

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

เครื่องคิดเลขใช้เสียงโดยวิธี RECOGNITION แบบขึ้นกับผู้พูด
SPEECH CONTROLLED CALCULATOR



โดย

นายชูเกียรติ	รักไทย	เลขประจำตัว 40013167
นายนิวัฒน์	ศรีแก้ว	เลขประจำตัว 40013175
นายนรินทร์	สายเสมา	เลขประจำตัว 40013216

อาจารย์ที่ปรึกษา

อาจารย์ สุรเดช ศรีไตรลักษณ์

ปริญญานิพนธ์นี้สำหรับปริญญาวិทยกรรมศาสตรบัณฑิต

สาขาวิชาอิเล็กทรอนิกส์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เลขที่.....

เลขทะเบียน 36961

วัน, เดือน, ปี 29 ส.ค. 2548

ปีการศึกษา 2542

สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่มีการรับผิดชอบต่อการใช้งานให้คัดลอกเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์ ปีการศึกษา 2542

ภาควิชา วิศวกรรมอิเล็กทรอนิกส์

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง เครื่องคิดเลขใช้เสียงโดยวิธี RECOGNITION แบบขึ้นกับบุคคล

ผู้จัดทำ

- 1) นายชูเกียรติ รักไทย
- 2) นายนิวัฒน์ ศรีแก้ว
- 3) นายนรินทร์ สายเสมา

..... 4 อาจารย์ที่ปรึกษา
(..... (อ.สุรเดช ศรีไพโรจน์))

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เครื่องคิดเลขใช้เสียงโดยวิธี RECOGNITION แบบขึ้นกับผู้พูด
SPEECH CONTROLLED CALCULATOR

นายชูเกียรติ รักไทย เลขประจำตัว 40013167

นายนิวัฒน์ ศรีแก้ว เลขประจำตัว 40013175

นายบุรินทร์ สายเสมา เลขประจำตัว 40013216

โครงการได้รับการตรวจสอบแล้ว พร้อมทั้งจะทำการสอบได้

(.....)
(.....)
(.....)

อาจารย์ที่ปรึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เครื่องคิดเลขใช้เสียงโดยวิธี RECOGNITION แบบขึ้นกับผู้พูด

นายชูเกียรติ รักไทย

นายนิวัฒน์ ศรีแก้ว

นายบุรินทร์ สายเสมา

อาจารย์สุรเดช ศรีไตรลักษณ์ อาจารย์ที่ปรึกษา

ปีการศึกษา 2542

บทคัดย่อ

ปริญญานิพนธ์นี้นำเสนอ หลักการเบื้องต้น และการพัฒนาเครื่องคิดเลขที่ควบคุมด้วยเสียงพูดแบบขึ้นกับผู้พูด ในกลุ่มคำ 16 คำ ดังนี้ “ศูนย์”, “หนึ่ง”, “สอง”, “สาม”, “สี่”, “ห้า”, “หก”, “เจ็ด”, “แปด”, “เก้า”, “สิบ”, “บวก”, “ลบ”, “คูณ”, “หาร”, และ “ชี้” เครื่องคิดเลขจะไม่ทำงานตามเวลาจริง ตัวโปรแกรมที่ใช้เขียนด้วยโปรแกรม Matlab เสียงพูดเก็บในไฟล์รูปแบบเวฟผ่านทางไมโครโฟนที่ต่ออยู่กับการ์ดเสียง ด้วยความถี่แซมปลิง 22050 Hz ความละเอียด 8 บิตต่อแซมปลิง

ส่วนรู้จำเสียงพูด ใช้วิธีการอัตราการตัดศูนย์และวิธีการหาจำนวนพิทช์ของสัญญาณเสียงเป็นพารามิเตอร์แทนลักษณะเฉพาะของเสียงพูด โดยวิธีการอัตราการตัดศูนย์จะทำการแบ่งลักษณะประเภทของคำออกเป็น แบบไม่มีเสียงเสียดสีฟัน (VOICE) และ แบบมีเสียงเสียดสีฟัน (UNVOICE) ส่วนวิธีการหาจำนวนพิทช์จะเป็นการแบ่งลักษณะของคำให้เป็นกลุ่มเล็กลงโดยแบ่งเป็น สระเสียงสั้น และ สระเสียงยาว และนำมาหาสเปกตรัมและทำการประมวลผลเพื่อหาผลลัพธ์ที่ถูกต้อง จากผลการทดสอบสามารถตัดสินใจถูกต้อง คิดเป็นร้อยละ 93.75%

SPEECH CONTROLLED CALCULATOR

Chukrait Rakthai

Niwat Srikraw

Burin Saisama

Suradaj Tritiraksana Project Report Advisor

Abstract

This research presents the speech controlled calculator. The speech recognition system is written in Matlab program. Speech is stored in wave format file by using microphone connected with sound card having sampling frequency of 22050 Hz and 8 bit per sample.

The speech recognition using zero crossing rate and length pitch as parameter of speech. This parameter is used as data for to determine word. The result show that find spectrum has an accuracy of 93.75 %

กิตติกรรมประกาศ

ผู้วิจัยขอขอบพระคุณ อาจารย์สุรเดช ศรีไตรลักษณ์ ที่ได้รับเป็นอาจารย์ที่ปรึกษา กรุณาให้คำแนะนำ ตลอดจนความช่วยเหลือในด้านต่าง ๆ

ขอขอบพระคุณอาจารย์ทุกท่าน ที่ได้เสียสละเวลาอันมีค่าในการอ่าน และสอบปรินญา
นิพนธ์

ขอขอบคุณ นายธม ทะการ ที่ให้ความช่วยเหลือด้วยดีมาตลอดและขอขอบคุณสำหรับผู้ที่
มอบรอยยิ้ม ซึ่งเป็นกำลังใจ ที่นำไปสู่ความสำเร็จ

ท้ายสุดนี้ ขอกราบขอบพระคุณ บิดา มารดา ที่มอบความอบอุ่น สนับสนุน และเป็นแรงใจ
จนกระทั่งปรินญานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี

นายชูเกียรติ รักไทย
นายนิวัฒน์ ศรีแก้ว
นายบุรินทร์ สายเสมา

สารบัญ

หน้า

บทคัดย่อภาษาไทย	1
บทคัดย่อภาษาอังกฤษ	2
กิตติกรรมประกาศ	3
บทที่ 1 บทนำ	8
ความสำคัญของปัญหา	8
วัตถุประสงค์	9
ขั้นตอนการวิจัย	9
ประโยชน์ที่คาดว่าจะได้รับ	10
บทที่ 2 หลักการและทฤษฎี	11
ทฤษฎีการสร้างเสียงพูด	11
ทฤษฎีการวิเคราะห์เสียงพูด	13
บทที่ 3 โครงสร้างงานวิจัย	23
ภาคอ่านไฟล์เสียง	23
ภาครู้จำเสียงพูด	24
ภาคประมวลผล	28
ภาคแสดงผลการคำนวณด้วยเสียงพูด	28

บทที่ 4	วิธีการดำเนินการวิจัย	29
	การเก็บข้อมูลกลุ่มตัวอย่าง และกลุ่มทดสอบเชิงคำพูด	29
	การเตรียมข้อมูลกลุ่มตัวอย่าง และกลุ่มทดสอบเพื่อใช้ในการตัดสินใจ	29
	การทดสอบความถูกต้องในภาคตัดสินใจ	30
	การทดสอบความถูกต้องกับโปรแกรมจริง	30
	การทดสอบการแยกพยางค์ออกจากประโยค	47
บทที่ 5	สรุปผลและอภิปรายผล	50
	สรุปผล	50
	อภิปรายผลความผิดพลาด	51
	ข้อเสนอแนะ	51
	การประยุกต์ใช้งาน	51
	เอกสารอ้างอิง	52
	ภาคผนวก	53

สารบัญรูป

หน้า

รูปที่ 2.1	อวัยวะภายในของระบบการพูดของมนุษย์	11
รูปที่ 2.2	บล็อกไดอะแกรมของระบบการสร้างเสียงพูด	12
รูปที่ 2.3	แบบจำลองอย่างง่ายของการกำหนดช่องทางเดินเสียง	13
รูปที่ 2.4	บล็อกไดอะแกรมของอัตราการตัดศูนย์เฉลี่ยในช่วงสั้น	16
รูปที่ 2.5	การแจกแจงของอัตราการตัดศูนย์สำหรับ แบบ ไม่มีเสียงเสียดสีฟัน (VOICE) และ แบบมีเสียงเสียดสีฟัน (UNVOICE)	17
รูปที่ 2.6	อัตราการเปลี่ยนแปลงเสียงพูดที่แตกต่างกัน 3 ครั้ง	18
รูปที่ 2.7	บล็อกไดอะแกรมของการถอดสัญญาณพีชแบบขนาน	20
รูปที่ 2.8	อินพุตสัญญาณขาขึ้นและการเกิดขบวนอิมพัลส์จาก PEAK และ VALEY	21
รูปที่ 3.1	ระบบการบันทึกเสียงพูด	23
รูปที่ 3.2	แสดงการตัดพยางค์ออกจากประโยค	24-26
รูปที่ 3.3	แสดงผลการคำนวณด้วยเสียงพูด	28
รูปที่ 4.1	แสดงการทดสอบความถูกต้องของคำกับโปรแกรมใช้งานจริง	31-46
รูปที่ 4.2	แสดงการตัดพยางค์ออกจากประโยค	47-49

สารบัญตาราง

หน้า

ตารางที่ 1	แสดงการแบ่งเสียงออกเป็นเสียงเสียดสีฟันและเสียงไม่เสียดสีฟัน	58
ตารางที่ 2	แสดงการแบ่งเสียงออกเป็นสระเสียงสั้นและสระเสียงยาว	59
ตารางที่ 3	แสดงผลการทดลองการรู้จำเสียง	60-62
ตารางที่ 4	การหาค่าจำนวนพิทช์เฉลี่ยของเสียงไม่เสียดสีฟัน	63
ตารางที่ 5	การหาค่าจำนวนพิทช์เฉลี่ยของเสียงเสียดสีฟัน	64
ตารางที่ 6	แสดงค่าหาจำนวนพิทช์ จำนวนเสียง 20 ชุด	65-66
ตารางที่ 7	แสดงค่าอัตราการตัดศูนย์ จำนวนเสียง 20 ชุด	67-68
ตารางที่ 8	แสดงค่าความถูกต้องในการประมวลผลผลลัพธ์	69-73
ตารางที่ 9	แสดงค่าความถูกต้องในการแยกพยางค์ออกจากประโยค	74

บทที่ 1

บทนำ

ความสำคัญของปัญหา

ความพยายามของมนุษย์ในการสร้างเครื่องมือเครื่องจักรต่าง ๆ ด้วยเสียงนั้นมีมานานแล้วทั้งนี้เพื่อเพิ่มความสะดวกและรวดเร็วในการติดต่อให้มากขึ้น ในอนาคตการสร้างงานด้วยเสียงคำพูดจะใช้กันอย่างแพร่หลาย คอมพิวเตอร์สามารถเข้าใจภาษามนุษย์ และสามารถพูดตอบโต้กับผู้ใช้ได้ ความต้องการให้เครื่องคอมพิวเตอร์สามารถเรียนรู้ถึง วิธีการในการรู้จำคำพูดได้ ทำให้เกิดศาสตร์แขนงหนึ่งขึ้นเรียกว่า การรู้จำเสียงคำพูด (Speech Recognition)

เราสามารถแบ่งการศึกษาตามลักษณะผู้พูดได้เป็น 3 แบบ คือ

- 1) แบบขึ้นกับผู้พูด (Speaker Dependence)
- 2) แบบหลายผู้พูด (Multi Speaker)
- 3) แบบไม่ขึ้นกับผู้พูด (Speaker independence)

โดยแต่ละแบบยังสามารถแบ่งย่อยการศึกษาตามลักษณะเสียงคำพูดได้อีก 3 แบบ

- 1) แบบเสียงคำพูด โดด (Isolated Speech)
- 2) แบบเสียงคำพูดต่อกัน (Connected Speech)
- 3) แบบเสียงคำพูดต่อเนื่อง (Continuous Speech)

ลักษณะการรู้จำเสียงคำพูดตัวเลขภาษาไทย สามารถแบ่งส่วนสำคัญได้ 2 ส่วน

- 1) ส่วนการหาพารามิเตอร์เสียงคำพูด

เป็นส่วนที่สำคัญของการรู้จำเสียงคำพูด เพราะ ความถูกต้องของการรู้จำเสียงคำพูด ขึ้นอยู่กับการเลือกพารามิเตอร์ ซึ่งใช้แทนลักษณะเฉพาะของเสียงคำพูด เช่น ความถี่ฟอร์แมนท์, สเปกตรัมดิสแตนท์, ค่าเฉลี่ยของระดับพลังงาน และสัมประสิทธิ์การประมาณพันธุเชิงสั้น

- 2) ส่วนตัดสินใจ

เป็นส่วนที่ทำการวิจัยกันมากในงานวิจัยที่ผ่านมา ต่างก็ใช้เทคนิคที่แตกต่างกันในส่วนนี้ เช่น ไดนามิกไทม์วาร์ปิง, เทคนิคการจัดกลุ่ม, ฮิดเดินมาร์คอฟโมเดล และนิวรัลเน็ตเวิร์ค

งานวิจัยนี้ ผู้วิจัยจึงมีแนวคิดที่จะศึกษาวิจัยการรู้จำเสียงคำพูดภาษาไทย และนำมาประยุกต์ใช้กับเครื่องคิดเลข โดยใช้สเปกตรัมของสัญญาณเสียงคำพูดเป็นพารามิเตอร์ ระบบรู้จำเสียงพูดภาษาไทยที่วิจัยนี้ จะเป็นแบบขึ้นกับผู้พูด (Speaker dependent) โดยเสียงคำพูดเป็นแบบคำโดด (Isolated Speech)

วัตถุประสงค์

- 1) ศึกษาการวิเคราะห์หาพารามิเตอร์ ที่เป็นลักษณะเด่นของเสียงคำพูดในกลุ่มคำ 16 คำ ดังนี้ “ศูนย์”, “หนึ่ง”, “สอง”, “สาม”, “สี่”, “ห้า”, “หก”, “เจ็ด”, “แปด”, “เก้า”, “สิบ”, “บวก”, “ลบ”, “คูณ”, “หาร”, และ “ชี้” ตามลำดับ
- 2) ศึกษาวิธีการแยกพยางค์จากประโยคที่พูด
- 3) ศึกษาวิธีการในการรู้จำเสียงคำพูดทั้ง 16 คำ
- 4) ศึกษาวิธีการการประมวลผลเพื่อให้ได้ผลลัพธ์ที่ถูกต้อง

ขอบเขตงานวิจัย

- 1) การวิจัยนี้มุ่งศึกษาพัฒนาบนเครื่องไมโครคอมพิวเตอร์ โดยมีอุปกรณ์เพิ่มเติม ได้แก่ การ์ดเสียง (Sound Card), ไมโครโฟน และลำโพง เป็นต้น
- 2) เสียงคำพูดถูกแปลงเป็นสัญญาณดิจิทัลด้วยความถี่แซมปลิง 22050 Hz และความละเอียด 8 Bits/Sample
- 3) เครื่องคิดเลขเขียนขึ้น โดยโปรแกรม Matlab บนวินโดวส์ และไม่ทำงานตามเวลาจริง (Real Time)
- 4) การวิจัยนี้สามารถคำนวณได้เพียงตัวเลขจำนวนเต็มบวกไม่เกิน 2 หลัก

ขั้นตอนการวิจัย

- 1) ศึกษากระบวนการรู้จำเสียงคำพูดในลักษณะต่าง ๆ ที่ได้มีการศึกษามาแล้ว
 - 2) ศึกษาทฤษฎีการสร้างเสียงพูด และอัลกอริทึมในการวิเคราะห์เสียงพูด
 - 3) กำหนดขอบเขตของงานวิจัย และแนวทางที่จะปฏิบัติ
 - 4) ศึกษาการเขียนโปรแกรมด้วยโปรแกรม Matlab
 - 5) ศึกษาอัลกอริทึมในการแปลงจำนวนในทางคณิตศาสตร์ กลับเป็นเสียงคำพูด
- จำนวนภาษาไทย
- 6) บันทึกเสียงพูดเป็นชุด ชุดหนึ่งมี 16 คำ คือ “ศูนย์”, “หนึ่ง”, “สอง”, “สาม”, “สี่”, “ห้า”, “หก”, “เจ็ด”, “แปด”, “เก้า”, “สิบ”, “บวก”, “ลบ”, “คูณ”, “หาร”, และ “ชี้” ตามลำดับ การบันทึกเสียงคำพูดจะบันทึกวันละ 8 ชุด เป็นเวลา 10 วัน รวม 1280 คำ
 - 7) เตรียมข้อมูล วิเคราะห์ และทดสอบ
 - 8) สรุปผลการวิจัย ประเมินผลในการวิจัย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย

- 1) ทราบถึงพารามิเตอร์ที่ควรใช้เป็นข้อมูล ในการตัดเสียงคำพูดของกลุ่มคำทั้ง 16 คำ
- 2) เป็นแนวทางการพัฒนาไปสู่ระบบรู้จำคำพูดภาษาไทยคำอื่น ๆ ต่อไป
- 3) สามารถนำไปประยุกต์ใช้งานในด้านต่าง ๆ เพื่อเพิ่มประสิทธิภาพการติดต่อกับเครื่อง

มือต่าง โดยใช้เสียงพูด



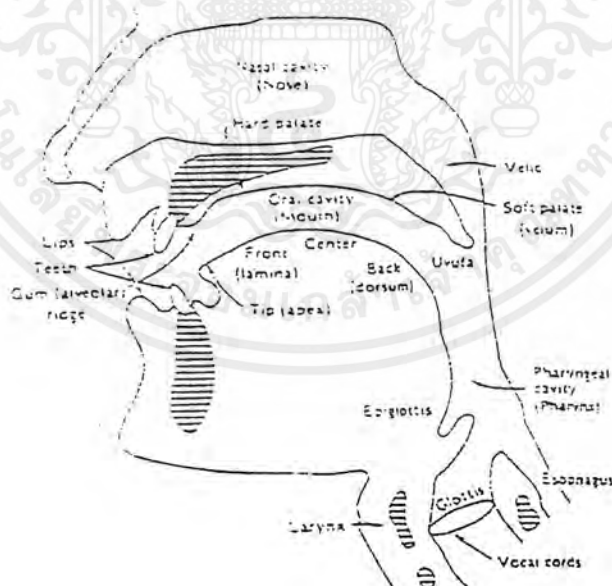
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

หลักการและทฤษฎี

ทฤษฎีการสร้างเสียงพูด

มนุษย์สร้างเสียงพูดขึ้นมาโดยมนุษย์จะมีเส้นเสียง (Vocal Cords) ในลำคอ ที่ทำให้เกิดเสียงซึ่งอยู่ภายในหลอดลม เส้นเสียงมีลักษณะเป็นกล้ามเนื้อ 2 แผ่น ปิดขวางอยู่ปากช่องหลอดลม จากด้านหลังมาด้านหน้า ระหว่างเส้นเสียงจะมีช่องว่างซึ่งเป็นทางให้ลมผ่านเข้าไปปอด และออกจากปอดได้ ช่องนี้เรียกว่า ช่องระหว่างเส้นเสียง (Glottis) เส้นเสียงทั้งสองสามารถจะดึงออกให้ห่างจากกัน หรือดึงเข้ามาประชิดกันก็ได้ เส้นเสียงเป็นส่วนที่ทำให้เกิดเสียงพูดขึ้น และจากการสั่นของเส้นเสียงจะทำให้เกิดคลื่นเสียงความถี่ต่าง ๆ ผ่านเข้าสู่ลำคอ ผ่านลำคอเข้าสู่ปาก หรือช่องจมูกออกไปภายนอก ซึ่งขนาด และรูปร่างของอวัยวะภายในช่องปากนี้ จะเป็นสิ่งกำหนดว่าคลื่นเสียงความถี่ไหนจะยอมให้ปรากฏออกมาให้ได้ยิน หรือคลื่นเสียงความถี่ไหนจะถูกดูดซับไว้ไม่ยอมให้ปรากฏออกมา อวัยวะส่วนที่มีหน้าที่ในการออกเสียงพูด แสดงดังรูปที่ 2.1

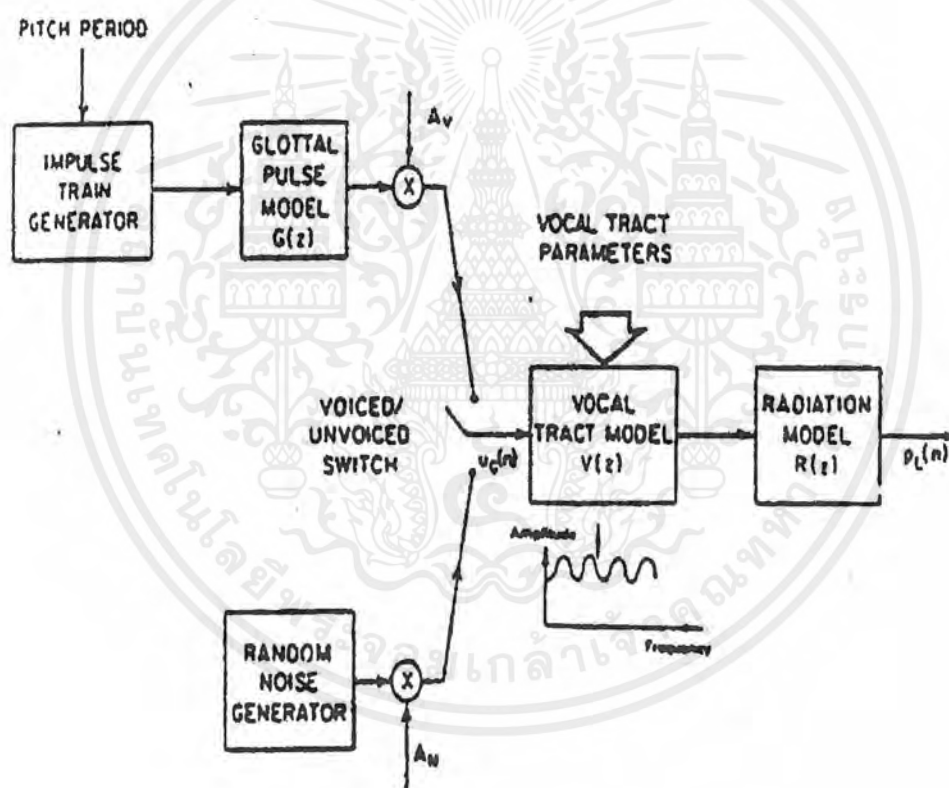


รูปที่ 2.1 อวัยวะภายในของระบบการพูดของมนุษย์

อวัยวะที่ใช้ในการออกเสียงนั้น อาจจะแบ่งคร่าว ๆ ได้เป็น 2 พวก คือ

1) อวัยวะที่เป็นส่วนกระทำอาการ (Articulator) หมายถึง อวัยวะส่วนที่เคลื่อนไหวเพื่อ พนักลมไปยังส่วนต่าง ๆ อวัยวะที่เป็นส่วนกระทำอาการที่สำคัญ คือ ลิ้น ซึ่งเป็นส่วนที่เคลื่อนไหว ได้มากที่สุด

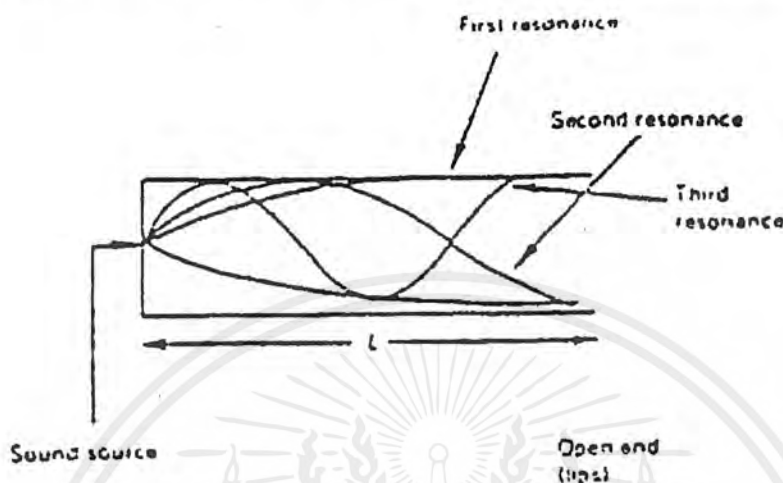
2) อวัยวะที่เป็นฐานที่เกิดเสียงต่าง ๆ (Point of articulator) หมายถึง ตำแหน่ง หรือฐานที่ เกิดเสียงต่าง ๆ เช่นริมฝีปาก ฟัน ปุ่มเหงือก เพดานส่วนต่าง ๆ เป็นต้น กระบวนการทำให้เกิดเสียงพูดของมนุษย์ที่กล่าวมา สามารถแสดงบล็อกไดอะแกรม (Block Diagram) อย่างง่าย ดังแสดงในรูปที่ 2.2



รูปที่ 2.2 บล็อกไดอะแกรมของระบบการสร้างเสียงพูด

ภาคสร้างขบวนอิมพัลส์ (Impulse Train Generator) จะสร้างขบวนอิมพัลส์ ซึ่งมีระยะห่างระหว่างอิมพัลส์เท่ากับ คาบเวลาพิทช์ (Pitch Period) ขบวนอิมพัลส์นี้จะถูกส่งไปกระตุ้น โมเดลช่องเส้นเสียง (Glottal Pulse Model) , $G(z)$ ทำให้สัญญาณออกจากโมเดลตามจังหวะคาบเวลาพิทช์ ผลทางสเปกตรัมของสัญญาณที่ออกจากโมเดลนี้มีความชันประมาณ -12 dB/Octave รูปแบบจำลองเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อย่างง่ายของช่องทางเดินเสียง อาจจะมองได้เป็นลักษณะของท่อทรงกลมตันที่มีต้นกำเนิดเสียงอยู่ที่ปลายข้างหนึ่ง (ส่วนของกล่องเสียง) และปลายอีกข้างหนึ่งจะเปิด (ส่วนของปาก) แสดงในรูปที่ 2.3



รูปที่ 2.3 แบบจำลองอย่างง่ายของการกำทอนช่องทางเดินเสียง

ดังนั้นมันจะเกิดการกำทอน (Resonance) ภายในท่อ การกำทอนของช่องทางเดินเสียงนี้จะสร้างรูปคลื่นเสียงที่มียอดสูงเด่นขึ้น ซึ่งเรียกว่า ฟอร์แมนท (Formants) ของเสียง ฟอร์แมนทที่มีความถี่ต่ำที่สุดจะเรียกว่า ฟอร์แมนทที่หนึ่ง ซึ่งจะมีค่าประมาณ 200 ถึง 1000 Hz ทั้งนี้ขึ้นกับขนาดของช่องทางเดินเสียงด้วย ส่วนฟอร์แมนทที่สองอยู่ถัดไปก็จะมีค่าประมาณ 500 ถึง 2500 Hz และฟอร์แมนทที่สามมีค่าประมาณ 1500 ถึง 3500 Hz เป็นต้น โดยที่ฟอร์แมนทที่หนึ่ง และสองนี้จะเป็นตัวแสดงคุณสมบัติที่สำคัญมากอันหนึ่งในการพิจารณาประยุกต์ใช้ในด้านการวิเคราะห์เสียงทางด้วยสเปกตรัม เสียงที่มีความถี่ตรงกับความถี่กำทอนจะปรากฏออกมาให้ได้ยินผ่านทางภาคโมเดลการแพร่ (Radiation Model) , $R(z)$ ซึ่งเป็นส่วนของปาก ผลทางสเปกตรัมของสัญญาณที่ออกจากโมเดลนี้มีความชันประมาณ -6 dB/Octave

ทฤษฎีการวิเคราะห์เสียงพูด

การวิเคราะห์เสียงพูด คือ พารามิเตอร์ต่าง ๆ ที่เป็นลักษณะเฉพาะ เพื่อใช้เป็นตัวแทนเสียง การวิเคราะห์พูดนั้นมีหลายวิธี ทั้งในด้านโดเมนเวลา (Time domain) คือ วิเคราะห์จากรูปคลื่นของสัญญาณเสียงตามแกนเวลาโดยตรง และด้านโดเมนความถี่ (Frequency domain) คือวิเคราะห์จากสเปกตรัมของเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) การวิเคราะห์สัญญาณเสียงในช่วงเวลาสั้น ๆ

เนื่องจากสัญญาณเสียงเป็นสัญญาณที่แปรตามเวลาที่มีการแปรเปลี่ยนที่ไม่แน่นอน ดังนั้นเทคนิคในการวิเคราะห์เสียงพูดส่วนใหญ่แล้ว จะสมมติให้สัญญาณเสียงมีคุณสมบัติ ที่เปลี่ยนแปลงสัมพันธ์กับเวลาอย่างช้าๆ นั่นก็คือเราจะต้องแบ่งทำการวิเคราะห์หาพารามิเตอร์ ของสัญญาณเสียงพูดในช่วงเวลาสั้น ๆ ที่เรียกว่า ช็อตไทม์วินโดว์ (Shot time Window) เพื่อจะได้มองเห็นเหมือนกับว่าเราหาพารามิเตอร์นั้น ๆ ได้มาจากสัญญาณเสียงที่อยู่ภายในวินโดว์ และมีความเสถียรภายในช่วงเวลาสั้น

การนำฟังก์ชันของวินโดว์ที่มีช่วงขนาดจำกัด $w(n)$ มาคูณเข้ากับสัญญาณ $s(n)$ จะทำให้ได้กลุ่มตัวอย่างของเสียงพูดที่ถูกกำหนดน้ำหนัก ให้แปรไปตามรูปร่างของวินโดว์ รูปแบบของวินโดว์ที่ใช้ในงานวิจัย เลือกกรอบแบบสี่เหลี่ยม (Rectangular Window) ซึ่งมีนิยามดังนี้

$$W(n) = \begin{cases} 1, n = 0, 1, \dots, n-1 \\ 0, n = \text{otherwise} \end{cases} \quad (2.1)$$

ในสมการที่ (1) นี้ คือการกำหนดช่องของการวิเคราะห์ให้มีจำนวนตัวอย่าง N ตัวอย่าง โดยปกติ การประยุกต์ส่วนมากจะใช้วินโดว์ที่มีความกว้าง 10 ถึง 30 มิลลิวินาที และเน้นเฉพาะช่องตรงกลางของช่องแคบให้เป็นส่วนของข้อมูลที่จะทำการวิเคราะห์

2) การวิเคราะห์ในโดเมนเวลา

การวิเคราะห์โดเมนเวลา คือ วิธีการประมวลสัญญาณเสียงพูดโดยยึดหลักการวิเคราะห์จากรูปคลื่นสัญญาณที่แปรตามเวลา ซึ่งมีอยู่หลายวิธีด้วยกัน การวิเคราะห์โดเมนเวลาที่ใช้ในงานวิจัยนี้จะกล่าวดังต่อไปนี้

2.1 ค่าอัตราการตัดศูนย์เฉลี่ยของเวลาช่วงสั้น (Short-Time Average Zero-Crossing Rate)

ในหัวข้อนี้ กล่าวถึงสัญญาณ Discrete-time, การตัดศูนย์ (Zero-Crossing) ที่พูดถึงถ้าเกิดมีสัญญาณตัวอย่างที่ทำให้เกิดเครื่องหมายทางคณิตศาสตร์ที่แตกต่างกัน อัตราการตัดศูนย์ เป็นการวัดสัญญาณความถี่ตัวอย่างซึ่งรวมกันอยู่ในรูปสัญญาณแคบๆ ตัวอย่างสัญญาณรูปซายน์ของความถี่ = F_0 , อัตรา sample = F_s , จะได้ F_s/F_0 คือจำนวน sample ต่อ สัญญาณซายน์ 1 รูปคลื่น แต่ละ 1 รูปคลื่นจะมีการตัดศูนย์ (Zero-Crossing) 2 ครั้ง จึงให้ค่าเฉลี่ยอัตราการตัดศูนย์ คือ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$Z = 2F_0/F_s \quad (\text{crossing / sample}) \quad (2.2)$$

ดังนั้น อัตราการเกิดการตัดศูนย์ (Zero-Crossing) จึงให้เหตุผลในการคาดเดา ความถี่ของสัญญาณไซน์ (Sine Signal) สัญญาณเสียงเป็นสัญญาณบรอดแบนด์ (Broadband) และการแปลงอัตราเฉลี่ยการตัดศูนย์ (Zero-Crossing Rate) มีความเที่ยงตรงน้อยมาก อย่างไรก็ตาม การคาดคะเนแบบหยาบๆ คุณสมบัติที่ได้รับ โดยการแทนฐานเวลาที่สั้นทำให้เกิดอัตราการตัดศูนย์ (Zero-Crossing Rate) ก่อนที่จะพูดถึงการแปลงอัตราการตัดศูนย์ สำหรับสัญญาณเสียงพูด อย่างแรกที่จะกำหนด คำนวณ และอธิบาย กำหนดนิยามได้ดังนี้

$$Z_n = \sum_{m=-\infty}^{\infty} | \text{Sgn}[X(m)] - \text{Sgn}[X(m-1)] | W(n-m) \quad (2.3)$$

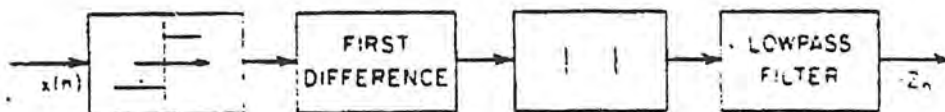
ซึ่ง

$$\text{Sgn}[X(m)] = \begin{cases} 1 & X(n) \geq 0 \\ -1 & X(n) < 0 \end{cases} \quad (2.4)$$

และ

$$W(n) = \begin{cases} 1/2N & 0 \leq n \leq N-1 \\ 0 & \text{elsewhere} \end{cases} \quad (2.5)$$

ดังนั้น การกระทำซึ่งเกี่ยวข้องกับ สมการ (2.2) จะถูกแสดงในรูปที่ 2.4 ซึ่งแทนการหาค่าอัตรา การตัดศูนย์ เฉลี่ยในช่วงเวลาสั้นๆ จะมีคุณสมบัติต่างๆ ไปเหมือนกับการหาค่าพลังงานและค่าสัมบูรณ์เฉลี่ยเวลาช่วงสั้น อย่างไรก็ตามในสมการ (2.2) และรูปที่ 2.4 จะทำการคำนวณค่า Z_n ซึ่งเกิดขึ้นยากกว่าความเป็นจริงซึ่งทั้งหมดต้องการตรวจสอบ 2 สัญญาณ ซึ่งจะถูกกำหนดค่าอัตราเฉลี่ยการตัดศูนย์ จะถูกคำนวณมากกว่า N สัญญาณตัวอย่างที่ติดๆกัน (ถูกแบ่งโดย N เป็นสัญญาณที่ไม่ต้องการ)

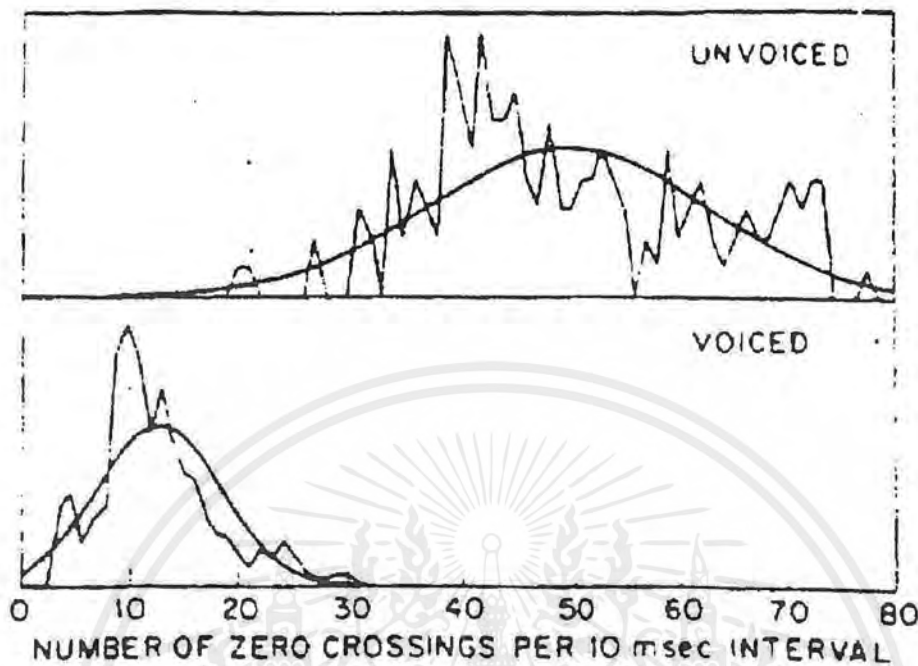


รูปที่ 2.4 บล็อกไดอะแกรมของอัตราการตัดศูนย์เฉลี่ยในเวลาช่วงสั้น

จากที่กล่าวมาสามารถบอกลำดับความสำคัญของค่าเฉลี่ย และถ้ากำหนดในรูปแบบการ โดยกำหนด ความยาวของหน้าต่าง(window)สมมาตร ,การถ่วงเวลาจะสามารถชดเชยได้อย่างถูกต้องตัวอย่างนี้เราว่าการเกิดอัตราการตัดศูนย์ (Zero-Crossing Rate) เฉลี่ยในช่วงเวลาสั้นจะนำไปประยุกต์ใช้เป็นสัญญาณการพูด รูปแบบของการกำเนิดเสียงพูด ซึ่งความถี่ต่ำกว่า 3 KHz การเกิดสเปกตรัม (spectrum) จะถูกบอกโดยคลื่นเสียง อย่างไรก็ตามเสียงเสียดสีฟัน (Unvoice) ส่วนมากพลังงานจะถูกพบที่ความถี่สูงมาก จึงกำหนดให้เป็นอัตราการตัดศูนย์สูง (High-zero-crossing) และที่ความถี่ต่ำเป็น อัตราการตัดศูนย์ต่ำ (Low-zero-crossing) ซึ่งมีความเกี่ยวข้องกันอย่างเหนียวแน่นระหว่างอัตราการตัดศูนย์ และการแจกแจงพลังงานกับความถี่ เป็นเหตุผลทั่วไปถ้าอัตราการตัดศูนย์ (Low-zero-crossing) มาก สัญญาณเสียงพูดจะเป็น เสียงเสียดสีฟัน (Unvoice) ขณะเดียวกันถ้าอัตราการตัดศูนย์ต่ำ สัญญาณเสียงพูดจะเป็น Voice อย่างใดก็ตามความคิดนี้อาจไม่เที่ยงตรงนักเพราะมนุษย์มีเสียงพูดทั้งสูงและต่ำ เป็นไปไม่ได้ที่จะถูกต้อง 100% รูปที่ 2.5 จะแสดง Histogram ของอัตราการตัดศูนย์ เฉลี่ยประมาณ >10 ms. สำหรับทั้งเสียงเสียดสีฟัน (Voice) และ เสียงไม่เสียดสีฟัน (Unvoice) สังเกตจากรูปเคิร์ฟของเกาส์เซียน เป็นการแจกแจงที่ดี , มีเหตุผล ค่าประมาณของอัตราการตัดศูนย์เฉลี่ยของ เสียงเสียดสีฟัน $\cong 40/10$ ms. และ $14/10$ ms. สำหรับเสียงไม่เสียดสีฟัน จะเห็นได้อย่างชัดเจนว่า อัตราการตัดศูนย์ จะเกิดการเหลื่อมล้ำ(overlap) ดังนั้น เสียงเสียดสีฟัน / เสียงไม่เสียดสีฟัน จะใช้เป็นพื้นฐานการตัดสินใจเพียงอย่างเดียวไม่ได้ แต่การแทนค่า เสียงเสียดสีฟัน / เสียงไม่เสียดสีฟัน จะเกิดประโยชน์ในการทำให้เกิดความแตกต่างได้

สำหรับตัวอย่างค่าอัตราการตัดศูนย์เฉลี่ย จะแสดงในรูปที่ 2.6 ช่วงเวลาที่เปลี่ยนแปลงจะถูกครอบด้วยวินโดว์ที่เวลา 15 ms. (150 ครั้ง ที่ 10 KHz) และ เอาท์พุทที่ได้จะถูกคำนวณที่ 100 Time/sec (window เลื่อนไป 100 ครั้ง) ข้อสังเกตในกรณีของพลังงานและค่าสัมบูรณ์เฉลี่ยต่ำมาก ดังนั้นอัตราการตัดศูนย์จะต้องพิจารณา เสียงเสียดสีฟันและ เสียงไม่เสียดสีฟัน มีลักษณะสูงๆ ต่ำๆ ดังรูปที่ 2.5 ตัวอย่าง อัตราการตัดศูนย์เฉลี่ยพิจารณาได้จากรูปที่ 2.6 ดังนั้นจึงใช้แนวความคิดพื้นฐาน (Algorithm) สำหรับคำนวณเกี่ยวกับค่าเปรียบเทียบกับสัญญาณการสุ่มทั้งคู่ต้องใส่ใจอย่างมากใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.5 การแจกแจงของอัตราการตัดศูนย์ สำหรับเสียงไม่เสียดสีฟัน และ เสียงเสียดสีฟัน

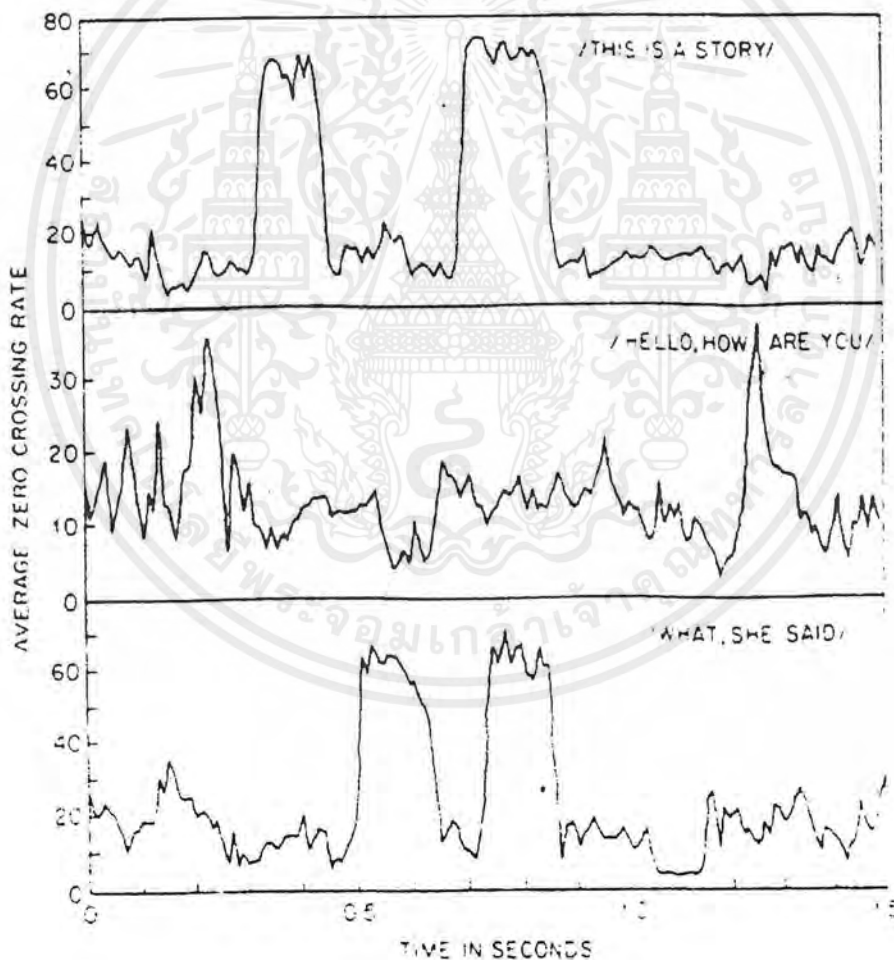
กระบวนการสุ่ม และเห็นได้ชัดเจนว่า อัตราการตัดศูนย์ จะมีผลอย่างมากในการเปลี่ยนแปลงสัญญาณอนาล็อก เป็นสัญญาณดิจิทัล (A/D converter) ซึ่งจะมีสัญญาณ ฮัม (hum) ออกมา 60 Hz ซึ่งแต่ละเสียงสามารถแทนในระบบดิจิทัล แต่ต้องระมัดระวังอย่างมากในการกระทำกระบวนการอนาล็อก สำหรับตัวอย่าง จะใช้วงจรกรองแถบความถี่ผ่าน (Band pass Filter) มากกว่า วงจรกรองความถี่ต่ำ (Low pass Filter) หรืออีกในหนึ่งเป็นการกรองที่บล็อกไฟกระสเตรง และ ส่วนประกอบ 60 Hz ในสัญญาณ นอกเหนือจากนี้เรายังใช้อัตราการตัดศูนย์ การสุ่มช่วงเวลาโดยให้เป็น T และช่วงเฉลี่ย N

ช่วงคาบของการสุ่มจะกำหนดเวลา (และความถี่) การแจกแจงโดยการแทนอัตราการตัดศูนย์ เป็นการเตรียมให้รายละเอียด 1 บิต (การเตรียมนี้จะเป็นการให้เครื่องหมายทางคณิตศาสตร์) ซึ่งทั้งหมดถูกต้องเพราะว่าการกำหนดจำนวนที่เหมาะสม จะมีหลากหลายมากความหลากหลายในความเหมือนกันในการแทนค่า จะถูกนำเสนอมาใช้งาน ความแตกต่างเหล่านี้จะทำให้เกิดความน่าสนใจในความไวของเสียงที่มีความไวน้อย แต่ละเสียงจะมีค่ากำหนดของแต่ละคน ข้อสังเกต เรา จะทำการแทนค่า การตัดช่วงบน (Up-crossing) การแสดงนี้จะพื้นฐานของการหาอัตราการตัดศูนย์ ที่เกิดในช่วงสไลปบวก

2.2 การเปรียบเทียบความแตกต่างของเสียงกับความเจ็บโดยใช้พลัง และอัตราการตัด

ศูนย์

ปัญหาของการหาตำแหน่งเริ่มต้น และตำแหน่งสุดท้ายของเสียงเป็นสิ่งที่สำคัญใน ขบวนการเสียง โดยทั่วไปการบันทึกเสียงอัตโนมัติของคำโคค มีความจำเป็นที่จะบอกตำแหน่ง กำเนิดสัญญาณ เสียง และการ ติดต่อแต่ละเสียง สำหรับการคำนวณหาตำแหน่งเริ่มต้น และตำแหน่งสุดท้าย ทำได้โดย การกำจัดการรบกวนของการคำนวณสัญญาณในระบบเวลาที่ไม่วาง โดย การทำให้เป็นอินพุท เพื่อบอกเป็นเสียงพูดปัญหาการแยกคำพูดจากเสียงไม่สำคัญ ยกเว้นในกรณีที่ว่า อัตราส่วนสัญญาณต่อสัญญาณรบกวนมีค่าสูง เช่น ในห้องอัด การบันทึกเสียงไฮไฟ ในห้องอัดที่ไม่มีเสียงเล็ดลอดออกไป



รูปที่ 2.6 อัตราการเปลี่ยนแปลงเสียงพูดที่แตกต่างกัน 3 ครั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับอัตราส่วนสัญญาณต่อสัญญาณรบกวนที่มีค่าสูงพลังงานของคำพูดเสียงก็จะต่ำที่สุด ดังนั้นวิธีการนี้จะใช้ได้กับการแยกเสียงอย่างง่าย โดยการบันทึกนี้จะไม่ได้เป็นวิธีการที่ดีที่สุด แนวความคิดพื้นฐาน(Algorithm) จะถูกอธิบายโดยพื้นฐานอย่างง่าย 2 ข้อ ในรูปเวลา-โดเมนคือ พลังงาน และอัตราการตัดศูนย์ ตัวอย่างง่ายๆทั้งหมดจะแสดงให้เห็นถึงความยากในการหาจุดเริ่มต้นและจุดสุดท้ายของเสียงพูดที่ออกมา กรณีของแนวความคิดทั้งหมด สามารถกำหนดจุดสิ้นสุดอย่างถูกต้องดังนั้นมันจึงมีเหตุผลเพียงพอสำหรับกำหนดช่วงเวลาของ UNVOICE ซึ่งรวมอยู่ภายในคำพูดทั้งหมด การประยุกต์ใช้งานของอัตราการตัดศูนย์ นี้ค่าสัมบูรณ์ขนาด (Magnitude) เหลือจะทำให้ไม่เกิน

2.3 การประมาณค่าช่วงเวลาพิทซ์ โดยวิธีการเข้าใกล้เส้นขนาน

ช่วงเวลาพิทซ์เป็นสัญญาณที่มีความสำคัญมากในกระบวนการพูด การถอดค่าพิทซ์ถูกใช้ในการกำหนดแยกแยะเสียงพูด ซึ่งมีหลักการพื้นฐานสำคัญดังแผนผังดังนี้

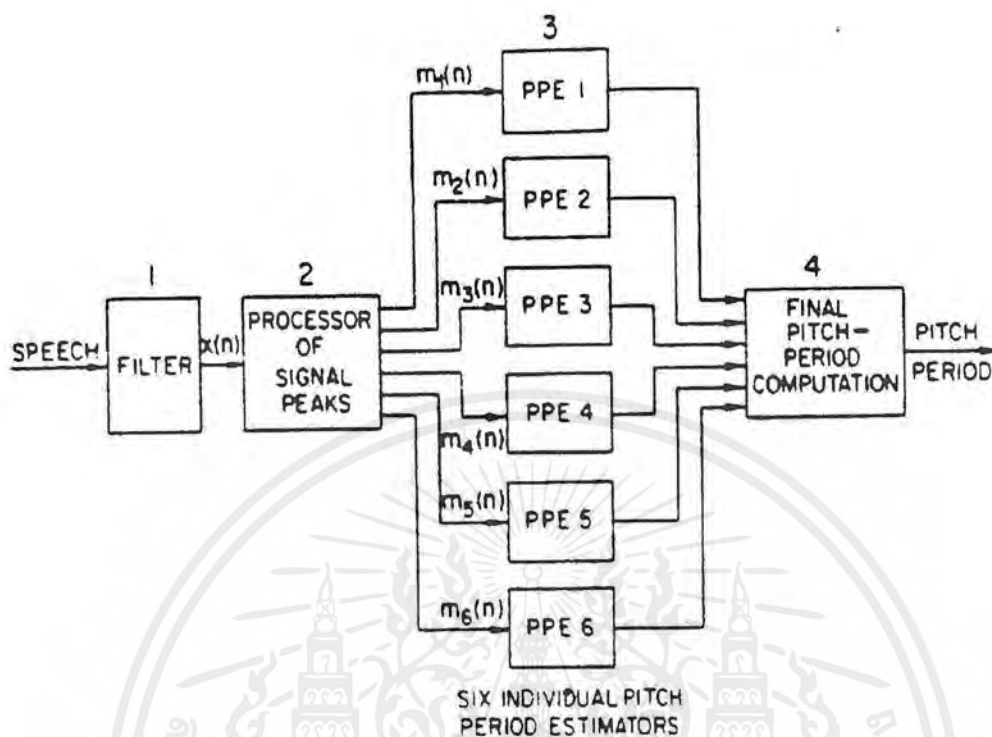
1) สัญญาณการพูดเป็นกรรมวิธีในการสร้างสัญญาณอิมพัลส์ ซึ่งจะรักษาช่วงเวลาเริ่มต้นและทั้งโครงสร้างที่ไม่ต้องการ โดยกรรมวิธีการถอดค่าพิทซ์

2) ขบวนการที่ยอมให้ถอดพิทซ์อย่างง่ายถูกใช้ประมาณช่วงเวลาของแต่ละขบวนการอิมพัลส์

3) การคาดคะเนของทุกๆสัญญาณการถอดค่าพิทซ์ทั้งหมดเป็นการรวมกันในสัญญาณลอจิกซึ่งจะสรุปรูปแบบช่วงเวลาของคำพูด

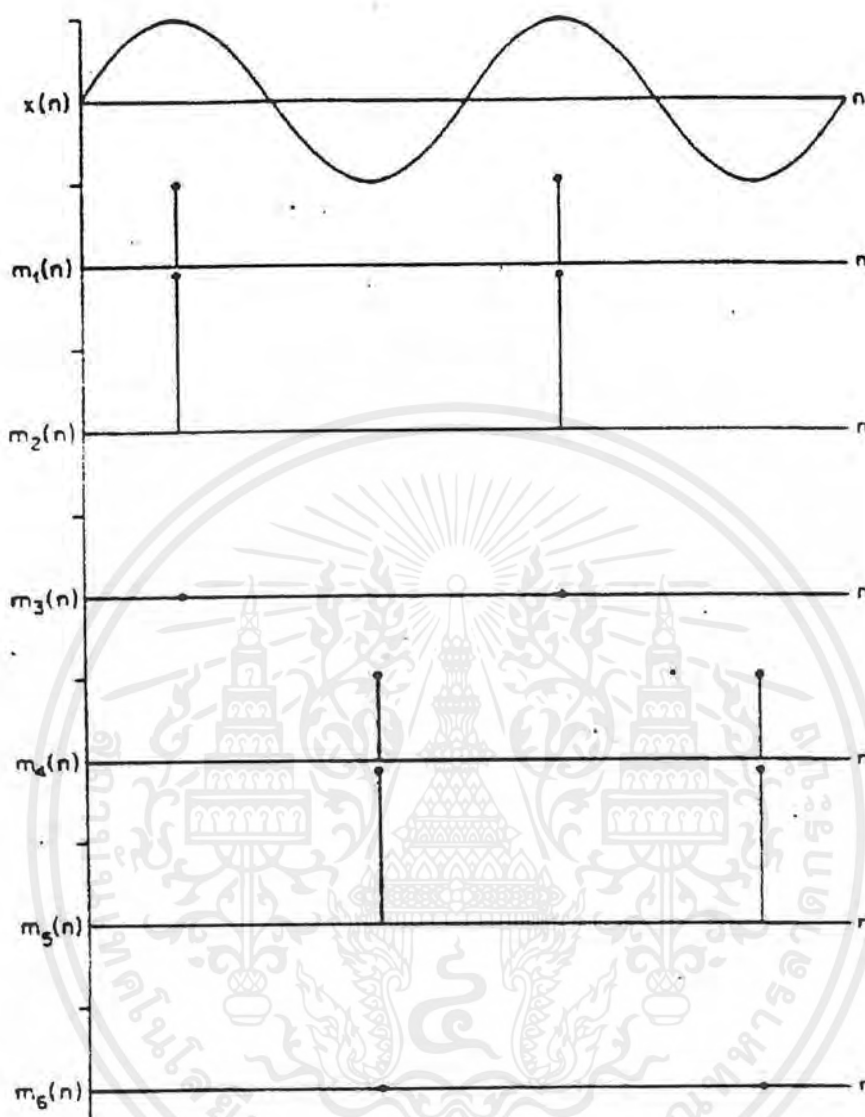
แผนผังนี้ถูกกำหนดโดยข้อคิดเห็นของ Gold และ Robiner ดังอธิบายในรูปที่ 2.7 รูปแบบของคำพูดถูกทดสอบด้วยอัตราส่วนที่เพียงพอ เพื่อให้จำนวนเวลาที่พอเพียง การทดสอบที่ 10 KHz ยอมให้ช่วงเวลาที่กำหนดภายใน 10^{-4} sec คำพูดนี้จะเป็นตัวกรองความถี่ต่ำผ่านมีจุดคัทออฟที่ 900 Hz เพื่อไปผลิตทำให้เกิดรูปแบบที่มีความสัมพันธ์กับตัวกรองแถบความถี่ผ่านความถี่ที่ 100Hz และ 900 Hz มีความจำเป็นในการขจัดสัญญาณเสียงที่ 60 Hz ในการใช้งานฟิลเตอร์นี้สามารถทำทั้งคู่ได้ด้วย

เป็นในการขจัดสัญญาณเสียงที่ 60 Hz ในการใช้งานฟิลเตอร์นี้สามารถทำทั้งคู่ได้ด้วย อนุลอกฟิลเตอร์ก่อนการสุ่มด้วยดิจิตอลฟิลเตอร์ จากการกรองที่ตำแหน่งสูงสุดจากตำแหน่งเหล่านี้และแอมพลิจูดทุกอิมพัลในรูปที่ 2.7 ได้มาจาก สัญญาณฟิลเตอร์แต่ละอิมพัลส์ประกอบไปด้วยอิมพัลส์บวกที่เกิดด้วยที่ตำแหน่งของ Peak และ Valey โดยทั้ง 6 กรณีนี้จะใช้หลักการของ Gold และ Robiner



รูปที่ 2.7 บล็อกไดอะแกรมของการถอดสัญญาณพิทช์แบบขนาน

- 1) $M_1(n)$: เมื่ออิมพัลส์มีขนาดแอมพลิจูดสูงสุดที่เท่ากันเกิดขึ้นที่แต่ละจุดของ Peak
- 2) $M_2(n)$: อิมพัลส์ที่เท่ากันแต่มีความแตกต่างกันที่แอมพลิจูดและเกิดด้านแอมพลิจูดของ Valey ซึ่งเกิดขึ้นแต่ละ Peak
- 3) $M_3(n)$: อิมพัลส์ที่เท่ากันและที่แตกต่างกันที่แอมพลิจูดที่สูงสุดและเกิดก่อนแต่ละ Peak
- 4) $M_4(n)$: อิมพัลส์ที่เท่ากันและเป็นลบที่แอมพลิจูดของ Valey เกิดขึ้นแต่ละ Valey
- 5) $M_5(n)$: อิมพัลส์ที่เท่ากันเป็นแอมพลิจูดลบที่ Valey บวก แอมพลิจูดที่อยู่หน้า Peak เกิดขึ้นแต่ละช่วงของ Valey
- 6) $M_6(n)$: อิมพัลส์ที่เท่ากันจะเป็นลบของแอมพลิจูดที่ Valey รวมกับแอมพลิจูดที่เกิดขึ้นก่อน ที่เกิดขึ้นที่จุดต่ำสุดที่แต่ละ Valey



รูปที่ 2.8 อินพุตสัญญาณชาวนั้นและการเกิดขบวนการอิมพัลส์จาก Peak และ Valey

ผลเหล่านี้จะเกิดการตัดสินใจของ แนวคิด ซึ่งต้องการช่วงเวลา 3 ช่วงขอดเพื่อให้อยู่ในกรรมวิธีนี้ ในการศึกษาเปรียบเทียบในปัจจุบันนี้ การถอดขอด แนวคิดจะสามารถทำงานภายใต้สถานะที่กว้างมาก การเปรียบเทียบวิธีนี้จะดีกว่าการกะประมาณขอด ซึ่งแสดงให้เห็นแล้ว

ในการสรุปวิธีการ ลายละเอียด กรรมวิธี อย่างแรกกรรมวิธีที่สัญญาณเสียงได้ทำการเซตอิมพัลส์ ซึ่งมีความจำเป็น(หรือการสูญเสียของช่วงเวลา)เพราะ โครงสร้างของสัญญาณอย่างง่าย ความง่ายในการประมาณค่าขอดทำให้เกิดการประมาณค่าขอดที่ดี สุดท้ายการประมาณก็จะรวมกันเพื่อเพิ่มความเชื่อถือ การประมาณค่าขอดให้เป็นที่ยอมรับดังนั้นจะเป็นกรรมวิธีอย่างง่ายที่จะบรรลุการคาดเดาที่ยาก เพื่อให้ได้สัญญาณคำพูดที่ต้องการ การทำงานทางลอจิก สามารถทำให้สำเร็จในอัตราที่ต่ำ (เป็นต้นว่า 100 ครั้ง/วินาที) มากกว่ากรรมวิธีของสัญญาณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับความรู้ทางวิชาการเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) การวิเคราะห์ตามแกนความดี

คือ การวิเคราะห์ทางสเปกตรัมนั่นเอง สาเหตุที่ต้องมีการวิเคราะห์ทางด้านสเปกตรัม เนื่องจากพารามิเตอร์ในการประมวลสัญญาณเสียงส่วนใหญ่จะสามารถหาได้จากโดเมนความถี่ การวิเคราะห์สเปกตรัมของสัญญาณในช่วงสั้นๆ นั้นจะเป็นฐานของเทคนิคการวิเคราะห์เสียงต่างๆ



บทที่ 3

โครงสร้างงานวิจัย

งานวิจัยนี้ได้แบ่งโครงสร้างของเครื่องคิดเลขเป็น 4 ภาคใหญ่ ๆ คือ

- 1) ภาคอ่านไฟล์เสียงพูด
- 2) ภาครู้จำเสียงพูด
- 3) ภาคประมวลผลผลลัพธ์
- 4) ภาคแสดงผลด้วยเสียงพูดภาษาไทย

ภาคอ่านไฟล์เสียงพูด

ระบบรู้จำเสียงพูดภาษาไทยที่วิจัยอยู่นี้ขึ้นกับผู้พูด (Speaker dependent) โดยเสียงคำพูดเป็นแบบคำโดด (Isolated Speech) เสียงพูดจะถูกบันทึกในรูปแบบเวฟ (Wave Format) โดยผ่านทางไมโครโฟนที่ต่ออยู่กับการ์ดเสียง (Sound Card) ด้วยความถี่แซมปลิง 22050 Hz และความละเอียด 8 บิตต่อแซมปลิง (Bit/Sample) การบันทึกเสียงพูดจะใช้โปรแกรมที่ให้มากับการ์ดเสียง หรือโปรแกรมที่สามารถจัดการเกี่ยวกับเสียงได้ ระบบการบันทึกเสียงพูดสามารถเขียนเป็นบล็อกไดอะแกรมดังแสดงในรูปที่ 3.1



รูปที่ 3.1 ระบบการบันทึกเสียงพูด

ภาคอ่านไฟล์เสียงคำพูด จะทำหน้าที่อ่านข้อมูลเสียงจากไฟล์ที่อยู่ในรูปแบบเวฟ โดยในการบันทึกไฟล์ในรูปแบบ mat.ไฟล์ เราจะเรียกใช้ฟังก์ชัน Recordsound ของโปรแกรม Matlab และเก็บข้อมูลที่อ่านได้ไว้ในไฟล์ C:\PROJ_DATA\y ; โดยมีรูปแบบการใช้งานดังนี้

```
y = RECORDSND(SECOND, SAMPLING);
```

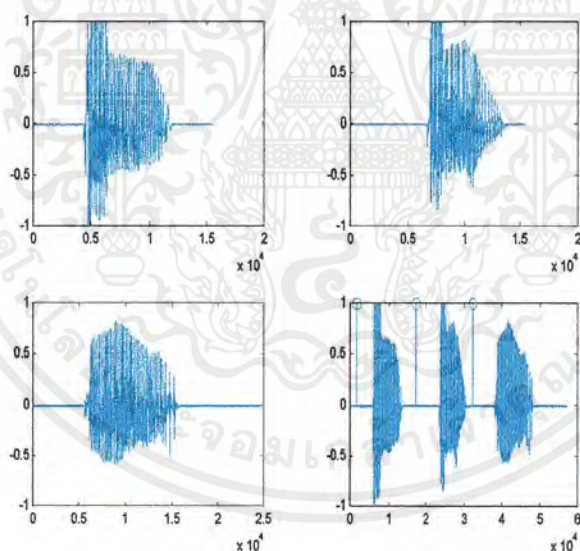
```
SAVE C:\PROJ_DATA\y y;
```

ภาครู้จำเสียงพูด

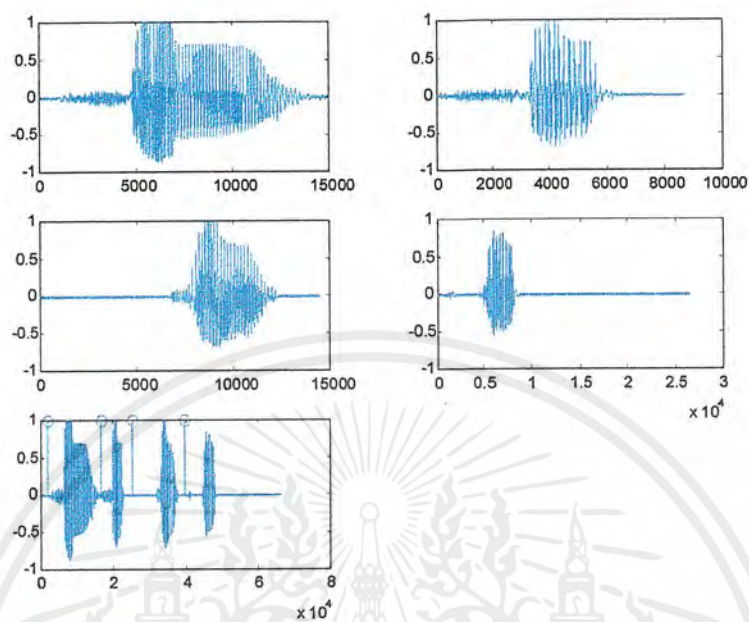
ทำหน้าที่วิเคราะห์หาความน่าจะเป็นของเสียงพูดว่าเหมือนคำใดมากที่สุด โดยจะประกอบด้วย ภาควิเคราะห์แยกพยางค์ออกจากประโยค ภาควิเคราะห์เริ่มต้นและจุดสุดท้ายของเสียงคำพูด, ภาควิเคราะห์พารามิเตอร์ของเสียงพูด และภาควิเคราะห์ตัดสินใจเสียงคำพูด ซึ่งแต่ละภาคมีหลักการทำงานดังนี้

1) ภาควิเคราะห์แยกพยางค์ออกจากประโยค

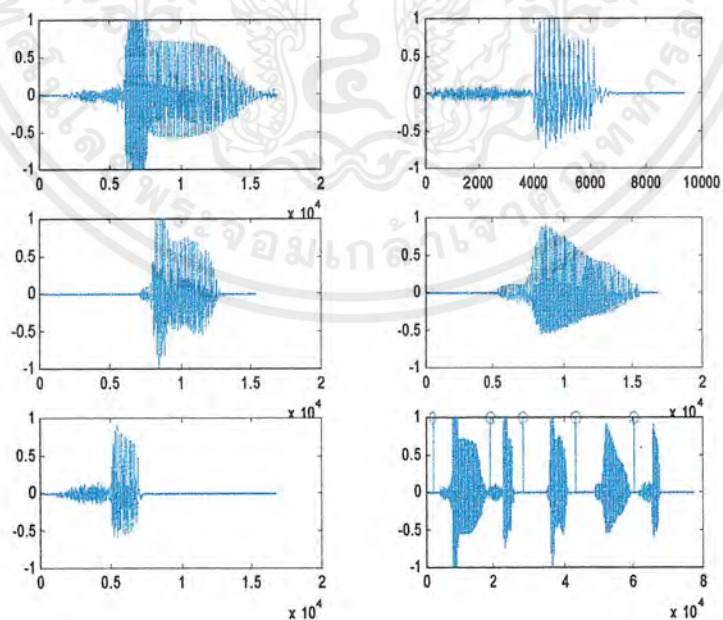
ภาคนี้จะทำหน้าที่แยกพยางค์แต่ละพยางค์ให้ออกจากประโยคที่ผู้ทดลองได้บันทึกเสียงเข้าไปไว้ในไฟล์เสียง ซึ่งในงานวิจัยนี้จะมีจำนวนพยางค์ที่แยกตั้งแต่ 3 พยางค์ จนถึง 7 พยางค์ ดังรูปที่ ซึ่งเราจะเรียกใช้ฟังก์ชัน `Cutword_1` ในการแยกพยางค์



รูปที่ 3.2ก แสดงการตัดพยางค์จำนวน 3 พยางค์ออกจากประโยค

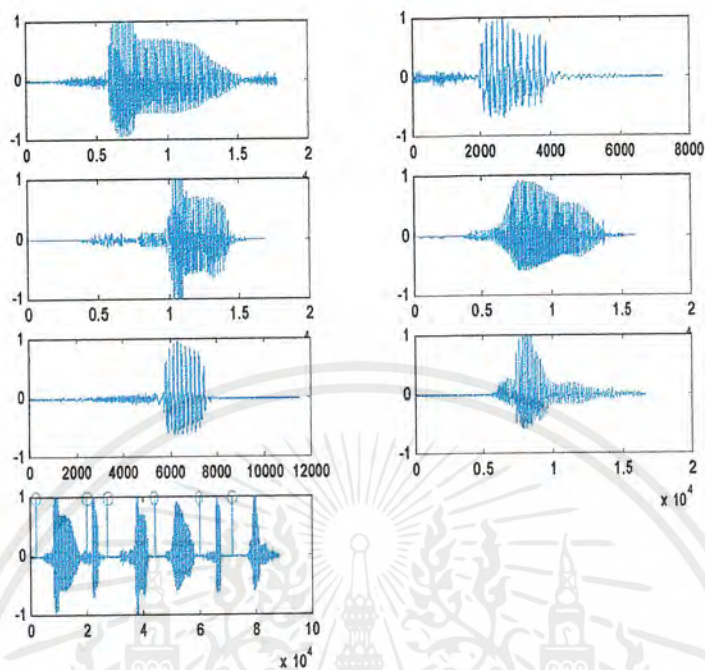


รูปที่ 3.2ข แสดงการตัดพยางค์จำนวน 4 พยางค์ออกจากประโยค

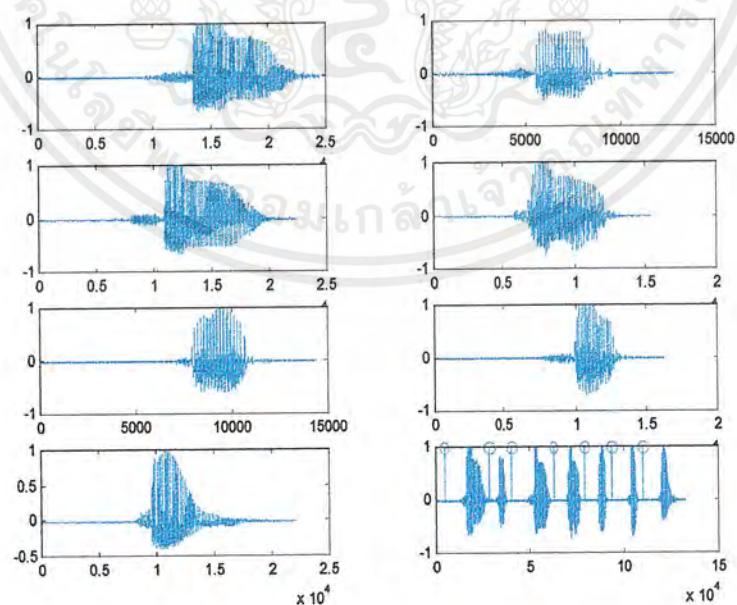


รูปที่ 3.2ค แสดงการตัดพยางค์จำนวน 5 พยางค์ออกจากประโยค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.2ง แสดงการตัดพยางค์จำนวน 6 พยางค์ออกจากประโยค



รูปที่ 3.2ข แสดงการตัดพยางค์จำนวน 7 พยางค์ออกจากประโยค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) ภาคหาจุดเริ่มต้น และจุดสุดท้ายของเสียงคำพูด

ภาคนี้มีหน้าที่หาจุดเริ่มต้น และจุดสุดท้ายของเสียงคำพูดแต่ละคำที่อ่านจากไฟล์เสียงพูด เพื่อนำเอาเฉพาะข้อมูลตรงส่วนที่เป็นสัญญาณเสียงคำพูดไปวิเคราะห์ในภาคอื่นถัดไป พร้อมทั้งหาจำนวนคำพูดว่ามีกี่คำ การหาจุดเริ่มต้น และจุดสุดท้ายของเสียงคำพูด จะเป็นการลดจำนวนข้อมูลที่ต้องนำไปวิเคราะห์ในภาคหาพารามิเตอร์เสียงคำพูดด้วย หลักการที่ใช้ในการหาจุดเริ่มต้น และจุดสุดท้ายของเสียงคำพูดคือ การหาค่าAmplitudeของสัญญาณเสียงที่ทำการ Nomalize Amplitudeในช่วงเวลาสั้น ๆ ในงานวิจัยนี้จะใช้วินโดว์แบบสี่เหลี่ยม (Rectangular window) ที่มีความกว้างเท่ากับ 8.16 มิลลิวินาที ทำการครอบแอมปริจูดของสัญญาณเสียงที่ทำการ Nomalize แอมปริจูด ในช่วงเวลาดังแต่เริ่มต้นถึงจุดสิ้นสุดซึ่งได้มาจากภาคการแยกพยางค์จากประโยค

3) ภาคการหาพารามิเตอร์ของเสียงพูด

นำสัญญาณเสียงคำพูดที่ผ่านการหาจุดเริ่มต้น และจุดสุดท้าย มาทำการแปลงฟูเรียร์อย่างรวดเร็ว (Fast Fourier Transfrom) ผลการแปลงจะได้สเปกตรัมของสัญญาณเสียง ที่มีจำนวนข้อมูลเท่ากับครึ่งหนึ่งของจำนวนแซมปลิงของสัญญาณเสียง โดยทั่วไปสเปกตรัมของสัญญาณเสียงในการวิจัยนี้จะมีจำนวนข้อมูลประมาณ 500 ถึง 2000 ข้อมูล ซึ่งเป็นจำนวนข้อมูลที่มากเกินไป ดังนั้นจำเป็นต้องมีการลดจำนวนข้อมูลลง เพื่อให้เหมาะสมกับการนำไปใช้ภาครู้จำเสียงพูดและในการทำวิจัยในครั้งนี้อัตราการเกิดจุดตัดศูนย์และ จำนวนยอดของเสียงใช้เป็นพารามิเตอร์ในการจำแนกเสียงออกเป็นกลุ่มๆซึ่งแบ่งได้ดังนี้

3.1) ค่าอัตราการเกิดจุดตัดศูนย์ ใช้ในการแบ่งประเภทเสียงออกเป็น เสียงไม่เสียดสีฟัน และ เสียงเสียดสีฟัน

3.2) ค่าจำนวนยอดของเสียง ใช้ในการแบ่งประเภทเสียงที่ได้จากการหาค่าอัตราการเกิดจุดตัดศูนย์ ออกเป็นสระเสียงสั้น (SHORT VOWEL) กับ สระเสียงยาว (LONG VOWEL)

4) ภาคการรู้จำเสียงคำพูด

ภาครู้จำเสียงคำพูดจะใช้การเปรียบเทียบค่าสเปกตรัมของเสียงที่บันทึกเข้าไปไว้ในไฟล์wav กับฐานข้อมูลของเสียงที่ได้บันทึกเก็บเอาไว้จำนวน 80 ฐานข้อมูล มาใช้ในการตัดสินใจว่าเป็นเสียงคำพูดใดในกลุ่มคำ 16 คำ

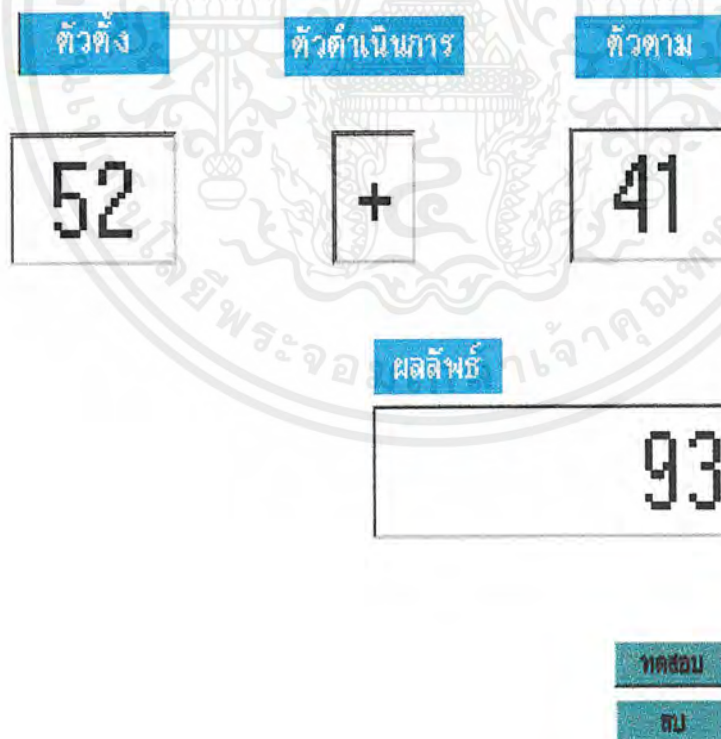
ภาคประมวลผล

ทำหน้าที่ คำนวณผลลัพธ์ที่ได้รับจากการตัดสินใจของภาครู้จำเสียงที่ได้ทำการตัดสินใจว่าเป็นคำใดจากกลุ่มคำทั้งหมด 16 คำ โดยการแปลความหมายเอาที่พหูพจน์ในแต่ละพยางค์ของประโยคให้เป็นจำนวนทางคณิตศาสตร์และทำการคำนวณผลลัพธ์ ตัวอย่างเช่น

$$79 + 11 = 100$$

ภาคแสดงผลการคำนวณด้วยเสียงพูด

ทำหน้าที่แปลงจำนวนในทางคณิตศาสตร์กลับเป็นตัวเลขอารบิกและคำพูดภาษาไทยโดยอาศัยฟังก์ชัน Test_cal ผลลัพธ์ที่ได้คือค่าตัวเลขจะถูกพูดออกมาเป็นจำนวนภาษาไทย ตัวอย่างเช่น ในรูปที่ 3.3 และเสียงพูดที่แสดงออกมาจะเป็นเสียงพูดของผู้วิจัยเอง โดยจะทำการบันทึกเสียงคำพูดในรูปแบบไฟล์เวฟ



รูปที่ 3.3 แสดงผลการคำนวณด้วยเสียงพูด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

วิธีดำเนินการวิจัย

งานวิจัยนี้ได้แบ่งวิธีดำเนินการวิจัยออกเป็น 4 ขั้นตอน ตามลำดับดังนี้ คือ

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

การเก็บข้อมูลกลุ่มตัวอย่างและกลุ่มทดสอบเสียงคำพูด

การเก็บข้อมูลกลุ่มตัวอย่าง และกลุ่มทดสอบเสียงคำพูด จะบันทึกเสียงคำพูดด้วยความถี่ 22050 Hz และความละเอียด 8 บิตต่อแซมปลิง คงสามารถฟังเสียงคำพูดที่บันทึกไว้ได้ การบันทึกเสียงทำในห้องที่ไม่มีเสียงรบกวน การเก็บข้อมูลกลุ่มตัวอย่างและกลุ่มทดสอบเสียงคำพูดจะเป็นเสียงคำพูดของผู้วิจัยเพียงคนเดียว โดยจะบันทึกเสียงคำพูดเป็นชุด ชุดหนึ่งมี 16 คำ คือ “ศูนย์”, “หนึ่ง”, “สอง”, “สาม”, “สี่”, “ห้า”, “หก”, “เจ็ด”, “แปด”, “เก้า”, “สิบ”, “บวก”, “ลบ”, “คูณ”, “หาร”, และ “เย้” การบันทึกเสียงคำพูดจะบันทึกวันละ 10 ชุด เป็นเวลา 10 วัน รวม 1600 คำ ทั้งนี้เพื่อลดผลกระทบจากเสียงคำพูดที่เปลี่ยนไปของผู้วิจัยเอง และแบ่งออกเป็น 2 กลุ่ม กลุ่มละ 50 ชุดเท่า ๆ กัน แต่ละกลุ่มประกอบด้วยเสียงคำพูด 800 คำ โดยใช้ชื่อ DATA_TEST และ DATA_BASE ตามลำดับ เพื่อใช้เป็นกลุ่มตัวอย่าง และกลุ่มทดสอบ กลุ่มตัวอย่าง และกลุ่มทดสอบที่สร้างขึ้นมานี้ มีไว้สำหรับการในการทดสอบภาคตัดสินใจเสียงพูด

การเตรียมข้อมูลกลุ่มตัวอย่าง และกลุ่มทดสอบเพื่อใช้ในการตัดสินใจ

ในการเตรียมข้อมูลที่ใช้ในการตัดสินใจ เราจะนำเสียงที่ถูกบันทึกเก็บไว้จำนวน 20 ชุด เสียง มาเป็นตัวแทนในการหาค่าพารามิเตอร์ 2 ชนิด คือ อัตราการตัดศูนย์ และ จำนวนยอดของเสียง เพื่อให้ได้ค่าพารามิเตอร์ที่เหมาะสม เพื่อที่จะใช้ในการแยกประเภทกลุ่มของคำใน 16 คำออกเป็นกลุ่มๆ ดังนี้

- 1) การหาอัตราการตัดศูนย์ เป็นพารามิเตอร์ในการแบ่งกลุ่มคำ 16 คำ ออกเป็น 2 กลุ่ม คือ

กลุ่มของเสียงไม่เสียดสีฟันและเสียงเสียดสีฟัน

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) การหาจำนวนยอดของเสียงเป็นพารามิเตอร์ในการแบ่งกลุ่มคำที่ถูกแบ่งมาจากการหาอัตราการจัดศูนย์ ออกอีกเป็น 2 กลุ่ม คือ กลุ่มสระเสียงสั้น (SHORT VOWEL) และ กลุ่มสระเสียงยาว (LONG VOWEL)

ทั้งนี้ ในการแบ่งกลุ่มแยกเป็นกลุ่มย่อยๆนั้น เพื่อให้จำนวนเสียงของฐานข้อมูลที่จะนำเสียงไปทดสอบมีจำนวนเสียงน้อยที่สุดเท่าที่จะเป็นไปได้ เพื่อให้การเปรียบเทียบมีค่าความถูกต้องสูงว่าการนำเสียงไปเปรียบเทียบกับเสียงทั้งหมด ซึ่งแสดงไว้ในตารางที่ 1 ในภาคผนวก

เมื่อกำหนดค่าอัตราการจัดศูนย์ และ จำนวนยอดของเสียง แล้ว เราจะสามารถสรุปได้ว่าในกลุ่มคำ 16 คำ เป็นกลุ่มคำประเภทไหน ซึ่งแสดงไว้ในตารางที่ 2 ภาคผนวก

การทดสอบความถูกต้องในภาคตัดสินใจ

ในการทดสอบความถูกต้องในการตัดสินใจนี้ ในขั้นแรกเราได้กำหนดจำนวนชุดเสียงของฐานข้อมูลของชุดเสียง (DATA BASE) 10 ชุด ทำการเปรียบเทียบกับชุดทดสอบ(DATA TEST) จำนวน 10 ชุด โดยใช้ฟังก์ชัน RECOGNITION โดยมีรูปแบบดังนี้

$$[\text{ค่าความถูกต้อง (\%)} , \text{เมตริกซ์ความถูกต้อง}] = \text{RECOGNITION ()};$$

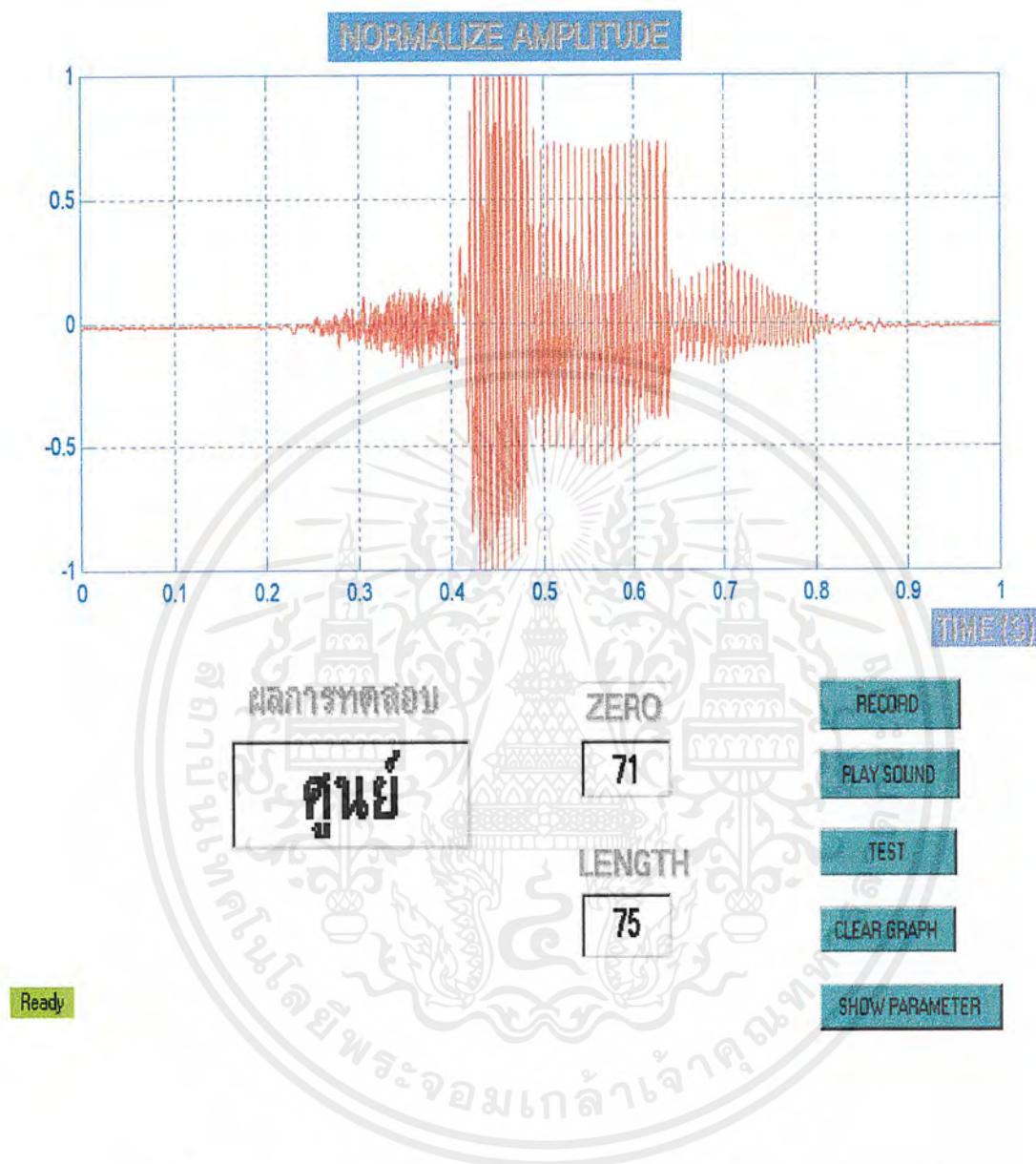
ต่อจากนั้นเราจะทำการเพิ่มชุดทดสอบเป็น 20,30,40 และ 50 ชุด ตามลำดับ แต่ยังคงจำนวนชุดของเสียงฐานข้อมูลเป็น 10 ชุด หลังจากนั้นเราก็จะทำการเพิ่มชุดเสียงของฐานข้อมูลเป็น 20,30,40 และ 50 ชุด ตามลำดับ โดยทำการทดสอบวิธีเดียวกับการใช้จำนวนชุดของเสียงฐานข้อมูล 10 ชุด

การทดสอบความถูกต้องกับโปรแกรมใช้งานจริง

ในการทดสอบความถูกต้องกับโปรแกรมใช้งานจริง โดยมีขั้นตอนในการทดสอบคือ ผู้วิจัยได้ทำการพูดเสียงทั้งหมดของกลุ่มคำจำนวน 16 คำ โดยพูดทีละคำ บันทึกเก็บไว้และทำการทดสอบโดยใช้ฟังก์ชัน

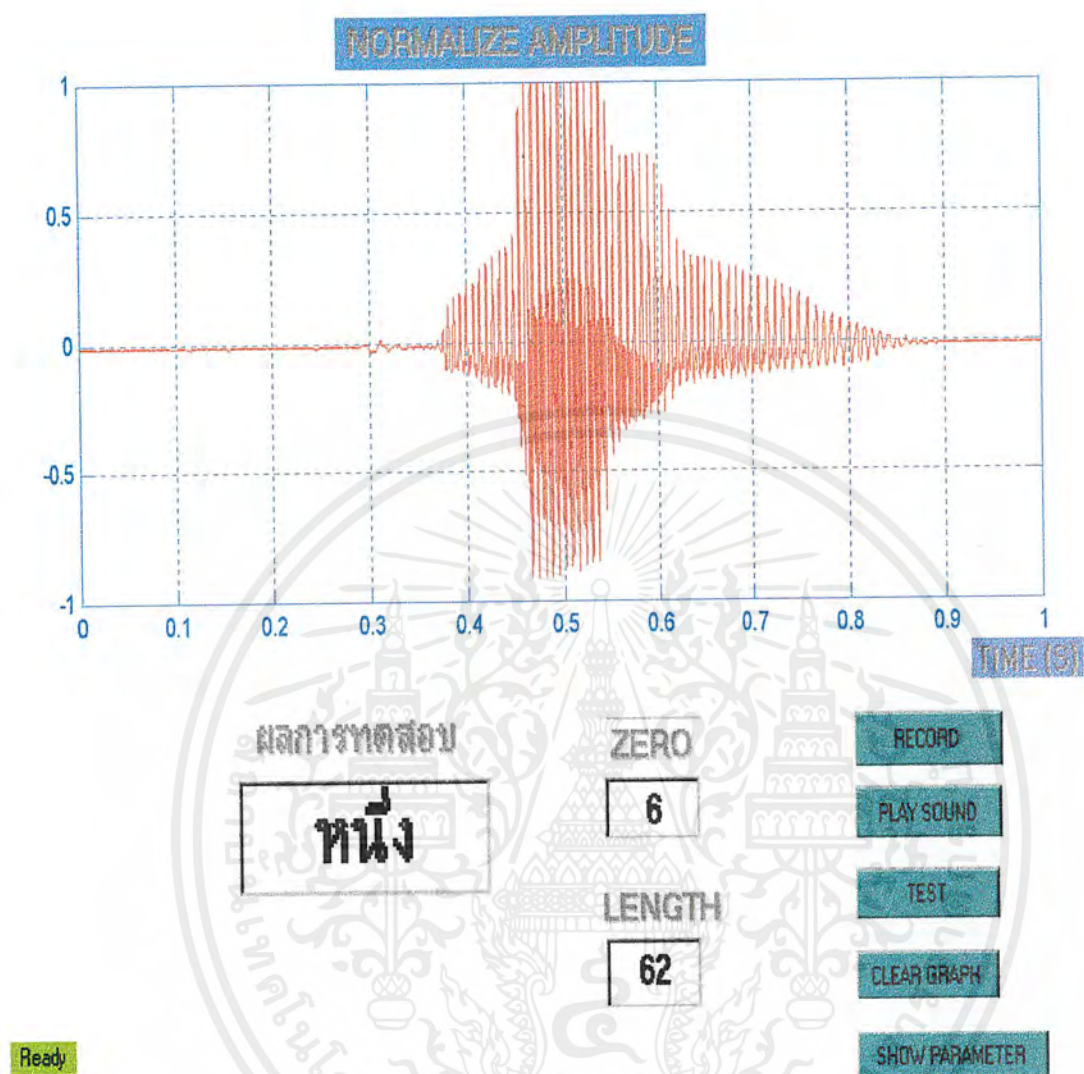
$$\text{RECORDSND (1 , SAMPLING RATE)};$$

ซึ่ง เริ่มแรกเราได้กำหนดจำนวนชุดเสียงของฐานข้อมูลจริงจำนวน 10 ชุด ทำการทดสอบความถูกต้องและบันทึกผลต่อจากนั้นเราได้เพิ่มจำนวนชุดเสียงของฐานข้อมูลเป็น 15,20,25,.....,80 ชุด ตามลำดับทำการทดสอบและบันทึกค่าตามลำดับ ดังแสดงในตารางที่3 ในภาคผนวก



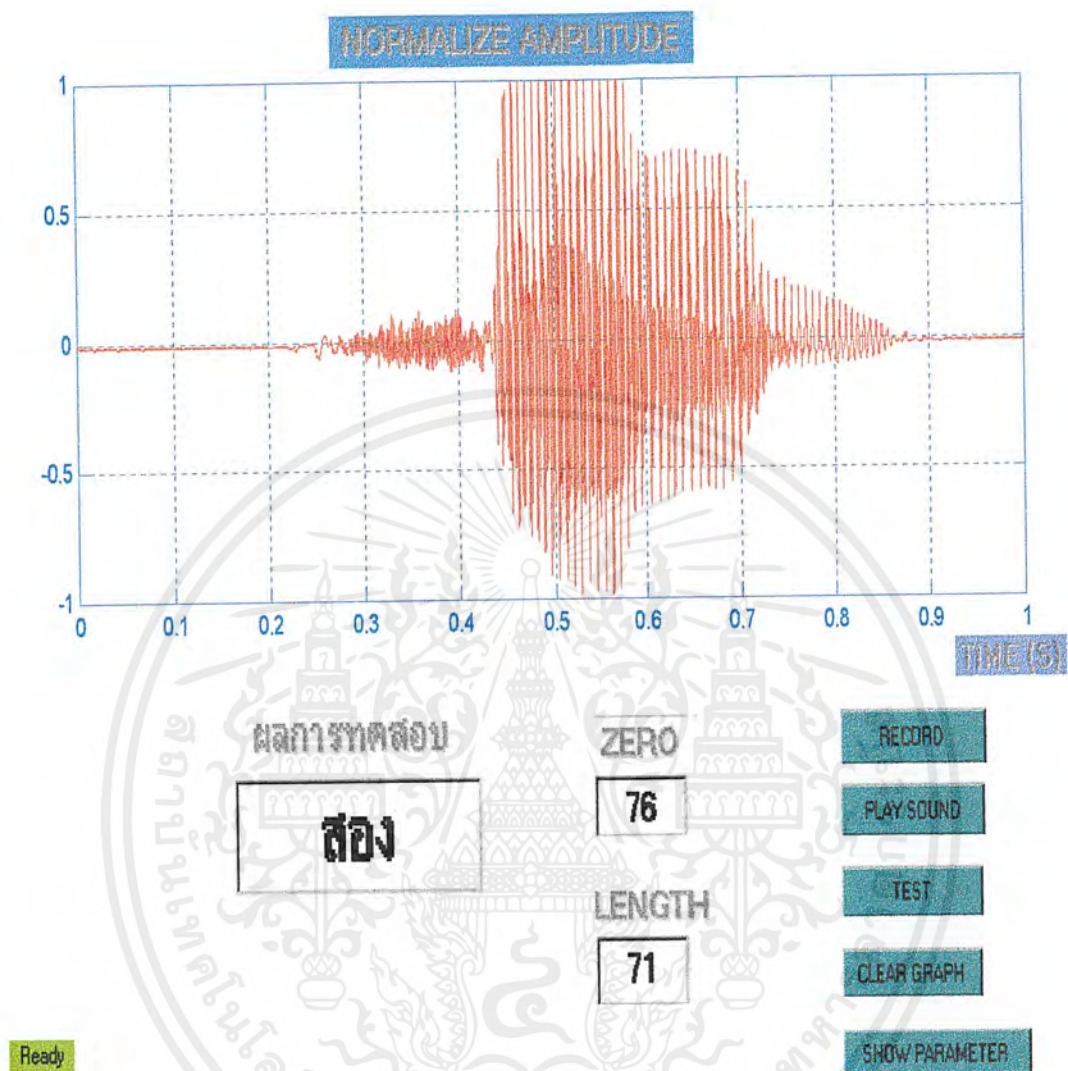
รูปที่ 4.1.1 แสดงการทดสอบความถูกต้องคำว่าศูนย์กับโปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



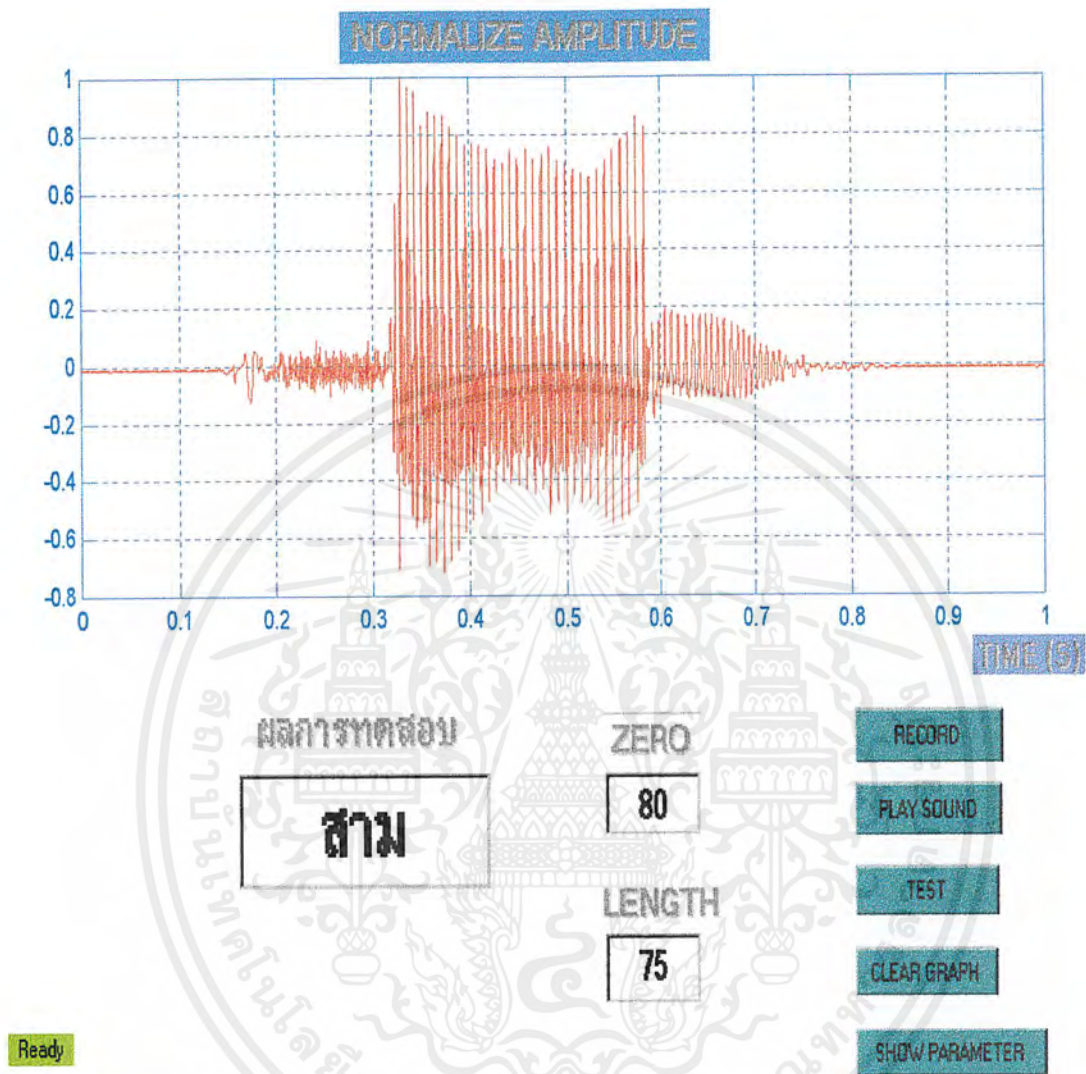
รูปที่ 4.1.2 แสดงการทดสอบความถูกต้องคำว่าหนึ่งกับ โปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



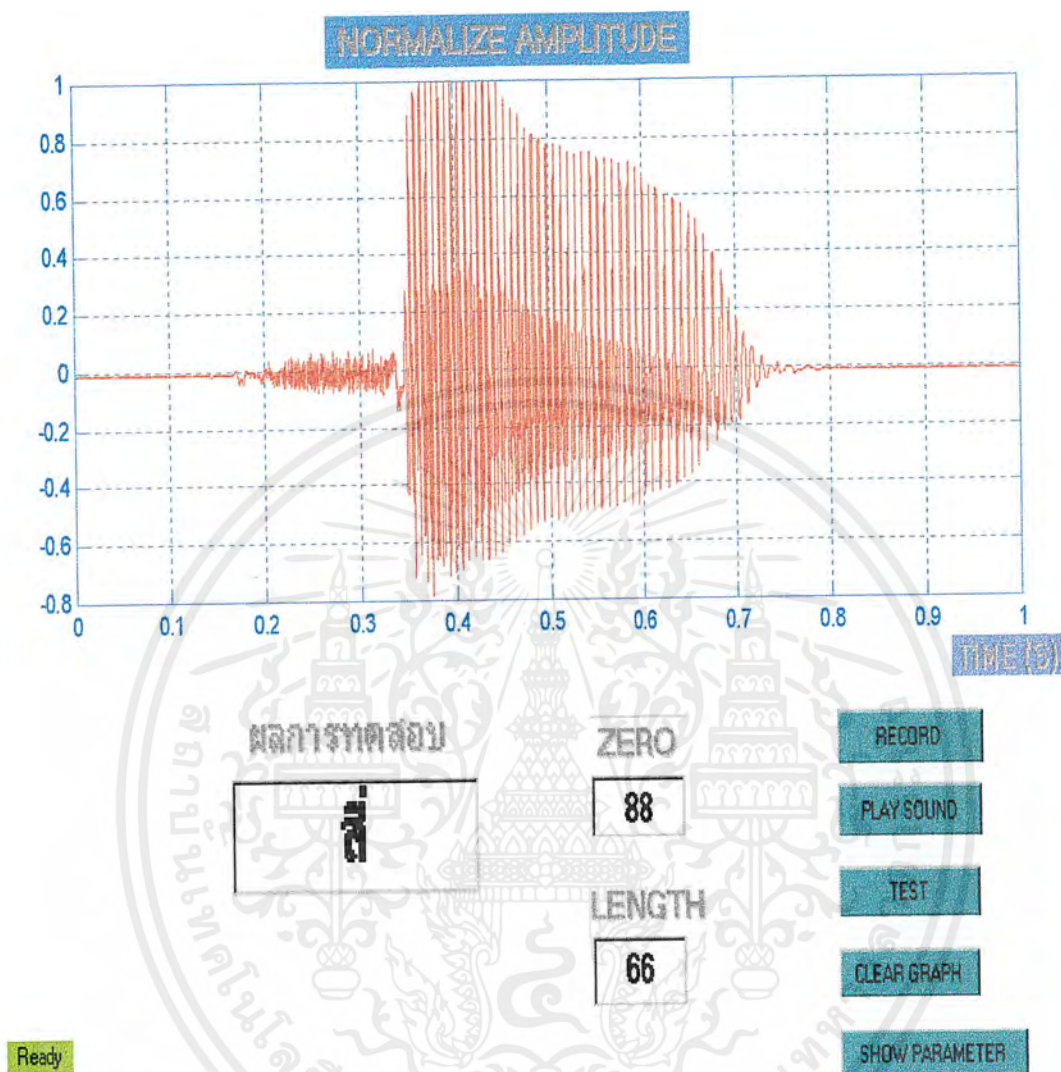
รูปที่ 4.1.3 แสดงการทดสอบความถูกต้องคำว่าสองกับ โปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



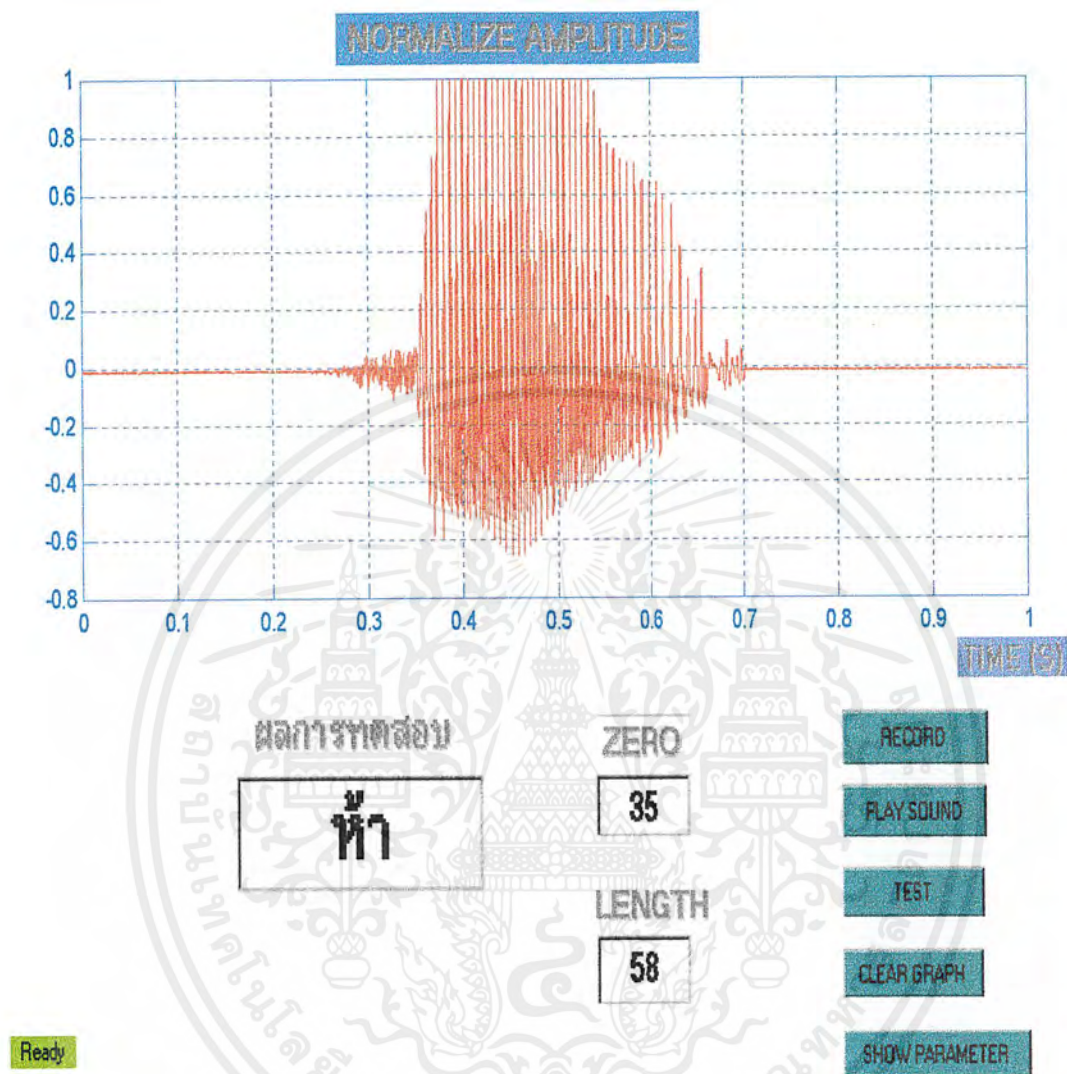
รูปที่ 4.1.4 แสดงการทดสอบความถูกต้องค่าว่าสามกับโปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



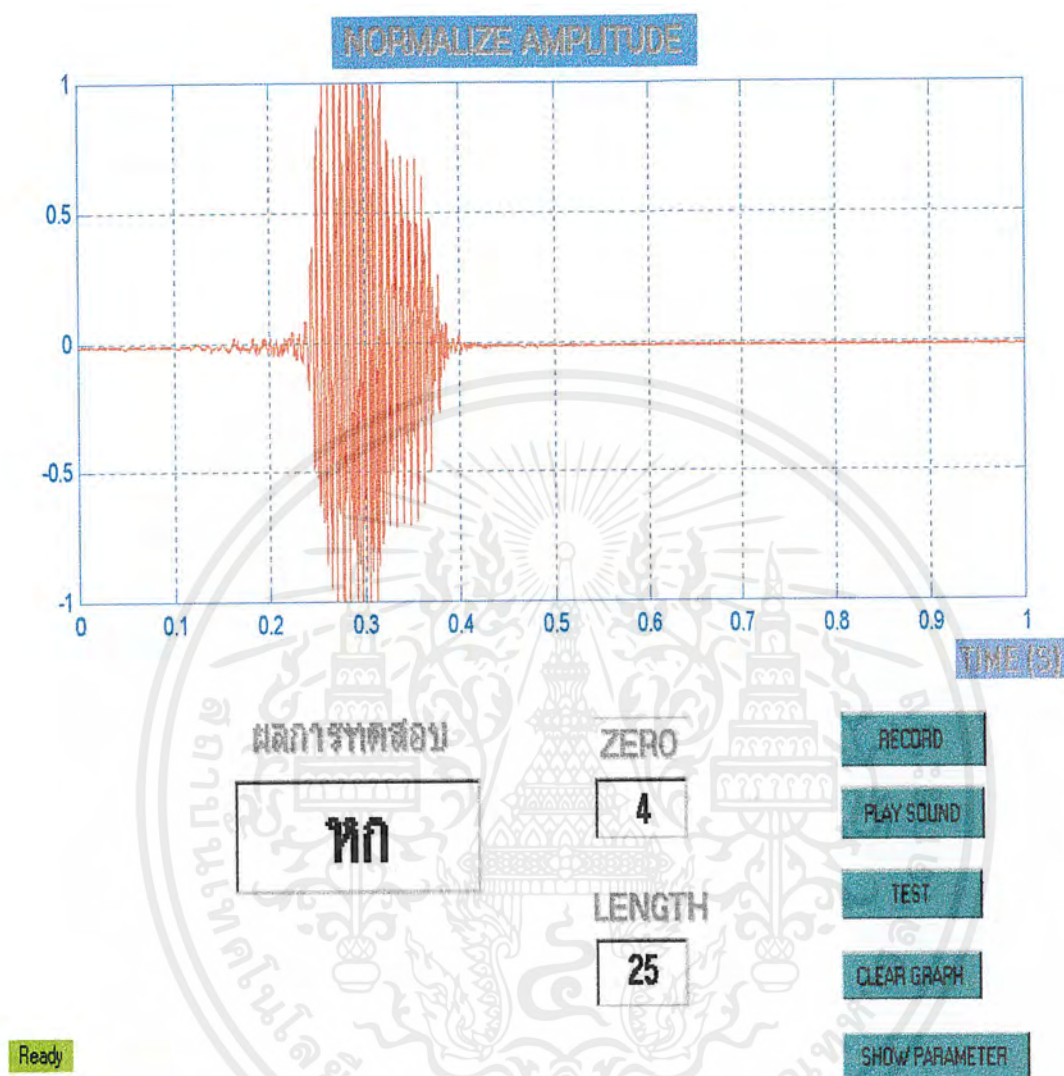
รูปที่ 4.1.5 แสดงการทดสอบความถูกต้องคำว่าสี่กับโปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



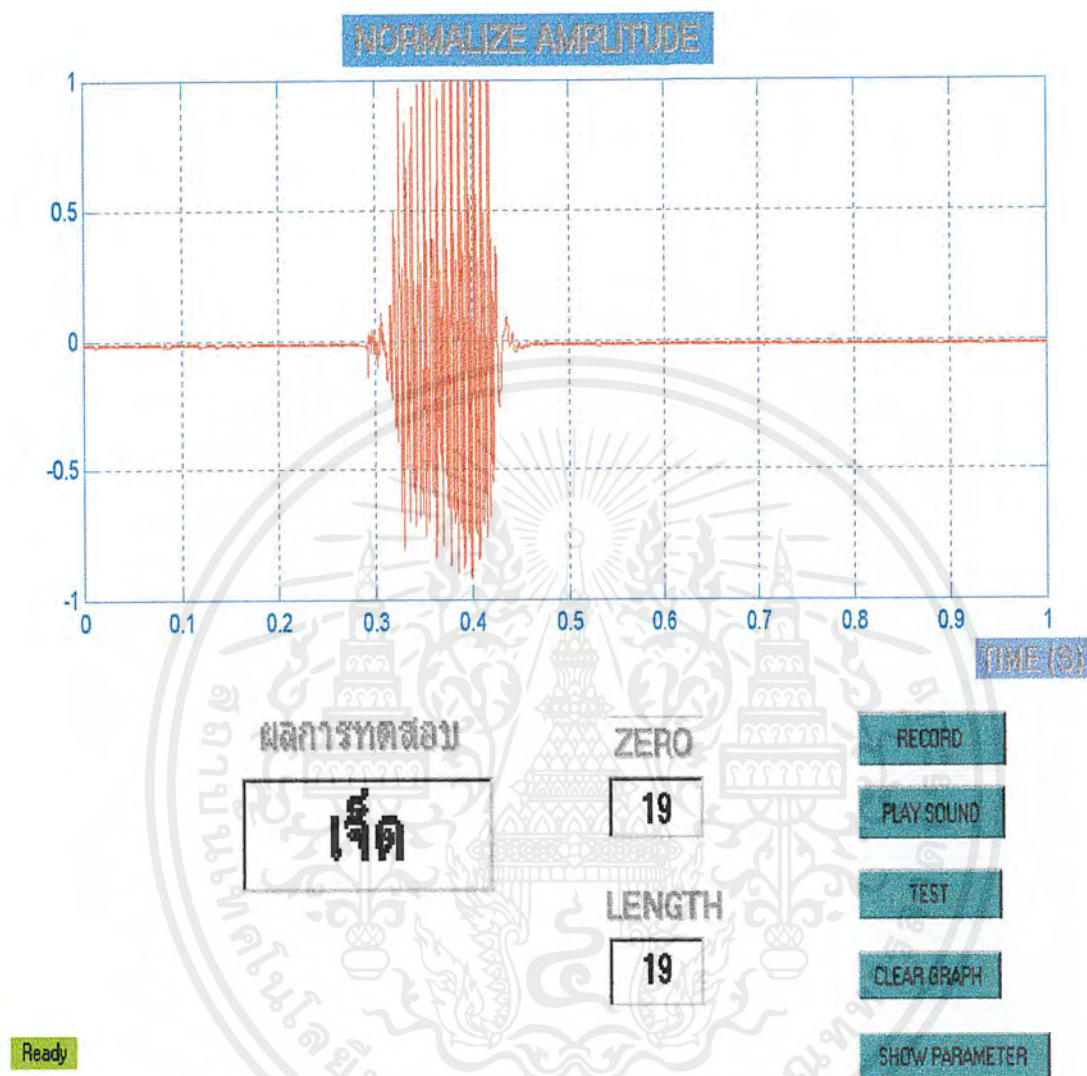
รูปที่ 4.1.6 แสดงการทดสอบความถูกต้องคำว่าห่ากับโปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



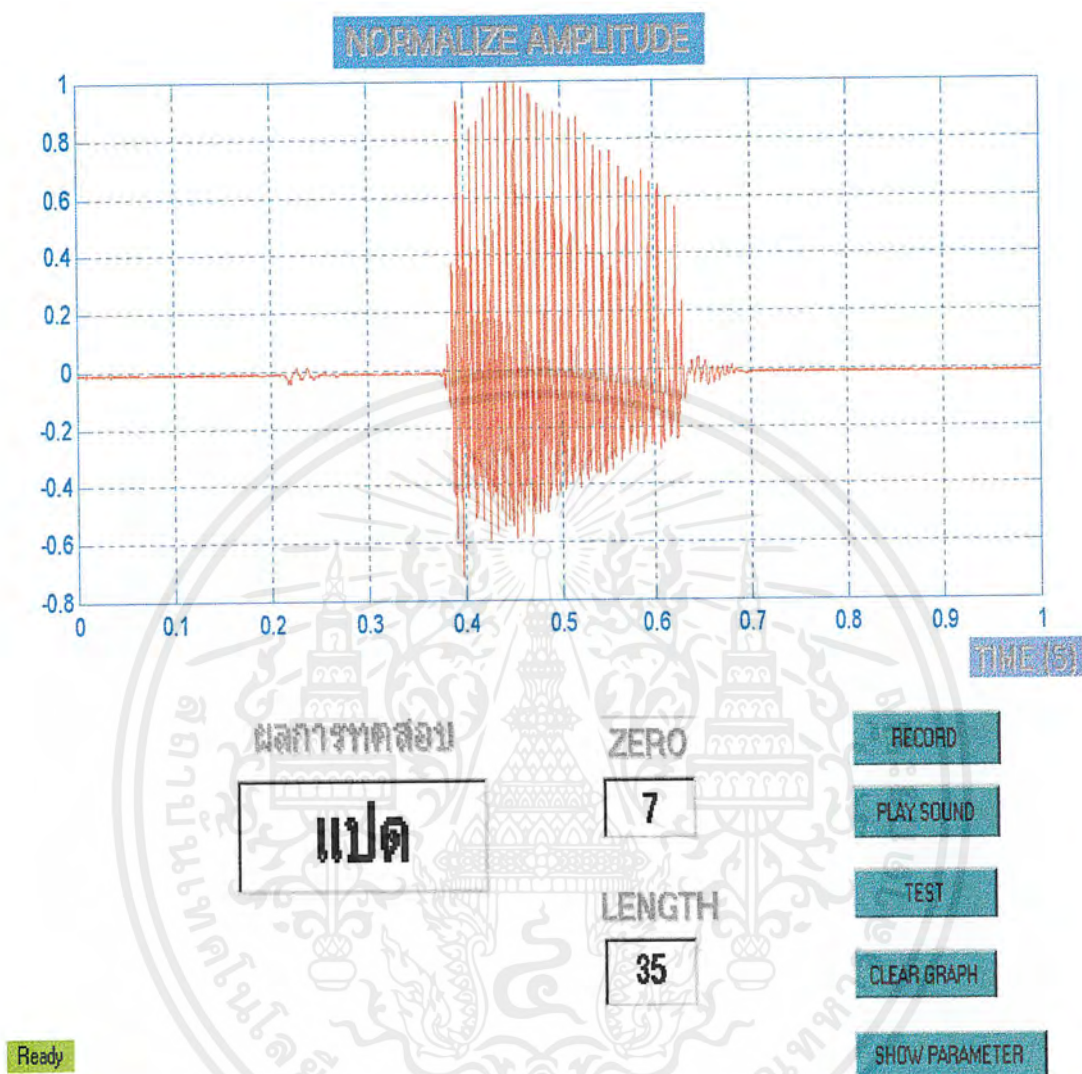
รูปที่ 4.1.7 แสดงการทดสอบความถูกต้องคำว่าหกกับโปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



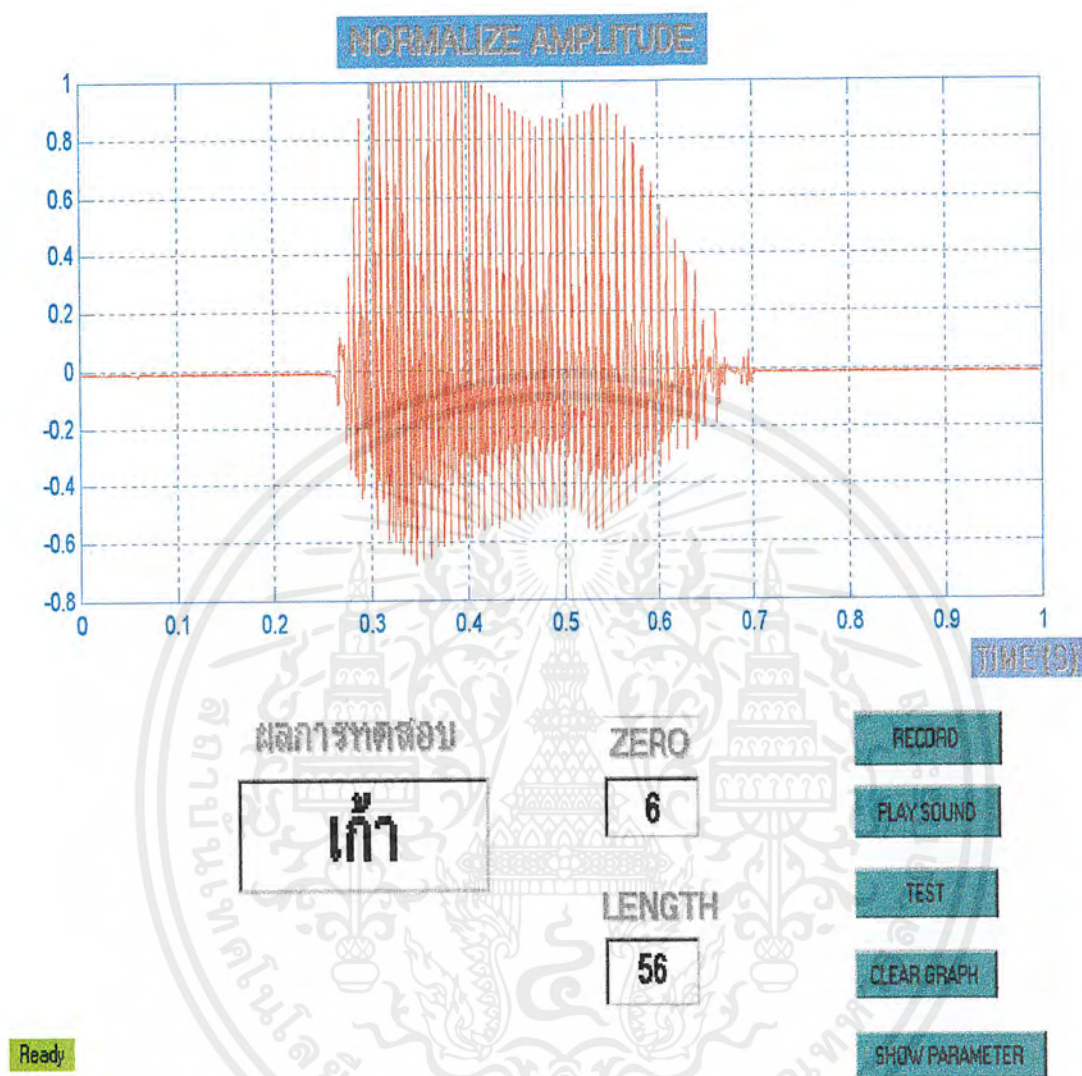
รูปที่ 4.1.8 แสดงการทดสอบความถูกต้องคำว่าเจ็ดกับโปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



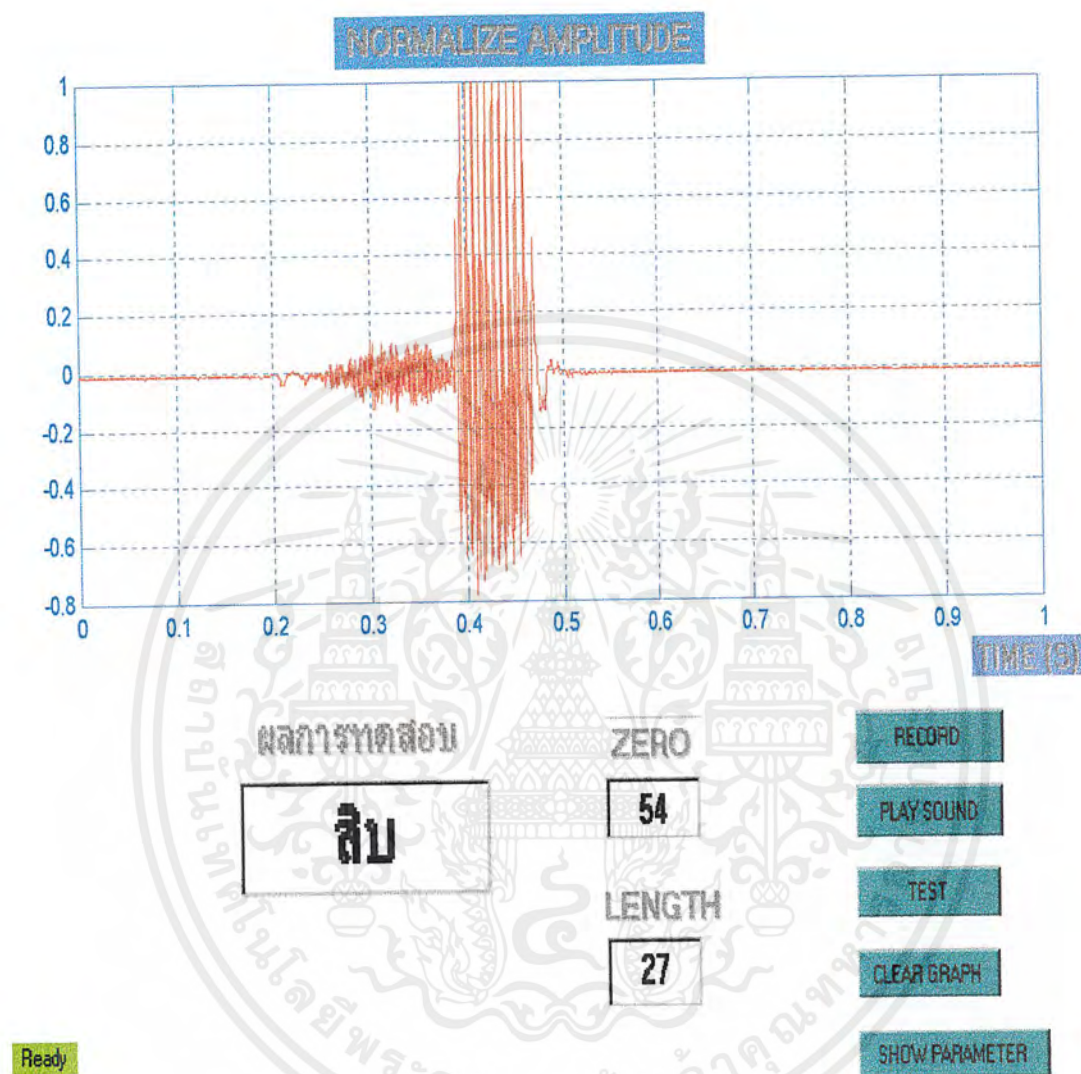
รูปที่ 4.1.9 แสดงการทดสอบความถูกต้องคำว่าแปดกับโปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



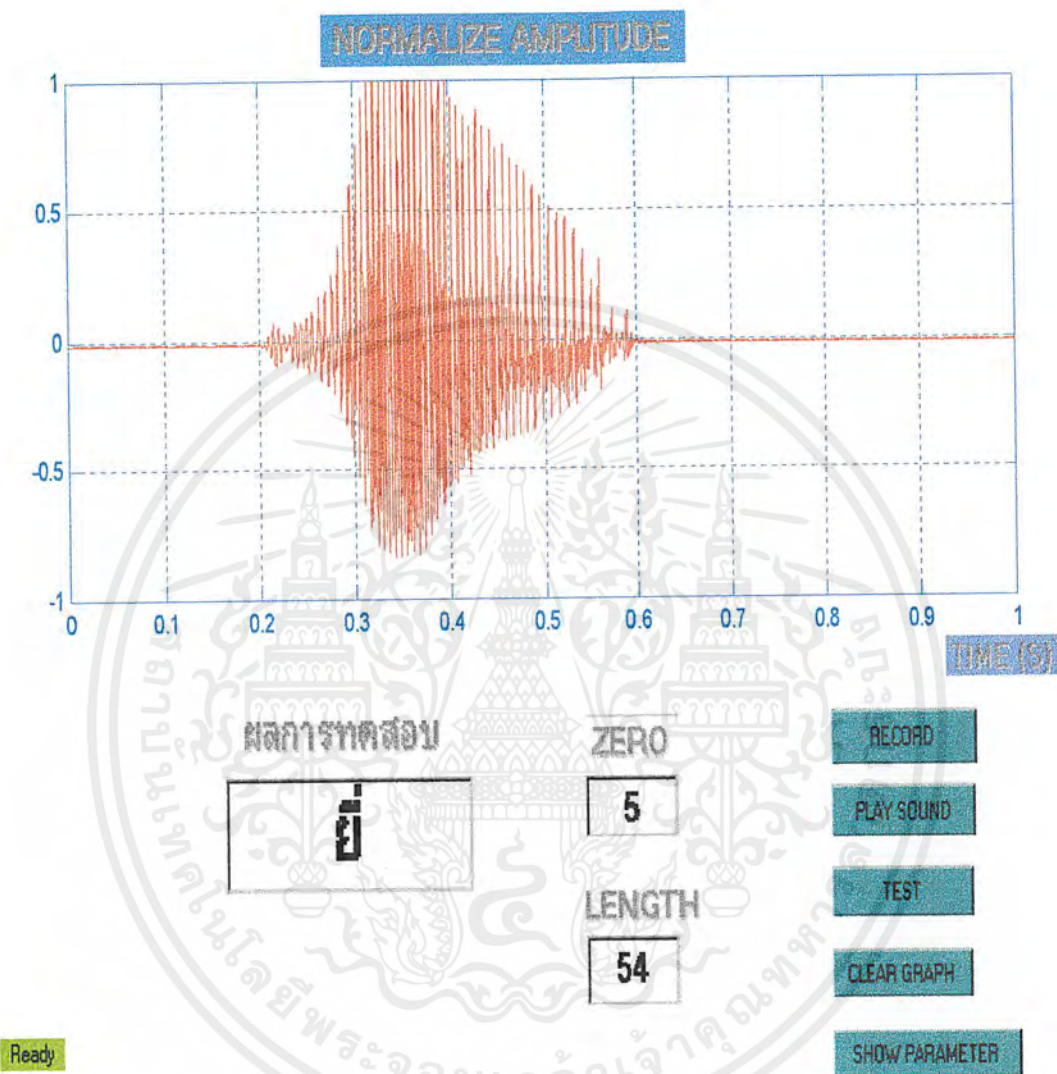
รูปที่ 4.1.10 แสดงการทดสอบความถูกต้องคำว่าเก้ากับโปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



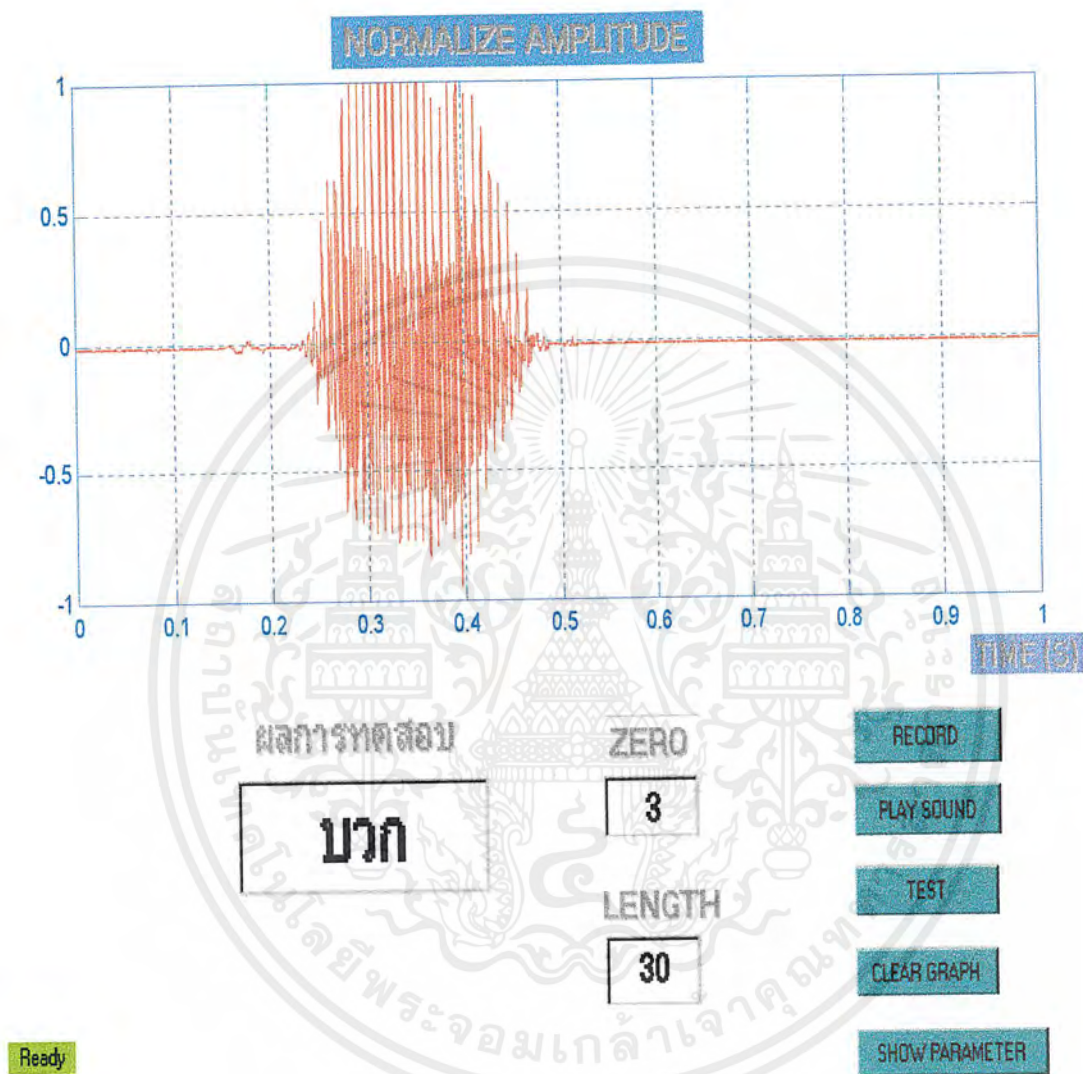
รูปที่ 4.1.11 แสดงการทดสอบความถูกต้องคำว่าสับกับโปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



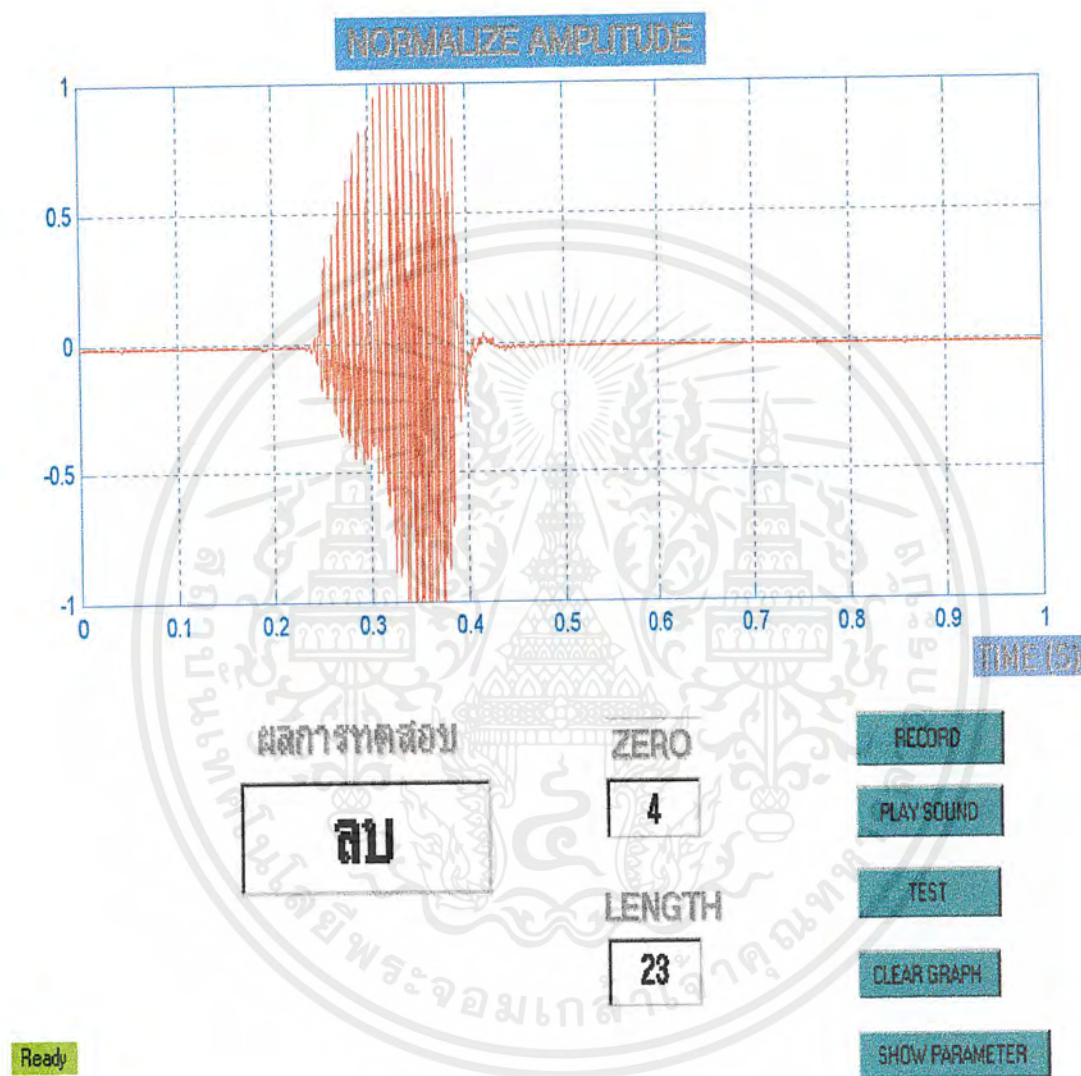
รูปที่ 4.1.12 แสดงการทดสอบความถูกต้องคำว่ายี่กับโปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



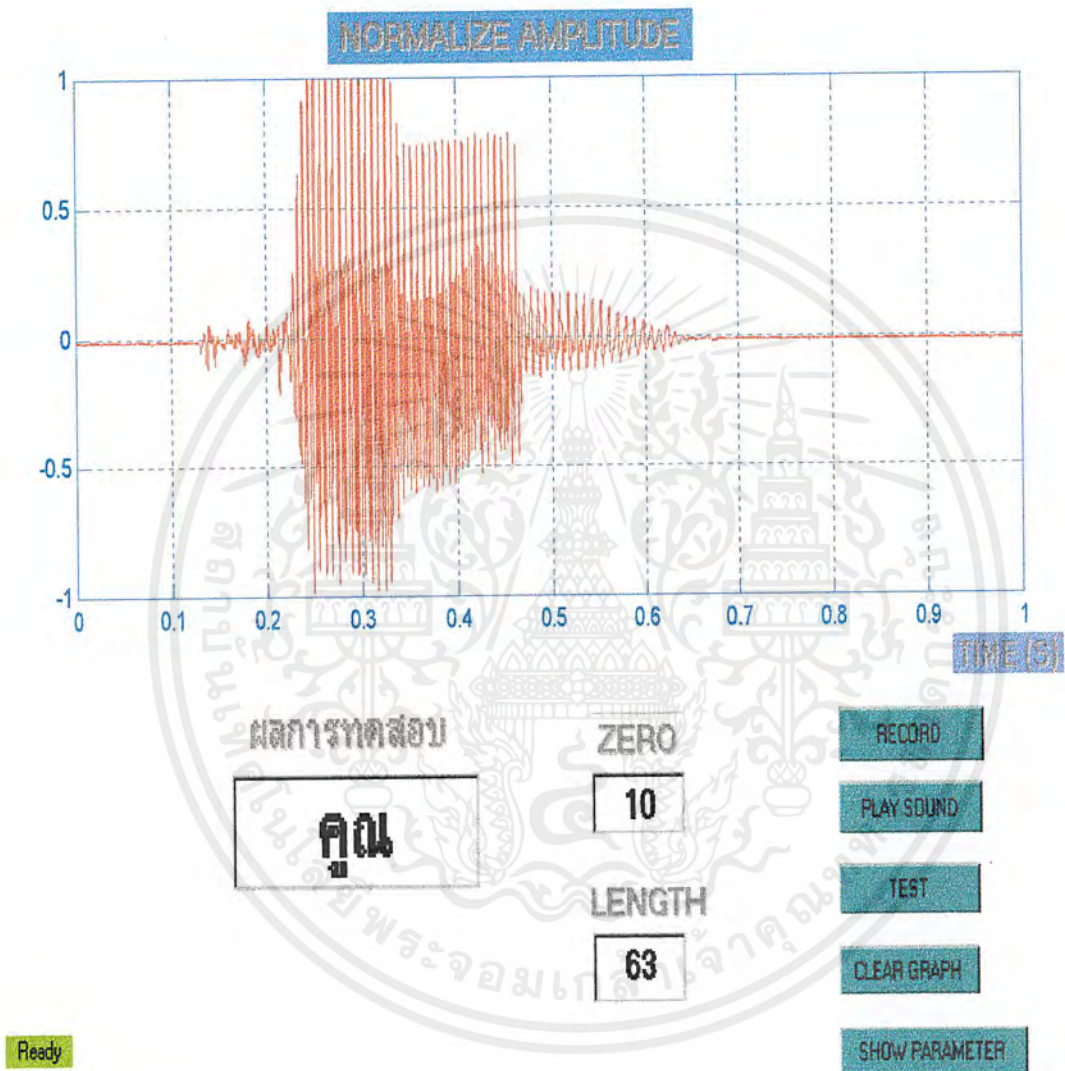
รูปที่ 4.1.13 แสดงการทดสอบความถูกต้องคำว่าบวกกับโปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



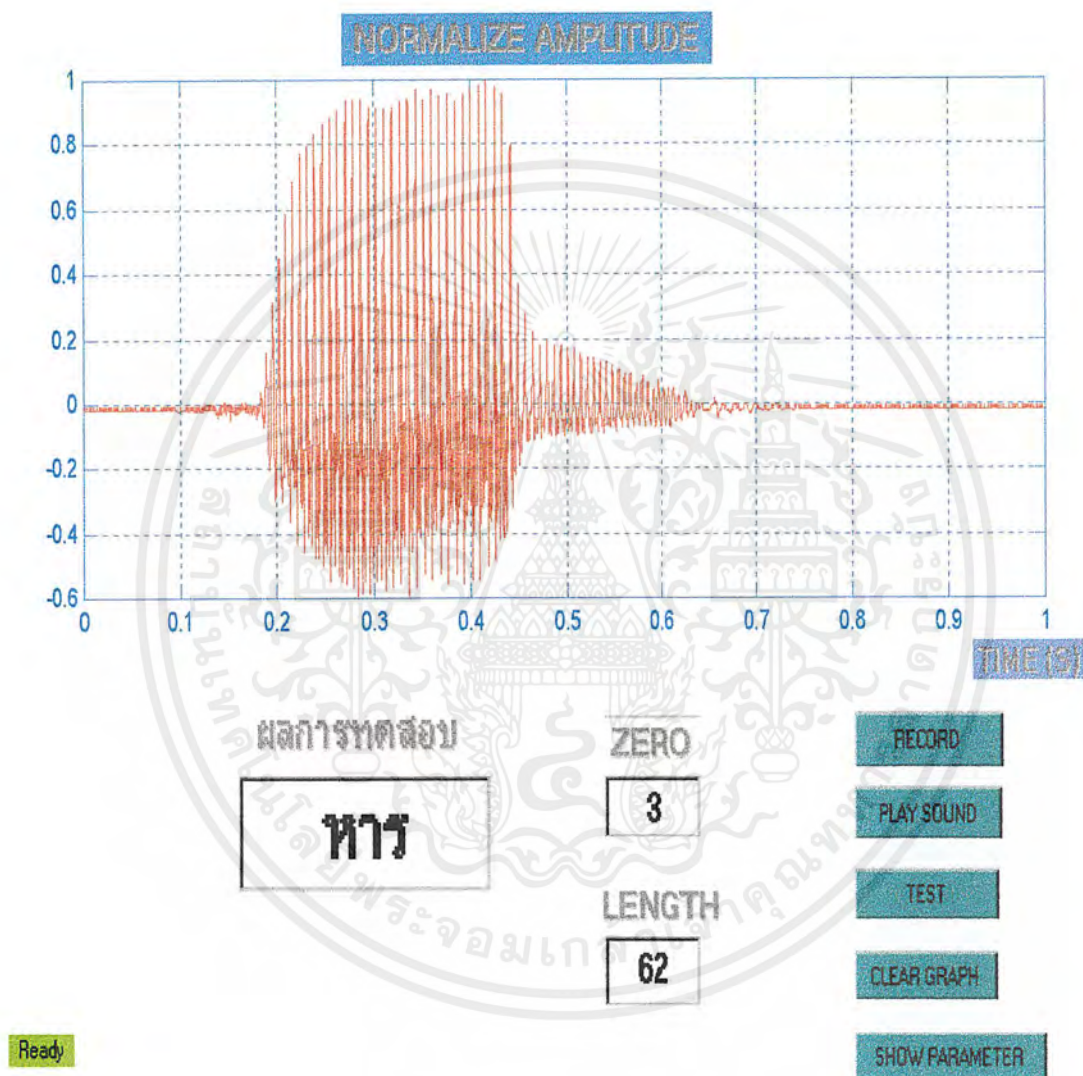
รูปที่ 4.1.14 แสดงการทดสอบความถูกต้องคำว่าลบกับโปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.1.15 แสดงการทดสอบความถูกต้องค่าว่าคูณกับ โปรแกรมใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

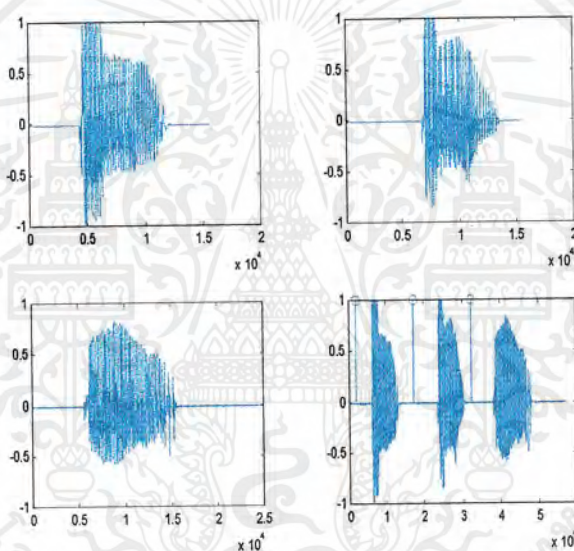


รูปที่ 4.1.16 แสดงการทดสอบความถูกต้องคำว่าหารกับโปรแกรมใช้งานจริง

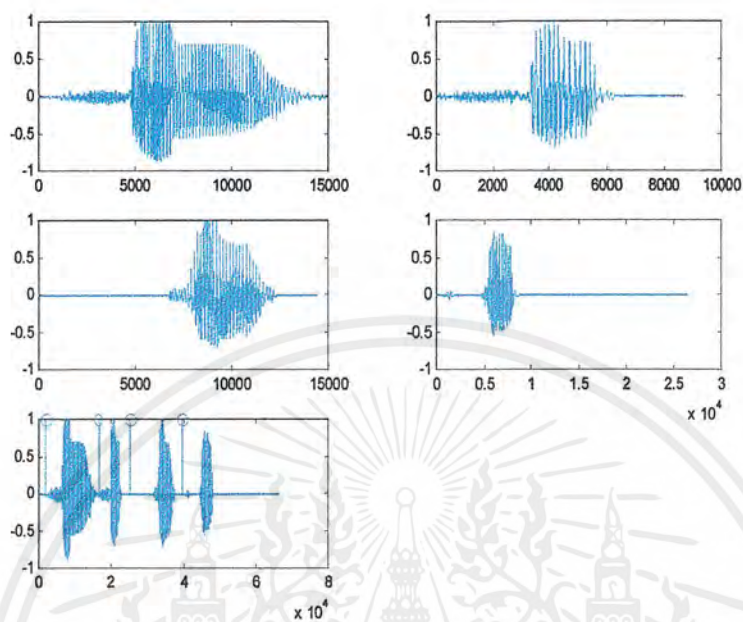
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทดสอบความถูกต้องในการแยกพยางค์ออกจากประโยค

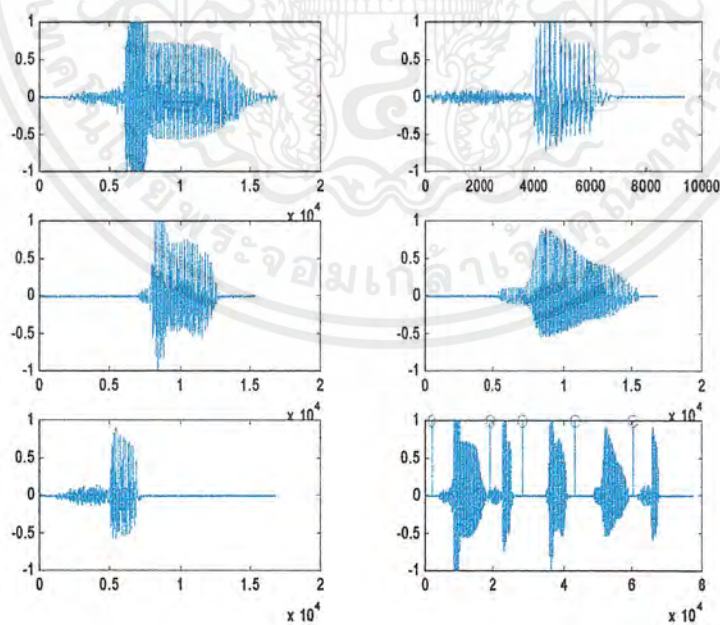
ในการทดสอบความถูกต้องในการแยกพยางค์ออกจากประโยค จะมีขั้นตอนในการทดสอบคือ ผู้ทำการทดลองจะทำการเก็บเสียงที่พูดเป็นประโยคไว้ในไฟล์เสียง แล้วทำการเรียกฟังก์ชันในการทดสอบคือฟังก์ชัน `test_cal` เพื่อทำการแยกพยางค์ออกจากประโยคซึ่งในงานวิจัยนี้จะมีจำนวนพยางค์ที่ใช้ตั้งแต่ 3 พยางค์ จนถึง 7 พยางค์ดังรูปที่ 4.2 และจากผลการทดลองในการแยกพยางค์นั้นได้ใช้ข้อมูลอย่างละ 20 ชุด ซึ่งสามารถทำการแยกพยางค์ได้อย่างถูกต้อง ดังตารางที่ 9



รูปที่ 4.2.1 แสดงการตัดพยางค์จำนวน 3 พยางค์ออกจากประโยค

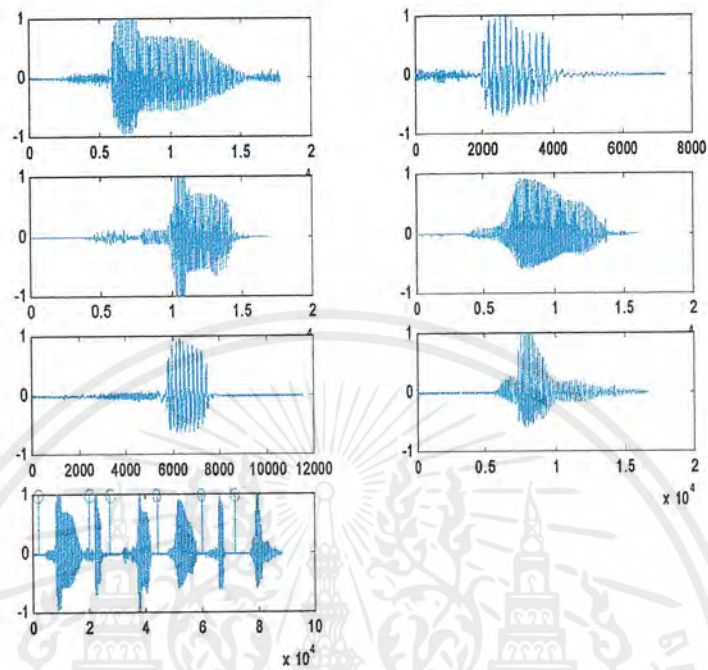


รูปที่ 4.2.2 แสดงการตัดพยางค์จำนวน 4 พยางค์ออกจากประโยค

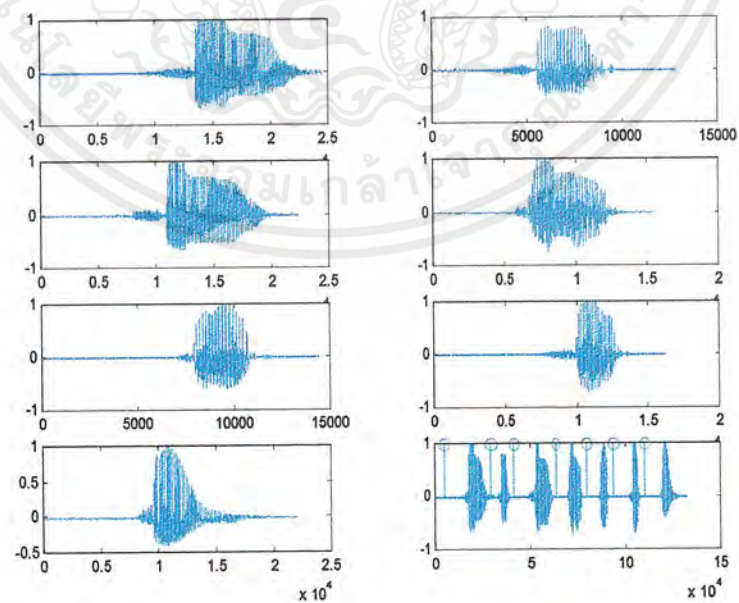


รูปที่ 4.2.3 แสดงการตัดพยางค์จำนวน 5 พยางค์ออกจากประโยค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.2.4 แสดงการตัดพยางค์จำนวน 6 พยางค์ออกจากประโยค



รูปที่ 4.2.5 แสดงการตัดพยางค์จำนวน 7 พยางค์ออกจากประโยค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผล และอภิปรายผล

สรุปผลการวิจัย

1) จากการทดลองหาอัตราการตัดสินใจเราจะได้ค่าที่เหมาะสมที่สุดคือ ค่าที่น้อยกว่าหรือเท่ากับ 46 จะพิจารณาให้อยู่ในกลุ่มของเสียงไม่เสียดสีฟัน และค่าที่มากกว่า 46 จะพิจารณาให้อยู่ในกลุ่มของเสียงเสียดสีฟัน

2) จากการทดลองหาค่าจำนวนยอดของเสียงเราจะได้ค่าที่เหมาะสมที่สุด คือ ในกลุ่มของเสียง VOICE ค่าที่มากกว่าหรือเท่ากับ 39 จะพิจารณาเป็นสระเสียงยาว ถ้าน้อยกว่า 39 จะพิจารณาเป็นสระเสียงสั้น ในกลุ่มของเสียง UNVOICE ค่าที่มากกว่า 37 จะพิจารณาเป็นสระเสียงยาว ถ้าน้อยกว่าหรือเท่ากับ 37 จะพิจารณาเป็นสระเสียงสั้น

3) จากการทดลองการแยกพยางค์ออกจากประโยคที่พูดนั้น สามารถทำการแยกพยางค์ออกจากประโยคได้อย่างถูกต้อง ซึ่งจำนวนพยางค์ที่ใช้ทั้งหมดในงานวิจัยนี้มีตั้งแต่ 3 พยางค์ จนถึง 7 พยางค์ ซึ่งเมื่อสามารถแยกพยางค์ได้แล้วก็จะนำเอาแต่ละพยางค์ไปทำการทดลองในการจดจำอีกต่อไป

4) ระบบรู้จำคำพูดแบบขึ้นตรงกับผู้พูด ผลการทดสอบในโปรแกรมใช้งานจริงค่าความถูกต้องที่มีค่ามากที่สุดอยู่ที่การใช้จำนวนชุดเสียงของฐานข้อมูลเป็นจำนวน 80 ชุด โดยมีเปอร์เซ็นต์ความถูกต้องเท่ากับ 93.75%

5) การเพิ่มจำนวนชุดเสียงของฐานข้อมูลให้มากขึ้น มีผลทำให้การตัดสินใจได้ความถูกต้องเพิ่มขึ้น โดยดูจากจำนวนความถูกต้องที่เพิ่มขึ้น

6) ในการประมวลผลผลลัพธ์ ผลลัพธ์ที่ได้ จะทำการแสดงค่าเป็นตัวเลข และมีเสียงพูดออกมา ซึ่งเป็นเสียงของผู้ทำงานวิจัยเอง ผลลัพธ์ที่ได้สามารถมีค่าคิดลบออกมาได้และยังสามารถแสดงเป็นเลขทศนิยม 6 หลัก ได้อีกด้วย ถ้าหากการประมวลผลไม่เจอตัวกระทำการหรือมีตัวกระทำการมากกว่า 1 ตัว ก็จะมีเสียงเตือนบอกออกมาว่าไม่สามารถแสดงผลได้เนื่องจากอะไร

อภิปรายความผิดพลาด

1) คำบางคำที่ใช้ในการทดลองจะมีค่าพารามิเตอร์ที่ไม่ชัดเจน ได้แก่ เสียง เจ็ด เนื่องจากจากมีค่าอัตราการตัดศูนย์ อยู่ในช่วงทั้งเสียงไม่เสียดสีฟัน และเสียงเสียดสีฟัน และเสียงบวก มีค่าจำนวนขอดของเสียง อยู่ในช่วงทั้งเสียงสระเสียงสั้น และเสียงสระเสียงยาว เป็นผลให้ภาคการตัดสินใจมีความผิดในค่าที่มีค่าพารามิเตอร์ที่ไม่ชัดเจน ดังนั้นเสียงคำพูดที่มีค่าพารามิเตอร์ไม่ชัดเจน อาจทำให้ระบบรู้จำผิดพลาดได้

2) ในการทดลองการประมวลผลผลลัพธ์จะมีการจำผิดพลาดของคำบางคำ คือ คำว่า แก้ว กับ คำว่า ห้า เนื่องจากว่ามีการออกเสียงสระที่คล้ายกันคือ สระอา จึงทำให้การจดจำเกิดการผิดพลาดได้ และคำว่าบวก กับ คูณ เนื่องจากทั้งสองคำนี้อยู่ในกลุ่มเดียวกัน และมีค่าใกล้เคียงกันจึงทำให้เกิดการผิดพลาด

3) การออกเสียงไม่ชัดเจนของผู้วิจัยอาจเป็นสาเหตุ ทำให้ระบบการรู้จำตัดสินใจผิดพลาด ข้อผิดพลาดนี้จะ ต้องแก้ไขที่ผู้ออกเสียงเอง ไม่จำเป็นต้องแก้ไขระบบ

ข้อเสนอแนะ

1) ในการทดลองจะต้องทำการทดลองในสถานที่ที่เงียบ ไม่ควรจะมีเสียงอื่นมารบกวน เพราะอาจทำให้เกิดความผิดพลาดได้

2) การเพิ่มความถูกต้องให้กับภาคการจดจำเสียงนั้น สามารถทำได้โดยการเพิ่มจำนวนชุดเสียงของฐานข้อมูล ให้มากขึ้นซึ่งจะทำให้ความถูกต้องมีค่าเพิ่มขึ้น

3) สำหรับภาคการรู้จำเสียงคำพูดในงานวิจัยนี้นั้น ผู้วิจัยได้ทำการคิดพารามิเตอร์ในการทดลองการจดจำเสียงขึ้นมาเอง ไม่ได้ทำการจดจำด้วยวิธีอื่นเช่น การสอนนิรอนเนคเวอร์ค เป็นต้น ซึ่งอาจจะยังไม่ใช่วิธีที่ดีที่สุด

การประยุกต์ใช้งาน

งานวิจัยนี้สามารถนำไปประยุกต์ใช้งานด้านต่างๆ เช่น ระบบสั่งงานเครื่องมือด้วยเสียงพูด ระบบติดต่อกับโปรแกรมด้วยเสียงพูด ระบบโทรศัพท์อัตโนมัติ เป็นต้น

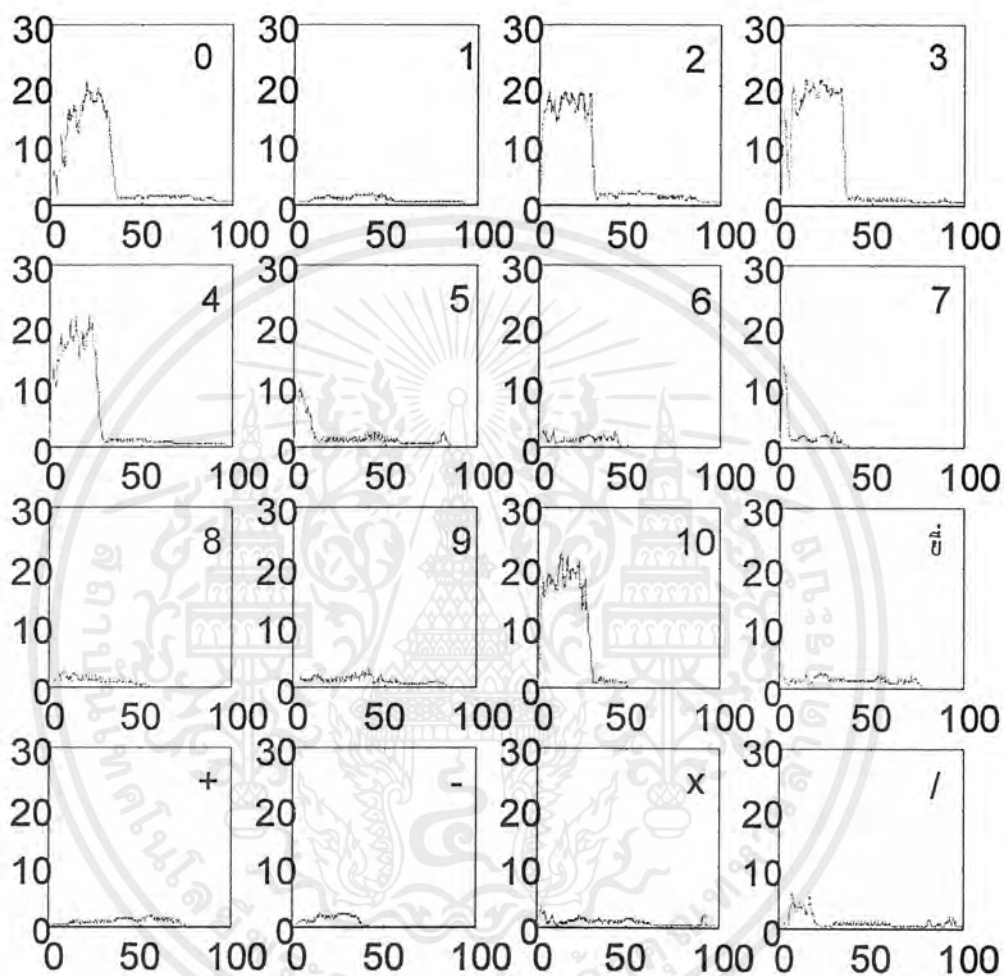
เอกสารอ้างอิง

- 1) ไชยันต์ สุวรรณชีวะศิริ “ การรู้จำเสียงพูดตัวเลขภาษาไทยแบบขึ้นกับผู้พูดด้วยนิวรัลเน็ตเวิร์ค ” วิทยานิพนธ์ปริญญาโทบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2539.
- 2) ณัฐกร ทับทอง “ การรู้จำเสียงคำพูดภาษาไทย โดยใช้ลักษณะบ่งความต่างของหน่วยเสียง ” วิทยานิพนธ์ปริญญาโทบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2538
- 3) กาญจนา นาคสกุล “ ระบบเสียงภาษาไทย ” กรุงเทพฯ , โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย , 2524
- 4) ชลธิ จันทรศิริวิไลกุล และนายณพคุณ คุณสมบัติ “ เครื่องคิดเลขที่ควบคุมด้วยเสียงพูด ” ปริญญานิพนธ์ ภาควิชาวิศวกรรมไฟฟ้า สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ , 2540
- 5) Rabiner, Lawrence R. and Schafer , Ronald W. Digital Processing of Speech Signals., USA.: Prentice-Hall,Inc. ,1978.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

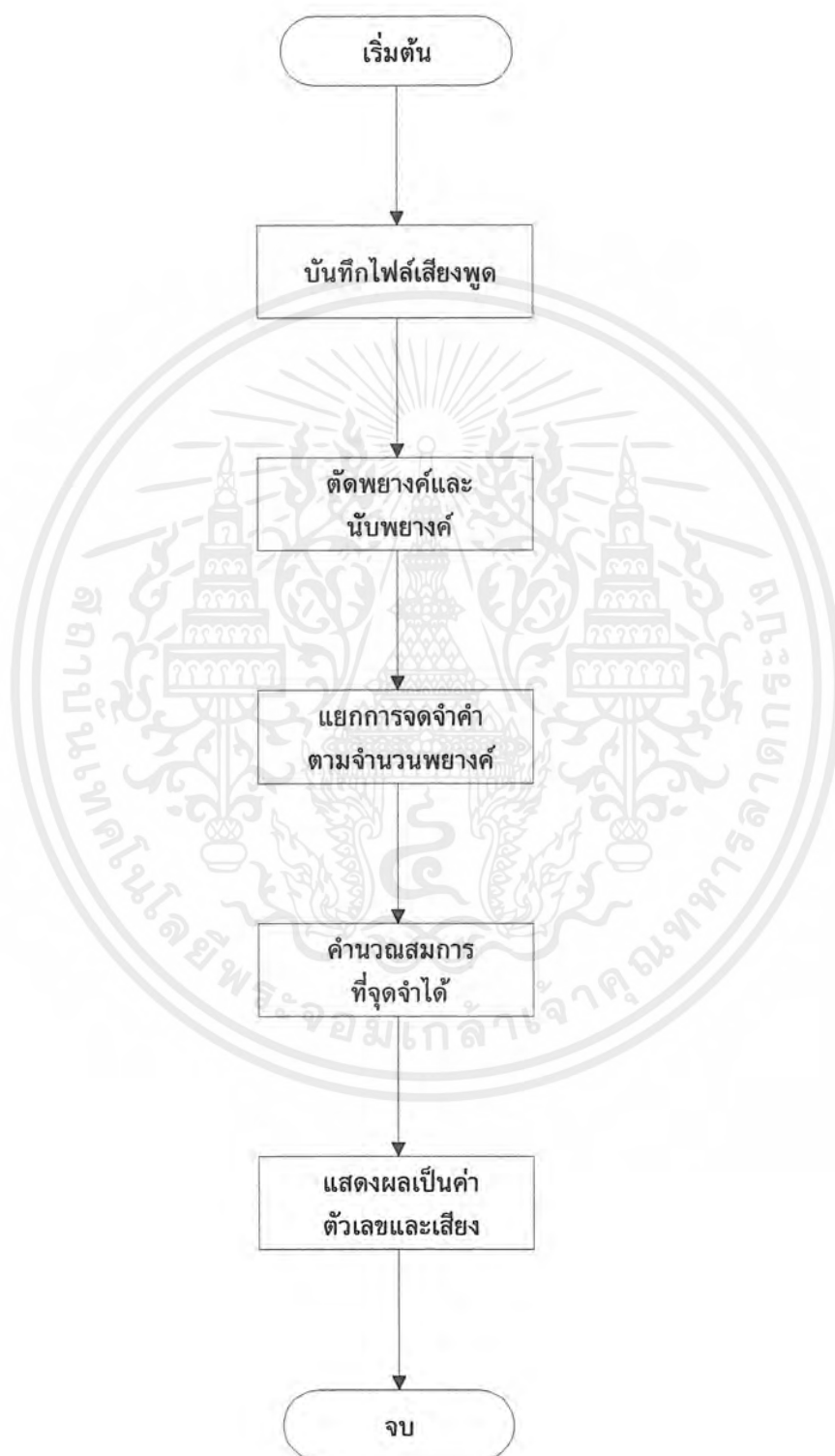
ภาคผนวก ก



รูปที่ 1 อัตราการตัดศูนย์ของเสียงที่ใช้ในการวิจัย

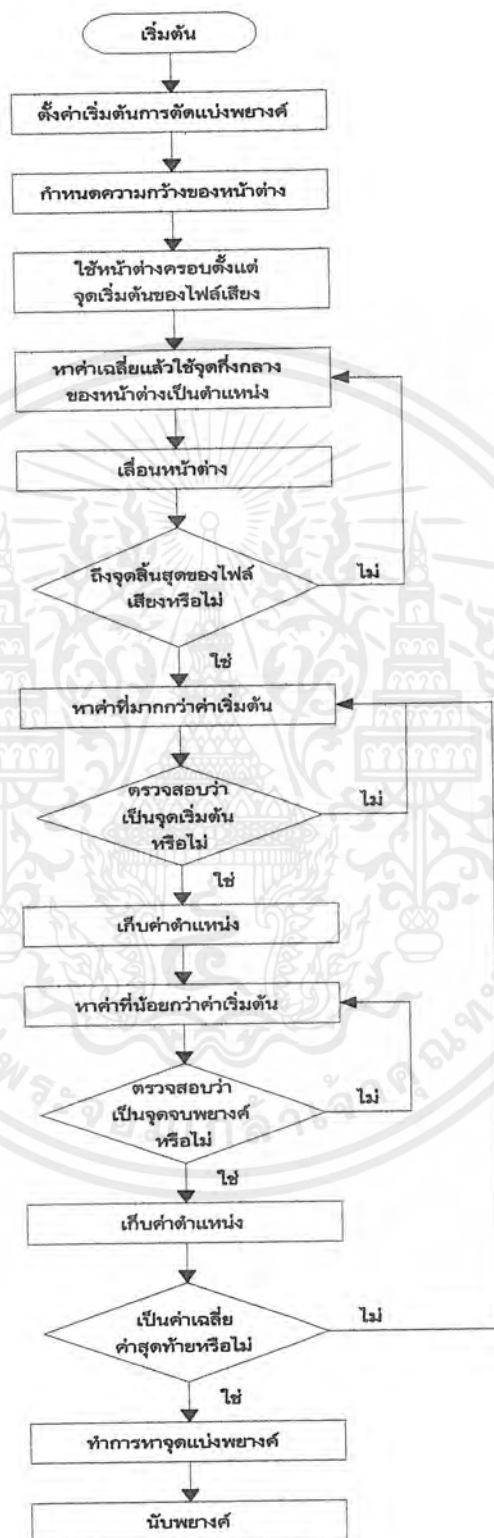
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข

**รูปที่ 2** อัลกอริทึมที่ใช้ในการทำเครื่องคำนวณด้วยเสียงแบบคำโดด

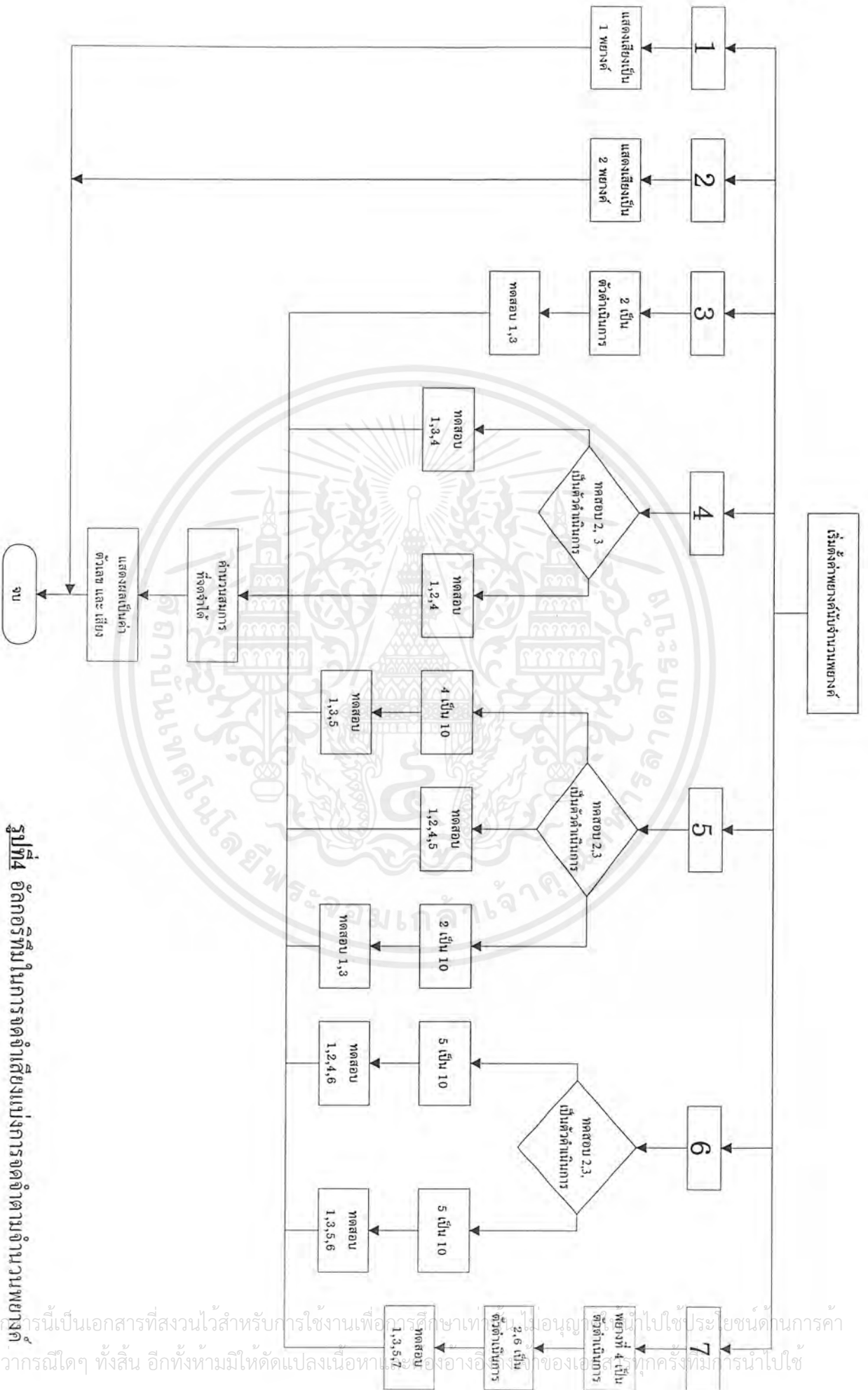
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค



รูปที่ 3 อัลกอริทึมในการตัดแบ่งพยางค์และนับจำนวนพยางค์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4 อัลกอริทึมในการจัดจำเสียงแบ่งการจัดจำตามจำนวนพยางค์

ตารางที่ 1 แสดงการแบ่งเสียงออกเป็นเสียงเสียดสีฟันและเสียงไม่เสียดสีฟัน

เสียงเสียดสีฟัน	เสียงไม่เสียดสีฟัน
ศูนย์	หนึ่ง
สอง	ห้า
สาม	หก
สี่	เจ็ด
	แปด
	เก้า
	ยี่
	บวก
	ลบ
	คูณ
	หาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2 แสดงการแบ่งเสียงออกเป็นสระเสียงสั้นและสระเสียงยาว

เสียงไม่เสียดสีฟัน		เสียงเสียดสีฟัน	
สระเสียงสั้น	สระเสียงยาว	สระเสียงสั้น	สระเสียงยาว
หก	หนึ่ง	สิบ	ศูนย์
เจ็ด	ห้า		สอง
แปด	เก้า		สาม
บวก	ยี่		สี่
ลบ	คูณ		
	หาร		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3 แสดงผลการทดลองการรู้จำเสียง

เสียง	10	15	20	25	30
ศูนย์	*	สี่	*	สี่	สี่
หนึ่ง	*	*	คูณ	*	คูณ
สอง	*	*	*	*	*
สาม	สอง	สอง	*	*	สอง
สี่	*	*	ยี่	*	*
ห้า	*	*	*	*	*
หก	*	*	ลบ	*	*
เจ็ด	*	*	*	*	*
แปด	หาร	หาร	เก้า	หาร	เก้า
เก้า	ห้า	ห้า	*	*	ห้า
สิบ	เจ็ด	เจ็ด	เจ็ด	เจ็ด	เจ็ด
ยี่	*	*	*	*	*
บวก	หนึ่ง	คูณ	คูณ	คูณ	คูณ
ลบ	*	*	*	*	*
คูณ	*	*	*	*	*
หาร	*	*	หนึ่ง	*	หนึ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3 แสดงผลการทดลองการรู้จำเสียง

เสียง	35	40	45	50	55
ศูนย์	สี่	*	*	*	*
หนึ่ง	คุณ	คุณ	*	*	*
สอง	*	*	*	*	*
สาม	*	*	*	*	*
สี่	*	*	*	*	*
ห้า	*	*	*	*	*
หก	*	*	*	หาร	*
เจ็ด	*	*	*	*	*
แปด	เก้า	เก้า	*	*	*
เก้า	ห้า	ห้า	ห้า	*	ห้า
สิบ	*	*	*	*	*
ยี่	*	*	*	*	*
บวก	คุณ	คุณ	คุณ	*	คุณ
ลบ	*	*	*	*	*
คูณ	*	*	*	*	*
หาร	*	*	*	*	*

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3 แสดงผลการทดลองการรู้จำเสียง

เสียง	60	65	70	75	80
ศูนย์	*	*	*	*	*
หนึ่ง	*	*	*	*	*
สอง	*	*	*	*	*
สาม	*	*	*	*	*
สี่	*	*	*	*	*
ห้า	*	*	*	*	*
หก	*	*	*	*	*
เจ็ด	*	*	*	*	*
แปด	*	*	*	*	*
เก้า	*	ห้า	*	*	*
สิบ	*	เจ็ด	*	เจ็ด	*
ยี่	*	*	*	*	*
บวก	*	*	*	*	หก
ลบ	*	*	*	*	*
คูณ	*	*	*	*	*
หาร	*	*	*	*	*

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4 การหาค่าจำนวนพิทช์เฉลี่ยของเสียงไม่เสียดสีฟัน

เสียงไม่เสียดสีฟัน	ค่าจำนวนพิทช์เฉลี่ย	การแบ่งประเภทเสียง
หนึ่ง	59	สระเสียงยาว
ห้า	54	สระเสียงยาว
หก	28	สระเสียงสั้น
เจ็ด	20	สระเสียงสั้น
แปด	34	สระเสียงสั้น
เก้า	51	สระเสียงยาว
ยี่	60	สระเสียงยาว
บวก	42	สระเสียงยาว
ลบ	29	สระเสียงสั้น
คูณ	66	สระเสียงยาว
หาร	63	สระเสียงยาว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5 การหาค่าจำนวนพิทช์เฉลี่ยของเสียงเสียดสีฟัน

เสียงเสียดสีฟัน	ค่าจำนวนพิทช์เฉลี่ย	การแบ่งประเภทเสียง
ศูนย์	69	สระเสียงยาว
สอง	74	สระเสียงยาว
สาม	65	สระเสียงยาว
สี่	66	สระเสียงยาว
สิบ	30	สระเสียงสั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 6 แสดงค่าหาจำนวนพิพจน์ของเสียงจำนวน 20 ชุด

จำนวนชุด	ศูนย์	หนึ่ง	สอง	สาม	สี่	ห้า	หก	เจ็ด
1	54	55	58	59	59	44	25	17
2	59	54	62	67	58	50	24	19
3	55	59	60	54	50	50	26	19
4	66	60	62	55	60	58	28	20
5	67	57	67	65	53	53	24	21
6	63	56	60	65	54	49	26	20
7	58	55	63	67	62	55	29	18
8	54	49	61	61	64	47	26	19
9	54	59	55	61	61	47	22	18
10	58	56	61	64	60	51	26	18
11	59	57	62	61	66	54	29	22
12	62	63	68	57	62	55	25	19
13	69	62	70	63	69	55	33	19
14	65	59	70	68	63	49	26	18
15	67	63	80	70	57	51	32	19
16	71	80	75	71	65	57	32	17
17	75	66	77	76	69	52	27	17
18	75	73	69	67	70	55	28	20
19	71	74	78	76	65	49	33	19
20	73	61	72	64	56	55	34	20

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 6 แสดงค่าหาจำนวนพิพจน์ของเสียงจำนวน 20 ชุด

จำนวนชุด	แปด	เก้า	สิบ	ยี่	บวก	ลบ	คูณ	หาร
1	32	52	25	62	31	27	54	65
2	31	52	30	57	38	26	68	49
3	31	46	22	53	32	21	57	48
4	31	47	26	59	37	23	59	53
5	31	49	24	61	37	27	62	58
6	32	51	27	60	32	25	57	61
7	32	50	31	61	36	30	60	51
8	33	54	30	56	32	27	61	60
9	29	42	28	52	31	21	59	52
10	29	50	30	57	35	27	54	48
11	34	54	24	60	34	26	61	59
12	36	48	27	60	39	24	62	65
13	30	48	29	69	29	30	67	74
14	32	50	27	55	35	21	62	67
15	33	49	26	55	34	29	82	71
16	34	54	33	53	41	25	70	67
17	33	46	33	58	39	30	64	66
18	38	52	24	61	37	27	69	63
19	32	48	33	59	34	24	72	76
20	32	51	35	58	36	24	84	78

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 7 แสดงค่าอัตราการตัดศูนย์จำนวนเสียง 20 ชุด

จำนวนชุด	ศูนย์	หนึ่ง	สอง	สาม	สี่	ห้า	หก	เจ็ด
1	74.3582	1.4188	56.4283	69.6587	70.563	31.2786	9.4399	18.8514
2	66.8514	4.7001	80.4935	81.6567	75.8879	23.6782	4.5693	43.7865
3	56.1966	3.1863	69.5191	58.6093	73.5042	32.4415	7.7138	32.4881
4	63.6016	4.5268	75.799	69.0766	85.8133	28.7745	16.4538	68.0125
5	65.6671	3.9651	69.7597	78.2506	74.7888	28.5184	22.1087	56.8987
6	54.0217	1.3896	73.5972	70.9577	67.314	14.5179	15.634	11.9831
7	67.8376	3.8277	73.0274	70.8929	74.2031	24.4608	14.9664	31.5103
8	52.5177	3.135	46.6288	76.1462	67.5888	17.297	7.3438	16.028
9	70.741	4.6619	89.9538	67.8147	86.9359	13.9937	11.8282	48.0839
10	76.6011	3.2599	69.3299	81.5154	55.1101	8.3911	11.3474	37.5413
11	75.9295	3.2679	81.3749	75.4962	83.4635	25.1154	9.9755	46.4309
12	70.2559	2.4381	69.824	79.325	85.7064	23.2324	6.7062	38.722
13	77.3597	3.1887	75.2654	77.6593	94.393	24.8482	4.0449	50.4852
14	66.2986	3.9272	73.9986	82.8858	65.2504	45.5235	10.315	42.3019
15	58.8733	3.9766	80.7823	89.2494	74.8901	41.3461	10.1251	35.9939
16	64.2078	5.9198	62.2299	68.0124	81.8512	22.8659	7.0935	24.7345
17	54.7289	1.7207	80.4616	76.6119	72.2539	18.1557	6.0784	50.5154
18	58.0129	2.5467	72.6505	66.4754	67.0612	12.7883	4.7782	47.5154
19	76.7266	3.1573	79.2718	74.7751	71.4796	32.3495	15.735	41.5146
20	52.8081	1.4875	72.0087	86.3718	75.3433	33.9842	7.6385	40.2629

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 7 แสดงค่าอัตราการจัดศูนย์จำนวนเสียง 20 ชุด

จำนวนชุด	แปด	เก้า	สิบ	ยี่	บวก	ลบ	คูณ	หาร
1	8.6484	6.3142	68.1417	4.6322	3.0694	3.9604	7.8393	10.3561
2	7.7918	6.9653	50.2669	5.4555	3.7657	4.878	4.3388	13.8361
3	7.4472	7.6648	57.8117	5.4141	7.3524	3.8162	6.8658	7.5273
4	7.8309	6.6963	57.5672	4.9587	2.9819	3.2657	5.928	5.5533
5	6.5493	8.0058	53.6814	5.424	3.024	1.8515	6.6038	14.533
6	6.8803	8.1721	47.5721	3.953	3.2685	4.3984	4.4767	14.804
7	7.5852	8.3582	51.6055	5.1383	3.146	3.2516	9.1146	19.4337
8	4.4264	6.3478	44.6503	4.5257	2.9733	2.4135	6.2125	26.7306
9	6.7127	5.9456	57.2094	4.3976	2.7421	3.275	6.6699	18.1916
10	6.7081	9.4766	54.0337	2.7934	3.1778	2.4323	4.6227	8.1555
11	6.4465	6.1695	51.6618	4.3811	3.6218	3.5032	5.2674	14.1063
12	7.852	5.9823	59.274	5.3303	2.5638	5.1694	4.3722	20.6393
13	6.3723	10.9193	47.3345	9.0597	3.1541	4.8465	4.9004	12.5208
14	8.4572	9.7275	49.7201	7.9416	2.9241	6.7136	7.6484	20.2945
15	8.0066	7.5895	50.8627	5.9133	3.2241	3.8293	5.899	16.645
16	7.6009	5.5247	48.4547	5.2802	2.3182	4.0377	4.9663	17.8708
17	7.9215	7.6453	51.0983	9.4685	5.691	4.449	3.6534	19.1852
18	7.9569	7.318	58.8206	9.7972	3.8512	4.0378	5.6233	25.9029
19	7.4282	7.4531	54.9021	5.4755	3.3144	6.1031	5.6257	33.0557
20	9.128	11.1238	50.4253	7.2076	2.9433	4.6431	5.3192	15.6998

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 8.1 ผลการทดลองความถูกต้องในการประมวลผลลัพธ์จากข้อมูล 3 พยางค์

ข้อมูล	ตัวตั้ง	ตัวกระทำ	ตัวตาม	ผลลัพธ์	ค่าความถูกต้อง
1+2	1	+	2	3	ถูก
3x9	2	x	9	18	ผิด
7-5	7	-	5	2	ถูก
8/4	8	/	4	2	ถูก
3+7	3	+	7	10	ถูก
5-1	5	-	1	4	ถูก
8x9	9	x	9	81	ผิด
10/5	10	/	5	2	ถูก
6+3	6	+	3	9	ถูก
2x10	2	x	10	20	ถูก
4+5	4	+	5	9	ถูก
9/3	9	/	3	3	ถูก
7-4	7	-	4	3	ถูก
10x5	7	x	5	50	ถูก
6+1	6	+	1	7	ถูก
9-5	9	-	5	4	ถูก
10+7	10	+	7	17	ถูก
3x3	3	x	3	9	ถูก
5/5	5	/	5	1	ถูก
4/2	4	/	2	2	ถูก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 8.2 ผลการทดลองความถูกต้องในการประมวลผลลัพธ์จากข้อมูล 4 พยางค์

ข้อมูล	ตัวตั้ง	ตัวกระทำ	ตัวตาม	ผลลัพธ์	ค่าความถูกต้อง
11+0	11	+	0	11	ถูก
12+2	12	+	1	13	ถูก
13-2	13	-	2	11	ถูก
14x3	14	x	3	21	ถูก
15/4	15	/	4	3.75	ถูก
16+5	16	+	5	21	ถูก
17-6	17	-	6	11	ถูก
18x7	15	x	7	105	ผิด
19/8	19	/	8	2.375	ถูก
20+9	20	+	5	25	ถูก
30+10	30	+	10	40	ถูก
40-1	40	-	1	39	ถูก
50x5	50	x	5	250	ถูก
60/10	60	/	10	6	ถูก
70+7	70	x	7	4900	ผิด
80-9	80	-	9	71	ถูก
90x2	50	x	2	100	ผิด
5x30	5	+	30	80	ถูก
7-19	7	-	19	-12	ถูก
8+60	8	+	60	68	ถูก

ตารางที่ 8.3 ผลการทดลองความถูกต้องในการประมวลผลลัพธ์จากข้อมูล 5 พยางค์

ข้อมูล	ตัวตั้ง	ตัวกระทำ	ตัวตาม	ผลลัพธ์	ค่าความถูกต้อง
71+1	71	+	1	71	ถูก
23-10	23	-	10	13	ถูก
35/5	35	/	5	7	ถูก
42x4	42	x	4	168	ถูก
57+3	57	+	3	60	ถูก
69-7	65	-	7	58	ผิด
72/6	72	/	6	12	ถูก
84x8	84	x	8	672	ถูก
99+1	99	+	1	100	ถูก
95x5	99	x	5	495	ผิด
9+27	5	x	27	135	ผิด
8-36	9	-	36	27	ผิด
7x45	7	x	45	315	ถูก
6/54	6	/	54	0.111111	ถูก
5+68	5	+	68	73	ถูก
4-71	4	-	71	-67	ถูก
3x88	3	x	88	264	ถูก
2/92	2	/	52	0.038462	ผิด
43+7	43	+	7	50	ถูก
0-37	0	-	37	-37	ถูก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 8.4 ผลการทดลองความถูกต้องในการประมวลผลลัพธ์จากข้อมูล 6 พยางค์

ข้อมูล	ตัวตั้ง	ตัวกระทำ	ตัวตาม	ผลลัพธ์	ค่าความถูกต้อง
17+25	17	+	25	42	ถูก
20-39	20	-	39	19	ถูก
30x41	30	x	41	1230	ถูก
40/29	40	/	29	1.37931	ถูก
50+53	50	+	53	103	ถูก
60+64	60	+	64	124	ถูก
70x79	70	+	75	145	ผิด
80/84	80	/	84	0.95238	ถูก
90+99	50	+	99	149	ผิด
15-24	15	-	24	9	ถูก
25+20	25	+	20	45	ถูก
39-30	39	-	30	19	ถูก
42x40	42	x	40	1680	ถูก
54/50	54	/	50	1.08	ถูก
66+14	66	+	14	80	ถูก
73-17	73	-	17	56	ถูก
89x18	88	x	18	1584	ผิด
93/13	53	/	13	4.07692	ผิด
26+60	26	+	60	86	ถูก
47-19	47	-	19	28	ถูก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ๑ แสดงค่าความถูกต้องในการแยกพยางค์ออกจากประโยค

จำนวนพยางค์	จำนวนชุดที่ทดสอบ	ค่าความถูกต้อง
3 พยางค์	20 ชุด	100%
4 พยางค์	20 ชุด	99%
5 พยางค์	20 ชุด	99%
6 พยางค์	20 ชุด	100%
7 พยางค์	20 ชุด	100%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้