ES019

การพัฒนาระบบควบคุม เปิด-ปิด ไฟฟ้าด้วยเสียงภาษาไทยโดยใช้ Arduino และ MATLAB

Development of a Thai Voice Control System using Arduino and MATLAB

นนร.รัชฎาะ หนันดูน[1], นนร.คชินทร์ บุญทรัพย์[2], นนร.ตันติกร ชูใจ[3], น.ท.พิศณุ คูมีชัย[4]
[1]สถานะผู้วิจัยหลัก (กองวิชาวิศวกรรมศาสตร์ ฝ่ายศึกษา โรงเรียนนายเรือ จ.สมุทรปราการ 10270)
[2]สถานะผู้วิจัยหลัก (กองวิชาวิศวกรรมศาสตร์ ฝ่ายศึกษา โรงเรียนนายเรือ จ.สมุทรปราการ 10270)
[3]สถานะผู้วิจัยหลัก (กองวิชาวิศวกรรมศาสตร์ ฝ่ายศึกษา โรงเรียนนายเรือ จ.สมุทรปราการ 10270)
[4]สถานะผู้วิจัยร่วม (กองวิชาวิศวกรรมศาสตร์ ฝ่ายศึกษา โรงเรียนนายเรือ จ.สมุทรปราการ 10270)

Correspondence Author : น.ท.พิศณุ คูมีชัย, 080-5246539
Email : pisanu41984198@hotmail.com

บทคัดย่อ

การพัฒนาระบบควบคุมเปิด-ปิดไฟฟ้าด้วยเสียงภาษาไทยโดยใช้ Arduino และ MATLAB มีวัตถุประสงค์เพื่อสร้างและพัฒนาระบบควบคุมการเปิด-ปิดเครื่องใช้ไฟฟ้าด้วยเสียงภาษาไทยโดยใช้ Arduino และ MATLAB โดยศึกษาการทำงานชุดควบคุมการเปิด-ปิดเครื่องใช้ไฟฟ้า และศึกษาลักษณะการทำงานของบอร์ด Arduino และ MATLAB พัฒนาระบบควบคุมการเปิด-ปิดไฟอัตโนมัติด้วยเสียงภาษาไทยด้วยเสียงโดยใช้ Arduino และ MATLAB เนื่องจากอุปกรณ์สำเร็จรูปด้านอินเทอร์เน็ตออฟติงส์ที่มีขายในปัจจุบันนั้นยังคงมีราคาที่ค่อนสูง และแอปพลิเคชันที่ใช้สำหรับการสั่งการด้วยเสียงที่มีอยู่ในตอนนี้ก็ยังไม่สามารถรองรับคำสั่งเสียงภาษาไทย ผู้พัฒนาจึงได้พัฒนาอุปกรณ์ไมโครคอนโทรลเลอร์ให้สามารถสั่งการด้วยคำสั่งเสียงภาษาไทยได้ ผลการวิจัยพบว่า จากการทดสอบการออกคำสั่งเปิด-ปิดไฟเสียงภาษาไทย 500 ครั้ง ระบบสามารถสั่งการได้สำเร็จ 385 ครั้ง หรือคิดเป็น 77 % จากการสั่งการทั้งหมดและมีจำนวน 115 ครั้งที่ไม่สามารถรับคำสั่งเสียงได้เนื่องจากมีเสียงภายนอกเข้ามารบกวนขณะการสั่งการด้วยคำสั่งเสียง

คำสำคัญ : ไมโครคอนโทรลเลอร์ , เสียงภาษาไทย , อินเทอร์เน็ตออฟติงส์

Abstract

Development of Thai Voice Electricity Control System Using Arduino and MATLAB, aims to create and develop Thai Voice Electrical Appliance Control System using Arduino and MATLAB, by studying the operation of the control set on-off electrical appliances and to study

the behavior of the Arduino and MATLAB boards, develop a control system for automatic light on-off with Thai voice by using Arduino and MATLAB. Since the Internet of Things ready-made equipment that is sold today is still quite expensive, and applications that are used for voice commands that exist at this time, it is not able to support Thai voice commands. The developer has developed a microcontroller device to be able to operate with Thai voice commands. The research results were found that from the test of issuing the command to turn on - off the Thai voice sound 500 times, the system was able to command successfully 385 times or 77% of all commands and 115 times the voice command cannot be received due to an external sound, interfering with the command with voice commands.

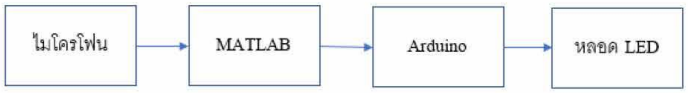
Keywords : Microcontroller, Thai language, Internet of things

1. บทนำ

ปัจจุบันอินเทอร์เน็ตได้เข้ามามีบทบาทในชีวิตประจำวันมากขึ้นผู้คนสามารถรับและส่งข้อมูลข่าวสารต่างๆ ได้อย่างมากมายจากเครือข่ายอินเทอร์เน็ตการที่จะใช้อินเทอร์เน็ตให้เกิดประโยชน์มากที่สุดในปัจจุบันสามารถประยุกต์ใช้ได้อย่างหลากหลาย อาทิเช่น แนวคิด Internet of Things หมายถึง การที่สิ่งต่างๆ ถูกเชื่อมโยงทุกสิ่งทุกอย่างสู่โลกอินเทอร์เน็ต ทำให้มนุษย์สามารถสั่งการควบคุมการใช้งานอุปกรณ์ต่าง ๆ ผ่านทางเครือข่ายอินเทอร์เน็ต เช่น การเปิด-ปิด อุปกรณ์เครื่องใช้ไฟฟ้า เครื่องมือสื่อสาร เครื่องมือทางการเกษตร อาคารบ้านเรือน เครื่องใช้ในชีวิตประจำวันต่าง ๆ ผ่านเครือข่ายอินเทอร์เน็ต เป็นต้น

บ้านอัจฉริยะ หรือ Smart home เป็นแนวคิดที่ต่อยอดมาจาก แนวคิด Internet of Things โดยจะเป็นการรวมโครงข่ายการสื่อสารของที่อยู่อาศัยรวมเข้าด้วยกันเพื่อเชื่อมต่อเครื่องใช้ไฟฟ้า การบริการ การตรวจตราดูแล รวมทั้ง สามารถเข้าถึงการควบคุมอุปกรณ์ต่าง ๆ ได้ ซึ่งการควบคุมอาจควบคุมได้ทั้งจากภายในที่บ้านเองหรือควบคุมจากภายนอกก็ได้โครงสร้างของบ้านอัจฉริยะจะต้องประกอบไปด้วย 3 ส่วนหลัก ได้แก่ ส่วนแรกต้องมีอุปกรณ์ Smart Device ใช้สำหรับเชื่อมโยงเข้ากับส่วนที่สอง คือ เครือข่ายอินเทอร์เน็ตภายในบ้าน และ ส่วนที่สาม คือ ไมโครคอนโทรลเลอร์ (Microcontroller) ส่วนควบคุมหลักที่เปรียบเสมือนสมองของบ้าน ซึ่งสามารถเขียนโปรแกรมให้อุปกรณ์ต่างๆ ในบ้านทำงาน ตามคำสั่งที่ต้องการได้เมื่อบ้านดังกล่าวมีองค์ประกอบครบทั้ง 3 ส่วนแล้วจึงจะถือได้ว่าเป็นบ้านอัจฉริยะ (ไอที 24 ชั่วโมง, 2555) ผู้วิจัยจึงเกิดแนวคิดที่จะพัฒนาระบบควบคุมเปิด-ปิดหลอดไฟที่สามารถรับคำสั่งเสียงภาษาไทยได้ และพัฒนาอุปกรณ์ไมโครคอนโทรลเลอร์ที่ควบคุมการจ่ายกระแสไฟฟ้าไปยังหลอดไฟตามคำสั่งเสียงจากระบบเพื่อความสะดวกสบายและความปลอดภัยในการใช้งาน งานวิจัยที่สอดคล้องมีดังนี้ เสฐียรพงษ์ สมบัติใหม่ และอาทร สิริกุลธมาลา (2557) ระบบควบคุมไฟฟ้าแสงสว่างผ่านเครือข่ายไร้สายคณะวิศวกรรมศาสตร์มหาวิทยาลัยเทคโนโลยีราชมงคล ลักษณะการทำงานของระบบจะทำงานโดยใช้ระบบ Wireless ในการส่งถ่ายข้อมูลในการรับคำสั่งสถานะเปิด-ปิด ร่วมกับแอปพลิเคชันของ ระบบปฏิบัติการ Android ระบบควบคุมไฟฟ้าแสงสว่างผ่าน

เครือข่ายไร้สาย สามารถควบคุมการทำงานได้ ที่ระยะ 30 เมตร ทั้งนี้ขึ้นอยู่กับขนาดของเสาสัญญาณของระบบ Wireless แต่ละขนาด และสิ่งกีดขวางใน การส่งถ่ายข้อมูล สัญชัย เอี่ยมสกุล และคณะ (2556) ชุดควบคุมเวลาการเปิด-ปิดเครื่องใช้ไฟฟ้า 8 จุดโดย ไมโครคอนโทรเลอร์ มหาวิทยาลัยสยาม โครงงานนี้สามารถเลือกการใช้งานได้ 8 เอาต์พุตแต่ละเอาต์พุต สามารถเลือกแรงดันการใช้งานได้ 3 ระดับ ควบคุมโดยซีเลคเตอร์สวิตช์ วณพันธ์ วัยวุฒิ (2555) ระบบควบคุมอุปกรณ์ไฟฟ้าแสงสว่างผ่านเครือข่ายไร้สาย สาขาวิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร คณะวิศวกรรมศาสตร์ ระบบ เครือข่ายไร้สายเป็นระบบที่มีข้อจำกัดเรื่องพลัง ระยะทางในการสื่อสารและความสามารถในการ ประมวลผลต่ำ แต่จะทำงานรวมกันเป็นเครือข่ายที่ใช้ได้ในพื้นที่กว้างรับ และส่งข้อมูลจากโหนด (Node) แต่ละโหนดในเครือข่าย สามารถรวบรวมและเก็บข้อมูล มีการส่งซ้ำเพื่อประมวลผลหรือเตือน การควบคุม ภาระแสงสว่างไม่ต้องใช้สายไฟ ด้วยเหตุผลดังกล่าว ผู้วิจัยได้นำมาเป็นแนวทางในการศึกษาระบบการติดต่อโดยผ่านระบบ เครือข่ายไร้สาย ใช้ไมโครคอนโทรเลอร์ในการคุมควบการทำงาน และควบคุมการเปิด-ปิดเครื่องใช้ไฟฟ้า ผ่านเว็บเบราว์เซอร์ เพื่อนำมาประยุกต์พัฒนาเข้าในงานวิจัยเรื่องระบบการควบคุมการเปิด-ปิดไฟด้วย ผ่านเว็บเบราว์เซอร์ วัสดุอุปกรณ์หลักในการวิจัย Arduino Mega 2560 คือ บอร์ดไมโครคอนโทรลเลอร์ชนิดหนึ่ง ที่ประกอบด้วย ATmaga 2560 ซึ่งมี 54 ดิจิตอล อินพุต เอาท์พุต โดยในส่วนประกอบที่สามารถใช้งานเป็น PWM ได้ 15 ขา อนาล็อคอินพุต 16 ขา UART 4 ชุด โดยความถี่คริสตรัลบนบอร์ดคือ 16 MHz เชื่อมต่อข้อมูลระหว่าง คอมพิวเตอร์ผ่านพอร์ต USB บนบอร์ดได้โดยตรง อีกทั้งรูปแบบการออกแบบยังออกแบบให้รองรับการ สวมกับ Shield ต่าง ๆ ได้โดยตรง ทำให้สามารถพัฒนาระบบต่าง ๆ ได้อย่างรวดเร็วและเรียบร้อยสวยงาม โดยรองรับการพัฒนาโปรแกรมบนแพลตฟอร์ม Arduino อย่างเต็มรูปแบบ ในงานวิจัยนี้เราใช้ MATLAB ต่อกับไมโครโฟนเพื่อรับเสียงและวิเคราะห์เสียงและ MATLAB ทำการต่อกับ Arduino เพื่อสั่งงานให้เปิด-ปิด หลอด LED



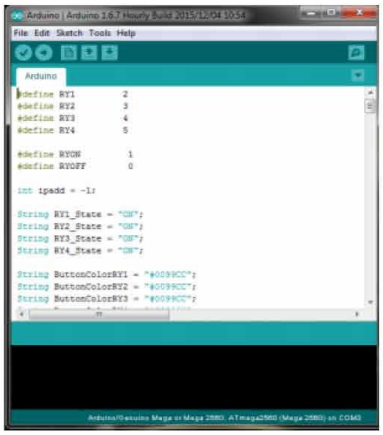
ภาพที่ 1 หลักการทำงาน

วัสดุอุปกรณ์หลักในการวิจัย

Arduino อ่านว่า (อา-ดู-อิ-โน่ หรืออาดุยโน่) เป็นบอร์ดไมโครคอนโทรเลอร์ตระกูล AVR ที่มี การพัฒนาแบบ Open Source คือ มีการเปิดเผยข้อมูลทั้งด้าน Hardwareและ Software ตัวบอร์ด Arduino ถูกออกแบบมาให้ใช้งานได้ง่าย ดังนั้นจึงเหมาะสำหรับผู้เริ่มต้นศึกษา ทั้งนี้ผู้ใช้งานยังสามารถดัดแปลง เพิ่มพัฒนาต่อยอด หรือโปรแกรมต่อได้อีกด้วย ความง่ายของบอร์ด Arduino ในการต่ออุปกรณ์เสริมต่าง ๆ คือผู้ใช้งานสามารถต่อวงจรอิเล็กทรอนิกส์จากภายนอกแล้วเชื่อมต่อเข้ามาที่ขา I/O ของบอร์ด ดังรูปที่ 2 หรือเพื่อความสะดวก สามารถเลือกต่อกับบอร์ดเสริม (Arduino Shield) ประเภทต่าง ๆ เช่น Arduino XBee Shield, Arduino Music Shield, Arduino Relay Shield, Arduino Wireless Shield, Arduino GPRS Shield เป็นต้นมาเสียบกับบอร์ดบนบอร์ด Arduino แล้วเขียนโปรแกรมพัฒนาต่อได้เลย

ภาพที่ 2 บอร์ดไมโครคอนโทรเลอร์ Arduino

โปรแกรม Arduino IDE เป็นโปรแกรมที่ใช้ในการควบคุมในการเขียนคำสั่ง ลงบน Arduino UNO R3 เพื่อใช้ในการทำงานควบคุมอุปกรณ์ต่าง ๆ ดังภาพที่ 3



ภาพที่ 3 โปรแกรม Arduino IDE

2. วัตถุประสงค์

2.1 เพื่อศึกษาการใช้ระบบคำสั่งเสียงตัวเลข

2.2 เพื่อพัฒนาการออกแบบอุปกรณ์ควบคุมการเปิด-ปิดอุปกรณ์ไฟฟ้ากับการใช้คำสั่งด้วยเสียง

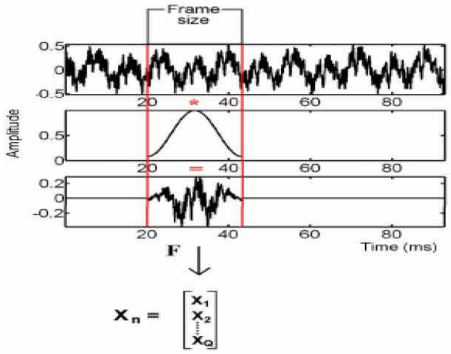
2.3 เพื่อศึกษาประสิทธิภาพการใช้งานของอุปกรณ์ควบคุมการเปิด-ปิดอุปกรณ์ไฟฟ้าด้วยเสียง

3. แนวคิดทฤษฎีและงานวิจัยที่เกี่ยวข้อง

เสียงของคำพูดมนุษย์นั้นจะสามารถแบ่งตามลักษณะเสียงของคำพูดได้สามแบบด้วยกัน [1] คือ แบบเสียงคำพูดโดด แบบเสียงคำพูดต่อกัน และแบบเสียงคำพูดต่อเนื่อง ซึ่งในการสื่อสารกันด้วยคำพูดนับเป็น

รูปแบบที่ใช้กันทั่วไปอย่างแพร่หลายในการสื่อสารระหว่างกัน มีงานวิจัยหลายชิ้นที่ได้มุ่งเน้นในการพัฒนาระบบที่สามารถสร้างความเข้าใจเสียงพูดของมนุษย์ด้วยคอมพิวเตอร์ จากการสร้างปฏิสัมพันธ์ระหว่างมนุษย์กับคอมพิวเตอร์ [2] คำพูดนั้นมีความสำคัญในการสร้างปฏิสัมพันธ์ระหว่างมนุษย์กับคอมพิวเตอร์ และความสนใจการพัฒนาคอมพิวเตอร์ที่สามารถตอบรับเสียงพูดได้เมื่อมีการป้อนข้อมูล มีการพัฒนาอย่างต่อเนื่อง มีความพยายามที่จะวิจัยที่พัฒนาการรับรู้ด้วยเสียงที่มีค่าใช้จ่ายที่ต่ำซึ่งจะส่งผลให้แอปพลิเคชันในการจดจำเสียงมีมากขึ้นในอนาคตอันใกล้ [3] ในส่วนของประเทศไทยก็มีงานวิจัยหลายงานที่มีความเกี่ยวข้องในการประยุกต์เสียงพูดในลักษณะต่างๆ กันอาทิเช่น การสั่งงานรถเข็นคนพิการกระบวนการแยกแยะคำสั่งในระบบรู้จำเสียงพูดสำหรับควบคุมรถเข็นคนพิการ ในการเคลื่อนที่ไปในทิศทางต่าง ๆ โดยสามารถแยกแยะคำพูดที่มีความหมายเป็นคำสั่งจากกลุ่มคำพูด หรือประโยคที่เป็นคำสั่ง หรือไม่ใช่ประโยคคำสั่ง แล้วแปลความหมายของคำสั่งนั้น ยกตัวอย่าง เช่น เดินหน้า เลี้ยวซ้าย เลี้ยวขวา เป็นต้น [4] ใช้การเข้ารหัสสัญญาณเสียงพูดโดยใช้โปรแกรม MATLAB มาช่วยศึกษาในการบีดอัดเสียงพูด [5] การรู้จำเสียงพูดภาษาไทยแบบหลายผู้พูดด้วยนิวรัลเน็ตเวอร์คสองชุดที่มีหนึ่งชั้นฮิดเดิ้ลที่เหมือนกัน [6] ระบบควบคุมบ้านอัจฉริยะช่วยให้คนพิการสามารถควบคุมอุปกรณ์ไฟฟ้าภายในบ้านเช่นโทรทัศน์ หลอดไฟ และพัดลมโดยใช้คำสั่งเสียงเพียงอย่างเดียวโดยไม่ต้องย้ายเพื่อเปิดหรือปิดอุปกรณ์ไฟฟ้า การรับรู้เสียงในอุปกรณ์ไฟฟ้ากำลังใช้แอปพลิเคชันของ Google Assistant บนสมาร์ทโฟน แอปพลิเคชัน Google Assistant จะยอมรับคำสั่งเสียงเมื่อการออกเสียงถูกต้อง คำสั่งเสียงใน Smart Home ที่ใช้ IoT (Internet of Things) นั้นง่ายต่อการใช้งานมากขึ้นโดยไม่ต้องพิมพ์ข้อความ [7] ใช้การผสมผสานระหว่างคุณสมบัติการควบคุมที่แตกต่างกันมีความสามารถในการตรวจจับสิ่งกีดขวางของรถเข็นคนพิการอัจฉริยะ ได้รับการพัฒนาโดยใช้ระบบจดจำเสียงเพื่อควบคุมการเคลื่อนไหวของรถเข็นและจอยสติ๊กแบบ Arduino นอกจากนี้ระบบอัลตร้าซาวด์ยังให้ความสะดวกในการตรวจจับสิ่งกีดขวางอัตโนมัติ ทำให้ได้สามารถอำนวยความสะดวกคนพิการได้ในราคาที่ต่ำ [8]

3.1 วินโดวส์และเฟรม



ภาพที่ 4 กราฟสัญญาณและการลดขนาดเฟรม

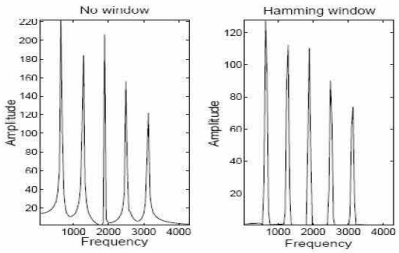
กราฟบนสุด เป็นสัญญาณเสียงดิบ ส่วนกราฟกลางเป็น hamming window และเป็นผลลัพธ์ของการคูณกันระหว่างรูปบนสุดและรูปกลาง ข้อดีที่ชัดเจนคือผลลัพธ์ของสัญญาณ ในรูปสุดท้ายเป็นการลดขนาดของเฟรมซึ่งลดปัญหา การรั่วของสเปกตรัม (spectral leakage) ด้วย ส่วน X_n คือเวกเตอร์คุณลักษณะเด่น

ลักษณะรูปคลื่นของสัญญาณเสียงนั้นเป็นสัญญาณที่มีการเปลี่ยนแปลงตามเวลา การหาค่าลักษณะสำคัญบนฐานของการแปลงฟูริเยร์นั้นจึงต้องใช้เทคนิคที่เรียกว่า การแปลงฟูริเยร์ในช่วงสั้น (Short-time Fourier Transform: STFT) เพื่อวิเคราะห์สัญญาณเสียงไปเป็นลำดับของเฟรม (Sequence of frames) นั้น ๆ วินโดวส์ (Window) ใช้เพื่อแบ่งเฟรมของสัญญาณเสียงในเชิงเวลา วินโดว์นั้นมีหลายแบบมากซึ่งสามารถเลือกใช้ได้ตามความเหมาะสมของงานนั้น ๆ วินโดวส์ (Window) ในเชิงเวลา กำหนดด้วย $w(n)$ ซึ่งเป็นลำดับจำกัด (Finite-length sequence) สำหรับใช้ในการเลือกเฟรมจากสัญญาณเดิม (s(n)) ด้วยการคูณในเชิงเวลา สมมติให้วินโดวส์เป็น Causal sequence ซึ่งมีจุดเริ่มต้นที่ n = 0 มีระยะ (Duration) คือ N_w และเป็น Linear phase sequence วินโดวส์ในเชิงความถี่จะถูกกำหนดตามสมการ

$$W(n)=\left|W(\omega)\right|e^{-j\omega\left(\frac{N_w-1}{2}\right)} \tag{1}$$

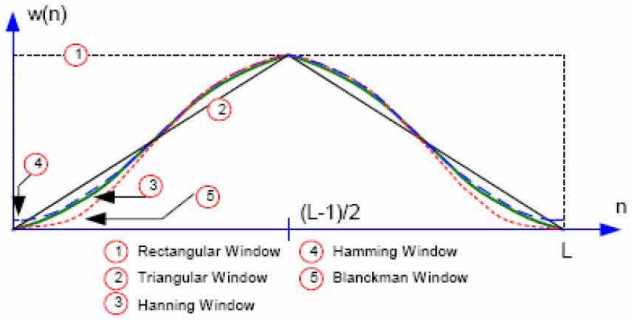
กระบวนการสร้างเฟรมโดยใช้วินโดวส์ทำได้โดยกลับรูป (Reverse) วินโดวส์ในเชิงเวลาและใช้การเลื่อน (Shift) ขอบของวินโดวส์ไปยังจุดเริ่มต้นซึ่งเป็นเวลาที่ต้องการ m จะได้ $w(m-n)$ และเฟรมของสัญญาณ $s(n)$ กล่าวคือ f_s(n;m) ซึ่งมีความยาวเท่ากับ N และสิ้นสุดที่ m จะได้จาก

$$f_s(n;m)=s(n)w(m-n),n=m-N_w+1,\ldots,m \tag{2}$$



ภาพที่ 5 แสดงสเปกตรัมความถี่ของสัญญาณฮาร์โมนิกส์ด้วยความถี่พื้นฐานและ สี่โอเวอร์โทน

สัญญาณมีการ sampling ความถี่เป็น 22 kHz และขนาดของเฟรมเป็น 512 samplesโดยทั่วไปฟังก์ชันวินโดวส์ที่ใช้จะเป็นลักษณะเรียว (Tapered window) เพื่อลดความไม่ต่อเนื่องที่จะเกิดขึ้นที่ขอบ ซึ่งส่งผลให้เกิดองค์ประกอบความถี่สูงขึ้นในสเปกตรัม นอกจากนี้ความยาวของวินโดวส์ที่ใช้ยังส่งผลต่อ Time resolution และ Frequency resolution ของสเปกตรัมอีก หากวินโดวส์มีความยาวมาก แต่ Frequency resolution ที่ต่ำจะทำให้ได้รายละเอียดการกระจายข้อมูลของสัญญาณน้อยกว่าที่เป็นอยู่จริงมาก สำหรับฟังชั่นวินโดวส์พื้นฐานที่มีอยู่ในปัจจุบันนั้นมีอยู่ 5 วินโดว์ (เห็นได้จากภาพที่ 6) คือ วินโดวส์ Rectangular วินโดวส์Triangular วินโดวส์ Hanning วินโดว์ hamming และ วินโดว์ blanckman ตามลำดับ



ภาพที่ 6 ฟังก์ชั่นวินโดว์

สำหรับวินโดวส์ที่นิยมใช้ที่รับมาจากพื้นฐานของวินโดวส์ hanning คือ วินโดวส์ hamming เนื่องจากมี Frequency resolution ที่ค่อนข้างสูงและการลดทอนที่ Side lobs ที่มากกว่า 43 dB ซึ่ง วินโดวส์ Hamming นี้มีความแตกต่างจากวินโดวส์ hanning คือ วินโดวส์ hamming มีขอบของวินโดวส์ที่สูงขึ้น 0.08 ฟังก์ชันวินโดวส์ hamming กำหนดตามสมการ

$$w(n) = \begin{cases} 0.5 - 0.46\cos\left(\dfrac{2\pi n}{N_w - 1}\right), n = 0, 1, ..., N_w \\ 0 \qquad\qquad\qquad\qquad , n \quad otherwise \end{cases} \tag{3}$$

3. การวิเคราะห์ระบบการจำแนกเสียง

การวิเคราะห์นี้มีโครงสร้างดังภาพที่ 7 โดยขั้นตอน 1 เรียกว่า Preprocessing เป็นการปรับปรุงคุณภาพสัญญาณเสียงโดยที่จะแปลงความถี่จาก 44.1 kHz เป็น 22 kHz ขั้นตอนที่ 2 เป็นการนำเสนอเครื่องมือการดึงคุณลักษณะเด่นของเสียง ซึ่งในงานวิจัยนี้ใช้ค่าสัมประสิทธ์เซปสตรัมบนสเกลเมล (Mel Frequency Cepstral Coefficient) เป็นวิธีการจำแนกเสียง

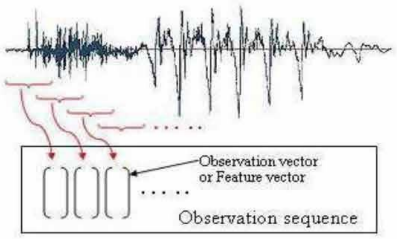
ภาพที่ 7 แสดงบล็อกไดอะแกรมของระบบการจำแนกเสียง

3.1 กระบวนการปรับปรุง (Preprocessing)

ในขั้นตอนนี้เป็นการทำ Down-sampling คือจะทำสัญญาณเสียง จาก 44.1 kHz เป็น 22.050 kHz ซึ่งการทำ down-sampling มีผลกระทบในทางที่ดีต่อการวิเคราะห์ในการค้นคว้าความคล้ายคลึงของเสียงการดึงสัญญาณเสียงดิจิตอลจากไฟล์เป็นการแทน $x[n] \in R$ จาก $n = 0, ..., N-1$ ที่ซึ่ง N เป็นจำนวนของการสุ่มในไฟล์เสียง

3.2 กระบวนการหาค่าเวกเตอร์คุณลักษณะเด่นของเสียง (Feature Extraction)

หัวข้อนี้เป็นการคำนวณเพื่อหาลักษณะต่าง ๆ ของรูปคลื่นสัญญาณ โดยที่สัญญาณมีการเปลี่ยนแปลงตามเวลา การหาค่าคุณลักษณะเด่นบนฐานของการแปลงฟูริเยร์นั้นจึงต้องใช้เทคนิคที่เรียกว่า การแปลงฟูริเยร์ในช่วงสั้น (Short-time Fourier Transform: STFT) เพื่อวิเคราะห์สัญญาณเสียงไปเป็นลำดับของเฟรม (Sequence of frames) ซึ่งแต่ละเฟรมแทนด้วยคอลัมน์เวกเตอร์ของเวกเตอร์ค่าคุณลักษณะเด่นของสเปกตรัมในเฟรมนั้น ๆ



ภาพที่ 8 แสดงการดึงคุณลักษณะเด่นของเสียง

Mel-Frequency Cepstral Coefficient (MFCC) ค่าสัมประสิทธิ์เซปสตรัมบนสเกลเมล (Mel Frequency Cepstral Coefficient: MFCC) ซึ่งเป็นลักษณะเด่นที่ปรับปรุงมาจาก สัมประสิทธิ์เซปสตรัมปกติ ด้วยการปรับสเกลของสเปกตรัมให้อยู่บนเสกลที่เหมาะสมสำหรับการรับฟังของมนุษย์ เพื่อให้สามารถเก็บรายละเอียดของสัญญาณเสียงช่วงความถี่ต่ำ ซึ่งมีนัยสำคัญมากกว่า ช่วงความถี่สูง เรียกว่า สเกลเมล (Mel Scale) ความถี่ของเสียงพูดสามารถปรับให้อยู่ในสเกลเมล (m) ดังสมการ

$$Mel(x) = 2595 \times \log_{10}(1 + \frac{f}{700}) \quad (4)$$

ที่ซึ่ง f เป็นความถี่ในหน่วย เฮิร์ต ซึ่งเป็นความถี่ในเชิงเส้น

3.3 การจำแนก

การทำงานของตัว Classifier นั้นคือการรับเอา เวกเตอร์คุณลักษณะเด่น ซึ่งเป็นการเขียนคุณสมบัติต่าง ๆ ของสัญญาณเสียงที่สามารถวัดได้หรือคำนวณได้ให้อยู่ในรูปของเวกเตอร์เข้าไป แล้วทำการตัดสินใจว่าสัญญาณเสียงนั้น ๆ อยู่ในกลุ่มหมายเลขใดหรือชื่อใดแล้วให้ชื่อ หรือหมายเลขกลุ่มที่สัญญาณเสียงเป็นสมาชิกอยู่ออกมา ซึ่งการที่ Classifier จะสามารถตัดสินใจได้ จะต้องมีสัญญาณเสียงตัวอย่างของแต่ละกลุ่มเสียก่อน ในงานวิจัยนี้ใช้ K-nearest neighborhood (K-NN) สำหรับการจำแนกชนิดเสียงในแต่ละบริเวณที่แตกต่างกัน

3.4 การตัดสินใจโดย K-nearest neighborhood

K-nearest neighborhood (K-NN) เป็นการเรียนรู้ที่ดีมากต่อ เทคนิคการจำแนกแบบ nonparametric และ parametric และมีประสิทธิภาพสำหรับเอกสารที่ซึ่งใช้ข้อมูลการสอนที่มาก ในวิทยานิพนธ์นี้จึงใช้ความเหมาะสมนี้ การจำแนกโดยการใช้ K-nearest neighborhood (K-NN) เป็นตัวเลือกที่ดี ในวิธี Nearest neighborhood เป็นความจำเป็นต่อการกำหนดค่าพารามิเตอร์ k อันซึ่ง k เป็นจำนวน

ของคุณลักษณะเด่นที่ใกล้เคียง ตัวอย่างเป็นการอธิบายในเทอมของ standard Euclidean distance ระยะห่างระหว่างเวกเตอร์สามารถแสดงเป็น $\|v_1-v_2\|$ คือ

$$\|v_1-v_2\|=\sqrt{\sum_{i=1}^{d}\left(v_1[i]-v_2[i]\right)^2} \quad (5)$$

4. ผลการทดลอง

ซอฟต์แวร์นี้จะทำงานบนระบบ window ที่มีไมโครโฟนอยู่ มันซึ่งใช้การเปรียบเทียบขั้นพื้นฐานเพื่อพิจารณาว่าคำสั่งเสียงนั้นเหมาะสมกับรูปแบบที่ระบุหรือไม่ ในการเปรียบเทียบเสียงใช้โปรแกรม MATLAB ในการประมวลผลและทำการสั่งหลอด LED โดยใช้ Arduino ฐานข้อมูลทั้งหมดทางผู้วิจัยได้จัดทำขึ้นมาโดยจัดทำในสิ่งแวดล้อมจริงประกอบไปด้วยเสียงพูดคำว่าเปิด 100 คำ และเสียงพูดคำว่าปิด 100 คำ จำนวน 5 คน คนละ 40 เสียง รวมทั้งสิ้น 200 เสียง มีการให้สัญญาณรบกวนที่แตกต่างกัน สำหรับข้อมูลการสอน การทดลองเป็นการพิจารณาการดึงเวกเตอร์คุณลักษณะเด่นโดยใช้วิธี MFCC และข้อมูลทดสอบมีเสียง 50 เสียง รวม 250 เสียง สำหรับวิธีการหาอัตราความถูกต้อง ในเอกสารนี้ใช้สมาการตามข้างล่าง

$$Accuracy(\%)=\left(\frac{In-Out}{Total}\right)\times 100 \quad (6)$$

ในตารางที่ 1 อธิบายผลการทดลองโดยวิธีการหาคุณลักษณะเด่นที่นำเสนอ

ตารางที่ 1 ผลการทดลองเสียงเปิด

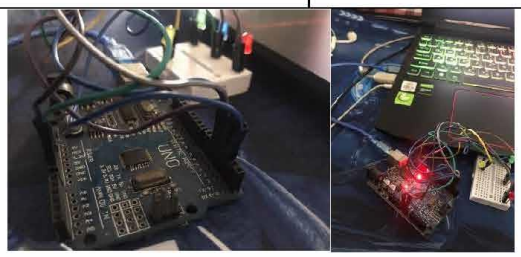
ครั้งที่	จำนวนเสียงที่ใช้ทดลอง	จำนวนเสียงเปิดที่ระบุตัวได้ถูกต้อง	%ความถูกต้อง	%ความถูกต้องเฉลี่ย
1	200	40	80	
2	200	41	82	
3	200	40	80	81.2
4	200	42	84	
5	200	40	80	

ตารางที่ 2 ผลการทดลองเสียงปิด

ครั้งที่	จำนวนเสียงที่ใช้ทดลอง	จำนวนเสียงเปิดที่ระบุตัวได้ถูกต้อง	%ความถูกต้อง	%ความถูกต้องเฉลี่ย
1	200	37	74	
2	200	35	70	
3	200	37	74	72.8
4	200	37	74	
5	200	36	72	

ตารางที่ 2 ผลการทดลองโดยรวม

วิธีการ	อัตราความถูกต้อง (%)
การทดลองโดยรวม	77



ภาพที่ 9 การต่อ LED เข้ากับ Arduino

5. สรุปผลการทดลอง

ในเอกสารฉบับนี้เราได้แสดงให้เห็นถึงการออกแบบและคุณสมบัติของระบบควบคุมด้วยเสียงอัตโนมัติ แบบเป็นออฟไลน์แบบมีสายและสามารถมีความยืดหยุ่นในแง่ของค่าใช้จ่าย ระบบนี้มีหลักคุณลักษณะสำหรับการตรวจหาคำพูดซึ่งจะถอดรหัสประโยคของผู้ใช้ลงในคำสั่งที่เหมาะสมมากขึ้นและเพิ่มระบบอัตโนมัติ

6. เอกสารอ้างอิง

[1] Morgan, David P., and Scofield, Christopher L. (1991). **"Neural Networks and Speech Processing." U.S.A.** : Kluwer Academic Publishers., 1991.

[2] P. P. Shrishrimal, R. R. Deshmukh, V. B. Waghmare. (2012). **"Indian language speech database: A review", International Journal of Computer Applications,** Vol. 47, No. 5, pp. 17-21, 2012.

[3] C. Huang, T. Chen, E. Chang. (2004). "Accent issues in large vocabulary continuous speech recognition**", International Journal of Speech Technology**, Vol. 7, No. 2-3, pp. 141-153, 2004.

[4] อนุพงษ์ ธรรมรักษาสิทธิ์. (2546) การแยกแยะคำสั่งควบคุมรถเข็นคนพิการในระบบรู้จำเสียงพูด . มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี/กรุงเทพฯ.

[5] รุ่งอรุณ คงดี. (2544). **การเข้ารหัสสัญญาณเสียงพูดโดยใช้โปรแกรม MATLAB ช่วยศึกษาการบีดอัดเสียง .** มหาวิทยาลัยเกษตรศาสตร์/กรุงเทพฯ.

[6] ไชยันต์ สุวรรณชีวะศิริ. (2540). **การรู้จำเสียงพูดภาษาไทยแบบหลายผู้พูดด้วยนิวรัลเน็ตเวอร์ค.** มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี/กรุงเทพฯ.

[7] Isyanto, Haris Arifin, Ajib Setyo Suryanegara, Muhammad. (2020). **"Design and** Implementation of IoT-Based Smart Home Voice Commands for disabled people using Google Assistant.**" International Conference on Smart Technology and Applications (ICoSTA) Smart Technology and Applications (ICoSTA),** 2020 International Conference on. :1-6 Feb, 2020.

[8] Md. Mamunur Rahman ; Swarup Chakraborty ; Avradip Paul ; Ali Mohammed Jobayer and Md. Azad Hossain. (2017). "Wheel therapy chair: A smart system for disabled person with therapy facility." **International Conference on Electrical, Computer and Communication Engineering (ECCE),** 2017.

[9] Chai, W. and Vercoe, B., (2001). "Folk music classification using hidden markov models", IEEE International Conference on Artificial Intelligence, USA, pp. 341-345.

[10] Soltau, H., Schultz, T., Westphal, M. and Waibel, A., 1998, "Recognition of music types," **IEEE International Conference on Acoustics, Speech, and Signal Processing, May** 12-15, USA, pp. 1137-1140.

[11] Matityaho, B. and Furst, M., (1995). "Neural network-based model for classification of music type", IEEE International Conference on Electrical and Electronic Engineers, Mar 7-8, Israel, pp.1-5.

[12] เครือข่ายเซ็นเซอร์ไร้สาย : http://www.thaitelecomkm.org/TTE/topic/attach/Wireless_Sensor_Network/index.php ค้นหาเมื่อ 11 พฤศจิกายน 2558

[13] เว็บเบราว์เซอร์ http://it.travel2mukdahan.com/2013/04/web-browser.html ค้นหาเมื่อ 2 ธันวาคม2558

[14] การเขียนโปรแกรมพื้นฐาน http://ilearn.lnwshop.com/article/3-การเขียนโปรแกรมพื้นฐาน-arduino.ค้นหาเมื่อ 10 มกราคม2559

[15] ศึกษาการใช้บอร์ดอุปกรณ์ของ Arduino และ http://thaieasyelec.com/ ค้นหาเมื่อ 10 มกราคม2559

[16] วณพันธ์ วัยวุฒิ.(2555). **ระบบควบคุมอุปกรณ์ไฟฟ้าแสงสว่างผ่านเครือข่ายไร้สายอัตตราต่ำ.** คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร.

[17] สัญชัย เอี่ยมสกุล และคณะ.(2556). **ชุดควบคุมเวลาการเปิด ปิดเครื่องใช้ไฟฟ้า 8 จุดโดยไมโครคอนโทรเลอร์.**คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสยาม.

[18] เสฐียรพงษ์ สมบัติใหม่ อาทร สิริกุลธมาลา.(2557). **ระบบควบคุมไฟฟ้าแสงสว่างผ่านเครือข่ายไร้สาย.** คณะวิศวกรรมศาสตร์มหาวิทยาลัยเทคโนโลยีราช.