

สารบัญ

	หน้า
บทคัดย่อ.....	(3)
ABSTRACT.....	(4)
กิตติกรรมประกาศ.....	(5)
สารบัญ.....	(6)
รายการตาราง.....	(8)
รายการภาพประกอบ.....	(9)
ตัวย่อและสัญลักษณ์.....	(12)
บัญชีคำศัพท์.....	(13)
บทที่	
1. บทนำ.....	1
1.1 บทนำต้นเรื่อง.....	1
1.2 การตรวจเอกสาร.....	2
1.3 วัตถุประสงค์.....	3
1.4 ขอบเขตของการวิจัย.....	4
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	4
2. ทฤษฎีและหลักการ.....	5
2.1 การรู้จำเสียงพูด.....	5
2.2 การสร้างแบบจำลองเสียงด้วยวิธีการของฮิดเดนมาร์คอฟ.....	7
2.3 การจำแนกหน่วยเสียง.....	8
2.4 การเปรียบเทียบผลการรู้จำ.....	9
2.5 การคำนวณค่าลักษณะเด่น.....	10
2.6 แบบจำลองฮิดเดนมาร์คอฟ.....	15
2.7 ขั้นตอนของ HTK ในการสร้างแบบจำลองเสียง.....	18
2.8 ชุดหน่วยเสียง.....	20
2.9 พจนานุกรม (dictionary).....	25
2.10 ไวยากรณ์ (grammar model).....	26

สารบัญ (ต่อ)

	หน้า
3. วิธีการวิจัย.....	29
3.1 ฐานข้อมูลเสียง.....	29
3.2 อุปกรณ์/เครื่องมือ/ซอฟต์แวร์.....	29
3.3 วิธีดำเนินการ.....	30
4. การดำเนินงาน.....	33
4.1 คัดเลือกคำลักษณะเด่น.....	33
4.2 คัดเลือกโครงสร้าง HMM.....	34
4.3 คัดเลือกชุดหน่วยเสียง.....	36
4.4 สร้างและหาค่าความแม่นยำของแบบจำลองเสียงต้นแบบ.....	37
4.5 วิเคราะห์ความสามารถในการรู้จำหน่วยเสียง พยางค์ และคำศัพท์.....	40
4.6 การพัฒนาไวยากรณ์ภาษาเบื้องต้น.....	44
4.7 การพัฒนาโปรแกรมคอมพิวเตอร์สำหรับรู้จำเสียงพูด.....	47
4.8 ทดสอบการรู้จำในสภาพแวดล้อมจริง.....	50
5. บทวิจารณ์.....	53
5.1 ปัญหา.....	53
5.2 แนวทางในการแก้ปัญหา.....	54
6. บทสรุป.....	55
7. บรรณานุกรม.....	56
8. ภาคผนวก.....	58
ก. การสร้างแบบจำลองเสียง.....	58
ข. การทดสอบแบบจำลองเสียง.....	61
ค. ตัวอย่างข้อมูลจากชุดหน่วยเสียง.....	64
ง. เอกสารที่ตีพิมพ์.....	67
9. ประวัติผู้เขียน.....	73

รายการตาราง

ตาราง	หน้า
บทที่ 2	
2.1 สัญลักษณ์หน่วยเสียงทั้งหมดในภาษาไทย	21
2.2 ชุดหน่วยเสียง $C_i+V_i+C_f$	22
2.3 ตัวอย่างรูปแบบของหน่วยเสียงสระวรรณยุกต์ V_i	22
2.4 หน่วยเสียงชุดตัวสะกดวรรณยุกต์ $C_i+V_i+C_n$	23
2.5 ชุดหน่วยเสียงภาษาไทยแบบ 41 หน่วยเสียง	23
2.6 ผลการรู้จำของหน่วยเสียงแบบจดจำเสียงรอบข้าง	24
2.7 ตัวอย่างพจนานุกรมคำศัพท์	25
2.8 ตัวอย่างพจนานุกรมหน่วยเสียง	26
2.9 ตัวอย่างพจนานุกรมพยางค์	26
บทที่ 4	
4.1 วิธีการปรับชุดหน่วยเสียง $C_i+V_i+C_n$ ให้ตรงกับชุดหน่วยเสียง $C_i+V_i+C_f$	36
4.2 แสดงค่าความแม่นยำของชุดหน่วยเสียง (%)	36
4.3 แสดงค่าความแม่นยำการจำแนกกลุ่มหน่วยเสียง	38
4.4 แสดงความสับสนในการจำแนกหน่วยเสียงกลุ่มพยัญชนะ	38
4.5 แสดงความสับสนในการจำแนกหน่วยเสียงสระ อะ และ สระ อา	38
4.6 แสดงความสับสนในการจำแนกหน่วยเสียงตัวสะกด	39
4.7 ตัวอย่างพจนานุกรมพยางค์	41
4.8 ตัวอย่างพจนานุกรมคำศัพท์	42
4.9 ตัวอย่างพจนานุกรมคำศัพท์ที่ใช้ในการทดลอง	43
4.10 แสดงค่าความแม่นยำของการรู้จำไวยากรณ์ภาษารูปแบบต่างๆ	46
4.11 พจนานุกรมไทยสำหรับ โปรแกรมรู้จำ	49

รายการภาพประกอบ

ภาพประกอบ	หน้า
บทที่ 1	
1.1 ระบบรู้จำเสียงพูดกับการประยุกต์ใช้ในงานต่าง ๆ	1
บทที่ 2	
2.1 หลักการทำงานของระบบรู้จำเสียงพูด	5
2.2 กระบวนการรู้จำเสียงพูด	6
2.3 ขั้นตอนการสร้างแบบจำลองเสียง	7
2.4 ขั้นตอนการจำแนกหน่วยเสียง	8
2.5 การเปรียบเทียบผลการรู้จำเสียงพูดและข้อมูลกำกับเสียง	9
2.6 ขั้นตอนการคำนวณหาค่าลักษณะเด่นแบบ MFCC	11
2.7 ขอบเขตช่วงของความถี่ของเมล	11
2.8 ตัวอย่างค่าลักษณะเด่นของเสียงในช่วงหนึ่ง	12
2.9 ความแตกต่างของสัญญาณเสียง หนึ่ง หนึ่ง สอง ตามลำดับ	13
2.10 ความแตกต่างของค่าลักษณะเด่นของเสียง หนึ่ง หนึ่ง สอง ตามลำดับ	14
2.11 ค่าพิตช์ของวรรณยุกต์ในภาษาไทย 5 ระดับวรรณยุกต์	15
2.12 หลักการทำงานของแบบจำลองเสียงแบบฮิดเดนมาร์คอฟ	16
2.13 ขั้นตอนภายในของการสร้างแบบจำลองเสียง	16
2.14 ตัวอย่างข้อมูลของชุดตัวแปร a_{ij} และ b_{jk}	16
2.15 การทำงานของแบบจำลองเสียงเพื่อรู้จำเสียงไม่ทราบค่า	17
2.16 ขั้นตอนการกำหนดค่าเริ่มต้นให้กับแบบจำลองเสียง	18
2.17 การสร้างแบบจำลองเสียงในขั้นตอน Single Re-estimation	19
2.18 ส่วนประกอบของข้อความในภาษาไทย	20
2.19 ตัวอย่างไวยากรณ์สำหรับการรู้จำตัวเลข	26
2.20 ตัวอย่างไวยากรณ์สำหรับการต่อโทรศัพท์	27
2.21 ตัวอย่างไวยากรณ์สำหรับการรู้จำหมายเลขโทรศัพท์แบบตัวเลข 9 ตัว	28

รายการภาพประกอบ (ต่อ)

ภาพประกอบ	หน้า
บทที่ 4	
4.1 กราฟเปรียบเทียบความแม่นยำของค่าลักษณะเด่นแบบต่างๆ	34
4.2 กราฟเปรียบเทียบความแม่นยำของชุด โครงสร้าง HMM	35
4.3 แผนภาพแสดงสถิติในการจำแนก สระ อะ เสียงสามัญ	39
4.4 รูปแบบการรู้จำหน่วยเสียง	40
4.5 รูปแบบการรู้จำหน่วยเสียงตามลำดับของพยางค์	41
4.6 รูปแบบการรู้จำพยางค์	41
4.7 รูปแบบการรู้จำคำศัพท์	42
4.8 ความแม่นยำในการรู้จำของแบบจำลองเสียงต้นแบบ	43
4.9 ไวยากรณ์สำหรับการรู้จำประโยคบอกเล่าพื้นฐาน	45
4.10 ไวยากรณ์สำหรับการรู้จำประโยคบอกเล่าซับซ้อน	45
4.11 ไวยากรณ์สำหรับการรู้จำคำศัพท์ทั่วไป	45
4.12 ผลการรู้จำของไวยากรณ์ประโยคบอกเล่าพื้นฐาน	46
4.13 ลักษณะการทำงานของโปรแกรมรู้จำ	47
4.14 วิธีการใช้งานชุดโปรแกรม HTK	48
4.15 ขั้นตอนการทำงานของโปรแกรมรู้จำ	49
4.16 หน้าจอโปรแกรมรู้จำที่ใช้แทนเครื่องคิดเลข	50
4.17 ผลการรู้จำจากระบบรู้จำเสียงพูดต้นแบบ	51
4.18 ลักษณะเสียงที่บันทึกภายในห้องบันทึกเสียง	52
4.19 ลักษณะเสียงที่บันทึกภายนอกห้องบันทึกเสียง	52
ภาคผนวก ก การสร้างแบบจำลองเสียง	
ก1 โครงสร้างข้อมูลในการสร้างแบบจำลองเสียง	58
ก2 ไฟล์ข้อมูลคอนฟิกของการสร้างแบบจำลองเสียง	58
ก3 ที่เก็บแบบจำลองเสียงในแต่ละขั้นตอน	59

รายการภาพประกอบ (ต่อ)

ภาพประกอบ	หน้า
ก4 ไฟล์ที่ผู้ใช้ต้องกำหนดเพื่อสร้างแบบจำลองเสียง	59
ก5 ไฟล์ข้อมูลกำกับเสียงในการสร้างแบบจำลองเสียง	59
ภาคผนวก ข การทดสอบการรู้จำ	
ข1 โครงสร้างหลักของระบบการทดสอบการรู้จำ	61
ข2 ไฟล์ข้อมูลคอนฟิกของการทดสอบการรู้จำ	61
ข3 ไฟล์ที่ผู้ใช้ต้องกำหนดเพื่อทดสอบการรู้จำ	62
ข4 ผลการรู้จำจากการทดสอบการรู้จำ	62
ข5 ไฟล์ข้อมูลกำกับเสียงในการทดสอบการรู้จำ	63
ภาคผนวก ค ตัวอย่างข้อมูลจากชุดหน่วยเสียง $C_i+V_i+C_r$	
ค1 ข้อมูลภายในไฟล์โครงสร้าง HMM (protomfc.txt)	64
ค2 ข้อมูลภายในไฟล์ที่เก็บข้อมูลฝึกฝน (mfccTrain.list)	64
ค3 ข้อมูลภายในไฟล์แบบจำลองเสียง (newmacros)	64
ค4 ข้อมูลภายในไฟล์รายชื่อหน่วยเสียงของชุดหน่วยเสียง $C_i+V_i+C_r$ (hmm.list)	65
ค5 ข้อมูลภายในไฟล์พจนานุกรมของชุดหน่วยเสียง $C_i+V_i+C_r$ (hmm.dict)	65
ค6 ข้อมูลภายในไฟล์ไวยากรณ์ของชุดหน่วยเสียง $C_i+V_i+C_r$ (hmm.gram)	65
ค7 ข้อมูลภายในไฟล์ผลการรู้จำแบบที่ 1 (resultTest1.list)	66
ค8 ข้อมูลภายในไฟล์ผลการรู้จำแบบที่ 2 (resultTest2.list)	66

ตัวย่อและสัญลักษณ์

- $C_i+V_i+C_f$ = ชุดหน่วยเสียง พยัญชนะ สระวรรณยุกต์ ตัวสะกด
- $C_i+V_i+C_n$ = ชุดหน่วยเสียง พยัญชนะ สระวรรณยุกต์ ตัวสะกดวรรณยุกต์
- LPC = Linear Prediction Coefficient
- LPC+D = LPC เพิ่มค่าอนุพันธ์อันดับที่ 1
- LPC+DA = LPC เพิ่มค่าอนุพันธ์อันดับที่ 1 และ 2
- MFCC = Mel Scale Frequency Cepstral Coefficient
- MFCC+D = MFCC เพิ่มค่าอนุพันธ์อันดับที่ 1
- MFCC+DA = MFCC เพิ่มค่าอนุพันธ์อันดับที่ 1 และ 2

บัญชีคำศัพท์

ข้อมูลกำกับเสียง (Transcription Label)	= ข้อมูลที่ระบุว่าเสียงในช่วงที่กำหนดเป็นเสียงใด
ข้อมูลทดสอบ (Testing Speech)	= ข้อมูลเสียงไม่ทราบค่า ใช้ในการจำแนกหน่วยเสียง และการรู้จำเสียงพูด
ข้อมูลฝึกฝน (Training Speech)	= ข้อมูลเสียงที่ใช้ในการสอนระบบรู้จำ ใช้คู่กับข้อมูลกำกับเสียง
ความถี่ในการสุ่มสัญญาณ (Sampling Rate)	= รูปแบบการบันทึกข้อมูลเสียงมีหน่วยเป็น Hz
ความละเอียดของข้อมูล (Resolution)	= รูปแบบการบันทึกข้อมูลเสียงมีหน่วยเป็น Bit
ค่าลักษณะเด่น (Feature Vector)	= ค่าที่แสดงถึงลักษณะเฉพาะของเสียงแต่ละเสียง
มิกซ์เจอร์ (Mixture)	= ส่วนประกอบภายในของโครงสร้างของ HMM ของแต่ละสถานะ
ตารางแสดงความสับสน (Confusion matrix)	= ค่าสถิติแสดงการเปรียบเทียบระหว่างผลการรู้จำที่ได้กับค่าที่ถูกต้อง
แบบจำลองเสียง (Phonetic Model)	= ฐานข้อมูลหน่วยเสียงเพื่อใช้ในการรู้จำเสียงพูด
โปรแกรมรู้จำ	= โปรแกรมที่พัฒนาขึ้นเพื่ออำนวยความสะดวกในการใช้งานระบบรู้จำ
พจนานุกรม (Dictionary)	= รายการคำศัพท์ควบคู่กับลำดับของหน่วยเสียง ที่รู้จำได้
พจนานุกรมไทย	= พจนานุกรมของโปรแกรมรู้จำ ใช้ในการแสดงผลภาษาไทย
พารามิเตอร์ (Parameter)	= ตัวแปรของระบบที่ต้องมีการกำหนดค่าก่อน
ระบบรู้จำเสียงพูดต้นแบบ	= ระบบรู้จำเสียงพูดในงานวิจัย สร้างจากส่วนประกอบคัดเลือกแล้ว
รูปเสียง	= ข้อมูลฝั่งซ้ายของพจนานุกรม แสดงรายชื่อคำศัพท์ที่เป็นผลการรู้จำ
ลู่เข้า (Converged)	= รูปแบบข้อมูลที่มีช่วงการเปลี่ยนแปลงลดลงเรื่อยๆ
ไวยากรณ์ (Grammar Model)	= กฎของการรู้จำหรือรูปแบบการจัดเรียงคำศัพท์ที่จะเป็นผลการรู้จำ

บัญชีคำศัพท์ (ต่อ)

- สถานะ (State)** = ส่วนประกอบภายในของโครงสร้างของ HMM ของแต่ละหน่วยเสียง
- สัมประสิทธิ์เคปสตรัมความถี่เมล (M²-l Scale Frequency Cepstral Coefficient)** = วิธีการคำนวณค่าลักษณะเด่นที่ใช้ในระบบรู้จำเสียงพูด
- เสียงเงียบ (Silence)** = เสียงในส่วนที่มีแอมพลิจูดเป็น 0 หรือเข้าใกล้ 0