การรู้จำตัวอักษรเขียนภาษาไทยที่เป็นคำแบบออฟไลน์โดยใช้หลักเกณฑ์ทางฟัซซีร่วมกับคุณลักษณะบ่งความต่าง

นายอิทธิพันธ์ เมธเศรษฐ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2543

ISBN 974–13–0870–1

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

OFF-LINE THAI WORD BASED HANDWRITTEN CHARACTER RECOGNITION USING FUZZY RULES WITH DISTINCTIVE FEATURES

Mr. Ithipan Methasate

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering in Electrical Engineering

Department of Electrical Engineering
Faculty of Engineering
Chulalongkorn University

Academic Year 2000
ISBN 974–13–0870–1

หัวข้อวิทยานิพนธ์	การรู้จำตัวอักษรเขียนภาษาไทยที่เป็นคำแบบออฟไลน์โดยใช้หลักเกณฑ์ทางฟัซซีร่วมกับคุณลักษณะบ่งความต่าง
โดย	นาย อิทธิพันธ์ เมธเศรษฐ
สาขาวิชา	วิศวกรรมไฟฟ้า
อาจารย์ที่ปรึกษา	รองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

…….. คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร. สมศักดิ์ ปัญญาแก้ว)

คณะกรรมการสอบวิทยานิพนธ์

…… ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. สุดาพร ลักษณียนาวิน)

…….. อาจารย์ที่ปรึกษา
(รองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล)

…… กรรมการ
(ดร. จุฬารัตน์ ตันประเสริฐ)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

อิทธิพันธ์ เมธเศรษฐ : การรู้จำตัวอักษรเขียนภาษาไทยที่เป็นคำแบบออฟไลน์โดยใช้หลักเกณฑ์ทางฟัซซีร่วมกับคุณลักษณะบ่งความต่าง. (OFF-LINE THAI WORD BASED HANDWRITTEN CHARACTER RECOGNITION USING FUZZY RULES WITH DISTINCTIVE FEATURES) อ. ที่ปรึกษา : รศ. ดร. สมชาย จิตะพันธ์กุล, 107 หน้า. ISBN 974–13–0870–1

วิทยานิพนธ์ฉบับนี้มีจุดมุ่งหมายเพื่อสร้างวิธีการรู้จำตัวอักษรเขียนภาษาไทยที่เป็นคำเรียกชื่อจังหวัด โดยประยุกต์ใช้ฟัซซีโลจิก ร่วมกับคุณลักษณะบ่งความต่างของอักษรภาษาไทย โดยระบบจะพิจารณาความคล้ายของตัวอักษรกับรูปแบบมาตรฐานที่อยู่ในรูปของกฎทางฟัซซี และเลือกรูปแบบที่มีความคล้ายมากที่สุด 5 รูป ไปตรวจสอบกับพจนานุกรมคำเรียกชื่อจังหวัด ในการตรวจสอบผลการรู้จำกับพจนานุกรมสามารถลดเวลาการทำงานได้โดยแบ่งกลุ่มคำออกตามความยาวของตัวอักษร

ผลการทดสอบกระทำบนเครื่องไมโครคอมพิวเตอร์ที่ใช้หน่วยประมวลผลกลางเป็นเพนเทียมความเร็ว 600 เมกะเฮิรตซ์ และมีหน่วยความจำแรม 128 เมกะไบต์ ตัวอักษรที่ใช้ทดสอบเป็นอักษรตัวเดี่ยว พยัญชนะ 44 ตัว จำนวน 100 แบบรวม 4400 ตัว สระและวรรณยุกต์ 19 ตัว จำนวน 40 แบบ รวม 5,160 ตัว ตัวอักษรระดับคำ 76 คำ จำนวน 40 แบบ 3,040 คำ รวม 23,720 ตัวอักษร อัตราการรู้จำผิดพลาดในกรณีตัวอักษรเดี่ยวร้อยละ 3.62 และอัตราการรู้จำผิดพลาดในกรณีคำภาษาไทย ร้อยละ 1.71

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา วิศวกรรมไฟฟ้า ลายมือชื่อนิสิต .....

สาขาวิชา วิศวกรรมไฟฟ้า ลายมือชื่ออาจารย์ที่ปรึกษา .....

ปีการศึกษา 2543

417 06478 21 : MAJOR ELECTRICAL ENGINEERING

KEY WORD: HANDWRITTEN / THAI WORD / FUZZY LOGIC / DISTINCTIVE FEATURE / RECOGNITION SYSTEM

ITHIPAN METHASATE : OFF-LINE THAI WORD BASED HANDWRITTEN CHARACTER RECOGNITION USING FUZZY RULES WITH DISTINCTIVE FEATURES. THESIS ADVISOR : ASSOC. PROF. SOMCHAI JITAPANKUL, Ph.D., 107 pp. ISBN 974–13–0870–1

The objective of this thesis is to propose a method of Thai word based handwritten character recognition by applying fuzzy logic with distinctive features. The system considers the similarity of the handwritten character and character patterns of the 76 province names. The system chooses 5 candidates in each character to check with the lexicon in postprocessing step. In the lexicon checking step, the length grouping is used to reduce the processing time.

The experiment was executed on a microcomputer with PentiumIII 600 MHz and 128 Mbytes RAM. The test includes 100 patterns of 44 Thai consonant characters totalling 4400 characters; 20 patterns of vowel and tonal marks totalling 760 characters, and 40 patterns of the 76 words totalling 23,720 characters. The error rate in character recognition is 3.62% and in word recognition the error rate is 1.71%.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Department Electrical Engineering Student's signature .....

Field of study Electrical Engineering Advisor's signature .....

Academic year 2000

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลืออย่างดียิ่งของรองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งท่านได้ให้คำแนะนำ ข้อคิดเห็นต่าง ๆ และสนับสนุนเครื่องมือในการทำวิจัยมาด้วยดีตลอด ผู้วิจัยจึงขอกราบขอบพระคุณมา ณ ที่นี้

ขอขอบคุณเพื่อนพี่น้องนิสิตที่อยู่ภายในห้องปฏิบัติการวิจัยกรรมวิธีสัญญาณดิจิทัล (Digital Signal Processing Research Laboratory) ที่ได้ช่วยเหลือ คำแนะนำ และกำลังใจตลอดระยะเวลาการทำวิจัยอย่างดียิ่ง

ขอขอบคุณทุกท่าน โดยเฉพาะโรงเรียนอรุณวิทยาที่ให้ความช่วยเหลือในการเก็บข้อมูลลายมือเขียนเป็นอย่างดี ซึ่งสำคัญต่อการทำวิทยานิพนธ์ฉบับนี้มาก

ท้ายนี้ผู้วิจัยขอกราบขอบพระคุณบิดามารดา ที่ให้การสนับสนุนแก่ผู้วิจัยเสมอมาจนสำเร็จการศึกษา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ฉ

สารบัญ ..... ช

สารบัญตาราง ..... ช

สารบัญภาพ ..... ช

บทที่

1 บทนำ 1

2 หลักการทฤษฎี และขั้นตอนวิธีการที่เกี่ยวข้อง 4

2.1 กระบวนการก่อนหน้า 4

2.1.1 การลดสัญญาณรบกวน 4

2.1.2 การทำขอบภาพให้เรียบ 5

2.1.3 การแก้ความเอียงของตัวอักษร 6

2.2 กระบวนการประมวลผลภาพดิจิทัล ที่ใช้ในระบบรู้จำตัวอักษร 8

2.2.1 การทำโครงร่างบาง 8

2.2.2 กระบวนการตรวจจับขอบเขตภาพ 11

2.2.3 กระบวนการเน้นขอบภาพ 12

2.2.4 การวิเคราะห์ขอบภาพ 13

2.2.5 การเปลี่ยนสายลำดับการลากผ่าน 14

2.2.6 การหาเส้นบรรทัด 15

2.2.7 การหาอัตราส่วนกรอบภาพ 16

2.3 ฟัซซีโลจิก 17

2.3.1 นิยามฟัซซีเซต 17

2.3.2 ฟังก์ชันความเป็นสมาชิก 17

2.3.2.1 ฟังก์ชันความเป็นสมาชิกแบบสามเหลี่ยม 18

2.3.2.2 ฟังก์ชันความเป็นสมาชิกแบบเอส
และฟังก์ชันความเป็นสมาชิกแบบพาย 18

2.3.3 ตัวดำเนินการฟัซซี 19

สารบัญ (ต่อ)

หน้า

2.3.3.1 ตัวเนินการยูเนียน 19

2.3.3.2 ตัวเนินการอินเตอร์เซค 20

2.3.3.3 ตัวเนินการคอมพลีเมนท์ 21

2.3.4 ตัวแปรฟัซซีเชิงภาษา 21

2.3.5 การประมาณเหตุผลด้วยฟัซซี 21

3 การรู้จำตัวอักษรภาษาไทยที่เป็นคำ 22

3.1 การเก็บข้อมูลภาพลายมือเพื่อใช้ในการวิเคราะห์และทดสอบระบบรู้จำตัวอักษร 22

3.2 กระบวนการก่อนหน้า 23

3.3 ขั้นตอนการรู้จำ 23

3.3.1 การสกัดคุณลักษณะสำคัญ และวัดค่าความเป็นสมาชิก 24

3.3.1.1 อัตราส่วนกรอบภาพ 24

3.3.1.2 การพิจารณาตำแหน่งอ้างอิงของตัวอักษร 25

3.3.1.3 การตรวจหาวงรอบ 27

3.3.1.4 การตรวจหาจุดปลาย 29

3.3.1.5 การตรวจหารอยหยัก 31

3.3.1.6 การตรวจหาหาง 32

3.3.1.7 การพิจารณาลำดับสายการลากผ่าน 33

3.3.2 การจำแนกตัวอักษร และการตัดสินใจ 33

3.3.2.1 การสร้างชุดกฎทางฟัซซี 34

3.3.2.2 การตัดสินใจ 36

3.4 กระบวนการภายหลัง 36

3.4.1 กระบวนการภายหลังที่แบ่งคำตามความยาวของคำ 37

3.4.2 กระบวนการภายหลังที่ใช้ความเหมือนกันของคำ 37

4 ขั้นตอนการทดลอง ผลการทดลอง และวิเคราะห์ผลการทดลอง 38

4.1 อุปกรณ์ที่ใช้ทำการทดสอบ 38

4.2 แหล่งที่มาของข้อมูลที่ใช้สร้างและทดสอบระบบ 38

4.3 การจัดกลุ่มข้อมูลเพื่อสร้าง และทดสอบระบบ 39

4.4 วิเคราะห์ผลกระทบเนื่องจากปากกาที่ใช้เขียนข้อมูล 39

4.5 วิเคราะห์ผลจากการฝึกฝนระบบรู้จำอักษรตัวเดี่ยว 41

สารบัญ (ต่อ)

หน้า

4.5.1 การฝึกฝนพยัญชนะ และสระ 41

4.6 ผลการทดสอบระบบรู้จำตัวอักษรตัวเดี่ยว 49

4.7 ผลการเปรียบเทียบผลการรู้จำโดยเลือกจำนวนคำตอบเป็น 1 ,3 และ 5 52

4.8 การฝึกฝนระบบรู้จำคำภาษาไทย 53

4.9 ผลการทดสอบระบบรู้จำคำภาษาไทย 57

4.10 ผลการทดสอบกระบวนการภายหลัง 61

4.10.1 กระบวนการภายหลังที่ตรวจสอบส่วนของคำย่อย 62

4.10.2 กระบวนการภายหลังที่ตรวจสอบส่วนของคำย่อย 63

5 สรุปผลการทดลองและข้อเสนอแนะ 66

รายการอ้างอิง 68

ภาคผนวก

ภาคผนวก ก 72

ภาคผนวก ข 73

ประวัติผู้เขียน 107

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

ตาราง หน้า

ตารางที่ 3.1 กลุ่มของอักษรตัวเดี่ยวแบ่งตามระดับของเส้นบรรทัด 34

ตารางที่ 4.1 ผลการทดสอบระบบเพื่อวิเคราะห์ ด้วยข้อมูลสำหรับการฝึกฝนระบบ 42

ตารางที่ 4.2 ผลการทดสอบการรู้จำสระ และวรรณยุกต์ เพื่อวิเคราะห์ ด้วยข้อมูล
สำหรับการฝึกฝนระบบ 43

ตารางที่ 4.3 ผลการทดสอบการรู้จำพยัญชนะ ด้วยข้อมูลกลุ่มฝึกฝน และทดสอบ 50

ตารางที่ 4.4 ผลการทดสอบการรู้จำสระ และวรรณยุกต์ ด้วยข้อมูลกลุ่มฝึกฝน และทดสอบ 51

ตารางที่ 4.5 ผลการทดสอบการรู้จำอักษรภาษาไทยที่เป็นคำด้วยข้อมูลชุดฝึกฝน 52

ตารางที่ 4.6 ผลการทดสอบระบบรู้จำอักษรที่เป็นคำ ด้วยชุดทดสอบ 57

ตารางที่ 4.7 กลุ่มของคำแบ่งตามจำนวนตัวอักษร 52

ตารางที่ 4.8 ส่วนของคำย่อยที่พบบ่อยในคำฐานข้อมูล 63

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

การเขียนเป็นวัฒนธรรมที่สำคัญอย่างหนึ่งของมนุษย์ ซึ่งมีการพัฒนามาหลายพันปี การเขียนช่วยให้มนุษย์สามารถสื่อสาร และบันทึกความรู้ และความคิดต่างๆ ได้ ในปัจจุบันเทคโนโลยีมีการพัฒนาไปมาก แต่การเขียนยังคงเป็นความชำนาญพื้นฐานของมนุษย์ แม้จะมีอุปกรณ์ทางอิเลคทรอนิกส์มากมายเป็นทางเลือกคนส่วนใหญ่ก็ยังใช้การจดบันทึก เป็นวิธีในการบันทึกความจำ จึงได้มีการพัฒนาและประยุกต์เทคโนโลยีทางด้านกรรมวิธีสัญญาณดิจิตอล กับลายมือเขียนของมนุษย์ เพื่อให้คอมพิวเตอร์สามารถวิเคราะห์และรู้จำตัวอักษรจากภาพลายมือผ่านทางเครื่องสแกนเนอร์ได้

งานวิจัยทางด้านรู้จำอักษรเขียน เป็นที่สนใจของนักวิจัยในช่วงศตวรรษที่ผ่านมา โดยระบบรู้จำอักษรเขียนที่เป็นคำสามารถแบ่งเป็น 2 กลุ่มได้ ดังนี้

1.ระบบรู้จำคำที่ไม่มีการตัดแยกคำออกเป็นตัวอักษร เป็นระบบรู้จำที่พิจารณาคำทั้งคำไม่มีการตัดแยกคำออกเป็นตัวอักษรย่อยๆ ในคำ วิธีนี้เหมาะกับการเขียนที่มีการติดกันของตัวอักษร (Cursive) ในการเขียนแบบนี้ลักษณะการเขียนตัวอักษรในคำมีรูปแบบเปลี่ยนแปลงไปจนไม่สามารถตัดแยกได้ เช่นในภาษาอังกฤษ ในระบบชนิดนี้ต้องมีฐานข้อมูลของคำที่จะรู้จำทั้งหมดซึ่งมีจำนวนมาก

2.ระบบรู้จำคำที่อาศัยการรู้จำอักษรภายในคำ เป็นระบบรู้จำที่รู้จำคำได้โดยอาศัยข้อมูลการรู้จำตัวอักษรภายในคำ วิธีนี้เหมาะที่จะใช้กับภาษาที่มีการติดกันของตัวอักษรไม่มาก และสามารถแยกตัวอักษรที่ติดกันได้ แล้วให้โครงสร้างของตัวอักษรเดี่ยวครบ

งานวิจัยทางด้านการรู้จำลายมือเขียนที่ผ่านมา มีการนำเสนอกรรมวิธีต่าง ๆ เพื่อใช้ในการรู้จำ และหาลักษณะสำคัญ (Feature) เช่น การสร้างรูปแบบการทาบเทียบรูปแบบที่จำลองขึ้นจากตัวอย่าง (Template matching)(Gwo-En and Wang,1994) การใช้แบบจำลองฮิดเดนมาร์คอฟ (Chen,Kundu,and Zhou, 1994) การวิเคราะห์ลายเส้น(Stroke Analysis)(Chang, and Yan, 1999) และการวิเคราะห์สายสัญลักษณ์(Symbolic description chain) (Simon, 1992) เป็นต้น

งานวิจัยทางด้านรู้จำตัวอักษรลายมือเขียนในภาษาที่ผ่านมา เป็นการรู้จำตัวพยัญชนะและสระแบบตัวเดี่ยวไม่มีการติดกันของตัวอักษรการพิจารณาหัวของตัวอักษร (Airphaiboon,and Sangworasil, 1989) ช่วยให้สามารถแยกตัวอักษรออกเป็นกลุ่มชัดเจน นอกจากนี้ยังมีผู้เสนอวิธีการเปลี่ยนลำดับสายการลากผ่าน (ประเสริฐ ฉอเรืองวิวัฒน์, 1998) และวิธีพิจารณาช่องว่างภายในตัวอักษร (Phokharatkul and Kimpan, 1998) เพื่อช่วยในการรู้จำ วิธีการรู้จำที่ได้มีการนำเสนอ

ยังมีข้อจำกัดอยู่ที่ตัวอักษรต้องมีหัวที่ชัดเจน ซึ่งยังไม่สามารถนำไปรู้จำลายมือเขียนแบบเป็นคำอย่างมีประสิทธิภาพได้

เนื่องจากปัญหาทางด้านการรู้จำลายมือเขียนเป็นปัญหาที่ไม่สามารถตัดสินใจได้ทันที (Fong and Yeung, 1998) จึงมีการนำฟัซซีโลจิกมาช่วยให้การตัดสินใจมีความยืดหยุ่นมากขึ้น งานวิจัยที่ผ่านมาได้แก่ วิธีประยุกต์ฟัซซีร่วมกับวิธีจำลองรูปแบบการลากเส้นเพื่อรู้จำตัวอักษรที่ติดกัน (Parizeau, and Plamondon, 1995), (Gader,and Keller, 1994), การพิจารณาลายเส้นย่อย (Substroke) ด้วยฟัซซี (Fong, and Yeung, 1998) และการใช้ฟัซซีร่วมกับรูปทรงพื้นฐาน (Poon, Man,and Chan, 1993) ซึ่งทุกงานวิจัยให้ผลการรู้จำที่เพิ่มขึ้นเมื่อเปรียบเทียบกับระบบเดิม ในภาษาไทยมีการนำฟัซซีโลจิกมาช่วยในการรู้จำตัวอักษรพิมพ์เช่น การรู้จำแบบรูปตัวอักษรพิมพ์ภาษาไทยโดยใช้เทคนิคแบบฟัซซีโลจิกและวิธีซินแทกติก (เดชา รัตนาธาร, 2538)

วัตถุประสงค์ 1. เพื่อพัฒนาวิธีการรู้จำตัวอักษรเขียนภาษาไทยที่เป็นคำ ด้วยการใช้กฎทางฟัซซีร่วมกับคุณลักษณะบ่งความต่าง

2. เพื่อสร้างโปรแกรมการรู้จำตัวอักษรลายมือเขียนภาษาไทยแบบเป็นคำ

ขั้นตอนการดำเนินงาน

1. ศึกษาระบบการรู้จำลายมือเขียนและลักษณะลายมือเขียนภาษาไทยแบบเป็นคำ
2. พัฒนาโปรแกรมสำหรับเก็บฐานข้อมูลลายมือเขียนภาษาไทยเป็นคำ
3. จัดเก็บตัวอย่างลายมือเขียนคำภาษาไทย
4. เขียนโปรแกรมทดสอบกรรมวิธีที่ใช้ในระบบการรู้จำลายมือเขียนแบบเป็นคำ
5. วิเคราะห์ผลที่ได้จากการทดสอบและหาวิธีแก้ไขข้อผิดพลาด
6. สรุปและรวบรวมผลงานวิจัยและปัญหาที่พบ

เป้าหมายและขอบเขตงานวิจัย

1. สร้างระบบรู้จำลายมือเขียนคำภาษาไทย
2. อัตราความถูกต้องในการรู้จำตัวอักษร แบ่งเป็น 2 กรณี
   - 2.1 กรณีรู้จำตัวอักษรและสระแบบตัวเดี่ยว อัตราความถูกต้องเฉลี่ยไม่ต่ำกว่าร้อยละ 95 โดยเก็บตัวอย่างจำนวน 40 ชุด จาก 40 คน ประกอบด้วยพยัญชนะ สระ และวรรณยุกต์รวมทั้งหมด 2,180 ตัวอักษร
   - 2.2 กรณีรู้จำแบบเป็นคำที่ผ่านการแยกตัวอักษรได้อย่างถูกต้อง อัตราความถูกต้องเฉลี่ยคิดเป็นคำไม่ต่ำกว่าร้อยละ 97 โดยเก็บตัวอย่างชื่อจังหวัดในประเทศไทย 76 จังหวัด จำนวน 40 ชุด จาก 160 คนเป็นจำนวน 3,040 คำ 23,720 ตัวอักษร

ประโยชน์ที่คาดว่าจะได้รับ

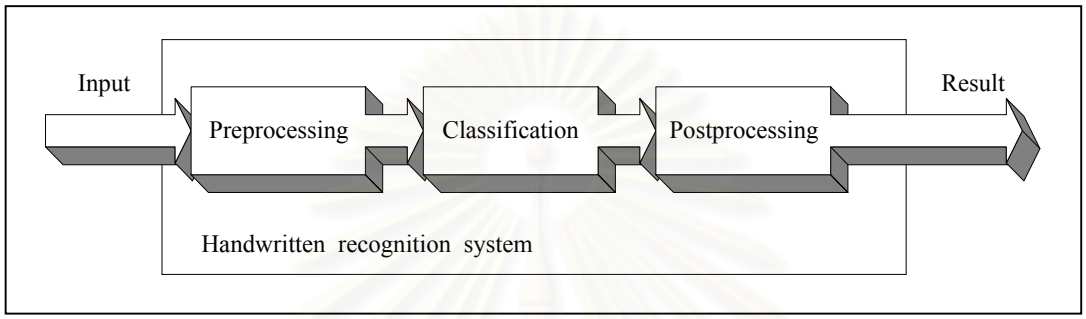
1. ทราบถึงวิธีการรู้จำคำภาษาไทยที่เป็นลายมือเขียน และลักษณะพิเศษของตัวอักษรภาษาไทย
2. ทราบถึงกระบวนการที่เหมาะสมสำหรับระบบรู้จำลายมือเขียนภาษาไทยที่เป็นคำ
3. ทำให้สามารถวางแนวทางที่เหมาะสมต่อไปในการพัฒนาระบบรู้จำลายมือเขียนภาษาไทยที่เป็นคำต่อไป

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

หลักการทฤษฎี และขั้นตอนวิธีการที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงหลักการ ทฤษฎี และขั้นตอนวิธีการ ที่นำมาประยุกต์ใช้กับกระบวนการรู้จำรูปแบบลายมือเขียน ซึ่งสามารถแบ่งเป็น 3 ขั้นตอนหลัก ได้ดังรูปที่ 2.1



รูปที่ 2.1 การรู้จำรูปแบบลายมือเขียน

2.1 กระบวนการก่อนหน้า (Preprocessing)

กระบวนการก่อนหน้าเป็นกระบวนการที่ทำเพื่อจัดการเตรียมข้อมูลภาพ (Input Image) ที่ได้จากการสแกน ให้มีความเหมาะสมสำหรับขั้นตอนต่อไป โดยมีจุดประสงค์หลักเพื่อลดสัญญาณรบกวน แก้ไขความผิดพลาดที่เกิดจากการสแกน และเน้นข้อมูลส่วนที่ต้องการ

กระบวนการก่อนหน้าแบ่งเป็นขั้นตอนต่าง ๆ ได้ดังนี้

1. การลดสัญญาณรบกวน (Noise Reduction)
2. การทำขอบภาพให้เรียบ (Edge Smoothing)
3. การแก้ความเอียง (Slant Correction)

2.1.1 การลดสัญญาณรบกวน

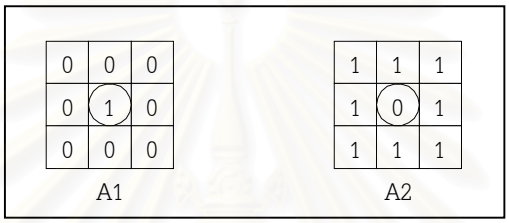
สัญญาณรบกวนที่เกิดขึ้นเนื่องจากการสแกนมีอยู่หลายชนิด(Yan and Leedham, 1997) แบ่งสัญญาณรบกวนได้เป็น 3 ชนิด คือ

1. สัญญาณรบกวนจากพื้นหลัง (Background Noise) เป็นสัญญาณรบกวนที่เกิดจากลายพื้นหลังของเอกสาร
2. สัญญาณรบกวนแบบเงา (Shadow Noise) เป็นสัญญาณรบกวนที่เกิดบริเวณขอบข้อมูลกับพื้นหลังของภาพ

3. สัญญาณรบกวนแบบซอล์ท แอน เปปเปอร์ (Salt and Pepper Noise) เป็นสัญญาณรบกวนแบบจุดขนาดเล็ก ที่เกิดขึ้นแบบสุ่มบนภาพทั่วทั้งภาพ

ในงานวิจัยนี้ใช้การเก็บภาพด้วยการสแกนจากแบบการเขียนที่เตรียมไว้ ซึ่งมีพื้นหลังเป็นสีขาว จึงไม่มีสัญญาณรบกวนชนิดพื้นหลัง สัญญาณรบกวนแบบเงาสามารถแก้ไขได้โดยการตั้งค่าเริ่มเปลี่ยนแปลง(Threshold)ให้เหมาะสมในขั้นตอนการประมวลผลภาพให้เป็นภาพไบนารี(Binary Image)

การลดสัญญาณรบกวนแบบซอล์ท แอน เปปเปอร์ ทำได้โดยการใช้หน้ากาก(Mask) ขนาด 3x3 ในรูปที่ 2.2 ตรวจสอบจุดภาพไปทุก ๆ จุดในภาพ

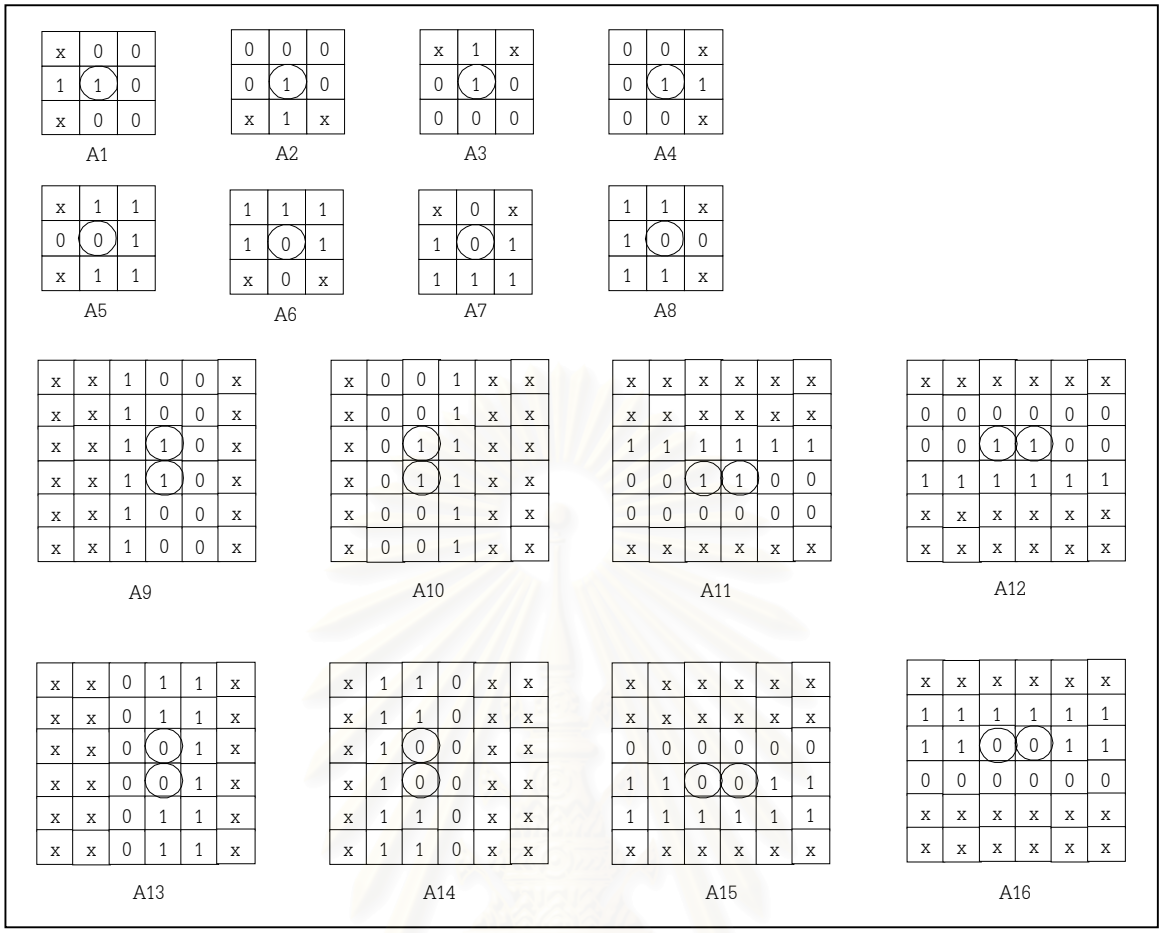


รูปที่ 2.2 หน้ากากสำหรับการลดสัญญาณรบกวนแบบซอล์ท แอน เปปเปอร์

ให้ 1 แทนจุดภาพข้อมูล และ 0 แทนพื้นหลังของภาพ ให้จุดภาพมีรูปแบบเดียวกับ หน้ากาก A1 จะเป็นการลบจุดภาพข้อมูลออกจากพื้นหลัง และในหน้ากาก A2 จะเป็นการเพิ่มจุดภาพเข้าไปถ้าตำแหน่งรอบข้างเป็นจุดภาพข้อมูล

2.1.2 การทำขอบภาพให้เรียบ

ข้อมูลภาพที่ได้จากการสแกนบางครั้งมีขอบที่ไม่เรียบ เนื่องจากสัญญาณรบกวนระหว่างขั้นตอนการสแกน หรือผลอันเนื่องมาจากกระดาษและปากกาที่ใช้เขียนในขั้นตอนการเขียนเก็บข้อมูล ดังนั้นจึงต้องมีการแก้ขอบภาพให้มีความเรียบเพียงพอต่อการนำไปพิจารณาหาคุณลักษณะสำคัญการทำขอบภาพให้เรียบสามารถทำได้ ในงานวิจัยที่ผ่านมามีการเสนอเทคนิค เพื่อทำขอบภาพของตัวอักษรให้เรียบโดย Phokharatkul และ Kimpan (2000) ได้เสนอวิธีการแก้ขอบภาพให้เรียบโดยการพิจารณาสายโซ่รหัส (Chain Code) เพื่อแก้ไขความเรียบของขอบแบบที่เป็น 2 จุดภาพติดกัน ในงานวิทยานิพนธ์นี้ไม่มีการใช้สายโซ่รหัส จึงได้ปรับมาเป็นใช้หน้ากาก 16 แบบ ดังแสดงในรูปที่ 2.3 ตรวจไปตามจุดภาพเพื่อปรับขอบภาพให้เรียบ ซึ่งให้ผลการทำงานเหมือนวิธีการสายโซ่รหัสทุกประการ



รูปที่ 2.3 หน้ากากสำหรับการทำขอบภาพให้เรียบ

ด้วยวิธีการใช้หน้ากากทั้ง 16 แบบทำให้สามารถลดสัญญาณรบกวนที่ทำให้ขอบภาพไม่เรียบจำนวน 1 จุดภาพและ 2 จุดภาพได้ซึ่งวิธีนี้มีความซับซ้อนน้อย และใช้เวลาในการประมวลผลต่ำซึ่งเหมาะสมในการใช้กับการลดสัญญาณรบกวนภาพตัวอักษรที่มีสัญญาณรบกวนขนาดเล็ก

2.1.3 การแก้ความเอียงของตัวอักษร

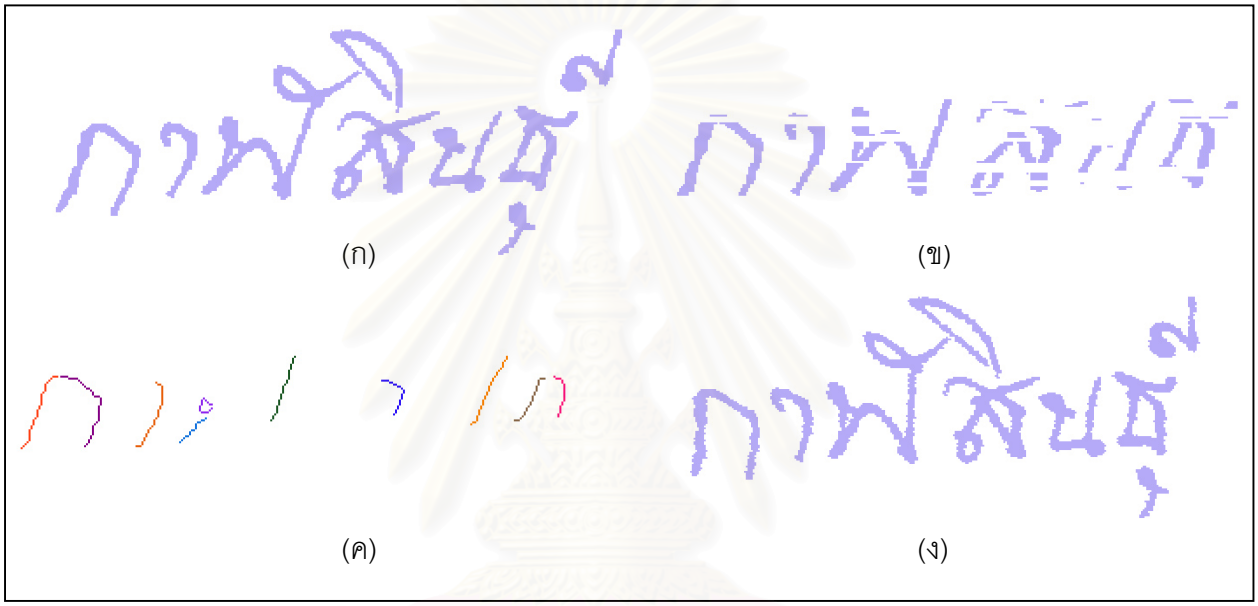
การแก้ความเอียงของตัวอักษรคือการบิดภาพตัวอักษรที่เอียงไปในทิศทางเดียวกันทั้งคำให้ตั้งตรงซึ่งจะทำได้เมื่อตัวอักษรทุกตัวในคำเอียงเหมือนกัน วิธีการแก้ปัญหาการเอียงของตัวอักษร (Bozinovic and Srihari, 1989) ซึ่งได้ทดลองใช้กับลายมือเขียนในภาษาอังกฤษที่มักมีการเขียนตัวหนังสือเอียง ดังแสดงในรูปที่ 2.4 ก่อนที่จะแก้ความเอียงต้องมีการหาค่ามุมความเอียงของคำก่อน แล้วจึงทำการหมุนตัวอักษรแต่ละตัวชดเชย การหามุมเอียงและแก้ความเอียงทำได้โดย

1. ลบข้อมูลในแถวแนวนอน ที่มีความหนาแน่นของจุดภาพมากกว่าค่าเริ่มเปลี่ยนแปลง (Threshold) ออกทำให้เหลือจุดภาพของเส้นในแนวตั้งเท่านั้น
2. ลบบริเวณภาพที่มีขนาดเล็กออก

3. ทำการหาโครงภาพตามวิธีในหัวข้อ 2.2.1 จากชิ้นภาพที่เหลืออยู่
4. ตีกรอบรอบโครงภาพแต่ละชิ้นภาพ แล้วหาค่าความชัน (Slope) ของเส้นทะแยงมุมของแต่ละกรอบรอบโครงภาพ
5. นำเส้นทะแยงมุมที่ได้ มาหาค่าเฉลี่ยจะได้ มุมเอียงของตัวอักษร θ
6. หมุนภาพอักษรตามสมการ

$$x' = x - y \cdot \tan\theta \qquad (2.1)$$

$$y' = y \qquad (2.2)$$



(ก) (ข)

(ค) (ง)

รูปที่ 2.4 วิธีการแก้ความเอียงของตัวอักษรตามลำดับ
(ก) คำลายมือเขียนที่มีการเอียง
(ข) ลบข้อมูลในแถวแนวนอน
(ค) หาโครงร่างบาง และลบส่วนที่มีขนาดเล็กออก
(ง) หมุนภาพด้วยมุมที่ประมาณได้

โดย (x', y') เป็นจุดภาพใหม่หลังจากแก้ความเอียง และ (x, y) เป็นจุดภาพเดิม และ θ เป็นมุมที่ตัวอักษรเอียง

