การรู้จำหน่วยเสียงสระภาษาไทยโดยใช้โครงข่ายประสาทเทียม

นาย เอกฤทธิ์ มณีน้อย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา. 2541

ISBN 974-331-422-9

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

I18994690

THAI VOWEL PHONEME RECOGNITION USING ARTIFICIAL NEURAL NETWORKS

Mr. Ekkarit Maneenoi

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 1998

ISBN 974-331-422-9

หัวข้อวิทยานิพนธ์	การรู้จำหน่วยเสียงสระภาษาไทยโดยใช้โครงข่ายประสาทเทียม
โดย	นาย เอกฤทธิ์ มณีน้อย
ภาควิชา	วิศวกรรมไฟฟ้า
อาจารย์ที่ปรึกษา	รองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล
อาจารย์ที่ปรึกษาร่วม	ผู้ช่วยศาสตราจารย์ ดร. สุดาพร ลักษณียนาวิน

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

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

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. วาทิต เบญจพลกุล)

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

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

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

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เอกฤทธิ์ มณีน้อย : การรู้จำหน่วยเสียงสระภาษาไทยโดยใช้โครงข่ายประสาทเทียม (THAI VOWEL PHONEME RECOGNITION USING ARTIFICIAL NEURAL NETWORKS) อาจารย์ที่ปรึกษา : รศ. ดร. สมชาย จิตะพันธ์กุล, อาจารย์ที่ปรึกษาร่วม : ผศ. ดร. สุดาพร ลักษณียนาวิน, 118 หน้า. ISBN 974-331-422-9.

วิทยานิพนธ์นี้มีจุดมุ่งหมายเพื่อพัฒนาระบบรู้จำหน่วยเสียงสระภาษาไทยโดยใช้โครงข่ายประสาทเทียม โครงข่ายประสาทเทียมที่ใช้ในงานวิจัยนี้มีโครงสร้างแบบ MLP ที่มีจำนวนชั้นซ่อนตัวหนึ่งชั้นและใช้ขั้นตอนวิธีการส่งค่าย้อนกลับในการฝึกฝน โครงข่ายประสาทเทียมจะถูกแบ่งออกเป็นโครงข่ายย่อยสำหรับรู้จำหน่วยเสียงสระเดี่ยวจากชุดคำศัพท์ที่ประกอบด้วยเสียงสระเดี่ยวภาษาไทย 9 หน่วยเสียงเพื่อการสร้างต้นแบบหน่วยเสียงที่มีประสิทธิภาพ นอกจากการรู้จำหน่วยเสียงสระเดี่ยวโดยโครงข่ายประสาทเทียมแล้วงานวิจัยนี้ได้สร้างแบบจำลองทางสถิติของช่วงความยาวเสียงสระเพื่อทำการรู้จำสระเสียงสั้น-ยาวกับชุดคำศัพท์ที่ประกอบด้วยสระเสียงสั้นยาวจำนวน 6 เสียง ค่าลักษณะสำคัญที่ใช้ในงานวิจัยนี้ได้แก่ สัมประสิทธิ์การประมาณพันธะเชิงเส้น สัมประสิทธิ์ cepstral และความถี่ฟอร์แมนท์ ผลการทดสอบอัตราการรู้จำมีค่าร้อยละ 90.34 เมื่อใช้การแบ่งกลุ่มหน่วยเสียงสระตามความถี่ฟอร์แมนท์และใช้ค่าสัมประสิทธิ์ cepstral เป็นค่าลักษณะสำคัญ สูงกว่ากรณีที่ไม่มีการแบ่งกลุ่มหน่วยเสียงสระตามความถี่ฟอร์แมนท์ที่มีอัตราการรู้จำร้อยละ 85.92 โดยที่ผู้พูดในชุดฝึกฝนมีจำนวน 30 คน และชุดทดสอบแบบไม่ขึ้นกับผู้พูด 20 คน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา	วิศวกรรมไฟฟ้า	**ลายมือชื่อนิสิต**	
สาขาวิชา	วิศวกรรมไฟฟ้า	**ลายมือชื่ออาจารย์ที่ปรึกษา**	
ปีการศึกษา	2541	**ลายมือชื่ออาจารย์ที่ปรึกษาร่วม**	

4070520621 : MAJOR ELECTRICAL ENGINEERING
KEY WORD: VOWEL PHONEME RECOGNITION / PHONEME RECOGNITION / ARTIFICIAL NEURAL NETWORKS

EKKARIT MANEENOI : THAI VOWEL PHONEME RECOGNITION USING ARTIFICIAL NEURAL NETWORKS. THESIS ADVISOR : ASSOC. PROF. SOMCHAI JITAPUNKUL, Dr. Ing. THESIS CO-ADVISOR : ASSIST. PROF. SUDAPORN LUKSANEEYANAVIN, Ph.D. 118 pp. ISBN 974-331-422-9.

The objective of this research is to develop the Thai vowel-phoneme recognition system using an artificial neural network. The single hidden layer MLP neural network with the backpropagation algorithm, is employed in this research. The artificial neural network is divided into subnetworks for recognizing the 9 vowel phonemes from the vocabulary set. Furthermore, this thesis has also constructed the statistical model for recognizing short and long vowels. The features used in this research are linear prediction coefficient, cepstral coefficient, and formant frequency. The recognition rate of this system is 90.34 percent when the subdivided artificial neural networks correspond to formant frequency and cepstral coefficient. The recognition rate of this subdivided network is higher than the single network of which the recognition rate is 85.92 percent. Thirty training speakers and twenty test speakers are employed in this research.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา วิศวกรรมไฟฟ้า ลายมือชื่อนิสิต .....

สาขาวิชา วิศวกรรมไฟฟ้า ลายมือชื่ออาจารย์ที่ปรึกษา .....

ปีการศึกษา 2541 ลายมือชื่ออาจารย์ที่ปรึกษาร่วม .....

เอกฤทธิ์ มณีน้อย : การรู้จำหน่วยเสียงสระภาษาไทยโดยใช้โครงข่ายประสาทเทียม (THAI VOWEL PHONEME RECOGNITION USING ARTIFICIAL NEURAL NETWORKS) อาจารย์ที่ปรึกษา : รศ. ดร. สมชาย จิตะพันธ์กุล, อาจารย์ที่ปรึกษาร่วม : ผศ. ดร. สุดาพร ลักษณียนาวิน, 118 หน้า. ISBN 974-331-422-9

วิทยานิพนธ์นี้มีจุดมุ่งหมายเพื่อพัฒนาระบบรู้จำหน่วยเสียงสระภาษาไทยโดยใช้โครงข่ายประสาทเทียม โครงข่ายประสาทเทียมที่ใช้ในงานวิจัยนี้มีโครงสร้างแบบ MLP ที่มีจำนวนชั้นซ่อนตัวหนึ่งชั้น และใช้ขั้นตอนวิธีการส่งค่าย้อนกลับในการฝึกฝน โครงข่ายประสาทเทียมจะถูกแบ่งออกเป็นโครงข่ายย่อยสำหรับรู้จำหน่วยเสียงสระเดี่ยวจากชุดคำศัพท์ที่ประกอบด้วยเสียงสระเดี่ยวภาษาไทย 9 หน่วยเสียงเพื่อการสร้างต้นแบบหน่วยเสียงที่มีประสิทธิภาพ นอกจากการรู้จำหน่วยเสียงสระเดี่ยวโดยโครงข่ายประสาทเทียมแล้วงานวิจัยนี้ได้สร้างแบบจำลองทางสถิติของช่วงความยาวเสียงสระเพื่อทำการรู้จำสระเสียงสั้น-ยาวกับชุดคำศัพท์ที่ประกอบด้วยสระเสียงสั้นยาวจำนวน 6 เสียง ค่าลักษณะสำคัญที่ใช้ในงานวิจัยนี้ได้แก่ สัมประสิทธิ์การประมาณพันธะเชิงเส้น สัมประสิทธิ์ cepstral และความถี่ฟอร์แมนท์ ผลการทดสอบอัตราการรู้จำมีค่าร้อยละ 90.34 เมื่อใช้การแบ่งกลุ่มหน่วยเสียงสระตามความถี่ฟอร์แมนท์และใช้ค่าสัมประสิทธิ์ cepstral เป็นค่าลักษณะสำคัญ สูงกว่ากรณีที่ไม่มีการแบ่งกลุ่มหน่วยเสียงสระตามความถี่ฟอร์แมนท์ที่มีอัตราการรู้จำร้อยละ 85.92 โดยที่ผู้พูดในชุดฝึกฝนมีจำนวน 30 คน และชุดทดสอบแบบไม่ขึ้นกับผู้พูด 20 คน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

4070520621: MAJOR ELECTRICAL ENGINEERING

KEY WORD: VOWEL PHONEME RECOGNITION / PHONEME RECOGNITION / ARTIFICIAL NEURAL NETWORKS

EKKARIT MANEENOI : THAI VOWEL PHONEME RECOGNITION USING ARTIFICIAL NEURAL NETWORKS. THESIS ADVISOR : ASSOC. PROF. SOMCHAI JITAPUNKUL, Dr. Ing. THESIS COADVISOR : ASSIST. PROF. SUDAPORN LUKSANEEYANAVIN, Ph.D. 118 pp. ISBN 974-331-422-9

The objective of this research is to develop the Thai vowel-phoneme recognition system using an artificial neural network. The single hidden layer MLP neural network with the backpropagation algorithm, is employed in this research. The artificial neural network is divided into subnetworks for recognizing the 9 vowel phonemes from the vocabulary set. Furthermore, this thesis has also constructed the statistical model for recognizing short and long vowels. The features used in this research are linear prediction coefficient, cepstral coefficient, and formant frequency. The recognition rate of this system is 90.34 percent when the subdivided artificial neural networks correspond to formant frequency and cepstral coefficient. The recognition rate of this subdivided network is higher than the single network of which the recognition rate is 85.92 percent. Thirty training speakers and twenty test speakers are employed in this research.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

กิตติกรรมประกาศ

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล อาจารย์ที่ปรึกษา ที่ได้กรุณาให้คำแนะนำในการดำเนินการวิจัย ให้คำปรึกษา จัดหาอุปกรณ์และทุนวิจัยเพื่อดำเนินการวิจัย และขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร. สุดาพร ลักษณียนาวิน อาจารย์ที่ปรึกษาร่วมที่ได้กรุณาให้คำแนะนำและให้คำปรึกษาทางด้านภาษาศาสตร์ซึ่งเป็นประโยชน์อย่างยิ่งต่อการทำวิจัย

ขอขอบคุณ คุณวิศรุต อาชุบุตร ที่ได้กรุณาสละเวลาให้คำแนะนำและคำปรึกษาที่สำคัญอย่างยิ่งต่อการดำเนินการวิจัย

ขอขอบคุณเพื่อนนิสิตปริญญาโทและนิสิตปริญญาตรีทุกท่านที่ได้ช่วยเหลือการดำเนินการวิจัยและเป็นกำลังใจตลอดมา

ขอขอบคุณโครงการกลุ่มวิจัยบัณฑิตศึกษาด้านโทรคมนาคมที่สนับสนุนทุนการศึกษาและค่าใช้จ่ายในการนำเสนอผลงานวิจัย

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....ง

บทคัดย่อภาษาอังกฤษ.....จ

กิตติกรรมประกาศ.....ฉ

สารบัญตาราง.....ฌ

สารบัญภาพ.....ฏ

บัญชีคำศัพท์.....ฒ

บทที่ 1 บทนำ.....1

1.1 แนวเหตุผล.....1

1.2 ปัญหาของการรู้จำเสียงพูดภาษาไทยที่ผ่านมา.....4

1.3 วัตถุประสงค์.....4

1.4 เป้าหมายและขอบเขตของงานวิจัย.....4

1.5 ขั้นตอนและวิธีการดำเนินการ.....4

1.6 ประโยชน์ที่คาดว่าจะได้รับ.....5

บทที่ 2 หลักการและทฤษฎีสำคัญ.....8

2.1 การประมวลผลสัญญาณเบื้องต้น.....8

2.2 การวิเคราะห์และวัดค่าลักษณะสำคัญ.....10

2.3 โครงข่ายประสาทเทียม.....21

2.4 กลสัทศาสตร์และภาษาศาสตร์.....32

2.5 การรู้จำแบบรูปโดยวิธีเชิงสถิติ.....34

บทที่ 3 วิธีดำเนินการวิจัย.....37

3.1 การเลือกคำศัพท์ภาษาไทย.....37

3.2 การสร้างระบบรู้จำหน่วยเสียงสระภาษาไทย.....37

3.3 ขั้นตอนการฝึกฝนระบบรู้จำหน่วยเสียงสระภาษาไทย.....39

3.4 ขั้นตอนการทดสอบระบบรู้จำหน่วยเสียงสระภาษาไทย.....46

บทที่ 4 ผลการวิจัย.....49

4.1 พารามิเตอร์ที่สำคัญในการฝึกฝนของโครงข่ายประสาทเทียมและข้อมูลเสียงพูด.....49

4.2 การแบ่งกลุ่มหน่วยเสียงสระภาษาไทย.....49

4.3 การทดลองเพื่อหาจำนวนโนดในชั้นซ่อนตัว.....50

4.4 ผลการรู้จำหน่วยเสียงสระภาษาไทยแต่ละกลุ่มย่อย.....53

4.5 ผลการรู้จำหน่วยเสียงสระภาษาไทยทั้งระบบ.....62

4.6 ผลการรู้จำสระเสียงสั้น-ยาว.....78
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....84
5.1 สรุปผลการวิจัย.....84
5.2 ข้อเสนอแนะ.....85
รายการอ้างอิง.....87
ภาคผนวก ก รายการชุดคำศัพท์ภาษาไทย.....91
ภาคผนวก ข สเปกตรัมของของหน่วยเสียงสระและความถี่ฟอร์แมนท์.....92
ภาคผนวก ค รายละเอียดผลการรู้จำหน่วยเสียงสระภาษาไทย.....101
ประวัติผู้วิจัย.....118

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

หน้า

ตารางที่ 1.1 งานวิจัยเกี่ยวกับระบบรู้จำเสียงพูดภาษาไทย.....5

ตารางที่ 2.1 ขั้นตอนวิธีการ Levinson-Durbin.....15

ตารางที่ 2.2 หน่วยเสียงสระเดี่ยวภาษาไทย.....33

ตารางที่ 4.1 กลุ่มหน่วยเสียงสระ.....50

ตารางที่ 4.2 ผลของจำนวนโนดในชั้นซ่อนตัวกับอัตราการรู้จำของหน่วยเสียงสระ
ในกลุ่มที่ 3 (ii, vv, uu).....51

ตารางที่ 4.3 ผลของจำนวนโนดในชั้นซ่อนตัวกับอัตราการรู้จำของหน่วยเสียงสระ
ในกลุ่มที่ 9 (สระทุกหน่วย).....52

ตารางที่ 4.4 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 3
หรือสระสูง (หน้า-กลาง-หลัง) ที่ใช้สัมประสิทธิ์ cepstral 16 อันดับเป็นค่าลักษณะสำคัญ.....60

ตารางที่ 4.5 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 4
หรือสระกลาง (หน้า-กลาง-หลัง) ที่ใช้สัมประสิทธิ์ cepstral 16 อันดับเป็นค่าลักษณะสำคัญ..60

ตารางที่ 4.6 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 5
หรือสระต่ำ (หน้า-กลาง-หลัง) ที่ใช้สัมประสิทธิ์ cepstral 16 อันดับเป็นค่าลักษณะสำคัญ.....60

ตารางที่ 4.7 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 6
หรือสระหน้า (สูง-กลาง-ต่ำ) ที่ใช้สัมประสิทธิ์ cepstral 16 อันดับเป็นค่าลักษณะสำคัญ.....61

ตารางที่ 4.8 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 7
หรือสระกลาง (สูง-กลาง-ต่ำ) ที่ใช้สัมประสิทธิ์ cepstral 16 อันดับเป็นค่าลักษณะสำคัญ.....61

ตารางที่ 4.9 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 8
หรือสระหลัง (สูง-กลาง-ต่ำ) ที่ใช้สัมประสิทธิ์ cepstral 16 อันดับเป็นค่าลักษณะสำคัญ.....61

ตารางที่ 4.10 ผลสรุปการจัดกลุ่มหน่วยเสียงสระตามความถี่ฟอร์แมนท์ที่ 1
ที่ใช้สัมประสิทธิ์ cepstral 16 อันดับ เป็นค่าลักษณะสำคัญ.....76

ตารางที่ 4.11 ผลสรุปการจัดกลุ่มหน่วยเสียงสระตามความถี่ฟอร์แมนท์ที่ 2
ที่ใช้สัมประสิทธิ์ cepstral 16 อันดับ เป็นค่าลักษณะสำคัญ.....76

ตารางที่ 4.12 อัตราการรู้จำหน่วยเสียงสระไม่มีการแบ่งกลุ่มหน่วยเสียงตามความถี่ฟอร์แมนท์
ที่ใช้สัมประสิทธิ์ cepstral 16 อันดับเป็นค่าลักษณะสำคัญ.....77

ตารางที่ 4.13 ค่าลักษณะสำคัญที่ให้อัตราการรู้จำสูงสุดสำหรับแต่ละกลุ่มหน่วยเสียงสระ.....77

ตารางที่ 4.14 อัตราการรู้จำสำหรับแต่ละกลุ่มหน่วยเสียงสระภายหลังจากการแบ่งกลุ่มหน่วยเสียงสระ.....78

ตารางที่ 4.15 ค่าทางสถิติของแบบจำลองช่วงเวลาความยาวของสระเสียงสั้น-ยาว.....82

หน้า

ตารางที่ 4.16 ความผิดพลาดของแบบจำลองช่วงเวลาความยาวของสระเสียงสั้น-ยาว
ที่มีการแจกแจงแบบปกติหรือแบบเกาส์....82
ตารางที่ 4.17 ผลการรู้จำสระเสียงสั้น-ยาว....82
ตารางที่ ก.1 รายละเอียดชุดคำศัพท์ที่ใช้สำหรับการรู้จำหน่วยเสียงสระ....91
ตารางที่ ก.2 รายละเอียดชุดคำศัพท์ที่ใช้สำหรับการจำแนกเสียงสระสั้น-ยาว....91
ตารางที่ ข.1 ความถี่ฟอร์แมนท์เฉลี่ยของหน่วยเสียงสระอี....92
ตารางที่ ข.2 ความถี่ฟอร์แมนท์เฉลี่ยของหน่วยเสียงสระอือ....93
ตารางที่ ข.3 ความถี่ฟอร์แมนท์เฉลี่ยของหน่วยเสียงสระอู....94
ตารางที่ ข.4 ความถี่ฟอร์แมนท์เฉลี่ยของหน่วยเสียงสระเอ....95
ตารางที่ ข.5 ความถี่ฟอร์แมนท์เฉลี่ยของหน่วยเสียงสระเออ....96
ตารางที่ ข.6 ความถี่ฟอร์แมนท์เฉลี่ยของหน่วยเสียงสระโอ....97
ตารางที่ ข.7 ความถี่ฟอร์แมนท์เฉลี่ยของหน่วยเสียงสระแอ....98
ตารางที่ ข.8 ความถี่ฟอร์แมนท์เฉลี่ยของหน่วยเสียงสระอา....99
ตารางที่ ข.9 ความถี่ฟอร์แมนท์เฉลี่ยของหน่วยเสียงสระออ....100
ตารางที่ ค.1 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ 3 หรือสระสูง (หน้า-กลาง-หลัง)....101
ตารางที่ ค.2 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ 4 หรือสระกลาง (หน้า-กลาง-หลัง)....102
ตารางที่ ค.3 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ 5 หรือสระต่ำ (หน้า-กลาง-หลัง)....103
ตารางที่ ค.4 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ 6 หรือสระหน้า (สูง-กลาง-ต่ำ)....104
ตารางที่ ค.5 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ ที่ 7 หรือสระกลาง (สูง-กลาง-ต่ำ)....105
ตารางที่ ค.6 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ 8 หรือสระหลัง (สูง-กลาง-ต่ำ)....106
ตารางที่ ค.7 ผลสรุปอัตราการจำแนกหน่วยเสียงสระแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
ตามความถี่ฟอร์แมนท์ที่ 1 (เสียงสระกลุ่มที่ 1)....107
ตารางที่ ค.8 ผลสรุปอัตราการจำแนกหน่วยเสียงสระแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
ตามความถี่ฟอร์แมนท์ที่ 2 (เสียงสระกลุ่มที่ 2)....108
ตารางที่ ค.9 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ 3 หรือสระสูง (หน้า-กลาง-หลัง) เมื่อพิจารณาทั้งระบบ....109

ตารางที่ ค.10 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ 4 หรือสระกลาง (หน้า-กลาง-หลัง) เมื่อพิจารณาทั้งระบบ.....110
ตารางที่ ค.11 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ 5 หรือสระต่ำ (หน้า-กลาง-หลัง) เมื่อพิจารณาทั้งระบบ.....111
ตารางที่ ค.12 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ 6 หรือสระหน้า (สูง-กลาง-ต่ำ) เมื่อพิจารณาทั้งระบบ.....112
ตารางที่ ค.13 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ 7 หรือสระกลาง (สูง-กลาง-ต่ำ) เมื่อพิจารณาทั้งระบบ.....113
ตารางที่ ค.14 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ 8 หรือสระหลัง (สูง-กลาง-ต่ำ) เมื่อพิจารณาทั้งระบบ.....114
ตารางที่ ค.15 ผลสรุปอัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระกลุ่มที่ 9.....115
ตารางที่ ค.16 ผลสรุปอัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระที่แบ่งกลุ่มตามความถี่ฟอร์แมนท์ที่ 1 เมื่อพิจารณาทั้งระบบ.....116
ตารางที่ ค.17 ผลสรุปอัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
สำหรับหน่วยเสียงสระที่แบ่งกลุ่มตามความถี่ฟอร์แมนท์ที่ 2 เมื่อพิจารณาทั้งระบบ.....117

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

หน้า

รูปที่ 1.1 แบบจำลองระบบรู้จำเสียงพูดภาษาไทย....7
รูปที่ 1.2 ระบบรู้จำเสียงพูดที่ใช้แบบจำลองของคำโดดและหน่วยเสียง....7
รูปที่ 2.1 ฟังก์ชันกรอบชนิด Hamming Window....9
รูปที่ 2.2 การแบ่งกรอบข้อมูลเสียงพูดด้วยฟังก์ชันกรอบชนิด Hamming Window....9
รูปที่ 2.3 การหาจุดเริ่มต้นและสิ้นสุดของพยางค์โดยวิธีพลังงาน....10
รูปที่ 2.4 แบบจำลองการประมาณพันธะเชิงเส้นสำหรับเสียงพูด....12
รูปที่ 2.5 แบบจำลองสังเคราะห์เสียงพูดจากแบบจำลองการประมาณพันธะเชิงเส้น....12
รูปที่ 2.6 พหุนามการประมาณพันธะเชิงเส้นที่มีอันดับต่างกัน....17
รูปที่ 2.7 กรรมวิธีการหา Real Cepstrum....18
รูปที่ 2.8 รูปร่างของช่องทางเดินเสียงและสเปกตรัม....19
รูปที่ 2.9 รูปคลื่นสัญญาณเสียงที่ออกจากปากและที่เส้นเสียง....19
รูปที่ 2.10 ขั้นตอนการคำนวณความถี่ฟอร์แมนท์....20
รูปที่ 2.11 โพลของฟังก์ชันถ่ายโอนของแบบจำลองการประมาณพันธะเชิงเส้น....20
รูปที่ 2.12 โครงสร้างของโครงข่ายประสาทเทียม....22
รูปที่ 2.13 แบบจำลองการทำงานที่ไม่เป็นเชิงเส้นของโนด....22
รูปที่ 2.14 ตัวอย่างโครงสร้างของโครงข่ายประสาทเทียมแบบ MLPที่มีจำนวนชั้นซ่อนตัว 2 ชั้น....24
รูปที่ 2.15 ทิศทางการไหลของสัญญาณในโครงข่ายประสาทเทียมแบบ MLP....25
รูปที่ 2.16 แสดงการไหลของสัญญาณของโนดที่ j....26
รูปที่ 2.17 การไหลของสัญญาณสำหรับโนด k ในชั้นข้อมูลออกที่เชื่อมต่อกับโนด j ในชั้นซ่อนตัว....28
รูปที่ 2.18 ขั้นตอนวิธีการ backpropagation ของโครงข่ายประสาทเทียมแบบ MLP....31
รูปที่ 2.19 กราฟการไหลของสัญญาณที่แสดงผลของค่าคงที่โมเมนตัม....31
รูปที่ 2.20 แผนผังการจำแนกรูปแบบ....32
รูปที่ 2.21 ความสัมพันธ์ของค่าความถี่ฟอร์แมนท์กับความยาวของทางช่องเดินเสียง
สำหรับเสียงสระภาษาไทย....34
รูปที่ 2.22 การแปลง a priori class probability $P(w_i)$
ไปเป็น a posteriori probability $P(w_i|x)$....35
รูปที่ 2.23 บริเวณที่มีความผิดพลาดในการตัดสินใจของการใช้
แบบจำลองทางสถิติที่มีการแจกแจงแบบเกาส์....36
รูปที่ 3.1 ระบบรู้จำหน่วยเสียงสระภาษาไทย....38
รูปที่ 3.2 ขั้นตอนการฝึกฝนระบบรู้จำหน่วยเสียงสระภาษาไทย....38

รูปที่ 3.3 ขั้นตอนการทดสอบระบบรู้จำหน่วยเสียงสระภาษาไทย.....39
รูปที่ 3.4 การขอบเขตกรอบข้อมูลเสียงพูดส่วนที่เป็นสระโดยใช้ค่าพลังงานและความถี่มูลฐาน.....41
รูปที่ 3.5 ความถี่ฟอร์แมนท์ที่คำนวณจากรากของพหุนามของสัมประสิทธิ์การประมาณพันธะเชิงเส้น.....43
รูปที่ 3.6 โครงข่ายประสาทเทียมแบบ MLP ที่มีจำนวนชั้นซ่อนตัว 1 ชั้น.....44
รูปที่ 3.7 รูปแบบข้อมูลเข้าที่ชั้นข้อมูลเข้าของโครงข่ายประสาทเทียม.....44
รูปที่ 3.8 การใช้โครงข่ายประสาทเทียมในการรู้จำหน่วยเสียงสระสำหรับงานวิจัยนี้.....47
รูปที่ 3.9 ความถี่ฟอร์แมนท์ของสระเสียงสั้น-ยาว.....48
รูปที่ 4.1 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 3
หรือสระสูง (หน้า-กลาง-หลัง) เมื่อใช้ผู้พูดฝึกฝน 30 คน.....54
รูปที่ 4.2 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 4
หรือสระกลาง (หน้า-กลาง-หลัง) เมื่อใช้ผู้พูดฝึกฝน 30 คน.....55
รูปที่ 4.3 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 5
หรือสระต่ำ (หน้า-กลาง-หลัง) เมื่อใช้ผู้พูดฝึกฝน 30 คน.....56
รูปที่ 4.4 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 6
หรือสระหน้า (สูง-กลาง-ต่ำ) เมื่อใช้ผู้พูดฝึกฝน 30 คน.....57
รูปที่ 4.5 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 7
หรือสระกลาง (สูง-กลาง-ต่ำ) เมื่อใช้ผู้พูดฝึกฝน 30 คน.....58
รูปที่ 4.6 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 8
หรือสระหลัง (สูง-กลาง-ต่ำ) เมื่อใช้ผู้พูดฝึกฝน 30 คน.....59
รูปที่ 4.7 แผนภาพการกระจายของความถี่ฟอร์แมนท์ที่ 1
และความถี่ฟอร์แมนท์ที่ 2 ของหน่วยเสียงสระ.....63
รูปที่ 4.8 แผนภาพแสดงการประมาณขอบเขตของของหน่วยเสียงสระ.....64
รูปที่ 4.9 อัตราการจำแนกหน่วยเสียงสระแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
ตามความถี่ฟอร์แมนท์ที่ 1 (เสียงสระกลุ่มที่ 1).....65
รูปที่ 4.10 อัตราการจำแนกหน่วยเสียงสระแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบ
ตามความถี่ฟอร์แมนท์ที่ 2 (เสียงสระกลุ่มที่ 2).....66
รูปที่ 4.11 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับ
หน่วยเสียงสระกลุ่มที่ 3 หรือสระสูง (หน้า-กลาง-หลัง) เมื่อพิจารณาทั้งระบบ.....67
รูปที่ 4.12 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับ
หน่วยเสียงสระกลุ่มที่ 4 หรือสระกลาง (หน้า-กลาง-หลัง) เมื่อพิจารณาทั้งระบบ.....68
รูปที่ 4.13 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับ
หน่วยเสียงสระกลุ่มที่ 5 หรือสระต่ำ (หน้า-กลาง-หลัง) เมื่อพิจารณาทั้งระบบ.....69

หน้า

รูปที่ 4.14 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับ
หน่วยเสียงสระกลุ่มที่ 6 หรือสระหน้า (สูง-กลาง-ต่ำ) เมื่อพิจารณาทั้งระบบ.....70
รูปที่ 4.15 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับ
หน่วยเสียงสระกลุ่มที่ 7 หรือสระกลาง (สูง-กลาง-ต่ำ) เมื่อพิจารณาทั้งระบบ.....71
รูปที่ 4.16 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับ
หน่วยเสียงสระกลุ่มที่ 8 หรือสระหลัง (สูง-กลาง-ต่ำ) เมื่อพิจารณาทั้งระบบ.....72
รูปที่ 4.17 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระกลุ่มที่ 9.....73
รูปที่ 4.18 อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบสำหรับหน่วยเสียงสระ
ของระบบที่มีการแบ่งกลุ่มตามความถี่ฟอร์แมนท์ที่ 1 ความถี่ฟอร์แมนท์ที่ 2
และระบบที่ไม่มีการแบ่งกลุ่มตามความถี่ฟอร์แมนท์.....74
รูปที่ 4.19 แบบจำลองและแผนภาพฮิสโตแกรมช่วงเวลาความยาวของเสียงสระอิ-อี.....79
รูปที่ 4.20 แบบจำลองและแผนภาพฮิสโตแกรมช่วงเวลาความยาวของเสียงสระอึ-อือ.....80
รูปที่ 4.21 แบบจำลองและแผนภาพฮิสโตแกรมช่วงเวลาความยาวของเสียงสระอะ-อา.....81
รูปที่ 4.22 ผลการรู้จำสระเสียงสั้น-ยาว.....83
รูปที่ ข.1 ตัวอย่างสเปกตรัมของหน่วยเสียงสระอี.....92
รูปที่ ข.2 ตัวอย่างสเปกตรัมของหน่วยเสียงสระอือ.....93
รูปที่ ข.3 ตัวอย่างสเปกตรัมของหน่วยเสียงสระอู.....94
รูปที่ ข.4 ตัวอย่างสเปกตรัมของหน่วยเสียงสระเอ.....95
รูปที่ ข.5 ตัวอย่างสเปกตรัมของหน่วยเสียงสระเออ.....96
รูปที่ ข.6 ตัวอย่างสเปกตรัมของหน่วยเสียงสระโอ.....97
รูปที่ ข.7 ตัวอย่างสเปกตรัมของหน่วยเสียงสระแอ.....98
รูปที่ ข.8 ตัวอย่างสเปกตรัมของหน่วยเสียงสระอา.....99
รูปที่ ข.9 ตัวอย่างสเปกตรัมของหน่วยเสียงสระออ.....100

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บัญชีคำศัพท์

ขั้นตอนวิธีการ	algorithm
โครงข่ายประสาทเทียม	artificial neural networks
อัตสหสัมพันธ์	autocorrelation
การส่งค่าย้อนกลับ	backpropagation
คำพูดต่อเนื่อง	connected words
น้ำหนักการเชื่อมต่อ	connection weight
พยัญชนะคู่	consonant cluster
หน่วยเสียงพยัญชนะ	consonantal phoneme
เสียงพูดต่อเนื่อง	continuous speech
สระผสม	diphthong
กรรมวิธีหาจุดสิ้นสุดเสียงพูด	endpoint detection
ค่าผิดพลาดผ่านระดับ	error threshold
ค่าลักษณะสำคัญ	feature
การสกัดค่าลักษณะสำคัญ	feature extraction
การวัดค่าลักษณะสำคัญ	feature measurement
ความถี่ฟอร์แมนท์	formant frequency
กรอบสัญญาณ	frame
ความถี่มูลฐาน	fundamental frequency
ชั้นซ่อนตัว	hidden layer
ชั้นข้อมูลเข้า	input layer
คำโดด	isolated word
การวนซ้ำ	iteration
อัตราการเรียนรู้	learning rate
จุดต่ำสุดเฉพาะที่	local minimum
สัมประสิทธิ์การประมาณพันธะเชิงเส้น	linear prediction coefficient
ค่าผิดพลาดกำลังสองเฉลี่ย	mean squared error
ค่าคงที่โมเมนตัม	momentum constant
สระเดี่ยว	monophthong

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

โนด	node
ความไม่เป็นเชิงเส้น	nonlinearity
ชั้นข้อมูลออก	output layer
การจำแนกรูปแบบ	pattern classification
การรู้จำรูปแบบ	pattern recognition
หน่วยเสียง	phoneme
การเน้นล่วงหน้า	preemphasis
สัมประสิทธิ์การสะท้อน	reflection coefficient
ความถี่ชักตัวอย่าง	sampling frequency
การประมวลผลสัญญาณเบื้องต้น	signal preprocessing
ผลตอบเป้าหมาย	target output
จุดเปลี่ยนระดับ	threshold
หน่วยเสียงวรรณยุกต์	tonal phoneme
ฟังก์ชันถ่ายโอน	transfer function
เสียงอโฆษะ	unvoiced
เส้นเสียง	vocal cords
ช่องทางเดินเสียง	vocal tract
เสียงโฆษะ	voiced
หน่วยเสียงสระ	vowel phoneme
ฟังก์ชันกรอบ	window function

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย