การรู้จำตัวอักษรพิมพ์ภาษาอังกฤษบนขอบเป็นสันหนังสือด้วยกล้อง CCD

Printed English Character on Edge book Recognition by CCD camera

โดย

นาย ทศพร พันภัยพาล

นาย นพพร พัฒนสุวรรณ

วัน เดือน ปี -5 ต.ค. 2541

เลขทะเบียน 038529

เลขเรียกหนังสือ T 40076 ท257ก

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตร์บัณฑิต

สาขาวิศวกรรมการวัดคุมทางอุตสาหกรรม

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2540

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

038529

การรู้จำตัวอักษรพิมพ์ภาษาอังกฤษบนสันหนังสือด้วยกล้อง CCD

Printed English Character on Edge book Recognition by CCD camera

จัดทำโดย

นาย ทศพร พันภัยพาล

นาย นพพร พัฒนสุวรรณ

อาจารย์ที่ปรึกษา

ผศ. เกษตร์ ศิริสันติสัมฤทธิ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์ปีการศึกษา 2540

ภาควิชา เทคโนโลยีการวัดคุมทางอุตสาหกรรม

สาขาวิชา วิศวกรรมการวัดคุมทางอุตสากรรม

คณะ วิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การรู้จำตัวอักษรพิมพ์ภาษาอังกฤษบนสันหนังสือด้วยกล้อง CCD

(Printed English Character on edge book Recognition by CCD camera)

ผู้จัดทำ

1. นาย ทศพร พันภัยพาล เลขประจำตัว 38012095

2. นาย นพพร พัฒนสุวรรณ เลขประจำตัว 38012100

..... อาจารย์ที่ปรึกษา

(ผศ. เกษตร์ ศิริสันติสัมฤทธิ์)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อปริญญานิพนธ์	การรู้จำตัวพิมพ์อักษรภาษาอังกฤษบนสันหนังสือด้วยกล้อง CCD		
นักศึกษา	นาย ทศพร	พันภัยพาล	38012095
	นาย นพพร	พัฒนสุวรรณ	38012100
อาจารย์ที่ปรึกษา	ผศ. เกษตร์	ศิริสันติสัมฤทธิ์	
ระดับการศึกษา	วิศวกรรมศาสตร์บัณฑิต สาขาวิศวกรรมการวัดคุมทางอุตสาหกรรม		
ปีการศึกษา	พ.ศ. 2540		

บทคัดย่อ

ในปัจจุบันคอมพิวเตอร์ได้ถูกนำมาประยุกต์ใช้งานแทบทุกสาขาวิชาชีพ ในปริญญานิพนธ์ฉบับนี้ ได้นำเอาคอมพิวเตอร์มาประยุกต์ใช้ในการวิเคราะห์ข้อมูลภาพสันหนังสือภาษาอังกฤษ เพื่อจดจำชื่อหนังสือ โดยข้อมูลภาพจะถูกจัดเก็บด้วยกล้อง CCD (Charge Coupled Device) แล้วบันทึกลงหน่วยความจำ หลังจากนั้นนำข้อมูลภาพจากหน่วยความจำมาทำการแปลงข้อมูลภาพเป็น 2 ระดับด้วยการเลือกค่า Threshold อัตโนมัติ เพื่อการวิเคราะห์ข้อมูลภาพดิจิตอล แล้วแยกองค์ประกอบของข้อมูลภาพบนสันหนังสือโดยใช้ทฤษฎี Cellular Automata หลังจากนั้นทำการวิเคราะห์หาลักษณะเด่นของตัว อักษร (Feature Extraction) โดยใช้คุณสมบัติทางโทโปโลยี จะได้จุดปลาย, จุดแยก, จุดต่อ, จุดตัดและจุดโดดเดี่ยว วิเคราะห์หารหัสตัวอักษรแล้วเก็บเป็นฐานข้อมูลการรู้จำ แล้วเปรียบเทียบกับฐานข้อมูลต้นแบบที่เก็บไว้ในหน่วยความจำว่าเป็นตัวหนังสืออะไร เมื่อวิเคราะห์ได้แล้วบันทึกค่าที่ได้เป็นรหัส ASCII Code เพื่อนำผลของข้อมูลไปประยุกต์ใช้งานต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis : Printed English Character on Edge book Recognition with CCD camera

Students: Tosaporn Ponpaipal 38012095

Nopporn Pattanasuwan 38012100

Advisor: Kaset Sirisantisumrit

Education Level: Bachelor of industrial instrument engineering

Education Year : 1997

Abstract

Now a days, Computers're applied to use in all occupations. In this thesis computers a applied to analyze image data on the edge of English books for remember the name of books. In this way, the image data are collected by CCD (Charge Coupled Device) and recorded in memory and then, the image data were modified to 2 levels by choosing automatic threshold value to analyze digital image data and synthesis the components of image data on the edge of English books by using Cellular Automata theory. After that analyze distinetive feature of each letters (Feature Extraction) by using Topology Property. In finally will find the end, junction, jointing, cutting and isolate. To analyze the letters code and record as recognition data base, and compare with the prototype data base which record in the memory what're the type of letters. When the analyzation are completed, record these values to ASCII Code for applying to use these data.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

		หน้า
บทคัดย่อ		
บทที่ 1	**บทนำ, วัตถุประสงค์และขอบเขตของโครงงาน**	1
บทที่ 2	**ทฤษฎีและหลักการของระบบการประมวลผลภาพ**	
	2.1 หลักการเบื้องต้นของการประมวลผลภาพ	3
	2.2 การจำลองภาพ	9
	2.3 พื้นฐานเกี่ยวกับความสัมพันธ์ระหว่างจุดภาพ	10
	2.4 การสุ่มแบบต่อเนื่อง(uniform sampling)	13
	2.5 การแปลงภาพ(image transform)	14
	2.6 การทำภาพให้ชัดเจนขึ้น(image enhancement)	24
บทที่ 3	**ทฤษฎีและหลักการที่จะนำมาใช้**	
	3.1 หลักการทำงานของกล้อง CCD	32
	3.2 การปรับปรุงภาพโดยใช้ Linear density Conversion	41
	3.3 ทฤษฎีที่ใช้ในการหาค่า Threshold	46
	3.4 การตรวจสอบขอบภาพโดยการติดตามขอบเขตของภาพ (Contour following)	54
	3.5 การหาจุดเปลี่ยนแปลงทิศทางของลายเส้นเพื่อใช้ในการหามุม (Corner detection)	55
	3.6 การหมุนภาพ	58
	3.7 คุณสมบัติทางโทโปโลยีของภาพดิจิตอล (Topology)	61
บทที่ 4	**การจัดเตรียมข้อมูลเพื่อการรู้จำ**	
	4.1 การแยกลักษณะของสันหนังสือออกจากกลุ่มสันหนังสือ	66
	4.2 การหมุนสันหนังสือให้อยู่ในแนวมาตรฐาน	66
	4.3 การปรับปรุงข้อมูลภาพตัวหลังจากหมุนข้อมูลภาพ	68
	4.4 การแยกข้อมูลภาพตัวอักษรออกจากสันหนังสือ	69
	4.5 การแยกตัวอักษรออกจากประโยค	75

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.6 การทำภาพตัวอักษรให้บาง (Thinning) 76
4.7 การกำจัดส่วนเกิน 80

บทที่ 5 การวิเคราะห์ภาพอักษร
5.1 การแบ่งภาพเพื่อการหารหัสแทนตัวอักษร 84
5.2 การกำหนดค่าบิตในแต่ละไบท์ 84
5.3 การกำหนดรหัสแทนตัวอักษร 86
5.4 การจัดเก็บฐานข้อมูลการรู้จำ 87
5.5 การแยกความแตกต่างของตัวอักษร 91

บทที่ 6 โปรแกรมและผลการทดลอง
6.1 โปรแกรมการรู้จำอักษรภาษาอังกฤษบนสันหนังสือ 103
6.2 ผลการทดลอง 108
6.3 ตัวอย่างผลการทดลอง 117
6.4 ผลสรุปการทดลอง 124

บทที่ 7 บทสรุป 125

กิตติกรรมประกาศ

หนังสืออ้างอิง

ภาคผนวก

ก. โครงสร้างไฟล์ .PCX

ข. ตารางแทนรหัส ASCII Code

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปประกอบ

รูปที่	หน้า
บทที่ 1	
1.1 แสดง flow chart การทำงานของการรู้จำตัวอักษร	2
บทที่ 2	
2.1 ความสัมพันธ์ของภาพโดยทั่วไปกับพิกเซลเมตริกซ์	3
2.2 ดัชนีของพิกเซลในเมตริกซ์	4
2.3 แสดงหน้าต่างของภาพและหน้าต่างของพิกเซล	4
2.4 การแปลงภาพให้เป็นพิกเซลเมตริกซ์	5
2.5 การแสดงผลของจุดภาพ	5
2.6 แสดงตัวอย่างของระดับเทา	6
2.7 ฮิสโตแกรมที่มี 8 ระดับเทา	7
2.8 แสดงการสร้างฮิสโตแกรมจากภาพ	8
2.9 แสดงจุดภาพที่โดนจุดภาพที่เราสนใจล้อมรอบ	11
2.10 แสดงลักษณะความต่อเนื่องของพิกเซล	11
2.11 แสดงการทำฟูเรียร์ทรานสฟอร์มของฟังก์ชั่น 1 มิติ	16
2.12 แสดงการทำฟูเรียร์ทรานฟอร์มของฟังก์ชั่น 2 มิติ	17
2.13 แสดงฟังก์ชั่นต่อเนื่องที่ถูกสุ่มในช่วงกว้างเท่าๆ กัน	18
2.14 กราฟแสดงการทำคอนโวลูชั่น	22-23
2.15 แสดงภาพย่อยขนาด 3x3 พิกเซลที่มีจุด (x,y) เป็นจุดศูนย์กลาง	24
2.16 ฟังก์ชั่นการแปลงระดับเทาสำหรับทำให้ภาพเป็น 2 ระดับ	25
2.17 รูปทั่วไปของมาส์ก 3x3 ที่จะไปกระทำกับพิกเซลของภาพตามจุดต่างๆ จะทำการปฏิบัติการได้	26
2.18 ฟิลเตอร์แบบอุดมคติกรองความถี่ต่ำ	28
2.19 แสดงการทำเกรเดียนท์ของฟังก์ชั่น 2 มิติ	30
2.20 แสดงฟิลเตอร์แบบอุดมคติกรองความถี่สูง	31
บทที่ 3	
3.1 แสดงให้เห็นถึงโครงสร้างทั่วไป	32

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 แสดงให้เห็นถึงลักษณะรูปคลื่นของแต่ละเฟสที่มีความสัมพันธ์กัน
และลักษณะการเคลื่อนที่ของประจุเมื่อมีการถ่ายเท 33
3.3 แสดงให้เห็นถึงลักษณะและขนาดของอุปกรณ์ CCD ในขนาดต่างๆ 34
3.4 แสดงส่วนประกอบหลักของระบบ CCD 34
3.5 แสดงให้เห็นถึงการติดตั้ง CCD แบบแสงโดยตรง 35
3.6 แสดงให้เห็นถึงภาพที่ถ่ายจากอุปกรณ์ CCD แต่ใช้ฟิลเตอร์ช่วงความยาวคลื่นต่างๆ 36
3.7 แสดงภาพถ่ายด้วยเทคนิคพิเศษ 37
3.8 แสดงโครงสร้างเบื้องต้นของเครื่องสเปกโตรกราฟที่ใช้ CCD ต่อร่วม 37
3.9 แสดงการรับภาพตัวอักษรจากกล้อง CCD 38
3.10 แสดงลักษณะของการสแกนของกล้อง CCD 39
3.11 การเปลี่ยนสัญญาณ Analog เป็น Digital 39
3.12 แสดงฮิสโตแกรมสำหรับการทำรูปแบบสองระดับ 40
3.13 แสดงลักษณะของรูปแบบสองระดับ 40
3.14 แสดง histogram ที่ใช้ในการคำนวณ 41
3.15 แสดงความแตกต่างของการปรับปรุงก่อนและหลังข้อ 1 42
3.16 แสดงความแตกต่างของการปรับปรุงก่อนและหลังข้อ 2 43
3.17 แสดงความแตกต่างของการปรับปรุงก่อนและหลังข้อ 3 44
3.18 แสดงความแตกต่างของการปรับปรุงก่อนและหลังข้อ 4 45
3.19 แสดง Fixed-point ในการทำวนรอบ 49
3.20 แสดงภาพลักษณะของ Threshold ค่าต่างๆ 50-51
3.21 การ SCAN หาจุดภาพที่เป็นจุดเริ่มต้น 54
3.22 จุดเริ่มต้นของการติดตามหาขอบภาพ 54
3.23 การติดตามหาขอบภาพที่เป็นขอบ 55
3.24 ขอบของภาพที่ได้จากการติดตามขอบเขต 55
3.25 แสดงการหาค่าความโค้ง 56
3.26 แสดงถึงมุม 56
3.27 กราฟแสดงค่าความโค้ง 57
3.28 แสดงจุดมุมทั้ง 4 ของสันหนังสือ 58
3.29 แสดงการกำหนดตำแหน่ง (x,y) มุมของสันหนังสือ 58
3.30 การหมุนภาพโดยที่จุดหมุนอยู่ที่จุดกำเนิด 59
3.31 การหมุนภาพรอบจุดหมุนใดๆ 60
3.32 แสดงหน้าต่างขนาด 3x3 บิต 61

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.33 แสดงลักษณะคุณสมบัติทางโทโปโลยี ของกลุ่มจุดภาพ 62

บทที่ 4

4.1 แสดงลักษณะของภาพและการ scan ภาพเพื่อหาสันหนังสือ 66
4.2 แสดงสันหนังสือทำมุมกับแนวระนาบ 67
4.3 แสดงการหมุนภาพของสันหนังสือที่ทำมุม $0^{\circ} < \theta \leq 90^{\circ}$ 67
4.4 แสดงข้อมูลสันหนังสือหลังจากหมุนข้อมูล 67
4.5 แสดงปรับมุมภาพของสันหนังสือที่ทำมุม $90^{\circ} < \theta < 180^{\circ}$ 67
4.6 แสดงข้อมูลสันหนังสือหลังจากหมุนข้อมูลแล้ว 67
4.7 แสดงตารางหน้าต่างที่กำหนดในการปรับปรุงข้อมูลภาพ 68
4.8 แสดงลักษณะของจุดภาพ 71
4.9 แสดงลักษณะของจุดเป็น 71-72
4.10 แสดงลักษณะของจุดเกิด 72
4.11 แสดงตัวอย่างของจุดตายที่มีค่าเป็น 1 ถูกเปลี่ยนให้มีค่าเป็น 0 72
4.12 แสดงตัวอย่างของจุดเป็นที่มีค่าเป็น 1 ยังคงมีค่าเป็น 1 ต่อไป 73
4.13 แสดงตัวอย่างของจุดเกิดที่เดิมมีค่าเป็น 0 ถูกเปลี่ยนให้มีค่าเป็น 1 73
4.14 แสดงพิกัดขอบเขตของสันหนังสือ 75
4.15 แสดงภาพตัวอักษรแต่ละตัวในรูปประโยค 75
4.16 แสดงการตรวจหาขนาดของตัวอักษรในแนวแกน X 76
4.17 แสดงการตรวจหาขนาดของตัวอักษรในแนวแกน Y 76
4.18 โครงร่างของตัวอักษร 77
4.19 ตารางหน้าต่างสำหรับการทำให้บาง 77
4.20 ตารางหน้าต่างสำหรับการตรวจสอบหาขอบ 78
4.21 แสดงหน้าต่างของจุดภาพโครงร่าง 78-79
4.22 โครงร่างของตัวอักษร 80
4.23 ขั้นตอนการทำงานของอัลกอริธึมกำจัดส่วนเกิน 81

บทที่ 5

5.1 แสดงอัลกอริทึมการทำงานของกระบวนการรู้จำตัวอักษร 83
5.2 แสดงการแบ่งภาพออกเป็น 9 ส่วน 84
5.3 แสดงการกำหนดรหัสในแต่ละบิต (x=0,1) 84
5.4 แสดงภาพลายเส้นของช่องที่ 1 ในรูปตัว E 85

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.5 แสดงการกำหนดรหัสแทนช่องที่ 1 ของรูปตัว E ขนาด 1 ไบต์ 85
5.6 แสดงอัลกอริธึมการวิเคราะห์หารหัสแทนตัวอักษร 9 ไบต์ 85
5.7 แสดงการจัดวางตำแหน่งรหัสแทนตัวอักษรทั้ง 9 ไบต์ 86
5.8 แสดงตัวอักษรตัว E 86
5.9 แสดงการแทนรหัสที่ได้จากการหาค่าในแต่ละช่องของรูปตัว E 87
5.10 แสดงรหัสแทนตัวอักษรที่สมบูรณ์ของตัวอักษร E 87
5.11 แสดงการเก็บฐานข้อมูลการรู้จำ 87
5.12 แสดงตารางการแบ่งส่วนของตัวอักษร 91
5.13 แสดงการแบ่งช่องของตัวอักษร D และ 0,O 91
5.14 แสดงการแบ่งช่องพิจารณาของตัวอักษร 92
5.15 แสดงการแบ่งช่องในการพิจารณาตัวอักษร 92
5.16 แสดงภาพการแบ่งช่องในการพิจารณาตัวอักษร 93
5.17 แสดงภาพส่วนที่ 1 ของตัวอักษร N, S และ 5 93
5.18 แสดงภาพส่วนที่ 2 ของตัวอักษร S และ 5 94
5.19 แสดงภาพการแบ่งภาพตัวอักษรเพื่อใช้ในการพิจารณา 94
5.20 แสดงภาพส่วนที่ 0 ของตัวอักษร 94
5.21 แสดงภาพการแบ่งตัวอักษรเพื่อใช้ในการพิจารณา 95
5.22 แสดงส่วนที่ 6 ของตัวอักษร 95
5.23 แสดงภาพการแบ่งตัวอักษรเพื่อใช้ในการพิจารณา 96
5.24 แสดงส่วนที่ 0 ของภาพอักษร 96
5.25 แสดง flow chart ในการแยกแยะความแตกต่างตัวอักษรขั้นตอนที่ 1 98-100
5.26 แสดง flow chart ในการแยกแยะตัวอักษรของ I,l 101

บทที่ 6

6.1 แสดงหน้าจอภาพก่อนทำการวิเคราะห์การรู้จำ 105
6.2 แสดงการเปิดแฟ้มข้อมูลภาพ .PCX 106
6.3 แสดงรูปภาพของแฟ้มข้อมูลภาพที่จะทำการรู้จำตัวอักษร 106
6.4 แสดงตัวอักษรที่สามารถวิเคราะห์ได้ 107
6.5 แสดงผลการรู้จำตัวอักษร 108
6.6 แสดงตัวอย่างตัวอักษรต้นแบบ CordiaUPC 109
6.7 แสดงรูปที่ใช้ในการทดลอง 110
6.8 แสดงภาพต้นแบบสันหนังสือถ่ายจากกล้อง CCD 111

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.9 แสดงข้อมูลภาพ 2 ระดับ 111
6.10 แสดงภาพการแยกสันหนังสือออกจากกลุ่ม 112
6.11 แสดงการหมุนภาพสันหนังสือเข้าหาแกนนอน 112
6.12 แสดงการรวมกลุ่มข้อมูลบนสันหนังสือ 113
6.13 แสดงการแยกกลุ่มตัวอักษรออกจากสันหนังสือ 113
6.14 แสดงการแยกตัวอักษรออกจากกลุ่มตัวอักษร 114
6.15 แสดงการทำตัวอักษรให้บาง 114
6.16 แสดงการแบ่งตัวอักษรออกเป็น 9 ช่อง 115
6.17 แสดงผลการรู้จำตัวอักษร 115
6.18 แสดงภาพตัวอักษรเมื่อไม่สามารถรู้จำตัวอักษรได้ 116
6.19 แสดงตัวอย่างผลการทดลองที่ 1 117
6.20 แสดงตัวอย่างผลการทดลองที่ 2 118
6.21 แสดงตัวอย่างผลการทดลองที่ 3 119
6.22 แสดงตัวอย่างผลการทดลองที่ 4 120
6.23 แสดงตัวอย่างผลการทดลองที่ 5 121
6.24 แสดงตัวอย่างผลการทดลองที่ 6 122
6.25 แสดงตัวอย่างผลการทดลองที่ 7 123

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
บทที่ 2	
2.1 จำนวนบิตที่ต้องใช้เมื่อมีการเปลี่ยนแปลงขนาดของ N และ m	14
2.2 จำนวนของไบท์ (8บิท) ที่ต้องใช้เมื่อมีการเปลี่ยนแปลงขนาดของ N และ m	14
บทที่ 3	
3.1 แสดงค่าคุณสมบัติทางโทโปโลยีคอลของกลุ่มจุดภาพ	63
บทที่ 5	
5.1 แสดงตารางค่ารหัสของตัวอักษรพิมพ์ใหญ่ภาษาอังกฤษต้นแบบ CordiaUPC	88-89
5.2 แสดงตารางค่ารหัสของตัวอักษรพิมพ์เล็กภาษาอังกฤษต้นแบบ CordiaUPC	89-90
5.3 แสดงตารางค่ารหัสของเลขพิมพ์ภาษาอังกฤษต้นแบบ CordiaUPC	90

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

ความต้องการในการใช้คอมพิวเตอร์ในการประมวลผลทางภาพมีเพิ่มมากขึ้น การที่ให้คอมพิวเตอร์ประมวลผลทางภาพเพื่อรู้จำตัวอักษรต่าง ๆ ก็เป็นอีกแนวคิดหนึ่งที่ต้องการนำไปใช้งาน ในที่นี้เราได้นำเอามาใช้ในการประมวลผลตัวพิมพ์อักษรภาษาอังกฤษที่อยู่บนสันหนังสือ ซึ่งจะทำการเก็บภาพเข้ามาแล้วทำการแปลงข้อมูลภาพที่ได้มาให้เป็นภาพ 2 ระดับ จึงนำภาพนั้นมาทำการประมวลผลโดยนำทฤษฎีทางโทโปโลยีมาช่วยในการแยกแยะลักษณะของจุดเด่นของตัวอักษรเพื่อนำข้อมูลที่ได้นั้นไปเปรียบเทียบกับฐานข้อมูลที่ได้ทำไว้ และแสดงผลออกมาว่าภาพที่เราจัดเก็บมานั้นมีตัวอักษรอะไรบ้าง ในที่นี้เราได้นำภาพตัวพิมพ์อักษรภาษาอังกฤษจากสันหนังสือ โดยเราจะทำการแยกแยะสันหนังสือออกจากกันก่อนแล้วทำการแยกแยะตัวอักษรอีกทีหนึ่ง และแสดงผลออกมา

วัตถุประสงค์ของโครงงาน

1. เพื่อศึกษาเกี่ยวกับทฤษฎีทางโทโปโลยีที่นำมาใช้ในการรู้จำตัวอักษร
2. เพื่อศึกษาเกี่ยวกับระบบภาพดิจิตอลที่นำมาประมวลผลโดยผ่านทางกล้อง CCD
3. เพื่อศึกษาเกี่ยวกับการประมวลผลภาพ
4. เพื่อศึกษาให้ไมโครคอมพิวเตอร์สามารถรู้จำข้อมูลทางภาพได้
5. เพื่อให้สามารถนำไปใช้กับงานในรูปแบบอื่นๆ ได้ เช่น ใช้กับแขนกลในการหยิบจับหนังสือที่ต้องการ

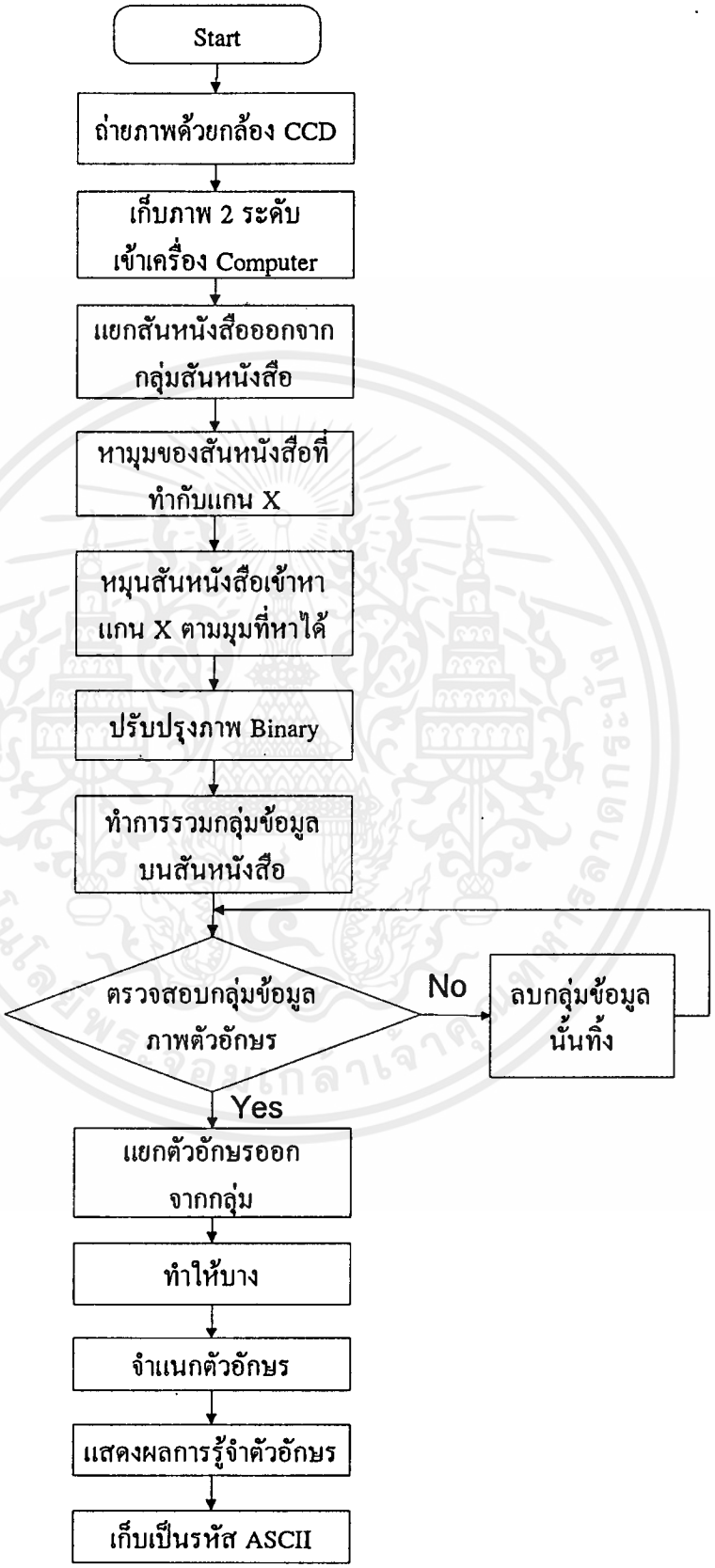
ขอบเขตของโครงงาน

1. สามารถนำภาพดิจิตอลเข้ามาทำการประมวลผล โดยผ่านทางกล้อง CCD
2. สามารถแบ่งแยกระหว่างสันหนังสือแต่ละเล่มได้
3. สามารถรู้จำตัวพิมพ์อักษรภาษาอังกฤษภายในสันหนังสือได้
4. สามารถนำภาพจริงที่เก็บเข้ามาประมวลผลการรู้จำตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลำดับขั้นตอนการทำงาน

สามารถแสดงในรูปของ flow chart การทำงาน 1 รอบการทำงานได้ต่อไปนี้



รูปที่ 1.1 แสดง flow chart การทำงานของการรู้จำตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและหลักการของระบบการประมวลผลภาพ

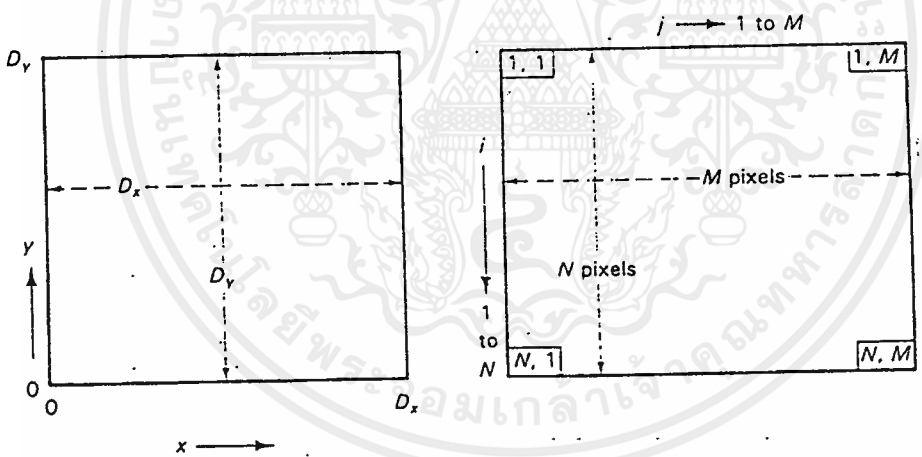
2.1 หลักการเบื้องต้นของการประมวลผลภาพ

2.1.1 พิกเซล (pixel)

การแสดงข้อมูลสภาพดิจิตอลสามารถอธิบายได้ด้วยเมตริกซ์ N x M และให้จุดต่างๆ ที่อยู่ใน เมตริกซ์เป็นจุด x,y ใดๆ ซึ่งเป็นส่วนประกอบของภาพในแต่ละจุด x,y ใดๆ เรียกว่าพิกเซลหรือจุดภาพ และ ในแต่ละพิกเซลจะแสดงให้เราเห็นได้ด้วยฟังก์ชั่นของความเข้มของแสง (องค์ประกอบของ p(i,j)) เมื่อเราเปรียบเทียบระหว่างภาพและพิกเซลเมตริกซ์ (pixel metrix) ดังรูปที่ 2.1 จะเห็นว่าจุดกำเนิดของภาพจะอยู่ที่มุมซ้ายล่าง แต่จุดกำเนิดของพิกเซลจะอยู่ที่มุมซ้ายบน ซึ่งจะเป็นลักษณะการประมวลผลภาพในกราฟฟิกของคอมพิวเตอร์

กล่าวคือ i = y เมื่อ 1 <= i <= N

j = x เมื่อ 1 <= j <= M



a) ลักษณะของรูปภาพ b) ลักษณะของพิกเซลเมตริกซ์

รูปที่ 2.1 ความสัมพันธ์ของภาพโดยทั่วไปกับพิกเซลเมตริกซ์

เมื่อ

x = Dx/M

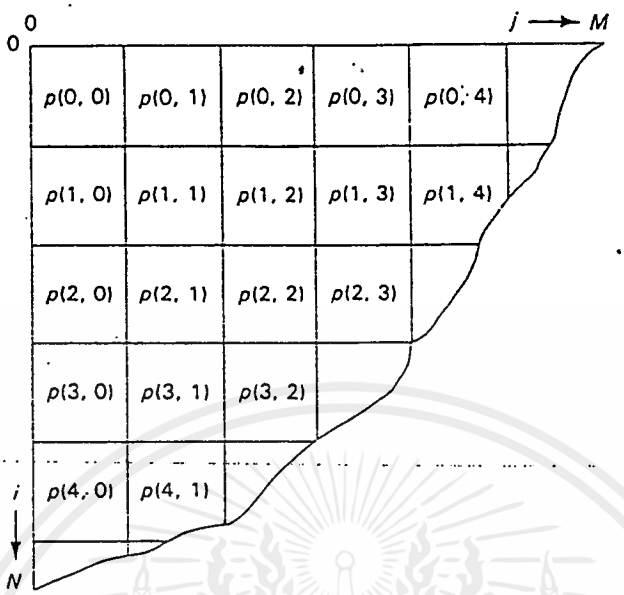
y = Dy/N

N = จำนวนสูงสุดของพิกเซลในแนวตั้ง

M = จำนวนสูงสุดของพิกเซลในแนวนอน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อเราให้จุดต่างๆ บนเมตริกซ์เป็น p(i,j) ใดๆ

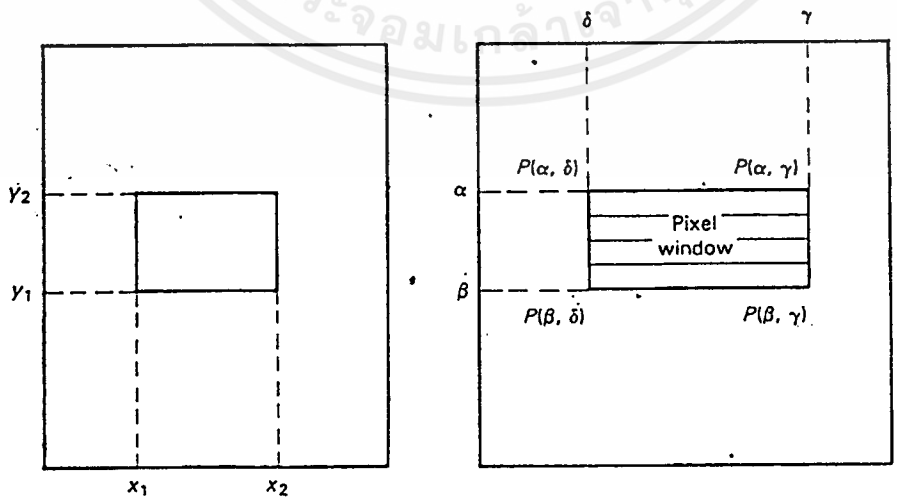


รูปที่ 2.2 ดัชนีของพิกเซลในเมตริกซ์

ค่าของพิกเซลหรือฟังก์ชั่น p(i,j) ณ. จุดใดๆ จะแสดงได้ด้วยค่าของความเข้มของแสงซึ่งอาจแบ่งได้หลายระดับ ถ้ามีแค่ 2 ระดับก็จะเป็นแค่ 0 กับ 1

2.1.2 หน้าต่าง (windows)

เป็นพื้นที่ส่วนย่อยของภาพหรือเรียกว่า หน้าต่าง และ สามารถกำหนดได้ด้วยมุม 4 มุม แสดงด้วยค่าของพิกเซล $P(\beta,\delta)$,$P(\beta,\gamma)$,$P(\alpha,\delta)$,$P(\alpha,\gamma)$ แสดงได้ดังรูปที่ 2.3



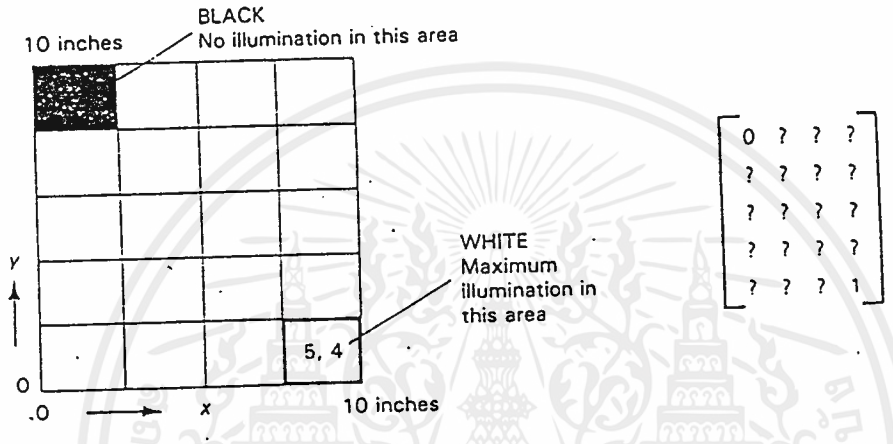
a) หน้าต่างของภาพ b) หน้าต่างของพิกเซล

รูปที่ 2.3 แสดงหน้าต่างของภาพและหน้าต่างของพิกเซล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.3 ตำแหน่งของพิกเซล (pixel position)

ตำแหน่งของจุดภาพหรือพิกเซลทุกจุดจะต้องอยู่ภายในพื้นที่ NxM เมื่อเราพิจารณาจากรูปที่ 2.4(a) จะเห็นว่าไม่มีแสง ณ. บริเวณมุมบนซ้าย และบริเวณที่สว่างที่สุดอยู่ที่มุมล่างขวาของภาพ ซึ่งมีขนาด 10 x 10 นิ้วพื้นที่ที่ไม่มีแสงแสดงได้ด้วยศูนย์ ส่วนพื้นที่ ที่สว่างที่สุดแสดงได้ด้วยหนึ่ง และภาพที่เห็นได้คือ 5 x 5 เมตริกซ์ (5 แถว,5 คอลัมน์) แต่ละส่วนของภาพจะกว้าง 2.5 นิ้ว ยาว 2 นิ้ว บริเวณมุมบนซ้ายจะเป็นศูนย์ และพื้นที่ 2.5 x 2 นิ้วจะเป็นซึ่งนำมาเขียนเป็นเมตริกซ์ได้ดัง รูปที่ 2.4(b)



a) ภาพ b) แปลงเป็นพิกเซลเมตริกซ์

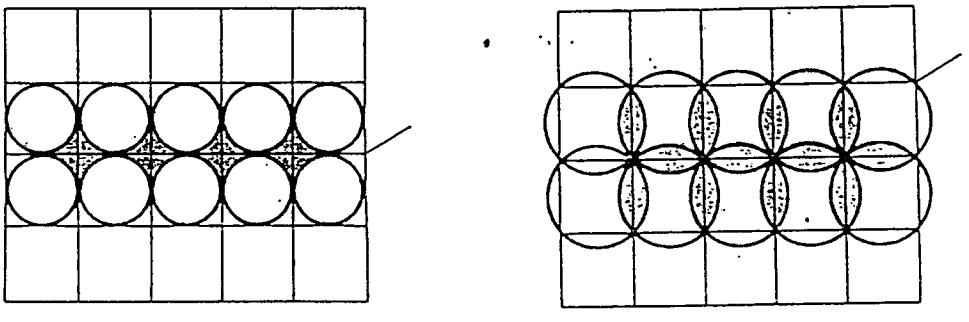
รูปที่ 2.4 การแปลงภาพให้เป็นพิกเซลเมตริกซ์

จากเมตริกซ์ในรูปที่ 2.4(b) ถ้ามีระดับเทา (gray scale) เป็น 16 ระดับค่าของพิกเซล ณ. บริเวณที่สว่างที่สุดจะมีค่าเป็น 15

ลักษณะการแสดงผลในแต่ละพิกเซลหรือจุดภาพจะขึ้นอยู่กับเซ็นเซอร์ซึ่งมีอยู่ 2 ลักษณะ

- แสดงผลเฉพาะพิกเซล
- แสดงผลเหลื่อมกัน

ซึ่งแสดงดังรูปที่ 2.5



a) แสดงผลเฉพาะพิกเซล b) แสดงผลเหลื่อมกัน

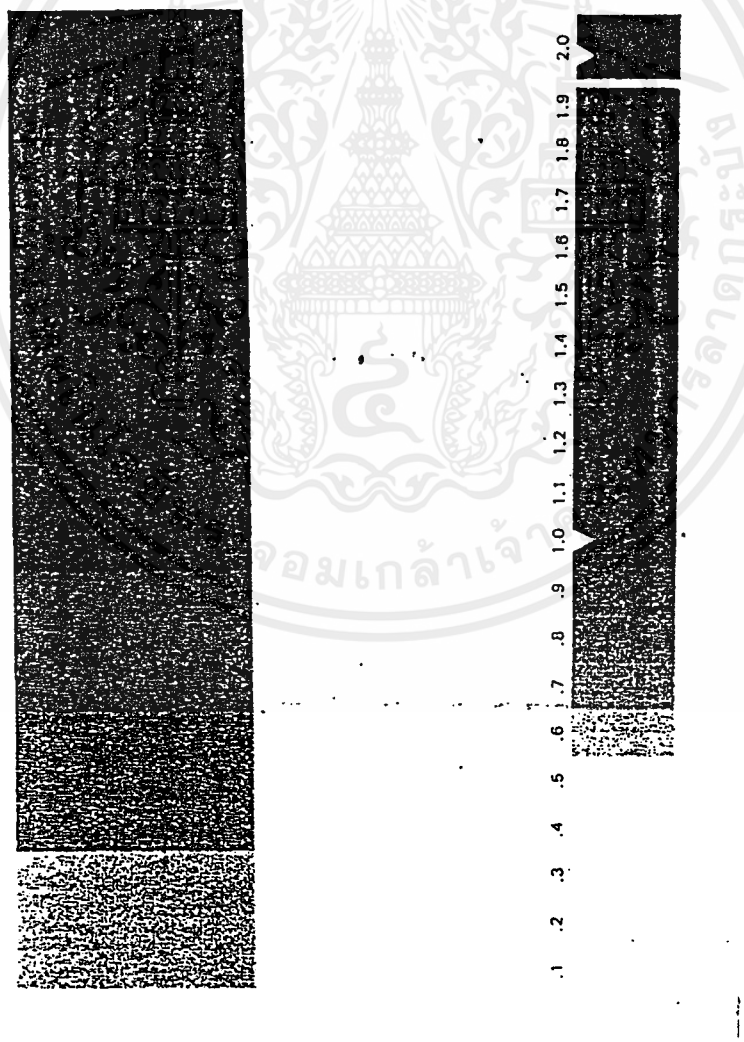
รูปที่ 2.5 การแสดงผลของจุดภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.4 ระดับเทา (gray scale)[1]

ระดับเทา เป็นค่าที่บอกถึงความสว่างของจุดภาพ เราสามารถเพิ่มความสว่างหรือระดับเทาให้กับจุดภาพได้หลายระดับ โดยการเพิ่มจำนวนของบิตในการนำเสนอค่าของพิกเซล ตัวอย่างเช่น ต้องการระดับความสว่างหาได้จาก 2 ยกกำลังด้วยจำนวนบิต

ระดับเทา		ค่าย่านของระดับเทา
2^1	2 ค่า	0 และ 1
2^3	8 ค่า	0 ถึง 7
2^4	16 ค่า	0 ถึง 15
2^8	256 ค่า	0 ถึง 255



รูปที่ 2.6 แสดงตัวอย่างของระดับเทา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.5 ฮิสโตแกรม (histograms)

ฮิสโตแกรมเป็นกราฟแท่งที่บอกถึงความถี่ของแต่ละความเข้มของแสง (gray scale) ของภาพ จากรูปในแนวแกน x เป็นค่าของระดับเทา และแกน y เป็นจำนวนของจุดภาพที่มีระดับต่างๆ

เราสามารถสร้างฮิสโตแกรมได้โดย

1. นำภาพที่มีระดับความเข้มของแสงที่แตกต่างกัน (0-255)
2. นับจำนวนจุดภาพที่มีระดับความเข้มของแสงเดียวกัน ทุกระดับความเข้มของแสงจาก (0-255)
3. พล๊อตความถี่ของจุดภาพที่แต่ละระดับความเข้มของแสงจาก 0-255

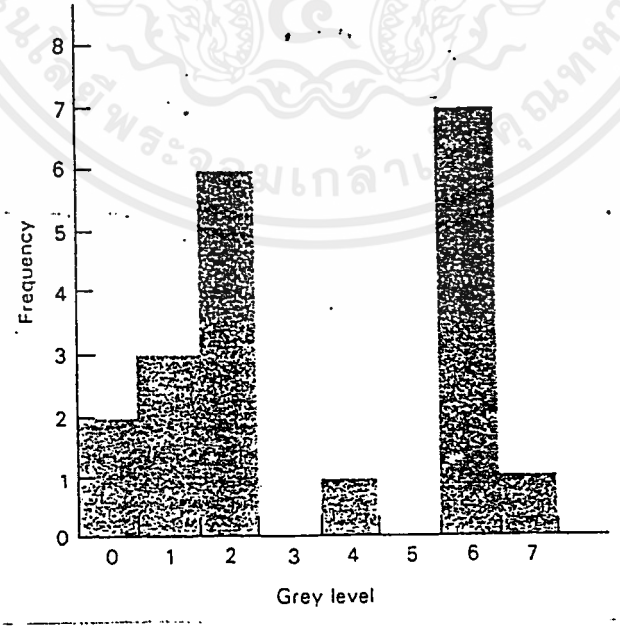
โดยเราสามารถหาความน่าจะเป็นของจุดภาพ ณ. ระดับเทาหนึ่งๆ ได้ตามสมการ

P(b) ณ. จุด (x,y) ในภาพ = ค่าของ b / จำนวนของพิกเซลทั้งหมดในภาพ

เช่น ที่ระดับเทา 6 ค่าของฮิสโตแกรมเท่ากับ 7 จะได้

$$P(6) = \frac{7}{20}$$

$$= 0.35$$

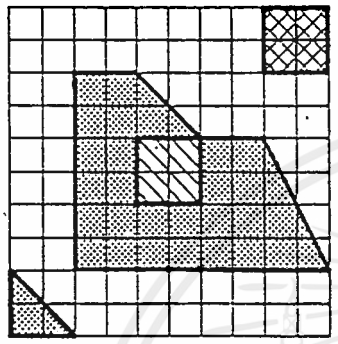


รูปที่ 2.7 ฮิสโตแกรมที่มี 8 ระดับเทา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปร่างของฮิสโตแกรมจะขึ้นอยู่กับคุณสมบัติของภาพ ประโยชน์ของฮิสโตแกรม คือ ใช้สำหรับปรับค่าเทรสโฮล (threshold) เพื่อจะแปลงระดับเทาของภาพให้เป็นภาพ 2 ระดับหรือใช้สำหรับปรับแต่งส่วนของสเปกตรัมระดับเทา

ตัวอย่างของภาพที่มีระดับเทาต่างกันเมื่อนำมาสร้างเป็นฮิสโตแกรม



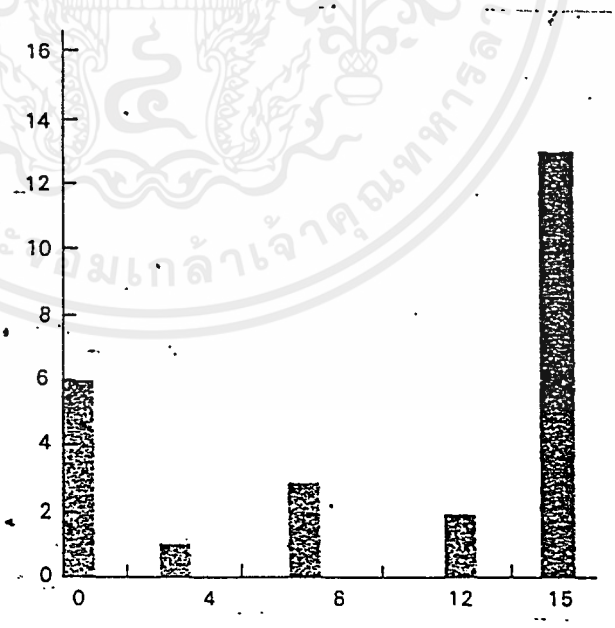
a) ขั้นที่ 1

$$\begin{bmatrix} 15 & 15 & 15 & 15 & 12 \\ 15 & 0 & 7 & 15 & 15 \\ 15 & 0 & 7 & 0 & 12 \\ 15 & 0 & 0 & 0 & 3 \\ 7 & 15 & 15 & 15 & 15 \end{bmatrix}$$

b) ขั้นที่ 2

Value	Number of Pixels
0	6
3	1
7	3
12	2
15	13
Total	25

c) ขั้นที่ 3



d) ขั้นที่ 4

รูปที่ 2.8 แสดงการสร้างฮิสโตแกรมจากภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีการ

1. หาจำนวนทั้งหมดของพิกเซลในเมตริกซ์ M x N

จะเห็นว่า M = 10, N = 10

พิกเซลทั้งหมด = 10 x 10 = 100

2. สร้างพื้นที่ของภาพแทนด้วยเมตริกซ์ จากตัวอย่างจะได้ เมตริกซ์ 5x5 จำนวนของพิกเซลทั้งหมดที่แทนในเมตริกซ์จะลดลงเหลือ 25 พิกเซล

3. ทำตารางความสัมพันธ์ระหว่างค่าระดับเทาและจำนวนของพิกเซล

4. สร้างฮิสโตแกรมเป็นกราฟแท่งโดยให้ระดับเทาเพิ่มทีละ 1 ระดับในแนวแกน X และเนื่องจากค่าระดับเทาสูงสุดคือ 15 จึงมีระดับเทาทั้งหมด 16 ระดับในแนวแกน Y

2.2 การจำลองภาพ

จากคำจำกัดความของ อิมเมจ (image) หมายถึง ฟังก์ชั่นความเข้มของแสงในระนาบ 2 มิติ คือ f(x,y) เมื่อค่าของ f ที่อยู่บนโคออดิเนท (x,y) คือความสว่างบนภาพ ณ. จุดนั้นๆ แสงจะอยู่ในรูปของพลังงาน ค่าฟังก์ชั่น f(x,y) จึงไม่เป็นศูนย์และมีค่าๆ หนึ่งนั่นคือ

$$0 < f(x,y) < \infty \quad \text{..... (2.1)}$$

เราเห็นภาพของวัตถุได้โดยการอาศัยการสะท้อนของแสงที่ไปกระทบวัตถุ ซึ่งสามารถอธิบายฟังก์ชั่น f(x,y) ได้ด้วยส่วนประกอบอันที่ 1 คือ แสงที่ส่องไปยังวัตถุ (illumination component) และส่วนประกอบที่ 2 คือ แสงที่สะท้อนกลับมาจากวัตถุ (reflectance component) และสามารถเขียนเป็นฟังก์ชั่น คือ i(x,y) และ r(x,y) ตามลำดับ ฟังก์ชั่น f(x,y) เกิดจากการคูณกันระหว่าง i(x,y) และ r(x,y)

$$f(x,y) = i(x,y)r(x,y) \quad \text{..... (2.2)}$$

เมื่อ

$$0 < i(x,y) < \infty \quad \text{.....(2.3)}$$

$$0 < r(x,y) < 1 \quad \text{.....(2.4)}$$

สมการที่ 2.4 เป็นฟังก์ชั่นการสะท้อนกลับของแสง ณ. จุด (x,y) มีค่าระหว่าง 0 (คือวัตถุดูดกลืนแสงทั้งหมด ไม่สะท้อนกลับเลย) กับ 1 (คือวัตถุ สะท้อนแสงกลับทั้งหมด ไม่มีการดูดกลืนแสง) ส่วนฟังก์ชั่น i(x,y) หาได้จากแหล่งกำเนิดแสง ส่วนฟังก์ชั่น r(x,y) หาได้จากคุณสมบัติการดูดแสงของวัตถุ ในวันที่มีแสงแดดจ้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดวงอาทิตย์จะให้ความสว่าง 9000 แรงเทียนส่องมายังผิวโลก ถ้าวันใดมีเมฆมากความสว่างจะลดลงเหลือ 1000 แรงเทียนโดยประมาณ ค่าของ r(x,y) จะมีค่า 0.01 สำหรับกำมะหยี่, 0.65 สำหรับสเตนเลส, 0.9 สำหรับเงินและ 0.93 สำหรับหิมะ

ความเข้มของแสงแบบโมโนโครม เรียกว่า ระดับเทา (gray level ; l) ของภาพ ณ. จุดนั้นๆ จะได้

$$L_{min} \leq l \leq L_{max}$$

โดย

$$L_{min} = i_{min}\, r_{min}$$

$$L_{max} = i_{max} r_{max}$$

โดยทั่วไป L_{min} จะมีค่าประมาณ 0.005 และ L_{max} มีค่าประมาณ 100 สำหรับใช้กับการประมวลผลภาพ

ช่วง $[L_{min}.L_{max}]$ เรียกว่า ระดับเทา เราอาจเขียนช่วงใหม่ให้เป็น [0,L], เมื่อ l = 0 นั่นคือสีดำและ l = L นั่นคือสีขาวตามสเกลของระดับเทา

2.3 พื้นฐานเกี่ยวกับความสัมพันธ์ระหว่างจุดภาพ

ในส่วนนี้จะแสดงให้เห็นถึงความสัมพันธ์ระหว่างจุดภาพ ในระบบการประมวลผลภาพดิจิตอล ฟังก์ชั่นของข้อมูลภาพแสดงได้ด้วยฟังก์ชั่น f(x,y) และจะแทนด้วยตัวอักษรตัวเล็กคือ p และ q ส่วนเซตย่อยของ จุดภาพของฟังก์ชั่น f(x,y) จะแทนด้วยตัว S

2.3.1 สิ่งรอบข้างพิกเซล (neighbors of a pixel)

พิกเซล p ณ. โคออดิเนท (x,y) จะโดนล้อมรอบด้วยพิกเซลอื่นทั้งแถวนอนและแถวตั้ง นั่นคือ

$$(x+1,y),(x-1,y),(x,y+1),(x,y-1)$$

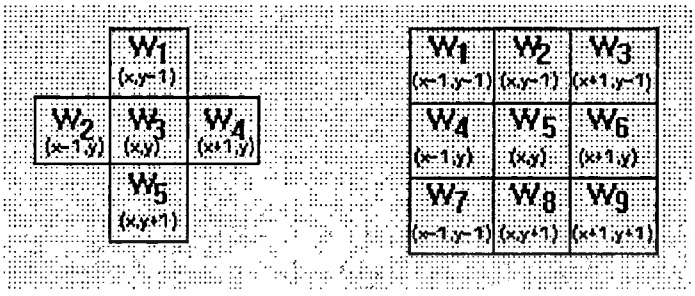
คือเซตของพิกเซล เรียกว่า การล้อมรอบทั้ง 4 ด้านของพิกเซล p (4-neighbors of p) ซึ่งแสดงได้โดย $N_4(p)$ นั่นคือจุดทั้ง 4 จะอยู่ห่างจากจุด (x,y) เป็นระยะ 1 หน่วย ในแนวทั้ง 4 จุด โดยจุดทั้ง 4 มีโคออดิเนทดังนี้

$$(x+1,y+1),(x+1,y-1),(x-1,y+1),(x-1,y-1)$$

ในกรณีเราคิดจุดล้อมรอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในแนวทะแยงด้วยจะเห็นว่าจุดภาพหรือพิกเซลจะโดนล้อมรอบด้วยจุดภาพอื่นๆ 8 พิกเซลด้วยกัน จึงเรียกว่า 8-neighbors ของ p หรือ $N_8(b)$ ซึ่งแสดงได้ดังรูปที่ 2.9



a) 4-neighbors b) 8-neighbors

รูปที่ 2.9 แสดงจุดภาพที่โดนจุดภาพที่เราสนใจล้อมรอบ

2.3.2 ความต่อเนื่องของจุดภาพ (connectivity)

ความต่อเนื่องระหว่างพิกเซลเป็นหลักการสำคัญที่จะใช้หาขอบเขตของวัตถุและส่วนประกอบพื้นที่ของภาพ พิกเซล 2 พิกเซลต่อถึงกันได้แสดงว่าพิกเซลนั้นๆ จะต้องอยู่ติดกันให้ V เป็นเซตของระดับเทาถ้าแต่ละความต่อเนื่องของพิกเซลมีความเข้ม 59,60,61 จะได้ V = {59,60,61} เราสามารถพิจารณาความต่อเนื่องได้ 3 วิธี

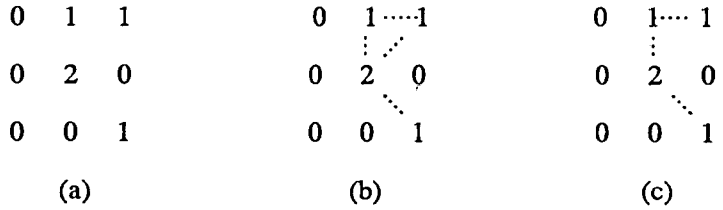
a) ความต่อเนื่อง 4 (4-connectivity) พิกเซล p และ q ที่มีค่าจากเซต V ต่อถึงกัน 4 ตัว ถ้า q อยู่ใน $N_4(p)$

b) ความต่อเนื่อง 8 (8-connectivity) พิกเซล p และ q ที่มีค่าจากเซต V ต่อถึงกัน 8 ตัว ถ้า q อยู่ใน $N_8(p)$

c) ความต่อเนื่องแบบผสม (mixed connectivity) พิกเซล p และ q ที่มีค่าจากเซต V จะต่อกันแบบผสม ถ้า

(i) q อยู่ใน $N_4(p)$ หรือ

(ii) q อยู่ใน $N_D(p)$ และไม่มีสมาชิกในเซต $N_4(p) \cap N_4(q)$



(a) (b) (c)

รูปที่ 2.10 แสดงลักษณะความต่อเนื่องของพิกเซล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

a) การเรียงกันของพิกเซล

b) พิกเซลหมายเลข 2 มีสมาชิกล้อมรอบ 8 ตัว

c) พิกเซลแบบล้อมรอบผสมของพิกเซลชนิดเดียวกัน

2.3.3 การวัดระยะระหว่างพิกเซล (distance measure)

ให้พิกเซล p,q และ z อยู่ที่โคออดิเนท (x,y),(s,t) และ (u,v) ตามลำดับ เรียก D ว่าฟังก์ชั่นระยะ (distance function) หรือ เมตริกซ์ ถ้า

a) $D(p,q) \geq 0$ $(D(p,q) = 0$ if $p=q)$,

b) $D(p,q) = D(q,p)$, and

c) $D(p,z) \leq D(p,q) + D(q,z)$

ระยะยูคลิด (Euclidean) ระหว่าง p และ q คือ

$$D_e(p,q) = [(x-s)^2 + (y-t)^2]^{1/2}$$

สำหรับระยะที่วัดได้ของพิกเซลจะมีระยะน้อยกว่าหรือเท่ากับค่าของ r จาก จุด (x,y) หรือจุดที่อยู่บนดิสก์ของรัศมี r เมื่อให้จุด (x,y) เป็นจุดศูนย์กลาง

ระยะ D4 ระหว่าง p และ q ถูกกำหนดโดย

$$D4(p,q) = |x - s| + |y - t|$$

ในกรณีนี้ระยะห่าง D4 เป็นระยะ (x,y) ซึ่งเป็นระยะที่น้อยกว่าหรือเท่ากับค่าของ r จากจุดศูนย์กลางของพิกเซลรูปเพชร ณ. จุด (x,y) จากตัวอย่าง พิกเซล D4 มีระยะน้อยกว่าเท่ากับ 2 จาก (x,y) หรือ จุดศูนย์กลาง

```
        2
    2   1   2
2   1   0   1   2
    2   1   2
        2
```

จะเห็นว่าพิกเซลล้อมรอบ 4 ตัวของพิกเซลที่มีค่า 0 จะเป็น 1 หรือ $D_4 = 1$

ค่าสำหรับระยะ D_8 ระหว่างจุด p และ q ถูกกำหนดโดย

$$D_8(p,q) = \max(|x- s| , |y-t|)$$

ในกรณีนี้พิกเซลที่มีระยะ D_8 จากจุด (x,y) จะมีค่าน้อยกว่าหรือเท่ากับค่าของ t จากจุดศูนย์กลางของพิกเซลรูปสี่เหลี่ยม ณ. จุด (x,y) จากตัวอย่างพิกเซล D_8 มีระยะน้อยกว่าเท่ากับ 2 จากจุด (x,y) หรือจุดศูนย์กลาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

ซึ่งจะเห็นว่าพิกเซลล้อมรอบ 8 ตัวของพิกเซลที่มีค่า 0 จะเป็น 1 หรือ $D_8 = 1$

2.4 การสุ่มแบบต่อเนื่อง (uniform sampling)

ในการประมวลผลภาพฟังก์ชั่นของข้อมูล f(x,y) จะต้องนำมาทำให้เป็นดิจิตอลทั้งสเปเทียลและแอมปลิจูด การทำ digitization ในโคออดิเนทสเปเทียลของจุด (x,y) จะหมายถึงการสุ่มข้อมูลภาพ (image sampling) ส่วนขนาดของการ digitization จะหมายถึงความสว่างของระดับเทา (gray - level quantization)

สมมติให้ฟังก์ชั่นของภาพต่อเนื่องจัดเรียงอยู่ในรูปของเมติกซ์ NxN

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ & & & \\ f(N-1,0) & f(N-1,1) & \cdots & f(N-1,n-1) \end{bmatrix} \quad \ldots\ldots\ldots\ldots (2.5)$$

จะต้องว่าด้านขวาของสมการของสมการเป็นภาพดิจิตอล (digital image) ในแต่ละจุดหรืออิลิเมนท์จะเรียกว่า พิกเซลหรือ เพล จากขบวนการ digitization เราจะต้องตัดสินใจเลือกขนาดของ N และจำนวนของระดับเทาว่าจะแยกได้กี่ระดับในแต่ละพิกเซลได้ใช้สมการ

$$N = 2^n \quad \ldots\ldots\ldots\ldots (2.6)$$

และ

$$G = 2^m \quad \ldots\ldots\ldots\ldots (2.7)$$

เมื่อ G คือจำนวนของระดับเทาโดยแยกได้ตั้งแต่ 0 ถึง L จากสมการที่ 2.6 และ 2.7 เราสามารถคำนวณจำนวนบิตที่ใช้ในการเก็บภาพที่ผ่านการ digitization แล้ว

$$b = N \times N \times m \quad \ldots\ldots\ldots\ldots (2.8)$$

เมื่อ b คือ จำนวนบิตที่ใช้ในการเก็บภาพที่ผ่านการ digitization

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Nm	1	2	3	4	5	6	7	8
32	1024	2048	3072	4096	5120	6144	7168	8192
64	4096	8192	12288	16384	20480	24576	28672	32768
128	16384	32768	49152	65536	81920	98304	124688	131072
256	65536	131072	196608	262144	327680	393216	458752	524088
512	262144	524288	786432	1048576	1310720	1572864	1835008	2097152

ตารางที่ 2.1 จำนวนบิตที่ต้องใช้เมื่อมีการเปลี่ยนแปลงขนาดของ N และ m

Nm	1	2	3	4	5	6	7	8
32	128	256	512	512	1024	1024	1024	1024
64	512	1024	2048	2048	4096	4096	4096	4096
128	2048	4096	8192	8192	16384	16384	16384	16384
256	8192	16384	32768	32768	65536	65536	65536	65536
512	32768	65536	131072	131072	262144	262144	262144	262144

ตารางที่ 2.2 จำนวนของไบท์ (8บิต) ที่ต้องใช้เมื่อมีการเปลี่ยนแปลงขนาดของ N และ m

2.5 การแปลงภาพ (image transform)

2.5.1 ฟูเรียร์ทรานฟอร์มเบื้องต้น (concept of the fourier transform)

ให้ f(x) เป็นฟังก์ชั่นต่อเนื่องของตัวแปรจริง x ฟูเรียร์ทรานฟอร์มของฟังก์ชั่น f (x) แสดงได้โดย $\mathfrak{F}\{f(x)\}$ จะถูกนิยามโดยสมการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\mathfrak{F}\{f(x)\} = F(u) = \int_{-\infty}^{\infty} f(x)\exp[-j2\pi ux]dx \quad \text{..... (2.9)}$$

เมื่อ $j = -1$

ฟังก์ชั่นที่ทำฟูเรียร์ทรานฟอร์มไปแล้ว สามารถแปลงกลับมาเป็นฟังก์ชั่นเดิมได้ โดยใช้อินเวอร์สฟูเรียร์ทรานสฟอร์ม (inverse fourier transform) โดย

$$\mathfrak{F}^{-1}\{F(u)\} = f(x)$$

$$= \int_{-\infty}^{\infty} F(u)\exp[-j2\pi ux]dx \quad \text{..... (2.10)}$$

จากสมการที่ 2.9 และ 2.10 จะเรียกว่า คู่สมการฟูเรียร์ทรานสฟอร์ม (fourier transforms pair) ซึ่งจะทำการทรานสฟอร์มได้ถ้าฟังก์ชั่น f(x) เป็นฟังก์ชั่นแบบต่อเนื่องและสามารถทำการอินติเกรทได้ และ F(u) สามารถอินติเกรทได้ด้วยสภาวะเช่นนี้โดยมากจะเกิดในทางปฏิบัติ เมื่อเราให้ f(x) เป็นจริงเราจะได้ฟูเรียร์ทรานสฟอร์มเป็นจำนวนเชิงซ้อน (complex) นั่นคือ

$$F(u) = R(u) + jI(u) \quad \text{..... (2.11)}$$

เมื่อ R(u) และ I(u) เป็นส่วนจริง (real component) และส่วนจินตภาพ (imaginary component) ของ F(u) เราสามารถเปลี่ยนสมการที่ 2.11 ให้อยู่ในฟอร์มของเอ็กซ์โปเนนเชี่ยล (exponential form) ได้

$$F(u) = |F(u)|\, e^{j\phi(u)} \quad \text{..... (2.12)}$$

เมื่อ

$$|F(u)| = \left[R^2(u) + I^2(u)\right]^{1/2} \quad \text{..... (2.13)}$$

และ

$$\phi(u) = \tan^{-1}\left[\frac{I(u)}{R(u)}\right] \quad \text{..... (2.14)}$$

แมกนิจูด (magnitude) ของฟังก์ชั่น $|f(u)|$ เรียกว่า ฟูเรียร์ สเปกตรัม (fourier spectrum) ของ f(x) และ $\phi(u)$ คือ มุมเฟส (phase angle)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สเปกตรัมกำลัง

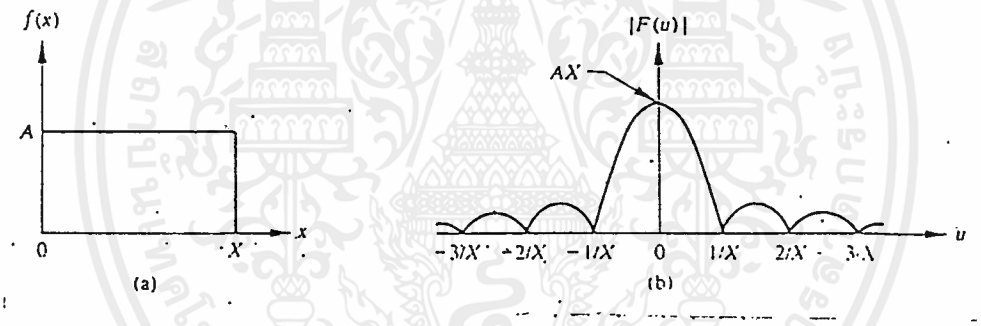
$$P(u) = |F(u)|^2$$
$$= R^2(u) + I^2(u) \quad \text{..... (2.15)}$$

P(u) คือ สเปกตรัมกำลัง (power spectrum) ของฟังก์ชั่น f(x)

ตัวแปร u ที่ใช้ในฟูเรียร์ทรานสฟอร์มเรียกว่าตัวแปรความถี่ (frequency variable) เมื่อใช้สูตรของออยเลอร์ในเทอมของเอ็กซ์โปเนนเชียล จะเปลี่ยนเป็น

$$\exp(-j2\pi ux) = \cos 2\pi ux - j\sin 2\pi ux \quad \text{..... (2.16)}$$

ถ้าเราอินติเกรทสมการที่ 2.9 จะได้ว่า F(u) ประกอบด้วยผลบวกอนันต์ของ sin และ cos และค่าแต่ละค่าของ u จะเป็นความถี่ของการตอบสนองของคู่ sin-cosin



รูปที่ 2.11 แสดงการทำฟูเรียร์ทรานสฟอร์มของฟังก์ชั่น 1 มิติ f(x)

ฟูเรียร์ทรานสฟอร์มสามารถขยายไปยังฟังก์ชั่น f(x,y) ซึ่งเป็นฟังก์ชั่น 2 มิติได้ ถ้า f(x,y) เป็นฟังก์ชั่นแบบต่อเนื่องและสามารถอินติเกรทได้ และ F(u) สามารถอินติเกรทได้ซึ่งจะได้ฟูเรียร์ทรานสฟอร์มของ f(x,y) คือ

$$\mathfrak{F}\{f(x,y)\} = F(u,v)$$
$$= \int_{-\infty}^{\infty}\int f(x,y)\exp[-j2\pi(ux + vy)]dxdy \quad \text{..... (2.17)}$$

และอินเวอร์สฟูเรียร์ทรานสฟอร์มคือ

$$\mathfrak{F}^{-1}\{f(u,v)\} = f(x,y)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$= \int_{-\infty}^{\infty} \int f(u, v)\exp[j2\pi (ux + vy)]dudv \quad \text{..... (2.18)}$$

เมื่อ u และ v เป็นตัวแปรคงที่

เช่นเดียวกันกับฟูเรียร์ทรานสฟอร์ม 1 มิติ เราจะได้

ฟูเรียร์ สเปกตรัม

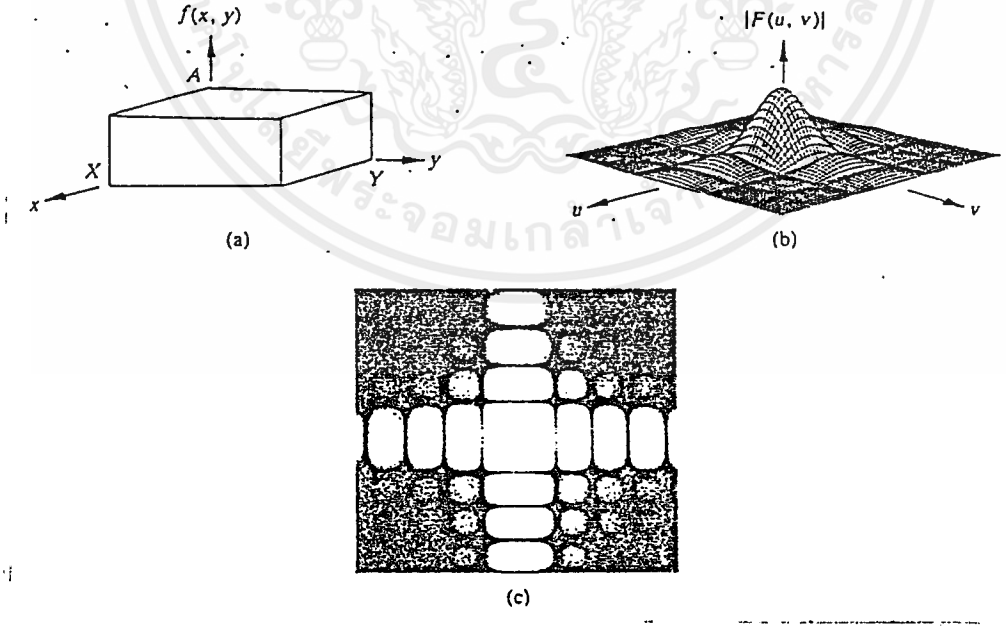
$$|F(u,v)| = [R^2(u,v) + I^2(u,v)]^{1/2} \quad \text{..... (2.19)}$$

มุมเฟส

$$\Phi(u,v) = \tan^{-1}\left[\frac{I(u,v)}{R(u,v)}\right] \quad \text{..... (2.20)}$$

และสเปกตรัมกำลัง

$$P(u,v) = |F(u,v)|^2 = R^2(u,v) + I^2(u,v) \quad \text{..... (2.21)}$$



รูปที่ 2.12 แสดงการทำฟูเรียร์ทรานฟอร์มของฟังก์ชั่น 2 มิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

b) ฟูเรียร์ สเปกตรัม

c) สเปกตรัมที่แสดงอยู่ในรูปของฟังก์ชั่นความเข้มของแสง

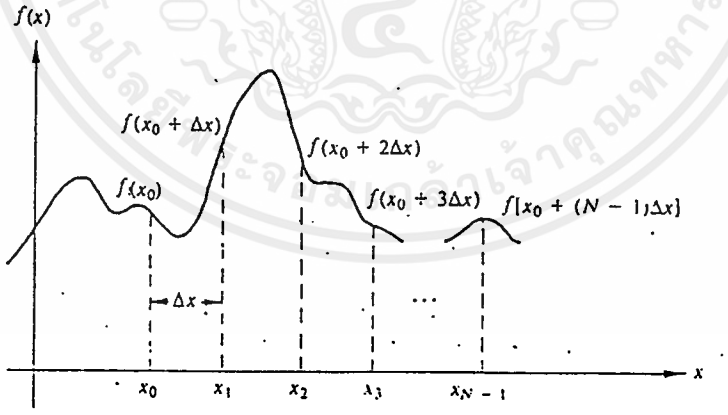
2.5.2 ฟูเรียร์ทรานสฟอร์มแบบไม่ต่อเนื่อง (discrete fourier transform)

สมมติให้ฟังก์ชั่น f(x,y) เป็นฟังก์ชั่นต่อเนื่องที่ถูกสุ่มออกเป็นลำดับ $\{f(x_0),f(x_0+\Delta x),f(x_0+2\Delta x),.....f(x_0+[N-1]\Delta x)$ โดย N เป็นจำนวนตัวอย่างในการสุ่มและสามารถเขียนฟังก์ชั่นได้

$$f(x) = f(x_0 + x\Delta x) \quad \text{..... (2.22)}$$

เมื่อ x คือ 0,1,2,...,N-1 และลำดับ { f(0),f(1),f(2),.....,f(N-1)} จะได้กำหนดช่องกว้างที่เท่ากันแต่นำมาแบ่งออกจากฟังก์ชั่นต่อเนื่องคือ

$$F(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\exp[-j2\pi ux / N] \quad \text{..... (2.23)}$$



รูปที่ 2.13 แสดงฟังก์ชั่นต่อเนื่องที่ถูกสุ่มในช่วงกว้างเท่าๆ กัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ u = 0,1,2,.....N-1 และ

$$f(x) = \frac{1}{N}\sum_{x=0}^{N-1} F(u)\exp[-j2\pi ux / N] \quad \text{..... (2.24)}$$

เมื่อ x = 0,1,2,.....N-1

ค่าของ u = 0,1,2,.....N-1 ในสมการที่ 2.24 จะสมนัยกับค่าที่สุ่มมา ณ.จุด $0,\Delta u,2\Delta u,...,(N-1)\Delta u$ ซึ่ง u และ x จะสัมพันธ์กัน

$$\Delta u = \frac{1}{N\Delta x} \quad \text{..... (2.25)}$$

ในกรณีเป็นฟังก์ชั่น 2 ตัวแปร สมการคู่ของฟูเรียร์ทรานสฟอร์มแบบไม่ต่อเนื่องจะกำหนดได้โดย

$$F(u,v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\exp[-j2\pi (ux / M + vy / N)] \quad \text{.....}$$

(2.26)

เมื่อ u = 0,1,2,...,M-1 , v = 0,1,2,....,N-1 และ

$$F(x,y) = \frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} f(u,v)\exp[-j2\pi (ux / M + vy / N)] \quad \text{..... (2.27)}$$

เมื่อ x = 0,1,2,...,M-1 , y = 0,1,2,....,N-1

การสุ่มของฟังก์ชั่นในตอนนี้จะเป็นแบบ 2 มิติ โดยจะถูกแบ่งเป็น Δx และ Δy ตามแนวแกน x และ แกน y ส่วนฟังก์ชั่น 1 มิติ ฟังก์ชั่นไม่ต่อเนื่อง f(x,y) จะถูกสุ่มออกเป็นฟังก์ชั่น f $(x0,x\Delta x,y\Delta y)$ สำหรับ x = 0,1,2,...,M-1 และ y = 0,1,2,...,N-1 เราจะได้ความสัมพันธ์ของพิกัดความถี่ (frequency domain) และพิกัดภาพ (spatial domain) ดังนี้

$$\Delta u = \frac{1}{M\Delta x} \quad \text{..... (2.28)}$$

และ

$$\Delta v = \frac{1}{N\Delta y} \quad \text{..... (2.29)}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้ง 038529

เมื่อภาพโดนสุ่มเป็นรูปสี่เหลี่ยมจัตุรัสเราจะได้ M = N

$$F(u,v) = \frac{1}{N}\sum_{x=o}^{N-1}\sum_{y=o}^{N-1} f(x,y)\exp[-j2\pi(ux + vy)/N] \quad \text{..... (2.30)}$$

เมื่อ u,v = 0,1,2,...,N-1

$$F(x,y) = \sum_{u=o}^{N-1}\sum_{v=o}^{N-1} f(u,v)\exp[-j2\pi(ux + vy)/N] \quad \text{..... (2.31)}$$

เมื่อ x,y = 0,1,2,...,N-1

สเปกตรัมฟูเรียร์ , มุมเฟส และ พลังงาน ของสเปกตรัมของฟังก์ชั่น 1 และ 2 มิติแบบไม่ต่อเนื่อง หาได้จากสมการที่ 2.13 ถึง 2.15 และสมการที่ 2.19 ถึง 2.21 ตามลำดับเพียงแต่ตัวแปรอิสระจะโดนสุ่ม

2.5.3 คุณสมบัติของฟูเรียร์ทรานสฟอร์ม 2 มิติ

2.5.3.1 การถ่ายค่าและสเกล (distribution and scaling)

จากนิยามของการทรานสฟอร์มแบบต่อเนื่องและการทรานสฟอร์มแบบไม่ต่อเนื่อง

$$\mathfrak{F}\{f_1(x,y) \cdot f_2(x,y)\} \neq \mathfrak{F}\{f_1(x,y)\} \cdot \{f_2(x,y)\} \quad \text{..... (2.32)}$$

ฟูเรียร์ทรานสฟอร์มของผลบวกของฟังก์ชั่น จะเท่ากับผลบวกของฟูเรียร์ทรานสฟอร์มของแต่ละฟังก์ชั่นส่วนปริมาณสเกลาของ a และ b

$$af(x,y) \Leftrightarrow aF(u,v) \quad \text{..... (2.33)}$$

$$f(ax,by) \Leftrightarrow \frac{1}{ab}F(u/a,v/b) \quad \text{..... (2.34)}$$

2.5.3.2 ค่าเฉลี่ย (average value)

การหาค่าเฉลี่ยของฟังก์ชั่นแบบไม่ต่อเนื่อง 2 มิติ จะหาได้โดย

$$\bar{f}(x,y) = \frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y) \quad \text{..... (2.35)}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แทนค่าของ u = v = 0 ในสมการที่ 2.30 จะได้

$$F(0,0) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y) \quad \text{..... (2.36)}$$

ดังนั้นเราจะเห็น $\bar{f}(x,y)$ สัมพันธ์กับฟูเรียร์ทรานสฟอร์มของ f(x,y) ดังสมการ

$$\bar{f}(x,y) = \frac{1}{N}F(0,0) \quad \text{..... (2.37)}$$

2.5.3.3 ลาปลาซ (laplacian)

ลาปลาซของฟังก์ชั่น 2 ตัวแปร f(x,y) กำหนดโดย

$$\nabla^2 f(x,y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad \text{..... (2.38)}$$

และจะได้ฟูเรียร์ทรานสฟอร์ม 2 มิติ

$$\Im\{\nabla^2 f(x,y)\} \Leftrightarrow -(2\pi)^2 (u^2 + v^2)F(u,v) \quad \text{..... (2.39)}$$

การแปลงลาปลาซจะใช้มากเกี่ยวกับเรื่องการแปลงภาพ

2.5.3.4 การคอนโวลูชั่น (convolution)

การคอนโวลูชั่นของฟังก์ชั่น 2 ฟังก์ชั่น f(x) และ g(x) คือ f(x)*g(x) จะถูกกำหนดโดย

$$f(x)*g(x) = \int_{-\infty}^{\infty} f(\alpha)g(x-\alpha)d\alpha \quad \text{..... (2.40)}$$

เมื่อ α เป็นตัวแปรหุ่นของการอินติเกรท จากตัวอย่างจะทำการคอนโวลูชั่นระหว่างฟังก์ชั่น f(x) และ g(x) ดังรูปที่ 2.14(a) และ 2.14(b) ตามลำดับก่อนที่จะทำการอินติเกรทจะต้องทำฟังก์ชั่นให้อยู่ในรูปของ $g(x-\alpha)$ ก่อน ตามขั้นตอนในรูปที่ 2.14(c) และ 2.14(d) จากนั้นคูณฟังก์ชั่น $g(-\alpha)$ ด้วย $g(x-\alpha)$ แล้วอินทิกรัล ผลคูณจาก ∞ ถึง $-\infty$

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ ได้รับอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลคูณของ $f(\alpha)$ และ $g(x-\alpha)$ คือส่วนที่แรเงาในรูปที่ 2.14(e) เมื่อให้ $0 \leq x \leq$ เมื่อผลคูณเป็นศูนย์สำหรับค่า α ที่อยู่นอกช่วง [0,x] เราจะพบว่า $f(x)*g(x) = x/2$ สำหรับในรูปที่ 2.14 (f) ค่า x จะอยู่ในช่วง [1,2] ในกรณีนี้จะได้ $f(x)*g(x) = (1-x/2)$ ดังนั้นแสดงว่า $f(\alpha)g(x-\alpha)$ จะเป็นศูนย์สำหรับค่า x ที่อยู่นอกช่วง [0,2] จะได้

$$f(x)*g(x) = \begin{cases} x/2 & 0 \leq x \leq 1 \\ 1 - x/2 & 1 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

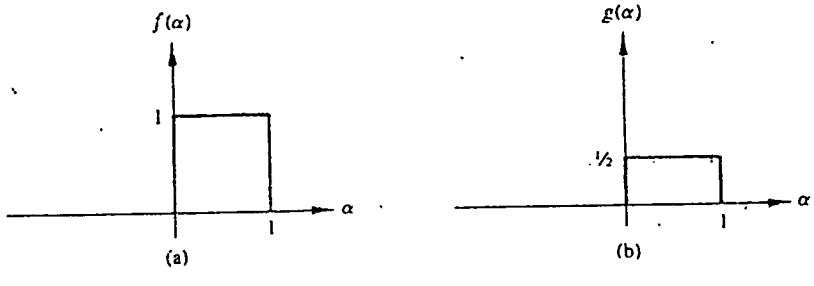
ผลที่ได้จะอยู่ในรูปที่ 2.14(g)

การคอนโวลูชั่นของฟังก์ชั่น f(x) ที่เป็นฟังก์ชั่นอิมพัลซ์ (impulse function) $\delta(x-x_0)$ จะนิยามโดยความสัมพันธ์

$$\int_{-\infty}^{\infty} f(x)\delta(x - x_0)dx = f(x_0) \quad \text{..... (2.41)}$$

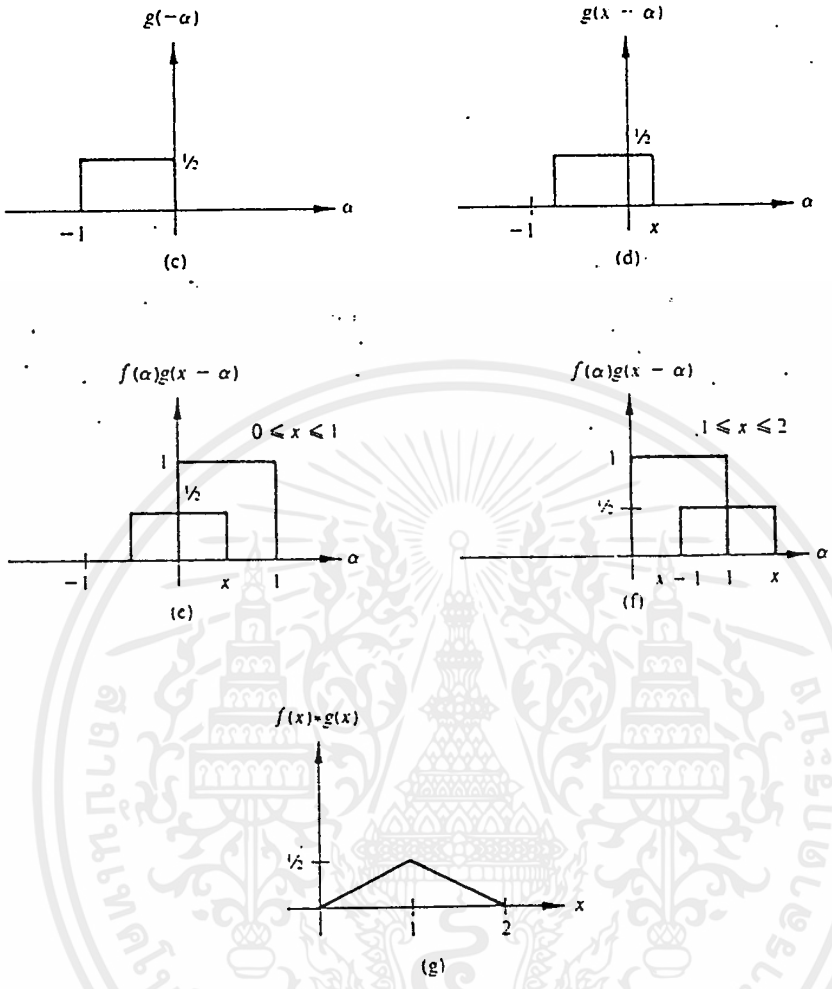
ฟังก์ชั่น $\delta(x-x_0)$ จะมีพื้นที่เท่ากับ 1

$$\int_{-\infty}^{\infty} \delta(x - x_0)dx = \int_{x_0^-}^{x_0^+} \delta(x - x_0)dx = 1 \quad \text{.....(2.42)}$$



รูปที่ 2.14 กราฟแสดงการทำคอนโวลูชั่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.14 กราฟแสดงการทำคอนโวลูชั่น (ต่อ)

ส่วนที่แรเงาแสดงถึงพื้นที่ที่ผลคูณไม่เท่ากับหนึ่งแบบ 1 มิติคือ

$$f(x)*g(x) \Leftrightarrow F(u)G(u) \quad \text{..... (2.43)}$$

และ

$$f(x)g(x) \Leftrightarrow F(u)*G(u) \quad \text{..... (2.44)}$$

เมื่อ F(u) **คือ ฟูเรียร์ทรานสฟอร์มของ** f(x)

G(u) **คือ ฟูเรียร์ทรานสฟอร์มของ** g(x)

ได้ฟังก์ชั่นเป็นแบบ 2 มิติจะได้กฎของการทำคอนโวลูชั่นดังนี้

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$f(x,y)*g(x,y) \Leftrightarrow F(u,v)G(u,v) \quad \text{..... (2.45)}$$

และ

$$f(x,y)g(x,y) \Leftrightarrow F(u,v)*G(u,v) \quad \text{..... (2.46)}$$

2.6 การทำภาพให้ชัดเจนขึ้น (image enhancement)

การทำให้ชัดเจนขึ้นมีวัตถุประสงค์ก็เพื่อต้องการให้ภาพที่จะนำมาประมวลผล มีผลลัพธ์ที่เหมาะสมกับงานที่จะนำไปประยุกต์ และการใช้เทคนิคหนึ่งเพื่อทำให้ภาพคมชัดขึ้นก็จะเหมาะกับงานหนึ่งๆ นั้น การพิจารณาจะแบ่งเป็น 2 ลักษณะคือ วิธีในโดเมนเชิงความถี่ (frequency-domain method) และวิธีในโดเมนสเปเทียล (spatial-domain method) เทคนิคอันแรกจะใช้เทคนิคของการแปลงฟูเรียร์ (fourier transform) ของภาพ ส่วนเทคนิคอย่างที่สองจะใช้การอ้างถึงระนาบภาพของตัวมันเองและการปฏิบัติการโดยตรงกับพิกเซลของภาพ

2.6.1 วิธีในโดเมนสเปเทียล (spatial-domain method)

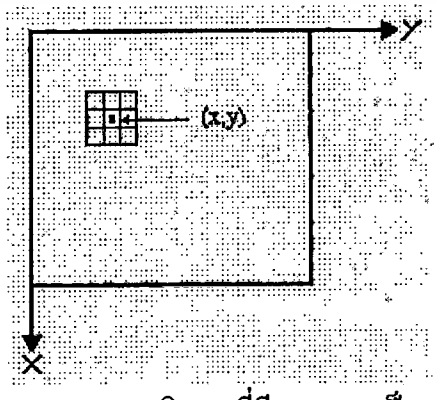
คำว่า โดเมนสเปเทียล หมายถึงการรวมกันของพิกเซลของภาพและวิธีในโดเมนสเปเทียลก็คือการปฏิบัติการโดยตรงกับพิกเซล ฟังก์ชั่นของการประมวลผลข้อมูลภาพในโดเมนสเปเทียลจะเขียนได้คือ

$$g(x,y) = T[f(x,y)] \quad \text{..... (2.47)}$$

เมื่อ f(x,y) คือ ฟังก์ชั่นของภาพอินพุท

g(x,y) คือ ภาพที่ผ่านการประมวลผลแล้ว

T คือ ตัวปฏิบัติการที่กระทำกับฟังก์ชั่น f(x,y) ซึ่งจะทำการปฏิบัติการในย่านขอบเขตของ (x,y)



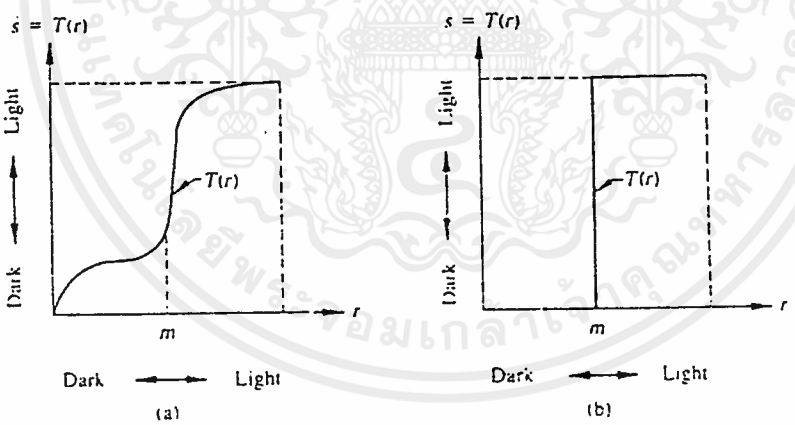
รูปที่ 2.15 แสดงภาพย่อยขนาด 3x3 พิกเซลที่มีจุด (x,y) เป็นจุดศูนย์กลาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ศูนย์กลางอยู่ที่ (x,y) ซึ่งแสดงดังรูปที่ 2.15 จุดศูนย์กลางอยู่ที่ของภาพย่อยจะเลื่อนจากพิกเซลหนึ่งไปยังพิกเซลหนึ่ง หลักในการกำหนดขอบเขตของ (x,y) โดยพิจารณาภาพย่อยที่เป็นสี่เหลี่ยมจัตุรัสหรือสี่เหลี่ยมผืนผ้าที่มีเซลหนึ่ง โดยเริ่มพิกเซลแรกสุดที่มุมบนซ้ายแล้วกระทำกับ จุด (x,y) แต่ละจุดซึ่งจะได้ค่าของ g ณ.ตำแหน่ง (x,y) นั้นๆ ส่วนรูปฟอร์มของ T เมื่อภาพมีขนาด 1x1 ในกรณีนี้ค่า g จะขึ้นอยู่กับค่าของฟังก์ชั่น f ที่จุด (x,y) และ T จะกลายเป็นฟังก์ชั่นของระดับเทา (gray level) นั่นคือ

$$s = T(r) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2.48)$$

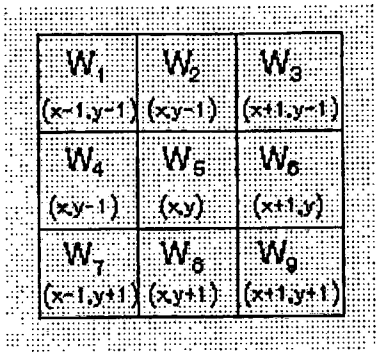
เมื่อ r และ s เป็นตัวแปรที่ใช้แทนระดับเทาของ f(x,y) และ g(x,y) ที่จุด (x,y) T(r) แสดงดังรูปที่ 2.16(a) ผลของการแปลงจะทำให้ระดับความสว่างของภาพสูงกว่าแบ่งออกเป็น 2 ส่วน คือ ส่วนที่สว่าง (light) และส่วนที่มืด (dark) โดยถูกแบ่งที่จุด m ส่วนที่ค่า r อยู่เหนือจุด m จะสว่าง และส่วนที่ค่า r ต่ำกว่าจุด m จุดมืด เทคนิคอันนี้จะเรียกว่า การแบ่งความแตกต่างของแสง (contrast stretching) ซึ่งจะเห็นดังรูปที่ 2.16(b) โดย T(r) จะให้ภาพออกมาเป็น 2 ระดับ (binary image)



รูปที่ 2.16 ฟังก์ชั่นการแปลงระดับเทาสำหรับทำให้ภาพเป็น 2 ระดับ

จากรูปที่ 2.17 ถ้าเราให้ w1,w2,...,w9 เป็นสัมประสิทธิ์ของมาส์ก (mask) และพิจารณาย่านรอบๆ จุด(x,y) 8 จุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.17 รูปทั่วไปของมาส์ก 3x3 ที่จะไปกระทำกับพิกเซลของภาพตามจุดต่างๆ จะทำการปฏิบัติการได้

$$T[f(x,y)] = w_1f(x-1,y-1) + w_2f(x,y-1) + w_3f(x+1,y-1) + w_4f(x-1,y) + w_5f(x,y) + w_6f(x+1,y) + w_7f(x-1,y+1) + w_8f(x,y+1) + w_9f(x+1,y+1) \quad \text{..... (2.49)}$$

ซึ่งถ้าเป็นมาส์กขนาดใหญ่กว่า รูปแบบของฟังก์ชั่นก็จะมีลักษณะคล้ายกับมาส์กแบบ 3x3 สมการที่ 2.49 จะเปลี่ยนสัมประสิทธิ์ฟังก์ชั่นของมาส์ก

2.6.2 วิธีในโดเมนเชิงความถี่ (frequency-domain method)

ทฤษฎีการคอนโวลูชั่นเป็นทฤษฎีที่นำมาใช้กับวิธีโดเมนเชิงความถี่ ให้ g (x,y) คือฟังก์ชั่นที่ทำการคอนโวลูชั่นโดยฟังก์ชั่นของ f(x,y) และตัวปฏิบัติการที่ไม่เปลี่ยนตำแหน่ง (position-invariant operator) นั่นคือ

$$g(x,y) = h(x,y)*f(x,y) \quad \text{..... (2.50)}$$

จากทฤษฎีของการคอนโวลูชั่นจะได้โดเมนเชิงความถี่ในรูปของความสัมพันธ์

$$G(u,v) = H(u,v)f(u,v) \quad \text{..... (2.51)}$$

เมื่อ G,H และ F คือฟูเรียร์ทรานสฟอร์มของ g,h และ f ตามลำดับ ตัวแปลง (transform) บางครั้งเรียกว่าทรานสเฟอร์ฟังก์ชั่น (transfer function) ของการประมวลผล ดังนั้น

$$g(x,y) = \mathfrak{I}^{-1}[H(u,v)F(u,v)] \quad \text{..... (2.52)}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6.3 การทำให้ภาพมีระดับที่ราบเรียบ (image smoothing)

2.6.3.1 การหาค่าเฉลี่ยรอบย่าน (neighborhood average)

การเฉลี่ยค่ารอบย่านเป็นเทคนิคโดยตรงที่ใช้ในโดเมนสเปเทียล เพื่อทำให้สัญญาณข้อมูลภาพเกิดความราบเรียบ ให้ภาพ f(x,y) มีขนาด NxN และ g(x,y) คือ ผลของการกระทำให้ภาพราบเรียบที่ระดับเทาของพิกเซล (x,y) ซึ่งจะได้ความสัมพันธ์

$$g(x,y) = \frac{1}{M} \sum_{(n,m) \in S} f(n, m) \quad \text{..... (2.53)}$$

เมื่อ x,y = 0,1,...,N-1

s คือ เซตของโคออดิเนทของจุดที่อยู่รอบๆ จุด (x,y) รวมทั้งจุด (x,y) ด้วย

M คือ ผลรวมของจำนวนของจุดในรอบๆ ย่านนั้น

2.6.3.2 ฟิลเตอร์แบบมัธยฐาน (median filtering)

ภาพที่เกิดความเบลอร์ที่ขอบภาพและมีสัญญาณรบกวนอาจแก้ไขได้โดยการตัดค่าเทรสโฮล (threshold) ซึ่งโดยทั่วไปจะใช้วิธีการลองผิดลองถูก (trial and error) แต่ถ้าเราใช้ฟิลเตอร์แบบมัธยฐานแทนระดับเทาของแต่ละพิกเซลที่อยู่รอบๆ ย่านแทนการเฉลี่ย จะทำให้ภาพมีประสิทธิภาพดียิ่งขึ้น การหาค่ามัธยฐาน (median,M) คือ เซตของค่าที่มีค่าในเซตครึ่งหนึ่งมากกว่า M และอีกครึ่งหนึ่งน้อยกว่า M สามารถทำได้โดยเรียงข้อมูลจากน้อยไปหามาก แล้วนำค่าที่อยู่ตรงกลางมาใช้แทนค่า M สมมุติให้ย่านขนาด 3x3 มีข้อมูล (10,20,20,20,15,20,20,25,100) จะสามารถเรียงข้อมูลจากน้อยไปหามากได้เป็น (10,15,20,20,20,20,20,25,100) จะได้ค่ามัธยฐานคือ 20 นั่นเอง

2.6.3.3 ฟิลเตอร์แบบความถี่ต่ำ (lowpass filtering)

ขอบภาพและส่วนที่คมของภาพนั้นในระดับเทาของภาพจะมีส่วนของสัญญาณความถี่สูงของการแปลงฟูเรียร์อยู่ ซึ่งจะทำให้ภาพเบลอร์ ดังนั้นจึงต้องกำจัดช่วงความถี่สูงที่ไม่ต้องการออกไป จากโดเมนเชิงความถี่

$$G(u,v) = H(u,v)F(u,v) \quad \text{..... (2.54)}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ F(u,v) คือภาพที่ต้องการทำให้เรียบโดยการแปลงฟูเรียร์ ปัญหาอยู่ที่การเลือกฟังก์ชั่น H(u,v) ให้ผลที่ได้ออกมาเป็น G(u,v) โดยการลดทอนส่วนประกอบของความถี่สูงของ F(u,v) ออกไปการแปลงอินเวอร์สฟูเรียร์ของ G(u,v) จะได้ g(x,y) ที่มีความราบเรียบขึ้นทั้งนี้ก็เพราะส่วนประกอบของความถี่สูงของจะโดนกรองออกไปและข้อมูลที่ผ่านไปได้จะเป็นช่วงของสัญญาณความถี่ต่ำ ซึ่งความถี่ต่ำที่ผ่านไปได้จะไม่โดนลดทอนลง วิธีนี้จึงเรียกว่าการกรองความถี่ต่ำ ฟังก์ชั่น H(u,v) หมายถึงฟังก์ชั่นการส่งถ่ายค่าของการกรอง (filter transfer function)

2.6.3.4 ฟิลเตอร์แบบอุดมคติ (ideal filter)

ฟิลเตอร์ 2 มิติ แบบอุดมคติที่จะมีการส่งถ่ายฟังก์ชั่น โดยมีความสัมพันธ์

$$H(u,v) = \begin{cases} 1 & D(u,v) \le D_0 \\ 0 & D(u,v) \ge D_0 \end{cases} \quad \text{..... (2.55)}$$

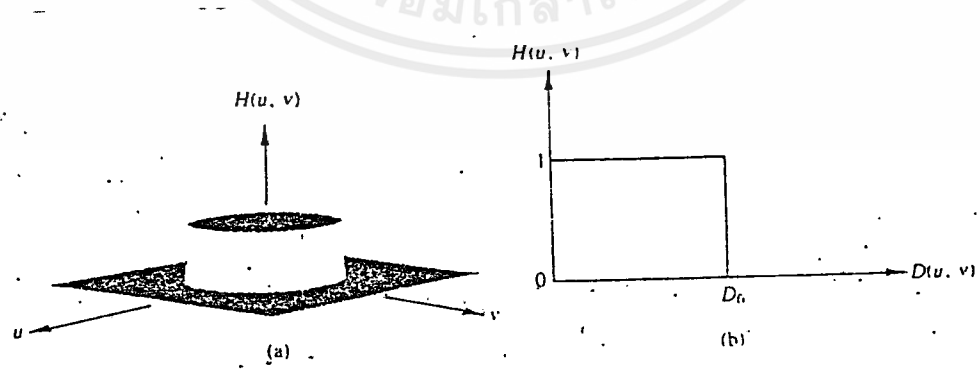
เมื่อ D_0 คือค่าจำเพาะที่ไม่เป็นลบ

D(u,v) คือระยะทางจากจุด (u,v) ถึงจุดกำเนิดในระนาบความถี่

นั่นคือ

$$D(u,v) = (u^2 + v^2)^{1/2} \quad \text{..... (2.56)}$$

สำหรับฟังก์ชั่น 3 มิติ จะสามารถพล็อตค่าของ H(u,v) เทียบกับ u และ v แสดงดังรูป 2.18(a) ตามอุดมคติความถี่ที่โดนกรองจะอยู่ในรัศมี D_0 ของวงกลมซึ่งความถี่ในช่วงนี้จะผ่านไปได้โดยไม่โดนลดทอนส่วนความถี่ที่อยู่นอกช่วงรัศมีของวงกลมจะโดนลดทอนหรือโดนตัดออก



รูปที่ 2.18 ฟิลเตอร์แบบอุดมคติกรองความถี่ต่ำ

a) ฟังก์ชั่น 3 มิติที่นำมากรองความถี่

b) รูปหน้าตัด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6.4 การทำให้ภาพมีความคม (image sharpening)

2.6.4.1 การทำภาพให้มีความคมโดยใช้ดิฟเฟอเรนเชียล (sharpening by differentiation)

โดยทั่วไปการใช้วิธีดิฟเฟอเรนเชียลในการประมวลผลภาพจะใช้การทำเกรเดียนท์ (gradient) ให้ภาพ คือ ฟังก์ชั่น f(x,y) เกรเดียนท์ของฟังก์ชั่น f ที่โคออดิเนท (x,y) จะนิยามโดยเวคเตอร์

$$G[f(x,y)] = \begin{bmatrix} \partial f \\ \partial x \\ \partial f \\ \partial y \end{bmatrix} \quad \text{..... (2.57)}$$

คุณสมบัติของการเกรเดียนท์คือ

1. เวคเตอร์ G[f(x,y)] จะเป็นทิศทางของอัตราการเพิ่มสูงสุดของฟังก์ชั่น f(x,y)
2. ขนาดของเวคเตอร์ G[f(x,y)] จะถูกกำหนดโดย G[f(x,y)] และจะได้

$$G[f(x,y)] = mag[G]$$

$$= [(\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2]^{1/2} \quad \text{..... (2.58)}$$

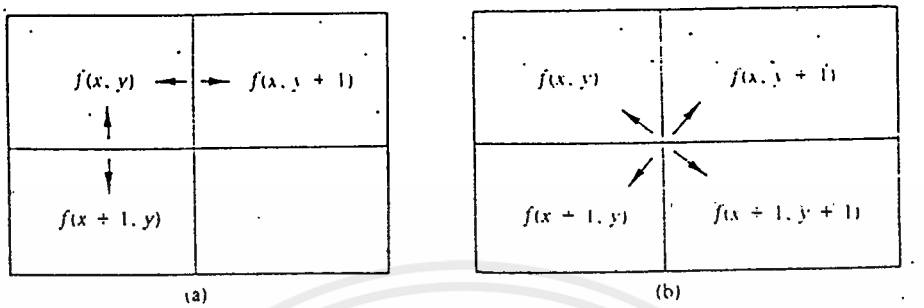
เท่ากับอัตราการเพิ่มสูงสุดของฟังก์ชั่น f(x,y) ต่อระยะทาง 1 หน่วยในทิศทางของ G ในการประมวลผลภาพอนุพันธ์ของสมการ 2.58 จะประมาณโดย

$$G[f(x,y)] \cong \{[f(x,y)-f(x+1,y)]^2 + [f(x,y)-f(x,y+1)]^2\}^{1/2} \quad \text{..... (2.59)}$$

ผลที่ได้จะเป็นค่าสัมบูรณ์ นั่นคือ

$$G[f(x,y)] \cong f(x,y)-f(x+1,y) + f(x,y)-f(x,y+1) \quad \text{..... (2.60)}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.19 แสดงการทำเกรเดียนท์ของฟังก์ชั่น 2 มิติ

ความสัมพันธ์ระหว่างพิกเซลของสมการ 2.59 และ 2.60 แสดงดังรูปที่ 2.19(a) ซึ่งให้ภาพมีขนาด NxN เราจะไม่สามารถทำเกรเดียนท์ที่แถวสุดท้าย (x=N) หรือคอลัมน์สุดท้าย (y=N) การใช้งานอื่นๆ โดยประมาณบางครั้งเรียกว่า Roberts gradient ใช้การดิฟเฟอเรนเชียลทแยง (cross differential) แสดงดังรูปที่ 2.19(b) ซึ่งจะได้ความสัมพันธ์

$$G[f(x,y)] \cong \{[f(x,y)-f(x+1,y+1)]^2 + [f(x+1,y)-f(x,y+1)]^2\}^{1/2} \quad \text{..... (2.61)}$$

หรือใช้ค่าสัมบูรณ์

$$G[f(x,y)] \cong f(x,y)-f(x+1,y+1) + f(x+1,y)-f(x,y+1) \quad \text{..... (2.62)}$$

2.6.4.2 ฟิลเตอร์แบบกรองความถี่สูง (highpass filtering)

จากการทำการกรองแบบความถี่ต่ำอาจทำให้ภาพเบลอร์ (blurred) ได้อันเนื่องมาจากเราลดทอนส่วนประกอบของความถี่สูงของการแปลงฟูเรียร์ซึ่งจะทำให้ขอบภาพเกิดการเปลี่ยนแปลงชองระดับเทา การทำให้ภาพมีความคมในโดเมนเชิงความถี่โดยวิธีการกรองแบบความถี่สูงจะทำการลดทอนส่วนประกอบของสัญญาณความถี่ต่ำ โดยปราศจากการกระทบกระเทือนต่อสัญญาณความถี่สูง ในการแปลงฟูเรียร์เราจะพิจารณาส่วนที่ต่ำกว่าความถี่สูงก่อนจะทำการพิจารณา เราจะทำการเลื่อนเฟสศูนย์ (zero-phase shift) เพื่อให้เกิดการสมมาตรและทำให้สมบูรณ์ได้โดย หน้าตัดจะขยายฟังก์ชั่นของระยะทางจากจุดกำเนิดของการแปลงฟูเรียร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

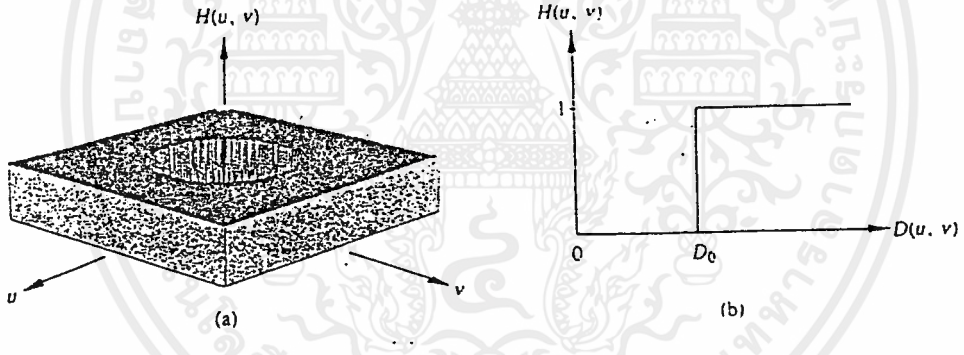
2.6.4.3 ฟิลเตอร์แบบอุดมคติ (ideal filter)

ฟังก์ชั่น 2 มิติของการกรองแบบอุดมคติ จะมีการส่งถ่ายค่าตามความสัมพันธ์

$$H(u,v) = \begin{cases} 0 & D(u,v) \le D_0 \\ 1 & D(u,v) > D_0 \end{cases} \quad \text{..... (2.63)}$$

เมื่อ D_0 คือระยะที่ตัดโดยวัดจากจุดกำเนิดของระนาบความถี่

โดยแสดงดังรูปที่ 2.20 การกรองแบบความถี่สูงจะมีคุณสมบัติที่ตรงกันข้ามกับการกรองแบบความถี่ต่ำโดยความถี่ที่อยู่นอกรัศมีของวงกลม D_0 จะผ่านไปได้โดยไม่โดนลดทอน ส่วนความถี่ที่อยู่ในช่วงของวงกลมจะผ่านไปไม่ได้



a) ฟังก์ชั่น 3 มิติที่นำมากรองความถี่สูง b) รูปหน้าตัด

รูปที่ 2.20 แสดงฟิลเตอร์แบบอุดมคติกรองความถี่สูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

ทฤษฎีและหลักการต่าง ๆ ที่นำมาใช้

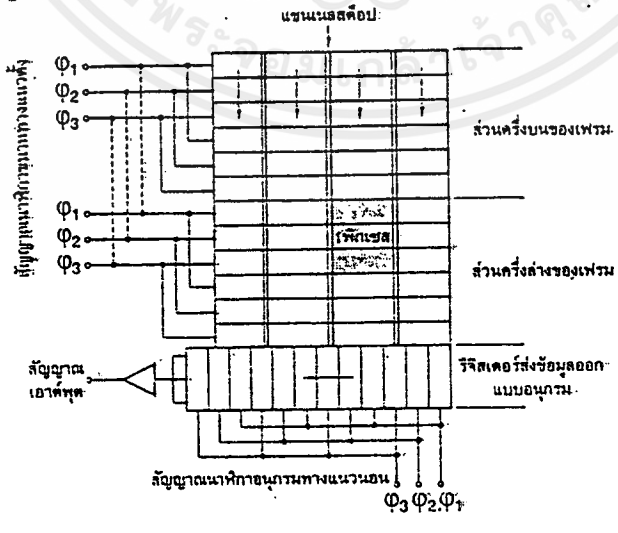
3.1 หลักการทำงานของกล้องCCD (Charge-Coupled Device) [6]

ทำหน้าที่เป็นตัวทำการเปลี่ยนพลังงานแสงไปสู่พลังงานไฟฟ้าเพื่อนำไปเข้าสู่ระบบอิมเมจโปรเซสซิ่ง ซึ่งถ้าพิจารณาดูแล้วอุปกรณ์จำพวกนี้มีอยู่มากมายหลายชนิดด้วยกันแต่ CCD มีคุณสมบัติที่ไม่ธรรมดารวมถึงการทำงานก็ไม่ธรรมดาไม่มีการใช้ปืนอิเล็กตรอนเป็นส่วนประกอบ แต่กับใช้การเปลี่ยนเฟสควบคุมการทำงานแทน

อุปกรณ์ CCD ถือกำเนิดขึ้นเมื่อประมาณ 20 ปีที่แล้วซึ่งขณะนั้นยังอยู่ในรูปแบบของอุปกรณ์หน่วยความจำรอม(random access memory)แต่ CCD มีประโยชน์นำไปใช้งานมากกว่านั้น ซึ่งเป็นได้ทั้งดีเลย์ไลน์ ,ส่วนประมวลผลของสัญญาณ,และที่สำคัญที่สุดเป็นอุปกรณ์ตรวจจับแสงคุณภาพสูง CCD มีหลักการทำงานต่อไปนี้

จากรูปที่ 3.1 แสดงให้เห็นถึงโครงสร้างทั่วไป และการคัปปลิ้งสัญญาณของอุปกรณ์ตัวนี้จะเห็นได้ทันทีว่ามีการแบ่งพื้นที่ออกเป็นส่วนเล็กๆมากมาย ซึ่งแต่ละชิ้นส่วนนั้นเรียกว่า พิกเซล (pixel)

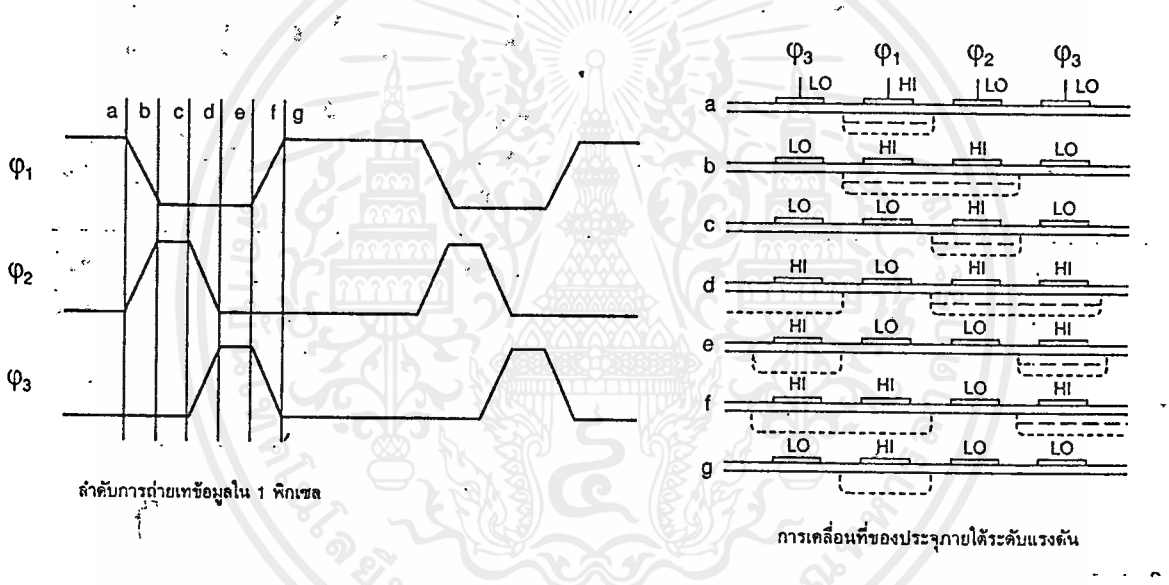
และถ้าสังเกตดูโครงสร้างแล้วจะเห็นได้ว่า CCD ไม่มีการแบ่งแต่ละพิกเซลออกเป็นตำแหน่งแอดเดสเหมือนหน่วยความจำชนิดอื่น แต่มีการแบ่งออกเป็นแถว(row) และหลัก(column) โดยมีส่วนขั้นกลางที่เรียกว่า แชลเนลสต็อป (channel stop) กันระหว่างแต่ละหลักอยู่ ภายในแต่ละพิกเซลจะมีประจุอยู่ตัวหนึ่งตัวที่ทำให้เกิดภาพ และจะมีการถ่ายเทเมื่อมีแสงมาตกกระทบการ



จากรูปที่ 3.1 แสดงให้เห็นถึงโครงสร้างทั่วไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

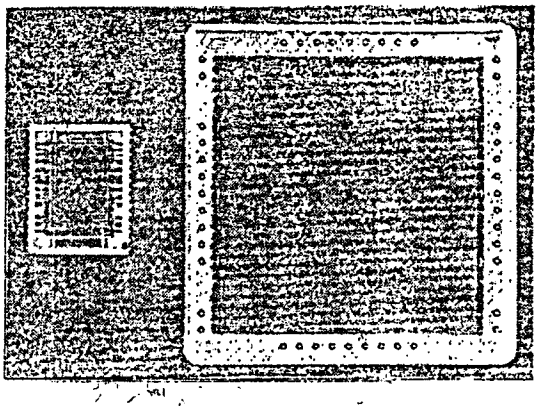
ถ่ายเทก็ย่อมมีตัวพาหรือขีดเหนี่ยวซึ่ง CCD ใช้หลักการหมุนเฟส(phase clock voltage) มาใช้ แต่ละแถวจะมีการกำหนดสถานะทางไฟฟ้าด้วยการหมุนเฟส 3 เฟส ที่มีความสัมพันธ์กันของช่วงเวลาและแรงดันและด้วยการทำงานเช่นนี้ประจุแต่ละพิกเซลจึงมีการถ่ายเทจากจุดหนึ่งไปอีกจุดหนึ่งได้ เมื่อประจุเคลื่อนมาจนถึงสุดขอบของแถวก็จะมีการส่งผ่านให้กับรีจิสเตอร์รับข้อมูลต่อไป เพื่อนำส่งผ่านข้อมูลสู่ภายนอก โดยการทยอยข้อมูลแบบอนุกรมของรีจิสเตอร์ และถูกควบคุมด้วยการหมุนเฟสเช่นเดียวกัน ดังนั้นผลรวมของการหมุนเฟสทั้งด้านแนวนอนและแนวตั้ง จึงสามารถนำประจุออกสู่ภายนอกได้เพื่อใช้ในการประมวณผลและสร้างภาพต่อไป



รูปที่ 3.2 แสดงให้เห็นถึงลักษณะรูปคลื่นของแต่ละเฟสที่มีความสัมพันธ์กันและลักษณะการเคลื่อนที่ของประจุเมื่อมีการถ่ายเท

สำหรับรายละเอียดของภาพและความไวต่อแสงสว่างน้อยๆ ขึ้นอยู่กับจำนวนพิกเซลบนตัว CCD ซึ่งมีให้เลือกหลายขนาดตามความต้องการตั้งแต่ 385x578 , 1500x1500 , 2048x2048 พิกเซลเป็นต้น และราคาของแต่ละขนาดต่างกัน ยิ่งมีจำนวนพิกเซลมากเท่าใด ประสิทธิ์ภาพก็ยิ่งสูงเท่านั้น และที่สำคัญราคาก็ยิ่งสูงตามไปด้วย

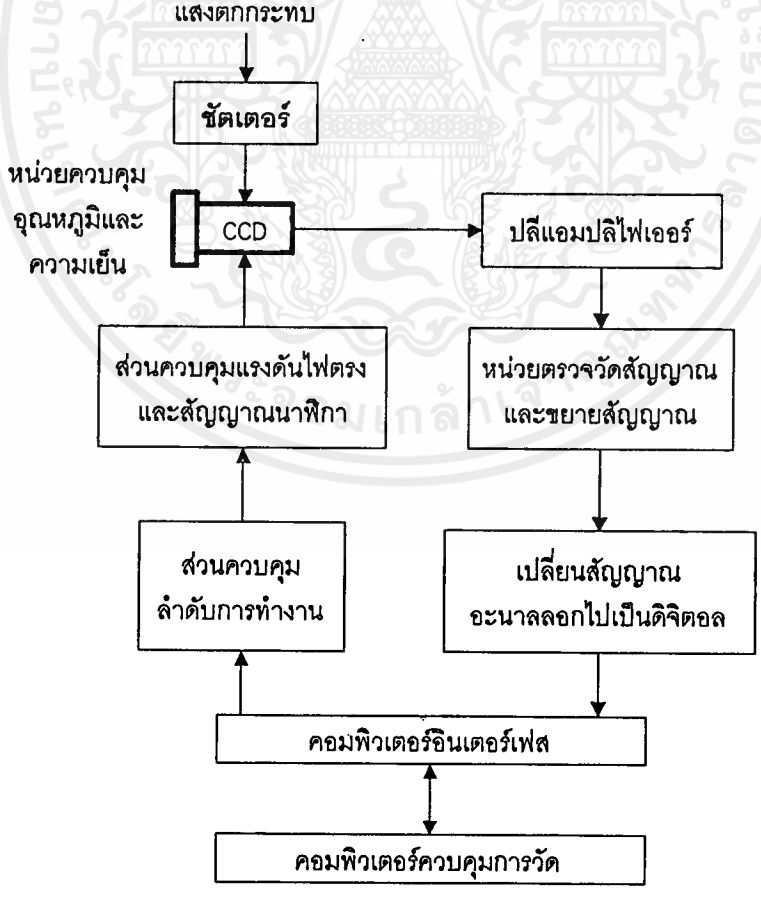
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.3 แสดงให้เห็นถึงลักษณะและขนาดของอุปกรณ์ CCD ในขนาดต่างๆ

3.1.1การนำภาพ CCDไปใช้ในงานถ่ายภาพ

โดยปกติอุปกรณ์ CCD สามารถทำงานได้ด้วยอุณหภูมิห้องปกติ(room temperature) สำหรับใช้ในงานถ่ายภาพที่มีแสงพอประมาณ เช่น การถ่ายภาพด้วยกล้องโทรทัศน์ กล้องถ่ายรูป และกล้องวีดีโอ ซึ่งก็ทำงานได้เป็นอย่างดี แต่โดยทั่วไปอุปกรณ์อิเล็กทรอนิกส์เกือบทุกชนิดโดยเฉพาะจำพวกสารกึ่งตัวนำ มักมีการรั่วไหลของกระแสส่วนน้อยๆเกิดขึ้น CCD ก็มีเช่นเดียวกัน



รูปที่ 3.4 แสดงส่วนประกอบหลักของระบบ CCD

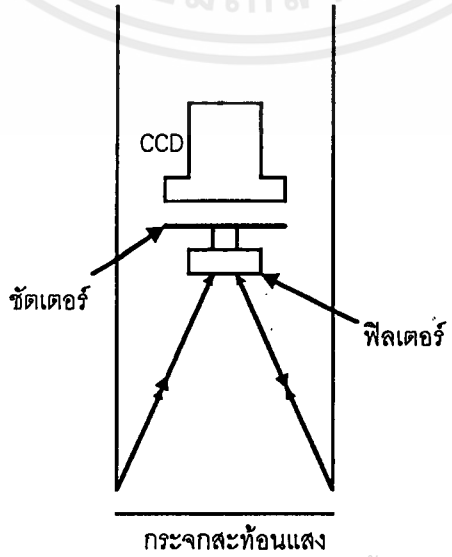
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การถ่ายภาพที่มีแสงสว่างน้อยๆนั้น จึงจำเป็นต้องคำนึงถึงกระแสรั่วไหลส่วนนี้ เพื่อให้ CCD สามารถเก็บแสงได้ดีที่สุด วิธีแก้ไขก็คือพยายามลดอุณหภูมิให้ต่ำมากๆ ประมาณ -120 องศาเซลเซียส โดยใช้ไนโตรเจนเหลวให้ความเย็นอย่างรุนแรงกับ CCD กระแสรั่วไหลจะลดการสูญเสียเหลือเพียง 1 อิเล็กตรอน/ชั่วโมง/พิกเซล ในขณะที่มีความจุอิเล็กตรอนสูงถึง 100,000 อิเล็กตรอน/พิกเซล นอกจากการแก้ปัญหาเนื่องจากกระแสรั่วไหลแล้ว ยังต้องมีส่วนอื่นอีกมากมายประกอบอยู่ด้วยดังรูปที่ 3.4 ระบบชัตเตอร์ที่ใช้ทำงานด้วยเครื่องกลไฟฟ้า ควบคุมโดยคอมพิวเตอร์ ซึ่งควบคุมการตั้งเวลาเปิดหน้ากล้อง ความไวในการถ่ายภาพเป็นอีกสิ่งหนึ่งที่สำคัญ ดังนั้นการตรวจวัดแสงจึงจำเป็นต้องใช้อุปกรณ์ที่มีสัญญาณรบกวนต่ำและมีระบบป้องกันที่ดี โดยเฉพาะใน ส่วนของวงจรภายนอก ซึ่งต้องมีการออกแบบวงจรให้ทำงานได้อย่างมีประสิทธิภาพ และที่สำคัญต้องทำให้เกิดสัญญาณรบกวนต่ำสุด อาทิเช่น วงจรปรีแอมปลิไฟร์ และวงจรแซมปลิ้ง

จากรูปที่ 3.4 เห็นได้ว่านอกจากวงจรที่กล่าวมาแล้ว ยังมีส่วนอื่นที่ร่วมอยู่ในกระบวนการอีกหลายส่วน เช่น ส่วนของวงจรขยายและวัดสัญญาณ ,วงจรแปลงสัญญาณอะนาลอกไปสู่ดิจิตอลเพื่ออินเตอร์เฟสกับเครื่องคอมพิวเตอร์ และวงจรกำเนิดแรงดันไบแอสไฟตรงและสัญญาณนาฬิกาเพื่อควบคุมการทำงานของ CCD ซึ่งก็ถูกควบคุมด้วยคอมพิวเตอร์ทั้งหมด

3.1.2 การติดตั้งและใช้งาน

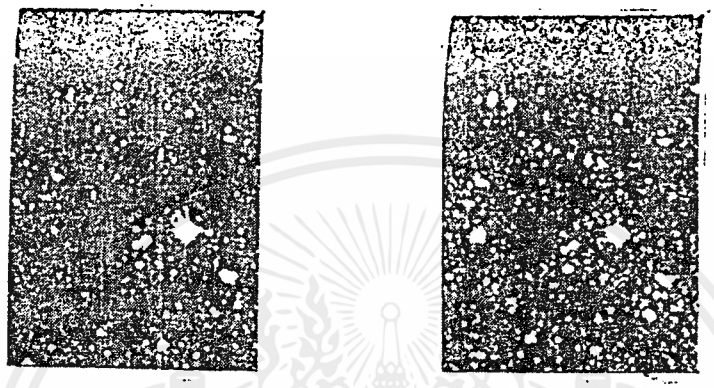
โดยปกติเราสามารถนำอุปกรณ์ CCD มาติดตั้ง และใช้งานแบบง่ายๆ ก็ได้เช่นกัน แต่ทั้งประสิทธิภาพความละเอียด และความไวต่อแสงนั้นจะด้อยกว่าเมื่อติดตั้งแบบเต็มระบบ ดังที่ได้กล่าวมาแล้ว การติดแบบง่ายๆ นี้คือ ออกแบบติดตั้งให้CCD รับแสงโดยตรงเลย แต่ต้องมีการปรับโฟกัสของกล้องโทรทัศน์ให้ดี เพื่อให้ได้ภาพที่คมชัดและมีรายละเอียดชัดเจนที่สุด



รูปที่ 3.5 แสดงให้เห็นถึงการติดตั้ง CCD แบบแสงโดยตรง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.5 แสดงให้เห็นถึงการติดตั้ง CCD แบบง่ายๆ ชัตเตอร์ที่อาจควบคุมด้วยมือหรือเครื่องกลก็ได้ ซึ่งต้องมีการจับเวลาเปิดหน้ากล้องที่เหมาะสมด้วยสำหรับฟิลเตอร์เป็นตัวทำหน้าที่กรองความถี่ของคลื่นแสงที่ไม่ต้องการทิ้งไป หรือเป็นตัวกำหนดความยาวคลื่นที่ต้องการนั้นเอง



(ก) (ข)

รูปที่ 3.6 แสดงให้เห็นถึงภาพที่ถ่ายจากอุปกรณ์ CCD แต่ใช้ฟิลเตอร์ช่วงความยาวคลื่นต่างๆ

รูปที่ 3.6 (ก) ถ่ายโดยใช้ฟิลเตอร์ย่านสีแดง(R-band) และจะเห็นได้ชัดว่าเมื่อถ่ายโดยใช้ฟิลเตอร์ย่านอินฟราเรด

รูปที่ 3.6 (ข) จะสังเกตเห็นจำนวนดวงดาวที่มีความยาวคลื่นในช่วงอินฟราเรดมากกว่า แดง และนี่คือประโยชน์ของฟิลเตอร์ที่นำมาใช้งาน

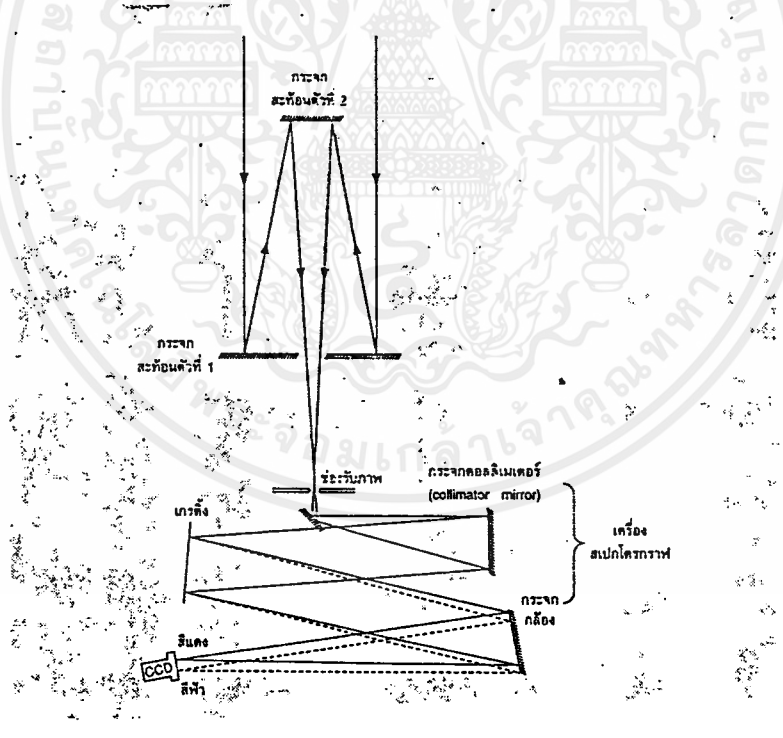
จากที่กล่าวมาแล้วว่า CCD มีช่วงไดนามิกเรนจ์ที่กว้างประโยชน์อีกอย่างหนึ่งที่นำมาใช้ คือ การถ่ายภาพของกลุ่มดาวที่เรียกว่า แกแลคซี่ (galaxy) ซึ่งมีลักษณะเป็นกระจุกดาวหนาแน่นอยู่ระหว่างใจกลางด้วยเทคนิคการถ่ายภาพของCCD อีกแบบหนึ่ง คือสามารถทำให้ CCD เก็บรายละเอียดในส่วนที่มีแสงน้อยๆของกลุ่มดาวทั้งหมดหรือถ่ายเฉพาะโครงสร้างของใจกลาง โดยลดความไวในการรับแสงของ CCD ลงซึ่งให้เห็นถึงความแตกต่างด้วยเทคนิคพิเศษนี้ดังรูปที่ 3.4

CCD ยังถูกนำมาใช้ในเครื่องมอวัดที่สำคัญอีกชิ้นหนึ่งนั้นคือ เครื่องวัด และบันทึกสเปกตรัม (spectroscopic)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.7แสดงภาพถ่ายด้วยเทคนิคพิเศษ

รูปที่ 3.8 แสดงให้เห็นถึงส่วนประกอบเบื้องต้นของเครื่องนี้เครื่องบันทึกสเปกตรัมมีประโยชน์ในการวัดความแรงของแต่ละความยาวคลื่นที่เกิดขึ้นบนดวงดาวแกแลคซี่ หรือท้องฟ้าในยามค่ำคืน



รูปที่ 3.8 แสดงโครงสร้างเบื้องต้นของเครื่องสเปกโตรกราฟที่ใช้ CCD ต่อร่วม

นอกจากการประยุกต์ใช้งานทั้งหมดที่กล่าวมา CCD ยังถูกนำมาใช้งานในด้านอื่นๆอีกมากมาย ตั้งแต่เครื่องมือวัดจนถึงเครื่องถ่ายภาพ จนในปัจจุบันอุปกรณ์ถ่ายภาพเกือบทุกชนิดต้องมี CCD ประกอบด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.3 ขีดจำกัดของ CCD

CCD ผลิตขึ้นจากสารกึ่งตัวนำชนิดซิลิกอน ซึ่งในขณะนี้ไม่สามารถผลิตอุปกรณ์ CCD ที่มีขนาดใหญ่ได้ อันเนื่องมาจากกระบวนการผลิตซิลิกอนนั่นเอง ทำให้การออกแบบจึงจำเป็นต้องนำ CCD หลายๆตัวมาต่อแบบเมตริกซ์กันเพื่อให้ได้ขนาดใหญ่ขึ้น

แต่ทั้งนี้ผลที่ได้ตามมาก็คือเกิดการสูญเสียขึ้นระหว่าง CCD แต่ละตัว ทำให้ทำงานได้ไม่เต็มประสิทธิภาพ การตอบสนองต่อความยาวคลื่นของ CCD เป็นอีกปัญหาหนึ่งที่สำคัญ คือ CCD ตอบสนองต่อความยาวคลื่นย่านสีน้ำเงินได้ไม่ดีนัก แต่สำหรับช่วงความยาวอื่นไม่เป็นปัญหาเท่าไร

ต่อไป CCD จะมีการพัฒนาทั้งด้านขนาดอุปกรณ์ซึ่งจะเหลือเพียง 15 ไมครอน และสามารถตรวจจับแสงได้ทั้งด้านหน้าและหลัง ปรับปรุงประสิทธิภาพการทำงานให้ได้ถึง 80 เปอร์เซ็นต์ เพิ่มความไวต่อความยาวคลื่นสีน้ำเงินและแสงที่มีความสว่างน้อย ตลอดจนการถ่ายเทประจุจนกระทั่งส่งข้อมูลออกจากตัว CCD ต้องมีการสูญเสียอิเล็กตรอนน้อยที่สุด และต้องออกแบบระบบป้องกันสัญญาณรบกวนได้ดีจนถึง 100 เปอร์เซ็นต์ การทำงานของ CCD ต้องมีลักษณะเป็นเชิงเส้น (linear) ตลอดย่านวัด ซึ่งในจุดนี้ต้องพัฒนาไปถึงกระบวนการผลิต รวมถึงการออกแบบโครงสร้างบน CCD

3.1.4 A/D CONVERTER

ประกอบด้วย A/D CONVERTER 2 ชุด ทำงานร่วมกัน โดยทำงานสลับกัน เพื่อให้ได้จำนวนจุดภาพเพิ่มขึ้นเป็น 2 เท่า แต่ใช้ความถี่เท่าเดิม โดยจะแยกเก็บเป็นจุดคู่-จุดคี่, ฟิลล์คู่-ฟิลล์คี่ โดย A/D CONVERTER แต่ละตัวจะเก็บภาพขนาด 256x256 จุด ซึ่งสัญญาณดิจิตอลที่ผ่านการแปลงจะถูกส่งไปเก็บยังใน Memory

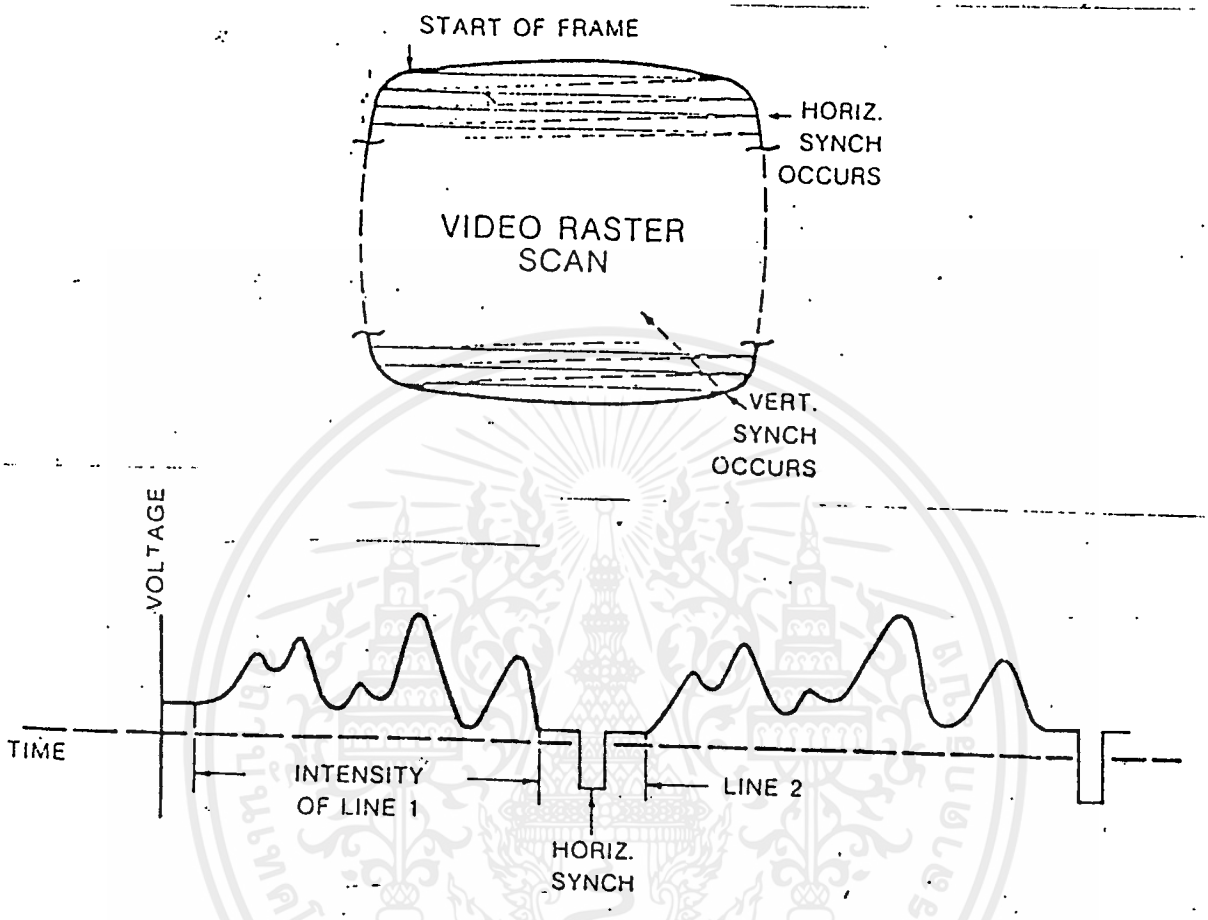
3.1.5 การเก็บรูปแบบตัวอักษรโดยการใช้กล้อง CCD

กล้อง CCD จะทำการเปลี่ยนสัญญาณแสงจากตัวอักษรที่ผ่านเข้ามา ให้กลายเป็นสัญญาณไฟฟ้าที่ต่อเนื่อง (Analog signal)

รูปที่ 3.9 แสดงการรับภาพตัวอักษรจากกล้อง CCD

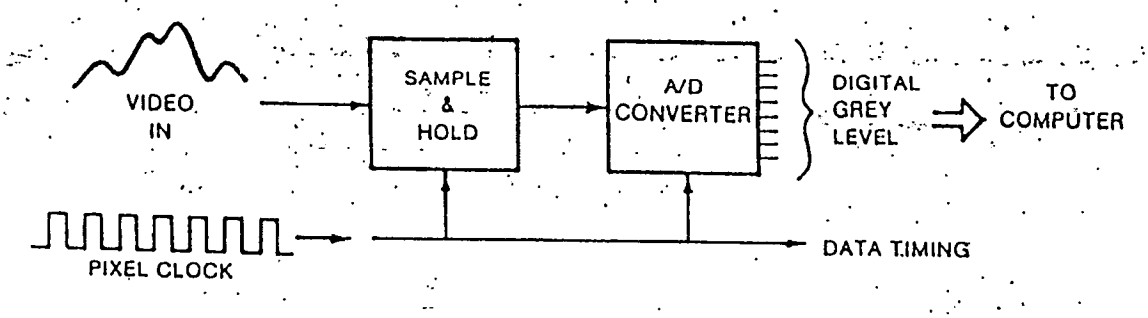
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัญญาณไฟฟ้าที่ต่อเนื่องนี้ก็คือ ค่าความเข้มของแสงแต่ละจุดบนตัวอักษรที่กล้องสแกน (Scan) ไปนั่นเอง ลักษณะการสแกนของกล้องแสดงในรูปที่ 3.10



รูปที่ 3.10 แสดงลักษณะของการสแกนของกล้อง CCD

สัญญาณไฟฟ้าที่ต่อเนื่องในรูปที่ 3.10 จะถูกนำมาสุ่ม (Sampling) แล้วทำการเปลี่ยนให้เป็นสัญญาณที่ไม่ต่อเนื่อง (Digital Signal) ด้วยอุปกรณ์เปลี่ยนสัญญาณ A/D Converter ดังรูปที่ 3.11

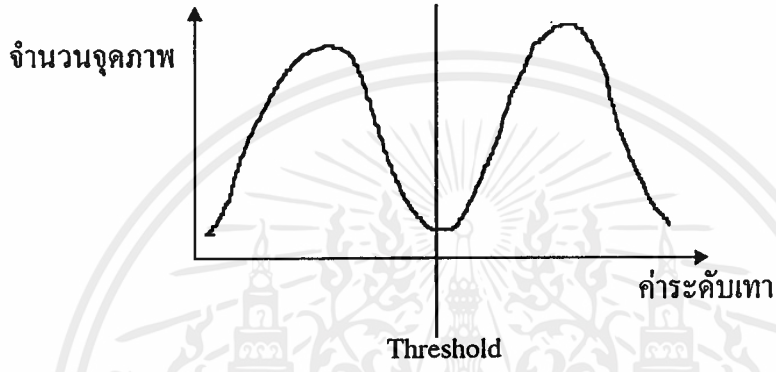


รูปที่ 3.11 การเปลี่ยนสัญญาณ Analog เป็น Digital

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ ... ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัญญาณในแต่ละเส้นสแกนก็จะถูกเปลี่ยน ให้เป็นสัญญาณแบบดิจิตอลหลาย ๆ จำนวน ขึ้นอยู่กับความถี่ของสัญญาณนาฬิกา แต่ละข้อมูลดิจิตอลที่ได้จะแทน 1 จุดภาพของการสแกน หรือ 1 Picture element (Pixel) ส่วนค่าของข้อมูลในแต่ละ Pixel นี้เรียกว่าค่าระดับเทา (Gray level)

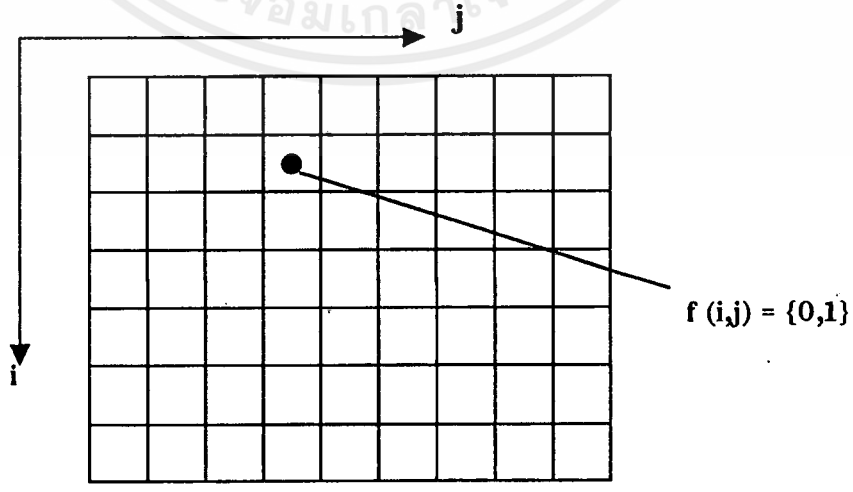
สำหรับรูปแบบที่ไม่ต้องการความละเอียดทางด้านระดับของแสงหรือระดับเทา ก็จะต้องทำการเปลี่ยนรูปแบบระดับเทา (Gray level pattern) ให้เป็นรูปแบบ 2 ระดับ (Binary Pattern) ด้วยการเลือกค่าขีดเริ่ม (Threshold) ที่เหมาะสม



รูปที่ 3.12 แสดงฮิสโตแกรมสำหรับการทำรูปแบบสองระดับ

ค่าขีดเริ่ม (Threshold) ที่เลือกอาจจะมีมากกว่า 1 ค่าก็ได้ โดยการพิจารณาค่าขีดเริ่ม ที่เหมาะสมจาก ฮิสโตแกรม (Histogram) ในรูปที่ 3.12 หรืออาจจะทำให้มีการคำนวณ ค่าขีดเริ่ม (Threshold) นี้ขึ้นมาเองโดยอัตโนมัติก็ได้

รูปแบบ 2 ระดับที่ได้จะมีค่าความสว่างในแต่ละจุดภาพได้สองค่า เช่น อาจจะใช้ค่า 1 แทนจุดภาพที่สว่าง หรือเป็นสีขาวและใช้ 0 แทนจุดภาพที่มืดหรือเป็นสีดำก็ได้ ดังแสดงในรูปที่ 3.13



รูปที่ 3.13 แสดงลักษณะของรูปแบบสองระดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 การปรับปรุงภาพโดยใช้ Linear density conversion [8]

สมการ Linear density conversion

$$Q(x,y) = a * P(x,y) + b \quad \text{..... (3.1)}$$

เมื่อ

P(x,y) = input image

Q(x,y) = output image

a = gain

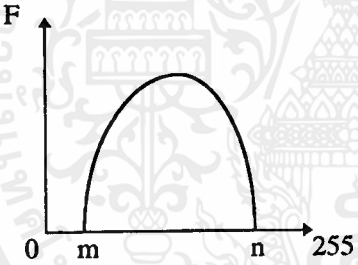
b = off-set

สมการหาค่า gian, off-set เมื่อพิจารณา histogram

$$gain = 255 / (n-m)$$

$$off\text{-}set = 255 - gain*n$$

ค่า m, n กำหนดโดยรูป histogram



รูปที่ 3.14 แสดง histogram ที่ใช้ในการคำนวณ

m = ค่า gray level ที่ความถี่เริ่มเปลี่ยนแปลงจาก 0 ไปหา 1

n = ค่า gray level ที่ความถี่เริ่มเปลี่ยนแปลงจาก 1 ไปหา 0

สมการ Linear density conversion ส่วนมากจะใช้ในการแปลงภาพเพื่อแสดงผลข้อมูลภาพดังนี้

1. การขยายข้อมูลภาพ (Stretching image)

ขยายข้อมูลภาพจาก 6 บิต 64 ระดับ เป็น 8 บิต 256 ระดับ สำหรับแสดงผล 8 บิต 256 ระดับ

2. การลดข้อมูลภาพ (Compression image)

ข้อมูลภาพ 10 บิต 1024 ระดับ ลดข้อมูลลงเป็น 8 บิต 256 ระดับ เพื่อทำการแสดงผลโดย 8 บิต 256 ระดับ

3. การปรับปรุงภาพ (Enhancement image)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทำการปรับปรุงส่วนที่เฉพาะเจาะจงของภาพข้อมูล เช่น ส่วนที่เป็น object หรือส่วนที่เป็น background

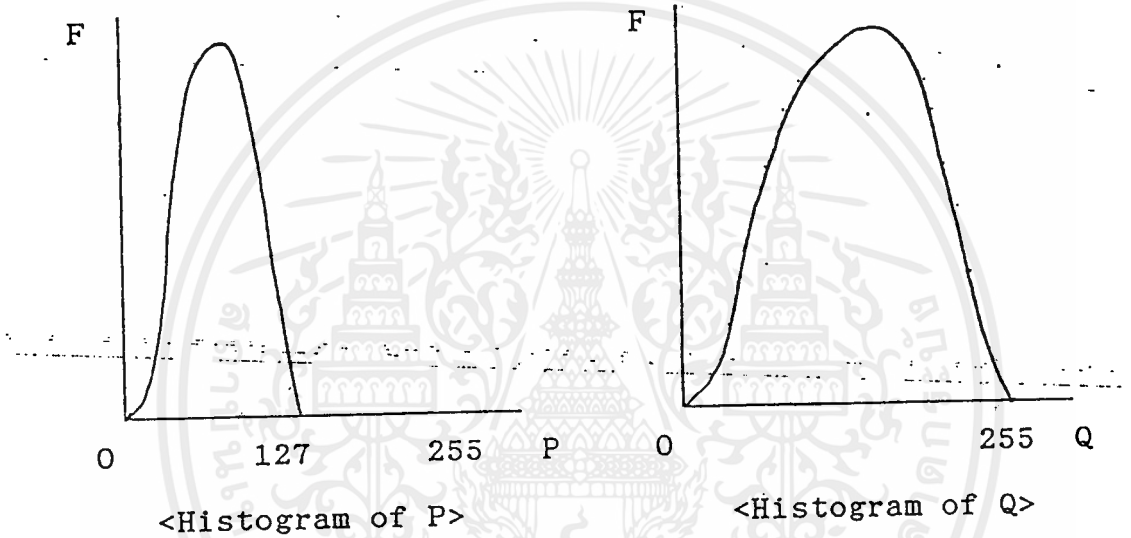
4. การแปลงภาพกลับ (Inversion image)

การแปลงภาพกลับนำมาใช้ในการแสดงภาพกลับระหว่างค่าความเข้มและความสว่างของภาพ ซึ่งค่าของการแปลงขึ้นอยู่กับการแปลงค่าความเข้มและความสว่างของภาพ

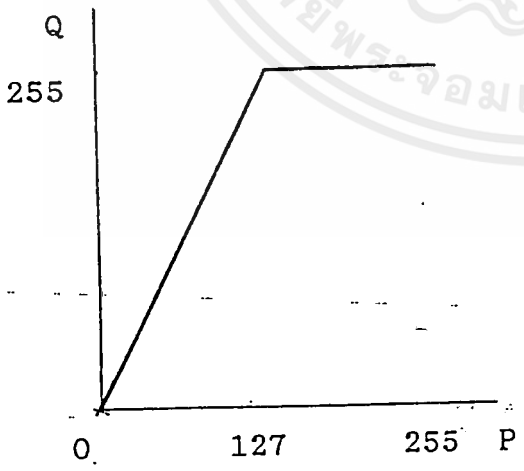
การแปลงภาพกลับทำโดย

เมื่อกำหนดให้ข้อมูล input 0 = สว่าง ; 255 = เข้ม

ข้อมูล output จะแสดงค่า 0 = เข้ม ; 255 = สว่าง



a = 2, b = 0



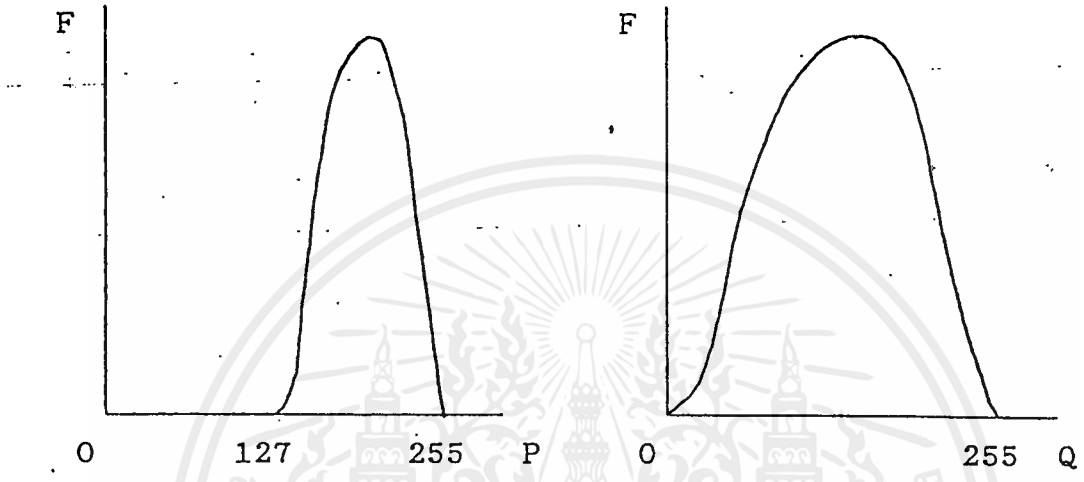
P	Q
0	0
⋮	⋮
127	254
128	255
⋮	⋮
255	255

<Conversion table>

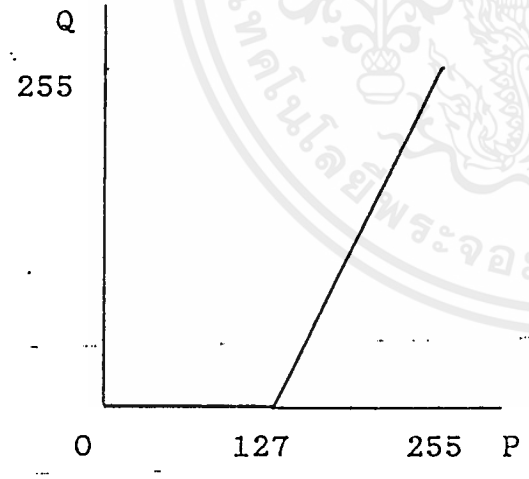
รูปที่ 3.15 แสดงความแตกต่างของการปรับปรุงก่อนและหลังข้อ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ปรับปรุงคุณภาพของภาพ เมื่อข้อมูลของภาพจาก histogram เบี่ยงเบนไปทางด้าน 255 gray scale



a = 2, b = -255



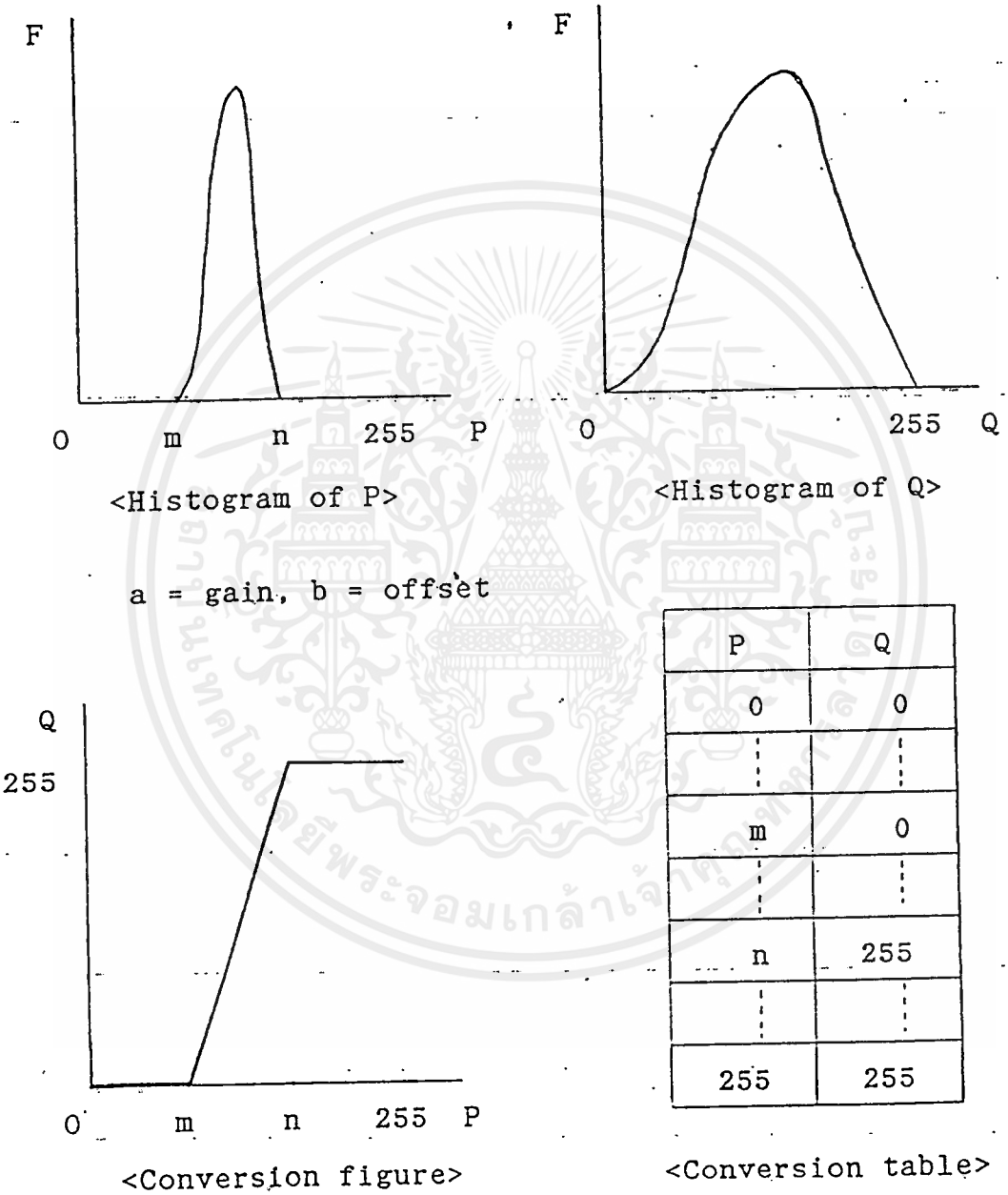
P	Q
0	0
⋮	⋮
127	0
128	1
⋮	⋮
255	255

<Conversion table>

รูปที่ 3.16 แสดงความแตกต่างของการปรับปรุงก่อนและหลังข้อ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

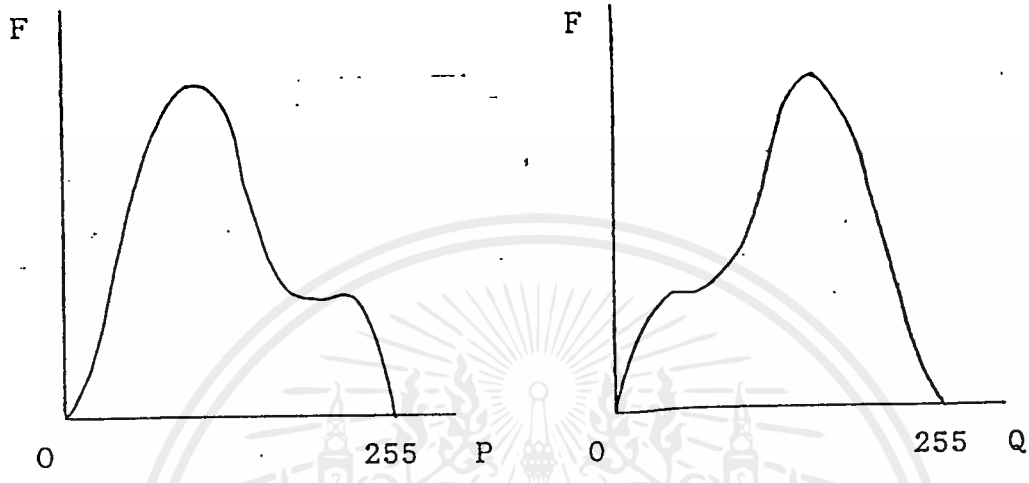
3. การปรับปรุงคุณภาพของภาพให้อยู่ในค่ากลางของภาพ โดยกำหนดให้ค่ากลางมีค่า m ≤ P ≤ n ค่า gain = 255 / (n - m) , ค่า offset = 255 - gain*n



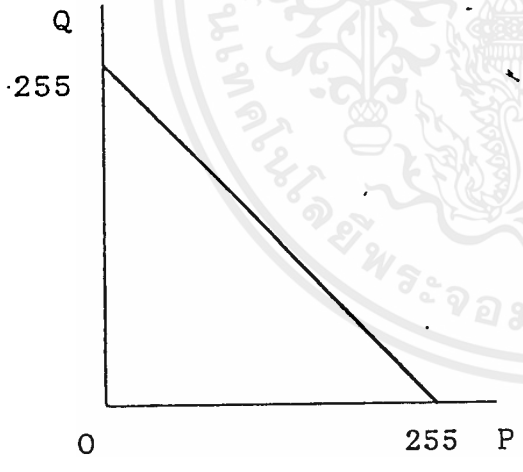
รูปที่ 3.17 แสดงความแตกต่างของการปรับปรุงก่อนและหลังข้อ 3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. ส่วนกลับของภาพ image ซึ่งมีค่า gain = -1 และ offset = 255



a = -1, b = 255



P	Q
0	255
⋮	⋮
127	128
128	127
⋮	⋮
255	0

<Conversion table>

รูปที่ 3.18 แสดงความแตกต่างของการปรับปรุงก่อนและหลังข้อ 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 ทฤษฎีที่ใช้ในการหาค่า Threshold [10]

การเลือกค่า Threshold นั้น มีวัตถุประสงค์เพื่อหาค่าแบ่งแยกระหว่างส่วนที่เป็น object และ ส่วนที่เป็น background เพื่อนำมาใช้ในการแปลงภาพ 2 ระดับ (binarization) โดยเฉพาะการประยุกต์ใช้ในการวิเคราะห์ภาพตัวอักษรที่เก็บข้อมูลภาพจากกล้อง CCD ที่มีค่าระดับ gray level 0-255 ระดับ เราจึงต้องพิจารณาเลือกค่า Threshold ที่มีความเหมาะสมกับทุกๆข้อมูลภาพที่นำมาวิเคราะห์ ด้วยเหตุนี้จึงต้องประยุกต์ใช้สมการหาค่า Threshold อัตโนมัติ เพื่อนำมาใช้ในการวิจัยในปริญญานิพนธ์ฉบับนี้

3.3.1 การเลือกค่า Threshold ระดับเดี่ยว (Single Level Threshold Selection)

การเลือกค่า Threshold เพื่อแยกความแตกต่างระหว่างส่วนขาวและดำ แบบระดับเดี่ยว กำหนดค่าสมการได้ดังนี้

$$T_{k+1} = \frac{1}{2}\left[\frac{\sum_{i=0}^{T_k} n(i)i}{\sum_{i=0}^{T_k} n(i)} + \frac{\sum_{i=T_k+1}^{N} n(i)i}{\sum_{i=T_k+1}^{N} n(i)}\right] \qquad (3.2)$$

เมื่อ n(i) ผลรวมของจำนวนจุดภาพที่ระดับเทาที่ i

T_{k+1} คือค่า Threshold ที่หาได้จากสมการที่ 3.2

สมการของการแบ่งแยกความแตกต่าง (σ^2) ของภาพ เพื่อทำการแบ่งภาพออกโดยค่า Threshold กำหนดโดย

$$\sigma^2(T) = p_D(\mu_D - T)^2 + p_B(\mu_B - \mu_T)^2 \qquad (3.3)$$

เมื่อ

$$p_D = \int_0^T p(x)dx \qquad p_B = \int_T^\infty p(x)dx$$

$$p_D\mu_D = \int_0^T xp(x)dx \qquad p_B\mu_B = \int_T^\infty xp(x)dx \qquad (3.4)$$

และ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\mu_T = p_D\mu_D + p_B\mu_B.$$

เมื่อ $p(x)$ = probability density function ของค่าระดับเทา

x = ค่าระดับเทา

p_D = ค่าความน่าจะเป็นของจุดภาพที่มีค่าน้อยกว่า T (ส่วนที่มืด)

p_B = ค่าความน่าจะเป็นของจุดภาพที่มีค่ามากกว่า T (ส่วนที่สว่าง)

$$\mu_D = \frac{\int_0^T xp(x)dx}{\int_T^\infty p(x)dx} = \text{mean of dark pixels.}$$

$$\mu_B = \frac{\int_T^\infty xp(x)dx}{\int_T^\infty p(x)dx} = \text{mean of bright pixels.}$$

$$\mu_T = \text{total mean} = \int_T^\infty xp(x)dx.$$

สมการที่ 3.3 สามารถลดรูปได้ดังนี้

$$\sigma^2(T) = P_D\mu_D^2 + P_B\mu_B^2 - \mu_T^2. \qquad (3.5)$$

การหาค่า Threshold T^* เป็นการเลือกค่าเหมือนตามที่ค่าสูงสุดของค่าความเบี่ยงเบน(variance) $\sigma^2(T)$, ค่าความเบี่ยงเบน(variance) $\sigma^2(T)$ เป็นการ differentiated กับ T เพื่อทำการหาค่าสภาวะสูงสุด

$$\frac{\partial\sigma^2(T)}{\partial T} = p_D(2\mu_D)\frac{\partial\mu_D}{\partial T} + \frac{\partial p_D}{\partial T}\mu_D^2 + p_B(2\mu_B)\frac{\partial\mu_B}{\partial T} + \frac{\partial\mu_B}{\partial T}\mu_B^2 = 0 \qquad (3.6)$$

แต่

$$\frac{\partial p_D}{\partial T} = p(T) \qquad (3.7)$$

$$\frac{\partial p_D}{\partial T} = -p(T) \qquad (3.8)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\frac{\partial \mu_D}{\partial T} = \frac{p_D T(T) - \left[\int_0^T xp(x)dx\right] p(T)}{p_D^2}$$

$$= \frac{1}{p_D}[T - \mu_D]p(T) \qquad (3.9)$$

และ

$$\frac{\partial \mu_B}{\partial T} = \frac{p_B(-Tp(T)) + \left[\int_T^\infty xp(x)dx\right] p(T)}{p_B^2}$$

$$= \frac{1}{p_B}[-T+\mu_B]p(T) \qquad (3.10)$$

แทนสมการ 3.7 – 3.10 ในสมการ 3.6, เราจะได้

$$2T(\mu_D - \mu_B) = \mu_D^2 - \mu_B^2$$

หรือ

$$T = \frac{\mu_D + \mu_B}{2} \qquad (3.11)$$

จากนั้น จะได้ค่าที่ดีที่สุดของ T ที่เราพึงพอใจ

$$T^* = \frac{1}{2}\left[\frac{\int_0^{T^*} xp(x)dx}{\int_0^{T^*} p(x)dx} + \frac{\int_{T^*}^\infty xp(x)dx}{\int_{T^*}^\infty p(x)dx}\right] \qquad (3.12)$$

หรือสำหรับการเพิ่มรูปแบบสมการ 3.12 สามารถเขียนได้ดังนี้

$$T^* = \frac{1}{2}\left[\frac{\sum_{i=0}^{T^*} n(i)i}{\sum_{i=0}^{T^*} n(i)} + \frac{\sum_{i=T^*+1}^{N} n(i)i}{\sum_{i=T^*+1}^{N} n(i)}\right] \qquad (3.13)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ N เป็นจำนวนของระดับ gray ในภาพ สำหรับภาพระดับ gray 8-bit นั้น N = 255 ระดับ ซึ่ง T^*สามารถทำการคำนวณได้จากสมการที่ 3.13 ใช้ในการ *fixed-point* เพื่อย้ำวิธีการ นั้น จะได้

$$T_{k+1} = \frac{1}{2}\left[\frac{\sum_{i=0}^{Tk} n(i)i}{\sum_{i=0}^{Tk} n(i)} + \frac{\sum_{i=Tk+1}^{N} n(i)i}{\sum_{i=Tk+1}^{N} n(i)}\right] \qquad (3.14)$$

ระดับ Threshold T เป็นค่าจำนวนเต็มที่มีขอบเขตอยู่ระหว่าง 0-255 สำหรับ ภาพระดับ gray 8-bit ตั้งแต่ด้านขวามือของสมการ 3.14 สามารถเป็นค่าที่ยอมให้ เป็นจำนวนจริง มันเป็นการทำแบบวนรอบกับค่าจำนวนเต็มที่ใกล้ที่สุดหลังจากทำวนรอบทุกครั้งแล้วผลรวมหาได้จาก

$$T_{k+1} = T_k \qquad (3.15)$$

หมายเหตุ จากเดิม $| T_{k+1} - T_k | \leq \varepsilon$ เป็นการลดรูปจากสมการที่ 3.15 ตั้งแต่สมการที่ 3.14 แสดงลักษณะของ algorithm แสดงในสมการที่ 3.3, เมื่อมันสามารถทำได้ตามสภาวะเป็นไปตามรูปแบบที่เราค้นหา ก็สามารถทำการลดการทำวนรอบ (iteration) ที่เหมือนกันให้เข้าใกล้ที่การคาดเดาเอาไว้ในสมการที่ 3.3

<u>ผลรวมของทุกสมการ</u>

จะได้สมการการทำวนรอบ (iterative) ของการเลือก Threshold หนึ่งระดับสามารถใช้กับรูปแบบทั่วไป ซึ่งแสดงได้ดังนี้

$$T_{n+1} = g(T_n) \qquad (3.16)$$

ในเชิงคณิตศาสตร์นี้การทำวนรอบ (iteration) สามารถแสดงได้จากภาพรูปที่ 3.19 โดยให้ T_n , โดย T_{n+1} เป็น x-coordinate ของจุดบนกราฟ y = T ได้มาโดยดูการฉายในแนวตั้งของจุด (T_n,g (T_n)) พิจารณาจุด A = (T_0,g(T_0)) ซึ่งถ้าเอนลงมาบนเส้นกราฟ y = g(T) ในแนวการตั้งของ g(T_0) บนเส้น y = T จะได้จุด B = (T_1,T_1) จาก T_1 เราสามารถคำนวณได้จาก g(T_1) และในการทำแบบเดิมทำให้เกิด T_2 โดยทำซ้ำในขบวนการนี้เราจะได้ T_3,T_4 ไปเรื่อยๆ

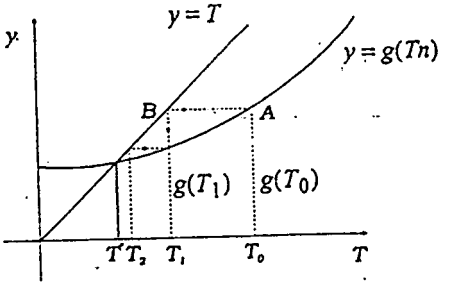
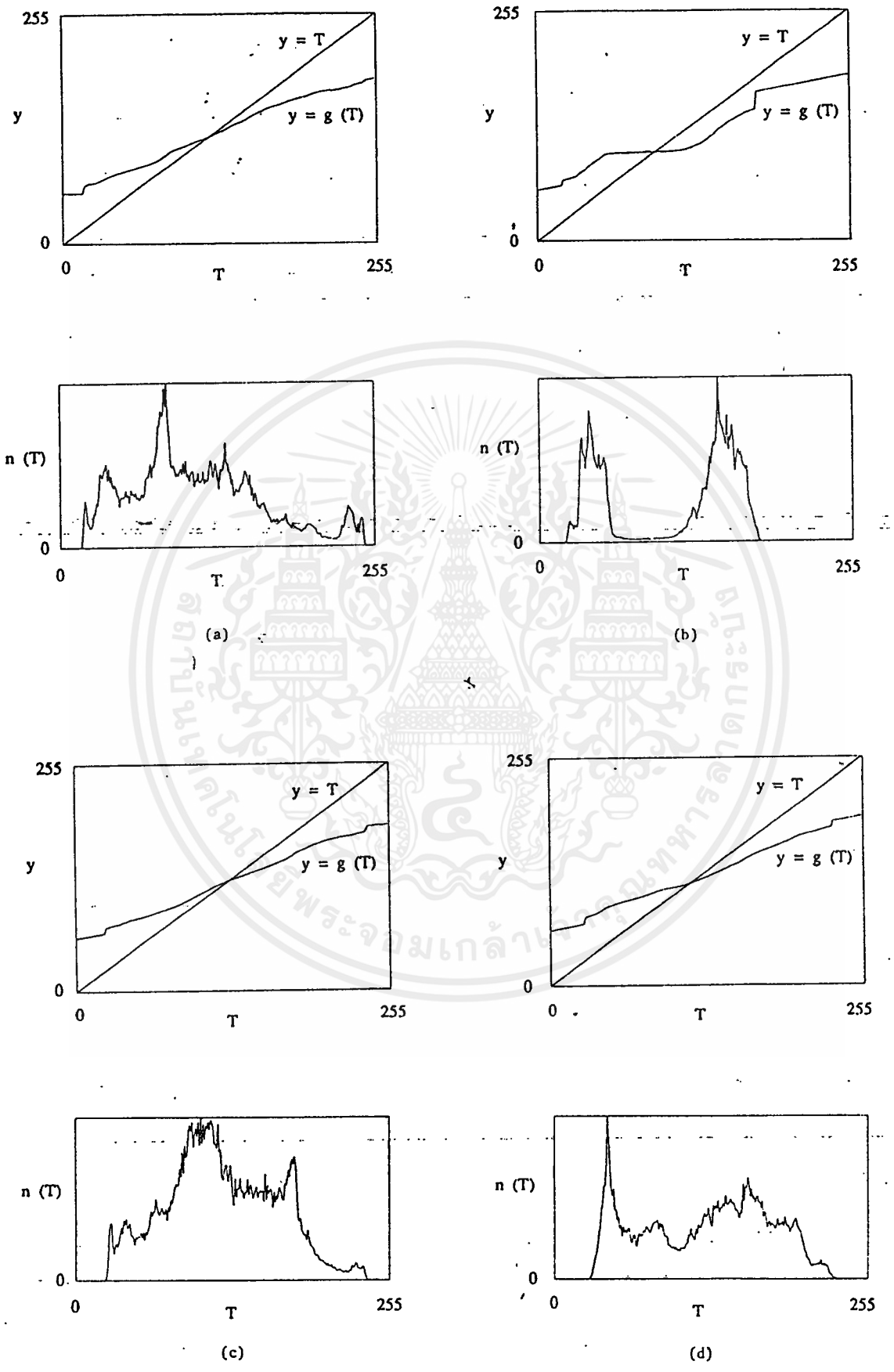


Fig. 1. Fixed-point iteration.

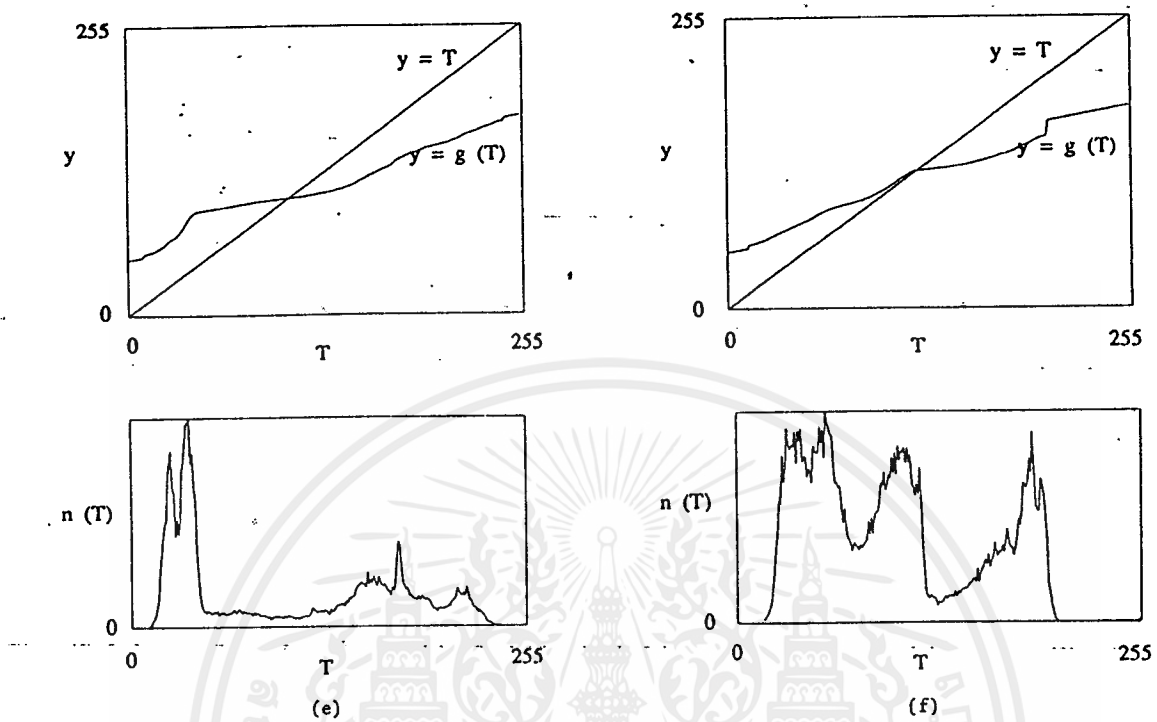
รูปที่ 3.19 แสดง Fixed-point ในการทำวนรอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.20 แสดงภาพลักษณะของ Threshold ค่าต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.20 แสดงภาพลักษณะของ Threshold ค่าต่างๆ (continued)

จากสมการ (3.14) มันแสดงอย่างชัดเจนดังนี้

$$g(1) = \frac{1}{2}\left[\frac{\sum_{i=0}^{I} n(j)i}{\sum_{i=0}^{I} n(i)} + \frac{\sum_{i=2}^{N} n(j)i}{\sum_{i=2}^{N} n(i)}\right] \qquad (3.17)$$

ตั้งแต่ $n(i) \geq 0$ เราสามารถเขียนได้เป็น

$$\sum_{i=0}^{I} n(i)i \leq \sum_{i=0}^{I} n(i)$$

และ

$$\sum_{i=2}^{N} n(i)i \leq N \sum_{i=2}^{N} n(i) \qquad (3.18)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้น

$$g(1) \leq \frac{1}{2}\left[\frac{\sum_{i=0}^{1} n(i)}{\sum_{i=0}^{1} n(i)} + \frac{N\sum_{i=2}^{N} n(i)}{\sum_{i=2}^{N} n(i)}\right] \qquad (3.19)$$

$$\leq \frac{1}{2}\,[\,1+N\,] > 0 \qquad (3.20)$$

เหมือนกับ

$$g(2) \leq \frac{1}{2}[2 + N]$$

$$g(N-1) \leq \frac{1}{2}[N - 1 + N] = N - \frac{1}{2} \qquad (3.21)$$

ในที่นี้ความหมายของ Curves y=g(T) และ y=T มีความสัมพันธ์กันโดยแสดงในรูปที่ 3.19, ในรูปภาพที่ 3.20(a)-(f) แสดงตัวอย่างของ Curves ของ g(T) และ histograms n(T) ได้มาจากการเปลี่ยนรูปภาพ บันทึกค่าของ Threshold ที่ได้มาโดย การนำมาใช้งานของรูปแบบ Single level iterative จะมีผลเพียงเท่านี้ สำหรับ histograms มีการจัดประเภทของ bimodal ตามรูปที่ 3.20 (b),(d) และ (e)

การรวมกันของสมการที่ 3.5 เป็นขอบเขตที่จะทำได้ในรูปแบบส่วนมากในสมการที่ 3.5-3.8 ในการวนรอบ ปรากฏความสัมพันธ์ของ Curves y=g(T) กับความสัมพันธ์ที่ y=T แสดงในตัวอย่างในรูปที่ 3.20 ซึ่งได้ทำการอธิบายถึงการเข้ากันอย่างเร็วของ algorithm

3.3.2 การเลือกค่า Threshold หลายระดับ (Multi-level Thresholding)

การเลือกค่า Threshold แบบ Multi-level Thresholding พัฒนาขึ้นเพื่อการเลือกค่า Threshold มากกว่า 1 ค่าใน 1 ข้อมูลภาพ และ ปรับปรุงการเลือกค่าแบบ Single-level Thresholding ให้มีความสามารถเลือกค่า Threshold ได้หลายๆ ค่าใน 1 ข้อมูลภาพดังนี้

พิจารณาสมการการแบ่งระดับขั้นของภาพออกเป็น 4 ระดับ โดยกำหนดค่า Threshold ออกเป็น T_1, T_2, T_3 ดังนี้

$$\sigma^2 (T_2,T_2,T_3) = p_1(\mu_1-\mu_T)^2 + p_2(\mu_2-\mu_T)^2 + p_3(\mu_3-\mu_T)^2 + p_4(\mu_4-\mu_T)^2 \qquad (3.22)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ μ_i คือ ค่าเฉลี่ยของระดับขั้นของภาพแต่ละระดับขั้น

p_i คือ ค่าความน่าจะเป็นของจำนวนจุดภาพที่อยู่ระหว่าง T_{i-1} และ T_i

กำหนดค่า i ตั้งแต่ 1 ค่า ค่าระดับชั้นขอบภาพ

โดยค่าตัวแปร μ_i และ p_i กำหนดค่าดังนี้

$$p_i = \int_{Ti-1}^{Tj} p(x)dx \qquad (3.23)$$

$$p_i\mu_i = \int_{Ti-1}^{Tj} xp(x)dx \qquad (3.24)$$

เมื่อ $T_0 = 0$ และ $T_4 = N$

ทำการดิฟเฟอเรนทิเอทสมการที่ (3.22) แสดงดังนี้

$$2T_i = \frac{\int_{Ti-1}^{Tj} xp(x)dx}{\int_{Ti-1}^{Tj} p(x)dx} + \frac{\int_{Ti}^{Ti+1} xp(x)dx}{\int_{Ti}^{Ti+1} p(x)dx} \qquad (3.25)$$

สำหรับค่า I = 1,2,3 และ T_0=0 , T_4=N

กระทำสมการต่างๆ เหมือนกับแบบ Single-level Thresholding และแสดงดังสมการดังนี้

$$2T_{i,k+1} = \frac{\sum_{i=Ti-1,k}^{Ti,k} n(i)i}{\sum_{i=Ti-1,k}^{Ti,k} n(i)} + \frac{\sum_{i=Ti,k+1}^{Ti,k} n(i)i}{\sum_{i=Ti,k+1}^{Ti+1,k} n(i)} \qquad (3.26)$$

เมื่อ i=1,2,....,M; M = จำนวนของค่า Threshold

ค่าองค์ประกอบของสมการนี้ (3.26) ตัวแรก (i) แสดงถึงจำนวนระดับค่า Threshold ตัวที่ 2 (k+1) แสดงถึงจำนวนของการวนรอบหาค่าตามสมการ แสดงได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

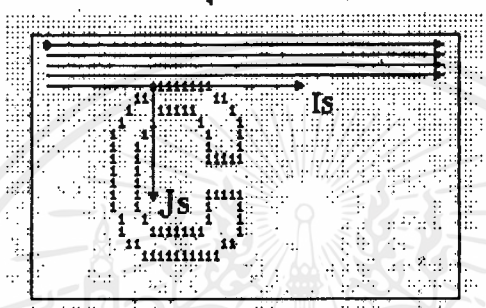
$$T_{M+1,k} = N \text{ and } T_{0,k} = 0, \quad \forall k,$$

สมการการวนรอบจะหยุดเมื่อ

$$|T_{i,k+1} - T_{i,k}| = 0, \quad \forall i,$$

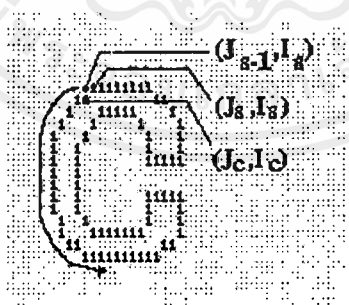
3.4 การติดตามขอบภาพ โดยการติดตามขอบเขตของภาพ (Contour following)

การหาขอบภาพโดยวิธีนี้ จะต้องหาจุดเริ่มต้นของขอบก่อน



รูปที่ 3.21 การ SCAN หาจุดภาพที่เป็นจุดเริ่มต้น

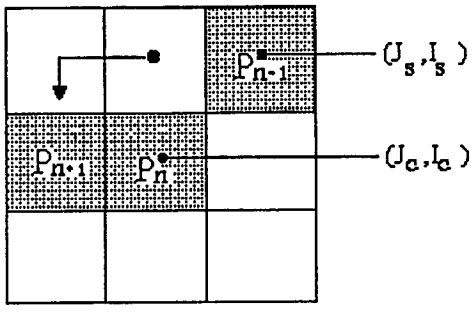
จากรูปแบบ 2 ระดับในรูปที่ 3.21 จะทำการ scan ไปบนรูปแบบจากบนลงล่างจนกระทั่งพบจุดที่เป็นภาพ ก็จะถือว่าจุดนั้นเป็นจุดเริ่มต้นของจุดภาพ (J_s,I_s) และจะกำหนดให้จุดก่อนหน้าที่จะถึงจุดเริ่มต้นของขอบนี้ (J_{s-1},I_s) เป็นตำแหน่งเริ่มต้นสำหรับการวนในทิศทวนเข็มนาฬิกา เพื่อติดตามหาขอบภาพจุดต่อไป (J_c,I_c) ดังรูปที่ 3.22



รูปที่ 3.22 จุดเริ่มต้นของการติดตามหาขอบของภาพ

จากนั้นวาง Window ขนาด 3x3 ที่ขอบภาพ (J_c,I_c) แล้วจะวนในทิศทางทวนเข็มนาฬิกา เพื่อติดตามหาขอบภาพจุดต่อไป ดังรูปที่ 3.23

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



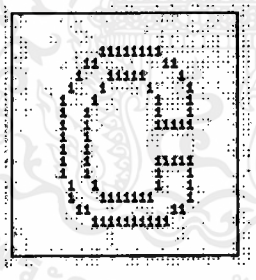
รูปที่ 3.23 การติดตามหาขอบภาพที่เป็นขอบ

เมื่อ P_{n-1} เป็นตำแหน่งของขอบภาพที่ผ่านมา

P_n เป็นตำแหน่งของขอบภาพที่กำลังพิจารณา

P_{n+1} เป็นตำแหน่งของขอบภาพต่อไป

จากรูปที่ 3.23 จะเห็นว่าจุดเริ่มต้นในการติดตามขอบจุดต่อไป (P_{n+1}) จะเป็นตำแหน่งที่ถัดจากขอบภาพที่ผ่านมา (P_{n-1}) ในทิศทางทวนเข็มนาฬิกาเสมอ ด้วยวิธีการดังกล่าวนี้เมื่อติดตามขอบเขตของภาพไปจนกระทั่งพบตำแหน่งที่เริ่มต้นของขอบภาพ (J_s,I_s) ก็จะได้ขอบของภาพออกมา ดังรูปที่ 3.24



รูปที่ 3.24 ขอบของภาพที่ได้จากการติดตามขอบเขต

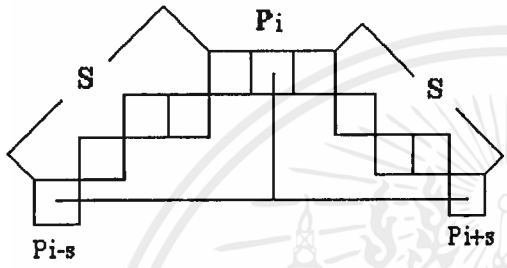
3.5 การหาจุดเปลี่ยนแปลงทิศทางของลายเส้นเพื่อใช้ในการหามุม (Corner Detection)[4]

เนื่องจากโครงร่างของขอบหนังสือ มีลักษณะเป็นสี่เหลี่ยมผืนผ้ามีจุดเปลี่ยนแปลงทิศทางของลายเส้นอยู่ 4 จุด เราจึงนำวิธีการนี้มาประยุกต์ใช้เพื่อหาองศาของสันหนังสือกับแนวระนาบ ซึ่งการหาจุดเปลี่ยนแปลงทิศทางของลายเส้นนี้เป็นสิ่งสำคัญมาก เพราะถ้าหากจุดเปลี่ยนแปลงทิศทางของลายเส้นผิดพลาดจะทำให้การหาองศาของสันหนังสือแนวระนาบผิดพลาดไปด้วย และรวมไปถึงการวิเคราะห์ในขั้นตอนต่อๆ ไปด้วย

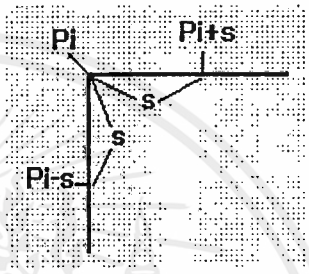
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5.1 การหามุมของลูป

มุมของลูป คือ จุดที่มีการเปลี่ยนแปลงทิศทางของลายเส้นอย่างรวดเร็ว หรือจุดที่มีค่าความโค้ง (k) มากๆ ซึ่งการหามุมหรือการหาค่าความโค้งที่นำเสนอในรูปที่ 3.25 คือ กำหนดช่วงของจุดภาพที่ติดต่อกัน โดยมีจุดที่ต้องการหาค่าความโค้งเป็นจุดกึ่งกลางของช่วง (P_i) และคำนวณหาค่าความโค้งได้จากระยะทางจากจุดกึ่งกลาง (P_i) ถึงเส้นตรงที่ลากจากจุดเริ่มต้นของช่วง (P_{i-s}) ไปยังจุดสุดท้ายของช่วง (P_{i+s}) ในแนวตั้งฉากดังรูปที่ 3.25



รูปที่ 3.25 แสดงการหาค่าความโค้ง



รูปที่ 3.26 แสดงถึงมุม

<u>สมการหาค่าความโค้ง</u>

$$K_i = \left| \frac{(y_{i+s} y_{i-s})\, x_i - (x_{i+s} - x_{i-s})\, y_i + x_{i+s} y_{i-s} - y_{i+s} x_{i-s}}{\sqrt{(y_{i+s} - y_{i-s})^2 + (x_{i+s} - x_{i-s})^2}} \right| \qquad \text{.....(3.27)}$$

เมื่อ i = จุดตรวจสอบใดๆ ของลูป (i = 0,1,2,...,n)

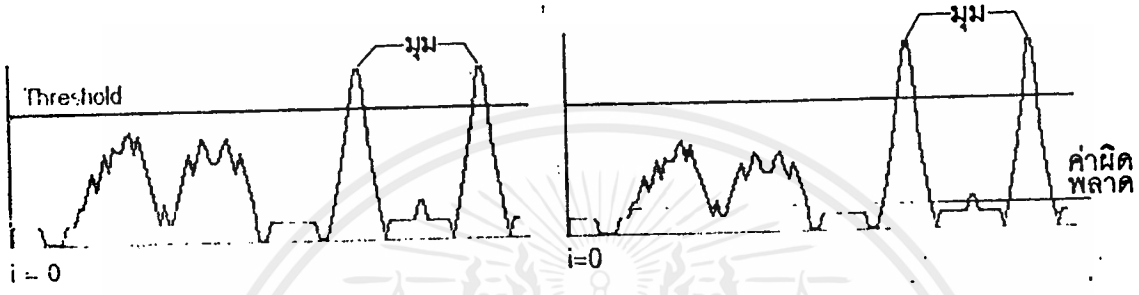
n = จำนวนจุดในลูป

จากสมการที่ 3.27 เราสามารถหาค่าความโค้งของทุก ๆ จุดในลูปได้ดังรูปที่ 3.27(ก) ซึ่งค่าความโค้งที่ได้นี้จะมีค่ามากหรือน้อยขึ้นอยู่กับการกำหนดช่วง (s) ถ้ากำหนดช่วง (s) มาก ค่าความโค้งก็จะมาก ในทางกลับกันถ้ากำหนดช่วง (s) น้อย ค่าความโค้งก็จะน้อย ซึ่งมีผลให้การกำหนดค่า Threshold ทำได้ยากหรือต้องกำหนดค่า Threshold หลายค่า ดังนั้นเพื่อให้การกำหนดค่า Threshold หนึ่งค่า สามารถครอบคุมการกำหนดช่วง (s) ได้หลาย ๆ ค่าสามารถทำได้โดยการทำให้ค่าความโค้งเป็นอัตราส่วนกับระยะห่างจากจุดเริ่มต้นของช่วง (P_{i-s}) กับจุดสุดท้ายของช่วง (P_{i+s}) ดังสมการที่ 3.28

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมการหาค่าความโค้ง

$$K_i = \left| \frac{(y_{i+s} - y_{i-s})x_i - (x_{i+s} - x_{i-s})y_i + x_{i+s}y_{i-s} - y_{i+s}x_{i-s}}{(y_{i+s} - y_{i-s})^2 + (x_{i+s} - x_{i-s})^2} \right| \quad \text{..... (3.28)}$$



(ก) กราฟแสดงค่าความโค้งที่ได้จากสมการที่ 3.27 (ข) กราฟแสดงค่าความโค้งที่ได้จากสมการที่ 3.28

รูปที่ 3.27 กราฟแสดงค่าความโค้ง

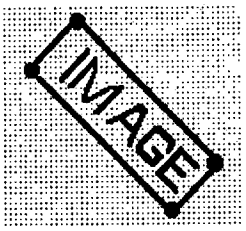
ค่าความโค้งที่หาได้จากสมการที่ 3.28 แสดงในรูปที่ 3.27(ข) ซึ่งจุดเห็นว่ารูปแบบของกราฟจะเหมือนเดิม แต่ค่าความโค้งบริเวณที่เป็นมุมจะสูงเด่นขึ้น เป็นขบวนการปรับปรุงสมการข้างต้นเพื่อขจัดความผิดพลาดอันอาจเกิดจากค่าที่ไม่ใช่มุมของลูป อาจถูกขจัดรวมเข้าไปในส่วนมุมของลูปด้วยโดยการกำหนดค่า Threshold ที่คงที่ใดๆ ค่าหนึ่ง

3.5.2 การวิเคราะห์หาจุดเปลี่ยนแปลงทิศทางของขอบของสันหนังสือ

เมื่อได้ค่าความโค้งขอบของสันหนังสือที่มีการเลือกค่าโดยค่า Threshold แล้ว ข้อมูลจะถูกจัดเป็นกลุ่มของข้อมูล 4 กลุ่มข้อมูล ตามจุดเปลี่ยนแปลงทิศทางของขอบของสันหนังสือ 4 จุดด้วยกัน

เราจะนำค่าที่สูงที่สุดของแต่ละกลุ่มมาเป็นตัวกำหนดจุดเปลี่ยนแปลงทิศทางของขอบสันหนังสือ ผลสรุปแล้วจะได้ค่าออกมา 4 จุดหรือ 4 ค่า ดังรูปที่ 3.28

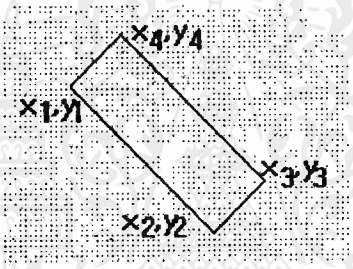
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.28 แสดงจุดมุมทั้ง 4 ของสันหนังสือ

3.5.3 การหาองศาของสันหนังสือกับแนวระนาบ

เมื่อเราทราบมุมของสันหนังสือทั้ง 4 มุม เราสามารถกำหนดองศาของสันหนังสือกับแนวระนาบ โดยกำหนดตำแหน่งของมุมสันหนังสือทั้ง 4 มุม ดังรูปที่ 3.29



รูปที่ 3.29 แสดงการกำหนดตำแหน่ง (x,y) มุมของสันหนังสือ

นำเส้นขอบของสันหนังสือที่มีความยาวของเส้นขอบสันหนังสือสูงสุดมาเป็นตัวกำหนดพิกัด เพื่อหาองศาของสันหนังสือกับแนวระนาบ

<u>สมการ</u>

$$\theta = \tan^{-1} \frac{Y_2 - Y_1}{X_2 - X_1} \qquad (3.29)$$

X_1, Y_1 = พิกัดจุด X,Y ด้านปลายด้านหนึ่งบนเส้นขอบสันหนังสือที่ยาวที่สุด

X_2, Y_2 = พิกัดจุด X,Y ด้านปลายที่ตรงข้ามกันบนเส้นขอบของสันหนังสือที่ยาวที่สุด

<u>3.6 การหมุนภาพ</u>

การหมุนภาพเป็นการแปลงอีกแบบหนึ่ง สำหรับการหมุนภาพนี้จะต้องกำหนดว่าจุดใดเป็นจุดหมุนเสมอ รูปที่ 3.30 แสดงการหมุนภาพสี่เหลี่ยมรอบจุดหมุน (pivot point) ซึ่งอยู่ที่จุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กำเนิดหรือจุด (0,0) นั่นเอง หลังจากที่ภาพถูกหมุนไปแล้ว ระยะห่างระหว่างจุดหมุนกับภาพจะยังคงมีค่าเท่าเดิม รูปร่าง ลักษณะของภาพก็ยังคงเดิม แต่ภาพจะมีการจัดวางที่ต่างไปจากเดิมอันเนื่องมาจากการหมุนนั่นเอง การหมุนภาพนี้อาจจะหมุนทีละหลาย ๆ ภาพก็ได้ จะหมุนแบบทวนเข็มนาฬิกาก็ได้ และจุดหมุนที่ใช้ก็อาจจะอยู่ภายในภาพหรืออยู่ภาพนอกภาพก็ได้

การอ้างถึงจุดพิกัด (x,y) นั้น นอกจากจะใช้ระบบพิกัดฉากแล้ว (คือกำหนดตำแหน่งจุดโดยบอกระยะทางในแนวนอนและในแนวดิ่ง) เราอาจจะใช้ระบบพิกัดโพลาร์ก็ได้ (ระบบพิกัดโพลาร์ คือการบอกตำแหน่งจุดโดยใช้เวกเตอร์) สองระบบนี้มีความสัมพันธ์กันดังนี้

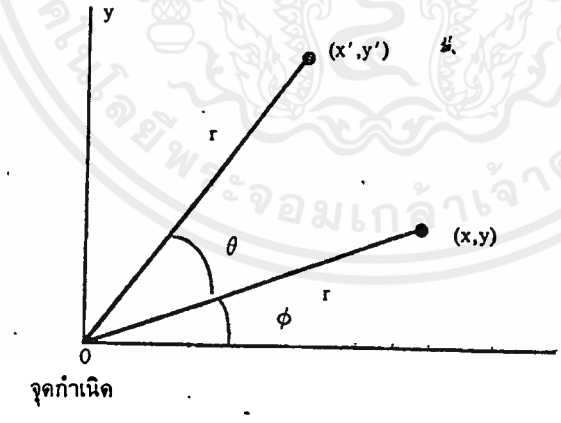
$$x = r \cos(\phi)$$

$$y = r \sin(\phi)$$

ถ้า (x,y) ถูกหมุนไปจากจุดเดิมเป็นมุม θ ในทิศทางทวนเข็มนาฬิกาจะได้จุดใหม่คือ (x',y') ดังรูป เราจะได้ว่า

$$x' = r \cos(\phi + \theta)$$

$$y' = r \sin(\phi + \theta)$$



รูปที่ 3.30 การหมุนภาพโดยที่จุดหมุนอยู่ที่จุดกำเนิด

ซึ่งอาจเขียนใหม่ได้ดังนี้

$$x' = r \cos(\phi) \cos(\theta) - r \sin(\phi) \sin(\theta)$$

$$y' = r \sin(\phi) \cos(\theta) + r \cos(\phi) \sin(\theta)$$

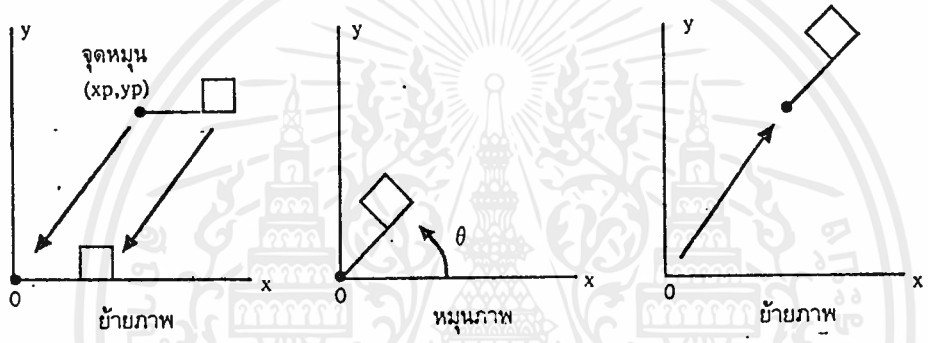
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และแทนค่า r cos(ϕ) ด้วย x, r sin(ϕ) ด้วย y จะได้

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

สมการข้างต้นก็คือสมการที่ใช้แปลงค่าพิกัด จากจุด (x,y) ไปเป็นจุดใหม่ (x',y') โดยการหมุนรอบจุดกำเนิดไปเป็นมุม θ ทิศทางทวนเข็มนาฬิกา การหมุนภาพทำได้โดยการแปลงค่าพิกัดทุกจุดที่ใช้นิยามภาพนั้นไปเป็นพิกัดใหม่โดยใช้สมการข้างต้นแล้วค่อยวาดภาพเดิมที่จุดพิกัดใหม่นั้น

ในทางปฏิบัติเราอาจจะต้องหมุนภาพรอบจุดใด ๆ ก็ได้ซึ่งไม่ใช่จุดกำเนิด สำหรับในกรณีนี้จะต้องใช้ 3 ขั้นดังนี้ (รูปที่ 3.31)



รูปที่ 3.31 การหมุนภาพรอบจุดหมุนใด ๆ

1. ย้ายจุดหมุน(xp,yp) ไปยังจุดกำเนิด (0,0) เมื่อย้ายแล้ว ทุก ๆ จุด(x,y) ที่ใช้นิยามภาพก็จะถูกย้ายไปยังจุดใหม่ (x',y') ด้วย โดยที่

$$x' = x - xp$$
$$y' = y - yp$$

ในตอนนี้จุดหมุนก็จะถูกย้ายไปยังจุดกำเนิด (0,0)

2. จัดการหมุนภาพรอบจุดกำเนิดนั่นคือจุด (x',y') ถูกย้ายไปเป็นมุม θ ได้จุดใหม่เป็น (x", y") โดยที่

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

แทนค่า x' และ y' จะได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$x'' = (x - xp)\cos(\theta) - (y - yp)\sin(\theta)$$

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

3. ย้ายจุดหมุนจากจุดกำเนิด (0,0) กลับไปยังจุดเดิม (xp,yp) ดังนั้นจุด (x",y") ก็จะถูกย้ายไปยังจุด (x",y") โดยที่

$$x'' = x'' + xp$$

$$y'' = y'' + yp$$

แทนค่า x" และ y" จะได้ว่า

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

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

สมการนี้คือสมการสำหรับหาค่าพิกัดใหม่ เมื่อหมุนภาพไปเป็นมุม θ ในทิศทางทวนเข็มนาฬิกา โดยหมุนรอบจุดหมุน (xp,yp) นั่นเอง

เรามักจะนิยมทำการหมุนภาพรอบจุดโดยหมุนทวนเข็มนาฬิกาเสมอ เนื่องจากว่าในทางคณิตศาสตร์จะกำหนดว่า ทิศทางการหมุนที่เป็นบวกจะเป็นทิศทวนเข็มนาฬิกา ถ้าเราต้องการจะหมุนตามเข็มนาฬิกาก็สามารถทำได้โดยใช้ค่า $-\theta$ แทน θ ในสมการ ผลที่ได้ก็จะเหมือนกัน

3.7 คุณสมบัติทางโทโปโลยีของภาพดิจิตอล (Topology)

คุณสมบัติทางโทโปโลยีของภาพดิจิตอล จะใช้จำแนกกลุ่มของจุดภาพ โดยวิธีการหาค่าคุณสมบัติทางโทโปโลยีของภาพดิจิตอลนี้ จะใช้วิธีการกำหนดหน้าต่างขนาด 3 x 3 ครอบบิตที่

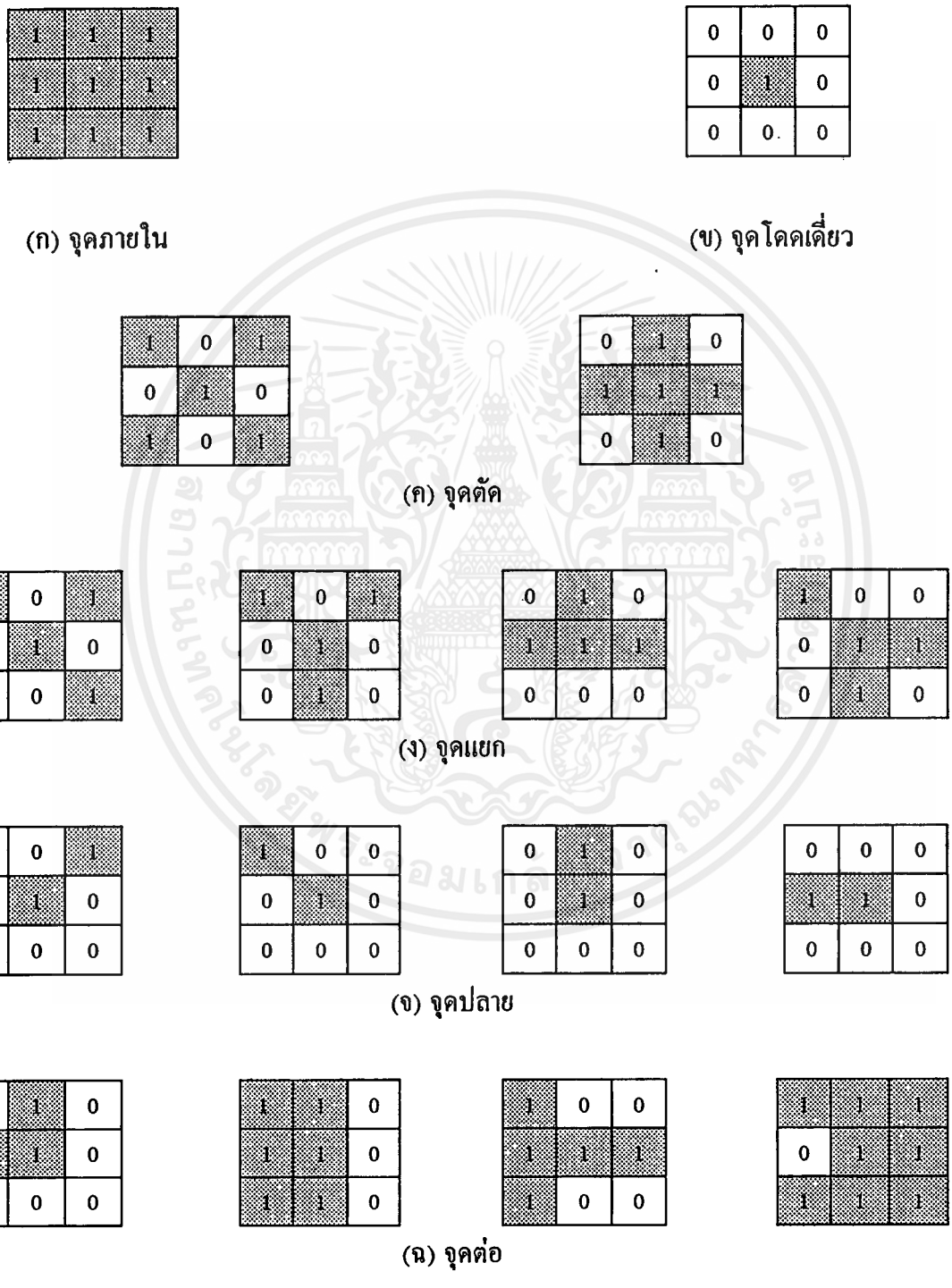
ต้องการหา ซึ่งหน้าต่างขนาด 3 x 3 มีลักษณะ ดังรูปที่ 3.32 ค่า X0, X1, X2, X3, X4, X5, X6, X7, X8 เป็นบิตทางดิจิตอลมีค่าเป็น 0 หรือ 1 โดย X0 คือจุดที่เราจะพิจารณาและจะพิจารณาเฉพาะเมื่อ X0 = 1 เท่านั้น

X4	X3	X2
X5	X0	X1
X6	X7	X8

รูปที่ 3.32 แสดงหน้าต่างขนาด 3x3บิต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการหาค่าคุณสมบัติทางโทโปโลยีสำหรับภาพดิจิตอลทั่วไป สามารถกำหนดความแตกต่างของกลุ่มจุดภาพในตารางหน้าต่าง 3 x 3 ออกได้เป็น 6 กลุ่มดังนี้ จุดภายใน (Internal) , จุดปลาย (End) , จุดตัด (Cross) , จุดต่อ (Connect) , จุดแยก (Branch) และจุดโดดเดี่ยว (Isolate) ดังแสดงในรูปที่ 3.33 (ก) ถึง (ฉ)



รูปที่ 3.33 แสดงลักษณะคุณสมบัติทางโทโปโลยี ของกลุ่มจุดภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.33(ก) เป็นจุดภายใน จะไม่มีส่วนประกอบรอบๆ ที่มีการเปลี่ยนแปลงจากค่า 1, 0 ไปเป็น 0, 1 เลย

รูปที่ 3.33(ข) เป็นจุดโดดเดี่ยว จะไม่มีส่วนประกอบรอบๆ ที่มีการเปลี่ยนแปลงจากค่า 0, 1 ไปเป็น 1, 0 เลย

รูปที่ 3.33(ค) เป็นจุดตัด ที่มีส่วนประกอบรอบๆ ที่มีการเปลี่ยนแปลงจากค่า 1, 0 ไปเป็น 0, 1 อยู่ 8 ช่วง

รูปที่ 3.33(ง) เป็นจุดแยก ที่มีส่วนประกอบรอบๆ ที่มีการเปลี่ยนแปลงจากค่า 1, 0 ไปเป็น 0, 1 อยู่ 6 ช่วง

รูปที่ 3.33(จ) เป็นจุดปลาย ที่มีส่วนประกอบรอบๆ ที่มีการเปลี่ยนแปลงจากค่า 1, 0 ไปเป็น 0, 1 อยู่ 2 ช่วง

ในกรณีของรูปที่ 3.33(ฉ) ได้จัดแสดงไว้เพียงบางส่วนเท่านั้น ซึ่งส่วนที่ไม่มีในรูปที่ 3.33 (ฉ) ก็คือ จุดต่อทั้งหมด เราสามารถสังเกตเห็นความแตกต่างได้ชัดเจน ในการไปจำแนกจุดภาพจากรูป

เมื่อเป็นเช่นนี้เราจะต้องหาฟังก์ชั่นที่จะตรวจสอบได้ว่า จุดที่เป็นส่วนประกอบรอบๆ มีการเปลี่ยนแปลง จากค่า 1,0 ไปเป็น 0,1 อยู่กี่ช่วง เราก็สามารถจำแนกลักษณะของคุณสมบัติทางโทโปโลยีในกลุ่มดังกล่าวได้ ซึ่งฟังก์ชั่นที่กำหนดให้ คือ

$$Ns = \sum(Xk \text{ XOR } Xk\text{-}1) \quad ; k=\{1,2,3,.....,8\}$$

$$Ng = \sum(Xk) \quad ; k=\{1,2,3,4,5,6,7,8\}$$

เมื่อ Ns คือ จำนวนครั้งของการเปลี่ยนแปลงของส่วนประกอบรอบๆ จากค่า 1 เป็น 0 หรือจากค่า 0 เป็น 1

Ng คือ ผลรวมของจำนวนค่าที่เป็น 1 ในแต่ละด้านของส่วนประกอบทั้งแปด

สมการ Ng จะถูกนำมาจำแนกค่าที่ Ns ไม่สามารถจำแนกได้ คือ ระหว่าง จุดภายในกับจุดโดดเดี่ยว และระหว่างจุดปลายกับจุดต่อ สมการทั้งสองสามารถแยกคุณสมบัติทางโทโปโลยีของ กลุ่มจุดภาพได้โดยการนำมาประกอบกัน ดังแสดงในตารางที่ 3.1

คุณสมบัติทางโทโปโลยีคอลของกลุ่มจุดภาพ

ฟังก์ชั่น	Internal	Isolate	Cross	Branch	End	Connect
Ns	0	0	8	6	2	2 4
Ng	8	0	4	3	1	2-7

ตารางที่ 3.1 แสดงค่าคุณสมบัติทางโทโปโลยีคอลของกลุ่มจุดภาพ

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อใช้ค่าจากตารางที่ 1 มาเขียนวิธีการทำงานในลักษณะการตัดสินใจ จะได้ดังนี้

1. เมื่อ Ns = 0 และ Ng = 8 P(X0) = internal
2. เมื่อ Ns = 0 และ Ng = 0 P(X0) = isolate
3. เมื่อ Ns = 8 P(X0) = cross
4. เมื่อ Ns = 6 P(X0) = branch
5. เมื่อ Ns = 2 และ Ng < 2 P(X0) = end
6. กรณีอื่นๆ P(X0) = connect

โดยที่ P(X0) คือ จุดภาพที่กำลังพิจารณาอยู่ว่าเป็นกลุ่มจุดภาพกลุ่มไหน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

การจัดเตรียมข้อมูลเพื่อการรู้จำ

ข้อมูลตัวอักษรภาพที่ใช้ในการรู้จำนี้ เราจะทำการนำมาจากการเก็บภาพจากสันหนังสือ โดยทั่วๆ ไปที่มีลักษณะเป็นพื้นที่ขาว และ ตัวอักษรที่ค่อนข้างเข้ม เพื่อให้กล้อง CCD สามารถแยกถึงความแตกต่างระหว่าง พื้นของสันหนังสือกับตัวอักษรบนสันหนังสือได้อย่างเด่นชัด แต่เนื่องจากการให้แสงที่จะต้องทำให้เกิดจุดรวมแสงน้อยที่สุด คือ ไม่จ้าเกินไปในจุดใดจุดหนึ่งบนสันหนังสือ เพราะอาจทำให้ข้อมูลที่จะนำไปวิเคราะห์ผิดพลาดได้ เมื่อได้ดังนั้นแล้วเราจะทำการเก็บข้อมูลภาพนั้นให้อยู่ในลักษณะของไฟล์ .PCX ซึ่งจะทำการบีบอัดข้อมูล ดังนั้นก่อนจะทำการรู้จำข้อมูล เราจำต้องทำการขยายไฟล์ .PCX ให้เป็นไฟล์ข้อมูลดิบที่ปราศจากการบีบอัด และเราจะทำการนำภาพที่ได้มาทำการแปลงให้เป็นภาพ 2 ระดับ (Binary) คือ มีแค่ ขาวกับดำ หรือ 1 กับ 0 นั่นเอง โดยการตัดค่า Threshold ให้อยู่ในระดับที่เราต้องการเพื่อให้เกิดความชัดเจนของภาพที่จะใช้ในการวิเคราะห์

การจัดเตรียมข้อมูลเพื่อการรู้จำที่จะอธิบายต่อไปนี้ เป็นการจัดเตรียมข้อมูลภาพที่อยู่ในหน่วยความจำแล้ว ซึ่งจะเห็นว่าข้อมูลที่อยู่ในหน่วยความจำนี้จะมีลักษณะอยู่ในสันหนังสือ และมีตัวอักษรอยู่ข้างใน แต่เนื่องจากการวิเคราะห์อักษรภาพเพื่อการรู้จำนั้นจะทำได้เพียงทีละตัวอักษร ซึ่งเราจะต้องทำการแยกสันหนังสือออกจากกันก่อน แล้วจึงทำการแยกตัวอักษรออกจากประโยค และทำการปรับปรุงข้อมูล เพื่อที่จะทำให้ขบวนการรู้จำตัวอักษรนั้นสะดวกขึ้น ดังนั้นมีขั้นตอน ดังต่อไปนี้ คือ

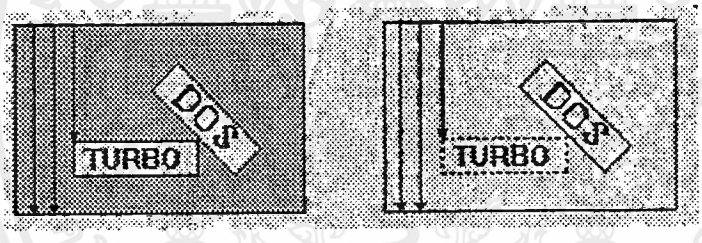
4.1 การแยกสันหนังสือออกจากกลุ่มสันหนังสือ
4.2 การหมุนสันหนังสือให้อยู่ในแนวมาตรฐาน
4.3 การปรับปรุงข้อมูลภาพหลังจากหมุนข้อมูลภาพ
4.4 การแยกข้อมูลภาพตัวอักษรออกจากสันหนังสือ
4.5 การแยกตัวอักษรออกจากประโยค
4.6 การทำภาพตัวอักษรให้บาง (Thinning)
4.7 การจำกัดส่วนเกิน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1 การแยกสันหนังสือออกจากกลุ่มของสันหนังสือ

การแยกภาพสันหนังสือแต่ละเล่มออกจากกลุ่มของสันหนังสือเป็นขั้นตอนแรกที่จะต้องกระทำเพื่อการวิเคราะห์ภาพตัวอักษร ข้อมูลภาพตัวอักษรที่ได้ มาจากกล้อง CCD ภาพ Black ground จะเป็นสีดำหรือ 1 ตัวอักษรสีดำหรือพื้นของตัวอักษรจะเป็นสีขาวหรือข้อมูลภาพสันหนังสือ แต่ละเล่มจะแบ่งแยกกันอย่างเด่นชัด และมีขอบเขตสันหนังสือแต่ละเล่มที่เป็นอิสระซึ่งกันและกัน

ดังนั้นการแยกสันหนังสือแต่ละเล่มออกจากกลุ่มของหนังสือ โดยการตรวจกวาดภาพจากมุมบนซ้ายจนถึงมุมล่างขวา จากด้านซ้ายมาด้านขวา เมื่อพบจุดภาพที่เป็น 0 ให้ติดตามขอบภาพของสันหนังสือพร้อมบันทึกตำแหน่งทุกตำแหน่งและเข้ารหัสขอบเขตของสันหนังสือ จนครบรอบของสันหนังสือที่อยู่ภายในขอบเขตนั้น ๆ มาและทำการลบข้อมูลภาพในขอบเขตของสันหนังสือเดิม หลังจากนั้นข้อมูลภาพที่คัดลอกมาวิเคราะห์ข้อมูลต่อไป รูปที่ 4.1 แสดงลักษณะของ ภาพและการ scan ภาพ



a) แสดงภาพข้อมูลภาพสันหนังสือและการ scan หาสันหนังสือ

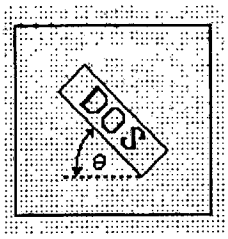
b) แสดงภาพสันหนังสือที่ถูกติดตามขอบภาพและถูกคัดลอกออกไป

รูปที่ 4.1 แสดงลักษณะของภาพและการ scan ภาพเพื่อหาสันหนังสือ

4.2 การหมุนสันหนังสือให้อยู่ในแนวมาตรฐาน

หลังจากคัดลอกข้อมูลสันหนังสือมาแล้วภาพสันหนังสือจะทำมุมต่างๆ กับแนวระนาบดังนั้นจึงต้องหามุมของสันหนังสือกับแนวระนาบ ดังทฤษฎีการหามุมของลูปมาประยุกต์ใช้การหามุมของสันหนังสือกับแนวระนาบ เมื่อได้มุมของสันหนังสือกับแนวระนาบแล้ว เราจะทำการหมุนภาพตามองศาของสันหนังสือเพื่อให้อยู่ในแนวระนาบสำหรับการรู้จำตัวอักษรต่อไป

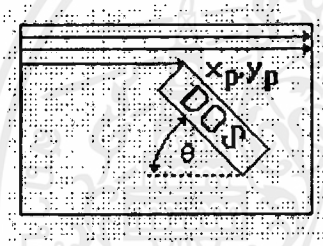
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



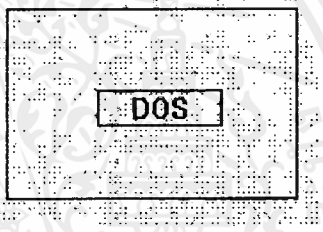
รูปที่ 4.2 แสดงสันหนังสือทำมุมกับแนวระนาบ

ข้อกำหนดของการหมุนข้อมูลภาพ

1. เมื่อข้อมูลภาพสันหนังสือทำมุม $0^\circ < \theta \leq 90^\circ$ เราจะตรวจกวาดภาพจากมุมบนซ้ายไปมุมขวาล่าง จากบนลงล่างเมื่อเจอจุดภาพให้จุดนั้นเป็นจุดหมุนของภาพ (x_p,y_p) แล้วหมุนข้อมูลภาพตามขนาดของมุมที่ทำกับแนวระนาบในทิศทางทวนเข็มนาฬิกา

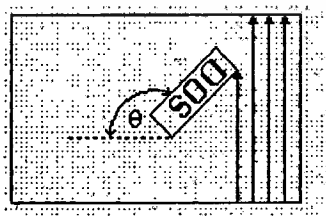


รูปที่ 4.3 แสดงการหมุนภาพของสันหนังสือที่ทำมุม $0^\circ < \theta \leq 90^\circ$

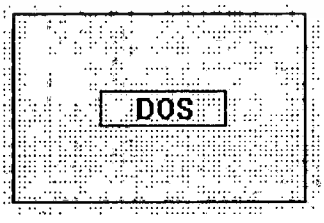


รูปที่ 4.4 แสดงข้อมูลสันหนังสือหลังจากหมุนข้อมูล

2. เมื่อข้อมูลภาพสันหนังสือทำมุม $90^\circ < \theta < 180^\circ$ เราจะตรวจกวาดข้อมูลภาพจากมุมล่างขวาไปมุมบนซ้ายทิศทางจากขวามาซ้าย เมื่อพบจุดภาพให้จุดนั้นเป็นจุดหมุนของจุดภาพ (x_p,y_p) แล้วหมุนข้อมูลภาพตามขนาดของมุมที่ทำกับระนาบในทิศทางทวนเข็มนาฬิกา



รูปที่ 4.5 แสดงปรับมุมภาพของสันหนังสือที่ทำมุม $90^\circ < \theta < 180^\circ$

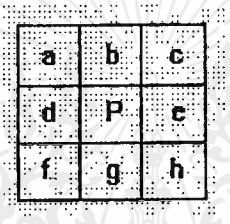


รูปที่ 4.6 แสดงข้อมูลสันหนังสือหลังจากหมุนข้อมูลแล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 การปรับปรุงภาพหลังจากหมุนข้อมูลภาพ

การหมุนภาพ (Rotate picture) นั้น จากสมการการหมุนภาพ เพื่อกำหนดตำแหน่งข้อมูลใหม่ตามขนาดของมุมที่ต้องการหมุนทวนเข็มนาฬิกา เนื่องจาก function ทางคณิตศาสตร์การกำหนดตำแหน่งใหม่ให้กับข้อมูลภาพอาจเกิดความผิดพลาดได้ในส่วนภาพตัวอักษรหลังจากหมุนภาพแล้วอาจเกิด ส่วนที่เป็นพื้นของตัวอักษรอาจมีส่วนที่เป็นภาพของตัวอักษรแทรกอยู่ก็ได้ และส่วนที่เป็นภาพตัวอักษรอาจจะมีส่วนที่เป็นพื้นของตัวอักษรแทรกอยู่ก็ได้ เพื่อขจัดความผิดพลาดที่อาจจะเกิดขึ้น ในขั้นตอนการวิเคราะห์ภาพตัวอักษรจึงต้องมีขบวนการ ปรับปรุงภาพก่อนการวิเคราะห์ภาพตัวอักษรพิจารณาดังสมการ



รูปที่ 4.7 แสดงตารางหน้าต่างที่กำหนดในการปรับปรุงข้อมูลภาพ

จากรูปกำหนดให้ P เป็นจุดที่กำลังทำการพิจารณา
การพิจารณามี 2 เงื่อนไขดังนี้

1. P เป็นจุดภาพ

(a|b|d)&(e|h|g)|(b|c|e)&(g|f|d)

- ถ้าเงื่อนไขเป็นจริง กำหนดให้ P เป็น 1
- ถ้าเงื่อนไขเป็นเท็จ กำหนดให้ P เป็น 0

2. P ไม่ใช่จุดภาพ

b&g&(d|e) | d&e&(d|g)

- ถ้าเงื่อนไขเป็นจริง กำหนดให้ P เป็น 1
- ถ้าเงื่อนไขเป็นเท็จ กำหนดให้ P เป็น 0

หลังจากปรับปรุงข้อมูลภาพจะได้ความสมบูรณ์และคมชัดขึ้น พร้อมที่จะนำไปวิเคราะห์ข้อมูลภาพต่อไปได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 การวิเคราะห์และระบุส่วนประกอบของสันหนังสือ

การวิเคราะห์และระบุส่วนประกอบของสันหนังสือ เป็นขบวนการทำงานหนึ่งในส่วนของการจัดการล่วงหน้า (Pre-processing process) ของระบบการรู้จำตัวอักษร (character recognition) ส่วนประกอบต่าง ๆ บนสันหนังสือได้กลายเป็นอุปสรรคหนึ่งของการรู้จำของระบบการรู้จำตัวอักษร ตัวอย่างเช่น บนสันหนังสือหนึ่งอาจจะไม่ได้มีเพียงชื่อของหนังสือล้วนๆ พิมพ์อยู่อย่างเดียว แต่อาจจะประกอบด้วยรูปภาพ ข้อความอื่นๆ ที่มิใช่ชื่อของหนังสือ ภาพโลโกกราฟฟิก ภาพสัญลักษณ์ พิมพ์รวมอยู่บนสันหนังสือด้วย ซึ่งกรณีนี้ถ้าไม่ได้มีการแยกแยะส่วนประกอบต่างๆ บนสันหนังสือให้ถูกต้องแล้วอาจทำให้การรู้จำตัวอักษรผิดพลาดได้ หรือทำไม่ได้เลยในกรณีที่ข้อมูลที่ส่งผ่านให้กับส่วนของการรู้จำตัวอักษรกลายเป็นรูปภาพไม่ใช่ภาพตัวอักษร เพราะฉะนั้นขบวนการวิเคราะห์และระบุส่วนประกอบของสันหนังสือ จึงเป็นขั้นตอนการ ทำงาน ที่สำคัญมากและจะขาดเสียไม่ได้ เพื่อการจัดเตรียมข้อมูลให้กับระบบการรู้จำตัวอักษรจะเป็นไปได้อย่างถูกต้อง

ทฤษฎีที่นำมาใช้

ทฤษฎีเซลลูล่าออโตมาตา (Cellular Automata)

เทคนิคที่นำมาประยุกต์ใช้ในการวิเคราะห์ และระบุส่วนประกอบของสันหนังสือโดยอาศัยทฤษฎีเซลลูล่าออโตมาตา ซึ่งเป็นทฤษฎีว่าด้วยการแก้ปัญหาทฤษฎีหนึ่ง อาศัยหลักการที่กล่าวว่าการแก้ปัญหาไม่ว่าจะเป็นเรื่องใดๆ ก็ตาม จะยึดหลักแบ่งปัญหาออกเป็นส่วนย่อยๆ แล้วค่อยๆ จัดการกับส่วนย่อยๆ เหล่านั้นก่อน เมื่อจัดการกับองค์ประกอบย่อยๆ เหล่านั้นได้ ก็จะทำให้สามารถจัดการกับปัญหาใหญ่ทั้งหมดได้ ส่วนย่อย ๆ ที่ถูกแบ่งออกจะเรียกอีกอย่างหนึ่งว่า เซล (Cell) การจัดการกับเซลแต่ละเซลแล้วจะสามารถจัดการกับระบบรวมหรือปัญหาใหญ่ๆ ได้ทั้งหมดนั้น มาจากความจริงที่ว่า ปัญหาใหญ่ ที่สามารถแบ่งแยกย่อยออกไปได้นั้น จะเกี่ยวข้องกับปัญหาย่อยหรือ เซลที่ถูกแยกออกมาและเมื่อไล่แก้ปัญหาย่อยก็เหมือนกับแก้ปัญหาใหญ่ที่ละเล็กทีละน้อย จนในที่สุดเมื่อแก้ปัญหาย่อยได้หมด ก็เป็นการแก้ปัญหาใหญ่ได้ทั้งหมดด้วย

การแก้ปัญหาส่วนย่อยหรือเซล นอกจากจะใช้สถานะ หรือรายละเอียดที่ได้จากภายในเซลเองแล้ว ก็ยังสามารถใช้สถานะของเซลอื่นๆ ที่อยู่รอบๆ เซลนั้นมาใช้ช่วยในการแก้ปัญหาได้ด้วย ซึ่งสุดท้ายจะนำไปสู่การแก้ปัญหาใหญ่ทั้งหมด

วิธีการแก้ปัญหาของทฤษฎีเซลลูล่าออโตมาตายึดหลักใหญ่ๆ อยู่ 2 ข้อคือ

1. แยกสิ่งที่สนใจและต้องการแก้ไขหรือหาคำตอบ ออกเป็นส่วนย่อยๆ หรือเซล
2. แก้ไขปัญหาส่วนย่อยทีละส่วนให้หมด โดยอาศัยข้อมูลจากภายในส่วนย่อย และข้อมูลจากส่วนย่อยที่อยู่รอบๆ นั้นด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4.1 การกำหนดขอบเขต (Blocking zones)

การทำงานในขั้นตอนนี้เริ่มต้นด้วยการเตรียมข้อมูลภาพสันหนังสือในลักษณะของข้อมูลไบนารี่ นั่นคือแต่ละจุดของภาพจะถูกแทนด้วยตัวเลขฐานสองเท่านั้น คือตัวเลข 0 กับ 1 โดยที่จุดภาพที่มีค่าเป็น 0 จะแทนพื้นสีขาวหรือช่องว่างบนหน้ากระดาษ และจุดภาพที่มีค่าเป็น 1 จะแทนส่วนที่เป็นลายเส้นของตัวอักษรหรือลายเส้นของรูปภาพ โดยจะเตรียมข้อมูลลักษณะนี้ไว้ 2 ชุดเพื่อให้ประกอบการวิเคราะห์

สำหรับข้อมูลชุดที่หนึ่ง จุดภาพแต่ละจุดจะถูกกำหนดสถานะของจุด โดยพิจารณาจากจุดภาพที่ล้อมรอบจุดที่กำลังพิจารณา นับจำนวนจุดภาพรอบๆ ที่มีค่าเป็น 1 และพิจารณาลักษณะของจุดภาพรอบๆ ที่ปรากฏ ซึ่งจะถูกนำมาใช้เป็นเงื่อนไขในการกำหนดสถานะของจุด โดยจะมีเงื่อนไขอยู่ด้วยกัน 4 เงื่อนไขดังต่อไปนี้

1. ถ้าจุดภาพที่กำลังพิจารณาอยู่มีค่าเป็น 1 และจำนวนจุดภาพรอบข้างมีจำนวนน้อยกว่าจำนวนที่กำหนดไว้ แล้วจุดภาพที่กำลังพิจารณาจะถูกกำหนดให้เป็น จุดตาย เนื่องจากมีจำนวนจุดภาพรอบข้างน้อยเกินไป

2. ถ้าจุดภาพที่กำลังพิจารณาอยู่มีค่าเป็น 1 และจำนวนจุดภาพรอบข้างมีจำนวนมากกว่าหรือเท่ากับจำนวนที่กำหนดไว้ แล้วจุดภาพที่กำลังพิจารณาจะถูกกำหนดให้เป็น จุดเป็น

3. ถ้าจุดภาพที่กำลังพิจารณาอยู่มีค่าเป็น 0 และจำนวนจุดภาพรอบข้างมีจำนวนมากกว่าหรือเท่ากับจำนวนที่กำหนดไว้ แล้วจุดภาพที่กำลังพิจารณาจะถูกกำหนดให้เป็น จุดเกิด

4. ถ้าจุดภาพที่กำลังพิจารณาอยู่มีค่าเป็น 0 และจำนวนจุดภาพรอบข้างมีจำนวนมากกว่าหรือเท่ากับจำนวนที่กำหนดไว้ แล้วจุดภาพที่กำลังพิจารณาอยู่จะยังคงสถานะเดิมต่อไปและค่าของจุดภาพยังคงมีค่าเหมือนเดิมไม่เปลี่ยนแปลง

เมื่อทราบสถานะของจุดภาพจุดหนึ่งแล้ว จะทำการเปลี่ยนค่าของจุดภาพให้มีค่าที่สอดคล้องกับสถานะของจุดภาพนั้นๆ แต่การเปลี่ยนค่าของจุดภาพจะกระทำกับจุดภาพในข้อมูลภาพชุดที่สอง ณ. ตำแหน่งพิกัดของจุดภาพเดียวกันกับพิกัดของจุดภาพในข้อมูลภาพชุดที่หนึ่ง ทั้งนี้ก็เพราะว่าจุดภาพที่ถูกเปลี่ยนค่าตามสถานะใหม่แล้วจะได้ไม่ส่งผลกระทบต่อการพิจารณากำหนดสถานะของจุดภาพต่อไปในข้อมูลภาพต้นฉบับสำหรับเงื่อนไขในการเปลี่ยนค่าของจุดภาพใหม่ถูกกำหนดโดยเงื่อนไข 3 ลักษณะดังนี้

1. จุดตาย ซึ่งเดิมมีค่าเป็น 1 จะถูกเปลี่ยนให้มีค่าเป็น 0

2. จุดเป็น ซึ่งเดิมมีค่าเป็น 1 ยังคงให้มีค่าเป็น 1 ต่อไป

3. จุดเกิด ซึ่งเดิมมีค่าเป็น 0 จะถูกเปลี่ยนให้มีค่าเป็น 1

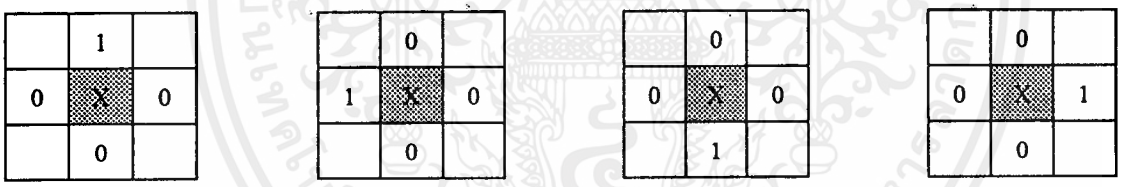
การพิจารณากำหนดสถานะของจุดภาพและเปลี่ยนค่าใหม่ให้กับจุดภาพจะกระทำกับทุกๆ จุดภาพในข้อมูลภาพ เริ่มจากจุดภาพที่มุมบนซ้ายมือ และกวาดตรวจไปบนคอลัมน์ในทิศทางจากซ้ายไปขวา และเลื่อนแถวลงมาในทิศทางจากบนลงล่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การประยุกต์ใช้งานทฤษฎีเซลลูล่าออโตมาตากับการกำหนดขอบเขตของส่วนประกอบของสันหนังสือในงานวิจัยเรื่องการรู้จำตัวอักษรของวิทยานิพนธ์ฉบับนี้ มีการปรับแต่งเงื่อนไขของการพิจารณาสถานะของจุดภาพบางประการให้เหมาะสมกับวัตถุประสงค์ของงานมากขึ้น เนื่องจากจุดมุ่งหมายของการกำหนดขอบเขตของสันหนังสือคือ ขอบเขตในลักษณะสี่เหลี่ยมที่เป็นตัวหนังสือหรือขอบเขตสี่เหลี่ยมที่เป็นรูปภาพ ดังนั้นการพิจารณาจุดภาพที่ล้อมรอบจุดภาพที่กำลังพิจารณา จะละเว้นไม่พิจารณาจุดภาพในแนวเส้นทะแยงมุมทั้งสอง คงเหลือพิจารณาเฉพาะจุดรอบข้างที่อยู่ด้านบน ด้านขวา ด้านล่าง และด้านซ้าย รวม 4 จุดเท่านั้น ซึ่งจะช่วยในการค้นหาขอบเขตในลักษณะสี่เหลี่ยมทำได้ง่ายขึ้น อีกทั้งยังช่วยลดการคำนวณลงได้อีกมาก เป็นผลให้การทำงานเป็น ไปได้รวดเร็วยิ่งขึ้นด้วย

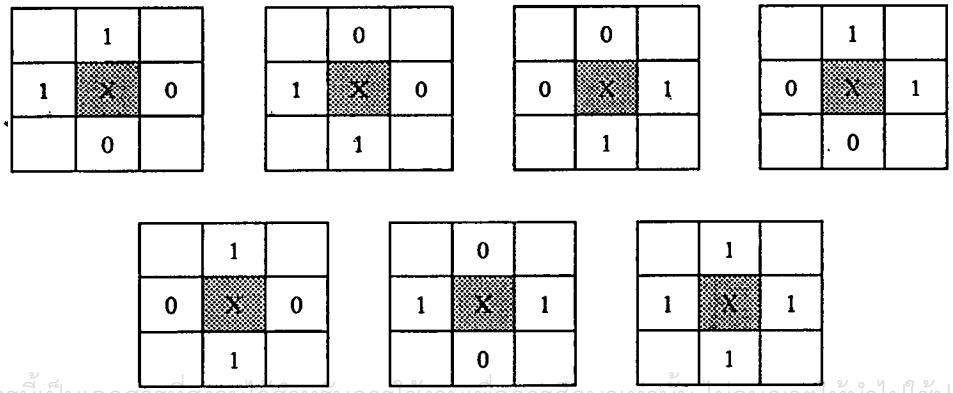
เงื่อนไขการพิจารณาสถานะของจุดของภาพที่ถูกประยุกต์ตามเหตุผลที่ได้กล่าวมาแล้วข้างต้นมีรายละเอียดของเงื่อนไขดังนี้

1. ถ้ากำหนดให้จุด x เป็นจุดใดๆ ที่มีค่าเป็น 1 แล้วจุด x จะมีสถานะเป็นจุดตาย ก็ต่อเมื่อมีจุดรอบข้างที่มีค่าเป็น 1 น้อยกว่า 2 จุด

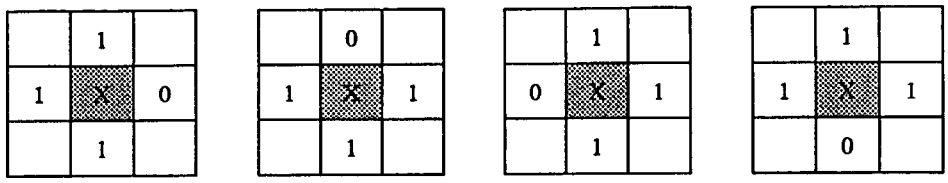


รูปที่ 4.8 แสดงลักษณะของจุดตาย

2. ถ้ากำหนดให้จุด x เป็นจุดใดๆ มีค่าเป็น 1 แล้ว จุด x จะมีสถานะเป็นจุดเป็น ก็ต่อเมื่อมีจุดรอบข้างที่มีค่าเป็น 1 มากกว่าหรือเท่ากับ 2 จุด

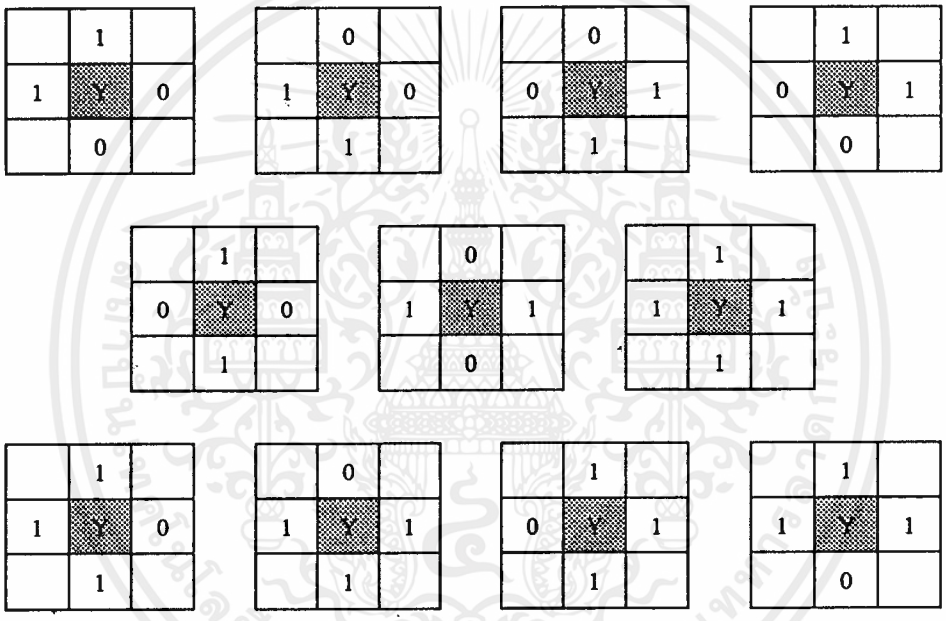


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.9 แสดงลักษณะของจุดเป็น

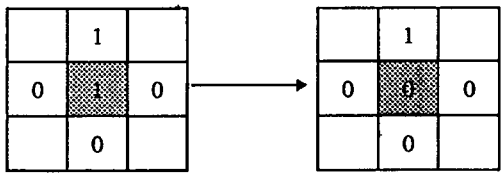
3. ถ้ากำหนดให้จุด y เป็นจุดใดๆ ที่มีค่าเป็น 0 แล้ว จุด y จะมีค่าสถานะเป็นจุดเกิด ก็ต่อเมื่อมีจุดรอบข้างที่มีค่าเป็น 1 มากกว่าหรือเท่ากับ 2 จุด



รูปที่ 4.10 แสดงลักษณะของจุดเกิด

สำหรับเงื่อนไขของการเปลี่ยนค่าใหม่ของจุดภาพตามสถานะของจุดภาพที่หาได้นั้นยังคงใช้เงื่อนไขเดิม ซึ่งมีรายละเอียดดังต่อไปนี้

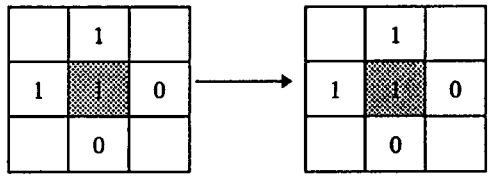
1. จุดตาย จุดภาพที่มีสถานะเป็นจุดตาย ซึ่งค่าของจุดภาพนั้นมีค่าเป็น 1 จะถูกเปลี่ยนให้มีค่าเป็น 0 ดังตัวอย่างในรูปที่ 4.11



รูปที่ 4.11 แสดงตัวอย่างของจุดตายที่มีค่าเป็น 1 ถูกเปลี่ยนให้มีค่าเป็น 0

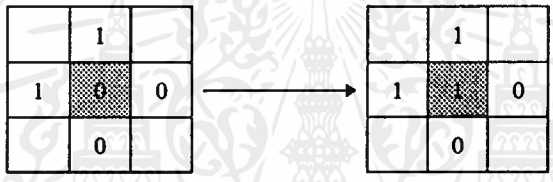
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. จุดเป็น จุดภาพที่มีสถานะเป็นจุดเป็น ซึ่งค่าของจุดภาพนั้นมีค่าเป็น 1 จะยังคงมีค่าเป็น 1 ต่อไปไม่เปลี่ยนแปลง ดังตัวอย่างในรูปที่ 4.12



รูปที่ 4.12 แสดงตัวอย่างของจุดเป็นที่มีค่าเป็น 1 ยังคงมีค่าเป็น 1 ต่อไป

3. จุดเกิด จุดภาพที่มีสถานะเป็นจุดเกิด ซึ่งค่าของจุดภาพนั้นมีค่าเป็น 0 จะถูกเปลี่ยนให้มีค่าเป็น 1 ดังตัวอย่างในรูปที่ 4.13



รูปที่ 4.13 แสดงตัวอย่างของจุดเกิดที่เดิมมีค่าเป็น 0 ถูกเปลี่ยนให้มีค่าเป็น 1

การประยุกต์ใช้ทฤษฎีเซลลูล่าออโตมาตา ในการกำหนดขอบเขตส่วนประกอบของสันหนังสือ จะสามารถกำหนดได้เพียงโครงร่าง (Layout) ของขอบเขตของส่วนประกอบของสันหนังสือเท่านั้น ขอบเขตต่างๆ ของส่วนประกอบบนสันหนังสือยังคงอยู่กระจัดกระจายบนสันหนังสือซึ่งยังไม่เหมาะที่จะแยกคัดลอกออกมา จึงจำเป็นต้องมีขั้นตอนการทำงานอีกขั้นตอนหนึ่งที่จะจับกลุ่มของขอบเขตต่างๆ ที่กระจัดกระจายกันอยู่ รวมเข้าด้วยกันให้ถูกต้องและเหมาะสม ตัวแปรตัวหนึ่งที่ถูกนำมาใช้เป็นตัวกำหนดการจับรวมกลุ่มของขอบเขตของส่วนประกอบของหน้าเอกสารที่กระจัดกระจายกันอยู่ คือ ค่าความหยาบของจุดภาพ

ค่าความหยาบของจุดภาพ (Coarsenness value) เป็นค่าคงที่ใดๆ ที่ใช้กำหนดระยะห่างของจุดภาพในขอบเขตต่างๆ ที่กระจัดกระจายอยู่ ตัวอย่างเช่น ถ้าจุดภาพของ 2 ขอบเขตใดๆ อยู่ในช่วงระยะห่างของค่าความหยาบของจุดภาพที่กำหนด ก็จะทำการรวม 2 ขอบเขตนั้นเข้าเป็นขอบเขตหรือพื้นที่เดียวกัน เป็นต้น ในการกำหนดค่าความหยาบของจุดภาพ ถ้ากำหนดให้ค่าความหยาบของจุดภาพมีค่าน้อยเกินไป จะทำให้ไม่สามารถจับรวมกลุ่มของขอบเขตที่กระจัดกระจายกันอยู่ได้ และถ้ากำหนดให้ค่าความหยาบของจุดภาพมีค่ามากเกินไป ก็จะทำให้ขอบเขตที่กระจัดกระจายอยู่รวมกลุ่มกันเป็นขอบเขตใหญ่เพียงขอบเขตเดียว ซึ่งจะไม่สามารถแยกส่วน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สันหนังสือได้ ดังนั้นการกำหนดขอบเขตของส่วนประกอบของสันหนังสือที่ถูกต้องและเหมาะสม จึงจำเป็นจะต้องกำหนดค่าความหยาบของจุดภาพที่ใช้ให้เหมาะสมด้วยเช่นกัน

ค่าความหยาบของจุดภาพที่เหมาะสม ที่จะสามารถกำหนดขอบเขตของส่วนประกอบของสันหนังสือได้อย่างถูกต้อง จะไปสอดคล้องกับตัวแปรอีกตัวหนึ่งนั่นก็คือ ค่าความละเอียดของจุดภาพของข้อมูลภาพ (Image resolution) ซึ่งค่าความละเอียดของจุดภาพเป็นค่าคงที่ ค่าความหยาบของจุดภาพที่สอดคล้องกับค่าความละเอียดของจุดภาพ สามารถกำหนดค่าขึ้นได้จากการทดลอง

ค่าความละเอียดของจุดภาพ และ ค่าความหยาบของจุดภาพที่เหมาะสม และสามารถนำไปใช้งานทั่วๆ ไปได้ จะขึ้นอยู่กับส่วนประกอบต่างๆ บนสันหนังสือ เช่น ภาพกราฟฟิก โลโก ข้อความ มีส่วนประกอบ กันอย่างหนาแน่นหรือเบาบาง ถ้ามีส่วนประกอบ อย่างหนาแน่น ก็ให้ค่าความหยาบน้อยหรือมีส่วนประกอบแบบเบาบางก็ให้ค่าความหยาบมากๆ ได้

4.4.2 การแยกและคัดลอกขอบเขต (Extracting rectangular zones)

การคัดลอกขอบเขตที่เป็นส่วนประกอบของสันหนังสือออกจากภาพสันหนังสือได้นำเทคนิคการติดตามรอยขอบของภาพมาใช้ในการแยกและคัดลอกขอบเขตออก การทำงานคือ ตรวจกวาดจุดภาพจากจุดมุมบนด้านซ้ายไปในทิศทางจากซ้ายไปขวา และเลื่อนลงจากบนลงล่าง เมื่อตรวจกวาดมาพบจุดภาพที่เป็นจุดใดจุดหนึ่งของขอบเขต ซึ่งมีค่าของจุดภาพเป็น 1 ก็จะเคลื่อนที่ติดตามรอยขอบของขอบเขตไป โดยอาศัยเทคนิคการติดตามรอยขอบของภาพ ในขณะที่เคลื่อนที่ติดตามรอยขอบของภาพไป ตำแหน่งหรือพิกัดของจุด (coordinate) ที่เป็นขอบของภาพจะถูกบันทึกเก็บไว้ด้วย และเมื่อเคลื่อนที่ไปจนรอบนั่นคือกลับมาอยู่ที่จุดเริ่มต้นแล้ว ก็จะนำพิกัดของจุดที่เป็นขอบภาพทั้งหมดมาคำนวณหาจุดพิกัดที่มีค่ามากที่สุดและน้อยที่สุดทั้งในแนวนอนและแนวตั้ง มากำหนดเป็นจุดมุมทั้งสี่ของรูปสี่เหลี่ยม แต่ละจุดมุมของรูปสี่เหลี่ยมอธิบายได้ดังต่อไปนี้

1. พิกัดของจุดมุมบนด้านซ้ายคือ พิกัดของจุดขอบที่มีค่าน้อยที่สุดในแนวตั้งหรือแนวแกน y กับพิกัดของจุดขอบที่มีค่าน้อยที่สุดในแนวนอนหรือแนวแกน x ($y_{min}x_{min}$)

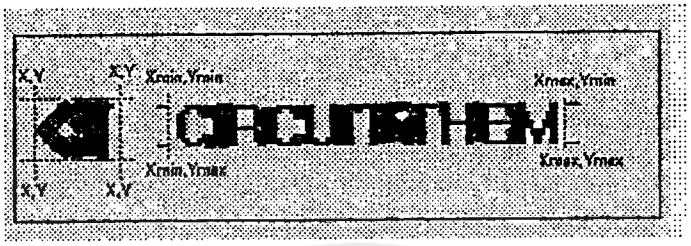
2. พิกัดของจุดมุมบนด้านขวาคือ พิกัดของจุดขอบที่มีค่าน้อยที่สุดในแนวตั้งหรือแนวแกน y กับพิกัดของจุดขอบที่มีค่ามากที่สุดในแนวนอนหรือแนวแกน x ($y_{min}x_{max}$)

3. พิกัดของจุดมุมล่างด้านซ้ายคือ พิกัดของจุดขอบที่มีค่ามากที่สุดในแนวตั้งหรือแนวแกน y กับพิกัดของจุดขอบที่มีค่าน้อยที่สุดในแนวนอนหรือแนวแกน x ($y_{max}x_{min}$)

4. พิกัดของจุดมุมล่างด้านขวาคือ พิกัดของจุดขอบที่มีค่ามากที่สุดในแนวตั้งหรือแนวแกน y กับพิกัดของจุดขอบที่มีค่ามากที่สุดในแนวนอนหรือแนวแกน x ($y_{max}x_{max}$)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

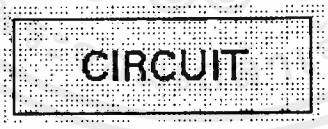
โดยกำหนดให้จุดมุมบนด้านซ้ายของข้อมูลภาพทั้งหมดเป็นจุดเริ่มต้น (Origin point) มีพิกัดของจุดเป็น (0,0) แนวแกน x คือ แกนอ้างอิงในแนวนอน และแนวแกน y คือแกนอ้างอิงในแนวตั้งดังแสดงในรูปที่ 4.14



รูปที่ 4.14 แสดงพิกัดขอบเขตของสันหนังสือ

เมื่อคำนวณหาจุดมุมทั้ง 4 ของรูป สี่เหลี่ยมได้แล้วก็สามารถกำหนดขอบเขตของส่วนประกอบของสันหนังสือได้ในลักษณะของพื้นที่รูปสี่เหลี่ยมจุดภาพทั้งหมดที่อยู่ภายใน บริเวณรูปสี่เหลี่ยม จะนำมาตัดสินใจว่าใช้จุดภาพที่เราจะทำการวิเคราะห์หรือไม่ โดยสามารถหาได้โดยหาค่าความยาว ในแนวแกน x ที่ทุกๆ ขอบเขตบนสันหนังสือขอบเขตใดมีค่าสูงสุดจะเป็นส่วนของจุดภาพที่เป็นชื่อของสันหนังสือ จุดภาพที่อยู่ภายในขอบเขตนั้นแล้วนำส่วนที่คัดลอกมาทำการวิเคราะห์ภาพตัวอักษรต่อไป

4.5 การแยกภาพตัวอักษรออกจากประโยค



รูปที่ 4.15 แสดงภาพตัวอักษรแต่ละตัวในรูปประโยค

การแยกภาพอักษรออกจากภาพประโยคจะไม่นำวิธีการติดตามขอบภาพตัวอักษรมาใช้ เพราะในกรณีของตัว i และตัว j เราไม่สามารถรวมสัญลักษณ์ทั้งสองอย่างเข้าเป็นรูปแบบในขอบเขตเดียวกันได้จึงได้ใช้การแยกตัวอักษรจากภาพประโยคดังนี้

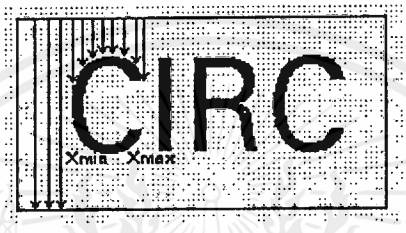
1. ทำกวาดภาพจากมุมบนซ้ายมือไปยังมุมขวาล่าง ทิศทางจากซ้ายไปขวาเมื่อพบจุดภาพกำหนดให้เป็น "x_{min}"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

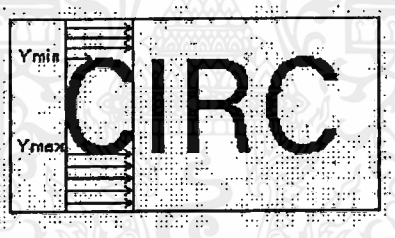
2. ทำการกวาดภาพต่อไปในลักษณะเดิมแต่จะเปลี่ยนรูปแบบของการตรวจสอบใหม่ คือ ตรวจสอบในแนวแกน y ทั้งแกน ถ้าไม่พบจุดภาพเลยก็จะลบค่าในแกน x ลง 1 แล้วกำหนดให้จุดนั้นเป็นจุด "x_{max}"

3. ทำการกวาดภาพจากมุมบนซ้ายไปยังมุมล่างขวามือทิศทางจากบนลงล่าง กำหนดให้ค่าเริ่มต้นในแกน $x = x_{min}$ จุดสิ้นสุดในแนวแกน $x = x_{max}$ เมื่อพบจุดภาพ กำหนดให้เป็น "y_{min}"

4. ทำการกวาดภาพจากมุมล่างซ้ายไปยังมุมบนขวาทิศทางจากล่างขึ้นบน กำหนดค่าเริ่มต้นในแกน $x = x_{min}$ จุดสิ้นสุดในแกน $x = x_{max}$ เมื่อกวาดพบจุดภาพ กำหนดให้เป็น "y_{max}"



รูปที่ 4.16 แสดงการตรวจหาขนาดของตัวอักษรในแนวแกน X

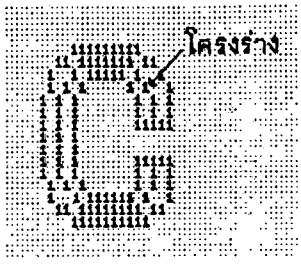


รูปที่ 4.17 แสดงการตรวจหาขนาดของตัวอักษรในแนวแกน Y

4.6 การทำให้บาง Thinning

รูปแบบ 2 ระดับที่ได้มาจากการ scan จะมีภาพของตัวอักษรในรูปแบบที่มีขนาดความกว้างมากกว่า 1 จุดภาพเสมอ แต่ในระบบการจดจำที่ใช้คุณสมบัติทาง Topology ของรูปแบบเป็นตัวจำแนกกลุ่มนั้น จะต้องการเฉพาะส่วนที่เป็นโครงร่าง (Skeleton) ของตัวอักษรเท่านั้น โครงร่างของตัวอักษรก็คือ ภาพของตัวอักษรที่มีความกว้างเพียง 1 จุดภาพ และอยู่ในตำแหน่งที่สมมาตรบนเนื้อของตัวอักษร ดังแสดงในรูปที่ 4.18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.18 โครงร่างของตัวอักษร

ดังนั้นจึงต้องทำตัวอักษรในรูปแบบให้บางลง ด้วยวิธีการทำให้บาง (Thinning) วิธีการทำให้บางนี้ก็มีการวิจัยและพัฒนากันหลาย ๆ วิธีแต่ในวิทยานิพนธ์เล่มนี้เราใช้ NC_8

n_3	n_2	n_1
n_4	P	n_0
n_5	n_6	n_7

รูปที่ 4.19 ตารางหน้าต่างสำหรับการทำให้บาง

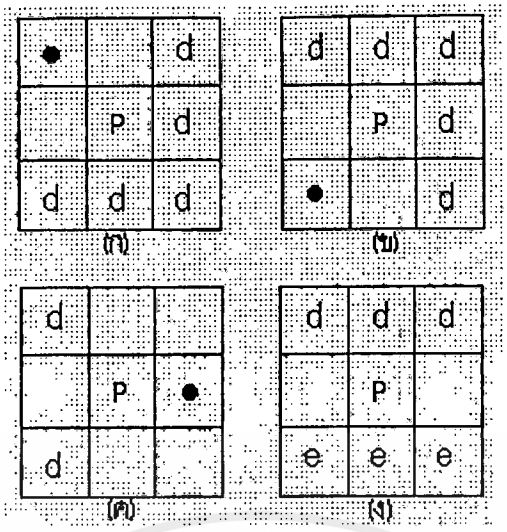
ตารางหน้าต่างในรูปที่ 4.19 ใช้นำมาพิจารณาส่วนของภาพตัวอักษรบนรูปแบบที่ต้องการทำให้บาง โดยที่จุดข้างเคียงของจุดภาพ P ซึ่งก็คือ จุดภาพ n_0 , n_1 , n_2 ,...., n_7 จะเป็นตัวกำหนดลักษณะขอบ (Edge) ของจุดภาพ P ใน 4 ลักษณะคือ

1. จุดภาพ P เป็นขอบซ้ายถ้ามี n_4 เป็นจุดสว่าง ($n_4 = 0$)
2. จุดภาพ P เป็นขอบขวาถ้ามี n_0 เป็นจุดสว่าง ($n_0 = 0$)
3. จุดภาพ P เป็นขอบบนถ้ามี n_2 เป็นจุดสว่าง ($n_2 = 0$)
4. จุดภาพ P เป็นขอบล่างถ้ามี n_6 เป็นจุดสว่าง ($n_6 = 0$)

ในขบวนการทำให้บางถือว่าขอบของภาพมิใช่โครงร่างของตัวอักษร ดังนั้น เมื่อพบขอบก็จะกำหนดค่าของจุดภาพนั้นให้มีค่าค่าหนึ่ง ที่ไม่เป็น 1 (flagging) แต่ถ้าพบจุดภาพที่มีความกว้างเพียง 1 จุดภาพก็จะกำหนดค่าของจุดภาพนั้นให้เป็นค่า 1 (unflagging)

การทดสอบจุดภาพที่เป็นขอบ กระทำโดยการเปรียบเทียบตารางหน้าต่างในรูปที่ 4.19 ที่สแกนจากบนลงล่างบนรูปแบบของตัวอักษร กับตารางหน้าต่างในรูปที่ 4.20

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.20 ตารางหน้าต่างสำหรับการตรวจสอบหาขอบ

จากรูปที่ 4.20 P และ ● เป็นจุดมืด (มีค่าเป็น 1) ส่วน d และ e เป็นค่าที่ไม่สนใจ (don't care)

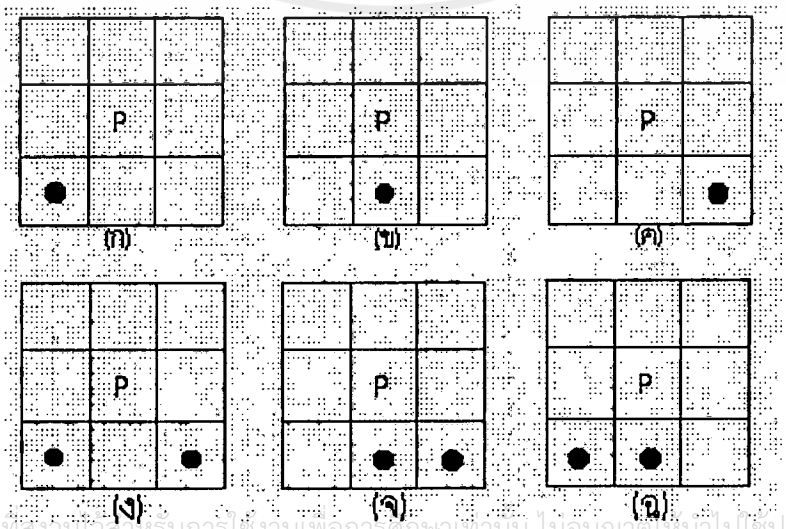
เมื่อพิจารณารูปที่ 4.20 (ก) , (ข) , (ค) จะแบ่งการพิจารณาเป็น 2 ลักษณะ คือ

1. ถ้า d ทุกตัวเป็นจุดสว่าง (มีค่าเป็น 0) จะได้ว่า P คือ จุดภาพที่เป็นจุดปลายของโครงร่าง

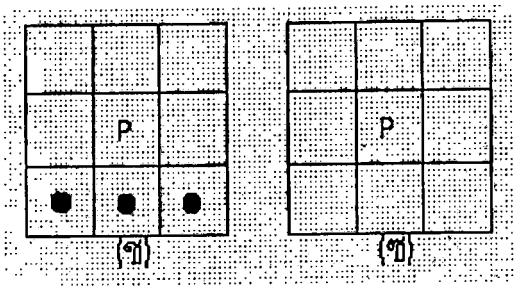
2. ถ้า d อย่างน้อย 1 จุดเป็นจุดมืด (มีค่าเป็น 1) จะได้ว่า P คือ จุดภาพที่เป็นโครงร่างของตัวอักษร

จุด P ใดๆ ที่มีคุณสมบัติตาม 2 ลักษณะดังกล่าว จะถูกกำหนดค่าให้เป็น 1 (unflaged)

ตารางหน้าต่างในรูปที่ 4.20(ง) จะมีการพิจารณาที่ซับซ้อนขึ้นโดยแยกการพิจารณาออกมาดังแสดงโดยตารางหน้าต่างในรูปที่ 4.21



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.21 แสดงหน้าต่างของจุดภาพโครงร่าง

จากรูปที่ 4.20(ง) ถ้า d และ e อย่างน้อย 1 จุด เป็นจุดมืดจะกำหนดให้จุด P เป็นจุดโครงร่าง ในกรณีนอกเหนือจากนี้จะพิจารณาตามตารางหน้าต่างในรูปที่ 4.21

จากรูปที่ 4.21 ตารางหน้าต่าง (ก) , (ข) , (ค) จะกำหนดให้จุด P เป็นจุดปลายของโครงร่าง ตารางหน้าต่าง (ง) กำหนดให้ P เป็นโครงร่างตารางหน้าต่าง (จ) และ (ฉ) กำหนดให้ P เป็นจุดที่ทำให้เกิดความกว้างขนาด 2 จุดภาพในแนวเอียง ตารางหน้าต่าง (ช) กำหนดให้จุด P เป็นกิ่งสั้น (protrusion) ของโครงร่าง ส่วนตารางหน้าต่าง (ซ) กำหนดให้จุด P เป็นจุดอิสระที่เกิดขึ้น

เพื่อลดความยุ่งยากในการพิจารณาเงื่อนไขของขอบและโครงร่างลง จึงได้นำเอาเงื่อนไขทั้งหมดของขอบมาเขียนเป็นนิพจน์ทางตรรกะ ทำให้ง่ายต่อการเขียน program ยิ่งขึ้น ดังแสดงในสมการที่ (1), (2), (3) และ (4) ตามลำดับ

ขอบขวา : $B0 = N4 \times (N2+N3+N5+N6) \times (N6+N7) \times (N1+N2)$ (1)

ขอบขวา : $B6 = N2 \times (N0+N1+N3+N4) \times (N4+N5) \times (N0+N7)$ (2)

ขอบขวา : $B4 = N0 \times (N1+N2+N6+N7) \times (N2+N3) \times (N5+N6)$ (3)

ขอบขวา : $B2 = N8 \times (N0+N4+N5+N7) \times (N0+N1) \times (N3+N4)$ (4)

เมื่อ x : operator AND

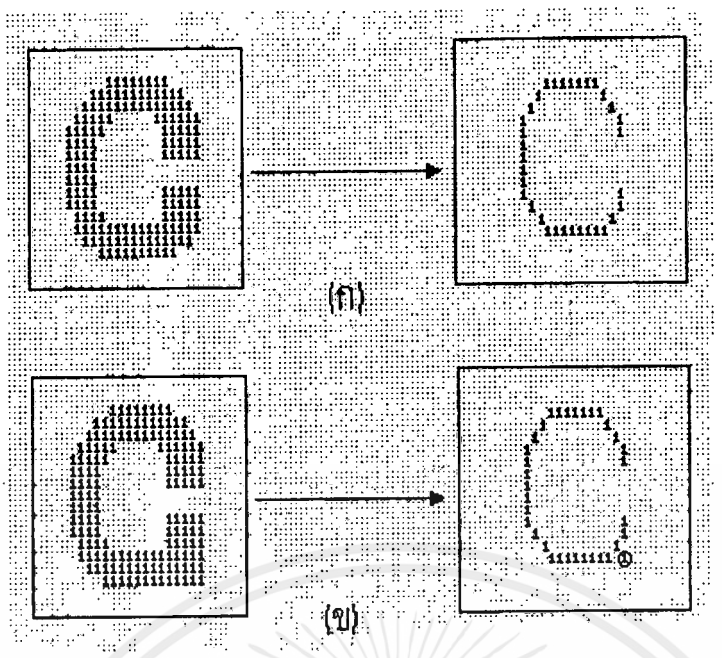
\+ : operator OR

\- : operator NOT

และ B4,B0,B2,B6 เป็นค่าของจุดภาพ P ที่มีลักษณะเป็นขอบซ้าย ขอบขวา ขอบบน และขอบล่างตามลำดับโดยการ scan ไปบนรูปแบบ 2 ระดับที่ต้องการทำให้บาง ทำการสแกนใน 2 ทิศทาง คือ scan จากบนลงล่าง และ scan จากซ้ายไปขวา ในทิศทางแรกจะพิจารณาขอบ B4 และ B0 และในทิศทางที่ 2 จะพิจารณาขอบ B2 และ B6 ตามลำดับ

ด้วยการกระทำแบบวนรอบ (Iteration) จนกระทั่งเหลือแต่จุดภาพที่เป็นโครงร่างก็จะได้ตัวอักษรที่บางออกมาดังแสดงในรูปที่ 4.22

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.22 โครงร่างของตัวอักษร
(ก) โครงร่างที่สมบูรณ์ (ข) โครงร่างที่มีส่วนเกิน

โครงร่างของตัวอักษรที่ได้จากการทำให้บางเก็บรูปแบบที่จะนำไปหาค่าคุณสมบัติทาง Topology ต่อไป

ข้อดีของการทำให้บาง คือ ลดจำนวนข้อมูลที่เป็นภาพตัวอักษรลงไปเป็นจำนวนมาก ทำให้ลดเวลาของการประมวนผลในขั้นตอนต่อไปได้ แต่ทั้งนี้การทำให้บางก็จะใช้เวลาในการประมวลผลนาน ถ้ารูปแบบ 2 ระดับที่เข้ามามีความหนามากๆ

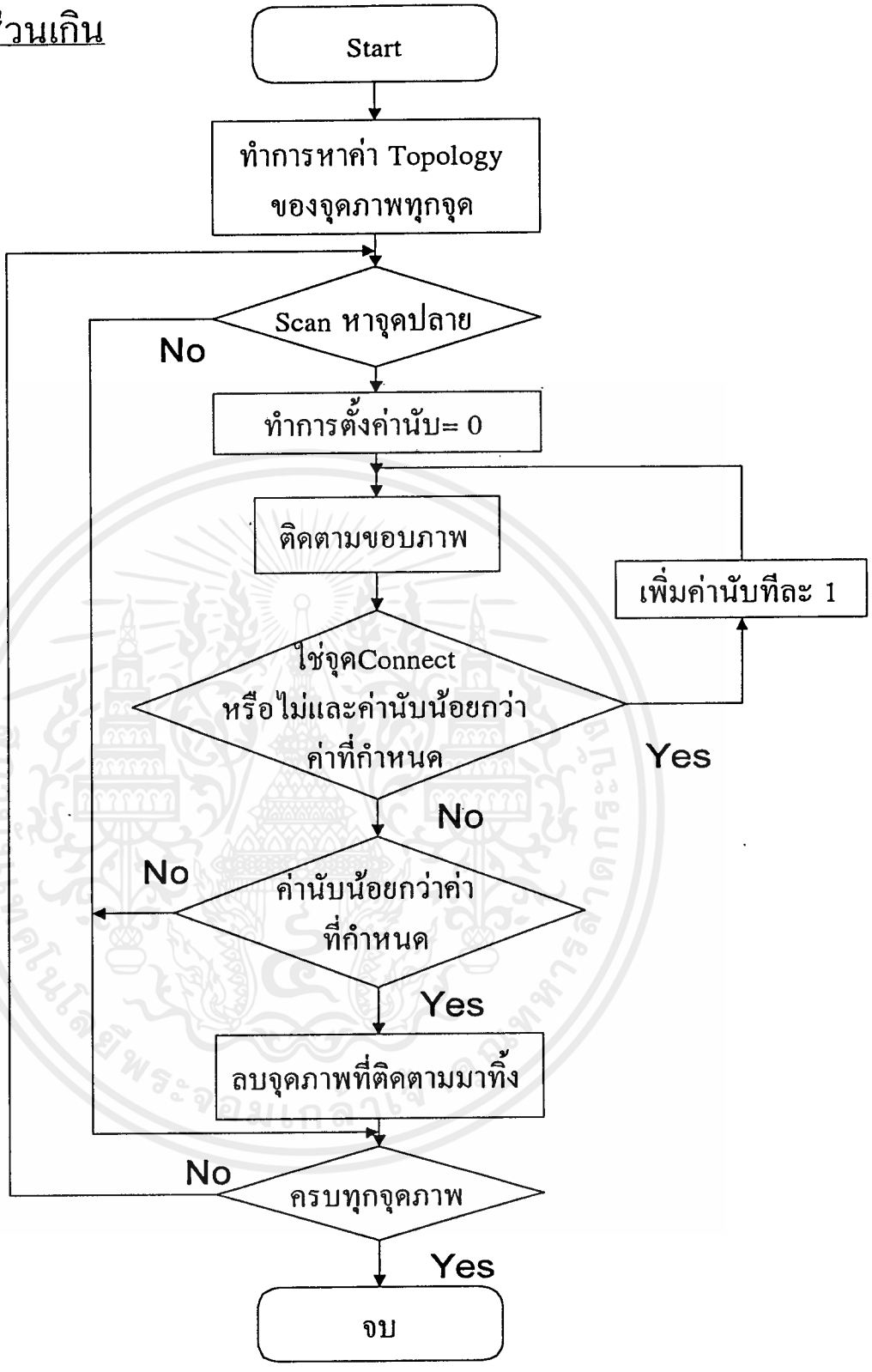
4.7 การกำจัดส่วนเกิน

ภาพตัวอักษรที่ผ่านการทำให้บางแล้วจะเหลือเฉพาะโครงร่าง ซึ่งรวมทั้งส่วนเกินที่ไม่ต้องการด้วย เราต้องทำการตัดเส้นกิ่งนั้นทิ้ง เส้นกิ่งสั้นๆ ที่ไม่ใช่ส่วนของภาพอักษรที่อาจเกิดเพราะวิธีการถ่ายภาพหรืออื่นๆ โดยจะ scan หาค่าโทโปโลยีของจุดภาพที่เป็นจุดปลายใดๆ แล้วติดตามเส้นนี้จนถึงจุดที่เป็นจุดแยกหรือจุดตัด ซึ่งจะมีการนับจำนวนจุดที่ติดตาม ถ้ามีค่าเกินที่กำหนดค่าหนึ่ง (หาความยาวของเส้นกิ่ง) เพราะเส้นกิ่งที่เป็นข้อผิดพลาดนี้ จะเป็นเส้นกิ่งที่มีลักษณะยาวไม่มาก หลังจากนั้นจึงลบจุดภาพที่ติดตามมาทิ้ง

ดังนั้นจึงมีความรวดเร็วและประหยัดหน่วยความจำอีกด้วย โดยสามารถเขียนอัลกอริธึมการกำจัดส่วนเกินดังรูปที่ 4.23 ได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การกำจัดส่วนเกิน



รูปที่ 4.23 ขั้นตอนการทำงานของอัลกอริธึมกำจัดส่วนเกิน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

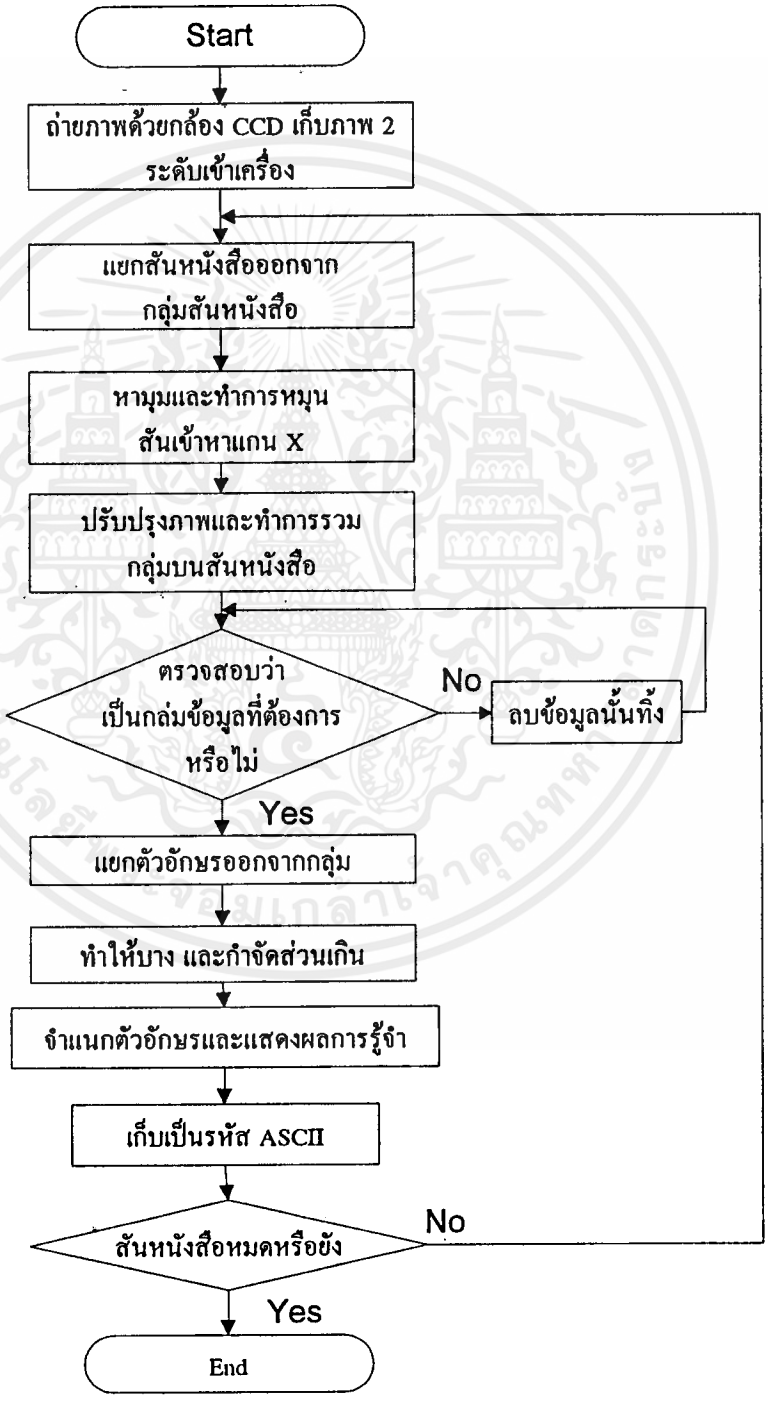
เมื่อได้ทำการทำให้บางและกำจัดส่วนเกินแล้วจะเห็นว่าข้อมูลภาพอักษรที่ได้จะมีความสมบูรณ์พอที่จะนำไปวิเคราะห์เพื่อแยกตัวอักษรต่อไปได้ ซึ่งสามารถเปรียบเทียบภาพก่อนทำให้บาง ก่อนกำจัดส่วนเกิน และหลังกำจัดส่วนเกิน ตามลำดับ โดยข้อมูลในขั้นตอนสุดท้ายนี้ก็พร้อมที่จะนำไปวิเคราะห์หารหัสแทนตัวอักษร (ASCII CODE) ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

การวิเคราะห์ภาพตัวอักษร

การวิเคราะห์ภาพตัวอักษร เป็นกระบวนการสุดท้าย ที่ใช้ในการจดจำตัวอักษร (Recognition Character) กระบวนการที่ผ่านมาสามารถเขียนอัลกอริทึมได้ดังรูปที่ 5.1 และจะกล่าวถึงขั้นตอน การวิเคราะห์ภาพตัวอักษรหลังจากได้มีการปรับปรุงภาพเพื่อการวิเคราะห์เรียบร้อยแล้วมีขั้นตอน ดังต่อไปนี้



รูปที่ 5.1 แสดงอัลกอริทึมการทำงานของกระบวนการรู้จำตัวอักษร

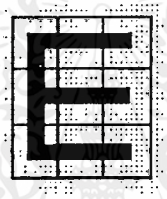
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนการวิเคราะห์ตัวอักษร

- การแบ่งภาพตัวอักษรเพื่อการหารหัสแทนตัวอักษร
- การกำหนดค่าบิตในแต่ละไบท์
- การกำหนดรหัสแทนตัวอักษร
- การแยกความแตกต่างของตัวอักษร

5.1 การแบ่งภาพเพื่อการหารหัสแทนตัวอักษร

การแบ่งภาพอักษรออกเป็น 9 ส่วน ในแต่ละส่วนจะมีการตรวจเช็คหาค่าคุณสมบัติทางโทโปโลยี เพื่อที่จะกำหนดเป็นค่าบิตใน 1 ไบต์รวมทั้งหมด 9 ไบต์



รูปที่ 5.2 แสดงการแบ่งภาพออกเป็น 9 ส่วน

5.2 การกำหนดค่าบิตในแต่ละไบต์

ในแต่ละบิตของแต่ละไบต์ถูกกำหนด โดยค่าคุณสมบัติทางโทโปโลยีที่เป็นจุดต่อ, จุดปลาย, จุดตัด, จุดแยก และจุดโดดเดี่ยว ดังนี้

บิตที่ 0 เป็นค่าจุดปลาย

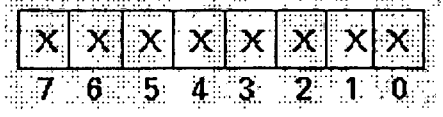
บิตที่ 1 เป็นค่าจุดต่อ

บิตที่ 2 เป็นค่าจุดแยก

บิตที่ 3 เป็นค่าจุดตัด

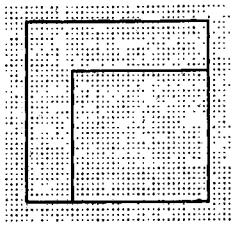
บิตที่ 4 เป็นค่าจุดโดดเดี่ยว

บิตที่ 5-7 จะเก็บค่า 0

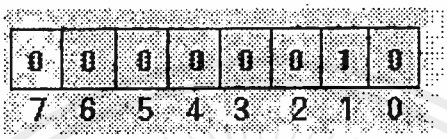


รูปที่ 5.3 แสดงการกำหนดรหัสในแต่ละบิต (x=0,1)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

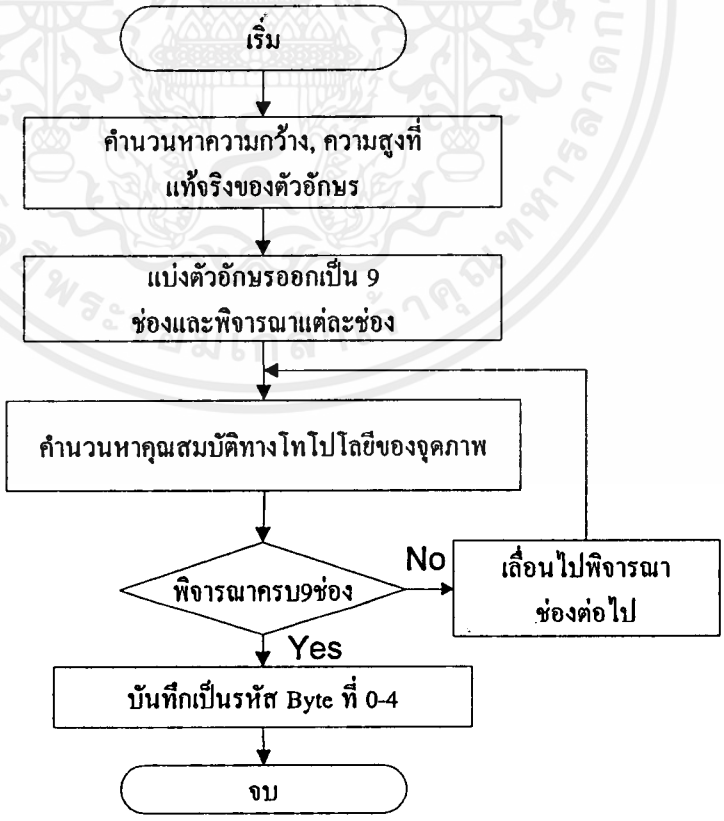


รูปที่ 5.4 แสดงภาพลายเส้นของช่องที่ 1 ในรูปตัว E



รูปที่ 5.5 แสดงการกำหนดรหัสแทนช่องที่ 1 ของรูปตัว E ขนาด 1 ไบต์

จากขั้นตอนต่างๆ ที่ได้กล่าวมาเริ่มต้นตั้งแต่การแบ่งภาพเพื่อหารหัสแทนตัวอักษรจนถึงการกำหนดค่าบิตในแต่ละไบต์สามารถแสดงอัลกอริทึมของขั้นตอนต่างๆ ได้ดังนี้

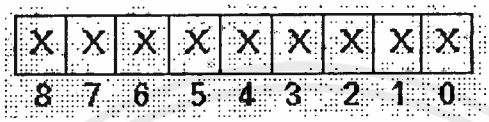


รูปที่ 5.6 แสดงอัลกอริทึมการวิเคราะห์หารหัสแทนตัวอักษร 9 ไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3 การกำหนดรหัสแทนตัวอักษร

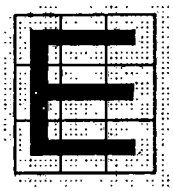
จากกระบวนการต่างๆ ที่ได้กล่าวมาในหัวข้อข้างต้นนี้ เมื่อนำมารวมเป็นรหัสที่ใช้แทนตัวอักษรแต่ละตัวจะใช้หน่วยความจำ 9 ไบต์ ดังแสดงในรูป โดยในไบต์ที่ 0-8 เป็นรหัส ที่ได้จากการหาคุณสมบัติทางโทโปโลยีของจุดภาพ



รูปที่ 5.7 แสดงการจัดวางตำแหน่งรหัสแทนตัวอักษรทั้ง 9 ไบต์

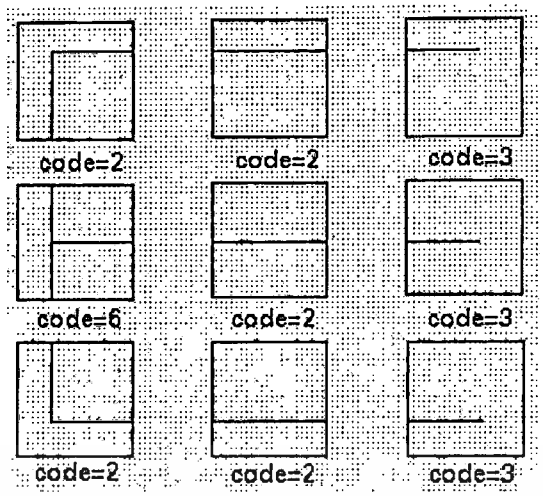
เมื่อไบต์ที่ 0 แทนรหัสที่ได้จากการหารหัสในช่องที่ 0
1 แทนรหัสที่ได้จากการหารหัสในช่องที่ 1
2 แทนรหัสที่ได้จากการหารหัสในช่องที่ 2
3 แทนรหัสที่ได้จากการหารหัสในช่องที่ 3
4 แทนรหัสที่ได้จากการหารหัสในช่องที่ 4
5 แทนรหัสที่ได้จากการหารหัสในช่องที่ 5
6 แทนรหัสที่ได้จากการหารหัสในช่องที่ 6
7 แทนรหัสที่ได้จากการหารหัสในช่องที่ 7
8 แทนรหัสที่ได้จากการหารหัสในช่องที่ 8

ตัวอย่างการหารหัสตัวอักษร



รูปที่ 5.8 แสดงตัวอักษรตัว E

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.9 แสดงการแทนรหัสที่ได้จากการหาค่าในแต่ละช่องของรูปตัว E

จากรูปที่ 5.8 อักษรได้ถูกแบ่งออกเป็น 9 ส่วนจะถูกตรวจกวาดหาค่าคุณสมบัติทางโทโปโลยีแล้วนำมากำหนดรหัสแทนตัวอักษรจำนวน 9 ไบต์ ตั้งแต่ไบต์ที่ 0-8 ดังรูปที่ 5.9 เมื่อนำรหัสมาเรียงกันก็จะได้รหัสแทนตัวอักษรที่สมบูรณ์ดังแสดงในรูปที่ 5.10

3	2	2	3	2	6	3	2	2
8	7	6	5	4	3	2	1	0

รูปที่ 5.10 แสดงรหัสแทนตัวอักษรที่สมบูรณ์ของตัวอักษร E

5.4 การจัดเก็บฐานข้อมูลการรู้จำ

ลักษณะการจัดเก็บฐานข้อมูลจะทำการจัดเก็บ 12 ไบต์ต่อ 1 ตัวอักษรโดย 9 ไบต์แรกนั้นจะเก็บรหัสที่ทำการวิเคราะห์ตามวิธีทางโทโปโลยีของจุดภาพส่วนไบต์ที่ 10-11 จะทำการเก็บค่าสภาวะคุณสมบัติประจำตัวของตัวอักษร และไบต์ที่ 12 จะเก็บรหัสแทนตัวอักษร (ASCII Code) และรหัสตัวอักษรตัวต่อไปก็จะเก็บเรียงต่อเนื่องไปเรื่อยๆ ดังแสดงในรูปที่ 5.11

QCODE	Characteristic code	ASCII	QCODE	Characteristic code	ASCII

รูปที่ 5.11 แสดงการเก็บฐานข้อมูลการรู้จำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ QCODE คือ รหัสของตัวอักษรแต่ละตัวที่วิเคราะห์ได้มีขนาด 9 ไบต์

Characteristic code คือ รหัสคุณสมบัติประจำตัวของตัวอักษรมีขนาด 2 ไบต์

ASCII คือ รหัสแทนตัวอักษร (ASCII Code) ของรหัส code ที่วิเคราะห์ได้มีขนาด 1 ไบต์

เริ่มแรกโปรแกรมยังไม่สามารถที่จะทำการรู้จำตัวอักษรได้ และจะรู้จำตัวอักษรได้ก็ต่อเมื่อได้เริ่มทำการวิเคราะห์ เมื่อวิเคราะห์ได้ก็จะได้รหัส (QCODE) ออกมา 9 ไบต์จากนั้นเราต้องกด key ตัดสินใจว่าภาพตัวอักษรตัวนี้เป็นตัวอักษรอะไร (เนื่องจากภาพตัวอักษรที่โปรแกรมยังไม่รู้จำจะแสดงปรากฏบนจอให้ผู้ใช้ตัดสินใจเลือกเอง) เมื่อเราป้อนค่าเข้าไปโปรแกรมจะเก็บรหัสที่วิเคราะห์ได้ 9 ไบต์เก็บจะเป็นฐานข้อมูลก่อน จากนั้นจะเก็บรหัสคุณสมบัติประจำตัวของตัวอักษรที่เราป้อนเข้าไปเป็นไบต์ที่ 10 และสุดท้ายก็จะทำการเก็บรหัสแทนตัวอักษรที่เราป้อนเข้าไปเป็นไบต์ที่ 12 เมื่อทำการวิเคราะห์ตัวอักษรต่อไปโปรแกรมจะอ่านรหัส QCODE 9 ไบต์ที่อยู่ในฐานข้อมูลขึ้นมาเปรียบเทียบกับรหัสที่วิเคราะห์ได้ ถ้าเท่ากันจะทำการอ่านไบต์ที่ 10-12 ซึ่งเป็นรหัสคุณสมบัติประจำตัว และรหัสแทนตัวอักษรของตัวอักษรออกมา แต่ถ้าค่ารหัสที่วิเคราะห์ได้ไม่เท่ากับในฐานข้อมูล โปรแกรมจะแสดงภาพตัวอักษรตัวที่ไม่สามารถวิเคราะห์ได้ขึ้นมาให้ผู้ใช้ป้อนรหัสคุณสมบัติประจำตัวอักษรและรหัสแทนตัวอักษรทางคีย์บอร์ด จากนั้นจะทำการเก็บรหัสที่วิเคราะห์กับรหัสคุณสมบัติประจำตัวอักษรและรหัสแทนตัวอักษรลงเป็นฐานข้อมูลเพิ่มเติมขึ้น เพิ่มการรู้จำที่มากขึ้นต่อไป

ตัวอักษร	รหัสที่ทำการวิเคราะห์ได้ (QCODE)									รหัสคุณสมบัติประจำตัว	
	1	2	3	4	5	6	7	8	9	10	11
A	2	2	2	6	2	6	3	0	3	1	0
B	2	2	2	6	6	2	2	2	2	1	0
C	2	2	3	2	0	0	2	2	3	0	1
D	2	2	2	2	0	2	2	2	2	1	0
E	2	2	3	6	2	3	2	2	3	1	0
F	2	2	3	6	2	3	3	0	0	1	0
G	2	2	3	2	3	2	2	2	2	1	0
H	3	0	3	6	2	6	3	0	3	1	0
I	0	3	0	0	2	0	0	3	0	0	3
J	0	0	3	0	0	2	3	2	2	1	0
K	3	2	3	6	2	0	3	2	3	1	0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

L	3	0	0	2	0	0	2	2	3	1	0
M	2	0	2	2	2	2	3	2	3	1	0
N	2	2	3	2	2	2	3	2	2	1	0
O	2	2	2	2	0	2	2	2	2	0	1
P	2	2	2	6	2	2	3	0	0	0	2
Q	2	2	2	2	0	2	2	2	7	1	0
R	2	2	2	6	6	2	3	0	3	1	0
S	2	2	3	2	2	2	3	2	2	0	1
T	3	6	3	0	2	0	0	3	0	1	0
U	3	0	3	2	0	2	2	2	2	0	1
V	3	0	3	2	0	2	2	2	2	0	1
W	3	2	3	2	2	2	2	0	2	0	1
X	3	2	3	0	10	0	3	2	3	0	1
Y	3	0	3	2	6	2	0	3	0	1	0
Z	3	2	2	0	2	2	2	2	3	0	1

ตารางที่ 5.1 แสดงตารางค่ารหัสของตัวอักษรพิมพ์ใหญ่ภาษาอังกฤษต้นแบบ CordiaUPC

ตัวอักษร	รหัสที่ทำการวิเคราะห์ได้ (QCODE)									รหัสคุณสมบัติประจำตัว	
	1	2	3	4	5	6	7	8	9	10	11
a	3	2	2	2	2	6	2	2	7	2	0
b	3	2	2	6	0	2	7	2	2	2	0
c	2	2	3	2	0	0	2	2	3	0	1
d	2	2	3	2	0	6	2	2	7	2	0
e	2	2	2	6	2	2	2	2	3	2	0
f	0	2	3	3	10	3	0	3	0	2	0
g	2	2	7	2	0	6	3	2	2	0	2
h	3	0	0	6	2	2	3	0	3	2	0
i	0	19	0	0	2	0	0	3	0	2	0
j	0	0	19	0	0	2	3	2	2	2	0
k	3	0	3	6	2	2	3	2	3	2	0
l	0	3	0	0	2	20	0	3	0	0	3
m	7	6	2	2	2	2	3	3	3	2	0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

n	7	2	2	2	0	2	3	0	3	2	0
o	2	2	2	2	0	2	2	2	2	0	1
p	7	2	2	6	0	2	3	2	2	0	2
q	2	2	7	2	0	6	2	2	3	2	0
r	7	2	3	2	0	0	3	0	0	2	0
s	2	2	3	2	2	2	3	2	2	0	1
t	3	11	3	0	2	0	0	2	3	2	0
u	3	0	3	2	0	2	2	2	7	0	1
v	3	0	3	2	0	2	2	2	2	0	1
w	3	2	3	2	2	2	2	0	2	0	1
x	3	2	3	0	10	0	3	2	3	0	1
y	3	0	3	2	6	2	3	2	0	2	0
z	3	2	2	0	2	2	2	2	3	0	1

ตารางที่ 5.2 แสดงตารางค่ารหัสของตัวอักษรพิมพ์เล็กภาษาอังกฤษต้นแบบ CordiaUPC

ตัวอักษร	รหัสที่ทำการวิเคราะห์ได้ (QCODE)									รหัสคุณสมบัติประจำตัว	
	1	2	3	4	5	6	7	8	9	10	11
0	2	2	2	2	0	2	2	2	2	1	0
1	3	2	2	0	0	2	0	0	3	1	0
2	3	2	2	0	2	2	2	2	3	1	0
3	3	2	2	0	3	6	3	2	2	1	0
4	2	2	2	2	0	2	2	2	7	1	0
5	2	2	3	2	2	2	3	2	2	1	0
6	2	2	3	6	2	2	2	2	2	1	0
7	3	2	2	0	2	2	0	3	0	1	0
8	2	2	2	6	6	2	2	2	2	1	0
9	2	2	2	2	2	6	3	2	2	0	2

ตารางที่ 5.3 แสดงตารางค่ารหัสของเลขพิมพ์ภาษาอังกฤษต้นแบบ CordiaUPC

<u>หมายเหตุ</u>

รหัสอักษรที่วิเคราะห์ได้พิจารณาจากตัวอักษรพิมพ์ที่ได้มาตรฐาน Font CordiaUPC

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.5 การแยกความแตกต่างของตัวอักษร

จากการวิเคราะห์จะพบว่ารหัสของตัวอักษรนั้นจะมีอักษรบางตัวที่มีรหัสเหมือนกันทุกไบต์เมื่อเป็นดังนี้ เราพิจารณาส่วนการแยกความแตกต่างของตัวอักษรออกเป็น 2 รูปแบบดังนี้

1. ภาพตัวอักษรที่รหัสตัวอักษรเหมือนกันแต่ภาพตัวอักษรแตกต่างกัน
2. ภาพตัวอักษรที่รหัสตัวอักษรเหมือนกันและภาพตัวอักษรก็เหมือนกันด้วย

5.5.1) ภาพตัวอักษรที่รหัสเหมือนกัน แต่ภาพตัวอักษรแตกต่างกัน

จากการพิจารณาภาพตัวอักษรที่มีลักษณะรหัสเหมือนกันทุกไบต์ แต่ภาพตัวอักษรแตกต่างกัน เพื่อให้การวิเคราะห์ได้ผลที่ถูกต้องจึงต้องมีการพิจารณา เฉพาะส่วนของตัวอักษรที่มีชุดรหัสเหมือนกันดังนี้

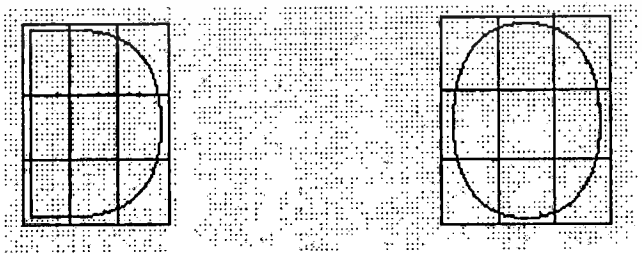
0	1	2
3	4	5
6	7	8

รูปที่ 5.12 แสดงตารางการแบ่งส่วนของตัวอักษร

5.5.1.1 กรณีชุดรหัสเหมือนกันของตัวอักษร 0, O และ D

เราจะแยกความแตกต่างของชุดรหัสอักษรเหมือนกัน พิจารณาออกเป็น 2 ขั้นตอนดังนี้

1. แยกพิจารณาภาพตัวอักษร D ออกจากตัวอักษร 0 และ O

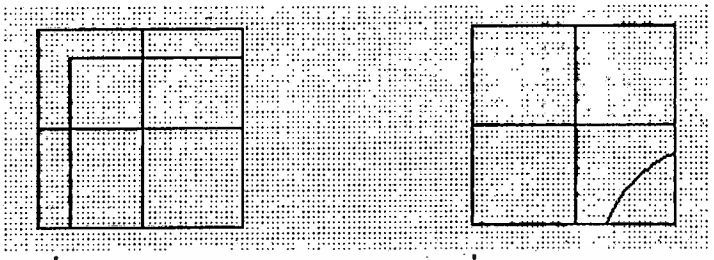


a)ภาพตัวอักษร D b)ภาพตัวอักษร O และ0

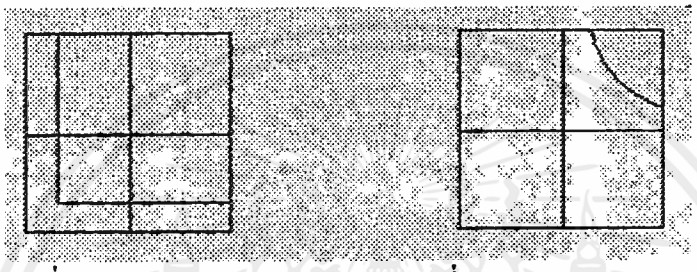
รูปที่ 5.13 แสดงการแบ่งช่องของตัวอักษร D และ 0,O

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 5.13 สามารถแยกความแตกต่างได้โดยพิจารณาส่วนที่ 0 และส่วนที่ 6 ของภาพตัวอักษร โดยพิจารณาแยกส่วนของตัวอักษรออกเป็น 4 ส่วนดังนี้



ส่วนที่ 0 ของตัวอักษร D ส่วนที่ 0 ของตัวอักษร 0 หรือ O

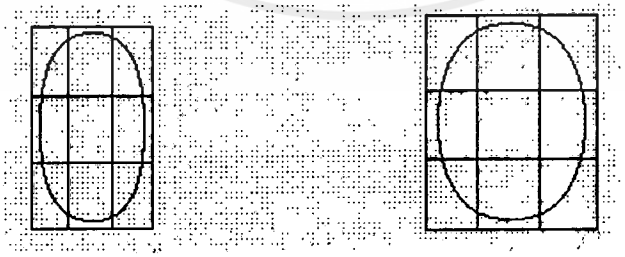


ส่วนที่ 6 ของตัวอักษร D ส่วนที่ 6 ของตัวอักษร 0 หรือ O

รูปที่ 5.14 แสดงการแบ่งช่องพิจารณาของตัวอักษร

จากรูปที่ 5.14 จะพิจารณาได้ว่าในส่วนที่ 0 และส่วนที่ 2 ของตัวอักษร D จะมีข้อมูลภาพอยู่แต่ตัวอักษร 0 หรือ O จะไม่มีข้อมูลภาพอยู่ เราจึงสามารถแยกความแตกต่างของข้อมูลภาพระหว่างตัว D กับตัว 0 และ O

2. แยกพิจารณาภาพตัวอักษร 0 กับ O



ภาพตัวอักษร 0 ภาพตัวอักษร O

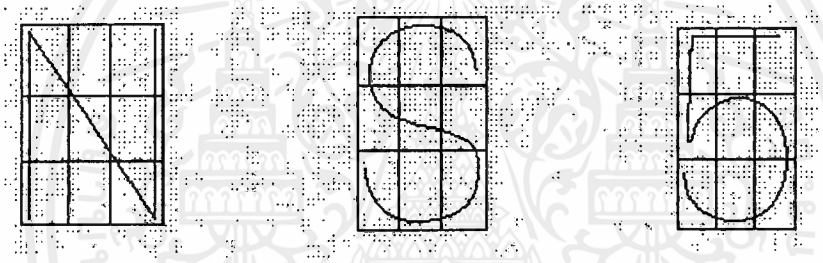
รูปที่ 5.15 แสดงการแบ่งช่องในการพิจารณาตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 5.15 เมื่อพิจารณารูปร่างของตัวอักษรเราแทบจะไม่สามารถแยกความแตกต่างระหว่างตัวอักษรระหว่าง 0 กับ O ได้เลย แต่เมื่อพิจารณาถึงอัตราส่วนความกว้างในแนวแกน X ต่อแกน Y ของภาพตัวอักษร เราสามารถแยกความแตกต่างของภาพตัวอักษรได้โดยที่ค่าอัตราส่วนของภาพอักษร 0 จะมีค่าน้อยกว่าภาพอักษร O ซึ่งอัตราส่วนของภาพอักษร O จะมีค่าประมาณ 0.75 และค่าอัตราส่วนของอักษร 0 จะประมาณ 0.55 ซึ่งมีค่าผลต่างประมาณ 0.20 เราจึงกำหนดค่า Threshold ค่าๆ หนึ่งประมาณ 0.65 เป็นค่าตัดสินใจภาพตัวอักษรว่าเป็นตัว 0 หรือ O โดยถ้าค่าอัตราส่วนมากกว่า 0.65 จะพิจารณาภาพตัวอักษรเป็น O ถ้าอัตราส่วนน้อยกว่า 0.65 จะพิจารณาภาพตัวอักษรเป็น 0

2. กรณีชุดรหัสเหมือนกันของตัวอักษร N,S และ 5 เราจะแยกความแตกต่างของชุดรหัสอักษรที่เหมือนกัน พิจารณาออกเป็น 2 ขั้นตอนดังนี้

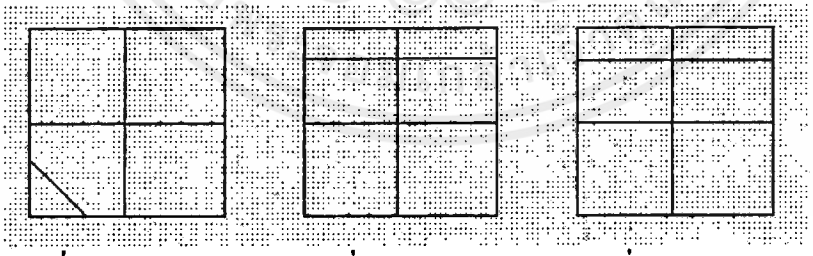
2.1 แยกพิจารณาภาพตัวอักษร N ออกจากภาพตัวอักษร S และ 5



ภาพตัวอักษร N ภาพตัวอักษร S ภาพตัวอักษร 5

รูปที่ 5.16 แสดงภาพการแบ่งช่องในการพิจารณาตัวอักษร

จากรูปที่ 5.16 สามารถแยกความแตกต่างได้โดยพิจารณาส่วนที่ 1 ของภาพตัวอักษรโดยพิจารณาแยกส่วนของตัวอักษรออกเป็น 4 ส่วน ดังนี้



ส่วนที่ 1 ของอักษร N ส่วนที่ 1 ของอักษร S ส่วนที่ 1 ของอักษร 5

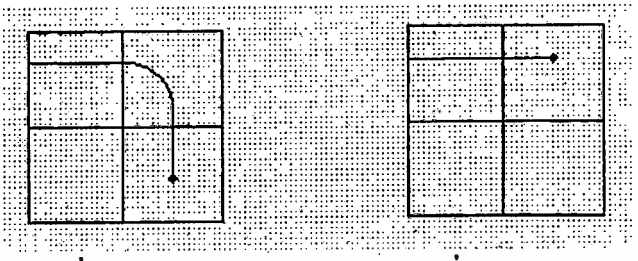
รูปที่ 5.17 แสดงภาพส่วนที่ 1 ของตัวอักษร N, S และ 5

จากรูปที่ 5.17 จะพิจารณาได้ว่าในส่วนที่ 1 ของตัวอักษร N จะไม่มีข้อมูลภาพอยู่แต่เมื่อพิจารณาเปรียบเทียบกับอักษร S และ 5 จะมีข้อมูลอยู่ เราจึงสามารถแยกแยะความแตกต่างของข้อมูลภาพระหว่างตัวอักษร N ออกจากตัวอักษร S และ 5 ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 แยกพิจารณาภาพตัวอักษร 5 ออกจากภาพตัวอักษร S

จากรูปที่ 5.17 สามารถแยกความแตกต่างได้โดยพิจารณาในส่วนที่ 2 ของตัวอักษร S และ 5 โดยพิจารณาแยกส่วนของตัวอักษรออกเป็น 4 ส่วนย่อยๆ ดังนี้

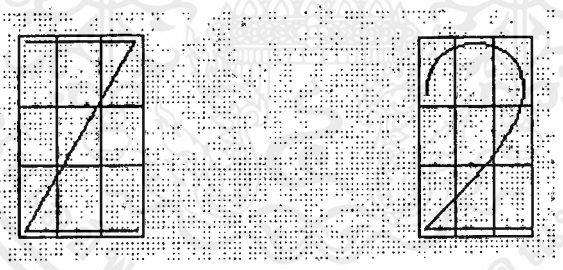


ส่วนที่ 2 ของตัวอักษร S ส่วนที่ 2 ของตัวอักษร 5

รูปที่ 5.18 แสดงภาพส่วนที่ 2 ของตัวอักษร S และ 5

จากรูปที่ 5.18 จะพิจารณาได้ว่าในส่วนที่ 1 ของตัวอักษร 5 จะมีจุดปลายอยู่แต่เมื่อพิจารณาเปรียบเทียบกับตัวอักษร S จะไม่มีจุดปลายอยู่ในส่วนนี้ แต่จะมีอยู่ในส่วนที่ 3 เราจึงสามารถแยกความแตกต่างของข้อมูลภาพระหว่างตัวอักษร S และ 5 ออกจากกันได้

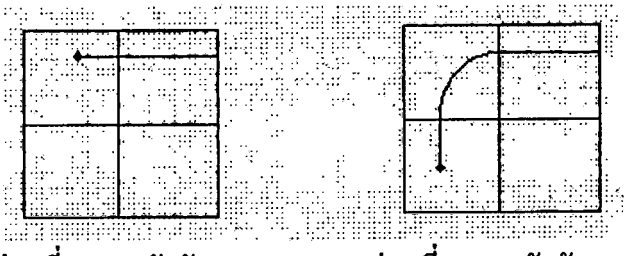
3. กรณีชุดรหัสเหมือนกันของตัวอักษร Z และ 2



ภาพตัวอักษร Z ภาพตัวอักษร 2

รูปที่ 5.19 แสดงภาพการแบ่งภาพตัวอักษรเพื่อใช้ในการพิจารณา

จากรูปที่ 5.19 สามารถแยกความแตกต่างได้โดยพิจารณา ส่วนที่ 0 ของภาพตัวอักษร Z และ 2 โดยพิจารณาแยกส่วนของตัวอักษรออกเป็น 4 ส่วนดังนี้



ส่วนที่ 0 ของตัวอักษร Z ส่วนที่ 0 ของตัวอักษร 2

รูปที่ 5.20 แสดงภาพส่วนที่ 0 ของตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 5.20 จะพิจารณาได้ว่าในส่วนที่ 0 ของตัวอักษร Z จะมีจุดปลายอยู่ แต่เมื่อพิจารณาเปรียบเทียบกับตัวอักษร 2 จะไม่มีจุดปลายอยู่ในส่วนนี้ แต่จะมีจุดปลายอยู่ในส่วนที่ 2 เราจึงสามารถแยกความแตกต่างของข้อมูลภาพระหว่างตัวอักษร Z และ 2 ออกจากกันได้

4. กรณีชุดรหัสเหมือนกันของตัวอักษร U และ V

ภาพตัวอักษร U ภาพตัวอักษร V

รูปที่ 5.21 แสดงภาพการแบ่งตัวอักษรเพื่อใช้ในการพิจารณา

จากรูปที่ 5.21 สามารถแยกความแตกต่างได้โดยพิจารณาส่วนที่ 6 ของภาพตัวอักษร โดยพิจารณาแยกส่วนของตัวอักษรออกเป็น 4 ส่วนดังนี้

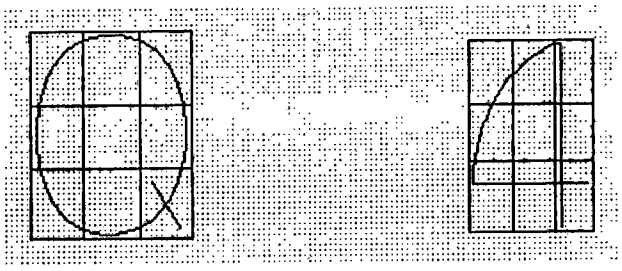
ส่วนที่ 6 ของตัวอักษร U ส่วนที่ 6 ของตัวอักษร V

รูปที่ 5.22 แสดงส่วนที่ 6 ของตัวอักษร

จากรูปที่ 5.22 จะพิจารณาได้ว่าในส่วนที่ 2 ของตัวอักษร U จะมีข้อมูลภาพอยู่ แต่ตัวอักษร V จะไม่มีข้อมูลภาพอยู่เราจึงสามารถแยกความแตกต่างของข้อมูลภาพระหว่างตัวอักษร U และ V ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

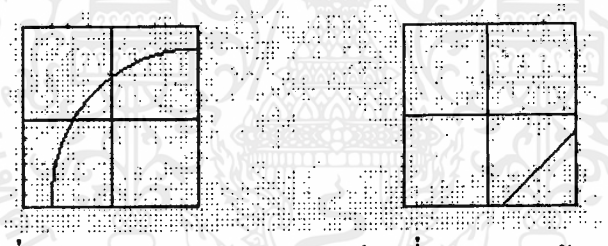
5. กรณีชุดรหัสเหมือนกันของตัวอักษร Q และ 4



ภาพตัวอักษร Q ภาพตัวอักษร 4

รูปที่ 5.23 แสดงภาพการแบ่งตัวอักษรเพื่อใช้ในการพิจารณา

จากรูปที่ 5.23 สามารถแยกความแตกต่างได้โดยพิจารณาส่วนที่ 0 ของภาพตัวอักษรโดยพิจารณาแยกส่วนของภาพตัวอักษรออกเป็น 4 ส่วนดังนี้



ส่วนที่ 0 ของภาพอักษร Q ส่วนที่ 0 ของภาพอักษร 4

รูปที่ 5.24 แสดงส่วนที่ 0 ของภาพอักษร

จากรูปที่ 5.24 จะพิจารณาได้ว่าในส่วนที่ 0,1,2 ของตัวอักษร Q จะมีข้อมูลภาพอยู่ แต่เมื่อพิจารณาเปรียบเทียบกับอักษร 4 จะไม่มีข้อมูลภาพอยู่เราจึงสามารถแยกความแตกต่างของข้อมูลภาพตัวอักษร Q และ 4 ออกจากกันได้

5.5.2) ภาพตัวอักษรที่รหัสตัวอักษรเหมือนกันและภาพตัวอักษรเหมือนกันด้วย

จากกระบวนการจำแนกตัวอักษรที่ผ่านมาข้างต้นยังไม่สามารถกำหนดตัวอักษรออกได้หมดทุกๆ ตัว ยังมีภาพตัวอักษรที่มีลักษณะเหมือนกัน เช่น ตัวอักษรที่มีลักษณะเหมือนกันทั้งตัวพิมพ์ใหญ่ และตัวพิมพ์เล็ก เมื่อพิจารณาภาพของตัวอักษรแล้วเราสามารถใช้วิธีที่ใกล้เคียงกับการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มองเห็นของสายตามนุษย์ ในการจำแนกความแตกต่างของตัวอักษรแต่ละตัว โดยวิธีพิจารณาตัวอักษรรอบๆ ตัวที่เรากำลังพิจารณาอยู่ โดยใช้ค่าสภาวะต่างๆ ของตัวอักษรรอบๆ ข้างที่พิจารณาเป็นตัวกำหนดความแตกต่างของตัวอักษรแต่ละตัว ซึ่งมีหลักการพิจารณาดังนี้

<u>หลักการที่ 1</u> คือ หลักการแยกความแตกต่างทางด้านความสูงของตัวอักษร สามารถแยกความแตกต่างของภาพตัวอักษรได้ ดังนี้ Cc; Oo, Ss, Uu, Vv, Ww, Xx, Zz

<u>หลักการที่ 2</u> คือ หลักการแยกความแตกต่างทางด้านตำแหน่งของตัวอักษร สามารถแยกภาพตัวอักษรได้ดังนี้ Pp,g9

<u>หลักการที่ 3</u> คือ หลักการแยกความแตกต่างทางด้านการจัดเรียงตัวอักษรของคำในภาษาอังกฤษ สามารถแยกภาพตัวอักษรได้ ดังนี้ คือ Il

<u>ค่าสภาวะต่างๆที่กำหนดแสดงดังนี้</u>

1. ค่าสภาวะคุณลักษณะของตัวอักษรพิมพ์ใหญ่, พิมพ์เล็ก หรือ ไม่สามารถกำหนดได้จากการวิเคราะห์ในขั้นตอนที่ผ่านมาว่าเป็นตัวพิมพ์ใหญ่หรือตัวพิมพ์เล็ก โดยเราได้กำหนดค่าเป็นตัวแปรดังนี้

flag-char[i]=0 หมายความว่า ภาพตัวอักษรยังไม่สามารถกำหนดได้จากการวิเคราะห์ในขั้นตอนที่ผ่านมาว่าเป็นตัวพิมพ์ใหญ่หรือตัวพิมพ์เล็ก

flag-char[i]=1 หมายความว่า ภาพตัวอักษรจากการวิเคราะห์ในขั้นตอนที่ผ่านมาเป็น ตัวพิมพ์ใหญ่

flag-char[i]=2 หมายความว่า ภาพตัวอักษรจากการวิเคราะห์ในขั้นตอนที่ผ่านมาเป็น ตัวพิมพ์เล็ก

2. ค่าสภาวะความสูงของตัวอักษร กำหนดเป็นตัวแปรดังนี้

flaghigh[i]=Valve ค่า Valve จะเป็นค่าที่แสดงถึงความกว้างของตัวอักษรในแนวแกน Y หรือความสูงของตัวอักษรแต่ละตัว

