



1.1 ประวัติความเป็นมาและความสำคัญของปัญหา

ในการใช้งานคอมพิวเตอร์ในปัจจุบันเรามักจะใช้ คีย์บอร์ด (Key board) หรือ เมาส์ (Mouse) เป็นอุปกรณ์รับข้อมูล (Input device) ในการสั่งงานให้คอมพิวเตอร์ทำงาน หรือรับข้อมูลตามที่เราต้องการ แต่ในบางสถานการณ์ หรือบางสภาวะ ผู้ใช้ไม่สามารถใช้อุปกรณ์ดังกล่าวในการสั่งงานคอมพิวเตอร์ได้ เช่นผู้พิการทางแขน มือ หรือนิ้ว หรืออาจทำงานอย่างอื่นอยู่และไม่สามารถใช้มือในการสั่งงานคอมพิวเตอร์ได้ นอกจากนั้นในบางระบบอาจจำเป็นต้องใช้เสียงในงานปฏิบัติงานเช่นระบบรักษาความปลอดภัยด้วยเสียง งานด้านจักรกลเช่นหุ่นยนต์ หรือเครื่องจักร

ในช่วง 25 ปีที่ผ่านมาตั้งแต่มีบทความ "Mad inventors and unreliable engineers" ของ John Pierce [15] ถ้ามองถึงความมุ่งหวังและความเป็นไปได้ ในเรื่อง การรู้จำเสียงพูด (Speech recognition) ซึ่งปัจจุบัน การรู้จำเสียงพูด (Speech Recognition) ถูกใช้อย่างกว้างขวางใน โปรแกรมสำเร็จรูป (Application) ต่างๆ หรือแม้แต่ระบบสื่อสาร ในระบบบริ่งมิ่ง (Cellular) ก็ใช้เพียงชิพตัวเดียวในการ รู้จำ (Recognize) คำสั่งเสียงโดยใช้ระบบ การรู้จำเสียง (Voice Recognition) ในขบวนการติดตั้งระบบโทรศัพท์ทางไกล และในคอมพิวเตอร์ส่วนบุคคล ก็สามารถเข้าใจคำพูดจากผู้ใช้ มีการผลิต Software ประเภทนี้เพื่อการค้าใน อเมริกา แคนาดา และในญี่ปุ่น และมีงานวิจัยในหัวข้อนี้ในห้องทดลองทั่วโลก สามารถ รู้จำ (Recognize) คำได้เป็นพันๆ คำ และในบางระบบให้ความถูกต้องสูงถึง 95% (ค่าจากผู้ผลิต) ในปัจจุบันมีการพัฒนาโปรแกรมที่เรียกว่า โปรแกรมรู้จำเสียงพูด (SPEECH RECOGNITION PROGRAM) ขึ้นมาหลายบริษัทเช่น Dragon system Inc. , L&H Speech Products หรือ IBM และในต่างประเทศก็มีการใช้งานอย่างกว้างขวาง แต่ในประเทศไทยถือได้ว่าน้อยมาก โดยเฉพาะงานวิจัยทางด้าน การรู้จำเสียงพูด (Speech Recognition) สำหรับภาษาไทย และคนไทย ซึ่งมีรูปแบบของเสียงและการพูดเฉพาะตัว

ดังนั้น ในการศึกษาและจัดทำโปรแกรมต้นแบบของการรู้จำเสียงพูด (Speech Recognition) ของคนไทยจะเป็นอีกก้าวหนึ่ง เพื่อส่งเสริมให้มีการใช้งานเทคโนโลยีดังกล่าว ในด้านต่างๆ ในประเทศไทย โดยอาศัยการ์ดเสียง (Sound Card) ซึ่งปัจจุบันเครื่องคอมพิวเตอร์ทุกเครื่องมักจะมี เพื่อให้เครื่องคอมพิวเตอร์รับเสียงจากผู้ใช้ และรับรู้ว่าผู้ใช้พูดคำว่าอะไร

นอกจากนี้ยังสามารถพัฒนาเป็นโปรแกรมเพื่อนำไปประยุกต์สำหรับผู้พิการแขนให้สามารถสั่งงานคอมพิวเตอร์ได้ และจะเป็นต้นแบบในการนำไปพัฒนา เพื่อเพิ่มความสามารถของโปรแกรมให้ทำงานอย่างอื่น ๆ หรือด้านอื่น เช่นระบบควบคุมอุปกรณ์ไฟฟ้าด้วยคำสั่งเสียง หรือการควบคุมเครื่องจักรกลเช่น หุ่นยนต์ การพัฒนาไปถึงการใช้เสียงแทนการพิมพ์ เช่น กรณีของโปรแกรมประมวลผลคำ (Word Processing)

1.2 วัตถุประสงค์

เพื่อพัฒนาขบวนการที่สามารถรับรู้เสียงพูดภาษาไทย โดยใช้เทคนิค Short time spectrum ในการวิเคราะห์สัญญาณ และโครงข่ายประสาทเทียม โดยพัฒนาโปรแกรมต้นแบบที่สามารถควบคุมระบบปฏิบัติการ Microsoft Windows 98

1.3 ขอบเขตงานวิจัย

- 1.3.1 พัฒนาโปรแกรมใช้เสียงพูดควบคุมการทำงานของระบบปฏิบัติการ Microsoft windows 98
- 1.3.2 พัฒนาโปรแกรมที่สามารถรับเสียงโดยผ่านทาง Sound Card ของคอมพิวเตอร์เพื่อนำไปประมวลผลเพื่อหาค่าประกอบด้านความถี่ของเสียง โดยใช้ทฤษฎี Short time Spectrum Analysis เพื่อนำไปเปรียบเทียบกับรูปแบบที่มีการฝึกไว้อยู่ในระบบ
- 1.3.3 โปรแกรมจะสามารถรู้จำเสียงที่ได้รับการฝึกไว้เท่านั้น โดยการพูดต้องพูดแบบ เป็นพยางค์ ไม่ต่อเนื่อง
- 1.3.4 โปรแกรมสามารถรู้จำคำสั่งภาษาไทยเป็นพยางค์ ดังต่อไปนี้
ชุดคำสั่งที่ระบบสามารถรู้จำ 50 ประเภทย่อย

ชุดตัวเลข : 0-9

ชุดเครื่องหมาย : + , - , x , / , ^ , = , % , จุด , เก็บ , เรียกคืน

ชุดคำสั่ง : ไทย , อังกฤษ , พิมพ์ , เปิด , ปิด , บันทึก , ย่อ , ขยาย , ปิดเครื่อง , คัดลอก , วาง , ล้าง , แทรก , ตกลง , ยกเลิก ,ช่วยเหลือ , เริ่ม , ค้นหา , วันที่ , คณิตเลข

ชุด Keyboard : ค ด้้น , ท้าย , ขึ้น , ลง , ซ้าย , ขวา , ถอย , เลื่อน , บรรทัด , ทั้งหมด

- 1.3.5 ผู้ทดสอบจะแบ่งเป็น 2 กลุ่มคือกลุ่มผู้ฝึก และกลุ่มนอกผู้ฝึก เป็นหญิงและชาย อายุในช่วง 15-30 ปี
- 1.3.6 โปรแกรมทำงานอยู่บนเครื่องคอมพิวเตอร์ Pentium II หน่วยความจำ มากกว่า 32 MB ระบบปฏิบัติการ Microsoft Windows 98
- 1.3.7 ระบบเสียง ใช้ Sound Card ที่เข้ากันได้กับ Sound Buster 16 ของบริษัท Creative
- 1.3.8 ระดับสัญญาณรบกวนต้องไม่เกิน 15 dB
- 1.3.9 การพูดในการฝึก และใช้งานโปรแกรมจะเป็นการพูดที่มีช่วงว่างระหว่างคำไม่น้อยกว่า 1 วินาที
- 1.3.10 การพูดในการฝึก และใช้งานโปรแกรมจะต้องมีระดับเสียงผ่านเข้าสู่ Sound Card ไม่น้อยกว่า 55 dB

1.4 ประโยชน์ที่จะได้รับ

- 1.4.1 เพื่อเปรียบเทียบผลของการใช้ เทคนิค Short time spectrum ในการวิเคราะห์สัญญาณ กับเทคนิคอื่นๆ ที่มีการวิจัยมาก่อน
- 1.4.2 เป็นแนวทางในการพัฒนาระบบรู้จำเสียงพูดภาษาไทยแบบต่อเนื่อง
- 1.4.3 ได้แนวทางในการพัฒนา ระบบ Software และ Hardware เพื่อนำไปพัฒนาระบบอื่นๆ ที่ต้องการใช้เสียงพูดในการควบคุม