

การตรวจรู้เสียงพูดภาษาไทย วิชาใช้หน่วยพยางค์



นายทวี ประทุมทาน

ศูนย์วิจัยทรัพยากร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

ภาควิชาวิศวกรรมคอมพิวเตอร์

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

พ. ศ. 2530


ISBN 974-568-196-2

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

012685

I10299440

THAI SPEECH RECOGNITION USING SYLLABLE UNIT



Mr. TAWI PRATHUMTHAN

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

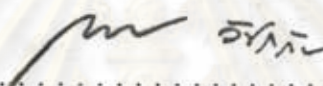
A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science
Department of Computer Engineering
Graduate School
Chulalongkorn University

1987


หัวข้อวิทยานิพนธ์ การตรวจรู้เสียงพูดภาษาไทย โดยใช้หน่วยพยางค์
โดย นาย ทวี ประทุมทาน
ภาควิชา วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษา รองศาสตราจารย์ สมชาย ทยานง
อาจารย์ที่ปรึกษาร่วม ผู้ช่วยศาสตราจารย์ ดร. วีระ ธีวพิทักษ์




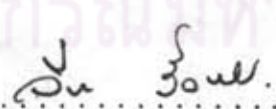
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัยนี้เป็น
ส่วนหนึ่งของการศึกษาคำหลักสูตรบริวัฒนามหาบัณฑิต

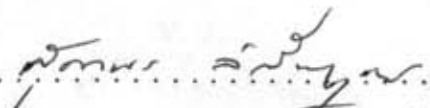

..... คณบดีบัณฑิตวิทยาลัย
(ศาสตราจารย์ ดร.ถาวร วัชรากษ์)

คณะกรรมการสอบวิทยานิพนธ์


..... ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ วันชัย ธีวเพบูลย์)


..... กรรมการ
(รองศาสตราจารย์ สมชาย ทยานง)


..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. วีระ ธีวพิทักษ์)


..... กรรมการ
(ดร.สุตาพร ลักษณินาวิน)

หัวข้อวิทยานิพนธ์	การตรวจรู้เสียงพูดภาษาไทย ภายใต้น่วยพยางค์
ชื่อนิสิต	นายทวี ประทุมทาน
อาจารย์ที่ปรึกษา	รองศาสตราจารย์ สมชาย พยานง
อาจารย์ที่ปรึกษาร่วม	ผู้ช่วยศาสตราจารย์ ดร. วีระ ธีวพิทักษ์
ภาควิชา	วิศวกรรมคอมพิวเตอร์
ปีการศึกษา	2529



บทคัดย่อ

การตรวจรู้เสียงพูดภาษาไทย ภายใต้น่วยพยางค์ เป็นการวิจัย เพื่อสร้างระบบการตรวจรู้เสียงพูดภาษาไทยในประโยคต่อเนื่อง ภายใต้น่วยเสียง พยางค์เป็นหน่วยเสียงที่จะใช้ในการเปรียบเทียบ ระบบที่ได้สร้างขึ้นนี้ แบ่งเป็น 3 ชั้นคอนใหญ่ ๆ คือ ชั้นคอนการวิเคราะห์หาค่าพารามิเตอร์, ชั้นคอนการแยกพยางค์ และชั้นคอนการนำพยางค์ที่แยกได้มาตรวจรู้

ในชั้นคอนการวิเคราะห์หาค่าพารามิเตอร์ ได้ใช้เทคนิคการวิเคราะห์แบบ การพยากรณ์เชิงเส้น (Linear Prediction) เพื่อหาค่าพารามิเตอร์ของเสียงพูด ซึ่งได้แก่ ค่าสัมประสิทธิ์ของการสะท้อนกลับ (Reflection Coefficient) , พลังงาน α และใช้เทคนิคของ SIFT (Simplified Inverse Filter Tracking) สำหรับหาค่าความถี่มูลฐาน (F_0 : fundamental frequency)

ชั้นคอนที่สอง เป็นการนำค่าพารามิเตอร์ ความถี่มูลฐาน และพลังงาน มาใช้ในการตัดพยางค์ ภายใต้อะไร่ระยะเวลาประกอบการศึกษาด้วย ซึ่งจากผลการทดลองโดยพูดในสภาวะแวดล้อมที่เงียบ และผู้พูดได้รับการฝึกฝนพอสมควร ภายมีความเร็วในการพูดระดับปานกลางแล้วได้ผลถูกต้องในการตัดพยางค์ 93.37 %

ชั้นคอนที่สาม เป็นการนำขอบเขตพยางค์ที่ตัดได้จากชั้นคอนที่สอง และค่าพารามิเตอร์ สัมประสิทธิ์การสะท้อนกลับ ที่ได้จากชั้นคอนที่หนึ่งของแต่ละพยางค์ มาทำการเปรียบเทียบหาค่าพารามิเตอร์ของพยางค์ที่เป็นต้นแบบ และแสดงผลการเปรียบเทียบออกมา ภายชั้นคอนการเปรียบเทียบนี้ ได้ใช้เทคนิค ไลเคนาไมคโปรแกรม

มิ่ง(Dynamic Programming) และจากผลการทดลองให้เปรียบเทียบเสียงพูด
เลขศูนย์ถึงเลขเก้า ปรากฏว่าได้ผลถูกต้องในการตรวจรู้ 96.87 %



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Thesis Title Thai Speech Recognition using Syllable Unit
Name Mr. Tawi Prathumthan
Thesis Advisor Associate Professor Somchai Thayanyong
Assistant Professor Dr. Weera Riewpituk
Department Computer Engineering
Academic Year 1986



ABSTRACT

Thai speech recognition using syllable unit is the system constructed for the recognition of Thai continuous speech. The system consists of 3 stages use.

- 1) The analysis of speech in terms of its overall energy, reflection coefficient, fundamental frequency.
- 2) The segmentation of syllables in continuous speech.
- 3) The recognition of segmented syllables.

In the analysis of speech, the techniques used are Linear Prediction Coding(LPC) and Simplified Inverse Filter Tracking(SIFT). LPC is used for the analysis of overall energy & reflection coefficient. SIFT is used for the analysis of pitch period.

The results from the analyses from the above techniques i.c. energy pattern, pitch period, duration are used for the segmentation of syllables in connected speech. The experiment for the segmentation speech is run with 3 subjects who are trained to speak with medium

speed. The percentage of accurate segmentation is 93.37.

For the recognition of segmented syllables reflection coefficient between syllable boundaries of the segmented syllables is used. The technique used in this stage is the dynamic programming.

The system constructed here is used for an experiment the recognition of Thai digits in connected speech. The percentage of accurate recognition is 96.87.



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



กิตติกรรมประกาศ

การทำวิทยานิพนธ์ครั้งนี้ ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ สมชาย ทยานิช และ รองศาสตราจารย์ ดร. วีระ ธีวพิทักษ์ เป็นอย่างยิ่งที่ได้กรุณาแนะนำหัวข้อวิทยานิพนธ์ ตลอดจนแนวทางในการวิจัย ให้คำปรึกษา และจัดหาเงินทุนในการทำวิจัย และอาจารย์ ดร. สุภาพร ลักษณะนิยนาวัน ที่ได้กรุณาแนะนำและให้คำปรึกษาเกี่ยวกับเรื่อง เสียงทางภาษาศาสตร์ จึงทำให้การวิจัยครั้งนี้สำเร็จลงได้ ตลอดจนคณาจารย์ทุกท่านในภาควิชาวิศวกรรมคอมพิวเตอร์ ที่ได้อบรมสั่งสอนความรู้ให้แก่ผู้วิจัย ผู้วิจัยต้องขอกราบขอบพระคุณมา ณ ที่นี้ด้วย นอกจากนี้แล้วผู้เขียนยังรู้สึกซาบซึ้งต่อน้ำใจของเพื่อนหลาย ๆ ท่าน ที่ได้ช่วยให้กำลังใจในการทำวิทยานิพนธ์มาตลอด โดยเฉพาะคุณเอิสรพล แก้วผลึก, คุณเชิดชัย เสริมคำรงค์ศักดิ์, คุณอาจิน จิรชีพพัฒนา ที่ได้ให้กำลังใจในการทำวิจัยครั้งนี้มาโดยตลอด ซึ่งผู้เขียนก็ต้องขอขอบคุณมา ณ โอกาสนี้ด้วย นอกจากนี้ต้องขอขอบคุณ คุณ อานันท์นิยะกุล ที่ช่วยเป็นธุระในการประกอบวงจร A/D ให้ และที่จะลืมเสียไม่ได้ก็คือคุณนพวรรณ เสริมคำรงค์ศักดิ์ และคุณรุ่ง เพ็ชร คันตีสุข ที่กรุณามาทูกเป็นตัวอย่างเสียงในการวิจัยครั้งนี้ด้วย รวมทั้งคุณบุศริน เจริญวัฒนคุณย์ ที่ได้กรุณามาช่วยพิมพ์งานวิจัย

ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ



	หน้า
บทคัดย่อภาษาไทย.....	ง.
บทคัดย่อภาษาอังกฤษ.....	ฉ.
กิตติกรรมประกาศ.....	ช.
รายการตารางประกอบ.....	ณ.
รายการรูปประกอบ.....	ด.
บทที่ 1. บทนำ.....	1
1.1 ความเป็นมาของปัญหา.....	1
1.2 วัตถุประสงค์.....	3
1.3 ประโยชน์ที่จะได้จากการวิจัย.....	4
1.4 ขอบเขตของการวิจัย.....	4
1.5 ขั้นตอนและวิธีการดำเนินการวิจัย.....	5
บทที่ 2. ประวัตติการค้นคว้าเกี่ยวกับเรื่องเสียง.....	6
2.1 ประวัตติการค้นคว้าเกี่ยวกับระบบตรวจรู้เสียง.....	6
2.2 ประวัตติการค้นคว้าเกี่ยวกับระบบสัทสังเคราะห์.....	8
บทที่ 3. หน่วยเสียงในภาษาไทย.....	12
3.1 หน่วยเสียงพยัญชนะ.....	12
3.2 หน่วยเสียงสระ.....	15
3.3 หน่วยเสียงวรรณยุกต์.....	19
บทที่ 4. พยางค์.....	20
4.1 พยางค์ในความหมายกว้าง.....	20
4.2 พยางค์ในภาษาไทย.....	24
บทที่ 5. เสียงพูด.....	28
5.1 อวัยวะที่ใช้ในการเปล่งเสียง.....	28
5.2 แบบจำลองของเสียงพูด.....	31
5.3 แบบจำลองของเสียงพูดตามวิธี Linear Prediction.....	34
5.4 วิธีในการหาค่าพารามิเตอร์.....	42

	หน้า
5.5 วิธีในการหาค่าความถี่มูลฐาน.....	44
บทที่ 6. ระบบในการคัดพยางค์และตรวจรู้.....	45
6.1 ผังงานของระบบ.....	45
6.1.1 ผังงานของระบบการวิเคราะห์.....	45
6.1.2 ผังงานของระบบการคัดพยางค์.....	46
6.1.3 ผังงานของระบบการตรวจรู้.....	47
6.2 วิธีการในการคัดพยางค์.....	48
6.3 วิธีในการตรวจรู้.....	52
บทที่ 7. สรุปผลการวิจัยและข้อ เสนอแนะ.....	64
7.1 สรุปผลการวิจัย.....	64
7.2 อภิปราย.....	65
7.3 ข้อ เสนอแนะ.....	66
บรรณานุกรม.....	68
ภาคผนวก	
ก. ผังงานการหาค่าพารามิเตอร์ , สัมประสิทธิ์การสะท้อนกลับ..	71
ข. ผังงานการหาค่าความถี่มูลฐาน.....	75
ค. ประโยคตัวอย่างที่ใช้ในการคัดพยางค์.....	81
ง. ผลของการคัดพยางค์.....	83
จ. ประโยคตัวอย่างที่ใช้ในการคัดพยางค์และตรวจรู้.....	116
ฉ. ผลของการตรวจรู้.....	117
ช. วงจร A/D.....	126
ซ. ตัวอย่าง เสียงพูด.....	128
ณ. รายละเอียดโปรแกรม.....	134
ประวัติผู้เขียน.....	220

รายการตารางประกอบ

ตารางที่	หน้า
1. หน่วยเสียงพยัญชนะในภาษาไทย.....	12
2. เสียงพยัญชนะในภาษาไทยที่สามารถควบกล้ำได้.....	14
3. หน่วยเสียงสระในภาษาไทย.....	15
4. แสดงอิตรากการ เกิดของหน่วยเสียงสระและหน่วยเสียง พยัญชนะในภาษาไทย.....	18
5. เสียงวรรณยุกต์ที่เกิดขึ้นได้กับพยางค์ในภาษาไทย.....	26
6. องค์ประกอบของพยางค์ที่เป็นไปได้ในภาษาไทย.....	27
7. แสดงผลในการตัดพยางค์.....	60
8. แสดงผลในการตัดพยางค์เพื่อใช้ในการตรวจรู้.....	62
9. แสดงผลในการตรวจรู้.....	63



คุุขยัวิทยัทรัพยักร
จุฬาลงกรณัมหัวิทยัาลัย

รายการรูปประกอบ

รูปที่	หน้า
1. ระบบการติดต่อกับเครื่องฉายใช้เสียงพูด.....	2
2. ระบบการสร้างเสียงของ Wheatstone.....	10
3. ระบบการสร้างเสียงของ J.Q. Stewart.....	10
4. ระบบการสร้างเสียงของ RR. Riesz.....	11
5. ระบบการสร้างเสียงของ RR. Riesz, H. W. Dudley, S. A. Watkins.....	11
6. หน่วยเสียงสระผสมในภาษาไทย.....	15
7. ช่วงความถี่ของ เสียงวรรณยุกต์ในภาษาไทย.....	19
8. แสดงลักษณะการ เกร็งของกล้ามเนื้อ.....	22
9. แสดงลักษณะช่องว่างภายในปาก.....	29
10. แสดงการ เคลื่อนไหวของ เส้นเสียง.....	29
11. อวัยวะที่ใช้ในการออกเสียง.....	30
12. แบบจำลองการ เกิดเสียง.....	31
13.1 ลักษณะของ เสียงพูด.....	34
13.2 ลักษณะเสียงพูดที่ใช้เทคนิค PARCOR.....	38
14. แสดงถึงหลักการของ inner product.....	39
15. แสดงหลักการทางานของ SIFT.....	44
16. แสดงการ คัดพยางค์.....	49
17. แสดงทาง เเคินของ c ในกรณีทีพารามิเตอร์หึ่งสองชุด ไม่มีความแตกต่างด้านเวลา.....	53
18. แสดงทาง เเคินของ c ในกรณีทีพารามิเตอร์หึ่งสองชุด มีความแตกต่างด้านเวลา.....	54
19. แสดงการด้วงน้ำหนักแบบสมมาตร.....	55
20. แสดงการด้วงน้ำหนักแบบไม่สมมาตร.....	55