3. ค่าสภาวะตำแหน่งของตัวอักษรกำหนดเป็นตัวแปรดังนี้

flagunder[i]=Valve ค่า Valve จะเป็นค่าที่แสดงถึงตำแหน่งล่างสุดในแนวแกน Y ของตัวอักษรแต่ละตัว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. ค่าสภาวะของตัวอักษร ที่พิจารณาว่ามีคุณลักษณะตรงตามหลักการวิเคราะห์ความแตกต่างของภาพแบบใด กำหนดดังนี้

flagrecog[i]=0 หมายความว่า ภาพตัวอักษรสามารถวิเคราะห์ได้จากขั้นตอนเบื้องต้นของการเปรียบเทียบกับฐานข้อมูล

flagrecopy[i]=1 หมายความว่า ภาพตัวอักษรที่จะนำมาวิเคราะห์ตามหลักการที่ 1

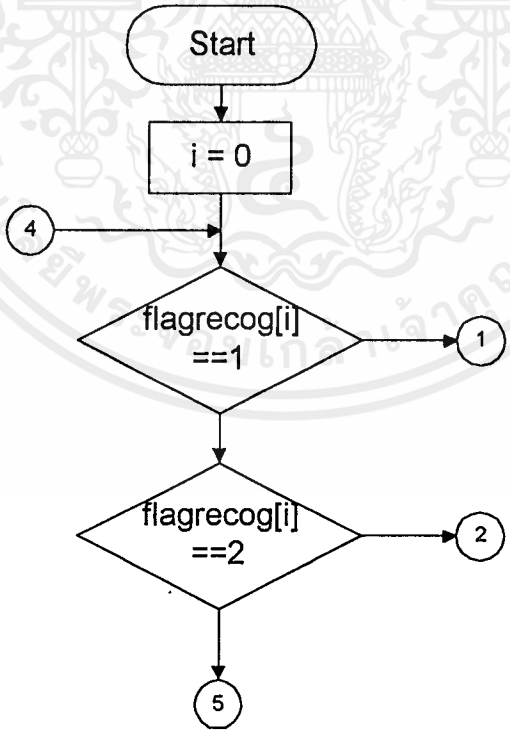
flagrecopy[i]=2 หมายความว่า ภาพตัวอักษรที่จะนำมาวิเคราะห์ตามหลักการที่ 2

flagrecopy[i]=3 หมายความว่า ภาพตัวอักษรที่จะนำมาวิเคราะห์ตามหลักการที่ 3

เราสามารถแบ่งขั้นตอนการแยกความแตกต่างของตัวอักษรได้ 2 ขั้นตอนดังนี้

<u>ขั้นตอนที่ 1</u> การพิจารณาแยกความแตกต่างของตัวอักษรโดยใช้หลักการที่ 1 และ 2 แสดงดัง flow chart ดังนี้

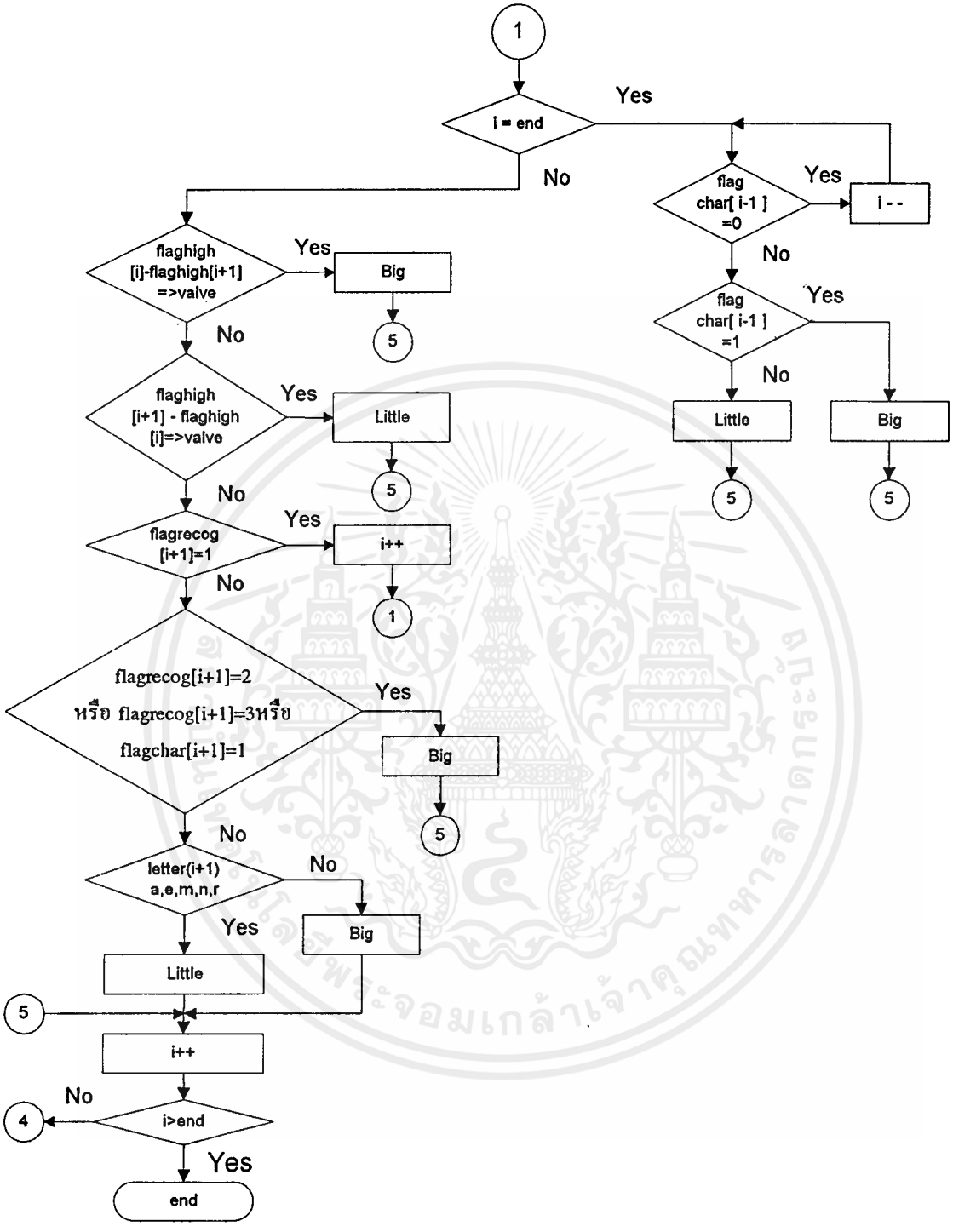
<u>อัลกอริทึมของการแยกแยะความแตกต่างของตัวอักษรขั้นที่ 1</u>



รูปที่ 5.25 แสดง flow chart ในการแยกแยะความแตกต่างตัวอักษรขั้นตอนที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

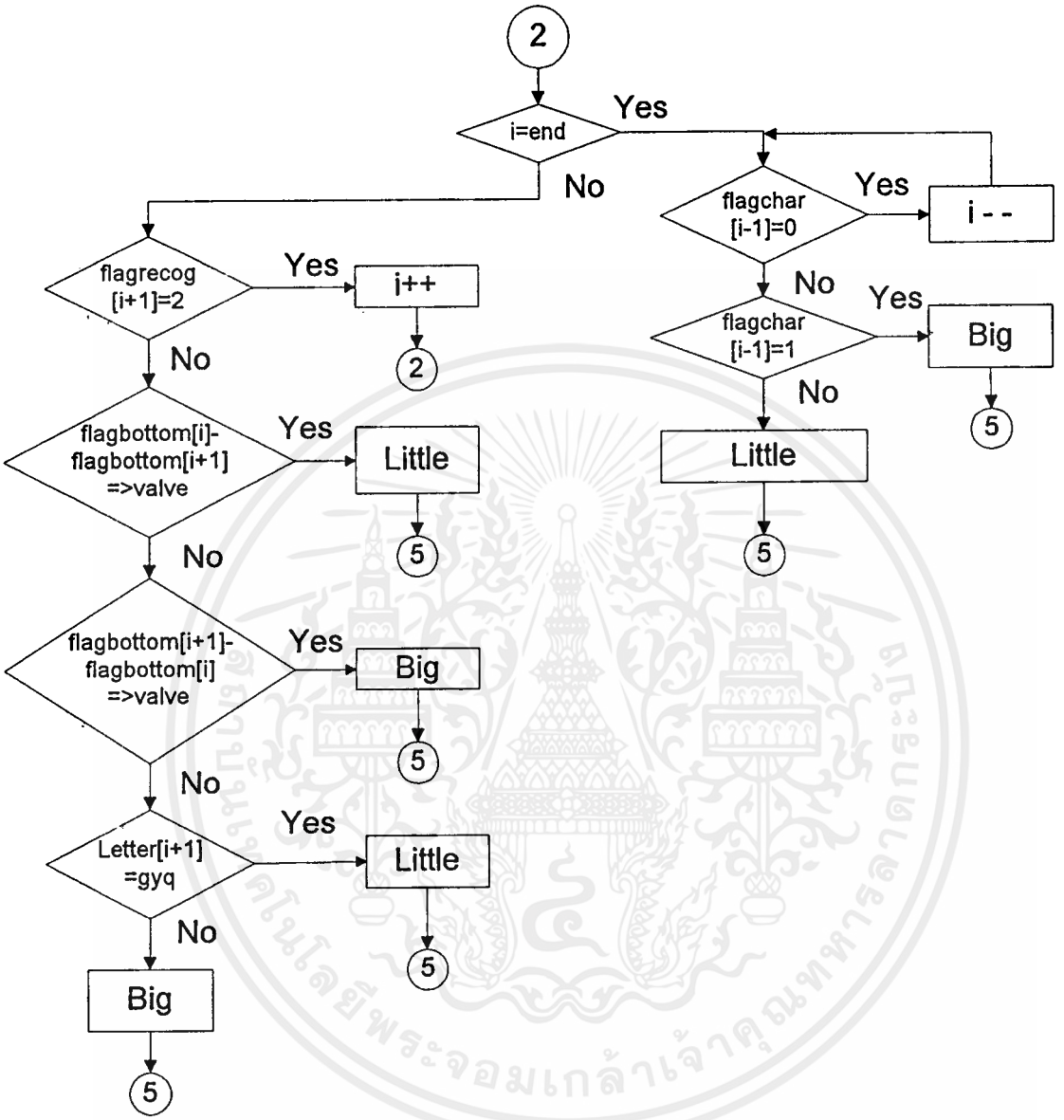
ขั้นตอนที่ 1.1



รูปที่ 5.25 (ต่อ) แสดง flow chart ในการแยกแยะความแตกต่างตัวอักษรขั้นตอนที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 1.2



รูปที่ 5.25 (ต่อ) แสดง flow chart ในการแยกแยะความแตกต่างตัวอักษรขั้นตอนที่ 1

อธิบายการทำงานของ flow chart ของขั้นตอนการทำงานที่ 1

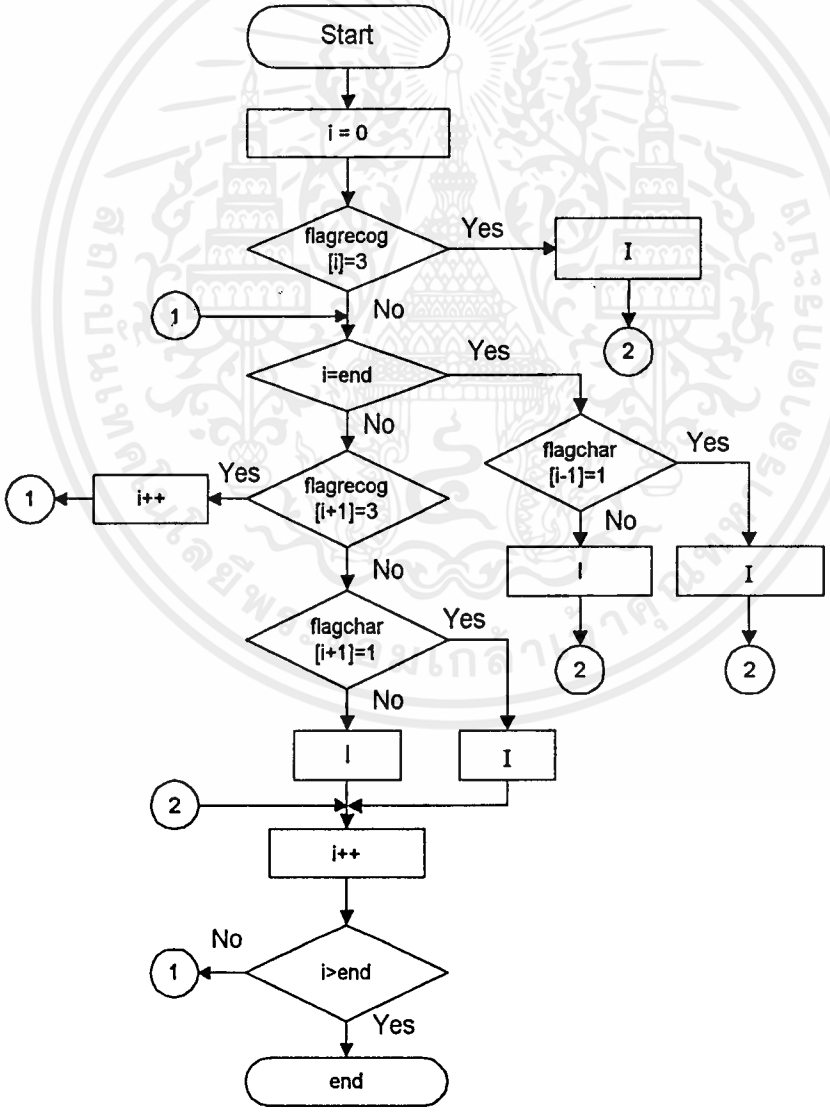
จากผังงานตามรูปที่ 5.25 เป็นขั้นตอนการแยกความแตกต่างของตัวอักษรที่มีรหัสเหมือนกันและตัวอักษรก็เหมือนกันด้วยเพื่อ พิจารณาแยกความแตกต่างโดยใช้หลักการที่ 1 และหลักการที่ 2 โดยเริ่มจากการพิจารณาตัวอักษรตัวแรกว่ามีสภาวะเป็นไปตามเงื่อนไขใด เพื่อตัดสินใจจะกระทำตามหลักการการที่ 1 หรือหลักการที่ 2 เมื่อตัดสินใจได้แล้วก็จะเข้าสู่การพิจารณาแต่ละส่วนของรูปแบบขั้นตอนที่กำหนด โดยใช้หลักการพิจารณาสภาวะแวดล้อมของตัวอักษรรอบๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้างในการพิจารณาแยกความแตกต่างของตัวอักษร เมื่อสามารถระบุได้ว่าตัวอักษรนั้นเป็นตัวอักษรพิมพ์เล็กหรือพิมพ์ใหญ่ก็จะทำการกำหนดค่าในตัวแปร flagchar[i] เสียใหม่ตามสภาวะที่วิเคราะห์ภาพตัวอักษรได้ เช่น ถ้าวิเคราะห์ภาพตัวอักษรได้ตัวพิมพ์ใหญ่ก็กำหนดให้ flagchar[i] เป็น 1 และถ้าเป็นตัวพิมพ์เล็กก็กำหนดให้เป็น 2 หลังจากนั้นทำการวิเคราะห์ภาพตัวอักษรต่อไปจนหมดภาพตัวอักษร

ขั้นตอนที่ 2 การพิจารณาแยกความแตกต่างของตัวอักษรโดยใช้หลักการที่ 3 แสดงดัง flow chart ดังนี้

ขั้นตอนที่ 2.0



รูปที่ 5.26 แสดง flow chart ในการแยกแยะตัวอักษรของ I,l

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<u>อธิบายหลักการทำงานของ flow chart ขั้นตอนการทำงานที่ 2</u>

จากผังงานตามรูปที่ 5.26 จะเป็นการแยกความแตกต่างของตัวอักษรระหว่าง I (ไอใหญ่) และ l (แอลเล็ก) ซึ่งรหัสตัวอักษรจะเหมือนกันและภาพตัวอักษรก็เหมือนกันด้วย เราไม่สามารถใช้ขั้นตอนเบื้องต้นในการแยกความแตกต่างของตัวอักษรสองตัวนี้ได้ เราจึงได้กำหนดเรื่องไขในการวิจัยในฉบับนี้ว่าให้ตัวอักษรตัวแรกของคำ เป็นตัวพิมพ์ใหญ่ส่วนตัวถัดมาจะเป็นพิมพ์ใหญ่หรือ เล็กก็ได้

การพิจารณาแยกความแตกต่างตามขั้นตอนนี้จะพิจารณาหลังจากขั้นตอนที่ 1 ได้กระทำเสร็จแล้ว เราจึงจะเริ่มพิจารณาตามขั้นตอนที่ 2 โดยเริ่มพิจารณาค่าตัวแปร flagrecog ของตัวอักษรตัวแรกถ้ามีค่าเท่ากับ 3 เราจะกำหนดให้ตัวอักษรนี้เป็นตัว I (ไอใหญ่) ถ้าไม่ใช่ให้พิจารณาค่าตัวแปร flagrecog ของตัวอักษรต่อไป ถ้ามีค่าเท่ากับ 3 ให้พิจารณาค่าสภาวะของตัวแปร flagchar ของตัวถัดไปว่าเป็นตัวพิมพ์เล็กหรือพิมพ์ใหญ่ ถ้าเป็นตัวใหญ่ให้กำหนดเป็นตัว I (ไอใหญ่) แต่ถ้าเป็นตัวพิมพ์เล็กกำหนดให้เป็นตัว l (แอลเล็ก) หรือถ้าตัวอักษรถัดไปค่าตัวแปร flagrecog เท่ากับ 3 ก็ให้วนลูปพิจารณาเหมือนเดิมจนกว่าจะตัดสินใจได้ หลังจากนั้นก็วิเคราะห์ภาพตัวอักษรต่อไปจนหมดภาพตัวอักษร ในกรณีวิเคราะห์ตัวอักษรได้เป็นตัวพิมพ์ใหญ่ให้กำหนดค่าตัวแปรใน flagchar เท่ากับ 1 หรือ ถ้าวิเคราะห์ภาพตัวอักษรได้เป็นตัวพิมพ์เล็กให้กำหนดค่าในตัวแปร flagchar เท่ากับ 2

เมื่อเราสามารถพิจารณาแยกความแตกต่างของตัวอักษรได้หมดทุกๆ ตัวแล้วเราจะกำหนดค่าของตัวอักษรที่เราพิจารณาเป็นรหัส ASCII เพื่อนำไปใช้ในงานต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6

โปรแกรมและผลการทดลอง

6.1 โปรแกรมการรู้จำอักษรพิมพ์ภาษาอังกฤษบนสันหนังสือ

โปรแกรมนี้เขียนด้วยภาษาซีเวอร์ชั่น 3.0 (Turbo C Version 3.0) ใช้สำหรับเครื่องไมโครคอมพิวเตอร์ทั่วไปโดยโปรแกรมนี้จะรับข้อมูลภาพจากแฟ้มข้อมูลที่มีฟอร์แมตเป็น .PCX เท่านั้น ดังนั้นการใช้งานโปรแกรมนี้ ต้องมีข้อมูลภาพอักษรที่เป็นฟอร์แมต .PCX ถ้าเป็นฟอร์แมตอื่นๆ เช่น .JPG, .BMP, .GIF ฯลฯ ก็อาจใช้โปรแกรมที่สามารถ Convert ฟอร์แมตให้เป็น .PCX เสียก่อนก็ได้ โดยแฟ้มข้อมูลภาพต้องถูกจัดเก็บลงแผ่นดิสก์หรือ ฮาร์ดดิสก์ ซึ่งจะสามารถนำแฟ้มข้อมูลภาพมาวิเคราะห์ในโปรแกรมนี้ได้

คุณสมบัติของโปรแกรม

1. อ่านชื่อไฟล์ข้อมูลภาพชนิด .PCX ได้
2. แสดงชื่อแฟ้มข้อมูลภาพบนจอภาพเพื่อสามารถเลือกแฟ้มข้อมูลภาพได้
3. โปรแกรมสามารถตัดค่า Threshold ได้อย่างอัตโนมัติ
4. โปรแกรมสามารถแยกสันหนังสือแต่ละเล่มออกจากกลุ่มของสันหนังสือได้
5. โปรแกรมสามารถแยกกลุ่มสันหนังสือเพื่อตัดสินใจนำกลุ่มข้อมูลภาพชื่อสันหนังสือนำมาวิเคราะห์ได้
6. โปรแกรมสามารถแยกภาพตัวอักษรออกจากกลุ่มชื่อสันหนังสือได้
7. โปรแกรมสามารถวิเคราะห์ภาพตัวอักษรโดยใช้คุณสมบัติทางโทโปโลยีของจุดภาพที่ประกอบไปด้วยจุดปลาย จุดต่อ จุดแยก จุดตัดและจุดโดดเดี่ยว มากำหนดเป็นตัวอักษร
8. ตัวอักษรใดที่ไม่สามารถวิเคราะห์ได้ จะแสดงภาพรูปร่างอักษรตัวนั้นให้ผู้ใช้ตัดสินใจเลือกรหัส ASCII ได้เอง
9. ผลลัพท์ที่ได้จากการวิเคราะห์จะถูกจัดเก็บเป็นรหัส ASCII Code

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การใช้งานโปรแกรม

โปรแกรมและไฟล์ที่เกี่ยวข้อง

ในการเตรียมโปรแกรมสำหรับใช้งานประกอบด้วย

1. โปรแกรมการรู้จำรูปแบบอักษรพิมพ์ภาษาอังกฤษบนสันหนังสือ ประกอบด้วยไฟล์ข้อมูลดังนี้

ENGBOOK.EXE	เป็นโปรแกรมหลักสำหรับการวิเคราะห์
ENGBOOK.DAT	เป็นฐานข้อมูลรหัสต้นแบบของตัวอักษร

2. ไฟล์ข้อมูลภาพที่มีฟอร์แมตเป็น .PCX ที่ได้มาจากกล้อง CCD ดังนี้

Book1.PCX

Book2.PCX

อุปกรณ์ที่ใช้ประกอบร่วม

อุปกรณ์ประกอบร่วมการทำงานของโปรแกรมวิเคราะห์การรู้จำตัวอักษรพิมพ์ภาษาอังกฤษบนสันหนังสือ ประกอบด้วย

1. เครื่องไมโครคอมพิวเตอร์ (Micro Computer)
2. จอภาพ (Monitor)
3. Keyboard
4. กล้อง CCD (Charge Coupled Device)
5. Card image
6. โปรแกรมวิเคราะห์การรู้จำตัวอักษรพิมพ์ภาษาอังกฤษบนสันหนังสือ

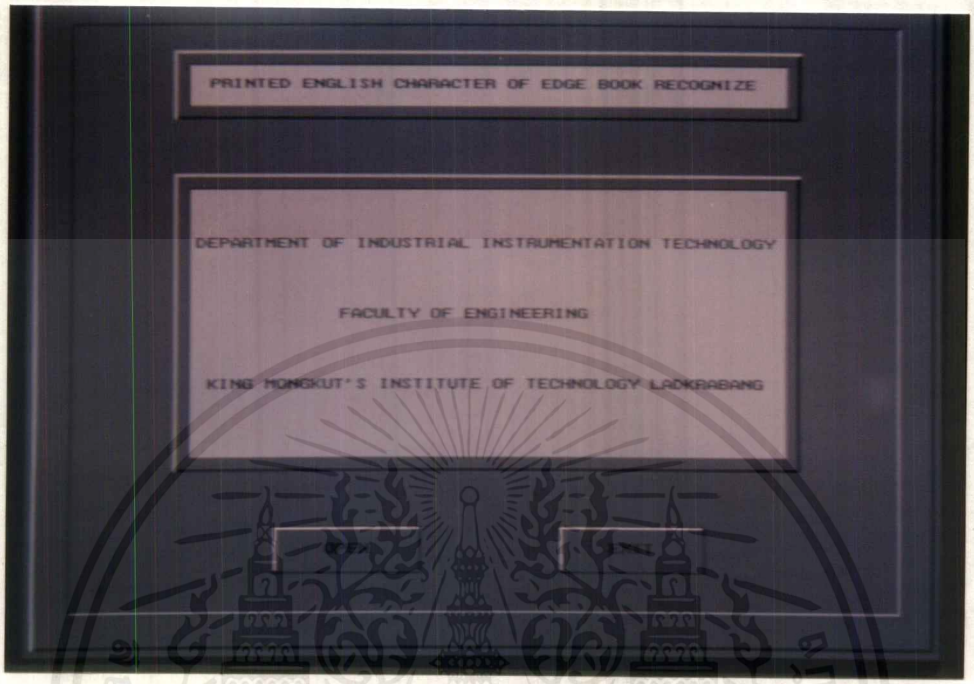
รายละเอียดการทำงาน

ทำการ RUN โปรแกรมการรู้จำโดยพิมพ์ ดังนี้

A:\ ENGBOOK <Enter>

จะปรากฏหน้าจอดังรูปที่ 6.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

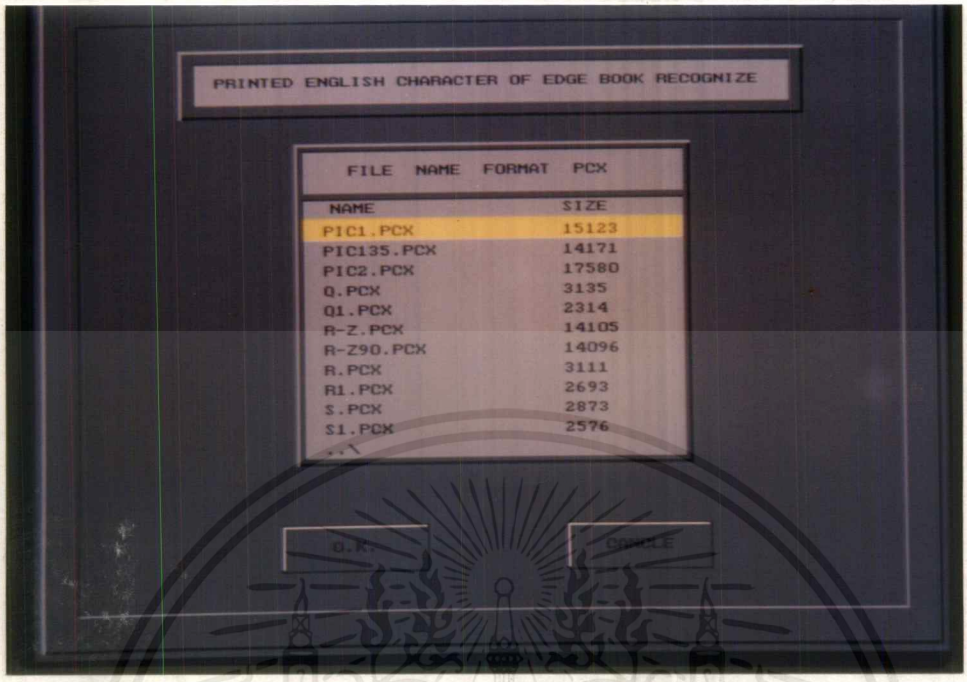


รูปที่ 6.1 แสดงหน้าจอภาพก่อนทำการวิเคราะห์การรู้จำ

<u>การเปิดแฟ้มข้อมูลภาพ</u>

เลือก "open" จะแสดงรายละเอียดชื่อแฟ้มข้อมูลภาพ ดังรูปที่ 6.2 สามารถใช้ Arrow Key เพื่อเลื่อนแฟ้มข้อมูลภาพขึ้นลงทีละแฟ้มข้อมูลหรือครั้งละ 1 หน้าจอภาพโดยใช้ PgUp หรือ PgDn หลังจากนั้น กด Enter จะปรากฏภาพสันหนังสือที่จะทำการวิเคราะห์ ดังรูปที่ 6.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.2 แสดงการเปิดแฟ้มข้อมูลภาพ .PCX

รูปที่ 6.3 แสดงรูปภาพของแฟ้มข้อมูลภาพที่จะทำการรู้จำตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

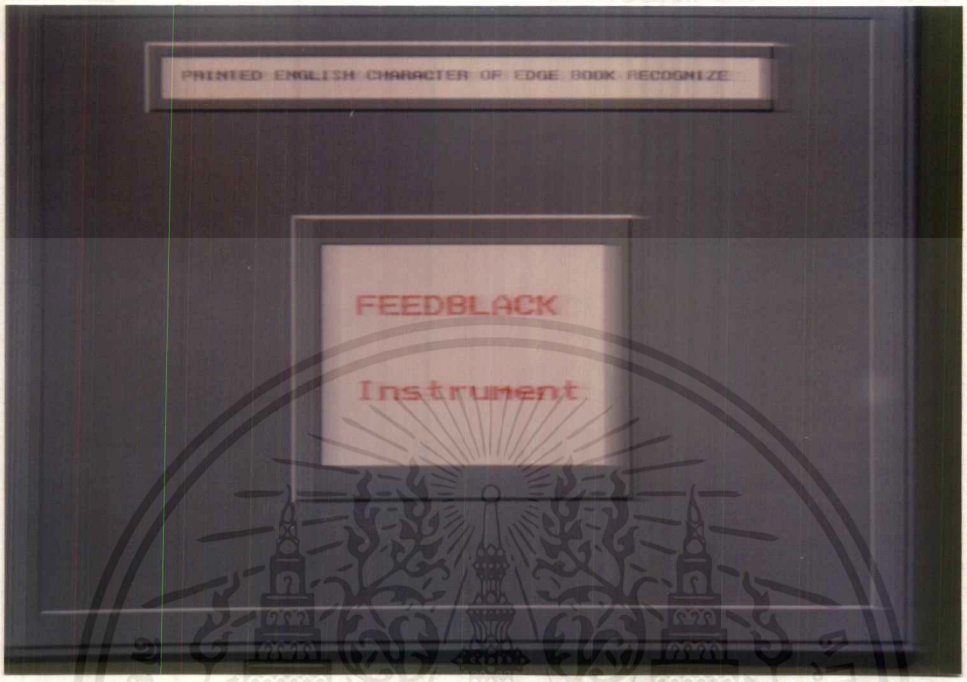
<u>การวิเคราะห์ภาพสันหนังสือ</u>

1. หลังจากเลือกแฟ้มข้อมูลภาพที่ต้องการและปรากฏภาพของข้อมูลที่ถูกเลือกแล้ว โปรแกรมจะถามเพื่อความแน่ใจก่อนเข้าสู่การรู้จำต่อไป
2. เมื่อตอบตกลงโดยเลือกปุ่ม “OK” แล้วกดปุ่ม “Enter” โปรแกรมจะเริ่มการรู้จำเมื่อโปรแกรมไม่สามารถวิเคราะห์ภาพอักษรได้ จะแสดงภาพอักษรเพื่อให้ผู้ใช้ตัดสินใจกำหนดค่าตัวอักษร ตามรูปที่แสดงบนหน้าจอดังรูปที่ 6.4

รูปที่ 6.4 แสดงภาพตัวอักษรที่ไม่สามารถวิเคราะห์ได้

เมื่อทำการวิเคราะห์ข้อมูลเสร็จแล้ว โปรแกรมจะแสดงผลการรู้จำหน้าจอ ดังรูปที่ 6.5 และจะถูกจัดเก็บเป็นรหัส ASCII Code เพื่อนำไปใช้งานต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.5 แสดงผลการรู้จำตัวอักษร

6.2 ผลการทดลอง

จากโปรแกรมที่สร้างขึ้นมา เพื่อให้แน่ใจว่าโปรแกรมสามารถใช้งานได้ และมีประสิทธิภาพมากน้อยเพียงใด จึงต้องมีการทดสอบโปรแกรมและทดลองการรู้จำอักษรพิมพ์ภาษาอังกฤษบนสันหนังสือในรูปแบบต่างๆ โดยดำเนินการทดลองโปรแกรมดังนี้

1. การสร้างการรู้จำให้โปรแกรม (Learning)

เริ่มแรกนั้นโปรแกรมจะไม่สามารถรู้จำภาพตัวอักษรได้เลยจนกว่าโปรแกรมจะได้ทำการรู้จำภาพตัวอักษรมาแล้วหนึ่งครั้ง ภาพตัวอักษรที่ผ่านการรู้จำครั้งแรกนี้จะเป็นอักษรต้นแบบดังรูปที่ 6.6 ที่ผ่านขบวนการรู้จำจนได้ค่ารหัสแทนตัวอักษรเก็บไว้เป็นฐานข้อมูล เพื่อการเปรียบเทียบกับชุดรหัสของอักษรตัวใหม่ที่ไม่ทราบว่าเป็นตัวอักษรอะไรให้สามารถระบุได้ถูกต้อง แต่ถ้าวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แล้วไม่ได้รหัสตรงกับฐานข้อมูลต้นแบบโปรแกรม ก็จะนำรหัสฐานข้อมูลใหม่นี้ ไปเก็บเป็นฐานข้อมูลต้นแบบเพิ่มขึ้นอีกต่อไป

ABCDEFGHIJKLMNOPQRSTUVWXYZ

abcdefghijklmnopqrstuvwxyz

1234567890

รูปที่ 6.6 แสดงตัวอย่างตัวอักษรต้นแบบ CordiaUPC

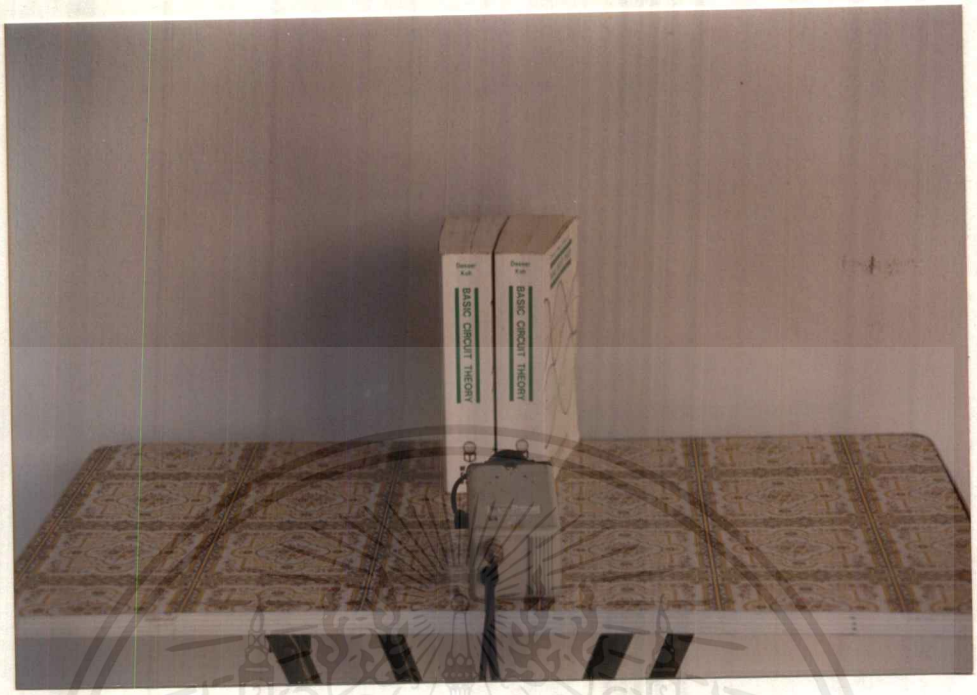
โดยการรู้จำครั้งแรกนั้น การที่จะให้รหัสภาพตัวอักษรตัวนั้นตรงกับรหัสแทนตัวอักษรใดนั้นขึ้นอยู่กับผู้ใช้โปรแกรมจะตัดสินใจ จากภาพอักษรที่วิเคราะห์ออกมาแล้ว

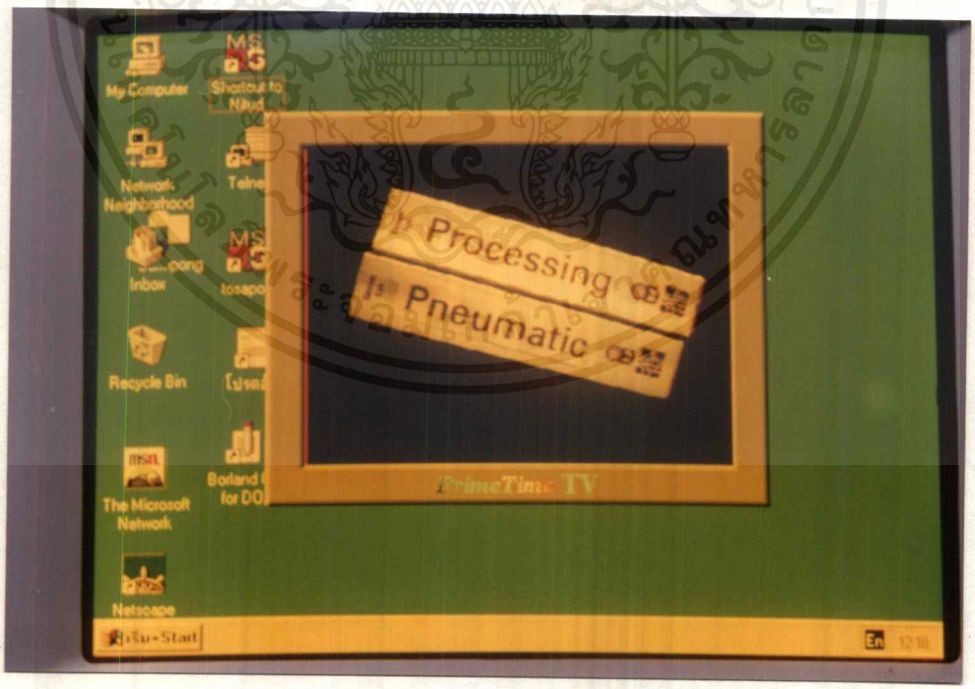
2. การทดลองการรู้จำอักษรพิมพ์ภาษาอังกฤษบนสันหนังสือ

เนื่องจากการทดลองการรู้จำกับสันหนังสือจริงๆ ยังไม่สามารถทำการรู้จำและวิเคราะห์ชื่อของหนังสือได้ เพราะมีเงื่อนไขและองค์ประกอบต่างๆ บนสันหนังสือที่โปรแกรมไม่สามารถแยกแยะชื่อหนังสือออกจากสันหนังสือได้ เราจึงต้องกำหนดเงื่อนไขในการทดลองสำหรับปริญญานิพนธ์ฉบับนี้ ดังนี้

1. สันหนังสือต้องมีพื้นสีขาว
2. ชื่อของหนังสือที่อยู่บนสันหนังสือต้องเป็นสีเข้ม
3. องค์ประกอบต่างๆ บนสันหนังสือต้องมีสีเข้ม
4. ชื่อของหนังสือต้องอยู่อย่างอิสระไม่มีองค์ประกอบอื่นๆ ที่ไม่ใช่ชื่อของหนังสือเกี่ยวข้อง
5. ชื่อของสันหนังสือควรจะมีขนาดใหญ่ประมาณ 1 cm. ขึ้นไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

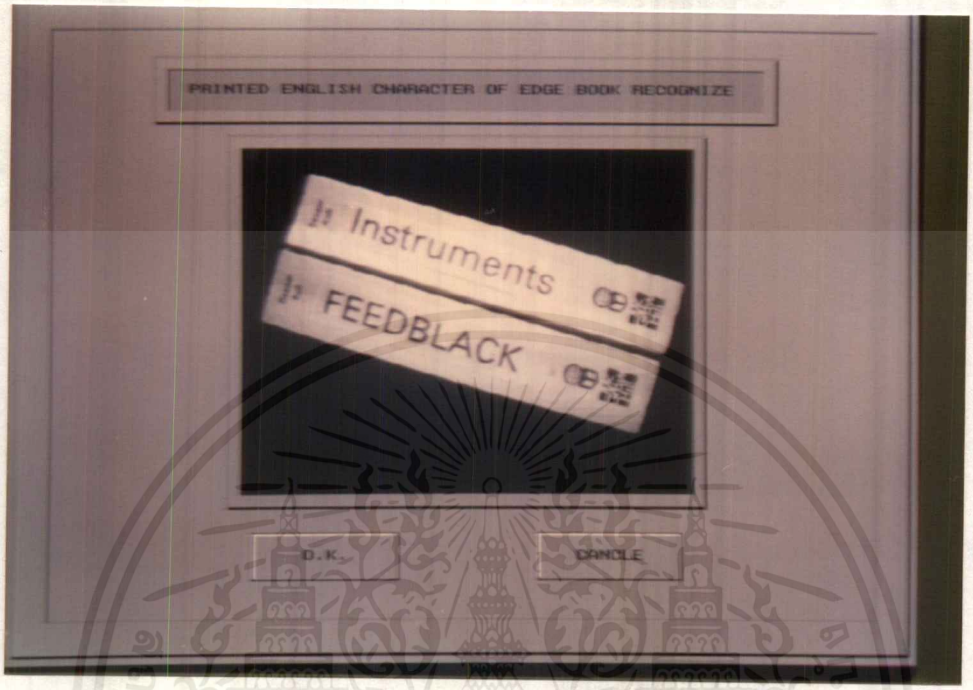




รูปที่ 6.7 แสดงรูปที่ใช้ในการทดลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อเงื่อนไขสภาวะต่างๆ เป็นไปตามที่กำหนดจึงเริ่มทำการทดลองโปรแกรมการรู้จำอักษรพิมพ์ภาษาอังกฤษดังนี้



รูปที่ 6.8 แสดงภาพต้นแบบสันหนังสือถ่ายจากกล้อง CCD

1. ถ่ายภาพสันหนังสือด้วยกล้อง CCD (Charge Coupled Device) แล้วบันทึกลงแฟ้มข้อมูลไฟล์ .PCX

รูปที่ 6.9 แสดงภาพข้อมูลภาพ 2 ระดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. แปลงภาพต้นแบบ gray scale 256 ระดับ เป็นภาพ binary 2 ระดับ เพื่อประมวลผลภาพดิจิตอล โดยการตัดค่า Threshold อัตโนมัติ

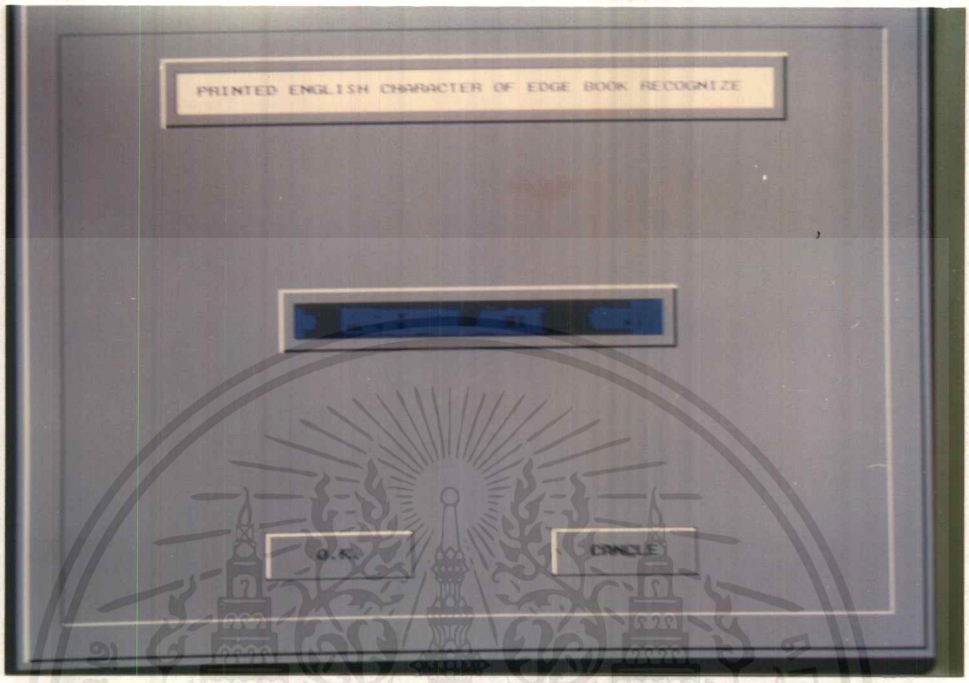
รูปที่ 6.10 แสดงภาพการแยกสันหนังสือออกจากกลุ่ม

3. เมื่อได้ภาพ binary 2 ระดับ พร้อมจะประมวลผลภาพและการวิเคราะห์การรู้จำขั้นตอนแรก คือ การแยกสันหนังสือออกจากกลุ่มของสันหนังสือ

รูปที่ 6.11 แสดงการหมุนภาพสันหนังสือเข้าหาแกนนอน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. เมื่อสันหนังสือมิได้อยู่ในแนวระนาบจึงต้องมีการหมุนสันหนังสือเข้าหาแนวระนาบเพื่อการวิเคราะห์ภาพตัวอักษร



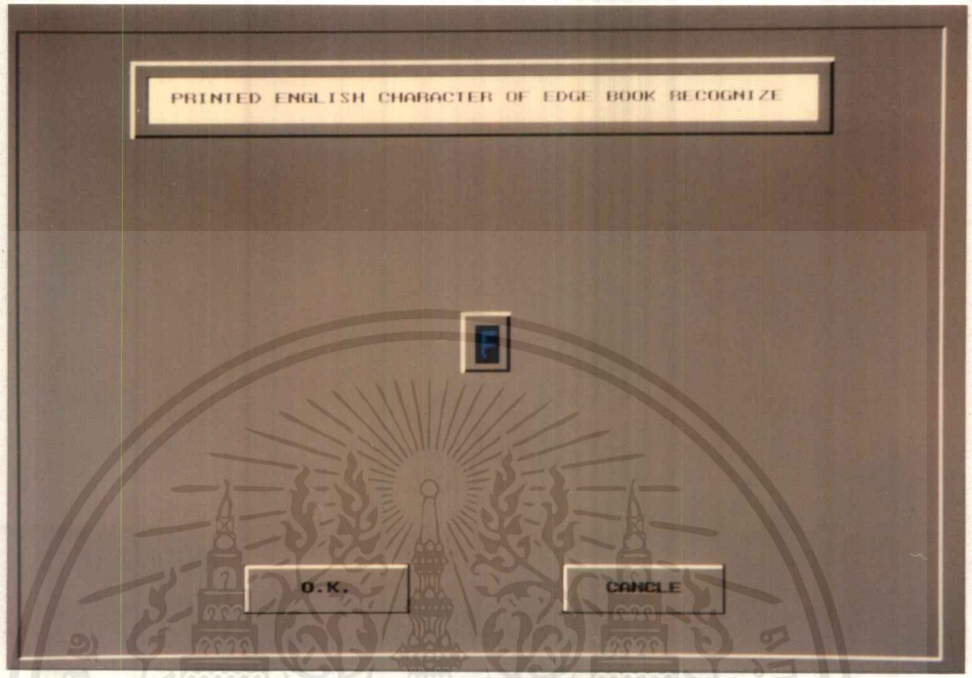
รูปที่ 6.12 แสดงการรวมกลุ่มข้อมูลบนสันหนังสือ

5. รวมกลุ่มข้อมูลบนสันหนังสือเพื่อตัดสินใจทำภาพตัวอักษรมาวิเคราะห์การรู้จำ

รูปที่ 6.13 แสดงการแยกกลุ่มตัวอักษรออกจากสันหนังสือ

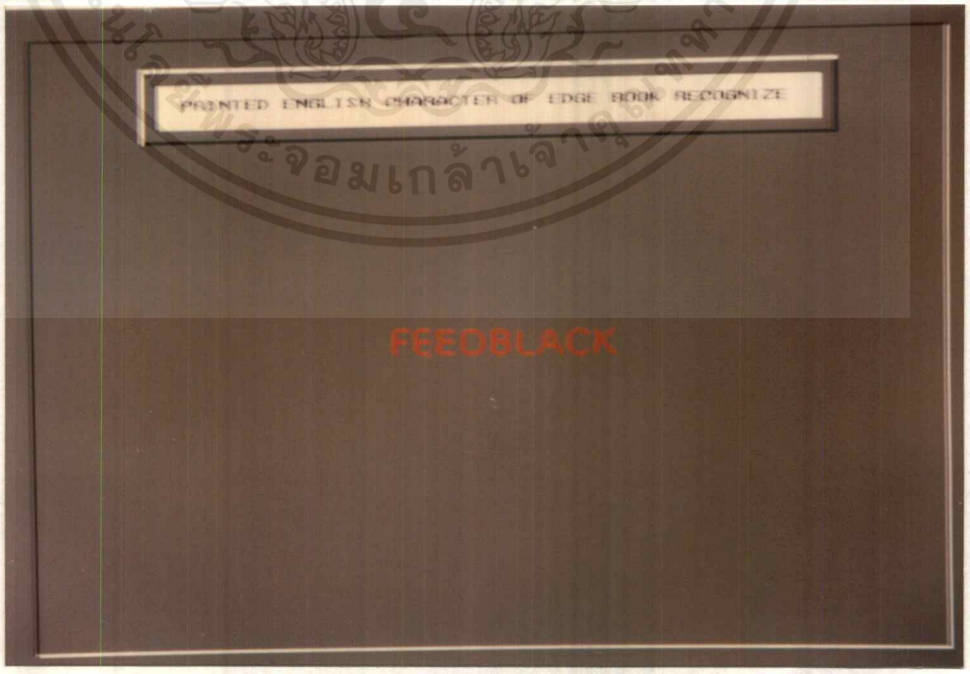
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. ตัดสินใจเลือกกลุ่มตัวอักษรบนสันหนังโดยใช้หลักการความยาวพอแนวระนาบเปรียบเทียบกับกลุ่มอื่นๆ



รูปที่ 6.14 แสดงการแยกตัวอักษรออกจากกลุ่มตัวอักษร

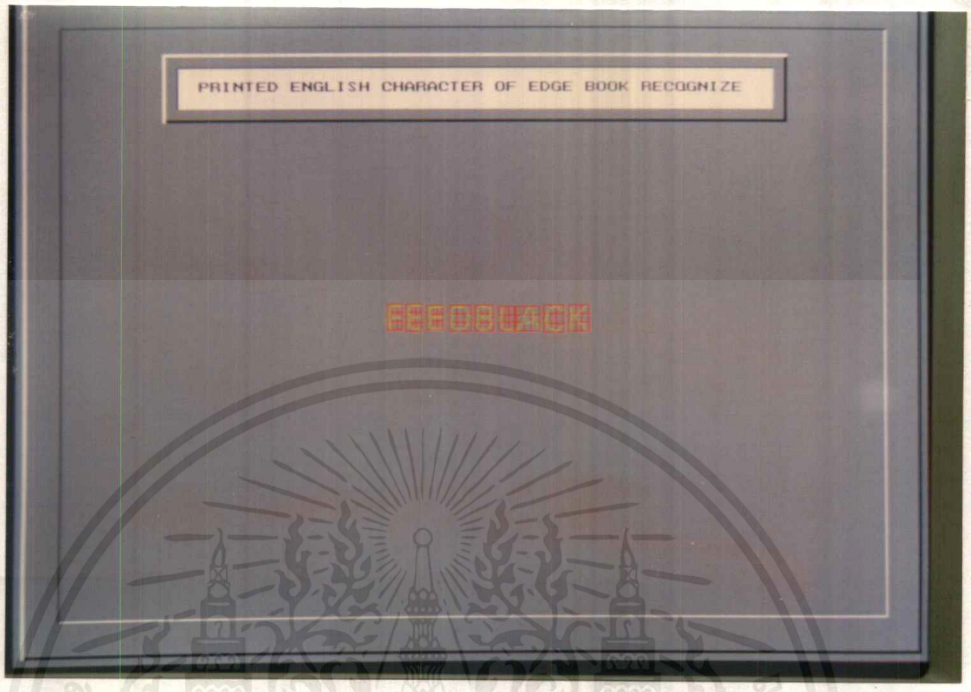
7. แยกตัวอักษรออกจากกลุ่มตัวอักษร



รูปที่ 6.15 แสดงการทำภาพตัวอักษรให้บาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8. ทำภาพตัวอักษรที่แยกออกมาให้บางก่อนนำมาวิเคราะห์ค่าคุณสมบัติทางโทโปโลยี

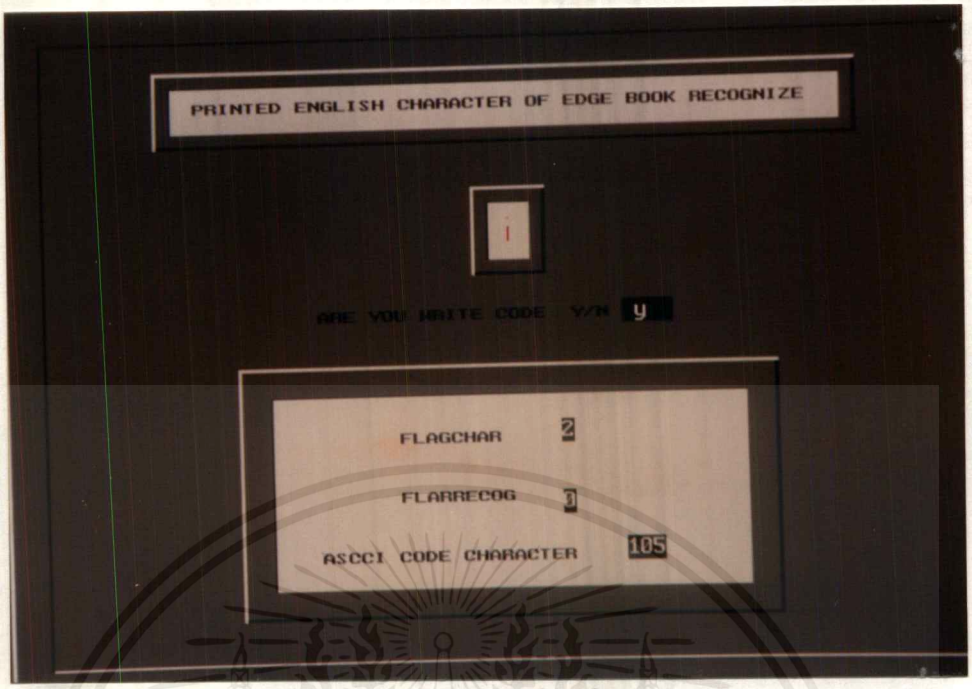


รูปที่ 6.16 แสดงการแบ่งภาพตัวอักษรออกเป็น 9 ช่อง

9. แบ่งภาพตัวอักษรออกเป็น 9 ช่องเพื่อเข้ารหัสภาพตัวอักษร

รูปที่ 6.17 แสดงผลการรู้จำตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



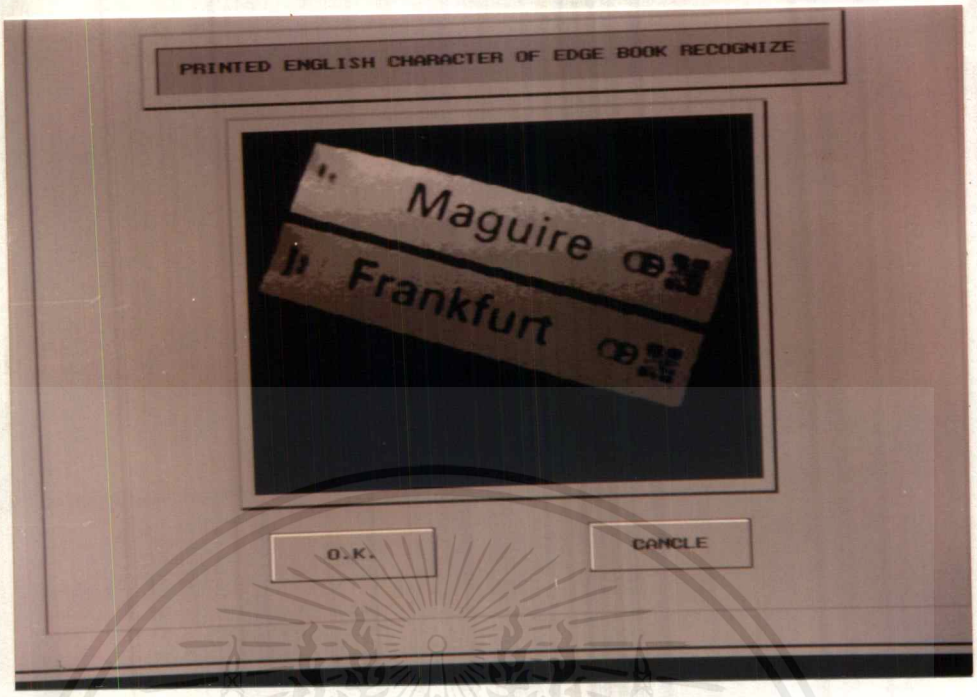
รูปที่ 6.18 แสดงภาพตัวอักษรเมื่อไม่สามารถรู้จำตัวอักษรได้

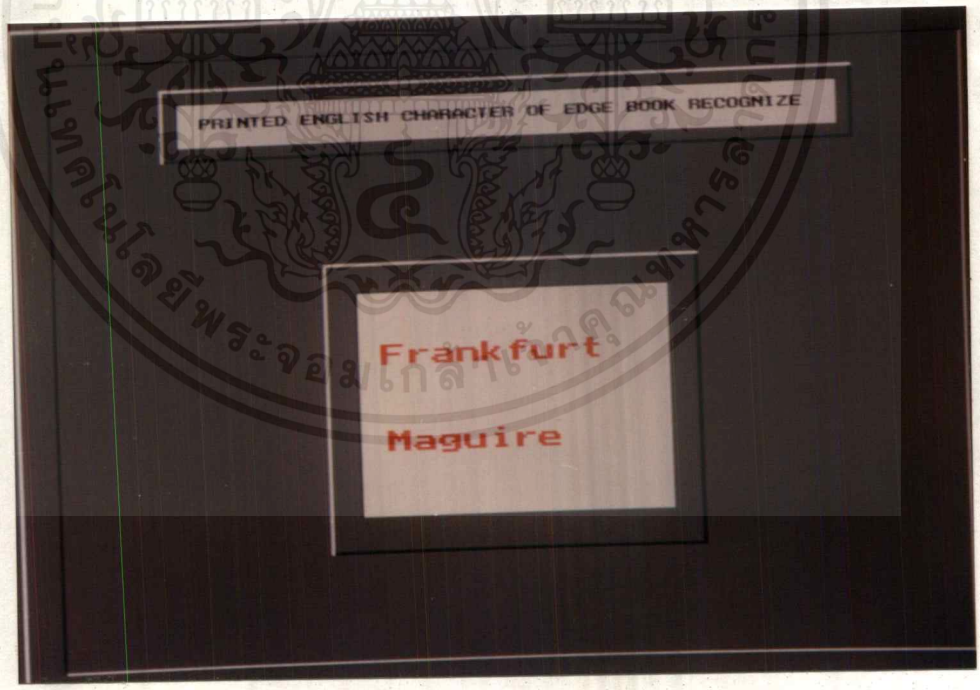
11. เมื่อรหัสต้นแบบในฐานข้อมูลไม่ตรงกับรหัสที่วิเคราะห์ได้จะแสดงภาพตัวอักษรพร้อมให้เราป้อนข้อมูลทาง Key board

หมายเหตุ ในการเพิ่มรหัส ASCII Code ของตัวอักษรสามารถเป็นฐาน 16 ต้องทำเป็นฐาน 10 เสียก่อน.

6.3 ตัวอย่างผลการทดลอง

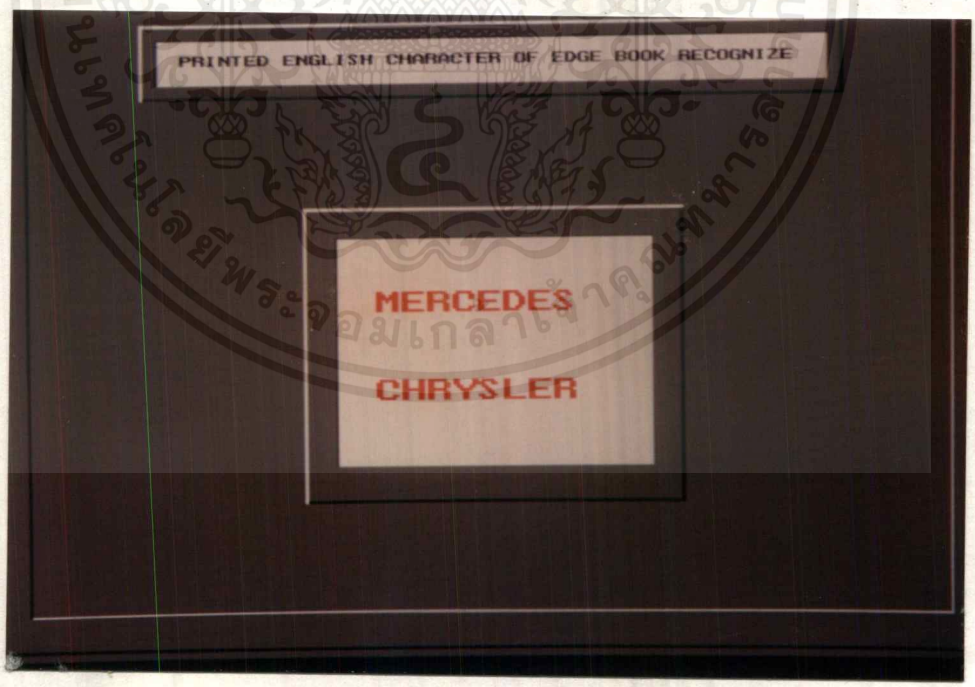
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้





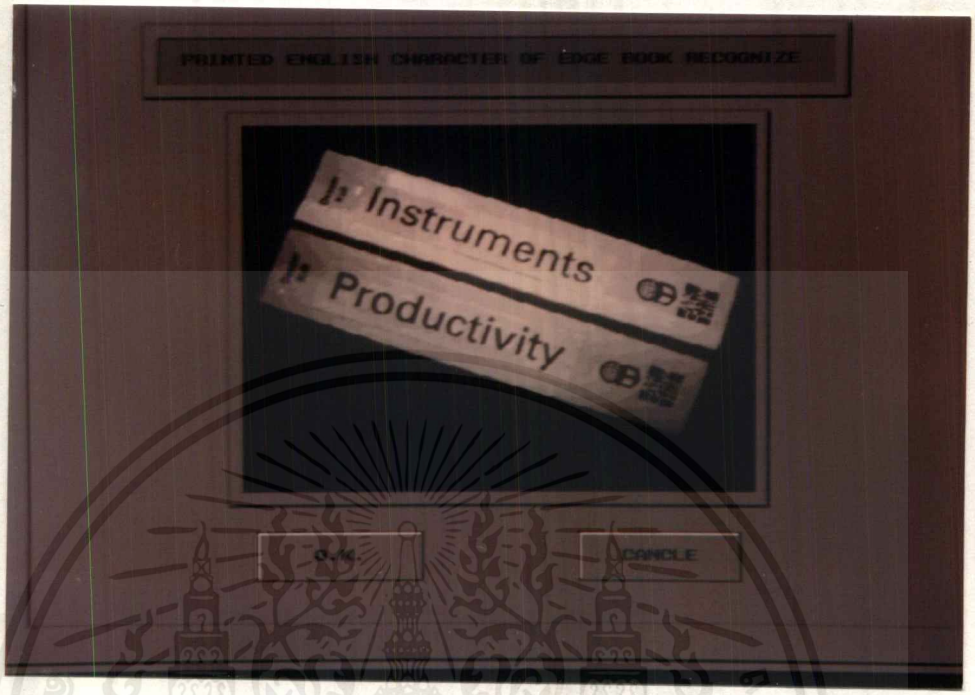
รูปที่ 6.19 แสดงตัวอย่างผลการทดลองที่ 1

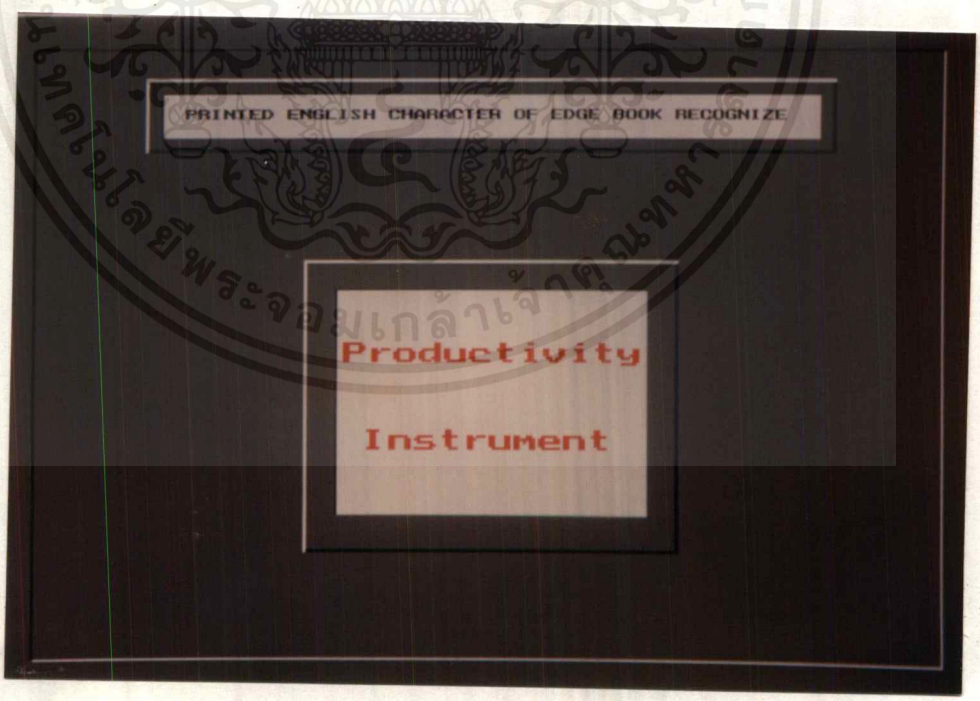
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.20 แสดงตัวอย่างผลการทดลองที่ 2

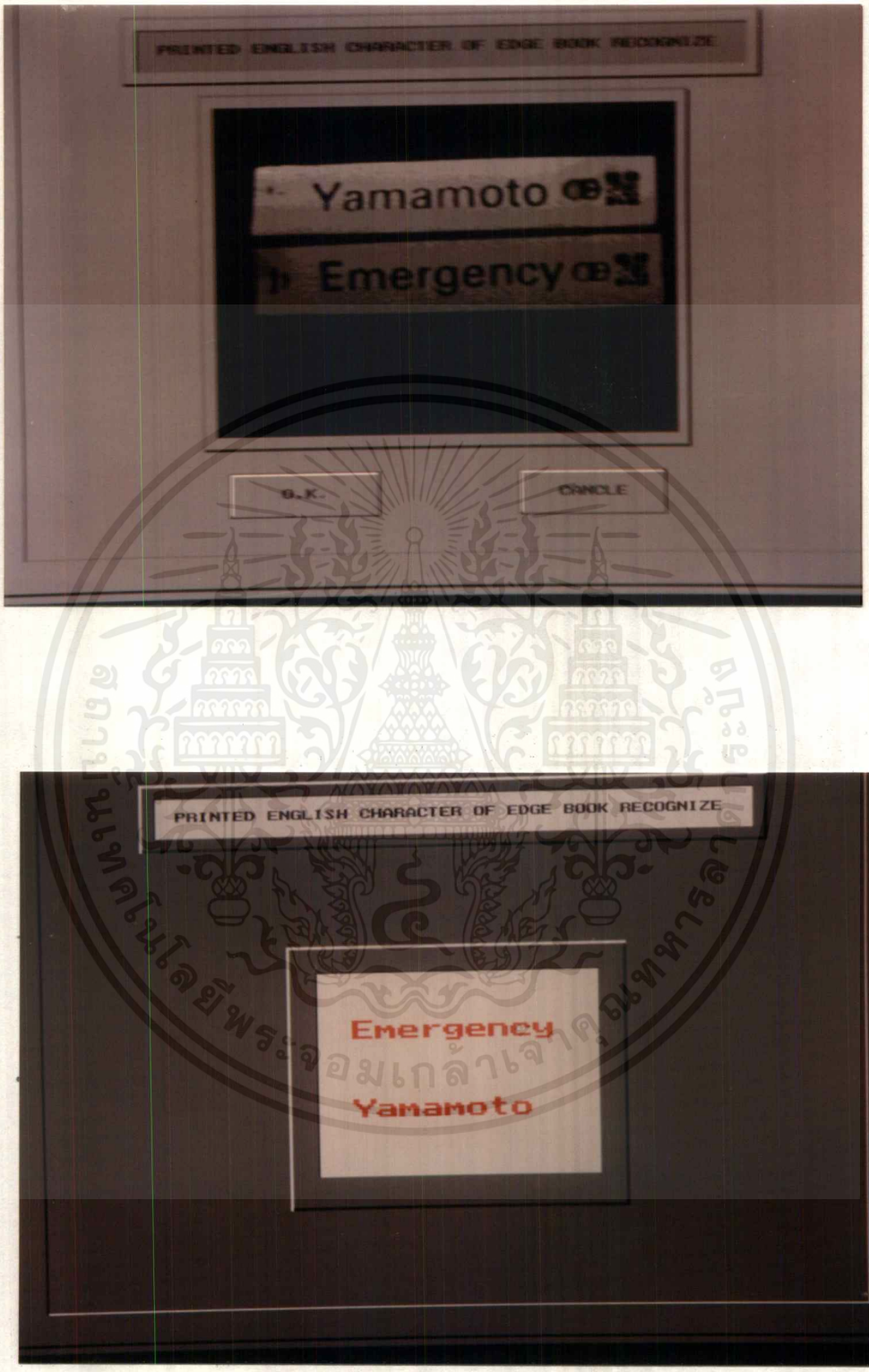
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้





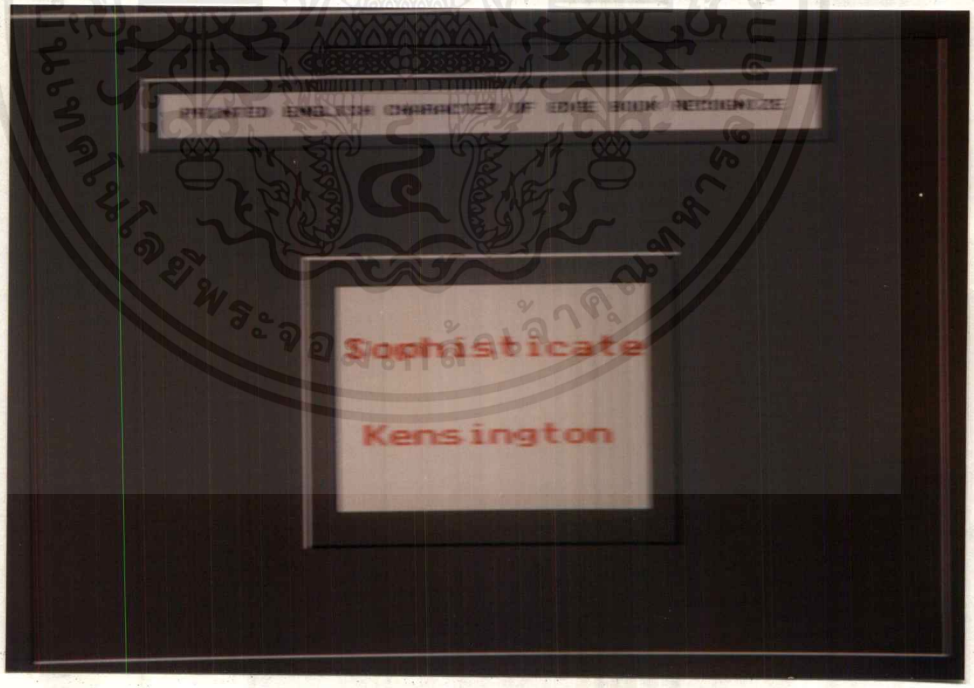
รูปที่ 6.21 แสดงตัวอย่างผลการทดลองที่ 3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



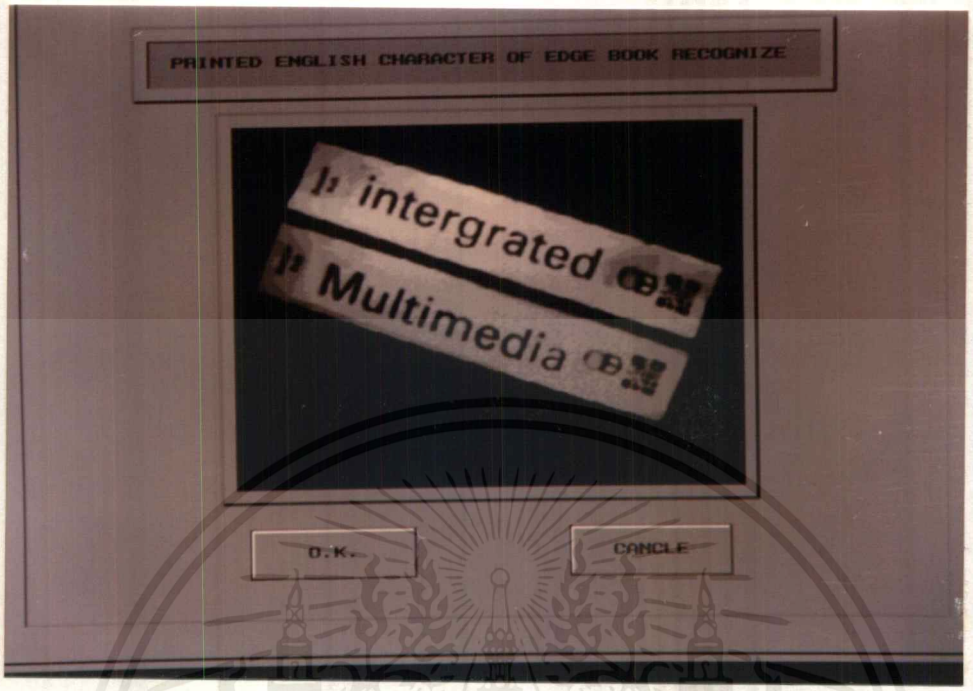
รูปที่ 6.22 แสดงตัวอย่างผลการทดลองที่ 4

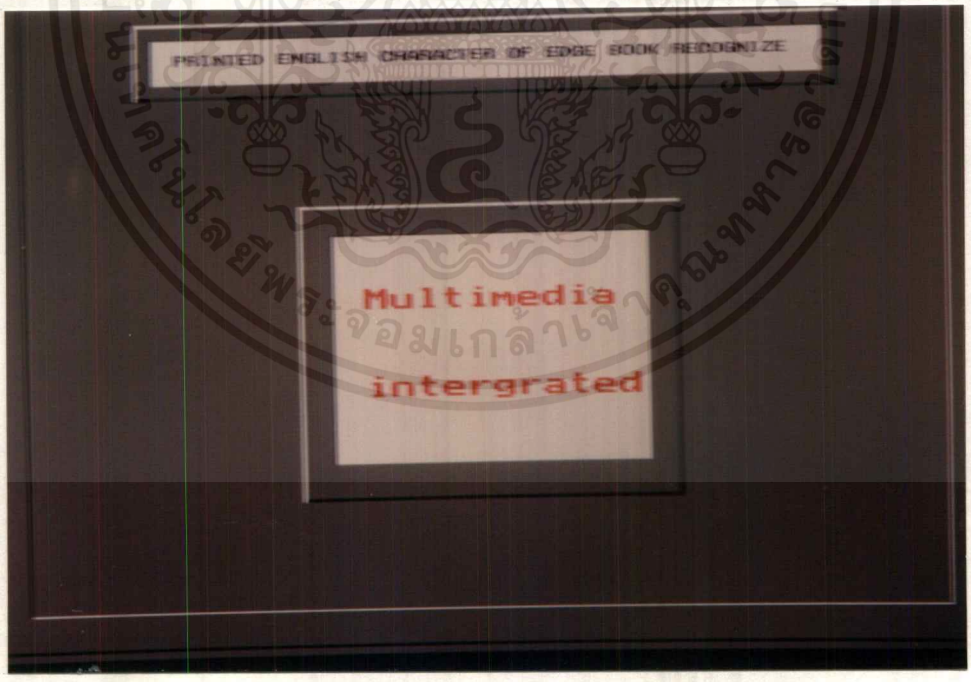
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.23 แสดงตัวอย่างผลการทดลองที่ 5

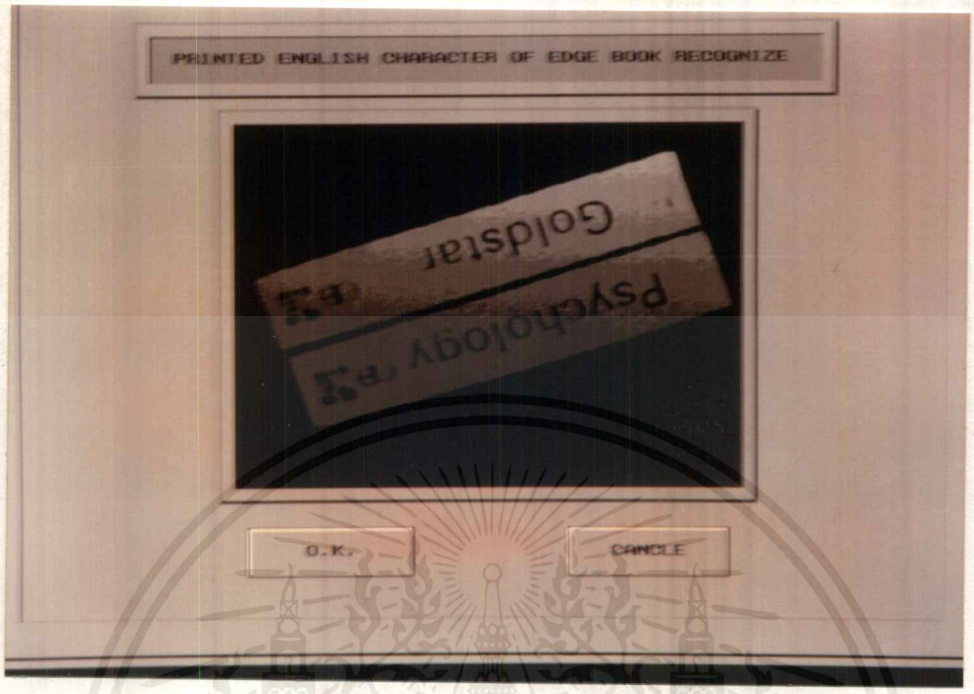
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

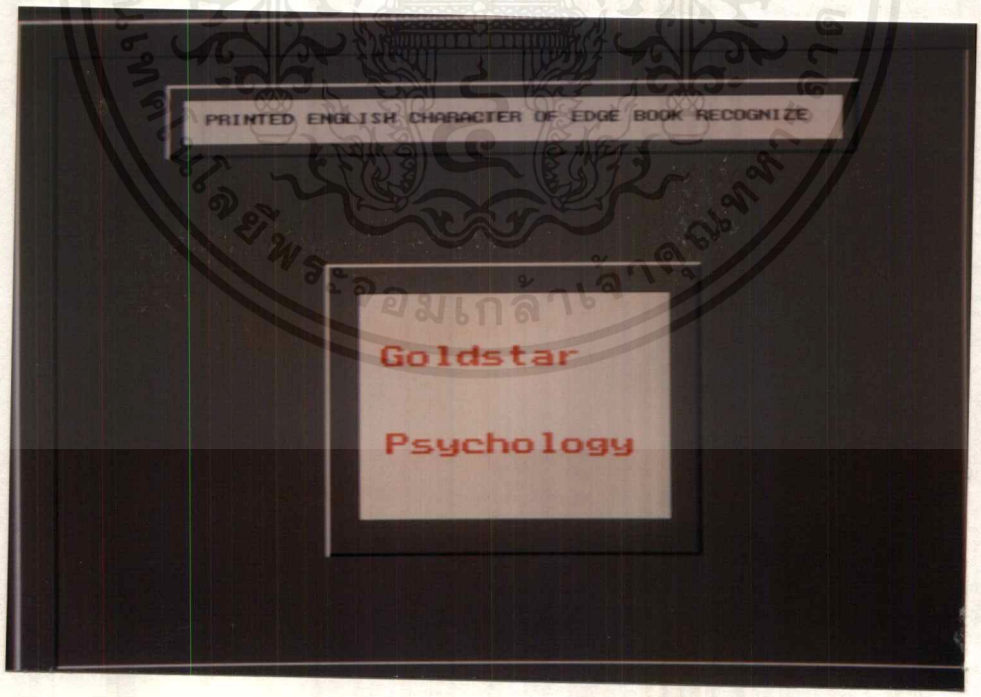




รูปที่ 6.24 แสดงตัวอย่างผลการทดลองที่ 6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้





รูปที่ 6.25 แสดงตัวอย่างผลการทดลองที่ 7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.4 สรุปผลการทดลอง

จากการทดลองได้ทำการทดลองกับสันหนังสือ 100 เล่มผลปรากฏว่าสามารถวิเคราะห์ชื่อของสันหนังสือได้ถูกต้องประมาณ 20 เล่ม และวิเคราะห์ชื่อของสันหนังสือได้บางส่วนประมาณ 60 เล่ม และไม่สามารถวิเคราะห์ชื่อของสันหนังสือได้เลยประมาณ 20 เล่ม จากผลการทดลองเราสามารถสรุปได้ว่าปัจจัยที่มีผลต่อการทดลองนี้ คือ คุณภาพของภาพข้อมูลที่จะทำการวิเคราะห์การรู้จำจะต้องมีความพอดี ไม่สว่างมากเกินไปหรือเข้มเกินไป ซึ่งจะมีผลต่อข้อมูลภาพหรือจากการทำให้บางแล้ว และปริมาณแสดงที่ตกกระทบบนสันหนังสือควรมีความสว่างที่เฉลี่ยเท่าๆ กัน ตลอดทั้งกลุ่มของสันหนังสือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 7
บทสรุป

สรุปการทำงานของระบบการรู้จำตัวอักษรพิมพ์ภาษาอังกฤษบนสันหนังสือ

1. ข้อจำกัดการทำงานของระบบ

1.1 ข้อจำกัดของข้อมูลภาพสันหนังสือที่ถ่ายจากกล้อง CCD

เนื่องจากการนำข้อมูลภาพสันหนังสือมาจากกล้อง CCD คุณภาพของข้อมูลภาพสันหนังสือขึ้นอยู่กับการเก็บข้อมูลภาพสันหนังสือ ลักษณะของการเก็บข้อมูลภาพด้วยกล้อง CCD ขึ้นอยู่กับองค์ประกอบหลายๆ องค์ประกอบ แต่องค์ประกอบที่สำคัญคือ การจัดแสงสำหรับการถ่ายภาพด้วยกล้อง CCD เป็นองค์ประกอบหลักของการเก็บภาพจะต้องทำให้แสงที่มาตกกระทบกับสันหนังสือ และตัวอักษรเป็นสีทึบ พื้นของตัวอักษรบนสันหนังสือเป็นสีขาว ตัดกับข้อมูลบนสันหนังสือ เมื่อจัดสภาวะของแสงที่ตกกระทบตรงตามเงื่อนไขแล้วยังต้องคำนึงถึงความเข้มและความคมชัดของข้อมูลภาพบนสันหนังสือด้วย เพราะจะมีผลกระทบไปถึงการทำภาพ 2 ระดับ ถ้าแสดงที่กระทบบนสันหนังสือมีความเข้มมากเกินไปหรือสว่างเกินไปจะทำให้ข้อมูลภาพหลังจากทำภาพเป็น 2 ระดับแล้วเป็นค่าข้อมูลเดียวกันหมดทั้งภาพ หรือได้ภาพข้อมูลที่เราไม่ต้องการ ดังนั้น การจัดแสดงจึงมีความสำคัญมาก กับข้อมูลภาพอินพุทที่จะนำเข้ามาวิเคราะห์การรู้จำถ้าได้ข้อมูลภาพที่ดี การรู้จำจะมีประสิทธิภาพสูงตามไปด้วย

1.2 ข้อจำกัดของการทำงานในการวิเคราะห์ระบุส่วนประกอบของสันหนังสือ

เนื่องจากลักษณะข้อมูลภาพที่อยู่บนสันหนังสือมีรูปแบบและองค์ประกอบที่ไม่แน่นอน ดังนั้นการวิจัยในปริญญานิพนธ์ฉบับนี้จึงต้องมีการกำหนดเงื่อนไขขององค์ประกอบของข้อมูลภาพบนสันหนังสือ ดังที่กำหนดข้างต้น เพื่อให้โปรแกรมสามารถทำการวิเคราะห์การรู้จำตัวอักษรได้ ซึ่งองค์ประกอบของข้อมูลภาพสันหนังสือจริงๆ เป็นข้อจำกัดการทำงานของระบบการรู้จำที่สร้างขึ้นมา

1.3 ข้อจำกัดของการทำงานในการวิเคราะห์การรู้จำชื่อหนังสือบนสันหนังสือ

เนื่องจากเทคนิคและวิธีการที่เลือกใช้ในการรู้จำตัวอักษรเป็นการใช้คุณสมบัติทางโทโปโลยี คุณสมบัติเหล่านี้เป็นตัวกำหนดค่ารหัสของตัวอักษรเพื่อนำมาเปรียบเทียบกับฐานข้อมูลการรู้จำ ดังนั้นคุณภาพของตัวอักษรที่นำมาวิเคราะห์ต้องมีคุณภาพดี ลายเส้นของตัวอักษรจะต้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมบูรณ์ไม่ขาดแหว่งไม่ผิดเพี้ยนจากมาตรฐานของตัวอักษรมากเกินไป รูปแบบของตัวอักษรต้องเป็นรูปแบบเดียวกันกับที่ใช้คำนวณหารหัสแทนตัวอักษรเก็บไว้ในพจนานุกรม สำหรับการรู้จำแล้วไม่ว่าจะเป็นขนาดเล็กหรือขนาดใหญ่ก็ดี ลักษณะตัวปกติหรือตัวเอียงก็ดี อย่างไรก็ตามการเตรียมข้อมูลที่ดีเป็นไปตามเงื่อนไขที่กำหนด ดังที่กล่าวมาข้างต้นยังพอจะทำให้ประสิทธิภาพของการรู้จำสามารถยอมรับได้

2. ข้อเสนอแนะสำหรับแนวทางพัฒนาในอนาคต

งานวิจัยนี้เป็นการวิเคราะห์และการรู้จำชื่อของหนังสือบนสันหนังสือ ซึ่งสามารถนำไปใช้งานได้จริงแต่ว่าการวิเคราะห์และการรู้จำยังมีข้อจำกัดอยู่หลายๆ ส่วนอาทิเช่น ข้อมูลภาพบนสันหนังสือที่นำมาใช้ในการวิเคราะห์และประมวลผลมีลักษณะเป็นข้อมูลภาพที่กำหนดตามเงื่อนไขของการวิจัยนี้เท่านั้น แต่ถ้าข้อมูลภาพสันหนังสือจริงๆ ยังมีส่วนประกอบบนสันหนังสืออีกมากมายหากได้มีการเพิ่มเติมอัลกอริทึมการแยกแยะชื่อหนังสือออกจากสันหนังสือ ก็จะเป็นการเพิ่มประสิทธิภาพให้กับการทำงานในส่วนนี้มากยิ่งขึ้น และในส่วนของการรู้จำตัวอักษรการเพิ่มเติมความสามารถทำข้อมูลภาพสันหนังสือที่ถ่ายจากกล้อง CCD ให้มีความคมชัดขึ้นและการกำจัดสัญญาณรบกวนของภาพข้อมูลสันหนังสือ จะทำให้การรู้จำมีความถูกต้องและแม่นยำสูงมากขึ้น ผู้วิจัยคาดหวังว่า แม้งานวิจัยชิ้นนี้จะไม่สามารถครอบคลุมในทุกสภาวะแวดล้อมของการทำงานได้ แต่ก็น่าจะเป็นแนวทางหรือจุดเริ่มต้นให้กับการพัฒนาต่อๆไปในการค้นคิดหาวิธีการ และเทคนิคเพื่อพัฒนาให้มีความสามารถทำการวิเคราะห์การรู้จำได้มีประสิทธิภาพยิ่งขึ้นซึ่งสามารถนำไปใช้งานได้จริง ต่อไปในอนาคต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

(ACKNOWLEDGEMENT)

ขอกราบขอบพระคุณท่านอาจารย์ที่ปรึกษา ผศ. เกษตร์ ศิริสันติสัมฤทธิ์ เป็นอย่างสูงที่ได้ประสาทวิชาความรู้ และให้คำปรึกษาเสนอแนะแนวทางในการทดลองโครงงาน และวิธีการแก้ปัญหาในการทำงานต่างๆ ตลอดจนอำนวยความสะดวกในเรื่องอุปกรณ์การทำงาน และหนังสือในการค้นคว้า อันเป็นผลทำให้โครงงานและปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี

ขอกราบขอบพระคุณอาจารย์ ทรงชัย วีระทวีมาศ ที่ได้ให้คำปรึกษา และแนะนำแนวทางในการปฏิบัติ

ขอกราบขอบพระคุณ พ่อแม่ และเพื่อนๆ ทุกคนที่ให้กำลังใจ และความสะดวกในการทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หนังสืออ้างอิง

1. นรินทร์ ทัศนลีลพร และ สมศักดิ์ รัตนฤชัตร "การรู้จำรูปแบบอักษรพิมพ์ภาษาไทย (Printed Thai Character Recognition)" ปริญญานิพนธ์ ภาควิชาเทคโนโลยีการวักคุมทางอุตสาหกรรม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ ปีการศึกษา 2539

2. สุรพันธ์ เอื้อไพบูรณ์ "การจดจำตัวอักษรลายมือเขียนภาษาไทยโดยพิจารณาหัวของตัวอักษร" วิทยานิพนธ์มหาบันฑิต คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ ปีการศึกษา 2531

3. สุรสิทธิ์ ราตรี "การรู้จำตัวอักษรตัวพิมพ์ภาษาไทย โดยวิธีค้นหาลักษณะโครงสร้างของลายเส้น (Recognition of Printed Thai Character using)" วิทยานินิพนธ์มหาบันฑิต คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ ปีการศึกษา 2533

4. สมคิด อุนการุณวงษ์ "การจดจำวงจรโลจิกเกทที่เขียนด้วยมือ โดยใช้หลักการแบ่งแยกลายเส้น" บทความของนักศึกษาปริญญาโท คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณ ทหารลาดกระบัง กรุงเทพฯ

5. สมศักดิ์ วลัยรัชต์ "การวิเคราะห์และระบุส่วนประกอบของหน้าเอกสารและการรู้จำตัวอักษร" วิทยานิพนธ์มหาบันฑิต คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ ปีการศึกษา 2537

6. รู้จักกับ CCD อุปกรณ์สำหรับระบบอิมเมจโปรเซสซิ่ง วารสารเซมิคอนดัคเตอร์ เล่มที่ 135 ประจำเดือน มกราคม 2537 หน้า 126-130

7. ขวัญนคร สุขเกษมสิลป์, ธนากร รอดประเสริฐ และ วิเชียร อาจองค์ "การรู้จำรูปแบบอักษรพิมพ์ภาษาอังกฤษ (Printed English Character Recognition)" ปริญญานิพนธ์ คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ ปีการศึกษา 2538

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8. G.J. AWCOCK and R.Thomas, "Applied Image Processing" McGraw-Hill International Edition 1996

9. Jia-Guu Leu and Hok-Lai Yau, "Detecting The Dislocations In Metal Crystals From Microscopic Images", Pattern Recognition, Vol. 24, No.1, pp.41-56, 1991

10. M.A. Sid-Ahmed , "A Hardware Structure For The Automatic Selection of Multi-level Thresholds in Digital Images" , Pattern Recognition, Vol.25, No.12,pp. 1517-1528, 1992

11. Rafael C. Gonzalez and Richard E. Woods , "Digital Image Processing" , Addison-Wesley Publishing Company

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก.

โครงสร้างไฟล์

โครงสร้างของ PCX ไฟล์

ไฟล์ PCX เริ่มต้นด้วยข้อมูล 128 ไบต์ ซึ่งจะทำการเก็บข้อมูลต่าง ๆ เพื่อนำไปใช้ในการ Restore ภาพ ดังรูปที่ 1

ข้อมูล Color Map ถูกใช้ในการแทนค่าเป็น Palette Register 16 ตัว โดยที่ Palette Register 1 ตัว ของระบบแสดงผลอีจีเอจะบรรจุข้อมูล 6 บิตเท่านั้น โดยแต่ละ 2 บิตจะแทนสีหลัก คือ แดง, น้ำเงิน, เขียว ดังรูปที่2 ตัวอักษรตัวใหญ่จะแทนความเข้มของสี 75 % ตัวอักษรเล็กแทน 25 % แต่ละสีหลักจะมี 4 ระดับของการแสดงสีคือ 0 %, 25 %, 75 %, 100 % (ค่าบิตของตัวอักษรตัวใหญ่เป็น '1' หมด) ดังรูปที่ 3 เช่น

- ค่ารีจิสเตอร์เป็น "00100001" หมายถึง แสดงสีแดงที่มีความเข้ม 75 % และ แสดงสีน้ำเงินที่มีความเข้ม 25 %
- ค่ารีจิสเตอร์เป็น "00100101" หมายถึง แสดงสีแดงที่มีความเข้ม 100 % และ แสดงสีหน้ำเงินที่มีความเข้ม 25 %
- ค่ารีจิสเตอร์เป็น "00000000" หมายถึง ไม่แสดงสีใดเลย เป็นต้น

ข้อมูล Color Map ในตอนต้นไฟล์ (Header File) ประกอบด้วยข้อมูล 16 ชุด ๆ ละ 3 ไบต์ ไบต์ที่ 1 ของแต่ละชุด จำนวน 3 ไบต์ย่อยนั้น ๆ เป็นค่าสำหรับแทนสีแดงโดยถูกนำมาสร้างเป็นเลข 4 จำนวนตั้งแต่ 0-3 (00,01,10,11) และตัวเลขนี้จะต้องคูณกับ 85 เพื่อจะทำการเก็บลงไฟล์ ในส่วนเฮดเดอร์ไฟล์ ในกรณีเดียวกันสีเขียวคือไบต์ที่ 2 และสีน้ำเงินคือไบต์ที่ 3 วิธีการนี้จะถูกทำซ้ำ ๆ กัน 16 รอบ เพื่อจะได้ค่าของสีแดง สีน้ำเงิน สีเขียว จนครบ 16 ชุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไบต์ที่	ขนาด (Byte)	ชื่อข้อมูล	รายละเอียด
0	1	Password	ค่า '0ah' = ไฟล์ .PCX
1	1	Version	เก็บค่า Version (รุ่น) ของ PC Paintbrush ดังนี้ 0=Version 2.5 2=Version 2.8 with Palette 3=Version 2.8without Palette 5=Version 3.0
2	1	Encoding	ค่า=1
3	1	Bits per Pixel	จำนวนของบิตที่ใช้แสดง 1 Pixel จาก1 Plane(ระนาบสี) 1 = EGA, VGA, or HERC 4 = CGA
4	8	Windows-Dimensions	ค่า integer 4 ค่า(ค่าละ 2 byte)ให้ค่ามุมบนซ้ายและมุมล่างขวาของภาพกำหนดให้รูปแบบ x1,y1,x2,y2
12	2	Horizontal-Resolution	ความละเอียดของการแสดงภาพในแนวนอน 640 = EGA,VGA 350 = CGA 720 = HERCULES
14	2	Vertical-Resolution	ความละเอียดของการแสดงภายในแนวตั้ง 480 = VGA 350 = EGA 200 = CGA 348 = HERC
16	48	Color Map	ข้อมูล Color Palette ดูรูปที่ 2
64	1	Reserved	
65	1	No. of Plane	จำนวนของ Plane ที่ใช้แสดง
66	2	Byte per Line	จำนวนของไบต์ต่อการสแกน 1 บรรทัด
68	2	Palette Info	How to interpret the palette
70	8	Maximum x value	Special for fractal files
78	8	Minimum x value	Special for fractal files
86	8	Maximum y value	Special for fractal files
94	8	Minimum y value	Special for fractal files
102	8	P Value	Special for fractal files
110	8	Q Value	Special for fractal files
118	10	Not used	

หมายเหตุ ข้อมูลไบต์ที่ 70 - 127 ใช้เฉพาะ Fractal Programming

รูปที่ 1 แสดงตารางโครงสร้าง PCXHEADER

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไบต์ที่	Palette	ไบต์ที่	Palette
16,17,18	0	40,41,42	8
19,20,21	1	43,44,45	9
22,23,24	2	46,47,48	10
25,26,27	3	49,50,51	11
28,29,30	4	52,53,54	12
31,32,33	5	55,56,57	13
34,35,36	6	58,59,60	14
37,38,39	7	61,62,63	15

รูปที่ 2 แสดงตำแหน่งในการเก็บสี จำนวน 3ไบต์ คือ แดง, เขียว, ฟ้า

บิต 7	บิต 6	บิต 5	บิต 4	บิต 3	บิต 2	บิต 1	บิต 0
-	-	R	G	B	r	g	b
x	x	75%	75%	75%	25%	25%	25%

รูปที่ 3 แสดงระดับการแสดงสี 4 ระดับ

<u>การแทนความหมาย Palette</u>

ถ้า Palette Register มีค่า xx100001 แล้ว (ใน Palette ใดๆ (0-15))

บิตที่ 5,2 -> แดง

บิตที่ 4,1 -> เขียว

บิตที่ 3,0 -> น้ำเงิน

แดง คือ '10' = 2 , เขียว '00' = 0, น้ำเงิน '01' = 1

ดังนั้นค่าที่จะเก็บใน Color Map ของ Palette ใดๆ (0-15) คือ (x85)

Red	Green	Blue
170	0	85

หรือถ้า Palette Register มีค่า 'xx111111' แล้ว ค่าที่จะเก็บคือ

Red	Green	Blue
255	255	255

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อแตกต่างพิเศษในจอวีจีเอ

จอภาพวีจีเอจะมีการควบคุมสีที่แตกต่างกับอีจีเอคือ ในแต่ละ Palette Register จะเก็บลำดับที่ของ Color Register ไว้ 1 ตัว และตัวของ Color Register จะใช้เนื้อที่จำนวน 6 บิต ดังนั้นค่าสีที่เป็นไปได้จะเป็น 64 ค่า (หมายถึง Shades ของสีนั้นเอง) ของแต่ละสี ในวีจีเอสามารถอ่านค่า (ค่า 6 บิตของ สีแดง น้ำเงิน เขียว) ข้อมูลใน Palette Register ใน Color Register เพื่อสร้าง Color Map ขึ้นมา โดยอ่านค่าแต่ละ Palette Register และนำไปแอดเดรส Color Register เพื่อทำการอ่าน Color Register อีกที, ทำการคูณสีแดง, น้ำเงิน, เขียว ด้วย 4 และเก็บผลลงในกลุ่ม 3 ไบต์ของ Palette นั้นๆ

ในวีจีเอจะมีสถานะสี 256 สีที่แตกต่างกัน รูปแบบการใช้งานเหมือนกับการแสดง 16 สี ข้อมูลใน Palette จะยาวกว่า 16 สี ข้อมูลนี้จะอยู่ที่ส่วนท้ายของไฟล์ .PCX การเข้าถึงข้อมูลเป็นดังนี้ต้องตรวจไบท์ตรงตำแหน่งเวอร์ชั่นของไฟล์ว่าเป็น 5 (เวอร์ชั่น 3.0) ต้องเป็น 0cH (12 DECIMAL) จากนั้นข้อมูลที่ท้ายไฟล์จะเป็นข้อมูล Palette 256 สี

พื้นฐานในการอ่านข้อมูลเก็บลงไฟล์

ภาพที่จะอ่านจะถูกอ่านตามแนวของจอภาพในแนวนอนจากซ้ายไปขวา เริ่มที่ Pixel ที่ตำแหน่งบน, ซ้ายไปทางขวาจนสุดภาพแล้วจึงอ่านในแถวถัดไปโดยในวีจีเอและอีจีเอ ซึ่งมีหน่วยความจำหลายๆ plane ดังนั้นใน 1 แถว (scanline) จะอ่านข้อมูลในทุก Plane เริ่มที่ plane 0,1,2,3 ตามลำดับจนครบแล้วจึงอ่านแถวถัดไปจนหมดภาพที่ต้องการเก็บ

การเข้ารหัสไฟล์ .PCX

ไฟล์ .PCX มีวิธีการเข้ารหัสที่เรียกว่า Run-Length Encoding (RLE) ซึ่งเป็นวิธีการหนึ่งในการลดขนาดไฟล์ดังรายละเอียดดังนี้

ถ้าไบต์ใดๆ มีค่าไม่เหมือนกับไบต์อื่นๆ และถ้าค่า 2 บิตบนไม่เท่ากับ "11" ไบต์นั้นจะถูกเก็บลงไฟล์เลยนอกนั้นจะมีตัวนับ (Counter) อยู่คอยนับว่าเหมือนกันกี่ไบต์ แต่ต้องไม่เกิน 63 ถ้าเกินให้นำค่าที่ได้ OR กับ C0h แล้วเก็บค่าตัวนับนั้นลงในไฟล์ แล้วตามด้วยค่าของไบต์นั้นๆ จากนั้นจึงเริ่มนับใหม่เป็น 1 ต่อไป ถ้าในกรณีของไบต์เดียวที่มีค่าบิตบนเป็น "11" ให้เขียนตัวนับ = 1 (OR with C0h ='C1h') ลงไปในไฟล์และตามด้วยค่า ไบต์นั้น ดังแสดงในรูปที่ 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เงื่อนไข	การกระทำ
ค่าที่เหมือนไบต์อื่น	นับ Counter = 1
ค่า < 'C0h' ไม่เหมือนไบต์อื่น	เก็บลงไฟล์เลย
ค่า > 'C0h' ไม่เหมือนไบต์อื่น	เก็บ 'C1h' ลงไฟล์ตามด้วยค่าไบต์นั้น
Counter > 63	Counter=1, เก็บค่า 'FFh' และค่าไบต์ลงไฟล์

รูปที่ 4 แสดงตารางเงื่อนไขและการกระทำ

<u>การถอดรหัสไฟล์ .PCX</u>

การถอดรหัสใช้วิธีการตรงกันข้ามกับการเข้ารหัส พอสรุปได้ดังนี้

อ่านข้อมูลมา 1 ไบต์ แล้วตรวจดูว่ามากกว่า 'C0h' หรือไม่ (ค่า 2 บิตบนเป็น 11) ถ้าใช่ค่านี้จะเป็นค่าตัวนับทันที (XOR ด้วย C0h จะได้ค่าจริงที่น้อยกว่า 63 ออกมา) ซึ่งไบต์ต่อไปก็จะเป็นค่าของไบต์ (value) แล้วทำการขยายข้อมูลออกมาตามจำนวนตัวนับนั้นๆ แล้วเก็บในหน่วยความจำ ถ้าค่าไบต์ที่อ่านน้อยกว่า C0h จะทำการเก็บข้อมูลไบต์นั้นตัวเดียว ลงหน่วยความจำที่มีอยู่ได้เลยโดยไม่ต้องขยายข้อมูล ทำเช่นนี้จนจบไฟล์ก็จะได้ข้อมูลของไฟล์ทั้งหมด จากนั้นนำข้อมูลในหน่วยความจำที่ได้มาแสดงบนจอภาพ โดยใช้ฟังก์ชั่น POKEB( ) ในการนำเอาข้อมูลไปไว้ในหน่วนความจำแสดงผล (display memory) หรืออาจจะใช้ฟังก์ชั่น PUTPIXEL( ) แสดงภาพออกมาก็ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข.

ตารางแทนรหัส ASCII Code

	0	1	2	3	4	5	6	7
0	NUL	DLE	SP	0	@	P	‘	p
1	SOH	DC1	!	1	A	Q	a	q
2	SXT	DC2	“	2	B	R	b	r
3	EXT	DC3	#	3	C	S	c	s
4	EOT	DC4	$	4	D	T	d	t
5	ENG	NAK	%	5	E	U	e	u
6	ACK	SYN	&	6	F	V	f	v
7	BEL	ETB	‘	7	G	W	g	w
8	BS	CAN	(	8	H	X	h	x
9	HT	EM	)	9	I	Y	i	y
A	LF	SUB	*	:	J	Z	j	z
B	VT	ESC	+	;	K	[	k	{
C	FF	FS	,	<	L	\	l	|
D	CR	GS	-	=	M	]	m	}
E	SO	RS	.	>	N	^	n	~
F	SI	US	/	?	O	_	o	DEL

หมายเหตุ

ตารางรหัส ASCII ภาษาอังกฤษใช้รหัสตามมาตรฐาน ISO 646-1983

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้