สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การรู้จำเสียง

RECOGNITION

โดย

นาย บุญคิด นิลประพัฒน์ 46010385
นาย พัลลภ กลิ่นขจร 46010513
นาย ภาคภูมิ ไพโรจน์ 46010567

ปพ.
ป4137
2549

เลขหมู่.....
เลขทะเบียน 72868
วัน,เดือน,ปี 25 มิ.ย. 2550

b 11223216
i.....

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
ภาควิชาวิศวกรรมระบบควบคุม สาขาวิชาวิศวกรรมระบบควบคุม
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2549

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์ ปีการศึกษา 2549

สาขาวิชาวิศวกรรมศาสตร์ระบบควบคุม

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การรู้จำเสียง

Recognition

ผู้จัดทำ

1. นาย บุญคิด นิลประพัฒน์ 46010385
2. นาย พัลลภ กลิ่นขจร 46010513
3. นาย ภาคภูมิ ไพโรจน์ 46010567

..... อาจารย์ที่ปรึกษา

(อาจารย์ วรรณดี เพชรมณีล้ำค่า)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การรู้จำเสียง

นาย บุญคิด นิลประพัฒน์ 46010385
นาย พัลลภ กลิ่นขจร 46010513
นาย ภาคภูมิ ไพโรจน์ 46010567
อาจารย์ที่ปรึกษา อ.วรรณดี เพชรมณีล้ำค่า
ปีการศึกษา 2549

บทคัดย่อ

วิทยานิพนธ์นี้ได้นำเสนอการรู้จำเสียงโดยใช้ dsPIC30F6014 เป็นตัวประมวลผล ซึ่งผู้จัดทำโครงงานเห็นว่ามีคุณสมบัติเพียงพอสำหรับการประมวลผลการรู้จำเสียง ใช้ภาษาซีในการเขียนโปรแกรม มีเสียงคำสั่งทั้งหมด 6 เสียง คือ Control,Play,Pause,Previous,Next และ Stop โดยในการสร้างเสียงต้นแบบนั้นเราจะใช้โปรแกรม Word Library Builder ในส่วนของการทำงานของระบบรู้จำเสียงนั้น จะเริ่มจากรับสัญญาณเสียงผ่านทางไมโครโฟนแล้วนำสัญญาณเสียงไปผ่านตัว Si3000 เพื่อแปลงสัญญาณเสียงเป็นสัญญาณดิจิตอล หลังจากนั้นเราจากนั้นเราจะนำข้อมูลสัญญาณดิจิตอลไปผ่านกระบวนการ MFCC เพื่อหาตัวแทนของเสียงเนื่องจากเราไม่ได้นำสัญญาณเสียงทั้งหมดมาประมวลผล หลังจากนั้นนำตัวแทนของเสียงไปเปรียบเทียบกับเสียงต้นแบบโดยใช้กระบวนการ HMM แล้วตัดสินใจว่าเสียงทดสอบใกล้เคียงกับเสียงต้นแบบใดมากที่สุดและแสดงผลบนจอ LCD ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

RECOGNITION

Mr. Bunkid Nilprapat 46010385

Mr. Pallop Klinkajorn 46010513

Mr. Pakpoom Pairoje 46010567

Miss. Wandee Petchmaneelumka Advisor

Academic Year 2006

Abstract

In this thesis, speech recognition by using dsPIC30F6014 as a processor is proposes. In our opinion this processor have enough ability to processing speech recognition. Using C language for programming. There are totally 6 control word which are "Control" , "Play" , "Pause" , "Previous" , "Next" and "Stop". In the process, MFCC method has been used feature attraction and HMM method for comparing tested speech with reference speech. Referenced speech will be created by a program called Word Library Builder. As for a speech recognition process, a tested speech signal will be received via microphone and will be converted into digital signal by Si3000. After that we will use MFCC method to extract a sound agent from digital signal . Once the signal has been extracted we will compare it with reference speech by using HMM method. The processor will decide that which referenced speech are the most matches with tested speech and then displayed it on LCD screen.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงได้อย่างดี ด้วยคำแนะนำ คำปรึกษาและคอยดูแลจากหลาย ๆ ฝ่ายด้วยกัน โดยเฉพาะอาจารย์ที่ปรึกษาที่คอยให้ความเอาใจใส่ แนะนำและช่วยเหลือเสมอมา ซึ่งก็คือ อาจารย์ธวัชชัย คำศรี และ อาจารย์วรรณดี เพชรมณีล้ำค่า

นอกเหนือจากนี้ต้องขอขอบพระคุณอาจารย์ทุกท่านในภาควิชาวิศวกรรมระบบควบคุม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังเป็นอย่างยิ่งที่ได้ช่วยประสิทธิ์ประสาทวิชาความรู้ให้แก่คณะผู้จัดทำ

ขอขอบคุณภาควิชาวิศวกรรมระบบควบคุมที่ได้จัดเตรียมสิ่งอำนวยความสะดวก เช่น อินเตอร์เน็ตความเร็วสูง บอร์ดทดลอง dsPicdem เงินทุน ซึ่งช่วยให้การวิจัย การค้นคว้าหาความรู้ต่าง ๆ และพัฒนาโปรแกรมเป็นไปได้ด้วยความสะดวกและรวดเร็ว

สุดท้ายต้องขอขอบคุณบุคคลที่สำคัญที่สุดในชีวิตที่ทำให้ข้าพเจ้ามีวันนี้ นั่นคือ บิดา มารดาและบุคคลในครอบครัว อันเป็นที่เคารพรัก ซึ่งได้เลี้ยงดู คอยสั่งสอนข้าพเจ้ามาเป็นอย่างดี พร้อมให้โอกาสในการศึกษาอย่างเต็มที่ และยังให้กำลังใจ ความรักเสมอมา ข้าพเจ้าขอกราบขอบพระคุณมา ณ ที่นี้ด้วย

นาย บุญคิต นิลประพัฒน์
นาย พัลลภ กลิ่นขจร
นาย ภาคภูมิ ไพโรจน์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อ	I
Abstract	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญรูป	VII
สารบัญตาราง	IX
บทที่ 1 บทนำ	1
1.1 ความสำคัญและที่มาของปัญหา	1
1.2 วัตถุประสงค์ของโครงงาน	2
1.3 ขอบเขตของโครงงาน	2
1.4 ขั้นตอนของโครงงาน	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ	3
1.6 รายละเอียดของปริญญานิพนธ์	3
บทที่ 2 ทฤษฎี	4
2.1 ทฤษฎีการสร้างเสียงพูด	4
2.1.1 อวัยวะที่ใช้ในการออกเสียงพูด	4
2.1.2 การเกิดเสียง	5
2.2 ทฤษฎีการรู้จำเสียง	6
2.2.1 การหาค่าพารามิเตอร์	6
2.2.2 การแบ่งช่วงสัญญาณ	7
2.2.3 การวินโดว์	8
2.2.4 การหาพลังงานของเสียง	9
2.2.5 การแปลงฟูเรียร์	9
2.2.6 การประมวลผลของเมลฟรีเควนซี่เซปตรัมโคเอฟฟิเชี่ยน	10
2.2.7 รูปแบบของมาร์คอฟ	12
2.2.7.1 นิยามและหลักการพื้นฐาน	12
2.2.7.2 สมมุติฐานสำหรับทฤษฎีรูปแบบของมาร์คอฟ	15
2.2.7.3 ปัญหาพื้นฐานของรูปแบบมาร์คอฟ	15

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.2.7.4 ปัญหาในการหาค่าความน่าจะเป็น	16
2.2.7.5 ปัญหาการถอดรหัส	19
2.2.7.6 ปัญหาการเรียนรู้	20
2.3 บอร์ด dsPICdem	22
2.4 dsPIC30F6014	24
2.4.1 สถาปัตยกรรมโดยสรุปของ dsPIC30F6014	27
2.4.2 สแต็กพอยน์เตอร์และเฟรมพอยน์เตอร์ทางซอฟต์แวร์	32
2.4.3 หน่วยประมวลผลสัญญาณดิจิตอล	34
2.4.4 การรองรับการหารเลขใน dsPIC	38
บทที่ 3 การคำนวณและการสร้าง	39
3.1 การสร้างสัญญาณเสียงอ้างอิง	39
3.2 การแปลงสัญญาณเสียงเป็นสัญญาณดิจิตอล	44
3.3 dsPIC30F6014 รับข้อมูล	45
3.4 การหาตัวแทนเสียง	46
3.5 การหาผลลัพธ์ของเสียง	46
3.5.1 ฮิดเดนมาร์คอฟโมเดล	46
3.5.2 การตัดสินใจ	47
3.6 การแสดงผล	48
3.7 การทดลองหาค่าเปอร์เซ็นต์ความถูกต้อง	49
3.8 การนำเสียงไปประยุกต์ใช้	50
บทที่ 4 การทดลองและผลการทดลอง	52
4.1 การทดลอง	53
4.2 ผลการทดลอง	54
4.3 สรุปผลการทดลอง	70
บทที่ 5 บทวิจารณ์และบทสรุป	71
5.1 สรุป	71
5.2 ปัญหาที่พบในการวิจัย	71
5.3 ข้อเสนอแนะ	72

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
ภาคผนวก ก การทดลองการใช้งานบอร์ด	73
ภาคผนวก ข การใช้งานโปรแกรม	78
ภาคผนวก ค โปรแกรม	82
เอกสารอ้างอิง	90

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ

รูปที่	หน้า
2.1 แสดงการแบ่งช่วงสัญญาณ	7
2.2 ตัวอย่างของฟังก์ชันหน้าต่างแบบแฮมมิ่ง	8
2.3 ตัวกรองความถี่สำหรับการสร้าง MFCC ที่มีความถี่ระหว่าง 300 ถึง 3400 Hz	11
2.4 ไดอะแกรมแสดงกระบวนการหา MFCC	11
2.5 แสดงส่วนประกอบต่าง ๆ ของ HMM	14
2.6 กระบวนการฝึกหัดกระบวนการแบ่งเสียงของเค	21
2.7 บอร์ดที่ใช้ในการทดลอง dsPICdem	22
2.8 อุปกรณ์ต่าง ๆ บนบอร์ด dsPICdem 1.1	23
2.9 ไดอะแกรมการทำงานและส่วนประทั้งหมดของ dsPIC30F6014	24
2.10 การจัดขาใช้งานของไมโครคอนโทรลเลอร์ dsPIC30F6014	25
2.11 การจัดสรรหน่วยความจำโปรแกรมของไมโครคอนโทรลเลอร์ dsPIC	29
2.12 โครงสร้างทางโปรแกรมของไมโครคอนโทรลเลอร์	30
2.13 โครงสร้างทางโปรแกรมของไมโครคอนโทรลเลอร์	32
2.14 แสดงตัวอย่างการดีความค่าตัวเลขข้อมูล ในกรณีที่บิต MSB เป็น “0” ทั้งในกรณีที่คิดเลขแบบจำนวนเต็มและเลขเศษส่วนทศนิยมในหน่วย DSP และแบบ ALU ของ dsPIC	35
2.15 แสดงตัวอย่างการดีความค่าตัวเลขข้อมูล ในกรณีที่บิต MSB เป็น “1” ทั้งในกรณีที่คิดเลขแบบจำนวนเต็มและเลขเศษส่วนทศนิยมในหน่วย DSP และแบบ ALU ของ dsPIC	36
3.1 แผนผังการทำงานโดยรวมของระบบรู้จำเสียง	39
3.2 โปรแกรม Word Library Builder	40
3.3 การเลือกสัญญาณเสียงอ้างอิง	41
3.4 การเรียนรู้การรู้จำเสียง	42
3.5 การสร้างข้อมูลของเสียงต้นแบบ	43
3.6 ไอซี SI 3000 รับสัญญาณเสียงจากการพูดทางไมโครโฟน	44
3.7 ตัวอย่างสัญญาณเสียง play ที่พูดผ่านไมโครโฟน	44
3.8 ตัวอย่างข้อมูลที่ผ่านการแปลงสัญญาณด้วย Si3000	45
3.9 บล็อคแสดงกระบวนการรับข้อมูลสัญญาณดิจิตอล	46

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ (ต่อ)

รูปที่	หน้า
3.10 บล็อคแสดงการหาตัวแทนของเสียง	46
3.11 จำลองกระบวนการ HMM	47
3.12 ตัวอย่างการแสดงผลของเสียง PLAY	48
3.13 ผังงานแสดงการทำงานโดยรวมตั้งแต่หัวข้อ 3.1 – 3.6	49
3.14 ตัวอย่างโปรแกรม MEDIA PLAYER	51
4.1 การติดตั้งอุปกรณ์	53
4.2 ข้อความที่ปรากฏบนจอ LCD	53
4.3 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 1	54
4.4 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 2	55
4.5 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 3	56
4.6 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 4	57
4.7 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 5	58
4.8 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 6	59
4.9 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 7	60
4.10 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 1	62
4.11 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 2	63
4.12 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 3	64
4.13 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 4	65
4.14 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 5	66
4.15 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 6	67
4.16 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 7	68
4.17 กราฟแสดงผลเฉลี่ยเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชาย	69
4.18 กราฟแสดงผลเฉลี่ยเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิง	69

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 แสดงรายชื่ออุปกรณ์ต่าง ๆ บนบอร์ด dsPICDEM 1.1	23
2.2 แสดงขอบเขตของข้อมูลในรีจิสเตอร์ที่มีขนาดต่างกันใน ไมโครคอนทรลเลอร์ dsPIC ทั้งในแบบเลขจำนวนเต็มและเศษส่วน/ทศนิยม	34
3.1 คำในไลบารี่และค่าของคำ	48
4.1 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 1	54
4.2 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 1	54
4.3 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 2	55
4.4 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 2	55
4.5 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 3	56
4.6 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 3	56
4.7 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 4	57
4.8 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 4	57
4.9 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 5	58
4.10 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 5	58
4.11 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 6	59
4.12ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 6	59
4.13 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 7	60
4.14 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 7	60
4.15 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 1	62
4.16 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 1	62
4.17 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 2	63
4.18 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 2	63
4.19 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 3	64
4.20 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 3	64
4.21 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 4	65
4.22 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 4	65
4.23 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 5	66
4.24 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 5	66

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.25 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 6	67
4.26 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 6	67
4.27 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 7	68
4.28 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 7	68
4.29 ผลการเฉลี่ยเปอร์เซ็นต์ความถูกต้องรู้จำเสียงทั้งผู้ชายและผู้หญิง	69

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความสำคัญและที่มาของปัญหา

ในปัจจุบันนี้เทคโนโลยีต่างๆ ได้เข้ามามีบทบาทในชีวิตประจำวันของมนุษย์มากขึ้น จึงทำให้เกิดการศึกษา ค้นคว้า วิจัยเพื่อค้นหาเทคโนโลยีใหม่ๆ มาใช้เพื่ออำนวยความสะดวก สบายต่อมนุษย์ โดยเฉพาะเทคโนโลยีที่เกี่ยวกับการสื่อสารในปัจจุบัน ซึ่งมนุษย์ได้เล็งเห็นว่าการสื่อสารนั้นเป็นสิ่งที่สำคัญและจำเป็นต่อมนุษย์เป็นอย่างมาก เนื่องจากมนุษย์เป็นสัตว์สังคมที่ต้องมีการติดต่อสื่อสารกันอยู่ตลอดเวลา โดยเฉพาะการพูดเป็นการสื่อสารที่ทำได้ง่าย ทำให้เทคโนโลยีการสื่อสารได้ถูกพัฒนาไปอย่างกว้างขวาง

การเข้าใจของสิ่งที่มนุษย์ได้ยินนั้น เป็นกระบวนการที่ซับซ้อน เกินที่จะจำลองเป็นรูปแบบหนึ่งๆ ได้ ถึงแม้ว่าเราจะรู้สึกว่า เราเข้าใจสิ่งที่เราได้ยินได้ง่าย แต่ในคอมพิวเตอร์จำเป็นต้องอาศัยการคำนวณเป็นจำนวนมหาศาลเพื่อวิเคราะห์คำพูดเพียงหนึ่งพยางค์ โดยอาศัยทฤษฎี ความรู้ด้านต่างๆ มาประยุกต์ใช้ในการสร้างระบบที่สามารถแยกแยะสิ่งที่เราพูดให้ได้

เทคโนโลยีทางด้านการรู้จำ เป็นเทคโนโลยีการสื่อสาร ซึ่งได้มีการพัฒนาและ ทำวิจัยต่างๆ กันมาเป็นระยะเวลานานแล้ว เทคโนโลยีนี้ทำให้การป้อนข้อมูลนั้นสามารถใช้เสียงเป็นอินพุตได้แทนที่จะใช้การคีย์ข้อมูลผ่านคีย์บอร์ดเท่านั้น ซึ่งเป็นสิ่งที่ทุกคนคาดหวังว่าในอนาคตจำเป็นต้องมีเทคโนโลยีเช่นนี้เกิดขึ้น อีกทั้งยังสามารถอำนวยความสะดวกแก่ผู้พิการให้สามารถติดต่อสื่อสารได้อย่างสะดวกยิ่งขึ้น โดยหลักการของการรู้จำระบบเป็นการวิเคราะห์ข้อมูลเสียงที่พูดออกไป เพื่อให้สามารถแปลงเป็นคำพูดให้กลายเป็นตัวอักษรที่ปรากฏบนหน้าจอคอมพิวเตอร์ ซึ่งการรู้จำระบบที่มีการวิจัยกันมานั้นมีทั้งรูปแบบที่เป็นทั้งภาษาอังกฤษและภาษาไทย หรือภาษาอื่นๆ

โดยในโครงงานนี้จะพัฒนาระบบการรู้จำเสียงพูดภาษาอังกฤษ ด้วยการใช้บอร์ด dsPicdem ซึ่งมี dsPIC30F6014 เป็นตัวประมวลผล เนื่องจากผู้จัดทำมีความเห็นว่า ในอดีตที่ผ่านมาการพัฒนาระบบรู้จำเสียงจะใช้คอมพิวเตอร์พัฒนาเป็นส่วนใหญ่ จึงทดลองนำบอร์ด ds30F6014 มาทำการรู้จำเสียงว่าจะมีประสิทธิภาพมากน้อยเพียงใด สามารถนำว่าประยุกต์ใช้ในชีวิตประจำวันได้หรือไม่ เช่นใช้เสียงในการกดเบอร์โทรศัพท์, การสั่งงานในคอมพิวเตอร์หรืออุปกรณ์อิเล็กทรอนิคส์ด้วยเสียง อีกทั้งยังสามารถนำไปเป็นแนวทางในการวิเคราะห์คำพูดแบบต่อเนื่อง หรือใช้ในการวิเคราะห์ภาษาอื่นได้อีกด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 วัตถุประสงค์ของโครงงาน

1. เพื่อศึกษาคุณลักษณะของเสียงว่ามีคุณสมบัติใดบ้าง
2. เพื่อศึกษาการทำงานของระบบรู้จำเสียง
3. เพื่อหาประสิทธิภาพการรู้จำเสียง
4. เพื่อนำการรู้เสียงพูดแบบไม่ต่อเนื่องไปใช้กับโปรแกรมประยุกต์ต่างๆ

1.3 ขอบเขตของโครงงาน

โครงงานนี้เน้นไปในเรื่องการแยกแยะคำออกจากคำพูดที่มนุษย์พูดแบบไม่ต่อเนื่อง เพื่อนำไปประยุกต์ใช้กับโปรแกรมประยุกต์บางอย่างโดยเฉพาะ โดยโครงงานนี้จะมีขอบเขตการทำงานดังต่อไปนี้

1. อินพุตของโปรแกรมจะเป็นเสียงที่ได้มาจากการพูดเข้าไมค์
2. ผลลัพธ์ของโปรแกรมจะเป็นตัวอักษรที่เป็นคำพูดเฉพาะที่มีอยู่ในฐานข้อมูลเท่านั้น
3. ผู้พูดจะต้องออกเสียงที่มีลักษณะใกล้เคียงกับคำเฉพาะที่มีในฐานข้อมูล
4. ระบบรู้จำคำศัพท์จะเป็นแบบสำหรับบุคคลทั่วไปที่มีความรู้ภาษาอังกฤษ ซึ่งหมายถึง ระบบยังไม่สามารถนำไปใช้ในการรู้จำเสียงภาษาอื่นได้
5. ระบบนี้ยังไม่สามารถรู้จำเสียงพูดที่มีลักษณะเป็นประโยคได้ ยังคงจำกัดอยู่ที่การรู้จำเสียงพูดที่มีลักษณะเป็นคำๆ เท่านั้น

1.4 ขั้นตอนของโครงงาน

1. ศึกษาธรรมชาติของเสียงพูด และทฤษฎีเบื้องต้นเกี่ยวกับเสียง
2. ศึกษาวิธีการบันทึกเสียงของบอร์ด dsPIC30F6014
3. ศึกษาและพัฒนาระบบการรู้จำเสียงบนบอร์ด dsPIC30F6014
4. ทำการทดลองหาเปอร์เซ็นต์ความถูกต้องของการรู้จำเสียง
5. เขียนโปรแกรมจำลองการสั่งงานด้วยเสียงบน Visual Basic 6.0 โดยการรับของมูลจากบอร์ด
6. วิเคราะห์ผลการทดลองที่ได้ โดยวัดผลจากอัตราความถูกต้องในการรู้จำคำพูด
7. สรุปผลของโครงงาน และหาแนวทางในการพัฒนาโครงงานต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 ประโยชน์ที่คาดว่าจะได้รับจากโครงงาน

1. สามารถนำการรู้จำเสียงพูดที่มีลักษณะไม่ต่อเนื่องไปประยุกต์ใช้กับโปรแกรมประยุกต์บางอย่างได้ เช่น การใช้เสียงในการเรียกหมายเลขโทรศัพท์ , การสั่งงานในคอมพิวเตอร์ด้วยเสียง นอกจากนี้ยังสามารถนำไปช่วยอำนวยความสะดวกให้กับคนพิการได้อีกด้วย
2. สามารถนำเอาทฤษฎีต่างๆ ที่เกี่ยวข้องมาใช้ในการแยกคำออกจากเสียงพูดได้ และสามารถนำไปใช้ในการแยกคำภาษาอื่นๆ ได้
3. สามารถนำเอาผลจากโครงงานนี้ไปพัฒนางานวิจัยอื่นๆ ที่เกี่ยวข้องกับโครงงานนี้ได้ เช่น ใช้เป็นแนวทางในการวิเคราะห์คำพูดแบบต่อเนื่อง

1.6 รายระเอียดของปริญญานิพนธ์

บทที่ 1 จะกล่าวถึง ความสำคัญและที่มาของปัญหา วัตถุประสงค์ของโครงงานขอบเขตของโครงงาน ขั้นตอนของโครงงาน และประโยชน์ที่คาดว่าจะได้รับจากโครงงาน

บทที่ 2 จะกล่าวถึง ทฤษฎีต่างที่เกี่ยวข้องกับการรู้จำเสียง เช่น ทฤษฎีของเสียง ทฤษฎีของ MFCC ทฤษฎีของ HMM คุณสมบัติต่างๆของ dsPIC30F6014

บทที่ 3 จะกล่าวถึง การออกแบบและการสร้าง ว่าเราจะสามารถทำการรู้จำเสียงโดยใช้ dsPIC30F6014 เป็นตัวประมวลผลได้อย่างไร

บทที่ 4 จะกล่าวถึง การทดลองและผลการทดลอง เพื่อดูว่าประสิทธิภาพของการรู้จำเสียงโดยใช้ dsPIC30F6014 เป็นตัวประมวลผลมีความสามารถหรือความถูกต้องในการรู้จำเสียงมากน้อยเพียงใด

บทที่ 5 จะกล่าวถึง สรุป ว่า dsPIC30F6014 มีความสามารถในการรู้จำเสียงได้หรือไม่ หรือมีความสามารถมากน้อยเพียงใด ปัญหาที่พบในระหว่างการทำโครงงานนี้และข้อเสนอแนะต่างๆ เพื่อนำไปพัฒนาต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีที่เกี่ยวข้อง

ในการทำกระบวนการการรู้จำเสียงนั้น เราจำเป็นที่จะต้องศึกษาทฤษฎีต่างๆที่เกียวข้อง เพื่อที่จะสามารถดำเนินโครงงานได้อย่างถูกต้อง เช่น ศึกษาเรื่องเสียงของคนเราว่ามีคุณสมบัติอย่างไร ศึกษาว่าเราจะสามารถนำเสียงมาประมวลผลหรือหาค่าตัวแทนของเสียงได้อย่างไร เราจะสามารถรู้ได้อย่างไรว่าเสียงที่เราพูดนั้นเป็นเสียงใด มีวิธีการเปรียบเทียบอย่างไร ซึ่งเป็นคำถามที่เราจะต้องหาคำตอบต่อไป

2.1 ทฤษฎีการสร้างเสียงพูด

การพูดของมนุษย์มิใช่อาการที่เกิดที่ปากอย่างเดียวเท่านั้น หากเริ่มจากลมหายใจเข้าของมนุษย์ที่นำลมเข้าสู่ปอดจากนั้นจะใช้ลมจากปอด ซึ่งก็คือลมหายใจออก มาทำให้เกิดเสียงพูดโดยลมจะถูกบังคับให้ผ่านอวัยวะต่าง ๆ ที่สำคัญ เราสามารถจำเสียงและวิธีการพูดของคน ๆ หนึ่งได้ เนื่องจากเมื่อพิจารณาเสียงที่พูดออกมาในแต่ละครั้งอาจจะสังเกตลักษณะที่แตกต่างกันได้ เสียงพูดที่จะอธิบายด้วยหลักเกณฑ์ทางวิทยาศาสตร์แม้ว่าในภาษาหนึ่ง ๆ จะมีเสียงต่างกันมากบ้างน้อยบ้าง แต่ละเสียงก็สามารถนำมาพิจารณาและอธิบายให้รู้ลักษณะการออกเสียงและตำแหน่งที่เกิดเสียงได้ คำอธิบายนี้จะทำให้เข้าใจลักษณะเสียงทุกเสียง วิชาที่ว่าด้วยเสียงพูดเรียกว่า วิชาสัทศาสตร์ (Phonetics)

2.1.1 อวัยวะที่ใช้ในการออกเสียงพูด

อวัยวะที่ใช้ในการออกเสียงมีอยู่หลายส่วน แต่ละส่วนสามารถทำให้เสียงพูดแตกต่างกันไป ได้อวัยวะที่ใช้ในการเปล่งเสียงพูดแบ่งออกเป็น 2 ประเภท คือ

ประเภท 1 อวัยวะที่ใช้ในการทำอาการ (Articulator) คืออวัยวะที่เคลื่อนไหวเพื่อผลักลมไปยังส่วนต่าง ๆ อวัยวะที่สำคัญ คือ ลิ้น ซึ่งเป็นส่วนที่เคลื่อนไหวได้มากที่สุด

ประเภท 2 ที่เป็นตำแหน่งที่เกิดเสียงต่าง ๆ (Point Of Articulator) คือตำแหน่งที่เกิดเสียงต่าง ๆ เช่น ริมฝีปาก ฟัน เพดาน ส่วนต่าง ๆ เป็นต้น

อวัยวะส่วนที่มีหน้าที่ในการออกเสียงโดยตรงมีดังนี้

1. ริมฝีปาก เป็นอวัยวะส่วนที่เคลื่อนไหวได้มากและทำให้เสียงแตกต่างกันได้มาก เราอาจบังคับปากอยู่ชิด กัน ห่างกัน ยื่นออกมา หรือห่อกลม เป็นต้น ลักษณะริมฝีปากแบบต่าง ๆ นี้ล้วนมีอิทธิพลต่อการออกเสียงและทำให้เสียงแตกต่างกันไปทั้งสิ้น

2. ฟัน เป็นอวัยวะที่ทำให้เกิดเสียงหลายชนิด เช่น เมื่อฟันกดลงบนริมฝีปากล่าง หรือฟันล่าง ลมที่ผ่านออกมาโดยแรงจะลอดช่องที่พอผ่านได้ออกมา ทำให้เกิดเป็นเสียงชนิดที่เรียกว่าเสียงเสียดแทรก เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. ปุ่มเหงือก เป็นส่วนนูนออกมาอยู่หลังฟันด้านบน ถ้าเอาลิ้นแตะดูจะรู้ว่าเป็นคลื่น

4. เพดานแข็ง หรือเพดานปาก คือส่วนเฉพาะที่โค้งเป็นกระดูกแข็ง

5. เพดานอ่อน คือส่วนที่อยู่ต่อจากเพดานแข็งไปข้างใน มีลักษณะเป็นกระดูกอ่อนที่ขยับขึ้น–ลงได้ เวลาหายใจเพดานอ่อนและลิ้นไก่ซึ่งอยู่ปลายเพดานอ่อนลดระดับลงมา เปิดช่องว่างให้ลมออกทางจมูกเวลาพูดส่วนใหญ่ปลายเพดานอ่อนและลิ้นไก่จะถูกยกขึ้นไปจรดหลังคอ นอกจากเวลาออกเสียงนาสิกเท่านั้นที่เพดานอ่อนจะลดระดับลงมา เพื่อให้ลมออกทางจม

6. ลิ้นไก่ เป็นก้อนเนื้อเล็ก ๆ อยู่ต่อปลายเพดานอ่อนตรงกลางปาก อวัยวะส่วนนี้สั่นรั่ว

7. ลิ้น เป็นส่วนที่เคลื่อนที่มากที่สุดในการออกเสียงพูด จึงต้องแบ่งออกเป็น 3 ส่วน ตามหน้าที่ในการออกเสียง คือ

7.1) ปลายลิ้น คือส่วนปลายลิ้นซึ่งสามารถยกขึ้นไปแตะกับอวัยวะส่วนต่าง ๆ ในปากตอนบนได้โดยง่าย

7.2) หน้าลิ้น คือส่วนที่อยู่ตรงข้ามกับเพดานแข็ง ถ้าวางลิ้นราบกับปาก

7.3) หลังลิ้น ถ้าวางลิ้นราบกับปาก ลิ้นส่วนนี้จะอยู่ตรงข้ามกับเพดานอ่อน

8. แผ่นเนื้อปากหลอดลม เป็นก้อนเนื้อเล็ก ๆ คล้ายลิ้นไก่ อยู่ต่อโคนลิ้นลงไปในคอ มีหน้าที่ปิดช่องลมในขณะรับประทานอาหาร และเปิดช่องลมเมื่อพูด

9. กรวยคอ หมายถึง โพรงคอที่อยู่ถัดจากช่องปากลงไปจนถึงเส้นเสียง

10. เส้นเสียง หรือสายเสียง เป็นอวัยวะสำคัญที่ทำให้เกิดเสียง เส้นเสียงมีลักษณะเป็นกล้ามเนื้อ 2 แผ่นปิดขวางอยู่บริเวณปากช่องหลอดลมจากด้านหลังมาด้านหน้า ระหว่างเส้นเสียงจะมีช่องว่างซึ่งเป็นช่องผ่านให้ลมไปถึงปอดและออกมาจากปอดได้ ช่องว่างนี้เรียกว่า ช่องว่างระหว่างเส้นเสียง (Glottis) เส้นเสียงทั้งสองสามารถดึงให้ห่างออกหรือดึงเข้าหากันได้ เส้นเสียงนี้เป็นส่วนสำคัญที่ทำให้เกิดเสียงพูดขึ้น

11. ช่องจมูก หมายถึง โพรงในช่องจมูกซึ่งอยู่เหนือลิ้นไก่ขึ้นไป เป็นช่องที่ลมซึ่งผ่านเส้นเสียงขึ้นมาจะผ่านออกไปทางจมูกได้เมื่อเวลาหายใจและเวลาออกเสียงนาสิก ในเวลาที่พูดเสียงอื่นลิ้นไก่จะถูกยกขึ้นไปปิดช่องจมูกเพื่อให้ลมออกมาทางปาก

เสียงที่เกิดขึ้นนั้นไม่ว่าจะเป็นเสียงประเภทใด จะผ่านไปตามทางเดินของเสียง โดยจะเริ่มตั้งแต่ช่องว่างระหว่างเส้นเสียงถึงริมฝีปาก ในทางวิทยาศาสตร์สรุปได้ส่าทางเดินของเสียงคือ ท่อนำเสียงที่มีรูปร่างไม่แน่นอน

12. เส้นเสียงปลอม เป็นอวัยวะ ที่มีลักษณะเหมือนเส้นเสียงแต่อยู่เหนือเส้นเสียงขึ้นไป เส้นเสียงปลอมนี้เข้าใจกันว่าจะดึงเข้าหากันเมื่อเวลาพูดเสียงกระซิบ

2.1.2 การเกิดของเสียง

การเกิดของเสียงแบ่งออกเป็น 3 ขั้นตอนคือ

ขั้นตอนที่ 1 จุดเริ่มต้น เป็นขั้นตอนที่ลมเริ่มถูกขับออกจากปอด ผ่านเข้าไปสู่ขั้นตอนที่2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 2 การดัดแปลงลมที่เส้นเสียง อวัยวะที่ใช้ในขั้นตอนนี้คือส่วนที่ต่อจากปอดขึ้นมาถึงกล่องเสียงและที่กล่องเสียง เส้นเสียงจะทำหน้าที่เป็นลิ้นปิดเปิดทำให้เกิดเสียง 2 ชนิดคือ

1. เสียงก้อง หรือเสียงโฆษะ (Voice) คือเสียงที่เกิดจากเส้นเสียงปิดกั้นลมไว้ ลมที่ผ่านออกมาจะเพิ่มแรงดันมากขึ้นจนเส้นเสียงปิด-เปิดสลับกันไป เป็นผลให้สัญญาณเสียงที่ได้ (Speech Waveform) มีลักษณะเป็นคาบ (Quasi-periodic) ซึ่งสามารถเรียกความถี่ในการปิด-เปิดของเส้นเสียงว่า ความถี่มูลฐาน (Fundamental Frequency : F_0)

2. เสียงไม่ก้อง หรือเสียงอโฆษะ (Unvoice หรือ Voiceless) คือเสียงที่เกิดในขณะที่เส้นเสียงเปิดอากาศที่ไหลผ่านอย่างรวดเร็วทำให้เกิดเสียงที่เป็นเสียงของสัญญาณรบกวน (Noise) ซึ่งไม่เป็นคาบ

ขั้นตอนที่ 3 การเปลี่ยนแปลงลักษณะเส้นเสียง อวัยวะที่ใช้คือส่วนที่ต่อจากกล่องเสียงจนถึงริมฝีปาก โดยลมที่ผ่านออกจากกล่องเสียงจะทำให้เกิดเสียงในลักษณะต่าง ๆ ซึ่งจะเกิดจากการเปลี่ยนแปลงของอวัยวะที่กล่าวไว้ในหัวข้อที่แล้ว

2.2 ทฤษฎีการรู้จำเสียง

การรู้จำ (Recognition) เป็นคุณสมบัติพื้นฐานของความเป็นมนุษย์ตลอดจนสิ่งมีชีวิตอื่นๆ แบบ (Pattern) คือสิ่งอธิบายคุณลักษณะหรือรูปร่างของวัตถุหรือสิ่งต่างๆ ในชีวิตจริงนั้นเรากระทำการรู้จำแบบทุก ๆ ย่างก้าวของเรา ตัวอย่างเช่น การเคลื่อนไหวของเราจะมีความสัมพันธ์กับวัตถุรอบ ๆ ตัวเรา ดังนั้นเราจึงจำเป็นจะต้องรูปว่าวัตถุที่อยู่รอบ ๆ ตัวเราคืออะไรเพื่อที่จะนำไปตัดสินใจว่าจะเคลื่อนไหวไปในทิศทางใด เช่น เราต้องการค้นหาเพื่อนเราในท่ามกลางฝูงชนเราจำเป็นต้องรู้จำเพื่อนเรา ซึ่งเราต้องจำรูปร่างลักษณะของเพื่อนของเราได้เพื่อที่จะนำไปเปรียบเทียบกับภาพที่เราเห็น นอกจากนั้นมนุษย์ยังทีกระบวนการที่ซับซ้อนยิ่งขึ้นไปกว่าความสามารถของการรู้จำเพียงอย่างเดียว เช่น เราสามารถแยกแยะได้ว่ารอยยิ้มที่เราเห็นเกิดจากความโกรธหรือความสุขเป็นต้น โดยตามธรรมชาติแล้วเราสามารถแบ่งการรู้จำได้เป็นสองชนิด

ชนิดแรก คือการรู้จำสิ่งที่มีตัวตนซึ่งอาจกล่าวได้ว่าเป็นการรู้จำของประสาทสัมผัสซึ่งรวมไปถึงการมองเห็นและการรับฟัง เช่น การรู้จำตัวอักขระ เสียงเพลง รูปภาพ และวัตถุต่าง ๆ รอบตัวเราเป็นต้น

ชนิดที่สอง คือรู้จำสิ่งที่ไม่มีตัวตน ตัวอย่างเช่นการรู้จำกระบวนการทางด้านความคิด โดยจะเป็นการรู้จำเหตุผล หรือวิธีการแก้ปัญหาต่าง ๆ ที่เราได้ประสบมา ซึ่งบางครั้งเรียกว่า Conceptual Recognition

2.2.1 การหาค่าพารามิเตอร์

การบีบอัดช่วงพิสัยพลวัตของสัญญาณเสียงพูด โดยการลดความแปรปรวนของสัญญาณด้วยการพิจารณาความเปลี่ยนแปลงของสัญญาณในแต่ละช่วง ซึ่งจะส่งผลให้ค่าอัตราส่วนสัญญาณ/

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช

นอยซ์ (Signal Noise Ratio) มีค่าสูงขึ้น ซึ่งทางปฏิบัติแล้วจะนำสัญญาณมาผ่านตัวกรอง Digital อันดับ1

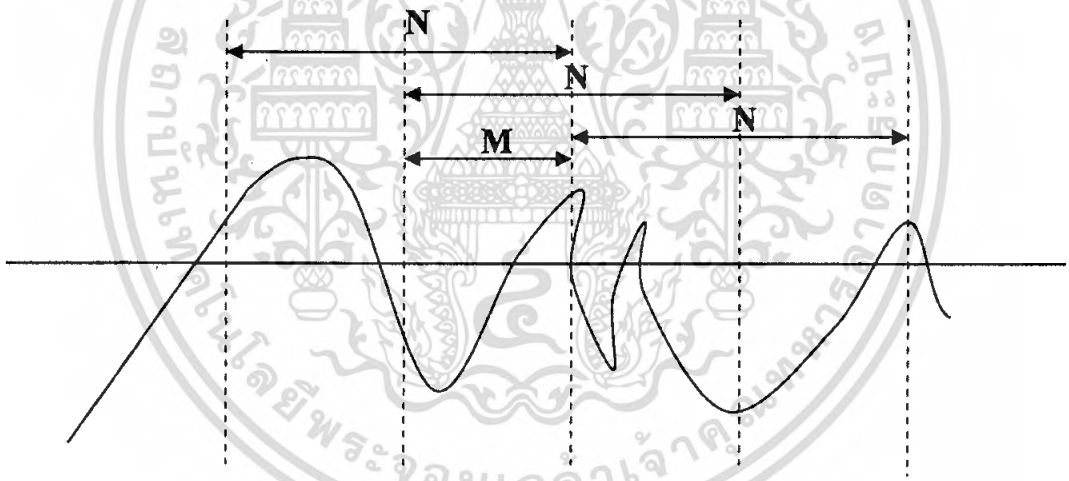
ให้สัญญาณที่เข้ามาเป็น s(n) และสัญญาณที่ผ่านการพรีเอมฟาซีสแล้วเป็น Y(n)เราสามารถเขียนความสัมพันธ์ของทั้งสองสัญญาณโดยสมการที่ (2.1)

$$Y(n) = x(n) - ax(n-1) \qquad (2.1)$$

ยิ่งค่า a เข้าใกล้ 1 เท่าใด ความถี่สูงจะถูกขยายมากขึ้นเท่านั้น ค่า a ที่นิยมใช้ในการหาพารามิเตอร์แอลพีซี คือค่า a = 0.95

2.2.2 การแบ่งช่วงสัญญาณ

ข้อมูลเสียงซึ่งอยู่ในโดเมนเวลา จะถูกแบ่งออกเป็นช่วงสัญญาณ (Block) ซึ่งแต่ละช่วงสัญญาณจะประกอบด้วยสัญญาณเสียง N สัญญาณ และในแต่ละช่วงสัญญาณจะถูกวิเคราะห์โดยการเลื่อนสัญญาณไปครั้งละ M สัญญาณ ดังแสดงในรูปที่ 2.1



รูปที่ 2.1 แสดงการแบ่งช่วงสัญญาณ

จะเห็นได้ว่าถ้าค่า M มีค่าน้อยกว่าค่า N มากเท่าใด จะทำให้การวิเคราะห์สัญญาณมีความแม่นยำมากเท่านั้น แต่ข้อเสียคือจะต้องใช้เวลาในการวิเคราะห์นานขึ้น และถ้าค่า M มีค่ามากกว่าค่า N แล้วจะทำให้สัญญาณบางส่วนไม่ถูกใช้ในการวิเคราะห์จะเกิดการผิดพลาดในการวิเคราะห์สัญญาณในส่วนต่อไป

การกำหนดขนาดของช่วงสัญญาณมีเงื่อนไขในการเลือกดังนี้

1. ค่า M จะต้องสั้นพอที่ทำให้คุณสมบัติของเสียงไม่เปลี่ยนแปลง
2. ค่า N จะต้องยาวพอที่จำนวนของตัวอย่างมีเพียงพอต่อการหาสัมประสิทธิ์
3. การเลื่อนการวิเคราะห์ (ค่า M) ต้องไม่ข้ามข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.3 การวินโดว์

เนื่องจากเราใช้วิธีตัดสัญญาณเพื่อมาหาสเปกตรัมเป็นบล็อก ๆ ซึ่งเท่ากับเป็นการหาสเปกตรัมของ "บล็อกของสัญญาณ" ไม่ใช่ตัวสัญญาณจริง ๆ ที่เข้ามาติดต่อกัน ไม่มีจุดเริ่มต้นหรือสิ้นสุด การตัดสัญญาณเป็นบล็อกนี้จะทำให้เกิดความคลาดเคลื่อนในสเปกตรัมที่ได้ วิธีลดผลของความคลาดเคลื่อนนี้ทำได้โดยคูณสัญญาณแต่ละบล็อก หรือ x_i (n) ด้วยฟังก์ชั่นหน้าต่าง (window function) ก่อนที่จะทำการหาสเปกตรัม

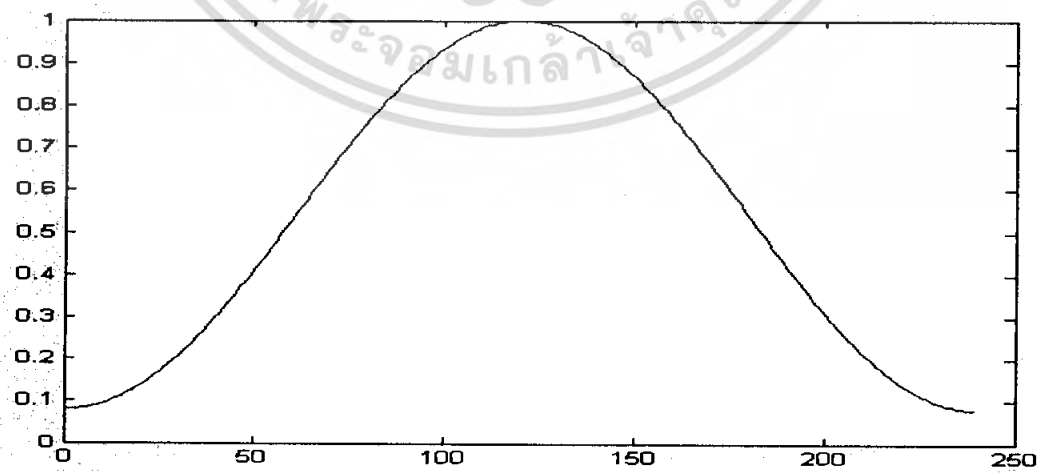
ฟังก์ชั่นหน้าต่างมีหลายแบบ เป็นฟังก์ชั่นที่มีลักษณะสมมาตร และมีค่าสูงสุดที่จุดกึ่งกลาง มีความยาวเท่าไหร่ก็ได้ตามต้องการ (ตามค่าที่แทนลงในสูตร) ตัวอย่างของฟังก์ชั่นหน้าต่างที่เป็นที่นิยมก็คือ หน้าต่างแฮมมิ่ง (Hamming window) ดังรูปที่ 2.2 ซึ่งมีสมการคือ

$$w(n) = 0.54 - 0.46\cos(\frac{2\pi n}{n-1}) \tag{2.2}$$

$$n = 0,1,2,...,N-1$$

$$x_{inew}(n) = x_i(n)w(n) \tag{2.3}$$

เราต้องการ w(n) ที่มีความยาวเท่ากับ 1 บล็อกสัญญาณ ดังนั้นต้องใช้ N เท่ากับความยาว 1 บล็อกสัญญาณ รูปร่างของหน้าต่างแบบแฮมมิ่งแสดงดังรูป ซึ่งจะสังเกตเห็นได้ว่า ฟังก์ชั่นหน้าต่างมีค่าเล็กที่ส่วนต้นและส่วนปลาย ซึ่งเป็นส่วนที่ลดผลของการเปลี่ยนแปลงที่จุดเริ่มต้น และจุดสิ้นสุดของบล็อก



รูปที่ 2.2 ตัวอย่างของฟังก์ชันหน้าต่างแบบแฮมมิ่ง ที่ N = 240

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.4 การหาพลังงานของเสียง

พลังงานของสัญญาณเป็นตัวแทนอันหนึ่งที่เรามักจะนำมาใช้ในการวิเคราะห์ลักษณะต่าง ๆ ของสัญญาณทั่ว ๆ ไป โดยพลังงานของสัญญาณ s(n) ใด ๆ ที่แปรตามเวลาสามารถนิยามได้ว่า

$$E = \sum_{n=-\infty}^{\infty} s^2(n) \tag{2.4}$$

แต่สำหรับสัญญาณเสียงซึ่งเป็นสัญญาณที่แปรเปลี่ยนอยู่ตลอดเวลา ไม่มีเสถียรภาพตามเวลา เราจะต้องแบ่งสัญญาณออกมาพิจารณาเป็นช่วงเล็ก ๆ ตามแกนเวลาหรือเรียกว่าแบ่งออกเป็นเฟรม ดังนั้นก็จะสามารถหาพลังงานของเสียงในแต่ละเฟรมได้เป็น

$$E_l(m) = \sum_{n=0}^{N-1} s^2(n) \tag{2.5}$$

เมื่อ N แทนจำนวนข้อมูลเสียงในแต่ละเฟรม

การวัดค่าพลังงานดังในสมการที่ (2.5) นั้นมีข้อจำกัดตรงที่ว่ามันจะมีความไวต่อสัญญาณที่มีขนาดใหญ่ ๆ เนื่องจากเราใช้วิธียกกำลังสองค่าของสัญญาณอินพุท ดังนั้นการแก้ปัญหาอย่างหนึ่งก็คือ วัดพลังงานของเสียงโดยใช้สมการดังนี้

$$E_l(m) = \sum_{n=0}^{N-1} |s(n)| \tag{2.6}$$

2.2.5 การแปลงฟูเรียร์

การแปลงฟูเรียร์เป็นทฤษฎีหนึ่งที่ใช้ในการวิเคราะห์สัญญาณ โดยจะแปลงสัญญาณจากโดเมนของเวลา (Time domain) ไปอยู่ในโดเมนของความถี่ (Frequency domain) โดยที่สมการที่ (2.7) เป็นสมการการแปลงฟูเรียร์จากสัญญาณที่อยู่ในโดเมนเวลาให้กลายเป็นโดเมนความถี่ ส่วนสมการที่ (2.8) เป็นสมการการแปลงกลับฟูเรียร์โดยแปลงสัญญาณจากโดนเมนความถี่ให้อยู่ในรูปโดเมนเวลา

การแปลงฟูเรียร์:

$$X(f) = \int_{-\infty}^{\infty} x(t) \cdot e^{-2j\pi ft} dt \tag{2.7}$$

เมื่อ X(f) คือ สัญญาณเสียงที่ได้จากการทำ Fourier Transform ซึ่งจะอยู่ในโดเมนของความถี่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

x(t) คือ สัญญาณเสียงในโดเมนของเวลา

f คือ ความถี่เสียง และ t คือ เวลา

การแปลงฟูเรียร์ย้อนกลับ:

$$x(t) = \int_{-\infty}^{\infty} X(f) \cdot e^{2j\pi ft} df \quad (2.8)$$

เมื่อ x(t) คือ สัญญาณเสียงที่ได้จากการทำ Inverse Fourier Transform ซึ่งจะอยู่ในโดเมนของเวลา

X(f) คือ สัญญาณเสียงที่ได้จากการแปลงฟูเรียร์ ซึ่งจะอยู่ในโดเมนของความถี่

f คือ ความถี่เสียง และ t คือ เวลา

จากสมการที่ (2.7) ถ้าสัญญาณ x(t) มีความถี่ตรงกับคลื่นไซนูซอยด์ (Sinusoidal Wave) จะให้ค่าสัมประสิทธิ์ฟูเรียร์ $X(f)$ มาก และในทางตรงข้ามถ้าค่าสัมประสิทธิ์ฟูเรียร์ $X(f)$ มีค่าน้อย ถ้าสัญญาณ x(t) ไม่ค่อยมีสัญญาณไซนูซอยด์ความถี่นั้นๆ อยู่

ผลลัพธ์การแปลงฟูเรียร์จะแสดงถึงสัญญาณไซนูซอยด์ที่ความถี่ต่างๆ ซึ่งประกอบอยู่ในสัญญาณ x(t) ใดๆ แต่ไม่สามารถแสดงถึงเวลาที่สัญญาณนั้นๆ ปรากฏ นั่นคือเราสามารถนำการแปลงฟูเรียร์ไปใช้ในการวิเคราะห์สัญญาณในโดเมนของความถี่ได้เพียงอย่างเดียว เนื่องจากข้อมูลทางด้านเวลาจะหายไป

2.2.6 การประมวลผลของเมลฟรีเควนซี่เซปตรัมโคเอฟฟีเชี่ยน

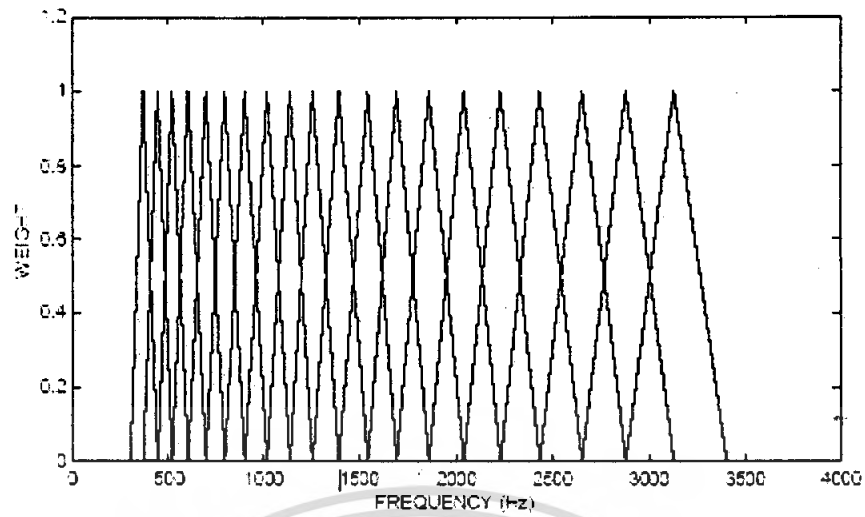
แรงบันดาลใจในการใช้ MFCC ก็เนื่องมาจากข้อเท็จจริงที่ว่าโสตประสาทในหูของมนุษย์รับความถี่แบบที่ไม่เป็นเชิงเส้น เราสามารถแปลงให้อยู่ในความถี่ของ Mel ได้ตามสมการ

$$\text{Mel}(f) = 2595 \log_{10}\left(1 + \frac{f}{700}\right) \quad (2.9)$$

เมื่อ f คือ ความถี่ของเสียง และ Mel(f) คือ ความถี่ของ Mel

ดังรูปที่ 2.3 ซึ่งแสดงการแบ่งช่วงความถี่ด้วยตัวกรองความถี่แบนด์พาส ที่มีลักษณะเป็นสามเหลี่ยมจำนวน 20 ตัว ที่ทำการแบ่งช่วงระยะห่างของแต่ละตัวด้วยความถี่ Mel ที่เท่าๆ กัน ในช่วง 300 ถึง 3400 Hz

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.3 ตัวกรองความถี่สำหรับการสร้าง MFCC ที่มีความถี่อยู่ระหว่าง 300 ถึง 3400 Hz

เราสามารถสรุปขั้นตอนในการคำนวณหา MFCC ได้ดังไดอะแกรมดังรูปที่ 2.4

รูปที่ 2.4 ไดอะแกรมแสดงกระบวนการหา MFCC

1. แปลงสัญญาณ x[n] ที่อยู่ในโดเมนเวลาให้กลายเป็นสัญญาณ X[k] ที่อยู่ในโดเมนความถี่โดยใช้ Discrete Fourier Transform (DFT) ตามสมการ

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi 2\pi k} \qquad (2.10)$$

เมื่อ N คือ จำนวนตัวอย่างของสัญญาณเสียง

2. นำ spectrum ที่ได้มาผ่านตัวกรองความถี่ตามสมการ

$$m_j = \sum_{k=0}^{N-1} |X[k]|^2 \, H_j[k] \qquad 0 \le j \le p \qquad (2.11)$$

เมื่อ $H_j[k]$ คือ transfer function ของตัวกรองความถี่ตัวที่ j

$|X[k]|^2$ คือ พลังงานของสัญญาณ

3. MFCC จะถูกคำนวณโดยการใช้ Discrete Cosine Transform (DCT)

$$MFCC_i = \sqrt{\frac{2}{N}} \sum_{j=1}^{N} m_j \cos\left(\frac{\pi i}{N}(j - 0.5)\right) \qquad (2.12)$$

เมื่อ N คือ จำนวนตัวกรองความถี่

m_j คือ ขนาดผลลัพธ์ของตัวกรองความถี่แต่ละช่อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.7 รูปแบบของมาร์คอฟ

รูปแบบของมาร์คอฟ เกิดขึ้นมาจากความพยายามในการจำลองการทำงานของการสร้างเสียงในทางสถิติ ดังนั้นวิธีการนี้จึงจัดอยู่ในประเภทแรก และในช่วง 10 ปีที่ผ่านมา วิธีการนี้ได้ประสบความสำเร็จกับการนำไปประยุกต์ใช้กับงานทางด้านการรู้จำเสียงพูดอัตโนมัติ โดยมีสาเหตุหลักมาจากความสามารถในการอธิบายลักษณะของสัญญาณเสียงในแนวทางทางคณิตศาสตร์ที่สามารถพิสูจน์ได้

2.2.7.1 นิยามและหลักการพื้นฐาน

รูปแบบของมาร์คอฟ คือ เซตของสเตท ที่มีจำนวนจำกัด โดยในแต่ละสเตท จะเกี่ยวเนื่องกับฟังก์ชันการแจกแจงความน่าจะเป็น (Probability Distribution Function) ในส่วนของการเปลี่ยนสเตท จะถูกกำกับโดยชุดข้อมูลที่เรียกว่าความน่าจะเป็นในการเปลี่ยนสเตท (State Transition Probabilities) แต่ละสเตท จะสามารถสร้างผลลัพธ์หรือที่เรียกว่า ข้อสังเกต ขึ้นมาตามการแจกแจงความน่าจะเป็น ซึ่งผลที่ได้นั้นเป็นเพียงแค่ผลลัพธ์ที่สร้างขึ้นมา โดยแบบจำลองเท่านั้น ไม่ได้มาจากการมองเห็นสเตท ของระบบภายในแต่อย่างใด ดังนั้นจะพบว่าสเตท เหล่านี้จะถูกซ่อน (Hidden) จากผู้สังเกต

ในการที่จะให้นิยามอย่างสมบูรณ์ เราจำเป็นต้องกล่าวถึง ส่วนประกอบต่างๆ ของ รูปแบบของมาร์คอฟ ดังต่อไปนี้

1. จำนวนสเตทของแบบจำลอง, N
2. จำนวนสัญลักษณ์ในการสังเกต (Observation Symbol) , M แต่ถ้าเป็นกรณีที่ผลจากการสังเกตมีลักษณะต่อเนื่อง (Continuous Observation) M จะมีค่าอนันต์
3. เซตของความน่าจะเป็นในการเปลี่ยนสเตท, $A = \{a_{ij}\}$

$$a_{ij} = \Pr\{q_{t+1} = j \mid q_t = i\}, \quad 1 \le i, j \le N \tag{2.13}$$

เมื่อ q_t คือ สเตทปัจจุบันของระบบ

q_{t+1} คือ สเตทถัดไปของระบบ

ความน่าจะเป็นในการเปลี่ยนสเตทจะต้องมีคุณสมบัติตามข้อกำหนดทางความน่าจะเป็นโดยทั่วไป คือ

$$a_{ij} \ge 0, \quad 1 \le i, j \le N \tag{2.14}$$

และ

$$\sum_{j=1}^{N} a_{ij} = 1, \quad 1 \le i, j \le N \tag{2.15}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. การแจกแจงความน่าจะเป็นในแต่ละสเตท, $B = \{b_{ij}(k)\}$

$$b_j(k) = \Pr\{o_t = v_k \mid q_t = j\}, \quad 1 \le j \le N,\ 1 \le k \le M \tag{2.16}$$

เมื่อ v_k คือ สัญลักษณ์ผลลัพธ์ที่สังเกตเห็นในตัวที่ k

o_t คือ ตัวแปรปัจจุบันที่ทำการสังเกต

q_t คือ สเตทปัจจุบันของระบบ

การแจกแจงความน่าจะเป็นในแต่ละสเตทจะต้องมีคุณสมบัติตามข้อกำหนด คือ

$$b_j(k) \ge 0, \quad 1 \le j \le N,\ 1 \le k \le M \tag{2.17}$$

และ

$$\sum_{k=1}^{M} b_j(k) = 1, \quad 1 \le j \le N \tag{2.18}$$

ในกรณีที่จะใช้ค่าจากการสังเกตเป็นค่าต่อเนื่อง เราจะใช้ฟังก์ชันความหนาแน่นความน่าจะเป็นแบบต่อเนื่อง แทนที่การใช้ความน่าจะเป็นแบบไม่ต่อเนื่อง ซึ่งโดยปกติแล้วเราสามารถประมาณค่าความหนาแน่นความน่าจะเป็นได้จากสมการต่อไปนี้

$$b_j(O) = \sum_{m=1}^{M} c_{jm}\, \eta(\mu_{jm}, U_{jm}, O) \tag{2.19}$$

เมื่อ O คือ เวกเตอร์ที่เราจะทำการสร้างแบบจำลองการทำงานของ

c_{jm} คือ สัมประสิทธิ์ส่วนผสม (Mixture) ของส่วนผสมที่ m ใน State j

η คือ ฟังก์ชันความหนาแน่นความน่าจะเป็นที่มีลักษณะเป็นรูปครึ่งวงรีสมมาตรที่มีเวกเตอร์ค่าเฉลี่ย (Mean Vector) μ_{jm} และมีเมตริกซ์ความแปรปรวนร่วม (Covariance Matrix) U_{jm} โดยสามารถหาได้ตามสมการ

$$\eta(\mu, U, O) = \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{|\det(U)|}} \exp\left\{-\frac{1}{2}(O-\mu)^{T} U^{-1} (O-\mu)\right\} \tag{2.20}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยปกติแล้วเราจะใช้ฟังก์ชันความหนาแน่นแบบกัสเสียนสำหรับ η ได้ ส่วนผสม c_{jm} จะต้องเป็นไปตามข้อกำหนดทางความน่าจะเป็นดังนี้

$$c_{jm} \geq 0, \quad 1 \leq j \leq N, 1 \leq m \leq M \tag{2.21}$$

และ

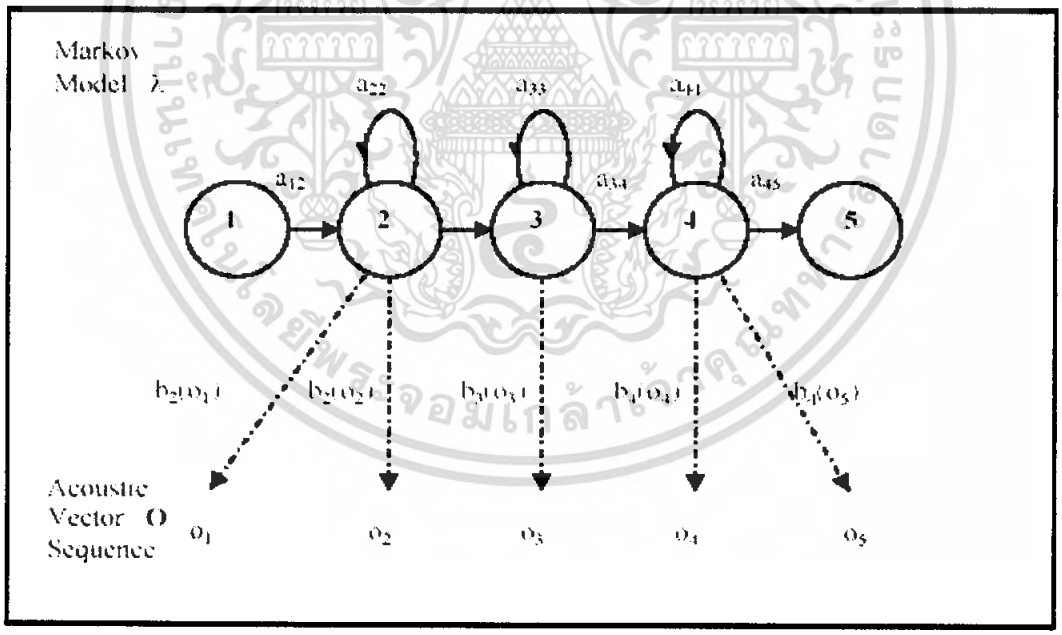
$$\sum_{m=1}^{M} c_{jm}, \quad 1 \leq j \leq N \tag{2.22}$$

5. การแจกแจงความน่าจะเป็นของสเตทเริ่มต้น, $\pi = \{\pi_i\}$

$$\pi_i = \Pr\{q_1 = i\}, \quad 1 \leq i \leq N \tag{2.23}$$

เมื่อ q_1 คือ สเตทแรกของระบบ

ดังนั้นเราจะใช้สัญลักษณ์ $\lambda = (A, B, \pi)$ แทนรูปแบบของมาร์คอฟที่มีการแจกแจงความหนาแน่นเป็นแบบไม่ต่อเนื่อง และ $\lambda = (A, c_{jm}, \mu_{jm}, U_{jm}, \pi)$ แทนรูปแบบของมาร์คอฟที่มีความหนาแน่นเป็นแบบต่อเนื่อง แสดงส่วนประกอบดังรูปที่ 2.5



รูปที่ 2.5 แสดงส่วนประกอบต่างๆ ของ รูปแบบของมาร์คอฟ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.7.2 สมมุติฐานสำหรับทฤษฎีรูปแบบของมาร์คอฟ

การที่เราจะนำรูปแบบของมาร์คอฟ มาใช้ในงานจริง จำเป็นจะต้องมีการตั้งสมมุติฐานเหล่านี้ขึ้นมาสำหรับทฤษฎีดังกล่าว เพื่อลดภาระทางด้านการคำนวณและการวิเคราะห์ สมมุติฐานดังกล่าว ได้แก่

1. สมมุติฐานของกระบวนการมาร์คอฟ (The Markov Assumption)

ดังที่เราได้กล่าวมาแล้วในนิยามของ HMM ความน่าจะเป็นในการเปลี่ยน State คือ

$$a_{ij} = \Pr\{q_{i+1} = j \mid q_t = i\} \tag{2.24}$$

จากสมการดังกล่าว เราจะพบว่าสเตทถัดไปจะเป็นสเตทใดขึ้นอยู่กับสเตทปัจจุบันซึ่งเราจะเรียกสิ่งนี้ว่าสมมุติฐานของมาร์คอฟ และจะทำให้เราได้ผลลัพธ์เป็นแบบจำลองรูปแบบของมาร์คอฟอันดับที่หนึ่ง อย่างไรก็ตามโดยทั่วไปแล้วเราสามารถกำหนดให้การพิจารณาสเตทถัดไปขึ้นอยู่กับ lสเตทก่อนหน้ามากกว่าหนึ่งสเตทซึ่งถ้าเราพิจารณาสเตทถัดไป โดยคำนึงถึงสเตทก่อนหน้าจำนวน k สเตทเราจะเรียกแบบจำลองนี้ว่ารูปแบบของมาร์คอฟอันดับที่ k แต่การทำเช่นนี้ก็จะส่งผลให้การคำนวณมีความซับซ้อนมากขึ้นเช่นกัน

2. สมมุติฐานของความไม่เปลี่ยนแปลงไปตามเวลา (The Stationary Assumption)

ความน่าจะเป็นในการเปลี่ยนสเตทจะเป็นอิสระต่อเวลาจริงๆ ที่การเปลี่ยน State นี้เกิดขึ้นโดยที่เราสามารถอธิบายให้อยู่ในเชิงคณิตศาสตร์ดังนี้

$$\Pr(q_{t_1+1} = j \mid q_{t_1} = i) = \Pr(q_{t_2+1} = j \mid q_{t_2} = i) \tag{2.25}$$

สำหรับทุกค่าของ t_1 และ t_2

3. สมมุติฐานของความอิสระของผลลัพธ์ (The Output Independence Assumption)

สมมุติฐานข้อนี้กล่าวว่าผลลัพธ์จากการสังเกตในตำแหน่งปัจจุบันจะมีค่าเป็นอิสระในทางสถิติกับผลลัพธ์ตัวก่อนหน้า โดยถ้าหากชุดของค่าจากการสังเกตคือ

$$O = o_1, o_2, ..., o_T$$

โดยถ้าเราพิจารณาตามสมมุติฐานที่เราได้กล่าวมาในข้างต้นจะได้ว่า สำหรับรูปแบบของมาร์คอฟ, λ

$$\Pr(O \mid q_1, q_2, ..., q_T, \lambda) = \prod_{t=1}^{T} \Pr(o_t \mid q_t, \lambda) \tag{2.26}$$

อย่างไรก็ตามสมมุติฐานข้อนี้ต่างจาก 2 ข้อที่ได้กล่าวมาในข้างต้น เพราะความจำกัดในการใช้งาน ซึ่งบางกรณีแล้วสมมุติฐานนี้อาจจะไม่เหมาะสม ซึ่งจะกลายเป็นข้อด้อยอย่างหนึ่งของรูปแบบของมาร์คอฟ

2.2.7.3 ปัญหาพื้นฐานของรูปแบบของมาร์คอฟ

จากที่เรากล่าวมาในข้างต้นทั้งหมด เราสามาระสรุปปัญหาพื้นฐานที่จะต้องมีการแก้เพื่อนำแบบจำลองนี้ไปทำงานจริงได้ทั้งหมด 3 ข้อด้วยกัน ได้แก่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. ทำอย่างไรเราจึงจะสามารถคำนวณค่า $Pr(O|\lambda)$ ได้อย่างมีประสิทธิภาพ โดยกำหนดชุดข้อมูลจากการสังเกตเป็น $O = o_1, o_2, ..., o_T$ และแบบจำลองของมาร์คอฟ เป็น $\lambda = (A, B, \pi)$

2. ทำอย่างไรเราจึงจะสามารถเลือกชุดของ State $Q = q_1, q_2, ..., q_T$ ที่มีความสอดคล้องกับแบบจำลองมากที่สุด โดยกำหนดชุดข้อมูลจากการสังเกตเป็น $O = o_1, o_2, ..., o_T$ และแบบจำลอง ของมาร์คอฟ เป็น $\lambda = (A, B, \pi)$

3. ทำอย่างไรเราจึงจะสามารถปรับค่าตัวแปรต่างๆ ของแบบจำลอง $\lambda = (A, B, \pi)$ เพื่อทำให้ค่า $Pr(O|\lambda)$ มีค่าสูงที่สุด

ปัญหาข้อที่ 1 คือ ปัญหาในการหาค่าความน่าจะเป็น (Evaluation Problem) ซึ่งจะเกี่ยวกับการหาค่าความน่าจะเป็นของชุดข้อมูลจากการสังเกตเมื่อเทียบกับแบบจำลอง หรือเราอาจกล่าวได้อีกมุมมองหนึ่งคือ เป็นปัญหาในการให้คะแนนในการเปรียบเทียบความเหมือนของแบบจำลองกับชุดข้อมูลจากการสังเกตที่เราป้อนเข้ามา ซึ่งเราจะพบว่ามุมมองหลังที่เราได้กล่าวถึงนั้นมีประโยชน์อย่างมากในการนำแบบจำลองนี้ไปใช้งาน ยกตัวอย่างเช่น ถ้าเราพยายามที่จะเลือกแบบจำลองที่มีความเหมือนมากที่สุด เทียบกับชุดข้อมูลจากการสังเกตออกมาจากชุดของแบบจำลอง คำตอบจากการแก้ปัญหาข้อแรกนี้จะสามารถทำให้เราเลือกแบบจำลองที่ถูกต้องออกมาได้อย่างมีประสิทธิภาพ

ปัญหาข้อที่ 2 เรียกอีกอย่างหนึ่งว่า ปัญหาการทอดรหัส ซึ่งจะเป็นปัญหาเกี่ยวกับการที่เราพยายามเปิดเผยส่วนที่ซ่อนอยู่ในแบบจำลอง กล่าวคือ เป็นการหาลำดับของสเตท ที่มีความเหมาะสมกับแบบจำลองนั้นๆ มากที่สุด โดยในการหาสเตท ดังกล่าวจะขึ้นอยู่กับข้อกำหนด ซึ่งในความเป็นจริงแล้วเราสามารถเลือกใช้งานข้อกำหนดได้ในหลายรูปแบบ ในแต่ละรูปแบบก็จะมีความแตกต่างกันไป ซึ่งจุดประสงค์ของการแก้ปัญหาข้อนี้มีหลายข้อ เช่น เพื่อเรียนรู้โครงสร้างของแบบจำลอง, เพื่อหาชุดของสเตท ที่มีความเหมาะสมกับการรู้จำเสียงพูดแบบต่อเนื่อง หรือเพื่อหาค่าทางสถิติของแต่ละแบบจำลอง เป็นต้น

ปัญหาสุดท้ายเรียกว่า ปัญหาการเรียนรู้ จะเกี่ยวกับการที่เราพยายามที่จะทำให้แบบจำลองมีความเหมาะสมกับชุดข้อมูลจากการสังเกตที่ให้เข้ามามากที่สุด ซึ่งเราจะเรียกชุดข้อมูลที่ใช้ในการปรับค่าตัวแปรของแบบจำลองนี้ว่า ชุดข้อมูลฝึกหัด (Training Sequence) เพราะว่าชุดข้อมูลเหล่านี้จะถูกนำไปฝึกแบบจำลองให้รู้จำข้อมูลเหล่านั้น ปัญหาการฝึกหัดนี้เป็นปัญหาที่มีความสำคัญมากอย่างหนึ่งในการประยุกต์ใช้งานแบบจำลองในหลายๆ ด้าน เนื่องจากการฝึกหัดนี้จะทำให้เราสามารถปรับค่าของแบบจำลอง ให้เหมาะสมกับชุดข้อมูลฝึกหัดจากการสังเกตได้ กล่าวอีกนัยหนึ่งคือ เพื่อสร้างแบบจำลองของปรากฏการณ์ที่ดีที่สุดนั่นเอง

2.2.7.4 ปัญหาในการหาค่าความน่าจะเป็น

ในการที่เราจะหาค่า $Pr(O|\lambda)$ เมื่อกำหนดแบบจำลอง $\lambda = (A, B, \pi)$ และชุดข้อมูลจากการสังเกต $O = o_1, o_2, ..., o_T$ สามารถทำได้โดยการคำนวณความน่าจะเป็นสำหรับทุกๆ ชุดของสเตท ความยาว T ที่เป็นไปได้ของแบบจำลองนั้นๆ ซึ่งสำหรับแต่ละลำดับของ สเตท

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$Q = q_1, q_2, ..., q_T$$

เมื่อ q_1 คือ สเตทเริ่มต้น

จะได้ความน่าจะเป็นของชุดข้อมูลที่สังเกตได้ O สำหรับแต่ละลำดับของสเตท คือ

$$\Pr(O \mid Q, \lambda) = \prod_{t=1}^{T} \Pr(O_t \mid q_t, \lambda) \tag{2.27}$$

และจากสมมุติฐานความเป็นอิสระของผลลัพธ์ เราจะได้

$$\Pr(O \mid Q, \lambda) = b_{q_1}(O_1) \cdot b_{q_2}(O_2) \cdot ... \cdot b_{q_T}(O_T) \tag{2.28}$$

เราสามารถเขียนความน่าจะเป็นของลำดับสเตท Q ได้คือ

$$\Pr(Q \mid \lambda) = \pi_{q_1} a_{q_1 q_2} a_{q_2 q_3} ... a_{q_{T-1} q_T} \tag{2.29}$$

ในส่วนของความน่าจะเป็นร่วม (Joint Probabilities) ของ O และ Q ซึ่งก็คือความน่าจะเป็นที่ทั้ง O และ Q เกิดขึ้นพร้อมกัน สามารถหาได้ง่ายๆ โดยการคูณทั้งสองสมการข้างบนเข้าด้วยกัน

$$\Pr(O, Q \mid \lambda) = \Pr(O \mid Q, \lambda) \bullet \Pr(Q \mid \lambda) \tag{2.30}$$

ดังนั้นเราสามารถหาความน่าจะเป็นของ O เมื่อให้แบบจำลอง λ ได้จากการรวมความน่าจะเป็นร่วมด้านบนสำหรับทุกๆ ชุดของ State ที่เป็นไปได้ทั้งหมดเข้าด้วยกัน

$$\begin{aligned} \Pr(O \mid \lambda) &= \sum_{\text{all } Q} \Pr(O \mid Q, \lambda) \cdot \Pr(Q \mid \lambda) \\ &= \sum_{q_1 q_2 \cdots q_T} \pi_{q_1} b_{q_1}(O_1) a_{q_1 q_2} b_{q_2}(O_2) ... a_{q_{T-2} q_T} b_{q_T}(O_T) \end{aligned} \tag{2.31}$$

เมื่อทำการวิเคราะห์ความซับซ้อนในการคำนวณพบว่าการหาค่าความน่าจะเป็นโดยตรงในแต่ละครั้ง จะเกี่ยวข้องกับการคำนวณถึง $2T \cdot N^T$ ครั้ง ซึ่งในการคำนวณแต่ละครั้งสำหรับแต่ละสเตท จะมีการคำนวณ 2T ครั้ง ส่วน N^T ก็คือจำนวนสเตทที่เป็นไปได้ทั้งหมด การคำนวณดังกล่าวไม่สามารถที่จะนำมาใช้งานได้ในความเป็นจริง แม้จะมีสเตท ในจำนวนไม่มากนักก็ตาม เช่น ถ้าแบบจำลองมี N=5, T=100 จำนวนครั้งในการคำนวณจะมีสูงถึง $2 \cdot 100 \cdot 5^{100} \approx 10^{72}$ ครั้ง

เพราะฉะนั้นการแก้ปัญหาในข้อนี้จะต้องใช้วิธีอื่นแทนที่การคำนวณโดยตรง วิธีการหนึ่งที่สามารถนำมาใช้แก้ปัญหาข้อนี้มีชื่อเรียกว่า ขบวนการก้าวหน้า-ถอยหลัง ซึ่งเป็นการใช้ตัวแปรช่วย ซึ่งมีชื่อเรียกว่า ตัวแปรก้าวหน้า (Forward Variable) โดยการนำตัวแปรดังกล่าวมาใช้จะช่วยลดความซับซ้อนในการคำนวณลงไปอย่างมาก

เราสามารถให้คำนิยามแก่ตัวแปรก้าวหน้า ได้ว่า ตัวแปรก้าวหน้า $\alpha_t(i)$ คือ ความน่าจะเป็นของชุดข้อมูลจากการสังเกตบางส่วน $o_1, o_2, ...o_t$ (ตั้งแต่เริ่มต้นจนถึงเวลา t) และออกจากสเตท i ที่เวลาดังกล่าว เราสามารถแสดงตัวแปรดังกล่าวในทางคณิตศาสตร์ได้ดังนี้คือ

$$\alpha_t(i) = \Pr(o_1, o_2, ..., o_t, q_t = i \mid \lambda) \tag{2.32}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

72868

เราสามารถหาค่าตัวแปรดังกล่าวได้จากการพิสูจน์ ดังต่อไปนี้

1. การเริ่มต้น:

$$\alpha_1(i) = \pi_i b_i(o_1), \quad 1 \le i \le N \tag{2.33}$$

2. การพิสูจน์:

$$\alpha_{t+1}(i) = \left[\sum_{i=1}^{N} a_t(i) a_{ij}\right] b_j(o_{t+1}), \quad 1 \le t \le T-1,\ 1 \le j \le N \tag{2.34}$$

3. การสิ้นสุด:

$$\Pr(O \mid \lambda) = \sum_{i=1}^{N} \alpha_T(i) \tag{2.35}$$

ถ้าเราพิจารณาความซับซ้อนของการคำนวณด้วยวิธีนี้พบว่า จะต้องมีการคำนวณเพียง N^2T ครั้ง แทนที่จะเป็น $2T \cdot N^T$ อย่างในการคำนวณโดยตรง

ในทำนองเดียวกัน เราสามารถนิยามตัวแปรถอยหลัง (Backward Variable) ซึ่งก็คือ ความน่าจะเป็นของชุดข้อมูลจากการสังเกตบางส่วน $o_{t+1}, o_{t+2}, \ldots o_T$ เมื่อกำหนด State เท่ากับ i เราอธิบายตัวแปรดังกล่าวในทางคณิตศาสตร์ได้ดังนี้

$$\beta_t(i) = \Pr(o_{t+1}, o_{t+2}, \ldots, o_T \mid q_t = i, \lambda) \tag{2.36}$$

ซึ่งเราสามารถหาค่าตัวแปรดังกล่าวได้ดังนี้

1. การเริ่มต้น:

$$\beta_t(i) = 1, \quad 1 \le i \le N \tag{2.37}$$

2. การพิสูจน์:

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(o_{t+1}) \beta_{t+1}(j), \quad t = T-1, T-2, \ldots 1 \quad 1 \le i \le N \tag{2.38}$$

ในความเป็นจริงแล้วเราได้แก้ปัญหาข้อนี้ไปด้วยการใช้ตัวแปรก้าวหน้าเรียบร้อยแล้ว อย่างไรก็ตามเราสามารถแก้ปัญหาข้อนี้ได้อีกแบบโดยพิจารณาตัวแปรถอยหลังเข้ามาร่วมด้วยซึ่งจะถูกนำไปใช้ในขั้นตอนการฝึกหัดระบบ ซึ่งเราจะพบว่า

$$\alpha_t(i)\beta_t(i) = \Pr(O, q_t = i \mid \lambda), \quad 1 \le i \le N,\ 1 \le t \le T \tag{2.39}$$

ซึ่งเราจะคำนวณค่าความน่าจะเป็น $\Pr(O \mid \lambda)$ ได้จาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\Pr(O \mid \lambda) = \sum_{i=1}^{N} \Pr(O, q_t = i \mid \lambda) = \sum_{i=1}^{N} \alpha_t(i)\beta_t(i) \tag{2.40}$$

ซึ่งสมการนี้จะถูกนำไปใช้อีกทีในการฝึกหัดระบบแบบการ์เดียนเบส(Gradient-Based)

2.2.7.5 ปัญหาการถอดรหัส

ในกรณีนี้เราต้องการที่จะหาลำดับของสเตท ที่มีความเหมาะสมที่สุด เมื่อกำหนดชุดข้อมูลจากการสังเกต $O = o_1, o_2, ..., o_T$ และแบบจำลอง λ

การหาคำตอบในข้อนี้จะขึ้นอยู่กับมุมมองของความเหมาะสมของลำดับสเตท ว่าถูกนิยามให้เป็นอย่างไร วิธีแรกอาจจะใช้วิธีการหาสเตท ที่มีความเหมาะสมในช่วงเวลาต่างๆ กันมาต่อกันเพื่อให้ได้ลำดับทั้งหมด แต่ในบางครั้งวิธีการนี้ก็ไม่สามารถให้คำอธิบายที่ดีได้เท่าไรนัก ดังนั้นเราจะพิจารณาวิธีอื่นที่ไม่มีปัญหาดังกล่าว

วิเทอร์บี อัลกอริทึม เป็นวิธีการในการหาลำดับของสเตท $Q = q_1, q_2, ..., q_T$ ที่ดีที่สุดเพียงทางเดินเดียว เมื่อกำหนดชุดข้อมูลจากการสังเกตเป็น $O = o_1, o_2, ..., o_T$ เราจำเป็นต้องกำหนดตัวแปรช่วยเพิ่มเติม

$$\delta_t(i) = \max_{q_1, q_2, ..., q_{t-1}} \Pr(q_1 q_2 ... q_t = i, o_1 o_2 ... o_t \mid \lambda) \tag{2.41}$$

กล่าวได้อีกนัยหนึ่งว่า $\delta_t(i)$ คือ คะแนนที่ดีที่สุด (ความน่าจะเป็นที่สูงที่สุด) ตามเส้นทางที่ค้นหามาที่เวลา t โดยพิจารณาข้อมูลจากการสังเกต t ข้อมูลแรกและจบลงที่สเตท i ซึ่งจากการพิสูจน์แล้วเราจะได้ว่า

$$\delta_{t+1}(j) = \left[\max_i \delta_t(i) a_{ij} \right] \cdot b_j(o_{t+1}) \tag{2.42}$$

ในการที่จะหาลำดับของสเตท ที่ดีที่สุดเราจะต้องติดตามพิจารณาสเตท ในทางเดินว่าสเตทใดบ้างที่ทำให้ค่าในสมการข้างต้นนี้มีค่ามากที่สุด สำหรับแต่ละค่า t และ j ซึ่งเราทำได้โดยกำหนดตัวแปรช่วยเพิ่มอีกตัวหนึ่งคือ $\psi_t(j)$ กระบวนการทั้งหมดสามารถแสดงได้ดังต่อไปนี้

1. การเริ่มต้น:

$$\delta_1(j) = \pi_i b_i(o_1), \qquad 1 \le i \le N \tag{2.43}$$

$$\psi_1(i) = 0 \tag{2.44}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. หลักการ:

$$\delta_t(j) = \max_{1\le i\le N}[\delta_{t-1}(i)a_{ij}]b_j(o_t), \qquad 2\le t\le T \qquad (2.45)$$

$$1\le j\le N$$

$$\psi_t(i) = \arg\max_{1\le i\le N}[\delta_{t-1}(i)a_{ij}], \qquad 2\le t\le T \qquad (2.46)$$

$$1 \le j\le N$$

3. การสิ้นสุด:

$$P^* = \max_{1\le i\le N}[\delta_T(i)] \qquad (2.47)$$

$$q_t^* = \arg\max_{1\le i\le N}[\delta_T(i)] \qquad (2.48)$$

4. เส้นทางการเดินย้อนกลับ:

$$q_t^* = \psi_{t+1}[q_{t+1}^*], \qquad t = T-1, T-2,\ldots,1. \qquad (2.49)$$

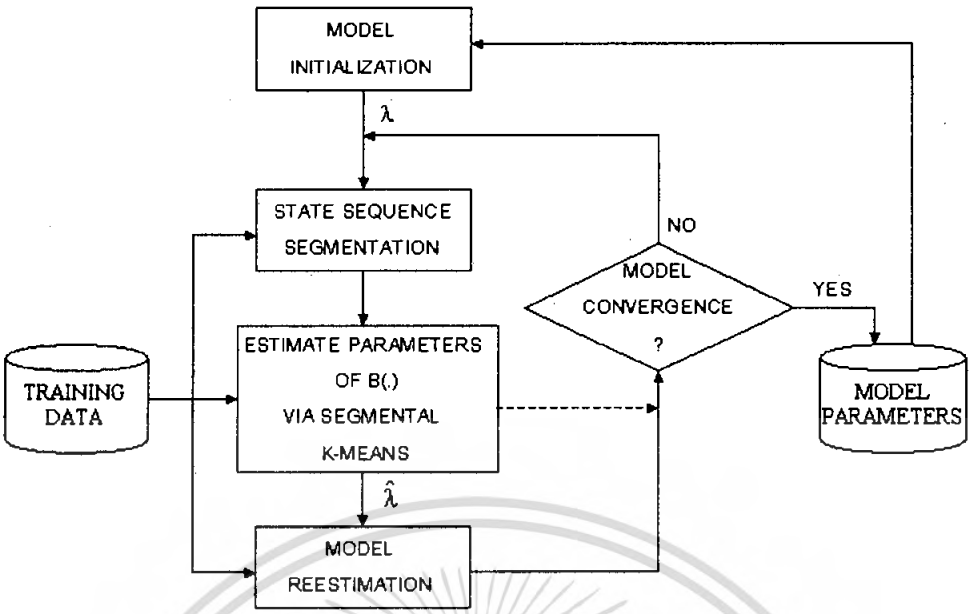
วิธีการที่ได้กล่าวมานี้จริงๆแล้วก็คือการค้นหาเส้นทางเดินที่ดีที่สุดบนกราฟ โดยที่ในแต่ละโหนดของกราฟจะเป็นสเตท ของ รูปแบบของมาร์คอฟ ในแต่ละช่วงเวลา t, $1\le t\le T$

2.2.7.6 ปัญหาการเรียนรู้

โดยทั่วไปแล้วปัญหาเกี่ยวกับการเรียนรู้ของรูปแบบมาร์คอฟ คือ จะทำอย่างไรจึงจะปรับค่าของตัวแปรของแบบจำลอง ให้สามารถอธิบายลักษณะของชุดข้อมูลฝึกหัดที่เราได้ให้เข้าไปในแบบจำลองเพื่อฝึกหัดตามแต่การประยุกต์ใช้งาน

สำหรับในโครงงานนี้เราได้ใช้วิธีการ แบ่งเสียงของเดให้กลายเป็นสเตท ซึ่งจะพิจารณาถึงการเริ่มต้นของค่า $b_j(o_i)$ เพื่อที่จะทำให้การฝึกหัดระบบสามารถทำได้อย่างมีประสิทธิภาพ ซึ่งในวิธีการนี้จะมีขั้นตอนดังรูปที่ 2.6 กระบวนการในการฝึกหัดเป็นวิธีการที่เป็นที่รู้จักกันในนามของกระบวนการหา K-Means สำหรับการรวมกลุ่มข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.6 กระบวนการฝึกหัด วิธีแบ่งเสียงของเค ซึ่งถูกใช้ในการหาค่าของตัวแปร เพื่อประมาณค่าส่วนผสมที่ต่อเนื่องให้เหมาะสมกับชุดข้อมูลฝึกหัดที่กำหนดให้

กระบวนการเริ่มต้นของแบบจำลองเริ่มจาก เซตของข้อมูลฝึกหัดจะถูกหาลำดับ สเตทที่เหมาะสมในปัจจุบัน โดยใช้แบบจำลองปัจจุบันเป็นฐาน ซึ่งการหาลำดับสเตท นี้จะใช้ ลำดับขั้นตอนของวิเตอร์บิ จากนั้นจะใช้กระบวนการ แบ่งเสียงของเค ซึ่งจะทำการแบ่งเวคเตอร์ชุดข้อมูลที่สังเกตในแต่ละสเตท ไปเป็น M ส่วนย่อยๆ โดยใช้การวัดความคลาดเคลื่อนของยูคลิด โดยในแต่ละส่วนย่อยๆ นั้นจะแสดงถึงแต่ละส่วนผสมของ $b_j(o_t)$ ซึ่งเราสามารถปรับค่าของตัวแปรของแบบจำลองได้ดังต่อไปนี้

$$\hat{c}_{jm} = \frac{\text{จำนวนของเวกเตอร์ที่แบ่งประเภทในกลุ่ม m ของสเตท j}}{\text{จำนวนของเวกเตอร์ในสเตท j}} \qquad (2.50)$$

$\hat{\mu}_{jm}$ = ตัวอย่างความหมายของเวกเตอร์ที่แบ่งประเภทในกลุ่ม m ของสเตท j (2.51)

$\hat{U}_{jm}$ = ตัวอย่างเมทตริกโควาเรียสของเวกเตอร์ที่แบ่งประเภทใน กลุ่ม m ของสเตท j (2.52)

และนอกจากนี้ในส่วนของความน่าจะเป็นในการเปลี่ยน State เราก็สามารถที่จะปรับค่าโดยใช้วิธีนี้ได้ โดยจะได้ว่า

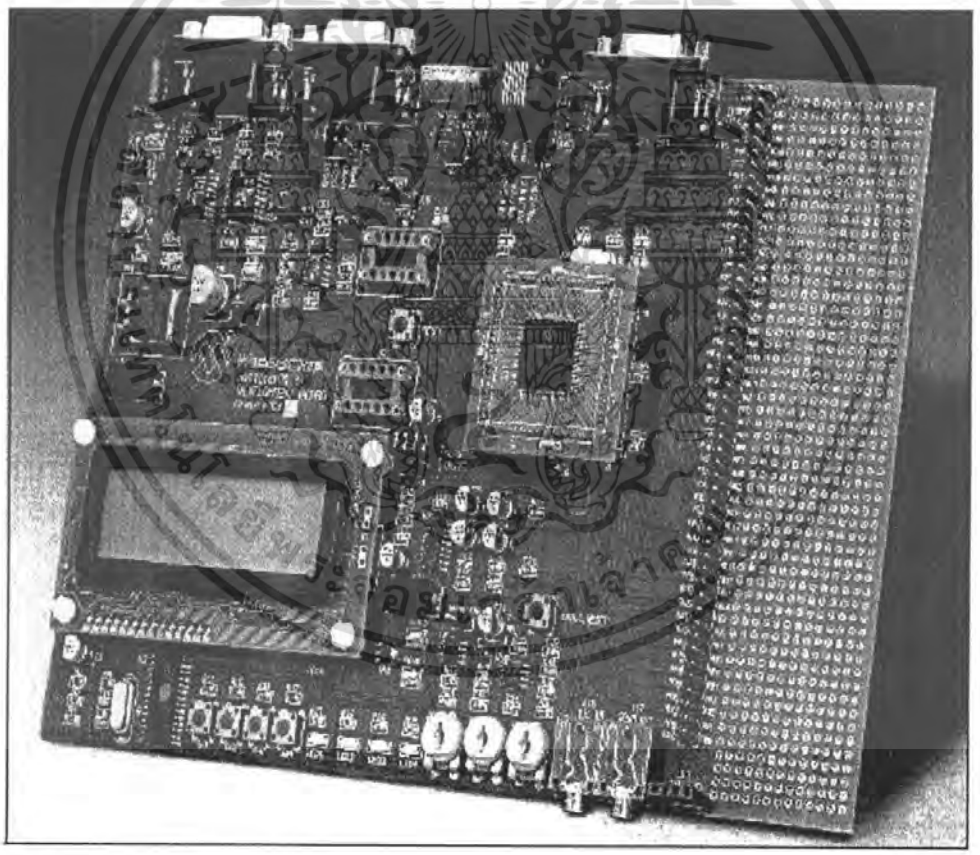
$$a_{ij} = \frac{\text{จำนวนของการส่งผ่านจากสเตท i ไปสเตท j}}{\text{จำนวนของการส่งผ่านจากสเตท i}} \qquad (2.53)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อทำการปรับปรุงแบบจำลองเรียบร้อยแล้ว แบบจำลองที่ได้มาใหม่ก็จะถูกเปรียบเทียบกับแบบจำลองเก่า ด้วยการหาคะแนนความพัฒนา ซึ่งถ้าความพัฒนานี้มีค่ามากกว่าระดับที่ยอมรับได้ก็จะทำการแทนแบบจำลองเก่าด้วยแบบจำลองใหม่ แต่ในทางกลับกันถ้าการพัฒนาน้อยกว่าระดับที่ยอมรับได้ ก็จะถือว่าแบบจำลองนั้นไม่สามารถฝึกหัดต่อไปได้อีกแล้ว ก็จะทำการบันทึกค่าสุดท้ายของแบบจำลองเอาไว้

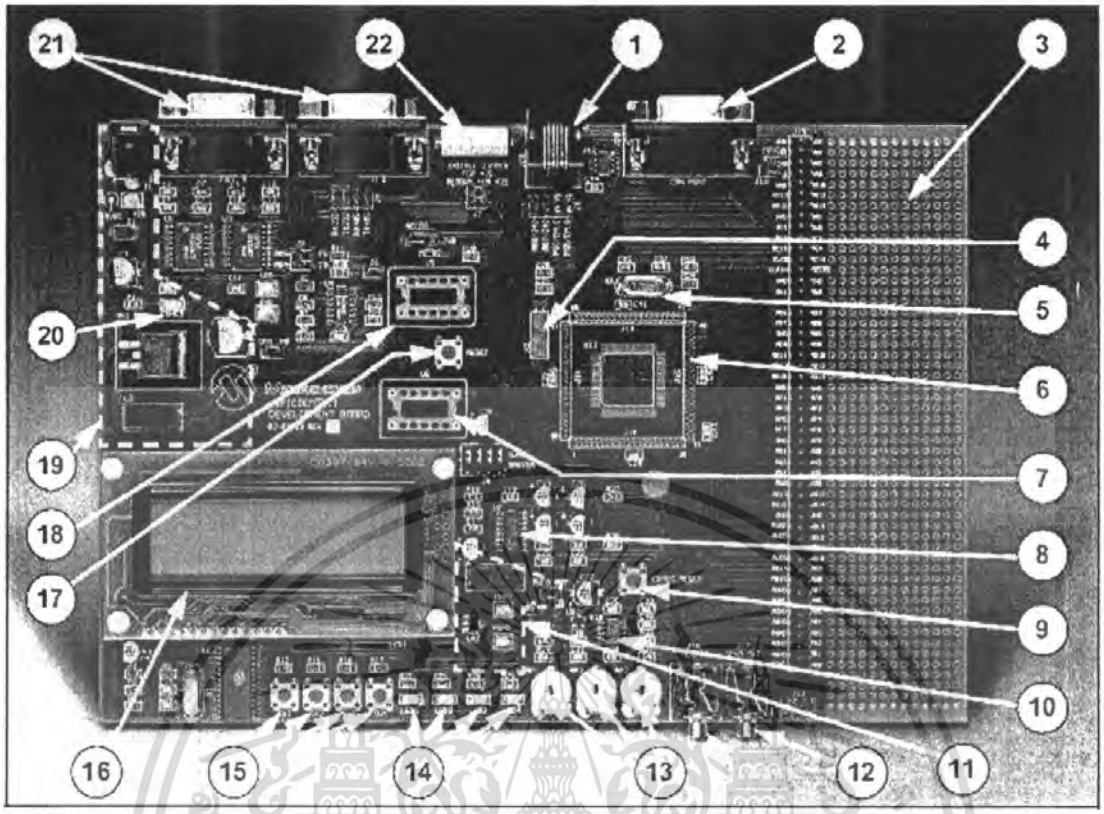
2.3 บอร์ด dsPICdem

ในการทำโปรเจคครั้งนี้เราเลือกบอร์ด dsPICdem ดังรูปที่ 2.7 เนื่องจากมี dsPIC30F6014 เป็นตัวประมวลผล ซึ่งทางผู้จัดทำโปรเจคเห็นว่ามีความสามารถเพียงพอในการประมวลผลสัญญาณ ดิจิตอล สำหรับกระบวนการรู้จำเสียง



รูปที่ 2.7 บอร์ดที่ใช้ในการทดลอง dsPICdem

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.8 อุปกรณ์ต่าง ๆ บนบอร์ด dsPICdem 1.1

ตารางที่ 2.1 แสดงรายชื่ออุปกรณ์ต่าง ๆ บนบอร์ด dsPICdem 1.1

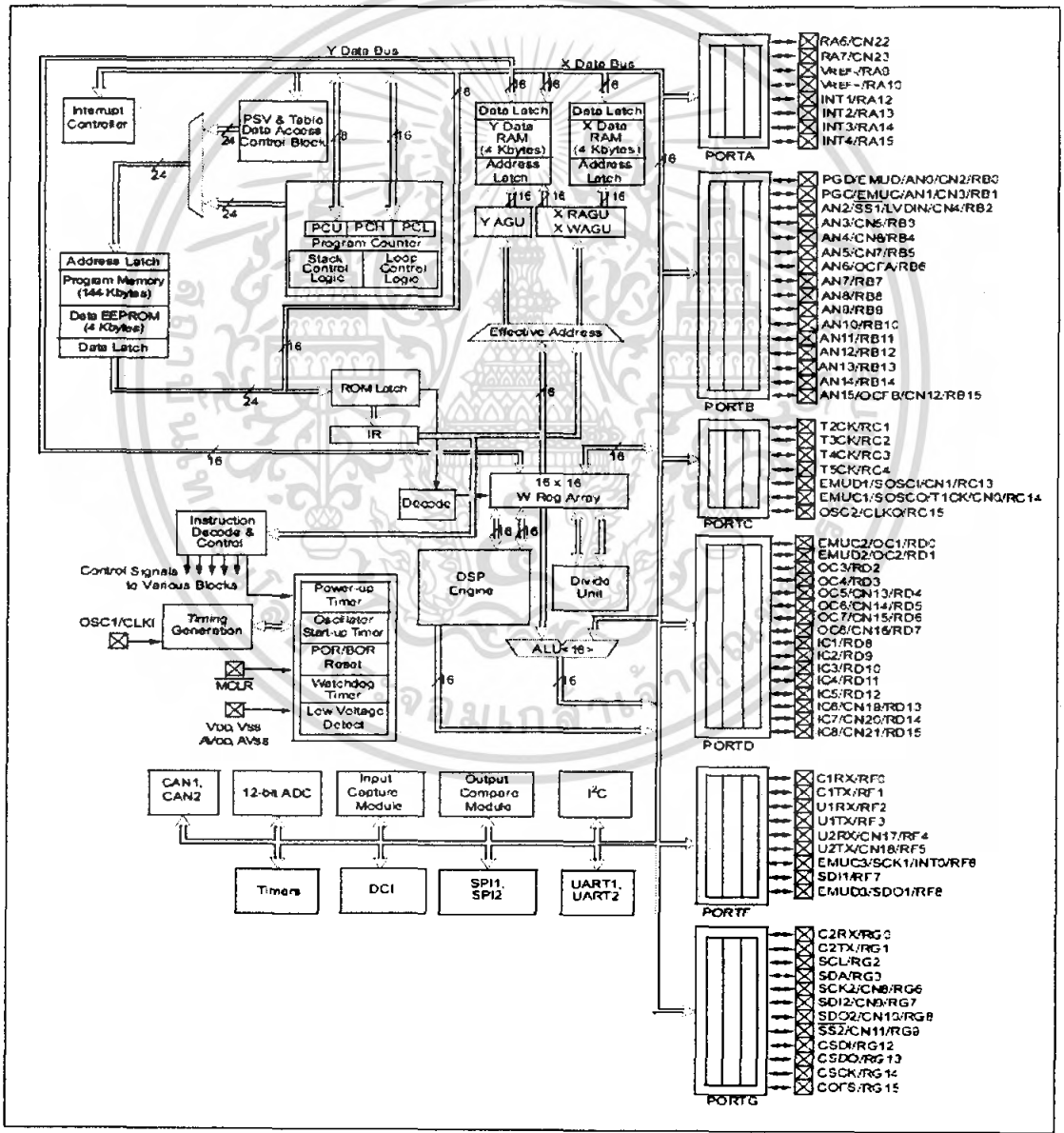
หมายเลข	ส่วนประกอบฮาร์ดแวร์	หมายเลข	ส่วนประกอบฮาร์ดแวร์
1	จุดเชื่อมต่อไอซีดี	12	ช่องเสียบโฟน
2	พอร์ทซีเอเอ็น	13	ตัวต้านทานปรับค่าได้
3	พื้นที่อเนกประสงค์	14	แอลอีดี
4	ออสซิลเลเตอร์ X2	15	สวิตซ์กดติดปล่อยดับ
5	ออสซิลเลเตอร์ X3	16	จอแสดงผลแอลซีดี
6	หัวเสียบอีมูเลชั่น	17	สวิตซ์รีเซต
7	ช่องเสียบสัญญาณนาฬิกาภายนอกเอสไอ3000	18	ช่องเสียบคริสตัล
8	เอสไอ3000โคเดค	19	วีดีดีเรกกูเลเตอร์
9	สวิตซ์รีเซตโคเดค	20	ไฟแสดงสถานะเครื่อง
10	เครื่องมือวัดอุ	21	พอร์ทอนุกรมอาร์เอส-232
11	เอวีดีดีเรกกูเลเตอร์	22	พอร์ทอาร์เอส-485และ422

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4 dsPIC30F6014

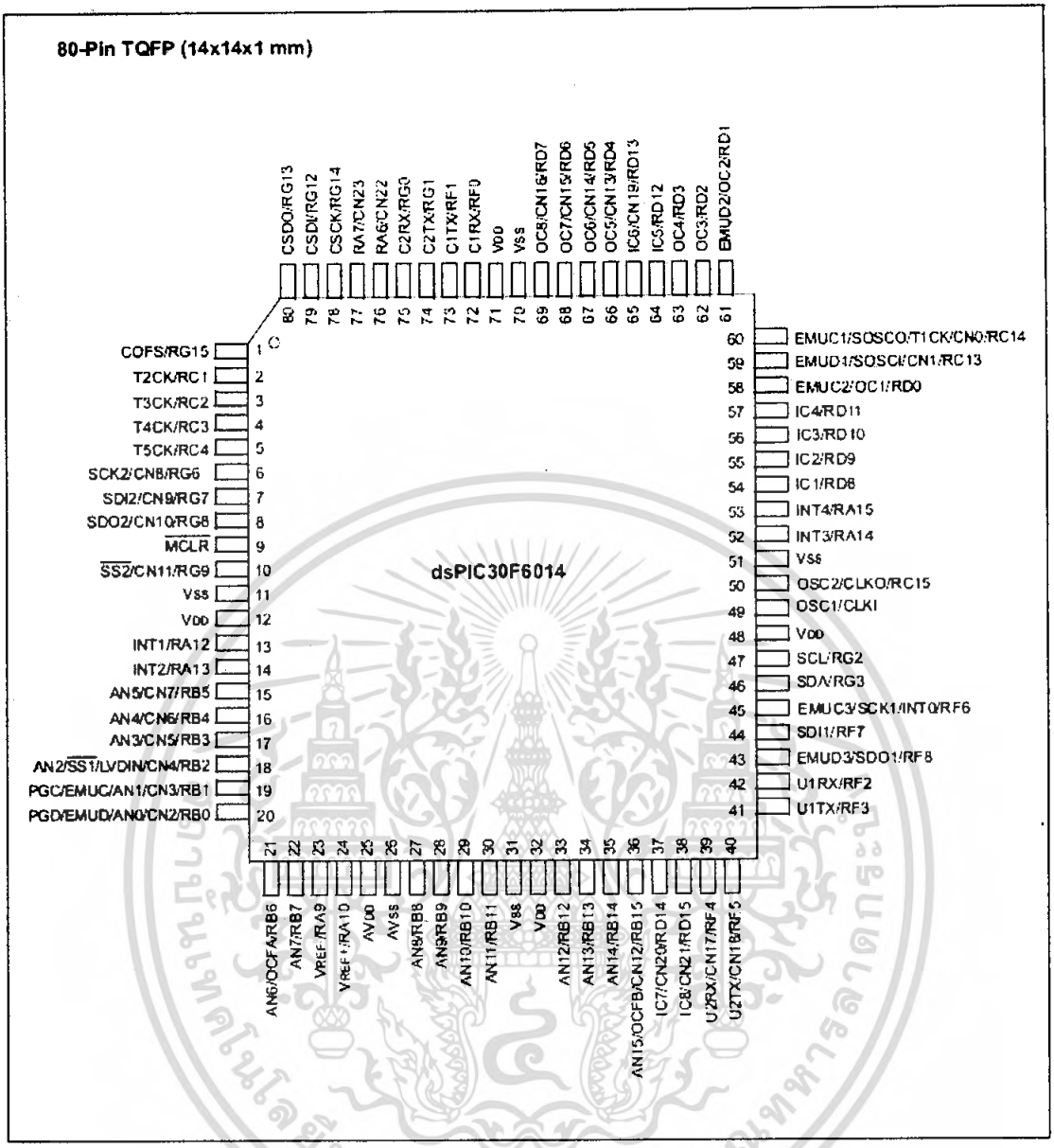
dsPIC คือชื่อของไมโครคอนโทรลเลอร์ 16 บิตจาก Microchip Technology Inc. ผู้ผลิตไมโครคอนโทรลเลอร์ PIC ซึ่งรู้จักกันเป็นอย่างดีในแวดวงนักพัฒนาระบบไมโครคอนโทรลเลอร์ โดย Microchip Technology ได้กำหนดชื่ออย่างเป็นทางการสำหรับไมโครคอนโทรลเลอร์อนุกรมใหม่นี้ว่า Digital Signal Controller หรือ DSC นั่นหมายความว่า dsPIC เป็นไมโครคอนโทรลเลอร์ที่ได้รับการออกแบบเป็นพิเศษเพื่องานประมวลผลสัญญาณดิจิตอลสำหรับสร้างระบบควบคุมอัตโนมัติที่มีความสามารถสูง

1. คุณสมบัติของซีพียู dsPIC30F6014



รูปที่ 2.9 ไดอะแกรมการทำงานและส่วนประทั้งหมดของ dsPIC30F6014

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.10 การจัดขาใช้งานของไมโครคอนโทรลเลอร์ dsPIC30F6014

* มี 84 คำสั่งมาตรฐาน สามารถรองรับรูปแบบการอ้างแอดเดรสได้อย่างอิสระ
* ชุดคำสั่งมีขนาด 24 บิต สามารถประมวลผลข้อมูลได้ 16 บิต
* มีหน่วยความจำโปรแกรมแบบแฟลชความจุ 196 กิโลไบต์ ลบและเขียนใหม่ได้ไม่น้อย กว่า 10,000 ครั้ง สามารถป้องกันการอ่านได้
* สามารถโปรแกรมหน่วยความจำโปรแกรมด้วยตัวเองโดยใช้กระบวนการทางซอฟต์แวร์
* มีหน่วยความจำข้อมูลอีอีพรอม 4 กิโลไบต์ ลบและเขียนใหม่ได้ไม่น้อยกว่า 100,000 ครั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

* มีหน่วยความจำข้อมูลแรม 8 กิโลไบต์
* รีจิสเตอร์ W จัดในรูปของอะเรย์ มีขนาด 16 บิต จำนวน 16 ตัว
* ความเร็วในการทำงานสูงถึง 30 ล้านคำสั่งต่อวินาที
* ความถี่สัญญาณนาฬิกาจากภายนอก ตั้งแต่ย่านไฟตรงจนถึง 40 MHz
* ความถี่สัญญาณนาฬิกาในกรณีใช้งานร่วมกับวงจรเฟลล็อกลูปภายใน ตั้งแต่ 4MHz ถึง 10 MHz เลือกได้ 3 ระดับคือ 4 , 8 หรือ 16 เท่า
* รองรับแหล่งกำเนิดสัญญาณอินเตอร์รัปต์ได้สูงสุด 41 แหล่ง
* สามารถกำหนดระดับความสำคัญในการตอบสนองอินเตอร์รัปต์ได้ 8 ระดับ
* มีอินเตอร์รัปต์เวกเตอร์ 62 ตำแหน่ง
* มีวงจรตรวจจับแรงดันไฟเลี้ยงต่ำกว่ากำหนดแบบโปรแกรมได้
* มีเพาเวอร์-ออนรีเซต , เพาเวอร์-อัปไทเมอร์ และออสซิลเลเตอร์สตาร์ต-อัปไทเมอร์
* มีวอตช์ด็อกไทเมอร์แบบโปรแกรมได้
* มีวงจรตรวจสอบการทำงานของวงจรกำเนิดสัญญาณนาฬิกา หากผิดพลาดจะเข้าสู่โหมดสัญญาณนาฬิกา RC พลังงานต่ำทันที
* รองรับการโปรแกรมในวงจรแบบอนุกรม (ICRP : In-Circuit Serial Programming)
* สามารถเลือกโหมดการใช้พลังงานได้
* ย่านไฟเลี้ยง 2.5 ถึง 5.5V กระแสไฟฟ้า 2.6 ถึง 44mA ที่ไฟเลี้ยง +5V ขึ้นอยู่กับการกำหนดความเร็วในการทำงาน

2. คุณสมบัติด้านการประมวลสัญญาณดิจิตอล
   * มีแอกคิวมูเลเตอร์ขนาด 40 บิต 2 ตัว รองรับการประมวลผลทางคณิตศาสตร์ได้เป็นอย่างดี
   * มีหน่วยประมวลผลด้านการคูณและหารเลข 17 บิตในรูปของฮาร์ดแวร์ จึงทำให้สามารถคูณและหารเลขได้อย่างรวดเร็ว
   * ทำการคูณเลข 16 บิตได้ภายในสัญญาณนาฬิกาเพียง 1 ไซเกิล
   * มีตัวเลื่อนข้อมูลบาร์เรล 40 สเตจ ช่วยให้การประมวลผลข้อมูลที่จำนวนบิตมากๆ สามารถทำได้อย่างรวดเร็ว
   * มีวงจรเฟตช์ข้อมูลคู่ จึงทำให้สามารถประมวลผลข้อมูลได้อย่างรวดเร็ว
3. คุณสมบัติของโมดูลฟังก์ชั่นพิเศษ
   * สามารถจ่ายกระแสออกทางขาพอร์ตได้ 25 mA ทั้งแบบกระแสซิงก์และซอร์ส
   * ไทเมอร์/เคาเตอร์มีขนาด 16 บิต 5 ตัว ต่อใช้งานร่วมกันเป็นไทเมอร์ 32 บิตได้
   * มีโมดูลตรวจจับและเปรียบเทียบสัญญาณดิจิตอลขนาด 16 บิต 8 ชุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

* มีโมดูลเปรียบเทียบข้อมูลและกำเนิดสัญญาณ PWM ความละเอียด 16 บิต 8 ชุด ในการเปรียบเทียบข้อมูลสามารถเลือกการทำงานได้ 2 โหมด
* มีส่วนเชื่อมต่ออุปกรณ์อนุกรมแบบ SPI
* มีส่วนเชื่อมต่ออุปกรณ์ผ่านระบบบัส I^2C ทั้งแบบ 7 และ 10 บิต กำหนดเป็นมาสเตอร์หรือสเลฟได้
* มีโมดูลสื่อสารข้อมูลอนุกรม UART พร้อมบัฟเฟอร์แบบ FIFO
* มีโมดูลสร้างสัญญาณ PWM สำหรับควบคุมมอเตอร์ 8 ช่อง
    - เลือกรูปแบบเอาต์พุตได้ทั้งแบบคอมพลีเมนต์และแบบอิสระ
    - มีโหมดปรับตำแหน่งการหมุนทั้งแบบปรับขอบสัญญาณและแบบกึ่งกลาง
    - มีส่วนกำเนิดดิวตี้ไซเกิล 4 ชุด
    - กำหนดฐานเวลาได้ 4 โหมด
    - สามารถเลือกขั้วของสัญญาณทางเอาต์พุตได้
    - มีสัญญาณกระตุ้นเพื่อให้ทำงานสัมพันธ์กับวงจรแปลงสัญญาณอนาลอกเป็นดิจิตอลภายในไมโครคอนโทรลเลอร์
    - สามารถควบคุมสัญญาณเอาต์พุตได้
* มีโมดูลเชื่อมต่อตัวเข้ารหัสแบบควอดราเจอร์
    - มีอินพุต Phase A , Phase B และรับสัญญาณพัลส์เพื่อกำหนดตำแหน่ง
    - มีตัวนับตำแหน่งขนาด 16 บิต นับได้ทั้งขึ้นและลง
    - แสดงสถานะของทิศทางการนับได้
    - กำหนดโหมดของการวัดตำแหน่งได้ 2 โหมดคือ x2 และ x4
    - มีวงจรกรองสัญญาณรบกวนแบบดิจิตอลจากอินพุตแบบโปรแกรมได้
    - สำหรับกำหนดให้ทำงานเป็นไทเมอร์/เคาเตอร์ขนาด 16 บิตได้
    - กำเนิดสัญญาณอินเตอร์รัปต์จากตำแหน่งที่นับเกิน (rollover) หรือนับขาด (underflow)
* มีวงจรแปลงสัญญาณอนาลอกเป็นดิจิตอล ความละเอียด 12 บิต 16 ช่อง

2.4.1 สถาปัตยกรรมโดยสรุปของ dsPIC30F6014

1. หน่วยประมวลผลกลาง

หน่วยประมวลผลของ dsPIC30F6014 ใช้คำสั่งที่มีความยาว 1 เวิร์ด ขนาด 24 บิต โดยมีโปรแกรมเคาเตอร์ขนาด 23 บิต (จริงๆ แล้วโดยโครงสร้างมี 24 บิต แต่ไม่สนใจบิต MSB ซึ่งก็คือบิต 23 และบิต LSB หรือบิต 0 กำหนดให้เป็น "0" จึงทำให้สามารถติดต่อหน่วยความจำโปรแกรมได้สูงสุด 4 เมกะเวิร์ด) เพื่อแจ้งแอดเดรสของหน่วยความจำโปรแกรมที่เข้าไปประมวลผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

dsPIC30F4016 มีความจุของหน่วยความจำโปรแกรม 144 กิโลไบต์ เมื่อคำสั่งมีความยาว 24 บิต จึงบรรจุคำสั่งได้จริง 16 กิโลเวิร์ด

รีจิสเตอร์หลักที่ใช้ในการทำงานคือ รีจิสเตอร์ W (Working register) สำหรับใน dsPIC จะแตกต่างจากไมโครคอนโทรลเลอร์ PIC อย่างมาก โดยรีจิสเตอร์ W ได้รับการจัดโครงสร้างเป็นอะเรย์ขนาด 16 บิต จึงทำให้สามารถรองรับทั้งข้อมูล, ค่าแอดเดรส หรือค่าของรีจิสเตอร์ใดๆ ที่ต้องนำมาประมวลผลโดยใน dsPIC มีรีจิสเตอร์ W ให้ใช้งานถึง 16 ตัว ส่วนใหญ่ใช้ในการประมวลผลหลัก ส่วนอีกตัวหนึ่งคือรีจิสเตอร์ W15 จะใช้ทำงานร่วมกับตัวชี้สแต็กในการทำงานของโปรแกรมย่อยและบริการอินเตอร์รัปต์

ด้านการตอบสนองอินเตอร์รัปต์นั้น dsPIC30F6014 มีการจัดสรรพื้นที่เก็บค่าอินเตอร์รัปต์เวกเตอร์ไว้มากถึง 61 ตำแหน่ง และยังสามารถกำหนดระดับความสำคัญได้อีก 8 ระดับด้วย

2. หน่วยความจำ

dsPIC30F6014 มีหน่วยความจำโปรแกรม 16 กิโลเวิร์ด แอดเดรสอยู่ในช่วง 0x000100 ถึง 0x007FFE สามารถโปรแกรมหรือเขียนข้อมูลลงไปได้ 2 วิธีคือ

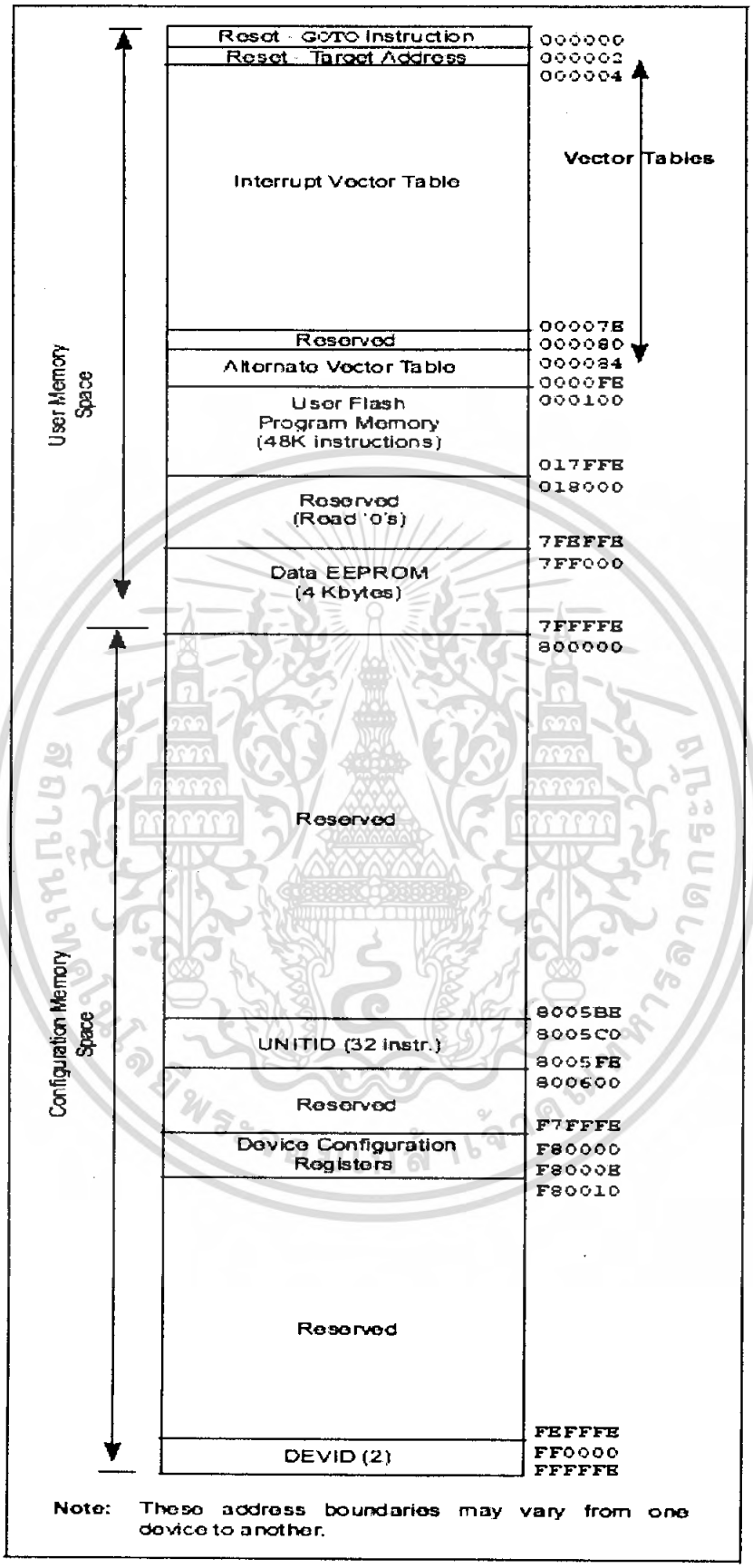
(1). โดยใช้การโปรแกรมในวงจรแบบอนุกรมหรือ ICSP ผ่านทางขา PGD และ PGC แล้วป้อนสัญญาณพัลส์แรงดันสูงสำหรับโปรแกรมผ่านเข้ามาทางขา MCLR

(2). โปรแกรมตัวเองในขณะทำงานหรือ RTSP (Run Time Self-Programming)

ส่วนหน่วยความจำข้อมูลแรมนั้น dsPIC30F6014 ได้จัดสรรเป็น 2 ส่วนคือ หน่วยความจำข้อมูลแรม X และ Y รวมเป็น 8192 ไบต์ โดยในแต่ละส่วนจะมีตัวกำหนดแอดเดรสแยกออกจากกัน เรียกว่า AGU (Address Generation Unit)

ในขณะที่หน่วยความจำข้อมูลอีอีพรอม dsPIC30F6014 จัดสรรไว้ที่แอดเดรส 0x7FF000 ถึง 0x7FFFFE มีความจุ 4 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

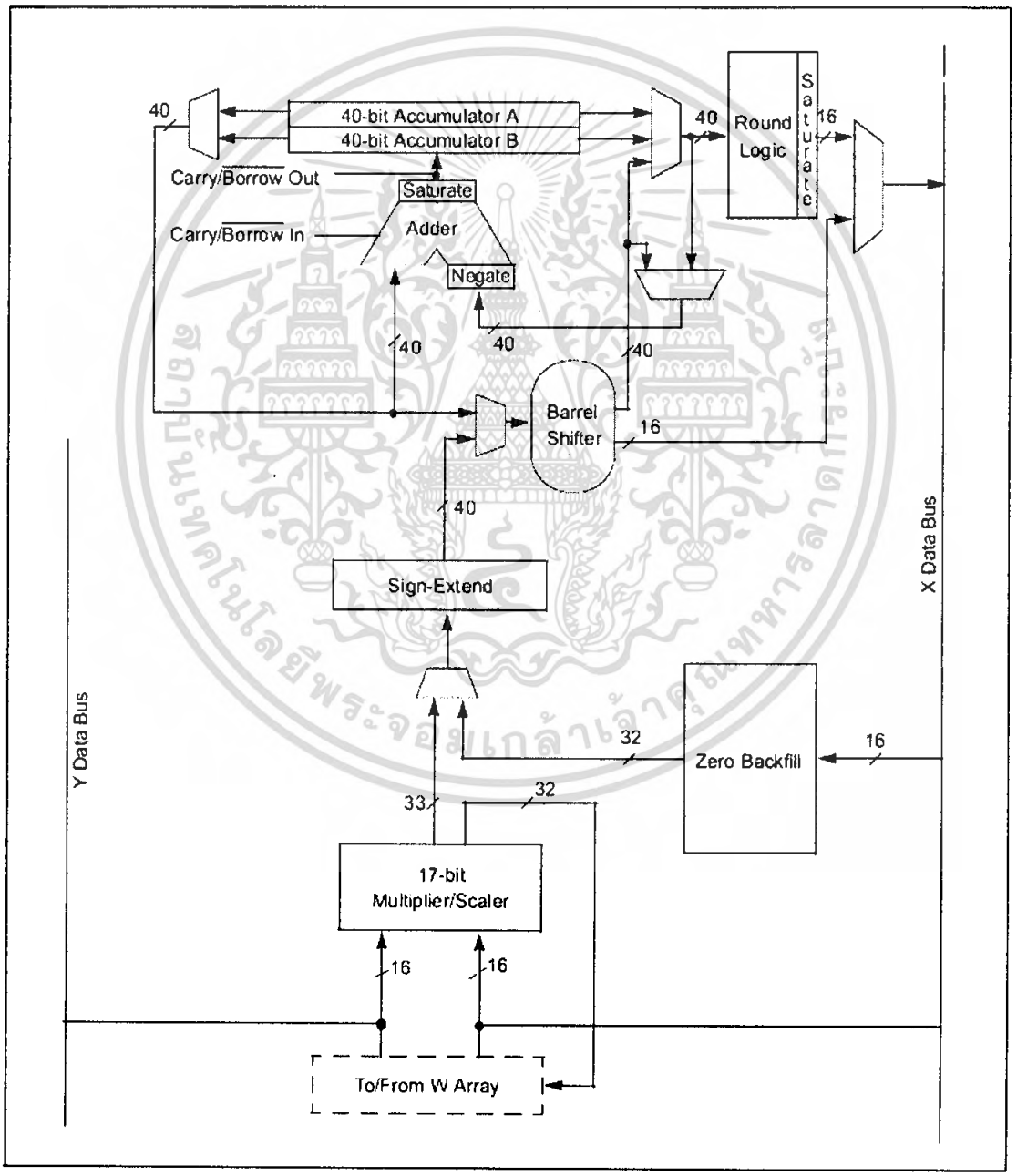


รูปที่ 2.11 การจัดสรรหน่วยความจำโปรแกรมของไมโครคอนโทรลเลอร์ dsPIC

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. ส่วนประมวลผลสัญญาณดิจิตอล (DSP Engine)

นับเป็นส่วนประกอบที่สำคัญอย่างยิ่งของ dsPIC เนื่องจาก dsPIC ได้รับการออกแบบมาให้ทำงานในด้านการประมวลผลสัญญาณดิจิตอลเป็นหลัก ดังนั้นจึงต้องมีการเพิ่มความสามรถในหน่วยคำนวณทางคณิตศาสตร์และลอจิกอย่างมาก โดยในส่วนประมวลผลสัญญาณดิจิตอลมีหน่วยจัดการคูณเลขขนาด 17 x 17 บิตความเร็วสูง, หน่วยประมวลผลทางคณิตศาสตร์และลอจิกหรือ ALU ขนาด 40 บิต, แอกคิวมูเลเตอร์ขนาด 40 บิต อีก 2 ตัว และตัวเลื่อนข้อมูล 2 ทิศทางแบบบาเรล (barrel shifter) ขนาด 40 บิต จึงทำให้สามารถจัดการข้อมูลขนาด 16 บิต ได้เสร็จสิ้นภายในสัญญาณนาฬิกาเพียงไซเกิลเดียว



รูปที่ 2.12 โครงสร้างทางโปรแกรมของไมโครคอนโทรลเลอร์

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. โมดูลฟังก์ชั่นพิเศษ

dsPIC30F6014 ได้รวมเอาโมดูลสำหรับทำงานเฉพาะทางไว้อย่างมากมาย ไม่ว่าจะเป็นโมดูลแปลงสัญญาณอนาลอกเป็นดิจิตอล ความละเอียด 12 บิต, โมดูลเชื่อมต่ออุปกรณ์อนุกรมหรือ SPI, โมดูลสื่อสารข้อมูลบนระบบบัส I^2C, โมดูลสื่อสารข้อมูลผ่านพอร์ตอนุกรมหรือ UART, ไทเมอร์ขนาด 16 บิตถึง 5 ตัว และที่เป็นพิเศษอีก 2 โมดูลคือ โมดูลสร้างสัญญาณ PWM เพื่อการควบคุมมอเตอร์และโมดูลเข้ารหัสแบบควอดราเจอร์ โดยสามารถใช้งานร่วมกันเพื่อสร้างระบบควบคุมมอเตอร์แบบปิดประสิทธิภาพสูง

โครงสร้างทางโปรแกรมที่ควรทราบ

ในรูปที่ 3.7 แสดงโครงสร้างทางโปรแกรมหรือ Programmer's model ของ dsPIC ที่ต้องทราบ ซึ่งประกอบไปด้วย

- รีจิสเตอร์ W ขนาด 16 บิตมากถึง 16 ตัว โดยที่ใช้งานเป็นหลักคือ W0 ส่วนตัวที่ถูกนำไปใช้ในส่วนประมวลผลสัญญาณดิจิตอลมี 10 ตัวคือ W4 – W13 ส่วน W14 ถูกนำไปใช้ร่วมเป็นตัวชี้เฟรม และ W15 ถูกนำไปใช้ร่มเป็นตัวชี้สแต็ก
- แอกคิวมูเลเตอร์ 40 บิต จำนวน 2 ตัว ใช้ในงานประมวลสัญญาณดิจิตอลเป็นหลัก
- โปรแกรมเคาน์เตอร์ ขนาด 24 บิต นำมาใช้ในการแจ้งแอดเดรส 23 บิต โดยไม่สนใจบิต MSB และบิต LSB ต้องเป็น " 0 "
- รีจิสเตอร์หลัก อันประกอบด้วย

  STATUS ซึ่งใช้แสดงสถานการณ์ทำงาน มีขนาด 16 บิต

  CORCON ใช้ควบคุมการทำงานของหน่วยประมวลผลกลาง มีขนาด 16 บิต

  TBLPAG เป็นรีจิสเตอร์กำหนดเพจของตารางข้อมูลในหน่วยความจำโปรแกรม มีขนาด 8 บิต

  PSVPAG เป็นรีจิสเตอร์แสดงเพจแอดเดรสของพื้นที่โปรแกรม มีขนาด 8 บิต

  RCOUNT เป็นรีจิสเตอร์เก็บค่าตัวนับจำนวนรอบของลูปที่ทำซ้ำ

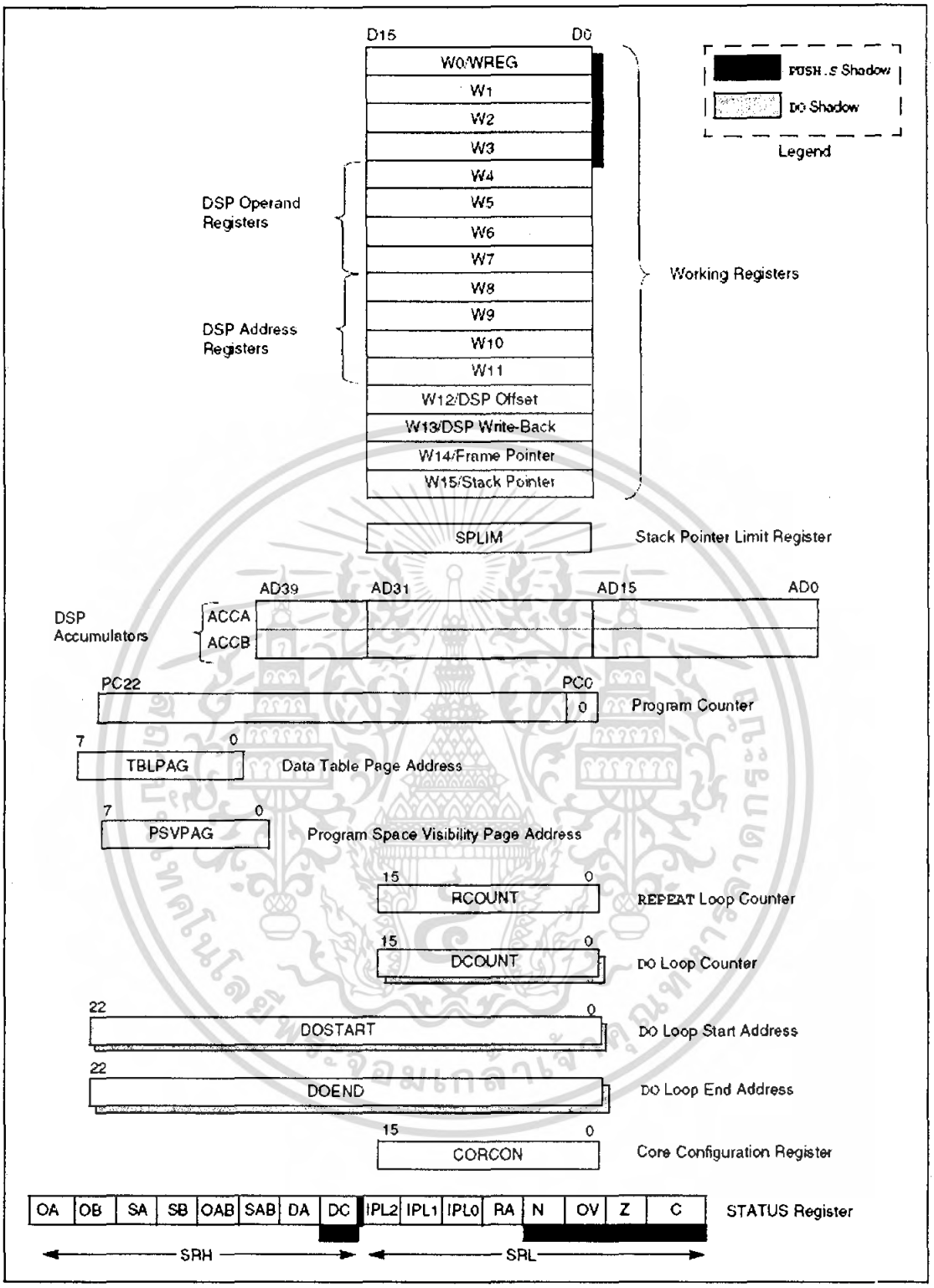
  DCOUNT เป็นรีจิสเตอร์เก็บค่าตัวนับจำนวนรอบของลูปที่ทำงาน

  DOSTART เป็นรีจิสเตอร์กำหนดแอดเดรสเริ่มต้นทำงานของโปรแกรมลูป

  DOEND เป็นรีจีสเตอร์กำหนดแอดเดรสปลายทางที่ต้องการทำโปรแกรมลูป

สำหรับรีจิสเตอร์ DCOUNT , DOSTART และ DOEND เป็นรีจิสเตอร์เงา (shadow register) หมายความว่า เป็นรีจิสเตอร์ที่ถูกสร้างขึ้นมาชั่วคราวเพื่อเก็บค่าก่อนที่จะมีการถ่ายทอดออกไปทำงาน จึงไม่สามารถเข้าถึงรีจิสเตอร์เหล่านี้ได้โดยตรง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.13 โครงสร้างทางโปรแกรมของไมโครคอนโทรลเลอร์

2.4.2 สแต็กพอยน์เตอร์และเฟรมพอยน์เตอร์ทางซอฟต์แวร์

ใน dsPIC ได้บรรจุสแต็กทางซอฟต์แวร์ขึ้นมา โดยใช้รีจิสเตอร์ W15 เป็นตัวชี้ตำแหน่งของสแต็กหรือสแต็กพอยเตอร์ (software attack pointer) ซึ่งค่าของ W15 จะเปลี่ยนแปลงอย่างอัตโนมัติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตามการกระโดดไปทำงานหรือออกจากโปรแกรมแกรมย่อย อย่างไรก็ตาม ก็ยังสามารถใช้งาน W15 เหมือนกับรีจิสเตอร์ W ตัวหนึ่งได้ด้วย แต่ถ้าไม่จำเป็นจริงๆ แนะนำให้สงวนรีจิสเตอร์ W15 นี้ไว้เพื่อทำหน้าที่เป็นตัวชี้สแต็กหรือสแต็กพอยเตอร์จะเหมาะสมกว่า

W15 จะมีค่าเท่ากับ 0x0800 หลังจากการรีเซต ผู้พัฒนาสามารถกำหนดค่าใหม่ได้ตามต้องการภายในขอบเขตของหน่วยความจำข้อมูล

นอกจากนั้นใน dsPIC ยังมีตัวชี้กรอบหรือขอบเขตของสแต็กหรือเรียกว่า เฟรมพอยน์เตอร์ (software frame pointer) ด้วย โดยใช้รีจิสเตอร์ W14 ทำหน้าที่นี้ โดยการกำหนดค่าของ W14 สามารถกระทำผ่านคำสั่ง LNK และ ULNK (คำสั่งภาษาแอสเซมบลี) เช่นเดียวกันก็ยังสามารถใช้งาน W14 เหมือนกับรีจิสเตอร์ W ตัวหนึ่งได้ด้วย แต่ควรสงวนรีจิสเตอร์ W14 นี้ไว้เพื่อทำหน้าที่เป็นเฟรมพอยน์เตอร์จะเหมาะสมกว่า

2. รีจิสเตอร์ STATUS

ใน dsPIC รีจิสเตอร์ตัวนี้มีความสำคัญมาก เพราะใช้ในการแสดงสถานะการทำงานของซีพียูโดยรีจิสเตอร์ STATUS (SR) มีขนาด 16 บิต แบ่งออกเป็น 2 ส่วนคือ SRH (บิต 8 ถึง15) และ SRL (บิต 0 ถึง 7)

โดยใน SRH บรรจุสถานะการบวกและลบค่าของข้อมูลในส่วนประมวลผลสัญญาณดิจิตอลเป็นหลักที่ 6 บิตบน ส่วนบิต DA ใช้แสดงสถานะการทำงานของกระบวนการวนทำงานหรือทำลูป (DO loop active) สุดท้ายเป็นบิตแฟลกแสดงการทดข้ามหลักหรือ Digit Carry (DC)

ส่วน SRL บรรจุสถานะการทำงานของหน่วยประมวลผลทางคณิตศาสตร์และลอจิกหรือ ALU โดยแสดงในรูปของแฟลก (flag) ที่บิต C (เกิดการทด), Z (ค่าเป็นศูนย์), 0V (ค่าเกิน), N (เกิดค่าลบ) นอกจากนั้นยังเก็บค่าของการจัดลำดับความสำคัญของการตอบสนองอินเตอร์รัปต์ไว้ที่บิต IPL0-IPL2 สุดท้ายมีบิตแสดงสถานะการทำงานของกระบวนการ REPEAT ซึ่งแสดงที่บิต RA

3. รีจิสเตอร์ควบคุมเพิ่มเติมที่เกี่ยวกับการทำงานของหน่วยประมวลผลกลางของ dsPIC ประกอบด้วย

(1). MODCON (Modulo Control Register) : เป็นรีจิสเตอร์ที่ใช้เอ็นเอเบิลและกำหนดการเข้าถึงหน่วยความจำข้อมูลแบบโมดูโล (modulo)

(2). XMODSRT และ XMODEND (X Modulo Start Address Register และ X Modulo End Address Register) : เป็นรีจิสเตอร์ที่ใช้เก็บค่าแอดเดรสตั้งต้น (XMODSRT) และปลายทาง (XMODEND) ของหน่วยความจำข้อมูล X ที่ใช้การเข้าถึงแบบโมดูโล

(3). YMODSRT และ YMODEND (Y Modulo Start Address Register และ Y Modulo End Address Register) : เป็นรีจิสเตอร์ที่ใช้เก็บค่าแอดเดรสตั้งต้น ZYMODSRT) และปลายทาง (YMODEND) ของหน่วยความจำข้อมูล Y ที่ใช้การเข้าถึงแบบโมดูโล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(4). XBREV (X Modulo Bit-Reverse Register) : เป็นรีจิสเตอร์ที่ใช้กำหนดขนาดของบัฟเฟอร์หรือหน่วยความจำข้อมูลชั่วคราวเพื่อใช้ในการเข้าถึงหน่วยความจำข้อมูลแบบกลับบิต

(5). DISICNT (Disable Interrupts Count Register) : เป็นรีจิสเตอรที่ถูกใช้ในการทำงานของคำสั่ง DISI เพื่อดิสเอเบิลการอินเตอร์รัปต์ที่มีระดับความสำคัญในระดับ 1 ถึง 6

2.4.3 หน่วยประมวลผลสัญญาณดิจิตอล

ตารางที่ 2.2 แสดงขอบเขตของข้อมูลในรีจิสเตอร์ที่มีขนาดต่างกันในไมโครคอนโทรลเลอร์ dsPIC ทั้งในแบบเลขจำนวนเต็มและเศษส่วน/ทศนิยม

ขนาดของรีจิสเตอร์	ย่านของตัวเลข กรณีเป็นเลขจำนวนเต็ม	ย่านของตัวเลข กรณีเป็นเลขเศษส่วนหรือทศนิยม	ความละเอียดของเลขเศษส่วนหรือทศนิยม
16 บิต	-32,768 ถึง 32,767	-1.0 ถึง $(1.0 - 2^{-15})$	3.052×10^{-5}
32 บิต	-2,147,483,648 ถึง 2,147,483,647	-1.0 ถึง $(1.0 - 2^{-31})$	4.657×10^{-10}
40 บิต	-549,755,813,888 ถึง 549,755,813,887	-256.0 ถึง $(256.0 - 2^{-31})$	4.657×10^{-10}

ข้อมูลที่นำมาประมวลผลในหน่วยประมวลผลสัญญาณดิจิตอล (DSP) นี้มาได้จาก 2 แหล่ง คือ

1. เข้ามาโดยตรงจากรีจิสเตอร์ W4 ถึง W7
2. จากบัสข้อมูลของหน่วยความจำข้อมูล X

ส่วนข้อมูลเอาต์พุตที่ได้จากหน่วย DSP นี้จะถูกส่งไปยัง 2 แหล่ง

1. ส่งไปยังแอกคิวมูเลเตอร์เป้าหมาย
2. หน่วยความจำข้อมูล X

ในตารางที่ 1-2 แสดงถึงขอบเขตของข้อมูลในรีจิสเตอร์ที่มีขนาดต่างกัน ทั้งในแบบเลขจำนวนเต็มและเลขเศษส่วน

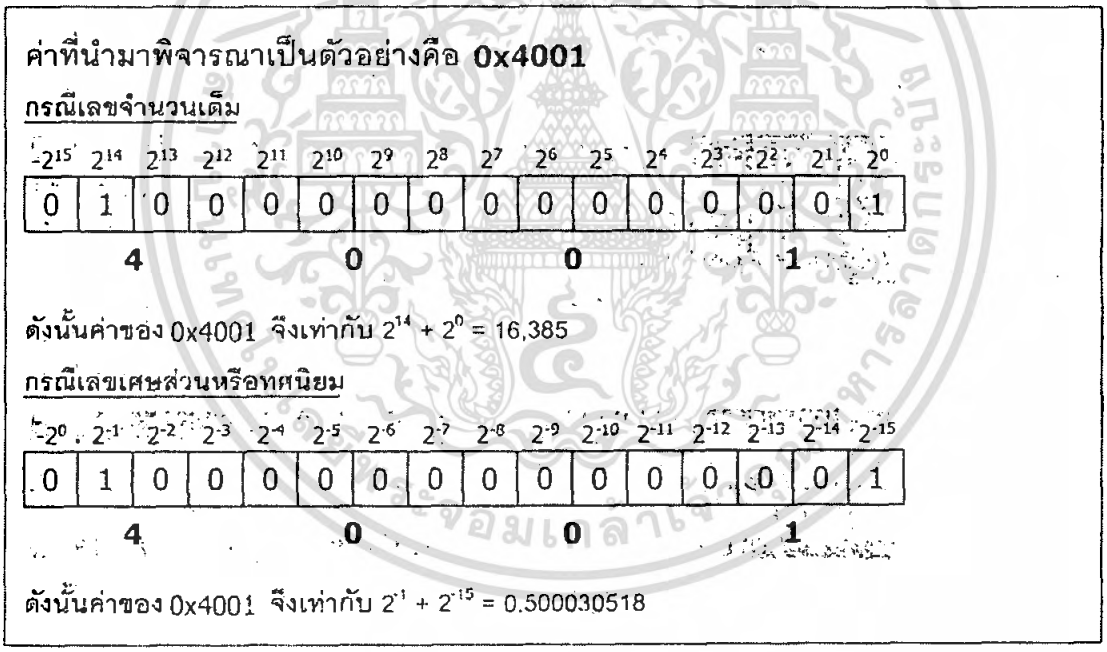
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. รูปแบบของตัวเลขที่ใช้ประมวลผลในหน่วย DSP และ ALU

ในรูปที่ 3.8 และ 3.9 แสดงให้เห็นถึงการกำหนดรูปแบบของข้อมูลเมื่อนำประมวลผลในหน่วย DSP และ ALU ทั้งในแบบเลขจำนวนเต็มและเลขเศษส่วนหรือทศนิยม

ในกรณีเลขจำนวนเต็ม จะคำนวณตามปกติ โดยบิต LSB จะมีค่าน้ำหนักประจำหลักเป็น 2^0 และที่บิตMSB หากเป็นเลขคิดเครื่องหมาย บิตนี้จะเป็นบิตแสดงเครื่องหมายว่าเป็นค่าบวกหรือลบ (ถ้าเป็น "0" หมายถึงเป็นค่าบวก ถ้าเป็น "1" หมายถึงเป็นค่าลบ) ส่วนน้ำหนักประจำหลักจะเท่ากับ 2^{15} ในกรณีเป็นเลข 16 บิต, 2^{31} กรณีเป็นเลข 32 บิต และ 2^{39} กรณีเป็นเลข 40 บิต

ในกรณีเป็นเลขเศษส่วนหรือทศนิยม บิตซ้ายสุดซึ่งปกติเป็นบิต MSB แต่ในกรณีเลขเศษส่วนหรือทศนิยมแล้ว บิตซ้ายสุดจะเป็นเลขจำนวนเต็มหลักหน่วยและถ้าเป็นเลขแบบคิดเครื่องหมายบิตนี้ก็จะใช้ในการแสดงเครื่องหมายด้วย นั่นคือ บิตซ้ายสุดจะมีค่าน้ำหนักประจำหลักเป็น 2^0 ในกรณีเป็นค่าบวก และ -2^0 ในกรณีเป็นเลขคิดเครื่องหมาย ส่วนบิตถัดมาทางขวาจะมีค่าน้ำหนักเป็น 2^{-1} ไล่เรียงไปจนถึง 2^{-15} ในกรณีเป็นเลข 16 บิต เป็นต้น



รูปที่ 2.14 แสดงตัวอย่างการดีความค่าตัวเลขข้อมูล ในกรณีที่บิต MSB เป็น "0" ทั้งในกรณีที่คิดเลขแบบจำนวนเต็มและเลขเศษส่วนทศนิยมในหน่วย DSP และแบบ ALU ของ dsPIC

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.15 แสดงตัวอย่างการตีความค่าตัวเลขข้อมูล ในกรณีที่บิต MSB เป็น "1" ทั้งในกรณีที่คิดเลขแบบจำนวนเต็มและเลขเศษส่วนทศนิยมในหน่วย DSP และแบบ ALU ของ dsPIC

2. แอกคิวมูเลเตอร์ 40 บิต

ในหน่วย DSP มีแอกคิวมูเลเตอร์ข้อมูลขนาด 40 บิต เพื่อใช้ในการคำนวณ 2 ตัวคือ AccA และ AccB โดยจัดแบ่งออกเป็น 3 ส่วนคือ

AccxL มีขนาด 16 บิต คือ แอกคิวมูเลเตอร์บิต 0 ถึง 15 (Accx0-Accx15 โดยที่ x หมายถึง A หรือ B)

AccxH มีขนาด 16 บิต คือ แอกคิวมูเลเตอร์บิต 16 ถึง 31 (Accx16-Accx31)

AccxU มีขนาด 8 บิต คือ แอกคิวมูเลเตอร์บิต 32 ถึง 39 (Accx32-Accx39)

เมื่อประมวลผลเชิงเศษส่วน แอกคิวมูเลเตอร์จะสามารถเก็บข้อมูลได้ตั้งแต่ -256.0 ถึง $256-2^{-31}$

เมื่อประมวลผลเลขจำนวนเต็ม แอกคิวมูเลเตอร์สามารถเก็บข้อมูลได้ตั้งแต่ -549,755,813,888 ถึง 549,755,813,887

3. ตัวคูณหรือมัลติพลายเออร์ 17 x 17 บิต ความเร็วสูง

เป็นส่วนประกอบที่มีบทบาทสำคัญอย่างยิ่งในหน่วย DSP และ ALU มัลติพลายเออร์นี้สามารถรองรับการคำนวณค่าตัวเลขทั้งแบบคิดและไม่คิดเครื่องหมาย รวมไปถึงข้อมูลที่เป็นเลขเศษส่วน ด้านการให้ผลลัพธ์สามารถให้ผลได้ทั้งเลขจำนวนเต็มขนาด 31 บิต และเลขเศษส่วน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มัลติพลายเออร์รับข้อมูลขนาด 16 บิตเข้ามาประมวลผล จากนั้นแปลงเป็นข้อมูล 17 บิต โดยการเพิ่มบิตเครื่องหมายเข้าไปเพื่อช่วยให้การคำนวณสามารถทำได้ง่ายและรวดเร็วขึ้น ด้วยการกำหนดสถานะที่บิต IF (บิต 0 ของรีจิสเตอร์ (CORCON) จะเป็นตัวกำหนดการทำงานว่า ต้องการให้ประมวลผลในรูปของเลขจำนวนเต็มหรือเศษส่วน นอกจากนั้นบิต IF ยังไม่ส่งผลกระทบใดๆ ต่อคำสั่งการคูณ 5 คำสั่งของ dsPIC คือ

MUL กับ MUL.UU	เป็นคำสั่งคูณเลขจำนวนเต็ม 2 จำนวนแบบไม่คิดเครื่องหมาย
MUL.SS	เป็นคำสั่งคูณเลขจำนวนเต็ม 2 จำนวนแบบคิดเครื่องหมาย
MUL.SU กับ MUL.US	เป็นคำสั่งคูณเลขจำนวนเต็มคิดเครื่องหมายกับเลขจำนวนเต็มที่ไม่คิดเครื่องหมาย

อย่างไรก็ตาม มัลติพลายเออร์นี้จะได้รับการกำหนดให้ทำงานกับเลขเศษส่วนเป็นหลักหลังจากเกิดการรีเซตไมโครคอนโทรลเลอร์

การให้ผลลัพธ์ของมัลติพลายเออร์เป็นดังนี้

(1). เมื่อทำงานในโหมดเลขจำนวนเต็ม จะให้ผลลัพธ์เป็นค่าทูคอมพลีเมนต์แบบคิดเครื่องหมาย (signed two's complement value) โดยบิตนัยสำคัญสูงสุดหรือ MSB ใช้แสดงเครื่องหมาย ดังนั้นค่าที่ได้จึงอยู่ระหว่าง -2^{N-1} ถึง $2^{N-1}-1$

(2). เมื่อทำงานในโหมดเลขเศษส่วน จะให้ผลลัพธ์เป็นค่าทูคอมพลีเมนต์ของเลขเศษส่วนแบบคิดเครื่องหมายเช่นกัน โดยบิตนัยสำคัญสูงสุดหรือ MSB ใช้แสดงเครื่องหมาย ค่าที่ได้จะอยู่ระหว่าง -1.0 ถึง $(1-2^{1-N})$

4. ตัวเลื่อนข้อมูลบาร์เรล (Barrel Shifter)

เป็นตัวเลื่อนข้อมูลขนาด 40 บิตจึงทำให้สามารถจัดการเลื่อนข้อมูลขนาด 16 บิต ไปทางซ้ายหรือขวาได้ภายในสัญญาณนาฬิกาเพียงไซเกิลเดียว นับเป็นคุณสมบัติที่โดดเด่นอีกประการหนึ่งของหน่วยคำนวณใน dsPIC อย่างไรก็ตามตัวเลื่อนข้อมูลนี้ต้องการข้อมูลที่มีการกำหนดเครื่องหมายเพื่อช่วยให้ทราบจำนวนบิตและทิศทางการเลื่อนข้อมูล ดังมีรายละเอียดต่อไปนี้

(1). ถ้าเป็นค่าบวก	จะเลื่อนข้อมูลไปทางขวา
(2). ถ้าเป็นค่าลบ	จะเลื่อนข้อมูลไปทางซาย
(3). ถ้าค่าเป็น "0"	จะไม่มีการทำงานใดๆ

ตัวเลื่อนข้อมูลจะให้ผลลัพธ์ออกมา 40 บิตเมื่อทำงานในหน่วย DSp และให้ผลลัพธ์ 16 บิเมื่อทำการคำนวณในหน่วย ALU ของไมโครคอนโทรลเลอร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.4 การรองรับการหารเลขใน dsPIC

คุณสมบัติเด่นอีกประการหนึ่งของไมโครคอนโทรลเลอร์ dsPIC คือ การหารเลข (divide) ในหน่วยคำนวณของไมโครคอนโทรลเลอร์ทั่วไปมักจะไม่รองรับการหารเลข ทำให้ต้องสร้างโปรแกรมย่อยขึ้นมาจัดการคำนวณในส่วนนี้เอง หรือถ้าจะรองรับก็เป็นแบบพื้นฐาน ทั้งยังต้องใช้เวลาในการคำนวณนานพอควร แต่สำหรับ dsPIC แล้วสามารถรองรับการหารเลขได้อย่างเต็มรูปแบบ โดยแบ่งเป็น 5 รูปแบบคือ

1. การหารเลขเศษส่วนหรือทศนิยม 16 บิตแบบคิดเครื่องหมาย โดยใช้คำสั่ง DIVF
2. การหารเลขตัวตั้ง 32 บิตด้วยตัวหาร 16 บิต แบบคิดเครื่องหมาย โดยใช้คำสั่ง DIV. SD
3. การหารเลขตัวตั้ง 32 บิตด้วยตัวหาร 16 บิต แบบไม่คิดเครื่องหมาย โดยใช้คำสั่ง DIV. UD
4. การหารเลขจำนวนเต็ม 16 บิต แบบคิดเครื่องหมาย โดยใช้คำสั่ง DIV. SW
5. การหารเลขจำนวนเต็ม 16 บิต แบบไม่คิดเครื่องหมาย โดยใช้คำสั่ง DIV. UW

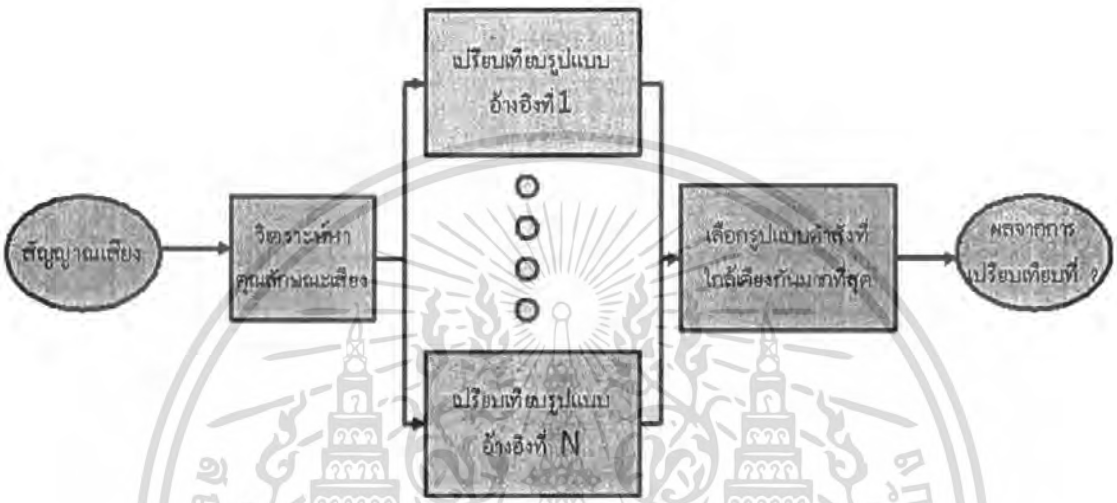
ผลหารของทุกคำสั่งหารเลขจะเก็บไว้ในรีจิสเตอร์ W0 ส่วนเศษของการหารเก็บไว้ที่รีจิสเตอร์ W1 ในขณะที่ค่าของตัวหาร 16 บิตสามารถเก็บไว้ในรีจิสเตอร์ W ตัวใดก็ได้ เช่นเดียวกับตัวตั้ง แต่ถ้าหากเป็นตัวตั้ง 32 บิตจะต้องนำค่าไปเก็บไว้ในรีจิสเตอร์ W จำนวน 2 ตัวที่อยู่ติดกัน เช่น W2 กับ W3 โดย W3 เก็บค่าตัวตั้ง 16 บิตบน และ W2 เก็บค่าตัวตั้ง 16 บิตล่าง

เวลาที่ใช้ในการคำนวณเพื่อหาค่าการนั้นจะต้องใช้มากถึง 18 เท่าของเวลาที่ใช้ในการกระทำคำสั่งปกติ สามารถใช้กระบวนการทำโปรแกรมลูปซ้ำด้วยคำสั่ง REPEAT 18 รอบ ด้วยการใช้กระบวนการ REPEAT ทำให้การเขียนโปรแกรมด้วยภาษาแอสเซมบลีทำได้ง่ายขึ้นอย่างมาก ดังนั้นรวมเวลาทั้งหมดที่ต้องใช้ในการหารเลขคือ 19 ไซเกิล

จากการศึกษาทฤษฎีที่เกี่ยวข้องกับกระบวนการรู้จำเสียงนั้น พบว่าเสียงของคนเราจะมีลักษณะของสัญญาณเสียงที่แตกต่างกันออกถึงแม้เสียงนั้นจะมาจากคำเดียวกัน หรือแม้แต่คนๆเดียวกันพูดเสียงเดียวกันยังได้สัญญาณเสียงที่แตกต่างกันอีกด้วย ซึ่งจะมีผลต่อการรู้จำเสียงด้วย ในส่วนของการรู้จำเสียงนั้นเมื่อเราได้สัญญาณเสียงแล้วเราจำเป็นที่จะต้องหาค่าตัวแทนของเสียงเนื่องจากเราไม่ได้นำสัญญาณเสียงมาประมวลผลทั้งหมดเนื่องจากจะต้องใช้เวลานานและสิ้นเปลืองหน่วยความจำ โดยเราเลือกใช้วิธี MFCC เนื่องจากเป็นวิธีที่ได้รับความนิยมสูงและให้ผลการรู้จำเสียงที่ดี หลังจากนั้นเราจะต้องทำการเปรียบเทียบตัวแทนเสียงกับเสียงต้นแบบ โดยใช้ วิธี HMM เพื่อหาว่าเสียงทดสอบเป็นเสียงใด และจากคุณสมบัติต่างของ dsPIC30F6014 นั้น มีคุณสมบัติที่เพียงพอต่อการประมวลผลสัญญาณดิจิตอล เนื่องจากกระบวนการรู้จำเสียงนั้นจะต้องมีการประมวลสูงและรวดเร็ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การคำนวณและการสร้าง



รูปที่ 3.1 แผนผังการทำงานโดยรวมของระบบรู้จำเสียง

จากรูปที่ 3.1 ในกระบวนการทำงานรู้จำเสียงนั้นเราจะเริ่มด้วยการนำสัญญาณเสียงมาวิเคราะห์หาคุณลักษณะของเสียงหรือตัวแทนของเสียง หลังจากนั้นเราจะนำไปเปรียบเทียบกับรูปแบบอ้างอิงหรือสัญญาณเสียงอ้างอิง แล้วเลือกสัญญาณเสียงอ้างอิงที่ใกล้เคียงที่สุด หลังจากนั้นเราจะแสดงผลเพื่อบอกให้รู้ว่าเสียงของผู้พูดเป็นเสียงใด เราสามารถสรุปขั้นตอนการทำงานของระบบรู้จำเสียงได้ดังนี้

3.1 การสร้างสัญญาณเสียงอ้างอิง

ในการทำการรู้จำเสียงนั้น เราจำเป็นที่ต้องมีสัญญาณเสียงอ้างอิงเพื่อเปรียบเทียบกับเสียงทดสอบ เราจะใช้โปรแกรม Word Library Builder ในการสร้างเสียงอ้างอิงโดยมีขั้นตอนการสร้างดังนี้

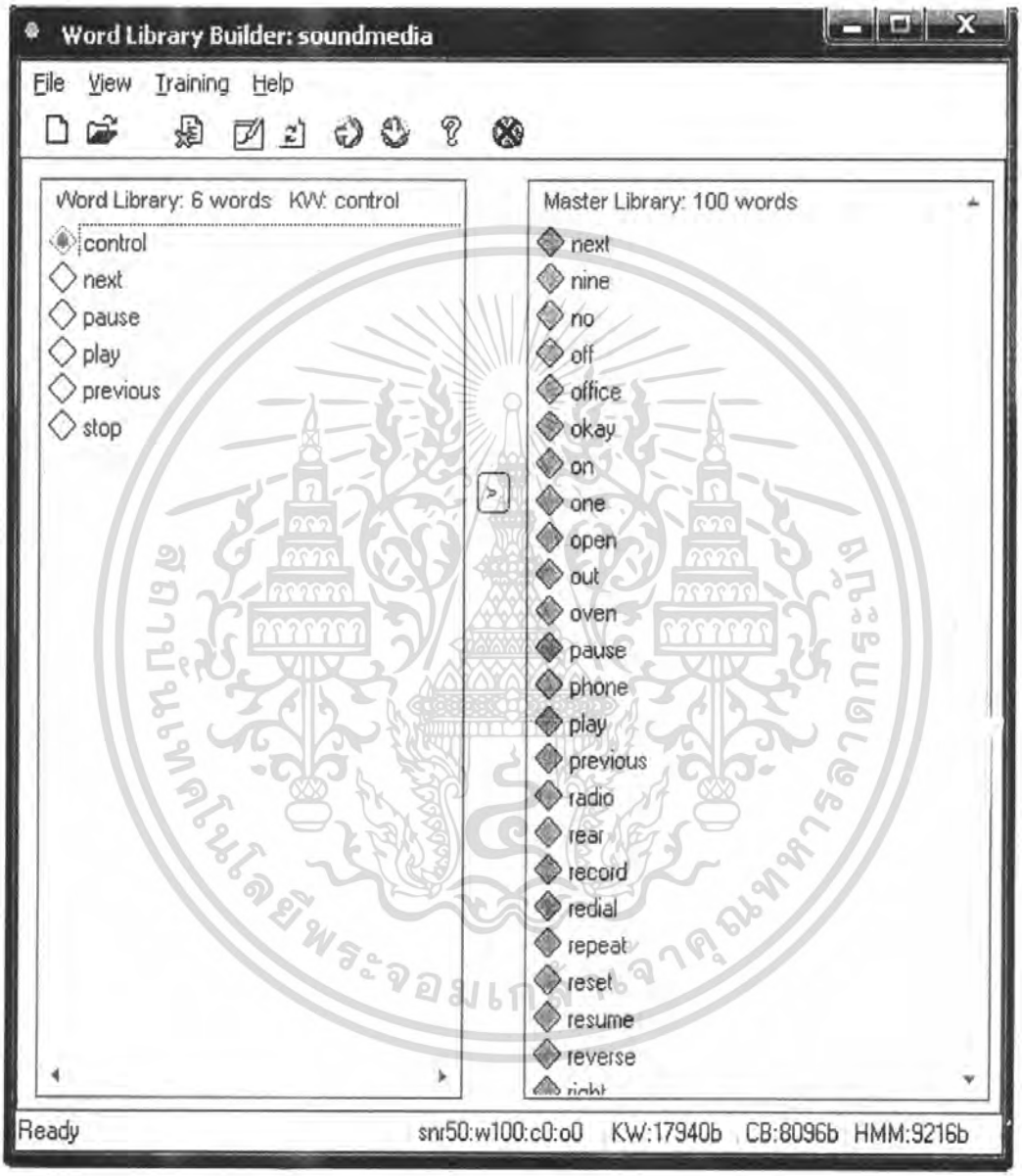
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. เปิดโปรแกรม Word Library Builder จะปรากฏดังรูปที่ 3.2 ซึ่งในโปรแกรมนี้มีสัญญาณเสียงอ้างอิงทั้งหมด 100 คำ

รูปที่ 3.2 โปรแกรม Word Library Builder

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

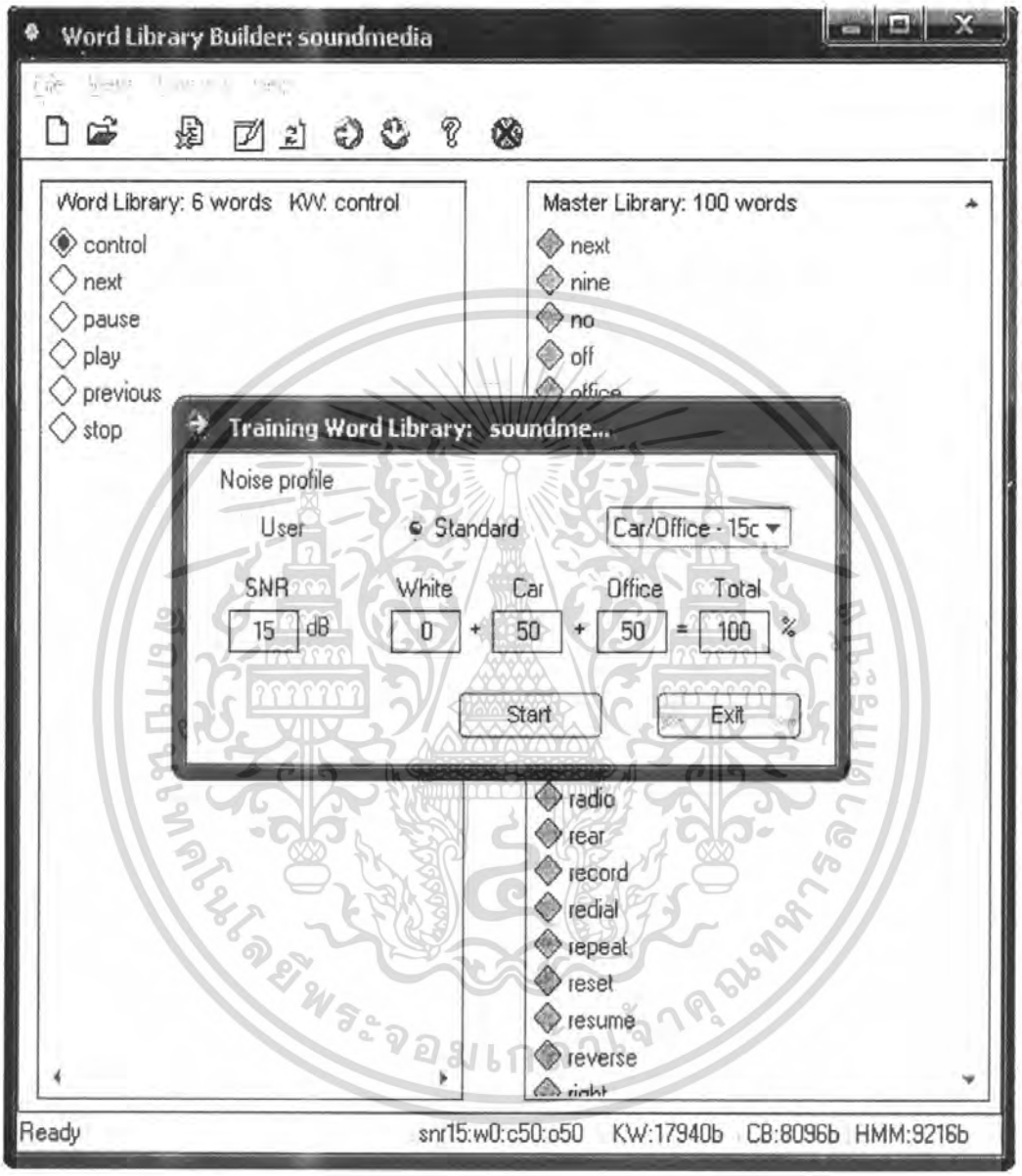
2. เลือกเสียงอ้างอิง คลิกที่ file แล้วเลือก new และเลือกคำดังนี้ control , play , pause , next , previous และ stop และดับเบิ้ลคลิกที่ control เพื่อเลือกเป็นคีย์เวิร์ดและบันทึกโดยตั้งชื่อ soundmedia ดังรูปที่ 3.3



รูปที่ 3.3 การเลือกสัญญาณเสียงอ้างอิง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

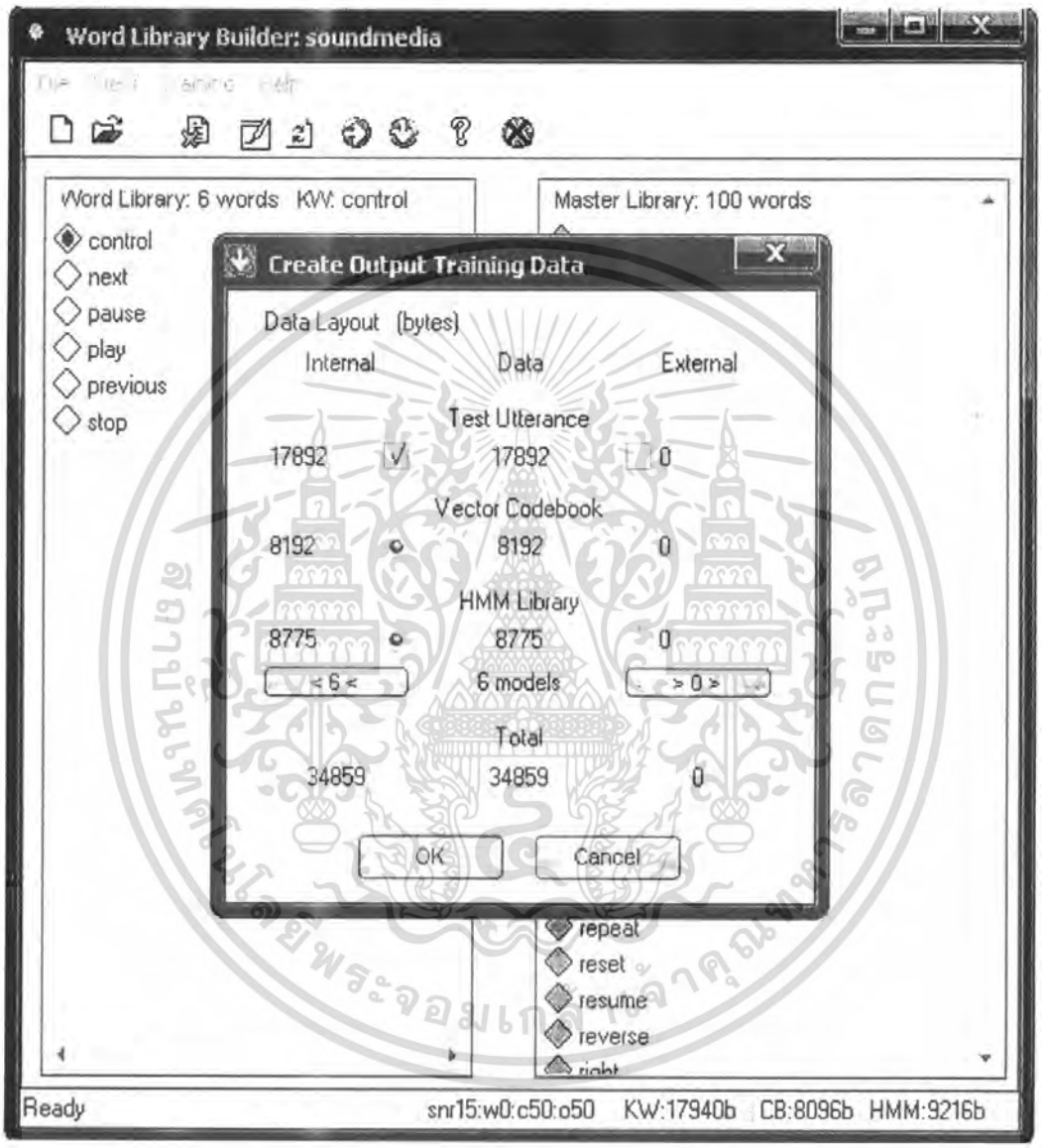
3. เริ่มกระบวนการเรียนรู้ โดยการคลิกที่ Training แล้วเลือก Train Word Library จะปรากฏดังรูปที่ 3.4 แล้วเลือกรูปแบบสัญญาณรบกวนตามต้องการ แล้วกด Start



รูปที่ 3.4 การเรียนรู้การรู้จำเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. สร้างข้อมูลสำหรับการทำการรู้จำเสียงบน dsPIC30F6014 โดยการคลิก Training แล้วเลือก Create Data และคลิก OK ดังรูปที่ 3.5 เราจะได้ ไฟล์ข้อมูล ดังนี้ soundmedia.h,soundmedia_ext.s,soundmedia_int.s และ soundmedia_voc.s



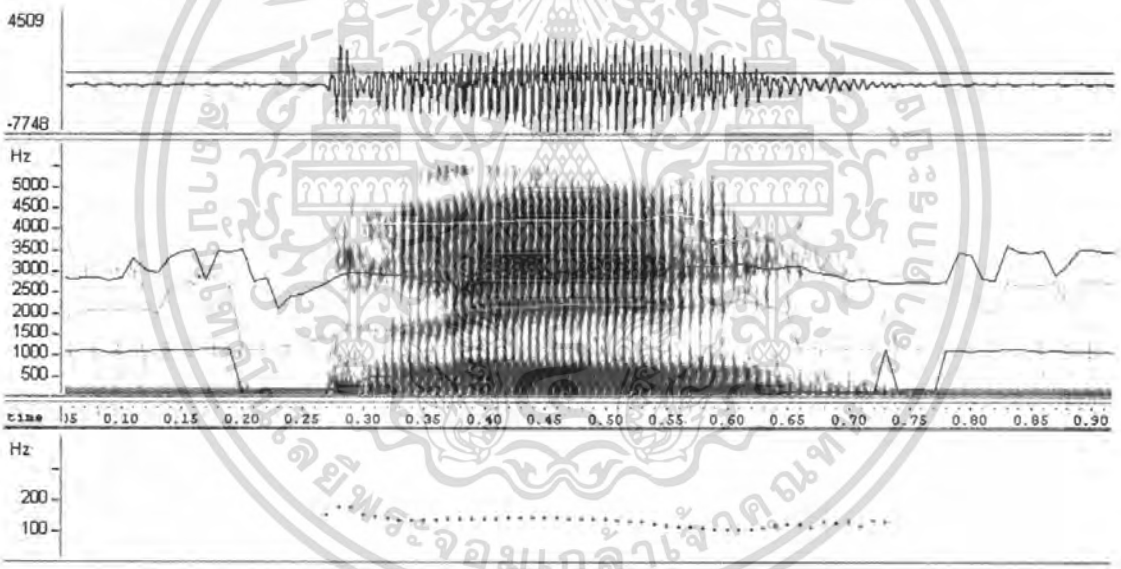
รูปที่ 3.5 การสร้างข้อมูลของเสียงต้นแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 การแปลงสัญญาณเสียงเป็นสัญญาณดิจิตอล

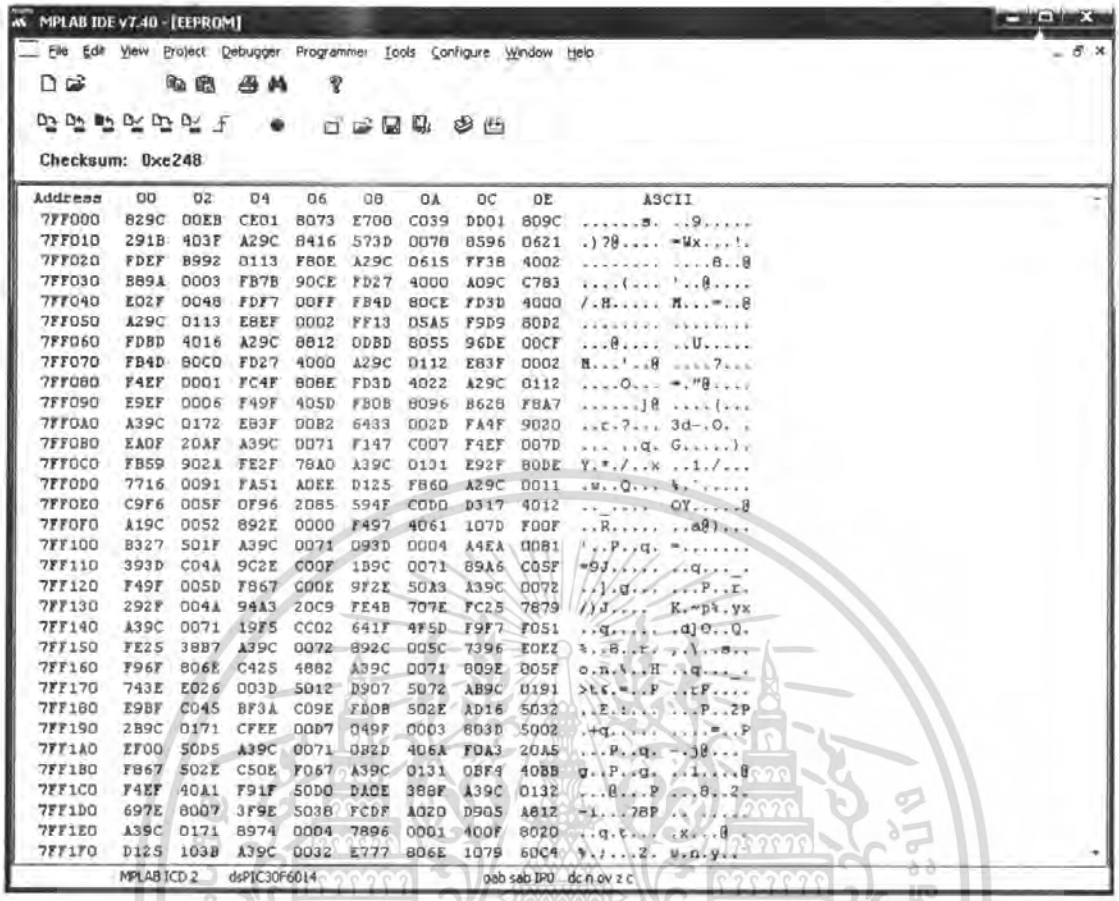
หลังจากที่เราได้สัญญาณเสียงอ้างอิงแล้ว ก็ต้องมีเสียงทดสอบเพื่อทำการรู้จำเสียง โดยเราจะทำการบันทึกเสียง เราจะใช้ Si3000 เป็นตัวแปลงสัญญาณเสียงให้เป็นสัญญาณดิจิตอล 16 บิต โดยที่อัตราการสุ่มสัญญาณอยู่ที่ 12 KHz เนื่องจากเป็นอัตราการสุ่มสัญญาณสูงสุดของ Si3000

รูปที่ 3.6 ไอซี SI 3000 รับสัญญาณเสียงจากการพูดทางไมโครโฟน



รูปที่ 3.7 ตัวอย่างสัญญาณเสียง play ที่พูดผ่านไมโครโฟน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

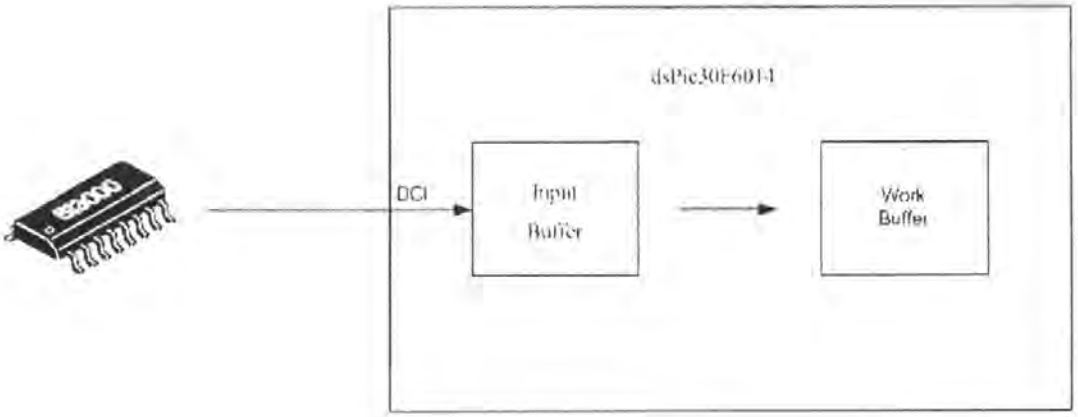


รูปที่ 3.8 ตัวอย่างข้อมูลที่ผ่านการแปลงสัญญาณด้วย Si3000

3.3 dsPIC30F6014 รับข้อมูล

สำหรับการรับข้อมูลสัญญาณดิจิตอลจาก Si3000 เราจะใช้โมดูล DCI (Data Converter Interface) โดยแต่ละครั้งของการอินเตอร์รับจะสามารถรับข้อมูลได้ 4 ตัวอย่าง ไปเก็บไว้ใน อินพุตบัฟเฟอร์(Input buffer) จนเต็มหน่วยความจำในอินพุตบัฟเฟอร์ซึ่งจะมีจำนวนข้อมูล ทั้งหมด 84 ตัวอย่าง หลังจากที่นั้นเราจะส่งข้อมูลต่อไปยัง เวิร์คบัฟเฟอร์ (Work buffer) เพื่อที่ อินพุตบัฟเฟอร์ จะสามารถเก็บข้อมูลสัญญาณเสียงที่เข้ามาใหม่ได้และนำข้อมูลใน เวิร์คบัฟเฟอร์ ไปประมวลผลต่อไป จำนวนข้อมูลทั้ง 84 ตัวอย่างนี้ เราจะเรียกว่า ซาวด์เฟรม(sound frame) โดยจำนวนเฟรมจะมากหรือน้อยขึ้นอยู่กับความยาวของเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.9 บล็อคแสดงกระบวนการรับข้อมูลสัญญาณดิจิตอล

3.4 การหาตัวแทนของเสียง

เนื่องจากเราไม่ได้นำข้อมูลสัญญาณเสียงทั้งหมดมาทำการประมวลผลเพราะต้องใช้หน่วยความจำมากและเวลาในการประมวลผลนาน ดังนั้น เราจำเป็นต้องหาตัวแทนของสัญญาณเสียง โดยวิธี MFCC ในการหาตัวแทนของเสียง เราจะนำ ซาวด์เฟรมแต่ละเฟรมที่อยู่ในเวิร์คบัฟเฟอร์ มาผ่านกระบวนการ MFCC ผลลัพธ์ที่ได้คือข้อมูลเวคเตอร์ ที่เรียกว่า ซาวด์เวคเตอร์ (sound vector) ทำไปเรื่อยจนครบทุกเฟรม

รูปที่ 3.10 บล็อคแสดงการหาตัวแทนของเสียง

3.5 การหาผลลัพธ์ของเสียง

ในการที่เราจะสามารถทราบว่าเสียงที่เราพูดไปนั้นเป็นเสียงใด ทำได้โดยการนำ ซาวด์เฟรมของเสียงมาผ่านกระบวนการ HMM ของแต่ละคำที่มีอยู่ในไลบรารี่ ทำไปจนครบทุกคำ ผลลัพธ์ที่ได้จะเป็นค่าตัวเลขที่จะใช้ในการอ้างอิงคำที่อยู่ในไลบรารี่ต่อไปดังในตารางที่ 3.1

3.5.1 ฮิดเดนมาร์คอฟโมเดล

ในการสร้างแบบจำลองและการรู้จำเสียงเพื่อเปรียบเทียบความคล้ายคลึงของสัญญาณเสียงแบบ HMM ของเสียงมีขั้นตอนดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. สุ่มค่าเริ่มต้น a , b และกำหนดให้ π = [1 0 0 0 0 0] ตามเงื่อนไขแบบจำลอง แบบซ้าย – ขวา

2. หาค่า α , β จากค่า a , b เริ่มต้น และลำดับค่าปรากฏ $\boldsymbol{O}$ = { $\boldsymbol{O}_1$, $\boldsymbol{O}_2$, ..., $\boldsymbol{O}_T$ }ซึ่งเรียกว่าลำดับเทรนนิ่ง ตามวิธีของ กระบวนการไปข้างหน้าและย้อนกลับ โดยใช้ลำดับของค่าปรากฏหลายๆ ลำดับเข้ามาเทรน เพื่อความถูกต้องมากขึ้น

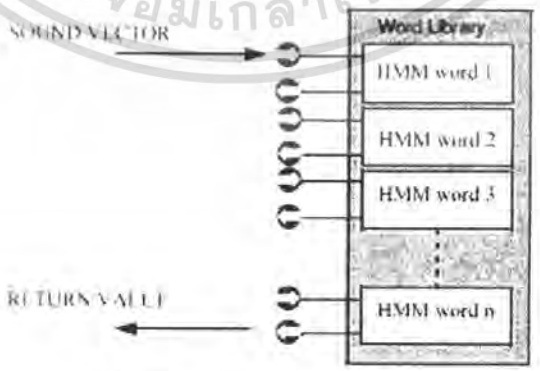
3. หาค่าพารามิเตอร์ a , b , π ที่ให้ความน่าจะเป็นสูงสุดที่จะเป็นแบบจำลอง λ ที่เหมาะสม

4. ตรวจสอบค่าพารามิเตอร์แบบจำลองที่ได้ว่าลู่เข้าหรือยัง โดยใช้วิธีการคำนวณค่า a , b ซ้ำๆ โดยในที่นี้ตั้งไว้ว่าเมื่อไม่มีการเปลี่ยนแปลงหรือค่าเท่าเดิมให้หยุดการทำซ้ำ ก็จะได้ค่าพารามิเตอร์ a , b , π ของแบบจำลองที่ต้องการ

5. เก็บค่าพารามิเตอร์ a , b , π ที่ได้จากข้อ 4. เป็นพารามิเตอร์ของแบบจำลองไว้ และหลังจากที่ได้แบบจำลอง HMM ของแต่ละคำศัพท์แล้วเมื่อมีลำดับค่าปรากฏ $\boldsymbol{O}$ = { $\boldsymbol{O}_1$, $\boldsymbol{O}_2$, ..., $\boldsymbol{O}_T$ } ของเสียงที่ไม่ทราบ ซึ่งเป็นเสียงที่ต้องการทดสอบเข้ามาเราจะทำการคำนวณหาค่าความน่าจะเป็น $\boldsymbol{P}$ ($\boldsymbol{O}$ | λ) ทุกแบบจำลองของแต่ละคำศัพท์โดยใช้วิธี วีเทอร์บี อัลกอริทึม แล้วเลือกเอาคำศัพท์ที่มีความน่าจะเป็นสูงสุดซึ่งก็คือ คำศัพท์ที่แบบจำลองจำได้นั่นเอง

3.5.2 การตัดสินใจ

เป็นขั้นตอนการตัดสินว่าเสียงนั้นเป็นเสียงอะไร โดยเปรียบเทียบกับสัญญาณเสียงอ้างอิง หลังจากนั้นก็ทำการเลือกผลที่ได้จากการเปรียบที่เหมือนกันที่สุดตามกฎการตัดสินใจ นั่นคือการตัดสินใจในการจำแนกเสียงพูดว่าเสียงที่นำมาทดลองเป็นคำอะไรนั่นเอง



รูปที่ 3.11 จำลองกระบวนการ HMM

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 คำในไลบารี่และค่าของคำ

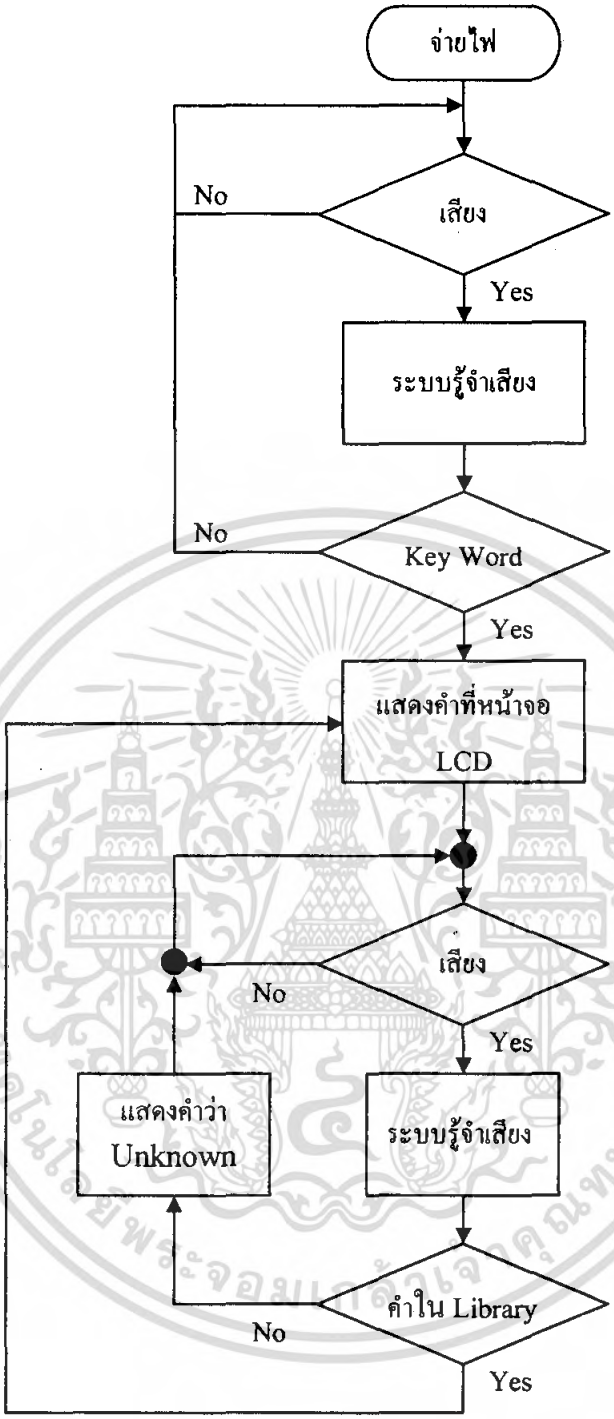
WORD	VALUE
Control	1
Next	2
Pause	3
Play	4
Previous	5
Stop	6
Unknown	7

3.6 การแสดงผล

เรามาแสดงผลการรู้จำเสียงบน LCD ได้โดย นำค่าตัวเลขจากการหาผลลัพธ์ ไปอ้างอิงกับตารางแล้วนำไปแสดงผล จะแสดงดังรูปที่ 3.12

รูปที่ 3.12 ตัวอย่างการแสดงผลของเสียง PLAY

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.13 ผังงานแสดงการทำงานโดยรวมตั้งแต่หัวข้อ 3.1 – 3.6

3.7 การทดลองหาค่าเปอร์เซ็นความถูกต้อง

เป็นการทดลองเพื่อค่าเปอร์เซ็นความถูกต้อง ทั้งผู้ชายและผู้หญิง ในสถานการณ์ต่าง คือ

- ห้องเงียบ
- มีสัญญาณรบกวน เช่น ในระหว่างเล่นเพลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.8 การนำเสียงไปประยุกต์ใช้

นำเสียงไปควบคุมการเล่นเพลง โดยได้เขียนโปรแกรม MEDIA PLAYER ดังรูปที่ 3.14 ขึ้นมา รับคำสั่งจาก dsPICแล้ว MEDIA PLAYER ปฏิบัติตามคำสั่งนั้น

รายละเอียดการทำงานของ MEDIA PLAYER

-ปุ่ม play ใช้ในการเล่นเพลง

-ปุ่ม stop ใช้ในการหยุดเล่นเพลง

-ปุ่ม pause ใช้ในการหยุดเล่นเพลงชั่วคราว

-ปุ่ม << ใช้ในการย้อนกลับไปเล่นเพลงก่อนหน้านี้

-ปุ่ม >> ใช้ในการเล่นเพลงถัดไป

-ปุ่ม add เพิ่มเพลงลงใน PLAYLIST

-ปุ่ม remove ลบเพลงที่เลือกออกจาก PLAYLIST

-ปุ่ม clear ลบเพลงใน PLAYLIST ทั้งหมด

- PLAYLIST แสดงรายชื่อเพลงทั้งหมด

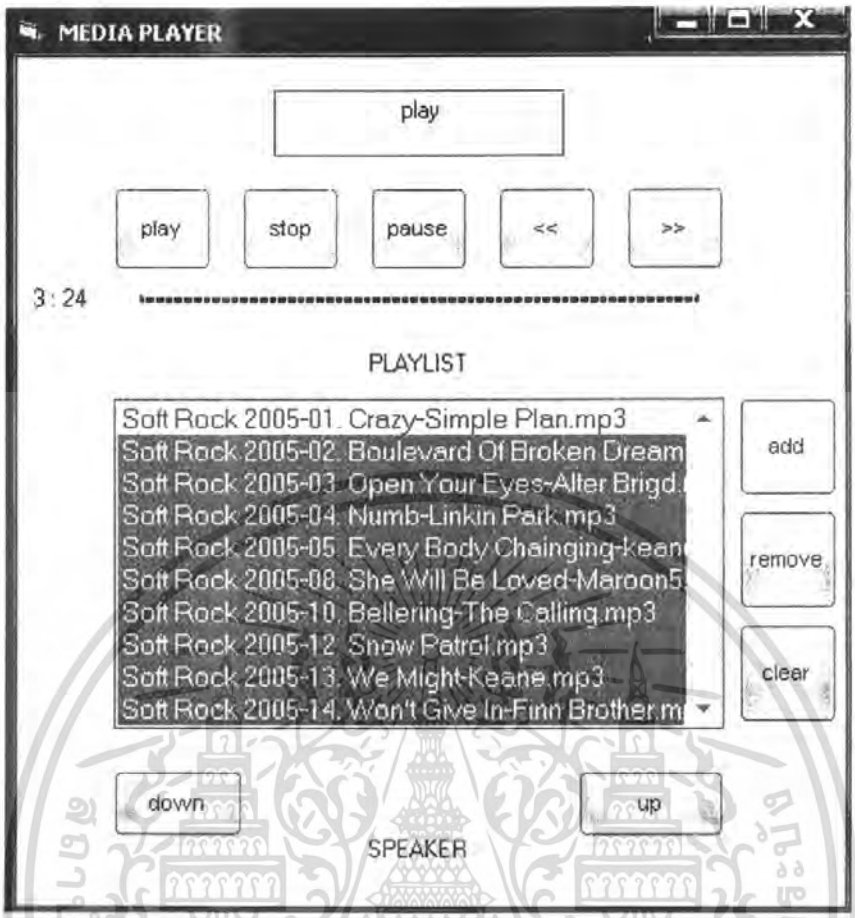
-ปุ่ม up และ down เพิ่มและลดระดับเสียง

-ช่องสีขาวด้านบนบอกว่าเสียงทดสอบเป็นเสียงใด

-สามารถใช้เสียง play แทนปุ่มplay เสียง stop แทนปุ่ม stop เสียง pause แทนปุ่ม pause เสียง next แทนปุ่ม next และเสียง previous แทนปุ่มprevious

-มีตัวเลขแสดงเวลาที่เหลือของเพลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.14 ตัวอย่างโปรแกรม MEDIA PLAYER

จากการออกแบบและการสร้างนั้น จะมีตัว Si3000 เป็นตัวแปลงสัญญาณเสียงของผู้ทดสอบให้เป็นสัญญาณดิจิตอล โดยมีอัตราการสุ่มสัญญาณที่ 12 kHZ และ ตัว dsPIC30F6014 จะเป็นตัวประมวลผลกลางโดยรับสัญญาณดิจิตอลจากตัว Si3000 หลังจากนั้นเราทำการหาค่าตัวแทนของเสียงจากสัญญาณดิจิตอลด้วย วิธี MFCC แล้วนำไปเปรียบเทียบกับเสียงต้นแบบที่เราได้สร้างไว้ก่อนแล้วโดยใช้วิธี HMM เพื่อที่จะหาว่าเสียงทดสอบใกล้เคียงกับเสียงใดมากที่สุด แล้วนำไปแสดงผลที่จอ LCD เพื่อบอกให้ผู้ทดสอบรู้ว่าผลการรู้จำเสียงนั้นเป็นคำใด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

การทดลองและผลการทดลอง

ในกระบวนการการรู้จำเสียงโดยใช้ dsPIC30F6014 เราจำเป็นที่จะต้องหาว่ามีประสิทธิภาพของการรู้จำเสียงมากน้อยเพียงใด สามารถที่จะปฏิบัติตามคำสั่งเราได้หรือไม่ โดยการหาเปอร์เซ็นต์ความถูกต้องของการรู้จำเสียง เราจะทำการทดลองกับบุคคลทั่วไปที่มีความรู้ทางด้านภาษาอังกฤษโดยทางคณะผู้จัดทำได้แบ่งการทดลองออกเป็น 2 กรณี คือ

กรณีที่ 1 ทดลองในห้องเงียบ

กรณีที่ 2 ทดลองในห้องที่มีเสียงรบกวน เช่น เสียงรบกวนจากการเปิดเพลง จากห้องทำงาน

การทดลองนี้มีเพื่อนร่วมการทดลองทั้งหมด14 คนโดยที่แบ่งออกเป็น ผู้ชาย 7 คน ผู้หญิง 7 คน มีรายชื่อผู้ร่วมการทดลองดังนี้ คือ

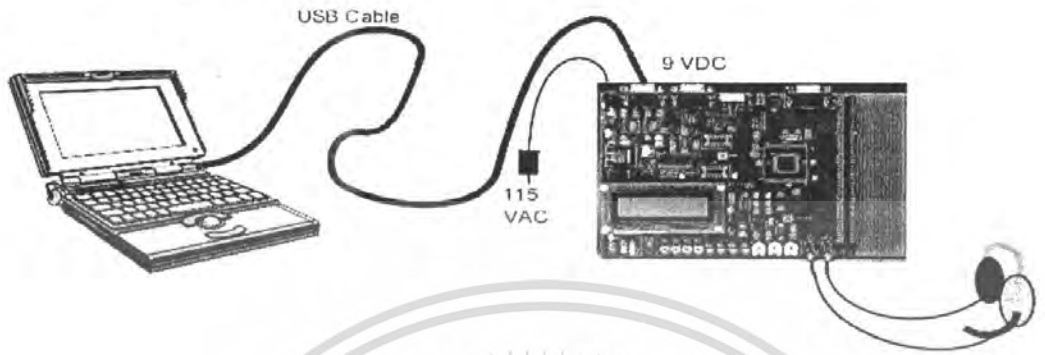
1. นาย พัลลภ กลิ่นขจร
2. นาย ภาคภูมิ ไพโรจน์
3. นาย บุญคิด นิลประพัฒน์
4. นาย วิศรุต ธรรมทัตโต
5. นาย คมสัน อนันต์วัฒนสุข
6. นาย วีรวัฒน์ ขวัญแก้ว
7. นาย เสน่ห์ จันทมุลตรี
8. นางสาว ธัญญรัตน์ รักทรง
9. นางสาว ชยาภรณ์ ศรีราชญา
10. นางสาว รัตนา จิรัฏฐิติกาล
11. นางสาว ปาริชาติ คงจุ้ย
12. นางสาว พิจิตตรา ศรีสุนทรไทย
13. นางสาว พีรยา สิริธาราธิกุล
14. นางสาว พันธิตรา ชัชวาลไกลวงศ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1 การทดลอง

กรณีที่ 1 ทดลองในห้องเงียบ

1.ติดตั้งอุปกรณ์ดังรูปที่ 4.1



รูปที่ 4.1 การติดตั้งอุปกรณ์

2.กดปุ่มรีเซ็ตบนบอร์ดจะปรากฏข้อความดังรูปที่ 4.2 ให้ผู้ทดลองออกเสียงคีเวิร์ดซึ่งก็คือเสียง Control

รูปที่ 4.2 ข้อความที่ปรากฏบนจอ LCD

3.ให้ผู้ทดลองทำการออกเสียงดังนี้ Control , Play , Pause , Stop , Previous , Next โดยให้พูดอย่างละ 10 คำ โดย จะเริ่มที่คำใดก่อนก็ได้ตามแต่ใจของผู้ทำการทดลองและไม่จำเป็นต้องพูดเรียงคำ

4.บันทึกผลการทดลองโดยหากเสียงของผู้ทดสอบตรงกับคำที่แสดงบนจอ LCD ถือว่าทำการรู้จำเสียงได้ถูกต้อง หากเสียงของผู้ทดสอบไม่ตรงกับคำที่แสดงบนจอ LCD ถือว่าทำการรู้จำเสียงได้ไม่ถูกต้อง

กรณีที่ 1 ทดลองในห้องที่มีสัญญาณรบกวน

ทำการทดลองเช่นเดียวกับกรณีแรก แต่ย้ายสถานที่ทำการทดลองไปที่ห้องทำงานและระหว่างการทดลองจะเปิดเพลงด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ผลการทดลอง

ตารางที่ 4.1 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 1

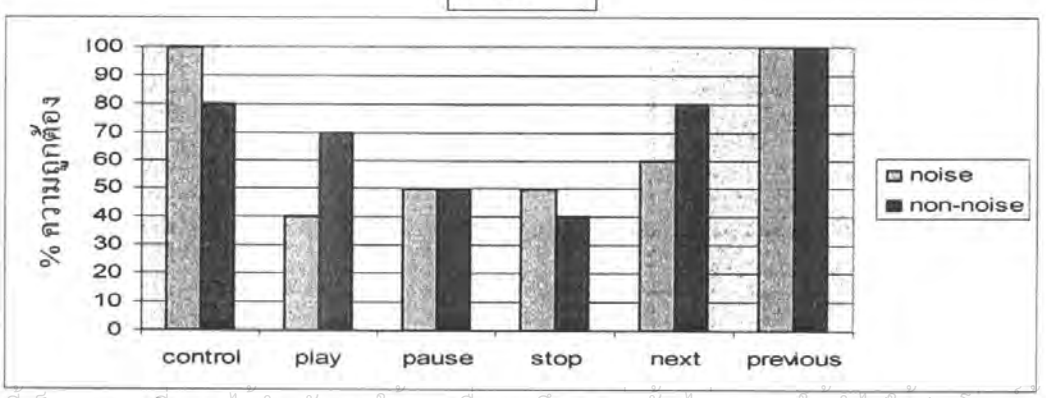
ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	x	x	√	√	√	√	√	80
play	√	√	√	x	√	√	√	x	x	√	70
pause	x	√	√	x	√	x	√	x	√	x	50
stop	x	√	√	x	x	x	x	x	√	√	40
next	x	√	√	√	√	√	x	√	√	√	80
previous	√	√	√	√	√	√	√	√	√	√	100

% เฉลี่ย
70

ตารางที่ 4.2 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 1

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	√	√	√	√	√	100
play	x	x	√	x	√	x	√	√	x	x	40
pause	√	√	√	x	x	√	x	√	x	x	50
stop	x	x	x	√	x	√	√	x	√	√	50
next	x	x	x	√	√	√	√	√	x	√	60
previous	√	√	√	√	√	√	√	√	√	√	100

%เฉลี่ย
67.67



รูปที่ 4.3 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 1

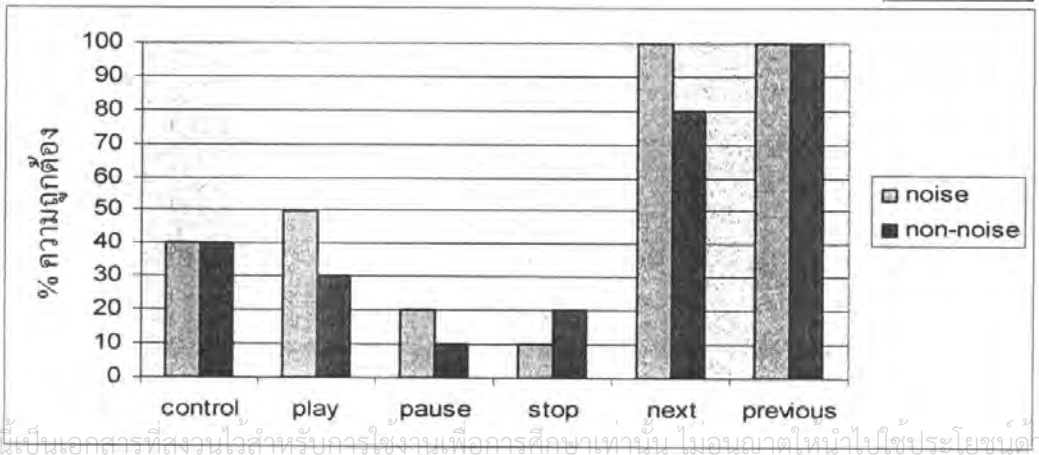
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 2

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	x	√	x	√	x	√	x	x	x	40
play	x	x	x	√	√	x	x	x	√	x	30
pause	x	x	x	x	√	x	x	x	x	x	10
stop	x	x	x	x	x	√	√	x	x	x	20
next	√	√	√	√	x	√	√	x	√	√	80
previous	√	√	√	√	√	√	√	√	√	√	100
											% เฉลี่ย
											46.67

ตารางที่ 4.4 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 2

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	x	√	x	√	x	√	x	x	x	40
play	x	x	x	√	√	√	√	x	√	x	50
pause	x	x	x	x	√	√	x	x	x	x	20
stop	x	x	x	x	x	√	x	x	x	x	10
next	√	√	√	√	√	√	√	√	√	√	100
previous	√	√	√	√	√	√	√	√	√	√	100
											% เฉลี่ย
											53.33



รูปที่ 4.4 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 2

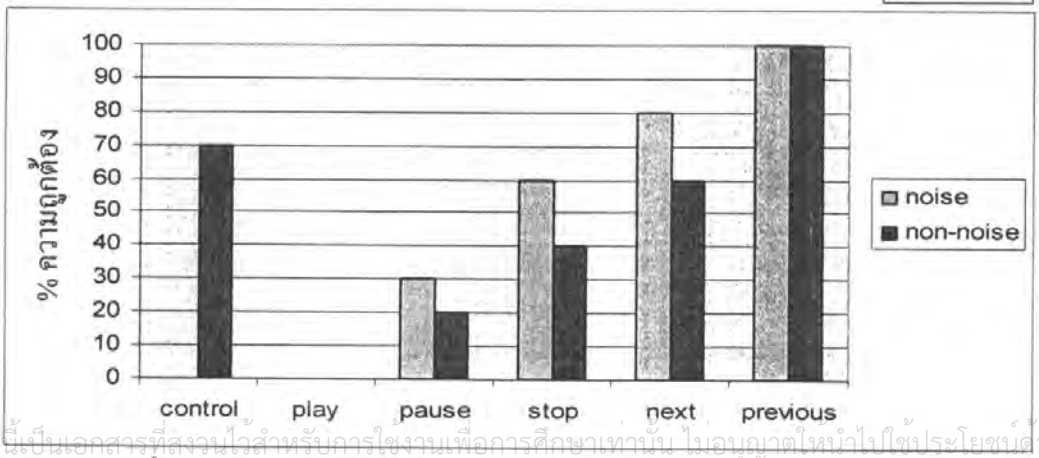
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.5 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 3

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	√	x	x	x	√	70
play	x	x	x	x	x	x	x	x	x	x	0
pause	√	x	x	x	x	x	x	x	x	√	20
stop	√	√	√	√	x	x	x	x	x	x	40
next	x	√	√	√	x	x	√	√	√	x	60
previous	√	√	√	√	√	√	√	√	√	√	100
											% เฉลี่ย
											48.33

ตารางที่ 4.6 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 3

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	x	x	x	x	x	x	x	x	x	x	0
play	x	x	x	x	x	x	x	x	x	x	0
pause	x	x	x	x	√	√	√	x	x	x	30
stop	√	√	√	√	x	√	√	x	x	x	60
next	√	√	√	x	x	√	√	√	√	√	80
previous	√	√	√	√	√	√	√	√	√	√	100
											% เฉลี่ย
											45



รูปที่ 4.5 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 3

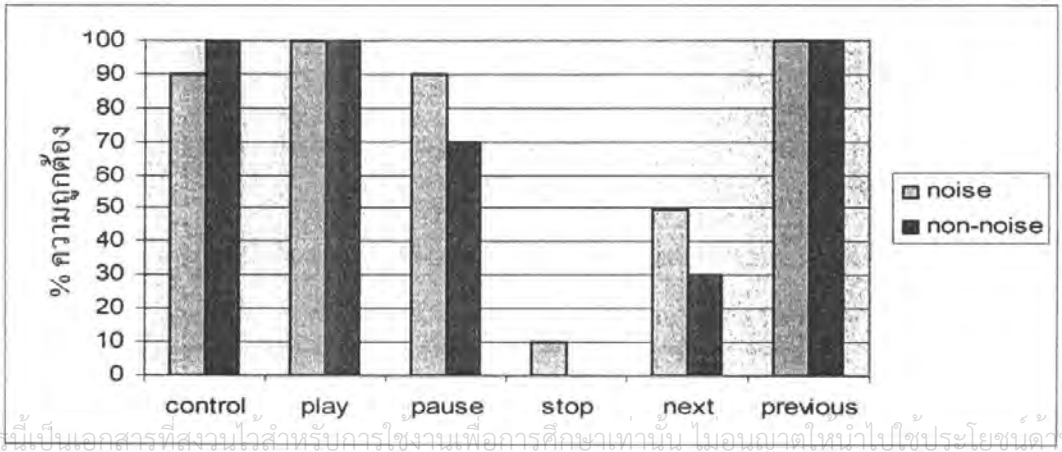
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.7 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 4

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	√	√	√	√	√	100
play	√	√	√	√	√	√	√	√	√	√	100
pause	√	x	x	x	√	√	√	√	√	√	70
stop	x	x	x	x	x	x	x	x	x	x	0
next	x	x	x	x	√	√	√	x	x	x	30
previous	√	√	√	√	√	√	√	√	√	√	100
											% เฉลี่ย
											66.67

ตารางที่ 4.8 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 4

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	x	√	√	√	√	90
play	√	√	√	√	√	√	√	√	√	√	100
pause	√	x	√	√	√	√	√	√	√	√	90
stop	x	x	x	x	x	√	x	x	x	x	10
next	x	√	√	x	√	√	√	x	x	x	50
previous	√	√	√	√	√	√	√	√	√	√	100
											% เฉลี่ย
											71.67



รูปที่ 4.6 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 4

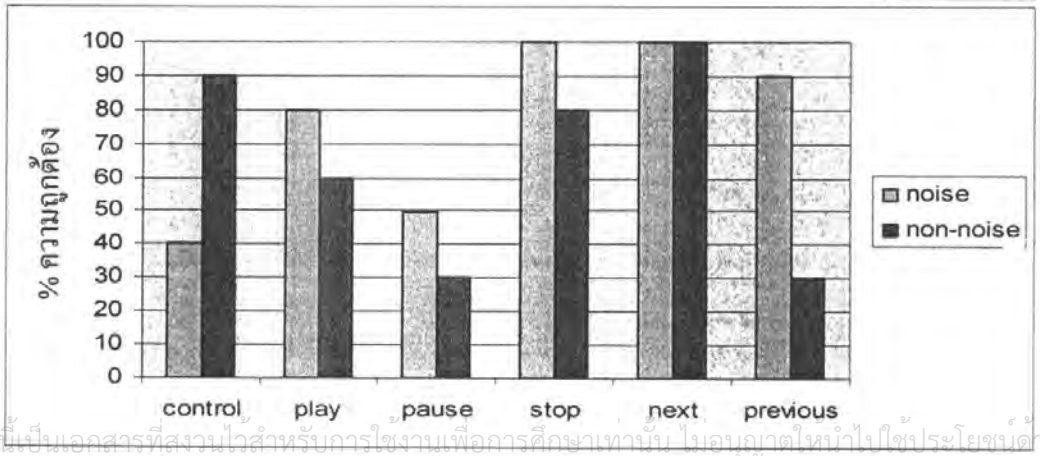
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.9 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 5

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	x	√	√	√	√	√	√	√	√	90
play	x	x	x	√	√	√	√	√	x	√	60
pause	x	x	√	√	x	x	√	x	x	x	30
stop	√	√	√	√	x	√	√	√	x	√	80
next	√	√	√	√	√	√	√	√	√	√	100
previous	x	x	x	x	√	√	√	x	x	x	30
											% เฉลี่ย
											71.67

ตารางที่ 4.10 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 5

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	x	x	x	x	√	√	x	x	40
play	√	√	√	x	√	√	√	x	√	√	80
pause	x	√	x	√	√	x	√	x	x	√	50
stop	√	√	√	√	√	√	√	√	√	√	100
next	√	√	√	√	√	√	√	√	√	√	100
previous	√	√	√	x	√	√	√	√	√	√	90
											% เฉลี่ย
											76.67



รูปที่ 4.7 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 5

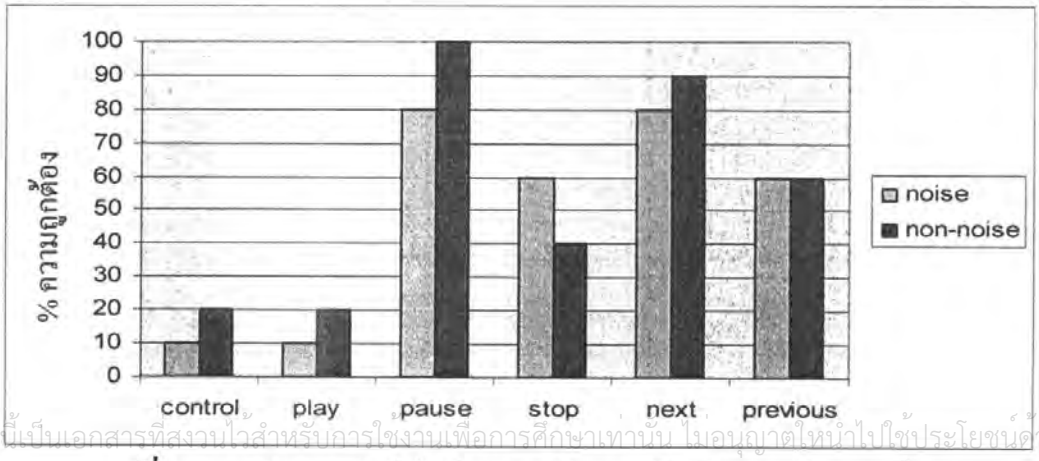
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.11 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 6

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	x	x	x	x	√	x	x	x	x	20
play	√	x	x	x	x	x	x	x	√	x	20
pause	√	√	√	√	√	√	√	√	√	√	100
stop	√	√	x	x	x	x	x	x	√	√	40
next	√	√	√	√	x	√	√	√	√	√	90
previous	√	√	√	x	√	√	x	x	x	√	60
											% เฉลี่ย
											55

ตารางที่ 4.12 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 6

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	x	x	x	x	x	x	x	x	x	10
play	x	x	x	x	x	x	x	x	√	x	10
pause	√	√	√	√	x	x	√	√	√	√	80
stop	√	√	√	√	x	x	x	x	√	√	60
next	√	√	√	√	x	x	√	√	√	√	80
previous	√	√	√	x	√	√	x	x	x	√	60
											% เฉลี่ย
											55



รูปที่ 4.8 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 6

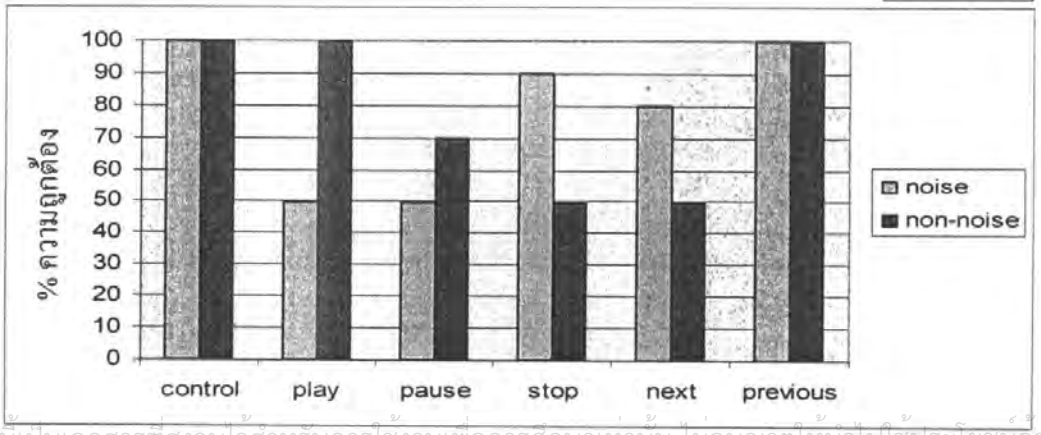
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของผลงานทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.13 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้ชายคนที่ 7

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	√	√	√	√	√	100
play	√	√	√	√	√	√	√	√	√	√	100
pause	√	x	√	√	x	√	√	x	√	√	70
stop	x	x	√	√	x	√	x	√	x	√	50
next	x	x	√	√	x	x	√	x	√	√	50
previous	√	√	√	√	√	√	√	√	√	√	100
											% เฉลี่ย
											78.33

ตารางที่ 4.14 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้ชายคนที่ 7

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	√	√	√	√	√	100
play	√	x	√	x	x	x	√	x	√	√	50
pause	x	x	√	√	x	√	x	√	x	√	50
stop	√	√	√	√	√	√	√	√	√	x	90
next	√	√	√	√	√	√	x	x	√	√	80
previous	√	√	√	√	√	√	√	√	√	√	100
											% เฉลี่ย
											78.33



รูปที่ 4.9 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชายคนที่ 7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในส่วนของการทดลองกระบวนการการรู้จำเสียงของผู้ชาย เราได้ให้ผู้ชายพูดเสียงผ่านทางไมโครโฟนโดยมีเสียงทดสอบทั้งหมด 6 คำ ดังนี้ Control , Play , Pause , Stop , Previous และ Next คำละ10 ครั้งโดยผู้ทำการทดลองจะพูดเสียงใดก่อนก็ได้ จะเรียงตามลำดับหรือจะไม่เรียงก็ได้ โดยได้ผลการทดลองตามตารางที่ 4.1 ถึงตารางที่ 4.14 เราจะใช้ √ หากเสียงทดสอบกับคำที่แสดงบนจอ LCD นั้นตรงกัน และจะใช้ x หากเสียงทดสอบกับคำที่แสดงบนจอ LCD นั้นไม่ตรงกัน จากผลการทดลองพบว่า ผู้ชายบางคนสามารถพูดเสียงทดสอบบางคำได้ตรงทุกครั้ง แต่บางคำผิดทุกครั้ง ดังเช่น ตารางที่ 4.5 สามารถออกเสียง Previous ได้เปอร์เซ็นความถูกต้องได้ 100% แต่ Play ได้ 0 % หากเปรียบเทียบผลการทดลองตารางที่ 4.2กับตารางที่ 4.6 พบว่าเสียง Control ของผู้ชายคนที่ 1 ได้เปอร์เซ็นความถูกต้องได้ 100 % แต่ผู้ชายคนที่ 3 ได้ 0 % แสดงคนแต่ละคนออกเสียงเดียวสัญญาณเสียงไม่เหมือนกันหรือแม้แต่คนๆเดียวกันก็ตาม หากเราเปรียบเทียบผลการทดลองในทั้งสองกรณีโดยพิจารณาจากกราฟ พบว่า ผลการทดลองไม่แตกต่างกันมากนักอาจจะเกิดจากมีสัญญาณรบกวนในสายไมโครโฟนทำให้ได้ผลการทดลองที่ใกล้เคียงกัน และจากการทดลองนั้น ประสิทธิภาพของการรู้จำเสียงของผู้ชายได้ผลการทดลองเป็นที่น่าพอใจ โดยมีเปอร์เซ็นความถูกต้องเฉลี่ย 62.83 %

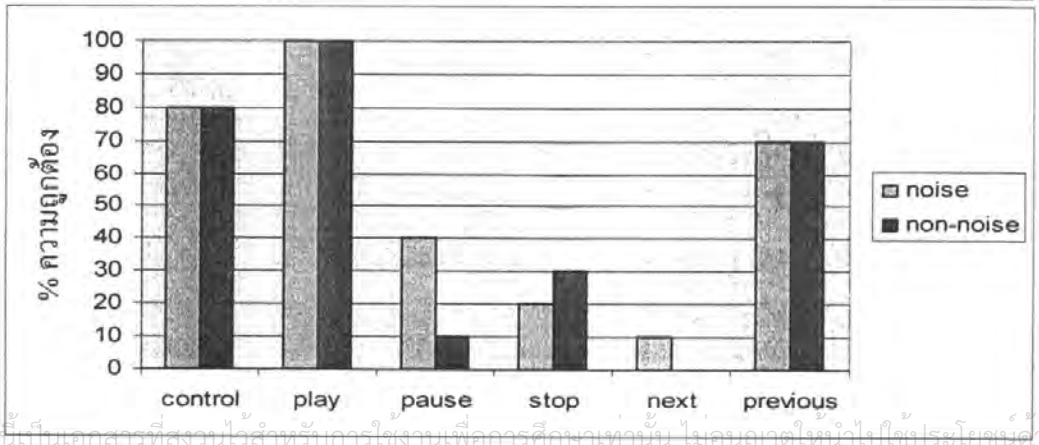
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.15 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 1

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	√	√	√	x	x	80
play	√	√	√	√	√	√	√	√	√	√	100
pause	x	x	√	x	x	x	x	x	x	x	10
stop	√	x	x	√	x	x	x	x	x	√	30
next	x	x	x	x	x	x	x	x	x	x	0
previous	x	√	√	√	x	√	x	√	√	√	70
											%เฉลี่ย
											48.33

ตารางที่ 4.16 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 1

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	x	x	√	√	√	√	√	√	80
play	√	√	√	√	√	√	√	√	√	√	100
pause	√	√	√	√	x	x	x	x	x	x	40
stop	x	√	x	x	x	x	x	√	x	x	20
next	x	√	x	x	x	x	x	x	x	x	10
previous	√	√	√	x	x	√	√	x	√	√	70
											% เฉลี่ย
											55.33



รูปที่ 4.10 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 1

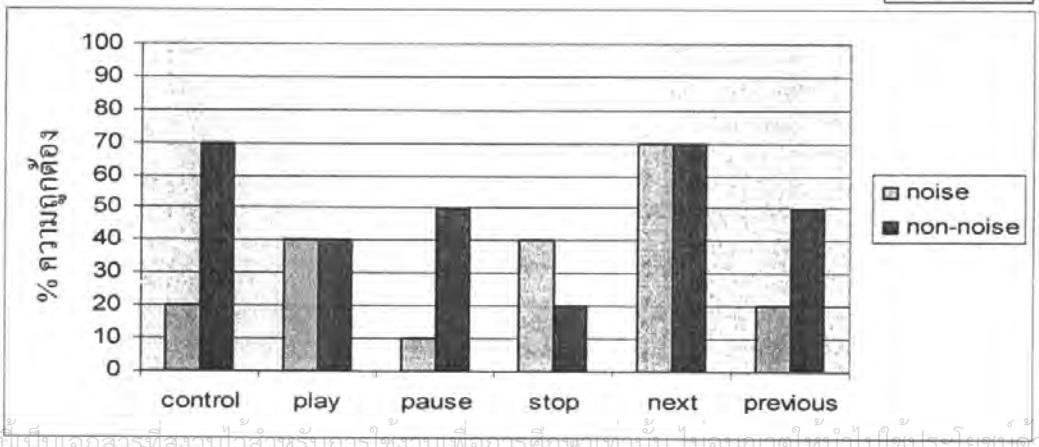
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.17 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 2

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	x	x	√	√	√	x	x	x	x	x	70
play	x	x	x	x	√	√	x	√	√	x	40
pause	x	√	√	x	x	√	√	√	x	x	50
stop	x	x	x	x	x	x	x	x	√	√	20
next	√	√	√	√	√	√	√	x	x	x	70
previous	√	√	√	√	√	x	x	x	x	x	50
											% เฉลี่ย
											50

ตารางที่ 4.18 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 2

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	x	x	x	x	x	x	x	x	√	√	20
play	√	√	x	√	√	x	x	x	x	x	40
pause	x	x	√	x	x	x	x	x	x	x	10
stop	x	x	√	x	x	x	√	√	√	x	40
next	x	√	√	x	√	x	√	√	√	√	70
previous	x	√	√	x	x	x	x	x	x	x	20
											% เฉลี่ย
											33.33



รูปที่ 4.11 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 2

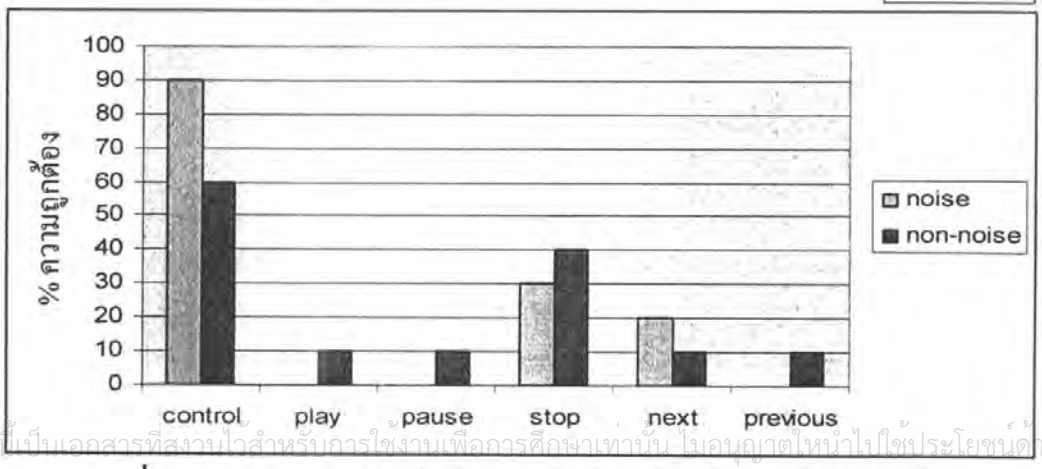
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.19 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 3

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	x	√	x	x	x	√	60
play	x	x	x	x	x	x	x	√	x	x	10
pause	x	x	√	x	x	x	x	x	x	x	10
stop	x	√	x	√	x	√	√	x	x	x	40
next	x	x	x	√	x	x	x	x	x	x	10
previous	√	x	x	x	x	x	x	x	x	x	10
											% เฉลี่ย
											23.33

ตารางที่ 4.20 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 3

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	√	√	√	√	x	90
play	x	x	x	x	x	x	x	x	x	x	0
pause	x	x	x	x	x	x	x	x	x	x	0
stop	x	√	x	√	x	x	x	x	x	√	30
next	x	√	x	√	x	x	x	x	x	x	20
previous	x	x	x	x	x	x	x	x	x	x	0
											% เฉลี่ย
											23.33



รูปที่ 4.12 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 3

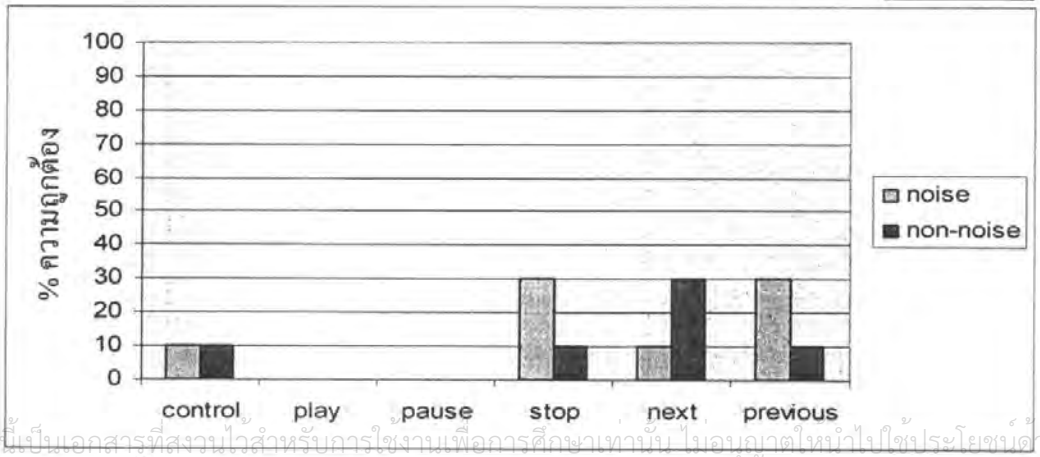
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของผลงานทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.21 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 4

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	√	√	√	√	√	10
play	x	x	x	x	x	x	x	x	x	x	0
pause	x	x	x	x	x	x	x	x	x	x	0
stop	x	x	x	x	√	x	x	x	x	x	10
next	x	x	√	x	x	√	√	x	x	x	30
previous	x	x	x	x	x	√	x	x	x	x	10
											% เฉลี่ย
											25

ตารางที่ 4.22 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 4

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	√	√	√	√	√	10
play	x	x	x	x	x	x	x	x	x	x	0
pause	x	x	x	x	x	x	x	x	x	x	0
stop	x	x	x	√	√	√	x	x	x	x	30
next	x	x	x	x	x	√	x	x	x	x	10
previous	x	x	x	x	x	x	x	√	√	√	30
											% เฉลี่ย
											28.33



รูปที่ 4.13 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 4

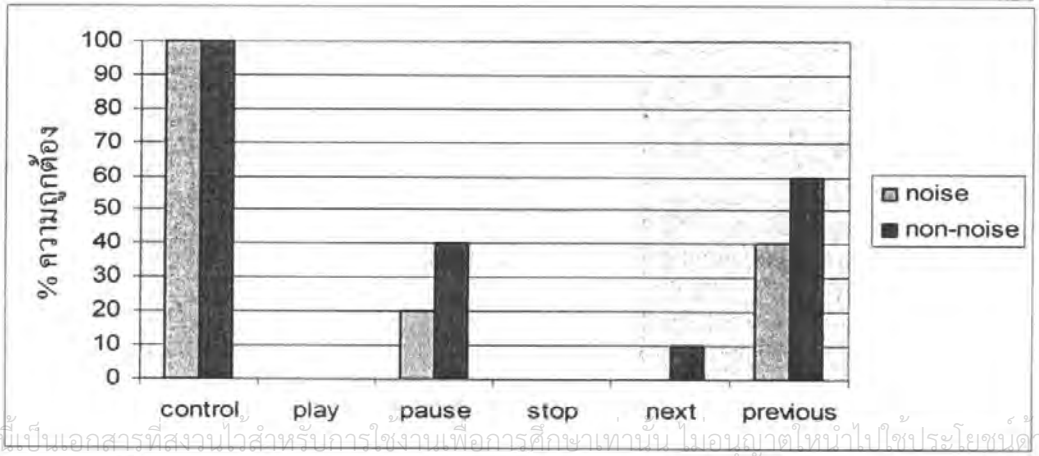
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.23 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 5

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	√	√	√	√	√	100
play	x	x	x	x	x	x	x	x	x	x	0
pause	x	x	√	x	√	√	√	x	x	x	40
stop	x	x	x	x	x	x	x	x	x	x	0
next	x	x	x	√	x	x	x	x	x	x	10
previous	x	x	x	√	√	x	√	√	√	√	60
											% เฉลี่ย
											35

ตารางที่ 4.24 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 5

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	√	√	√	√	√	100
play	x	x	x	x	x	x	x	x	x	x	0
pause	x	x	√	x	x	x	√	x	x	x	20
stop	x	x	x	x	x	x	x	x	x	x	0
next	x	x	x	x	x	x	x	x	x	x	0
previous	x	x	x	√	√	x	√	x	x	√	40
											% เฉลี่ย
											26.67



รูปที่ 4.14 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 5

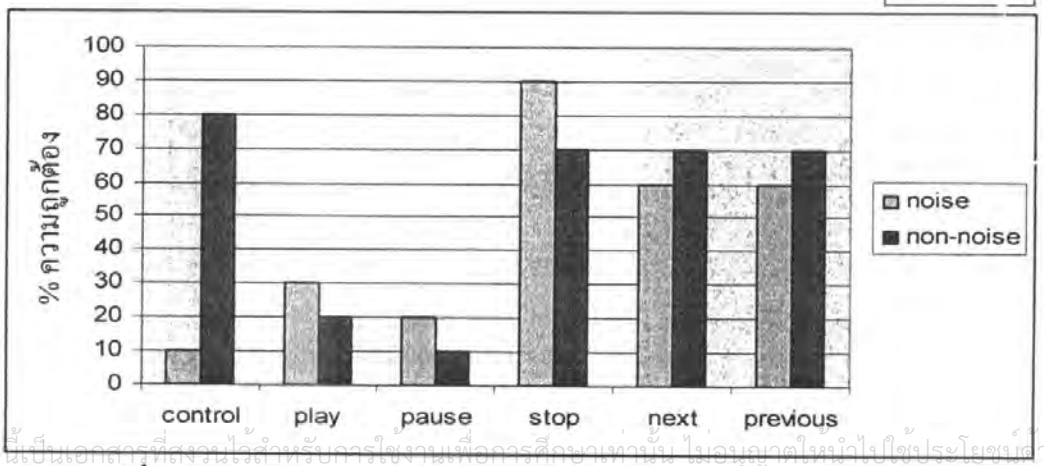
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.25 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 6

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	√	√	√	√	√	x	x	80
play	x	x	√	x	x	x	√	x	x	x	20
pause	x	x	x	x	x	x	√	x	x	x	10
stop	x	x	x	√	√	√	√	√	√	√	70
next	x	x	√	√	x	x	√	x	x	x	30
previous	x	x	√	√	x	x	√	x	x	x	30
											% เฉลี่ย
											53.33

ตารางที่ 4.26 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 6

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	x	x	x	x	x	x	x	x	x	10
play	x	x	√	√	x	x	√	x	x	x	30
pause	x	x	x	√	x	x	√	x	x	x	20
stop	x	x	x	√	x	x	x	x	x	x	10
next	x	x	x	√	√	x	√	x	x	√	40
previous	√	√	√	√	√	√	√	√	√	√	100
											% เฉลี่ย
											51.67



รูปที่ 4.15 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 6

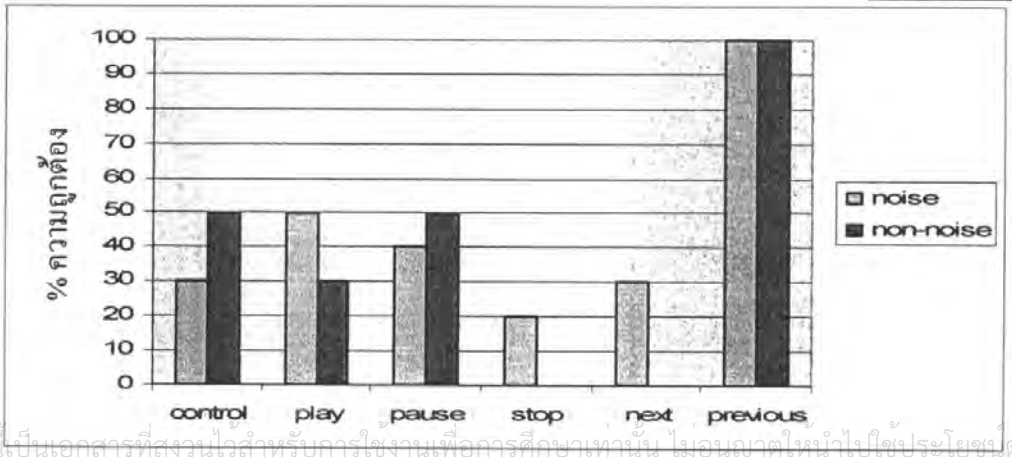
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.27 ผลการทดลองการรู้จำเสียงในห้องเงียบของผู้หญิงคนที่ 7

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	√	√	x	√	x	x	x	x	√	50
play	x	√	√	x	x	x	x	x	√	x	30
pause	√	√	√	x	√	x	x	x	x	√	50
stop	x	x	x	x	x	x	x	x	x	x	0
next	x	x	x	x	x	x	x	x	x	x	0
previous	√	√	√	√	√	√	√	√	√	√	100
											% เฉลี่ย
											38.33

ตารางที่ 4.28 ผลการทดลองการรู้จำเสียงในห้องที่มีสัญญาณรบกวนของผู้หญิงคนที่ 7

ครั้งที่ / คำสั่ง	1	2	3	4	5	6	7	8	9	10	%ความถูกต้อง
control	√	x	√	x	√	x	x	x	x	x	30
play	x	√	√	x	√	x	x	x	√	√	50
pause	x	√	√	x	x	x	x	x	√	√	40
stop	x	x	√	x	x	x	x	x	x	√	20
next	x	√	√	x	x	x	x	x	√	x	30
previous	√	√	√	√	√	√	√	√	√	√	100
											% เฉลี่ย
											45

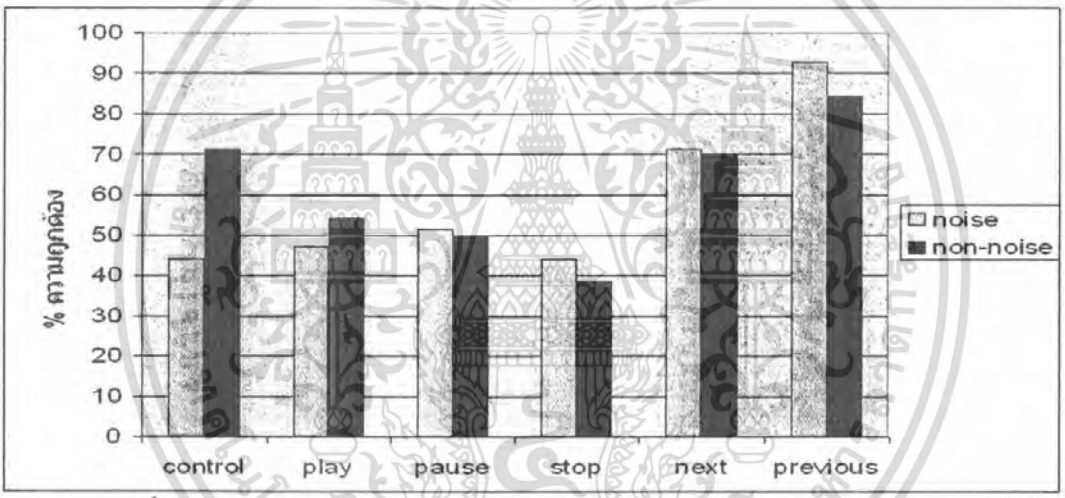


รูปที่ 4.16 กราฟแสดงเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิงคนที่ 7

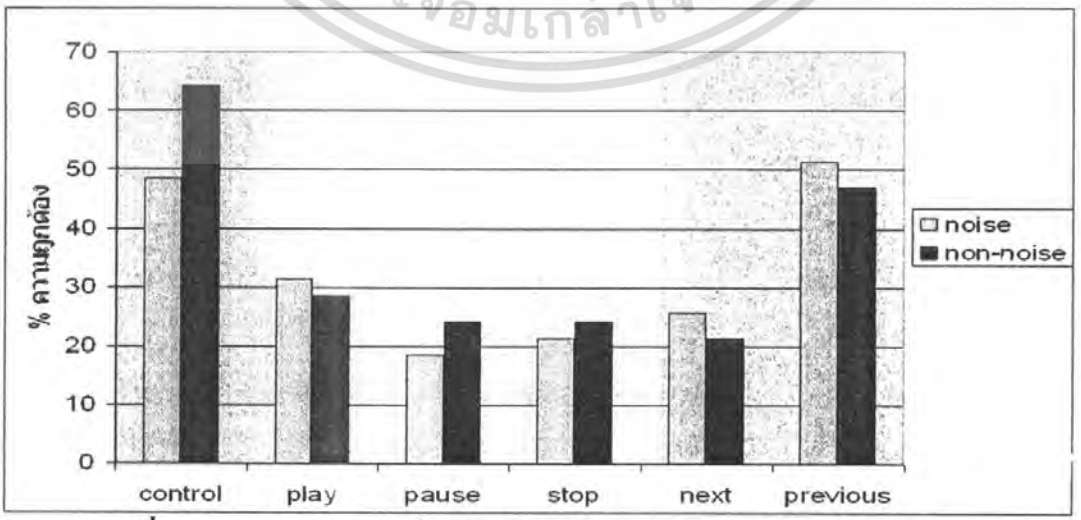
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.29 ผลการเฉลี่ยเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงทั้งผู้ชายและผู้หญิง

เพศ / ค่า	ชาย		หญิง	
	ห้องเงียบ	ห้องมีสัญญาณรบกวน	ห้องเงียบ	ห้องมีสัญญาณรบกวน
control	71.43	44.29	64.29	48.57
play	54.29	47.14	28.57	31.43
pause	50	51.43	24.29	18.57
stop	38.57	44.29	24.29	21.43
next	70	71.43	21.43	25.71
previous	84.29	92.86	47.14	51.43



รูปที่ 4.17 กราฟแสดงผลเฉลี่ยเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้ชาย



รูปที่ 4.18 กราฟแสดงผลเฉลี่ยเปอร์เซ็นต์ความถูกต้องการรู้จำเสียงของผู้หญิง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในส่วนของการทดลองกระบวนการการรู้จำเสียงของผู้หญิงนั้น เราได้ทำการทดลองเช่นเดียวกับผู้ชาย โดยได้ผลการทดลองตามตารางที่ 4.15 ถึงตารางที่ 4.28 เราจะใช้ √ หากเสียงทดสอบกับคำที่แสดงบนจอ LCD นั้นตรงกัน และจะใช้ x หากเสียงทดสอบกับคำที่แสดงบนจอ LCD นั้นไม่ตรงกัน จากผลการทดลองพบว่า ผู้หญิงส่วนใหญ่มีเปอร์เซ็นความถูกต้องของเสียง Previous ค่อนข้างสูงแต่เสียงอื่นนั้นเปอร์เซ็นความถูกต้องค่อนข้างต่ำ หากเราเปรียบเทียบผลการทดลองในทั้งสองกรณีโดยพิจารณาจากกราฟ พบว่า ผลการทดลองไม่แตกต่างกันมากนักเช่นเดียวกับผู้ชาย อาจจะเกิดจากมีสัญญาณรบกวนในสายไมโครโฟนทำให้ได้ผลการทดลองที่ใกล้เคียงกัน และจากการทดลองนั้น ประสิทธิภาพของการรู้จำเสียงของผู้หญิงได้ผลการทดลองเป็นที่ไม่ค่อยน่าพอใจ เนื่องจากมีเปอร์เซ็นความถูกต้องโดยรวมค่อนข้างต่ำ โดยมีเปอร์เซ็นความถูกต้องเฉลี่ย 38.18 %

4.3 สรุปผลการทดลอง

จากการทดลอง พบว่า อัตราการรู้จำเสียงของผู้ชายในห้องเงียบเท่ากับ 62.38 % ในห้องที่มีสัญญาณรบกวนเท่ากับ 63.82 % ในส่วนของผู้หญิงนั้นอัตราการรู้จำเสียงในห้องเงียบเท่ากับ 39 % ในห้องที่มีสัญญาณรบกวนเท่ากับ 37.36 % ดังนั้นอัตราการรู้จำเสียงทั้งหมดเท่ากับ 50.64% โดยคำที่สามารถรู้จำเสียงได้ดีที่สุดในผู้ชาย คือ previous ส่วนผู้หญิง คือ control เราสามารถสรุปผลการทดลองได้ดังนี้

1. เสียงพูดของคนๆเดียวกัน คำเดียวกันพบว่าเสียงที่พูดออกมาแต่ละครั้งไม่เหมือนกัน

2.จำนวนเสียงที่นำมาสร้างแบบอ้างอิงมีผลต่อความถูกต้องของการรู้จำเสียง โดยยิ่งจำนวนเสียงต้นแบบมาก เปอร์เซ็นต์ความถูกต้องจะน้อยลง

3. สำเนียงเสียงที่ใกล้เคียงกันมีผลต่อความถูกต้องของการรู้จำเสียง จากการทดลอง เสียงที่ใกล้เคียงกัน คือ เสียง play กับ pause (เสียงพยัญชนะ พอ) พบว่าเปอร์เซ็นต์การรู้จำเสียงค่อนข้างต่ำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
บทวิจารณ์และบทสรุป

5.1 สรุป

โครงงานนี้ได้นำเสนอ การรู้จำเสียงโดยใช้ dsPIC30F6014 เป็นตัวประมวลผลกลาง เนื่องจากผู้จัดทำโครงงงานนี้มีความเห็นว่า มีคุณสมบัติและความสามาถที่เพียงพอในการประมวลผล ในส่วนกระบวนการรู้จำเสียงนั้น จะรับสัญญาณเสียงผ่านตัวทางไมโครโฟน แล้วนำสัญญาณเสียงไปผ่านตัว Si3000 ซึ่งเป็นตัวแปลงสัญญาณเสียงอนาลอกเป็นสัญญาณเสียงดิจิตอล ซึ่งมีอัตราการสุ่มสัญญาณที่ 12 kHz. ซึ่งเป็นอัตราการสสุ่มสัญญาณที่สูงสุดของตัว Si3000 และเป็นอัตราการส่งสัญญาณที่เหมาะสมในการทำโครงงานนี้ เมื่อได้สัญญาณเสียงดิจิตอลแล้ว จะนำสัญญาณเสียงไปผ่านกระบวนการ MFCC เพื่อหาตัวแทนของเสียง เนื่องจากจะไม่นำสัญญาณเสียงทั้งหมดมาประมวลผล เพราะจะทำให้ใช้เวลาในการประมวลผลนานและต้องใช้หน่วยความจำมาก เมื่อได้ตัวแทนของเสียงแล้ว จะนำไปเปรียบเทียบกับเสียงต้นแบบ โดยใช้กระบวนการ HMM เพื่อหาว่าเสียงทดสอบใกล้เคียงกับเสียงต้นแบบใดมากที่สุด แล้วนำไปแสดงผลบนจอ LCD เพื่อบอกว่าเสียงทดสอบเป็นเสียงใด

ในส่วนของการหาประสิทธิภาพของการรู้จำเสียงนั้น ได้ทำการทดลอง ใน 2 กรณี คือ

1.) ทดลองในห้องเงียบ

2.) ทดลองในห้องที่มีสัญญาณรบกวน

ซึ่งผลการทดลองโดยรวมถือว่าเป็นที่น่าพอใจ แต่ถ้าดูอย่างละเอียดจะเห็นว่า ผลกาทดลองของผู้ชายได้ผลการทดลองที่ดีกว่าผู้หญิง อาจเป็นเพราะเสียงของผู้ชายใกล้เคียงกับเสียงต้นแบบมากกว่า จึงควรมีการวิจัยและพัฒนาต่อไป เพื่อให้ประสิทธิภาพในการรู้จำเสียงโดยใช้ dsPIC30F6014 เป็นตัวประมวลผลดีขึ้น

5.2 ปัญหาที่พบในการวิจัย

- ปัญหาที่เกิดจากวิธีการ ผู้ทำการทดลองวางตำแหน่งของไมโครโฟนนั้นไม่ถูกต้องก็ส่งผลให้การรับสัญญาณเสียงที่ไม่ถูกต้องเข้าไป เช่น การไมโครโฟนใกล้กับปากเกินไปทำให้มีการพ่นลมเข้าไปในไมโครโฟนด้วย หรือวางใกล้กับจมูกเกินไปก็จะทำให้มีเสียงลมหายใจเข้าไปหรืออาจทำให้เสียงที่ออกทางปากเข้าน้อยกว่าเสียงนาสิก หรือวางไมโครโฟนในระดับต่ำกว่าปากก็จะทำให้ไมโครโฟนไม่สามารถรับเสียงนาสิกได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ปัญหาที่เกิดจากอุปกรณ์ คือในการใช้งานแต่ละครั้งต้องมีการปรับอัตราขยายของไมโครโฟนอยู่เรื่อย ๆ เนื่องมาจากคุณภาพไมโครโฟน การใช้อุปกรณ์ที่มีคุณลักษณะและคุณภาพต่างกันมากๆ จะส่งผลต่อความถูกต้อง เช่น ใช้ไมโครโฟนคนละตัวกับที่ใช้ในครั้งแรก ก็จะส่งผลกระทบทั้งสิ้น

- ปัญหาเกิดจากผู้ทดสอบ คือ ผู้ทดสอบเกิดความประหม่าขณะทำการทดสอบทำให้การพูดไม่เป็นไปตามธรรมชาติ และการพูดที่ผิดหลักภาษา ดังนั้น การออกเสียงของผู้ใช้จะต้องมีความชัดเจน

5.3 ข้อเสนอแนะ

- เนื่องจากเสียงต้นแบบเป็นเสียงภาษาอังกฤษแต่ผู้ทำการทดลองเป็นคนไทย จึงทำให้เปอร์เซ็นความถูกต้องของการรู้จำเสียงค่อนข้างต่ำ จึงควรวิจัยและพัฒนาเสียงต้นแบบที่เป็นภาษาไทยขึ้นมาซึ่งจะช่วยให้เปอร์เซ็นความถูกต้องสูงขึ้น

- ควรใช้ไมโครโฟนที่มีคุณภาพสูง เพื่อในการป้องกันสัญญาณรบกวนที่เกิดจากสายส่งสัญญาณจะช่วยเปอร์เซ็นความถูกต้องสูงขึ้น

- ก่อนการทดลองเก็บข้อมูลจริงของผู้ทำการทดลอง ควรมีการซ้อมเพื่อความไม่ประหม่าที่เกิดขึ้นในตอนทดลองจะช่วยให้ผู้ทำการทดลองออกเสียงได้เป็นธรรมชาติในตอนทดลอง

- การเลือกเสียงต้นแบบควรหลีกเลี่ยงเสียงที่มีลักษณะใกล้เคียงกันจะช่วยให้เปอร์เซ็นต์ความถูกต้องสูงขึ้น

- จำนวนคำของเสียงต้นแบบไม่ควรมากเกินไปเพราะจะทำให้เปอร์เซ็นต์ความถูกต้องต่ำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] Lawrence Rabiner and Biing-Hwang Juang, "Fundamentals of Speech Recognition", Prentice-Hall signal processing series, 1993.

[2] F.J Owens, "Signal Processing of Speech", The Macmillan Press , 1993.

[3] Li Tan and Montri Karnjanadecha, "MODIFIED MEL-FREQUENCY CEPSTRUM COEFFICIENT", Department of Computer Engineering Faculty of Engineering, Prince of Songkhla University.

[4] NUTTAKORN THUBTHONG and BOONSERM KIJSIRIKUL, "TONE RECOGNITION OF CONTINUOUS THAI SPEECH UNDER TONAL ASSIMILATION AND DECLINATION EFFECTS USING HALF-TONE MODEL", Department of Computer Engineering, Chulalongkorn University.

[5] Prom-On S., Techathammawong S. and Piriyabenchawat P., "Consonant Phoneme Recognition using Hidden Markov Model and Vowel Classification using Formant", Bachelor Degree Project, Department of Computer Engineering, King Mongkut's University of Technology Thonburi, 2001.

[6] Mangklawisarut K., Boonchoochuey C. and Thouyjaroen T., "Thai Consonant Speech Signal Analysis using Continuous Wavelet Transform", Bachelor Degree Project, Department of Computer Engineering, King Mongkut's University of Technology Thonburi, 2000.

[7] Alpay Koc, "Acoustic Feature Analysis for Robust Speech Recognition", BS. In EE., Bilkent University.

[8] Hung-Ju Huang1 and Chun-Nan HsuRecognizing, "100 Speakers using Homologous Naïve Bayes", Department of Computer and Information Science, National Chiao-Tung University.

[9] รวินทร์ วิรัชพินทุ, สุวารี เหลือเพิ่มพร "ระบบสั่งงานด้วยเสียง" ปริญญานิพนธ์วิศวกรรมศาสตรบัณฑิต สาขาวิชาคอมพิวเตอร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.2546.

[10] N. Jittiwarangkul, S. Luksaneeyanawin, S. Jitapunkul, "Syllable Segmentation Algorithm for Thai Connected Speech", Department of Electrical Engineering Faculty of Engineering, Chulalongkorn University.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

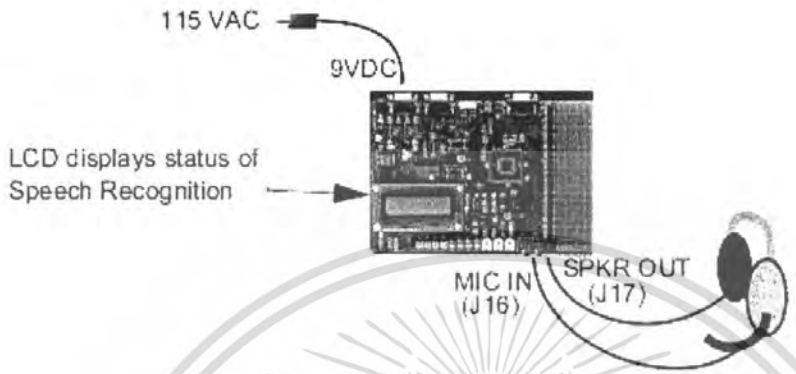
ภาคผนวก ก.

การทดลองการใช้งานบอร์ด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีการใช้งานบอร์ด dsPICdem

1. เซ็ทอุปกรณ์ตามรูปที่ ก.1



รูปที่ ก.1 แสดงการติดตั้งอุปกรณ์

2. ทันทีที่จ่ายไฟให้กับบอร์ดจะปรากฏข้อความบนจอ LCD ตามรูปที่ ก.2

รูปที่ ก.2 หน้าจอแรกของโปรแกรม

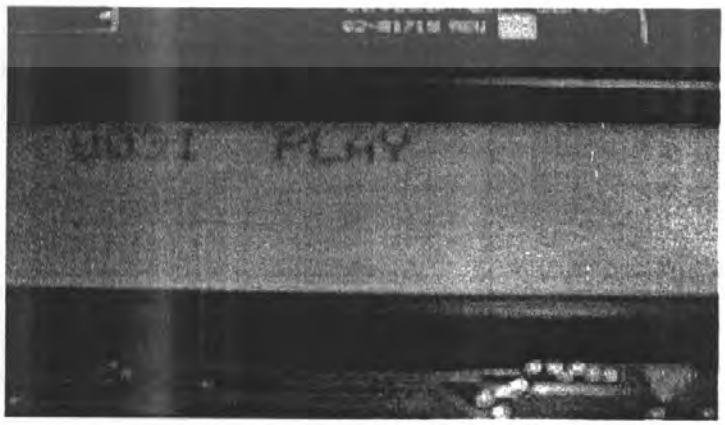
3. ให้ผู้ทำการทดลองออกเสียง Keyword คำว่า " Control " หากผลการเปรียบเทียบการรู้จำเสียงถูกต้องจะแสดงผลตามรูปที่ ก.3 และหากเสียง Keyword ไม่ถูกต้องจะแสดงผลดังรูปที่ ก.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ก.3 แสดงหน้าจอเมื่อ Keyword ถูกต้อง

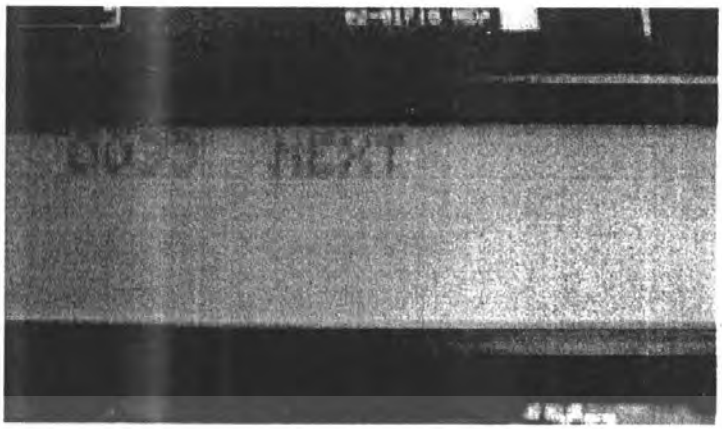
รูปที่ ก.4 แสดงหน้าจอเมื่อ Keyword ไม่ถูกต้อง

4. หลังจาก Keyword ถูกต้องแล้ว ให้ผู้ทดลองออกเสียงตามคำใน Library



รูปที่ ก.5 แสดงหน้าจอเมื่อออกเสียง Play ถูกต้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

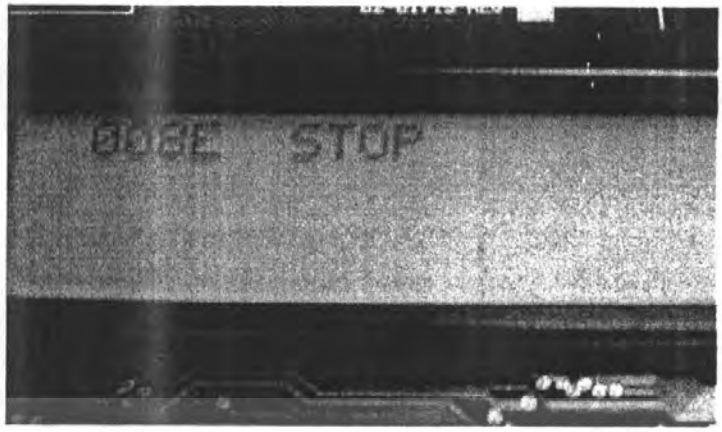


รูปที่ ก.6 แสดงหน้าจอเมื่อออกเสียง Next ถูกต้อง

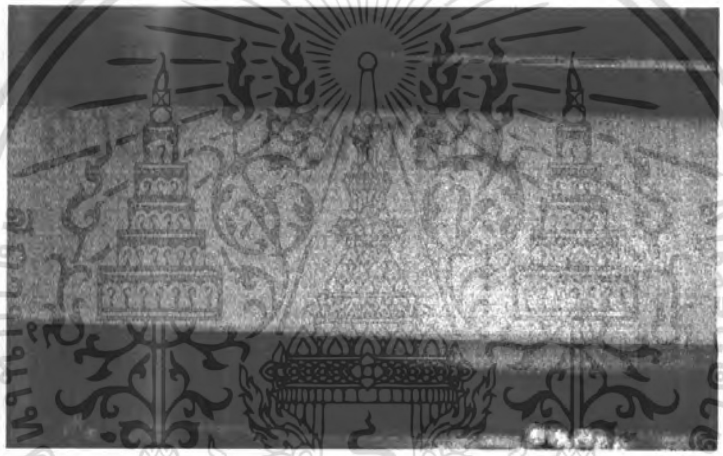
รูปที่ ก.7 แสดงหน้าจอเมื่อออกเสียง Pause ถูกต้อง

รูปที่ ก.8 แสดงหน้าจอเมื่อออกเสียง Previous ถูกต้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ก.9 แสดงหน้าจอเมื่อออกเสียง Stop ถูกต้อง



รูปที่ ก.10 แสดงหน้าจอเมื่อออกเสียง Control ถูกต้อง

รูปที่ ก.11 แสดงหน้าจอเมื่อออกเสียงไม่ถูกต้องตามคำใน Library

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

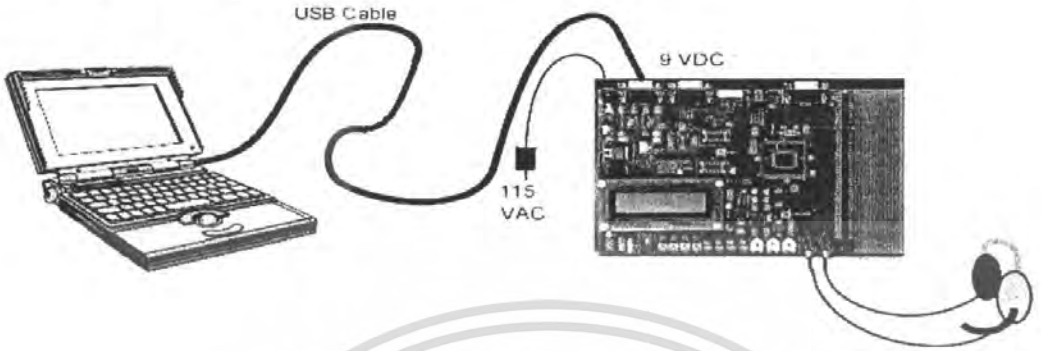
ภาคผนวก ข

การใช้งานโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีการใช้งานโปรแกรม MEDIA PLAYER

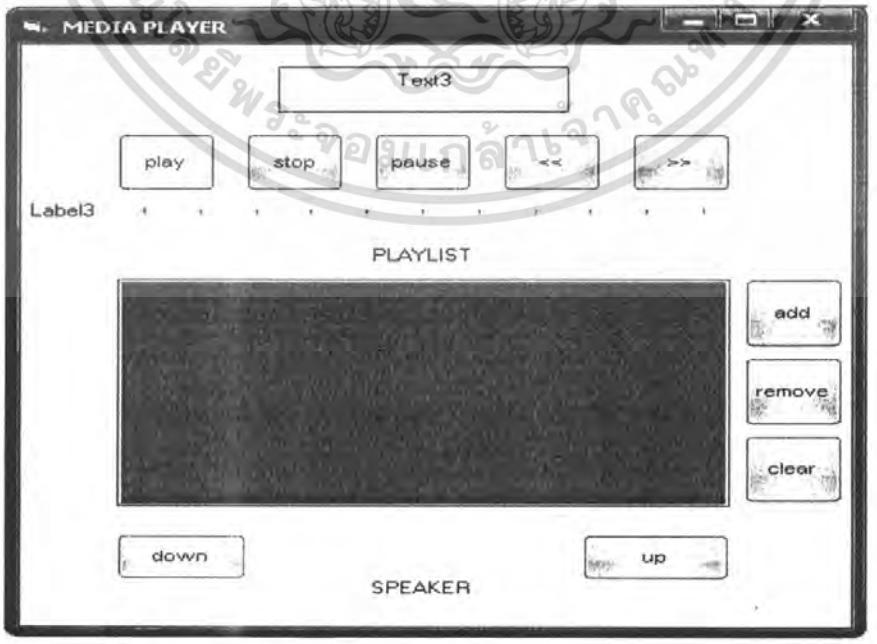
1. ติดตั้งอุปกรณ์ตามรูปที่ ข.1



รูปที่ ข.1 แสดงการติดตั้งอุปกรณ์

2. เรียกโปรแกรมโดยดับเบิลคลิกที่ไอคอน multimdia ดังรูปที่ ข.2 จะปรากฏฟอร์ม MEDIA PLAYER ดังรูปที่ ข.3 ตามลำดับ

รูปที่ ข.2 แสดงรูปไอคอน multimedia



รูปที่ ข.3 แสดงโปรแกรม MEDIA PLAYER

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. เลือกเพลงลงใน PLAYLIST โดยการกดปุ่ม add จะปรากฏฟอร์มให้เลือกเพลงตามใจชอบดังรูปที่ ข.4 สามารถลบเพลงที่ไม่ต้องการได้โดยการกด remove และสามารถลบรายการเพลงทั้งหมดโดยการกด clear

รูปที่ ข.4 แสดงตัวอย่างการเลือกเพลง

4. เล่นเพลง เมื่อทำการเลือกเพลงแล้วดังรูปที่ ข.5

ในการเล่นเพลงนั้น สามารถเล่นเพลงได้สองทาง คือ

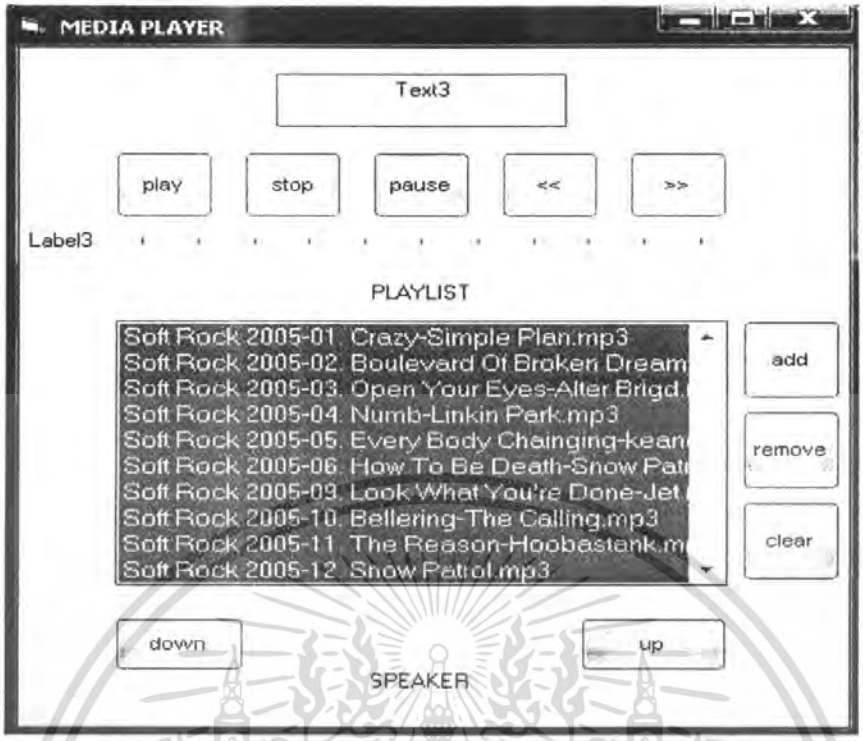
- โดยการใช้เมาส์
- โดยการใช้เสียง

ในส่วนของการใช้เสียงนั้น เสียงที่ใช้ในการเล่นเพลงมีดังต่อไปนี้

a. เสียง Play -เมื่อต้องการเล่นเพลง
b. เสียง Pause –เมื่อต้องการหยุดเพลงชั่วคราว
c. เสียง Stop – เมื่อต้องการจบเพลงนั้น
d. เสียง Next – เมื่อต้องการข้ามไปเล่นเพลงต่อไป
e. เสียง Previous- เมื่อต้องการกลับไปเล่นเพลงที่ผ่านมา

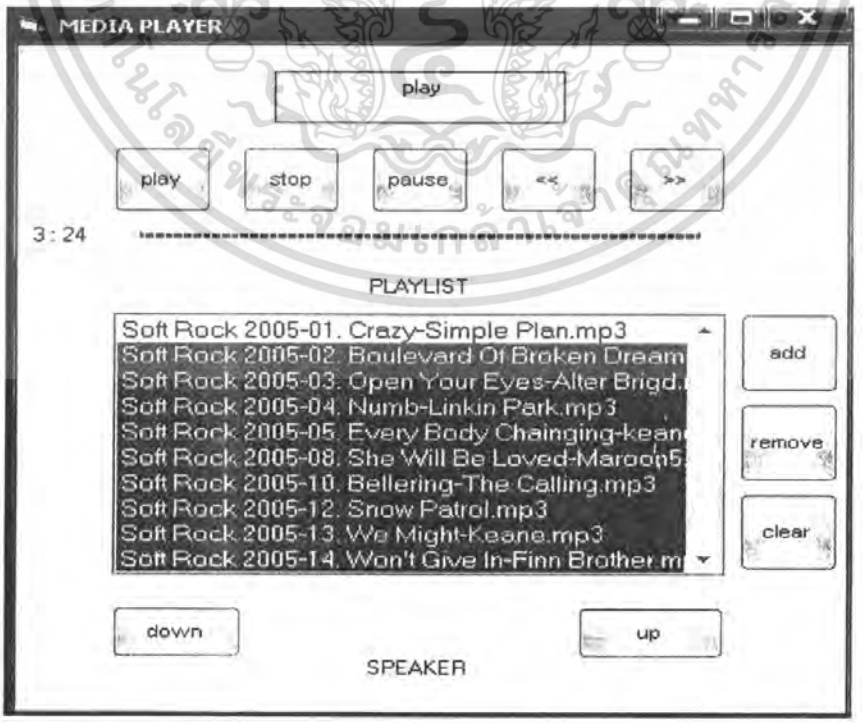
เสียงแต่ละเสียงที่พูดไปนั้นจะไปปรากฏที่ข้อความด้านบน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ข.5 แสดงโปรแกรม MEDIA PLAYER ที่เลือกเพลงเรียบร้อยและพร้อมที่จะเปิดเพลง

5. เพิ่ม-ลดระดับเสียงด้วยการเลื่อนแถบตรง SPEAKER หรือกด down ดังรูปที่ ข.6



รูปที่ข.6 แสดงโปรแกรม MEDIA PLAYER ขณะปรับระดับเสียงตอนเปิดเพลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

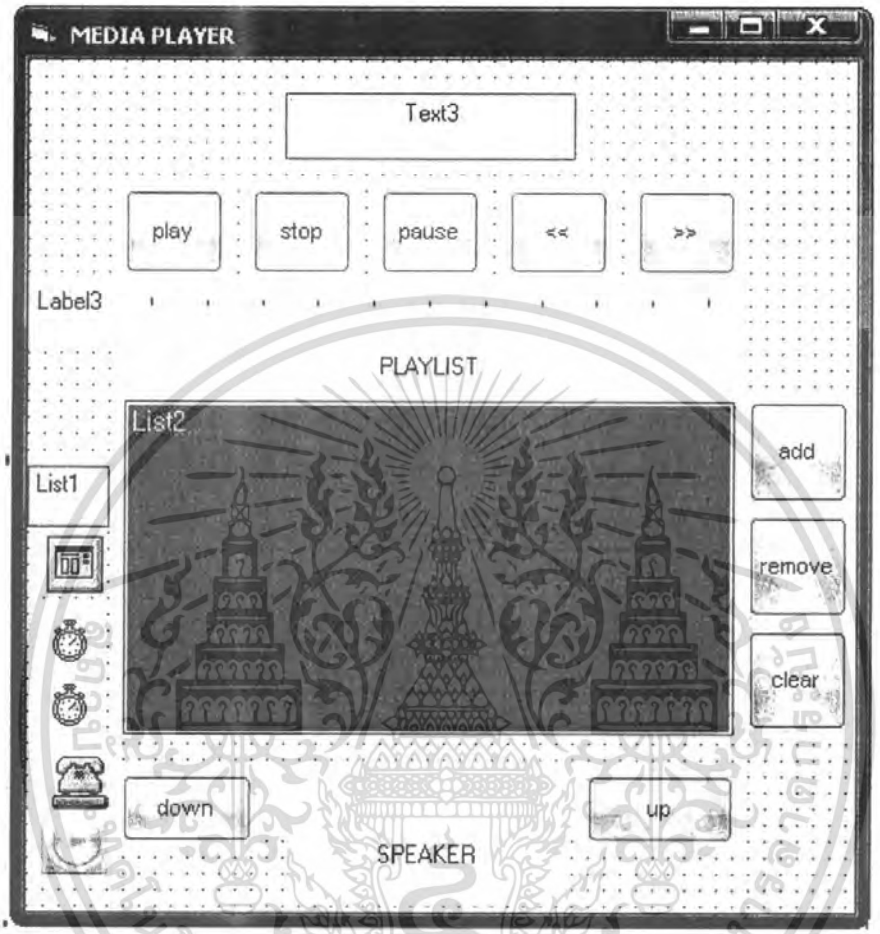
ภาคผนวก ค

โปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การสร้างโปรแกรม MEDIA PLAYER

รูปแบบฟอร์มของโปรแกรม MEDIA PLAYER ดังรูปที่ ค.1



รูปที่ ค.1 ตัวอย่างโปรแกรม MEDIA PLAYER

```
Dim  action As String, t As String, actio As Integer
Private Sub Command1_Click()
 If actio = 0 Then
WindowsMediaPlayer1.Url = List1.Text
WindowsMediaPlayer1.Controls.Play
 Else
 WindowsMediaPlayer1.Controls.Play
 End If
End Sub
Private Sub Command10_Click()
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
Slider2.Value = Slider2.Value - 5
WindowsMediaPlayer1.Settings.Volume = Slider2.Value
End Sub
Private Sub Command2_Click()
actio = 0
WindowsMediaPlayer1.Controls.Stop
End Sub
Private Sub Command3_Click()
actio = 1
WindowsMediaPlayer1.Controls.Pause
End Sub
Private Sub Command4_Click()
Dim i As Integer
CommonDialog1.InitDir = "D:\song"
CommonDialog1.Filter = "mp3Files|*.mp3| Wav files|*.wav| All files|*.*"
CommonDialog1.ShowOpen
t = CommonDialog1.FileName
List1.AddItem CommonDialog1.FileName
For i = 0 To 3
t = Right(t, Len(t) - InStr(t, "\"))
Next
List2.AddItem Right(t, Len(t) - InStr(t, "\"))
End Sub
Private Sub Command5_Click()
Slider2.Value = Slider2.Value + 5
WindowsMediaPlayer1.Settings.Volume = Slider2.Value
End Sub
Private Sub Command6_Click()
If List2.ListIndex >= 0 Then
   List1.RemoveItem List1.ListIndex
   List2.RemoveItem List2.ListIndex
   Else
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    'MsgBox "can't remove item from list."
    End If
End Sub
Private Sub Command7_Click()
List1.Clear
List2.Clear
End Sub
Private Sub Command8_Click()
 actio = 0
 If List2.ListCount <> 0 Then
     If List2.ListIndex = List2.ListCount - 1 Then
     List2.ListIndex = 0
     List1.ListIndex = List2.ListIndex
      Else
      List2.ListIndex = List2.ListIndex + 1
      List1.ListIndex = List2.ListIndex
      End If
   Else
   MsgBox "no list."
   End If
End Sub
Private Sub Command9_Click()
  actio = 0
If List2.ListCount <> 0 Then
 If List2.ListIndex = 0 Then
      List2.ListIndex = List2.ListCount - 1
      List1.ListIndex = List2.ListIndex
      Else
      List2.ListIndex = List2.ListIndex - 1
      List1.ListIndex = List2.ListIndex
      End If
      Else
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    MsgBox "no list."
    End If
End Sub
Private Sub Form_Load()
mscreceiver.CommPort = 5
mscreceiver.Settings = "9600,n,8,1"
mscreceiver.PortOpen = True
Timer1.Interval = 100
Timer1.Enabled = True
Timer2.Interval = 1000
Timer2.Enabled = False
Slider2.Value = 50
WindowsMediaPlayer1.Settings.Volume = Slider2.Value
End Sub
Private Sub Label4_Click()
End Sub
Private Sub List2_DblClick()
List1.ListIndex = List2.ListIndex
WindowsMediaPlayer1.Url = List1.Text
Timer2.Enabled = True
End Sub
Private Sub List2_Click()
List1.ListIndex = List2.ListIndex
actio = 0
End Sub
Private Sub Slider1_Scroll()
WindowsMediaPlayer1.Controls.CurrentPosition = Slider1.Value
End Sub
Private Sub Slider2_Click()
WindowsMediaPlayer1.Settings.Volume = Slider2.Value
End Sub
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
Private Sub Timer1_Timer()
mscreceiver.InputLen = 0
Data = mscreceiver.Input
If Data = "play " Then
    Text3.Text = "play"
    If actio = 0 Then
      WindowsMediaPlayer1.Url = List1.Text
      WindowsMediaPlayer1.Controls.Play
    Else
      WindowsMediaPlayer1.Controls.Play
    End If
   ElseIf Data = "close " Then
   Text3.Text = "close"
   ElseIf Data = "stop " Then
   Text3.Text = "stop"
   WindowsMediaPlayer1.Controls.Stop
   actio = 0
   ElseIf Data = "pause " Then
   Text3.Text = "pause"
   WindowsMediaPlayer1.Controls.Pause
   actio = I
   ElseIf Data = "control " Then
   Text3.Text = "control"
   ElseIf Data = "unknown " Then
   Text3.Text = "unknown"
   ElseIf Data = "previous " Then
   actio = 0
   Text3.Text = "previous"
   WindowsMediaPlayer1.Controls.Stop
   If List2.ListCount <> 0 Then
     If List2.ListIndex = 0 Then
       List2.ListIndex = List2.ListCount - 1
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        List1.ListIndex = List2.ListIndex
        Else
        List2.ListIndex = List2.ListIndex - 1
        List1.ListIndex = List2.ListIndex
        End If
        WindowsMediaPlayer1.Url = List1.Text
        WindowsMediaPlayer1.Controls.Play
        Else
       MsgBox "no list."
       End If
      ElseIf Data = "next " Then
      actio = 0
      Text3.Text = "next"
      WindowsMediaPlayer1.Controls.Stop
      If List2.ListCount <> 0 Then
       If List2.ListIndex = List2.ListCount - 1 Then
       List2.ListIndex = 0
       List1.ListIndex = List2.ListIndex
        Else
        List2.ListIndex = List2.ListIndex + 1
        List1.ListIndex = List2.ListIndex
        End If
        WindowsMediaPlayer1.Url = List1.Text
        WindowsMediaPlayer1.Controls.Play
     Else
     MsgBox "no list."
     End If
      ElseIf Data = "open " Then
      Text3.Text = "open"
   End If
End Sub
Private Sub Timer2_Timer()
```

เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
 Slider1.Max = WindowsMediaPlayer1.currentMedia.Duration
 Slider1.Value = WindowsMediaPlayer1.Controls.CurrentPosition
Dim lengths, lengths1, min, sec As Long
lengths = WindowsMediaPlayer1.Controls.currentItem.Duration
tinseconden = WindowsMediaPlayer1.Controls.CurrentPosition
lengths1 = lengths - tinseconden
min = lengths1 \ 60
sec = lengths1 - min * 60
Label3.Caption = min & " : " & sec
min = tinseconden \ 60
sec = tinseconden - (min * 60)
 If WindowsMediaPlayer1.Controls.CurrentPosition + 2 >=
WindowsMediaPlayer1.Controls.currentItem.Duration + 1 Then
 List2.ListIndex = List2.ListIndex + 1
    List1.ListIndex = List2.ListIndex
    End If
End Sub
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้