การดึงความเร็วการลากเส้นตัวอักษรลายมือเขียนจากภาพระดับสีเทา
โดยใช้ตัวแบบแหล่งกำเนิด

นาย นวภัทร์ สุจิรชาโต

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า

คณะวิศวกรรมศาตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2544

ISBN 974-17-0010-5

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

STROKE VELOCITY EXTRACTION OF GRAY-SCALE

HANDWRITTEN CHARACTER BY THE USING SOURCE MODEL

Mr. Navapat Sujirachato

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 2001

ISBN 974-17-0010-5

หัวข้อวิทยานิพนธ์ การดึงความเร็วการลากเส้นตัวอักษรลายมือเขียนจากภาพระดับสีเทาโดยใช้ตัวแบบแหล่งกำเนิด

โดย นาย นวภัทร์ สุจิรชาโต

สาขาวิชา วิศวกรรมไฟฟ้า

อาจารย์ที่ปรึกษา อาจารย์ สุวิทย์ นาคพีระยุทธ

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

....คณบดีคณะวิศวกรรมศาสตร์

(ศาสตราจารย์ ดร.สมศักดิ์ ปัญญาแก้ว)

คณะกรรมการสอบวิทยานิพนธ์

....ประธานกรรมการ

(รองศาสตราจารย์ ดร.สมชาย จิตะพันธ์กุล)

....อาจารย์ที่ปรึกษา

(อาจารย์ สุวิทย์ นาคพีระยุทธ)

....กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.เจษฎา ชินรุ่งเรือง)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

นวภัทร์ สุจิรชาโต : การดึงความเร็วการลากเส้นตัวอักษรลายมือเขียนจากภาพระดับสีเทาโดยใช้ตัวแบบแหล่งกำเนิด. (STROKE VELOCITY EXTRACTION OF GRAY-SCALE HANDWRITTEN CHARACTER BY USING THE SOURCE MODEL) อ.ที่ปรึกษา : อาจารย์ สุวิทย์ นาคพีระยุทธ 53 หน้า. ISBN 974-17-0010-5.

วิทยานิพนธ์ฉบับนี้มีจุดมุ่งหมายเพื่อศึกษาวิธีการดึงข้อมูลการลากเส้นจากภาพระดับสีเทา ออกมาเป็นความเร็ว เวลา และพารามิเตอร์อื่นๆ ของการเคลื่อนที่ โดยพิจารณาว่าภาพลายมือเขียน เกิดจากแบบแหล่งกำเนิดที่ประกอบด้วย ตัวแบบการเคลื่อนที่ของมือซึ่งเป็นสมการอนุพันธ์อันดับสอง จะทำหน้าที่ สร้างตัวอักษร ขึ้นจากความเร็วแนวนอน และความเร็วแนวตั้ง ทั้งความเร็วของแนวนอนและความเร็วของแนวตั้งจะสัมพันธ์กันและเป็นคุณลักษณะของตัวอักษรนั้นๆ ร่วมกับตัวแบบการเกิดภาพของปากกา ตัวแบบแหล่งกำเนิดนี้จะกำหนดความเข้มและความหนาของเส้น โดยแปรเปลี่ยนตามความเร็วในการลากเส้นจากตัวแบบการเคลื่อนที่ของมือ

ในการศึกษานี้จะใช้ตัวแบบและชุดพารามิเตอร์ที่กำหนดขึ้น ในการสร้างภาพลายมือเขียน และทดลองดึงข้อมูลการลากเส้นกลับคืนมาจากภาพระดับสีเทาเพื่อเปรียบเทียบกัน ผลการทดสอบพบว่าการดึงข้อมูลความเร็วการลากเส้นในแนวตั้งและแนวนอนได้ดีพอสมควร แต่การดึงพารามิเตอร์ต่างๆ ของการเคลื่อนที่ ได้แก่ ความถี่และความหน่วง ที่เวลาต่างๆ ไม่มีเสถียรภาพพอสำหรับการรู้จำผู้เขียนได้

ภาควิชา.......วิศวกรรมไฟฟ้า.......... ลายมือชื่อนิสิต..
สาขาวิชา......วิศวกรรมไฟฟ้า......... ลายมือชื่ออาจารย์ที่ปรึกษา.....................................
ปีการศึกษา....2544.....................

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

4170364121 : MAJOR ELECTRICAL ENGINEERING

KEY WORD : STROKE / SOURCE MODEL

NAVAPAT SUJIRACHATO : STROKE VELOCITY EXTRACTION OF GRAY-SCALE HANDWRITTEN CHARACTER BY USING THE SOURCE MODEL. THESIS ADVISOR : SUVIT NAKPEERAYUTH 53 pp. ISBN 974-17-0010-5.

The purpose of this thesis is to study the method of extracting the stroke information from gray-scale image. The extracted information are the velocity, timing and other motion parameters. The main idea is to model the handwritten image generation by two source models. The first one is the hand movement model which are second order differential equations. The model gives the horizontal and vertical velocities of the stroke with proper relationship to form a character. The pen's image generating model render the image using the velocity from hand movement model. This model will determine the intensity and width of the stroke from the ink quantity deposited on the paper which is a function of stroke velocity.

In this study, the test gray-scale image will be generated from predefined model and parameters set. Then the extraction of all parameters using the inverse model are preformed on the test image for comparison. The test results show quite acceptable level for both vertical and horizontal velocity retrieved from gray-scale image. But other parameters such as frequency and damping are not stable enough for furher usage.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

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

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

Academic year..2001………………………

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลืออย่างดียิ่งของ อาจารย์สุวิทย์ นาคพีระยุทธ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งท่านได้ให้คำแนะนำ ข้อคิดเห็นต่างๆ และสนับสนุนเครื่องมือในการทำวิจัยมา อีกทั้งยังได้ผลักดันจนงานสำเร็จได้อย่างดี ผู้วิจัยขอกราบขอบพระคุณมา ณ ที่นี้ด้วย

ขอขอบคุณ เพื่อน พี่น้องนิสิตที่อยู่ภายในห้องปฏิบัติการสื่อสาร ที่ได้ช่วยเหลือและเป็นกำลังใจที่ดียิ่งต่อผู้วิจัย ท้ายนี้ ผู้วิจัยใคร่ขอกราบขอบพระคุณ บิดามารดา และครอบครัว ซึ่งสนับสนุนและให้กำลังใจแก่ผู้วิจัยเสมอมาจนสำเร็จการศึกษา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย .... ง

บทคัดย่อภาษาอังกฤษ .... จ

กิตติกรรมประกาศ .... ฉ

สารบัญ .... ช

สารบัญภาพ .... ญ

บทที่

1 บทนำ 1

1.1 แนวเหตุผล 1

1.2 วัตถุประสงค์ 2

1.3 ขั้นตอนการดำเนินการ 3

1.4 ขอบเขตของโครงงาน 3

1.5 ประโยชน์ที่คาดว่าจะได้รับ 3

2 แนวคิดทฤษฎีและผลงานที่ผ่านมา 4

2.1 งานวิจัยที่เกี่ยวข้อง 4

2.2 งานวิจัยที่ใช้ระดับสีเทาในการแก้ปัญหา 4

2.3 งานวิจัยที่เกี่ยวข้องกับตัวแบบแหล่งกำเนิด 6

2.4 การทำโครงกระดูกของภาพขาวดำ 7

2.4.1 อัลกอริทึมการทำโครงกระดูกภาพแบบ Local 8

2.5 การแปลงระยะทาง 10

2.5.1 อัลกอริทึมในการแปลงระยะทาง 10

2.5.2 Local Distance Mask 10

2.5.3 วิธีการหาระยะทางสะสมแบบ sequential 11

2.5.4 ตัวอย่างของการแปลงระยะทาง 12

2.6 การค้นหาโครงกระดูกจากแผนที่ระยะทาง 13

2.6.1 วิธีการเดินขึ้นเขาเพื่อค้นหาโครงกระดูกภาพจากแผนที่ระยะทาง 14

2.6.2 อัลกอริทึมในการเดินหาโครงกระดูกจากแผนที่ระยะทาง 15

2.6.3 ตัวอย่างการค้นหาโครงกระดูก 16

สารบัญ (ต่อ)

บทที่ หน้า

2.7 การทำโครงกระดูกกับภาพระดับสีเทา (Gray-scale Skeletonization) 18
3 ตัวแบบแหล่งกำเนิด 20
3.1 ตัวแบบแหล่งกำเนิด.. 20
3.2 ตัวแบบการเคลื่อนที่ของมือ... 22
3.2.1 วิธีหาพารามิเตอร์โดยใช้ค่าเฉลี่ยความผิดพลาดกำลังสอง............................... 25
3.2.2 การหาพารามิเตอร์โดยใช้การปรับจำนวนเฟรมโดยอัตโนมัติ แบบ Binary Search
.... 25
3.2.3 ตัวอย่างการหาขอบเฟรม 26
3.2.4 การหาพารามิเตอร์โดยใช้การแบ่งเฟรมที่ตำแหน่งความเร็วการลากเส้นแนวตั้งเป็นศูนย์ 26
3.2.5 ข้อดีของการใช้ตัวแบบการเคลื่อนที่ของมือ 27
3.3 ตัวแบบการเกิดภาพของปากกาและตัวแบบย้อนกลับแหล่งกำเนิด.......................... 27
3.3.1 การใช้ตัวแบบการเกิดภาพของปากกาไปสร้างภาพ 30
3.3.2 หน้าตัดระดับความเข้มจากภาพจริงและจากภาพจำลอง 31
3.3.3 วิธีการคำนวณปริมาณหมึกจากภาพระดับสีเทา... 32
3.3.4 ตัวแบบย้อนกลับการเกิดภาพของปากกา... 33
4 การทดลอง ผลการทดลอง และวิเคราะห์ผล 34
4.1 เครื่องมือที่ใช้ในการทดลอง.. 34
4.2 แหล่งที่มาของข้อมูล 34
4.3 การทดสอบตัวแบบการเกิดภาพของปากกา 34
4.3.1 การเลือกค่าระยะทางสำหรับทำ Sampling by Distance 35
4.3.2 ตัวอย่างของภาพลายมือเขียนที่ผ่านแสกน.. 35
4.3.3 ความเร็วในการลากเส้นเขียนตัวอักษร.. 36
4.3.4 การสร้างตัวแบบการเกิดภาพของปากกา 36
4.4 การดึงข้อมูลการลากเส้นโดยตัวแบบย้อนกลับการเกิดภาพของปากกา...................... 39
4.4.1 การหาโครงกระดูกด้วยตัวแบบย้อนกลับการเกิดภาพของปากกา..................... 39

สารบัญ (ต่อ)

บทที่ หน้า

4.4.2 การใช้ตัวแบบย้อนกลับเพื่อดึงความเร็วจากปริมาณหมึก 40

4.5 การหาพารามิเตอร์ของตัวแบบการเคลื่อนที่ของมือ 41

4.5.1 การหาพารามิเตอร์โดยใช้การปรับจำนวนเฟรมโดยอัตโนมัติแบบ Binary Search 41

4.5.2 การหาพารามิเตอร์โดยใช้การแบ่งเฟรมที่ตำแหน่งความเร็วการลากเส้นแนวตั้งเป็นศูนย์ 42

4.6 การทดลองดึงข้อมูลการลากเส้นจากภาพระดับสีเทาที่สังเคราะห์ขึ้น 42

5 บทสรุปและข้อเสนอแนะ 48

5.1 สรุปผลงานวิจัย 48

5.2 การประยุกต์ใช้ 50

5.3 ข้อเสนอแนะ 50

รายการอ้างอิง 51

ประวัติผู้เขียนวิทยานิพนธ์ 53

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

หน้า

รูปที่ 1.1 ระบบ OCR โดยทั่วไป 1
รูปที่ 1.2 ภาพลายมือเขียนที่เกิดจากตัวแบบแหล่งกำเนิด 2
รูปที่ 2.1 กระบวนการทำโครงกระดูกภาพ (ก) ภาพต้นแบบ (ข) ภาพที่ผ่านการทำโครงกระดูกภาพ 8
รูปที่ 2.2 แบบรูป (Templates) การทำโครงกระดูกภาพ 9
รูปที่ 2.3 Local Distance Mask 5x5 และ Local Distance Mask 3x3 11
รูปที่ 2.4 Forward Mask และ Backward Mask และทิศทางการสแกนภาพ 12
รูปที่ 2.5 ภาพขาวดำ และแผนที่ระยะทาง 12
รูปที่ 2.6 แผนที่ระยะทางสะสมที่ได้จาก Forward Mask และ Backward Mask 13
รูปที่ 2.7 ขั้นตอนการแปลงระยะทาง 13
รูปที่ 2.8 ตัวอย่างของ Base Point ภายในหน้าต่าง 3x3 14
รูปที่ 2.9 ตัวอย่างของ Directional Neighborhood of P 15
รูปที่ 2.10 แผนที่ระยะทาง 16
รูปที่ 2.11 Apex Point บนแผนที่ระยะทาง 16
รูปที่ 2.12 การเดินขึ้นเขาบนแผนที่ระยะทาง 16
รูปที่ 2.13 การเดินลงเขาบนแผนที่ระยะทาง 17
รูปที่ 2.14 ตัวอย่างของการหาโครงกระดูก 17
รูปที่ 3.1 ตัวอย่างของแบบแหล่งกำเนิดเอกสารจากเครื่องคอมพิวเตอร์ 20
รูปที่ 3.2 แบบแหล่งกำเนิดภาพลายมือเขียน 21
รูปที่ 3.3 การใช้แบบแหล่งกำเนิดย้อนกลับของอุปกรณ์การเขียน 21
รูปที่ 3.4 การจับปากกาเขียน และตัวแบบการเคลื่อนที่ของมือ 22
รูปที่ 3.5 ลักษณะความเป็นเชิงเส้นและความต่อเนื่องของความถี่และค่าความหน่วง 24
รูปที่ 3.6 ตัวอย่างการหาขอบเฟรม 26
รูปที่ 3.7 ภาพประกอบแนวการลากเส้นจาก x1 ไป x2 27
รูปที่ 3.8 หน้าตัดของระดับความเข้มเป็นรูป Raised Cosine 28
รูปที่ 3.9 การสร้างภาพจำลองโดยการซ้อนทับด้วยวงกลม 30
รูปที่ 3.10 หน้าตัดระดับความเข้มของปากกาเมจิค 31
รูปที่ 3.11 หน้าตัดระดับความเข้มที่ได้จาก Raised Cosine 31

รูปที่ 3.12 Forward Mask และ Backward Mask และทิศทางการสแกนภาพ 32
รูปที่ 4.1 Sampling By Distance ที่ความยาวแตกต่างกัน 35
รูปที่ 4.2 ตัวอย่างของความกว้างและระดับความเข้มของแนวการลากเส้นที่แตกต่างกัน...... 36
รูปที่ 4.3 แนวโน้มการเปลี่ยนแปลงระดับความเข้มตามความเร็วการลากเส้น 37
รูปที่ 4.4 แนวโน้มการเปลี่ยนแปลงระดับความกว้างตามความเร็วการลากเส้น 37
รูปที่ 4.5 ปริมาณหมึกที่สัมพันธ์กับความเร็วของการลากเส้น 38
รูปที่ 4.6 ตัวอย่างภาพขนาดจริงที่สร้างจากตัวแบบการเกิดภาพของปากกา 38
รูปที่ 4.7 ตัวอย่างภาพ ก) ตำแหน่งต่างๆที่ถูกบันทึกความเร็ว ข) ภาพโครงกระดูกและตำแหน่งที่ถูกนำประมาณหาความเร็วการลากเส้น.. 39
รูปที่ 4.8 ความเร็วแนวนอนและแนวตั้งการลากเส้นที่เกิดขึ้นจริง.. 40
รูปที่ 4.9 ความเร็วแนวนอนและแนวตั้งที่เกิดจากการประมาณด้วยตัวแบบย้อนกลับการเกิดภาพของปากกา.. 40
รูปที่ 4.10 ขั้นตอนการสังเคราะห์ภาพลายมือเขียน.. 42
รูปที่ 4.11 ตัวอย่างของการหาพารามิเตอร์ของตัวแบบการเคลื่อนที่ของมือ.......................... 42
รูปที่ 4.12 ความเร็วการลากเส้นในแนวนอนและแนวตั้งที่ได้จากการเก็บข้อมูลทาง Tablet.... 43
รูปที่ 4.13 ความเร็วการลากเส้นในแนวนอนและแนวตั้งที่ได้จากการสังเคราะห์โดยพารามิเตอร์ 43
รูปที่ 4.14 ลายมือเขียนที่เกิดขึ้นจริงและตัวอักษรที่ได้จากการสังเคราะห์ซึ่งมีความใกล้เคียงกัน 43
รูปที่ 4.15 ภาพลายมือเขียนระดับสีเทาที่สังเคราะห์ขึ้นจากตัวแบบการเกิดภาพของปากกา .. 44
รูปที่ 4.16 ภาพโครงกระดูกลายมือเขียนที่หามาได้จากภาพข้างต้น.................................... 44
รูปที่ 4.17 ความเร็วการลากเส้นที่ได้จากตัวแบบย้อนกลับการเกิดภาพของปากกา 45
รูปที่ 4.18 ความเร็วการลากเส้นที่ได้จากสังเคราะห์ด้วยพารามิเตอร์.................................. 45
รูปที่ 4.19 ภาพลายมือเขียนที่เกิดจากพารามิเตอร์และภาพลายมือเขียนที่ได้จากการดึงข้อมูลย้อนกลับจากภาพระดับสีเทา 45
รูปที่ 4.20 แนวโน้มของความถี่ของความเร็วการลากเส้นแนวนอนระหว่างพารามิเตอร์ต้นแบบและพารามิเตอร์ที่ได้จากการหาย้อนกลับ.. 46
รูปที่ 4.21 แนวโน้มของค่าความหน่วงของความเร็วการลากเส้นแนวนอนระหว่างพารามิเตอร์ต้นแบบและพารามิเตอร์ที่ได้จากการหาย้อนกลับ 46

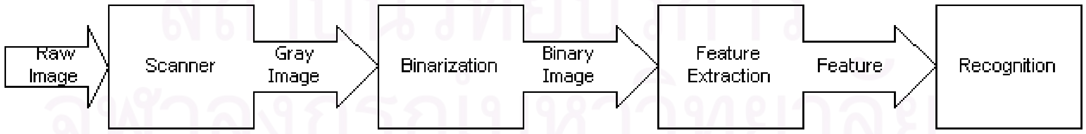
รูปที่ 4.22 แนวโน้มของความถี่ของความเร็วการลากเส้นแนวตั้งระหว่างพารามิเตอร์ต้นแบบและพารามิเตอร์ที่ได้จากการหาย้อนกลับ 46
รูปที่ 4.23 แนวโน้มของค่าความหน่วงของความเร็วการลากเส้นแนวตั้งระหว่างพารามิเตอร์ต้นแบบและพารามิเตอร์ที่ได้จากการหาย้อนกลับ 46

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1
บทนำ

1.1 แนวเหตุผล

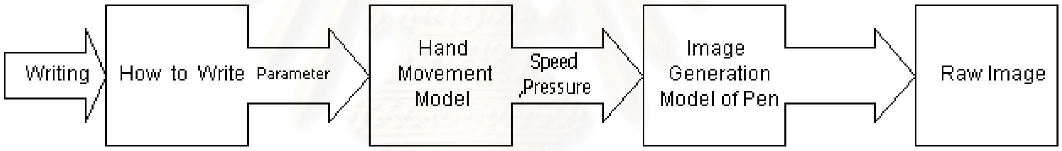
ในปัจจุบันงานวิจัยในด้านการรู้จำ (pattern recognition) ได้เข้ามามีบทบาทสำคัญอย่างมาก เนื่องจากการพัฒนาในเทคโนโลยีด้านคอมพิวเตอร์และอิเลคทรอนิกส์ ทำให้งานด้านการรู้จำมีการนำมาใช้ในชีวิตประจำวันมากขึ้นเพราะความสะดวกสบายและรวดเร็วนั่นเอง โดยเฉพาะงานOCR ซึ่งมีการนำมาใช้งานอย่างกว้างขวาง ดังตัวอย่างเช่น การรู้จำรหัสไปรษณีย์และการรู้จำตัวเลขจำนวนเงินบนเช็คธนาคาร เป็นต้น ระบบ OCR ที่ใช้วิธีดูลักษณะการลากเส้นเขียนตัวอักษรนั้นจะต้องหาแนวการลากเส้น(stroke)ของตัวอักษรให้ถูกต้องเพื่อการรู้จำที่ถูกต้อง โดยปกติการเขียนตัวอักษร จะประกอบไปด้วยสิ่งแวดล้อมหนึ่ง ได้แก่ อุปกรณ์การเขียน ความสัมพันธ์ระหว่างการเปลี่ยนทิศทางและความเร็วของการลากเส้นและน้ำหนักของการกดของปากกากับความเข้มของภาพตัวอักษร เป็นต้น ในที่นี้จะเรียกรูปแบบที่ก่อเกิดเป็นตัวอักษรว่า ตัวแบบแหล่งกำเนิด (source model) ซึ่งถ้าเราสามารถจำลองตัวแบบแหล่งกำเนิด ได้ใกล้เคียงจะทำให้เราทราบตัวอักษรถูกเขียนขึ้นมาอย่างไร โดยเฉพาะการรู้จำแบบ offline ตัวแบบแหล่งกำเนิดจึงเปรียบเสมือนเป็น a priori knowledge ของระบบเพื่อนำมาประกอบการพิจารณาเพื่อหาลักษณะสำคัญของภาพตัวอักษร ในที่นี้จะหมายถึง stroke (การลากเส้น) ทำให้การหา stroke ถูกต้องมากขึ้น และจะช่วยให้การรู้จำทำได้ดีขึ้นด้วย ซึ่งจะเปรียบเสมือนการที่ตำรวจไปตรวจที่เกิดเหตุ แล้วสามารถสรุปเรื่องราวให้เข้ากับสมมติฐานได้จากการพิจารณาพยานและหลักฐานในที่เกิดเหตุ แม้ว่าจะไม่ได้จะอยู่เหตุการณ์จริงก็ตาม ซึ่งบางครั้งก็อาจจะผิดบ้าง และยิ่งมีองค์ประกอบหรือหลักฐานมาพิจารณามากขึ้นเท่าไรก็ยิ่งถูกมากขึ้น



รูปที่ 1.1 ระบบ OCR โดยทั่วไป

อย่างไรก็ตามในระบบOCRโดยทั่วไป(รูปที่1.1) จะมีขั้นตอนโดยพื้นฐาน เริ่มต้นจากการสแกนภาพตัวอักษร แล้วนำไป binarization ซึ่งจะได้ภาพที่มีระดับความเข้ม 2 ระดับ(ขาวดำ) แล้วจึงหาลักษณะสำคัญของภาพตัวอักษร ก่อนเข้าสู่กระบวนการรู้จำในลำดับถัดมา ซึ่งขั้นตอน binarization สามารถเปรียบเทียบกับการทำลายหลักฐานในที่เกิดเหตุ ส่งผลให้การพิจารณาคดีผิดเพี้ยนดังนั้นระบบ OCR ยังมีข้อจำกัดและความผิดพลาดที่เกิดขึ้นอยู่เสมอ จึงได้มีความพยายามจะแก้ไข ตัวอย่างเช่น การนำพจนานุกรมมาตรวจสอบ หรือ การจำกัดลักษณะการเขียนและการลากเส้น เป็นต้น แต่นั่นก็เป็นเพียงการแก้ปัญหาที่ปลายเหตุ สาเหตุในลำดับแรกๆที่มักถูกละเลย คือ การ binarization ภาพระดับสีเทา เพราะการ binarization จะทำลายข้อมูลดิบซึ่งประกอบด้วยคุณลักษณะและความผิดเพี้ยน ได้แก่

- ผิดเพี้ยนอันเนื่องจาก point spread function ของสแกนเนอร์
- ผิดเพี้ยนอันเนื่องจาก noise ในขณะสแกน
- คุณลักษณะจากความแตกต่างพื้นหลังและพื้นหน้า พื้นหลังที่เป็นลวดลาย
- คุณลักษณะของ source model ที่ให้ภาพระดับสีเทาที่ไม่สม่ำเสมอตลอดแนวการลากเส้น



รูปที่ 1.2 ภาพลายมือเขียนที่เกิดจากตัวแบบแหล่งกำเนิด

1.2 วัตถุประสงค์

1. เพื่อศึกษากรรมวิธีในการใช้ source model ที่เหมาะสมเพื่อสร้างภาพลายมือเขียนจากแนวการลากเส้น

2. เพื่อศึกษากรรมวิธีการดึงแนวการลากเส้น ของภาพอักษรลายมือเขียนระดับสีเทาโดยใช้ source model

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

1.3 ขั้นตอนการดำเนินการ

1. ศึกษาปัญหาและแนวทางในการแก้ปัญหา โดยเน้นปัญหาที่เกิดขึ้นกับภาพลายมือเขียน
2. พัฒนากรรมวิธีเพื่อแก้ปัญหาที่เกิดขึ้น
3. สร้างภาพจากแบบจำลอง source model
4. สร้างโปรแกรมเพื่อทดสอบกรรมวิธี
5. เก็บตัวอย่างข้อมูลภาพตัวอักษร
6. วิเคราะห์ผลที่ได้จากการทดสอบและหาวิธีแก้ไขข้อผิดพลาด
7. สรุปผลและเขียนวิทยานิพนธ์

1.4 ขอบเขตของโครงงาน

1. หา source model ที่เหมาะสมกับการเขียนตัวอักษรเพื่อนำไปสู่การสร้าง stroke ที่คล้ายจริง
2. พัฒนาโปรแกรมเพื่อหาข้อมูลการลากเส้นของภาพลายมือเขียนที่สร้างขึ้นจาก source model
3. ทดสอบและสรุปเมื่อนำใช้งานกับภาพลายมือที่สร้างขึ้นมา

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถรู้ถึงข้อมูลการลากเส้นที่เป็นไปตามธรรมชาติของการเขียนตัวอักษรได้ใกล้เคียง
2. ใช้ในการรู้จำลายมือเขียนที่ทำได้ยาก เช่น ลายมือที่หวัด ตัวอักษรที่ติดกัน
3. เป็นแนวทางในการพัฒนาการรู้จำลายเซ็น เมื่อสามารถปรับปรุง source model ให้เหมาะสำหรับแต่ละคน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2
แนวคิดทฤษฎีและผลงานที่ผ่านมา

2.1 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องการรู้จำด้านลายมือเขียน ได้ถูกพัฒนาอย่างต่อเนื่อง ด้วยกรรมวิธีที่หลากหลาย และมีข้อดีข้อเสียที่แตกต่างตามแต่ละกรรมวิธี อย่างไรก็ตามการแก้ปัญหาการรู้จำทั้งหมด โดยอาศัยความสามารถของกรรมวิธีรู้จำเพียงอย่างเดียวคงเป็นไปไม่ได้ ระบบการรู้จำในปัจจุบันนิยมปรับภาพระดับสีเทา ให้กลายเป็นภาพขาวดำแล้วทำการดึงคุณลักษณะจากภาพดังกล่าว ซึ่งการปรับระดับดังกล่าว ยังไม่ถูกต้องมากนักในขณะนี้ การดึงข้อมูลคุณลักษณะ(Feature) จากภาพระดับสีเทาจึงถูกพัฒนาขึ้นมา ข้อมูลคุณลักษณะที่เราต้องการดึงในที่นี้ก็คือ ข้อมูลการลากเส้น เพราะยังคงเค้าโครงเดิมของภาพต้นฉบับได้ และนิยมใช้ในระบบรู้จำ จากความสัมพันธ์ระหว่างระดับความเข้มสีเทากับพฤติกรรมของการเขียน ณ ตอนนั้น ทำให้เราพยายามดึงข้อมูลที่เกิดขึ้นในระหว่างการเขียนจากภาพระดับสีเทา อันได้แก่ ความเร็ว ทิศทางการลากเส้น แรงกด ซึ่งถือเป็นคุณลักษณะที่สำคัญอีกอย่างหนึ่ง ซึ่งข้อมูลนี้มีส่วนช่วยในการวิเคราะห์แนวการลากเส้นเช่นกัน งานวิจัยในด้านนี้มีมากมาย สามารถสรุปได้ดังนี้

2.2 งานวิจัยที่ใช้ระดับสีเทาในการแก้ปัญหา

งานวิจัยด้านลายมือเขียน และด้านอื่นๆที่เกี่ยวข้อง นิยมพิจารณาภาพลายมือเขียนให้เป็นการลากเส้น ดังนั้นภาพลายมือเขียนดังกล่าวจะต้องถูกทำให้บางเป็นโครงกระดูกของตัวอักษร อย่างไรก็ตาม ในระบบรู้จำส่วนมากจะทำการปรับระดับภาพระดับสีเทาให้เป็นภาพขาวดำ แล้วจึงทำให้ภาพของลายมือเขียน เป็นโครงกระดูกของตัวอักษร และทำการหาคุณลักษณะ แล้วเข้าสู่การรู้จำ งานวิจัยในกลุ่มนี้จึงได้เสนอว่า การปรับระดับความเข้มของภาพให้เหลือ 2 ระดับ ทำให้เกิดปัญหาบางอย่างได้ เช่น การมีตัวอักษรติดกัน ตัวอักษรที่ขาดกัน ดังนั้นจึงได้เสนอการทำโครงกระดูกกับภาพระดับสีเทาโดยตรง ซึ่งได้แก่งานวิจัยต่อไปนี้

Li Wang,Theo Pavlidis (1993) ได้เสนองานวิจัยที่ริเริ่มการทำหาโครงกระดูกกับภาพระดับสีเทา โดยการมองภาพระดับสีเทาให้เสมือนแผนที่ทางภูมิศาสตร์ โดยใช้ระดับสีเทาจากภาพที่แต่ละตำแหน่งของภาพ นำมาคำนวณหา eigen-vector และ eigen-value ของ Hessian Matrix ที่แต่ละจุด ค่าที่คำนวณได้จะนำมาเปรียบเทียบและกำหนดคุณลักษณะทาง topographic อันได้แก่ สันเขา หลุม ยอดเขา หุบเหว ร่องเขา คุณลักษณะดังกล่าวของทุกตำแหน่งภาพจะ

ประกอบเป็นแผนที่ทางภูมิศาสตร์ แผนที่ทางภูมิศาสตร์จะถูกประมวลผลโดยวิธีการทางกราฟ เพื่อให้ได้โครงกระดูกของลายมือเขียน คุณลักษณะทาง topographic ถูกนำมาพิจารณาร่วมในการหาโครงกระดูกได้เป็นอย่างดี

Seong-Whan Lee ,Young Joon Kim(1995) ได้เสนองานวิจัยที่พัฒนาจาก Li Wang, Theo Pavlidis โดยบอกว่าการหา eigen-vector และ eigen-value เป็นการคำนวณที่ยุ่งยากและเสียเวลา ได้เสนอวิธีการหา principal curvature direction และ principal orthogonal element แล้วนำ orthogonal element ร่วมกับการตรวจสอบทิศทางของ zero crossing ซึ่งสามารถใช้ระบุคุณลักษณะทาง topographic ได้เช่นเดียวกัน

Shy-Shyan,Chen Frank Y Shih(1996) ได้เสนองานวิจัยที่พัฒนาจาก Li Wang, Theo Pavlidis โดยยังคงหา eigen-vector และ eigen-value แล้วนำค่าดังกล่าวไปใช้ใน membership grade of ridge point เพื่อให้คะแนน ridge ที่มีความสำคัญ ridge ที่ได้คะแนนมากจะเป็นโครงกระดูกของภาพ ridge ที่ได้คะแนนน้อย อาจเป็นส่วนเชื่อมต่อระหว่างลายมือเขียน

Hirobumi Nishida (1998) ได้เสนองานวิจัยที่ใช้ระดับสีเทามาช่วยในการดึงคุณลักษณะของตัวอักษรโดยใช้ Triangle Approximation ในการหา close-boundaryของตัวอักษร ได้ทดลองกับระบบรู้จำรหัสไปรษณีย์ โดยในขั้นแรกระบบรู้จำจะทำการรู้จำกับภาพที่ผ่านการ binarization ถ้าพบว่าภาพดังกล่าวถูกคัดออกมา จะทำการดึงคุณลักษณะจากภาพระดับสีเทา ซึ่งช่วยเพิ่มความถูกต้อง และ วิธีดึงคุณลักษณะจากภาพระดับสีเทาสามารถใช้กับภาพที่มีความละเอียดต่ำได้ดี

G. Boccignoise ,A Chianese,L.P Cordella(1993) ได้เสนองานวิจัยที่เกี่ยวข้องกับการดึงแนวการลากเส้นที่เกิดขึ้นตามการลากจริงๆ โดยทำการหา MAT (Medial Axis Transform) แล้วพยายามแทน MAT ด้วย polygonal segment เพราะว่า MAT นั้นค่อนข้างอ่อนไหวต่อการรบกวนของขอบภาพ และข้อดีของ polygonal segment คือ เราทราบความยาวและมุมของ segment และ เป็นการลดปริมาณข้อมูลลง segment สามารถรวมกันเป็น segment เมื่อมีการคำนวณความเหมือนกับอีก segment วิธีดังกล่าวแสดงแนวการลากเส้นในตัวอักษร A X Y ได้อย่างถูกต้องเหมือนจริง

David S. Doermann, Azriel Rosenfield (1993) ได้เสนองานวิจัยที่พยายามดึงข้อมูล online จากข้อมูล offline โดยเสนอว่า ข้อมูล online ดังกล่าวแฝงในรูปของระดับความเข้ม และความกว้างของแนวการลากเส้น โดยข้อมูลที่ต้องการดึงกลับคืน ได้แก่ ความเร็วสัมพัทธ์ แนวการลากเส้น และแรงกด ขณะเดียวกันเราควรจะทราบข้อมูลเบื้องต้นเหล่านี้ก่อนที่จะดึงข้อมูลดังกล่าว ได้แก่ การเปลี่ยนของระดับความเข้ม และความกว้าง จุดเริ่มต้นของการลากเส้น วิธีการสังเคราะห์ลายมือเขียน อุปกรณ์การเขียน งานวิจัยนี้เสนอว่าผลการเปลี่ยนแปลงของระดับความเข้มและ

ความกว้างของแนวการลากเส้นเกิดจากปัจจัยทางแรงกด มากกว่าปัจจัยด้านความเร็วและความเร่ง การดึงข้อมูลด้านความเร็วและความเร่งนั้นยากมาก โดยแรงกดที่ต่างกันเกิดจากพฤติกรรมการเขียนที่ต่างกัน การวิเคราะห์จะต้องทำในบริเวณนั้นๆเพื่อหาพฤติกรรมนั้นๆ อย่างไรก็ตามได้พบปัญหาหลัก คือ การหาแนวการลากเส้นที่ซ้อนทับ ตัดกัน หรือ มีการลากซ้ำ งานวิจัยนี้ได้วิเคราะห์พฤติกรรมของปากการ่วมกับการแปรเปลี่ยนของระดับความเข้มด้วย

นอกจากนี้ ในงานวิจัยนี้ มีสมมติฐานเกี่ยวกับปากกา คือ ปากกาจะต้องมีการเคลื่อนที่ก่อนที่หมึกจะเริ่มไหล ตัวอย่างเช่น ปากกาลูกลื่น ดังนั้นตอนหยุดปากกาจากการเขียนตำแหน่งนั้นจะมีระดับความเข้มมากกว่า ในขณะที่ตำแหน่งเริ่มต้นจรดปากกา อาจจะไม่มีหมึกอยู่เลย

K. K.Lau .Pong C Yuen,Y. Y. Tang (2000) ได้เสนองานวิจัยที่พยายามหาโครงกระดูกจากภาพระดับสีเทา มีหลักการคือ อุปกรณ์การเขียนที่แตกต่างกัน จะมีผลทำให้ รูปร่างของหน้าตัดของระดับความเข้มต่างกัน ปากกาบางชนิดมีรูปร่างหน้าตัดของระดับความเข้มสูงสุด มีช่วงกว้าง ซึ่งจะไม่มี local maximum เกิดขึ้น จึงต้องทำการปรับรูปร่างหน้าตัด ให้มี local maximum เกิดขึ้น เพื่อใช้ในการหาโครงกระดูกต่อไป

2.3 งานวิจัยที่เกี่ยวข้องกับตัวแบบแหล่งกำเนิด

งานวิจัยในกลุ่มนี้กล่าวถึงการหาตัวแบบแหล่งกำเนิดโดยนัย งานวิจัยกลุ่มนี้มาจากระบบรู้จำแบบ online มีพื้นฐานความคิดที่ว่า การลากเส้นใดๆเกิดจากความเร็วการลากเส้นในแนวดิ่งและแนวนอน โดยสำหรับแต่ละตัวอักษร ความเร็วทั้งสองแนวจะสัมพันธ์กัน ความเร็วในการลากเส้นทั้งสองแนวจะประกอบขึ้นเป็นคุณลักษณะ ของแต่ละตัวอักษร สามารถนำไปใช้ในการรู้จำได้

Hollerbach (1961) ได้เสนอ Oscillator model เป็นงานวิจัยริเริ่มในการมองว่าตัวอักษร หรือ การลากเส้นเกิดจากความเร็วในการลากเส้นแนวดิ่งและแนวนอน ความเร็วทั้งสองอยู่ในลักษณะของ sinusoidal

$$\dot{y}(t) = A_y(t)\sin(\omega_y(t)t + \phi_y) \qquad (2\text{-}1)$$

$$\dot{x}(t) = A_x(t)\sin(\omega_x(t)t + \phi_x) + c \qquad (2\text{-}2)$$

A_x และ A_y คือ แอมพลิจูด ω_x และ ω_y คือ ความถี่ ϕ_x และ ϕ_y คือ เฟส เป็นพารามิเตอร์ที่กำหนดรูปร่างของแนวการลากเส้นที่เกิดขึ้น โดยจากสมการข้างต้นความเร็วในการลากเส้นแนวนอนจะมีค่าคงที่ c และc>0 ซึ่งหมายถึง การเขียนจะเริ่มจากซ้ายไปขวาเสมอ อย่างไรก็ตาม oscillator model ยังมีปัญหาจากความเร็วการลากเส้นในบางครั้งเพิ่มขึ้นอย่างรวดเร็ว

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Y.Singer และ Tishby(1994) ได้เสนอการหาพารามิเตอร์ดังกล่าว โดยการแบ่งแนวการลากเส้นของลายมือเขียนเป็นเฟรม ขอบของเฟรมหาจากจุด zero-crossing ของความเร็วในการเขียนแนวตั้ง

Orly Stettiner , Dan Chazan (1994) ได้เสนอ Damping Sinusoidal เป็นงานวิจัยที่พัฒนาจาก Oscillator model เนื่องจากในตัวแบบแหล่งกำเนิดข้างต้นไม่เหมาะสมกับความเร็วของการลากเส้นที่บางครั้งมีการเพิ่มขึ้น ตัวแบบแหล่งกำเนิดนี้มีความใกล้เคียงมากกว่าแบบเดิม การหาพารามิเตอร์ของตัวแบบแหล่งกำเนิดนี้ ทำโดยการ optimization แบบnonlinear ในตัวแบบแหล่งกำเนิดนี้จะมองความเร็วของการลากเส้นเป็นเฟรม พารามิเตอร์จะเปลี่ยนอย่างเป็นเชิงเส้นอย่างช้าๆ ภายใน frame และ จะต้องต่อเนื่องระหว่างเฟรมด้วย

$$V_x(t) = \frac{A_x}{\sqrt{1-\xi(t)^2}} \sin(\omega_x(t)\sqrt{1-\xi(t)^2}\,t + \phi_x) \cdot \exp(-\xi(t)\,\omega_x(t).t) + c \qquad (2\text{-}3)$$

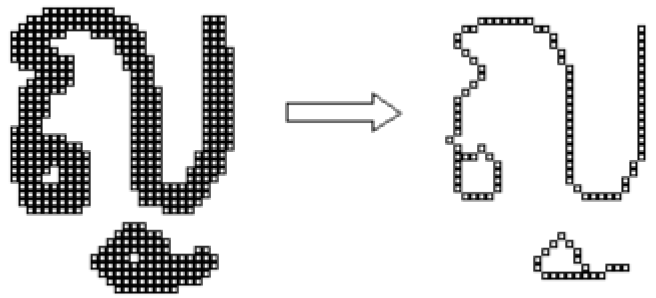
A_x คือ แอมพลิจูด ω_x คือ ความถี่ ϕ คือ เฟส และ ξ คือ ค่าความหน่วง เป็นพารามิเตอร์ที่กำหนดรูปร่างของแนวการลากเส้น สิ่งที่แตกต่าง คือ ξ สามารถมีค่าเป็นลบได้ และ $|\xi| < 1$ ค่านี้ใช้บังคับให้ความเร็วเพิ่มขึ้นอย่างรวดเร็วได้

2.4 การทำโครงกระดูกของภาพขาวดำ

การทำโครงกระดูกของภาพตัวอักษร เป็นการทำให้ความหนาของภาพต้นแบบเปลี่ยนแปลงไป โดยความหนาจะเหลือเพียงหนึ่งจุดภาพ และยังคงเค้าโครงเดิมอยู่ วิธีในการทำโครงกระดูกมีหลายวิธี โดยสามารถแบ่งกรรมวิธีในการทำโครงกระดูก ตามวิธีประมวลผลได้ 2 วิธี

- LOCAL เป็นวิธีที่นิยมใช้อย่างแพร่หลายมาก โดยมีหลักการคือ ใช้หน้าต่างสี่เหลี่ยมขนาดต่างๆกันทาบไปตามภาพ แล้วทำการลบจุดออกไปเรื่อยๆ เสมือนการปอกเปลือกผลไม้ นิยมใช้เนื่องจากความง่ายและรวดเร็วในประมวลผล แต่มีข้อเสียคือ ไม่ได้พิจารณาข้อมูลทั้งหมดในภาพ
- NON-LOCAL เป็นวิธีที่ซับซ้อนโดยใช้หลักการของ MAT(Medial Axis Transform) ประกอบด้วย การแปลงระยะทาง และการค้นหาโครงกระดูกจากแผนที่ระยะทาง ซึ่งจะกล่าวอย่างละเอียด ในลำดับถัดไป

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



(ก) (ข)

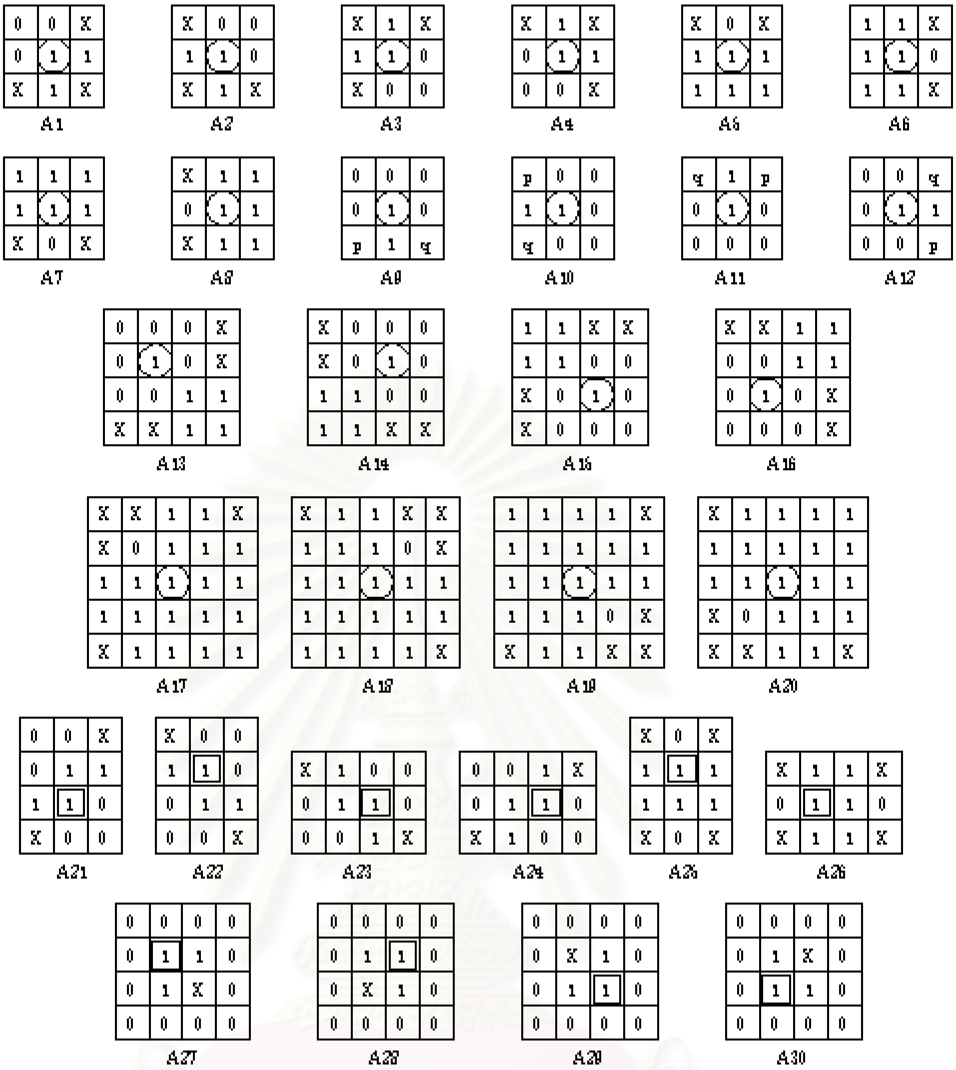
รูปที่ 2.1 กระบวนการทำโครงกระดูกภาพ (ก) ภาพต้นแบบ (ข) ภาพที่ผ่านการทำโครงกระดูก

2.4.1 อัลกอริทึมการทำโครงกระดูกภาพแบบ Local

เนื่องจากการทำโครงกระดูกภาพวิธีนี้มีอยู่ด้วยกันหลายวิธี ในที่นี้เลือกใช้การทำให้บางโดยกรรมวิธี One-Pass Parallel Thinning ของ Ben K. Jang และ Roland T. Chin (1992) เพราะให้รายละเอียดคุณลักษณะของจุดต่อภาพได้ดีเพียงพอ โดยมีแบบรูป (Template) ที่ใช้ในการทำโครงกระดูกภาพทั้งหมด 30 รูปแบบ ซึ่งแบบรูป (Templates) A1 ถึง A20 เป็นแบบรูปการทำโครงกระดูกภาพ

ในขณะที่ A21 ถึง A30 เป็นแบบรูปการเรียกกลับคืน (Restoring) p และ q เป็นตัวดำเนินการทางตรรกศาสตร์ : p or q = 1 และ X เป็นค่าที่ไม่สนใจ ดังแสดงในรูปที่ 2.2

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.2 แบบรูป (Templates) การทำโครงกระดูกภาพ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

- **ขั้นตอนการหาโครงกระดูกภาพ**

i. จุดภาพจะถูกทำการตรวจสอบตามแบบรูปการทำโครงกระดูกภาพ (Template) ทั้งหมด

ii. หากจุดภาพที่ถูกตรวจอยู่ในแบบรูป A1 ถึง A20 จะถูกลบจุดภาพ ในขณะที่จุดภาพอยู่ในอยู่แบบ A21 ถึง A30 จะถูกกู้กลับคืนมา

iii. เมื่อภาพผ่านกระบวนการทำโครงกระดูกบางจะเหลือความกว้างของภาพเพียง 1 จุดภาพ

2.5 การแปลงระยะทาง (Distance Transform)

การแปลงระยะทาง คือ การหาระยะทางที่สั้นที่สุดจากตำแหน่งหนึ่งบนภาพ(พื้นหน้า) ไปยังพื้นหลัง แล้วนำระยะทางที่คำนวณได้ดังกล่าวของทุกตำแหน่งบนภาพ มาประกอบแผนที่ระยะทาง (Distance Map) ดังนั้นวิธีนี้จะต้องทำการปรับระดับภาพให้เหลือเพียงขาวดำ คือ พื้นหน้าและพื้นหลัง แผนที่ระยะทางที่ได้จะมีลักษณะดังนี้ มีตำแหน่งภาพที่ใกล้กับพื้นหลัง จะมีระยะทางไปสู่พื้นหลังที่น้อยที่สุด ตามความหมายของการแปลงระยะทาง และ ตำแหน่งภาพที่อยู่ถัดมาชั้นหนึ่งจะมีระยะทางไปสู่พื้นหลังมากขึ้นอีกเรื่อยๆ จนกระทั่งค่าระยะทางดังกล่าวมีค่ามากถึงค่าหนึ่งเท่านั้น ซึ่งเป็นแนวสันที่อยู่ห่างจากพื้นหลังสองด้านเท่ากัน

2.5.1 อัลกอริทึมในการแปลงระยะทาง

การหาระยะทางที่น้อยที่สุด ที่ตำแหน่งภาพเดียว จะต้องคำนวณระยะทางจากตำแหน่งนั้นไปยังพื้นหลังกับทุกจุด แล้วเปรียบเทียบระยะทางเหล่านั้น ซึ่งจะใช้การคำนวณหลายรอบ การแปลงระยะทางของทั้งภาพจะใช้การคำนวณมาก และ การแปลงระยะทาง โดยใช้ Euclidean Distance จะให้ความถูกต้องมากที่สุด แต่จะใช้การคำนวณมากเช่นกัน อัลกอริทึมที่ถูกพัฒนาขึ้นเพื่อลดการคำนวณ และ จำนวนครั้งของการคำนวณลง โดยมีหลักการดังนี้

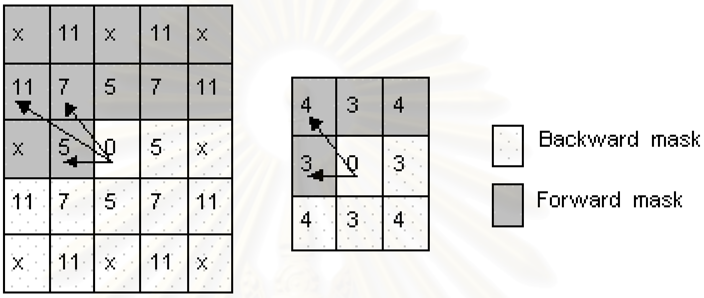
ระยะทางที่ใช้อยู่ในรูป pseudo-Euclidean ซึ่งเป็นตัวเลขจำนวนเต็ม เพราะเป็นการประมาณที่ใกล้เคียง การวัดระยะทางเกิดจากการสะสมระยะทางจากตำแหน่งภาพใกล้เคียงทำให้มีแต่การบวกเท่านั้น และ ลดจำนวนครั้งของการคำนวณระยะทาง การสะสมระยะทางมีทั้งแบบ sequential และ แบบ parallel ทำโดยใช้ maskที่มี local distance ใช้ในการสะสมระยะทาง ในที่นี้จะเสนอแบบ sequential เท่านั้น

2.5.2 Local Distance Mask

mask ใช้ในการคำนวณระยะทางสะสม ภายใน mask จะมีค่า local distance ซึ่งเป็นค่าระยะทางที่ใกล้เคียงที่สุดจากศูนย์กลางของ mask ไปยังตำแหน่งต่างๆบน mask local distance ถูกประมาณขึ้นมาเป็นเลขจำนวนเต็ม Local Distance mask มีหลายขนาด ได้แก่ 3x3 5x5 7x7 ขนาดของ mask ที่มีขนาดใหญ่ จะให้ค่าระยะทางที่ถูกต้องมากกว่าขนาดเล็ก เพราะมีการประมาณค่า local distance ในตำแหน่งต่างๆมากกว่า การเลือกใช้เลขจำนวนเต็มแทน local distance ทำให้มีวิธีวัดระยะทางหลายวิธี ได้แก่

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

- CityBlock Distance $D = |X - Xo| + |Y - Yo|$ เป็นการประมาณที่ให้ความถูกต้องน้อยที่สุด โดยให้ระยะทางระหว่างตำแหน่งภาพในแนวระนาบและแนวตั้งเป็น 1 หน่วยเท่านั้น ดังนั้น ระยะทางในแนวทะแยงมุม 45 องศาได้จากระยะทางแนวระนาบบวกกับแนวตั้ง ซึ่งเท่ากับ 2 หน่วย
- Chessboard Distance $D = \max(|X - Xo|, |Y - Yo|)$
- Chamfer 5-7-11 , Chamfer 3-4 ใช้ใน mask 5x5 , 3x3 ตามลำดับ (Gunilla Borgefos (1986))



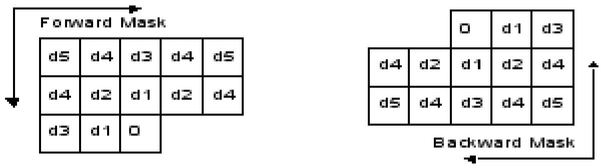
รูปที่ 2.3 Local Distance Mask 5x5 และ Local Distance Mask 3x3

จากตัวอย่างข้างต้นแสดง local distance mask ที่นิยมใช้กัน mask 5x5 จะมี local distance จากจุดศูนย์กลางที่ประมาณขึ้นมา 3 ค่า อันได้แก่ แนวระนาบหรือแนวตั้งมีค่า local distance เท่ากันคือ 5 หน่วย แนว 45 องศามีค่า local distance เท่ากับ 7 หน่วย แนว 26.5 องศา มีค่า local distance เท่ากับ 11 หน่วย และ mask 3x3 จะมี local distance ที่ประมาณขึ้นมา 3 ค่า อันได้แก่ แนวระนาบหรือแนวตั้งมีค่า local distance เท่ากันคือ 3 หน่วย แนว 45 องศามีค่า local distance เท่ากับ 4 หน่วย

2.5.3 **วิธีการหาระยะทางสะสมแบบ** sequential

เริ่มต้นจากสร้างแผนที่ระยะทางสะสมที่มีระยะทางสะสมเป็นศูนย์ทุกตำแหน่ง ทำการสแกนภาพในทิศทาง forward จากซ้ายไปขวา และ บนลงล่าง ให้ค้นหาตำแหน่งของภาพที่เป็นพื้นหน้าให้นำ mask ไปวางทาบบนแผนที่ระยะทางสะสม โดยให้ศูนย์กลางของ mask อยู่ที่เดียวกับตำแหน่งภาพที่เป็นพื้นหน้าดังกล่าว เลือกค่าที่น้อยที่สุดจากการนำ local distance ไปบวกกับค่าระยะทางสะสมบนแผนที่ระยะทางสะสม บันทึกค่านั้นลงแผนที่ระยะทางสะสม สามารถเขียนในรูปสมการดังนี้

สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.4 Forward Mask และ Backward Mask และทิศทางการสแกนภาพ

$$v_{i,j} = \min_{(k,l)\in mask} (v_{i+k,j+l} + d(k,l)) \qquad (2\text{-}4)$$

$V_{i,j}$ คือ ระยะทางสะสมในแผนที่ระยะทางสะสมที่ตำแหน่ง (i,j)

d(k,l) คือ ค่า local distance ที่ตำแหน่ง (k,l) ภายใน mask ที่ใช้ในการสแกน

ทำการสะสมในอีกทิศทางกลับกันด้วย (Backward scan) แล้วเก็บระยะทางสะสมในอีกแผนที่หนึ่ง เมื่อทำการสแกนภาพครบทั้ง 2 ทิศทาง จะได้แผนที่ระยะทางสะสมในทิศทางที่ต่างกัน 2 ชุด แล้วจึงนำแผนที่ระยะทางสะสมทั้งสองมาเปรียบเทียบ ที่ตำแหน่งเดียวกันเลือกค่าระยะทางสะสมที่มีค่าน้อยที่สุด แล้วบันทึกค่าระยะทางดังกล่าวในแผนที่ระยะทางสุดท้ายให้ทำจนกระทั่งครบทุกตำแหน่งภาพ ด้วยวิธีการเปรียบเทียบนี้ ทำให้ลดจำนวนครั้งของคำนวณระยะทาง

2.5.4 **ตัวอย่างของการแปลงระยะทาง**

0	0	0	0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	0
0	1	1	1	0	0	0	1	1	1	0
0	1	1	1	0	0	0	1	1	1	0
0	1	1	1	0	0	0	1	1	1	0
0	0	0	0	0	0	0	0	0	0	0

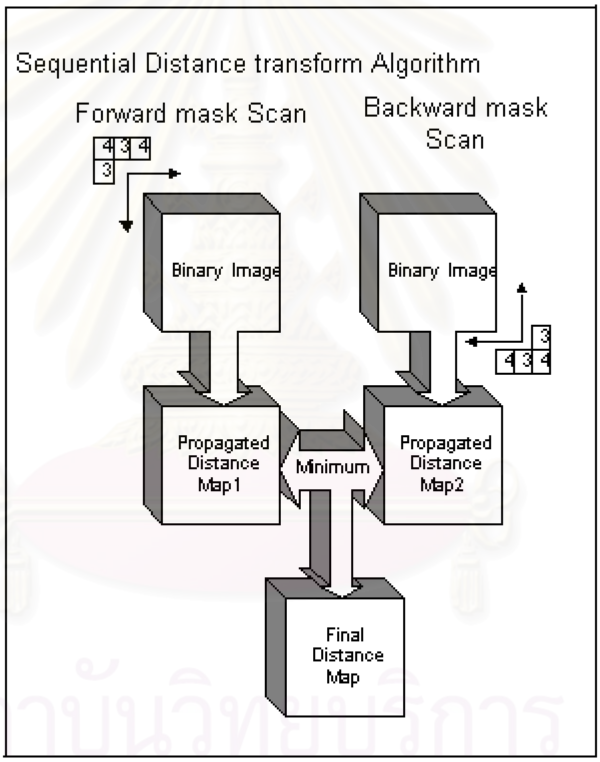
0	0	0	0	0	0	0	0	0	0	0
0	3	3	3	3	3	3	3	3	3	0
0	3	6	6	6	6	6	6	6	3	0
0	3	6	4	3	3	3	4	6	3	0
0	3	6	3	0	0	0	3	6	3	0
0	3	6	3	0	0	0	3	6	3	0
0	3	3	3	0	0	0	3	3	3	0
0	0	0	0	0	0	0	0	0	0	0

รูปที่ 2.5 ภาพขาวดำ และแผนที่ระยะทาง

0	0	0	0	0	0	0	0	0	0	0
0	3	3	3	3	3	3	3	3	3	0
0	3	6	6	6	6	6	6	6	4	0
0	3	6	9	9	9	9	9	9	4	0
0	3	6	9	0	0	0	3	6	4	0
0	3	6	4	0	0	0	3	6	4	0
0	3	6	4	0	0	0	3	6	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	4	11	10	9	9	9	9	6	3	0
0	3	8	7	6	6	6	7	6	3	0
0	3	7	4	3	3	3	4	6	3	0
0	3	7	3	0	0	0	4	6	4	0
0	3	6	3	0	0	0	4	6	3	0
0	3	3	3	0	0	0	3	3	3	0
0	0	0	0	0	0	0	0	0	0	0

รูปที่ 2.6 แผนที่ระยะทางสะสมที่ได้จาก Forward Mask และ Backward Mask



รูปที่ 2.7 ขั้นตอนการแปลงระยะทาง

2.6 การค้นหาโครงกระดูกจากแผนที่ระยะทาง

แผนที่ระยะทางจะมีลักษณะเหมือนแผนที่ภูมิศาสตร์ ได้แก่ สันเขา หลุม ยอดเขา หุบเหว ร่องเขา สังเกตว่าสำหรับแผนที่ระยะทางใดๆมีแนวสันเขาเกิดขึ้นอยู่บริเวณตรงกลาง และมีเค้าโครงเหมือนภาพต้นฉบับก่อนผ่านการแปลงระยะทาง ที่ตำแหน่งดังกล่าวถ้าสร้างวงกลมที่รัศมี

เท่ากับระยะทางที่ระบุในแผนที่ระยะทาง พบว่าวงกลมจะสัมผัสพื้นหลังอย่างน้อยสองจุดซึ่งหมายความว่า เมื่อวัดระยะทางจากตำแหน่งดังกล่าวไปยังพื้นหลัง 2 ตำแหน่งนั้น จะได้ระยะทางเท่ากัน ดังนั้นสันเขาดังกล่าวมีความสมมาตรระหว่างขอบของภาพ จึงถูกนำมาใช้เป็นโครงกระดูกของภาพ โครงกระดูกที่ได้จะต้องมีความต่อเนื่อง จึงต้องรวมตำแหน่งที่อื่นๆที่ไม่ใช่สันเขาแต่มีความสมมาตรเข้าไปด้วย วิธีการรวบรวมโครงกระดูกจากแผนที่ระยะทางในที่นี้จะอ้างอิงจาก Maxima-Tracking Method for Skeletonization from Euclidean Distance Transform ของ Frank Y Shih(1991)

กำหนด คุณลักษณะทาง topographic เหล่านี้ลงในแผนที่

- Base Point ตามแผนที่ทางภูมิศาสตร์ จะเป็นตำแหน่งภาพอยู่ติดกับพื้นหลัง และอยู่ที่มุมของภาพ พิจารณาภายในหน้าต่าง 3x3 จุดนี้จะมีลักษณะดังรูป ซึ่งแต่ละรูปยังหมุนไป 45 องศา ได้อีก 8 แบบ
-

1	0	0
1	1	0
0	0	0

1	1	0
1	1	0
0	0	0

1	1	1
1	1	0
0	0	0

รูปที่ 2.8 ตัวอย่างของ Base Point ภายในหน้าต่าง 3x3

- Apex Point ตามแผนที่ทางภูมิศาสตร์ เมื่อพิจารณาภายในหน้าต่าง 3x3 จะเป็นตำแหน่งภาพที่มีค่าระยะมากที่สุดเปรียบเทียบกับของตำแหน่งอื่นๆภายในหน้าต่าง 3x3 ซึ่งเป็นจุด local maximum

2.6.1 วิธีการเดินขึ้นเขาเพื่อค้นหาโครงกระดูกภาพจากแผนที่ระยะทาง

การเดินเพื่อค้นหาโครงกระดูกจำเป็นจะตรวจสอบคุณสมบัติของตำแหน่งถัดไปในการเดิน และตำแหน่งนั้นควรจะอยู่ในทิศทางที่ใกล้เคียงกับโครงกระดูกเดิมที่มีอยู่แล้ว

- Directional Neighborhood of P (Dp) เป็นการกำหนดทิศทางของการค้นหาโครงกระดูก P เป็นโครงกระดูกที่กำลังพิจารณาอยู่ จากหน้าต่างแรก P5 เป็นโครงกระดูก ดังนั้นในการค้นหาจุดต่อไปในการเดินควรจะต้องอยู่ในทิศทางที่น่าเป็นไปได้ ซึ่งได้แก่ P2 P1 P8

0	0	P2
P5	P	P1
0	0	P8

P4	P3	P2
0	P	0
0	P7	0

P4	P3	0
P5	P	0
0	0	P8

0	P3	P2
0	P	P1
P6	0	0

รูปที่ 2.9 ตัวอย่างของ Directional Neighborhood of P

- วิธีการเดินขึ้นเขา (Uphill Tracing)

  จากตำแหน่งที่เริ่มต้น ให้ค้นหาภายในหน้าต่าง 3x3 ร่วมกับ Dp ตำแหน่งอื่นๆที่มีระยะทางมากกว่า หรือ เท่ากับ ระยะทางที่จุดเริ่มต้น ใช้ตำแหน่งนั้นในการตามหาโครงกระดูก ในการเดินขึ้นเขา ตลอดทางที่เดิน จะมีค่าระยะทางที่เพิ่มขึ้นตลอด

$$P_{next}^{uphill} = \{P_i \subset D_p \cup \{P\} \mid Max(Dist(P_i))\} \qquad (2\text{-}5)$$

- วิธีการเดินลงเขา (Downhill Tracing)

  จากตำแหน่งที่เริ่มต้น ให้ค้นหาภายในหน้าต่าง 3x3 ร่วมกับ Dp โดยยกเว้นตำแหน่งเริ่มต้น ตำแหน่งอื่นๆที่มีระยะทางมากกว่าหรือเท่ากับระยะทางที่จุดเริ่มต้น ใช้ตำแหน่งนั้นในการตามหาโครงกระดูกในการเดินลงเขาจะพบว่า บางครั้งตลอดทางที่เดินจะมีค่าระยะทางที่เพิ่ม หรือ ลดลงได้

$$P_{next}^{downhill} = \{P_i \subset D_p \mid Max(Dist(P_i))\} \qquad (2\text{-}6)$$

2.6.2 อัลกอริทึมในการเดินหาโครงกระดูกจากแผนที่ระยะทาง

i. ทำการหา Apex point แล้วบันทึกไว้

ii. เริ่มต้นที่ Apex point ทำการเดินขึ้นเขาจนกระทั่งพบ Apex point จุดอื่นๆ ให้บันทึกจุดสิ้นสุดการเดินขึ้นเขาไว้

iii. จากจุด Apex ที่มีการบันทึกในขั้นตอนที่ 2 ให้เริ่มต้นที่จุดนี้ในการเดินลงเขา สุดท้ายจะได้โครงกระดูกตามต้องการ

iv. สำหรับจุด Base เอาไว้ตรวจสอบส่วนเกิน โครงกระดูกที่สิ้นสุดที่จุด Base ถือว่าเป็นส่วนเกิน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.6.3 การค้นหาโครงกระดูก

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	3	3	3	3	3	0	0	0	0	0	0	0
0	3	3	3	3	4	6	6	6	3	0	0	0	0	0	0	0
0	3	6	6	6	7	8	9	6	3	0	3	3	3	3	3	0
0	3	3	3	3	4	6	6	7	4	3	4	6	6	6	3	0
0	0	0	0	0	3	3	3	4	6	6	7	8	9	6	3	0
0	0	0	0	0	0	0	0	3	3	3	4	6	6	6	3	0
0	0	0	0	0	0	0	0	0	0	0	3	3	3	3	3	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

รูปที่ 2.10 แผนที่ระยะทาง

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	3	3	3	3	3	0	0	0	0	0	0	0
0	3	3	3	3	4	6	6	6	3	0	0	0	0	0	0	0
0	3	6	6	6	7	8	9	6	3	0	3	3	3	3	3	0
0	3	3	3	3	4	6	6	7	4	3	4	6	6	6	3	0
0	0	0	0	0	3	3	3	4	6	6	7	8	9	6	3	0
0	0	0	0	0	0	0	0	3	3	3	4	6	6	6	3	0
0	0	0	0	0	0	0	0	0	0	0	3	3	3	3	3	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

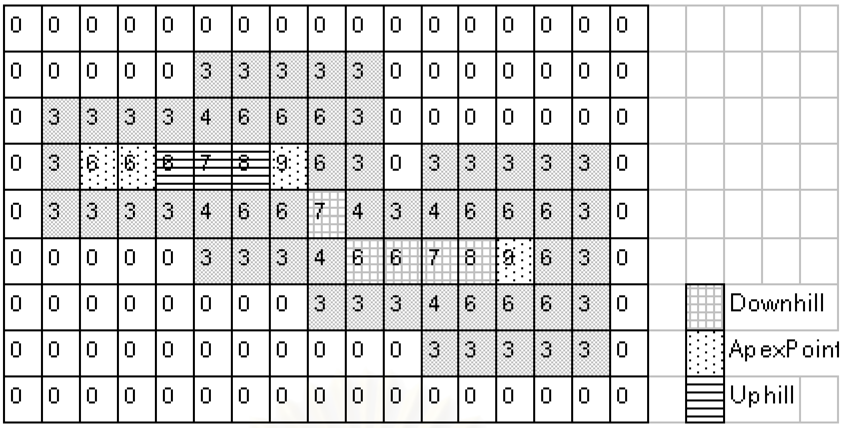
ApexPoint

รูปที่ 2.11 Apex Point บนแผนที่ระยะทาง

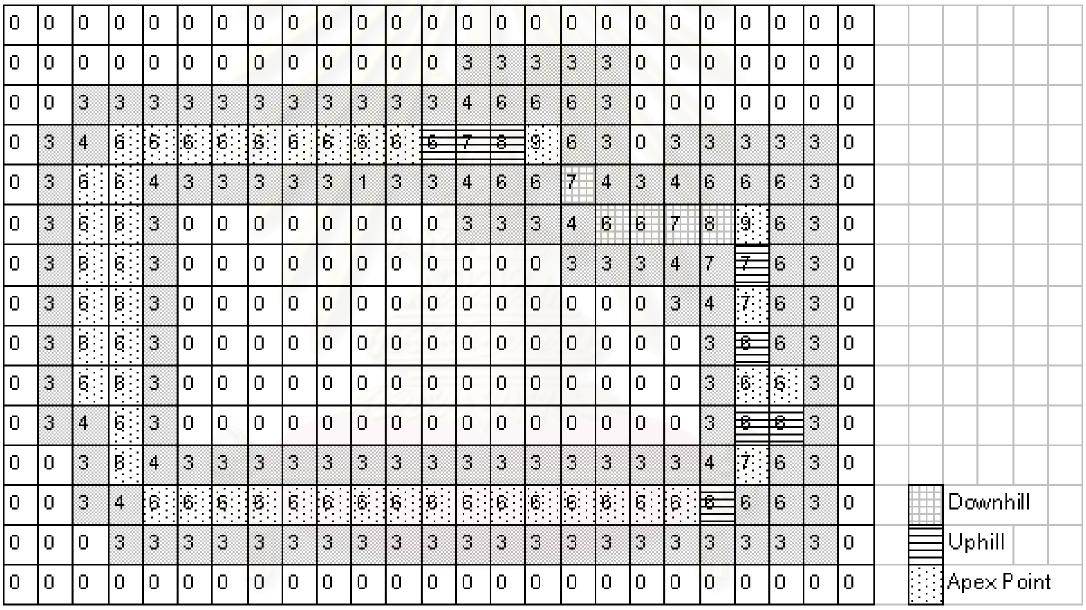
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	3	3	3	3	3	0	0	0	0	0	0	0
0	3	3	3	3	4	6	6	6	3	0	0	0	0	0	0	0
0	3	6	6	6	7	8	9	6	3	0	3	3	3	3	3	0
0	3	3	3	3	4	6	6	7	4	3	4	6	6	6	3	0
0	0	0	0	0	3	3	3	4	6	6	7	8	9	6	3	0
0	0	0	0	0	0	0	0	3	3	3	4	6	6	6	3	0
0	0	0	0	0	0	0	0	0	0	0	3	3	3	3	3	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Uphill
ApexPoint

รูปที่ 2.12 การเดินขึ้นเขาบนแผนที่ระยะทาง



รูปที่ 2.13 การเดินลงเขาบนแผนที่ระยะทาง



รูปที่ 2.14 ตัวอย่างของการหาโครงกระดูก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จากตัวอย่างพบว่ามีโครงกระดูกบางส่วนที่มีความหนาเป็น2 จุดภาพ เนื่องจากในแผนที่ระยะทางมีค่าระยะทางเท่ากันทำให้ไม่สามารถตัดสินใจได้ชัดเจนจึงตัดสินใจ เลือกทั้งสองตำแหน่ง

2.7 **การทำโครงกระดูกกับภาพระดับสีเทา** (Gray-scale Skeletonization)

การหาโครงกระดูกจากภาพระดับสีเทานั้นจะอาศัยหลักการใกล้เคียงกับ MAT ด้วยเหตุผลที่ว่าภาพระดับสีเทาจะมีพื้นผิว 3 มิติที่มีคุณลักษณะทาง topographic ที่ใกล้เคียงกับภาพขาวดำที่ได้ผ่านการแปลงระยะทาง เราจะทำการกำหนดคุณลักษณะทาง topographic มีหลายวิธี ในที่นี้ใช้วิธีของ Li Wang, Theo Pavlidis (1993) โดยคำนวณจาก eigen vector และ eigen vector ของ Hessian Matrix ในที่นี้จะมองภาพระดับสีเทาเป็นพื้นผิว 3 มิติ g(x,y) ให้ ∇g เป็นเวคเตอร์ของความชันสูงสุด (gradient vector) $\|\nabla g\|$ เป็นขนาดของความชัน H เป็น Hessian matrix

$$H = \begin{bmatrix} \frac{\partial^2 g}{\partial x^2} & \frac{\partial^2 g}{\partial x \partial y} \\ \frac{\partial^2 g}{\partial x \partial y} & \frac{\partial^2 g}{\partial y^2} \end{bmatrix} \qquad (2\text{-}7)$$

และ ω_1 และ ω_2 เป็น eigenvector หนึ่งหน่วย λ_1 และ λ_2 เป็น eigenvalue
คุณลักษณะทางtopographicสามารถคำนวณได้จาก eigenvector และ eigenvalue ดังนี้

- Peak (จุดยอด) คือ จุดที่มี local maximum ในทุกทิศทาง หรือ เป็นไปตามเงื่อนไข

  $\|\nabla g\| = 0, \lambda_1 < 0, \lambda_2 < 0$ (2-8)

- Pit (หลุม)คือ จุดที่มี local minimum (ตรงข้ามกับ peak) ในทุกทิศทางหรือเป็นไปตามเงื่อนไข

  $\|\nabla g\| = 0, \lambda_1 > 0, \lambda_2 > 0$ (2-9)

- Ridge (สันเขา) คือ จุดที่มี local maximum ในทิศทางหนึ่ง หรือเป็นไปตามเงื่อนไข

  $\|\nabla g\| =, \lambda_1 < 0, \lambda_2 = 0$ (2-10)

  $\|\nabla g\| = 0, \lambda_1 = 0, \lambda_2 < 0$ (2-11)

  $\|\nabla g\| \neq 0, \lambda_2 < 0, \nabla g \cdot \omega_2 = 0$ (2-12)

  $\|\nabla g\| \neq 0, \lambda_1 < 0, \nabla g \cdot \omega_1 = 0$ (2-13)

- Ravine (คูคลอง) คือ จุดที่มี local minimum ในทิศทางหนึ่ง หรือเป็นไปตามเงื่อนไข

  $\|\nabla g\| = 0, \lambda_1 > 0, \lambda_2 = 0$ (2-14)

  $\|\nabla g\| = 0, \lambda_1 = 0, \lambda_2 > 0$ (2-15)

  $\|\nabla g\| \neq 0, \lambda_1 > 0, \nabla g \cdot \omega_1 = 0$ (2-16)

  $\|\nabla g\| \neq 0, \lambda_2 > 0, \nabla g \cdot \omega_2 = 0$ (2-17)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

- Saddle (อานม้า) คือ จุดที่มี local maximum ในทิศทางหนึ่งและ local minimum ในทิศทางที่ตั้งฉากกับทิศทางแรก

  $\|\nabla g\| = 0, \lambda_1 \cdot \lambda_2 < 0$

- flat (ที่ราบ) คือ จุดที่บนพื้นราบ เป็นไปตามเงื่อนไข

  $\|\nabla g\| = 0, \lambda_1 = 0, \lambda_2 = 0$

- hillside (เชิงเขา) คือ จุดที่นอกเหนือจากเงื่อนไขข้างต้น

จากการกำหนดคุณลักษณะทาง topographic ให้กับภาพระดับสีเทาแล้วจึงนำคุณลักษณะดังกล่าว เรียบเรียงเป็นโครงกระดูก ซึ่งโดยปกติจะหาจากจุดที่เป็น Ridge ,Pit ,Peak นอกจากนี้ยังนำคุณลักษณะดังกล่าวแก้ปัญหาเส้นที่ติดกันได้อีกด้วย

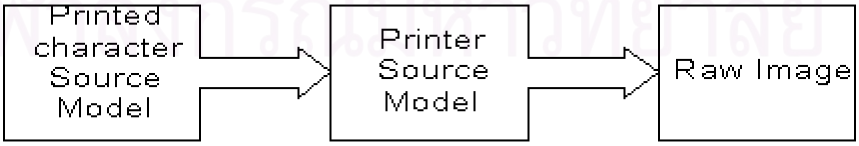
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3
ตัวแบบแหล่งกำเนิด

เนื้อหาในบทนี้จะกล่าวถึง ตัวแบบการเคลื่อนที่ของมือ และตัวแบบการเกิดภาพของปากกา การประยุกต์ตัวแบบย้อนกลับแหล่งกำเนิดเพื่อดึงแนวการลากเส้นและการดึงความเร็วการลากเส้น

3.1 ตัวแบบแหล่งกำเนิด

ในระบบรู้จำเกี่ยวกับลายมือเขียนแบ่งตามการรับข้อมูลและประมวลผลได้เป็น 2 ระบบ คือ online และ offline การเขียนใดๆจะเกิดจากเคลื่อนไหวของมือ ปากกา ข้อศอก และ อื่นๆ ดังนั้นระบบรู้จำ online ซึ่งรับรู้การเคลื่อนไหวดังกล่าวผ่านทาง coordinate ของ Tablet หรือ วิธีรับข้อมูลแบบอื่นๆ ข้อมูลดังกล่าวจะอยู่ในฟังชันของเวลา การรู้จำประมวลผลกับสัญญาณ 1 มิติ ระบบรู้จำแบบ online สามารถให้ผลการทดลองที่ถูกต้องมากกว่า ซึ่งหมายความว่า การรู้ข้อมูลที่เกี่ยวข้องกับ การเคลื่อนไหวอาจทำให้ระบบรู้จำแบบ offline มีความถูกต้องมากขึ้นได้ อย่างไรก็ตามข้อมูลที่เกิดขึ้นในระหว่างการเขียน เช่น ความเร็ว ความเร่ง หรือ แรงกด ยังคงแฝงตัวอยู่บ้างในภาพระดับสีเทา อันได้แก่ ความเข้ม หรือ ความกว้างของแนวการลากเส้น ซึ่งก็มีงานวิจัยที่เกี่ยวข้องกับวิธีการดึงข้อมูลเหล่านี้ดังที่ได้กล่าวในบทที่ 2 แล้ว วิธีการดึงข้อมูลดังกล่าวที่นิยมใช้มากที่สุด คือ การอ้างถึงเงื่อนไขของการเขียน ซึ่งเป็น a priori knowledge ในขณะนั้น เพื่อใช้ในการพิจารณาร่วมกับการประมวลผลกับภาพลายมือเขียนและใช้ประโยชน์ในการรู้จำ ดังนั้นถ้าเงื่อนไขการเขียนที่เรานำมาใช้อ้างมีความถูกต้อง การใช้ข่าวสารดังกล่าวจะส่งผลประโยชน์ให้การประมวลผลใดๆของเราดีขึ้นได้ มีการสร้างสมมติฐานเกี่ยวกับเงื่อนไขการเขียนแบบต่างๆ ตัวอย่างเช่น ความกว้างของปากกา เป็นต้น งานวิจัยนี้ขอเรียกสมมติฐานต่างๆว่า ตัวแบบแหล่งกำเนิด (Source Model) ตัวอย่าง ตัวแบบแหล่งกำเนิดเอกสารจากเครื่องพิมพ์

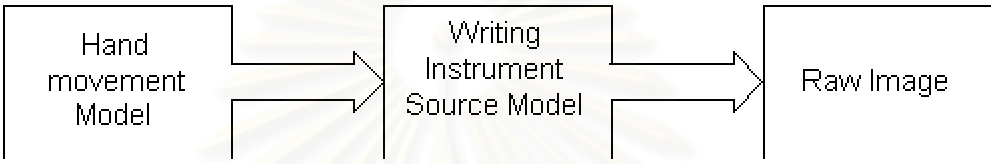


รูปที่ 3.1 ตัวแบบแหล่งกำเนิดเอกสารจากเครื่องคอมพิวเตอร์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

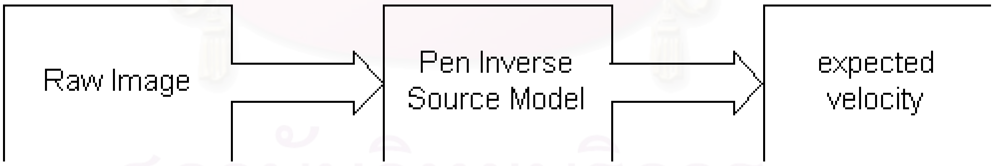
จากตัวอย่างข้างต้น ภาพตัวอักษรพิมพ์ที่เกิดขึ้นจากเครื่องคอมพิวเตอร์อาจจะมีหลายแบบ แต่ถ้าเราทราบว่าภาพดังกล่าวเกิดจากตัวอักษรแบบ Angsana UPC จะทำการรู้จำ หรือการหาโครงกระดูกจะทำได้ง่าย ทำนองเดียวกันภาพลายมือเขียนก็มีตัวแบบแหล่งกำเนิดเช่นเดียวกัน ส่วนเครื่องพิมพ์ก็สามารถใช้ประโยชน์จากตัวแบบได้ เช่น ใช้เครื่องพิมพ์ Dot Matrix 9 เข็มที่มีเข็มหนึ่งเสีย

งานวิจัยนี้มีแนวความคิดว่า ภาพลายมือเกิดจากตัวแบบแหล่งกำเนิดที่ประกอบด้วยตัวแบบการเคลื่อนที่ของมือ และ ตัวแบบการเกิดภาพของปากกา



รูปที่ 3.2 ตัวแบบแหล่งกำเนิดภาพลายมือเขียน

ตัวแบบการเคลื่อนที่ของมือ ทำการสร้างแนวการลากเส้น Trajectories ของตัวอักษร หรือ การลากเส้นใดๆ แนวการลากเส้นดังกล่าวซึ่งจะอยู่ในรูปของความเร็วถูกส่งไปให้ ตัวแบบการเกิดภาพของปากกา แปลงเป็นรูปภาพข้อมูลดิบ



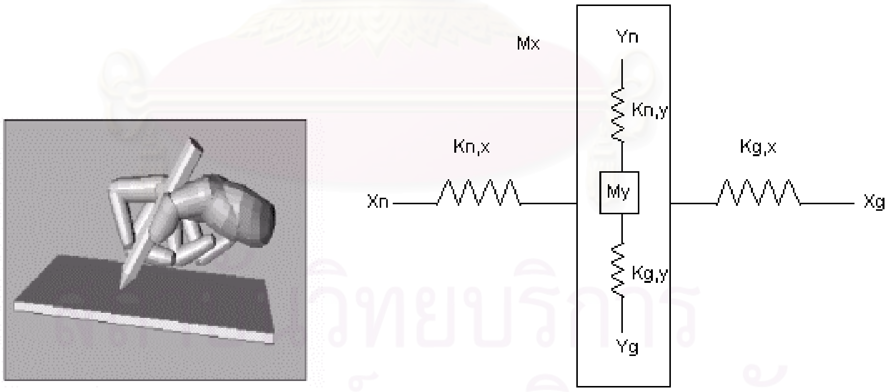
รูปที่ 3.3 การใช้แบบแหล่งกำเนิดย้อนกลับของอุปกรณ์การเขียน

และงานวิจัยนี้ยังนำเสนอการดึงแนวการลากเส้นด้วยตัวแบบแหล่งกำเนิดย้อนกลับของอุปกรณ์การเขียน แนวการลากเส้นอาจจะอยู่ได้หลายรูปแบบ ในที่นี้จะได้ความเร็วการลากเส้น

3.2 ตัวแบบการเคลื่อนที่ของมือ

ตัวแบบการเคลื่อนที่ของมือนี้จะพิจารณาลายมือเขียนในรูปของความเร็วในการลากเส้น ซึ่งอ้างอิงจากงานวิจัยของ Bruno Simard , Birenda Prasada และ R. Mahesh K Sinha(1993) การเขียนตัวอักษรตัวหนึ่ง ก็จะมีการลากเส้นด้วยความเร็วแบบหนึ่งๆ ซึ่งมีที่มาจากการจำลองการเคลื่อนไหวของมือขณะทำการเขียน โดยปกติขณะที่เราเขียนจะใช้ข้อมือในการเขียนแนวนอน และใช้นิ้วชี้และนิ้วโป้งในการเขียนแนวตั้ง มีสมมติฐานว่า การเขียนในแนวนอน หรือแนวตั้ง จะมีกล้ามเนื้อเป็นคู่ที่ทำงานแบบปฏิสัมพันธ์กันอยู่ โดยที่กล้ามเนื้อที่ทำหน้าที่บังคับการเคลื่อนในแนวนอน และ แนวตั้งจะทำงานอย่างอิสระจากกัน

Hollerbach(1961)ได้สร้างสมมติฐานของการเขียนจากตัวแบบจำลองกล้ามเนื้อดังนี้ พิจารณาว่า กล้ามเนื้อทำงานเสมือนสปริงซึ่งสามารถยืดหดได้ เมื่อกล้ามเนื้อหนึ่งมีการยืดตัว กล้ามเนื้อที่คู่กันจะหดตัว และ เป็นเช่นนั้นในทำนองกลับกัน คู่กล้ามเนื้อที่ทำหน้าที่บังคับการเขียนในแนวตั้งจะทำให้เกิดการเคลื่อนตัวของมวลของปากกาและมวลของนิ้ว M_y โดยที่ค่าการยืดหยุ่นของกล้ามเนื้อคู่ดังกล่าว มีค่า $K_{n,x}$ และ $K_{g,x}$ และจุดตรึงของกล้ามเนื้อที่ Y_n และ Y_g ในขณะที่คู่กล้ามเนื้อที่ทำหน้าที่บังคับการเขียนในแนวนอนจะทำให้เกิดการเคลื่อนตัวของมวลของมือและมวลของปากกา M_x โดยที่ค่าการยืดหยุ่นของกล้ามเนื้อคู่ดังกล่าว มีค่า $K_{n,y}$ และ$K_{g,y}$ และจุดตรึงของกล้ามเนื้ออยู่ที่ X_n และ X_g



รูปที่ 3.4 การจับปากกาเขียน และตัวแบบการเคลื่อนที่ของมือ

เมื่อนำตัวแบบดังกล่าวมาวิเคราะห์โดยสมดุลของแรง

$$m_y \ddot{y}(t) = k_{g,y}(y_g - y(t)) - k_{n,y}(y(t) - y_n)$$

$$m_y \ddot{y}(t) + (k_{g,y} + k_{n,y})y(t) = (k_{g,y}y_g + k_{n,y}y_n)$$

$$m_y \ddot{y}(t) + k_1 y(t) = k_2 \qquad (3\text{-}1)$$

โดย $k_1 = k_{g,y} + k_{n,y}$ และ $k_2 = k_{g,y}y_g + k_{n,y}y_n$ สังเกตว่าตัวแบบดังกล่าวอยู่ในรูปของสมการอนุพันธ์อันดับที่สอง ซึ่งตรงกับงานวิจัยของ Yashuhara ซึ่งเสนอว่า มือมีพฤติกรรมในรูปของ สมการอนุพันธ์อันดับที่สอง สำหรับการแก้สมการดังกล่าว ทำได้โดยกำหนดรูปแบบคำตอบที่น่าจะเป็นไปได้เพื่อแก้สมการ ซึ่งได้แก่

$$\dot{y}(t) = a\sin\left(\sqrt{\frac{k_1}{m_y}}(t-to)+\phi\right) \quad (3\text{-}2)$$

$$y(t) = -a\sqrt{\frac{m_y}{k_1}}\cos\left(\sqrt{\frac{k_1}{m_y}}(t-to)+\phi\right)+\frac{k_2}{k_1} \quad (3\text{-}3)$$

ซึ่งสุดท้ายสามารถจัดรูปให้ง่ายขึ้นดังนี้ และทำนองเดียวกันกับความเร็วแนวนอน

$$\dot{y}(t) = A_y \sin(\omega_y(t)\cdot t + \phi_y) \quad (3\text{-}4)$$

$$\dot{x}(t) = A_x \sin(\omega_x(t)\cdot t + \phi_x) + c \quad (3\text{-}5)$$

จากสมการดังกล่าวอาจกล่าวได้ว่า ความเร็วของการเขียนในแนวนอน(sweep) จะมีค่าเท่ากับค่าคงที่ c c>0 ซึ่งมาจากการเขียนที่เริ่มจากซ้ายไปขวา และสรุปได้ว่าการเขียนใดๆ สามารถสร้างขึ้นจากพารามิเตอร์ ซึ่งได้แก่ แอมปลิจูด ความถี่ เฟส และความเร็วแนวนอนคงที่ และการเขียนใดๆ สามารถแทนด้วยความเร็วของการลากเส้นทั้งในแนวตั้งและแนวนอน โดยที่ความเร็วทั้งสองแนวมีความสัมพันธ์กัน การหาพารามิเตอร์ของตัวแบบแหล่งกำเนิดนี้ถูกเสนอโดย Singer และ Tishby โดยการแบ่งลายมือเขียนเป็นเฟรม ขอบของเฟรมหาจากจุด zero-crossing ของความเร็ว ในแนวตั้งของการเขียน ซึ่งเป็นจุดที่มีการลากเส้นกลับทิศทางจากบนลงล่าง หรือล่างขึ้นบน อย่างไรก็ตาม Oscillator model ไม่สามารถใช้กับการเขียนที่ความเร็วในการลากเพิ่มขึ้นอย่างรวดเร็วได้ ดังนั้น Orly Stettiner, Dan Chazan (1994) ได้เปลี่ยนตัวแบบแหล่งกำเนิดใหม่ให้สอดคล้องขึ้นดังนี้

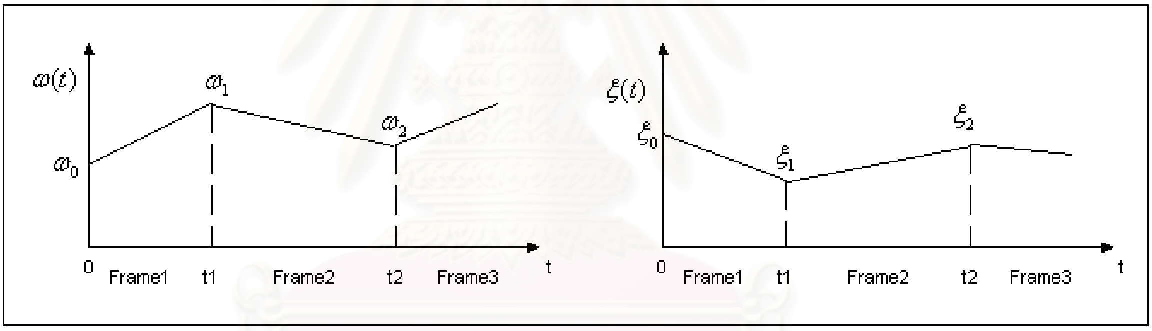
$$V_x(t) = \frac{A_x}{\sqrt{1-\xi_x(t)^2}} \sin(\omega_x(t)\sqrt{1-\xi_x(t)^2} \cdot t + \phi_x) \cdot \exp(-\xi_x(t)\omega_x(t) \cdot t) + c$$

(3-6)

$$V_y(t) = \frac{A_y}{\sqrt{1-\xi_y(t)^2}} \sin(\omega_y(t)\sqrt{1-\xi_y(t)^2} \cdot t + \phi_y) \cdot \exp(-\xi_y(t)\omega_y(t) \cdot t)$$

(3-7)

model แบบใหม่นี้อาศัยค่า ξ (damping) ซึ่งสามารถมีค่าเป็นลบได้ และ ξ ค่านี้ใช้บังคับให้ความเร็วเพิ่มขึ้นอย่างรวดเร็วได้ จึงมีความเหมาะสมมากกว่ากับการเขียนที่เกิดขึ้นจริงๆ การหาพารามิเตอร์ทำได้โดยแบ่งแนวการลากเส้นเป็นเฟรม พารามิเตอร์ ξ และ ω จะมีการเปลี่ยนแปลงอย่างช้าๆ อย่างเป็นเชิงเส้นภายในเฟรมและมีความต่อเนื่องกันด้วยที่ขอบของเฟรมดังรูป ในขณะที่แอมปลิจูด เฟส และ ความเร็วการเคลื่อนที่แนวนอนมีค่าคงที่ตลอด



รูปที่ 3.5 ลักษณะความเป็นเชิงเส้นและความต่อเนื่องของความถี่และความหน่วง

เราสามารถเขียน feature vector จากพารามิเตอร์ได้ดังนี้

$\theta = \{A, \phi, c, (\xi_i, \omega_i)\}$ โดย i =P+1 P คือ จำนวนเฟรมทั้งหมด

3.2.1 วิธีการหาพารามิเตอร์โดยใช้ค่าเฉลี่ยความผิดพลาดกำลังสอง

การหาพารามิเตอร์ ภายในเฟรมหนึ่งๆทำได้โดย nonlinear optimize โดยให้ตั้งจุดประสงค์ของการ optimization คือ การหาพารามิเตอร์ที่ทำให้ค่าเฉลี่ยความผิดพลาดกำลังสอง (mean square error) ระหว่างแนวการลากเส้นสังเคราะห์กับแนวการลากเส้นของจริงมีค่าน้อยที่สุด ให้ V เป็นความเร็วการลากเส้นจริง $V(\theta)$ เป็นความเร็วการลากเส้นที่เกิดจากการประมาณด้วยพารามิเตอร์ และ v เป็นความผิดพลาดระหว่างความเร็วจริง กับความเร็วจากการประมาณ

$$V = V(\theta) + v$$

$$S = \int V dt = \int (V(\theta) + v) dt = S' + s_{err}$$

เราจะต้องหา θ ที่ทำให้ค่าเฉลี่ยความผิดพลาดกำลังสองของระยะทางน้อยที่สุด

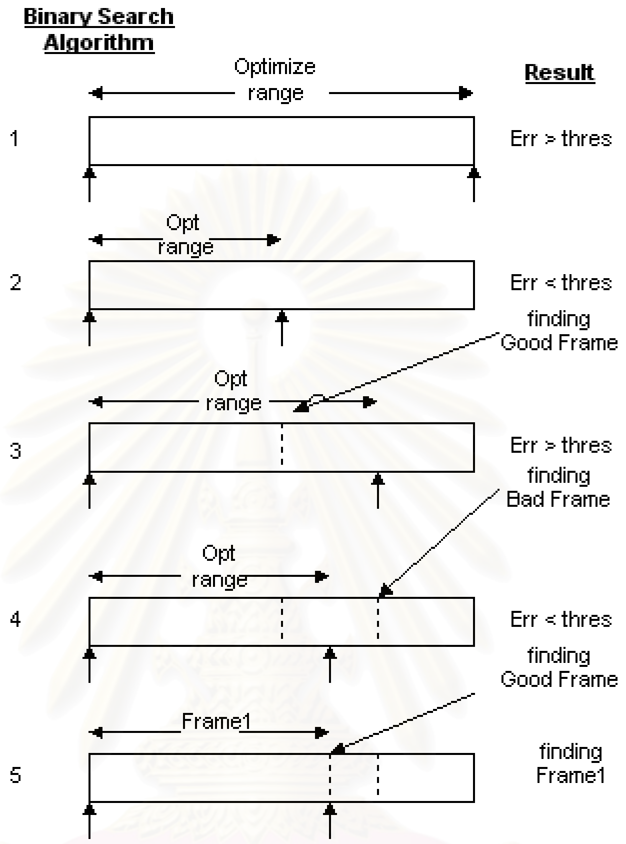
$$\theta_{ls} = Min(s^2{}_{err}) \qquad (3\text{-}8)$$

3.2.2 การหาพารามิเตอร์โดยใช้การปรับจำนวนเฟรมโดยอัตโนมัติแบบ Binaray Search

เนื่องจากเราไม่สามารถทราบว่าขอบเขตของเฟรมอยู่ที่ใด เราจะต้องหาขอบเฟรมโดยอาศัยค่าเฉลี่ยของความผิดพลาดกำลังสอง (mean square error) ค่าเฉลี่ยของความผิดพลาดกำลังสองจะมีค่าเพิ่มมากขึ้นเรื่อยเมื่อมีการเพิ่มขนาดเฟรม (monotonic increasing) เราจะต้องกำหนดค่า threshold เพื่อจำกัดขอบเขตของเฟรม โดยวิธีง่ายๆ ดังนี้

- ค่าความเฉลี่ยของความผิดพลาดกำลังสองน้อยกว่า threshold ให้บันทึกว่าขอบเฟรมดังกล่าวเป็นขอบเฟรมที่ดี ให้ทำการขยายเฟรม
- ค่าความเฉลี่ยของความผิดพลาดกำลังสองมากกว่า threshold ให้บันทึกว่าขอบเฟรมดังกล่าวเป็นขอบเฟรมทีไม่ดี ให้ทำการลดเฟรม
- การเพิ่มขอบหรือลดขอบเฟรม ให้เลือกขอบเฟรมที่อยู่ตรงกลางระหว่างขอบเฟรมที่ดีและไม่ดี
- หยุดการเพิ่มหรือลดขอบเฟรม เมื่อขอบเฟรมที่ดีและขอบเฟรมที่ไม่ดีอยู่ห่างกันหนึ่งเฟรม

3.2.3 ตัวอย่างการหาขอบเฟรม



รูปที่ 3.6 ตัวอย่างการหาขอบเฟรม

3.2.4 การหาพารามิเตอร์โดยใช้การแบ่งเฟรมที่ตำแหน่งความเร็วการลากเส้นแนวตั้งเป็นศูนย์

การใช้จุดแบ่งเฟรมเป็นจุดที่ตำแหน่งความเร็วการลากเส้นแนวตั้งเป็นศูนย์ เนื่องจากความเร็วการลากเส้นแนวตั้งเป็นศูนย์เสมือนจุดตั้งต้นของการลากเส้นเขียน เพราะ การเขียนของมือเรานั้นมักจะมีการเคลื่อนที่กลับไปมาในแนวตั้งมากที่สุด ข้อดีของวิธีการนี้ คือ การหาพารามิเตอร์ โดย optimize ทุกตัวแปรพร้อมกันอันได้แก่ A, ξ_i , ω_i , ϕ_i ,c ทำให้พารามิเตอร์แต่ละตัวมีผลต่อ ความคลาดเคลื่อนจากแนวการลากเส้น การปรับพารามิเตอร์ตัวหนึ่งจะมีผลต่อพารามิเตอร์อื่นทันที ดังนั้นทุกเฟรมจะมีความสำคัญเท่ากันหมด ดังนั้นจำนวนเฟรมในวิธีการนี้จะขึ้นกับลักษณะของวิธีการเขียนด้วย

3.2.5 ข้อดีของการใช้ตัวแบบการเคลื่อนที่ของมือ

- ช่วยลดปริมาณข้อมูล เพราะสามารถแทนแนวการลากเส้นด้วยพารามิเตอร์เพียงหนึ่งชุด
- สามารถใช้ได้กับการเขียนใดๆ ได้
- สามารถสร้างกลับคืนจากพารามิเตอร์ และ พารามิเตอร์เป็นคุณลักษณะเฉพาะของการเขียน สามารถพัฒนาไปสู่การรู้จำได้

3.3 ตัวแบบการเกิดภาพของปากกาและตัวแบบย้อนกลับแหล่งกำเนิด

จากงานวิจัยที่ผ่านมาได้มีการพยายามตั้งสมมติฐานเกี่ยวกับอุปกรณ์ในการเขียน เพื่อช่วยประกอบ ในการดึงแนวการลากเส้น หรือ หาโครงกระดูกของตัวอักษร หรือ การหาจุดเริ่มต้นและสิ้นสุดของการเขียน เพราะอุปกรณ์การเขียนจะทำให้เกิดภาพขึ้นมา ดังนั้นสิ่งที่สังเกตได้จากภาพ ส่วนหนึ่งเกิดขึ้นจากอุปกรณ์การเขียน ตัวอย่างเช่น ภาพลายมือเขียนมีระดับสีเทาไม่สม่ำเสมอตลอดแนวการลากเส้น และ ความกว้างของแนวการลากเส้นก็มีขนาดไม่คงที่ เหตุการณ์ดังกล่าวเกิดจากอุปกรณ์การเขียนซึ่งทำหน้าที่โดยตรงในการส่งปริมาณหมึกไปบนเนื้อกระดาษ และ เมื่อหมึกติดไปบนเนื้อของกระดาษหมึก ความเข้มของหมึกจะส่งผลให้ระดับความเข้มมาก หรือ กรณีที่หมึกซึมมากปริมาณความกว้างของแนวการลากเส้นก็จะมากตามไปด้วย จะต้องมีพารามิเตอร์ในการกำหนดเหตุการณ์ที่เกิดขึ้นดังกล่าว การศึกษาถึงอุปกรณ์การเขียนจึงเป็นเรื่องสำคัญอย่างยิ่ง เรามีความคิดว่าอุปกรณ์การเขียนจะส่งผลต่อความสามารถในการคำนวณหาความเร็วของการเขียนจากภาพ

รูปที่ 3.7 ภาพประกอบแนวการลากเส้นตรงแนวนอนจาก x1 ไป x2

กำหนดอัตราการไหลของหมึกเท่ากับ K volume/sec คงที่

ในการเขียนเส้นตามรูป ปริมาณหมึกที่ไหลออกมาเท่ากับ $Q = K(t_2 - t_1)$

ความเร็วในการลากเส้นเท่ากับ $$V = \frac{(x_2 - x_1)}{(t_2 - t_1)}$$

จัดรูปสมการปริมาณหมึก $$Q = \frac{K(x_2 - x_1)}{V}$$

จัดรูปสมการปริมาณหมึกต่อหน่วยความยาวเส้น $$q = \frac{Q}{(x_2 - x_1)} = \frac{K}{V}$$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

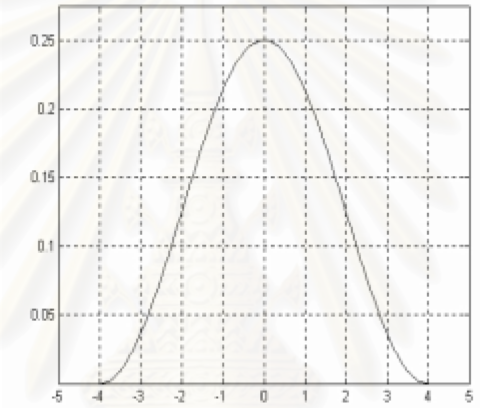
q มีหน่วยเป็น volume/ mm ปริมาณหมึกต่อหนึ่งมิลลิเมตร สังเกตว่าปริมาณหมึกมีความสัมพันธ์แบบผกผันกับความเร็วในการลากเส้น

$$q \propto \frac{1}{V} \tag{3-9}$$

ดังนั้นเราสามารถคำนวณปริมาณหมึกจากภาพลายมือเขียน โดยเราประมาณว่า ปริมาณหมึกสามารถคำนวณจากพื้นที่ใต้หน้าตัดของระดับความเข้ม $f(x)$ เราจะได้ปริมาณหมึก

$$q = \int_{-\sigma}^{\sigma} f(x)dx \tag{3-10}$$

เราสมมติว่า หน้าตัดของระดับความเข้มจะมีรูปดังนี้ โดยค่าที่มากหมายถึง ค่าความเข้มมาก



รูปที่ 3.8 หน้าตัดของระดับความเข้ม เป็นรูป Raised Cosine

$$C(x) = \frac{1}{\sigma}\cos^2\left(\frac{\pi x}{2\sigma}\right) \tag{3-11}$$

จากสมการข้างต้น σ เป็นความกว้างของหน้าตัดระดับความเข้ม ถ้าเราหาพื้นที่ใต้กราฟ ด้วยการอินทิเกรต เพื่อคำนวณหาปริมาณหมึก

$$\int_{-\sigma}^{\sigma} C(x)dx = 1 \tag{3-12}$$

ดังนั้นเราจะเปลี่ยนหน้าตัดของระดับความเข้มเพื่อให้มีความกว้างและความสูงใดๆ โดยให้ H ระดับความเข้ม และ W เป็นความกว้างของหน้าตัดระดับความเข้ม

$$C(x) = H\cos^2\left(\frac{\pi x}{2\sigma}\right) \tag{3-13}$$

$$q = \int_{-\sigma}^{\sigma} C(x)dx = H\sigma \cong H \cdot W \tag{3-14}$$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ดังนั้นถ้าหน้าตัดระดับความเข้มเป็น Raised Cosine และปริมาณหมึกสามารถคำนวณจากพื้นที่ใต้หน้าตัดระดับความเข้ม ปริมาณหมึกสามารถหาได้จาก H และ W จากความเป็นที่จริงที่ว่า เมื่อเขียนด้วยความเร็วน้อยๆ ระดับความเข้ม(H)จะสูง และ ความกว้าง(W) ของแนวการลากจะมาก ดังนั้นปริมาณหมึก(q) จะมากตามไปด้วย แต่ถ้าเขียนด้วยความเร็วมากๆ ระดับความเข้ม(H)จะต่ำ และ ความกว้างของเส้น(W)จะน้อยด้วย เพราะปริมาณหมึก(q) จะน้อยลงไปด้วยแสดงว่า ผลคูณของความกว้างและระดับความเข้มแปรผกผันกับความเร็วการลากเส้น

ในงานวิจัยนี้กำหนดความสัมพันธ์ระหว่างความกว้างของแนวการลากเส้นกับความเร็วการลากเส้น และ ระดับความเข้มของแนวการลากเส้นกับความเร็วการลากเส้น ไว้ตามลำดับดังนี้

$$W = \begin{cases} W_o & ,V < V_{thres1} \\ \dfrac{W_o}{\left(\dfrac{V}{V_{thres1}}\right)^{\frac{2}{3}}} & ,V > V_{thres1} \end{cases} \qquad (3\text{-}15)$$

จากความสัมพันธ์ข้างต้น V_{thres1} มีหน้าที่จำกัดว่า ในช่วงที่ความเร็วต่ำๆ จะมีความกว้างสูงสุดไม่เกิน W_o แต่ความกว้างจะมีค่าน้อยเมื่อความเร็วการลากเส้นมีค่ามากกว่า V_{thres1}

$$H = \begin{cases} H_o & ,V < V_{thres\,2} \\ \dfrac{H_o}{\left(\dfrac{V}{V_{thres\,2}}\right)^{\frac{1}{3}}} & ,V > V_{thres\,2} \end{cases} \qquad (3\text{-}16)$$

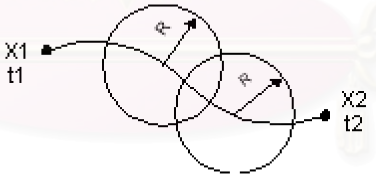
จากความสัมพันธ์ข้างต้น V_{thres2} มีหน้าที่จำกัดว่า ในช่วงที่ความเร็วต่ำๆ จะมีระดับความเข้มไม่เกิน H_o แต่ระดับความเข้มจะมีค่าน้อยเมื่อความเร็วการลากเส้นมีค่ามากกว่า V_{thres2} ข้อดีของการสมมติแบบนี้ คือ เนื่องจากการเปลี่ยนแปลงของความกว้างและระดับความเข้ม ไม่ได้เกิดที่ความเร็วเดียวกัน นั่นคือ ระดับความเข้มน่าจะมีช่วงของการเปลี่ยนแปลงมากกว่า ระดับความเข้มมีการเปลี่ยนแปลงแม้ว่าจะเขียนด้วยความการลากเส้นต่ำๆ ในขณะที่ความกว้างจะเริ่มเปลี่ยนแปลงเมื่อเขียนด้วยความเร็วการลากเส้นที่สูงกว่า ดังนั้นอาจจะสรุปได้ว่า $V_{thres1} > V_{thres2}$ และในกรณีที่เราให้หน้าตัดระดับความเข้มเป็น Raised Cosine ทำให้เราคำนวณหาปริมาณหมึก ได้ง่ายขึ้น คือ

$$q = \begin{cases} H_o W_o & , V < V_{thres\,2} \\ \dfrac{W_o H_o}{\left(\dfrac{V}{V_{thres\,2}}\right)^{\frac{1}{3}}} & , V_{thres\,2} < V < V_{thres\,1} \\ \dfrac{W_o H_o}{V} (V_{thres\,2})^{\frac{1}{3}} (V_{thres\,1})^{\frac{2}{3}} & , V > V_{thres\,1} \end{cases} \quad (3\text{-}17)$$

จากสมมติฐานข้างต้นปริมาณหมึกแปรผกผันกับความเร็วการลากเส้น เมื่อความเร็วการลากเส้นมากกว่าV_{thres1} และมีปริมาณคงที่ในช่วงความเร็วการลากเส้นที่ต่ำกว่าV_{thres2} อย่างไรก็ตาม ความสัมพันธ์ระหว่างปริมาณหมึกกับความเร็วการลากเส้น เมื่อความเร็วการลากเส้นมีค่าเป็นศูนย์จะปริมาณหมึกไหลออกมาเป็นอนันต์ ดังนั้นจะเห็นว่าความกว้างจะถูกกำหนดให้ค่ามากที่สุด เมื่อความเร็วการลากเส้นมีค่าเป็นศูนย์หรือมีค่าน้อยๆ เช่นเดียวกับ ระดับความเข้มจะถูกกำหนดให้มีค่ามากที่สุด เมื่อความเร็วการลากเส้นมีค่าเป็นศูนย์หรือมีค่าน้อยๆ และที่ช่วงความเร็วการลากเส้นมีค่าน้อยๆ นี้จะมีปริมาณหมึกไหลอยู่ค่าหนึ่งเสมอ คือ ปริมาณหมึกที่อิ่มตัว q_{sat}

3.3.1 **การใช้ตัวแบบการเกิดภาพของปากกาไปสร้างภาพ**

เนื่องจากความสัมพันธ์ระหว่างความกว้างกับความเร็วการลากเส้น และ ความสัมพันธ์ระหว่างระดับความเข้มกับความเร็วการลากเส้น เรามีวิธีการสร้างมีดังนี้



รูปที่ 3.9 การสร้างภาพจำลองโดยการซ้อนทับด้วยวงกลม

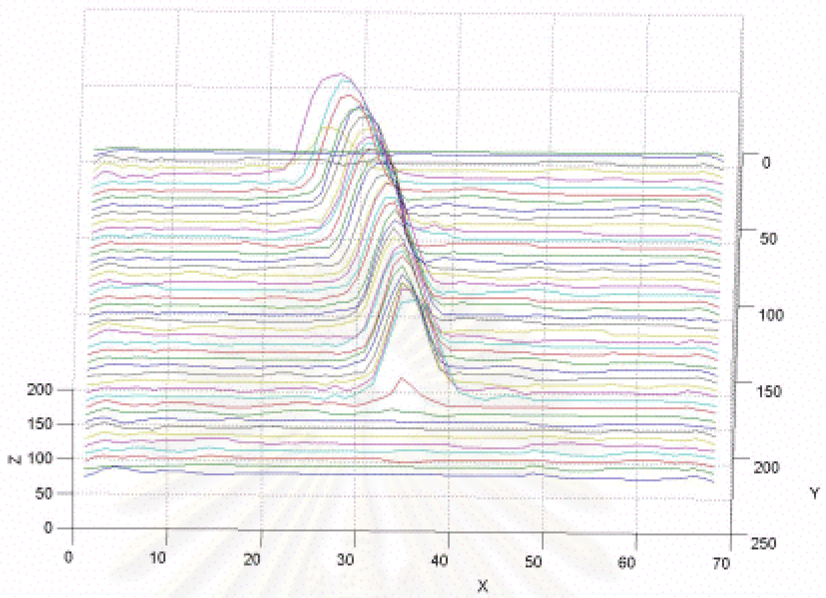
เมื่อเราทราบความเร็วการลากเส้น จะทำให้ทราบขนาดความกว้างของแนวการลากเส้น Wและระดับความเข้ม H จากความสัมพันธ์ที่ได้กล่าวไปแล้ว ให้นำขนาดความกว้างไปสร้างวงกลมที่มีรัศมีเท่ากับ R=W/2 ทุกตำแหน่งภายในวงกลมที่มีค่าระดับความเข้มเป็นไปตาม Raised Cosine แต่ละตำแหน่งมีระดับความเข้มดังนี้

$$h(x,y) = C(x)C(y) \quad (3\text{-}18)$$

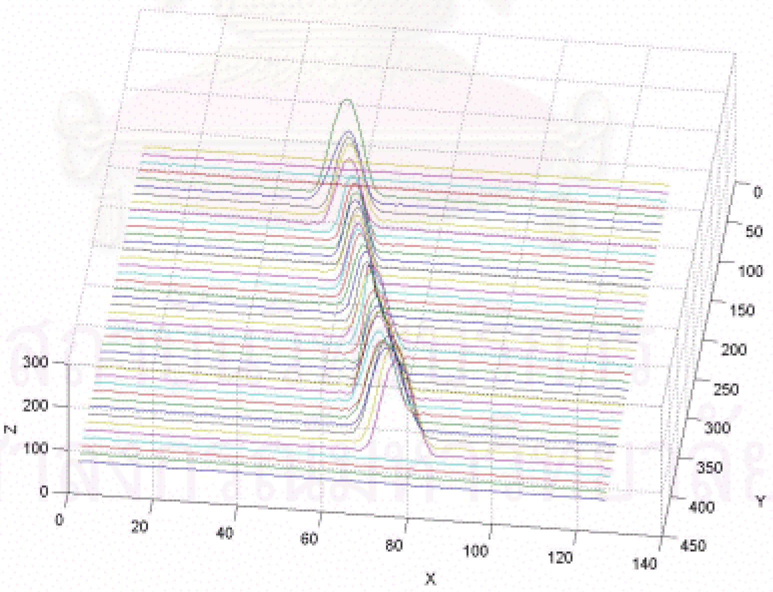
$$C(x) = \begin{cases} H \cos^2\left(\dfrac{\pi x}{2\sigma}\right), & |x| \le \sigma \\ 0 \end{cases} \quad (3\text{-}19)$$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

3.3.2 หน้าตัดระดับความเข้มจากภาพจริงและจากภาพจำลอง



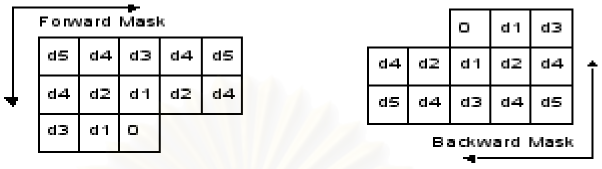
รูปที่ 3.10 หน้าตัดระดับความเข้มของปากกาเมจิค



รูปที่ 3.11 หน้าตัดระดับความเข้มที่ได้จาก Raised cosine

3.3.3 วิธีการคำนวณปริมาณหมึกจากภาพระดับสีเทา

เราสามารถคำนวณปริมาณหมึกจากภาพระดับสีเทาได้ โดยใช้วิธีคล้ายๆกับการแปลงระยะทาง โดยให้แต่ละจุดภาพมีน้ำหนักในการคำนวณแปรตามระดับความเข้มของจุดนั้นๆ ผลลัพธ์จะเป็นการอินทิเกรตพื้นที่ของหน้าตัด W-H ที่แต่ละจุด



รูปที่ 3.12 Forward Mask และ Backward และทิศทางการแสกนภาพ

$$q_{i,j} = \min_{(k,l)\in mask} (q_{i+k,j+l} + I_{i,j}d(k,l)) \qquad (3\text{-}20)$$

$q_{i,j}$ คือ ปริมาณหมึกสะสมที่ตำแหน่ง (i,j)

d(k,l) คือ ค่า local distance ที่ตำแหน่ง (k,l) ภายใน mask ที่ใช้ในการแสกน

$I_{i,j}$ คือ ระดับความเข้มที่ตำแหน่ง (i,j)

ทำการสะสมปริมาณหมึกทั้ง 2 ทิศทาง จะได้แผนที่ปริมาณหมึกสะสมในทิศทางที่ต่างกัน 2 ชุด แล้วจึงนำแผนที่ปริมาณหมึกสะสมทั้งสองมาเปรียบเทียบ ที่ตำแหน่งเดียวกันเลือกปริมาณหมึกสะสมที่มีค่าน้อยที่สุด แล้วบันทึกปริมาณหมึกดังกล่าวในแผนที่ปริมาณหมึกสุดท้าย ให้ทำจนกระทั่งครบทุกตำแหน่งภาพ จากการทำแผนที่ปริมาณหมึกจะมีลักษณะคล้ายแผนที่ระยะทางทำให้เราสามารถหาโครงกระดูกของภาพได้อีกด้วย

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

3.3.4 ตัวแบบย้อนกลับการเกิดภาพของปากกา

$$V = \begin{cases} V_{thres2} & ,q \geq q_{sat} \\ \dfrac{(H_oW_o)^3}{q^3} \cdot V_{thres2} & ,q_{sat} > q \geq q_{pt} \\ \dfrac{H_oW_o}{q} \cdot (V_{thres1})^{\frac{2}{3}} (V_{thres2})^{\frac{1}{3}} & ,q < q_{pt} \end{cases} \qquad (3\text{-}21)$$

$$q_{sat} = H_oW_o$$
$$q_{pt} = \frac{H_oW_o}{(\frac{V_{thres1}}{V_{thres2}})^{\frac{1}{3}}} \qquad (3\text{-}22)$$

จากความสัมพันธ์ (3-17) เราสามารถหาตัวแบบย้อนกลับการเกิดภาพ หรือ ความสัมพันธ์ระหว่างปริมาณหมึกที่คำนวณได้จากหัวข้อ 3.3.3 และ ความเร็วการลากเส้นได้ดังสมการ (3-21 และ 3-22) เนื่องจากตัวแบบการเกิดภาพจะให้ปริมาณหมึก(q) ที่มากที่สุดเท่ากับ H_oW_o ที่ความเร็วการลากเส้นต่ำๆ ทำนองเดียวกัน ในตัวแบบย้อนกลับการเกิดภาพ ปริมาณหมึก q จะถือว่าอิ่มตัวที่ค่า H_oW_o เช่นกัน เราจึงกำหนดให้ความเร็วการลากเส้นเท่ากับ V_{thres2} ซึ่งถือว่าเป็นความเร็วต่ำสุด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4
การทดลอง ผลการทดลอง และวิเคราะห์

4.1 เครื่องมือที่ใช้ในการทดลอง

- เครื่องคอมพิวเตอร์ส่วนบุคคล ใช้หน่วยประมวลผลรุ่น Pentium ความเร็ว 400 MHz หน่วยความจำ 64 Mbyte
- โปรแกรมดึงข้อมูลการลากเส้น ที่พัฒนาโดยภาษาคอมพิวเตอร์ Visual C++ เวอร์ชัน 6 ของบริษัทไมโครซอฟท์ และทำงานภายใต้ระบบปฏิบัติการวินโดว์ 2000
- อุปกรณ์รับข้อมูลการเขียน Tablet ของบริษัท Wacom ซึ่งเชื่อมต่อเข้ากับเครื่องคอมพิวเตอร์ส่วนบุคคลผ่านทางพอร์ตอนุกรม
- โปรแกรมดึงพารามิเตอร์ของตัวแบบการเคลื่อนที่ของมือ ที่พัฒนาโดยภาษาคอมพิวเตอร์ Matlab เวอร์ชัน 6.0

4.2 แหล่งที่มาของข้อมูล

ข้อมูลที่ได้จากการเขียนตัวอักษรตัวเดียว ผ่านทาง Tablet ได้แก่ ตำแหน่งของ X และ Y ค่า (X,Y) ที่ได้จะเริ่มจาก (0,0) ที่มุมบนซ้ายของ Tablet และมีค่ามากขึ้นเรื่อยๆ จนมากที่สุดที่มุมล่างขวา (Absolute mode) และ เวลาของการลากผ่านตำแหน่ง (X,Y) ซึ่งมีหน่วย คือ millisecond ในที่นี้เราได้กำหนดความละเอียดของการรับข้อมูลที่ 300 จุด ต่อ 1 นิ้ว แต่เนื่องจากปากกาของ Tablet จะไม่มีหมึกออกมา หรือถ้ามีก็จะไม่ใช่อุปกรณ์การเขียนที่ต้องการ จึงแยกมาใช้ปากกาเขียนภาพ เพื่อนำมาแสกนต่างหากอีกครั้งหนึ่ง ภาพที่นำมาเป็นต้นแบบในการเปรียบเทียบได้จากการแสกนที่ความละเอียด 300 จุดภาพต่อ 1นิ้ว

4.3 การทดสอบตัวแบบการเกิดภาพของปากกา

การทดสอบนี้มีวัตถุประสงค์ เพื่อหาวิธีการสร้างภาพจากตัวแบบการเกิดภาพของปากกา ที่คล้ายกับปากกาที่มีอยู่จริง การสร้างภาพจากตัวแบบการเกิดภาพของปากกา จะใช้ความเร็วการลากเส้นของการเขียนจริงๆเป็นอินพุต และได้เอาต์พุทเป็นภาพระดับสีเทา โดยตัวแบบดังกล่าวจะต้องสอดคล้องกับความเร็วการลากเส้น การทดสอบมีขั้นตอนดังนี้

- ทำการเก็บข้อมูลของภาพสแกนลายมือเขียน ที่แสกนที่ความละเอียด 300 จุดภาพ ต่อ 1 นิ้ว และ มีระดับความเข้ม 256 ระดับ เนื่องจากที่ความละเอียดดังกล่าวเป็นที่นิยมใช้ในระบบการรู้จำโดยทั่วไป เราจึงต้องทำให้ตัวแบบสอดคล้องกันด้วย สังเกตระดับความเข้ม และ ความกว้างของเส้น
- รับข้อมูลการเขียนจาก Tablet ตัวอักษรเดี่ยวตัวหนึ่ง ได้แก่ ตำแหน่ง (X,Y) และ เวลา

- ทำการ Sampling by Distance ค่า (X,Y) และ เวลา คำนวณความเร็วให้สังเกตความเร็วมากที่สุด และ น้อยที่สุด
- ทำการสร้างภาพโดยสร้างเส้นตรงเชื่อมจุด X-Y ที่ผ่านการ Sampling ใช้ระดับความเข้ม และความกว้างของเส้นเป็นไปตามตัวแบบการเกิดภาพของปากกา
- เปรียบเทียบตัวแบบการเกิดภาพของปากกา แบบใดๆ คล้ายกับปากกาจริง

4.3.1 **การเลือกค่าระยะทางสำหรับทำ** Sampling by Distance

เราจำเป็นต้องจำกัดจำนวนตำแหน่ง (X,Y) ในการนำมาสร้างภาพ เนื่องจากการรับข้อมูลตำแหน่ง (X,Y) โปรแกรมคอมพิวเตอร์ที่เขียนขึ้นมามีความอ่อนไหวในการเก็บข้อมูล ในบางครั้งจะบันทึกตำแหน่ง (X,Y) ซ้ำกัน แต่ที่เวลาต่างกันเข้ามาด้วย เราต้องการเปลี่ยนให้เก็บข้อมูลทุกครั้งที่มีการเคลื่อนที่ไปเป็นระยะทางหนึ่งๆ ที่เหมาะสมเท่านั้น

รูปที่ 4.1 Sampling By Distance ที่ความยาวแตกต่างกัน

การเลือกค่าระยะทางเพื่อทำ Sampling By Distance ที่มีค่ามาก จะทำให้การเก็บข้อมูลเกิดขึ้นแบบหยาบๆ จากตัวอย่างแสดงให้เห็นว่า แต่ละตำแหน่งที่ถูกบันทึกเมื่อนำสร้างเป็นรูป มีความเป็นเหลี่ยมเป็นมุมมากเกินจริง ทั้งที่ควรจะเป็นส่วนโค้ง ในขณะที่การเลือกค่าระยะทางที่มีค่าน้อยๆ จะให้ภาพที่มีความโค้งเหมือนจริง แต่ทำให้ต้องเก็บข้อมูลที่มากกว่า ในที่นี้เราเลือกค่าระยะทางนี้ให้เท่ากับ 20 จุดภาพ หรือ 1.69 มิลลิเมตร

4.3.2 **ตัวอย่างของภาพลายมือเขียนที่ผ่านสแกน**

ภาพลายมือเขียนนี้ได้นำมาจากการสแกนที่ความละเอียด 300 จุดต่อ 1 นิ้ว ที่เกิดจากปากกาเมจิค โดยปากกาที่เราเลือกขึ้นมา แสดงให้เห็นถึงความหลากหลายของความกว้างของแนวการลากเส้น และ ระดับความเข้มด้วย

รูปที่ 4.2 ตัวอย่างของความกว้างและระดับความเข้มของแนวการลากเส้นที่แตกต่างกัน

จากการสังเกตภาพที่แสกนเข้ามามีความกว้างของแนวการลากเส้นมากที่สุด 12 จุดภาพ และน้อยที่สุด 6 จุดภาพ โดยระดับความเข้มและความกว้างของแนวการลากเส้นจะถูกนำมาสร้างเป็นตัวแบบการเกิดภาพของปากกา

4.3.3 **ความเร็วในการลากเส้นเขียนตัวอักษร**

เนื่องจากการสร้างตัวแบบการเกิดภาพของปากกาจะต้องรู้ความเร็วในการลากเส้นที่ค่ามากที่สุดและน้อยที่สุด ซึ่งเราสามารถสังเกตได้จากการรับข้อมูลจาก Tablet พบว่าความเร็วที่ใช้ในการลากเส้นที่มีค่ามากที่สุดและน้อยที่สุดประมาณ 17 mm/sec และ 180 mm/sec ตามลำดับ

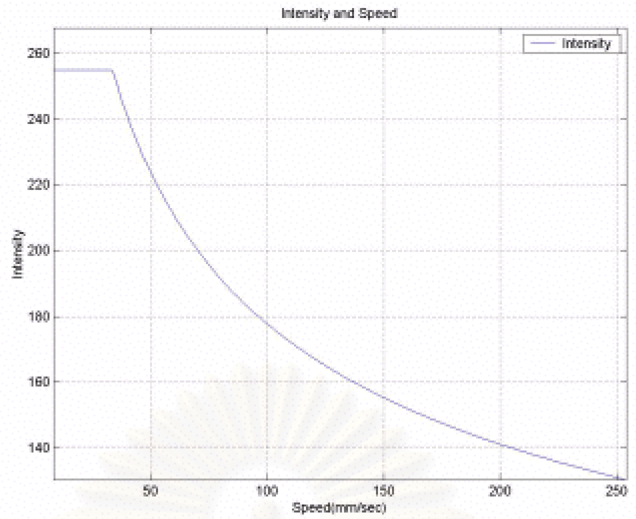
4.3.4 **การสร้างตัวแบบการเกิดภาพของปากกา**

จากสมมติฐานในบทที่ 3 เราเลือกความสัมพันธ์ระหว่างระดับความเข้มและความเร็ว ความกว้างและความเร็ว ดังสมการต่อไปนี้

- ความสัมพันธ์ระหว่างระดับความเข้มและความเร็วการลากเส้น

$$H = \frac{H_o}{\left(\frac{V}{V_{thres\,2}}\right)^{\frac{1}{3}}} \qquad (4\text{-}1)$$

ในที่นี้เราเลือกให้ V_{thres2} เท่ากับ 33.88 mm/sec ซึ่งหมายความว่า ในช่วงที่ความเร็วของการลากเส้น มีค่าน้อยกว่านี้ ระดับความเข้มจะไม่มีการเปลี่ยนแปลงและมีค่าเท่ากับ H_o ในที่นี้เราเลือกให้ค่าระดับความเข้มเท่ากับ 255

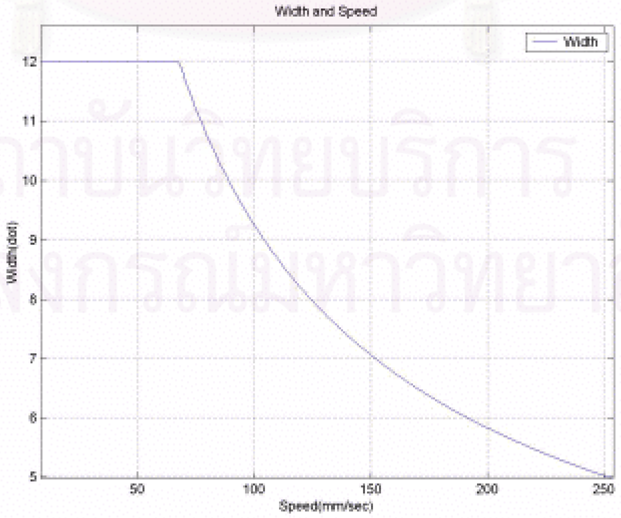


รูปที่ 4.3 แนวโน้มการเปลี่ยนแปลงระดับความเข้มตามความเร็วการลากเส้น

- ความสัมพันธ์ระหว่างระดับความกว้างและความเร็ว

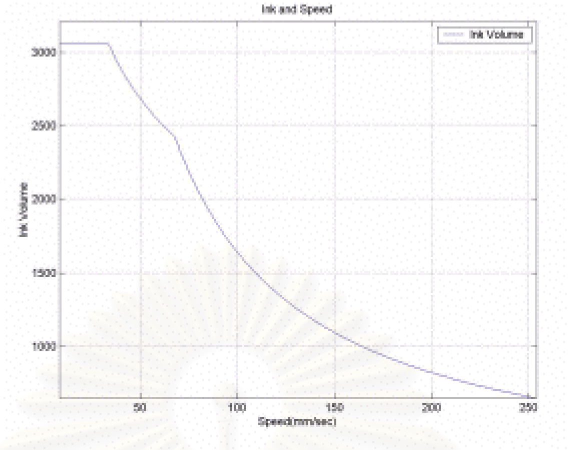
$$W = \frac{W_o}{\left(\frac{V}{V_{thres1}}\right)^{\frac{2}{3}}} \qquad (4\text{-}2)$$

ขณะเดียวกันเลือก V_{thres1} เท่ากับ 67.76 mm/sec ซึ่งหมายความว่า ในช่วงที่ความเร็วของการลากเส้นน้อยกว่านี้ ระดับความกว้างจะไม่เปลี่ยนแปลงและมีค่าเท่ากับ W_o ซึ่งเราเลือกให้ความกว้างนี้เท่ากับ 12 จุดภาพ ซึ่งความกว้างที่มากที่สุด



รูปที่ 4.4 แนวโน้มการเปลี่ยนแปลงระดับความกว้างตามความเร็วการลากเส้น

- ปริมาณที่คำนวณจากความสัมพันธ์ระหว่างระดับความเข้ม ความกว้าง และความเร็วการลากเส้น



รูปที่ 4.5 ปริมาณหมึกที่สัมพันธ์กับความเร็วของการลากเส้น

$$q=\begin{cases} H_oW_o & ,V<V_{thres\,2}(33.88\,mm/\sec) \\ \dfrac{W_oH_o}{\left(\dfrac{V}{V_{thres2}}\right)^{\frac{1}{3}}} & ,V_{thres2}(33.88\,mm/\sec)<V<V_{thres\,1}(67.76\,mm/\sec) \\ \dfrac{W_oH_o}{V}(V_{thres2})^{\frac{1}{3}}(V_{thres1})^{\frac{2}{3}} & ,V>V_{thres\,1}(67.76\,mm/\sec) \end{cases} \tag{4-3}$$

รูปที่ 4.6 ตัวอย่างภาพที่สร้างจากตัวแบบการเกิดภาพของปากกา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จากรูปที่สร้างขึ้นมาพบว่ามีความใกล้เคียงกับภาพจริง ซึ่งสังเกตได้จากที่บริเวณที่เขียนช้า ได้แก่ หัวของตัวอักษร บริเวณดังกล่าวมีระดับความเข้มและความกว้างมากกกว่าส่วนอื่นๆ ในส่วนปลายของการเขียน เกิดการตวัด ความเร็วการลากเส้นมีค่ามาก ระดับความเข้มและความกว้างของเส้นจะลดลงด้วย

4.4 การดึงความเร็วการลากเส้นโดยตัวแบบย้อนกลับการเกิดภาพของปากกา

การทดสอบนี้มีวัตถุประสงค์ เพื่อดึงความเร็วการลากเส้นจากภาพลายมือเขียน ที่เกิดขึ้นโดยตรงจากตัวแบบการเกิดภาพของปากกา โดยทดสอบกับภาพที่สร้างขึ้นจากหัวข้อ 4.3 ข้อมูลการลากเส้นอยู่ในรูปของแนวโน้มของความเร็วการลากเส้น ตัวแบบย้อนกลับของการเกิดภาพของปากกาจะนำระดับความเข้ม และ ความกว้างของแนวการลากเส้น ไปประเมินหาความเร็วของการลากเส้น การทดสอบมีขั้นตอนดังนี้

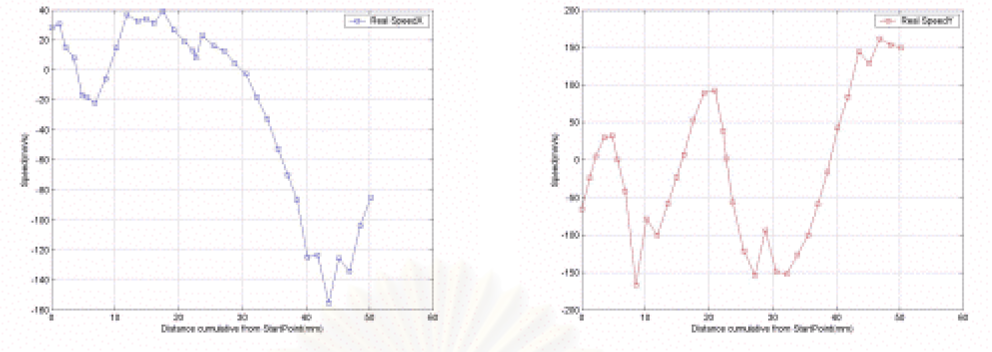
- นำภาพที่ได้จากขั้นตอนในหัว 4.3 มาประมวลโดยตัวแบบย้อนกลับการเกิดภาพของปากกา
- ทำการหาโครงกระดูกของภาพ โดยใช้แผนที่ระยะทางที่คำนวณจากปริมาณหมึกดังกล่าวถึงในหัวข้อที่ 3.3.3 ตำแหน่งของภาพที่เป็นโครงกระดูกเท่านั้นที่นำมาวิเคราะห์ประเมินหาแนวโน้มของความเร็วการลากเส้น
- เนื่องจากภาพที่สร้างขึ้นมาขาดลำดับของการลาก ต้องทำการกำหนดตำแหน่งของการลากเส้น ตามลำดับให้เหมือนกับการลากเส้นจริง
- นำความเร็วของที่ประเมินได้จากตัวแบบย้อนกลับไปเปรียบเทียบกับความเร็วที่เกิดขึ้นจริงของการลากเส้นที่ได้เก็บ

4.4.1 การหาโครงกระดูกด้วยตัวแบบย้อนกลับการเกิดภาพของปากกา

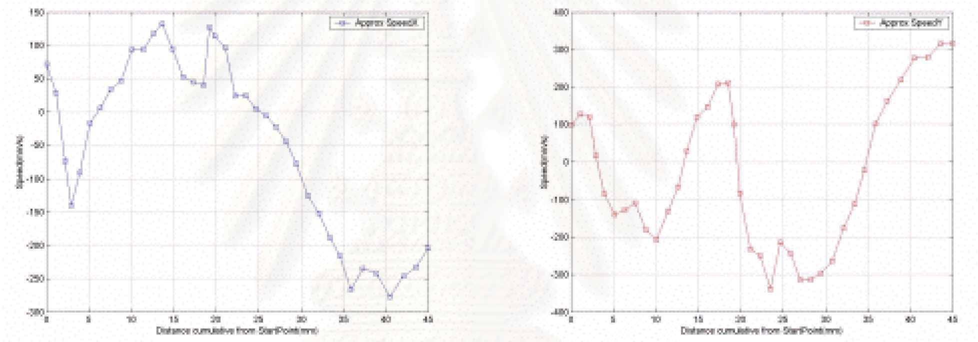
ก) ข)

รูปที่ 4.7 ตัวอย่างภาพ ก) ตำแหน่งต่างๆที่ถูกบันทึกความเร็ว ข) ภาพโครงกระดูกและตำแหน่งที่ถูกนำมาประมาณหาความเร็วการลากเส้น

4.4.2 การใช้ตัวแบบย้อนกลับเพื่อดึงความเร็วจากปริมาณหมึก



รูปที่ 4.8 ความเร็วแนวนอนและแนวตั้งของการลากเส้นที่เกิดขึ้นจริง



รูปที่ 4.9 ความเร็วแนวนอนและแนวตั้งที่เกิดจากการประมาณด้วยตัวแบบย้อนกลับการเกิดภาพของปากกา

การใช้ตัวแบบย้อนกลับแหล่งกำเนิดภาพของปากกาสามารถหาความเร็วการลากเส้นที่ใกล้เคียงกับความเร็วการลากเส้นที่เกิดขึ้นจริง ทั้งความเร็วการลากเส้นแนวนอนและแนวตั้ง ตัวแบบย้อนกลับในการหาความเร็วจากปริมาณหมึก (q) ที่คำนวณได้

$$V = \begin{cases} V_{thres2} & , q \geq q_{sat} \\ \dfrac{(H_o W_o)^3}{q^3} \cdot V_{thres2} & , q_{sat} > q \geq q_{pt} \\ \dfrac{H_o W_o}{q} (V_{thres1})^{\frac{2}{3}} (V_{thres2})^{\frac{1}{3}} & , q < q_{pt} \end{cases} \qquad (4\text{-}4)$$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

โดยที่

$$q_{sat} = H_o W_o = 255*12 = 3060$$

$$q_{pt} = \frac{H_o W_o}{(\frac{V_{thres1}}{V_{thres2}})^{\frac{1}{3}}} = \frac{255*12}{\left(\frac{67.76}{33.88}\right)^{\frac{1}{3}}} = 2428.76 \qquad (4\text{-}5)$$

4.5 การหาพารามิเตอร์ของตัวแบบการเคลื่อนที่ของมือ

ในการทดลองนี้เราต้องการทดสอบประสิทธิภาพของตัวแบบการเคลื่อนที่ของมือ เพื่อนำไปใช้โดยทำการทดลองการข้อมูลที่เกิดจากการเขียนจริงที่สุ่มขึ้นมาจาก Tablet ด้วยความถี่ 50 เฮิร์ต ข้อมูลที่ได้ คือ การเขียนตัวอักษรตัวเดี่ยวที่ประกอบด้วยการลากเส้นเพียงหนึ่งครั้ง หรือ โดยไม่มีการยกปากกา พารามิเตอร์ที่ได้จะต้องสามารถสร้างภาพได้ใกล้เคียงจริง เราต้องการหาพารามิเตอร์ที่ทำให้เกิดการเขียนที่เป็นตัวอักษรลายมือเขียนของแบบดังกล่าว วิธีการหาพารามิเตอร์ของตัวแบบการเคลื่อนที่ของมือ มี 2 วิธีหลัก คือ

4.5.1 การหาพารามิเตอร์โดยใช้การปรับจำนวนเฟรมโดยอัตโนมัติแบบ Binary Search

การแบ่งลายมือเขียนเป็นเฟรม แล้วทำการโดยพิจารณาจากค่าเฉลี่ยความผิดพลาดกำลังสองของระยะทาง ในงานวิจัยได้ทดลองและพบความผิดพลาดของวิธีการนี้ ซึ่งจะอธิบายได้ดังนี้ คือ

1. การหาพารามิเตอร์โดยเริ่มต้นหาพารามิเตอร์ในเฟรมแรกก่อน ซึ่งได้แก่ A ξ_1 ξ_2 ω_1 ω_2 ϕ และ sweep ได้พารามิเตอร์ของเฟรมแรกที่ทำให้ค่าเฉลี่ยผิดพลาดกำลังสองน้อย เนื่องจากมีจำนวนตัวแปรที่ใช้ในการ optimize มาก พารามิเตอร์ที่ได้ที่ตำแหน่งสุดท้ายของเฟรมแรก ξ_2 ω_2 จะเป็นค่าพารามิเตอร์ต้นของเฟรมที่สอง จากนั้นทำการหาพารามิเตอร์ปลายของเฟรมสอง ξ_3 และ ω_3 การหาพารามิเตอร์ในขั้นตอนนี้มีจำนวนตัวแปรน้อยเหลือเพียงสองตัว การหาพารามิเตอร์มีความจำกัด (degree of freedom น้อย) แนวการลากเส้นที่พยายามสร้างขึ้นมีความแตกต่างจากแนวการลากเส้นของจริง และโดยอัลกอริทึมของการหาขอบเฟรมจะทำการลดขนาดความยาวของเฟรม ในบางครั้งจนกระทั่งเหลือความยาวเฟรมน้อยมากจนเป็น 1 จุด
2. การใช้ค่าเฉลี่ยความผิดพลาดกำลังสองที่เพิ่มขึ้นอย่างรวดเร็วเพื่อกำหนดจำนวนเฟรม จะให้คำตอบที่ไม่ดีนัก ซึ่งจากการทดลองพบว่า ค่าเฉลี่ยผิดพลาดกำลังสองจะไม่ใช่ monotonic increasing function ตามความยาวของเฟรม ทำให้มีคำตอบได้หลายชุด

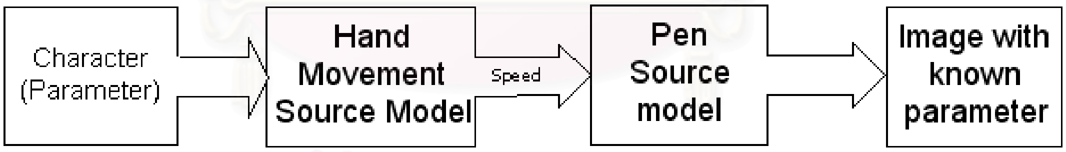
3. ความไม่มีเสถียรภาพของพารามิเตอร์ การใช้ค่าเริ่มต้นในการหาค่าพารามิเตอร์ที่แตกต่างกันได้คำตอบของมากกว่าหนึ่งชุด ซึ่งสาเหตุมาจากการที่มีตำแหน่งของการตัดเฟรมมาก ทำให้เราไม่สามารถแทนข้อมูลของการเขียนด้วยพารามิเตอร์แบบนี้ เพราะอาจจะมีมากกว่าหนึ่งแบบ นอกจากนี้ในงานวิจัยที่เป็นต้นฉบับได้ระบุถึงความจำเป็นในการศึกษาเสถียรภาพของตัวแบบแหล่งกำเนิดนี้

4.5.2 การหาพารามิเตอร์โดยใช้การแบ่งเฟรมที่ตำแหน่งความเร็วการลากเส้นแนวตั้งเป็นศูนย์

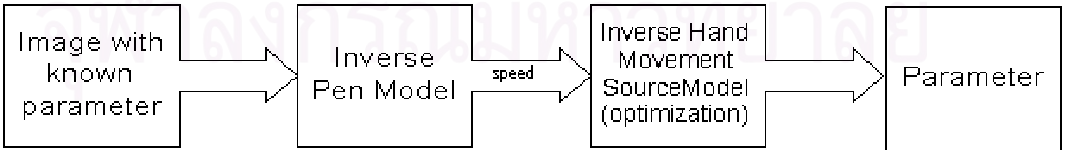
การใช้จุดแบ่งเฟรมเป็นจุดที่ตำแหน่งความเร็วการลากเส้นแนวตั้งเป็นศูนย์ การหาพารามิเตอร์ โดย optimize ทุกตัวแปรพร้อมกันอันได้แก่ A ξ_i ω_i ϕ_i และ c ทำให้พารามิเตอร์แต่ละตัวมีผลต่อ ความคลาดเคลื่อนจากแนวการลากเส้น การปรับพารามิเตอร์ตัวหนึ่งจะมีผลต่อพารามิเตอร์อื่นทันที ดังนั้นทุกเฟรมจะมีความสำคัญเท่ากันหมด การหาพารามิเตอร์แบบนี้จะทำให้ได้แนวการลากเส้นสังเคราะห์ ที่ใกล้เคียงกับของจริง การทดลองของวิธีนี้จะอยู่ในหัวข้อถัดไป

4.6 การทดลองดึงข้อมูลการลากเส้นจากภาพระดับสีเทาที่สังเคราะห์ขึ้น

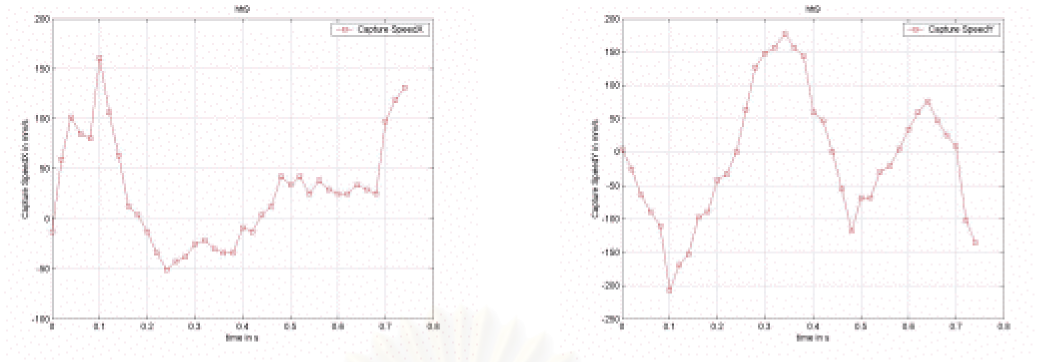
ในที่นี้เราได้สร้างภาพระดับสีเทาจากตัวแบบการเคลื่อนที่ของมือ และตัวแบบแหล่งกำเนิดภาพของปากกา การทดลองนี้ต้องการทดสอบการดึงข้อมูลแบบโดยตรงจากภาพที่สร้างขึ้น เพื่อตรวจสอบข้อมูลที่เกิดขึ้นในตอนเขียนว่ายังคงเหลืออยู่หรือไม่



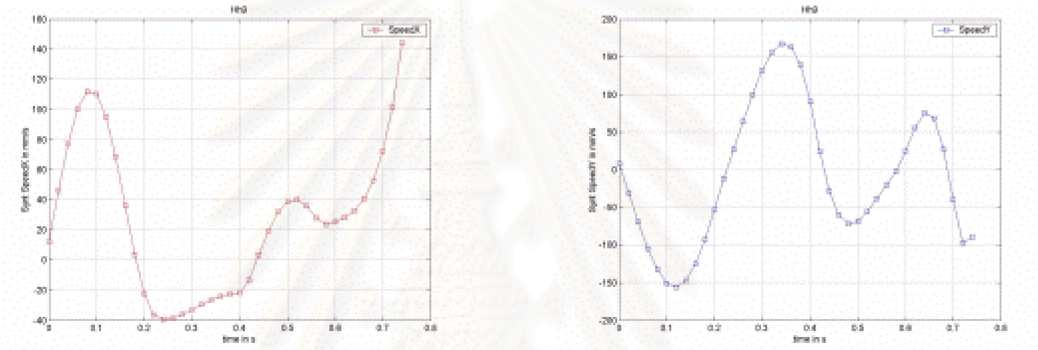
รูปที่ 4.10 ขั้นตอนการสังเคราะห์ภาพลายมือเขียน



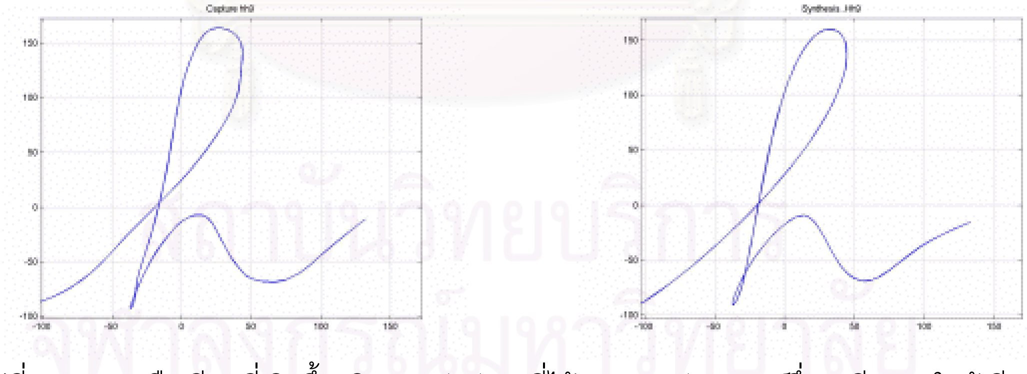
รูปที่ 4.11 ตัวอย่างของการหาพารามิเตอร์ของตัวแบบการเคลื่อนที่ของมือ



รูปที่ 4.12 ความเร็วการลากเส้นในแนวนอนและแนวตั้ง ที่ได้จากการเก็บข้อมูลทาง Tablet



รูปที่ 4.13 ความเร็วการลากเส้นในแนวนอนและแนวตั้งที่ได้จากการสังเคราะห์โดยพารามิเตอร์



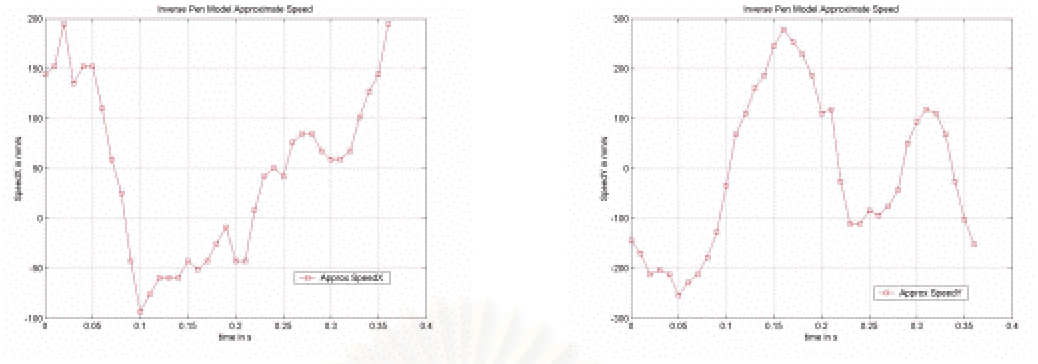
รูปที่ 4.14 ลายมือเขียนที่เกิดขึ้นจริงและตัวอักษรที่ได้จากการสังเคราะห์ซึ่งจะมีความใกล้เคียงกัน

รูปที่ 4.15 ภาพลายมือเขียนระดับสีเทาที่สังเคราะห์ขึ้นจากตัวแบบแหล่งกำเนิดภาพของปากกา

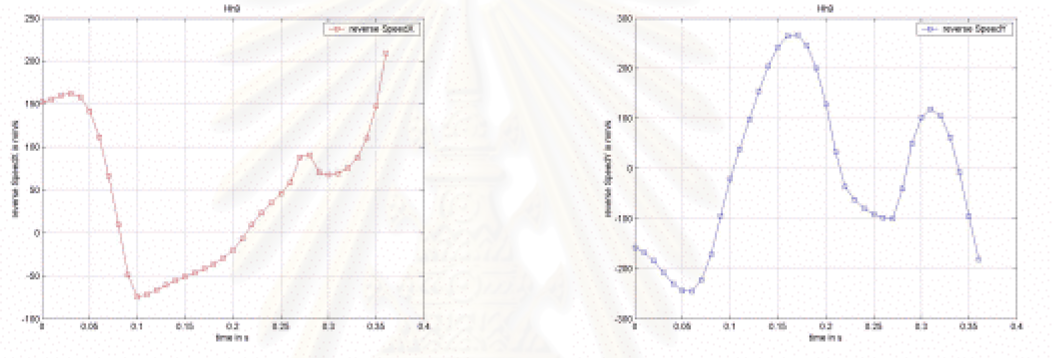
ภาพลายมือเขียนระดับสีเทาที่ได้จากการนำความเร็วการลากเส้นแนวนอนและแนวตั้งเป็นพารามิเตอร์ในการสร้างภาพให้กับตัวแบบแหล่งกำเนิดภาพของปากกา ในขั้นตอนนี้เรารู้พารามิเตอร์ทั้งหมดของภาพ

รูปที่ 4.16 ภาพโครงกระดูกลายมือเขียนที่หามาจากภาพข้างต้น

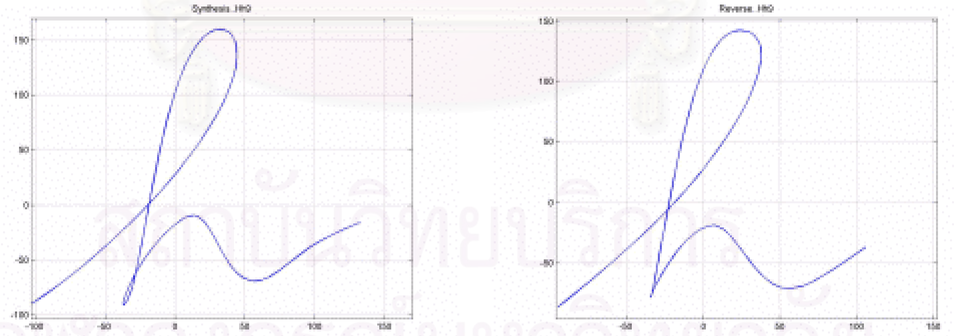
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 4.17 ความเร็วการลากเส้นที่ได้จากตัวแบบย้อนกลับการเกิดภาพของปากกา

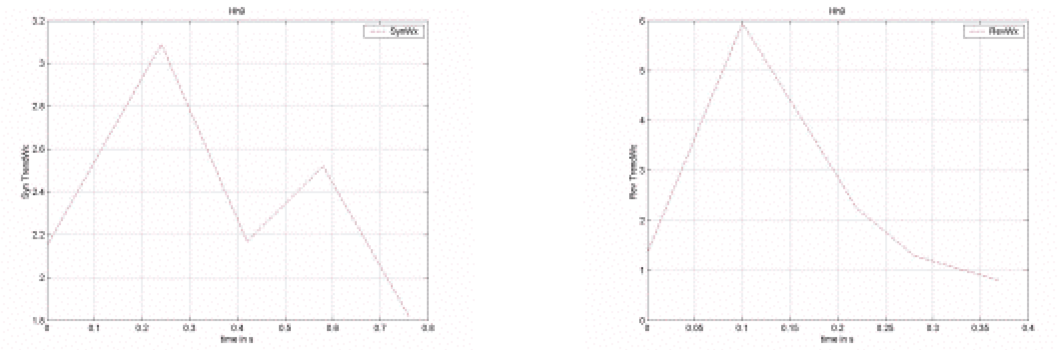


รูปที่ 4.18 ความเร็วการลากเส้นที่ได้จากการสังเคราะห์ด้วยพารามิเตอร์

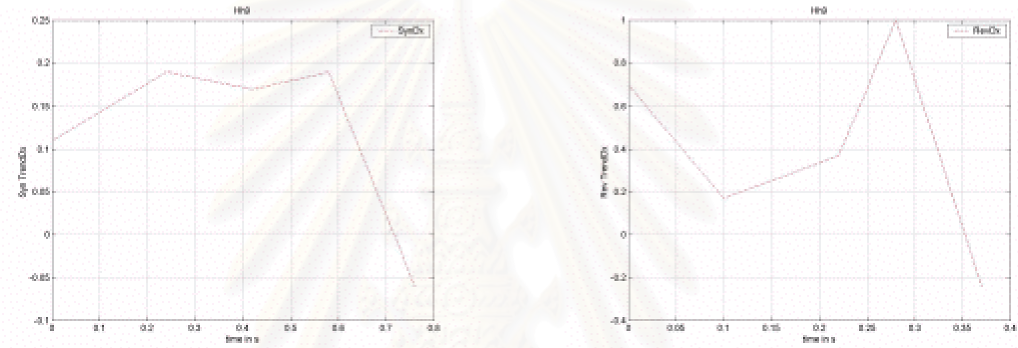


รูปที่ 4.19 ภาพลายมือเขียนที่เกิดจากพารามิเตอร์และภาพลายมือเขียนที่ได้จากดึงข้อมูลย้อนกลับจากภาพลายมือเขียนระดับสีเทา

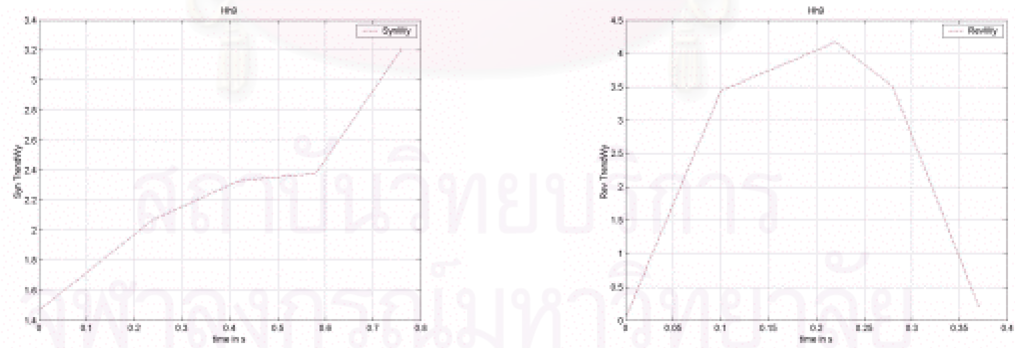
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



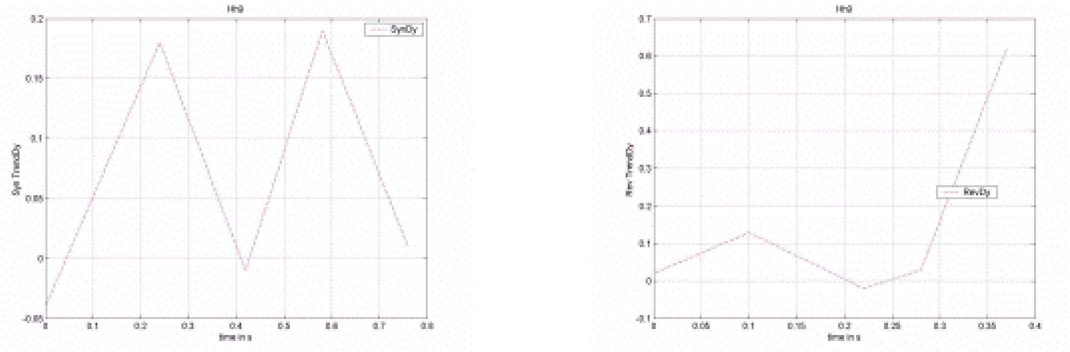
รูปที่ 4.20 แนวโน้มของความถี่ของความเร็วการลากเส้นแนวนอนระหว่างพารามิเตอร์ต้นแบบและพารามิเตอร์ที่ได้จากการหาย้อนกลับ



รูปที่ 4.21 แนวโน้มของค่าความหน่วงของความเร็วการลากเส้นแนวนอนระหว่างพารามิเตอร์ต้นแบบและพารามิเตอร์ที่ได้จากการหาย้อนกลับ



รูปที่ 4.22 แนวโน้มของความถี่ของความเร็วการลากเส้นแนวตั้งระหว่างพารามิเตอร์ต้นแบบและพารามิเตอร์ที่ได้จากการหาย้อนกลับ



รูปที่ 4.23 แนวโน้มของค่าความหน่วงของความเร็วการลากเส้นแนวตั้งระหว่างพารามิเตอร์ต้นแบบและพารามิเตอร์ที่ได้จากการหาย้อนกลับ

ผลการทดสอบพบว่าค่าความเร็วในแนวดิ่งและแนวนอนมีแนวโน้มที่หาจากภาพระดับสีเทาคล้ายต้นแบบที่สามารถนำไปรู้จำผู้เขียนได้ แต่ยังมีปริมาณข้อมูลมากอยู่ เมื่อดึงพารามิเตอร์ของการเขียนออกมาพบว่า เมื่อเปรียบเทียบกับแนวโน้มความถี่ต้นแบบ (พารามิเตอร์ต้นแบบ) พบว่าแนวโน้มความถี่ของความเร็วการลากเส้นทั้งสองแนว ไม่มีความสัมพันธ์กับแนวโน้มความถี่ซึ่งเป็นต้นแบบ เช่นเดียวกับแนวโน้มของค่าความหน่วง การนำพารามิเตอร์การเคลื่อนที่แบบ Damping Sinusoidal ที่หากลับมาได้จากภาพระดับสีเทาจึงไม่มีเสถียรภาพพอสำหรับการใช้งานโดยตรง จำเป็นต้องศึกษาวิธีการดึง feature อื่นที่มีเสถียรภาพต่อการหาข้อมูลจากภาพระดับสีเทาที่ดีกว่านี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5
บทสรุปและข้อเสนอแนะ

5.1 สรุปผลงานวิจัย

วิธีการดึงแนวการลากเส้นจากภาพลายมือเป็นขั้นตอนที่สำคัญของระบบรู้จำ ในปัจจุบันการดึงแนวการลากเส้นจากภาพลายมือเขียนโดยทั่วไปจะประมวลผลกับภาพขาวดำ ในบางครั้งภาพขาวดำที่ได้จากการปรับระดับความเข้มถูกทำให้ผิดเพี้ยนมาก การประมวลผลกับภาพขาวดำจะได้ผลที่ไม่ดี และวิธีดังกล่าวเป็นการมองข้ามการใช้ระดับสีเทา เพราะ ภาพลายมือเขียนระดับสีเทาเป็นภาพที่มีความคงอยู่ของข้อมูลมากที่สุด นอกจากนี้ระดับสีเทานั้นได้บ่งบอกถึง ข่าวสารข้อมูลการลากเส้น ดังนั้นการใช้ระดับสีเทาทำให้เราได้แนวการลากเส้นและข้อมูลการลากเส้น ในการดึงข้อมูลการลากเส้นจากภาพระดับสีเทา เราต้องทราบว่าระดับสีเทาที่เกิดบนภาพลายมือเขียนนั้นเกิดจากขึ้นได้อย่างไร ในงานวิจัยได้เสนอว่า ภาพลายมือเขียนเกิดตัวแบบแหล่งกำเนิด ที่ประกอบด้วย ตัวแบบแหล่งกำเนิดการเคลื่อนที่ของมือ และ ตัวแบบแหล่งกำเนิดภาพของปากการ่วมกัน

1. ตัวแบบแหล่งกำเนิดการเคลื่อนที่ของมือ ซึ่งเป็นคำตอบสมการอนุพันธ์อันดับสอง ทำหน้าที่สร้างตัวอักษร ขึ้นจากความเร็วแนวนอน และความเร็วแนวตั้ง ความเร็วทั้งสองจะสัมพันธ์กันและเป็นคุณลักษณะของตัวอักษรนั้นๆ ความเร็วในแต่ละแนวจะอยู่ในสมการ Damping Sinusoidal โดยมีความถี่,ความหน่วง,แอมปลิจูด,เฟส เป็นพารามิเตอร์ โดยพารามิเตอร์เหล่านี้เปลี่ยนแปลงอย่างเป็นเชิงเส้น ดังนั้นการลากเส้นใดๆสามารถอธิบายได้ด้วยพารามิเตอร์ชุดนี้ ตัวแบบแหล่งกำเนิดนี้ได้จำลองพฤติกรรมของมือ
2. ตัวแบบแหล่งกำเนิดภาพของปากกา ทำหน้าที่สร้างระดับสีเทาและความกว้างของเส้นที่ลากที่เปลี่ยนแปลงตามความเร็วการลากเส้น ตัวแบบแหล่งกำเนิดนี้สร้างจากปากกาที่อัตราการไหลของหมึกต่อเวลาจะคงที่(K) ดังนั้นเมื่อการลากเส้นเกิดขึ้นปากกาเคลื่อนที่ไปภายในเวลา t ปริมาณหมึกสุทธิที่ออกมาทั้งหมด Q จะเท่ากับผลคูณของอัตราการไหลกับเวลา (K*t) ดังนั้นการลากเส้นที่ความเร็วมากจะผลทำให้ปริมาณหมึกสุทธิออกมาน้อย เพราะใช้เวลาน้อย และตรงกันข้ามถ้าลากเส้นด้วยความเร็วน้อย และ จากความสัมพันธ์ข้างต้น เราได้ความสัมพันธ์อีกอันหนึ่งที่สอดคล้อง คือ ปริมาณหมึก(q) ซึ่งมีหน่วยเป็นปริมาตรต่อความยาวแปรผกผันกับกับความเร็วการลากเส้น เมื่อเรารู้ความเร็วการลากเส้น ทำให้เรารู้ค่าปริมาณหมึก(q) และเนื่องจากเราได้เสนอวิธีการคำนวณปริมาณหมึก (q) ด้วยการอินทิเกรตหน้าตัดความเข้มหรือ

เสมือนผลคูณของความกว้างกับระดับความเข้ม ดังนั้นปริมาณหมึก(q) ที่เปลี่ยนแปลงตามความเร็วการลากเส้นนั้น จะส่งผลระดับความเข้มและความกว้างเปลี่ยนแปลงตามกันด้วย งานวิจัยนี้ได้เสนอว่า ความกว้างจะถูกลดทอนลงด้วยค่ากำลังหนึ่งส่วนสามของอัตราส่วนของความเร็วการลากเส้นกับความเร็ว threshold1 ทำนองเดียวกับ ระดับความเข้ม จะถูกลดทอนด้วยค่ากำลังสองส่วนสามของอัตราส่วนความเร็วการลากเส้นกับความเร็ว threshold2 โดยเมื่อความเร็วการลากเส้นมากกว่าศูนย์จนกระทั่งเท่ากับความเร็ว threshold1 ความกว้างจะมีค่ามากที่สุด แต่เมื่อความเร็วการลากเส้นมากกว่านั้น ความกว้างของเส้น จะถูกลดทอนลงด้วยความสัมพันธ์ที่ได้กล่าวมาแล้ว เช่นเดียวกับ ถ้าความเร็วการลากเส้นมากกว่าศูนย์จนกระทั่งเท่ากับความเร็ว threshold2 ระดับความเข้มจะมีค่ามากที่สุด และ เมื่อความเร็วมากกว่านั้นระดับความเข้มจะถูกลดทอนลงด้วยความสัมพันธ์ที่ได้กล่าวมาแล้ว การกำหนดให้ความสัมพันธ์ระหว่างความกว้างและความเร็วการลากเส้น และความสัมพันธ์ระหว่างระดับความเข้มและความเร็วการลากเส้น ยังคงทำให้ความสัมพันธ์ปริมาณหมึกแปรผกผันตามความเร็วการลากเส้นเป็นจริงอยู่ เมื่อนำความสัมพันธ์ของระดับความเข้มและความกว้างของเส้นมาสร้างภาพ จะได้ภาพที่มีปริมาณหมึกสัมพันธ์กับความเร็วการลากเส้น

งานวิจัยได้เสนอกรรมวิธีการคำนวณปริมาณหมึก(q) ซึ่งมีหน่วยเป็นปริมาตรต่อหน่วยความยาว จากภาพระดับสีเทา ซึ่งนำใช้เป็นส่วนหนึ่งของการดึงข้อมูลการลากเส้นและเป็นการนำระดับสีเทามาใช้ประโยชน์ด้วย โดยกำหนดให้คำนวณปริมาณหมึกจากผลคูณของระยะทางกับระดับความเข้ม ซึ่งเสมือน การอินทิเกรตพื้นที่หน้าตัดระดับความเข้ม จากวิธีการดังกล่าวเราจะได้แผนที่ปริมาณหมึก(q) ซึ่งจะมีความใกล้เคียงกับแผนที่ระยะทางที่ได้จากการแปลงระยะทาง ปริมาณหมึก(q) ที่จะนำไปใช้ได้ คือ ตำแหน่งที่เป็นโครงกระดูกของแผนที่ปริมาณหมึก เพราะที่ตำแหน่งดังกล่าวเกิดจากความสมมาตรของปริมาณหมึก (q) ค่าที่ปริมาณหมึก(q) ที่อ่านได้จากตำแหน่งโครงกระดูกของภาพเมื่อนำคูณสองจะแสดงถึงปริมาณหมึกที่ทำให้เกิดแนวการลากเส้นดังกล่าวขึ้นมา ซึ่งปริมาณหมึกที่ตำแหน่งนี้จะสัมพันธ์กับความเร็วการลากเส้น

วิธีการทดสอบวิธีการดึงข้อมูลการลากเส้น ทำได้โดยการสร้างภาพลายมือเขียนขึ้นจากชุดพารามิเตอร์ โดยใช้ตัวแบบแหล่งกำเนิดการเคลื่อนที่ของมือและตัวแบบแหล่งกำเนิดภาพของปากกาว่มกัน ซึ่งภาพที่ได้จากสังเคราะห์มีความใกล้เคียงกับภาพจริง เมื่อได้ภาพสังเคราะห์ที่รู้ค่าพารามิเตอร์เป็นต้นแบบ ชุดพารามิเตอร์ต้นแบบจะถูกนำไปเปรียบเทียบกับพารามิเตอร์ที่ได้จากการดึงข้อมูลการเขียนจากภาพดังกล่าว

เมื่อนำภาพสังเคราะห์มาคำนวณปริมาณหมึก(q) แล้วใช้ตัวแบบย้อนกลับการเกิดภาพของปากกาในการหาความเร็วการลากเส้น ความเร็วที่หากลับคืนมามีความใกล้เคียงกับความเร็วการลากเส้นจริง และเมื่อนำความเร็วที่หากลับคืนได้มาเพื่อนำไปหาพารามิเตอร์ที่เกี่ยวข้องกับตัว

แบบการเคลื่อนที่ของมือ ซึ่งได้แก่ ความถี่ ความหน่วง พบว่าพารามิเตอร์ที่หาได้ไม่มีความสัมพันธ์กับพารามิเตอร์ต้นแบบ พารามิเตอร์ที่หากลับคืนมาไม่มีเสถียรภาพเพียงพอกับการใช้งาน จึงจำเป็นต้องศึกษาวิธีการนำไปใช้ต่อไป

5.2 การประยุกต์ใช้

1. การนำระดับสีเทาและความกว้างมาช่วยพิจารณาร่วมในการรู้จำลายเซ็น เนื่องจากส่วนมากในระบบรู้จำแบบ offline มักใช้โครงกระดูกเป็นหลัก ในการวิเคราะห์ลายเซ็น ซึ่งการปลอมลายเซ็นสามารถทำให้โครงกระดูกของลายเซ็นปลอมเหมือนจริงมากๆ การนำระดับสีเทาและความกว้างมาพิจารณาสามารถช่วยได้ คือ เนื่องจากผู้ที่ทำการปลอมลายเซ็นจะลากเส้นได้ช้ากว่าเจ้าของลายเซ็นตัวจริง ดังนั้นจะพบว่าที่ภาพลายเซ็นปลอมจะมีระดับสีเทาและความกว้างของแนวการลากเส้นมากกว่าภาพลายเซ็นจริง
2. การหาลำดับการลากเส้นที่เป็นไปตามจริง เนื่องจากการลากเส้นใดๆ ย่อมเกิดจากตัวแบบการเคลื่อนที่ของมือ ซึ่งมีพารามิเตอร์ทำหน้าที่กำกับรูปร่างอยู่ ในกรณีการลากเส้นใดๆ จะต้องสัมพันธ์กับความถนัดของมือ หรือ การเขียนโดยทั่วไปจะไม่ใช่พฤติกรรมที่ขัดกับธรรมชาติของมือ ดังนั้นการหาลำดับลากเส้น คือ การหาพารามิเตอร์ที่ทำให้ตัวแบบการเคลื่อนที่ของมือมีความเรียบง่ายที่สุด

5.3 ข้อเสนอแนะ

1. การศึกษาแนวทางในการนำพารามิเตอร์ของตัวแบบการเคลื่อนที่ของมือไปใช้ในงานรู้จำ เนื่องจากพารามิเตอร์ดังกล่าวยังไม่มีเสถียรภาพในการนำไปใช้งาน ณ ตอนนี้
2. การออกแบบตัวแบบแหล่งกำเนิดภาพของปากกาที่ถูกต้องที่สุด จะต้องมีอุปกรณ์ที่สามารถรับความเร็วของการเขียน และ ขณะเดียวกันอุปกรณ์นั้น จะต้องเขียนออกมาเป็นภาพข้อมูลดิบได้ด้วย เพื่อเราจะได้ตรวจสอบความเร็ว ณ ตอนนั้น สัมพันธ์กับระดับความเข้มได้อย่างไร และตัวแบบการเกิดภาพของปากกาที่สร้างขึ้นมาจะถูกต้องมากกว่า และ ตัวแบบย้อนกลับการเกิดภาพของปากกาที่ดีอาจจะทำให้หาความเร็วการลากเส้นได้ใกล้เคียงขึ้นอีกด้วย
3. เนื่องจากระดับความเข้มของภาพลายมือเขียนเกิดจากปัจจัยด้านอื่นๆที่ไม่ใช่ความเร็วการลากเส้นเพียงอย่างเดียว การที่เราพยายามสรุปว่าเกิดจากความเร็วนั้น แล้วนำตัวแบบย้อนกลับการเกิดภาพไปทดลองหาความเร็วการลากเส้นกับภาพจริงอาจจะเกิดความผิดพลาดได้ ภาพจริงนั้นมีปัจจัยด้านอื่นๆ เช่น แรงกด เป็นต้น

รายการอ้างอิง

1. Bruno Simard, Birenda Prasada and R Mahesh K Sinha. On-line Character Recognition Using Handwriting Modeling. Pattern Recognition ,Vol 26 (1993): 993-1007
2. Orly Stettiner and Dan Chazan. A Statistical Parameter Model for Recognition and Synthesis of Handwriting. Pattern Recognition , Conference-B ,Computer Vision and Image Processing. Proceedings of the 12th IAPR International. Vol.2 (1994): 34-38
3. H Chen ,O. E. Agazzi and C. Y. Suen. Piecewise Linear Modulation Model of Handwriting. [Online] Available from http://citeseer.nj.nec.com/404021.html
4. Shy-Shyan Chen and Frank Y Shih. Skeletonization For Fuzzy Degraded Character Image. IEEE Transcation on Image Processing ,Vol 5(October 1996):1481-1485
5. Seong-Whan Lee and Young Jooh Kim. Direct Extraction of Topographic Features for Gray Scale Character Recognition. IEEE Transaction on Pattern Analysis and Machine Intelligence. Vol 17(July 1995):724-729
6. Li Wang and Theo Pavlidis. Direct Gray-Scale Extraction of Features for Character Recogntion. IEEE Transaction on Pattern Analysis and Machine Intelligence. Vol 15 (October 1993):1053-1067
7. Y. Singer ,N. Tishby. Dynamical Encoding of Cursive Handwriting. Biological Cybernetics. (18 February 1994)
8. David Doermannn ,Venogopal Varma and Azriel Rosenfield. Instrument Grasp:A Model and Its Effects on Handwritting Strokes. Pattern Recognition. Vol 27(1994):233-245
9. G. Boccignone , A. Chianese ,L. P. Cordella and Marcelli. Recovering Dynamic Information from Static Handwriting. Pattern Recognition.Vol 26 (1993):
10. Gunilla Borgefors. Distance Transformations in Digital Images. Computer Vision,Graphics And Image Processing Vol 34 (1986):334-371

11. K. K. Lau,Pong C. Yuen and Y. Y. Tang EDT Based Tracing Maximum Thinging A Algorithm on Gray Scale Images. Pattern Recognition,2000.Proceeding 15th International Conference ,Vol 2 (2000):863-866

12. Hirobumi Nishida.A note on practical uses of gray-scale image analysis in document recogntion Pattern Recognition Letters. Vol. 19 (1998):889-897.

13. David Doermann .Document Image Understanding :Integrating Recovery and Interpretation.PhD. Thesis ,University of Maryland ,College Park,MD,1993

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นาย นวภัทร์ สุจิรชาโต เกิดเมื่อวันที่ 9 กุมภาพันธ์ พ.ศ. 2520 ที่จังหวัดกรุงเทพมหานครฯ สำเร็จการศึกษาระดับปริญญาตรีวิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมไฟฟ้าอิเลคทรอนิกส์ คณะวิศวกรรมไฟฟ้า สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารฯ ลาดกระบัง เมื่อปี พ.ศ. 2541

ปี พ.ศ. 2541 เข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย