สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การสั่งงาน SWITCH ON – OFF ด้วยเสียงพูด

T042372

โดย

นาย วชิรพงศ์ ศรีรี

นาย สมเจตน์ จูประวัติ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เลขหมู่.....
เลขทะเบียน 42372
วัน, เดือน, ปี 20 พ.ค. 2545

.b.....
.i.....

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาอุตสาหกรรมศาสตรบัณฑิต

ภาควิชาเทคนิคอุตสาหกรรม

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2544

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

D11207875

หัวข้อปริญญานิพนธ์

โดย นาย วชิรพงศ์ ศรีรี

นาย สมเจตน์ จูประวัติ

อาจารย์ที่ปรึกษา รศ. นิกร สุขุตมตันติ

ภาควิชา เทคนิคอุตสาหกรรม

ปีการศึกษา 2544

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง อนุมัติให้ปริญญานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาอุตสาหกรรมศาสตรบัณฑิต

คณะกรรมการสอบปริญญานิพนธ์

_____ ประธานกรรมการ

( )

_____ กรรมการ

( )

_____ กรรมการ

( )

_____ กรรมการ

( )

_____ กรรมการ

( )

_____ กรรมการ

( )

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การสั่งงาน SWITCH ON-OFF ด้วยเสียงพูด

โดย นาย วชิรพงศ์ ศรีรี
นาย สมเจตน์ จูประวัติ

อาจารย์ที่ปรึกษา รศ.นิกร สุขุตมตันติ

ปีการศึกษา 2544

บทคัดย่อ

ในปัจจุบันวิทยาการทางด้านคอมพิวเตอร์ได้เจริญก้าวหน้ามากและได้ก้าวไปอย่างรวดเร็วซึ่งปริญญานิพนธ์นี้นำคอมพิวเตอร์นำมาประมวลผลสัญญาณเสียงที่ผ่านมาจากไมโครที่ต่ออยู่กับการ์ดเสียง ด้วยความถี่แซมปลิ้ง16000 Hz ความละเอียด 8 บิทต่อแซมปลิ้ง หลักการเบื้องต้น และหลักการในการนำควบคุมสวิทช์ด้วยเสียงพูดผ่านทางคอมพิวเตอร์ซึ่ง แบบขึ้นกับผู้พูดในกลุ่มคำ 7 คำ ดังนี้ "หนึ่ง", "สอง ", "สาม", " สี่ ", "ห้า"" ON", "OFF" โปรแกรมที่ใช้คือ MATLAB

ส่วนรู้จำเสียงพูด (Speech Recognition) ใช้วิธีการหาอัตราการตัดศูนย์ของเสียงเป็นพารามิเตอร์แทนลักษณะของเสียงพูด โดยวิธีการหาอัตราการตัดศูนย์จะทำการแบ่งลักษณะประเภทออกเป็น แบบไม่มีเสียงเสียดสีฟัน (VOICE) และแบบเสียงเสียดสีฟัน (UNVOICED) เพื่อมาทำการประมวลผลหาผลลัพธ์ที่ถูกต้อง จากผลการทดสอบสามารถตัดสินใจความถูกต้องคิดเป็นร้อยละ 67 %

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

SPEECH RECOGNITION CONTROL SWITCH

By VACHIRAPONG SRIREE
SOMJATE JUPAWAT
Advisor ASSOC.PROF. NIKORN SUKUTAMATANTI
Year 2001

Abstract

This research present the speech control switch. This speech recognition system is write in MATLAB program. Speech is stored in wave format file by using microphone connected with soundcard having sampling frequency of 16000 Hz and 8 bit per sampling

The speech recognition using zero crossing rate as parameter of speech. This parameter is used as data for to determine word .The result show that find zero crossing rate has and accuracy of 67%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

คณะผู้จัดทำขอขอบพระคุณ รศ. นิกร สุขุตมตันติ ที่ให้คำแนะนำในการทำปริญญานิพนธ์ และได้ให้คำปรึกษาเพื่อเป็นแนวทางในการทำปริญญานิพนธ์ให้สำเร็จลุล่วงไปได้ด้วยดี และบุคคลที่มีส่วนเกี่ยวข้องกับปริญญานิพนธ์ที่มิได้กล่าวนามในที่นี้ด้วย ที่ลืมมิได้ก็คือ จ.อ. ธม ทะการ ที่ให้คำปรึกษาและแนะนำแนวทางในการทำปริญญานิพนธ์

ขอขอบพระคุณอย่างสูง

คณะผู้จัดทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

เรื่อง	หน้า
บทคัดย่อ	I
Abstract	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญรูป	VI
บทที่ 1 บทนำ	1
1.1 ความสำคัญของปัญหา	1
1.2 วัตถุประสงค์	2
1.3 ขอบเขตของงานวิจัย	2
1.4 ขั้นตอนของการวิจัย	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ	3
บทที่ 2 หลักการและทฤษฎี	
2.1 ทฤษฎีการสร้างเสียงพูด	4
2.2 การเกิดของเสียง (SPEECH PRODUCTION)	5
2.3 ทฤษฎีการวิเคราะห์เสียงพูด	8
2.4 การวิเคราะห์สัญญาณเสียงในช่วงสั้นๆ	8
2.5 การวิเคราะห์ในโดเมนเวลา	9
2.6 การประเมินค่าช่วงเวลาพิทช์โดยวิธีการเข้าใกล้เส้นขนาน	13
2.7 การวิเคราะห์ตามแกนความถี่	16
บทที่ 3 โครงสร้างงานวิจัย	17
3.1 ภาคอ่านเสียงไฟล์	17
3.2 ภาครู้จำเสียงพูด	18
3.3 ภาคประมวลผลลัพธ์	19

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4 ภาคแสดงผลโดยการนำไปควบคุมสวิทช์ 19

บทที่ 4 วิธีการดำเนินการงานวิจัย 20
4.1 ขั้นตอนการวิจัย 20
4.2 การเก็บข้อมูลกลุ่มตัวอย่างและกลุ่มตัวอย่างทดสอบเสียงพูด 20
4.3 การเตรียมข้อมูลกลุ่มตัวอย่างและกลุ่มทดสอบ 20
เพื่อใช้ในการตัดสินใจ
4.4 การวิเคราะห์สัญญาณเสียงในช่วงเวลาสั้น ๆ 21
4.5 รูปแบบช่องแคบ (WINDOW) 22

บทที่ 5 ผลการทดลอง 23
5.1 รูปแสดงลักษณะของเสียงและการหาจุดเริ่มต้น 23
และจุดสุดท้ายของสัญญาณเสียง
5.2 การหาสเปกตรัมของสัญญาณเสียงต่างๆ 29
5.3 รูปแสดงผลการทดลองที่แสดงผลทางจอ Computer 32
ปัญหาในการทำงานและข้อเสนอแนะ 38
บรรณานุกรม 39
ภาคผนวก 40

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ

หน้า

รูปที่2.1 แสดงอวัยวะในการออกเสียง 4
รูปที่2.2 บล็อกไดอะแกรมของระบบการสร้างเสียงพูด 6
รูปที่2.3 แบบจำลองอย่างง่ายของการกำทอนช่องทางเดินเสียง 7
รูปที่2.4 บล็อกไดอะแกรมของอัตตราการตัดศูนย์เฉลี่ยในวลาช่วงสั้น 10
รูปที่2.5 อัตตราการเปลี่ยนแปลงของเสียงที่แตกต่างกัน3ครั้ง 12
รูปที่2.6 บล็อกไดอะแกรมของการถอดสัญญาณพิทช์แบบขนาน 14
รูปที่2.7 อินพุทสัญญาณซายน์และการเกิดขบวนอิมพัลส์จากPeak Valley 15
รูปที่3.1 ระบบการบันทึกเสียง 17
รูปที่5.1 รูปแสดงสัญญาณเสียง " หนึ่ง " 23
รูปที่5.2 รูปแสดงสัญญาณเสียง " สอง " 23
รูปที่5.3 รูปแสดงสัญญาณเสียง " สาม " 24
รูปที่5.4 รูปแสดงสัญญาณเสียง " สี่ " 24
รูปที่5.5 รูปแสดงสัญญาณเสียง " ห้า " 25
รูปที่5.6 รูปแสดงสัญญาณเสียง " ON " 25
รูปที่5.7 รูปแสดงสัญญาณเสียง " OFF " 26
รูปที่5.8 สเปกตรัมของเสียงหนึ่ง 29
รูปที่5.9 สเปกตรัมของเสียงสอง 29
รูปที่5.10 สเปกตรัมของเสียงสาม 30
รูปที่5.11 สเปกตรัมของเสียงสี่ 30
รูปที่5.12สเปกตรัมของเสียงห้า 31
รูปที่5.13 สเปกตรัมของเสียง ON 31
รูปที่5.14 สเปกตรัมของเสียง OFF 32
รูปที่5.15 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.1 ON 32
รูปที่5.16 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.1 OFF 33

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ

หน้า

รูปที่5.17 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.2 ON 33
รูปที่5.18 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.2 OFF 34
รูปที่5.19 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.3 ON 34
รูปที่5.20 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.3 OFF 35
รูปที่5.21 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.4 ON 35
รูปที่5.22 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.4 OFF 36
รูปที่5.23 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.5 ON 36
รูปที่5.24 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.5 OFF 37

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความสำคัญของปัญหา

ความพยายามของมนุษย์ในการสั่งงานเครื่องมือเครื่องจักรต่างๆด้วยเสียงนั้นมีมานานแล้ว ทั้งนี้เพื่อความสะดวกและรวดเร็วในการติดต่อให้มากขึ้น ในอนาคตการสั่งงานด้วยเสียงคำพูดจะใช้กันอย่างแพร่หลาย คอมพิวเตอร์สามารถเข้าใจมนุษย์และสามารถโต้ตอบกับผู้ใช้ได้ความต้องการ ใช้คอมพิวเตอร์สามารถเรียนรู้ถึง วิธีการในการรู้จักคำพูดได้ ทำให้เกิดศาสตร์แขนงหนึ่งขึ้นเรียกว่าการรู้จำเสียงคำพูด (Speech Recognition)

เราสามารถแบ่งการศึกษาตามลักษณะผู้พูดได้เป็น 3 แบบ คือ

1. แบบขึ้นกับผู้พูด (Speaker Dependence)
2. แบบหลายผู้พูด (Multi Speaker)
3. แบบไม่ขึ้นกับผู้พูด (Speaker Independence)

โดยแต่ละแบบยังสามารถแบ่งย่อยการศึกษาตามลักษณะเสียงคำพูดได้อีก 3 แบบคือ

1. แบบเสียงคำพูดโดด (Isolate Speech)
2. แบบคำพูดต่อกัน (Connected Speech)
3. แบบเสียงคำพูดต่อเนื่อง (Continuous Speech)

ลักษณะการรู้จำเสียงคำพูดตัวเลขในภาษาไทย สามารถแบ่งส่วนสำคัญได้ 2ส่วน

1. ส่วนการหาพารามิเตอร์เสียงคำพูด

เป็นส่วนที่สำคัญของการรู้จักคำพูด เพราะความถูกต้องของการรู้จำเสียงคำพูดขึ้นอยู่กับการเลือกพารามิเตอร์ ซึ่งใช้แทนลักษณะเฉพาะของเสียงคำพูด เช่น ความถี่ฟอร์แมนท์ สเปกตรัมดิสแตนซ์, ค่าเฉลี่ยของระดับพลังงาน

2. ส่วนการตัดสินใจ

เป็นส่วนที่ทำการวิจัยกันมากในงานวิจัยที่ผ่านมา ต่างก็ใช้เทคนิคที่ต่างกันในส่วนนี้ เช่น ไดนามิคไทม์วาร์ปปิง, เทคนิคการจัดกลุ่ม

งานวิจัยนี้ผู้วิจัยจึงมีแนวคิดที่จะศึกษาวิจัยการรู้จักเสียงคำพูด และนำมาประยุกต์เพื่อนำเอาท์พุทไปควบคุมสวิทช์ เปิด-ปิด อุปกรณ์ไฟฟ้า โดยใช้สเปกตรัมของสัญญาณเสียงคำพูดเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พารา-มิเตอร์ ระบบรู้จำเสียงพูด เป็นแบบขึ้นกับผู้พูด (Speaker dependence) โดยเสียงคำพูดเป็นแบบคำโดด (Isolate Speech)

1.2 วัตถุประสงค์

1. ศึกษาการวิเคราะห์หาพารามิเตอร์ ที่เป็นลักษณะเด่นของเสียงคำพูดในกลุ่มคำ 7 คำ ดังนี้ " หนึ่ง ", "สอง", "สาม", "สี่", "ห้า", "ON", "OFF"
2. ศึกษาการแยกพยางค์จากประโยคที่พูด
3. ศึกษาวิธีการในการรู้จำเสียงคำพูดทั้ง 7 คำ
4. ศึกษาวิธีการประมวลผลเพื่อให้ได้ผลลัพธ์ที่ถูกต้อง

1.3 ขอบเขตงานวิจัย

1. การวิจัยนี้มุ่งเน้นศึกษาบนเครื่องคอมพิวเตอร์ โดยมีอุปกรณ์เพิ่มเติม ได้แก่ การ์ดเสียง (Sound Card) , ไมโครโฟน และลำโพง เป็นต้น
2. เสียงคำพูดถูกแปลงเป็นสัญญาณดิจิตอลด้วยความถี่แซมปลิ้ง 1,6000 Hz และความละเอียด 8 bit / Sample
3. การวิจัยนี้สามารถคำนวณและนำเอาท์พุทออกไปควบคุมสวิทซ์ได้ 5 ช่อง

1.4 ขั้นตอนการวิจัย

1. ศึกษาระบบการรู้จำเสียงคำพูดในลักษณะต่างๆ ที่ได้มีการศึกษามาแล้ว
2. ศึกษาทฤษฎีการสร้างเสียงพูด และอัลกอริทึมในการวิเคราะห์เสียงพูด
3. กำหนดขอบเขตของงานวิจัยและแนวทางในการปฏิบัติ
4. ศึกษาการเขียนโปรแกรมด้วยโปรแกรม MATLAB
5. ศึกษาหาอัลกอลิทึมในการแปลงจำนวนในทางคณิตศาสตร์ กลับเป็นเสียงคำพูด
6. บันทึกเสียงพูดชุดหนึ่งมี 7 คำ คือ "หนึ่ง " , "สอง", "สาม", "สี่", "ห้า" , "ON", "OFF" ตามลำดับ การบันทึกเสียง 1 คำใช้ Data Base 70 คำซึ่งใช้ 7*70=490 คำ และ Data Test 7*70=490 คำ
7. เตรียมข้อมูลวิเคราะห์ และ ทดสอบ
8. สรุปผลการวิจัยประเมินผลในการวิจัย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย

1. ทราบถึงพารามิเตอร์ที่ควรใช้เป็นข้อมูลในการตัดเสียงคำพูดของกลุ่มคำทั้ง 7 คำ
2. เป็นแนวทางการพัฒนาไปสู่ระบบรู้จำคำพูดคำอื่นๆต่อไป
3. สามารถนำไปประยุกต์ใช้งานในด้านต่างๆ เพื่อเพิ่มประสิทธิภาพการติดต่อกับเครื่องมือต่างๆ

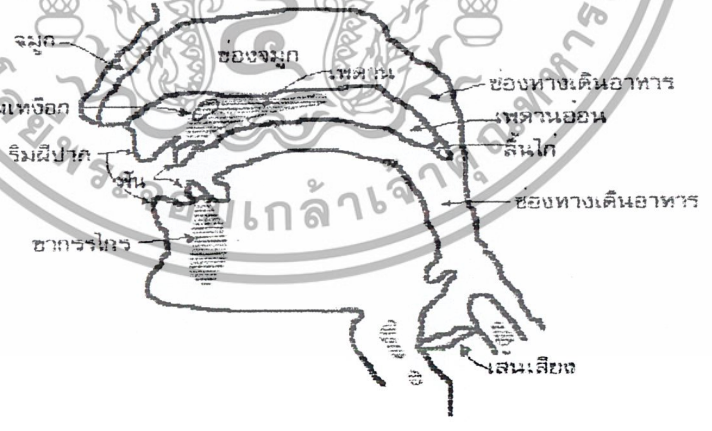
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

หลักการและทฤษฎี

2.1 ทฤษฎีการสร้างเสียงพูด

มนุษย์สร้างเสียงพูดขึ้นมาโดยมนุษย์จะมีเส้นเสียง (Vocal Cord) ในลำคอ ที่ทำให้เกิดเสียง ซึ่งอยู่ในหลอดลม เส้นเสียงมีลักษณะเป็นกล้ามเนื้อ 2 แผ่น ปิดขวางอยู่ปากช่องหลอดลมจากด้านหลังมาด้านหน้า ระหว่างเส้นเสียงจะมีช่องว่างซึ่งเป็นทางให้ลมผ่านเข้าไปปอด และออกจากปอดได้ ช่องนี้เรียกว่า ช่องระหว่างเส้นเสียง (Glottis) เส้นเสียงทั้งสองเส้นสามารถจะดึงออกให้ห่างจากกัน หรือดึงเข้ามาประชิดกันได้ เส้นเสียงเป็นส่วนที่ทำให้เกิดเสียงพูด และจากการสั่นของเส้นเสียงทำให้เกิดคลื่นเสียงความถี่ต่างๆ ผ่านเข้าสู่ลำคอ ผ่านลำคอสู่ปาก หรือช่องจมูกออกไปภายนอก ซึ่งขนาดและรูปร่างของอวัยวะภายในช่องปากนี้ จะเป็นสิ่งกำหนดว่าคลื่นเสียงความถี่ไหนจะยอมให้ปรากฏออกมาให้ได้ยินหรือคลื่นความถี่ไหนจะถูกดูดซับไว้ไม่ยอมให้ปรากฏออกมาอวัยวะส่วนที่มีหน้าที่ในการออกเสียงพูด แสดงดังรูป 2.1



รูปที่ 2.1 แสดงอวัยวะที่ใช้ในการออกเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 การเกิดของเสียง (Speech Production)

ขั้นตอนการเกิดของเสียงแบ่งได้เป็น 3 ขั้นตอนใหญ่ๆ คือ

1. ขั้นตอนการเริ่มต้น (Initiation) ขั้นตอนนี้เป็นขั้นตอนที่ลมเริ่มขับออกมาจากปอดเพื่อเข้าสู่ขั้นตอนที่ 2 ต่อไป

2. ขั้นตอนดัดแปลงลมที่เส้นเสียง (Probation) เป็นขั้นตอนที่ลมออกจากปอดจะผ่านมายังหลอดลมและกล่องเสียง ซึ่งที่กล่องเสียงนี้ เส้นเสียงจะทำหน้าที่เป็น ลิ้น เปิด-ปิด ทำให้เกิดเสียงได้ 2 ชนิด คือเสียง ก้อง (Voiced Sound) และ เสียงไม่ก้อง (Unvoiced- Sound) อวัยวะที่ใช้ในตอนนี้คือ ส่วนที่ต่อจากปอดขึ้นมาที่ริมฝีปาก

เสียงที่เกิดจากขั้นตอนดังกล่าว จะถูกแบ่งเป็น 2 ประเภท คือ

1. เสียงก้อง (Voiced Sound) เกิดขึ้นเนื่องจากการออกเสียงในลักษณะที่เส้นเสียงถูกดึงมาใกล้กันจนเกือบปิดช่องทางลมเสียสนิท ลมที่ดันขึ้นมาจากปอดจะทำให้เส้นเสียงสั่นลมที่ออกมาไม่สะดวก เพราะต้องบีบผ่านช่องเป็นจังหวะจึงทำให้เป็นเสียงขึ้นเรียกว่า " เสียงก้อง "

2. เสียงไม่ก้อง (Unvoiced Sound) เป็นการออกเสียงในขณะที่เส้นเสียงยังเปิดกว้างโดยเปิดช่องระหว่างเส้นเสียงหรือช่องคอหอย (Glottis) ให้ลมหายใจผ่านเข้าออกสะดวก เรียกเสียงเหล่านี้ว่า " เสียงไม่ก้อง "

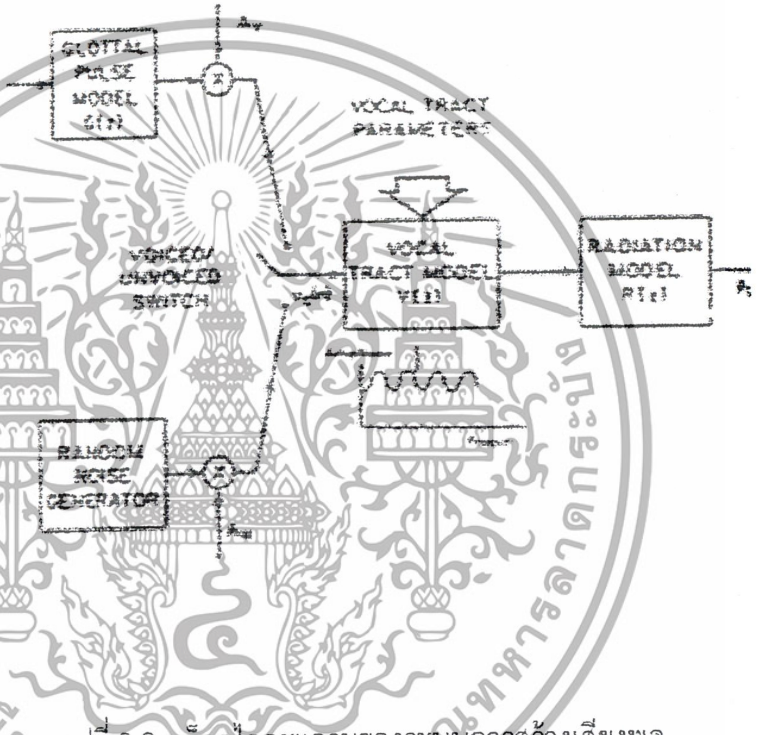
อวัยวะที่ใช้ในการออกเสียงนั้น อาจจะแบ่งคร่าวๆ ได้เป็น 2 พวก คือ

1. อวัยวะส่วนที่กระทำอาการ (Articulator) หมายถึง อวัยวะส่วนที่เคลื่อนไหวเพื่อผลักลมไปยังส่วนต่างๆ อวัยวะที่เป็นส่วนการะทำอาการที่สำคัญ คือ ลิ้น ซึ่งเป็นส่วนที่เคลื่อนไหวได้มากที่สุด

2. อวัยวะที่เป็นฐานที่เกิดเสียงต่างๆ (Point of articulator) หมายถึง ตำแหน่ง หรือฐานที่ เกิดเสียงต่างๆ เช่น ริมฝีปาก ฟัน ปุ่มเหงือก เพดานส่วนต่างๆ เป็นต้น

กระบวนการเกิดเสียงพูดของมนุษย์ที่กล่าวมา โดยสามารถแสดงบล็อกไดอะแกรม (Block Diagram) อย่างง่าย ดังแสดงดังรูป 2.2

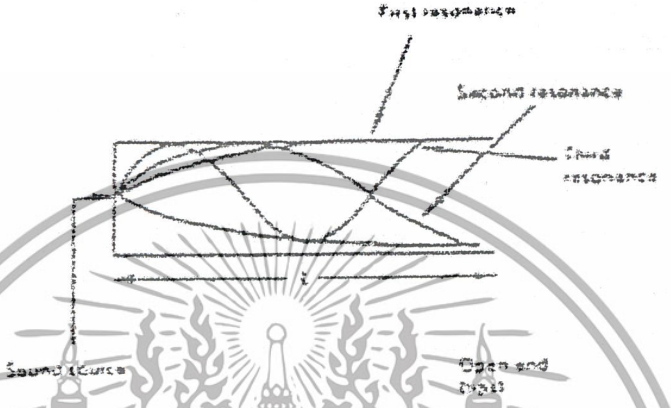
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.2 บล็อคไดอะแกรมของระบบการสร้างเสียงพูด

ภาคสร้างขบวนอิมพัลส์ (Impulse Train Generator) จะสร้างขบวนอิมพัลส์ ซึ่งมีระยะห่างระหว่างอิมพัลส์เท่ากับ คาบเวลาพิทช์ (Pitch Period) ขบวนอิมพัลส์ นี้จะถูกส่งไปกระตุ้นโมเดลช่องเส้นเสียง (Glottal Pulse Model) , G(z) ทำให้สัญญาณออกจากโมเดลตามจังหวะค่าเวลาพิทช์ ผลทางสเปกตรัมของ สัญญาณที่ออกจากโมเดลนี้มีความชันประมาณ –12 dB/Octave รูปแบบจำลองอย่างง่ายๆของช่องทางเดินเสียง อาจจะมองได้เป็นลักษณะของท่อทรงกลมต้นที่มีต้นกำเนิดอยู่ที่ปลายข้างหนึ่ง (ส่วนของกล่องเสียง) และปลายอีกข้างหนึ่งจะเปิด (ส่วนของปาก) แสดงดังรูป 2.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.3 แบบจำลองอย่างง่ายของการกำทอนช่องทางเดินเสียง

ดังนั้นจะเกิดการกำทอน (Resonance) ภายในท่อ การกำเนิดของช่องทางเดินเสียงนี้จะสร้างรูปคลื่นเสียงที่มียอดสูงเด่นขึ้น ซึ่งเรียกว่า ฟอร์แมนท์ (Formants) ของเสียง ฟอร์แมนท์ที่มีความถี่ต่ำที่สุดจะเรียกว่า ฟอร์แมนที่หนึ่งซึ่งจะมีค่าประมาณ 200 ถึง 1000 Hz ทั้งนี้ขึ้นกับขนาดของช่องทางเดินของเสียงด้วย ส่วนฟอร์แมนท์ที่สองอยู่ถัดไปก็จะมีค่าประมาณ 500 ถึง 2500 Hz และฟอแมนท์ที่สามมีค่าประมาณ 1500 ถึง 3500 Hz เป็นต้นโดยที่ฟอแมนท์ที่หนึ่ง และสองนี้จะเป็นตัวแสดงคุณสมบัติที่สำคัญมากอันดับหนึ่งในการพิจจารณาประยุกต์ใช้ในด้านการวิเคราะห์เสียงด้วยสเปกตรัม เสียงที่มีความถี่กำทอนตรงกับความถี่กำทอนจะปรากฏออกมาให้ได้ยินผ่านทางภาคโมเดลการแพร่ (Radiation Model) , R(z) ซึ่งเป็นส่วนของปาก ผลทางสเปกตรัมของสัญญาณที่ออกจากโมเดลที่มีความชันประมาณ –6 dB/Octave

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 ทฤษฎีการวิเคราะห์เสียงพูด

การวิเคราะห์เสียงพูด คือ พารามิเตอร์ต่างๆ ที่มีลักษณะเฉพาะ เพื่อใช้เป็นตัวแทนเสียงการวิเคราะห์เสียงพูดนั้นมีหลายวิธี ทั้งในด้านโดเมนเวลา (Time domain) คือ วิเคราะห์จากรูปคลื่นของสัญญาณเสียงตามแกนเวลาโดยตรง และด้านโดเมนความถี่ (Frequency domain) คือการวิเคราะห์จากสเปกตรัมของเสียง

2.4 การวิเคราะห์สัญญาณเสียงในช่วงสั้น ๆ

เนื่องจากสัญญาณเสียงเป็นสัญญาณที่แปรตามเวลาที่มีการเปลี่ยนแปลงไม่แน่นอน ดังเทคนิคในการวิเคราะห์เสียงพูดส่วนใหญ่แล้ว จะสมมติให้สัญญาณเสียงมีคุณสมบัติที่เปลี่ยนแปลงสัมพันธ์กับเวลาอย่างช้าๆ นั่นก็คือเราจะทำการแบ่งวิเคราะห์พารามิเตอร์ ของสัญญาณเสียงพูดในช่วงสั้นๆ ที่เรียกว่า ช็อตทามวินโดว์ (Shot time Window) เพื่อจะมอง เห็นเหมือนกับว่าเราหาพารามิเตอร์นั้นๆ ได้มาจากสัญญาณเสียงที่อยู่ในวินโดว์ และมีความเสถียรภายในช่วงสั้นๆ

การนำฟังก์ชันของวินโดว์ที่มีช่วงขนาดจำกัด W(n) มาคูณเข้ากับสัญญาณ S(n) จะทำให้ได้กลุ่มตัวอย่างของเสียงที่ถูกกำหนดน้ำหนัก ให้แปรไปตามรูปร่างของวินโดว์ รูปแบบของวินโดว์ที่ใช้ในงานวิจัย เลือกแบบกรอบสี่เหลี่ยม

(Rectangular Window) ซึ่งมีนิยามดังนี้

$$W(n) = \begin{cases} 1, n = 0,1,2,3,\ldots n \\ 0, n = \text{otherwise} \end{cases}$$

ในสมการที่(1) นี้คือการกำหนดช่องวิเคราะห์ให้มีจำนวนตัวอย่าง N ตัวอย่างโดยปกติ การประยุกต์ส่วนมากจะใช้วินโดว์ที่มีความกว้าง 10 ถึง 30 มิลลิวินาทีและเน้นเฉพาะช่องตรงกลางของช่องแคบให้เป็นส่วนของข้อมูลที่จะทำการวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5 การวิเคราะห์โดเมนเวลา

การวิเคราะห์โดเมนเวลา คือ วิธีการประมวลสัญญาณเสียงพูดโดยยึดหลักการวิเคราะห์จากรูปคลื่นสัญญาณที่แปรตามเวลา ซึ่งมีอยู่หลายวิธีด้วยกัน การวิเคราะห์โดเมนเวลาที่ใช้ในงานวิจัยดังต่อไปนี้ 2 ค่าอัตราการตัดศูนย์เฉลี่ยของเวลาช่วงสั้น (Shot time Average Zero-Crossing Rate) ในหัวข้อนี้กล่าวถึงสัญญาณ Discrete-time ,อัตราการตัดศูนย์ (Zero-crossing) ที่กล่าวถึงถ้าเกิดมีสัญญาณตัวอย่างที่ทำให้เกิดเครื่องหมายทางคณิตศาสตร์ที่แตกต่างกัน อัตราการตัดศูนย์เป็นการวัดสัญญาณความถี่ตัวอย่างซึ่งรวมกันอยู่ในรูปสัญญาณแคบๆ ตัวอย่าง สัญญาณรูปซายน์ของ ความถี่ = F0, อัตรา Sample = Fs จะได้ Fs/F0 คือจำนวน sample ต่อ สัญญาซายน์ 1 รูปคลื่น แต่ละ 1ลูกคลื่นจะมีการตัดศูนย์ (Zero-Crossing) 2 ครั้ง จึงให้ค่าเฉลี่ยอัตราการตัดศูนย์ คือ

$$Z=2*F0/Fs \quad (\text{ zero crossing / sample })$$

ดังนั้น อัตราการตัดศูนย์ จึงให้เหตุผลในการคาดเดาความถี่ของสัญญาณซายน์ (Sine Signal) สัญญาณเสียงเป็นสัญญาณบรอดแบนด์ (Broadband) และการแปลงอัตราเฉลี่ยการตัดศูนย์มีความเที่ยงตรงน้อยมาก อย่างไรก็ตาม การคาดคะเนแบบหยาบๆคุณสมบัติที่ได้รับโดยการแทนฐานเวลาที่สั้นทำให้เกิดอัตราการตัดศูนย์ก่อนที่จะกล่าวถึงการเปลี่ยนแปลงอัตราการตัดศูนย์ สำหรับสัญญาณเสียงพูด อย่างแรกที่จะกำหนดคำนวณ และอธิบาย กำหนดนิยามได้ดังนี้

$$Zn = \sum_{m=-\alpha}^{\alpha} |Sgn[X(m) - Sgm[X(m-1)]| W(n-m)$$

ซึ่ง

$$Sgn\ [\ X\ (m)\ ] = 1 \quad X(n) \geq$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

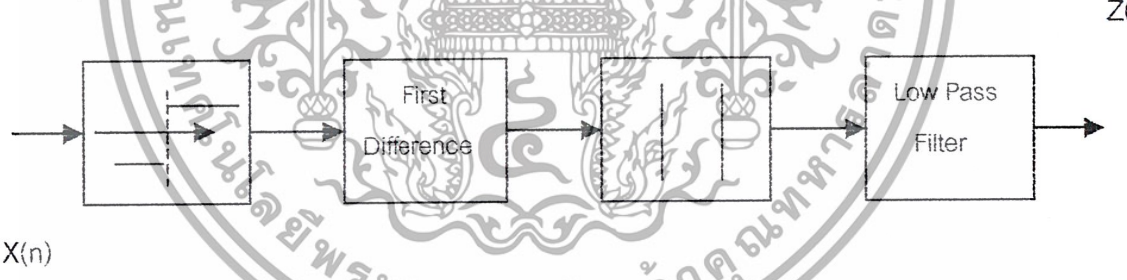
$$= -1\ X(n) < 0$$

และ

$$W(n) = 1/2N \qquad 0 \leq n\ N\text{-}1 \leq N\text{-}1$$

$$= 0$$

ดังนั้น การกระทำที่เกี่ยวข้องกับ สมการ (2.2) จะถูกแสดงในรูปที่ 2.4 ซึ่งแทนการหาค่าอัตราการตัดศูนย์เฉลี่ยในช่วงเวลาสั้นๆไปเหมือนกับการหาค่าพลังงานและค่าสมบูรณ์เฉลี่ยช่วงเวลาสั้นๆ อย่างไรก็ตามในสมการ (2.2) และรูปที่ 2.4 จะทำการคำนวณค่า Zn ซึ่งเกิดขึ้นยากกว่าความเป็นจริงซึ่งทั้งหมดต้องการตรวจสอบ 2 สัญญาณ ซึ่งจะถูกกำหนดค่าอัตราเฉลี่ยการตัดศูนย์จะถูกคำนวณมากกว่า N สัญญาณตัวอย่างที่อยู่ติดๆกัน (ถูกแบ่งโดยเป็นสัญญาณที่ไม่ต้องการ)



รูปที่2.4บล็อกไดอะแกรมของอัตราการตัดศูนย์เฉลี่ยในเวลาช่วงสั้น

จากที่กล่าวมาสามารถบอกลำดับความสำคัญของค่าเฉลี่ย และถ้ากำหนดในรูปสมการ โดยกำหนดความยาวของหน้าต่าง (window) สมมาตร การถ่วงเวลาจะสามารถชดเชยได้อย่างถูกต้อง ตัวอย่างนี้เรารู้ว่าเกิดอัตราการตัดศูนย์ (Zero-Crossing Rate) เฉลี่ยในช่วงสั้นๆ จะนำไปประยุกต์ใช้เป็นสัญญาณการพูด ซึ่งมีความถี่ต่ำกว่า 3 KHz การเกิดสเปกตรัม (Spectrum) จะถูกบอกโดยคลื่นเสียง อย่างไรก็ตามเสียงสีฟัน (Unvoiced) ส่วนมากพลังงานจะถูกพบที่ความถี่สูงมากๆ จึงกำหนดให้เป็นอัตราการตัดศูนย์สูง (High –zero-crossing) และที่ความถี่ต่ำเป็นอัตราการตัดศูนย์ต่ำ (Low-zero-crossing) ซึ่งมีความสัมพันธ์กัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อย่างเหนียวแน่นระหว่างอัตราการตัดศูนย์และการแจกแจงพลังงานกับความถี่ เป็นเหตุผลทั่วๆ ไปถ้าอัตราการตัดศูนย์มาก (Low-zero-crossing) สัญญาณเสียงพูดจะเป็น เสียงเสียดสีฟัน (Unvoiced) ขณะเดียวกันถ้าอัตราการตัดศูนย์ต่ำสัญญาณเสียงพูดจะเป็น (Voice) อย่างไรก็ตามความคิดนี้อาจไม่เที่ยงตรงนักเพราะมนุษย์มีเสียงพูดทั้งสูงและต่ำเป็นไปไม่ได้ที่จะถูกต้อง 100% รูปที่ 2.5 จะแสดง Histogram ของอัตราการตัดศูนย์ เฉลี่ยประมาณมากกว่า 10 ms สำหรับทั้งเสียงเสียดสีฟัน (Voice) และไม่เสียดสีฟัน (Unvoiced) สังเกตจากรูปเคิร์ปของเกาส์เซียน เป็นการแจกแจงที่ดีมีเหตุผลค่าประมาณของอัตราการตัดศูนย์เฉลี่ยของ เสียงเสียดสีฟันประมาณ 40/10 ms สำหรับเสียงไม่เสียดสีฟันประมาณ 14/10 จะเห็นได้ชัดเจนว่าอัตราการตัดศูนย์ จะเกิดเหลื่อมล้ำ (Overlap) ดังนั้นเสียงเสียดสีฟัน และเสียงไม่เสียดสีฟัน จะใช้เป็นพื้นฐานการตัดสินใจเพียงอย่างเดียวไม่ได้ แต่การแทนค่า เสียงเสียดสีฟัน และเสียงไม่เสียดสีฟัน จะทำให้เกิดประโยชน์ในการทำให้เกิดความแตกต่างได้

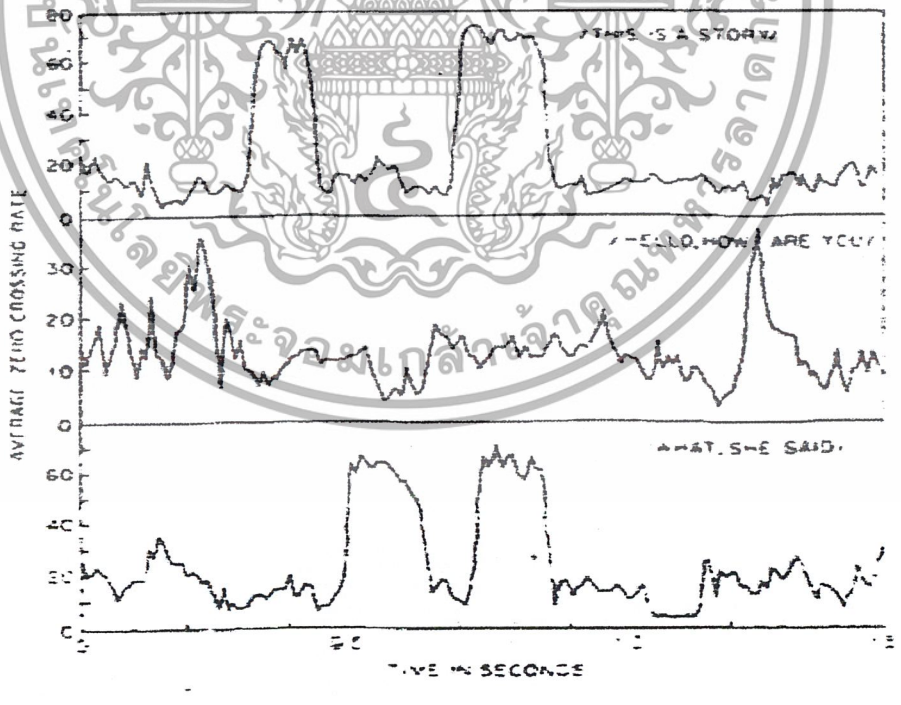
สำหรับตัวอย่างค่าอัตราการตัดศูนย์เฉลี่ย จะแสดงในรูปที่ 2.6 ช่วงเวลาที่เปลี่ยนแปลงจะครอบด้วย วินโดว์ที่เวลา 15 ms (150 ครั้ง ที่ 10 KHz) เอาท์พุทที่ได้จะถูกคำนวณที่ 100 Time/ sec (window เลื่อนไป 100 ครั้ง) ข้อสังเกตในกรณีของพลังงานและค่าสัมบูรณ์เฉลี่ยต่ำมาก ดังนั้นอัตราการตัดศูนย์จะต้องพิจารณา เสียงเสียดสีฟันและเสียงไม่เสียดสีฟัน มีลักษณะสูงๆ ต่ำๆ ดังรูปที่ 2.5 ตัวอย่างอัตราการตัดศูนย์เฉลี่ยพิจารณาได้จากรูปที่ 2.6 ดังนั้นจึงใช้ความคิดพื้นฐาน (Algorithm) สำหรับคำนวณเกี่ยวกับค่าเปรียบเทียบสัญญาณการสุ่มทั้งคู่ต้องใส่ใจอย่างมากในกระบวนการสุ่ม และเห็นได้ชัดเจนว่า อัตราการตัดศูนย์ จะมีผลอย่างมากในการเปลี่ยนแปลงสัญญาณอะนาลอกเป็นสัญญาณดิจิตอล (A/C converter) ซึ่งจะมีสัญญาณ ฮัม (Hum) ออกมา 60 Hz ซึ่งแต่ละเสียงสามารถแทนในระบบดิจิตอล แต่ต้องระมัดระวังอย่างมากในการกระทำกระบวนการอะนาลอก สำหรับตัวอย่างจะใช้วงจรกรองแถบความถี่ผ่าน (Band pass filter)มากกว่าวงจรความถี่ต่ำผ่าน(Low pass-filter) หรืออีกในหนึ่งเป็นการกรองที่บล็อคไฟกระแสตรง และส่วนประกอบ 60 Hz ในสัญญาณ นอกเหนือจากนี้เรายังใช้อัตราการตัดศูนย์ การสุ่มช่วงเวลาโดยให้เป็น T และช่วงเฉลี่ย N

ช่วงคาบเวลาของการสุ่มจะกำหนดเวลา (และความถี่) การแจกแจงโดยการแทนอัตราการตัดศูนย์ เป็นการเตรียมให้รายละเอียด 1 บิต (การเตรียมนี้จะเป็นการให้เครื่องหมายทางคณิตศาสตร์) ซึ่งทั้งหมดถูกต้องเพราะว่าการกำหนดจำนวนอย่างเหมาะสม จะมีหลากหลายมากในความเหมือนกันในการแทนค่า จะถูกเสนอมาใช้งานความแตกต่างเหล่านี้จะทำให้เกิดความน่าสนใจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในความไวของเสียงที่มีความไวน้อย แต่ละเสียงจะทีค่ากำหนดของแต่ละคน ข้อสังเกตเราจะทำการแทนค่าการตัดช่วงบน (Up-Crossing) การแสดงนี้จะเป็นพื้นฐานของการหาอัตราการตัดศูนย์ที่เกิดในช่วงสโลปบวก

การเปรียบเทียบความแตกต่างของเสียงกับความเงียบโดยใช้พลังงาน และอัตราการตัดศูนย์ปัญหาของการหาตำแหน่งเริ่มต้น และตำแหน่งสุดท้ายของเสียงเป็นสิ่งที่สำคัญในขบวนการเสียง โดยทั่วไปการบันทึกเสียงอัตโนมัติของคำโดด มีความจำเป็นที่จะบอกตำแหน่งกำเนิดสัญญาณ เสียง และการติดต่อแต่ละเสียงสำหรับการคำนวณหาตำแหน่งเริ่มต้นและ ตำแหน่งสุดท้ายทำได้โดย การกำจัดเครื่องหมายของการคำนวณสัญญาณในระบบเวลาที่ไม่จริง โดยการกระทำให้เป็นอินพุท เพื่อบอกเป็นเสียงพูดปัญหาการแยกคำพูดจากเสียงไม่สำคัญยกเว้นในกรณีที่ว่า อัตราส่วนสัญญาณต่อสัญญาณรบกวนมีค่าสูง เช่น ในห้องอัดเสียง การบันทึกเสียงไฮไฟในห้องอัดเสียงที่ไม่มีเสียงเล็ดลอดออกไปได้เลย



รูปที่ 2.5 อัตราการเปลี่ยนแปลงของเสียงพูดที่แตกต่างกัน3ครั้ง

สำหรับอัตราส่วนสัญญาณต่อสัญญาณรบกวนที่มีค่าสูงพลังงานของคำพูดเสียงก็จะต่ำที่สุด ดังนั้นวิธีการนี้จะใช้ได้กับการแยกเสียงอย่างง่าย โดยการบันทึกนี้จะไม่ได้เป็นวิธีที่ดีที่สุด แนวความคิดพื้นฐาน

เอกสารนี้มีลิขสิทธิ์ ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(Algorithm) จะถูกอธิบายโดยพื้นฐานอย่างง่าย 2 ข้อ ในรูปโดเมนเวลา คือพลังงาน และอัตราการตัดศูนย์ ตัวอย่างง่ายๆ ทั้งหมดจะแสดงให้เห็นถึงความยากในการหาจุดเริ่มต้นและจุดสุดท้ายของของเสียงพูดที่ออกมา กรณีของแนวความคิดทั้งหมด สามารถกำหนดจุดสิ้นสุดอย่างถูกต้องดังนั้นมันจึงมีเหตุผลเพียงพอสำหรับกำหนดช่วงเวลาของ Unvoiced ซึ่งรวมอยู่ภายในคำพูดทั้งหมด การประยุกต์ใช้งานของอัตราการตัดศูนย์ นี้ค่าสมบูรณ์ขนาด (Magnitude) เฉลี่ยจะทำให้ไม่เกิน

2.6 การประเมินค่าช่วงเวลาพิทซ์ โดยวิธีการเข้าใกล้เส้นขนาน

ช่วงเวลาพิทซ์เป็นสัญญาณที่มีความสำคัญมากในกระบวนการพูด การถอดค่าพิทซ์ถูกใช้ในการกำหนดแยกแยะเสียงพูด ซึ่งมีหลักการพื้นฐานสำคัญดังนี้

1. สัญญาณการพูดเป็นกรรมวิธีในการสร้างสัญญาณอิมพัลส์ ซึ่งจะรักษาช่วงเวลาเริ่มต้นและทั้งโครงสร้างที่ไม่ต้องการโดยกรรมวิธีถอดค่าพิทซ์

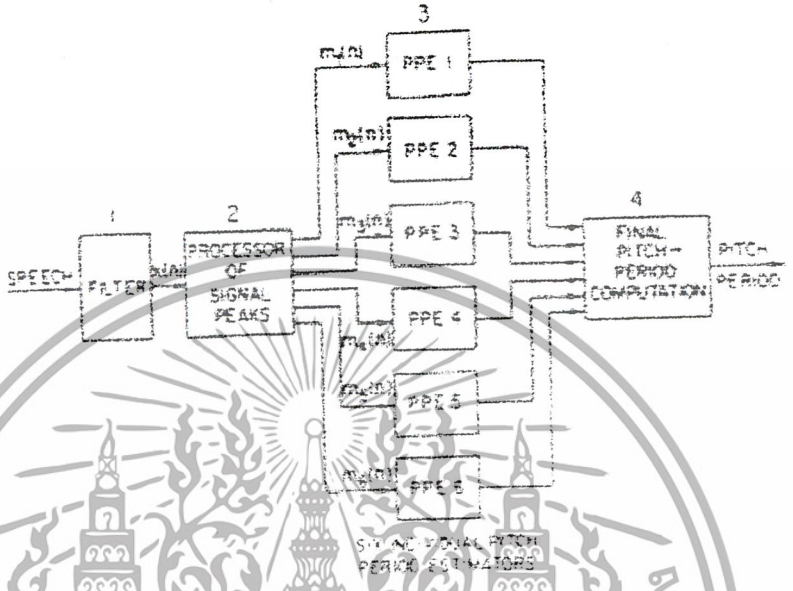
2.ขบวนการที่ยอมให้ถอดค่าพิทซ์อย่างง่ายถูกใช้ประมาณช่วงเวลาของแต่ละขบวนอิมพัลซ์

3.การคาดคะเนของทุกๆสัญญาณการถอดค่าพิทซ์ทั้งหมดเป็นการรวมกันในสัญญาณลอจิกซึ่งจะสรุปรูปแบบช่วงเวลาของคำพูด

แผนผังนี้ถูกกำหนดโดยข้อคิดเห็นของ Gold และ Robiner ดังอธิบายในรูปที่ 2.7 รูปแบบของคำพูดถูกทดสอบด้วยอัตราส่วนที่เพียงพอ เพื่อให้จำนวนเวลาที่พอเพียง การทดสอบที่ 10 KHz ยอมให้ช่วงเวลาที่กำหนดภายใน 1/10000 sec คำพูดนี้จะเป็นตัวกรองความถี่ต่ำผ่านมีจุดคัทออฟที่ 900 Hz เพื่อไปผลิตทำให้เกิดรูปแบบที่มีความสัมพันธ์กับตัวกรองแถบความถี่ผ่านความถี่ที่ 100 Hz และ 900 Hz มีความจำเป็นในการขจัดสัญญาณเสียงที่ 60 Hz ในการใช้งานฟิลเตอร์นี้สามารถทำได้ทั้งคู่ด้วย

การขจัดสัญญาณเสียงที่ 60 Hz ในการใช้งานฟิลเตอร์นี้สามารถใช้งานทั้งคู่ได้ด้วยอนาลอกฟิลเตอร์ก่อนการสุ่มด้วยอนาลอกฟิลเตอร์ จากการกรองที่ตำแหน่งสูงสุดจากตำแหน่งเหล่านี้และแอมปริจูดทุกๆอิมพัลซ์ในรูปที่ 2.7 ได้มาจาก สัญญาณฟิลเตอร์แต่ละอิมพัลซ์ประกอบไปด้วยอิมพัลส์บวกที่เกิดด้วยตำแหน่งของ Peak และ Valley โดยทั้ง 6 กรณีนี้โดยจะใช้หลักการของ Gold และ Robiner

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.6 บล็อกไดอะแกรมของการถอดสัญญาณพิทช์แบบขนาน

1. M1(n): เมื่ออิมพัลส์มีขนาดแอมปลิจูดสูงสุดที่เท่ากันเกิดขึ้นที่แต่ละจุดของ Peak

2.M2(n): อิมพัลส์ที่เท่ากันแต่มีความแตกต่างกันที่แอมปลิจูดและเกิดด้านแอมปลิจูดของ Valley ซึ่งเกิดขึ้นแต่ละ Peak

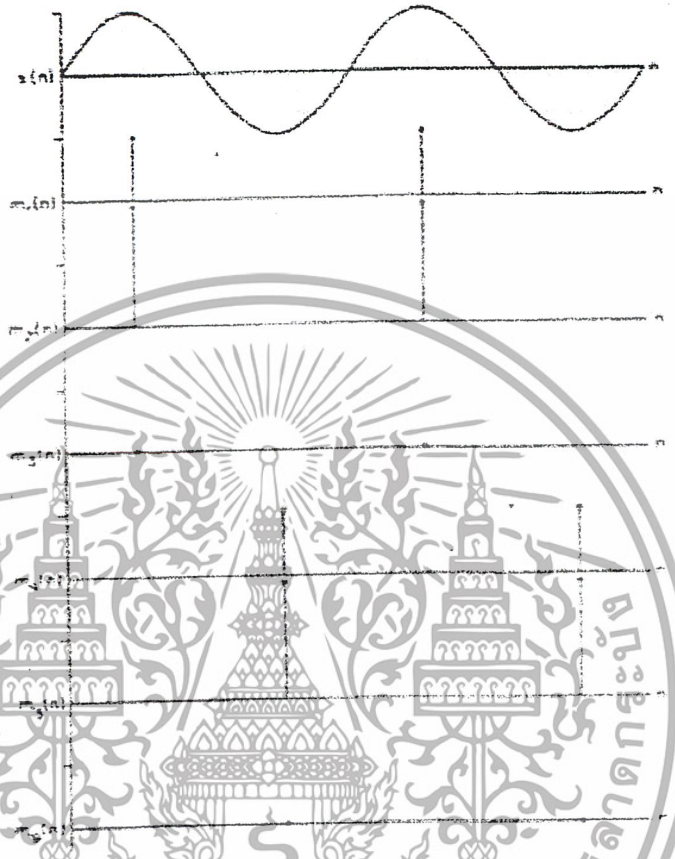
3. M3(n):อิมพัลส์ที่กันและแตกต่างกันที่แอมปลิจูดที่สูงสุดและเกิดก่อนแต่ละ Peak

4. M4(n):อิมพัลส์ที่เท่ากันและเป็นลบที่แอมปลิจูดของ Valley

5. M5(n): อิมพัลส์ที่เท่ากันเป็นแอมปลิจูดที่ลบของ Valley บวกแอมปลิจูดที่อยู่หน้า Peak เกิดขึ้นแต่ละช่วงของ Valley

6 .M6(n): อิมพัลส์ที่เท่ากันจะเป็นลบของแอมปลิจูดที่ Valley รวมกับแอม ปลิจูดที่เกิดขึ้นก่อน ที่เกิดขึ้นที่จุดต่ำสุดที่แต่ละ Valley

เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.7 อินพุทสัญญาณซายน์และการเกิดขบวนการอิมพัลส์จาก peak และ valley

ผลเหล่านี้จะเกิดการตัดสินใจของแนวคิด ซึ่งต้องถ่วงเวลา 3 ช่วงยอดเพื่อให้อยู่ในกรรมวิธีนี้ ในการศึกษาการเปรียบเทียบในปัจจุบันนี้ การถอดยอดแนวคิดจะสามารถทำงานภายใต้สภาวะที่กว้างมาก การเปรียบเทียบวิธีนี้จะดีกว่าการกะประมาณยอดซึ่งแสดงให้เห็นแล้ว

ในการสรุปวิธีรายละเอียด กรรมวิธีอย่างแรก กรรมวิธีที่สำคัญที่สัญญาณเสียงได้ทำการเซตอิมพัลส์ ซึ่งมีความจำเป็น (หรือการสูญเสียของช่วงเวลา)เพราะโครงสร้างของสัญญาณอย่างง่ายในการประมาณค่ายอดทำให้เกิดการประมาณค่ายอดที่ดีสุดท้ายการประมาณการก็จะรวมกันเพื่อเพิ่มความเชื่อถือ การประมาณค่ายอดให้เป็นที่ยอมรับดังนั้นจะเป็นกรรมวิธีที่อย่างง่ายที่จะบรรลุการคาดเดาที่ยาก เพื่อให้ได้สัญญาณที่ถูกต้อง การดำเนินการทางลอจิก สามารถทำให้ได้ในอัตราที่ต่ำมากกว่ากรรมวิธีของสัญญาณ

เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.7 การวิเคราะห์ตามแกนความถี่

คือ การวิเคราะห์ทางสเปกตรัมนั่นเอง สาเหตุที่ต้องมีการวิเคราะห์ทางด้านสเปกตรัมเนื่องมาจากพารามิเตอร์ในการประมวลสัญญาณเสียงส่วนใหญ่จะหามาได้จากโดเมนความถี่ กาวิเคราะห์สเปกตรัมของสัญญาณในช่วงเวลาสั้นๆ นั้นจะเป็นฐานทางเทคนิคการวิเคราะห์ของเสียงต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

บทที่ 3

โครงสร้างงานวิจัย

โครงสร้างงานวิจัย

งานวิจัยนี้ได้แบ่งโครงสร้างเป็น 4 ส่วนใหญ่ๆ คือ

3.1 ภาคอ่านไฟล์ของเสียงพูด

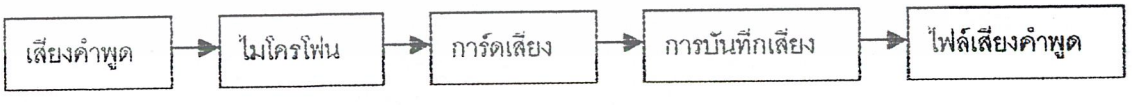
3.2 ภาครู้จำเสียงพูด

3.3 ภาคประมวลผลดัฟซ์

3.4 ภาคแสดงผลโดยการนำไปควบคุมสวิทช์

3.1 ภาคอ่านไฟล์เสียงข้อมูล

ระบบรู้จำเสียงพูดนี้ขึ้นอยู่กับผู้พูด (Speaker Dependent) โดยเสียงคำพูดเป็นแบบคำโดด (Isolated Speech) เสียงพูดจะถูกบันทึกในรูปแบบเวฟ (Wave Format) โดยผ่านทางไมโครโฟนที่ต่อกันอยู่กับการ์ดเสียง (Sound Card) ด้วยความถี่แซมปลิ้ง 16,000 Hz และความละเอียด 8 บิท ต่อแซมปลิ้ง การบันทึกเสียงพูดจะใช้โปรแกรมที่ให้มากับการ์ดเสียงหรือโปรแกรมที่สามารถจัดการเกี่ยวกับเสียงได้ ระบบการบันทึกเสียงพูดสามารถเขียนเป็นบล็อคไดอะแกรมดังแสดงในรูป3.1



รูปที่ 3.1 ระบบการบันทึกเสียง

ภาคอ่านไฟล์เสียงคำพูด จะทำหน้าที่อ่านข้อมูลเสียงจากไฟล์ที่อยู่ในรูปแบบเวฟ โดยในการบันทึกไฟล์ในรูป แบบ mat.file เราจะเรียกใช้ฟังก์ชั่น Record sound ของโปรแกรม MATLAB และเก็บข้อมูลที่อ่านได้ไว้ในไฟล์ C:\pro_data\computer_projectข้อจำกัดของฟังก์ชัน Record sound คือ สามารถบันทึกเสียงได้เพียง 1 วินาที เท่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

42372

3.2 ภาครู้จำเสียงพูด

ทำหน้าที่วิเคราะห์หาความน่าจะเป็นของเสียงพูดว่าเหมือนคำใดมากที่สุด โดยจะประกอบด้วย ภาคแยกพยางค์ออกจากประโยค ภาคหาจุดเริ่มต้นและจุดสุดท้ายของเสียงคำพูด ภาคหาพารามิเตอร์ของเสียงพูด และภาคการตัดสินใจเสียงคำพูด ซึ่งแต่ละภาคมีหลักการทำงานดังนี้

1. ภาคการแยกพยางค์ออกจากประโยค

ภาคนี้จะทำการแยกพยางค์แต่ละพยางค์ให้ออจากประโยคที่ผู้ทดลองได้บันทึกเสียงเข้าไปไว้ในไฟล์เสียง ซึ่งในงานวิจัยนี้จะมีจำนวนพยางค์ 2 พยางค์ ซึ่งเราเรียกใช้ฟังก์ชั่น cut_word ในการแยกพยางค์

2. ภาคหาจุดเริ่มต้นและหาจุดสุดท้าย

ภาคนี้มีหน้าที่หาจุดเริ่มต้นและจุดสุดท้ายของเสียงคำพูดแต่ละคำที่อ่านจากไฟล์เสียงพูด เพื่อนำเอาเฉพาะข้อมูลตรงส่วนที่เป็นสัญญาณเสียงคำพูด จะเป็นการลดจำนวนข้อมูลที่ต้องนำไปวิเคราะห์ในภาคหาพารามิเตอร์เสียงคำพูดด้วย หลักการที่ใช้ในการหาจุดเริ่มต้นและจุดสุดท้ายของเสียงคำพูดคือ การหาค่า Amplitude ของสัญญาณเสียงที่ทำการ Normalize แอมปลิจูด ในช่วงเวลาสั้นๆ ในงานวิจัยนี้จะใช้วินโดว์แบบสี่เหลี่ยม (Rectangular Window) ที่มีความกว้างเท่ากับ 8.16 มิลลิวินาที ทำการครอบแอมปลิจูดของสัญญาณเสียงที่ทำการ Normalize แอมปลิจูดในเวลาตั้งแต่เริ่มต้นถึงจุดสิ้นสุดซึ่งได้มาจากภาคแยกพยางค์จากประโยค

3. ภาคการหาพารามิเตอร์ของเสียงพูด

นำสัญญาณเสียงพูดที่ผ่านการหาจุดเริ่มต้น และจุดสุดท้ายมาทำการแปลงฟู-เรียร์อย่างเร็ว (Fast Fourier Transform) ผลการแปลงจะได้สเปกตรัมของสัญญาณเสียง ที่มีจำนวนข้อมูลเท่ากับข้อครึ่งหนึ่งของจำนวนแซมปลิ้งของสัญญาณเสียง โดยทั่วไปสเปกตรัมของสัญญาณเสียงในการวิจัยนี้จะมีจำนวนข้อมูลประมาณ 500 ถึง 2000 ข้อมูล ซึ่งเป็นจำนวนข้อมูลที่มากเกินไป ดังนั้นจำเป็นต้องมีการลดลง เพื่อให้เหมาะสมกับการนำไปใช้ภาครู้จำเสียงพูดและในการทำวิจัยในครั้งนี้อัตราการเกิดการตัดศูนย์และจำนวนยอดของเสียงใช้เป็นพารามิเตอร์ในการจำแนกเสียงออกเป็นกลุ่มๆซึ่งแบ่งได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

วิธีการดำเนินการวิจัย

4.1 ขั้นตอนการวิจัย

งานวิจัยนี้ได้แบ่งวิธีดำเนินการวิจัยออกเป็น 5 ขั้นตอน ดังนี้

1. การเก็บข้อมูลกลุ่มตัวอย่าง และกลุ่มทดสอบเสียงคำพูด
2. การเตรียมข้อมูลกลุ่มตัวอย่าง และกลุ่มทดสอบเพื่อใช้ในการตัดสินใจ
3. การทดสอบความถูกต้องในภาคตัดสินใจ
4. การทดสอบความถูกต้องในภาคตัดสินใจกับโปรแกรมใช้งานจริง
5. การทดสอบความถูกต้องในการตัดพยางค์ออกจากประโยค

4.2 การเก็บข้อมูลกลุ่มตัวอย่างและกลุ่มทดสอบเสียงพูด

การเก็บข้อมูลตัวอย่าง และ กลุ่มทดสอบเสียงคำพูดจะบันทึกเสียงคำพูดด้วยความถี่ 16,000 Hz และความละเอียด 8 บิท ต่อแชมปลิ้ง การบันทึกเสียงทำในห้องที่ไม่มีเสียงรบกวน การเก็บข้อมูลกลุ่มตัวอย่างและกลุ่มทดสอบเสียงคำพูดจะเป็นเสียงคำพูดของผู้วิจัยคนเดียว โดยจะบันทึกเสียงคำพูดเป็นชุด 7 คำ คือ “หนึ่ง”, “สอง”, “สาม”, “สี่”, “ห้า”, “ON”, “OFF” ทั้งนี้ลดผลกระทบจากเสียงคำพูดที่เปลี่ยนไปของผู้วิจัยเองและแบ่งออกเป็น 2 กลุ่ม คำละ 70 ชุด แต่ละกลุ่มประกอบด้วยเสียงคำพูด 490 คำ โดยใช้ชื่อ DATA_BASE และ DATA_TEST เพื่อใช้เป็นกลุ่มตัวอย่างและกลุ่มทดสอบ ซึ่งมีไว้สำหรับในการทดสอบภาคตัดสินใจเสียงพูด

4.3 การเตรียมข้อมูลกลุ่มตัวอย่าง และ กลุ่มทดสอบเพื่อใช้ในการตัดสินใจ

ในการเตรียมข้อมูลที่ใช้ในการตัดสินใจ เราจะนำเสียงพูดที่ถูกบันทึกไว้จำนวน 20 เสียง มาเป็นตัวแทนในการหาค่าพารามิเตอร์ 2 ชนิด คือ อัตราการตัดศูนย์ และ จำนวนยอดของการตัดเสียง เพื่อให้ได้ค่าพารามิเตอร์ที่เหมาะสม เพื่อที่จะใช้ในการแยกประเภทกลุ่มของคำใน 7 คำ เป็นกลุ่มๆ ดังการหาอัตราการตัดศูนย์ เป็นพารามิเตอร์ในการแบ่งกลุ่มคำ 7 คำ ออกเป็น 2 กลุ่ม คือ กลุ่มของเสียงเสียดสีฟัน และ กลุ่มไม่เสียดสีฟัน

เอกสารนี้สงวนลิขสิทธิ์ ห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้ ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1 การหาจำนวนยอดของเสียงเป็นพารามิเตอร์ในการหากลุ่มคำที่ถูกแบ่งมาจากการหาอัตราการตัดศูนย์ ออกอีกเป็น 2 กลุ่ม คือ กลุ่ม สระเสียงสั้น (Shot vowel) และ กลุ่มสระเสียงยาว (Long vowel)ทั้งนี้ ในการแบ่งกลุ่มแยกกลุ่มเป็นกลุ่มย่อยๆ เพื่อให้จำนวนเสียงของฐานข้อมูลที่จะนำเสียงไปทดสอบมีจำนวนเสียงที่น้อยที่สุดที่จะเป็นไปได้ เพื่อให้ได้การเปรียบเทียบมีค่าความถูกต้องสูงกว่าการนำเสียงไปเปรียบเทียบกับเสียงทั้งหมด

เมื่อกำหนดอัตราการตัดศูนย์ และจำนวนยอดของเสียงแล้ว เราก็จะสรุปได้ว่า ในกลุ่มคำ 7 คำเป็นกลุ่มคำประเภทใด

4.4 การวิเคราะห์สัญญาณเสียงในช่วงเวลาสั้นๆ

เนื่องจากว่าเสียงเป็นสัญญาณที่เปลี่ยนตามเวลา มีการเปลี่ยนแปลงตลอดเวลา เช่น ในขณะพูดช้าๆรูปร่างโพรงเสียงรวมทั้งรูปแบบการกระตุ้นอาจจะไม่มีการเปลี่ยนแปลงนานที่สุดประมาณ 200 มิลลิวินาที ดังนั้นเทคนิคในการวิเคราะห์ส่วนใหญ่แล้ว จะสมมติให้สัญญาณเสียงมีคุณสมบัติที่ไม่เปลี่ยนแปลงสัมพันธ์กับเวลาที่เชื่องช้า นั่นคือเราจะทำการวิเคราะห์หาพารามิเตอร์ของสัญญาณเสียงพูดในช่วงเวลาสั้นๆ เหมือนผ่านช่องแคบๆ ที่เรียกว่า วินโดว์ (WINDOW) เมื่อเทียบกับการเปลี่ยนแปลงเวลา เพื่อจะได้มองเห็นกับว่าเราหาพารามิเตอร์นั้นๆ ได้มาจากสัญญาณเสียงที่มีอยู่ในช่องแคบๆ และมีความเสถียรภายในช่วงเวลานั้น

เทคนิคส่วนใหญ่จะกำหนดให้พารามิเตอร์ได้มีจากค่าเฉลี่ยจากการวิเคราะห์ภายในช่วงเวลาแคบๆ สำหรับในกรณีที่ต้องพิจารณาพารามิเตอร์ต่างๆที่มีการเปลี่ยนแปลงก็จะทำการแบ่งสัญญาณออกเป็นหลายๆช่อง หรือเรียกว่า กรอบการวิเคราะห์ (ANALYSIS FRAME) ดังนั้นพารามิเตอร์ต่างๆ จะสามารถวิเคราะห์ได้ทันเพียงพอที่จะติดตามการเปลี่ยนแปลงของสัญญาณ สำหรับช่วงสัญญาณที่มีการเปลี่ยนแปลงช้าๆ อาจกำหนดให้มีช่องแคบๆมีขนาดใหญ่ประมาณ200 มิลลิวินาที แต่ในทางตรงกันข้ามที่มีการเปลี่ยนแปลงเร็ว ก็ต้องใช้ช่องแคบที่มีขนาดเล็กเพื่อป้องกันการสูญเสียของรายละเอียดของสัญญาณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5 รูปแบบช่องแคบ (WINDOW)

การกำหนดขนาดช่องแคบ

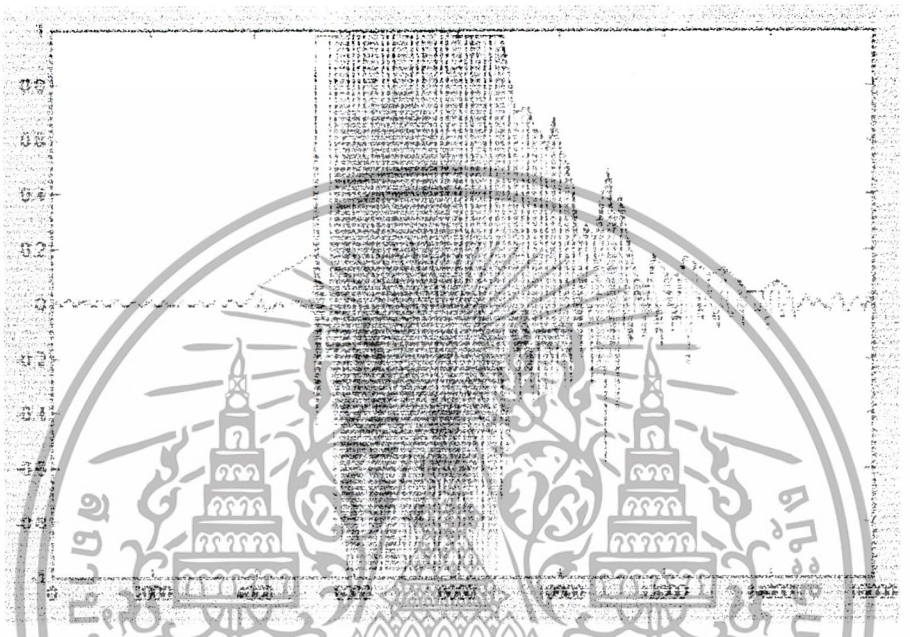
1. ช่องจะต้องสั้นพอควร ที่จะทำให้คุณสมบัติของเสียงที่กำลังพิจารณาไม่มีการเปลี่ยนแปลง
2. ช่องแคบจะต้องมีความยาวพอ ที่จะทำให้การจัดเตรียมตัวอย่างของเสียงที่จะนำไปคำนวณหาพารามิเตอร์ให้ได้ตามต้องการ
3. การทดสอบความถูกต้องในภาคตัดสินใจ

ในการทดสอบความถูกต้องในการตัดสินใจ ในขั้นแรกเราได้กำหนดจำนวนชุดเสียงของฐานข้อมูล (DATA_BASE) 10 ชุด ทำการเปรียบเทียบกับชุดทดสอบ (DATA_TEST) จำนวน 10 ชุด ต่อจากนั้นเราจะทำการเพิ่มชุด ทดสอบเป็น 20,30,40...70 ชุดตามลำดับแต่ยังคงจำนวนชุดของเสียงฐานข้อมูลเป็น 10 ชุด หลังจากนั้นเราก็จะทำการเพิ่มชุดเสียงของฐานข้อมูลเป็น 20,30,40,...,70 ชุด ตามลำดับ

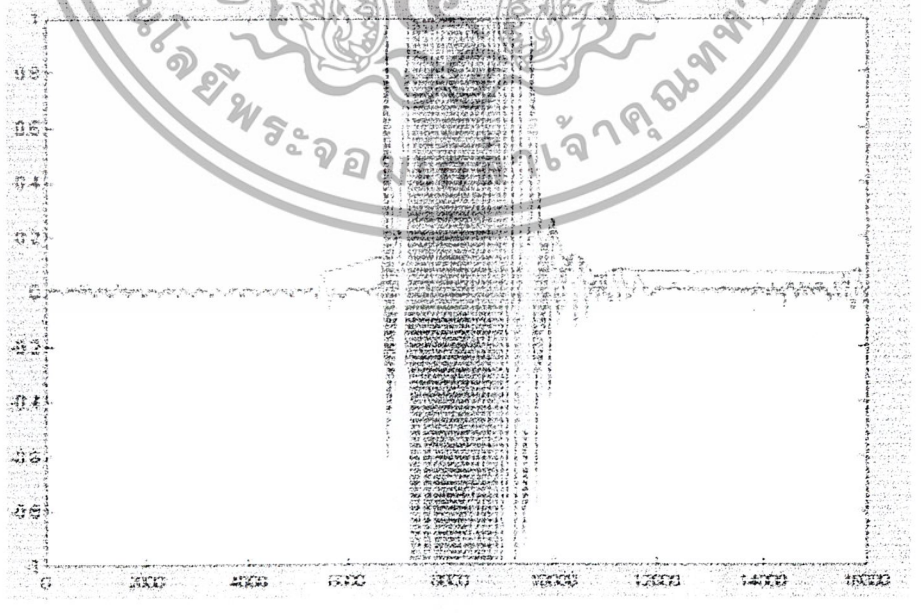
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
ผลการทดลอง

5.1 รูปแสดงลักษณะของเสียงและการหาจุดเริ่มต้นและจุดสุดท้ายของสัญญาณเสียง

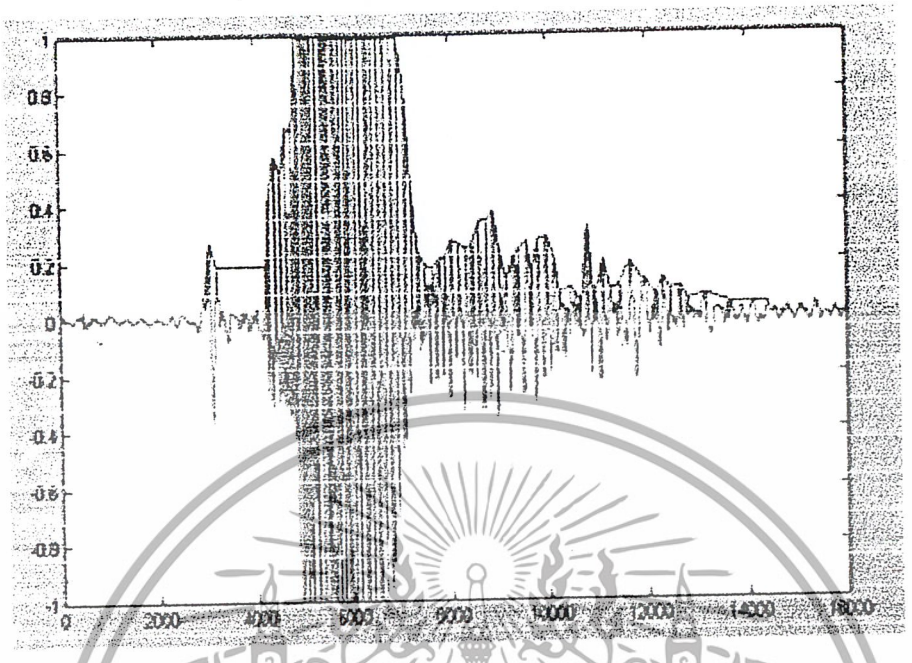


รูปที่ 5.1 รูปแสดง สัญญาณเสียง " หนึ่ง "

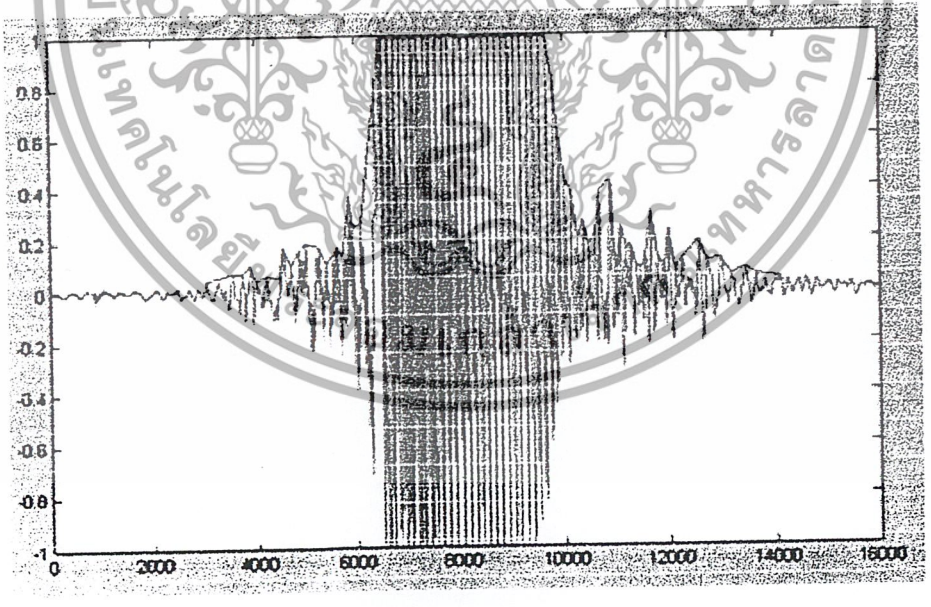


รูปที่ 5.2 รูปแสดง สัญญาณเสียง " สอง "

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

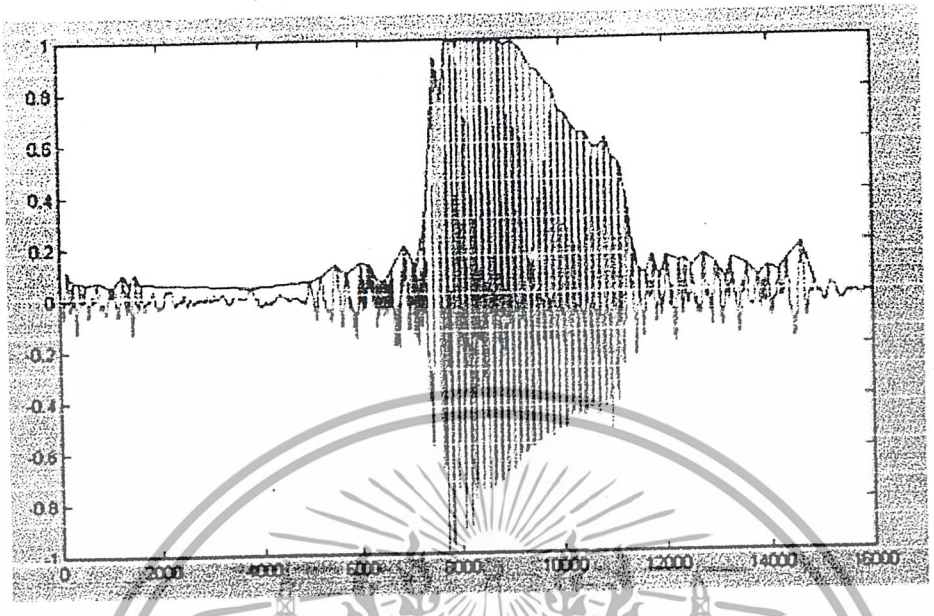


รูปที่ 5.3 รูปแสดง สัญญาณเสียง " สาม "

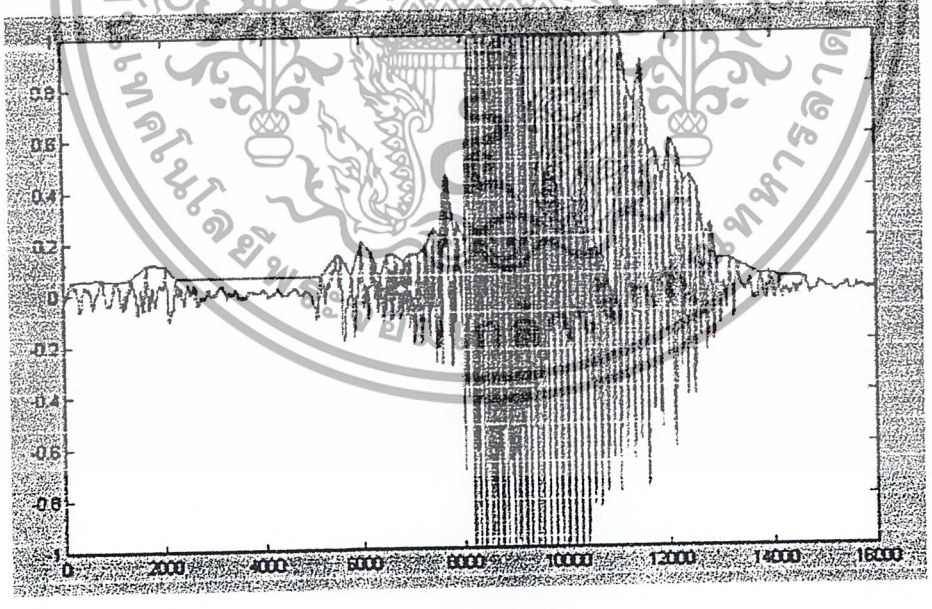


รูปที่ 5.4 รูปแสดง สัญญาณเสียง " สี่ "

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

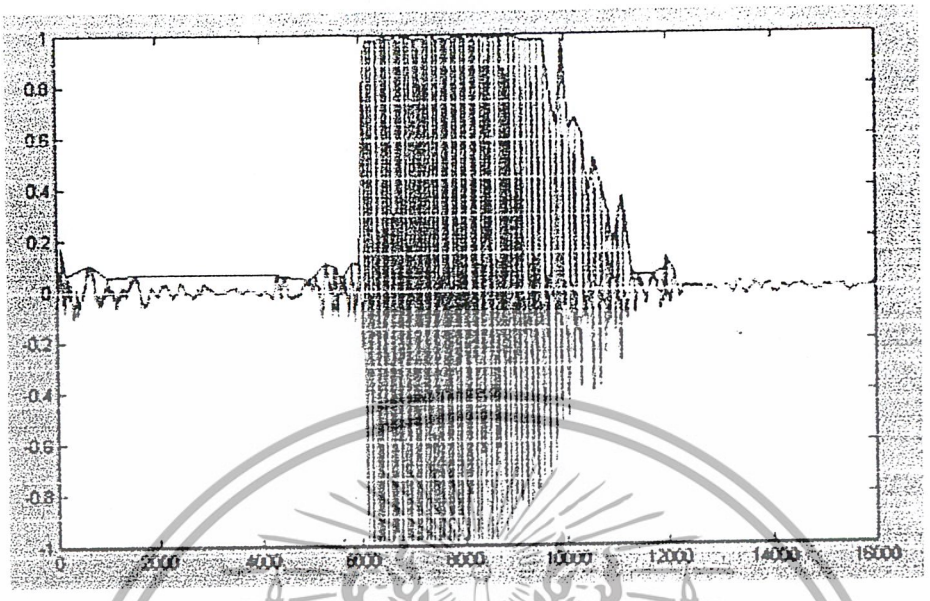


รูปที่ 5.5 รูปแสดง สัญญาณเสียง " ห้า "



รูปที่ 5.6 รูปแสดง สัญญาณเสียง " ON "

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.7 รูปแสดง สัญญาณเสียง " OFF "

การหาอัตราการตัดศูนย์ของเสียง 1,2,3,4,5,ON,OFF

ลำดับที่	เสียงหนึ่ง	เสียงสอง	เสียงสาม	เสียงสี่	เสียงห้า	เสียงON	เสียงOFF
1	5.2287	27.1354	73.436	77.4504	7.5866	12.4369	15.2314
2	5.398	24.1583	50.5826	22.3542	11.7619	11.3691	13.2624
3	5.1686	32.6603	11.2287	29.6161	6.1532	10.87	12.7495
4	3.97	82.2845	56.6371	14.094	5.1377	11.9902	12.9765
5	3.8743	31.4552	46.351	70.462	11.4886	12.8806	13.2465
6	3.5337	43.9395	20.2816	60.6075	6.537	11.4384	12.6547
7	3.9947	34.56	84.4037	49.421	7.0141	11.8244	14.1856
8	3.8812	28.1439	45.9798	29.8654	5.5319	8.7136	14.0155
9	4.9516	34.7986	53.3989	38.6934	8.2173	12.0134	14.1357
10	4.5476	46.2597	25.9625	79.4155	9.0544	10.4539	13.8501
11	5.5024	58.7991	14.8435	56.6545	5.5488	10.4549	12.3681
12	5.0049	69.8614	73.7753	56.116	8.1216	12.0327	13.3195

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

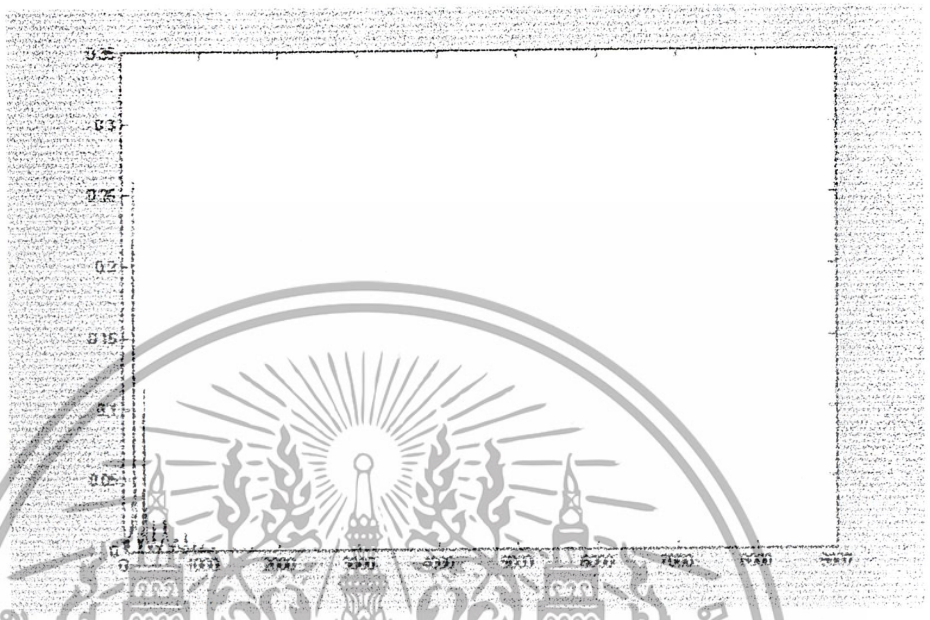
ลำดับที่	เลียงหนึ่ง	เลียงสอง	เลียงสาม	เลียงสี่	เลียงห้า	เลียงON	เลียงOFF
13	4.6302	12.5118	22.5574	65.4752	9.3113	9.8241	12.5674
14	5.4552	29.9216	34.5817	55.351	12.51	12.802	13.5052
15	4.9933	32.7075	40.6668	84.4573	8.5874	10.247	16.5239
16	6.2292	35.9506	4.9755	63.623	7.1806	12.8823	15.0411
17	4.3172	19.6896	30.8083	57.0474	11.2686	8.8459	12.86789
18	5.3216	54.8008	32.7524	67.4235	7.6669	9.6619	15.2764
19	3.6389	35.0787	29.6552	74.2647	4.7578	11.7686	14.2456
20	3.9622	40.8448	28.253	61.8274	7.7389	11.5463	17.0042
21	2.843	46.6836	63.7782	49.6972	6.5708	12.191	13.9973
22	2.7019	58.0045	17.5035	72.2991	402471	12.2577	15.0014
23	7.4411	14.5736	49.4537	93.3292	5.1409	11.9719	16.4193
24	4.9597	33.6226	8.5689	74.9002	7.9483	12.8433	12.6404
25	4.8184	81.366	23.185	89.6347	14.2753	11.5294	15.7484
26	5.4423	55.4022	30.9549	59.3745	8.777	11.3136	14.5146
27	3.667	56.7954	17.7899	39.5601	9.0599	11.4389	15.7494
28	4.8056	39.7154	43.0402	64.4092	14.431	11.0892	14.6097
29	5.3809	24.3071	51.8447	79.1315	5.5654	10.3657	15.6204
30	5.5775	54.9964	7.3432	57.7863	5.8534	10.3652	14.6331
31	4.3923	60.89	15.775	62.7271	11.5257	11.2513	13.0787
32	5.5098	42.9247	27.0917	64.2681	6.3252	11.3462	16.2556
33	4.453	18.2861	5.6444	86.3664	5.083	10.2546	18.7663
34	4.6667	28.5746	53.9352	65.308	6.2209	10.3657	14.3801
35	4.4935	56.4056	11.1872	54.6902	7.0849	10.8244	15.4457
36	2.7527	9.9798	43.4554	66.7481	7.2678	13.0763	13.4207
37	6.4172	38.3301	25.9459	35.6357	7.094	9.0121	15.6582
38	4.4049	83.2983	21.7785	82.4343	9.2228	9.7746	14.1464
39	4.735	40.1789	21.6912	83.5166	3.5583	11.6959	14.3386
40	6.7646	34.9571	25.9677	58.6447	8.6909	13.3262	13.619
41	4.9458	62.0868	26.1984	61.3523	7.8037	8.92	13.6894

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

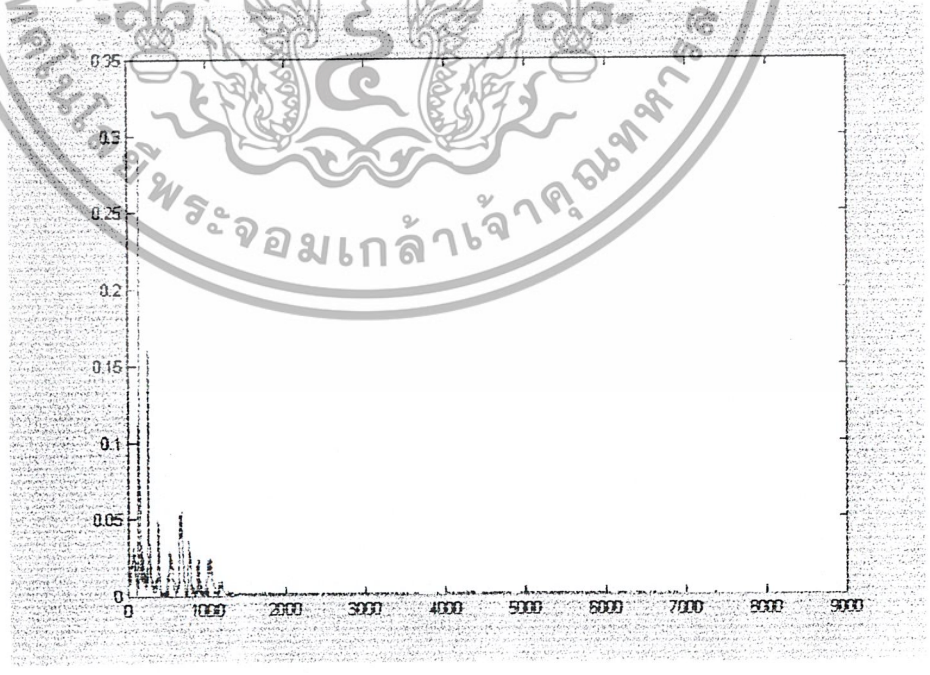
ลำดับที่	เสียงหนึ่ง	เสียงสอง	เสียงสาม	เสียงสี่	เสียงห้า	เสียงON	เสียงOFF
42	3.4218	49.5235	23.7786	52.9022	17.5207	12.1647	15.1343
43	4.8839	47.3056	18.2641	51.4719	6.0434	12.4466	14.5938
44	6.4895	43.4042	8.9302	69.2527	4.5853	9.5191	13.4114
45	4.0698	40.7871	46.7737	58.3982	6.9683	11.2972	16.3389
46	3.7256	44.2696	53.9977	71.7039	8.8357	12.017	12.8523
47	5.8549	50.5877	32.7783	92.4114	10.7279	12.8382	11.9204
48	3.5637	57.8464	61.395	89.6285	10.9449	13.6532	14.2953
49	4.771	45.9979	12.0591	53.0422	6.4671	11.2543	15.3306
50	4.9948	82.3237	42.1436	47.5058	8.0463	9.7995	12.4309
51	3.6735	80.478	69.9762	40.3353	4.663		
52	2.7684	27.9976	40.4007	25.7696	4.1077		
53	2.866	64.5263	34.1139	44.4783	7.9926		
54	3.9782	52.7934	23.6851	53.056	9.5802		
55	2.9249	68.9219	25.4737	53.9538	7.4433		
56	2.885	43.3675	35.1626	47.907	12.9049		
57	4.7098	58.2746	22.5088	50.7565	5.1376		
58	2.7809	37.2535	23.8921	66.196	8.8217		
59	2.7679	43.9567	34.108	79.2646	9.1219		
60	4.2477	61.2491	38.3858	18.9201	8.8208		
61	3.0968	25.4006	10.8637	57.7388	9.1382		
62	2.8236	24.6136	18.9721	70.9995	10.0551		
63	4.8102	40.9197	32.1944	15.3315	12.1522		
64	3.4381	47.6654	7.2401	66.545	6.953		
65	4.0167	35.8638	52.9311	45.1945	6.7597		
66	3.4175	26.8697	23.9647	39.0544	8.0624		
67	3.4723	37.5314	10.7381	24.6469	7.9633		
68	2.7293	49.0473	9.0734	65.3774	10.6476		
69	2.79686	49.0473	26.0485	26.019	5.1538		
70	3.6835	30.4052	42.1415	37.261	6.1322		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 การหาสเปกตรัมของสัญญาณเสียงต่าง ๆ

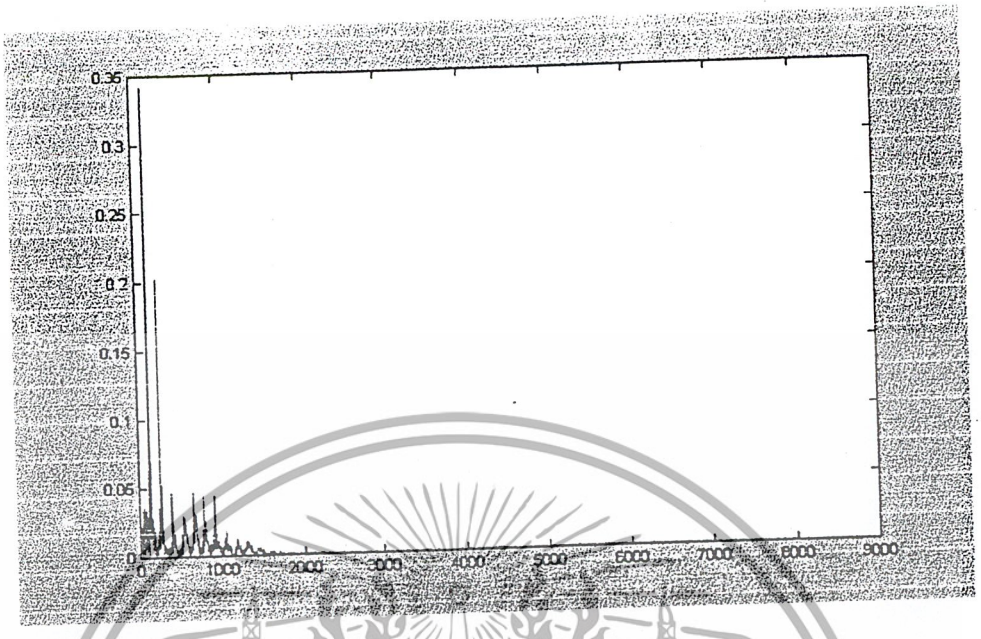


รูปที่ 5.8 สเปกตรัมของเสียงหนึ่ง



รูปที่ 5.9 สเปกตรัมของเสียงลอง

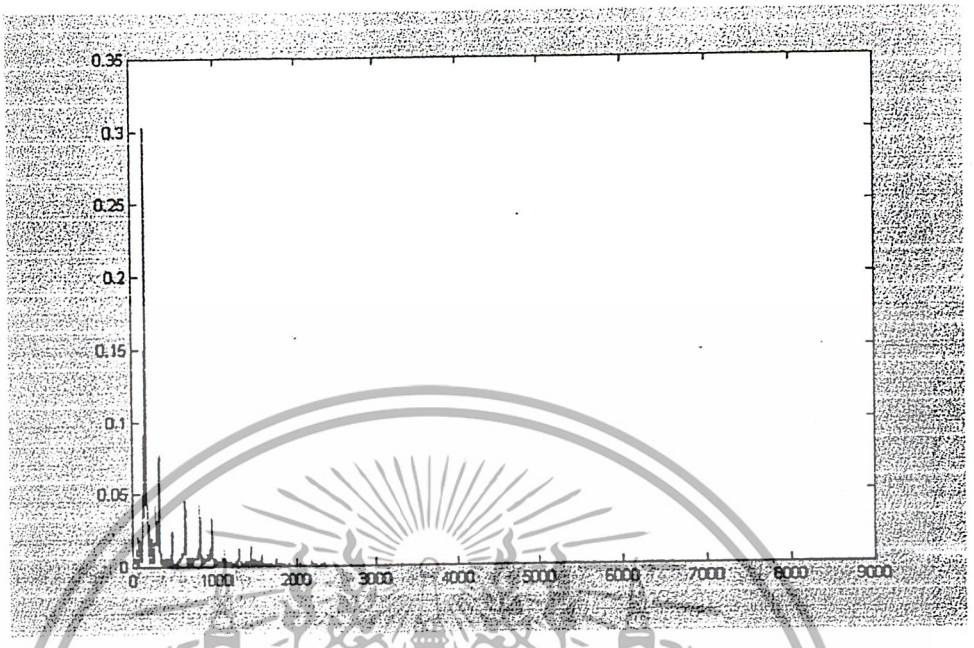
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



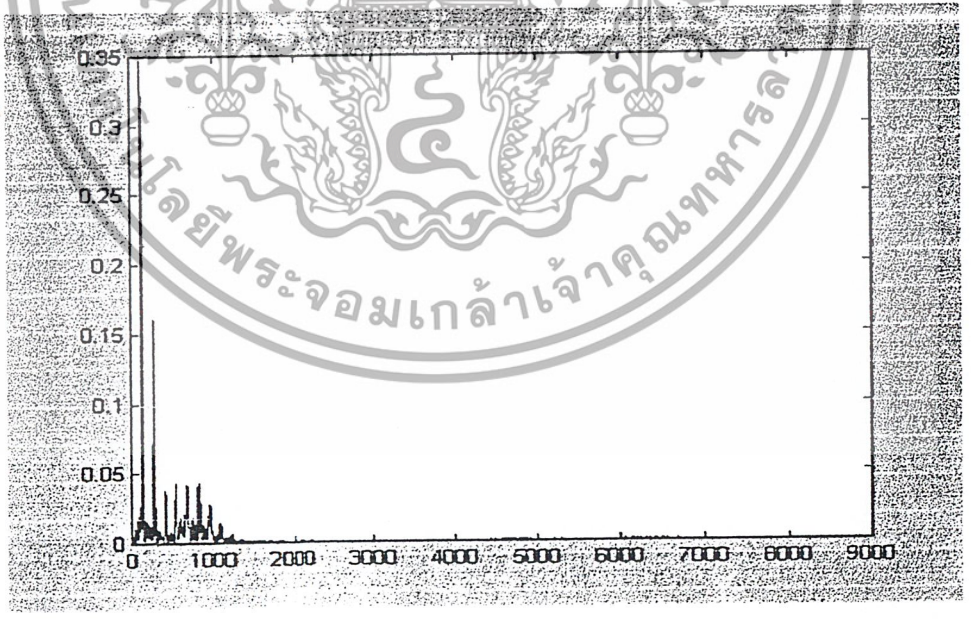
รูปที่ 5.10 สเปกตรัมของเสียงสาม

รูปที่ 5.11 สเปกตรัมของเสียงสี่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

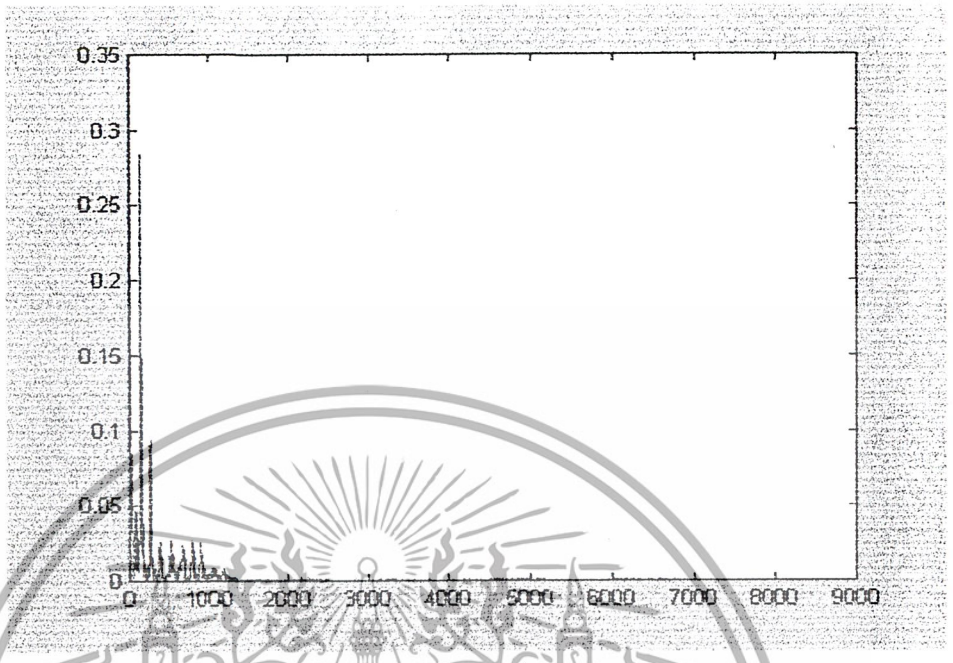


รูปที่ 5.12 สเปกตรัมของเสียงห้า



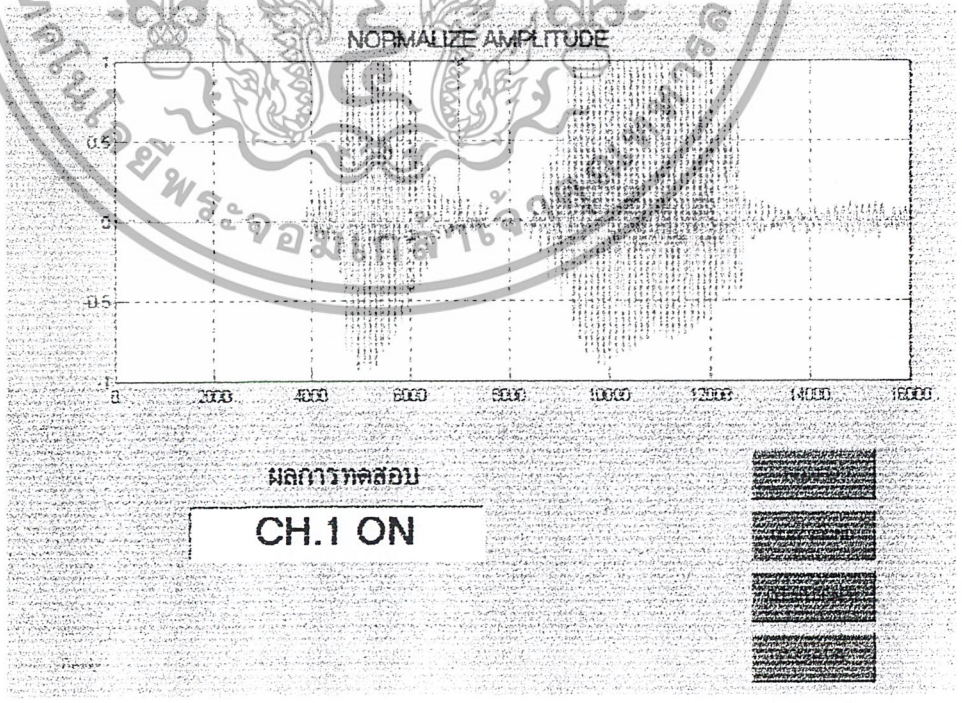
รูปที่ 5.13 สเปกตรัมของเสียง ON

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



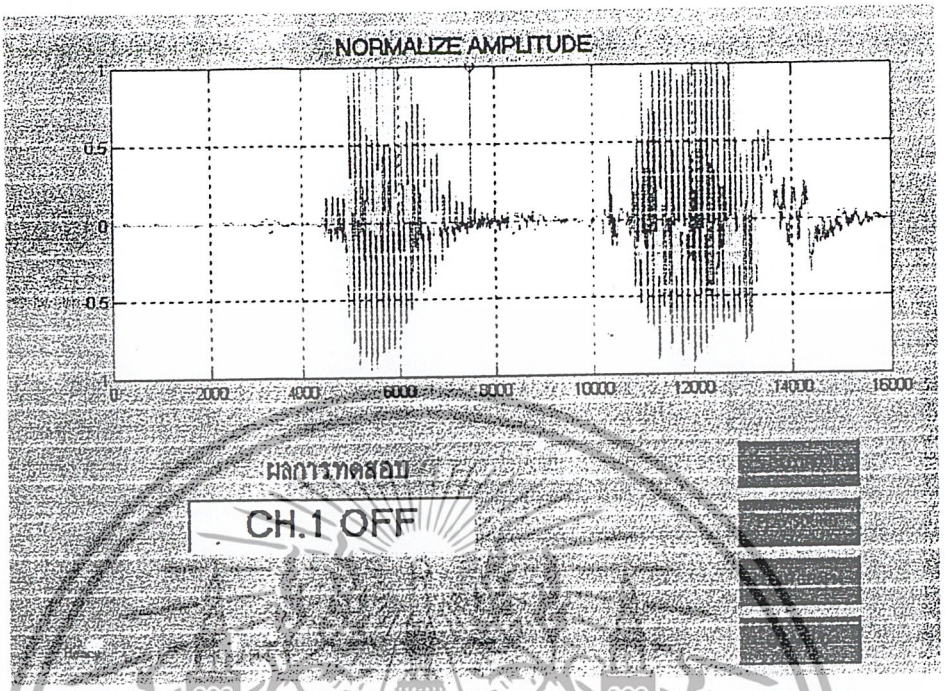
รูปที่ 5.14 สเปกตรัมของเสียง OFF

5.3 รูปแสดงผลการทดลองที่แสดงผลทางจอ Computer

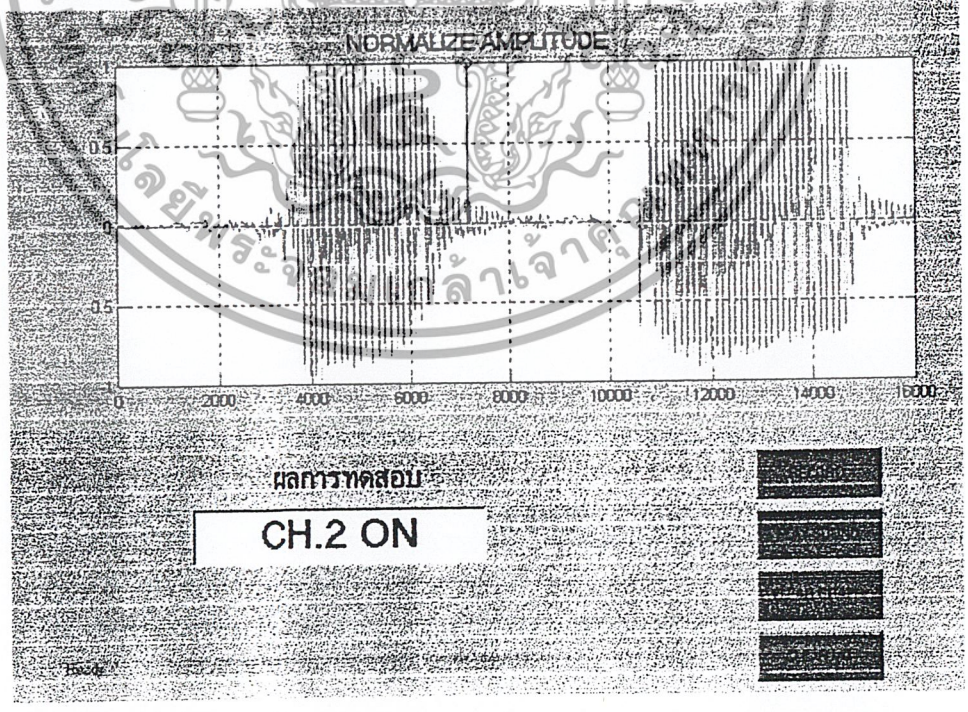


รูปที่ 5.15 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.1 ON

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

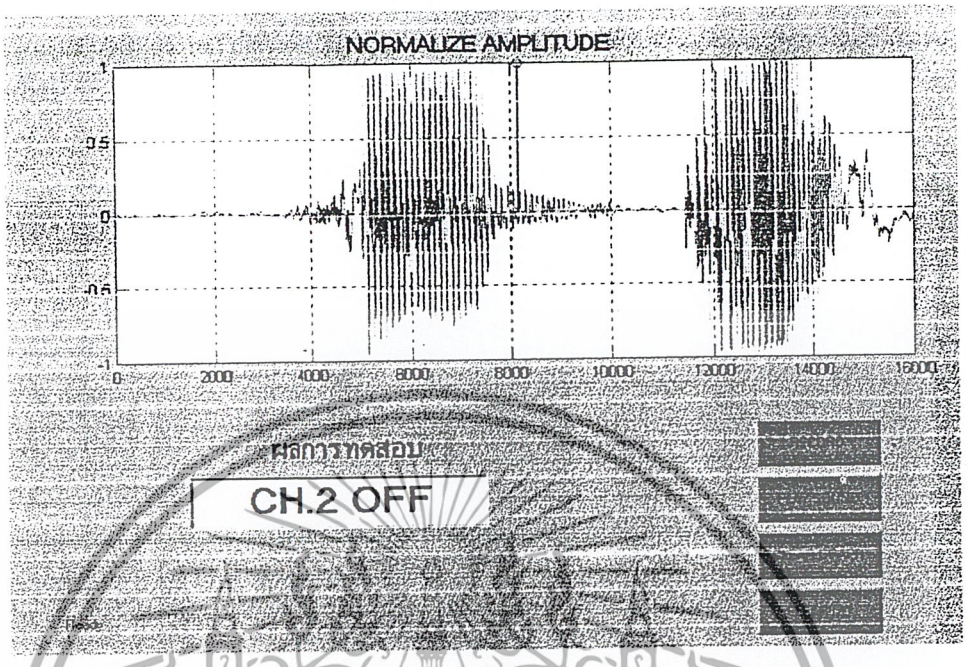


รูปที่ 5.16 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.1 OFF

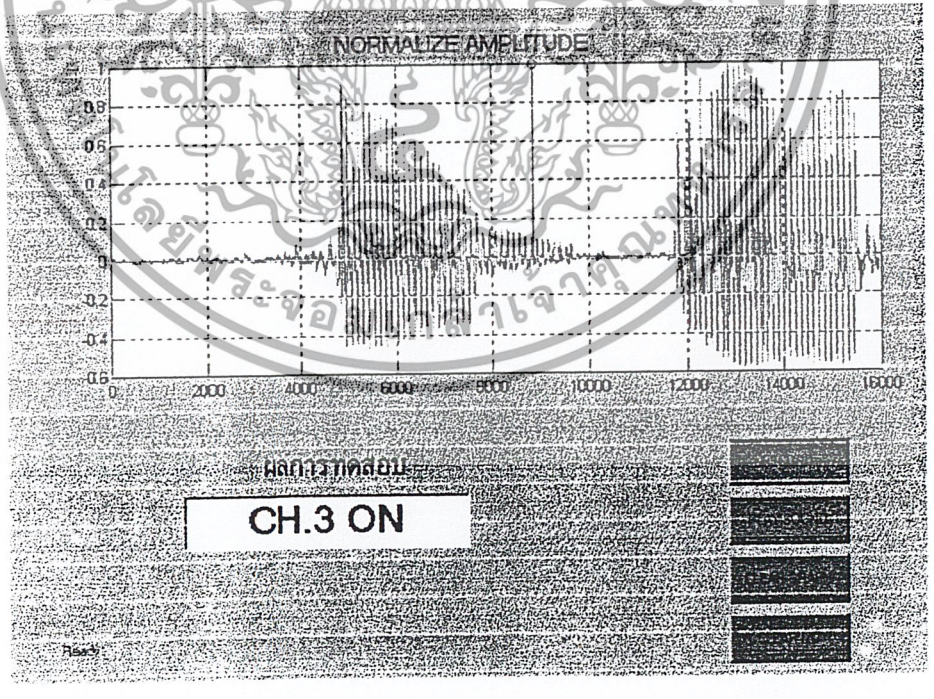


รูปที่ 5.17 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.2 ON

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

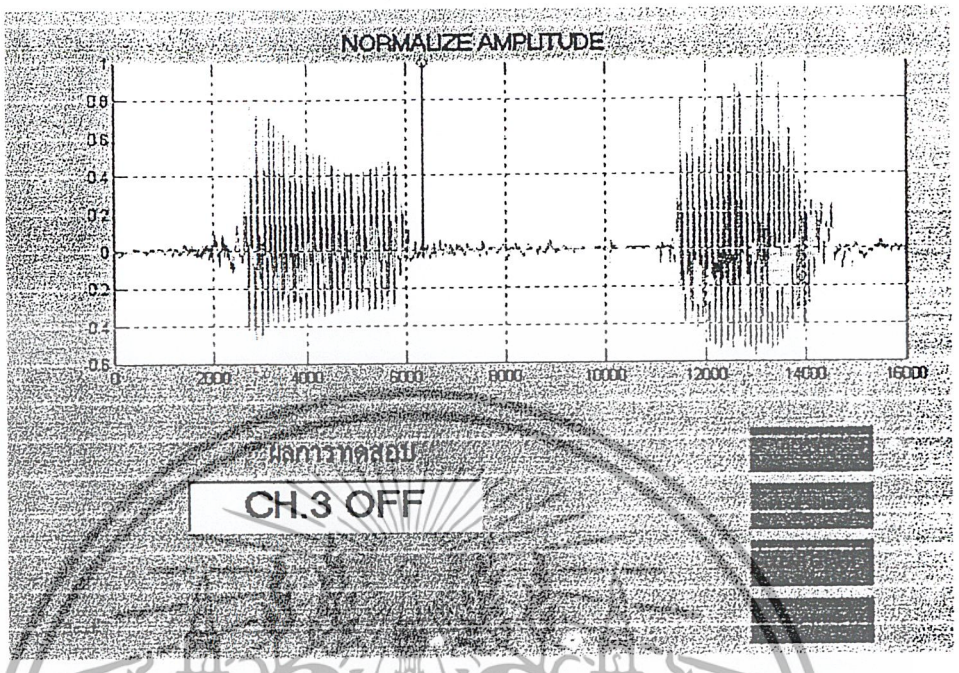


รูปที่ 5.18 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.2 OFF

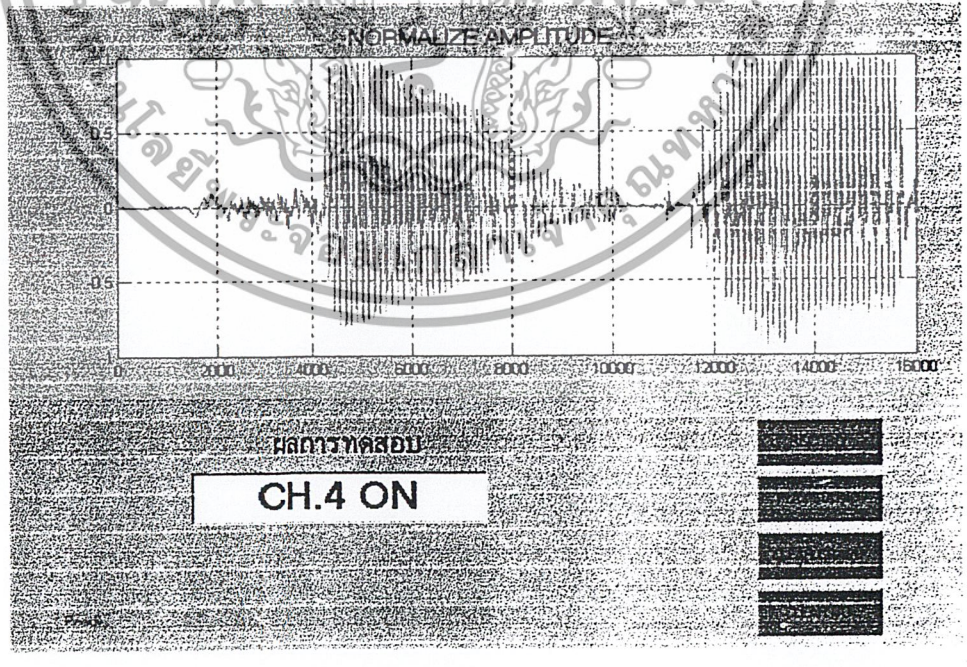


รูปที่ 5.19 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.3 ON

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

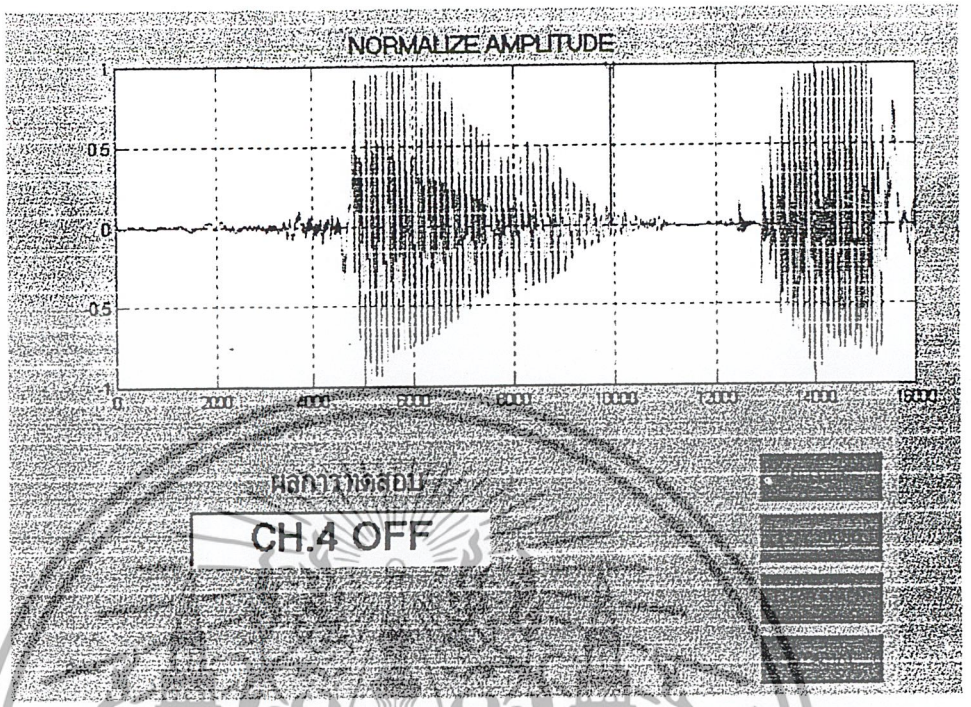


รูปที่ 5.20 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.3 OFF

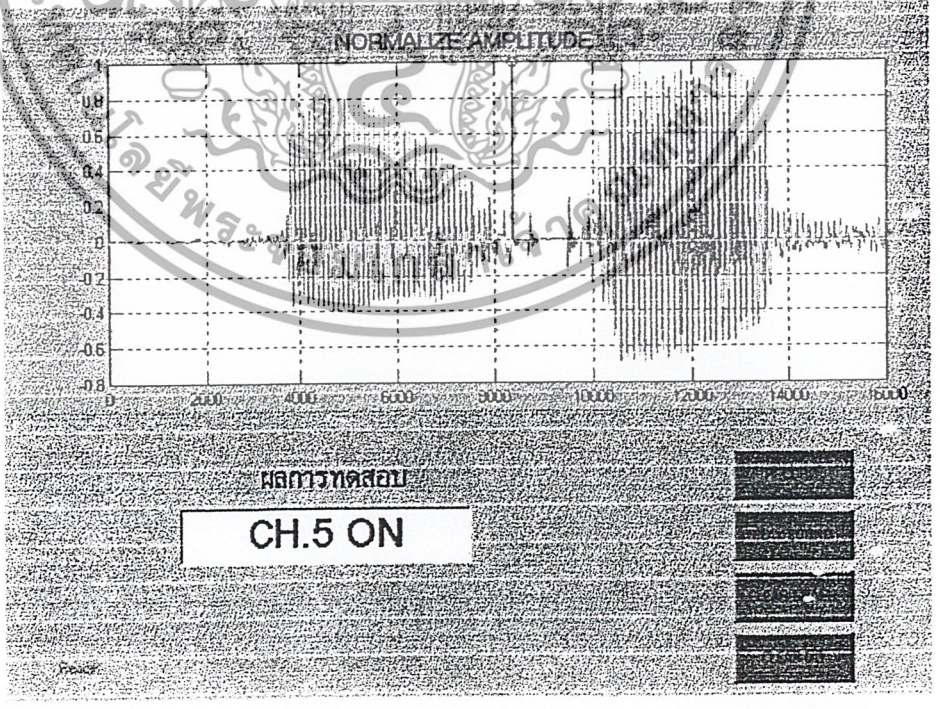


รูปที่ 5.21 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.4 ON

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

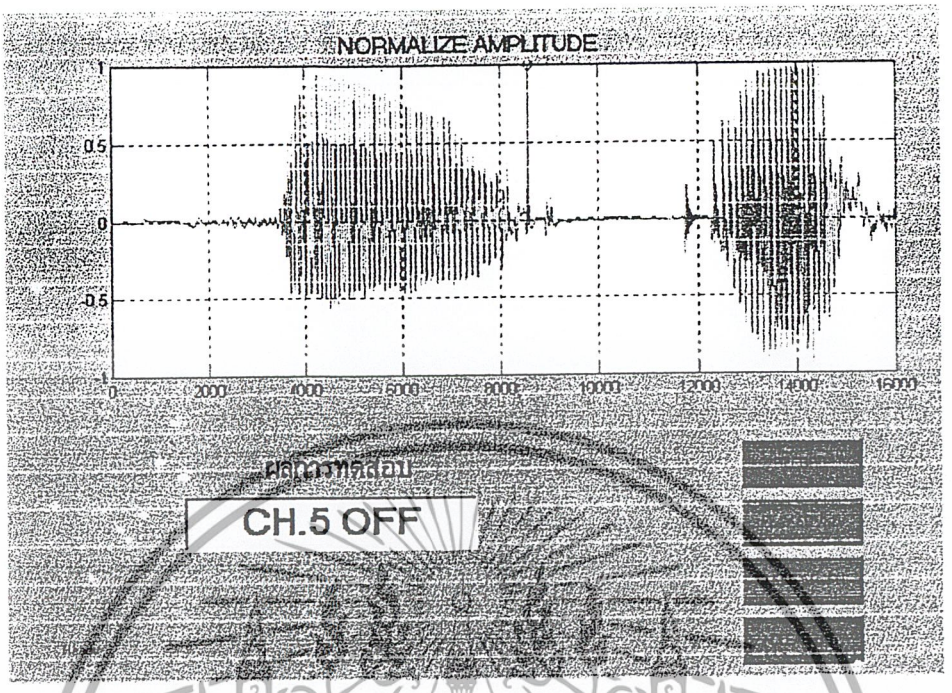


รูปที่ 5.22 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.4 OFF



รูปที่ 5.23 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.5 ON

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.24 แสดงการทดสอบการทำงานที่ใช้งานจริง CH.5 OFF

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัญหาในการทำงาน

1. Computer ที่ใช้ต้องมีความเร็วสูง เนื่องจากต้องใช้เวลาในการประมวลผลมาก ถ้าใช้เครื่องที่มีความเร็วต่ำจะทำให้ช้าต่อการทำงานมาก

2. Sound card ต้องใช้ที่มีประสิทธิภาพที่ดีในการที่จะเก็บเสียงที่พูด ถ้าใช้ Sound Card ที่ไม่ดีจะทำให้ความถูกต้องน้อยลงไป

3. ไมโครโฟน จะต้องใช้ไมโครโฟนที่มี Sensitivity สูง เพื่อนำมาบันทึกเสียงที่สามารถที่จะรับความละเอียดของลักษณะของสัญญาณเสียง ถ้าใช้ไมโครโฟนที่มี Sensitivity ต่ำ จะทำให้ลักษณะของสัญญาณเสียงไม่ดีการคำนวณก็จะผิดพลาด

4. ในการบันทึกเสียง จะต้องไม่มีเสียงรบกวน ถ้ามีเสียงรบกวนจะทำให้เกิดสัญญาณรบกวน (Noise) เช่น เสียงลมพัดก็มีผลกับสัญญาณเสียง เมื่อนำมาคำนวณก็จะเกิดการผิดพลาด

5. ในส่วนของโปรแกรม MATLAB ซึ่งมีความสามารถในการเก็บข้อมูลของเสียงได้ 1วินาที ทำให้ไม่สามารถบันทึกคำพูดที่ต่อเนื่องได้

6. ในการบันทึกเสียงแต่ละครั้งจะมีความแตกต่างกันทำให้เกิดความไม่แน่นอนของการจดจำของคำพูด ซึ่งจะทำให้มีความผิดพลาดได้

ข้อเสนอแนะ

1) ในการทดลองจะต้องทำการทดลองในสถานที่ที่เงียบ ไม่ควรจะมีเสียงอื่นมารบกวน เพราะอาจทำให้เกิดความผิดพลาดได้
2) การเพิ่มความถูกต้องให้กับภาคการจดจำเสียงนั้น สามารถทำได้โดยการเพิ่มจำนวนชุดเสียงของฐานข้อมูล ให้มากขึ้นซึ่งจะทำให้ความถูกต้องมีเพิ่มมากขึ้น
3) สำหรับภาคการรู้จำเสียงคำพูดในงานวิจัยนี้นั้น ผู้วิจัยได้ทำการคิดพารามิเตอร์ในการทดลองการจดจำเสียงขึ้นมาเอง ไม่ได้ทำการจดจำด้วยวิธีอื่นเช่น การสอนนิวตรอนเนคเวอร์คเป็นต้น ซึ่งอาจจะไม่ใช่วิธีที่ดีที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฟังก์ชันการทำงานของโปรแกรม MATLAB

```
function [n]=call_test_on_off_w1(yw1,ip1,nf)
                % THIS PROGRAM ONLY USE FOR VIOCE
SIGNALS %
              % LOAD DATA BASE

              lp1=length(ip1);
              lz1=round(0.2*lp1);
              z=zcross(yw1(ip1(1):ip1(lz1)));
              wf=find_fft(yw1,ip1);

  if(z<30)
              for w=1:nf

                    St1=['load
d:\data_base\chanel\wave' int2str(w) '\f1;'];
                    eval(St1);
                    St2=['d(w,1)=mean(abs(wf-
f1));'];
                    eval(St2);
                    clear St1;
                    clear St2;
                    St1=['load
d:\data_base\chanel\wave' int2str(w) '\f5;'];
                    eval(St1);
                    St2=['d(w,2)=mean(abs(wf-
f5));'];
                    eval(St2);

                                % CLEAR VARIABLE%

                    clear f1;
                    clear f5;
                    clear St1;
                    clear St2;


                end
          % FIND THE MIN VALUE %
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                    mm=min(d);
                    [m n]=min(mm);
                    if(n==2)
                       n=5;
                    end
  else

          for w=1:nf
           for uv=2:4
                    St1=['load
d:\data_base\chanel\wave' int2str(w) '\f' int2str
(uv) ';'];
                    eval(St1);
                    St2=['d(w,uv-1)=mean(abs(wf-f'
int2str(uv) '));'];
                    eval(St2);

       % CLEAR VARIABLE%

                    St3=['clear f' int2str(uv)
';'];
                    eval(St3);
                    clear St1;
                    clear St2;
                    clear St3;

             end
          end

     % FIND THE MIN VALUE %
                    mm=min(d);
                    [m n]=min(mm);
                    n=n+1;

  end
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
function [n]=call_test_on_off_w2(wf,nf)
                % THIS PROGRAM ONLY USE FOR VIOCE
SIGNALS %
                % LOAD DATA BASE


              for w=1:nf
                  for k=1:2
                      %St1=['load
d:\data_base\chanel\wave' int2str(w) '\f' int2str
(k) ';'];
                      %eval(St1);
                      St1=['load
d:\data_base\on_off\wave' int2str(w) '\f' int2str
(k) ';'];
                     eval(St1);



                      St2=['d(w,k)=mean(abs(wf-f'
int2str(k) '));'];
                      eval(St2);
                              % CLEAR VARIABLE%
                      St3=['clear f' int2str(k) ';'];
                      eval(St3);
                      clear St1;
                      clear St2;
                      clear St3;

                  end
              end
                        % FIND THE MIN VALUE %
                       mm=min(d);
                       [m n]=min(mm);
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
function    find_formant(a)
t1=a;
      fs=22050;
      % NUMBER OF DATA TEST %
       for w=1:5
        for k=1:5

                        st1=['[y,fs]=wavread
(''d:\voice\wave' int2str(w) '\w' int2str(k)
''');'];
                        eval(st1);
                         % FIND START AND END POINT %
                        [sx,ip]=start_end(y);
                        lp=length(ip);
                        zx=sx(ip(1):ip(lp));
                        f=abs(fft(zx,16384));
                        ff1=f(1:8192);
                        [m,i]=max(ff1);
                        ff2=ff1/m;
                        st2=['f' int2str(k) '=ff2;'];
                        eval(st2);
                        st3=['save d:\database\wave'
int2str(w) '\f' int2str(k) ' f' int2str(k) ';'];
                        eval(st3);

                        clear y;
                        clear sx;
                        clear zx;
                        clear f;
                        clear ff1;
                        clear ff2;
                        clear lp;
                        clear ip;
                        clear m;
                        clear st1;
                        clear st2;
                        clear st3;
                        st4=['clear f' int2str(k) ';'];
                        eval(st4);
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
function fig = ON_OFF_GUI()
%My job.
load GUI
global chanel
chanel=0;
h0 = figure('Color',[0.8 0.8 0.8], ...
   'Colormap',mat0, ...
   'FileName','C:\computer_project\GUI.m', ...
   'PaperPosition',[18 180 576 432], ...
   'PaperUnits','points', ...
   'Position',[1 1 800 562], ...
   'Tag','Fig1', ...
    'ToolBar','none');
global graph

h1 = axes('Parent',h0, ...
   'Units','pixels', ...
   'CameraUpVector',[0 1 0], ...
   'Color',[1 1 1], ...
   'ColorOrder',mat1, ...
   'Position',[93 256 670 260], ...
   'Tag','Axes1', ...
   'XColor',[0 1 1], ...
   'XGrid','on', ...
   'YColor',[0 1 1], ...
   'YGrid','on', ...
    'ZColor',[0 0 0]);
graph=h1;

h2 = text('Parent',h1, ...
   'Color',[0 0 1], ...
   'HandleVisibility','off', ...
   'HorizontalAlignment','center', ...
   'Position',[0.4992526158445441 -
0.0926640926640927 9.160254037844386], ...
   'Tag','Axes1Text4', ...
   'VerticalAlignment','cap');
set(get(h2,'Parent'),'XLabel',h2);
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
h2 = text('Parent',h1, ...
   'Color',[0 1 1], ...
   'HandleVisibility','off', ...
   'HorizontalAlignment','center', ...
   'Position',[-0.043348281016442445
0.4942084942084941 9.160254037844386], ...
   'Rotation',90, ...
   'Tag','Axes1Text3', ...
   'VerticalAlignment','baseline');
set(get(h2,'Parent'),'YLabel',h2);

h2 = text('Parent',h1, ...
   'Color',[0 0 0], ...
   'HandleVisibility','off', ...
   'HorizontalAlignment','right', ...
   'Position',mat2, ...
   'Tag','Axes1Text2', ...
   'Visible','off');
set(get(h2,'Parent'),'ZLabel',h2);

h2 = text('Parent',h1, ...
   'Color',[0 0 0], ...
   'HandleVisibility','off', ...
   'HorizontalAlignment','center', ...
   'Position',[0.4992526158445441 1.0270270270270270
9.160254037844386], ...
   'Tag','Axes1Text1', ...
   'VerticalAlignment','bottom');
set(get(h2,'Parent'),'Title',h2);
global display status

h1 = uicontrol('Parent',h0, ...
   'Units','points', ...
   'BackgroundColor',[1 1 1], ...
   'Callback','A', ...
   'FontSize',24, ...
    'FontWeight','demi', ...
   'ListboxTop',0, ...
    'Position',[115 82.5 189.75 33.75], ...
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
   'String','', ...
   'Style','edit', ...
    'Tag','EditText1');
     display=h1;


h1 = uicontrol('Parent',h0, ...
   'Units','points', ...
    'BackgroundColor',[0.7804 0.7804 0.7804], ...
    'ForegroundColor',[0 0 0.6725], ...
   'ListboxTop',0, ...
    'Position',[150 124.75 130 20.5], ...
    'FontSize',19, ...
   'String','[illegible]', ...
   'Style','text', ...
    'Tag','StaticText2');


h1 = uicontrol('Parent',h0, ...
   'Units','points', ...
    'BackgroundColor',[0.7804 0.7804 0.7804], ...
    'ForegroundColor',[1 0 0], ...
   'ListboxTop',0, ...
   'Position',[29.25 16.25 34.5 11], ...
   'String','Ready', ...
   'Style','text', ...
    'Tag','StaticText2');
status=h1;

global error_status
h1 = uicontrol('Parent',h0, ...
   'Units','points', ...
    'BackgroundColor',[0.7804 0.7804 0.7804], ...
    'ForegroundColor',[1 0 0], ...
   'ListboxTop',0, ...
   'Position',[100 16.25 230 11], ...
   'String','', ...
   'Style','text', ...
    'Tag','StaticText2');
error_status=h1;
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
global enable_play enable_clear enable_record
enable_clear_io

h1 = uicontrol('Parent',h0, ...
   'Units','points', ...
   'BackgroundColor',[0.250980392156863
0.501960784313725 0.501960784313725], ...
   'Callback','record', ...
   'ListboxTop',0, ...
   'Position',[471.75 119.25 77.25 30.25], ...
   'String','RECORD', ...
    'Tag','Pushbutton1');
    enable_record=h1;

h1 = uicontrol('Parent',h0, ...
   'Units','points', ...
   'BackgroundColor',[0.250980392156863
0.501960784313725 0.501960784313725], ...
   'Callback','play_sound', ...
   'ListboxTop',0, ...
   'Position',[471.75 82.25 77.25 30.25], ...
   'String','PLAY SOUND', ...
    'Tag','Pushbutton3');
    enable_play=h1;

h1 = uicontrol('Parent',h0, ...
   'Units','points', ...
   'BackgroundColor',[0.250980392156863
0.501960784313725 0.501960784313725], ...
   'Callback','clear_graph', ...
   'ListboxTop',0, ...
   'Position',[471.75  45.25 77.25 30.25], ...
   'String','CLEAR GRAPH', ...
    'Tag','Pushbutton4');
    enable_clear=h1;

h1 = uicontrol('Parent',h0, ...
   'Units','points', ...
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
   'BackgroundColor',[0.250980392156863
0.501960784313725 0.501960784313725], ...
   'Callback','clear_io', ...
   'ListboxTop',0, ...
   'Position',[471.75  8.25 77.25 30.25], ...
   'String','CLEAR I/O', ...
    'Tag','Pushbutton4');
    enable_clear_io=h1;


h1 = uicontrol('Parent',h0, ...
   'Units','points', ...
    'BackgroundColor',[0.7804 0.7804 0.7804], ...
    'ForegroundColor',[0 0 0.6725], ...
   'ListboxTop',0, ...
    'Position',[204 388.5 193.5 19.5], ...
    'FontSize',15, ...
   'String','NORMALIZE AMPLITUDE', ...
   'Style','text', ...
    'Tag','StaticText2');


if nargout > 0, fig = h0; end
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
lobal status graph display enable_clear enable_play
;
global  enable_record error_status enable_clear_io;


set(status,'String','Busy');
set(display,'String',' ');
set(error_status,'String',' ');
set(enable_clear,'Enable','off');
set(enable_play,'Enable','off');
set(enable_record,'Enable','off');
set(enable_clear_io,'Enable','off');

clear_graph;

fs=16000;
y=recordsnd(1,fs);
ly=1:1:length(y);
t=ly/fs;
length_y=length(y);
x=y;
[mx nx]=max(abs(x));
x=x/mx;


[end_w1,enrun,le]=cut_word_on_off(x);

if((end_w1<length_y)&(mx>0.1)&(enrun==1))

w1=y(1:end_w1);
w2=y((end_w1+1):length_y);

plot(x,'r');
hold on;
stem(end_w1,1);
hold off;

set(graph,'XColor',[0 0 1]);
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
set(graph,'XGrid','on');
set(graph,'YColor',[0 0 1]);
set(graph,'YGrid','on');
set(graph,'ZColor',[0 0 0]);
save d:\data_play\y y;


[yw1 ip1]=start_end(w1);
[yw2 ip2]=start_end(w2);
lp1=length(ip1);
lz1=round(0.2*lp1);

num_database1=40;
sound1=call_test_on_off_w1(yw1,ip1,num_database1);

freq2=find_fft(yw2,ip2);
num_database2=30;
sound2=call_test_on_off_w2(freq2,num_database2);

if((sound1==1)&(sound2==1))
   set(display,'String','CH.1 ON');
    Outport(1,1);

elseif((sound1==1)&(sound2==2))
       set(display,'String','CH.1 OFF');
        Outport(1,2);

elseif((sound1==2)&(sound2==1))
      set(display,'String','CH.2 ON');
      Outport(2,3);

elseif((sound1==2)&(sound2==2))
     set(display,'String','CH.2 OFF');
     Outport(2,4);

elseif((sound1==3)&(sound2==1))
       set(display,'String','CH.3 ON');
       Outport(3,5);
elseif((sound1==3)&(sound2==2))
       set(display,'String','CH.3 OFF');
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        Outport(3,6);

elseif((sound1==4)&(sound2==1))
        set(display,'String','CH.4 ON');
        Outport(4,7);
elseif((sound1==4)&(sound2==2))
        set(display,'String','CH.4 OFF');
        Outport(4,8);

elseif((sound1==5)&(sound2==1))
        set(display,'String','CH.5 ON');
        Outport(5,9);
elseif((sound1==5)&(sound2==2))
        set(display,'String','CH.5 OFF');
        Outport(5,10);

end


else
   set(error_status,'String','Error Message :
Uncomplete word,please record again ');

end

set(enable_clear,'Enable','on');
set(enable_play,'Enable','on');
set(enable_record,'Enable','on');
set(enable_clear_io,'Enable','on');
set(status,'String','Ready');
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
function [yyx,ip2]=start_end(yy)

  yyx=yy;
          % NORMALIZE THE WAVE
   mmx=max(abs(yyx));
  if mmx~=0
     yyx=yyx/mmx;
       end
       %nwp=180;
       nwp=100;
   L=length(yyx);
   %Leng=L
     step=1;
      ok=0;
       h=1;
    % FIND THE START AND END OF WAVE
while ok==0
     if ((step+nwp)<=L)
         y1=yyx((step:step+nwp));
         [mx,ix]=max(y1);
            if ix==(nwp/2+1)
               n=step+nwp/2;
                 if yyx(n)>=0.04
                    ip(h)=n;
                  m(h)=yyx(n);
                  h=h+1;
                end
           end
            step=step+1;
     else
      ok=1;
     end
     clear y1;
     clear n;
     clear ix;
 end
 %s=step
 %index=ip
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
%maxni=m
    % DETECT THE REDUNDANT LENGTH PITCH
    lep=length(ip);
    hf=round(lep/2);
    start=1;
    last=lep;
 for kk=1:hf
   if(kk<(hf-1))
     if (ip(kk+1)-ip(kk))>500
        start=kk+1;
     end
   end
 end

for kl=lep:-1:(hf+2)
        if (ip(kl)-ip(kl-1))>500
          last=kl-1;
        end
  end

        ip2=ip(start:last);

        % plot(yyx,'r')
        % hold on
       % plot(ip,m);
        %hold off
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
function [standard]=zcross(x)

    lx=length(x);
    N=260;
        % N=220;
              % FIND ZERO CROSSING
    for n=1:lx
        if x(n)>=0
           sgn(n)=1;
        else
           sgn(n)=-1;
        end
    end
           w=ones(N,1);
           for k=2:lx
              sgnx(k)=abs(sgn(k)-sgn(k-1));
           end
                % FIND CROSS POINT
                 z=conv(sgnx,w);
                %plot(z);
                 standard=std(z);
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้