2.2 กระบวนการประมวลผลภาพดิจิทัล ที่ใช้ในระบบรู้จำตัวอักษร

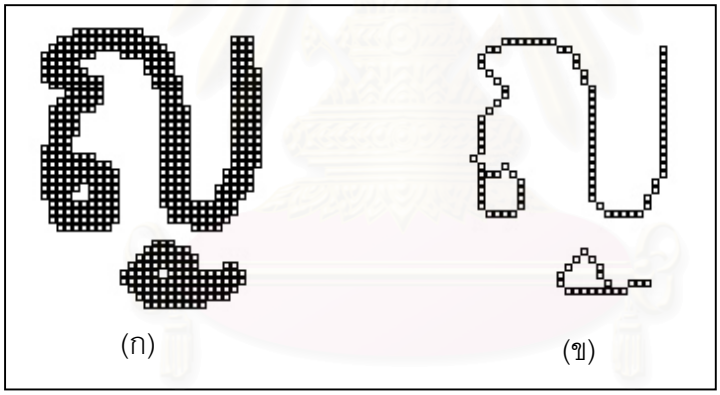
กระบวนการประมวลผลภาพดิจิทัล ถูกนำมาประยุกต์เพื่อใช้ในการหาคุณลักษณะสำคัญ และบ่งบอกความแตกต่างของตัวอักษร กระบวนการประมวลผลภาพที่นำมาใช้ในกระบวนการรู้จำ

มีหลายขั้นตอนโดยมักจะขึ้นอยู่กับคุณลักษณะสำคัญ และวิธีการรู้จำที่เลือกใช้ ในงานวิจัยนี้ใช้คุณลักษณะสำคัญที่เป็นโครงสร้างของตัวอักษร ดังนั้นกระบวนการประมวลผลภาพจึงใช้เพื่อการค้นหาลักษณะของตัวอักษรที่มีอยู่ การประมวลผลภาพดิจิทัลที่นำมาใช้ในระบบรู้จำมี 5 วิธีดังนี้

1. การทำโครงร่างบาง (Thinning)
2. การตรวจจับขอบเขตภาพ (Boundary Detection)
3. การเน้นขอบภาพ (Edge Enhancement)
4. การวิเคราะห์ขอบภาพ (Profile Analysis)
5. การเปลี่ยนสายลำดับการลากผ่าน (Stroke Changing Sequence)
6. การหาเส้นบรรทัด (Sentence Line Detection)
7. การหาอัตราส่วนของโครงร่าง (Aspect Ratio)

2.2.1 การทำโครงร่างบาง

การทำโครงร่างบางเป็นกระบวนการที่สำคัญเพื่อตรวจสอบคุณลักษณะสำคัญ(Feature) ของตัวอักษร ซึ่งกระบวนการนี้เป็นการทำให้ภาพที่รับเข้ามามีความหนาของภาพเพียงหนึ่งจุดภาพดังแสดงในรูปที่ 2.5



รูปที่ 2.5 กระบวนการทำโครงร่างบาง (ก) ภาพต้นแบบ (ข) ภาพที่ผ่านการทำโครงร่างบาง

ภาพที่ผ่านกระบวนการทำโครงร่างบางจะได้คุณลักษณะของจุดต่อในภาพโครงร่างซึ่งประกอบด้วย

- จุดปลาย คือจุดที่มีผลรวมของจุดรอบข้างเท่ากับหนึ่ง
- จุดต่อเนื่อง คือจุดที่มีผลรวมของจุดรอบข้างเท่ากับสอง
- จุดแยกสาม คือจุดที่มีผลรวมของจุดรอบข้างเท่ากับสาม
- จุดแยกสี่ คือจุดที่มีผลรวมของจุดรอบข้างเท่ากับสี่

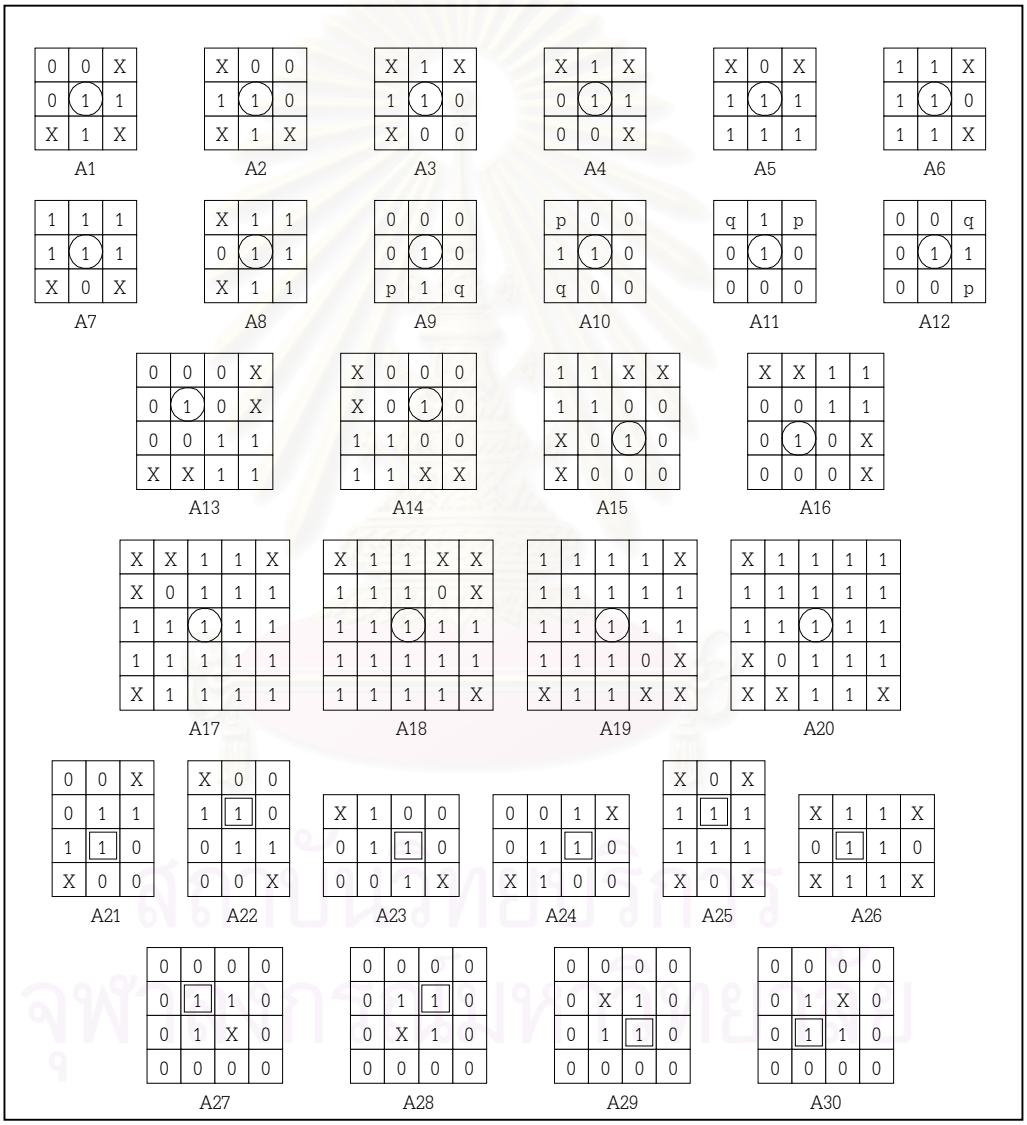
เนื่องจากการทำโครงร่างบางมีอยู่ด้วยกันหลายวิธีได้เลือกใช้การทำโครงร่างโดยกรรมวิธี One-Pass Parallel Thinning (Jang and Chin,1992) เพราะให้รายละเอียดคุณลักษณะของจุดต่อภาพได้ดีเพียงพอ

กรรมวิธี One-Pass Parallel Thinning มีหน้ากาก (Mask) ที่ใช้ในการทำโครงร่างบางทั้งหมด 30 รูปแบบ ซึ่งหน้ากาก A1 ถึง A20 เป็นหน้ากากการทำโครงร่างบาง ในขณะที่ A21 ถึง A30 เป็นหน้ากากการเรียก

กลับคืน (Restoring) สมมุติให้ p และ q เป็นตัวดำเนินการทางตรรกศาสตร์ : p or q = 1 และ X เป็นค่าที่ไม่สนใจ ดังแสดงในรูปที่ 2.6

ขั้นตอนการทำโครงร่างมีดังนี้

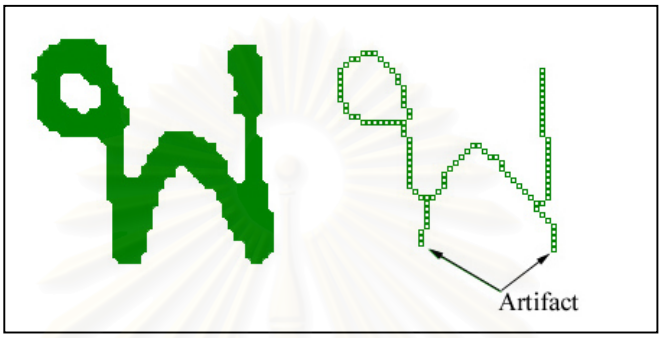
1. จุดภาพจะถูกทำการตรวจสอบตามหน้ากาก(Mask) สำหรับการทำโครงร่างบางทั้งหมด
2. หากจุดภาพที่ถูกตรวจอยู่ในหน้ากาก A1 ถึง A20 จะถูกลบจุดภาพ ในขณะที่จุดภาพอยู่ในอยู่แบบ A21 ถึง A30 จะถูกกู้กลับคืนมา
3. เมื่อภาพผ่านกระบวนการทำโครงร่างบางจะเหลือความกว้างของภาพเพียง 1 จุดภาพ



รูปที่ 2.6 หน้ากาก(mask) สำหรับการทำโครงร่างบาง

ในการทำโครงร่างบางจะมีความไม่สมบูรณ์เนื่องจากการวกกลับของปากกาเวลาเขียนหรือการเปลี่ยนทิศทางของปากกาอย่างรวดเร็ว ทำให้เกิดเส้นหักมุมที่มีความหนาเมื่อทำโครงร่างบาง จะเกิดจุดปลายเทียม (Artifact) ได้ โดยจุดปลายเทียมจะอยู่ใกล้กับจุดเชื่อมต่อ(Junction Point) เสมอ ดังรูปที่ 2.7 ปัญหานี้มักเกิดขึ้นในการทำโครงร่างบางกับตัวอักษรในภาษาต่าง ๆ แต่

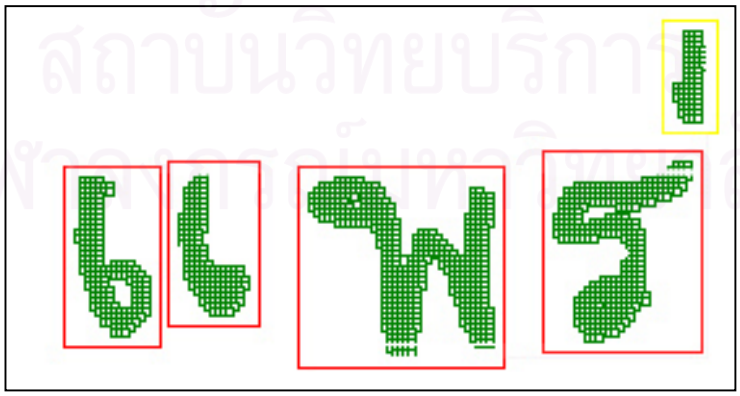
ระบบสามารถตรวจสอบจุดปลายที่เกิดขึ้นนี้ได้โดย จุดปลายที่เกิดขึ้นจะอยู่ในตำแหน่งที่ใกล้เคียงกับจุดแยก ซึ่งสามารถตรวจสอบได้



รูปที่ 2.7 จุดปลายเทียมที่เกิดขึ้นเมื่อจุดปลายเกิดขึ้นใกล้กับจุดเชื่อมต่อ

2.2.2 **กระบวนการตรวจจับขอบเขตภาพ** (Boundary Detection)

การตรวจจับขอบเขตภาพจะตรวจจับขอบเขตกรอบของภาพตัวอักษรไทยโดยจะแยกชิ้นภาพที่ไม่ติดออกจากกัน แล้วจึงนำแต่ละส่วนที่แยกออกจากกันได้ส่งเข้าสู่ระบบรู้จำทีละส่วน ซึ่งในงานวิจัยนี้ถือว่าแต่ละส่วนที่แยกได้จะเป็นตัวอักษรหนึ่งตัวอักษรที่ระบบสามารถรู้จำได้ ภาพตัวอักษรไทยที่ผ่านกระบวนการนี้จะให้ขอบเขตของภาพประกอบด้วยองค์ประกอบของขอบด้านซ้าย ขอบด้านขวา ขอบด้านบน และขอบด้านล่าง แสดงในรูปที่ 2.8 ซึ่งผลที่ได้จากการจับขอบเขตภาพจะทำให้การดำเนินการในกระบวนการถัดไปกระทำเพียงเฉพาะส่วนที่เป็นเนื้อของภาพตัวอักษรเท่านั้นเป็นการลดเวลาในการทำงาน



รูปที่ 2.8 กระบวนการตรวจจับขอบภาพตัวอักษรในคำ

วิธีการตรวจจับขอบภาพ

1. หาจุดภาพข้อมูลให้พบ กำหนดให้เป็นจุดเริ่ม

2. ทำการกำหนดค่าที่จุดเริ่มนั้นเป็นจุดพิเศษ และเปลี่ยนจุดข้อมูลที่ติดกับจุดพิเศษนั้นให้เป็นจุดพิเศษด้วย ทำซ้ำจนไม่พบจุดข้อมูลอยู่ติดกับจุดพิเศษอีก
3. ทำการเก็บค่าจุดบนสุด จุดขวาสุด จุดซ้ายสุด และจุดล่างสุด ของกลุ่มจุดที่เป็นจุดพิเศษนั้นทำให้ได้กรอบของวัตถุ
4. เริ่มหาจุดข้อมูลเริ่มต้นใหม่ แล้วทำขั้นตอนในข้อ 2 และ 3 ซ้ำ จนไม่พบจุดข้อมูลในภาพอีก
5. ผลลัพธ์ที่ได้ จะได้กรอบของภาพอักษรซึ่งประกอบด้วย ขอบด้านซ้าย ขอบด้านขวา ขอบด้านบน และขอบด้านล่างของทุกวัตถุในภาพ

ส่วนในกรณีที่ตัวอักษรเกิดการติดกันจะต้องมีการตัดแยก ก่อนส่งเข้าระบบรู้จำเพื่อรู้จำตัวอักษรทีละตัว

2.2.3 กระบวนการเน้นขอบภาพ (Edge Detection)

การเน้นขอบภาพเป็นกระบวนการปรับปรุงคุณภาพของภาพโดยเน้นส่วนที่เป็นขอบภาพ ส่วนบริเวณที่ไม่ใช่ขอบของภาพจะถูกกำจัด ซึ่งผลจากการเน้นขอบภาพจะได้ขอบภาพประกอบด้วยขอบภาพด้านในและขอบภาพด้านนอก ดังแสดงในรูปที่ 2.9 ภาพที่ได้ผ่านกระบวนการเน้นขอบภาพจะถูกส่งต่อไปยังกระบวนการลบขอบภาพด้านนอก ซึ่งจะกล่าวในหัวข้อ 2.2.4 เพื่อให้ได้องค์ประกอบในการพิจารณาการดึงคุณลักษณะ

Lee (1996) ได้นำกระบวนการเน้นขอบภาพมาใช้ในการดึงคุณลักษณะ โดยได้ทำการศึกษาและสรุปไว้ว่า ในกรรมวิธีการเน้นขอบภาพ (Edge Detection) ทั้งหมดอันได้แก่ Frei Chen edge detection, Kirsch edge detection, Prewitt edge detection, Sobel edge detection, Hough Transform และ อื่นๆ Kirsch edge detection เป็นกระบวนการที่รู้จักกันดีว่าจะทำการเน้นขอบภาพทั้งสี่ทิศทางอันได้แก่ แนวดิ่ง แนวราบ แนวทะแยงซ้าย และแนวทะแยงขวา ซึ่งมีความถูกต้องมากกว่าการเน้นขอบภาพโดยกรรมวิธีอื่นๆ โดย Kirsch compass operator มีตัวดำเนินการ 8 ตัวดังนี้

$$W_0=\begin{bmatrix}5&5&5\\-3&0&-3\\-3&-3&-3\end{bmatrix}\quad W_1=\begin{bmatrix}5&5&-3\\5&0&-3\\-3&-3&-3\end{bmatrix}\quad W_2=\begin{bmatrix}5&-3&-3\\5&0&-3\\5&-3&-3\end{bmatrix}\quad W_3=\begin{bmatrix}-3&-3&-3\\5&0&-3\\5&5&-3\end{bmatrix}$$

$$W_4=\begin{bmatrix}-3&-3&-3\\-3&0&-3\\5&5&5\end{bmatrix}\quad W_5=\begin{bmatrix}-3&-3&-3\\-3&0&5\\-3&5&5\end{bmatrix}\quad W_6=\begin{bmatrix}-3&-3&5\\-3&0&5\\-3&-3&5\end{bmatrix}\quad W_7=\begin{bmatrix}-3&5&5\\-3&0&5\\-3&-3&-3\end{bmatrix}$$

สำหรับตัวดำเนินการ compass ขอบของภาพสามารถคำนวณได้จากสมการ

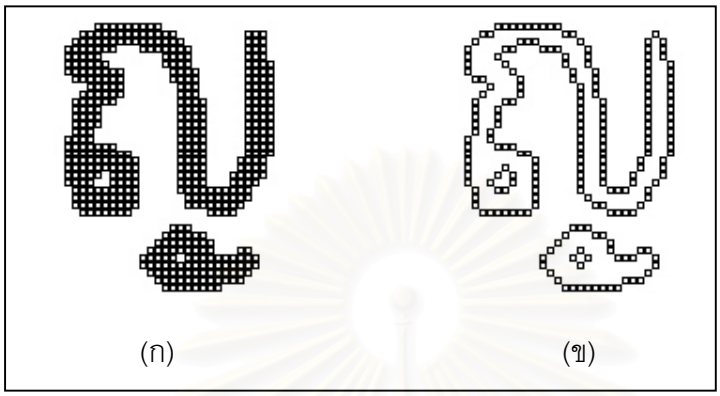
$$y(i,j)=\max\{|y_0(i,j)|,|y_1(i,j)|,...,|y_7(i,j)|\} \qquad (2.3)$$

ขณะที่ y_0, y_1 และอื่นๆ เป็น cross correlation ของ masks W_0, W_1 และอื่น ๆ กับภาพ

วิธีการเน้นขอบภาพ

1. ทำการเน้นขอบภาพโดยเลือกใช้ Operator Kirsch ซึ่งประกอบด้วยหน้าต่าง (Mask) 8 รูปแบบด้วยกัน

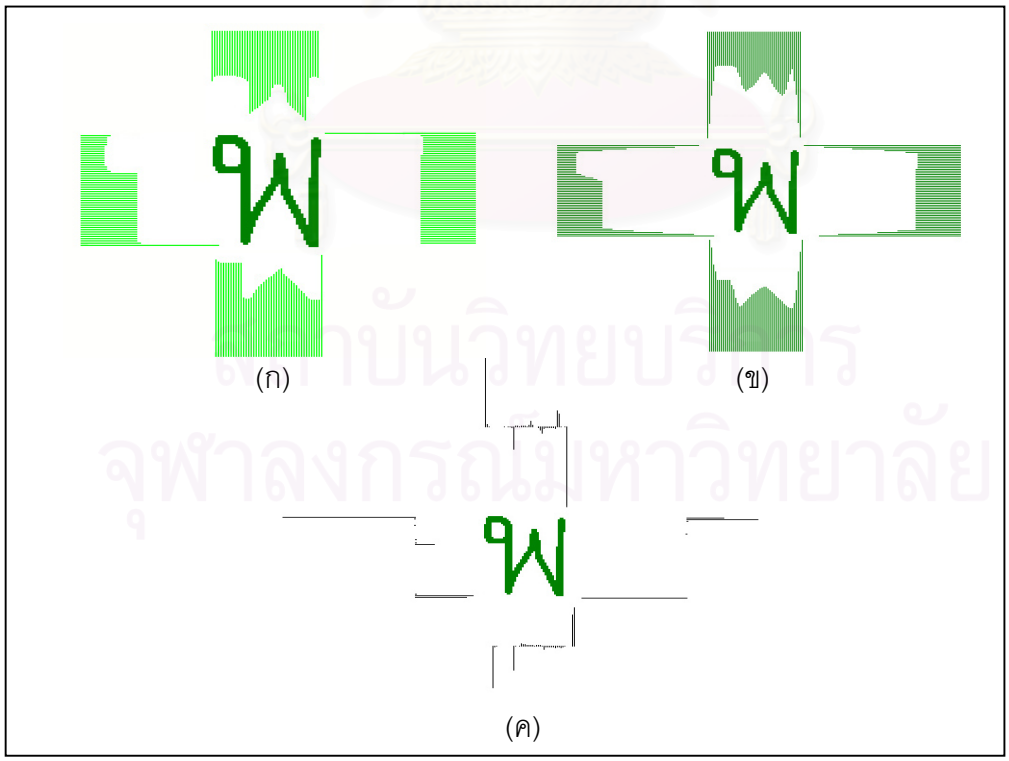
2. จุดภาพจะถูกคำนวณโดยผ่านหน้าต่างทั้ง 8 รูปแบบ ซึ่งกำหนดค่าระดับ (Threshold) ในการตัดสิน
3. พิจารณาหน้ากากผลลัพธ์ที่ให้ค่ามากที่สุดตามสมการที่ 2.3 จะได้จุดที่เป็นขอบภาพ



รูปที่ 2.9 กระบวนการเน้นขอบภาพตัวอักษร (ก) ภาพต้นแบบ (ข) ภาพที่ผ่านการเน้นขอบภาพ

2.2.4 **การวิเคราะห์ขอบภาพ** (Profile Analysis)

การวิเคราะห์ขอบภาพ คือการพิจารณาระยะห่างระหว่างขอบภาพ กับกรอบภาพอ้างอิงทั้ง 4 ทิศทาง ผลของการวิเคราะห์นี้แสดงถึงการเปลี่ยนแปลงของเส้นขอบภาพในแต่ละด้าน ซึ่งจะนำไปสู่การหารอยหยัก (Curl) ที่เกิดขึ้นภาพในตัวอักษรได้ ดังแสดงในรูปที่ 2.10



รูปที่ 2.10 กระบวนการวิเคราะห์ขอบภาพ

(ก) ระยะห่างระหว่างขอบภาพกับกรอบภาพอ้างอิง

(ข) ระยะห่างเฉลี่ยระหว่างขอบภาพกับกรอบภาพอ้างอิง

(ค) อัตราการเปลี่ยนแปลงของระยะห่างระหว่างขอบภาพกับกรอบภาพอ้างอิง

ในงานวิจัยของ วิชา พานิช (2536) ได้ใช้วิธีการหาค่าเฉลี่ยของการฉายเงาของจุดภาพในแนวแกนตั้ง เพื่อตัดแยกบรรทัดของตัวอักษรพิมพ์ภาษาไทยแล้วให้ผลที่ถูกต้อง ในงานวิจัยนี้จึงนำวิธีหาค่าเฉลี่ยนี้มาใช้ร่วมในการวิเคราะห์ขอบภาพ โดยระบบจะพิจารณาระยะทางเฉลี่ยระหว่างขอบภาพกับกรอบภาพอ้างอิงซึ่งเป็นการเฉลี่ยค่าระยะห่างของขอบภาพกับระยะห่างที่ตำแหน่งข้างเคียงโดยใช้สมการ

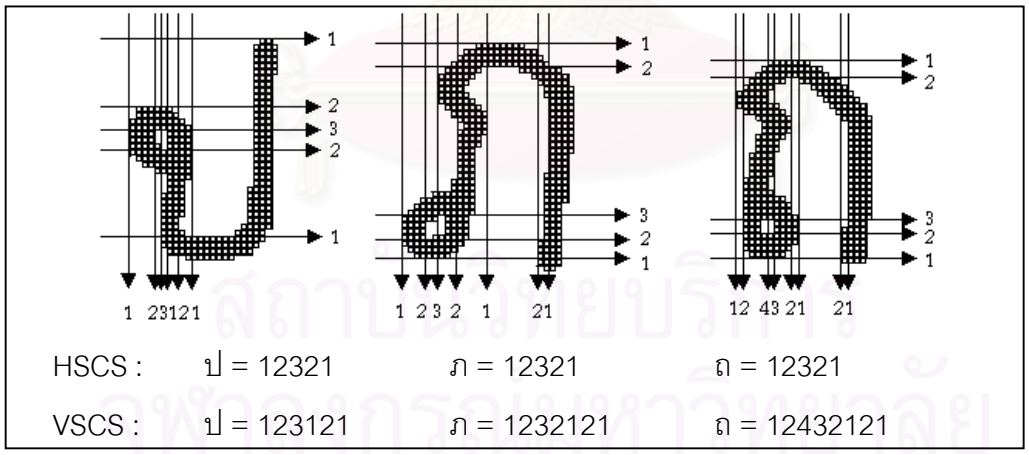
$$x_{ave} = \left(x_{i-2} + x_{i-1} + 2 \cdot x_i + x_{i+1} + x_{i+2}\right) \Big/ 6 \qquad (2.4)$$

$$y_{ave} = \left(y_{i-2} + y_{i-1} + 2 \cdot y_i + y_{i+1} + y_{i+2}\right) \Big/ 6 \qquad (2.5)$$

2.2.5 การเปลี่ยนสายลำดับการลากผ่าน (Stroke Changing Sequence : SCS)

Wang and Wang (1993;1994) ได้เสนอการเปลี่ยนสายลำดับการลากผ่าน (Stroke Changing Sequence : SCS) โดยใช้กับตัวเลขอารบิก ซึ่งพิจารณาการเปลี่ยนลำดับการลากผ่านแนวราบ (Horizontal Stroke Changing Sequence : HSCS) โดยการรวมการพิจารณาการเปลี่ยนสายลำดับการลากผ่านแนวดิ่ง (Vertical Stroke Changing Sequence : VSCS) มาพิจารณาร่วมด้วย

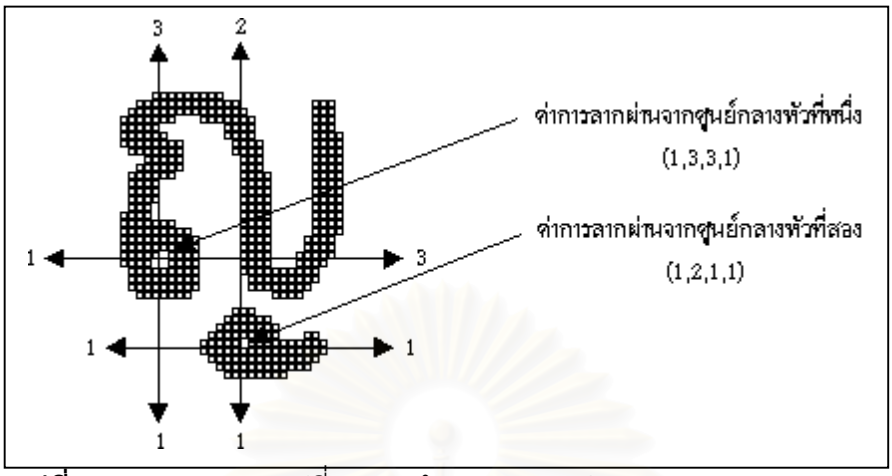
ตัวอย่างในรูปที่ 2.11 พบว่าการพิจารณาเฉพาะ HSCS จะได้ค่า HSCS ที่เหมือนกันในขณะที่เป็นตัวอักษรที่ต่างกัน แต่เมื่อพิจารณา VSCS ร่วมด้วยจะพบว่าตัวอักษรทั้ง 3 ตัวสามารถจะแยกออกจากกันได้



รูปที่ 2.11 การเปลี่ยนสายลำดับการลากผ่านแนวราบและแนวดิ่ง

ประเสริฐ ฉอเรืองวิวัฒน์(2541) ได้เสนอวิธีการประยุกต์ SCS เพื่อใช้ช่วยในการรู้จำตัวอักษรไทย ด้วยเหตุที่ว่าตัวอักษรพยัญชนะไทยนั้นเป็นตัวอักษรที่มีหัว ดังนั้นจึงได้ดัดแปลงการเปลี่ยนสายลำดับการลากผ่านโดยใช้ตำแหน่งศูนย์กลางของหัวตัวอักษรเป็นจุดเริ่มต้นในการหา SCS โดยทำการลากผ่านในแนวราบ ได้แก่ ซ้ายและขวาจากจุดศูนย์กลาง และลากผ่านในแนวดิ่งได้แก่ บนและล่างจากจุดศูนย์กลาง จะทำให้ได้ค่า SCS ของแต่ละหัวตัวอักษรซึ่งประกอบด้วย ซ้าย บน ขวา และ ล่าง ตามลำดับ(ซ้าย, บน, ขวา, ล่าง) ดังแสดงในรูปที่ 2.12 ซึ่งในกรณีนี้จะลดการคำนวณลงได้อย่างมากกล่าวคือ หัวตัวอักษร 1 หัวจะคำนวณค่า SCS เพียง 4 ค่าเท่า

นั้น อีกทั้งยังช่วยลดปัญหาการเกิดค่า SCS มากกว่าปกติเนื่องจากไม่ได้หาค่าลำดับการลากผ่านทุกเส้นมาต่อกัน

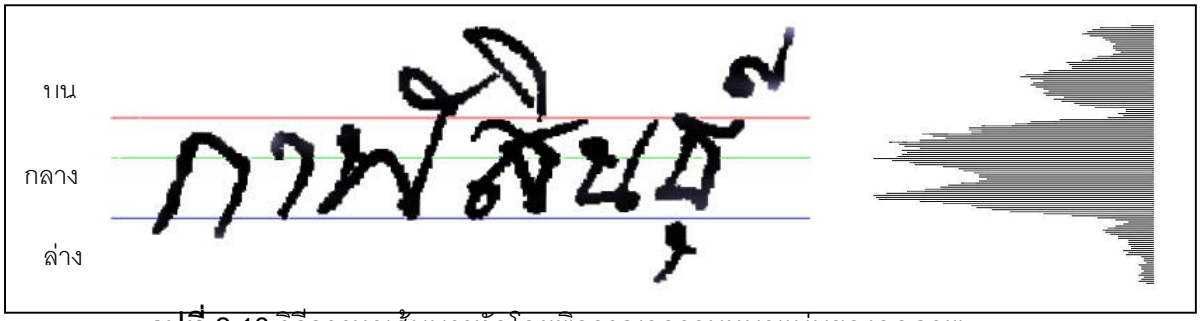


รูปที่ 2.12 กระบวนการเปลี่ยนสายลำดับการลากผ่านจากหัวของตัวอักษร

โดยใช้ตำแหน่งจุดศูนย์กลางของหัวตัวอักษรเป็นจุดเริ่มต้นในการหาค่า SCS ได้แก่ค่า SCS ในแนวราบ ซึ่งประกอบด้วยทางด้านซ้ายและทางด้านขวา และค่า SCS ในแนวดิ่ง ซึ่งประกอบด้วยทางด้านบน และทางด้านล่าง โดยเรียงลำดับค่าที่ได้จาก SCS ในสี่ทิศทางดังนี้ (ซ้าย, บน, ขวา, ล่าง) ตามลำดับ

2.2.6 **การหาเส้นบรรทัด** (Sentence Line Detection)

ในงานวิจัยเกี่ยวกับการหาเส้นบรรทัดมีผู้เสนอกันหลายวิธีในภาษาไทย มีการเสนอวิธีการหาเส้นบรรทัดของตัวพิมพ์โดย วิชา พานิช (2536) ซึ่งหาเส้นบรรทัดของประโยคซึ่งมีความแตกต่างของจำนวนจุดภาพมาก ในงานวิจัยด้านลายมือเขียน Bin, Hui และ Shaowei (1994) ได้เสนอวิธีการแยกบรรทัดไว้ โดยการใช้การฉายเงาเช่นกัน ซึ่งนำมาประยุกต์ใช้กับการหาเส้นบรรทัดของคำในภาษาไทยได้ Lohakan, Airphaiboon และ Sangworasil (2000) ได้เสนอวิธีการแบ่งเส้นบรรทัดออกเป็น 4 ระดับด้วยกันตามตำแหน่งการเกิดของอักขระ แต่ในงานวิจัยนี้จะทำการแบ่งระดับของเส้นบรรทัดโดยคำนึงถึงตัวอักษรที่เกิดในตำแหน่งบนบรรทัดด้วยทำให้สามารถแบ่งระดับของตัวอักษรในคำได้เป็น 3 ระดับ ดังแสดงในรูป 2.13 ในกรณีที่มีสระด้านบน และวรรณยุกต์ จะรวมกันอยู่ในระดับเดียว ซึ่งการหาเส้นบรรทัดนี้สามารถทำได้โดยพิจารณาความหนาแน่นของจุดภาพในแนวนอน มาพิจารณาโดยตำแหน่งมีความหนาแน่นของจุดภาพน้อยที่สุดและมีค่าต่ำกว่าค่าเริ่มเปลี่ยนแปลง (Threshold) จะถือว่าเป็นเส้นบรรทัดดังแสดงในรูป 2.13 โดยเส้นบรรทัดที่ได้จะมีอยู่ 2 เส้นคือ เส้นบรรทัดด้านบน (Upper Baseline) และเส้นบรรทัดด้านล่าง (Lower Baseline)



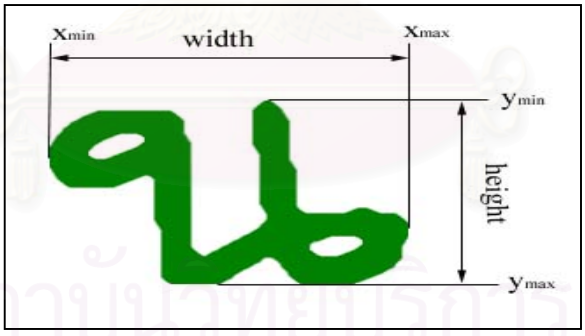
รูปที่ 2.13 วิธีการหาเส้นบรรทัดโดยพิจารณาความหนาแน่นของจุดภาพ

2.2.7 **การหาอัตราส่วนกรอบภาพ** (Aspect ratio)

Malaviya และ Peters (1997) ได้เสนอการหาอัตราส่วนขอบภาพเพื่อบ่งบอกลักษณะภายนอกของกรอบตัวอักษร ซึ่งสามารถนำมาใช้กับตัวอักษรไทยได้ เนื่องจากตัวอักษรพยัญชนะไทย บางตัวมีโครงสร้างที่ใกล้เคียงกัน ทำให้ลักษณะสำคัญที่วิเคราะห์ได้มีความเหมือนกัน แต่จะมีความแตกต่างกันในด้านความกว้างของตัวอักษร ดังนั้นการพิจารณาอัตราส่วนของกรอบภาพจึงเป็นสิ่งจำเป็นในการแยกตัวอักษร ตัวอย่างได้แก่ "ข" กับ "บ" เป็นต้น การพิจารณาหาอัตราส่วนกรอบภาพสามารถคำนวณได้ตามสมการ

$$\mu_{ap} = \frac{y_{\max} - y_{\min}}{\sqrt{\left(x_{\max} - x_{\min}\right)^2 + \left(y_{\max} - y_{\min}\right)^2}} \qquad (2.6)$$

โดย $y_{\max} - y_{\min}$ เป็นความสูงของตัวอักษรและ $x_{\max} - x_{\min}$ เป็นความกว้างของตัวอักษร ดังแสดงในรูปที่ 2.14



รูป 2.14 การวัดความกว้างและความสูงของตัวอักษร

2.3 **ฟัซซีโลจิก** (Fuzzy Logic)

ทฤษฎีทางฟัซซีโลจิกได้ถูกเสนอขึ้นโดย Lofti Zadeh ในปี 1965 เป็นการใช้ทฤษฎีของเซตมาประยุกต์เพื่อแสดงลักษณะที่คลุมเครือของสิ่งต่างที่มีความไม่แน่นอน หรือมีความซับซ้อน ซึ่งยากต่อการสร้างแบบจำลองทางคณิตศาสตร์ ฟัซซีโลจิกถูกนำมาเพื่อสร้างแบบจำลองที่มีความใกล้เคียงความเป็นจริง โดยมีการพิจารณาลักษณะของปัญหาที่สนใจ แล้วหาค่าความเป็นสมาชิกเพื่อเป็นตัวแปรที่นำมาตัดสินใจ ซึ่งการประยุกต์ใช้ฟัซซีโลจิกทำให้ระบบสามารถบ่งบอกถึงความไม่แน่นอน และเมื่อมีการนำมาใช้ร่วมกับการตัดสินใจทำให้มีความยืดหยุ่นในการตัดสินใจมากยิ่ง

ขึ้น นอกจากนี้ได้มีการนำเอาวิธีทางฟัซซีโลจิกมาประยุกต์ใช้กับงานหลาย ๆ ด้านเช่น ระบบควบคุม, ระบบทางชีววิทยา และระบบรู้จำ

ในงานวิจัยด้านการรู้จำรูปแบบได้มีการนำเทคนิคทางฟัซซีโลจิกมาประยุกต์อย่างกว้างขวาง เพื่อเพิ่มความยืดหยุ่นในการตัดสินใจ และแก้ปัญหาความไม่แน่นอนของรูปแบบที่เกิดขึ้น รวมถึงลดความซับซ้อนของระบบจึงเหมาะที่จะนำมาใช้กับปัญหาที่สร้างแบบจำลองทางคณิตศาสตร์ได้ยาก หรือมีความซับซ้อนของสูง ซึ่งเป็นปัญหาสำคัญในการรู้จำรูปแบบ โดยเฉพาะการรู้จำอักษรลายมือเขียน Lee, Huang และ Sheu (1998) ได้นำเสนอวิธีการหาคุณลักษณะสำคัญโดยใช้ฟัซซีเพื่อพิจารณาโครงสร้างของลายเส้นในการรู้จำตัวอักษรภาษาจีน ซึ่งเป็นภาษาที่ตัวอักษรมีลายเส้นอยู่มาก และตัดสินใจโดยพิจารณาการเชื่อมต่อของลายเส้นด้วยกฎทางฟัซซี ในงานเกี่ยวกับตัวอักษรภาษาไทย เดชา รัตนาธาร (2538) ได้นำฟัซซีโลจิกมาประยุกต์กับการรู้จำโครงสร้างของตัวเลขไทยโดยใช้วิธีการแบบซินแทคติก

2.3.1 **นิยาม ฟัซซีเซต**

กำหนดให้ X เป็นกลุ่มของวัตถุซึ่งแทนด้วยเซต $\{x\}$ ซึ่งเป็นได้ทั้งแบบต่อเนื่องและไม่ต่อเนื่อง ดังนั้นฟัซซีเซต F ใน X ถูกแทนด้วยเซตของคู่ลำดับ

$$F = \{(x, \mu_X(x)) \mid x \varepsilon X\} \qquad (2.7)$$

โดย $\mu_F(x)$ เป็นฟังก์ชันความเป็นสมาชิก หรือ ระดับความเป็นสมาชิก ของสมาชิก x ในฟัซซีเซต F และ ในฟัซซีเซต F ในช่วงค่าจำนวนจริงระหว่าง [0,1]

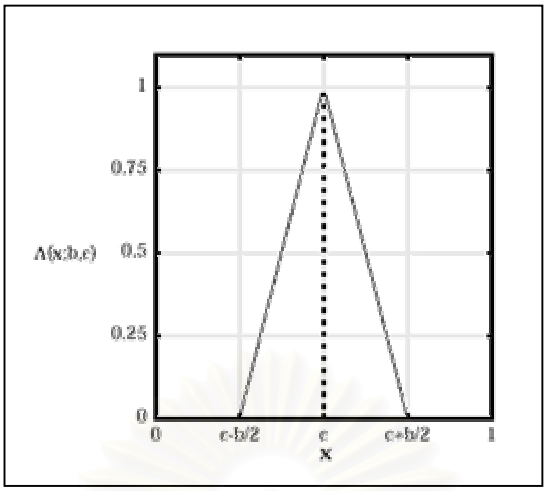
2.3.2 **ฟังก์ชันความเป็นสมาชิก**

ฟังก์ชันความเป็นสมาชิกในฟัซซีเซตคือฟังก์ชันที่บ่งบอกถึงระดับของคุณลักษณะ โดยมีค่าเป็นค่าจริงอยู่ระหว่าง 0 กับ 1 ฟังก์ชันความเป็นสมาชิกเป็นได้ทั้งฟังก์ชันต่อเนื่อง และไม่ต่อเนื่องฟังก์ชันความเป็นสมาชิกที่ใช้กันทั่วไปมีหลายรูปทรงไม่ว่าจะเป็น สามเหลี่ยม (Triangular shape), รูปแบบเอส และพาย (S- and π- shape) ในงานวิจัยนี้มีฟังก์ชันความเป็นสมาชิกทั้งที่ต่อเนื่องในรูปแบบสามเหลี่ยม และไม่ต่อเนื่องซึ่งจะกล่าวโดยละเอียดในบทต่อไป

2.3.2.1 **ฟังก์ชันความเป็นสมาชิกแบบสามเหลี่ยม**

ฟังก์ชันความเป็นสมาชิกแบบสามเหลี่ยมเป็นฟังก์ชันที่นิยมใช้กันมากที่สุด ซึ่งมีรูปทรงดังรูปที่ 2.15 และมีสมการดังนี้

$$\Delta(x; b, c) = \begin{cases} 1 - 2\left|\left(\dfrac{x-c}{b}\right)\right| & \text{สำหรับ } \left(c - \dfrac{b}{2}\right) \le x \le \left(c + \dfrac{b}{2}\right) \\ 0 & \text{ค่าอื่น} \end{cases} \qquad (2\text{-}8)$$



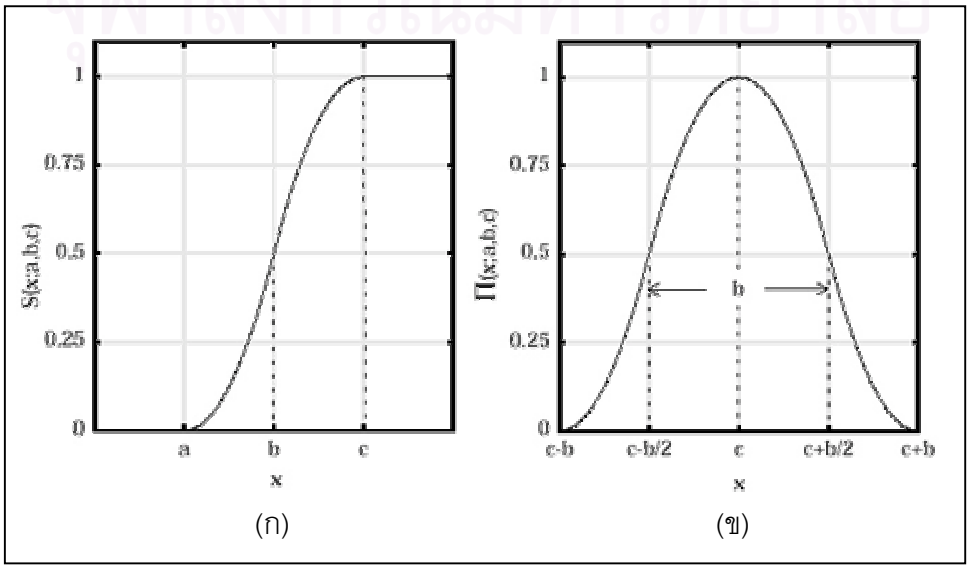
รูปที่ 2.15 ฟังก์ชันความเป็นสมาชิกแบบสามเหลี่ยม

2.3.2.2 **ฟังก์ชันความเป็นสมาชิกแบบเอส และฟังก์ชันความเป็นสมาชิกแบบพาย** (S- and π- functions)

ฟังก์ชันเอสและพาย กำหนดได้จากสมการ

$$S(x;a,b,c) = \begin{cases} 0 & ; x \le a \\ 2\cdot\left(\frac{x-a}{c-a}\right)^2 & ; a<x\le b \\ 1-2\cdot\left(\frac{x-c}{c-a}\right)^2 & ; b<x<c \\ 1 & ; x>c \end{cases} \quad (2\text{-}9)$$

$$\Pi(x;b,c) = \begin{cases} S\left(x;c-b,c-\frac{b}{2},c\right) & ; x < c \\ 1-s\left(x;c,c+\frac{b}{2},c+b\right) & ; x>c \end{cases} \quad (2\text{-}10)$$



รูปที่ 2.16 ฟังก์ชันความเป็นสมาชิก ก) แบบเอส และ ข) แบบพาย

2.3.3 **ตัวดำเนินการฟัซซี** (Fuzzy Operators)

ตัวดำเนินการพื้นฐานของฟัซซีเซทมีอยู่ 3 อย่างคือ ยูเนียน(union), อินเตอร์เซคชัน (intersection), และ คอมพลีเมนท์(complement)

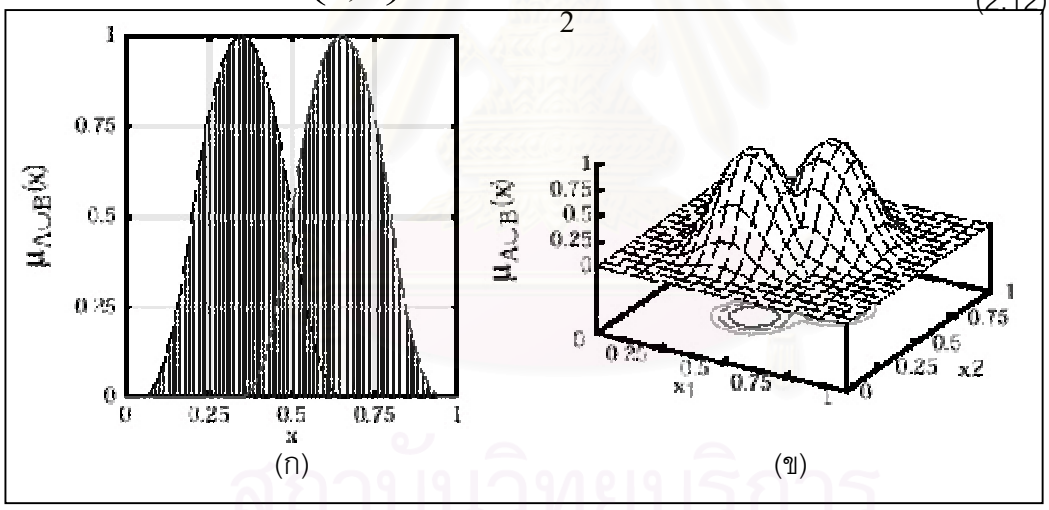
2.3.3.1 **ตัวดำเนินการยูเนียน** (Union Operator)

ฟังก์ชันความเป็นสมาชิก $\mu_{A\cup B}(x)$ ถูกกำหนดสำหรับทุก ๆ x ที่เป็นสมาชิกของ X โดย

$$\mu_{A\cup B}(x) = \text{MAX}\ \{\mu_A(x), \mu_B(x)\} \qquad (2.11)$$

โดยที่ MAX(a,b) คือค่าที่มากสุดระหว่าง a และ b ซึ่งสามารถเขียนเป็นสมการทางพีชคณิตได้ดังนี้

$$\text{MAX}\ (a,b) = \frac{a+b+|a-b|}{2} \qquad (2.12)$$



รูปที่ 2.17 ก) ผลการยูเนียนของเซตสองเซตที่อยู่บนโดเมนเดียวกัน
ข) ผลการยูเนียนของเซตสองเซตที่อยู่บนคนละโดเมน

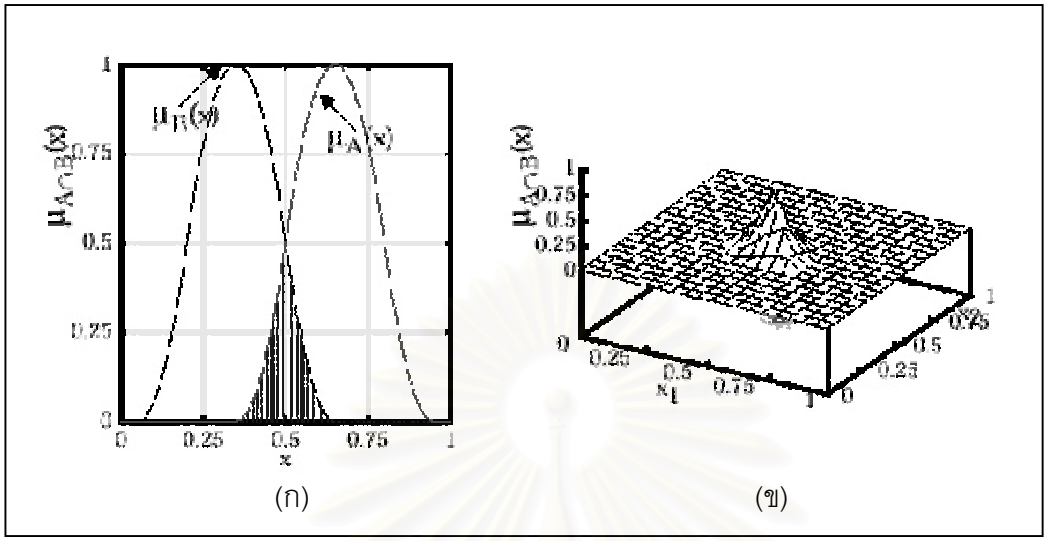
2.3.3.2 **ตัวดำเนินการอินเตอร์เซคชัน** (Intersection Operator)

ฟังก์ชันความเป็นสมาชิก $\mu_{A\cap B}(x)$ ถูกกำหนดสำหรับทุก ๆ x ที่เป็นสมาชิกของ X โดย

$$\mu_{A\cap B}(x) = \text{MIN}\ \{\mu_A(x), \mu_B(x)\} \qquad (2.13)$$

โดยที่ MIN(a,b) คือค่าที่มากสุดระหว่าง a หรือ b ซึ่งสามารถเขียนเป็นสมการทางพีชคณิตได้ดังนี้

$$\text{MIN}\ (a, b) = \frac{a + b - |a - b|}{2} \qquad (2.14)$$

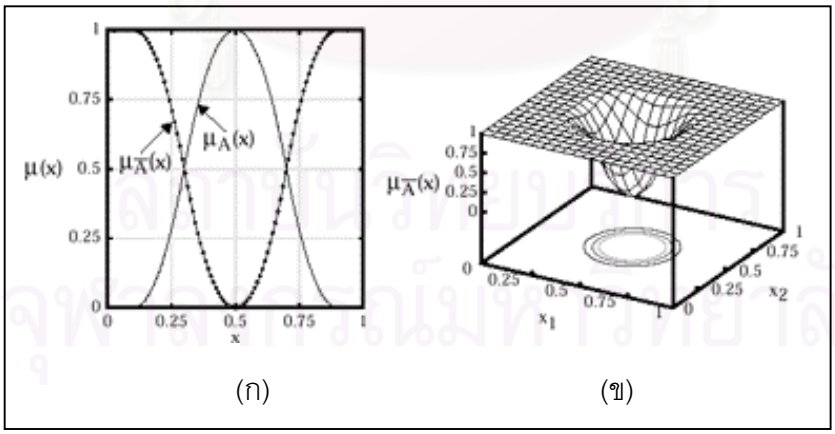


รูปที่ 2.18 ก) ผลการอินเตอร์เซคชันของเซตสองเซตที่อยู่บนโดเมนเดียวกัน
ข) ผลการอินเตอร์เซคชันของเซตสองเซตที่อยู่บนคนละโดเมน

2.3.3.3 **ตัวดำเนินการคอมพลีเมนท์** (Complement Operator)

ฟังก์ชันความเป็นสมาชิก $\mu_{\bar{A}}(x)$ ถูกกำหนดสำหรับทุก ๆ x ที่เป็นสมาชิกของ X โดย

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \qquad (2.15)$$



รูปที่ 2.19 ก) ผลการคอมพลีเมนท์ของเซตสองเซตที่อยู่บนโดเมนเดียวกัน
ข) ผลการคอมพลีเมนท์ของเซตสองเซตที่อยู่บนคนละโดเมน

2.3.4 **ตัวแปรฟัซซีเชิงภาษา**

ส่วนที่สำคัญในแบบจำลองฟัซซี คือการกำหนดตัวแปรเชิงภาษา ในวิธีเชิงภาษาจะใช้คำแทนค่าของตัวเลขเพื่ออธิบายเหตุการณ์ที่ซับซ้อนเกินไป หรือยากที่จะกำหนดคุณลักษณะในเชิงปริมาณ ในการเปรียบเทียบหน่วยทำได้โดยใช้ฟัซซีเซตมากกว่าหนึ่งตัว เช่น “มาก” “ค่อนข้างมาก” “ค่อนข้างน้อย” “น้อย” คุณสมบัติตัวอย่าง เช่น ความกลม สามารถวัดในลักษณะ“กลมเล็กน้อย” “กลมมาก” “กลมโดยสมบูรณ์”

2.3.5 การประมาณเหตุผลด้วยฟัซซี

ฟังก์ชันความเป็นสมาชิกจะเชื่อมโยงตัวแปรค่าจริงไปยังตัวแปรเชิงภาษา ดังนั้นข้อมูลสามารถประมาณสรุปได้ ในระบบฟัซซีความไม่แน่นอนนี้ถูกเรียกว่าความเป็นฟัซซี เพราะความไม่แน่นอนทำให้ไม่สามารถกำหนดกลุ่มสมาชิกอย่างเด็ดขาด

ในการตัดสินใจด้วยกฎทางฟัซซี ทำได้โดยใช้การแสดงเงื่อนไข “ถ้า ….. แล้ว…….”

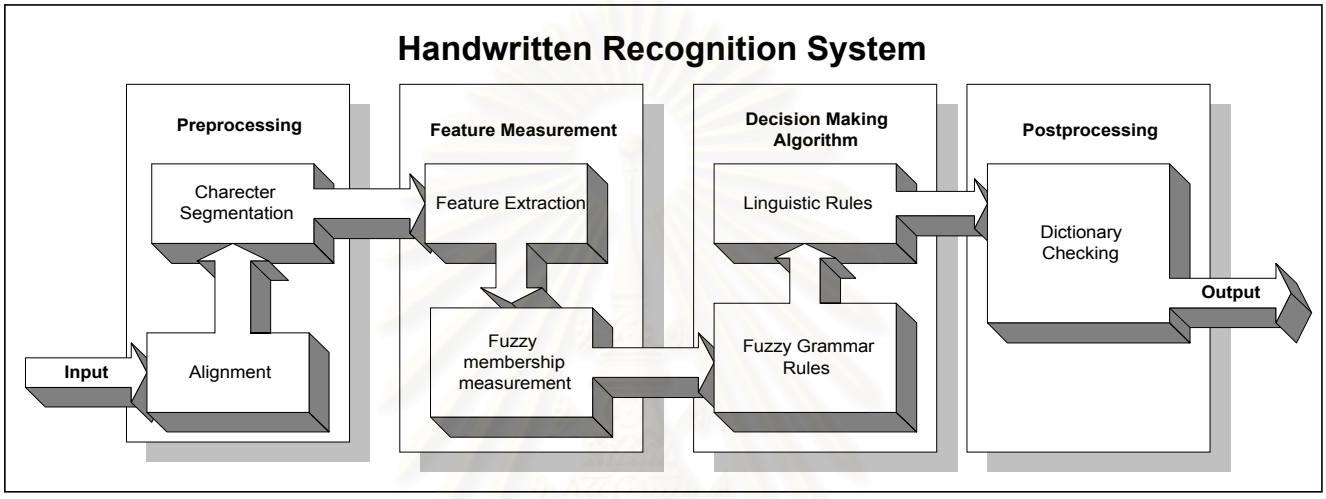
“ ถ้า x คือ A แล้ว y คือ B” สามารถเขียนเป็นสัญลักษณ์ได้เป็น “A $\rightarrow$ B” โดยที่ A และ B เป็นฟัซซีเซตในเอกภพสัมพัทธ์(Universe) X ด้วย ฟังก์ชันความเป็นสมาชิก μ_A และ μ_B

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

การรู้จำตัวอักษรภาษาไทยที่เป็นคำ

ในบทนี้กล่าวถึงการรู้จำตัวอักษรภาษาไทยที่เป็นคำ โดยระบบรู้จำสามารถแบ่งออกเป็น 3 ขั้นตอนหลักดังแสดงในรูปที่ 3.1



รูปที่ 3.1 โครงสร้างของระบบรู้จำ

3.1 การเก็บข้อมูลภาพลายมือเพื่อใช้ในการวิเคราะห์และทดสอบระบบรู้จำตัวอักษร

ในการเก็บข้อมูลภาพลายมือเขียน แบ่งได้เป็น 2 กลุ่ม คือ

1. ภาพอักษรตัวเดี่ยว ได้แก่ พยัญชนะ สระ และวรรณยุกต์ เก็บข้อมูลโดยการเขียนบนกระดานอิเลคทรอนิกส์

2. ลายมือเขียนที่เป็นคำ ได้มาจากเขียนลงบนกระดาษเก็บข้อมูล ด้วยปากกาหมึกสีดำ แล้วนำภาพที่ได้จากการเขียนมาสแกนที่ความละเอียด 300 จุดภาพต่อความยาว 1นิ้ว (dpi) จากนั้นภาพที่ได้จะถูกนำมาแปลงให้เป็นภาพแบบ 2 ระดับ (Binary) และเก็บที่ขนาด 510x150 จุดภาพในรูปแบบของ Portable Gray Map (pgm)

3.2 กระบวนการก่อนหน้า

กระบวนการก่อนหน้า เป็นขั้นตอนที่สำคัญเพื่อปรับคุณภาพของข้อมูลภาพให้มีความเหมาะสมก่อนส่งเข้าไปในระบบรู้จำ ซึ่งมี 5 ขั้นตอนดังนี้

1. การลดสัญญาณรบกวน เป็นขั้นตอนการลดสัญญาณรบกวนที่มีขนาดเล็ก (salt and pepper noise) ที่เกิดขึ้นเนื่องจากการสแกนภาพ วิธีการลดสัญญาณรบกวนสามารถทำได้ตามหัวข้อ 2.1.1
2. การทำขอบภาพให้เรียบ เป็นขั้นตอนการปรับขอบภาพที่มีความเปลี่ยนแปลงมากเนื่องจากสัญญาณรบกวนที่เกิดขึ้นที่ขอบของภาพตัวอักษร และผลเนื่องจากการซึมของหมึกในกระดาษในระหว่างการเขียน วิธีการทำขอบของภาพให้เรียบสามารถทำได้ดังแสดงไว้ในหัวข้อ 2.1.2
3. การแก้ความเอียง เป็นขั้นตอนที่ปรับตัวอักษรให้ตั้งตรง ในการแก้ตัวอักษรเอียง ตัวอักษรทุกตัวในคำต้องมีการเอียงในทิศทางเดียวกัน การแก้ความเอียงจะทำการหาค่ามุมที่ตัวอักษรทุกตัวในคำที่เอียง แล้วทำการชดเชยมุมที่เกิดขึ้น เพื่อปรับตัวอักษรให้ตรง ดังแสดงไว้ในหัวข้อ 2.1.3
4. การตัดแยกคำ เป็นขั้นตอนในการตัดแยกตัวอักษรเดี่ยวออกจากคำ และตัดตัวอักษรที่ติดกันให้แยกออก เพื่อส่งไปรู้จำทีละตัวอักษร ซึ่งการแยกตัวอักษรที่ไม่มีการติดกันแสดงไว้ในหัวข้อที่ 2.2.6 ส่วนคำที่เกิดการติดกันของตัวอักษรภายในคำจะต้องตัดแยกเอง
5. การหาเส้นบรรทัด เป็นขั้นตอนในการแบ่งระดับบรรทัด และแยกตัวอักษรในแต่ละระดับดังแสดงไว้ในหัวข้อ 2.2.3

3.3 ขั้นตอนการรู้จำ

ตัวอักษรภาษาไทยมีลักษณะพิเศษแตกต่างจากภาษาอื่นอยู่หลายอย่าง เช่น ตัวอักษรภาษาไทยโดยส่วนใหญ่มีหัวเป็นส่วนประกอบ บางตัวอักษรมีรอยหยัก ซึ่งพอจะแยกเป็นลักษณะพิเศษได้เป็น หัว จุดปลาย หยักที่ติดกับหัว หยักที่อยู่บริเวณกลางตัวอักษร และหาง ตัวอักษรไทยทั้งหมดพอจะกล่าวได้ว่าเกิดจากการประกอบกันขึ้นของลักษณะพิเศษข้างต้น โดยเชื่อมต่อกันด้วยเส้นตรง หรือเส้นโค้ง

เพราะฉะนั้นในขั้นตอนการรู้จำ หากระบบสามารถค้นหาลักษณะเฉพาะของตัวอักษร ก็สามารถจำแนกความแตกต่างของตัวอักษรในแต่ละตัวได้ แต่เนื่องจากในอักษรลายมือเขียนลักษณะดังกล่าวอาจเกิดการเปลี่ยนรูปร่างและลักษณะไป ตามผู้เขียน และสภาพแวดล้อม จึงจำเป็นต้องหาความเป็นไปได้ที่จะเกิดโครงสร้างต่าง ๆ ขึ้นเพื่อจำแนกตัวอักษร

ขั้นตอนการรู้จำสามารถแบ่งย่อยออกเป็น 3 ขั้นตอนหลัก คือ การหาคุณลักษณะสำคัญ, การคำนวณค่าความเป็นสมาชิกของคุณลักษณะสำคัญ และ การจำแนกตัวอักษร

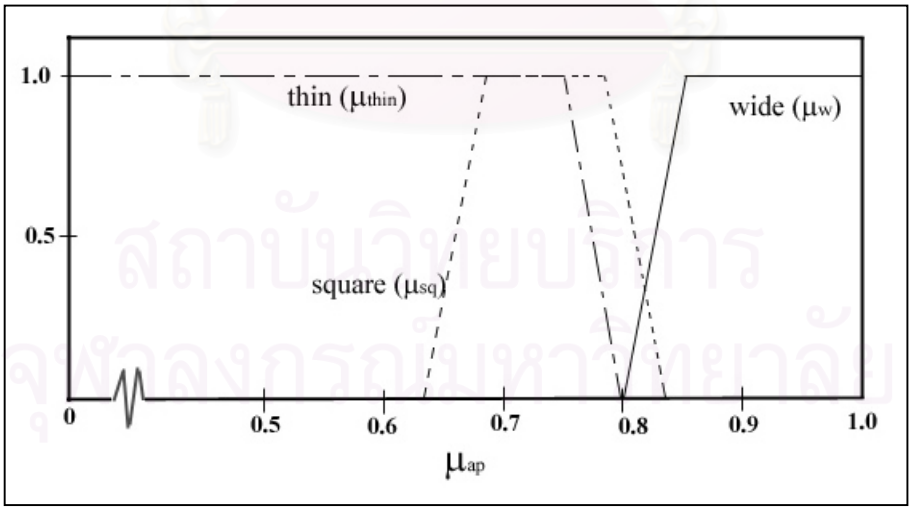
3.3.1 **การสกัดคุณลักษณะสำคัญ** (Feature Extraction) **และวัดค่าความเป็นสมาชิก** (fuzzy membership measurement)

การสกัดลักษณะสำคัญ คือขั้นตอนการวิเคราะห์ข้อมูลภาพเพื่อบ่งบอกลักษณะและความแตกต่างของตัวอักษรและสระ ลักษณะสำคัญที่นำมาใช้ในการรู้จำได้แก่ อัตราส่วนกรอบภาพ (Aspect Ratio), วงรอบ (Loop), จุดปลาย(End Point), รอยหยัก (Curl), หาง (Tail), และลำดับสายการลากผ่าน(Stroke Changing Sequence)

3.3.1.1 **อัตราส่วนกรอบภาพ** (Aspect Ratio)

เนื่องจากตัวอักษรพยัญชนะไทย บางตัวมีโครงสร้างที่ใกล้เคียงกัน ทำให้ลักษณะสำคัญที่วิเคราะห์ได้มีความเหมือนกัน แต่จะมีความแตกต่างกันในด้านความกว้างของตัวอักษร ดังนั้นการพิจารณาอัตราส่วนของกรอบภาพจึงเป็นสิ่งจำเป็นในการแยกตัวอักษร ตัวอย่างได้แก่ “ข” กับ “บ” เป็นต้น การพิจารณาหาอัตราส่วนกรอบภาพสามารถคำนวณได้ดังหัวข้อ 2.2.3

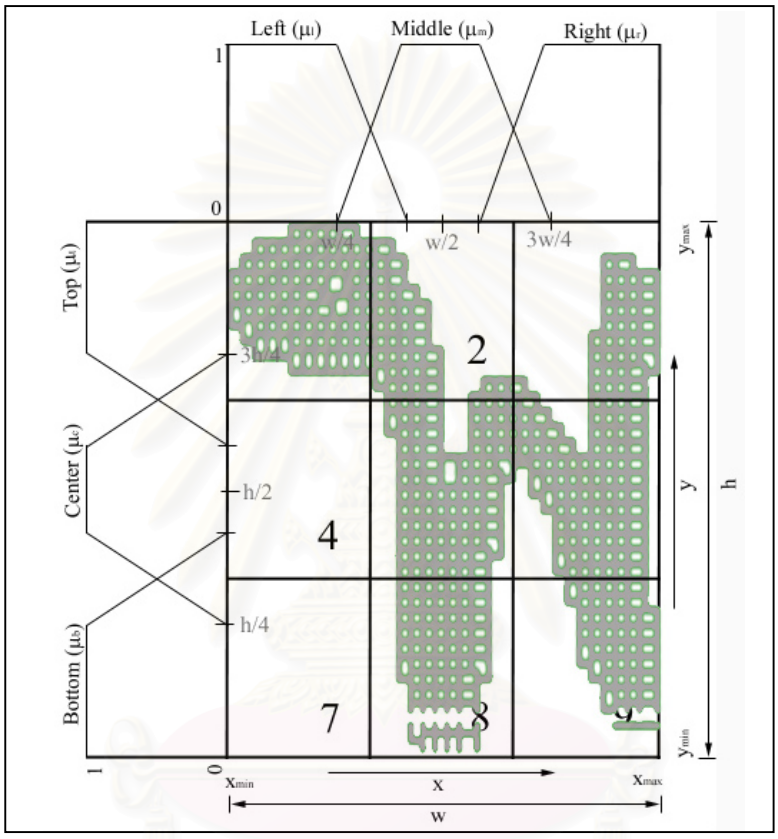
อัตราส่วนของกรอบภาพที่คำนวณได้บ่งบอกถึงรูปร่างภายนอกของตัวอักษรซึ่งสามารถใช้ตัวแปรเชิงภาษา 3 แบบคือ “กว้าง” “สี่เหลี่ยม” “สูง” เพื่อบ่งบอกลักษณะของตัวอักษรได้ ซึ่งมีฟังก์ชันความเป็นสมาชิกดังรูปที่ 3.2



รูปที่ 3.2 ฟังก์ชันความเป็นสมาชิกของอัตราส่วนของกรอบตัวอักษร

3.3.1.2 **การพิจารณาตำแหน่งอ้างอิงของตัวอักษร** (Positional Feature)

เนื่องจากตัวอักษรไทยเกิดจากการผสมกันของคุณลักษณะดังที่ได้กล่าวในตอนต้น ซึ่งในแต่ละตัวอักษรจะมีความแตกต่างกันในเชิงตำแหน่ง และทิศทางเป็นหลัก ระบบจึงต้องพิจารณาตำแหน่งของคุณลักษณะเทียบกับกรอบอ้างอิงของตัวอักษรเพื่อใช้จำแนกความต่างของตัวอักษร ระบบจะทำการแบ่งบริเวณต่าง ๆ ของตัวอักษรออกเป็น 9 ส่วนดังรูปที่ 3.3 และกำหนดตัวแปรเชิงภาษาเพื่อบอกถึงตำแหน่งของคุณลักษณะต่าง ๆ ด้วยตัวแปร 6 ตัวดังนี้ ในแนวนอนมีตัวแปร 3 ตัว คือ “ซ้าย” “กลาง” “ขวา” และในแนวตั้งมีตัวแปร 3 ตัวคือ ”บน” “กลาง” “ล่าง”



รูปที่ 3.3 ฟังก์ชันความเป็นสมาชิกของตัวแปรเชิงตำแหน่ง

จากรูปที่ 3.3 ตัวแปรทางภาษาทั้ง 6 ตัวสามารถคำนวณค่าความเป็นสมาชิกได้ด้วยสมการต่อไปนี้

$$\mu_{\text{Left}} = \begin{cases} 1 & ; \ x < w/4 \\ \dfrac{\left(5w/12\right) - x}{w/6} & ; \ w/4 < x < 5w/12 \\ 0 & ; \ x > 5w/12 \end{cases} \qquad (3.1n)$$

$$\mu_{\text{Middle}} = \begin{cases} 1 - \dfrac{\left| x - \left(5w/12 \right) \right|}{w/6} & ;\ w/4 < x < 5w/12 \\ 1 & ;\ 5w/12 < x < 7w/12 \\ 1 - \dfrac{\left| x - \left(7w/12 \right) \right|}{w/6} & ;\ 7w/12 < x < 3w/4 \\ 0 & ;\ x < w/4 \text{ or } x > 3w/4 \end{cases} \quad (3.1\text{ข})$$

$$\mu_{\text{Right}} = \begin{cases} 0 & ;\ x < 7w/12 \\ \dfrac{x - \left(7w/12 \right)}{w/6} & ;\ 7w/12 < x < 3w/4 \\ 1 & ;\ x > 3w/4 \end{cases} \quad (3.1\text{ค})$$

$$\mu_{\text{Bottom}} = \begin{cases} 1 & ;\ y < h/4 \\ \dfrac{\left(5h/12 \right) - y}{h/6} & ;\ h/4 < y < 5h/12 \\ 0 & ;\ y > 5h/12 \end{cases} \quad (3.1\text{ง})$$

$$\mu_{\text{Center}} = \begin{cases} 1 - \dfrac{\left| y - \left(5h/12 \right) \right|}{h/6} & ;\ h/4 < y < 5h/12 \\ 1 & ;\ 5h/12 < y < 7h/12 \\ 1 - \dfrac{\left| y - \left(7h/12 \right) \right|}{h/6} & ;\ 7h/12 < y < 3h/4 \\ 0 & ;\ y < h/4 \text{ or } y > 3h/4 \end{cases} \quad (3.1\text{จ})$$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

$$\mu_{Top} = \begin{cases} 0 & ; y < 7h/12 \\ \dfrac{y - \left(7h/12\right)}{h/6} & ; 7h/12 < y < 3h/4 \\ 1 & ; y > 3h/4 \end{cases} \quad (3.1ฉ)$$

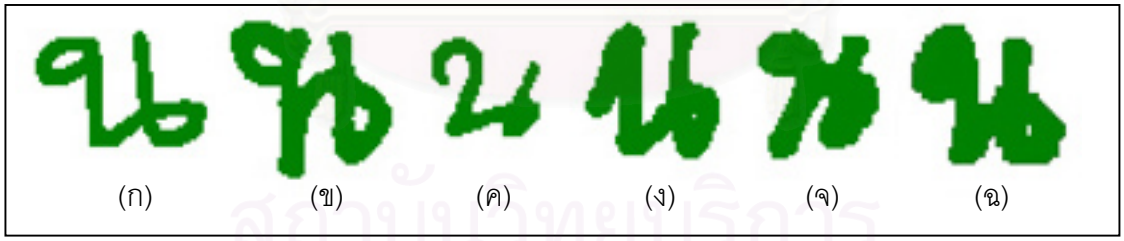
โดย $w = x_{max} - x_{min}$, $h = y_{max} - y_{min}$ และ

$\mu_{Left}, \mu_{Middle}, \mu_{Right}, \mu_{Top}, \mu_{Center}, \mu_{Bottom}$ เป็นตัวแปรฟัซซีที่บอกลักษณะเชิงตำแหน่ง

3.3.1.3 **การตรวจหาวงรอบ** (Loop Feature)

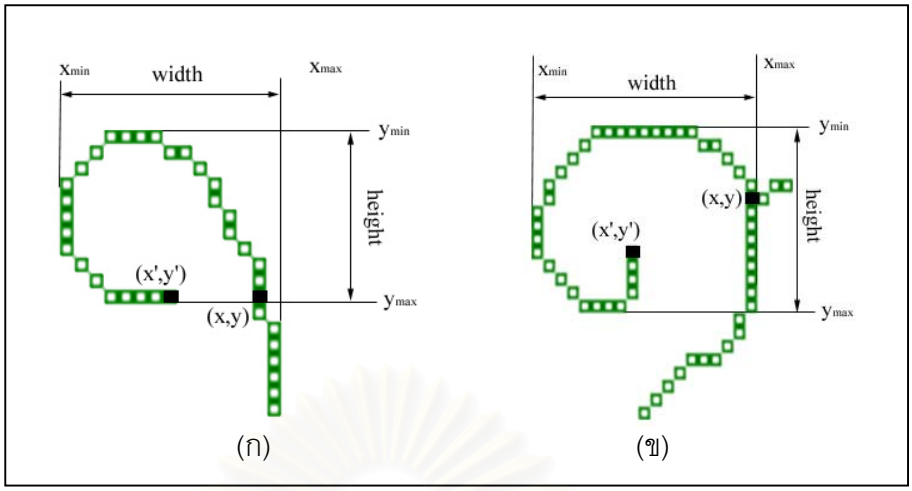
เนื่องจากตัวอักษรภาษาไทยส่วนใหญ่มีหัวเป็นส่วนประกอบที่สำคัญ ซึ่งสามารถตรวจสอบได้โดยพิจารณาวงรอบที่เกิดขึ้นในตัวอักษร การบอกความแตกต่างของตัวอักษรแต่ละชนิดจึงทำได้ง่ายเมื่อพิจารณาจากวงรอบของตัวอักษร แต่ในลายมือเขียนวงรอบอาจมีความไม่สมบูรณ์หรือมีลักษณะที่แตกต่างออกไป แต่เนื่องจากวงรอบมีความสำคัญในตัวอักษรภาษาไทย จึงได้หาวิธีการตรวจสอบหาวงรอบทั้งที่สมบูรณ์ และไม่สมบูรณ์

ในกรณีที่ตัวอักษรมีวงรอบที่สมบูรณ์สามารถตรวจหาได้โดยการเน้นขอบภาพ จากนั้นทำการลบขอบภาพทางด้านนอกออก ด้วยวิธีที่กล่าวมาแล้วในหัวข้อที่ 2.3 และ 2.4 จุดภาพที่เหลือจะเป็นขอบของวงรอบที่สมบูรณ์ของตัวอักษร ดังแสดงในรูปที่ 2.10



รูปที่ 3.4 หัวของตัวอักษร “น” ที่เกิดขึ้นในลายมือเขียน

ในรูปที่ 3.4 แสดงลักษณะของหัวของตัวอักษร “น” ที่เกิดขึ้นในการเขียนซึ่งรูปที่ 3.4 (ก) และ (ข) เป็นรูป “น” ที่มีหัวสมบูรณ์ ในรูป (ค) และ (ง) เป็นหัวของตัวอักษรที่ไม่สมบูรณ์ และในรูป (จ) และ (ฉ) เป็นลักษณะตัวอักษรที่มีหัวตัน ซึ่งในกรณีหัวของตัวอักษรที่ไม่สมบูรณ์ หรือหัวตันสามารถตรวจสอบได้โดยทำโครงร่างบางของตัวอักษร และทำการตรวจสอบตามเส้นโครงร่างโดยการนับความยาวของทางเดินไปเรื่อยๆ และจะหยุดนับเมื่อตำแหน่งที่พบมีความสูงระดับเดียวกันกับจุดปลายที่ตรวจสอบ หรือพบจุดแยก ซึ่งพิจารณากรอบของจุดภาพที่เดินผ่านจะได้ดังรูปที่ 3.5



