ระบบการรู้จำผู้พูด : การบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูด

นายควิต กาสุริยะ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2542

ISBN 974-333-506-4

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

I19254015

SPEAKER RECOGNITION SYSTEM: TEXT-DEPENDENT SPEAKER IDENTIFICATION

Mr. Sawit Kasuriya

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 1999

ISBN 974-333-506-4

หัวข้อวิทยานิพนธ์	ระบบการรู้จำผู้พูด : การบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูด
โดย	นายควิต กาสุริยะ
ภาควิชา	วิศวกรรมไฟฟ้า
อาจารย์ที่ปรึกษา	รองศาสตราจารย์ ดร.สมชาย จิตะพันธ์กุล
อาจารย์ที่ปรึกษาร่วม	ผู้ช่วยศาสตราจารย์ ดร.สุดาพร ลักษณียนาวิน

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

..... คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร.ธัชชัย สุมิตร)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(รองศาสตราจารย์ ดร.วาทิต เบญจพลกุล)

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

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

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

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ควิต กาสุริยะ : ระบบการรู้จำผู้พูด : การบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูด (SPEAKER RECOGNITION SYSTEM: TEXT-DEPEDENT SPEAKER IDENTIFICATION) อาจารย์ที่ปรึกษา : รศ. ดร.สมชาย จิตะพันธ์กุล, อาจารย์ที่ปรึกษาร่วม : ผศ. ดร.สุดาพร ลักษณียนาวิน, 133 หน้า. ISBN 974-333-506-4.

วิทยานิพนธ์ฉบับนี้มีวัตถุประสงค์เพื่อพัฒนาระบบการรู้จำผู้พูดชนิดการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูดภาษาไทย 3 ประโยคโดยใช้แบบจำลองฮิดเดนมาร์คอฟแบบดิสครีตร่วมกับการควอนไทซ์แบบเวกเตอร์ในการรู้จำเสียงของผู้พูดจำนวน 12 คนซึ่งประกอบด้วยเพศชาย 6 คน และเพศหญิง 6 คน วิธีการสำคัญที่ทดลองใช้ในการวิเคราะห์เสียงพูดของระบบการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูดนี้ ได้แก่ สัมประสิทธิ์การประมาณพันธะเชิงเส้น สัมประสิทธิ์เซปสตรอล และสัมประสิทธิ์เซปสตรอลบนความถี่เมล ในการวิจัยนี้ได้ทดลองปรับเปลี่ยนค่าพารามิเตอร์ของระบบอันได้แก่ อันดับของลักษณะสำคัญ ขนาดของต้นแบบอ้างอิงชุดรหัส และจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟแบบดิสครีต ในงานวิจัยนี้ได้เปรียบเทียบประสิทธิภาพของต้นแบบอ้างอิงชุดรหัสที่สร้างและฝึกฝนจากขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน กับขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด จากผลการทดสอบระบบการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูดสามารถสรุปได้ว่าสัมประสิทธิ์เซปสตรอลบนความถี่เมลเป็นลักษณะสำคัญที่ให้อัตราการบ่งชี้ผู้พูดมากที่สุดเท่ากับร้อยละ 100 ที่อันดับของสัมประสิทธิ์เท่ากับ 14 เมื่อต้นแบบอ้างอิงชุดรหัสมีขนาดเท่ากับ 50 และฝึกฝนด้วยขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด โดยจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟแบบดิสครีตเท่ากับ 5 ในประโยคคำพูดทุกประโยค และสัมประสิทธิ์การประมาณพันธะเชิงเส้นเป็นลักษณะสำคัญที่ให้อัตราการบ่งชี้ผู้พูดน้อยที่สุดเท่ากับร้อยละ 80 85 และ 78.33 ตามลำดับประโยค เมื่ออันดับของสัมประสิทธิ์เท่ากับ 14 ขนาดของต้นแบบอ้างอิงชุดรหัสเท่ากับ 50 และฝึกฝนด้วยขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน โดยที่จำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟแบบดิสครีตเท่ากับ 5

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา วิศวกรรมไฟฟ้า	ลายมือชื่อนิสิต ควิต กาสุริยะ
สาขาวิชา วิศวกรรมไฟฟ้า	ลายมือชื่ออาจารย์ที่ปรึกษา
ปีการศึกษา 2542	ลายมือชื่ออาจารย์ที่ปรึกษาร่วม

4070228421: MAJOR ELECTRICAL ENGINEERING
KEYWORD: SPEAKER RECOGNITION / SPEAKER IDENTIFICATION / EXPECTATION MAXIMIZATION / DISCRETE HIDDEN MARKOV MODELS
SAWIT KASURIYA : SPEAKER RECOGNITION SYSTEM : TEXT-DEPENDENT SPEAKER IDENTIFICATION. THESIS ADVISOR : ASSOC. PROF. SOMCHAI JITAPUNKUL, Dr. Ing. THESIS CO-ADVISOR : ASSIST. PROF. SUDAPORN LUKSANEEYANAVIN, Ph.D. 133 pp. ISBN 974-333-506-4.

The objective of this thesis is to develop a speaker recognition system for text-dependent speaker identification using the speakers' (6 males and 6 females) speech of three Thai sentences. Discrete hidden Markov models and vector quantization are used for this recognition system. Linear prediction coefficients, Cepstral coefficients and Mel frequency cepstral coefficients are evaluated in term of analysis tools. In this research, the order of features, size of codebook reference template and the numbers of discrete hidden Markov model states are varied in order to obtain the optimal system parameters and the best recognition rate. This research also compares the efficiency of codebook of K-means and expectation maximization in the algorithm for the construction of codebook reference template. The results of this text-dependent speaker identification system, can be concluded as follows: the best identification rate is 100 percent, obtained from the 14^{th} order of Mel frequency cepstral coefficients, 50 codebook reference template constructed from the expectation maximization algorithm and the 5 states discrete hidden Markov models for all sentences. And the lowest identification rate is 80, 85 and 78.33 percent in the 3 sentences accordingly. They are obtained from the 14^{th} order of linear prediction coefficients, 50 codebook reference template constructed from K-means algorithm and the 5 states discrete hidden Markov models.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา วิศวกรรมไฟฟ้า ลายมือชื่อนิสิต Sawit K.
สาขาวิชา วิศวกรรมไฟฟ้า ลายมือชื่ออาจารย์ที่ปรึกษา Somchai
ปีการศึกษา 2542 ลายมือชื่ออาจารย์ที่ปรึกษาร่วม Sudaporn L.

กิตติกรรมประกาศ

ในการศึกษา ค้นคว้าการวิจัย และจัดทำวิทยานิพนธ์ฉบับนี้ ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.สมชาย จิตะพันธ์กุล หัวหน้าภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ อาจารย์ที่ปรึกษาที่กรุณาสละเวลาอันมีค่าให้คำแนะนำ เสนอแนะแนวทางและการดำเนินการทำวิจัยนี้เป็นอย่างดีเยี่ยม รวมทั้งการจัดหาอุปกรณ์และทุนในการทำวิจัยตลอดมา และผู้ช่วยศาสตราจารย์ ดร.สุดาพร ลักษณียนาวิน หัวหน้าภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ อาจารย์ที่ปรึกษาร่วมที่กรุณาเสนอแนะ แนะนำให้คำปรึกษาทางด้านภาษาศาสตร์ที่เป็นประโยชน์อย่างมากต่อการทำวิจัย และดูแลการทำวิจัยของผู้ทำวิจัยเป็นอย่างดียิ่ง

ขอขอบคุณ อาจารย์สุวิทย์ นาคพีระยุทธ์ อาจารย์ประจำภาควิชาวิศวกรรมไฟฟ้าและ Mr. Shotaro Akaho นักวิจัยประจำห้องปฏิบัติการ Electrotechnical สถาบันวิจัยของรัฐบาลญี่ปุ่น ที่ให้คำแนะนำและเสนอแนะการแก้ปัญหาในขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด ซึ่งเป็นส่วนหนึ่งของงานวิจัยฉบับนี้เป็นอย่างดีมาโดยตลอด

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดีด้วยความช่วยเหลือจากนายวิศรุต อาชุบุตร นายเอกฤทธิ์ มณีน้อย นางสาวอุมาสี ทาทอง และนายพงษ์ไท ทาสระคู นิสิตปริญญาเอกและโทกลุ่มวิจัยเสียงพูดภาษาไทย ห้องปฏิบัติการวิจัยกรรมวิธีสัญญาณดิจิตอล ที่กรุณาร่วมกันพัฒนาโปรแกรมและเครื่องมือในการวิเคราะห์เสียงพูดที่เป็นประโยชน์อย่างมากต่อการทำวิจัย

