สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การประยุกต์ใช้เวฟเลทแยกเสียงพยัญชนะ เสียงสระ ภาษาไทย

APPLICATION OF WAVELET TRANSFORMS FOR CONSONANT VOWEL SEGMENTATION ON THAI SPEECH SIGNAL

T 063676

นริศ บุญศักดิ์เฉลิม

NARIS BUNSAKCHALERM

QW.
น251ก
2549

เลขหมู่.....
เลขทะเบียน 63676
วัน,เดือน,ปี 30 ส.ค. 2549

b 11039578
i.....

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมโทรคมนาคม

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2549

ISBN 974-15-2633-4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

APPLICATION OF WAVELET TRANSFORMS FOR CONSONANT VOWEL SEGMENTATION ON THAI SPEECH SIGNAL

NARIS BUNSAKCHALERM

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ENGINEERING IN TELECOMMUNICATION ENGINEERING

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2006

ISBN 974-15-2633-4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2006

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์ การประยุกต์ใช้เวฟเลทแยกเสียงพยัญชนะ เสียงสระ ภาษาไทย
นักศึกษา นายนริศ บุญศักดิ์เฉลิม
รหัสประจำตัว 44061736
ปริญญา วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา วิศวกรรมโทรคมนาคม
พ.ศ. 2549
อาจารย์ผู้ควบคุมวิทยานิพนธ์ รศ. ดร. ไกรสิน ส่งวัฒนา

บทคัดย่อ

วิทยานิพนธ์นี้นำเสนอวิธีการแยกเสียงพยัญชนะและเสียงสระสำหรับเสียงพูดในภาษาไทย โครงสร้างของเสียงในภาษาไทย เกิดจากการรวมกันของพยัญชนะแล้วตามด้วยสระ การแยกพยัญชนะเป็นส่วนสำคัญในระบบการรู้จำเสียงในภาษาไทย การแปลงเวฟเลทเป็นวิธีที่นำเสนอการค้นหาตำแหน่ง การแยกพยัญชนะและสระโดยตรงโดยใช้สมการ ผลคูณและรูปแบบของพลังงาน สมการผลคูณได้จากการประมาณค่าเวฟเลท และ รายละเอียดของสัญญาณเสียงพูดที่เข้ามา และสามารถนำมาใช้บ่งชี้ตำแหน่งของพยัญชนะและสระ สมการผลคูณและการตรวจสอบรูปแบบของพลังงาน การแยกพยัญชนะและสระสามารถแยกตำแหน่งได้อย่างเหมาะสมด้วยการคำนวณที่ไม่ซับซ้อน การทดลองใช้ระยะเวลาน้อย ไม่สิ้นเปลืองในการคำนวณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	APPLICATION OF WAVELET TRANSFORMS FOR CONSONANT VOWEL SEGMENTATION ON THAI SPEECH SIGNAL
Student	Mr. Naris Bunsakchalerm
Student ID	44061736
Degree	Master of Engineering
Programme	telecommunication Engineering
Year	2006
Thesis Advisor	Assoc. Prof. Dr. Kraisin Songwatana

ABSTRACT

This thesis proposes Consonant/Vowel (C/V) segmentation algorithm for Thai speech signal. Since the Thai phoneme structure is a combination of a consonant followed by a vowel ,the C/V segmentation is an important part in the Thai speech recognition system. Based on the wavelet transforms, the proposed method can directly search for the C/V segmentation point by using a product function and energy profile. The product function is generated from the appropriate approximation wavelet and detail of input speech signal, and it can be applied to indicate the C/V segmentation point. With this product function and the additional verification of energy profile, the C/V segmentation can be accurately pointed out with a low computation complexity. Experiment result demonstrates the performance of the proposed algorithm .

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี ด้วยความกรุณาของ รศ.ดร. ไกรสิน ส่งวัฒนา อาจารย์ที่ปรึกษา ที่กรุณาแนะนำ ให้คำปรึกษาในการวิจัยครั้งนี้ และให้ความเมตตาแก่ศิษย์ด้วยดีเสมอมา ผู้เขียนรู้สึกซาบซึ้งในความกรุณาเป็นอย่างยิ่ง และขอกราบขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณอาจารย์ทุกท่านที่ได้ประสิทธิ์ประสาทวิชาความรู้ต่างๆให้คำแนะนำปรึกษาที่ดีมาโดยตลอดให้แก่ศิษย์

ขอขอบคุณ พี่ๆ เพื่อนๆ และน้องๆ ทุกท่านทั้งที่ห้องวิจัย ที่ทำงาน และที่เรียนด้วยกัน สำหรับความเป็นมิตร และความช่วยเหลือต่างๆ ที่มีให้กันมาโดยตลอด

ขอขอบคุณเจ้าของเสียงทุกๆ ท่าน ที่ได้เสียสละเวลามาบันทึกเสียงให้

ขอกราบขอบพระคุณ บิดา มารดา และครอบครัว ที่ให้ทุกสิ่งทุกอย่าง และที่สำคัญกำลังใจแก่ลูกคนนี้

สุดท้ายขอขอบคุณบัณฑิตวิทยาลัย ที่ได้ให้ทุนสนับสนุนการทำวิทยานิพนธ์ครั้งนี้

ประโยชน์อันพึงมีจากวิทยานิพนธ์ฉบับนี้ ผู้เขียนขอมอบแด่ผู้มีพระคุณทุกท่าน

นริศ บุญศักดิ์เฉลิม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย I

บทคัดย่อภาษาอังกฤษ..... II

กิตติกรรมประกาศ..... III

สารบัญ..... IV

สารบัญตาราง..... VIII

สารบัญรูป..... IX

บทที่ 1 บทนำ..... 1

1.1 ความเป็นมาและความสำคัญของปัญหา..... 1

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา..... 2

1.3 สมมติฐานของการศึกษา..... 3

1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย..... 3

1.5 ขอบเขตการวิจัย..... 3

1.6 ขั้นตอนการศึกษา..... 4

1.7 โครงประกอบของวิทยานิพนธ์..... 4

บทที่ 2 ทฤษฎีการสร้างเสียงพูดในภาษาไทย..... 5

2.1 กล่าวนำ..... 5

2.2 ทฤษฎีภาษาศาสตร์..... 5

2.2.1 อวัยวะออกเสียง..... 6

2.2.2 กระบวนการผลิตเสียงพูด..... 9

2.2.3 เสียงพูดของมนุษย์..... 10

2.2.4 ความยาวของเสียง..... 11

2.2.5 ระดับเสียงสูง-ต่ำ..... 11

2.2.6 ความดัง..... 11

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.2.7 การลงน้ำหนัก..... 11

2.2.8 ช่วงต่อของเสียง..... 11

2.3 หน่วยเสียงสำคัญในภาษาไทย..... 12

2.3.1 หน่วยเสียงสระ..... 12

2.3.2 หน่วยเสียงพยัญชนะ..... 14

2.3.3 หน่วยเสียงวรรณยุกต์..... 16

2.4 ลักษณะ พยางค์ และ คำในภาษาไทย..... 17

บทที่ 3 ความเป็นมาและทฤษฎีเวฟเลท..... 19

3.1 การวิเคราะห์หลายระดับความละเอียด..... 21

3.2 ลักษณะของฟังก์ชันสเกลลิงและฟังก์ชันเวฟเลท..... 23

3.3 ตัวกรองฟิลเตอร์แบงก์ (Filter Bank)..... 27

3.4 การแปลงเวฟเลทแบบแบ่งช่วง (Discrete Wavelet Transform) 29

3.4.1 การแจกแจงสัมประสิทธิ์เวฟเลท..... 29

3.4.2การแปลงเวฟเลทโดยใช้หลักการAnalysis Filter Bank..... 30

3.4.3 การแปลงเวฟเลทโดยใช้หลักการ Synthesis Filter Bank..... 31

3.5 สรุป..... 32

บทที่ 4 การทดลองและผลการทดลอง..... 33

4.1 วิธีการแยกเสียงพยัญชนะและสระ..... 33

4.2 วิธีการทดลอง..... 34

4.2.1 การเก็บข้อมูลเสียง..... 34

4.2.2 ขั้นตอนในการแยกเสียงพยัญชนะและสระ..... 37

4.3 ผลการทดลอง..... 38

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

บทที่ 5 สรุปผลและข้อเสนอแนะ 54

5.1 สรุปผลการทดลอง 55

5.2 ข้อสังเกต และปัญหาที่พบในการทดลอง 55

5.3 ข้อเสนอแนะ 55

เอกสารอ้างอิง 56

ภาคผนวก ก.บทความเกี่ยวข้องที่ได้รับการตีพิมพ์ 57

ประวัติผู้เขียน 63

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 อวัยวะส่วนที่มีหน้าที่ในการออกเสียงพูด.....	7
2.2 ลักษณะเสียงพูดของมนุษย์.....	12
2.3 เสียงสระเดี่ยวในภาษาไทย.....	13
2.4 เสียงสระเดี่ยวในภาษาไทยแบ่งตามความสั้น-ยาว.....	14
2.5 เสียงพยัญชนะจำแนกตามการออกเสียง.....	15
2.6 เสียงพยัญชนะ ในภาษาไทย.....	15
2.7 แสดงหน้าที่ของหน่วยเสียงพยัญชนะในภาษาไทย.....	16
2.8 แสดงลักษณะของเสียงวรรณยุกต์ในภาษาไทย.....	17
4.1 แสดงผลของความถูกต้องของการแยกเสียงพยัญชนะและเสียงสระ.....	40

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่	หน้า
1.1 ส่วนประกอบของการรู้จำเสียงพูดในภาษาไทยโดยวิธีแยกแยะการรู้จำหน่วยเสียง.....	2
2.1 วงจรการสื่อสารภาษาพูด.....	5
2.2 อวัยวะในระบบการพูดของมนุษย์.....	9
2.3 แผนภาพระบบเสียงพูดของมนุษย์.....	10
2.4 ตำแหน่งของลิ้นเมื่ออยู่ต่ำ ส่วนหน้าสูง ส่วนหลังสูง.....	13
2.5 การเกิดเสียงสระผสมในภาษาไทย.....	14
2.6 องค์ประกอบของหน่วยเสียงต่างๆ ในพยางค์.....	18
3.1 ฟังก์ชันชายน์และฟังก์ชันเวฟเลท.....	20
3.2 การซ้อนทับของปริภูมิเวกเตอร์สเปซ.....	24
3.3 สเกลลิงฟังก์ชันและเวปเลทเวกเตอร์สเปซ.....	26
3.4 ขั้นตอนการวิเคราะห์สัญญาณเสียงพูดเบื้องต้น.....	27
3.5 Octave Band Synthesis Filter Banks.....	27
3.6 ลักษณะการแยกแบนด์ความถี่ (Frequency Bands) ของ Octave Filter Banks.....	28
3.7 Octave Banks Synthesis Filter Banks.....	28
4.1 แสดงรูปแบบพลังงานพยัญชนะและสระ.....	34
4.2 แสดงสัญญาณเสียงที่เก็บตัวอย่างสัญญาณเสียง คำว่า “ขา”.....	35
4.3 ผลตอบสนองความถี่ของสัญญาณเสียงคำว่า “ขา”.....	36
4.4 ผังการทำงานของการแยกเสียงพยัญชนะเสียงสระ.....	37
4.5 เสียงพยัญชนะเสียงระเบิดตามด้วยสระ คำว่า “ขอ”.....	39
4.6 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงระเบิดตามด้วยเสียงสระ””.....	39

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.7เสียงพยัญชนะเสียงเสียดแทรกตามด้วยสระ คำว่า “แซะ””””	40
4.8 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงเสียดแทรกตามด้วยเสียงสระ	40
4.9 เสียงพยัญชนะเสียงกึ่งเสียดแทรกตามด้วยสระ คำว่า “ชี””””	41
4.10 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงกึ่งเสียดแทรกด้วยเสียงสระ	41
4.11 เสียงพยัญชนะเสียงนาสิกตามด้วยสระ คำว่า “นอ”)	42
4.12 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงนาสิกแทรกด้วยเสียงสระ	42
4.13 เสียงพยัญชนะเสียงข้างตามด้วยสระ คำว่า “ลือ”	43
4.14 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงข้างด้วยเสียงสระ	43
4.15 เสียงพยัญชนะเสียงรัวตามด้วยสระ คำว่า “รอ”	44
4.16 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงรัวด้วยเสียงสระ	44
4.17 เสียงพยัญชนะเสียงครึ่งสระตามด้วยสระ คำว่า “วี”	45
4.18 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงครึ่งสระด้วยเสียงสระ	45

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันทั่วโลกมีการพัฒนาเทคโนโลยีขึ้นมามากมาย เนื่องจากเทคโนโลยีทางด้านคอมพิวเตอร์ได้ถูกพัฒนาให้มีความสามารถมากขึ้นทั้งด้านความเร็ว และความถูกต้อง และได้เข้ามามีบทบาทในชีวิตประจำวันของมนุษย์เป็นอย่างมาก ซึ่งโดยปกติการติดต่อระหว่างคอมพิวเตอร์กับมนุษย์จะทำได้โดยจะรับคำสั่งเข้าทางแป้นพิมพ์ และจะแสดงผลออกทางจอภาพหรือเครื่องพิมพ์ ดังนั้นการติดต่อกับคอมพิวเตอร์ผ่านทางแป้นพิมพ์ ต้องอาศัยความชำนาญและการฝึกฝนเป็นเวลานานพอควร ทำให้เกิดความคิดในการพัฒนาการจะเพิ่มความสามารถของมนุษย์ในการติดต่อกับเครื่อง(Man-Machine Communication) จึงเป็นที่มาของการศึกษาวิจัย และพัฒนาวิชาแขนงหนึ่งเรียกว่า "การรู้จำเสียงพูด (Speech Recognition)" เป็นพัฒนาการติดต่อกับคอมพิวเตอร์โดยใช้เสียงพูดแทน

ในปัจจุบันการรู้จำเสียงพูด (Speech Recognition) ซึ่งได้ถูกวิจัยและพัฒนามามากมายในเสียงภาษาอังกฤษถูกนำมาประยุกต์ใช้งานกันอย่างกว้างขวาง ได้แก่ การกรอกข้อมูลใส่ฐานข้อมูลที่สำคัญต่างๆ ด้วยเสียง การควบคุมคุณภาพด้วยเสียงในกระบวนการผลิต การสั่งอุปกรณ์และเครื่องใช้ต่างๆ ด้วยเสียง โดยเฉพาะในระบบสื่อสารโทรคมนาคมนำมาใช้ตั้งแต่ การหมุนเลขหมายโทรศัพท์ การบริการความช่วยเหลืออัตโนมัติที่สั่งการด้วยเสียงทางโทรศัพท์ การใช้บริการต่างๆ ด้วยเสียงผ่านโทรศัพท์ทั้งทางการรักษาทางการแพทย์ด้วย จุดมุ่งหมายหลักของการรู้จำเสียงพูดก็คือ การเพิ่มพูนความสามารถของอุปกรณ์ต่างๆในการรับรู้และสื่อสารโต้ตอบกับมนุษย์ได้ เพื่อเพิ่มทางเลือกในการควบคุมสั่งการอุปกรณ์เครื่องมือต่างๆ โดยเฉพาะเครื่องคอมพิวเตอร์ ซึ่งการใช้เสียงพูดควบคุมสั่งการนี้ถือได้ว่าเป็นวิธีการที่เป็นธรรมชาติมากที่สุดของมนุษย์

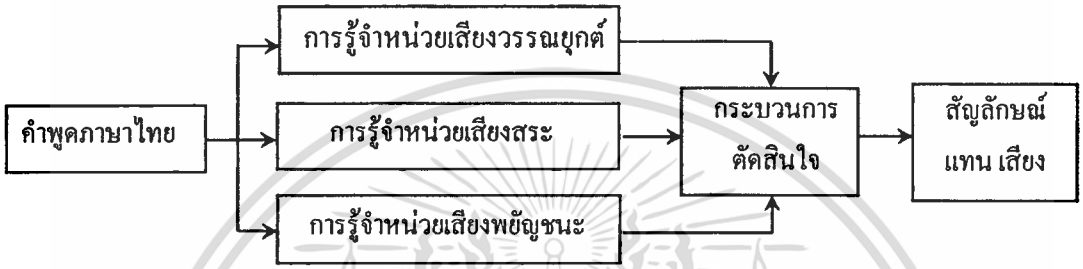
การรู้จำเสียงพูดนั้นได้มีการศึกษาและวิจัยกันอย่างแพร่หลาย มีทั้งการรู้จำเสียงแบบขึ้นกับผู้พูด (Speaker-Dependent Speech Recognition) และการรู้จำเสียงแบบไม่ขึ้นกับผู้พูด (Speaker-Independent Speech Recognition) ซึ่งการรู้จำเสียงพูดนั้นมีความยุ่งยากและซับซ้อนมาก เนื่องจากความแตกต่างในการออกเสียงของผู้พูดแต่ละบุคคล จึงทำให้การพัฒนาทางด้านการรู้จำเสียงพูดเป็นไปอย่างล่าช้า การศึกษาและวิจัยการรู้จำเสียงพูดนั้นยังแบ่งตามลักษณะการรับรู้เสียงพูดออกเป็น 2 วิธีคือ

1.1 พิจารณาทั้งหน่วยภาษาที่เปล่งเสียงออกมาทั้งชุด มีทั้งระบบการรู้จำคำพูดเดี่ยว [1] (Isolated word Recognition) และระบบรู้จำคำพูดต่อเนื่อง(Continuous word Recognition) ซึ่งข้อดีของระบบเหล่านี้คือง่าย เนื่องจากมีการหลีกเลี่ยงผลกระทบอันเนื่องจากฐานรากของเสียงภายในคำ

เอกสารนี้มีลิขสิทธิ์ ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หรือกลุ่มคำนั้น แต่ข้อเสียคือสามารถรู้จำคำได้จำนวนคำที่จำกัด เนื่องจากต้องใช้เนื้อที่จำนวนมากในการจัดเก็บแบบจำลองอ้างอิง และต้องใช้เวลาในการคำนวณเพื่อเปรียบเทียบมากตามจำนวนของแบบจำลองอ้างอิงที่มีอยู่

1.2 พิจารณาโดยการแยกแยะรายละเอียดของหน่วยเสียง (Phonetic Recognition) วิธีนี้จะพิจารณาลักษณะของหน่วยเสียงพยัญชนะ หน่วยเสียงสระ และหน่วยเสียงวรรณยุกต์ ดังแสดงในรูปที่ 1.1 โดยจะใช้หน่วยเสียงย่อยเหล่านี้เป็นหลักในการรู้จำเสียงพูด ซึ่งวิธีนี้เหมาะสำหรับการพัฒนาไปสู่ระบบการรู้จำคำจำนวนมาก



รูปที่ 1.1 ส่วนประกอบของการรู้จำเสียงพูดในภาษาไทยโดยวิธีแยกแยะการรู้จำหน่วยเสียง

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

1.2.1 เพื่อศึกษาลักษณะของหน่วยเสียงพยัญชนะและสระ

1.2.2 เพื่อศึกษาระบบการรู้จำเสียงในภาษาไทย

1.2.3 เพื่อศึกษาวิธีการแยกเสียงพยัณชนะเสียงสระแบบต่างๆ

1.2.4 เพื่อศึกษาคุณสมบัติของเวฟเลท

1.2.5 เพื่อศึกษาในการนำเวฟเลทมาประยุกต์ใช้งาน

1.2.6 เพื่อพัฒนาระบบการรู้จำเสียงพูดในภาษาไทย โดยนำเวฟเลทมาแยกเสียงพยัญชนะเสียงสระ

1.3 สมมติฐานของการศึกษา

คำในภาษาไทยประกอบด้วยหน่วยเสียงวรรณยุกต์ หน่วยเสียงสระ และหน่วยเสียงพยัญชนะ คำในภาษาไทยนั้นจะมีหน่วยเสียงพยัญชนะหน่วยเสียงสระเป็นองค์ประกอบที่สำคัญ เนื่องจากหน่วยเสียงพยัญชนะหน่วยเสียงสระสระเป็นส่วนที่มีมากที่สุดของคำ และเนื่องจากในภาษาไทยมีเสียงพยัญชนะในภาษาไทยมีทั้งหมด 21 เสียง จากรูป พยัญชนะ 44 รูป แบ่งออกเป็นเสียงระเบิด (Plosive) เสียงเสียดแทรก (Fricative) เสียงกึ่งเสียดแทรก (Affricate) เสียงนาสิก (Nasal) เสียงข้าง (Lateral) เสียงรัว (Trill)เสียงครึ่งสระ (Semi- Vowel) เสียงสระแบ่งออกเป็นเสียงสระเดี่ยว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สระผสม และยังแบ่งเป็นสระเสียงสั้นและสระเสียงยาว แต่องค์ประกอบของเสียงสระทั้งหมดที่กล่าวนั้นเกิดจากหน่วยเสียงย่อยที่สุดของเสียงสระคือหน่วยเสียงสระเสียงเดี่ยว ดังนั้นการแยกเสียงพยัญชนะเสียงสระสำหรับภาษาไทยทั้งหมดนั้นก็เป็นจุดเริ่มต้นจากการรู้จำเสียงพยัญชนะ เสียงสระ ซึ่งงานวิจัยนี้ได้นำการแปลงเวฟเลทมาแยกเสียงพยัญชนะ สียงสระ ออกจากกันเป็นวิธีการพัฒนาให้มีประสิทธิภาพมากขึ้นเช่น ลดเวลาในกระบวนการประมวลผล ลดขนาดแบบอ้างอิงที่ใช้ในการรู้จำ และความถูกต้องในการรู้จำที่ดีขึ้น

1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย

คำในภาษาไทยประกอบด้วยพยัญชนะตามด้วยสระและระดับเสียงวรรณยุกต์ ซึ่งกระบวนการรู้จำนั้นแบ่งออกเป็นรู้จำเสียง เสียงพยัญชนะ เสียงสระ และวรรณยุกต์ ขั้นตอนการรู้จำเสียงพยัญชนะและเสียงสระต้องมีการแยกเสียงทั้งสองออกจากกันดังนั้นความถูกต้องของการรู้จำเสียงพยัญชนะและเสียงสระต้องแยกเสียงดังกล่างออกจากกัน วิธีการแยกเสียงพยัญชนะและเสียงสระออกจากกัน โดยอาศัยกำลังงานของเสียงและความถี่ที่ต่างกันมาเป็นตัวแยก วิธีการแยกเสียงพยัญชนะและเสียงสระมีหลายวิธีเช่น Neural Network Model [1] วิธีนี้ใช้การเขียนโปรแกรมสร้างแบบจำลองให้ครอบคลุมกับเงื่อนไขของเสียง ซึ่งได้เปอร์เซ็นต์ความถูกต้องของค่าเฉลี่ยอยู่ที่ 90 เปอร์เซ็นต์ วิธีการนี้จะใช้ระยะเวลาในการคำนวณมากหรือน้อย ขึ้น อยู่กับรายละเอียดของแบบจำลองที่เราสร้างขึ้นมาและความแม่นยำนั้นขึ้นอยู่กับการสร้างแบบจำลองให้ครอบคลุมกับรายระเอียดของเสียงและข้อเสียหนึ่งคือต้องใช้เวลาในการจำแนกเสียงมากยิ่งขึ้น และอีกวิธีนั้นคือวิธี Zero-Crossing rate[2] เป็นการชี้จุดตัดของกำลังงานเสียงของพยัญชนะและเสียงสระได้เปอร์เซ็นต์ความถูกต้องอยู่ที่ 95 เปอร์เซ็นต์ ซึ่งสามารถนำมาพัฒนาต่อให้ได้ความเร็วและความแม่นยำเพิ่มมากขึ้นโดยใช้การแปลงเวฟเลท[3],[4]

สุดท้ายงานวิจัยนี้จะทำให้เราได้การการแยกเสียงพยัญชนะและเสียงสระในภาษาไทยที่มีประสิทธิภาพเพิ่มขึ้นอย่างมาก

1.5 ขอบเขตการวิจัย

งานวิจัยนี้ได้เลือกศึกษาคำในภาษาไทย โดยแบ่งออกเป็นคำที่ประกอบไปด้วยเสียงพยัญชนะที่ผันตามเสียงพยัญชนะต้น 21 รูปเสียง ผสมกับสระในภาษาไทยประกอบด้วยเสียงสระแท้ทั้งหมด 24 เสียง จำนวนคำทั้งหมด จำนวน 504 พยางค์ (พยัญชนะ 21 เสียง และ สระ 24 เสียง) ทั้งชายและหญิงเก็บด้วยความละเอียด 8000 Hz 8 บิต เฉพาะเสียงพูด

1.6 ขั้นตอนการศึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.6.1 ศึกษาระบบการรู้จำภาษาในลักษณะต่างๆ ที่ได้มีการศึกษามาแล้ว

1.6.2 ศึกษาทฤษฎีทางด้านภาษาศาสตร์ และอัลกอลิธึม (Algorithm) ต่างๆ ที่ใช้ในการวิเคราะห์เสียงพูด

1.6.3 กำหนดขอบเขต และแนวทางการวิจัย

1.6.4 บันทึกข้อมูลเสียง โดยใช้ผู้พูดทั้งหมด 8 คน เพศชาย 4 คน เพศหญิง คน

1.6.5 ทำการวิเคราะห์เสียง และทำการวิเคราะห์สัมประสิทธ์ของเสียง

1.6.6 ทำการเก็บข้อมูล การทดลองและผลความถูกต้องในการรู้จำ

1.6.6 สรุปผลการวิจัย และเสนอแนะแนวทางในการวิจัยต่อไป

1.7 โครงประกอบของวิทยานิพนธ์

แบ่งออกเป็น 5 บทดังนี้

บทที่ 1 ดังได้กล่าวมาแล้วข้างต้น

บทที่ 2 ทฤษฎีการสร้างเสียงพูดในภาษาไทย เกี่ยวกับการออกเสียง หน่วยเสียงต่างๆ ในภาษาไทย

บทที่ 3 ความเป็นมาและทฤษฎีของเวฟเล็ท

บทที่ 4 การทดลองและผลการทดลอง แสดงขั้นตอนในการทดลอง และผลจากการทดลอง

บทที่ 5 สรุป และข้อเสนอแนะ เป็นการสรุปเกี่ยวกับการทดลองทั้งหมด ค่าที่มีประโยชน์ ข้อสังเกตจากการทดลองที่พบปัญหาและข้อเสนอแนะสำหรับเป็นแนวความคิดในการพัฒนางานวิจัยต่อไป

ภาคผนวก

ภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

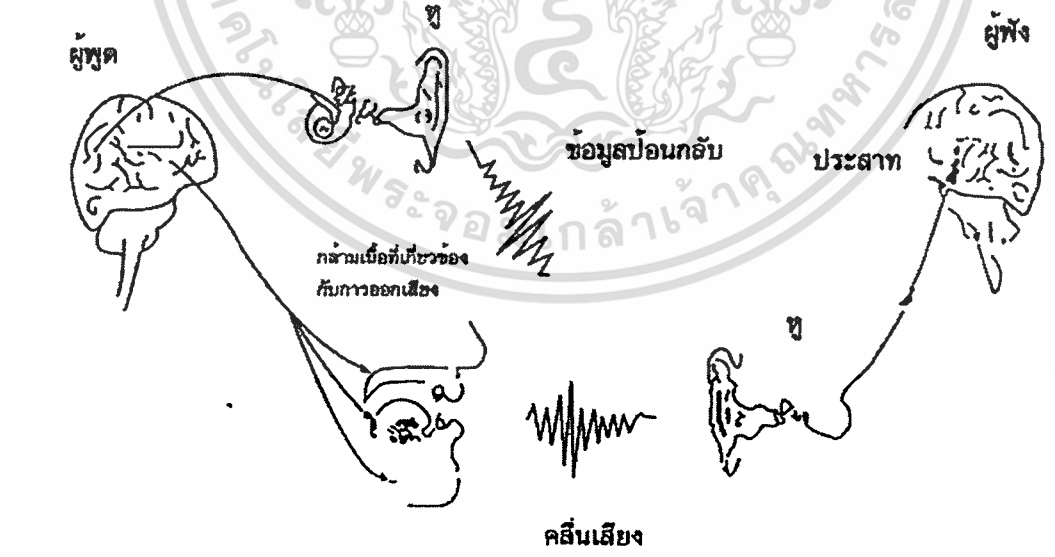
บทที่ 2
ทฤษฎีการสร้างเสียงพูดในภาษาไทย

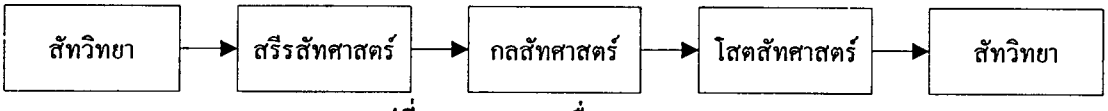
2.1 กล่าวนำ

ในบทนี้จะกล่าวถึงทฤษฎีต่างๆที่เกี่ยวข้องกับงานวิจัยนี้ โดยจะเริ่มจากทฤษฎีทางภาษาศาสตร์ ซึ่งจะกล่าวถึงวงจรการสื่อสารทางภาษาพูด เรื่องทฤษฎีการสร้างเสียงพูดที่ใช้อวัยวะกำเนิดหน่วยเสียงต่างๆ ประกอบด้วย หน่วยเสียงสระ หน่วยเสียงพยัญชนะ หน่วยเสียงวรรณยุกต์ และลักษณะคำ และพยางค์ ในภาษาไทย จากนั้นบทต่อไปจะกล่าวถึงทฤษฎีการวิเคราะห์เสียงประกอบด้วย การคำนวณสเปกตรัมแอลพีซี แถบความถี่วิกฤติ และความเข้มสัญญาณในแถบความถี่วิกฤติ

2.2 ทฤษฎีภาษาศาสตร์

ในการสื่อสาร โดยภาษาพูด ผู้พูดเริ่มวงจรจากความนึกคิดว่าต้องการจะสื่อข้อความใด แล้วสมองจึงสั่งการส่งสัญญาณให้อวัยวะต่างๆที่เกี่ยวข้องทำงานและออกเสียงทีละเสียงเป็นคำและถ้อยความตามที่ต้องการ คลื่นเสียงเดินทางผ่านตัวกลางคือ อากาศมากระทบหูผู้ฟัง ผู้ฟังรับคลื่นเสียงและถอดเสียงในถ้อยความและตีความหมายรับรู้สาร





รูปที่ 2.1 วงจรการสื่อสารภาษาพูด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วงจรนี้เริ่มต้นที่สามัตถิยะทางภาษา (Linguistic competence) คือความรู้ความสามารถทางภาษาที่เป็นธรรมชาติของมนุษย์ โดยเฉพาะในส่วนที่เกี่ยวข้องกับเสียงและระบบเสียงซึ่งเป็นเรื่องทาง สัทวิทยา (Phonology) การออกเสียงและเปล่งเสียงเป็นเรื่องทางสรีรสัทศาสตร์ (Articulatory phonetics) กายภาพของเสียงหรือคลื่นเสียงเป็นเรื่องทางกลสัทศาสตร์ (Acoustic phonetics) และการได้ยินและการรับรู้เสียงเป็นเรื่องทางโสตสัทศาสตร์ (Auditory phonetics) การถอดเสียงในถ้อยความของผู้ฟังเป็นเรื่องของสามัตถิยะทางภาษาที่เกี่ยวข้องกับเสียงและระบบเสียงซึ่งสัมพันธ์กับหน่วยคำหรือคำซึ่งเป็นเรื่องของสัทวิทยา โดยสัทวิทยาเป็นการศึกษาเรื่องเสียงและระบบเสียงของภาษาซึ่งเป็นนามธรรมและปรากฏเป็นพฤติกรรมทางภาษาที่เป็นรูปธรรมกล่าวคือสัทวิทยาศึกษาโครงสร้างและหน้าที่ของเสียงซึ่งสัมพันธ์กับความหมายในภาษา [8]

มนุษย์สร้างเสียงพูดขึ้นมาได้โดยมนุษย์จะมีเส้นเอ็นในลำคอที่เรียกว่าเส้นเสียง (Vocal Cords) ซึ่งอยู่ในช่องหลอดลม และจากการสั่นของเส้นเสียงจะทำให้เกิดคลื่นเสียงความถี่ต่างๆ ผ่านเข้าสู่ลำคอ ผ่านจากลำคอเข้าสู่ปากหรือช่องจมูกออกไปภายนอกซึ่งขนาดและรูปร่างของอวัยวะภายในช่องปากนี้จะเป็นสิ่งกำหนดว่าคลื่นเสียงความถี่ไหนจะยอมให้ปรากฏออกมาให้ได้ยิน (Passed Through) หรือคลื่นเสียงความถี่ไหนที่จะถูกดูดซับไว้ไม่ยอมให้ปรากฏออกมา (Absorbed) ปรากฏการณ์ที่ทำให้เกิดเสียงพูดออกมาได้นี้ เนื่องจากรูปร่างของช่องคอและช่องปาก ของมนุษย์ที่สามารถเปลี่ยนแปลงไปมาได้อย่างอิสระ ทำให้ผู้พูดมีความสามารถในการกำหนดคลื่นเสียงระดับต่างๆ เหล่านี้นั้นเองที่ทำให้เกิดการแปลงเป็นเสียงพูดออกมา [9]

2.2.1 อวัยวะออกเสียง (The Organs of Speech)

ธรรมชาติได้กำหนดตำแหน่งของอวัยวะที่ใช้ในการทำให้เกิดเสียงของมนุษย์เอาไว้ นั้นคือ ริมฝีปาก ฟัน เพดาน และ ลิ้น ที่อยู่ตามลำดับตลอดช่องของกล้ามเนื้อช่องคอ เพื่อทำให้สามารถเปลี่ยนขนาดและรูปร่างได้ตามต้องการ ดังรูปที่ 2.1 เราจะเปล่งคำพูดออกมาโดยการเปล่งออกมาเป็นลำดับขั้นของความถี่คลื่นเสียงที่ได้คัดเลือกไว้ออกมา นั้นคือ จะมีการเปลี่ยนแปลงตำแหน่งของช่องทางเดินเสียง (Vocal Tract) ของมนุษย์เป็นลำดับขั้นอย่างต่อเนื่อง

อวัยวะที่ใช้ในการออกเสียงนั้น อาจจะแบ่งคร่าวๆ ได้เป็น 2 พวก คือ (กาญจนา นาคสกุล, 2524)

2.2.1.1 อวัยวะที่เป็นส่วนกระทำอาการ (Articulator) หมายถึงอวัยวะส่วนที่เคลื่อนไหวเพื่อผลักลมไปยังส่วนต่างๆ อวัยวะที่เป็นส่วนกระทำการที่สำคัญ คือลิ้น ซึ่งเป็นส่วนที่เคลื่อนไหวได้มากที่สุด

2.2.1.2 อวัยวะที่เป็นฐานที่เกิดเสียงต่างๆ (Point of articulator) หมายถึง ตำแหน่งหรือฐานที่เกิดของเสียงต่างๆ เช่น ริมฝีปาก ฟัน เพดานส่วนต่างๆ เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อวัยวะส่วนที่มีหน้าที่ในการออกเสียงพูดมีดังแสดงในตารางที่2.1

ตารางที่ 2.1 อวัยวะส่วนที่มีหน้าที่ในการออกเสียงพูด

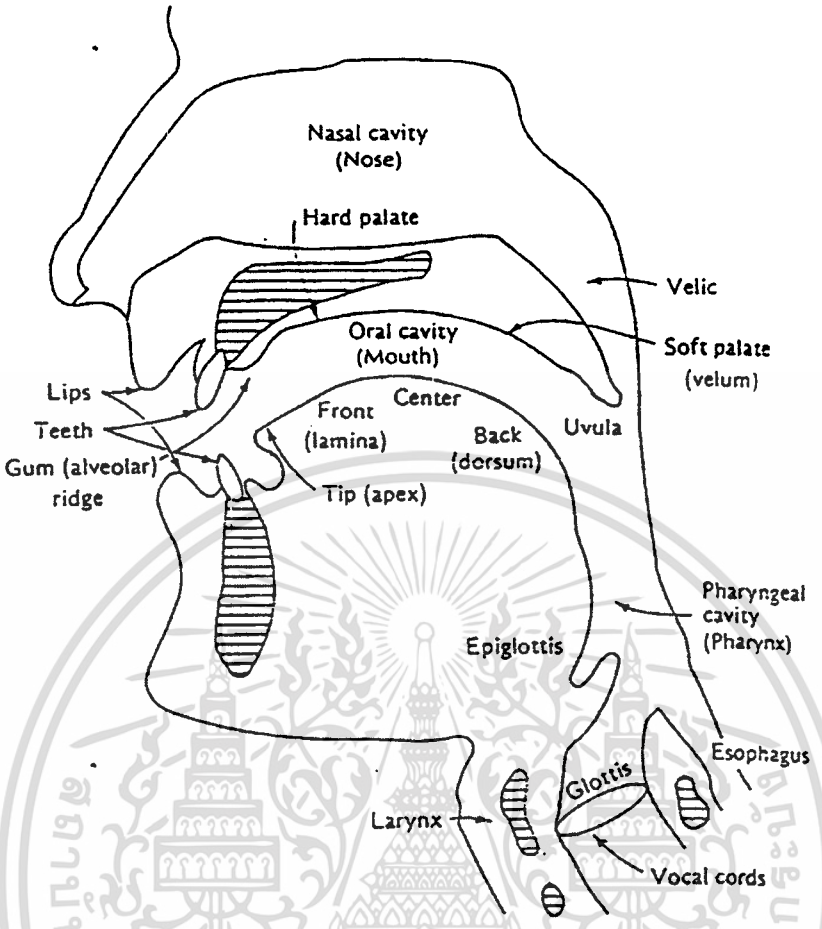
อวัยวะ	ลักษณะ / หน้าที่
ริมฝีปาก (Lips)	เป็นอวัยวะส่วนที่สามารถเคลื่อนไหวได้มาก และทำให้เสียงแตกต่างกันได้มาก เราอาจบังคับริมฝีปากให้อยู่ชิดกัน ห่างกัน ยื่นออกมา หรือห่อกลม ฯลฯ ก็ได้ ลักษณะริมฝีปากต่างๆ นี้ ล้วนมีอิทธิพลต่อการออกเสียง และการทำให้เสียงแตกต่างกันไปทั้งสิ้น
ฟัน (Teeth)	เป็นอวัยวะที่เกิดของเสียงหลายชนิด เช่น เมื่อฟันบนกดลงบนริมฝีปากล่างหรือกดกับฟันล่าง ลมที่ผ่านออกมาโดยแรงจะลอดช่องที่พอผ่านได้ออกมาทำให้เกิดเป็นเสียงชนิดที่เรียกว่าเสียงเสียดแทรกที่เกิดที่ฟัน เป็นต้น
ปุ่มเหงือก (Gum ridge)	เป็นส่วนที่อยู่ต่อจากฟันบน
เพดานแข็ง หรือเพดานปาก (Hard palate)	เฉพาะเพดานที่โค้งเป็นกระดูกแข็งเท่านั้น
เพดานอ่อน (Soft palate)	เป็นส่วนของเพดานที่อยู่ต่อเพดานแข็งไปข้างใน มีลักษณะเป็นกระดูกอ่อนที่ขยับขึ้นลงได้เวลาหายใจเพดานอ่อนและลิ้นไก่ซึ่งอยู่ปลายเพดานอ่อนจะลดระดับลงมาเปิดช่องให้ลมออกไปทางจมูก เวลาพูดส่วนใหญ่ปลายเพดานอ่อนและลิ้นไก่จะถูกยกขึ้นไปจรดกับหลังคอ นอกจากเวลาออกเสียงนาสิก เท่านั้นที่เพดานอ่อนจะลดระดับลงมา
ลิ้น (Tongue)	เป็นส่วนที่เคลื่อนไหวมากที่สุดในการออกเสียงพูด จึงต้องแบ่งเป็น 3 ส่วนด้วยกันตามหน้าที่ในการออกเสียง
ปลายลิ้น (Blade of the tongue)	เป็นส่วนปลายซึ่งสามารถยกขึ้นไปแตะอวัยวะส่วนต่างๆ ในปากตอนบนได้ง่าย
หน้าลิ้น (Front of the tongue)	หมายถึงลิ้นที่อยู่ตรงข้ามกับเพดานแข็ง ขณะที่วางลิ้นราบกับปากตอนที่ไม่ได้พูด
หลังลิ้น (Back of the tongue)	หมายถึงส่วนของลิ้นที่อยู่ตรงข้ามกับเพดานอ่อน ขณะที่วางลิ้นราบกับปาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.1 (ต่อ)

อวัยวะ	ลักษณะ / หน้าที่
ช่องปาก (Oral cavity)	ช่องปากนี้เป็นอวัยวะอีกส่วนหนึ่งที่มีความสำคัญในการพูด กล่าวคือ ช่องปากจะทำหน้าที่เป็นช่องกำทอน (Resonant chamber) ซึ่งมีบทบาทในการกำทอนเสียงที่เดินทางผ่านมาถึงบริเวณนี้ ทั้งนี้เพราะช่องปากสามารถจะเปลี่ยนแปลงเป็นช่องกำทอนรูปร่างต่างๆ กัน เนื่องจากรูปร่างของมันแปรผันไปตามการจัดท่าทางของลิ้น ริมฝีปาก และขากรรไกร
โพรงจมูก(Nasal cavity)	คุณสมบัติสักษณะของเสียงที่พูดที่เกิดขึ้นจะแปรผันไปตามการปิด – เปิดของช่องทางออกสู่โพรงจมูก ซึ่งเป็นผลจากการยกขึ้นหรือเลื่อนลงของเพดานอ่อน
ช่องคอ (Pharynx)	หมายถึงโพรงคอซึ่งอยู่ถัดจากปากลงไปจนถึงเส้นเสียง
ลิ้นปิดหลอดลม (Epiglottis)	เป็นก้อนเนื้อเล็กๆ คล้ายลิ้นไก่ อยู่ต่อจากโคนลิ้นลงไปในลำคอ มีหน้าที่ปิดช่องลมขณะรับประทานอาหาร และปิดช่องลมและจมูก
เส้นเสียง (Vocal cords)	เป็นอวัยวะสำคัญที่เกิดของเสียง เส้นเสียงมีลักษณะเป็นกล้ามเนื้อ 2 แผ่น ปิดขวางอยู่ปากช่องหลอดลมจากด้านหลังมาด้านหน้า ระหว่างเส้นเสียงจะมีช่องว่างระหว่างเส้นเสียง (glottis) เส้นเสียงทั้งสองสามารถจะดึงออกให้ห่างจากกัน หรือดึงเข้ามาประชิดกันก็ได้ เส้นเสียงเป็นส่วนทำให้เกิดเสียงพูดขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.2 แสดงอวัยวะในระบบการพูดของมนุษย์

2.2.2 กระบวนการผลิตเสียงพูด

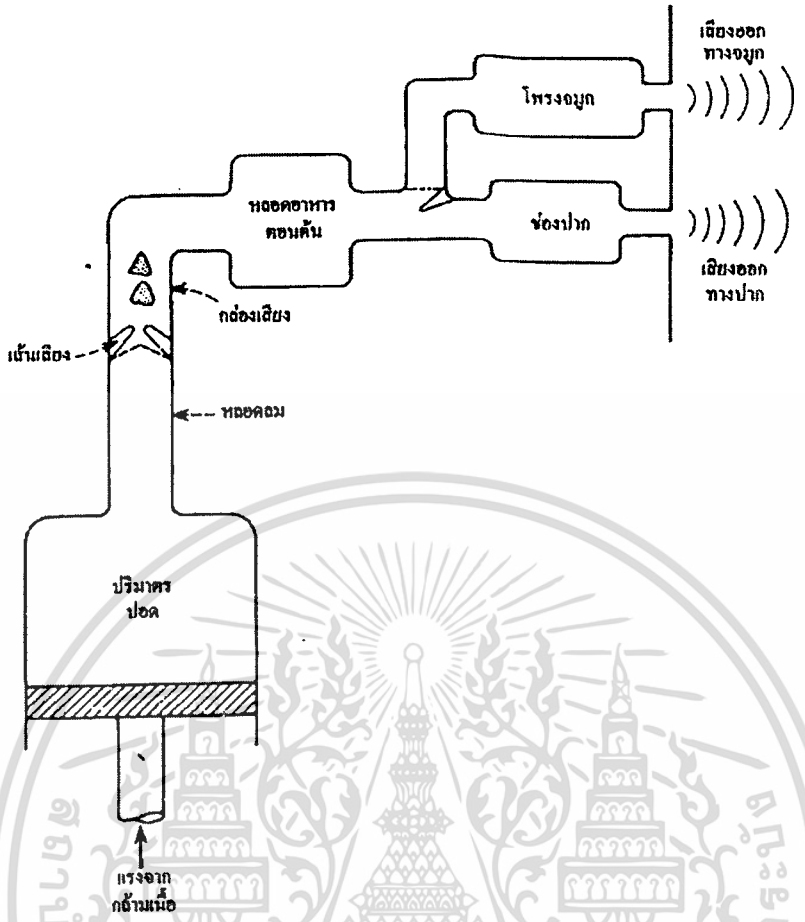
จากระบบเสียงพูดของมนุษย์ สามารถแสดงเป็นแผนภาพของระบบกำเนิดเสียงดังรูปที่ 2.2 ซึ่งเราสามารถจำแนกกลไกสร้างเสียงพูด ของมนุษย์ได้ 3 แบบ ดังนี้

2.2.2.1 อากาศไหลจากปอดจะถูกมอดูเลท (Modulate) โดยการสั่นของเส้นเสียงทำให้เกิดคลื่นเสียงลักษณะคล้ายพัลส์ที่มีคาบเวลาแบบควอไซ (Quasi_periodic pulse – like excitation)

2.2.2.2 อากาศไหลจากปอดถูกทำให้ปั่นป่วน ด้วยการบังคับให้ไหลผ่านช่องแคบอันเกิดจากการบีบตัวของอวัยวะในช่องปากทำให้เกิดเสียงลักษณะคล้ายเสียงรบกวน (Noise – like excitation)

2.2.2.3 อากาศที่ไหลถูกกัก และเกิดแรงดันอยู่ภายในส่วนของช่องปากที่ปิด จากนั้นจึงปล่อยให้อากาศที่มีแรงดันพุ่งออกไปอย่างรวดเร็ว ทำให้เกิดการกระตุ้นเป็นเสียงในช่วงเริ่มต้น (Transient excitation)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.3 แผนภาพระบบเสียงพูดของมนุษย์

2.2.3 เสียงพูดของมนุษย์

เสียงพูดเป็นคลื่นความยาว (Longitudinal wave) เกิดจากการสั่นของอนุภาคตัวกลาง นั้นคือ อากาศ และทิศทางการสั่นของอนุภาคจะอยู่ในทิศเดียวกันกับทิศทางของการเคลื่อนที่ คลื่นเสียงเป็นคลื่นที่เปลี่ยนแปลงไปตามเวลา

เสียงพูดแบ่งได้เป็น 2 ชนิด ตามการกำเนิดเสียงหรือโหมด (Mode) การกระตุ้น คือ

2.2.3.1 เสียงวอยซ์ หรือเสียงโฆษะ (Voiced sound) เกิดจาการบังคับอากาศให้ผ่านสายเสียงทำให้มีการเปลี่ยนแปลงความตึงหย่อนของเส้นเสียง โดยเส้นเสียงจะสั่นและเกิดเป็นพัลส์ (Pulse) ของอากาศไปกระตุ้นอวัยวะกำทอนเกิดเป็นเสียงวอยซ์ ตัวอย่างเสียงวอยซ์ ได้แก่ เสียงสระ เสียงพยัญชนะ ที่ออกเสียงจากลำคอ (Voiced consonants)

2.2.3.2 เสียงอันวอยซ์หรือเสียงอโฆษะ (Voiceless หรือ Unvoiced sounds) เป็นเสียงที่เกิดในช่องปากหรือโพรงจมูกโดยอวัยวะภายในช่องปาก ริมฝีปาก ขวางการไหลของอากาศได้เป็นช่องเล็กๆ อากาศจึงไหลผ่านอย่างรวดเร็วและปั่นป่วนจนกระทั่งสร้างเป็นเสียงรบกวนช่วงความถี่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กว้าง (Broad – spectrum noise) ตัวอย่างเสียงอันวอยซ์ ได้แก่ เสียงพยัญชนะที่ไม่ได้เกิดจากลำคอ (Voiceless consonants)

2.2.4 ความยาวของเสียง (Length)

การเกิดเสียงใดเสียงหนึ่งเปล่งออกมาได้นานเท่าใด เสียงพูดบางเสียงอาจจะเปล่งออกมาได้ติดต่อกันได้นาน เช่น เสียงสระ เสียงพยัญชนะ หรือเสียงพยัญชนะเสียดแทรก

ในภาษาไทย เสียงพูดที่มีความยาว – สั้น ก็มีเพียงเสียงสระเท่านั้น เช่น อะ อิ อุ เป็นเสียงสั้น อา อี อู เป็นเสียงยาวเป็นต้น

2.2.5 ระดับเสียง สูง – ต่ำ (Pitch)

เสียงพูดจะมีระดับ สูงหรือต่ำ อยู่ที่ความถี่พื้นฐานของเสียง (Fundamental frequency) ถ้าความถี่ต่ำเสียงก็จะต่ำ อวัยวะส่วนที่ทำให้เสียงมีระดับ สูง – ต่ำ คือ เส้นเสียง ดังนั้นระดับเสียงสูง – ต่ำ ก็คืออัตราการสั่นสะบัดของเส้นเสียงนั้นเอง

ในการพูดเสียงที่มีระดับสูง – ต่ำ ได้คือเสียงก้องเท่านั้น เพราะมีการสั่นสะเทือนของเส้นเสียง ที่ทำให้เกิดมีความถี่ระดับต่างๆ ได้ ในภาษาไทยระดับเสียงสูง – ต่ำ ของคำเราเรียกว่า "วรรณยุกต์"

2.2.6 ความดัง (Loudness)

ความดังขึ้นอยู่กับปริมาณของลม ที่ผู้พูดเปล่งเสียงออกมาในช่วงเวลาหนึ่งๆ

2.2.7 การลงน้ำหนัก (Stress)

การออกเสียงพยางค์ใดพยางค์หนึ่งให้ดังเน้นมากหรือน้อยกว่าพยางค์อื่นที่อยู่ข้างเคียง (เพื่อต้องการเรียกร้องความสนใจเป็นพิเศษ หรือแสดงอารมณ์ อย่างใดอย่างหนึ่ง)

2.2.8 ช่วงต่อของเสียง (Juncture)

ช่วงระยะที่ผู้พูดเปล่งเสียงหนึ่งแล้วต่อไปเปล่งเสียงอีกเสียงหนึ่ง ซึ่งเรียงกันมาเป็นลำดับเสียงที่ประกอบกันเข้าเป็นพยางค์ จะมีช่วงต่อของเสียงแนบสนิทจนไม่เห็นร่องรอย (Close juncture) แต่ถ้าเสียงปรากฏอยู่คนละพยางค์ หรือคนละคำจะมีช่วงต่อ "ห่าง" จนสังเกตได้ชัด (Open juncture) ดังนั้นช่วงต่อของเสียง โดยเฉพาะช่วงต่อห่างจะมีความสำคัญมากในการแบ่งคำในภาษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 หน่วยเสียงสำคัญในภาษาไทย

"หน่วยเสียง" (Phoneme) เป็นหน่วยเล็กที่สุดของภาษา หน่วยดังกล่าวได้แก่เสียงสำคัญๆ ในภาษาใดภาษาหนึ่งซึ่งทำหน้าที่ให้ความหมายของคำที่ใช้ในภาษานั้น และทำให้ความหมายของคำนั้นๆ มีความหมายแตกต่างจากคำอื่นๆ หน่วยเสียงสำคัญในภาษาไทยมี 3 ประเภทใหญ่ๆ คือ เสียงพยัญชนะ เสียงสระ และเสียงวรรณยุกต์ หน่วยเสียงทั้ง 3 นี่เองที่ประกอบกันเข้าเป็นคำที่ใช้ในภาษาไทย

เสียงพูดของมนุษย์ซึ่งมีความแตกต่างกันมากมายนั้น ถ้าเราพิจาณาอย่างถ้วนๆ จะพบว่าสามารถแบ่งออกเป็น 2 ประเภทใหญ่ ดังตารางที่ 2.2

ตารางที่ 2.2 ลักษณะเสียงพูดของมนุษย์

เสียง	ลักษณะ
เสียงเรียง (Segmental sound)	เป็นหน่วยเสียงที่แยกออกจากเสียงอื่นได้โดยเด็ดขาด เพราะมีลักษณะเด่นเฉพาะตัว ในภาษาไทยได้แก่เสียงสระ และเสียงพยัญชนะ
เสียงซ้อน (Supra – segmental feature)	เป็นเสียงที่ทำหน้าที่เป็นส่วนประกอบของเสียงอื่น เพราะไม่สามารถแยกเปล่งเสียงได้ตามลำพัง ในภาษาไทยได้แก่เสียงวรรณยุกต์ และทำนองเสียง เป็นต้น

2.3.1 หน่วยเสียงสระ

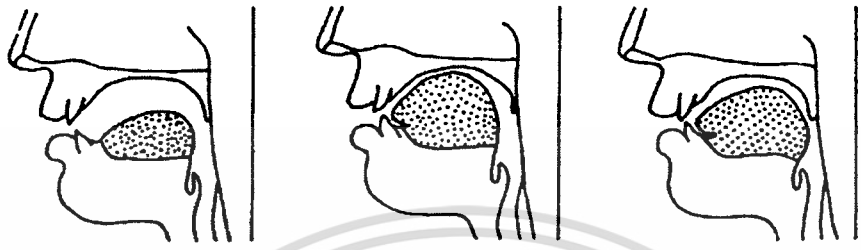
2.3.1.1 ลักษณะของเสียงสระ

ลักษณะสำคัญของเสียงสระก็คือ " เป็นเสียงก้องที่เปล่งเสียงออกมาโดยให้ลมออกทางช่องปากโดยไม่ถูกลิ้นกักหรือขัดขวาง" ดังนั้นเวลาเราออกเสียงสระจะออกเสียงได้สะดวกและออกเสียงได้นาน ทั้งนี้เพราะคุณสมบัติของเสียงสระมีความดังเด่นกว่าเสียงอื่นๆ ที่เรียงอยู่ข้างเสมอ อวัยวะที่เกี่ยวข้องกับการออกเสียงสระได้แก่ลิ้น กับริมฝีปาก ถ้าลิ้นส่วนใดทำหน้าที่เพียงส่วนเดียว เสียงที่เกิดขึ้นก็จะมีเพียงเส้นเสียงเดียว เสียงเช่นนี้เรียกว่า "สระเดี่ยว" แต่ถ้าลิ้นส่วนอื่นทำหน้าที่ร่วมด้วย เสียงสระนั้นเรียกว่า "สระผสม"

ฉะนั้นการอธิบายเรื่องการออกเสียงสระจึงเป็นเรื่องที่ยากกว่าการออกเสียงพยัญชนะเพราะไม่มีจุดสัมผัสในปากที่เราสามารถใช้เป็นหลักในการอธิบายได้ การที่สระมีเสียงต่างๆ กันก็เพราะช่องคอ ช่องปากมีรูปร่างต่างกันในขณะที่อากาศกำลังผ่านออกจากปากไป สิ่งที่มีส่วนทำให้ช่องปากช่องคอมีรูปร่างต่างๆ กันออกไปก็คือ ลิ้น และริมฝีปาก เพื่อความสะดวกในการอธิบายเสียงเราจะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบ่งลิ้นออกเป็น 3 ส่วน คือ ส่วนหน้า (Front) ส่วนกลาง (Center) และส่วนหลัง (Back) ในขณะที่เราหุบปากอยู่นิ่งๆ ส่วนหน้าของลิ้นอยู่ใต้เพดานแข็ง (hard palate) ส่วนหลังอยู่ใต้เพดานอ่อน (Velum) และส่วนกลางอยู่ระหว่างกลางของสองส่วนนี้ รูปต่อไปนี้แสดงตำแหน่งของลิ้นเมื่ออยู่ต่ำ ส่วนหน้าสูง ส่วนหลังสูง ดังรูปที่ 2.4



[a] "◌ะ" [i] "◌ิ" [u] "◌ุ"

รูปที่ 2.4 แสดงตำแหน่งของลิ้นเมื่ออยู่ต่ำ ส่วนหน้าสูง ส่วนหลังสูง

สำหรับภาษาไทยมีหน่วยเสียงสระทั้งหมด 24 หน่วยเสียง แยกออกเป็นสระเดี่ยว 18 หน่วยเสียง และสระผสม 6 หน่วยเสียง [8]

2.3.1.1.1สระเดี่ยว

เสียงสระเดี่ยว 18 หน่วยเสียง พิจารณาการเกิดเสียงได้เป็น 2 กรณีใหญ่ๆ คือ

2.3.1.1.1.1 การเกิดจากส่วนต่างๆ ของลิ้นหมายถึง ลมผ่านส่วนหน้า ส่วนกลาง หรือส่วนหลังของลิ้น

2.3.1.1.1.2 การเกิดจากลมผ่านลิ้นในขณะที่ลิ้นอยู่ในระดับ สูง กลาง หรือ ต่ำ

ตารางที่ 2.3 เสียงสระเดี่ยวในภาษาไทย

สระ	ระดับลิ้น	หน่วยเสียง	สัญลักษณ์
หน้า	สูง	อิ, อี	" i ", " i: "
	กลาง	เอะ, เอ	" e ", " e: "
	ต่ำ	แอะ, แอ	" ɛ ", " ɛ: "
กลาง	สูง	อึ อือ	" ɯ", " ɯ: "
	กลาง	เออะ เออ	" ɤ ", " ɤ: "
	ต่ำ	อะ อา	" a ", " a: "
หลัง	สูง	อุ อู	" u ", " u: "
	กลาง	โอะ โอ	" o ", " o: "
	ต่ำ	เอาะ ออ	" ɔ ", " ɔ: "

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

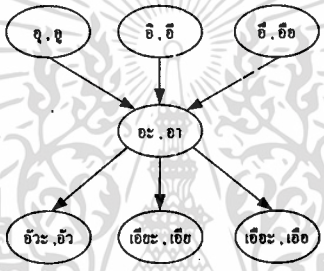
นอกจากนี้ หน่วยเสียงสระเดี่ยว 18 หน่วย สามารถแบ่งตามความสั้น – ยาว ของการออกเสียงได้เป็น

ตารางที่ 2.4 เสียงสระเดี่ยวในภาษาไทยแบ่งตามความสั้น-ยาว

สระ	จำนวน (หน่วย)	หน่วยเสียง								
สระเดี่ยวเสียงสั้น	9	อะ	อิ	อึ	อุ	เอะ	แอะ	โอะ	เอาะ	เออะ
สระเดี่ยวเสียงยาว	9	อา	อี	อือ	อู	เอ	แอ	โอ	ออ	เออ

2.3.1.1.2 สระผสม

เสียงผสม 6 หน่วยเสียง เกิดจากลมผ่านกระทบลิ้น 2 ส่วนคือ ส่วนบน และส่วนล่าง ซึ่งในขณะออกเสียงลิ้น จะอยู่ในระดับสูง แล้วลดลงต่ำโดยเสียงหลังเป็นสระ (อะ อา) เสมอ ดังรูปที่ 2.5 นี้



รูปที่ 2.5 แสดงการเกิดเสียงสระผสมในภาษาไทย

2.3.1.2 หน้าที่ของหน่วยเสียงสระในภาษาไทย

หน่วยเสียงสระในภาษาไทยทั้ง 24 หน่วยเสียงนี้ ทำหน้าที่เป็นแกนกลางของพยางค์ หรือคำ กล่าวคือ คำทุกคำในภาษาไทยจะต้องมีเสียงสระอยู่ด้วย และเสียงสระในภาษาไทยจะสามารถเกิดกับเสียงพยัญชนะต้นได้ทุกเสียง และสามารถเกิดกับหน่วยเสียงวรรณยุกต์ได้ทุกหน่วย แต่ไม่สามารถเกิดกับหน่วยเสียงพยัญชนะสะกดได้ทุกหน่วย หน่วยเสียงสระที่ทำให้เกิดคำหรือพยางค์ใช้ได้มากที่สุดในภาษามักเป็นหน่วยเสียงยาว

2.3.2 หน่วยเสียงพยัญชนะ

เสียงพยัญชนะในภาษาไทยมีทั้งหมด 21 หน่วยเสียง (44 รูป) หน่วยเสียงพยัญชนะออกเสียงได้ไม่สะดวกเท่าหน่วยเสียงชนิดอื่น เพราะเวลาออกเสียงลมหายใจที่พุ่งออกมาจากหลอดลมจะถูกขัดขวางตามส่วนต่างๆของปาก เสียงพยัญชนะจึงออกเสียงให้ยาวนานอย่างเสียงสระไม่ได้ และเสียงพยัญชนะก็ไม่ใช่เสียงก้องเสมอไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.2.1 ลักษณะของเสียงพยัญชนะ

หน่วยเสียงพยัญชนะ 21 หน่วยเสียงนี้จำแนกตามลักษณะเสียงเป็น เสียงก้อง เสียงไม่ก้อง เสียงหนัก เสียงเบา ตามตารางที่ 2.5 และลักษณะการออกเสียง ตามตารางที่ 2.6

ตารางที่ 2.5 เสียงพยัญชนะจำแนกตามการออกเสียง

ลักษณะเสียง	พยัญชนะ
เสียงก้อง (โฆษะ)	/ง//ย//บ//ด//ม//น// ร//ล//ว/
เสียงไม่ก้อง (อโฆษะ)	/ก//ค//จ//ช//ซ//ท//ต//ป//พ//ฟ//อ//ฮ/
เสียงหนัก (ธนิต)	/ค//ช//ท//พ/
เสียงเบา (ถวิล)	/ก//จ//ต//ป/

ตารางที่ 2.6 เสียงพยัญชนะในภาษาไทย

ที่เกิดเสียง / ลักษณะของเสียง		ริมฝีปาก	ปุ่มเหงือก	เพดานแข็ง	เพดานอ่อน	เส้นเสียง
พยัญชนะกัก	ไม่พ่นลม	ป*	ฏ, ต *	จ	ก*	อ*
	พ่นลม	ผ พ ภ	ฐ ท ฒ ถ ท ธ	ฉ ช ฌ	ข ค ฆ	
	ก้อง	บ	ฎ ด			
พยัญชนะไม่กัก	นาสิก	ม *	น ณ *		ง *	
	เสียดแทรก	ฝ ฟ	ซ ศ ษ ส			ฮ ห
	กระทบ		ร			
	ข้างลิ้น		ล ฬ			
	กึ่งสระ	ว *		ญ, ย*		

<u>หมายเหตุ</u> หน่วยเสียง พยัญชนะที่มีเครื่องหมายดอกจันทร์ (*) คือหน่วยเสียงที่สามารถปรากฏ ที่ตำแหน่งท้ายพยางค์ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.2.2 หน้าที่ของหน่วยเสียงพยัญชนะในภาษาไทย

เสียงพยัญชนะในภาษาไทย 21 หน่วยเสียงนี้สามารถทำหน้าที่ได้ตามตารางที่ 2.7

ตารางที่ 2.7 แสดงหน้าที่ของหน่วยเสียงพยัญชนะในภาษาไทย

หน้าที่	พยัญชนะ
พยัญชนะต้นเดี่ยว	/ ก / / ค / / ง / / ย / / บ / / ด / / ม / / น / / ร / / ล / / ว / / จ / / ช / / ซ / / ท / / ต / / ป / / พ / / ฟ / / อ / / ฮ /
พยัญชนะต้นควบ	หน่วยเสียงแรก / ก / / ค / / ต / / ป/ และ / พ / หน่วยเสียงที่สอง / ร / / ล / หรือ / ว /
พยัญชนะสะกดของพยางค์	/ ป / (แม่กบ) / ต / (แม่ กด) / ก / (แม่กด) / ม / (แม่กม) / ง / (แม่กง) / น / (แม่กน) / ว / (แม่เกอว) / ย / (แม่เกย) และ ไม่มีเสียงพยัญชนะ สะกด (แม่กา)

2.3.3 หน่วยเสียงวรรณยุกต์

หน่วยเสียงวรรณยุกต์ คือ ระดับเสียงสูง – ต่ำ ของคำในภาษาไทย เช่นเดียวกับภาษาจีนและภาษาอื่นๆ ที่เป็นภาษาคำโดด ซึ่งมีการกำหนดเสียงสูง – ต่ำ ไว้ตายตัวในแต่ละคำ ถ้าออกเสียงสูง – ต่ำ ผิดไปความหมายย่อมผิดตามไปด้วย

ในภาษาไทยหน่วยเสียงวรรณยุกต์เป็นหน่วยเสียงสำคัญ มีความหมายต่างกัน ดังนั้นอาจกล่าวได้ว่าหน้าที่ ของหน่วยเสียงวรรณยุกต์ ก็คือ การทำให้เกิดคำขึ้นใช้ในภาษามากขึ้น และเป็นวิธีการสร้างคำขึ้นใช้เพิ่มในภาษาเป็นวิธีแรก ทั้งนี้เพราะถ้าเราเปลี่ยนเสียงวรรณยุกต์ ก็จะทำให้คำเกิดความหมายเพิ่มขึ้นใหม่ นั้นเอง

เสียงสูง – ต่ำ ในภาษาพูด เกิดจากการสั่นสะเทือนของเส้นเสียงในอัตราต่างๆ กัน โดยเสียงที่เปล่งออกมาในขณะที่เส้นเสียงสั่นนั้น จะต้องเป็นเสียงก้อง ดังนั้นหน่วยเสียงวรรณยุกต์ ในภาษาไทยจึงจัดเป็นหน่วยเสียงซ้อน กับอักษรที่ใช้จึงเป็นรูปเครื่องหมายเขียนซ้อนข้างบนหน่วยเสียงสระ (ซึ่งเป็นเสียงก้อง) ซึ่งรูปวรรณยุกต์อยู่ 4 รูป แทนเสียงวรรณยุกต์ทั้งหมด 5 หน่วยเสียง โดยเสียงสามัญไม่มีรูปวรรณยุกต์

ซึ่งลักษณะของเสียงวรรณยุกต์ สามารถแบ่งออกตามลักษณะระดับเสียงได้ตามตารางที่ 2.8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.8 แสดงลักษณะของเสียงวรรณยุกต์ในภาษาไทย

หน่วยเสียง		สัญลักษณ์	เสียง	ลักษณะ
กลุ่มวรรณยุกต์ระดับ (Level tone)	หน่วยเสียงวรรณยุกต์ระดับ ต่ำ (Low tone)	/ \ /	เอก	หน่วยเสียงนี้จะปรากฏในพยางค์ของภาษาไทย ได้ทุกแบบ
	หน่วยเสียงวรรณยุกต์ระดับกลาง (Mid tone)	/ - /	สามัญ	หน่วยเสียงนี้จะไม่ปรากฏในพยางค์ ที่มีตัวสะกดเป็น พยัญชนะกัก (พยางค์คำตาย)
	หน่วยเสียงวรรณยุกต์ระดับสูง (High tone)	/ / /	ตรี	หน่วยเสียงนี้จะไม่ปรากฏในพยางค์ที่ผสมด้วยสระเสียงยาวซึ่งมีตัวสะกดเป็นเสียงกัก
กลุ่มวรรณยุกต์เปลี่ยนระดับ (Contour tone)	หน่วยเสียงวรรณยุกต์เปลี่ยนตก (Falling tone)	/ ^ /	โท	หน่วยเสียงนี้จะไม่ปรากฏในพยางค์ที่มีสระเสียงสั้น และมีเสียงพยัญชนะสะกดเป็นพยัญชนะกัก
	หน่วยเสียงวรรณยุกต์เปลี่ยนขึ้น (Rising tone)	/ v /	จัตวา	หน่วยเสียงนี้จะไม่ปรากฏในพยางค์ที่มีเสียงพยัญชนะสะกด เป็นเสียงกัก

2.4 ลักษณะ พยางค์ และคำในภาษาไทย

กาญจนา นาคสกุล (2520 : 104) ได้ให้ความหมายของพยางค์ในระบบเสียงภาษาไทยว่า "พยางค์ หมายถึง จำนวนเสียงที่ดังเด่นซึ่งปรากฏในกลุ่มเสียงที่เรียงกันเป็นคำพูด ส่วนเสียงอื่นๆ ที่อยู่ข้างเคียงก็จะประกอบกันเข้าเป็นส่วนของพยางค์" เสียงที่ดังเด่นในกลุ่มเสียงก็คือเสียงสระ ซึ่งมีลักษณะประจำตัวก็คือเป็นเสียงก้อง ซึ่งดังเด่นกว่าเสียงอื่นๆ ดังนั้นเสียงสระจึงมักเป็นเสียงที่ทำให้เกิดพยางค์ ถ้ามีเสียงสระเด่นอยู่กี่เสียง พยางค์ก็จะมีจำนวนเท่านั้นด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

63676

พยางค์ที่เปล่งออกมาครั้งหนึ่งๆ อาจมีความหมายหรือไม่ก็ได้ แต่เมื่อใดพยางค์ที่ประกอบขึ้นจากเสียงสระ พยัญชนะ และวรรณยุกต์ เป็นอย่างน้อยที่สุด และกลุ่มเสียงเหล่านี้มีความหมายและสามารถปรากฏได้โดยลำพังพยางค์นั้นๆ ก็กลายเป็นคำในภาษาไทย

คำในภาษาไทยส่วนใหญ่จะเป็นคำพยางค์เดียว ซึ่งเป็นคำพื้นฐาน (Base words) ของภาษา ภาษาไทยจึงจัดอยู่ในตระกูลภาษา คำโดด หรือ คำพยางค์เดียว (Monosyllabic language) หน่วยเสียงที่ประกอบกันเข้าเป็นพยางค์ จะต้องมีอย่างน้อย 3 หน่วยคือ หน่วยเสียงพยัญชนะต้น 1 หน่วย หน่วยเสียงสระ 1 หน่วย และหน่วยเสียงวรรณยุกต์ 1 หน่วย และมีหน่วยเสียงอย่างมากไม่เกิน 5 หน่วย คือ เพิ่มหน่วยเสียงพยัญชนะต้นที่เป็นเสียงควบกล้ำอีก 1 หน่วย และหน่วยเสียงพยัญชนะสะกด อีก 1 หน่วย โดยมีองค์ประกอบของหน่วยเสียงต่างๆ ในพยางค์ แสดงได้ ดังรูปที่ 2.6

		วรรณยุกต์	
พยัญชนะต้น	(ควบ)	สระ	(พยัญชนะสะกด)

รูปที่ 2.6 แสดงองค์ประกอบของหน่วยเสียงต่างๆ ในพยางค์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

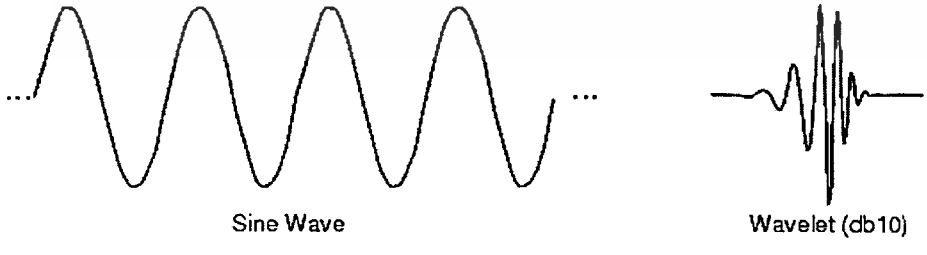
ความเป็นมาและทฤษฎีของเวฟเล็ท

บทนี้เป็นการอธิบายถึงหลักการพื้นฐานของเวฟเล็ทโดยกล่าวถึง การวิเคราะห์สัญญาณแบบหลายระดับความละเอียด (MRA : Multi Resolution Analysis) การแปลงเวฟเล็ทบนขอบเขตจำกัด (Discrete Wavelet Transform) การแจกจงสัมประสิทธิเวฟเล็ท ขบวนการแปลงเวฟเล็ทโดยหลักการ Analysis Filter Banks การแปลงกลับเวฟโดยหลักการ Synthesis Filter Banks ลักษณะสมบัติของฟังก์ชันสเกลลิงและฟังก์ชันเวฟเล็ทเบื้องต้น และอธิบายถึงหลักการของสเตชันนารีเวฟเล็ททรานสฟอร์ม (Stationary Wavelet Transform) เบื้องต้น

วิธีการในการวิเคราะห์สัญญาณ (Signal Analysis) มีอยู่หลายวิธีด้วยกันเช่น ฟูเรียร์ทรานสฟอร์ม (Fourier Transform) , เวฟเล็ททรานสฟอร์ม (Wavelet Transform) เป็นต้นซึ่งการที่จะเลือกนำวิธีการแบบใดมาใช้นั้นก็ขึ้นอยู่กับความเหมาะสมในการใช้งาน ฟูเรียร์ทรานสฟอร์มเป็นการทรานสฟอร์มที่สามารถทำการทรานสฟอร์มกลับได้ (Inverse Transform) นั่นคือสามารถที่จะทำการแปลงกลับไปกลับมาระหว่างสัญญาเดิม (Original Signal) และสัญญาณที่ผ่านการทรานสฟอร์มแล้ว (Transformed Signal) ได้อย่างไรก็ตามจะไม่มีข้อมูลเกี่ยวกับเวลา (Time Information) ในสัญญาณที่ผ่านการทรานสฟอร์มมาแล้ว นั่นก็คือฟูเรียร์ทรานสฟอร์มจะบอกได้ว่ามีความถี่แต่ละค่ามากน้อยเท่าไหร่ประกอบอยู่ในสัญญาณนั้น (Spectral Components) แต่บอกไม่ได้ว่าความถี่เหล่านี้เกิดในเวลาใดยกตัวอย่างเช่นการทำฟูเรียร์ทรานสฟอร์มสัญญาณ Stationary Signal ซึ่งประกอบด้วยความถี่ 10, 20 และ 50 Hz ตลอดเวลาและเปรียบเทียบผลลัพธ์จากฟูเรียร์ทรานสฟอร์มของ Non-stationary signal ซึ่งประกอบด้วยความถี่ 10 Hz ที่ช่วงเวลา 0-100 mS, 20 Hz จากช่วงเวลา 100-300 ms และ 50 Hz ในช่วงเวลา 300 mS-600 mS พบว่าสเปกตรัมสัญญาณที่ได้รับจากทั้งสองกรณ์มีความคล้ายคลึงกันแต่จะแตกต่างกันบ้างในรายละเอียดบางประการเช่นการกระเพื่อมของสัญญาณในบริเวณที่มีการเปลี่ยนจากความถี่ค่าหนึ่งไปยังอีกค่าหนึ่งในกรณีของ Non-stationary signal ดังนั้นจะเห็นได้ว่าการวิเคราะห์สัญญาณด้วยการแปลงฟูเรียนร์มีความแม่นยำด้านความถี่และเหมาะสมในการวิเคราะห์สัญญาณที่มีลักษณะเป็นคาบเวลาที่แน่นอน (Stationary Signal) หรือเมื่อไม่ต้องการทราบข้อมูลทางเวลาจากการวิเคราะห์ข้อมูลนั้น อีกทั้งการที่ข้อมูลซึ่งได้จากฟูเรียร์ทรานสฟอร์มมีตลอดย่านความถี่จาก $+\infty$ จนถึง $-\infty$ เมื่อเกิดการเปลี่ยนแปลงของสัญญาณอินพุทที่ขณะเวลาใด ๆ ก็มีผลกระทบตลอดย่านสเปกตรัมของสัญญาณด้วยและฟูเรียร์ทรานสฟอร์มยังให้ข้อมูลที่มากเกินความจำเป็นหากต้องการวิเคราะห์สัญญาณเพียงเฉพาะในช่วงเวลาหรือช่วงความถี่ใดความถี่หนึ่งเนื่องจากฟูเรียร์ทรานสฟอร์มไม่สามารถที่จะเลือกวิเคราะห์สัญญาณเฉพาะช่วงเวลาหรือช่วงความถี่ใด ๆ ได้ต่อมาจึงได้มีการพัฒนาการแปลงฟูเรียร์ในช่วงเวลาสั้น ๆ ขึ้น (Short-Time Fourier Transform: STFT) ซึ่งการทรานสฟอร์มแบบนี้สามารถที่จะพิจารณาเป็นการทำฟูเรียร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทรานสฟอร์มของสัญญาณคูณอยู่กับวินโดว์ฟังก์ชัน (window function) โดยสามารถกำหนดตำแหน่งเวลาและช่วงความถี่ในการวิเคราะห์ได้ ดังนั้นผลที่ได้จากการทรานสฟอร์มจึงแสดงถึงความถี่ต่าง ๆ ที่ประกอบอยู่ในสัญญาณที่วิเคราะห์และยังบอกถึงช่วงเวลา (time intervals) ที่ปรากฏอยู่ของความถี่เหล่านี้ด้วย อนึ่งจากคุณลักษณะของสัญญาณทั่วไปพบว่าสัญญาณที่ความถี่สูงจะมีคาบเวลาที่แคบในขณะเดียวกันสัญญาณความถี่ต่ำก็มีคาบเวลาที่กว้าง ดังนั้นในการวิเคราะห์สัญญาณช่วงเวลาและช่วงความถี่ที่ทำการวิเคราะห์ควรมีการเปลี่ยนแปลงที่สัมพันธ์กันคือช่วงความถี่สูงควรใช้ช่วงเวลาในการวิเคราะห์น้อยแต่ในช่วงความถี่ต่ำควรที่จะใช้เวลาในการวิเคราะห์ยาวขึ้นแต่สำหรับการแปลงฟูเรียร์ในช่วงเวลาสั้น ๆ แล้วยังคงเป็นวิธีการวิเคราะห์ซึ่งมีช่วงเวลาและช่วงความถี่ในการวิเคราะห์ที่คงที่ (Fixed resolution transform) ข้อมูลผลลัพธ์ที่ได้จึงอาจมากหรือน้อยไป [4-5] และหากในการวิเคราะห์เลือกใช้วินโดว์ฟังก์ชันแคบ (Narrow window) ก็จะให้ผลการวิเคราะห์ทางเวลาที่ดีแต่ผลการวิเคราะห์ทางความถี่ไม่ดี (good time resolution, poor frequency resolution) ในขณะที่การใช้วินโดว์ฟังก์ชันขนาดกว้าง (Wide window) กลับให้ผลการวิเคราะห์ทางความถี่ที่ดีแต่ผลการวิเคราะห์ทางเวลาที่ไม่ค่อยดีนัก (good frequency resolution, poor time resolution) เพราะฉะนั้นการเลือกวินโดว์ฟังก์ชันจึงขึ้นอยู่กับความเหมาะสมในการใช้งานต่อมาได้มีการพัฒนาเวฟเล็ททรานสฟอร์มขึ้นซึ่งเป็นการทรานสฟอร์มรูปแบบหนึ่งที่กระทำผ่านวินโดว์ฟังก์ชันและสามารถเลื่อนตำแหน่งและปรับเปลี่ยนช่วงเวลาและช่วงความถี่ที่จะวิเคราะห์ได้ การที่สามารถอธิบายสัญญาณหรือฟังก์ชันใด ๆ ได้ว่าเกิดจากการรวมเชิงเส้น (Linear Combination) ของฟังก์ชันมูลฐาน (Basis Function) เป็นผลให้สามารถทำการวิเคราะห์สัญญาณต่าง ๆ ได้สะดวกยิ่งขึ้น โดยเป็นที่ทราบกันว่าฟังก์ชันซายน์และโคซายน์เป็นฟังก์ชันพื้นฐานเวฟเล็ท (Wavelet Basis Function) ซึ่งเวฟเล็ทฟังก์ชันมีลักษณะที่สำคัญ ๆ คือมีการเปลี่ยแปลงอย่างต่อเนื่อง (Oscillatory) และมีขนาดลดลงสู่ศูนย์อย่างรวดเร็วโดยมีความแตกต่างจากฟังก์ชันซายน์และโคซายน์ ดังรูปที่ 2.1



รูปที่ 3.1 ฟังก์ชันซายน์และฟังก์ชันเวฟเล็ท

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1 การวิเคราะห์หลายระดับความละเอียด (Multi-Resolution Analysis): MRA

ก่อนที่จะอธิบายถึงทฤษฎีการวิเคราะห์แบบหลายระดับความละเอียดจะกล่าวถึงทฤษฎีที่เกี่ยวข้องในการวิเคราะห์ก่อนซึ่งก็คือ สเปซของเวกเตอร์ (Vector Space) หรือปริภูมิของฟังก์ชัน โดยกำหนดให้สเปซหรือปริภูมิที่สำคัญสำหรับขบานการวิเคราะห์สัญญาณคือ $L^2(R)$

สเปซของเวกเตอร์ถ้าพิจารณาความหมายในด้านการวิเคราะห์สัญญาณแล้วก็หมายถึงสเปซที่เกิดจากการรวมกันของเวกเตอร์พื้นฐาน (Basis Vector) โดยเวกเตอร์พื้นฐานนี้จะเป็นเซตของเวกเตอร์ที่เป็นอิสระต่อกันและอยู่ในเวกเตอร์สเปซดังนั้นเวกเตอร์ใด ๆ ที่สามารถเขียนให้อยู่ในรูปการรวมเชิงเส้น (Linear Combination) ของเวกเตอร์พื้นฐานได้ เวกเตอร์นั้นก็จะอยู่ในสเปซของเวกเตอร์เสมอ นั่นก็คือเวกเตอร์ใด ๆ สามารถที่จะเขียนเป็นการรวมเชิงเส้นของเวกเตอร์พื้นฐานที่ค่าสัมประสิทธิ์ต่าง ๆ ได้ ดังตัวอย่างแสดงในสมการที่ (3.1) โดยในที่นี้กำหนดให้ V เป็นเวกเตอร์ใด ๆ b แสดงถึงเวกเตอร์พื้นฐานและ υ แทนค่าสัมประสิทธิ์

$$V = \sum_k \upsilon^k \, bk \tag{3.1}$$

ในทำนองเดียวกันถ้าทำการพิจารณาสมการที่ (3.1) ในรูปของสัญญาณโดยแทนเวกเตอร์พื้นฐานด้วยฟังก์ชันฐาน $\varphi_k(t)$ และแทนเวกเตอร์ V ใด ๆ ด้วยสัญญาณ f(t) และ a_k แทนค่าสัมประสิทธิ์ ดังนั้นจะได้ว่าสัญญาณใด ๆ สามารถที่จะเขียนในรูปผลรวมเชิงเส้นของฟังก์ชันพื้นฐานได้ดังแสดงในสมการที่ (3.2)

$$f(t) = \sum_k a_k \varphi_k(t) \qquad ; k \in Z \tag{3.2}$$

โดยกำหนดให้ z แทนเซตของจำนวนเต็มและเมื่อให้สัญญาณ f(t) เป็นฟังก์ชันใด ๆ ที่อยู่ในปริภูมิ $L^2(R)$ แล้วจะได้ว่า $\int_R |f(t)|^2 \, dt < \infty$ โดยที่ R เป็นเซตของจำนวนจริง สำหรับการวิเคราะห์สัญญาณแบบหลายระดับความละเอียดนั้นจะมีลักษณะและข้อกำหนดที่สำคัญหลายประการด้วยกัน ดังจะได้กล่าวต่อไป

(1) ภายในปริภูมิ $L^2(R)$ จะมีการซ้อนทับกันของปริภูมิย่อย ๆ ซึ่งแสดงในสมการที่ (3.3)

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots \subset L^2(R) \tag{3.3}$$

นั่นคือ

$$V_j \subset V_{j+1} \qquad ; j \in Z \tag{3.4}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อกำหนดให้ V_j หมายถึงสเปซของเวกเตอร์ , j หมายถึงระดับความละเอียดของสัญญาณและจำนวนฟังก์ชันพื้นฐานที่ประกอบอยู่ในฟังก์ชันนั้น ๆ ดังนั้นถ้า j มีค่ามากขึ้นก็จะมีจำนวนของฟังก์ชันพื้นฐานมากขึ้น สัญญาณใด ๆ ก็ตามจะสามารถพิจารณาที่ระดับความละเอียดใด ๆ ก็ได้โดยที่ยังคงอยู่ภายใต้สัญญาณเดิม

(2) ฟังก์ชันทุกฟังก์ชันใน $L^2(R)$ สามารถที่จะประมาณให้อยู่ในสเปซของเวกเตอร์ที่ระดับความละเอียดใด ๆ ก็ได้ในลักษณะการโปรเจคชั่นสัญญาณใด ๆ ลงใน V_j การประมาณสัญญาณไปที่ระดับความละเอียดใดความละเอียดหนึ่งจะทำให้ข่าวสารของสัญญาณนั้น ๆ หายไปบางส่วนตามระดับความละเอียดนั้นแต่หากระดับความละเอียดเพิ่มขึ้นถึงอนันต์สัญญาณนั้นก็ยังคงลู่เข้าสู่สัญญาณต้นกำเนิดซึ่งสามารถที่จะแสดงได้ดังสมการที่ (3.5)

$$\text{Close } L^2 = L^2(R) \tag{3.5}$$

(3) ในการโปรเจคชั่นสัญญาณแต่ละครั้งที่ค่าของ j ลดน้อยลงจะมีผลให้สัญญาณที่ถูกทำการประมาณนั้นมีข่าวสารสูญหายไปมากขึ้น ดังนั้นรายละเอียดของสัญญาณจะน้อยลง ๆ จนลู่เข้าสู่ศูนย์ดังแสดงในสมการที่ (3.6)

$$\left(\bigcap_{j \in z} V_j \right) = \{ 0 \} \tag{3.6}$$

(4) เมื่อทำการโปรเจคสัญญาณจากระดับหนึ่งไปยังระดับที่ต่ำกว่าหนึ่งระดับจะส่งผลให้ข่าวสารหรือพลังงานของสัญญาณหายไปส่วนหนึ่ง นั่นคือเกิดการสูญหายของแถบความถี่ขณะสัญญาณไปบางส่วนซึ่งพลังงานในแบนด์ความถี่ที่หายไปดังกล่าวนี้จะถูกรวมใน Complementary Subspace (W_j) สามารถพิจารณาได้ดังสมการที่ (3.7)

$$V_j + W_j = V_{j+1} \tag{3.7}$$

(5) คุณสมบัติของ Two scale property ซึ่งเป็นการลดระดับความละเอียดของสัญญาณลงมาหนึ่งระดับกล่าวว่าสัญญาณที่อยู่ใน V_{j+1} จะมีความถี่เป็นสองเท่าของสัญญาณใน V_j ดังแสดงในสมการ (3.8)

$$f(t) \in V_j \Leftrightarrow f(2t) \in V_{j+1} \tag{3.8}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 ลักษณะของฟังก์ชันสเกลลิงและฟังก์ชันเวฟเล็ท

จำนวนของฟังก์ชันพื้นฐานที่ประกอบขึ้นเป็นสัญญาณใด ๆ ในแต่ละสเปซจะมีจำนวนแตกต่างกันโดยสัญญาณในสเปซที่ระดับความละเอียดสูงกว่าก็จะประกอบด้วยฟังก์ชันเพื้นฐานจำนวนมากกว่า การที่สัญญาณประกอบด้วยฟังก์ชันพื้นฐานจำนวนมาก หมายถึงสัญญาณมีระดับความละเอียดสูงซึ่งถ้ากำหนดให้ช่วงเวลาที่ใช้ในการวิเคราะห์สัญญาณในสเปซที่ระดับความละเอียดต่าง ๆ เหล่านี้คงที่แล้ว เมื่อทำการเปรียบเทียบจำนวนฟังก์ชันพื้นฐานของสัญญาณในแต่ละสเปซพบว่าสเปซที่ระดับความละเอียดสูงมีฟังก์ชันพื้นฐานบรรจุอยู่เป็นจำนวนมากกว่าในสเปซที่ระดับความละเอียดต่ำกว่าเมื่อพิจารณาในช่วงเวลาเท่ากันซึ่งหมายความว่าฟังก์ชันพื้นฐานของสัญญาณที่ระดับความละเอียดสูงกว่าจะมีความถี่มากกว่าฟังก์ชันพื้นฐานของสัญญาณนั้นเมื่อพิจารณาที่ระดับความละเอียดต่ำกว่าตัวอย่างเช่นฟังก์ชันพื้นฐานที่อยู่ใน V_{j+1} จะมีความถี่เป็นสองเท่าของฟังก์ชันพื้นฐานที่อยู่ใน V_j โดยฟังก์ชันพื้นฐานที่ประกอบกันเป็นสัญญาณการประมาณในสเปซที่ระดับความละเอียด V_j ใด ๆ นี้คือฟังก์ชันสเกลลิง (Scaling function: $\varphi(t)$ ซึ่งแสดงในสมการที่ (3.9)

$$\varphi_k(t) = \varphi(t-k); \varphi_k(t) \in L^2(R), k \in Z \tag{3.9}$$

จากสมการที่ (3.9) กำหนดให้ k คือตำแหน่งที่เกิดฟังก์ชั่นสเกลลิง เมื่อให้ปริภูมิย่อยของ $L^2(R)$ ที่มี $\varphi_k(t)$ เป็นฟังก์ชันมูลฐานแสดงดังสมการที่ (3.10) โดย $\varphi_k(t)$ ที่ตำแหน่งต่าง ๆ กันเหล่านี้จะมีคุณสมบัติตั้งฉากเชิงปกติซึ่งกันและกัน (Orthonormal Property) จากการที่ฟังก์ชันพื้นฐานในปริภูมิย่อยใด ๆ มีคุณสมบัติตั้งฉากเชิงปกติกันจะเป็นผลให้ Inner product ของฟังก์ชันพื้นฐานที่ตำแหน่งเดียวกันและระดับความละเอียดเดียวกันมีค่าเป็นหนึ่งดังแสดงในสมการที่ (3.11)

$$V_0 = \text{span}\{\varphi_k(t)\} \tag{3.10}$$

$$\langle \varphi_{j,k}\, \varphi_{j,l} \rangle = \delta_{k,l}, \; j,k,l \in Z$$

$$\delta_{k,l} \begin{cases} = 1, & k = l \\ 0, & k \neq l \end{cases} \tag{3.11}$$

การเพิ่มขนาดปริภูมิย่อยใด ๆ ทำได้โดยการเปลี่ยนขนาดของเวลของฟังก์ชันสเกลลิงดังนั้นจะสามารถกำหนดฟังก์ชันพื้นฐานสำหรับปริภูมิย่อยใด ๆ ได้ดังสมการที่ (3.12)

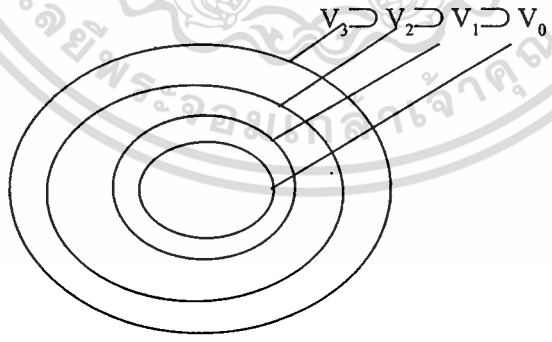
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\varphi_{j,k}(t) = 2^{j/2}\ \varphi(2^j t - k)\quad j,k \in Z \tag{3.12}$$

เมื่อ $2^{j/2}$ ในสมการที่ (2.12) มีไว้เพื่อเป็นการนอร์มอลไลซ์ (Normalized) ให้พลังงานของฟังก์ชันพื้นฐานทุกฟังก์ชันในทุกปริภูมิย่อยมีค่าเท่ากับหนึ่งเสมอ จากลักษณะสมบัติของการวิเคราะห์แบบหลายระดับความละเอียดทำให้สามารถทำการประมาณสัญญาณ f(t) ซึ่งอยู่ในปริภูมิ $L^2(R)$ ไปยังสเปซที่ระดับความละเอียดใด ๆ V_j ได้ดังแสดงในสมการต่อไปนี้

$$fj(t) = \sum_k c_j(k) \cdot \varphi_{j,k}(t) \tag{3.13}$$

เมื่อ c_j^k คือสัมประสิทธิ์หรือน้ำหนักที่คูณอยู่กับฟังก์ชันสเกลลิงที่ตำแหน่ง k ใด ๆ ค่าสัมประสิทธิ์ c_j^k จะหาได้จากสมการ $c_j^k = \langle f(t), \varphi_{j,k}(t) \rangle$ ซึ่งเป็นการหาค่าผลคูณภายในระหว่าง f(t) กับ $\varphi_{j,k}(t)$ แล้วประกอบเป็นสัญญาณที่ระดับความละเอียด V_j นั้น กำหนดให้ปริภูมิของเวกเตอร์มีการซ้อนทับกันดังสมการ $\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots \subset L^2(R)$ หรือ $V_j \subset V_{j+1}$; $j \in Z$ เมื่อ $V_{-\infty} = \{0\}$, $V_{\infty} = L2$ พบว่าปริภูมิของเวกเตอร์ที่บรรจุสัญญาณที่ระดับความละเอียดสูงก็จะบรรจุสัญญาณที่ระดับความละเอียดต่ำกว่าอยู่ด้วย นั่นคือฟังก์ชันพื้นฐานที่อยู่ในปริภูมิเวกเตอร์ V_j จะเป็นการสเกลของฟังก์ชันพื้นฐานที่อยู่ในปริภูมิเวกเตอร์ที่สูงกว่าซึ่งการซ้อนทับกันของปริภูมิเวกเตอร์แสดงในรูปที่ 3.2



รูปที่ 3.2 การซ้อนทับกันของปริภูมิเวกเตอร์ที่ระดับความละเอียดต่างๆ

กำหนดให้ $\varphi(t) \in V_0$ และ $V_0 \subset V_1$ แล้วฟังก์ชันสเกลลิงที่ระดับความละเอียด $j = 1$ คือ $2^{1/2}\ \varphi(2t - n)$ ดังนั้นความสัมพันธ์ระหว่างฟังก์ชันสเกลลิงที่ระดับความละเอียด V_0 และ V_1 แสดงดังสมการที่ (2.14) สัมประสิทธิ h(n) สามารถหาได้จากสมการ (3.15)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\varphi(t) = \sum_{n} h(n) \cdot 2^{1/2} \varphi(2t - n);\ n \in Z, h(n) \in L^2(Z) \qquad (3.14)$$

$$h(n) = \langle \varphi(t), \varphi(2t - n) \rangle ;\ n \in Z \qquad (3.15)$$

เมื่อ สัมประสิทธิ์ h(n) เป็นลำดับของจำนวนจริงหรือจำนวนเชิงซ้อนที่เรียกว่าสัมประสิทธิของสเกลลิงฟังก์ชันหรือสเกลลิงฟิลเตอร์หรือสเกิลลิงเวกเตอร์ (Scaling function coefficient or scaling filter or the scaling vector)

การวิเคราะห์สัญญาณที่ระดับความละเอียดต่ำลงมาจะมีผลให้พลังงานของสัญญาณบางส่วนหายไปอยู่ในอีกสเปซหนึ่งนั่นคือสเปซของเวกเตอร์เวฟเล็ท (Wavelet vector space: W_j) ซึ่งมีฟังก์ชันพื้นฐานคือฟังก์ชันเวฟเล็ท (Wavelet function: $\psi(t)$ การที่ปริภูมิ V_j และ W_j เป็นส่วนเติมเชิงตั้งฉากปกติซึ่งกันและกันแล้วความสัมพันธ์ระหว่างฟังก์ชันพื้นฐานสเกลลิงในปริภูมิ V_j และฟังก์ชันพื้นฐานเวฟเล็ทในปริภูมิ W_j จะสามารถแสดงได้ดังนี้

$$\langle \varphi_{j,k}(t), \psi_{j,l}(t) \rangle ; \int \varphi_{j,k}(t) \psi_{j,k}(t) dt = 0 \quad ; j, k, l \in Z \qquad (3.16)$$

ในทำนองเดียวกันกับฟังก์ชันสเกลลิงจะสามารถเขียนสมการของฟังก์ชันเวฟเล็ทที่ระดับความละเอียดต่าง ๆ ดังนี้

$$\psi_{j,k}(t) = 21/2 \cdot \psi(2^j t - k) ; \quad j,k,l \in Z \qquad (3.17)$$

โดยที่ฟังก์ชันพื้นฐานในปริภูมิย่อยเวฟเล็ทจะมีคุณสมบัติตั้งฉากเชิงปกติซึ่งกันและกันดังนี้

$$W_j = \text{span}\ \{\psi_{j,k}(t)\} ; \quad j,k,l \in Z \qquad (3.18)$$

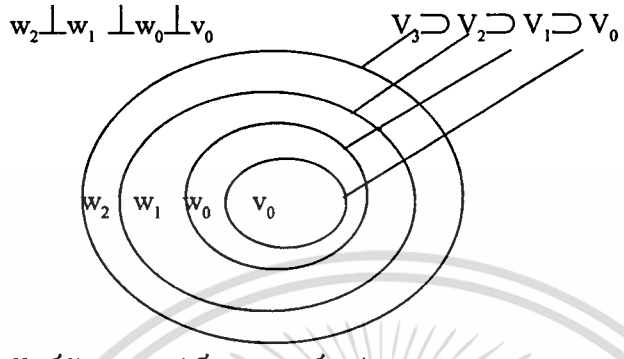
$$\langle \psi_{j,k}(t)\, \psi_{,j,l}(t) \rangle = \delta_{k,l} \ j,k,l \in Z$$

$$\delta_{k,l} \begin{cases} = 1, k = l \\ 0, k \neq l \end{cases} \qquad (2.19)$$

ความสัมพันธ์ระหว่างสเปซย่อยอธิบายได้จาก $\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots \subset L^2$ เห็นได้ว่าสามารถเริ่มอธิบายสัญญาณที่ระดับความละเอียดใด ๆ V_j ก็ได้และถ้าเริ่มที่ระดับความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ละเอียด j = 0 จะได้ว่า $V_0 \subset V_1 \subset V_2 \subset \cdots \subset L^2$ อาศัยคุณสมบัติของการวิเคราะห์แบบหลายระดับความละเอียดข้อที่ (4) คือ $V_j + Wj = V_{j+1}$ ดังนั้นจะสามารถเขียน $V_1 = V_0 \oplus W_0$ แล้วจะได้ว่า $V_2 = V_0 \oplus W_0 \oplus W_1$ นั่นคือ $L^2 = V_0 \oplus W_0 \oplus W_1 \oplus \ldots$ เมื่อให้ V_0 เป็นสเปซเริ่มต้นลักษณะความสัมพันธ์ระหว่างสเปซเวกเตอร์เวฟเล็ทสเกลลิงฟังก์ชันแสดงดังรูปที่ 3.3



รูปที่ 3.3 สเกลลิงฟังก์ชันและเวฟเล็ทเวกเตอร์สเปซ

โดยระดับความละเอียดของเวกเตอร์สเปซเวฟเล็ทจะเป็นที่ค่าใด ๆ ก็ได้ซึ่งถ้าระดับความละเอียดเริ่มที่ระดับ $j = -\infty$ แล้วจะได้ $L^2 = \ldots \oplus W_{-2} \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \ldots$ จะเห็นว่าประกอบด้วยเฉาะเทอมของเว็ฟเล็ทเวกเตอร์สเปซทั้งนี้เนื่องมาจากความสัมพันธ์ระหว่าง V_0 กับเวฟเล็ทเวกเตอรืสเปซดังนี้ $W_{-\infty} \oplus W_{-1} = V_0$ และจาก $W_0 \subset V_1$ สามารถเขียนได้เป็นผลรวมของค่าน้ำหนักของสเกลลิงฟังก์ชัน $\varphi(2t)$ ที่ถูกสเกลไปดังสมการที่ (3.20) และค่าสัมประสิทธิ์ g (n) หาได้จากสมการที่ (3.21)

$$\psi(t) = \sum_n g(n) .2^{1/2} \varphi(2t-n); \qquad n \in Z, g(n) \in L^2(Z) \qquad (3.20)$$

$$g(n) = \langle \psi(t), 2^{1/2} \varphi(2t-n); \qquad n \in Z \qquad (3.21)$$

โดยเทอมของ h(n) และ g(n) จะใช้สำหรับขบวนการลดจำนวนฟังก์ชันพื้นฐานลงครึ่งหนึ่งซึ่งมีผลให้ความถี่ของฟังก์ชันพื้นฐานลดลง

สำหรับสัญญาณใด ๆ gj(t) ที่เกิดจากการรวมกันของฟังก์ชันพื้นฐานเวฟเล็ท $\psi(t)$ ภายในสเปซของเวกเตอร์เวฟเล็ทW_j ที่ระดับความละเอียดใด ๆ สามารถเขียนได้ดังสมการที่ (3.22) เมื่อ d_j^k คือค่าสัมประสิทธิ์หรือค่าน้ำหนักที่คูณกับฟังก์ชันเวฟเล็ทที่ตำแหน่งที่ k นั้น ๆ ซึ่งหาค่าได้จากสมการที่ (2.23)

$$g^j(t) = \sum_n d_j^k \psi_{,j,l}(t) \qquad (3.22)$$

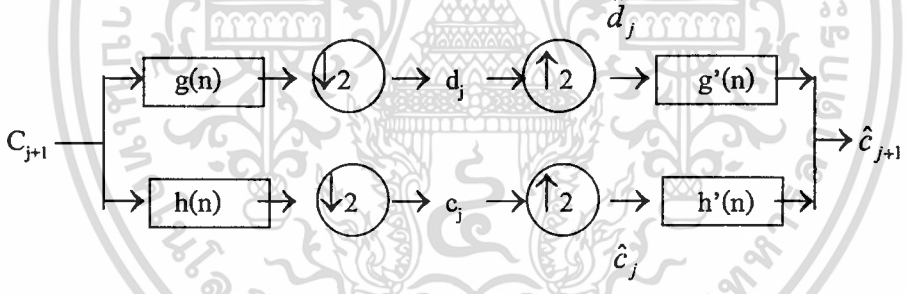
$$d_j^k = \langle g(t), \psi_{,j,l}(t) \rangle; \; j,k \in Z \qquad (3.23)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

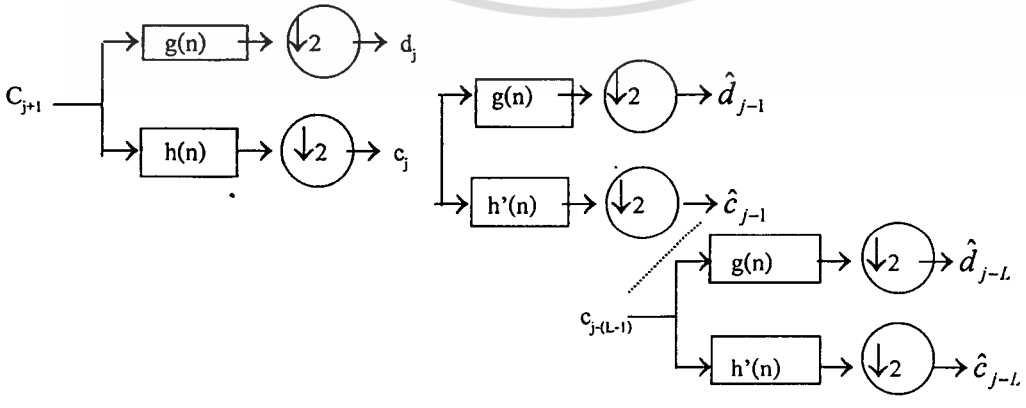
3.3 ตัวกรองฟิลเตอร์แบงก์ (Filter Bank)

ทฤษฎีที่สำคัญในการอธิบายการแปลงเวฟเล็ทแบบแบ่งช่วงจากหลักการวิเคราะห์แบบหลายระดับความละเอียดแล้ว หลักการของฟิลเตอร์แบงก์ (Filter Bank) ก็นับว่ามีบทบาทสำคัญหลักการหนึ่ง ดังนั้นจะได้อธิบายถึงหลักการของฟิลเตอร์แบงก์ในเบื้องต้นดังนี้

ฟิลเตอร์แบงก์แบบสองช่องสัญญาณ (Two Channel Filter Bank) [8] ดังแสดงในรูปที่ 2.4 ประกอบด้วยชุดตัวกรองในการวิเคราะห์ (Analysis Filter Bank) และชุดตัวกรองสำหรับการสังเคราะห์ (Synthesis Filter Bank) ตัวกรองในการวิเคราะห์ทำหน้าที่ในการแยกสัญญาณอินพุทเป็นสองส่วนคือความถี่สูงด้วยตัวกรองความถี่สูง g(n) และความถี่ต่ำด้วยตัวกรองความถี่ต่ำ h(n) ตามด้วย Decimators หรือ Downsampling สัญญาณ เมื่อนำฟิลเตอร์แบงก์แบบสองช่องสัญญาณมาเรียงต่อกันในลักษณะ Tree structure ก็จะได้ Octave analysis filter banks ดังภาพที่ 2.5 ฟิลเตอร์แบงก์แบบสองช่องสัญญาณชุดแรกจะแยกสัญญาณ C_{j+1} เป็นสัญญาณความถี่สูงครึ่งบน d_j และสัญญาณที่ความถี่ต่ำครึ่งล่าง C_j ออกจากกัน สัญญาณเอาท์พุท C_j เป็นส่วนที่ถูกแยกความถี่ต่อไป ดังนั้นส่วนของความถี่ต่ำที่ได้จากแต่ละขั้นตอนจะถูกทำการแยกแบนด์ต่อไปเรื่อย ๆ ตามระดับที่ต้องการ

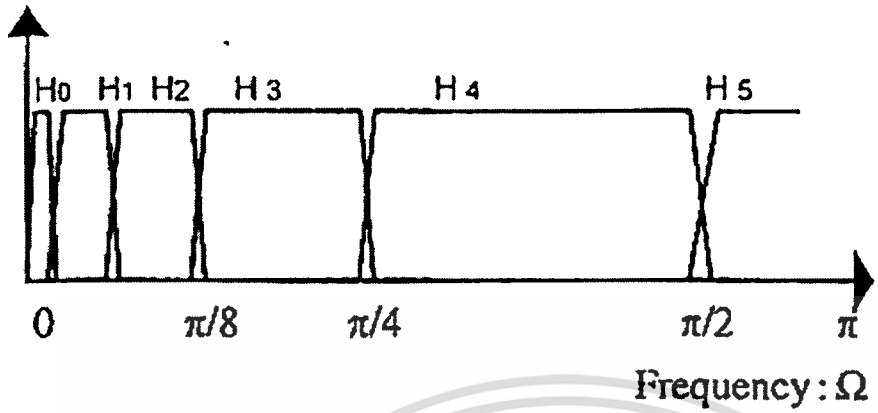


รูปที่ 3.4 ฟิลเตอร์แบงก์แบบสองช่องสัญญาณ

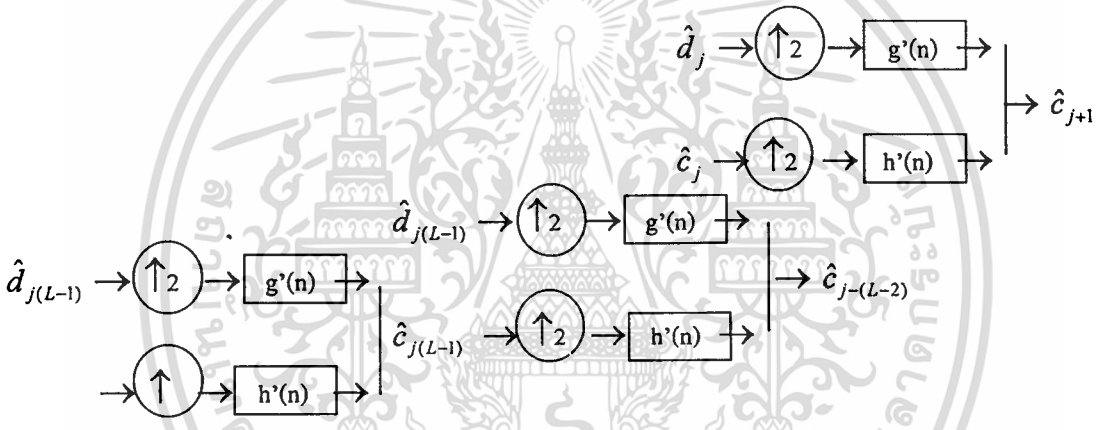


รูปที่ 3.5 Octave Band Analysis Filter Banks

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.6 ลักษณะการแยกแบนด์ความถี่ (Frequency Bands) ของ Octave Filter Banks



รูปที่ 3.7 Octave Band Synthesis Filter Banks

ในการแยกแบนด์ความถี่ของ Octave Filter Banks แบนด์ของความถี่ต่ำจะถูกแบ่งออกไปเรื่อย ๆ สัมพันธ์กับอัตราการสุ่มที่ลดลงครั้งละสองเท่าดังนั้นเอาท์พุทในส่วนของความถี่ต่ำจะลดลงในอัตราสองเท่าในแต่ละสเตทและความถี่คัทออฟ (cut off frequency) จะมีค่าลดลงครั้งละสองเท่าโดยลดลงไปทางด้านความถี่ที่ต่ำลง ทรานสเฟอร์ฟังก์ชัน $H_0(z)$ คือส่วนของความถี่ต่ำผ่านครั้งสุดท้ายในขณะที่ $H_1(z)$, $H_2(z)$, $H_3(z)$,$H_4(z)$ มีลักษณะเป็นแถบความถี่ผ่านและ $H_5(z)$ เป็นส่วนของความถี่สูงป่านแสดงดังรูปที่ 3.6 ดังนั้น Octave Filter Banks จึงสามารถแยกแบนด์ความถี่ต่าง ๆ ออกได้ (Band-separating Filter)

ฟิลเตอร์แบงก์ในการสังเคราะห์หรือสร้างกลับสัญญาณ (Synthesis Filter Banks) จะประกอบด้วย Expander หรือ Upsampling และตัวกรองในการสังเคราะห์ (Synthesis Filter) h'(n) กับ g'(n) ลักษณะของ Octave Banks Synthesis Filter Banks แสดงในรูปที่ 2.7 การสร้างกลับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัญญาณ (Reconstruction) เป็นการหาค่าของ C_{j+1} ตัวกรองในการวิเคราะห์และตัวกรองในการสังเคราะห์มีความสัมพันธ์กันในลักษณะของวงจรคลวอดราเจอร์มิเรอร์

3.4 การแปลงเวฟเล็ทแบบแบ่งช่วง (Discrete Wavelet Transform)

3.4.1 การแจกแจงสัมประสิทธิ์เวฟเล็ท (Wavelet Series Expansion)

การที่สัญญาณ $f(t) \in L^2(R)$ สามารถลดระดับความละเอียดลงได้และส่วนรายละเอียดของสัญญาณที่หายไปจะไปอยู่ในสเปซซึ่งเป็น (Orthogonal Complement กับ V_j ซึ่งคือ W_j ดังกล่าวมาแล้ว โดยรายละเอียดของ f(t) จะปรากฏอยู่ในทั้งสเปซ V_j และ W_j นั่นคือในสเปซ W_j ก็จะเกิดสัญญาณ $g_j(t)$ ที่เป็นส่วนประกบอของ f(t) ด้วยดังในสมการที่ (3.22) และค่าของ d_j^k ในสมการที่ (2.23) เป็นสัมประสิทธิ์ที่เกิดจากการโปรเจคชันเชิงตั้งฉาก (Orthogonal Projection) สัญญาณ f(t) ด้วย $\psi_{j,k}(t)$ นั้นเอง จากความสัมพันธ์ของ $V_j + W_j = V_j + 1$ จะได้ว่า $f_j+1 = f_j + g_i$ ถ้าให้ $f(t) \in V_{j+1}$ แล้วจะสามารถแตกกระจายสัญญาณให้มีความละเอียดน้อยลงได้โดยที่ V_j สามารถแตกต่อไปได้จนถึงระดับความละเอียด j = 0 แล้วจะได้ $V_{j+1} = V_0 \oplus W_0 \oplus W_1 \oplus \ldots \oplus W_j$

ดังนั้นจะเขียนความสัมพันธ์ของการแตกกระจายสเปซ (Decomposition Vector Spnce) ได้ดังสมการที่ (3.24) ซึ่งก็คือสัญญาณใด ๆ $f(t) \in L^2(R)$ จะสามารถแตกกระจายในลักษณะผลรวมของสัญญาณในสเปซย่อยที่ประกอบด้วยฟังก์ชันสเกลลิงกับสัญญาณในสเปซย่อยหลาย ๆ ชุดที่แต่ละชุดประกอบด้วยฟังก์ชันเวฟเล็ทที่มีการหดหรือขยายความถี่ (Dyadic Wavelet)

$$L^2(R) = V_j \oplus W_j \oplus W_{j+1} \oplus \ldots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \ldots \; j = Z \tag{3.24}$$

การโปรเจคชันสัญญาณ $f(t) \in L^2(R)$ ไปในสเปซย่อย ๆ มีความสัมพันธ์กับการแปลงเวฟเล็ทซึ่งเป็นการแจกแจงหาสัมประสิทธิ์ในแต่ละสเปซ (Wavelet Series Expansion) โดยสัญญาณ f(t) สามารถแจกแจงไปอยู่ในรูปแบบของฟังก์ชันสเกลลิงและฟังก์ชันเวฟเล็ทได้ดังต่อไปนี้

$$f'(t) = f_j + g_j + g_{j+1} + g_{j+2} + \ldots + g_\infty \tag{3.25}$$

$$f(t) = \sum_k c_j(k)\varphi_{j,k}(t) + \sum_{i=j}^{\infty}\sum_k d_i(k)\varphi_{j,k}(t)$$

จากสมการที่ (3.25) สัมประสิทธิ์ $c_j(k)$ และ $d_j(k)$ สามารถหาได้จากสมการการโปรเจคชันดังนี้ $c_j(k) = \langle f(t), \varphi_{j,k}(t) \rangle$ และ $d_j(k) = \langle f(t), \varphi_{j,k}(t) \rangle$ โดยพลังงานของสัญญาณ f(t) ที่ถูกแตกกระจายไปอยู่ในสเปซย่อยต่าง ๆ ยังคงเท่ากับสัญญาณต้นกำเนิด เนื่องจากสเปซย่อย V_j และ W_j มีลักษณะตั้งฉากซึ่งกันและกันและฟังก์ชันพื้นฐานที่อยู่ภายในแต่ละ สเปซนี้จะมีความสัมพันธ์เชิงตั้งฉากกันด้วยดังแสดงในสมการต่อไปนี้

เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\langle \varphi_{j,m}(t), \varphi_{j,n}(t) \rangle = \delta_{m,n} \quad (3.26)$$

$$\langle \psi_{i,m}(t), \psi_{i,n}(t) \rangle = \delta_{iln}\delta_{m,n} \quad (3.27)$$

$$\langle \varphi_{j,m}(t), \psi_{j,n}(t) \rangle = 0 \quad (3.28)$$

ดังนั้นการคำนวณหาค่ากลุ่มย่อยของสัมประสิทธิ์ $c_j(k)$, $d_j(k)$ (Coarse and Detail Expansion Coefficients) ก็คือวิธีการแปลงเวฟเล็ทแบบแบ่งช่วง (Discrete Wavelet Transform : DWT) นั่นเอง

3.4.2 การแปลงเวฟเล็ทโดยใช้หลักการ Analysis Filter Banks

ถ้า $f(t) \in V_{j+1}$ ซึ่งสามารถแสดงด้วยอนุกรมของฟังก์ชันสเกลลิงในปริภูมิย่อย V_{j+1} ได้ ดังนี้

$$f(t) = \sum_k c_j(k), \varphi_{j,k}(t) + \sum_k d_j(k) \psi_{j,k}(t) \quad (3.29)$$
$$= \sum_k c_j(k) 2^{j/2} \varphi(2^j t - k) + \sum_k d_j(k) 2^{j/2} \psi(2^j t - k)$$

แล้วสามารถพิจารณาสัมประสิทธิ์การสเกลที่ระดับ j ได้ว่า

$$c_j(k) = \langle f(t), \varphi_{j,k}(t) \rangle = \int f(t)\, 2^{j/2} \varphi(2^j t - k) dt \quad (3.30)$$

จากสมการ $\varphi(t) = \sum_n h(n) \cdot 2^{j/2} \varphi(2^j t - k)$; $n \in Z$, $h(n) = L^2(Z)$ ถ้าแทน t ด้วย t' โดยกำหนดให้ $t' = 2^j t - k$ ดังนั้นจากสมการที่ (2.30) จะได้

$$c_j(k) = \int f(t)\, 2^{j/2} \sum_k h(n) \sqrt{2\varphi}\ (2^j 2t - 2k - n)\, dt \quad (3.31)$$

$$c_j(k) = \sum_m h(m - 2k) \int f(t)\, 2^{(j+1)/2} \varphi(2^{(j+1)} t - m)\, dt \ ; \ m = 2k + n$$

ดังนั้นจากสมการที่ (3.30) และ (3.31) จะได้สมการสำหรับหาค่าของ $c_j(k)$ ดังนี้

$$c_j(k) = \sum_m h(m - 2k)\, c_{j+1}(m) \quad (3.32)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในทำนองเดียวกันสามารถคำนวณหาค่า $d_j(k)$ ได้ จาก

$$d_j(k) = \sum_m g(m-2k)\, c_{j+1}(m) \tag{3.33}$$

3.4.3 การแปลงกลับเวฟเล็ทโดยหลักการ Synthesis Filter Bank

การสร้างกลับสัญญาณที่ระดับความละเอียดสูงเป็นการรวมกันของสเกลลิงฟังก์ชันและเวฟเล็ทฟังก์ชันที่ระดับความละเอียดต่ำ การรวมสัญญาณในสเปซ V_j, W_j กลับไปเป็นสัญญาณในสเปซ V_{j+1} เป็นการเพิ่มรายละเอียดของสัญญาณ f(t) จากระดับ j ไปยังระดับ j+1 คล้ายกับขบวนการ Two channel Synthesis Filter Bank คือเป็นการรวมสัมประสิทธิ์ของฟังก์ชันสเกลลิงและฟังก์ชันเวฟเล็ทที่ระดับ j ไปเป็นสัมประสิทธิ์ฟังก์ชันสเกลลิงที่ระดับ j+1 โดยพิจารณาจากสัญญาณในสเปซ V_{j+1} ให้ $f(t) \in V_{j+1}$ โดยสัญญาณ f(t) นี้สามารถเขียนในเทอมของสเกลลิงฟังก์ชันได้ดังนี้

$$f(t) = \sum_k c_{j+1}(k), 2^{(j-1)/2} \varphi(2^{j+1}t - k) \tag{3.34}$$

หรือเขียนในเทอมของสเกลลิงฟังก์ชันที่ระดับความละเอียดต่ำลงมาได้เป็น

$$f(t) = \sum_k c_j(k), 2^{j/2} \varphi(2^{j+1}t - k) + \sum_k d_j(k) 2^{j/2} \psi(2^j t - k) \tag{3.35}$$

เมื่อแทนสมการที่ (3.20) และ (3.22) ลงในสมการที่ (3.35) จะได้

$$\begin{aligned} f(t) &= \sum_j c_j(k) \sum_n h(n)\, 2^{(j+1)/2} \varphi(2^{j+1}t - 2k - n) \\ &\quad + \sum_k d_j(k) \sum_n g(n)\, 2^{(j+1)/2} \varphi(2^{(j+1)}t - 2k - n) \end{aligned} \tag{3.36}$$

เนื่องจากทุกฟังก์ชันในที่นี้มีคุณสมบัติเชิงตั้งฉากเมื่อทำการคูณสมการที่ (3.34) และสมการที่ (3.36) ด้วยสมการ $\varphi(2^{j+1}t - k')$ แล้วทำการอินติเกตจะสามารถหาค่าของ c_{j+1} ได้ดังนี้

$$c_j(k) = \sum_m c_j(m)\, h(k-2m) + \sum_m d_j(m)\, h(k-2m) \tag{3.37}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5 สรุป

ทฤษฎีการวิเคราะห์สัญญาณแบบหลายระดับความละเอียดชี้ให้เห็นว่าสัญญาณใดๆ สามารถทำการวิเคราะห์ที่ระดับความละเอียดใด ๆ ก็ได้โดยยังคงอยู่ภายใต้สัญญาณเดิม และอาศัยหลักการของ Octave Filter Bank ที่แยกสัญญาณอินพุทออกเป็นแบนด์ความถี่สูงและแบนด์ความถี่ต่ำและนำสัญญาณในส่วนแบนด์ความถี่ต่ำไปแยกต่อไปเรื่อย ๆ จึงทำให้สามารถแยกแบนด์ความถี่ต่าง ๆ ออกได้ (Band separating filter) การแปลงเวฟเล็ทเป็นวิธีการคำนวณหาค่ากลุ่มย่อยของสัมประสิทธิ์ cj(k) , dj(k) (Coarse and Detail Expansion Coefficients) โดยอาศัยหลักการของการวิเคราะห์สัญญาณแบบหลายระดับความละเอียดและการวิเคราะห์สัญญาณด้วยฟิลเตอร์แบงก์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
วิธีการทดลอง

4.1วิธีการแยกเสียงพยัญชนะและสระ

การรู้จำเสียงพูดมีขั้นตอนการแยกเสียงพยัญชนะและ แยกเสียงสระออกจากกัน วิธีที่เราจะใช้แยกนั้นเราจะมาดูคุณสมบัติของสียงพูดออกมาเป็นคำๆในภาษาไทยลักษณะของเสียงแบ่งออกเป็นกำลังงานและความถี่ซึ่งเสียงของพยัญชนะจะมีเวลาการเกิดเสียงสั้นเมื่อเทียบกับเสียงสระใช้เวลานานและความถี่คงที่เราจึงสามารถแบ่งเสียงออกเป็นประเภทด้วยกำลังงานของเสียงดังรูป 4.1 คุณสมบัติของเสียงพูดออกมาเป็นคำๆเราพิจรณาลักษณะทางด้านกำลังงานของเสียงนั้นมี 4 แบบ โดยจะแสดงไว้ดังรูป โดยมีลูกศรเป็น ตัวขั้นกลางของ พยัญชนะและสระ

-แบบที่1 พลังงานพยัญชนะและสระมีค่าต่ำ ด้วยกันทั้งคู่

- **แบบที่ 2** พลังงานของพยัญชนะมีค่าต่ำ และพลังงานของ สระมีค่าสูง

-แบบที่ 3 พลังงานพยัญชนะและสระมีค่าสูง ต่ำด้วยกันทั้งคู่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบบที่ 4 พลังงานของพยัญชนะมีค่าสูงและพลังงานของสระ มีค่าต่ำ

รูป 4.1 แสดงรูปแบบ พลังงานพยัญชนะและสระ

วิธีการแยกเสียงพยัญชนะและเสียงสระมีหลายวิธีเช่น Neural Network Mode วิธีนี้ใช้การเขียนโปรแกรมสร้างแบบจำลองให้ครอบคลุมกับเงื่อนไขของเสียง ซึ่งได้เปอร์เซ็นต์ความถูกต้องของค่าเฉลี่ยอยู่ที่ 90 เปอร์เซ็นต์ วิธีการนี้ จะใช้ระยะเวลาในการคำนวณมากหรือน้อยขึ้นอยู่กับรายละ เอียดของแบบจำลองที่เราสร้างขึ้นมาและความแม่นยำนั้นขึ้น อยู่กับการสร้างแบบจำลองให้ครอบคลุมกับรายระเอียดของเสียงและข้อเสียหนึ่งคือต้องใช้เวลาในการจำแนกเสียงมากยิ่ง ขึ้นและอีกวิธีนั้นคือวิธี Zero-Crossing rate เป็นการใช้ จุดตัดของกำลังงานเสียงของพยัญชนะและเสียงสระได้เปอร์ เซ็นต์ความถูกต้องอยู่ที่ 95 เปอร์เซ็นต์ ซึ่งสามารถนำมาพัฒนาต่อให้ได้ความเร็วและความแม่นยำเพิ่มมากขึ้นโดยใช้การแปลงเวฟเลท

4.2 วิธีการทดลอง

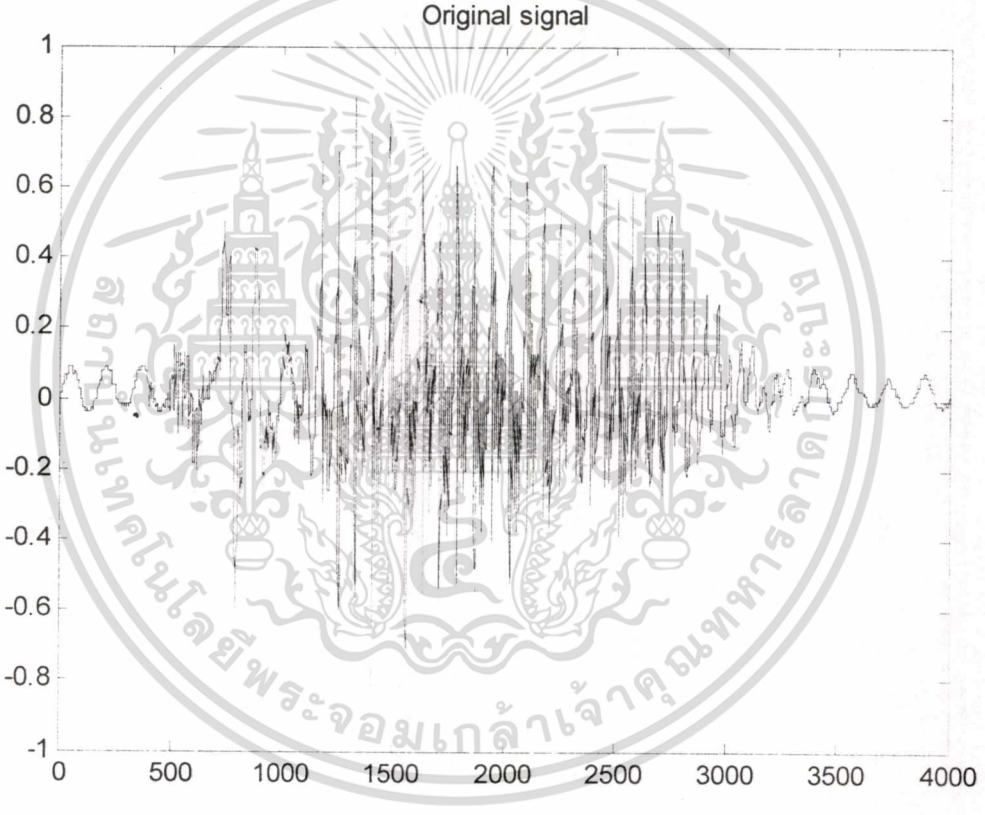
เริ่มจากการเก็บเสียงตัวอย่างที่พูดออกมาเป็นคำๆมาฟังเสียงว่าส่วนประกอบของเสียงพยัญชนะและเสียงสระว่ามีจุดเริ่มต้นและสิ้นสุดของคำว่าอยู่ตำแหน่งใดเพื่อมาบันทึกเก็บข้อมูลเอาไว้เปรียนเทียบกับเสียงที่ได้จากการคำนวน และการทดลองนี้ใช้โปรแกรมคำนวณทางคณิตศาสตร์ MATLAB เวอร์ชั่น 7.0.4 ทำงานบนเครื่องไมโครคอมพิวเตอร์ที่มีหน่วยประมวลผล "Celeron Processor" ความเร็ว 1.6 เมกกะเฮิร์ต และมีหน่วยความจำขนาด 256 เมกกะไบต์

4.2.1 การเก็บข้อมูลเสียง

สัญญาณเสียงพูดที่ใช้ในวิทยานิพนธ์นี้เป็นคำพยางค์เดียวที่ได้จากการเก็บตัวอย่างข้อมูลเสียงโดยใช้เครื่องคอมพิวเตอร์แบบพกพา การเก็บข้อมูลเสียงจะใช้คนทั้งหมด 8 คน ประกอบด้วยผู้ชาย 4 คน ผู้หญิง 4 คน โดยจะให้ผู้ทดสอบแต่ละคนอ่านออกเสียงเป็นคำๆที่ประกอบไปด้วยพยัญชนะ 21 รูปเสียง กับสระ 24 รูปเสียง ผู้ทดสอบแต่ละคนจะอ่านออกเสียงรวมคนละ 504 คำ 1 รอบ ดังนั้น

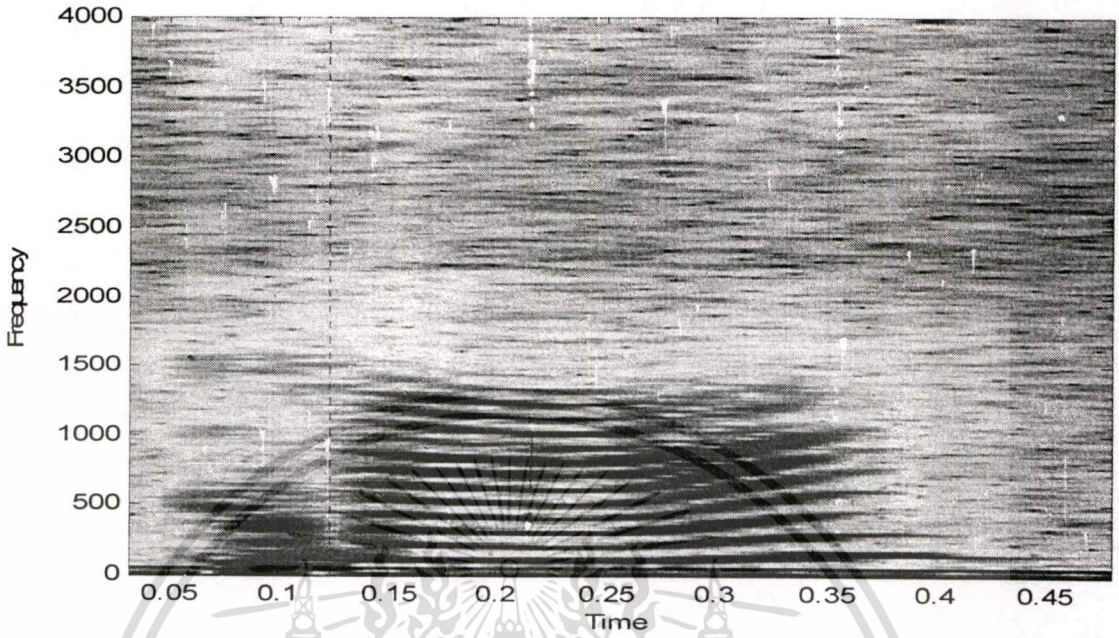
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิทยานิพนธ์นี้จะมีจำนวนคำที่ใช้ทั้งหมด 24 x 21 x 8 = 4,032 คำ ซึ่งจะนำไปสร้างแบบอ้างอิง จำนวน 4,032 คำ และสร้างแบบทดสอบจำนวน 4,032 คำ โดยการอ่านจะให้อ่านออกเสียงแบบไม่ต่อเนื่องคือจะให้อ่านแล้วหยุดสลับกันเพื่อให้คำที่พูดมีลักษณะไม่ต่อเนื่องกัน คำที่จะอ่านจะใช้เสียงสระ 1 เสียง แล้วเปลี่ยนเสียงพยัญชนะไปจนครบ 21 เสียง จากนั้นก็เปลี่ยนเสียงสระใหม่จนครบ 24 เสียง ซึ่งเสียงที่อ่านนั้นจะมีเสียงวรรณยุกต์รวมอยู่ด้วย ซึ่งข้อมูลจะถูกบันทึกอยู่ในรูปแบบของไฟล์ข้อมูล ".wav" ซึ่งข้อมูล 1 ตัวอย่างของเสียงจะถูกแทนด้วยข้อมูลขนาด 8 บิท โดยใช้ความถี่ในการแซมปลิ้ง เสียงเท่ากับ 8.0 กิโลเฮิร์ต และไฟล์ข้อมูล ".wav" นี้จะถูกใช้เป็นข้อมูลอินพุทสำหรับการคำนวณของโปรแกรมที่เขียนขึ้น



รูป ที่ 4.2 แสดงสัญญาณเสียงที่เก็บตัวอย่างสัญญาณเสียง คำว่า "ขา"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.3 แสดงสเปคโตแกรมของสัญญาณเสียง คำว่า "ขา"

จากสัญญาณเสียงของคำว่า "ขา" มีส่วนประกอบของเสียงพยัญชนะ "ข" อยู่ส่วนหน้าของเสียงพูดมีส่วนประกอบของความถี่เสียงต่างจากส่วนหลังซึ่งเป็นเสียงสระ "อา"

$$f_j(t) = \sum_m C_j(m)\phi_{jk}(t) + \sum_{i=j}^{\alpha}\sum_m D_j(m)\lambda\psi_{j,m}(t) \tag{4.1}$$

$$C_j(m) = \left\langle f(t), \phi_{j,m}(t) \right\rangle \tag{4.2}$$

$$D_j(m) = \left\langle f(t), \psi_{j,m}(t) \right\rangle \tag{4.3}$$

สมการที่ (4.1) เป็นการหาค่าสัมประสิทธิ์การประมาณค่าของสัญญาณเสียงซึ่งจะเก็บส่วนค่าสัมประสิทธิ์ที่สำคัญ สมการที่ 4.2 เป็นการเก็บค่าสัมประสิทธิ์รายละเอียดของสัญญาณเสียงมีความถี่สูง จากนั้นนำค่าสัมประสิทธิ์ทั้งสองมาคูณกันดังสมการที่ 4.4

$$P_j(m) = \frac{C_j(m) \cdot D_j(m)}{2^{-2j}} \tag{4.4}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากสมการที่ 4.5 หาค่า E(n) Short time Energy กำหนดให้ M เป็น 40 แล้วเพิ่มค่าจนหมดข้อมูล n

$$E(n) = \sum_{t=n-M+1}^{n} |P(t)| \tag{4.5}$$

จากสมการที่ 4.6 เมื่อค่า E(n) ผลของสัญญาณที่ได้ไม่ราบเรียบจึงนำสัญญาณผ่านวงจรกรองความถี่ต่ำสัญญาณ y(n) ในการทดลองจะกำหนดความถี่อยู่ใน 0~30 เฮิร์ต

$$y(n) = b(1)*E(n) + b(2)*E(n-1) + \ldots + b(nb+1)*E(n-nb) - a(2)*y(n-1) - \ldots - a(na+1)*y(n-na) \tag{4.6}$$

จากสมการที่ 4.7 เมื่อค่า Ps คือจุดเริ่มต้นของเสียงและ Pe จุดสิ้นสุดของ เสียงผลการคำนวณจะได้ระดับ Threshold มาใช้ในการชี้ตำแหน่งรอยต่อของเสียงพยัญชนะและเสียงสระ

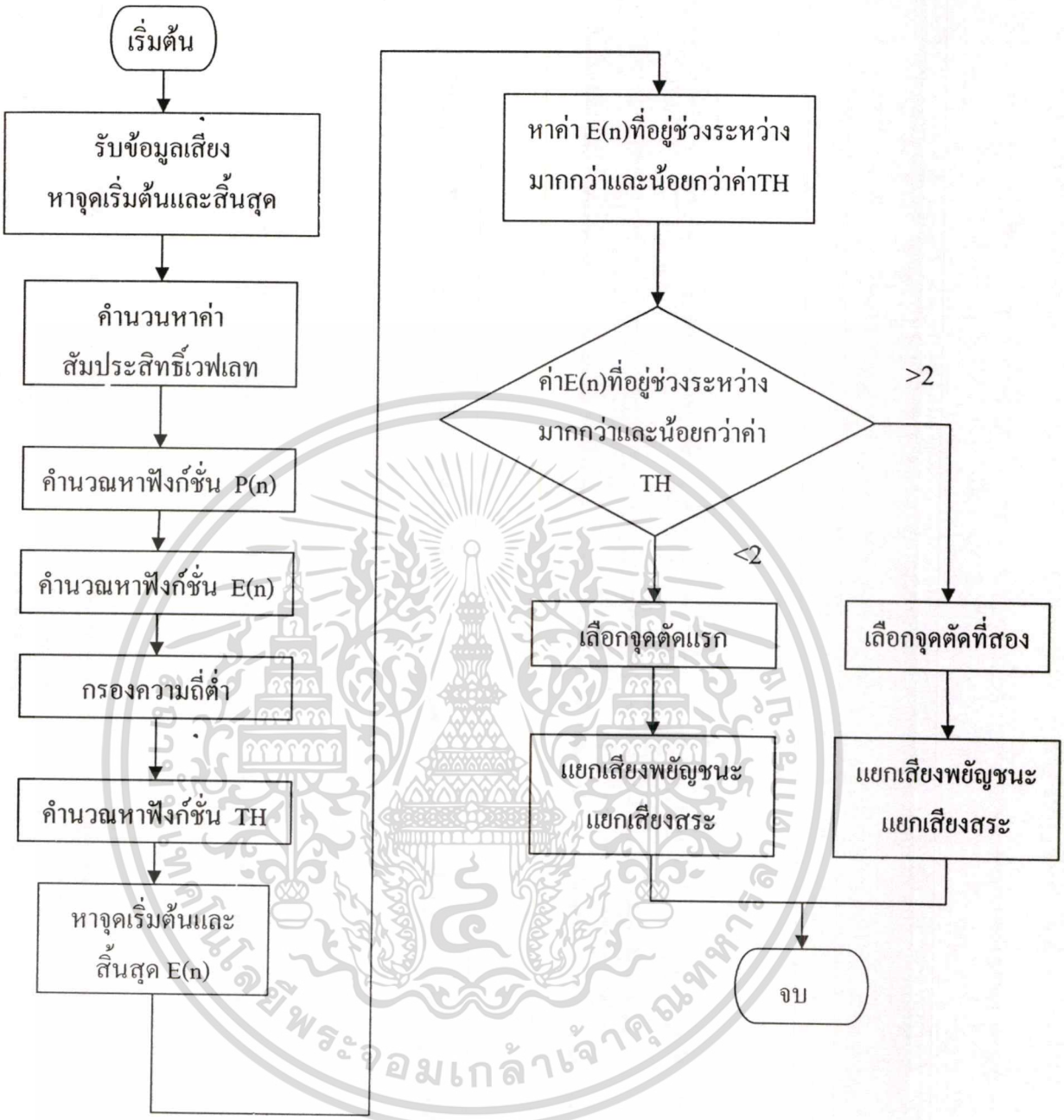
$$Th = \sum_{t=Ps}^{Pe} \frac{|P(t)|}{(Pe - Ps)} \tag{4.7}$$

จากผลการคำนวณได้ค่าจุดตัดของสัญญาณเสียงพยัญชนะและเสียงสระจะคิดแยกออกเป็นสี่แบบตามรูปแบบของพลังงานการเกิดของเสียงโดยเสียงนั้นจะแยกจุดขอบขาขึ้นของระดับพลังงานตัดของระดับ Threshold ถ้ามีจุดตัดดังกลางหนึ่งจุดเลือกจุดนั้นแต่ถ้ามีจุดตัดเป็นสองจุดเลือกจุดที่สอง

4.2.2 ขั้นตอนการในการแยกเสียงพยัญชนะและเสียงสระ

จากผังการทำงานเริ่มจากการอ่านข้อมูลเสียงไฟล์ข้อมูล " ****.wav " เก็บค่าเริ่มต้นและสิ้นสุดของข้อมูลไฟล์เสียง คำนวณหาค่าสัมประสิทธ์ของเวฟเลท เตทที่ 2 จะเก็บค่าสัมประสิทธ์การประมาณค่า และสัมประสิทธ์ลายละเอียดของเสียงเข้าสมการผลคูณคำนวณหาค่ากำลังงานแล้วกรองความถี่ต่ำผ่านหาจุดเริ่มต้นและสิ้นสุดของกำลังงาน ตรวจสอบนับจุดค่ากำลังงานเทียบกับค่า Threshold ว่าค่ากำลังงานในช่วงขาขึ้งที่อยู่ระหว่างค่า Threshold มีหนึ่งจุดถือว่าจดนั้นเป็นจุดรอยต่อของเสียงพยัญชนะ และถ้ามีสองจุดจุดที่สองเป็นจุดรอยต่อของเสียงพยัญชนะและเสียงสระเมื่อได้ผลดังกล่าวแล้วมาเทียบกับกราฟสเปคโตแกรมของสัญญาณเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



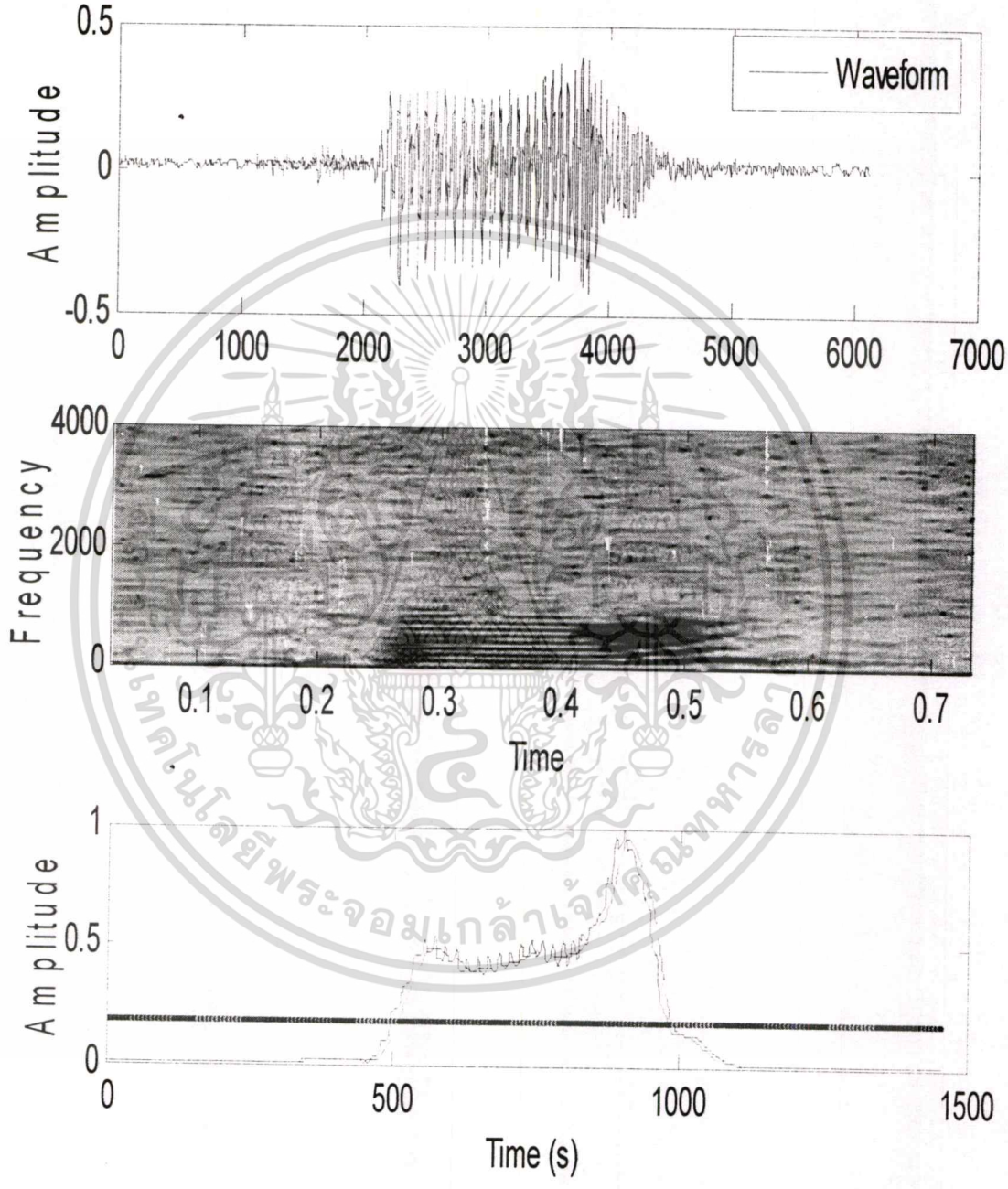
รูปที่ 4.4 ผังการทำงานของการแยกเสียงพยัญชนะเสียงสระ

4.3 ผลการทดลอง

การทดลองโดยนำเสียงที่เก็บตัวอย่างแบ่งแยกออกเป็นกลุ่มของลักษณะการกำเนิดเสียง พยัญชนะ เสียงของคำในการพูดที่ประกอบไปด้วยพยัญชนะตามด้วยสระนั้นจุดรอยต่อจะมีขอบเขต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

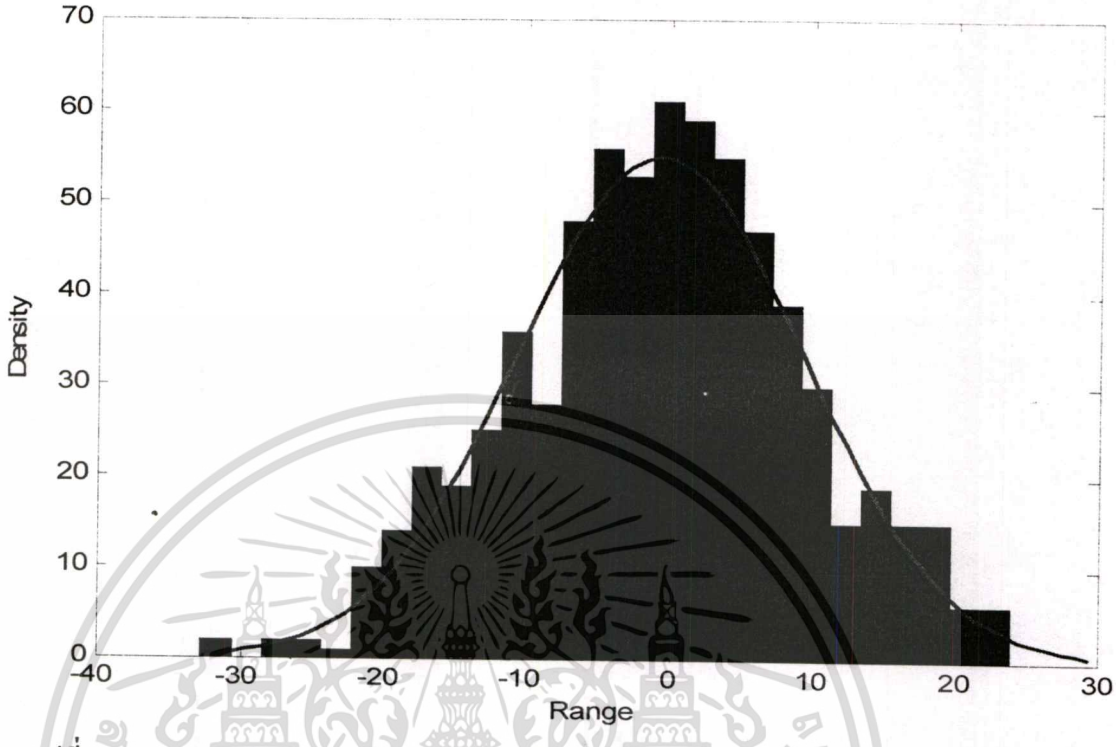
ของแต่ละกลุ่มมีขอบเขตไม่เท่ากันเสียงของพยัญชนะเสียงประเภทเสียงระเบิดจะมีค่ากำลังงานของเสียงพยัญชนะต่ำดังนั้นจุดตัดของเสียงพยัญชนะและเสียงสระจะอยู่ในจุดแรกช่วงของรอยต่อของเสียงพยัญชนะ



รูปที่ 4.5 เสียงพยัญชนะเสียงระเบิดตามด้วยสระ คำว่า "ขอ"

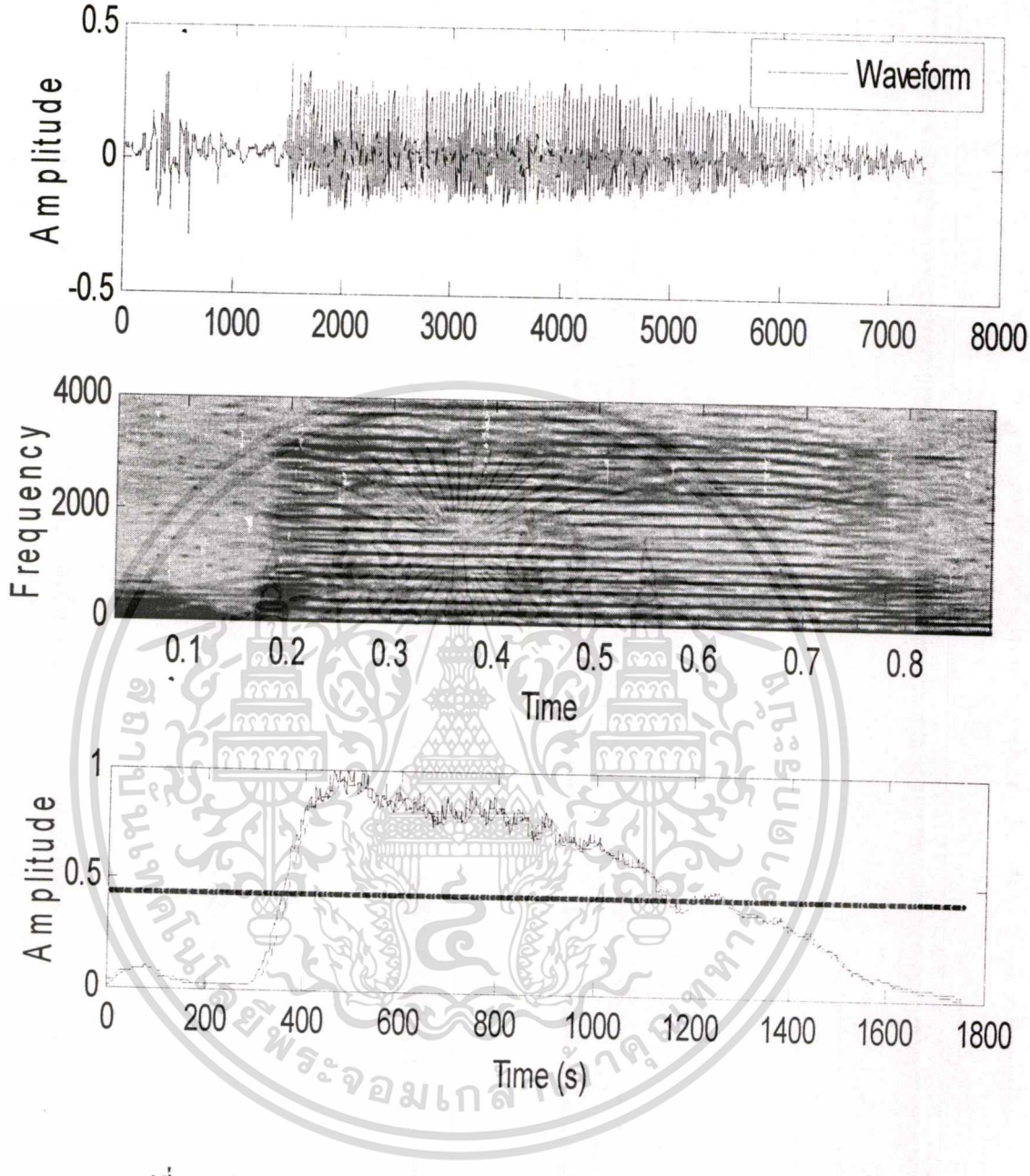
ในรูปที่ 4.5 แสดงตัวของของเสียงระเบิดของคำว่า "ขอ" รูปแรกเป็นเสียงที่เราได้บันทึกรูปที่สองเป็นสเปคโตแกรมของสัญญาณเสียง รูปล่างเป็นแสดงจุดรอยต่อของเสียงพยัญชนะเสียงสระ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.6 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงระเบิดตามด้วยเสียงสระ จากการทดลองของเสียงพยัญชนะของเสียงระเบิดทั้งหมดเก็บค่าแล้วมาหาค่าช่วงกว้างของรอยต่อของเสียงพยัญชนะเสียงสระ

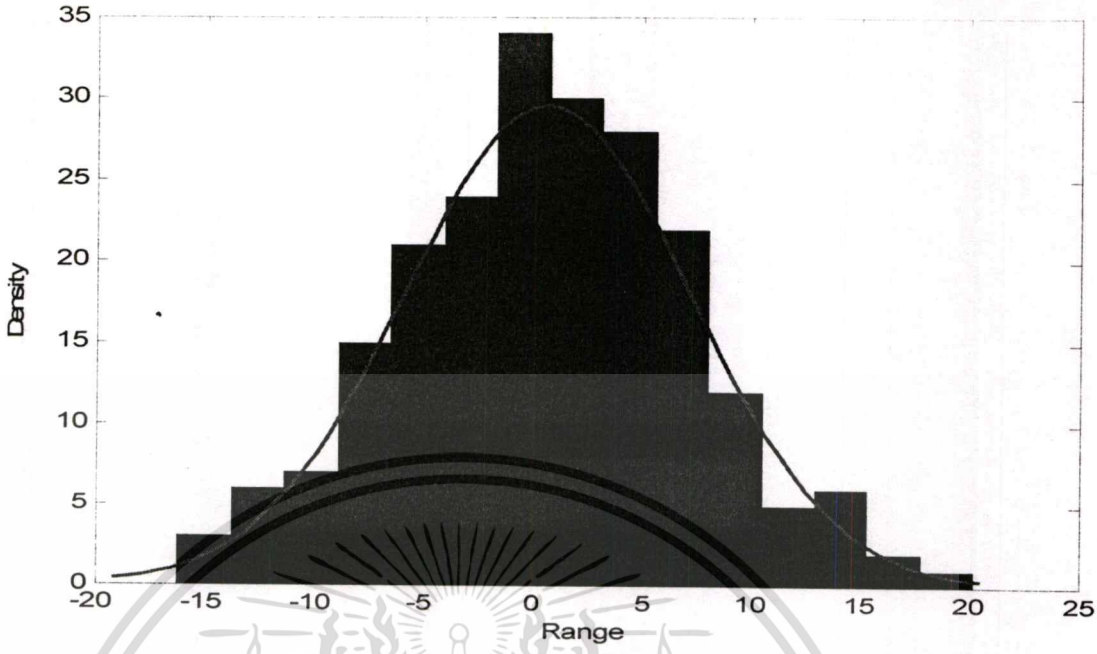
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.7 เสียงพยัญชนะเสียงเสียดแทรกตามด้วยสระ คำว่า "แซะ"

เสียงพยัญชนะเสียงเสียดแทรกตามด้วยสระนั้นมีจุดแยกที่เห็นชัด จากการเกิดของเสียงเสียดแทรกทำให้ความถี่ของเสียงพยัญชนะและเสียงสระนั้นมีความถี่ที่ต่างกันจึงทำให้รอยต่อของเสียงเด่นชัดจึงทำให้แยกได้ง่ายกว่าเสียงพยัญชนะกว่าประเภทอื่น ในรูปที่ 4.7 รูปแรกแสดงสัญญาณเสียงของคำว่า "แซะ" เสียงพยัญชนะจะมีความถี่ต่ำและเกิดชั่วขณะส่วนของเสียงสระมีคาบเวลาและความถี่ที่คงที่ รูปล่างจะเป็นการหาจุดต่อของเสียงพยัญชนะและเสียงสระ

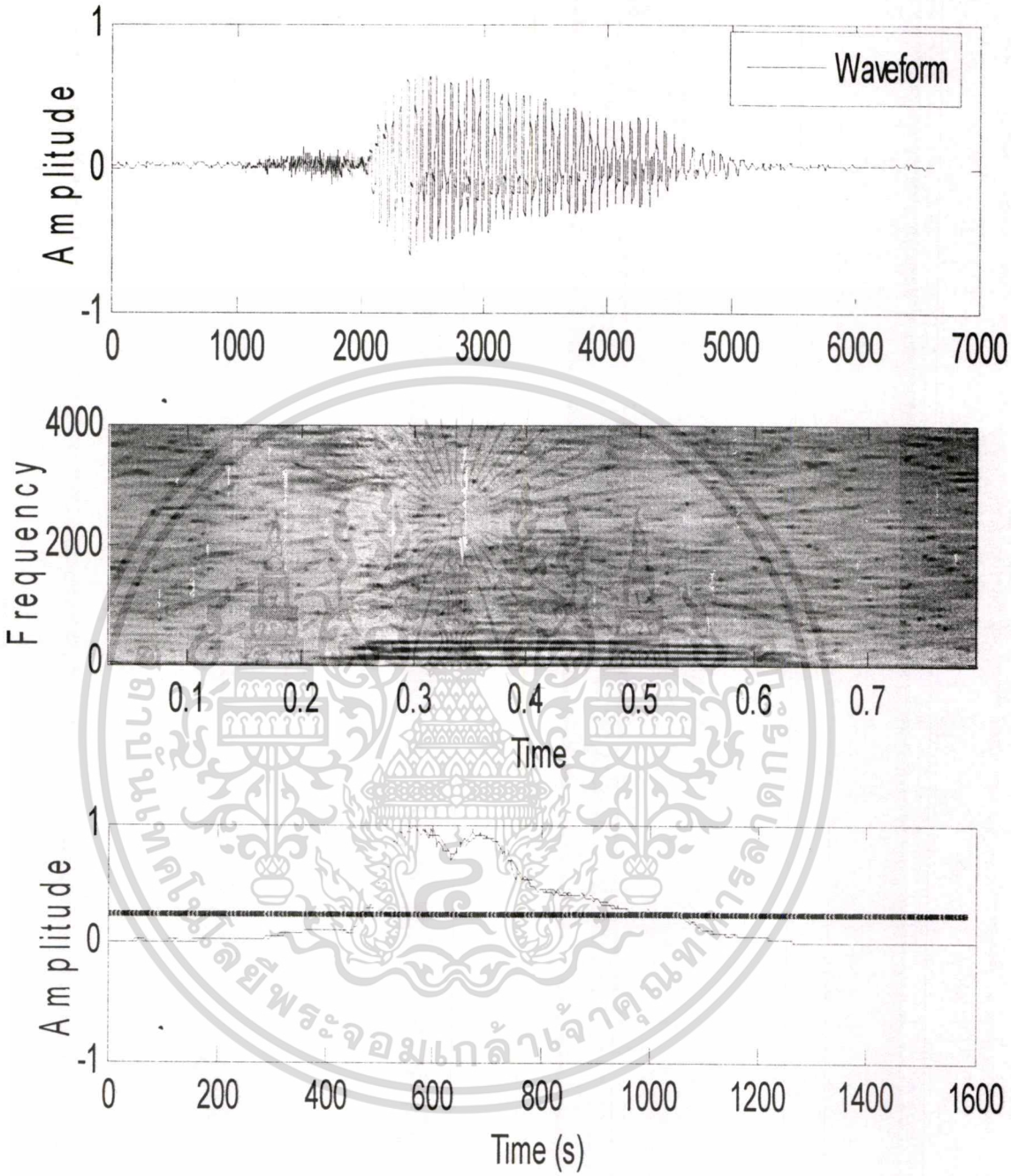
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.8 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงเสียดแทรกตามด้วยเสียงสระ

เมื่อนำเสียงพยัญชนะเสียงเสียดแทรกตามด้วยเสียงสระทั้งหมดเก็บช่วงกว่างของเสียงพยัญชนะและเสียงสระ ช่วงรอยต่อของเสียงจะแคบ

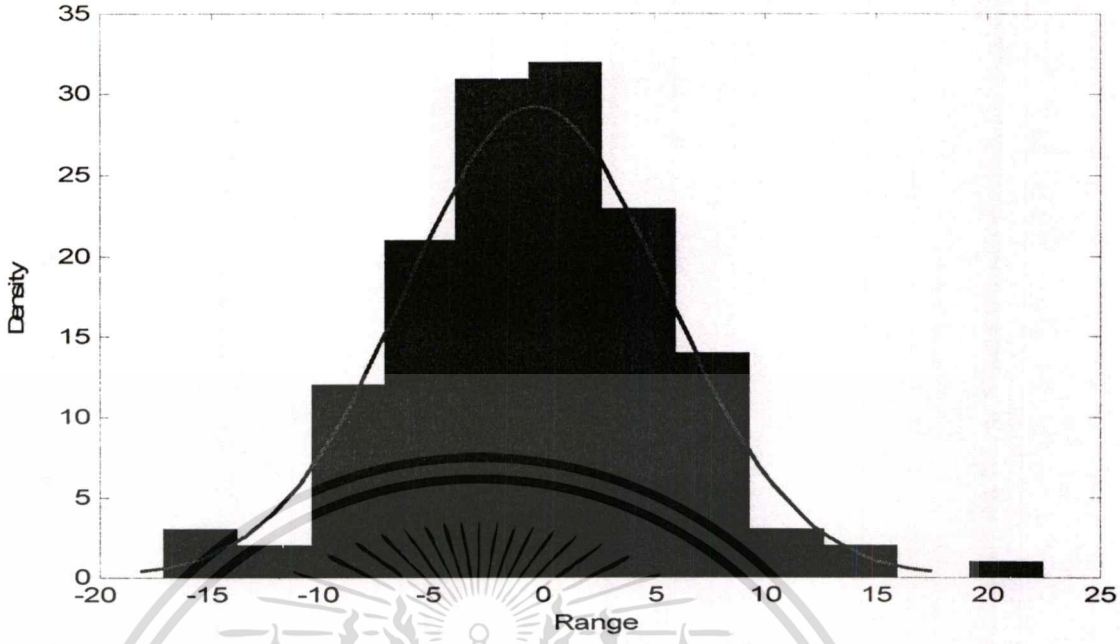
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.9 เสียงพยัญชนะเสียงกึ่งเสียดแทรกตามด้วยสระ คำว่า "ชี"

เสียงพยัญชนะเสียงกึ่งเสียดแทรกตามด้วยสระมีจุดแยกที่เห็นชัด จากการเกิดของเสียงกึ่งเสียดแทรกทำให้ความถี่ของเสียงพยัญชนะสูงและเสียงสระนั้นมีความถี่ต่ำจึงทำให้รอยต่อของเสียงเด่นชัดจึงทำให้แยกได้ง่าย ในรูปที่ 4.9 รูปแรกแสดงสัญญาณเสียงของคำว่า "ชี" เสียงพยัญชนะจะมีความถี่สูงและเกิดชั่วขณะส่วนของเสียงสระมีคาบเวลาและความถี่ต่ำที่คงที่ รูปล่างจะเป็นการหาจุดต่อของเสียงพยัญชนะและเสียงสระ

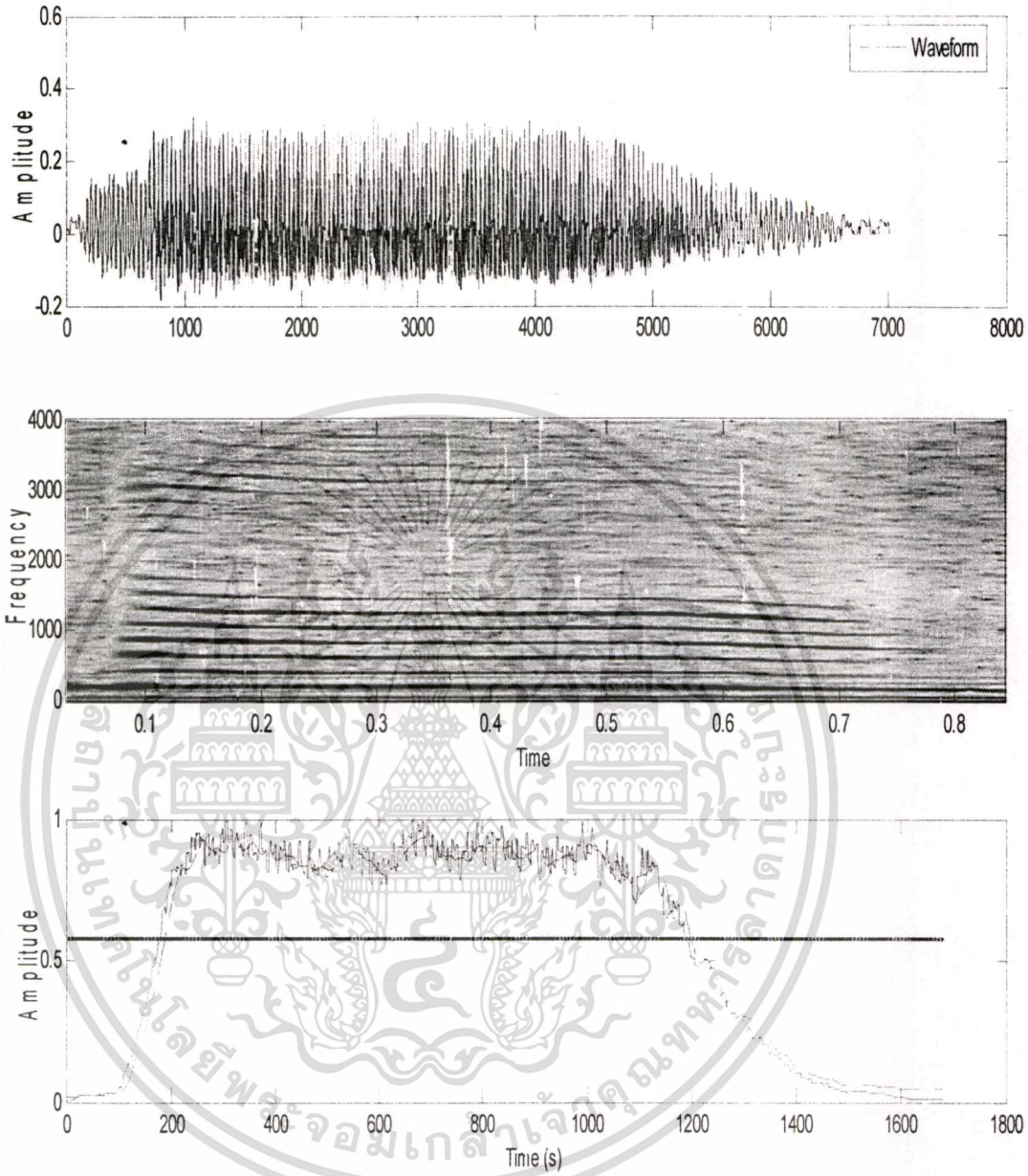
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.10 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงกึ่งเสียดแทรกตามด้วยเสียงสระ

เมื่อนำเสียงพยัญชนะเสียงกึ่งเสียดแทรกตามด้วยเสียงสระทั้งหมดเก็บช่วงกว่างของเสียงพยัญชนะ และเสียงสระ ช่วงรอยต่อของเสียงจะแคบ

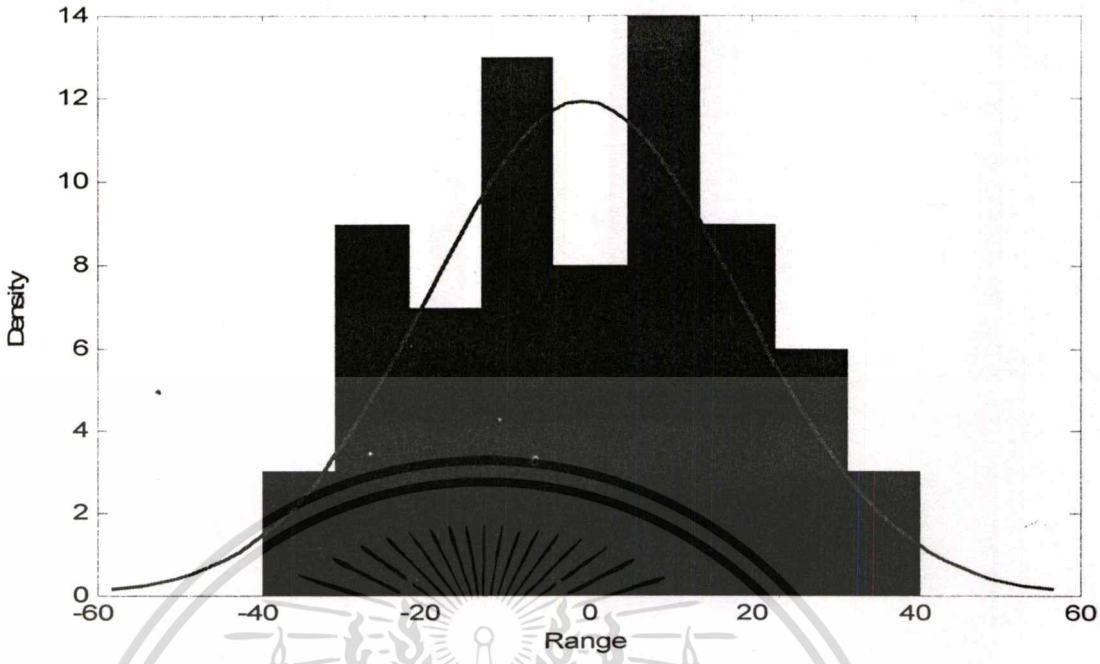
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.11 เสียงพยัญชนะเสียงนาสิกตามด้วยสระ คำว่า “นอ”

เสียงพยัญชนะเสียงนาสิกตามด้วยสระมีจุดแยกที่ไม่ชัดเจนเนื่องจากการเกิดของเสียงเสียงนาสิกทำให้ความถี่ของเสียงพยัญชนะกลมกลืนกับเสียงสระจึงทำให้รอยต่อของเสียงไม่เด่นชัดจึงทำให้แยกได้ยาก ในรูปที่ 4.11 รูปแรกแสดงสัญญาณเสียงของคำว่า “นอ” เสียงพยัญชนะจะมีความถี่ต่ำและเกิดชั่วขณะส่วนของเสียงสระมีคาบเวลาและความถี่สูงที่คงที่ รูปล่างจะเป็นการหาจุดต่อของเสียงพยัญชนะและเสียงสระเส้นปะจะเป็นส่วนที่ผ่านวงจรกรอง

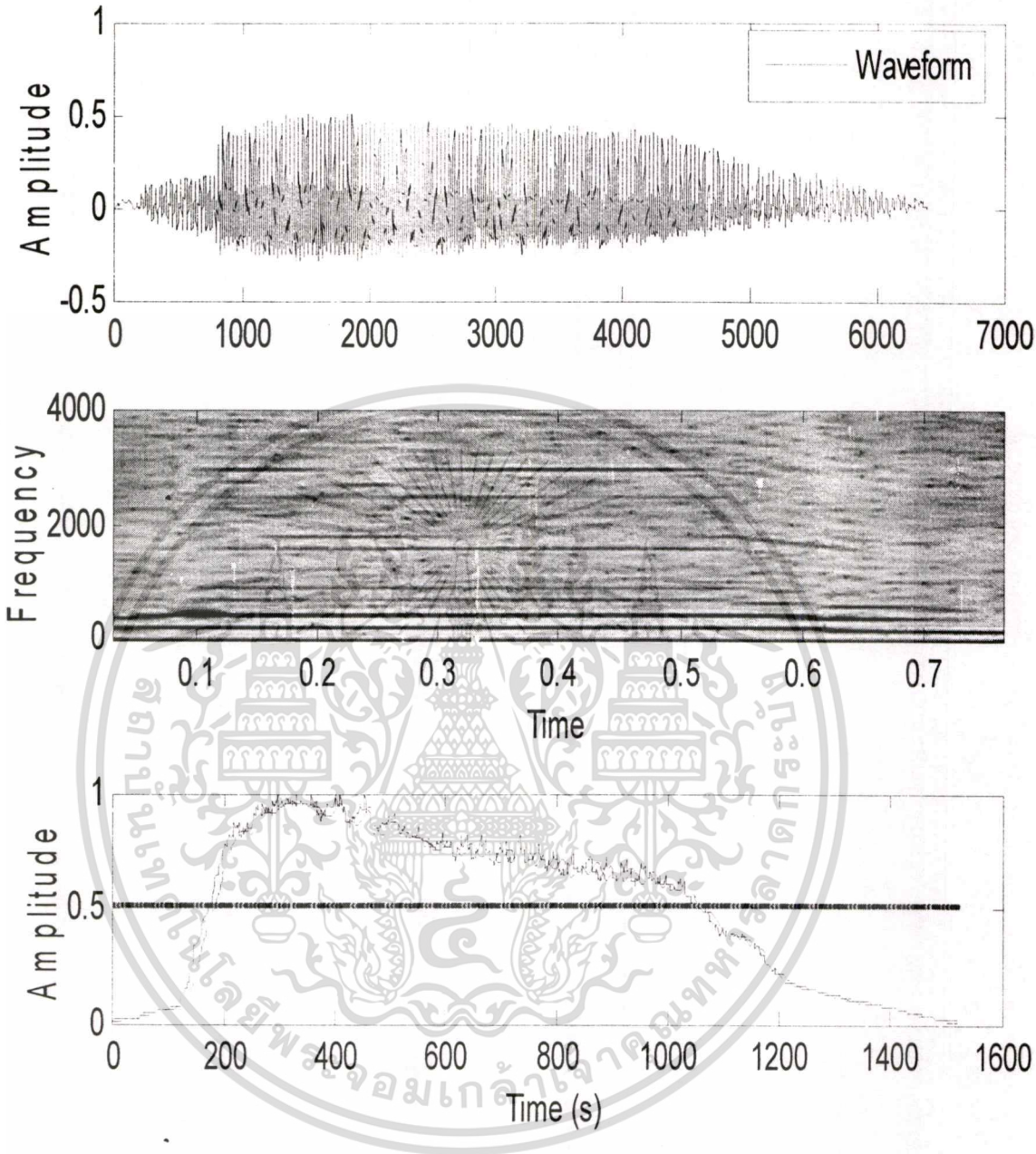
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.12 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงนาสิกตามด้วยเสียงสระ

เมื่อนำเสียงพยัญชนะเสียงนาสิกตามด้วยเสียงสระทั้งหมดเก็บช่วงกว่างของเสียงพยัญชนะและเสียงสระ ช่วงรอยต่อของเสียงจะกว้างเนื่องจากเสียงทั้งสองผสมกลมกลืนกัน

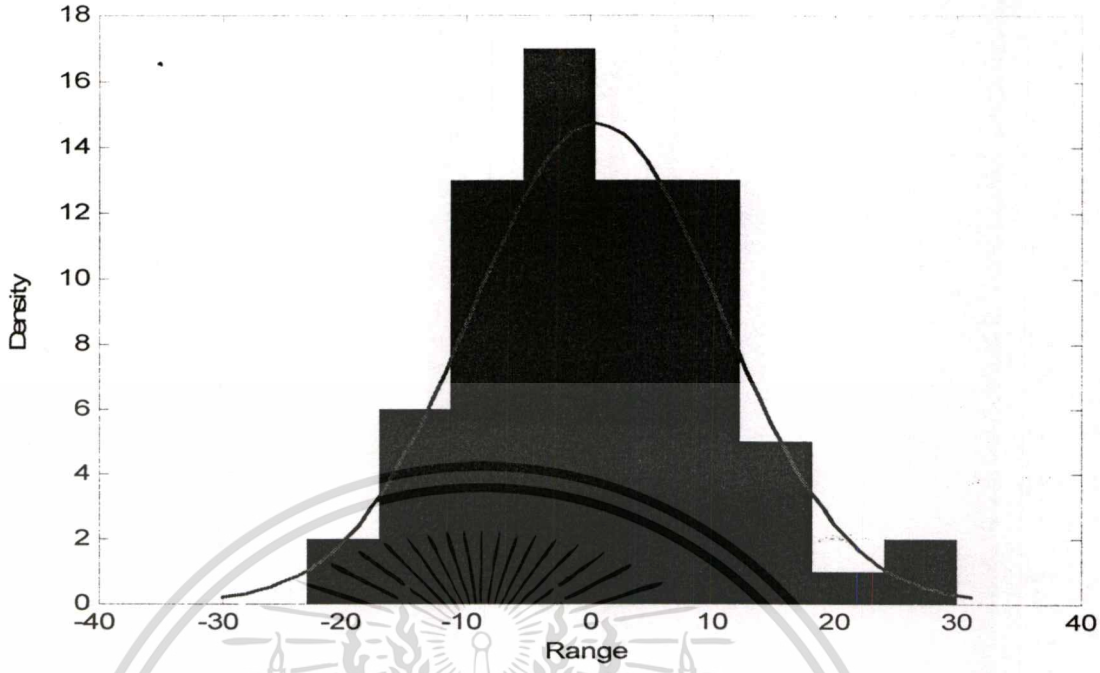
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.13 เสียงพยัญชนะเสียงข้างตามด้วยสระ คำว่า “ลือ”

เสียงพยัญชนะเสียงข้างตามด้วยสระมีจุดแยกที่ไม่ชัดเจนเนื่องจากการเกิดของเสียงเสียงข้างทำให้ความถี่ของเสียงพยัญชนะกลมกลืนกับเสียงสระจึงทำให้รอยต่อของเสียงไม่เด่นชัดจึงทำให้แยกได้ยาก ในรูปที่ 4.13 รูปแรกแสดงสัญญาณเสียงของคำว่า “ลือ” เสียงพยัญชนะจะมีความถี่ต่ำและเกิดชั่วขณะส่วนของเสียงสระมีคาบเวลาและความถี่สูงที่คงที่ รูปล่างจะเป็นการหาจุดต่อของเสียงพยัญชนะและเสียงสระเส้นปะจะเป็นส่วนที่ผ่านวงจรกรอง

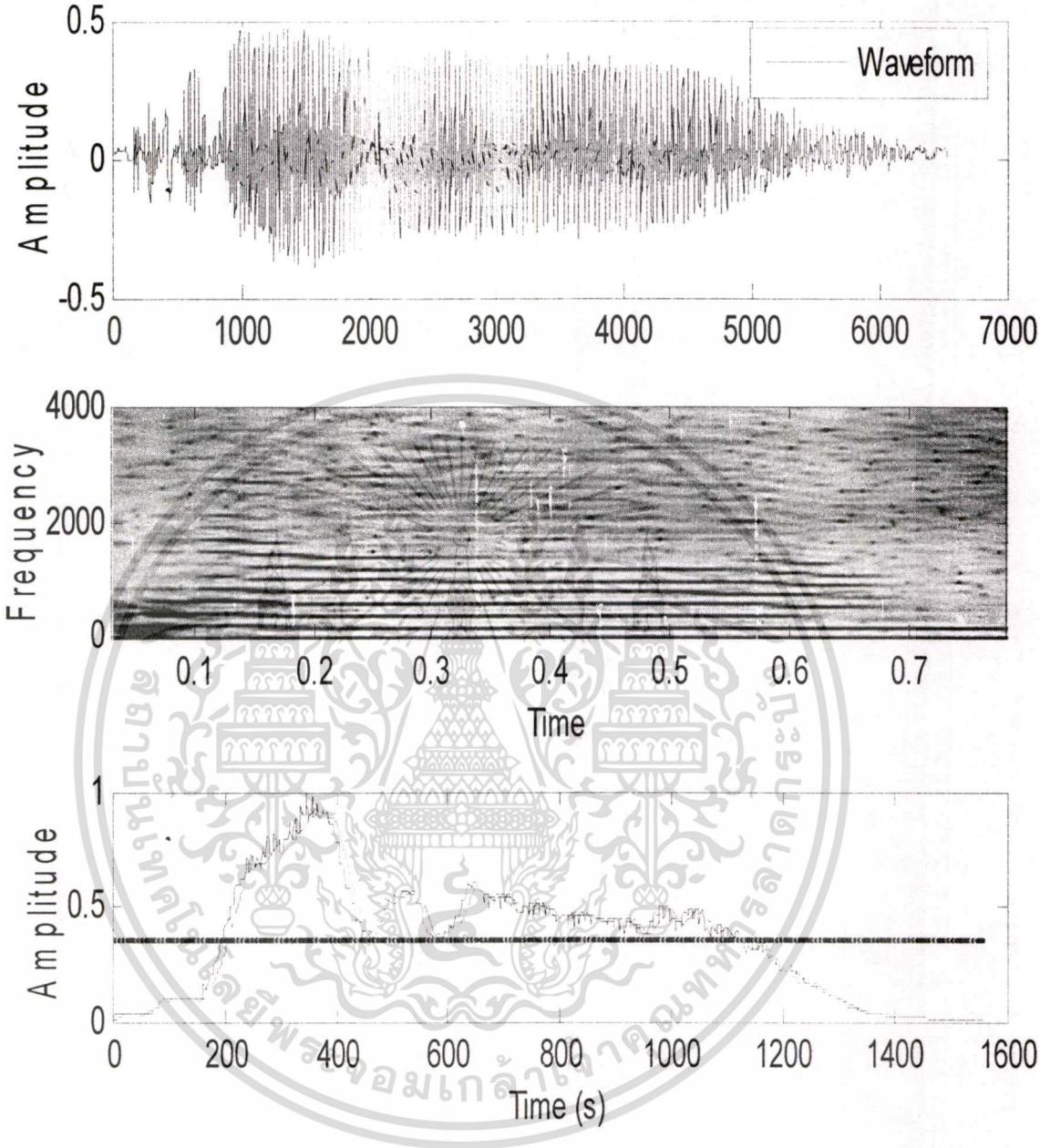
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.14 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงข้างตามด้วยเสียงสระ

เมื่อนำเสียงพยัญชนะเสียงข้างตามด้วยเสียงสระทั้งหมดเก็บช่วงกว่างของเสียงพยัญชนะและเสียงสระ ช่วงรอยต่อของเสียงจะกว้างเนื่องจากเสียงทั้งสองผสมกลมกลืนกัน

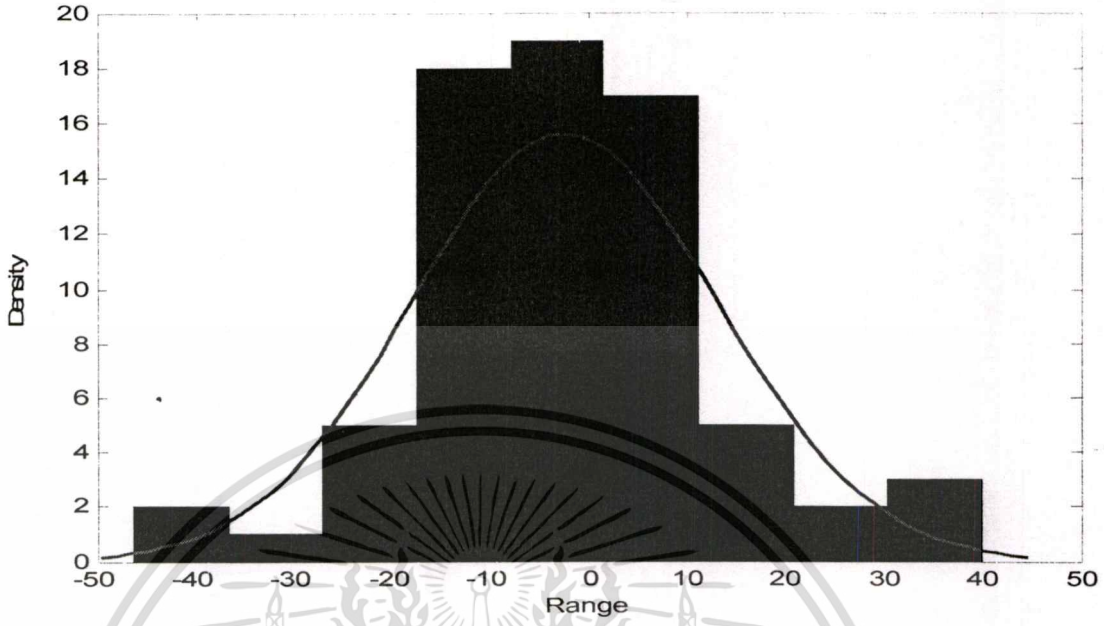
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.15 เสียงพยัญชนะเสียงรัวตามด้วยสระ คำว่า “รอ”

เสียงพยัญชนะเสียงรัวตามด้วยสระมีจุดแยกที่ไม่ชัดเจนเนื่องจากการเกิดของเสียงเสียงรัวทำให้ความถี่ของเสียงพยัญชนะถี่ใกล้เคียงกับเสียงสระจึงทำให้รอยต่อของเสียงไม่เด่นชัดจึงทำให้แยกได้ยาก ในรูปที่ 4.15 รูปแรกแสดงสัญญาณเสียงของคำว่า “รอ” เสียงพยัญชนะจะมีความถี่ต่ำและเกิดชั่วขณะส่วนของเสียงสระมีคาบเวลาที่คงที่ รูปล่างจะเป็นการหาจุดต่อของเสียงพยัญชนะและเสียงสระเส้นปะจะเป็นส่วนที่ผ่านวงจรกรอง

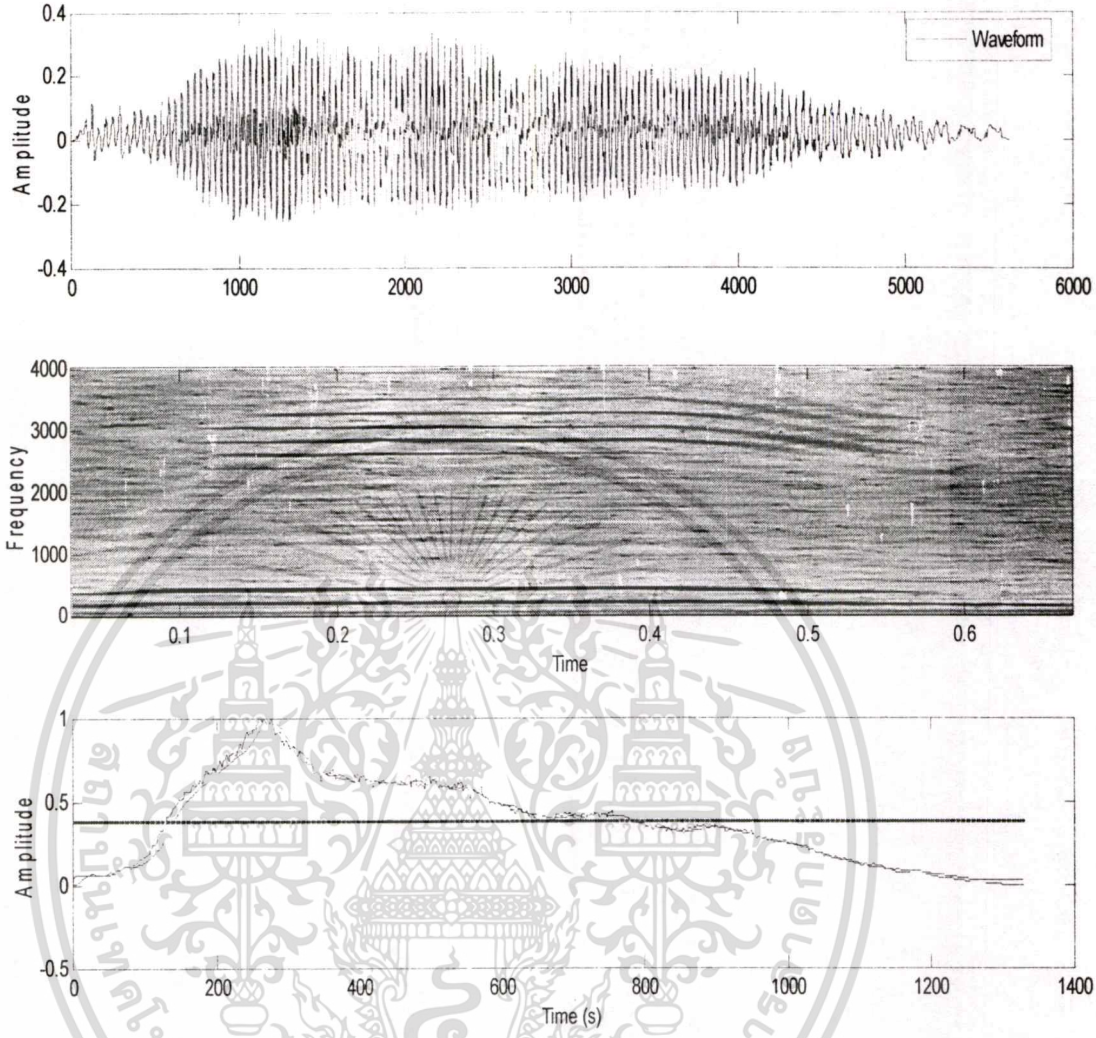
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.16 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงรัวด้วยตามเสียงสระ

เมื่อนำเสียงพยัญชนะเสียงรัวตามด้วยเสียงสระทั้งหมดเก็บช่วงกว่างของเสียงพยัญชนะและเสียงสระ ช่วงรอยต่อของเสียงจะกว้างเนื่องจากเสียงทั้งสองมีความถี่ใกล้เคียงกัน

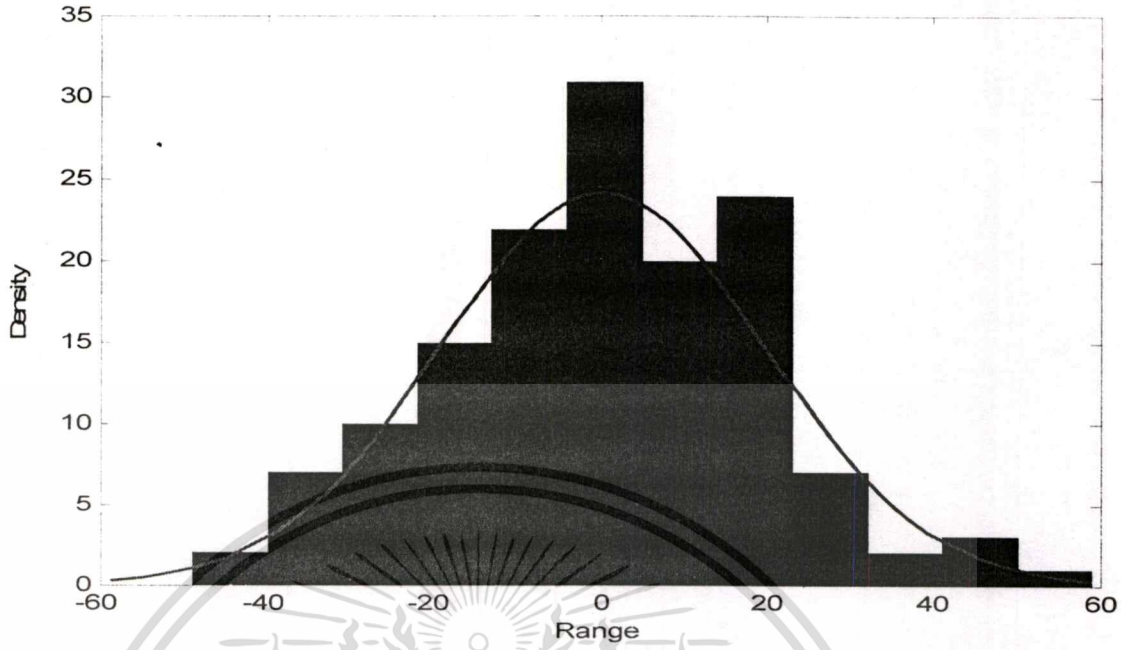
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.17 เสียงพยัญชนะเสียงครึ่งสระตามด้วยสระ คำว่า "วี"

เสียงพยัญชนะเสียงครึ่งสระตามด้วยสระมีจุดแยกที่ไม่ชัดเจนเนื่องจากการเกิดของเสียงเสียงครึ่งสระทำให้ความถี่ของเสียงพยัญชนะถี่ผสมกลมกลืนใกล้เคียงกับเสียงสระจึงทำให้รอยต่อของเสียงไม่เด่นชัดจึงทำให้แยกได้ยาก ในรูปที่ 4.17 รูปแรกแสดงสัญญาณเสียงของคำว่า "วี" เสียงพยัญชนะจะมีความถี่ต่ำและเกิดชั่วขณะส่วนของเสียงสระมีคาบเวลาที่คงที่ รูปล่างจะเป็นการหาจุดต่อของเสียงพยัญชนะและเสียงสระเส้นประจะเป็นส่วนที่ผ่านวงจรกรอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



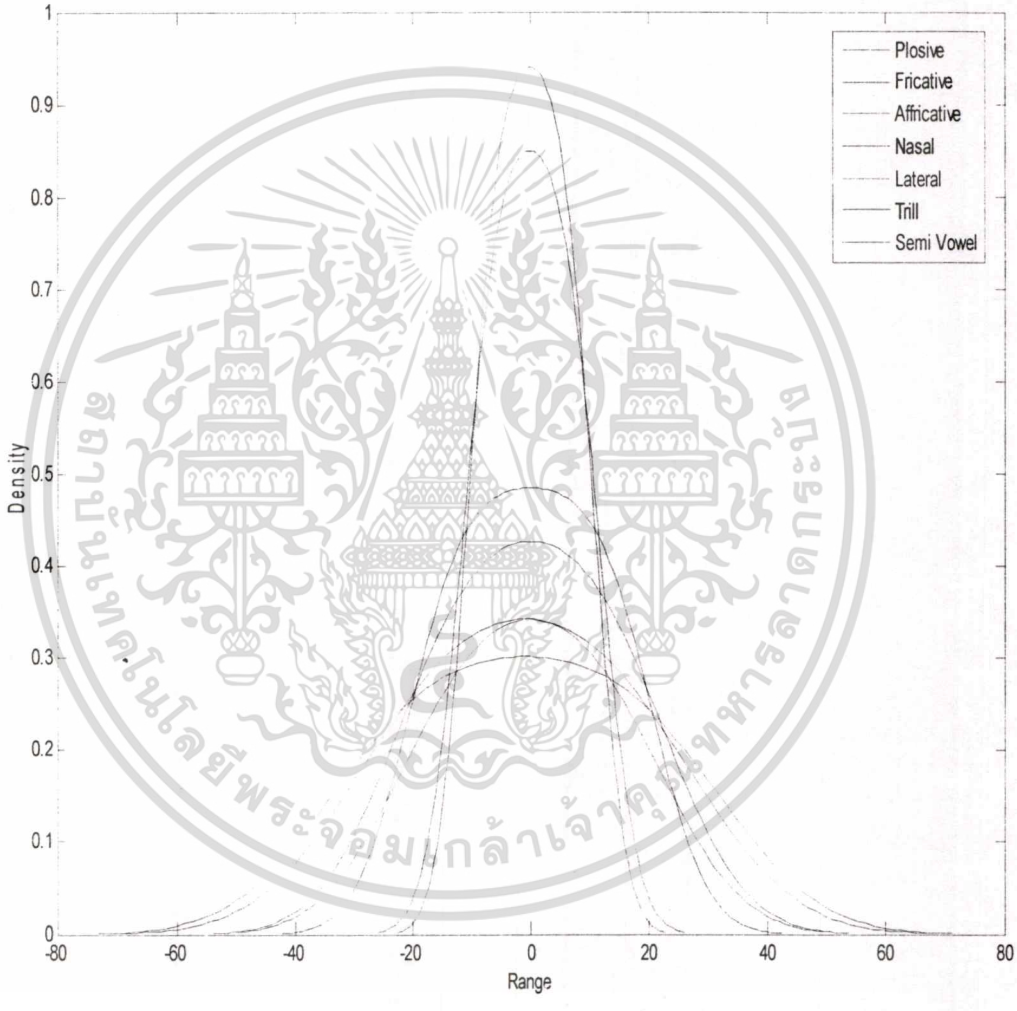
รูปที่ 4.18 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะเสียงครึ่งสระตามด้วยเสียงสระ

เมื่อนำเสียงพยัญชนะเสียงครึ่งสระตามด้วยเสียงสระทั้งหมดเก็บช่วงกว่างของเสียงพยัญชนะและเสียงสระ ช่วงรอยต่อของเสียงจะกว้างเนื่องจากเสียงทั้งสองมีความถี่ผสมกลมกลืนกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 สรุปผลการทดลอง

เมื่อเราได้ทำการทดสอบเสียงของข้อมูลที่คัดมาทั้งหมดแล้วใช้ฟังก์ชั่นในการคำนวณการแปลงเวฟเลทผลของความถูกต้องของการแยกเสียงพยัญชนะและเสียงสระได้ความถูกต้องของเสียงแต่ละประเภทของเสียงพยัญชนะมีขอบเขตของรอยต่อเสียงพยัญชนะและเสียงสระเสียงพยัญชนะเสียดแทรก และพยัญชนะกึ่งเสียดแทรกขอบเขตของรอยต่อเสียงที่แคบส่วนของเสียงที่มีขอบเขตของจุดรอยต่อที่กว่างที่สุดคือเสียงของพยัญชนะเสียงครึ่งสระเพราะมีความถี่ของเสียงใกล้เคียงกัน



รูปที่ 4.19 ขอบเขตเสียงจุดรอยต่อของเสียงพยัญชนะและเสียงสระ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลและข้อเสนอแนะ

วิทยานิพนธ์นี้เป็นการนำเสนอวิธีการแยกเสียงพยัญชนะและเสียงสระเพิ่มประสิทธิภาพในการรู้จำเสียงเดี่ยวในภาษาไทย โดยใช้การฟังก์ชันการแปลงเวฟเลทวิธีการแยกเสียงพยัญชนะและเสียงสระสำหรับเสียงพูดในภาษาไทย โครงสร้างของเสียงในภาษาไทย เกิดจากการรวมกันของพยัญชนะแล้วตามด้วยสระ การแยกพยัญชนะเป็นส่วนสำคัญในระบบการรู้จำเสียงในภาษาไทย การแปลงเวฟเลท เป็นวิธีนำเสนอการค้นหาตำแหน่ง การแยกพยัญชนะและสระโดยตรงโดยใช้สมาการผลคูณและรูปแบบพลังงานมาการผลคูณได้จากการประมาณค่าเวฟเลท approximate และ detail ของสัญญาณเสียงพูดที่เข้ามา และสามารถนำมาใช้บ่งชี้ตำแหน่งของพยัญชนะและสระสมาการผลคูณและการตรวจสอบรูปแบบของพลังงาน การแยกพยัญชนะและสระสามารถแยกตำแหน่งได้อย่างเหมาะสมด้วยการคำนวณที่ไม่ซับซ้อน

5.1 สรุปผลการทดลอง

เริ่มการทดลองด้วยการนำคำที่ประกอบด้วยเสียงสระที่ได้จากการบันทึกอัดเสียงมาวิเคราะห์สัญญาณเสียงเบื้องต้น เพื่อหาจุดแยกรอยต่อของพยัญชนะและเสียงสระการใช้เวฟเลทสามารถแยกเสียงพยัญชนะและเสียงสระออกจากกัน จากการทดลองเสียงพูดของผู้ชายและผู้หญิงทั้งหมดจำนวน 8 คน เพื่อศึกษาหารอยต่อของสัญญาณระหว่างเสียงพยัญชนะและเสียงสระทั้งหมด 4,032 เสียงปรากฏว่าจากเสียงพูดของผู้ชายและผู้หญิงได้เปอร์เซ็นต์ความถูกต้อง 96.1 เปอร์เซ็นต์ ส่วนเสียงข้าง (leteral) หรือเสียงรัว (Trill) มีความแม่นยำที่ได้คิดเป็นเปอร์เซ็นต์ ของการแยกเสียงได้ต่ำกว่าเนื่องจากมีสัญญาณเสียงพยัญชนะและเสียงสระยาวต่อเนื่องกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ข้อสังเกตและปัญหาที่พบการทดลอง

1) การอัดเสียงสำหรับมาทำวิจัยนั้น มีความผิดพลาดเยอะทั้งการออกเสียง สภาพแวดล้อม ความดังของเสียง เมื่อนำมาใช้ในการรู้จำแล้วทำให้ได้บางตัวอย่างเสียงได้คุณภาพต่างกัน
2) ลักษณะของคุณสมบัติเสียงข้าง (Leteral) หรือเสียงรัว (Trill) มีความถี่ปนไปกลับเสียงสระทำให้มีความแม่นยำลดลง
3) ในเสียงคำพูดเดียวกันเสียงของผู้ชายจะมีความถี่ต่ำกว่าเสียงของผู้หญิง จึงทำให้ผลของการทดลองเสียงของผู้หญิงจะมีความถูกต้องที่ต่ำกว่า

5.3 ข้อเสนอแนะ

1) ในส่วนของการอัดเสียงน่าจะคิดกระบวนการที่มีมาตรฐานเช่นขั้นตอนขจัดเสียงรบกวน ซึ่งจะช่วยให้มีการอัดเสียงที่มีคุณภาพ
2) ฐานข้อมูลเสียงภาษาไทยเป็นสิ่งที่จำเป็นในการพัฒนาการรู้จำเสียงในภาษาไทย ควรจะมีการรับผิดชอบในการจัดตั้งฐานข้อมูลเสียงภาษาไทยแห่งชาติขึ้น เพื่อใครก็ตามจะทำการวิจัยทางด้านเสียงจะได้นำเสียงมาจากแหล่งเดียวกัน สามารถตั้งบรรทัดฐานในค่าต่าง ๆ ของการรู้จำเสียงขึ้นได้ และมีประโยชน์เป็นแนวทางการพัฒนาต่อไป
3) นำงานวิจัยที่ได้ไปพัฒนาเพื่อให้ใช้งานในชีวิตประจำวันได้จริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] Jhing-Fa Wang and Shi-Huang Chen, Wu,Shin-Hung Chang,and Jau-Yien Lee "A Hierarchical Neural Network Model Based on a C/V Segment Algorithm for Isolated . Mandarin Speech Recognition" IEEE Transaction on signal processing Vol.39 N. 9 . . . September 1991 ,pp. 2141-2146

[2] Stephen W.K Fu,C.H.Lee Orville Leverne Clubb"A robust C/V Segmentation . a ' ' Algorithm for Cantonese" IEEE TENCON-Digital Signal Processing Application1996,pp 42-45

[3] S.-H.Chen and J.-F.Wang, "Application of wavelet transforms for C/V segmentation on speech signals", IEE Proc.-Image Signal Process., Vol. 148, No. 2, April 2001, pp. 133-139

[4] Jhing-Fa Wang and Shi-Huang Chen, " A C/V segmentation algorithm for mandarin speech signal based on wavelet transforms", 1999 IEEE, pp. 417-420

[5] จิตรลดา จารุมิศร์และไกรสิน ส่งวัฒนา , "การออกแบบ แบบจำลองในการรู้จำเสียงวรรณยุกต์ สำหรับภาษาไทย โดยใช้เทคนิคการควอนไตซ์พิทช์ และ Hidden Markov Modeling", วิทยานิพนธ์ มหาบัณฑิต, สถาบันเทคโนโลยี พระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2542

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก.
ผลงานวิจัยที่ได้รับการตีพิมพ์

1. นริศ บุญศักดิ์เฉลิม, ไกรสิน ส่งวัฒนา "การประยุกต์ใช้เวฟเลทแยกเสียงพยัญชนะ เสียงสระภาษาไทย", วิศวกรรมสารลาดกระบัง, ปีที่ 23, ฉบับที่ 2 เดือนมิถุนายน 2549.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การประยุกต์ใช้เวฟเลทแยกเสียงพยัญชนะ เสียงสระ ภาษาไทย

Application of Wavelet Transforms for Consonant/Vowel Segmentation on Thai Speech Signal

นริศ บุญศักดิ์เฉลิม ไกรสิน ส่งวัฒนา

ภาควิชาวิศวกรรมโทรคมนาคม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

บทคัดย่อ

บทความนี้นำเสนอวิธีการแยกเสียงพยัญชนะและเสียงสระสำหรับเสียงพูดในภาษาไทย โครงสร้างของเสียงในภาษาไทย เกิดจากการรวมกันของพยัญชนะแล้วตามด้วยสระ การแยกพยัญชนะเป็นส่วนสำคัญในระบบการรู้จำเสียงในภาษาไทย การแปลงเวฟเลท เป็นวิธีที่นำเสนอการค้นหาตำแหน่ง การแยกพยัญชนะและสระโดยตรงโดยใช้สมการ ผลคูณและรูปแบบของพลังงาน สมการผลคูณได้จากการประมาณค่าเวฟเลท และ รายละเอียดของสัญญาณเสียงพูดที่เข้ามา และสามารถนำมาใช้บ่งชี้ตำแหน่งของพยัญชนะและสระ สมการผลคูณและการตรวจสอบรูปแบบของพลังงาน การแยกพยัญชนะและสระสามารถแยกตำแหน่งได้อย่างเหมาะสมด้วยการคำนวนที่ไม่ซับซ้อน การทดลองใช้ระยะเวลาน้อย ไม่สิ้นเปลืองในการคำนวณ จากการทดลองได้ค่าเปอร์เซ็นต์ ความถูกต้อง 96.1 เปอร์เซ็นต์

Abstract

This paper proposes Consonant/Vowel (C/V) segmentation algorithm for Thai speech signal. Since the Thai phoneme structure is a combination of a consonant followed by a vowel ,the C/V segmentation is an important part in the Thai speech recognition system. Based on the wavelet transforms, the proposed method can directly search for the C/V segmentation point by using a product function and energy profile. The product function is generated from the appropriate approximation wavelet and detail of input speech signal, and it can be applied to indicate the C/V segmentation point. With this product function and the additional verification of energy profile, the C/V segmentation can be accurately pointed out with a low computation complexity. Experiment result demonstrates the performance of the proposed algorithm to be better. They have shown the average accuracy 96.1 percent.

1.บทนำ

การรู้จำเสียงพูดมีขั้นตอนการแยกเสียงพยัญชนะและแยกเสียงสระออกจากกัน ซึ่งเป็นขั้นตอนที่มีความสำคัญต่ออัตราความถูกต้องและต่อการรู้จำเพราะจะได้ส่วนของเสียงพยัญชนะและเสียงสระ เพื่อนำไปสู่กระบวนการรู้จำเสียงพยัญชนะและ เสียงสระต่อไป

วิธีการแยกเสียงพยัญชนะและเสียงสระมีหลายวิธีเช่น Neural Network Model [1] วิธีนี้ใช้การเขียนโปรแกรมสร้างแบบจำลองให้ครอบคลุมกับเงื่อนไขของเสียง ซึ่งได้เปอร์เซ็นต์ความถูกต้องของค่าเฉลี่ยอยู่ที่ 90 เปอร์เซ็นต์ วิธีการนี้จะใช้ระยะเวลาในการคำนวณมากหรือน้อยขึ้นอยู่กับรายละเอียดของแบบจำลองที่เราสร้างขึ้นมาและความแม่นยำนั้นขึ้น อยู่กับการสร้างแบบจำลองให้ครอบคลุมกับรายระเอียดของเสียงและข้อเสียหนึ่งคือต้องใช้เวลาในการจำแนกเสียงมากยิ่ง ขึ้นและอีกวิธีนั้นคือวิธี Zero-Crossing rate [2] เป็นการใช้ จุดตัดของกำลังงานเสียงของพยัญชนะและเสียงสระได้เปอร์ เซ็นต์ความถูกต้องอยู่ที่ 95 เปอร์เซ็นต์ ซึ่งสามารถนำมาพัฒนาต่อให้ได้ความเร็วและความแม่นยำเพิ่มมากขึ้นโดยใช้การแปลงเวฟเลท[3],[4]

1.1 เสียงพยัญชนะภาษาไทย

เสียงพยัญชนะในภาษาไทยมีทั้งหมด 21 เสียง จากรูปพยัญชนะทั้งหมด 44 รูป ดังตารางที่ 1 ซึ่งได้แสดงลักษณะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของเสียงพยัญชนะคือเสียงระเบิด เสียงแทรก เสียงกึ่งเสียดแทรก เสียงนาสิก เสียงข้าง เสียงรัวและเสียงครึ่งสระ ซึ่งแต่ละเสียงนั้นจะมีความแตกต่างขณะที่เปล่งเสียงออกมาเพราะขณะเปล่งเสียงลมที่พุ่งออกมา จากหลอดลมผ่านเส้นเสียงนั้นจะถูกกีดขวางตามส่วน ต่างๆภายในช่องปากอย่างเช่นลิ้นไก่ ลิ้น ฟัน เพดานปากริมฝีปากหรือลักษณะการเปิดหรือปิดช่องว่างภายในปากที่จะบีบเสียงออกมาตามที่ต้องการไม่ว่าเสียงสูงหรือต่ำเพราะฉะนั้นจึงเกิดเป็นเสียงสั้นๆและมีความถี่ไม่คงที่

ลักษณะของเสียงพยัญชนะ	ประกอบด้วยพยัญชนะ								
เสียงระเบิด	ป	พ	บ	ด	ท	ค	ก	ล	อ
เสียงเสียดแทรก	ช	ฟ	ฮ						
เสียงกึ่งเสียดแทรก	จ	ช							
เสียงนาสิก	ม	น	ง						
เสียงข้าง	ล								
เสียงรัว	ร								
เสียงครึ่งสระ	ว	ย							

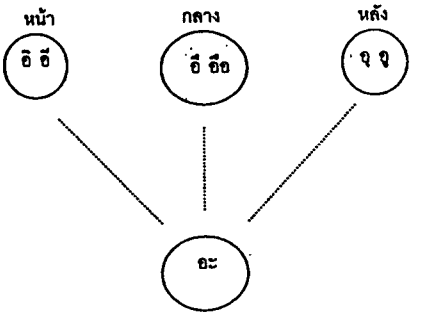
ตารางที่ 1 ลักษณะของเสียงพยัญชนะจำแนกตามลักษณะรูปเสียง

1.2 เสียงสระภาษาไทย

สระในภาษาไทยมีสระทั้งหมด 2 ชนิดซึ่งประกอบด้วยสระแท้และสระผสมสระแท้มีเสียงทั้งหมดจำนวน 24 เสียง โดยมีเสียงสระเสียงสั้นจำนวน 9 เสียง คือ อิ, เอะ, แอะ, อึ, เออะ, อะ, อุ, โอะ และเอาะ และเสียงสระเสียงยาว 9 เสียง คือ อี, เอ, แอ, อือ, เออ, อา, อู, โอ และออ และเสียงสระผสมอีก 6 เสียง คือ เอียะ, เอีย, เอือะ, เอือ, อัวะ และ อัว สระเสียงยาวนั้นเกิดจากการเปล่งเสียงสระเสียงสั้นในช่วงเวลาที่ยาวขึ้นและมีการหยุดของเสียงที่ช้ากว่าเสียงสระเสียงสั้นดังแสดงในตารางที่ 2 ส่วนสระผสมนั้นเกิดจากการรวมกันของหน่วยเสียงสระเสียงเดี่ยว 2 ชนิด [5]ดังรูปที่ 1

สระ	หน้า	กลาง	หลัง
สูง	อิ , อี	เออะ , เออ	อุ , อู
กลาง	เอะ , เอ	เออะ , เออ	โอะ , โอ
ต่ำ	แอะ , แอ	อะ , อา	เอาะ , ออ

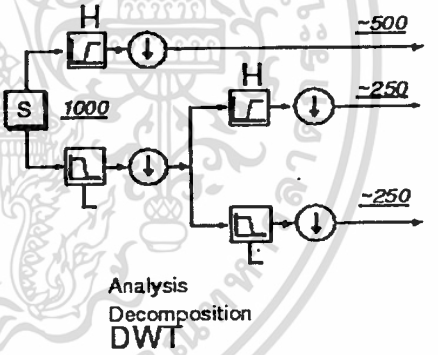
ตารางที่ 2. แสดงเสียงสระเดี่ยว



รูปที่ 1 การรวมเสียงสระเดี่ยวเป็นสระประสม

2. การแปลงเวฟเลทแบบไม่ต่อเนื่อง (Discrete Wavelet Transforms : DWT)

การแปลงเวฟเลทโดยใช้หลักการสังเคราะห์ความถี่ (Analysis Filter Banks) เป็นขบวนการกรองข้อมูลจากระดับความละเอียดสูงสู่ระดับความละเอียดต่ำกว่าโดยอาศัยหลักการ two channel analysis filter banks ที่มีการเปลี่ยนแปลงความ ถี่ในอัตราครั้งละสองเท่าคือจำนวนข้อมูลสุ่มจะลดลง (Down Sampling) สองเท่าในแต่ละสเตทนั้นดังรูปที่ 2 โดยที่ L เป็น Lowpass Filter และ H เป็น Highpass Filter



รูปที่ 2 ลักษณะของ DWT กับกระบวนการ Two channel analysis filter bank

$$\psi_{m,n}(t) = 2^{m/2}\psi(2^m t - n) \quad (1)$$

$\psi_{m,n}(t)$ เป็นเวฟเลทแม่โดยกำหนดให้ m และ n เป็นค่าคงที่ $\psi(t)$ สามารถแยกกระจายออกเป็น $\varphi(t)$ และ $\psi(t)$

$$\varphi(t) = \sqrt{2}\sum_k h(k)\varphi(2t-k) \quad (2)$$

$$\psi(t) = \sqrt{2}\sum_k g(k)\varphi(2t-k) \quad (3)$$

โดยที่ค่า h(k) และ g(k) ในสมการ (2) และ (3) เป็นตัวกรองความถี่ต่ำและความถี่สูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$g(k) = (-1)^K h(N-k-1) \quad (4)$$

ดังนั้นสัมประสิทธิ์เวฟเลทสามารถหาได้จากสมการการโปรเจคชั่น j เป็นลำดับขั้นการสเกล ดังสมการที่ (5), (6) และ (7)

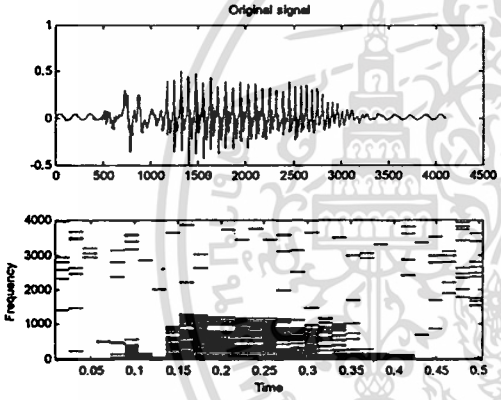
$$f(x) = \sum_n S_{j+1}(n)\varphi_{j+1,n}(t) + \sum_n \sum_{j=0}^{J} W(n)\psi(t) \quad (5)$$

$$s_j(n) = \sum_m h(m-2n)S_{j-1}(m) \quad (6)$$

$$w_j(n) = \sum_m g(m-2n)S_{j-1}(m) \quad (7)$$

3.การแยกเสียงพยัญชนะและเสียงสระ

เสียงพยัญชนะส่วนใหญ่มีความถี่ต่ำและความถี่ของพยัญชนะอยู่ในช่วงสั้นๆและมีความถี่ของคลื่นเสียงไม่คงที่ แต่เสียงสระมีคาบเวลาพีทช์ที่ต่อเนื่องและมีความถี่สูงกว่าความถี่ของพยัญนะดังรูปที่ 3 ได้แสดงตัวอย่างเสปคตรัมของสัญญาณ เสียง ดังกล่าว



รูปที่ 3 สัญญาณเสียงพยัญชนะและสระในโดเมนเวลาและความถี่

เนื่องจากคุณสมบัติของเสียงพยัญชนะกับเสียงสระมีความถี่ที่แตกต่างกัน เราใช้คุณสมบัติของเวฟเลทสมการที่ (6) เก็บค่าสัมประสิทธิ์ของเสียงพยัญชนะและสมการที่ (7) ใช้เก็บค่าสัมประสิทธิ์ของเสียงสระ

$$P_j(n) = \frac{S_j(n) \cdot w_j(n)}{2^{-2j}} \quad (8)$$

สมการที่ (8) เกิดจากการนำค่าสัมประสิทธิ์ของสมการที่ (6) Product function [3] และค่าสัมประสิทธิ์ของสมการที่ (7) หาค่าฟังก์ชั่น $P_j(n)$ เพื่อกำจัดส่วนต่างระหว่างเสียงพยัญชนะและเสียงสระ

$$E(n) = \sum_{t=n-M+1}^{n} |P(n)| \quad (9)$$

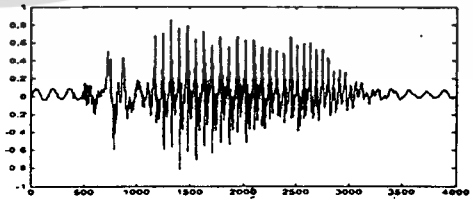
จากสมการที่ (9) หาค่า E(n) Energy profile [3] กำหนดให้ M เป็น 40 แล้วเพิ่มค่าจนหมดข้อมูล n

$$Th = \sum_{n=Ps}^{Pe} \frac{|P(n)|}{(Pe - Ps)} \quad (10)$$

จากสมการที่ (10) เมื่อค่า Ps คือพลังงานเริ่มต้นของเสียงและ Pe คือค่าพลังงานสิ้นสุดของเสียง ผลการคำนวณจะได้ระดับ Threshold มาใช้ในการชี้ตำแหน่งรอยต่อของเสียงพยัญชนะและเสียงสระ

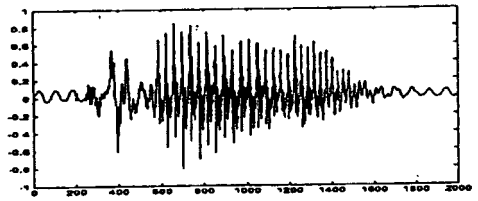
4.การทดลองและผลการทดลอง

การทดลองนี้ได้เก็บตัวอย่างเสียงพูดจากผู้ชาย 4 คน และผู้หญิง 4 คน รวมทั้งหมดจำนวน 8 คน ซึ่งอยู่ในสภาพแวดล้อมปกติที่ไม่มีเสียงรบกวนโดยให้แต่ละคนออกเสียงเป็นคำๆ โดยแต่ละคำนั้นมีเสียงพยัญชนะและเสียงสระรวมอยู่ด้วยโดยมีเสียงสระเดี่ยวทั้งหมดจำนวน 18 เสียง แบ่งเป็นเสียงสระเสียงสั้นจำนวน 9 เสียง และเสียงสระเสียงยาวจำนวน 9 เสียง เสียงสระ ผสมอีกจำนวน 6 เสียง ดังนั้นรวมเสียงสระทั้งหมดเป็น 24 เสียง และรวมเสียงด้วยพยัญชนะทั้งหมดมี 21 เสียง ฉะนั้นจำนวนคำทั้งหมดที่ใช้ในการทดลองรวมเป็นจำนวน 24 x 21 x 8 = 4,032 เสียงจากสัญญาณเสียงพูด ความถี่มีค่าประมาณ 0 ถึง4 กิโลเฮิร์ต ความถี่ในการสุ่มตัวอย่าง 8 กิโลเฮิร์ตส่วนการแยกสัญญาณนั้น ได้นำสัญญาณเสียงผ่านขบวนการแปลงเวฟเลทเป็นขั้นตอนขั้นแรกจะได้สัญญาณเสียงออกเป็นสองส่วน ส่วนแรกคือส่วนของความถี่ต่ำจะได้ค่าสัมประสิทธิ์การประมาณค่าสัญญาณ 0 ถึง2 กิโลเฮิร์ต และส่วนของความถี่สูงจะได้ค่า สัมประสิทธิ์สัญญาณ 2 ถึง4 กิโลเฮิร์ต ดังรูปที่ 4, 5 และ6 จะ เป็นสัญญาณเสียงพูด "ขา"

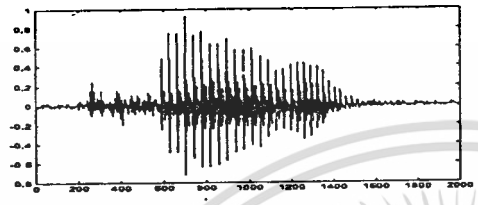


รูปที่ 4 แสดงสัญญาณเสียง "ขา"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

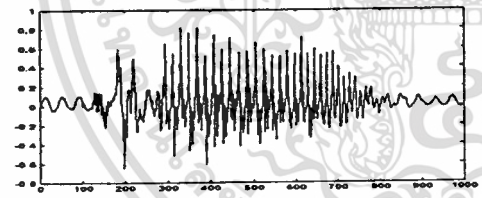


รูปที่ 5 สัมประสิทธิ์การประมาณค่าสัญญาณเวฟเลท 0 ถึง 2 กิโลเฮิร์ต

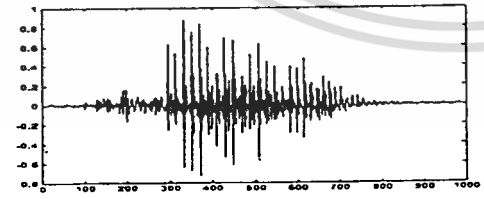


รูปที่ 6 ค่าสัมประสิทธิ์การประมาณค่าสัญญาณเวฟเลท 2 ถึง 4 กิโลเฮิร์ต

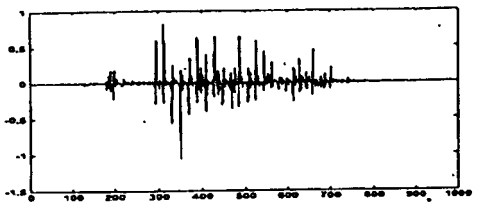
จากนั้นได้นำส่วนของความถี่ต่ำที่มีค่าสัมประสิทธิ์การประมาณค่าสัญญาณ appropriate 0 ถึง 2 กิโลเฮิร์ต มาแปลงด้วยเวฟเลทอีกครั้งจะได้ค่าสัมประสิทธิ์การประมาณค่าสัญญาณเวฟเลท 0 ถึง 1 กิโลเฮิร์ต และส่วนของความถี่สูงจะได้ ค่าสัมประสิทธิ์สัญญาณเวฟเลท 1 ถึง 2 กิโลเฮิร์ต ดังรูปที่ 7, 8 เป็นการแปรสัญญาณเสียงพูด"ขา"



รูปที่ 7 ค่าสัมประสิทธิ์การประมาณค่าสัญญาณเวฟเลท 0 ถึง 1 กิโลเฮิร์ต

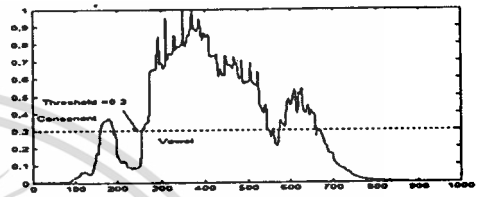


รูปที่ 8 ค่าสัมประสิทธิ์การประมาณค่าสัญญาณเวฟเลท 1 ถึง 2 กิโลเฮิร์ต



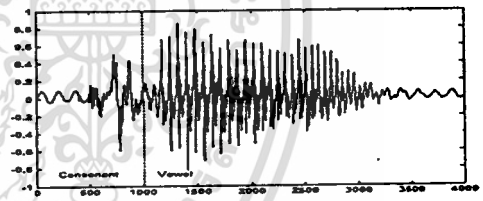
รูปที่ 9 สัญญาณเสียงของฟังก์ชั่น P(n)

จากรูปที่ 9 เป็นการแสดงผลตัวอย่างของสมการที่ (8) เพื่อกำจัดส่วนต่างระหว่างเสียงพยัญชนะและเสียงสระ



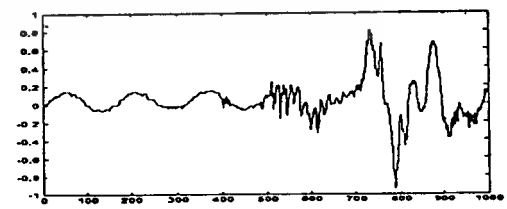
รูปที่ 10 สัญญาณเสียงของฟังก์ชั่น E(n)

จากรูปที่ 10 แสดงผลการคำนวนจากสมการที่ (9) โดยแสดงระดับกำลังงานของสัญญาณเสียงและพยัญชนะซึ่งระดับ Threshold ที่ได้มีค่า 0.3 จึงสามารถหาตำแหน่งของรอยต่อเสียงได้ที่ตำแหน่ง 250



รูปที่ 11 แสดงสัญญาณจุดรอยต่อเสียงพยัญชนะกับ สระ

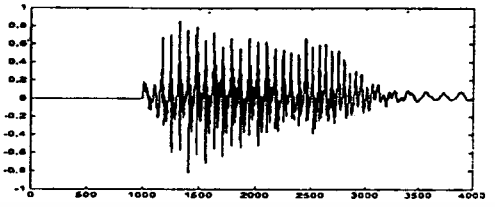
ในรูปที่ 11 ตำแหน่งของรอยต่อเสียงคือตำแหน่งที่1,000 โดยคำนวณย้อนกลับจากการใช้เวฟเลทมาสองสเตทเป็นการย่อข้อมูลลงมาสี่เท่า



รูปที่ 12 สัญญาณเสียงพยัญชนะ "ข"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากค่าตำแหน่งรอยต่อของสัญญาณเสียงพยัญชนะและเสียงสระ คือตำแหน่งที่1,000 สามารถแยกพยัญชนะ "ข" ที่แสดงไว้ในรูปที่ 12 และสามารถแยก "า" ไว้ในรูปที่ 13



รูปที่ 13 สัญญาณเสียงสระ "า"

เสียงที่ทดสอบ	ความถูกต้อง(%)
พยัญชนะเสียงระเบิด (Plosive) , สระ(Vowel)	99
พยัญชนะเสียงเสียดแทรก . (Fricativ e) ,สระ(Vowel)	97
พยัญชนะเสียงกึ่งเสียดแทรก (Affricative) ,สระ(Vowel)	96
พยัญชนะเสียงนาสิก (Nasal) ,สระ(Vowel)	96
พยัญชนะเสียงข้าง (Lateral) ,สระ(Vowel)	95
พยัญชนะเสียงรัว (Trill) ,สระ(Vowel)	94
พยัญชนะเสียงกึ่งสระ (Semi Vowel) ,สระ(Vowel)	96
สรุป	96.1

ตารางที่ 3

แสดงความถูกต้องในการแยกเสียงพยัญชนะออกจากสระ

จากตารางที่ 3 คือผลการทดลองจากการแบ่งกลุ่มของพยัญชนะและจำแนกตามลักษณะของรูปเสียง สามารถแบ่งออกได้ทั้งหมด 7 ประเภท แล้วจึงนำมาผสมกับสระเป็นจำนวนทั้งหมด 4,032 เสียง เพื่อจะทำให้ทราบว่าเสียงของพยัญชนะแต่ละกลุ่มเสียงมีความแม่นยำในการแยกเสียงพยัญชนะ เสียงสระแตกต่างกันอย่างไร

5. สรุปผลการทดลอง

การใช้เวฟเลทสามารถแยกเสียงพยัญชนะและเสียงสระออกจากกันจากการทดลองเสียงพูดของผู้ชายและผู้หญิง ทั้งหมดจำนวน 8 คน เพื่อศึกษาหารอยต่อของสัญญาณระหว่างเสียงพยัญชนะและเสียงสระทั้งหมด 4,032 เสียงปรากฏว่าจากเสียงพูดของผู้ชายและผู้หญิงได้เปอร์เซ็นต์ความถูกต้อง 96.1 เปอร์เซ็นต์ ส่วนเสียงข้าง (Lateral) หรือเสียงรัว (Trill) มีความแม่นยำที่ได้คิดเป็นเปอร์เซ็นต์ ของการแยกเสียงได้ต่ำกว่าเนื่องจากมีสัญญาณเสียงพยัญชนะและเสียงสระยาวต่อเนื่องกัน ผลงานวิจัยนี้สามารถแยกสัญญาณเสียงของพยัญชนะและเสียงของสระหรือนำไปใช้ในการรวบรวมเสียงเป็นพยางค์ต่างๆ ได้และสามารถที่จะนำไปวิเคราะห์เสียงพยัญชนะเสียง ระในการ รู้จำเสียงต่อไปได้

6. เอกสารอ้างอิง

[1] J:F Wang and S.-H Chen, W.-H Chang, and J.-Y Lee "A Hierarchical Neural Network Model Based on a C/V Segment Algorithm for Isolated Mandarin Speech Recognition," IEEE Trans. Signal Processing., Vol.39, No. 9, pp 2141-2146, September 1991.

[2] S.W.K. Fu, "A robust C/V Segmentation Algorithm for Cantonese,"Proc. TENCON'96., Vol. 1, pp. 42-45, November 1996.

[3] S.-H.Chen and J.-F.Wang, "Application of Wavelet Transforms for C/V Segmentation on Speech Signal," Proc.-Image Signal Process., Vol. 148, No. 2, pp. 133-139, April 2001.

[4] J:F Wang and S.-H Chen, "A C/V Segmentation Algorithm for Mandarin Speech Signal Based on Wavelet Transforms," Proc. ICASSP '99., Vol. 1, pp.417 – 420, 15-19 March 1999.

[5] จิตรลดา จารุมิศร์และไกรสิน ส่งวัฒนา "การออกแบบแบบจำลองในการรู้จำเสียง วรรณยุกต์ สำหรับภาษาไทย โดยใช้เทคนิคการควอนไตซ์พิทช์ และ Hidden Markov Modeling," วิทยานิพนธ์ วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2542

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ผู้เขียน	นายนริศ บุญศักดิ์เฉลิม
วันเดือนปีเกิด	15 ตุลาคม 2517
สถานที่เกิด	จังหวัด ฉะเชิงเทรา
ปริญญา	สาขาวิศวกรรมโทรคมนาคม คณะครุศาสตร์อุตสาหกรรม สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้