

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

สรุปผลการวิจัย

งานวิจัยนี้เป็นการศึกษาเกี่ยวกับระบบการรู้จำคำไทยหลายพยางค์ บนพื้นฐานของแบบจำลองฮิดเดน มาร์คอฟร่วมกับขั้นตอนวิธีการควอนไทซ์แบบเวกเตอร์และวิธีการหาจุดสิ้นสุดเสียงพูด ระบบได้รับการออกแบบให้ทำการฝึกฝนเพื่อหาต้นแบบอ้างอิง และทำการทดสอบหาอัตราความแม่นยำในการรู้จำคำพูดภาษาไทย จากชุดคำศัพท์ภาษาไทย 70 คำ โดยแบ่งเป็นชุดคำศัพท์พยางค์เดียว ชุดคำศัพท์สองพยางค์ ชุดคำศัพท์สามพยางค์ชุดละ 20 คำ และชุดคำศัพท์ตัวเลขศูนย์ถึงเก้าจำนวน 10 คำ ดังมีรายละเอียดในภาคผนวก ก เสียงพูดคำศัพท์ภาษาไทยที่บันทึกไว้แบ่งเป็น 3 ชุด ได้แก่ ชุดเสียงพูดเพื่อฝึกฝน ชุดเสียงพูดเพื่อฝึกฝนพูดครั้งที่สอง และชุดเสียงพูดเพื่อทดสอบ

ในขั้นตอนการวิเคราะห์หาจุดสิ้นสุดเสียงพูดและจุดสิ้นสุดพยางค์เพื่อหาจำนวนพยางค์นั้น จะอาศัยเส้นระดับพลังงานในการวิเคราะห์หาจุดเริ่มต้นและจุดสิ้นสุดพยางค์ โดยการกำหนดระดับจุดเริ่มเปลี่ยนที่มีค่าร้อยละ 5 และร้อยละ 10 ของค่าความเข้มเสียงสูงสุดของเสียงพูด เพื่อเป็นระดับอ้างอิงในการวิเคราะห์หาจุดเริ่มต้นและจุดสิ้นสุดพยางค์

ในขั้นตอนการควอนไทซ์แบบเวกเตอร์จะอาศัยใช้ชุดรหัสขนาด 256 เวกเตอร์จำนวน 70 ชุดสำหรับคำศัพท์แต่ละคำที่ได้มาจากการสร้างและฝึกฝนชุดรหัส เพื่อเป็นตัวแทนเวกเตอร์ของสัมประสิทธิ์ของการประมาณพหุระเชิงเส้นที่ 10 ลำดับ โดยการเปรียบเทียบและแทนที่ด้วยค่าของชุดรหัสซึ่งเวกเตอร์ของชุดรหัสนั้นซึ่งให้ค่าความผิดพลาดกำลังสองเฉลี่ยมีค่าต่ำที่สุด

ในขั้นตอนการสร้างต้นแบบของเสียงพูดสำหรับคำศัพท์แต่ละคำโดยใช้แบบจำลองฮิดเดน มาร์คอฟเป็นการจัดเก็บชุดพารามิเตอร์ของแบบจำลองสำหรับคำศัพท์แต่ละคำ เพื่อใช้ในการรู้จำด้วยขั้นตอนวิธีการ Viterbi ผลลัพธ์ที่ได้จากขั้นตอนการรู้จำจะเป็นค่าความน่าจะเป็นของคำศัพท์แต่ละคำ โดยคำศัพท์ที่รู้จำได้จะเป็นคำศัพท์ที่มีค่าความน่าจะเป็นสูงที่สุด

ผลการทดสอบอัตราการรู้จำแบบขึ้นกับผู้พูดของชุดเสียงพูดเพื่อฝึกฝนมีค่าเฉลี่ยร้อยละ 100.000 ชุดเสียงพูดเพื่อฝึกฝนพูดครั้งที่สองมีค่าเฉลี่ยร้อยละ 95.000 และผลการทดสอบอัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของชุดเสียงพูดเพื่อทดสอบมีค่าเฉลี่ยร้อยละ 89.906

ในงานวิจัยนี้ได้ทำการเปลี่ยนแปลงค่าพารามิเตอร์ของระบบ เพื่อศึกษาถึงผลกระทบที่มีต่ออัตราการรู้จำของระบบ ค่าพารามิเตอร์ของระบบที่ทำการเปลี่ยนแปลงได้แก่ จำนวนเสียงพูดเพื่อฝึกฝน ขนาดชุดรหัส

และจำนวนสถานะ ในการเปรียบเทียบจะแบ่งออกเป็น 3 ชุด โดยจะทำการเปรียบเทียบเมื่อเพิ่มขนาดชุดรหัสทำการเปรียบเทียบเมื่อเพิ่มจำนวนสถานะ และทำการเปรียบเทียบเมื่อเพิ่มจำนวนเสียงพูดเพื่อฝึกฝนตามลำดับ

ผลการเปรียบเทียบเมื่อเพิ่มขนาดของชุดรหัสให้สูงขึ้น จะส่งผลให้อัตราการรู้จำมีค่าเพิ่มสูงขึ้นอย่างมาก ผลการเปรียบเทียบเมื่อเพิ่มจำนวนสถานะของแบบจำลองฮิดเดน มาร์คอฟเป็น 10 15 20 และ 25 สถานะจะช่วยให้อัตราการรู้จำเพิ่มสูงขึ้นด้วย แต่อย่างไรก็ตามยิ่งจำนวนสถานะมากขึ้นเท่าใดเวลาที่ใช้ในการฝึกฝนและการรู้จำก็จะนานมากขึ้นเท่านั้น และอัตราการรู้จำก็เพิ่มขึ้นเพียงเล็กน้อยเท่านั้น ดังนั้นในงานวิจัยนี้จึงเลือกใช้จำนวนสถานะ 15 สถานะเป็นหลัก ผลการเปรียบเทียบเมื่อเพิ่มจำนวนเสียงพูดเพื่อฝึกฝนให้สูงขึ้น จะช่วยให้้อตราการรู้จำเพิ่มสูงขึ้นเพียงเล็กน้อย

ความผิดพลาดในการรู้จำคำพูดภาษาไทย อาจเกิดขึ้นจากความผิดพลาดในการวิเคราะห์หาจำนวนพยางค์และจุดสิ้นสุดพยางค์หรือจุดสิ้นสุดเสียงพูด และลำดับค่าชุดรหัสของเสียงพูดซึ่งมาจากชุดรหัสต่างกัน แต่มีลำดับค่าชุดรหัสคล้ายคลึงกันก็อาจทำให้เกิดการรู้จำข้ามกลุ่มขึ้น ซึ่งเป็นสาเหตุหนึ่งที่ทำให้เกิดการรู้จำผิดพลาดขึ้น

ข้อเสนอแนะเพื่อการแก้ไขข้อผิดพลาดในการรู้จำ

ข้อเสนอแนะสำหรับแก้ไขข้อผิดพลาดของการรู้จำคำไทยหลายพยางค์โดยอาศัยแบบจำลองฮิดเดน มาร์คอฟมีดังนี้

1) ในการวิเคราะห์และวัดค่าลักษณะสำคัญของเสียงพูดนั้น ควรทำการวิเคราะห์ในเชิงความถี่นอกเหนือไปจากการวิเคราะห์ในเชิงเวลา เนื่องจากลักษณะเฉพาะของเสียงพูดภาษาไทยซึ่งเป็นภาษาที่มีวรรณยุกต์หรือมีเสียงที่มีความถี่สูงกว่า จึงจะเป็นต้องอาศัยลักษณะสมบัติทางความถี่ในการแยกแยะความแตกต่างของเสียงพูดด้วยลักษณะเฉพาะตัวของแต่ละเสียง

2) กรรรมวิธีในการหาจุดสิ้นสุดเสียงพูดและจุดสิ้นสุดพยางค์นั้น ควรปรับปรุงเพิ่มเติมวิธีการอื่นๆ เสริมเข้าไปเพื่อช่วยให้สามารถหาจุดสิ้นสุดเสียงพูดและจุดสิ้นสุดพยางค์ได้มีประสิทธิภาพมากยิ่งขึ้น เช่น การคำนวณหาค่า Pitch Period ของเสียงพูด Zero-Crossing เป็นต้น

3) การขยายขนาดของชุดรหัสจากขนาด 128 เวกเตอร์ไปเป็น 256 เวกเตอร์หรือมากกว่า จะช่วยให้ค่าความผิดพลาดในการควอนไทซ์แบบเวกเตอร์มีค่าลดลง เนื่องจากขนาดของชุดรหัสที่ใหญ่ขึ้นจะสามารถครอบคลุมความหลากหลายของข้อมูลเวกเตอร์สัมประสิทธิ์การประมาณพันธะเชิงเส้นได้มากขึ้น

4) แบบจำลองฮิดเดน มาร์คอฟจำเป็นต้องอาศัยข้อมูลเสียงพูดเพื่อการฝึกฝนเป็นจำนวนมาก เพื่อเป็นการเพิ่มประสิทธิภาพในการรู้จำจึงควรเพิ่มจำนวนเสียงพูดให้มากยิ่งขึ้น หรือนำขั้นตอนวิธีการอื่นเข้ามาประยุกต์ใช้งานร่วมด้วยเพื่อเสริมประสิทธิภาพในการรู้จำของระบบให้ดียิ่งขึ้น

5) ในการบันทึกเสียงพูดคำศัพท์ภาษาไทยแต่ละคำ ผู้แปลงเสียงจำเป็นต้องอ่านออกเสียงคำศัพท์แต่ละคำอย่างชัดเจนและถูกต้องตามหลักภาษาไทย โดยเฉพาะคำควบกล้ำและเสียงพยัญชนะ “ร” และ “ล”

6) ชุดคำศัพท์ที่ใช้ในการบันทึกเสียง ควรปรับปรุงเพิ่มเติมจำนวนคำศัพท์ให้มีความหลากหลายมากยิ่งขึ้น หรืออาจเพิ่มเติมคำศัพท์ที่ใช้เฉพาะงานก็ได้

7) ในระบบการรู้จำคำพูดภาษาไทยควรเพิ่มเติมพื้นฐานความรู้ทางด้านภาษาศาสตร์ไว้ในระบบ เพื่อช่วยเสริมประสิทธิภาพในการรู้จำให้ดียิ่งขึ้น

ข้อเสนอแนะเพื่อการประยุกต์ใช้งาน

การประยุกต์ใช้งานระบบการรู้จำคำไทยหลายพยางค์ สามารถนำไปประยุกต์ใช้งานได้มากมาย ตัวอย่างเช่น การสั่งงานเครื่องคอมพิวเตอร์ด้วยเสียงพูดภาษาไทย ระบบต่อหมายเลขโทรศัพท์ด้วยเสียงพูด เป็นต้น สำหรับการสั่งงานเครื่องคอมพิวเตอร์ด้วยเสียงพูดภาษาอังกฤษนั้นมีความก้าวหน้าไปมากโดยมีฐานข้อมูลคำศัพท์นับแสนคำ สำหรับใช้ในการรู้จำคำโดด คำต่อเนื่องและเสียงพูดต่อเนื่องได้เป็นอย่างดี แต่สำหรับเสียงพูดภาษาไทยยังคงอยู่ในช่วงเริ่มต้นเท่านั้นและยังต้องการการพัฒนาอีกมาก

สำหรับการประยุกต์ใช้งานเฉพาะอย่าง ตัวอย่างเช่น ระบบต่อหมายเลขโทรศัพท์ด้วยเสียงพูดภาษาไทยก็สามารถนำไปประยุกต์ใช้งานได้อย่างมีประสิทธิภาพ ระบบเช่นนี้เป็นการรู้จำคำโดดพยางค์เดียว และจำนวนคำศัพท์จำกัดแต่เพียงตัวเลขศูนย์ถึงเก้า ซึ่งในปัจจุบันสามารถรู้จำได้แต่เพียงเสียงพูดตัวเลขภาษาอังกฤษเท่านั้น จึงจำเป็นต้องมีการพัฒนาวิธีการในการรู้จำคำพูดตัวเลขภาษาไทยเพิ่มเติม

สำหรับการพัฒนาระบบการรู้จำเสียงพูดบนเครื่องไมโครคอมพิวเตอร์ส่วนบุคคลด้วยระบบปฏิบัติการ Microsoft Windows 95 และ Microsoft Windows NT นั้น ได้มีการแจกจ่ายชุดพัฒนาระบบการรู้จำเสียงพูดให้แก่ผู้พัฒนาระบบและผู้พัฒนาโปรแกรมประยุกต์ ซึ่งช่วยให้เกิดการตื่นตัวในการพัฒนาทางด้านระบบการรู้จำเสียงพูด และในอนาคตจะผนวกรวมระบบการรู้จำเสียงพูดเข้าไว้ภายในระบบปฏิบัติการเพื่อให้เป็นมาตรฐานในการปฏิสัมพันธ์กับเครื่องคอมพิวเตอร์ นอกเหนือไปจากวิธีการเดิมที่ใช้งานอยู่ในปัจจุบัน โดยจะช่วยให้การประยุกต์ใช้งานระบบรู้จำเสียงพูดเป็นรูปเป็นร่างมากยิ่งขึ้น