การรู้จำเสียงคำไทยหลายพยางค์แบบไม่ขึ้นกับผู้พูด โดยใช้เทคนิคแบบฟัซซีและนิวรอลเน็ตเวิร์ก

นายชัย วุฒิวิวัฒน์ชัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2540

ISBN 974-638-434-1

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

I17855078

SPEAKER INDEPENDENT THAI POLYSYLLABIC WORD RECOGNITION USING FUZZY-TECHNIQUE AND NEURAL NETWORK

Mr. Chai Wutiwiwatchai

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
Graduate School
Chulalongkorn University
Academic Year 1997

ISBN 974-638-434-1

หัวข้อวิทยานิพนธ์	การรู้จำเสียงคำไทยหลายพยางค์แบบไม่ขึ้นต่อผู้พูด โดยใช้เทคนิคแบบฟัซซี และนิวรอลเน็ตเวอร์ค
โดย	นาย ชัย วุฒิวิวัฒน์ชัย
ภาควิชา	วิศวกรรมไฟฟ้า
อาจารย์ที่ปรึกษา	รองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล
อาจารย์ที่ปรึกษาร่วม	ผู้ช่วยศาสตราจารย์ ดร. สุดาพร ลักษณียนาวิน

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

..... คณบดีบัณฑิตวิทยาลัย
(ศาสตราจารย์ นายแพทย์ ศุภวัฒน์ ชุติวงศ์)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(รองศาสตราจารย์ ดร. เทียนชัย ประดิสถายน)

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

..... อาจารย์ที่ปรึกษาร่วม
(ผู้ช่วยศาสตราจารย์ ดร. สุดาพร ลักษณียนาวิน)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. วาทิต เบญจพลกุล)

..... กรรมการ
(ดร. จุฬารัตน์ ตันประเสริฐ)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

พิมพ์ต้นฉบับบทคัดย่อวิทยานิพนธ์ภายในกรอบสีเขียวนี้เพียงแผ่นเดียว

ชัย วุฒิวิวัฒน์ชัย : การรู้จำเสียงพูดคำไทยหลายพยางค์แบบไม่ขึ้นกับผู้พูด โดยใช้เทคนิคแบบฟัซซี และนิวรอลเน็ตเวิร์ก (SPEAKER INDEPENDENT THAI POLYSYLLABIC WORD RECOGNITION USING FUZZY-TECHNIQUE AND NEURAL NETWORK) อาจารย์ที่ปรึกษา : รศ. ดร. สมชาย จิตะพันธ์กุล, อาจารย์ที่ปรึกษาร่วม : ผศ. ดร. สุดาพร ลักษณียนาวิน, 95 หน้า. ISBN 974-638-434-1

วิทยานิพนธ์นี้ มีจุดมุ่งหมายเพื่อพัฒนาวิธีการรู้จำคำไทยหลายพยางค์แบบไม่ขึ้นต่อผู้พูด โดยใช้นิวรอลเน็ตเวอร์ก และใช้เทคนิคแบบฟัซซีในการปรับปรุงข้อมูลที่ใช้ในการฝึกฝน โดยแทนที่จะใช้ข้อมูลเข้าเป็นค่าสมาชิกภาพแบบฟัซซี ร่วมกับข้อมูลออกที่ต้องการแบบค่าสมาชิกภาพของแต่ละคำศัพท์ในการฝึกฝน ดังที่เคยใช้ในงานวิจัยหลายๆ งานที่ผ่านมา จะใช้ข้อมูลเข้าเป็นค่าสมาชิกภาพแบบฟัซซี ร่วมกับข้อมูลออกที่ต้องการแบบเลขฐานสอง ชุดคำศัพท์ประกอบด้วยคำศัพท์ตัวเลข 0-9 ชุดคำศัพท์หนึ่งพยางค์อื่นๆ ที่ไม่ใช่ตัวเลข 20 คำ ชุดคำศัพท์สองพยางค์ 20 คำ และชุดคำศัพท์สามพยางค์ 20 คำ โดยใช้วิธีการตรวจสอบจำนวนพยางค์ และการตรวจสอบเสียงวรรณยุกต์ ในการแบ่งกลุ่มคำศัพท์เบื้องต้น เพื่อลดจำนวนคำศัพท์ที่นิวรอลเน็ตเวอร์กหนึ่งๆ จะต้องรู้จำ ผู้พูดในชุดฝึกฝนมีจำนวน 50 คน และชุดทดสอบแบบไม่ขึ้นต่อผู้พูด 10 คน ผลการทดสอบได้อัตราการรู้จำแบบขึ้นต่อผู้พูด และแบบไม่ขึ้นต่อผู้พูดเฉลี่ย 94.4 และ 93.3 เปอร์เซ็นต์ ตามลำดับ ซึ่งสูงกว่ากรณีที่ใช้ข้อมูลเข้าเป็นค่าสัมประสิทธิ์การประมาณพันธะเชิงเส้น 3.3 และ 3.4 เปอร์เซ็นต์ ตามลำดับ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา วิศวกรรมไฟฟ้า ลายมือชื่อนิสิต .....

สาขาวิชา วิศวกรรมไฟฟ้า ลายมือชื่ออาจารย์ที่ปรึกษา .....

ปีการศึกษา 2540 ลายมือชื่ออาจารย์ที่ปรึกษาร่วม .....

C819152 : MAJOR ELECTRICAL ENGINEERING
KEY WORD:
SPEECH RECOGNITION / THAI POLYSYLLABIC WORD / FUZZY MEMBERSHIP FUNCTION / NEURAL NETWORK

CHAI WUTIWIWATCHAI : SPEAKER INDEPENDENT THAI POLYSYLLABIC WORD RECOGNITION USING FUZZY-TECHNIQUE AND NEURAL NETWORK. THESIS ADVISOR : ASSO. PROF. SOMCHAI JITAPUNKUL, Ph.D. THESIS COADVISOR : ASSIST. PROF. SUDAPORN LUKSANEEYANAVIN, Ph.D. 95 pp. ISBN 974-638-434-1

The objective of this thesis is to develop algorithms for a speaker independent Thai polysyllabic word recognition system using Neural Network with improvement training data by Fuzzy technique. Instead of using Fuzzy membership input data and class membership desired-output data during training as seen in several works, we use Fuzzy membership input data and binary desired-output. Vocabulary set contains numeral figures 0-9, including other 20 single-syllabic words, 20 double-syllabic words and 20 triple-syllabic words. The syllable detection and tone detection algorithms are used for vocabulary pre-classification in order to decrease the number of vocabularies to be feeding to the Neural Network. With 50 training subjects and 10 independent test subjects, the average recognition rates of speaker dependent and speaker independent test are 94.4 and 93.3% respectively. There is an increase of 3.3 and 3.4% respectively compared to the recognition using only LPC input data.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา วิศวกรรมไฟฟ้า ลายมือชื่อนิสิต ชัย วุฒิวิวัฒน์ชัย

สาขาวิชา วิศวกรรมไฟฟ้า ลายมือชื่ออาจารย์ที่ปรึกษา Somchai Jitapunkul

ปีการศึกษา 2540 ลายมือชื่ออาจารย์ที่ปรึกษาร่วม Sudaporn Luksaneeyanavin

กิตติกรรมประกาศ

ข้าพเจ้าขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.สมชาย จิตะพันธ์กุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ และผู้ช่วยศาสตราจารย์ ดร.สุดาพร ลักษณียนาวิน อาจารย์ที่ปรึกษาร่วม ที่เปิดโอกาสให้ข้าพเจ้าเข้ามารับความรู้ และประสบการณ์ชีวิตที่มิอาจลืมเลือนได้ ทั้งยังเสียสละเวลาให้คำปรึกษาชี้แนะแนวทางเสมอมา ข้าพเจ้าขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.เทียนชัย ประดิสถายน ผู้ช่วยศาสตราจารย์ ดร.วาทิต เบญจพลกุล และ ดร.จุฬารัตน์ ตันประเสริฐ ที่ได้ให้คำแนะนำ และคำวิจารณ์ที่เป็นประโยชน์

วิทยานิพนธ์นี้จะไม่สามารถสำเร็จได้ถ้าไม่ได้รับคำปรึกษา คำแนะนำ และการช่วยเหลือจาก นายวิศรุต อาชุบุตร นายวุฒิพงษ์ พรสุขจันทรา นายเอกฤทธิ์ มณีน้อย และนางสาวณัฏฐชา จิตติวรางกูล ข้าพเจ้าขอขอบพระคุณพี่ๆ เพื่อนๆ และน้องๆ ทุกท่านที่มีส่วนช่วยให้คำแนะนำ เป็นกำลังใจ และสร้างบรรยากาศการทำงานให้ครึกครื้นอยู่เสมอ

ข้าพเจ้าขอขอบพระคุณพี่ชายของข้าพเจ้าที่คอยรับฟังปัญหา ให้กำลังใจ และคอยให้คำปรึกษาในฐานะวิศวกรด้วยกัน สุดท้ายนี้ข้าพเจ้าขอกราบขอบพระคุณคุณพ่อคุณแม่ โดยเฉพาะคุณแม่ที่เปิดโอกาสให้ข้าพเจ้าได้เรียนตามที่ข้าพเจ้าต้องการ ทั้งยังคอยเอาใจใส่ ช่วยเหลือ และไม่ต่อว่าใดๆ แม้ว่าข้าพเจ้าจะไม่มีเวลาให้ท่านเท่าที่ควร และแม้ว่าจะมีปัญหาทางเศรษฐกิจก็ตาม

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญตาราง ..... ฌ
สารบัญรูปประกอบ ..... ญ
คำอธิบายศัพท์ ..... ฐ
บทที่ 1 บทนำ 1
แนวเหตุผล 1
วัตถุประสงค์ 3
เป้าหมาย และขอบเขต 3
ประโยชน์ที่คาดว่าจะได้รับ 3
บทที่ 2 หลักการ และทฤษฎีสำคัญ 4
2.1 การรู้จำเสียงพูด 4
2.2 นิวรอลเน็ตเวอร์ก 7
2.3 ทฤษฎีเกี่ยวกับฟัซซี 11
2.4 การใช้ทฤษฎีเกี่ยวกับฟัซซีร่วมกับนิวรอลเน็ตเวอร์กในการรู้จำ 16
2.5 ทฤษฎีอื่นๆ ที่สำคัญ 19
บทที่ 3 กระบวนการรู้จำที่เสนอ 28
ช่วงฝึกฝน 29
3.1 การประมวลผลเบื้องต้น 29
3.2 การสกัดค่าลักษณะเด่น 31
3.3 การสกัดค่าลักษณะเด่นแบบฟัซซี 33
3.4 เวกเตอร์ข้อมูลออกที่ต้องการ 35
3.5 นิวรอลเน็ตเวอร์กแบบ MLP 36
ช่วงทดสอบการรู้จำ 38
3.6 การแบ่งกลุ่มเบื้องต้น 38

หน้า

3.7 กฎเกณฑ์การตัดสินใจ....40
บทที่ 4 การทดลอง และการวิเคราะห์ผลการทดลอง....41
4.1 ข้อมูล และการเก็บข้อมูล....41
4.2 การทดลองเพื่อวิเคราะห์หาค่าพารามิเตอร์ที่เหมาะสม....41
4.3 การวิเคราะห์ค่าข้อมูลออกที่ต้องการในกรณีกำกวมที่สุด (Fuzziest Case)....51
4.4 ผลการทดลองการรู้จำ และการวิเคราะห์....52
บทที่ 5 สรุปผลการวิจัย และข้อเสนอแนะ....72
5.1 สรุปผลการวิจัย....72
5.2 ปัญหา และข้อเสนอแนะ....73
รายการอ้างอิง....75
ภาคผนวก ก รายการชุดคำศัพท์ที่ใช้ และการแยกกลุ่มคำศัพท์....79
ภาคผนวก ข แผนภาพขั้นตอนกระบวนการที่สำคัญ....85
ภาคผนวก ค ผลการทดลองเพื่อหาระยะบวกเผื่อที่ต้น และท้ายคำ....91
ประวัติผู้เขียน....95

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

หน้า

ตารางที่ 2.1 การแบ่งแยกชนิดของการรู้จำเสียงพูด.....5

ตารางที่ 4.1 อัตราการรู้จำที่ได้จากชุดทดสอบที่ 2
เพื่อหาจำนวนส่วนย่อยที่เหมาะสมสำหรับเสียงคำ 1 พยางค์.....44

ตารางที่ 4.2 อัตราการรู้จำที่ได้จากชุดทดสอบที่ 2
เพื่อหาจำนวนส่วนย่อยที่เหมาะสมสำหรับเสียงคำ 2 พยางค์.....44

ตารางที่ 4.3 อัตราการรู้จำที่ได้จากชุดทดสอบที่ 2
เพื่อหาจำนวนส่วนย่อยที่เหมาะสมสำหรับเสียงคำ 3 พยางค์.....45

ตารางที่ 4.4 อัตราการรู้จำที่ได้จากชุดทดสอบที่ 2 สำหรับกลุ่มคำศัพท์ที่ใช้
จำนวนโหนดในระดับชั้นข้อมูลออกเท่ากันคือ 5 โหนด.....46

ตารางที่ 4.5 อัตราการรู้จำที่ได้จากชุดทดสอบที่ 2 สำหรับกลุ่มคำศัพท์ที่ใช้
จำนวนโหนดในระดับชั้นข้อมูลเข้าเท่ากันคือ 460 โหนด.....46

ตารางที 4.4 จำนวนโหนดในระดับชั้นซ่อนตัวที่เหมาะสมสำหรับแต่ละกลุ่มคำศัพท์.....47

ตารางที่ 4.7 ผลอัตราการรู้จำจากชุดทดสอบที่ 2 สำหรับหาค่า *fdenom* ที่เหมาะสม.....48

ตารางที่ 4.8 ผลอัตราการรู้จำจากชุดทดสอบที่ 2 สำหรับหาค่า f_d และ f_e ที่เหมาะสม.....49

ตารางที่ 4.9 อัตราการรู้จำในกรณีที่ใช้ข้อมูลออกที่ต้องการเลขฐานสอง.....54

ตารางที่ 4.10 อัตราการรู้จำในกรณีที่ใช้ข้อมูลออกที่ต้องการแบบ
ค่าสมาชิกภาพของแต่ละคำศัพท์.....54

ตารางที่ 4.11 อัตราการรู้จำในกรณีที่ใช้ข้อมูลเข้าเป็นค่าสัมประสิทธิ์ LPC.....58

ตารางที่ 4.12 อัตราการรู้จำในกรณีที่ใช้ข้อมูลเข้าเป็น
ค่าสมาชิกภาพแบบฟัซซีชนิดสามเหลี่ยม (Tg).....59

ตารางที่ 4.13 อัตราการรู้จำในกรณีที่ใช้ข้อมูลเข้าเป็น
ค่าสมาชิกภาพแบบฟัซซีชนิด Pi.....59

ตารางที่ 4.14 ตัวอย่างผลการรู้จำเสียงจำนวน 100 เสียงจากชุดทดสอบที่ 2 เมื่อใช้ข้อมูลเข้า
เป็นค่าสัมประสิทธิ์ LPC โดยใช้ผู้พูดจำนวน 50 คนในการฝึกฝน.....63

หน้า

ตารางที่ 4.15 ตัวอย่างผลการรู้จำเสียงจำนวน 100 เสียงจากชุดทดสอบที่ 2 เมื่อใช้ข้อมูลเข้าเป็นค่าสมาชิกภาพแบบฟัซซีชนิดสามเหลี่ยม โดยใช้ผู้พูดจำนวน 50 คนในการฝึกฝน 64
ตารางที่ 4.16 ตัวอย่างผลการรู้จำเสียงจำนวน 100 เสียงจากชุดทดสอบที่ 2 เมื่อใช้ข้อมูลเข้าเป็นค่าสมาชิกภาพแบบฟัซซีชนิดสี่เหลี่ยมคางหมู โดยใช้ผู้พูดจำนวน 50 คนในการฝึกฝน 65
ตารางที่ 4.17 ตัวอย่างผลการรู้จำเสียงจำนวน 100 เสียงจากชุดทดสอบที่ 2 เมื่อใช้ข้อมูลเข้าเป็นค่าสมาชิกภาพแบบฟัซซีชนิด Pi โดยใช้ผู้พูดจำนวน 50 คนในการฝึกฝน 66
ตารางที่ 4.18 ผลอัตราการรู้จำสำหรับกลุ่มคำศัพท์ที่ 2-5 เมื่อใช้ผู้พูด 50 คนในการฝึกฝน 68
ตารางที่ 4.19 ผลอัตราการรู้จำสำหรับกลุ่มคำศัพท์ที่ 6-19 ยกเว้นกลุ่มคำศัพท์ที่ 17 และ 19 เมื่อใช้ผู้พูด 50 คนในการฝึกฝน 70
ตารางที่ ก.1 ชุดคำศัพท์ตัวเลข 79
ตารางที่ ก.2 ชุดคำศัพท์หนึ่งพยางค์.....80
ตารางที่ ก.3 ชุดคำศัพท์สองพยางค์.....81
ตารางที่ ก.4 ชุดคำศัพท์สามพยางค์.....82
ตารางที่ ก.5 รายละเอียดกลุ่มคำศัพท์ย่อย 83
ตารางที่ ค.1 จำนวนจุดข้อมูลที่ควรบวกเผื่อที่หัว และท้ายคำตัวอย่างจากผู้พูด 5 คน 91

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญรูปประกอบ

หน้า

รูปที่ 2.1 โครงสร้างของการรู้จำเสียงคำเดี่ยว โดยใช้คำเป็นตัวทาบเทียบ 5
รูปที่ 2.2 โครงสร้างของนิวรอลเน็ตเวอร์กแบบ MLP 7
รูปที่ 2.3 รายละเอียดของโหนดในนิวรอลเน็ตเวอร์กแบบ MLP 9
รูปที่ 2.4 โครงสร้างขั้นตอนในการฝึกฝนแบบ Back-propagation 10
รูปที่ 2.5 แผนภาพแสดงสมาชิกในเซต A, B และ $\widetilde{A}, \widetilde{B}$ 13
รูปที่ 2.6 ตัวอย่างกราฟความสัมพันธ์ของค่าสมาชิกภาพแบบฟัซซีกับค่าของสมาชิก
ในเซต $\widetilde{A}$ ซึ่งเป็นเซตของความสูงของคนที่จัดว่าสูงปานกลาง 14
รูปที่ 2.7 กราฟแสดงฟังก์ชันสมาชิกภาพแบบฟัซซีทั้ง 3 ชนิด 14
รูปที่ 2.8 ตัวอย่างกราฟแสดงฟังก์ชันสมาชิกภาพแบบฟัซซี
บนคุณสมบัติทางภาษา 3 ระดับ.....15
รูปที่ 2.9 แผนภาพแสดงขั้นตอนการสกัดเวกเตอร์ข้อมูลเข้าเป็นค่าสมาชิกภาพแบบฟัซซี...17
รูปที่ 2.10 การนอร์มอลไลซ์ทางเวลาโดยใช้การเปลี่ยนอัตราการชักตัวอย่าง.....22
รูปที่ 2.11 ตัวอย่างการนอร์มอลไลซ์ทางเวลาโดยการประมาณค่าในช่วงเชิงเส้น 22
รูปที่ 3.1 โครงสร้างของระบบรู้จำเสียงโดยใช้เทคนิคแบบฟัซซีร่วมกับนิวรอลเน็ตเวอร์ก..28
รูปที่ 3.2 โครงสร้างขั้นตอนในกระบวนการประมวลผลเบื้องต้น 29
รูปที่ 3.3 ตัวอย่างรูปคลื่น และพลังงานของสัญญาณเสียง 30
รูปที่ 3.4 โครงสร้างขั้นตอนกระบวนการสกัดค่าลักษณะเด่น 31
รูปที่ 3.5 กราฟแสดงฟังก์ชันสมาชิกภาพแบบฟัซซีชนิดสี่เหลี่ยมคางหมู
บนคุณสมบัติทางภาษา 3 ระดับ.....34
รูปที่ 3.6 กราฟแสดงฟังก์ชันสมาชิกภาพแบบฟัซซีชนิดสามเหลี่ยม
บนคุณสมบัติทางภาษา 3 ระดับ.....34
รูปที่ 3.7 กราฟแสดงฟังก์ชันสมาชิกภาพแบบฟัซซีชนิด Pi
บนคุณสมบัติทางภาษา 3 ระดับ.....36
รูปที่ 3.8 กราฟแสดงค่าความถี่มูลฐานเทียบกับเวลา.....39
รูปที่ 4.1 กราฟการกระจายของเสียงคำที่มีจำนวนส่วนย่อยต่างๆ กัน.....43

หน้า

รูปที่ 4.2 ตัวอย่างเวกเตอร์ข้อมูลออกที่ต้องการแบบค่าสมาชิกภาพของแต่ละคำศัพท์ ในกรณี f_d และ f_e ต่างๆ กัน 50
รูปที่ 4.3 ตัวอย่างเวกเตอร์ข้อมูลออกที่ต้องการแบบต่างๆ.....51
รูปที่ 4.4(ก) กราฟแสดงอัตราการรู้จำเมื่อใช้ข้อมูลออกที่ต้องการแบบเลขฐานสอง.....55
รูปที่ 4.4(ข) กราฟแสดงอัตราการรู้จำเมื่อใช้ข้อมูลออกที่ต้องการแบบ สมาชิกภาพของแต่ละคำศัพท์.....55
รูปที่ 4.5 ตัวอย่างเวกเตอร์ข้อมูลออกที่ต้องการ เพื่อจากวิเคราะห์ผลการทดลองในกราฟรูปที่ 4.4.....56
รูปที่ 4.6 ตัวอย่างรูปคลื่นของคำศัพท์หมายเลข 410 ซึ่งไปคล้ายกับคำศัพท์หมายเลข 403...57
รูปที่ 4.7(ก) กราฟเปรียบเทียบอัตราการรู้จำจากชุดทดสอบที่ 1 (ขึ้นต่อผู้พูด).....60
รูปที่ 4.7(ก) กราฟเปรียบเทียบอัตราการรู้จำจากชุดทดสอบที่ 2 (ไม่ขึ้นต่อผู้พูด).....60
รูปที่ 4.8 กราฟแสดงผลอัตราการรู้จำที่เพิ่มขึ้นเฉลี่ย เมื่อใช้ข้อมูลเข้าเป็นค่าสมาชิกภาพแบบฟัซซีชนิดต่างๆ เทียบกับเมื่อใช้ข้อมูลเข้าแบบค่าสัมประสิทธิ์ LPC 61
รูปที่ 4.9(ก) ตัวอย่างรูปคลื่นของเสียงตัวเลขสอง ที่วิเคราะห์ผิดเป็นเลขสาม แต่วิเคราะห์ได้ถูกต้องเมื่อใช้ข้อมูลเข้าเป็นค่าสมาชิกภาพแบบฟัซซี.....67
รูปที่ 4.9(ข) ตัวอย่างรูปคลื่นของเสียงตัวเลขเจ็ด ที่วิเคราะห์ผิดเป็นเลขหก แต่วิเคราะห์ได้ถูกต้องเมื่อใช้ข้อมูลเข้าเป็นค่าสมาชิกภาพแบบฟัซซี.....67
รูปที่ 4.10 ตัวอย่างรูปคลื่นของคำศัพท์ของคำศัพท์หมายเลข 111 ซึ่งใกล้เคียงกับคำศัพท์หมายเลข 407 และอยู่ในกลุ่มคำศัพท์เดียวกัน 71
รูปที่ ข.1 แผนภาพกระบวนการตัดหัวท้ายหน่วยในการรู้จำ.....85
รูปที่ ข.2(ก) แผนภาพกระบวนการแปลงเป็นค่าสมาชิกภาพแบบฟัซซี.....86
รูปที่ ข.2(ข) แผนภาพกระบวนการแปลงเป็นค่าสมาชิกภาพแบบฟัซซีชนิดสี่เหลี่ยมคางหมู...87
รูปที่ ข.2(ค) แผนภาพกระบวนการแปลงเป็นค่าสมาชิกภาพแบบฟัซซีชนิดสามเหลี่ยม 87
รูปที่ ข.2(ง) แผนภาพกระบวนการแปลงเป็นค่าสมาชิกภาพแบบฟัซซีชนิด Pi 88
รูปที่ ข.3 แผนภาพกระบวนการสกัดข้อมูลออกแบบค่าสมาชิกภาพของแต่ละคำศัพท์ 89
รูปที่ ข.4 แผนภาพกระบวนการฝึกฝนนิวรอลเน็ตเวอร์ก.....90

คำอธิบายคำศัพท์

ขั้นตอนกระบวนการ	Algorithm
การเคลือบแฝง	Aliasing
อัตสหสัมพันธ์	Autocorrelation
การแพร่กระจายย้อนกลับ	Back-propagation
สมาชิกภาพของแต่ละคำศัพท์	Class Membership
น้ำหนักการเชื่อมต่อ	Connection Weight
ขึ้นต่อบริบท	Context-dependent
เสียงต่อเนื่อง	Continuous Speech
ความแปรปรวนร่วม	Covariance
เซตแบบแน่นอน	Crisp Set
ข้อมูลออกที่ต้องการ	Desired Output
ลักษณะบ่งความต่าง	Distinctive Feature
พิสัยพลวัต	Dynamic Range
การตัดหัวท้ายหน่วยในการรู้จำ	Endpoint Detection
การสกัด	Extraction
ลักษณะเด่น	Feature
แพร่กระจายไปข้างหน้า	Feed Forward
ส่วนย่อย	Frame
ความถี่มูลฐาน	Fundamental Frequency
กรณีกำกวมที่สุด	Fuzziest Case
การแปลงเป็นค่าสมาชิกภาพแบบฟัซซี	Fuzzification
ความกำกวม	Fuzziness
ระดับชั้นซ่อนตัว	Hidden Layer
คำเดี่ยว	Isolated Word
การเรียนรู้	Learning
การประมาณค่าในช่วงเชิงเส้น	Linear Interpolation
การประมาณพันธะเชิงเส้น	Linear Prediction

ทางภาษา	Linguistic
สมาชิกภาพ	Membership
การกระตุ้นแบบไม่เป็นเชิงเส้น	Nonlinear Activation
การซ้อนทับ	Overlapping
หน่วยเสียง	Phoneme
รูปเสียง	Phonetic
คำหลายพยางค์	Polysyllabic Word
การรู้จำ	Recognition
รากที่สองของกำลังสองเฉลี่ย	Root Mean Square
อัตราการชักตัวอย่าง	Sampling Rate
ค่าที่กำหนด	Threshold
การฝึกฝน	Training
เสียงก้อง	Voiced Speech
ฟังก์ชันหน้าต่าง	Window Function
การตัดแกนศูนย์	Zero Crossing

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย