สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การควบคุมอุปกรณ์ไฟฟ้าโดยใช้เสียง

Controlling Electrical Equipment

With Speech Recognition

โดย

นางสาวนงนาจ เด่นเพชรกุล

นายนนทรัฐ บำรุงเกียรติ

นายนพเก้า โสธนกุล

เลขหมู่.....
เลขทะเบียน.....62724
วัน,เดือน,ปี 21 ส.ค. 2549

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

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมโทรคมนาคม

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2548

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การควบคุมอุปกรณ์ไฟฟ้าโดยใช้เสียง

Controlling Electrical Equipment

With Speech Recognition

โดย

นางสาวนงนาจ	เด่นเพชรกุล	45010362
นายนนทรัฐ	บำรุงเกียรติ	45010365
นายนพเก้า	โสธนกุล	45010368

อาจารย์ที่ปรึกษา

รศ.ดร.ไกรสิน ส่งวัฒนา

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมโทรคมนาคม

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2548

ผ่านการตรวจชิ้นงานแล้ว

(ลงชื่อ).....ผู้ตรวจ

ผ่านการตรวจรูปเล่มแล้ว

(ลงชื่อ).....ผู้ตรวจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์ปีการศึกษา 2548

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง **การควบคุมอุปกรณ์ไฟฟ้าโดยใช้เสียง**

CONTROLLING ELECTRICAL EQUIPMENT

WITH SPEECH RECOGNITION

ผู้จัดทำ

1. นางสาวนงนาจ เด่นเพชรกุล 45010362

2. นายนนทรัฐ บำรุงเกียรติ 45010365

3. นายนพเก้า โสธนกุล 45010368

อาจารย์ที่ปรึกษา

(รศ.ดร.ไกรสิน ส่งวัฒนา)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การควบคุมอุปกรณ์ไฟฟ้าโดยใช้เสียง

ELECTRICAL EQUIPMENT CONTROLLING WITH SPEECH RECOGNITION

โดย	นางสาวนงนาจ	เด่นเพชรกุล	45010362
	นายนนทรัฐ	บำรุงเกียรติ	45010365
	นายนพเก้า	โสธนกุล	45010368

อาจารย์ที่ปรึกษา รศ.ดร.ไกรสิน ส่งวัฒนา

บทคัดย่อ

ปริญญานิพนธ์ฉบับนี้มีจุดประสงค์ เพื่อที่จะทำการศึกษาการควบคุมอุปกรณ์ไฟฟ้าโดยใช้เสียงพูด โดยใช้คอมพิวเตอร์ส่วนบุคคลในการรับเสียงนำไปประมวลผลเพื่อทำการแยกแยะคำสั่ง และผลที่ได้ส่งออกไปเป็นอินพุทของวงจรโซลิดสเตทรีเลย์เพื่อควบคุมอุปกรณ์ไฟฟ้า

ABSTRACT

This project studies a method of using human voice to command electrical appliance by using a personal computer to receive speeches, process and make decision. The result is used as input to relay circuit for controlling the electrical appliances.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี ด้วยความกรุณาของ รศ.ดร. ไกรสิน ส่งวัฒนา อาจารย์ที่ปรึกษา ที่กรุณาแนะนำ ให้คำปรึกษาในการทำปริญญานิพนธ์ฉบับนี้ และให้ความเมตตาแก่ศิษย์ด้วยดีเสมอมา ผู้จัดทำรู้สึกซาบซึ้งในความกรุณาเป็นอย่างยิ่ง และขอกราบขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณอาจารย์ทุกท่านที่ได้ประสิทธิ์ประสาทวิชาความรู้ต่างๆ ให้คำแนะนำ คำปรึกษา ที่ดีมาโดยตลอดระยะเวลาที่ศึกษาอยู่ให้แก่ศิษย์

ขอขอบคุณ พี่ๆ เพื่อนๆ ทุกท่านทั้งที่ห้องวิจัย และที่เรียนด้วยกัน สำหรับความเป็นมิตร และความช่วยเหลือต่างๆ ที่มีให้กันระหว่างศึกษาอยู่

ขอขอบคุณเจ้าของเสียงทุกๆ ท่าน ที่ได้เสียสละเวลามาบันทึกเสียงให้

สุดท้ายขอกราบขอบพระคุณ บิดา มารดา และครอบครัว ที่ให้ทุกสิ่งทุกอย่าง กับลูก

ประโยชน์อันถึงมีจากวิทยานิพนธ์ฉบับนี้ ผู้จัดทำขอมอบแด่ผู้มีพระคุณทุกท่าน

ขอขอบคุณด้วยใจจริง

คณะผู้จัดทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

		หน้า
บทที่ 1	บทนำ	1
บทที่ 2	ทฤษฎีและหลักการ	3
2.1	ระบบการพูดและเสียงพูดของมนุษย์	3
2.1.1	กายวิภาคของระบบการเปล่งเสียงของมนุษย์	3
2.1.2	กระบวนการผลิตเสียงพูด	4
2.1.3	ความถี่ฟอร์แมนท์	5
2.1.4	แบบจำลองระบบกำเนิดเสียงพูด	6
2.2	รูปแบบการรู้จำเสียงพูด	7
2.3	การประมวลผลสัญญาณเบื้องต้น	7
2.3.1	การพรีเอมฟาซีส	8
2.3.2	การแบ่งช่วงสัญญาณ	9
2.3.3	การวินโดว์	10
2.4	การวิเคราะห์และวัดค่าลักษณะสำคัญ	11
2.4.1	สัมประสิทธิ์ของ การประมาณพันธะเชิงเส้น	12
2.4.2	การหาค่าสัมประสิทธิ์ LPC และอัตราขยาย	15
2.5	แถบความถี่วิกฤต และความเข้มสัญญาณในแถบความถี่วิกฤต	15
2.5.1	สเกลบาร์คหรือ Critical Band Rate Scale	16
2.5.2	Critical Band Intensity	18
2.6	การจำแนกรูปแบบด้วยแบบจำลองมาร์คอฟแบบซ่อนเร้น	18
2.6.1	ความหมายของแบบจำลองฮิดเด็นมาร์คอฟ	19
2.6.2	กระบวนการแบบจำลองมาร์คอฟ แบบไม่ต่อเนื่องทางเวลา	19
2.6.3	ส่วนประกอบของแบบจำลองมาร์คอฟ	21
2.6.4	ชนิดของ HMM	21
2.6.5	ปัญหาพื้นฐานของแบบจำลองมาร์คอฟ	22
2.6.6	ระบบแบบจำลองมาร์คอฟ	23
2.7	โครงสร้างและสถาปัตยกรรมของไมโครคอนโทรลเลอร์ MCS-51	23
2.7.1	คุณสมบัติของไมโครคอนโทรเลอร์ตระกูล MCS-51 อนุกรม	24
2.7.2	การจัดขาของไมโครคอนโทรเลอร์ MCS-51	24
2.7.3	จังหวะการทำงานของไมโครคอนโทรเลอร์ MCS-51	25
2.7.4	หลักการสื่อสารข้อมูลผ่านพอร์ตอนุกรม	26
2.7.5	รูปแบบของการสื่อสารข้อมูลอนุกรม	27

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.8 การเชื่อมต่อพอร์ตอนุกรมมาตรฐานRS-232 27
2.9 ไอซี MAX232 , L232 28
บทที่ 3 การออกแบบและการสร้าง 30
3.1 การออกแบบส่วนของฮาร์ดแวร์ 30
3.1.1 วงจรสวิตซ์เปิด-ปิด 4 ช่องควบคุมผ่านพอร์ตอนุกรม 30
3.1.2 ส่วนโปแกรมของไมโครคอนโทรลเลอร์ควบคุมฮาร์ดแวร์ 31
3.2 การออกแบบและการสร้างส่วนของการรู้จำเสียงพูด 33
3.2.1 การออกแบบสร้างคำต้นแบบอ้างอิงและคำทดสอบ 33
3.2.2 การเก็บข้อมูล 33
3.2.3 เครื่องมือและอุปกรณ์ในการวิเคราะห์สัญญาณเสียง 33
3.2.4 หลักการรู้จำเสียงพูดในปริญญานิพนธ์นี้ 33
3.2.5 การหาค่าพารามิเตอร์ที่ใช้ในการแยกเสียงพูด 34
บทที่ 4 การทดลองและผลการทดลอง 36
4.1 การทดลองส่วนฮาร์ดแวร์เพื่อส่งข้อมูลผ่านพอร์ตอนุกรม 36
4.1.1 ทดสอบวงจร MAX 232 36
4.1.2 ทดสอบวงจรสวิตซ์เปิดปิด 4 ช่องผ่านพอร์ตอนุกรม 36
4.2 ผลการทดลองส่วนฮาร์ดแวร์ 37
4.2.1 ทดสอบวงจร MAX 232 37
4.2.2 ทดสอบวงจรสวิตซ์เปิดปิด 4 ช่องผ่านพอร์ตอนุกรม 39
4.3 การทดลองส่วนการรู้จำเสียง 42
4.4 ผลการทดลองส่วนการรู้จำเสียง 43
บทที่ 5 บทสรุปและวิจารณ์ 51
ภาคผนวก 52

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ

หน้า

รูปที่ 2.1 แสดงภาพตัดขวางร่างกายเพื่อแสดงอวัยวะในระบบการพูดของมนุษย์ 3

รูปที่ 2.2 แสดงภาพกล่องเสียงในขณะต่างๆ 4

รูปที่ 2.3 แสดงแผนภาพระบบเสียงพูดของมนุษย์ 5

รูปที่ 2.4 แสดงสเปกตรัมของเสียง /a/ ที่ได้จากสัมประสิทธิ์ LPC 6

รูปที่ 2.5 แสดงแผนภาพกรอบจำลองระบบกำเนิดเสียงเบื้องต้น 6

รูปที่ 2.6 แสดงกระบวนการของขั้นตอน Signal Preprocessing 7

รูปที่ 2.7 แสดงขนาดสเปกตรัมของฟังก์ชั่นถ่ายโอนของการพรีเอมฟาซีส 9

รูปที่ 2.8 แสดงการแบ่งช่วงสัญญาณที่ใช้ในการวิเคราะห์ 9

รูปที่ 2.9 แสดงส่วนของสัญญาณที่ตัดมาวิเคราะห์ 10

รูปที่ 2.10 แสดงวินโดว์สี่เหลี่ยม 11

รูปที่ 2.11 แสดงวินโดว์แบบแฮมมิ่ง 11

รูปที่ 2.12 แสดงความสัมพันธ์ระหว่างค่าความผิดพลาดกับจำนวนอันดับ
ของการประมาณเชิงเส้น 12

รูปที่ 2.13 แสดงบล็อกไดอะแกรมแสดงโมเดลการสร้างสัญญาณเสียงพูดอย่างง่าย 13

รูปที่ 2.14 ระดับของสัญญาณโทนทดสอบที่ถูกบดบังโดยสัญญาณ white noise
ที่ระดับสัญญาณต่างๆและเส้นประเป็นระดับ Threshold in quiet 16

รูปที่ 2.15 แสดงความสัมพันธ์ระหว่างลำดับของ Bark Scale กับความถี่ Hz 17

รูปที่ 2.16 แสดงการคำนวณ Critical Band Intensity 18

รูปที่ 2.17 แสดงตัวอย่างแบบจำลองมาร์คอฟ 5 สเตท 20

รูปที่ 2.18 แสดงการส่งข้อมูลขนาด 8 บิตแบบอนุกรมพร้อมด้วย บิตเริ่มต้น,
บิตพาริตี้,บิตหยุด ด้วยความเร็ว 9600 บิตต่อวินาที 27

รูปที่ 2.19 แสดงระดับแรงดันสัญญาณของพอร์ตอนุกรม RS-232 กับ TTL
ในสถานะลอจิก "1" และ "0" 28

รูปที่ 2.20 แสดงตำแหน่งขาของไอซี MAX232, L232 และการต่อใช้งาน 29

รูปที่ 3.1 แสดงวงจรแปลงแรงดัน 220 V 30

รูปที่ 3.2 วงจรสวิตซ์เปิด-ปิด 4 ช่องควบคุมผ่านพอร์ตอนุกรม 31

รูปที่ 3.3 แผนผังแสดงขั้นตอนการทำงานของโปรแกรมสวิตซ์เปิดปิด 4 ช่อง 32

รูปที่ 3.4 แสดงขั้นตอนในการวิเคราะห์เสียง 34

รูปที่ 4.1 แสดงผลเมื่อทำการส่ง 1 จากคีย์บอร์ดผ่านวงจร MAX232 37

รูปที่ 4.2 แสดงผลเมื่อทำการส่ง 2 จากคีย์บอร์ดผ่านวงจร MAX232 37

รูปที่ 4.3 แสดงผลเมื่อทำการส่ง 3 จากคีย์บอร์ดผ่านวงจร MAX232 38

รูปที่ 4.4 แสดงผลเมื่อทำการส่ง 4 จากคีย์บอร์ดผ่านวงจร MAX232 38

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.5 แสดงรูปหน้าจออินเตอร์เฟสหลังจากกดสวิตซ์ Open Port 39
รูปที่ 4.6 แสดงการทำงานของวงจรหลังจากกดสวิตซ์ Open Port 39
รูปที่ 4.7 แสดงรูปหน้าจออินเตอร์เฟสหลังจากกดสวิตซ์ OPEN 1 40
รูปที่ 4.8 แสดงการทำงานของวงจรหลังจากกดสวิตซ์ OPEN 1 40
รูปที่ 4.9 แสดงรูปหน้าจออินเตอร์เฟสหลังจากกดสวิตซ์ OPEN 2,OPEN 3 และ OPEN 4 41
รูปที่ 4.10 แสดงการทำงานของวงจรหลังจากกดสวิตซ์ OPEN 2,OPEN 3 และ OPEN 4 41
รูปที่ 4.12 แสดงสัญญาณเสียง "เปิด หนึ่ง" 43
รูปที่ 4.13 แสดงสัญญาณเสียง "เปิด หนึ่ง" ในช่วง 900 แซมเปิ้ลแรก 43
รูปที่ 4.14 แสดงสัญญาณเสียง "เปิด หนึ่ง" หลังจากผ่านกระบวนการพรีเอมฟาซีส 44
รูปที่ 4.15 แสดงสัญญาณเสียง "เปิด หนึ่ง" หลังจากผ่านกระบวนการแฮมมิ่งวินโดว์ 44
รูปที่ 4.16 แสดงสัมประสิทธิ์ LPC เฟรมที่ 1 ของแต่ละสัญญาณเสียง 45
รูปที่ 4.17 แสดงสเปกตรัม LPC เฟรมที่ 1 ของแต่ละสัญญาณเสียง 46
รูปที่ 4.18 แสดง Critical Band Intensity บน สเกลบาร์ค เฟรมที่ 1 ของแต่ละสัญญาณเสียง 47

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

	หน้า
ตารางที่ 2.1 ขอบเขตและความถี่ศูนย์กลาง Critical Rate Scale	17
ตารางที่ 4.1 ผลการทดสอบแบบอ้างอิงจากผู้พูดในกรณีต่างๆ	49

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

ปัจจุบันเทคโนโลยี ต่างๆมี การพัฒนาอย่างรวดเร็ว รวมถึงเทคโนโลยีด้านคอมพิวเตอร์ ซึ่งจะนำไปรวมไปใช้งานร่วมกับ เทคโนโลยีด้านอื่นๆอีกมากมาย เช่น ด้านการสื่อสาร ระบบควบคุม เป็นต้น คอมพิวเตอร์ได้ถูกพัฒนาขึ้นมาเรื่อยๆ เพื่อทำงานต่างๆแทนมนุษย์ได้ การรู้จำเป็นการพัฒนาอีกด้านของคอมพิวเตอร์

ในอดีตมีการวิเคราะห์ว่าการสอนให้คอมพิวเตอร์สามารถรู้จำเสียงพูดของมนุษย์นั้นไม่สามารถทำได้ เนื่องจากคำพูดของมนุษย์มีความซับซ้อน แต่มีความแตกต่างกันในแต่ละบุคคล ความเร็ว รวมถึง เสียงสูง เสียงต่ำ แต่ได้มาการศึกษาพัฒนาให้คอมพิวเตอร์สามารถรู้จำเสียงพูดได้เรื่อยมา จนในปัจจุบันการรู้จำเสียงพูดสามารถนำไปใช้งานกับอุปกรณ์ต่างๆ เช่น โทรศัพท์มือถือ และ ATM (Automatics Teller Machine) ถึงแม้ว่าคอมพิวเตอร์จะไม่สามารถรู้จำเสียงได้เท่ามนุษย์ แต่เราสามารถสอนให้คอมพิวเตอร์รู้จำได้ในขีดจำกัดระดับหนึ่ง เช่น จำกัดคำที่สามารถรู้จำได้ โดยมีการคาดคะเนกันว่าในอนาคตอุปกรณ์เครื่องใช้เกือบทุกประเภทจะสามารถสั่งงานได้ด้วยเสียงมนุษย์

การรู้จำเสียงพูดนั้นมีทั้งการรู้จำเสียงแบบขึ้นกับผู้พูด(Speaker-Dependent Speech Recognition) และการรู้จำเสียงแบบไม่ขึ้นกับผู้พูด (Speaker-Independent Recognition) ซึ่งการรู้จำเสียงพูดนั้นมีความยุ่งยากละซับซ้อนมาก เนื่องจากความแตกต่างในการออกเสียงของผู้พูดแต่ละบุคคล จึงทำให้การพัฒนาทางด้านการรู้จำเสียงพูดเป็นไปอย่างล่าช้า การศึกษาและวิจัยการรู้จำเสียงพูดนั้นยังแบ่งตามลักษณะการรับรู้เสียงพูดออกเป็น 2 วิธีคือ

1. พิจารณาทั้งหน่วยภาษาที่เปล่งออกมาทั้งชุด มีทั้งระบบการรู้จำคำพูดเดี่ยว (Isolate word Recognition) และระบบการรู้จำคำพูดต่อเนื่อง (Continuous word Recognition) ซึ่งข้อดีของระบบเหล่านี้คือง่าย เนื่องจากมีการหลีกเลี่ยงผลกระทบอันเนื่องมาจากฐานรากของเสียงภายในคำหรือกลุ่มคำนั้น แต่ข้อเสียคือสามารถรู้จำคำได้จำนวนจำกัด เนื่องจากต้องใช้เนื้อที่จำนวนมากในการจัดเก็บแบบจำลองอ้างอง และต้องใช้เวลาในการคำนวณเพื่อเปรียบเทียบมากตามแบบจำลองที่มีอยู่

2. พิจารณาโดยการแยกแยะรายละเอียดของหน่วยเสียง (Phonetic Recognition) วิธีนี้จะพิจารณาลักษณะของหน่วยเสียงพยัญชนะ หน่วยเสียงสระ และหน่วยเสียงวรรณยุกต์ โดยจะใช้หน่วยเสียงย่อยเหล่านี้เป็นหลักในการรู้จำเสียงพูด ซึ่งวิธีนี้เหมาะสำหรับการพัฒนาไปสู่ระบบการรู้จำคำจำนวนมาก

โครงงานนี้เป็นโครงงานที่ศึกษาเกี่ยวกับการสั่งงานอุปกรณ์เครื่องใช้ไฟฟ้า ด้วยเสียงมนุษย์ โดยต่อเข้ากับพอร์ตอนุกรมของคอมพิวเตอร์ส่วนบุคคล โดยส่วนประกอบของโครงงาน คือส่วนซอฟแวร์และส่วนฮาร์ดแวร์ โดยส่วนซอฟแวร์นั้นเป็นการศึกษาถึงกระบวนการของการรู้จำเสียงพูดมนุษย์แบบพิจารณาทั้งหน่วยภาษา เพื่อนำผลที่ได้จากกระบวนการนี้ เข้าสู่ส่วนฮาร์ดแวร์ที่เป็นวงจรสวิตช์เปิดปิด 4 ช่อง ผ่านพอร์ตอนุกรม เพื่อใช้เป็นวงจรในการควบคุมการเปิดปิดอุปกรณ์เครื่องใช้ไฟฟ้าได้ 4 ช่อง ให้สามารถทำงานได้จริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วัตถุประสงค์

เพื่อนำระบบรู้จำเสียงพูดมาประยุกต์ใช้ควบคุมการเปิดปิดอุปกรณ์ไฟฟ้าผ่านพอร์ตนุกรมของคอมพิวเตอร์ส่วนบุคคล

ขอบเขตโครงงาน

ออกแบบและสร้างวงจรสวิตซ์เปิดปิด 4 ช่อง ผ่านพอร์ตอนุกรม ซึ่งสามารถควบคุมการเปิดปิดอุปกรณ์ไฟฟ้าโดยใช้เสียงพูดให้สามารถใช้งานได้จริง โดย

1.วงจรมีสวิตซ์กดเพื่อควบคุมการเปิดปิด อุปกรณ์ไฟฟ้าบนบอร์ดได้โดยตรง

2.วงจรสามารถควบคุมการเปิดปิดอุปกรณ์ไฟฟ้าได้ด้วยการสั่งงานผ่านพอร์ตอนุกรมของคอมพิวเตอร์ส่วนบุคคล

3.วงจรสามารถควบคุมการเปิดปิดอุปกรณ์ไฟฟ้าได้โดยใช้เสียงสั่งงานผ่านพอร์ตอนุกรมของคอมพิวเตอร์ส่วนบุคคล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

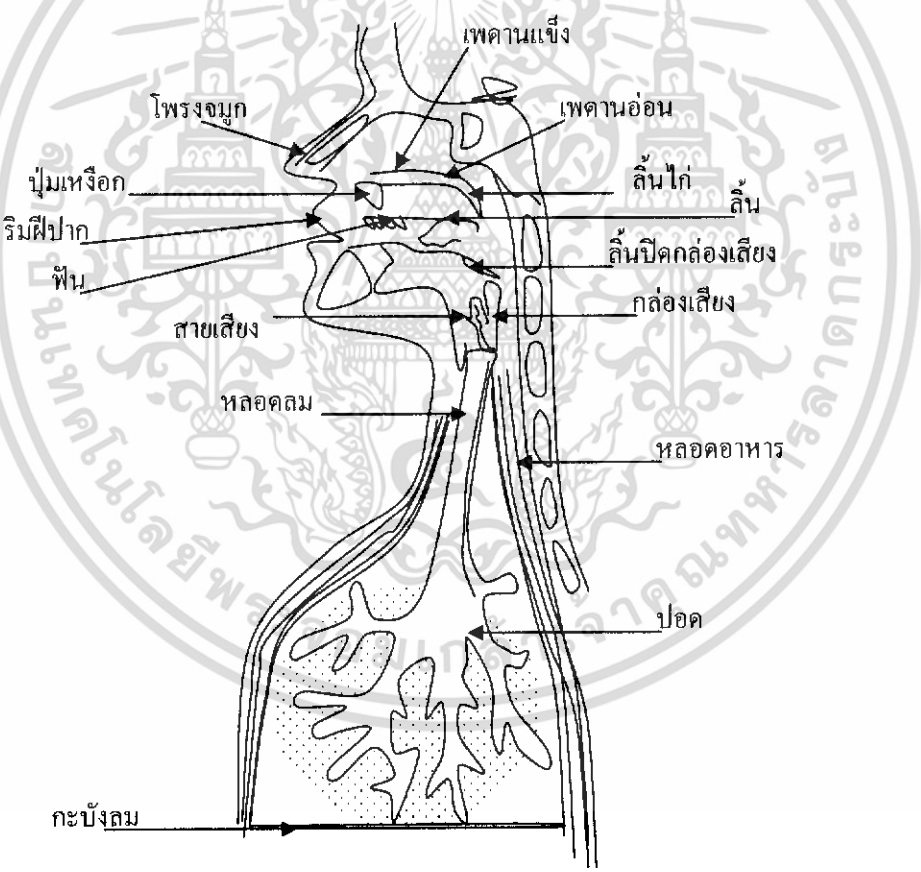
บทที่ 2
ทฤษฎีและหลักการ

2.1 ระบบการพูดและเสียงพูดของมนุษย์

2.1.1 กายวิภาคของระบบการเปล่งเสียงของมนุษย์

การศึกษาด้าน กายวิภาคศาสตร์ของมนุษย์ (Human anatomy) วิชาที่ว่าด้วย เสียงของภาษา (phonetics) และศาสตร์ทางด้านเสียง (acoustics) ช่วยให้เข้าใจขั้นตอนการทำงานร่วมกันของอวัยวะต่างๆใน การเปล่งเสียงพูด ตลอดจนลักษณะทางกายภาพของเสียงพูด เพื่อนำมาวิเคราะห์และสร้างแบบจำลองเลียนแบบเสียงพูดมนุษย์

การทำให้เกิดเสียงเป็นหน้าที่หนึ่งของระบบหายใจ การออกเสียงหรือการพูดของมนุษย์แต่ละครั้งจะต้อง มีการทำงานร่วมกันของอวัยวะต่างๆ ของร่างกาย ดังรูปที่ 2.1



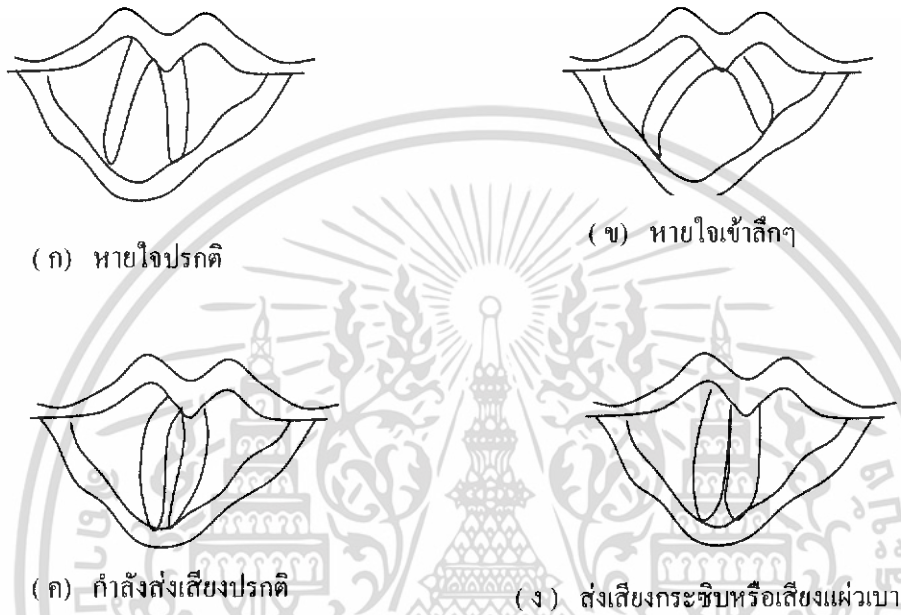
รูปที่ 2.1 แสดงภาพตัดขวางร่างกายเพื่อแสดงอวัยวะในระบบการพูดของมนุษย์

1. ปอดและกะบังลม ทำหน้าที่สำคัญในการหายใจ และเป็นต้นกำเนิดการไหลของอากาศใน กระบวนการผลิตเสียง

เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. หลอดลม ทำหน้าที่นำอากาศออกจากปอดผ่านกล่องเสียง และเป็นอวัยวะที่อยู่ด้านหน้าของหลอดอาหาร

3. กล่องเสียง เป็นอวัยวะพิเศษที่ทำหน้าที่เป็นทางเดินอากาศเวลาหายใจ และเป็นตัวผลิตพัลส์ (Pulse) ของอากาศขณะเปล่งเสียง ซึ่งประกอบด้วยเส้นเสียง (vocal cords) และช่องสายเสียง (glottis) รูปร่างของกล่องเสียงและเส้นเสียงในลักษณะต่างๆ แสดงดังรูปที่ 2.2



(ก) หายใจปรกติ (ข) หายใจเข้าลึกๆ

(ค) กำลังส่งเสียงปรกติ (ง) ส่งเสียงกระซิบหรือเสียงแผ่วเบา

รูปที่ 2.2 แสดงภาพกล่องเสียงในขณะต่างๆ

4. ช่องปากและส่วนของหลอดอาหารตอนต้น อวัยวะกลุ่มนี้อยู่ต่อจากกล่องเสียง อาจเรียกว่าอวัยวะกำทอนเสียง (vocal tract) ทำหน้าที่กำทอนเสียง โดยให้กำทอนทั้งเสียงที่เกิดจากกล่องเสียงและเสียงที่เกิดภายในช่องปาก ขนาดของอวัยวะกำทอนเสียง ขึ้นอยู่กับตำแหน่งของลิ้น ริมฝีปาก ขากรรไกร และเพดานอ่อน และเปลี่ยนแปลงไปตามการออกเสียง

5. โพรงจมูก เริ่มจากเพดานอ่อน จนถึงรูจมูกทั้งสองทำหน้าที่กำทอนเสียงร่วมกับช่องปาก เมื่อมีการเปล่งเสียงที่ออกทางจมูก (nosal sounds) เช่นเสียง /ม/ , /น/ และ /ง/ เป็นต้น

2.1.2 กระบวนการผลิตเสียงพูด

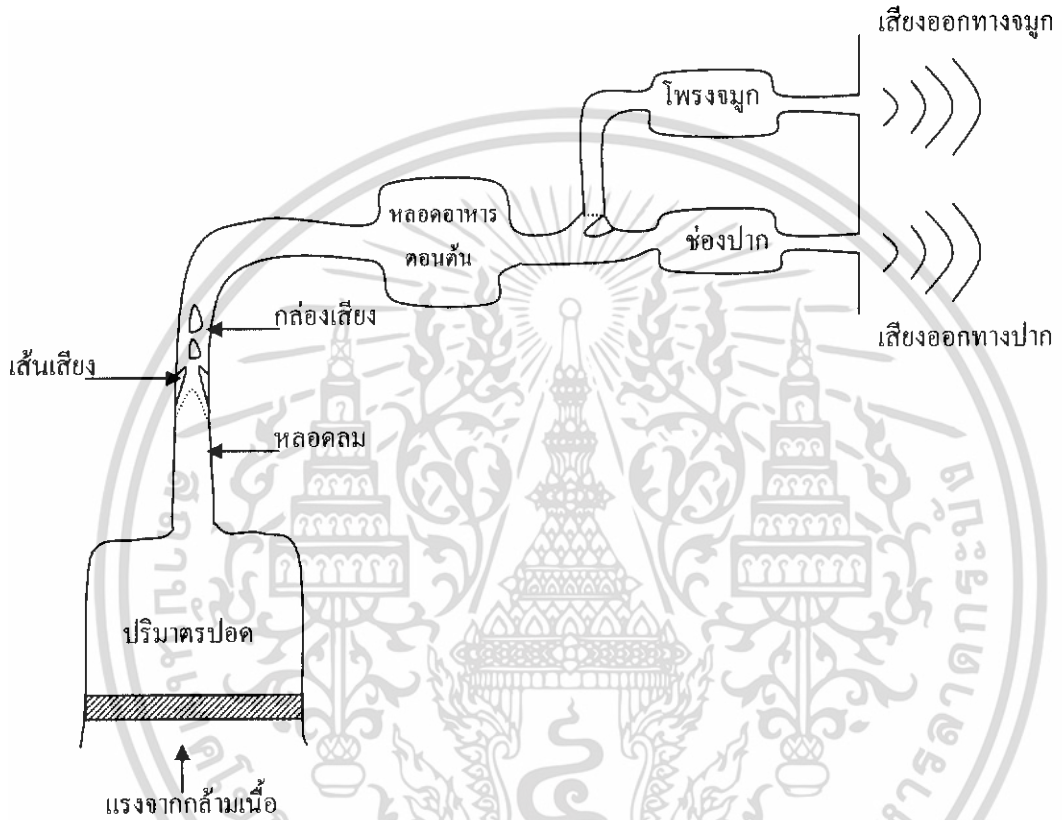
จากกายวิภาคและการเปล่งเสียงพูดของมนุษย์ สามารถแสดงแผนภาพของระบบกำเนิดเสียง ดังรูปที่ 2.3 ซึ่งสามารถจำแนกกลไกการสร้างเสียงมนุษย์ได้ 3 แบบ ดังนี้

1. อากาศไหลจากปอดจะถูกมอดูเลต (modulate) โดยการสั่นของเส้นเสียง ทำให้เกิดคลื่นเสียงลักษณะคล้ายพัลส์ ที่มีคาบเวลาแบบควอไซด์ (quasi periodic pulse like excitation)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. อากาศที่ไหลจากปอดถูกทำให้ปั่นป่วน โดยการบังคับให้ไหลผ่านช่องแคบอันเกิดจากการบีบตัวของอวัยวะในช่องปาก ทำให้เกิดเสียงในคล้ายเสียงรบกวน (noise like excitation)

3. อากาศที่ไหลถูกกัก และเกิดแรงดันอยู่ภายในส่วนของช่องปากที่ปิดจากนั้นจึงปล่อยให้อากาศที่มีแรงดัน พุ่งออกไปอย่างรวดเร็ว ทำให้เกิดการกระตุ้นเป็นเสียงในช่วงเริ่มต้น (transient excitation)



รูปที่ 2.3 แสดงแผนภาพระบบเสียงพูดของมนุษย์

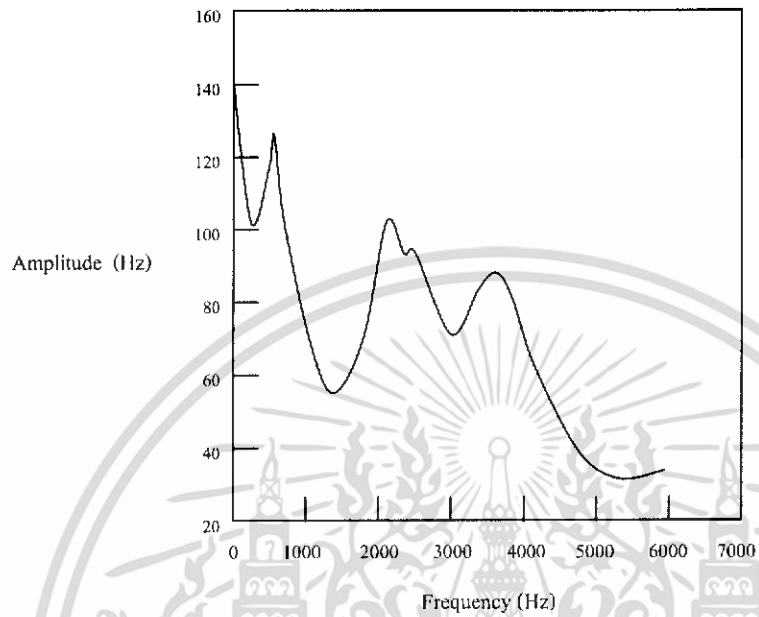
2.1.3 ความถี่ฟอร์แมนท์ (formant frequency)

คลื่นเสียงพูดของมนุษย์นั้นเป็นคลื่นตามยาว (Longitude wave) เกิดจากการสั่นของอนุภาคตัวกลางนั่นคืออากาศ และทิศทางการสั่นของอนุภาคจะอยู่ในทิศทางเดียวของการเคลื่อนที่ คลื่นเสียงเป็นคลื่นที่เปลี่ยนแปลงตามเวลา

ส่วนความถี่ของเสียงมนุษย์นั้นจะมีค่าเท่าใดขึ้นอยู่กับ การบังคับขนาดและรูปทรงของกลุ่มอวัยวะกำทอนเสียง ซึ่งหมายถึง ช่องปากและส่วนของหลอดอาหารตอนต้น ซึ่งอวัยวะกลุ่มนี้จะอยู่ต่อจากกล่องเสียง และความถี่นั้นจะเรียกว่า ความถี่กำทอน (resonance frequency) หรือความถี่ฟอร์แมน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

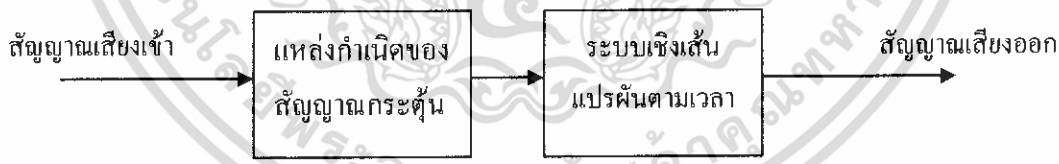
ความถี่ฟอร์แมนท์ สำหรับเสียงพูดคำหนึ่งๆ อาจจะมีได้หลายค่า ดังรูปที่ 2.4 เป็นสเปกตรัมของเสียง /a / ซึ่งมีความถี่ฟอร์แมนท์ที่ 1, 2, 3 และ 4 อยู่ ณ ตำแหน่งที่ 550, 2150, 2600, 3750 เฮิร์ต ตามลำดับ



รูปที่ 2.4 แสดงสเปกตรัมของเสียง /a/ ที่ได้จากสัมประสิทธิ์ LPC

2.1.4 แบบจำลองระบบกำเนิดเสียงพูด

จากรูปที่ 2.3 สามารถแสดงแผนภาพกรอบจำลองระบบกำเนิดเสียงเบื้องต้น ดังรูปที่ 2.5



รูปที่ 2.5 แสดงแผนภาพกรอบจำลองระบบกำเนิดเสียงเบื้องต้น

จากรูป 2.5 แหล่งกำเนิดสัญญาณกระตุ้นทำหน้าที่แทนการทำงานของปอด และกล่องเสียงส่วนนี้จะผลิตขบวนพัลส์ที่มีคาบเวลาพิตช์ขณะเปล่งเสียงโฆษะ (voiced) และให้กำเนิดเสียงคล้ายเสียงรบกวนขณะเปล่งเสียงอโฆษะ (unvoiced) ส่วนที่สองเป็นท่อกำทอนเสียง จะแทนการทำงานของช่องปากและโพรงจมูก ทำหน้าที่เหมือนตัวกรองสัญญาณ (filter) ที่ยอมให้ความถี่ฟอร์แมนท์ผ่านได้ ซึ่งสามารถแทนด้วยระบบเชิงเส้นแปรผันตามเวลา (time-varying linear system)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 รูปแบบการรู้จำเสียงพูด (Speech recognition pattern)

การรู้จำเสียงพูด (speech recognition) จุดมุ่งหมายหลักของการรู้จำเสียงพูดก็คือ การเพิ่มพูนความสามารถของอุปกรณ์ต่างๆในการรับรู้สื่อสารและโต้ตอบกับมนุษย์ได้ เพื่อเพิ่มทางเลือกในการควบคุมสั่งการอุปกรณ์เครื่องมือต่างๆ โดยเฉพาะเครื่องคอมพิวเตอร์ ซึ่งการใช้เสียงพูดควบคุมการสั่งการนี้ถือเป็นวิธีการที่เป็นธรรมชาติมากที่สุดของมนุษย์

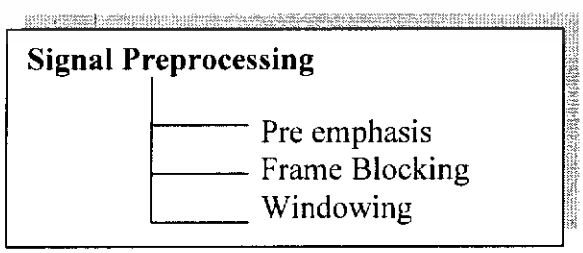
การรู้จำเสียงพูดนั้นมีการศึกษาทั้งชนิดขึ้นกับผู้พูด (speaker Dependent) และชนิดไม่ขึ้นกับผู้พูด (speaker Independent) และในกรณีการรู้จำโดยไม่ขึ้นกับผู้พูดนั้นจำแนกได้เป็น แบบคำโดด (isolate Word) แบบคำต่อเนื่อง (connected Word) และแบบเสียงพูดต่อเนื่อง (continuous Speech)

ระบบการรู้จำเสียงพูดจะอาศัยการฝึกฝน (training) เพื่อจดจำรูปแบบอ้างอิง (reference pattern) ไว้ใช้ในการเปรียบเทียบกับเสียงพูดที่ยังไม่ทราบรูปแบบ และใช้รูปแบบอ้างอิงเหล่านั้นในการตัดสินใจรูปแบบที่ใกล้เคียงกับเสียงพูดที่นำมาเปรียบเทียบมากที่สุด ขั้นตอนการรู้จำเสียงพูดโดยทั่วไปแบ่งได้เป็น 3 ขั้นตอน ได้แก่ การประมวลสัญญาณเบื้องต้น (signal Preprocessing) การวิเคราะห์และวัดค่าลักษณะสำคัญ (feature measurement) และการจำแนกรูปแบบ (pattern classification)

2. 3 การประมวลผลสัญญาณเบื้องต้น (Signal Preprocessing)

การประมวลผลสัญญาณเบื้องต้นเป็นขั้นตอนกรรมวิธีในการจัดเตรียมข้อมูลจากข้อมูลดิบของเสียงพูดที่ได้จากการบันทึกเสียงนำมาผ่านกรรมวิธีประมวลผลสัญญาณเชิงเลข เพื่อใช้ในการประมวลผลในขั้นตอนต่อไป และเนื่องจากสัญญาณเสียงพูดโดยรวมจะแปรเปลี่ยนตามเวลา อีกทั้งสัญญาณเสียงพูดนั้นไม่เสถียร (Nonstationary) ดังนั้น ในการประยุกต์ใช้งานกรรมวิธีการประมวลผลสัญญาณเชิงเลขกับสัญญาณเสียงพูด จึงต้องแบ่งสัญญาณเสียงพูดออกเป็นส่วนย่อย เรียกว่า กรอบเสียงพูด (Speech Frame) โดยแต่ละกรอบเสียงพูดจะมีความยาวประมาณ 10 – 40 มิลลิวินาที ซึ่งอนุโลมได้ว่าสัญญาณเสียงพูดในแต่ละกรอบเสียงพูดมีความเสถียรไม่แปรเปลี่ยนตามเวลา (Stationary) หรือมีลักษณะที่คงที่นั่นเอง จากนั้นจึงสามารถทำการประมวลผลสัญญาณเชิงเลขกับสัญญาณเสียงพูดในแต่ละกรอบเสียงพูดได้

การประมวลผลสัญญาณเบื้องต้นแบ่งได้เป็น 3 กระบวนการย่อยได้แก่ การพรีเอมฟาซีส (preemphasis) การแบ่งช่วงสัญญาณ (frame blocking) การวินโดว์ (windowing)



รูปที่ 2.6 แสดงกระบวนการของขั้นตอน Signal Preprocessing

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.1 การพรีเอมฟาซีส

เนื่องจากสัญญาณเสียงพูดของมนุษย์จะมีส่วนประกอบใหญ่อยู่ที่ความถี่ต่ำเมื่อเทียบกับแถบความถี่ที่ปฏิบัติงาน (Bandwidth) ไม่เกิน 5 กิโลเฮิรตซ์ ดังนั้นเพื่อให้อัตราส่วนสัญญาณเสียงต่อสัญญาณรบกวน (signal to noise ratio : SNR) มีค่าค่อนข้างคงที่ตลอดช่วงความถี่ที่ปฏิบัติงานนี้ เราจึงต้องมีการ พรีเอมฟาซิสหรือการเน้นล่วงหน้า โดยเน้นความถี่สูงให้มีขนาดสูงขึ้น นั่นคือ การพรีเอมฟาซิสก็คือการกรองสัญญาณด้วยวงจรกรองความถี่สูงผ่าน (high pass filter) ซึ่งมักนิยมใช้ในวงจรกรองอันดับเชิงเลขอันดับหนึ่ง (First-Order Digital Filter) ซึ่งมีฟังก์ชั่นถ่ายโอนเป็น

$$H(z) = 1 - a.z^{-1} \quad ; \quad 0.9 \leq a \leq 1.0$$

กรณีนี้เอาท์พุตการพรีเอมฟาซีส $s'(n)$ มีความสัมพันธ์กับอินพุต $s(n)$ ดังสมการผลต่างสืบเนื่อง

$$s'(n) = s(n) - a.s(n-1)$$

เมื่อ a เป็นสัมประสิทธิ์ตัวกรอง

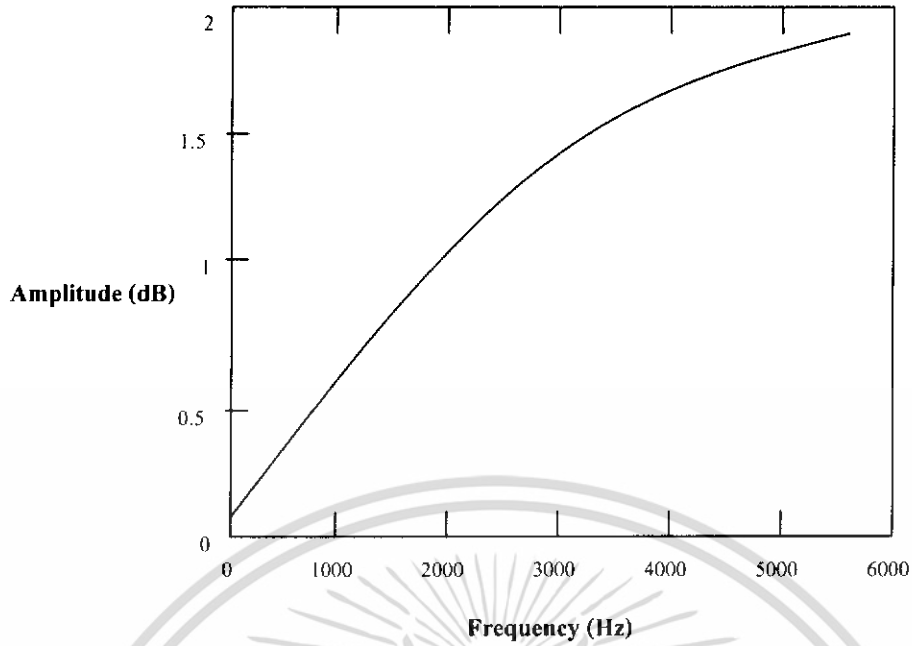
$s'(n)$ เป็นค่าสัญญาณเสียงพูดขาออกที่ผ่านการพรีเอมฟาซีสที่ n

$s(n)$ เป็นค่าสัญญาณเสียงพูดขาเข้าที่ n

$s(n-1)$ เป็นค่าสัญญาณเสียงพูดขาเข้าที่ n-1

โดยยิ่งค่า a มีค่าเข้าใกล้ 1 มาก ที่ความถี่สูงก็จะถูกขยายมากขึ้นด้วย ค่า a ที่นิยมใช้ร่วมกับการวิเคราะห์หาค่าสัมประสิทธิ์การประมาณพันธะเชิงเส้นหรือ สัมประสิทธิ์ของ LPC (Linear Prediction Coefficient) ในขั้นตอนการวิเคราะห์และวัดค่าลักษณะสำคัญซึ่งจะกล่าวถึงต่อไปนั้น กำหนดมีค่าประมาณ 0.9375 เมื่อนำฟังก์ชั่นถ่ายโอนมาพล็อตกราฟของขนาดเทียบกับความถี่จะได้ รูปที่ 2.7

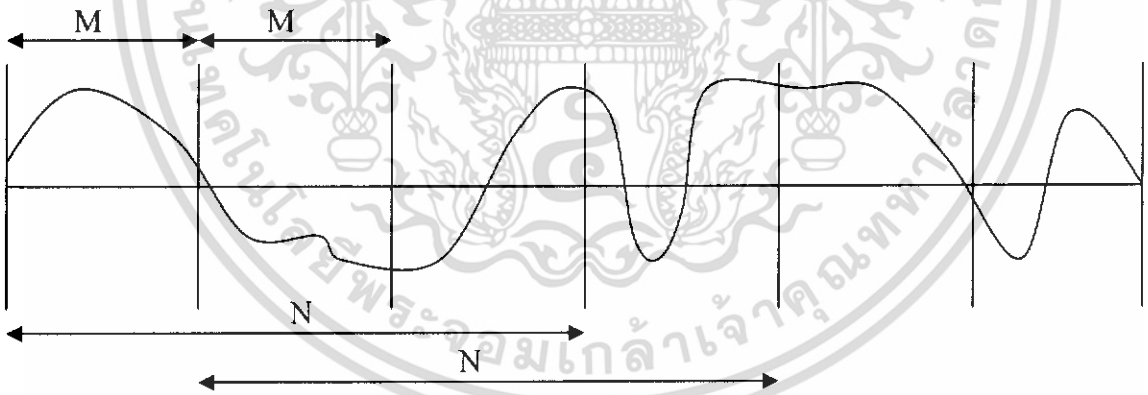
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.7 แสดงขนาดสเปกตรัมของฟังก์ชั่นถ่ายโอนของการพรีเอมฟาซีส

2.3.2 การแบ่งช่วงสัญญาณ

สัญญาณที่ผ่านการพรีเอมฟาซิสแล้ว $s'(n)$ จะถูกตัดออกเป็นช่วง ๆ หรือเฟรม ช่วงละ N ตัวอย่างสัญญาณ การวิเคราะห์จะวิเคราะห์ทีละช่วงของ N ตัวอย่างสัญญาณ ดังรูปที่ 2.8



รูปที่ 2.8 แสดงการแบ่งช่วงสัญญาณที่ใช้ในการวิเคราะห์

โดยช่วงในการวิเคราะห์แต่ละช่วงจะถูกเลื่อนระยะไปเป็น M ช่วงสัญญาณ จะเห็นได้ว่า ถ้าค่า M มากกว่าค่า N ในการเลื่อนของช่วงในการวิเคราะห์จะทำให้บางสัญญาณไม่ถูกใช้ในการวิเคราะห์ ก็จะเป็นการสูญเสียส่วนหนึ่งทำให้ผลที่ได้ไม่ถูกต้องเท่าที่ควร ถ้าค่า M เล็กกว่า N ก็จะทำให้ตัวอย่างสัญญาณทุกตัวถูกนำมาวิเคราะห์ ยิ่งค่า M เล็กเท่าใด ความแม่นยำในการวิเคราะห์ก็จะยิ่งสูงเท่านั้น แต่ก็จะทำให้การคำนวณช้าลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลักการกำหนดขนาดวินโดว์

1.วินโดว์จะต้องสั้นพอที่คุณสมบัติของเสียงที่เราสนใจวิเคราะห์นั้นยังไม่เปลี่ยนแปลงในวินโดว์

2.วินโดว์จะต้องยาวพอที่ จำนวนตัวอย่างสัญญาณในวินโดว์สามารถนำมาคำนวณหาคุณ สมบัติที่ต้องการได้

3.วินโดว์ที่ติดกัน ไม่ควรสั้นจนกระโดดข้ามข้อมูลบางส่วนไป แต่ควรเลื่อนวินโดว์ให้น้อยกว่าขนาดของเฟรม

เนื่องจากเราใช้ความถี่ในการสุ่มสัญญาณ 11.025 กิโลเฮิรตซ์ ในการวิเคราะห์นี้เราเลือกใช้ค่า N = 300 และค่า M=100 [1] นั่นคือ ช่วงในการวิเคราะห์คือ 27.21 มิลลิวินาที และระยะเวลาในการเลื่อนเฟรมประมาณ 9 มิลลิวินาที

2.3.3 การวินโดว์

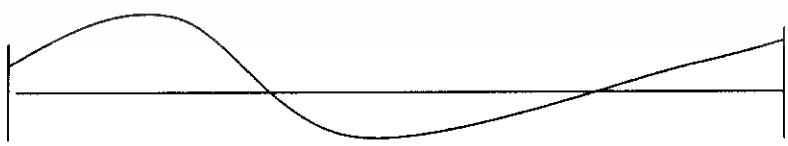
ขั้นตอนการวินโดว์หรือขั้นตอนการวางกรอบขนาดสัญญาณ โดยการคูณแต่ละค่าของสัญญาณในกรอบเสียงพูดด้วยค่าฟังก์ชั่นกรอบ (window function) ซึ่งมีหลายประเภท ได้แก่ Rectangular window , Hamming window , Hanning window, Blackman window , Kaiser window เป็นต้น

พิจารณาช่วงสัญญาณ N ตัวอย่างของช่วงสัญญาณใด ๆ ที่ตัดมาวิเคราะห์ ดัง รูปที่ 2.10 จะเห็นว่าที่ขอบของเฟรมที่ตัดมานี่มีความไม่ต่อเนื่องของสัญญาณ ถ้ามองในโดเมนคงวามถี่สูงก็จะมีความถี่สูงเกิดขึ้น ดังนั้นเพื่อที่จะลดองค์ประกอบทางความถี่สูงเหล่านี้เราจะต้องคูณด้วยฟังก์ชันวินโดว์เพื่อลดความไม่ต่อเนื่องของสัญญาณที่ขอบ และไม่ทำให้สเปคตรัมของสัญญาณในช่วงความถี่ต่ำเปลี่ยนไปมากนัก ในที่นี้จะใช้วินโดว์แฮมมิ่ง (Hamming window function) ซึ่งนิยามโดยสมการดังนี้คือ

$$w(n) = 0.54 - 0.46\cos(\frac{2\pi n}{N-1}) \quad \text{เมื่อ } n = 0,1,2,\ldots,N-1 \qquad (2.1)$$

โดย N เป็นจำนวนข้อมูลในแต่ละกรอบเสียงพูด

n เป็นข้อมูลที่ n ของข้อมูลทั้งหมด N ค่า ในแต่ละกรอบเสียงพูด



รูปที่ 2.9 แสดงส่วนของสัญญาณที่ตัดมาวิเคราะห์

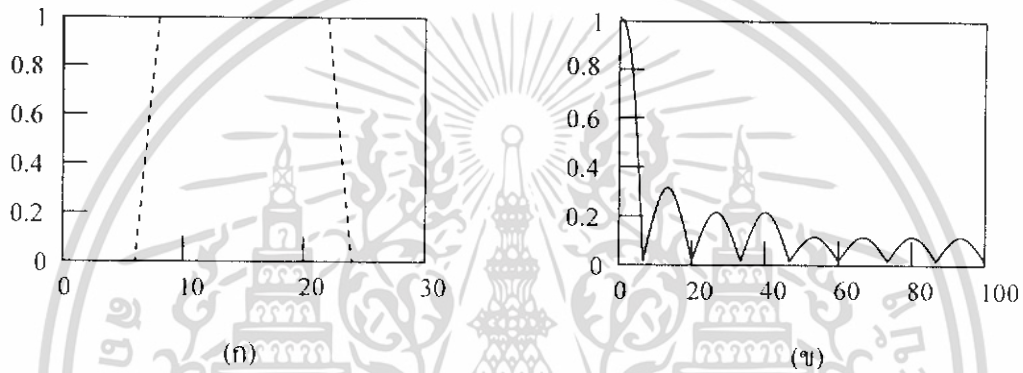
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และรูปที่ 2.10 และ รูปที่ 2.11 แสดงองค์ประกอบทางเวลา และทางความถี่ของฟังก์ชั่นวินโดว์แบบสี่เหลี่ยม และแบบแฮมมิ่ง ตามลำดับ จะเห็นได้ว่าสเปกตรัมของวินโดว์แฮมมิ่งจะมีริพเพิล (ripple) น้อยกว่าวินโดว์สี่เหลี่ยม และจะได้

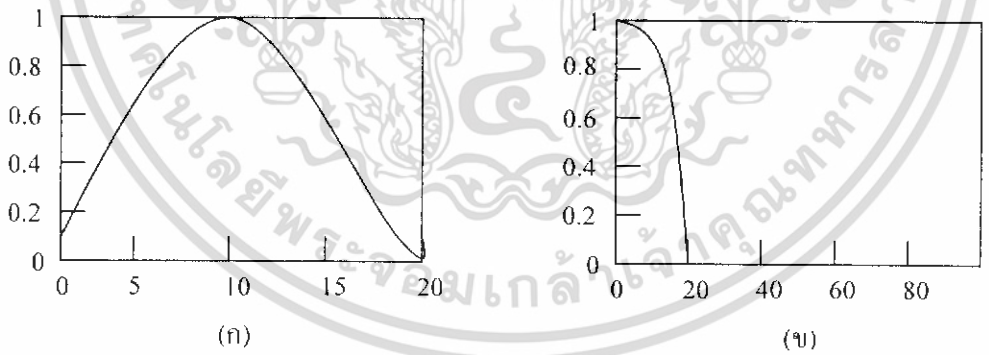
$$x'_l(n) = x_l(n).w(n) \qquad \text{เมื่อ } l = 0,1,2,\ldots L\text{-}1 \tag{2.2}$$

โดย L เป็นจำนวนกรอบเสียงพูดทั้งหมด

l เป็นกรอบที่ l ของกรอบเสียงพูด L ทั้งหมด



รูปที่ 2.10 แสดงวินโดว์สี่เหลี่ยม (ก) ในโดเมนเวลา (ข) ในโดเมนความถี่



รูปที่ 2.11 แสดงวินโดว์แบบแฮมมิ่ง (ก) ในโดเมนเวลา (ข) ในโดเมนความถี่

2.4 การวิเคราะห์และวัดค่าลักษณะสำคัญ (Feature Measurement)

การวิเคราะห์และวัดค่าลักษณะสำคัญ เป็นการวิเคราะห์สัญญาณเสียงพูด เพื่อเก็บรวบรวมลักษณะสำคัญของเสียงพูดแต่ละเสียง สำหรับการฝึกฝนระบบให้รับรู้ถึงความแตกต่างของเสียงพูดแต่ละเสียงและเพื่อใช้ในการเปรียบเทียบแบ่งแยก ความแตกต่างของเสียงพูดแต่ละเสียงออกจากกัน ในการวัดค่าลักษณะสำคัญจะใช้วิธีการหาสัมประสิทธ์ของการประมาณพันธะเชิงเส้น ส่วนการวิเคราะห์ลักษณะสำคัญจะอาศัยหลักการของความเข้มสัญญาณในแถบความถี่วิกฤต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

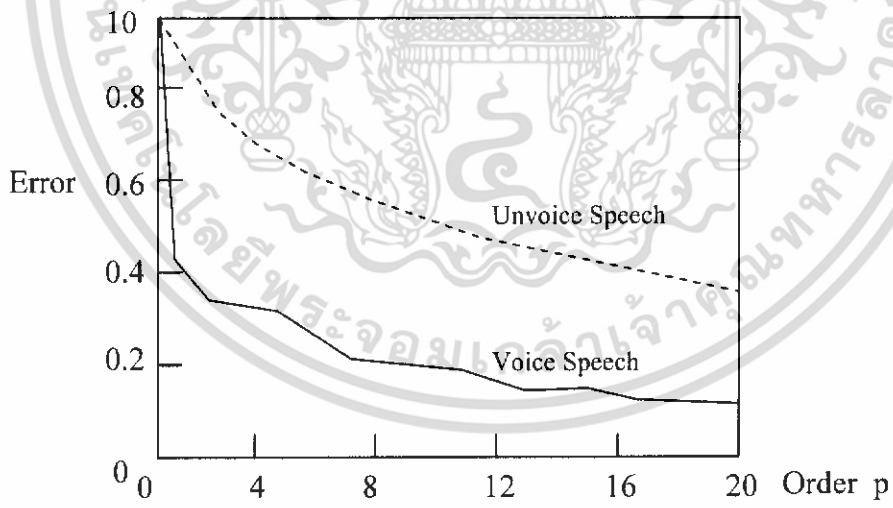
2.4.1 สัมประสิทธิ์ของ การประมาณพันธะเชิงเส้น

การประมาณพันธะเชิงส้น หรือ การประมาณเชิงเส้น มีความสำคัญคือ การที่รูปคลื่นและลักษณะสมบัติทางความถี่ของเสียงพูดสามารถแสดงด้วยค่าพารามิเตอร์ไม่กี่ค่าได้อย่างแม่นยำและมีประสิทธิภาพ นอกจากนี้พารามิเตอร์ดังกล่าวยังสามารถคำนวณหาได้ง่ายอีกด้วย

หลักการพื้นฐานของการประมาณเชิงเส้นคือ การประมาณค่าสัญญาณจากผลรวมเชิงเส้นของสัญญาณก่อนหน้านี้ สมมุติว่าสัญญาณเดิมเป็น $s(n)$ การประมาณค่าสัญญาณเป็น $s'(n)$ ดังนั้นเราสามารถอธิบาย การประมาณเชิงเส้น ได้ดังสมการต่อไปนี้

$$s'(n) = \sum_{k=1}^{p} \alpha_k s(n-k) \qquad (2.3)$$

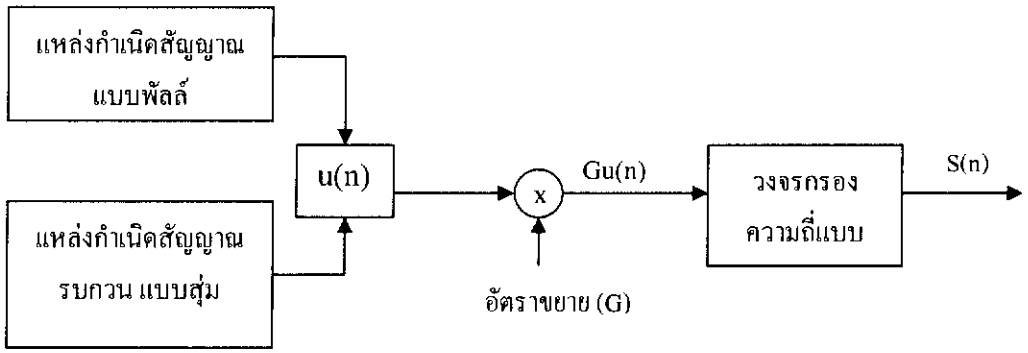
เมื่อ α_k เป็นค่าคงที่ เรียกวิธีการนี้ว่าการประมาณเชิงเส้นอันดับ p โดยมีเงื่อนไขว่า ค่า α_k ที่ใช้ในการประมาณจะต้องทำให้ ผลรวมของกำลังสองของความคลาดเคลื่อน $\{s(n)-s'(n)\}^2$ มีค่าน้อยที่สุด นั่นคือ $\sum e^2(n) = \sum \{s(n)-s'(n)\}^2$ ที่ค่าต่ำที่สุด แสดงดังรูปที่ 2.12 เห็นได้ว่า ที่ p ประมาณ 19 เป็นต้นไปนั้น จะมีค่าความผิดพลาดจะค่อนข้างคงที่ ซึ่งการประมาณมีอยู่หลายวิธีได้แก่ วิธีโควาเลียนซ์ วิธีออโตคอร์รีเลชัน วิธีแลตทิส วิธีวงจรกรองผกผัน เป็นต้น ซึ่งในที่นี้จะใช้การประมาณเชิงเส้น วิธีออโตคอร์รีเลชั่น (Autocorrelation Method) หรือ วิธีอัตตสัมพันธ์



รูปที่ 2.12 แสดงความสัมพันธ์ระหว่างค่าความผิดพลาดกับจำนวนอันดับของการประมาณเชิงเส้น

จากหลักการพื้นฐานของการประมาณเชิงเส้น และแบบจำลองระบบสร้างสัญญาณเสียง เราสามารถเขียน บล็อกไดอะแกรม การทำการประมาณเชิงเส้น มาสร้างสัญญาณเสียงพูด ได้ดังรูปที่ 2.13

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.13 แสดงบล็อคไดอะแกรมแสดงโมเดลการสร้างสัญญาณเสียงพูดอย่างง่าย

จากรูป 2.13 สามารถเขียนสมการได้เป็น

$$s(n) = Gu(n) + \sum_{k=1}^{p} \alpha_k s(n-k) \tag{2.4}$$

การประมาณเชิงเส้นโดนใช้สัมประสิทธิ์ (α_k) คือ

$$s'(n) = \sum_{k=1}^{p} \alpha_k s(n-k) \tag{2.5}$$

ดังนั้นความคลาดเคลื่อน คือ

$$\begin{aligned} e(n) &= s(n) - s'(n) \\ &= s(n) - \sum_{k=1}^{p} \alpha_k s(n-k) \end{aligned} \tag{2.6}$$

ฟังก์ชั่นถ่ายโอนระหว่าง e(n) และ s(n) คือ

$$\begin{aligned} A(Z) &= \frac{E(Z)}{S(Z)} \\ &= 1 - \sum_{k=1}^{p} \alpha_k z^{-k} \end{aligned} \tag{2.7}$$

จะเห็นได้ว่า ถ้า $\alpha_k = a_k$ แล้ว

$$e(n) = Gu(n) \tag{2.8}$$

ในการประมวลผลสัญญาณจะมีการแบ่งช่วงสัญญาณออกเป็นช่วงสั้น ๆ (Windowing) ช่วงละประมาณ 20-40 มิลลิวินาที ดังนั้น ค่าผลรวมของกำลังสองของความคลาดเคลื่อน

$$\begin{aligned} E_n &= \sum_m e_n^2(m) \\ E_n &= \sum_m \left[s(m) - s'(m) \right]^2 \end{aligned} \tag{2.9}$$

โดยที่ n คือ ช่วงที่ n ของสัญญาณที่ใช้คำนวณ เพราะฉะนั้นเพื่อให้ได้ค่า E_n ต่ำที่สุดจะต้องมีเงื่อนไขว่า

$$\frac{\partial E_n}{\partial \alpha_i} = 0 \qquad \text{เมื่อ } i = 1,2,3,\ldots,p$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากสมการ (2.9)

$$\frac{\partial E_n}{\partial \alpha_i} = -2s_n(m-1)\sum_m \left[s_n(m) - \sum_{k=1}^{p} \alpha_k s_n(m-k) \right] \quad \text{เมื่อ i= 1,2,3,...,p}$$

$$= -2\left[\sum_m s(m)s(m-i) - \sum_{k=1}^{p} \sum_m \alpha_k s(m-k)s(m-i) \right]$$

$$\frac{\partial E_n}{\partial \alpha_i} = 0 \quad \text{ก็ต่อเมื่อ}$$

$$\sum_{k=1}^{p} \alpha_k \sum_m s_n(m-k)s_n(m-i) = \sum_m s_n(m)s_n(m-i) \quad \text{เมื่อ i= 1,2,3,...,p} \qquad (2.10)$$

ถ้าเรากำหนดให้ $\phi_n(i,j) = \sum s_n(m-k)s_n(m-i)$ เพราะฉะนั้น

$$\sum_{k=1}^{p} \alpha_k \phi_n(i,j) = \phi_n(i,0) \qquad (2.11)$$

โดยสมการ (2.9) – (2.10) จะได้ว่า

$$E_n = \sum_m s_n^2(m) - \sum_{k=1}^{p} \alpha_k \sum_m s_n(m)s_n(m-k)$$

และจาก $\phi_n(i,j) = \sum s_n(m-k)s_n(m-i)$

$$E_n = \phi_n(0,0) - \sum_{k=1}^{p} \alpha_k \phi_n(0,k) \qquad (2.12)$$

สมมุติว่าใน 1 เฟรม ของสัญญาณ ที่ตัดมาคำนวณมี N ตัวอย่าง คือ $s_n(0), s_n(1), s_n(2), ..., s_n(N-1)$
ในที่นี้เราให้ $s_n(m) = 0$ เมื่อ m < 0 หรือ m > N-1 เพราะฉะนั้น

$$\phi_n(i,j) = \sum_m s_n(m-k)s_n(m-i)$$

$$= \sum_{m=0}^{N-1-(1-k)} s_n(m)s_n(m+1-k) \qquad 0 \le k \le p, 1 \le i \le p$$

ให้ $R_n(k) = \sum_m^{N-1-k} s_n(m)s_n(m+k)$ เมื่อ k = 0,1,2,...,p (2.13)

ดังนั้น จากสมการ (2.12) และ (2.13) จะได้ว่า

$$R_n(k) = \sum_{m=0}^{N-1-k} x'(m)x'(m+k) \qquad (2.14)$$

จากสมการ (2.11) จะได้ว่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\sum_{k=1}^{p} \alpha_k R_n(|i-k|) = R_n(i) \qquad \text{เมื่อ } i = 1,2,3\ldots,p \qquad (2.15)$$

เขียนอยู่ในรูปเมทริกซ์ได้เป็น

$$\begin{bmatrix} R_n(0) & R_n(1) & \ldots & R_n(p-1) \\ R_n(1) & R_n(0) & \ldots & R_n(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_n(p-1) & R_n(p-2) & \ldots & R_n(0) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix} = \begin{bmatrix} R_n(1) \\ R_n(2) \\ \vdots \\ R_n(p) \end{bmatrix} \qquad (2.16)$$

ให้ $$R_n \cdot \alpha = r_n \qquad (2.17)$$

เมื่อ $$R_n = \begin{bmatrix} R_n(0) & R_n(1) & \ldots & R_n(p-1) \\ R_n(1) & R_n(0) & \ldots & R_n(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_n(p-1) & R_n(p-2) & \ldots & R_n(0) \end{bmatrix}, \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix}, r_n = \begin{bmatrix} R_n(1) \\ R_n(2) \\ \vdots \\ R_n(p) \end{bmatrix}$$

2.4.2 การหาค่าสัมประสิทธิ์ LPC (α) และอัตราขยาย (G)

เมื่อได้ค่า $R_n(0), R_n(1), R_n(2), \ldots, R_n(p)$ แล้ว ก็สามารถหาค่าสัมประสิทธิ์ LPC (α) ได้ นั่นคือ

$$\alpha = R_n^{-1} \cdot r_n \qquad (2.18)$$

และ จะได้ว่า

$$e(n) = Gu(n)$$

$$E_n = \sum_{m=0}^{N-1} e^2(m) = G \sum_{m=0}^{N-1} u^2(m)$$

$$E_n = \phi_n(0,0) - \sum_{k=1}^{p} \alpha_k \phi_n(0,k)$$

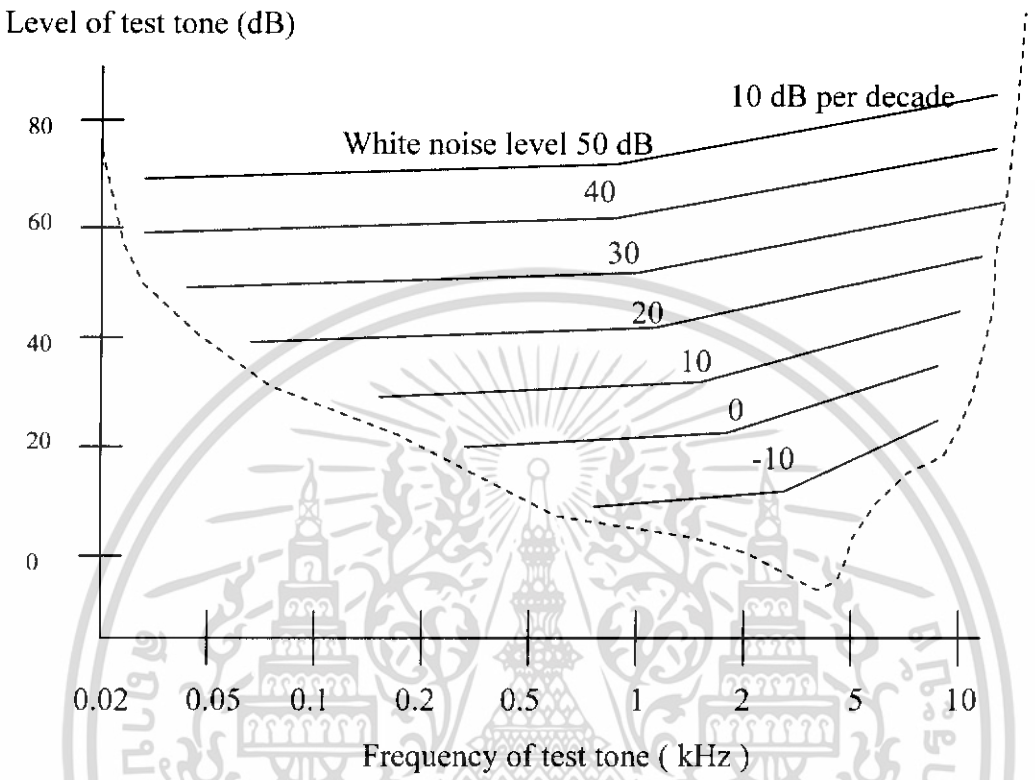
$$E_n = R_n(0) - \sum_{k=1}^{p} \alpha_k \phi_n(k) \qquad (2.19)$$

2.5 แถบความถี่วิกฤต และความเข้มสัญญาณในแถบความถี่วิกฤต

คำจำกัดความของ แถบความถี่วิกฤตนั้นได้ถูกกล่าวขึ้นโดย Fletcher โดยที่เขาได้ตั้งสมมุติฐานไว้ว่า ในกรณีที่สัญญาณรบกวน (noise) ซ้อนทับจนสามารถบดบังสัญญาณโทนทดสอบ (Test tone) ได้นั้นส่วนของสัญญาณที่มีผลต่อการบดบังสัญญาณโทนทดสอบจริงๆจะเป็นส่วนที่มีความถี่ใกล้เคียงกับสัญญาณโทนทดสอบเท่านั้น ในการที่จะทำการหาค่าสัมประสิทธิ์ต่างของแถบความถี่วิกฤตได้นั้น ยังต้องอาศัยสมมุติฐาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่ว่า การบดบังสัญญาณจะเป็นผลก็ต่อเมื่อกำลังงานของสัญญาณรบกวนในแถบสัญญาณเดียวกับสัญญาณโทนทดสอบนั้นเท่ากับกำลังงานของสัญญาณโทนทดสอบ และถือว่ากำลังงานของสัญญาณรบกวนที่อยู่นอกเหนือจากแถบสัญญาณที่ใกล้กับสัญญาณโทนทดสอบไม่มีผลต่อการบดบังสัญญาณ



รูปที่ 2.14 ระดับของสัญญาณโทนทดสอบที่ถูกบดบังโดยสัญญาณ white noise ที่ระดับสัญญาณต่างๆและเส้นประเป็นระดับ Threshold in quiet

2.5.1 สเกลบาร์คหรือ Critical Band Rate Scale

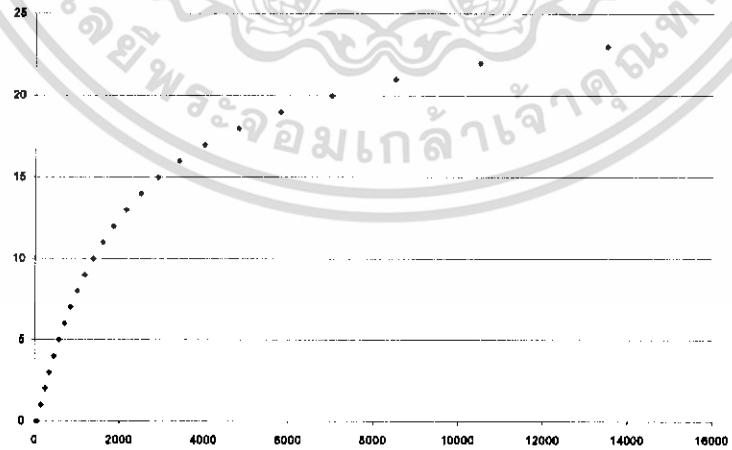
แถบความถี่วิกฤตนั้นมีความสำคัญต่อการอธิบายลักษณะในการได้ยินของมนุษย์อย่างมาก โดยจะถูกใช้ในแบบจำลองต่าง ๆ ในรูปแบบของสเกลที่เรียกว่า Critical Band Rate Scale ซึ่งสเกลนี้อ้างอิงจากหลักความจริงที่ว่าระบบการรับฟังเสียงของคนเรานั้นจะทำการแบ่งการวิเคราะห์เสียงออกเป็นส่วนต่าง ๆ ตามแถบความถี่วิกฤต และถ้ายึดเอาแถบความถี่วิกฤตแรกเป็นจุดอ้างอิงแล้วทำการคำนวณแถบความถี่วิกฤตต่อมาเรื่อย ๆ โดยให้ขอบเขตบนของแถบความความถี่วิกฤตปัจจุบันเป็นขอบเขตล่างของแถบความถี่วิกฤตถัดไปก็จะทำให้ได้สเกลที่เรียกว่า Critical Band Rate Scale ซึ่งสามารถแสดงได้ดังตารางที่ 2.1 ส่วนในรูปที่ 2.15 เป็นแผนภูมิที่นำข้อมูลในตารางที่ 2.1 มาแสดง ซึ่งใน Critical Band Rate Scale นั้นอันดับของ Critical Band ถูกเรียกเป็น “Bark”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

ตารางที่ 2.1 ขอบเขตและความถี่ศูนย์กลาง Critical Band Rate Scale

Z Bark	flower/fupper Hz	fc Hz	fc Bark	Bandwidth Hz	Z Bark	flower/fupper Hz	fc Hz	fc Bark	Bandwidth Hz
0	0 100	50	0.5	100	12	1720 2000	1850	12.5	280
1	100 200	150	1.5	100	13	2000 2320	2150	13.5	320
2	200 300	250	2.5	100	14	2320 2700	2500	14.5	380
3	300 400	350	3.5	100	15	2700 3150	2900	15.5	450
4	400 510	450	4.5	110	16	3150 3700	3400	16.5	550
5	510 630	570	5.5	120	17	3700 4400	4000	17.5	700
6	630 770	700	6.5	140	18	4400 5300	4800	18.5	900
7	770 920	840	7.5	150	19	5300 6400	5800	19.5	1100
8	920 1080	1000	8.5	160	20	6400 7700	7000	20.5	1300
9	1080 1270	1170	9.5	190	21	7700 9500	8500	21.5	1800
10	1270 1480	1370	10.5	210	22	9500 12000	10500	22.5	2500
11	1480 1720	1600	11.5	240	23	12000 15500	13500	23.5	3500



รูปที่ 2.15 แสดงความสัมพันธ์ระหว่างลำดับของ Bark Scale กับความถี่ Hz

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

62724

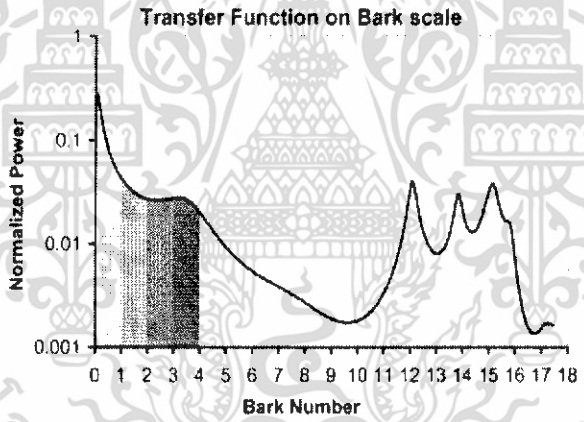
2.5.2 Critical Band Intensity

ระดับการตอบสนองต่อเสียงที่ความถี่ใด ๆ ของระบบการได้ยินเสียงของมนุษย์นั้น สามารถคาดคะเนได้จากการหาความเข้มของสัญญาณเสียงในแต่ละ Critical Band โดยสามารถคำนวณได้จากสเปกตรัมบนสเกลบาร์คดังสมการต่อไปนี้

$$I_B = \int_{B-05\,Bark}^{B+0.5\,Bark} \frac{dI}{dB} dB \qquad (2.22)$$

โดย B คือ Bark Number

ซึ่งการหา Critical Band Intensity จากสมการที่ 2.22 ก็คือพื้นที่ใต้กราฟของแต่ละแถบความถี่ใน Transfer function บนสเกลบาร์ค ดังรูปที่ 2.16 ซึ่งในปริญญานิพนธ์นี้เสียงจะถูกสุ่มที่ความถี่ 11.025 KHz ดังนั้นความถี่สูงสุดของเสียงที่สุ่มมาได้คือ 5.50125 KHz ซึ่งอยู่ในบาร์คที่ 18 ดังนั้น Critical Band Intensity ทั้ง 18 ตัว (บาร์คที่ 1 ถึง 18) จะถูกนำไปใช้เป็นองค์ประกอบของ Feature Vector 18 มิติ ที่ใช้เป็นตัวแทนของเสียงที่ใช้ในการทดลอง



รูปที่ 2.16 แสดงการคำนวณ Critical Band Intensity

2.6 การจำแนกรูปแบบด้วยแบบจำลองฮิดเด็นมาร์คอฟ (Hidden Markov Models)

แบบจำลองฮิดเด็นมาร์คอฟ

คำพูดเสียงเป็นสิ่งที่แสดงถึงคำหนึ่งคำหรือหลาย ๆ คำ ที่มีลักษณะการเปลี่ยนแปลงของสัญญาณความถี่(spectral envelop) ที่ช้า มนุษย์ได้ยินสัญญาณความถี่นี้และเปลี่ยนมันให้เป็นคำและสามารถเข้าใจถึงความหมายได้ จุดประสงค์ของการประมวลเสียงคำพูดคือการลอกเลียนระบบดังกล่าวของมนุษย์เพื่อที่จะสร้างเครื่องมือสำหรับพูดคุยกับมนุษย์เองได้ ก้าวแรกของการเริ่มต้นระบบนี้คือการแปลงคำพูดที่ได้ยินหรือสัญญาณเสียงให้เป็นเครื่องหมายที่เหมาะสม เครื่องหมายที่เหมาะสมในที่นี้คือลักษณะพิเศษที่สามารถหาได้จากสัญญาณเสียง เช่น ความถี่มูลหลัก (fundamental frequency) สัญญาณความถี่ (spectral envelop) เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีการหนึ่งที่นิยมนำมาใช้และให้ผลดีที่สุดในขณะนี้ในการแปลงคำพูดให้เป็นเครื่องหมายคือการใช้วิธีการสถิติของแบบจำลองฮิดเด็นมาร์คอฟ

2.6.1 ความหมายของแบบจำลองฮิดเด็นมาร์คอฟ

แบบจำลองฮิดเด็นมาร์คอฟสามารถสร้างโมเดลคำหรือหน่วยเสียงก็ได้ แต่เหตุผลของการที่จะโมเดลว่าให้เป็นคำหรือหน่วยเสียงจะขึ้นอยู่กับการใช้งานของระบบการรู้จำเสียง เช่น ถ้าระบบรู้จำเสียงมีข้อจำกัดคือให้รู้จำคำศัพท์ที่น้อยและรู้จำคำพูดที่ไม่ติดต่อกัน การโมเดลเสียงให้เป็นคำเสียงจะง่ายกว่าโมเดลหน่วยเสียงเนื่องจากการโมเดลเป็นคำจะให้ความแม่นยำมากกว่าและวิธีการของระบบรู้จำเสียงจะง่ายกว่าการโมเดลหน่วยเสียง การโมเดลหน่วยเสียงจะต้องการโมเดลแต่ละหน่วยเสียงและจะต้องรู้จำหน่วยเสียงนั้น ๆ ก่อน เมื่อรู้จำหน่วยเสียงได้แล้วก็จะต้องมีวิธีการสร้างคำจากหน่วยเสียงที่ได้รู้จำมา ถ้ามีการรู้จำหน่วยเสียงหนึ่งผิดไปก็จะทำให้คำคำนั้นรู้จำผิดได้ แต่สำหรับการสร้างโมเดลคำเสียง ความแม่นยำในการรู้จำจะขึ้นโดยตรงกับโมเดลคำเสียงที่ได้ฝึกฝนมาและไม่ต้องการวิธีสร้างคำจากหน่วยเสียง แต่สำหรับระบบที่มีคำศัพท์มาก(มากกว่า1000คำ)และรู้จำเสียงแบบพูดติดต่อกันระบบรู้จำที่มีโมเดลหน่วยเสียงจะมีประสิทธิภาพมากกว่าเพราะจำนวนโมเดลของหน่วยเสียงที่ต้องการจะมีน้อยกว่าโมเดลของคำเสียงและทำให้ระบบรู้จำเสียงใช้เวลาในการค้นหาคำเสียงที่มีความน่าจะเป็นมากที่สุดมากกว่า อีกประการหนึ่งที่สำคัญก็คือการรู้จำเสียงใช้เวลาในการค้นหาคำเสียงที่มีความน่าจะเป็นมากที่สุดมากกว่า อีกประการหนึ่งที่สำคัญก็คือการรู้จำคำพูดแบบติดต่อกัน มีความเป็นไปได้สูงที่คำหนึ่งคำจะเปลี่ยนไปเมื่อใช้กับคำ ๆ อื่น ถ้าระบบมีการฝึกฝนที่ไม่ดีก็จะทำให้ระบบรู้จำมีความแม่นยำต่ำ แต่สำหรับหน่วยเสียงจะมีการเปลี่ยนแปลงน้อยกว่าเมื่อใช้กับคำคำอื่นจึงทำให้มีความแม่นยำมากกว่า

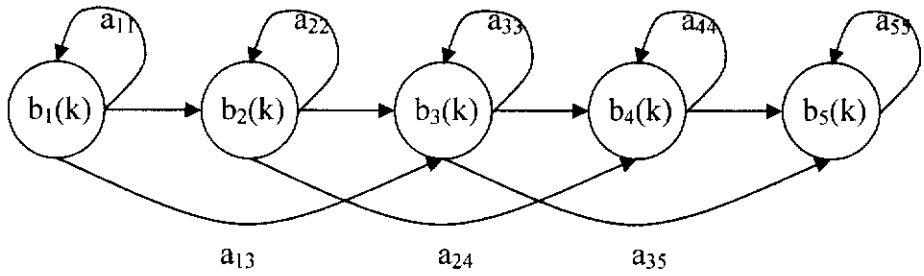
ตามทฤษฎีแบบจำลองฮิดเด็นมาร์คอฟเป็นแบบจำลองทางสถิติ คือ การเก็บรวบรวมของสถานะหลาย ๆ สถานะ ที่ถูกเชื่อมโยงโดยการเปลี่ยนสถานะ ในที่นี้คือ ลักษณะของสัญญาณเสียงหรือเครื่องหมายที่หมายถึงสัญญาณเสียงนั้น ณ เวลาหนึ่ง ๆ ส่วนการเปลี่ยนสถานะคือการเลื่อนหรือเปลี่ยนแปลงของสัญญาณเสียงจากเวลาหนึ่งไปหาอีกเวลาหนึ่ง แบบจำลองฮิดเด็นมาร์คอฟมีค่าความน่าจะเป็นอยู่ 2 ชนิดคือ ค่าความน่าจะเป็นที่การเปลี่ยนสถานะหนึ่งที่จะเกิดขึ้น(transition probability) และ ค่าความน่าจะเป็นของผลลัพธ์ (output symbol) เมื่อมีการเปลี่ยนสถานะหนึ่งเกิดขึ้น ผลลัพธ์ที่ว่านี้จะนำไปใช้ในการตัดสินใจว่าเสียงที่ได้ยินนั้นเป็นเสียงอะไร (output probability)

แบบจำลองมาร์คอฟ แบ่งออกเป็น 2 ประเภท คือ แบบต่อเนื่อง (Continuous) และแบบไม่ต่อเนื่อง (Discrete-time) ในปริญญานิพนธ์นี้ได้เลือกใช้แบบไม่ต่อเนื่อง เพราะเป็นวิธีการที่ซับซ้อนน้อยกว่าและใช้งานได้ดีกับคำพูดสั้น ๆ ที่ได้ทำการทดลองในปริญญานิพนธ์นี้

2.6.2 กระบวนการแบบจำลองมาร์คอฟ แบบไม่ต่อเนื่องทางเวลา

รูปที่ 2.17 เป็นรูปที่แสดงระบบที่มี 5 สเตท (state) การอธิบายระบบ ณ เวลาหนึ่ง ๆ จะอธิบายด้วยเซตของดัชนีสเตท {1,2,…,N} โดยในแต่ละช่วงเวลาสามารถเปลี่ยนสเตทไป ด้วยเซตของความน่าจะเป็นที่เกี่ยวข้องกับสเตท

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.17 แสดงตัวอย่างแบบจำลองมาร์คอฟ 5 สเตท

เราจะแทนค่าเวลาที่เกี่ยวข้องกับการเปลี่ยนสเตทไป ด้วย t = 1,2,...T และแทนสเตทที่เวลา t ด้วย q_t ในห่วงโซ่ Markov (Markov Chain) นั้นจะแสดงความน่าจะเป็นในการย้ายสเตทไป โดยให้ความสนใจเพียงสเตทก่อนหน้านั้น เท่านั้น คือ

$$P[q_t = j \mid q_{t-1} = k,..] = P[q_t = j \mid q_{t-1} = i] \qquad (2.31)$$

และ a_i จะแทนความน่าจะเป็นของการย้ายสเตทจาก i ไป j

$$a_i = P[q_t = j \mid q_{t-1} = i] \qquad 1 \le i, j \le N \qquad (2.32)$$

โดยมีคุณสมบัติว่า

$$a_{ij} \ge 0 \qquad \forall_{i,j}$$

$$\sum_{j=1}^{N} a_{ij} = 1 \qquad \forall_i \qquad \text{(ผลรวมความน่าจะเป็นทั้งหมดที่ออก=1)} \qquad (2.33)$$

ถ้าหากเอาท์พุทของกระบวนการเป็นเซตของสเตท ที่แต่ละเวลาหนึ่งซึ่งแต่ละ สเตท หมายถึงเหตุการณ์ที่สังเกตได้แล้วจะเรียกมาร์คอฟแบบจำลองนี้ว่า มาร์คอฟแบบจำลองที่สังเกตเห็นได้ (Observable Markov Model)

ส่วนแบบจำลองมาร์คอฟนั้น เหตุการณ์จะเป็นฟังก์ชันของความน่าจะเป็นของสเตท คือ จะไม่ทราบเหตุการณ์ที่จะเกิดในแต่ละสเตทโดยตรง แต่จะรู้เหตุการณ์ได้โดยวิธีทางสถิติ ซึ่งจะทำให้เราได้ลำดับของเหตุการณ์ออกมา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6.3 ส่วนประกอบของแบบจำลองมาร์คอฟ

1. N: จำนวนสเตทในแบบจำลอง โดยทั่วไปแล้วสเตทจะเกี่ยวข้องซึ่งกันและกัน ในแง่ที่ว่า สเตทจะสามารถย้ายมาได้จากสเตทอื่น ๆ เราจะให้เซทของสเตทเป็น {1,2,...,N} และแทนสเตทที่เวลา t เป็น q_t

2. M: จำนวนของเหตุการณ์ต่าง ๆ ต่อหนึ่งสเตท เราจะแทนแต่ละสัญลักษณ์ เป็น $V = \{V_1, V_2, ..., V_M\}$

3 ความน่าเป็นของการย้ายสเตท A=$\{a_{ij}\}$ เมื่อ $a_i = P[q_t = j \mid q_{t-1} = i] \quad 1 \le i, j \le N$ สำหรับกรณีพิเศษที่สเตทใด ๆ จะย้ายไปยังสเตทอื่น ๆ เราจะได้ $a_{ij} \ge 0$ สำหรับทุก ๆ i,j สำหรับแบบจำลองมาร์คอฟ ชนิดอื่น ๆ เราจะได้ $a_{ij} = 0$ สำหรับคู่ {i,j} หนึ่งคู่หรือมากกว่า

4. การกระจายความน่าจะเป็นของสัญลักษณ์ที่สังเกตได้ $B = \{b_j(k)\}$

โดย $b(k) = P[o_t = v_t \mid q_t = j], 1 \le k \le m$ (2.34)

เป็นนิยามการกระจายสัญลักษณ์ในสเตท j ; j = 1,2,...,N

จะเห็นได้ว่า แบบจำลองมาร์คอฟ ต้องการพารามิเตอร์ของแบบจำลอง 2 ตัว คือ N และ M และกลุ่มของความน่าจะเป็น A,B,π และเพื่อความสะดวกจะใช้สัญลักษณ์เป็น { A,B,π } เพื่อแสดงเซตของพารามิเตอร์ที่สมบูรณ์ของแบบจำลอง

2.6.4 ชนิดของ HMM

แบ่งชนิดตามลักษณะของการย้ายสเตท (Transition) ของเมตริกซ์ A ได้ 3 ชนิด

1. HMM แบบ Egordic Model หรือ Fully Connected Model : การย้ายสเตทสามารถย้ายไปยังทุกๆ สเตทของแบบจำลอง

2. HMM แบบ Left-Right Model หรือ Bakis Model : การย้ายสเตทจะย้ายจากซ้ายไปขวาซึ่งจะมีคุณสมบัติของสัมประสิทธิ์การย้ายสเตทดังนี้

$$Wj = 0 \quad ; j < i$$

คือจะไม่มีการย้ายสเตทที่ต่ำกว่าสเตทปัจจุบัน และนอกจากนี้ก็ยังมีความน่าจะเป็นของสเตทเริ่มต้นดังนี้

$$\pi_r = \begin{cases} 0, i \neq 1 \\ 1, i = 1 \end{cases}$$

คือ ลำดับของสเตทต้องเริ่มที่สเตท 1 และ Left-Right Model นี้มักจะมีกฎบังคับการย้ายสเตท เพื่อไม่ให้มีการเปลี่ยนแปลงดัชนีของ สเตทมากนัก กล่าวคือ

$$a_{ij} = 0 \qquad i > i + \Delta i$$

ค่าของ $\Delta i = 2$ คือจะไม่กระโดด (ย้าย) ไปเกิน 2 สเตท จะได้เมตริกซ์การย้าย สเตทเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

จะเห็นว่าสเตทสุดท้าย สัมประสิทธิ์การย้ายสเตทจะเป็น

$$a_{NN} = 1$$
$$a_{Ni} = 0, i < N$$

แบบจำลองแบบนี้จะเหมาะกับสัญญาณที่มีลักษณะเปลี่ยนแปลงตามเวลาอย่างต่อเนื่อง เช่น เสียงพูด

3. HMM แบบ parallel Left-Right Model : เป็นแบบจำลอง ที่มีความยืดหยุ่นมากกว่าแบบที่ 2 Left-Right Model

2.6.5 ปัญหาพื้นฐานของแบบจำลองมาร์คอฟ

2.6.5.1 ปัญหาที่ 1

เมื่อกำหนดลำดับเหตุการณ์ O = $\{o_1 o_2 \ldots o_t\}$ และแบบจำลอง λ = { A,B, π } จะคำนวณหา P(O, λ) ซึ่งเป็นความน่าจะเป็นของลำดับเหตุการณ์นี้ได้อย่างไร การแก้ปัญหานี้จะสามารถทำให้เราเลือกแบบจำลองที่สอดคล้องกับเหตุการณ์ที่ดีที่สุดได้จากแบบจำลองหลายอันที่มีอยู่

2.6.5.2 ปัญหาที่ 2

เมื่อกำหนดลำดับของเหตุการณ์ O = $\{o_1 o_2 \ldots o_t\}$ และแบบจำลอง λ = { A,B, π } เราจะเลือกลำดับสเตท q = $\{q_1 q_2 \ldots q_t\}$ ที่เหมาะกับเหตุการณ์ที่สุดได้อย่างไร (อธิบายเหตุการณ์ได้ดีที่สุด) การแก้ปัญหานี้ช่วยหาลำดับของสเตทที่ถูกต้อง

2.6.5.3 ปัญหาที่ 3

เราจะปรับพารามิเตอร์ของแบบจำลอง λ = { A,B, π } เพื่อให้ค่าสูงสุดของ P(O| λ) ได้อย่างไร การแก้ปัญหานี้ช่วยให้ได้ λ ที่เหมาะที่สุด เพื่ออธิบายการได้มาซึ่งลำดับเหตุการณ์ที่กำหนดให้อย่างดีที่สุด

ลำดับเหตุการณ์ที่ใช้ปรับค่าพารามิเตอร์ของแบบจำลอง เรียกว่าลำดับเทรนนิ่ง (training sequence) เนื่องจากเป็นลำดับเหตุการณ์ที่ถูกใช้เพื่อเป็นเทรนนิ่ง แบบจำลองมาร์คอฟ และด้วยการแก้ปัญหานี้เราจะได้แบบจำลองที่ดีที่สุดสำหรับคำนั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่าง สำหรับตัวรู้จำเสียงพูด (isolate word speech recognizer) ง่าย ๆ เช่น สำหรับคำว่า "หนึ่ง" เราต้องออกแบบแบบจำลองมาร์คอฟ แบบ N สเตท เราจะแทนสัญญาณคำพูดของคำด้วยลำดับเวลาของเวคเตอร์สเปคตรัม สมมุติว่าข้อมูลได้ถูกเข้ารหัสมาแล้วโดยใช้ สเปคตรัมโค้ดบุค (spectrum codebook) ที่มีเวคเตอร์สเปคตรัมอยู่ที่ M แบบ นั่นคือ แต่ละเหตุการณ์แทนด้วยดัชนีของเวคเตอร์สเปคตรัมที่ใกล้เคียงกับสัญญาณเสียงเดิมมากที่สุด ดังนั้นแต่ละคำศัพท์จะแทนด้วยลำดับเทรนนิ่ง ซึ่งประกอบด้วย ลำดับของดัชนีโค้ดบุคของคำจำนวนมาก สิ่งแรกที่ต้องทำคือการสร้างแบบจำลองของแต่ละคำ โดยใช้ การแก้ปัญหาที่ 3 หาพารามิเตอร์ λ ของแบบจำลองที่เหมาะสมที่สุด สำหรับแต่ละแบบจำลองของคำ จากนั้นใช้การแก้ปัญหาที่ 2 เพื่อแบ่งแต่ละลำดับเทรนนิ่งของคำไปยังสเตทต่าง ๆ และเรียนรู้ถึงคุณสมบัติของเวคเตอร์สเปคตรัม ซึ่งจะทำให้ทราบถึงเหตุการณ์ที่เกิดขึ้นในแต่ละสเตท จุดประสงค์ก็เพื่อทำให้ แบบจำลองมีความสามารถในการจำลองเสียงพูดเข้าไปนั้นเอง

เมื่อได้แบบจำลองมาร์คอฟ ของคำว่า "หนึ่ง" ที่เหมาะสมแล้ว ก็จะใช้ การแก้ปัญหาที่ 1 ในการหาความน่าจะเป็นไปได้ของแบบจำลองในการเกิดเหตุการณ์ ตามลำดับของเหตุการณ์เสียง ที่ไม่รู้ว่าเป็นเสียงใดเข้ามา ถ้าแบบจำลองใดของคำว่า "หนึ่ง" ให้ค่าความน่าจะเป็นของคำสูงกว่าแบบจำลองของคำอื่น ๆ ("สอง", "สาม" , "สี่" ,...) แสดงว่าคำที่ไม่ทราบว่าเป็นเสียงใดนี้คือคำว่า "หนึ่ง" นั่นเอง

2.6.6 ระบบแบบจำลองมาร์คอฟ

เมื่อมีคำศัพท์อยู่ V คำ ในการทำการรู้จำได้ จะต้องทำการสร้างแบบจำลองของคำแต่ละคำด้วยแบบจำลองมาร์คอฟ ที่แตกต่างกัน คำแต่ละคำจะมีลำดับเทรนนิ่งที่ได้มาจากคุณลักษณะเฉพาะของคำศัพท์นั้นๆ

การที่เรารู้จำคำพูดได้ ต้องทำดังนี้

1. สำหรับแต่ละคำศัพท์ V ต้องสร้างแบบจำลองมาร์คอฟ: λ_v นั่นคือ การหาค่า(A, B, π)ที่เหมาะสมกับลำดับเทรนนิ่งของคำ V นั้น

2. ในการจะรู้จำคำศัพท์แต่ละคำ หลังจากที่มีลำดับเหตุการณ์ของคำพูดเข้ามา (เสียงพูดที่ผ่าน LPC และ การจัดระดับเวกเตอร์มาแล้ว) จะทำการหาค่าความน่าจะเป็น ของทุกๆแบบจำลอง แล้วเลือกแบบจำลองที่มีค่าความน่าจะเป็นในการเกิดเหตุการณ์สูงสุด

คำศัพท์ที่สอดคล้องกับแบบจำลองดังกล่าวนี้ จะเป็นคำศัพท์คำเดียวกับคำศัพท์ที่ต้องการจะรู้จำนั่นเอง โดยขั้นตอนการคำนวณหาค่าความน่าจะเป็น จะใช้วิธีเวอเทอร์บีดังที่ได้กล่าวแล้วในตอนต้น

2.7 โครงสร้างและสถาปัตยกรรมของไมโครคอนโทรลเลอร์ MCS-51

ไมโครคอนโทรลเลอร์ตระกูล MCS-51 ที่ใช้ในการทดลองจะใช้ไมโครคอนโทรเลอร์ตระกูล MCS-51 ซึ่งมีหน่วยความจำโปรแกรมภายในเป็นแบบแฟลชของ Atmel Corporation มีเบอร์ขึ้นต้นด้วย AT89

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.7.1 คุณสมบัติของไมโครคอนโทรเลอร์ตระกูล MCS-51 อนุกรม

- เป็นไมโครคอนโทรเลอร์ที่ใช้ซีพียูขนาด 8 บิต
- ภายในมีหน่วยความจำที่เป็นแบบแฟลชสามารถลบและเขียนใหม่ได้พันครั้ง
- หน่วยความจำข้อมูลพื้นฐานเป็นหน่วยความจำแบบแรมในบางเบอร์จะมีหน่วยความจำแบบอีอีพรอมเพิ่มเติม
- ขาพอร์ตเป็นแบบสองทิศทาง สามารถใช้งานเป็นได้ทั้งอินพุตและเอาต์พุต
- มีวงจรสื่อสารอนุกรมแบบฟูลดูเพล็กซ์
- ไทเมอร์/เคาน์เตอร์ขนาด 16 บิต อย่างน้อย 2 ตัว
- สามารถรองรับแหล่งกำเนิดอินเตอร์รัปต์ได้ 6 ประเภท
- สามารถขยายหน่วยความจำภายนอกเพิ่มเติมได้สูงสุด 64 กิโลไบต์
- มีวงจรกำเนิดสัญญาณนาฬิกาอยู่ภายในชิป
- มีวงจรสื่อสารอนุกรมแบบ SPI สำหรับในอนุกรม AT89Sxx
- มีวอตช์ด็อกไทเมอร์ในตัวสำหรับอนุกรม AT89Sxx

โครงสร้างของ AT89Cxx จะเหมือนกับไมโครคอนโทรเลอร์ตระกูล MCS-51 พื้นฐาน หากแต่จะแตกต่างกันเฉพาะหน่วยความจำแบบเฟลชที่เพิ่มเข้ามา หากเป็นไมโครคอนโทรลเลอร์อนุกรม 87xx หน่วยความจำโปรแกรมภายในจะเป็นแบบอีพรอมและบางเบอร์สามารถโปรแกรมได้เพียงครั้งเดียว

2.7.2 การจัดขาของไมโครคอนโทรเลอร์ MCS-51

ไมโครคอนโทรลเลอร์ MCS-51 ทุกตัวจะมีสถาปัตยกรรมและขาใช้งานพื้นฐานเหมือนกัน โดยมีรายละเอียดขั้นต้นดังนี้

ขา Vcc ใช้สำหรับต่อไฟเลี้ยง +5 V

ขา GND เป็นขากราวด์สำหรับต่อกับกราวด์ของระบบ

ขาพอร์ต 0 (P0.0-P0.7) มี 8 ขา แต่ละขาสามารถกำหนดให้เป็นได้ทั้งอินพุตและเอาต์พุตสำหรับใช้งานทั่วไป หากต้องการให้ขาพอร์ต 0 ขาใดขาหนึ่งเป็นอินพุตสามารถทำได้โดยการเขียนข้อมูล "1" ไปยังแต่ละบิตของพอร์ตที่ต้องการติดต่อด้วย ส่งผลให้ขานั้นมีสถานะปล่อยลอย (float) จึงมีอินพุตอิมพิแดนซ์สูง สามารถใช้งานเป็นขาพอร์ตอินพุตได้ นอกจากนั้นขาพอร์ตนี้ยังถูกใช้งานการติดต่อกับขาแอดเดรสไบต์ต่ำของหน่วยความจำภายนอก (A0-A7) และขาข้อมูล (D0-D7) โดยใช้กระบวนการมัลติเพล็กซ์เข้าช่วยเพื่อสลับการทำงานเป็นได้ทั้งขาติดต่อแอดเดรสและขาข้อมูล

ขาพอร์ต 1 (P1.0-P1.7) มี 8 ขาแต่ละขาสามารถกำหนดให้เป็นได้ทั้งอินพุตและเอาต์พุตสำหรับใช้งานทั่วไป

ขาพอร์ต 2 (P2.0-P2.7) มี 8 ขา แต่ละขาสามารถกำหนดให้เป็นทั้งอินพุตและเอาต์พุตสำหรับใช้งานทั่วไป นอกจากนี้ขาพอร์ตนี้ยังถูกใช้งานในการติดต่อกับขาแอดเดรสไบต์สูงของหน่วยความจำภายนอก (A8-A15)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขาพอร์ต 3 (P3.0-P3.7) มี 8 ขา แต่ละขาสามารถกำหนดให้เป็นทั้งอินพุตและเอาต์พุตสำหรับใช้งานทั่วไป นอกจากนี้ขาพอร์ต 3 ยังเป็นขาที่มีหน้าที่การใช้งานพิเศษ ดังมีรายละเอียดขั้นต้นต่อไปนี้

P3.0 ใช้เป็นขาอินพุตสำหรับรับข้อมูลจากการสื่อสารอนุกรม หรือขา RxD

P3.1 ใช้เป็นขาอินพุตสำหรับส่งข้อมูลจากการสื่อสารอนุกรม หรือขา TxD

P3.2 ใช้เป็นขาอินพุตสำหรับรับสัญญาณอินเตอร์รัปต์จากภายนอกช่อง 0 หรือขา $\overline{INT0}$

P3.3 ใช้เป็นขาอินพุตสำหรับรับสัญญาณอินเตอร์รัปต์จากภายนอกช่อง 1 หรือขา $\overline{INT1}$

P3.4 ใช้เป็นขาอินพุตสำหรับรับสัญญาณไทเมอร์จากภายนอกช่อง 0 หรือขา T0

P3.5 ใช้เป็นขาอินพุตสำหรับรับสัญญาณไทเมอร์จากภายนอกช่อง 1 หรือขา T1

P3.6 ใช้เป็นขาสัญญาณ $\overline{WR}$ ใช้ในกรณีที่ใช้เชื่อมต่อกับหน่วยความจำภายนอก

P3.7 ใช้เป็นขาสัญญาณ $\overline{RD}$ ใช้ในกรณีที่ใช้เชื่อมต่อกับหน่วยความจำภายนอก

ขารีเซต (Reset) ใช้ในการรีเซตการทำงานของไมโครคอนโทรลเลอร์

ขา $\overline{ALE}/PROG$ (Address Latch Enable/Program Pulse Input) เป็นขาที่ใช้ในการควบคุมการแลตช์ของขาพอร์ต 0 เมื่อมีการใช้งานหน่วยความจำภายนอก

ขา $\overline{PSTN}$ (Program Store Enable) ขานี้ใช้ในการส่งสัญญาณเพื่อร้องขอติดต่อกับหน่วยความจำโปรแกรมภายนอก

ขา $\overline{EA}/V_{pp}$ (External Access enable/Programming voltage input) ใช้สำหรับการเลือกติดต่อหน่วยความจำโปรแกรมจากภายนอกหรือภายในตัวไมโครคอนโทรเลอร์

ขา XTAL1 และ XTAL2 เป็นขาสำหรับต่อคริสตอลเพื่อสร้างสัญญาณนาฬิกาในการกำหนดจังหวะการทำงานของไมโครคอนโทรเลอร์

2.7.3 จังหวะการทำงานของไมโครคอนโทรเลอร์ MCS-51

ในการใช้งานไมโครคอนโทรเลอร์ MCS-51 จะต้องทำความเข้าใจถึงจังหวะการทำงานของซีพียูและลำดับขั้นตอนการประมวลผลคำสั่ง ในการประมวลผลคำสั่งของซีพียูนั้นจะมีขั้นตอนหลัก ๆ 2 ขั้นตอน คือ

1. กระบวนการเฟตช์ (fetch) เป็นการเรียกคำสั่งออกจากหน่วยความจำโปรแกรมแล้วทำการแปลงรหัสคำสั่งนั้นเป็นภาษาเครื่องเพื่อเตรียมการประมวลผล
2. กระบวนการเอ็กซิคิวต์ (execute) เป็นการกระทำตามคำสั่งที่กำหนดหรือตามที่เฟตช์ขึ้นมาโดยกระบวนการก่อนหน้านี้

เมื่อทำกระบวนการเอ็กซิคิวต์คำสั่งเรียบร้อยแล้วก็จะไปเริ่มกระบวนการเฟตช์คำสั่งต่อไป

เมื่อเริ่มจ่ายไฟให้แก่ไมโครคอนโทรเลอร์จะเกิดการรีเซตในลักษณะที่เรียกว่า พาวเวอร์ออนรีเซต (power on reset) ซีพียูเริ่มต้นทำงานที่แอดเดรส 0000H ของหน่วยความจำโปรแกรม จังหวะการทำงานของซีพียูจะเป็นไปตามรูปแบบ โดยได้รับการกำหนดมาจากรอบการทำงานหรือแมชชีนไซเคิล (machine cycle) จังหวะการทำงานของไมโครคอนโทรเลอร์ MCS-51 ใน 1 รอบการทำงานหรือ 1 แมชชีนไซเคิลจะแบ่งออกเป็น 6 สเตท (state) กำหนดชื่อเป็น S1-S6 ในแต่ละสเตทมีค่าเวลาเท่ากับ 2 คาบเวลาของสัญญาณนาฬิกา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถ้าสัญญาณนาฬิกามีความถี่ 12 MHz จะมีคาบเวลาเท่ากับ 1 ms คาบเวลาทั้งสองภายในหนึ่งสเตทจะเรียกว่า เฟส 1 (phase 1) และเฟส 2 (phase 2)

จากไดอะแกรมเวลาสามารถสรุปได้ว่าในการทำงาน 1 รอบหรือ 1 แมชชีนไซเคิล ซีพียูในไมโครคอนโทรเลอร์ MCS-51 จะใช้เวลา 12 คาบเวลาของสัญญาณนาฬิกานั่นคือเวลาในการทำงาน 1 แมชชีนไซเคิลมีค่าเท่ากับ 1 ms หรือมีความเร็วในการทำงานภายใน 1 MHz ในกรณีที่ใช้ความถี่สัญญาณนาฬิกา 12 MHz ดังนั้นถ้าต้องการทราบความเร็วในการทำงานของไมโครคอนโทรเลอร์ MCS-51 สามารถหาได้จาก ค่าความถี่สัญญาณนาฬิกาหารด้วย 12 และถ้าต้องการหาค่าเวลาของ 1 รอบการทำงานหรือ 1 แมชชีนไซเคิลสามารถทำได้โดยการหาส่วนกลับของความเร็วในการทำงานภายในไมโครคอนโทรเลอร์ MCS-51 สามารถสรุปได้เป็นสูตรทางคณิตศาสตร์ได้ดังนี้

ความเร็วในการทำงานของไมโครคอนโทรเลอร์ = ความถี่ของสัญญาณนาฬิกา (ค่าคริสตอลที่ต่ออยู่ที่ขา XTAL1 และ XTAL2)/12

เวลา 1 แมชชีนไซเคิล = 1/ความเร็วในการทำงานภายในของไมโครคอนโทรเลอร์

โหมดการทำงานของพอร์ตอนุกรมใน MCS-51

พอร์ตอนุกรมในไมโครคอนโทรเลอร์ MCS-51 สามารถเลือกการทำงานได้ถึง 4 โหมดคือ

1. โหมด 0 เป็นการกำหนดให้พอร์ตอนุกรมทำงานในลักษณะชิฟรีจีสเตอร์
2. โหมด 1 เป็นการกำหนดให้เป็น UART ขนาด 8 บิต สามารเลือกอัตราบอดได้
3. โหมด 2 เป็นการกำหนดให้เป็น UART ขนาด 9 บิต โดยมีอัตราบอดคงที่
4. โหมด 3 เป็นการกำหนดให้เป็น UART ขนาด 9 บิตสามารเลือกอัตราบอดได้

การเลือกโหมดในการทำงานทำได้ด้วยการกำหนดข้อมูลให้แก่บิต SM0 และ SM1 ในรีจีสเตอร์ SCON

2.7.4 หลักการสื่อสารข้อมูลผ่านพอร์ตอนุกรม

การสื่อสารข้อมูลผ่านพอร์ตอนุกรมของไมโครคอนโทรเลอร์นั้นทำได้ 2 วิธี คือ

1. ใช้อินเตอร์รัปต์ เป็นวิธีการที่ให้ผลการทำงานเร็วที่สุด แต่มีความยุ่งยากในการทำงานกว่า เนื่องจากตำแหน่งของการอินเตอร์รัปต์ทั้งการรับและส่งข้อมูลนั้นอยู่ที่ตำแหน่งเดียวกัน ต้องพิจารณาจากแฟลก TI หรือ RI ก่อนว่าเกิดการอินเตอร์รัปต์จากสาเหตุใด และต้องพิจารณาการใช้รีจีสเตอร์ในช่วงเวลานั้น ๆ ด้วยว่ามีโอกาสซ้อนทับกันหรือไม่ ทำให้โปรแกรมของการทำงานในส่วนนี้มีความซับซ้อนมากกว่า

2. วนโปรแกรมตรวจสอบแฟลก เป็นวิธีที่มีความซับซ้อนน้อยกว่า โดยเขียนโปรแกรมให้วนตรวจสอบแฟลกอยู่ตลอกเวลาจนกว่าจะเกิดการเปลี่ยนแปลง ยกตัวอย่าง เมื่อต้องการตรวจสอบการส่งข้อมูลให้ทำการวนตรวจสอบแฟลก TI ว่าถูกเซตหรือไม่ เมื่อถูกเซตแสดงว่ามีการส่งข้อมูลเกิดขึ้นเรียบร้อยแล้วจากนั้นให้ทำการเคลียร์แฟลก TI แล้วทำการส่งข้อมูลตัวถัดไปหรือทำงานในคำสั่งต่อไปได้
ในกรณีที่ต้องการตรวจสอบการรับข้อมูล ให้ทำการตรวจสอบแฟลก RI ว่าถูกเซตหรือไม่ เมื่อตรวจสอบด้วยว่าถูกเซตแสดงว่าเกิดการรับข้อมูลขึ้น ให้ทำการเคลียร์แฟลก RI แล้วนำค่าจากรีจีสเตอร์ SBUF มาใช้ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทันที แต่วิธีการนี้มีข้อเสียตรงที่มีการทำงานแบบเรียงลำดับ ทำให้ขั้นตอนในการทำงานช้ากว่าการใช้อินเตอร์รัปต์

2.7.5 รูปแบบของการสื่อสารข้อมูลอนุกรม

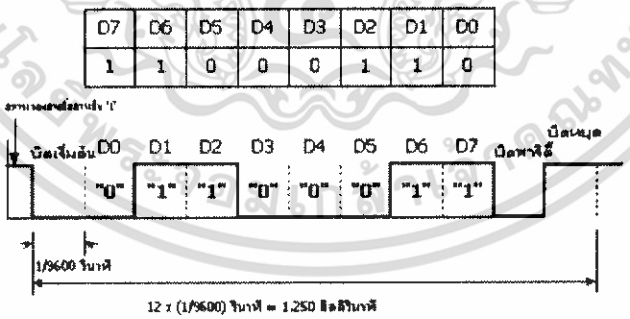
การสื่อสารข้อมูลอนุกรมแบบอะซิงโคนัส เป็นวิธีการรับและส่งข้อมูลโดยไม่ต้องอาศัยสัญญาณนาฬิกาส่งร่วมไปด้วย แต่จะใช้อัตราความเร็วของจำนวนข้อมูลต่อวินาที และจะทำการเพิ่มบิตข้อมูลบางอย่างร่วมไปกับการส่งข้อมูลจริง เพื่อจะได้ทำการตรวจสอบข้อมูลได้อย่างถูกต้องมากยิ่งขึ้นแสดงดังรูปที่ 2.18 ซึ่งประกอบด้วยกัน 4 ส่วนคือ

1 บิตเริ่มต้น (Start bit) จะมีขนาด 1 บิต จะเป็นระดับลอจิกตรงกันข้ามกับระดับลอจิกของสภาวะสายสื่อสาร ขณะที่ยังไม่มีการส่งข้อมูล

2 บิตข้อมูล (Data bit) จะเริ่มจากบิตที่มีนัยสำคัญต่ำสุดก่อนหรือ บิต LSB ก่อน โดยข้อมูลที่จะส่งอาจจะมีขนาด 5 ,6, 7 หรือ 8 บิตก็ได้

3 บิตแสดงสภาวะเลขคู่หรือเลขคี่ (Parity bit) มีขนาด 1 บิตโดยบิตนี้จะนำไปต่อท้ายกับบิตข้อมูล ค่าของบิตนี้ขึ้นอยู่กับจำนวนค่าของข้อมูลที่เป็น "1" โดยเลือกการส่งข้อมูลเป็นแบบ พาริตี้คู่ หรือ พาริตี้คี่ ตัวอย่าง ถ้ากำหนดให้มีการส่งข้อมูลแบบพาริตี้คู่ แต่ข้อมูลมีเลข 1 เป็นจำนวนคี่ ก็จะให้บิตพาริตี้นี้เป็น "1" เพื่อจะได้จำนวนเลข "1" เป็นคู่นั้นเอง ทำนองเดียวกันทางด้านรับเองก็ต้องมีการตรวจสอบจำนวนข้อมูลที่รับเข้ามาเป็น "1" รวมทั้งบิตพาริตี้ 1 บิต ถ้ามีค่า "1"เป็นจำนวนคู่ แสดงว่าข้อมูลที่รับเข้ามาถูกต้อง สามารถกำหนดการรับและส่งข้อมูลเป็นแบบ NONE โดยไม่ต้องมีการตรวจสอบพาริตี้บิตก็ได้

4 บิตสุดท้ายหรือบิตหยุด (Stop bit) เป็นการระบุถึงขอบเขตของการสิ้นสุดข้อมูล โดยจะทำให้ขาข้อมูลมีสถานะ ลอจิกเป็น "1" ซึ่งอาจมีจำนวนมากกว่า หนึ่งบิตก็ได้ เช่น 1 บิต 1.5 บิต หรือ 2 บิต



รูป 2.18 แสดงการส่งข้อมูลขนาด 8 บิตแบบอนุกรมพร้อมด้วย บิตเริ่มต้น, บิตพาริตี้, บิตหยุด ด้วยความเร็ว 9600 บิตต่อวินาที

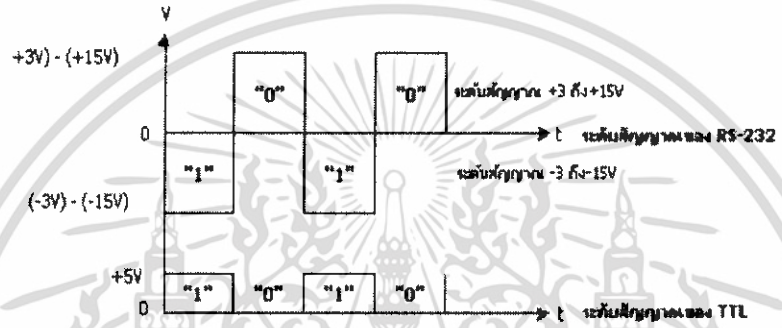
2.8 การเชื่อมต่อพอร์ตอนุกรมมาตรฐานRS-232

การกำหนดมาตรฐานการเชื่อมต่อแบบอนุกรม EIA RS-232 (x) เป็นมาตรฐานอุตสาหกรรม โดยคณะกรรมการสมาคมอุตสาหกรรม อิเล็กทรอนิกส์ (Electronic Industries Association) ออกแบบมาเพื่อใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการส่งข้อมูลอนุกรมแบบอะซิงโครนัส 2 ทิศทาง เพื่อให้มีการใช้งาน ในการเชื่อมต่อที่สอดคล้องกัน ระหว่างอุปกรณ์คอมพิวเตอร์ต่างๆ การรับส่งสัญญาณจะกำหนดความยาวสูงสุดไว้ที่ไม่เกิน 50 ฟุตโดยมีระดับ สัญญาณตั้งแต่ 3 โวลท์ จนถึง 15 โวลท์ สำหรับลอจิก "0" และมีระดับแรงดันที่ -3 โวลท์ จนถึง-15 โวลท์สำหรับลอจิก"1"ดังแสดงในรูป 2.19

ดังนั้นสังเกตได้ว่าจะมีระดับแรงดันที่ใช้ในสถานะลอจิก "0" และ ลอจิก "1" แตกต่างออกไปจากระบบไอซีดิจิตอลทั่วๆไปการต่อใช้งานจึงต้อง มีอุปกรณ์ที่ทำหน้าที่เปลี่ยนระดับแรงดันจาก 0 - 5 โวลท์ จากไมโครคอนโทรลเลอร์ ให้เป็นระดับแรงดันที่สูงกว่า +3 หรือต่ำกว่า - 3 โดยจะมีไอซีสำเร็จรูปพร้อมใช้งานหรืออาจจะต่อวงจรจากทรานซิสเตอร์ได้

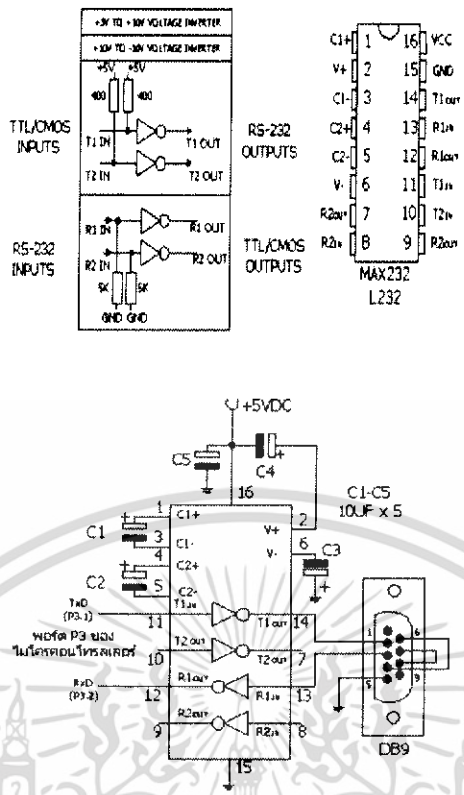


รูปที่ 2.19 แสดงระดับแรงดันสัญญาณของพอร์ตอนุกรม RS-232 กับ TTL ในสถานะลอจิก "1" และ "0"

2.9 ไอซี MAX232, L232

ไอซี MAX232, L232 เป็นไอซีที่แปลงระดับสัญญาณจากระดับ TTL ไปเป็นระดับของ RS-232 และในทำนองเดียวกันก็รับระดับสัญญาณจาก RS-232 เพื่อแปลงเป็นระดับสัญญาณจากระดับ TTL ให้กับไมโครคอนโทรลเลอร์ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.20 แสดงตำแหน่งขาของไอซี MAX232, L232 และการต่อใช้งาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การออกแบบและการสร้าง

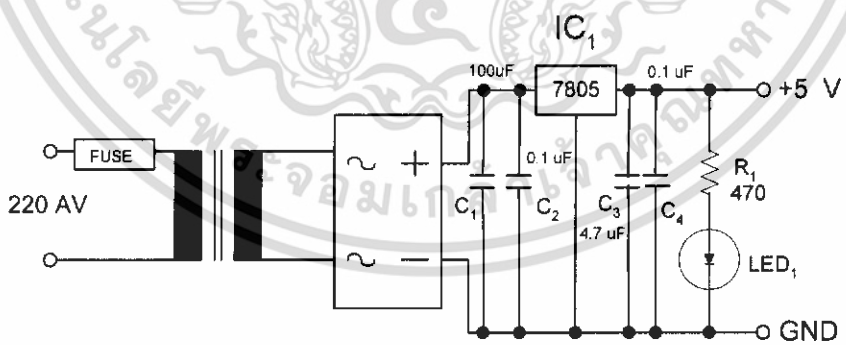
ในโครงงานนี้ ได้แบ่งการออกแบบและการสร้างเป็น 2 ส่วน คือ ในส่วนของฮาร์ดแวร์และส่วนของซอฟแวร์ โดยส่วนของฮาร์ดแวร์ นั้นได้มีการออกแบบวงจรเป็นสวิตซ์ 4 ช่องควบคุมอุปกรณ์ไฟฟ้า โดยแสดงผลด้วย LED ซึ่งสามารถเลือกการควบคุมได้ทั้งจากการกดสวิตซ์บนบอร์ดหรือควบคุมผ่านพอร์ตอนุกรมของเครื่องคอมพิวเตอร์ส่วนบุคคล โดยข้อมูลที่ส่งผ่านพอร์ตอนุกรมนั้น จะได้จากส่วนของซอฟแวร์ ที่ทำงานตามกระบวนการรู้จำเสียง

3.1 การออกแบบส่วนของฮาร์ดแวร์

3.1.1 วงจรสวิตซ์เปิด-ปิด 4 ช่องควบคุมผ่านพอร์ตอนุกรม

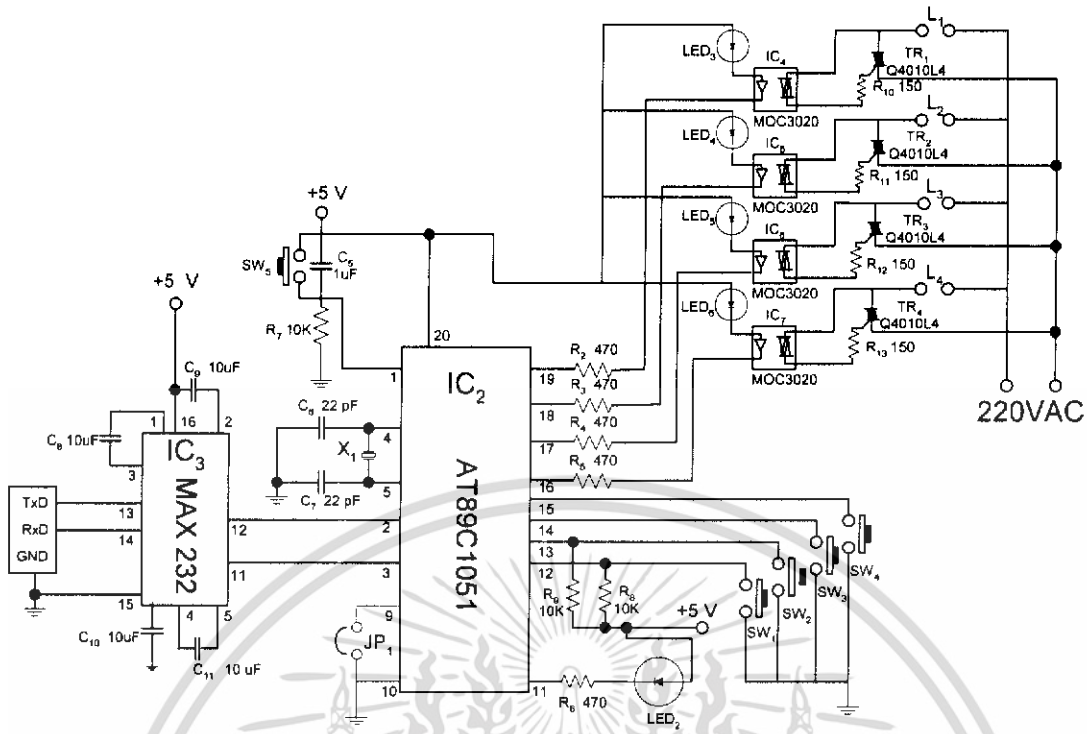
จากรูปที่ 3.1 เมื่อป้อนไฟกระแสสลับแรงดัน 220 โวลต์ เข้าหม้อแปลงแล้ว จะได้แรงดันเอาท์พุตที่ออกจากหม้อแปลงเป็นไฟกระแสสลับแรงดัน 6 โวลต์ โดยผ่านไดโอดบริดจ์ และต่อเข้ากับ ไอซี เบอร์ LM7805 ทำให้ได้แรงดันเอาท์พุตคงที่ 5 โวลต์ เพื่อป้อนเป็นไฟเลี้ยงให้กับวงจร สถานะของ C_1, C_2 , C_3 และ C_4จะทำหน้าที่กรองกระแสให้เรียบ

ไอซี MAX 232 ทำหน้าที่ปรับระดับสัญญาณ ให้เหมาะสมกับมาตรฐานของ RS-232 ซึ่งเป็นมาตรฐานของการรับส่งข้อมูลผ่านพอร์ตอนุกรมของเครื่องคอมพิวเตอร์ส่วนบุคคล โดยเชื่อมต่อกับ ขา R×D และ T×D ของพอร์ต COM1 และ COM2 ด้านหนึ่งและเชื่อมต่อกับ ไอซี ไมโครคอนโทรลเลอร์เบอร์ AT89C1051 ของ ATMEL อีกด้านหนึ่ง ซึ่งได้มีการโปรแกรมการทำงานไว้ภายใน ซึ่งจะกล่าวในหัวข้อต่อไป



รูปที่ 3.1 แสดงวงจรแปลงแรงดัน 220 V

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



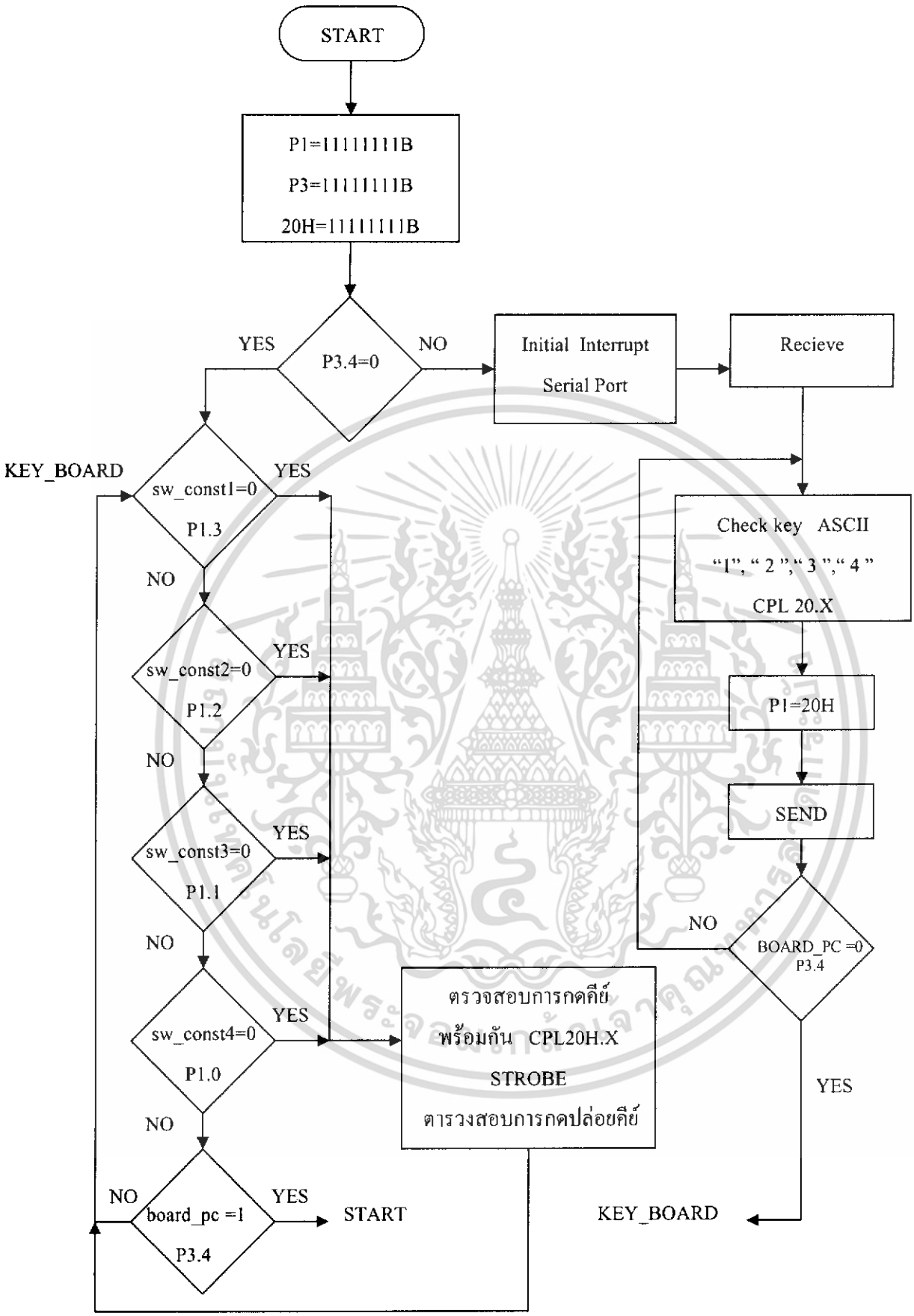
รูปที่ 3.2 วงจรสวิตซ์เปิด-ปิด 4 ช่องควบคุมผ่านพอร์ตอนุกรม

C_5 , R_7 และ SW_5 จะเป็นส่วนของการรีเซต และ C_6 , C_7 และ X_1 เป็นการสร้างสัญญาณนาฬิกาให้กับไอซี ไมโครคอนโทรลเลอร์ ที่จะทำหน้าที่ควบคุมและสั่งงานการทำงานของวงจร โดยเริ่มจากทำการตรวจสอบการเลือกส่วนของการควบคุมจากบิตที่ P3.4 โดยเลือกที่จัมเปอร์ JP_1 ส่วนสวิตซ์ SW_1 , SW_2 SW_3 และ SW_4 จะทำหน้าที่ทำการเลือกการใช้ งาน เปิด ปิด ในขณะเลือกการควบคุมจากวงจรโดยตรง และที่ P3.7 จะต่อกับ LED_2 โดยมี R_6 ทำหน้าที่เป็นตัวจำกัดกระแส วงจรในส่วนของการติดต่อกับอุปกรณ์ไฟฟ้าภายนอก ได้ทำเป็นโซลิดสเตทรีเลย์ 4 ช่อง ได้ต่อ LED_3 , LED_4 , LED_5 และ LED_6 เป็นส่วนแสดงสถานะของการเปิดปิด ซึ่งได้ต่ออนุกรมไว้กับ LED ภายในของ IC_4-IC_7

3.1.2 ส่วนโปแกรมของไมโครคอนโทรลเลอร์ควบคุมฮาร์ดแวร์

ส่วนของโปรแกรมของไมโครคอนโทลเลอร์ ได้แบ่งออกเป็น 2 ส่วน คือ ส่วนควบคุมการเปิดปิดโดยสวิตซ์จากวงจรโดยตรง และส่วนควบคุมการเปิดปิดโดยข้อมูลที่ส่งผ่านพอร์ตอนุกรม โดยมีการเลือกทำงานจากจัมเปอร์ JP_1 รูปที่ 3.3 แสดงแผนผังการทำงานของโปรแกรมที่ใช้ในโครงงานนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.3 แผนผังแสดงขั้นตอนการทำงานของโปรแกรมสวิตช์เปิดปิด 4 ช่อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 การออกแบบและการสร้างส่วนของการรู้จำเสียงพูด

3.2.1 การออกแบบสร้างคำต้นแบบอ้างอิงและคำทดสอบ

การสร้างคำต้นแบบอ้างอิงใน ปริญญานิพนธ์นี้ จะใช้คนทั้งหมด 4 คน ประกอบด้วย ผู้ชาย 2 คน ผู้หญิง 2 คน โดยจะให้ผู้สร้างคำต้นแบบ แต่ละคน อ่านออกเสียง คำ 9 คำ ได้แก่ เปิดหนึ่ง , เปิดสอง , เปิดสาม , เปิดสี่ , ปิดหนึ่ง , ปิดสอง , ปิดสาม , ปิดสี่ และ รีเซต คำละ 20 ครั้ง

และในการสร้างคำทดสอบ จะใช้คนทั้งหมด 20 คน ประกอบด้วยผู้ชาย 10 คน ผู้หญิง 10 คน โดยให้ผู้สร้างคำทดสอบ แต่ละคนอ่านออกเสียง คำ 9 คำ ต้นแบบ คำละ 2 ครั้ง

3.2.2 การเก็บข้อมูล

การเก็บข้อมูลในปริญญานิพนธ์นี้ จะใช้คำพูดเป็นภาษาไทยโดยจะให้ผู้ทดสอบ แต่ละคนอ่านออกเสียง 9 คำ ในการบันทึกเสียง ซึ่งข้อมูลที่ถูกบันทึกจะถูกบันทึกอยู่ในรูปแบบของไฟล์ข้อมูล ".wav" ซึ่งข้อมูล 1 ตัวอย่างของเสียงจะถูกแทนด้วยข้อมูลเสียงขนาด 8 บิทโดยใช้ความถี่ในการแซมปลิ้งเสียงเท่ากับ 11.025 KHz ไฟล์ข้อมูลเสียง ".wav" จะถูกนำไปใช้เป็นอินพุทสำหรับโปรแกรมการคำนวณการรู้จำเสียงต่อไป

3.2.3 เครื่องมือและอุปกรณ์ในการวิเคราะห์สัญญาณเสียง

เครื่องมือและอุปกรณ์ในการวิเคราะห์กระบวนการรู้จำเสียงในปริญญานิพนธ์นี้ จะใช้คอมพิวเตอร์ 1 เครื่อง ไมโครโฟน 1 ตัว ลำโพง 1 ชุด โปรแกรมการบันทึกเสียงและโปรแกรมที่ใช้ในการรู้จำเสียงพูดใช้โปรแกรม Visual Studio C++

ข้อจำกัดในการเก็บเสียงพูด

1. เป็นเสียงพูดภาษาไทยภาคกลาง พูดด้วยความเร็วปกติ
2. น้ำเสียงของผู้พูดไม่แสดงอารมณ์ใด ๆ
3. พูดอย่างชัดถ้อยชัดคำ
4. สภาพแวดล้อมขณะเก็บเสียงเงียบสงบ หรือมีเสียงรบกวนที่ไม่ดังเช่นเสียงแอร์ได้
5. สภาพร่างกายผู้พูดต้องอยู่ในสภาวะปกติไม่เป็นโรคที่มีผลต่อเสียงพูดเช่นหวัด
6. คำพูดที่ใช้ในการรู้จำเป็นตัวเลข 0-9 และคำสั้น ๆ เช่น เปิด ปิด
7. ในช่วงเวลาหนึ่ง ๆ สามารถรองรับผู้ใช้เพียง 1 คน

3.2.4 หลักการรู้จำเสียงพูดในปริญญานิพนธ์นี้

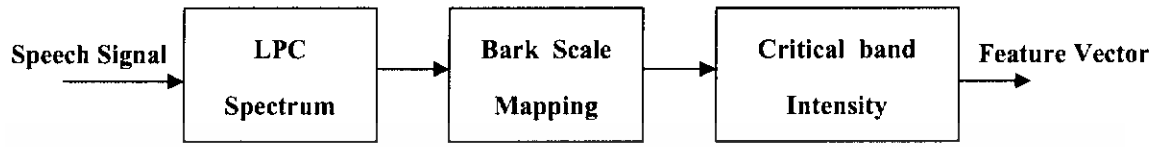
เสียงพูดที่ใช้ในปริญญานิพนธ์นี้ ถูกออกแบบให้มีความสอดคล้องกับส่วนฮาร์ดแวร์ที่ใช้ในการควบคุมการเปิด ปิด อุปกรณ์ไฟฟ้า ซึ่งได้ใช้คำทั้งหมด 9 คำ ได้แก่ เปิดหนึ่ง , เปิดสอง , เปิดสาม, เปิดสี่, ปิดหนึ่ง, ปิดสอง, ปิดสาม, ปิดสี่ และรีเซต ซึ่งคำทั้งหมดนี้ มีลักษณะเป็นเสียงก้อง (Voice) และมีลักษณะสัญญาณเป็นคาบ ซึ่งเป็นลักษณะที่เห็นได้ชัดในโดเมนของเวลา จากนั้น นำเสียงที่ได้ ทั้ง 9 คำ มาทำการวิเคราะห์หาคุณลักษณะของแต่ละคำ แบ่งเป็น 3 ขั้นตอน ตามรูปที่ 3.4 ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. การหาสเปกตรัม LPC

2. การแปลงจากสเกลความถี่ไปเป็นสเกลบาร์ค หรือ Critical Band Rate Scale และการคำนวณหา Critical Band Intensity

3. ใช้ Critical Band Intensity เป็น Feature Vector ในการสร้างแบบอ้างอิง



รูปที่ 3.4 แสดงขั้นตอนในการวิเคราะห์เสียง

3.2.5 การหาค่าพารามิเตอร์ที่ใช้ในการแยกเสียงพูด

3.2.5.1 การหาค่าสเปกตรัมแอลซีพี

วิธีการหาสเปกตรัมแอลซีพีนั้นเป็นการนำสัญญาณตัวอย่างที่ต้องการนำมาเป็น คำต้นแบบอ้างอิง ผ่านตามขั้นตอนต่างๆ ที่ได้กล่าวแล้วในบทที่ 2 การคำนวณหาค่าสเปกตรัมแอลพีซีของสัญญาณเสียง เริ่มจากสัญญาณเสียงที่ผ่านการอัดเสียง แล้วนำมาทำพรีเอมฟาซีส เพื่อให้อัตราส่วนของสัญญาณเสียง ต่อสัญญาณ รบกวน มีค่าค่อนข้างคงที่ตลอดช่วงความถี่ โดยเน้นความถี่สูงให้มีขนาดที่สูงขึ้น สัญญาณที่ผ่านการพรีเอมฟาซีสแล้ว จะถูกแบ่งออกเป็นเฟรม เฟรมละ 300 แซมเปิ้ล การวิเคราะห์สัญญาณเสียงจะวิเคราะห์ทีละเฟรม โดยแต่ละเฟรมจะถูกเลื่อนไปที่ 100 แซมเปิ้ล จากนั้นนำสัญญาณแต่ละเฟรมมาคูณ ด้วยฟังก์ชั่นวินโดว์ เพื่อลดความไม่ต่อเนื่องของสัญญาณที่ขอบ และไม่ทำให้สเปกตรัมของสัญญาณในช่วงความถี่ต่ำเปลี่ยนแปลงไป แล้วนำสัญญาณที่ผ่านการวินโดว์ไปคำนวณหาสัมประสิทธิ์แอลพีซี เมื่อได้ค่าสัมประสิทธิ์แอลพีซีแล้วก็จะสามารถคำนวณหาสเปกตรัมแอลซีพีได้

การคำนวณหาสัมประสิทธิ์แอลพีซีนั้น จำเป็นจะต้องเลือกอันดับของแอลพีซีให้เหมาะสม ถ้าอันดับของแอลพีซีมีค่าต่ำ จะทำให้ การคำนวณเร็วขึ้นแต่ค่าสเปกตรัมแอลพีซีที่ได้จะมีความละเอียดน้อยทำให้การคำนวณหา Critical Band Intensity เกิดความผิดพลาด แต่ถ้าอันดับแอลพีซีสูงจะใช้เวลานาน ทำให้การคำนวณยุ่งยากและซับซ้อนเกินความจำเป็น ดังนั้นจากการศึกษาพบว่า ค่าอันดับ 19 นั้นเป็นค่าที่ให้ประสิทธิภาพสูงสุด

3.2.5.2 การแปลงจากสเกลความถี่ ไปเป็นสเกลบาร์ค หรือ Critical Band Rate Scale และการคำนวณหา Critical Band Intensity

สเกลบาร์คนั้นช่วยให้ความสามารถในการแสดงคุณสมบัติต่อการตอบสนองความถี่ต่างๆ ของมนุษย์เป็นเชิงเส้นตลอดช่วงความถี่ที่มนุษย์สามารถได้ยิน ดังนั้นจึงทำการแปลงสเปกตรัมแอลพีซีที่เป็นความถี่ไปบนสเกลบาร์ค แล้วนำค่าสเปกตรัมแอลพีซีที่อยู่บนสเกลบาร์คไปคำนวณหาค่าความเข้มของสัญญาณในแต่ละความถี่วิกฤต ซึ่งค่าที่ได้จะนำไปใช้เป็นองค์ประกอบของเวกเตอร์แทนหน่วยเสียงต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.5.3 การสร้างและการทดสอบแบบอ้างอิง

การสร้างแบบอ้างอิง หรือ การสร้างโมเดล นั้น จะใช้หลักการของแบบจำลองมาร์คอฟ โดยเรียกขั้นตอนการสร้างแบบอ้างอิง ว่า การเทรนนิ่ง และเรียกการทดสอบแบบอ้างอิงว่า การเทสติ้ง

การเทรนนิ่ง เป็นการนำค่าจากขั้นตอน Feature Measurment ซึ่งก็คือ ค่า Critrical Band Intensity บนสเกลบาร์ค เข้าสู่กระบวนการของฮิดเด็นมาร์คอฟเพื่อสร้างเป็นโมเดลอ้างอิง ของแต่ละสัญญาณเสียงพูด ในปริญญานิพนธ์นี้ ใช้สัญญาณเสียงพูด 9 คำ และได้ค่า Critrical Band Intensity บนสเกลบาร์คจากผู้สร้างโมเดล 4 คน คนละ 20 ครั้ง รวมทั้งหมดได้ค่า Critrical Band Intensity บนสเกลบาร์ค 80 ค่า ต่อสัญญาณเสียง 1 คำ

การเทสติ้ง เป็นการทดสอบแบบอ้างอิงที่ได้สร้างขึ้น โดยวิธีการเปรียบเทียบความน่าจะเป็นว่าใกล้เคียงกับแบบอ้างอิงใดที่ได้จากกระบวนการเทรนนิ่งมากที่สุด

3.2.5.4 การเชื่อมต่อระหว่างส่วนฮาร์ดแวร์และส่วนของการรู้จำเสียงพูด

ในส่วนของวงจรสวิตซ์ เปิด ปิด 4 ช่องผ่านพอร์ตอนุกรมนั้นได้กำหนดพารามิเตอร์ในการทำการเปิดปิด ไว้ทั้งหมด 9 ค่า ดังนี้

“เปิด หนึ่ง”	ส่งค่า “1”.
“เปิด สอง”	ส่งค่า “2”.
“เปิด สาม”	ส่งค่า “3”.
“เปิด สี่”	ส่งค่า “4”.
“ปิด หนึ่ง”	ส่งค่า “1”.
“ปิด สอง”	ส่งค่า “2”.
“ปิด สาม”	ส่งค่า “3”.
“ปิด สี่”	ส่งค่า “4”.
“รีเซต”	ส่งค่า “0”.

และได้ออกแบบให้มีการตรวจเช็คสถานการณ์เปิดหรือปิดของสวิตซ์ก่อน เพื่อป้องกันการทำงานที่ผิดพลาดอันเนื่องมาจากการส่งค่าเปิด ปิด ที่เป็นค่าเดียวกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
การทดลองและผลการทดลอง

4.1 การทดลองส่วนฮาร์ดแวร์เพื่อส่งข้อมูลผ่านพอร์ตอนุกรม

4.1.1 ทดสอบวงจร MAX 232

วัดสัญญาณที่ออกจากขา 12 ของ MAX 232 ซึ่งเป็นขา TxD ที่เข้าสู่ขา 2 ของไอซีไมโครคอนโทรลเลอร์ ด้วยออสซิโลสโคป ต่อสาย DB9 ของวงจร MAX 232 เข้าสู่พอร์ตอนุกรม COM1 ของคอมพิวเตอร์

1.เรียกโปรแกรมหน้าจออินเตอร์เฟสที่ได้ออกแบบขึ้น โดยตั้งค่าที่ COM1 และ Baudrate มีค่า 9600 Kbps แล้วทำการเชื่อต่อพอร์ตอนุกรมโดยกดสวิตซ์ Open Port

2. กดสวิตซ์ Open1, Open2 ,Open3 และ Open4 สังเกตสัญญาณที่ได้จากออสซิโลสโคป

3. บันทึกผลการทดลองในการกดแต่ละครั้ง

4.1.2 ทดสอบวงจรสวิตซ์เปิดปิด 4 ช่องผ่านพอร์ตอนุกรม

จากวงจรรูปที่ 3.2 ต่ออุปกรณ์ไฟฟ้า 4 ชนิด ณ ตำแหน่ง L_1, L_2, L_3 และ L_4 ต่อสาย DB9 เข้าสู่พอร์ตอนุกรม COM1 ของคอมพิวเตอร์

1. เปิดโปรแกรมหน้าจออินเตอร์เฟสที่ได้ออกแบบขึ้น โดยตั้งค่าที่ COM1 และ Baudate มีค่า 9600 Kbps แล้วทำการเชื่อต่อพอร์ตอนุกรมโดยกดสวิตซ์ Open Port

2. กดสวิตซ์ Open1 สังเกตผลที่เกิดขึ้นที่วงจรและอุปกรณ์ไฟฟ้าทั้ง 4 ชนิด

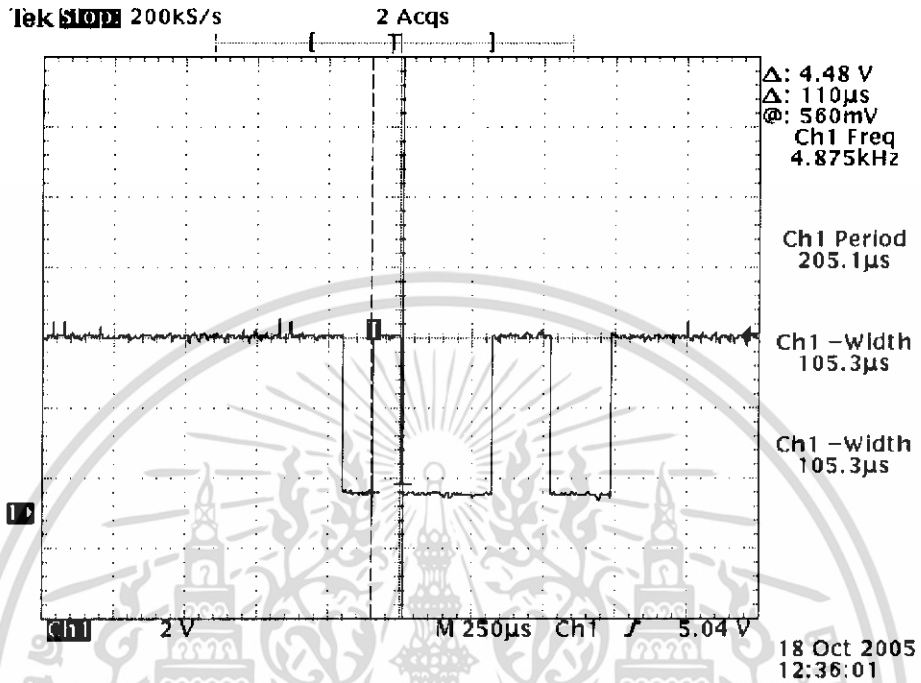
3. กด สวิตซ์ Open2, Open3 และ Open4 ตามลำดับ สังเกตผลที่เกิดขึ้นที่วงจรและอุปกรณ์ไฟฟ้าทั้ง 4 ชนิด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

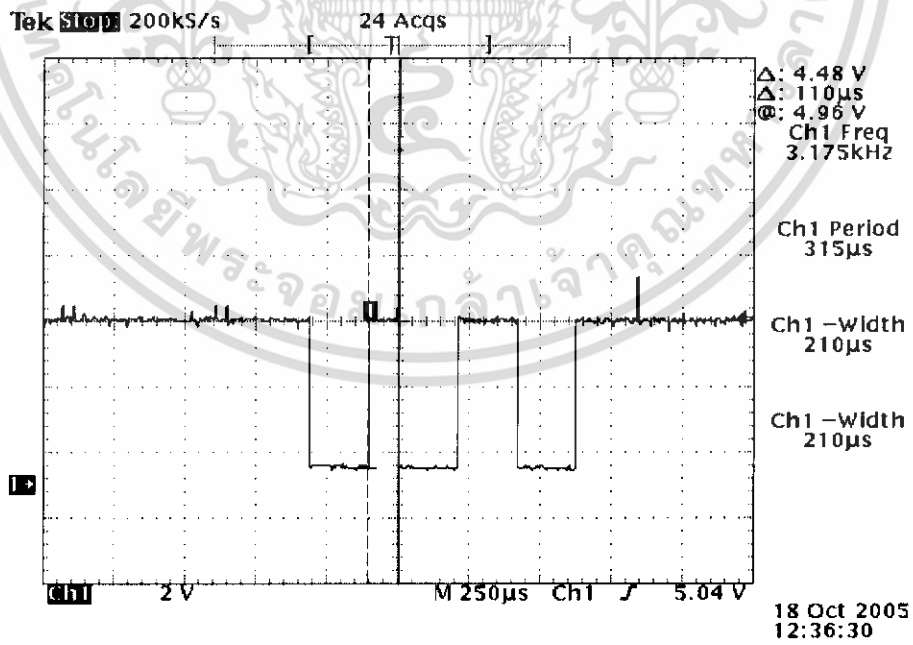
4.2 ผลการทดลองส่วนฮาร์ดแวร์

4.2.1 ทดสอบวงจร MAX 232

ส่งข้อมูลผ่านพอร์ตอนุกรมด้วย Baud rate 9600 b/s ได้คาบ 104 μs

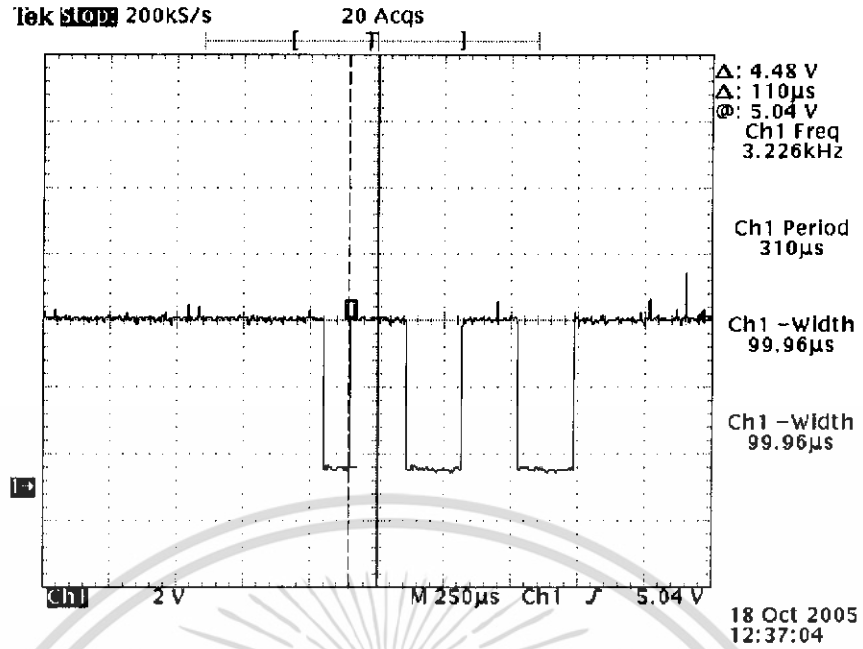


รูปที่ 4.1 แสดงผลเมื่อทำการส่ง 1 จากคีย์บอร์ดผ่านวงจรแปลงแรงดัน

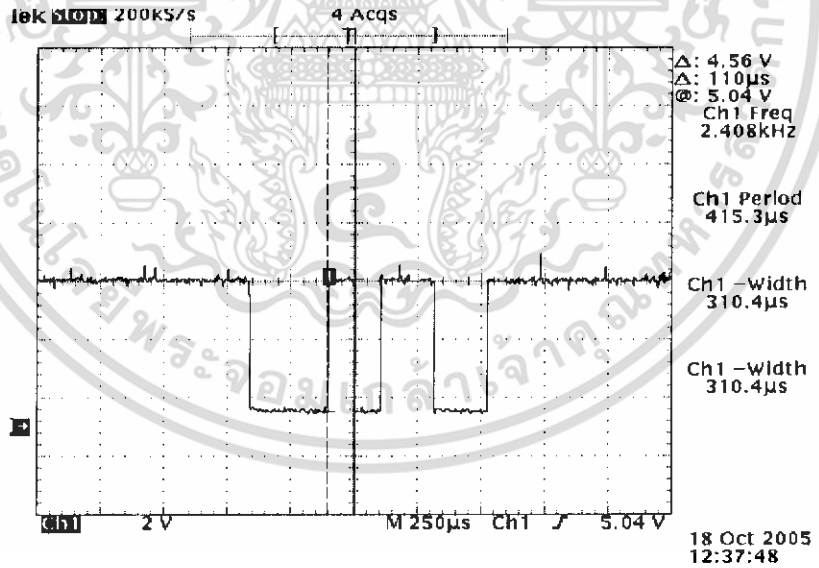


รูปที่ 4.2 แสดงผลเมื่อทำการส่ง 2 จากคีย์บอร์ดผ่านวงจรแปลงแรงดัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.3 แสดงผลเมื่อทำการส่ง 3 จากคีย์บอร์ดผ่านวงจรแปลงแรงดัน

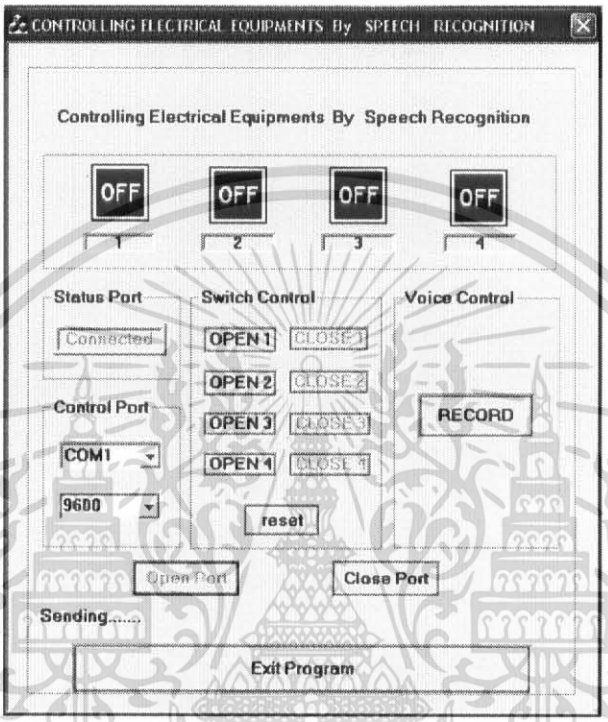


รูปที่ 4.4 แสดงผลเมื่อทำการส่ง 4 จากคีย์บอร์ดผ่านวงจรแปลงแรงดัน

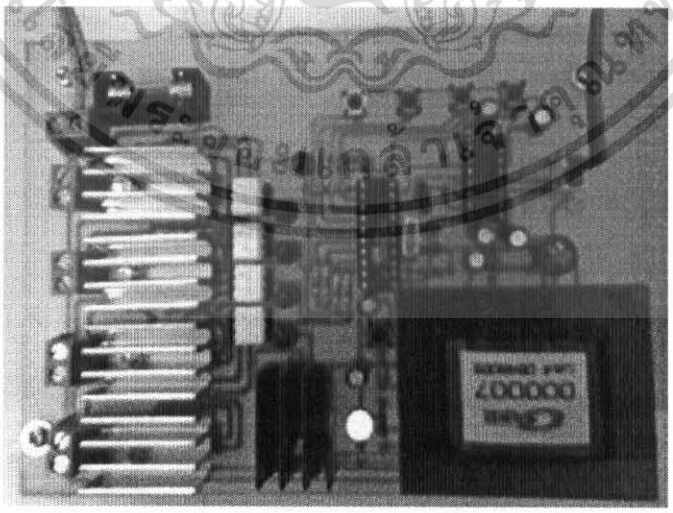
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2 ทดสอบวงจรสวิตซ์เปิดปิด 4 ช่องผ่านพอร์ตอนุกรม

1. หลังจากกดสวิซต์ Open Port พบว่า ที่ Status Port แสดงว่า Connected แสดงว่าขณะนี้พอร์ต ที่ตั้งไว้คือ COM1 ค่า Baud rate 9600 Kbps ได้เชื่อมต่อกับโปรแกรมอินเตอร์เฟสเรียบร้อยแล้ว ในรูป 4.5 และ LED ที่สว่างในรูป 4.6 แสดงสถานะว่าในวงจรการแปลงไฟกระแสสลับจาก 220 V เป็นไฟกระแสตรง 5 V ที่ใช้ในวงจรเป็นส่วนใหญ่ สามารถทำงานได้ปกติ



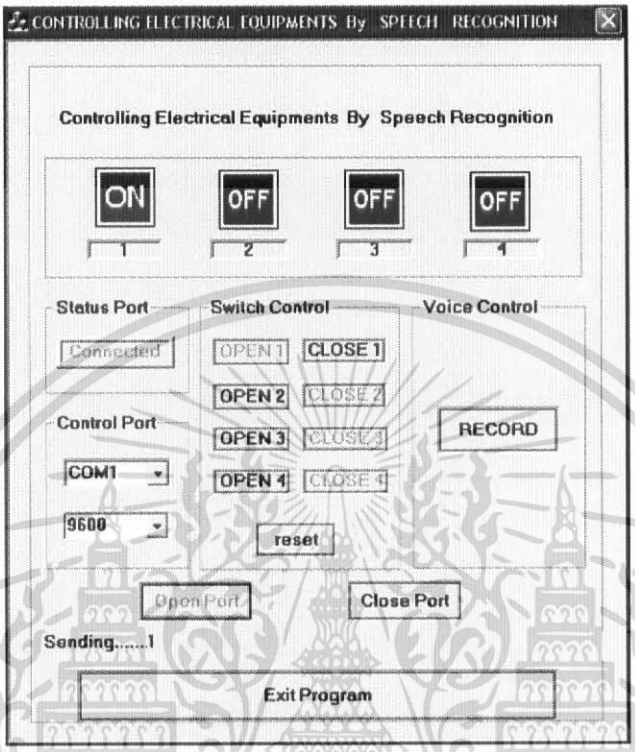
รูปที่ 4.5 แสดงรูปหน้าจออินเตอร์เฟสหลังจากกดสวิตซ์ Open Port



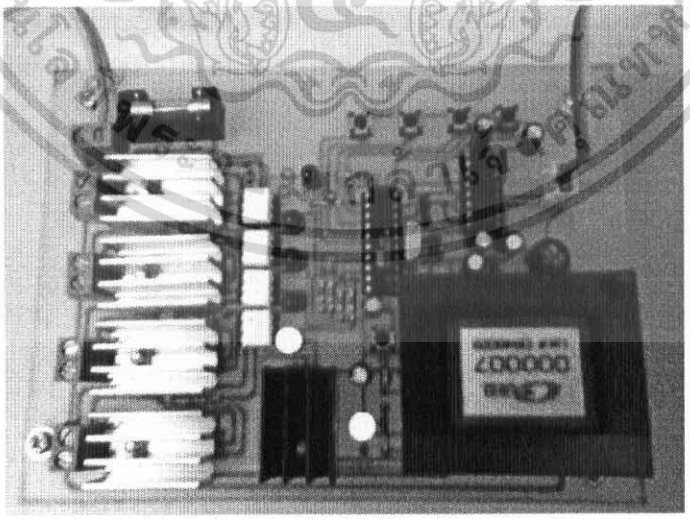
รูปที่ 4.6 แสดงการทำงานของวงจรหลังจากกดสวิตซ์ Open Port

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. หลังจากการ กดสวิตช์ OPEN 1 สังเกตผลที่ได้จากโปรแกรมหน้าจออินเตอร์เฟส และสามารถทราบได้ว่าขณะนี้ส่งค่า 1 ออกพอร์ตอนุกรม จากข้อความที่ปรากฏด้านล่าง “Sending.....1”ดังรูปที่ 4.7 และในรูปที่ 4.8 LED ตัวที่ 1 สว่างแสดงสถานะว่าสวิตช์ช่องที่ 1 ทำงาน



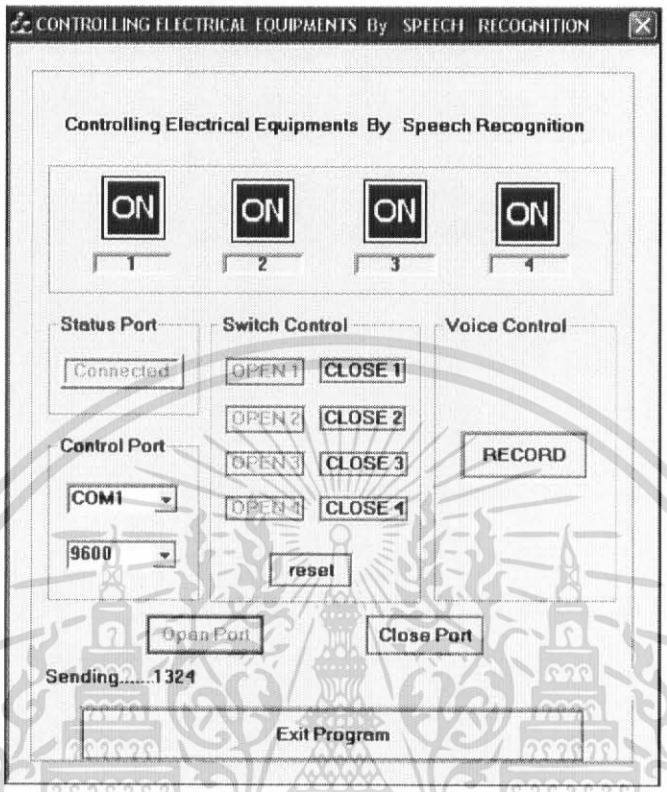
รูปที่ 4.7 แสดงรูปหน้าจออินเตอร์เฟสหลังจากกดสวิตช์ OPEN 1



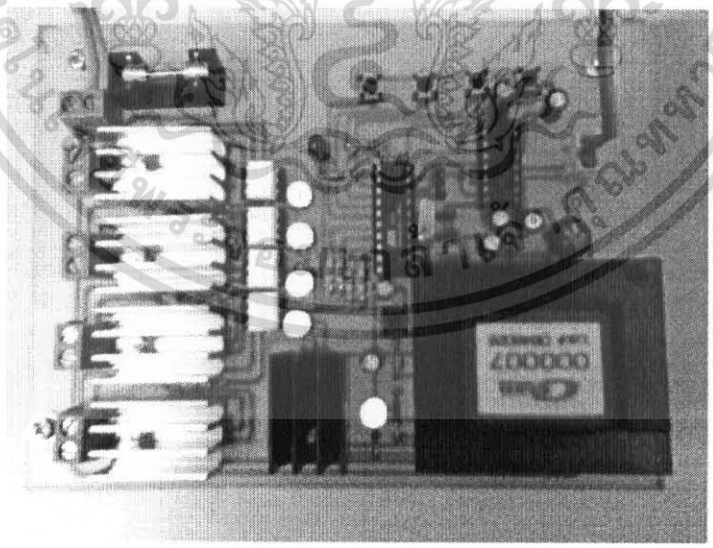
รูปที่ 4.8 แสดงการทำงานของวงจรหลังจากกดสวิตช์ OPEN 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. หลังจากกดสวิตซ์ OPEN3 , OPEN2 และ OPEN 4 สังเกตที่ "Sending... 1324" แสดงว่ามีการส่งค่า 1,3,2 และ 4 ตามลำดับ



รูปที่ 4.9 แสดงรูปหน้าจออินเตอร์เฟสหลังจากกดสวิตซ์ OPEN 2,OPEN 3,OPEN 4



รูปที่ 4.10 แสดงการทำงานของวงจรหลังจากกดสวิตซ์ OPEN 2,OPEN 3,OPEN 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 การทดลองส่วนของการรู้จำเสียงพูด

ขั้นตอนต่างๆในส่วนของการรู้จำ แบ่งเป็น 4 ส่วนใหญ่ๆ ดังนี้

1. ขั้นตอนในการเตรียมสัญญาณเสียง หรือ Signal Preprocessing

ในปริญญานิพนธ์นี้ ใช้สัญญาณเสียงคำพูดทั้งหมด 9 คำ ได้แก่ เปิดหนึ่ง, เปิดสอง, เปิดสาม , เปิดสี่ , ปิดหนึ่ง , ปิดสอง , ปิดสาม , ปิดสี่ และรีเซต จากผู้ชาย 2 คน และผู้หญิง 2 คน คำละ 20 ครั้ง ผ่านกระบวนการพรีเอมฟาซีส การแบ่งเฟรม และการวินโดว์

2 ขั้นตอนการหาคุณลักษณะพิเศษของเสียง หรือ Feature Measurement

ในปริญญานิพนธ์นี้ใช้หลักการของ Linear Predictive Coding (LPC) โดยวิธีอัตตสัมพันธ์ หรือ Autocorrelation เพื่อหาสัมประสิทธิ์ LPC และสเปกตรัม LPC จากนั้นจะทำการแปลงสเปกตรัม LPC จากสเกลความถี่ ไปเป็นสเกลบาร์ค หรือ Critical Band Rate Scale และคำนวณ Critical Band Intensity และใช้ Critical Band Intensity เป็น Feature Measurement ในการสร้างแบบอ้างอิงด้วยวิธีการของ HMM

3 ขั้นตอนในการสร้างแบบอ้างอิง และทดสอบแบบอ้างอิงที่ได้สร้างขึ้น

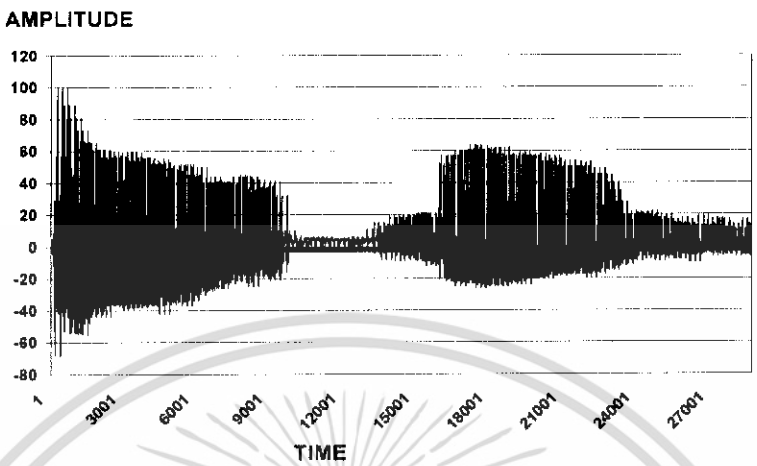
ขั้นตอนนี้จะแบ่งการทดสอบแบบอ้างอิงเป็นสองการทดลองคือ ทดสอบแบบอ้างอิงที่ได้สร้างขึ้นจากผู้สร้างแบบอ้างอิงเอง และทดสอบแบบอ้างอิงที่ได้สร้างขึ้นจากผู้ที่ไม่ได้เป็นผู้สร้างแบบอ้างอิง

4 ขั้นตอนในการเชื่อมต่อการรู้จำเสียงพูดกับส่วนของฮาร์ดแวร์สวิตซ์เปิด ปิด สี่ช่องผ่านพอร์ตอนุกรม

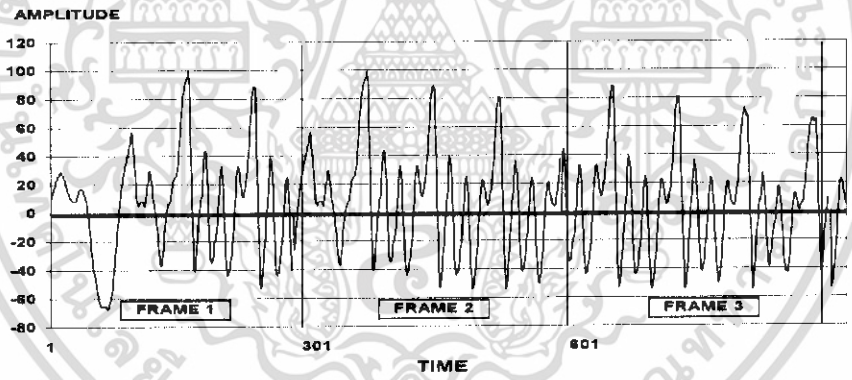
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 ผลการทดลองส่วนของการรู้จำเสียงพูด

1. ขั้นตอนในการเตรียมสัญญาณเสียง หรือ Signal Preprocessing

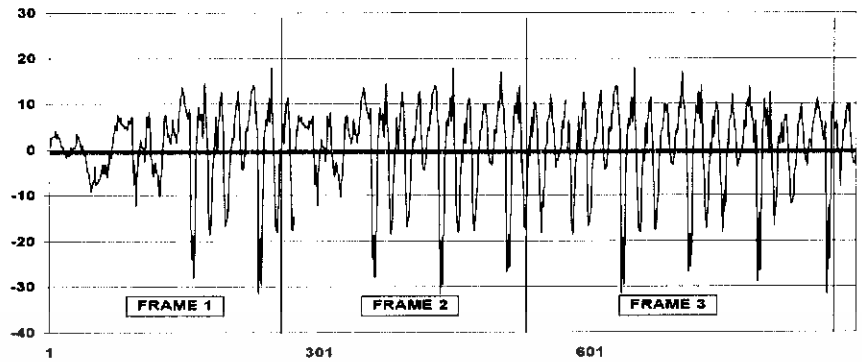


รูปที่ 4.12 แสดงสัญญาณเสียง "เปิด หนึ่ง"

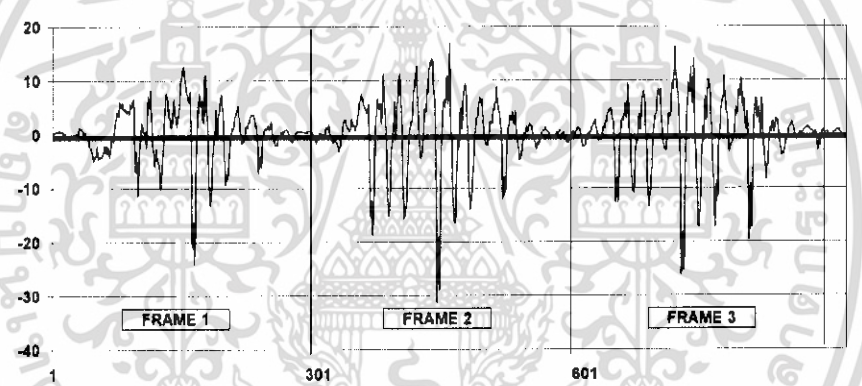


รูปที่ 4.13 แสดงสัญญาณเสียง "เปิด หนึ่ง" ในช่วง 900 แซมเปิ้ลแรก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



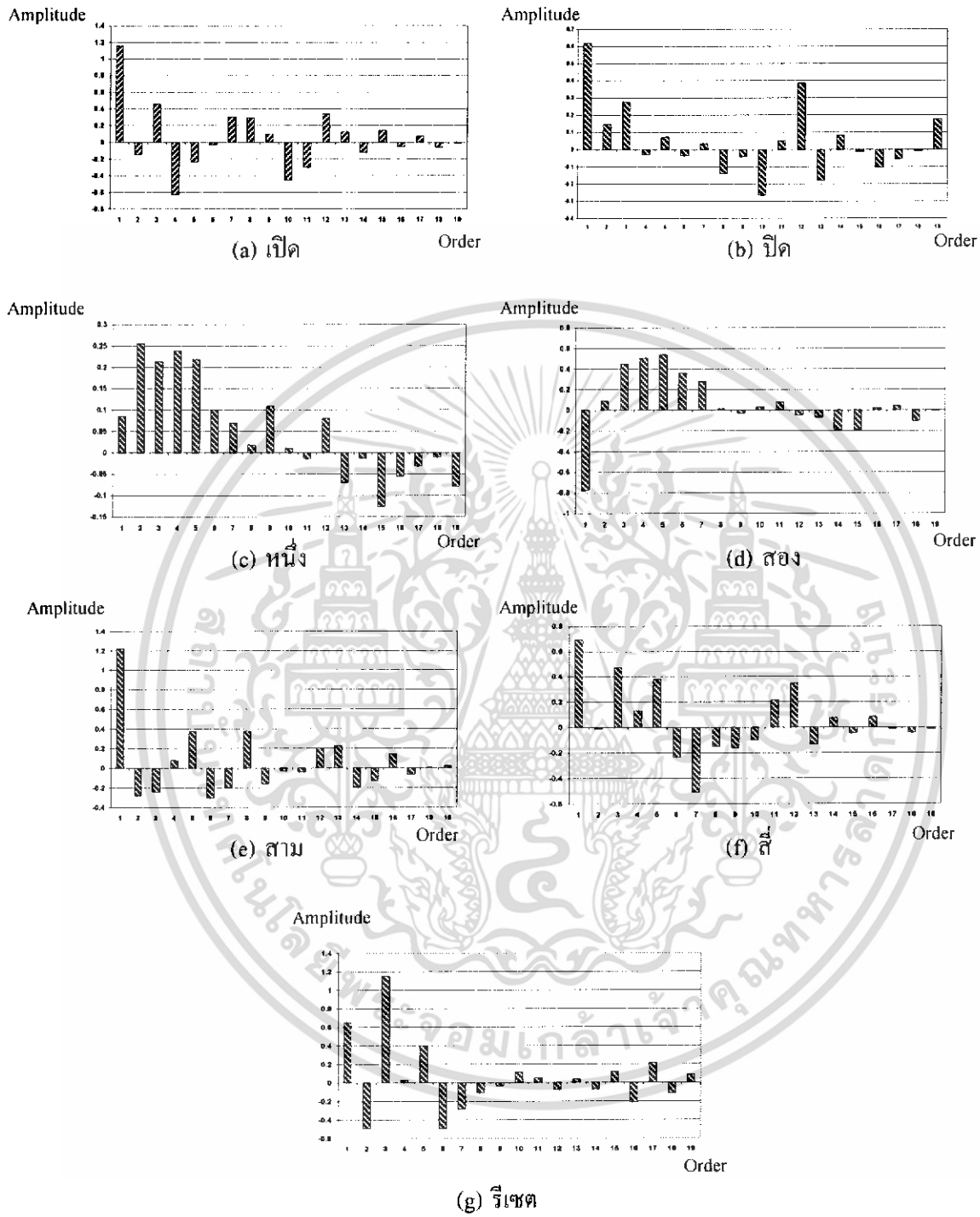
รูปที่ 4.14 แสดงสัญญาณเสียง “เปิด หนึ่ง” หลังจากผ่านกระบวนการพรีเอมฟาซีส



รูปที่ 4.15 แสดงสัญญาณเสียง “เปิด หนึ่ง” หลังจากผ่านกระบวนการแฮมมิ่งวินโดว์

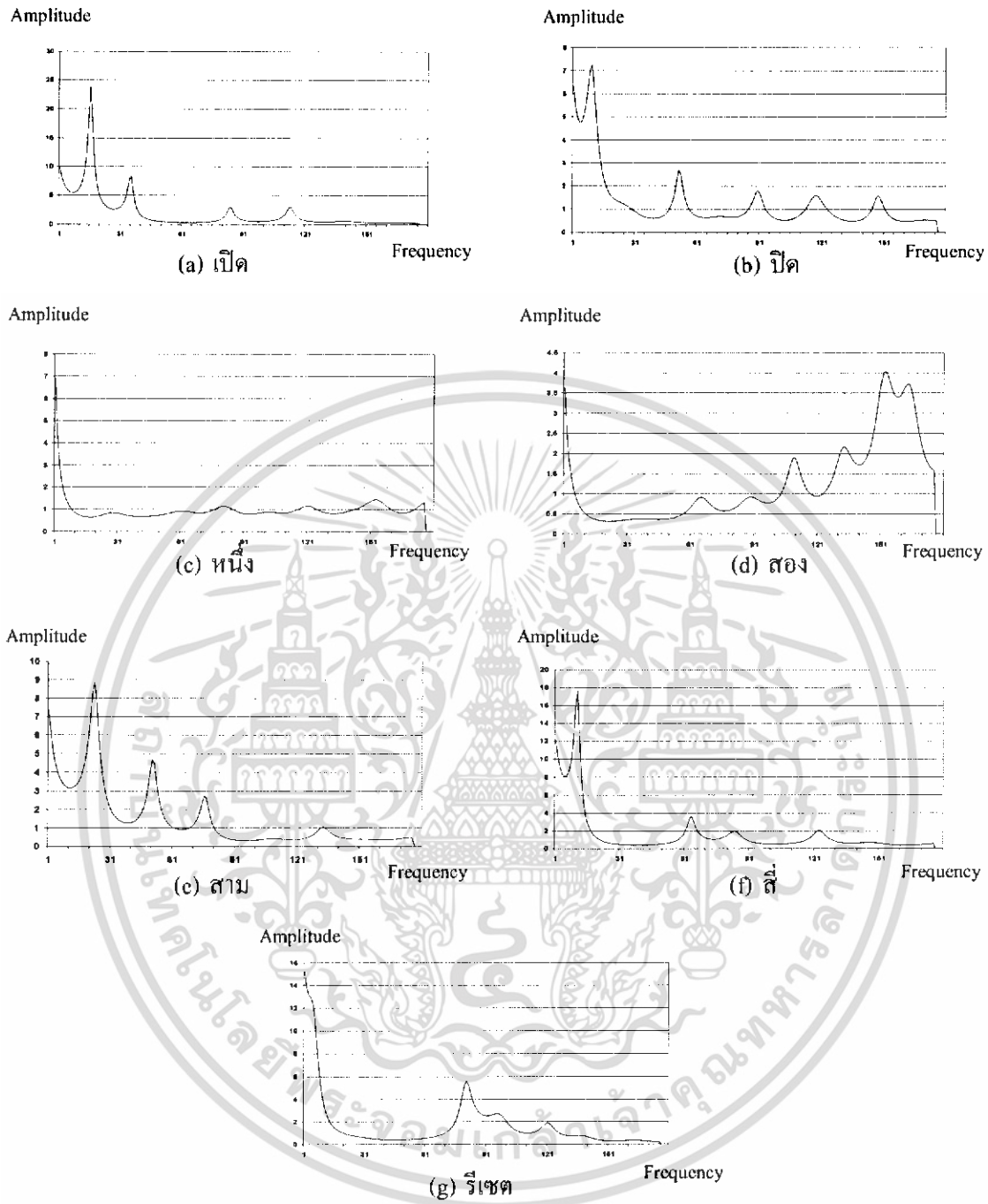
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2 ขั้นตอนการหาคุณลักษณะพิเศษของเสียง หรือ Feature Measurement



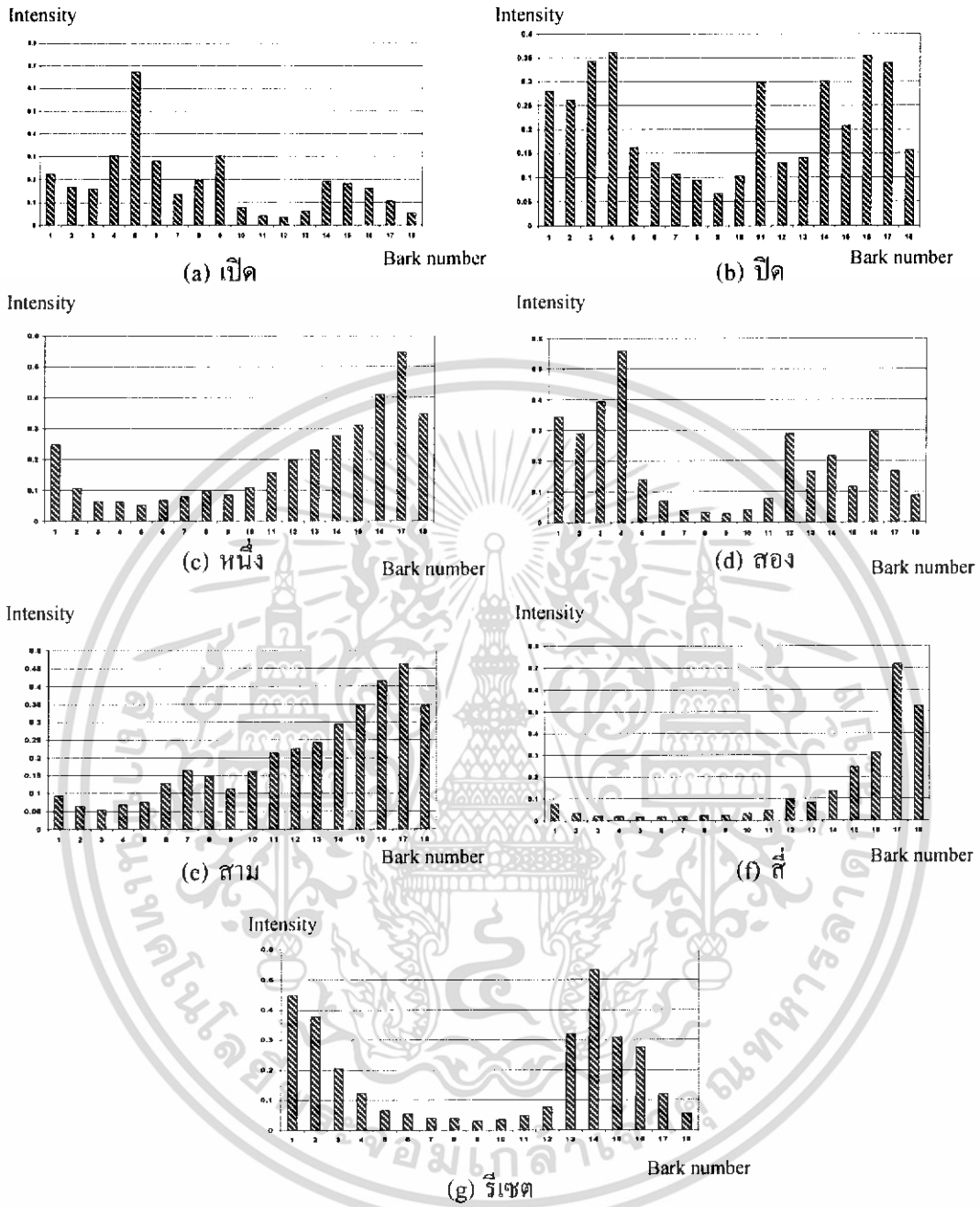
รูปที่ 4.16 แสดงสัมประสิทธิ์ LPC เฟรมที่ 1 ของแต่ละสัญญาณเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.17 แสดงสเปกตรัม LPC เฟรมที่ 1 ของแต่ละสัญญาณเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.18 แสดง Critical Band Intensity บน สเกลบาร์ค เฟรมที่ 1 ของแต่ละสัญญาณเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.12 แสดงตัวอย่างสัญญาณเสียงที่เราต้องทำการหาสเปกตรัมแอลพีซีโดยเริ่มต้นเราจะทำการพรีเอมฟาซีสเพื่อให้ความถี่สูงมีขนาดสูงขึ้นดังรูปที่ 4.14 สัญญาณที่ผ่านการพรีเอมฟาซิสแล้วจะถูกตัดออกเป็นช่วง ๆ หรือเฟรม เฟรมละ 300 แซมเปิ้ล โดยช่วงในการวิเคราะห์แต่ละช่วงจะถูกเลื่อนระยะไปครั้งละ 100 แซมเปิ้ล เนื่องมาจากเสียงพูดช่วงประมาณ 10 – 40 มิลลิวินาที จะอนุโลมได้ว่าสัญญาณเสียงพูดมีความเสถียรไม่แปรเปลี่ยนตามเวลา (Stationary) หรือมีลักษณะที่คงที่นั่นเอง เนื่องจากเราใช้ความถี่ในการสุ่มสัญญาณ 11.025 กิโลเฮิร์ต ดังนั้น

ช่วง 30 มิลลิวินาที $N = 30\times10^{-3}\times1\times11.025\times10^{3} \approx 330.75 \approx 300$

ช่วง 10 มิลลิวินาที $M = 10\times10^{-3}\times1\times11.025\times10^{3} \approx 110.25 \approx 100$

จากนั้นจึงทำการวินโดว์เพื่อลดความไม่ต่อเนื่องของสัญญาณที่บริเวณขอบโดยการใช้แฮมมิ่งวินโดว์ซึ่งจะได้ผลตามรูปที่ 4.15 แล้วจึงนำสัญญาณที่ผ่านการวินโดว์ไปทำการคำนวณหาสัมประสิทธิ์แอลพีซีตามที่กล่าวมาแล้วในหัวข้อ 2.4 ซึ่งจะได้สัมประสิทธิ์แอลพีซีดังรูป 4.16

สำหรับอันดับในการคำนวณสเปกตรัมแอลพีซีนั้นหากใช้อันดับต่ำก็จะทำให้การคำนวณเร็วขึ้นแต่ก็จะทำให้สเปกตรัมแอลพีซีแสดงรายละเอียดน้อยลงและส่งผลให้การคำนวณ Critical Band Intensity ผิดพลาดด้วย แต่ถ้าคำนวณแอลพีซีที่อันดับสูงมากก็จะทำให้การคำนวณซับซ้อนเกินความจำเป็น ได้มีผู้ทำการทดลองเพื่อหาความสัมพันธ์ระหว่างความผิดพลาดของการคำนวณ Critical Band Intensity กับอันดับการคำนวณแอลพีซี ดังแสดงตามรูปที่ 2.12 ซึ่งเห็นได้ว่าให้ผลที่สอดคล้องกันคือการคำนวณแอลพีซีที่ก่อให้เกิดประสิทธิภาพสูงสุดควรอยู่ที่อันดับที่ 19 (p = 19)

สัญญาณที่อยู่ในย่านความถี่ 5.5 กิโลเฮิร์ตนั้นจะอยู่ในบาร์คที่ 18 ซึ่งแสดงว่าจำนวนแถบความถี่ที่สามารถนำมาใช้ในการรู้จำมีทั้งหมด 18 แถบความถี่ และทำการคำนวณความเข้มในแต่ละแถบความถี่วิกฤตทั้ง 18 แถบความถี่วิกฤตก็จะได้จำนวนบาร์คทั้งหมด 18 ค่า สำหรับสัญญาณเสียงแต่ละเฟรมซึ่งเราสามารถนำไปใช้เป็นองค์ประกอบของ Feature Vector เพื่อที่จะใช้ในการสร้างแบบอ้างอิง

การใช้สเกลบาร์คนั้นช่วยให้สามารถแสดงการตอบสนองต่อความถี่ต่างๆของการได้ยินของมนุษย์ ได้เป็นเชิงเส้นตลอดช่วงความถี่ที่มนุษย์สามารถได้ยิน จึงทำการแสดงสเปกตรัมแอลพีซีบนสเกลบาร์คและใช้ความเข้มของสัญญาณในแต่ละแถบความถี่วิกฤตเป็นองค์ประกอบของเวกเตอร์แทนหน่วยเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. ขั้นตอนตอนการทดสอบแบบอ้างอิง

ตารางที่ 4.1 ผลการทดสอบแบบอ้างอิงจากผู้พูดในกรณีต่างๆ

	เปิด1	เปิด2	เปิด3	เปิด4	ปิด1	ปิด2	ปิด3	ปิด4	รีเซต	%
เปิด1	80	0	0	0	0	0	0	0	0	100
เปิด2	0	73	3	0	0	2	2	0	0	91.25
เปิด3	0	2	74	1	0	2	1	0	0	92.5
เปิด4	0	1	2	71	2	0	3	1	0	88.75
ปิด1	3	0	0	1	76	0	0	0	0	95
ปิด2	0	4	0	0	0	73	3	0	0	91.25
ปิด3	0	2	5	0	0	2	69	2	0	91.25
ปิด4	2	1	0	3	0	0	0	70	4	87.5
รีเซต	1	0	0	0	0	0	0	2	77	96.25
MEAN										92.18

(a) ผู้พูดเป็นผู้สร้างแบบเสียงอ้างอิง

	เปิด1	เปิด2	เปิด3	เปิด4	ปิด1	ปิด2	ปิด3	ปิด4	รีเซต	%
เปิด1	33	2	1	0	4	0	0	0	0	82.5
เปิด2	0	25	5	3	0	4	1	2	0	62.5
เปิด3	2	0	27	4	0	2	3	2	0	67.5
เปิด4	0	0	2	30	0	1	3	4	0	75
ปิด1	2	0	0	0	32	0	0	2	4	80
ปิด2	0	4	3	1	0	29	3	0	0	72.5
ปิด3	0	0	5	0	3	0	28	3	1	70
ปิด4	0	0	0	4	0	5	2	22	0	72.5
รีเซต	0	0	0	2	0	2	0	0	27	90
MEAN										72.81

(b) ผู้พูดไม่ได้เป็นผู้สร้างแบบเสียงอ้างอิง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดสอบแบบอ้างอิงโดยผู้พูดที่เป็นผู้สร้างแบบอ้างอิงในตารางที่ 4.1(a) นั้นได้เปอร์เซ็นต์ความถูกต้องรวมอยู่ที่ประมาณ 92.18 เปอร์เซ็นต์ โดยคำว่า เปิด หนึ่ง และ รีเซต มีความถูกต้องสูง ส่วนคำที่มีลักษณะใกล้เคียงกันมาก เช่น ปิดสอง ปิดสาม หรือ เปิดสอง จะมีความถูกต้องที่ต่ำกว่า อาจเนื่องมาจากลักษณะของคำที่มีสัญญาณเสียงที่แตกต่างกันอย่างไม่ชัดเจนนั่นเอง

ส่วนผลการทดสอบแบบอ้างอิงโดยผู้พูดที่ไม่ได้เป็นผู้สร้างแบบอ้างอิง นั้นในตารางที่ 4.1(b) จะให้เปอร์เซ็นต์ความถูกต้อง ที่ต่ำกว่า คือ ให้เปอร์เซ็นต์ความถูกต้องอยู่ที่ประมาณ 72.81 เปอร์เซ็นต์ แต่ในคำว่าเปิดหนึ่งและรีเซต ยังคงให้เปอร์เซ็นต์ความถูกต้องที่สูงเช่นเดิม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
บทสรุปและวิจารณ์

ในส่วนของฮาร์ดแวร์ จากการทดลองพบว่าวงจรที่ออกแบบสามารถควบคุมการเปิดปิดอุปกรณ์ไฟฟ้า เช่น หลอดไฟ, พัดลม หรือ โคมไฟได้จริง ซึ่งในวงจรนี้ สามารถเลือกการควบคุมได้ 2 ทางคือ เลือกการควบคุมโดยสวิตซ์ในวงจร หรือควบคุมด้วยเสียงผ่านพอร์ตอนุกรมของคอมพิวเตอร์ส่วนบุคคล

ในส่วนของซอฟแวร์ เมื่อสัญญาณเสียงที่เป็นข้อมูลดิบในการวิเคราะห์ผ่านกระบวนการต่างๆจนได้ค่าสัมประสิทธิ์แอลพีซีแล้ว จากนั้นจะนำสัมประสิทธิ์แอลพีซีไปแสดงบนสเกลบาร์ค และทำการคำนวณหาความเข้มสัญญาณในแต่ละแถบความถี่วิกฤตบนสเกลบาร์คแล้วจึงนำความเข้มของสัญญาณในทุกแถบความถี่วิกฤต ซึ่งในจำนวนนี้มี 18 แถบความถี่วิกฤต มาใช้เป็น Feature Vector

ส่วนในการทดสอบโมเดลอ้างอิงที่สร้างขึ้นด้วยวิธีการของ HMM นั้นได้แบ่งเป็น 2 แบบ คือ การทดสอบจากผู้พูดที่เป็นผู้สร้างแบบอ้างอิง และ การทดสอบจากผู้พูดที่ไม่ได้เป็นผู้สร้างแบบอ้างอิง ซึ่งให้ค่าความถูกต้อง ในการรู้จำประมาณ 92.18 และ 72.81 เปอร์เซ็นต์ ตามลำดับ โดยความผิดพลาดส่วนใหญ่จะเกิดขึ้นจากคำพูดที่มีลักษณะที่ใกล้เคียงกัน ซึ่งจากผลการทดลองจะพบความผิดพลาดในคำว่า ปิดสี่ ปิดสอง ซึ่งคำว่า สี่ และสอง มีลักษณะที่คล้ายกันมาก เป็นต้น

ปัญหาที่พบในการทดลอง

ในส่วนของฮาร์ดแวร์

1. ฟิวส์ที่ใช้ในวงจรจะต้องมีการคำนวณให้เหมาะสมกับกระแสของโหลดทั้งหมด

2. ควรใช้แผ่นระบายความร้อน ที่มีขนาดเหมาะสมเพื่อป้องกันไม่ให้ไตรแอกร้อนจนเกิดความเสียหาย

3.ในการใช้งานจริง ถ้าระดับแรงดันไม่คงที่ เช่น การเกิดไฟตก หรือการเปลี่ยนระดับแรงดันของอุปกรณ์ไฟฟ้า จะทำให้วงจรไม่สามารถทำงานได้

ในส่วนซอฟแวร์

1. ปัญหาที่สำคัญที่สุดคือ วิธีการในการอัดเสียงที่ใช้เป็นข้อมูลดิบนั้น ต้องมีสัญญาณรบกวนที่น้อยที่สุด อีกทั้งยังต้องพูดให้อยู่ในช่วงเวลาตามที่กำหนดไว้ด้วย

2. การใช้คำในการทดลองมีความใกล้เคียงกันบ้างบางคำทำให้การตัดสินใจเลือกแบบอ้างอิงมีความถูกต้องลดลง

3.ในการออกเสียงคำสั่งจะต้องพูดด้วยน้ำเสียงปกติและชัดเจนเพื่อความถูกต้องในการประมวลผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หนังสืออ้างอิง

[1] วรา คงคาวิชูรย์. "การรู้จำเสียงสระเสียงเดี่ยวในภาษาไทยโดยใช้ทรานส์เฟอร์ฟังก์ชันบนสเกลบาร์ค", วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง. 2542.

[2] ทัศเวท วีระวัฒน์. "การรู้จำเสียงคำไทยเฉพาะบุคคล" วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2541.

[3] วิไลวรรณ ขนิษฐานันท์. "ภาษาและภาษาศาสตร์", พิมพ์ครั้งที่ 5. กรุงเทพฯ : สำนักพิมพ์มหาวิทยาลัยธรรมศาสตร์, 2533.

[4] สุนทร อรอินทร์, อัฐ เครือฟัก. "การประมวลเสียงพูดโดยการประมาณเชิงเส้น", วิทยานิพนธ์ภาควิศวกรรมไฟฟ้า สาขาวิชาวิศวกรรมโทรคมนาคม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง. 2538.

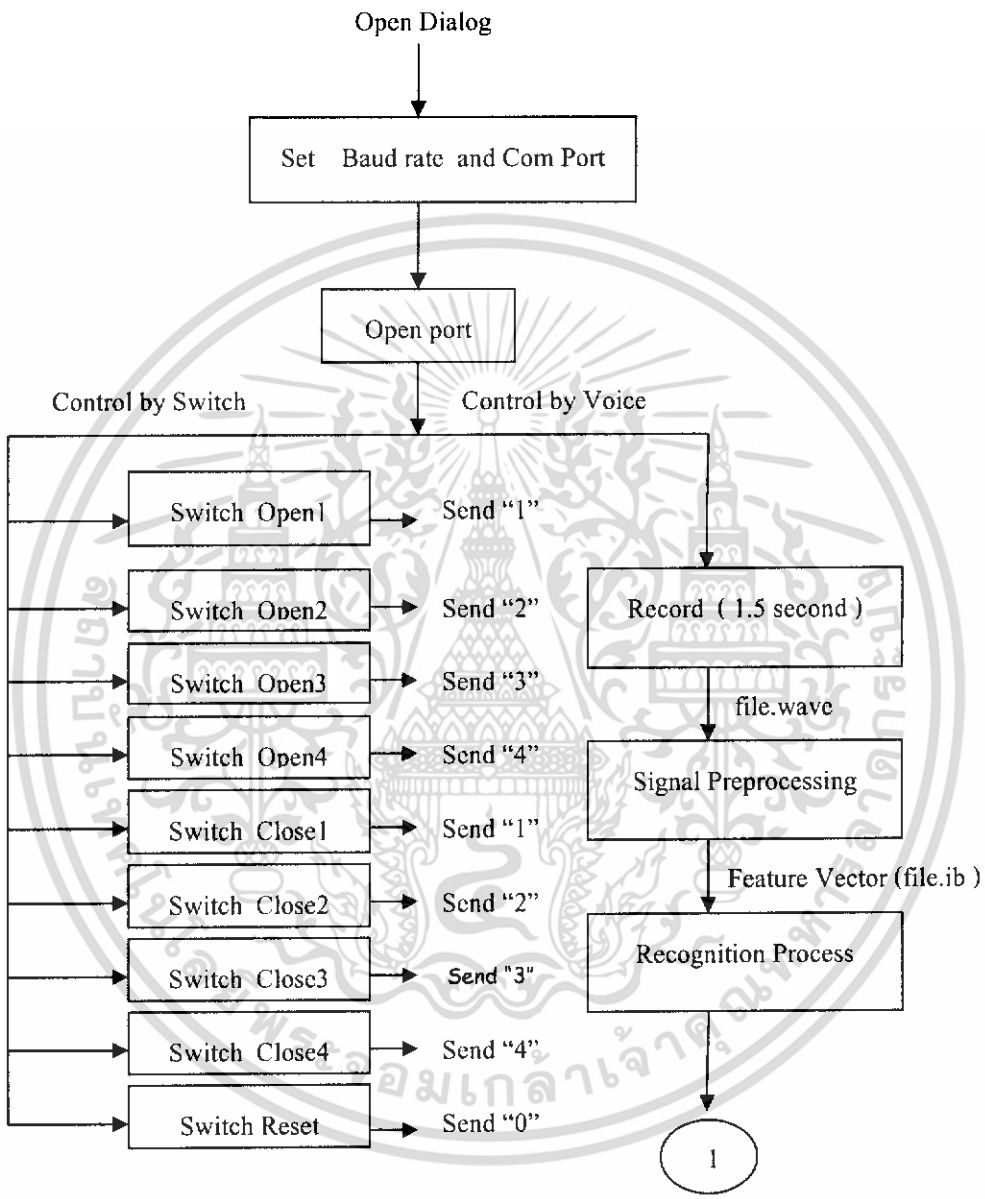
[5] กัญน์ สิทธิประเสริฐ. "การแยกสระเสียงสั้น และสระเสียงยาวในภาษาไทยโดยใช้การฟิตเคร์ฟพหุนามอันดับสอง บนฟังก์ชันพลังงานเสียง", วิทยานิพนธ์วิศวกรรมศาสตร์มหาบัณฑิต สาขาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2547

[6] นงนุช สุขตั้งมั่น. "การรู้จำสระเสียงเดี่ยวและสระเสียงประสมในภาษาไทยโดยการใช้สเปกตรัม แอลพีซีบนสเกลบาร์ค", วิทยานิพนธ์วิศวกรรมศาสตร์มหาบัณฑิต สาขาวิศวกรรมไฟฟ้า บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2547

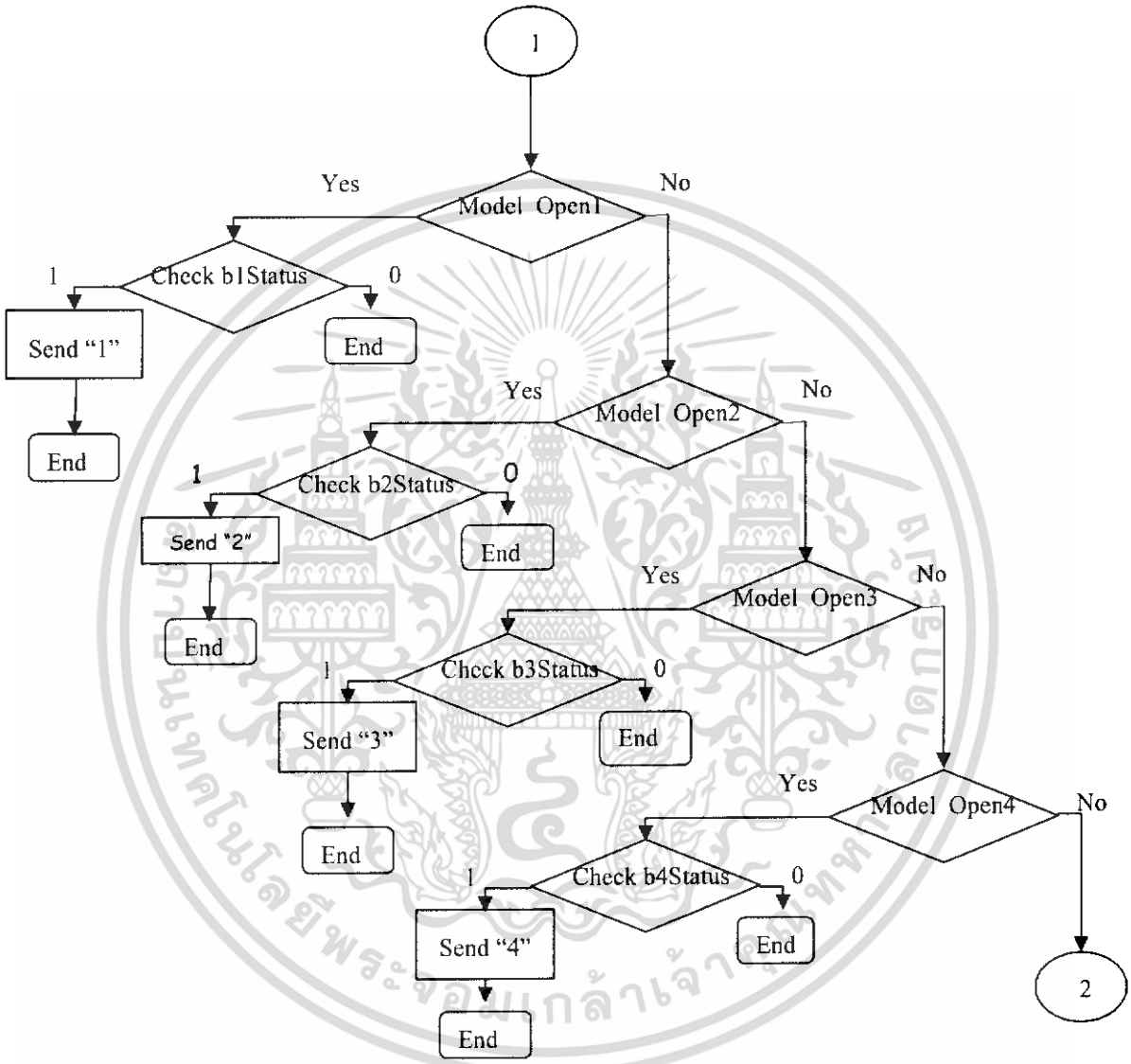
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

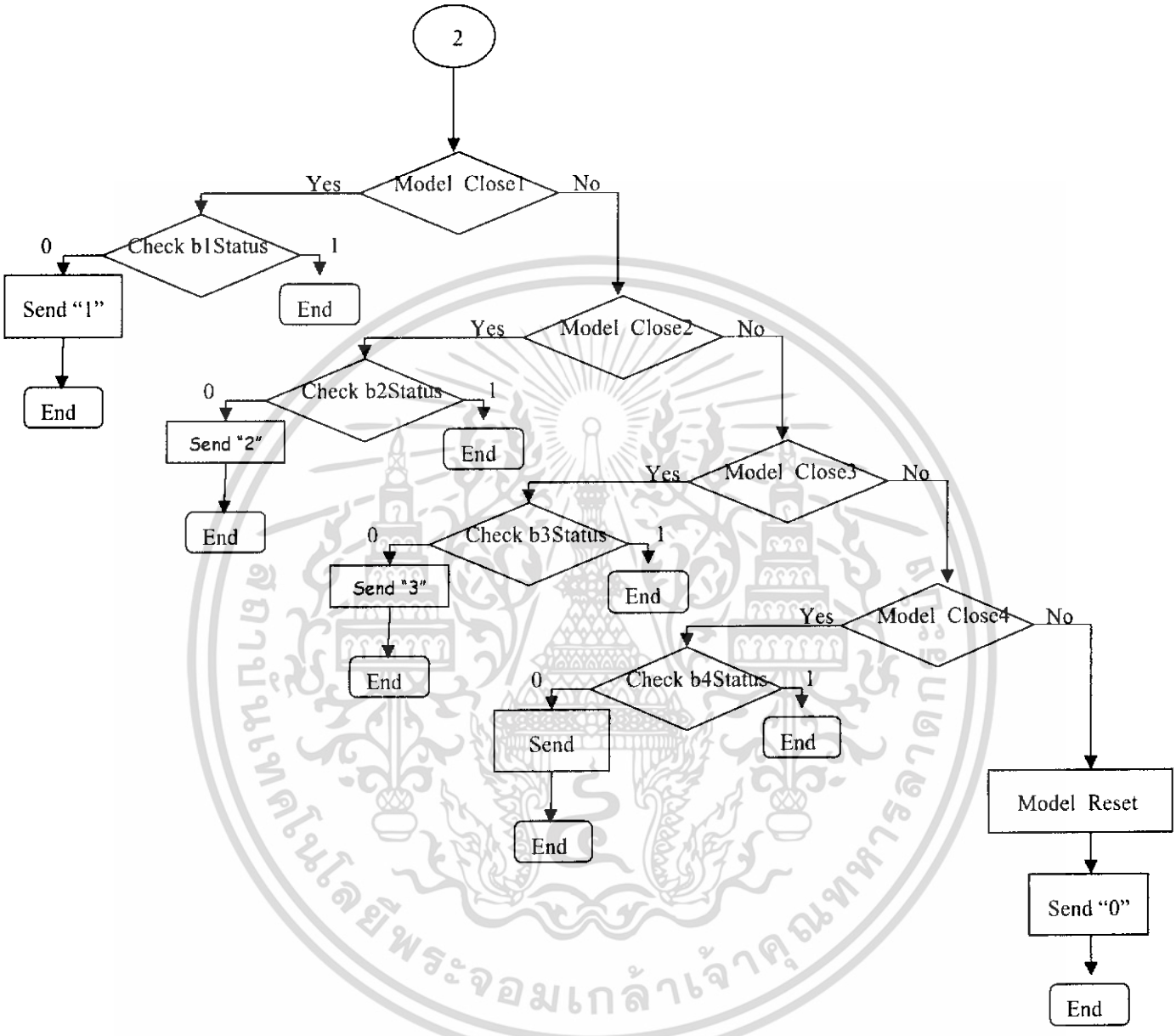
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

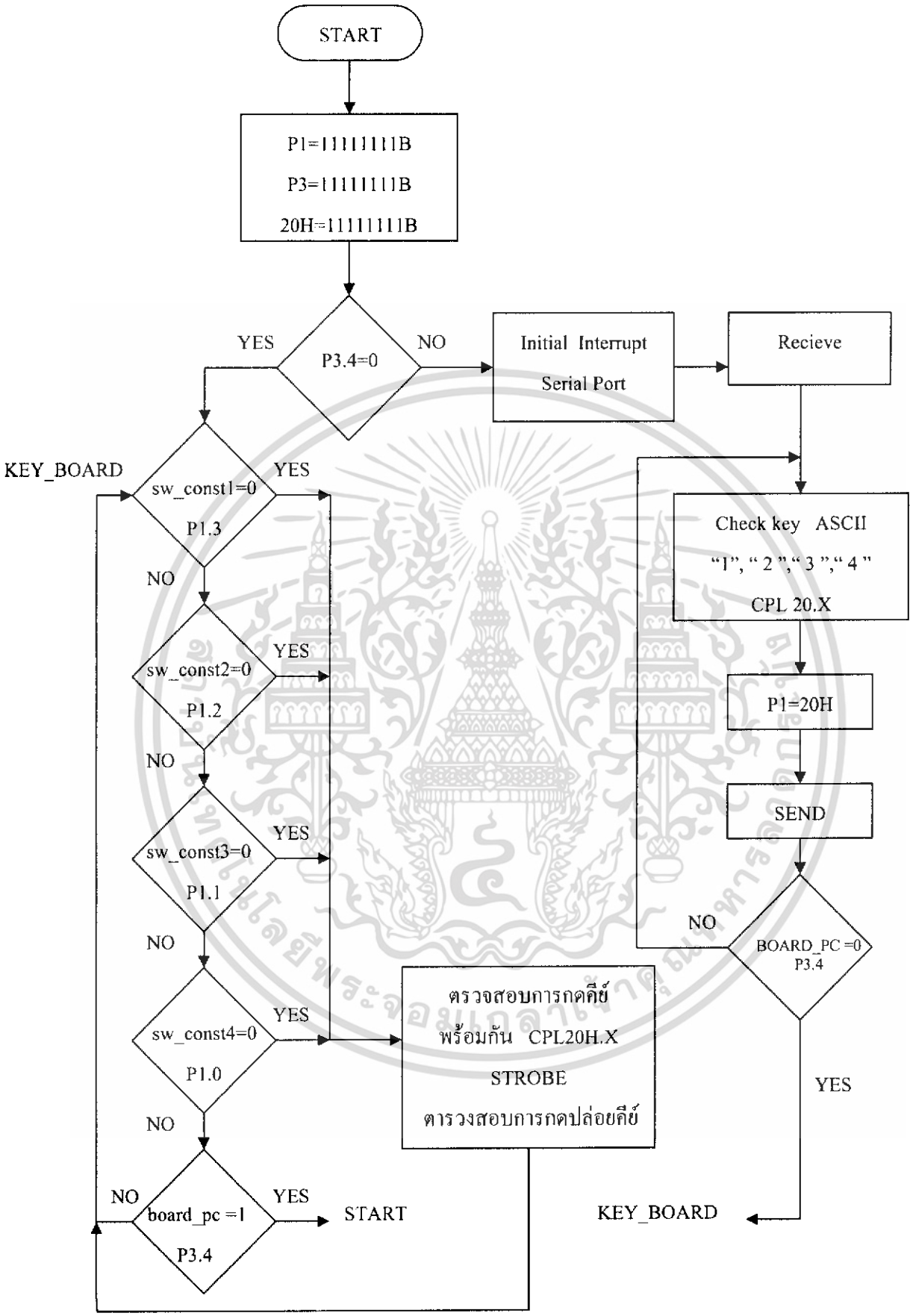


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



Flow chart แสดงขั้นตอนการทำงานของโครงงานนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



Flow chart แสดงขั้นตอนการทำงานของโปรแกรมสวิตช์เปิดปิด 4 ช่อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
      ;*****************************************************
      ;*****  RS232 AND MCS 51 CONTROL PORT  ************
      ;*****************************************************
LOAD1       BIT         07H
LOAD2       BIT         06H
LOAD3       BIT         05H
LOAD4       BIT         04H
STORBE            BIT         P3.7
SW_CONT1    BIT         P1.3
SW_CONT2    BIT         P1.2
SW_CONT3    BIT         P1.1
SW_CONT4    BIT         P1.0
BOARD_PC    BIT         P3.4
CR          EQU                         0DH
LF          EQU         0AH
      ;*****************************************************
            ORG         0000H
START:            MOV         P1,#0FFH
            MOV         P3,#0FFH
            MOV         20H,#0FFH
            JNB         BOARD_PC,KEY_BOARD
            ACALL       INIT
            MOV         DPTR,#SHOW
            ACALL       DISPLAY
MAIN:       ACALL       RECEIVE
            ;ACALL            SEND
            JNB         BOARD_PC,KEY_BOARD
            SJMP        MAIN
      ;*****************************************************
            ;MOV         IE,#00000000B
INIT:       ;MOV        TL1,#0FDH
            MOV         PCON,#00H
            MOV         SCON,#50H
            MOV         TMOD,#20H
            MOV         TH1,#0FDH
            SETB        TR1
            RET
CHECK_KEY:  CJNE        R1,#31H,LOAD_2
            CPL         LOAD1
            ACALL       SIG_STORBE
            SJMP        END_CHK
LOAD_2:           CJNE        R1,#32H,LOAD_3
            CPL         LOAD2
            ACALL       SIG_STORBE
            SJMP        END_CHK
LOAD_3:           CJNE        R1,#33H,LOAD_4
            CPL         LOAD3
            ACALL       SIG_STORBE
            SJMP        END_CHK
LOAD_4:           CJNE        R1,#34H,CLR_DATA
            CPL         LOAD4
CLR_DATA:   CJNE        R1,#30H,END_CHK
            MOV         20H,#0FFH
            ACALL       SIG_STORBE
END_CHK:    MOV         P1,20H
            RET
      ;*****************************************************
```

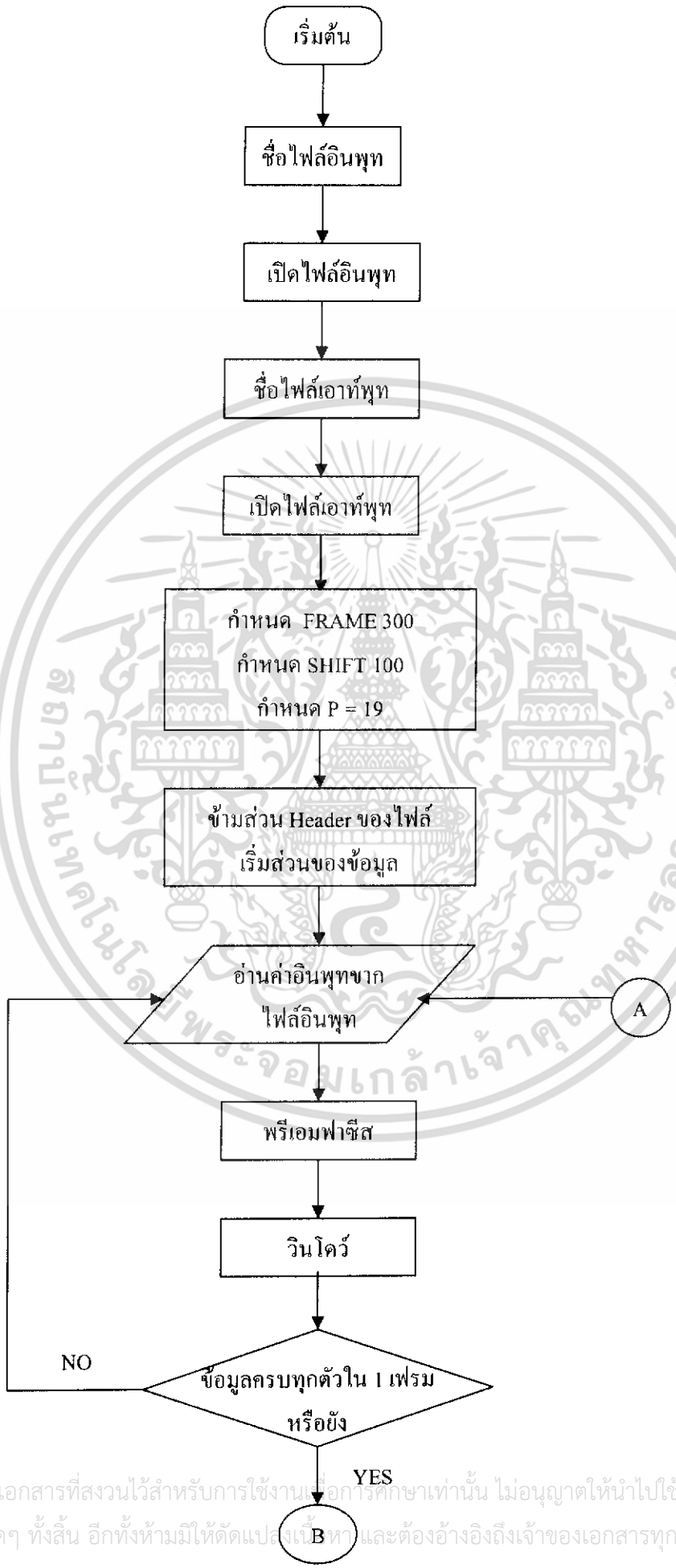
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
KEY_BOARD:  MOV         P1,20H
SW_1:       JNB         SW_CONT1,CHK_LOAD1
SW_2:       JNB         SW_CONT2,CHK_LOAD2
SW_3:       JNB         SW_CONT3,CHK_LOAD3
SW_4:       JNB         SW_CONT4,CHK_LOAD4
            JB          BOARD_PC,START
            SJMP        SW_1
CHK_LOAD1:  ACALL       DELAY_D
            JB          SW_CONT1,SW_2
            JNB         SW_CONT2,SW_1
            JNB         SW_CONT3,SW_1
            JNB         SW_CONT4,SW_1
            CPL         LOAD1
            ACALL             SIG_STORBE
            JNB         SW_CONT1,$
            SJMP        KEY_BOARD
CHK_LOAD2:  ACALL       DELAY_D
            JB          SW_CONT2,SW_3
            JNB         SW_CONT1,SW_1
            JNB         SW_CONT3,SW_1
            JNB         SW_CONT4,SW_1
            CPL         LOAD2
            ACALL       SIG_STORBE
            JNB         SW_CONT2,$
            SJMP        KEY_BOARD
CHK_LOAD3:  ACALL       DELAY_D
            JB          SW_CONT3,SW_4
            JNB         SW_CONT1,SW_1
            JNB         SW_CONT2,SW_1
            JNB         SW_CONT4,SW_1
            CPL         LOAD3
            ACALL       SIG_STORBE
            JNB         SW_CONT3,$
            SJMP        KEY_BOARD
CHK_LOAD4:  ACALL       DELAY_D
            JB          SW_CONT4,SW_1
            JNB         SW_CONT1,SW_1
            JNB         SW_CONT2,SW_1
            JNB         SW_CONT3,SW_1
            CPL         LOAD4
            ACALL       SIG_STORBE
            JNB         SW_CONT4,$
            SJMP        KEY_BOARD
      ;*************************************************
SIG_STORBE: CLR         STORBE
            ACALL       DELAY_D
            SETB        STORBE
            RET
      ;*****************************************************
      ;*** RECEIVE DATA FORM RS232 TO MCS51  *************
      ;*****************************************************
RECEIVE:    JNB   RI,$
            CLR         RI
            MOV         A,SBUF
            ACALL       SEND
            MOV         R1,A
            ACALL       CHECK_KEY
            RET
```

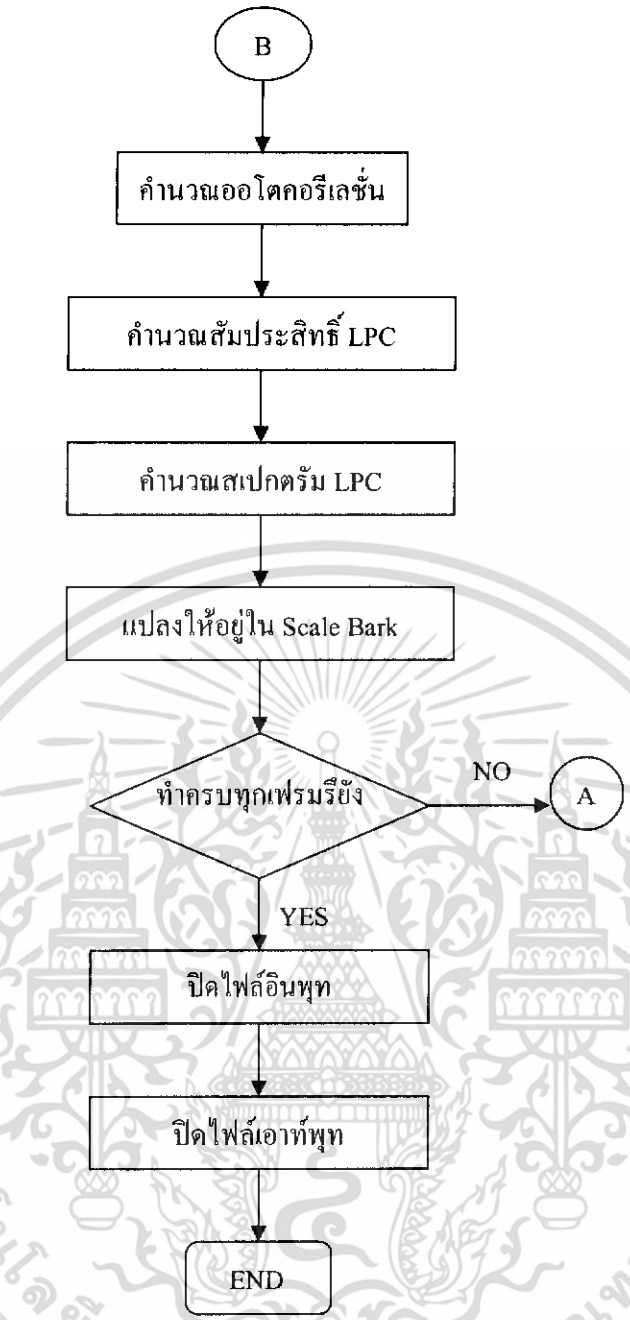
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
      ;*******************************************************
      ;******** SENT DATA FROM MCS51 TO RS232 *************
      ;*******************************************************
SEND:       MOV         SBUF,A
            JNB         TI,$
            CLR         TI
            RET
      ;*******************************************************
DELAY_D:    MOV         R1,#90H
DELAY_1:    MOV   R2,#0FFH
            DJNZ         R2,$
            DJNZ        R1,DELAY_1
            RET
      ;*******************************************************
DISPLAY:    MOV   A,#00
            MOVC        A,@A+DPTR
            ACALL       SEND
            INC         DPTR
            JZ          END_SHOW
            SJMP  DISPLAY
END_SHOW:   RET
SHOW:       DB          CR,LF,'VOICE RECONITION BY 4A',CR,LF
            DB          00H
            END
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



Flow chart แสดงขั้นตอนการหาค่า Feature Vector

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Features

- Compatible with MCS-51™ Products
- 2K Bytes of Reprogrammable Flash Memory
  - Endurance: 1,000 Write/Erase Cycles
- 2.7V to 6V Operating Range
- Fully Static Operation: 0 Hz to 24 MHz
- Two-Level Program Memory Lock
- 128 x 8-Bit Internal RAM
- 15 Programmable I/O Lines
- Two 16-Bit Timer/Counters
- Six Interrupt Sources
- Programmable Serial UART Channel
- Direct LED Drive Outputs
- On-Chip Analog Comparator
- Low Power Idle and Power Down Modes

8-Bit Microcontroller with 2K Bytes Flash

AT89C2051

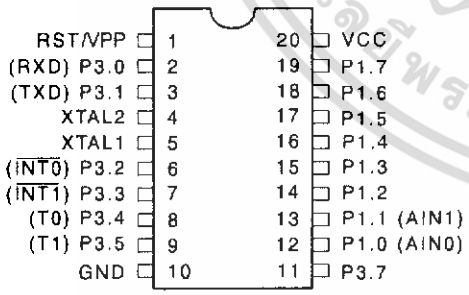
Description

The AT89C2051 is a low-voltage, high-performance CMOS 8-bit microcomputer with 2K Bytes of Flash programmable and erasable read only memory (PEROM). The device is manufactured using Atmel's high density nonvolatile memory technology and is compatible with the industry standard MCS-51™ instruction set. By combining a versatile 8-bit CPU with Flash on a monolithic chip, the Atmel AT89C2051 is a powerful microcomputer which provides a highly flexible and cost effective solution to many embedded control applications.

The AT89C2051 provides the following standard features: 2K Bytes of Flash, 128 bytes of RAM, 15 I/O lines, two 16-bit timer/counters, a five vector two-level interrupt architecture, a full duplex serial port, a precision analog comparator, on-chip oscillator and clock circuitry. In addition, the AT89C2051 is designed with static logic for operation down to zero frequency and supports two software selectable power saving modes. The Idle Mode stops the CPU while allowing the RAM, timer/counters, serial port and interrupt system to continue functioning. The Power Down Mode saves the RAM contents but freezes the oscillator disabling all other chip functions until the next hardware reset.

Pin Configuration

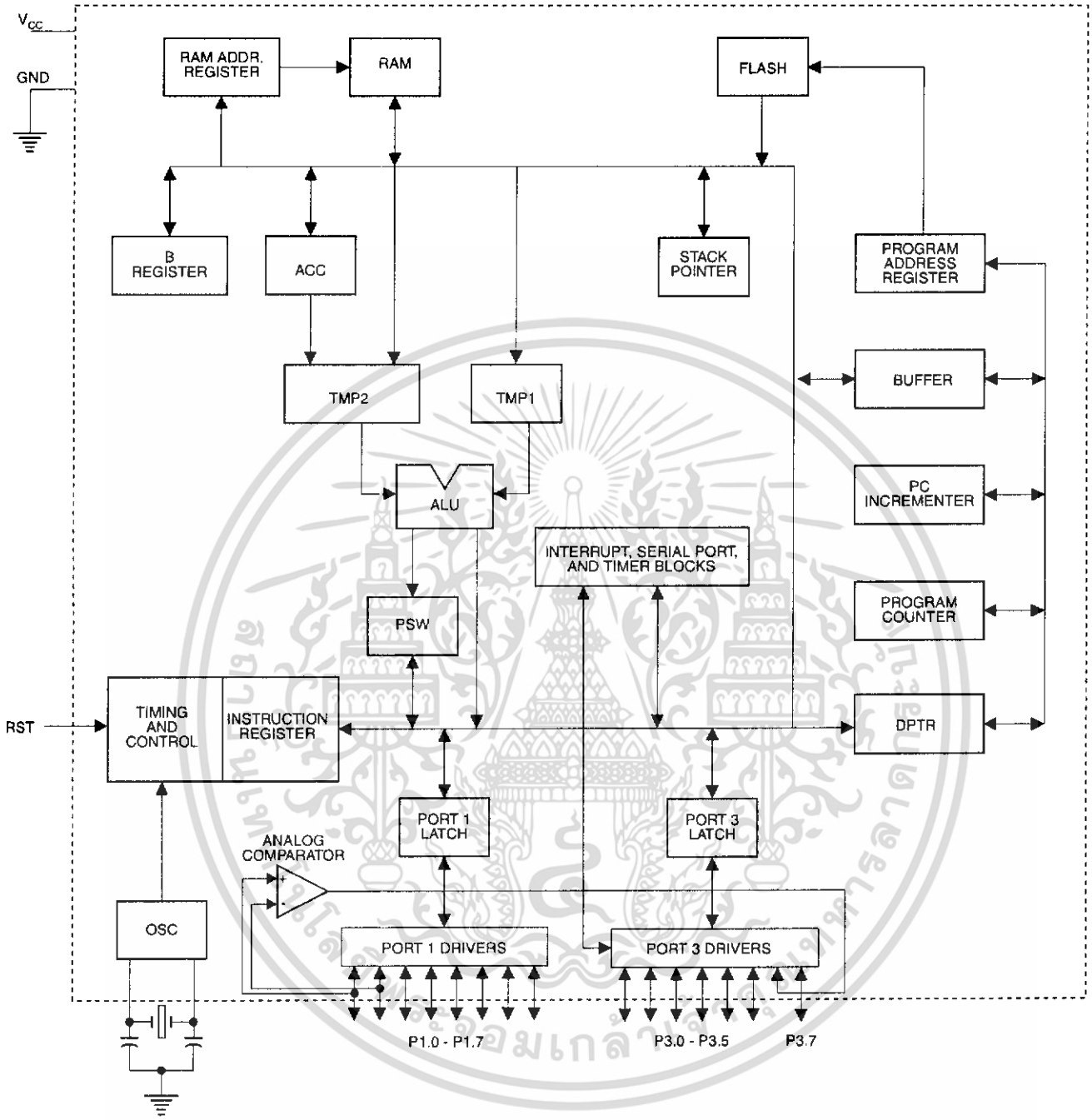
PDIP/SOIC



0368D-B–12/97

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Block Diagram



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Pin Description

V_{CC}
Supply voltage.

GND
Ground.

Port 1
Port 1 is an 8-bit bidirectional I/O port. Port pins P1.2 to P1.7 provide internal pullups. P1.0 and P1.1 require external pullups. P1.0 and P1.1 also serve as the positive input (AIN0) and the negative input (AIN1), respectively, of the on-chip precision analog comparator. The Port 1 output buffers can sink 20 mA and can drive LED displays directly. When 1s are written to Port 1 pins, they can be used as inputs. When pins P1.2 to P1.7 are used as inputs and are externally pulled low, they will source current (I_{IL}) because of the internal pullups.

Port 1 also receives code data during Flash programming and verification.

Port 3
Port 3 pins P3.0 to P3.5, P3.7 are seven bidirectional I/O pins with internal pullups. P3.6 is hard-wired as an input to the output of the on-chip comparator and is not accessible as a general purpose I/O pin. The Port 3 output buffers can sink 20 mA. When 1s are written to Port 3 pins they are pulled high by the internal pullups and can be used as inputs. As inputs, Port 3 pins that are externally being pulled low will source current (I_{IL}) because of the pullups.

Port 3 also serves the functions of various special features of the AT89C2051 as listed below:

Port Pin	Alternate Functions
P3.0	RXD (serial input port)
P3.1	TXD (serial output port)
P3.2	$\overline{INT0}$ (external interrupt 0)
P3.3	$\overline{INT1}$ (external interrupt 1)
P3.4	T0 (timer 0 external input)
P3.5	T1 (timer 1 external input)

Port 3 also receives some control signals for Flash programming and verification.

RST
Reset input. All I/O pins are reset to 1s as soon as RST goes high. Holding the RST pin high for two machine cycles while the oscillator is running resets the device.

Each machine cycle takes 12 oscillator or clock cycles.

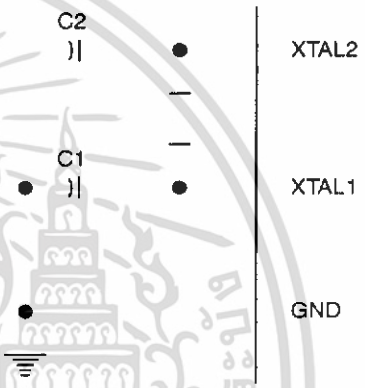
XTAL1
Input to the inverting oscillator amplifier and input to the internal clock operating circuit.

XTAL2
Output from the inverting oscillator amplifier.

Oscillator Characteristics

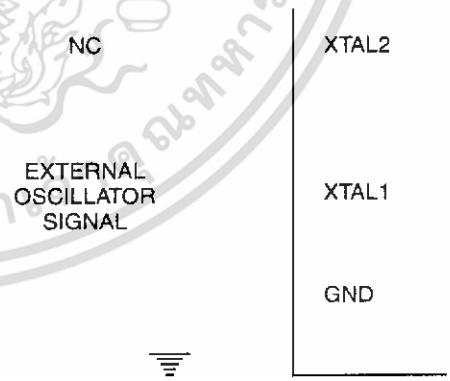
XTAL1 and XTAL2 are the input and output, respectively, of an inverting amplifier which can be configured for use as an on-chip oscillator, as shown in Figure 1. Either a quartz crystal or ceramic resonator may be used. To drive the device from an external clock source, XTAL2 should be left unconnected while XTAL1 is driven as shown in Figure 2. There are no requirements on the duty cycle of the external clock signal, since the input to the internal clocking circuitry is through a divide-by-two flip-flop, but minimum and maximum voltage high and low time specifications must be observed.

Figure 1. Oscillator Connections



Note: C1, C2 = 30 pF ± 10 pF for Crystals
= 40 pF ± 10 pF for Ceramic Resonators

Figure 2. External Clock Drive Configuration



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Special Function Registers

A map of the on-chip memory area called the Special Function Register (SFR) space is shown in the table below.

Note that not all of the addresses are occupied, and unoccupied addresses may not be implemented on the chip. Read accesses to these addresses will in general return random data, and write accesses will have an indeterminate effect.

User software should not write 1s to these unlisted locations, since they may be used in future products to invoke new features. In that case, the reset or inactive values of the new bits will always be 0.

Table 1. AT89C2051 SFR Map and Reset Values

0F8H									0FFH
0F0H	B 00000000								0F7H
0E8H									0EFH
0E0H	ACC 00000000								0E7H
0D8H									0DFH
0D0H	PSW 00000000								0D7H
0C8H									0CFH
0C0H									0C7H
0B8H	IP XXX00000								0BFH
0B0H	P3 11111111								0B7H
0A8H	IE 0XX00000								0AFH
0A0H									0A7H
98H	SCON 00000000	SBUF XXXXXXXX							9FH
90H	P1 11111111								97H
88H	TCON 00000000	TMOD 00000000	TL0 00000000	TL1 00000000	TH0 00000000	TH1 00000000			8FH
80H		SP 00000111	DPL 00000000	DPH 00000000				PCON 0XXX0000	87H

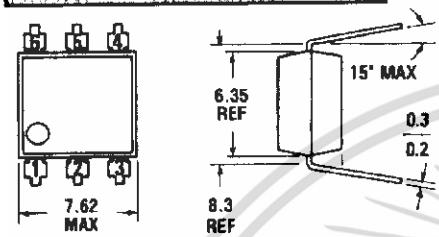
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

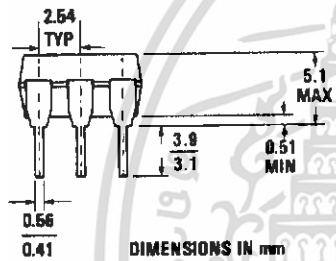
QT OPTOELECTRONICS

NON-ZERO-CROSSING TRIACS

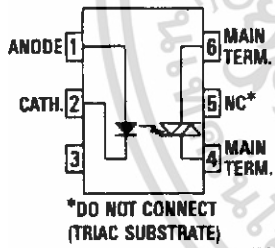
MOC3020 MOC3021 MOC3022 MOC3023

PACKAGE DIMENSIONS





DIMENSIONS IN mm
PACKAGE CODE E
ST1603



Equivalent Circuit

DESCRIPTION

The MOC3020, MOC3021, MOC3022 and MOC3023 are optically isolated triac driver devices. These devices contain a GaAs infrared emitting diode and a light activated silicon bilateral switch, which functions like a triac. This is designed for interfacing between electronic controls and power triacs to control resistive and inductive loads for 240 VAC operations.

FEATURES

- Excellent I_F stability—IR emitting diode has low degradation
- High isolation voltage—minimum 7500 VAC peak
- Underwriters Laboratory (UL) recognized—File #E90700

APPLICATIONS

- European applications for 240 VAC
- Triac driver
- Industrial controls
- Traffic lights
- Vending machines
- Motor control
- Solid state relay

ABSOLUTE MAXIMUM RATINGS

TOTAL PACKAGE

Storage temperature	−55°C to 150°C
Operating temperature	−40°C to 100°C
Lead temperature (soldering, 10 sec)	260°C

INPUT DIODE

Forward DC current	50 mA
Reverse voltage	3 V
Peak forward current (1 μs pulse, 300 pps)	3.0 A
Power dissipation (25°C ambient)	100 mW
Derate linearly (above 25°C ambient)	1.33 mW/°C

OUTPUT DRIVER

Off-state output terminal voltage		400 Volts
On-state RMS current (Full cycle, 50 to 60 Hz)	T_A=25°C	100 mA
	T_A=70°C	50 mA
Peak nonrepetitive surge current (PW=10 ms, DC=10%)		1.2 A
Total power dissipation (25°C ambient)		300 mW
Derate above 25°C		4.0 mW/°C

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ELECTRO-OPTICAL CHARACTERISTICS (25°C Temperature Unless Otherwise Specified)

INDIVIDUAL COMPONENT CHARACTERISTICS

CHARACTERISTIC	SYMBOL	MIN.	TYP.	MAX.	UNITS	TEST CONDITIONS
INPUT DIODE						
Forward voltage	V_f		1.2	1.50	V	I_F=10 mA
Junction capacitance	C_J		50		pF	V_F=0 V, f=1 MHz
Reverse leakage current	I_R			100	μA	V_R=3.0 V
OUTPUT DETECTOR						
Peak blocking current, either direction	I_{DRM}	—	10	100	nA	V_{DRM}=400 V, Note 1
Peak on-state voltage, either direction	V_{TM}	—	2.5	3.0	Volts	I_{TM}=100 mA Peak

Note 1. Test voltage must be applied within dv/dt rating.

TRANSFER CHARACTERISTICS

DC CHARACTERISTICS		SYMBOL	MIN.	TYP.	MAX.	UNITS	TEST CONDITIONS
LED trigger current (current required to latch output)	MOC3020	I_{FT}	—	—	30	mA	Main terminal voltage=3.0 V, R_L=150Ω
	MOC3021	I_{FT}	—	—	15	mA	
	MOC3022	I_{FT}	—	—	10	mA	
	MOC3023	I_{FT}	—	—	5	mA	
Holding current		I_H	—	100	—	μA	Either direction

TRANSFER CHARACTERISTICS

CHARACTERISTICS	SYMBOL	MIN.	TYP.	MAX.	UNITS	TEST CONDITIONS
dv/dt RATING Critical rate of rise of off-state voltage	dv/dt	—	12	—	V/μs	Static dv/dt, T_A=85°C (see Fig. 3)
Critical rate of rise of commutating voltage	dv/dt	—	0.2	—	V/μs	Commutating dv/dt I_{LOAD}=15 mA (see Fig. 4)

ISOLATION CHARACTERISTICS

CHARACTERISTICS	SYMBOL	MIN.	TYP.	MAX.	UNITS	TEST CONDITIONS
Isolation voltage	V_{iso}	5300			V_{AC}RMS	I_{I-O} ≤ 1 μA, 1 Minute
	V_{iso}	7500			V_{AC}PEAK	I_{I-O} ≤ 1 μA, 1 Minute
Isolation resistance	R_{iso}	10^{11}			ohms	V_{I-O}=500 VDC
Isolation capacitance	C_{iso}		0.5		pF	f=1 MHz

Note 1: Ratings apply to either polarity of pin 6 — referenced to pin 4. Voltages must be applied within dv/dt rating.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

TYPICAL ELECTRICAL CHARACTERISTIC CURVES

(25°C Free Air Temperature Unless Otherwise Specified)

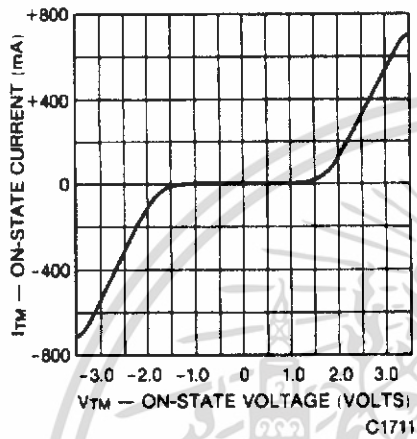


Fig. 1. On-State Characteristics

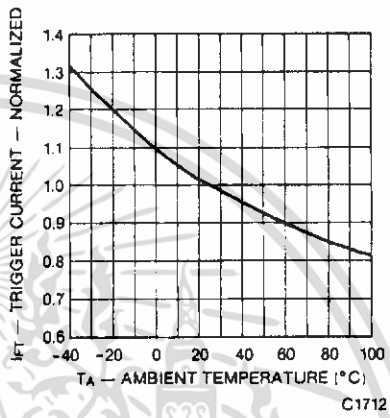
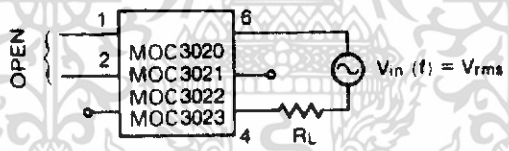


Fig. 2. Trigger Current vs. Temperature

TEST CIRCUITS FOR dV/dt MEASUREMENTS



$$\frac{dV}{dt} = wV_{pack} = 2\pi f \times 1.414\ V_{rms}$$
$$= 8.88\ f\ V_{rms}$$

Fig. 3. Static dV/dt

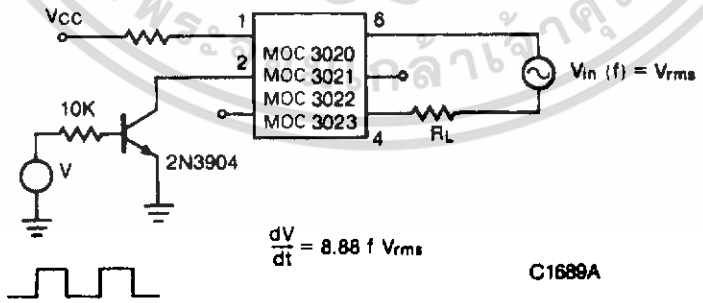


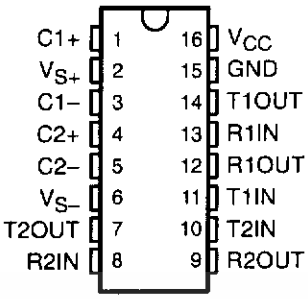
Fig. 4. Commutating dV/dt

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

SLLS047I – FEBRUARY 1989 – REVISED OCTOBER 2002

- **Meet or Exceed TIA/EIA-232-F and ITU Recommendation V.28**
- **Operate With Single 5-V Power Supply**
- **Operate Up to 120 kbit/s**
- **Two Drivers and Two Receivers**
- **±30-V Input Levels**
- **Low Supply Current . . . 8 mA Typical**
- **Designed to be Interchangeable With Maxim MAX232**
- **ESD Protection Exceeds JESD 22**
  - **2000-V Human-Body Model (A114-A)**
- **Applications**
  - **TIA/EIA-232-F**
  - **Battery-Powered Systems**
  - **Terminals**
  - **Modems**
  - **Computers**

MAX232 . . . D, DW, N, OR NS PACKAGE
MAX232I . . . D, DW, OR N PACKAGE
(TOP VIEW)



description/ordering information

The MAX232 is a dual driver/receiver that includes a capacitive voltage generator to supply EIA-232 voltage levels from a single 5-V supply. Each receiver converts EIA-232 inputs to 5-V TTL/CMOS levels. These receivers have a typical threshold of 1.3 V and a typical hysteresis of 0.5 V, and can accept ±30-V inputs. Each driver converts TTL/CMOS input levels into EIA-232 levels. The driver, receiver, and voltage-generator functions are available as cells in the Texas Instruments LinASIC™ library.

ORDERING INFORMATION

T_A	PACKAGE†		ORDERABLE PART NUMBER	TOP-SIDE MARKING
0°C to 70°C	PDIP (N)	Tube	MAX232N	MAX232N
	SOIC (D)	Tube	MAX232D	MAX232
		Tape and reel	MAX232DR	
	SOIC (DW)	Tube	MAX232DW	MAX232
		Tape and reel	MAX232DWR	
	SOP (NS)	Tape and reel	MAX232NSR	MAX232
–40°C to 85°C	PDIP (N)	Tube	MAX232IN	MAX232IN
	SOIC (D)	Tube	MAX232ID	MAX232I
		Tape and reel	MAX232IDR	
	SOIC (DW)	Tube	MAX232IDW	MAX232I
		Tape and reel	MAX232IDWR	

† Package drawings, standard packing quantities, thermal data, symbolization, and PCB design guidelines are available at www.ti.com/sc/package.

Please be aware that an important notice concerning availability, standard warranty, and use in critical applications of Texas Instruments semiconductor products and disclaimers thereto appears at the end of this data sheet.

LinASIC is a trademark of Texas Instruments.

PRODUCTION DATA information is current as of publication date. Products conform to specifications per the terms of Texas Instruments standard warranty. Production processing does not necessarily include testing of all parameters.

Copyright © 2002, Texas Instruments Incorporated

TEXAS INSTRUMENTS
POST OFFICE BOX 655303 • DALLAS, TEXAS 75265

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Function Tables

EACH DRIVER

INPUT TIN	OUTPUT TOUT
L	H
H	L

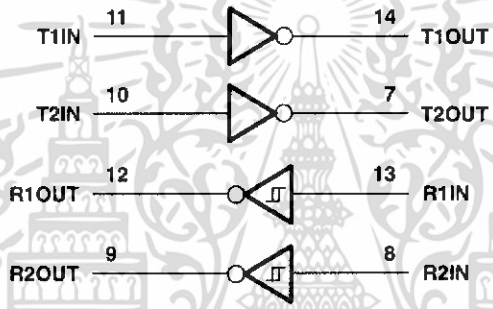
H = high level, L = low level

EACH RECEIVER

INPUT RIN	OUTPUT ROUT
L	H
H	L

H = high level, L = low level

logic diagram (positive logic)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

absolute maximum ratings over operating free-air temperature range (unless otherwise noted)†

Input supply voltage range, V_{CC} (see Note 1) . . . . . –0.3 V to 6 V
Positive output supply voltage range, V_{S+} . . . . . V_{CC} – 0.3 V to 15 V
Negative output supply voltage range, V_{S-} . . . . . –0.3 V to –15 V
Input voltage range, V_I: Driver . . . . . –0.3 V to V_{CC} + 0.3 V
Receiver . . . . . ±30 V
Output voltage range, V_O: T1OUT, T2OUT . . . . . V_{S-} – 0.3 V to V_{S+} + 0.3 V
R1OUT, R2OUT . . . . . –0.3 V to V_{CC} + 0.3 V
Short-circuit duration: T1OUT, T2OUT . . . . . Unlimited
Package thermal impedance, θ_{JA} (see Note 2): D package . . . . . 73°C/W
DW package . . . . . 57°C/W
N package . . . . . 67°C/W
NS package . . . . . 64°C/W
Lead temperature 1,6 mm (1/16 inch) from case for 10 seconds . . . . . 260°C
Storage temperature range, T_{stg} . . . . . –65°C to 150°C

† Stresses beyond those listed under "absolute maximum ratings" may cause permanent damage to the device. These are stress ratings only, and functional operation of the device at these or any other conditions beyond those indicated under "recommended operating conditions" is not implied. Exposure to absolute-maximum-rated conditions for extended periods may affect device reliability.

NOTE 1: All voltage values are with respect to network ground terminal.
2. The package thermal impedance is calculated in accordance with JESD 51-7.

recommended operating conditions

			MIN	NOM	MAX	UNIT
V_{CC}	Supply voltage		4.5	5	5.5	V
V_{IH}	High-level input voltage (T1IN,T2IN)		2			V
V_{IL}	Low-level input voltage (T1IN, T2IN)				0.8	V
R1IN, R2IN	Receiver input voltage				±30	V
T_A	Operating free-air temperature	MAX232	0		70	°C
		MAX232I	–40		85	

electrical characteristics over recommended ranges of supply voltage and operating free-air temperature (unless otherwise noted) (see Note 3 and Figure 4)

PARAMETER		TEST CONDITIONS		MIN	TYP‡	MAX	UNIT
I_{CC}	Supply current	V_{CC} = 5.5 V, T_A = 25°C	All outputs open,		8	10	mA

‡ All typical values are at V_{CC} = 5 V and T_A = 25°C.
NOTE 3: Test conditions are C1–C4 = 1 µF at V_{CC} = 5 V ± 0.5 V.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

DRIVER SECTION

electrical characteristics over recommended ranges of supply voltage and operating free-air temperature range (see Note 3)

PARAMETER			TEST CONDITIONS	MIN	TYP†	MAX	UNIT
V_{OH}	High-level output voltage	T1OUT, T2OUT	R_L = 3 kΩ to GND	5	7		V
V_{OL}	Low-level output voltage‡	T1OUT, T2OUT	R_L = 3 kΩ to GND		−7	−5	V
r_o	Output resistance	T1OUT, T2OUT	V_{S+} = V_{S-} = 0, V_O = ±2 V	300			Ω
I_{OS}§	Short-circuit output current	T1OUT, T2OUT	V_{CC} = 5.5 V, V_O = 0		±10		mA
I_{IS}	Short-circuit input current	T1IN, T2IN	V_I = 0			200	µA

† All typical values are at V_{CC} = 5 V, T_A = 25°C.
‡ The algebraic convention, in which the least positive (most negative) value is designated minimum, is used in this data sheet for logic voltage levels only.
§ Not more than one output should be shorted at a time.
NOTE 3: Test conditions are C1–C4 = 1 µF at V_{CC} = 5 V ± 0.5 V.

switching characteristics, V_{CC} = 5 V, T_A = 25°C (see Note 3)

PARAMETER		TEST CONDITIONS	MIN	TYP	MAX	UNIT
SR	Driver slew rate	R_L = 3 kΩ to 7 kΩ, See Figure 2			30	V/µs
SR(t)	Driver transition region slew rate	See Figure 3		3		V/µs
	Data rate	One TOUT switching		120		kbit/s

NOTE 3: Test conditions are C1–C4 = 1 µF at V_{CC} = 5 V ± 0.5 V.

RECEIVER SECTION

electrical characteristics over recommended ranges of supply voltage and operating free-air temperature range (see Note 3)

PARAMETER			TEST CONDITIONS	MIN	TYP†	MAX	UNIT
V_{OH}	High-level output voltage	R1OUT, R2OUT	I_{OH} = −1 mA	3.5			V
V_{OL}	Low-level output voltage‡	R1OUT, R2OUT	I_{OL} = 3.2 mA			0.4	V
V_{IT+}	Receiver positive-going input threshold voltage	R1IN, R2IN	V_{CC} = 5 V, T_A = 25°C		1.7	2.4	V
V_{IT-}	Receiver negative-going input threshold voltage	R1IN, R2IN	V_{CC} = 5 V, T_A = 25°C	0.8	1.2		V
V_{hys}	Input hysteresis voltage	R1IN, R2IN	V_{CC} = 5 V	0.2	0.5	1	V
r_i	Receiver input resistance	R1IN, R2IN	V_{CC} = 5, T_A = 25°C	3	5	7	kΩ

† All typical values are at V_{CC} = 5 V, T_A = 25°C.
‡ The algebraic convention, in which the least positive (most negative) value is designated minimum, is used in this data sheet for logic voltage levels only.
NOTE 3: Test conditions are C1–C4 = 1 µF at V_{CC} = 5 V ± 0.5 V.

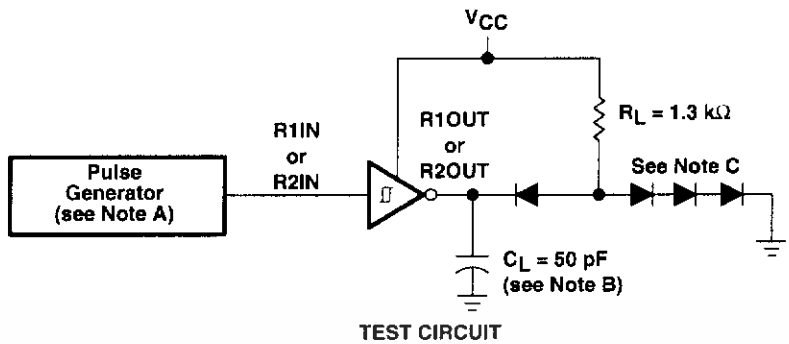
switching characteristics, V_{CC} = 5 V, T_A = 25°C (see Note 3 and Figure 1)

PARAMETER		TYP	UNIT
$t_{PLH(R)}$	Receiver propagation delay time, low- to high-level output	500	ns
$t_{PHL(R)}$	Receiver propagation delay time, high- to low-level output	500	ns

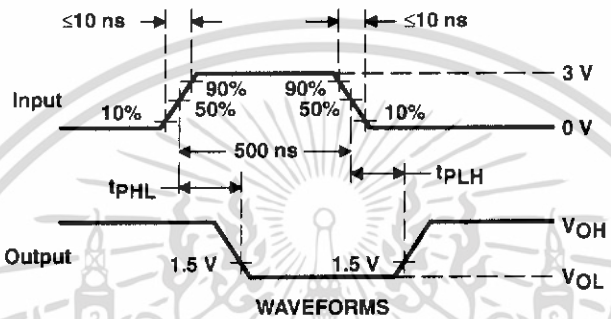
NOTE 3: Test conditions are C1–C4 = 1 µF at V_{CC} = 5 V ± 0.5 V.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

PARAMETER MEASUREMENT INFORMATION



TEST CIRCUIT



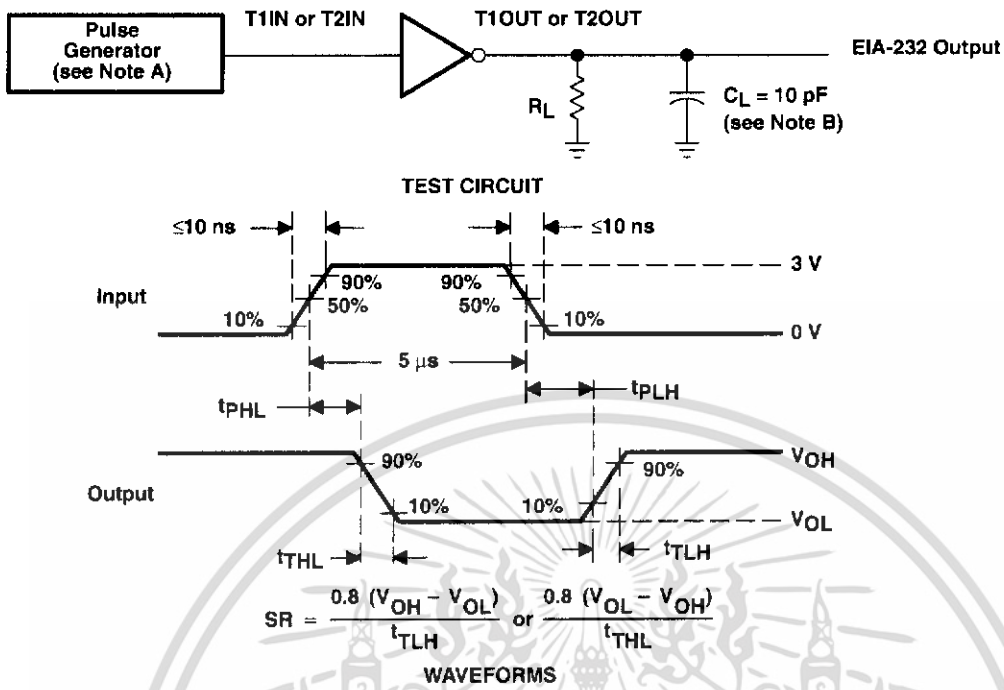
WAVEFORMS

NOTES: A. The pulse generator has the following characteristics: Z_O = 50 Ω, duty cycle ≤ 50%.
B. C_L includes probe and jig capacitance.
C. All diodes are 1N3064 or equivalent.

Figure 1. Receiver Test Circuit and Waveforms for t_{PHL} and t_{PLH} Measurements

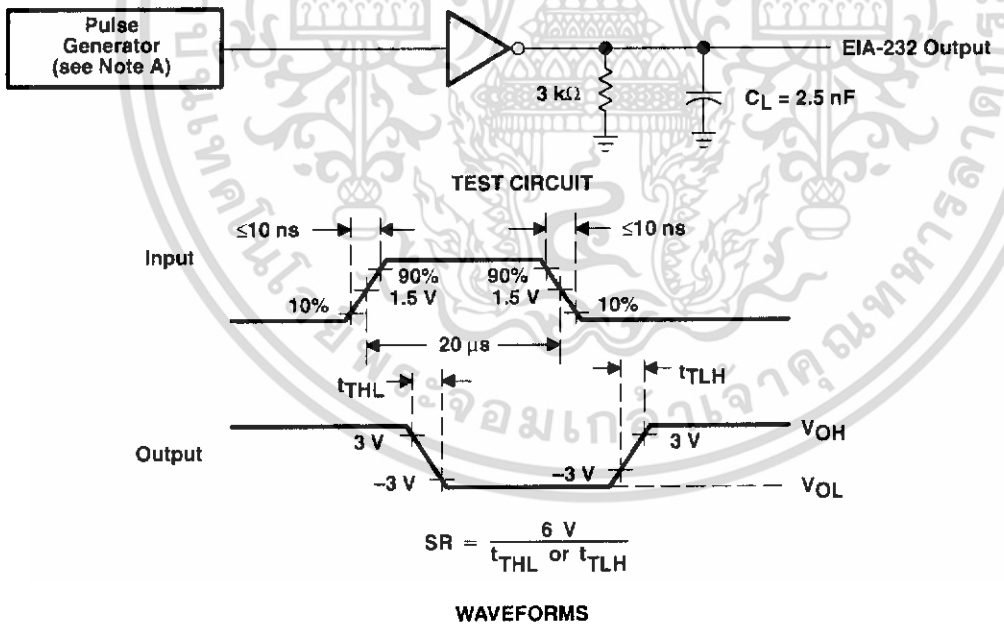
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

PARAMETER MEASUREMENT INFORMATION



NOTES: A. The pulse generator has the following characteristics: Z_O = 50 Ω, duty cycle ≤ 50%.
B. C_L includes probe and jig capacitance.

Figure 2. Driver Test Circuit and Waveforms for t_{PHL} and t_{PLH} Measurements (5-µs Input)

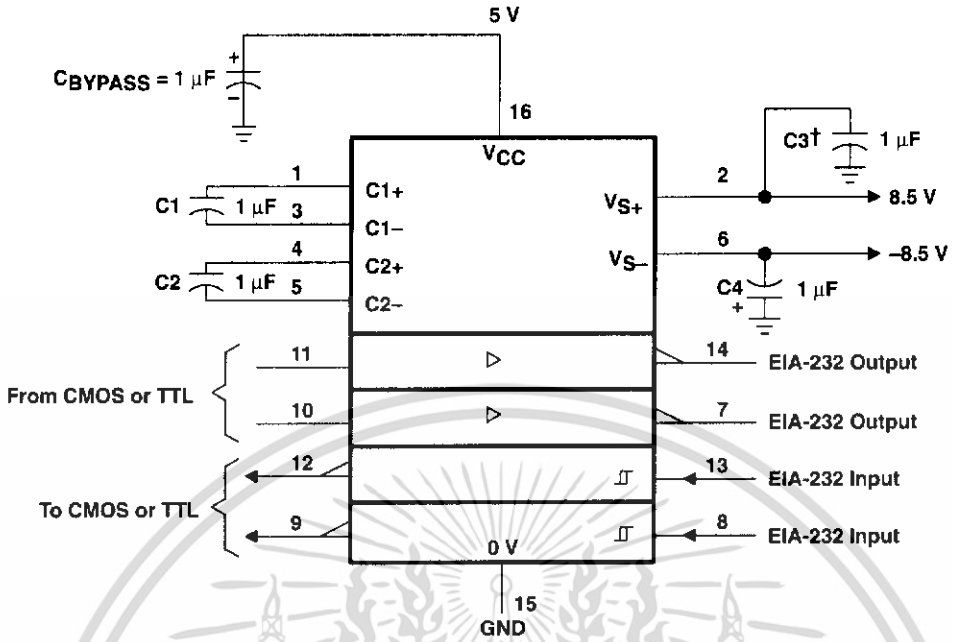


NOTE A: The pulse generator has the following characteristics: Z_O = 50 Ω, duty cycle ≤ 50%.

Figure 3. Test Circuit and Waveforms for t_{THL} and t_{TLH} Measurements (20-µs Input)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

APPLICATION INFORMATION



† C3 can be connected to V_{CC} or GND.

Figure 4. Typical Operating Circuit

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

IMPORTANT NOTICE

Texas Instruments Incorporated and its subsidiaries (TI) reserve the right to make corrections, modifications, enhancements, improvements, and other changes to its products and services at any time and to discontinue any product or service without notice. Customers should obtain the latest relevant information before placing orders and should verify that such information is current and complete. All products are sold subject to TI's terms and conditions of sale supplied at the time of order acknowledgment.

TI warrants performance of its hardware products to the specifications applicable at the time of sale in accordance with TI's standard warranty. Testing and other quality control techniques are used to the extent TI deems necessary to support this warranty. Except where mandated by government requirements, testing of all parameters of each product is not necessarily performed.

TI assumes no liability for applications assistance or customer product design. Customers are responsible for their products and applications using TI components. To minimize the risks associated with customer products and applications, customers should provide adequate design and operating safeguards.

TI does not warrant or represent that any license, either express or implied, is granted under any TI patent right, copyright, mask work right, or other TI intellectual property right relating to any combination, machine, or process in which TI products or services are used. Information published by TI regarding third–party products or services does not constitute a license from TI to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property of the third party, or a license from TI under the patents or other intellectual property of TI.

Reproduction of information in TI data books or data sheets is permissible only if reproduction is without alteration and is accompanied by all associated warranties, conditions, limitations, and notices. Reproduction of this information with alteration is an unfair and deceptive business practice. TI is not responsible or liable for such altered documentation.

Resale of TI products or services with statements different from or beyond the parameters stated by TI for that product or service voids all express and any implied warranties for the associated TI product or service and is an unfair and deceptive business practice. TI is not responsible or liable for any such statements.

Mailing Address:

Texas Instruments
Post Office Box 655303
Dallas, Texas 75265

Copyright © 2002, Texas Instruments Incorporated

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

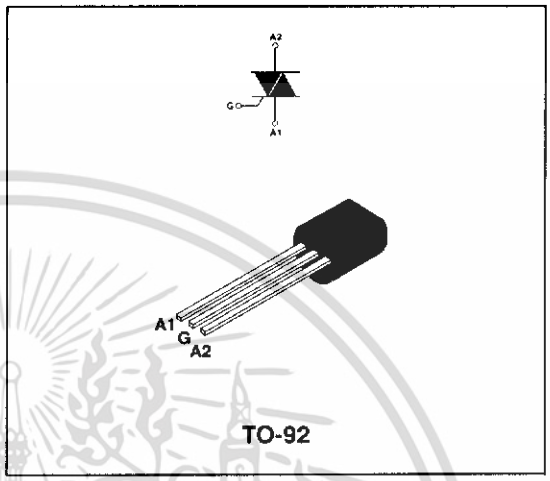
Z00607MA

STANDARD

0.8A TRIAC

Table 1: Main Features

Symbol	Value	Unit
$I_{T(RMS)}$	0.8	A
V_{DRM}/V_{RRM}	600	V
$I_{GT\ (Q_1)}$	5	mA



TO-92

DESCRIPTION

The **Z00607MA** is suitable for low power AC switching applications, such as fan speed, small light controllers...
Thanks to low gate triggering current, it can be directly driven by microcontrollers.

Table 2: Order Codes

Part Numbers	Marking
Z00607MA 1BA2	Z0607MA
Z00607MA 2BL2	Z0607MA
Z00607MA 5BL2	Z0607MA

Table 3: Absolute Maximum Ratings

Symbol	Parameter			Value	Unit
$I_{T(RMS)}$	RMS on-state current (full sine wave)		T_j = 50°C	0.8	A
I_{TSM}	Non repetitive surge peak on-state current (full cycle, T_j initial = 25°C)	F = 50 Hz	t = 20 ms	9	A
		F = 60 Hz	t = 16.7 ms	9.5	
I^2t	I^2t Value for fusing	t_p = 10 ms		0.45	A^2s
dI/dt	Critical rate of rise of on-state current $I_G = 2 \times I_{GT}$, $t_r \leq 100$ ns	F = 120 Hz	T_j = 110°C	20	A/µs
I_{GM}	Peak gate current	t_p = 20 µs	T_j = 110°C	1	A
$P_{G(AV)}$	Average gate power dissipation		T_j = 110°C	0.1	W
T_{stg} T_j	Storage junction temperature range Operating junction temperature range			- 40 to + 150 - 40 to + 110	°C

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Tables 4: Electrical Characteristics (T_j = 25°C, unless otherwise specified)

Symbol	Test Conditions	Quadrant		Value	Unit
I_{GT} (1)	V_D = 12 V R_L = 30 Ω	I - II - III	MAX.	5	mA
		IV		7	
V_{GT}		ALL	MAX.	1.3	V
V_{GD}	$V_D = V_{DRM}$ R_L = 3.3 kΩ T_j = 110°C	ALL	MIN.	0.2	V
I_H (2)	I_T = 200 mA		MAX.	5	mA
I_L	I_G = 1.2 I_{GT}	I - III - IV	MAX.	10	mA
		II		20	
dV/dt (2)	V_D = 67 %V_{DRM} gate open T_j = 110°C		MIN.	10	V/µs
(dI/dt)c (2)	(dV/dt)c = 0.35 A/ms T_j = 110°C		MIN.	1.5	A/ms

Table 5: Static Characteristics

Symbol	Test Conditions			Value	Unit
V_{TM} (2)	I_{TM} = 1.1 A t_p = 380 µs	T_j = 25°C	MAX.	1.5	V
V_{to} (2)	Threshold voltage	T_j = 110°C	MAX.	0.95	V
R_d (2)	Dynamic resistance	T_j = 110°C	MAX.	420	mΩ
I_{DRM}	$V_{DRM} = V_{RRM}$ = 600 V	T_j = 25°C	MAX.	5	µA
I_{RRM}		T_j = 110°C		0.1	mA

Note 1: minimum I_{GT} is guaranted at 5% of I_{GT} max.

Note 2: for both polarities of A2 referenced to A1.

Table 6: Thermal resistances

Symbol	Parameter	Value	Unit
$R_{th(j-l)}$	Junction to lead (A.C.)	60	°C/W
$R_{th(j-a)}$	Junction to ambient	150	°C/W

Figure 1: Maximum power dissipation versus RMS on-state current (full cycle)

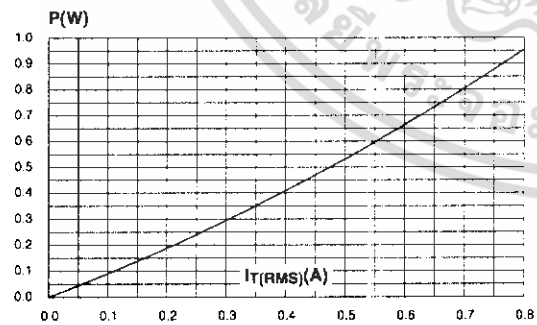
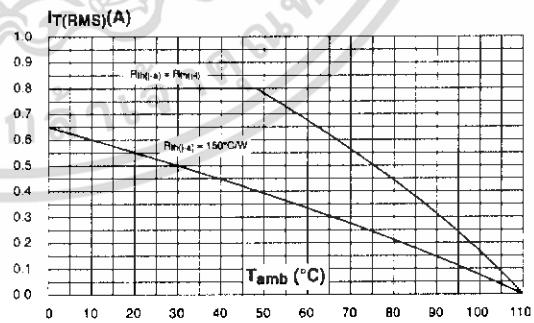


Figure 2: RMS on-state current versus ambient temperature (full cycle)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Figure 3: Relative variation of thermal impedance versus pulse duration

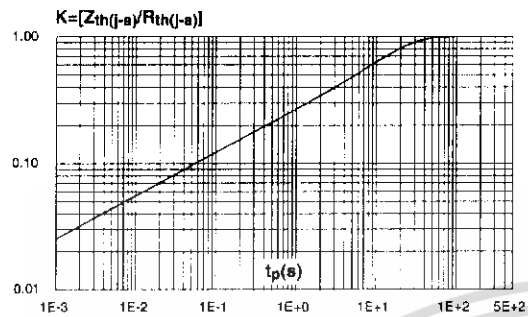


Figure 4: Relative variation of gate trigger current, holding current and latching current versus junction temperature (typical values)

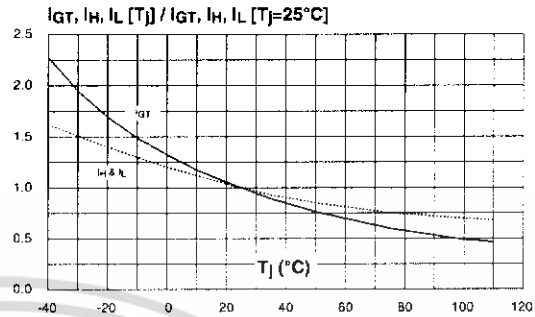


Figure 5: Surge peak on-state current versus number of cycles

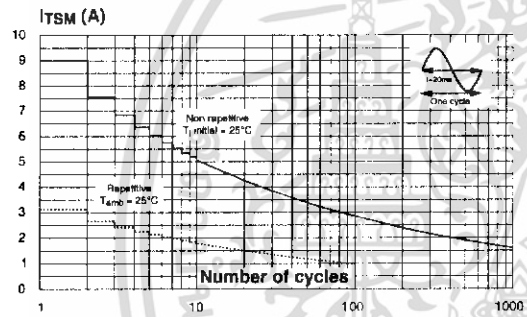


Figure 6: Non-repetitive surge peak on-state current for a sinusoidal pulse with width $t_p < 10$ ms and corresponding value of I^2t

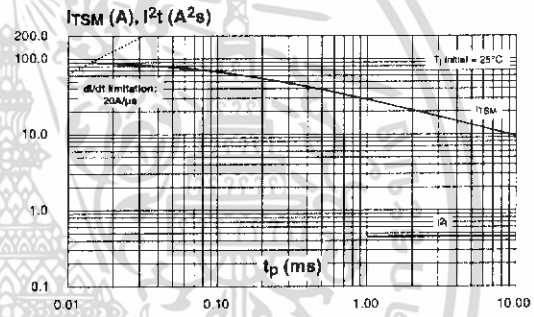


Figure 7: On-state characteristics (maximum values)

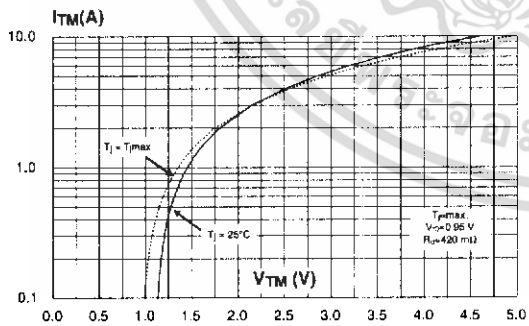
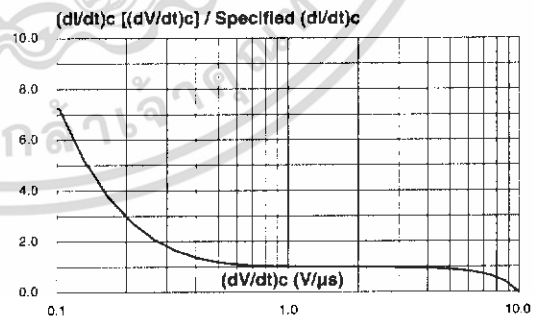


Figure 8: Relative variation of critical rate of decrease of main current versus (dV/dt)c (typical values)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Figure 9: Relative variation of critical rate of decrease of main current versus junction temperature

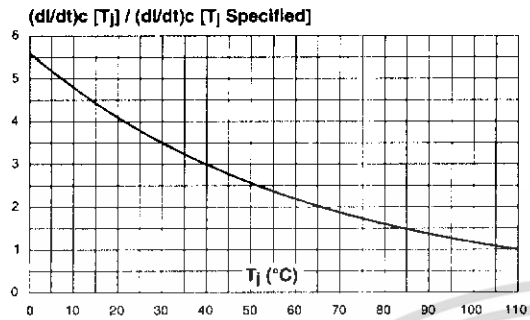


Figure 10: Ordering Information Scheme

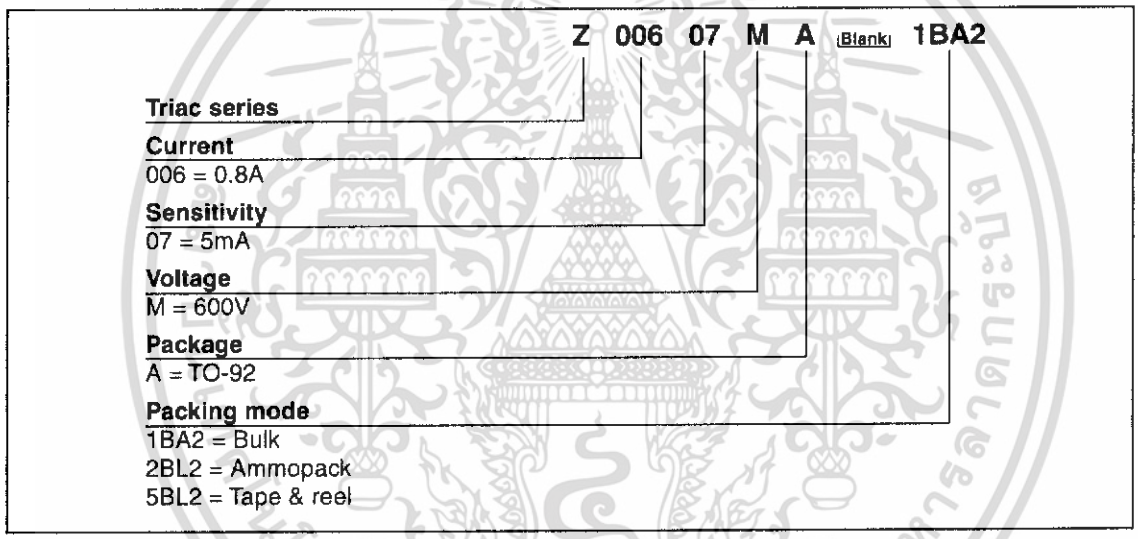
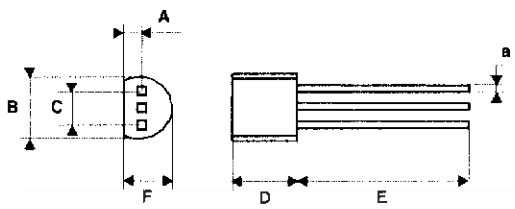


Table 7: Product Selector

Part Number	Voltage	Sensitivity	Type	Package
Z00607MA	600 V	5 mA	Standard	TO-92

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Figure 11: TO-92 Package Mechanical Data



REF.	DIMENSIONS					
	Millimeters			Inches		
	Min.	Typ.	Max.	Min.	Typ.	Max.
A		1.35			0.053	
B			4.70			0.185
C		2.54			0.100	
D	4.40			0.173		
E	12.70			0.500		
F			3.70			0.146
a			0.50			0.019

Table 8: Ordering Information

Ordering type	Marking	Package	Weight	Base qty	Delivery mode
Z00607MA 1BA2	Z0607MA	TO-92	0.2 g	2500	Bulk
Z00607MA 2BL2	Z0607MA			2000	Ammopack
Z00607MA 5BL2	Z0607MA			2000	Tape & reel

Table 9: Revision History

Date	Revision	Description of Changes
Oct-2001	4	Last update.
25-Mar-2005	5	Package: TO-92 tape & reel delivery mode 5BL2 added.
21-Jun-2005	6	Markings updated from Z006xxxx to Z06xxxx
13-Sep-2005	7	Z00607MA 2BL2: marking corrected from 00607mA to Z0607MA

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้