ขอขอบคุณพี่ๆ เพื่อนๆ น้องๆ นิสิตวิศวกรรมไฟฟ้า สาขาสื่อสารทุกท่านที่เป็นกำลังใจและให้ความช่วยเหลือตลอดมา และขอบคุณกลุ่มตัวอย่างทุกท่านที่กรุณาสละเวลาและตั้งใจมาบันทึกเสียงพูดเพื่อเป็นข้อมูลในการทำวิจัย

ขอขอบคุณโครงการกลุ่มวิจัยบัณฑิตศึกษาด้านโทรคมนาคมที่สนับสนุนทุนการศึกษาและค่าใช้จ่ายในการนำเสนอผลงานวิจัย

สุดท้ายนี้ ผู้วิจัยใคร่ขอกราบขอบพระคุณบิดา มารดา และขอบคุณน้องชายและผู้มีพระคุณทุกท่านที่ให้การสนับสนุน เอาใจใส่ ดูแล และเร่งรัดผู้วิจัยเสมอมาจนสำเร็จการศึกษา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ฉ

สารบัญ ..... ช

สารบัญรูป ..... ฌ

สารบัญตาราง ..... ฏ

สารบัญคำศัพท์ ..... ณ

บทที่ 1 บทนำ 1

1.1 วัตถุประสงค์ 2

1.2 หลักการและเหตุผล 2

1.3 ปัญหาของการบ่งชี้ผู้พูด 5

1.4 เป้าหมายและขอบเขตของงานวิจัย 5

1.5 ขั้นตอนและวิธีการดำเนินการ 5

1.6 ประโยชน์ที่คาดว่าจะได้รับ 5

บทที่ 2 หลักการ ทฤษฎี และขั้นตอนวิธีการที่เกี่ยวข้อง 6

2.1 การประมวลผลสัญญาณเบื้องต้น 6

2.2 การสกัดลักษณะสำคัญ 8

2.3 การควอนไทซ์แบบเวกเตอร์ 16

2.4 แบบจำลองฮิดเดนมาร์คอฟแบบดิสครีต 25

บทที่ 3 ขั้นตอนวิธีการดำเนินการวิจัย 40

3.1 การเลือกบทคำพูด 40

3.2 การเก็บตัวอย่างข้อมูลสัญญาณเสียงพูด 40

3.3 การสร้างและรูปลักษณ์ของระบบการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูด 41

3.4 ขั้นตอนการฝึกฝนระบบการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูด 45

3.5 ขั้นตอนการทดสอบระบบการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูด 48

บทที่ 4 ผลการวิจัยและวิเคราะห์ผล 50

4.1 พารามิเตอร์และขั้นตอนวิธีการที่สำคัญในระบบการบ่งชี้ผู้พูด 50

4.2 การทดลองปรับเปลี่ยนค่าพารามิเตอร์ต่างๆ 51

4.3 วิเคราะห์ผลการบ่งชี้ผู้พูดที่ผิดพลาด 95

หน้า

บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ 120
5.1 สรุปผลการวิจัย 120
5.2 ข้อเสนอแนะ 122
รายการอ้างอิง 123
ภาคผนวก 126
ประวัติผู้เขียน 133