รูปที่ 3.5 การพิจารณาหัวที่ไม่สมบูรณ์

(ก) เมื่อพบจุดภาพที่ระดับเดียวกับจุดปลาย

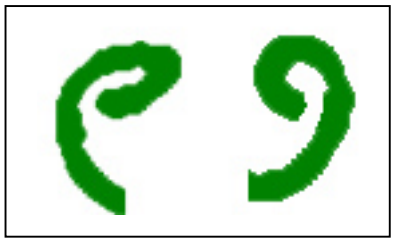
(ข) เมื่อพบจุดแยก

ความสมบูรณ์ของวงรอบของตัวอักษรสามารถคำนวณได้จากสมการ

$$\mu_{Loop} = 1 - \left| 0.5 \cdot \frac{x' - x}{x_{max} - x_{min}} \right| - \left| 0.5 \cdot \frac{y' - y}{y_{max} - y_{min}} \right| \qquad (3.2)$$

โดยที่ (x', y') คือตำแหน่งของจุดปลาย และ (x, y) คือจุดที่หยุดการนับเส้นทาง

นอกจากจะพิจารณาความสมบูรณ์ของวงรอบแล้ว ลักษณะของวงรอบที่ต้องพิจารณาอีก 2 ประการ คือ ตำแหน่งและทิศทางของวงรอบ ในการพิจารณาตำแหน่งทำได้โดยหาค่าจุดศูนย์กลางของ วงรอบ แล้วนำไปเทียบกับบริเวณอ้างอิงในหัวข้อที่ 3.2.2 เพื่อบ่งบอกตำแหน่ง ส่วนการพิจารณาทิศทางของวงรอบทำได้โดยการ พิจารณาจุดแยก (Junction Point) ที่ติดกับวงรอบ ว่าอยู่บริเวณด้านซ้ายหรือขวา ของจุดศูนย์กลางของวงรอบ ถ้าจุดกึ่งกลางอยู่ทางด้านซ้ายให้ถือว่าเป็นวงรอบที่หันไปทางซ้าย แต่ถ้าจุดกึ่งกลางอยู่ทางด้านขวาให้ถือว่าเป็นวงรอบที่หันไปทางขวา ส่วนในกรณีวงรอบที่ไม่ครบวง พิจารณาจากจุดปลายกับเส้นทางที่ตรวจสอบ ถ้าจุดปลายอยู่ด้านซ้ายของจุดภาพที่ใกล้ที่สุดในระดับเดียวกัน แสดงว่าวงรอบหันไปทางด้านซ้าย แต่ถ้าจุดปลายอยู่ด้านขวาของจุดภาพที่ใกล้ที่สุดในระดับเดียวกัน แสดงว่าวงรอบหันไปทางด้านขวา ดังแสดงในรูปที่ 3.6



รูปที่ 3.6 หัวของตัวอักษรที่หันไปด้านขวาและซ้าย

ในการพิจารณาหัวของตัวอักษรเราจะใช้คุณลักษณะ 2 ประการมาพิจารณาคือ ความสมบูรณ์ของวงรอบและตำแหน่งของวงรอบ เนื่องจากทิศทางของหัวตัวอักษรสามารถกำหนดได้ชัดเจนจากการพิจารณาข้างต้น ดังนั้นในการพิจารณาค่าความเป็นสมาชิกของหัวสามารถคำนวณได้จากสมการ

$$\mu_{\text{Head}} = \mu_{\text{Loop}} \cdot \left(w_x \mu_H + w_y \mu_V \right) \tag{3.3}$$

โดย w_x, w_y เป็นค่าถ่วงน้ำหนักของตำแหน่งที่เกิดหัวของตัวอักษร ค่าถ่วงน้ำหนักมีค่ามากแสดงว่าหัวของตัวอักษรมีความคลาดเคลื่อนเชิงตำแหน่งมีค่าน้อย μ_x, μ_y คือค่าความเป็นสมาชิกในเชิงตำแหน่งในแนวแกน x และ แนวแกน y

ค่าถ่วงน้ำหนัก w_x, w_y สามารถคำนวณได้จากค่าส่วนเบี่ยงเบนมาตรฐาน(Standard Deviation: S.D.)ของตำแหน่งของคุณลักษณะทั้งในแนวแกน x และแกน y ที่ได้จากข้อมูลตัวอักษรในฐานข้อมูลซึ่งเป็นไปตามสมการดังนี้

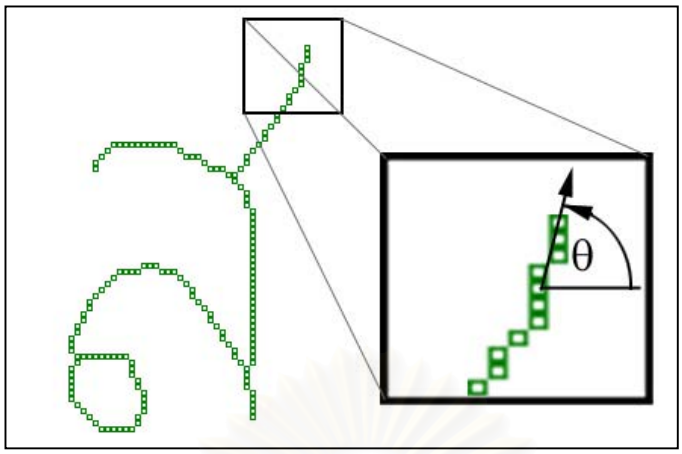
$$w_x = 1 - \frac{\text{S.D.}_x}{\text{S.D.}_x + \text{S.D.}_y} \tag{3.4}$$

$$w_y = 1 - \frac{\text{S.D.}_y}{\text{S.D.}_x + \text{S.D.}_y} \tag{3.5}$$

3.3.1.4 การตรวจหาจุดปลาย (End Point Feature)

จุดปลายของตัวอักษรของแต่ละตัวอักษรจะมีตำแหน่งของจุดปลายที่แตกต่างกันออกไป ในการตรวจหาจุดปลายทำได้โดยการทำขอบโครงร่างบางดังแสดงในหัวข้อที่ 2.2.2 แล้วจุดภาพที่มีผลรวมของจุดภาพรอบข้างเป็น 1 จะเป็นจุดปลาย

ในการพิจารณาจุดปลายมี 2 สิ่งที่ต้องให้ความสนใจคือ ทิศทางของจุดปลาย และตำแหน่งของจุดปลาย ในการพิจารณาตำแหน่งของจุดปลายจะทำการหาตำแหน่งเทียบกับบริเวณอ้างอิงทั้ง 9 บริเวณดังแสดงในหัวข้อ 3.2.2 ในการพิจารณาทิศทางของจุดปลายทำได้โดยการคำนวณหามุมของจุดปลายเทียบกับจุดถัดที่เลื่อนเข้ามา ดังแสดงในรูปที่ 3.7



รูปที่ 3.7 การวัดมุมของจุดปลาย

การบ่งบอกทิศทางโดยใช้ตัวแปรเชิงภาษา 4 ตัวคือ “บน” “ขวา” “ล่าง” “ซ้าย” มีฟังก์ชันความเป็นสมาชิกดังสมการ

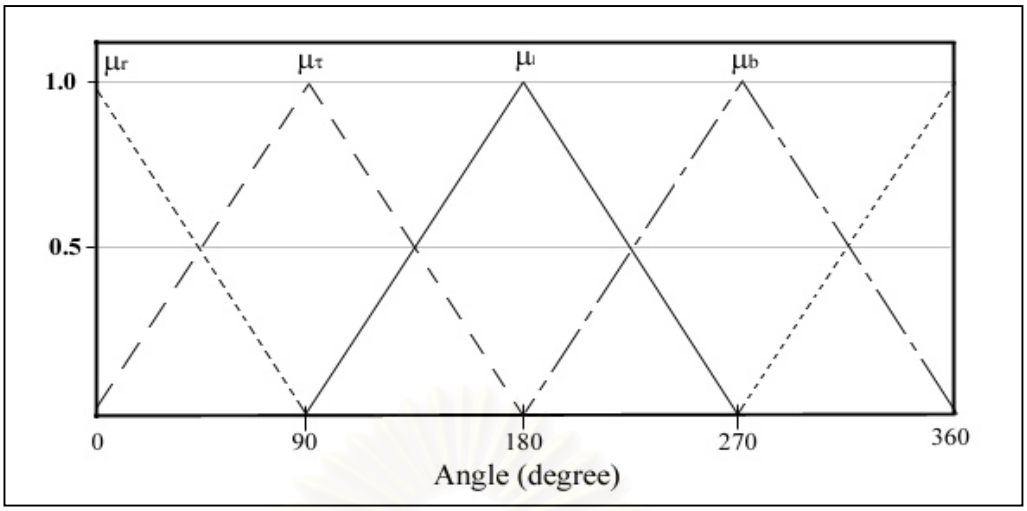
$$\mu_U = \begin{cases} 0 & ; \ \theta < 0 \\ 1 - \left| \dfrac{\theta - 90}{90} \right| & ; \ \theta > 0 \end{cases} \qquad (3.6ก)$$

$$\mu_D = \begin{cases} 1 - \left| \dfrac{\theta + 90}{90} \right| & ; \ \theta < 0 \\ 0 & ; \ \theta > 0 \end{cases} \qquad (3.6ข)$$

$$\mu_L = \begin{cases} 1 - \left| \dfrac{\theta}{90} \right| & ; \ |\theta| < 90 \\ 0 & ; \ |\theta| > 90 \end{cases} \qquad (3.6ค)$$

$$\mu_R = \begin{cases} 0 & ; \ |\theta| < 90 \\ 1 - \dfrac{180 - |\theta|}{90} & ; \ |\theta| > 90 \end{cases} \qquad (3.6ง)$$

โดย θ คือค่ามุมของจุดปลายเทียบกับแกนนอน และ $\mu_U, \mu_D, \mu_R, \mu_L$ คือค่าความเป็นสมาชิกของทิศทางด้านบน ,ล่าง ,ขวา และซ้ายตามลำดับ



รูปที่ 3.8 ฟังก์ชันความเป็นสมาชิกแสดงลักษณะมุมของจุดปลาย

ในการพิจารณาค่าจุดปลายของตัวอักษรจะคำนวณได้จากสมการ

$$\mu_{EP} = \mu_{Direction} \cdot \left(w_x \mu_x + w_y \mu_y \right) \quad (3.7)$$

โดย $\mu_{Direction}$ เป็นค่าความเป็นสมาชิกของทิศทางที่ต้องการ w_x, w_y เป็นค่าถ่วงน้ำหนักของตำแหน่งที่เกิดหัวของตัวอักษร ค่าถ่วงน้ำหนักมีค่ามากแสดงว่าหัวของตัวอักษรมีความคลาดเคลื่อนเชิงตำแหน่งมีค่าน้อยซึ่งหาได้จากสมการ (3.5) และ (3.6) μ_x, μ_y คือค่าความเป็นสมาชิกในเชิงตำแหน่งในแนวแกน x และ แนวแกน y

3.3.1.5 **การตรวจหารอยหยัก** (Curl Feature)

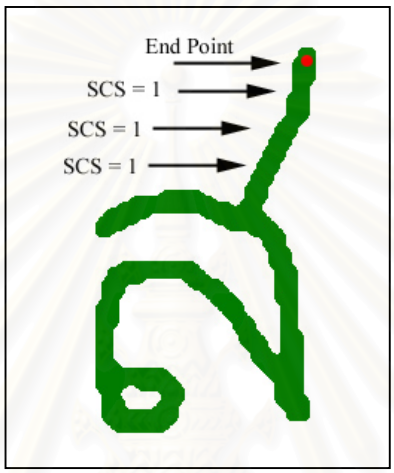
ในภาษาไทยรอยหยักเป็นความแตกต่างที่สำคัญระหว่างตัวอักษรบางคู่เช่น “ด" และ "ต” ซึ่งการหารอยหยักทำได้โดยการวิเคราะห์ขอบภาพ (Profile Analysis) แล้วพิจารณาหาค่าระยะห่างสูงสุดระหว่างขอบภาพกับกรอบภาพอ้างอิง ในแต่ละทิศทางดังแสดงในรูปที่ 3.10

รูปที่ 3.9 การพิจารณารอยหยักจากการวิเคราะห์การเปลี่ยนแปลงของขอบภาพ

การหาค่าสูงสุดของระยะห่างขอบภาพ จะค้นหาด้านละ 2 ค่าจากการนำจุดที่เกิดรอยหยักในแต่ละด้านนั้นมาหาค่าความเป็นสมาชิกเทียบกับบริเวณอ้างอิงดังแสดงในหัวข้อ 2.2.5

3.3.1.6 **การตรวจหาหาง** (Tail Feature)

จากการทดสอบกับฐานข้อมูลตัวอักษรภาษาไทย พบว่าความยาวหางของตัวอักษรมักจะยาวไม่ต่ำกว่า 10 จุดภาพ ดังนั้นจึงทำการตรวจสอบหางของตัวอักษรโดยทำการลากเส้นตัดตามแนวนอนจำนวน 10 จุดภาพบนสุดที่ติดกับขอบด้านบนของตัวอักษรแล้วพบลายเส้นข้อมูลเพียง 1 เส้นเท่านั้นดังแสดงในรูปที่ 3.11



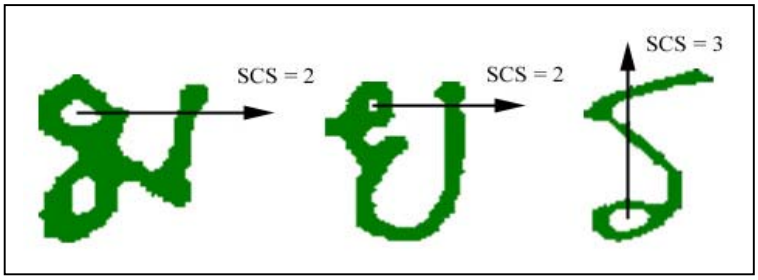
รูปที่ 3.10 การตรวจหาหางของตัวอักษร

ฟังก์ชันความเป็นสมาชิกของหางของตัวอักษรเป็นดังนี้

$$\mu_{Tail} = \begin{cases} 0 & ; \text{not found} \\ 1 & ; \text{found} \end{cases} \qquad (3.8)$$

3.3.1.7 **การพิจารณาลำดับสายการลากผ่าน** (Stroke Changing Sequence Feature)

ได้นำเอาวิธีการประยุกต์ลำดับสายการลากผ่านที่ ประเสริฐ ฉอเรืองวิวัฒน์ (1997) ได้เสนอไว้ ให้ทำการพิจารณาสายการลากผ่านใน 4 ทิศทาง โดยมีจุดเริ่มจากหัวของตัวอักษรมาปรับแก้ โดยทำการเลือกเฉพาะทิศทางที่เหมาะสม และมีจุดเริ่มที่วงรอบของตัวอักษรบางบริเวณเท่านั้น และการพิจารณาสายการลากผ่านนี้จะใช้เพื่อช่วยในการรู้จำในตัวอักษรบางตัวเท่านั้นเพื่อบ่งบอกความแตกต่างของการลากลายเส้นในบางทิศทางเพื่อแยกอักษรบางตัวอักษรที่แยกได้ยากด้วยคุณลักษณะสำคัญที่กล่าวมาแล้ว วิธีดังแสดงในรูปที่ 3.12



รูปที่ 3.11 การพิจารณาลำดับสายการลากผ่านของตัวอักษรบางตัว

ฟังก์ชันความเป็นสมาชิกของสายการลากผ่านเป็นดังนี้

$$\mu_{SCS} = \begin{cases} 0 & ; \text{not found} \\ 1 & ; \text{found} \end{cases} \qquad (3.9)$$

3.3.2 **การจำแนกตัวอักษร และการตัดสินใจ** (Classification and Decision Making)

เพื่อให้มีความยืดหยุ่นในการตัดสินใจมากยิ่งขึ้นจึงได้นำฟัซซีโลจิกเข้ามาประยุกต์ใช้ร่วมกับการรู้จำด้วยวิธีซินแทคติกที่อาศัยคุณลักษณะบ่งความต่าง ซึ่งการรู้จำจะแบ่งกลุ่มตัวอักษรโดยใช้เส้นบรรทัดด้านบน และเส้นบรรทัดด้านล่างตามที่เสนอไว้ในหัวข้อที่ 2.2.6 ก่อน จากนั้นจะทำการรู้จำตัวอักษรด้วยวิธีซินแทคติค

คำในภาษาไทยสามารถแบ่งตามระดับของตัวอักษรในบรรทัดได้ 3 ระดับดังได้กล่าวมาแล้วในหัวข้อที่ 2.2.6 ซึ่งแบ่งเป็นกลุ่มตามระดับของบรรทัดได้ดังตารางที่ 3.1

ตารางที่ 3.1 กลุ่มของอักษรตัวเดี่ยวแบ่งตามระดับของเส้นบรรทัด

ระดับของตัวอักษร	ตัวอักษร
ระดับบน	◌ิ ◌ี ◌ึ ◌ื ◌์ ◌ั ◌่ ◌้ ◌๊ ◌๋ ◌็
ระดับกลาง	ก ข ฃ ค ฅ 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 ใ ไ โ
ระดับล่าง	◌ุ ◌ู

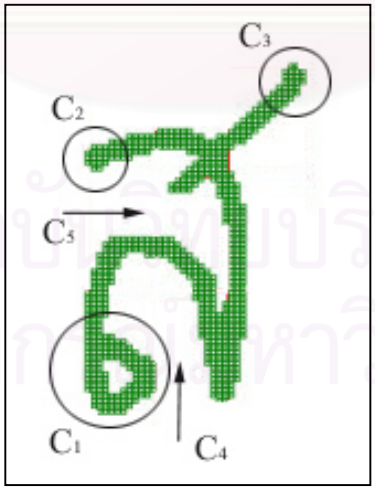
ในการพิจารณาระดับบรรทัดทำได้โดยใช้เส้นบรรทัดด้านบน (Upper Baseline) และเส้นบรรทัดด้านล่าง(Lower Baseline) มาใช้พิจารณาดังแสดงในรูปที่ 2.10 โดย

- ตัวอักษรที่ถือว่าอยู่ในกลุ่มของตัวอักษรที่อยู่เหนือเส้นบรรทัดด้านบนจะต้องมีจุดศูนย์กลางของตัวอักษรอยู่เหนือเส้นบรรทัดด้านบน และมีขอบล่างของตัวอักษรอยู่เหนือเส้นกึ่งกลางบรรทัด
- ตัวอักษรที่ถือว่าอยู่ในกลุ่มของตัวอักษรที่อยู่ใต้เส้นบรรทัดด้านล่างจะต้องมีจุดศูนย์กลางของตัวอักษรอยู่ใต้เส้นบรรทัดด้านล่าง และมีขอบบนของตัวอักษรอยู่ใต้เส้นกึ่งกลางบรรทัด

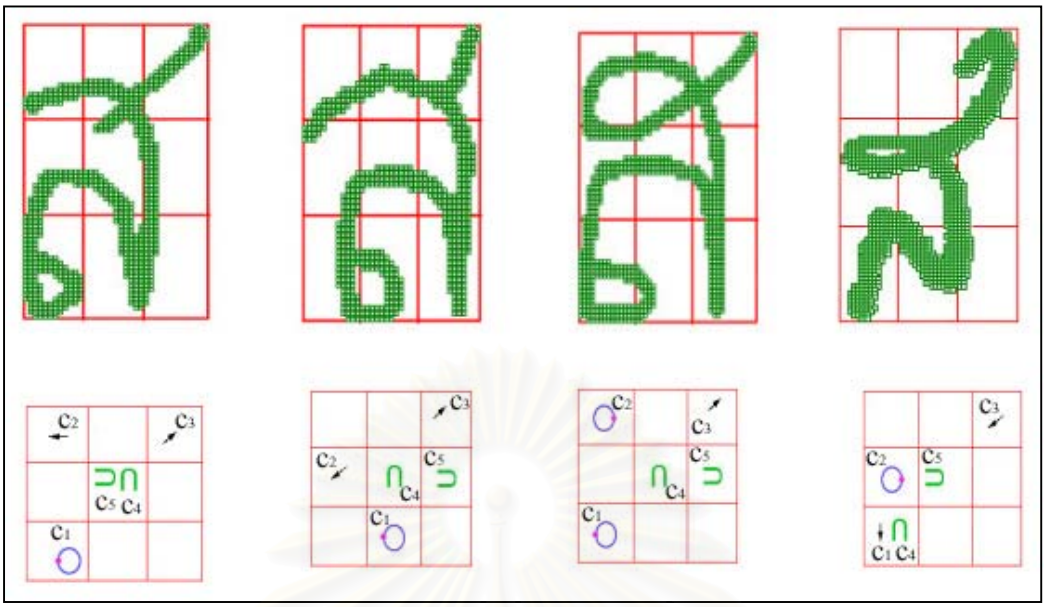
3.3.2.1 **การสร้างชุดกฎทางฟัซซี**

การสร้างชุดกฎทางฟัซซีทำได้โดยพิจารณาจากโครงสร้างของตัวอักษร และผลลัพธ์ที่ได้จากการหาคุณลักษณะ และคำนวณค่าความเป็นสมาชิกของคุณลักษณะเหล่านั้น ในตัวอักษรแต่ละตัวจะมีลักษณะที่เป็นโครงสร้างเฉพาะที่ไม่ซ้ำกัน เช่น “ด” จะมีหัวอยู่ตรงกลางตัวหันไปทางด้านซ้าย มีจุดปลายทางด้านขวาล่างชี้ลงด้านล่าง และด้านบนไม่มีรอยหยัก หรือ “ม” มีหัวทางด้านมุมซ้ายบนหันไปทางซ้าย และหัวทางมุมซ้ายล่างหันไปทางซ้าย และมีจุดปลายทางมุมขวาด้านบนชี้ขึ้นด้านบน จากนั้นนำค่าความเป็นสมาชิกของคุณลักษณะมาพิจารณาความเบี่ยงเบนและความหลากหลายของรูปแบบของลายมือเขียน

ในการพิจารณาตัวอักษรแต่ละตัวจะทำการหาองค์ประกอบเฉพาะ (c_i) ที่ประกอบกันขึ้นเป็นตัวอักษร ดังแสดงในรูปที่ 3.13 “ส” ประกอบด้วยส่วนประกอบที่สำคัญซึ่งจะพบองค์ประกอบเหล่านี้เสมอ ไม่ว่าจะเขียน “ส” ในรูปแบบใดก็ตาม



รูปที่ 3.12 แสดงองค์ประกอบที่พบในตัวอักษร “ส”



รูปที่ 3.14 แสดงโครงสร้างของ “ส” ที่แตกต่างกันและ รูปแบบของลักษณะสำคัญที่ตรวจสอบได้

จากรูปที่ 3.14 พบว่าแต่ละองค์ประกอบ “ส” มีรูปแบบต่างๆ กันออกไปตามแต่ลักษณะการเขียนดังนั้นตัวดำเนินการทางฟัซซีจึงถูกใช้เพื่อรวมรูปแบบที่คลาดเคลื่อนไปเหล่านั้นเข้าไว้ในองค์ประกอบเดียวกันโดยใช้ตัวดำเนินการ “หรือ” ($\vee$) ดังสมการ

$$C_j = \mu_1 \vee \mu_2 \vee ... \vee \mu_n \qquad (3.10)$$

โดย C_i คือองค์ประกอบของตัวอักษร และ $\mu_1, \mu_2, ..., \mu_n$ คือคุณลักษณะสำคัญขององค์ประกอบ C_i ในรูปแบบต่าง ๆ

กฎทางฟัซซี (R_i) แต่ละกฎจะแทนโครงสร้างของตัวอักษรหนึ่งตัวซึ่งเป็นการรวมกันของคุณลักษณะสำคัญที่ประกอบกันในตัวอักษร สามารถแสดงได้ดังสมการ

$$R_i = \frac{\sum_{i=1}^{n} C_i}{n} \qquad (3.11)$$

โดย R_i คือกฎทางฟัซซีที่บ่งบอกถึงลักษณะตัวอักษรแต่ละรูปแบบ

กฎทางฟัซซีแต่ละกฎจะแทนโครงสร้างของตัวอักษรหนึ่งตัว ในบางตัวอักษรที่มีโครงสร้างของลายมือหลายรูปแบบต้องใช้กฎทางฟัซซีหลายๆ กฎเพื่อแทนรูปแบบของโครงสร้างเหล่านั้นดังตัวอย่าง

ตัวอย่างกฎ “ส” ในรูปแบบที่ 1 และ 3 ของรูป 3.14

Rule1 “ส” : Loop # b-l & (EP # t-l # Down || Left) & (EP # t-l # Up || Right)

ความหมาย: ตัวอักษร “ส” มีวงรอบด้านซ้าย-ล่าง หันไปด้านขวาและมีจุดปลายชี้ไปทางด้านซ้ายหรือด้านล่างบริเวณซ้าย-บน และมีจุดปลายชี้ไปทางด้านบน หรือทางด้านขวา บริเวณขวา-บน

Rule2 “ส” : Loop # b-l & (Loop # m-c) & (EP # t-l # Up || Right)

ความหมาย: ตัวอักษร “ส” มี วงรอบด้านซ้าย-ล่าง หันไปด้านขวาและ มีวงรอบบริเวณกลาง-กลาง หันไปด้านซ้าย และมีจุดปลายชี้ไปทางด้านบน หรือทางด้านขวา บริเวณขวา-บน

กฎทางฟัซซีของระบบรู้จำตัวอักษรไทยทั้งหมดจะแสดงในภาคผนวก ข

3.3.2.2 **การตัดสินใจ** (Decision Making)

ระบบจะนำค่าความเป็นสมาชิกของคุณลักษณะต่างๆ มาคำนวณในกฎทางฟัซซีแต่ละกฎ กฎทางฟัซซีแต่ละกฎจะให้ค่าความคล้ายกับรูปแบบของตัวอักษรที่กฎ(R_i)ข้อนั้นแสดงอยู่ รูปแบบของตัวอักษรที่มีค่าความคล้ายสูงที่สุดจะถือว่าเป็นผลของการรู้จำ

$$\text{Answer} = \text{Max}\left(R_i\right) \tag{3.12}$$

3.4 **กระบวนการภายหลัง** (Post processing)

หลังจากใช้ระบบการรู้จำตัวอักษรแต่ละตัวจนครบทุกตัว ระบบจะนำชุดคำตอบมาตรวจสอบกับฐานข้อมูลพจนานุกรมเพื่อตรวจสอบความถูกต้องในการรู้จำ และแก้ไขความผิดพลาดที่เกิดจากการรู้จำในแต่ละตัวอักษร กระบวนการค้นหาคำจากพจนานุกรมมีผู้ออกแบบระบบเพื่อให้สามารถค้นหาและตรวจสอบคำได้อย่างมีระเบียบและมีความรวดเร็วในการหา และสามารถใช้กับฐานข้อมูลคำที่มีขนาดใหญ่ได้ดี แต่มีการค้นหาคำในทิศทางเดียว และแก้ไขความผิดพลาดของตัวอักษรในคำได้ในจำนวนอักษรที่จำกัด เช่นการจัดเก็บคำด้วยวิธีตรัยโยง (Linked Tried) มีการค้นหาได้จากด้านหน้าไปด้านหลัง และ ไม่สามารถแก้ไขตัวอักษรที่ผิดพลาดหลายๆ ตัวได้ และโอกาสที่จะตรวจสอบผิดพลาดขึ้นกับตำแหน่งของอักษรที่ผิด แต่ในระบบรู้จำตัวอักษรเขียน การรู้จำผิดพลาดสามารถเกิดขึ้นได้ทุกตำแหน่ง และมักเกิดความผิดพลาดพร้อมกันหลายๆ ตำแหน่ง

เพื่อให้การตรวจสอบคำที่รู้จำผิด และมีการแก้ไขอย่างมีประสิทธิภาพ จึงได้นำเสนอวิธีการตรวจสอบคำตอบกับระบบรู้จำ โดยการนำผลการรู้จำ 5 คำตอบสูงสุดของแต่ละตัวอักษร มารวมกันเป็นกลุ่มคำตอบ โดยแยกแต่ละตัวอักษรออกจากกัน แล้วทำการเทียบผลการรู้จำกับคำในพจนานุกรมทีละตัวอักษร คำที่มีจำนวนตัวอักษรตรงกับผลการรู้จำมากที่สุดจะถูกเลือกเป็นคำตอบ แต่เนื่องจากการตรวจสอบพจนานุกรมแบบนี้ใช้เวลาในการตรวจสอบมากจึงเสนอวิธีการสองวิธีเพื่อลดเวลาในการตรวจสอบดังนี้

1.4.1 กระบวนการภายหลังที่แบ่งคำตามความยาวของคำ

เนื่องจากคำที่ใช้ทดสอบกับระบบเป็นระบบปิดที่เป็นคำที่ใช้ในงานเฉพาะด้าน ทำให้สามารถแบ่งกลุ่มคำตามจำนวนตัวอักษรได้ และเลือกตรวจสอบกับคำเฉพาะที่มีความยาวของคำเท่ากัน

3.4.2 กระบวนการภายหลังที่ใช้ความเหมือนกันของคำ

เนื่องจากกลุ่มคำที่ใช้ในการทดสอบมีส่วนที่เหมือนกันอยู่ เช่นในคำ “ปทุมธานี สุราษฎร์ธานี อุดรธานี อุทัยธานี อุบลราชธานี” มี “ธานี” อยู่ตอนท้ายของคำเหมือนกันทุกคำจึงอาศัยความเหมือนกันนี้จัดคำเป็นกลุ่ม ๆ แล้วทำการตรวจสอบกับคำตอบที่ได้จากการรู้จำว่ามีคำย่อย ๆ เหล่านี้อยู่หรือไม่ แล้วจึงตรวจสอบคำในกลุ่มที่พบคำย่อยต่อไป

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4

ขั้นตอนการทดลอง ผลการทดลอง และวิเคราะห์ผลการทดลอง

ในบทนี้จะกล่าวถึงวิธีการทดลอง ผลการทดลอง รวมถึงการวิเคราะห์ผลที่ได้จากการทดลอง

4.1 อุปกรณ์ที่ใช้ทำการทดสอบ

1. เครื่องคอมพิวเตอร์ส่วนบุคคล ใช้หน่วยประมวลผลกลางรุ่น Pentium III ความเร็ว 600 เมกกะเฮิร์ท(MHz) หน่วยความจำหลัก 128 เมกกะไบต์(MBytes)
2. โปรแกรมรู้จำลายมือเขียนที่พัฒนาด้วยภาษา C++

4.2 แหล่งที่มาของข้อมูลที่ใช้สร้างและทดสอบระบบ

- ข้อมูลที่เป็นอักษรตัวเดี่ยวได้จากการเขียนข้อมูลด้วยปากกาบนกระดานอิเล็กทรอนิกส์ ในกรอบข้อมูลที่กำหนดให้
- ข้อมูลที่เป็นคำในภาษาไทย ได้จากการเขียนคำในกระดาษที่เตรียมไว้ ดังแสดงในรูปที่ 4.1 ซึ่งแบ่งไว้เป็นช่องๆ และเขียนด้วยปากกาสีดำที่มีหัวปากกาขนาด 0.1 มิลลิเมตรและ 0.3 มิลลิเมตร โดยเขียนชื่อจังหวัดในประเทศไทย 76 จังหวัด โดยแบ่งตามความคล้ายกันของคำ ซึ่งสามารถแบ่งได้เป็น 4 ชุดดังนี้

ชุดที่ 1 ประกอบด้วย ตรัง สตูล ชุมพร กรุงเทพมหานคร สมุทรสงคราม สมุทรสาคร สมุทรปราการ ร้อยเอ็ด พิจิตร ลำพูน กระบี่ นครปฐม แม่ฮ่องสอน ฉะเชิงเทรา นราธิวาส นครสวรรค์ สระแก้ว ลำปาง ยโสธร ชัยนาท

ชุดที่ 2 ประกอบด้วย นนทบุรี ลพบุรี อุทัยธานี จันทบุรี พัทลุง ปทุมธานี ภูเก็ต หนองบัวลำภู อุดรธานี ชลบุรี อุตรดิตถ์ กาญจนบุรี อำนาจเจริญ เพชรบูรณ์ สุรินทร์ ราชบุรี สระบุรี สุพรรณบุรี เชียงราย นครราชสีมา

ชุดที่ 3 ประกอบด้วย เพชรบุรี อุบลราชธานี สุโขทัย พิษณุโลก ชัยภูมิ ปราจีนบุรี อุบลราชธานี สุราษฎร์ธานี ปัตตานี นครศรีธรรมราช สิงห์บุรี เชียงใหม่ กาฬสินธุ์ บุรีรัมย์ ประจวบคีรีขันธ์ รัฐธรรมนูญ อาสาฬหบูชา พฤหัสบดี เปาโลเมโมเรียล มิชชัน

ชุดที่ 4 ประกอบด้วย เลย ตาก ตราด นครนายก นครพนม หนองบัวลำภู มหาสารคาม สกลนคร สงขลา พะเยา ระนอง ระยอง ยะลา น่าน อ่างทอง ขอนแก่น แพร่ พังงา กำแพงเพชร อยุธยา มุกดาหาร

แล้วนำมาสแกนด้วยความละเอียดขนาด 300 จุดภาพต่อนิ้ว (dpi) จัดเก็บเป็นข้อมูลภาพ 2 ระดับสีเทาตามมาตรฐานแบบ Portable Gray Map (pgm)

เนื่องจากมีการเก็บข้อมูลที่เป็นคำทำให้รูปแบบของตัวอักษรที่ได้มีความหลากหลาย โดยบางตัวอักษรมีการเขียนติดกัน ทั้งที่มีการยกปากกา และไม่มีการยกปากกา ทำให้ได้ข้อมูลคำที่มีการเขียนใกล้เคียงกับความเป็นจริงกับการเขียนที่เกิดขึ้นในชีวิตประจำวัน

รูปที่ 4.1 แบบสำหรับเก็บข้อมูลคำภาษาไทย

4.3 การจัดกลุ่มข้อมูลเพื่อสร้าง และทดสอบระบบ

ข้อมูลพยัญชนะภาษาไทยตัวเดี่ยว 44 แบบ (ก ถึง ฮ) แบบอักษรละ 100 ตัว แบ่งเป็น 2 กลุ่ม สำหรับสร้างและวิเคราะห์ระบบ 60 รูปแบบ และสำหรับทดสอบระบบ 40 รูปแบบ

ข้อมูลสระและวรรณยุกต์ภาษาไทยตัวเดี่ยว แบบอักษรละ 40 แบบ แบ่งเป็น 2 กลุ่มสำหรับสร้างและวิเคราะห์ระบบ 20 รูปแบบ และสำหรับทดสอบระบบ 20 รูปแบบ

ข้อมูลคำภาษาไทยเป็นชื่อจังหวัดจำนวน 76 จังหวัด จังหวัดละ 40 คำ ประกอบด้วย แบ่งเป็น 2 กลุ่ม สำหรับสร้างและวิเคราะห์ระบบ 20 รูปแบบ และสำหรับทดสอบระบบ 20 รูปแบบ

เนื่องจากโครงสร้างของตัวพยัญชนะไทย มีความซับซ้อนและมีรูปแบบที่เกิดขึ้นในการเขียนมากกว่าสระและวรรณยุกต์จึงใช้จำนวนตัวอักษรในการสร้างและทดสอบระบบจำนวนมากกว่าเพื่อจะได้เรียนรู้รูปแบบได้ครบถ้วนยิ่งขึ้น

4.4 วิเคราะห์ผลกระทบเนื่องจากปากกาที่ใช้เขียนข้อมูล

ในการเก็บข้อมูลลายมือเขียนเพื่อใช้ในการสร้างและทดสอบระบบรู้จำนี้ ได้ให้ผู้เขียนใช้ปากกาหมึกสีดำที่มีลายเส้น 2 ขนาด คือปากกาที่มีหัวขนาด 0.3 มิลลิเมตรและ ปากกาที่มีหัวขนาด 0.1 มิลลิเมตร ซึ่งหลังจากที่นำลายมือผ่านการสแกน และเก็บตัวอย่างในรูปแบบไฟล์ดิจิทัลด้วยค่าผ่านระดับ (Threshold) ที่เหมาะสม ซึ่งได้จากการพิจารณาภาพทีละภาพ เนื่องจากกระดาษแต่ละแผ่นที่ใช้เขียน และความเข้มของเส้นกรอบในแบบสำหรับเก็บข้อมูลมีความแตกต่างกันในแต่ละแผ่น ซึ่งผลที่ได้จากการสแกนมีลักษณะดังนี้

1. ข้อมูลที่เขียนด้วยปากกาขนาด 0.1 มิลลิเมตร ลายเส้นมีขนาดเล็ก มีขนาดความกว้างของลายเส้นประมาณ 1 ถึง 4 จุดภาพ ภาพถูกรบกวนจากสัญญาณรบกวนได้ง่าย ทำให้ขอบของตัวอักษรขรุขระ ไม่เรียบ และลายเส้นมักไม่ต่อเนื่องเมื่อเขียนตัวอักษรขนาดใหญ่ แต่รายละเอียดของตัวอักษรคงเหลืออยู่มาก และชัดเจน
2. ข้อมูลที่เขียนด้วยปากกาขนาด 0.3 มิลลิเมตร ลายเส้นมีขนาดปานกลาง มีขนาดความกว้างของลายเส้นประมาณ 3 ถึง 6 จุดภาพ สัญญาณรบกวนมีผลต่อข้อมูลภาพน้อย แต่ลายเส้นมักติดกัน ลายละเอียดของตัวอักษรบางอย่างหายไป เนื่องจากลายเส้นมีความกว้างมากเทียบกับขนาดตัวอักษร

(ก)

(ข)

รูปที่ 4.2 ตัวอย่างภาพที่สแกนจากการเขียนด้วยปากกาขนาดต่างกัน

ก) ปากกาขนาด 0.1 มิลลิเมตร

ข) ปากกาขนาด 0.3 มิลลิเมตร

เนื่องจากผลของขนาดปากกาที่ต่างกัน ทำให้กระบวนการก่อนหน้าจำเป็นมากเมื่อใช้ปากกาที่มีขนาด 0.1 มิลลิเมตร โดยเฉพาะในขั้นตอนการลดสัญญาณรบกวน แต่จะให้ผลการรู้จำที่ดี ส่วนกรณีเมื่อใช้ปากกาที่มีขนาด 0.3 มิลลิเมตร ผลกระทบจากสัญญาณรบกวนจะน้อยกว่า แต่ทำให้การรู้จำทำได้ยากเนื่องจากรายละเอียดของตัวอักษรหายไป ทั้งยังไม่สามารถแก้ไขได้ด้วยขั้นตอนของกระบวนการก่อนหน้า

4.5 วิเคราะห์ผลจากการฝึกฝนระบบรู้จำอักษรตัวเดี่ยว

4.5.1 การฝึกฝนพยัญชนะ และสระ

การฝึกฝนระบบรู้จำเพื่อรู้จำพยัญชนะ ตัวเดี่ยวในภาษาไทย เนื่องจากตัวอักษรมีรูปแบบที่หลากหลายจึงใช้ข้อมูลเพื่อวิเคราะห์ระบบ 60 แบบ จากผู้เขียนจำนวน 12 คน และในการฝึกฝนระบบรู้จำสระ และวรรณยุกต์ตัวเดี่ยวใช้ข้อมูล 20 แบบ จากผู้เขียนจำนวน 10 คน ทั้งพยัญชนะ สระและวรรณยุกต์ตัวเดี่ยวมีผลการรู้จำดังตารางที่ 4.1

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.1 ผลการทดสอบระบบเพื่อวิเคราะห์ ด้วยข้อมูลสำหรับการฝึกฝนระบบ

	ก	ข	ฃ	ค	ฅ	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	ถูก (%)	ผิด(%)
ก	58	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	96.667	3.3333
ข	0	50	2	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	83.333	16.667
ฃ	0	1	57	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95	5
ค	0	0	0	57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95	5
ฅ	0	0	0	1	56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	93.333	6.6667
ง	0	0	0	0	0	58	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	96.667	3.3333
จ	0	0	0	0	0	0	56	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	93.333	6.6667
ฉ	0	0	0	0	0	0	3	51	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	85	15
ช	0	5	0	0	0	0	0	0	43	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	2	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	71.667	28.333
ซ	0	0	0	0	0	0	0	0	1	59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98.333	1.6667
ฌ	0	0	0	0	0	0	0	0	0	0	58	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	96.667	3.3333
ญ	0	0	0	0	0	0	0	0	0	0	2	58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	96.667	3.3333
ฎ	0	0	0	0	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
ฏ	0	0	0	0	0	0	0	0	0	0	0	0	2	57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	95	5
ฐ	0	0	0	0	0	0	0	0	0	0	0	0	0	4	54	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	90	10
ฑ	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	53	2	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	88.333	11.667
ฒ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
ณ	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	56	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	93.333	6.6667
ด	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	56	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	93.333	6.6667
ต	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98.333	1.6667
ถ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
ท	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	58	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	96.667	3.3333
ธ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	57	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	95	5
น	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	53	0	0	0	0	3	0	0	2	0	0	0	0	0	0	0	0	0	0	0	88.333	11.667
บ	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	54	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	90	10
ป	0	0	0	0	0	0	0	0	6	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	86.667	13.333
ผ	0	0	2	0	0	0	3	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	2	0	47	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	78.333	21.667
ฝ	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	4	43	0	3	0	0	0	0	0	0	0	0	0	0	0	0	3	71.667	28.333
พ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	2	0	53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	88.333	11.667
ฟ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	59	0	0	0	0	0	0	0	0	0	0	0	0	0	98.333	1.6667
ภ	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	58	0	0	0	0	0	0	0	0	0	0	0	0	96.667	3.3333
ม	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	100	0
ย	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0	83.333	16.667
ร	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	4	0	0	0	0	1	0	0	0	0	0	0	0	51	0	2	0	0	0	0	0	0	0	85	15
ล	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	1	0	0	0	0	1	0	0	0	47	0	0	0	8	0	0	0	0	78.333	21.667
ว	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	54	0	0	0	0	3	0	0	90	10
ศ	0	0	0	0	2	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	52	0	4	0	0	0	0	86.667	13.333
ษ	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	2	0	0	0	0	0	0	0	0	51	0	0	0	0	0	85	15
ส	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	3	0	53	0	0	0	0	88.333	11.667
ห	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	10	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	47	0	0	0	78.333	21.667
ฬ	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	57	0	4	95	5
อ	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	50	0	83.333	16.667
ฮ	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	7	46	76.667	23.333
																																												อัตราการรู้จำเฉลี่ย			90.114	9.8864

ตารางที่ 4.2 ผลการทดสอบการรู้จำสระ และวรรณยุกต์ เพื่อวิเคราะห์ ด้วยข้อมูลสำหรับการฝึกฝนระบบ

	า	ิ	ี	ึ	ื	ุ	ู	เ	็	่	้	๊	๋	ั	์	ๆ	ฯ	ๅ	+	ก	จ	ถูก(%)	ผิด(%)
า	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	90	10
ิ	0	13	2	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	65	35
ี	0	1	14	0	3	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	70	30
ึ	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
ื	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
ุ	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
ู	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
เ	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
็	0	3	0	0	0	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0	85	15
่	0	0	0	0	0	0	0	0	0	18	0	2	0	0	0	0	0	0	0	0	0	90	10
้	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	100	0
๊	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	100	0
๋	0	0	1	0	0	0	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	95	5
ั	0	0	1	0	0	0	0	0	0	0	0	0	2	17	0	0	0	0	0	0	0	85	15
์	0	0	1	0	0	0	0	0	0	0	0	0	2	0	17	0	0	0	0	0	0	85	15
ๆ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	100	0
ฯ	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	16	0	0	0	0	80	20
ๅ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	19	0	0	0	95	5
+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	100	0
																			อัตราการรู้จำเฉลี่ย			91.58	8.421

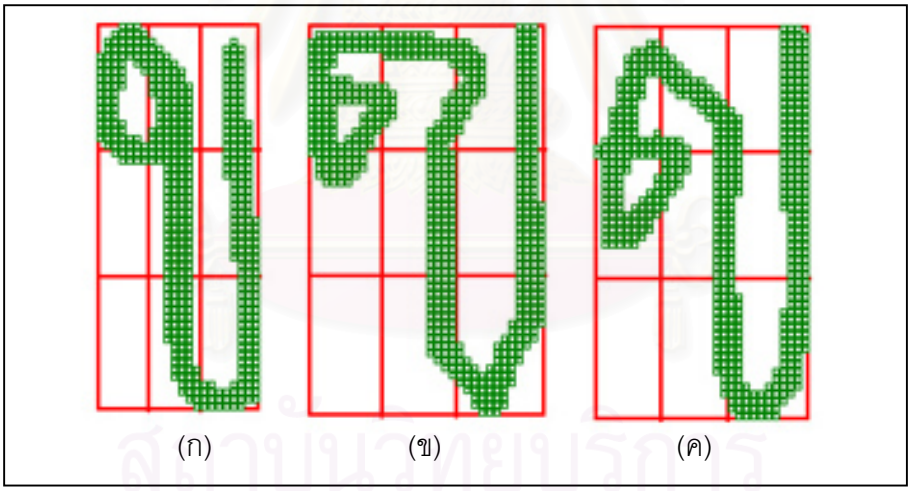
วิเคราะห์ผลการฝึกฝนระบบรู้จำตัวพยัญชนะและสระ

เนื่องจากการประยุกต์ทฤษฎีทางฟัซซีเข้ามาในระบบรู้จำแบบซินแทคติก(Syntactic) ทำให้การตัดสินใจแบบฉับพลัน (Hard Decision) ซึ่งจะตัดสินใจโดยทันทีเมื่อพบลักษณะที่บ่งบอกความแตกต่าง เป็นการตัดสินใจแบบนุ่มนวล (Soft Decision) ซึ่งพิจารณาองค์ประกอบรอบข้าง แล้วตัดสินใจเอาตัวอักษรที่มีลักษณะเหมือนกับรูปแบบที่กำหนดไว้ในกฎทางฟัซซีมากที่สุด มาเป็นคำตอบของระบบ ซึ่งระบบเช่นนี้จะยอมให้ตัวอักษรมีรูปแบบที่ผิดเพี้ยนไปได้บ้าง

ความผิดพลาดในการรู้จำตัวอักษรตัวเดี่ยวจึงมักเกิดจากรูปแบบที่หลากหลายของตัวอักษร ซึ่งแก้ไขได้ด้วยการเพิ่มรูปแบบที่เกิดขึ้นเข้าในกฎทางฟัซซีของตัวอักษรแต่ละตัว โดยกฎทางฟัซซีของตัวอักษรแต่ละตัวมีแสดงไว้ใน ภาคผนวก ก

ความผิดพลาดเนื่องจากรูปแบบลักษณะสำคัญที่เกิดขึ้นมีความหลากหลาย

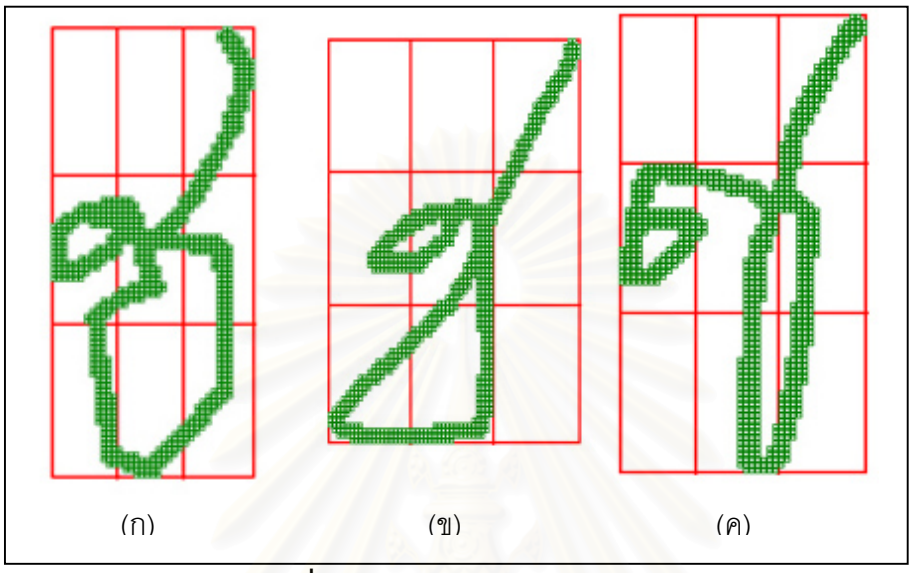
หัวเป็นลักษณะสำคัญของตัวอักษรภาษาไทย และมีรูปแบบที่หลากหลายทั้งในเชิงตำแหน่ง ทิศทาง และรูปแบบ



รูปที่ 4.3 ลักษณะหัวของตัวอักษร “ข”

จากภาพ 4.3 จะพบว่าตัวอักษร “ข” มีรูปแบบหัวหลายรูปแบบ เช่นในรูปที่ 4.3 (ก) มีหัวของตัวอักษรหันไปทางด้านซ้าย ในรูปที่ 4.3 (ข) มีหัวของตัวอักษรหันไปทางด้านขวา เมื่อพิจารณาด้วยวิธีการในหัวข้อที่ 3.3.3 และในรูปที่ 4.3 (ค) หัวของตัวอักษรอยู่ในบริเวณที่ต่างออกไปจากรูป 4.3 (ก) และ 4.3 (ข)

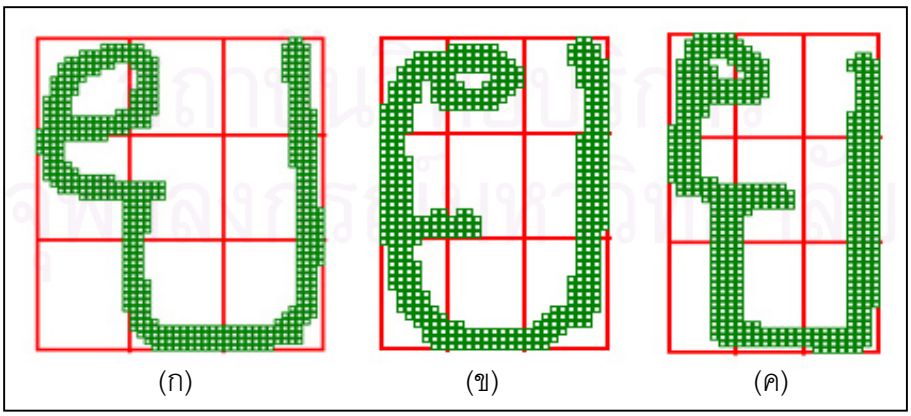
แก้ไขโดยการเพิ่มรูปแบบให้กับกฎทางฟัซซีของตัวอักษร “ข” โดยให้เพิ่มรูปแบบจากหัวของ “ข” มีทิศทางไปทางด้านซ้ายบริเวณด้านบนซ้าย เป็นหัวของ “ข” มีทิศทางไปทางด้านซ้ายหรือขวา บริเวณด้านบนซ้าย หรือหัวหันไปทางขวา บริเวณกลางซ้าย



รูปที่ 4.4 ลักษณะหัวของตัวอักษร “ข”

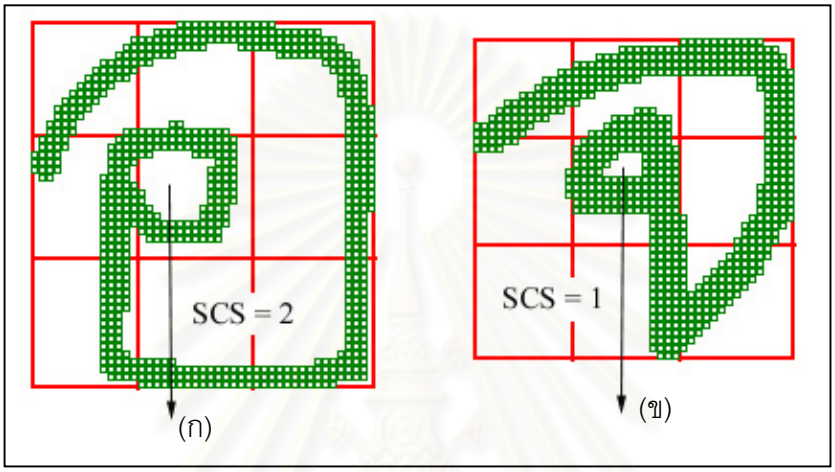
จากภาพ 4.4 จะพบว่าตัวอักษร “ข” มีหัวที่พบอยู่หลายตำแหน่ง เช่นในรูปที่ 4.4 (ก) มีหัวของตัวอักษรหันไปทางซ้ายอยู่ในบริเวณกลางซ้าย ซึ่งในรูปที่ 4.4 (ข) มีหัวของตัวอักษรในบริเวณกลางกลาง และในรูปที่ 4.4 (ค) มีหัวไปทางด้านขวา

แก้ไขโดยการเพิ่มรูปแบบให้กับกฎทางฟัซซีของ “ช” โดยให้เปลี่ยนจากหัวมีทิศทางไปทางด้านซ้ายบริเวณด้านกลางซ้าย เป็น หัวของ “ช” มีทิศทางไปทางด้านซ้าย หรือ ขวา บริเวณด้านกลางซ้าย หรือ บริเวณกลางของระดับกลาง



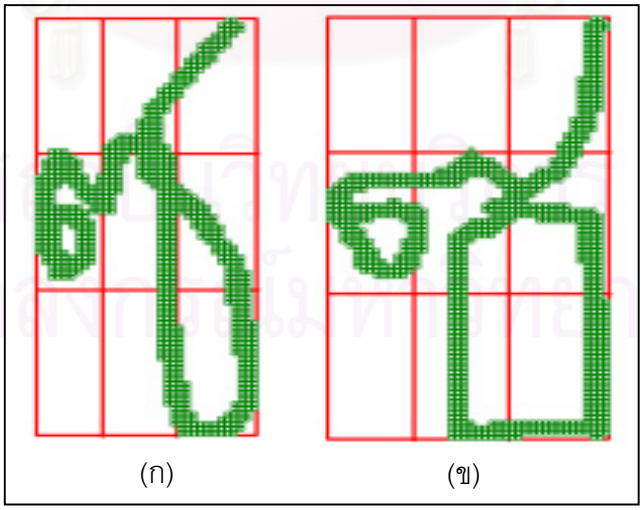
รูปที่ 4.5 ลักษณะของตัวอักษร “ย”

ตัวอักษรในรูปที่ 4.5 รู้จำผิดเป็น “ข” เนื่องจาก รอยหยักบริเวณด้านซ้ายไม่ชัดเจน แต่เกิดจุดปลายบริเวณกลางตัวอักษร สามารถแก้ไขได้โดย เพิ่มองค์ประกอบของจุดปลายบริเวณกลางตัวอักษรเข้าในกฎทางฟัซซีของ “ย” และ ใช้ตัวดำเนินการคอมพลีเมนต์ กับกฎของตัวอักษร “ข” ดังนี้ “ตัวอักษร “ข” มีหัวหันไปทางซ้าย หรือขวา ทางด้านบนซ้าย หรือกลางซ้าย มีจุดปลายชี้ด้านบน บริเวณบนขวาและมีหยักแรกจากด้านบนบริเวณล่างกลาง และ*ไม่มีจุดปลายชี้ทางขวา บริเวณกลางของระดับกลาง*”



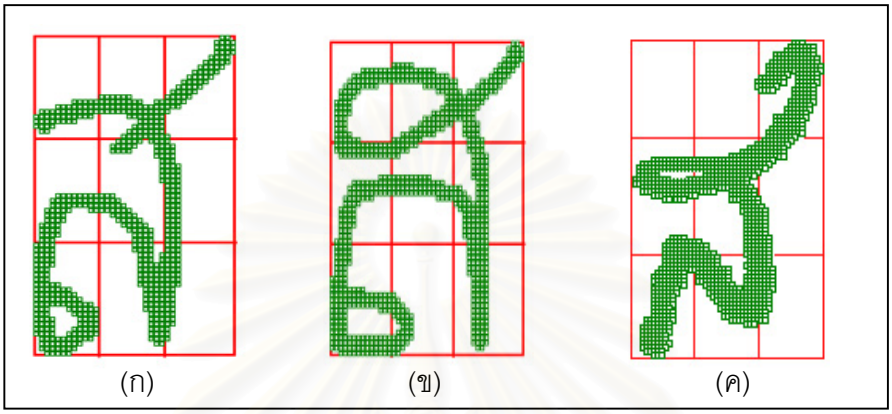
รูปที่ 4.6 ลักษณะของตัวอักษร "อ" และ “จ”

ตัวอักษรในรูปที่ 4.6 รู้จำผิดเป็น “อ” เนื่องจากมีความแตกต่างเพียงทิศทางของหัวเท่านั้น ดังนั้นจึงใช้ลำดับสายการลากผ่านมาบ่งชี้ความชัดเจนของความแตกต่างนี้โดย พิจารณาสายการลากผ่านในทิศทางลงด้านล่างโดยเริ่มจากหัวของตัวอักษร ในกรณี “จ” จะมีจำนวนการตัดของเส้นเท่ากับ 1 ส่วนกรณี “อ” จะมีจำนวนการตัดของเส้นเท่ากับ 2



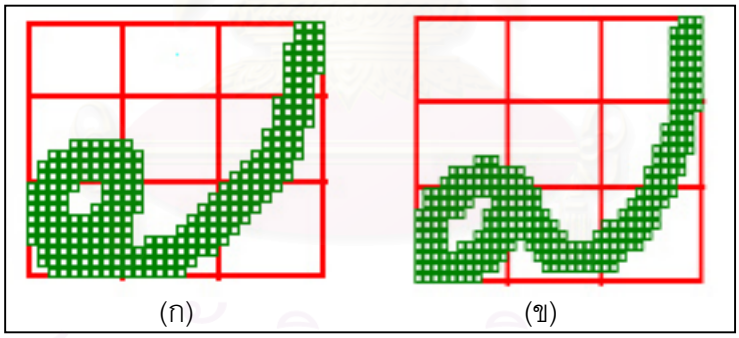
รูปที่ 4.7 ลักษณะของตัวอักษร “ซ”

ตัวอักษรในรูปที่ 4.7 รู้จำผิดเป็น “ซ” เนื่องจากรอยหยักบริเวณที่ติดกับหัวของตัวอักษร ไม่ชัดเจน ทำให้รู้จำผิดพลาด ตัวอักษรในรูปที่ 4.7(ก) บอกได้ว่าเป็น “ซ” เนื่องจากพิจารณาตำแหน่งของรอยหยัก ส่วนตัวอักษรในรูปที่ 4.7(ข) รู้จำเป็น ”ซ“ เนื่องจากตรวจไม่พบรอยหยัก แต่ตัวอักษร “ซ” มีค่าความคล้ายรองลงมาจากตัวอักษร "ซ"



รูปที่ 4.8 ลักษณะของตัวอักษร “ส”

ตัวอักษรในรูปที่ 4.8 รู้จำผิดพลาดเนื่องจากมีรูปแบบของตัวอักษรหลายรูปแบบ และมีความไม่แน่นอนของตำแหน่งคุณลักษณะสำคัญอยู่มาก ซึ่งสามารถแก้ไขได้โดยการเพิ่มกฎทางฟัซซีของตัวอักษร “ส” ในรูปแบบ 4.8 (ข) เข้าไปดังนี้

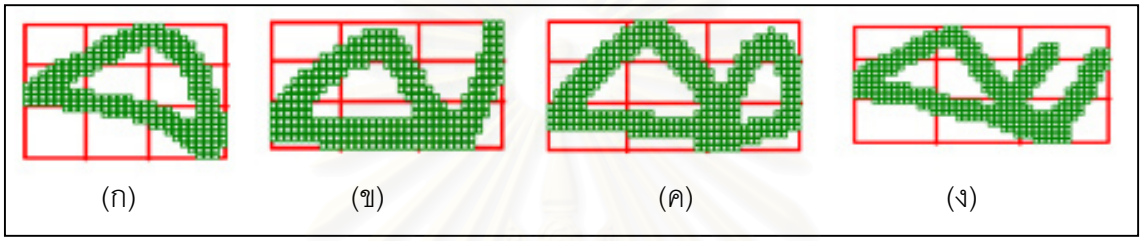


รูปที่ 4.9 ลักษณะของตัวอักษร ไม้หันอากาศ “ ั" และ ไม้โท ” ้ ”

เนื่องจากสระทั้งสองตัว อยู่ในระดับบนของบรรทัดทั้งสองตัว และทั้งสองตัวมีลักษณะทางโครงสร้างคล้ายกัน โดยในรูปที่ 4.9 (ข) ไม้โท ” ้ ” รู้จำเป็นไม้หันอากาศ “ ั “ เนื่องจากมีหัวหันไปทางซ้าย และมีหยักจากทางด้านบนเหมือนกัน ดังนั้นจึงทำการแยกสระทั้งสองโดยไม้โท ” ้ ” จะมีหยักจากทางด้านล่าง แต่ไม้หันอากาศ “ ั “ ไม่มีหยักทางด้านล่าง กฎทางฟัซซีของสระทั้งสองกำหนดได้ดังนี้

สระ ” ้ ” ไม้โทมีหัวหันไปทางซ้าย บริเวณกลางซ้าย หรือบริเวณล่างซ้าย หรือบริเวณกลางกลาง มีจุดปลายชี้ขึ้น บริเวณบนขวา มีหยักแรกจากทางด้านบน บริเวณล่างกลาง และมีหยักแรกจากด้านซ้ายบริเวณกลางซ้าย หรือ ล่างซ้าย หรือหยักแรกจากด้านล่างบริเวณล่าง กลาง”

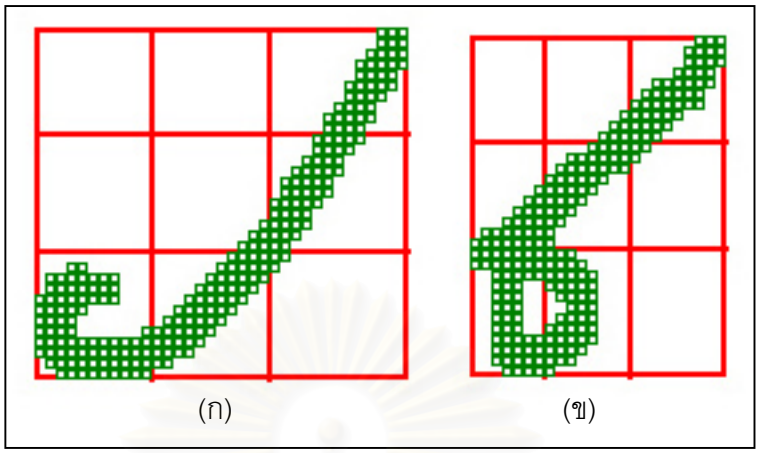
”สระ “ ั “ ชิดไม้หันอากาศ มีหัวหันไปทางซ้ายหรือขวา ทางด้านบนซ้าย หรือกลางซ้าย หรือล่างซ้าย มีจุดปลายชี้ขึ้น บริเวณบนขวา มีหยักแรกจากทางด้านบน บริเวณล่างซ้าย และไม่มีหยักแรกจากด้านซ้ายบริเวณกลางซ้าย หรือ ล่างซ้าย”



รูปที่ 4.10 ลักษณะของสระอิ “ ิ ” ,สระอี “ ี ” ,สระอึ “ ึ ” ,สระอือ “ ื ”

เนื่องจากสระทั้งสี่ชนิดมีโครงสร้างที่คล้ายกัน และมีความแตกต่างกันเพียงบริเวณด้านขวาของสระเท่านั้นจึงกำหนด กฎสำหรับสระทั้งสี่ชนิดดังนี้

- ”สระอิ มีหัวหันไปทางซ้ายหรือขวา ทางด้านกลางกลาง ไม่มีจุดปลายชี้ขึ้น บริเวณบนขวา ไม่มีหัวหันไปทางขวา บริเวณขวาบน”
- ”สระอี มีหัวหันไปทางซ้าย ทางด้านกลางกลาง มีจุดปลายชี้ขึ้น มีหยักบริเวณล่างขวา บริเวณบนขวา ไม่มีหัวหันไปทางขวา บริเวณขวาบน”
- ”สระอึ มีหัวหันไปทางซ้าย ทางด้านกลางกลาง มีหัวหันไปทางขวา บริเวณขวาบน ไม่มีจุดปลายชี้ขึ้น มีหยักบริเวณล่างขวา บริเวณบนขวา”
- ”สระอือ มีหัวหันไปทางซ้าย ทางด้านกลางกลาง มีหยักแรกบริเวณล่างขวา หรือ กลางขวา มีหยักที่สองบริเวณล่างขวา มีจุดปลายชี้ขึ้น บริเวณบนขวา ไม่มีหัวหันไปทางขวา บริเวณขวาบน”



รูปที่ 4.11 ลักษณะของสระ ไม้หันอากาศ " ั " และ การันต์ " ์ "

เนื่องจากสระทั้งสองตัวมีลักษณะคล้ายกัน ทำการแยกโครงสร้างทั้งสองชนิดโดยใช้รอยหยักทางด้านขวามาพิจารณา โดยการันต์จะมีรอยหยักทางด้านขวา บริเวณกลางกลาง ส่วนไม้หันอากาศจะมีรอยหยักทางด้านบนบริเวณล่างขวา

ในการรู้จำตัวอักษรเดี่ยวระบบจะทำงานเหมือนกันในทุกๆ ตัวอักษร ดังนั้นทำให้เวลาในการตัดสินใจตัวอักษรแต่ละตัวไม่ต่างกันมาก จากการทดลองพบว่าเวลาในการตัดสินใจเฉลี่ยต่อหนึ่งตัวอักษรเป็น 0.037 วินาที

สาเหตุที่ทำให้เกิดการรู้จำที่ผิดพลาดของการรู้จำตัวอักษรเดี่ยวสามารถสรุปได้ดังนี้

1. กฎฟัซซีไม่ครอบคลุมถึงรูปแบบของตัวอักษรที่เกิดขึ้น
2. ตัวอักษรมีรูปแบบที่คล้ายกันทำให้ระบบสับสนในการตัดสินใจ
3. ความคลุมเครือของตัวอักษร ที่มีคุณลักษณะสำคัญที่ไม่ชัดเจน หรือผิดเพี้ยนไป
4. คุณลักษณะเฉพาะที่สนใจมีรูปแบบที่หลากหลาย

4.6 ผลการทดสอบระบบรู้จำตัวอักษรตัวเดี่ยว

ในการทดสอบระบบรู้จำตัวอักษรตัวเดี่ยว ทำการทดสอบกับตัวอักษรพยัญชนะจำนวน 100 แบบ ที่ได้จากการเขียนของคน 20 คน สระและวรรณยุกต์ จำนวน 40 แบบที่ได้จากเขียนของคน 20 คน ได้ผลดังนี้

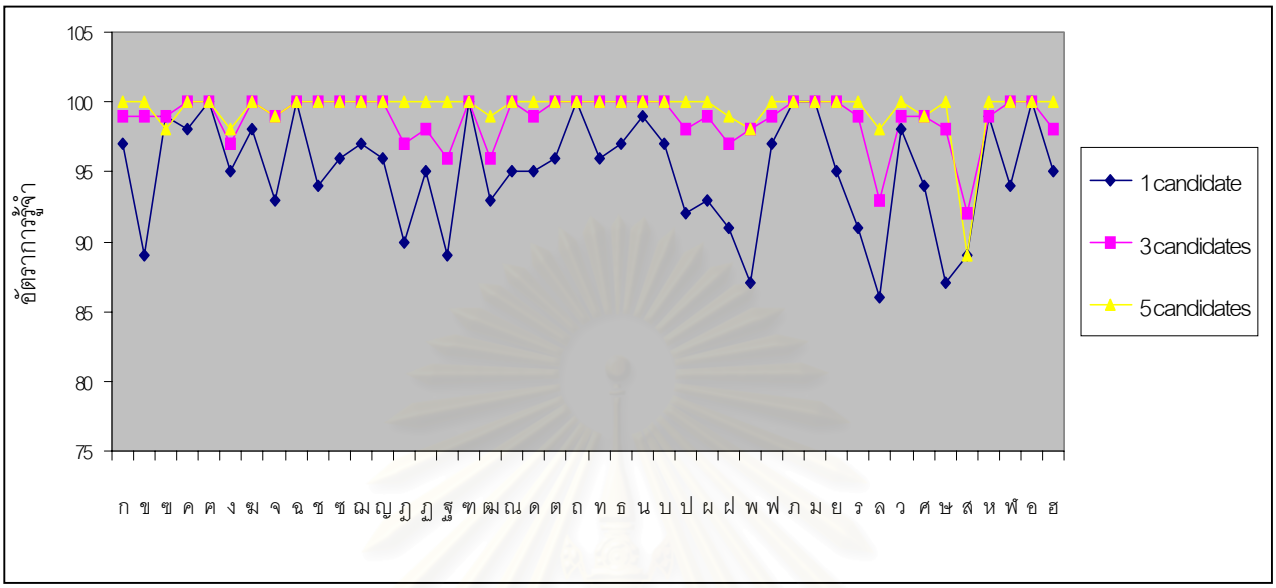
ตารางที่ 4.3 ผลการทดสอบการรู้จำพยัญชนะ ด้วยข้อมูลกลุ่มฝึกฝน และทดสอบ

	ก	ข	ฃ	ค	ฅ	ง	ฆ	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	ถูก (%)	ผด(%)
ก	97	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	97	3
ข	0	89	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	4	0	0	0	0	0	0	0	1	0	0	3	89	11
ฃ	0	0	99	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	99	1
ค	0	0	0	98	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	2
ฅ	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
ง	0	0	0	0	1	95	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	95	5
ฆ	0	0	0	0	0	0	98	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	98	2
จ	0	0	0	0	0	0	0	93	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	2	0	1	0	93	7
ฉ	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
ช	0	0	0	0	0	0	0	0	0	94	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	94	6
ซ	0	0	0	0	0	0	0	0	0	4	96	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	96	4
ฌ	0	0	0	0	0	0	0	0	0	0	0	97	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	97	3
ญ	0	0	0	0	0	0	0	0	0	0	0	4	96	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	96	4
ฎ	0	0	0	0	0	0	0	0	1	0	0	0	0	90	3	0	0	0	0	3	0	1	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	90	10
ฏ	0	0	0	0	0	0	0	0	0	0	0	0	0	2	94	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	2		95	5
ฐ	0	0	0	0	0	0	0	0	0	0	0	0	0	3	83	0	0	0	0	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	89	11
ฑ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
ฒ	0	0	2	0	0	1	0	0	0	0	0	0	0	0	0	0	93	0	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0		93	7
ณ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95	5
ด	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	95	5
ต	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	96	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	96	4
ถ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
ท	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	96	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	96	4
ธ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	97	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	97	3
น	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	99	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	99	1
บ	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	97	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	97	3
ป	0	0	0	0	0	0	0	0	0	6	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	92	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	92	8
ผ	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	93	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	93	7
ฝ	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	91	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	91	9
พ	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	5	0	4	0	87	0	0	0	0	0	0	0	0	0	2	0	0	0	87	13
ฟ	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	97	0	0	0	0	0	0	0	0	0	0	0	0	97	3
ภ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	100	0
ม	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	100	0
ย	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	95	0	0	0	0	0	0	0	0	0	95	5
ร	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	4	0	0	0	0	1	0	0	0	0	0	0	0	0	91	0	2	0	0	0	0	0	0	91	9
ล	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	2	0	0	0	2	0	0	0	0	2	0	0	0	0	86	0	0	0	0	0	0	1	0	86	14
ว	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	0	0	0	0	0	2	0	98	2
ศ	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	94	0	4	0	0	0	0	94	6
ษ	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	87	0	0	0	0	0	87	13
ส	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	4	0	3	0	89	0	0	1	0	0	89	11
ห	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	99	0	0	0	99	1
ฬ	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	94	0	4	94	6
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	100	0
ฮ	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	96	95	5
																																												อัตราการรู้จำเฉลี่ย		95.045	4.9545

ตารางที่ 4.4 ผลการทดสอบการรู้จำสระ และวรรณยุกต์ ด้วยข้อมูลกลุ่มฝึกฝน และทดสอบ

	า	ิ	ี	ึ	ื	ุ	ู	เ	แ	ใ	ไ	โ	็	ั	์	่	้	๊	๋	ก	จ	ถูก(%)	ผิด(%)
า	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	92.5	7.5
ิ	0	32	2	0	0	0	0	0	4	0	0	0	0	2	0	0	0	0	0	0	0	80	20
ี	0	0	34	0	3	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	85	15
ึ	0	0	1	39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	97.5	2.5
ื	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
ุ	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
ู	0	0	0	0	0	2	38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95	5
เ	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
แ	0	3	0	0	0	0	0	0	35	0	0	0	0	2	0	0	0	0	0	0	0	87.5	12.5
ใ	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	100	0
ไ	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	100	0
โ	0	0	0	0	0	0	0	0	0	2	0	38	0	0	0	0	0	0	0	0	0	95	5
็	0	0	1	0	0	0	0	0	0	0	0	0	37	2	0	0	0	0	0	0	0	92.5	7.5
ั	0	0	1	0	0	0	0	0	0	0	0	0	1	38	0	0	0	0	0	0	0	95	5
์	0	0	0	0	0	0	0	0	0	0	0	0	2	0	38	0	0	0	0	0	0	95	5
่	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	100	0
้	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	37	0	0	0	0	92.5	7.5
๊	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	40	0	0	0	100	0
๋	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	100	0
อัตราการรู้จำเฉลี่ย																						95.13	4.8684

4.7 ผลการเปรียบเทียบผลการรู้จำโดยเลือกจำนวนคำตอบเป็น 1 ,3 และ 5



รูปที่ 4.12 ผลการรู้จำตัวอักษรพยัญชนะเมื่อระบบตอบ 1 ,3 และ 5 คำ

รูปที่ 4.12 แสดงผลการรู้จำจากตัวอักษรพยัญชนะจำนวนตัวอักษรละ 100 แบบ ของระบบเมื่อระบบให้คำตอบจำนวน 1 ,3 และ 5 ตามลำดับ อัตราการรู้จำเฉลี่ยของระบบเป็นดังนี้

อัตรารู้จำเฉลี่ยของระบบที่ตอบ 1 คำตอบเป็น 95.04545 %

อัตรารู้จำเฉลี่ยของระบบที่ตอบ 3 คำตอบเป็น 98.79545 %

อัตรารู้จำเฉลี่ยของระบบที่ตอบ 5 คำตอบเป็น 99.47727 %

จากผลการรู้จำพบว่า เมื่อระบบตอบ 5 คำตอบอัตราการรู้จำจะสูงขึ้นกว่า ระบบที่ตอบคำตอบเดียวมาก ซึ่งระบบใช้เวลาในการรู้จำใกล้เคียงกัน เนื่องจากระบบจะต้องเปรียบเทียบความคล้ายกับรูปแบบของตัวอักษรทั้งหมด ซึ่งเวลาที่เพิ่มขึ้นจะใช้กับการเลือก และจัดลำดับค่าคำตอบที่มีความคล้ายมากที่สุดเท่านั้น ซึ่งการส่งผลการรู้จำตัวอักษรเข้าไปส่วนกระบวนการภายหลังเพื่อตรวจสอบกับพจนานุกรมทำให้ระบบสามารถแก้ไขความผิดพลาดให้คำตอบในการรู้จำที่ถูกต้องมากยิ่งขึ้น

เนื่องจากตัวอักษรภาษาไทยที่มีการเขียนติดกันเป็นคำ มีความหลากหลายของรูปแบบมากกว่าการเขียนตัวอักษรเดี่ยวทีละตัวแยกกัน นอกจากนี้บางตัวอักษรมีการเขียนติดกันโดยไม่ยกปากกา ทำให้ตัวอักษรที่เขียนมีรูปร่างที่เปลี่ยนไป ซึ่งโดยมากมักเป็นรูปแบบของการที่ลักษณะสำคัญบางอย่างหายไปหรือไม่ชัดเจน ทำให้ระบบรู้จำที่ได้รับการฝึกฝนไว้ด้วยฐานข้อมูลของรูปแบบตัวอักษรเดี่ยวที่เกิดจากการเขียนทีละตัว ไม่สามารถทำงานได้ครอบคลุมรูปแบบที่เกิดขึ้นในการเขียนที่เป็นคำได้จึงต้องมีการฝึกฝนระบบใหม่เพิ่มเติมเพื่อเรียนรู้รูปแบบที่เพิ่มเติมเข้ามา นอกจากนี้ยังเพื่อหาข้อผิดพลาดจากของระบบรู้จำในขั้นตอนอื่นๆ อีกด้วย

4.8 การฝึกฝนระบบรู้จำคำภาษาไทย

ในการรู้จำเพื่อฝึกฝนระบบของคำในภาษาไทย ใช้ฐานข้อมูลซึ่งเป็นชื่อจังหวัดรวม 76 คำ จำนวนคำละ 20 รูปแบบ มาฝึกฝนระบบเพื่อวิเคราะห์ผลการรู้จำ โดยระบบที่ใช้ในการฝึกฝนมีทั้งส่วนรู้จำตัวอักษรภาษาไทยแบบตัวเดี่ยว และส่วนการตรวจสอบคำผิดเพื่อแก้ไขคำที่รู้จำไม่ได้ด้วย

เพื่อให้ระบบทำงานได้ถูกต้องจึงได้มีการปรับเปลี่ยนรูปแบบของตัวพยัญชนะไทย 2 ตัวจาก “ญ” เป็น “ถ” และ "ญ" "และ “ฐ " เป็น “ฐ” และ " ฐ " และสระ จาก “ะ” เป็น “ ั ” และ “ะ” และ “แ” เป็น “เ” และ “แ” และ ” ำ ” เป็น “ ำ” และ “า” โดยจะพิจารณาส่วนที่ไม่ติดกันของตัวอักษรเสมือนตัวอักษรอีกตัวอักษรหนึ่งเพื่อให้ใช้ในระบบรู้จำแบบเป็นคำได้สะดวกผลการรู้จำได้ผลดังตารางที่ 4.5 ดังนี้

ตารางที่ 4.5 ผลการทดสอบการรู้จำอักษรภาษาไทยที่เป็นคำด้วยข้อมูลชุดฝึกฝน

คำ	รู้จำถูก (คำ)	รู้จำถูก (%)	ผิดพลาดเป็น (คำ)
กระบี่	20	100	-
กรุงเทพมหานคร	20	100	-
กาญจนบุรี	20	100	-
กาฬสินธุ์	18	90	สิงห์บุรี (2)
กำแพงเพชร	20	100	-
ขอนแก่น	20	100	-
จันทบุรี	19	95	สระบุรี(1)
ฉะเชิงเทรา	19	95	กำแพงเพชร (1)
ชลบุรี	17	85	ลพบุรี (3)
ชัยนาท	20	100	-
ชัยภูมิ	19	95	ปัตตานี (1)

คำ	รู้จำถูก (คำ)	รู้จำถูก (%)	ผิดพลาดเป็น (คำ)
ชุมพร	20	100	-
เชียงราย	20	100	-
เชียงใหม่	18	90	เพชรบูรณ์(2)
ตรัง	20	100	-
ตราด	20	100	-
ตาก	20	100	-
นครนายก	20	100	-
นครปฐม	19	95	นครนายก (1)
นครพนม	20	100	-
นครราชสีมา	20	100	-
นครศรีธรรมราช	20	100	-

ตารางที่ 4.5 ผลการทดสอบการรู้จำอักษรภาษาไทยที่เป็นคำด้วยข้อมูลชุดฝึกฝน (ต่อ)

คำ	รู้จำถูก (คำ)	รู้จำถูก (%)	ผิดพลาดเป็น (คำ)
นครสวรรค์	20	100	-
นนทบุรี	17	85	นครปฐม (3)
นราธิวาส	20	100	-
น่าน	20	100	-
บุรีรัมย์	19	95	อุทัยธานี (1)
ปทุมธานี	18	90	อุดรธานี(2)
ปราจีนบุรี	20	100	-
ประจวบคีรีขันธ์	20	100	-
ปัตตานี	20	100	-
พะเยา	17	85	พิจิตร (2),ระนอง(1)
พังงา	20	100	-
พัทลุง	19	95	พังงา(1)
พิจิตร	18	90	สกลนคร (1)
พิษณุโลก	20	100	-
เพชรบุรี	20	100	-
เพชรบูรณ์	20	100	-
แพร่	20	100	-
ภูเก็ต	20	100	-
มหาสารคาม	20	100	-
แม่ฮ่องสอน	20	100	-
มุกดาหาร	19	95	-
ยะลา	18	90	ยโสธร(2)
ยโสธร	19	95	ยะลา (1)
ร้อยเอ็ด	20	100	-
ระนอง	17	85	ระยอง (3)
ระยอง	19	95	ระนอง(1)

คำ	รู้จำถูก (คำ)	รู้จำถูก (%)	ผิดพลาดเป็น (คำ)
ราชบุรี	20	100	-
ลพบุรี	19	95	ชลบุรี (1)
เลย	20	100	-
ลำปาง	17	85	ลำพูน (3)
ลำพูน	19	95	ลำปาง (1)
ศรีสะเกษ	20	100	-
สกลนคร	20	100	-
สงขลา	18	90	ยะลา (2)
สตูล	20	100	-
สมุทรปราการ	18	90	สมุทรสงคราม(2)
สมุทรสงคราม	20	100	-
สมุทรสาคร	19	95	สระแก้ว (1)
สระแก้ว	19	95	ศรีสะเกศ (1)
สระบุรี	20	100	-
สิงห์บุรี	19	95	อุทัยธานี (1)
สุพรรณบุรี	19	95	ปราจีนบุรี(1)
สุพรรณบุรี	19	95	ปราจีนบุรี(1)
สุราษฎร์ธานี	20	100	-
สุรินทร์	16	80	สระบุรี(4)
สุโขทัย	20	100	-
หนองคาย	20	100	-
หนองบัวลำภู	20	100	-
อยุธยา	20	100	-
อ่างทอง	20	100	-
อุดรธานี	19	95	ปทุมธานี(1)
อุตรดิตถ์	19	95	บุรีรัมย์ (1)

ตารางที่ 4.5 ผลการทดสอบการรู้จำอักษรภาษาไทยที่เป็นคำด้วยข้อมูลชุดฝึกฝน (ต่อ)

คำ	รู้จำถูก (คำ)	รู้จำถูก (%)	ผิดพลาดเป็น (คำ)
อุทัยธานี	15	75	อุตรดิตถ์(5)
อุบลราชธานี	20	100	-

คำ	รู้จำถูก (คำ)	รู้จำถูก (%)	ผิดพลาดเป็น (คำ)
อำนาจเจริญ	20	100	-
อัตราการรู้จำผิดพลาด			3.61842105

วิเคราะห์ผลการฝึกฝนระบบด้วยข้อมูลที่เป็นคำ

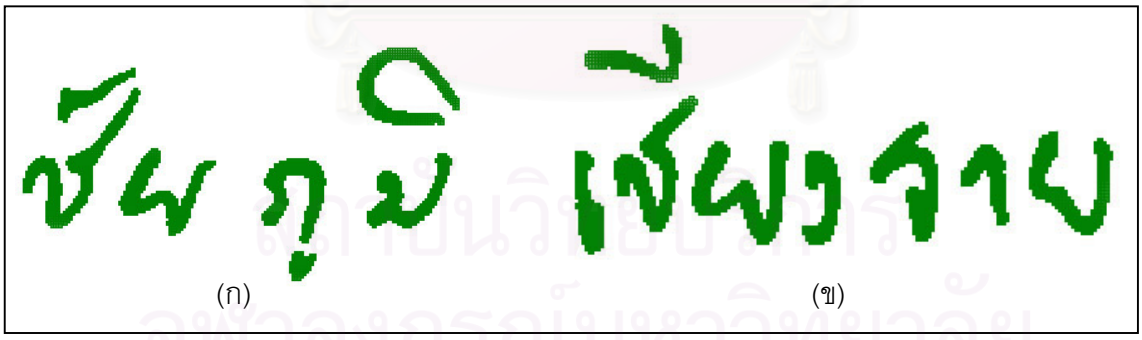
1. ความหลากหลายของรูปของตัวอักษรเมื่อมีการเขียนตัวอักษรต่อกันเป็นคำ ซึ่งต่างจากการเขียนตัวอักษรทีละตัว คุณลักษณะสำคัญที่เคยพบในการฝึกฝนระบบรู้จำตัวอักษรแบบตัวเดี่ยว มีการเปลี่ยนแปลง และขาดหายไปเป็นผลให้การรู้จำอักษรเกิดความผิดพลาด

   โดยมีการแบ่งประเภทของรูปแบบของปัญหาออกเป็นกลุ่ม ๆ ได้ดังนี้

   1.1 รูปแบบของลักษณะสำคัญของตัวอักษร เนื่องจากตัวอักษรไทยมีลักษณะเฉพาะในแต่ละตัวอักษร การที่ลักษณะเฉพาะบางประการหายไปมีผลต่อการรู้จำอย่างมาก ดังนั้นรูปแบบที่เปลี่ยนไปของตัวอักษรจึงเป็นสิ่งที่สำคัญ มีผลต่อการตัดสินใจ

   ในรูปที่ 4.13 แสดงตัวอย่างของตัวอักษรที่ลักษณะสำคัญบางอย่างของตัวอักษรหายไป ซึ่งมีทั้งหัวของตัวอักษร จุดปลาย และรอยหยักที่หายไป

   การแก้ไข ทำได้โดยการเพิ่มรูปแบบของจุดปลายที่ตำแหน่งหัวเดิมที่เคยพบ เข้าไปในกฎทางฟัซซีของตัวอักษรนั้นด้วย



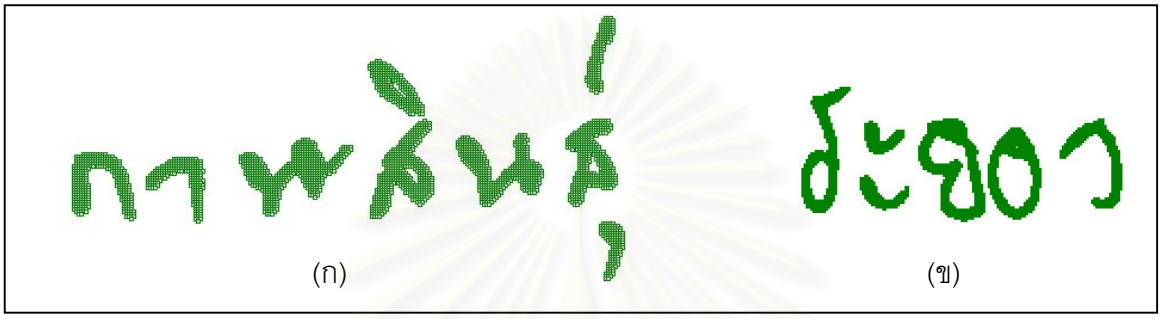
(ก) (ข)

รูปที่ 4.13 รูปลายมือเขียนที่ลักษณะสำคัญของตัวอักษรเปลี่ยนรูปแบบไป

   1.2 การทับกันของเส้นภายในตัวอักษร เนื่องจากการลากเส้นของผู้เขียนทำให้เกิดรูปแบบอยู่ 2 ลักษณะ คือ เกิดรูปแบบของตัวอักษรใหม่ และลักษณะสำคัญหายไป โดยทั้งสองรูปแบบเป็นผลเนื่องจากลักษณะการเขียนของผู้เขียนเอง และขนาดของตัวอักษรที่เขียน เมื่อตัวอักษรมีขนาดเล็ก ลักษณะสำคัญบางอย่างมักขาดหายไป

รูปที่ 4.14 แสดงรูปแบบบางของตัวอักษรบางแบบเมื่อเกิดการเชื่อมกันของเส้นภายในตัวอักษรตัวเดียวกัน ซึ่งมักเกิดขึ้นในการเขียน บางครั้งโครงสร้างอาจเปลี่ยนไปจากเดิมอย่างมากทำให้ไม่สามารถรู้จำได้ด้วยกฎทางฟัซซีที่มีอยู่

การแก้ไข เมื่อเกิดรูปแบบของตัวอักษรขึ้นมาให้สามารถแก้ไขได้โดยเพิ่มกฎของรูปแบบตัวอักษรนั้น ๆ เพิ่มขึ้นจะทำให้สามารถรู้จำรูปแบบใหม่นั้นได้

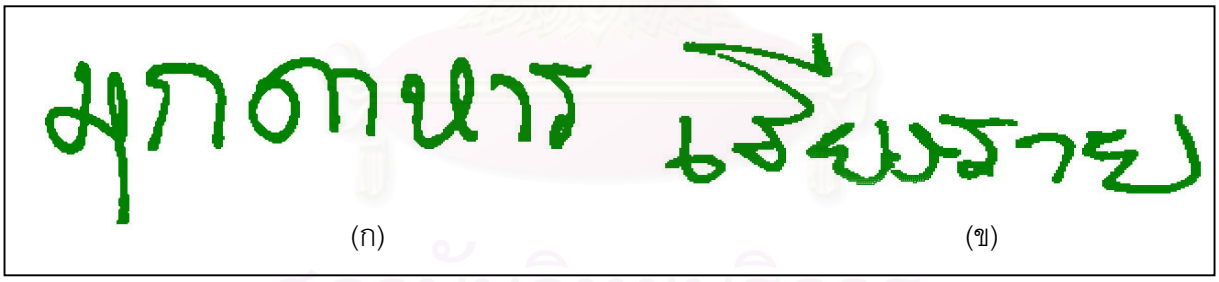


(ก) (ข)

รูปที่ 4.14 รูปลายมือเขียนที่เกิดตัวอักษรในรูปแบบใหม่

1.2 ความเปลี่ยนแปลงของตัวอักษรเนื่องจากการเขียนตัวอักษรติดกัน โดยไม่ยกปากกา ทำให้รูปแบบของตัวอักษรที่ติดกันเกิดเปลี่ยนแปลงไป ดังแสดงในรูปที่ 4.15

การแก้ไข เพิ่มรูปแบบที่เปลี่ยนไปเข้ากับกฎทางฟัซซีของตัวอักษรนั้นเดิมที่มีอยู่ เนื่องจากการเขียนต่อเนื่องมักเกิดขึ้นบ่อยในการเขียนโดยปกติ

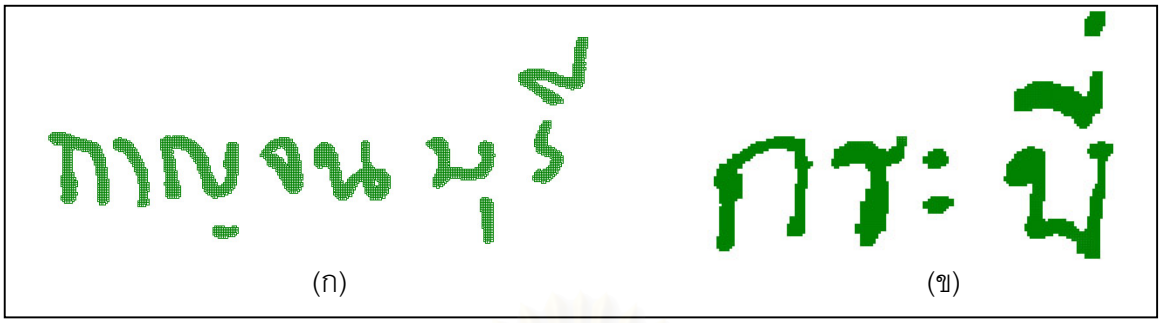


(ก) (ข)

รูปที่ 4.15 รูปลายมือเขียนที่มีการเปลี่ยนรูปแบบของตัวอักษรที่เกิดจากการเขียนที่ต่อเนื่องกัน

1.3 รูปแบบของสระที่เกิดขึ้นในการเขียนคำในภาษาไทย ลักษณะสำคัญของสระมักจะถูกละเลยไปในการเขียนทำให้การรู้จำสระในคำเป็นเรื่องที่ลำบาก ดังแสดงในรูปที่ 4.16

การแก้ไข เนื่องจากมีการแบ่งกลุ่มของสระเทียบกับเส้นบรรทัดอ้างอิงไว้ ทำให้จำนวนสระมีจำนวนรูปแบบลดลง ดังนั้นการที่ระบบรู้จำให้ตัวเลือกหลายคำตอบมาเพื่อตรวจสอบกับกระบวนการภายหลัง ในแต่ละตัวอักษรทำให้สามารถแก้อักษรที่รู้จำผิดพลาดได้



(ก) (ข)

รูปที่ 4.16 รูปลายมือเขียนที่มีรูปแบบของสระเปลี่ยนแปลงไป

2. ผลเนื่องจากขนาดของสระทำให้การลำดับตัวอักษรผิดพลาด ดังแสดงในรูปที่ 4.17 ตำแหน่งของสระขยับไปทางตัวอักษรด้านหลังทำให้ระบบเรียงลำดับของตัวอักษรเป็น “ชยัภูมิ” จาก “ชัยภูมิ” ซึ่งทำให้ตรวจสอบกับพจนานุกรมผิดพลาดได้

การแก้ไข ทำได้โดยการเพิ่มคำที่มีการลำดับตัวอักษรแบบอื่นลงในฐานข้อมูลของพจนานุกรม

รูปที่ 4.17 ภาพอักษรที่เกิดการลำดับอักษรในการรู้จำผิดพลาด

4.9 ผลการทดสอบระบบรู้จำคำภาษาไทย

ในการทดสอบระบบรู้จำคำรายชื่อจังหวัดในภาษาไทยที่ได้ผ่านการปรับปรุงให้ระบบมีการรู้จำที่ถูกต้องมากขึ้น ใช้ฐานข้อมูลซึ่งเป็นชื่อจังหวัดรวม 76 คำ จำนวนคำละ 40 รูปแบบ ซึ่งประกอบด้วยชุดคำที่ใช้ฝึกฝน 20 รูปแบบ และชุดคำที่ระบบยังไม่เคยพบจำนวน 20 รูปแบบ และระบบรู้จำที่ใช้มีการใช้ พจนานุกรมคำศัพท์ชื่อจังหวัดในการตรวจสอบความผิดพลาดของการรู้จำ ซึ่งผลการรู้จำเป็นดังตารางที่ 4.6 ดังนี้

ตารางที่ 4.6 ผลการทดสอบระบบรู้จำอักษรที่เป็นคำ ด้วยชุดทดสอบ

คำ	รู้จำถูก (คำ)	รู้จำถูก (%)	ผิดพลาดเป็น (คำ)	คำ	รู้จำถูก (คำ)	รู้จำถูก (%)	ผิดพลาดเป็น (คำ)
กระบี่	20	100	-	บุรีรัมย์	19	95	อุทัยธานี (1)
กรุงเทพมหานคร	20	100	-	ปทุมธานี	20	100	-
กาญจนบุรี	20	100	-	ปราจีนบุรี	20	100	-
กาฬสินธุ์	20	100	-	ประจวบคีรีขันธ์	20	100	-
กำแพงเพชร	20	100	-	ปัตตานี	20	100	-
ขอนแก่น	20	100	-	พะเยา	18	90	พิจิตร (2)
จันทบุรี	20	100	-	พังงา	20	100	-
ฉะเชิงเทรา	19	95	กำแพงเพชร (1)	พัทลุง	19	95	พังงา(1)
ชลบุรี	19	95	ลพบุรี (1)	พิจิตร	19	95	สกลนคร (1)
ชัยนาท	20	100	-	พิษณุโลก	20	100	-
ชัยภูมิ	19	95	ปัตตานี (1)	เพชรบุรี	20	100	-
ชุมพร	20	100	-	เพชรบูรณ์	20	100	-
เชียงราย	20	100	-	แพร่	20	100	-
เชียงใหม่	19	95	เพชรบูรณ์(1)	ภูเก็ต	20	100	-
ตรัง	20	100	-	มหาสารคาม	20	100	-
ตราด	20	100	-	แม่ฮ่องสอน	20	100	-
ตาก	20	100	-	มุกดาหาร	20	100	-
นครนายก	20	100	-	ยะลา	18	90	ยโสธร(2)
นครปฐม	20	100	-	ยโสธร	19	95	ยะลา (1)
นครพนม	20	100	-	ร้อยเอ็ด	20	100	-
นครราชสีมา	20	100	-	ระนอง	17	85	ระยอง (3)
นครศรีธรรมราช	20	100	-	ระยอง	20	100	-
นครสวรรค์	20	100	-	ราชบุรี	20	100	-
นนทบุรี	19	95	นครปฐม (1)	ลพบุรี	20	100	-
นราธิวาส	20	100	-	เลย	20	100	-
น่าน	20	100	-	ลำปาง	19	95	ลำพูน (1)

ตารางที่ 4.6 ผลการทดสอบระบบรู้จำอักษรที่เป็นคำ ด้วยชุดทดสอบ (ต่อ)

คำ	รู้จำถูก (คำ)	รู้จำถูก (%)	ผิดพลาดเป็น (คำ)
ลำพูน	20	100	-
ศรีสะเกษ	20	100	-
สกลนคร	20	100	-
สงขลา	20	100	-
สตูล	20	100	-
สมุทรปราการ	20	100	-
สมุทรสงคราม	20	100	-
สมุทรสาคร	19	95	สระแก้ว (1)
สระแก้ว	19	95	ศรีสะเกศ (1)
สระบุรี	20	100	-
สิงห์บุรี	19	95	อุทัยธานี (1)
สุพรรณบุรี	20	100	-

คำ	รู้จำถูก (คำ)	รู้จำถูก (%)	ผิดพลาดเป็น (คำ)
สุราษฎร์ธานี	20	100	-
สุรินทร์	18	90	สระบุรี(2)
สุโขทัย	20	100	-
หนองคาย	20	100	-
หนองบัวลำภู	20	100	-
อยุธยา	20	100	-
อ่างทอง	20	100	-
อุดรธานี	19	95	ปทุมธานี(1)
อุตรดิตถ์	19	95	บุรีรัมย์ (1)
อุทัยธานี	18	90	อุตรดิตถ์(2)
อุบลราชธานี	20	100	-
อำนาจเจริญ	20	100	-
อัตราการรู้จำผิดพลาดเฉลี่ย			1.71053

วิเคราะห์ผลการทดสอบระบบรู้จำ

คำที่ได้รับการแก้ไขให้มีความถูกต้องมากขึ้นเนื่องจากได้มีการปรับปรุงกฎทางฟัซซีให้รู้จักกับรูปแบบของตัวอักษรที่เกิดขึ้น ซึ่งระบบไม่เคยพบรูปแบบของตัวอักษรเหล่านี้ในการทดสอบกับตัวอักษรตัวเดี่ยว ตัวอักษรที่พบในการเขียนเป็นคำมักมีความไม่ชัดเจนของคุณลักษณะสำคัญ ซึ่งยากต่อการรู้จำโดยปกติ

ความผิดพลาดที่ยังคงอยู่ภายหลังการฝึกฝนระบบ

ความผิดพลาดของการรู้จำที่เกิดขึ้นหลังการฝึกฝนระบบ เมื่อทดสอบด้วยคำรายชื่อจังหวัด 76 จังหวัด จำนวน 40 รูปแบบ เกิดขึ้นเนื่องจาก 3 สาเหตุดังต่อไปนี้

1. ความผิดพลาดเนื่องจากความคล้ายกันของตัวอักษรภายในคำ ทำให้มีโอกาสที่ระบบจะรู้จำผิดไปเป็นคำอื่นที่มีตัวอักษรภายในคำเหมือนกัน ในกรณีนี้ ตัวอักษรที่แตกต่างกันเกิดการรู้จำที่ผิดหรือมีการรู้จำเป็นตัวอักษรที่อยู่ในคำที่คล้ายกันด้วย ทำให้ไม่สามารถบ่งบอกความแตกต่างของคำทั้งสองได้ คำที่มีความผิดพลาดเนื่องจากสาเหตุนี้ได้แก่ พัทลุง ผิดเป็น พังงา ชลบุรี ผิดเป็น ลพบุรี ลำปาง ผิดเป็น ลำพูน อุดรธานี ผิดเป็น ปทุมธานี อุทัยธานี ผิดเป็น อุตรดิตถ์

รูปที่ 4.18 แสดงตัวอย่างของตัวอักษรที่มีความไม่ชัดเจน ทำให้เกิดการผิดพลาดไปเป็นคำที่มีตัวอักษรภายในคำที่คล้ายกัน ความผิดพลาดที่เกิดขึ้นมักเกิดในตำแหน่งของตัวอักษรที่บ่งบอกความแตกต่างของคำ ในรูปที่ 4.18 (ก) มีความผิดพลาดเป็นจาก ชลบุรี ผิดพลาดเป็น ลพบุรี เนื่องจากตัวอักษร “ช” และ “ล” ภายในคำมีความผิดเพี้ยนของรู้แบบตัวอักษรผิดไปจากลักษณะของตัวอักษรปกติมากทำให้ระบบไม่สามารถรู้จำได้ถูก ทำให้เกิดความผิดพลาดขึ้น ในรูปที่ 4.18 (ข) ตัวอักษร “ป” ในคำว่า “ลำปาง” มีความคล้ายกับตัวอักษร “ฟ” และ “พ” ทำให้ระบบรู้จำผิดพลาดเป็นคำว่า “ลำพูน”

(ก) (ข)

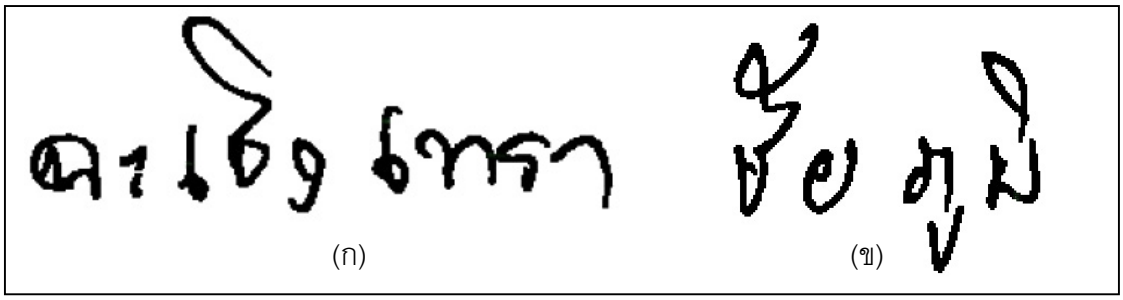
รูปที่ 4.18 ตัวอักษรที่เกิดการรู้จำผิดพลาด เนื่องจากคำมีความคล้ายกัน

(ก) คำชื่อจังหวัด “ชลบุรี” ผิดพลาดเป็น “ลพบุรี”

(ข) คำชื่อจังหวัด “ลำปาง” ผิดพลาดเป็น “ลำพูน”

2. ความผิดพลาดเนื่องจากตัวอักษรในคำมีความไม่ชัดเจน ทำให้ไม่สามารถรู้จำตัวอักษรภายในคำนั้นได้ ทำให้ระบบเลือกคำอื่นที่มีความคล้ายที่สุดขึ้นมาแทน ในกรณีนี้คำที่มักผิดพลาด มักมีสระอยู่มากซึ่งเวลาเขียนเป็นคำ สระมักจะถูกละเลยรายละเอียดไปทำให้เกิดการรู้จำที่ผิดพลาด นอกจากนี้ ปัจจัยที่ช่วยทำให้กรณีนี้มีความผิดพลาดมากขึ้นคือ ตำแหน่งของสระในคำที่คลาดเคลื่อนทำให้การจัดลำดับตัวอักษร และสระเพื่อเปรียบเทียบกับพจนานุกรม มีความผิดพลาดตามไปด้วย คำที่มีความผิดพลาดเนื่องจากสาเหตุนี้ได้แก่ ฉะเชิงเทรา ผิดเป็น กำแพงเพชร นนทบุรี ผิดเป็น นครปฐม พะเยา ผิดเป็น พิจิตร สระแก้ว ผิดเป็น ศรีสะเกศ สิงห์บุรี ผิดเป็น อุทัยธานี ชัยภูมิ ผิดเป็น ปัตตานี พิจิตร ผิดเป็น สกลนคร ยะลา ผิดเป็น ยโสธร ยโสธร ผิดเป็น ยะลา

รูปที่ 4.19 แสดงตัวอย่างคำที่เกิดการรู้จำผิด เนื่องจากตัวอักษรภายในคำมีความไม่ชัดเจนและมีความผิดเพี้ยนไปจากรูปแบบในปกติมาก ทำให้ระบบไม่สามารถรู้จำตัวอักษรภายในคำได้ และระบบเลือกคำที่มีความคล้ายกับผลการรู้จำตัวอักษรมากสุด ความผิดพลาดที่เกิดขึ้น



(ก) (ข)

รูปที่ 4.19 คำที่เกิดการรู้จำผิดพลาด เนื่องจากความไม่ชัดเจนของตัวอักษร

(ก) คำชื่อจังหวัด “ฉะเชิงเทรา” ผิดพลาดเป็น “กำแพงเพชร”

(ข) คำชื่อจังหวัด “ชัยภูมิ” ผิดพลาดเป็น “ปัตตานี”

3. ความผิดพลาดเนื่องจากระดับบรรทัดมีความผิดพลาดทำให้การบ่งบอกตัวอักษรในระดับบรรทัดเพื่อแยกกลุ่มมีความผิดพลาดตาม ซึ่งความผิดพลาดในกรณีนี้ มาจาก 2 ประการคือ ตัวอักษรในคำมีสระทั้งด้านบนและด้านล่าง และการเขียนตัวอักษรใกล้กัน ทำให้การแยกระดับบรรทัดด้วยวิธีการปกติไม่สามารถหาตำแหน่งการแบ่งได้คำที่มีความผิดพลาดเนื่องจากสาเหตุนี้ได้แก่ เชียงใหม่ ผิดเป็น เพชรบูรณ์ สุรินทร์ ผิดเป็น สระบุรี อุตรดิตถ์ ผิดเป็น บุรีรัมย์

รูปที่ 4.20 แสดงตัวอย่างของคำที่มีการแบ่งเส้นบรรทัดผิดพลาด เป็นเหตุให้เกิดการรู้จำผิด โดยคำว่า “สุรินทร์” มีสระทั้งในระดับบนและระดับล่าง และหางของ “ส” ทำให้ยากต่อการหาเส้นบรรทัดเมื่อมีการเขียนตัวอักษรใกล้กันมาก ระบบจึงไม่สามารถหาระดับการแบ่งของบรรทัดจนทำให้รู้จำผิดได้

รูปที่ 4.20 คำที่เกิดการรู้จำผิดพลาด เนื่องจากตัวอักษรติดกันมาก

4.10 ผลการทดสอบกระบวนการภายหลัง

กระบวนการค้นหาคำจากพจนานุกรมมีผู้ออกแบบระบบเพื่อให้สามารถค้นหาและตรวจสอบคำได้อย่างมีระเบียบและมีความรวดเร็วในการหา และสามารถใช้กับฐานข้อมูลคำที่มีขนาดใหญ่ได้ดี แต่มีข้อจำกัดเนื่องจากคุณสมบัติในการการทำงานดังนี้

1. การค้นหาคำในทิศทางเดียว ซึ่งส่วนใหญ่มักเรียกจากอักษรซ้ายไปขวา
2. แก้ไขความผิดพลาดของตัวอักษรในคำได้ในจำนวนอักษรที่จำกัด และไม่สามารถแก้คำผิดเมื่อเกิดความผิดพลาดหลาย ๆ ตัวได้

แต่ในระบบรู้จำตัวอักษรเขียน การรู้จำผิดพลาดสามารถเกิดขึ้นได้ทุกตำแหน่งเท่าๆ กัน และมักเกิดความผิดพลาดพร้อมกันหลายๆ ตำแหน่งพร้อมๆ กันดังนั้นจึงต้องการระบบการตรวจสอบคำที่ผิดพลาดที่มีคุณสมบัติดังต่อไปนี้

1. ตำแหน่งของความผิดพลาดที่แตกต่างกัน ไม่มีผลต่อการแก้ไขคำที่ผิด
2. แก้ไขความผิดพลาดจากการรู้จำได้หลายตำแหน่ง
3. มีการตรวจสอบที่รวดเร็วเมื่อเทียบกับเวลาที่ใช้ในการรู้จำ

ดังนั้นจึงได้เสนอวิธีการตรวจสอบและแก้ไขคำที่รู้จำผิดพลาด ดังแสดงในหัวข้อต่อไปนี้

4.10.1 กระบวนการภายหลังที่ตรวจสอบส่วนของคำย่อย

เป็นวิธีการแยกกลุ่มของคำตามส่วนย่อยในคำที่เหมือนกัน ซึ่งในกลุ่มคำที่ใช้ทดสอบมีกลุ่มคำที่เหมือนกันดังแสดงในตารางที่ 4.7

ตารางที่ 4.7 กลุ่มของคำแบ่งตามจำนวนตัวอักษร

จำนวนตัวอักษร	คำ
3	ตาก เลย
4	ตรัง ตราด น่าน สตูล
5	ชุมพร พังงา พัทลุง แพร่ ยะลา ยโสธร สงขลา
6	ชลบุรี ชัยนาท นครพนม พะเยา พิจิตร ภูเก็ต ระนอง ระยอง ลพบุรี ลำปาง ลำพูน สกลนคร อยุธยา
7	กระบี่ ชัยภูมิ นครนายก นครปฐม นนทบุรี ปัตตานี ราชบุรี สุโขทัย หนองคาย อ่างทอง
8	ขอนแก่น จันทบุรี เชียงราย นราธิวาส ปทุมธานี พิษณุโลก เพชรบุรี มุกดาหาร ร้อยเอ็ด สระบุรี สุรินทร์ อุดรธานี
9	กาฬสินธุ์ เชียงใหม่ นครสวรรค์ บุรีรัมย์ เพชรบูรณ์ มหาสารคาม ศรีสะเกศ สมุทรสาคร สระแก้ว สิงห์บุรี อุตรดิตถ์ อุทัยธานี
10	กาญจนบุรี นครราชสีมา ปราจีนบุรี สุพรรณบุรี

ตารางที่ 4.7 กลุ่มของคำแบ่งตามจำนวนตัวอักษร (ต่อ)

11	กำแพงเพชร ฉะเชิงเทรา แม่ฮ่องสอน สมุทรปราการ สมุทรสงคราม อุบลราชธานี
12	สุราษฎร์ธานี หนองบัวลำภู อำนาจเจริญ
13	กรุงเทพมหานคร นครศรีธรรมราช
16	ประจวบคีรีขันธ์

ในการแบ่งกลุ่มนี้ตัวอักษรจะต้องมีการตัดแยกได้อย่างถูกต้องเท่านั้น ซึ่งระบบจะมีโอกาสผิดพลาดเมื่อคำมีจำนวนอักษรในคำเท่ากัน และไม่สามารถรู้จำตัวอักษรในคำได้

4.10.2 **กระบวนการภายหลังที่ตรวจสอบส่วนของคำย่อย**

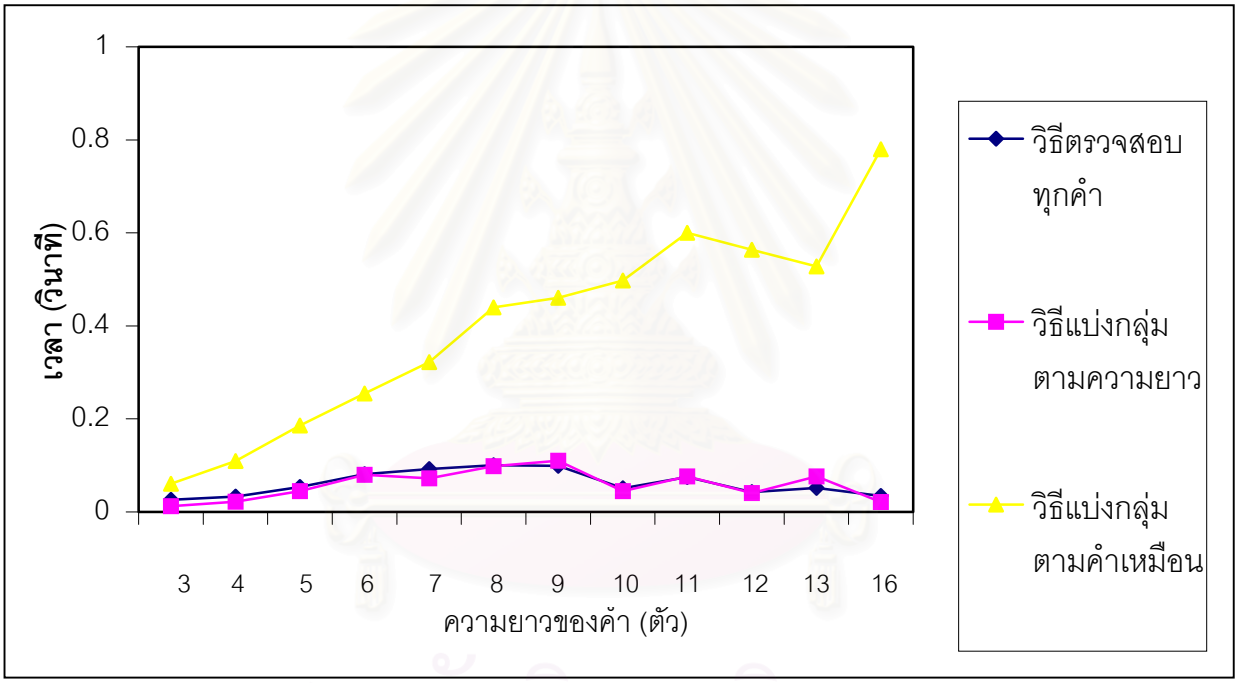
เป็นวิธีการแยกกลุ่มของคำตามส่วนย่อยในคำที่เหมือนกัน ซึ่งในกลุ่มคำที่ใช้ทดสอบมีกลุ่มคำที่เหมือนกันดังแสดงในตาราง

ตารางที่ 4.8 ส่วนของคำย่อยที่พบบ่อยในคำฐานข้อมูล

ส่วนของคำ	คำ
ชัย	ชัยนาท ชัยภูมิ
เชียง	เชียงใหม่ เชียงราย
ทัย	สุโขทัย อุทัยธานี
ธานี	ปทุมธานี สุราษฎร์ธานี อุดรธานี อุทัยธานี อุบลราชธานี
นคร	กรุงเทพมหานคร นครนายก นครปฐม นครพนม นครราชสีมา นครศรีธรรมราช นครสวรรค์ สกลนคร
บุรี	กาญจนบุรี จันทบุรี ชลบุรี นนทบุรี บุรีรัมย์ ปราจีนบุรี เพชรบุรี ราชบุรี ลพบุรี สระบุรี สิงห์บุรี สุพรรณบุรี
ปรา	ปราจีนบุรี สมุทรปราการ
เพชร	กำแพงเพชร เพชรบุรี เพชรบูรณ์
มหา	กรุงเทพมหานคร มหาสารคาม
ศรี	นครศรีธรรมราช ศรีสะเกศ
สมุทร	สมุทรปราการ สมุทรสงคราม สมุทรสาคร
สระ	สระแก้ว สระบุรี
หนอง	หนองคาย หนองบัวลำภู

วิธีนี้จะทำการหาคำย่อย ๆ ในชุดคำตอบ แล้วจึงตรวจสอบกับคำในกลุ่ม ถ้าระบบไม่เจอคำย่อย จะต้องตรวจสอบกับจำนวนคำที่เหลือทั้งหมดอีกครั้งหนึ่ง เนื่องจากเป็นไปได้ที่ระบบจะรู้จำผิดพลาดที่คำย่อยพอดี ในการแบ่งกลุ่มคำชนิดนี้ จะมีประสิทธิภาพเมื่อมีการรู้จำที่มีความถูกต้องสูง และมีจำนวนคำค่อนข้างมาก

ผลการทดสอบกระบวนการภายหลังที่ใช้การแบ่งกลุ่มคำตามความยาวของคำ กับกระบวนการภายหลังที่ไม่มีการแบ่งกลุ่มคำตามความยาว และการทดสอบกระบวนการภายหลังที่ใช้การแบ่งกลุ่มคำตามความยาวของคำ และเทียบกับคำย่อย



รูปที่ 4.21 เวลาในการทำงานของกระบวนการภายหลัง

ค่าเฉลี่ยของทั้งสามวิธีเป็นดังนี้

วิธีตรวจสอบกับทุกคำ ใช้เวลาเฉลี่ยเท่ากับ 0.0617 วินาทีต่อคำ

วิธีตรวจสอบโดยแบ่งกลุ่มตามความยาวของคำ ใช้เวลาเฉลี่ยเท่ากับ 0.058 วินาทีต่อคำ

วิธีตรวจสอบโดยแบ่งกลุ่มคำตามคำเหมือน ใช้เวลาเฉลี่ยเท่ากับ 0.4 วินาทีต่อคำ

โดยคำที่ใช้มีความยาวเฉลี่ย 7.8 ตัวอักษรต่อคำ

รูปที่ 4.21 แสดงเวลาเฉลี่ยในกระบวนการตรวจสอบกับพจนานุกรมที่ใช้ในแต่ละความยาว และจากค่าเวลาเฉลี่ยพบว่า วิธีการตรวจสอบคำโดยใช้การแบ่งกลุ่มคำตามความยาว ให้เวลาในการ

ตรวจสอบต่ำสุด เนื่องจากใช้จำนวนการเปรียบเทียบน้อยที่สุด วิธีการนี้ใช้ได้ดีเมื่อมีจำนวนคำในแต่ละกลุ่มใกล้เคียงกัน แม้ว่าจำนวนคำที่ใช้จะมากกว่านี้ ส่วนในวิธีการแบ่งกลุ่มคำจากคำที่เหมือนกัน จากกราฟแสดงเวลาเราพบว่า ใช้เวลาในการตรวจสอบมากกว่าทั้งสองวิธีมาก เนื่องจาก ต้องตรวจสอบหาคำย่อยก่อน แล้วจึงค้นหาคำเต็ม หากในขั้นตอนหาคำย่อยไม่พบ เนื่องจากไม่มีคำในตัวอักษรเขียนหรือเกิดการรู้จำผิดพลาด ระบบจะต้องทำการตรวจสอบกับคำทั้งหมดอีกครั้งหนึ่งทำให้เกิดการเปรียบเทียบคำมากกว่าทั้งสองวิธีข้างต้น ระบบในวิธีที่สามจะทำได้ดีเมื่อการรู้จำมีความถูกต้องสูง และมีคำจำนวนมากพอ ซึ่งไม่เหมาะกับกลุ่มตัวอย่างที่ใช้ทดสอบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

สรุปผลการทดลองและข้อเสนอแนะ

วิทยานิพนธ์ฉบับนี้ได้เสนอการรู้จำตัวอักษรภาษาไทย โดยแบ่งออกเป็น 2 ส่วนด้วยกันได้แก่ การรู้จำอักษรภาษาไทยตัวเดี่ยวโดยใช้กฎทางฟัซซีร่วมกับคุณลักษณะบ่งความต่างเพื่อรู้จำอักษรภาษาไทยแบบตัวเดี่ยว และกระบวนการภายหลังที่ใช้การแบ่งกลุ่มคำออกตามความยาวของคำในระบบรู้จำตัวอักษรภาษาไทย ได้ยืนยันถึงแนวความคิดที่ว่าสามารถทำงานได้ด้วยอัตราการรู้จำผิดพลาดเฉลี่ยของตัวอักษรตัวเดี่ยวร้อยละ 3.62 และอัตราการรู้จำผิดพลาดเฉลี่ยของคำภาษาไทยร้อยละ 1.71

ในการเขียนตัวอักษรภาษาไทยที่เป็นคำ ปัจจัยต่างๆ มีผลต่อรูปแบบของตัวอักษรไม่ว่าจะเป็นกระดาษ ปากกา ความละเอียดในการสแกนภาพ และลักษณะลายมือของผู้เขียนเอง ในผลการทดลองแสดงให้เห็นว่า ตัวอักษรที่เขียนเป็นคำ จะเกิดความผิดเพี้ยนของรูปแบบไปจากโครงสร้างตามปกติมาก ซึ่งเป็นผลให้การตัดสินใจอย่างถูกต้องเป็นไปได้ยาก ดังนั้นการนำฟัซซีโลจิกเข้าช่วยในระบบรู้จำทำให้ระบบสามารถตัดสินใจได้อย่างถูกต้องมากขึ้น

เนื่องจากระบบรู้จำลายมือแบบออฟไลน์ ไม่สามารถรู้ได้ว่าการลากเส้นของตัวอักษรมีลำดับอย่างไร ดังนั้นระบบจะต้องพิจารณาตัวอักษรจากโครงสร้าง ซึ่งตัวอักษรไทยแต่ละตัวมีลักษณะเฉพาะที่เป็นเอกลักษณ์ของตัวเอง ดังนั้นการเปรียบเทียบลักษณะเฉพาะที่เกิดขึ้น ทำให้สามารถแยกแยะตัวอักษรออกจากกันได้

ในระบบรู้จำตัวอักษรภาษาไทยตัวเดียวได้เสนอวิธีการรู้จำที่พิจารณาโครงสร้างหลักที่เป็นลักษณะเฉพาะของตัวอักษรภาษาไทย และเลือกตัวอักษรที่มีโครงสร้างใกล้เคียงกันที่สุดมาเป็นคำตอบ ซึ่งด้วยการประยุกต์นำเอาฟัซซีโลจิกมาบ่งบอกลักษณะของคุณลักษณะสำคัญทำให้สามารถยอมรับคุณลักษณะในรูปแบบที่หลากหลาย เนื่องจากตัวอักษรยังมีโครงสร้างหลักที่คงเดิม แต่ลักษณะเฉพาะที่ใช้พิจารณานั้นไม่ชัดเจนหรือเปลี่ยนรูปแบบไปเท่านั้น

ในระบบรู้จำคำในภาษาไทย การจัดลำดับของตัวอักษรในภาษาไทยจะพิจารณากันจากขนาดและตำแหน่งของตัวอักษรเป็นหลักซึ่งในบางครั้ง ขนาดที่เปลี่ยนไปตามแต่ลักษณะการเขียนของแต่ละคนทำให้ ระบบไม่สามารถจัดลำดับคำได้อย่างถูกต้อง

กระบวนการภายหลังที่ใช้ตรวจแก้ผลการรู้จำที่ผิดพลาดโดยการใช้ผลการรู้จำตัวอักษร 5 อันดับแรกมาเป็นตัวตรวจสอบทำให้สามารถเลือกคำตอบของการรู้จำได้ถูกต้อง แต่การตรวจสอบกับ

คำตอบทั้งหมดใช้เวลานาน จึงใช้วิธีการแบ่งกลุ่มตัวอักษรตามความยาวทำให้จำนวนการเปรียบเทียบข้อมูลลดลง และลดเวลาการทำงานของระบบลง

ข้อเสนอแนะเพื่องานวิจัยในอนาคต

1. การศึกษา และพัฒนากระบวนการภายหลังที่ใช้กับระบบรู้จำ โดยใช้ฐานข้อมูลที่มีจำนวนคำศัพท์มากขึ้น
2. ปรับปรุงระบบโดยให้มีการป้อนกลับผลการรู้จำ กับกระบวนการตัดคำ เพื่อตัดคำที่ติดกันในการเขียน เพื่อให้ระบบตัดสินใจในการตัดได้ถูกต้องมากขึ้น
3. พัฒนาและเปรียบเทียบ วิธีการรู้จำที่เสนอ กับวิธีการรู้จำที่ใช้เทคนิคอื่น ๆ เช่น ฮิดเดนมาร์คอฟ (Hidden Makov) และ นิวรอลเนทเวิร์ค (Neural Network)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รายการอ้างอิง

กิตติพงษ์ เจนวิถีสุข. การรู้จำอักษรพิมพ์ภาษาไทยโดยใช้นิวรอลเน็ตเวิร์ก และวิธีซินแทกติก. วิทยานิพนธ์ ปริญญามหาบัณฑิต, ภาควิชาวิศวกรรมไฟฟ้า, จุฬาลงกรณ์มหาวิทยาลัย, 2538.

เดชา รัตนาธาร. การรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้เทคนิคแบบฟัซซีโลจิก และวิธีซินแทกติก. วิทยานิพนธ์ ปริญญามหาบัณฑิต, ภาควิชาวิศวกรรมไฟฟ้า, จุฬาลงกรณ์มหาวิทยาลัย, 2538.

วิชา พานิช. ระบบการรู้จำรูปแบบอักษรภาษาไทย โดยใช้ลักษณะบ่งความต่างของตัวอักษร. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาวิศวกรรมไฟฟ้า, จุฬาลงกรณ์มหาวิทยาลัย, 2539.

สนธยา เมรินทร์. การศึกษาการรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้วิธีซินแทกติก. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาวิศวกรรมไฟฟ้า, จุฬาลงกรณ์มหาวิทยาลัย, 2537.

สุเจตน์ จันทรังษ์, ศุภชัย นำเกียตรติสกุล และ สุริยา วิทยาประดิษฐ์. การรู้จำอักษรลายมือเขียนภาษาไทย. การประชุมวิชาการวิศวกรรมไฟฟ้า ครั้งที่ 16 , 2536.

เสรี ปานซาง, สุรสิทธิ์ ราตรี และ ชม กิ้มปาน. การหารหัสเบื้องต้นของรูปร่างตัวพิมพ์อักษรภาษาไทย. การประชุมวิชาการวิศวกรรมไฟฟ้า, ครั้งที่ 11 เล่ม 1, 2531.

อนันต์ เอกวงค์วิริยะ. การศึกษาการรู้จำตัวเลขไทยแบบตัวพิมพ์โดยวิธีซินแทกติก. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาวิศวกรรมไฟฟ้า, จุฬาลงกรณ์มหาวิทยาลัย, 2537.

Airphaiboon, S.and Sangworasil, M. Recognition of Handwritten Thai Character Considering The Head of Character. IEEE Inter conference on Image Proceeding Singapore, 1989

Airphaiboon S., Sangworasil, M. and Kondo Off-line Handwritten Thai Characters from Word. Script Proceedings of the 12th IAPR International. Conference on Pattern Recognition vol. 2 ,1994

Ali, M.B.H. Background noise detection and cleaning in document images. IEEE Transactions on Volume: 11 , 1989

Bin Z., Hui, S. and Shaowei X. Document Analysis and Recognition. Proceedings of the Fourth International Conference on Vol 2, 1997

Bin Z.; Hui, S. and Shaowei X. A new method for segmenting unconstrained handwritten numeral string. Document Analysis and Recognition Proceedings of the Fourth International Conference on Vol 2 ,1997

Bozinovic, R.M. and Srihari, S.N. Off - line cursive script word recognition. Pattern Analysis and Machine Intelligence, IEEE Transactions on Vol 11 ,1989

Choruengwiwat P., Jitapunkul P., Wuttisittikulkij L. and Seehapan P. Distinctive Feature Analysis for Thai Handwritten Character Recognition based on Modified Stroke Changing Sequence. IEEE Asia-Pacific Conference on Circuit and Systems ,1998

Fong, H.S. and Yeung, D.S. An Extension of A Fuzzy Substroke Extractor. IEEE International Conference on Systems, Man, and Cybernetics Vol 5 ,1998

Gonzalez, R. C. and Wood,R.E. Digital Image Processing. 2 nd ed . USA : Addison Wesley, 1993.

Gyeonghwan K. and Govindaraju, V. A lexicon driven approach to handwritten word recognition for real-time applications Pattern Analysis and Machine Intelligence, IEEE Transactions on Vol19 ,1997

Hong, T.and Hull, J.J. Algorithms for postprocessing OCR results with visual inter-word constraints Image Processing International Conference on Vol 3, 1995

Joe C. H. P. and Gary M.T. M. and C.H. Chan. The application of fuzzy grammar to Handwritten character recognition, TENCON'93,IEEE Conference on Computer, Communication, Control and Power Engineering, 1993

Jung, H. C. and Gader, P.D Hybrid fuzzy-neural systems in handwritten word recognition Fuzzy Systems ,IEEE Transactionson Vol 5, 1997

Lee, S.-W., Kim, E.-S and Min, B.-W. Efficient postprocessing algorithms for error correction in handwritten Hangul address and human name recognition. Document Analysis and Recognition, Proceedings of the Second International Conference, 1993

Malaviya, A. Peters, L. and Theissinger, M. FOHDEL-a new fuzzy language for online handwriting recognition Fuzzy Systems, IEEE World Congress on Computational Intelligence, Proceedings of the Third IEEE Conference, 1994

Malaviya, A.. On-line handwriting Recognition with a Fuzzy Feature Description Language. Ph.D. Thesis, Computer Science Computer and Technology Technische University Berlin, 1996.

Malaviya, A.; Leja, C.; Peters, L Multi-script handwriting recognition with FOHDEL. Fuzzy

Information Processing Society, 1996. NAFIPS., 1996 Biennial Conference of the North American , 1996: 147 -151.

Marc P. and Rejean P. A Fuzzy-Syntactic Approach to Allograph Modeling for Cursive Script Recognition. IEEE Trans. On Pattern Analysis and Machine Intelligence vol.17 , 1995

Olivier de vil, Wangsuya S. and Coolmans D. On Thai Character Recognition. IEEE International Conference on Neural Networks, 1995

Paul D. Gader and James M. Keller. Application of Fuzzy Set Theory to Handwriting Recognition. ,IEEE World Congress on Computational Intelligence. Proceedings of the Third IEEE Conference on Fuzzy Systems, 1994

Phokharatkul P. and Kimpan C. Recognition of Handprinted Thai Characters Using the Cavity Features of Character Based on Neural Network, IEEE Asia-Pacific Conference on Circuit and Systems,1998

Poon, J.C.H. Man, G.M.T. and Chan, C.H. The application of fuzzy grammar to handwritten character recognition Computer, Communication, Control and Power Engineering. Proceedings. TENCON '93., 1993

Pyeoung K. K.. Improving handwritten numeral recognition using fuzzy logic TENCON '97. IEEE Region 10 Annual Conference. Speech and Image Technologies for Computing and Telecommunications., Proceedings of IEEE Volume: 2 , 1997

Schalkoff, R. J. Pattern Recognition:Statistical, Structural and Neural Approaches. 1 st ed .USA : John Wiley & Sons, 1992.

Sherkat, N. , Whitrow, R.J. and Evans, R.G. Wholistic recognition of handwriting using structural features Document Image Processing and Multimedia, IEE Colloquium on , 1999

Sun, C.-T. and Jang, J.S. A neuro-fuzzy classifier and its applications. Fuzzy Systems, Second IEEE International Conference, 1993

Tappert, C. Suen, Y. and Wakahara T. The state of art in on-line handwriting recognition.. IEEE Trans. Patt. Anal. Mach. Intell., vol.12, 1990

Wang, L.-X. and Mendel, J.M. A fuzzy approach to hand-written rotation-invariant character recognition Acoustics, Speech, and Signal Processing, 1992. ICASSP-

92, IEEE International Conference on vol 3 , 1992

Yan S. and Leedham, C.G. Noise and background removal from handwriting images. Intelligent Information Systems, IIS '97. Proceedings , 1997

Yuan Y. T., Lo-Ting T., Jiming L., Seong-Whan L.and Win-Win L. Off-line recognition of Chinese handwriting by multifeature and multilevel classification Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume: 20,1998

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก

สัญลักษณ์ของตัวแปรความเป็นสมาชิกของคุณลักษณะสำคัญ

คุณลักษณะสำคัญเชิงตำแหน่ง

ตัวแปร	สัญลักษณ์	หัวข้ออ้างอิง
top	_Top	3.3.1.2
middle(vertical)	_MidV	3.3.1.2
bottom	_Bot	3.3.1.2
left	_Left	3.3.1.2
middle(horizontal)	_MidH	3.3.1.2
right	_Rig	3.3.1.2

คุณลักษณะสำคัญของรูปทรง

ตัวแปร	สัญลักษณ์	หัวข้ออ้างอิง
thin	Thin	3.3.1.1
square	Square	3.3.1.1
wide	Wide	3.3.1.1

คุณลักษณะสำคัญของวงรอบ

ตัวแปร	สัญลักษณ์	หัวข้ออ้างอิง
Loop	Loop	3.3.1.3

คุณลักษณะสำคัญเชิงมุมของจุดปลาย

ตัวแปร	สัญลักษณ์	หัวข้ออ้างอิง
Up	EP_Up	3.3.1.4
Down	EP_D	3.3.1.4
Left	EP_Lft	3.3.1.4
Right	EP_Rgt	3.3.1.4

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

คุณลักษณะสำคัญของรอยหยัก

ตัวแปร	สัญลักษณ์	หัวข้ออ้างอิง
Up Direction	Curl_Up	3.3.1.5
Down Direction	Curl_Down	3.3.1.5
Left Direction	Curl_Left	3.3.1.5
Right Direction	Curl_Right	3.3.1.5

คุณลักษณะสำคัญของหางตัวอักษร

ตัวแปร	สัญลักษณ์	หัวข้ออ้างอิง
Tail	Tail	3.3.1.6

คุณลักษณะสำคัญของลำดับสายการลากผ่าน

ตัวแปร	สัญลักษณ์	หัวข้ออ้างอิง
SCS	SCS	3.3.1.7

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ข

โครงสร้างและกฎทางฟัซซีของตัวอักษรภาษาไทย

ในการอ้างสัญลักษณะในภาคผนวกขนี้จะใช้สัญลักษณ์ที่ได้กำหนดไว้แล้วในภาคผนวกกมา ประกอบซึ่งในการบ่งบอกถึงคุณลักษณะสำคัญที่ประกอบกันของค่าความเป็นสมาชิกจะให้สัญลักษณะดังนี้

คุณลักษณะสำคัญหัวของตัวอักษร

“หัว#ตำแหน่งในแนวนอนตำแหน่งในแนวตั้ง#ทิศทางของหัว”

เช่นหัวในตำแหน่งซ้ายบนมีทิศหันไปทางซ้ายเขียนเป็นสัญลักษณ์ได้เป็นLoop#_Left_Top#_L

คุณลักษณะสำคัญจุดปลายของตัวอักษร

“จุดปลาย#ตำแหน่งในแนวนอนตำแหน่งในแนวตั้ง#ทิศทางของจุดปลาย”

เช่นจุดปลายในตำแหน่งขวาบนมีทิศชี้ขึ้นเขียนเป็นสัญลักษณ์ได้เป็นEP#_Left_Top#_U

คุณลักษณะสำคัญหยักของตัวอักษร

“หยักจากทิศทาง#ตำแหน่งในแนวนอนตำแหน่งในแนวตั้ง”

เช่นหยักจากทางด้านบนในตำแหน่งกลางบนเขียนเป็นสัญลักษณ์ได้เป็นCurl_Up#_MidH_Top

คุณลักษณะสำคัญสายการลากผ่านของตัวอักษร

“สายการลากผ่าน#จากลักษณะสำคัญ#ทิศในการลาก#จำนวนการผ่านเส้น”

เช่นสายการลากผ่านจากวงรอบในทิศทางขวาตัดเส้น3ครั้งเขียนเป็นสัญลักษณ์ได้เป็น SCS#LP#_Rgt#3

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ก		
จุดปลายบริเวณซ้ายล่าง	EP_LB	EP#_Left_Bot#_D
จุดปลายบริเวณขวาล่าง	EP_RB	EP#_Rig_Bot#_D
รอยหยักจากด้านซ้ายบริเวณซ้ายบน	C_Left_LT	Curl_Lft#_Left_Top
รอยหยักจากด้านล่างบริเวณกลางบน	C_Down_MT	Curl_Down#_MidH_Top
โครงสร้างที่ไม่ต้องการ	Complement Element	- Curl_Up#_MidH_MidV - (Loop#_MidH_MidV#Right) V (Loop#_Left_MidV#Right)
Rule"ก1" : (EP_LB + EP_RB + C_LT + C_Down_MT + Complement Element)/4 Rule"ก2" : (EP_LB + EP_RB + C_Down_MT + Complement Element)/3		

ข		
รูปทรง	Thin	Thin
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_R) V (LP#_Left_Top#_L) V (EP#_Left_Top#_LT) V (EP#_Left_Top#_LT)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rig_Top#_U
รอยหยักจากด้านบนบริเวณกลางล่าง	C_Up_MB	(Curl_Up#_Mid_Top) V (Curl_Up#_Rig_Bot) V (Curl_Up#_Mid_Bot)
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	(SCS#LP#_R#2) V (SCS#LP#_R#3)
โครงสร้างที่ไม่ต้องการ	Complement Element	-(Curl_Up#_Rig_Bot) V (Curl_Up#_Rig_Top) V (Curl_Down#_Mid_Bot) V (EP#_Mid_Mid#_R)
Rule"ข" : (Thin + LP_LT + EP_RT + C_Up_MB + SCS_LP_R + Complement Element)/5		

ฃ		
รูปทรง	Thin	Thin
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_R) V (LP#_Left_Top#_L) V (LP#_Left_Mid#_R) V (EP#_Left_Top#_L)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rig_Top#_U
รอยหยักจากด้านบนบริเวณซ้ายบน	C_Up_LU	(Curl_Up#_Left_Top) V (Curl_Up#_Mid_Top)
รอยหยักจากด้านบนบริเวณกลางล่าง	C_Up_MB	(Curl_Up#_Rig_Top) V (Curl_Up#_Rig_Bot) V (Curl_Up#_Mid_Bot)
โครงสร้างที่ไม่ต้องการ	Complement Element	-(LP#_Left_Bot#_L)
Rule"ฃ" : (Thin + LP_LT + EP_RT + C_Up_LU + C_Up_MB + SCS_LP_R + Complement Element)/5		

ค		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_R) V (LP#_Lft_MidV#_R) V (EP#_MidH_MidV#_RT)
จุดปลายบริเวณขวาล่าง	EP_RB	EP#_Rig_Bot#_D
รอยหยักจากด้านล่างบริเวณกลางกลาง	C_Down_MM	(Curl_Down#_MidH_MidV) V (Curl_Down#_Lft_MidV) V (Curl_Down#_Rig_T)

โครงสร้างที่ไม่ต้องการ	Complement Element	-(Curl_Up#_Lft_Top) V (Curl_Up#_MidV_Top) V (Curl_Up#_Lft_MidH) V (Curl_Up#_Lft_Bot)
Rule"ค" : (LP_MM + EP_RB + C_Down_MM + Complement Element)/3		

ฅ		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_R) V (LP#_Lft_MidV#_R) V (EP#_MidH_MidV#_RT)
จุดปลายบริเวณขวาล่าง	EP_RB	EP#_Rig_Bot#_D
รอยหยักจากด้านบนบริเวณกลางบน	C_Up_MT	Curl_Up#_MidH_Top
รอยหยักจากด้านล่างบริเวณกลางกลาง	C_Down_MM	(Curl_Down#_MidH_MidV) V (Curl_Down#_Lft_MidV) V (Curl_Down#_Rig_T)
โครงสร้างที่ไม่ต้องการ	Complement Element	-(Curl_Down#_Rig_Top) V (Curl_Down#_Lft_MidH) V (Curl_Down#_MidV_MidH))
Rule"ฅ" : (LP_MM + EP_RB + C_Down_MM + Complement Element)/3		

ฆ		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Lft_Top#_L) V (LP#_Lft_MidV#_R) V (LP#_Lft_Top#_H)
หัวบริเวณซ้ายล่าง	LP_LB	(LP#_Lft_Bot#_L) V (LP#_MidH_Bot#_L)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rig_Top#_U
รอยหยักจากด้านบนบริเวณซ้ายบน	C_Up_LT	(Curl_Up#_Lft_Top) V (Curl_Up#_MidH_Top)
รอยหยักจากด้านบนบริเวณขวาล่าง	C_Down_LB	(Curl_Up#_Rig_Bot)

รอยหยักจากด้านล่างบริเวณกลางล่าง	C_Down_MB	(Curl_Down#_MidH_Bot) V (Curl_Down#_Rig_Bot)
รอยหยักจากด้านซ้ายบริเวณซ้ายกลาง	C_Left_LM	(Curl_Left#_MidH_MidV)
Rule"ฆ" : (LP_LT + LP_LB + EP_RT + C_Up_LT + C_Down_LB + C_Down_MB + C_Left_LM)/7		

ช		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Lft_MidV#_L) V (LP#_Lft_MidV#_H) V (EP#_Lft_MidV#_D)
หัวบริเวณกลางกลาง	LP_LB	(LP#_Lft_Bot#_L) V (LP#_MidH_Bot#_L)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rig_Top#_U
รอยหยักจากด้านบนบริเวณกลางกลาง	C_Up_LT	(Curl_Up#_Lft_Top) V (Curl_Up#_MidH_Top)
รอยหยักจากด้านขวาบริเวณขวากลาง	C_Up_RB	(Curl_Up#_Rig_Bot)
รอยหยักจากด้านซ้ายบริเวณกลางกลาง	C_Down_MB	(Curl_Down#_MidH_Bot) V (Curl_Down#_Rig_Bot)
หาง	Tail	Tail
โครงสร้างที่ไม่ต้องการ	Complement Element	-(Curl_Down#_Rig_Top) V (Curl_Down#_Lft_MidH) V (Curl_Down#_MidV_MidH))

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Rule"ช1" : (LP_LT + LP_LB + EP_RT + C_Up_LT + C_Up_MM + C_Down_MB + Tail + Complement Element)/7 Rule"ช2" : (LP_LT + EP_RT + C_Up_LT + C_Down_LB + C_Down_MB + Tail + Complement Element)/6 Rule"ช3" : (LP_LT + LP_LB + EP_RT + C_Up_LT + C_Down_MB + Tail + Complement Element)/6 Rule"ช4" : (LP_LT + EP_RT + C_Up_LT + C_Up_MM + Tail + Complement Element)/5

ช		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_MidV #_L) V (LP#_Left_Mid_V#_R) V (EP#_MidH_MidV#_LD)
หัวบริเวณกลางกลาง	LP_LB	LP#_MidH_MidV#_R
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#RU
รอยหยักจากด้านบนบริเวณกลางกลาง	C_Up_LT	(Curl_Up#_MidH_Top) V (Curl_Up#_Left_Top) V (Curl_Up#_Left_MidV) V (Curl_Up#_MidH_MidV)
รอยหยักจากด้านขวาบริเวณขวากลาง	C_Up_MM	(Curl_Up#_MidH_MidV) V (Curl_Up#_MidH_Bot)
รอยหยักจากด้านซ้ายบริเวณกลางกลาง	C_Right_MB	(Curl_Right#_Rig_MidV) V (Curl_Right#_MidH_MidV) V (Curl_Rgt#_Rgt_Top)
หาง	Tail	Tail
โครงสร้างที่ไม่ต้องการ	Complement Element	- ((Curl_Down#_Rig_Top) V (Curl_Down#_Lft_MidH) V (Curl_Down#_MidH_MidV))

Rule"ซ1" : (LP_LT + LP_LB + EP_RT + C_Up_LT + C_Up_MM + C_Right_MB + Tail + Complement Element)/7 Rule"ซ2" : (LP_LT + EP_RT + C_Up_LT + C_Down_LB + C_Right_MB + Tail + Complement Element)/6 Rule"ซ3" : (LP_LT + LP_LB + EP_RT + C_Up_LT + C_Right_MB + Tail + Complement Element)/6 Rule"ซ4" : (LP_LT + EP_RT + C_Up_LT + C_Up_MM + Tail + Complement Element)/5

ฌ		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Bot#R) V (EP#_Left_Bot#LU)
หัวบริเวณกลางกลาง	LP_LB	LP#_MidH_MidV#_R
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U
รอยหยักจากด้านบนบริเวณกลางกลาง	C_Up_LT	(Curl_Left#_Left_Top) V (Curl_Left#_Left_MidV)
รอยหยักจากด้านขวาบริเวณขวากลาง	C_Down_LB	Curl_Up#_Rig_Bot
รอยหยักจากด้านซ้ายบริเวณกลางกลาง	C_Down_MB	(Curl_Down#_Rgt_Bot) V (Curl_Down#_Lft_MidV)
Rule"ฌ" : (LP_LT + LP_LB + EP_RT + C_Up_LT + C_Down_LB + C_Down_MB)/6		

ญ		
หัวบริเวณซ้ายล่าง	EP_LB	EP#_Rgt_Top#U
รอยหยักจากด้านล่างบริเวณซ้ายบน	C_Down_LT	(Curl_Down#_Left_Top) V (Curl_Down#_MidH_Top)
รอยหยักจากด้านซ้ายบริเวณซ้ายบน	C_Left_LT	(Curl_Left#_Left_Top) V (Curl_Left#_Left_MidV) V (Curl_Down#_MidH_Top)

รอยหยักจากด้านซ้ายบริเวณซ้ายล่าง	C_Left_LB	(LP#_Left_Bot#R) V (EP#_Left_Bot V #LU)
รอยหยักจากด้านบนบริเวณกลางกลาง	C_Up_MM	(Curl_Up#_MidH_MidV) V (Curl_Up#_Right_MidV)
จุดปลายบริเวณขวาล่าง	EP_RB	(EP#_Rgt_Bot#RU) V (EP#_Rgt_MidV#_RT)
หัวบริเวณซ้ายล่าง	LP_LB	(LP#_Left_Bot#_R) V (EP#_Left_Bot#RU)
หัวบริเวณซ้ายกลาง	LP_LM	(LP#_Left_MidV#_R) V (EP#_Left_MidV#RU)
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#R) V (LP#_MidH_MidV#_R)
รอยหยักจากด้านล่างบริเวณขวาบน	C_Down_MB	(EP#_Rgt_Bot#RU) V (EP#_Rgt_MidV#_RT)
Rule"ญ1" : (LP_LM + EP_LB + C_Down_LT + C_Left_LT + C_Down_MB + EP_RB)/6 Rule"ญ2" : (LP_LB + EP_LB + C_Down_LT + C_Left_LT + C_Down_MB)/5		

ฎ		
หัวบริเวณซ้ายกลาง	LP_LM	(LP#_Left_MidV#_L) V (LP#_MidH_MidV#_L) V (EP#_MidH_MidV#L) V (EP#_Left_MidV#L)
หัวบริเวณซ้ายล่าง	LP_LB	LP#_Left_Bot#_L
รูปทรง	Thin	Thin
รอยหยักจากด้านซ้ายบริเวณซ้ายบน	C_Left_LT	(Curl_Left#_Left_Top) V (Curl_Down#_MidH_Top)
รอยหยักจากด้านซ้ายบริเวณขวาล่าง	C_Left_RB	(Curl_Left#_Rig_Bot) V (Curl_Left#_Rig_MidV)

จุดปลายบริเวณขวากลาง	EP_RM	(EP#_Rgt_MidV#_RT) V (EP#_MidH_Bot#_RT) V (EP#_Rgt_Bot#RU)
รอยหยักจากด้านล่างบริเวณซ้ายกลาง	C_Down_LM	Curl_Down#_Lft_MidV
โครงสร้างที่ไม่ต้องการ	Complement Element	-(Curl_Down#_Rgt_Bot) V (Curl_Down#_Lft_MidV)
Rule"ฎ1" : (LP_LM + LP_LB + Thin + EP_RM + C_Left_LT + C_Down_LM + Complement Element)/6 Rule"ฎ2" : (LP_LM + LP_LB + Thin + EP_RM + C_Left_LT + C_Left_RB + C_Down_LM + Complement Element)/7		

ฎ		
หัวบริเวณซ้ายกลาง	LP_LM	(LP#_Left_MidV#_L) V (LP#_MidH_MidV#_L) V (EP#_MidH_MidV#L) V (EP#_Left_MidV#L)
หัวบริเวณซ้ายล่าง	LP_LB	LP#_Left_Bot#_L
รูปทรง	Thin	Thin
รอยหยักจากด้านซ้ายบริเวณซ้ายบน	C_Left_LT	(Curl_Left#_Left_Top) V (Curl_Down#_MidH_Top)
รอยหยักจากด้านล่างบริเวณขวาล่าง	C_Down_RB	(Curl_Down#_Rgt_Bot) V (Curl_Down#_Lft_MidV)
รอยหยักจากด้านล่างบริเวณซ้ายกลาง	C_Down_LM	(Curl_Down#_Lft_MidV) V (Curl_Down#_Lft_Bot)
จุดปลายบริเวณขวากลาง	EP_RM	(EP#_Rgt_MidV#_RT) V (EP#_MidH_Bot#_RT) V (EP#_Rgt_Bot#RU)

Rule"ฏ1" : (LP_LM + LP_LB + Thin + EP_RM + C_Left_LT + C_Left_RB + C_Down_LM + Complement Element)/7 Rule"ฏ2" : (LP_LM + LP_LB + EP_RM + C_Left_LT + C_Left_RB + C_Down_LM + Complement Element)/6 Rule"ฏ3" : (LP_LM + LP_LB + C_Left_LT + C_Left_RB + C_Down_LM + Complement Element)/5

จ		
หัวบริเวณกลางบน	LP_MT	(LP#_MidH_Top#L) V (LP#_MidH_MidV#_L)
หัวบริเวณขวาล่าง	LP_RB	LP#_Rig_Bot#_L
หัวบริเวณซ้ายล่าง	LP_LB	LP#_Left_Bot#_L
จุดปลายบริเวณกลางล่าง	EP_MB	EP#_MidH_Bot#_RT
จุดปลายบริเวณขาวบน	EP_RT	EP#_Rgt_Top#RU
รอยหยักจากด้านบนบริเวณขวาบน	C_Up_RT	Curl_Up#_Rgt_Top
รอยหยักจากด้านขวาบริเวณขวาบน	C_Right_RT	(Curl_Rgt#_Rgt_Top) V (Curl_Rgt#_MidH_Top) V (Curl_Rgt#_Lft_Top)
รอยหยักจากด้านล่างบริเวณซ้ายล่าง	C_Down_LB	(Curl_Down#_Lft_Bot) V (Curl_Down#_Lft_MidV)
รอยหยักจากด้านล่างบริเวณซ้ายกลาง	C_Down_LM	(Curl_Down#_Lft_MidV) V (Curl_Down#_Rgt_Bot)
Rule" จ1 " : (LP_MT + LP_RB + LP_LB + EP_MB + EP_RT + C_Up_RT + C_Right_RT + C_Down_LB + C_Down_LM)/9 Rule" จ2 " : (LP_MT + LP_RB + LP_LB + EP_RT + C_Up_RT + C_Right_RT + C_Down_LB + C_Down_LM)/8 Rule" จ3 " : (LP_MT + LP_RB + LP_LB + EP_RT + C_Right_RT + C_Down_LB + C_Down_LM)/7		

ฑ		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_L) V (LP#_Left_Top#_R) V (LP#_Left_MidV#_R)
จุดปลายบริเวณซ้ายล่าง	EP_LB	(EP#_Left_Bot#_D) V (EP#_MidH_Bot#_D)
จุดปลายบริเวณขวาล่าง	EP_RB	(EP#_Rig_Bot#_D) V (EP#_MidH_Bot#_D)
รอยหยักจากด้านบนบริเวณซ้ายบน	C_Up_LT	(Curl_Up#_Left_Top) V (Curl_Up#_MidH_Top)
รอยหยักจากด้านบนบริเวณกลางกลาง	C_Up_MM	(Curl_Up#_MidH_MidV)
รอยหยักจากด้านล่างบริเวณกลางบน	C_Down_MT	(Curl_Down#_MidH_Top) V Curl_Down#_Rig_Top) V (Curl_Down#_MidH_Top) V Curl_Down#_Rig_Top)
รอยหยักจากด้านขวาบริเวณขวาบน	C_Right_RT	(Curl_Rgt#_Rgt_Top) V (Curl_Rgt#_MidH_Top) V (Curl_Rgt#_Lft_Top)
รอยหยักจากด้านล่างบริเวณซ้ายล่าง	C_Down_LB	(Curl_Down#_Lft_Bot) V (Curl_Down#_Lft_MidV)
Rule” ฑ ” : (LP_LT + EP_LB + EP_RB + C_Up_LT + C_Up_MM + C_Down_MT + C_Right_RT + C_Down_LB)/7		

ฒ		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_L) V (LP#_Left_MidV#_L)(EP#_Left_MidV#RU) V (EP#_MidH_MidV#_RU)
หัวบริเวณกลางล่าง	LP_MB	LP#_MidH_Bot#L
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U
รอยหยักจากด้านบนบริเวณกลางบน	C_Up_MT	Curl_Up#_MidH_Top
รอยหยักจากด้านล่างบริเวณซ้ายกลาง	C_Down_LM	Curl_Down#_Lft_MidV
Rule” ฒ ” : (LP_MM + LP_MB + EP_RT + C_Up_MT + C_Down_LM)/5		

ณ		
หัวบริเวณซ้ายล่าง	LP_LT	(LP#_Left_Bot#R) V (EP#_Left_Bot V #LU)
หัวบริเวณขวาล่าง	LP_RB	LP#_Right_Bot#_R
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U
รอยหยักจากด้านบนบริเวณกลางล่าง	C_Up_MB	(Curl_Up#_MidH_Bot) V (Curl_Up#_MidH_MidV)
รอยหยักจากด้านล่างบริเวณกลางบน	C_Down_MT	(Curl_Down#_MidH_Top) V (Curl_Down#_Left_Top)
รอยหยักจากด้านซ้ายบริเวณซ้ายบน	C_Left_LT	(Curl_Left#_Left_Top) V (Curl_Left#_Left_MidV)
Rule” ณ ” : (LP_LT + LP_RB + EP_RT + C_Up_MB + C_Down_MT + C_Left_LT)/6		

ด		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_L) V (LP#_Left_MidV#_L) V (LP#_MidH_Top#L) V (EP#_MidH_MidV#_RU)
จุดปลายบริเวณขวาล่าง	EP_RB	EP#_Rig_Bot#_D
รอยหยักจากด้านล่างบริเวณขวาบน	C_Down_RT	(Curl_Down#_Rig_Top) V (Curl_Down#_MidH_Top)
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	(SCS#LP#_R#2)
โครงสร้างที่ไม่ต้องการ	Complement Element	- ((Curl_Up#_MidH_Top) V (Curl_Up#_Left_Top))
Rule” ด ” : (LP_MM + EP_RB + C_Down_RT + SCS_LP_R + Complement Element)/4		

ต		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_L) V (LP#_Left_MidV#_L) V (EP#_MidH_MidV#_RU)
จุดปลายบริเวณขวาล่าง	EP_RB	EP#_Rig_Bot#_D
รอยหยักจากด้านล่างบริเวณขวาบน	C_Down_RT	Curl_Down#_Rig_Top
รอยหยักจากด้านบนบริเวณกลางบน	C_Up_MT	(Curl_Up#_MidH_Top) V (Curl_Up#_Left_Top)
Rule” ต ” : (LP_MM + EP_RB + C_Down_RT + C_Up_MT)/4		

ถ		
หัวบริเวณซ้ายล่าง	LP_LB	(LP#_Left_Bot#R) V (LP#_MidH_MidV#_R) V (EP#_Left_Bot#_R)
จุดปลายบริเวณกลางกลาง	EP_RB	EP#_Rig_Bot#_D

รอยหยักจากด้านซ้ายบริเวณขวาบน	C_Left_LT	(Curl_Left#_Left_Top) V (Curl_Down#_MidH_Top) V (Curl_Left#_Left_MidV) V Curl_Left#_MidH_MidV)
รอยหยักจากด้านล่างบริเวณกลางบน	C_Down_MT	Curl_Down#_MidH_Top
Rule” ถ1 ” : (LP_LB + EP_RB + C_Left_LT + C_Down_MT)/4 Rule” ถ2 ” : (LP_LB + EP_RB + C_Down_MT)/3		

ท		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_L) V (EP#_Left_Top#_LT)
จุดปลายบริเวณซ้ายล่าง	LP_LB	(EP#_Left_Bot#_D) V (EP#_MidH_Bot#_D)
จุดปลายบริเวณขวาล่าง	EP_RT	EP#_Rig_Bot#_D
รอยหยักจากด้านบนบริเวณซ้ายกลาง	C_Up_LM	(Curl_Up#_Left_MidV) V (Curl_Up#_Left_Bot) V (Curl_Up#_MidH_MidV) V (Curl_Up#_MidH_Bot)
รอยหยักจากด้านล่างบริเวณกลางบน	C_Down_MT	(Curl_Down#_MidH_Top) V (Curl_Down#_Rig_Top) V (Curl_Down#_MidH_Top) V (Curl_Down#_Rig_Top)
โครงสร้างที่ไม่ต้องการ	Complement Element	-((Curl_Up#_Rig_Bot) V Curl_Up#_MidH_Bot))
Rule” ท ” : (LP_LT + LP_LB + EP_RT + C_Up_LM + C_Down_MT + Complement Element)/5		

ธ		
หัวบริเวณกลางกลาง	LP_MM	LP#_MidH_MidV#_L
จุดปลายบริเวณซ้ายกลาง	EP_LM	EP#_Left_MidV#_U
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#RU
รอยหยักจากด้านซ้ายบริเวณขวากลาง	C_Left_RM	Curl_Left#_Rig_MidV
รอยหยักจากด้านขวาบริเวณซ้ายบน	C_Right_LT	Curl_Rgt#_Lft_Top
รอยหยักจากด้านบนบริเวณขวาบน	C_Up_RT	Curl_Up#_Rgt_Top
ลำดับสายลากผ่านของจุดปลายไปทางด้านบน	SCS_EP_U	SCS#_EP#_U#2
Rule” ธ1 ” : (EP_LM + EP_RT + C_Left_RM + C_Right_LT + C_Up_RT + SCS_EP_U)/6 Rule” ธ2 ” : (LP_MM + EP_RT + C_Left_RM + C_Right_LT + C_Up_RT + SCS_EP_U)/6 Rule” ธ3 ” : (EP_LM + EP_RT + C_Left_RM + C_Right_LT + SCS_EP_U)/5 Rule” ธ4 ” : (LP_MM + EP_RT + C_Left_RM + C_Right_LT + SCS_EP_U)/5		

น		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_L) V (EP#_Left_Top#_LT)
หัวบริเวณขวาล่าง	LP_RB	LP#_Right_Bot#_R
จุดปลายบริเวณขวาบน	EP_RT	(EP#_Rgt_Top#U) V (EP#_MidH_Top#_U)
รอยหยักจากด้านบนบริเวณซ้านล่าง	C_Up_LB	(Curl_Up#_Left_Bot) V (Curl_Up#_MidH_Bot)
Rule” น ” : (LP_LT + LP_RB + EP_RT + C_Up_LB)/4		

ป		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_L) V (LP#_Left_Top#_R)(EP#_Left_Top#_LT)

หัวบริเวณกลางกลาง	EP_RT	EP#_Rgt_Top#U
จุดปลายบริเวณขวาบน	C_Up_MB	(Curl_Up#_MidH_Bot) V (Curl_Up#_Rig_Bot)
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	SCS#_LP#_R#2
ลำดับสายลากผ่านของหัวไปทางด้านบน	SCS_LP_U	SCS#_LP#_U#1
รูปทรง	Square	(Square) V (Wide)
โครงสร้างที่ไม่ต้องการ	Complement Element	-((EP#_Left_MidV#_R) V (EP#_MidH_MidV#_R) V (Curl_Left#_Left_MidV) V (LP#_Right_Bot#_R) V (Curl_Down#_MidH_Top) V (Curl_Down#_MidH_Top) V (LP#_Rig_MidV#_L))
Rule" บ " : (LP_LT + EP_RT + C_Up_MB + SCS_LP_R + SCS_LP_U + Square + Complement Element)/6		

ป		
หัวบริเวณซ้ายกลาง	LP_LM	(LP#_Left_MidV#_L) V (EP#_Left_MidV#_LU)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U
รอยหยักจากด้านบนบริเวณกลางล่าง	C_Up_MB	Curl_Up#_MidH_Bot
หาง	Tail	Tail
โครงสร้างที่ไม่ต้องการ	Complement Element	-Curl_Up#_Rig_Bot
Rule" ป " : (LP_LM + EP_RT + C_Up_MB + Tail + Complement Element)/4		

ผ		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_R) V (LP#_MidH_Top#_L) V (EP#_Left_Top#_RU) V (EP#_MidH_Top#_RU)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U
รอยหยักจากด้านบนบริเวณกลางกลาง	C_Up_MM	(Curl_Up#_MidH_MidV) V (Curl_Up#_MidH_Bot)
รอยหยักจากด้านบนบริเวณขวาล่าง	C_Up_RB	(Curl_Up#_Rig_Bot) V (Curl_Up#_Rig_Bot)
รอยหยักจากด้านบนบริเวณล่างกลาง	C_Down_LM	(Curl_Down#_Lft_MidV) V (Curl_Down#_Rig_Bot) V (Curl_Down#_MidH_MidV)
ลำดับสายลากผ่านของหัวไปทางด้านล่าง	SCS_LP_D	SCS#_LP#_D#2
โครงสร้างที่ไม่ต้องการ	Complement Element	-(Curl_Up#_Left_Top) V (LP#_Left_Bot#_L)
Rule” ผ ” : (LP_LT + EP_RT + C_Up_MM + C_Up_RB + C_Down_LM + SCS_LP_D + Complement Element)/6 Rule” ผ ” : (LP_LT + EP_RT + C_Up_MM + C_Down_LM + SCS_LP_D + Complement Element)/5		

ฝ		
หัวบริเวณซ้ายกลาง	LP_LM	(LP#_Left_MidV#_R) V (EP#_Left_MidV#_R)
จุดปลายบริเวณกลางกลาง	EP_MM	(EP#_MidH_MidV#_U) EP#_MidH_Bot#_U)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U

รอยหยักจากด้านบนบริเวณซ้ายล่าง	C_Up_LB	Curl_Up#_Left_Bot
รอยหยักจากด้านบนบริเวณขวาล่าง	C_Up_RB	(Curl_Up#_Rig_Bot) V (Curl_Up#_Rig_Bot)
รอยหยักจากด้านล่างบริเวณซ้ายกลาง	C_Down_LM	Curl_Down#_Lft_MidV
หาง	Tail	Tail
Rule" ฝ1 " : (LP_LM + EP_MM + EP_RT + C_Up_LB + C_Up_RB + C_Down_LM + Tail)/7 Rule" ฝ2 " : (LP_LM + EP_RT + C_Up_LB + C_Up_RB + C_Down_LM + Tail)/6 Rule" ฝ3 " : (LP_LM + EP_RT + C_Up_RB + C_Down_LM + Tail)/5		

พ		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_L)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U
รอยหยักจากด้านบนบริเวณซ้ายล่าง	C_Up_LB	Curl_Up#_Left_Bot
รอยหยักจากด้านบนบริเวณขวากลาง	C_Up_RB	(Curl_Up#_Rig_Bot) V (Curl_Up#_Rig_Bot)
รอยหยักจากด้านล่างบริเวณซ้ายกลาง	C_Down_LM	Curl_Down#_Lft_MidV
Rule" พ " : (LP_LT + EP_RT + C_Up_LB + C_Up_RB + C_Down_LM)/5		

ฟ		
หัวบริเวณซ้ายบน	LP_LM	(LP#_Left_MidV#_L) V (EP#_Left_MidV#_LU)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U
จุดปลายบริเวณซ้ายล่าง	EP_LB	EP#_Left_Bot#_D
รอยหยักจากด้านบนบริเวณซ้ายล่าง	C_Up_LB	(Curl_Up#_Left_Bot) V (Curl_Up#_MidH_Bot)
รอยหยักจากด้านบนบริเวณขวาล่าง	C_Up_RB	Curl_Up#_Rig_Bot

รอยหยักจากด้านล่างบริเวณกลางกลาง	C_Down_MM	(Curl_Down#_MidH_MidV) V (Curl_Down#_MidH_MidV) V (Curl_Down#_Lft_MidV) V (Curl_Down#_Lft_MidV)
หาง	Tail	Tail
Rule" ฟ1 " : (LP_LM + EP_RT + EP_LB + C_Up_LB + C_Up_RB + C_Down_MM + Tail)/7 Rule" ฟ2 " : (LP_LM + EP_RT + C_Up_LB + C_Up_RB + C_Down_MM + Tail)/6		

ภ		
หัวบริเวณซ้ายล่าง	LP_LB	(LP#_Left_Bot#_L) V (EP#_Left_Bot#_R)
จุดปลายบริเวณขวาล่าง	EP_RB	EP#_Rig_Bot#_D)
รอยหยักจากด้านซ้ายบริเวณซ้ายบน	C_Left_LT	(Curl_Left#_Left_Top) V (Curl_Down#_MidH_Top) V (Curl_Left#_Left_MidV) V (Curl_Left#_MidH_MidV)
รอยหยักจากด้านล่างบริเวณกลางบน	C_Down_MT	(Curl_Down#_MidH_Top) V (Curl_Down#_MidH_MidV)
Rule" ภ1 " : (LP_LB + EP_RB + C_Left_LT + C_Down_MT)/4 Rule" ภ2 " : (LP_LB + EP_RB + C_Down_MT)/3		

ม		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_L) V (LP#_MidH_Top#L)(EP#_Left_Top#_LT)
หัวบริเวณซ้ายล่าง	LP_LB	(LP#_Left_Bot#_L) V (LP#_Left_MidV#_L)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U
รอยหยักจากด้านบนบริเวณขวาล่าง	C_Up_RB	Curl_Up#_Rig_Bot

รอยหยักจากด้านบนบริเวณกลางกลาง	C_Left_MM	(Curl_Left#_MidH_MidV) V (Curl_Left#_Left_MidV)
รอยหยักจากด้านล่างบริเวณซ้ายกลาง	C_Down_LM	(Curl_Down#_Lft_MidV) V (Curl_Down#_MidH_MidV) V (Curl_Left#_Left_Bot)
โครงสร้างที่ไม่ต้องการ	Complement Element	-2*LP#_Right_Bot#_R)- Curl_Down#_Left_Top)- Curl_Up#_Right_MidV)
Rule" ม " : (LP_LT + LP_LB + EP_RT + C_Up_RB + C_Left_MM + C_Down_LM + Complement Element)/6		

ย		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_R) V (LP#_MidH_Top#_L) V (EP#_Left_Top#_R)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U
จุดปลายบริเวณซ้ายกลาง	EP_LM	(EP#_Left_MidV#_R) V (EP#_MidH_MidV#_R)
รอยหยักจากด้านบนบริเวณกลางล่าง	C_Up_MB	Curl_Up#_MidH_Bot
รอยหยักจากด้านซ้ายบริเวณซ้ายกลาง	C_Left_LM	Curl_Left#_Left_MidV
ลำดับสายลากผ่านของหัวไปทางด้านขวา	SCS_LP_R	SCS#_LP#_R#2
Rule" ย1 " : (LP_LT + EP_RT + EP_LM + C_Up_MB + C_Left_LM + SCS_LP_R)/6 Rule" ย2 " : (LP_LT + EP_LM + C_Up_MB + C_Left_LM + SCS_LP_R)/5 Rule" ย3 " : (LP_LT + EP_RT + EP_LM + C_Up_MB + C_Left_LM)/5 Rule" ย4 " : (LP_LT + EP_LM + C_Up_MB + C_Left_LM)/4		

ร		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#R) V (LP#_Rig_Bot#_L)(EP#_MidH_Bot#_L)
จุดปลายบริเวณขวาบน	EP_RT	(EP#_Rig_Top#_R) V EP#_Rgt_Top#RU)
รอยหยักจากด้านบนบริเวณกลางบน	C_Up_MT	(Curl_Up#_MidH_Top) V (Curl_Up#_Rgt_Top)
รอยหยักจากด้านซ้ายบริเวณกลางกลาง	C_Left_MM	(Curl_Left#_MidH_MidV) V (Curl_Left#_Rig_MidV) V (Curl_Left#_Rig_Bot)
รอยหยักจากด้านขวาบริเวณซ้ายบน	C_Right_LT	(Curl_Rgt#_Lft_Top) V (Curl_Rgt#_MidH_Top) V (Curl_Right#_Left_MidV)
ลำดับสายลากผ่านของหัวไปทางด้านบน	SCS_LP_U	SCS#_LP#_U#3
โครงสร้างที่ไม่ต้องการ	Complement Element	-(LP#_Left_Top#_R) V (LP#_Left_MidV#_R) (LP#_MidH_MidV#_L)
Rule” ร1 ” : (LP_MM + EP_RT + C_Up_MT + C_Left_MM + SCS_LP_U + Complement Element)/5 Rule” ร2 ” : (LP_MM + EP_RT + C_Left_MM + SCS_LP_U + Complement Element)/4		

ล		
หัวบริเวณซ้ายล่าง	LP_LB	(LP#_Left_Bot#R) V (LP#_MidH_MidV#_R)
จุดปลายบริเวณซ้ายบน	EP_LT	(EP#_Left_Top#_LD) V (EP#_Rig_MidV#_LD)

รอยหยักจากด้านซ้ายบริเวณขวาบน	C_Left_RT	(Curl_Left#_Rig_Top) V (Curl_Left#_Rig_MidV)
รอยหยักจากด้านล่างบริเวณกลาง กลาง	C_Down_MM	(Curl_Down#_MidH_MidV) V (Curl_Down#_Rig_MidV) V (Curl_Down#_MidH_MidV) V (Curl_Down#_Rig_MidV)
โครงสร้างที่ไม่ต้องการ	Complement Element	-Tail
Rule” ล ” : (LP_LB + EP_LT + C_Left_RT + C_Down_MM + Complement Element)/4		

ว		
หัวบริเวณขวาล่าง	LP_RB	(LP#_Rig_Bot#_L) V (LP#_MidH_MidV#R) V (EP#_Rig_Bot#_L)
จุดปลายบริเวณขวากลาง	EP_RM	(EP#_Rig_MidV#_LD) V (EP#_Left_Top#_LD)
รอยหยักจากด้านซ้ายบริเวณขวาบน	C_Left_RT	(Curl_Left#_Rig_Top) V (Curl_Left#_Rig_MidV)
ลำดับสายลากผ่านของหัวไปทาง ด้านบน	SCS_LP_U	SCS#_LP#_U#2
Rule” ว ” : (LP_RB + EP_RM + C_Left_RT + SCS_LP_U)/4		

ศ		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_R) V (LP#_Left_MidV#_R) V (EP#_MidH_MidV#_R)
จุดปลายบริเวณขวาล่าง	EP_RB	EP#_Rig_Bot#_D
จุดปลายบริเวณซ้ายล่าง	EP_LB	EP#_Left_Bot#_D

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#RU
รอยหยักจากด้านล่างบริเวณขวาบน	C_Down_RT	(Curl_Down#_Rig_Top) V (Curl_Down#_MidH_MidV) V (Curl_Down#_Rig_Top) V (Curl_Down#_Rig_MidV)
รอยหยักจากด้านล่างบริเวณกลางบน	C_Down_MT	(Curl_Down#_MidH_Top) V (Curl_Down#_MidH_Top)
รอยหยักจากด้านขวาบริเวณขวากลาง	C_Right_RM	(Curl_Right#_Rig_MidV) V (Curl_Right#_MidH_MidV) V (Curl_Rgt#_Rgt_Top)
หาง	Tail	Tail
โครงสร้างที่ไม่ต้องการ	Complement Element	-LP#_Right_Bot#_R)- LP#_MidH_MidV#_R)– LP#_Left_Bot#R)-LP#_Left_Top#_L)
Rule” ศ1 ” : (LP_MM + EP_LB + EP_RT + EP_RB + C_Down_RT + C_Right_RM + Tail + Complement Element)/7 Rule” ศ2 ” : (LP_MM + EP_RT + EP_RB + C_Down_RT + C_Right_RM + Tail + Complement Element)/6 Rule” ศ3 ” : (LP_MM + EP_RT + EP_RB + C_Right_RM + Tail + Complement Element)/5 Rule” ศ4 ” : (LP_MM + EP_RT + EP_RB + C_Down_RT + Tail + Complement Element)/5		

ษ		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_L) V (EP#_Left_Top#_L)
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_L) V (LP#_MidH_Top#L)

จุดปลายบริเวณขวากลาง	EP_RM	(EP#_Rgt_MidV#_RT) V (EP#_Rig_MidV#_R) V (EP#_Rig_Top#_R)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U
รอยหยักจากด้านบนบริเวณกลางล่าง	C_Up_MB	Curl_Up#_MidH_Bot
รอยหยักจากด้านบนบริเวณขวากลาง	C_Up_RM	(Curl_Up#_Right_MidV) V (CurlFound (23)
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	SCS#_LP#_R#3
โครงสร้างที่ไม่ต้องการ	Complement Element	-((LP#_Right_Bot#_R) V (LP#_Left_MidV#_L) V (LP#_MidH_Top#L) V (Curl_Up#_Left_Top))
Rule” ษ ” : (LP_LT + LP_MM + EP_RM + EP_RT + C_Up_RM + C_Up_MB + SCS_LP_R + Complement Element)/7		

ส		
หัวบริเวณซ้ายบน	LP_LB	(LP#_Left_Bot#R) V (LP#_MidH_MidV#_R) (EP#_MidH_Bot#_R) V (EP#_Left_Bot#_R) V (EP#_Left_Bot#_D)
หัวบริเวณกลางกลาง	EP_RM	(EP#_Rig_MidV#_LD) V (EP#_Left_Top#_LD)
จุดปลายบริเวณขวาบน	EP_RT	(EP#_Rgt_Top#RU) V (EP#_MidH_Top#_LU) V (EP#_MidH_Top#_LT)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รอยหยักจากด้านบนบริเวณขวาบน	C_Up_RT	(Curl_Up#_Rgt_Top) V (Curl_Up#_MidH_Top)
รอยหยักจากด้านซ้ายบริเวณขวากลาง	C_Left_RM	Curl_Left#_Rig_MidV
รอยหยักจากด้านขวาบริเวณขวากลาง	C_Right_RM	(Curl_Right#_Rig_MidV) V (Curl_Right#_MidH_MidV)
หาง	Tail	Tail
โครงสร้างที่ไม่ต้องการ	Complement Element	-((Curl_Up#_Rig_Bot) V (Curl_Up#_MidH_Bot))
Rule" ส1 " : (LP_LB + EP_RM + EP_RT + C_Up_RT + C_Left_RM + C_Right_RM + Tail + Complement Element)/7 Rule" ส2 " : (LP_LB + EP_RM + EP_RT + C_Left_RM + C_Right_RM + Tail + Complement Element)/6 Rule" ส3 " : (LP_LB + EP_RM + EP_RT + C_Right_RM + Tail + Complement Element)/5 Rule" ส4 " : (LP_LB + EP_RM + EP_RT + C_Right_RM + Complement Element)/4		

ห		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_L) V (EP#_Left_Top#_LD)
หัวบริเวณขวาบน	LP_RT	(LP#_Rig_Top#_L) V (LP#_Rig_Top#_R)
จุดปลายบริเวณซ้ายล่าง	EP_LB	(EP#_Left_Bot#_D) V (EP#_MidH_Bot#_D)
จุดปลายบริเวณขวาล่าง	EP_RB	EP#_Rig_Bot#_D
รอยหยักจากด้านบนบริเวณซ้ายล่าง	C_Up_LB	(Curl_Up#_Left_Bot) V (Curl_Up#_MidH_Bot)
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	SCS#_LP#_R#3

โครงสร้างที่ไม่ต้องการ	Complement Element	-(Curl_Up#_Rig_Bot)-Curl_Up#_MidH_Bot)
Rule" ห1 " : (LP_LT + LP_RT + EP_LB + EP_RB + SCS_LP_R + Complement Element)/5 Rule" ห2 " : (LP_LT + LP_RT + EP_LB + EP_RB + C_Up_LB + Complement Element)/5		

ฟ		
หัวบริเวณซ้ายกลาง	LP_LM	(LP#_Left_MidV#_L) V (EP#_Rig_MidV#_LD)
หัวบริเวณกลางบน	LP_MT	LP#_MidH_Top#L
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#RU
จุดปลายบริเวณขวาล่าง	EP_RB	EP#_Rig_Bot#_D
ลำดับสายลากผ่านของหัวไปทางด้านล่าง	SCS_LP_D	SCS#_LP#_D#2
รอยหยักจากด้านบนบริเวณซ้ายล่าง	C_Up_LB	Curl_Up#_Left_Bot
รอยหยักจากด้านล่างบริเวณกลางกลาง	C_Down_MM	(Curl_Down#_MidH_MidV) V (Curl_Down#_Lft_MidV) V (Curl_Down#_MidH_MidV) V (Curl_Down#_Lft_MidV)
Rule" ฟ1 " : (LP_LM + LP_MT + EP_RT + EP_RB + C_Up_LB + C_Down_MM + SCS_LP_D)/7 Rule" ฟ2 " : (LP_LM + LP_MT + EP_RT + C_Up_LB + C_Down_MM + SCS_LP_D)/6 Rule" ฟ3 " : (LP_LM + LP_MT + EP_RT + C_Up_LB + SCS_LP_D)/5		

อ		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_R) V (LP#_Left_MidV#_R) (EP#_MidH_MidV#_RU) V (EP#_MidH_MidV#_R)

จุดปลายบริเวณซ้ายบน	EP_LT	(EP#_Left_Top#_LD) V (EP#_Rig_MidV#_LD) V (EP#_Left_Bot#_LD)
รอยหยักจากด้านซ้ายบริเวณขวาบน	C_Left_RT	(Curl_Left#_Rig_Top) V (Curl_Left#_MidH_MidV)
รอยหยักจากด้านซ้ายบริเวณซ้ายบน	C_Down_LT	Curl_Down#_Left_Top
Rule” อ1 ” : (LP_MM + EP_LT + C_Left_RT)/3 Rule” อ2 ” : (LP_MM + EP_LT + C_Down_LT)/3		

ฮ		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_R) V (LP#_Left_MidV#_R) V (EP#_MidH_MidV#_RU) V (EP#_MidH_MidV#_R)
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_L) V (LP#_MidH_Top#_L)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#RU
จุดปลายบริเวณขวากลาง	C_Left_RM	(Curl_Left#_Rig_MidV) V (Curl_Left#_MidH_MidV)
Rule” ฮ ” : (LP_MM + LP_LT + EP_RT + C_Left_RM)/4		

ั		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_R) V (LP#_Left_Top#_L) V (EP#_Left_Top#_LT)
จุดปลายบริเวณขวากลาง	EP_RM	(EP#_Rgt_MidV#_RT) V (EP#_Rgt_Top#RU)

รอยหยักจากด้านบนบริเวณกลาง กลาง	C_Up_MM	Curl_Up#_MidH_MidV
รูปทรง	Squared	(Wide) V (Squared)
Rule“ ั ”: (LP_LT + EP_RM + EP_RT + Squared)/4		

า		
จุดปลายบริเวณซ้ายบน	EP_LT	(EP#_Left_Top#_LD) V (EP#_Left_Top#_L)
จุดปลายบริเวณกลางกลาง	EP_RB	EP#_Rig_Bot#_D
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	SCS#_EP#_R#1
Rule“ า ”: (EP_LT + EP_RB + SCS_LP_R)/3		

ิ		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_R) V (LP#_MidH_MidV#_L)
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	SCS#_LP#_R#1
รูปทรง	Square	(Wide) V (Squared)
Rule“ ิ ”: (LP_MM + SCS_LP_R + Square)/3		

ี		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_L) V (LP#_MidH_MidV#R)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U)
รอยหยักจากด้านบนบริเวณขวาล่าง	C_Up_RB	Curl_Up#_Rig_Bot)
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	SCS#_LP#_R#2
จุดปลายบริเวณขวาบน	Squared	(Wide V (Squared)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Rule“ ๔ ”: (LP_MM + EP_RT + C_Up_RB + SCS_LP_R + Squared)/5

๔		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_L) V (LP#_MidH_MidV#R)
หัวบริเวณขวากลาง	LP_RM	(LP#_Rig_MidV#_R) V (LP#_Rig_Top#_R)
รอยหยักจากด้านบนบริเวณขวากลาง	C_Up_RM	Curl_Up#_Right_MidV
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	SCS#_LP#_R#3
รูปทรง	Squared	(Wide) V (Squared)
Rule“ ๕ ”: (LP_MM + LP_RM + C_Up_RM + SCS_LP_R + Squared)/5		

๕		
หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_L) V (LP#_MidH_MidV#R)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U
จุดปลายบริเวณกลางบน	EP_MT	EP#_MidH_Top#_U
รอยหยักจากด้านบนบริเวณขวากลาง	C_Up_RM	Curl_Up#_Right_MidV
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	SCS#_LP#_R#3
รูปทรง	Squared	(Wide) V (Squared)
Rule“ ๖ ”: (LP_MM + EP_RT + EP_MT + C_Up_RM + SCS_LP_R + Squared)/6		

ๆ		
หัวบริเวณกลางบน	LP_MT	(LP#_MidH_Top#L) V (LP#_Left_Top#_L) V (LP#_Left_MidV#_L) V (LP#_MidH_MidV#_L)
จุดปลายบริเวณขวาล่าง	LP_RB	EP#_Rig_Bot#_D
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	SCS#_LP#_R#1
รูปทรง	Squared	(Wide) V (Squared)
Rule“ ๆ ”: (LP_MT + LP_RB + SCS_LP_R + Squared)/4		

ฆ		
หัวบริเวณกลางบน	LP_MT	(LP#_MidH_Top#L) V (LP#_Left_Top#_L)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#U
รอยหยักจากด้านบนบริเวณขวาล่าง	C_Up_RB	Curl_Up#_Rig_Bot
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	SCS#_LP#_R#2
รูปทรง	Squared	(Wide) V (Squared)
Rule“ ฆ ”: (LP_MT + EP_RT + C_Up_RB + SCS_LP_R + Squared)/5		

เ		
หัวบริเวณขวาล่าง	LP_MM	(LP#_MidH_MidV#_R) V (LP#_Right_Bot#_R)
จุดปลายบริเวณซ้ายบน	LP_LT	EP#_Left_Top#_U
ลำดับสายลากผ่านของหัวไปทางขวา	SCS_LP_R	SCS#_LP#_R#1
รูปทรง	Thin	Thin
Rule“ เ ”: (LP_MM + LP_LT + SCS_LP_R + Thin)/4		

โ		
หัวบริเวณขวาล่าง	LP_RB	LP#_Right_Bot#_R
หัวบริเวณขวาบน	LP_RT	EP#_Rgt_Top#RU
รอยหยักจากด้านขวาบริเวณซ้ายบน	C_Right_LT	(Curl_Rgt#_Lft_Top) V (Curl_Rgt#_MidH_Top)
รูปทรง	Thin	Thin
Rule“ โ ”: (LP_RB + LP_RT + C_Right_LT + Thin)/4		

ใ		
หัวบริเวณขวาล่าง	LP_RB	LP#_Right_Bot#_R)
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_L) V (LP#_Left_Top#_R)
รอยหยักจากด้านล่างบริเวณกลางบน	C_Down_MT	(Curl_Down#_MidH_Top) V (Curl_Down#_Rig_Top)
รูปทรง	Thin	Thin
Rule“ ใ ”: (LP_RB + LP_LT + C_Down_MT + Thin)/4		

ไ		
หัวบริเวณซ้ายบน	LP_RB	LP#_Right_Bot#_R)
จุดปลายบริเวณกลางกลาง	EP_LT	(EP#_Left_Top#_LT) V (EP#_Left_MidV#_LU)
รอยหยักจากด้านล่างบริเวณขวาบน	C_Down_RT	(Curl_Down#_Rig_Top) V (Curl_Down#_Rig_Top)
รอยหยักจากด้านบนบริเวณกลางบน	C_Up_MT	(Curl_Up#_MidH_Top) V (Curl_Up#_MidH_MidV)
รูปทรง	Thin	Thin
Rule“ ไ ”: (LP_RB + EP_LT + C_Down_RT + C_Up_MT + Thin)/5		

๘		
หัวบริเวณขวาล่าง	LP_RB	(LP#_Right_Bot#_R) V (LP#_MidH_MidV#_R)
จุดปลายบริเวณขวาบน	EP_RT	EP#_Rgt_Top#RU
รอยหยักจากด้านขวาบริเวณซ้ายกลาง	C_Right_LM	Curl_Right#_Left_MidV
รอยหยักจากด้านล่างบริเวณซ้ายกลาง	C_Down_LM	Curl_Down#_Lft_MidV
Rule“ ็ ”: (LP_RB + EP_RT + C_Right_LM + C_Down_LM)/4		

'		
จุดปลายบริเวณกลางบน	EP_MT	EP#_MidH_Top#_U
จุดปลายบริเวณกลางกลาง	EP_MB	EP#_MidH_Bot#_D
รูปทรง	Thin	Thin
Rule“ ' ”: (EP_MT + EP_MB + Thin)/3		

ข		
หัวบริเวณซ้ายบน	LP_LT	(LP#_Left_Top#_L) V (LP#_MidH_Top#L)
จุดปลายบริเวณซ้ายล่าง	EP_LB	EP#_Rgt_Top#RU
รอยหยักจากด้านบนบริเวณกลางล่าง	C_Up_MB	Curl_Up#_MidH_Bot
รอยหยักจากด้านบนบริเวณกลางกลาง	C_Left_MM	Curl_Left#_MidH_MidV
Rule“ ้ ”: (LP_LT + EP_LB + C_Up_MB + C_Left_MM)/4		

๗		
หัวบริเวณซ้ายล่าง	LP_LB	LP#_Left_Bot#R

จุดปลายบริเวณขวาบน	EP_RT	(EP#_Rgt_Top#U) V (EP#_MidH_Top#_LT)
รอยหยักจากด้านบนบริเวณกลางกลาง	C_Up_MM	Curl_Up#_MidH_MidV
รอยหยักจากด้านล่างบริเวณกลางกลาง	C_Down_MM	Curl_Down#_MidH_MidV
Rule“ ”: (LP_LB + EP_RT + C_Up_MM + C_Down_MM)/4		

+		
จุดปลายบริเวณกลางบน	EP_MT	EP#_MidH_Top#_U
จุดปลายบริเวณซ้ายกลาง	EP_LM	EP#_Left_MidV#_R
จุดปลายบริเวณขวากลาง	EP_RM	EP#_Rig_MidV#_D
จุดปลายบริเวณกลางล่าง	EP_MB	EP#_MidH_Bot#_L
Rule“ + ”: (EP_MT + EP_LM + EP_RM + EP_MB)/4		

หัวบริเวณซ้ายล่าง	LP_LT	LP#_Left_Bot#_R
จุดปลายบริเวณขวาบน	EP_LB	EP#_Rig_Top V #_RU
Rule“ ”: (LP_LT + EP_LB)/2		

หัวบริเวณกลางกลาง	LP_MM	(LP#_MidH_MidV#_L) V (LP#_MidH_MidV#_R)
Rule“ ”: LP_MM		

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นายอิทธิพันธ์ เมธเศรษฐ เกิดวันที่ 7 มีนาคม พ.ศ. 2519 ที่กรุงเทพมหานคร เข้ารับการศึกษาในหลักสูตรวิศวกรรมศาสตรบัณฑิต คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2537 สำเร็จการศึกษาปริญญาตรีวิศวกรรมศาสตร์บัณฑิต ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2540 และเข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต ที่ห้องปฏิบัติการวิจัยประมวลผลสัญญาณดิจิทัล ภาควิชาวิศวกรรมไฟฟ้า ที่จุฬาลงกรณ์มหาวิทยาลัยในปีการศึกษา 2541

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย