กองห้องสมุดกลาง พระจอมเกล้าลาดกระบัง

การรู้จำอักษรตัวพิมพ์ภาษาไทย โดยวิธีค้นหาลักษณะโครงสร้างของลายเส้น

Recognition Of Printed Thai Character Using topological Feature Extraction

นายสุรสิทธิ์ ราตรี

Mr.Surasit Ratree

อาจารย์ที่ปรึกษา

รองศาสตราจารย์ ดร.ชม กิ้มปาน

ADVISOR

ASSOCIATE PROFESSOR Dr.Chom Kimpan

วิทยานิพนธ์สำหรับปริญญาวิศวกรรมศาสตร์มหาบัณฑิต

สาขาวิศวกรรมไฟฟ้า

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2533

DISSERTATION FOR THE DEGREE OF MASTER OF ENGINEERING

IN ELECTRICAL ENGINEERING

KING MONGUT'S INSTITUTE OF TECHNOLOGY

CHAOKHUN TAHARN LADKRABANG

ACADEMIC YEAR 1990

ISBN 974-8155-18-8

เลขหมู่
เลขทะเบียน 15939
วัน, เดือน, ปี 19 มิ.ย. 2534

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทคัดย่อ

เรามักจะพบเห็นตัวพิมพ์อักษรภาษาไทย ในลักษณะรูปแบบและขนาดต่าง ๆ ในสิ่งตีพิมพ์ทั่วไป ดังนั้นหากมีการนำวิธีการรู้จำรูปแบบตัวพิมพ์อักษรภาษาไทยของคอมพิวเตอร์มาใช้เก็บข้อมูลเข้าคอมพิวเตอร์อัตโนมัติจำเป็นจะต้องมีความสามารถในการวิเคราะห์อักษรลักษณะรูปแบบและขนาดต่าง ๆ กันได้

วิทยานิพนธ์ฉบับนี้ได้เสนอวิธีที่เหมาะสมวิธีหนึ่ง คือการค้นหาลักษณะโครงสร้างลายเส้น (Topological Feature Extraction) สำหรับการวิเคราะห์ในขบวนการรู้จำอักษรตัวพิมพ์ภาษาไทย โดยใช้เครื่องตรวจกวาดภาพ (Scanner) ทดสอบอ่านข้อมูลภาพอักษรตัวพิมพ์ภาษาไทยทั้งหมด 15 รูปแบบที่มีขนาดต่าง ๆ กัน 5 ขนาด (ดูภาคผนวก ค.) ในการรู้จำของวิธีการนี้ประกอบด้วย ภาคจัดการล่วงหน้าอาศัยเทคนิคการแยกภาพตัวอักษรออกจากภาพประโยค ด้วยวิธีการหาขอบภาพของตัวอักษร (Edge Detection) ส่วนภาคการรู้จำได้แบ่งการพิจารณาตัวอักษรออกเป็น 8 ส่วน แต่ละส่วนจะถูกค้นหาลักษณะโครงสร้างลายเส้นของตัวอักษร เพื่อให้ได้ชุดรหัสเฉพาะของอักษรแต่ละตัว แต่เพื่อความสะดวกในการหารหัส และเพื่อไม่ให้ฐานความรู้มีขนาดโตเกินไป ได้จัดแบ่งกลุ่มอักษร(Rough classification)ออกเป็นกลุ่ม ๆ ทั้งหมด 7 กลุ่ม โดยใช้ค่ารหัสลักษณะเด่นของตัวอักษร (Concentrate Code) ณ. ตำแหน่งจุดศูนย์ถ่วงของตัวอักษรเพื่อจัดแบ่งกลุ่มตัวอักษร ปรากฏว่าตัวอักษรส่วนใหญ่อยู่ที่กลุ่ม F (หน้า 47) ส่วนในการแยกตัวอักษรออกจากกลุ่ม (Fine classification) ซึ่งใช้วิธีหารหัสจากลักษณะโครงสร้างลายเส้นทั้ง 8 ส่วนของรูปร่างตัวอักษรแต่ละส่วน จะถูกประมาณให้เป็นเส้นตรงเป็นส่วนใหญ่ จึงทำให้อักษรบางกลุ่มเมื่อประมาณค่าแล้ว ลักษณะส่วนโค้งบางส่วนจะหายไป จำเป็นต้องใช้เทคนิคเฉพาะมาใช้พิจารณาเฉพาะส่วนเพื่อนำไปสร้างเป็นตารางการตัดสินใจประจำกลุ่มอักษรนั้น

สรุปผลการทดสอบ ในทุกขั้นตอนได้ใช้ข้อมูลอักษรตัวพิมพ์จากโรงพิมพ์ ที่มีลักษณะรูปร่างแยกต่างกัน 15 รูปแบบที่ขนาดต่างกัน 5 ขนาด จากจำนวนตัวอักษร 9,750 ตัว สามารถวิเคราะห์ตัวอักษรได้ถูกต้องร้อยละ 95.5 ถูกตัดทิ้งไปร้อยละ 4 และวิเคราะห์ผิดพลาดร้อยละ 0.5 ด้วยความเร็วในการวิเคราะห์เฉลี่ยประมาณตัวละ 1 วินาที โปรแกรมที่ใช้ในขบวนการรู้จำเขียนขึ้นจากภาษาแอ็ดแซมบลิ้นเครื่อง IBM Compatible ทำงานที่ความเร็ว 16 MHz

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Abstract

Printed Thai characters in general printed matter usually have multifont and any size. Thus, The pattern recognition of printed Thai characters for computer should provide for solving this feature.

In the case study submitted to this thesis, the system of pattern recognition of Printed Thai characters by Topology Feature Exaction is presented.(Annex ก)Printed Thai characters in 15 multifonts at 5 sizes are read by scanner. The methodology contains two parts, preprocessing process and recognition process. Edge detection, technique of segmentation characters from sentences is used in the preprocessing process.The recognition process shows two important steps, Rough Classification and Fine Classification. All character is grouped in 7 groups (Rough classification) by concentrate code from its center of gravity in each character. The result shows that most of the characters is classified in group F.(page 47) Topological feature from 8 octants (Fine Classification) is searched. Since the code from each octant is approximated to be straight line, i.e. curves in some group of character disappeared, special technique are used in order to make the decision table for each group of charactor.

Printed Thai characters for testing the system are contributed from 15 diffference patterns from newspapers.Total 9,750 characters show accurate rate 95.5%, reject 4% and error 0.5% with an average time taken of 1 second per character. Language used in pattern recognition is Assembly on IBM compatible with clock at 16 MHz.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อ	ก
Abstract	ข
สารบัญ	ค
สารบัญรูปประกอบ	ง
สารบัญตารางประกอบ	จ
บทที่ 1 บทนำ	1
บทที่ 2 วัตถุประสงค์ และผลงานวิจัยที่เกี่ยวข้อง	4
2.1 วัตถุประสงค์ของวิทยานิพนธ์	4
2.2 ผลงานวิจัยที่เกี่ยวข้อง	4
2.2.1 การจดจำรูปแบบตัวพิมพ์อักษรภาษาไทย	4
2.2.2 การจดจำลายมือเขียนตัวอักษรภาษาไทยในระบบออนไลน์	5
2.2.3 การจดจำตัวอักษรลายมือเขียนภาษาไทยโดยการพิจารณาหัวของตัวอักษร	6
2.2.4 การจดจำรูปแบบตัวอักษรคัดลายมือภาษาไทยโดยวิธีแยกลักษณะเด่น	6
2.2.5 การออกแบบพจนานุกรมสำหรับการเรียนรู้อักษรคัดลายมือไทย-อังกฤษอัตโนมัติบนเครื่องไมโครคอมพิวเตอร์	7
2.3 ลักษณะของวิทยานิพนธ์ที่แตกต่างจากผลงานวิจัยที่ผ่านมา	7
2.3.1 วิธีแยกภาพตัวอักษรออกจากภาพประโยค	7
2.3.2 ลักษณะของตัวอักษรที่ใช้วิเคราะห์	8
2.3.3 เทคนิควิธีการ	8
2.4 เนื้อหาวิทยานิพนธ์	8
บทที่ 3 ทฤษฎีที่นำมาใช้	11
3.1 หลักการทำงานของเครื่อง Scanner	11
3.1.1 เครื่อง Scanner	11
3.1.2 เทคนิคของเครื่อง Scanner	11

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

หน้า

3.2 ทฤษฎีเฟียเจอร์คอนเซนเทร 14
3.2.1 การหาค่า Q-Code 14
3.2.2 การกำหนดค่า Q-Code ของตัวอักษร 15
3.3 ทฤษฎีโทโบโลจิคอลเฟียเจอร์เอ็กแทร็กชั่น 17
3.4 การหาจุดศูนย์ถ่วง 30
3.5 การหาขอบภาพ 31
3.5.1 ตัวอย่างการคำนวณของภาพ พิจารณาที่เป็นขอบ 31
3.5.2 ตัวอย่างการคำนวณขอบภาพ พิจารณาที่ไม่ใช่ขอบ 32
3.6 การหาขนาดความกว้างและความสูงของตัวอักษร 34
บทที่ 4 การจัดเตรียมข้อมูลเพื่อการรู้จำ 36
4.1 การจัดเก็บข้อมูลภาพ 36
4.2 ลักษณะเฉพาะของข้อความที่เครื่อง Handy Scanner สามารถอ่านได้ 37
4.3 เทคนิคการแยกภาพตัวอักษรออกจากประโยค 38
4.3.1 การทำขอบภาพ 38
4.3.2 การคัดลอกและย้ายจุดพิกัดตัวอักษร 39
4.3.3 การจัดส่งข้อมูลไปภาควิเคราะห์ 41
4.3.4 การจัดเรียงรูปประโยคหลังการวิเคราะห์ 42
บทที่ 5 การจัดกลุ่มตัวอักษร 45
5.1 การให้รหัสกลุ่มตัวอักษร 45
5.2 กลุ่มตัวอักษร 46
บทที่ 6 การแยกตัวอักษรออกจากกลุ่ม 48
6.1 การวิเคราะห์แต่ละตัวอักษร 48
6.2 การย่อรหัส 53
6.3 การวิเคราะห์เฉพาะส่วนของตัวอักษร 54
6.3.1 กรณีชุด A-Code เหมือนกันของอักษร ด และ ต 54
6.3.2 กรณีชุด A-Code เหมือนกันของอักษร ' และ ' 56
6.3.3 กรณีชุด A-Code เหมือนกันของเลข 4 และ 0 58

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

หน้า

บทที่ 7 โปรแกรมการรู้จำรูปแบบด้วพิมพ์อักษรภาษาไทย 64
7.1 โปรแกรมสำเร็จรูปควบคุมการอ่านภาพจากเครื่อง Scanner 64
7.2 โปรแกรมที่พัฒนาขึ้นสำหรับวิเคราะห์การรู้จำอักษรภาษาไทย 64
7.2.1 คุณสมบัติของโปรแกรม 66
7.2.2 โปรแกรมและ File ที่เกี่ยวข้อง 66
7.2.3 การใช้งานของโปรแกรม 67
7.2.4 อุปกรณ์ที่ใช้ประกอบ 74
บทที่ 8 การทดสอบโปรแกรม 75
8.1 การรู้จำให้กับโปรแกรม (Learning) 75
8.2 การทดสอบอักษรจากบทความกับอักษรต้นแบบที่มีรูปแบบเดียวกัน 75
8.3 การทดสอบกลุ่มอักษรหลายรูปแบบและหลายขนาดกับอักษรต้นแบบ 76
8.4 การทดสอบบทความกับอักษรต้นแบบหลายรูปแบบและหลายขนาด 78
บทที่ 9 บทสรุปและข้อเสนอแนะ 81
9.1 เปรียบเทียบผลงานวิจัยที่เกี่ยวข้อง 82
9.2 การแยกภาพอักษรออกจากภาพประโยค 82
9.3 การจัดและแยกกลุ่มอักษร 82
9.3.1 ขั้นการจัดกลุ่ม 82
9.3.2 ขั้นการแยกอักษรออกจากกลุ่ม 83
9.4 ผลการทดลอง 84
9.5 ข้อเสนอแนะ 84
กิตติกรรมประกาศ 85
เอกสารอ้างอิง 86
ภาคผนวก ก ตารางรหัส ASCII CODE 89
ภาคผนวก ข การใช้งานโปรแกรมควบคุม Handy Scanner 90
ภาคผนวก ค ตัวอักษรและข้อความรูปแบบต่าง ๆ 93

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปประกอบ

หน้า

รูปที่ 1 ก. ตัวอย่างลักษณะการใช้งานระบบการรู้จำตัวพิมพ์อักษรภาษาไทย 2
รูปที่ 1 ข. ลำดับการทำงานของระบบการรู้จำอักษรภาษาไทย 3
รูปที่ 2. ขบวนการรู้จำตัวพิมพ์อักษรภาษาไทยหลายรูปแบบและหลายขนาด 10
รูปที่ 3. แสดงลักษณะของเครื่อง Scanner 11
รูปที่ 4. แสดงหลักการทำงานของเครื่อง Scanner 12
รูปที่ 5. ภาพแสดงลักษณะข้อมูลที่ได้จากอ่านด้วยเครื่อง Scanner 13
รูปที่ 6. ลักษณะทิศทางที่ใช้พิจารณาการหาค่ารหัส Feature Concentrant 14
รูปที่ 7. แสดงการแทนค่ารหัส Q-Code อักษร ก. 16
รูปที่ 8. แสดงการแทนค่ารหัส Q-code อักษร ม. 16
รูปที่ 9. ลักษณะทิศทางทั้ง 8 ซึ่งพื้นที่ถูกแบ่งเป็น 8 Octant 17
รูปที่ 10 ก.,รูปที่ 10 ข.,รูปที่ 10 ค.,รูปที่ 10 ง. ลักษณะของรหัส a_0 19
รูปที่ 11.ก.,รูปที่ 11 ข.,รูปที่ 11 ค.,รูปที่ 11 ง. ลักษณะของรหัส a_1 20
รูปที่ 12.ก.,รูปที่ 12 ข.,รูปที่ 12 ค.,รูปที่ 12 ง. ลักษณะของรหัส a_2 20
รูปที่ 13.ก.,รูปที่ 13 ข.,รูปที่ 13 ค.,รูปที่ 13 ง. ลักษณะของรหัส a_3 21
รูปที่ 14.ก.,รูปที่ 14 ข.,รูปที่ 14 ค.,รูปที่ 14 ง. ลักษณะของรหัส a_4 21
รูปที่ 15.ก.,รูปที่ 15 ข.,รูปที่ 15 ค.,รูปที่ 15 ง. ลักษณะของรหัส a_5 22
รูปที่ 16.ก.,รูปที่ 16 ข.,รูปที่ 16 ค.,รูปที่ 16 ง. ลักษณะของรหัส a_6 22
รูปที่ 17.ก.,รูปที่ 17 ข.,รูปที่ 17 ค. ลักษณะของรหัส a_7 23
รูปที่ 18.ก.,รูปที่ 18 ข.,รูปที่ 18 ค. ลักษณะของรหัส a_8 23
รูปที่ 19.ก.,รูปที่ 19 ข.,รูปที่ 19 ค.,รูปที่ 19 ง. ลักษณะของรหัส a_9 24
รูปที่ 20. แผนผังแสดงวิธีการหาค่ารหัส A-Code จาก Octant 28
รูปที่ 21. ตัวอย่างการหารหัสชุด A-Code จากลักษณะStorkeของอักษร ง. 29
รูปที่ 22. ภาพแสดงจุดศูนย์ถ่วงของตัวอักษร ก.ซึ่งคำนวณได้จากสมการ 30
รูปที่ 23. ภาพตัวอย่างจุดที่พิจารณาบริเวณที่เป็นขอบ 31
รูปที่ 24. ภาพตัวอย่างจุดที่พิจารณาบริเวณที่ไม่ใช่ขอบ 32
รูปที่ 25. แสดงลักษณะภาพที่ยังไม่ถูกคำนวณหาขอบภาพ 33
รูปที่ 26. แสดงลักษณะผลการคำนวณขอบภาพของตัวอักษร 33
รูปที่ 27. ลักษณะจุด Coordinate ของตัวอักษรที่ถูกนำไปคำนวณ หาค่าความกว้างและความสูง 35

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปประกอบ (ต่อ)

หน้า

รูปที่ 28. ลักษณะรูปร่างของเครื่อง Handy Scanner 36
รูปที่ 29. ภาพแสดงลักษณะการใช้งานของเครื่อง Handy Scanner 37
รูปที่ 30. ลักษณะข้อความที่เครื่อง Handy Scanner สามารถอ่านได้ 37
รูปที่ 31. ภาพประโยคที่ปรากฏในจอมอนิเตอร์ที่อ่านเข้ามาด้วยเครื่อง Scanner 38
รูปที่ 32. ผลการหาขอบภาพของภาพประโยค 39
รูปที่ 33 ก. ลักษณะการกวาดหาพิกัดจุดภาพเพื่อคัดลอก 40
รูปที่ 33 ข. การค้นหาจุดภาพอักษรตัวถัดไป 40
รูปที่ 34. ตัวอย่างลักษณะข้อมูลภาพตัวอักษรที่ถูกคัดลอกแยกออกจากภาพประโยคเตรียมส่งให้ภาควิเคราะห์การรู้จำต่อไป 41
รูปที่ 35 รูปตำแหน่งจุดพิกัดกึ่งกลางฐานตัวอักษร 42
รูปที่ 36. ตัวอย่างกลุ่มอักษรที่ตำแหน่งจุดศูนย์ถ่วงมีรหัส Q-Code เป็นรหัส E เหมือนกัน 45
รูปที่ 37. ตัวอย่างกลุ่มอักษรที่ตำแหน่งจุดศูนย์ถ่วงมีรหัส Q-Code เป็นรหัส B เหมือนกัน 46
รูปที่ 38. ภาพตัวเลขและตัวอักษรที่ใช้ตรวจสอบหารหัส Q-Code ที่จุดศูนย์ถ่วง 46
รูปที่ 39 ก. แสดงตัวอย่างการวิเคราะห์แยกอักษร ก ออกจากกลุ่ม E 48
รูปที่ 39 ข. แสดงตัวอย่างการวิเคราะห์แยกอักษร ภ ออกจากกลุ่ม E 49
รูปที่ 40. เปรียบเทียบส่วนที่แตกต่างกันของอักษร ด และ ต 54
รูปที่ 41. บริเวณส่วนครึ่งบนของอักษร ด และ ต ที่เลือกพิจารณาหาค่าชุดรหัส A-Code ใหม่ 55
รูปที่ 42. การหาค่าชุดรหัส A-Code บริเวณครึ่งบนของอักษร ด และ ต 55
รูปที่ 43. เปรียบเทียบลักษณะที่แตกต่างกันของอักษร ◌่ และ ◌๋ 57
รูปที่ 44. ลักษณะที่ถูกพิจารณาเป็น Storke เดี่ยวไม่มี Storke คู่ขนานตามเงื่อนไขในทฤษฎีโทโบโลจิคอลเฟียเจอร์เอ็กแทร็กชั่น 57
รูปที่ 45. บริเวณส่วนครึ่งบนของอักษร ◌่ และ ◌๋ ที่เลือกพิจารณาหาค่าชุดรหัส A-Codeใหม่ 57
รูปที่ 46. การหาค่าชุดรหัส A-Code บริเวณครึ่งบนของอักษร ◌่ และ ◌๋ 58

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปประกอบ (ต่อ)

หน้า

รูปที่ 47. เปรียบเทียบลักษณะที่แตกต่างกันของตัวเลข 4 และ 0 59
รูปที่ 48. ลักษณะที่ถูกพิจารณาหาค่ารหัส A-Code จาก Octant ทั้ง 8 59
รูปที่ 49. บริเวณส่วนครึ่งบนของตัวเลข 4 และ 0 ที่เลือกพิจารณา หาค่าชุดรหัส A-Codeใหม่ 60
รูปที่ 50. การหาค่าชุดรหัส A-Code บริเวณครึ่งบนของตัวเลข 4 และ 0 60
รูปที่ 51. ภาพแสดงสรุปการแบ่งกลุ่มอักษร 61
รูปที่ 52. ขบวนการวิเคราะห์การรู้จำรูปแบบอักษรภาษาไทย 62
รูปที่ 53. แสดงภาพการใช้งานของโปรแกรมที่พัฒนาขึ้น 65
รูปที่ 54. แสดงตัวอย่างการใช้งานฟังค์ชั่น "Input file." 68
รูปที่ 55. แสดงตัวอย่างการใช้งานฟังค์ชั่น "Output file." 68
รูปที่ 56 ก. แสดงตัวอย่างการใช้ฟังค์ชั่น "Switch to grahics mode" 69
รูปที่ 56 ข. ตัวอย่างการแสดงลักษณะภาพกลุ่มข้อความที่ถูกเลือกจากฟังค์ชั่นนี้ 69
รูปที่ 57. ภาพตัวอย่างหน้าจอขณะกำลังวิเคราะห์ตัวอักษรเมื่อได้เลือกฟังค์ชั่น "Recognition in graph mode: เป็น "ON" 70
รูปที่ 58. ภาพที่ปรากฏขณะทำการวิเคราะห์ตัวอักษร โดยเลือกฟังค์ชั่น "Start recognition" 71
รูปที่ 59. ภาพตารางรหัส ASCII CODE ที่ปรากฏขึ้นกรณีที่ขบวนการไม่สามารถวิเคราะห์ได้ โดยให้ผู้ใช้ตัดสินใจเลือกรหัสเอง 72
รูปที่ 60. ภาพที่ปรากฏเมื่อกดคีย์ F1 เพื่อพิจารณาลักษณะรูปร่างอักษรที่ไม่สามารถวิเคราะห์ได้ ก่อนกลับไปตัดสินใจเลือกรหัส ASCII 73
รูปที่ 61. ภาพบทความที่ใช้ทดสอบ มีรูปแบบและขนาดเดียวกันกับอักษรต้นแบบ 75
รูปที่ 62. ภาพอักษรที่มีรูปแบบและขนาดแตกต่างกันจำนวน 14 ชุด 76
รูปที่ 63. ภาพตัวอย่างบางส่วนของบทความทั้ง 15 ชุด 78

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญประกอบตาราง

หน้า

ตารางที่ 1. แสดงลักษณะของ Q-Code Table 15
ตารางที่ 2. แสดงตัวอย่างการพิจารณาลักษณะ Storke ในแต่ละ Octant 25
ตารางที่ 3. ตารางเมตริกซ์ 3x3 31
ตารางที่ 4. กลุ่มตัวอักษรที่ถูกจัดแยกกลุ่มด้วยรหัส Q-Code 47
ตารางที่ 5. เปรียบเทียบชุดรหัส A-Code ของกลุ่ม E 49
ตารางที่ 6. เปรียบเทียบชุดรหัส A-Code ของกลุ่ม B 50
ตารางที่ 7. เปรียบเทียบชุดรหัส A-Code ของกลุ่ม C 50
ตารางที่ 8. เปรียบเทียบชุดรหัส A-Code ของกลุ่ม D 51
ตารางที่ 9. เปรียบเทียบชุดรหัส A-Code ของกลุ่ม F 51
ตารางที่ 10. ชุดรหัส A-Code เฉพาะส่วนครึ่งบนของอักษร ด และ ต 56
ตารางที่ 11. ชุดรหัส A-Code เฉพาะส่วนครึ่งบนของอักษร ' และ + 58
ตารางที่ 12. ชุดรหัส A-Code เฉพาะส่วนครึ่งบนของตัวเลข 4 และ 0 60
ตารางที่ 13. เปรียบเทียบผลการทดสอบอักษรรูปแบบต่าง ๆ ทั้ง 14 ชุด 77
ตารางที่ 14. เปรียบเทียบผลการทดสอบบทความอักษร รูปแบบต่างกัน 15 ชุด 79
ตารางที่ 15. สรุปผลการทดสอบ 80
ตารางที่ 16. เปรียบเทียบผลงานวิจัยที่เกี่ยวข้อง 81
ตารางที่ 17. เปอร์เซนต์เวลาในการแยกตัวอักษรออกจากภาพประโยค 82
ตารางที่ 18. กลุ่มรหัส Q-Code ที่เพิ่มจากเดิม 83

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

ปัจจุบัน แม้ว่าความสามารถของเครื่องคอมพิวเตอร์จะได้ถูกพัฒนา และมีการประยุกต์ใช้งานในวงการต่าง ๆ ทุกสาขาอาชีพ แต่ไม่ว่าคอมพิวเตอร์จะถูกนำไปใช้สำหรับงานลักษณะใดก็ตาม ผู้ใช้เครื่องจะติดต่อกับเครื่องได้ก็โดยการป้อนข้อมูลแบบใดแบบหนึ่งเช่น Key to Card , Key to Disk หรือ Key to Magnetic Tape เป็นต้น การใช้แป้นพิมพ์ อาจเป็นวิธีหนึ่งที่สะดวกสำหรับผู้ที่มีความชำนาญการใช้แป้นกด แต่สำหรับผู้ที่ยังไม่ชำนาญโอกาสผิดพลาดย่อมเกิดขึ้นได้ง่าย ปัญหาการป้อนข้อมูลจะลำบากมากขึ้นและถ้าข้อมูลที่จะป้อนเหล่านั้นมีจำนวนมากขึ้น ก็จำเป็นต้องใช้กำลังคน ระยะเวลา รวมทั้งงบประมาณค่าใช้จ่ายมากขึ้นตามด้วย

เนื่องจากการป้อนข้อมูลจากเอกสารต่างๆ เข้าเครื่องคอมพิวเตอร์เป็นสิ่งที่หลีกเลี่ยงไม่ได้ดังกล่าวแล้ว เราจึงพยายามหาวิธีการป้อนข้อมูลที่สะดวกรวดเร็วและประหยัดทั้งแรงงานและงบประมาณได้ และที่สำคัญกว่านั้นคือ จะต้องถูกต้องแม่นยำไม่มีที่ผิดพลาดอีกด้วย จึงมีผู้คิดค้นวิธีพัฒนาเครื่องคอมพิวเตอร์ ให้สามารถอ่านตัวอักษรจากเอกสารต่าง ๆ ได้เองแทนการใช้แป้นกด แบบเดิมป้อนข้อมูล เราเรียกวิธีการป้อนข้อมูลแบบใหม่นี้ว่า Page to Disk โดยการใช้เครื่องอ่านตัวอักษรด้วยแสง (Optical Character Read หรือ OCR) สำหรับเปลี่ยนภาพรูปร่างอักษรตัวพิมพ์ต่าง ๆ จากเอกสารให้เป็นสัญญาณทางไฟฟ้าที่เครื่องคอมพิวเตอร์สามารถวิเคราะห์และวินิจฉัยสัญญาณไฟฟ้านั้นได้ว่าเป็นตัวอักษรอะไร วิธีการนี้นอกจากจะแก้ปัญหาการป้อนข้อมูลจำนวนมากเข้าเครื่องแทนการกดแป้นแล้ว อาจประยุกต์ใช้งานที่เป็นประโยชน์ด้านอื่น ๆ ได้อีกเช่น เป็นเครื่องอ่านหนังสือสำหรับคนตาบอดหรือ เป็นเครื่องตรวจสอบเอกสารต่าง ๆ เป็นต้น

ลักษณะการทำงานของเครื่องคอมพิวเตอร์ ที่จะทำให้สามารถอ่านตัวพิมพ์อักษรภาษาไทยได้นั้น มีวิธีการดังนี้คือ เมื่อป้อนแผ่นกระดาษเอกสารใด ๆ ที่มีข้อความอักษรตัวพิมพ์ภาษาไทยเข้าเครื่อง OCR ภาพของตัวอักษรเหล่านั้นจะถูกแปลงไปเป็นสัญญาณไฟฟ้าและถูกจัดเก็บในลักษณะค่าเลขดิจิตอล ลงในหน่วยความจำของเครื่องคอมพิวเตอร์ จากนั้นคอมพิวเตอร์จะแยกภาพอักษรออกจากประโยคทีละตัวโดยอาศัยช่องว่างระหว่างตัวอักษรแต่จะยังไม่ทราบว่าอักษรตัวที่ถูกแยกออกไปนั้นคือตัวอักษรอะไร ขั้นตอนนี้จะเรียกว่าเป็นภาคการจัดการล่วงหน้า (pre processing) ขั้นตอนต่อมาคอมพิวเตอร์จะต้องวิเคราะห์วินิจฉัยให้ได้ว่าอักษรตัวนั้นคือตัวอักษรอะไร ในขั้นตอนหลังนี้เรียกว่าเป็นขบวนการรู้จำ[1](recognition) ตัวพิมพ์อักษรภาษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไทยซึ่งถือว่าเป็นขั้นตอนที่สำคัญยิ่ง เนื่องจากถ้าวิธีการรู้จำไม่สมบูรณ์พอจะทำให้การวิเคราะห์ผิดพลาดได้ง่าย เมื่อวิเคราะห์ได้แล้วว่าเป็นอักษรอะไรก็จะส่งระหัส ASCII ภาษาไทย ของอักษรตัวนั้น เก็บไว้ในแผ่นเก็บข้อมูลหรือพิมพ์ผลลัพธ์ออกทางจอภาพ ดังภาพตัวอย่างลักษณะการใช้งานในรูป 1ก. และลำดับการทำงานของระบบดังแสดงในรูป 1ข.

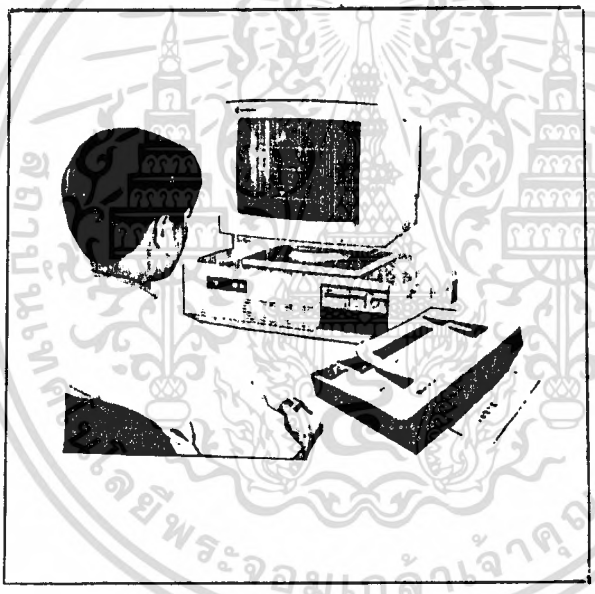
ปัญหาที่ตามมาสำหรับขบวนการรู้จำตัวพิมพ์อักษรภาษาไทยก็คือ ลักษณะของตัวอักษรไทยซึ่งมีหลายรูปแบบ ตัวอย่างที่เห็นชัดคือหนังสือพิมพ์ จะมีลักษณะอักษรรูปแบบแตกต่างกันในฉบับเดียวกัน ทั้งตัวตรง ตัวเอียง ตัวหนา ตัวบางรวมทั้งรูปแบบของตัวอักษรที่มีผู้ออกแบบขึ้นใช้เฉพาะงานเช่น ตัวพิมพ์อักษรลายไทย และรูปแบบตัวพิมพ์อักษรลวดลายอื่นๆทั้งนี้ยังประกอบด้วยหลายขนาดในรูปแบบเดียวกัน ขบวนการจดจำรูปแบบตัวพิมพ์อักษรเหล่านั้นจำเป็นต้องหาเทคนิคที่จะสามารถรู้จำรูปแบบของตัวอักษร ที่แตกต่างกันหลายรูปแบบและหลายขนาดได้

รูป 1ก. ตัวอย่างลักษณะการใช้งานระบบการรู้จำตัวพิมพ์อักษรภาษาไทย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

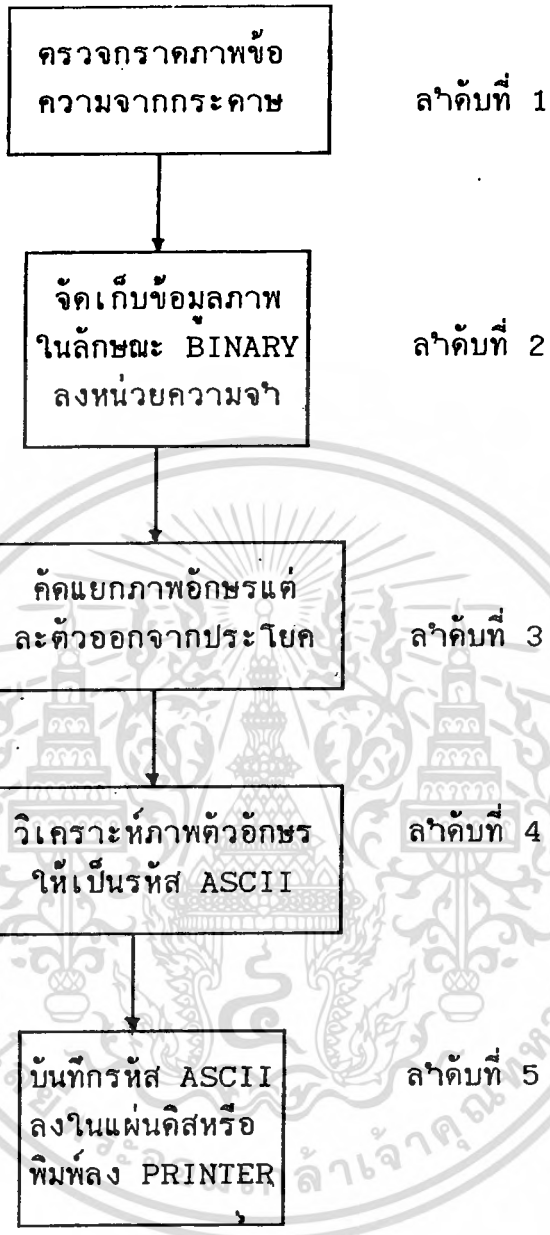
ไทยซึ่งถือว่าเป็นขั้นตอนที่สำคัญยิ่ง เนื่องจากถ้าวิธีการรู้จำไม่สมบูรณ์พอจะทำให้การวิเคราะห์ผิดพลาดได้ง่าย เมื่อวิเคราะห์ได้แล้วว่าเป็นอักษรอะไรก็จะส่งระหัส ASCII ภาษาไทย ของอักษรตัวนั้น เก็บไว้ในแผ่นเก็บข้อมูลหรือพิมพ์ผลลัพธ์ออกทางจอภาพ ดังภาพตัวอย่างลักษณะการใช้งานในรูป 1ก. และลำดับการทำงานของระบบดังแสดงในรูป 1ข.

ปัญหาที่ตามมาสำหรับขบวนการรู้จำตัวพิมพ์อักษรภาษาไทยก็คือ ลักษณะของตัวอักษรไทยซึ่งมีหลายรูปแบบ ตัวอย่างที่เห็นชัดคือหนังสือพิมพ์ จะมีลักษณะอักษรรูปแบบแตกต่างกันในฉบับเดียวกัน ทั้งตัวตรง ตัวเอียง ตัวหนา ตัวบางรวมทั้งรูปแบบของตัวอักษรที่มีผู้ออกแบบขึ้นใช้เฉพาะงานเช่น ตัวพิมพ์อักษรลายไทย และรูปแบบตัวพิมพ์อักษรลวดลายอื่นๆทั้งนี้ยังประกอบด้วยหลายขนาดในรูปแบบเดียวกัน ขบวนการจดจำรูปแบบตัวพิมพ์อักษรเหล่านั้นจำเป็นต้องหาเทคนิคที่จะสามารถรู้จำรูปแบบของตัวอักษร ที่แตกต่างกันหลายรูปแบบและหลายขนาดได้



รูป 1ก. ตัวอย่างลักษณะการใช้งานระบบการรู้จำตัวพิมพ์อักษรภาษาไทย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 1ข. ลำดับการทำงานของระบบการรู้จำอักษรภาษาไทย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

วัตถุประสงค์ และผลงานวิจัยที่เกี่ยวข้อง

2.1 วัตถุประสงค์ของวิทยานิพนธ์ฉบับนี้

1. จัดการเชื่อมต่อเครื่องตรวจกวาดภาพ (Image Scanner) เข้ากับเครื่องไมโครคอมพิวเตอร์ เพื่อให้สามารถเก็บบันทึกข้อมูลภาพอักษรตัวพิมพ์ภาษาไทยรูปแบบต่าง ๆ ไว้ในเครื่อง ไมโครคอมพิวเตอร์ได้
2. ศึกษาวิธีการที่เหมาะสม ให้สามารถแยกภาพตัวอักษรออกจากภาพประโยคได้
3. ศึกษาวิธีการ วิเคราะห์วินิจฉัย ลักษณะของอักษรตัวพิมพ์ภาษาไทยรูปแบบต่าง ๆ ให้สามารถทราบได้ว่า คือตัวอักษรอะไร
4. จัดแสดงตัวอักษรภาษาไทยให้ปรากฏที่หน้าจอภาพ ตามอักษรที่วิเคราะห์ได้

2.2 ผลงานวิจัยที่เกี่ยวข้อง

2.2.1 ชื่อผลงานวิจัย "การจดจำรูปแบบตัวพิมพ์อักษรภาษาไทย"[3] ผู้วิจัย **รศ.ดร.ชม กิ้มปาน** ได้เสนอวิธีการไว้ 2 วิธีคือ การรู้จำแบบซ้อนทับ (matching method) และวิธีการแบบวิเคราะห์โครงสร้าง(structural analysis method)

วิธีที่ 1.การจดจำแบบซ้อนทับ[4]ได้แบ่งขั้นตอนออกเป็น 2 ขั้นตอนคือการแบ่งกลุ่มตัวอักษร (rough classification stage)ในขั้นตอนแรกนี้ผู้วิจัยได้พิจารณารูปร่างปลีกย่อยของตัวอักษรที่ไม่จำเป็นและสัญญาณรบกวนต่างๆ ที่ถูกอ่านเข้ามาจะถูกกำจัดออกไปโดยการทำตัวอักษรให้เบลอ ตัวอักษรที่เบลอแล้วจะถูกแบ่งออกเป็นกลุ่มๆ เงื่อนไขสำหรับการแบ่งกลุ่ม ผู้วิจัยได้ใช้วิธีการวัดค่าสัมประสิทธิ์ของความเหมือนของแต่ละรูปแบบแล้วใช้การกระจายแบบคาร์ฮูเนนโลบ(Karhunen Loeve expansion) หารูปแบบมาตราฐานสำหรับเก็บไว้เป็นตัวแทนของแต่ละกลุ่ม จากนั้นในขั้นตอนซึ่งเป็นการแยกตัวอักษรออกจากกลุ่ม (Fine classification stage) มีวิธีการอยู่สองวิธีสำหรับการแยกตัวอักษรออกจากกัน วิธีแรกใช้การซ้อนทับเป็นส่วนๆ (subpattern matching) ในการแยกตัวอักษรออกจากกัน วิธีที่สองใช้วิธีการสร้างฟังค์ชั่นการตัดสินใจแบบเชิงเส้น (linear decision functions) บนระนาบของไอเกนเวคเตอร์ที่ได้จากการกระจายแบบคาร์ฮูเนนโลบ สำหรับแยกตัวอักษรในแต่ละกลุ่มออกจากกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับวิธีที่ 2. การจดจำแบบวิเคราะห์โครงสร้าง[5]ผู้วิจัยได้แบ่งระบบการรู้จำออกเป็นสองขั้นตอนเหมือนเช่นกันวิธีแรก กล่าวคือในขั้นตอนการแบ่งกลุ่มตัวอักษรอาศัยเทคนิคการทำตัวอักษรให้บางกับข้อมูลตัวอักษรที่จะวิเคราะห์ ทั้งนี้เพื่อจะได้กำจัดสัญญาณรบกวน และลักษณะปลีกย่อยที่ไม่ต้องการต่างๆ ออกไปก่อน ตัวอักษรที่ถูกทำให้บางแล้วจะถูกแบ่งออกเป็นกลุ่ม ๆ โดยอาศัยฟังค์ชั่นการตัดสินใจแบบเชิงเส้นที่ยึดหลักความกว้างและความสูงของตัวอักษร และในขั้นตอนที่สองเป็นการแยกตัวอักษรออกจากกลุ่ม ผู้วิจัยได้ใช้ค่าความสัมพันธ์ระหว่างจุด โดยแทนที่ค่าที่ได้จากการคำนวณของแต่ละจุดในรูปร่างตัวอักษร จะเกิดคุณสมบัติทางเรขาคณิตศาสตร์ของตัวอักษรที่อยู่ในฟอร์มของค่าความสัมพันธ์ระหว่างจุดที่แทนจุดปลายจุดแยก จุดต่อเนื่องและอื่นๆจะให้ความแตกต่างระหว่างตัวอักษรต่าง ๆ ซึ่งคุณสมบัติเหล่านี้จะนำมาใช้ในการแยกตัวอักษรออกจากกันในแต่ละกลุ่ม

จากผลงานวิจัยทั้งสองวิธีนี้ การรู้จำแบบการซ้อนทับเป็นวิธีที่มีประสิทธิภาพสูงกว่า เพราะไม่เกิดความกำกวมในกระบวนการรู้จำ ทั้งนี้เพราะผลของการกระจายแบบคาร์ฮูเนน-โลบ แต่ขบวนการทั้งหมดจะใช้เวลาในการประมวลผลนานซึ่งใช้ได้ผลเฉพาะในการทดลองเท่านั้น ยังไม่เหมาะสำหรับที่จะนำไปประมวลผลกับเครื่องไมโครคอมพิวเตอร์ เพื่อใช้ในเชิงพานิชได้จริง

2.2.2 ชื่อผลงานวิจัย "การจดจำลายมือเขียนตัวอักษรภาษาไทยในระบบออนไลน์"[6],[7] ผู้วิจัย ดร.พิพัฒน์ หิรัญวณิชชากร ได้หาวิธีการจำอักษรตัวลายมือเขียนภาษาไทย โดยวิธีวิเคราะห์โครงสร้างจากคอนทัวร์ (Contour) ของตัวอักษรในลักษณะดิจิตอลคอนทัวร์ เพื่อหารหัสที่สอดคล้องกับทิศทางตามลักษณะของฟรีแมนแชนโค้ด (Freeman Chain Code) แล้วขจัดสิ่งรบกวน (Noise) ที่ไม่ต้องการออกไปจากนั้นค้นหาส่วนนูน (Convex) และส่วนเว้า (Concave) ของโครงร่างตัวอักษรและนำลักษณะมุมทางเรขาคณิตของส่วนนูนและส่วนเว้า มาคำนวณหาค่าความเหมือนของมุมทั้งจากภายในตัวอักษรเอง และความเหมือนระหว่างตัวอักษรแต่ละตัว แล้วนำค่าที่ได้ทั้งหมดเก็บเป็นค่าต้นแบบไว้ในตาราง

เมื่อมีตัวอักษรที่เป็นอินพุทเข้ามาจะถูกนำมาผ่านขบวนการเดียวกัน แล้วนำค่าที่คำนวณได้ไปเปรียบเทียบกับค่าต้นแบบในตาราง ลักษณะของงานวิจัยเน้นการใช้งานในแบบออนไลน์ โดยการเขียนตัวอักษรบนแผ่นดิจิไตร์เซอร์ ซึ่งจะเหมาะกับวิธีการนี้เท่านั้น ทั้งนี้เพราะการคำนวณหาค่ามุมส่วนนูนและส่วนเว้าจะดำเนินไปพร้อม ๆ กับการลากเส้นโครงร่างของตัวอักษร แต่ในการนำไปใช้งานจริง ยังไม่มีอุปกรณ์อินพุทที่เหมาะสมกับการใช้งาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.3 ชื่อผลงานวิจัย "การจดจำตัวอักษรลายมือเขียนภาษาไทยโดยการพิจารณาหัวของตัวอักษร"[8],[9] ผู้วิจัย นายสุรพันธ์ เอื้อไพบูลย์ ได้เลือกพิจารณาส่วนหัวของตัวอักษรซึ่งแตกต่างไปจากอักษรภาษาอื่น ๆ ซึ่งมักไม่มีส่วนหัว ทั้งนี้ผู้วิจัยได้ทำการตรวจสอบหัวของตัวอักษร แล้วจัดแบ่งกลุ่มตัวอักษรออกเป็นกลุ่ม ๆ คือ กลุ่มอักษรที่ไม่มีหัว กลุ่มอักษรที่มี 1 หัว กลุ่มอักษรที่มี 2 หัว และกลุ่มอักษรที่มี 3 หัว ภายในกลุ่มอักษรเหล่านั้นจะถูกพิจารณาแยกตัวอักษรออกจากกลุ่ม ซึ่งผู้วิจัยได้ใช้ค่าความสัมพันธ์ระหว่างจุดโดยแทนที่ค่าที่ได้จากการคำนวณของแต่ละจุดในรูปร่างตัวอักษรจะเกิดคุณสมบัติทางเรขาคณิตของตัวอักษร ที่อยู่ในฟอร์มของค่าความสัมพันธ์ระหว่างจุดที่แทนจุดปลายจุดแยก จุดต่อเนื่อง และอื่น ๆ จะให้ความแตกต่างระหว่างตัวอักษรต่าง ๆ ซึ่งคุณสมบัติเหล่านี้จะนำมาใช้ในการแยกตัวอักษรออกจากกันในแต่ละกลุ่มหรือการพิจารณาเฉพาะขอบเขตย่อย กรณีที่ลักษณะของตัวอักษรที่มีความคล้ายคลึงกันนี้ จนสามารถวิเคราะห์อักษรออกทีละตัวได้

จุดเด่นของวิธีการนี้อยู่ที่การพิจารณาหัวของตัวอักษรเป็นส่วนสำคัญ ทำให้ตัวอักษรที่มีลักษณะหัวทับ ต้น วิธีการนี้จะวิเคราะห์ผิดพลาดได้ง่าย โดยเฉพาะตัวอักษรพิมพ์ดีดหรืออักษรตัวพิมพ์ทั่วไปที่มักจะมีหัวทับต้นเสมอ

2.2.4 ชื่อผลงานวิจัย "การจดจำรูปแบบตัวอักษรคัดลายมือภาษาไทยโดยวิธีแยกลักษณะเด่น"[10] ผู้วิจัย นายประสาน ต้นติสานนท์ เป็นวิธีการที่พิจารณาถึงโครงสร้างของรูปแบบตัวอักษร ทั้งลายเส้นและพื้นเบื้องหลังของลายเส้นควบคู่กันไป วิธีการพิจารณาลายเส้นและพื้นเบื้องหลังลายเส้น ผู้วิจัยได้แบ่งออกเป็นสามขั้นตอนด้วยกันคือขั้นตอนแรก เป็นการเปลี่ยนลักษณะลายเส้นและลักษณะของพื้นเบื้องหลังลายเส้นไปเป็นรหัสเบื้องต้น (initial feature extraction) ขั้นตอนที่สองเป็นการลดทอนรหัสเบื้องต้นของพื้นเบื้องหลังลายเส้นที่ซ้ำซ้อนกันออกไป (unification) ขั้นตอนที่สามเป็นวิธีการรวมรหัสของลายเส้นและรหัสของพื้นเบื้องหลังลายเส้นเข้าด้วยกันเพื่อให้เป็นคุณสมบัติของรูปแบบตัวอักษร (concentration) เมื่อได้รหัสรวมของอักษรแต่ละตัวจะถูกจัดแยกกลุ่ม แล้วพิจารณาหาคุณสมบัติที่เป็นลักษณะเด่นเฉพาะของตัวอักษรนั้น ๆ ที่ไม่เหมือนกันกับอักษรตัวอื่น ทำการแยกตัวอักษรออกจากกลุ่มออกไป รหัสที่เป็นคุณสมบัติเด่นของอักษรแต่ละตัวจะถูกพิจารณาจากบริเวณพื้นที่รอบ ๆ ตัวอักษร โดยผู้วิจัยจะพิจารณาด้วยสายตา แล้วจึงกำหนดเป็นเงื่อนไข วิธีการนี้จึงใช้ได้ผลดีกับตัวอักษรต้นแบบ แต่หากลักษณะของตัวอักษรที่ต้องการวิเคราะห์เปลี่ยนรูปร่างแตกต่างจากต้นแบบแล้วจะไม่สามารถวิเคราะห์ได้ อีกทั้งยังยากต่อการเพิ่มเงื่อนไขใหม่ ๆ เพื่อแยกตัวอักษรออกจากกลุ่ม หากอักษรที่วิเคราะห์ไม่สามารถวินิจฉัยได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.5 ชื่อผลงานวิจัย "การออกแบบพจนานุกรมสำหรับการเรียนรู้อักษรคัดลายมือไทย-อังกฤษอัตโนมัติบนเครื่องไมโครคอมพิวเตอร์"[12] ผู้วิจัย นายชาย เกษมอมรกุล ผลงานวิจัยชิ้นนี้ได้พัฒนาปรับปรุงจากผลงานวิจัยในหัวข้อ 2.2.4 โดยผู้วิจัยยังคงอาศัยหลักการพิจารณาการหารหัสเป็น 3 ขั้นตอนเช่นเดียวกัน แต่วิธีการค้นหาคุณสมบัติเด่นของตัวอักษร จากค่าความถี่ของรหัสต่าง ๆ 5 อันดับแรกเป็นคุณสมบัติประจำตัวอักษรแต่ละตัว แล้วบันทึกเก็บเพิ่มเติมไว้เป็นอักษรต้นแบบ หรือหากอักษรใดที่วิเคราะห์ได้แตกต่างจากอักษรต้นแบบ ก็จะเพิ่มเป็นคุณสมบัติใหม่สำหรับการเป็นต้นแบบในครั้งต่อไป

วิธีการนี้นับว่าได้แก้ปัญหาการเพิ่มความสามารถการรู้จำของโปรแกรมได้เป็นอย่างดี แต่ต้องเพิ่มค่าต้นแบบทุก ๆ ตัวอักษรที่วิเคราะห์ได้ใหม่ ดังนั้นขนาดของตารางฐานข้อมูลเพื่อการตัดสินใจจะมีขนาดใหญ่มากตามไปด้วยทำให้โปรแกรมทำงานได้ช้าลง ยิ่งค่าต้นแบบมากเท่าใดความเร็วในการวิเคราะห์ก็จะยิ่งช้าลงมากด้วยนอกจากนั้นในการเขียนลายมือคัดเพื่อเป็นข้อมูลในการวิเคราะห์ ยังจำกัดขนาดตัวอักษรและจำนวนบรรทัด จึงยังไม่เหมาะที่จะนำไปใช้งานจริง

2.3 ลักษณะของวิทยานิพนธ์ที่แตกต่างจากผลงานวิจัยที่ผ่านมา

2.3.1 วิธีแยกภาพตัวอักษรออกจากภาพประโยค[17]

ผลงานวิจัยต่าง ๆ ที่ได้กล่าวมาแล้ว จะปรากฏข้อดีข้อเสียแตกต่างกัน แต่ขั้นตอนหนึ่งที่ผู้วิจัยส่วนใหญ่มักไม่เน้นความสำคัญหรือกล่าวถึงคือ ขั้นตอนการแยกภาพตัวอักษรออกจากภาพประโยค (segmentation) ทั้งนี้ในการใช้งานจริงการแยกกลุ่มอักษรแต่ละบรรทัดออกจากกันและแยกภาพตัวอักษรทีละตัวออกจากประโยคนับเป็นขั้นตอนที่สำคัญอันดับแรก ไม่น้อยไปกว่าขั้นตอนการรู้จำ จะมีบางผลงานวิจัยที่กล่าวถึงการแยกภาพอักษรออกจากประโยค แต่ก็เพียงจำกัดขอบเขตให้ผู้ใช้เขียนตัวอักษรลงในกรอบที่กำหนดให้เท่านั้น ซึ่งในกรณีเอกสารหรือสิ่งตีพิมพ์ทั่วไปจะไม่มีการกำหนดกรอบของอักษรแต่ละตัว แต่จะเรียงชิดกันเป็นประโยคยาวเต็มบรรทัด และเต็มทั้งหน้ากระดาษ ดังนั้นวิทยานิพนธ์ฉบับนี้จึงได้กล่าวถึงวิธีการแยกภาพตัวอักษรออกจากภาพประโยคอย่างละเอียด โดยได้ทดลองนำไปแยกภาพอักษรจากภาพประโยคของหน้าหนังสือพิมพ์ สิ่งพิมพ์และเอกสารทั่วไป ซึ่งให้ผลดีเป็นที่น่าพอใจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.2 ลักษณะของตัวอักษรที่ใช้วิเคราะห์

จากผลงานวิจัยที่ผ่านมาทั้ง 5 ชิ้นมีผู้วิจัยได้เลือกใช้อักษรคัดลายมือเป็นข้อมูลในการวิเคราะห์ถึง 4 ผลงาน อาจกล่าวได้ว่าวิธีการของผลงานวิจัยส่วนใหญ่ที่ผ่านมาอาจใช้งานได้เหมาะกับอักษรคัดลายมือ ซึ่งสามารถคัดตัวบรรจงและขนาดตามความเหมาะสมได้ ซึ่งจะแตกต่างจากอักษรตัวพิมพ์ที่ อาจมีจุดหมึก รอยเบื้อน หรือการขาดเว้าแหว่งของอักษรตัวพิมพ์จากสิ่งพิมพ์ อีกทั้งยังมีขนาดที่แตกต่างกันจากขนาดใหญ่จนถึงเล็กสุด ปัญหาในการพิจารณาวิธีการวิเคราะห์จึงซับซ้อนแตกต่างออกไป ในวิทยานิพนธ์ฉบับนี้ได้เลือกอักษรตัวพิมพ์ภาษาไทยหลายรูปแบบและหลายขนาดมาทำการวิเคราะห์ เพื่อให้ได้ผลในการพัฒนาเพื่อไปใช้งานจริง

2.3.3 เทคนิควิธีการ

ในด้านวิธีการ ผู้เขียนนี้ได้ศึกษาวิธีการของผลงานวิจัยที่ผ่านมา พิจารณาเห็นว่าการหาค่าคอนเซ็นแทร็คโค้ต ในผลงานหัวข้อ 2.2.4 ซึ่งให้ผลที่แสดงถึงคุณสมบัติเฉพาะได้ดี แต่ใช้เวลาในการประมวลผลเพื่อกำหนดรหัสให้ครบทั้งสามขั้นตอนนานเกินความจำเป็น จึงได้ดัดแปลงวิธีการดังกล่าวให้สามารถหารหัสได้อย่างรวดเร็วขึ้น นำมาเป็นส่วนหนึ่งของการแยกกลุ่มตัวอักษร เป็นผลทำให้การจัดแยกกลุ่มเพื่อการวิเคราะห์ได้รวดเร็วยิ่งขึ้น และในขั้นตอนการแยกตัวอักษรออกจากกลุ่มผู้วิจัยได้เสนอวิธีการที่แตกต่างจากผลงานวิจัยที่ผ่านมา โดยการพิจารณาลักษณะเส้นประกอบเป็นโครงร่างของตัวอักษร ซึ่งตัวอักษรแต่ละตัวจะถูกแบ่งเป็นพื้นที่ออกเป็น 8 ส่วน แล้วประมาณลักษณะเส้นโครงร่างตัวอักษรแต่ละส่วน เป็นเส้นตรงแนวตั้งหรือแนวนอน เป็นส่วนโค้งหันไปทางซ้ายหรือทางขวา เป็นส่วนโค้งหงายขึ้นหรือคล่ำลง เหล่านี้เป็นต้น หลังจากนำค่าประมาณลักษณะของเส้นโครงร่างตัวอักษรทั้ง 8 ส่วนมาประกอบรวมอีกครั้งก็จะเป็นค่าคุณสมบัติเฉพาะของตัวอักษร

2.4 เนื้อหาของวิทยานิพนธ์

สำหรับวิทยานิพนธ์นี้ เป็นการเสนอวิธีการรู้จำอักษรตัวพิมพ์ภาษาไทยหลายรูปแบบและหลายขนาด โดยจะกล่าวถึงขั้นตอนต่าง ๆ ที่จะทำให้เครื่องคอมพิวเตอร์สามารถอ่านตัวพิมพ์อักษรภาษาไทยได้ ดังสรุปเนื้อหาของวิทยานิพนธ์ได้ดังนี้

บทที่ 1 เป็นบทนำ ซึ่งกล่าวโดยย่อ ๆ ถึงประโยชน์และวิธีการของการรู้จำตัวอักษร

บทที่ 2 กล่าวถึงวัตถุประสงค์ ความเป็นมา และผลงานวิจัยที่เกี่ยวข้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3 ทฤษฎีที่นำมาใช้ หลักการทำงานของเครื่อง OCR ซึ่งเป็นการแปลงสัญญาณแสง ไปเป็นสัญญาณไฟฟ้า แล้วถูกเปลี่ยนเป็นค่าเลขดิจิตอล เพื่อจัดเก็บไว้ในหน่วยความจำของเครื่องคอมพิวเตอร์ ทฤษฎีเฟียเจอร์คอนเซนเทรชั่น ทฤษฎีโทโบโลจิคอลเฟียเจอร์เอ็กแทร็กชั่น การหาขนาดความกว้างความสูงของตัวอักษร การหาจุดศูนย์ถ่วงและการหาขอบภาพเพื่อนำไปใช้ในการแยกตัวอักษรออกจากประโยค

บทที่ 4 การจัดเตรียมข้อมูลล่วงหน้า ตั้งแต่การจัดเก็บข้อมูลภาพตัวอักษรลงในหน่วยความจำและเทคนิคการแยกตัวอักษรออกจากประโยค

บทที่ 5 การจัดกลุ่มตัวอักษร โดยอาศัยจุดศูนย์ถ่วงของตัวอักษรเป็น ตำแหน่งอ้างอิง และที่ตำแหน่งจุดศูนย์ถ่วงนั้นเองจะทำการตรวจสอบหาค่ารหัสคอนเซนแทตโค๊ตซึ่งเป็นค่ารหัสจากลักษณะเด่นของตัวอักษร ค่ารหัสที่ได้จะสามารถจัดแบ่งกลุ่มของตัวอักษรออกเป็นกลุ่มใหญ่ ๆ ได้ 7 กลุ่ม

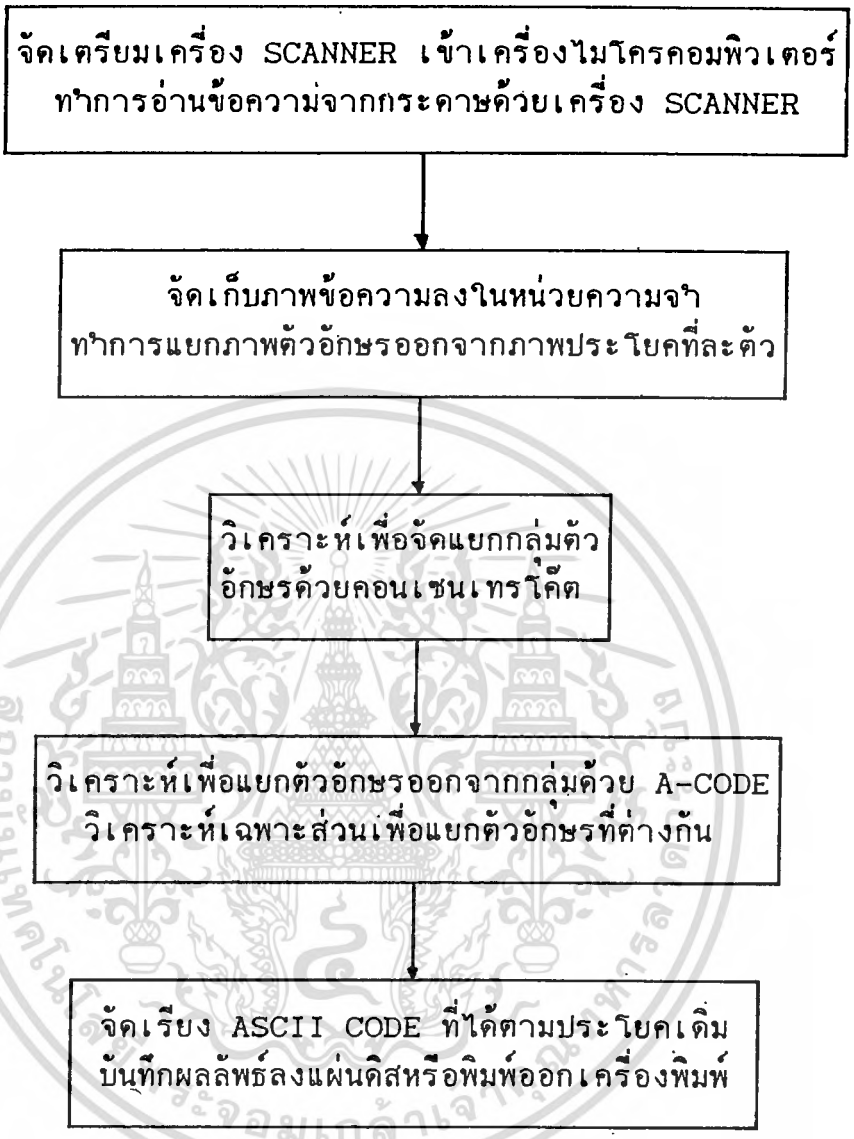
บทที่ 6 การแยกตัวอักษรออกจากกลุ่ม ภายในกลุ่มอักษรย่อยจะถูกนำมาพิจารณาแยกอักษรออกจากกัน โดยจะแยกวิเคราะห์ 2 ลักษณะคือวิเคราะห์แต่ละตัวอักษร แต่หากยังมีตัวอักษรใดภายในกลุ่มย่อยมีลักษณะใกล้เคลียงกันอยู่จะทำการวิเคราะห์เฉพาะส่วน เทคนิคในการวิเคราะห์นี้เริ่มจากการแบ่งส่วนประกอบตัวอักษรออกเป็น 8 ส่วน ตามลักษณะเส้นทะแยงมุมทั้ง 8 ทิศ แล้วค้นหาค่ารหัส A-Code ตามลักษณะทางโทโบโลจิคอลเพื่อแทนค่าเป็นรหัสออกมา ชุดรหัส A-code ที่วิเคราะห์แล้วจะเป็นคุณสมบัติเฉพาะประจำตัวอักษรเหล่านั้นและสามารถใช้เทคนิควิธีการเดียวกันกับตัวอักษรอินพุทที่เข้ามา ก็จะสามารถแยกอักษรออกจากกลุ่มและวิเคราะห์ได้ว่าเป็นอักษรตัวอะไร และสามารถนำ ASCII CODE ที่ได้มาเรียงเป็นประโยคตามตำแหน่งเดิมเหมือนต้นฉบับเดิม ดังแสดงขบวนการรู้จำไว้ในรูปที่ 2

บทที่ 7 โปรแกรมการรู้จำรูปแบบตัวพิมพ์อักษรภาษาไทย เป็นโปรแกรมที่พัฒนาขึ้นสำหรับงานวิจัยครั้งนี้ โดยใช้งานร่วมกับโปรแกรมสำเร็จรูปที่ทำหน้าที่ควบคุมการทำงานของเครื่อง Handy Scanner โปรแกรมที่พัฒนาขึ้นนี้มีฟังค์ชั่นการใช้งานที่เป็นประโยชน์ต่อผู้ใช้ สามารถนำไปใช้งานได้จริง

บทที่ 8 การทดสอบโปรแกรม โดยนำข้อมูลภาพตัวอักษรทั้งหมด 15 แบบมาทำการทดสอบกับโปรแกรมที่พัฒนาขึ้น ข้อทั้งหมดเป็นอักษรต้นแบบ 896 ตัวและข้อมูลภาพที่เป็นบทความจากอักษรรูปแบบต่าง ๆ กันทั้ง 15 แบบรวมตัวอักษรที่ใช้ทดสอบ 9750 ตัว

บทที่ 9 เป็นการสรุป และวิจารณ์ผลการทดลอง รวมทั้งข้อเสนอแนะต่าง ๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2. ขบวนการรู้จำตัวพิมพ์อักษรภาษาไทยหลายรูปแบบและหลายขนาด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
ทฤษฎีที่นำมาใช้

3.1 หลักการทำงานของเครื่อง Scanner

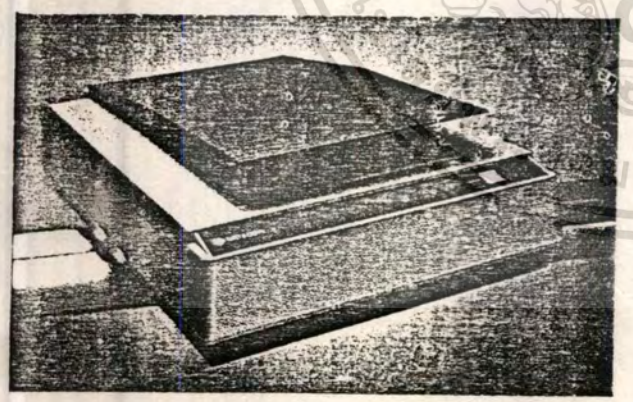
3.1.1 เครื่อง Scanner

เครื่อง Scanner จัดเป็นเครื่อง OCR ชนิดหนึ่งที่มีผู้นิยมนำมาใช้สำหรับงานด้าน Pattern Recognition กันอย่างแพร่หลาย ในอนาคตข้างหน้าเป็นที่คาดกันว่ากระดาษอยู่บนโต๊ะทำงานนั้นเราสามารถแปลงข้อความบนกระดาษเป็นตัวอักษรให้เป็น ASCII Code และเปลี่ยนรูปภาพบนกระดาษ ให้อยู่ในรูปของ BIT MAPPED GRAPHICS ที่สามารถแก้ไข ก่อนพิมพ์ลงในเครื่องพิมพ์เลเซอร์ และส่งเป็นข้อความหรือรูปภาพที่มีคุณภาพนั้น ให้กับคนอื่นที่มีจอบนโต๊ะได้อย่างรวดเร็ว

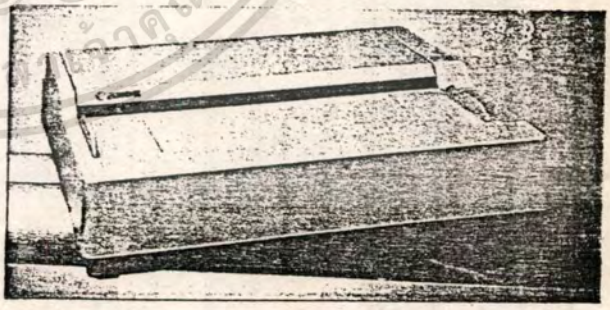
ปัจจุบันอุปกรณ์ดังกล่าว เป็นที่รู้จักกันดีในชื่อที่เรียกว่าเครื่อง Scanner ส่วนมากจะมีอยู่ในสำนักงานอยู่บ้างแล้ว เครื่อง Scanner จะทำหน้าที่แปลงภาพจากกระดาษให้ปรากฏรูปภาพนั้นบนจอภาพคอมพิวเตอร์ ซึ่งหมายถึงมีข้อมูลรูปภาพอยู่ในหน่วยความจำของคอมพิวเตอร์แล้วนั่นเอง

3.1.2 เทคนิคของเครื่อง Scanner

ในตอนแรกของการ scan เป็นขั้นตอนที่สำคัญที่สุด ขึ้นอยู่กับบริษัทผู้ออกแบบว่าจะให้กระดาษเคลื่อนผ่านชิ้นส่วนของการ scan ที่อยู่กับที่ หรือจะให้ชิ้นส่วนของการ scanเคลื่อนผ่ารกระดาษแทน วิธีธรรมดาที่ใช้กันมากก็คือ ให้กระดาษที่จะอ่านเคลื่อนที่ไปทางช่องผ่านของกระดาษ (paper path) ที่สามารถปรับตำแหน่งและป้อนกระดาษที่จะอ่านได้



ชนิดที่ให้กระดาษอยู่กับที่

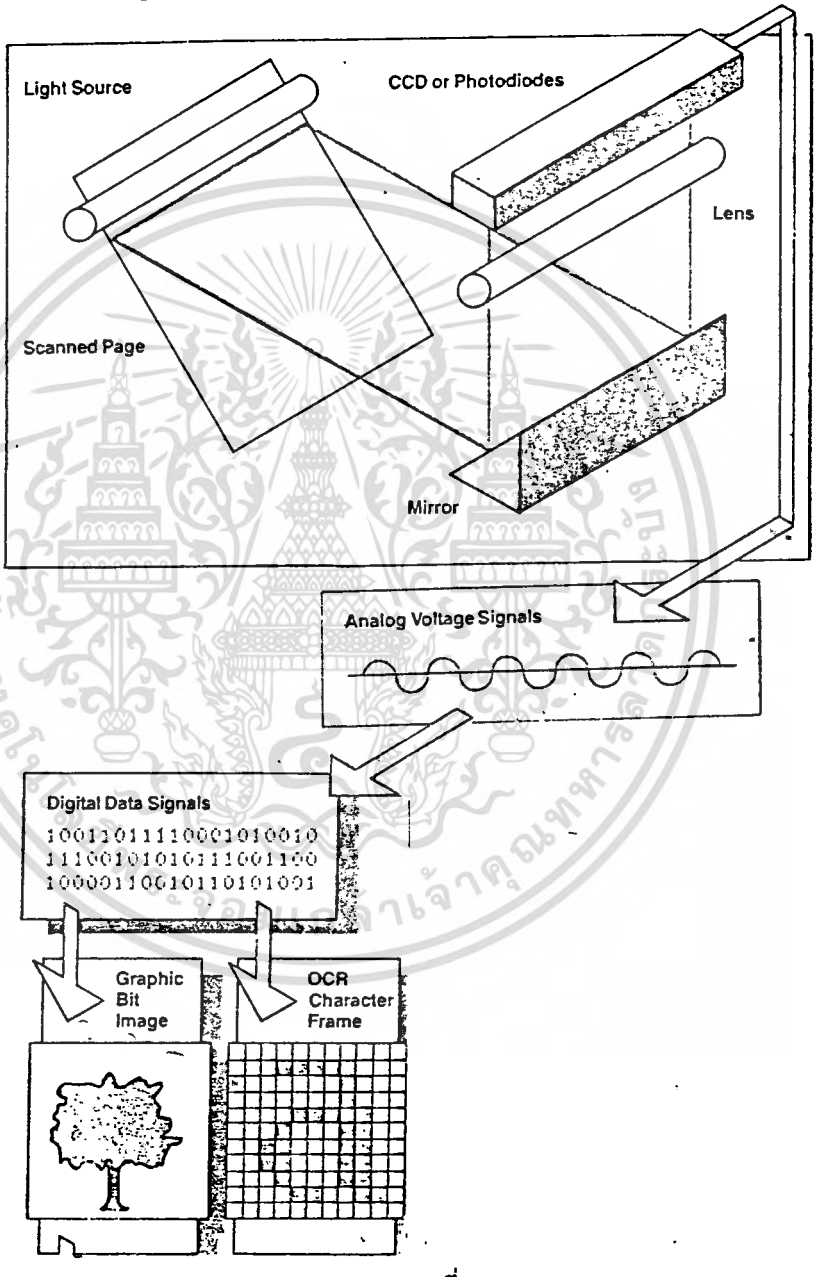


ชนิดให้กระดาษเคลื่อนผ่าน

รูปที่ 3. แสดงลักษณะของเครื่อง Scanner

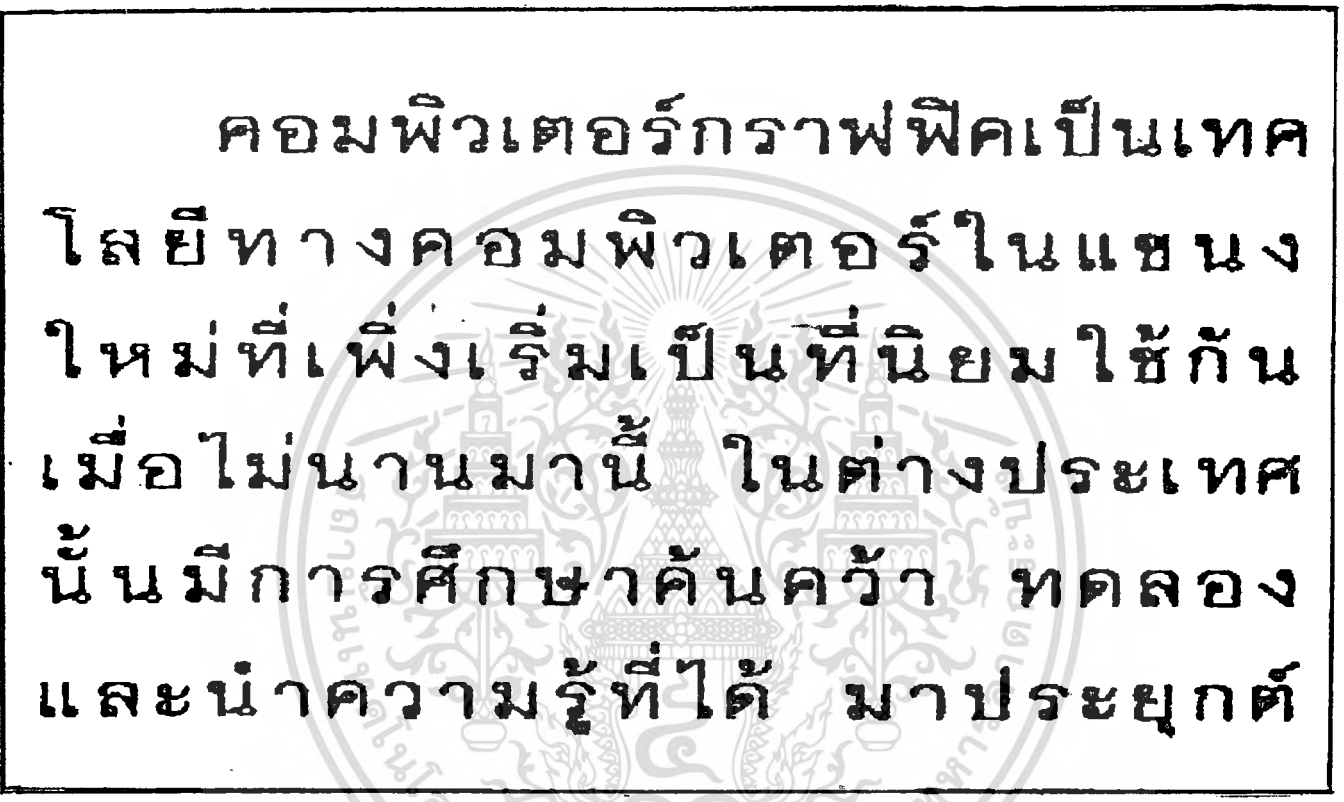
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อกระดาษที่จะอ่านอยู่ในตำแหน่งที่จะ scan ได้ จะใช้แสงความถี่ต่ำ (low-frequency light) สีแดงหรือเหลืองนีออนส่องกระทบกระดาษ พื้นกระดาษส่วนที่มืดหรือดำจะซึมซับแสงเอาไว้ แต่ส่วนที่สว่างหรือขาวแสงจะสะท้อนออกมาพร้อมกับจะปรับจุดโฟกัสของแสงให้พุ่งไปที่ photodiode array หรือที่ CCD(charged coupled device) ซึ่งจะแปลงส่วนของแสงสว่างและมืด เป็นbinary data (0 หรือ 1) จากจุดดังกล่าว ข้อมูลที่ได้จากตัวหนังสือหรือรูปภาพจะถูกเก็บใน RAM เพื่อการจำตัวอักษร



รูปที่ 4. แสดงหลักการทำงานของเครื่อง Scanner

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

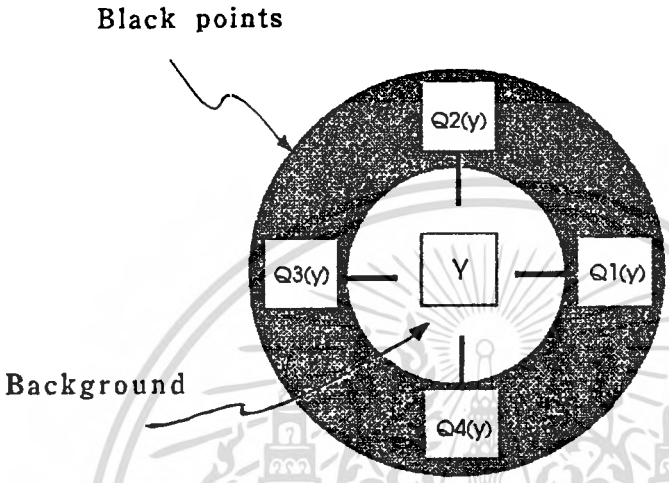


รูปที่ 5. ภาพแสดงลักษณะข้อมูลที่ได้จากอ่านด้วยเครื่อง Scanner

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 <u>ทฤษฎีเฟียเจอร์คอนเซนเทร</u>[13]

การพิจารณาหารหัสของ อินนิเชี่ยนเฟียเจอร์เอ็กแทร็กชั่น ซึ่งต่อไปจะเรียกว่าค่า Q-Code จะอาศัยการพิจารณาจากจุดที่อยู่รอบบริเวณตำแหน่งปัจจุบัน ทั้ง 4 ทิศ ดังรูปที่ 6.



รูปที่ 6. ลักษณะทิศทางที่ใช้พิจารณาการหาค่ารหัส
Feature Concentrat Code (Q-Code)

<u>3.2.1 การหาค่า Q-CODE</u>

จากรูปที่ 6. กำหนดให้ Qi(x) เป็นฟังค์ชั่นหนึ่งที่มีค่าเป็น 1 เมื่อมีจุดค่าอย่างน้อย 1จุดที่เกิดขึ้นในทิศทางหนึ่ง โดยพิจารณาจากจุดขาวจุด ๆ หนึ่ง (จุดy) และจะมีค่าเป็น 0 สำหรับกรณีอื่น จากนั้นจะถูกเข้าโค๊ดเป็น Qk (k = 0,1,2,3..

$$K(y) = 8*Q_1(y) + 4*Q_2(y) + 2*Q_3(y) + 1*Q_4(y) \quad .....(1)$$

หรืออาจนำค่าจากการสังเกตไปเทียบหาค่า Q-code ในตารางที่ 1 (ตารางแสดงลักษณะของ Q-code) ซึ่งจะใช้แทนในจุดที่กำลังพิจารณาอยู่ ด้วยโค้ดต่างๆ กันจากQ_1 , Q_2 ,..., Q_{15}

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Q-code	Pictorial Symbol
Q_0	
Q_1	
Q_2	
Q_3	
Q_4	
Q_5	
Q_6	
Q_7	

Q-code	Pictorial Symbol
Q_8	
Q_9	
Q_{10}	
Q_{11}	
Q_{12}	
Q_{13}	
Q_{14}	
Q_{15}	

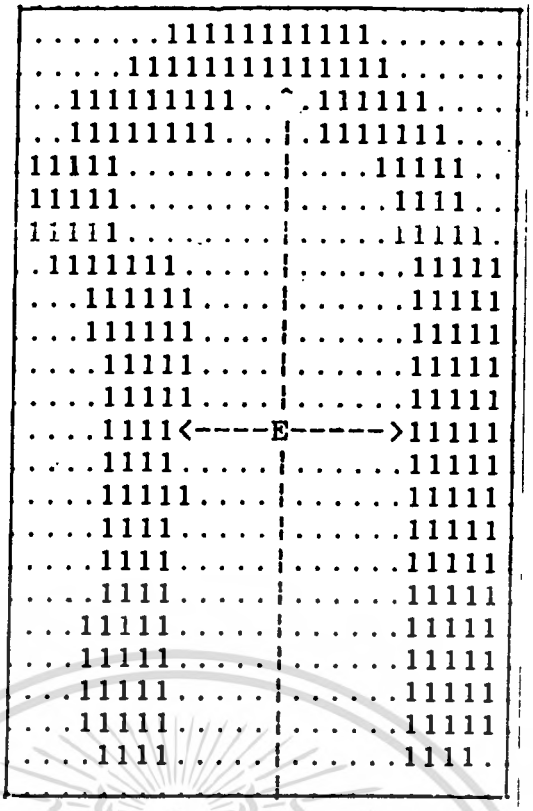
ตารางที่ 1 แสดงลักษณะของ Q-CODE Table

3.2.2 การกำหนดค่า Q-Code ของตัวอักษร

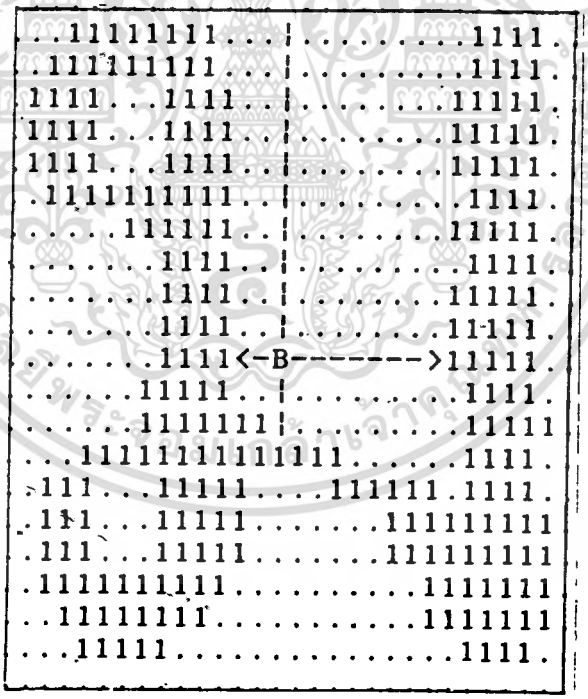
จุดที่จะหารหัสอยู่ส่วนใด ๆ ของตัวอักษร เราจะตรวจสอบลักษณะแวดล้อมทั้ง4 ทิศทางของจุดหารหัสนั้น แล้วแทนด้วยค่าของ Q-Code ที่คำนวณได้

ดังตัวอย่างในรูปที่ 7. จะเห็นว่าตำแหน่งที่จะหารหัสของอักษร ก มีตำแหน่งจุดศูนย์ถ่วงอยู่บริเวณช่องว่างเมื่อคำนวณค่า K(y) หรือพิจารณาเทียบกับตารางที่ 1 แล้วณ ที่ตำแหน่งจุดหารหัสจะถูกแทนค่าด้วย ระหัส E หรือกรณีรูปที่ 8. จะเห็นว่าตำแหน่งที่จะหารหัสของอักษร ม มีตำแหน่งจุดศูนย์ถ่วงอยู่บริเวณช่องว่างเมื่อคำนวณค่า K(y) หรือพิจารณาเทียบกับตารางที่ 1 แล้วณ ที่ตำแหน่งจุดหารหัสจะถูกแทนค่าด้วย รหัส B

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 7. แสดงการแทนค่ารหัส Q-Code อักษร ก.



รูปที่ 8. แสดงการแทนค่ารหัส Q-Code อักษร ม.

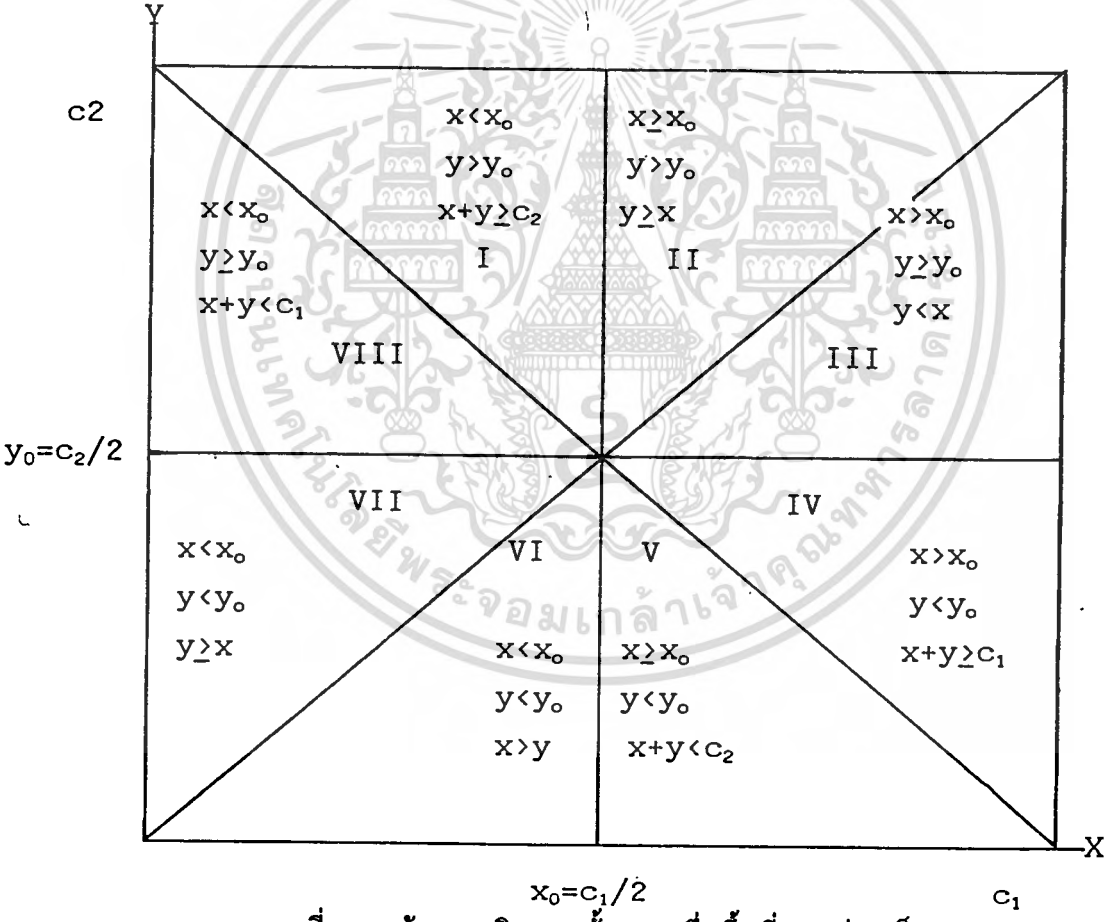
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ห้องสมุดกลาง พระจอมเกล้าลาดกระบัง

3.3 ทฤษฎีโทโปโลจิคอลเพียเจอร์เอ็กแทร็กชั่น[24]

จากธรรมชาติของลักษณะอักษรภาษาไทย ส่วนใหญ่รูปร่างตัวอักษรจะเป็นลักษณะ stroke เดียว (เขียนรูปร่างอักษรต่อเนื่องจนครบตัวโดยไม่ยกปากกา) ซึ่งทำให้ปรากฏส่วนโค้งส่วนเว้าของตัวอักษรหลายตำแหน่งในอักษรหนึ่งตัว ดังนั้นในการกำหนดรหัสจากรูปร่างลักษณะของตัวอักษรเหล่านั้น เราได้กำหนดวิธีการค้นหาลักษณะรูปร่างของตัวอักษรเหล่านั้นให้ละเอียดและครอบคลุมทุกด้าน รหัสที่หาได้ก็จะแสดงค่าลักษณะเฉพาะที่เป็นส่วนประกอบเป็นตัวอักษรเหล่านั้นออกมาได้อย่างชัดเจน

ในการพิจารณาลักษณะตัวอักษรให้ครบทุกด้านนี้ จะทำการแบ่งส่วนต่าง ๆ ของรูปร่างตัวอักษรออกเป็น 8 ส่วนในลักษณะเส้นทะแยงมุม 8 เส้น โดยใช้เลขโรมันสำหรับส่วนที่ถูกแบ่งออกไปเป็น 8 ส่วนซึ่งจะเรียกว่าเป็น 8 Octant มีจุดศูนย์กลางอยู่บริเวณกึ่งกลางรูปร่างตัวอักษรดังรูปที่ 9.



รูปที่ 9. ลักษณะทิศทางทั้ง 8 ซึ่งพื้นที่ถูกแบ่งเป็น 8 Octant

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสาร 15939 นำไปใช้

จากรูปที่ 9. กำหนดให้

x : ค่าโคออดิแนทของจุดใด ๆ ทางแกน X

y : ค่าโคออดิแนทของจุดใด ๆ ทางแกน Y

c_1 : ขนาดความกว้างของตัวอักษร

c_2 : ขนาดความสูงของตัวอักษร

$x_o = c_1/2$

$y_o = c_2/2$

การกำหนดขอบเขตของแต่ละ octant เป็นดังนี้

จุดโคออดิแนทใด ๆ จะอยู่ภายใน octant I เมื่อ

$$x < x_o$$
$$y > y_o$$
$$x+y \geq c_2$$

จุดโคออดิแนทใด ๆ จะอยู่ภายใน octant II เมื่อ

$$x \geq x_o$$
$$y > y_o$$
$$y \geq x$$

จุดโคออดิแนทใด ๆ จะอยู่ภายใน octant III เมื่อ

$$x > x_o$$
$$y \geq y_o$$
$$x < x$$

จุดโคออดิแนทใด ๆ จะอยู่ภายใน octant IV เมื่อ

$$x > x_o$$
$$y < y_o$$
$$x+y \geq c_1$$

จุดโคออดิแนทใด ๆ จะอยู่ภายใน octant V เมื่อ

$$x \geq x_o$$
$$y < y_o$$
$$x+y < c_2$$

จุดโคออดิแนทใด ๆ จะอยู่ภายใน octant VI เมื่อ

$$x < x_o$$
$$y < y_o$$
$$x > y$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จุดโคออดิแนทใด ๆ จะอยู่ภายใน octant VII เมื่อ

$$x < x_0$$

$$y < y_0$$

$$y \geq x$$

จุดโคออดิแนทใด ๆ จะอยู่ภายใน octant VIII เมื่อ

$$x < x_0$$

$$y \geq y_0$$

$$x+y < c_1$$

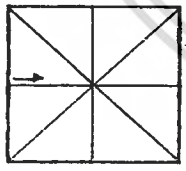
จากแต่ละ Octant จะพิจารณาจุดใด ๆ ที่อยู่ภายใน Octant หากมีส่วนใดของรูปร่างตัวอักษรอยู่ภายใน Octant ใด กล่าวได้ว่าที่ Octant นั้นมี stroke และในทางตรงข้ามหากไม่มีส่วนใดของรูปร่างตัวอักษรอยู่ภายใน Octant เลยถือได้ว่า Octant นั้นไม่มี stroke ลักษณะของ stroke จะปรากฏเป็นรูปร่างหลายลักษณะ ซึ่งลักษณะเหล่านั้นเราจะกำหนดเป็นรหัส A-Code และเนื่องจากลักษณะตัวอักษรที่นำมาพิจารณา ได้พิจารณาเฉพาะส่วนที่เป็นขอบของตัวอักษร ดังนั้นลักษณะของ Stroke ที่เกิดขึ้นจึงประกอบด้วยเส้นต่อเนื่องสองเส้นขนานกัน จึงเรียก stroke ลักษณะดังกล่าวว่าเป็น Stroke คู่ขนาน

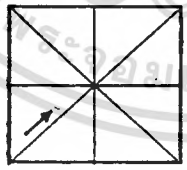
สำหรับในการกำหนดลักษณะของ stroke ให้เป็นรหัส A-Code ในแต่ละ Octant ได้กำหนดเงื่อนไขดังนี้

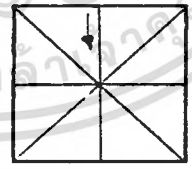
1. หากไม่มี stroke ใน Octant ใด ๆ หรือเป็น stroke ที่ไม่มี Stork คู่ขนานจาก octant ใด ๆ ลากผ่าน octant ที่อยู่ติดกัน กำหนดรหัส A-Code เป็น a_0

ตัวอย่างดังรูปที่ 10.ก แสดง octant ที่ไม่มี stroke และรูปที่ 10 ข,ค แสดงลักษณะ stroke ที่ลากผ่าน octant โดยที่ไม่มี stroke คู่ขนาน

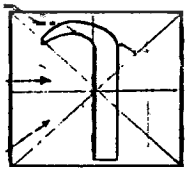
ลักษณะ stroke

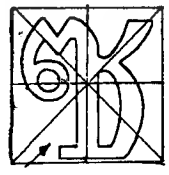


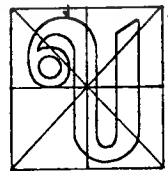




ตัวอย่างอักษร





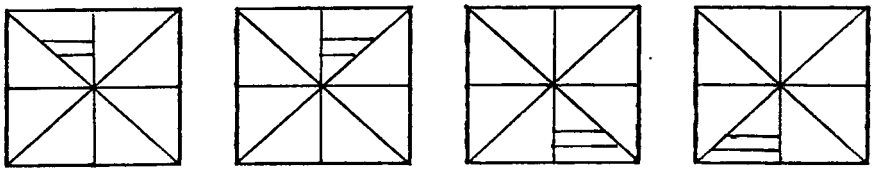


รูปที่ 10.ก รูปที่ 10.ข รูปที่ 10.ค

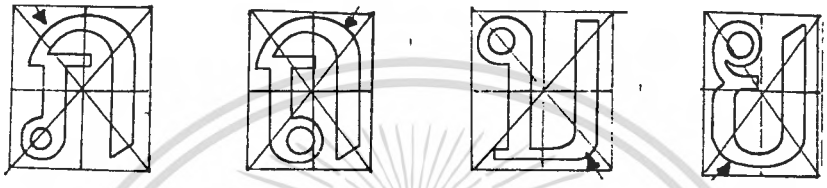
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. stroke คู่ขนานใด ๆ ที่ผ่าน octant 1,2,5 หรือ 6 จะถูกประมาณค่าเป็น horizontal stroke และให้รหัส A-Code เป็น a_1 ตัวอย่างดังรูปที่ 11.

ลักษณะ stroke



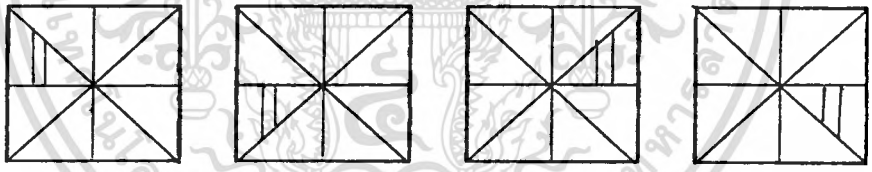
ตัวอย่าง อักษร



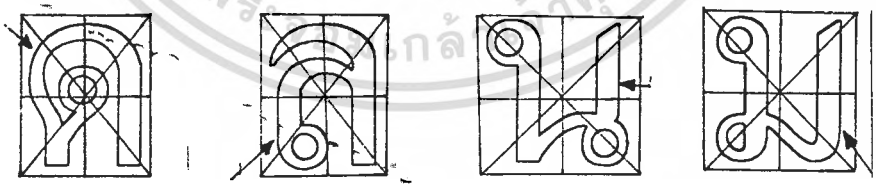
รูปที่ 11.ก รูปที่ 11.ข รูปที่ 11.ค รูปที่ 11.ง

3. stroke คู่ขนานใด ๆ ที่ผ่าน octant 3,4,7 หรือ 8 ถูกประมาณค่าเป็น vertical stroke และให้รหัส A-Code เป็น a_2 ตัวอย่างดังรูปที่ 12.

ลักษณะ stroke



ตัวอย่าง อักษร

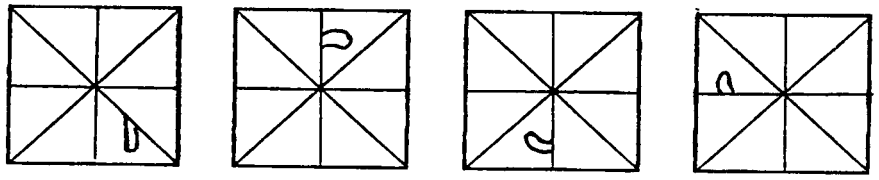


รูปที่ 12.ก รูปที่ 12.ข รูปที่ 12.ค รูปที่ 12.ง

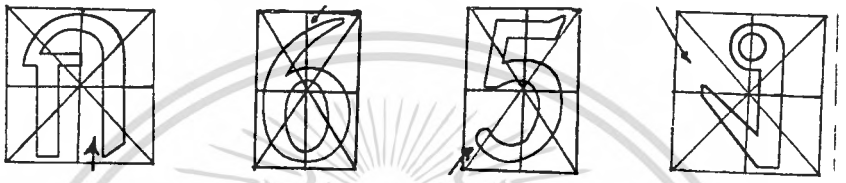
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. stroke ที่ไม่มี stroke คู่ขนาน เริ่มต้นใน Octant ใด ๆ มีลักษณะพุ่งเข้าหา Octant ที่อยู่ติดกัน แล้วย้อนกลับเข้ามาหา Octant เดิม เห็นเป็นลักษณะตามเข็มนาฬิกา ให้ค่ารหัส A-Code เป็น a_3 ตัวอย่างดังรูปที่ 13.

ลักษณะ stroke



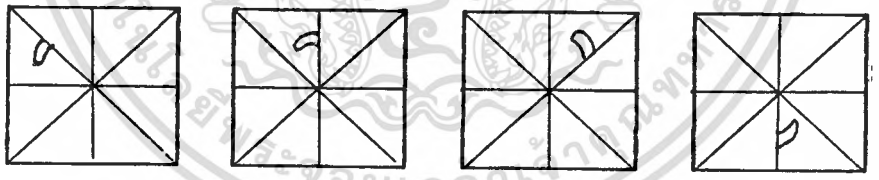
ตัวอย่างอักษร



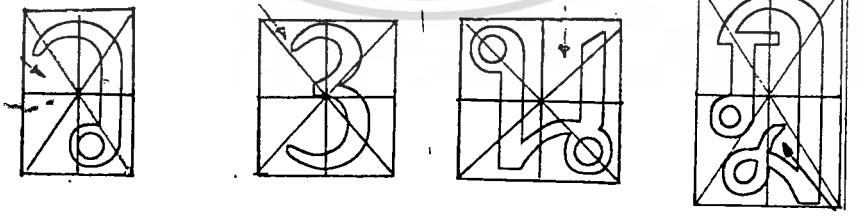
รูปที่ 13.ก รูปที่ 13.ข รูปที่ 13.ค รูปที่ 13.ง

5. stroke ที่ไม่มี stroke คู่ขนาน เริ่มต้นใน Octant ใด ๆ มีลักษณะพุ่งเข้าหา Octant ที่อยู่ติดกัน แล้วย้อนกลับเข้ามาหา Octant เดิม เห็นเป็นลักษณะทวนเข็มนาฬิกา ให้ค่ารหัส A-Code เป็น a_4 ตัวอย่างดังรูปที่ 14.

ลักษณะ stroke



ตัวอย่างอักษร

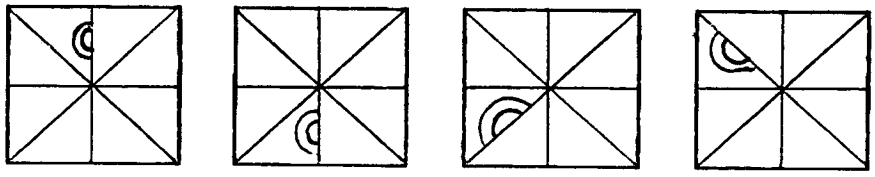


รูปที่ 14.ก รูปที่ 14.ข รูป 14.ค รูป 14.ง

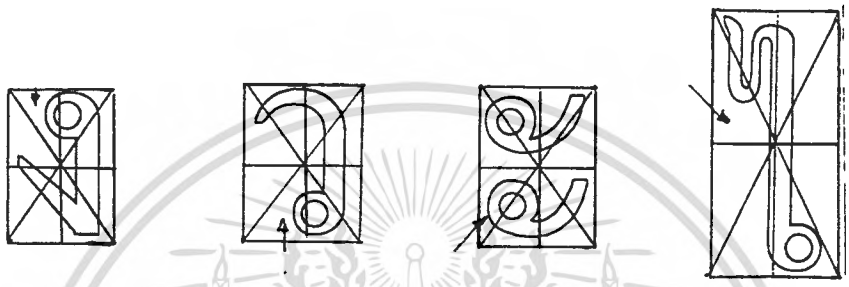
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. stroke คู่ขนานใดที่กำลังพิจารณาเป็น loop ภายใน octant ที่ 1,6,7 และ 8 มีลักษณะ stroke ที่ หันไปด้านซ้าย ให้ค่ารหัส A-Code เป็น a_5 ตัวอย่างดังรูปที่ 15.

ลักษณะ stroke



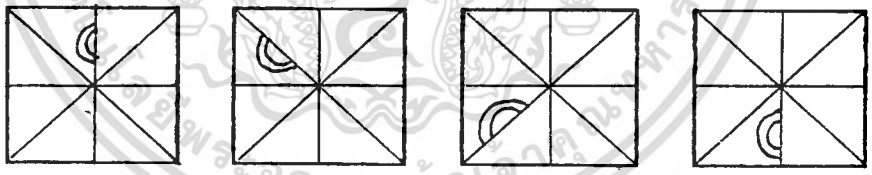
ตัวอย่าง อักษร



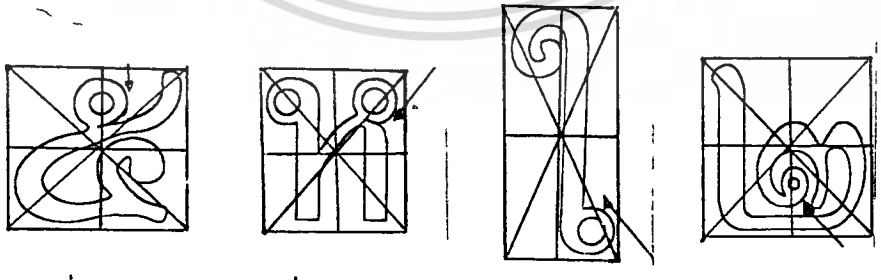
รูปที่ 15.ก รูปที่ 15.ข รูป 15.ค รูป 15.ง

7. stroke คู่ขนานใด ที่กำลังพิจารณาเป็น loop ภายใน octant ที่ 2,3,4 และ 5 มีลักษณะ stroke ที่ หันไปด้านขวา ให้ค่ารหัส A-Code เป็น a_6 ตัวอย่างดังรูปที่ 16.

ลักษณะ stroke



ตัวอย่าง อักษร

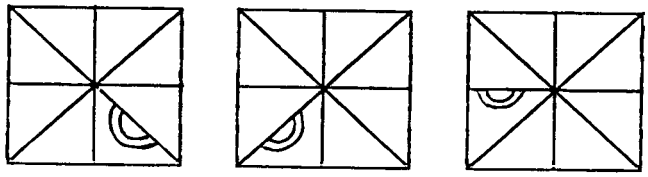


รูปที่ 16.ก รูปที่ 16.ข รูป 16.ค รูป 16.ง

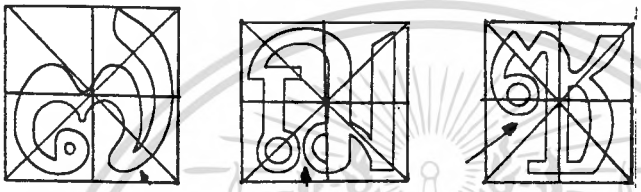
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8. stroke คู่ขนานใดที่กำลังพิจารณาใน octant ที่ 4,5,6 และ 7 หากมีลักษณะเป็น Loop ที่วน ขึ้นไปด้านล่าง ให้ค่ารหัส A-Code เป็น a_7 ตัวอย่างดังรูปที่ 17.

ลักษณะ stroke



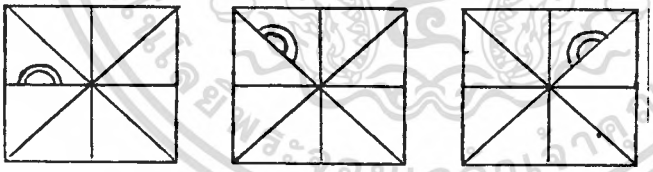
ตัวอย่างอักษร



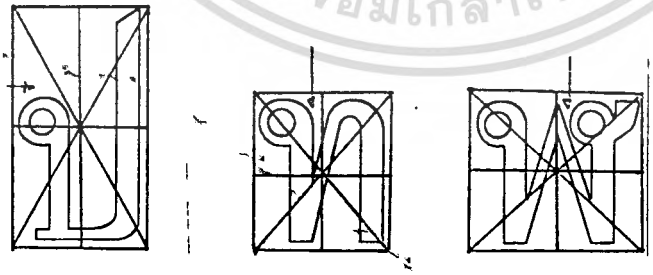
รูปที่ 17.ก รูปที่ 17.ข รูป 17.ค

9. stroke คู่ขนานใดที่กำลังพิจารณาใน octant ที่ 1,2,3 และ 8 หากมีลักษณะเป็น Loop ที่วน ขึ้นไปด้านบน ให้ค่ารหัส A-Code เป็น a_8 ตัวอย่างดังรูปที่ 18.

ลักษณะ stroke



ตัวอย่างอักษร

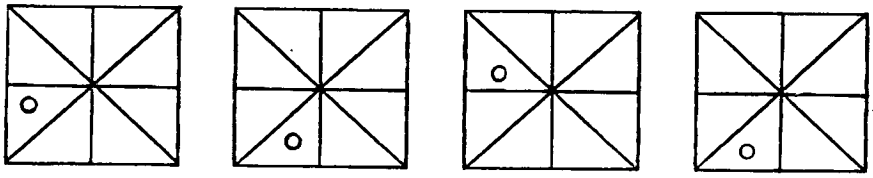


รูปที่ 18.ก รูปที่ 18.ข รูป 18.ค

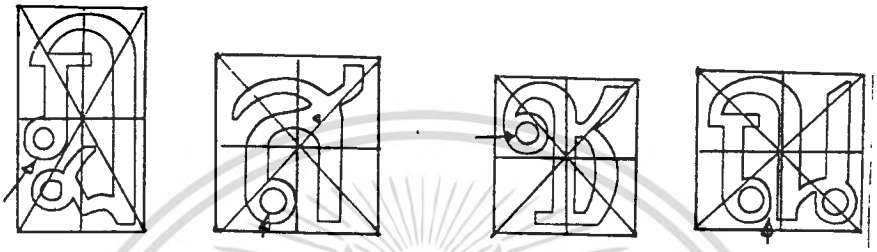
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

10. stroke ที่เกิดขึ้นภายใน octant และสิ้นสุดที่ Octant นั้น ให้ค่ารหัส เป็น a_0 ตัวอย่างดังรูปที่ 19.

ลักษณะ stroke



ตัวอย่าง อักษร

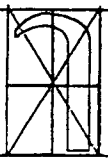


รูปที่ 19.ก รูปที่ 19.ข รูป 19.ค รูป 19.ง

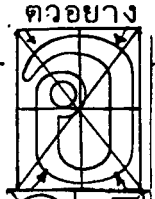
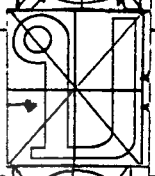
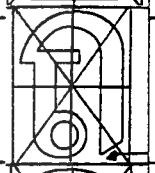
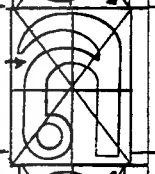
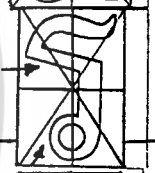
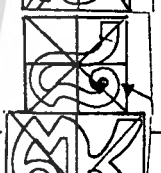
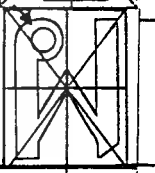
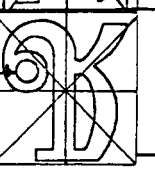
ความหมายของ stroke ที่มีลักษณะเป็น Loop ที่ถูกแทนรหัส A-Code เป็น $a_5, a_6, a_7,$ และ a_8 นั้น โดยทั่วไปมีลักษณะเป็นดังนี้ stroke คู่ขนานใด ๆ ที่เริ่มต้นใน octant พุ่งเข้าหาoctant ที่ติดกันแล้วย้อนกลับมาหา octantเดิมจะถูกกำหนดเป็น loop

สำหรับการนำเงื่อนไขตามทฤษฎีดังกล่าวไปเขียนโปรแกรม ได้แสดงเป็นแผนผังไว้ ดังในรูปที่ 20.(หน้า 28)

และเพื่อให้การพิจารณาได้ชัดเจน ได้แสดงลักษณะของ stroke ที่เกิดขึ้นในแต่ละ Octant แล้วกำหนดให้เป็นค่ารหัส A-Code ดังในตารางที่2.

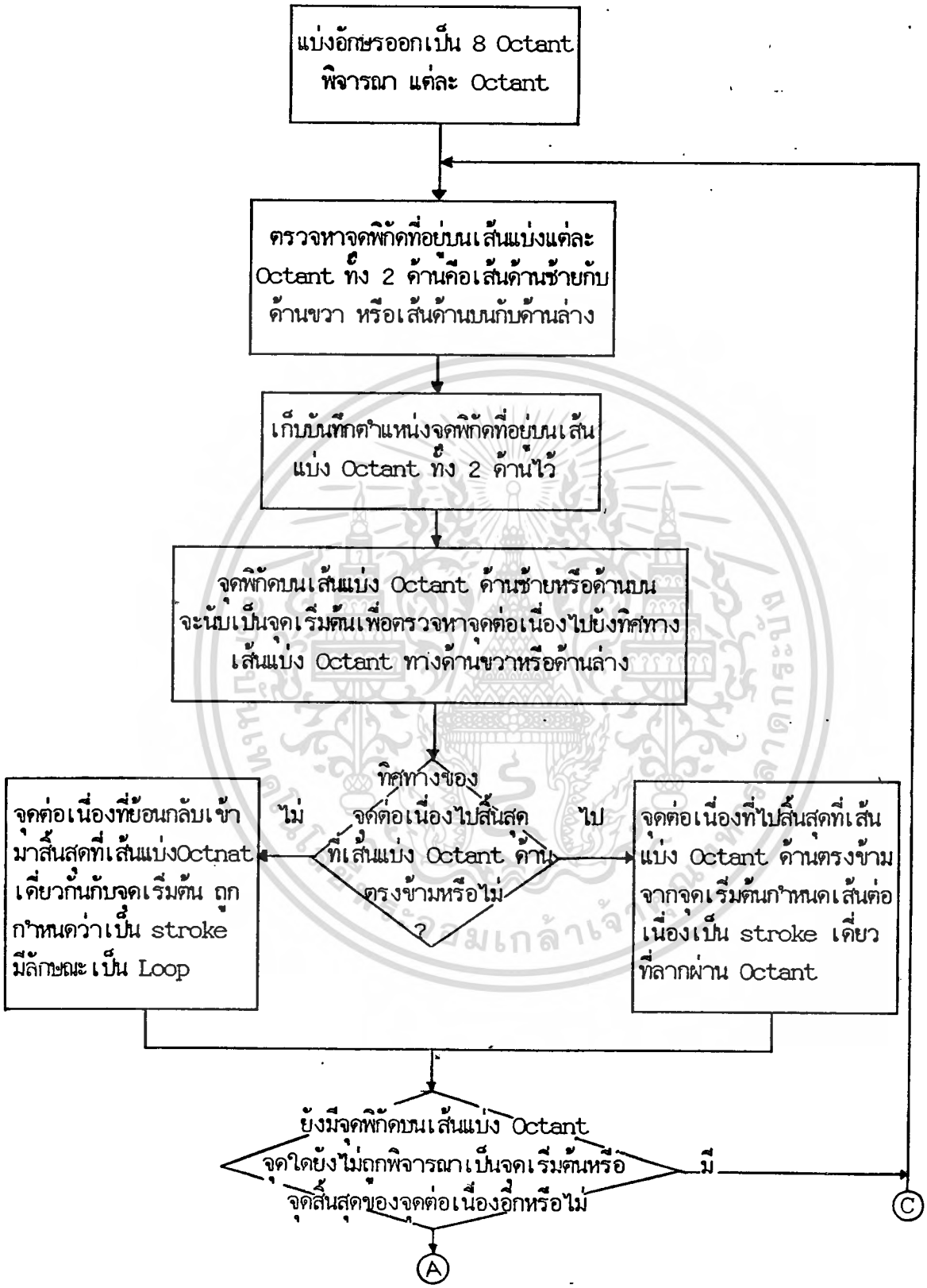
ลักษณะของโทโบโลจิคอล	สัญลักษณ์	คำอธิบาย	ตัวอย่าง
ไม่ stroke	a_0	ไม่มีข้อมูลใด ๆ ใน octant	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

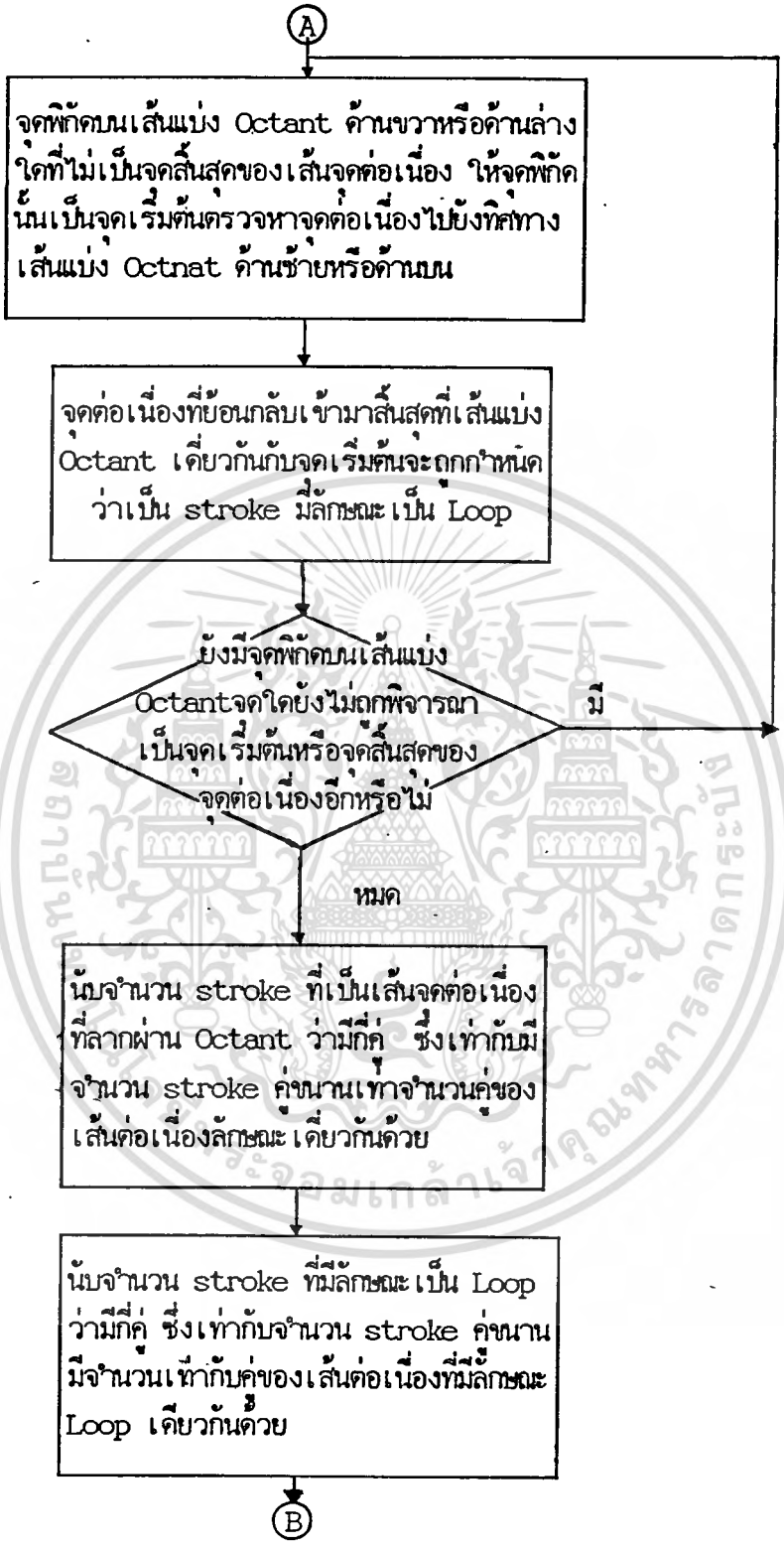
ลักษณะของโทโบโลจิคอล	สัญลักษณ์	คำอธิบาย	ตัวอย่าง
	a_1	stroke คู่ขนานที่ข้าม octant 1,2, 5,6 ประมาณได้ว่าเป็น เส้นแนวตั้ง	
	a_2	stroke คู่ขนานที่ข้าม octant 3,4, 7,8 ประมาณได้ว่าเป็น เส้นแนวนอน	
	a_3	stroke ที่ไม่มีคู่ขนานใน octant และเป็น Loop ในทิศทางตามเข็มนาฬิกา	
	a_4	stroke ที่ไม่มีคู่ขนานใน octant และเป็น Loop ในทิศทางทวนเข็มนาฬิกา	
	a_5	Loop ภายใน octant หันด้านซ้าย พิจารณาเฉพาะ octant 1,6,7,8	
	a_6	Loop ภายใน octant หันด้านขวา พิจารณาเฉพาะ octant 2,3,4,5	
	a_7	Loop ภายใน octant หงายขึ้น พิจารณาเฉพาะ octant 2,3,4,5	
	a_8	Loop ภายใน octant คล่ำลง พิจารณาเฉพาะ octant 1,2,3,8	
	a_9	stroke ที่อยู่ภายใน octant และสิ้นสุดใน Octant นั้น	

ตารางที่ 2. แสดงตัวอย่างการพิจารณาลักษณะ Stroke ในแต่ละ Octant

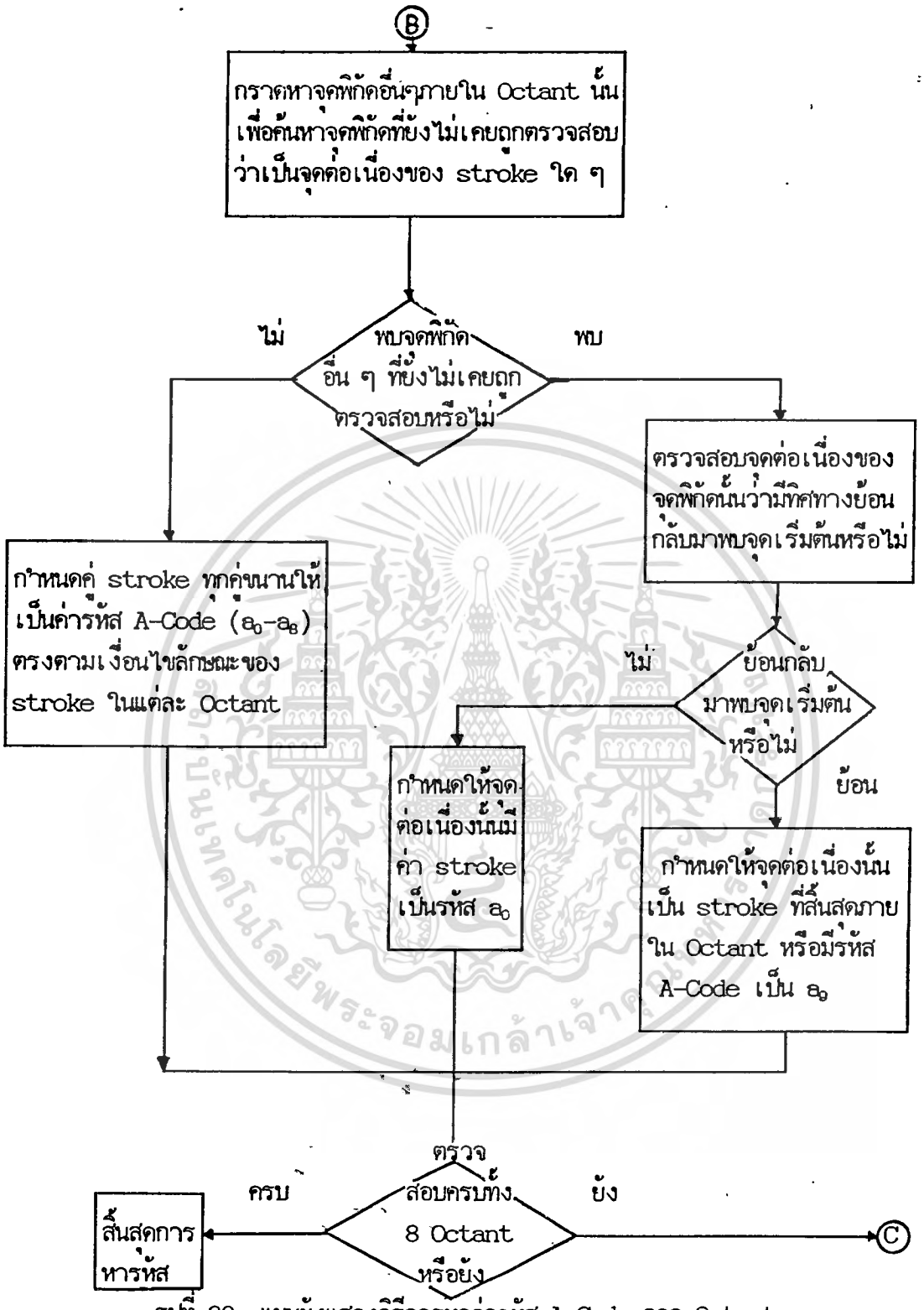
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 20. แผนผังแสดงวิธีการหาค่ารหัส A-Code จาก Octant

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อใช้ basic topological feature เพื่อค้นหา A-Code ในรูปร่างอักษรหนึ่งตัว จะถูกกำหนดรหัสด้วยชุดของรหัสซึ่งเรียกว่า feature vector ซึ่งมีแบบฟอร์ทดังนี้

$$X = (x_{11}, x_{12}, x_{13}; x_{21}, x_{22}, x_{23}; \ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots x_{11}, x_{12}, x_{13}; x_{81}, x_{82}, x_{83}) \ldots\ldots(2)$$

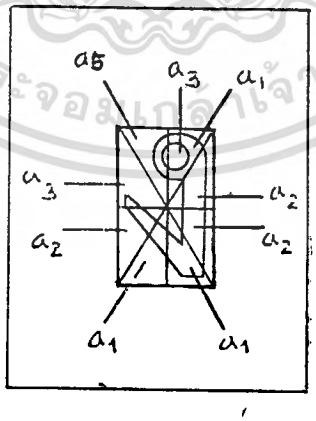
เมื่อ

$$x_{ij} \quad \{ a_k \}, \qquad i = 1,2,3,\ldots 8$$
$$j = 1,2,3$$
$$k = 0,1,2,\ldots 9$$

หรือ $x_{ij} = 0$

ในแต่ละ octant จะยอมให้มีเพียง 3 feature เท่านั้น ค่ารหัสที่ได้จะเก็บเป็นค่าไว้ในค่าคงที่ โดยแทรก 0 ในส่วนที่เหลือจากการแทน Feature vector เพื่อให้ครบจำนวน 3 หลัก และการเรียงรหัสภายในแต่ละ feature จะเรียงเลขน้อยไปหาเลขมากยกเว้นเลข 0 ซึ่งจะต้องแทรกต่อท้าย feature เลขที่มีค่ามากกว่าเสมอ ตัวอย่าง feature Vector ของอักษรตัว ง. ในรูปที่ 21.และค่าของfeature Vector ในชุดของ A-CODE ปรากฏในสมการที่ 3.

$$X = a_5,0,0; a_1,a_3,0; a_2,0,0; a_1,0,0;$$
$$a_1,0,0; a_1,0,0; a_2,0,0; a_3,0,0) \qquad \ldots\ldots(3)$$



รูปที่ 21. ตัวอย่างการหารหัสชุด A-CODE จากลักษณะ stroke ของอักษรตัว ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4 การหาจุดศูนย์ถ่วง (Center of Gravity)[18]

เมื่อทำการอ่านภาพตัวอักษรจากเครื่อง Scanner ได้ภาพของตัวอักษรที่มีลักษณะเป็น binary เก็บอยู่ในหน่วยความจำแล้ว (จากรูปที่ 3.) ข้อมูลประกอบด้วยเลข 0 ซึ่งเป็นค่าแทนจุดที่เป็นพื้น และเลข 1 เป็นค่าแทนจุดที่เป็นเนื้อของตัวอักษร การคำนวณจุดศูนย์ถ่วงของตัวอักษร หาได้จากสมการดังนี้

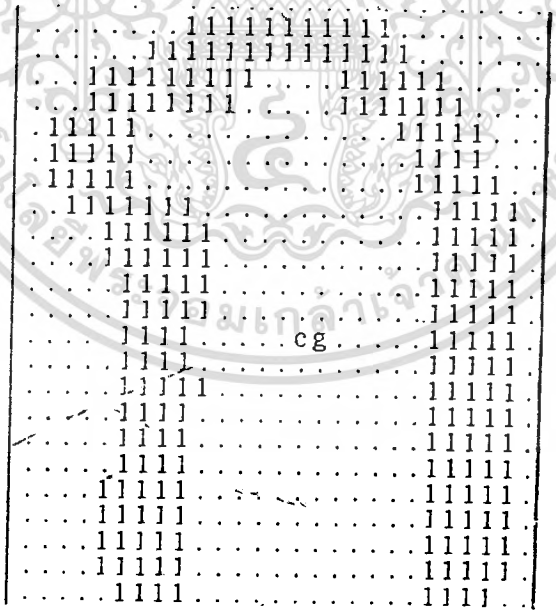
$$Ix = \frac{\sum_{j}\sum_{i} I.F(i,j)}{\sum_{j}\sum_{i} F(i,j)} \qquad \ldots\ldots (4)$$

$$Iy = \frac{\sum_{i}\sum_{j} J.F(i,j)}{\sum_{i}\sum_{j} F(i,j)} \qquad \ldots\ldots (5)$$

เมื่อ F(i,j) คือจุดใด ๆ ของตัวอักษรที่มีค่าเป็น 0 หรือ 1

Ix , Iy คือตำแหน่งของจุดศูนย์ถ่วง

จุดศูนย์ถ่วงที่คำนวณได้ นำไปเป็นอ้างอิงในการตรวจสอบหาค่า Concentrat Code (Q-Code) ดังตัวอย่างภาพจุดศูนย์ถ่วงที่คำนวณได้ในรูปที่ 22.



รูปที่ 22.ภาพแสดงจุดศูนย์ถ่วงของตัวอักษร ก.ซึ่งคำนวณได้จากสมการ 4 และ 5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5 การหาขอบภาพ (EDGE DETECTION)[23]

เทคนิคการหาขอบภาพ เริ่มโดยการกวาดภาพทางแนวนอน จากมุมบนซ้ายไปยังมุมล่างขวา โดยมีเงื่อนไขของจุดที่จะเป็นขอบว่า อย่างน้อย 1 ในจุดที่ติดกันจะต้องเป็นพื้นขาว และเป็นส่วนหนึ่งของภาพตัวมันเอง

ให้ P_0 เป็นจุดใด ๆ ของภาพเนื้อตัวอักษรที่มีค่าบิทเป็น 1 พิจารณาค่าของบิทที่เป็น 1 และ 0 ที่อยู่รอบ ๆ จุด P_0 ตามลักษณะตารางเมตริกซ์ขนาด 3 x 3 ดังตารางที่ 3 และตัวอย่างการคำนวณหาขอบภาพดังรูปที่ 23. และรูปที่ 24.

P_4	P_3	P_2
P_5	P_0	P_1
P_6	P_7	P_8

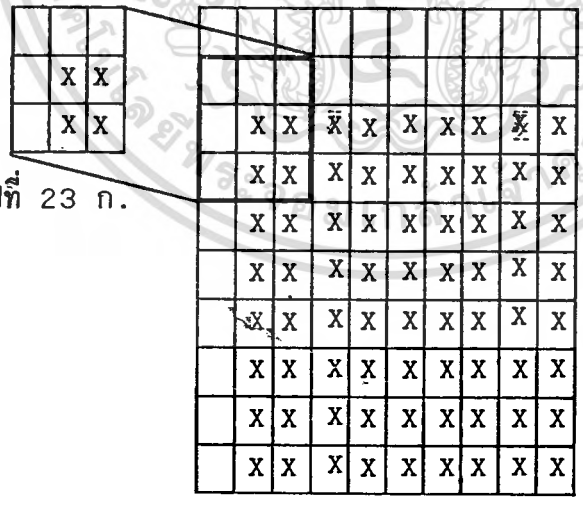
ตารางที่ 3. ตารางเมตริกซ์ 3 x 3

$$X = P_0.\overline{(P_1.P_3.P_5.P_7)}.(P_1+P_2+P_3+P_4+P_5+P_6+P_7+P_8) \quad .....(6)$$

เมื่อ $P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8$ เป็นจุดที่อยู่รอบ ๆ จุด P_0 มีค่าเป็น 0 หรือ 1

X : เป็นจุดขอบของภาพ ที่ถูกพิจารณาตามเงื่อนไขสมการที่ 6 แล้วนำไปแทนที่ตำแหน่งเดียวกันกับจุด P_0

3.5.1 ตัวอย่างการคำนวณขอบภาพ พิจารณาจุดที่เป็นขอบ



รูปที่ 23 ก.

รูปที่ 23 ข.

รูปที่ 23. ภาพตัวอย่างจุดที่พิจารณาบริเวณที่เป็นขอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 23 ก. แทนค่าจุด P_0 ถึง P_8 ลงในสมการที่ 6.

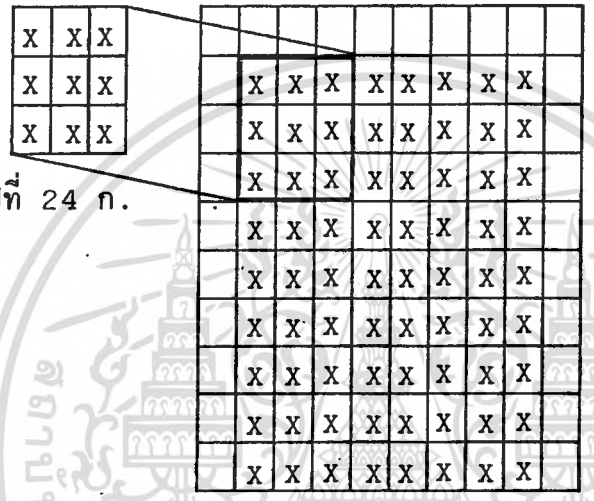
$X = P_0.(\overline{P_1.P_3.P_5.P_7}).(P_1+P_2+P_3+P_4+P_5+P_6+P_7+P_8)$

$X = 1\ .(\overline{1.0.0.1}).(1+0+0+0+0+0+1+1)$

$X = 1.(\overline{0}).(1)$

$X = 1$

3.5.2 <u>ตัวอย่างการคำนวณขอบภาพ พิจารณาจุดที่ไม่ใช่ขอบ</u>



รูปที่ 24 ก.

รูปที่ 24 ข.

รูปที่ 24. ภาพตัวอย่างจุดที่พิจารณาบริเวณที่ไม่ใช่ขอบ

จากรูปที่ 24 ก. แทนค่าจุด P_0 ถึง P_8 ลงในสมการที่ 6.

$X = P_0.(\overline{P_1.P_3.P_5.P_7}).(P_1+P_2+P_3+P_4+P_5+P_6+P_7+P_8)$

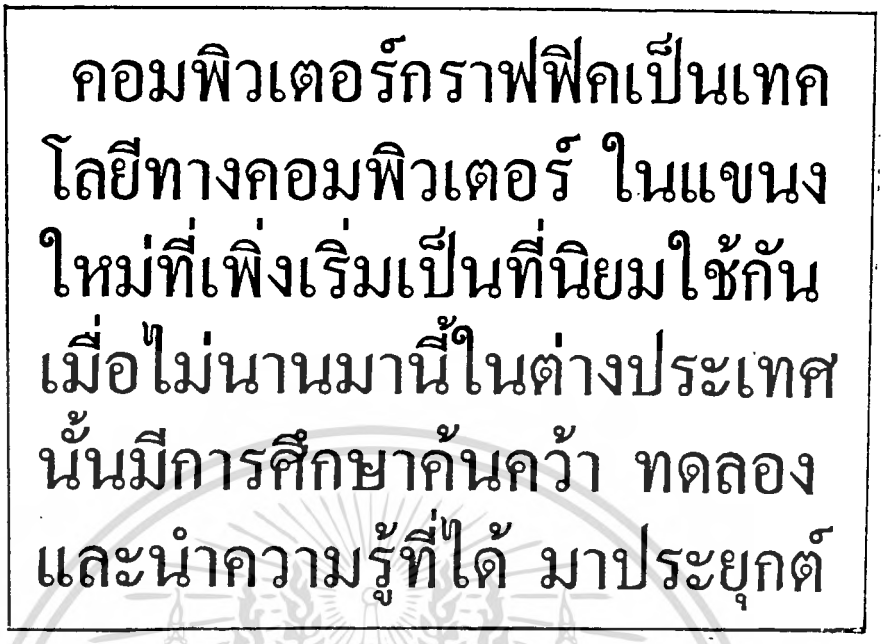
$X = 1\ .(\overline{1.1.1.1}).(1+1+1+1+1+1+1+1)$

$X = 1.(\overline{1}).(1)$

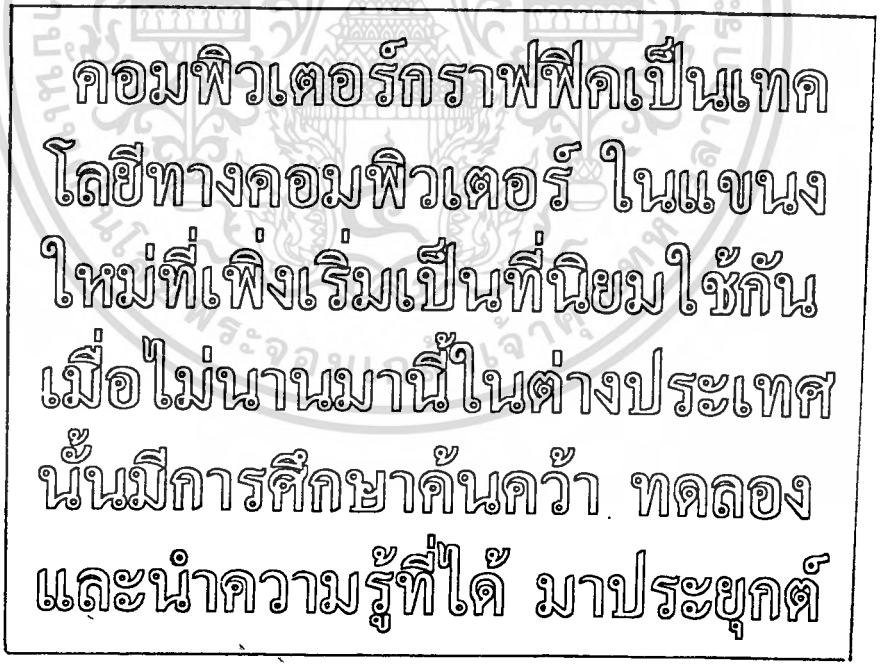
$X = 0$

ด้วยเทคนิคการหาขอบภาพด้วยวิธีนี้ จุดรบกวน (niose) ซึ่งกระจายอยู่ตามส่วนต่าง ๆ ของภาพ จะถูกกำจัดไปด้วย ตัวอย่างภาพที่ยังไม่นำไปคำนวณหาขอบภาพดังแสดงในรูปที่ 25. และตัวอย่างภาพผลการหาขอบภาพ แสดงไว้ในรูปที่ 26.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 25. แสดงลักษณะภาพที่ยังไม่ถูกคำนวณหาขอบภาพ



รูปที่ 26. แสดงลักษณะผลการคำนวณขอบภาพของตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.6 การหาขนาดความกว้างและความสูงของตัวอักษร[18]

ภาพตัวอักษรในหน่วยความจำจะถูกตรวจสอบทีละตัว โดยการกวาดทางแนวตั้งจากมุมล่างซ้าย ไปยังมุมบนขวาเมื่อพบจุดแรกจะถือว่าเป็นส่วนหนึ่งของอักษรตัวแรก ตรวจหาจุด Coordinate ที่มีค่าต่ำสุดและสูงสุดของอักษรตัวนั้น เพื่อหาขนาดความกว้างและความสูง โดย

ให้ L_n : คือจุดปัจจุบันใด ๆ บนเส้นขอบรูปอักษรแต่ละตัว

$n = 1,2,3,4,5 \ldots\ldots N$

$n = 1$: ถ้าเป็นจุดแรกของตัวอักษร

$n = N$: ถ้าเป็นจุดสุดท้ายของอักษรตัวเดียวกัน

x,y : เป็นจุด Coordinate ของ L_n จะต้องบันทึกไว้ทุกจุด

L_{n+1} : เป็นจุดที่จะบันทึกค่า x,y จุดต่อไป

L_{n-1} : เป็นจุดอดีต ที่เพิ่งผ่านการบันทึกค่า x,y

X_{max} : ค่า Coordinate ของ x ที่มีค่าสูงสุด

X_{min} : ค่า Coordinate ของ x ที่มีค่าต่ำสุด

Y_{max} : ค่า Coordinate ของ y ที่มีค่าสูงสุด

Y_{min} : ค่า Coordinate ของ y ที่มีค่าต่ำสุด

เขียนโปรแกรม เพื่อหาค่า X_{max} ,X_{min} ,Y_{max} ,Y_{min} ของอักษรแต่ละตัวพร้อมทั้งบันทึกค่า Coordinate x,y ของ L_n ทุกจุด ซึ่งสามารถหาขนาดความกว้างและความสูงของอักษรแต่ละตัวได้จาก สมการที่ 7 และ 8

$$I = X_{max} - X_{min} + 1 \qquad \ldots\ldots\ldots(7)$$

$$J = Y_{max} - Y_{min} + 1 \qquad \ldots\ldots\ldots(8)$$

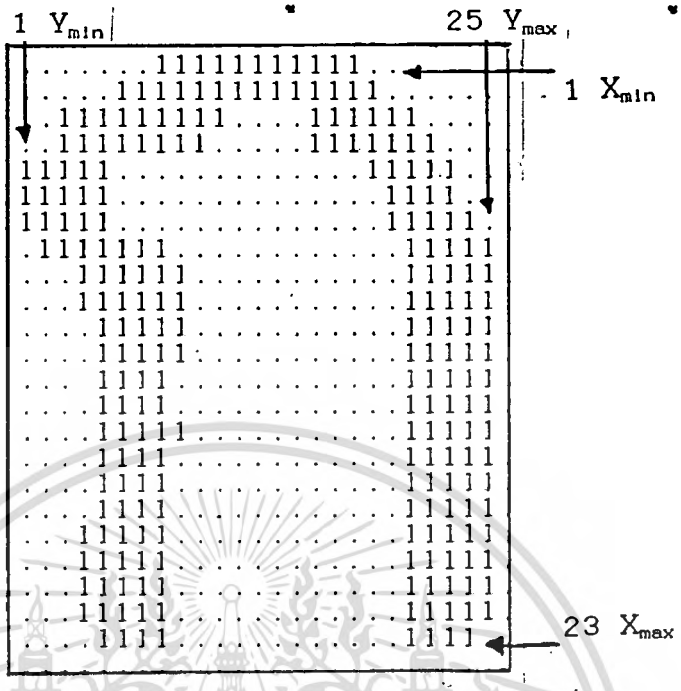
เมื่อ

I : คือขนาดความกว้างของอักษรแต่ละตัว

J : คือขนาดความสูงของอักษรแต่ละตัว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างการคำนวณหาค่าขนาดความกว้างความสูงของตัวอักษร ดังแสดงในรูปที่ 27.



รูปที่ 27. ลักษณะจุด Coordinate ของตัวอักษร ที่ถูกนำไปคำนวณหาค่าความกว้างและความสูง

จากรูปที่ 27. ได้ค่าที่จุด Coordinate สูงสุดและต่ำสุดดังนี้

$X_{max} = 25$

$X_{min} = 1$

$Y_{max} = 23$

$Y_{min} = 1$

แทนค่าลงในสมการที่ 7.

$$I = X_{max} - X_{min} + 1$$
$$= 25 - 1 + 1$$
$$= 25$$

และ
$$J = Y_{max} - Y_{min} + 1$$
$$= 23 - 1 + 1$$
$$= 23$$

ดังนั้นอักษรจากรูปที่ 27. จึงมีขนาดความกว้างเท่ากับ 25 และความสูงเท่ากับ 23

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
การจัดเตรียมข้อมูลเพื่อการรู้จำ

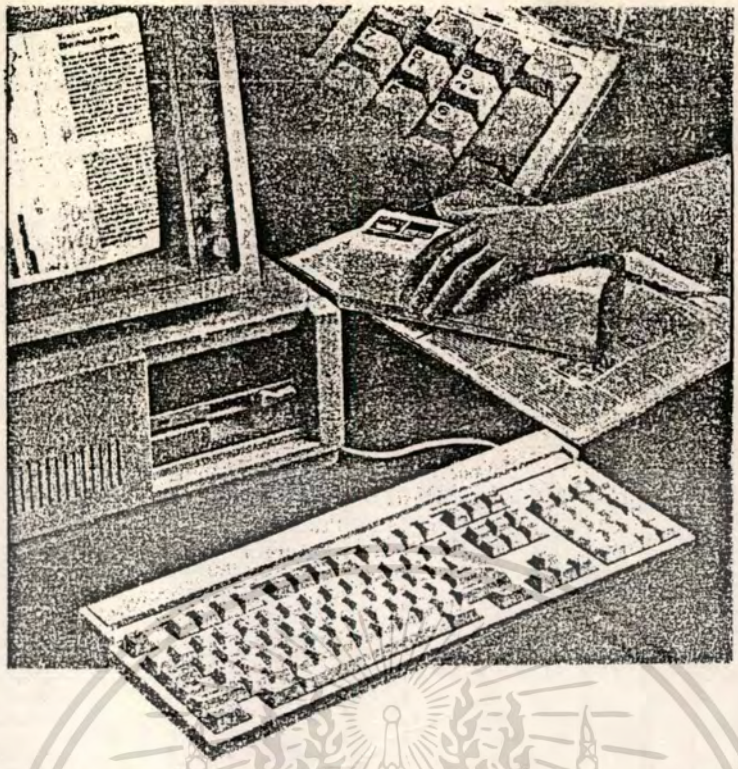
ข้อมูลภาพตัวอักษรที่นำมาใช้ในการรู้จำวิธีการนี้ สามารถเลือกภาพอักษรตัวพิมพ์ภาษาไทยที่มีอยู่ทั่วไป ในลักษณะที่เป็นรูปประโยคหรือบทความ ตามหน้าหนังสือพิมพ์ วารสาร หรือเอกสารทั่วไป ดังนั้นในการนำภาพข้อความเหล่านั้นไปดำเนินการในขบวนการรู้จำ จะต้องผ่านขั้นตอนการอ่านภาพข้อความด้วยเครื่อง Scanner ลงสู่หน่วยความจำของเครื่องไมโครคอมพิวเตอร์ และหรือนำข้อมูลภาพเหล่านั้นบันทึกเก็บในรูปของแฟ้มข้อมูลภาพลงในแผ่นดิส ภาพข้อความเหล่านั้นถึงแม้ได้เก็บบันทึกลงในหน่วยความจำทั้งบทความแล้ว แต่การประมวลผลตามวิธีการนี้จะกระทำได้เพียงทีละตัวอักษรเท่านั้น จำเป็นต้องทำการแยกภาพตัวอักษรแต่ละตัวออกจากบทความ ซึ่งเมื่อสามารถวิเคราะห์ตัวอักษรได้ครบทั้งหมดแล้วจะต้องนำรหัส ASCII ของตัวอักษรเหล่านั้นกลับมาเรียงลำดับตำแหน่งเดิมตามลักษณะไวยากรณ์ที่ถูกต้องเหมือนต้นฉบับด้วย

4.1 การจัดเก็บข้อมูลภาพ

เครื่อง Scanner ที่ใช้ในการทดลองครั้งนี้ เป็น Scanner ชนิด Handy Scanner มีขนาดเล็กกระทัดรัดสามารถจับเคลื่อนย้ายเพื่อลบผ่านข้อความที่ต้องการ จากสิ่งพิมพ์ต่าง ๆ ได้อย่างสะดวกและรวดเร็ว โดยให้รายละเอียดของจุดภาพจาก 100 - 400 dpi ลักษณะของเครื่อง Handy Scanner ดังกล่าวแสดงไว้ในรูปที่ 28. และลักษณะการใช้งานแสดงไว้ในรูปที่ 29.

รูปที่ 28. ลักษณะรูปร่างของเครื่อง Handy Scanner

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 29. ภาพแสดงลักษณะการใช้งานของเครื่อง Handy Scanner

บทความจากสิ่งพิมพ์ จะถูกอ่านจากเครื่อง Handy Scanner โดยมีโปรแกรมคอยควบคุม ให้ได้ภาพปรากฏที่จอไมโครคอมพิวเตอร์ที่สัมพันธ์กับการลากเครื่อง Scanner ผ่านข้อความเหล่านั้น พร้อมกับภาพข้อความเหล่านั้นจะถูกจัดเรียงเป็นลักษณะ bit map ไว้ในหน่วยความจำไปด้วย เมื่อสิ้นสุดหน้ากระดาษ หรือหยุดการอ่านภาพ เราสามารถจัดเก็บภาพข้อความเหล่านั้นลงในแผ่นดิส โดยโปรแกรมได้

4.2 ลักษณะเฉพาะของข้อความที่เครื่อง Handy Scanner สามารถอ่านได้

เนื่องจากเครื่อง Handy Scanner ถูกออกแบบให้ใช้งานอ่านข้อความในพื้นที่ความกว้างจำกัด เพียง 5 เซนติเมตรเท่านั้น ดังนั้นลักษณะข้อความที่จะนำมาให้เครื่อง Handy Scanner อ่าน ควรเป็นลักษณะดังตัวอย่าง แสดงในรูปที่ 30.

รูปที่ 30.
ลักษณะข้อความที่เครื่อง
Handy Scanner
สามารถอ่านได้

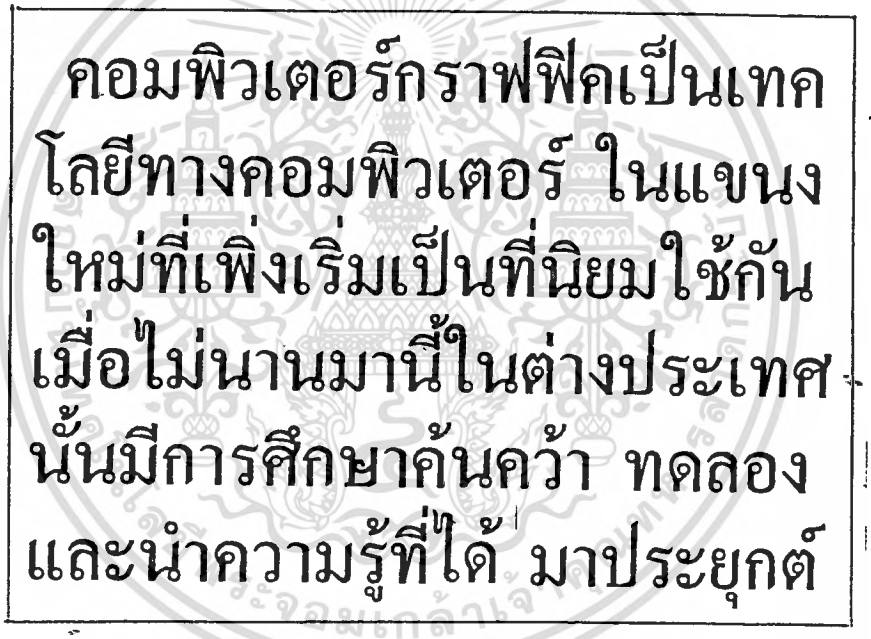
คอมพิวเตอร์กราฟฟิคเป็นเทคโลยีทางคอมพิวเตอร์ในแขนงใหม่ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้า ทดลองและนำความรู้ที่ได้ มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพ มีความเด่นชัดสา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 เทคนิคการแยกภาพตัวอักษรออกจากประโยค[17]

การแยกภาพตัวอักษรภาษาไทยออกจากภาพของประโยค เป็นขั้นตอนหนึ่งสำหรับการจัดการล่วงหน้าให้กับขบวนการรู้จำ ทั้งนี้เพราะเมื่อเครื่อง Scanner ได้อ่านภาพข้อความจากกระดาษเข้ามา ภาพของข้อความเหล่านั้นจะยังเป็นลักษณะกลุ่มภาพอักษรเรียงต่อเนื่องกันเป็นประโยค แต่ขบวนการรู้จำจะทำการวิเคราะห์ข้อมูลได้เพียงที่ละหนึ่งภาพตัวอักษรเท่านั้น[16] ดังนั้นการแยกภาพตัวอักษรออกจากภาพประโยค จำเป็นต้องศึกษาวิธีการและเทคนิคเฉพาะ เพื่อให้การจัดการล่วงหน้าได้เตรียมข้อมูลสำหรับการวิเคราะห์ให้แก่ขั้นตอนการรู้จำได้ถูกต้องและรวดเร็ว

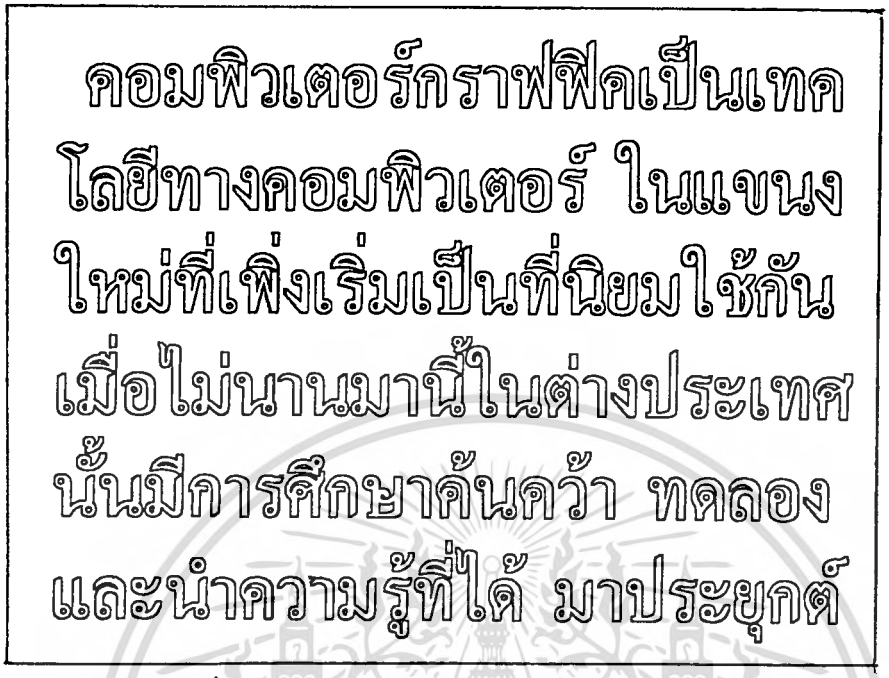
4.3.1 การทำขอบภาพ ภาพอักษรที่เรียงกันเป็นประโยคข้อความซึ่งได้จากการอ่านด้วยเครื่องScannerจะถูกนำมาแสดงให้ปรากฏขึ้นที่จอภาพมอนอเตอร์ดังรูป 31.



รูปที่ 31.ภาพประโยคที่ปรากฏในจอมอนิเตอร์ที่อ่านเข้ามาด้วยเครื่องScanner

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทำการหาขอบภาพของกลุ่มภาพประโยค จากรูปที่ 31. จะปรากฏผลดังรูปที่ 32.

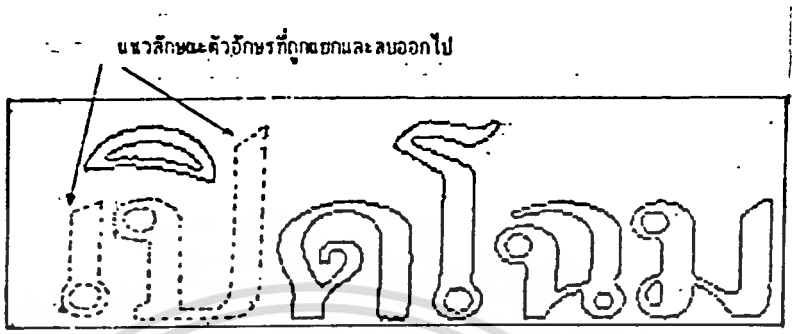


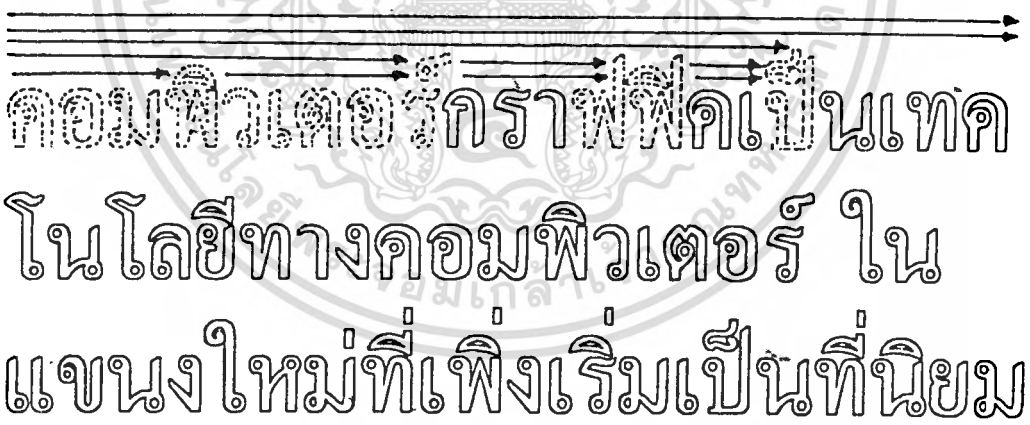
รูปที่ 32. ผลการหาขอบภาพของภาพประโยค

4.3.2 <u>การคัดลอกและย้ายจุดพิกัดของตัวอักษร</u>

โดยทั่วไปลักษณะของอักษรภาษาไทยจะเรียงกันอยู่หลายระดับ เช่นอักษรสระ สระ จะอยู่ระดับล่างสุด อักษรที่เป็นพยัญชนะจะอยู่ระดับกลาง และสระที่อยู่ระดับบนหรือบนสุดเป็นต้น การพิจารณาหาขอบเขตกรอบของตัวอักษรโดยวิธีกวาดภาพธรรมดาอาจไม่สามารถคำนวณหาขนาดความกว้างความสูงของตัวอักษรที่แท้จริงได้ ตัวอย่างเช่น คำว่า "เปิด" ในรูป 33 ก. ขอบเขตของอักษร ป และอักษร ิ ซึ่งซ้อนทับกันอยู่หากกวาดหาขนาดความกว้างความสูงของตัวอักษรโดยตรงจะได้ค่าขนาดของอักษรทั้งสองตัวที่ซ้อนทับกันในบริเวณเดียวกัน[17] แต่ถ้าพิจารณาเฉพาะเส้นที่เป็นของอักษรจากอักษรแต่ละตัว และไต่ตามขอบตัวอักษรเพื่อทราบค่าพิกัดจุดภาพทุกจุดของอักษรตัวนั้นทำการคัดลอกและย้ายพิกัดจุดภาพแยกไว้ต่างหากและลบจุดพิกัดบริเวณภาพอักษรเดิมนั้นให้ว่างไป ภาพอักษรที่ถูกคัดลอกมาจะถูกนำไปคำนวณหาขนาดความกว้างความสูงของตัวอักษร และบริเวณส่วนว่างก็สามารถกวาดภาพค้นหาพิกัดจุดภาพอักษรอื่นต่อไปได้โดยไม่มีผลต่อตัวอักษรที่อยู่ข้างเคียงดังแสดงไว้ในรูปที่ 33ก. ในการกวาดหาจุดใด ๆ ของตัวอักษรจะเริ่มกวาดหาจากมุมบนซ้าย ไปยังมุมล่างขวาพบส่วนใดส่วนหนึ่งของอักษรตัว ก็ทำการคัดลอกและย้ายจุดพิกัดออกไป ดังแสดงในรูป 33ข.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

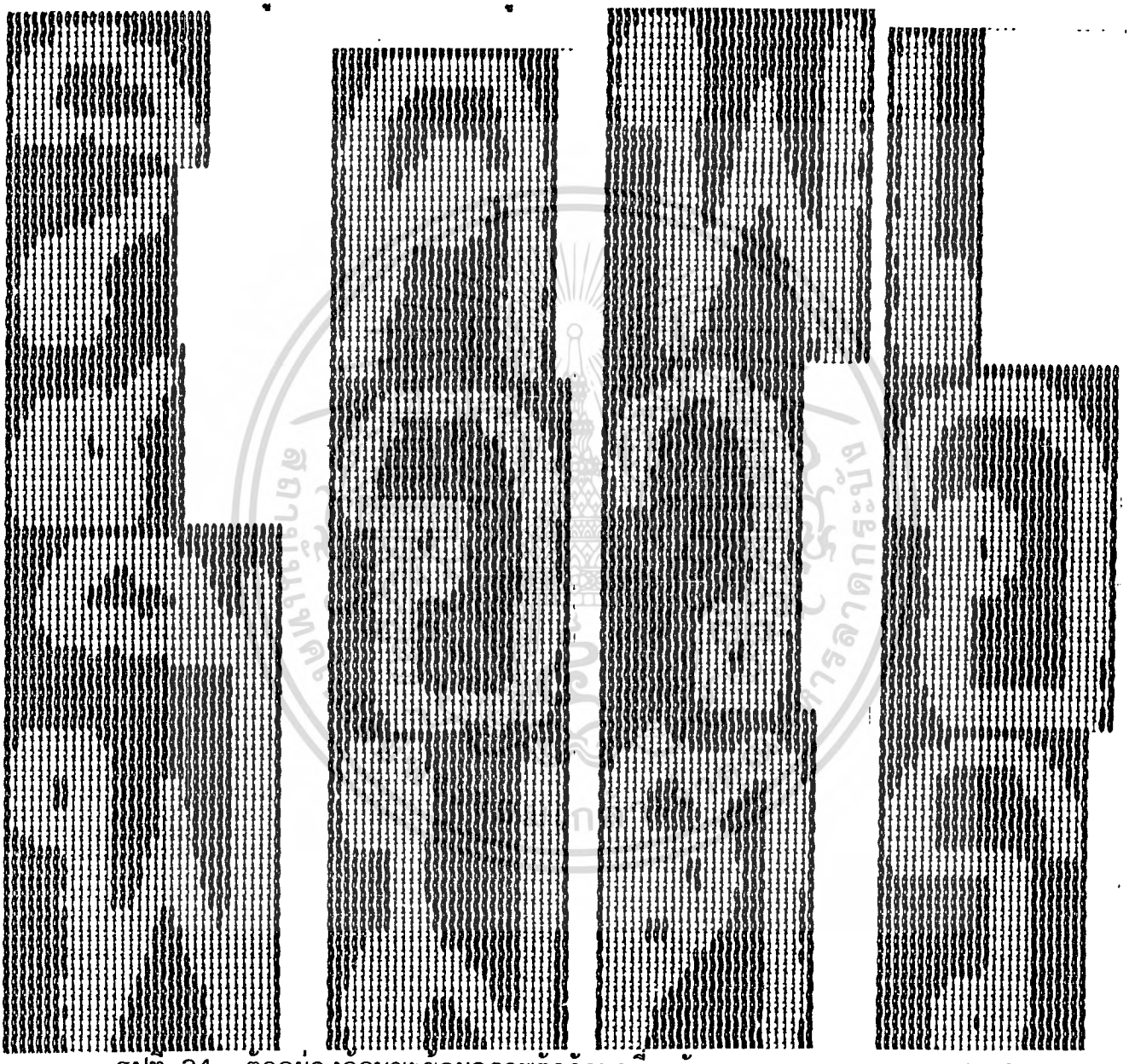


รูปที่ 33 ก. ลักษณะการกวาดหาพิกัดจุดภาพเพื่อคัดลอก

รูปที่ 33 ข. การค้นหาจุดภาพอักษรตัวถัดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.3 <u>การจัดส่งข้อมูลไปภาคการวิเคราะห์</u>

เมื่อทำการแยกภาพอักษรแต่ละตัวออกจากภาพประโยคได้แล้ว อักษรแต่ละตัวจะถูกส่งไปให้ภาควิเคราะห์การรู้จำ เพื่อจัดแยกกลุ่มและแยกตัวอักษรออกจากกลุ่ม ลักษณะของข้อมูลที่ใช้วิเคราะห์ จะเว้นขอบรอบตัวอักษรด้านละ 1 จุดพิกัด ดังตัวอย่างลักษณะข้อมูลตัวอักษร แสดงในรูปที่ 34.

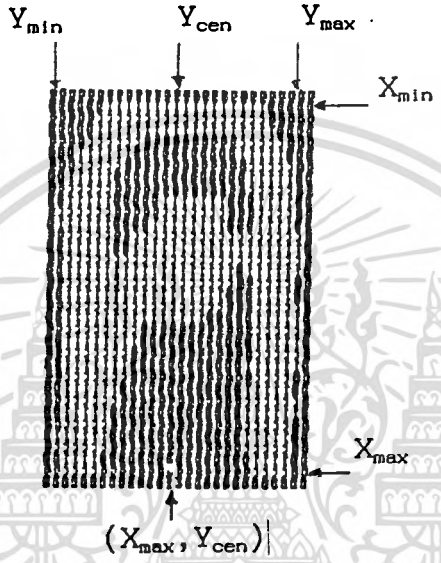


รูปที่ 34. ตัวอย่างลักษณะข้อมูลภาพตัวอักษรที่ถูกคัดลอกแยกออกจากภาพประโยค เตรียมส่งให้ภาควิเคราะห์การรู้จำต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.4 <u>การจัดเรียงรูปประโยคหลังการวิเคราะห์</u>

หลังจากแยก ข้อมูลภาพตัวอักษรแต่ละตัวออกมาได้แล้ว ก่อนที่จะถูกส่งต่อไปยังภาคการวิเคราะห์ตัวอักษร จะต้องบันทึกค่าตำแหน่งจุดพิกัด (X_{max}, Y_{cen}) ซึ่งเป็นจุดพิกัด บริเวณกึ่งกลางฐานข้อมูลภาพตัวอักษร เก็บไว้ทุกตัว ดังแสดงในรูปที่ 35.



X_{man} : ค่า Coordinate ของ X ที่มีค่าสูงสุด

$Y_{cen} = (Y_{max} - Y_{min})/2$

รูปที่ 35. ตำแหน่งจุดพิกัดบริเวณ กึ่งกลางฐานข้อมูลภาพตัวอักษร

และเมื่อตัวอักษรถูกส่งไปวิเคราะห์ จนทราบว่าเป็นอักษรตัวอะไรแล้ว รหัส ASCII ของอักษรตัวนั้นจะถูกนำมาจัดเรียงลำดับ ตามค่า Y_{cen} ของอักษรแต่ละตัวจากน้อยไปหามากเพื่อจัดเรียงตัวอักษรให้อยู่เรียงกันคล้ายตำแหน่งรูปประโยคเดิม และจัดเรียงลำดับค่า X_{man} จากน้อยไปหามาก สำหรับเรียงบรรทัดตั้งแต่บรรทัดบนสุดไปหาล่างสุด เพื่อให้ตัวอักษรอยู่เรียงบรรทัดคล้ายรูปประโยคเดิม ดังได้แสดงลำดับขั้นและเงื่อนไขการจัดเรียงประโยคหลังการวิเคราะห์ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นที่ 1. ตัวอย่างข้อความเดิมเมื่อยังไม่ถูกจัดแยกตัวอักษรออกจากประโยค

คอมพิวเตอร์กราฟฟิคเป็นเทค (บรรทัดที่ 1)
โนโลยีทางคอมพิวเตอร์ในแขนง (บรรทัดที่ 2)
ใหม่ ที่เพิ่งเริ่มเป็นที่นิยมใช้กัน (บรรทัดที่ 3)
สามารถสื่อความหมาย ข้อมูลที่ (บรรทัดที่ 4)
เป็นตัวเลขหรือตัวหนังสือ (บรรทัดที่ 5)

ขั้นที่ 2. ตัวอย่างข้อความบรรทัดแรก เมื่อถูกแยกตัวอักษรออกจากประโยคและหลังจากวิเคราะห์จนทราบรหัส ASCII ของตัวอักษรแล้ว ตัวอักษรจะยังถูกจัดเรียงตามลำดับก่อนหลังการถูกวิเคราะห์ จนถึงอักษรตัวสุดท้ายของบรรทัด

[illegible]ฟฟปคอมพวเตอรกราคเนเทค

ขั้นที่ 3. และเมื่อเริ่มแยกอักษรตัวแรกของบรรทัดที่ 2 ค่าจุดพิกัด X_{max} ของอักษรตัว "โ" จะแตกต่างจากค่าจุดพิกัด X_{max} ของอักษร "ค" ซึ่งเป็นอักษรตัวสุดท้ายของบรรทัดแรก ทำให้ทราบได้ว่าอักษรตัว "โ" จะต้องขึ้นต้นบรรทัดใหม่เป็นบรรทัดที่ 2 จากนั้นทำการจัดเรียงอักษรทุกตัวในบรรทัดที่ 1 ตามค่าจุดพิกัด Y_{cen} ของอักษรแต่ละตัวจะได้ประโยคข้อความในบรรทัดแรกคล้ายประโยคเดิม

คอมพิวเตอร์กราฟฟิคเป็นเทค (บรรทัดที่ 1)
โ (บรรทัดที่ 2)

ในการตัดสินใจเพื่อขึ้นบรรทัดใหม่จะกำหนดเงื่อไขดังนี้

1. เมื่อค่าจุดพิกัด X_{max} ของพยัญชนะอักษรตัวหลังต้องมีค่าต่างจากจุดพิกัดของพยัญชนะอักษรตัวหน้ามากว่า 1 ส่วน 4 ของความสูงตัวอักษรตัวหน้า

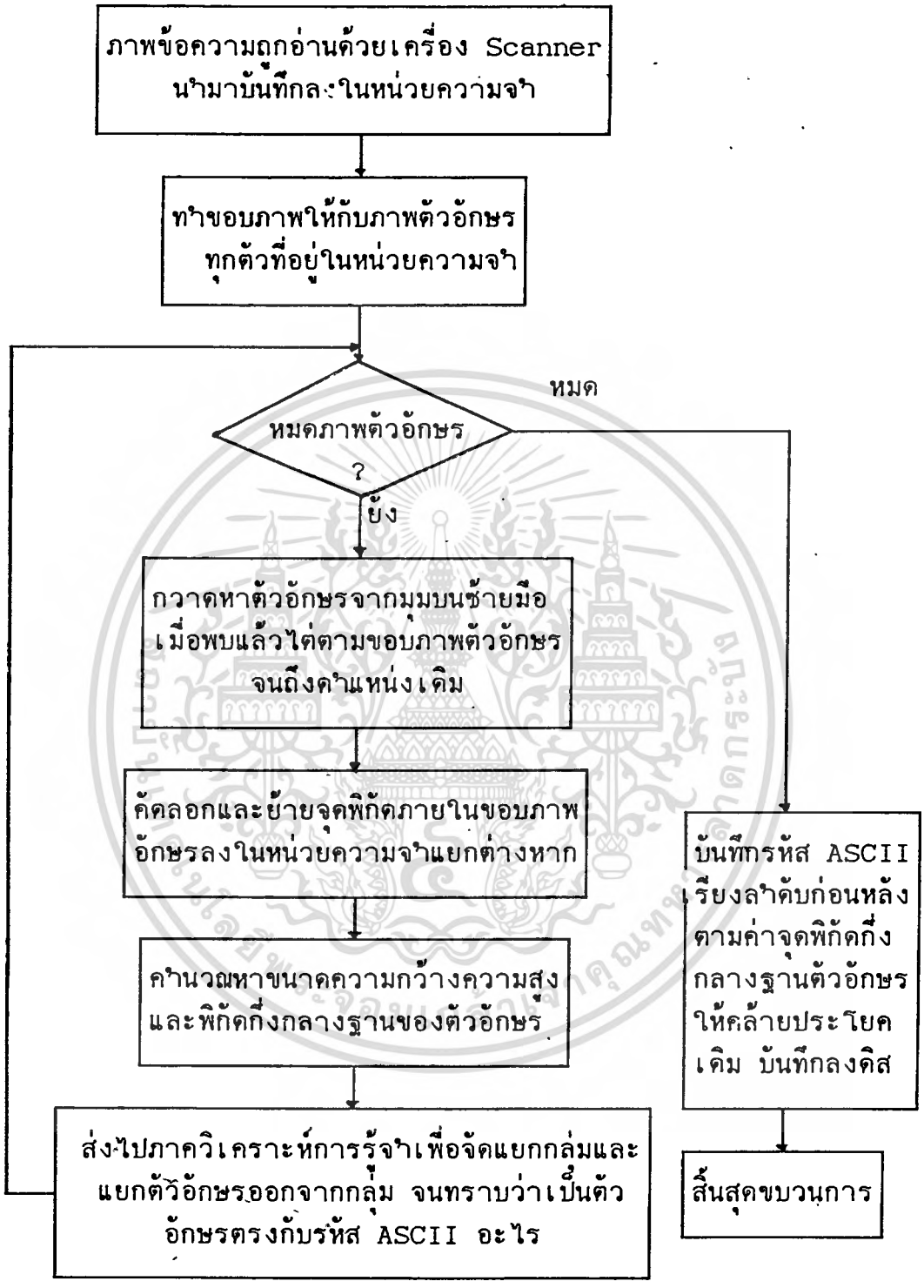
2. เมื่ออักษรตัวหน้าเป็นพยัญชนะและอักษรตัวหลังเป็นสระบน เช่น

[illegible]ชหมทเพงเรมเนทนยมกัน (บรรทัดที่ 3)
(บรรทัดที่ 4)

3. เมื่ออักษรตัวหน้าเป็นสระล่างและอักษรตัวหลังเป็นสระบน เช่น

[illegible]สามารถสอความหมายขอมลท (บรรทัดที่ 4)
(บรรทัดที่ 5)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



สรุปลำดับขั้นตอนการแยกภาพตัวอักษรออกจากภาพประโยค

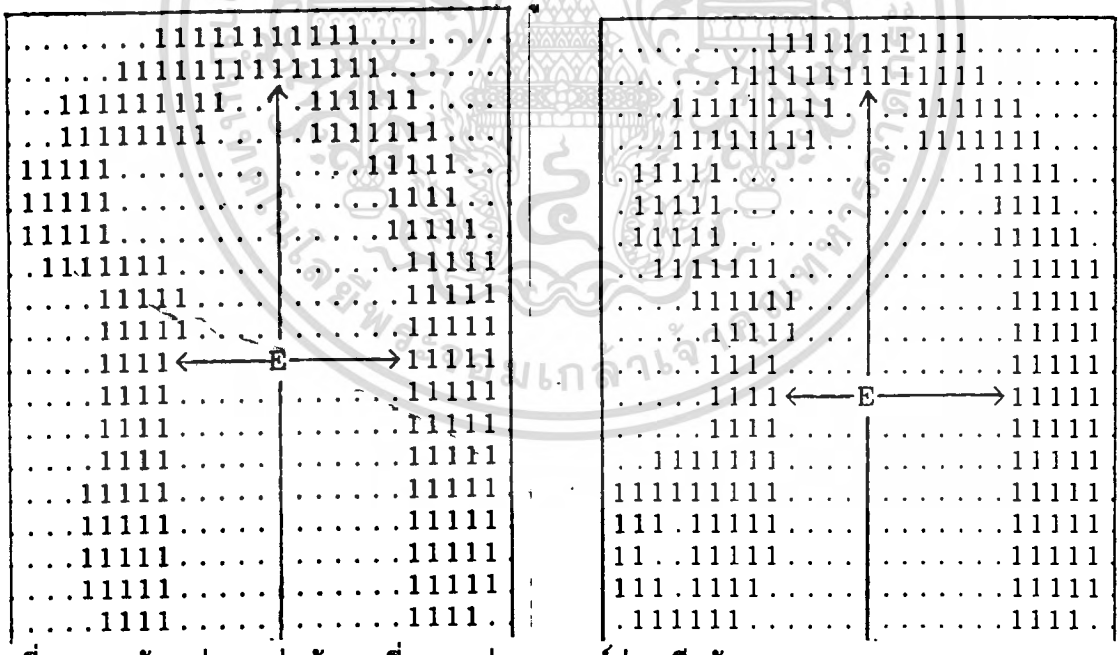
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
การจัดกลุ่มตัวอักษร

เนื่องจากลักษณะรูปร่างของอักษรภาษาไทย มีอักษรหลายตัวที่มีลักษณะรูปร่างคล้ายกัน เช่น ค ด ต หรือบางกลุ่มมีลักษณะเฉพาะกลุ่มแตกต่างออกไปเป็นคู่ๆ เช่นกลุ่ม ช ซ , บ ป , พ ฟ , ผ ฝ และขณะเดียวกันก็มีอักษรบางตัวมีลักษณะที่ไม่คล้ายกับตัวอักษรใดเลยเช่น เ เป็นต้น ในการวิเคราะห์เพื่อการรู้จำนั้นจำเป็นต้องให้การประมวลผลเป็นไปอย่างรวดเร็วและถูกต้องดังนั้นการพิจารณาจัดแยกกลุ่มอักษรออกเป็นกลุ่ม ๆ เพื่อสะดวกต่อการพิจารณาตัดสินใจของโปรแกรมและเพื่อความรวดเร็วในการวิเคราะห์ จึงจัดแยกกลุ่มอักษรโดยอาศัยทฤษฎี Initail Feature Extraction เพื่อให้รหัสเบื้องต้นสำหรับการจัดแยกกลุ่มตัวอักษร

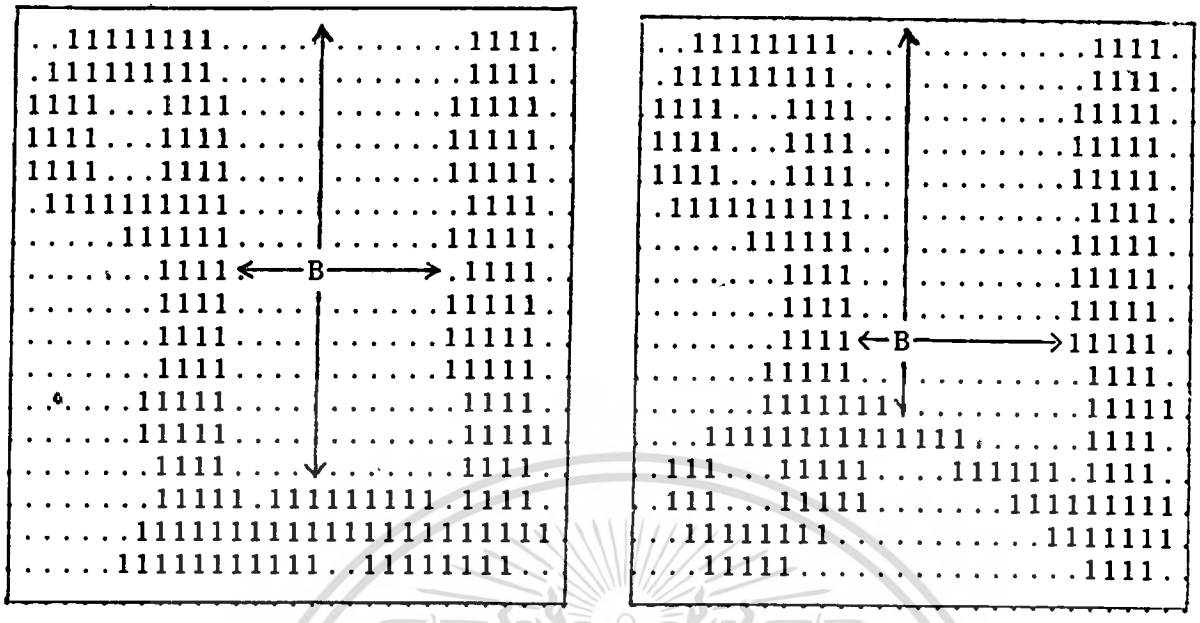
5.1 การให้รหัสกลุ่มตัวอักษร

ตัวอักษรแต่ละตัวเมื่อถูกแยกออกจากประโยค จะนำมาคำนวณหาจุดศูนย์ถ่วงของตัวอักษร ตำแหน่งจุดศูนย์ถ่วงที่คำนวณได้จะถือเป็นจุดปัจจุบันที่จะพิจารณาหาค่ารหัส Q-Code ตัวอย่างเช่น อักษรตัว ก ภ เมื่อถูกพิจารณาหาค่ารหัส Q-Code ที่ตำแหน่งจุดศูนย์ถ่วงแล้ว จะปรากฏรหัส Q-Code เป็น รหัส E เหมือนกัน ดังตัวอย่างรูปที่ 36. หรือกลุ่มอักษร บ ป ม ที่จุดศูนย์ถ่วงของตัวอักษรก็จะปรากฏรหัส Q-Code เป็นรหัส B เหมือนกันดังตัวอย่างรูปที่ 37.



รูปที่ 36. ตัวอย่างกลุ่มอักษรที่ตำแหน่งจุดศูนย์ถ่วงมีรหัส Q-Code เป็นรหัส E เหมือนกัน

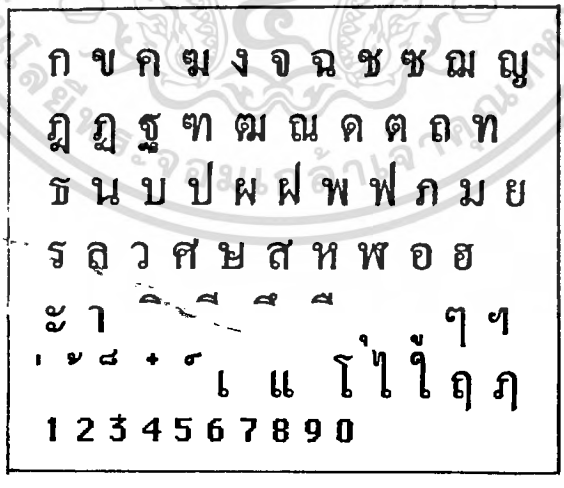
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 37. ตัวอย่างกลุ่มอักษรที่ตำแหน่งจุดศูนย์ถ่วงมีรหัส Q-Code เป็นรหัส B เหมือนกัน

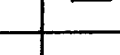
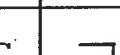
5.2 กลุ่มตัวอักษร

จากหัวข้อที่ 5.1 เมื่อนำตัวเลขและตัวอักษรทั้งสระ พยัญชนะทุกตัว ดังในภาพที่ 38. จำนวน 71 ตัว มาตรวจสอบหารหัส Q-Code ที่ตำแหน่งจุดศูนย์ถ่วงแล้ว ปรากฏกลุ่มอักษรต่าง ๆ จำนวน 7 กลุ่มดังในตารางที่ 4



รูปที่ 38. ภาพตัวเลขและตัวอักษรที่ใช้ตรวจสอบหารหัส Q-Code ที่จุดศูนย์ถ่วง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลำดับที่	ชื่อกลุ่ม	ลักษณะ Q-CODE	รหัสQ-CODE ที่ตำแหน่งCG	ตัวอักษร	จำนวน
1.	3		3	เ	1
2.	7		7	(	1
3.	B		B	น บ ป ม ย อ้ อ้ แ	8
4.	C		C	า ว ๆ ๆ โ ใ ไ	7
5.	D		D	ง ร 7 ฐ) 2 5 3	8
6.	E		E	ก ภ	2
7.	F		F	ข ค ม จ ฉ ช ซ ฌ ญ ฏ ฏ ฑ ฒ ณ ด ต ถ ท ธ พ ฟ ผ ฝ ล ศ ษ ส ห ฬ อ ฮ ่ ้ ๊ ๋ 1 4 6 8 9 0 ๆ ๆ	44

ตารางที่ 4. กลุ่มตัวอักษรที่ถูกจัดแยกกลุ่มด้วยรหัส Q-Code

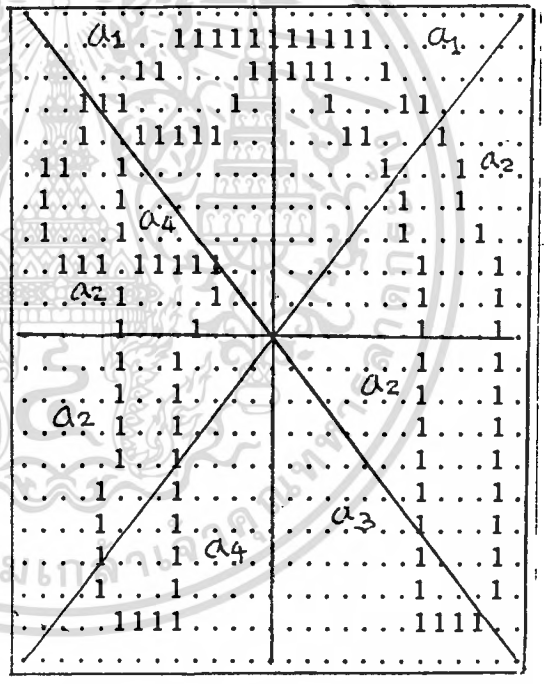
จากตารางที่ 4 ปรากฏว่ามีกลุ่มลำดับที่ 1 และกลุ่มลำดับที่ 2 เพียงตรวจสอบหาค่ารหัส Q-Code ก็สามารถวิเคราะห์ได้ทันทีว่าเป็นอักษรอะไร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6
การแยกตัวอักษรออกจากกลุ่ม

6.1 การวิเคราะห์แต่ละตัวอักษร

ข้อมูลตัวอักษรจากรูปที่ 38. เมื่อตัวอักษรถูกจัดแยกกลุ่มกันแล้ว ภายในกลุ่มจะมีรูปร่างลักษณะที่คล้ายกันเป็นส่วนใหญ่ หากจะพิจารณาวิธีที่จะวิเคราะห์เพื่อแยกตัวอักษรทุกตัวออกจากกลุ่ม วิธีการนั้นต้องครอบคลุมลักษณะรูปร่างอย่างละเอียดจากทฤษฎี Topological Feature Extraction ภายในแต่ละ Octant จะมีส่วนต่างๆ ของตัวอักษรที่ลากผ่านเส้นทะแยงมุมทั้ง 8 และสามารถตรวจค้นหาค่ารหัส A-Code จากตัวอักษรในกลุ่มได้อย่างมีประสิทธิภาพ ดังนั้นตัวอักษรทุกตัวจึงถูกนำมาพิจารณาค้นหาค่ารหัสชุด A-Code ดังตัวอย่างกลุ่มรหัส E ซึ่งประกอบด้วยอักษร ก ภ แสดงไว้ในรูปที่ 39 ก. และ 39 ข.



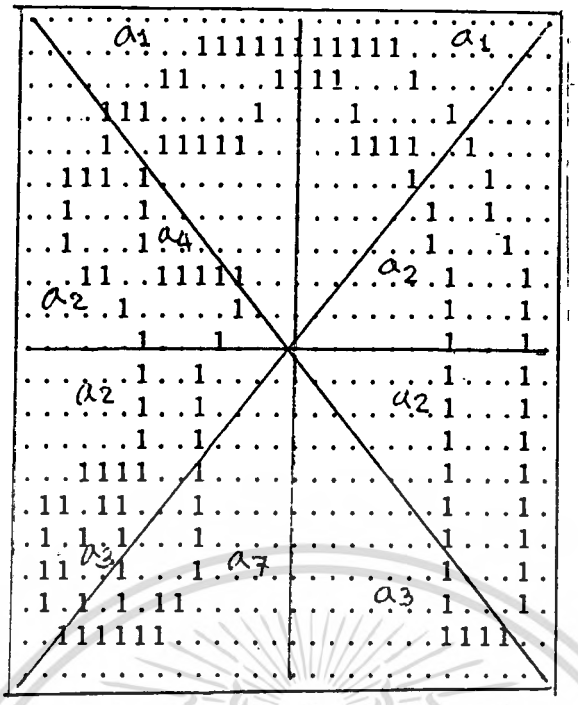
รูปที่ 39 ก.
แสดงตัวอย่างการ
วิเคราะห์แยกอักษร
ตัว ก ออกจากกลุ่ม E

ชุดรหัส A-code ของภาพอักษร ก ประกอบด้วย

A-code = $a_1,0,0;a_1,0,0;a_2,0,0;a_2,0,0;a_3,0,0;$
$a_4,0,0;a_2,0,0;a_2,a_4,0;$

และเมื่อนำจัดเรียงรหัสใหม่เพื่อให้เหมาะสมต่อการจัดเก็บเป็นฐานข้อมูลจะได้ค่าเป็นชุด A-code ของอักษร ก คือ 100 100 200 200 300 400 200 240

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 39 ข. แสดงตัวอย่างการวิเคราะห์แยกอักษร ภ ออกจากกลุ่ม E

ชุดรหัส A-code ของภาพอักษร ภ ประกอบด้วย

A-code = $a_1,0,0;a_1,0,0;a_2,0,0;a_2,0,0;a_3,0,0;$
$a_7,0,0;a_2,a_3,0;a_2,a_4,0;$

และเมื่อนำจัดเรียงรหัสใหม่เพื่อให้เหมาะสมต่อการจัดเก็บเป็นฐานข้อมูลจะได้ค่าเป็นชุด A-code ของอักษร ภ คือ 100 100 200 200 300 700 230 240

ดังนั้นการแยกตัวอักษรออกจากกลุ่มสำหรับกลุ่ม E สามาแยกได้โดยเปรียบเทียบชุดรหัส A-Code ของตัวอักษรภายในกลุ่มดังตารางที่ 5

ชื่อกลุ่ม	ตัวอักษร	ชุดรหัส A-code							
		Octant							
		1	2	3	4	5	6	7	8
E	ก	100	100	200	200	300	400	200	240
	ภ	100	100	200	200	300	700	230	240

ตารางที่ 5. เปรียบชุดรหัส A-code ของกลุ่ม E

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และกลุ่มอักษรอื่น ๆ ได้เปรียบเทียบชุดรหัส A-code ดังแสดงในตารางที่ 6,7,8,9

ชื่อกลุ่ม	อักษร	ชุดรหัส A-code							
		Octant							
		1	2	3	4	5	6	7	8
B	น	800	400	200	240	170	100	200	240
	บ	800	400	200	200	100	100	200	240
	ป	000	400	200	200	100	100	240	800
	ม	800	400	200	200	100	140	230	240
	ย	830	400	200	200	100	100	200	240
	◌ั	800	400	200	200	100	100	400	200
	◌้	300	400	200	200	100	100	200	290
	แ	300	400	200	260	300	190	200	200

ตารางที่ 6. เปรียบชุดรหัส A-code ของกลุ่ม B

ชื่อกลุ่ม	อักษร	ชุดรหัส A-code							
		Octant							
		1	2	3	4	5	6	7	8
C	า	100	100	200	200	300	000	000	400
	ว	100	100	200	200	140	500	000	400
	ๆ	113	100	200	200	300	000	000	240
	ๆ	130	100	200	200	300	000	000	500
	เ	340	200	300	400	300	000	400	400
	ใ	110	100	200	260	370	000	000	500
	ไ	130	100	200	260	370	000	000	500

ตารางที่ 7. เปรียบชุดรหัส A-code ของกลุ่ม C

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อกลุ่ม	อักษร	ชุดรหัส A-code Octant							
		1	2	3	4	5	6	7	8
D	ง	500	130	200	200	100	100	100	300
	ร	100	100	230	200	140	500	000	500
	7	100	100	200	200	000	300	000	400
	ฐ	113	130	200	240	230	140	300	440
	)	100	100	200	200	200	200	300	400
	2	100	100	200	240	220	500	000	400
	5	100	130	200	200	100	100	300	500
	3	400	300	200	200	100	300	000	400

ตารางที่ 8. เปรียบชุดรหัส A-code ของกลุ่ม D

ชื่อกลุ่ม	อักษร	ชุดรหัส A-code Octant							
		1	2	3	4	5	6	7	8
F	ข	800	400	200	200	300	000	000	490
	ค	110	110	200	200	300	400	300	240
	ม	300	400	200	200	300	490	700	230
	จ	130	100	200	200	100	100	200	240
	ฉ	130	100	200	240	230	200	240	240
	ช	800	400	200	200	300	000	000	200
	ซ	300	400	200	200	300	000	700	230
	ฌ	130	140	220	220	240	570	230	240
	ญ	130	140	220	340	200	400	200	240

ตารางที่ 9. เปรียบชุดรหัส A-code ของกลุ่ม F

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อกลุ่ม	อักษร	ชุดรหัส A-code								
		Octant								
		1	2	3	4	5	6	7	8	
F	ฎ	130	140	200	200	240	100	200	200	
	ฏ	130	140	200	200	240	100	200	200	
	ฑ	130	800	220	220	440	000	700	230	
	ฒ	130	140	220	220	240	450	244	220	
	ณ	800	400	200	240	330	490	200	240	
	ด	110	110	220	220	240	240	220	220	*1
	ต	110	110	220	220	240	240	220	220	*1
	ถ	110	130	200	200	460	130	200	240	
	ท	800	400	200	200	400	000	200	240	
	ธ	110	130	200	200	100	100	200	000	
	พ	800	140	220	230	300	100	200	240	
	ฟ	000	440	220	340	240	200	240	800	
	ผ	800	400	200	300	300	400	400	200	
	ฝ	000	400	200	200	100	100	230	800	
	ล	110	110	230	200	400	100	200	240	
	ศ	110	140	230	200	300	400	340	220	
	ษ	140	140	220	240	120	130	200	240	
	ส	110	139	230	200	300	490	200	240	
	ห	180	140	260	200	300	400	200	224	
	ฬ	180	184	220	220	300	400	220	224	
	อ	130	130	200	200	100	100	200	240	
	ฮ	110	130	200	200	100	100	340	250	
	'	000	000	000	000	000	000	000	000	*2
	๗	100	140	220	220	300	700	240	200	
	.	000	000	000	000	000	000	000	000	*2

ตารางที่ 9.(ต่อ) เปรียบชุดรหัส A-code ของกลุ่ม F

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อกลุ่ม	อักษร	ชุดรหัส A-code								
		Octant								
		1	2	3	4	5	6	7	8	
F	๋	100	140	600	300	100	100	200	200	
	์	000	400	300	000	000	400	200	300	
	็	100	100	000	000	300	000	000	400	
	่	100	400	200	200	400	000	000	300	
	้	300	000	000	000	000	000	000	400	
	4	100	100	200	200	100	100	200	200	$\star^3$
	6	110	130	200	200	100	100	200	240	
	8	400	300	000	400	100	100	200	000	
	9	100	100	200	230	110	130	200	200	
	0	100	100	200	200	100	100	200	200	$\star^3$

ตารางที่ 9.(ต่อ) เปรียบชุดรหัส A-code ของกลุ่ม F

หมายเหตุ $\star^1$ หมายถึง ในกลุ่ม F ตัวอักษร ด และ ต มีชุดรหัส A-code เหมือนกัน
$\star^2$ หมายถึง ในกลุ่ม F ตัวอักษร ่ และ ้ มีชุดรหัส A-code เหมือนกัน
$\star^3$ หมายถึง ในกลุ่ม F เลข 4 และ 0 มีชุดรหัส A-code เหมือนกัน

6.2 การย่อรหัส

ในการบันทึกค่ารหัส ซึ่งเป็นคุณสมบัติเด่นของตัวอักษรแต่ละตัว เก็บไว้เป็นฐานข้อมูลต้นแบบเพื่อการรู้จำ หากทำการบันทึกรหัสชุด A-Code เป็นตัวเลขฐานสิบเรียงกันดังได้กล่าวมาแล้ว อักษรต้นแบบ หนึ่งตัวจะต้องใช้พื้นที่ในการเก็บรหัสอย่างน้อย 24 Byte ซึ่งทำให้เปลืองพื้นที่หน่วยความจำมาก เพื่อแก้ปัญหาดังกล่าวจึงได้ลดพื้นที่การเก็บชุดรหัส A-Code ของอักษรต้นแบบโดย แปลงค่ารหัสจากเลขฐานสิบให้เป็นเลขฐานสอง โดยเก็บหลักละ 4 bit แล้วแปลงเลขฐานสองเป็นเลขฐานสิบหก ดังตัวอย่าง รหัสอักษร ก. 100 100 200 200 300 400 200 240 รวมเป็น 24 Byte แปลงให้อยู่ในรูปเลขฐานสอง หลักละ 4 bit จำนวน Byte จะลดลงเหลือเพียง 12 Byte ซึ่งสามารถมองเห็นเป็นลักษณะเลขฐานสิบหกในหน่วยความจำของเครื่องได้ดังนี้ 10 01 00 20 02 00 30 04 00 20 02 40

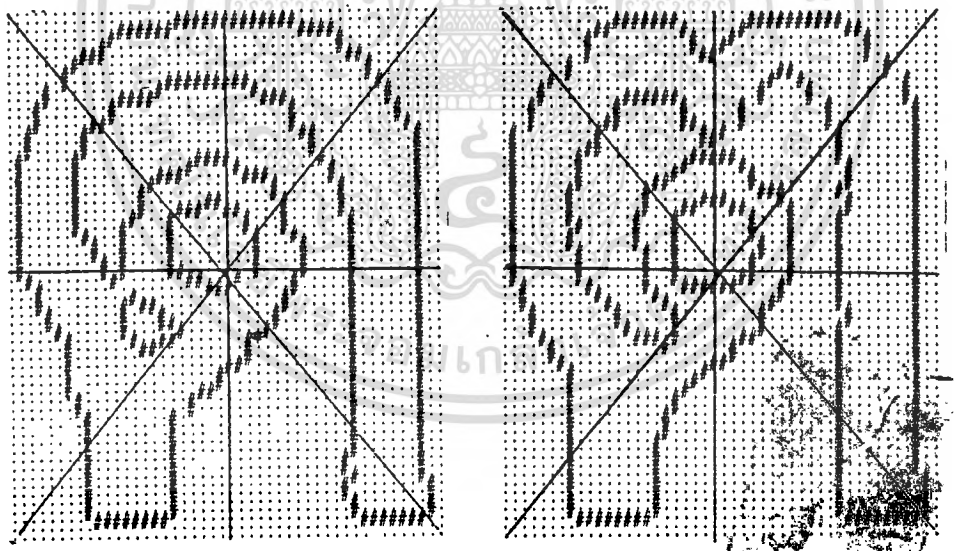
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.3 การวิเคราะห์เฉพาะส่วนของตัวอักษร

จากหัวข้อ 6.1 อักษรในกลุ่ม F ปรากฏมีอักษรที่มีชุดรหัส A-Code เหมือนกันทุก Octant อยู่ 3 คู่คือ ด ต , [illegible] และ 4 0 ดังนั้นเพื่อให้การวิเคราะห์ได้ผลถูกต้องจึงได้พิจารณาเฉพาะส่วนของกลุ่มอักษรที่เกิดปัญหาเหล่านั้นดังนี้

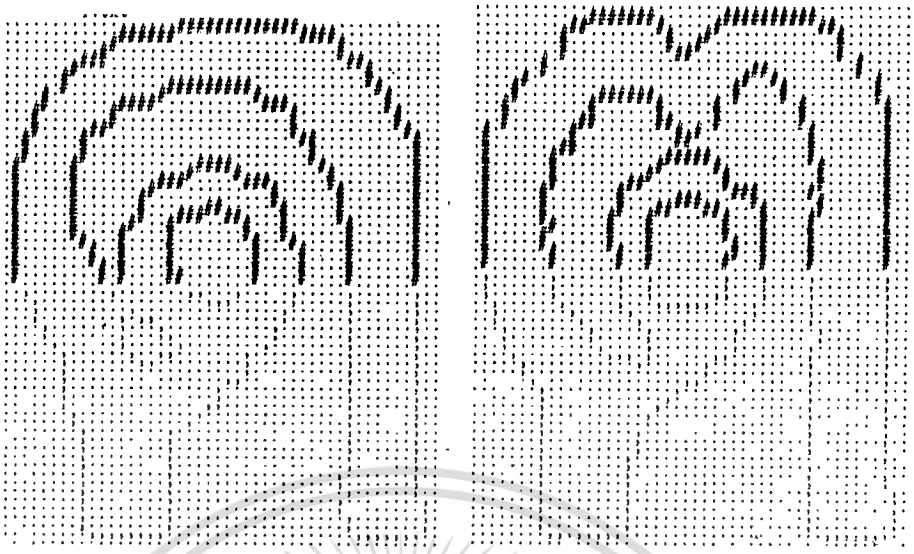
6.3.1 กรณีชุด A-Code เหมือนกันของอักษรตัว ด และ ต

ข้อแตกต่างของอักษร ด กับ ต มีจุดที่แตกต่างเห็นได้ชัดคือบริเวณส่วนบนของอักษร ดังรูปที่ 40. ดังนั้นจึงได้พิจารณาเฉพาะบริเวณส่วนบนของอักษรคู่นี้ โดยเริ่มจากจุดกึ่งกลางของอักษรขึ้นไปถึงส่วนบนสุด ดังรูปที่ 41. เพื่อใช้เป็นพื้นที่สำหรับวิเคราะห์เฉพาะส่วน โดยแบ่งส่วนพื้นที่ตามเส้นทะแยงมุมออกเป็น 8 Octant ดังรูปที่ 42. เมื่อพิจารณาตามทฤษฎีโทโบโลจิคอลเพียเจอร์เอ็กแทร็กชั่น แล้วปรากฏว่าชุดรหัส A-Code ของอักษร ด และ ต เฉพาะบริเวณครึ่งบนให้รหัสแตกต่างกันดังแสดงในตารางที่ 10.

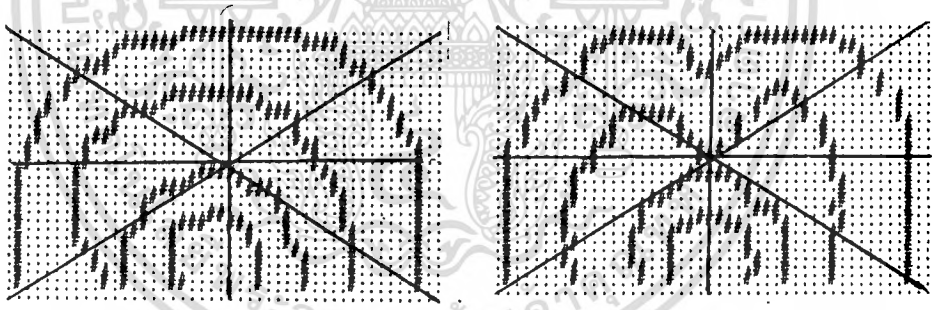


รูปที่ 40. เปรียบเทียบส่วนที่แตกต่างกันของอักษร ด และ ต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 41. บริเวณส่วนครึ่งบนของอักษร ด และ ต ที่เลือกพิจารณาหาค่าชุดรหัส A-Code ใหม่



รูปที่ 42. การหาค่าชุดรหัส A-Code บริเวณครึ่งบนของอักษร ด และ ต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของผลงานทุกครั้งที่มีการนำไปใช้

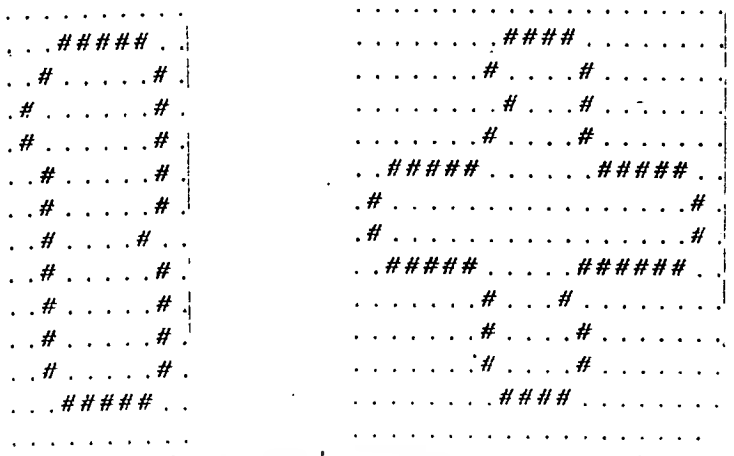
ส่วนครึ่งบนของ	ชุดรหัส A-code Octant							
อักษร	1	2	3	4	5	6	7	8
ด	100	100	200	240	300	400	230	200
ต	300	400	200	200	000	400	200	200

ตารางที่ 10. ชุดรหัส A-code เฉพาะส่วนครึ่งบนของอักษร ด และ ต

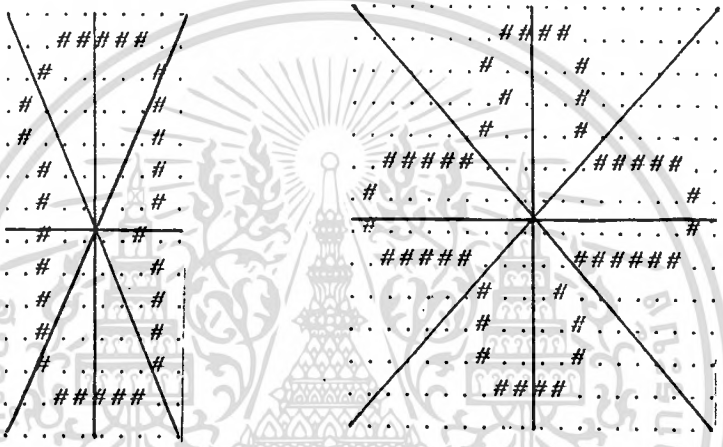
6.3.2 กรณีชุด A-Code เหมือนกันของอักษรตัว ' และ +

ข้อแตกต่างของอักษร ' กับ + มีจุดที่แตกต่างเห็นได้ชัดตลอดทั้ง 4 ด้านของอักษร ดังรูปที่ 43. แต่เนื่องจากเงื่อนไขที่ได้กำหนดไว้ในทฤษฎีโทโบโลจิคอลเฟียเจอร์เอ็กแทร็กชั่น ตั้งแต่เบื้องต้นว่าให้พิจารณา Storke ที่ลากผ่าน Octant หนึ่งไปยัง Octant ที่อยู่ติดกัน โดยไม่มีเส้น Storke คู่ขนาน จะให้รหัส A-Code เป็น a_0 ดังนั้นค่าชุดรหัส A-Code ของอักษร ' และ + จึงให้ค่าออกมาเหมือนกัน เพราะมีลักษณะเป็นเส้น Storke เดี่ยวลากผ่าน Octant คล้ายกัน ดังแสดงในรูปที่ 44. ในการพิจารณาวิเคราะห์เฉพาะส่วนของอักษรคู่นี้ ได้ใช้หลักการเดียวกันกับการแก้ปัญหาของอักษร ด และ ต กล่าวคือ ได้เลือกพิจารณาเฉพาะบริเวณส่วนครึ่งบนของอักษรคู่นี้ โดยเริ่มจากจุดกึ่งกลางของอักษรขึ้นไปถึงส่วนบนสุด ดังรูปที่ 45. เพื่อใช้เป็นพื้นที่สำหรับวิเคราะห์เฉพาะส่วน โดยแบ่งส่วนพื้นที่ตามเส้นทะแยงมุมออกเป็น 8 Octant ดังรูปที่ 46. เมื่อพิจารณาตามทฤษฎีโทโบโลจิคอลเฟียเจอร์เอ็กแทร็กชั่น แล้วปรากฏว่าชุดรหัส A-Code ของอักษร ' และ + เฉพาะบริเวณครึ่งบนให้รหัสแตกต่างกันดังแสดงในตารางที่ 11.

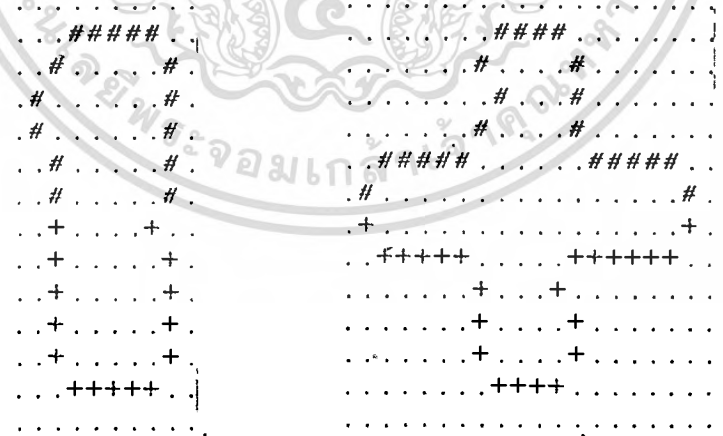
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 43. เปรียบเทียบลักษณะที่แตกต่างกันของอักษร ฺ และ ฺ

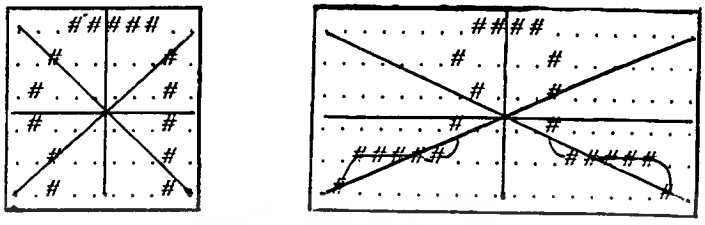


รูปที่ 44. ลักษณะที่ถูกพิจารณาเป็น Storke เดี่ยวที่ไม่มี Storke ค่ขนาน ตามเงื่อนไขในทฤษฎีทฤษฎีโทโบโลจิคอลเฟียเจอร์เอ็กแทร็กชั่น



รูปที่ 45. บริเวณส่วนครึ่งบนของอักษร ด และ ต ที่เลือกพิจารณาหาค่าชุดรหัส A-Code ใหม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 46. การหาค่าชุดรหัส A-Code บริเวณครึ่งบนของอักษร ' และ +

ส่วนครึ่งบนของ	ชุดรหัส A-code Octant							
อักษร	1	2	3	4	5	6	7	8
'	000	000	000	000	000	000	000	000
+	000	000	000	400	300	400	300	000

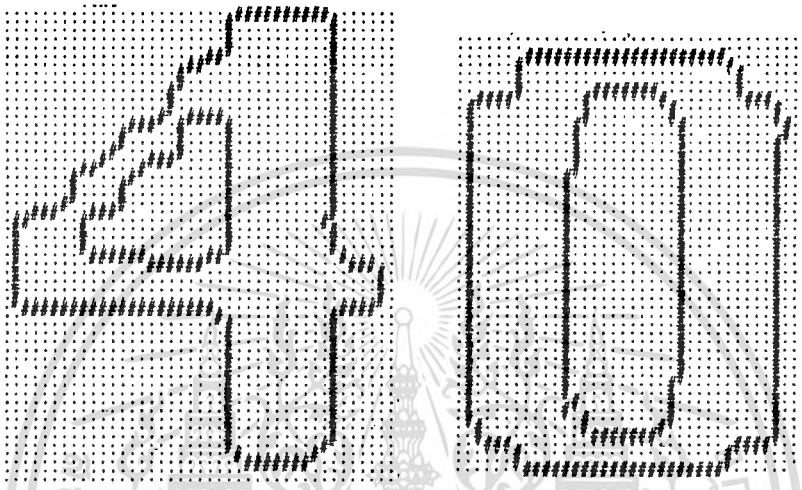
ตารางที่ 11. ชุดรหัส A-code เฉพาะส่วนครึ่งบนของอักษร + และ +

6.2.3 <u>กรณีชุด A-Code เหมือนกันของตัวเลข 4 และ 0</u>

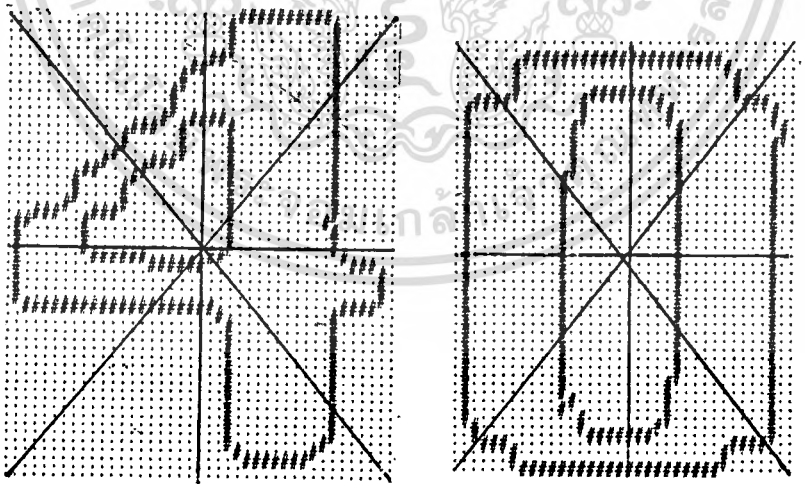
ข้อแตกต่างของตัวเลข 4 กับ 0 ดูด้วยตาเปล่าจะเห็นลักษณะที่แตกต่างโดยทั่วไปอย่างเห็นได้ชัด ดังรูปที่ 47. แต่เนื่องจากเงื่อนไขที่ได้กำหนดไว้ในทฤษฎีโทโบโลจิคอลเฟียเจอร์เอ็กแทร็กชั่น ดังนั้นค่าชุดรหัส A-Codeของตัวเลข 4 และ 0 จึงให้ค่าออกมาเหมือนกันอีก ดังแสดงในรูปที่ 48. ในการพิจารณาวิเคราะห์เฉพาะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนของตัวเลขคู่นี้ได้ใช้หลักการเดียวกันกับการแก้ปัญหาของอักษรคู่ ด ต และคู่ ' + คือได้เลือกพิจารณาเฉพาะบริเวณส่วนครึ่งบนของตัวเลขคู่นี้ โดยเริ่มจากจุดกึ่งกลางของอักษรขึ้นไปถึงส่วนบนสุด ดังรูปที่ 49. เพื่อใช้เป็นพื้นที่สำหรับวิเคราะห์เฉพาะส่วน โดยแบ่งส่วนพื้นที่ตามเส้นทะแยงมุมออกเป็น 8 Octant ดังรูปที่ 50. เมื่อพิจารณาตามทฤษฎีโทโบโลจิคอลเฟียเจอร์เอ็กแทร็กชั่นแล้วปรากฏว่าชุดรหัสA-Code ของตัวเลข 4 และ 0 เฉพาะบริเวณครึ่งบนให้รหัสแตกต่างกันดังแสดงในตารางที่ 12.

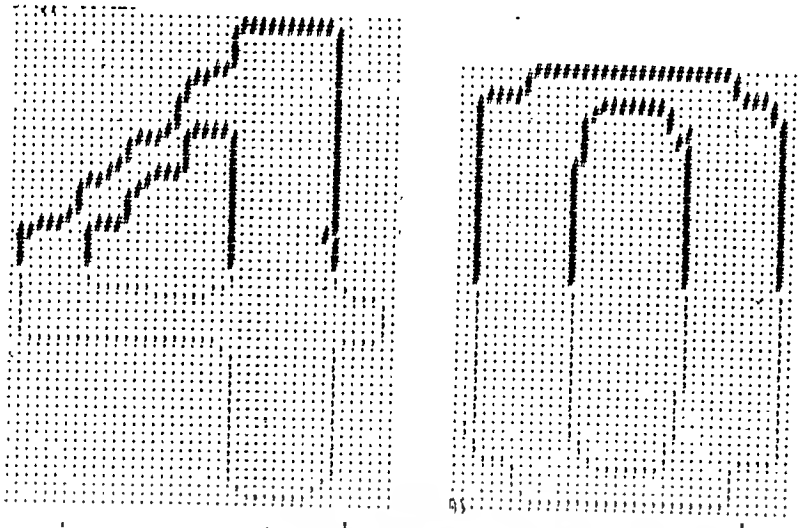


รูปที่ 47. เปรียบเทียบลักษณะที่แตกต่างกันของตัวเลข 4 และ 0

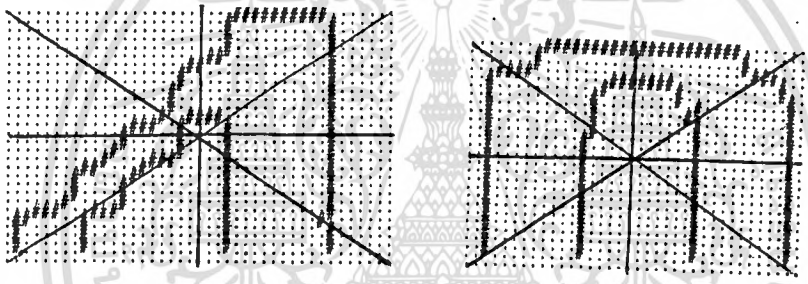


รูปที่ 48. ลักษณะที่ถูกพิจารณาหาค่ารหัส A-Code จาก Octant ทั้ง 8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 49. บริเวณส่วนครึ่งบนของตัวเลข 4 และ 0 ที่เลือกพิจารณาหาค่าชุดรหัส A-Code ใหม่

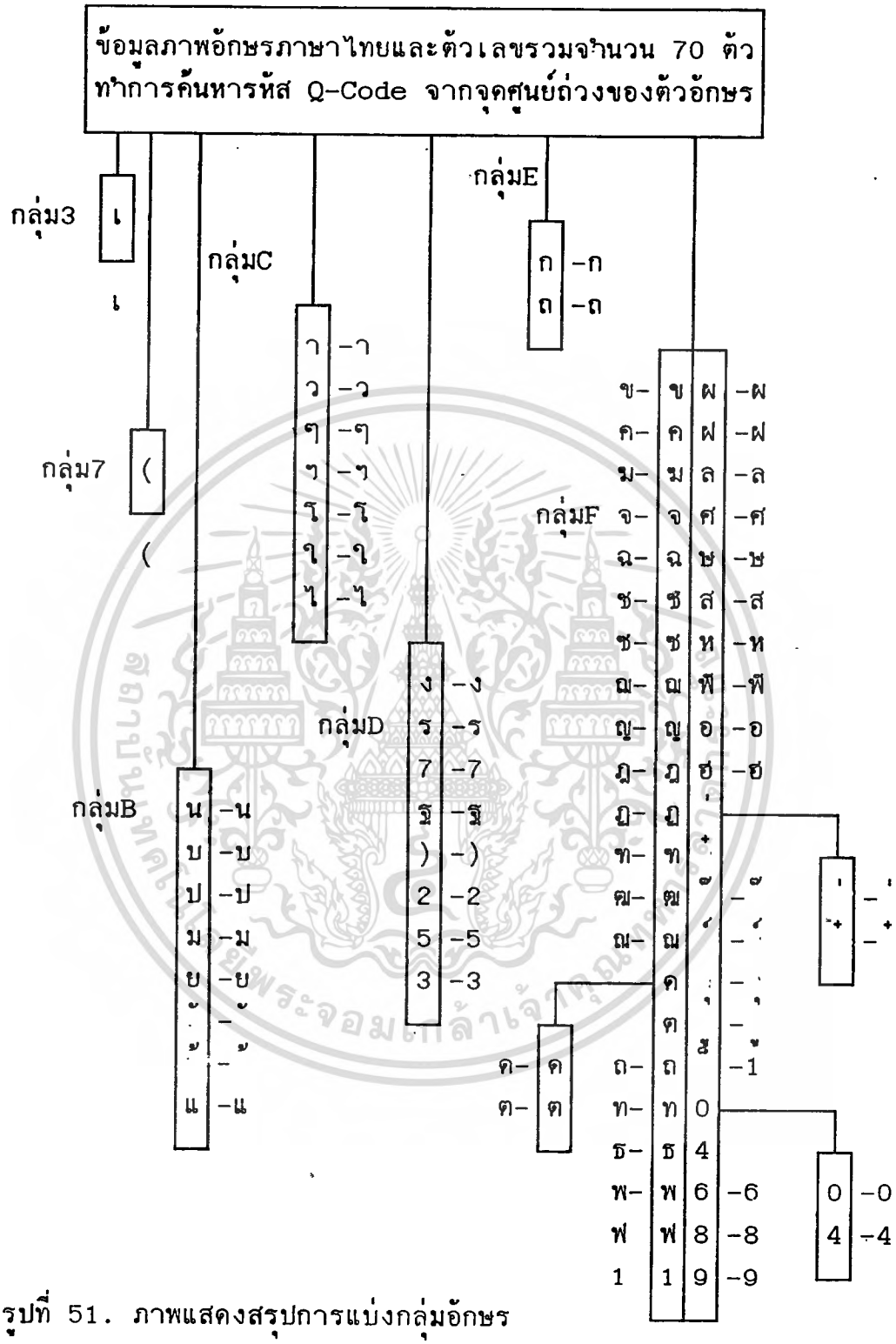


รูปที่ 50. การหาค่าชุดรหัส A-Code บริเวณครึ่งบนของตัวเลข 4 และ 0

ส่วนครึ่งบนของ	ชุดรหัส A-code Octant							
อักษร	1	2	3	4	5	6	7	8
4	100	100	200	200	000	700	230	200
0	100	100	200	200	000	000	200	200

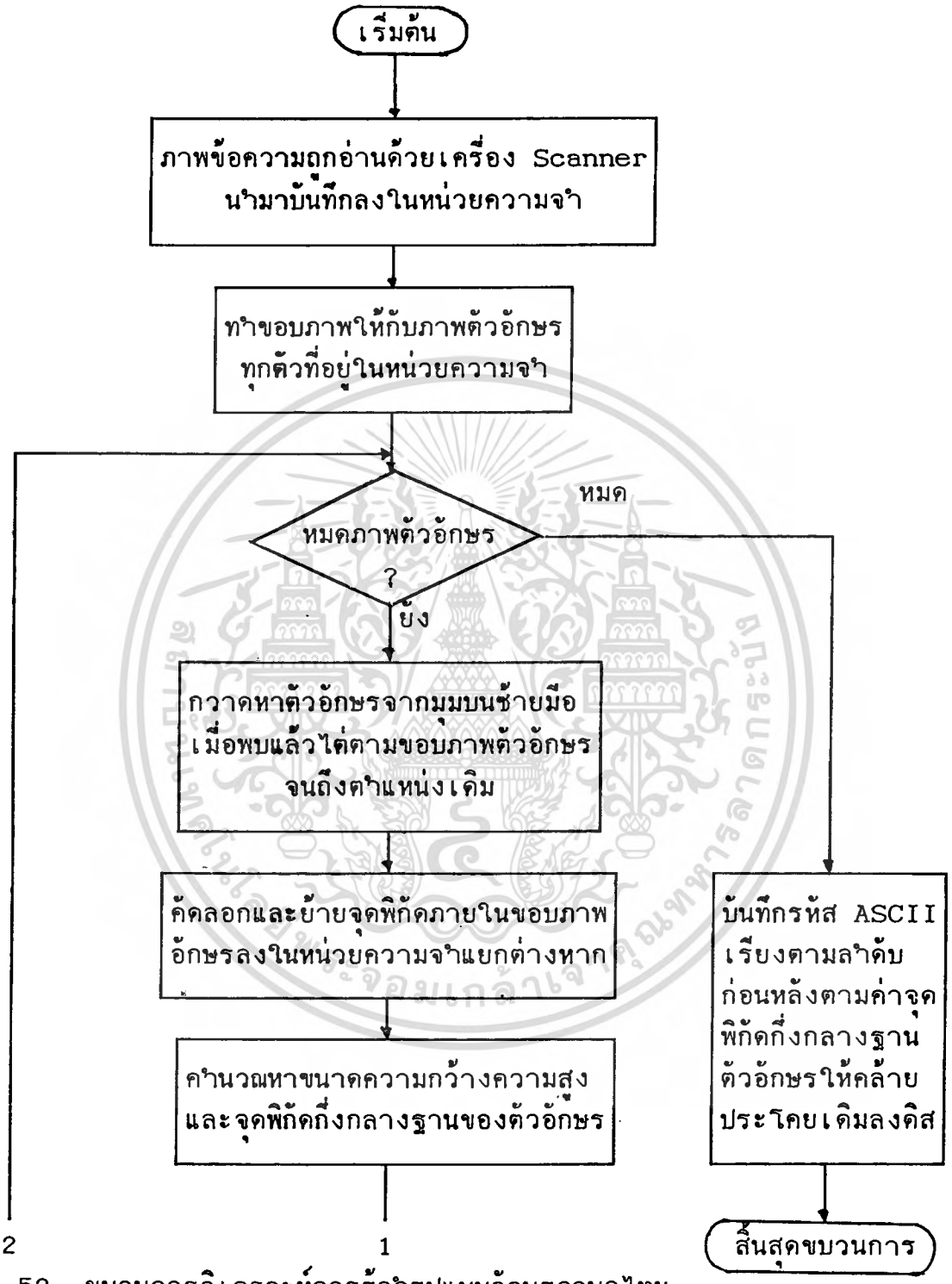
ตารางที่ 12. ชุดรหัส A-code เฉพาะส่วนครึ่งบนของตัวเลข 4 และ 0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



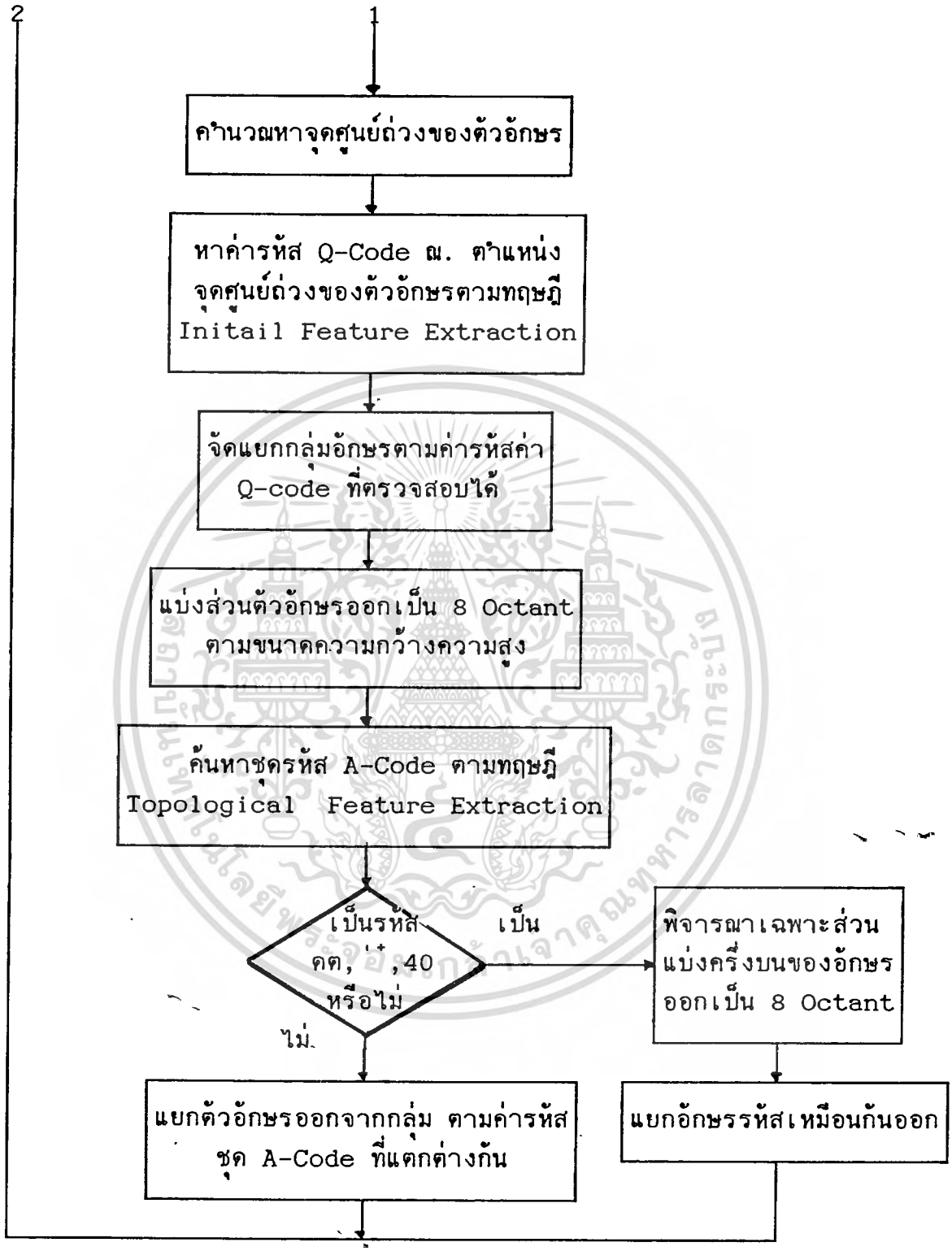
รูปที่ 51. ภาพแสดงสรุปการแบ่งกลุ่มอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 52. ขบวนการวิเคราะห์การรู้จำรูปแบบอักษรภาษาไทย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 7
โปรแกรมการรู้จำรูปแบบตัวพิมพ์อักษรภาษาไทย

โปรแกรมที่ใช้และพัฒนาขึ้นเพื่อใช้ทดสอบการทำงานของขบวนการ ตามวิธีการของวิทยานิพนธ์ฉบับนี้ ได้แบ่งโปรแกรมที่ใช้งานเป็น 2 ส่วนคือ

1. โปรแกรมสำเร็จรูปควบคุมการอ่านภาพจากเครื่อง Handy Scanner
2. โปรแกรมที่พัฒนาขึ้นสำหรับวิเคราะห์การรู้จำอักษรภาษาไทย

7.1 โปรแกรมสำเร็จรูปควบคุมการอ่านภาพจากเครื่อง Handy Scanner

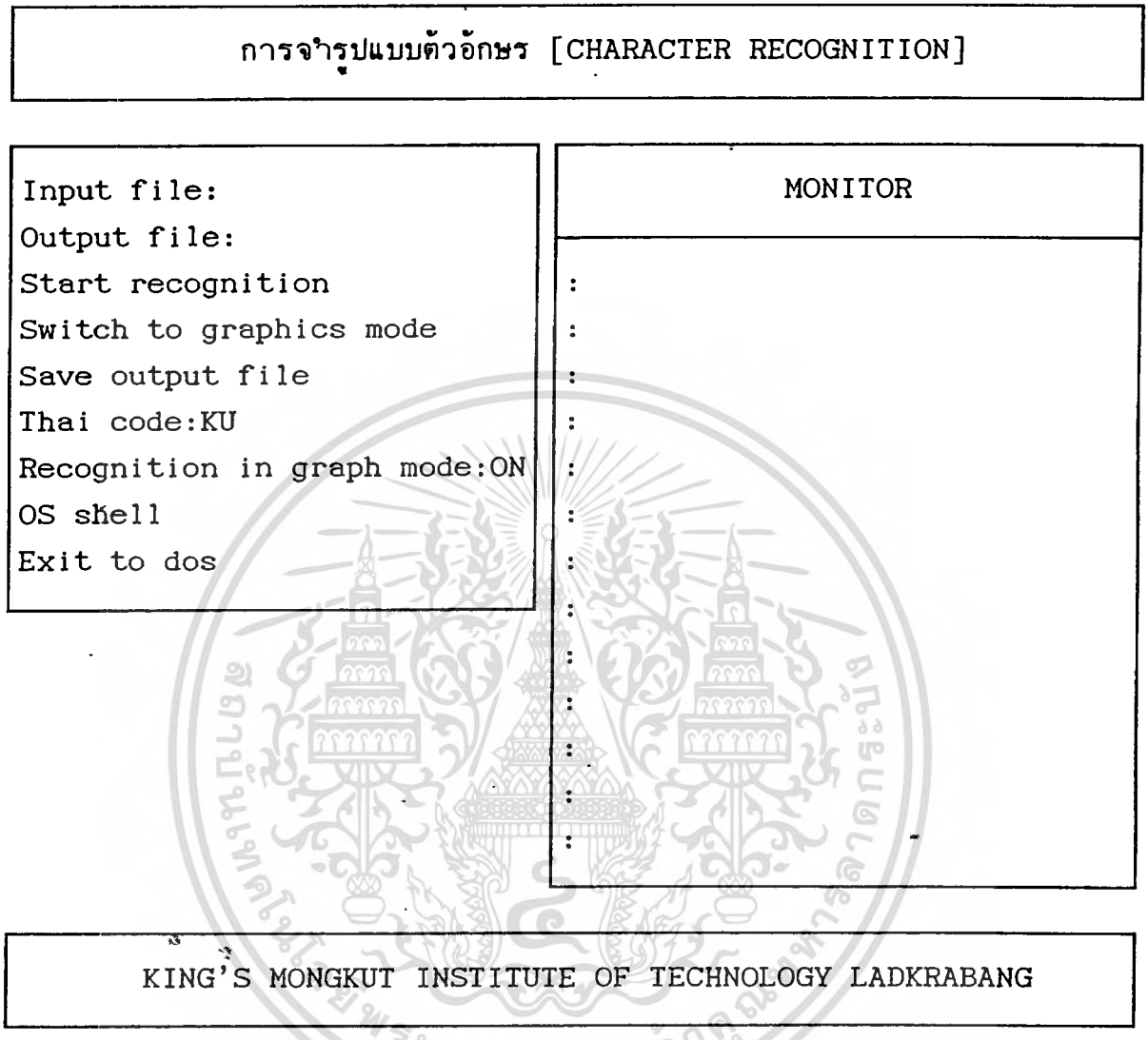
โปรแกรมชุดนี้ เป็นโปรแกรมสำเร็จรูปที่บริษัทผู้ผลิตเครื่อง Handy Scanner พัฒนาขึ้น สำหรับควบคุมการทำงานของเครื่อง Handy Scanner โดยทำงานควบคู่กับ Interface Card ทำการอ่านข้อมูลภาพจากเครื่อง Handy Scanner ตั้งแต่ผู้ใช้เริ่มวางเครื่อง Handy Scanner แล้วลากผ่านหน้ากระดาษที่มีข้อความ ซึ่งกลุ่มข้อความเหล่านั้นจะไปปรากฏที่หน้าจอภาพคอมพิวเตอร์ เรียงตามแนวตั้งของจอภาพ โดยจะสัมพันธ์กับการลากเครื่อง Handy Scanner ที่กำลังผ่านข้อความบนกระดาษ จนเต็มหน่วยความจำส่วน Buffer Scan ของโปรแกรม การอ่านภาพจึงจะสิ้นสุด กลุ่มภาพข้อความที่ปรากฏในหน้าจอพร้อมจะนำไปบันทึกลงในแผ่นดิสค์ ซึ่งรายละเอียดการใช้งานของโปรแกรมชุดนี้ ได้กล่าวไว้ในภาคผนวก ข. File ที่บันทึกลงในแผ่นดิสค์สามารถเลือก Format ของ File ได้หลายชนิด ในงานวิจัยครั้งนี้ได้เลือก Format ชนิด OCR ซึ่งเป็น Format File สำหรับการนำไปใช้ประโยชน์ในการวิเคราะห์แบบตัวอักษรโดยเฉพาะ

7.2 โปรแกรมที่พัฒนาขึ้นสำหรับวิเคราะห์การรู้จำอักษรภาษาไทย

โปรแกรมที่พัฒนาขึ้นเขียนด้วยภาษาแอ็ดแซมบลี้ ที่ใช้สำหรับเครื่องไมโครคอมพิวเตอร์ตระกูล IBM ทั่วไป โดยโปรแกรมนี้จะรับข้อมูลภาพจาก File ที่มี Format เป็นแบบ OCR เท่านั้น ดังนั้นก่อนจะเริ่มใช้งานโปรแกรมชุดนี้ จะต้องอ่านภาพข้อความจากสิ่งพิมพ์ด้วยเครื่อง Handy Scanner โดยอาศัยโปรแกรมจากหัวข้อ7.1 ทำการบันทึกภาพข้อความลงในแผ่นดิสค์ก่อน จึงจะสามารถนำ File ข้อมูลภาพมาวิเคราะห์ในโปรแกรมชุดนี้ได้

ในการใช้งาน โปรแกรมได้จัดทำเป็นเมนูสำหรับเลือก ฟังค์ชั่นการใช้งานที่เป็นประโยชน์และสะดวกต่อผู้ใช้ดังแสดงตัวอย่างในรูปที่ 53.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 53. แสดงภาพการใช้งานของโปรแกรมที่พัฒนาขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7.2.1 <u>คุณสมบัติเฉพาะของโปรแกรม</u>

โปรแกรมได้ถูกพัฒนาให้เหมาะต่อการใช้งานจริง ซึ่งประกอบด้วยคุณสมบัติดังนี้

1. อ่านชื่อ File ข้อมูลภาพชนิด OCR ได้
2. แสดงภาพข้อมูลให้ปรากฏที่จอภาพ พร้อมบังคับเลื่อนภาพขึ้นลงหรือซ้ายขวาได้
3. ในการวิเคราะห์ตัวอักษร สามารถเลือกให้แสดงภาพข้อมูลที่กำลังวิเคราะห์พร้อมไปด้วยก็ได้
4. โปรแกรมสามารถแยกภาพตัวอักษรออกจากภาพกลุ่มข้อความ ที่ประกอบด้วยภาพอักษรหลายบรรทัดได้
5. ผลลัพธ์ที่ได้จากการวิเคราะห์อักษรแต่ละตัวจะปรากฏให้เห็นเรียงเป็นประโยค
6. อักษรใดที่ไม่สามารถวิเคราะห์ได้ จะแสดงภาพรูปร่างอักษรตัวนั้นให้ผู้ใช้ตัดสินใจเลือกรหัส ASCII ได้เอง
7. สามารถเลือกผลลัพธ์เป็นรหัส ASCII Code ภาษาไทยของ รหัสเกษตร หรือ รหัส สมอ. ได้
8. ผลลัพธ์ที่วิเคราะห์ได้จะปรากฏที่หน้าจอให้สามารถแก้ไขเปลี่ยนแปลงข้อความได้
9. สามารตั้งชื่อ File ของผลลัพธ์ที่วิเคราะห์เสร็จแล้วเตรียมบันทึกลงแผ่นดิสค์ได้
10. สามารถยกเลิกการใช้งานของโปรแกรมชั่วคราว เพื่อใช้งานโปรแกรมอื่น ๆ ก่อน แล้วย้อนกลับมาใช้โปรแกรมวิเคราะห์การจดจดอักษรอีกครั้ง โดยไม่ต้อง RUN โปรแกรมใหม่

7.2.2 <u>โปรแกรมและ File ที่เกี่ยวข้อง</u>

โปรแกรมเมื่อพัฒนาเสร็จ ได้ทันทึกเก็บไว้ในแผ่นดิสค์ชื่อ "THAI OCR" พร้อมกับ File อื่น ๆ ที่เกี่ยวข้อง ในการเตรียมโปรแกรมสำหรับการใช้งานประกอบด้วย File 3 กลุ่มคือ

1.โปรแกรมการรู้จำรูปแบบอักษร

โปรแกรมชุดนี้ประกอบด้วย File สำคัญ 2 File คือ

OCR.EXE : เป็นโปรแกรมหลักสำหรับการวิเคราะห์ตัวอักษร

THAICODE.DAT : เป็นฐานข้อมูลรหัสต้นแบบของตัวอักษร

2.กลุ่ม File ข้อมูลภาพ

File กลุ่มนี้เป็นข้อมูลภาพที่ถูกอ่านจากเครื่อง Handy Scanner โดยมี Format ของ File เป็นชนิด OCR เช่น

TEXT.OCR หรือ

DATA.OCR เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.โปรแกรม Driver ภาษาไทย

เพื่อให้การปรากฏผลลัพธ์ของรหัส ASCII ภาษาไทย ที่หน้าจอภาพปรากฏเป็นภาษาไทยที่สามารถอ่านได้ จำเป็นต้อง RUN โปรแกรมที่เป็น Driver ภาษาไทยสำหรับ Card ควบคุมจอภาพก่อน โดยใช้กับ Card ภาษาไทยของยี่ห้อ Super Dyna ดังนั้นจะต้องมี File ชื่อ THAID.COM รวมอยู่ในแผ่น "THAI OCR"

7.2.3 <u>การใช้งานของโปรแกรม</u>

เมื่อนำแผ่นดิสค์ "THAI OCR" มาทำการ Directory จะปรากฏ File อย่างน้อยดังนี้

COMMAND	COM		
THAID	COM	:	THAI DRIVER ของ Super Dyna
OCR	EXE	:	โปรแกรมวิเคราะห์การรู้จำตัวอักษร
THAICODE	DAT	:	ฐานข้อมูลรหัสอักษรต้นแบบ
TEXT	OCR	:	ตัวอย่างข้อมูลภาพกลุ่มข้อความที่
DATA	OCR	:	อ่านได้จากเครื่อง Handy Scanner

เมื่อนำแผ่นดิสค์ "THAI OCR" ใส่ในช่อง Drive A: แล้วก็พร้อมเริ่มการใช้งานตามลำดับดังนี้

<u>ขั้นที่ 1.</u> ทำการ RUN โปรแกรม THAI DRIVER ก่อนดังนี้

A>THAID (แล้วกด ENTER)

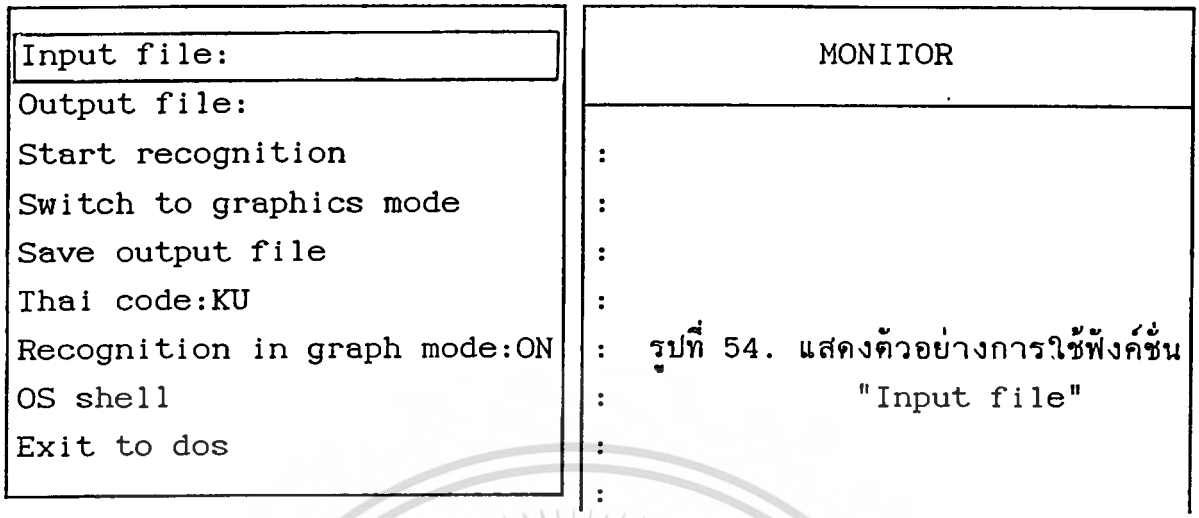
<u>ขั้นที่ 2.</u> แล้ว RUN โปรแกรมวิเคราะห์การรู้จำตัวอักษร ดังนี้

A>OCR (แล้วกด ENTER)

จะปรากฏภาพเมนูการใช้งาน ดังรูปที่ 53.

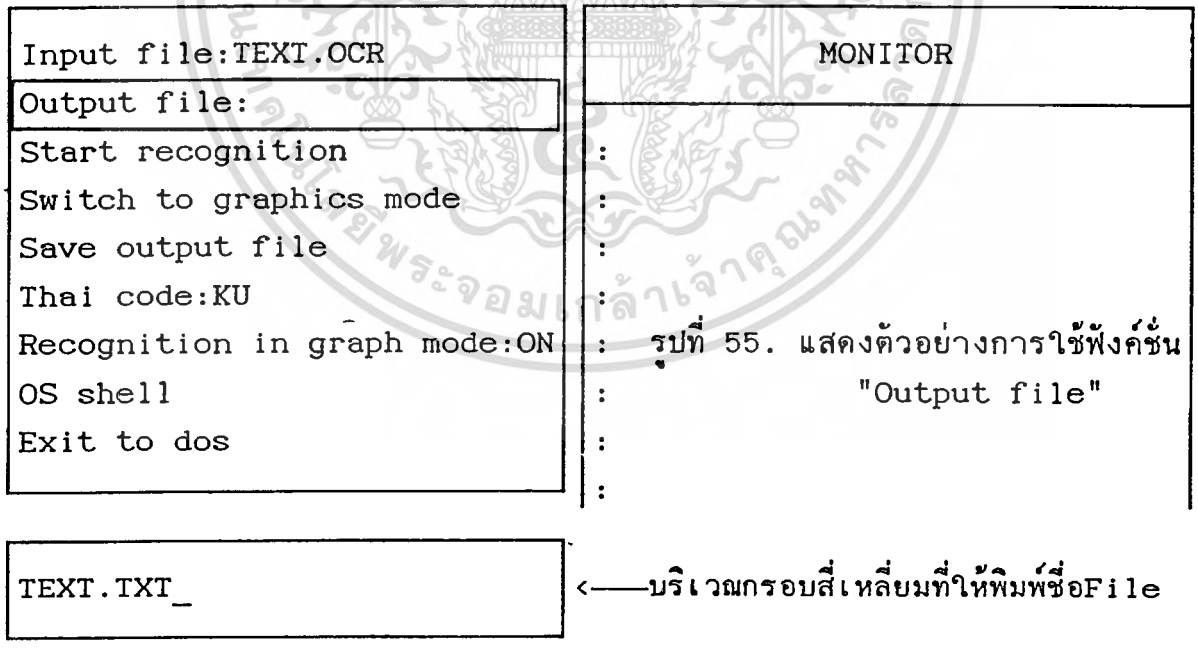
<u>ขั้นที่ 3.</u> เลื่อนแถบเคอร์เซอร์ให้ตรงกับฟังค์ชั่น "INPUT FILE" แล้วกด Eneter จะปรากฏช่องว่างกรอบสี่เหลี่ยมให้พิมพ์ชื่อ File ข้อมูลภาพที่ต้องการวิเคราะห์ดังตัวอย่างรูปที่ 54.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 54. แสดงตัวอย่างการใช้ฟังค์ชั่น "Input file"

ขั้นที่ 4. เลื่อนแถบเคอร์เซอร์ให้ตรงกับฟังค์ชั่น "Output file" แล้วกด Eneter จะปรากฏช่องว่างกรอบสี่เหลี่ยมให้พิมพ์ชื่อ File Output ผลลัพธ์จากการวิเคราะห์เสร็จแล้ว จะถูกบันทึกลงแผ่นดิสค์ในชื่อ File ที่ได้ระบุนี้ ดังรูป 55.



รูปที่ 55. แสดงตัวอย่างการใช้ฟังค์ชั่น "Output file"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นที่ 5. หากต้องการตรวจสอบ หรือดูลักษณะภาพข้อความของ Input file ที่อ่านเข้ามาแล้ว สามารถกระทำได้โดยเลื่อนแถบเคอร์เซอร์มาตรงกับฟังค์ชั่น "Switch to graphics mode" แล้วกด Enter จอภาพจะเปลี่ยนไปเป็นภาพกลุ่มข้อความ ที่เคยถูกอ่านจากเครื่อง Handy Scanner ผู้ใช้สามารถใช้กลุ่มคีย์เคอร์เซอร์เลื่อนภาพได้ทั้ง 4 ทิศทาง เพื่อตรวจสอบหรือดูลักษณะของภาพกลุ่มข้อความทั้ง File ได้ ดังแสดงตัวอย่างในรูปที่ 56 ก. และ 56 ข.

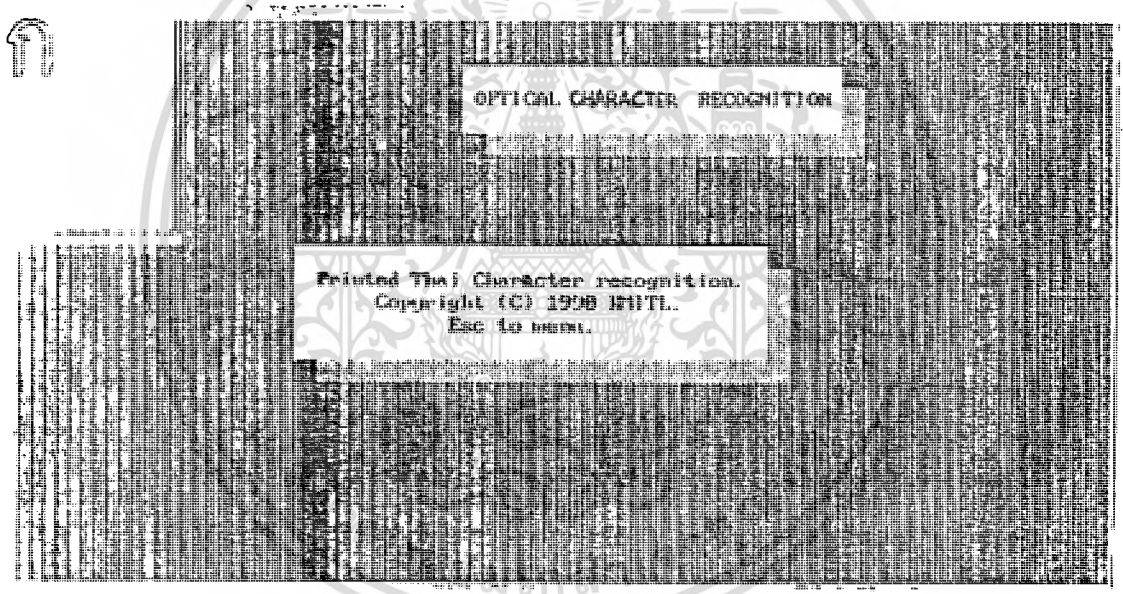
Input file:TEXT.OCR	MONITOR
Output file:TEXT.TXT	
Start recognition	:
Switch to graphics mode	:
Save output file	:
Thai code:KU	:
Recognition in graph mode:ON	:รูปที่ 56ก. แสดงตัวอย่างการใช้ฟังค์ชั่น
OS shell	: "Swithch to graphics mode"
Exit to dos	:
	:

รูปที่ 56 ข.ตัวอย่างการแสดงลักษณะภาพกลุ่มข้อความที่ถูกเลือกจากฟังค์ชั่นนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นที่ 6. หากต้องการเปลี่ยน รหัส ASCII ภาษาไทยที่จะเป็นผลลัพธ์จากรหัสเกษตร (KU) เป็น รหัส สมอ. (TIS) ให้เลื่อนแถบเคอร์เซอร์ตรงกับฟังค์ชั่น "Thai Code" แล้วกด Enter รหัสจะเปลี่ยนจาก KU เป็น TIS หรือเปลี่ยนจาก TIS เป็น KU (ภาคผนวก ก.) โดยปกติโปรแกรมจะตั้งไว้เป็นรหัส KU

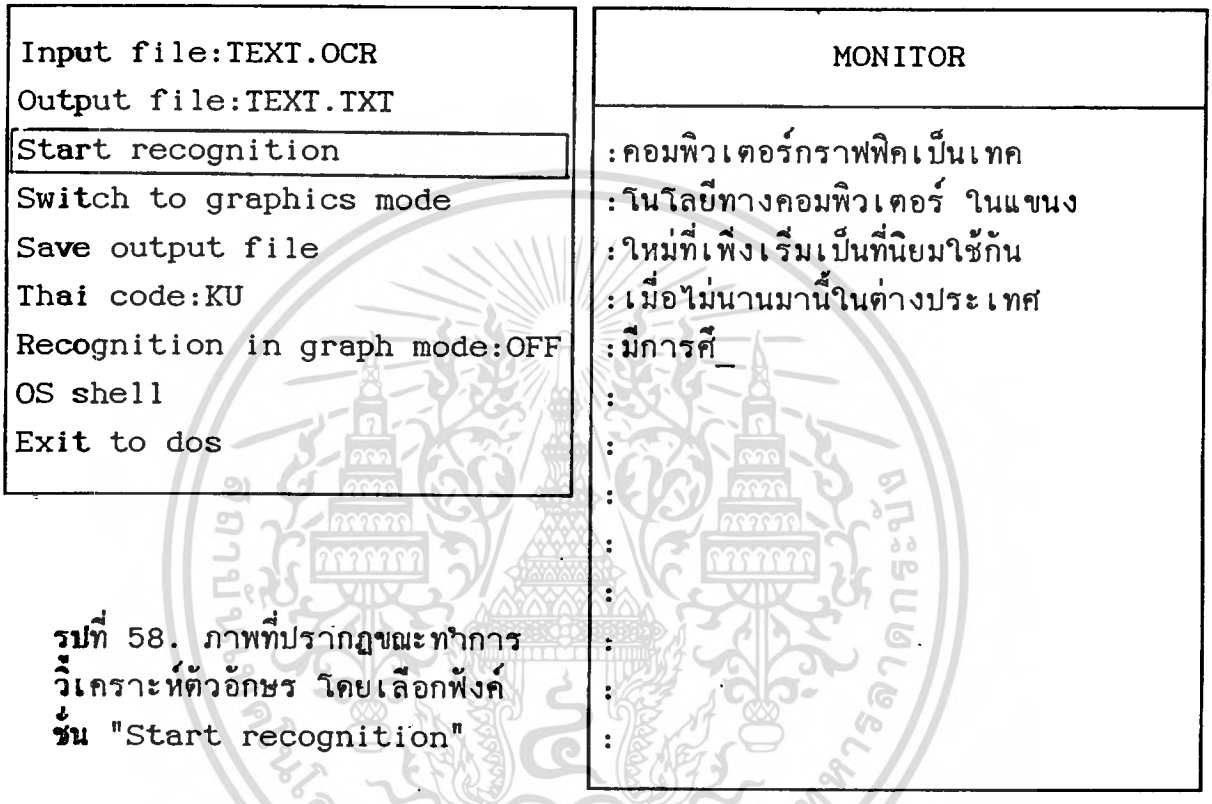
ขั้นที่ 7. หากไม่ต้องการแสดงภาพตัวอักษรที่กำลังวิเคราะห์ให้เห็นสามารถกำหนดได้โดยเลื่อนแถบเคอร์เซอร์ให้ตรงกับฟังค์ชั่น"Recognition in graph mode:" แล้วกด Enter คำว่า "ON" จะเปลี่ยนเป็น "OFF" นั่นคือขณะทำการวิเคราะห์ข้อมูลจอภาพจะปรากฏเมนูเช่นเดียวกับรูปที่ 53. คือพร้อมจะแสดงผลลัพธ์ที่วิเคราะห์แล้วให้ปรากฏออกมาในส่วน MONITOR แต่หากฟังค์ชั่นนี้เป็น "ON" ขณะทำการวิเคราะห์จอภาพจะเปลี่ยนเป็นการแสดงภาพลักษณะรูปร่างตัวอักษรที่กำลังวิเคราะห์ให้เห็นด้วย ดังตัวอย่างภาพในรูปที่ 57.



รูปที่ 57. ภาพตัวอย่างหน้าจอขณะกำลังวิเคราะห์ตัวอักษรเมื่อได้เลือกฟังค์ชั่น "Recognition in graph mode: เป็น "ON"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นที่ 8. เมื่อได้เลือกฟังค์ชั่นสำหรับเตรียมการใช้งานพร้อมแล้วก็สามารถเริ่มทำการวิเคราะห์การรู้จำตัวอักษรได้ โดยเลื่อนแถบเคอร์เซอร์ให้ตรงกับฟังค์ชั่น "Start recognition" แล้วกด Enter ขบวนการวิเคราะห์การรู้จำรูปแบบอักษรภาษาไทย ก็จะเริ่มต้นทันที อักษรที่วิเคราะห์แล้วจะปรากฏผลลัพธ์ขึ้นในบริเวณส่วน "MONITOR" ดังแสดงตัวอย่างภาพขณะทำการวิเคราะห์ และปรากฏผลลัพธ์ในรูปที่ 58.



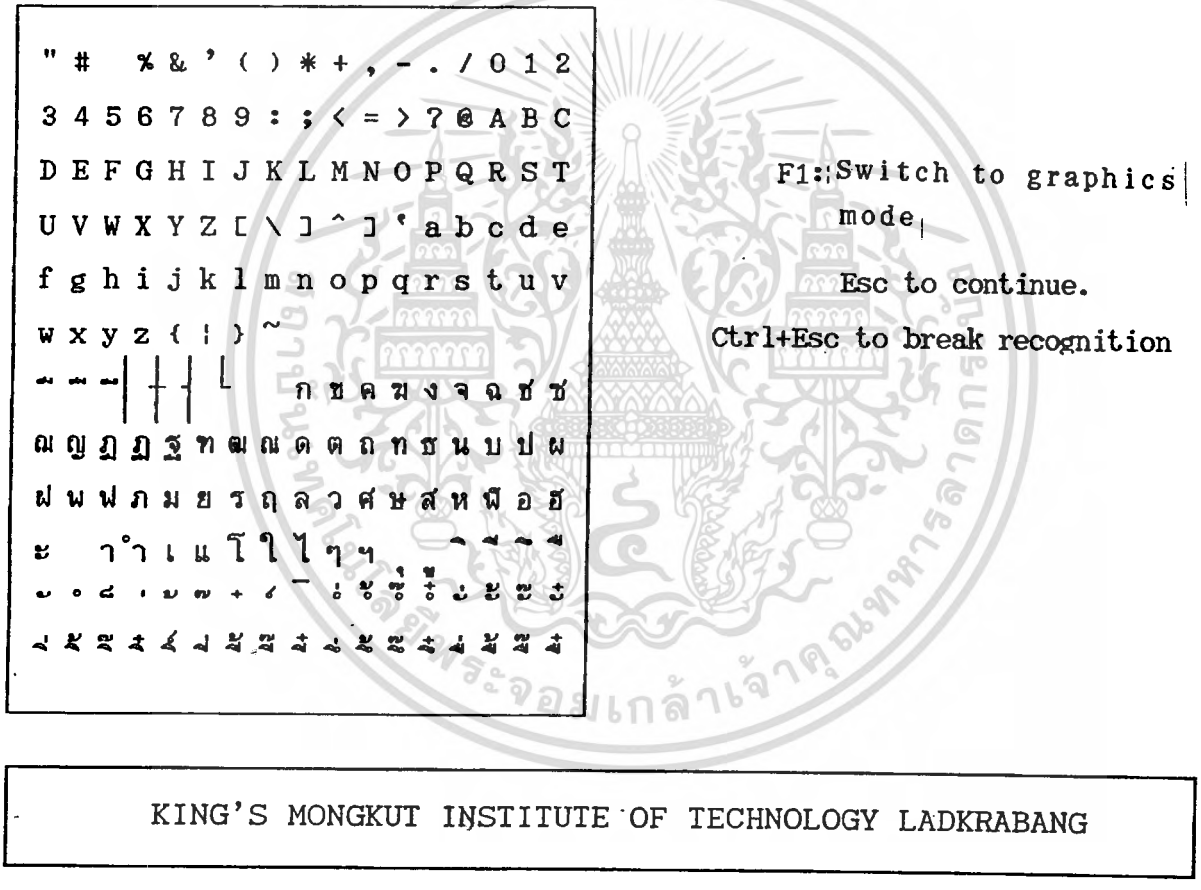
รูปที่ 58. ภาพที่ปรากฏขณะทำการวิเคราะห์ตัวอักษร โดยเลือกฟังค์ชั่น "Start recognition"

หากขบวนการวิเคราะห์ได้พบภาพอักษรรูปแบบอื่น ที่ไม่เคยวิเคราะห์ภาพต้นแบบมาก่อน จอภาพจะเปลี่ยนปรากฏเป็นภาพตาราง ASCII CODE ให้ผู้ใช้ทำการเลือกรหัส ASCII เอง ภาพตารางรหัส ASCII ที่ปรากฏขึ้นดังแสดงในรูปที่ 59. ขณะเดียวกันผู้ใช้สามารถกดคีย์ F1 สำหรับให้โปรแกรมปรากฏภาพรูปร่างอักษรที่ไม่สามารถวิเคราะห์ได้ให้เห็น ดังตัวอย่างแสดงในรูปที่ 60. และหลังจากพิจารณาภาพรูปร่างอักษรแล้ว ให้กด Esc สำหรับกลับมายังภาพแสดงตารางรหัส ASCII CODE อีกครั้งเพื่อเลือกตัดสินใจว่ารูปร่างอักษรที่กำลังพิจารณานั้นตรงกับรหัส ASCII อักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

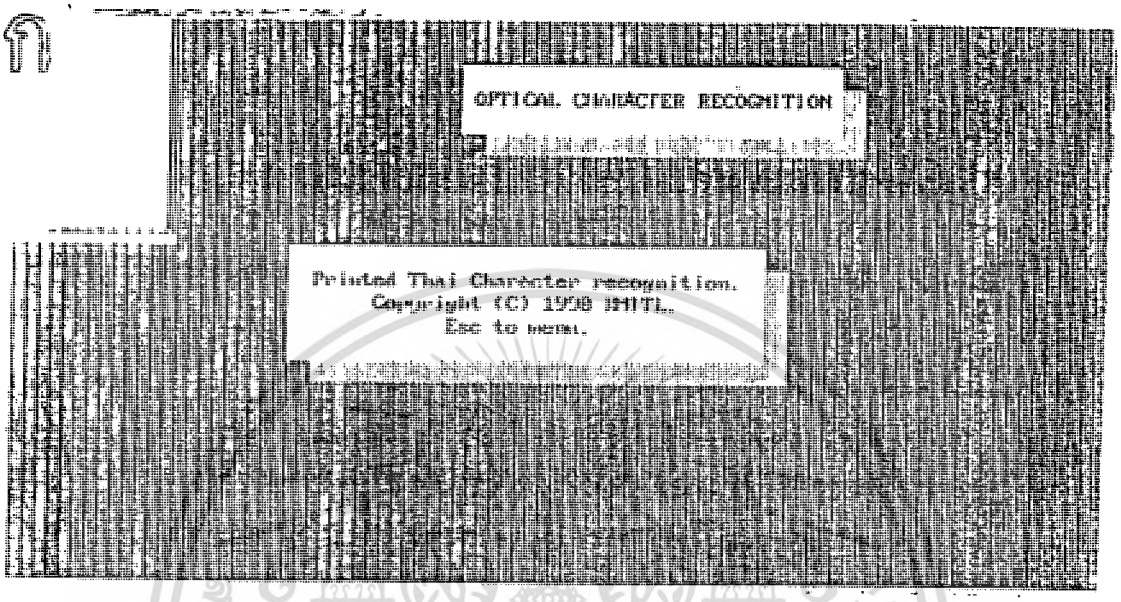
ตัวใด เมื่อผู้ใช้ได้เลือกรหัส ASCII โดยการกด Enter แล้วค่ารหัสที่เป็นคุณสมบัติเฉพาะของอักษรตัวนั้น จะถูกบันทึกรวมใน File ฐานข้อมูลรหัสอักษรต้นแบบที่ชื่อ THAICODE.DAT แล้วดำเนินขบวนการวิเคราะห์ภาพอักษรตัวต่อไป

หรือหากไม่ต้องการบันทึกค่ารหัส ที่เป็นคุณสมบัติเฉพาะของอักษรตัวนั้นอาจเนื่องจากภาพที่กำลังพิจารณานั้นไม่ใช่รูปร่างที่เป็นตัวอักษร หรือเป็นภาพตัวอักษรที่ไม่สมบูรณ์ ก็สามารถกำหนดไม่ให้มีการบันทึกรหัสเพิ่มเติมได้ โดยการกดคีย์ Esc โปรแกรมก็จะเลิกวิเคราะห์ภาพรูปร่างนั้น แล้วดำเนินขบวนการวิเคราะห์ภาพอักษรที่เหลือต่อไปจนหมดภาพตัวอักษรจึงสิ้นสุดขบวนการวิเคราะห์ หรือหากต้องการยกเลิกการวิเคราะห์ ขณะกำลังดำเนินการวิเคราะห์อยู่ ให้กดคีย์ Ctrl กับคีย์ Esc พร้อมกัน การวิเคราะห์ก็จะหยุดทันที



รูปที่ 59. ภาพตารางรหัส ASCII CODE ที่ปรากฏขึ้นกรณีที่ขบวนการไม่สามารถวิเคราะห์ได้ โดยให้ผู้ใช้ตัดสินใจเลือกรหัสเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 60. ภาพที่ปรากฏเมื่อกดคีย์ F1 เพื่อพิจารณาลักษณะรูปร่างอักษร ที่ไม่สามารถวิเคราะห์ได้ ก่อนกลับไปตัดสินใจเลือกรหัส ASCII

ขั้นที่ 9. เมื่อสิ้นสุดขบวนการวิเคราะห์ และได้ผลลัพธ์ปรากฏที่ส่วน "MONITOR" แล้ว ลักษณะของกลุ่มคำอาจจัดเรียงกันไม่ถูกหลักไวยากรณ์ ทั้งนี้เนื่องจาก ในระหว่างขบวนการวิเคราะห์ ได้ทำการแยกภาพอักษรออกจากภาพประโยชน์ทีละตัวนั้น ได้ไล่เรียงตามลำดับก่อนหลัง ไม่ได้วิเคราะห์เรียงลำดับตามรูปแบบไวยากรณ์ที่ถูกต้อง แต่ข้อความซึ่งเป็นผลลัพธ์ที่วิเคราะห์ได้ จะปรากฏใน MONITOR โดยจัดเรียงอักษรคล้ายข้อความเดิม และเมื่อได้ผลลัพธ์ครบทุกตัวอักษรแล้ว ให้เลือกฟังค์ชั่น Save output file เพื่อบันทึกผลลัพธ์ที่วิเคราะห์ได้ ลงในแผ่นบันทึกข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นที่ 10. เมื่อทำการวิเคราะห์ File ข้อมูลภาพสิ้นสุดแต่ละครั้งแล้ว หากต้องการใช้งานโปรแกรมอื่น เช่น โปรแกรมที่จะทำหน้าที่อ่านภาพข้อความจากเครื่อง Handy Scanner จากบทความอื่นเข้ามาใหม่ ก็สามารถยกเลิกการใช้โปรแกรมชั่วคราวได้ โดยเลื่อนแถบเคอร์เซอร์ให้ตรงกับฟังค์ชั่น "Os Shell"โปรแกรมจะSwitch การทำงาน ออกมายังเครื่องหมาย A> ซึ่งเป็นส่วนการทำงานของ DOS พร้อมที่จะให้ RUN โปรแกรมอื่นๆ ก่อนได้ และเมื่อสิ้นสุดการใช้งานโปรแกรมอื่นแล้ว และต้องการกลับมาทำการวิเคราะห์อักษรชุดใหม่อีก ก็สามารถกลับเข้ายังโปรแกรมได้ใหม่โดยไม่ต้อง RUN โปรแกรม OCR ซ้ำอีกครั้ง เพียงพิมพ์ข้อความดังนี้

A>EXIT (แล้วกด Enter)

จะปรากฏภาพเมนูของโปรแกรมดังรูปที่ 53. ตามเดิม

ขั้นที่ 11. และเมื่อสิ้นสุดการใช้โปรแกรมแล้ว ต้องการยกเลิกโปรแกรมจริงๆ ก็สามารถทำได้โดยเลื่อนแถบเคอร์เซอร์ให้ตรงฟังค์ชั่น "Exit to Dos" แล้วกด Enter ก็จะสิ้นสุดการทำงานของโปรแกรมทันที

7.2.4 อุปกรณ์ที่ใช้ประกอบร่วม

อุปกรณ์ที่ใช้ประกอบร่วมการทำงาน ของโปรแกรมวิเคราะห์การรู้จำตัวอักษรภาษาไทยนี้ ประกอบด้วย

1. เครื่องไมโครคอมพิวเตอร์ตระกูล IBM รุ่น XT ,AT286 หรือ AT386 หน่วยความจำหลัก ไม่น้อยกว่า 640 Kb และ Disk Drive 1 ตัว
2. จอภาพโมโนโครม พร้อม Card ภาษาไทยยี่ห้อ Super Dyna
3. เครื่อง Handy Scanner พร้อม Interface Card
4. เมาส์
5. เครื่องพิมพ์
6. โปรแกรมควบคุมการทำงาน Handy Scanner
7. แผ่นโปรแกรมวิเคราะห์การรู้จำอักษรภาษาไทย "THAI OCR"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 8
การทดสอบโปรแกรม

เมื่อได้พัฒนาโปรแกรมให้สามารถใช้งานได้ตามทฤษฎีที่ได้กล่าวมาแล้ว เพื่อให้แน่ใจว่าโปรแกรมที่พัฒนาขึ้นสามารถใช้งานได้ผล และมีประสิทธิภาพมากน้อยเพียงใด จึงได้ทำการทดสอบโปรแกรมกับอักษรข้อความจริงที่เลือกจากเอกสารที่เป็นสิ่งพิมพ์ต่าง ๆ โดยเลือกรูปแบบของตัวอักษรหลายลักษณะจากขนาดแตกต่างกัน ในการทดสอบโปรแกรม ได้ดำเนินการทดสอบดังนี้

8.1 การรู้จำให้กับโปรแกรม (Learning)

ภาพตัวอักษรทั้งพยัญชนะและสระทุกตัวตามรูปที่ 38. จะถูกอ่านด้วยเครื่อง Handy Scanner แล้วนำภาพข้อมูลอักษรเหล่านั้น ผ่านขั้นตอนการแยกตัวอักษรออกจากประโยค และผ่านขบวนการแยกกลุ่มตลอดจนแยกตัวอักษรออกจากกลุ่ม จนได้ค่ารหัสชุด A-Code ซึ่งเป็นคุณสมบัติของอักษรแต่ละตัว บันทึกรหัสชุด A-Code เก็บเป็นฐานข้อมูลเพื่อการเปรียบเทียบกับชุดรหัส A-Code ของอักษรตัวใหม่ที่ไม่ทราบว่าเป็นอักษรอะไร ให้สามารถระบุได้ถูกต้อง

8.2 การทดสอบอักษรจากบทความกับอักษรต้นแบบที่มีรูปแบบเดียวกัน

จากอักษรต้นแบบที่ได้เคยทำการทดลองวิเคราะห์เพื่อการรู้จำให้โปรแกรมตามขบวนการในหัวข้อ 8.1 เมื่อนำบทความหนึ่ง ที่มีอักษรจำนวน 650 ตัว ซึ่งเป็นอักษรรูปแบบและขนาดเดียวกันกับอักษรต้นแบบ ดังแสดงภาพบทความไว้ในรูปที่ 61. มาทำการทดสอบกับโปรแกรม ผลปรากฏสามารถวิเคราะห์ได้ถูกต้อง 634 ตัวและไม่สามารถวิเคราะห์ได้ 16 ตัว ทั้งนี้เพราะข้อความบางประโยคเช่นคำว่า "กราฟฟิค" หรือคำว่า "เป็น" มีอักษรที่เป็นสระบนติดกับอักษรที่เป็นพยัญชนะคือ ฟิ หรือ ป็ จึงถูกวิเคราะห์ว่าเป็นอักษรรูปแบบใหม่ที่ไม่มีในอักษรต้นแบบและไม่สามารถวิเคราะห์ได้

คอมพิวเตอร์กราฟฟิคเป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้าทดลองและนำความรู้ที่ได้ มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพ มีความเด่นชัดสามารถใช้สื่อความหมายได้ดีกว่า ข้อมูลที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอนและความยุ่งยากต่าง ๆ ที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลข จนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อย และเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้น และมีโอกาสได้ใช้สมองกับความคิดและการสร้างสรรค์ใหม่ ๆ แทน

รูปที่ 61. ภาพบทความที่ใช้ทดสอบ มีรูปแบบและขนาดเดียวกันกับอักษรต้นแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8.3 การทดสอบกลุ่มอักษรหลายรูปแบบและหลายขนาดกับอักษรต้นแบบ

และเพื่อให้การวิจัยครั้งนี้ ตรงตามวัตถุประสงค์ที่จะให้โปรแกรมสามารถวิเคราะห์การรู้จำอักษรภาษาไทยได้หลายรูปแบบและหลายขนาด จึงได้นำภาพอักษรที่มีรูปแบบและขนาดแตกต่างกันจำนวน 14 ชุด ดังแสดงในรูปที่ 62. แต่ละชุดซึ่งประกอบด้วยสระ พยัญชนะจำนวนชุดละ 64 ตัวรวมทั้งสิ้น 896 ตัว มาทำการทดสอบเปรียบเทียบกับอักษรต้นแบบที่เคยวิเคราะห์ผ่านมาแล้วในบทที่5. และบทที่6. โดยทำการทดสอบทีละชุด และเมื่อชุดอักษรใดที่กำลังทำการวิเคราะห์ หากปรากฏว่ามีอักษรตัวใดไม่สามารถวิเคราะห์ได้ก็จะถูกเลือกรหัส ASCII ให้ และถือเป็นอักษรต้นแบบที่เพิ่มขึ้นใหม่ด้วย ดังได้แสดงตารางผลการทดสอบกลุ่มอักษรรูปแบบและขนาดแตกต่างกันแสดงไว้ในตารางที่ 13.

แบบ	ตัวอย่างอักษร	แบบ	ตัวอย่างอักษร	แบบ	ตัวอย่างอักษร
1	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ	6	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ	11	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ
2	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ	7	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ	12	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ
3	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ	8	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ	13	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ ะ ิ ี ึ ื ุ ู เ แ ๆ ่ ้ ๊ ๋ ์ ๎ ๏ ๐ ๑ ๆ ๆ
4	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ	9	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ	14	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ ะ ิ ี ึ ื ุ ู เ แ ๆ ่ ้ ๊ ๋ ์ ๎ ๏ ๐ ๑ ๆ ๆ
5	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ	10	ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ		

รูปที่ 62. ภาพอักษรที่มีรูปแบบและขนาดแตกต่างกันจำนวน 14 ชุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชุดอักษรลำดับที่	ชื่อแบบตัวอักษร	จำนวนอักษร	วิเคราะห์ถูกต้อง		วิเคราะห์ไม่ได้		วิเคราะห์ผิดพลาด	
1.	DENDROBIUM (6)	64	12	18.7%	52	81.3%	-	-
2.	ROYAL ORCHID (8)	64	32	50.0%	32	50.0%	-	-
3.	CATTALEYA (8)	64	38	59.4%	26	40.6%	-	-
4.	DENDROBIUM (8)	64	35	54.5%	29	45.5%	-	-
5.	GIGANTEA (8)	64	39	60.9%	25	39.1%	-	-
6.	RAINY (9)	64	53	82.8%	11	17.2%	-	-
7.	AIREDES (10)	64	57	89.0	7	11.0%	-	-
8.	CATTALEYA (10)	64	48	75.0	16	25.0%	-	-
9.	DENDROBIUM (10)	64	54	84.3	10	15.7%	-	-
10.	GIGANTEA (10)	64	61	95.3	3	4.7%	-	-
11.	GOLDEN ORCHID (10)	64	59	92.2	5	7.8%	-	-
12.	RAINY (10)	64	54	84.4	10	15.6%	-	-
13.	ARLUG (12)	64	32	50.0%	31	46.7%	2	3.%
14.	ARLUG (12I)	64	7	11.0%	57	89.0%	-	-

ตารางที่ 13. เปรียบเทียบผลการทดสอบอักษรรูปแบบต่าง ๆ ทั้ง 14 ชุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8.4 การทดสอบบทความกับอักษรต้นแบบหลายรูปแบบและหลายขนาด

เมื่อได้ทำการวิเคราะห์ภาพอักษรต้นแบบรูปแบบต่าง ๆ ทั้ง 14 ชุด และรวมกับรหัสต้นแบบเดิมเป็นทั้งหมด 15 ชุด ซึ่งเท่ากับมีฐานข้อมูลรหัสของอักษรต้นแบบเพิ่มเติมจากเดิมในส่วนของฐานข้อมูลรหัสต้นแบบทั้งหมด 15 ต้นแบบ ซึ่งเป็นการเพิ่มการเรียนรู้การรู้จำรูปแบบอักษรต่าง ๆ ของโปรแกรมได้มากขึ้น

ดังนั้นเมื่อนำบทความที่อักษรมีรูปแบบเช่นเดียวกันกับอักษรชุดต่าง ๆ ทั้ง15 ชุดมาทำการทดสอบอีกครั้งก็สามารถสรุปผลการทดสอบได้ ลักษณะของข้อมูลภาพที่เป็นบทความและนำมาใช้ทดสอบทั้ง 15 ชุด แสดงไว้ในภาคผนวก ค. และรูปที่ 63. แสดงตัวอย่างบางส่วนของข้อมูลที่เป็นบทความทั้ง 15 ชุด ข้อมูลภาพบทความแต่ละชุดประกอบด้วยอักษร 650 ตัว รวมจำนวนตัวอักษรข้อมูลภาพบทความที่ใช้ทดสอบทั้งสิ้น 9750 ตัว สามารถเปรียบเทียบผลการทดสอบได้ดังในตารางที่ 14.

แบบที่	ตัวอย่างบทความ	แบบที่	ตัวอย่างบทความ
1	ทางคอมพิวเตอร์ในแขนงใหม่ที่เพิ่ นิยมใช้กันเมื่อไม่นานมานี้ ในต่าง	9	เตอร์ในแขนงใหม่ ที่เพิ่งเริ่ม ไม่นานมานี้ ในต่างประเทศนั้
2	ทางคอมพิวเตอร์ในแขนงใหม่ที่เพิ่ง นิยมใช้กันเมื่อไม่นานมานี้ ในต่างปร	10	ทางคอมพิวเตอร์ในแขนงใ ที่นิยมใช้กันเมื่อไม่นานมา
3	ทางคอมพิวเตอร์ในแขนงใหม่ที่เพิ่ นิยมใช้กันเมื่อไม่นานมานี้ ในต่าง	11	ทางคอมพิวเตอร์ในแขนงใ ที่นิยมใช้กันเมื่อไม่นานมา
4	ทางคอมพิวเตอร์ในแขนงใหม่ที่ นิยมใช้กันเมื่อไม่นานมานี้ ในต่	12	แขนงใหม่ ที่เพิ่งเริ่มเป็นที่นิย ต่างประเทศนั้นมีการศึกษาค้น
5	ทางคอมพิวเตอร์ในแขนงใหม่ นิยมใช้กันเมื่อไม่นานมานี้ ใน	13	และความสามารถในการ จากข้อมูลที่เป็นตัวเลข จนเกิน
6	ทางคอมพิวเตอร์ในแขนงใหม่ที่ นิยมใช้กันเมื่อไม่นานมานี้ ในต่	14	ทางคอมพิวเตอร์ในแขน เป็นที่นิยมใช้กันเมื่อไม่น
7	เตอร์ในแขนงใหม่ ที่เพิ่งเริ่มเ ไม่นานมานี้ ในต่างประเทศนั้	15	เตอร์ในแขนงใหม่ ที่เพิ่งเริ่มเ ไม่นานมานี้ ในต่างประเทศนั้
8	เตอร์ในแขนงใหม่ ที่เพิ่งเริ่มเ ไม่นานมานี้ ในต่างประเทศนั้		

รูปที่ 63. ภาพตัวอย่างบางส่วนของบทความทั้ง 15 ชุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชุดอักษรลำดับที่	ชื่อแบบตัวอักษร	จำนวนอักษร	วิเคราะห์ถูกต้อง		วิเคราะห์ไม่ได้		วิเคราะห์ผิดพลาด	
1.	DENDROBIUM (6)	650	557	85.7%	82	12.6%	11	1.7%
2.	ROYAL ORCHID (8)	650	624	96.0%	24	3.7%	2	0.3%
3.	CATTALEYA (8)	650	613	94.3%	37	5.7%	-	-
4.	DENDROBIUM (8)	650	607	93.8%	40	6.2%	-	-
5.	GIGANTEA (8)	650	618	95.0%	31	4.8%	1	0.2%
6.	RAINY (9)	650	631	97.0%	19	3.0%	-	-
7.	AIREDES (10)	650	633	97.4%	17	2.6%	-	-
8.	CATTALEYA (10)	650	639	98.3%	11	1.7%	-	-
9.	DENDROBIUM (10)	650	632	97.2%	18	2.8%	-	-
10.	GIGANTEA (10)	650	634	97.5%	16	2.5%	-	-
11.	GOLDEN ORCHID 10	650	636	97.8%	14	2.2%	-	-
12.	RAINY (10)	650	640	98.5%	10	1.5%	-	-
13.	ARLUG (12)	650	601	92.5%	45	7.0%	4	0.5%
14.	ARLUG (12I)	650	614	94.5%	29	4.5%	7	1.0%

ตารางที่ 14. เปรียบเทียบผลการทดสอบบทความอักษรรูปแบบต่างกัน 15 ชุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชุดอักษรลำดับที่	ชื่อแบบตัวอักษร	จำนวนอักษร	วิเคราะห์ถูกต้อง		วิเคราะห์ไม่ได้		วิเคราะห์ผิดพลาด	
15.	ROYAL ORCHID (10)	650	634	97.5%	16	2.5%	-	-
	รวม	9750	9313	95.5%	413	4.3%	24	0.2%

ตารางที่ 14.(ต่อ) เปรียบเทียบผลการทดสอบบทความอักษรรูปแบบต่างกัน 15 ชุด

กลุ่มอักษรที่ใช้ทดสอบ	อักษรต้นแบบชุดแรก	บทความชุดแรก	อักษร 14 แบบ	บทความ 15 แบบ
จำนวนตัวอักษร	74	650	896	9750
วิเคราะห์ถูกต้อง	100%	97.5%	65%	95.5%
วิเคราะห์ไม่ได้	-	2.5%	34.5%	4%
วิเคราะห์ผิดพลาด	-	-	0.5%	0.5%
รวม	100%	100%	100%	100%

ตารางที่ 15. สรุปผลการทดสอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 9
บทสรุปและข้อเสนอแนะ

9.1 เปรียบเทียบผลงานวิจัยที่เกี่ยวข้อง

จากการดำเนินศึกษาวิวัยวิธีการนี้ ได้ผลที่แตกต่างจากงานวิจัยอื่น ๆ ที่ผ่านมาดังสรุปได้ดังนี้

ลักษณะที่เปรียบเทียบ	1.	2.	3.	4.	5.	งานวิจัยที่ทำ
ก. เทคนิควิธีการ	A	B	C	D	D	E
ข. ลักษณะตัวอักษรที่ใช้	ตัวพิมพ์	ลายมือ	ลายมือ	ลายมือ	ลายมือ	ตัวพิมพ์
ค. วิธีการแยกตัวอักษร	ไม่เน้น	ไม่กล่าว	แยกเอง	ไม่เน้น	1บรรทัด	หลายบรรทัด
ง. ความถูกต้อง	100%	98%	99%	99%	98%	95.5%
จ. ความเร็ว	ไม่กล่าว	ไม่กล่าว	1.35	ไม่กล่าว	1.1	1.0

ตารางที่ 16. เปรียบเทียบผลงานวิจัยที่เกี่ยวข้อง

หมายเหตุ ชื่อผลงานวิจัยและวิธีการที่ใช้

1. การจดจำรูปแบบตัวพิวพ์อักษรภาษาไทย
2. การจดจำลายมือเขียนตัวอักษรภาษาไทยในระบบออนไลน์
3. การจดจำตัวอักษรลายมือเขียนภาษาไทย
4. การจัดจำรูปแบบตัวอักษรคัดลายมือภาษาไทยโดยวิธีแยกลักษณะเด่น
5. การออกแบบพจนานุกรมสำหรับการเรียนรู้อักษรคัดลายมือไทย-อังกฤษอัตโนมัติบนเครื่องไมโครคอมพิวเตอร์

A : วัดค่าสัมประสิทธิ์ของความเหมือนแล้วกระจายแบบคาร์ฮูเนนโลบ

B : วิเคราะห์โครงสร้างจากคอนทัวร์ตามลักษณะของฟรีแมนแชนโค้ด

C : แบ่งกลุ่มตามจำนวนหัวของตัวอักษร และใช้ความสัมพันธ์ระหว่างจุด

D : แยกลักษณะเด่นของรหัสคอนเซนเทรดโค้ด

E : ค้นหาลักษณะลายเส้นของตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

9.2 การแยกภาพอักษรออกจากภาพประโยค

การแยกภาพอักษรออกจากภาพประโยคทีละตัวนั้น นับเป็นปัญหาสำคัญอันดับแรกของขบวนการการรู้จำภาษาไทย ทั้งนี้เพราะลักษณะการเรียงประโยคของภาษาไทยมีหลายระดับ การแยกแต่ละบรรทัดออกจากกันมักทำได้ยุ่งยาก วิธีการที่ใช้ในงานวิจัยครั้งนี้ไม่ได้คำนึงถึงตำแหน่งบรรทัดแต่ละบรรทัด แต่จะถือเอาค่าจุดพิกัดกึ่งกลางฐานของอักษรแต่ละตัวเป็นตำแหน่งอ้างอิง หลังจากวิเคราะห์จนทราบว่าเป็นอักษรอะไรแล้วจึงจัดเรียงผลลัพธ์ที่ได้ตามลำดับค่าจุดพิกัดที่เคยเป็นของอักษรตัวนั้น ก็สามารถให้ผลออกมาได้เช่นเดียวกับการแยกข้อความทีละบรรทัดก่อนนับเป็นผลดีต่อการแก้ปัญหา

เนื่องจากการแยกภาพอักษรออกจากภาพประโยค วิธีการนี้อาศัยความถูกต้องจากการวิเคราะห์เป็นหลัก หากการวิเคราะห์ถูกต้องแม่นยำ ผลการจัดเรียงรูปประโยคจะถูกต้องตรงตามประโยคเดิมทุกประการ แต่ถ้าการวิเคราะห์ผิดพลาด เช่น วิเคราะห์สระอุ " ุ " เป็นไม้เอก " ่ " ซึ่งเงื่อนไขการขึ้นบรรทัดใหม่ได้กล่าวว่า สระบนที่ตามหลังพยัญชนะให้ขึ้นบรรทัดใหม่ อาจทำให้การจัดเรียงรูปประโยคและข้อความผิดจากรูปประโยคเดิมได้

และการแยกอักษรออกจากภาพประโยควิธีนี้ ไม่สามารถแยกภาพอักษรที่ติดกันได้ซึ่งมักเกิดขึ้นบ่อยมากในลักษณะของสิ่งตีพิมพ์ทั่วไป

สำหรับความเร็วเฉลี่ยในขั้นตอนการแยกตัวอักษรออกจากประโยคเฉลี่ยตัวอักษรละ 0.25 วินาที แบ่งเป็นเปอร์เซนต์ของขั้นตอนต่าง ๆ ดังนี้

เวลาเฉลี่ยในการแยกตัวอักษรออกจากประโยคต่อหนึ่งตัว ใช้เวลาประมาณ 0.25 วินาที แบ่งเป็น	
ขั้นการทำขอบภาพ(เฉลี่ยทั้งหน้า)	25%
ขั้นการคัดลอกและย้ายจุดภาพ	60%
ขั้นการจัดเรียงตัวอักษร	15%

ตารางที่ 17. เปอร์เซนต์เวลาในการแยกตัวอักษรออกจากภาพประโยค

9.3 การจัดและแยกกลุ่มอักษ

9.3.1 ขั้นการจัดกลุ่ม

ในการทดลองได้นำอักษรรูปแบบและขนาดแตกต่างกัน 15 แบบ แต่ละแบบประกอบด้วยพยัญชนะ ก ถึง ฮ และสระทุกตัวรวมแบบละ 64 ตัวอักษร รูปแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต่าง ๆ ของตัวอักษรเหล่านี้จะนำมาเป็นต้นแบบในการเรียนรู้ ให้กับขบวนการรู้จำ รวมตัวอักษรต้นแบบทั้งหมด 960 ตัว ทำการจัดแยกกลุ่มอักษรเหล่านี้ตามวิธีการในบทที่5. ปรากฏจำนวนกลุ่มเพิ่มขึ้นจากเดิม 7 กลุ่มเป็น 9 กลุ่มทั้งนี้เพราะมีลักษณะตัวอักษรบางตัวที่รูปร่างแตกต่างไปจากเดิม เนื่องจากการเอียงหรือโค้งมากไปของตัวอักษร กลุ่มที่เพิ่มขึ้น 2 กลุ่มแสดงไว้ในตารางที่ 18.

นอกจากนั้นอักษรที่เคยประจำแต่ละกลุ่ม เมื่อเพิ่มรูปแบบตัวอักษรมากขึ้นจะมีอักษรบางตัวเข้าไปปะปนอยู่กับอักษรกลุ่มอื่น ที่ไม่ใช่กลุ่มประจำของอักษรเหล่านั้น เป็นผลให้การวิเคราะห์เกิดการผิดพลาดได้ ทั้งนี้เนื่องจากขนาดของอักษรมีขนาดเล็กเกินไป หรือมีลักษณะโค้งเอียงมากเกินไป

ลำดับที่	ชื่อกลุ่ม	ลักษณะ Q-Code	รหัสQ-code ที่ตำแหน่งCG	ตัวอักษร
8.	6	⌜	6	ๆ ช
9.	9	⌟	9	ๅ

ตารางที่ 18. กลุ่มรหัส Q-Code ที่เพิ่มจากเดิม

9.3.2 ขั้นการแยกตัวอักษรออกจากกลุ่ม

ในขั้นการแยกตัวอักษรออกจากกลุ่มโดยการวิเคราะห์ที่ละตัวอักษร จากจำนวนอักษรทั้งหมด 9750 ตัว นั้นสามารถวิเคราะห์แยกอักษรออกจากกลุ่มได้ผล 95.5% ทั้งนี้เพราะมีบางกลุ่ม เมื่อเพิ่มแบบตัวอักษรมากขึ้น อักษรจากกลุ่มอื่นได้เข้ามาปะปนในกลุ่ม และได้ปรากฏรหัสชุด A-Code ที่เหมือนกันเกิดขึ้นในกลุ่มเดียวกัน ทำให้เกิดการวิเคราะห์ผิดพลาดถึง 0.5% ซึ่งจะสามารถแก้ไขลดเปอร์เซนต์ความผิดพลาดลงได้โดยการเพิ่มเงื่อนไขเฉพาะของอักษรคู่ที่เกิดปัญหา และอีก 4%ไม่สามารถวิเคราะห์ได้ เนื่องจากอักษรที่อ่านเขามามีลักษณะไม่สมบูรณ์เช่น ขาด แหว่ง หรือติดกัน ทำให้ค้นหาชุดรหัส A-Code ของอักษรเหล่านั้นไม่พบ จึงไม่สามารถวิเคราะห์ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

9.4 ผลการทดลอง

ตัวอย่างแบบทดสอบที่ใช้ทั้ง 15 รูปแบบได้จัดพิมพ์ข้อความจำนวนรูปแบบละ 650 ตัวอักษร โดยมีข้อความเหมือนกันทั้ง 15 แบบรวมทั้งสิ้น 9750 ตัว ทำการทดสอบกับขบวนการรู้จำทั้งหมด ผลปรากฏว่าสามารถวิเคราะห์ได้ถูกต้องร้อยละ 95.5 วิเคราะห์ผิดพลาดร้อยละ 0.5 และไม่สามารถวิเคราะห์ได้ร้อยละ 4 ทั้งนี้เพราะลักษณะตัวอักษรที่ติดกัน ซึ่งไม่สามารถกำหนดเป็นรหัส ASCII ตัวใดตัวหนึ่งได้ อีกทั้งถ้าหากภาพข้อความมีขนาดเล็กเกินไปมักทำให้การวิเคราะห์ผิดพลาดง่าย

9.5 ข้อเสนอแนะ

การรู้จำอักษรตัวพิมพ์ภาษาไทยหลายรูปแบบและหลายขนาด ในวิธีการนี้สามารถนำไปใช้งานได้จริง และได้ผลเป็นที่น่าพอใจ และมีข้อเด่นอีกประการคือ อักษรตัวใดที่ไม่สามารถสามารถวิเคราะห์ได้ จะปรากฏรูปร่างของตัวอักษรให้ผู้ใช้ตัดสินใจเองได้ว่าคืออักษรอะไร พร้อมกับจัดเก็บบันทึกรหัสอักษรที่ไม่ทราบนั้นไว้ในกลุ่มเพื่อจะสามารถวิเคราะห์ได้เองในการวิเคราะห์ครั้งต่อไป

ในขั้นตอนการวิเคราะห์อักษรแต่ละตัว ลักษณะของ Storke ที่นำมาวิเคราะห์หาค่ารหัส A-Code ไม่ได้ทำตัวอักษรให้บาง (Tinning) แต่ใช้ขอบภาพของตัวอักษรที่เป็นเส้นคู่นำไปวิเคราะห์เลย ทำให้ขบวนการนี้ประมวลผลได้เร็วขึ้นแต่จะเป็นปัญหาการเขียนเงื่อนไขในโปรแกรม เพื่อให้ได้ Feature vector ตรงตามทฤษฎีของ Topological Feature Extraction นั้นทำได้ยุ่งยากซับซ้อน แต่ทั้งนี้จะเป็นผลดีต่อการวิเคราะห์เฉพาะส่วนเพราะจำนวนเส้น Stroke ที่มากกว่าการใช้ตัวอักษรที่ถูกทำให้บางแล้วมาวิเคราะห์ ทำให้เกิดลักษณะเฉพาะของค่า A-Code ที่แตกต่างกันออกไปได้ชัดเจนดียิ่งขึ้น หรือหากหัวของอักษรตัวใดมีลักษณะทึกหรือต้นการนำเส้นคู่ที่เป็นขอบภาพมาวิเคราะห์ จะได้ผลดีกว่าการนำตัวอักษรที่ทำให้บางแล้วมาวิเคราะห์เพราะลักษณะของหัวที่ทึบต้น เมื่อถูกทำให้บางจะมีลักษณะเหมือนเส้นตรงดังนั้นค่า Feature vector จะให้ผลที่ไม่สามารถแยกได้ชัดเจน ในทำนองเดียวกันในขั้นของการจัดกลุ่มนั้น เมื่อนำตัวอักษรมาคำนวณหาจุดศูนย์ถ่วงอักษรส่วนใหญ่จะปรากฏมีจุดศูนย์ถ่วงบริเวณที่เป็นส่วนเนื้อของตัวอักษร ดังนั้นหากมีการนำวิธีการนี้ไปพัฒนาต่ออาจทำตัวอักษรให้บางก่อนในขั้นการวิเคราะห์หาค่า Feature vector (Q-Code) แล้วจึงนำขอบภาพตัวอักษรไปวิเคราะห์ในขั้นวิเคราะห์หาค่าชุดรหัส A-Code หรือในขั้นตอนหาวิเคราะห์เฉพาะส่วน อาจทำให้การทำงานโปรแกรมมีประสิทธิภาพมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ขอบพระคุณ รองศาสตราจารย์ ดร.ชม กิ้มปาน ได้ให้ความอนุเคราะห์กรุณาให้คำปรึกษาและแนะนำการจัดทำวิทยานิพนธ์ฉบับนี้จนสำเร็จลุล่วงด้วยดี

ขอบพระคุณคณาจารย์ภาควิชาวิศวกรรมคอมพิวเตอร์

และขอบคุณ ผศ.ณหทัย ราศรี ที่ช่วยเหลือและให้กำลังใจโดยตลอด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

1. ผศ.ชม กิ้มปาน," ทฤษฎีการจดจำรูปแบบเบื้องต้น (Pattern Recognition Principles)",ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าวิทยาเขตเจ้าคุณทหารลาดกระบัง กรุงเทพฯ,พิมพ์ครั้งที่ 1,กรกฎาคม 2525
2. KIMPAN,C., A.ITOH and PROF.K.KAWANISHI , " Recognition of Printed Thai Characters Using a Matching Methods",IEE Procedding, vol.130,Pt.E,No.6,November 1983.
3. KIMPAN,C.,"Printed Thai Characters Recognition",Dissertation of Doctor of Engineering, King Mogkut's Institute of technology Chaokhuntaharn Ladkrabang,1986.
4. KIMPAN,C.,A.ITOH and PROF.K.KAWANISHI,"Fine Classification of Printed Thai Character Recognition Using the Karhunen-love Expansion " , IEE Procedding, vol.134, Pt.E, No.5, September 1987.
5. KIMPAN,C., " Printed Thai Characters Recognition Using Topological Properties Method",INT.J.Electronics,1986,vol.60,No.3,303-329.
6. HIRANVANICHAKORN,P.AGUI,T.,and NAKAJIMA,M." A Recognition Method of Thai Characters ". The Transaction of IECE of japan,vol.E65,no.12,December 1982.
7. HIRANVANICHAKORN,P.AGUI,T.,and NAKAJIMA,M." Recognition of Thai Characters by Using Local Features ". The Transaction of the IECE of japan,vol.E67,no.8,August 1984.
8. สุรพันธ์ เอื้อไพบูลย์, ดร.บุญวัฒน์ อัตชู และ รศ.มนัส สังวรศิลป์,"การจำแนกกลุ่มของตัวอักษรภาษาไทยโดยการพิจารณาหัวของตัวอักษร "บทความประชุมทางวิชาการ "การประยุกต์ใช้คอมพิวเตอร์ในทางวิศวกรรม",มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่ 11-12 กุมภาพันธ์ 2531
9. สุรพันธ์ เอื้อไพบูลย์,"การจดจำอักษรลายมือเขียนภาษาไทยโดยการพิจารณาหัวของตัวอักษร (Recognition of Hand-written Thai Character Considering the Head of Characters)",วิทยานิพนธ์มหาบัณฑิต คณะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้า เจ้าคุณทหาร ลาดกระบัง, ประจำปีการศึกษา 2531
10. ประสาร ตั้งติสานนท์,"การจดจำรูปแบบตัวอักษรค้ดลายมือภาษาไทย โดยวิธีแยกลักษณะเด่น " วิทยานิพนธ์มหาบัณฑิต คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้า เจ้าคุณทหารลาดกระบัง ,ประจำปีการศึกษา 2529
11. ชาย เกษมอมรกุล และ ดร.ชม กิ่มปาน ,"การจดจำอักษรค้ดลายมือภาษาไทยสำหรับเครื่องไมโครคอมพิวเตอร์โดยการพิจารณาลักษณะเด่นของอักษร " การประชุมวิชาการวิศวกรรมไฟฟ้าครั้งที่ 12 ณ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์,16-17 พฤศจิกายน 2532
12. ชาย เกษมอมรกุล," การออกแบบพจนานุกรมสำหรับการเรียนรู้อักษรค้ดลายมือไทย-อังกฤษอัตโนมัติบนเครื่องไมโครคอมพิวเตอร์ (A Design of Dictionary for automatic Handprinted Thai-English Characters Recognition on Microcomputer)",วิทยานิพนธ์มหาบัณฑิต คณะ วิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้า เจ้าคุณทหาร ลาดกระบัง, ประจำปีการศึกษา 2532
13. SOMSAK WALAIRACHT , CHAI KASEMAMORNKUL and DR.CHOM KIMPAN, "Handprinted Thai Character Recognition by Feature Concentrate Method" , COMPUTER PROCESSING OF ASIAN LANGUAGE (CPAL) , ASIAN INSTITUTE OF TECHNOLOGY ,September 26-28,1989
14. กำจร วิศลดิลก วิเชียร สุภานันท์ และ วิเชียร เสิศธนพงศ์, "การพัฒนาซอฟแวร์สำหรับเครื่องอ่านอักษรภาษาไทย", ปริญญานิพนธ์ คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้า เจ้าคุณทหาร ลาดกระบัง, ประจำปีการศึกษา 2527
15. วัชระ ฉัตรวิริยะ กิตติพัทธ์ เจียรพงษ์ และศจ.ดร.ศรีศักดิ์ จามรมาน, .การแยกแยะตัวพิมพ์ภาษาไทยออกจากภาพตัวอักษรของประโยค",การประชุมวิชาการวิศวกรรมไฟฟ้า ครั้งที่ 10 ,คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย , 24-25พฤษภาคม 2530
16. ประกาศิต ชาติบุรุษ,ธนา หงษ์สุวรรณ, วัชระ ฉัตรวิริยะ และ ดร.บุญวัฒน์ อัตชู ,"การแยกแยะตัวพิมพ์ภาษาไทยทีละตัวจากคำ",บทความการประชุมทางวิชาการ " การประยุกต์ใช้คอมพิวเตอร์ในทางวิศวกรรม " ,มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่,11-12 กุมภาพันธ์ 2531

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

17. อนุชิต จารุวนารัตน์, สุรสิทธิ์ ราตรี และรศ.ดร.ชม กิ้มปาน,"การแยกภาพตัวอักษรภาษาไทยออกจากภาพประโยค",การประชุมวิชาการวิศวกรรมไฟฟ้า ครั้งที่ 10,คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย , 24-25 พฤษภาคม 2530
18. สุรสิทธิ์ ราตรี, รศ.ดร.ชม กิ้มปาน , "การจดจำรูปแบบตัวพิมพ์อักษรภาษาไทยโดยใช้คุณสมบัติทางเรขาคณิตของตัวอักษร" ,วิศวกรรมสารลาดกระบัง,ปีที่ 8 , ฉบับที่1, มิถุนายน 2529
19. สุรสิทธิ์ ราตรี, รศ.ดร.ชม กิ้มปาน , "การจดจำรูปแบบตัวพิมพ์อักษรภาษาไทยโดยการวิเคราะห์โครงสร้างแบบมีไวยากรณ์",การประชุมวิชาการวิศวกรรมไฟฟ้า ครั้งที่10,คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย,24-25 พฤษภาคม 2530
20. ศมิทธ์ เอมสมบัติ, สุรสิทธิ์ ราตรี และรศ.ดร.ชม กิ้มปาน,"การจดจำรูปแบบอักษรตัวพิมพ์ภาษาไทยด้วยการกระจายแบบคาร์ฮูเนน-โลษ",บทความประชุมทางวิชาการและนิทรรศการ "การประยุกต์ใช้คอมพิวเตอร์ในทางวิศวกรรม", มหาวิทยาลัยสงขลา นครินทร์ วิทยาเขตหาดใหญ่ 11-12 กุมภาพันธ์ 2531
21. สมศักดิ์ วลัยรัชต์, สุรสิทธิ์ ราตรี และรศ.ดร.ชม กิ้มปาน ,"การจดจำรูปแบบอักษรตัวพิมพ์ภาษาไทย",บทความประชุมทางวิชาการและนิทรรศการ "การประยุกต์ใช้ คอมมพิวเตอร์ในทางวิศวกรรม",มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่, 11-12 กุมภาพันธ์ 2531
22. เสรี ปานชาง, สุรสิทธิ์ ราตรี, รศ.ดร.ชม กิ้มปาน "การหารหัสเบื้องต้นโดยอัตโนมัติของรูปร่างตัวพิมพ์อักษรภาษาไทย", การประชุมวิชาการวิศวกรรมไฟฟ้า ครั้งที่ 11 ณ คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีราชมงคล,2531
23. สุรสิทธิ์ ราตรี, รศ.ดร.ชม กิ้มปาน , "การจดจำรูปแบบตัวพิมพ์อักษรภาษาไทยหลายรูปแบบ(Multi Fonts Thai Printed Character Recognition)" ,การประชุมวิชาการวิศวกรรมไฟฟ้า ครั้งที่ 12 ,คณะวิศวกรรมศาสตร์ ,มหาวิทยาลัยเกษตรศาสตร์, 16-17 พฤศจิกายน 2532
24. J.T.TOU,FELLOW and R.C.GONZALEZ,"Recognition of Handwritten Characters by Topological Feature Extraction and Multilevel Categorization", IEEE Trans.computers, July 1972.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก.

ตารางรหัส ASCII CODE

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0				0	@	P	`	p	┌	ไ		ฐ	ภ	ะ	เ	o
1			!	1	A	Q	a	q	┐	ใ	ก	ท	ม	ั	แ	๑
2			"	2	B	R	b	r	└	ๅ	ข	ฒ	ย	า	โ	๒
3			#	3	C	S	c	s	┘	ะ	ข	ณ	ร	ำ	ใ	๓
4			$	4	D	T	d	t	├	ะ	ค	ด	ฤ	ิ	ไ	๔
5			%	5	E	U	e	u	─	ะ	ค	ต	ล	ี	า	๕
6			&	6	F	V	f	v	┼	ๆ	ม	ถ	ฦ	ึ	ๆ	๖
7			'	7	G	W	g	w	┤	ะ	ง	ท	ว	ื	ะ	๗
8			(	8	H	X	h	x	┴	ะ	จ	ธ	ศ	่	.	๘
9			)	9	I	Y	i	y	┬	ะ	ฉ	น	ษ	้	ะ	๙
A			*	:	J	Z	j	z	┼	ะ	ช	บ	ส	๊	ๆ	๚
B			+	;	K	[	k	{	▌	ะ	ซ	ป	ห	๋	ฯ	๛
C			,	<	L	\	l	¦		ะ	ฌ	ผ	ฬ	์	ะ	๜
D			-	=	M	]	m	}		ะ	ญ	ฝ	อ	๎	.	๝
E			.	>	N	^	n	~		ะ	ฎ	พ	ฮ	๏	ไ	๞
F			/	?	O	_	o			ไ	ฏ	ฟ	ฯ	ฺ	๏	

รูป ก-1. ASCII CODE รหัส สมอ.

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0				0	@	P	`	p		o		ฒ	ย	เ	ฺ	๟
1			!	1	A	Q	a	q		๑	ก	ณ	ร	แ	๑	๟
2			"	2	B	R	b	r		๒	ข	ค	ฤ	โ	๒	๠
3			#	3	C	S	c	s		๓	ค	ต	ล	ใ	๓	๡
4			$	4	D	T	d	t		๔	ม	ถ	ว	ไ	๔	๢
5			%	5	E	U	e	u		๕	ง	ท	ศ	ๆ	●	๣
6			&	6	F	V	f	v		๖	จ	ธ	ษ	ๆ		๤
7			'	7	G	W	g	w		๗	ฉ	น	ส	่		๥
8			(	8	H	X	h	x		๘	ช	บ	ห	๊		๦
9			)	9	I	Y	i	y	┼	๙	ซ	ป	ฬ	๋		๧
A			*	:	J	Z	j	z	┌	▌	ฌ	ผ	อ	์	๏	๨
B			+	;	K	[	k	{	┐	┼	ญ	ฝ	ฮ	๎	๞	๩
C			,	<	L	\	l	¦	└	┴	ฎ	พ	ะ	๏	๟	๪
D			-	=	M	]	m	}	┘	┬	ฏ	ฟ	๐	๐	๠	๫
E			.	>	N	^	n	~	▏	├	ฐ	ภ	า	.	๡	๬
F			/	?	O	_	o		─	┤	ฑ	ม	ำ	๏	๢	

รูป ก-2. ASCII CODE รหัส เกษตร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

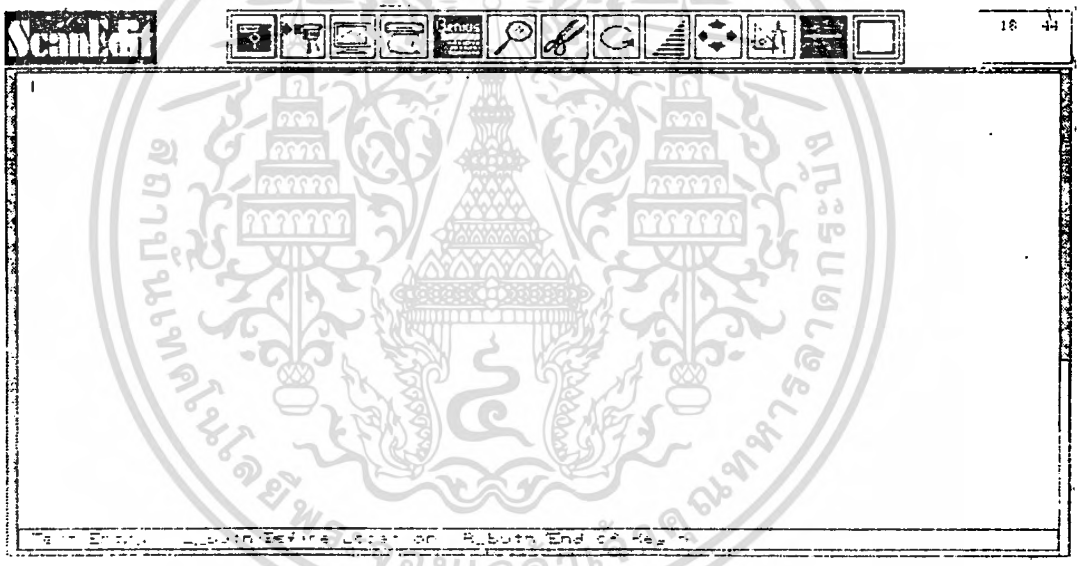
ภาคผนวก ข
การใช้งานโปรแกรมควบคุม HANDY SCANNER

โปรแกรมควบคุมการทำงานของเครื่อง Handy Scanner ประกอบด้วยโปรแกรมหลักชื่อ SCANEDIT.COM และโปรแกรมสนับสนุนอีก 2 โปรแกรมคือ MOUSE.EXE ซึ่งเป็น Drevier สำหรับ Mouse และ GSCAN.EXE สำหรับเป็น Graphic Editor แต่ในการใช้งานสำหรับการทดลองครั้งนี้ จะใช้เพียงโปรแกรมหลัก SCANEDIT.COM ร่วมกับ Driver ของ mouse โดยจัดการทำงานรวมกันในแบ็ทไฟล์ชื่อ SE.BAT ดังนั้นเมื่อเริ่มใช้งาน จึงมีขั้นตอนดังนี้

ขั้นที่ 1. เมื่อ Boot Dos และพร้อมจะเริ่มใช้งานให้พิมพ์

A>SE แล้วกด ENTER

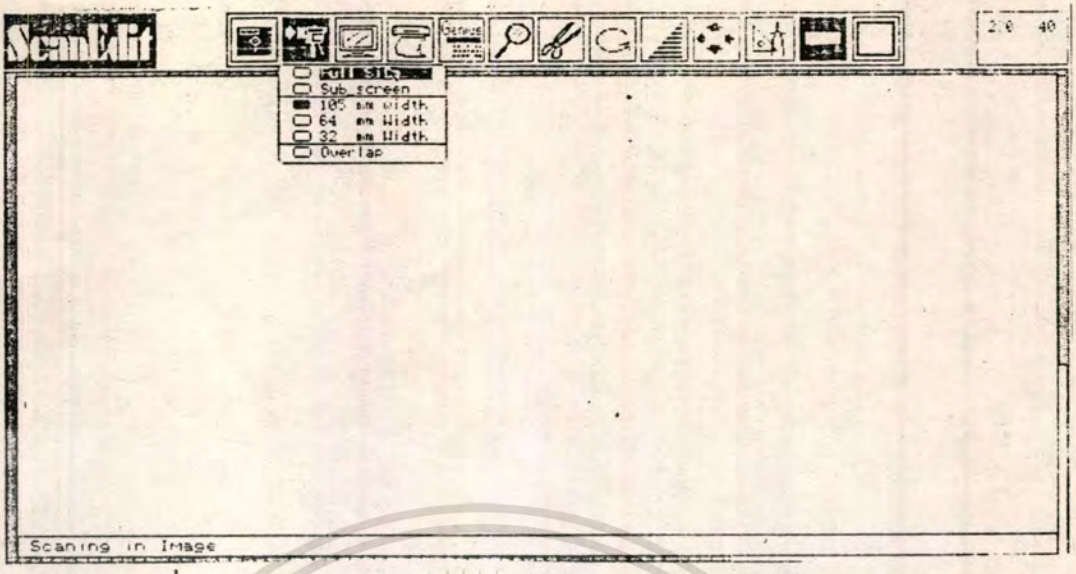
จะปรากฏภาพที่หน้าจอ ดังรูปที่ ข-1.



รูปที่ ข-1. ภาพปรากฏที่จอ พร้อมจะเริ่มใช้งาน

ขั้นที่ 2. เลื่อน mouse ไปชี้ที่เครื่องหมายที่เป็นรูปกล้องถ่าย กดปุ่ม mouse ด้านซ้าย 1 ครั้ง จะปรากฏเป็น กรอบให้เลือกขนาดความกว้างของการ Scan ภาพ อาจเลือกหัวข้อ Full Size หรือหัวข้ออื่น ตามความต้องการ ดังแสดงในรูปที่ ข-2.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ข-2. แสดงลักษณะการเลือกขนาดความกว้างของการ Scan

ขั้นที่ 3. เมื่อเลื่อน mouse ชี้เพื่อเลือกขนาดความกว้างแล้วกด ปุ่มซ้าย ของ mouse จะปรากฏแสงเรืองสีเขียวบริเวณส่วนหัวของเครื่อง Handy Scanner ให้นำเครื่อง Handy Sanner วางทาบลงบนกระดาษที่มีข้อความที่ต้องการ กดปุ่ม START ของเครื่อง Handy Scanner ลากผ่านข้อความที่ต้องการอย่างช้า ๆ จะปรากฏข้อความที่เครื่อง Handy Scanner ลากผ่าน ปรากฏที่หน้าจอ ดังรูปที่ ข-3.

รูปที่ ข-3. แสดงภาพข้อความที่เครื่อง Handy Scanner ลากผ่าน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นที่ 4. เมื่อลากเครื่อง Handy Scanner ผ่านข้อความได้ประมาณ 10-12 บรรทัด ข้อมูลภาพจะเต็มหน่วยความจำ Buffer หรือหากผู้ใช้ต้องการหยุดการอ่านภาพก่อนหน่วยความจำ Buffer เต็ม ให้กดปุ่ม mouse ด้านขวา 2 ครั้ง การบังคับ Mouse จะกลับเข้ามาส่วนหน้าจอภาพอีกครั้ง

ขั้นที่ 5. ให้เลื่อน mouse ไปชี้ที่ช่องเครื่องหมาย แผ่นดิส แล้วกดปุ่ม mouse ด้าน ซ้าย 1 ครั้ง จะปรากฏช่องกรอบ ให้เลือกชนิดของ File ข้อมูลรูปภาพที่ต้องการเก็บบันทึก ให้เลื่อน Mouse ไปชี้ที่ช่อง OCR (.OCR) format แล้วเลื่อน mouse ไปชี้ที่ช่อง Save Scan Buffer ในช่องกรอบเหนือขึ้นไป ดังแสดงในรูปที่ ข-4. แล้วกดปุ่มซ้ายของ mouse 1 ครั้ง เพื่อตั้งชื่อ File ข้อมูลที่เตรียมจะบันทึก เก็บเป็นข้อมูลภาพสำหรับนำไปวิเคราะห์เพื่อการจะจำรูปแบบตัวอักษรต่อไป

รูปที่ ข-4. แสดงการเลือกชนิดของ File เพื่อบันทึกข้อมูลภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค.
ตัวอักษรและข้อความรูปแบบต่าง ๆ

ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ
ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ ั ็ ่ ้ ๊ ๋ ์ ๆ เ แ ฤ
อ่ อ้ อ๊ อ๋ อ์ ไ ๅ ๆ ๆ

ตัวอักษรรูปแบบที่ 1
ชื่อ DENDROBIUM
ขนาด 6 point size

คอมพิวเตอร์กราฟฟิค เป็นเทคโนโลยี
ทางคอมพิวเตอร์ในแขนงใหม่ที่เพิ่งเริ่มเป็นที่
นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้น
มีการศึกษาค้นคว้าทดลองและนำความรู้ที่ได้
มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเห
ตุที่รูปภาพมีความเด่นชัด สามารถใช้สื่อความ
หมายได้ดีกว่าข้อมูลที่เป็นตัวเลขหรือตัวหนัง
สือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิค
จึงมีบทบาทสำคัญในการลดขั้นตอนและความ
ยุ่งยากต่างๆที่เกี่ยวข้องกับการนำเสนอผลงาน
แต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความ
ไตร่ตรองและความสามารถในการแปลความ
หมายจากข้อมูลที่เป็นตัวเลขจนเกินความจำ
เป็น ทั้งในแง่ของความเหน็ดเหนื่อยและเวลาที่
ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระ
มากขึ้นและมีโอกาสได้ใช้สมองกับความคิด
และการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษร
รูปแบบที่ 1 ชื่อ
DENDROBIUM ขนาด
6 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ
ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ า ิ ี ึ ื ุ ู เ แ ฤ
อ่ อ้ อ๊ อ๋ อ์ ไ ใ ๆ ฯ

ตัวอักษรรูปแบบที่ 2
ชื่อ ROYAL ORCHID
ขนาด 8 point size

คอมพิวเตอร์กราฟฟิค เป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้าทดลองและนำความรู้ที่ได้มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพมีความหมายชัดเจน สามารถใช้สื่อความหมายได้ดีกว่าข้อมูลที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอนและความยุ่งยากต่างๆที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลขจนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อยและเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้นและมีโอกาสได้ใช้สมองกับความคิด และการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษรรูปแบบที่ 2 ชื่อ ROYAL ORCHID ขนาด 8 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ฆ ง จ ฉ ช ฌ ญ
ฏ ฐ ฑ ฒ ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ า ิ ี ึ ื ุ ู เ แ ๆ
อ่ อ้ อ๊ อ๋ อ์ ไ ใ ๆ ๆ

ตัวอักษรรูปแบบที่ 2
ชื่อ CATTALEYA
ขนาด 8 point size

คอมพิวเตอร์กราฟฟิค เป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้าทดลองและนำความรู้ที่ได้มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพมีความเด่นชัด สามารถใช้สื่อความหมายได้ดีกว่าข้อมูลที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอนและความยุ่งยากต่างๆที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลขจนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อยและเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้นและมีโอกาสได้ใช้สมองกับความคิด และการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษรรูปแบบที่ 2 ชื่อ CATTALEYA ขนาด 8 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ฆ ง จ ฉ ช ฌ ญ
ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ า ิ ี ึ ื ุ ู เ แ ฤ
อ่ อ้ อ๊ อ๋ อ์ ไ ใ ๆ ฯ

ตัวอักษรรูปแบบที่ 4
ชื่อ DENDROBIUM
ขนาด 8 point size

คอมพิวเตอร์กราฟฟิค เป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้าทดลองและนำความรู้ที่ได้มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพมีความเด่นชัด สามารถใช้สื่อความหมายได้ดีกว่าข้อมูลที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอนและความยุ่งยากต่างๆที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลขจนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อยและเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้นและมีโอกาสได้ใช้สมองกับความคิด และการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษรรูปแบบที่ 4 ชื่อ DENDROBIUM ขนาด 8 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ม ง จ ฉ ช ฌ ญ
ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ า ิ ี ึ ื ุ ู เ แ ๆ
อ่ อ้ อ๊ อ๋ อ็ อ์ ไ ใ โ ฯ ๆ

ตัวอักษรรูปแบบที่ 5
ชื่อ GIGANTEA
ขนาด 8 point size

คอมพิวเตอร์กราฟฟิค เป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้าทดลองและนำความรู้ที่ได้มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพมีความเด่นชัด สามารถใช้สื่อความหมายได้ดีกว่าข้อมูลที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอนและความยุ่งยากต่างๆที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลขจนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อยและเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้นและมีโอกาสได้ใช้สมองกับความคิดและการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษรรูปแบบที่ 5 ชื่อ GIGANTEA ขนาด 8 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ฆ ง จ ฉ ช ฌ ญ
ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ า ิ ี ึ ื ุ ู เ แ ๅ
อ่ อ้ อ๊ อ๋ อ์ ไ ๆ ๆ

ตัวอักษรรูปแบบที่ 6
ชื่อ RAINY
ขนาด 9 point size

คอมพิวเตอร์กราฟฟิค เป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้าทดลองและนำความรู้ที่ได้มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพมีความเด่นชัด สามารถใช้สื่อความหมายได้ดีกว่าข้อมูลที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอนและความยุ่งยากต่างๆที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลขจนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อยและเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้นและมีโอกาสได้ใช้สมองกับความคิดและการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษรรูปแบบที่ 6 ชื่อ RAINY ขนาด 9 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ฆ ง จ ฉ ช ฌ ญ
ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ า ิ ี ึ ื ุ ู เ แ ๆ
อ่ อ้ อิ อ๋ อ็ อ์ ไ ใ โ ฯ ๆ

ตัวอักษรรูปแบบที่ 7
ชื่อ AIREDES
ขนาด 10 point size

คอมพิวเตอร์กราฟฟิคเป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้าทดลองและนำความรู้ที่ได้ มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพ มีความเด่นชัด สามารถใช้สื่อความหมายได้ดีกว่า ข้อมูลที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอนและความยุ่งยากต่าง ๆ ที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลข จนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อย และเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้น และมีโอกาสได้ใช้สมองกับความคิดและการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษรรูปแบบที่ 7 ชื่อ AIREDES ขนาด 10 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ม ง จ ฉ ช ฌ ญ
ฎ ฏ ฐ ท ฒ ณ ด ต ถ ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ ิ ี ึ ื ุ ู เ แ ฤ
อ่ อ้ อิ อ๋ อ็ อ์ ไ ใ ๆ ฯ

คอมพิวเตอร์กราฟฟิคเป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้าทดลองและนำความรู้ที่ได้ มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพ มีความเด่นชัดสามารถใช้สื่อความหมายได้ดีกว่า ข้อมูลที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอนและความยุ่งยากต่าง ๆ ที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลข จนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อย และเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้น และมีโอกาสได้ใช้สมองกับความคิดและการสร้างสรรค์ใหม่ ๆ แทน

ตัวอักษรรูปแบบที่ 8 ชื่อ CATTALEYA ขนาด 10 point size

ข้อความจากตัวอักษรรูปแบบที่ 8 ชื่อ CATTALEYA ขนาด 10 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ม ง จ ฉ ช ช ฌ ญ
ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ ิ ี ึ ื ุ ู เ แ ฤ
อ่ อ้ อ๊ อ๋ อ์ ไ ๆ ๆ ๆ

ตัวอักษรรูปแบบที่ 9
ชื่อ DENDROBIUM
ขนาด 10 point size

คอมพิวเตอร์กราฟฟิคเป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้าทดลองและนำความรู้ที่ได้ มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพ มีความเด่นชัดสามารถใช้สื่อความหมายได้ดีกว่า ข้อมูลที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอนและความยุ่งยากต่าง ๆ ที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลข จนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อย และเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้น และมีโอกาสได้ใช้สมองกับความคิดและการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษรรูปแบบที่ 9 ชื่อ DENDROBIUM ขนาด 10 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ม ง จ ฉ ช ซ ฌ ญ
ฎ ฏ ฐ ท ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ า ิ ี ึ ื ุ ู เ แ ฤ
อ่ อ้ อ๋ อ็ อ์ ไ ใ โ ๆ ๆ

ตัวอักษรรูปแบบที่ 10
ชื่อ GIGANTEA
ขนาด 10 point size

คอมพิวเตอร์กราฟฟิคเป็นเทคโนโลยี
ทางคอมพิวเตอร์ในแขนงใหม่ ที่เพิ่งเริ่มเป็น
ที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศ
นั้นมีการศึกษาค้นคว้า ทดลองและนำความรู้
ที่ได้ มาประยุกต์ใช้งานกันอย่างแพร่หลาย
ด้วยเหตุที่รูปภาพ มีความเด่นชัด สามารถใช้
สื่อความหมายได้ดีกว่า ข้อมูลที่เป็นตัวเลข
หรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิว
เตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลด
ขั้นตอน และความยุ่งยากต่าง ๆ ที่เกี่ยวข้อง
กับการนำเสนอผลงานแต่ละสาขาวิชา
มนุษย์อาจไม่ต้องใช้ความไตร่ตรอง
และความสามารถในการแปลความหมาย
จากข้อมูลที่เป็นตัวเลข จนเกินความจำเป็น
ทั้งในแง่ของความเหน็ดเหนื่อย และเวลาที่ถูก
ใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระ
มากขึ้น และมีโอกาสได้ใช้สมองกับความคิด
และการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษร
รูปแบบที่ 10 ชื่อ
GIGANTEA ขนาด
10 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ
ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ ิ ี ึ ื ุ ู เ แ ฤ
อ่ อ้ อ๊ อ๋ อ์ ไ ไ ๆ ๆ

ตัวอักษรรูปแบบที่ 11
ชื่อ GOLDEN ORCHID
ขนาด 10 point size

คอมพิวเตอร์กราฟฟิคเป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้า ทดลองและนำความรู้ที่ได้ มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพ มีความเด่นชัด สามารถใช้สื่อความหมายได้ดีกว่า ข้อมูลที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอน และความยุ่งยากต่าง ๆ ที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลข จนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อย และเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้น และมีโอกาสได้ใช้สมองกับความคิดและการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษรรูปแบบที่ 11 ชื่อ GOLDEN ORCHIDขนาด 10 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ
ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ ิ ี ึ ื ุ ู เ แ ฤ
อ่ อ้ อ๊ อ๋ อ์ ไ ใ ๆ ๆ

ตัวอักษรรูปแบบที่ 12
ชื่อ RAINY
ขนาด 10 point size

คอมพิวเตอร์กราฟฟิคเป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้า ทดลองและนำความรู้ที่ได้ มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพ มีความเด่นชัด สามารถใช้สื่อความหมายได้ดีกว่า ข้อมูลที่เป็นตัวเลข หรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอน และความยุ่งยากต่าง ๆ ที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลข จนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อย และเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้น และมีโอกาสได้ใช้สมองกับความคิดและการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษรรูปแบบที่ 12 ชื่อ RAINY ขนาด 10 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ฆ ง จ ฉ ช ฌ ญ
ฎ ฏ ฐ ฑ ฒ ฑ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ า ิ ี ึ ื ุ ู เ แ ๅ
อ่ อ้ อ๊ อ๋ อ์ ไ ใ ๆ ๆ

ตัวอักษรรูปแบบที่ 13
ชื่อ ARLUG
ขนาด 12 point size

คอมพิวเตอร์กราฟฟิคเป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้า ทดลองและนำความรู้ที่ได้ มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพ มีความเด่นชัด สามารถใช้สื่อความหมายได้ดีกว่า ข้อมูลที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอน และความยุ่งยากต่าง ๆ ที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลข จนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อย และเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้นและมีโอกาสได้ใช้สมองกับความคิด และการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษรรูปแบบที่ 13 ชื่อ ARLUG ขนาด 12 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ม ง จ ฉ ช ฌ ญ
ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ ิ ี ึ ื ุ ู เ แ ๅ
อ่ อ้ อ๋ อ๊ อ็ อ์ ไ ใ ๆ ๆ

คอมพิวเตอร์กราฟฟิคเป็นเทคโนโลยี ทางคอมพิวเตอร์ในแขนงใหม่ ที่เพิ่งเริ่ม เป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่าง ประเทศนั้นมีการศึกษาค้นคว้า ทดลองและ นำความรู้ที่ได้ มาประยุกต์ใช้งานกันอย่าง แพร่หลาย ด้วยเหตุที่รูปภาพ มีความเด่นชัด สามารถใช้สื่อความหมายได้ดีกว่า ข้อมูล ที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่าง เดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาท สำคัญในการลดขั้นตอน และความยุ่งยาก ต่าง ๆ ที่เกี่ยวข้องกับการนำเสนอผลงาน แต่ละสาขาวิชา มนุษย์อาจไม่ต้อง ใช้ความไตร่ตรองและความสามารถใน การแปลความหมายจากข้อมูลที่เป็นตัว เลข จนเกินความจำเป็น ทั้งใน แง่ของความเหน็ดเหนื่อย และเวลาที่ถูก ใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระ มากขึ้น และมีโอกาสได้ใช้สมองกับความ คิด และการสร้างสรรค์ใหม่ ๆ แทน

ตัวอักษรรูปแบบที่ 14
ชื่อ ARLUG (เอียง)
ขนาด 12 point size

ข้อความจากตัวอักษร
รูปแบบที่ 14 ชื่อ
ARLUG (อียง) ขนาด
12 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก ข ค ฆ ง จ ฉ ช ซ ฌ ญ
ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ
น บ ป ผ ฝ พ ฟ ภ ม ย ร
ล ว ศ ษ ส ห ฬ อ ฮ
ะ า ิ ี ึ ื ุ ู เ แ ๅ
อ่ อ้ อ๋ อ๊ อ็ อ์ ไ ใ โ ฯ ๆ

ต้วอักษรรูปแบบที่ 15
(ต้นแบบ)
ชื่อ ROYAL ORCHID
ขนาด 10 point size

คอมพิวเตอร์กราฟฟิคเป็นเทคโนโลยีทางคอมพิวเตอร์ในแขนงใหม่ ที่เพิ่งเริ่มเป็นที่นิยมใช้กันเมื่อไม่นานมานี้ ในต่างประเทศนั้นมีการศึกษาค้นคว้าทดลองและนำความรู้ที่ได้ มาประยุกต์ใช้งานกันอย่างแพร่หลาย ด้วยเหตุที่รูปภาพ มีความเด่นชัดสามารถใช้สื่อความหมายได้ดีกว่า ข้อมูลที่เป็นตัวเลขหรือตัวหนังสือแต่เพียงอย่างเดียว คอมพิวเตอร์กราฟฟิคจึงมีบทบาทสำคัญในการลดขั้นตอนและความยุ่งยากต่าง ๆ ที่เกี่ยวข้องกับการนำเสนอผลงานแต่ละสาขาวิชา มนุษย์อาจไม่ต้องใช้ความไตร่ตรองและความสามารถในการแปลความหมายจากข้อมูลที่เป็นตัวเลข จนเกินความจำเป็น ทั้งในแง่ของความเหน็ดเหนื่อย และเวลาที่ถูกใช้ไปอย่างฟุ่มเฟือย มนุษย์เราจึงมีอิสระมากขึ้น และมีโอกาสได้ใช้สมองกับความคิดและการสร้างสรรค์ใหม่ ๆ แทน

ข้อความจากตัวอักษร
รูปแบบที่ 15 (ต้นแบบ)
ชื่อ ROYAL ORCHID
ขนาด 10 point size

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้