การรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้เทคนิคด้านการวิเคราะห์ตัวประกอบสำคัญและนิวรอลเน็ตเวิร์ก

นาย ธเนศ ศรีวิรุฬห์ชัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2541

ISBN 974-331-513-6

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

I 19236761

THAI PRINTED CHARACTERS RECOGNITION USING PRINCIPAL COMPONENT ANALYSIS AND NEURAL NETWORKS

Mr. Tanes Sriviroolchai

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science in Computer Science

Department of Computer Engineering

Graduate School

Chulalongkorn University

Academic Year 1998

ISBN 974-331-513-6

หัวข้อวิทยานิพนธ์ การรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้เทคนิคด้านการวิเคราะห์ตัวประกอบสำคัญและนิวรอลเน็ตเวิร์ก

โดย นาย ธเนศ ศรีวิรุฬห์ชัย

ภาควิชา วิศวกรรมคอมพิวเตอร์

อาจารย์ที่ปรึกษา อ. ดร. บุญเสริม กิจศิริกุล

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

..... คณบดีบัณฑิตวิทยาลัย
(ศาสตราจารย์ นพ. ศุภวัฒน์ ชุติวงศ์)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(อาจารย์ ดร. ยรรยง เต็งอำนวย)

..... อาจารย์ที่ปรึกษา
(อาจารย์ ดร. บุญเสริม กิจศิริกุล)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ นงลักษณ์ โควาวิสารัช)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. สมชาย ประสิทธิ์จูตระกูล)

พิมพ์ต้นฉบับบทคัดย่อวิทยานิพนธ์ภายในกรอบสีเขียวนี้เพียงแผ่นเดียว

ธเนศ ศรีวิรุฬห์ชัย: การรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้เทคนิคด้านการวิเคราะห์ตัวประกอบสำคัญและนิวรอลเน็ตเวิร์ก (THAI PRINTED CHARACTERS RECOGNITION USING PRINCIPAL COMPONENT ANALYSIS AND NEURAL NETWORKS) อ. ที่ปรึกษา : ดร. บุญเสริม กิจศิริกุล, 66 หน้า. ISBN 974-331-513-6

วิทยานิพนธ์นี้มีวัตถุประสงค์เพื่อประยุกต์ใช้เทคนิคด้านการวิเคราะห์ตัวประกอบสำคัญและนิวรอลเน็ตเวิร์กในการรู้จำตัวอักษรพิมพ์ภาษาไทย และพัฒนาโปรแกรมสำหรับรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้เทคนิคด้านการวิเคราะห์ตัวประกอบสำคัญและนิวรอลเน็ตเวิร์ก

การรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้เทคนิคด้านการวิเคราะห์ตัวประกอบสำคัญและนิวรอลเน็ตเวิร์กมีขบวนการทำงานคือ ทำการอ่านภาพตัวอักษรภาษาไทยหนึ่งภาพต่อหนึ่งตัวอักษร ลักษณะภาพแบบขาวดำ แล้วทำการเปลี่ยนขนาดภาพให้อยู่ในขนาด 32 x 32 จุด ทำการแปลงแบบเค-แอลของเมตริกซ์ของจุดภาพและทำการวิเคราะห์เมตริกซ์รูปแบบที่ได้จากการแปลงด้วยนิวรอลเน็ตเวิร์กแบบแบคพรอพาเกชัน

ในการวิจัยได้นำข้อมูลภาพตัวอักษรที่ได้จากการพิมพ์ด้วยเครื่องพิมพ์เลเซอร์ที่ความละเอียด 600 จุดต่อนิ้ว นำเอกสารมาอ่านผ่านเครื่องสแกนเนอร์ความละเอียด 200 จุดต่อนิ้วจำนวน 3264 ตัวอักษรซึ่งประกอบด้วยตัวอักษรแบบ AngsanaUPC, BrowaliaUPC, CordiaUPC, DilleniaUPC, EucrosiaUPC และ FreesiaUPC แต่ละแบบประกอบด้วยตัวอักษรขนาด 16, 18, 20, 22, 24, 26, 28 และ 36 จุด เป็นตัวอักษรต้นแบบ และทำการทดสอบด้วยข้อมูลภาพตัวอักษรที่ได้จากการพิมพ์ด้วยเครื่องพิมพ์เลเซอร์ที่ความละเอียด 600 จุดต่อนิ้ว นำเอกสารที่ได้จากเครื่องพิมพ์มาถ่ายเอกสารให้จางลง และนำเอกสารที่ได้จากเครื่องพิมพ์มาถ่ายเอกสารให้เข้มขึ้น แล้วนำเอกสารทั้ง 2 ฉบับมาอ่านผ่านเครื่องสแกนเนอร์ที่ความละเอียด 200 จุดต่อนิ้วจำนวน 6528 ตัวอักษรซึ่งประกอบด้วยตัวอักษรแบบ AngsanaUPC, BrowaliaUPC, CordiaUPC, DilleniaUPC, EucrosiaUPC และ FreesiaUPC แต่ละแบบประกอบด้วยตัวอักษรขนาด 16, 18, 20, 22, 24, 26, 28 และ 36 จุด ได้ผลการรู้จำมีความถูกต้องเฉลี่ยร้อยละ 96.84

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต .....

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่ออาจารย์ที่ปรึกษา .....

ปีการศึกษา 2541 ลายมือชื่ออาจารย์ที่ปรึกษาร่วม .....

3970672921 : MAJOR COMPUTER SCIENCE
KEYWORD: PRINCIPAL COMPONENT ANALYSIS / NEURAL NETWORKS / BACKPROPAGATION / CHARACTER RECOGNITION

TANES SRIVIROOLCHAI : THAI PRINTED CHARACTERS RECOGNITION USING PRINCIPAL COMPONENT ANALYSIS AND NEURAL NETWORKS. THESIS ADVISOR : BOONSERM KIJSIRIKUL, Ph.D. 66 pp. ISBN 974-331-513-6.

The objective of this thesis is to apply principal component analysis and neural networks to Thai printed characters recognition, and develop a program for this task.

The process of recognizing Thai printed characters using principal component analysis and neural networks are as follows. First, read bi-level character image at one picture per character. Second, for each character image normalize its size into 32x32 dots. Next, transform the normalized image by K-L transform. Finally, classify transformed matrix by backpropagation neural networks.

In this research, training prototype characters were generated from the outputs from laser printer at 600 dots per inch. The originals composed of 3264 characters in 6 fonts: AngsanaUPC, BrowaliaUPC, CordiaUPC, DilleniaUPC, EucrosiaUPC, and FreesiaUPC, each of which is composed of size 16, 18, 20, 22, 24, 26, 28, and 36 points. The originals were scanned at 200 dots per inch and used as the training prototype characters. The program was tested with two sets of samples. The first sample set was generated by photocopying the original printout by a photocopy machine with lightened mode and the second set was generated with darkened mode. The total number of samples is 6528 characters. The samples were then scanned at 200 dots per inch and tested with the program. The recognition rate is about 96.84 %.

ภาควิชา วิศวกรรมคอมพิวเตอร์	**ลายมือชื่อนิสิต** .....
สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์	**ลายมือชื่ออาจารย์ที่ปรึกษา** .....
ปีการศึกษา 2541	**ลายมือชื่ออาจารย์ที่ปรึกษาร่วม** .....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ได้สำเร็จลุล่วงไปได้โดยความช่วยเหลือของ อาจารย์ ดร.บุญเสริม กิจศิริกุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้แนวทาง ข้อเสนอแนะ มาตลอด ขอขอบคุณ คุณชาลี เจริญลาภนพรัตน์, คุณบัญชา คิ้วคมขำ, อาจารย์จักรพันธ์ แซ่ลี่ ที่เป็นธุระช่วยหาเอกสารอ้างอิงบางฉบับ ขอขอบคุณ คุณธเนศ วัฒนโกสิน ที่ให้ความอนุเคราะห์ให้ใช้เครื่องสแกนเนอร์ในการเก็บข้อมูลภาพตัวอักษร ตลอดจนเพื่อน ๆ และ น้อง ๆ ที่คอยเป็นกำลังใจมาตลอด

