

การสร้าง เครื่องอ่านตัวอักษร เพื่อใช้ในการจำแนกตัว เลขลายมือเขียน



นาย พิเชฐ อุดมประเสริฐ

วิทยานิพนธ์นี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

ภาควิชาวิศวกรรมไฟฟ้า

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

พ.ศ. 2529

ISBN 974 - 566 - 674 - 2

013397

IN667989X

CONSTRUCTION OF A CHARACTER READER FOR CLASSIFYING  
HANDWRITTEN NUMERALS

Mr. Pichet Udomprasert

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

Graduate School  
Chulalongkorn University

1986

หัวข้อวิทยานิพนธ์

การสร้างเครื่องอ่านตัวอักษรเพื่อใช้ในการจำแนกตัวเลขลายมือเขียน

โดย

นาย พิเชฐ อุดมประเสริฐ

ภาควิชา

วิศวกรรมไฟฟ้า

อาจารย์ที่ปรึกษา

รองศาสตราจารย์กฤษดา วิศวธีรานนท์



บัณฑิตวิทยาลัยจุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัยนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

.....

(รองศาสตราจารย์ ดร.สรชัย พิศาลบุตร)

รักษาการในตำแหน่งรองคณบดี ฝ่ายวิชาการ

ปฏิบัติราชการแทนรักษาการในตำแหน่งคณบดี

บัณฑิตวิทยาลัย

คณะกรรมการสอบวิทยานิพนธ์

.....

.....กรรมการ

(รองศาสตราจารย์ ดร.สมชาย จิตะพันธ์กุล)

.....กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.พิชิต โพธารามิก)

.....กรรมการ

(รองศาสตราจารย์ ดร.สุรียัน ดิษยาธิคม)

.....กรรมการ

(รองศาสตราจารย์กฤษดา วิศวธีรานนท์)

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

หัวข้อวิทยานิพนธ์      การสร้างเครื่องอ่านตัวอักษร เพื่อใช้ในการจำแนกตัวเลขลายมือ เขียน  
ชื่อนิลิต                    นาย พิเชฐ อุดมประเสริฐ  
อาจารย์ที่ปรึกษา        รองศาสตราจารย์กฤษดา วิศวกรรมนท์  
ภาควิชา                    วิศวกรรมไฟฟ้า  
ปีการศึกษา                2528



#### บทคัดย่อ

วิทยานิพนธ์นี้ กล่าวถึงการสร้างเครื่องอ่านตัวอักษร เพื่อใช้เป็นอุปกรณ์ในการศึกษา การจดจำรูปแบบ โดยเฉพาะอย่างยิ่งการจำแนกตัวเลขอารบิกลายมือเขียน

เครื่องอ่านตัวอักษรที่สร้างขึ้นประกอบด้วย กล้องวิดีโอคอน อินเตอร์เฟสบอร์ด ชุดไมโคร คอมพิวเตอร์ อินเตอร์เฟสบอร์ดทำหน้าที่แปลงสัญญาณวิดีโอจากกล้องให้เป็นสัญญาณภาพดิจิทัล 2 ระดับ ส่งเข้าเครื่องไมโครคอมพิวเตอร์ โดยสามารถอ่านได้ละเอียดดีกว่า 90 x 75 จุด ต่อตารางเซนติเมตร ได้พัฒนาซอฟต์แวร์สำหรับขบวนการจดจำตัวเลขอารบิกซึ่งประกอบด้วย การจัดการล่วงหน้า การนำเสนอรูปแบบ การจำแนกรูปแบบได้ทำการทดสอบการจำแนกตัวเลข ลายมือเขียน และตัวเลขตัวพิมพ์ดีด ผลปรากฏว่ามีความถูกต้อง ประมาณร้อยละ 90 สำหรับ ตัวเลขลายมือเขียนจากการสุ่มตัวอย่างลายมือเขียน 1,000 ตัว และร้อยละ 100 พอดี สำหรับ ตัวเลขตัวพิมพ์ดีดจากตัวอย่าง 9 แบบ





กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ สำเร็จได้ด้วยความช่วยเหลือจาก รองศาสตราจารย์กฤษดา  
วิเศษรัตน์ ที่ได้ให้คำแนะนำที่เป็นประโยชน์และเป็นอาจารย์ที่ปรึกษา อีกทั้งช่วยจัดหาตำรา  
และอุปกรณ์ตลอดจนค่าใช้จ่ายในการทดลองสร้าง เครื่องต้นแบบ จึงใคร่ขอกราบขอบพระ-  
คุณมา ณ ที่นี้ด้วย นอกจากนี้ที่ลืมมิได้เลย คือ บิดา มารดา ที่คอยให้กำลังใจและการสนับสนุน  
นอกจากนี้ ผู้เขียนขอขอบคุณทุกท่านที่ไม่ได้กล่าวในที่นี้อันมีส่วนช่วยในการพิมพ์และ  
จัดรูปเล่ม



สารบัญ

ช

หน้า

บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญตาราง.....	ญ
สารบัญรูปภาพ.....	ฎ

บทที่

1. บทนำ .....	1
1.1 ความเบื้องต้น .....	1
1.2 วัตถุประสงค์และเป้าหมายการวิจัย.....	2
1.3 ขั้นตอนการวิจัย.....	2
2. ทฤษฎีของระบบจดจำ.....	4
2.1 ทฤษฎีเบื้องต้น.....	4
2.2 การตัดสินใจเลือกวิธีจำแนกข้อมูล.....	6
2.3 วิธี Heuristic.....	7
2.3.1 Boolean functions.....	7
2.3.2 Zone features.....	10
2.3.3 Sequentially detected strip features..	11
2.3.4 Polygonal approximation.....	15
2.4 โครงสร้างของระบบจดจำตัวอักษร.....	16
2.4.1 ภาคเปลี่ยนข้อมูลเป็นสัญญาณไฟฟ้า.....	17
2.4.2 ภาคจัดการล่วงหน้า.....	17
2.4.3 ภาคนำเสนอรูปแบบ.....	23
2.4.4 ภาคจำแนกรูปแบบ.....	25

สารบัญ (ต่อ)

	หน้า
3. การออกแบบและสร้างเครื่องอ่านตัวอักษร.....	26
3.1 OCR คืออะไร .....	26
3.2 เทคนิคการอ่านข้อมูลด้วยแสง.....	28
3.2.1 แมคคาทรอนิกส์.....	28
3.2.2 ฟลายอิงสปอต.....	29
3.2.3 ไฟโตเซล.....	30
3.2.4 วิดีคอน.....	31
3.3 การพัฒนา OCR แบบใช้กล้องวีดิคอน.....	31
3.3.1 เหตุผลที่ใช้กล้องวีดิคอน.....	31
3.3.2 หลักการและแผนภูมิของระบบ .....	32
3.3.3 การออกแบบอินเตอร์เฟสบอร์ด.....	34
3.3.4 โปรแกรมการควบคุมวงจรรีดิเตอร์เฟสและแสดงผล....	40
3.4 การทดสอบ OCR ที่สร้างขึ้น.....	44
3.4.1 ระยะเวลาของตัวอักษร.....	44
3.4.2 ความละเอียดของข้อมูล.....	44
3.4.3 ความเป็นเชิงเส้น.....	45
3.4.4 การอ่านตัวอักษร.....	45
4. การออกแบบโปรแกรมระบบจดจำตัวเลขอารบิก .....	49
4.1 ลักษณะข้อมูลอินพุต.....	49
4.2 โครงสร้างของระบบ โปรแกรมจดจำตัวเลขอารบิก.....	51
4.3 MEMORY MAP ของระบบ โปรแกรม.....	54
4.4 โปรแกรมจัดการล่วงหน้า.....	55
4.4.1 โปรแกรมการหาขอบเขต.....	56
4.4.2 โปรแกรมกำจัดสัญญาณรบกวน.....	59



## สารบัญ (ต่อ)

	หน้า
4.3.3 โปรแกรมหาขอบตัวอักษร.....	61
4.5 โปรแกรมจดจำลักษณะ.....	64
4.5.1 โปรแกรมนำเสนอรูปแบบ .....	65
4.5.2 โปรแกรมจำแนกรูปแบบ.....	70
5. การทดลองระบบจดจำที่พัฒนาขึ้น.....	75
5.1 บทนำ.....	75
5.2 วิธีเก็บตัวอย่าง.....	75
5.3 ขั้นตอนการใช้โปรแกรม .....	78
5.3.1 การติดตั้งอินเตอร์เฟสบอร์ด.....	78
5.3.2 ศึกษาควบคุมการทำงาน.....	79
5.3.3 การป้อนข้อมูลภาพตัวเลข.....	80
5.3.4 การใช้โปรแกรมทำงานแบบ Batch file.....	80
5.4 ผลการทดลอง.....	80
5.5 การปรับปรุงแก้ไข.....	82
6. สรุปและข้อเสนอแนะ.....	83
6.1 สรุปผลการวิจัย.....	83
6.2 ข้อเสนอแนะในการพัฒนาต่อไป.....	83
6.2.1 จุดอ่อนของระบบจำแนกตัวอักษร.....	83
6.2.2 ข้อเสนอแนะ.....	84
เอกสารอ้างอิง.....	86
ภาคผนวก ก. รายละเอียดและคุณสมบัติของกล่องวิติคอนที่ใช้.....	91
ภาคผนวก ข. PROGRAM LISTING .....	96
ประวัติผู้เขียน .....	115

## สารบัญตาราง

ตารางที่		หน้า
1.1	ตารางเวลาของแต่ละขั้นตอนการวิจัย.....	3
2.1	เปรียบเทียบข้อดีข้อเสียวิธีต่าง ๆ ในกลุ่ม Heuristic approach...	16
6.1	แสดงผลการจำแนกตัวเลขลายมือเขียนจากการสุ่ม 100 คน.....	85

## สารบัญรูปภาพ

รูปที่	หน้า
2.1 แสดงบล็อกของระบบจดจำรูปแบบ.....	1
2.2 แสดงแผนภูมิแสดงวิธีในการจำแนกข้อมูล.....	5
2.3 แสดงตัวอย่างวิธีการจำแนกตัวอักษรของวิธี Heuristic.....	7
2.4 แสดงตำแหน่งของชิพดีวีซีเตอร์ของ IBM 1418 .....	9
2.5 แสดงรูปแบบของ stroke segment direction.....	12
2.6 แสดงการสแกนของหน้าต่าง 3 x 3 .....	12
2.7 แสดงรหัสแทนลักษณะทางแนวนอน .....	13
2.8 แสดงตัวอย่างรหัสที่ได้จากการสแกนตัวเลข '8' .....	13
2.9 แสดงตัวอย่างแผนภูมิทรานซิชันสำหรับเลข '8' .....	14
2.10 แสดงโครงสร้างระบบจดจำตัวอักษร.....	16
2.11 แสดงการโปรเจคลงบนแกน x และ y .....	18
2.12 แสดงผลการลดสัญญาณรบกวนโดยวิธี Dineen .....	19
2.13 แสดงผลจากขบวนการขยายและหด.....	19
2.14 แสดงสัญลักษณ์ของจุดข้างเคียงที่ใช้ใน Sobel's gradient...	21
2.15 แสดงการหาขอบโดยวิธีกวาดคอนทัวร์ .....	22
2.16 แสดงจุดข้างเคียงที่ได้รับการยกเว้นโดยวิธี Sherman .....	23
2.17 แสดงพารามิเตอร์ที่ใช้ในสมการ (2).....	24
3.1 แสดงลักษณะภายนอกของ CDC 915 page reader .....	27
3.2 แสดงโครงสร้างการอ่านข้อมูลแบบแมคคานิกส์.....	28
3.3 แสดงลำดับของการกวาดข้อมูล.....	29
3.4 แสดงโครงสร้างการอ่านข้อมูลแบบ ฟลายอิงสปอต .....	30
3.5 แสดงการอ่านข้อมูลแบบโฟโต เซล.....	31
3.6 แสดงแผนภูมิ OCR ที่สร้างขึ้น.....	32

## สารบัญรูปภาพ (ต่อ)

รูปที่	หน้า
3.7	แสดงระบบ OCR ที่สร้างขึ้น..... 33
3.8	แสดงวิธีในการสุ่มข้อมูลจากสัญญาณภาพ..... 34
3.9	แสดงแผนภูมิของส่วนอินเตอร์เฟส ..... 35
3.10	แสดงวงจรส่วนอินเตอร์เฟส..... 37
3.11	แสดงวงจรส่วนอินเตอร์เฟส (ต่อจากรูปที่ 3.10)..... 38
3.12	แสดงรูปบอร์ดอินเตอร์เฟสที่สร้างขึ้น..... 39
3.13	แสดงส่วนประกอบของโปรแกรมควบคุมวงจรอินเตอร์เฟสและ แสดงผล..... 40
3.14	แสดงไฟลว์ชาร์ทของโปรแกรมหลักในส่วนควบคุมวงจรอินเตอร์เฟส และแสดงผล..... 41
3.15	แสดงไฟลว์ชาร์ทการอ่านข้อมูลภาพ..... 42
3.16	แสดงไฟลว์ชาร์ทของโปรแกรมแสดงผล ..... 43
3.17	แสดงรูปที่ใช้ทดสอบความเป็นเชิงเส้น..... 45
3.18	แสดงผลการอ่านรูปที่ 3.17 ซึ่งพิมพ์ออกทางเครื่องพิมพ์..... 46
3.19	กราฟแสดงการทดสอบความเป็นเชิงเส้นแนวนอน..... 46
3.20	กราฟแสดงการทดสอบความเป็นเชิงเส้นในแนวตั้ง..... 47
3.21	แสดงภาพที่อ่านได้จากกล้องวิดีโอเป็นสัญญาณอนาล็อก..... 48
3.22	แสดงภาพที่หลังจากสุ่มข้อมูลเข้ามาเก็บเป็นสัญญาณดิจิทัล..... 48
4.1	แสดงโครงสร้างของโปรแกรมจดจำตัวเลขอารบิกที่เขียนขึ้น..... 50
4.2	แสดงตัวอย่างข้อมูลที่พิมพ์ในแต่ละขบวนการประมวลผล..... 53
4.3	แสดง MEMORY MAP ของระบบโปรแกรมจดจำตัวเลขอารบิกที่ ออกแบบ..... 54
4.4	แสดงส่วนประกอบของโปรแกรมจัดการลวงหน้า..... 56
4.5	แสดงไฟลว์ชาร์ทโปรแกรมการหาขอบเขต..... 56

## สารบัญรูปภาพ (ต่อ)

รูปที่		หน้า
4.6	แสดงการไปรเจคสร้างฮิสโตแกรมบนแกน $x$ และ $y$ .....	58
4.7	แสดงจุดที่มีความชัน เป็นศูนย์ที่เลือก.....	58
4.8	แสดงความผิดพลาดที่เกิด เนื่องจาก เขียนตัวเลขขีดกรอบ.....	59
4.9	แสดงตำแหน่งอ้างอิงหน้าค่า $3 \times 3$ .....	60
4.10	แสดงไฟล์ชาร์ทไปรแกรมกำจัดสัญญาณรบกวน.....	60
4.11	แสดงตัวอย่างผลการกำจัดสัญญาณรบกวนจากตัวเลขในรูปที่ 4.6.	61
4.12	แสดงไฟล์ชาร์ทในขบวนการ กวาดคอนทัวร์ .....	63
4.13	แสดงเงื่อนไขจุดขอยไม่เป็นเส้นปิด.....	64
4.14	แสดงโครงสร้างการทำงานไปรแกรมจดจำลักษณะ.....	64
4.15	แสดงไฟล์ชาร์ทการทำงานในไปรแกรม polygonal approximation.....	66
4.16	แสดงจุดอ่อนของขบวนการ split .....	66
4.17	แสดงผลขบวนการ polygonal approximation.....	67
4.18	แสดงการแบ่งพื้นที่ตัวเลขเป็นโซน.....	68
4.19	แสดงการอธิบายทิศทางของเชกเมนต์หรือเส้นแบ่งครึ่งมุม.....	68
4.20	แสดงนิยามเส้นแบ่งครึ่งส่วนเว้า.....	69
4.21	แสดง classification tree ที่สร้างขึ้น.....	71
4.22	แสดงสมาชิกใน class ของตัวเลข 0 ถึง 9 .....	74
5.1	แสดงแบบตัวพิมพ์เลขอารบิกที่ใช้ในการทดสอบ .....	76
5.2	แสดงแบบฟอร์มในการทดสอบการจำแนกตัวเลขลายมือเขียน....	77
5.3	แสดงโมดูลีความคุม.....	79
5.4	แสดงการแจกแจงเปอร์เซ็นต์ความผิดพลาดในแต่ละตัวเลขในการทดสอบครั้งแรก.....	81

## สารบัญรูปภาพ (ต่อ)

รูปที่		หน้า
5.5	แสดงรูปแบบที่จำแนกผิดในแต่ละ class ของตัวเลข.....	82
5.6	แสดงการแจกแจง เปอร์เซนต์ความผิดพลาดหลังจากได้แก้ไข ปรับปรุง.....	82