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญรูป

หน้า

รูปที่ 1.1 การแบ่งระบบการรู้จำผู้พูดตามลักษณะการใช้งาน 1
รูปที่ 1.2 แบบจำลองของระบบการบ่งชี้ผู้พูด 3
รูปที่ 1.3 แผนภาพของระบบการบ่งชี้ผู้พูดโดยใช้แบบจำลองฮิดเดนมาร์คอฟแบบดิสครีต 4
รูปที่ 2.1 ขั้นตอนการประมวลผลสัญญาณเบื้องต้น 6
รูปที่ 2.2 ฟังก์ชันกรอบชนิด Hamming Window 7
รูปที่ 2.3 แบบจำลอง All-Pole ที่ใช้วิเคราะห์หาค่าสัมประสิทธิ์การประมาณพันธะเชิงเส้น 10
รูปที่ 2.4 Mel scale ของความถี่สัญญาณเสียง 14
รูปที่ 2.5 วงจรกรองแบบผ่านแถบความถี่ที่ให้ค่าสัมประสิทธิ์เซปสตรอลบนความถี่เมล 15
รูปที่ 2.6 การควอนไทซ์แบบเวกเตอร์ 16
รูปที่ 2.7 การแบ่งปริภูมิที่มีขนาด 2 มิติ 17
รูปที่ 2.8 การแบ่งปริภูมิที่มีขนาด 2 มิติโดยใช้แบบจำลองการกระจายแบบเกาส์ 22
รูปที่ 2.9 กระบวนการมาร์คอฟที่มี 3 สถานะ 27
รูปที่ 2.10 รายละเอียดของตัวแปรในการคำนวณค่าตัวแปรไปหน้า $\alpha_t(i)$ 32
รูปที่ 2.11 รายละเอียดของตัวแปรในการคำนวณค่าตัวแปรย้อนกลับ $\beta_t(i)$ 33
รูปที่ 2.12 ลำดับของเหตุการณ์ที่ต้องใช้ในการคำนวณเหตุการณ์ร่วมของระบบในสถานะ i ที่เวลา t และสถานะ j ที่เวลา $t+1$ 34
รูปที่ 2.13 ระบบการรู้จำเสียงพูดโดยใช้แบบจำลองฮิดเดนมาร์คอฟ 37
รูปที่ 2.14 แบบจำลองฮิดเดนมาร์คอฟ 3 ประเภท 38
รูปที่ 3.1 ระบบการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูดโดยใช้แบบจำลองฮิดเดนมาร์คอฟแบบดิสครีต 42
รูปที่ 3.2 ช่วงการฝึกฝนชุดรหัสต้นแบบอ้างอิง 46
รูปที่ 3.3 ช่วงการฝึกฝนพารามิเตอร์ของแบบจำลองฮิดเดนมาร์คอฟ 47
รูปที่ 3.4 แบบจำลองฮิดเดนมาร์คอฟ 5 สถานะชนิดซ้ายไปขวา 48
รูปที่ 3.5 ช่วงการทดสอบระบบการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูด 49
รูปที่ 4.1 อัตราการบ่งชี้ผู้พูดในประโยคที่ 1 เมื่อใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 50 61
รูปที่ 4.2 อัตราการบ่งชี้ผู้พูดในประโยคที่ 2 เมื่อใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 50 61
รูปที่ 4.3 อัตราการบ่งชี้ผู้พูดในประโยคที่ 3 เมื่อใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 50 62

หน้า

รูปที่ 4.4 อัตราการบ่งชี้ผู้พูดในประโยคที่ 1 เมื่อใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 100 62
รูปที่ 4.5 อัตราการบ่งชี้ผู้พูดในประโยคที่ 2 เมื่อใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 100 63
รูปที่ 4.6 อัตราการบ่งชี้ผู้พูดในประโยคที่ 3 เมื่อใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 100 63
รูปที่ 4.7 อัตราการบ่งชี้ผู้พูดในประโยคที่ 1 เมื่อใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 150 64
รูปที่ 4.8 อัตราการบ่งชี้ผู้พูดในประโยคที่ 2 เมื่อใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 150 64
รูปที่ 4.9 อัตราการบ่งชี้ผู้พูดในประโยคที่ 3 เมื่อใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 150 65
รูปที่ 4.10 อัตราการบ่งชี้ผู้พูดในประโยคที่ 1 เมื่อใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 50 67
รูปที่ 4.11 อัตราการบ่งชี้ผู้พูดในประโยคที่ 2 เมื่อใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 50 68
รูปที่ 4.12 อัตราการบ่งชี้ผู้พูดในประโยคที่ 3 เมื่อใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 50 68
รูปที่ 4.13 อัตราการบ่งชี้ผู้พูดในประโยคที่ 1 เมื่อใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 100 69
รูปที่ 4.14 อัตราการบ่งชี้ผู้พูดในประโยคที่ 2 เมื่อใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 100 69
รูปที่ 4.15 อัตราการบ่งชี้ผู้พูดในประโยคที่ 3 เมื่อใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 100 70
รูปที่ 4.16 อัตราการบ่งชี้ผู้พูดในประโยคที่ 1 เมื่อใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 150 70
รูปที่ 4.17 อัตราการบ่งชี้ผู้พูดในประโยคที่ 2 เมื่อใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 150 71
รูปที่ 4.18 อัตราการบ่งชี้ผู้พูดในประโยคที่ 3 เมื่อใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่มีขนาดเท่ากับ 150 71
รูปที่ 4.19 อัตราการบ่งชี้ผู้พูดในประโยคที่ 1 จากข้อมูลตารางที่ 4.13 77
รูปที่ 4.20 อัตราการบ่งชี้ผู้พูดในประโยคที่ 2 จากข้อมูลตารางที่ 4.13 78

หน้า

รูปที่ 4.21 อัตราการบ่งชี้ผู้พูดในประโยคที่ 3 จากข้อมูลตารางที่ 4.13 78
รูปที่ 4.22 อัตราการบ่งชี้ผู้พูดในประโยคที่ 1 จากข้อมูลตารางที่ 4.14 79
รูปที่ 4.23 อัตราการบ่งชี้ผู้พูดในประโยคที่ 1 จากข้อมูลตารางที่ 4.14 80
รูปที่ 4.24 อัตราการบ่งชี้ผู้พูดในประโยคที่ 1 จากข้อมูลตารางที่ 4.14 81
รูปที่ 4.25 อัตราการบ่งชี้ผู้พูดของประโยคที่ 1 จากข้อมูลในตารางที่ 4.16 88
รูปที่ 4.26 อัตราการบ่งชี้ผู้พูดของประโยคที่ 2 จากข้อมูลในตารางที่ 4.16 89
รูปที่ 4.27 อัตราการบ่งชี้ผู้พูดของประโยคที่ 3 จากข้อมูลในตารางที่ 4.16 89
รูปที่ 4.28 อัตราการบ่งชี้ผู้พูดของประโยคที่ 1 จากข้อมูลในตารางที่ 4.17 90
รูปที่ 4.29 อัตราการบ่งชี้ผู้พูดของประโยคที่ 2 จากข้อมูลในตารางที่ 4.17 90
รูปที่ 4.30 อัตราการบ่งชี้ผู้พูดของประโยคที่ 3 จากข้อมูลในตารางที่ 4.17 91
รูปที่ 4.31 อัตราการบ่งชี้ผู้พูดของประโยคที่ 1 จากข้อมูลในตารางที่ 4.18 91
รูปที่ 4.32 อัตราการบ่งชี้ผู้พูดของประโยคที่ 2 จากข้อมูลในตารางที่ 4.18 92
รูปที่ 4.33 อัตราการบ่งชี้ผู้พูดของประโยคที่ 3 จากข้อมูลในตารางที่ 4.18 92
รูปที่ 4.34 อัตราการบ่งชี้ผู้พูดของประโยคที่ 1 จากข้อมูลในตารางที่ 4.19 93
รูปที่ 4.35 อัตราการบ่งชี้ผู้พูดของประโยคที่ 2 จากข้อมูลในตารางที่ 4.19 93
รูปที่ 4.36 อัตราการบ่งชี้ผู้พูดของประโยคที่ 3 จากข้อมูลในตารางที่ 4.19 94

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

หน้า

ตารางที่ 4.1 อัตราการบ่งชี้ผู้พูดของระบบที่ใช้สัมประสิทธิ์การประมาณพันธะเชิงเส้นเป็นลักษณะสำคัญ เมื่อจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟเท่ากับ 5 และฝึกฝนต้นแบบอ้างอิงชุดรหัสโดยใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน 53

ตารางที่ 4.2 อัตราการบ่งชี้ผู้พูดของระบบที่ใช้สัมประสิทธิ์เซปสตรอลเป็นลักษณะสำคัญ เมื่อจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟเท่ากับ 5 และฝึกฝนต้นแบบอ้างอิงชุดรหัสโดยใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน 54

ตารางที่ 4.3 อัตราการบ่งชี้ผู้พูดของระบบที่ใช้สัมประสิทธิ์เซปสตรอลบนความถี่เมลเป็นลักษณะสำคัญ เมื่อจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟเท่ากับ 5 และฝึกฝนต้นแบบอ้างอิงชุดรหัสโดยใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน 55

ตารางที่ 4.4 อัตราการบ่งชี้ผู้พูดของระบบที่ใช้สัมประสิทธิ์การประมาณพันธะเชิงเส้นเป็นลักษณะสำคัญ เมื่อจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟเท่ากับ 5 และฝึกฝนต้นแบบอ้างอิงชุดรหัสโดยใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด 57

ตารางที่ 4.5 อัตราการบ่งชี้ผู้พูดของระบบที่ใช้สัมประสิทธิ์เซปสตรอลเป็นลักษณะสำคัญ เมื่อจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟเท่ากับ 5 และฝึกฝนต้นแบบอ้างอิงชุดรหัสโดยใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด 58

ตารางที่ 4.6 อัตราการบ่งชี้ผู้พูดของระบบที่ใช้สัมประสิทธิ์เซปสตรอลบนความถี่เมลเป็นลักษณะสำคัญ เมื่อจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟเท่ากับ 5 และฝึกฝนต้นแบบอ้างอิงชุดรหัสโดยใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด 59

ตารางที่ 4.7 อัตราการบ่งชี้ผู้พูดเฉลี่ยในแต่ละประโยคของสัมประสิทธิ์การประมาณพันธะเชิงเส้น เมื่อใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วนในการฝึกฝนต้นแบบอ้างอิงชุดรหัส 66

ตารางที่ 4.8 อัตราการบ่งชี้ผู้พูดเฉลี่ยในแต่ละประโยคของสัมประสิทธิ์เซปสตรอล เมื่อใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วนในการฝึกฝนต้นแบบอ้างอิงชุดรหัส 66

ตารางที่ 4.9 อัตราการบ่งชี้ผู้พูดเฉลี่ยในแต่ละประโยคของสัมประสิทธิ์เซปสตรอลบนความถี่เมล เมื่อใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วนในการฝึกฝนต้นแบบอ้างอิงชุดรหัส 67

ตารางที่ 4.10 อัตราการบ่งชี้ผู้พูดเฉลี่ยในแต่ละประโยคของสัมประสิทธิ์การประมาณพันธะเชิงเส้น เมื่อใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุดในการฝึกฝนต้นแบบอ้างอิงชุดรหัส 72

ตารางที่ 4.11 อัตราการบ่งชี้ผู้พูดเฉลี่ยในแต่ละประโยคของสัมประสิทธิ์เซปสตรอล เมื่อใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุดในการฝึกฝนต้นแบบอ้างอิงชุดรหัส 73

ตารางที่ 4.12 อัตราการบ่งชี้ผู้พูดเฉลี่ยในแต่ละประโยคของสัมประสิทธิ์เซปสตรอลบนความถี่เมล เมื่อใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุดในการฝึกฝนต้นแบบอ้างอิงชุดรหัส 73

หน้า

ตารางที่ 4.13 อัตราการบ่งชี้ผู้พูดเมื่อจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟเท่ากับ 5 โดยใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วนในการฝึกฝนต้นแบบอ้างอิงชุดรหัส 75
ตารางที่ 4.14 อัตราการบ่งชี้ผู้พูดเมื่อจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟเท่ากับ 5 โดยใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุดในการฝึกฝนต้นแบบอ้างอิงชุดรหัส 76
ตารางที่ 4.15 อัตราการบ่งชี้ผู้พูดเฉลี่ยในการทดลองปรับเปลี่ยนขนาดของต้นแบบอ้างอิงชุดรหัส 82
ตารางที่ 4.16 อัตราการบ่งชี้ผู้พูดเมื่อขนาดของต้นแบบอ้างอิงชุดรหัสเท่ากับ 50 โดยใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วนในการฝึกฝนต้นแบบอ้างอิงชุดรหัส 83
ตารางที่ 4.17 อัตราการบ่งชี้ผู้พูดเมื่อขนาดของต้นแบบอ้างอิงชุดรหัสเท่ากับ 100 โดยใช้ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วนในการฝึกฝนต้นแบบอ้างอิงชุดรหัส 84
ตารางที่ 4.18 อัตราการบ่งชี้ผู้พูดเมื่อขนาดของต้นแบบอ้างอิงชุดรหัสเท่ากับ 50 โดยใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุดในการฝึกฝนต้นแบบอ้างอิงชุดรหัส 86
ตารางที่ 4.19 อัตราการบ่งชี้ผู้พูดเมื่อขนาดของต้นแบบอ้างอิงชุดรหัสเท่ากับ 100 โดยใช้ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุดในการฝึกฝนต้นแบบอ้างอิงชุดรหัส 87
ตารางที่ 4.20 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนอันดับของสัมประสิทธิ์การประมาณพันธะเชิงเส้น และฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน 96
ตารางที่ 4.21 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนอันดับของสัมประสิทธิ์เซปสตรอล และฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน 98
ตารางที่ 4.22 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนอันดับของสัมประสิทธิ์เซปสตรอลบนความถี่เมล และฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน 100
ตารางที่ 4.23 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนอันดับของสัมประสิทธิ์การประมาณพันธะเชิงเส้น และฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด 101
ตารางที่ 4.24 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนอันดับของสัมประสิทธิ์เซปสตรอล และฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด 102
ตารางที่ 4.25 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนอันดับของสัมประสิทธิ์เซปสตรอลบนความถี่เมล และฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด 103
ตารางที่ 4.26 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนขนาดของต้นแบบอ้างอิงชุดรหัสและฝึกฝนด้วย ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน เมื่อใช้สัมประสิทธิ์การประมาณพันธะเชิงเส้น เป็นลักษณะสำคัญ 105
ตารางที่ 4.27 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนขนาดของต้นแบบอ้างอิงชุดรหัสและฝึกฝนด้วย ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน เมื่อใช้สัมประสิทธิ์เซปสตรอลเป็นลักษณะสำคัญ 106

ตารางที่ 4.28 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนขนาดของต้นแบบอ้างอิงชุดรหัสและฝึกฝนด้วยขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน เมื่อใช้สัมประสิทธิ์เซปสตรอลบนความถี่เมล เป็นลักษณะสำคัญ....107
ตารางที่ 4.29 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนขนาดของต้นแบบอ้างอิงชุดรหัสและฝึกฝนด้วยขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด เมื่อใช้สัมประสิทธิ์การประมาณพันธะเชิงเส้น เป็นลักษณะสำคัญ....108
ตารางที่ 4.30 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนขนาดของต้นแบบอ้างอิงชุดรหัสและฝึกฝนด้วยขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด เมื่อใช้สัมประสิทธิ์เซปสตรอลเป็นลักษณะสำคัญ....109
ตารางที่ 4.31 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนขนาดของต้นแบบอ้างอิงชุดรหัสและฝึกฝนด้วยขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด เมื่อใช้สัมประสิทธิ์เซปสตรอลบนความถี่เมล เป็นลักษณะสำคัญ....110
ตารางที่ 4.32 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟ เมื่อใช้สัมประสิทธิ์การประมาณพันธะเชิงเส้นเป็นลักษณะสำคัญและฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน....112
ตารางที่ 4.33 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟ เมื่อใช้สัมประสิทธิ์เซปสตรอลเป็นลักษณะสำคัญและฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน....114
ตารางที่ 4.34 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟ เมื่อใช้สัมประสิทธิ์เซปสตรอลบนความถี่เมลเป็นลักษณะสำคัญและฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน....115
ตารางที่ 4.35 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟ เมื่อใช้สัมประสิทธิ์การประมาณพันธะเชิงเส้นเป็นลักษณะสำคัญและฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด....116
ตารางที่ 4.36 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟ เมื่อใช้สัมประสิทธิ์เซปสตรอลเป็นลักษณะสำคัญและฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด....117
ตารางที่ 4.37 วิเคราะห์เสียงที่บ่งชี้ผิดจากการปรับเปลี่ยนจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟ เมื่อใช้สัมประสิทธิ์เซปสตรอลบนความถี่เมลเป็นลักษณะสำคัญและฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด....118
ตารางที่ 5.1 เวลาที่ใช้ในการประมวลผลจากการทดลองปรับเปลี่ยนค่าพารามิเตอร์ของระบบ....121
ตารางที่ 5.1 เปรียบเทียบเวลาที่ใช้ในการประมวลผลตามกระบวนการต่างๆ ของระบบ....122
ตารางที่ ก.1 รายละเอียดและรูปเสียงของบทคำพูด....125
ตารางที่ ก.2 รายชื่อผู้พูดที่ใช้ในระบบการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูด....125

ตารางที่ ก.3 ตัวอย่างเสียงพูดของผู้พูดกลุ่มอายุตั้งแต่ 18 ถึง 27 ปี.....126
ตารางที่ ก.4 ตัวอย่างเสียงพูดของผู้พูดกลุ่มอายุตั้งแต่ 28 ถึง 37 ปี.....128
ตารางที่ ก.5 ตัวอย่างเสียงพูดของผู้พูดกลุ่มอายุตั้งแต่ 18 ถึง 27 ปี.....130

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญคำศัพท์

หลอดเสียง	acoustic tube
ขั้นตอนวิธีการ	algorithm
อัตสหสัมพันธ์	Autocorrelation
เวกเตอร์จุดศูนย์ถ่วง	Centroid Vector
ชุดรหัส	Codebook
ต้นแบบอ้างอิงชุดรหัส	Codebook Reference Template
เวกเตอร์รหัส	Codeword
แบบจำลองฮิดเดนมาร์คอฟแบบต่อเนื่อง	Continuous Hidden Markov Models
เสียงพูดแบบต่อเนื่อง	continuous speech
ข้อมูลบริบูรณ์	complete data
สัมประสิทธิ์เซปสตรอล	Cepstral Coefficients
ปัญหาการถอดรหัส	decoding problem
มิติ	dimension
แบบจำลองฮิดเดนมาร์คอฟแบบดิสครีต	Discrete Hidden Markov Models
การแจกแจงความน่าจะเป็นขาออกแบบดิสครีต	discrete output probability distribution
ขั้นตอนวิธีการแบบสุนัขกับกระต่าย	Dog Rabbit Algorithm
การเปรียบเทียบทางเวลาแบบพลวัต	dynamic time warping
ช่วงพิสัยพลวัต	dynamic range
กระบวนการแบบพลวัต	dynamic process
ปัญหาการประมาณค่า	estimation problem
ปัญหาการประเมินค่า	evaluation problem
ขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด	expectation maximization algorithm
ค่าความอ่อนล้า	fatigue
ลักษณะสำคัญ	feature
การสกัดลักษณะสำคัญ	feature extraction
ขั้นตอนวิธีการไปหน้าและย้อนกลับ	forward-backward algorithm
ส่วนย่อย, กรอบ	frame
วงจรกรองดิจิตอลอันดับที่หนึ่ง	first-order digital filter
ชุดวงจรกรอง	filter bank
ค่าเหมาะสมที่สุดที่ครอบคลุมโดยรวม	global optimum

ลักษณะสำคัญของเส้นเสียง	glottal feature
การแจกแจงแบบเกาส์	Guassian Distribution
แบบจำลองฮิดเดนมาร์คอฟ	Hidden Markov Models
อัตราการบ่งชี้ผู้พูด	identification rate
ข้อมูลไม่บริบูรณ์	incomplete data
ผู้พูดนอกระบบ	impostor
การแบ่งกลุ่มแบบวนซ้ำ	iterative clustering algorithm
ค่าเหมาะสมที่สุดเฉพาะแห่ง	local optimum
สัมประสิทธิ์การประมาณพันธะเชิงเส้น	linear prediction coefficients
แบบจำลองการประมาณพันธะเชิงเส้น	linear prediction model
ความน่าเป็นจริง	likelihood
ขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน	K-means algorithm
ลูกโซ่มาร์คอฟ	Markov Chain
กระบวนการมาร์คอฟ	Markov Process
คุณสมบัติของมาร์คอฟ	Markov Property
สัมประสิทธิ์เซปสตรอลบนความถี่เมล	Mel Frequency Cepstral Coefficients
ลำดับ	order
แบบจำลองที่ใช้พารามิเตอร์ในการจำลอง	parametric model
การรู้จำแบบรูป	pattern recognition
รูปเสียง	phonetic
การเน้นล่วงหน้า	Preemphasis
ฟังก์ชันความหนาแน่นของความน่าจะเป็น	probability density function
ลักษณะสำคัญของรูปแบบการพูด	prosodic feature
ข้อมูลที่สังเกต	observable data
ลำดับการสังเกต	observation sequence
การควอนไทซ์	Quantization
อัตราการชักตัวอย่าง	sampling rate
แบบจำลองฮิดเดนมาร์คอฟแบบกึ่งต่อเนื่อง	Semi-Continuous Hidden Markov Models
ปริภูมิสัญญาณ	signal space
ข้อมูลเสียงพูด	speech data
สัญญาณเสียงพูด	speech signal
การประมวลผลสัญญาณเบื้องต้น	signal preprocessing
อัตราส่วนสัญญาณต่อสัญญาณรบกวน	signal to noise ratio

การวางกรอบขนาดสัญญาณ	smoothing window
การบ่งชี้ผู้พูด	speaker identification
การรู้จำผู้พูด	speaker recognition
การตรวจสอบผู้พูด	speaker verification
ลำดับของสถานะ	state sequence
ไม่เปลี่ยนแปลงตามเวลา	stationary
แบบจำลองทางสถิติ	statistic model
เงื่อนไขเฟ้นสุ่ม	stochastic constraint
กระบวนการเฟ้นสุ่ม	stochastic process
ลำดับฝึกฝน	training sequence
ขึ้นกับบทคำพูด	text-dependent
ไม่ขึ้นกับบทคำพูด	text-independent
ข้อมูลชุดฝึกฝน	testing data
นอร์มอลไลซ์ทางเวลา	time normalization
จุดเริ่มเปลี่ยน	threshold
ข้อมูลชุดฝึกฝน	training data
ข้อมูลที่ไม่ได้สังเกต	unobservable data
เสียงอโฆษะ	unvoiced
การควอนไทซ์แบบเวกเตอร์	Vector Quantization
ลักษณะสำคัญของช่องทางเดินเสียง	vocal tract feature
เสียงโฆษะ	voiced
เส้นเสียง	vocal cord
สัญญาณรบกวนสีขาว	white noise
ฟังก์ชันกรอบ	window function

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย