ห้องสมุดคณะเทคโนโลยีสารสนเทศ พระจอมเกล้าลาดกระบัง

ระบบควบคุมการทำงานภายในบ้านผ่านเสียงและมือถือ

HOME AUTOMATION SYSTEM VIA VOICE AND MOBILE

H006751

โดย

อิมรอน สมานพิทักษ์

IMRON SAMARNPITAK

อาจารย์ที่ปรึกษา

ดร. ปานวิทย์ ธุวะนุติ

เลขหมู่.....
เลขทะเบียน.....6751
วัน,เดือน,ปี.....11 ตค 2555

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

รายงานนี้เป็นส่วนหนึ่งของวิชาการศึกษาอิสระ 2
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ภาคเรียนที่ 2 ปีการศึกษา 2553

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

HOME AUTOMATION SYSTEM VIA VOICE AND MOBILE

IMRON SAMARNPITAK

A REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS OF THE COURSE

INDEPENDENT STUDY 2

MASTER OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2/ 2010

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2011

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใบรับรองการศึกษาอิสระ 2 (Independent Study 2)

เรื่อง

ระบบควบคุมการทำงานภายในบ้านผ่านเสียงและมือถือ

HOME AUTOMATION SYSTEM VIA VOICE AND MOBILE

นายอิมรอน สมานพิทักษ์

รหัสประจำตัว 52660519

ขอรับรองว่ารายงานฉบับนี้ ข้าพเจ้าไม่ได้คัดลอกมาจากที่ใด
รายงานฉบับนี้ได้รับการตรวจสอบและอนุมัติให้เป็นส่วนหนึ่งของ
การศึกษาวิชาอิสระ หลักสูตรวิทยาศาสตรมหาบัณฑิต (เทคโนโลยีสารสนเทศ)
ภาคเรียนที่ 2 ปีการศึกษา 2553

.....อาจารย์ที่ปรึกษา
(ดร. ปานวิทย์ ธุวะนุติ)

.....กรรมการสอบ
(รศ.ดร. นพพร โชติกำธร)

.....กรรมการสอบ
(ดร.นล เปรมัษเฐียร)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อ	ระบบควบคุมการทำงานภายในบ้านผ่านเสียงและมือถือ
นักศึกษา	นายอิมรอน สมานพิทักษ์
รหัสนักศึกษา	52660519
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
แขนงวิชา	วิทยาการสารสนเทศ
ปีการศึกษา	2553
อาจารย์ที่ปรึกษา	ดร.ปานวิทย์ ธุวะนุติ

บทคัดย่อ

รายงานฉบับนี้นำเสนอการพัฒนาระบบควบคุมการทำงานภายในบ้านสำหรับผู้ที่ขาดภาวะการควบคุมอุปกรณ์ภายในบ้านได้สะดวก เช่น ผู้สูงอายุ ผู้พิการ ผู้ป่วย หรือ ผู้ที่ไม่สามารถควบคุมอุปกรณ์ได้ด้วยตนเองโดยตรง การพัฒนาจะแบ่งประเภทการควบคุมอยู่ 2 ประเภท คือ การควบคุมโดยใช้เสียง และ การควบคุมโดยใช้มือถือ และกำหนดให้ควบคุมอุปกรณ์ทั้งหมดแปดชิ้น

ประเภทการควบคุมโดยใช้เสียงจะเป็นการควบคุมโดยไม่ขึ้นกับผู้พูด โดยขั้นตอนกระบวนการรู้จำเสียงพูดประกอบด้วย การประมวลผลสัญญาณเสียงเบื้องต้น การสกัดลักษณะเด่นของคำด้วยวิธีการประมาณพันธะเชิงเส้น (LPC) และ การเรียนรู้ลักษณะของคำใช้เครือข่ายประสาทเทียม (Neural Network) โดยมีการอัลกอริทึมแบบแพร่กระจายกลับ ส่วนขั้นตอนในกระบวนการตัดสินใจจะใช้เครือข่ายประสาทเทียมโดยมีอัลกอริทึมแบบแพร่ไปข้างหน้าในการวิเคราะห์สัญญาณเสียงที่เข้ามาและส่งค่าผลลัพธ์ที่ได้เข้าไมโครคอนโทรลเลอร์เพื่อควบคุมอุปกรณ์ภายในบ้านต่อไป ประเภทการควบคุมโดยใช้มือถือจะเป็นการควบคุมที่ผู้ใช้งานสามารถควบคุมอุปกรณ์ภายในบ้านจากระยะไกลโดยผ่านช่องทางเครือข่ายโทรศัพท์มือถือ โดยเทคโนโลยีที่นำมาใช้งานคือเทคโนโลยีเว็บ โดยสามารถส่งผลของคำสั่งผ่านทางเครือข่ายผู้ให้บริการมือถือเพื่อควบคุมและตรวจสอบอุปกรณ์ภายในบ้านตามที่ต้องการได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Title	Home Automation System via Voice and Mobile
Student	Mr.Imron Samarnpitak
Student ID.	52660519
Degree	Master of Science
Program	Information Technology
Major	Information Science
Academic Year	2010
Advisor	Dr. Panwit Thuwanut

ABSTRACT

This report describes the development of control systems work at home for the lack of state control devices within the home for example the elderly person ,persons with disability , patients or who can not control their own devices. The development will be classified in two types of control is to control using voice and control using cell phones. And assigned to control all eight pieces of equipment.

Type of control equipment using voice, regardless of the speaker The process steps of voice recognition which were consisted of The basic audio signal processing. , Feature Extraction by Linear Predictive Coding(LPC) and recognition to use artificial neural network with a back propagation algorithm.The steps in the decision-making process to use artificial neural network with a feed forward algorithm in analyzing the incoming audio signal and send the results to the microcontroller to control the device to your home. Type of control equipment using mobile phone . User can control home devices remotely via the network by mobile phone .The Web technology is used to control Home devices. The results of the command to send through a network of mobile service providers to control and monitor devices within the home.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

โครงงานหัวข้อการควบคุมอุปกรณ์ภายในบ้านผ่านเสียงและมือถือ ประสบความสำเร็จได้อย่างดี โดยท่านผู้ที่ทำให้โครงงานนี้ดำเนินการสำเร็จตามวัตถุประสงค์และเวลาที่ได้ตั้งไว้ คือ ท่านอาจารย์ที่ปรึกษาในโครงงาน ดร.ปานวิทย์ ธุวะนุติ ซึ่งให้คำแนะนำ คำปรึกษาปัญหา ข้อสงสัย ข้อติดขัด ในตัวรายละเอียดของโครงงาน ในทุกช่วงเวลา

สำหรับท่านอาจารย์สามท่านที่ขอกล่าวถึงคือ รศ.ดร.นพพร โชติกำธร รศ.ดร.โชติพัชร์ ภรณวลัย และดร.นล เปรมัษเฐียร ซึ่งเป็นกรรมการการนำเสนอโครงงาน ซึ่งทุกท่านที่กล่าวถึงได้ให้แนวทางการพัฒนา ชี้จุดบกพร่องในโครงงาน เพื่อเติมเต็มทำให้โครงงานสมบูรณ์ตามวัตถุประสงค์ที่ได้ตั้งไว้

สำหรับสุดท้ายนี้ ต้องกราบขอบพระคุณ บิดา มารดา ที่เลี้ยงดู ให้กำลังใจ ชี้นำแนวทางในการดำเนินชีวิต และกราบขอบพระคุณอาจารย์ในคณะเทคโนโลยีสารสนเทศทุกท่าน ที่ได้อบรมสั่งสอน ประสิทธิ์ประสาทวิชาความรู้ความสามารถ คุณธรรม จริยธรรม และขอบคุณเพื่อนร่วมรุ่นทุกท่าน ที่คอยช่วยเหลือในยามที่มีปัญหาเสมอมา

อิมรอน สมานพิทักษ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย I

บทคัดย่อภาษาอังกฤษ II

กิตติกรรมประกาศ III

สารบัญ IV

สารบัญตาราง VII

สารบัญรูป VIII

บทที่ 1 บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์ของการศึกษา 1

1.3 ขอบเขตของการศึกษา 2

1.4 ขั้นตอนของการศึกษา 2

1.5 ประโยชน์ที่คาดว่าจะได้รับ 3

1.6 ส่วนประกอบของรายงาน 3

บทที่ 2 ทฤษฎีและหลักการที่เกี่ยวข้อง

2.1 ทฤษฎีเสียง 4

2.1.1 อวัยวะที่สัมพันธ์กับเสียงพูด 4

2.1.2 ลักษณะเสียงพูดของมนุษย์ 6

2.2 การแปลงรูปแบบสัญญาณเสียงพูด 7

2.3 กระบวนการในการรู้จำเสียง 8

2.3.1 การประมวลผลสัญญาณเบื้องต้น (Signal Preprocessing) 9

2.3.2 การสกัดลักษณะเด่น (Feature Extraction) 14

2.3.3 เครือข่ายประสามเทียม 16

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

หน้า

บทที่ 3 การออกแบบระบบ

3.1 การออกแบบระบบรู้จำเสียงเบื้องต้น 28

3.1.1 การออกแบบและสร้างคำเพื่อเป็นส่วนที่ใช้ในการรู้จำและทดสอบ 28

3.1.2 รายละเอียดของไฟล์เสียงข้อมูล 28

3.1.3 เครื่องมือและอุปกรณ์ที่ใช้ในการบันทึกไฟล์เสียง 28

3.2 ขั้นตอนกระบวนการรู้จำเสียง 29

3.2.1 กระบวนการประมวลผลสัญญาณเสียงเบื้องต้น 30

3.2.2 การสกัดลักษณะเด่นของคำด้วยวิธีการประมาณพันธะเชิงเส้น (Linear Predictive Coding)..... 35

3.2.3 เครือข่ายประสาทเทียม (Neural Network) 36

3.2.4 กระบวนการตัดสินใจ 39

3.2.5 การออกแบบการทดสอบของการรู้จำด้านเสียง 41

3.3 การออกแบบระบบของการควบคุมอุปกรณ์ด้วยมือถือ 42

3.3.1 เครื่องมือที่ใช้ 42

3.3.2 ผังการทำงานหลักของระบบ 42

3.3.3 ยูสเคสไดอะแกรม (Use case diagram) 43

3.3.4 แอคทิวิตีไดอะแกรม (Activity diagram) 45

3.4 การออกแบบส่วนของฮาร์ดแวร์ 46

3.4.1 รายละเอียดของอุปกรณ์ที่ใช้และการเชื่อมต่อ 46

3.4.2 ผังการทำงานของไมโครคอนโทรลเลอร์ 47

บทที่ 4 การทดลองและวิเคราะห์ผลการทดลอง

4.1 ข้อมูลที่ใช้ทดสอบ 49

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

หน้า

4.2 การหาจุดสิ้นสุดของคำและการตัดคำ 49
4.3 การปรับบรรทัดฐานทางเวลา 51
4.4 การสกัดลักษณะเด่นของคำ 51
4.5 เครือข่ายประสาทเทียม 53
4.6 ส่วนติดต่อผู้ใช้งานของระบบการควบคุมอุปกรณ์ด้วยเสียง (Voice User-Interface) 54
4.7 ส่วนติดต่อผู้ใช้งานของระบบการควบคุมอุปกรณ์ผ่านมือถือ (Web User-Interface) 58
บทที่ 5 สรุปผลและข้อเสนอแนะ
5.1 สรุปผลการทดลอง 60
5.2 ปัญหาและอุปสรรค 61
5.3 ข้อเสนอแนะ 61
บรรณานุกรม 62
ประวัติผู้เขียน 63

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ หน้า

2.1 แบบจำลองของสมองมนุษย์เทียบกับเครือข่ายประสาทเทียม 19

3.1 คำอธิบายยูสเคสไดอะแกรมของ Control Device 43

3.2 คำอธิบายยูสเคสไดอะแกรมของ Read Device Status 44

3.3 คำอธิบายยูสเคสไดอะแกรมของ Login 44

4.1 ผลสรุปการทดลองในการหาจุดสิ้นสุดของคำ 50

4.2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ในค่าอันดับต่างๆ 51

4.3 ตารางผลการทดสอบเครือข่ายประสาทเทียม 53

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่ หน้า

2.1 แสดงภาพของอวัยวะในร่างกายที่มีความสัมพันธ์กับเสียงพูด 4

2.2 ภาพแสดงกระบวนการสุ่มสัญญาณ 8

2.3 แสดงแผนภาพขั้นตอนการทำงานของกระบวนการรู้จำ 9

2.4 ฟังก์ชั่นกรอบสัญญาณสี่เหลี่ยม 11

2.5 ฟังก์ชั่นกรอบสัญญาณแฮมมิ่ง 11

2.6 ฟังก์ชั่นกรอบสัญญาณแฮนนิ่ง 12

2.7 แสดงการเพิ่มจุดสัญญาณของข้อมูล 14

2.8 Analysis Filter 16

2.9 ส่วนประกอบของนิวรอนในสมองมนุษย์ 17

2.10 การทำงานของนิวรอน 18

2.11 แบบจำลองนิวรอนที่มีอินพุตเดียว 19

2.12 กราฟของสมการซิกมอยด์(A) และไฮเปอร์โบลิกแทนเจนท์(B) 20

2.13 แผนภาพนิวรอนที่มีหลายอินพุต 21

2.14 เครือข่ายประสาทเทียมแบบชั้นเดียว 22

2.15 เครือข่ายประสาทเทียมแบบหลายชั้น 22

2.16 เครือข่ายประสาทเทียมแบบหลายชั้น โดยใช้อัลกอริธึมแบบ Backpropagation 24

3.1 กระบวนการรู้จำเสียง 29

3.2 แผนภาพกระบวนการประมวลผลเสียงเบื้องต้น 30

3.3 แผนภาพกระบวนการหาจุดสิ้นสุดของคำ 31

3.4 แผนภาพของค่าพลังงานที่ได้จากการวิธีการหาค่าพลังงานสัมบูรณ์ 32

3.5 กระบวนการปรับบรรทัดฐานทางเวลาของคำ 34

3.6 แผนภาพขั้นตอนหลังจากกระบวนการประมวลผลเสียงเบื้องต้น 35

3.7 แผนภาพการวางกรอบหน้าต่างสัญญาณและการซ้อนทับของหน้าต่าง 36

3.8 กระบวนการในเครือข่ายประสาทเทียม 38

3.9 กระบวนการตัดสินใจ 40

3.10 กระบวนการทำงานการควบคุมอุปกรณ์ผ่านมือถือ 42

3.11 ยูสเคสไดอะแกรมของระบบการควบคุมอุปกรณ์ผ่านมือถือ 43

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป(ต่อ)

รูปที่ หน้า

3.12 แอคทิวิตีไดอะแกรมการควบคุมอุปกรณ์ 45

3.13 แอคทิวิตีไดอะแกรมการอ่านค่าสถานะปัจจุบันของอุปกรณ์ 45

3.14 แผนภาพการต่ออุปกรณ์ทางด้านกายภาพ 46

3.15 กระบวนการการทำงานของไมโครคอนโทรลเลอร์ 47

4.1 เปรียบเทียบการตัดเสียงด้วยวิธีหาค่าพลังงานในช่วงเวลาสั้นๆ 50

4.2 กราฟในโดเมนความถี่ระหว่างสัญญาณเสียงเดิมกับสัญญาณเสียงที่ได้จากค่าสัมประสิทธิ์ การประมาณพันธะเชิงเส้น 52

4.3 ส่วนติดต่อกับผู้ใช้งานของระบบควบคุมอุปกรณ์ไฟฟ้าด้วยเสียง 54

4.4 ส่วนติดต่อผู้ใช้งานในการรับข้อมูลเข้า 54

4.5 ส่วนที่แสดงผลค่าและขั้นตอนในการประมวลผลเสียงที่รับเข้ามา 55

4.6 ส่วนที่เป็นเอาท์พุทที่ได้จากการประมวลผลเสียง 56

4.7 ส่วนของการตรวจสอบและแสดงสถานะของอุปกรณ์ 56

4.8 ส่วนของการเรียนรู้คำสำหรับผู้ใช้งาน 57

4.9 หน้าสำหรับการยืนยันตัวตน 58

4.10 หน้าสำหรับการควบคุมอุปกรณ์ภายในบ้าน 58

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันเทคโนโลยีได้มีการพัฒนาไปอย่างรวดเร็วมาก รวมถึงเทคโนโลยีด้านคอมพิวเตอร์และการสื่อสาร ซึ่งจากเทคโนโลยีสองตัวนี้สามารถนำไปใช้พัฒนาร่วมกับระบบอื่นๆได้ เช่น ระบบไฟฟ้า ระบบรักษาความปลอดภัย เป็นต้น เพื่อให้มนุษย์ได้รับประโยชน์ของเทคโนโลยีที่นำมาประยุกต์ใช้และพัฒนาร่วมกัน

โดยในโครงงานนี้จะเป็นการศึกษาเกี่ยวกับการควบคุมอุปกรณ์ภายในบ้าน ซึ่งเทคโนโลยีที่จะนำมาใช้พัฒนาจะประกอบด้วยสองส่วนคือ เทคโนโลยีทางด้านคอมพิวเตอร์ผ่านการรู้จำเสียงพูดของมนุษย์ และเทคโนโลยีระบบสื่อสารผ่านมือถือ ซึ่งทั้งสองเทคโนโลยีจะนำมาพัฒนาร่วมกับระบบอุปกรณ์ไฟฟ้าภายในบ้านเพื่ออำนวยความสะดวกสำหรับผู้อยู่อาศัยภายในบ้าน

ปัญหาที่เกิดขึ้นสำหรับผู้สูงอายุ,ผู้พิการทางแขนขาหรือแม้กระทั่งผู้ป่วย จะมีปัญหากับการเคลื่อนไหวร่างกายเพื่อที่จะไปควบคุมอุปกรณ์ไฟฟ้า เทคโนโลยีทางด้านการรู้จำเสียงพูดจะช่วยผู้สูงอายุ,ผู้พิการทางแขนขาและผู้ป่วยให้สามารถที่จะควบคุมไฟฟ้าได้สะดวกขึ้น ในส่วนของการควบคุมผ่านระบบมือถือนั้นจะเป็นส่วนที่ทำให้ผู้อยู่อาศัย สามารถที่จะควบคุมอุปกรณ์ไฟฟ้าภายในบ้านโดยสามารถที่จะควบคุมได้จากระยะไกล เช่น เจ้าของบ้านไม่อยู่บ้านหลายวัน ต้องการเปิดไฟหน้าบ้านในตอนกลางคืนและปิดตอนเช้าเพื่อประหยัดค่าไฟฟ้าและทำให้เสมือนมีคนอยู่บ้านป้องกันการโจรกรรมทางจิตวิทยาได้ หรือแม้แต่การรดน้ำต้นไม้ ในเวลาที่เจ้าของบ้านไม่อยู่ โดยสามารถสั่งมาทางโทรศัพท์ว่าให้มีการรดน้ำในช่วงเวลาที่ต้องการได้ เป็นต้น

1.2 วัตถุประสงค์ของการศึกษา

1. เพื่อศึกษาและพัฒนาระบบการรู้จำเสียงพูดแบบไม่ขึ้นกับผู้พูดโดยใช้การสกัดลักษณะเด่นและโครงข่ายประสาทเทียม
2. เพื่อศึกษาและพัฒนาระบบการสื่อสารผ่านระบบมือถือมาประยุกต์ใช้ควบคุมอุปกรณ์ไฟฟ้าภายในบ้านโดยอาศัยเทคโนโลยีเว็บ
3. เพื่อนำเทคโนโลยีที่กล่าวถึงมาใช้กับผู้พิการทางแขนขา ผู้สูงอายุ หรือผู้ที่ต้องการได้รับประโยชน์จากการใช้งานจากโครงงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3 ขอบเขตของการศึกษา

ในส่วนของขอบเขตของโครงงานจะแบ่งออกได้เป็นสองส่วน ดังนี้

1. การควบคุมอุปกรณ์ไฟฟ้าผ่านการรู้จำเสียงพูดของมนุษย์ โดยไม่ขึ้นกับผู้พูด โดยจะมีขอบเขตของโครงงานดังนี้
    1.1 จำนวนคำทั้งหมด 10 คำเป็นคำโดดสองพยางค์ เช่น เปิดไฟ ปิดไฟ เป็นต้น
    1.2 สภาพแวดล้อมขณะที่ทำการพูดอยู่ในลักษณะเงียบสงบ
    1.3 สำเนียงที่พูดเป็นภาษากลาง โดยการพูดต้องพูดเป็นพยางค์ไม่ต่อเนื่องและมีช่องว่างระหว่างคำไม่น้อยกว่า 1 วินาที
    1.4 ระบบการรู้จำเสียงพูดจะไม่ทำงานตามเวลาจริง โดยจะควบคุมบนโปรแกรมคอมพิวเตอร์ก่อนจะรับสัญญาณเสียงเข้าเพื่อควบคุมอุปกรณ์
    1.5 ระบบการประมวลผลจะกระทำอยู่บนเครื่องคอมพิวเตอร์โดยมีการส่งค่าจากการประมวลผลเพื่อไปควบคุมอุปกรณ์ภายในบ้าน
2. การควบคุมอุปกรณ์ไฟฟ้าผ่านระบบมือถือ
    2.1 ควบคุมผ่านมือถือโดยใช้เว็บในการเข้าถึงเพื่อควบคุมอุปกรณ์
    2.2 ระบบการประมวลผลจะกระทำอยู่บนคอมพิวเตอร์โดยมีการส่งค่าจากการประมวลผลเพื่อไปควบคุมอุปกรณ์ภายในบ้าน
3. ผลที่ได้จากการประมวลผลจากคอมพิวเตอร์ จะส่งผ่านทางพอร์ตอนุกรมไปยังไมโครคอนโทรลเลอร์เพื่อนำไปควบคุมอุปกรณ์ไฟฟ้าต่อไป

1.4 ขั้นตอนของการศึกษา

ในขั้นตอนการศึกษานี้จะแบ่งออกเป็นสองช่วง ดังนี้

1. ศึกษาและพัฒนาระบบการควบคุมอุปกรณ์ภายในบ้านผ่านเสียงพูดมนุษย์
    1.1 กำหนดขอบเขตของคำที่ใช้ในโครงงาน
    1.2 ศึกษาลักษณะโครงสร้างของเสียงพูดของมนุษย์
    1.3 ศึกษาระบบของการรู้จำเสียงพูด
    1.4 บันทึกเสียงพูดจากกลุ่มคำที่กำหนดไว้
    1.5 ออกแบบและพัฒนาโปรแกรมเพื่อใช้ในการรู้จำเสียงพูดของมนุษย์
    1.6 นำข้อมูลเสียงพูดที่ได้จากการบันทึกมาทำการเข้าสู่โปรแกรมเพื่อรู้จำเสียงพูดและทดสอบการรู้จำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.7 ผลที่ได้จากการประมวลผลด้วยคอมพิวเตอร์ จะส่งผ่านทางพอร์ตอนุกรมไปยังไมโครคอนโทรลเลอร์เพื่อไปควบคุมอุปกรณ์ไฟฟ้า

1.8 สรุปผลการศึกษาและทดสอบ ประเมินผล และเสนอแนะแนวทางในการพัฒนาต่อไป

2. ศึกษาและพัฒนาระบบการควบคุมอุปกรณ์ภายในบ้านผ่านมือถือ

2.1 กำหนดขอบเขตของการพัฒนาบนระบบมือถือ

2.2 ศึกษาเทคโนโลยีที่จะนำมาใช้พัฒนาบนระบบมือถือ

2.3 ทำการออกแบบและพัฒนาโปรแกรมเพื่อใช้ในมือถือ

2.4 ผลที่ได้จากการประมวลผลจะส่งผ่านทางพอร์ตอนุกรมไปยังไมโครคอนโทรลเลอร์เพื่อไปควบคุมอุปกรณ์ไฟฟ้าต่อไป

2.5 สรุปผลการศึกษาและทดสอบ ประเมินผล และเสนอแนะแนวทางในการพัฒนาต่อไป

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ทำให้ได้มาซึ่งระบบที่ช่วยให้เกิดประโยชน์ต่อผู้ใช้งานไม่ว่าจะเป็นผู้สูงอายุ ผู้พิการทางแขนขาหรือผู้อยู่อาศัยทั่วไป
2. สามารถนำเทคโนโลยีที่พัฒนาไปประยุกต์ใช้กับด้านต่างๆเพิ่มขึ้นได้ เช่น นำมาประยุกต์กับระบบเซนเซอร์เพื่อตรวจสอบด้านความปลอดภัย โดยอาจจะมีการแจ้งเตือนผ่านระบบมือถือ เป็นต้น

1.6 ส่วนประกอบของรายงาน

ในรายงานโครงงานนี้ จะมีโครงสร้างทั้งหมด 5 บท โดยที่

บทที่ 2 จะกล่าวถึงหลักทฤษฎี ที่ศึกษาและนำมาใช้พัฒนาโครงงาน โดยจะแบ่งเป็นส่วนที่เกี่ยวข้องกับทฤษฎีในการรู้จำเสียง เช่น ทฤษฎีการสกัดคุณลักษณะเด่นของเสียง , ทฤษฎีเครือข่ายประสาทเทียม เป็นต้น และอีกส่วนหนึ่งก็จะเป็นทฤษฎีการพัฒนาบนระบบมือถือ

บทที่ 3 กล่าวถึงการวิเคราะห์และออกแบบสำหรับการพัฒนาโครงงานโดยจะใช้หลักทฤษฎีในบทที่สองเพื่อสร้างขั้นตอนในการพัฒนา

บทที่ 4 กล่าวถึงการพัฒนา โดยจะพัฒนาตามขั้นตอนที่ได้มีการวิเคราะห์และออกแบบไว้ในบทที่ 3

บทที่ 5 กล่าวถึงผลการทดลองและสรุปประเด็นสำคัญที่ได้รับจากการพัฒนา

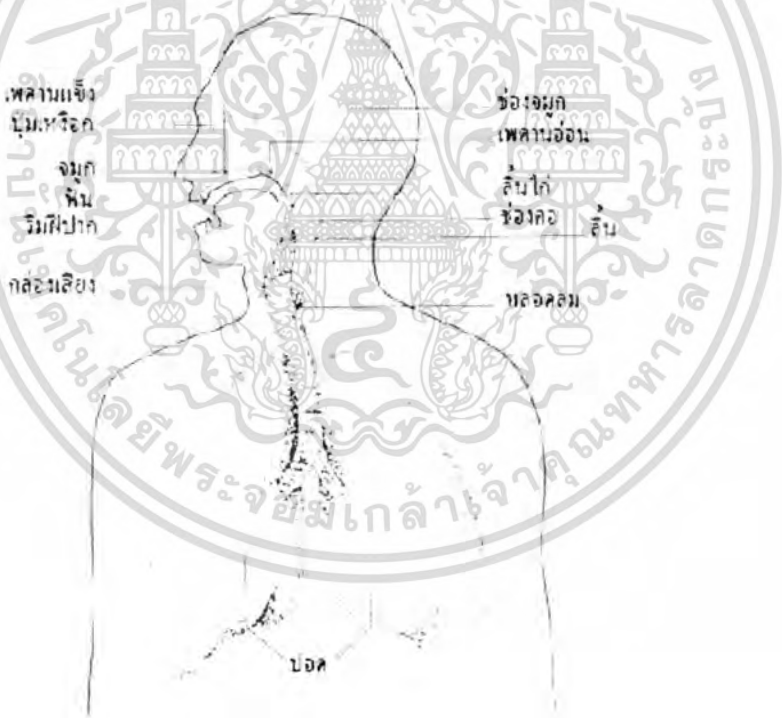
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีและหลักการที่เกี่ยวข้อง

2.1 ทฤษฎีเสียง

2.1.1 อวัยวะที่สัมพันธ์กับเสียงพูด

อวัยวะในส่วนต่างๆของร่างกายที่มีความสำคัญในการผลิตเสียงพูดมี 3 ส่วนสำคัญ ส่วนแรกคือส่วนของลำตัวซึ่งเกี่ยวข้องกับระบบหายใจ ประกอบไปด้วย ปอด กล้ามเนื้อ หลอดลม หรือเรียกง่ายๆว่าการหายใจ, ส่วนที่สองคือส่วนบริเวณลำคอที่เกี่ยวข้องกับระบบของการเปล่งเสียงพูด โดยมีกล่องเสียงทำหน้าที่หลักในการเปล่งเสียงพูด ซึ่งไม่เพียงแต่มนุษย์ สัตว์ต่างๆเปล่งเสียงโดยใช้กล่องเสียงเช่นกัน สุดท้ายคืออวัยวะในส่วนของศีรษะซึ่งเป็นอวัยวะที่ใช้ผลิตเสียงพูด โดยเริ่มจากริมฝีปาก, ปาก, ฟันและลิ้น แม้ว่าแต่ละส่วนของร่างกายที่กล่าวมาข้างต้นนั้นจะทำหน้าที่ต่างกัน แต่อวัยวะทุกส่วนนั้นมีการทำงานไปพร้อมๆกันในขณะที่มีการผลิตเสียงพูด[2] สามารถอธิบายอวัยวะและหน้าที่ของอวัยวะที่สัมพันธ์ของการพูดได้ดังนี้



รูปที่ 2.1 แสดงภาพของอวัยวะในร่างกายที่มีความสัมพันธ์กับเสียงพูด

ปอด ปอดเป็นอวัยวะที่สำคัญ เนื่องจากเป็นต้นกำเนิดของกระแสลมที่ทำให้เกิดเสียงพูด คือมีการใช้ลมหายใจที่เกิดจากการเคลื่อนไหวของกล้ามเนื้อ ทำให้ลมหายใจมี 2 ทิศทางคือ ทิศทางลมออก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(egressive airstream) และทิศทางลมเข้า (ingressive airstream) เสียงพูดโดยทั่วไป ในภาษาต่างๆใช้กระแสลมออก และบางเสียงในภาษาที่มีการใช้กระแสลมเข้าแต่ทั้งนี้ไม่มีบทบาททางภาษา

กล่องเสียง เมื่อลมผ่านจากปอดแล้วจะเคลื่อนผ่านออกมาทางหลอดลม ผ่านต่อมายังกล่องเสียง (larynx) ซึ่งอยู่ด้านบนของหลอดลม กล่องเสียงประกอบด้วยกล้ามเนื้อและกระดูกอ่อน 3 ส่วนดังนี้ กระดูกอ่อนชิ้นแรกคือกระดูกไทรอยด์ (thyroid) เป็นกระดูกอ่อนชิ้นใหญ่ มีลักษณะโป่งนูน หรือที่เรียกว่า "ลูกกระเดือก" โดยมีหน้าที่สำคัญคือป้องกันไม่ให้กล่องเสียงด้านหน้าได้รับความกระทบกระเทือนและเป็นส่วนยึดกล้ามเนื้อที่ทำให้เกิดการทำงานของเส้นเสียงกระดูกอ่อนชิ้นที่สองคือ กระดูกไครคอยด์ (cricoid) จะอยู่ด้านล่างของกระดูกไทรอยด์ มีลักษณะคล้ายวงแหวน บางครั้งเรียกว่ากระดูกรูปวงแหวน เมื่อทำหน้าที่ร่วมกับกระดูกไทรอยด์จะช่วยป้องกันกล่องเสียงได้เป็นอย่างดี กระดูกอ่อนชิ้นที่สามคือกระดูกแอรีทีนอยด์ (arythenoid) จะอยู่ส่วนบนด้านหลังของกระดูกไทรอยด์และกระดูกไครคอยด์ ส่วนสุดท้ายของเส้นเสียงยึดติดกับด้านหน้าของกระดูกแอรีทีนอยด์ กระดูก 2 ชิ้นนี้เคลื่อนที่ขึ้นลงหรือเคลื่อนที่ไปทางซ้ายทางขวาได้และเป็นกระดูกชิ้นสำคัญของการทำงานของเส้นเสียง

เส้นเสียง (หรือเส้นเสียงแท้) เส้นเสียงอยู่ในบริเวณส่วนประกอบของกระดูกอ่อนทั้งสามที่กล่าวมาข้างต้น มีลักษณะเป็นเส้นเอ็นบางๆ ยืดหยุ่นและเคลื่อนไหวได้ด้วยกล้ามเนื้อไทโรแอรีทีนอยด์ มี 2 เส้นโยงจากด้านหน้าไปด้านหลังของกล่องเสียง สามารถเคลื่อนมาชิดติดกันหรือยืดห่างออกจากกันเพื่อเปิดเป็นช่องว่าง (glottis) โดยปกติแล้วเส้นเสียงของผู้ชายจะยาวกว่าเส้นเสียงของผู้หญิง เสียงผู้ชายจึงทุ้มต่ำกว่าเสียงผู้หญิง

ช่องคอ มีลักษณะเป็นท่อรูปกรวยเหนือกล่องเสียง มีกล้ามเนื้อบุตามผนังช่องคอลักษณะของช่องคออาจเปลี่ยนไปได้ขึ้นอยู่กับรูปร่างของลิ้นที่เปลี่ยนรูปไปเนื่องจากการออกเสียงเช่น การยกส่วนหลังของลิ้นขึ้นสู่เพดานอ่อน หรือถดลิ้นส่วนหลังหรือโคนลิ้นสู่ผนังช่องคอในการออกเสียงบางเสียง ช่องคอยังแบ่งออกเป็นส่วนช่องคอด้านช่องจมูก (nasapharynx) หรือส่วนช่องด้านช่องปาก (oropharynx) ทั้งหมดมีหน้าที่เป็นเนื้อที่รองรับให้ลิ้นส่วนหลังและโคนลิ้นเคลื่อนไหวในการออกเสียง ตลอดจนเป็นช่องผ่านให้ลมจากปอดออกไปสู่ช่องปากหรือช่องจมูกได้ นอกจากนี้ส่วนของกล้ามเนื้อตามผนังคอยังมีบทบาททางกายภาพเกี่ยวกับการกลืนคือมีการขยายและหดตัวเมื่อมีการกลืนเกิดขึ้น

ลิ้น เป็นอวัยวะที่มีบทบาทสำคัญในการออกเสียงพยัญชนะและเสียงสระ มีกล้ามเนื้อซับซ้อนเพื่อให้เกิดการเคลื่อนไหวหลายรูปแบบ เมื่อลิ้นวางอยู่ในลักษณะปกติไม่เคลื่อนไหว ส่วนตรงกลาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของลิ้น (หรือลิ้นส่วนหน้า) วางอยู่ในลักษณะตั้งฉากกับส่วนต่อระหว่างเพดานแข็งและเพดานอ่อน กล้ามเนื้อลิ้นช่วยให้ปลายลิ้นและปลายสุดของลิ้นเคลื่อนไปยังบริเวณปาก ฟัน แนวปุ่มเหงือก หรือพับปลายลิ้นไปด้านหลังบริเวณหลังแนวปุ่มเหงือกหรือเกือบถึงเพดานแข็งได้ ส่วนบริเวณกลางลิ้นหรือเรียกว่าลิ้นส่วนหน้าเคลื่อนสู่บริเวณหน้าของเพดานแข็งหรือยกขึ้นสู่เพดานแข็ง สำหรับส่วนหลังของลิ้นเคลื่อนสู่บริเวณเพดานอ่อน และโคนลิ้นเคลื่อนลงสู่ช่องคอได้แต่ละจุดของอวัยวะในปากที่กล่าวมานั้น เมื่อลิ้นเคลื่อนที่ไปตามจุดแต่ละจุด จะทำให้เกิดเสียงที่แตกต่างกันไป

เพดานปาก แบ่งออกเป็น 2 ส่วน คือส่วนที่ประกอบไปด้วยกระดูกและส่วนที่เป็นกล้ามเนื้อ ส่วนที่ประกอบด้วยกระดูกเริ่มจากแนวปุ่มเหงือกหรือปุ่มเหงือกที่ห่อหุ้มรากฟันตามด้วยส่วนที่แข็งเรียกว่าเพดานแข็ง ส่วนหลังของเพดานปากถัดจากเพดานแข็งเป็นส่วนที่ประกอบด้วยกล้ามเนื้อเรียกว่าเพดานอ่อน และยังมีกล้ามเนื้อรูปร่างคล้ายอักษร "V" ติดกับเพดานอ่อน ห้อยตัวสู่
ช่องคอเรียกว่าลิ้นไก่ บริเวณกล้ามเนื้อที่ประกอบด้วยเพดานอ่อนและลิ้นไก่นี้นอกจากจะมีบทบาทในการเป็นจุดที่เกิดเสียงแล้ว ยังมีบทบาททำให้เกิดเป็นช่องปากหรือช่องจมูก คือเมื่อเพดานอ่อนรวมตลอดถึงลิ้นไก่ยกขึ้นจรดผนังคอปิดกั้นไม่ให้ลมออกทางช่องจมูก ทำให้กระแสลมที่ใช้ในการออกเสียงผ่านไปทางช่องปากทางเดียว แต่เมื่อกล้ามเนื้อของเพดานอ่อนและลิ้นไก่ลดตัวลงทำให้กระแสลมจากช่องคอขึ้นมาทางจมูกได้ด้วย ทำให้เกิดเสียงต่างประเภทกันคือ เสียงระเบิด (stopหรือ plosive) ใช้ลมที่ออกมาทางช่องปาก ส่วนเสียงนาสิก (nasal) เป็นเสียงลมที่ออกทางช่องจมูก

ฟัน ในทางทันตกรรมได้แบ่งฟันไว้หลายส่วนตามหน้าที่ในการบดเคี้ยวอาหาร จากด้านหน้าไปทางด้านหลังตามลำดับคือ ฟันหน้า, ฟันเขี้ยว, ฟันกรามน้อย, ฟันกราม ถ้าพิจารณาฟันในลักษณะที่เป็นจุดที่เกิดเสียงแล้วพบว่าฟันหน้าบนมีบทบาทมากกว่าฟันล่าง

ริมฝีปาก เป็นอวัยวะที่มีบทบาทในการออกเสียง เช่น การใช้ริมฝีปากทั้งคู่บน-ล่างจรดกันสนิทเพื่อออกเสียง 'บ, ป, ม' ในภาษาไทย หรือใช้ริมฝีปากล่างเคลื่อนไปใกล้เกือบชิดฟันบนเพื่อออกเสียง 'ฟ'

2.1.2 ลักษณะเสียงพูดของมนุษย์

เสียงพูด เป็นคลื่นตามยาว (Longitudinal Wave) เกิดจากการสั่นของอนุภาคตัวกลาง นั่นคืออากาศและทิศทางการสั่นของอนุภาคจะอยู่ในทิศเดียวกันกับทิศทางของการเคลื่อนที่ คลื่นเสียงเป็นคลื่นที่เปลี่ยนแปลงไปตามเวลา เสียงพูดแบ่งออกได้เป็น 2 ชนิดตามการกำเนิดเสียง (Mode) หรือการกระตุ้น คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.2.1 เสียงก้องหรือเสียงโฆษะ (Voiced) เกิดจากการบังคับอากาศให้ผ่านช่องสายเสียงทำให้มีการเปลี่ยนแปลงความตึงหย่อนของเส้นเสียง โดยเส้นเสียงจะสั่นและเกิดเป็นพัลส์ (Pulse) ของอากาศไปกระตุ้นอวัยวะกำทอนเกิดเป็นเสียงก้อง ตัวอย่างเสียงก้องได้แก่ เสียงสระ เสียงพยัญชนะ ที่ต้องออกเสียงจากลำคอ (Voiced Consonants)

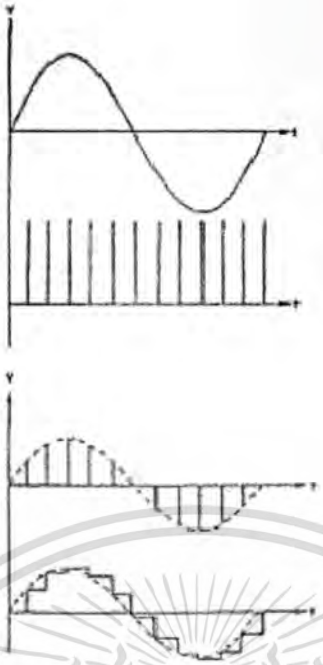
2.1.2.2 เสียงไม่ก้องหรืออโฆษะ (Unvoiced หรือ Voiceless) เป็นเสียงที่ไม่เกิดจากการสั่นของเส้นเสียง แต่เกิดในช่องปากหรือโพรงจมูก โดยอวัยวะภายในช่องปาก ริมฝีปาก ขวางการไหลของอากาศให้ผ่านได้เป็นช่องเล็กๆ อากาศจึงไหลผ่านอย่างรวดเร็วและปั่นป่วนจนกระทั่งสร้างเป็นเสียงรบกวน ช่วงความถี่กว้าง (Broad-spectrum Noise) ตัวอย่างเสียงไม่ก้องได้แก่ เสียงพยัญชนะที่ไม่ได้เกิดจากลำคอ (Voiceless Consonants)

2.2 การแปลงรูปแบบสัญญาณเสียงพูด

รูปแบบของสัญญาณไฟฟ้าที่เราพบเห็นและคุ้นเคยในชีวิตประจำวันจะอยู่ในรูปแบบของสัญญาณเสียงต่อเนื่อง หรือที่เรียกว่าสัญญาณอนาล็อก ซึ่งแต่เดิมการเอาสัญญาณไฟฟ้าดังกล่าวมาประมวลผลในแบบอนาล็อก แต่เมื่อมีการคิดค้นเทคนิคการประมวลผลด้วยดิจิตอล จึงได้มีการเปลี่ยนการประมวลผลมาเป็นแบบดิจิตอล เนื่องจากพบว่าการสื่อสารและแสดงผลในรูปแบบดิจิตอลสามารถกระทำได้ง่ายและมีประสิทธิภาพมากกว่า โดยวงจรแปลงสัญญาณอนาล็อคเป็นดิจิตอล (Analog to Digital Converters) หรือ ADC และนำมาประมวลด้วยตัวประมวลผลทางดิจิตอล (Digital Processors) เช่น คอมพิวเตอร์หรือวงจรดิจิตอล (Digital Circuit)

หลักการพื้นฐานที่ใช้ในการแปลงสัญญาณอนาล็อกเป็นสัญญาณดิจิตอล จะใช้ระบบการสุ่มสัญญาณ โดยสัญญาณอนาล็อกจะถูกสุ่มเป็นระยะด้วยค่าคงที่ ดังแสดงในรูป 2.2 กลุ่มของสัญญาณสุ่มจะแทนด้วยแบนด์วิธที่ทำงานด้วยความเร็วสูง ซึ่งจะทำการตัดต่อสัญญาณอนาล็อกในช่วงเวลาอันสั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



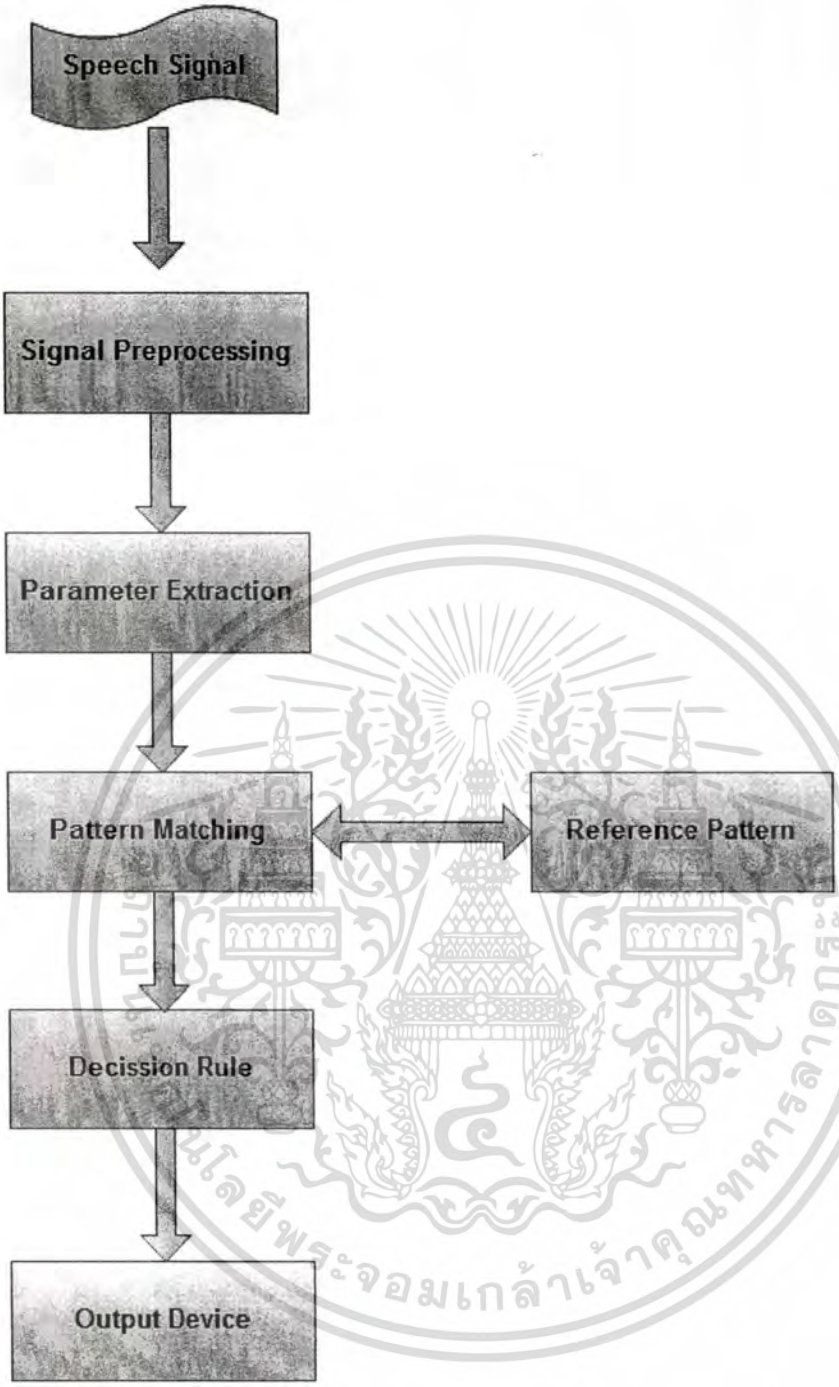
รูปที่ 2.2 ภาพแสดงกระบวนการสุ่มสัญญาณ

ผลของการสุ่มสัญญาณด้วยความเร็วจะเสมือนกับการคูณขบวนสัญญาณพัลส์แคบๆกับสัญญาณอนาล็อก ซึ่งจะได้เป็นสัญญาณที่เกิดการมอดูเลทระหว่างขบวนการพัลส์กับสัญญาณอนาล็อกดังแสดงในรูป 2.2(ค) โดยสัญญาณอนาล็อกจะขี่มาบนขบวนพัลส์ ถ้าหากเอาสวิทช์และตัวเก็บประจุแทนสวิทช์แล้วสัญญาณอนาล็อกที่ถูกสุ่มจะถูกเก็บไว้ในตัวเก็บประจุจนกว่าสัญญาณค่าใหม่ถูกสุ่มเข้ามา ซึ่งลักษณะของเอาต์พุตที่แสดงในรูป 2.2(ง) สำหรับปัญหาที่ว่าอัตราการสุ่มสัญญาณนั้นควรมีขนาดเท่าใดนั้นจึงจะไม่ทำให้ข้อมูลสูญเสียไปเมื่อสัญญาณนั้นถูกเปลี่ยนกลับมาเป็นเช่นเดิมคำตอบก็คือขึ้นอยู่กับความถี่ของสัญญาณอนาล็อค โดยทฤษฎีของการสุ่มกล่าวไว้ว่า " ถ้าสัญญาณต่อเนื่องซึ่งมีความถี่และฮาร์โมนิคไม่เกิน f_c แล้วสัญญาณดังกล่าวจะสามารถเปลี่ยนกลับมาเป็นอย่างเดิมโดยไม่สูญเสียรายละเอียดหรือผิดเพี้ยนไป ถ้าอัตราการสุ่มไม่น้อยกว่า $2f_c$ ต่อวินาที "

2.3 กระบวนการในการรู้จำเสียง

กระบวนการในการรู้จำเสียงพูดโดยทั่วไปจะสามารถแบ่งเป็นส่วนต่างๆได้ดังแผนภาพที่ 2.3 จะประกอบด้วยขั้นตอนการดำเนินงานสามขั้นตอนดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.3 แสดงแผนภาพขั้นตอนการทำงานของกระบวนการรู้จำ

2.3.1การประมวลผลสัญญาณเบื้องต้น (Signal Preprocessing)

เป็นขั้นตอนการประมวลผลสัญญาณเบื้องต้น[6] เพื่อจัดเตรียมข้อมูลให้เหมาะกับการประมวลผลในขั้นต่อๆไป โดยในขั้นตอนนี้ประกอบด้วย

2.3.1.1 กรรมวิธีเน้นล่วงหน้า (Pre-Emphasis)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เนื่องจากสัญญาณเสียงของมนุษย์ จะมีองค์ประกอบส่วนใหญ่อยู่ในบริเวณความถี่ต่ำ เมื่อเทียบกับแถบความถี่แบนด์วิดธ์ (Bandwidth) ไม่เกิน 4 กิโลเฮิรตซ์ เพื่อให้อัตราส่วนเสียงต่อสัญญาณรบกวน (signal to noise ratio:SNR) มีค่าค่อนข้างคงที่ตลอดช่วงความถี่แบนด์วิดธ์นี้ เราจึงต้องมีกรรมวิธีเน้นล่วงหน้า โดยเน้นที่ความถี่ที่สูงให้มีขนาดสูงขึ้น นั้นคือ การพรีแอมพฟาซิส ก็คือ การกรองสัญญาณด้วยวงจรความถี่สูงผ่าน(High pass filter) ซึ่งมักนิยมใช้วงจรกรองอันดับหนึ่งซึ่งจะเป็นการเน้นให้สัญญาณช่วงความถี่สูงมีขนาดสูงขึ้นนั้นเอง มีฟังก์ชั่นถ่ายโอนเป็น

$$H(z) = 1 - a \cdot z^{-1} \qquad 0 \le a \le 1 \tag{2.1}$$

ซึ่งสัมพันธ์กับสัญญาณอินพุตดังสมการ

$$x'(n) = x(n) - a(x-1) \tag{2.2}$$

2.3.1.2 การวางกรอบสัญญาณหน้าต่าง (Windowing)

เนื่องจากสัญญาณเสียงพูดโดยรวมมีลักษณะไม่คงที่และมีการเปลี่ยนแปลงอย่างช้าๆไปตามเวลา จึงจำเป็นต้องมีการวิเคราะห์สัญญาณเสียงที่ละช่วงสั้นๆเฉพาะที่อยู่ภายในกรอบสัญญาณที่กำหนดขึ้น ขนาดของกรอบสัญญาณที่นิยมใช้ในการวิเคราะห์สัญญาณเสียงเพื่อการรู้จำมีค่าประมาณ 10-30 มิลลิวินาทีและควรวางกรอบสัญญาณถัดไปให้มีการเหลื่อมกันของสัญญาณประมาณ ½ หรือ 1/3 ของขนาดกรอบสัญญาณเพื่อให้ข้อมูลที่วิเคราะห์มีความต่อเนื่องกัน กรอบสัญญาณหลักๆมีสามชนิด

2.3.1.2.1 กรอบสัญญาณสี่เหลี่ยม (Rectangular Window) ดังสมการที่ 2.3

$$w(n) = \begin{cases} 1, 0 \le n \le N-1 \\ 0, otherwise \end{cases} \tag{2.3}$$

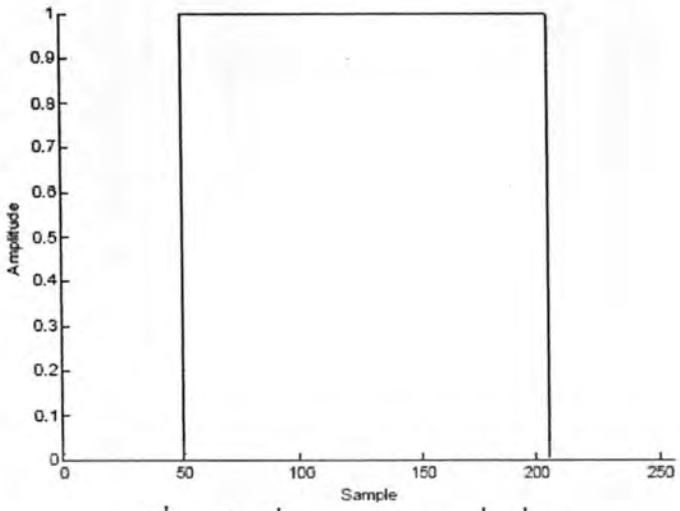
เมื่อ $w(n)$ คือผลลัพธ์ของฟังก์ชั่นกรอบตำแหน่งที่ n

N คือความกว้างหน้าต่าง

n คือข้อมูลในหน้าต่างมีค่าตั้งแต่ 0 จนถึง $N-1$

ซึ่งมีลักษณะดังภาพที่ 2.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.4 ฟังก์ชั่นกรอบสัญญาณสี่เหลี่ยม

2.3.1.2.2 กรอบสัญญาณแฮมมิ่ง (Hamming Window) ดังสมการที่ 2.4

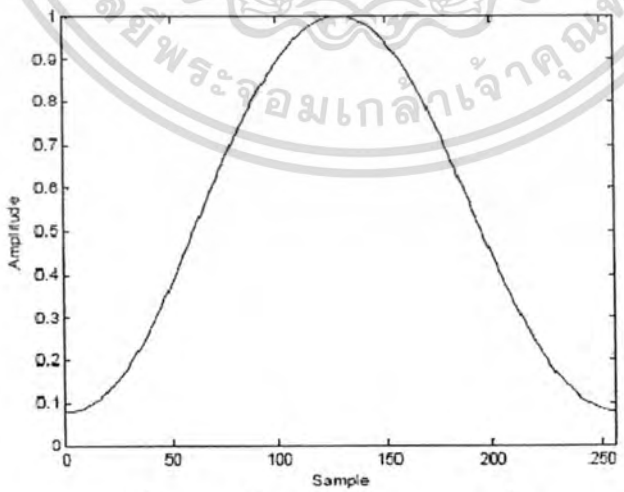
$$w(n) = \begin{cases} 0.54 - 0.46\cos\left[\dfrac{2\pi n}{N-1}\right], 0 \le n \le N-1 \\ 0, otherwise \end{cases} \quad (2.4)$$

เมื่อ $w(n)$ คือผลลัพธ์ของฟังก์ชั่นกรอบตำแหน่งที่ n

N คือความกว้างหน้าต่าง

n คือข้อมูลในหน้าต่างมีค่าตั้งแต่ *0* จนถึง *N-1*

ซึ่งมีลักษณะดังภาพที่ 2.5



รูปที่ 2.5 ฟังก์ชั่นกรอบสัญญาณแฮมมิ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.1.2.3 กรอบสัญญาณแฮนนิ่ง (Hanning Windowing) ดังสมการที่ 2.5

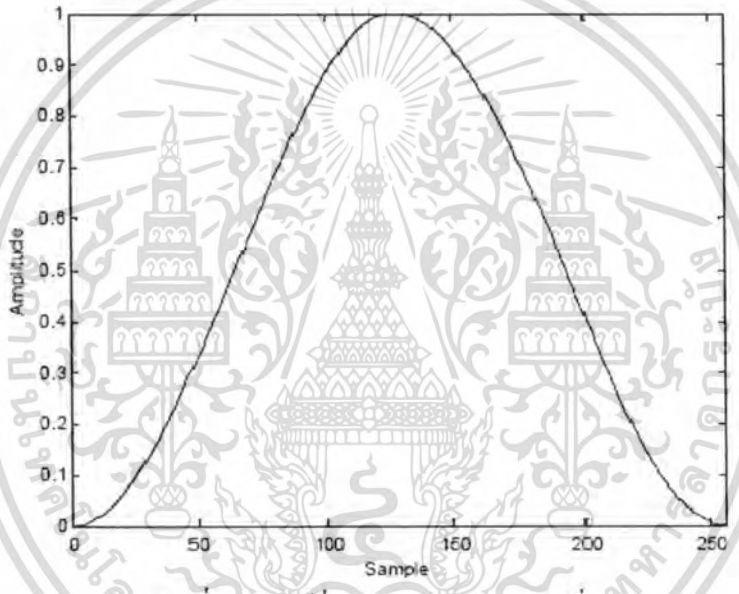
$$w(n)=\begin{cases}0.5\left[1-\cos\left(\dfrac{2\pi n}{N-1}\right)\right], 0\le n\le N-1\\ 0, otherwise\end{cases} \quad (2.5)$$

เมื่อ $w(n)$ คือผลลัพธ์ของฟังก์ชั่นกรอบตำแหน่งที่ n

N คือความกว้างหน้าต่าง

n คือข้อมูลในหน้าต่างมีค่าตั้งแต่ *0* จนถึง *N-1*

ซึ่งมีลักษณะดังภาพที่ 2.6



รูปที่ 2.6 ฟังก์ชั่นกรอบสัญญาณแฮนนิ่ง

กรอบสัญญาณแบบแฮมมิ่ง แฮนนิ่ง เหมาะสำหรับประมวลผลสัญญาณเสียงเพราะ สามารถเน้นเสียงในกรอบที่กำลังพิจารณาให้มีความสำคัญสูงสุด โดยลดความสำคัญของสัญญาณเสียงที่อยู่ในกรอบรอบข้าง แต่ยังคงความต่อเนื่องของสัญญาณเสียงให้มีความต่อเนื่องกันในแต่ละกรอบสัญญาณ ทำให้เสียงที่ผ่านการวางกรอบสัญญาณยังคงข้อมูลครบถ้วยอยู่

2.3.1.3 การหาจุดสิ้นสุดของเสียงพูด (End-Point Detection)

การหาจุดสิ้นสุดของเสียงพูดที่ใช้ในขบวนการรู้จำ เป็นกระบวนการที่ใช้แยกคำพูดออกมาจากส่วนที่ไม่ใช่คำพูด การตัดคำเป็นกระบวนการที่สำคัญ มีผลต่อการรู้จำค่อนข้างมาก โดยวิธีการหาจุดสิ้นสุดของเสียงจะใช้หลักการการวิเคราะห์หาพลังงานในช่วงเวลาสั้นๆ (Short-Time Energy)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยหลักการของการวิเคราะห์พลังงานก็คือ ในช่วงเสียงพูดจะมีพลังงานมากกว่าช่วงที่ไม่ใช่คำพูด การหาค่าพลังงานเขียนเป็นสมการได้ดังนี้

$$E_n = \sum_{m=-\infty}^{\infty} [x(m)w(n-m)]^2 \tag{2.6}$$

โดยค่า E_n คือค่าพลังงาน E ในช่วงเวลาที่ n

ค่า w คือหน้าต่างกรอบสัญญาณ

การหาค่าพลังงานตามวิธีการนี้จะมีความไวต่อสัญญาณที่มีแอมปลิจูดขนาดใหญ่ๆและใช้เวลาในการคำนวณมากเนื่องจากการยกกำลังสอง ดังนั้นจึงแก้ปัญหาโดยการเปลี่ยนมาใช้ค่าสัมบูรณ์ (Absolute) แทนค่ายกกำลังสองตามสมการที่ 2.7 ผลจากการวิเคราะห์ที่ได้จะเรียกว่าค่าพลังงานสัมบูรณ์

$$M_n = \sum_{m=-\infty}^{\infty} |x(m)|w(n-m) \tag{2.7}$$

เมื่อค่า M_n คือ พลังงานสัมบูรณ์ที่เฟรม n

$w(n)$ คือ หน้าต่างกรอบสัญญาณ

$x(n)$ คือ สัญญาณเสียง

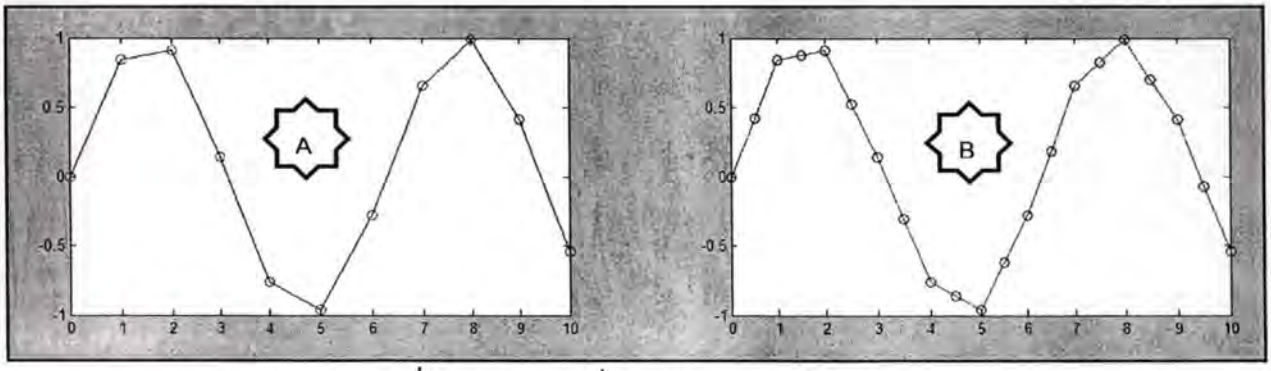
2.3.1.4 การปรับบรรทัดฐานทางเวลา (Time-Normalize)

เป็นกระบวนการที่ทำให้เสียงพูดแต่ละคำมีจำนวนจุดสัญญาณในแกนเวลาเป็นจำนวนเท่ากัน กระบวนการนี้เป็นกระบวนการที่สำคัญก่อนที่จะทำในขั้นตอนต่อๆไปเพราะคำที่ตัดมาแต่ละคำมีความยาวไม่เท่ากัน[1] โดยการปรับบรรทัดฐานทางเวลาทำโดยใช้วิธีการดังนี้

- **การประมาณค่าในช่วงเชิงเส้น (Linear Interpolation)**

จะทำการประมาณค่าแอมปลิจูดของสัญญาณที่จุดที่ไม่ทราบค่าจากความสัมพันธ์แบบเชิงเส้นของจุดสัญญาณเดิมก่อนการปรับบรรทัดฐานทางเวลาที่อยู่รอบจุดที่ไม่ทราบค่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.7 แสดงการเพิ่มจุดสัญญาณของข้อมูล

จากรูปที่ 2.7 เป็นการเพิ่มขนาดของจุดข้อมูลโดยใช้วิธีการประมาณค่าในช่วงเชิงเส้น จากรูป 2.7-A มีจุดข้อมูลอยู่ 11 จุด ต้องการจะเพิ่มจุดข้อมูลเป็น 22 จุด สามารถทำได้โดยการคำนวณค่าแอมพลิจูดของจุดข้อมูลใหม่ โดยจากเดิมจะเลื่อนข้อมูลไปจุดละ 1 เป็นการเลื่อนไปจุดละ 0.5 และใช้วิธีการประมาณค่าในช่วงเชิงเส้นระหว่างค่าแอมพลิจูดของจุดข้อมูลเดิมที่อยู่ข้างๆจุดนี้ เมื่อผ่านการประมาณเชิงเส้นแล้วจะได้ดังรูป 2.7-B ที่มีจุดข้อมูลเพิ่มเป็น 22 จุด ตามต้องการ

2.3.2 การสกัดลักษณะเด่น (Feature Extraction)

การนำสัญญาณเสียงมาเปรียบเทียบกันจุดต่อจุดทำให้ต้องเสียเวลาในการคำนวณและสิ้นเปลืองเนื้อที่หน่วยความจำมาก การเปรียบเทียบเฉพาะจุดสำคัญหรือจุดเด่นจะช่วยลดเวลาในการประมวลผลลงได้มาก การสกัดลักษณะเด่นจะเป็นเทคนิคที่ช่วยลดจำนวนข้อมูลลงโดยที่ข้อมูลจำนวนมากจะถูกแปลงให้เป็นชุดข้อมูลขนาดน้อยลงแต่ยังมีคุณสมบัติสำคัญของคลื่นสัญญาณเสียงได้อย่างถูกต้อง สัญญาณเสียงพูดจะถูกเก็บลักษณะสำคัญของเสียงพูดแต่ละเสียงไว้ การสกัดลักษณะเด่นจะเป็นขั้นตอนที่สำคัญมาก เพราะถ้าสามารถสกัดลักษณะเด่นที่บ่งบอกถึงความแตกต่างของสัญญาณเสียงแต่ละเสียงได้ดี อัตราการรู้จำก็จะมีประสิทธภาพมากขึ้นโดยในโครงงานจะใช้การประมาณพันธะเชิงเส้น (Linear Predictive Coding) โดยการประมาณพันธะเชิงเส้นเป็นวิธีการที่ใช้กันอย่างแพร่หลายให้การหาลักษณะเด่นของเสียง เป็นขบวนการทางคณิตศาสตร์ที่ใช้ในการหาเอกลักษณ์ของระบบ โดยพิจารณาว่าเสียงเกิดจากผลรวมเชิงเส้น (linear combination) ของสัญญาณที่ทราบค่าแล้วโดยวิธีกำลังสองน้อยที่สุด (least-square method) ในการเลือกค่าพารามิเตอร์ของระบบ การประมาณพันธะเชิงเส้นสามารถแสดงคุณสมบัติได้ใกล้เคียงกับพื้นฐานการกำเนิดเสียงของมนุษย์ ความสำคัญของเทคนิคการประมาณพันธะเชิงเส้น คือ การที่รูปคลื่นและลักษณะสมบัติทางความถี่ของเสียงพูดสามารถแสดงด้วยค่าพารามิเตอร์เพียงไม่กี่ค่าได้อย่างแม่นยำและมีประสิทธิภาพ[3]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การประมาณพันธะเชิงเส้น (Linear Predictive Coding , LPC)

กำหนดให้สัญญาณเสียงเดิมคือ $x(n)$ ให้ค่าที่ได้จากการประมาณพันธะเชิงเส้นของสัญญาณเดิม $\tilde{x}$ และสมมติให้สัญญาณเสียงที่ทำการวิเคราะห์เกิดจากการนำเอาสัญญาณเสียงในอดีตของสัญญาณเสียงจริงมาจำนวน i ตัวอย่างดังนั้นจะได้ว่า

$$\tilde{x} = a_1 x(n-1) + a_2 x(n-2) + \ldots + a_M x(n-M) = \sum_{i=1}^{M} a_i x(n-i) \tag{2.8}$$

เมื่อค่า a_i คือค่าสัมประสิทธิ์ของ LPC และค่าความคลาดเคลื่อนระหว่างสัญญาณจริงกับสัญญาณที่ประมาณได้คือ

$$\varepsilon(n) = x(n) - \tilde{x}(n) = x(n) - \sum_{i=1}^{M} a_i x(n-i) \tag{2.9}$$

ซึ่งผลรวมของค่าความคลาดเคลื่อนกำลังสองก็คือ

$$E = \sum_{n} \varepsilon^2(n) = \sum_{n} (x(n) - \sum_{i=1}^{M} a_i x(n-i))^2 \tag{2.10}$$

ต้องการทำให้ผลรวมของค่าความคลาดเคลื่อนกำลังสองมีค่าน้อยสุดโดยการให้ค่า E จากสมการที่ 2.10 เท่ากับศูนย์เทียบสัมประสิทธิ์ a_i (ใช้กฎ chain rule) ได้ดังนี้

$$2 \sum_{n} x(n-k) \left[x(n) - \sum_{i=1}^{M} a_i x(n-i) \right] = 0 \quad for\ k = 1,2,3,\ldots,M. \tag{2.11}$$

จากสมการที่ 2.11 จะประกอบด้วย M สมการ M ตัวแปรที่ไม่ทราบค่าดังแสดงได้ดังนี้

$$\begin{aligned} & a_1 \sum_{n} x(n-k)x(n-1) + a_2 \sum_{n} x(n-k)x(n-2) + \cdots \\ & + a_M \textstyle\sum_n x(n-k)x(n-M) \\ & + a_M \sum_{n} x(n-k)x(n-M) \\ & = \textstyle\sum_n x(n-k)x(n-1) \qquad for\ k = 1,2,3,\ldots,M \end{aligned} \tag{2.12}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถ้ากำหนดให้

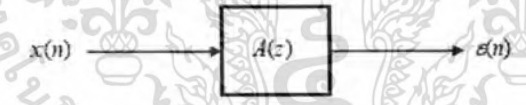
$$r(k) = \sum_{n=0}^{N-1-k} x(n)x(n+k) \tag{2.13}$$

จากสมการที่ 2.13 ค่า $r(k)$ คือค่าอัตสหสัมพันธ์ (Autocorrelation) และเมื่อนำสมการที่ 2.12 มาเขียนใหม่ในรูปเมตริกจะได้สมการดังนี้

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(M-2) & r(M-1) \\ r(1) & r(0) & \cdots & r(M-3) & r(M-2) \\ \vdots & & \ddots & \vdots & \\ r(M-2) & r(M-3) & \cdots & r(0) & r(1) \\ r(M-1) & r(M-2) & \cdots & r(1) & r(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{M-1} \\ a_M \end{bmatrix} = \begin{bmatrix} r(1) \\ r(2) \\ \vdots \\ r(M-1) \\ r(M) \end{bmatrix} \tag{2.14}$$

หรือ $Ra = r$

ค่าสัมประสิทธิ์ที่คำนวณได้ออกมาหนึ่งครั้ง สัมประสิทธิ์ที่ได้จะถูกแทนลงไปในสมการที่ 2.9 แล้วจะได้ค่าความผิดพลาดออกมา ซึ่งขบวนการในสมการที่ 2.9 จะถูกเรียกว่า Analysis filter โดยที่ $x(n)$ คือ อินพุตและ $\varepsilon(n)$ คือ เอาท์พุท แสดงภาพในรูปด้านล่าง



รูปที่ 2.8 Analysis Filter

โดยมีฟังก์ชั่นถ่ายโอน (transfer function) ดังนี้

$$A(z) = 1 - \sum_{i=1}^{M} a_i z^{-i} \tag{2.15}$$

2.3.3 เครือข่ายประสามเทียม

เครือข่ายประสาทเทียม (Artifical Neural Network หรือ ANN) [4][5]เป็นรูปแบบที่มีโครงสร้างและการทำงานประมวลผลเหมือนกับสมองในสิ่งมีชีวิต ที่ซึ่งมีการปรับเปลี่ยนการตอบสนองของอินพุตตามกฎการเรียนรู้ (Learning Rule) หลังจากเครือข่ายได้เรียนรู้สิ่งที่ต้องการแล้ว เครือข่ายนั้นจะจะสามารถทำงานตามที่กำหนดไว้ได้ เครือข่ายประสาทเทียมได้ถูกพัฒนา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

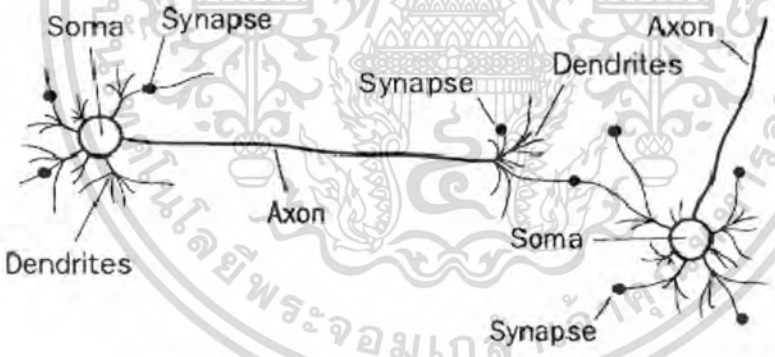
ห้องสมุดคณะเทคโนโลยีสารสนเทศ พระจอมเกล้าลาดกระบัง

คิดค้นจากการทำงานของสมองมนุษย์ โดยสมองมนุษย์ประกอบด้วยหน่วยประมวลผลเรียกว่า นิวรอน (เซลล์ประสาท หรือ neuron) จำนวนนิวรอนในสมองมนุษย์มีอยู่ประมาณ 10^{11} และมีการเชื่อมต่อกันอย่างมากมาย สมองมนุษย์จึงเปรียบเหมือนคอมพิวเตอร์ที่มีการปรับตัวเอง (adaptive) ไม่เป็นเชิงเส้น (nonlinear) และทำงานแบบขนานในการทำงานจัดการการทำงานของนิวรอนในสมอง เพราะฉะนั้นการคำนวณเชิงนิวรอนเป็นการทำงานที่เลียนแบบการทำงานในสมองนั้นเอง

2.2.3.1 เครือข่ายประสาท (Neural Network) คืออะไร

ระบบประสาทเป็นระบบที่มีขนาดใหญ่และซับซ้อน สมองเป็นศูนย์กลางของระบบประสาท โดยจะเชื่อมต่อกับประสาทสัมผัส เพื่อรับข้อมูลจากภายนอกเข้าระบบประสาท แล้วทำการประมวลผลเป็นคำสั่งเพื่อสั่งไปยังอวัยวะต่างๆภายในร่างกาย โดยปกติสมองจะประกอบด้วยเครือข่ายประมาณ 10^{11} ของนิวรอนและเชื่อมต่อกันภายในเครือข่ายย่อยเรียกว่านิวคลีไอ (nuclei) เครือข่ายย่อยนี้จะทำการแยกและปรับข้อมูลจากประสาทสัมผัสก่อนที่จะส่งไปยังเครือข่ายย่อยอื่นๆ และสุดท้ายกลายเป็นรูปสัญญาณที่ผ่านการประมวลส่งเพื่อส่งไปยังอวัยวะส่วนต่างๆ

นิวรอนในสมองมนุษย์ประกอบด้วยองค์ประกอบหลักอยู่ 3 อย่าง คือ เด็นไดรต์ (Dendrite) ร่างเซลล์(Cell Body หรือ Soma) และแกนประสาท (Axon) ดังรูปที่ 2.9

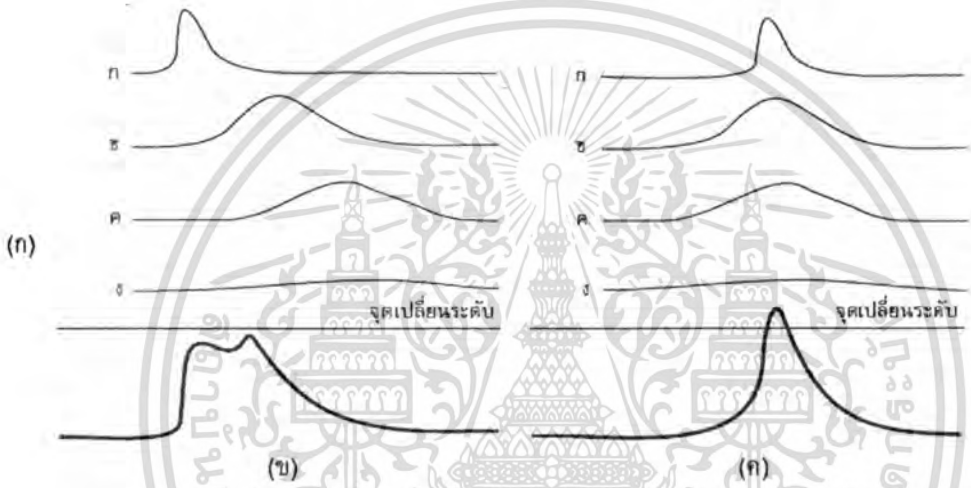


รูปที่ 2.9 ส่วนประกอบของนิวรอนในสมองมนุษย์

เด็นไดรต์จะรับสัญญาณจากนิวรอลอื่นในสมอง แอกซอนของนิวรอลอื่นจะเชื่อมต่อกับเด็นไดรต์และหน้าสัมผัสของตัวเซลล์ด้วยจุดประสานประสาท (synapse) จำนวนการเชื่อมต่อของจุดประสานประสาทขึ้นกับจำนวนชนิดของนิวรอล ซึ่งอาจจะมีตั้งแต่สองสามร้อยจนถึงหลักหมื่น เนื่องจากเยื่อหุ้มเซลล์(membrane) ในสมองมีคุณสมบัติทางไฟฟ้า ขนาดของสัญญาณที่มาถึงเด็นไดรต์จะสลายอย่างรวดเร็ว โดยจะขึ้นอยู่กับเวลาและระยะทาง ดังนั้นการที่นิวรอนหนึ่งๆ จะถูกกระตุ้นจนถึงขั้นทำงานได้นั้นจะต้องประกอบไปด้วยสัญญาณจากหลายๆ นิวรอนในเวลาและตำแหน่งที่ใกล้เคียงกัน ดังรูปที่ 2.9 ร่างเซลล์หรือโซมา (soma) ทำหน้าที่รวมสัญญาณที่เข้ามาใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
6751

เซลล์จากเด็นไดรต์เมื่อผลรวมของระดับสัญญาณสูงถึงระดับจุดเปลี่ยนที่จะกระตุ้นนิวรอน ตัวนิวรอนจะให้กำเนิดศักย์กระทำ (action potential) นั่น คือ นิวรอนทำงานพร้อมกับส่งค่าศักย์ดังกล่าวผ่านแกนประสาทไปยังนิวรอนอื่นๆ หรือ เซลล์อื่นๆ ที่อยู่นอกระบบประสาท เช่น เซลล์กล้ามเนื้อ เป็นต้น อย่างไรก็ตาม ถ้าผลรวมของสัญญาณไม่ถึงระดับจุดเปลี่ยน สัญญาณที่มีอยู่จะสลายไปอย่างรวดเร็วและไม่ทำให้เกิดการกระตุ้นใดๆ ของนิวรอน ดังนั้นนิวรอนจะมีอยู่เพียงสองสถานะคือทำงานและไม่ทำงาน ความเข้มแข็งของสัญญาณอินพุตที่เข้ามาสู่นิวรอนจะไม่ขึ้นอยู่กับขนาด แต่จะขึ้นอยู่กับความถี่ กล่าวคืออินพุตที่เข้มแข็งกว่าจะสร้างศักย์กระทำในความถี่ที่สูงกว่าอินพุตที่อ่อนแอกว่า



รูปที่ 2.10 การทำงานของนิวรอน

(ก) นิวรอนที่ใช้ในการตรวจจับความเคลื่อนไหว (ข) สัญญาณ ก ข ค ง ถูกกระตุ้นพร้อมๆกันแต่ผลรวมของสัญญาณไม่ถึงระดับที่ทำให้นิวรอนถูกกระตุ้น (ค) สัญญาณถูกกระตุ้นในระดับ ง ค ข ก ทำให้เกิดช่วงยอดของสัญญาณพร้อมกันที่เนินแกนประสาทซึ่งเกินระดับจุดเปลี่ยน จึงทำให้นิวรอนถูกกระตุ้นให้เกิดการทำงาน

2.2.3.2 สถาปัตยกรรมของเครือข่ายประสาทเทียม

สิ่งแรกที่จะต้องพิจารณาในการใช้เครือข่ายประสาทเทียมคือ การศึกษารูปแบบของเครือข่าย โดยเครือข่ายประสาทเทียมที่มีโครงสร้างแตกต่างกันจะมีคุณลักษณะหรือพฤติกรรมแตกต่างกันด้วย โครงสร้างของโครงข่ายประสาทเทียมจะประกอบด้วยสองส่วนใหญ่ๆ ดังนี้

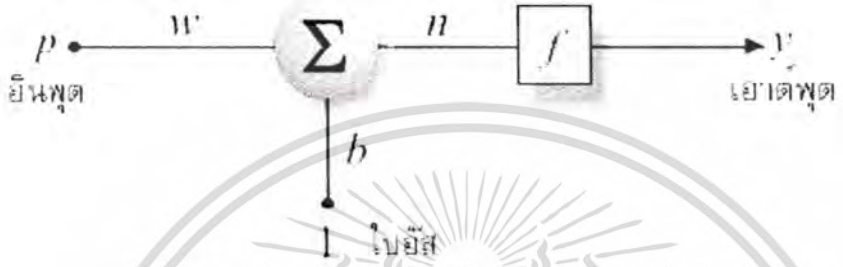
- แบบจำลองของนิวรอน
- สถาปัตยกรรมการเชื่อมต่อเป็นระบบเครือข่ายของนิวรอน

โดยสามารถอธิบายโครงสร้างของโครงข่ายประสาทเทียมที่กล่าวถึงได้ ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. แบบจำลองของนิวรอน : นิวรอนที่มีอินพุตเดียว

โครงสร้างของนิวรอนที่มีอินพุตเดียวแสดงได้ดังรูปที่ 2.11 อินพุต p จะถูกคูณด้วยค่าน้ำหนักประสาท (weight) w โดยมีไบอัส b (bias หรือ อ๊อฟเซ็ต - offset) เป็นอีกอินพุตหนึ่งที่มีค่าน้ำหนักประสาทคงที่เท่ากับ 1 อินพุตที่ได้ทั้งสองจะถูกรวมกันได้เอาต์พุตเป็น n มักจะเรียกเอาต์พุตนี้ว่า เน็ตอินพุต (net input) ซึ่งจะเป็นอินพุตของฟังก์ชั่นถ่ายโอน f (Transfer Function หรือ Activation Function) และจะได้เอาต์พุตของนิวรอนเป็น y



รูปที่ 2.11 แบบจำลองนิวรอนที่มีอินพุตเดียว

เราสามารถคำนวณเอาต์พุตของนิวรอนได้ดังนี้

$$y = f(wp + b) \tag{2.16}$$

เอาต์พุตของนิวรอนขึ้นกับฟังก์ชั่นถ่ายโอน พารามิเตอร์ w , p จะสามารถปรับค่าได้ โดยปกติแล้วฟังก์ชั่นถ่ายโอนจะถูกกำหนดโดยผู้ใช้งาน ส่วนพารามิเตอร์ w และ b จะถูกปรับจากค่า กฎการเรียนรู้ จากแบบจำลองข้างต้น เราสามารถนำมาเปรียบเทียบกับสมองมนุษย์ได้ดังตารางที่ 2.1

ตารางที่ 2.1 แบบจำลองของสมองมนุษย์เทียบกับเครือข่ายประสาทเทียม

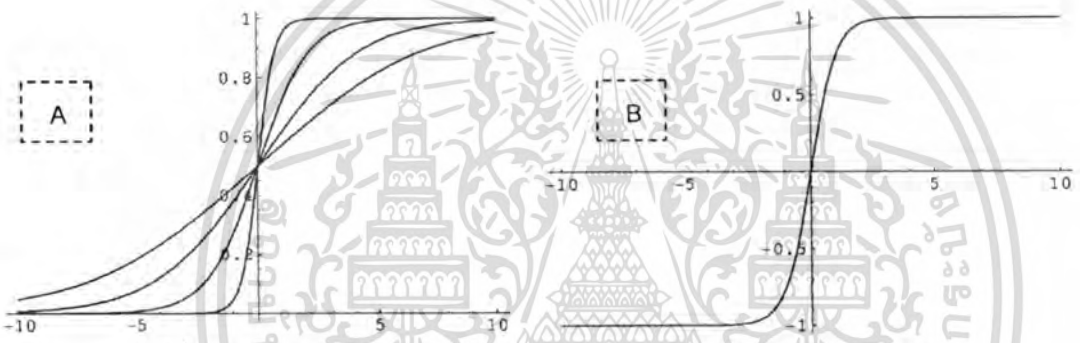
แบบจำลองของสมองมนุษย์	แบบจำลองเครือข่ายประสาทเทียม
ร่างเซลล์ (Soma)	นิวรอน
เด็นไดรต์ (Dendrite)	อินพุต
แกนประสาท (Axon)	เอาต์พุต
จุดประสานประสาท (Synapse)	ค่าน้ำหนัก

2. ฟังก์ชันถ่ายโอนหรือฟังก์ชันกระตุ้น (Transfer Function หรือ Activation Function)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฟังก์ชันถ่ายโอนเป็นส่วนที่ทำหน้าที่รวมค่าเชิงตัวเลขจากเอาต์พุตของนิวรอน แล้วทำการตัดสินใจว่าจะส่งสัญญาณเอาต์พุตออกไปในรูปแบบใด ฟังก์ชั่นถ่ายโอนเป็นได้ทั้งแบบเชิงเส้นและไม่เป็นเชิงเส้น การเลือกฟังก์ชันถ่ายโอนจะขึ้นอยู่กับรูปแบบของระบบ ที่จะนำเอาเครือข่ายประสาทเทียมไปประยุกต์ใช้ โดยในโครงงานนี้จะใช้ฟังก์ชันถ่ายโอน คือ ไฮเปอร์โบลิกแทนเจนท์ (Hyperbolic tangent) ซึ่งเป็นสมการที่รูปร่างเหมือนกับซิกมอยด์ฟังก์ชั่น คือมีรูปร่างเป็นตัว S เช่นเดียวกันแต่จะมีความสมมาตรกว่าคือให้เอาต์พุตอยู่ระหว่าง -1 ถึง 1 โดยค่าเอาต์พุตจะเป็น 0 เมื่อค่า *net* เป็น 0 ค่าเอาต์พุตจะเข้าใกล้ 1 เมื่อค่า *net* เป็นค่าบวก และค่าเอาต์พุตเข้าใกล้ -1 เมื่อค่า *net* เป็นลบ เขียนเป็นสมการได้ดังนี้

$$a = \tanh(net) \tag{2.17}$$



รูปที่ 2.12 กราฟของสมการซิกมอยด์(A) และไฮเปอร์โบลิกแทนเจนท์(B)

3. นิวรอนที่มีหลายอินพุต

แบบจำลองนิวรอนที่มีหลายอินพุต จากรูปที่ 2.13 จะประกอบด้วยอินพุต $p_1, p_2, p_3, \ldots p_n$ โดยแต่ละอินพุตจะมีค่าน้ำหนักเป็นของตัวเองคือ $w_{11}, w_{12}, w_{13}, w_{14}, \ldots w_{1n}$ จากอินพุตและค่านำหนักสามารถเขียนเป็นสมการเน็ทอินพุต n ได้ดังนี้

$$n = w_{11}p_1 + w_{12}p_2 + w_{13}p_3 + \cdots + w_{1n}p_n + b \tag{2.18}$$

ซึ่งจากค่าน้ำหนักประสาทเราสามารถเขียนเป็น เมตริกซ์น้ำหนักประสาท (weight metrix) มีสัญลักษณ์ W

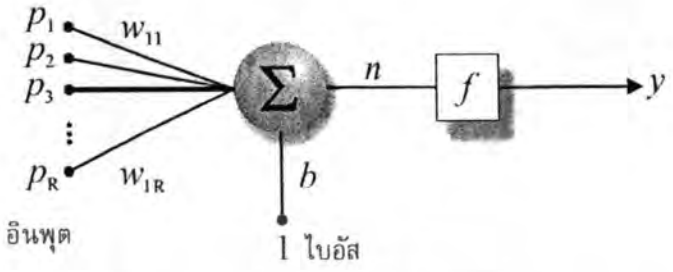
ทำให้เราสามารถเขียนค่าสมการเน็ทอินพุตในรูปเมตริกซ์ได้ดังนี้

$$n = Wp + b \tag{2.19}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นเอาต์พุตของนิวรอนสามารถเขียนได้ดังนี้

$$Y = f(Wp + b) \tag{2.20}$$



รูปที่ 2.13 แผนภาพนิวรอนที่มีหลายอินพุต

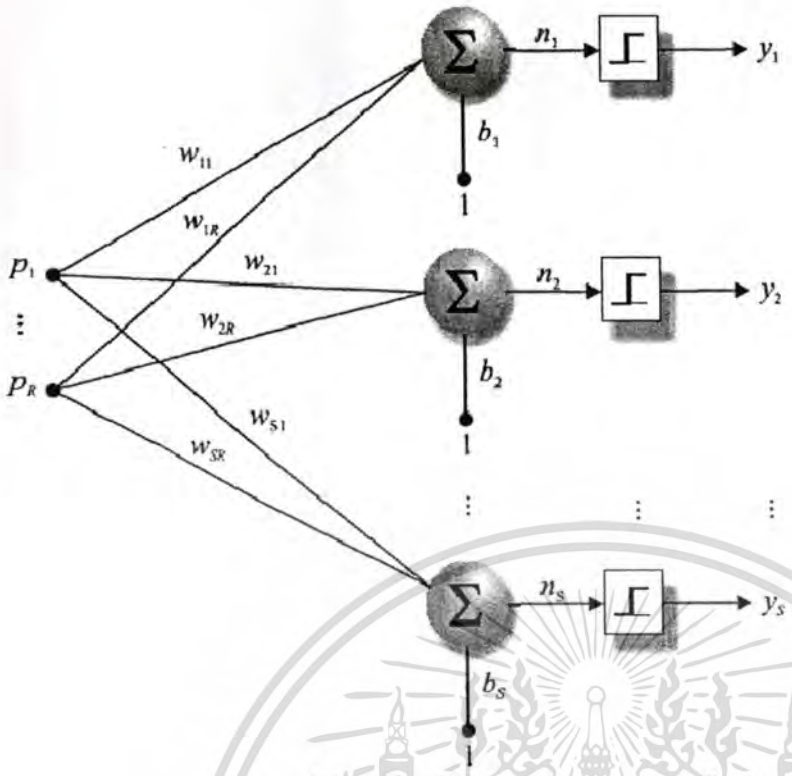
2.2.3.3 รูปแบบการเชื่อมโยงของเครือข่ายประสาทเทียม

จากที่กล่าวผ่านมาเป็นหลักการและเหตุผลของการสร้างเซลล์ประสาทเพียงเซลล์เดียว การที่จะนำเซลล์ประสาทเทียมมาใช้งานได้นั้นจะต้องใช้เซลล์ที่มีคุณลักษณะต่างๆกัน (ค่าน้ำหนักประสาทจะทำให้เซลล์มีคุณลักษณะที่ต่างกัน) มาเชื่อมโยงเป็นเครือข่าย ลักษณะการเชื่อมโยงของเครือข่ายจะมีอยู่สองประเภท คือ

ก. เป็นเครือข่ายประสาทเทียมแบบชั้นเดียว (Single Layer Artificial Neural Network)

เครือข่ายประสาทเทียมแบบชั้นเดียวจะประกอบด้วย เซลล์ประสาทเทียมแบบง่ายๆหลายชุดที่เชื่อมต่อกันระหว่างชั้นข้อมูลเข้า(input layer) และชั้นข้อมูลออก(output layer) เท่านั้น เพราะฉะนั้นโครงข่ายประสาทเทียมแบบชั้นเดียวจะมีประสิทธิภาพในการคำนวณน้อยแต่จะได้ประสิทธิภาพทางด้านเวลาที่มากกว่า โดยลักษณะของโครงข่ายสามารถแสดงได้ดังรูปที่

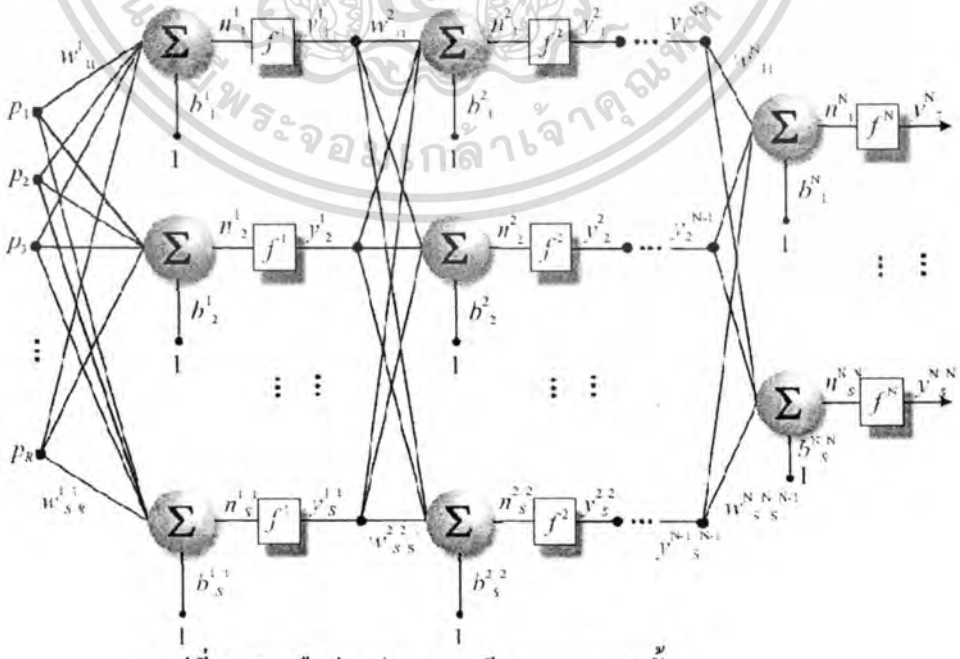
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.14 เครือข่ายประสาทเทียมแบบชั้นเดียว

ข. เครือข่ายประสาทเทียมแบบหลายชั้น (Multilayer Artificial Neural Network)

เครือข่ายประสาทเทียมแบบหลายชั้นจะมีโครงสร้างที่ซับซ้อนมากยิ่งขึ้นแต่ความสามารถคำนวณก็จะดีขึ้น โครงข่ายแบบหลายชั้นอาจจะสร้างจากกลุ่มของโครงข่ายแบบชั้นเดียว โดยเอาต์พุตขั้นหนึ่งจะไปเป็นอินพุตให้กับอีกชั้นถัดไป ดังรูป



รูปที่ 2.15 เครือข่ายประสาทเทียมแบบหลายชั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.15 แสดงเครือข่ายประสาทเทียมแบบหลายชั้น โดย ในเครือข่ายประสาทเทียมมีการเรียกชื่อชั้นต่างๆดังนี้ คือ ชั้นที่ต่อโดยตรงกับอินพุต เรียกว่าชั้นข้อมูลเข้า ชั้นที่อยู่ขวามือสุด คือ ชั้นข้อมูลออกจะเป็นชั้นที่เครือข่ายจะให้ผลลัพธ์ และชั้นที่อยู่ระหว่างชั้นข้อมูลขาเข้าและขาออกไม่ว่าจะกี่ชั้นก็ตามจะเรียกว่า ชั้นซ่อนตัว (Hidden Layer)

2.2.3.4 การเรียนรู้ของเครือข่ายประสาทเทียม

การเกิดความรู้ความเข้าใจในสมองของมนุษย์นั้นจะขึ้นกับการนำความรู้ความเข้าใจเดิมมาสัมพันธ์กับความรู้ใหม่ที่ได้รับ โดยเราสามารถสร้างความสัมพันธ์ของการเรียนรู้ของมนุษย์กับการเรียนรู้ของเครือข่ายประสาทเทียมได้ว่า เด็กแรกเกิดเริ่มแรกนั้นยังไม่มีการเรียนรู้ใด มีเพียงแต่สัญชาติญาณที่ธรรมชาติสร้างให้มาเท่านั้น แต่เด็กจะมีการพัฒนาการเรียนรู้เป็นขั้นเป็นตอนต่อเพื่อตอบสนองต่อสิ่งเร้า เมื่อสมองของเด็กได้รับการฝึกสอนและเจริญเติบโตขึ้น เซลล์สมองก็จะได้รับการปรับคุณลักษณะสอดคล้องกับการฝึกสอน และเจริญเป็นโครงข่ายที่สอดคล้องกัน

ในเครือข่ายประสาทเทียมก็เช่นเดียวกัน คือ ค่าน้ำหนักจะเป็นค่าที่ใช้แทนความรู้ โดยเมื่อเริ่มต้นเซลล์ทุกเซลล์จะไม่มีคุณลักษณะใดๆเลย เนื่องมาจากว่ายังไม่มีการกำหนดค่าน้ำหนักที่เหมาะสมให้กับเซลล์เพื่อนำไปใช้กับงานที่ต้องการ โดยในตอนเริ่มต้น ค่าน้ำหนักอาจจะเป็นค่าสุ่มขึ้นมาหรือให้เป็นไปตามสมการใดสมการหนึ่ง เนื่องจากค่าน้ำหนักเป็นค่าสุ่มทำให้เครือข่ายไม่แสดงคุณลักษณะใดๆออกมา การฝึกสอนเครือข่ายคือการปรับค่าน้ำหนักทุกๆจุดให้สอดคล้องกับอินพุตหลายๆแบบเพื่อให้ได้เอาท์พุตตามความต้องการ ซึ่งจะคล้ายกับการเรียนรู้ของมนุษย์ที่ในบางเรื่องไม่สามารถที่จะเข้าใจหรือจดจำในการอ่านเพียงรอบเดียว จึงต้องมีการนำเสนอตัวอย่างและทบทวนหลายครั้งการเรียนรู้จึงจะดีขึ้น ในทำนองเดียวกันบนเครือข่ายประสาทเทียมก็จะมีการนำค่าความผิดพลาดที่ได้จากการเรียนรู้รอบที่แล้วมาอ้างอิงสำหรับการปรับปรุงค่าน้ำหนักในรอบถัดไป หลังจากปรับค่าน้ำหนักจนได้ค่าผิดพลาดที่เอาต์พุตเทียบกับเป้าหมายน้อยลงจนเป็นที่พอใจแล้ว เครือข่ายประสาทเทียมก็พร้อมที่จะวิเคราะห์อินพุตและให้เอาต์พุตตามลักษณะตัวอย่างที่เคยได้เรียนรู้มาจากการฝึกสอน

ดังนั้นการเรียนรู้ของเครือข่ายประสาทเทียมจะมีประสิทธิภาพมากน้อยเพียงใดขึ้นอยู่กับค่าน้ำหนักของเครือข่าย ซึ่งการฝึกสอนก็คือการหาค่าน้ำหนักที่เหมาะสมให้แก่โครงข่ายนั้นๆ โดยการฝึกสอนจะมีอยู่สองประเภท คือ

1. การเรียนรู้แบบมีผู้ฝึกสอน (Supervised Learning)

คุณลักษณะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- เรียนรู้ที่จะสร้างผลลัพธ์ที่ต้องการให้ได้ตามตัวอย่างที่ได้รับหรือตามเป้าหมายที่กำหนดไว้

2. การเรียนรู้แบบไม่มีผู้ฝึกสอน (Unsupervised Learning)

คุณลักษณะ

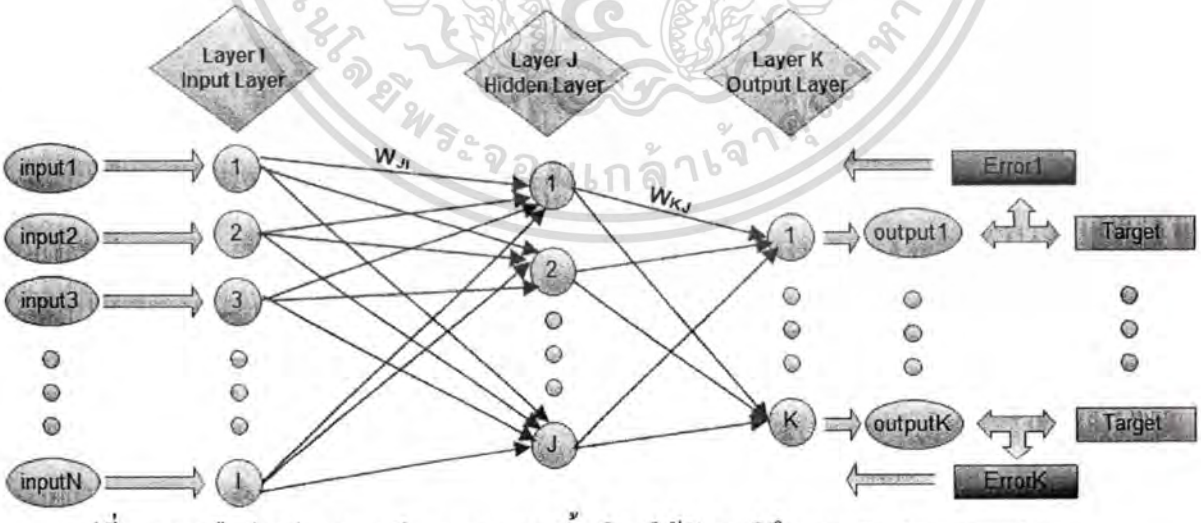
- จัดการข้อมูลอินพุตของระบบด้วยตัวเอง
- ค้นหาคุณลักษณะของตัวเองจากอินพุต

2.2.3.5 เครือข่ายประสาทเทียมที่ใช้ในโครงงาน

ในโครงงานนี้จะใช้การเรียนรู้แบบ Supervised Learning โดยมีรูปแบบการเชื่อมโยงของเครือข่ายแบบ Multilayer Neural Network และมีอัลกอริธึมในการเรียนรู้แบบ Backpropagation โดยมีรายละเอียดดังนี้

จากรูปที่ 2.16 จะประกอบด้วยสามชั้นคือ ชั้น input layer (Layer I) , ชั้น Hidden Layer (Layer J) , ชั้น output layer(Layer K)

ค่าน้ำหนัก (W) จะใช้การเรียกชื่อจากขวามือมาทางซ้ายมือ เช่น ในชั้น input-hidden : W_{JI} ในชั้น Hidden-Output : W_{KJ}



รูปที่ 2.16 เครือข่ายประสาทเทียมแบบหลายชั้น โดยใช้อัลกอริธึมแบบ Backpropagation

1. **เครือข่ายที่ส่งสัญญาณไปข้างหน้า (FeedForward)** การทำงานของ Feed Forward จะเป็นการเริ่มจากชั้นอินพุตไปจนถึงชั้นเอาต์พุต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- **หาเอาต์พุต ชั้นซ่อนตัว**

จากรูปที่ 2.16 (แทนอินพุตด้วย X) อินพุต x_1 จนถึง x_i จะถูกคูณด้วยเมตริกซ์ค่าน้ำหนัก ซึ่งผลรวมของค่าอินพุตคูณกับค่าน้ำหนักของแต่ละเซลล์ประสาทเทียมเรียกว่า NET สามารถเขียนเป็นสมการได้ดังนี้

$$NET_j = x_1 w_{j1} + x_2 w_{j2} + \cdots + x_i w_{ji} \tag{2.21}$$

หรือ

$$NET_j = \sum_{i=1}^{I} x_i \, w_{ji}$$

โดย X_i เป็นค่าที่รับเข้ามาของอินพุตโหนดที่ i โดยที่ $i=1,2,\ldots,I$

W_{ji} คือค่าน้ำหนักประจำตัวของแต่ละอินพุตโหนด i ที่ส่งเข้าไปยังโหนด j ในชั้นถัดไป โดย $j=1,2,\ldots,J$

เมื่อได้ค่า NET ของแต่ละเซลล์ประสาทเทียมแล้วจะปรับให้ได้เอาท์พุตอยู่ในช่วงที่ต้องการ ซึ่งจะเป็นค่าที่จะเป็นอินพุตให้กับเซลล์ประสาทเทียมตัวถัดไป โดยผ่านฟังก์ชันถ่ายโอนที่กำหนดในรูปแบบไฮเปอร์บอลิคแทนเจนท์ (Hyperbolic tangent) ดังจะได้สมการตามนี้

$$Output_j = O_j = F_j(NET_j) = \tanh(NET_j) \tag{2.22}$$

- **หาเอาต์พุตชั้นเอาต์พุต** จะมีขั้นตอนการทำเช่นเดียวกับชั้นอินพุต-ชั้นซ่อนตัว

$$NET_k = \sum_{j=1}^{J} O_j \, W_{kj} \tag{2.23}$$

โดย O_j เป็นค่าที่รับเข้ามาของอินพุต(คือค่าเอาท์พุตของชั้นซ่อนตัว)โหนดที่ j โดยที่ $j=1,2,\ldots,J$

W_{kj} คือค่าน้ำหนักประจำตัวของแต่ละอินพุตโหนด j ที่ส่งเข้าไปยังโหนด k ในชั้นถัดไป โดย $k=1,2,\ldots,K$

$$Output_k = O_k = F_k(NET_k) = \tanh(NET_k) \tag{2.24}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. เครือข่ายที่ส่งสัญญาณย้อนกลับ (Reverse Pass)

เป็นส่วนที่ป้อนค่าความผิดพลาดกลับ จากชั้นข้อมูลออกย้อนกลับมาปรับปรุงค่าน้ำหนักที่ชั้นก่อนหน้า โดยสามารถเขียนเป็นลำดับขั้นตอนเครือข่ายที่มีการป้อนกลับสรุปได้ดังนี้

- **การปรับปรุงค่าน้ำหนักในชั้น Hidden - Output**

1. นำเอาท์พุตที่ได้จากชั้นข้อมูลออกไปลบออกจากค่าเป้าหมายได้เป็นค่าความผิดพลาด

$$\delta = Target - output_k \tag{2.25}$$

2. นำค่าความผิดพลาดในสมการที่ 2.25 มาคูณกับ F' ที่ได้จากชั้นข้อมูลออก โดย F' จากในโครงงานคือการ differential ของ tanh(x) ซึ่งเท่ากับ $sech^2(x)$

$$Error_k = \delta \times sech^2(NET_k) \tag{2.26}$$

3. นำค่า $Error_k$ มาคูณกับ $output_j$ แล้วคูณกับค่าอัตราการเรียนรู้ (η)

$$\Delta W_{kj} = Error_k \times output_j \times \eta \tag{2.27}$$

4. ค่าน้ำหนักค่าใหม่ จะเท่ากับ ค่าน้ำหนักค่าเดิมบวกกับค่าในสมการที่ 2.27

$$Weigth^{new} = \Delta W_{kj} + Weigth^{old} \tag{2.28}$$

- **การปรับปรุงค่าน้ำหนักในชั้น Input - Hidden**

1. นำค่า Error มาคูณกับค่าน้ำหนักของชั้น Hidden-Output

$$\delta = \sum_{k=0}^{K} Error_k W_{jk} \tag{2.29}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. นำค่าความผิดพลาดในสมการที่ 2.29 มาคูณกับ F' ที่ได้จากชั้นข้อมูลออก โดย F' จากในโครงงานคือการ differential ของ tanh(x) ซึ่งเท่ากับ $\text{sech}^2(x)$

$$Error_j = \delta \times sech^2(NET_j) \tag{2.30}$$

3. นำค่า Error_j มาคูณกับ input แล้วคูณกับค่าอัตราการเรียนรู้ (η)

$$\Delta W_{ji} = Error_j \times X_j \times \eta \tag{2.31}$$

4. ค่าน้ำหนักค่าใหม่ จะเท่ากับ ค่าน้ำหนักค่าเดิมบวกกับค่าในสมการที่ 2.31

$$Weigth^{new} = \Delta W_{ji} + Weigth^{old} \tag{2.32}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การออกแบบระบบ

3.1 การออกแบบระบบรู้จำเสียงเบื้องต้น

ในส่วนการออกแบบเบื้องต้นก่อนเข้าสู่กระบวนการรู้จำเสียง สามารถแบ่งเป็นขั้นตอนได้ดังนี้

3.1.1 การออกแบบและสร้างคำเพื่อเป็นส่วนที่ใช้ในการรู้จำและทดสอบ

- ในขั้นตอนนี้จะเป็นขั้นตอนการสร้างคำต้นแบบ โดยใช้คนทดลองทั้งหมด 100 คน โดยในจำนวน 100 คนจะประกอบด้วยเพศหญิงและเพศชาย โดยจะเป็นคำโดดพยางค์เดียวที่เป็นต้นฉบับทั้งหมด 10 คำ คือ เปิด ปิด ไฟ แอร์ ปั๊ม ปลั๊ก เพลง รั้ว ม่าน เตา โดยจากทั้ง 10 คำนี้ให้ทุกคน ออกเสียงคำละ 5 ครั้ง เพราะฉะนั้นจะมีจำนวนคำที่อัดทั้งหมด

  100 คน * 10 คำ * 5 ครั้ง = 5000 คำ

- ในจำนวนที่อัด 100 คนนี้ จะแบ่งเป็น 70:30 คือ 70 คือส่วนที่จะนำมาใช้ในกระบวนการรู้จำเสียง และ 30 จะนำมาใช้สำหรับทดสอบ

3.1.2 รายละเอียดของไฟล์เสียงข้อมูล

รายละเอียดของไฟล์เสียงที่นำมาใช้เก็บข้อมูลหลังจากสร้างคำมาแล้ว เป็นดังนี้

รูปแบบของไฟล์เสียงข้อมูล : ".WAV"

จำนวนบิทข้อมูล : 16 บิท

ความถี่ในการชักตัวอย่าง (sampling frequency) : 44,100 Hz

โดยไฟล์ข้อมูลเสียงที่ทำการบันทึกจะนำมาเข้าสู่กระบวนการรู้จำเสียงและทดสอบ

3.1.3 เครื่องมือและอุปกรณ์ที่ใช้ในการบันทึกไฟล์เสียง

ใช้เครื่องอัดเสียงของ Sony รุ่น ICD-SX700

คอมพิวเตอร์ 1 เครื่องเพื่อใช้ในการถ่ายโอนเสียงจากเครื่องอัดเสียง

จำนวนที่บันทึกต่อ 1 คน คือ 10 คำ คำละ 5 ครั้ง

โดยการบันทึกเสียงมีขอบเขตดังนี้

สภาพแวดล้อมขณะทำการบันทึกต้องเงียบ ไม่มีเสียงรบกวน

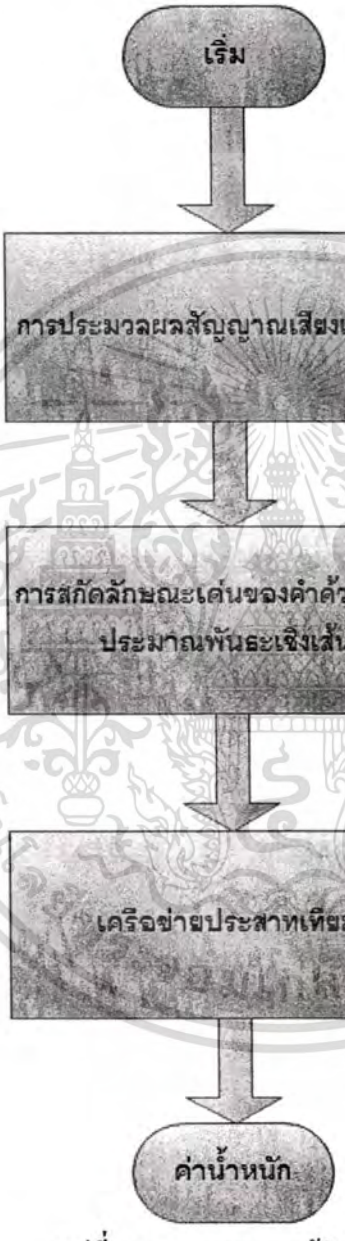
ลักษณะของสำเนียงที่พูดเป็นภาษากลาง

ไม่มีการแสดงอารมณ์ขณะทำการบันทึก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 ขั้นตอนกระบวนการรู้จำเสียง

หลังจากทำการบันทึกเสียงได้แล้ว กระบวนการนี้จะเป็นการนำเสียงที่ได้มาวิเคราะห์หาลักษณะสำคัญและนำลักษณะสำคัญมาเข้าเครือข่ายประสาทเทียม โดยหลักการที่ใช้ในการทำขั้นตอนดังแผนภาพมีดังนี้

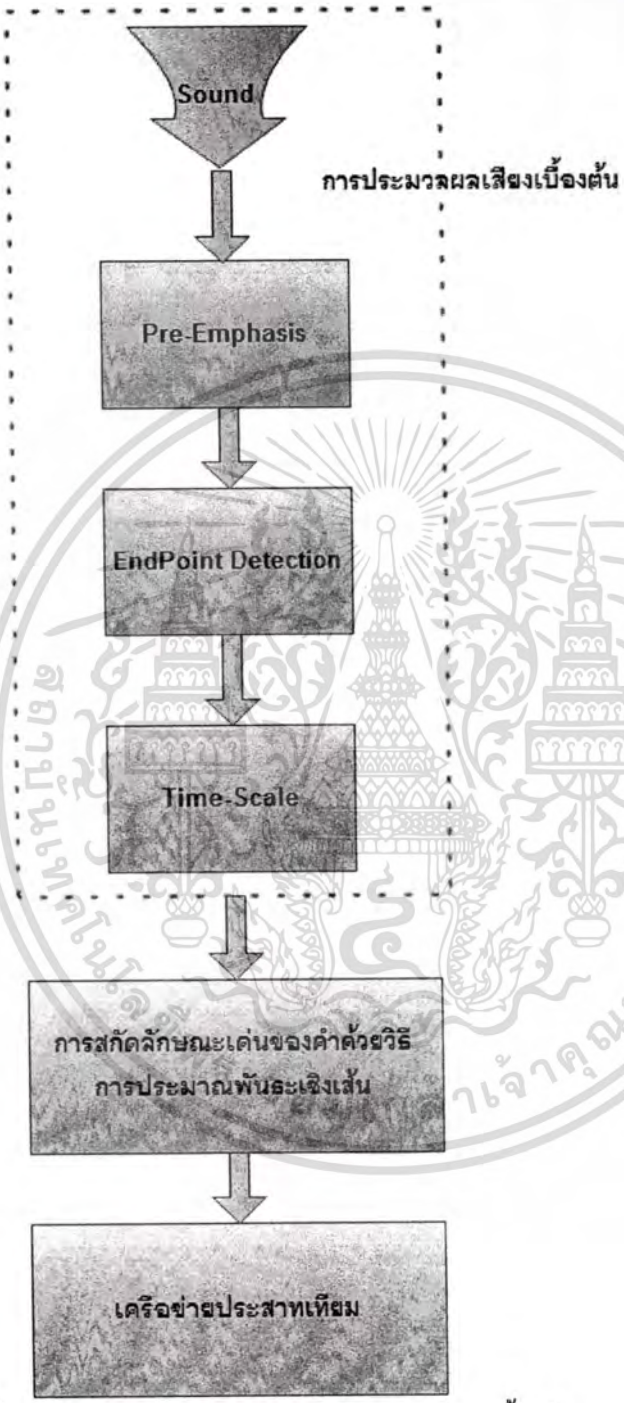


รูปที่ 3.1 กระบวนการรู้จำเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.1 กระบวนการประมวลผลสัญญาณเสียงเบื้องต้น

ในกระบวนการประมวลผลสัญญาณเสียงเบื้องต้น มีขั้นตอนย่อยภายในดังนี้



รูปที่ 3.2 แผนภาพกระบวนการประมวลผลเสียงเบื้องต้น

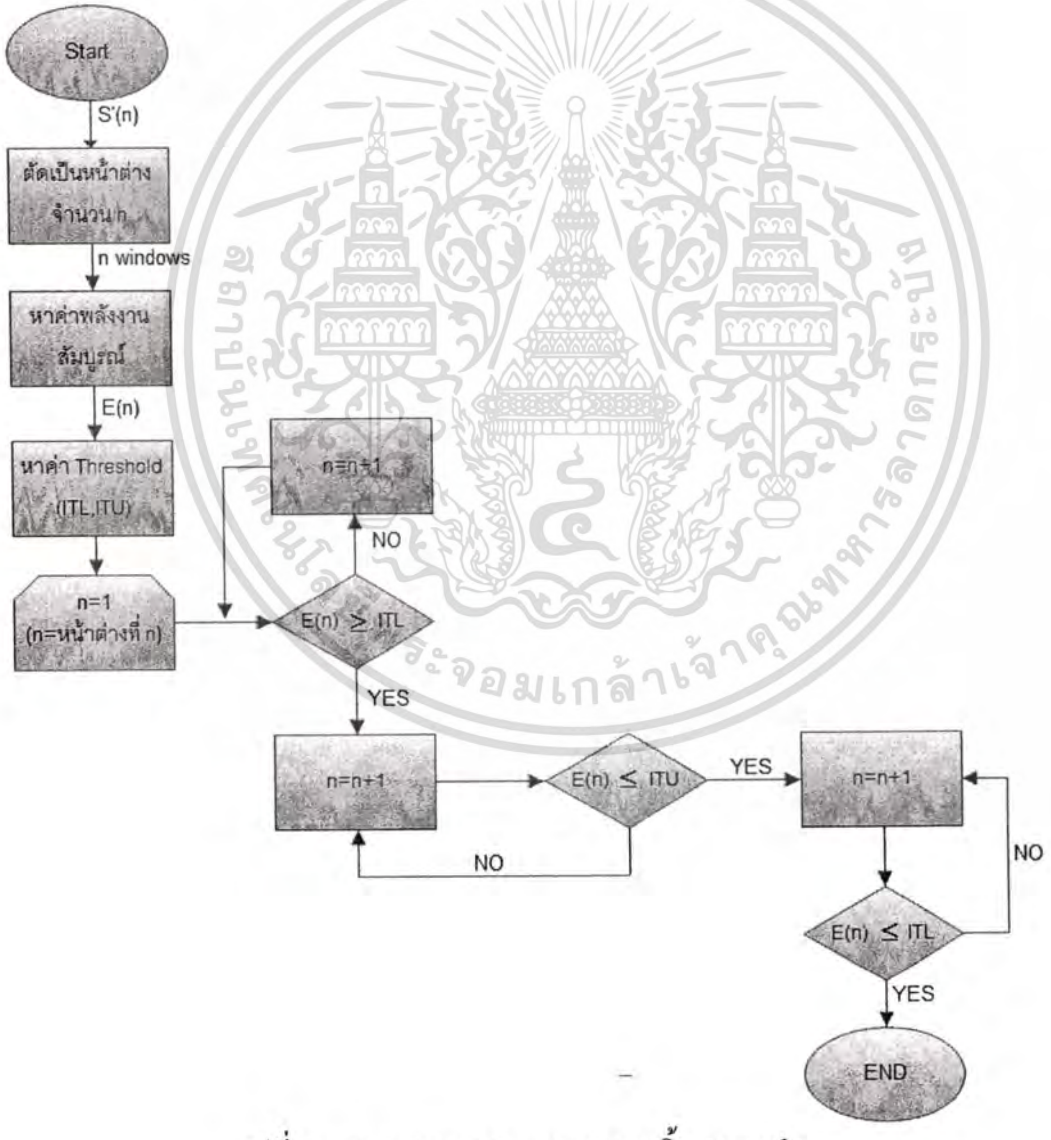
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- **กรรมวิธีเน้นล่วงหน้า(Pre-Emphasis)**

ในกระบวนการนี้เราจะนำเสียงมาผ่านกรรมวิธีเน้นล่วงหน้า โดยจากสมการที่ 2.1 เราจะกำหนดค่าสัมประสิทธิ์ของวงจรกรอง a โดยค่า a ยิ่งค่าเข้าใกล้ 1 เท่าใด สัญญาณที่ความถี่สูงก็จะถูกขยายมากขึ้นเท่านั้น ค่า a ที่นิยมใช้สำหรับการหาค่าสัมประสิทธิ์ของ LPC คือ $\frac{15}{16} = 0.9375$

- **กระบวนหาจุดสิ้นสุดของคำพูด (Endpoint Detection)**

ในกระบวนการหาจุดสิ้นสุดของคำพูดจะใช้การตัดคำโดยใช้ค่าพลังงาน(Short-Time Energy) โดยใช้โปรแกรมแมทแลปในการเขียนอัลกอริธึมขึ้นมา โดยมีการออกแบบไว้ตามแผนภาพดังนี้



รูปที่ 3.3 แผนภาพกระบวนการหาจุดสิ้นสุดของคำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. นำสัญญาณเสียงที่ผ่านการเอมฟาซิสแล้ว แทนด้วย $S'(n)$ นำมาตัดแบ่งออกเป็นหน้าต่าง หน้าต่างละ 441 ตัวอย่างสัญญาณ โดยค่า 441 ตัวอย่างสัญญาณต่อ 1 หน้าต่าง มาจาก การที่ได้กำหนดเสียงที่อัดมีความถี่ในการชักตัวอย่าง 44100 เฟรม/วินาที และต้องการให้ 1 หน้าต่างมีความยาว 10 ms เพราะฉะนั้น $44100\ frams/\sec \times 10ms = 441$
2. นำสัญญาณเสียงมาหาค่าพลังงานสัมบูรณ์จากสมการที่ 2.7 ของแต่ละหน้าต่างแล้วทำการเก็บค่าไว้
3. ทำการกำหนดค่า Threshold ซึ่งประกอบด้วยค่า ITL และ ITU เพื่อนำมาหาจุดหัวและท้ายของสัญญาณเสียง โดยมีขั้นตอนดังนี้

กำหนดค่า ITL จาก

IMIN = เป็นการหาค่าเฉลี่ยของค่าพลังงานที่น้อยที่สุด 10 หน้าต่าง

IMAX = เป็นการหาค่าเฉลี่ยของค่าพลังงานที่มากที่สุด 10 หน้าต่าง

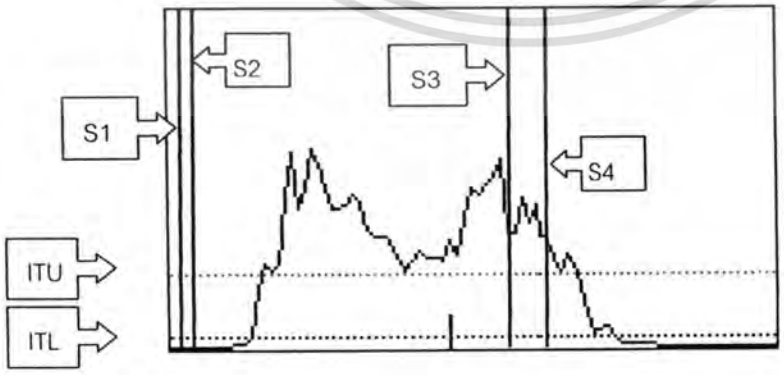
$$I_1 = 0.05 \times (IMX - IMN) + IMN$$
$$I_2 = 6IMN$$

แล้วนำค่า I_1 และ I_2 มาหาค่าน้อยที่สุดแล้วกำหนดให้เป็นค่า ITL

ดังนั้น จะกำหนดค่า ITU,ITL จาก

$$ITL = \min(I_1, I_2)$$
$$ITU = 5 \cdot ITL$$

4. วิธีการตัดคำหลังจากได้ทำการหาค่า Threshold มาเรียบร้อยแล้วสามารถอธิบายได้ดังนี้ หลังจากได้คำนวณค่า Threshold ขึ้นมาแล้ว เมื่อกรอบหน้าต่างของสัญญาณมีค่าเท่ากับค่า Thershold จะทำให้เกิดจุดตัดเกิดขึ้น โดยจุดตัดที่เกิดขึ้นทำให้สามารถคำนวณหัวท้ายของคำได้ จะมีอยู่ด้วยกัน 4 จุด จากรูปที่ 3.4 จะเกิดจุดตัดระหว่างเส้น Threshold กับ สัญญาณเสียงที่เข้ามา โดยจุด S1,S4 จะเกิดจากสัญญาณตัดเส้น ITL ส่วน S2,S3 เกิดจากสัญญาณตัดเส้น ITU



รูปที่ 3.4 แผนภาพของค่าพลังงานที่ได้จากการวิธีการหาค่าพลังงานสัมบูรณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

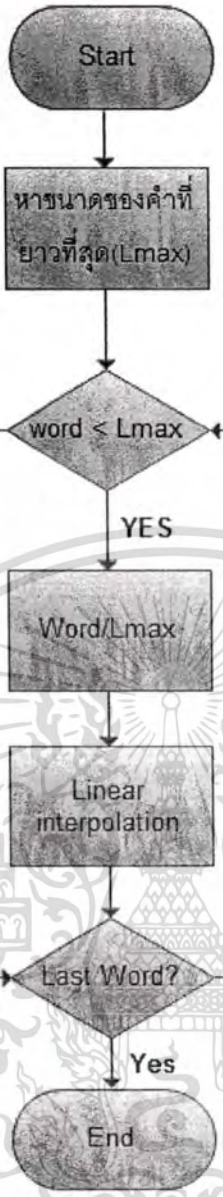
หาจุดเริ่มต้นของเสียง : จะเป็นค่าแรกของพลังงานที่เกินค่า *ITL* ขึ้นมาโดยเงื่อนไขที่ทำให้เป็นค่าเริ่มแรกคือ หลังจากเจอค่าแรกที่เกินค่า *ITL* (จุด S1) ขึ้นมาให้เก็บไว้ในตัวแปรไว้ก่อนแล้วตรวจสอบค่าพลังงานว่าได้ผ่านตำแหน่ง S2,S3,S4 ตามลำดับหรือไม่ ถ้าเป็นไปตามลำดับให้ถือว่าค่าตรงจุด s1 เป็นจุดเริ่มต้น ซึ่งถ้าในกรณีที่ ผ่านจุด S1 แล้ว ไม่ผ่านจุด S2,S3 ไปเจอจุด S4 จะไม่ถือว่าเป็นคำที่ต้องการ ให้ยกเลิกจุด S1 ในตัวแปรทิ้งไปแล้วทำการหาค่าจุด S1 ใหม่อีกครั้ง

จุดสิ้นสุดของคำ : หลังจากได้จุดเริ่มต้นของคำมาแล้ว แสดงว่าจุด S4 ถือว่าเป็นจุดสิ้นสุดของคำ เพราะฉะนั้นคำๆนี้จะเริ่มตั้งแต่จุด S1 จนถึง S4 หลังจากได้จุดเริ่มต้นและสิ้นสุดของคำแล้วให้ทำการบันทึกเสียงที่ได้จากการหาจุดสิ้นสุดของคำในรูปแบบ WAV ที่มีจำนวนบิทข้อมูล 16 บิท และความถี่ในการชักค่าตัวอย่างข้อมูลเท่ากับ 44100 Hz

- **การปรับบรรทัดฐานทางเวลาของคำ (Time Scale)**

ในการปรับบรรทัดฐานทางเวลาจะใช้วิธีการการประมาณค่าในช่วงเชิงเส้น(Linear interpolation) ที่ได้กล่าวถึงในบทที่สอง โดยมีวิธีการดังแผนภาพนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



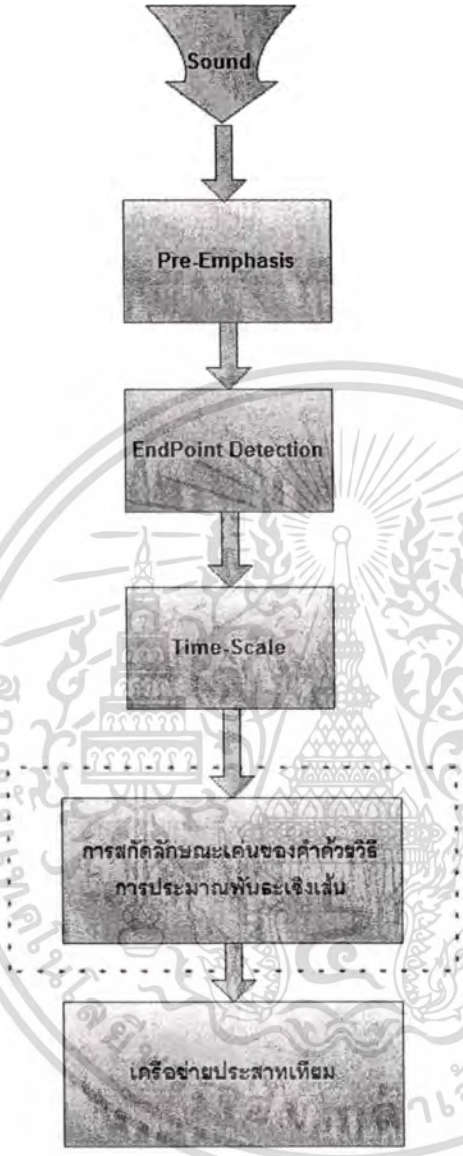
รูปที่ 3.5 กระบวนการปรับบรรทัดฐานทางเวลาของคำ

ขั้นแรก นำคำที่ผ่านการหาจุดสิ้นสุดของคำ มาวิเคราะห์หาว่าคำได้มีความยาวของคำมากที่สุดจากคำทั้งหมด และทำการหาค่าตัวเลขความยาวของคำๆนั้น(Lmax)

ขั้นที่สอง นำ Lmax ที่ได้มาเป็นมาตรฐานเพื่อปรับค่าให้คำอื่นๆ ให้มีบรรทัดฐานทางเวลาของคำให้เท่ากับคำที่มีความยาวมากที่สุด โดยการปรับจะใช้วิธีการประมาณค่าในช่วงเชิงเส้น(Linear Interpolation)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากเสร็จสิ้นขั้นตอนกระบวนการประมวลผลสัญญาณเสียงเบื้องต้นแล้วจะนำเสียงที่ได้มาทำการสกัดลักษณะเด่นของคำ (Feature Extraction)



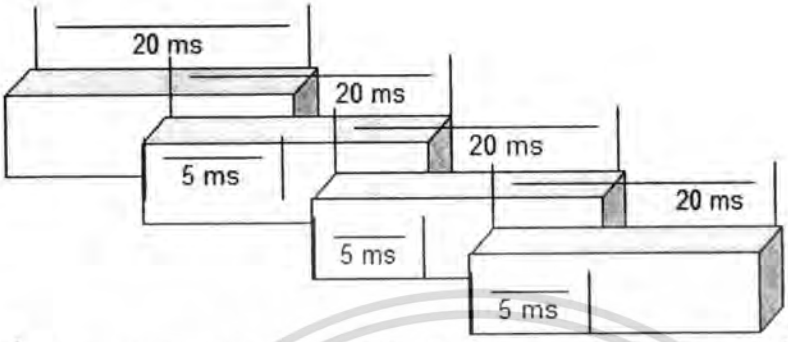
รูปที่ 3.6 แผนภาพขั้นตอนหลังจากกระบวนการประมวลผลเสียงเบื้องต้น

3.2.2 การสกัดลักษณะเด่นของคำด้วยวิธีการประมาณพันธะเชิงเส้น(Linear Predictive Coding)

หลังจากผ่านกระบวนการประมวลผลสัญญาณเสียงเบื้องต้นแล้ว จะนำค่าเสียงที่ได้มาสกัดลักษณะเด่นของคำนั้นๆด้วยวิธีการประมาณพันธะเชิงเส้น(LPC) โดยจะมีขั้นตอนย่อยๆภายในดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- กำหนดการแบ่งคำเป็นหน้าต่างสัญญาณ โดยใช้การวางกรอบหน้าต่างสัญญาณแบบแฮมมิ่ง โดยกำหนดค่ากรอบสัญญาณแต่ละหน้าต่างยาว 20 ms และให้มีการซ้อนทับกันระหว่างหน้าต่างก่อนหน้ากับหน้าต่างปัจจุบัน โดยช่วงของการซ้อนทับระหว่างหน้าต่างยาว 5 ms



รูปที่ 3.7 แผนภาพการวางกรอบหน้าต่างสัญญาณและการซ้อนทับของหน้าต่าง

- ทำการทดสอบค่าอันดับของการประมาณพันธะเชิงเส้น หลายๆอันดับเพื่อทำการหาค่าที่เหมาะสม ที่ให้ค่าสัมประสิทธิ์ของการประมาณพันธะเชิงเส้นเมื่อกระตุ้นด้วยสัญญาณ impulse แล้วได้สัญญาณใกล้เคียงกับสัญญาณเสียงที่นำมาวิเคราะห์มากที่สุด

หลังจากได้ค่าสัมประสิทธิ์การประมาณพันธะเชิงเส้นจากแต่ละอันดับ วิธีการที่จะดูว่าค่าอันดับของการประมาณพันธะเชิงเส้นค่าไหนที่เหมาะสม คือ นำค่าสัมประสิทธิ์การประมาณพันธะเชิงเส้นมาทำการกระตุ้นด้วยค่า impulse แล้วทำการแปลงค่านั้นให้อยู่ในโดเมนความถี่เปรียบเทียบกับเสียงต้นฉบับในโดเมนความถี่เช่นกัน ซึ่งเมื่อได้ค่าของอันดับของสัมประสิทธิ์ที่เหมาะสมแล้วก็ทำการสกัดค่าสัมประสิทธิ์ออกมา เพื่อเข้าสู่กระบวนการจดจำรูปแบบและลักษณะของคำแต่ละคำโดยใช้เครือข่ายประสาทเทียม

3.2.3 เครือข่ายประสาทเทียม (Neural Network)

ในกระบวนการเครือข่ายประสาทเทียม เป็นการหาความคล้ายคลึงกันของรูปแบบของข้อมูลสัมประสิทธิ์การประมาณพันธะเชิงเส้นในคำแต่ละคำ โดยในขั้นตอนการสกัดลักษณะเด่นนั้นจะมีจำนวนสัมประสิทธิ์ของข้อมูลเท่ากับ ค่าอันดับของการประมาณพันธะเชิงเส้น× จำนวนกรอบสัญญาณหน้าต่าง ซึ่งค่าที่ได้จะมีจำนวนสัมประสิทธ์จากการประมาณพันธะเชิงเส้นที่เท่ากันทั้งหมด เพราะว่าได้มีการปรับบรรทัดฐานทางเวลาของคำแต่ละคำให้เท่ากันแล้ว ซึ่งจากข้อมูลทั้งหมดสามารถออกแบบเครือข่ายประสาทเทียมได้ดังนี้

จำนวนอินพุตในชั้นข้อมูลเข้า : ค่าอันดับของการประมาณพันธะเชิงเส้น × จำนวนกรอบหน้าต่างสัญญาณทั้งหมด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จำนวนชั้นซ่อนตัว : 1 ชั้น จำนวนโหนดขึ้นกับการทดสอบ

จำนวนเอาท์พุตในชั้นข้อมูลออก : 10 โหนด เนื่องมาจากมีจำนวนคำที่ใช้ในการทดสอบเท่ากับ 10 คำ คือ ปิด เปิด ไฟ แอร์ เพลง ปั๊ม ปลั๊ก รั้ว เตา ม่าน

ค่าเป้าหมายที่ใช้เป็นส่วนของการฝึกสอน สามารถกำหนดค่าตามคำต้นแบบได้ดังนี้

ตารางที่ 3.3 ตารางกำหนดค่าเป้าหมาย

	โหนดของข้อมูลออกที่									
	1	2	3	4	5	6	7	8	9	10
เปิด	1	0	0	0	0	0	0	0	0	0
ปิด	0	1	0	0	0	0	0	0	0	0
ไฟ	0	0	1	0	0	0	0	0	0	0
แอร์	0	0	0	1	0	0	0	0	0	0
เพลง	0	0	0	0	1	0	0	0	0	0
ปั๊ม	0	0	0	0	0	1	0	0	0	0
ปลั๊ก	0	0	0	0	0	0	1	0	0	0
เตา	0	0	0	0	0	0	0	1	0	0
รั้ว	0	0	0	0	0	0	0	0	1	0
ม่าน	0	0	0	0	0	0	0	0	0	1

เมื่อได้กำหนดค่าเป้าหมาย กำหนดโหนดในแต่ละชั้นแล้วก็จะนำค่าสัมประสิทธิ์การประมาณพันธะเชิงเส้นของแต่ละคำมาเข้าสู่เครือข่ายประสาทเทียม และให้เครือข่ายประสาทเทียมทำงานโดยการหาค่าเอาต์พุตเทียบกับค่าเป้าหมายแล้วปรับค่าน้ำหนักให้ได้ค่าที่ทำให้เอาต์พุตที่ได้มีค่าใกล้เคียงกับค่าเป้าหมายที่สุด โดยเงื่อนไขของการหยุดฝึกฝนมีดังนี้

เงื่อนไขการหยุดฝึกฝน กำหนดเป็นสองกรณี คือ

<u>กรณีที่ 1</u> เมื่อผลรวมของค่าความผิดพลาดเฉลี่ยระหว่างเอาต์พุตกับเป้าหมายทั้งหมด มีค่าลดลงน้อยกว่าค่าที่กำหนดไว้ โดยสมการเพื่อหาค่าเฉลี่ยของเอาต์พุตทั้งหมดมีดังนี้

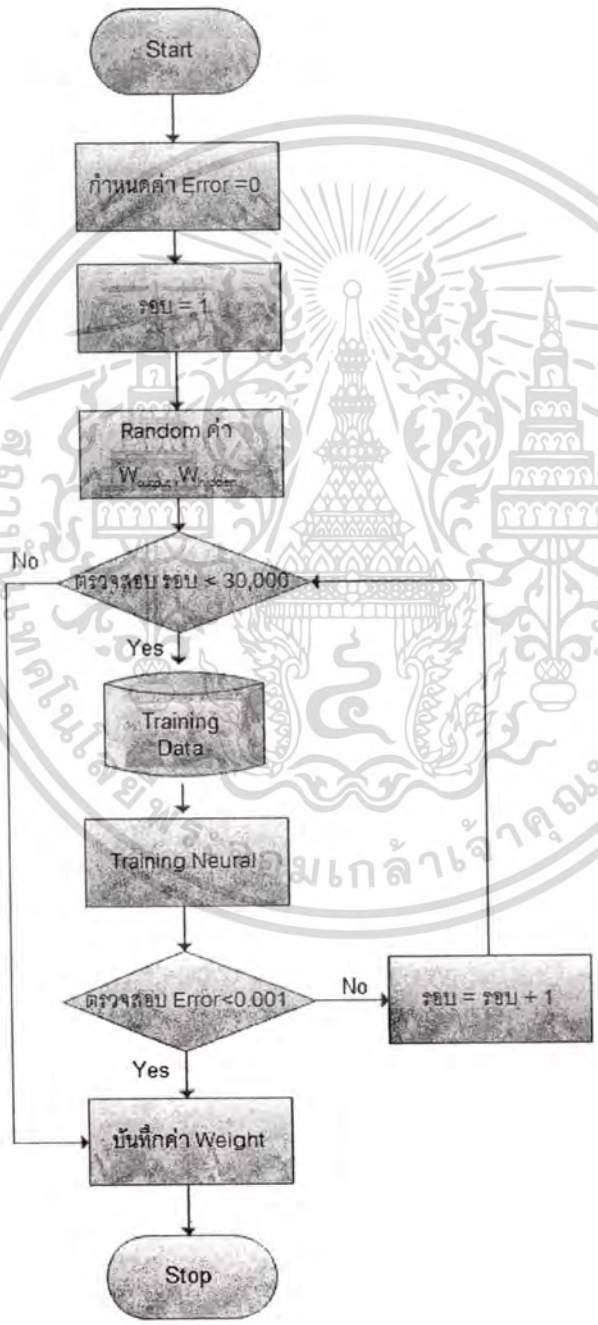
$$MSQERR = \frac{1}{n}\sum(Target - Output)^2 \qquad (3.1)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

MSQERR คือค่า Mean Square Error ที่ได้จากเอาต์พุตของโครงข่ายประสาทเทียมเทียบกับค่าเป้าหมายที่ใช้สำหรับบ่งชี้ผลของการฝึกฝนว่าเพียงพอกับการใช้งานหรือไม่

กรณีที่ 2 จำนวนรอบของการฝึกฝนดำเนินไปจนถึงค่าที่ตั้งไว้ ซึ่งจำนวนที่เหมาะสมได้จากการทดลอง

โดยขั้นตอนในกระบวนการเครือข่ายประสาทเทียม เป็นดังรูปที่ 3.8



รูปที่ 3.8 กระบวนการในเครือข่ายประสาทเทียม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยจากรูปที่ 3.8 สามารถอธิบายขั้นตอนได้ดังนี้

1. ทำการกำหนดค่าตัวแปรเบื้องต้น คือ ค่าความผิดพลาด (Error) = 0 , ค่าจำนวนรอบเริ่มต้น = 1 และ ทำการสุ่มค่าน้ำหนักระหว่างชั้นอินพุต-ชั้นซ่อนตัว และ ระหว่างชั้นซ้อนตัว-ชั้นเอาต์พุท
2. วนลูปข้อมูลอินพุตที่รับเข้ามา โดยเงื่อนไขการหยุดกำหนดไว้สองแบบ คือ ที่ จำนวนรอบมากกว่า 30,000 รอบ และ ค่าความผิดพลาด (Error) น้อยกว่า 0.001
3. ถ้ากรณีที่ยังไม่เข้าเงื่อนไขการหยุดให้ทำการเทรนนิ่งข้อมูลในเครือข่ายประสาทเทียม แต่กรณีที่เข้าเงื่อนไขการหยุดให้ทำการ บันทึกค่าน้ำหนักไว้

หลังจากสิ้นสุดกระบวนการฝึกสอนแล้วจะได้ค่าน้ำหนักที่เหมาะสมเพื่อนำไปใช้ในกระบวนการรู้จำต่อไป

3.2.4 กระบวนการตัดสินใจ

ในขั้นตอนนี้จะเป็นขั้นตอนสุดท้ายหลังจากเสร็จสิ้นกระบวนการฝึกสอนด้วยเครือข่ายประสาทเทียมจนได้ค่าน้ำหนักที่เหมาะสมซึ่งเป็นค่าน้ำหนักที่นำมาใช้ในกระบวนการรู้จำเพื่อแยกแยะสัญญาณเสียงที่เข้ามาว่าเป็นคำใด โดยแผนภาพกระบวนการตัดสินใจดังรูปที่ 3.9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนในกระบวนการตัดสินใจเพื่อควบคุมการเปิดปิดอุปกรณ์ไฟฟ้ามีดังนี้

1. เริ่มจากรับสัญญาณเสียงมาจากผู้พูดหรือผู้ออกคำสั่ง โดยเสียงที่รับเข้ามาจะประกอบด้วยคำที่เป็นคำเป้าหมายของเราที่เป็นคำหลักมีสองคำคือ เปิดและปิด ส่วนคำที่นำมาประกอบเข้ากับคำหลัก มีอีกแปดคำ

เช่น ไฟ แอร์ ม่าน เป็นต้น โดยเมื่อประกอบเข้าสองคำ จะเป็น เปิดไฟ ปิดไฟ เปิดแอร์ ปิดแอร์ รวมเป็นสองพยางค์

เปิด ปิด	ไฟ แอร์ เพลง ปั๊ม ปลั๊ก รั้ว เตา ม่าน

2. ทำการตัดเสียงของคำด้วยวิธีหาจุดสิ้นสุดของคำโดยใช้พลังงานสัมบูรณ์
3. ตรวจสอบเสียงที่ตัดมาว่าประกอบด้วยคำเท่ากับสองคำหรือไม่ ถ้าสองคำทำขั้นตอนต่อไป แต่ถ้าไม่ใช่ให้จบโปรแกรม
4. ทำการปรับบรรทัดฐานทางเวลาของคำให้เป็นมาตรฐาน
5. นำคำที่ผ่านการปรับบรรทัดฐานทางเวลา มาทำการหาค่าสัมประสิทธิ์การประมาณพันธะเชิงเส้น
6. ค่าสัมประสิทธิ์ที่ได้จะมาเป็นค่าตัวแปรอินพุตให้กับโครงข่ายประสาทเทียมแบบ Feed Forward
7. ตรวจสอบเอาต์พุตจากโครงข่ายประสาทเทียมทั้งสองคำ ว่าคำแรกต้องเป็นคำว่าเปิดหรือปิด และอีกคำต้องเป็นคำประกอบ (ไฟ แอร์ เพลง ปั๊ม ปลั๊ก รั้ว เตา ม่าน) ถ้าไม่ประกอบด้วยคำหลักและคำที่นำมาประกอบให้จบโปรแกรม
8. ส่งค่าเอาต์พุตที่ได้ ไปยัง ไมโครคอนโทรลเลอร์ เพื่อให้ไมโครคอนโทรลเลอร์ควบคุมอุปกรณ์ไฟฟ้าต่อไป

3.2.5 การออกแบบการทดสอบของการรู้จำด้านเสียง

การทดสอบของกระบวนการรู้จำผ่านเสียง จะมีการออกแบบการทดลองสองแบบ ดังนี้

1. การทดสอบโดยใช้เสียงที่มีสัญญาณเสียงเหมือนกับคำที่อยู่ในระบบ (คำที่อยู่ในระบบคือ คำว่า เปิด ปิด ไฟ แอร์ ปลั๊ก ปั๊ม เตา ม่าน รั้ว เพลง)
2. การทดสอบโดยใช้เสียงที่มีสัญญาณเสียงไม่เหมือนคำที่มีในระบบ เช่น นำคำที่ออกเสียงคล้ายกัน ยกตัวอย่างเช่น ไป ผ่าน เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

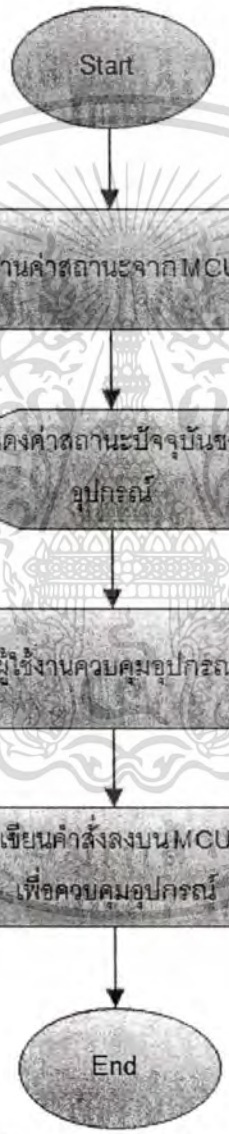
3.3 การออกแบบระบบของการควบคุมอุปกรณ์ด้วยมือถือ

3.3.1 เครื่องมือที่ใช้

การควบคุมอุปกรณ์ภายในบ้านด้วยมือถือใช้ลักษณะเว็บแอพพลิเคชั่น

- ภาษาที่ใช้คือ ภาษา PHP ,HTML เป็นหลัก
- เครื่องมือที่ใช้จำลองเป็นเว็บเซิร์ฟเวอร์ใช้ APACHE WEB SERVER

3.3.2 ผังการทำงานหลักของระบบ



รูปที่ 3.10 กระบวนการทำงานการควบคุมอุปกรณ์ผ่านมือถือ

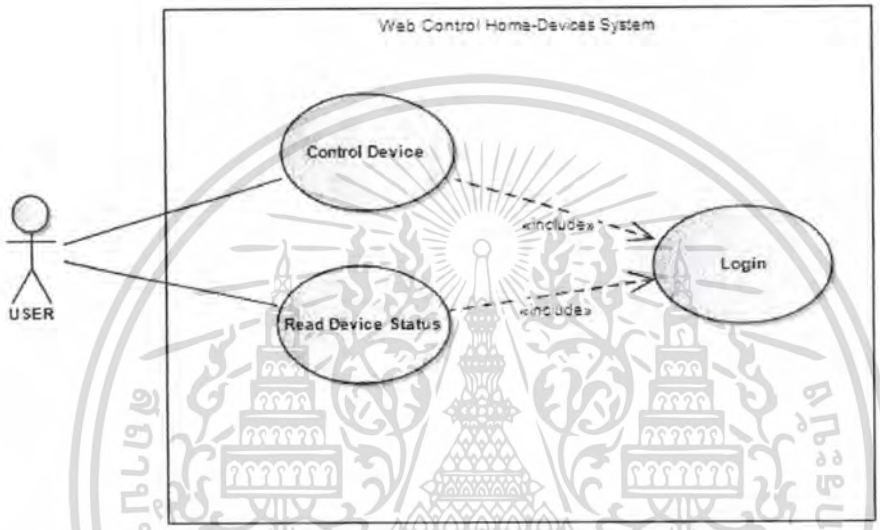
จากรูปที่ 3.10 เป็นรูปกระบวนการการทำงานหลักของการควบคุมอุปกรณ์ภายในบ้านผ่านมือถือ โดยขั้นแรกเมื่อเริ่มต้นหน้าเว็บที่ใช้ในการควบคุมอุปกรณ์ขึ้นมา เว็บจะทำการไปอ่านค่าสถานะล่าสุดของอุปกรณ์ที่ถูกเขียนไว้ใน EEPROM บนไมโครคอนโทรลเลอร์ขึ้นมา เพื่อแสดงให้เห็นว่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สถานะปัจจุบันของอุปกรณ์มีสถานะเปิดหรือปิดอุปกรณ์ เมื่อเสร็จสิ้นกระบวนการอ่านค่าสถานะ ผู้ใช้งานสามารถที่จะควบคุมอุปกรณ์แต่ละชนิดตามที่ต้องการให้เปิดหรือปิด โดยค่าที่ได้จากการควบคุมจะถูกเขียนลงบนไมโครคอนโทรลเลอร์เพื่อให้ไมโครคอนโทรลเลอร์ทำการนำค่าที่ได้ไปควบคุมอุปกรณ์ภายในบ้านที่ต่อไว้ต่อไป

3.3.3 ยูสเคสไดอะแกรม (Use case diagram)

จากขั้นตอนการกระบวนการทำงานในการควบคุมอุปกรณ์ภายในบ้านผ่านมือถือสามารถสร้างเป็นยูสเคสไดอะแกรมได้ดังนี้



รูปที่ 3.11 ยูสเคสไดอะแกรมของระบบการควบคุมอุปกรณ์ผ่านมือถือ

จากยูสเคสไดอะแกรมสามารถอธิบายได้ดังนี้

ตารางที่ 3.1 คำอธิบายยูสเคสไดอะแกรมของ Control Device

ยูสเคส	Control Device
วัตถุประสงค์	เพื่อให้ผู้ใช้งานควบคุมอุปกรณ์ภายในบ้าน
เมื่อทำงานเสร็จ	อุปกรณ์ภายในบ้านจะมีสถานะตามที่ผู้ใช้งานควบคุม
Actor	User (ผู้ใช้งาน)
อินพุต	ข้อมูลของผู้ใช้งานในการควบคุมอุปกรณ์
เอาท์พุต	ค่าคำสั่งที่จะส่งให้ไมโครคอนโทรลเลอร์
รายละเอียด	1.ผู้ใช้งานทำการควบคุมอุปกรณ์ภายในบ้านโดยมีสองสถานะการควบคุมคือ เปิดและปิด 2.ระบบจะทำการแปลงค่าข้อมูลจากผู้ใช้งานเข้าสู่ไมโครคอนโทรลเลอร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.2 คำอธิบายยูสเคสไดอะแกรมของ Read Device Status

ยูสเคส	Read Device Status
วัตถุประสงค์	เพื่อให้ผู้ใช้งานทราบถึงสถานะปัจจุบันของอุปกรณ์ภายในบ้าน
เมื่อทำงานเสร็จ	ระบบจะแสดงผลสถานะปัจจุบันของอุปกรณ์ภายในบ้านให้แก่ผู้ใช้งาน
Actor	User (ผู้ใช้งาน)
อินพุต	ค่าคำสั่งเพื่ออ่านข้อมูลสถานะ
เอาท์พุต	ค่าคำสั่งที่จะส่งให้ไมโครคอนโทรลเลอร์
รายละเอียด	1.ระบบจะทำการสถานะปัจจุบันของอุปกรณ์ภายในบ้านโดยจะอ่านค่าจาก EEPROM ที่อยู่ในไมโครคอนโทรลเลอร์

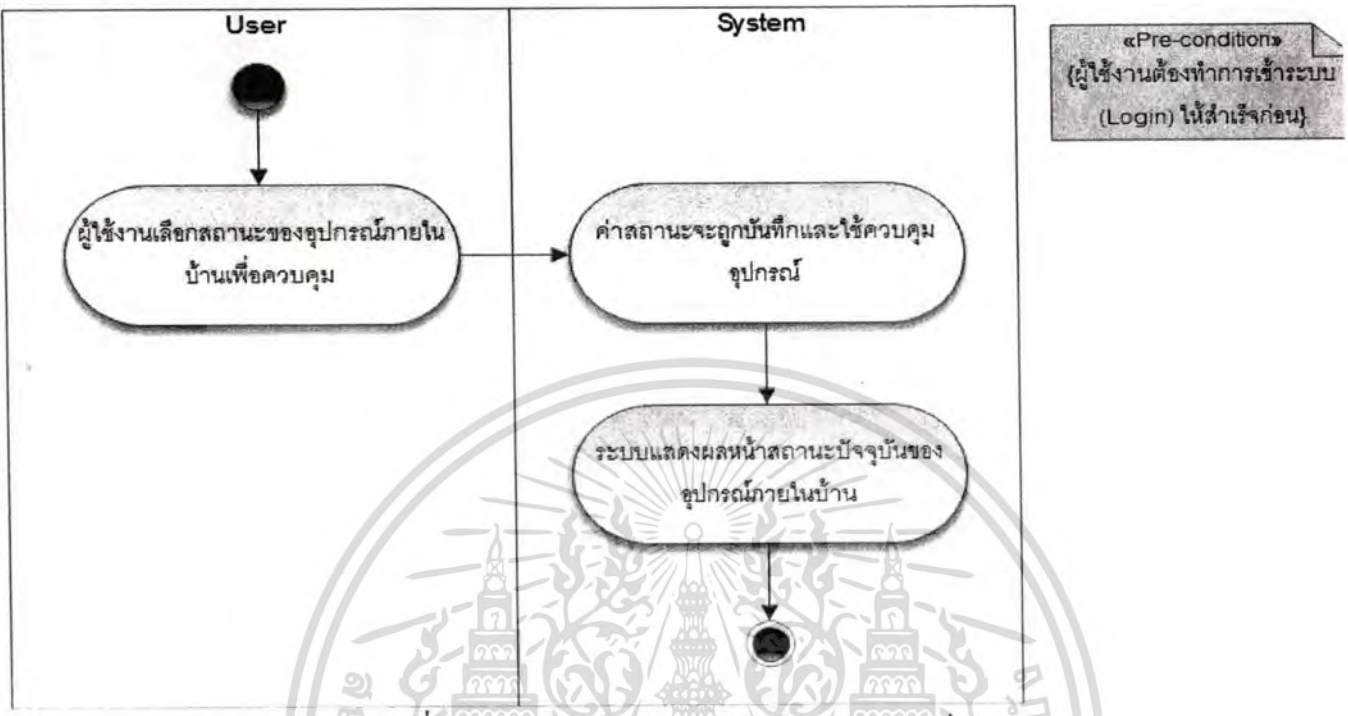
ตารางที่ 3.3 คำอธิบายยูสเคสไดอะแกรมของ Login

ยูสเคส	Login
วัตถุประสงค์	เพื่อให้ผู้ใช้งานทำการยืนยันตัวตนก่อนที่จะใช้งานระบบ
เมื่อทำงานเสร็จ	ระบบจะทำการแสดงผลหน้าที่ใช้ควบคุมอุปกรณ์ภายในบ้าน
Actor	User (ผู้ใช้งาน)
อินพุต	ข้อมูล Username,Password
เอาท์พุต	ระบบแสดงผลความถูกต้องของการยืนยันตัวตน
รายละเอียด	1.ผู้ใช้งานต้องทำการยืนยันตัวตนให้สำเร็จก่อนใช้งานควบคุมอุปกรณ์ภายในบ้านทุกครั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

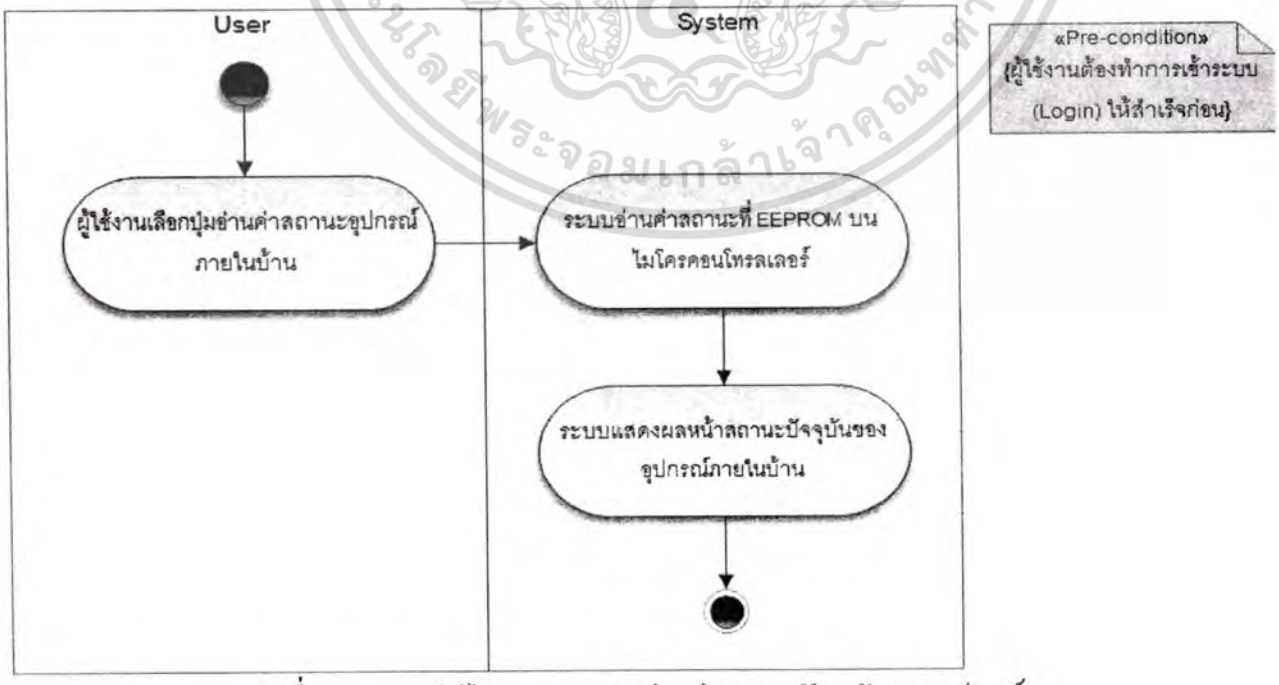
3.3.4 แอคทิวิตีไดอะแกรม (Activity diagram)

3.3.4.1 แอคทิวิตีไดอะแกรมการ Control Device



รูปที่3.12 แอคทิวิตีไดอะแกรมการควบคุมอุปกรณ์

3.3.4.2 แอคทิวิตีไดอะแกรมการ Read Device Status

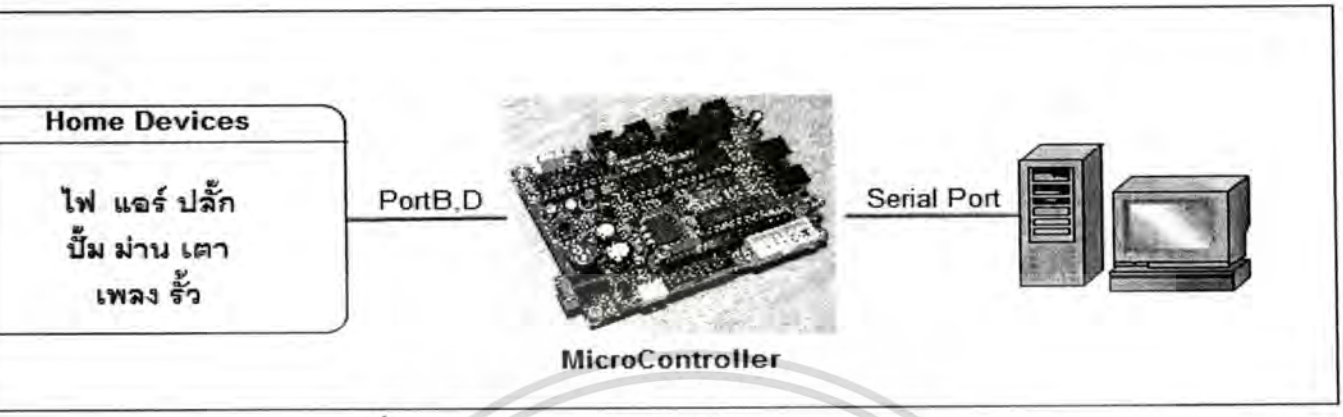


รูปที่3.13 แอคทิวิตีไดอะแกรมการอ่านค่าสถานะปัจจุบันของอุปกรณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4 การออกแบบส่วนของฮาร์ดแวร์

3.4.1 รายละเอียดของอุปกรณ์ที่ใช้และการเชื่อมต่อ



รูปที่ 3.14 แผนภาพการต่ออุปกรณ์ทางด้านกายภาพ

1. ไมโครคอนโทรเลอร์ (Microcontroller) : ที่ใช้งานในโครงงานคือ ET-BASE AVR EASY328 โดยมีคุณสมบัติดังนี้

- ใช้ MCU ตระกูล AVR เบอร์ ATMEGA328
- 32 KBYTE FLASH SRAM 2 KBYTE, EEPROM 1 KBYTE, RUN ความถี่ 19.6608 MHz
- มี PORT I/O ขนาด 20 BIT จำนวน 3 PORT (PB 6 BIT), (PC 6 BIT), (PD 8 BIT) โดยเป็น RS232, SPI, I2C, TIMER
- RS232 PORT แบบ 4 PIN ET
- POWER SUPPLY 7 - 10 VDC, ใช้ LM2940 (LOW DROP) ON BOARD

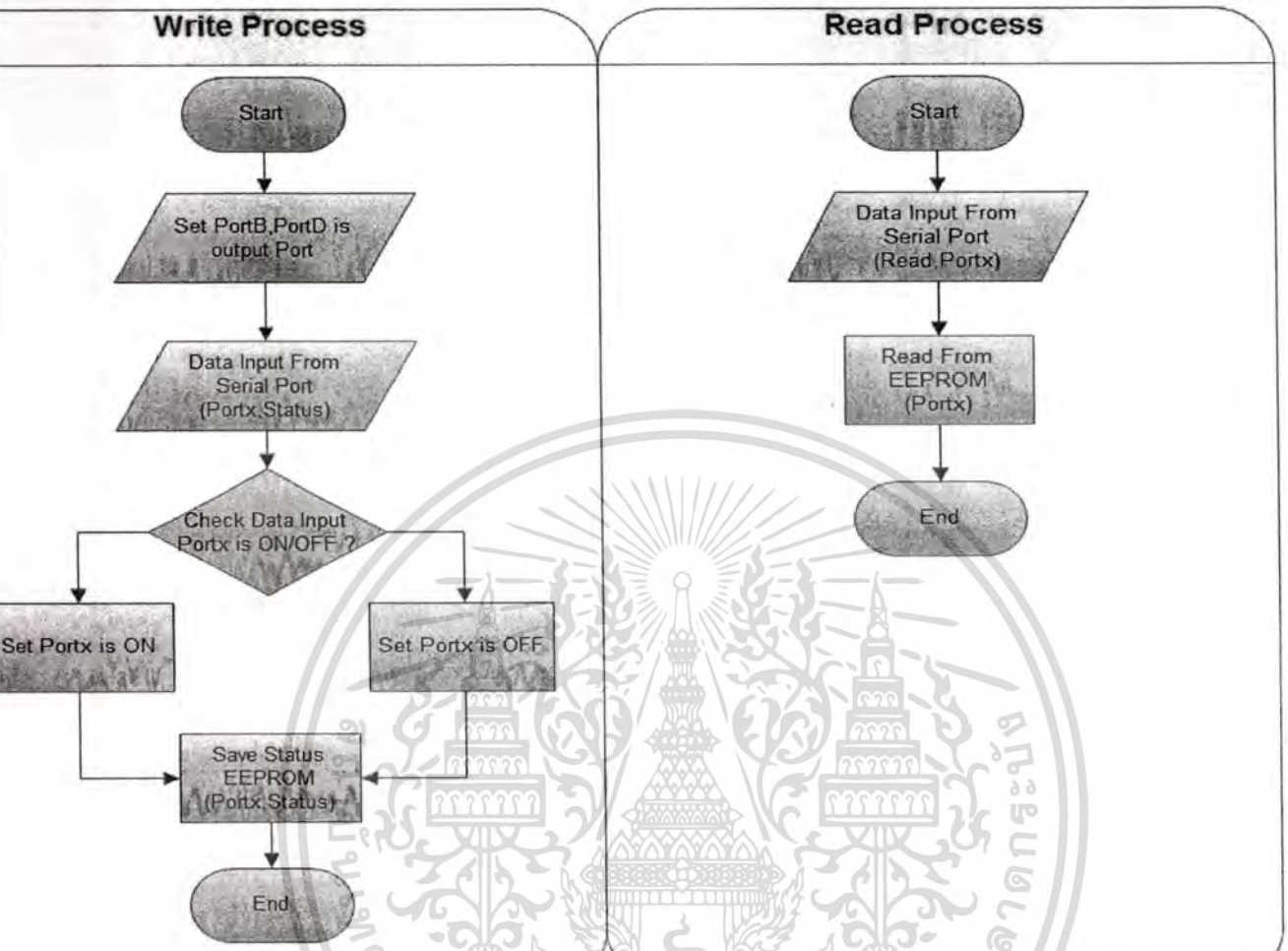
2. การเชื่อมต่อจากรูปที่ 3.14 เป็นการเชื่อมต่อทางด้านกายภาพ โดยจะมีการเชื่อมต่ออธิบายได้ดังนี้

- เครื่องคอมพิวเตอร์ที่ใช้ในการประมวลผลจะเชื่อมต่อไมโครคอนโทรลเลอร์ผ่านทาง Serial Port เพื่อที่จะทำการส่งและรับข้อมูลระหว่างคอมพิวเตอร์กับไมโครคอนโทรลเลอร์
- ไมโครคอนโทรลเลอร์จะเชื่อมต่อกับอุปกรณ์ภายในบ้านผ่านทาง Port I/O ซึ่งจากการออกแบบเลือกใช้ Port B,D ของไมโครคอนโทรลเลอร์ เพื่อส่งค่าสถานะที่ได้รับจากผู้ใช้งานมาควบคุมการเปิด ปิดอุปกรณ์

3. การทำงานจะทำโดยผู้ใช้งานจะทำการส่งค่าให้คอมพิวเตอร์ประมวลผล เครื่องคอมพิวเตอร์ที่ใช้ในการประมวลผลจะส่งค่าที่ประมวลผลและเปลี่ยนอยู่ในรูปแบบที่ถูกต้องผ่านทาง Serial Port ไมโครคอนโทรลเลอร์จะรับค่าไปประมวลผลเพื่อสั่งการในการควบคุมอุปกรณ์ภายในบ้านหรืออ่านสถานะของอุปกรณ์ภายในบ้านต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4.2ผังการทำงานของไมโครคอนโทรลเลอร์



รูปที่3.15 กระบวนการการทำงานของไมโครคอนโทรลเลอร์

จากรูปที่3.14 จะมีกระบวนการหลักๆภายในไมโครคอนโทรลเลอร์อยู่สองกระบวนการด้วยกัน คือ

1. **กระบวนการเขียนข้อมูล** เป็นกระบวนการควบคุมอุปกรณ์ที่ต่อเข้ากับพอร์ตหมายเลขต่างๆของไมโครคอนโทรลเลอร์ให้มีสถานะตามที่ผู้ใช้งานได้ทำการกำหนด โดยกระบวนการภายในเป็นดังนี้

   ขั้นตอนที่หนึ่ง ระบบจะทำการกำหนดพอร์ตบนไมโครคอนโทรลเลอร์(PortD,PortB) ให้เป็นเอาท์พุตพอร์ตเพื่อที่จะใช้ในการส่งค่าควบคุมอุปกรณ์ที่ต่อกับพอร์ตนั้นๆ

   ขั้นตอนที่สอง ไมโครคอนโทรลเลอร์จะทำการรับค่าอินพุตจาก Serial port ที่เชื่อมต่อกับเครื่องประมวลผล โดยค่าที่รับเข้ามาจากผู้ใช้งานในกระบวนการเขียนข้อมูล จะส่งค่าในรูปแบบ หมายเลขพอร์ตซึ่งแทนอุปกรณ์ภายในบ้านและค่าสถานะว่าจะให้ทำการเปิดหรือปิด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่สาม ไมโครคอนโทรลเลอร์จะทำการตรวจสอบค่าอินพุตที่รับเข้ามาและนำค่านั้นเพื่อควบคุมอุปกรณ์ภายในบ้านที่ต่อกับพอร์ตนั้นๆ

ขั้นตอนที่สี่ ไมโครคอนโทรลเลอร์จะทำการเก็บค่าสถานะไว้ใน EEPROM ของไมโครคอนโทรลเลอร์ เพื่อที่จะใช้ในการจดจำค่าสถานะปัจจุบันไว้

2. **กระบวนการอ่านข้อมูล** กระบวนการนี้จะเป็นการอ่านสถานะของอุปกรณ์ที่ถูกเก็บไว้บน EEPROM โดยมีขั้นตอนย่อยๆภายในดังนี้

ขั้นตอนที่หนึ่ง ไมโครคอนโทรลเลอร์จะทำการอ่านค่าข้อมูลที่เข้ามาผ่าน Serial port เช่นเดียวกัน โดยรูปแบบที่ส่งเข้ามาจะเป็น R ซึ่งแทน read และ หมายเลขพอร์ตที่ต้องการอ่านค่า

ขั้นตอนที่สอง ไมโครคอนโทรลเลอร์จะนำอินพุตที่รับมาทำการอ่านค่าที่เก็บอยู่ใน EEPROM แล้วส่งสถานะกลับไปทาง Serial Port ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
การทดลองและวิเคราะห์ผลการทดลอง

ในการทดลอง ผู้วิจัยได้ทำการเก็บข้อมูลที่ใช้ในการทดสอบและพัฒนาปรับกฎเกณฑ์การตัดสินใจที่ใช้ในโครงข่ายประสาทเทียมเป็นเครื่องตรวจสอบ โดยภายในโครงข่ายทางผู้วิจัยได้ทำการปรับปรุงพารามิเตอร์ที่เหมาะสม โดยรายละเอียดส่วนต่างๆมีดังนี้

4.1 ข้อมูลที่ใช้ทดสอบ

ในกระบวนการฝึกสอนและกระบวนการทดสอบการรู้จำเสียงพูด ผู้วิจัยได้ทำการจัดกลุ่มเสียงตัวอย่างเป็นสองชุดดังนี้

ชุด A : เป็นชุดเสียงพูดของคำว่า **เปิด ปิด ไฟ แอร์ ปั๊ม ปลั๊ก เพลง รั้ว ม่าน เตา** จำนวน 70 คน

ชุด B : เป็นชุดเสียงพูดของคำว่า **เปิด ปิด ไฟ แอร์ ปั๊ม ปลั๊ก เพลง รั้ว ม่าน เตา** จำนวน 30 คน

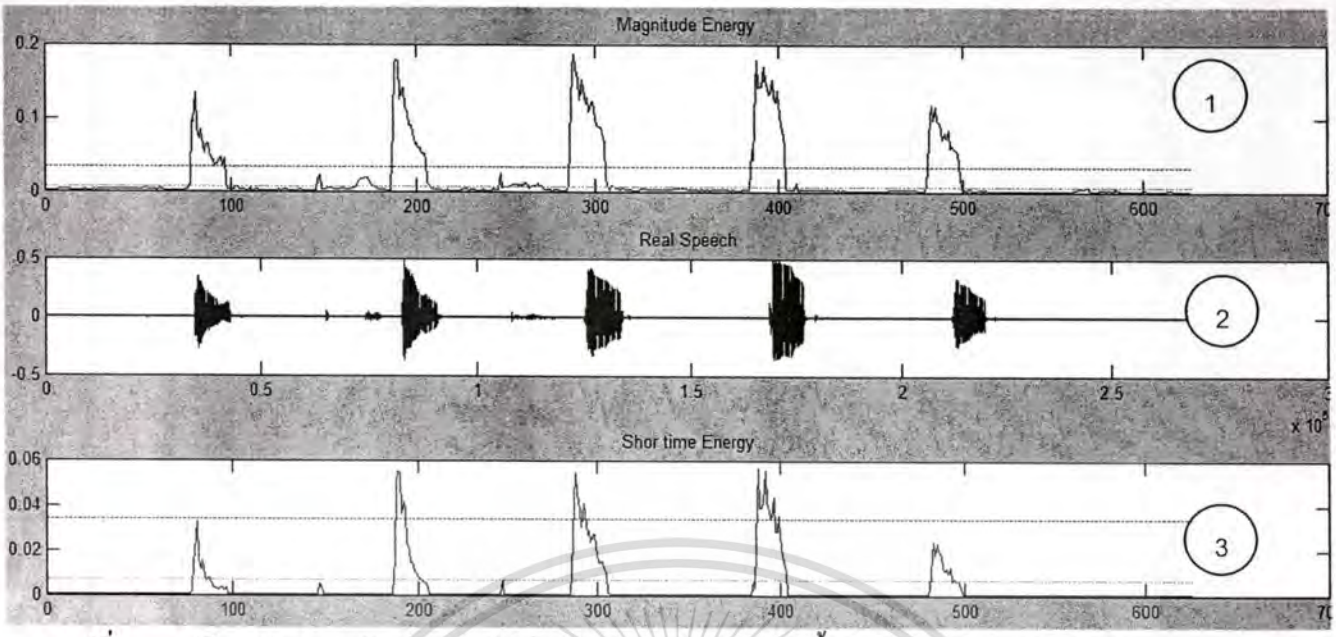
ชุด A จะนำมาใช้เพื่อเป็นต้นแบบของคำในการหาค่าน้ำหนักที่เหมาะสม โดยส่วนนี้จะนำเข้าสู่โครงข่ายประสาทเทียม เพื่อทำการเรียนรู้และจดจำลักษณะเด่นของคำแต่ละคำที่นำมาทดสอบ

ชุด B จะนำมาใช้เพื่อทำทดสอบการรู้จำแบบไม่ขึ้นกับผู้พูด หลังจากที่โครงข่ายประสาทเทียมได้ทำการเรียนรู้ลักษณะเด่นของคำแต่ละคำจากชุด A เรียบร้อยแล้ว

4.2 การหาจุดสิ้นสุดของคำและการตัดคำ

จากรูปที่ 4.1 เป็นการหาจุดสิ้นสุดของคำและตัดคำด้วยการหาค่าพลังงาน โดยจะเป็นการเปรียบเทียบวิธีการหาสองวิธี วิธีแรกหาด้วยค่าพลังงานในช่วงเวลาสั้นๆ และ วิธีที่สองหาด้วยค่าพลังงานสัมบูรณ์ จากรูป 4.1(3) จะเห็นได้ว่า การหาจุดสิ้นสุดของคำด้วยค่าพลังงานในช่วงเวลาสั้นๆ(Short time Energy) การหาด้วยวิธีนี้จะไวต่อค่าแอมปลิจูดที่มีขนาดใหญ่ ทำให้เมื่อมีสัญญาณเสียงมีแอมปลิจูดแตกต่างกันมากระหว่างคำที่มาประมวลผล ทำให้ค่าของพลังงานที่ได้แตกต่างกันมาก และยากต่อการหาจุดเริ่มต้นและสิ้นสุดของคำเพื่อตัดเป็นคำออกมา แต่ถ้าเป็นวิธีการหาด้วยค่าพลังงานสัมบูรณ์รูปที่ 4.1(1) ค่าที่ได้ในแอมปลิจูดที่แตกต่างกันของสัญญาณเสียงเมื่อแปลงเป็นค่าพลังงานจะไม่ส่งผลแตกต่างกันมาก ทำให้การหาจุดเริ่มต้นและจุดสิ้นสุดของคำทำได้ง่ายกว่าและถูกต้องมากกว่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.1 เปรียบเทียบการตัดเสียงด้วยวิธีหาค่าพลังงานในช่วงเวลาสั้นๆ 4.1(3) และ วิธีค่าพลังงานสัมบูรณ์4.1(1) ส่วน 4.1(2) จะเป็นสัญญาณเสียงที่รับเข้ามาโดยเป็นสัญญาณเสียงในโดเมนเชิงเวลา

จากการทดสอบการหาจุดสิ้นสุดของคำด้วยวิธีหาค่าพลังงานสัมบูรณ์ ได้ผลการทดลองดังนี้

ตารางที่ 4.1 ผลสรุปการทดลองในการหาจุดสิ้นสุดของคำ

คำ	ความยาวต่ำสุด (sample)	ความยาวสูงสุด (sample)	เปอร์เซนต์ความถูกต้องในการตัดคำ(%)
เปิด	7,939	18,082	81.6
ปิด	4,852	88,21	83.2
ไฟ	12,349	25,579	70
แอร์	14,113	25,579	74.4
ปั๊ม	8,821	22,492	79.2
ปลั๊ก	4,852	11,026	86.4
เพลง	10,585	25,579	82
รั๊ว	11,467	25,579	71.6
ม่าน	10,585	24,256	74
เตา	10,144	24,256	79

จากผลการทดลอง ในการหาจุดสิ้นสุดของคำโดยวิธีหาค่าพลังงานสัมบูรณ์ เปอร์เซ็นต์ความถูกต้องเฉลี่ยเท่ากับ 78.12 % ซึ่งปัญหาที่ทำให้การหาจุดสิ้นสุดของคำมีความถูกต้องไม่ร้อย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เปอร์เซนต์ คือ เรื่องมาตรฐานการพูดของแต่ละบุคคลไม่เท่ากันบางคนพูดสั้น ยาว เสียงหนัก เบา มีเสียงรบกวน และลักษณะของคำที่มาใช้ในโครงงาน ซึ่งจะสังเกตุได้ว่าคำที่มีเปอร์เซนต์ความถูกต้องสูงจะสะกดด้วยคำตาย เช่น แม่กด แม่กก เป็นต้น

4.3 การปรับบรรทัดฐานทางเวลา

จากตารางที่ 4.1 ได้ทำการสรุปขนาดความยาวของคำ ทั้งที่เป็นขนาดของคำที่ยาวที่สุดและสั้นที่สุดของแต่ละคำไว้ ซึ่งการปรับบรรทัดฐานทางเวลาโดยการประมาณค่าในช่วงเชิงเส้น จะใช้ค่าขนาดความยาวเพื่อเป็นมาตรฐานให้กับคำอื่นๆ โดยต้องเลือกคำที่มีความยาวสูงที่สุด ซึ่งค่าขนาดความยาวสูงที่สุดคือ 25,579 sample ซึ่งจะนำมาเป็นค่ามาตรฐานที่ปรับให้คำอื่นมีความยาวเท่ากัน

4.4 การสกัดลักษณะเด่นของคำ

ในการหาลักษณะเด่นของคำจะใช้สัมประสิทธิ์การประมาณพันธะเชิงเส้น โดยในการหาสัมประสิทธิ์การประมาณพันธะเชิงเส้น ต้องทำการหาค่าอันดับที่เหมาะสมที่ทำให้สัญญาณเสียงที่ได้จากสัมประสิทธิ์การประมาณพันธะเชิงเส้นหลังจากมากระตุ้นด้วยความถี่ค่าหนึ่งมีลักษณะของสัญญาณใกล้เคียงกับสัญญาณเสียงต้นฉบับ จากการทดลองได้ทำการทดสอบค่าอันดับต่างๆได้ผลสรุปดังนี้

- ทดสอบค่าอันดับ และทำการหาความคลาดเคลื่อนระหว่าง ค่าสัญญาณเดิมกับสัญญาณที่ได้จากค่าสัมประสิทธิ์การประมาณพันธะเชิงเส้นตามค่าอันดับ ผลที่ได้ตามตารางที่ 4.2

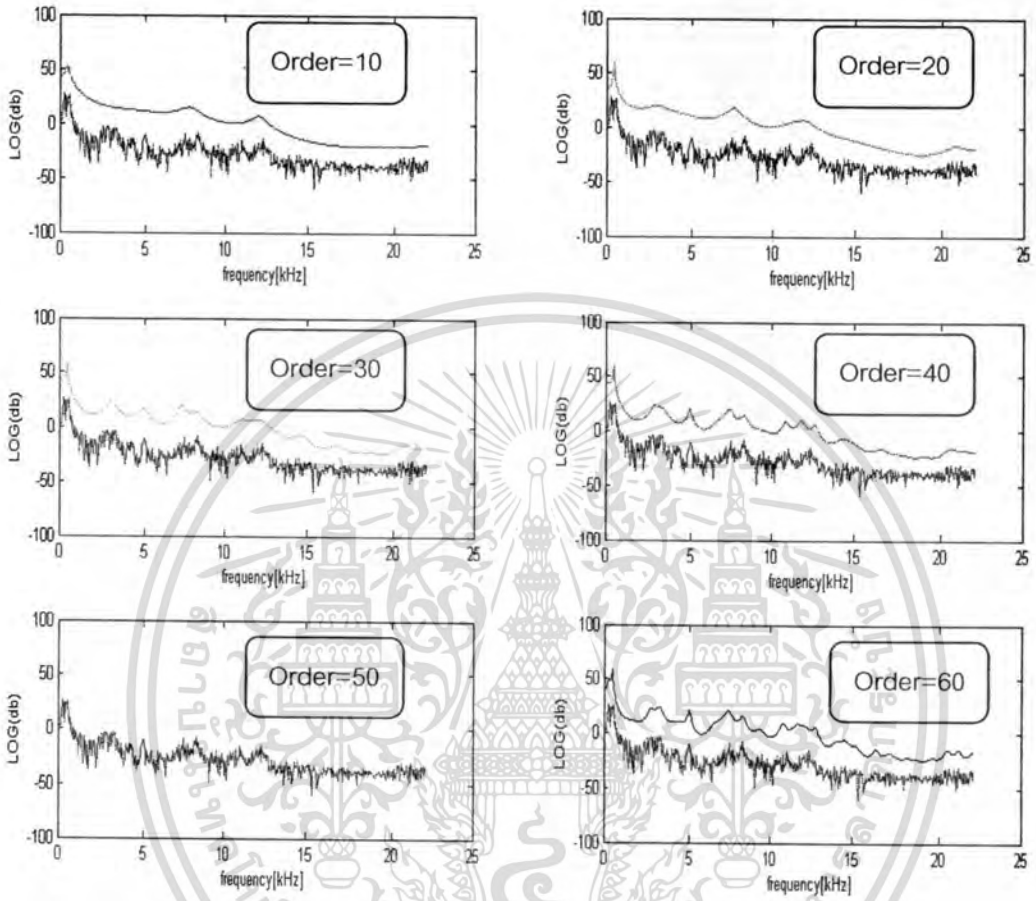
ตารางที่ 4.2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ในค่าอันดับต่างๆ

ค่าอันดับ LPC	ค่า Mean Square Error
10	0.00051396
20	0.00013632
30	0.00010952
40	0.00012637
50	0.00010421
60	0.00008089

จากตารางที่ 4.2 จะเห็นได้ว่าที่ค่าอันดับการประมาณพันธะเชิงเส้นที่ค่าอันดับตั้งแต่ 30 ขึ้นไปให้ค่าความคลาดเคลื่อกำลังสองเฉลี่ยในค่าที่ใกล้เคียงกัน และจากรูปที่ 4.2 เป็นการเปรียบเทียบให้เห็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระหว่างสัญญาณเสียงต้นฉบับกับสัญญาณเสียงที่ได้จากค่าสัมประสิทธิ์การประมาณพันธะเชิงเส้นในค่าอันดับต่างกัน



รูปที่ 4.2 กราฟโดเมนความถี่ระหว่างสัญญาณเสียงเดิมกับสัญญาณเสียงที่ได้จากค่าสัมประสิทธิ์การประมาณพันธะเชิงเส้น

จากผลการทดลองเพื่อหาค่าอันดับที่เหมาะสมสำหรับใช้ในโครงงาน ได้ข้อสรุปที่จะใช้ ค่าอันดับเท่ากับ 30 ด้วยเหตุผลที่ว่า ที่ค่าอันดับที่ 30 ก็ให้ค่าความคลาดเคลื่อนเฉลี่ยใกล้เคียงกับค่าอันดับสูงๆ ซึ่งการใช้ค่าอันดับที่ต่ำลงมาจะช่วยลดเรื่องจำนวนอินพุตที่จะเข้าไปประมวลผลในโครงข่ายประสาทเทียมทำให้เวลาในการประมวลผลลดลงไปได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5 เครือข่ายประสาทเทียม

จากที่ได้มีการออกแบบรูปแบบโครงข่ายประสาทเทียมไว้ในบทที่สาม หัวข้อเรื่องโครงข่ายประสาทเทียม โดยได้ทำการทดสอบสองกลุ่ม โดยผลการทดสอบค่าพารามิเตอร์กับโครงสร้างที่ได้จากการออกแบบไว้ว่าค่าใดเหมาะสม ได้ผลสรุปดังนี้

ค่าพารามิเตอร์ที่ใช้ในการทดสอบ

จำนวนอินพุตโหนด เท่ากับ 1,140

จำนวนโหนดชั้นซ่อนตัว เท่ากับ 20,30,50,100

อัตราการเรียนรู้ เท่ากับ 0.01,0.02,0.05,0.1,0.2

จำนวนรอบที่ใช้ทดสอบ เท่ากับ 30,000-200,000 รอบ

เงื่อนไขการหยุด

1. เมื่อถึงจำนวนรอบที่ได้กำหนดไว้

เงื่อนไขการสรุปผลทดสอบแต่ละคำ

1. ตำแหน่งคำที่ต้องการทดสอบมีค่าผลเอาต์พุตมากกว่าค่าคำอื่นๆ และ มีค่ามากกว่า 0.9 (ค่าที่ได้จากการทดสอบจะอยู่ระหว่าง 0 ถึง 1) และจะทดสอบคำนอกเหนือจากที่กำหนดไว้คือคำว่า ไป, ผ่าน

ตารางที่ 4.3 ตารางผลการทดสอบเครือข่ายประสาทเทียม

ผลการทดสอบคำว่าเปิด-ปิด (โหนดเอาต์พุต 2 โหนด)

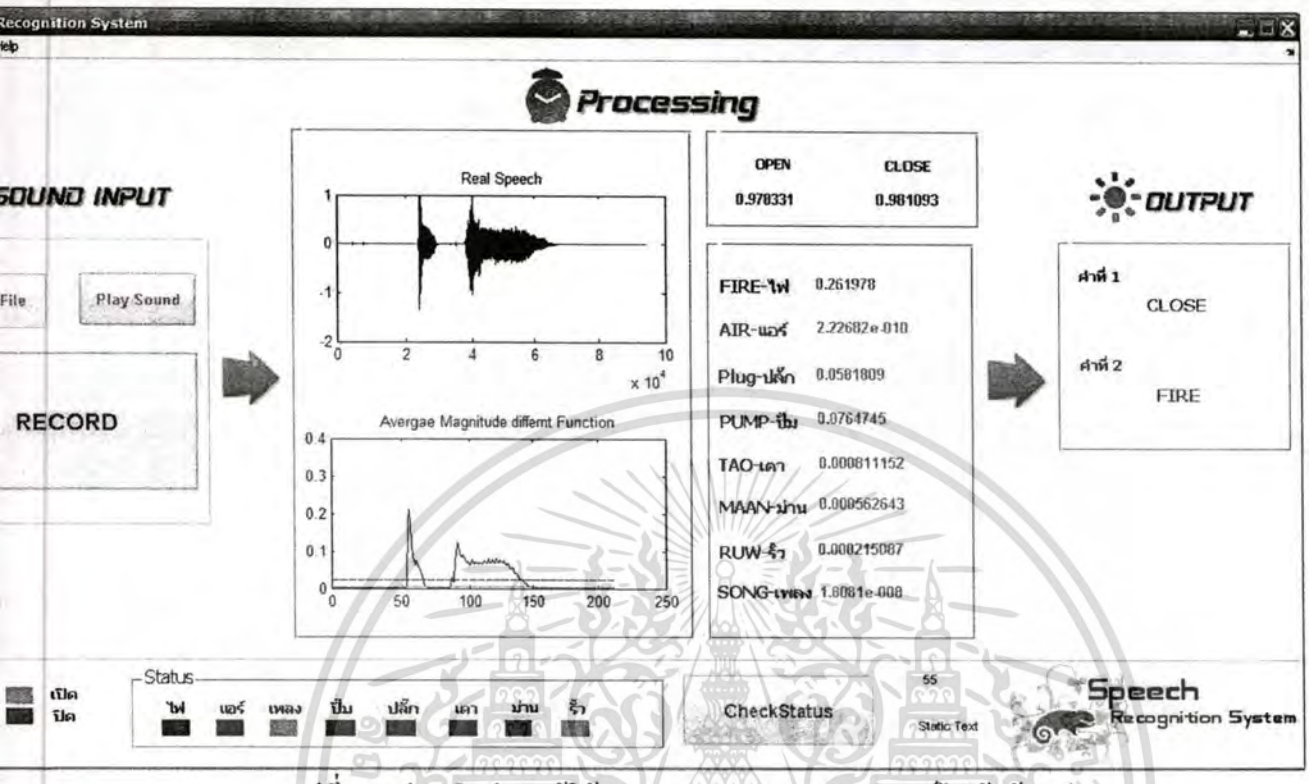
โหนดชั้นอินพุต	โหนดชั้นซ่อนตัว	จำนวนรอบ	อัตราการเรียนรู้	% ความถูกต้องคำว่า **เปิด**	% ความถูกต้องคำว่า **ปิด**	% ความถูกต้องคำว่า **ไป**
1,140	30	30,000	0.02	94.61%	92.30%	11.25%
1,140	50	35,000	0.02	93.84%	90.0%	12.89%
1,140	100	40,000	0.02	72.30%	95.38%	9.29%
1,140	**30**	**70,000**	**0.01**	**98.20%**	**97.85%**	**4.39%**

ผลการทดสอบคำว่า ไฟ แอร์ ปั๊ม ปลั๊ก เพลง รั้ว ม่าน เตา (โหนดเอาต์พุต 8 โหนด)

จำนวน			% ความถูกต้องของคำ								
ตัว	รอบ	อัตราการเรียนรู้	ไฟ	แอร์	ปั๊ม	ปลั๊ก	เพลง	รั้ว	ม่าน	เตา	ผ่าน
20	200,000	0.1	50.8	52.3	55.4	100	46.2	70	40.8	59.2	26.29
30	130,000	0.2	55.4	55.4	53.1	99.2	46.2	75.4	42.3	60	28.2
30	150000	0.05	86.2	82.3	83.1	100	78.5	86.2	77.7	85.4	14.2
20	**150000**	**0.1**	**87.7**	**85.4**	**83.1**	**100**	**79.2**	**87.7**	**77.7**	**85.4**	**11.02**
30	100000	0.1	54.6	50	54.6	100	45.4	70.8	40.8	56.2	27.22
30	100000	0.05	48.5	57.7	57.7	100	42.3	68.5	40.8	63.1	27.59

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.6 ส่วนติดต่อผู้ใช้งานของระบบการควบคุมอุปกรณ์ด้วยเสียง (Voice User-Interface)



รูปที่ 4.3 ส่วนติดต่อกับผู้ใช้งานของระบบควบคุมอุปกรณ์ไฟฟ้าด้วยเสียง

จากรูปที่ 4.3 เป็นส่วนที่ติดต่อกับผู้ใช้งานของระบบควบคุมอุปกรณ์ไฟฟ้าด้วยเสียง โดยจากส่วนติดต่อผู้ใช้งานนี้จะมีส่วนประกอบที่สำคัญอยู่ด้วยกัน 5 ส่วน คือ ส่วนของที่รับอินพุต ส่วนของกระบวนการประมวลผล ส่วนของเอาท์พุท ส่วนของตรวจสอบสถานะอุปกรณ์ ส่วนของการเรียนรู้ โดยสามารถอธิบายได้ดังนี้

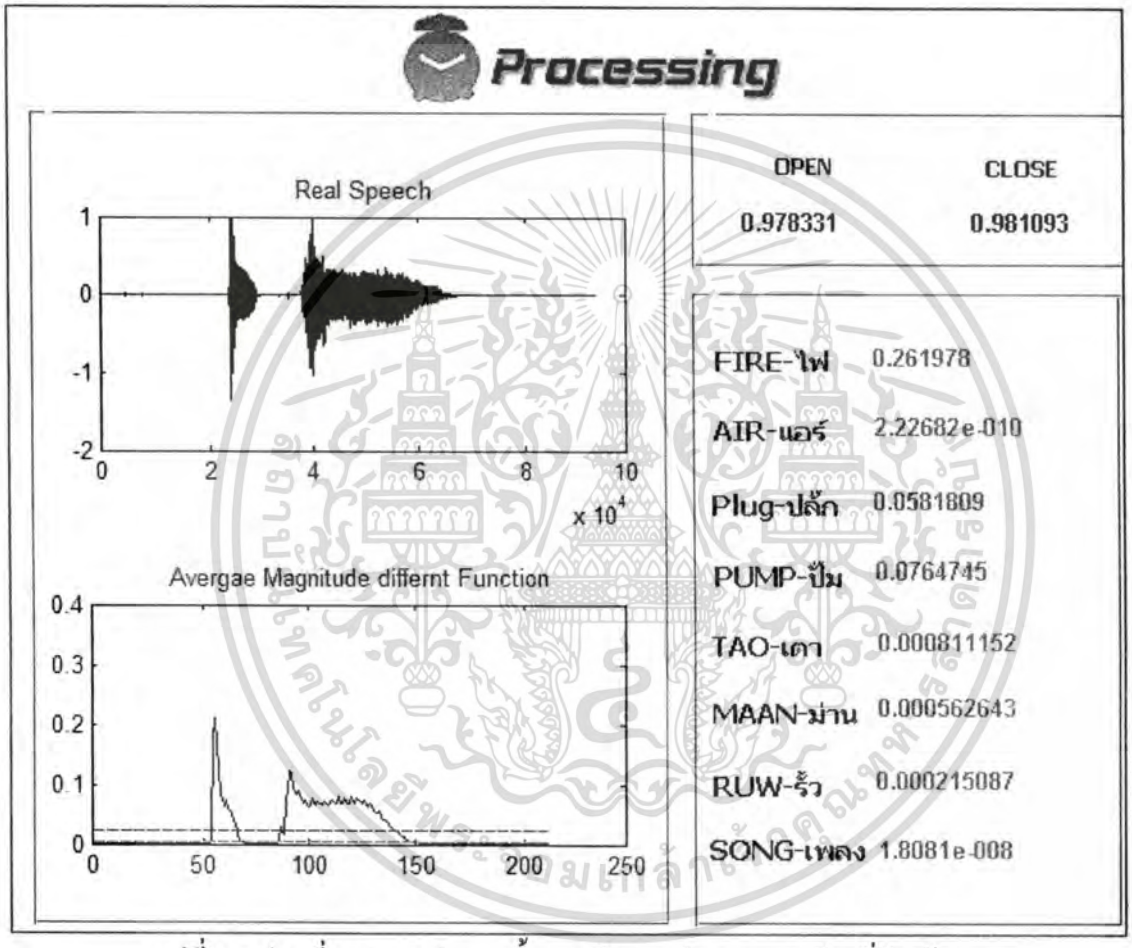
1. ส่วนที่รับอินพุต (Sound Input)

รูปที่ 4.4 ส่วนติดต่อผู้ใช้งานในการรับข้อมูลเข้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนที่รับอินพุตเข้า จะเป็นส่วนที่ผู้ใช้งานนำเสียงที่ได้จากการบันทึกเข้ามาทำการประมวลผลในระบบโดย จะสามารถนำข้อมูลเข้ามาประมวลผลได้สองวิธีด้วยกัน คือ วิธี Record หรือ การบันทึกข้อมูลจากไมโครโฟน และ วิธีนำไฟล์เสียงที่บันทึกไว้จากอุปกรณ์บันทึกเสียงอย่างอื่นมาทำการทดสอบ ซึ่งทั้งสองวิธีสามารถฟังเสียงที่ทำการนำเข้ามาด้วยปุ่มเล่นเสียง

2. ส่วนของกระบวนการประมวลผล

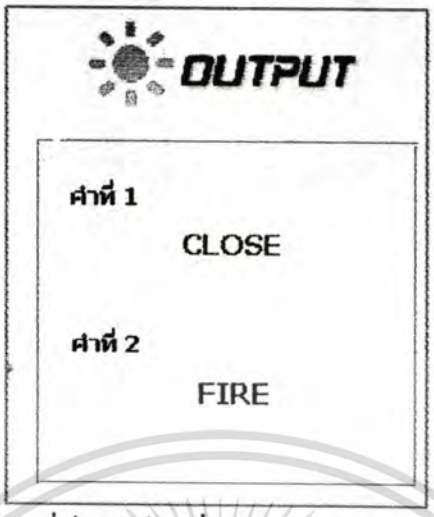


รูปที่ 4.5 ส่วนที่แสดงผลค่าและขั้นตอนในการประมวลผลเสียงที่รับเข้ามา

ในส่วนติดต่อผู้ใช้งานในกระบวนการการประมวลผลนี้จะทำการนำเสนอ กระบวนการที่ใช้และค่าที่ได้จากการประมวลผล โดยจะประกอบด้วยกราฟที่ได้จากขั้นตอนการนำเสียงเข้ามาและขั้นตอนการแปลงค่าเสียงเป็นค่าพลังงานเพื่อใช้ในการตัดแบ่งคำ ส่วนค่าตัวเลขจะมีทั้งหมด 10 ค่าด้วยกัน โดยจะเป็นค่าของคำที่ได้จากการประมวลผลด้วยวิธีการเครือข่ายประสาทเทียมซึ่งจะนำค่าของคำที่ได้จากการแปลงเสียงที่นำเข้ามาเป็นค่าสัมประสิทธิ์การประมาณพันธะเชิงเส้นเทียบกับรูปแบบของเสียงที่ได้ทำการเรียนรู้ไว้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. ส่วนของเอาท์พุท



รูปที่ 4.6 ส่วนที่เป็นเอาท์พุทที่ได้จากการประมวลผลเสียง

ในส่วนของเอาท์พุทจะเป็นส่วนที่นำข้อมูลที่ได้จากการประมวลผลมาตัดสินใจว่า เสียงที่เข้ามาซึ่งเมื่อผ่านการประมวลผลแล้วเป็นคำใด

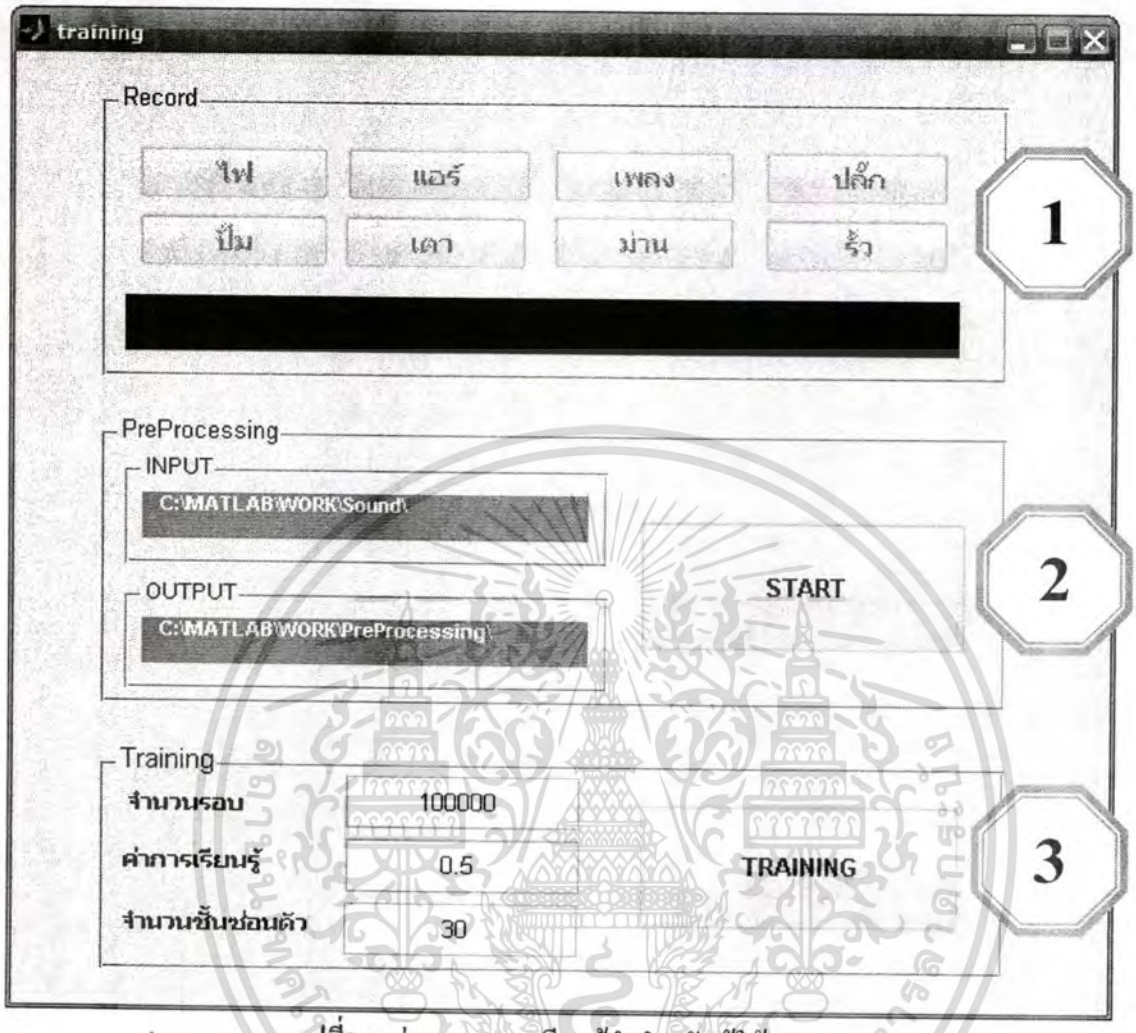
4. ส่วนของการตรวจสอบสถานะของอุปกรณ์

รูปที่ 4.7 ส่วนของการตรวจสอบและแสดงสถานะของอุปกรณ์

ในส่วนนี้จะเป็นส่วนที่ผู้ใช้งานสามารถที่จะทำการตรวจสอบสถานะของอุปกรณ์ที่กำหนดไว้ว่ามีสถานะปัจจุบันอย่างไร โดยกดปุ่ม CheckStatus

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5. **ส่วนของการเรียนรู้** (Training) **สำหรับผู้ใช้งาน**



รูปที่ 4.8 ส่วนของการเรียนรู้คำสำหรับผู้ใช้งาน

ส่วนของการเรียนรู้คำสำหรับผู้ใช้งาน เป็นส่วนที่ให้ผู้ใช้งานสามารถนำคำมาเรียนรู้เพื่อใช้ในการสั่งงานอุปกรณ์ภายในบ้านด้วยเสียง โดยในส่วนติดต่อผู้ใช้งานนี้จะมีการทำงานย่อยๆภายในด้วยกันสามประเภท

1. ส่วนของการบันทึกเสียง : เป็นส่วนที่ให้ผู้ใช้งานทำการบันทึกเสียงของคำ

2. ส่วนของการประมวลผลคำเบื้องต้น : เป็นส่วนที่ให้ผู้ใช้งานทำการนำคำที่ทำการบันทึกไว้มาประมวลสัญญาณเสียง ตามขั้นตอนที่ได้กล่าวถึงในบทที่สาม

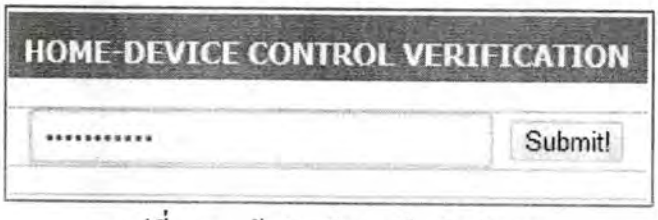
3. ส่วนของการเรียนรู้สัญญาณเสียง : เป็นส่วนที่ให้ผู้ใช้งานทำการเรียนรู้สัญญาณเสียง โดยผู้ใช้งานสามารถกำหนดจำนวนรอบ ค่าอัตราการเรียนรู้ และ จำนวนชั้นซ่อนตัวตามที่ต้องการได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.7 ส่วนติดต่อผู้ใช้งานของระบบการควบคุมอุปกรณ์ผ่านมือถือ (Web User-Interface)

ในส่วนของการติดต่อผู้ใช้งานสำหรับการควบคุมอุปกรณ์ผ่านมือถือ จะใช้เว็บเป็นตัวในการควบคุมอุปกรณ์ภายในบ้าน โดยจะประกอบด้วยส่วนติดต่อผู้ใช้งานหลักๆดังนี้

1.หน้าสำหรับการยืนยันตัวตน



รูปที่ 4.9 หน้าสำหรับการยืนยันตัวตน

จากรูปที่ 4.9 หน้าสำหรับการยืนยันตัวตน จะเป็นหน้าแรกของการเข้าถึงการควบคุมภายในบ้าน ซึ่งในหน้านี้จะมีส่วนให้ผู้ใช้งานทำการใส่รหัสผ่านที่ได้ตั้งไว้ เพื่อเป็นการยืนยันตัวตนและควบคุมการเข้าถึงหน้าสำหรับการควบคุมอุปกรณ์ภายในบ้าน

2. หน้าสำหรับควบคุมอุปกรณ์ภายในบ้าน

รูปที่ 4.10 หน้าสำหรับการควบคุมอุปกรณ์ภายในบ้าน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.10 เป็นหน้าสำหรับการควบคุมอุปกรณ์ภายในบ้าน โดยหลังจากทำการยืนยันตัวตนได้สำเร็จแล้วก็จะเข้ามาหน้าควบคุมอุปกรณ์ภายในบ้าน ซึ่งจะประกอบด้วยอุปกรณ์ภายในบ้านจำนวน 8 อย่าง (ไฟ แอร์ ปั๊ม ปลั๊ก เตา ม่าน รั้ว เพลง) โดยแต่ละอุปกรณ์ก็จะมีปุ่มเพื่อทำการเปิด(ON) และปิด(OFF) อุปกรณ์เพื่อใช้ในการสั่งการ

โดยเมื่อเข้ามาหน้าควบคุมอุปกรณ์ ตัวระบบจะทำการเริ่มต้นโดยไปอ่านค่าสถานะของอุปกรณ์แต่ละชนิดจาก EEPROM บนไมโครคอนโทรลเลอร์ซึ่งทำการบันทึกค่าไว้ มาแสดงผลบนหน้าควบคุมว่าแต่ละอุปกรณ์มีการเปิดหรือปิดไว้ หลังจากนั้นผู้ใช้งานจึงสามารถที่จะควบคุมอุปกรณ์ผ่านหน้าควบคุมว่าจะให้อุปกรณ์เปิดหรือปิดได้ตามต้องการ โดยกดปุ่ม SEND (ส่ง)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลและข้อเสนอแนะ

5.1 สรุปผลการทดลอง

การพัฒนาโปรแกรมควบคุมอุปกรณ์ไฟฟ้าภายในบ้าน มีวัตถุประสงค์เพื่ออำนวยความสะดวกในการไปเปิดปิดอุปกรณ์ไฟฟ้าในรูปแบบการควบคุมอุปกรณ์ด้วยเสียงและควบคุมผ่านระบบมือถือ โดยขั้นตอนในการพัฒนาโปรแกรมของส่วนการควบคุมอุปกรณ์ไฟฟ้าด้วยเสียง เริ่มจากการนำสัญญาณเสียงมาผ่านการประมวลผลสัญญาณเบื้องต้น การสกัดค่าลักษณะสำคัญของเสียงด้วยเทคนิคการประมาณพันธะเชิงเส้น ส่วนในขั้นตอนการทดสอบความคล้ายคลึงกันของรูปแบบและกฎเกณฑ์การตัดสินใจ ใช้เครือข่ายประสาทเทียมในการฝึกฝน เรียนรู้คำต้นแบบ และตัดสินใจเปรียบเทียบผลการออกเสียง ส่วนขั้นตอนในการควบคุมอุปกรณ์ไฟฟ้าด้วยมือถือ จะใช้เทคโนโลยีเว็บซึ่งพัฒนาด้วยภาษาพีเอชพี (PHP) ในการเปิดปิดอุปกรณ์จากระยะไกล

จากผลการทดลองพบว่าในขั้นตอนการหาจุดเริ่มต้นและสิ้นสุดของคำได้ใช้การหาค่าพลังงานสัมบูรณ์ ของสัญญาณ ซึ่งผลความถูกต้องของการตัดคำที่ได้อยู่ที่ประมาณร้อยละ 78.12 และ ผลการทดสอบหาค่าอันดับของการประมาณ โดยการหาค่าความผิดพลาดระหว่างสัญญาณเสียงจริงกับสัญญาณเสียงที่ได้จากสัมประสิทธิ์การประมาณพันธะเชิงเส้น ผลที่ได้มีค่าความผิดพลาด ตั้งแต่ค่าอันดับที่ 30 ขึ้นไปใกล้เคียงกัน ดังนั้นจึงเลือกใช้ค่าอันดับที่ 30 ซึ่งเป็นลำดับที่ให้ค่าความผิดพลาดในระดับที่ยอมรับได้และยังให้จำนวนสัมประสิทธิ์ในปริมาณที่ไม่มาก ซึ่งทำให้ลดเวลาในการประมวลผลของขั้นตอนในเครือข่ายประสาทเทียม

และจากการทดลองเครือข่ายประสาทเทียมเพื่อหารูปแบบของคำตัวอย่าง จะเห็นได้ว่าที่จำนวนรอบสูงขึ้น ค่าอัตราการเรียนรู้ที่ละเอียดขึ้น จำนวนชั้นซ่อนตัวที่เหมาะสมจะทำให้ได้ค่าความผิดพลาดลดต่ำลง ซึ่งจะส่งผลให้ค่าน้ำหนักที่ได้เมื่อนำมาทดสอบในกระบวนการตัดสินใจได้ผลความถูกต้องสูงอยู่ในเกณฑ์ที่ยอมรับได้ ซึ่งจากผลการทดสอบที่ได้กล่าวไว้ในบทที่สี่ ถึงความถูกต้องของคำ และนำมาวิเคราะห์ความเป็นไปได้ของการใช้งาน สามารถสรุปได้ว่า สามารถนำมาใช้งานได้จริงในบางกรณี เช่น ในภาวะที่เสียงเงียบ ไม่มีเสียงรบกวน สำเนียงหรือการสั่งงานะได้ถูกต้อง โดยได้วิเคราะห์ถึงสิ่งที่จะทำให้สามารถใช้งานได้ในทุกกรณี คือ การทำการเรียนรู้เฉพาะสมาชิกภายในบ้านนั้นๆเท่านั้น เพื่อช่วยเพิ่มความถูกต้องของการสั่งงาน และ ทำการสร้างส่วนของสัญญาณที่จะเข้ามาก่อนจะมีเสียงสั่งงานจากผู้ใช้งานเพื่อช่วยให้สั่งงานในภาวะที่มีเสียงรบกวนได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ปัญหาและอุปสรรค

1. การเก็บสัญญาณเสียงตัวอย่างของกลุ่มต้นแบบ บางเสียงมีลักษณะการออกเสียงของคำไม่ถูกต้อง เช่นเสียง ร ล หรือ คำควบกล้ำต่างๆ หรือ มีเสียงรบกวนเข้ามายังระบบ หรือ ออกเสียงคำสั้นบ้างยาวบ้าง ทำให้เกิดข้อผิดพลาดในการทดสอบการรู้จำ

2. ปัญหาเรื่องการประมวลผลในขั้นตอน เครือข่ายประสาทเทียม ใช้เวลาในการประมวลผลยาวนานเพื่อให้ได้ผลค่าน้ำหนักเพื่อนำมาทดสอบความถูกต้อง

3. สืบเนื่องจากข้อสองเรื่องระยะเวลาที่ใช้ในการประมวลผลที่ยาวนาน ปัญหาที่ตามมาคือเรื่องการหาเครื่องคอมพิวเตอร์ที่ใช้รองรับในการประมวลผลที่ต้องเปิดประมวลผลตลอดเวลาเพื่อให้ได้ผลค่าน้ำหนักตามต้องการ

5.3 ข้อเสนอแนะ

1. ในขั้นตอนการเก็บตัวอย่างเสียงมาใช้ในโครงงาน ต้องทำการควบคุมลักษณะสภาพแวดล้อมและการออกเสียงคำแต่ละคำของผู้พูด เพราะจะส่งผลต่อการรู้จำ

2. เพิ่มจำนวนคำที่ใช้ควบคุมอุปกรณ์ให้มากขึ้น และ เพิ่มจำนวนพยางค์ของคำสั่งแต่ละคำสั่งให้ยาวขึ้นมากกว่าสองพยางค์ เพื่อให้สามารถควบคุมอุปกรณ์ไฟฟ้าได้ครอบคลุมยิ่งขึ้น

3. ทำให้เป็นระบบตอบสนองได้ทันที คือเมื่อตรวจพบสัญญาณเสียง แล้วให้นำสัญญาณเสียงนั้นมาวิเคราะห์ประมวลผลได้ทันที

4. สร้างแอพพลิเคชั่นบนมือถือ โดยอาจจะส่งคำสั่งผ่านทาง SMS หรือ ระบบเครือข่ายขนาดเล็กภายในบ้าน ใช้ในกรณีที่ระบบอินเตอร์เน็ตใช้การไม่ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

กาญจนา ทองบุญนาค. 2544. **การรู้จำเสียงคำโดดโดยใช้โครงข่ายประสาทเทียม**. วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต สาขาวิทยาการคอมพิวเตอร์ บัณฑิตวิทยาลัย,มหาวิทยาลัยเชียงใหม่.

เยาวลักษณ์ ชาติสุขศิริเดช. 2548. **การใช้เสียงในภาษาไทย**. กรุงเทพฯ: สำนักพิมพ์อักษรเจริญทัศน์

Ian McLoughin. 2009. **Applied Speech and Audio Processing With Matlab Examples**. England : Cambridge University Press.

Joe Tebelskis.1995. **Speech Recognition using Neural Networks**. Ph.D.Thesis , Carnegie Mellon University.

Micahel Negnevitsky. 2002. **Artificial Intelligence A Guide to Intelligent Systems**. England : Addison Wesley.

Rabiner L.W. and Schafer R.W. 1978. **Digital Processing of Speech Signals**. New Jersey : Prentice-Hall Inc.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อผู้เขียน	นาย อิมรอน สมานพิทักษ์
วันเดือนปีเกิด	8 ธันวาคม พ.ศ. 2526
สถานที่เกิด	สงขลา
การศึกษา	ระดับปริญญาตรี
	วศบ.(วิศวกรรมศาสตรบัณฑิต)
	สาขาวิศวกรรมคอมพิวเตอร์
	มหาวิทยาลัยสงขลานครินทร์
ประสบการณ์ทำงาน	บริษัทแอดวานซ์อิน โฟร์เซอร์วิส (มหาชน)
	ตำแหน่งวิศวกร
	การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย
	ตำแหน่งวิศวกร

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้