รายงานวิจัย

เรื่อง

การจำแนกข้อความและภาพถ่าย เพื่อสร้างภาพไบนารีของเอกสาร

หัวหน้าโครงการ : ดร. ยุทธพงษ์ ชังสรรค์เสรี
ผู้ร่วมวิจัย : นายสิทธิชัย บุษหมั่น
นางรพีพร ต่ำชอง

สำนักวิจัยและบริการคอมพิวเตอร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีงบประมาณ 2539

รายงานวิจัย

เรื่อง

การจำแนกข้อความและภาพถ่าย เพื่อสร้างภาพไบนารีของเอกสาร

หัวหน้าโครงการ : ดร. ยุทธพงษ์ รังสรรค์เสรี
ผู้ร่วมวิจัย : นายสิทธิชัย บุษหมั่น
นางรพีพร ช่ำชอง

สำนักวิจัยและบริการคอมพิวเตอร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีงบประมาณ 2539

การจำแนกข้อความและภาพถ่ายเพื่อสร้างภาพไบนารีของเอกสาร

(Text/Photograph Segmentation for Document Image Binarization)

บทคัดย่อ

ข้อมูลที่ปรากฏบนเอกสารใดๆ อาจจำแนกได้เป็นสองส่วนที่สำคัญคือ ส่วนที่เป็นข้อความตัวอักษร ซึ่งมักมีสีเข้มบนพื้นกระดาษสีขาวหรือสีที่อ่อนกว่า และส่วนที่เป็นภาพถ่ายซึ่งมีสีที่แตกต่างกันหลายระดับ การสร้างภาพหน้าเอกสารแบบไบนารีจำเป็นต้องมีความสามารถในการจำแนกหน้าเอกสารทั้งสองส่วนได้โดยอัตโนมัติ โดยที่ในแต่ละส่วนจะใช้วิธีในการสร้างภาพไบนารีที่แตกต่างกัน การสร้างภาพหน้าเอกสารแบบไบนารีที่นำเสนอนี้ เริ่มจากข้อมูลภาพขนาด 256 ระดับเทาที่ได้จากการสแกนเอกสารใดๆ ด้วยสแกนเนอร์ขาวดำที่มีรายละเอียด 300 จุดต่อนิ้ว ทำการจำแนกส่วนของข้อความและภาพถ่าย แล้วนำส่วนของข้อความไปสร้างภาพแบบไบนารีโดยเทคนิคเธรชโชลด์และส่วนของภาพถ่ายนำไปสร้างภาพแบบไบนารีโดยเทคนิคฮาฟท์โทน จากนั้นนำทั้งสองส่วนมารวมกันได้เป็นภาพหน้าเอกสารแบบไบนารี วิธีนี้สามารถสร้างภาพไบนารีที่มีคุณภาพสูงสำหรับเอกสารใดๆ ที่อาจประกอบไปด้วยข้อความและภาพถ่ายปะปนอยู่ในหน้าเดียวกัน

ABSTRACT

The content of any documents may contain two types of regions : text regions representing dark objects located on a clear background, and photograph regions representing continuous-tone areas. A critical function of document processing systems is the automatic segmentation of digitized documents into regions of text and photographs. Each of region is produce binary images by another method. This thesis will present an algorithm for Mixed Document Binariztion. Documents are scanned as 256 gray-level images with commercial scanner at 300 dots per inch. It separates the document into two types of regions, which we call text regions and photograph regions. Each type of those regions will be binarized by appropriate technique : thresholding for text regions, and halftoning for photograph regions. The final binary image is then obtained by merging the results from those two processes. Hence this technique can produce high-quality binary representation of any documents consisting of text and photograph in same page.

สารบัญ

หน้า

บทคัดย่อ ..... ก

สารบัญ ..... ข

สารบัญภาพ ..... ค

บทที่

1. บทนำ 1
   - ความเป็นมาและความสำคัญของงานวิจัย 1
   - วัตถุประสงค์ของการวิจัย 3
2. ทฤษฎีเบื้องต้น 4
   - การจำแนกภาพ 4
   - การสร้างภาพไบนารี 12
3. การจำแนกข้อความและภาพถ่ายอัตโนมัติ 16
   - บทนำ 16
   - ขั้นตอนการทำงานของกระบวนการจำแนกส่วนของข้อความและภาพถ่าย 18
   - การทดลองการจำแนกข้อความและภาพถ่าย 23
   - บทสรุป 28
4. การสร้างภาพไบนารีสำหรับภาพเอกสาร 29
   - ขั้นตอนการทำงาน. 30
   - การทดลองการสร้างภาพไบนารี 35
   - บทสรุป 42
5. สรุปผลและข้อเสนอแนะ 43
   - สรุปผลการวิจัย 43
   - ปัญหาที่พบและข้อเสนอแนะ 44

บรรณานุกรม 45

สารบัญภาพ

หน้า

1. แสดงขั้นตอนในการจำแนกข้อความและภาพถ่ายเพื่อสร้างภาพไบนารี 2
2. (ก) และ (ข) แสดงวินโดว์สำหรับการคำนวณหาค่ากราเดียนท์ 6
3. (ก) และ (ข) แสดงวินโดว์ขนาด 2x2 สำหรับการคำนวณหาค่ากราเดียนท์ 7
4. (ก) และ (ข) แสดงวินโดว์ขนาด 2x2 สำหรับการคำนวณหาค่ากราเดียนท์ โดยวิธีของ Robert 8
5. (ก) และ (ข) แสดงวินโดว์ขนาด 3x3 สำหรับการคำนวณหาค่ากราเดียนท์ โดยวิธีของ Sobel 8
6. (ก) และ (ข) แสดงวินโดว์ขนาด 3x3 สำหรับการคำนวณหาค่ากราเดียนท์ โดยวิธีของ Prewitt 9
7. แสดงตัวดำเนินการหาขอบภาพ โดยวิธีของ Laplacian 10
8. แสดงหลักการสร้างภาพไบนารี 12
9. (ก) (ข) (ค) และ(ง) แสดงการสร้างภาพไบนารี โดยเทคนิคเธรชโชลด์ 13
10. (ก) (ข) (ค) และ(ง) แสดงการสร้างภาพไบนารี โดยเทคนิคฮาฟท์โทน 15
11. แสดงภาพหน้าเอกสารที่ประกอบไปด้วยส่วนของข้อความและภาพถ่าย 17
12. แสดงขั้นตอนของการจำแนกข้อความและภาพถ่ายอัตโนมัติ 12
13. แสดงการแบ่งข้อมูลภาพออกเป็นบล็อกขนาด 4x4 จุดภาพ 19
14. แสดงภาพหน้าเอกสารต้นฉบับที่ใช้ในการจำแนกส่วนของข้อความและภาพถ่าย ... 24
15. แสดงผลลัพธ์ที่ได้จากการพิจารณาว่าบล็อกใดเป็นบล็อกข้อความหรือ บล็อกใดเป็นบล็อกของภาพถ่าย 25
16. แสดงผลลัพธ์ที่ได้จากการคำนวณหาพื้นที่ภาพถ่าย โดยการพิจารณาจากจำนวนบล็อกของภาพถ่าย 26
17. แสดงพื้นที่ภาพถ่ายที่ได้จากการจำแนกข้อความและภาพถ่าย 27
18. แสดงขั้นตอนการทำงานของระบบการจำแนกข้อความและภาพถ่าย เพื่อสร้างภาพไบนารี 30
19. แสดงหลักการทำงานของวิธีการกระจายความผิดพลาด 33
20. แสดงรูปแบบตัวกรองความผิดพลาด โดยวิธีของ Floyd and Steinberg 34

สารบัญภาพ(ต่อ)

หน้า

21. แสดงแนวความคิดในการคำนวณหาค่าเธรชโชลด์
โดยการพิจารณาจากฮิสโตแกรม 34
22. แสดงภาพหน้าเอกสารต้นฉบับที่ใช้ในการสร้างภาพไบนารี
สำหรับภาพเอกสาร 36
23. แสดงภาพหน้าเอกสารแบบไบนารีที่ได้จากวิธีการที่นำเสนอ 37
24. แสดงภาพหน้าเอกสารต้นฉบับที่มีระดับความเข้ม 256 ระดับเทา 38
25. แสดงภาพหน้าเอกสารแบบไบนารีที่ได้จากวิธีการที่นำเสนอ 39
26. แสดงภาพหน้าเอกสารแบบไบนารีที่ได้จากเทคนิคเธรชโชลด์ 40
27. แสดงภาพหน้าเอกสารแบบไบนารีที่ได้จากเทคนิคเฮาฟท์โทน 41

บทที่ 1

บทนำ

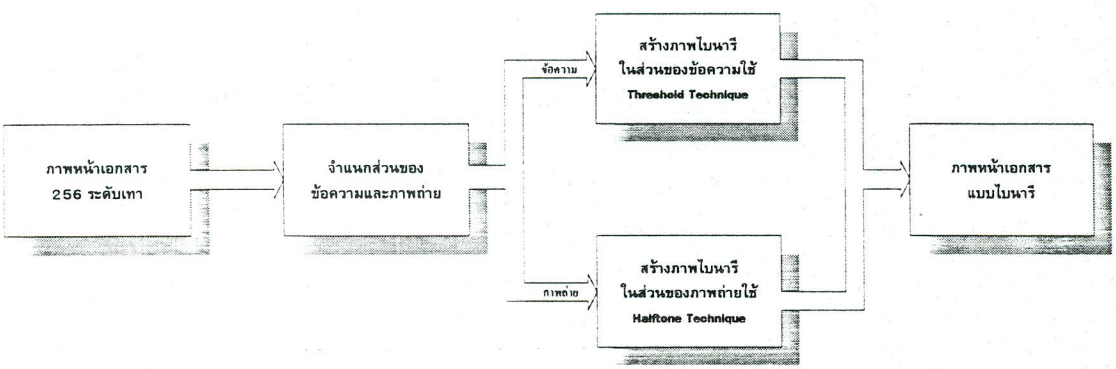
ความเป็นมาและความสำคัญของงานวิจัย

เนื่องจากในปัจจุบันเอกสารเป็นสิ่งสำคัญและจำเป็นในการใช้ติดต่อสื่อสาร ไม่ว่าจะเป็นภาคเอกชนทางด้านธุรกิจหรือทางด้านรัฐบาล ต่างก็มีปัญหาในการจัดเก็บเอกสาร และในปัจจุบันก็จะยิ่งเพิ่มมากขึ้นเรื่อยๆ สิ่งเหล่านี้เป็นปัญหาในการจัดเก็บเอกสารที่มีมากเกินไป ปัจจุบันได้มีการพยายามที่จะทำให้ที่ทำงานหรือสำนักงานเป็นสถานที่ที่มีกระดาษน้อยที่สุด (Paperless Office) โดยการพัฒนาระบบคอมพิวเตอร์ให้สามารถจัดเก็บเอกสารอย่างมีระบบและง่ายต่อการค้นหา เช่นระบบจัดเก็บเอกสารอิเล็กทรอนิกส์ (Automatic Filling System) ซึ่งเป็นระบบที่ต้องจัดเก็บเอกสารลงดิสก์ โดยที่จะต้องสามารถเรียกกลับมาพิมพ์ลงบนกระดาษได้ในคุณภาพใกล้เคียงกับต้นฉบับ การเก็บข้อมูลภาพหน้าเอกสารในลักษณะของภาพสีเทา 256 ระดับ สามารถให้ผลเป็นที่น่าพอใจแต่ปัญหาที่เกิดขึ้นก็คือจะกินเนื้อที่ในดิสก์มาก การสืบค้นหรือการเรียกกลับมาพิมพ์จะเสียเวลามาก ดังนั้นการแก้ปัญหาก็สามารถทำได้โดยต้องพยายามแปลงข้อมูลภาพนั้นให้เป็นภาพแบบไบนารี (Binary Image) ซึ่งเป็นภาพที่มีความเข้มเพียง 2 ระดับคือ ขาวกับดำ เพื่อที่จะใช้เนื้อที่ในการเก็บน้อยลง (สามารถลดลงได้ถึง 8 เท่า) หรืออีกตัวอย่างคือ ระบบการวิเคราะห์เอกสาร (Document Analysis System) ซึ่งเป็นระบบที่มีจุดมุ่งหมายเพื่อแปลงสิ่งที่ปรากฏบนหน้าเอกสารที่มนุษย์สามารถเข้าใจความหมายและสามารถมองเห็นไปสู่สัญลักษณ์ที่เครื่องคอมพิวเตอร์สามารถประมวลผลได้ ในระบบนี้มีขั้นตอนที่สำคัญขั้นตอนหนึ่งคือ ขั้นตอนการทำไบนาไรเซชั่น (Image Binarization) หรือขั้นตอนการสร้างภาพไบนารีจากภาพที่มีหลายระดับเทา เพื่อที่จะได้นำภาพหน้าเอกสารแบบไบนารีที่ได้ไปทำการประมวลผลในขั้นตอนต่อไป

ปัจจุบันได้มีเทคนิคที่สำคัญในการสร้างภาพไบนารี 2 เทคนิคหลักๆคือ เทคนิคเธรชโชลด์ (Thresholding Technique) วิธีนี้จะทำการคำนวณค่าตัวแบ่งขึ้นมาเพื่อใช้แยกว่าจุดภาพแต่ละจุดนั้นควรจะเป็นจุดขาวหรือจุดดำ แต่วิธีนี้จะเหมาะกับข้อมูลที่เป็นข้อความตัวหนังสือเท่านั้น สำหรับข้อมูลที่เป็นภาพถ่ายนั้นวิธีที่เหมาะสมคือ เทคนิคฮาฟท์โทน (Halftoning Technique) เทคนิคนี้จะใช้การกระจายจุดขาวกับดำตามระดับความเข้มในพื้นที่ส่วนเล็กๆ จะเห็นได้ว่าเราจำเป็นจะต้องเลือกใช้วิธีการที่เหมาะสมสำหรับการสร้างภาพไบนารีคือ เทคนิคเธรชโชลด์สำหรับข้อความ และเทคนิคฮาฟท์โทนสำหรับภาพถ่าย เมื่อเป็นเช่นนี้แล้วในการสร้าง

ภาพหน้าเอกสารแบบไบนารีให้มีประสิทธิภาพนั้น จำเป็นที่จะต้องทำให้เครื่องคอมพิวเตอร์เข้าใจถึงรายละเอียดภายในเอกสารนั้นก่อน โดยทั่วไปแล้วในหน้าเอกสารหนึ่งๆประกอบไปด้วยสองส่วนใหญ่ๆคือ ส่วนที่เป็นข้อความ มักมีสีเข้มบนพื้นกระดาษสีขาวหรือสีที่อ่อนกว่า และอีกส่วนหนึ่งก็คือภาพถ่าย มักมีสีที่แตกต่างกันหลายระดับ ซึ่งปัญหาก็คือว่าทำอย่างไรจึงจะให้คอมพิวเตอร์สามารถจำแนกได้ว่าส่วนใดในหน้าเอกสารเป็นข้อความและส่วนใดในหน้าเอกสารเป็นภาพถ่าย เพื่อที่จะได้นำแต่ละส่วนที่ได้จากการจำแนกมาใช้ในการสร้างภาพหน้าเอกสารแบบไบนารี อีกทั้งปัญหาในการจำแนกภาพหน้าเอกสารเป็นปัญหาที่น่าสนใจ และมีประโยชน์มากในหลายๆด้านไม่เฉพาะการสร้างภาพแบบไบนารี แต่อาจจะนำเฉพาะส่วนของข้อความไปทำการรู้จำตัวอักษรก็ได้

ดังนั้น งานวิจัยนี้จึงได้ทำการศึกษาวิธีการในการจำแนกข้อความและภาพถ่ายเพื่อสร้างภาพไบนารี โดยการทำงานจะเริ่มจากการจำแนก (Segmentation) ภาพหน้าเอกสารให้ได้ว่าส่วนใดในหน้าเอกสารเป็นข้อความ และส่วนใดในหน้าเอกสารเป็นภาพถ่าย จากนั้นในส่วนของข้อความจะนำเทคนิคเธรชโชลด์มาประยุกต์ใช้ในการสร้างภาพไบนารี และส่วนของภาพถ่ายนำเทคนิคฮาฟท์โทนนิงมาประยุกต์ใช้ในการสร้างภาพไบนารี ผลลัพธ์ที่ได้คือภาพหน้าเอกสารแบบไบนารีที่มีความถูกต้องสวยงามเป็นที่น่าพอใจและมีคุณภาพใกล้เคียงกับภาพหน้าเอกสารต้นฉบับ



ภาพที่ 1 แสดงขั้นตอนในการจำแนกข้อความและภาพถ่ายเพื่อสร้างภาพไบนารี

วัตถุประสงค์ของงานวิจัย

1. เพื่อเป็นการศึกษาหาแนวทางในการพัฒนาและสร้างภาพแบบไบนารี โดยได้มุ่งเน้นศึกษาวิธีการในการจำแนกข้อความและภาพถ่ายเพื่อสร้างภาพไบนารี สำหรับภาพหน้าเอกสารที่ประกอบไปด้วยข้อความและภาพถ่าย

2. เพื่อเป็นการศึกษาหาแนวทางในการจำแนกภาพหน้าเอกสาร โดยมุ่งเน้นศึกษาในเรื่องของการจำแนกส่วนที่เป็นข้อความและส่วนที่เป็นภาพถ่าย

3. สามารถลดขนาดของเนื้อที่ในการเก็บรักษาข้อมูลภาพหน้าเอกสาร โดยที่รายละเอียดของภาพหน้าเอกสารมีคุณภาพใกล้เคียงกับภาพต้นฉบับ

4. เพื่อที่จะนำไปประยุกต์ใช้กับระบบประยุกต์อื่นๆได้ เช่น ระบบจัดเก็บเอกสารอัตโนมัติ ในแง่ของการลดเนื้อที่ในการจัดเก็บภาพหน้าเอกสาร ระบบรู้จำตัวอักษรอันเป็นผลเนื่องมาจากการจำแนกส่วนของข้อความและภาพถ่าย โดยการนำส่วนที่เป็นข้อความมาใช้ในการประมวลผลเป็นต้น

บทที่ 2

ทฤษฎีเบื้องต้น

การจำแนกภาพ (Image Segmentation)

การจำแนกภาพ หมายถึงการแยกองค์ประกอบหรือส่วนต่างๆที่อยู่ในภาพออกไปเป็นส่วนประกอบย่อยๆ ที่สัมพันธ์กับลักษณะทางกายภาพของรูปนั้นออกมาให้เด่นชัดขึ้น โดยที่จะทำการแยกส่วนพื้นที่ต่างๆ ในภาพหรือขอบรอยต่อของภาพ (edge) ออกเป็นส่วนๆ ส่วนประกอบที่ถูกแยกออกมานั้นจะถูกนำไปตีความและทำการรู้จำในขั้นตอนต่อไป ในการจำแนกภาพจะถือว่าเนื้อของวัตถุที่อยู่ในบริเวณเดียวกันหรือใกล้เคียงกันจะมีลักษณะของจุดภาพที่มีคุณสมบัติพื้นฐานเหมือนกันหรือมีความเป็นเอกพันธ์ (Homogeneous) ตัวอย่างเช่น สีและความสว่างของภาพที่อยู่ในบริเวณเดียวกัน โดยที่ระดับความเข้มของแต่ละจุดภาพจะมีค่าใกล้เคียงกัน นอกจากนี้แล้วจะเกิดการเปลี่ยนระดับความเข้มทันทีทันใดในบริเวณที่เป็นขอบรอยต่อของวัตถุ อย่างไรก็ตามสมมุติฐานนี้ก็ไม่สามารถใช้ได้กับทุกกรณี ในบางครั้งขอบภาพทางกายภาพระหว่างวัตถุที่คล้ายกันก็จะไม่ปรากฏเป็นขอบของภาพได้อย่างชัดเจน

โดยทั่วไปแล้ว การจำแนกภาพได้ถูกนำมาใช้งานอย่างกว้างขวางในฐานะที่เป็นตัวเลือกวัตถุมาทำการวิเคราะห์ภาพ (Image Analysis) ในขั้นตอนต่อไป การจำแนกภาพทำงานในวิธีเดียวกันกับที่ตาของคนแยกลักษณะเด่นออกมาจากภาพที่มองเห็น นอกจากนี้การจำแนกยังทำการแบ่งภาพออกเป็นส่วนๆที่มีความหมายทางกายภาพและแยกพื้นที่ๆ เป็นเอกพันธ์สำหรับลักษณะบางอย่างอีกด้วย ถึงแม้ว่าจะมีลักษณะหลายอย่างที่สามารถนำมาใช้ในการตัดสินใจความเป็นเอก-พันธ์ในพื้นที่ก็ตาม ในการจำแนกภาพส่วนใหญ่จะสนใจเพียงลักษณะสำคัญ 2 อย่างคือ ความคล้ายกัน (Similarity) กล่าวคือมีความคล้ายคลึงกันภายในบริเวณที่ใกล้เคียงกัน และ ความแตกต่างกัน (Discontinuity) จะมีความแตกต่างกันขององค์ประกอบที่อยู่ในบริเวณเดียวกัน

วิธีการในการจำแนกภาพ สามารถแบ่งเป็นหลักๆ ได้ 2 วิธีคือ วิธีอาศัยขอบเป็นหลัก และวิธีอาศัยพื้นที่เป็นหลัก ซึ่งหลักการของการจำแนกภาพทั้ง 2 แบบอธิบายได้ดังนี้

1. การจำแนกภาพโดยอาศัยขอบรอยต่อของภาพเป็นหลัก

วิธีนี้จะใช้หลักการของการตรวจหาขอบรอยต่อของภาพเป็นหลักในการจำแนก ในการหาขอบรอยต่อของภาพจะสันนิษฐานว่าที่ขอบรอยต่อของภาพจะมีการเปลี่ยนแปลงความเข้มจนทำให้เกิดความเข้มไม่ต่อเนื่องเป็นลักษณะของ step function ในภาพ 2 มิติขอบรอยต่อของภาพจะมีทิศทางเช่นเดียวกับภาพและความเข้มในแนวขอบรอยต่อจะเป็นยูนิฟอร์ม (Uniform) วิธีการหาขอบรอยต่อของภาพสามารถหาได้หลายวิธีเช่น

1.1. การหาขอบภาพโดยใช้อนุพันธ์เชิงเส้นอันดับที่หนึ่ง การหาขอบภาพเป็นขั้นตอนที่สำคัญในการสืบหาบริเวณที่มีการเปลี่ยนแปลงระดับความเข้มของจุดภาพเกิดขึ้น การหาขอบภาพสามารถทำได้โดยการหาค่ากราเดียนท์ (Gradient) บนข้อมูลภาพ ซึ่งกราเดียนท์เป็นฟังก์ชั่นที่วัดอัตราการเปลี่ยนแปลงระดับความเข้มที่เกิดขึ้นของจุดภาพ ซึ่งอยู่บนพื้นฐานการหาอนุพันธ์เชิงเส้นอันดับที่หนึ่ง และกราเดียนท์ $\nabla f(x,y)$ ที่เป็นฟังก์ชั่นต่อเนื่องในระนาบ 2 มิติสามารถกำหนดอยู่ในรูปของเวคเตอร์ได้ดังสมการต่อไปนี้

$$\nabla f(x,y)=\begin{bmatrix} G_x \\ G_y \end{bmatrix}=\begin{bmatrix} \partial f/\partial x \\ \partial f/\partial y \end{bmatrix} \qquad (2.1)$$

ขนาดของกราเดียนท์ (Magnitude of Gradient) จะเท่ากับ อัตราการเปลี่ยนแปลงระดับความเข้มของจุดภาพในทิศทาง G ที่กำหนด (ทิศทาง G ในที่นี้ คือ ทิศทางในแนวนอน และแนวตั้ง) ซึ่งขนาดของกราเดียนท์สามารถหาได้จากสมการดังต่อไปนี้

$$\nabla f(x,y)=\sqrt{G_x^2+G_y^2} \qquad (2.2)$$

ขนาดกราเดียนท์โดยประมาณ (Approximate Magnitude of Gradient) สามารถหาได้จากสมการค่าสัมบูรณ์ดังต่อไปนี้

$$\nabla f(x,y)\approx |G_x|+|G_y| \qquad (2.3)$$

ทิศทางของกราเดียนท์จะเท่ากับ

$$\alpha(x,y) = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad \text{.....(2.4)}$$

สำหรับภาพเชิงตัวเลขการหาค่ากราเดียนท์ของฟังก์ชั่นสองมิติ $g(x,y)$ ซึ่งเป็นฟังก์ชั่นไม่ต่อเนื่องที่พิกัด (x,y) สามารถกระทำในรูปของสมการดังต่อไปนี้

$$G_x = g(x+1,y) - g(x,y) \quad \text{.....(2.5)}$$

$$G_y = g(x,y) - g(x,y+1) \quad \text{.....(2.6)}$$

จากสมการที่ (2.5) และ (2.6) สามารถพิจารณาในรูปของการนำวินโดว์ขนาด 1x2 และ 2x1 ที่แสดงดังภาพที่ 2 มาคอนโวลูชันหรือการประสานงาน (Convolution) กับข้อมูลภาพเชิงตัวเลขเพื่อใช้สำหรับคำนวณหาค่า G_x และ G_y ได้ตามลำดับ

ภาพที่ 2 แสดงวินโดว์สำหรับการคำนวณหาค่ากราเดียนท์
(ก) วินโดว์ขนาด 1x2 ที่ใช้สำหรับหาค่า G_x
(ข) วินโดว์ขนาด 2x1 ที่ใช้สำหรับหาค่า G_y

จากขั้นตอนการนำวินโดว์ขนาด 1x2 และ 2x1 ข้างต้นมาคอนโวลูชันกับข้อมูลภาพเชิงตัวเลขเพื่อหาค่ากราเดียนท์ในทิศทาง x และ y ตามลำดับนั้น พบว่าค่า G_x ที่คำนวณได้นั้นเปรียบเสมือนค่าที่เกิดจากการอินโทโปรเลตหรือการทำนายเชิงเส้น (interpolate) ที่พิกัดจุดภาพ $(x+\frac{1}{2},y)$ โดยอ้างอิงระดับความเข้มของจุดภาพข้างเคียงที่มีพิกัดสอดคล้องกับพิกัดบนวินโดว์ ส่วนค่า G_y ที่คำนวณได้นั้นเปรียบเสมือนค่าที่เกิดจากการอินโทโปเลตที่พิกัดจุดภาพ $(x,y+\frac{1}{2})$ โดยอ้างอิงระดับความเข้มของจุดภาพข้างเคียงที่มีพิกัดสอดคล้องกับพิกัดบนวินโดว์เช่นกัน นอกจากนี้ยังสามารถที่นำวินโดว์ขนาด 2x2 ดังภาพที่ 3 (ก) และ (ข) มาทำคอนโวลูชั่นกับภาพเชิงตัวเลขเพื่อหาค่า G_x และ G_y ได้ตามลำดับ และขนาดของกราเดียนท์ที่เกิด

ขึ้นจะสูงกว่าการคอนโวลูชั่นด้วยวินโดว์ขนาด 1x2 และ 2x1 และจากการคอนโวลูชั่นข้อมูลภาพด้วยวินโดว์ขนาด 2x2 พบว่าค่ากราเดียนท์ที่คำนวณได้ทั้งในทิศทาง x และ ทิศทาง y เปรียบเสมือนค่าที่เกิดจากการอินโทโปรเลตที่พิกัดจุดภาพ $g(x+\frac{1}{2},y+\frac{1}{2})$ โดยอ้างอิงระดับความเข้มของจุดภาพรอบข้างที่มีพิกัดสอดคล้องกับพิกัดบนวินโดว์ หรืออาจจะใช้วินโดว์ขนาด 3 x3 มาคำนวณหากราเดียนท์ของจุดภาพที่พิกัดศูนย์กลางของวินโดว์ก็ได้ ในการหาขนาดของกราเดียนท์มีหลายวิธีเช่น

-1	1
-1	1

(ก)

1	1
-1	-1

(ข)

ภาพที่ 3 แสดงวินโดว์ขนาด 2x2 ที่ใช้สำหรับหาค่ากราเดียนท์
(ก) สำหรับหาค่า G_x (ข) สำหรับหาค่า G_y

1.1.1 การหาขอบภาพโดยวิธีของ Roberts การหาขนาดของกราเดียนท์แบบไม่ต่อเนื่อง (Magnitude of Gradient) โดยวิธีของ Roberts สามารถคำนวณจากสมการดังต่อไปนี้

$$\nabla g(x,y) = |g(x,y) - g(x+1,y+1)| + |g(x,y+1) - g(x+1,y)| \quad \ldots\ldots\ldots\ldots\ldots \quad (2.7)$$

จากสมการการที่ (2.7) สามารถพิจารณาในรูปของการคอนโวลูชั่นข้อมูลภาพเชิงตัวเลขด้วยวินโดว์ โดยวิธีของ Robert จะใช้วินโดว์ขนาด 2x2 ดังภาพที่ 4 (ก) และ (ข) สำหรับการหากราเดียนท์ของภาพในทิศทาง x และ y ตามลำดับ

1	0
0	-1

(ก)

0	-1
1	0

(ข)

ภาพที่ 4 แสดงวินโดว์ขนาด 2x2 ที่ใช้สำหรับหาค่ากราเดียนท์โดยวิธีของ Robert
(ก) สำหรับหาค่า G_x (ข) สำหรับหาค่า G_y

1.1.2 การหาขอบภาพโดยวิธีของ Sobel สำหรับการหาขนาดของกราเดียนท์แบบไม่ต่อเนื่องโดยวิธีของ Sobel สามารถคำนวณจากสมการดังต่อไปนี้

$$\nabla g(x,y)=\begin{pmatrix}|(g(x+1,y-1)+c*g(x+1,y)+g(x+1,y+1))-(g(x-1,y-1)+c*g(x-1,y)-g(x-1,y+1))|-\\(g(x-1,y+1)+c*g(x,y+1)+g(x+1,y+1))-((g(x-1,y-1)+c*g(x,y-1)-g(x+1,y-1))|\end{pmatrix}..(2.8)$$

โดยค่าคงที่ C เท่ากับ 2

จากสมการการหาขนาดกราเดียนท์ข้างต้น สามารถพิจารณาในรูปของการคอนโวลูชั่นข้อมูลภาพเชิงตัวเลขด้วยวินโดว์ โดยวิธีของ Sobel จะใช้วินโดว์ขนาด 3x3 ดังภาพที่ 5 (ก) และ (ข) สำหรับการหากราเดียนท์ของภาพในทิศทาง x และ y ตามลำดับ

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

(ก)

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

(ข)

ภาพที่ 5 วินโดว์ขนาด 3x3 ที่ใช้สำหรับหาค่ากราเดียนท์โดยวิธีของ Sobel
(ก) สำหรับหาค่า G_x (ข) สำหรับหาค่า G_y

1.1.3 การหาขอบภาพโดยวิธีของ Prewitt สมการที่ใช้ในการคำนวณหาขนาดกราเดียนท์โดยวิธีของ Prewitt จะเป็นสมการเดียวกับสมการที่ใช้หาขนาดของกราเดียนท์โดยวิธีของ Sobel แต่ต่างกันตรงที่สมการของ Prewitt ค่าคงที่ C เท่ากับ 1 ดังนั้น วินโดว์ที่ใช้ในการหากราเดียท์ของภาพ ทั้งในทิศทาง x และ y แสดงดังภาพที่ 6

-1	0	1
-1	0	1
-1	0	1

(ก)

1	1	1
0	0	0
-1	-1	-1

(ข)

ภาพที่ 6 แสดงวินโดว์ขนาด 3x3 ที่ใช้สำหรับหาค่ากราเดียนท์โดยวิธีของ Prewitt
(ก) สำหรับหาค่า G_x (ข) สำหรับหาค่า G_y

จากการหาขอบภาพโดยวิธีต่าง ๆ ที่กล่าวข้างต้น พบว่าวินโดว์ที่นำมาใช้คอนโวลูชั่นกับข้อมูลภาพในแต่ละวิธีจะมีค่าถ่วงน้ำหนักต่างไปจากการหากราเดียนท์ของสมการเริ่มต้น คือ สมการที่ (2.7) และ (2.8) ซึ่งมีค่าถ่วงน้ำหนักเพียงสองระดับเท่านั้น (คือ 1,-1) ส่วนการหาขอบภาพโดยวิธีของ Roberts และ Prewitt จะใช้ค่าถ่วงน้ำหนักถึง 3 ระดับ (คือ -1,0,1) และวิธีของ Sobel จะใช้ค่าถ่วงนำหนักถึง 5 ระดับ (คือ -2,-1,0,1,2)

1.2 การหาขอบภาพโดยใช้อนุพันธ์เชิงเส้นอันดับที่สอง จากวิธีการหาขอบโดยใช้อนุพันธ์อันดับที่หนึ่งที่ได้กล่าวมามักมีปัญหาเมื่อใช้กับลักษณะความเข้มลาดเอียงที่ไม่เป็น Step ซึ่งจะพบเมื่อวัตถุลาดเอียงถูกส่องด้วยแสงจากแหล่งเดียว ปัญหานี้จะหมดไปเมื่อใช้วิธีการหาขอบภาพโดยใช้อนุพันธ์เชิงเส้นอันดับที่สอง การหาขอบภาพโดยวิธีนี้อาศัยหลักการที่ว่า ที่ตำแหน่งของจุดภาพใดมีการเปลี่ยนแปลงระดับความเข้มอย่างทันทีทันใด และเมื่อหาอนุพันธ์เชิงเส้นอันดับที่สองที่ตำแหน่งจุดภาพนั้นจะพบว่ามีซีโร่ครอสซิ่ง (Zero Crossing) เกิดขึ้น สำหรับขอบที่มีลักษณะกระโดดแล้วอนุพันธ์อันดับที่สองที่จุดขอบนั้นจะเป็นศูนย์แต่จะมีค่าบวกและค่าลบอยู่คนละด้านของขอบ จากนั้นการหาขอบทำได้จากการทำซีโร่ครอสซิ่ง สำหรับขั้นตอนการหาขอบภาพที่อยู่บนพื้นฐานของหลักการดังกล่าวมีอยู่หลายวิธี แต่วิธีหนึ่งที่นิยมใช้กันมากก็คือ ตัวดำเนินการหาขอบภาพ Laplacian (Laplacian Operator) โดยที่ตัวดำเนินการหาขอบภาพ Laplacian ของฟังก์ชั่นสองมิติ $g(x,y)$ ซึ่งเป็นฟังก์ชั่นไม่ต่อเนื่องที่พิกัด (x,y) จะได้มาจากการประมาณอนุพันธ์เชิงเส้นอันดับสองของฟังก์ชั่นต่อเนื่องดังสมการที่ (2.9)

$$\nabla^2 f(x,y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad \text{.....(2.9)}$$

สำหรับในกรณีของ $\nabla^2 g(x,y)$ ที่ได้จากการประมาณสมการที่ (2.10) เพื่อใช้ในการหาขอบภายในภาพเชิงตัวเลข สามารถกำหนดได้ดังนี้

$$\nabla^2 g(x,y) = 4g(x,y) - ((g(x+1,y) + g(x-1,y) + g(x,y+1) + g(x,y-1)) \quad \text{.....(2.10)}$$

จากสมการการที่ (2.10) สามารถพิจารณาในรูปของการคอนโวลูชั่นข้อมูลภาพเชิงตัวเลขด้วยวินโดว์ขนาด 3x3 โดยสัมประสิทธิ์ในวินโดว์ที่ใช้คูณอยู่ที่พิกัดที่ต้องการหาขอบภาพจะต้องมีค่าเป็นบวกเสมอ ส่วนสัมประสิทธิ์ในตำแหน่งอื่นๆ อาจมีค่าเป็นบวกหรือลบก็ได้ เนื่องจากอนุพันธ์ของ Laplacian ผลรวมของสัมประสิทธิ์ทั้งหมดจะต้องเป็นศูนย์เสมอ ซึ่งในกรณีนี้ถ้าจุดภาพที่อยู่ในพื้นที่ขนาด 3x3 มีค่าเท่ากันหมด เมื่อนำมาคอนโวลูชั่นกับตัวดำเนินการนี้ผลลัพธ์ที่ได้ก็คือ 0 ซึ่งแสดงว่าจุดนี้ไม่ใช่ขอบภาพอย่างแน่นอน ในภาพที่ 7 แสดงวินโดว์ของ Laplacian ในรูปแบบต่าง ๆ ทั้งที่ใช้จุดภาพในบริเวณใกลเคียง 4 จุด หรือ 8 จุด ดังรูป ซึ่งจะมีผลของขอบภาพต่างกัน

0	-1	0
-1	4	-1
0	-1	0

(ก)

-1	-1	-1
-1	8	-1
-1	-1	-1

(ข)

1	-2	1
-2	4	-2
1	-2	1

(ค)

ภาพที่ 7 แสดงตัวดำเนินการหาขอบภาพโดยวิธีของ Laplacian

2. การจำแนกภาพโดยอาศัยพื้นที่เป็นหลัก

วิธีนี้คล้ายกับวิธีการแบ่งพื้นที่ หลักการนี้จะหาเซตของจุดภาพที่เชื่อมต่อกัน ซึ่งข้อมูลในเซตนี้มีลักษณะพื้นฐานบางอย่างเหมือนกัน โดยที่ภายในบริเวณเดียวกันจะต่อถึงกันและมีค่าระดับสีเทาใกล้เคียงกัน จึงอาจสันนิษฐานว่ามีส่วนสัมพันธ์กับวัตถุ วิธีการนี้จะเริ่มต้นด้วยการเลือกจุดภาพขึ้นมาก่อนหนึ่งจุด ต่อจากนั้นก็จะทำการพิจารณาจุดภาพที่อยู่ใกล้เคียงเพื่อที่จะรวมเป็นส่วนเดียวกัน จุดภาพที่อยู่ในบริเวณใกล้เคียงทุกจุดที่อยู่ภายใต้เกณฑ์การรวมจะกลาย

เป็นส่วนของภาพที่เกิดขึ้นในภาพนั้น แต่ถ้าจุดภาพที่อยู่ใกล้เคียงนั้นตรวจสอบแล้วไม่อยู่ในเกณฑ์การรวม จุดภาพนั้นจะไม่ถูกรวมเข้าไปในส่วนนั้นของภาพแต่จะถูกเลือกให้เป็นจุดเริ่มต้นส่วนอื่นๆต่อไป และหลังจากที่จุดภาพทุกจุดได้รวมตัวกันเป็นกลุ่มเรียบร้อยแล้ว การที่จะทำให้ได้ส่วนของภาพสมบูรณ์ขึ้น ก็สามารถที่จะเริ่มใหม่โดยการแยกแต่ละส่วนออกแล้วใช้หลักการรวมกันของจุดภาพที่กล่าวมาแล้วอีกครั้งหนึ่ง ในปัจจุบันมีวิธีการที่ใช้หลักการอันนี้เช่น

2.1 Thresholding เป็นวิธีการในการจำแนกพื้นที่ในภาพที่ง่ายที่สุด ทุกจุดบนภาพที่มีคุณสมบัติของภาพที่อยู่ในช่วงใดช่วงหนึ่งจะถูกจัดเป็นกลุ่มๆหนึ่ง ลักษณะเช่นนี้เหมาะกับลักษณะภาพที่ประกอบด้วยวัตถุที่เป็นเนื้อเดียวกันที่วางอยู่บนฉากหลังที่มีระดับความเข้มที่สูงกว่าและมีความเป็นเนื้อเดียวกัน เช่นตัวอักษรบนพื้นกระดาษสีขาวหรืออ่อนกว่า โดยการทำงานจะเลือกใช้ค่าคงที่ที่เรียกว่าค่าเธรชโชลด์มาใช้ในการแยกวัตถุออกจากกัน ซึ่งค่าเธรชโชลด์นี้อาจถูกเลือกมาจากประสบการณ์หรืออาจกำหนดให้โดยอ้างอิงจากฮิสโตแกรมของระดับความเข้ม โดยที่ค่าเธรชโชลด์จะเป็นค่าที่อยู่บริเวณรอยต่อระหว่างกลุ่มของข้อมูลในฮิสโตแกรม

2.2 Region Growing ในการจำแนกภาพโดยวิธีนี้ได้ถือว่าบริเวณที่อยู่ข้างเคียงเป็นบริเวณที่สำคัญมาก จุดภาพที่อยู่เคียงข้างกันมักจะมีคุณสมบัติทางสถิติที่คล้ายกันหรือใกล้เคียงกันสำหรับจุดรอบข้างที่มาเชื่อมต่อกัน ในวิธีนี้จะทำการพิจารณาภาพบริเวณย่อยๆจำนวนมาก จากนั้นพื้นที่ที่ติดกันจะถูกนำมาพิจารณาความเป็นเนื้อเดียวกันร่วมกัน การรวมตัวกันจะสิ้นสุดลงเมื่อพื้นที่ข้างเคียงไม่สามารถพิจารณาถึงความเป็นเนื้อเดียวกันได้ แต่ถ้าจุดภาพที่อยู่ใกล้เคียงนั้นตรวจสอบแล้วไม่อยู่ในเกณฑ์การรวม จุดภาพนั้นจะไม่ถูกรวมเข้าไปในส่วนนั้นของภาพแต่จะถูกเลือกให้เป็นจุดเริ่มต้นส่วนอื่นๆต่อไป และหลังจากที่จุดภาพทุกจุดได้รวมตัวกันเป็นกลุ่มเรียบร้อยแล้ว การที่จะทำให้ได้ส่วนของภาพสมบูรณ์ขึ้น ก็สามารถที่จะเริ่มใหม่โดยการแยกแต่ละส่วนออกแล้วใช้หลักการรวมกันของจุดภาพที่กล่าวมาแล้วอีกครั้งหนึ่ง วิธีที่ใช้หลักเกณฑ์ดังกล่าวคือ วิธี Split and Merge แต่สำหรับการจำแนกภาพโดยอาศัยพื้นที่เป็นหลักนี้ มีข้อจำกัดคือไม่สามารถที่จะสร้างส่วนของภาพที่เป็นเส้นที่มีขนาดเล็กในภาพได้ดี

การสร้างภาพไบนารี

การสร้างภาพไบนารี หมายถึงการแปลงข้อมูลภาพที่มีระดับความเข้มหลายระดับให้เป็นภาพที่มีความเข้มสองระดับหรือภาพไบนารี สำหรับเทคนิคต่างๆที่ใช้ในการสร้างภาพไบนารีมีหลายวิธี ขั้นตอนการสร้างภาพไบนารีจากภาพที่มีระดับความเข้มหลายระดับ สามารถแสดงได้ดังภาพที่ 8

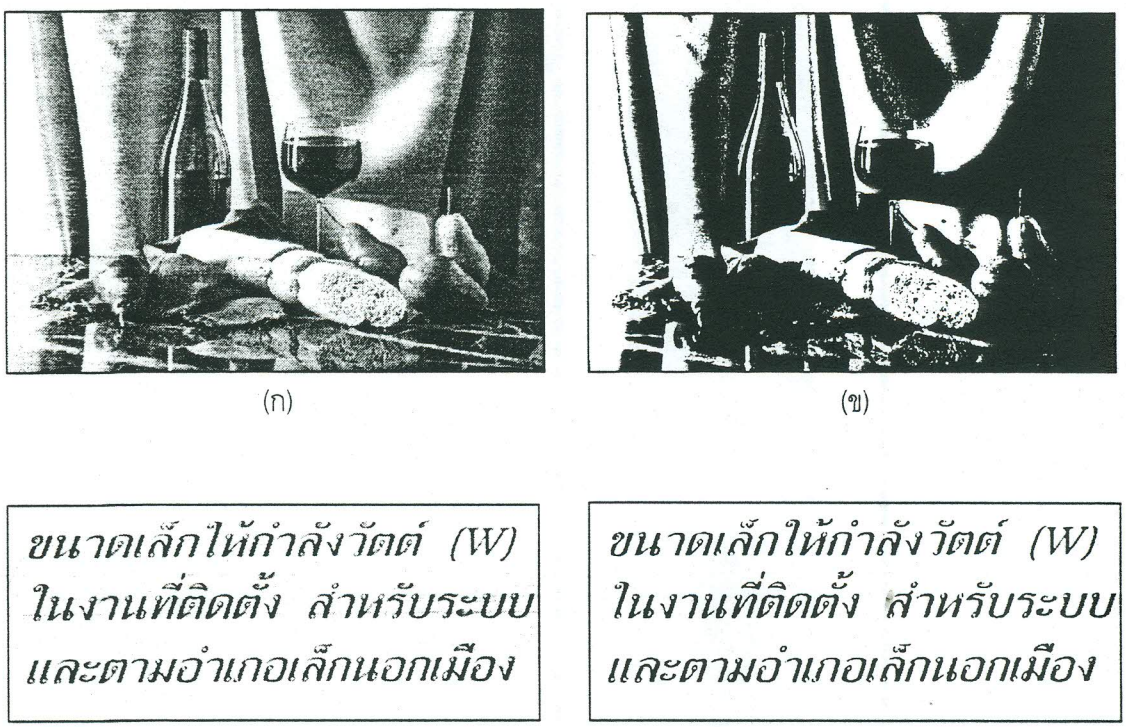
ภาพที่ 8 แสดงหลักการสร้างภาพไบนารี

จากภาพสามารถอธิบายหลักการทำงานของการสร้างภาพไบนารีได้ดังนี้ ภาพเริ่มต้นเป็นภาพที่มีแต่ละจุดภาพมีระดับเทาหลายระดับ ตั้งแต่ 0 จนถึง L ระดับ จากนั้นภาพเริ่มต้นนี้จะถูกนำเข้าไปสู่กระบวนการที่เรียกว่า "อิมเมจไบนาไรเซชั่น" (Image Binarization) ซึ่งเป็นกระบวนการในการแปลงข้อมูลภาพที่มีระดับความเข้มหลายระดับเป็นภาพไบนารี ภาพผลลัพธ์ที่ได้จากกระบวนการนี้จะได้เป็นภาพไบนารีที่แต่ละจุดภาพจะมีเพียง 2 ระดับคือ 1 (จุดดำ) กับ 0 (จุดขาว)

สำหรับกระบวนการสร้างภาพไบนารีนี้ โดยทั่วไปแล้วมีเทคนิคในการทำงานอยู่ 2 เทคนิคหลักๆ โดยที่แต่ละเทคนิคจะมีขั้นตอนวิธีการทำงานและวัตถุประสงค์ในการใช้งานแตกต่างกัน สำหรับรายละเอียดแต่ละเทคนิคจะได้อธิบายดังต่อไปนี้

1. เทคนิคการทำเธรชโชลด์ (Thresholding Technique)

เป็นเทคนิคที่ใช้ในการสร้างภาพไบนารี ซึ่งการพิจารณาว่าจุดภาพใดควรจะเป็นจุดขาวหรือจุดดำ จะกระทำโดยการเปรียบเทียบระหว่างจุดภาพเริ่มต้นกับค่าคงที่ค่าหนึ่งที่เรียกว่า "ค่าเธรชโชลด์" (Threshold Value) เทคนิคนี้ใช้กันมากในกรณีที่ข้อมูลภาพมีลักษณะแตกต่างกันมากระหว่างวัตถุ (object) และพื้นหลัง (background) เช่นภาพตัวอักษรที่ปรากฏบนพื้นกระดาษสีขาวหรืออ่อนกว่าเป็นต้น ตัวอย่างการสร้างภาพไบนารีโดยเทคนิคเธรชโชลด์แสดงได้ดังภาพที่ 9 สำหรับรายละเอียดและวิธีการในการทำเธรชโชลด์จะกล่าวถึงในหัวข้อเทคนิคการทำเธรชโชลด์



(ก) (ข)

(ค) (ง)

ภาพที่ 9 แสดงการสร้างภาพไบนารีโดยเทคนิคเธรชโชลด์
โดยที่ (ก) และ (ค) เป็นภาพอินพุต 256 ระดับเทา
ส่วนภาพ (ข) และ (ง) เป็นภาพไบนารีที่ได้จากการทำเธรชโชลด์

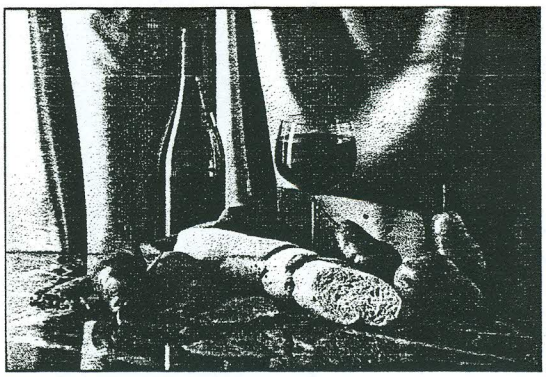
จากภาพที่ 9 จะเห็นได้ว่าภาพไบนารีที่ได้จากเทคนิคเธรชโชลด์สามารถให้ผลลัพธ์ที่คมชัดและสวยงาม โดยเฉพาะภาพที่มีความแตกต่างกันระหว่างวัตถุและพื้นหลัง ตัวอย่างเช่นภาพ (ง) และเมื่อเปรียบเทียบกันระหว่างภาพ (ข) และ ภาพ (ง) จะพบว่าภาพ (ง) มีความคมชัดสวยงามมากกว่า เนื่องจากภาพอินพุตเป็นภาพที่มีความแตกต่างระหว่างวัตถุหรือสิ่งที่ปรากฏในภาพได้เป็นสองส่วน ในกรณีนี้ก็คือส่วนตัวอักษรและส่วนที่เป็นพื้นหลังที่มีสีที่อ่อนกว่า เมื่อเป็นเช่นนี้แล้วเทคนิคนี้จึงสามารถแยกวัตถุเหล่านั้นได้อย่างดี และภาพที่ได้มีความเหมาะสมสวยงาม ส่วนภาพ (ข) นั้นเป็นภาพผลลัพธ์ที่ได้ไม่ค่อยมีความสวยงามนัก เนื่องจากความแตกต่างระหว่างวัตถุในภาพมีมากกว่าสองส่วน ค่าเธรชโชลด์จึงไม่สามารถแยกแต่ละส่วนหรือแต่ละวัตถุได้อย่างชัดเจน ภาพที่ได้จึงดูมืด (จุดดำมากเกินไป) หรือสว่าง (จุดขาวมากเกินไป) เกินไป

2. เทคนิคการทำฮาฟท์โทน (Halftoning Technique)

เทคนิคนี้เป็นอีกเทคนิคหนึ่งที่ใช้สร้างภาพไบนารี ซึ่งจุดประสงค์ของการทำฮาฟท์โทนนี้ก็เพื่อให้ได้ภาพไบนารีที่มีลักษณะของการมองดูคล้ายกับการไล่ระดับของความเข้มของสีเทา คือดูเหมือนมีการไล่ระดับจากสีขาวไปจนถึงสีดำ ทำให้ภาพที่ได้คล้ายกับภาพต้นฉบับหรือมีคุณภาพใกล้เคียงกับภาพต้นฉบับ โดยการทำงานจะพิจารณาจากจุดภาพที่อยู่รอบข้างมาใช้เพื่อตัดสินใจว่าจุดที่กำลังพิจารณาอยู่นั้นควรจะเป็นจุดขาวหรือจุดดำ โดยทั่วไปแล้วเทคนิคนี้จะใช้กันมากในการพิมพ์ข้อมูลภาพที่มีความเข้มหลายระดับทางเครื่องพิมพ์, การแสดงผลภาพบนจอภาพแบบโมโนโครม เป็นต้น ตัวอย่างการสร้างภาพโดยเทคนิคฮาฟท์โทนนี้สามารถแสดงได้ดังภาพที่ 10

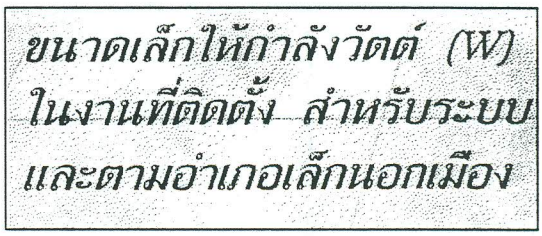
จากภาพจะเห็นได้ว่าภาพไบนารีที่ได้จากเทคนิคฮาฟท์โทนสามารถให้ผลลัพธ์ที่สวยงามดูเหมือนมีการไล่ระดับของสีเทาได้ดีทั้งสองภาพ และเมื่อเปรียบเทียบกันระหว่างภาพ (ข) และภาพ (ง) จะพบว่าทั้งสองภาพให้ภาพ (ข) มีความสวยงามมากกว่าภาพ (ง) เนื่องจากภาพอินพุตเป็นภาพที่มีการไล่ระดับของสีเทา ส่วนภาพ (ง) ภาพที่ได้จะไม่สวยงามเท่าที่ควร เนื่องจากภาพที่ใช้เราสามารถแยกความแตกต่างภายในภาพได้เป็นสองส่วนคือวัตถุและพื้นหลัง ในกรณีนี้ก็คือส่วนตัวอักษรและส่วนที่เป็นพื้นหลังที่มีสีที่อ่อนกว่า เมื่อเป็นเช่นนี้แล้วเทคนิคภาพผลลัพธ์ที่ได้จะทำให้เกิดจุดสีดำเล็กเกิดขึ้นมากมาย ซึ่งจุดเหล่านี้เป็นสิ่งที่เราไม่ต้องการทำให้ภาพผลลัพธ์ที่ได้ดูเหมือนมีรอยเปื้อนเกิดขึ้น

(ก)



(ข)

(ค)



(ง)

ภาพที่ 10 แสดงการสร้างภาพไบนารีโดยใช้เทคนิคฮาฟท์โทน
โดยที่ภาพ (ก) และ (ค) เป็นภาพอินพุต 256 ระดับเทา
ส่วนภาพ (ข) และ (ง) เป็นภาพไบนารีที่ได้จากการทำฮาฟท์โทน

จากภาพที่ 9 และ 10 จะเห็นได้ว่า การสร้างภาพไบนารีโดยเทคนิคทั้งสองให้ผลลัพธ์สวยงามเป็นที่น่าพอใจ แต่เมื่อเปรียบเทียบกันแล้วจะพบว่า เทคนิคเธรชโชลด์จะเหมาะกับข้อมูลภาพที่มีความแตกต่างระหว่างวัตถุและพื้นหลัง (เช่นภาพตัวอักษรบนพื้นกระดาษ) มากกว่าเทคนิคฮาฟท์โทน เนื่องจากเทคนิคนี้จะใช้ค่าเธรชโชลด์เป็นตัวแบ่งระหว่างวัตถุกับพื้นหลัง ภาพที่ได้จึงมีความคมชัดสวยงาม ส่วนเทคนิคฮาฟท์โทนจะเหมาะกับภาพที่มีลักษณะที่มีความต่อเนื่องของระดับความเข้มของภาพ คือดูเหมือนมีการไล่ระดับของสีเทาได้ดี ภาพที่ได้จึงดูสวยงามใกล้เคียงกับภาพต้นฉบับ

บทที่ 3

การจำแนกข้อความและภาพถ่ายอัตโนมัติ

บทนำ

ในการสร้างภาพไบนารีสำหรับภาพหน้าเอกสารนั้นจำเป็นอย่างยิ่งที่จะต้องทำให้คอมพิวเตอร์สามารถจำแนก (Segmentation) ได้ว่าส่วนใดที่ปรากฏในหน้าเอกสารนั้นเป็นส่วนของข้อความและส่วนใดเป็นของภาพถ่าย เพื่อที่จะได้นำแต่ละส่วนที่ได้จากการจำแนกไปทำการประมวลผลในขั้นตอนต่อไป ซึ่งในกรณีนี้ถ้าไม่มีการจำแนกส่วนประกอบของภาพหน้าเอกสารแล้ว อาจทำให้การประมวลผลในขั้นตอนของการของการสร้างภาพไบนารีผิดพลาดไป ทำให้คุณภาพของภาพผลลัพธ์ที่ได้ไม่ดีเท่าที่ควร เนื่องจากในขั้นตอนการสร้างภาพไบนารีของแต่ละส่วนใช้เทคนิคในการประมวลผลที่แตกต่างกัน เพราะฉะนั้นขบวนการจำแนกข้อความและภาพถ่ายของหน้าเอกสารจึงเป็นขั้นตอนที่สำคัญมากและจะขาดเสียไม่ได้ จึงจำเป็นอย่างยิ่งที่จะต้องทำการจำแนกให้ได้ว่าส่วนใดเป็นส่วนของข้อความและส่วนใดเป็นของภาพถ่าย เพื่อเป็นการจัดเตรียมข้อมูลให้กับการสร้างภาพเอกสารแบบไบนารีในขั้นตอนต่อไป

จากเหตุผลข้างต้น การจำแนกข้อความและภาพถ่ายของภาพหน้าเอกสารจึงเป็นขั้นตอนที่สำคัญขั้นตอนหนึ่งของการสร้างภาพไบนารีสำหรับภาพเอกสาร โดยทั่วไปแล้วหน้าเอกสารหนึ่งหน้าจะประกอบไปด้วยสองส่วนที่สำคัญคือ ส่วนที่เป็นข้อความ (Text Region) โดยทั่วไปมักจะมีสีดำหรือสีเข้ม (ตัวอักษร) บนพื้นกระดาษที่มีสีขาวหรืออ่อนกว่า และอีกส่วนหนึ่งก็คือส่วนที่เป็นภาพถ่าย (Photograph Region) ซึ่งส่วนนี้จะมีระดับสีเทาแตกต่างกันหลายระดับ ดังแสดงในภาพที่ 11

ภาพที่ 11 แสดงภาพหน้าเอกสารที่ประกอบไปด้วยส่วนของข้อความและภาพถ่าย

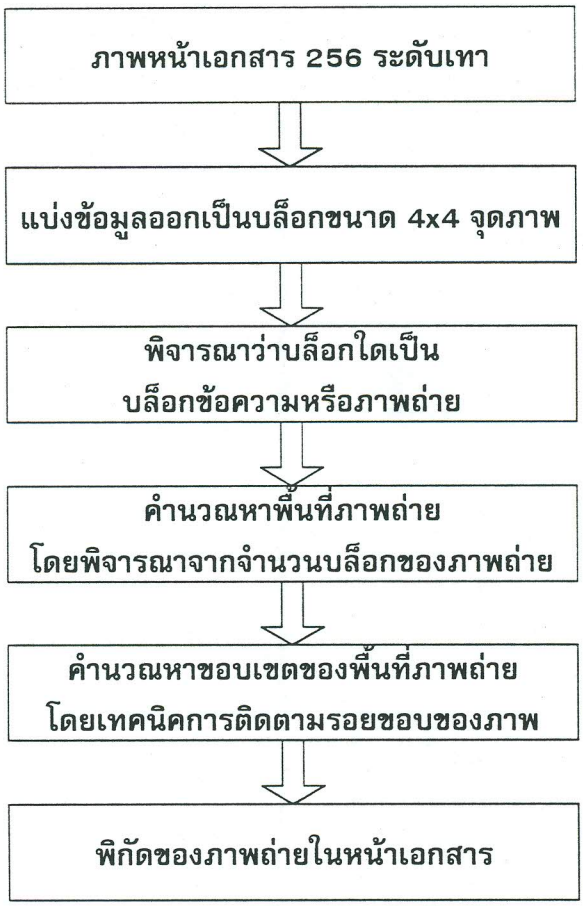
จากภาพที่ 11 จะแสดงให้เห็นถึงส่วนประกอบของหน้าเอกสารที่สามารถแยกออกได้เป็น 2 ส่วน โดยที่พื้นที่ที่ลูกศรชี้หมายถึงบริเวณที่เป็นภาพถ่าย และส่วนที่เหลือทั้งหมดคือส่วนที่เป็นข้อความ

สำหรับบทนี้จะเป็นการนำเสนอวิธีการในการจำแนกข้อความและภาพถ่ายในหน้าเอกสาร โดยการนำเทคนิคเธรชโชลด์ มาประยุกต์ใช้ในการวิเคราะห์ว่าส่วนใดเป็นส่วนของข้อความและส่วนใดเป็นภาพถ่าย และเทคนิคการติดตามรอยขอบของภาพมาประยุกต์ใช้ในการวิเคราะห์หาพิกัดของส่วนที่เป็นภาพถ่ายที่ปรากฏในภาพหน้าเอกสาร

64379

ขั้นตอนการทำงานของกระบวนการจำแนกส่วนของข้อความและภาพถ่าย

ในหัวข้อนี้เป็นการนำเสนอขั้นตอนและรายละเอียดการทำงานของกระบวนการจำแนกส่วนของข้อความและภาพถ่ายในหน้าเอกสาร ขั้นตอนการทำงานสามารถสรุปการทำงานออกได้ดังภาพที่ 12

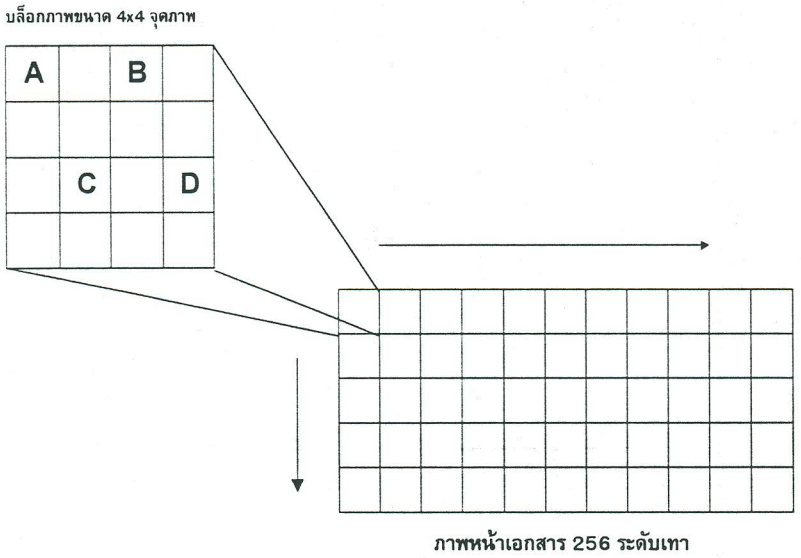


ภาพที่ 12 แสดงขั้นตอนของการจำแนกข้อความและภาพถ่ายอัตโนมัติ

จากภาพจะเห็นว่า การจำแนกข้อความและภาพถ่ายจากภาพหน้าเอกสารนั้น สามารถแบ่งออกเป็นขั้นตอนการทำงานย่อยๆหลายขั้นตอน ซึ่งสามารถอธิบายรายละเอียดในแต่ละขั้นตอนได้ดังนี้

1. แบ่งข้อมูลภาพเป็นบล็อกขนาด 4x4 จุดภาพ

การทำงานในขั้นตอนนี้เริ่มต้นด้วยการเตรียมข้อมูลภาพหน้าเอกสารหนึ่งหน้ากระดาษที่ได้จากเครื่องสแกนเนอร์ขาวดำ โดยกำหนดที่ความละเอียด 300 จุดต่อนิ้ว ภาพที่ได้เป็นภาพขาวดำที่มีระดับความเข้ม 256 ระดับ กล่าวคือแต่ละจุดภาพมีค่าระดับความเข้มตั้งแต่ 0 - 255 โดยที่จุดภาพที่มีค่าระดับความเข้มมีค่าต่ำๆ จะหมายถึงจุดภาพนั้นมีสีดำหรือสีเข้ม ส่วนจุดภาพที่มีระดับความเข้มมีค่าสูงๆจะหมายถึงจุดภาพนั้นมีสีขาว เมื่อได้ภาพหน้าเอกสารแล้วก็สามารถทำการจำแนกส่วนของข้อความและภาพถ่ายได้ทำการแบ่งข้อมูลภาพออกเป็นบล็อก (Block) เล็กๆ ซึ่งบล็อกเล็กๆนี้ต้องมีขนาดเล็กเพียงพอที่จะแสดงความเป็นเนื้อเดียวกันภายในได้อย่างชัดเจน ในงานวิจัยนี้ได้เลือกใช้บล็อกขนาด 4x4 จุดภาพ เพื่อที่จะใช้ในการจำแนกส่วนของข้อความและภาพถ่ายดังกล่าว โดยที่แต่ละบล็อกจะต้องไม่มีส่วนที่ซ้อนทับกัน และทิศทางการประมวลผลจะกระทำในทิศทางจากซ้ายไปขวาและจากบนลงล่าง ดังแสดงในภาพที่ 13



ภาพที่ 13 แสดงการแบ่งข้อมูลภาพออกเป็นบล็อกขนาด 4x4 จุดภาพ

ในการที่เลือกใช้บล็อกขนาด 4x4 จุดภาพเพื่อการจำแนกส่วนของข้อความและภาพถ่ายนั้น ได้มีการกำหนดจุดอ้างอิงขึ้นมา 4 จุดคือ จุด A, B, C และ D เพื่อใช้ในการตัดสินใจว่าบล็อกที่กำลังพิจารณาอยู่นั้นเป็นบล็อกของข้อความ (Text Block) หรือบล็อกของภาพถ่าย (Photograph Block) สำหรับจุดอ้างอิงที่กำหนดขึ้นมา 4 จุดนี้เราสามารถที่จะกำหนดขึ้นมาให้มากว่าหรือน้อยกว่า 4 จุดก็ได้ ถ้ากำหนดมากกว่า 4 จุดก็จะทำให้การพิจารณาแยกระหว่างข้อความและภาพถ่ายนั้นละเอียดและครอบคลุมยิ่งขึ้น แต่จะเสียเวลามากในการประมวลผล และ

ถ้ากำหนดจุดในการพิจารณาน้อยก็จะทำให้การพิจารณาไม่ครอบคลุมเท่าที่ควร ซึ่งจะเป็นผลทำให้การจำแนกข้อความและภาพถ่ายได้ผลลัพธ์ที่ไม่ถูกต้อง

เมื่อพิจารณาจากการกำหนดจุดที่ใช้อ้างอิง 4 จุดเพื่อการจำแนกข้อความและภาพถ่าย พบว่าเพียงพอสำหรับใช้ในการจำแนก เนื่องจากโดยทั่วไปแล้วบริเวณที่เป็นภาพถ่าย จุดภาพที่อยู่ใกล้เคียงกันจะมีความคล้ายคลึงกันมาก (ระดับความเข้มไม่แตกต่างกันมากนัก) จากเหตุผลนี้เองเมื่อเรานำแต่ละบล็อกมาวางเรียงกันทั้งในแนวนอนและแนวตั้ง จะพบว่าจุดที่ใช้ในการจำแนกทั้ง 4 จุดในแต่ละบล็อกจะอยู่กระจัดกระจายต่อเนื่องกันอย่างสม่ำเสมอ ดังนั้นจะเห็นได้ว่าการกำหนดจุดอ้างอิงเพียง 4 จุดในแต่ละบล็อกก็เพียงพอสำหรับใช้ในการจำแนกข้อความและภาพถ่าย

2. พิจารณาว่าบล็อกใดเป็นบล็อกข้อความหรือภาพถ่าย

หลังจากที่ทำการแบ่งข้อมูลออกเป็นบล็อกเล็กๆขนาด 4x4 จุดภาพแล้ว ในขั้นตอนนี้จะเป็นการพิจารณาว่าแต่ละบล็อกนั้น บล็อกใดควรจะเป็นบล็อกของข้อความหรือบล็อกใดควรจะเป็นบล็อกของภาพถ่าย สำหรับการทำงานของขั้นตอนนี้ได้นำเอาเทคนิคเธรชโชลด์มาประยุกต์ใช้ เนื่องจากโดยทั่วไปแล้วส่วนที่เป็นข้อความ จะเป็นส่วนที่ประกอบไปด้วยตัวอักษรสีเข้มหรือสีดำบนพื้นกระดาษสีขาว และส่วนที่เป็นภาพถ่ายจะเป็นส่วนที่มีสีเข้มกว่าสีขาวขึ้นไป แต่จะไม่เข้มมากเท่ากับความเข้มของส่วนที่เป็นตัวอักษร กล่าวคือระดับความเข้มของระดับสีเทาในส่วนที่เป็นภาพถ่ายจะมีระดับความเข้มอยู่ระหว่างสีขาวและสีดำ ด้วยเหตุนี้จึงได้มีการกำหนดค่าเธรชโชลด์ขึ้นมา 2 ค่าเพื่อใช้เป็นเกณฑ์ในการพิจารณาว่าเป็นบล็อกของข้อความหรือบล็อกของภาพถ่าย โดยเรียกว่า ค่าขอบเขตบน (*Upper Bar*) และค่าขอบเขตล่าง (*Lower Bar*) ซึ่งสามารถคำนวณได้จากสมการที่ 3.1 และ 3.2 โดยในการคำนวณหาค่าเธรชโชลด์ทั้งสองค่านี้จะทำการคำนวณทุกๆ 64 บรรทัดของภาพหน้าเอกสาร

$$Upper\ Bar = Maximum\ Gray\ Level \times 0.9 \quad \text{.....(3.1)}$$

$$Lower\ Bar = Maximum\ Gray\ Level \times 0.1 \quad \text{.....(3.2)}$$

โดยที่

Upper Bar = ค่าขอบเขตบน

Lower Bar = ค่าขอบเขตล่าง

Maximum Gray Level = ค่าระดับเทาสูงสุดของภาพหน้าเอกสาร

ในการพิจารณาว่าบล็อกใดเป็นข้อความหรือบล็อกใดเป็นภาพถ่าย สามารถพิจารณาโดยใช้เงื่อนไขดังนี้

- ถ้าจุด A, B, C และ D จุดใดจุดหนึ่งมากกว่าหรือเท่ากับค่าขอบเขตบน แสดงว่าบล็อกนั้นเป็นบล็อกของข้อความ ถ้าไม่ใช่แสดงว่าเป็นบล็อกของภาพถ่าย

- ถ้าจุด A, B หรือจุด A, C หรือจุด B, D หรือจุด C, D มีค่าน้อยกว่าหรือเท่ากับค่าขอบเขตล่างแสดงว่าบล็อกนั้นเป็นบล็อกของข้อความ ถ้าไม่ใช่แสดงว่าเป็นบล็อกของภาพถ่าย

สำหรับในขั้นตอนนี้เราจะได้ภาพหน้าเอกสารขึ้นมาอีก 1 ภาพ ที่มีขนาดความกว้างและความยาวเท่ากับหน้าเอกสารต้นฉบับ แต่ภาพที่ได้ใหม่นั้นเป็นภาพแบบไบนารี นั่นคือแต่ละจุดภาพจะแทนด้วยเลขฐานสองเท่านั้น โดยกำหนดให้จุดภาพที่แทนด้วย 1 (จุดดำ) หมายถึงจุดภาพที่เป็นภาพถ่าย ส่วนจุดภาพที่แทนด้วย 0 (จุดขาว) จะหมายถึงจุดภาพที่เป็นข้อความ จากการพิจารณาว่าบล็อกใดควรจะเป็นข้อความหรือบล็อกใดควรจะเป็นภาพถ่ายแล้วจะกำหนดให้บล็อกที่เป็นภาพถ่ายนั้นแทนด้วย 1 ส่วนบล็อกที่เป็นข้อความจะถูกกำหนดให้เป็น 0 เมื่อสิ้นสุดขั้นตอนนี้แล้วจะได้ข้อมูลภาพไบนารี เพื่อใช้ในการประมวลผลในขั้นตอนต่อไป ซึ่งแสดงดังภาพที่ 15

1.3 คำนวณหาพื้นที่ภาพถ่ายโดยการพิจารณาจากจำนวนบล็อกภาพถ่าย

โดยทั่วไปแล้วพื้นที่ที่เป็นภาพถ่ายนั้นจะประกอบไปด้วยบล็อกของภาพถ่ายหลายๆ บล็อกที่มีความคล้ายคลึงกันรวมเป็นพื้นที่ที่ติดต่อกัน จากเหตุผลนี้เราสามารถหาพื้นที่ที่ประกอบกันเป็นพื้นที่ของภาพถ่าย โดยการพิจารณาจากจำนวนบล็อกของภาพถ่ายที่อยู่ติดต่อกัน สำหรับการทดลองได้กำหนดให้ภาพถ่ายที่ปรากฏอยู่ในภาพหน้าเอกสารมีขนาดตั้งแต่ 2/3 นิ้วขึ้นไป (ประมาณ 50 บล็อกขึ้นไป) ดังนั้นเงื่อนไขในการปรับปรุงการจำแนกสามารถกำหนดได้ดังนี้

- ถ้าจำนวนบล็อกของภาพถ่ายที่เรียงติดกันมากกว่า 50 บล็อก แสดงว่าเป็นพื้นที่ของภาพถ่าย (Photograph Region) ถ้าไม่ใช่แสดงว่าเป็นพื้นที่ของข้อความ (Text Region)

จากเงื่อนไขดังกล่าวจะพบว่าบล็อกของภาพถ่ายที่เรียงติดต่อกันน้อยกว่าค่าที่กำหนดจะถูกกำจัดออกไปและบล็อกของภาพถ่ายที่ปะปนอยู่ในส่วนของพื้นที่ของข้อความก็จะถูกกำจัดออกไปด้วย

สำหรับในขั้นตอนนี้ภาพไบนารีที่ได้จากขั้นอนที่ 1.2 จะเปลี่ยนแปลงไปเนื่องจากการปรับปรุงการจำแนกเพื่อหาพื้นที่ที่เป็นภาพถ่าย กล่าวคือเมื่อได้พื้นที่ของภาพถ่ายแล้ว จะกำหนดให้พื้นที่ที่เป็นภาพถ่ายเหล่านั้นด้วย 1 นอกนั้น (พื้นที่ที่เป็นข้อความ) ด้วย 0 เมื่อสิ้นสุดขั้นตอนนี้แล้วจะได้ข้อมูลภาพไบนารี เพื่อใช้ในการประมวลผลในขั้นตอนต่อไป ซึ่งแสดงดังภาพที่ 16

1.4 คำนวณหาพิกัดของพื้นที่ภาพถ่าย

ขั้นตอนการแยกและคัดลอกพื้นที่ที่เป็นภาพถ่ายเป็นขั้นตอนของว่าการวิเคราะห์ว่าหน้าเอกสารนั้นประกอบไปด้วยภาพถ่ายกี่ภาพ แต่ละภาพมีพิกัดหรือตำแหน่งอยู่ที่ใด ในขั้นตอนนี้ได้นำเทคนิคของการติดตามรอยขอบของภาพมาประยุกต์ใช้ โดยการทำงานจะกระทำกับภาพไบนารีที่ได้จากขั้นตอนที่ 1.3 การทำงานเริ่มจากตรวจกวาดจุดภาพจากจุดมุมบนด้านซ้ายไปขวาและบนลงล่าง จนกระทั่งพบจุดภาพที่เป็นจุดใดจุดหนึ่งของขอบเขตที่เป็นพื้นที่ภาพถ่าย ซึ่งมีค่าของจุดภาพเป็น 1 (จุดดำ) ก็จะเคลื่อนที่ติดตามรอยขอบของภาพไป ตำแหน่งหรือพิกัดของจุด (Coordinate) ที่เป็นขอบภาพจะถูกบันทึกเก็บไว้ด้วย และเมื่อเคลื่อนที่ไปจนรอบ นั่นคือกลับมาอยู่ที่จุดเริ่มต้นแล้ว ก็จะนำพิกัดของจุดที่เป็นขอบภาพทั้งหมดมาคำนวณหาจุดพิกัดที่มีค่ามากที่สุดและน้อยที่สุดทั้งในแนวนอนและแนวตั้ง มากำหนดเป็นจุดมุมทั้งสี่ของพื้นที่ที่เป็นภาพถ่าย เมื่อได้จุดมุมทั้งสี่ของภาพถ่ายแต่ละภาพแล้วทำการตรวจสอบภาพถ่ายเหล่านั้น ด้วยเงื่อนไขต่อไปนี้

- พื้นที่ภาพถ่ายที่ได้ต้องมีขนาดความกว้างและความยาวไม่น้อยกว่า 1/2 นิ้ว (150 บรรทัดในแนวตั้งและ 150 จุดภาพในแนวนอน) หากไม่เป็นไปตามเงื่อนไขพื้นที่ภาพถ่ายนั้นจะถูกเปลี่ยนให้เป็นพื้นที่ข้อความแทน
- พื้นที่ภาพถ่ายแต่ละภาพ ต้องห่างกันมากกว่า 1/3 นิ้ว (100 จุดภาพในแนวนอนและ 100 บรรทัดในแนวตั้ง) ถ้าน้อยกว่าจะทำการรวมพื้นที่ภาพนั้นเข้าด้วยกัน

เมื่อคำนวณหาจุดมุมทั้ง 4 ของรูปสี่เหลี่ยมแล้ว ก็จะสามารถกำหนดขอบเขตของส่วนที่เป็นภาพถ่ายในหน้าเอกสารนั้นๆ ได้ในลักษณะของพื้นที่รูปสี่เหลี่ยม จุดภาพทั้งหมดที่อยู่ภายในบริเวณรูปสี่เหลี่ยมนั้นจะถูกเปลี่ยนค่าของจุดภาพภายในบริเวณนั้นๆให้มีค่าเป็น 0 หรือกล่าว

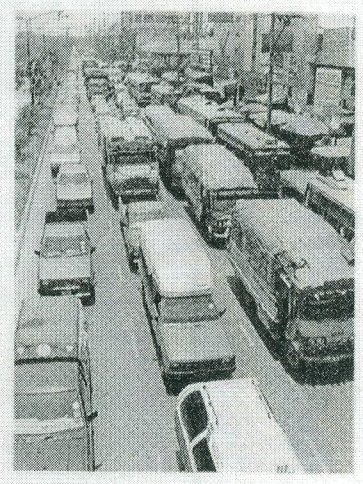
อีกนัยหนึ่งว่าเปลี่ยนจุดดำของจุดภาพให้กลายเป็นจุดขาวหรือจุดพื้นให้หมดจากนั้นก็จะย้อนกลับไปยังจุดเริ่มต้นที่พบขอบเขตก่อนหน้านี้ที่เพิ่งจะคัดลอกจุดภาพออกไป เริ่มตรวจกวาดหาจุดภาพที่เป็นขอบเขตอื่นต่อไป เพื่อที่จะกำหนดขอบเขตและคัดลอกขอบเขตนั้นๆออกมา การทำงานจะทำงานวนไปในลักษณะเช่นนี้ไปเรื่อยๆ จนสามารถกำหนดและแยกคัดลอกขอบเขตที่เป็นส่วนประกอบของหน้าเอกสารออกมาได้ทั้งหมด

เมื่อภาพหน้าเอกสารได้ทำงานผ่านขั้นตอนการแยกและคัดลอกพื้นที่ภาพถ่าย ผลลัพธ์ที่ได้คือ พิกัดของพื้นที่ที่เป็นภาพถ่ายในหน้าเอกสาร โดยที่พิกัดที่ได้จะเป็นจุดมุมของรูปสี่เหลี่ยม คือมุมบนซ้ายและมุมล่างขวาของพื้นที่ภาพถ่าย ส่วนพื้นที่ที่นอกเหนือจากพื้นที่ๆเป็นภาพถ่ายนั้นก็คือพื้นที่ๆเป็นข้อความ เมื่อได้พิกัดของพื้นที่ภาพถ่ายแล้วจะถูกส่งต่อไปประมวลผลในขั้นตอนต่อไป นั่นก็คือสร้างภาพไบนารีของหน้าเอกสาร ซึ่งจะใช้เทคนิคในการสร้างภาพแตกต่างกัน โดยพื้นที่ๆเป็นภาพถ่ายจะสร้างภาพไบนารีโดยใช้เทคนิคฮาฟท์โทน ส่วนพื้นที่ๆเป็นข้อความ (พื้นที่ๆอยู่นอกพิกัดของพื้นที่ภาพถ่าย) จะสร้างภาพโดยใช้เทคนิคเธรชโชลด์

การทดลองการจำแนกข้อความและภาพถ่ายอัตโนมัติ

ในการทดลองการจำแนกข้อความและภาพถ่ายสำหรับภาพหน้าเอกสาร ได้ทำการทดลองกับภาพหน้าเอกสารที่ประกอบไปด้วยข้อความและภาพถ่ายปะปนอยู่ในหน้าเดียวกัน โดยภาพหน้าเอกสารได้มาจากเครื่องสแกนเนอร์ขาวดำ ที่มีความละเอียด 300 จุดต่อนิ้วและเป็นภาพที่มีระดับความเข้ม 256 ระดับ ทำการพัฒนาโปรแกรมด้วยภาษา C และทดลองด้วยเครื่องไมโคร-คอมพิวเตอร์ รุ่น 486DX2-50 ซึ่งปรากฏผลการทดลองจำแนกภาพหน้าเอกสารในแต่ละขั้นตอนดังภาพที่ 14 - 17

1. แก้ปัญหาจราจรให้คลายจลาจล



ด้วยการดำเนินโครงการทั้ง ๕ ให้เสร็จสิ้นในเวลารวดเร็ว คือ ถนนวงแหวนรอบกลาง-รอบนอก และถนนใยแมงมุม (2-3 ปี) ทางด่วนพิเศษให้ครบทั้ง 5 ขั้นตอน (3 ปี) รถขนส่งมวลชนทั้งติดดินยกระดับและใต้ดิน (4-6 ปี) รถไฟชานเมืองทั้ง 3 ทิศของ กทม. (3 ปี) และเมืองบริวารรอบกรุงเทพ 20 เมือง (3 ปี)

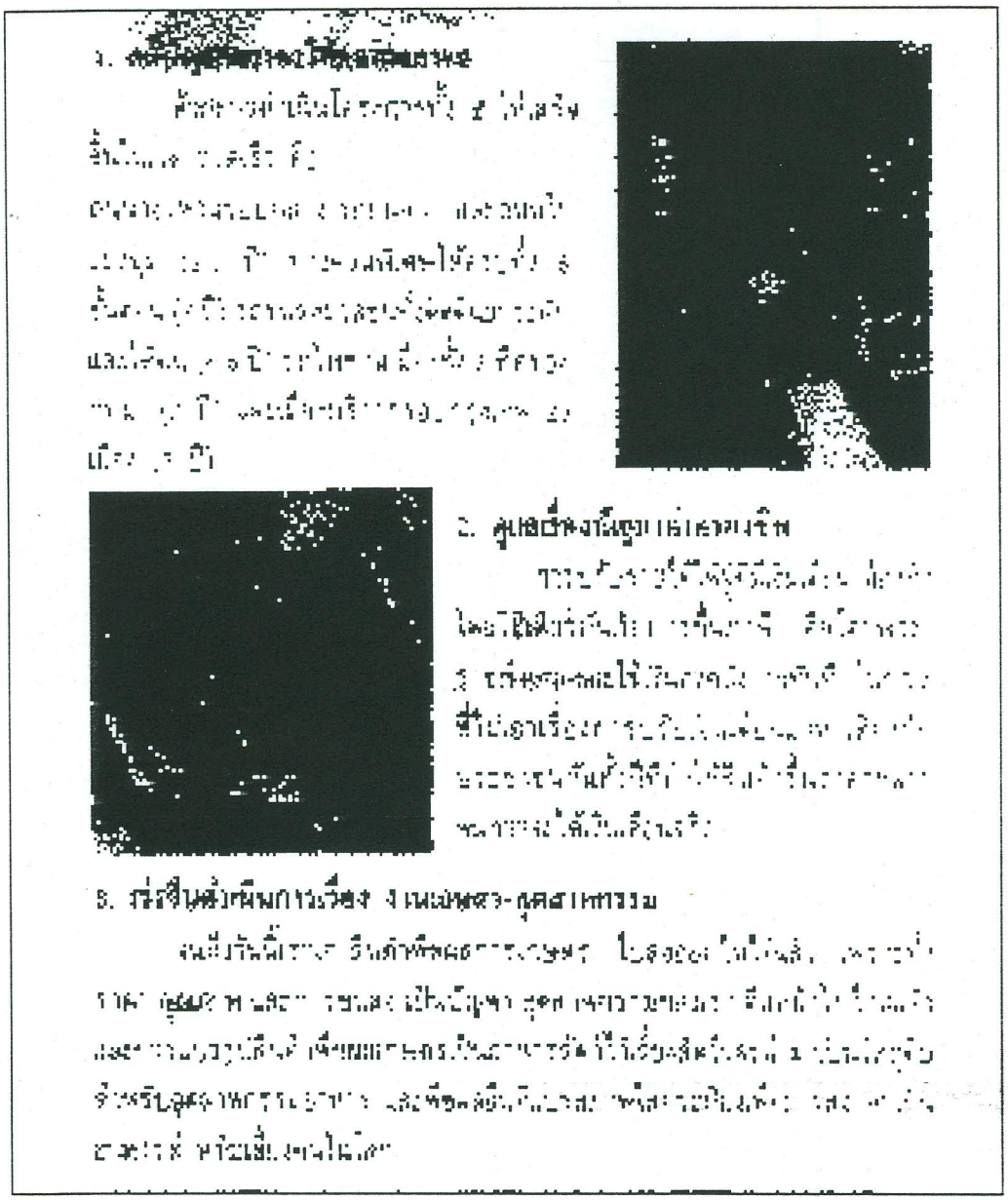
2. ดูแลเรื่องปัญหาค่าครองชีพ

การปรับรายได้ให้ผู้ที่มีเงินเดือน ต้องทำโดยวิธีเดียวกันกับการขึ้นภาษี คือโดยพระราชกำหนดและใช้เงินคงคลังจ่ายทันที ในขณะที่ไม่เอาเรื่องการปรับเงินเดือนมาหาเสียงกับประชาชนกันทั้งปีที่ทำให้สินค้าขึ้นราคาหลายหนกว่าจะได้เงินเดือนจริง

3. เร่งรีบดำเนินการเรื่อง งานเกษตร-อุตสาหกรรม

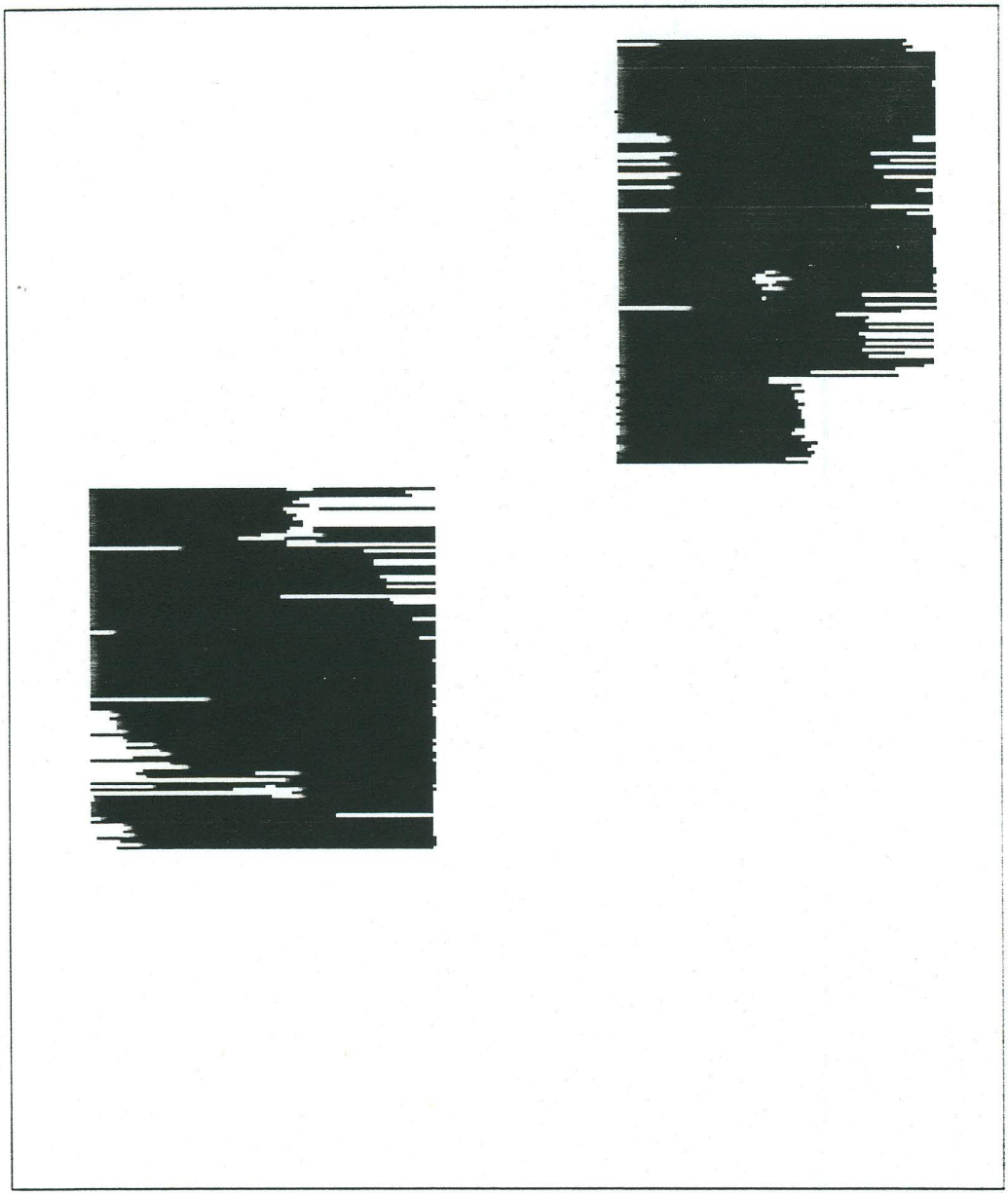
จนถึงวันนี้เราเอาสินค้าพืชผลการเกษตร ไปส่งออกไม่ได้แล้ว เพราะทั้งราคา คุณภาพ และการขนส่ง เป็นปัญหา อุตสาหกรรมของเราเดินหน้าไปไกลแล้ว และการแปรรูปสินค้าพืชผลเกษตรเป็นอาหารสัตว์ไว้เลี้ยงสัตว์และนำมาเป็นวัตถุดิบสำหรับอุตสาหกรรมอาหาร และพืชผลอื่นที่แปรสภาพใส่กระป๋องเพื่อการส่งออกเป็นอาหารสำหรับเลี้ยงคนในโลก

ภาพที่ 14 แสดงภาพหน้าเอกสารต้นฉบับที่ใช้ในการจำแนกส่วนของข้อความและภาพถ่าย



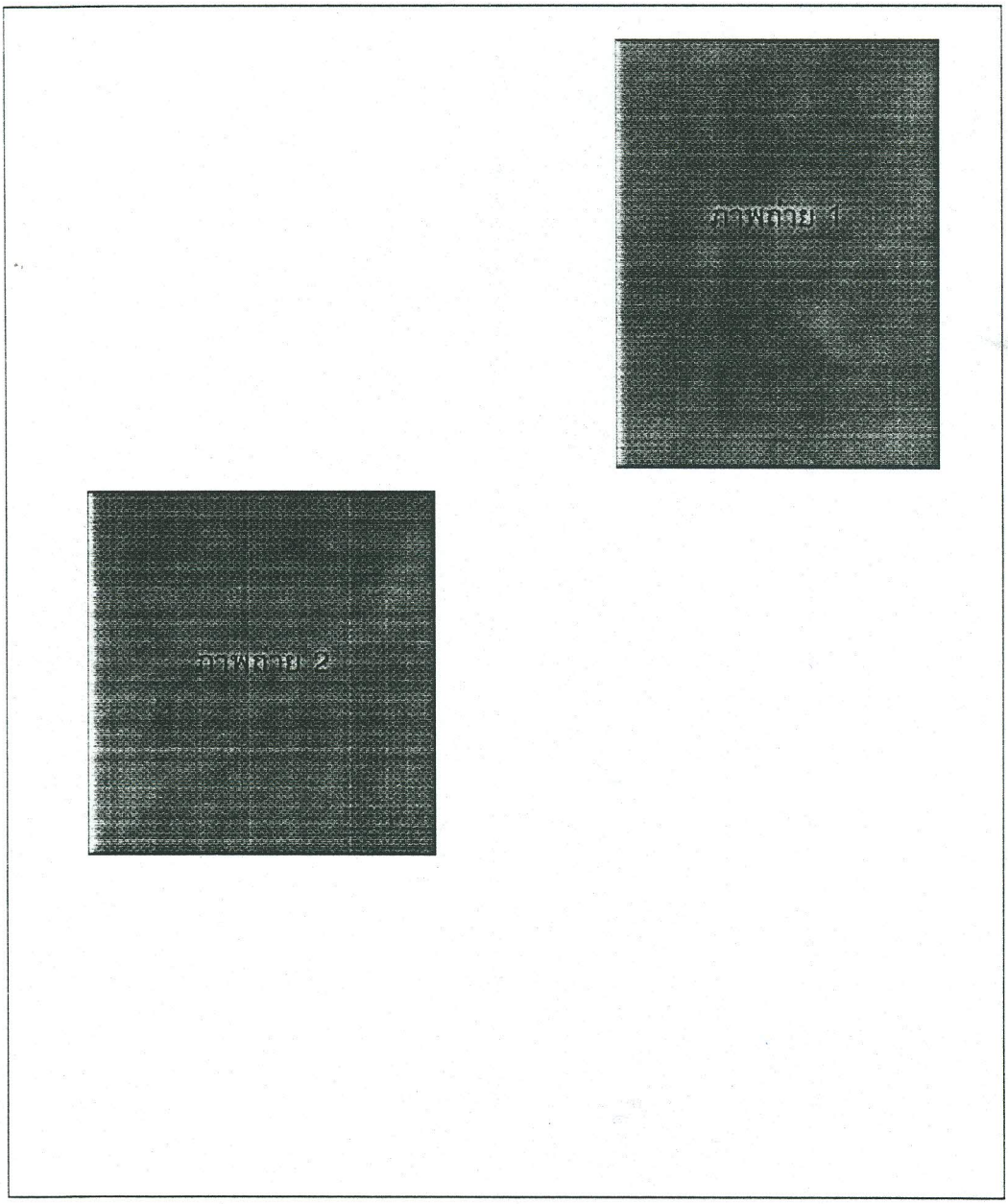
ภาพที่ 15 แสดงผลลัพธ์ที่ได้จากการพิจารณาว่าบล็อกใดเป็นบล็อกข้อความหรือ
บล็อกใดเป็นบล็อกของภาพถ่าย

จากภาพจะแสดงให้เห็นถึงผลลัพธ์ของการจำแนกในขั้นตอนที่ 2 ในการพิจารณาแต่ละบล็อกว่าบล็อกใดเป็นบล็อกข้อความหรือบล็อกใดเป็นบล็อกของภาพถ่าย โดยบริเวณที่เป็นสีดำหมายถึงบล็อกที่เป็นภาพถ่าย ส่วนที่เหลือ (สีขาว) หมายถึงบล็อกของข้อความ



ภาพที่ 16 แสดงผลลัพธ์ที่ได้จากการคำนวณหาพื้นที่ภาพถ่าย
โดยการพิจารณาจากจำนวนบล็อกของภาพถ่าย

จากภาพจะแสดงให้เห็นถึงผลลัพธ์ของการจำแนกในขั้นตอนที่ 3 ในการพิจารณาหาพื้นที่ที่เป็นภาพถ่าย โดยการพิจารณาจากจำนวนบล็อกของภาพถ่าย โดยที่บริเวณที่เป็นสีดำหมายถึงบริเวณที่เป็นภาพถ่าย ส่วนที่เหลือ (สีขาว) หมายถึงบริเวณที่เป็นข้อความ



ภาพที่ 17 แสดงพื้นที่ภาพถ่ายที่ได้จากการจำแนกข้อความและภาพถ่าย

จากภาพจะแสดงให้เห็นถึงผลลัพธ์ของการจำแนกในขั้นตอนที่ 4 ในการพิจารณาหาพื้นที่ที่เป็นภาพถ่าย โดยการนำเทคนิคของการติดตามรอยขอบของภาพมาคำนวณหาพิกัดของภาพถ่ายแต่ละภาพในหน้าเอกสาร โดยที่บริเวณที่เป็นสีดำหมายถึงบริเวณที่ถูกจำแนกว่าเป็นภาพถ่าย ซึ่งในที่นี้จะพบว่ามีภาพถ่ายจำนวน 2 ภาพ ส่วนที่เหลือ (สีขาว) หมายถึงบริเวณที่เป็นข้อความ

บทสรุป

ในการจำแนกข้อความและภาพถ่ายอัตโนมัติ สำหรับภาพหน้าเอกสารที่ประกอบไปด้วยข้อความและภาพถ่ายที่ปะปนอยู่ในหน้าเดียวกันนั้น จะเห็นได้ว่า ภาพผลลัพธ์ที่ได้นั้นจะได้เป็นพิกัดจุดมุมสี่เหลี่ยมของภาพถ่ายแต่ละภาพที่ปรากฏอยู่ในหน้าเอกสาร ซึ่งการประมวลผลจะทำการประมวลผลหลายขั้นตอน โดยจะเริ่มจากภาพหน้าเอกสารที่มีระดับความเข้ม 256 ระดับเทา ถูกแบ่งออกเป็นบล็อกเล็กๆขนาด 4x4 และในแต่ละบล็อกก็จะถูกกำหนดจุดขึ้นมาเพื่อใช้อ้างอิงว่าแต่ละบล็อกบล็อกใดควรจะเป็นบล็อกของข้อความบล็อกใดควรจะเป็นภาพถ่ายจำนวน 4 จุด จากนั้นเมื่อทำการวิเคราะห์แต่ละบล็อกแล้วก็จะรวมบล็อกที่เป็นภาพถ่ายเข้าด้วยกัน โดยการพิจารณาจากจำนวนบล็อก ซึ่งก็จะได้พื้นที่ๆเป็นภาพถ่ายออกมาและในขั้นตอนสุดท้ายก็จะทำการหาพิกัดมุมทั้ง 4 ของภาพถ่ายแต่ละภาพในหน้าเอกสาร โดยการใช้การติดตามรอยขอบของภาพมาใช้ในการหาพิกัดครั้งนี้ ซึ่งจะเห็นได้ว่าเมื่อได้พิกัดของพื้นที่ที่เป็นภาพถ่ายแล้ว เราสามารถเอาเฉพาะส่วนที่เป็นภาพถ่ายไปประมวลผลในระบบประยุกต์อื่นๆได้ เช่นนำส่วนของภาพถ่ายไปจัดเก็บในรูปแบบของภาพกราฟฟิค เช่น TIFF, BMP, GIF ฯลฯ เพื่อจัดทำเป็นระบบจัดเก็บข้อมูลภาพเป็นต้น อีกทั้งยังสามารถนำเฉพาะส่วนที่เป็นข้อความที่อยู่นอกพิกัดของภาพถ่ายไปประมวลผลในการรู้จำตัวอักษรได้อีกด้วย

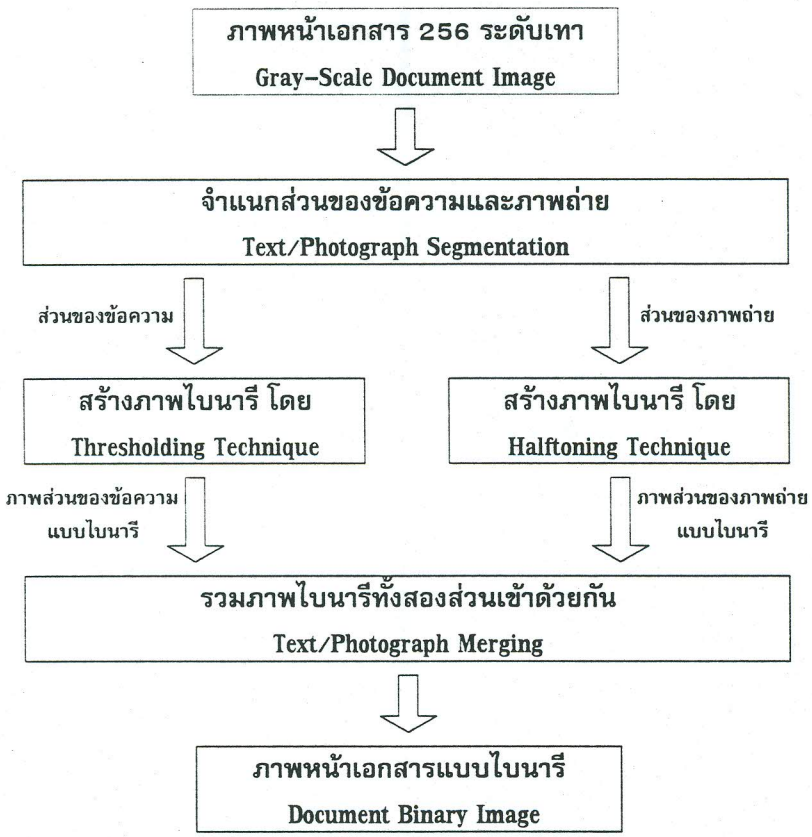
บทที่ 4

การสร้างภาพไบนารีสำหรับภาพเอกสาร

การสร้างภาพหน้าเอกสารแบบไบนารี เป็นขั้นตอนที่สำคัญอย่างหนึ่งในระบบการประมวลผลเอกสาร (Document Analysis System) ซึ่งขั้นตอนนี้เป็นขั้นตอนที่ทำการแปลงข้อมูลภาพหน้าเอกสารที่มีระดับความเข้มหลายระดับไปเป็นภาพที่มีระดับความเข้มเพียง 2 ระดับ โดยทั่วไปแล้วเอกสารหนึ่งหน้าจะประกอบไปด้วย 2 ส่วนที่สำคัญคือ ส่วนที่เป็นข้อความ (Text Region) และส่วนที่เป็นภาพถ่าย (Photograph Region) จากบทที่ผ่านมาจะเห็นได้ว่ามีเทคนิคในการสร้างภาพไบนารีที่สำคัญ 2 เทคนิค โดยแต่ละเทคนิคมีความเหมาะสมในการสร้างภาพไบนารีที่แตกต่างกัน ซึ่งจะพบว่าเทคนิคเธรชโชลด์เป็นเทคนิคที่เหมาะสมสำหรับการสร้างภาพไบนารีในส่วนที่เป็นข้อความ ส่วนเทคนิคฮาฟท์โทนเป็นเทคนิคที่เหมาะสมสำหรับส่วนที่เป็นภาพถ่าย เนื่องจากสามารถให้ผลลัพธ์ที่เป็นภาพไบนารีที่ดูเหมือนมีการไล่ระดับของสีเทาที่ต่อเนื่อง ดังนั้นในงานวิจัยนี้จึงได้นำเสนอวิธีการในการสร้างภาพไบนารีคุณภาพสูงสำหรับภาพหน้าเอกสารหลายองค์ประกอบ โดยทำการจำแนกภาพหน้าเอกสารออกเป็น 2 ส่วนคือส่วนที่เป็นข้อความและส่วนที่เป็นภาพถ่าย จากนั้นทำการสร้างภาพไบนารีในแต่ส่วนด้วยวิธีที่แตกต่างกัน คือเทคนิคเธรชโชลด์สำหรับส่วนที่เป็นข้อความและเทคนิคฮาฟท์โทนสำหรับส่วนที่เป็นภาพถ่าย สำหรับรายละเอียดในการทำงานจะได้กล่าวในหัวข้อต่อไป

ขั้นตอนการทำงาน

สำหรับการสร้างภาพไบนารีคุณภาพสูงสำหรับภาพเอกสารหลายองค์ประกอบนี้ มีขั้นตอนการทำงานดังแสดงในภาพที่ 18



ภาพที่ 18 แสดงขั้นตอนการทำงานของระบบการจำแนกข้อความและภาพถ่ายเพื่อสร้างภาพไบนารี

จากภาพสามารถอธิบายขั้นตอนการทำงานของการสร้างภาพไบนารีสำหรับภาพหน้าเอกสาร โดยการทำงานจะเริ่มจากข้อมูลภาพหน้าเอกสารที่ได้จากเครื่องสแกนเนอร์ขาวดำที่มีความละเอียด 300 จุดต่อนิ้ว ระดับความเข้ม 256 ระดับ ถูกนำเข้ามาสู่กระบวนการจำแนกข้อความและภาพถ่ายอัตโนมัติ ซึ่งจะได้ผลลัพธ์ออกมา 2 ส่วนคือ ส่วนของข้อความและส่วน

ของภาพถ่าย จากนั้นในแต่ละส่วนจะถูกนำไปสร้างเป็นภาพไบนารีด้วยวิธีที่แตกต่างกัน กล่าวคือ ส่วนที่เป็นข้อความจะถูกนำไปสร้างเป็นภาพไบนารีโดยใช้เทคนิคเธรซโฮลด์ และส่วนที่ภาพถ่ายก็จะถูกส่งไปสร้างเป็นภาพไบนารีโดยเทคนิคฮาฟท์โทน หลังจากนั้นเมื่อทำการสร้างภาพไบนารีทั้งสองส่วนแล้ว จะเป็นการรวมทั้งสองส่วนที่สร้างเป็นภาพไบนารีแล้วเข้าด้วยกัน ซึ่งจะได้ภาพไบนารีที่มีความถูกต้องสวยงามและมีคุณภาพใกล้เคียงกับภาพต้นฉบับ สำหรับรายละเอียดในการทำงานแต่ละขั้นตอนจะได้กล่าวถึงรายละเอียดในหัวข้อต่อไป

1. การจำแนกส่วนของข้อความและภาพถ่าย

ในขั้นตอนนี้จะเป็นการจำแนกภาพหน้าเอกสารออกเป็น 2 ส่วน คือ ส่วนของข้อความและส่วนของภาพถ่าย ซึ่งโดยทั่วไปแล้วหน้าเอกสารอาจประกอบด้วยสองส่วนใหญ่ๆคือ ส่วนของข้อความและส่วนของภาพถ่าย เมื่อทำการจำแนกหน้าเอกสารได้แล้วก็จะส่งผลลัพธ์ที่ได้จากการจำแนกในแต่ละส่วนไปประมวลผลในขั้นตอนต่อไป สำหรับขั้นตอนและรายละเอียดต่างๆ ของการจำแนกส่วนของข้อความและภาพถ่าย ได้แสดงไว้ในบทที่ 4 ซึ่งสามารถสรุปการทำงานได้ 5 ขั้นตอน กระทำได้โดย

1.1 ข้อมูลภาพที่ใช้ในการทดลองเป็นข้อมูลภาพขนาด A4 ที่มีระดับความเข้ม 256 ระดับเทาที่ได้จากการสแกนเอกสารใดๆ ด้วยสแกนเนอร์ขาวดำที่มีรายละเอียด 300 จุดต่อนิ้ว

1.2 ข้อมูลภาพจะถูกแบ่งออกเป็นบล็อกเล็กๆ ขนาด 4 x 4 จุดภาพ จากบล็อกภาพขนาด 4 x 4 จุดภาพ กำหนดจุดสำหรับอ้างอิงขึ้นมา 4 จุด เพื่อใช้ในการตัดสินว่าแต่ละบล็อกเป็นบล็อกของข้อความหรือเป็นบล็อกของภาพถ่าย โดยการพิจารณาจากระดับเทาของข้อมูล 4 จุดภาพนั้น

1.3 พิจารณาว่าบล็อกใดเป็นข้อความหรือบล็อกใดเป็นภาพถ่าย โดยการเปรียบเทียบกับค่าเธรซโฮลด์ที่ใช้ในการจำแนกจำนวน 2ค่า คือ ค่าขอบเขตบน (*Upper Bar*) และค่าขอบเขตล่าง (*Lower Bar*) ซึ่งคำนวณหาได้จากสมการ (4.1) และ (4.2)

$$Upper\ Bar = Maximum\ Gray\ Level \times 0.9 \quad \text{.....(4.1)}$$

$$Lower\ Bar = Maximum\ Gray\ Level \times 0.1 \quad \text{.....(4.2)}$$

โดยการพิจารณาจากเงื่อนไข ถ้า (จุดใดจุดหนึ่งใน 4 จุดนั้น มากกว่าหรือเท่ากับค่า *Upper Bar*) หรือ (มีอย่างน้อย 2 จุดใน 4 จุดน้อยกว่าหรือเท่ากับค่า *Lower Bar*) แสดงว่าบล็อกนั้นเป็นบล็อกข้อความ นอกนั้นเป็นบล็อกของภาพถ่าย

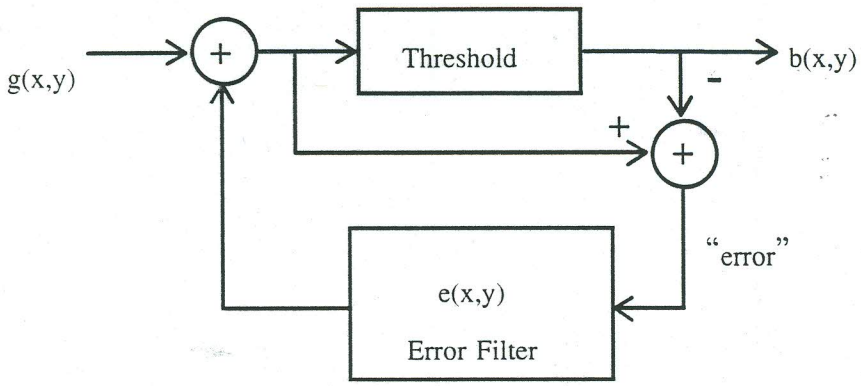
1.4 ทำการรวมกลุ่มบล็อกของภาพถ่ายเพื่อคำนวณหาพื้นที่ภาพถ่าย โดยทั่วไปแล้วพื้นที่ที่เป็นภาพถ่ายนั้นจะประกอบไปด้วยบล็อกของภาพถ่ายหลายๆบล็อกเรียงเป็นพื้นที่ติดต่อกัน จากเหตุผลนี้เราสามารถคำนวณหาพื้นที่ของภาพถ่ายได้โดยพิจารณาจากจำนวนบล็อกของภาพถ่าย โดยในการทดลองได้กำหนดให้ ถ้าจำนวนบล็อกของภาพถ่ายที่ติดกันมีจำนวนมากกว่า 50 บล็อก แสดงว่าเป็นพื้นที่เป็นภาพถ่าย นอกนั้นจะถูกกำหนดให้เป็นพื้นที่ของข้อความ

1.5 เก็บพิกัดของพื้นที่ภาพถ่ายที่ได้จากการจำแนก ขั้นตอนนี้เป็นขั้นตอนของการ วิเคราะห์ว่าหน้าเอกสารนั้ประกอบไปด้วยภาพถ่ายกี่ภาพ แต่ละภาพมีพิกัดหรือตำแหน่งอยู่ที่ใด โดยภาพถ่ายแต่ละภาพนั้น จะถูกทำการเก็บบันทึกตำแหน่งหรือพิกัดของพื้นที่ภาพถ่ายนั้นๆ โดยใช้เทคนิคการติดตามรอยขอบของภาพ เมื่อทำการจำแนกครบทั้งหน้าเอกสารแล้ว ก็จะนำพิกัดของจุดที่เป็นขอบภาพของแต่ละภาพมาคำนวณหาจุดพิกัดที่มีค่ามากที่สุดและน้อยที่สุดทั้งในแนวนอนและแนวตั้ง มากำหนดเป็นจุดมุมทั้งสี่ของรูปสี่เหลี่ยม

2. การสร้างภาพไบนารีโดยเทคนิคฮาฟท์โทน

การทำฮาฟท์โทน (Halftoning) เป็นวิธีการสำหรับแปลงภาพที่มีระดับความเข้มต่อเนื่อง (ภาพถ่าย) ให้เป็นภาพไบนารี ภาพที่ได้จากการทำฮาฟท์โทนนี้จะให้ความรู้สึกในการมองเห็นด้วยสายตาใกล้เคียงกับภาพเริ่มต้นคือดูคล้ายกับมีการไล่ระดับของสีเทา เทคนิคนี้เป็นเทคนิคที่ใช้กันมากในการแสดงผลภาพบนอุปกรณ์ที่สามารถแสดงผลได้ 2 ระดับ ขั้นตอนการทำฮาฟท์โทนนี้จะถูกนำมาใช้ในส่วนที่เมื่อทำการจำแนกส่วนที่เป็นข้อความและภาพถ่ายได้แล้ว ส่วนของภาพถ่ายจะถูกนำมาทำการสร้างภาพไบนารีโดยเทคนิคฮาฟท์โทนนี้

สำหรับในงานวิจัยนี้ได้เลือกใช้วิธีการกระจายความผิดพลาด ซึ่งวิธีนี้จะเป็นการสร้างภาพไบนารีโดยอาศัยจุดภาพรอบข้างมาช่วยในการพิจารณาว่าจุดภาพที่กำลังพิจารณาอยู่นั้นควรจะเป็นจุดขาวหรือจุดดำ โดยภาพไบนารีที่สร้างขึ้นจะให้ความรู้สึกในการมองเห็นด้วยสายตาของมนุษย์ได้ใกล้เคียงกับภาพเริ่มต้น โดยหลักการทำงานสามารถแสดงได้ดังภาพที่ 19



ภาพที่ 19 แสดงหลักการทำงานของวิธีการกระจายความผิดพลาด

หลักการทำงานวิธีการกระจายความผิดพลาด อธิบายได้ตามแผนภาพในภาพที่ 19 เมื่อ $g(x,y)$ คือข้อมูลภาพเริ่มต้น แต่ละพิกเซลมีช่วงระดับความเข้มต่อเนื่องตั้งแต่ 0 จนถึง L ระดับ เช่นถ้าภาพที่เป็นข้อมูลอินพุตเป็นภาพ 256 ระดับเทา จะมีระดับความเข้มตั้งแต่ 0 - 255 ถูกนำมาเปรียบเทียบกับเธรชโชลด์ (Threshold) ซึ่งเป็นค่าคงที่เท่ากับ $\frac{L}{2}$ เพื่อตัดสินใจให้กำหนดค่าเอาต์พุตสำหรับภาพผลลัพธ์ $b(x,y)$ เป็น 0 (จุดดำ) หรือ 1 (จุดขาว) เพียง 2 ระดับเท่านั้น ความแตกต่างระหว่างข้อมูลทั้งสองเรียกว่า "ค่าความผิดพลาด" (Error) ที่เกิดขึ้นจากการทำเธรชโชลด์ จะถูกนำมากรองความผิดพลาด เพื่อกระจายค่าความผิดพลาดนี้ให้กับจุดภาพที่อยู่รอบข้างจำนวนหนึ่ง ซึ่งอยู่ถัดจากจุดภาพที่กำลังพิจารณาเทียบกับทิศทางของการประมวลผลด้วยสัดส่วนที่ไม่เท่ากัน ซึ่งข้อมูลที่ได้จากการกรองความผิดพลาดนี้คือ $e(x,y)$

บนพื้นฐานของการทำงานของวิธีนี้ สิ่งสำคัญที่ทำให้ภาพผลลัพธ์ที่ได้ดูเหมือนมีการไล่ระดับก็คือ "ตัวกรองความผิดพลาด" (Error Filter) ซึ่งในงานวิจัยนี้ได้เลือกใช้รูปแบบตัวกรองความผิดพลาดของ Floyd and Steinberg ซึ่งจะใช้การกรองด้วยตัวถ่วงน้ำหนักจำนวน 4 ตัว ซึ่งจะทำให้การประมวลผลรวดเร็วขึ้น สามารถแสดงรูปแบบตัวกรองความผิดพลาดของ Floyd และ Steinberg ดังภาพที่ 20 ตัวเลขต่างๆ แสดงถึงค่าถ่วงน้ำหนัก (Weight) ที่จะนำมาคูณกับค่าความผิดพลาดและกระจายให้กับแต่ละพิกเซล จะสังเกตได้ว่าผลรวมของตัวถ่วงน้ำหนักทั้งหมดเท่ากับ 1 เสมอ แสดงว่าไม่มีการเพิ่มหรือลดค่าความผิดพลาดที่จะกระจายให้กับพิกเซลรอบข้าง

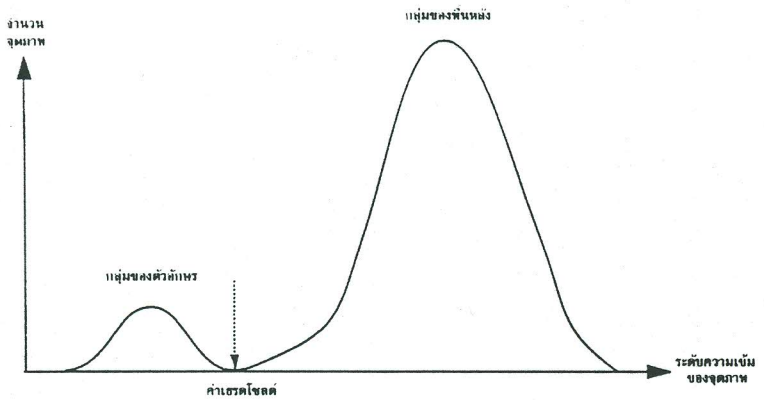
	●	7
3	5	1

$\times \frac{1}{16}$

ภาพที่ 20 แสดงรูปแบบตัวกรองความผิดพลาด โดยวิธีของ Floyd and Steinberg

3. การสร้างภาพไบนารีโดยเทคนิคเธรชโชลด์

การทำเธรชโชลด์ (Thresholding) เป็นวิธีการแปลงภาพที่มีหลายระดับเทาให้เป็นภาพแบบไบนารี การทำเธรชโชลด์นี้เป็นขั้นตอนสำหรับการสร้างภาพเอกสารในส่วนของข้อความให้อยู่ในรูปของภาพไบนารี อัลกอริธึมสำหรับการคำนวณหาค่าเธรชโชลด์ของภาพเอกสารนั้นมีหลายวิธี สำหรับงานวิจัยนี้จะใช้วิธีคำนวณหาค่าเธรชโชลด์ โดยการพิจารณาจากฮิสโตแกรมระดับเทาของภาพ เนื่องจากวิธีนี้จะมีความยืดหยุ่นมากในการคำนวณค่าเธรชโชลด์ สำหรับแนวความคิดในการคำนวณหาค่าเธรชโชลด์โดยวิธีนี้ แสดงดังภาพที่ 21



ภาพที่ 21 แสดงแนวความคิดในการคำนวณหาค่าเธรชโชลด์ โดยการพิจารณาจากฮิสโตแกรม

โดยหลักการทำงานจะเริ่มจากข้อมูลภาพหน้าเอกสารในส่วนของข้อความที่ได้จากการจำแนก นำมาสร้างเป็นฮิสโตแกรมระดับเทา จากฮิสโตแกรมจะเห็นว่าเกิดกลุ่มของระดับเทา (Peak) 2 กลุ่ม กลุ่มหนึ่งคือกลุ่มของข้อความหรือวัตถุ (Object) และอีกกลุ่มหนึ่งคือ กลุ่มของพื้นหลัง (Background) จากนั้นจะเลือกค่าที่ต่ำที่สุด (Valley) ที่อยู่ระหว่าง 2 จุดนั้น เป็นค่าเธรชโชลด์ ซึ่งค่าเธรชโชลด์นี้จะใช้ในการแบ่งว่าจุดภาพแต่ละจุดๆใดควรจะเป็นจุดขาวหรือจุดใดควรจะเป็นจุดดำ เมื่อได้ค่าเธรชโชลด์แล้วสามารถสร้างภาพไบนารีได้ โดยจะพิจารณาว่าระดับ

ความเข้มใดๆของข้อมูลภาพอินพุต $(g(x,y))$ จะมีค่าความเข้มสูงกว่าค่าเธรชโชลด์หรือไม่ หากมีค่าน้อยกว่าก็จะให้ค่าเป็น "1" (จุดดำ) นอกนั้นจะให้เป็นค่า "0" (จุดขาว) ในภาพไบนารีเอาต์พุต $(b(x,y))$ ดังสมการที่ (4.3)

$$b(x,y)=\begin{cases}1; & g(x,y)<Thr\\ 0; & g(x,y)\geq Thr\end{cases} \quad \text{.....(4.3)}$$

4. การรวมส่วนของข้อความและภาพถ่ายที่เป็นภาพไบนารีเข้าด้วยกัน

ขั้นตอนนี้เป็นขั้นตอนของการรวมส่วนของข้อความและส่วนของภาพถ่ายที่ทำการสร้างเป็นภาพแบบไบนารีแล้วโดยวิธีที่กล่าวมาข้างต้นในแต่ละส่วนเข้ามาไว้ด้วยกัน จะได้เป็นภาพหน้าเอกสารแบบไบนารีที่เหมาะสมใกล้เคียงกับภาพเอกสารต้นฉบับ

การทดลอง

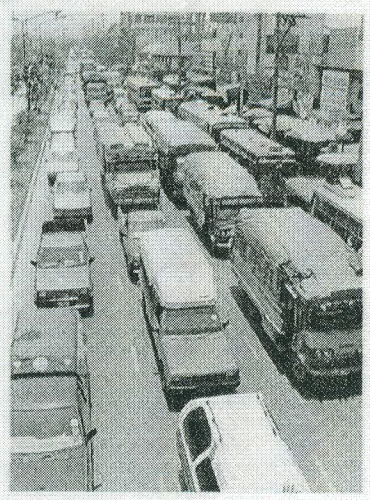
ในการทดลองสร้างภาพไบนารีสำหรับภาพหน้าเอกสาร ได้ทำการทดลองกับภาพหน้าเอกสารที่ได้จากเครื่องสแกนเนอร์ขาวดำที่มีความละเอียด 300 จุดต่อนิ้ว ที่มีระดับความเข้ม 256 ระดับเทา ได้ทำการทดลองบนเครื่องไมโครคอมพิวเตอร์ รุ่น 486DX2-50 โดยแบ่งการทดลองออกเป็นสองส่วนคือ การสร้างภาพไบนารีคุณภาพสูงจากวิธีการที่นำเสนอ และอีกส่วนหนึ่งจะเป็นการศึกษาเปรียบเทียบผลลัพธ์ที่ได้ระหว่างวิธีการที่นำเสนอและเทคนิคการสร้างภาพไบนารีที่นิยมใช้กันโดยทั่วไป

1. การสร้างภาพไบนารีจากวิธีการที่นำเสนอ

การทดลองในครั้งนี้จะเป็นการทดลองนำภาพหน้าเอกสารที่มีระดับความเข้ม 256 ระดับเทามาสร้างเป็นภาพไบนารีโดยวิธีการที่นำเสนอข้างต้น ปรากฏว่าให้ผลลัพธ์เป็นที่น่าพอใจ กล่าวคือภาพผลลัพธ์ที่ได้ในส่วนที่เป็นข้อความนั้นจะมีความคมชัดของตัวหนังสือ และส่วนที่เป็นภาพถ่ายจะดูเหมือนมีการไล่ระดับของสีเทาได้ดีทำให้ภาพที่ได้มีความสวยงามเป็นที่น่าพอใจและมีคุณภาพใกล้เคียงกับภาพต้นฉบับ ดังแสดงผลจากการทดลองได้ในภาพที่ 23

1. แก้ปัญหาจราจรให้คลายจลาจล

ด้วยการดำเนินโครงการทั้ง ๕ ให้เสร็จสิ้นในเวลารวดเร็ว คือ ถนนวงแหวนรอบกลาง-รอบนอก และถนนใยแมงมุม (2-3 ปี) ทางด่วนพิเศษให้ครบทั้ง 5 ขั้นตอน (3 ปี) รถขนส่งมวลชนทั้งติดดินยกระดับและใต้ดิน (4-6 ปี) รถไฟชานเมืองทั้ง 3 ทิศของ กทม. (3 ปี) และเมืองบริวารรอบกรุงเทพ 20 เมือง (3 ปี)



2. ดูแลเรื่องปัญหาค่าครองชีพ

การปรับรายได้ให้ผู้ที่มีเงินเดือน ต้องทำโดยวิธีเดียวกันกับการขึ้นภาษี คือโดยพระราชกำหนดและใช้เงินคงคลังจ่ายทันที ในขณะที่ไม่เอาเรื่องการปรับเงินเดือนมาหาเสียงกับประชาชนกันทั้งปีที่ทำให้สินค้าขึ้นราคาหลายหนกว่าจะได้เงินเดือนจริง

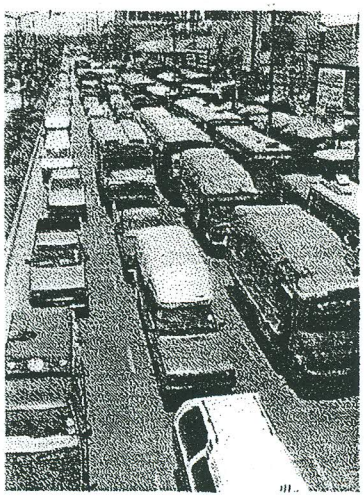
3. เร่งรีบดำเนินการเรื่อง งานเกษตร-อุตสาหกรรม

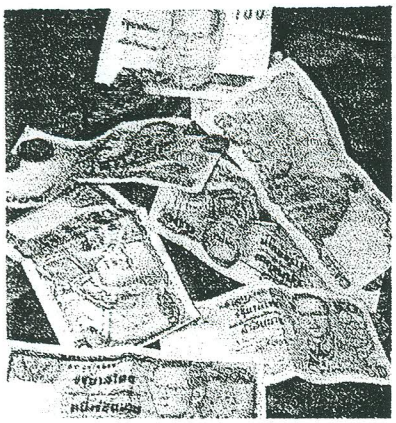
จนถึงวันนี้เราเอาสินค้าพืชผลการเกษตร ไปส่งออกไม่ได้แล้ว เพราะทั้งราคา คุณภาพ และการขนส่ง เป็นปัญหา อุตสาหกรรมของเราเดินหน้าไปไกลแล้ว และการแปรรูปสินค้าพืชผลเกษตรเป็นอาหารสัตว์ไว้เลี้ยงสัตว์และนำมาเป็นวัตถุดิบสำหรับอุตสาหกรรมอาหาร และพืชผลอื่นที่แปรสภาพใส่กระป๋องเพื่อการส่งออกเป็นอาหารสำหรับเลี้ยงคนในโลก

ภาพที่ 22 แสดงภาพหน้าเอกสารต้นฉบับที่ใช้ในการสร้างภาพไบนารีสำหรับภาพเอกสาร

1. แก้ปัญหาจราจรให้คลายจลาจล

ด้วยการดำเนินโครงการทั้ง ๕ ให้เสร็จสิ้นในเวลารวดเร็ว คือ ถนนวงแหวนรอบกลาง-รอบนอก และถนนใยแมงมุม (2-3 ปี) ทางด่วนพิเศษให้ครบทั้ง 5 ขั้นตอน (3 ปี) รถขนส่งมวลชนทั้งติดดินยกระดับและใต้ดิน (4-6 ปี) รถไฟชานเมืองทั้ง 3 ทิศของ กทม. (3 ปี) และเมืองบริวารรอบกรุงเทพ 20 เมือง (3 ปี)





2. ดูแลเรื่องปัญหาค่าครองชีพ

การปรับรายได้ให้ผู้ที่มีเงินเดือน ต้องทำโดยวิธีเดียวกันกับการขึ้นภาษี คือโดยพระราชกำหนดและใช้เงินคงคลังจ่ายทันที ในขณะที่ไม่เอาเรื่องการปรับเงินเดือนมาหาเสียงกับประชาชนกันทั้งปีที่ทำให้สินค้าขึ้นราคาหลายหนกว่าจะได้เงินเดือนจริง

3. เร่งรีบดำเนินการเรื่อง งานเกษตร-อุตสาหกรรม

จนถึงวันนี้เราเอาสินค้าพืชผลการเกษตร ไปส่งออกไม่ได้แล้ว เพราะทั้งราคา คุณภาพ และการขนส่ง เป็นปัญหา อุตสาหกรรมของเราเดินหน้าไปไกลแล้ว และการแปรรูปสินค้าพืชผลเกษตรเป็นอาหารสัตว์ไว้เลี้ยงสัตว์และนำมาเป็นวัตถุดิบสำหรับอุตสาหกรรมอาหาร และพืชผลอื่นที่แปรสภาพใส่กระป๋องเพื่อการส่งออกเป็นอาหารสำหรับเลี้ยงคนในโลก

ภาพที่ 23 แสดงภาพหน้าเอกสารแบบไบนารีที่ได้จากวิธีการที่นำเสนอ

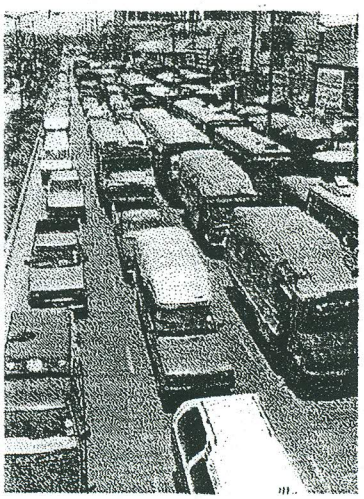
2. การศึกษาเปรียบเทียบกับการสร้างภาพไบนารีเทคนิคอื่น

ในการศึกษาเปรียบเทียบกับเทคนิคการสร้างภาพไบนารีกับเทคนิคการสร้างภาพไบนารีเทคนิคอื่นที่นิยมใช้กันโดยทั่วไปนี้ โดยในการเปรียบเทียบครั้งนี้ได้ทำการเปรียบเทียบระหว่างผลลัพธ์จากวิธีที่นำเสนอกับเทคนิคเธรชโชลด์และเทคนิคฮาฟท์โทน เพื่อที่จะเปรียบเทียบผลลัพธ์ที่ได้จากวิธีที่นำเสนอและเทคนิคทั้งสอง สำหรับภาพผลลัพธ์ที่ได้จากการสร้างภาพไบนารีจากวิธีการทั้ง 3 สามารถแสดงผลการเปรียบเทียบได้ดังภาพที่ 24 - 27

ภาพที่ 24 แสดงภาพหน้าเอกสารต้นฉบับที่มีระดับความเข้ม 256 ระดับเทา

1. แก้ปัญหาจราจรให้คลายจลาจล

ด้วยการดำเนินโครงการทั้ง ๕ ให้เสร็จสิ้นในเวลารวดเร็ว คือ ถนนวงแหวนรอบกลาง-รอบนอก และถนนใยแมงมุม (2-3 ปี) ทางด่วนพิเศษให้ครบทั้ง 5 ขั้นตอน (3 ปี) รถขนส่งมวลชนทั้งติดดินยกระดับและใต้ดิน (4-6 ปี) รถไฟชานเมืองทั้ง 3 ทิศของ กทม. (3 ปี) และเมืองบริวารรอบกรุงเทพ 20 เมือง (3 ปี)



2. ดูแลเรื่องปัญหาค่าครองชีพ

การปรับรายได้ให้ผู้ที่มีเงินเดือน ต้องทำโดยวิธีเดียวกันกับการขึ้นภาษี คือโดยพระราชกำหนดและใช้เงินคงคลังจ่ายทันที ในขณะที่ไม่เอาเรื่องการปรับเงินเดือนมาหาเสียงกับประชาชนกันทั้งปีที่ทำให้สินค้าขึ้นราคาหลายหนกว่าจะได้เงินเดือนจริง

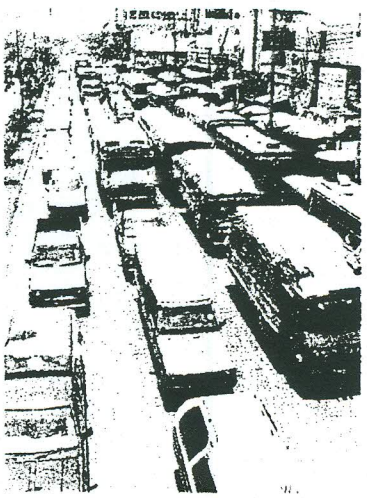
3. เร่งรีบดำเนินการเรื่อง งานเกษตร-อุตสาหกรรม

จนถึงวันนี้เราเอาสินค้าพืชผลการเกษตร ไปส่งออกไม่ได้แล้ว เพราะทั้งราคา คุณภาพ และการขนส่ง เป็นปัญหา อุตสาหกรรมของเราเดินหน้าไปไกลแล้ว และการแปรรูปสินค้าพืชผลเกษตรเป็นอาหารสัตว์ไว้เลี้ยงสัตว์และนำมาเป็นวัตถุดิบสำหรับอุตสาหกรรมอาหาร และพืชผลอื่นที่แปรสภาพใส่กระป๋องเพื่อการส่งออกเป็นอาหารสำหรับเลี้ยงคนในโลก

ภาพที่ 25 แสดงภาพหน้าเอกสารแบบไบนารีที่ได้จากวิธีการที่นำเสนอ

1. แก้ปัญหาจราจรให้คลายจลาจล

ด้วยการดำเนินโครงการทั้ง ๕ ให้เสร็จสิ้นในเวลารวดเร็ว คือ ถนนวงแหวนรอบกลาง-รอบนอก และถนนใยแมงมุม (2-3 ปี) ทางด่วนพิเศษให้ครบทั้ง 5 ขั้นตอน (3 ปี) รถขนส่งมวลชนทั้งติดดินยกระดับและใต้ดิน (4-6 ปี) รถไฟชานเมืองทั้ง 3 ทิศของ กทม. (3 ปี) และเมืองบริวารรอบกรุงเทพ 20 เมือง (3 ปี)



2. ดูแลเรื่องปัญหาค่าครองชีพ

การปรับรายได้ให้ผู้ที่มีเงินเดือน ต้องทำโดยวิธีเดียวกันกับการขึ้นภาษี คือโดยพระราชกำหนดและใช้เงินคงคลังจ่ายทันที ในขณะที่ไม่เอาเรื่องการปรับเงินเดือนมาหาเสียงกับประชาชนกันทั้งปีที่ทำให้สินค้าขึ้นราคาหลายหนกว่าจะได้เงินเดือนจริง

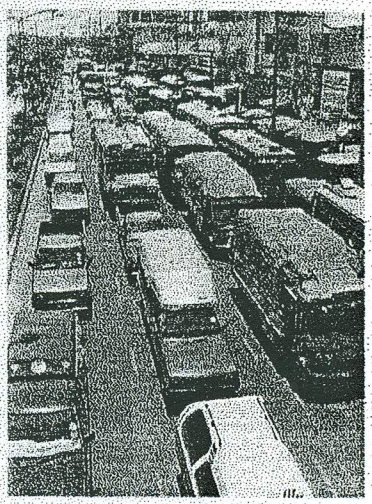
3. เร่งรีบดำเนินการเรื่อง งานเกษตร-อุตสาหกรรม

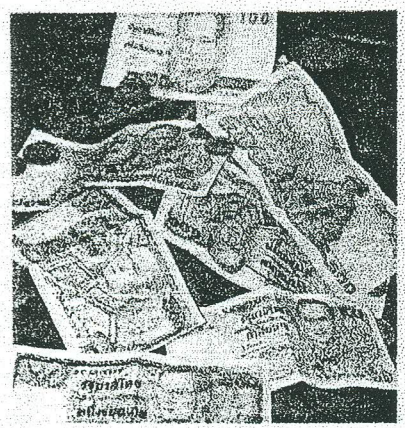
จนถึงวันนี้เราเอาสินค้าพืชผลการเกษตร ไปส่งออกไม่ได้แล้ว เพราะทั้งราคา คุณภาพ และการขนส่ง เป็นปัญหา อุตสาหกรรมของเราเดินหน้าไปไกลแล้ว และการแปรรูปสินค้าพืชผลเกษตรเป็นอาหารสัตว์ไว้เลี้ยงสัตว์และนำมาเป็นวัตถุดิบสำหรับอุตสาหกรรมอาหาร และพืชผลอื่นที่แปรสภาพใส่กระป๋องเพื่อการส่งออกเป็นอาหารสำหรับเลี้ยงคนในโลก

ภาพที่ 26 แสดงภาพหน้าเอกสารแบบไบนารีที่ได้จากเทคนิคเธรชโชลด์

1. แก้ปัญหาจราจรให้คลายจลาจล

ด้วยการดำเนินโครงการทั้ง ๕ ให้เสร็จสิ้นในเวลารวดเร็ว คือ ถนนวงแหวนรอบกลาง-รอบนอก และถนนใยแมงมุม (2-3 ปี) ทางด่วนพิเศษให้ครบทั้ง 5 ขั้นตอน (3 ปี) รถขนส่งมวลชนทั้งติดดินยกระดับและใต้ดิน (4-6 ปี) รถไฟชานเมืองทั้ง 3 ทิศของ กทม. (3 ปี) และเมืองบริวารรอบกรุงเทพ 20 เมือง (3 ปี)





2. ดูแลเรื่องปัญหาค่าครองชีพ

การปรับรายได้ให้ผู้ที่มีเงินเดือน ต้องทำโดยวิธีเดียวกันกับการขึ้นภาษี คือโดยพระราชกำหนดและใช้เงินคงคลังจ่ายทันที ในขณะที่ไม่เอาเรื่องการปรับเงินเดือนมาหาเสียงกับประชาชนกันทั้งปีที่ทำให้สินค้าขึ้นราคาหลายหนกว่าจะได้เงินเดือนจริง

3. เร่งรีบดำเนินการเรื่อง งานเกษตร-อุตสาหกรรม

จนถึงวันนี้เราเอาสินค้าพืชผลการเกษตร ไปส่งออกไม่ได้แล้ว เพราะทั้งราคา คุณภาพ และการขนส่ง เป็นปัญหา อุตสาหกรรมของเราเดินหน้าไปไกลแล้ว และการแปรรูปสินค้าพืชผลเกษตรเป็นอาหารสัตว์ไว้เลี้ยงสัตว์และนำมาเป็นวัตถุดิบสำหรับอุตสาหกรรมอาหาร และพืชผลอื่นที่แปรสภาพใส่กระป๋องเพื่อการส่งออกเป็นอาหารสำหรับเลี้ยงคนในโลก

ภาพที่ 27 แสดงภาพหน้าเอกสารแบบไบนารีที่ได้จากเทคนิคฮาฟท์โทน

จากการทดลองเปรียบเทียบการสร้างภาพไบนารีโดยวิธีที่นำเสนอกับเทคนิคเธรชโชลด์และเทคนิคฮาฟท์โทน ในภาพที่ 24 - 27 จะพบว่าภาพที่ได้จากวิธีที่นำเสนอจะมีความคมชัดสวยงามและมีคุณภาพของภาพผลลัพธ์มากกว่าเทคนิคอื่น กล่าวคือ มีคุณภาพใกล้เคียงกับภาพต้นฉบับ โดยภาพที่ 24 แสดงบางส่วนของภาพหน้าเอกสารต้นฉบับ รูปที่ 25 แสดงบางส่วนของภาพหน้าเอกสารแบบไบนารีโดยเทคนิคฮาฟท์โทน วิธีการกระจายความผิดพลาด จะเห็นได้ว่าในส่วนของภาพถ่ายที่อยู่บนหน้าเอกสารนั้น ดูสวยงามเหมือนกับมีการไล่ระดับสีเทา แต่สำหรับในส่วนที่เป็นตัวอักษรนั้นจะเห็นได้ว่ามีการกระจายจุดดำอยู่ทั่วไป ซึ่งเมื่อนำไปประยุกต์ใช้กับระบบประยุกต์อื่นเช่น ระบบการรู้จำตัวอักษร แล้ว จุดดำที่ปรากฏเหล่านี้จะถือว่าเป็นสิ่งรบกวน (Noise) ที่เกิดขึ้น อีกทั้งตัวอักษรที่ปรากฏก็ไม่คมชัดเท่าที่ควร รูปที่ 26 แสดงบางส่วนของภาพหน้าเอกสารแบบไบนารีโดยวิธีเธรชโชลด์ โดยกำหนดค่าเธรชโชลด์เท่ากับ 128 จะเห็นได้ว่าส่วนของตัวอักษรจะมีความคมชัด และไม่มีสิ่งรบกวนเกิดขึ้น แต่สำหรับส่วนที่เป็นภาพถ่ายนั้นไม่สวยงามเท่าที่ควร และไม่มีการไล่ระดับสีเทา และรูปที่ 27 แสดงบางส่วนของภาพหน้าเอกสารแบบไบนารี โดยวิธีการที่นำเสนอ จะเห็นได้ว่าในส่วนของตัวอักษรมีความคมชัด ไม่มีสิ่งรบกวนและส่วนของภาพถ่ายนั้นดูเหมือนมีการไล่ระดับสีเทาและภาพที่ได้มีความสวยงาม

บทสรุป

ในการสร้างภาพไบนารีสำหรับภาพเอกสารที่นำเสนอในบทนี้ เป็นการสร้างภาพไบนารีวิธีหนึ่งที่ได้นำข้อดีของการสร้างภาพไบนารีโดยเทคนิคฮาฟท์โทน ที่มีความสามารถในการสร้างภาพไบนารีได้สวยงามดูเหมือนมีการไล่ระดับของสีเทาได้ดีเหมาะสำหรับภาพที่เป็นภาพถ่ายและเทคนิคเธรชโชลด์ที่มีความสามารถในการสร้างภาพไบนารีที่เหมาะกับภาพที่เป็นข้อความบนพื้นกระดาษที่สีอ่อนกว่า โดยการทำงานจะเริ่มจากการจำแนกภาพหน้าเอกสารออกเป็น 2 ส่วนคือ ส่วนที่เป็นข้อความและส่วนที่เป็นภาพถ่าย จากนั้นนำแต่ละส่วนไปสร้างภาพไบนารีด้วยวิธีที่แตกต่างกัน กล่าวคือ ส่วนที่เป็นข้อความใช้เทคนิคเธรรชโชลด์และส่วนที่เป็นภาพถ่ายจะใช้เทคนิคฮาฟท์โทน จากนั้นจึงนำทั้ง 2 ส่วนมารวมกันจะได้เป็นภาพไบนารีที่สวยงาม โดยภาพหน้าเอกสารแบบไบนารีที่ได้จะมีความสวยงามเหมาะสมและมีคุณภาพใกล้เคียงกับภาพต้นฉบับมากกว่าการสร้างภาพไบนารีโดยใช้เทคนิคเธรชโชลด์หรือเทคนิคฮาฟท์โทนเพียงอย่างเดียว

บทที่ 5

สรุปผลและข้อเสนอแนะ

สรุปผลการวิจัย

ในการวิจัยครั้งนี้ได้ทำการศึกษาและวิจัยเกี่ยวกับการสร้างภาพไบนารีสำหรับภาพหน้าเอกสาร เพื่อที่จะได้ภาพไบนารีที่มีความเหมาะสมสวยงามและมีคุณภาพใกล้เคียงกับภาพหน้าเอกสารต้นฉบับ ในการสร้างภาพไบนารีโดยวิธีที่นำเสนอนี้จะเป็นการนำเอาจุดเด่นของการสร้างภาพไบนารีโดยเทคนิคฮาฟท์โทนและเทคนิคเธรชโชลด์มาประยุกต์ใช้ร่วมกัน สำหรับการวิจัยในครั้งนี้ได้แบ่งขั้นตอนของการวิจัยออกเป็น 2 ส่วนด้วยกัน คือ

1. การจำแนกข้อความและภาพถ่ายในหน้าเอกสาร

ในการจำแนกข้อความและภาพถ่ายในหน้าเอกสาร จะเป็นขั้นตอนที่สำคัญในการสร้างภาพไบนารีสำหรับภาพหน้าเอกสาร เนื่องจากว่าโดยทั่วไปแล้วหน้าเอกสารหนึ่งหน้านั้นจะประกอบไปด้วย 2 ส่วนหลักๆคือ ส่วนที่เป็นข้อความ และส่วนที่เป็นภาพถ่าย ดังนั้นในขั้นตอนนี้จึงได้ทำการจำแนกภาพหน้าเอกสารออกเป็นสองส่วน เพื่อที่จะได้นำผลลัพธ์ในแต่ละส่วนไปประมวลผลในขั้นตอนต่อไป ในการทดลองจะพบว่า ผลลัพธ์ที่ได้จะเป็นการคำนวณหาพิกัดหรือขอบเขตสี่เหลี่ยมของภาพถ่ายที่ปรากฏในหน้าเอกสาร ส่วนบริเวณที่เหลือนั้นจะเรียกว่าเป็นบริเวณที่เป็นพื้นที่ข้อความ ในการประมวลผลนั้นจะพิจารณาหาพื้นที่ภาพถ่ายโดยใช้บล็อกเล็กๆ ขนาด 4x4 ร่วมกับการพิจารณาจากจำนวนของบล็อกภาพถ่ายที่ได้จากการจำแนก ส่วนการคำนวณหาพิกัดมุมทั้ง 4 ของภาพถ่ายนั้นได้นำหลักการของการติดตามรอยขอบของภาพมาใช้

2. การสร้างภาพไบนารีสำหรับภาพเอกสาร

ในส่วนนี้จะเป็นการสร้างภาพไบนารีสำหรับภาพเอกสาร เป็นการสร้างภาพไบนารีวิธีหนึ่งที่ได้นำข้อดีของการสร้างภาพไบนารีโดยเทคนิคฮาฟท์โทน ที่มีความสามารถในการสร้างภาพไบนารีได้สวยงามดูเหมือนมีการไล่ระดับของสีเทาได้ดีเหมาะสำหรับภาพที่เป็นภาพถ่ายและเทคนิคเธรชโชลด์ ที่มีความสามารถในการสร้างภาพไบนารีที่เหมาะกับภาพที่เป็นข้อความ

บนพื้นกระดาษที่สีอ่อนกว่า โดยการทำงานจะเริ่มจากการจำแนกภาพหน้าเอกสารออกเป็น 2 ส่วนคือ ส่วนที่เป็นข้อความและส่วนที่เป็นภาพถ่าย จากนั้นนำแต่ละส่วนไปสร้างภาพไบนารีด้วยวิธีที่แตกต่างกัน กล่าวคือ ส่วนที่เป็นข้อความใช้เทคนิคเธรรชโซลด์และส่วนที่เป็นภาพถ่ายจะใช้เทคนิคฮาฟท์โทน จากนั้นจึงนำทั้ง 2 ส่วนมารวมกันจะได้เป็นภาพไบนารีที่สวยงาม โดยภาพหน้าเอกสารแบบไบนารีที่ได้จะมีความสวยงามเหมาะสมและมีคุณภาพใกล้เคียงกับภาพหน้าเอกสารต้นฉบับมากกว่าการสร้างภาพไบนารีโดยใช้เทคนิคเธรรชโซลด์หรือเทคนิคฮาฟท์โทนเพียงอย่างเดียว

ปัญหาที่พบและข้อเสนอแนะ

สำหรับปัญหาที่พบในการสร้างภาพไบนารีคุณภาพสูงสำหรับภาพเอกสารหลายองค์ประกอบก็คือ ปัญหาเรื่องข้อมูลภาพอินพุต ในที่นี้ก็คือภาพหน้าเอกสารที่ได้จากการเครื่องสแกนเนอร์ เนื่องจากหน้าเอกสารที่ใช้เป็นหน้าเอกสาร 1 หน้ากระดาษ ซึ่งโดยทั่วๆไปนิยมใช้กระดาษขนาด A4 มีขนาดกว้างประมาณ 8 นิ้วและยาวประมาณ 11 นิ้ว ในกรณีนี้ข้อมูลภาพจะเป็นข้อมูลภาพ 256 ระดับเทา นั่นก็หมายถึง 1 ไบท์ (Byte) ก็คือ 1 จุดภาพ และข้อมูลภาพนั้นจะมีความละเอียด 300 จุดต่อนิ้ว ดังนั้นภาพหน้าเอกสารหนึ่งภาพจะประกอบไปด้วยจุดภาพเท่ากับ 2,400 x 3,300 จุดภาพ ทำให้ภาพที่ใช้เป็นข้อมูลอินพุตจะมีขนาดประมาณ 8 เมกกะไบต์ จากขนาดของภาพที่ใหญ่นี้เองทำให้การประมวลผลต้องใช้เวลามาก ซึ่งการแก้ปัญหาในเรื่องของเวลาที่ใช้ในการประมวลผลสามารถทำให้ได้โดย การเพิ่มประสิทธิภาพทางฮาร์ดแวร์ให้เครื่องคอมพิวเตอร์ ไม่ว่าจะเป็นการเพิ่มหน่วยความจำ (RAM) หรือการเลือกใช้หน่วยประมวลผลกลาง (CPU) ที่มีความเร็วในการประมวลผลที่เร็วขึ้น จากการที่ข้อมูลภาพอินพุตมีขนาดใหญ่นี้เองทำให้การพัฒนาโปรแกรมบนระบบปฏิบัติการ DOS มีความยุ่งยากและมีข้อจำกัดในการจัดการหน่วยความจำเพื่อใช้ในการประมวลผล ดังนั้นผู้วิจัยจึงได้พัฒนาโปรแกมบนระบบปฏิบัติการ WINDOWS แทน ซึ่งระบบนี้จะไม่มีปัญหาเรื่องของหน่วยความจำดังกล่าว เนื่องจากมีระบบปฏิบัติการนี้จะจัดการหน่วยความจำให้ เราไม่จำเป็นต้องเขียนโปรแกรมในการจัดการหน่วยความจำเอง จึงทำให้การพัฒนาโปรแกรมดังกล่าวง่ายขึ้น

สำหรับแนวทางในการพัฒนาขั้นต่อไปของการสร้างภาพไบนารีโดยวิธีนี้นั้น เนื่องจากว่าภาพหน้าเอกสารหนึ่งหน้านั้นมีขนาดใหญ่มาก ทำให้การประมวลผลโดยใช้ซอฟท์แวร์จะต้องใช้เวลาในการประมวลผลมาก ดังนั้นจึงน่าที่จะประยุกต์ระบบดังกล่าวในเชิงของฮาร์ดแวร์ จะทำให้การประมวลผลใช้เวลาในการประมวลผลน้อยลง

เอกสารอ้างอิง

[1] ดร.ยุทธพงษ์ รังสรรค์เสรี และสิทธิชัย บุษหมั่น, "การจำแนกข้อความและภาพถ่ายอัตโนมัติ", *วารสารพระจอมเกล้าลาดกระบัง*, ปีที่ 3, ฉบับที่ 2 (กันยายน 2538) หน้า 18-22.

[2] สาธิต อินทจักร์. *การหาขอบภาพโดยใช้แบบจำลองฟัซซี่และนิวรอลเน็ตเวิร์ค*, วิทยานิพนธ์วิศวกรรมศาสตร์มหาบัณฑิต. กรุงเทพฯ : บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง. 2537.

[3] สมศักดิ์ วลัยลักษณ์. *การวิเคราะห์และระบุส่วนประกอบหน้าเอกสารและการรู้จำตัวอักษร*. วิทยานิพนธ์ วิศวกรรมศาสตร์มหาบัณฑิต. กรุงเทพฯ : บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง. 2537.

[4] A. Low. *INTRODUCTORY COMPUTER VISION AND IMAGE PROCESSING.* New York : McGraw-Hill Inc. 1991.

[5] H.S. Baird, *Structured Document Image Analysis.* New York : Springer-Verlag, 1992.

[6] H.R. Myler and A.R. Weeks., *Computer Imaging Recipes in C*, Prentice Hall, 1993.

[7] HEWLETT-PACKARD. *A Guide to the Tag Image File Format Version. 5.0.* Hewlette-Packard Company.

[8] K. Y. Wong, R. G. Casey and F. M. Wahl. "Document Analysis System." *IBM J. Res. Develop*, no. 6 (1982) PP. 647-656.

[9] Lawrence O'Gorman and Rangachar Kasturi. *DOCUMENT IMAGE ANALYSIS.* Los Alamitos : IEEE Computer Society Press, 1995.

[10] N. Otsu, "A Threshold Selection Method From Gray-Level Histograms." *IEEE Trans. Systems Man Cybernet,* SMC-9 (1979) PP.62-66.

[11] P. Ioannis., *Digital Image Processing Algorithms.*, Prentice Hall., 1993.

[12] R.C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley, 1992.

[13] R. Jain, R. Kasturi and B.G. Schunck, *MACHINE VISION*. Singapore. McGraw-Hill Inc. 1995.

[14] R. Ulichney. *Digital halftoning.* London : MIT Press, 1987.

[15] S. Rimmer. *WINDOWS BITMAPPED GRAPHICS.* New York : McGraw-Hill Inc. 1992.

[16] S. Rimmer. *Bit-Mapped Graphics.* 2nd ed. New York : McGraw-Hill Inc. 1993.

[17] S. Rimmer. *Supercharged Bitmapped Graphics.* New York : McGraw-Hill Inc. 1992.

[18] Schurmann, Jurgen and others., "Document Analysis-From Pixels to Contents" *Proceeding of The IEEE*, Vol. 80 no. 7 (1992) PP. 1101-1119.

[19] Y. Chen, F.C. Mintzer and K.S. Pennington, "PANDA : Processing Algorithm for Noncoded Document Acquisition." *IBM J. Res. Develop*, no. 1 (1987) PP. 32-43.