ท้ายที่สุดนี้ ผู้วิจัยใคร่กราบขอบพระคุณ บิดา มารดา ซึ่งให้การสนับสนุนและเป็นกำลังใจให้แก่ผู้วิจัยจนสำเร็จการศึกษา

สารบัญ

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญตาราง ..... ฌ
สารบัญภาพ ..... ญ
บทที่ 1 บทนำ 1
ความเป็นมาและความสำคัญของปัญหา 1
งานวิจัยที่เกี่ยวข้อง 2
วัตถุประสงค์ของการวิจัย 3
ขอบเขตของการวิจัย 4
ขั้นตอนและวิธีการดำเนินการวิจัย 5
ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย 5
บทที่ 2 แนวคิดและทฤษฎี 6
ผังการทำงาน 6
ข้อมูลภาพ 9
การหาเมตริกซ์ค่าเฉลี่ยและเมตริกซ์ของไอเกนเวกเตอร์ (EIGENVECTOR) 9
การแปลงแบบเค-แอล (K-L TRANSFORM) 10
การสอนนิวรอลเน็ตเวิร์ก 11
การป้อนเวกเตอร์รูปแบบที่ต้องการสอนไปข้างหน้า 12
การคำนวนและกระจายป้อนกลับของค่าความผิดพลาดของผลลัพธ์ 13
การปรับค่าน้ำหนัก 14
การใช้นิวรอลเน็ตเวิร์กทำการรู้จำ 15
บทที่ 3 การออกแบบและพัฒนา 16
การนำเข้าข้อมูลภาพ 16
การปรับกรอบภาพและการปรับขนาดภาพ 17
การปรับกรอบภาพ 17
การเปลี่ยนขนาดภาพ 18

การหาคุณลักษณะของภาพของตัวอักษร (FEATURE EXTRACTION) 20
การตัดสินใจแยกแยะ (CLASSIFICATION) 22
บทที่ 4 ผลการวิจัย 27
แหล่งที่มาของข้อมูลภาพ 27
วิธีการทดสอบ 31
ผลการทดสอบ 36
ผลการเปรียบเทียบ 42
วิเคราะห์ผลการวิจัย 42
ปัญหาและข้อจำกัด 43
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ 45
สรุปผลการวิจัย 45
ข้อเสนอแนะ 45
รายการอ้างอิง 47
ภาคผนวก ก. การใช้งานโปรแกรมรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้เทคนิคด้านการวิเคราะห์ตัวประกอบสำคัญและนิวรอลเน็ตเวิร์ก 49
ภาคผนวก ข. ตัวอักษรต้นแบบแบบ AngsanaUPC 54
ภาคผนวก ค. ตัวอักษรต้นแบบแบบ BrowaliaUPC 56
ภาคผนวก ง. ตัวอักษรต้นแบบแบบ CordiaUPC 58
ภาคผนวก จ. ตัวอักษรต้นแบบแบบ DilleniaUPC 60
ภาคผนวก ฉ. ตัวอักษรต้นแบบแบบ EucrosiaUPC 62
ภาคผนวก ช. ตัวอักษรต้นแบบแบบ FreesiaUPC 64
ประวัติผู้เขียน 66

สารบัญตาราง

ตารางที่ 3.1 เปรียบเทียบผลของการรู้จำเมื่อนิวรอลเน็ตเวิร์กมีจำนวนบัพในชั้นที่ถูกซ่อน 73 บัพ 100 บัพ 128 บัพ และ 200 บัพ 23
ตารางที่ 4.1 ผลการทดสอบการรู้จำแยกตามรูปแบบตัวอักษร 36
ตารางที่ 4.2 ผลการทดสอบการรู้จำแยกตามตัวอักษรที่ทดสอบ 38
ตารางที่ 4.3 เปรียบเทียบเทคนิคที่ใช้ในขั้นตอนต่าง ๆ ของวิทยานิพนธ์ที่นำมาเปรียบเทียบ 41
ตารางที่ 4.4 ผลการเปรียบเทียบการรู้จำของวิทยานิพนธ์ฉบับนี้เทียบกับวิทยานิพนธ์ฉบับอื่น ๆ .. 42

สารบัญภาพ

รูปที่ 2.1 ผังการทำงานของกระบวนการเรียนรู้ 7
รูปที่ 2.2 ผังการทำงานของกระบวนการรู้จำ 8
รูปที่ 2.3 ลักษณะการแทนจุดภาพด้วยเวกเตอร์ 9
รูปที่ 2.4 นิวรอลเน็ตเวิร์กแบบแบคพรอพาเกชันที่มีชั้นที่ถูกซ่อน 1 ชั้น 11
รูปที่ 2.5 ซิกมอยด์ฟังก์ชัน ที่มีผลลัพธ์ในช่วง (0,1) 12
รูปที่ 3.1 ผังการทำงานของการรู้จำตัวอักษร 16
รูปที่ 3.2 กรอบจุดภาพของตัวอักษร 17
รูปที่ 3.3 การหาจุดดำจุดแรกด้านซ้าย 17
รูปที่ 3.4 การหาจุดดำจุดแรกด้านขวา 17
รูปที่ 3.5 การหาจุดดำจุดแรกด้านบน 18
รูปที่ 3.6 การหาจุดดำจุดแรกด้านล่าง 18
รูปที่ 3.7 กรอบภาพใหม่ของตัวอักษรซึ่งมีจุดขาวล้อม 1 จุด 18
รูปที่ 3.8 การเพิ่มจุดขาวในด้านซ้ายและด้านขวาเพื่อให้ภาพมีขนาด 32x32 20
รูปที่ 3.9 การเพิ่มจุดขาวในด้านบนและด้านล่างเพื่อให้ภาพมีขนาด 32x32 20
รูปที่ 3.10 ค่าของไอเกน (eigenvalue) 128 ตัวแรกที่ได้จากการเรียนรู้ตัวอักษรต้นแบบ 21
รูปที่ 3.11 ตัวอย่างเวกเตอร์รูปแบบที่ได้จากการแปลงเค-แอลของตัวอักษร "อ" แบบ AngsanaUPC ขนาด 18 จุด 22
รูปที่ 3.12 ผลลัพธ์ที่ได้จากนิวรอลเน็ตเวิร์กเมื่อให้ทำการรู้จำเวกเตอร์รูปแบบที่ได้จากการแปลงเค-แอลของตัวอักษร "อ" แบบ AngsanaUPC ขนาด 18 จุด 26
รูปที่ 4.1 ตัวอย่างข้อมูลภาพชุดที่ 1 ที่ได้จากการอ่านเอกสารที่พิมพ์ด้วยเครื่องพิมพ์เลเซอร์ 28
รูปที่ 4.2 ตัวอย่างข้อมูลภาพชุดที่ 2 ที่ได้จากจากการอ่านเอกสารที่ได้จากการถ่ายเอกสารแบบจาง 29
รูปที่ 4.3 ตัวอย่างข้อมูลภาพชุดที่ 3 ที่ได้จากการอ่านเอกสารที่ได้จากการถ่ายเอกสารแบบเข้ม ... 30
รูปที่ 4.4 ตัวอย่างผลลัพธ์จากนิวรอลเน็ตเวิร์กที่มีบัพใดบัพหนึ่งเท่านั้นที่มีค่าสูง (มากกว่า 0.5) ... 32
รูปที่ 4.5 ตัวอย่างผลลัพธ์จากนิวรอลเน็ตเวิร์กที่ไม่มีบัพใดบัพหนึ่งเป็นค่าสูง (มากกว่า 0.5) และใช้บัพที่มีค่าสูงสุดแทน 33
รูปที่ 4.6 ตัวอย่างผลลัพธ์จากนิวรอลเน็ตเวิร์กที่มีบัพที่มีค่าสูง (มากกว่า 0.5) เพียงบัพเดียวแต่รู้จำผิด 34
รูปที่ 4.7 ผลลัพธ์จากนิวรอลเน็ตเวิร์กซึ่งมีบัพที่มีค่าสูง (มากกว่า 0.5) มากกว่า 1 บัพ 35