

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การรู้จำเสียงพูด (Speech Recognition) คือ การทำให้คอมพิวเตอร์เข้าใจความหมายคำพูดของมนุษย์ โดยการแปลงสัญญาณเสียงพูดให้เป็นข้อมูลซึ่งมีรูปแบบ (Pattern) ที่แตกต่างกันตามการออกเสียง และมีกระบวนการแบ่งแยก (Classification) รวมทั้งจัดจำรูปแบบของข้อมูลซึ่งแทนคำพูดแต่ละคำ ด้วยเหตุนี้ระบบรู้จำเสียงพูดจึงทำให้มนุษย์สามารถควบคุมและสั่งงานคอมพิวเตอร์รวมถึงอุปกรณ์รอบข้างด้วยเสียงพูดผ่านทางไมโครโฟนได้ เมื่อนำระบบรู้จำเสียงพูดไปใช้งานร่วมกับอุปกรณ์รับข้อมูลอย่างแป้นพิมพ์และเมาส์จะทำให้คอมพิวเตอร์มีช่องทางในการติดต่อกับมนุษย์มากขึ้น จึงมีการพัฒนาวิธีการรู้จำเสียงพูดอย่างต่อเนื่องเพื่อเพิ่มความถูกต้องในการรู้จำ จนกระทั่งปัจจุบันในต่างประเทศมีบริษัทผลิตซอฟต์แวร์ (Software) เกี่ยวกับการรู้จำเสียงพูด เช่น โปรแกรมพิมพ์ตามเสียงพูด โปรแกรมที่ทำงานโดยใช้คำสั่งเสียง ซึ่งเป็นการเพิ่มความสะดวกสบายให้กับผู้ใช้ทั่วไป และยังช่วยให้ผู้ใช้ที่พิการทางด้านร่างกาย เช่น แขนหรือมือสามารถใช้คอมพิวเตอร์ได้ง่ายขึ้น

เทคโนโลยีเกี่ยวกับการรู้จำเสียงพูดส่วนใหญ่ถูกสร้างมาในรูปแบบของซอฟต์แวร์เพื่อใช้กับคอมพิวเตอร์ส่วนบุคคล (PC) ที่มีหน่วยความจำ (Memory) ขนาดใหญ่และตัวประมวลผลกลาง (CPU) ที่มีความเร็วสูงซึ่งออกแบบมาเพื่อใช้ทำงานหลายๆ อย่าง แต่ในกรณีที่ต้องการสร้างเครื่องมือที่ใช้งานเฉพาะอย่างโดยนำเอาคอมพิวเตอร์มาประมวลสัญญาณจะทำให้ใช้งานได้ไม่สะดวกเนื่องจากภายในเครื่องคอมพิวเตอร์ประกอบด้วยอุปกรณ์ที่มีขนาดใหญ่และมีน้ำหนักมาก ดังนั้นหากมีการออกแบบระบบสมองกลฝังตัว (Embedded system) เพื่อใช้ในการรู้จำเสียงพูดซึ่งเป็นการลดขนาดของหน่วยความจำและหน่วยประมวลผลจะทำให้เทคโนโลยีการรู้จำเสียงพูดสามารถนำไปประยุกต์ใช้งานได้หลากหลายยิ่งขึ้น

FPGA (Field Programmable Gate Array) เป็นอุปกรณ์สารกึ่งตัวนำที่มีโครงสร้างภายในประกอบด้วยส่วนของวงจรถอดจิกที่สามารถกำหนดการทำงานได้ (Configurable Logic Blocks) จึงสามารถออกแบบ FPGA ให้ทำงานเหมือนวงจรถอดจิกพื้นฐาน ตั้งแต่ระดับลอจิกเกตหรือรวมกันเพื่อทำงานที่มีความซับซ้อนมากขึ้น เช่น Decoders หรือ Multiplexers จนถึงการออกแบบในระดับสูงโดยใช้ภาษา HDL (Hardware Description Language) เช่น การประมวล

สัญญาณดิจิทัล (Digital Signal Processing) ซึ่งใน FPGA ทั่วไป นอกจากจะประกอบด้วยส่วนของ Logic blocks แล้ว จะยังมีส่วนของหน่วยความจำ ซึ่งอาจจะสร้างโดยใช้ฟลิปฟลอป (Flip-Flops) หรือใช้พื้นที่ของสารกึ่งตัวนำสร้างเป็นหน่วยความจำจริงๆ อยู่ภายในก็ได้

ในการออกแบบวงจรดิจิทัลอิเล็กทรอนิกส์ที่มี FPGA อยู่บนแผงวงจรด้วยนั้น จะช่วยให้ผู้ออกแบบสามารถลดจำนวนอุปกรณ์ลงได้ ทำให้แผงวงจรมีขนาดเล็กลงและสามารถออกแบบแผงวงจรได้โดยไม่ต้องทดสอบรายละเอียดทั้งหมดภายในชิ้นงาน ซึ่ง FPGA มีข้อได้เปรียบเรื่องเวลาที่ใช้ในการพัฒนาระบบเนื่องจากวงจรที่ออกแบบทั้งหมดสามารถนำไปจำลองการทำงาน (Simulation) ผ่านทางโปรแกรมคอมพิวเตอร์ เมื่อได้ผลที่ถูกต้องก็ทำการบรรจุ (Download) วงจรที่ออกแบบไว้ลงบน FPGA แล้วนำไปใช้งานจริงได้ทันที และยังสามารถปรับเปลี่ยนหรือแก้ไขการทำงานของวงจรได้หลังจากที่ใช้งานไปแล้ว

การลดความซับซ้อนในการออกแบบหน่วยประมวลผลสามารถทำได้โดยใช้เครื่องมือ (Design tools) ในการออกแบบ Embedded system ช่วยสร้างไมโครโปรเซสเซอร์ขึ้นจาก Configurable Logic Blocks (CLBs) ภายใน FPGA และกำหนดอุปกรณ์ต่อร่วม (Peripherals) ที่ใช้ในระบบ แล้วจึงทำการออกแบบระบบในลักษณะการออกแบบร่วมกันระหว่างฮาร์ดแวร์-ซอฟต์แวร์ (Hardware-software codesign) โดยในการออกแบบส่วนของซอฟต์แวร์ที่ทำงานในไมโครโปรเซสเซอร์จะเขียนโดยใช้ภาษา C และส่วนของฮาร์ดแวร์จะใช้ภาษา HDL ซึ่งทำให้ได้ระบบที่มีประสิทธิภาพภายในระยะเวลาที่สั้นลง และความยืดหยุ่นในการออกแบบทางด้านฮาร์ดแวร์ของ FPGA จะทำให้ระบบมีความน่าสนใจมากขึ้น

แนวทางของงานวิจัยเกี่ยวกับการรู้จำเสียงพูดคำโดดที่ผ่านมาส่วนใหญ่เป็นการศึกษาค้นคว้าและปรับปรุงวิธีการในส่วนต่างๆ เพื่อเปรียบเทียบประสิทธิภาพของระบบ ต่อมาเมื่อระบบมีอัตราการเรียนรู้ที่สูงขึ้น แนวทางของงานวิจัยจึงเป็นการพัฒนาระบบเพื่อนำไปประยุกต์ใช้ในชีวิตประจำวัน เช่น ควบคุมสั่งงานอุปกรณ์ต่างๆ ซึ่งเป็นการพัฒนาโปรแกรมคอมพิวเตอร์และส่วนเชื่อมต่อกับอุปกรณ์ที่ต้องการควบคุม สำหรับงานวิจัยนี้มุ่งเน้นที่การออกแบบและพัฒนาระบบรู้จำเสียงพูดคำโดดบนอุปกรณ์ FPGA โดยสร้างไมโครโปรเซสเซอร์ภายใน FPGA เพื่อทำหน้าที่ประมวลสัญญาณเสียงแทนหน่วยประมวลผลกลางของเครื่องคอมพิวเตอร์ ซึ่งการออกแบบระบบบน FPGA เป็นการรวมกันของการออกแบบฮาร์ดแวร์และการออกแบบโปรแกรม

1.2 วัตถุประสงค์ของวิทยานิพนธ์

- 1.2.1 เพื่อศึกษาและพัฒนาาระบบรู้จำเสียงพูดคำโดด (Isolated Word Recognition)
- 1.2.2 เพื่อหาวิธีการที่เหมาะสมในการออกแบบระบบรู้จำเสียงพูดคำโดดบนอุปกรณ์

FPGA

1.3 ขอบเขตของวิทยานิพนธ์

สร้างระบบรู้จำเสียงพูดคำโดดที่ทำงานบนอุปกรณ์ FPGA ซึ่งมีการทำงานแบบกดสวิตช์แล้วพูดคำสั่ง โดยสามารถรู้จำเสียงพูดคำสั่งจำนวน 7 คำได้ในสภาพแวดล้อมที่ไม่มีเสียงรบกวน และหาประสิทธิภาพความถูกต้องในการรู้จำของระบบ

1.4 ประโยชน์ที่คาดว่าจะได้รับจากวิทยานิพนธ์

- 1.4.1 ได้ศึกษาการทำงานของระบบรู้จำเสียงพูดคำโดด
- 1.4.2 สามารถนำวิธีการต่างๆ ไปประยุกต์ใช้กับการรู้จำเสียงอื่นๆ ได้
- 1.4.3 สามารถนำไปออกแบบสร้างเครื่องมือที่มีระบบรู้จำเสียงพูดเป็นส่วนควบคุมการทำงานได้