mbps up to 20 kilometers

The 802.3ah standard provides for Ethernet over three different topologies over specific distances. The real-world advantage is in the fiber optic implementations.

...ATM packets and other technologies...  
 ...they could instead use standard equip...  
 ...ment, the cost of equipment...  
 ...could drop...  
 ...However...  
 ...that welcome the reduced costs of...  
 ...running Ethernet from...  
 ...office to...  
 ...head

...the central office, they have little incen...  
 ...to adopt Ethernet for their entire ne...  
 ...twork. In the United States, the incumb...  
 ...carriers own the network and have histo...  
 ...rically been slow to add new service

...they must give their competitors access...  
 ...to the network, says Alex Saunders, pres...  
 ...ident and CEO, Metrobility Optical Sol...  
 ...utions, an optical Ethernet vendor. Pho...  
 ...n companies also want to wring as muc...  
 ...revenue as they can from their cur...  
 ...equipment, adds Easley. The phone...  
 ...companies are quick to change.

...What's more, optical Ethernet...  
 ...is somewhat of a world history. Metro...  
 ...bility vendor Yipes made a splash in 200...  
 ...and 2001, when it built its own fiber opt...  
 ...network specifically to sell Ethernet to...  
 ...businesses in 14 U.S. markets. Then th...  
 ...telecom bubble burst and the compa...  
 ...went into Chapter 11. It has since...  
 ...emerged from bankruptcy and is getting...  
 ...back on the path to profitability...  
 ...according to its new CEO, John P...

...Asia is the first place optical Eth...  
 ...ernet is being widely deployed, says Michael...  
 ...Howard, principal analyst at Infonet...  
 ...Research. Because much of China does...  
 ...n't have a legacy infrastructure, it ca...  
 ...start from scratch and install new-gene...  
 ...ration fiber networks, notes Kamra

...Shantizadeh, chief...  
 ...technology officer...  
 ...and...  
 ...at...  
 ...New...  
 ...Howard believe...  
 ...that U.S. pho...  
 ...n companies wi...  
 ...eventually bu...  
 ...ll-scale Eth...

...equipment or at least suppl...  
 ...ment their existing networks, because it...  
 ...going to be much cheaper to operate a...  
 ...Ethernet network. "We think that in 10...  
 ...to 15 years, a lot of carriers will have don...  
 ...e it." The new standa...  
 ...will...  
 ...China

รูปที่ 4.6 การแพร่โดยกำหนดเป้าหมาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## First Looks

### Publisher 95

Continued from page 67  
to impart the correct look to your emerging document, you select a classic, jazzy, or modern type, enter some text for headings and personal data, and watch Publisher 95 create a document with sample graphics and text.

Just click on theummy text and type to replace it with your own. To replace sample graphics, you right-click and choose the replace option that leads to Publisher 95's 16MB clip-art collection. Other options let you make the text flow around

the image. Alternatively, because Publisher 95 is window-compliant, you can scan a page directly into the program.

You can rotate text boxes to any angle by clicking on a toolbar and following the prompts. We found that with some display drivers, if you create your own text box instead of using one created by a wizard, Publisher 95 turns the background of the rotated text dark blue so you can't read your words; the blue background doesn't print, though.

Everything in Publisher 95 is set up to let beginners use design secrets normally reserved for experts. For instance, when you select an object or an object handle, the mouse arrow displays a word and icon that tells you whether your next action will move, resize, or rotate the object. Preset color palettes include 12 palettes of six coordinated colors and a shade of gray. Publisher 95's pop-up tooltips give the name of the

### ALSO AVAILABLE

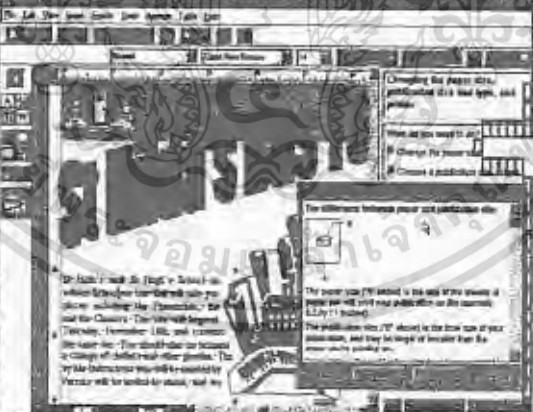
The menus use plain English terminology instead of jargon. Fancy first letter" instead of "drop caps" and "Spacing between characters" instead of "letter-spacing." If you choose the letter-spacing option, you'll be able to change the setting for either the current paragraph

or the five preset spacing options. The word and character spacing in Publisher 95 isn't ideal, but you can improve on it with an option that uses font-based kerning data. The drop-cap feature produces letters that appear slightly mispositioned when but print correctly.



**AIRBORNE:** Publisher 95 will print paper airplanes using with folding directions.

of designing brochures. Publisher 95 uses the feature that advanced programs call "master pages" and that Publisher 95 calls Background. This feature creates template-style backgrounds for all pages in a publication, with graphics, running heads, column layouts, and other features. You can now customize column grids on the background pages—a feature usually found only in more expensive



**DEMONSTRATIVE:** Publisher 95's help screens include detailed instructions.

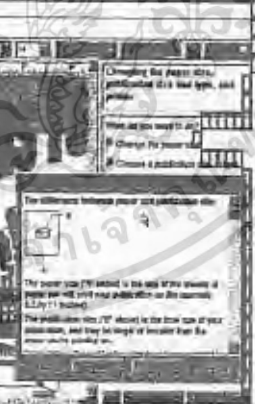
packages—and switch off the template for individual pages. Publisher 95 does not let you create individual background pages or different

backgrounds, but it includes a percentage of Windows 95's system-wide color-control features so you can tell Publisher 95 to warn you when your printer can't produce a requested color. The program includes a gallery of preset objects that you can use to display quotes (or callouts), partially created tables of contents and other features. Publisher 95's most welcome feature is its ability to add objects that you create yourself, so you can establish

### HOW OLD AND NEW

Like most Microsoft products, Publisher 95 is built partly from modules from other programs, some of which are more advanced than others. For example, the text-editing engine is a reduced version of Microsoft Word, though it can handle paragraph styles, bullets and numbering and automatic replacement of two hyphens with an *em dash* without a problem. Publisher 95, unlike Word, isn't smart enough to replace the hyphens with the shorter *en dash* when there are spaces on either side of the hyphen.

**DRIVING OTHER SOFTWARE:** From the standard Microsoft Draw OLE server, which is good



enough for simple line drawings, you click on the Help menu's "How to" item, a dialog box tells you that the help topic doesn't exist.

throughout the program to create shaded and distorted font effects badly needs updating. Unlike everything else in the program, it doesn't display tooltips and its color menus don't correspond with the color palette you select.

### THE BOTTOM LINE

Unfortunately, Publisher 95 lets you edit one document at a time. Instead of automatic timed backups, the program pops up a reminder to save at intervals that you specify. The reminder dialog box doesn't warn you that if you save the file, you'll lose your previous action.

The help system is highly detailed—and intrusive—system in Microsoft Works. Unless you switch it off, a large list of help topics obscures the right-hand quarter of the window, and displaying the help index takes up another quarter of the screen. Whenever you do something unusual or complicated, a pop-up box appears at the cursor location, giving you hints about what to do next. Happily, you can turn this pop-up box off, but you'll have to go deep into the Options menu to do so.

The message boxes are provided somewhat awkwardly between Publisher 95's own tabbed dialog box of recently used files and a standard Windows File Open dialog box that you can access from the menu. Both dialog boxes include preview screens, but you can only use Windows 95's right-click file management pop-up menus from the standard dialog box.

Publisher 95's earlier version established its reputation as the easiest-to-use desktop publishing package. This new version goes further in that direction and adds so many advanced features that DTP novices will find it flexible, powerful, and enjoyable.

**Microsoft Publisher CD-ROM for Windows 95.** List price: \$79.95. Requires: 6MB RAM, 2MB hard disk space, Microsoft Windows 95 or Windows NT. Microsoft Corp., Redmond, WA, 200-426-9400, 206-882-8090, 206-936-7329.

## รูปที่ 4.7 ปัญหาของการแพร่โดยกำหนดเป้าหมาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.5 การแบ่งซ้ำตามแนวแกน XY

แนวทางสุดท้ายได้แนวคิดมาจากการแบ่งซ้ำแนวแกนซึ่งวิธีการนี้ได้แสดงไว้แล้วในบทที่ 3 และเป็นวิธีการที่ตรงไปตรงมา ข้อดีหลักของวิธีการนี้คือใช้เวลาในการคำนวณค่อนข้างจะน้อย ประกอบกับวิธีการนี้สามารถที่จะทำงานได้ดีกับภาพเอกสารที่มีการจัดแบ่งคอลัมน์ชัดเจน โดยเฉพาะอย่างยิ่งภาพเอกสารที่มีรูปแบบโครงสร้างเป็นแบบสี่เหลี่ยม โดยมีแนวความคิดว่า ถ้าสามารถแทรกเข้าไปในระหว่างจุดผิดพลาดด้วยข้อมูลที่มีความละเอียดสูงและมีการกำหนดทิศทาง การแบ่ง และกำหนดระดับค่าความเหมาะสมให้กับการวางแต่ละหน้าต่างเพื่อแบ่งซ้ำได้ น่าจะให้ผลในการแบ่งที่ดีขึ้น

วิธีนี้จะนำการแบ่งซ้ำในแนว XY เข้ามาช่วยปรับปรุงภาพที่ได้จากการแบ่งส่วนประกอบใหญ่ ให้มีความถูกต้องมากขึ้น ซึ่งจะเน้นในส่วนของคอลัมน์เพื่อที่จะแก้ปัญหาแบ่งคอลัมน์ที่ติดกันผิดพลาด โดยอาศัยการตัดภาพฉายเงาทั้งแนวแกน X และแนวแกน Y ณ จุดผิดพลาดหรือระหว่างโครงสร้างที่ติดกันของส่วนประกอบ ซึ่งผลลัพธ์จากการตัดภาพฉายเงาจะ ได้จะ ได้เป็นขอบเขตของบล็อกสี่เหลี่ยมย่อยๆ และสามารถแยก โครงสร้างที่ติดกันออกจากกันได้เป็นอย่างดี

การทำงานจะเริ่มการระบุจุดผิดพลาด ถ้าจุดนั้น ได้รับการยืนยันว่าเป็นจุดผิดพลาดที่แท้จริงก็จะดำเนินการแบ่งความถูกต้องลักษณะเฉพาะของจุดผิดพลาดนั้นๆ ซึ่งโดยทั่วไปจะสามารถแบ่งได้สองลักษณะทั้งแนวนอนและแนวตั้ง โดยเพื่อความถูกต้องการแบ่งแต่ละแนวจะมีการกำหนดหน้าต่างที่เป็นเป็นเป้าหมายปลายทางในการแบ่ง และเงื่อนไขที่ใช้จบการทำงาน คือ ย่อหน้าใดๆ ที่มีจุดผิดพลาดได้รับการแบ่งจุดผิดพลาดจนครบทุกจุดแล้ว และทุกย่อหน้าในภาพเอกสารถูกดำเนินการจนเสร็จสิ้น

จากการทดลองพบว่าวิธีการดังกล่าวสามารถควบคุมจุดผิดพลาดได้ดี และใช้เวลาในการประมวลผลน้อย โดยตัวอย่างของผลลัพธ์ที่ได้จากการทดลองแสดงในรูปที่ 4.8 พื้นที่ที่สนใจจะถูกล้อมรอบด้วยกรอบสี่เหลี่ยม ซึ่งพบว่าสามารถที่จะแบ่งแยกระหว่างโครงสร้างของส่วนประกอบแต่ละส่วนได้เป็นอย่างดี และผลจากการแบ่งจะให้ลักษณะของผลลัพธ์เป็นขอบเขตของบล็อกสี่เหลี่ยมย่อยๆ ดังจะเห็นได้จากตอนล่างของภาพเอกสารมีรูปภาพที่วางระหว่างคอลัมน์ที่ 2 และ 3 ผลจากการแบ่งจุดผิดพลาดจะให้ส่วนของรูปภาพและส่วนประกอบที่เป็นกลุ่มของตัวอักษรในบล็อกสี่เหลี่ยมขนาดเล็กแยกจากกัน สำหรับข้อด้อยของวิธีการ คือ หากจุดผิดพลาดมีวิถึรูปทรงใดๆ ที่ไม่ใช่สี่เหลี่ยมจะทำให้การดำเนินการแบ่งตามวิธีการนี้ไม่สามารถที่จะทำได้ เพราะการแบ่งในแนว XY สามารถที่จะดำเนินการได้เฉพาะกับเอกสารที่มีจุดผิดพลาดเป็นสี่เหลี่ยมเท่านั้น



## บทที่ 5

# ผลการทดลอง วิเคราะห์ และสรุปผลการทดลอง

ในบทนี้จะกล่าวถึงผลที่ได้จากการทดลองตามวิธีการคิดตามเส้นขอบและการแบ่งในแนว ที่ได้กล่าวมาแล้วในบทที่ 3 และ 4 การเปรียบเทียบผลการทดลองกับวิธีการอื่นที่ใช้ในการแบ่งส่วนประกอบเอกสาร พร้อมทั้งวิเคราะห์และสรุปผลการทดลอง

### 5.1 ผลการทดลอง

สำหรับข้อมูลที่นำมาทดลองประกอบด้วยภาพเอกสาร 3 ประเภท คือ ภาพของกลุ่มข้อความเพียงอย่างเดียว (Text) ภาพกลุ่มข้อความรวมอยู่กับรูปภาพ (Text and Picture) และภาพกลุ่มข้อความรวมกับรูปภาพที่มีโครงสร้างของส่วนประกอบภายในติดกัน (Connected Structure) โดยภาพที่ใช้ในการทดสอบแต่ละประเภทจะแบ่งออกเป็น 3 ขนาด คือ ภาพขนาดเล็ก (800 x 800 – 1600 x 2200 พิกเซล) ภาพขนาดกลาง (1200 x 1800 – 1500 x 2500 พิกเซล) และภาพขนาดใหญ่ (2000 x 3000 – 2600 x 3500 พิกเซล) ซึ่งภายในแต่ละขนาดยังรวมภาพเอกสารที่ไม่สามารถประมวลผลได้เนื่องจากข้อจำกัดของวิธีการอีกจำนวนหนึ่ง รวมรูปภาพที่ใช้ในการทดสอบทั้งหมด 212 ภาพ ที่มาของข้อมูลรูปภาพส่วนหนึ่งได้จากรายชื่อของงานวิจัยเดิม [1] และอีกส่วนได้จากการสแกนภาพเอกสารและจัดเตรียมข้อมูลขึ้นมาใหม่ แหล่งที่มาของภาพเอกสารส่วนใหญ่ นำมาจาก หนังสือตำราเรียน วารสารทางวิชาการ หนังสือพิมพ์ แผ่นพับ ฯลฯ โดยภาพเอกสารที่ใช้ในงานวิจัยเป็นเอกสารที่ไม่มีกรอบถูกสแกนด้วยความละเอียด 300 จุดต่อนิ้ว มีการแก้ไขความเอียง และจัดเก็บในรูปแบบไฟล์บีตแมป (ภาพขาวดำ)

ในการทดลองจะวัดความเร็วในการประมวลผลของอัลกอริทึม จำนวนจุดภาพทั้งหมดที่ถูกใช้ในการประมวลผลของแต่ละภาพเอกสาร และความถูกต้องในการแบ่งภาพเอกสารที่ประเมินจากเปอร์เซ็นต์ของความถูกต้องในการแบ่ง โดยการวัดเวลาของการประมวลผลเริ่มตั้งแต่การเดินวนรอบส่วนประกอบใหญ่จนกระทั่งแบ่งส่วนประกอบของเอกสารเสร็จสิ้นทุกขั้นตอนที่ได้กล่าวไว้ในหัวข้อที่ 3.2 ซึ่งจะถูกวัดในหน่วยวินาที (Second) ในตารางที่ 5.1 แสดงตารางสรุปผลการทดลอง โดยภาพเอกสารจะแบ่งเป็นสามประเภท คือ เอกสารขนาดเล็ก (Small หรือ 'S') กลาง (Medium หรือ 'M') และใหญ่ (Large หรือ 'L') ชนิดของเอกสารแต่ละขนาดประกอบไปด้วย 3 ชนิดด้วยกัน คือ ภาพตัวอักษรอย่างเดียว ('T') รูปภาพกับตัวอักษร ('T&P') และ รูปภาพกับตัวอักษรที่มีโครงสร้างที่ติดกัน ('C') ส่วนภาพเอกสารขนาดใหญ่จะเพิ่มเอกสารที่มีโครงสร้างเป็นสี่เหลี่ยม ('R') เข้ามาเพื่อเปรียบเทียบผลลัพธ์ที่ได้ ซึ่งเวลาเฉลี่ยที่ใช้ในการประมวลผลของวิธีการจะแสดงในหน่วยของวินาที และจำนวนจุดภาพที่เข้าถึง โดยเฉลี่ยจะแสดงในหน่วยพิกเซล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.1 แสดงผลการทดลองด้วยวิธีการวิธีการการติดตามเส้นขอบและการแบ่งในแนว XY

ขนาดภาพ เอกสาร	ชนิดภาพ	จำนวนภาพ (ภาพ)	เวลาเฉลี่ยในการ ประมวลผลจาก งานวิจัยเดิม (วินาที)	เวลาเฉลี่ยในการ ประมวลผล (วินาที)	จำนวนจุดโดยเฉลี่ยที่ เข้าถึง (พิกเซล)
S	T	6	0.223	0.065	569,218
	T&P	6	0.218	0.074	808,780
	C	4	0.233	0.070	469,016
M	T	6	0.494	0.180	2,272,318
	T&P	13	0.496	0.154	1,544,275
	C	4	0.508	0.146	1,304,385
L	T	50	1.085	0.327	3,795,709
	T&P	52	1.131	0.320	3,649,575
	C	47	1.216	0.356	2,709,402
	R	90	1.106	0.310	3,343,381

## 5.2 เปรียบเทียบวิธีการอื่นๆ

### 5.2.1 การเปรียบเทียบเวลาและจำนวนจุดภาพที่ใช้ในการแบ่งส่วนประกอบ

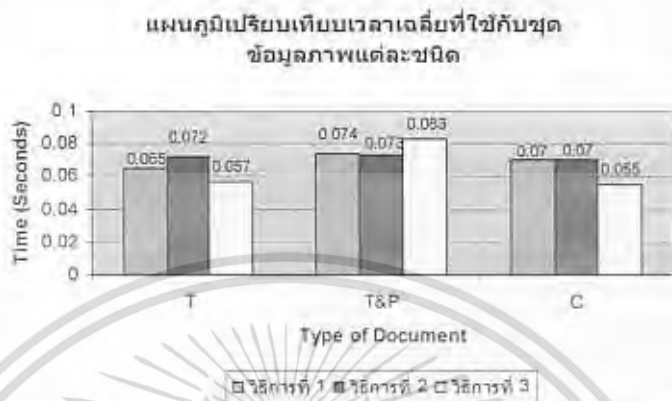
เนื้อหาในส่วนนี้จะแสดงแผนภูมิเปรียบเทียบผลการทดลองระหว่าง 3 วิธีการด้วยกันอันได้แก่ วิธีการที่ 1 คือ การติดตามเส้นขอบและการแบ่งในแนว XY ซึ่งเป็นวิธีการที่ได้นำเสนอในวิทยานิพนธ์ วิธีการที่ 2 คือ วิธีการแบ่งจำด้วยวิธีการแพร่อย่างมีทิศทาง และวิธีการที่ 3 คือ วิธีการแบ่งซ้ำในแนวแกน X-Y สาเหตุที่เลือกวิธีการนี้มาเปรียบเทียบเพราะเป็นวิธีการที่เรียบง่าย สามารถที่จะเขียน โปรแกรม ได้ง่ายเนื่องจากมีอัลกอริทึมที่ง่าย ไม่ซับซ้อนและเป็นที่ยอมรับกันแพร่หลาย โดยในการทดลองจะใช้เวลาเป็นค่าประเมินผลการทำงานของแต่ละวิธีการด้วย แต่อย่างไรก็ตามการใช้เวลาประเมินความเร็วของแต่ละวิธีการนั้นจะขึ้นอยู่กับโครงสร้างข้อมูลและวิธีการในการดำเนินการ (implement) ด้วย ดังนั้นในการทดลองจึงเสนอวิธีนับว่าวิธีการใดมีการเข้าถึงจุดภาพมากน้อยเพียงใดมาเป็นตัวประเมินร่วมด้วย เพราะว่ารายละเอียดและวิธีการในการดำเนินการอาจจะทำให้เวลาผิดพลาดได้ ถึงแม้ว่าจะเป็นวิธีการที่ดีก็ตาม ในแผนภูมิเปรียบเทียบเวลาที่ใช้ในการประมวลผลรูปภาพและจำนวนจุดภาพที่ถูกใช้ในการแบ่งขอบเขตในแต่ละภาพ เริ่มจากเอกสารขนาดเล็ก เอกสารขนาดกลาง และเอกสารขนาดใหญ่ ภายในเอกสารแต่ละขนาดจะมี 2 แผนภูมีย่อย ซึ่งเป็นแผนภูมิระหว่างเวลาที่ใช้ในการประมวลผล โดยเฉลี่ยกับชนิดของภาพเอกสาร และจำนวนจุดภาพเฉลี่ยที่ถูกใช้กับชนิดของภาพเอกสาร ในแผนภูมิเปรียบเทียบเวลาที่ใช้ในการประมวลผลภาพนั้นแกน X จะแสดงชนิดของภาพที่ใช้ ส่วนแกน Y จะแสดงเวลาเฉลี่ยที่ใช้สำหรับจุดของภาพเอกสารที่ใช้ทดสอบชนิดนั้นๆ ในหน่วยวินาที ส่วนแผนภูมิเปรียบเทียบการเข้าถึงจุดภาพนั้น แกน

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาค้นคว้าเท่านั้น ไม่อนุญาติให้นำไปใช้ประโยชน์ด้านการค้า

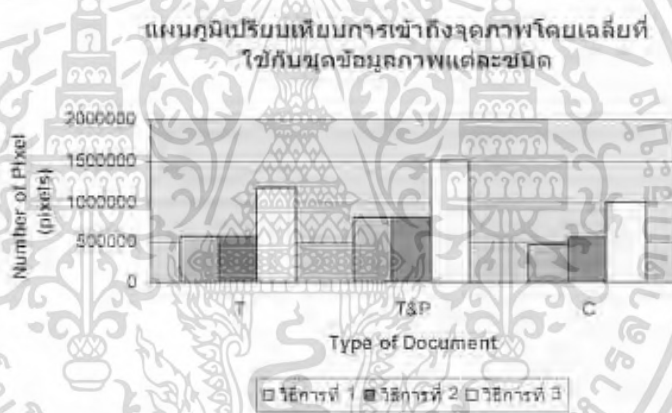
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จะแทนชนิดของภาพที่ใช้ทดสอบเช่นกัน และแกน Y จะแสดงถึงจำนวนจุดภาพเฉลี่ยที่ถูกใช้เพื่อระบุขอบเขตของภาพเอกสารแต่ละชนิด ดังแสดงดังต่อไปนี้

### 1. ภาพเอกสารขนาดเล็ก

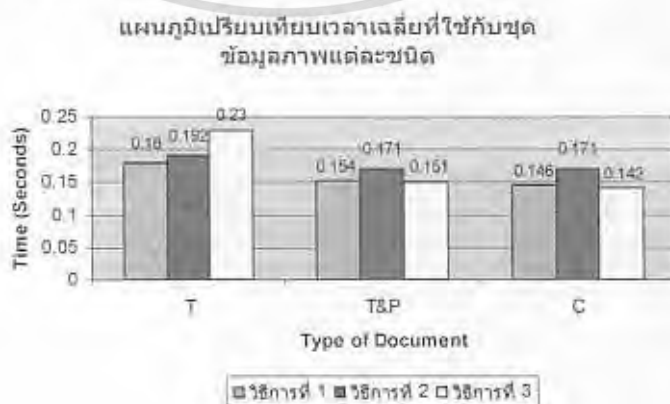


รูปที่ 5.1 แสดงผลการเปรียบเทียบเวลาเฉลี่ยที่ใช้ในการประมวลผลของภาพเอกสารแต่ละชนิด



รูปที่ 5.2 แสดงผลการเปรียบเทียบจำนวนจุดภาพที่เข้าถึงในการประมวลผลของภาพเอกสารแต่ละชนิด

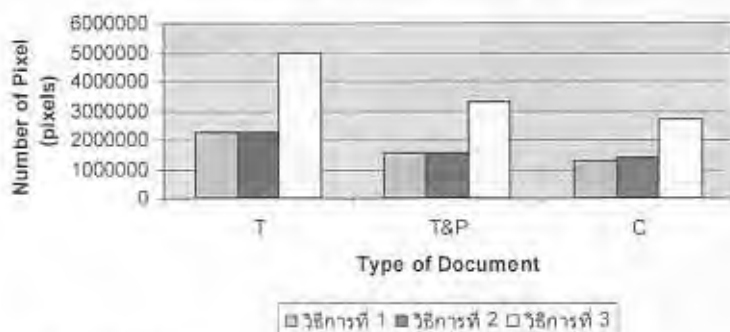
### 2. ภาพเอกสารขนาดกลาง



รูปที่ 5.3 แสดงผลการเปรียบเทียบเวลาเฉลี่ยที่ใช้ในการประมวลผลของภาพเอกสารแต่ละชนิด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

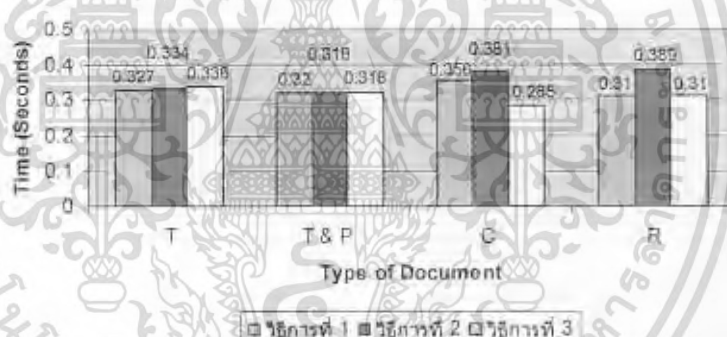
แผนภูมิเปรียบเทียบการเข้าถึงจุดภาพโดยเฉลี่ยที่  
ใช้กับข้อมูลภาพแต่ละชนิด



รูปที่ 5.4 แสดงผลการเปรียบเทียบจำนวนจุดภาพที่เข้าถึงในการประมวลผลของภาพเอกสารแต่ละชนิด

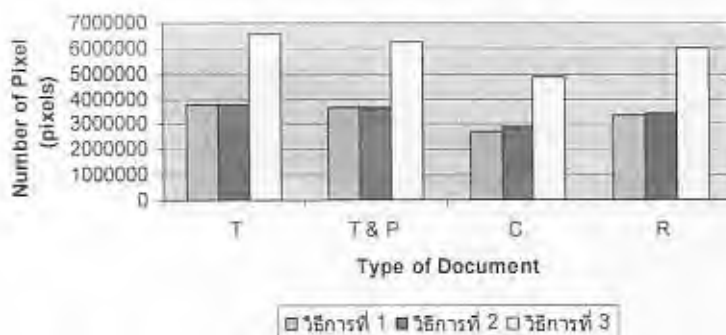
3. ภาพเอกสารขนาดใหญ่

แผนภูมิเปรียบเทียบเวลาเฉลี่ยที่ใช้กับชุด  
ข้อมูลภาพแต่ละชนิด



รูปที่ 5.5 แสดงผลการเปรียบเทียบเวลาเฉลี่ยที่ใช้ในการประมวลผลของภาพเอกสารแต่ละชนิด

แผนภูมิเปรียบเทียบการเข้าถึงจุดภาพโดยเฉลี่ยที่  
ใช้กับชุดข้อมูลภาพแต่ละชนิด



รูปที่ 5.6 แสดงผลการเปรียบเทียบจำนวนจุดภาพที่เข้าถึงในการประมวลผลของภาพเอกสารแต่ละ

ชนิด เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากผลการทดลองจากแผนภูมิของเวลาและจำนวนจุดที่ถูกใช้ในการแบ่งภาพเอกสารในรูปที่ 5.1-5.6 สามารถสรุปผลการทดลองได้ว่าเวลาที่ใช้ในการประมวลผลทั้งสามวิธีการค่อนข้างจะใกล้เคียงกัน โดยวิธีการแบ่งซ้ำในแนวแกน X-Y จะเร็วกว่าวิธีการอื่นเล็กน้อย แต่ในเรื่องของการเข้าถึงจุดภาพวิธีการแบ่งซ้ำในแนวแกน X-Y จะมีการเข้าถึงจุดภาพมากกว่าวิธีการอื่น ซึ่งวิธีการที่ได้นำเสนอในวิทยานิพนธ์ใช้เวลาในการประมวลผลใกล้เคียงกับวิธีการแบ่งซ้ำในแนวแกน X-Y มีสาเหตุมาจากอัลกอริทึมที่ใช้มีการเปลี่ยนแปลงโครงสร้างข้อมูลมาก เช่น โครงสร้างของย่อหน้า (paragraph structure) ที่เป็นลิงก์เน็ต ถ้ามีการแก้ไขโครงสร้างข้อมูลให้กับข้อมูลเท่าที่จำเป็นน่าจะทำให้เวลาของวิธีการที่ได้นำเสนอรวดเร็วยิ่งขึ้น

## 5.2.2 การประเมินผลลัพธ์เชิงปริมาณในการแบ่งส่วนประกอบภาพเอกสาร (Quantitative evaluation of segmentation results)

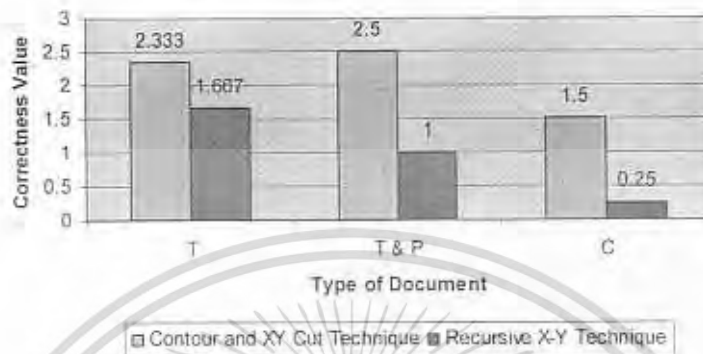
การทำงานวิจัยที่ดเนินจำเป็นต้องมีการประเมินผลการทำงานเสมอ เพื่อให้เกิดความน่าเชื่อถือของข้อมูลงานวิจัย สำหรับในการแบ่งส่วนประกอบภาพเอกสารก็เช่นกันต้องมีการประเมินผลเพื่อนำเสนอและเปรียบเทียบกับงานวิจัยอื่น ซึ่งการประเมินผลเชิงปริมาณน่าจะแสดงให้เห็นถึงประสิทธิภาพของงานวิจัยเมื่อเปรียบเทียบกับงานวิจัยอื่น ได้ดีกว่าหรือยกตัวอย่างจากผลลัพธ์ของประมวลผลที่ดีเพียงไม่กี่ภาพ ดังนั้นในวิทยานิพนธ์นี้จึงนำเสนอการเปรียบเทียบเชิงปริมาณระหว่างวิธีการที่นำเสนอกับการวิธีการแบ่งซ้ำตามแนวแกน X-Y โดยแผนภูมิเปรียบเทียบข้อแตกต่างระหว่างค่าของตัวแปร 3 ตัวด้วยกัน ซึ่งได้แก่ ค่าความถูกต้องเฉลี่ย (Correctness), ค่าผิดพลาดเฉลี่ย (error) และค่า Average Normalize โดยแกน X จะแสดงชนิดของภาพที่ใช้ ส่วนแกน Y จะแสดงค่าความถูกต้องโดยเฉลี่ย ค่าความผิดพลาดเฉลี่ย และค่า Average Normalize ตามลำดับค่าทั้ง 3 จะนับจากจำนวนส่วนประกอบโดยเฉลี่ยที่ได้จากการทดลองกับชุดของภาพเอกสารแต่ละชนิดดังแสดงดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 1. ภาพเอกสารขนาดเล็ก

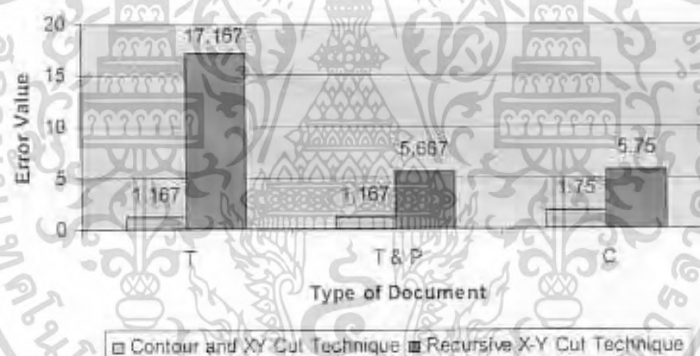
ในรูปที่ 5.7-5.9 แสดงแผนภูมิการเปรียบเทียบค่าความถูกต้องเฉลี่ย (Average Correctness) ค่าผิดพลาดเฉลี่ย (Average error) และค่า Average Normalize ของภาพเอกสารขนาดเล็ก

แผนภูมิแสดงค่าความถูกต้องเฉลี่ย



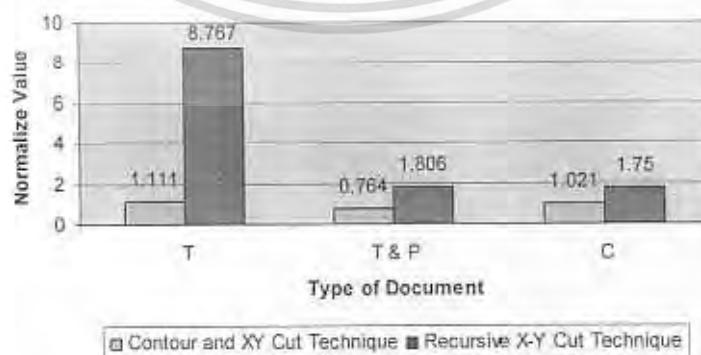
รูปที่ 5.7 แสดงแผนภูมิเปรียบเทียบค่าความถูกต้องเฉลี่ย

แผนภูมิแสดงค่าผิดพลาดเฉลี่ย



รูปที่ 5.8 แสดงแผนภูมิเปรียบเทียบค่าผิดพลาดเฉลี่ย

แผนภูมิแสดงค่า Average Normalize



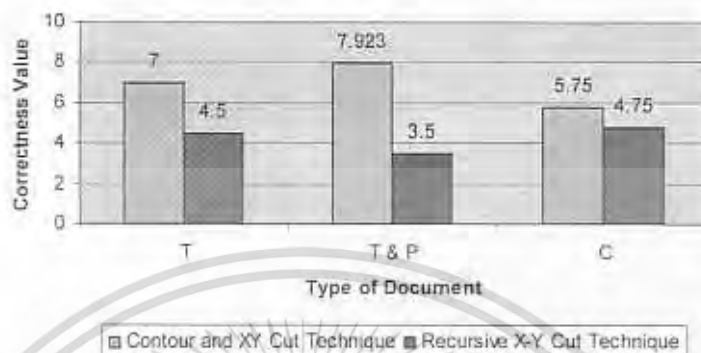
รูปที่ 5.9 แสดงแผนภูมิเปรียบเทียบค่า Average Normalize

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2. ภาพเอกสารขนาดกลาง

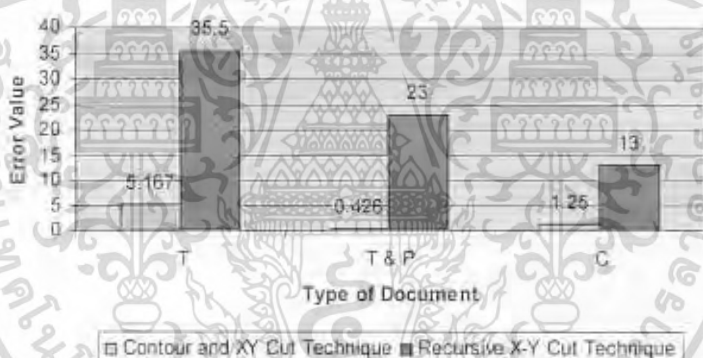
ในรูปที่ 5.10-5.12 แสดงแผนภูมิการเปรียบเทียบค่าความถูกต้องเฉลี่ย (Average Correctness) ค่าผิดพลาดเฉลี่ย (Average error) และค่า Average Normalize ของภาพเอกสารขนาดกลาง

แผนภูมิแสดงความถูกต้องเฉลี่ย



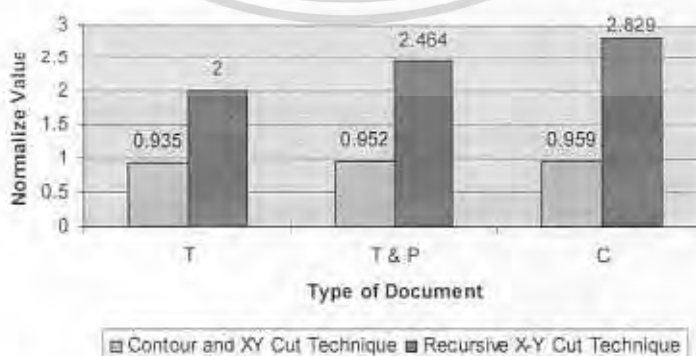
รูปที่ 5.10 แสดงแผนภูมิเปรียบเทียบความถูกต้องเฉลี่ย

แผนภูมิแสดงค่าผิดพลาดเฉลี่ย



รูปที่ 5.11 แสดงแผนภูมิเปรียบเทียบค่าผิดพลาดเฉลี่ย

แผนภูมิแสดงค่า Average Normalize



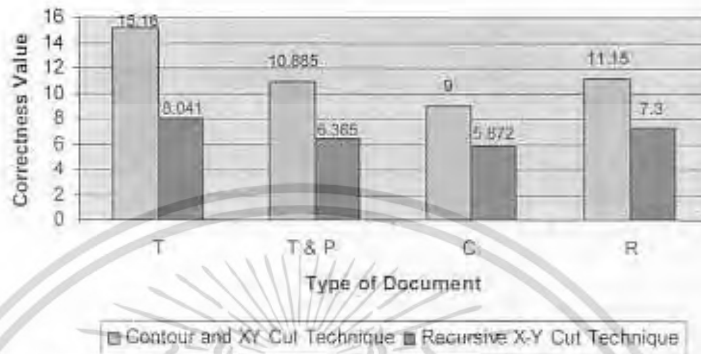
รูปที่ 5.12 แสดงแผนภูมิเปรียบเทียบค่า Average Normalize

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3. ภาพเอกสารขนาดใหญ่

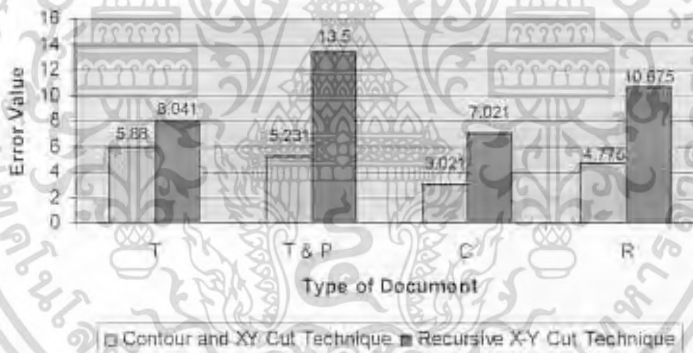
ในรูปที่ 5.13-5.15 แสดงแผนภูมิการเปรียบเทียบค่าความถูกต้องเฉลี่ย (Average Correctness) ค่าผิดพลาดเฉลี่ย (Average error) และค่า Average Normalize ของภาพเอกสารขนาดใหญ่

แผนภูมิแสดงค่าความถูกต้องเฉลี่ย



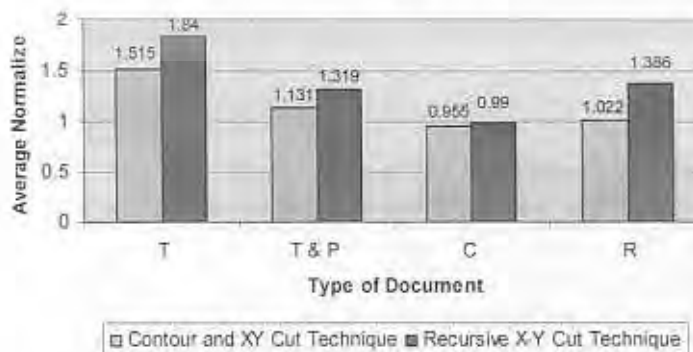
รูปที่ 5.13 แสดงแผนภูมิเปรียบเทียบค่าความถูกต้องเฉลี่ย

แผนภูมิแสดงค่าผิดพลาดเฉลี่ย



รูปที่ 5.14 แสดงแผนภูมิเปรียบเทียบค่าผิดพลาดเฉลี่ย

แผนภูมิแสดงค่า Average Normalize



รูปที่ 5.15 แสดงแผนภูมิเปรียบเทียบค่า Average Normalize

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากผลการทดลองจากแผนภูมิของเวลาและจำนวนจุดที่ถูกใช้ในการแบ่งภาพเอกสารในรูปแบบที่ 5.7-5.115 สามารถสรุปผลการทดลองได้ว่าวิธีการในการติดตามเส้นขอบและการแบ่งในแนว XY ให้ค่าความถูกต้องเฉลี่ยที่สูงกว่า ค่าผิดพลาดเฉลี่ยที่น้อยกว่า และค่า Average Normalize โค้ดก็เคียงกับ 1 มากที่สุด นั่นคือบ่งบอกว่าผลการแบ่งไม่เกิด Over- หรือ Under-Segmentation ส่วนวิธีการแบ่งซ้ำแนวแกน X-Y ให้ค่าทั้ง 3 ในทางตรงข้าม

### 5.3 สรุปผลการทดลอง

ผลการทดลองที่ได้จากการแบ่งภาพเอกสารโดยใช้เส้นขอบและการแบ่งในแนว XY วิธีการแบ่งซ้ำด้วยวิธีการแพรอย่างมีทิศทาง และตัดแบ่งซ้ำในแนวแกน XY จะทำการเปรียบเทียบข้อแตกต่างในแต่ละวิธีการในเรื่องของเวลาที่ใช้ในการประมวลผล จำนวนจุดข้อมูลที่ใช้ในการแบ่งขอบเขตส่วนประกอบ ซึ่งทั้งหมดจะสามารถสรุปได้ในตารางที่ 5.2 ซึ่งจะเปรียบเทียบวิธีการทำงาน จะทำการเปรียบเทียบเวลาที่ใช้ในการประมวลผลและจำนวนจุดภาพที่เข้าถึงสำหรับในแต่ละวิธีการ วิธีการที่ 1 คือ วิธีการการติดตามเส้นขอบและการแบ่งในแนว XY ซึ่งเป็นวิธีที่ใช้ในงานวิจัย วิธีการที่ 2 คือ วิธีการแพรโดยอาศัยพิกัดตำแหน่งเป้าหมาย ซึ่งเป็นวิธีการที่พัฒนาขึ้นมาเพื่อเปรียบเทียบผลกับวิธีการแรก และวิธีการสุดท้าย คือ การแบ่งซ้ำตามแนวแกน XY โดยประเภทของเอกสารจะแบ่งออกเป็น 3 ประเภทดังได้กล่าวมาแล้ว คือ เอกสารขนาดเล็ก (Small หรือ 'S') กลาง (Medium หรือ 'M') และใหญ่ (Large หรือ 'L') ชนิดของเอกสารประกอบด้วย 4 ชนิด คือ ภาพตัวอักษรอย่างเดี่ยว ('T') รูปภาพกับตัวอักษร ('T&P') รูปภาพกับตัวอักษรที่มีโครงสร้างที่ติดกัน ('C') และเอกสารที่มีโครงสร้างเป็นสี่เหลี่ยม ('R') เวลาที่ใช้ในตารางจะเป็นเวลาเฉลี่ยสำหรับภาพเอกสารในประเภทนั้นๆ แสดงในหน่วยวินาที โดยจะแยกตามชนิดของภาพเอกสาร ทั้งยังรวมถึงจำนวนพิกเซล โดยเฉลี่ยสำหรับภาพเอกสาร ในประเภทนั้นๆ ซึ่งแสดงในหน่วยเป็นพิกเซล ซึ่งผลการทดลองโดยละเอียดได้แสดงในภาคผนวก ข

ในตารางสรุปผลการทดลองที่ 5.2 และภาพตัวอย่างจากการทดลองแบ่งขอบเขตของภาพเอกสารพบว่า

1. ถ้าภาพเอกสารที่เป็นตัวอักษรหรือภาพเอกสารที่ประกอบด้วยรูปภาพกับตัวอักษรมีโครงสร้างง่าย ๆ มีการแบ่งแยกระหว่างส่วนประกอบค่อนข้างจะชัดเจนเวลาที่ใช้ในการประมวลผลสำหรับวิธีการที่นำเสนอในวิทยานิพนธ์และวิธีการแพรอย่างมีทิศทางใช้เวลาในการประมวลผลใกล้เคียงกัน ส่วนวิธีการแบ่งซ้ำแนวแกน XY จะใช้เวลาประมวลผลที่รวดเร็วกว่า แต่ใช้จำนวนจุดภาพมากตามไปด้วย ส่วนสองวิธีแรกใช้จำนวนจุดภาพเท่ากัน

2. ถ้าภาพเอกสารที่เป็นตัวอักษรหรือประกอบด้วยภาพและตัวอักษรที่มีโครงสร้างค่อนข้างซับซ้อนแต่ระยะห่างระหว่างส่วนประกอบชัดเจน การประมวลผลด้วยวิธีการแบ่งซ้ำแนวแกน XY จะใช้เวลาน้อยที่สุดเพราะไม่สามารถแทรกเข้าไประหว่างโครงสร้างที่ซับซ้อนได้ รองลงมาคือเอกสารนี้เป็นเอกสารที่สแกนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์การค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีการที่ได้นำเสนอ และวิธีการแพร่จะใช้เวลานานที่สุด สำหรับจำนวนจุดภาพที่ใช้ในการแบ่งส่วนประกอบสองวิธีแรกจะได้จำนวนจุดภาพใกล้เคียงกัน วิธีการแบ่งซ้ำจะใช้จำนวนจุดภาพมากที่สุด

3. ถ้าภาพเอกสารที่ประกอบด้วยภาพและตัวอักษรมีโครงสร้างซับซ้อนและมีโครงสร้างระหว่างส่วนประกอบติดกันวิธีการแบ่งซ้ำแนวแกน XY ใช้เวลาน้อยสุดเพราะไม่สามารถแทรกเข้าไประหว่างโครงสร้างที่ซับซ้อนได้ วิธีการที่ได้นำเสนอจะใช้เวลาประมวลผลอยู่ในระดับกลาง และวิธีการแพร่จะใช้เวลานานที่สุด สำหรับจำนวนจุดภาพที่ใช้ในการแบ่งส่วนประกอบสองวิธีแรกจะได้จำนวนจุดภาพใกล้เคียงกัน วิธีการแบ่งซ้ำแนวแกน XY จะใช้จำนวนจุดภาพมากที่สุด

ตารางที่ 5.2 สรุปผลการทดลองพร้อมทั้งเปรียบเทียบผลการทดลองกับวิธีการอื่นๆ

ประเภทเอกสาร	ชนิดเอกสาร	จำนวนภาพ (ภาพ)	วิธีการ					
			1		2		3	
			เวลาเฉลี่ย (วินาที)	จำนวนทิกจอกเฉลี่ยที่เข้าถึง	เวลาเฉลี่ย (วินาที)	จำนวนทิกจอกเฉลี่ยที่เข้าถึง	เวลาเฉลี่ย (วินาที)	จำนวนทิกจอกเฉลี่ยที่เข้าถึง
S	T	6	0.065	569,218	0.072	571,766	0.057	1,163,285
	T&P	6	0.074	806,780	0.073	808,780	0.083	1,499,898
	C	4	0.070	469,116	0.070	553,404	0.055	987,968
M	T	6	0.180	2,272,318	0.192	2,275,449	0.230	4,951,595
	T&P	13	0.154	1,544,275	0.171	1,544,275	0.151	3,322,257
	C	4	0.146	1,304,385	0.171	1,410,759	0.142	2,694,248
L	T	50	0.327	3,795,709	0.314	3,797,784	0.336	6,540,115
	T&P	52	0.320	3,649,575	0.318	3,649,575	0.318	6,250,985
	C	47	0.356	2,709,402	0.381	2,909,073	0.285	4,866,924
	R	40	0.310	1,343,381	0.389	3,434,214	0.310	6,010,254

ในตารางที่ 5.3 จะเป็นบทสรุปของการประเมินผลสัมฤทธิ์ประสิทธิผลระหว่างวิธีการคิดค้นเส้นขอบและการแบ่งในแนว XY และวิธีการแบ่งซ้ำตามแนวแกน X-Y และผลการทดลองโดยละเอียดจะแสดงในภาคผนวก ข (ตารางที่ ข.6-ข.8) โดยใช้ค่าที่ประเมิน 3 ค่าด้วยกันและค่าเหล่านั้นได้แก่

- Normalize (N) เป็นค่าอัตราส่วนระหว่างผลลัพธ์จากการแบ่งส่วนประกอบภาพเอกสารที่ได้จากวิธีการนั้นๆ กับผลที่ได้จากการแบ่งส่วน โดยใช้ผู้เชี่ยวชาญประเมิน (Human vision)
- Correctness (Cor.) เป็นค่าจำนวนผลลัพธ์ที่ถูกต้องของส่วนประกอบที่ได้จากการแบ่งส่วนประกอบด้วยวิธีการนั้นๆ เทียบกับส่วนประกอบที่ได้จากการแบ่งโดยผู้เชี่ยวชาญ
- error เป็นค่าจำนวนส่วนประกอบที่แบ่งผิดพลาดโดยคิดจากจำนวนผลลัพธ์ในการแบ่งส่วนประกอบที่ได้จากวิธีการนั้นๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.3 แสดงผลสรุปของค่าเฉลี่ยการประเมินผลลัพธ์เชิงปริมาณเมื่อเปรียบเทียบกับวิธีการอื่น

ประเภทเอกสาร	ชนิดเอกสาร	จำนวนภาพ	วิธีการ					
			Contour and XY Cut Technique			Recursive X-Y Cut Technique		
			Cor.	error	N	Cor.	error	N
S	T	6	2,333	1,167	1,111	1,667	17,167	8,767
	T&P	6	2,500	1,167	0,764	1,000	5,667	1,086
	C	4	1,500	1,750	1,021	0,250	5,750	1,750
M	T	6	7,000	5,167	0,935	4,500	35,500	2,000
	T&P	13	7,923	0,426	0,952	3,500	23,000	2,464
	C	4	5,750	1,250	0,959	4,750	13,000	2,829
L	T	50	15,160	5,880	1,515	8,041	17,18	1,840
	T&P	52	10,885	5,231	1,131	6,365	13,500	1,319
	C	47	9,000	3,021	0,955	5,872	7,021	0,990
	R	40	11,150	4,775	1,022	7,300	10,675	1,386

จากตารางที่ 5.3 ซึ่งสรุปผลค่าเฉลี่ยการประเมินผลลัพธ์เชิงปริมาณพบว่าวิธีการที่ได้นำเสนอในวิทยานิพนธ์ให้ค่าความถูกต้องเฉลี่ย ค่าผิดพลาดเฉลี่ย และค่า Average Normalize ที่ดีกว่าวิธีการแบ่งซ้ำตามแนวแกน X-Y

ตารางที่ 5.4 แสดงผลสรุปของเปอร์เซ็นต์ค่าความถูกต้องเฉลี่ยที่ได้จากการประเมินผลลัพธ์เชิงปริมาณเมื่อเปรียบเทียบกับวิธีการอื่น

ประเภทเอกสาร	ชนิดเอกสาร	จำนวนภาพ	วิธีการ	
			Contour and XY Cut Technique	Recursive X-Y Cut Technique
			Percentage of Correctness (%)	Percentage of Correctness (%)
S	T	6	70	50
	T&P	6	55	22,222
	C	4	40	6,667
M	T	6	54,454	38,961
	T&P	13	88,793	41,379
	C	4	88,461	73,077
L	T	50	77,744	44
	T&P	52	77,534	45,324
	C	47	70,149	45,771
	R	40	83,054	54,376

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในตารางที่ 5.4 แสดงการเปรียบเทียบเปอร์เซ็นต์ความถูกต้องเฉลี่ย (Percentage of Correctness) ที่ได้จากการประเมินผลลัพธ์เชิงปริมาณระหว่าง 2 วิธีการ สำหรับเปอร์เซ็นต์ความถูกต้องคำนวณจากสูตร

$$\text{Percentage of Correctness} = \frac{\text{Result of Segmentation from Algorithm}}{\text{Result of Segmentation from Expert}} \times 100$$

ตัวอย่างการคำนวณ ถ้าผลลัพธ์ส่วนประกอบที่ได้จากการแบ่งโดยใช้วิธีการทดสอบมีจำนวนส่วนประกอบเท่ากับ 13 ส่วน และจำนวนส่วนประกอบที่ประเมินโดยใช้ผู้เชี่ยวชาญแบ่งได้ 14 ส่วน ดังนั้นจะได้ว่าเปอร์เซ็นต์ความถูกต้อง คือ

$$\begin{aligned} \text{Percentage of Correctness} &= \frac{13}{14} \times 100 \\ &= 92.857\% \end{aligned}$$

นอกจากนี้แล้วในวิทยานิพนธ์นี้ได้ทดลองแบ่งช่วงเวลาการประมวลผลออกเป็น 3 ช่วง ได้แก่ ช่วงที่ 1 (T<sub>1</sub>) คือ ช่วงเวลาดังแต่เปิดไฟรูปภาพจนกระทั่งเดินวนรอบส่วนประกอบใหญ่เสร็จสิ้น ช่วงที่ 2 (T<sub>2</sub>) คือ ช่วงเวลาในการระบุจุดสีจุดพลุด ช่วงที่ 3 (T<sub>3</sub>) คือ ช่วงเวลาในการดำเนินการแบ่งส่วนประกอบที่ติดกันออกจากกัน และช่วงที่ 3 (T<sub>3</sub>) คือ ช่วงเวลาในการแสดงผลออกหน้าจอ ซึ่งสำหรับภาพเอกสารที่ใช้ทดสอบนั้นคือภาพเอกสารลำดับที่ 38 ในข้อมูลเอกสารขนาดใหญ่ที่มีโครงสร้างติดกัน โดยรายละเอียดต่าง ได้แสดงดังตารางที่ 5.5 เวลาจะถูกแสดงในหน่วยวินาที (S)

ตารางที่ 5.5 แสดงเวลาในการประมวลผลภาพเอกสารในแต่ละช่วง

จำนวนครั้ง	Total Time From Contour and XY Cut Technique (S)	(T <sub>1</sub> ) (S)	(T <sub>2</sub> ) (S)	(T <sub>3</sub> ) (S)	(T <sub>4</sub> ) (S)	Total Time (S)
1	0.340	0.260	0.010	0.010	0.060	0.340
2	0.331	0.260	0.010	0.010	0.070	0.350
3	0.320	0.261	0.010	0.010	0.060	0.341
4	0.320	0.271	0.010	0.010	0.060	0.351
5	0.321	0.270	0.010	0.010	0.070	0.360
6	0.320	0.270	0.010	0.010	0.060	0.350
7	0.311	0.270	0.010	0.010	0.060	0.350
8	0.311	0.281	0.010	0.010	0.060	0.361
9	0.321	0.260	0.010	0.010	0.061	0.341
10	0.331	0.270	0.090	0.010	0.060	0.349
11	0.311	0.270	0.009	0.010	0.070	0.359
12	0.310	0.270	0.010	0.010	0.060	0.350
Avg.	0.321	0.268	0.010	0.010	0.063	0.350

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 5.4 วิเคราะห์ผลการทดลอง

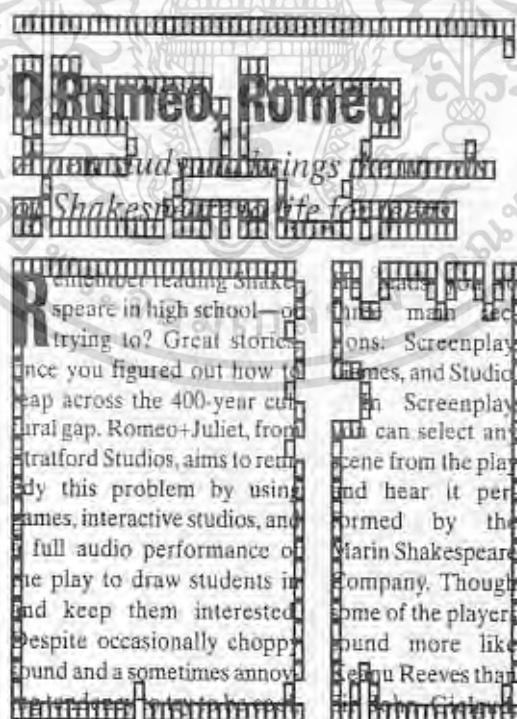
จากผลการทดลองพบว่าข้อมูลภาพที่ใช้ทดสอบมีการกระจายของข้อมูลมาก ข้อมูลในแต่ละชุดมีทั้งภาพเอกสารที่มีโครงสร้างซับซ้อนและไม่ซับซ้อนปะปนอยู่ด้วยกัน ทำให้ผลการทดลองในเรื่องของเวลาต้องพิจารณาโดยละเอียดเปรียบเทียบกันทีละภาพ สำหรับเวลาในการประมวลผลระหว่างสามวิธีการในการแบ่งเอกสารที่มีโครงสร้างติดกันจะพบว่าวิธีการแบ่งข้ามแนวแกน XY ใช้เวลาน้อยสุดจึงน่าจะเป็นวิธีการที่เหมาะสม แต่หากพิจารณารูปทรงที่ได้จากวิธีการแบ่งข้ามแนวแกน XY จะพบว่าได้เฉพาะรูปทรงสี่เหลี่ยมเท่านั้น จึงพิจารณาว่าขาดความเหมาะสมในเรื่องรูปทรงของส่วนประกอบ และมักจะมี Over-Segmentation แต่ถึงอย่างไรก็ตามในบางกรณีที่เกิด Under-Segmentation จะทำให้ต้องเสียพื้นที่หน่วยความจำ พื้นที่ว่างเปล่าที่ไม่ใช่เนื้อหาสาระที่แท้จริง

## 5.5 ภาพตัวอย่างผลลัพธ์จากการประมวลผล

ต่อไปเป็นตัวอย่างบางส่วนที่ได้จากการประมวลผลภาพเอกสารประเภทต่างๆ มีทั้งที่สามารถประมวลผลได้ และภาพที่ไม่สามารถประมวลผลได้สืบเนื่องมาจากข้อจำกัดของวิธีการ ซึ่งสามารถแสดงได้ดังรูปที่ 5.16-5.33 ตามลำดับ

### 5.3.1 ภาพเอกสารขนาดเล็ก (รูปที่ 5.16-5.19)

#### 1. ตัวอย่างการแบ่งแถบแนวรูปภาพตัวอักษรอย่างเดี่ยว

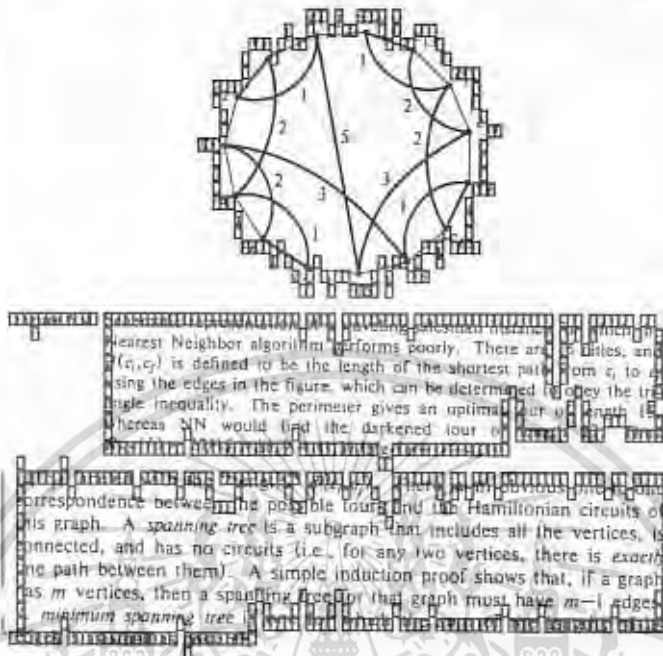


รูปที่ 5.16 ผลการแบ่งภาพเอกสารที่ประกอบด้วยตัวอักษรอย่างเดี่ยว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

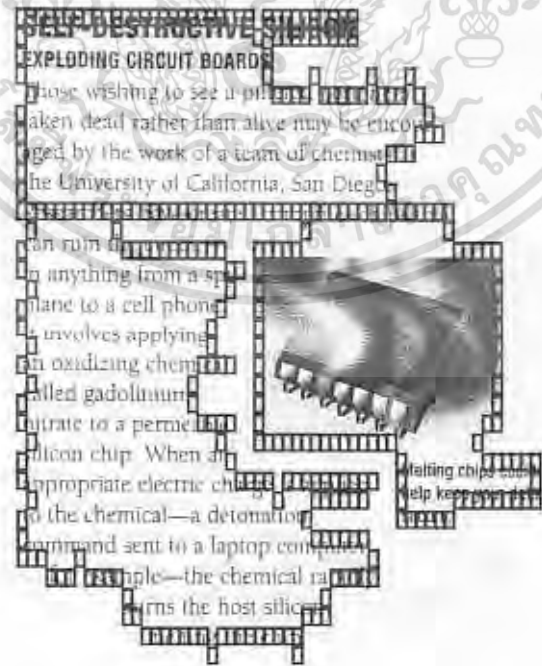
2. ตัวอย่างการแบ่งขอบเขตภาพที่ประกอบด้วยรูปภาพและตัวอักษร

รูปที่ 5.16 ผลการแบ่งภาพเอกสารที่ประกอบด้วยรูปภาพและตัวอักษร



รูปที่ 5.17 ผลการแบ่งภาพเอกสารที่ประกอบด้วยรูปภาพและตัวอักษร

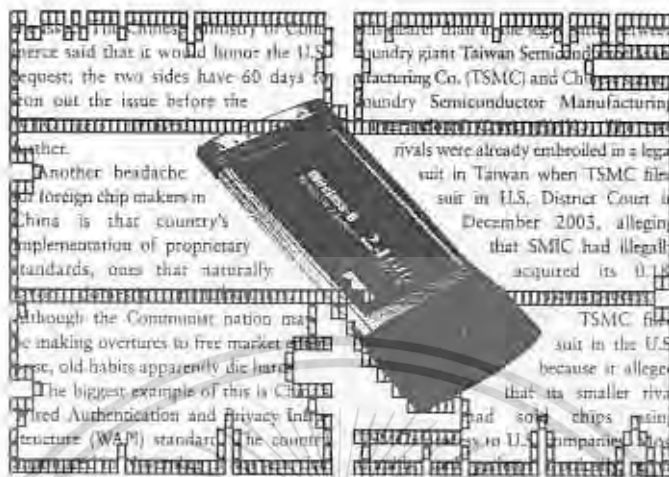
1. ตัวอย่างการแบ่งขอบเขตภาพที่ประกอบด้วยรูปภาพและตัวอักษรที่มีโครงสร้างติดกัน



รูปที่ 5.18 ผลการแบ่งภาพเอกสารที่มีโครงสร้างติดกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

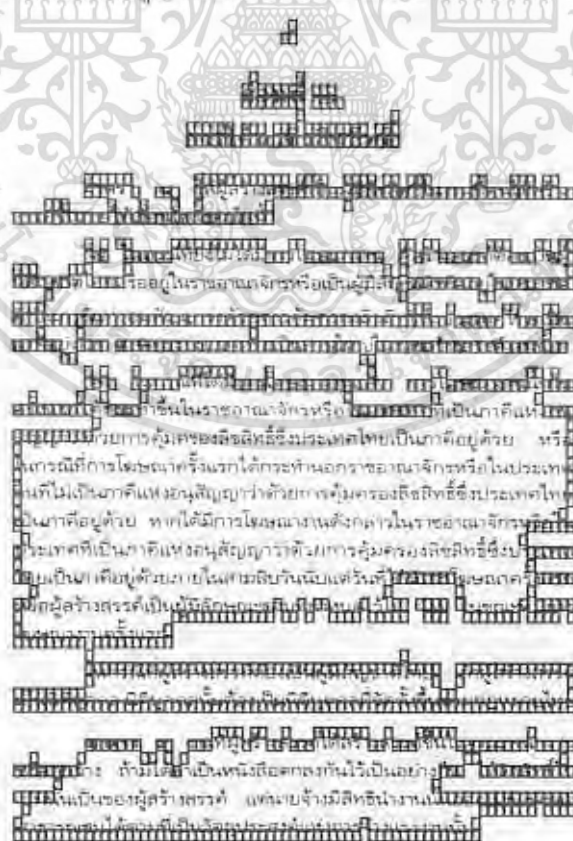
2. ตัวอย่างการแบ่งขอบเขตภาพที่ประกอบด้วยรูปภาพที่ไม่สามารถค้นคืนการได้



รูปที่ 5.19 ผลการแบ่งภาพเอกสารที่ไม่สามารถค้นคืนได้

5.3.2 ภาพเอกสารขนาดกลาง (รูปที่ 5.20-5.23)

1. ตัวอย่างการแบ่งขอบเขตรูปภาพตัวอักษรอย่างเดียว



รูปที่ 5.20 ผลการแบ่งภาพเอกสารที่ประกอบด้วยตัวอักษรอย่างเดียวน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



#### 4. ตัวอย่างการแบ่งขอบเขตภาพที่ประกอบด้วยรูปภาพที่ไม่สามารถดำเนินการได้



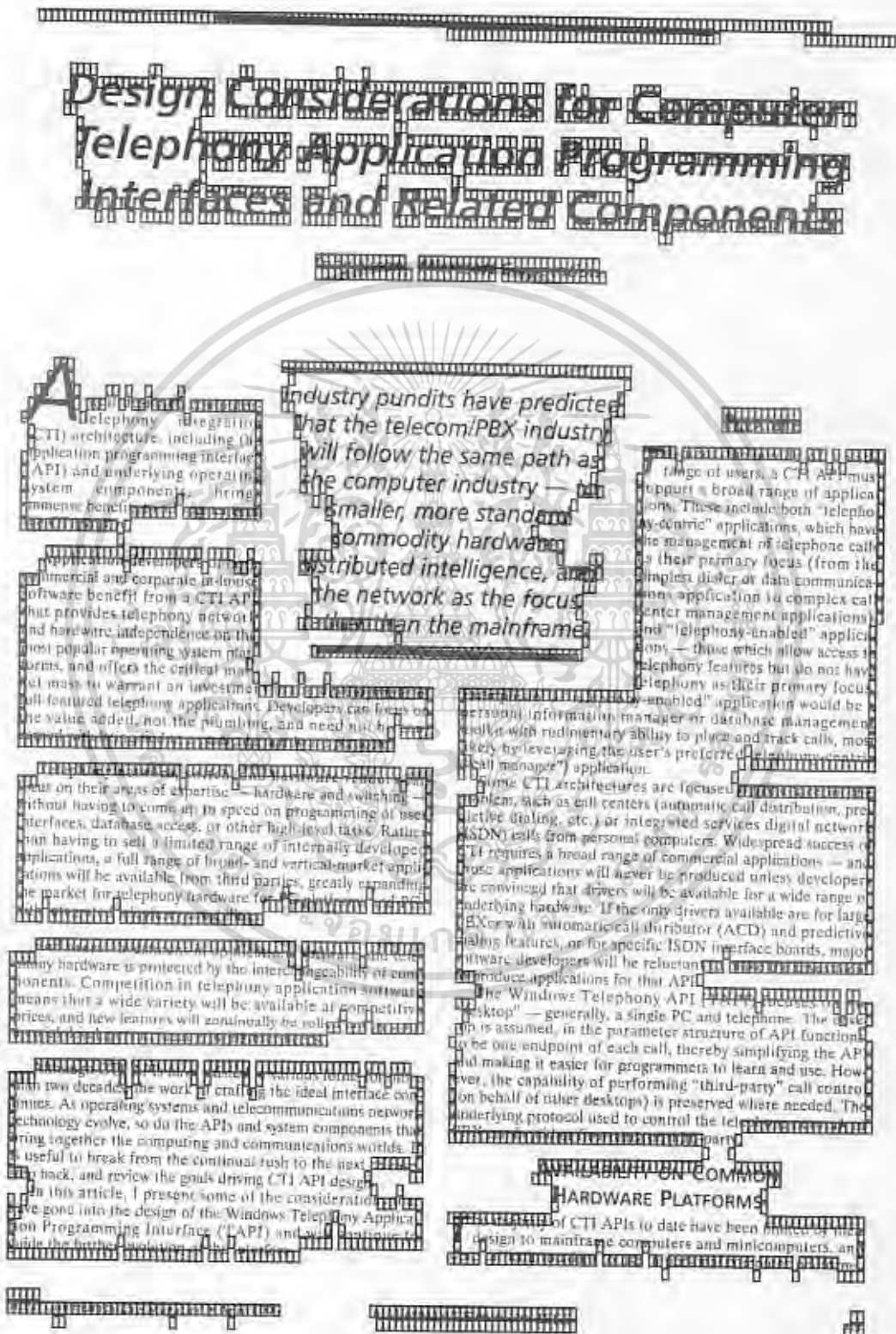
รูปที่ 5.23 ผลการแบ่งภาพเอกสารที่ไม่สามารถแบ่งซ้ำได้

จากรูปที่ 5.23 วรริในรูปแบบแสดงบริเวณที่มีการแบ่งผืนนี้เองมาจากข้อจำกัดของวิธีการที่ใช้ใน  
วิทยานิพนธ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 5.3.3 ภาพเอกสารขนาดใหญ่ (รูปที่ 5.24-5.33)

#### 1. ตัวอย่างการแบ่งขอบเขตรูปภาพตัวอักษรอย่างเดียว (รูปที่ 3.24-5.25)



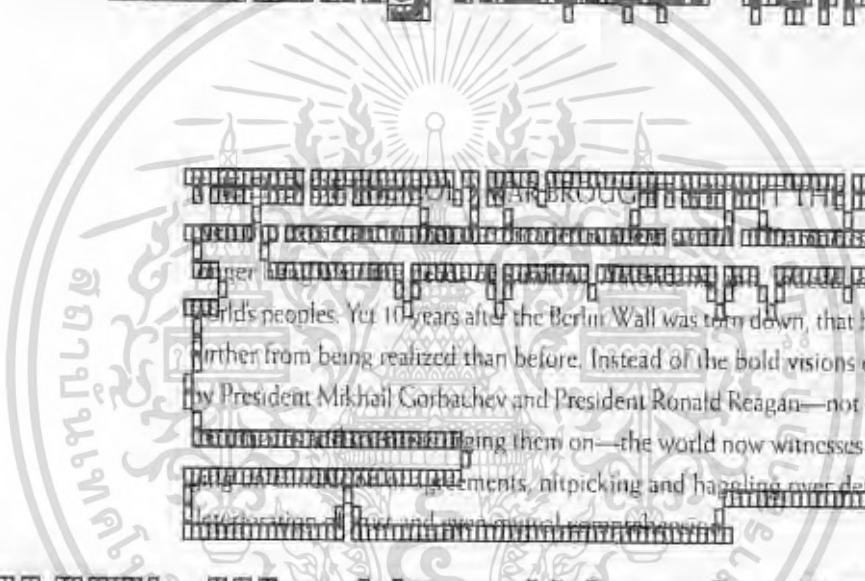
รูปที่ 5.24 ผลการแบ่งภาพเอกสารที่ประกอบด้วยตัวอักษรอย่างเดียว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นับว่าตีพิมพ์ไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารที่ 5.25

THE INSIDIOUS DECAY OF ARMS CONTROL COULD BE REVERSED IF THE TWO COUNTRIES WERE TO INTENSIFY THEIR COOPERATION IN NEW AREAS AND TO TACKLE KEY ISSUES THROUGH PARALLEL ACTIONS

# Mending nuclear fences



...the world's peoples. Yet 10 years after the Berlin Wall was torn down, that hope seems further from being realized than before. Instead of the bold visions enunciated by President Mikhail Gorbachev and President Ronald Reagan—not to speak of being them on—the world now witnesses foot-dragging, nitpicking and haggling over details...

STANFORD UNIVERSITY  
Hoover Institute  
California

...in treaties and tacit understandings during the Cold War is very nearly defunct. The U.S. Senate has rejected the Comprehensive Test Ban Treaty (see *Defining Terms*, p. 36), and the Russian Duma has refused to ratify the second Strategic Arms Reduction Treaty (Sartar 2). Meanwhile, the two countries are sparring acrimoniously over whether to renegotiate the 1972 treaty banning missile defenses, which both Boris Yeltsin and Bill Clinton only in 1997 described as a "cornerstone of arms control"—as did their fellow...

...economic stakes, and clashing values...

Of course, relations between the United States and the Soviet Union during the Cold War period were poisoned by a fundamental ideological hostility. But at the same time both superpowers operated within a common system of geopolitical priorities, so that a set of concepts and doctrines was developed to maintain strategic stability and avoid global nuclear war. Most of these concepts were codified in the strategic nuclear arms control agreements. From the late 1960s onward, these always had the purpose of keeping channels open between adversaries...

11

11

11

รูปที่ 5.25 ผลการแบ่งภาพเอกสารที่ประกอบด้วยตัวอักษรอย่างเดียว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

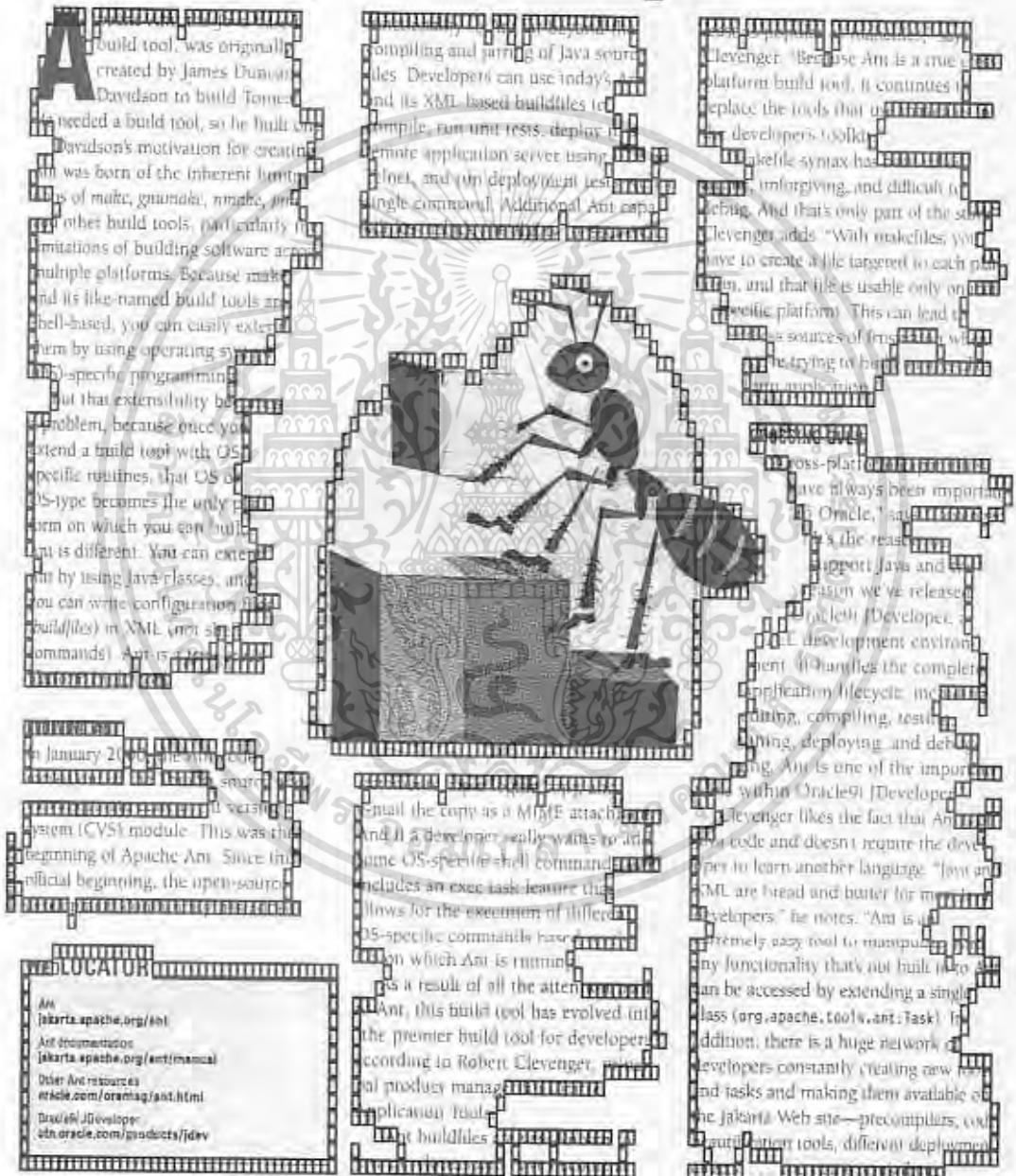
## 2. ตัวอย่างการแบ่งขอบเขตภาพที่ประกอบด้วยรูปภาพและตัวอักษร (รูปที่ 3.26-5.37)



รูปที่ 5.26 ผลการแบ่งภาพเอกสารที่ประกอบด้วยรูปภาพและตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

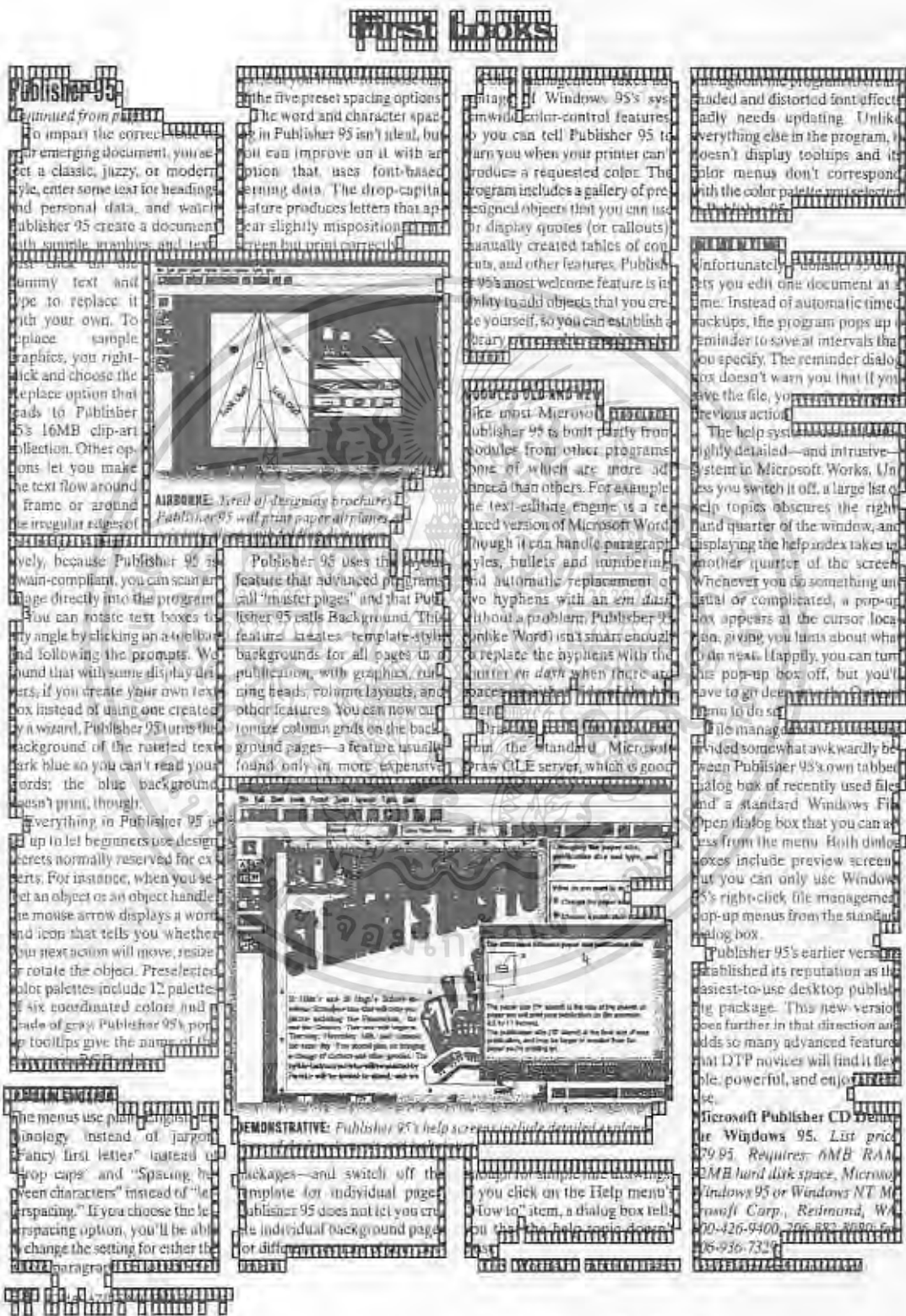
# Ant Builds On



รูปที่ 5.27 ผลการแบ่งภาพเอกสารที่ประกอบด้วยรูปภาพและตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3. ตัวอย่างการแบ่งขอบเขตภาพที่ประกอบด้วยรูปภาพและตัวอักษรที่มีโครงสร้างที่จัดกัน (รูปที่ 3.28-5.30)



รูปที่ 5.28 ผลการแบ่งภาพเอกสารที่มีโครงสร้างที่จัดกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# That long, lonesome last mile

...for the "last mile"—the place where it joins up with other network traffic? Theoretically, it makes sense to use it, but it will probably take a while to catch on, especially in the United States. Most U.S. businesses use Ethernet for their local-area networks, but when they want to connect those networks to PCs at other locations, they have to go through the telephone company. The telephone company uses technologies, such as asynchronous optical network (SONET) and asynchronous transfer mode (ATM), designed primarily to handle voice. So it has to translate packets from Ethernet networks into ATM packets that can travel over SONET links and then translate them back into Ethernet packets when they reach their destination. It's inefficient and expensive.

A new standard, expected to be approved this summer, may help Ethernet make some headway in metro-area networks. What's more, the standard's backers think Ethernet can eventually replace current residential technologies, such as DSL, over copper and cable modems for the broadband connections to the home.

The Ethernet in the First Mile Alliance (EFMA), a group of networking vendors, has been promoting the standard since late 2001. (Whether it's the first or the last mile depends on the direction from which you view it. The phone companies call it the last mile; the Ethernet vendors prefer to call it the first mile.) The group expects the IEEE to ratify the proposed standard, 802.3ah, this summer, according to Craig Easley, president of the EFMA and a director at Extreme Networks. The standard specifies the protocols and interfaces for using Ethernet over access links, for both copper and optical fiber, in the first mile, and puts in place certain quality-of-service specifications to ensure that the service is carrier class.

Using Ethernet in the first mile offers several advantages, says EFMA's Easley. First, Ethernet is a standard, so

...are less expensive than other technologies, because vendors sell more of them. There are two billion Ethernet ports worldwide, says Easley, compared to perhaps 200,000 ports each for SONET and ATM. Second, adoption of Ethernet throughout the entire broadband



networks would eliminate the need to translate packets from Ethernet to another standard, such as ATM (for phone lines) or DOCSIS for cable. This would reduce not only total equipment costs but also the expense of specialized staff to manage the other technologies, says Easley.

Optical Ethernet, which runs over fiber, promises greater speeds at lower costs than other available technologies. The standard is likely to reduce the complexity and cost of transmitting data from the home or the business to the phone company's central office. Today, phone companies use customized, proprietary

Network topology	Speed
Copper coaxial point-to-point	10 Mbps up to 750 meters or 2 Mbps up to 2,700 meters
Fiber optic point-to-point	100 Mbps and 1,000 Mbps up to 10 kilometers
Fiber optic point-to-multipoint	1,000 Mbps up to 20 kilometers

The 802.3ah standard provides for Ethernet over three different topologies over specific distances. The real speed advantage is in the fiber optic implementations.

...ATM packets and other technologies, they could instead use standard equipment, the cost of which would drop. However, carriers are not likely to welcome the reduced costs of running Ethernet from their office to the home

...to adopt Ethernet for their entire network. In the United States, the incumbent carriers own the network and have historically been slow to add new services because under current regulatory policies they must give their competitors access to the network, says Alex Saunders, president and CEO, Metrobility Optical Systems, an optical Ethernet vendor. Phone companies also want to wring as much revenue as they can from their current equipment, adds Easley. It's not likely to change.

What's more, Ethernet has been somewhat of a jaded hero. Metrobility's first vendor Yipes made a splash in 2001 and 2004, when it built its own fiber optic network specifically to sell Ethernet to businesses in 14 U.S. markets. Then the telecom bubble burst and the company went into Chapter 11. It has since emerged from bankruptcy and is getting back "on the path to profitability," according to its new CEO, John P.

Asia is the first place optical Ethernet is being widely deployed, says Michael Toward, principal analyst at Infonetics Research. Because much of China doesn't have a legacy infrastructure, it can start from scratch and install next-generation fiber networks, notes Kumar

...technology office and... forward believe, but U.S. phone companies will eventually build full-scale Ether

...ent their existing networks, because it's going to be much cheaper to operate an Ethernet network. "We think that in 10 to 15 years, a lot of carriers will have done this." The new standard is expected to

รูปที่ 5.29 ผลการแบ่งภาพเอกสารที่มีโครงสร้างติดกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# NEWS ANALYSIS

## Will the Gridlock Remain Unbroken?

Will the United States be able to address the problems of the 21st century? The answer is uncertain.

One of the major problems facing the United States is whether President George W. Bush can break the political gridlock that has largely paralyzed its federal government for most of the last century. The gridlock has even prevented the administration from effectively addressing a huge energy and electricity crisis, one that has brought California's three major utilities to the brink of bankruptcy and threatened the rest of the country.

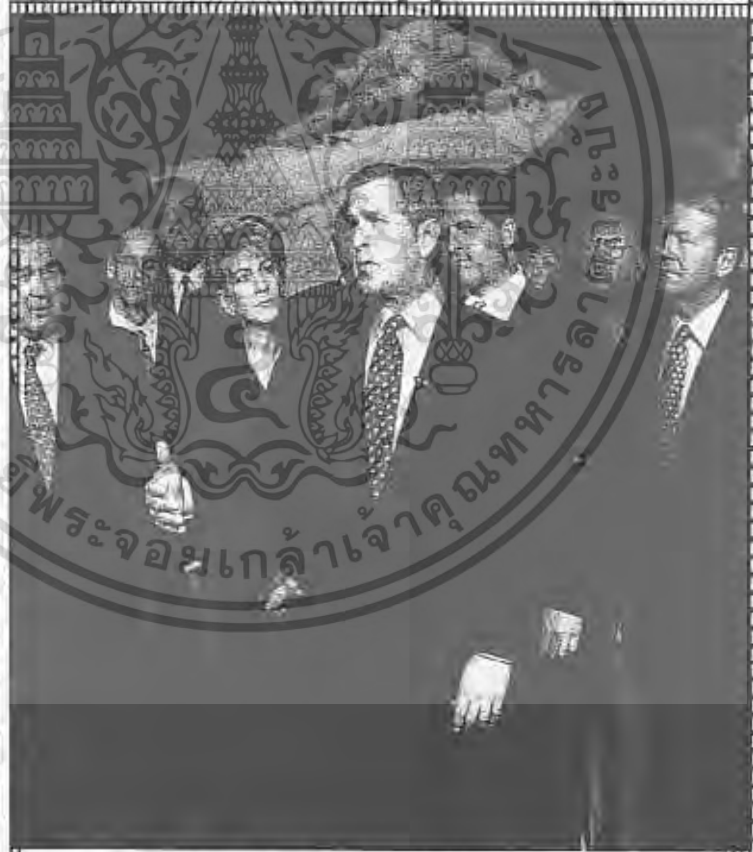
In telecommunications, the President somehow broke the gridlock through a regulatory "loophole" that has failed to produce significant competition in local telephone service. He has left executives of the long-distance carriers openly wondering whether it still pays to invest in the industry.

When it comes to global warming, nuclear power, and privacy, the President has not yet taken any action. These are issues that candidate Bush professed to take most seriously, but was not very specific about what he would do in the areas where Bush made definite promises. The promotion of high technology, for example, and pre-college technical education can he garner enough legislative support to deliver an answer to some of the questions put to him during the presidential campaign?

At the IEEE-USA, the volunteer-driven arm of the IEEE that works on policy issues in Washington, D.C., candidate Bush was asked to address the

governor of Texas. He made four leaders tied for the highest gains in math from 1992 to 1996, that African-American fourth graders made the largest gains in the country in the same period, its Hispanic fourth graders the second largest, and that white fourth graders scored highest in math overall. Bush already has signaled that education will be his first legislative priority, and if he can generalize his Texas

experience to the rest of the country, there is something to write home about. So, in technology policy, prospects for the future seem good. Because both he and candidate Al Gore promised to make permanent an R&D investment tax credit, he ought to be able to garner support for that. The same can hardly be said for telecommunications and energy, where deep philosophical differences divide Republicans and Democrats, and



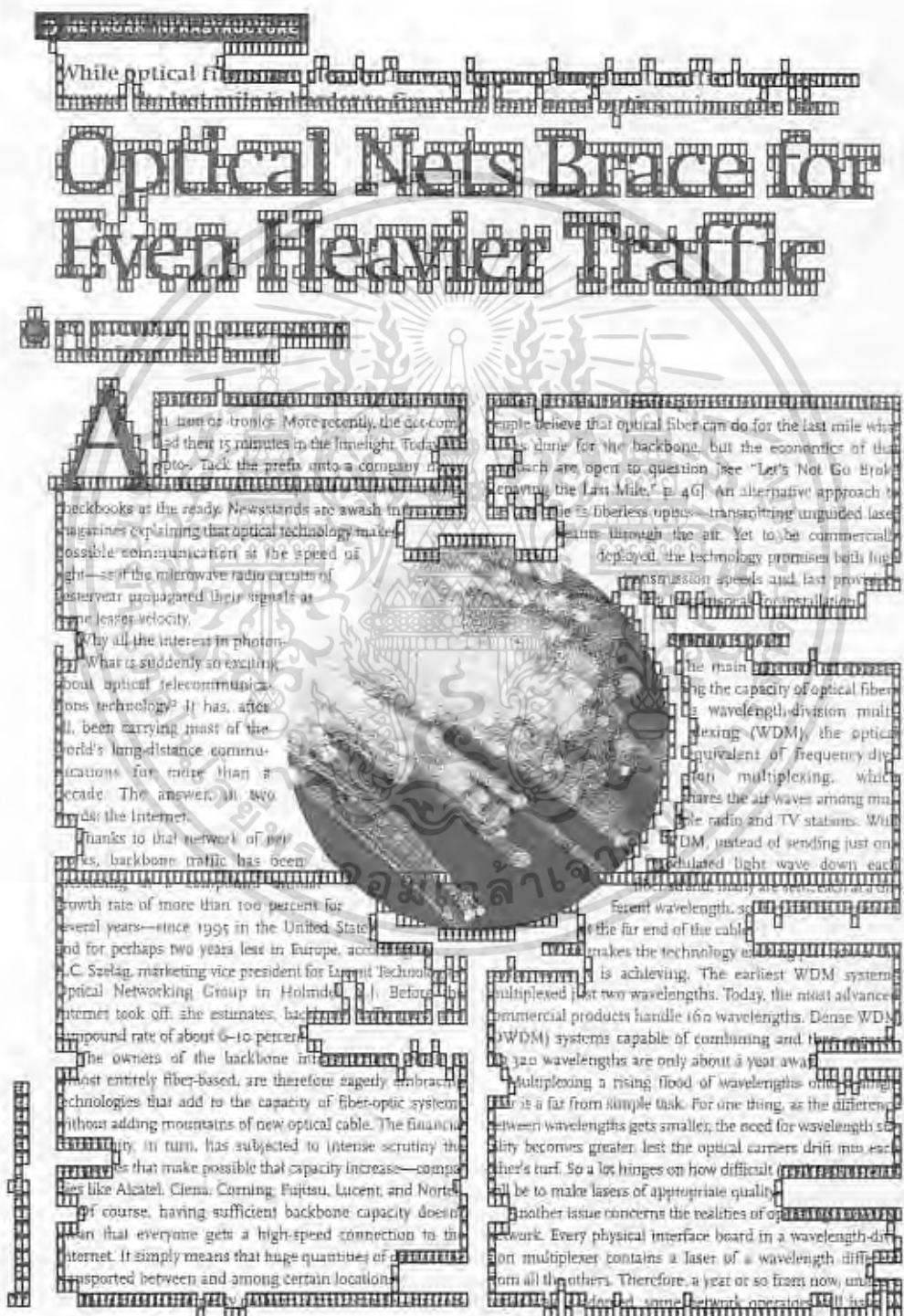
President-elect George W. Bush met with leaders of high-tech companies on 4 January in Austin, Texas. He is flanked [front row, left to right] by IBM's Louis V. Gerstner Jr., Hewlett-Packard's Carly Fiorina, Bell Computer's Michael Dell and Cisco Systems' John Chambers.

### รูปที่ 5.30 ผลการแบ่งภาพเอกสารที่มีโครงสร้างติดกัน

เอกสารนี้เป็นเอกสารที่สแกนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. ตัวอย่างการแบ่งขอบเขตภาพที่ประกอบด้วยรูปภาพที่ไม่สามารถดำเนินการได้

ในรูปที่ 5.31 และ 5.32 แสดงภาพเอกสารที่ไม่สามารถประมวลผลได้เนื่องจากตรงรอยต่อของจุดผิดพลาดมีรูปทรงใดๆ ที่ไม่ใช่สี่เหลี่ยม ส่วนในภาพที่ 5.33 ไม่สามารถประมวลผลได้เนื่องจากขนาดของช่องไฟกว้างเกินไปทำให้อัลกอริทึมวิเคราะห์จุดผิดพลาดไม่ได้ผล โดยวงกลมทั้ง 3 วงจะแสดงบริเวณที่มีการแบ่งผิด



รูปที่ 5.31 ผลการแบ่งภาพเอกสารที่ไม่สามารถแบ่งซ้ำได้

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์การแบ่งซ้ำเพื่อทำซ้ำนั้นไม่อนุญาตให้ทำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## Fuel Cells for the Long Haul, Batteries for the Sprints

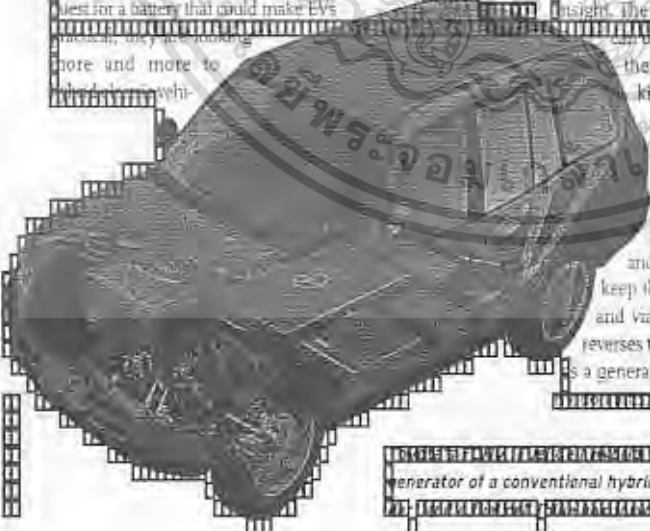
BY MICHAEL J. WATKINS

NEVER MIND THE HYDROGEN CARS. Montreal's Grand Prix racetrack, the Circuit Gilles-Villeneuve, is not the obvious place to look for the latest in environmentally benign vehicles. Still, it was there that Ford Motor Co., Dearborn, Mich., unveiled its first fuel-cell demonstration vehicle less than three months ago: a full-sized, four-door sedan powered by a hydrogen-fuel-cell system. The choice of venue was not random: Montreal was the host city for EVS-17, the 17th International Electric Vehicle Symposium, held 16-18 October 2006, during which the vehicle made its debut. But it may also have been symbolic—acknowledgment by a major automaker that environmental vehicles were about ripe for the fast-track product of Ford Motor's new environment.

Frank Group, the hydrogen-carrying vehicle, like Toyota's (CHV [below]) is not and probably never will be commercially available. But it marks a decisive change in the automotive industry's thinking about the future, making it one of the more important technology developments of recent times.

Until only a few years ago, when environmentalists, automakers, and other concerned parties spoke about low-emission vehicles, they were almost always referring to electric vehicles (EVs)—cars, trucks, and buses powered by batteries.

EVs are still the best bet for a battery that could make EVs a reality. They are also the most and more to be developed.



EVs are still the best bet for a battery that could make EVs a reality. They are also the most and more to be developed. Batteries are not actually dead, but after more than a century of development, they are still not satisfactory for automotive propulsion. They weigh too much, cost too much, and take too long to charge. But hybrids based on a combination of electric motors and internal combustion engines are attractive for two main reasons: they require no technology breakthroughs and no new infrastructure. They work with existing infrastructure; they do not rely on them for primary energy. Rather, they can obtain fuel at any service station.

The battery in a modern HEV serves essentially as a capacitor, providing bursts of power for acceleration and hill climbing. That allows the internal combustion engine (ICE) to run in the vicinity of the "sweet spot" in its operating map, where its efficiency is high and emissions low. Interestingly, it is the high efficiency, translated into good fuel economy, that appeals to consumers today even though the motivation behind fuel development initially was the reduction of air pollution.

The first major automaker with a hybrid electric vehicle to come to the general public was Toyota Motor Corp., of Toyota, Japan. Its Prius HEV went on sale in 1999 in Japan. There about 40,000 of them have so far been sold. It is now also available in North America. Although the Prius uses its ICE most of the time, it operates as a "pure" (battery-only) EV at low speeds—when creeping along in stop-and-go traffic, for example—unless its battery is low on charge.

Ford Motor Co., Tokyo, beat Toyota into the U.S. market with its answer to the Prius, which it called the Insight. The two cars are alike in being parallel hybrids—the ICE can deliver power to the wheels from both the ICE and the electric motor. In a series hybrid—the other kind of HEV—all traction power comes from the electric motor. The ICE never drives the wheels directly, but instead powers the battery charger.

The advantages of HEVs are clear. They can be gassed up like any other car and need no special charging infrastructure since they keep their batteries charged from the engine/generator and via regenerative braking: pressing the brake pedal reverses the action of a vehicle's electric motor, so that it acts as a generator, converting the kinetic energy of the vehicle's

generator of a conventional hybrid with a hydrocarbon fuel.

### รูปที่ 5.32 ผลการแบ่งภาพเอกสารที่ไม่สามารถแบ่งซ้ำได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



## บทที่ 6

### สรุปผลการวิจัยและข้อเสนอแนะ

จากผลการทดลองที่ได้ในบทที่ 5 แสดงให้เห็นว่าการแบ่งขอบเขตของส่วนประกอบอย่างเหมาะสม มีรูปทรงตามรูปทรงของส่วนประกอบจริง และใช้เวลาในการประมวลผลที่รวดเร็ว แต่ถึงอย่างไรก็ตามวิธีการดังกล่าวยังมีข้อจำกัดอยู่หลายประการด้วยกัน ดังนั้นจึงมีข้อสรุปสำหรับผลการวิจัยและข้อเสนอแนะสำหรับแนวทางการทำวิจัยต่อดังนี้

#### 6.1 สรุปผลการวิจัย

วิทยานิพนธ์นี้นำเสนอแนวคิดของกริดตามเส้นขอบและการแบ่งซ้ำในแนวแกน XY มาใช้เพื่อสามารถแบ่งส่วนประกอบของภาพเอกสาร์ เพื่อให้ได้ผลลัพธ์ในการแบ่งส่วนประกอบที่ถูกต้อง รวดเร็ว และมีประสิทธิภาพ ซึ่งจะช่วยให้ประสิทธิภาพของกระบวนการรู้จำภาพเอกสาร์ให้ดียิ่งขึ้น เพราะถ้ามีกระบวนการแบ่งส่วนประกอบภาพเอกสาร์ที่มีประสิทธิภาพจะช่วยประหยัดทั้งเวลาในการประมวลผลเพื่อรู้จำและลดจำนวนหน่วยความจำที่ต้องใช้งานให้น้อยลง โดยเริ่มแรกจะให้หน้าต่างขนาด  $16 \times 32$  พิกเซลวางไปตามขอบของภาพเอกสาร์เริ่มจากตำแหน่งบนซ้ายวนตามเข็มนาฬิกาตามรหัสที่ถูกใจที่ถนัดจนกระทั่งวนกลับมาที่หน้าต่างแรกถือว่าเป็นรอบรอบส่วนประกอบที่ต้องการระบุขอบเขต ดำเนินการแบ่งซ้ำจนครบทุกย่อหน้าในภาพเอกสาร์ จากนั้นพิจารณาส่วนประกอบของโครงสร้างที่ติดกันที่มีการแบ่งคิดในแต่ละย่อหน้า โดยอัตโนมัติ เพื่อที่จะทำให้ผลลัพธ์ของการแบ่งภาพเอกสาร์มีความถูกต้องมากที่สุด ซึ่งผลลัพธ์ที่ดีมากคือการเพิ่มประสิทธิภาพในการทำงานของระบบ โดยการลดพื้นที่หน่วยความจำในการรู้จำส่วนที่ไม่ใช่เนื้อหาสาระ หากพบว่ามีการแบ่งส่วนประกอบซ้ำด้วยกระบวนการแบ่งซ้ำในแนวแกน XY ซึ่งหน้าต่างเหล่านี้จะถูกเก็บในรูปแบบโครงสร้างข้อมูลที่เป็นแบบไดนามิกที่รู้จักกันดี คือ ลิงค์ลิสต์ จากการศึกษาวิจัยได้ผลลัพธ์เป็นที่น่าพอใจ ทั้งในด้านเวลาที่ใกล้เคียงกับวิธีการแบ่งซ้ำตามแนวแกน X-Y ซึ่งเป็นวิธีการที่มีความรวดเร็วในการแบ่งส่วนประกอบภาพเอกสาร์วิธีหนึ่ง นอกจากนั้นแล้ววิธีการที่ได้นำเสนอยังให้ผลลัพธ์ของรูปทรงส่วนประกอบหลังการแบ่งส่วนประกอบได้ใกล้เคียงกับความเป็นจริงมากที่สุดจึงส่งผลให้สามารถประหยัดเนื้อที่หน่วยความจำได้มากกว่าวิธีการแบ่งซ้ำตามแนวแกน X-Y เมื่อเปรียบเทียบกับวิธีการแพร่อย่างมีทิศทางพบว่าวิธีการที่นำเสนอใช้เวลาในการประมวลผลน้อยกว่าเล็กน้อย แต่การแพร่อย่างมีทิศทางจะให้ผลลัพธ์ใกล้เคียงกับความเป็นจริงมากกว่าเนื่องจากสามารถทำงานกับบริเวณจุดผิดพลาดที่มีลักษณะใดๆ ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 6.2 ข้อเสนอแนะแนวทางการทำวิจัยต่อ

วิทยานิพนธ์นี้เป็นการแบ่งส่วนประกอบภาพเอกสาร โดยใช้เทคนิคการติดตามเส้นขอบและการแบ่งในแนว XY แต่อย่างไรก็ตามแม้ว่าวิธีการนี้จะสามารถแบ่งส่วนประกอบได้อย่างถูกต้องและเหมาะสม แต่ยังคงมีบางประเด็นที่ต้องปรับปรุง ซึ่งผู้วิจัยขอเสนอแนะดังนี้

1. ในกระบวนการแบ่งโครงสร้างที่ติดกันจะทำให้ข้อมูลขอบเขตเล็กลงไปเรื่อยๆ เป็นผลให้ต้องมีการเก็บข้อมูลเพิ่มมากขึ้น ทำให้ต้องมีการดูแลและจัดการหน่วยความจำมากขึ้นเป็นลำดับ ดังนั้นถ้ามีการลดการจัดเก็บขอบเขตของข้อมูลลง เลือกเก็บเฉพาะข้อมูลที่จำเป็นจะสามารถลดการใช้หน่วยความจำลงได้ โดยใช้การออกแบบโครงสร้างและการจัดการข้อมูลที่ดี มีการคืนหน่วยความจำเมื่อไม่ใช้แล้ว

2. ปรับเปลี่ยนอัลกอริทึมในการแบ่งซ้ำ โดยเปลี่ยนมาใช้อัลกอริทึมของการแพร่อย่างมีทิศทางจะทำให้ได้ส่วนประกอบได้ถูกต้องยิ่งขึ้น ซึ่งงานในส่วนนี้จะต้องมีการปรับปรุงอยู่ 3 ส่วนหลักๆ คือ

2.1 การปรับปรุงอัลกอริทึมในการจำแนกประเภทของหน้าต่างให้มีความถูกต้องแม่นยำในการจำแนกยิ่งขึ้น

2.2 ทดลองหาตัวจำแนกจุดสีดำที่เหมาะสมที่ใช้ในการแบ่งสันใจหยุดการวางหน้าต่างในการแพร่ ซึ่งค่าดังกล่าวต้องมีความชัดเจนระหว่างการแพร่ในทิศทางแนวตั้งและแนวนอน

2.3 ประยุกต์ใช้หลักการของการหาระยะทางที่สั้นที่สุดเข้ามาช่วยในการแพร่สู่เป้าหมาย

3. หาวิธีการใหม่เพื่อจัดการกับภาพเอกสารที่มีระยะห่างระหว่างตัวอักษรมากกว่าเอกสารปกติ หรือเอกสารที่มีตัวอักษรขนาดใหญ่มากๆ โดยหาขนาดของช่องไฟก่อนเลือกใช้ขนาดของหน้าต่างโดยอัตโนมัติ

4. นอกเหนือจากการเตรียมเอกสารตามแนวทางการจัดหมวดหมู่ในวิทยานิพนธ์นี้แล้ว ผู้วิจัยมีความเห็นว่าในการจัดหมวดหมู่ในเอกสารแต่ละขนาด แต่ละประเภทควรมีการแยกภาพเอกสารที่มีโครงสร้างซับซ้อนและไม่ซับซ้อนออกจากกัน ทั้งนี้เพื่อที่จะได้เปรียบเทียบผลการทดลองเกี่ยวกับเวลาได้ชัดเจนยิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## เอกสารอ้างอิง

- [1] Kruatrachue B., and Suthaphan P. "A fast and efficient method for document segmentation for OCR", Electrical and Electronic Technology, 2001. Proceeding of IEEE Region 10 International conference on, Volume:1., 2001. pp.381- 383.
- [2] Nagy G., and Seth S. "Hierarchical representation of optically scanned documents," Proceedings of the ICPR, 1984. pp.347-349.
- [3] Ha J., Haralick R.M., and Phillips I.T. "Recursive X-Y Cut using Bounding Boxes of connected components", Proceedings of the 3<sup>rd</sup> ICDAR, Montréal, Canada, 1995. pp.952-955.
- [4] Hori O. "A Video Text Extraction Method for Character Recognition," Proceedings of the 5<sup>th</sup> ICDAR, Bangalore, India, 1999. pp. 25-28.
- [5] Li H., Doermann D., and Kia Q. "Automatic Text Detection and Tracking in Digital Video." IEEE Trans. on Image Processing, vol. 9, 2000. pp. 147-156.
- [6] Okun O., and Pietikäinen M. "A Survey of Texture-based Methods for Document Layout Analysis." Proceedings of Workshop on Texture Analysis in Machine Vision, Oulu, Finland, 1999. pp. 137-148.
- [7] Wang S.Y., and Yagasaki T. "Block Selection: A Method for Segmenting Page Image of Various Editing Styles." Proceedings of the 3<sup>rd</sup> ICDAR, Montréal, Canada, 1995. pp.128-133.
- [8] Sauvola J., and Pietikäinen M. "Page Segmentation and Classification using Fast Feature Extraction and Connectivity Analysis." Proceedings of the 3<sup>rd</sup> ICDAR, Montréal, Canada, 1995. pp.1127-1131.
- [9] Payne J.S., Stonham T.J., and Patel D. "Document Segmentation using Texture Analysis." Proceedings of the 12<sup>th</sup> ICPR, Jerusalem, Israel, 1994, pp. 380-382.
- [10] Antonacopoulos A., and Ritchings R.T., "Flexible Page Segmentation using the Background." Proceedings of the 12<sup>th</sup> ICPR, Jerusalem, Israel, 1994. pp. 339-344.
- [11] Antonacopoulos A. "Page Segmentation using the Description of the Background." Computer Vision and Image Understanding, vol. 70, 1998. pp. 350-369.
- [12] Baird H.S. "Background Structure in Document Images." Advances in Structural and Syntactic Pattern Recognition, World Scientific, Singapore, 1992. pp. 253-269.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [13] Normand H.S., and Viard-Gaudin C. "A Background Based Adaptive Page Segmentation Algorithm." Proceedings of the 3<sup>rd</sup> ICDAR, Montréal, Canada, 1995. pp.138-141.
- [14] Cinque L., Lombardi L., and Manzini G. "A Multiresolution Approach for Page Segmentation." Pattern Recognition Letters, vol. 19, 1998. pp. 217-225.
- [15] Etemad K., Doermann D., and Chellappa R. "Multiscale Segmentation of Unstructured Document Page using Soft Decision Integration." IEEE Trans. on PAMI, vol. 19, 1997. pp. 92-96.
- [16] Jain A.K., and Zhong Y. "Page Segmentation using Texture Analysis." Pattern Recognition, vol. 29, 1996. pp. 743-770.
- [17] Fan K.C., Liu C.H., and Wang Y.K. "Segmentation and Classification of Mixed Text/Graphics/Image Documents." Pattern Recognition Letters, vol. 15, 1994. pp. 1201-1209.
- [18] Fan K.C., and Wang L.S. "Classification of Document Blocks using Density Feature and Connectivity Histogram." Pattern Recognition Letters, vol. 16, 1995. pp. 955-962.
- [19] Gatos B., Mantzaris S.I., Chandrinos K.V., Tsigris A., and Perantonis S.JJ. "Integrated Algorithms for Newspaper Page Decomposition and Article Tracking." Proceedings of the 5<sup>th</sup> ICDAR, Bangalore, India, 1999. pp.559-562.
- [20] Le D.X., Thoma G.R., and Wechsler H. "Classification of Binary Document Image into Textual or Non-textual Data Blocks using Neural Network Models." Machine Vision and Applications, vol. 8, 1995. pp. 289-304.
- [21] Shih F.Y., and Chen S.S., "Adaptive Document Block Segmentation and Classification." IEEE Trans. on SMC-Part B, Cybernetics, vol. 26, 1996. pp. 797-802.
- [22] Etemad K., Doermann D., and Chellappa R. "Page Segmentation using Decision Integration and Wavelet Packets." Proceedings of the 12<sup>th</sup> ICPR, Jerusalem, Israel, 1994. pp.345-349.
- [23] Chen W.Y., and Chen S.Y. "Adaptive Page Segmentation for Color Technical Journals' Cover Images." Image and Vision Computing, vol. 16, 1998. pp. 855-877.
- [24] Jain A.K., and Yu B. "Automatic Text Location in Images and Video Frames." Pattern Recognition, vol. 31, 1998. pp. 2055-2076.
- [25] Messelodi S., and Medena C.M. "Automatic Identification and Skew Estimation of Text Lines in Real Scene Images." Pattern Recognition, vol. 32, 1999. pp. 791-810.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [26] Park T., Kim D., and Chung K. "Orientation and Scale Invariant Text Region Extraction in WWW Images." *Proceedings IARP Workshop on Machine Vision and Applications*, Chiba, Japan, 1998. pp. 290-293.
- [27] Zhong Y., Karu K., and Jain A.K. "Locating Text in Complex Color Images." *Proceedings of the 3<sup>rd</sup> ICDAR*, Montréal, Canada, 1995. pp.146-149.
- [28] Zhou J., and Lopresti D. "Extracting Text from WWW Images." *Proceedings of the 4<sup>th</sup> ICDAR*, Ulm, Germany, 1997. pp.248-252.
- [29] Akiyama T., and Hagita N. "Automated Entry System for Printed Documents." *Pattern Recognition*, vol. 23, 1990, pp. 1141-1154.
- [30] Ha J., Haralick R., Phillips I. "Document Page Decomposition by the Bounding-Box Projection Technique." *Proceedings of the 3<sup>rd</sup> Int'l Conf. Document Analysis and Recognition*, Montréal, Canada, 1995. pp. 1,119-1,122.
- [31] Jain A.K., and Bhattacharjee S. "Text Segmentation using Gabor Filters for Automatic Document Processing." *Machine Vision and Application*, vol. 5, 1992. pp. 169-184.
- [32] Chen J.L. "A Simplified Approach to HMM Based Texture Analysis and Its Application to Document Segmentation." *Pattern Recognition Letters*, vol. 18, 1997. pp. 993-1007.
- [33] Wang D., and Srihari S.N. "Classification of Newspaper Image Blocks using Texture Analysis." *Computer Vision, Graphics, and Image Processing*, vol. 47, 1989. pp. 327-352.
- [34] Liu J., Tang Y.Y., and Suen C.Y. "Chinese Document Layout Analysis Based on Adaptive Split-and-Merge and Qualitative Spatial Reasoning." *Pattern Recognition*, vol. 30, 1997. pp. 1265-1278.
- [35] Srihari S.N., Hong T., and Srikantan G. "Machine Printed Japanese Document Recognition." *Pattern Recognition*, vol. 30, 1997. pp. 1301-1313.
- [36] Sobottka K., Bunke H., and Kronenberg H. "Identification of Text on Colored Book and Journal Covers," *Proceedings of the 5<sup>th</sup> ICDAR*, Bangalore, India, 1999. pp. 57-62.
- [37] Drivas D., and Amin A. "Page Segmentation and Classification Utilizing Bottom-Up Approach." *Proceedings of the 3<sup>rd</sup> ICDAR*, Montréal, Canada, 1995. pp.610-614.
- [38] Hönes F., and Lichter J. "Layout Extraction of Mixed Mode Documents." *Machine Vision and Applications*, vol. 7, 1994. pp. 237-246.
- [39] Jain A.K., and Yu B. "Page Segmentation using Document Model." *Proceedings of the 4<sup>th</sup> ICDAR*, Ulm, Germany, 1997. pp.34-38.

เอกสารนี้เป็นเอกสารลิขสิทธิ์สงวนไว้สำหรับใช้ภายในห้องสมุดเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [40] Zhong Y., Karu K., and Jain A.K. "Locating Text in Complex Color Images." *Pattern Recognition*, vol. 28, 1995. pp. 1523-1535.
- [41] Sivaramakrishnan R., Phillips I.T., Ha J., Subramaniam S., and Haralick R.M. "Zone Classification in a Document using the Method of Feature Vector Generation." *Proceedings of the 3<sup>rd</sup> ICDAR, Montréal, Canada, 1995*. pp.541-544.
- [42] Antonacopoulos A., and Ritchings R.T. "Representation and Classification of Complex-shaped Printed Regions using White Tiles." *Proceedings of the 3<sup>rd</sup> ICDAR, Montréal, Canada, 1995*. pp.1132-1135.
- [43] Pavlidis T., and Zhou H. "Page Segmentation and Classification." *Computer Vision Graphics and Image Processing*, vol. 54, 1992. pp. 484-496.
- [44] Okun O., Doermann D., and Pietikäinen M. "Page Segmentation and Zone Classification: the State of Art." *Technical Report LAMP-TR-036, University of Maryland, Language and Media Processing Laboratory, 1999*, 34 p.
- [45] Kise K., Sato A., and Iwata M. "Segmentation of Page Images using the Area Voronoi Diagram." *Computer Vision and Image Understanding*, vol. 70, 1998. pp. 370-382.
- [46] Tang Y.Y., Ma H., Mao X., Liu D., and Suen C.Y. "A New Approach to Document Analysis Based on Modified Fractal Signature." *Proceedings of the 3<sup>rd</sup> ICDAR, Montréal, Canada, 1995*. pp.567-570.
- [47] Ishitani Y. "Document Layout Analysis Based on Emergent Computation." *Proceedings of the 4<sup>th</sup> ICDAR, Ulm, Germany, 1997*. pp.45-50.
- [48] Hase H., Shinokawa T., Yoneda M., Sakai M., and Maruyama H. "Character String Extraction from a Color Document." *Proceedings of the 5<sup>th</sup> ICDAR, Bangalore, India, 1999*. pp. 75-78.
- [49] Jain A.K., and Yu B. "Document Representation and Its Application to Page Decomposition." *IEEE Trans. on PAMI*, vol. 20, 1998. pp. 294-308.
- [50] Kanai J., Rice S.V., Nartker T.A., and Nagy G., "Automatic Evaluation of OCR Zoning." *IEEE Trans. on PAMI*, vol. 17, 1995. pp. 86-90.
- [51] Antonacopoulos A., and Brough A. "Methodology for Flexible Analysis of the Performance of Page Segmentation Algorithms." *Proceedings of the 5<sup>th</sup> ICDAR, Bangalore, India, 1999*. pp. 451-454.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [52] Kanungo T., and Haralick R.M. "An Automatic Closed-loop Methodology for Generating Character Groundtruth for Scanned Documents." *IEEE Trans. on PAMI*, vol. 21, 1999. pp. 179-183.
- [53] Yanikoglu B.A., and Vincent L. "Ground-truthing and Benchmarking Document Page Segmentation." *Proceedings of the 3<sup>rd</sup> ICDAR*, Montréal, Canada, 1995. pp.601-604.
- [54] Yanikoglu B.A., and Vincent L. "Pink Panther: A Complete Environment for Ground-truthing and Benchmarking Document Page Segmentation." *Pattern Recognition*, vol. 31, 1998. pp. 1191-1204.
- [55] Junker M., Hoch R., and Dengel A. "On the Evaluation of Document Analysis Components by Recall Precision and Accuracy." *Proceedings of the 5<sup>th</sup> ICDAR*, Bangalore, India, 1999. pp. 713-716.
- [56] Wahl F., Wong K., Casey R. "Block Segmentation and Text Extraction in Mixed Text/Image Document." *Computer Vision Graphic and Image Processing*, vol. 20, 1982. pp.375-390.
- [57] Fujisawa H., Nakano Y. "Top-Down Approach for the Analysis of Document." *Proceedings of the 10<sup>th</sup> Int'l Conf. Pattern Recognition*, Atlantic City, N.J., 1990. pp. 113-122.
- [58] Fisher J., Hinds S., D' Amato K. "A Rule Based System for Document Image Segmentation." *Proceedings of the 10<sup>th</sup> Int'l Conf. Pattern Recognition*, Atlantic City, N.J., 1990. pp. 567-572.
- [59] Pavlidis T., Zhou J. "Page Segmentation by White Streams." *Proceedings of the First Int'l Conf. Document Analysis and Recognition*, Saint-Malo, France, 1991. pp. 945-953.
- [60] Baird H. "Anatomy of a Versatile Page Reader." *Proceedings of the IEEE*, vol. 80, 1992. pp. 1,059-1,065.
- [61] Legourgeois F., Bublinski Z., Emptoz H. "A Fast and Efficient Method for Extracting Text Paragraphs and Graphics Form Unconstrained Documents," *Proceedings of the 11<sup>th</sup> Int'l Conf. Document Analysis and Recognition*, The Hague, 1992. pp. 272-276.
- [62] Akidele O., Belaid A. "Page Segmentation by Segment Tracing." *Proceedings of the Second Int'l Conf. Pattern Recognition*, Tsukuba, Japan, 1993. pp. 341-344.
- [63] Amamoto N., Torigoe S., Hirogaki Y. "Block Segmentation and Text Area Extraction of Vertically/Horizontally Written Document." *Proceedings of the Second Int'l Conf. Pattern Recognition*, Tsukuba, Japan, 1993. pp. 739-742.

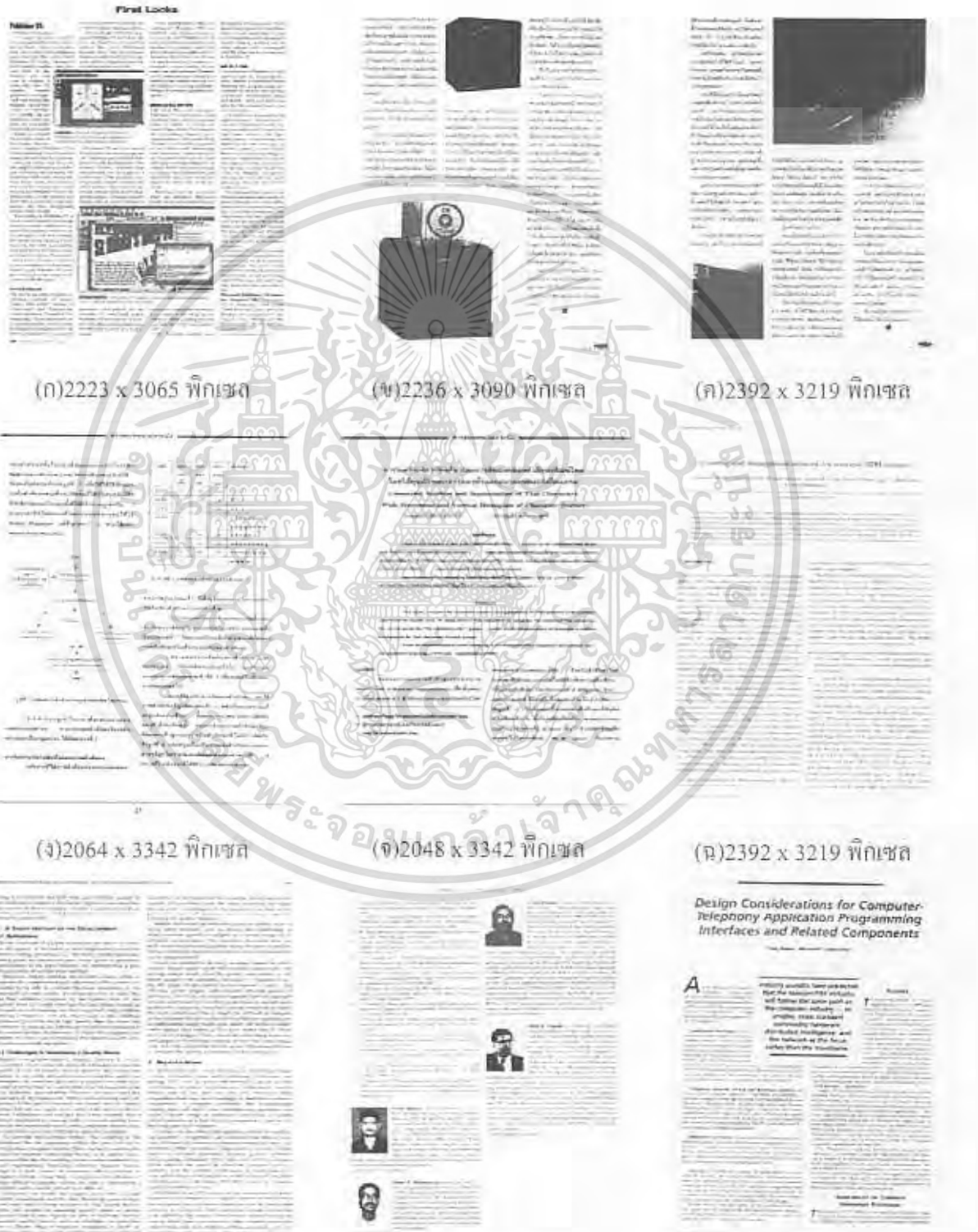
- [64] Ittner D., Baird H. "Language-Free Layout Analysis." Proceedings of the 'Second Int'l Conf. Pattern Recognition, Tsukuba, Japan, 1993. pp. 336-340.
- [65] O' Gorman L. "The Document Spectrum for Page Layout Analysis." Proceedings of the IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 15, 1993. pp. 1,162-1,173.
- [66] Zlatopolsky A. "Automated Document Segmentation." Pattern Recognition Letter, vol. 15, 1994. pp. 699-704.
- [67] Doermann D. "Page Decomposition and Related Research." Proceedings of the Symp. Document Image Understanding Technology, Bowie, Md, 1995. pp. 39-55.
- [68] Sylwester D., Seth S. "A trainable, Single-Pass Algorithm for Column Segmentation." Proceedings of the 3<sup>rd</sup> Int'l Conf. Document Analysis and Recognition, Montréal, Canada, 1995. pp. 615-618.
- [69] Jain A., Karu K. "Learning Texture Discrimination Masks." Proceedings of the IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, 1995. pp. 195-205.
- [70] Kise K., Yanagida O., Takamatsu S. "Page Segmentation Based on Thinning of Background." Proceedings of the 13<sup>th</sup> Int'l Conf. Pattern Recognition, Vienna, 1996. pp. 788-792.
- [71] Lui J., Fang Y., He Q., Sun C. "Adaptive Document Segmentation and Geometric Relation Labeling: Algorithms and Experimental Results." Proceedings of the 13<sup>th</sup> Int'l Conf. Pattern Recognition, Vienna, 1996. pp. 763-767.
- [72] Yamachita A., Amano T., Hirayama Y., Itoh N., Katoh S., Mano T., Toyokawa K. "A Document Recognition System and Its Applications." IBM J. Research and Development, vol. 40, 1996. pp. 341-352.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับภาพเอกสารที่ใช้ในการทดสอบความเหมาะสมของขนาดหน้าตานั้น จะใช้ภาพเอกสารที่ถูกสแกนในลักษณะภาพขาวดำ 30 ภาพ แต่ละมีความละเอียด 300 จุดต่อนิ้ว เป็นภาพเอกสารที่ไม่มียกขอบและได้รับการแก้ไขความเอียงเรียบร้อยแล้ว สำหรับตัวอย่างในภาคผนวกนี้จะนำเสนอเพียงบางส่วน (9 ภาพ) จากภาพ (ก)-(ฉ)



เอกสารนี้เป็น (ข) 2275 x 3238 พิกเซล สำหรับการใช้งาน (ข) 2387 x 3153 พิกเซล อนุญาตให้ (ฉ) 2437 x 3355 พิกเซล ค่าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.1 แสดงผลการทดลองแบ่งภาพเอกสาร โดยใช้หน้าต่างขนาดต่างๆ ดังนี้

เอกสาร ที่	ขนาดภาพ เอกสาร (พิกเซล)	ขนาดของหน้าต่าง (กว้าง x ยาว)							
		32 x 32	32 x 16	32 x 8	16 x 32	16 x 16	8 x 32	8 x 16	4 x 3 2
1.	2223x3065	3	0	0	1	0	2	0	0
2.	2236x3090	2	0	0	1	0	2	0	3
3.	1295x981	4	0	0	1	0	2	0	3
4.	1134x1359	4	0	0	1	0	2	0	3
5.	2064x3342	2	0	0	1	0	3	0	-1
6.	2048x3342	2	0	0	1	0	3	-1	-1
7.	2275x3238	2	1	0	3	4	3	-1	5
8.	2349x1319	4	0	0	3	0	2	0	1
9.	2387x3153	4	0	0	2	1	3	-1	5
10.	2437x3355	3	1	0	2	6	4	0	5
11.	2392x3219	5	0	0	3	4	2	1	2
12.	2311x3355	2	0	0	1	0	3	-1	-1
13.	2273x3129	3	0	0	1	0	2	0	0
14.	2531x3477	2	3	0	1	0	1	-1	0
15.	2531x3478	2	0	0	1	0	1	0	0
16.	2388x3340	2	0	0	1	0	3	0	0
17.	2244x2432	1	0	0	2	0	0	-1	0
18.	2531x3478	1	0	-1	2	0	3	-1	0
19.	2531x3478	3	0	0	1	0	2	-1	0
20.	2571x3478	3	0	0	1	-1	2	-1	0
21.	2531x3478	3	0	0	0	0	2	-1	0
22.	2544x3295	4	0	0	2	1	3	0	0
23.	2528x3295	0	0	-1	2	0	1	3	0
24.	4223x3300	0	0	0	1	3	2	0	0
25.	2352x3000	0	0	0	1	0	2	0	0
26.	2349x3111	3	0	0	2	1	4	-1	0
27.	2190x2988	3	0	0	2	1	1	-1	0
28.	1119x1533	0	0	0	1	1	2	1	0
29.	1758x2609	1	0	0	1	0	0	0	0
30.	2362x3057	1	0	0	1	2	2	-1	0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับการทดลองเลือกขนาดของหน้าต่างที่เหมาะสมในวิทยานิพนธ์จะทดลองใช้หน้าต่างตามขนาดความกว้างของหน้าต่าง คือ 32, 16, 8, และ 4 พิกเซลตามลำดับ ส่วนสูงของขนาดหน้าต่างนั้นจะแปรผันลดลงเป็นจำนวนเท่าของ 32 เช่นกัน ซึ่งสามารถแบ่งออกได้ 4 กลุ่มดังนี้

1. 32x32, 32x16, 32x8, 32x4
2. 16x32, 16x16, 16x8, 16x4
3. 8x32, 8x16, 8x8, 8x4
4. 4x32, 4x16, 4x8, 4x4

จากการทดลองพบว่าหน้าต่างบางขนาดไม่เหมาะสมที่จะนำมาแบ่งภาพเอกสารเลข ซึ่งสืบเนื่องมาจากเหตุผลหลายประการด้วยกันตามเงื่อนไขที่ได้ระบุไว้ด้านล่าง จึงทำให้ในตารางแสดงผลถูกลดรูปเหลือเพียง 8 ขนาดลงในตารางที่ ก.1

จากตารางที่ ก.1 ค่าตัวเลขต่างๆ ที่ระบุในตารางการที่วงวนมีความหมายดังนี้

- 1 แสดงระดับความเหมาะสมมากที่สุด
- 2 แสดงระดับความเหมาะสมมาก
- 3 แสดงระดับความเหมาะสมปานกลาง
- 4 แสดงระดับความเหมาะสมน้อย
- 5 และ 6 แสดงระดับความเหมาะสมน้อยสุด
- 0 ไม่เหมาะสมที่จะนำมาใช้

-1 ไม่เหมาะสมเนื่องจากไม่สามารถประมวลผลได้เนื่องจากใช้หน่วยความจำมากเกินไป โดยการพิจารณาเลือกใช้หน้าต่างมีเงื่อนไขหลักๆ ดังนี้

1. แบ่งแล้วไม่ทำให้เกิดการแบ่งเกิน (Over segmentation) คือการแบ่งภาพเอกสารใดๆ แล้วไม่ทำให้เกิดข้อหน้ามากเกินไปจนข้อหน้าที่ยาวจะมี
2. แบ่งแล้วไม่ทำให้เกิดการแบ่งขาด (Under segmentation) คือการแบ่งภาพเอกสารใดๆ แล้วไม่ทำให้เกิดการแบ่งข้อหน้าน้อยกว่าที่ควรจะมี
3. มีความเหมาะสมในการแบ่งแยกระหว่างลอสมันท์ที่ติดกันออกจากกันได้
4. มีความเหมาะสมที่จะไม่รวมเอาส่วนของสัญญาณรบกวนเล็กๆ น้อยๆ เข้ามารวมกับส่วนของเนื้อหาที่ต้องการ ซึ่งในหัวข้อนี้ต้องพิจารณาร่วมกับค่าเรโซไซท์ที่เหมาะสม

จากการทดลองจะพบว่าขนาดของหน้าต่างที่สามารถที่จะแบ่งส่วนประกอบภาพเอกสารได้ดีมีอยู่ 4 ขนาดด้วยกันคือ 32x32 , 16x32, 16x16 และ 8x32 พิกเซล จากการพิจารณาผลการทดลองหน้าต่างขนาด 16x32 พิกเซลจะให้ความถูกต้องในการแบ่งลอสมันท์ได้ดี ทั้งสามารถประมวลผลภาพเอกสารที่หลากหลายได้ ดังนั้นในวิทยานิพนธ์นี้จึงเลือกใช้หน้าต่างดังกล่าวในการทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.2 แสดงการเลือกค่าเรโซลูชันที่ใช้ในการทดลอง

เอกสารที่	ขนาดภาพเอกสาร (พิกเซล)	ค่า Threshold ที่ใช้				
		0.01	0.02	0.03	0.04	0.05
1.	2223x3065	1	1	2	0	0
2.	2236x3090	2	1	0	0	0
3.	1295x981	1	1	2	0	0
4.	1134x1359	1	2	0	0	0
5.	2064x3342	1	1	2	3	0
6.	2048x3342	2	1	3	0	0
7.	2275x3238	1	2	0	0	0
8.	2349x1319	1	1	0	0	0
9.	2387x3153	2	1	0	0	0
10.	2437x3355	1	1	0	0	0
11.	2392x3219	1	1	0	0	0
12.	2311x3355	1	1	0	0	0
13.	2273x3129	2	1	3	0	0
14.	2531x3477	2	1	3	0	0
15.	2531x3478	3	1	2	0	0
16.	2388x3340	2	1	3	0	0
17.	2244x2432	1	2	3	0	0
18.	2531x3478	2	1	0	0	0
19.	2531x3478	2	1	0	0	0
20.	2571x3478	1	2	0	0	0
21.	2531x3478	1	1	3	4	0
22.	2544x3295	2	1	3	0	0
23.	2528x3295	3	2	1	3	0
24.	4223x3300	3	2	1	3	0
25.	2352x3000	2	1	3	0	0
26.	2349x3111	1	1	2	0	0
27.	2190x2988	2	1	0	0	0
28.	1119x1533	1	2	0	0	0
29.	1758x2609	1	1	2	0	0
30.	2362x3057	1	1	0	0	0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ ก.2 ค่าตัวเลขต่างๆ ที่ระบุในตารางอธิบายได้ดังนี้

- 1 แสดงระดับความเหมาะสมมากที่สุด
- 2 แสดงระดับความเหมาะสมปานกลาง
- 3 และ 4 แสดงระดับความเหมาะสมน้อย
- 0 ไม่เหมาะสมที่จะนำมาใช้

จากตารางพบว่ามีค่าเรโซไซร์ที่เหมาะสมจะนำมาใช้ 2 ค่าด้วยกันคือ 0.01 และ 0.02 สำหรับในวิทยานิพนธ์นี้เลือกใช้ค่าเรโซไซร์เท่ากับ 0.02 เพราะที่ค่าเรโซไซร์เท่ากับ 0.01 นั้นเนื่องจากมีค่าเรโซไซร์ในการทำงานต่ำ ดังนั้นจึงมีการรวมเอาส่วนของสัญญาณรบกวนเข้ามาด้วยจึงไม่เหมาะสมงานวิจัย

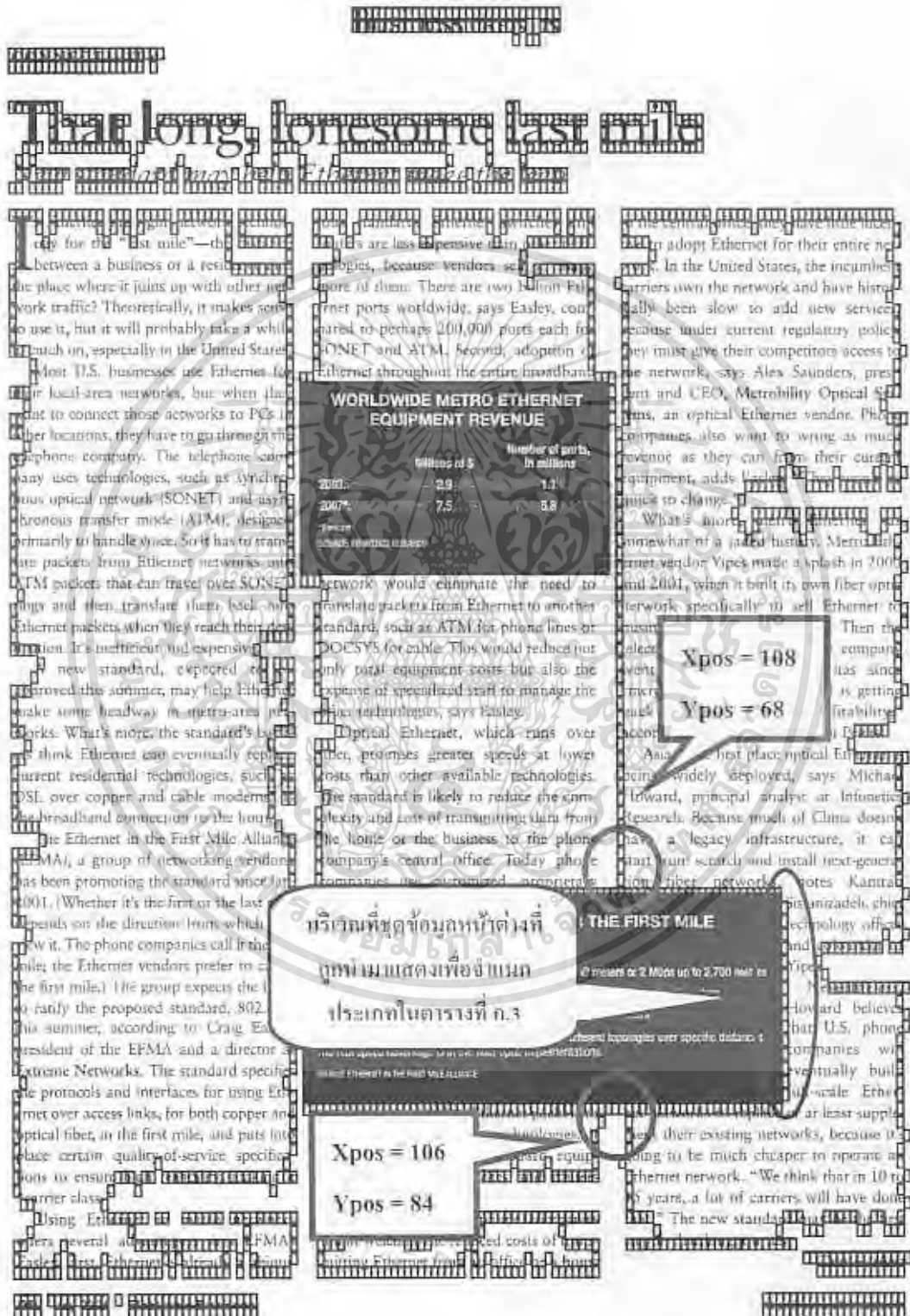
### ผลการทดลองการแบ่งประเภทของหน้าต่าง

สำหรับในตารางที่ ก.3 ได้แสดงตัวอย่างบางส่วนของกรดำเนินการจำแนกประเภทของหน้าต่างผิดพลาด ซึ่งเป็นผลมาจากการทดลองเพื่อหาวิธีการจำแนกประเภทของหน้าต่างตามวิธีการเพื่ออย่างมีทิศทาง ซึ่งตัวย่อต่างๆ ในตารางมีดังนี้ Xpos คือ พิกัดคัมหมงหน้าต่างในแนวแกน x, Ypos คือ พิกัดคัมหมงหน้าต่างในแนวแกน y, Block ID คือ หมายเลขคัมหมงหน้าต่าง, TV คือ Transition of pixel in vertical ซึ่งบ่งบอกค่าการเปลี่ยนแปลงระหว่างจุดขาวและดำที่เกิดขึ้นภายในหน้าต่างตามแนวตั้ง, TH คือ Transition of pixel in horizontal ซึ่งบ่งบอกค่าการเปลี่ยนแปลงระหว่างจุดขาวและดำที่เกิดขึ้นภายในหน้าต่างตามแนวนอน, TR คือ Transition of pixel in oblique right ซึ่งบ่งบอกค่าการเปลี่ยนแปลงระหว่างจุดขาวและดำที่เกิดขึ้นภายในหน้าต่างตามแนวทแยงขึ้น, TL คือ Transition of pixel in oblique left ซึ่งบ่งบอกค่าการเปลี่ยนแปลงระหว่างจุดขาวและดำที่เกิดขึ้นภายในหน้าต่างตามแนวทแยงลง, SV คือ Sum of black vertically เป็นผลรวมของจุดดำทั้งหมดตามแนวตั้ง, SH คือ Sum of black horizontal เป็นผลรวมของจุดดำทั้งหมดตามแนวนอน, OR คือ Sum of black oblique right เป็นผลรวมของจุดดำทั้งหมดตามแนวทแยงขึ้น, OL คือ Sum of black oblique left เป็นผลรวมของจุดดำทั้งหมดตามแนวทแยงลง โดยค่าในวงเล็บหลังตัวแปรนั้นๆ คือค่าสูงสุดที่เป็นไปได้ และ TB คือ Type of block or Type of window เป็นค่าที่ระบุประเภทของหน้าต่างอธิบายได้ดังนี้

- 1 คือ หน้าต่างนั้นเป็นหน้าต่างที่แทนขาว
- 2 คือ หน้าต่างนั้นเป็นหน้าต่างที่แทนส่วนประกอบของตัวอักษร
- 3 คือ หน้าต่างนั้นเป็นหน้าต่างที่แทนตัวอักษร
- 4 คือ หน้าต่างนั้นเป็นหน้าต่างที่แทนรูปภาพ
- 1 คือ หน้าต่างที่ไม่อาจตัดสินใจว่าจัดอยู่ในประเภทใด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในรูปที่ ก.1 แสดงภาพตัวอย่างที่ใช้ในการทดสอบการจำแนกประเภทของหน้าต่าง โดยในวงรี จะแสดงบริเวณที่ชุดหน้าต่างถูกนำมาแยกประเภท ซึ่งจะอยู่ในคอลัมน์ที่ 132 และ 133 โดยจะมีการปรับเปลี่ยนค่าหมายเลขแถวตามแนวตั้ง



เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ของบริษัทฯ ขอสงวนสิทธิ์ในเนื้อหา และขอสงวนสิทธิ์ในชื่อและชื่อประโยชน์ด้านการค้า ไม่ว่าการณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.3 แสดงตัวอย่างการจำแนกประเภทชุดข้อมูลหน้าต่างในแนวดิ่ง

Xpos	Ypos	BlockID	TV(16)	TH(32)	TR(16)	TL(16)	SV(50)	SH(50)	OR(16)	OL(16)	TB
132	68	10,876	9	27	1	4	147	147	6	6	1
132	69	11,034	3	37	2	1	229	229	7	7	-1
132	70	11,192	2	35	1	1	227	227	6	6	-1
132	71	11,350	0	32	2	1	224	224	7	7	1
132	72	11,508	4	41	1	1	233	233	6	6	-1
132	73	11,666	2	34	1	1	226	226	6	6	-1
132	74	11,824	0	32	1	1	224	224	6	6	1
132	75	11,982	0	32	1	1	224	224	6	6	1
132	76	12,140	4	33	1	1	226	226	6	6	-1
132	77	12,298	4	33	1	1	226	226	6	6	-1
132	78	12,456	4	35	1	1	228	226	6	6	-1
132	79	12,614	0	32	1	1	224	224	6	6	1
132	80	12,772	0	32	1	1	224	224	6	6	1
132	81	12,930	0	32	1	1	224	224	6	6	1
132	82	13,088	0	32	2	1	244	244	7	7	1
132	83	13,246	4	34	1	1	226	226	6	6	-1
132	84	13,404	10	13	0	2	55	55	0	0	1
133	68	10,877	23	19	1	0	34	34	1	1	-1
133	69	11,035	42	35	2	3	89	89	5	5	-1
133	70	11,193	33	45	3	2	109	109	3	3	-1
133	71	11,351	19	12	2	1	30	30	1	1	-1
133	72	11,509	56	53	2	4	97	97	4	4	-1
133	73	11,667	33	43	2	2	92	92	2	2	-1
133	74	11,825	15	14	2	0	29	29	1	1	3
133	75	11,983	0	0	0	0	0	0	0	0	4
133	76	12,141	34	25	1	1	65	65	3	3	2
133	77	12,299	12	22	0	4	50	50	0	0	1
133	78	12,457	42	76	4	4	168	168	4	4	-1
133	79	12,615	7	2	0	0	7	7	0	0	4
133	80	12,773	33	35	2	4	64	64	1	1	-1
133	81	12,931	46	30	5	2	66	66	3	3	2
133	82	13,089	25	23	1	1	44	44	3	3	-1
133	83	13,247	33	54	2	2	128	128	2	2	-1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับภาพเอกสารที่ใช้ในการทดสอบ จะแบ่งออกเป็น 3 ขนาดด้วยกัน คือ ภาพเอกสาร ขนาดเล็ก กลาง และขนาดใหญ่ ซึ่งแต่ละขนาดจะประกอบไปด้วยภาพเอกสารที่จัดไว้สำหรับการทดสอบสามประเภท คือ ภาพตัวอักษรอย่างเดียว (T) ภาพตัวอักษร และรูปภาพอยู่ร่วมกัน (T&P) และรูปภาพที่มีส่วนติดกันเมื่อเกิดการแบ่ง (C)

ตารางเปรียบเทียบวิธีการทำงานจะทำการเปรียบเทียบเวลาที่ใช้ในการประมวลผลและจำนวนจุดภาพที่เข้าถึงสำหรับในแต่ละวิธีการ วิธีการที่ 1 คือ วิธีการในการติดตามเส้นขอบและการแบ่งในแนว XY ที่ใช้ในงานวิจัย วิธีการที่ 2 คือ วิธีการแพร่โดยอาศัยพิกัดตำแหน่งเป้าหมาย ซึ่งเป็นวิธีการที่พัฒนาขึ้นมาเพื่อเปรียบเทียบผลกับวิธีการแรก และวิธีการสุดท้าย คือ การแบ่งซ้ำตามแนว แกน XY ซึ่งเป็นวิธีการทำงานดั้งเดิมที่มีความเร็วในการประมวลผลสูง แต่มีการเข้าถึงข้อมูลของจุดภาพมากเกินไป

ตารางที่ ข.1 แสดงผลการสำหรับการแบ่งขอบเขตภาพเอกสารขนาดเล็ก

ภาพ ที่	ประเภท เอกสาร	ขนาดของ ภาพ (พิกเซล)	วิธีการ					
			1		2		3	
			เวลา (วินาที)	จำนวน พิกเซลที่ เข้าถึง	เวลา (วินาที)	จำนวน พิกเซลที่ เข้าถึง	เวลา (วินาที)	จำนวน พิกเซลที่ เข้าถึง
1	T	840 x 1314	0.050	353,952	0.060	353,952	0.050	1,045,424
2	T	942 x 1149	0.050	354,402	0.050	354,402	0.040	519,824
3	T	1216x1504	0.110	766,622	0.140	781,908	0.070	1,578,880
4	T	1225x1053	0.060	626,157	0.060	626,157	0.060	1,366,512
5	T	1235x1067	0.060	789,237	0.070	789,237	0.060	1,463,200
6	T	1290x1149	0.060	524,941	0.080	524,941	0.060	1,005,872
7	T&P	1049x877	0.040	245,336	0.040	245,336	0.070	379,840
8	T&P	1087x1161	0.080	483,042	0.070	483,042	0.050	724,416
9	T&P	1260x1307	0.050	317,991	0.060	317,991	0.040	526,574
10	T&P	1105x1210	0.050	531,917	0.050	531,917	0.050	1,250,096
11	T&P	1482x1668	0.120	1,337,257	0.120	1,337,257	0.120	2,309,232
12	T&P	1578x1488	0.101	1,937,136	0.100	1,937,136	0.170	3,809,232
13	C	1057x991	0.100	413,022	0.090	499,380	0.060	843,808
14	C	1332x1303	0.050	275,169	0.050	467,164	0.050	852,352
15	C	1260x1307	0.070	478,327	0.080	480,483	0.050	877,024
16	C	1062x1255	0.060	709,547	0.060	766,588	0.060	1,378,688

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตารางที่ ข.2 แสดงผลการสำหรับการแบ่งขอบเขตภาพเอกสารขนาดกลาง

ภาพ ที่	ประเภท เอกสาร	ขนาดของ ภาพ (พิกเซล)	วิธีการ					
			1		2		3	
			เวลา (วินาที)	จำนวน พิกเซลที่ เข้าถึง	เวลา (วินาที)	จำนวน พิกเซลที่ เข้าถึง	เวลา (วินาที)	จำนวน พิกเซลที่ เข้าถึง
1	T	1535x2417	0.200	1,885,192	0.211	1,885,192	0.231	4,667,056
2	T	1583x2472	0.190	2,214,476	0.200	2,214,476	0.210	4,937,952
3	T	1543x2377	0.201	2,957,658	0.241	2,957,658	0.291	5,993,088
4	T	1511x2093	0.170	1,987,274	0.180	2,005,985	0.201	4,005,424
5	T	1527x2361	0.160	2,217,258	0.170	2,217,258	0.230	4,982,256
6	T	1455x2249	0.160	2,372,427	0.150	2,372,427	0.220	5,123,792
7	T&P	1542x2336	0.180	2,161,530	0.210	2,161,530	0.211	4,959,664
8	T&P	1607x2334	0.151	2,148,598	0.190	2,148,598	0.190	4,783,152
9	T&P	1519x2449	0.170	2,216,355	0.210	2,216,355	0.220	5,398,496
10	T&P	1407x2086	0.150	1,128,231	0.150	1,128,231	0.120	1,695,392
11	T&P	1367x2056	0.141	1,589,744	0.160	1,589,744	0.130	3,151,872
12	T&P	1351x2009	0.140	1,214,447	0.141	1,214,447	0.110	2,673,824
13	T&P	1367x2065	0.130	1,053,019	0.150	1,053,019	0.100	1,966,320
14	T&P	1407x2049	0.140	1,051,193	0.160	1,051,193	0.100	1,936,480
15	T&P	1426x2047	0.150	1,112,478	0.150	1,112,478	0.130	2,543,200
16	T&P	1443x1943	0.140	1,109,173	0.150	1,109,173	0.140	3,223,664
17	T&P	1470x2364	0.170	2,083,765	0.181	2,083,765	0.170	3,921,360
18	T&P	1727x2234	0.170	1,446,883	0.190	1,446,883	0.170	3,293,264
19	T&P	1526x2065	0.170	1,760,162	0.180	1,760,162	0.170	3,642,656
20	C	1495x2329	0.160	1,553,133	0.221	1,639,424	0.180	3,349,088
21	C	1422x2053	0.140	1,206,930	0.150	1,287,950	0.130	2,062,176
22	C	1459x2099	0.160	1,532,265	0.170	1,706,024	0.150	3,287,584
23	C	1380x1981	0.130	925,212	0.141	1,009,638	0.110	2,078,144

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ตารางที่ ข.3 แสดงผลสำหรับการแบ่งขอบเขตภาพเอกสารขนาดใหญ่ที่เป็นตัวอักษรอย่างเดียว

ภาพ ที่	ประเภท เอกสาร	ขนาดของ ภาพ (พิกเซล)	วิธีการ					
			1		2		3	
			เวลา (วินาที)	จำนวน พิกเซล ที่เข้าถึง	เวลา (วินาที)	จำนวน พิกเซล ที่เข้าถึง	เวลา (วินาที)	จำนวน พิกเซล ที่เข้าถึง
1	T	2254x3209	0.320	3,397,000	0.321	3,397,000	0.351	6,471,680
2	T	2532x3479	0.400	6,249,046	0.412	6,249,046	0.580	11,870,768
3	T	2382x3079	0.330	5,882,836	0.361	5,882,836	0.571	11,362,336
4	T	2224x3103	0.320	4,307,628	0.340	4,307,628	0.450	8,447,680
5	T	2280x3031	0.311	3,846,340	0.291	3,846,340	0.350	6,871,952
6	T	2316x3001	0.271	5,303,954	0.311	5,303,954	0.490	10,766,480
7	T	2350x3073	0.300	4,929,882	0.310	4,929,882	0.491	10,480,944
8	T	1470x2449	0.280	2,475,176	0.280	2,489,882	0.260	5,471,968
9	T	2132x3040	0.280	4,056,797	0.300	4,056,797	0.360	6,824,528
10	T	2072x3080	0.311	4,288,309	0.310	4,288,309	0.431	8,781,552
11	T	2145x3063	0.290	3,745,860	0.310	3,745,860	0.350	6,387,040
12	T	2308x3500	0.360	4,002,661	0.370	4,002,661	0.301	4,742,448
13	T	2332x3500	0.341	3,879,065	0.360	3,879,065	0.321	6,042,560
14	T	2528x3295	0.330	5,087,957	0.370	5,087,957	0.461	10,339,968
15	T	2048x3342	0.300	5,132,890	0.320	5,132,890	0.421	10,710,064
16	T	2437x3355	0.380	4,023,597	0.391	4,023,597	0.310	6,435,184
17	T	2043x3273	0.311	2,577,372	0.330	2,577,372	0.250	3,713,408
18	T	2275x3238	0.310	2,021,342	0.331	2,021,342	0.221	3,010,672
19	T	2311x3355	0.330	3,757,608	0.340	3,757,608	0.320	6,714,768
20	T	2368x3328	0.330	3,731,340	0.361	3,731,340	0.290	4,986,048
21	T	2176x3124	0.320	1,458,368	0.320	1,458,368	0.300	2,845,674
22	T	2304x3271	0.320	3,004,557	0.340	3,004,557	0.271	3,467,856
23	T	2304x3271	0.310	3,004,557	0.331	3,004,557	0.251	3,467,856
24	T	2368x3289	0.340	3,099,358	0.340	3,099,358	0.270	4,534,752
25	T	2368x3289	0.330	3,329,210	0.350	3,341,657	0.300	5,542,604
26	T	2272x3276	0.320	4,151,180	0.351	4,151,180	0.300	6,347,040
27	T	2304x3252	0.310	3,368,385	0.321	3,368,385	0.281	4,606,704
28	T	2368x3289	0.340	3,341,087	0.330	3,344,883	0.291	5,166,496
29	T	2304x3245	0.320	2,353,497	0.311	2,353,497	0.230	3,441,408

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.3 (ต่อ)

ภาพ ที่	ประเภท เอกสาร	ขนาดของ ภาพ (พิกเซล)	วิธีการ					
			1		2		3	
			เวลา (วินาที)	จำนวน พิกเซล ที่เข้าถึง	เวลา (วินาที)	จำนวน พิกเซล ที่เข้าถึง	เวลา (วินาที)	จำนวน พิกเซล ที่เข้าถึง
30	T	2368x3289	0.321	3,714,951	0.331	3,714,951	0.320	6,258,192
31	T	2368x3289	0.350	3,902,923	0.351	3,902,923	0.350	7,025,600
32	T	2368x3289	0.330	2,667,251	0.331	2,667,251	0.281	4,587,056
33	T	2368x3289	0.341	2,667,251	0.330	2,667,251	0.290	4,587,056
34	T	2304x3258	0.330	3,092,092	0.321	3,092,092	0.230	3,669,552
35	T	2352x3243	0.310	3,296,900	0.330	3,296,900	0.280	4,570,800
36	T	2544x3295	0.350	5,331,201	0.341	5,331,201	0.460	10,544,144
37	T	2240x3116	0.310	3,223,983	0.300	3,223,983	0.290	5,589,568
38	T	2275x3238	0.311	2,085,072	0.320	2,132,227	0.220	3,010,672
39	T	2362x3238	0.330	3,380,872	0.330	3,380,872	0.311	6,191,168
40	T	2104x3342	0.330	4,973,452	0.320	4,973,452	0.390	10,076,112
41	T	2128x3342	0.330	4,495,382	0.341	4,495,382	0.361	8,342,624
42	T	2048x3342	0.301	5,132,890	0.310	5,132,890	0.431	10,710,064
43	T	2399x3153	0.320	3,682,399	0.321	3,682,399	0.290	5,824,240
44	T	2311x3355	0.340	3,757,608	0.341	3,757,608	0.310	6,714,768
45	T	2437x3355	0.380	3,962,737	0.360	3,962,737	0.331	6,357,632
46	T	2437x3355	0.371	4,082,002	0.361	4,082,002	0.330	6,756,144
47	T	2437x3355	0.370	3,861,576	0.361	3,861,576	0.310	6,277,632
48	T	2180x3329	0.340	4,876,419	0.340	4,900,251	0.351	8,066,048
49	T	2374x3128	0.321	3,694,382	0.340	3,694,382	0.270	4,972,208
50	T	2249x3393	0.360	4,097,262	0.350	4,097,262	0.331	7,022,032

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างภาพเอกสารขนาดใหญ่ (ภาพตัวอักษรอย่างเดี่ยว)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.4 แสดงผลสำหรับการแบ่งขอบเขตภาพเอกสารขนาดใหญ่ที่ประกอบด้วยรูปภาพและตัวอักษร

ภาพที่	ประเภทเอกสาร	ขนาดของภาพ (พิกเซล)	วิธีการ					
			1		2		3	
			เวลา (วินาที)	จำนวนพิกเซลที่เข้าถึง	เวลา (วินาที)	จำนวนพิกเซลที่เข้าถึง	เวลา (วินาที)	จำนวนพิกเซลที่เข้าถึง
1	T&P	2349x3111	0.301	3,509,514	0.320	3,509,514	0.301	5,804,032
2	T&P	2190x2988	0.271	2,483,155	0.261	2,483,155	0.220	4,190,448
3	T&P	2362x3057	0.301	2,792,629	0.290	2,792,629	0.270	4,860,816
4	T&P	2310x3081	0.300	3,107,954	0.280	3,107,954	0.281	5,661,488
5	T&P	2357x3097	0.301	3,126,434	0.301	3,126,434	0.280	2,593,328
6	T&P	2532x3479	0.360	5,710,224	0.360	5,710,224	0.400	8,917,984
7	T&P	2302x3185	0.321	3,534,899	0.310	3,534,899	0.300	6,026,912
8	T&P	2538x3483	0.371	4,226,391	0.381	4,226,391	0.410	7,674,960
9	T&P	2532x3479	0.371	3,913,912	0.371	3,913,912	0.430	7,043,808
10	T&P	2539x3484	0.371	3,889,198	0.350	3,889,198	0.420	9,076,160
11	T&P	2286x3133	0.301	2,441,040	0.321	2,441,040	0.240	3,969,616
12	T&P	2313x3041	0.310	3,272,103	0.301	3,272,103	0.250	4,852,748
13	T&P	2400x3247	0.361	2,404,634	0.340	2,404,634	0.270	4,195,536
14	T&P	2508x3264	0.350	3,370,606	0.351	3,370,606	0.300	5,446,480
15	T&P	2532x3517	0.321	4,558,969	0.320	4,558,969	0.370	8,770,544
16	T&P	2231x3049	0.271	4,029,541	0.271	4,029,541	0.350	7,941,744
17	T&P	2539x3156	0.340	2,687,684	0.330	2,687,684	0.290	4,549,536
18	T&P	2532x3276	0.370	3,768,810	0.351	3,768,810	0.330	6,183,088
19	T&P	2367x3156	0.350	2,674,089	0.330	2,674,089	0.331	5,575,648
20	T&P	2295x3144	0.311	2,567,776	0.300	2,567,776	0.260	4,340,496
21	T&P	2387x3154	0.311	2,708,916	0.311	2,708,916	0.361	7,776,880
22	T&P	2334x3097	0.290	3,114,934	0.300	3,114,934	0.291	5,389,744
23	T&P	2334x3175	0.360	3,110,227	0.351	3,110,227	0.310	5,083,456
24	T&P	2054x3034	0.360	2,742,573	0.291	2,742,573	0.240	4,034,896
25	T&P	2222x3113	0.320	2,085,780	0.291	2,085,780	0.361	4,291,440
26	T&P	2254x3201	0.300	2,404,639	0.301	2,404,639	0.280	5,096,096
27	T&P	2240x3025	0.280	1,994,998	0.280	1,994,998	0.210	3,221,184
28	T&P	2326x3297	0.300	3,861,127	0.310	3,861,127	0.330	6,932,864

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.4 (ต่อ)

ภาพ ที่	ประเภท เอกสาร	ขนาดของ ภาพ (พิกเซล)	วิธีการ					
			1		2		3	
			เวลา (วินาที)	จำนวน พิกเซลที่ เข้าถึง	เวลา (วินาที)	จำนวน พิกเซลที่ เข้าถึง	เวลา (วินาที)	จำนวน พิกเซลที่ เข้าถึง
29	T&P	2206x2985	0.270	1,949,262	0.270	1,949,262	0.211	2,850,688
30	T&P	2273x3080	0.300	3,078,658	0.291	3,078,658	0.270	4,930,576
31	T&P	2453x3329	0.330	3,997,451	0.330	3,997,451	0.360	7,556,144
32	T&P	2493x3273	0.321	4,628,590	0.331	4,628,590	0.431	9,588,608
33	T&P	2406x3233	0.321	4,825,011	0.320	4,825,011	0.370	8,191,136
34	T&P	2342x3305	0.321	3,926,280	0.330	3,926,280	0.311	6,411,984
35	T&P	2109x3065	0.280	4,060,428	0.280	4,060,428	0.250	4,923,872
36	T&P	2234x3169	0.340	3,611,030	0.350	3,611,030	0.300	5,204,016
37	T&P	2128x3342	0.311	4,483,948	0.331	4,483,948	0.331	7,961,248
38	T&P	2104x3342	0.300	4,370,763	0.331	4,370,763	0.311	7,087,168
39	T&P	2072x3342	0.301	4,585,986	0.331	4,585,986	0.330	7,061,248
40	T&P	2040x3342	0.290	3,678,538	0.291	3,678,538	0.310	6,907,264
41	T&P	2088x3342	0.300	4,436,135	0.331	4,436,135	0.311	7,074,032
42	T&P	2040x3342	0.300	5,018,986	0.310	5,018,986	0.321	7,717,904
43	T&P	2064x3342	0.320	5,198,139	0.300	5,198,139	0.390	8,735,952
44	T&P	2048x3342	0.311	4,587,079	0.321	4,587,079	0.341	7,103,712
45	T&P	2064x3342	0.290	4,334,893	0.300	4,334,893	0.350	7,378,816
46	T&P	2088x3342	0.321	5,098,762	0.321	5,098,762	0.360	9,112,640
47	T&P	2349x3468	0.351	3,819,529	0.360	3,819,529	0.280	5,132,048
48	T&P	2202x3408	0.331	2,889,918	0.340	2,889,918	0.270	3,111,424
49	T&P	2324x3342	0.320	3,888,979	0.341	3,888,979	0.310	6,433,136
50	T&P	2387x3308	0.340	4,500,574	0.340	4,500,574	0.441	9,310,240
51	T&P	2168x3321	0.310	4,727,778	0.311	4,727,778	0.361	8,637,824
52	T&P	2096x3342	0.391	3,988,410	0.310	3,988,410	0.330	7,127,600

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างภาพเอกสารขนาดใหญ่ (รูปภาพกับตัวอักษร)

**SCIENCE**

**Stopping Rogue Nukes**  
BY MICHAEL

**Ant Builds On**

**China Development Center Opens**  
BY MICHAEL

**Power Struggle**  
BY MICHAEL

**Vintage Video Games**

**Rapidly integrate all the benefits of Autonomy inside your Oracle application**

(ก) 2190 x 2988 พิกเซล (ข) 2400 x 3247 พิกเซล (ค) 2539 x 3156 พิกเซล

(ง) 2453 x 3329 พิกเซล (จ) 2342 x 3305 พิกเซล (ฉ) 2206 x 2985 พิกเซล

(ช) 2326 x 3297 พิกเซล (ซ) 2273 x 3080 พิกเซล (ฌ) 2190 x 3065 พิกเซล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.5 แสดงผลสำหรับการแบ่งขอบเขตภาพเอกสารขนาดใหญ่ที่ประกอบด้วยรูปภาพและ  
ตัวอักษรที่มีโครงสร้างติดกัน

ภาพที่	ประเภทเอกสาร	ขนาดของภาพ (พิกเซล)	วิธีการ					
			1		2		3	
			เวลา (วินาที)	จำนวนพิกเซลที่เข้าถึง	เวลา (วินาที)	จำนวนพิกเซลที่เข้าถึง	เวลา (วินาที)	จำนวนพิกเซลที่เข้าถึง
1	C	2209x3066	0.341	2,277,414	0.371	2,446,044	0.221	2,783,904
2	C	2214x3065	0.321	2,490,885	0.330	2,523,599	0.211	3,899,264
3	C	2278x3073	0.330	2,697,072	0.370	2,856,244	0.271	5,118,912
4	C	2238x3097	0.360	2,074,106	0.381	2,442,221	0.240	3,398,240
5	C	2174x2913	0.340	1,700,795	0.361	2,251,353	0.201	2,205,696
6	C	2437x3252	0.431	2,275,230	0.441	2,462,274	0.321	4,169,120
7	C	2463x3016	0.411	2,078,922	0.481	2,197,809	0.300	3,731,824
8	C	2222x3483	0.310	2,983,807	0.341	3,330,025	0.260	4,889,152
9	C	2532x3129	0.370	3,184,292	0.400	3,475,408	0.301	5,128,304
10	C	2413x3224	0.351	2,774,473	0.400	3,069,267	0.270	3,993,616
11	C	2230x3081	0.331	3,033,449	0.350	3,200,810	0.241	4,420,992
12	C	2275x3090	0.330	3,007,063	0.371	3,189,181	0.250	4,855,392
13	C	2341x3071	0.330	2,149,964	0.340	2,149,964	0.240	3,351,200
14	C	2250x3035	0.321	1,650,387	0.341	1,664,927	0.220	2,869,024
15	C	2283x3210	0.371	2,422,377	0.401	2,687,599	0.230	3,613,712
16	C	2402x3240	0.381	2,508,772	0.411	2,887,390	0.290	4,907,760
17	C	2310x3113	0.350	2,411,207	0.360	2,584,029	0.230	3,811,040
18	C	2532x3479	0.421	3,606,303	0.451	3,919,262	0.401	7,551,872
19	C	2280x3229	0.351	3,148,688	0.361	3,205,765	0.300	5,368,544
20	C	2322x3199	0.391	2,249,230	0.401	2,406,984	0.310	4,273,616
21	C	2350x3265	0.370	2,979,567	0.401	3,210,409	0.311	5,987,952
22	C	2314x3136	0.340	2,337,980	0.371	2,586,894	0.280	4,351,248
23	C	2234x3096	0.361	1,830,758	0.360	1,911,986	0.300	3,139,936
24	C	2253x3154	0.330	2,981,445	0.350	2,960,147	0.290	5,277,536
25	C	2244x3103	0.321	1,917,513	0.351	2,172,938	0.250	3,519,920
26	C	2214x3086	0.331	2,691,699	0.370	2,714,192	0.300	5,342,016
27	C	2451x3223	0.370	2,773,572	0.411	3,102,003	0.340	5,780,688
28	C	2300x3192	0.361	3,116,836	0.451	3,219,500	0.260	5,009,376
29	C	2302x3121	0.340	2,347,172	0.371	2,520,214	0.270	4,807,936

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.5 (ต่อ)

ภาพที่	ประเภทเอกสาร	ขนาดของภาพ (พิกเซล)	วิธีการ					
			1		2		3	
			เวลา (วินาที)	จำนวนพิกเซล ที่เข้าถึง	เวลา (วินาที)	จำนวนพิกเซล ที่เข้าถึง	เวลา (วินาที)	จำนวนพิกเซล ที่เข้าถึง
30	C	2413x3109	0.371	2,812,406	0.381	2,983,646	0.331	6,090,272
31	C	2302x3233	0.361	2,488,242	0.360	2,760,824	0.301	5,337,184
32	C	2285x3033	0.330	2,173,213	0.331	2,452,619	0.250	3,838,624
33	C	2309x3149	0.341	3,115,501	0.340	3,207,036	0.301	5,105,232
34	C	2350x3160	0.381	3,107,589	0.371	3,252,896	0.380	6,937,360
35	C	2109x3065	0.331	2,116,751	0.350	2,415,423	0.251	3,490,160
36	C	2342x3073	0.350	2,068,817	0.410	2,220,636	0.270	3,770,752
37	C	2387x3153	0.371	3,999,538	0.381	3,983,968	0.320	5,977,936
38	C	2392x3219	0.370	2,395,491	0.400	2,951,318	0.271	4,523,856
39	C	2531x3478	0.421	3,878,318	0.471	4,203,344	0.391	7,156,032
40	C	2531x3478	0.421	3,829,634	0.460	4,040,888	0.411	7,663,248
41	C	2160x2860	0.300	1,966,151	0.311	2,252,382	0.220	3,602,288
42	C	2211x3090	0.371	2,864,120	0.380	2,878,456	0.380	5,282,704
43	C	2237x3393	0.371	3,733,729	0.390	3,813,716	0.281	5,866,592
44*	C	2299x3267	0.360	3,609,924	0.371	3,824,318	0.310	5,656,576
45	C	2201x3172	0.340	3,205,302	0.361	3,448,979	0.301	5,600,928
46	C	2352x3000	0.331	2,242,911	0.390	2,622,169	0.241	3,109,184
47	C	2349x3173	0.361	3,833,309	0.370	4,065,394	0.251	3,718,688

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างภาพเอกสารขนาดใหญ่ (ภาพที่เกิดการติดกันระหว่างภาพกับกลุ่มตัวอักษร)



(ก) 2223 x 3065 พิกเซล



(ข) 1511 x 2093 พิกเซล



(ค) 2392 x 3219 พิกเซล



(ง) 2285 x 3033 พิกเซล



(จ) 2309 x 3149 พิกเซล



(ฉ) 2350 x 3160 พิกเซล



(ช) 2230 x 3081 พิกเซล



(ซ) 2262 x 3097 พิกเซล



(ฌ) 2463 x 3016 พิกเซล

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้ในเชิงการค้าโดยไม่ได้รับอนุญาต ไม่อนุญาตให้ทำซ้ำโดยไม่ได้รับอนุญาต หรือเผยแพร่โดยไม่ได้รับอนุญาต  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.6 แสดงผลสำหรับการแบ่งขอบเขตภาพเอกสารขนาดใหญ่ที่มีโครงสร้างส่วนประกอบ  
เป็นสี่เหลี่ยม

ภาพ ที่	ประเภท เอกสาร	ขนาดของ ภาพ (พิกเซล)	วิธีการ					
			1		2		3	
			เวลา (วินาที)	จำนวนพิกเซล ที่เข้าถึง	เวลา (วินาที)	จำนวนพิกเซล ที่เข้าถึง	เวลา (วินาที)	จำนวนพิกเซล ที่เข้าถึง
1	R	2538x3483	0.380	4,226,391	0.451	4,226,391	0.411	7,674,960
2	R	2190x2988	0.260	2,483,155	0.320	2,483,155	0.270	4,190,448
3	R	2367x3156	0.360	2,674,089	0.410	2,674,089	0.330	5,575,648
4	R	2234x3175	0.350	3,110,227	0.430	3,110,227	0.290	5,083,456
5	R	1119x1533	0.071	377,918	0.080	377,918	0.050	577,440
6	R	2362x3057	0.291	2,792,629	0.361	2,792,629	0.271	4,860,816
7	R	2350x3073	0.290	4,929,882	0.381	4,929,882	0.421	10,480,944
8	R	2180x3329	0.321	4,876,419	0.401	4,900,251	0.351	8,606,048
9	R	2308x3500	0.351	4,002,661	0.441	4,002,661	0.291	4,742,448
10	R	2332x3500	0.341	3,879,065	0.411	3,879,065	0.320	6,042,560
11	R	2528x3295	0.330	5,087,957	0.410	5,087,957	0.471	10,339,968
12	R	2304x3273	0.320	2,577,372	0.380	2,577,372	0.250	3,713,408
13	R	2275x3238	0.300	2,021,342	0.381	2,021,342	0.210	3,010,672
14	R	2311x3355	0.331	3,757,608	0.401	3,757,608	0.320	6,714,768
15	R	2304x3258	0.321	3,092,092	0.401	3,092,092	0.241	3,669,552
16	R	2544x3295	0.330	5,331,201	0.450	5,331,201	0.451	10,544,144
17	R	2040x3342	0.270	3,162,578	0.351	3,678,538	0.301	6,907,264
18	R	2362x3238	0.310	3,380,872	0.401	3,380,872	0.310	6,191,168
19	R	2311x3355	0.330	3,757,608	0.411	3,757,608	0.320	6,714,768
20	R	2254x3209	0.310	3,397,000	0.370	3,397,000	0.310	6,471,680
21	R	2532x3479	0.390	6,249,046	0.420	6,249,046	0.491	11,870,768
22	R	2382x3097	0.330	5,882,836	0.411	5,882,836	0.481	11,362,336
23	R	2224x3103	0.310	4,307,628	0.391	4,307,628	0.391	8,447,680
24	R	2280x3031	0.281	3,846,340	0.340	3,846,340	0.301	6,871,952
25	R	2316x3001	0.291	5,303,954	0.340	5,303,954	0.410	10,766,480
26	R	2209x3066	0.310	1,700,795	0.391	2,251,353	0.170	2,205,696
27	R	2262x3097	0.290	3,033,449	0.381	3,200,810	0.240	4,420,992
28	R	2275x3090	0.290	3,007,063	0.381	3,189,181	0.250	4,855,392

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.6 (ต่อ)

ภาพ ที่	ประเภท เอกสาร	ขนาดของ ภาพ (พิกเซล)	วิธีการ					
			1		2		3	
			เวลา (วินาที)	จำนวนพิกเซล ที่เข้าถึง	เวลา (วินาที)	จำนวนพิกเซล ที่เข้าถึง	เวลา (วินาที)	จำนวนพิกเซล ที่เข้าถึง
29	R	2250x3035	0.291	1,650,387	0.350	1,664,927	0.210	2,869,024
30	R	2402x3240	0.321	2,508,772	0.461	2,887,390	0.300	4,907,760
31	R	2532x3479	0.381	3,606,303	0.471	3,919,262	0.421	7,551,872
32	R	2214x3065	0.280	2,490,885	0.360	2,523,599	0.220	3,899,264
33	R	2322x3199	0.330	2,249,203	0.430	2,406,984	0.300	4,273,616
34	R	2314x3136	0.311	2,337,980	0.410	2,586,894	0.271	4,351,248
35	R	2244x3103	0.281	1,917,513	0.381	2,172,938	0.220	3,519,920
36	R	2302x3233	0.310	2,488,242	0.401	2,760,824	0.301	5,337,184
37	R	2342x3073	0.291	2,068,817	0.371	2,220,636	0.250	3,770,752
38	R	2238x3097	0.311	2,074,106	0.421	2,442,221	0.250	3,398,240
39	R	2299x3267	0.331	3,609,924	0.401	3,609,924	0.301	5,656,576
40	R	2174x2913	0.300	4,483,948	0.401	4,483,948	0.350	7,961,248

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# ตัวอย่างภาพเอกสารขนาดใหญ่ (โครงสร้างส่วนประกอบเป็นสี่เหลี่ยม)

## AUTOMOTIVE

### 2001 Volvo S40

The new Volvo S40 is the first car in the world to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque.



The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque.

The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque.

The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque.

The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque.

The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque.

The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque. The car is also the first to feature a 2.0-liter engine with 150 hp and 140 lb-ft of torque.



## Nuclear Power Gets a Second Look

By Tom Ichniowski, Editor in Chief, ENR

The nuclear industry is looking for a second chance. After a decade of setbacks, the industry is now looking for a second chance. The industry is now looking for a second chance. The industry is now looking for a second chance.



A photograph of a nuclear power plant facility with large cooling towers.

(ก) 2238 x 3097 พิกเซล

(ข) 2040 x 3342 พิกเซล

(ค) 2334 x 3175 พิกเซล

Polycarbonate with Core-Shell Structured Latex Particles. This article discusses the use of polycarbonate with core-shell structured latex particles. It covers the properties of these materials and their applications in various industries.

Polycarbonate with Core-Shell Structured Latex Particles. This article discusses the use of polycarbonate with core-shell structured latex particles. It covers the properties of these materials and their applications in various industries.

Polycarbonate with Core-Shell Structured Latex Particles. This article discusses the use of polycarbonate with core-shell structured latex particles. It covers the properties of these materials and their applications in various industries.

(ง) 2304 x 3273 พิกเซล

(ฉ) 2311 x 3355 พิกเซล

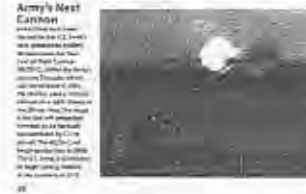
(ช) 2362 x 3238 พิกเซล



Planet Watch. This is a screenshot from the game Planet Watch, showing a space-themed interface with various elements and text.

Planet Watch. This is a screenshot from the game Planet Watch, showing a space-themed interface with various elements and text.

Planet Watch. This is a screenshot from the game Planet Watch, showing a space-themed interface with various elements and text.



Army's Next. This is a screenshot from the game Army's Next, showing a military-themed interface with various elements and text.



World. This is a screenshot from the game World, showing a globe and various data points.

World. This is a screenshot from the game World, showing a globe and various data points. It includes a small advertisement for 'Put your Power Quality Analyzer'.

(ซ) 2254 x 3031 พิกเซล

(ฅ) 2314 x 3136 พิกเซล

(ด) 2244 x 3103 พิกเซล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



## ผลการทดลองการประเมินผลลัพธ์เชิงปริมาณ

สำหรับในตารางที่ ข.7-ข.10 แสดงการเปรียบเทียบประสิทธิภาพการทำงานระหว่างวิธีการในการติดตามเส้นขอบและการแบ่งในแนว XY กับวิธีการแบ่งซ้ำตามแนวแกน X-Y ซึ่งรายละเอียดในตารางประกอบด้วยค่าต่างๆ ดังนี้

-Pic.No. คือ หมายเลขลำดับของภาพที่ใช้ทดสอบ

-T คือ ภาพเอกสารทดสอบที่ทั้งภาพประกอบไปด้วยตัวอักษรอย่างเดียว

-T & P คือ ภาพเอกสารทดสอบที่ประกอบด้วยรูปภาพและตัวอักษรที่มีโครงสร้างไม่ติดกัน

-C คือ ภาพเอกสารทดสอบที่ประกอบด้วยภาพตัวอักษรอย่างเดียว หรือประกอบด้วยรูปภาพและตัวอักษรที่มีโครงสร้างติดกัน

- N คือ ค่าอัตราส่วน Normalize ที่จะใช้บ่งบอกผลในการตัดแบ่งว่าให้ผลเป็น Over-Segmentation หรือ Under-Segmentation โดยที่ค่าตัวเลขตัวเลขแสดงผลที่ได้จากการแบ่งด้วยวิธีการนั้นๆ ตัวส่วนแสดงจำนวนส่วนประกอบที่ได้จากการประเมินด้วยสายคา ซึ่งค่าของ N แบ่งเป็น 3 ระดับดังนี้

1. ค่า  $N = 1$  ถือว่าผลการแบ่งภาพเอกสาร ได้เท่ากับส่วนประกอบที่ต้องการจากการประเมินด้วยสายคา

2. ค่า  $N < 1$  ถือว่าผลของการแบ่งส่วนประกอบภาพเอกสารเกิด Under-Segmentation

3. ค่า  $N > 1$  ถือว่าผลของการแบ่งส่วนประกอบภาพเอกสารเกิด Over-Segmentation

- Cor. คือ ค่าความถูกต้อง (Correctness) ของส่วนประกอบที่เกิดจากผลลัพธ์ของการแบ่งด้วยวิธีการนั้นๆ เมื่อเทียบกับส่วนประกอบที่ได้จากการประเมินด้วยสายคา

- error คือ ค่าความผิดพลาดของผลลัพธ์ที่ได้จากวิธีการนั้นๆ

ตัวอย่างเช่น ภาพเอกสารหมายเลข 1 ได้รับการประเมินส่วนประกอบโดยอาศัยการลดขนาดภาพให้เล็กลงแล้วแบ่งส่วนประกอบ โดยการประเมินด้วยสายคาว่ามีส่วนประกอบเท่ากับ 9 ส่วน และผลการแบ่งส่วนประกอบโดยใช้การติดตามเส้นขอบและการแบ่งในแนว XY ได้ส่วนประกอบทั้งหมด 10 ส่วน และภายใน 10 ส่วนที่ได้มีความถูกต้องตามการประเมินด้วยสายคาอยู่ 8 ส่วน มีการแบ่งผิดพลาดอีก 2 ส่วน ดังนั้นเราจึงสรุปค่าต่างๆ ได้คือ  $N = 10/9$ ,  $Cor. = 8$  และ  $error = 2$  เป็นต้น

ทั้งนี้การพิจารณาความถูกต้องของวิธีการนั้นจำเป็นที่จะต้องพิจารณาค่าทั้งสามประกอบกัน เพราะในบางครั้งผลของการแบ่งส่วนประกอบภาพเอกสารวิธีการนั้นๆ ได้จำนวนที่ใกล้เคียงกับความเป็นจริง แต่ที่จริงแล้วเป็นการแบ่งผิดส่วนจากที่ต้องการอันเป็นผลมาจาก Under หรือ Over-Segment ก็ตาม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.7 เปรียบเทียบผลการแบ่งส่วนประกอบเชิงปริมาณ โดยเทียบกับจำนวนส่วนประกอบที่ประเมินด้วยสายตาสำหรับภาพเอกสารขนาดเล็ก

Pic No.	Contour and XY Cut Technique									Recursive X-Y Cut Technique								
	T			T&P			C			T			T&P			C		
	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N
1	1	2	3/2	6	0	6/6	4	1	5/4	1	29	30/2	2	5	7/6	1	1	2/4
2	3	2	5/4	4	0	4/4	2	1	3/2	2	16	18/4	0	3	3/4	0	1	½
3	3	2	5/4	0	2	2/4	0	2	2/6	1	19	20/4	0	4	4/4	0	6	6/6
4	3	1	4/6	2	2	4/6	0	3	3/3	6	2	8/6	2	1	3/6	0	15	15/3
5	3	0	3/3	1	2	2/3	-	-	-	0	6	6/3	0	20	20/3	-	-	-
6	1	0	1/1	2	1	3/4	-	-	-	0	31	31/1	2	1	3/1	-	-	-
Avg.	2.333	1.167	1.111	2.500	1.167	0.764	1.500	1.750	1.021	1.667	17.167	8.767	1.000	5.667	1.806	0.250	5.750	1.750

ตารางที่ ข.8 เปรียบเทียบผลการแบ่งส่วนประกอบเชิงปริมาณ โดยเทียบกับจำนวนส่วนประกอบที่ประเมินด้วยสายตาสำหรับภาพเอกซารขนาดกลาง

Pic No.	Contour and XY Cut Technique									Recursive X-Y Cut Technique								
	T			T&P			C			T			T&P			C		
	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N
1	5	2	7/9	8	1	9/10	4	2	6/9	3	63	66/9	2	48	50/10	8	9	17/9
2	4	7	11/8	7	0	7/7	5	1	6/5	3	44	47/8	4	14	18/7	3	2	5/5
3	12	15	27/25	6	0	6/6	9	1	10/9	4	46	50/25	0	70	70/6	4	21	25/5
4	5	1	6/9	4	0	4/4	5	1	6/7	4	14	18/9	0	10	10/4	4	20	24/7
5	4	4	8/9	9	0	9/9	-	-	-	1	30	31/9	5	15	20/9	-	-	-
6	12	2	14/17	15	2	17/21	-	-	-	12	16	28/17	3	38	41/21	-	-	-
7	-	-	-	6	0	6/6	-	-	-	-	-	-	1	26	27/6	-	-	-
8	-	-	-	8	0	8/8	-	-	-	-	-	-	3	8	11/8	-	-	-
9	-	-	-	6	0	6/6	-	-	-	-	-	-	3	20	23/6	-	-	-
10	-	-	-	8	0	8/8	-	-	-	-	-	-	3	4	7/8	-	-	-
11	-	-	-	13	0	13/13	-	-	-	-	-	-	9	16	25/13	-	-	-
12	-	-	-	2	1	3/4	-	-	-	-	-	-	2	7	9/4	-	-	-
13	-	-	-	11	2	13/14	-	-	-	-	-	-	10	4	14/14	-	-	-
Avg.	7.000	5.167	0.935	7.923	0.426	0.952	5.750	1.250	0.959	4.500	35.500	2.000	3.500	23.000	2.464	4.750	13.000	2.829

ตารางที่ ข.9 เปรียบเทียบผลการแบ่งส่วนประกอบเชิงวิวัฒนาการ โดยเทียบกับจำนวนส่วนประกอบที่ประเมินด้วยสายตาสำหรับภาพเอกสารขนาดใหญ่

Pic No.	Contour and XY Cut Technique									Recursive X-Y Cut Technique								
	T			T&P			C			T			T&P			C		
	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N
1	38	0	38/38	10	1	11/12	5	4	9/10	35	13	48/38	6	2	8/12	1	2	3/10
2	62	15	77/93	9	2	11/13	9	4	13/13	68	8	76/93	7	3	10/13	9	1	10/13
3	8	4	12/10	6	1	7/8	7	5	9/12	6	11	17/12	4	1	5/8	3	14	17/12
4	12	4	16/14	4	1	5/8	8	2	10/10	7	22	29/14	5	7	12/8	3	1	4/10
5	13	3	16/14	8	2	10/12	6	5	11/9	8	12	20/14	9	1	10/12	2	4	6/9
6	7	7	14/8	12	0	12/12	7	1	8/9	5	15	20/8	10	2	12/12	3	1	4/9
7	14	1	15/14	13	0	13/13	5	3	8/10	6	68	74/14	7	2	9/13	5	2	7/10
8	6	14	20/8	18	5	24/18	6	4	10/9	4	63	67/8	8	12	20/18	5	1	6/9
9	10	42	52/15	9	12	21/14	7	3	10/9	3	18	21/15	8	18	26/14	4	1	5/9
10	12	15	27/20	14	4	18/17	8	3	11/14	6	71	77/20	8	11	11/17	4	2	6/14
11	3	12	15/5	9	4	13/14	7	5	12/10	2	11	13/5	6	7	13/14	5	2	7/10
12	24	0	24/24	8	3	11/11	8	1	9/13	3	3	6/24	4	119	123/14	3	3	6/13
13	14	2	16/17	16	3	19/17	7	1	8/9	8	7	15/17	9	5	14/17	3	2	5/9
14	11	5	17/11	13	2	15/13	2	2	4/6	3	16	21/11	7	2	9/12	6	3	9/6
15	8	15	23/12	11	0	11/11	14	5	14/19	8	36	44/12	7	11	19/11	17	7	24/19
16	13	4	17/26	16	2	18/19	17	1	18/19	3	21	24/26	5	10	15/19	16	4	20/19

ตารางที่ ข.9 (ต่อ)

Pic No.	Contour and XY Cut Technique									Recursive X-Y Cut Technique								
	T			T&P			C			T			T&P			C		
	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N
17	9	3	12/15	9	1	10/11	8	6	14/13	0	7	7/15	8	12	20/11	13	12	25/13
18	7	1	8/7	11	1	12/11	9	7	16/28	2	2	4/7	8	5	13/11	16	12	28/28
19	22	7	29/22	9	3	12/15	21	1	22/21	11	11	22/22	4	6	10/15	2	21	23/21
20	5	3	8/10	13	4	17/19	6	2	8/10	1	14	15/10	8	4	12/19	2	12	15/10
21	3	11	14/15	9	1	10/11	6	4	11/9	3	11	14/5	6	16	22/11	4	2	6/9
22	15	3	18/22	9	4	13/17	7	3	10/10	4	3	7/22	5	4	9/17	3	3	6/10
23	18	6	24/25	8	2	10/9	12	3	15/15	4	3	7/25	3	35	38/9	10	1	11/15
24	16	2	18/18	19	1	20/21	19	0	19/19	2	8	10/18	12	7	19/21	11	2	13/19
25	13	6	19/21	10	3	13/14	6	1	7/8	6	11	17/21	8	1	9/14	4	1	5/8
26	25	3	28/27	17	3	20/22	4	1	5/9	8	12	20/31	6	10	16/22	3	28	31/9
27	29	1	30/31	6	0	6/6	9	0	9/9	9	6	15/31	5	6	11/6	5	2	7/9
28	17	5	22/23	4	1	5/6	10	2	12/12	3	15	18/23	2	23	25/6	5	1	6/12
29	6	2	8/10	7	0	7/7	13	2	15/14	3	2	5/10	1	3	4/7	7	45	52/14
30	20	3	23/26	17	0	17/17	10	2	12/11	21	2	23/26	9	4	13/17	7	14	21/11
31	20	7	27/25	5	3	8/10	8	0	8/8	11	16	27/25	0	2	2/10	5	7	12/8
32	15	1	16/17	17	1	18/19	8	2	10/9	8	7	15/17	6	7	13/19	2	14	16/9

ตารางที่ ข.9 (ต่อ)

Pic No.	Contour and XY Cut Technique									Recursive X-Y Cut Technique								
	T			T&P			C			T			T&P			C		
	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N
33	15	1	16/17	14	5	19/20	7	2	9/8	8	7	15/17	3	23	26/20	4	13	17/8
34	23	3	26/23	6	2	8/10	18	3	21/19	6	2	8/23	9	12	21/10	14	8	22/19
35	10	4	14/15	19	3	21/22	7	0	7/7	6	4	10/15	8	26	34/22	3	1	4/7
36	5	6	11/8	21	1	22/26	9	0	9/9	5	20	25/8	9	16	25/26	4	1	5/9
37	21	6	27/24	21	8	29/34	14	6	20/24	14	13	27/24	9	30	39/34	1	6	7/24
38	5	1	6/7	14	4	18/18	12	7	19/18	2	2	4/7	5	19	24/18	3	4	7/18
39	17	4	21/16	12	5	17/11	16	8	24/34	20	5	25/16	3	43	46/11	21	8	29/34
40	16	14	30/20	8	1	9/9	10	10	20/21	19	42	61/20	5	7	13/9	4	10	14/21
41	12	4	16/12	4	27	31/12	15	1	16/18	6	35	41/12	7	23	30/12	13	13	26/18
42	20	4	24/30	6	15	21/12	8	0	8/8	20	24	44/30	5	19	24/12	7	2	9/8
43	16	2	18/16	8	5	13/11	7	0	7/7	11	2	13/16	9	5	14/11	2	7	9/9
44	28	2	30/28	7	12	19/12	9	0	9/9	12	9	21/28	6	13	19/12	2	7	9/9
45	13	6	19/27	16	22	38/20	1	5	6/9	4	50	54/27	9	14	23/20	2	18	16/9
46	17	2	19/20	10	15	25/13	4	8	12/11	6	8	14/20	10	18	28/13	3	1	4/11
47	14	3	17/27	4	31	35/10	7	7	14/13	2	48	50/27	7	15	23/10	5	4	9/13
48	11	14	25/15	12	4	16/14	-	-	-	4	25	29/15	6	9	15/14	-	-	-

ตารางที่ ๗.๑ (ต่อ)

Pic No.	Contour and XY Cut Technique									Recursive X-Y Cut Technique								
	T			T&P			C			T			T&P			C		
	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N	Cor.	error	N
49	10	7	17/10	12	22	34/15	-	-	-	1	9	10/10	9	25	36/14	-	-	-
50	10	4	14/12	9	0	9/9	-	-	-	10	29	39/12	2	16	18/9	-	-	-
51	-	-	-	10	5	15/16	-	-	-	-	-	-	3	2	5/16	-	-	-
52	-	-	-	7	15	23/9	-	-	-	-	-	-	6	9	15/9	-	-	-
Avg.	15.160	5.880	1.515	10.885	5.231	1.131	9.000	3.021	0.955	8.041	17.180	1.840	6.365	13.500	1.319	5.872	7.021	0.990

ตารางที่ ข.10 เปรียบเทียบผลการแบ่งส่วนประกอบเชิงปริมาณ โดยเทียบกับจำนวนส่วนประกอบที่  
ประเมินด้วยสายตาสำหรับภาพเอกสารถายขนาดใหญ่ที่มีโครงสร้างสี่เหลี่ยม

Pic No.	Contour and XY Cut Technique			Recursive X-Y Cut Technique		
	Cor.	Error	N	Cor.	Error	N
1	14	5	19/17	8	1	11/17
2	9	2	11/13	7	1	10/13
3	11	0	11/11	5	0	22/11
4	9	4	13/14	8	1	9/14
5	3	4	7/13	4	1	5/13
6	12	1	12/13	9	1	10/13
7	15	0	16/15	6	67	73/15
8	12	13	25/14	3	25	28/14
9	20	2	22/24	3	3	6/24
10	10	6	16/14	8	7	15/14
11	12	4	16/12	5	15	20/12
12	10	2	12/14	1	6	7/14
13	5	1	6/7	2	2	4/7
14	20	9	29/21	14	6	22/21
15	23	3	26/24	16	12	28/24
16	8	0	8/8	7	17	24/24
17	7	4	10/11	9	5	14/11
18	10	11	21/13	9	16	25/13
19	20	9	29/21	13	9	22/21
20	38	0	38/38	32	16	48/38
21	10	62	72/13	3	76	79/13
22	8	4	12/11	10	7	17/11
23	13	0	13/13	10	21	31/13
24	12	8	20/15	5	15	20/15
25	7	7	14/9	7	14	20/9
26	6	5	11/9	2	4	6/9
27	7	6	12/9	5	2	7/11
28	9	0	9/9	5	1	6/9
29	4	1	5/6	3	2	5/6
30	18	1	19/20	18	6	24/20
31	14	2	16/19	17	8	25/19

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.10 (ต่อ)

Pic No.	Contour and XY Cut Technique			Recursive X-Y Cut Technique		
	Cor.	Error	N	Cor.	Error	N
32	13	1	14/15	7	3	10/15
33	7	1	8/9	2	20	22/9
34	7	3	10/10	4	2	6/10
35	6	1	7/8	4	1	5/8
36	8	0	8/8	5	7	12/8
37	9	0	9/9	4	1	5/9
38	8	2	10/9	3	1	4/9
39	6	0	6/6	4	16	20/6
40	6	7	13/9	5	9	14/9
Avg.	11.15	4.775	1.022	7.3	10.675	1.386



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# ECTI-CON 2005

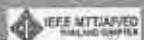
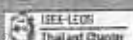
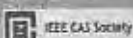
The 2005 ECTI International Conference



Proceedings of The 2005 Electrical Engineering/Electronics, Computer, Telecommunications, and Information Technology (ECTI) International Conference

May 12-13, 2005

Asia Pattaya Beach Hotel, Pattaya, Choburi, THAILAND



เอกสารนี้เป็นเอกสารต้นฉบับที่จัดทำขึ้นเพื่อใช้ในการประชุมวิชาการและเผยแพร่ความรู้แก่สมาชิกและผู้สนใจในสาขาที่เกี่ยวข้อง

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# Fast Document Segmentation Using Contour and X-Y Cut Technique

Boontee Kruatrachue, Narongchai Moongfangklang, Kritawan Siriboon

Department of Computer Engineering, Faculty of Engineering,  
King Mongkut's Institute of Technology Ladkrabang Bangkok, Thailand 10520  
kkboontee@kmitl.ac.th, nanfa\_blue@yahoo.com, kritawan@diamond.ce.kmitl.ac.th

## ABSTRACT

This paper describes fast and efficient method for page segmentation of document containing nonrectangular block. The segmentation is based on edge following algorithm using small window of 16 by 32 pixels. This segmentation is very fast since only border pixels of paragraph are used without scanning the whole page. Still, the segmentation may contain error if the space between them is smaller than the window used in edge following. Consequently, this paper reduce this error by first identify the missed segmentation point using direction information in edge following then, using X-Y cut at the missed segmentation point to separate the connected columns. The advantage of the proposed method is the fast identification of missed segmentation point. This methodology is faster with fewer overheads than other algorithms that need to access much more pixel of a document.

**Keywords:** Contour Direction Technique, Missed Segmentation Points, Page Segmentation, Recursive X-Y Cut Technique

## 1. INTRODUCTION

Page segmentation [1] is the process to identify the areas of interest in the image of a document page. For a conventional document page with material printed in dark ink on a light colored paper, the areas of interest in the (binary) image will be neighborhoods of black pixels. Page segmentation produces a description of the geometrical aspects of the areas of interest. The most common aspects are spatial extent and position on the page. Page segmentation can be thought of as a mapping from the pixel-based image data to a description of the areas of interest.

Several methods have been explored to solve the problem of automated document segmentation. There are three main methods for automatically document segmentation, top-down method [2], [4], [5], bottom-up method [6], [1] and mixed method [3]. A successful method must cope with as many variations (e.g. shapes of regions, skew) in the document as possible. In doing so, it must also not compromise in processing speed. This paper tries to improve the segmentation speed by using contour following technique. This technique is very fast since only border pixels information is used in segmentation. However, the segmentation correctness depends on the window size used in edge following a paragraph. If the window is too small it

will extract letters or words but not paragraph (over segment). On the other hand, if the window is too big it will segment multiple structures (columns) together (under segment). This paper proposed a fast connected point identification to handle the under segment problem. Once connected point is identified, X-Y cut at the point will separate the connected blocks.

## 2. PREVIOUS WORKS

Several methods and their variants have been employed in page segmentation approaches. Some of them include our previous work are briefly reviewed below.

The well known top-down page segmentation technique is the recursive X-Y cut [2]. This approach decomposes a document image by projection profile cuts recursively in to a set of rectangular block. The disadvantage of this method is the rectangular structure of a block.

The hybrid (mixed) approach was presented in [3] using contour following technique. This technique used the square window 32 by 32 pixel followed the rim of paragraph, to specified the boundaries of paragraph in image document. To separate the potential connected blocks, vertical cuts on space between characters are used to separate connected blocks. The advantages of method are the fast computation time, and nonrectangular format block in the contour segmentation. Due to vertical cuts, it can not separate connected columns with nonrectangular format.

## 3. FAST DOCUMENTATION USING CONTOUR AND X-Y CUT TECHNIQUE

Even though many researches have been studied the automated document segmentation, there is no appropriate and complete method for every document. Our approach focus on the columns with narrow gab between them.

The proposed page segmentation consists of three main processes as shown in Fig.1. Firstly, we used the modified contour following technique in clockwise direction [3]. This technique is very fast since only border pixels information is used in segmentation. Then, we identify the missed segmentation point by looking at the turning point of window between columns. The missed segmentation point is the point where the direction of window is counter clockwise.

Finally, we apply X-Y Cut Technique at the missed segment points to separate the connected columns.

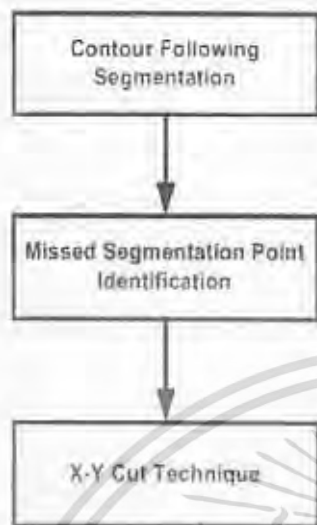


Fig. 1: the overview of document segmentation.

### 3.1 Contour Following Segmentation

The first stage page segmentation is a top-down approach using modified contour following algorithm. The algorithm starts with the searching of the first black pixel. It was used as the starting point of the 16 by 32 window. Then normal edge following algorithm was performed, where the window was considered as located on the black pixels if 10 black pixels were in the window. In spite of its fast execution time, the segmented block may contain multiple columns or paragraphs. In order to reduce the multiple connected columns, the window use an contour following is rectangular of 16 by 32 pixel instead of square 32 by 32 pixel as in our previous paper [3]. The main advantage of contour following segmentation is the shape of block can be in any form (non rectangular) and its execution speed.

### 3.2 Missed segmentation point identification

In order to make use of the fast segmentation in the previous stage, we must be able to also identify the connected column fast. Since the block segmentation is performed by contour following in the clockwise direction. The single column will contain only windows in clockwise direction only. Hence, the potential missed segmentation point is at the window that has counter clockwise direction.

Some window that has counter clockwise direction may not be the connected point as shown below by the dash circled in Fig2. In order to identify the real connected point, the height or depth of the connected point has to be greater than some threshold.

The connection between top and bottom (row) paragraph is allowed. Since a paragraph may not has the justified format (center, align right, align left), it is hard to separate the connected top and bottom paragraph.

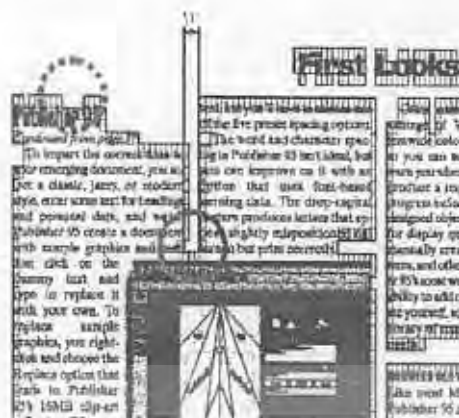


Fig. 2: show the candidate (dash circle) and real (circle) missed segmentation point.

### 3.3 X-Y Cut Technique

Once we find the real missed segmentation point, the next step is to separate the connected structures. There are two types of separation, horizontally and vertically. If the space ( $w$ ) between structures at the connected point is wider than the window width (16 - pixels), the separation starts horizontally, as shown in Fig. 2 (by bold line). On the other hands, if the width is smaller or equal to the width, the separation starts vertically.

Vertical separation can be done by projection cut vertically direction in top or bottom until facing the border window of structure. If the projection hits black pixel before border windows, the projection stops and turns horizontally to both left and right directions. This projection continues until it hits the black or border of window. If it hits black pixel, it changes to perpendicular direction. If it hits the border window, the projections stop. Horizontal separation also performs in the same manner.

## 4. EXPERIMENT AND RESULT

Hundred of documents have been used in segmentation test. Some of the difficult segmentation samples are shown in Fig. 3. The boundaries result of Fig. 3 show in Fig. 4. Since the connected column separation use X-Y cut, the algorithm still has problem with non rectangular shaped of connected structures, as shown in Fig. 5.

## 5. CONCLUSION AND RESULT

Our paper discusses an improved version of the contour following technique using rectangular window, fast missed segmentation identification and correct separation of connected block. This technique is very fast since only border pixels are accessed. We improved the speed of this technique with fast automatically identification of the connected point using window directions. Then apply an X-Y cut

technique to split the connected structure at the connected point. If the connected structures are nonrectangular shape, they still can not be separated due in the X-Y cut technique.



# Capital Equipment

## Tariffs, standards and IP their oh my!



Fig. 3: Missed segmentations point and issue of projection X-Y cut technique

Fig.3: Show the error segmentation result from missed segmentations of the skew object inside paragraph, missed segmentation points are circled



Fig. 4: Show the result region of structure of the original image in Fig. 3. A space between each column is very small due to noise and layout of document. In this figure, the contour direction only is not enough to correct in the part of column

### 6. REFERENCES

- [1] A. Antonopoulos, "Page Segmentation using the Description of the Background Computer Vision and Image Understanding, Vol. 70, (1998) 350-369.
- [2] G. Nagy and S. Seth, "Hierarchical representation of optically scanned documents," *Proc. of IJPR*, (1984) 347-349.
- [3] B. Krutrochug, P. Suthaphan, "A fast and efficient method for document segmentation (or OCR)," *Electrical and Electronic Technology, 2001. Proceeding of IEEE Region 10 International Conference on, Volume 1, 19-22 Aug. (2001) 381-383 vol.1*.
- [4] Jaekyu Ha, R.M. Haralick, I.T. Phillips, "Recursive X-Y Cut Using Bounding Boxes of connected components", *Proceedings of the Third International Conference on Document Analysis and Recognition, Volume 2, 14 - 15 Aug. (1995) 952-954*.
- [5] Jaekyu Ha, R.M. Haralick, I.T. Phillips, "Document Page Decomposition by the Bounding-Box Projection Technique", *Proceedings of the Third International Conference on Document Analysis and Recognition, Volume 2, 14 - 15 Aug. (1995) 1119-1122*.
- [6] T. Saitoh, T. Pavlidis, "Page Segmentation without Rectangle Assumption", *Pattern Recognition Methodology and Systems, Proceedings 11th IAPR International Conference on, 30 Aug-3 Sept. (1992) 277-280*.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ประวัติผู้เขียน

ชื่อ-นามสกุล	นาย ณรงค์ชัย มุ่งแสงกลาง
วันเดือนปีเกิด	วันที่ 25 กันยายน 2517 ที่จังหวัดนครราชสีมา
ที่อยู่	166 หมู่ 15 หมู่บ้าน โนนสะอาด ต.ขามสะแกแสง อ. ขามสะแกแสง จังหวัดนครราชสีมา
ประวัติการศึกษา	2540 จบการศึกษาคณะวิศวกรรมศาสตร์สาขาวิชาวิศวกรรมโทรคมนาคม มหาวิทยาลัยเทคโนโลยีสุรนารี
ประสบการณ์ทำงาน	
พ.ศ. 2544-2545	ผู้ช่วยอาจารย์คณะวิศวกรรมศาสตร์สาขาวิชาวิศวกรรมไฟฟ้า มหาวิทยาลัยเทคโนโลยีสุรนารี
พ.ศ. 2546-ปัจจุบัน	ผู้ช่วยอาจารย์คณะวิศวกรรมศาสตร์สาขาวิชาวิศวกรรมคอมพิวเตอร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ผู้ช่วยอาจารย์คณะวิศวกรรมศาสตร์สาขาวิชาวิศวกรรมคอมพิวเตอร์  
มหาวิทยาลัยรามคำแหง



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้