บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

1. งานวิจัยนี้เป็นการพัฒนาระบบรู้จำหน่วยเสียงสระภาษาไทยบนพื้นฐานของโครงข่ายประสาทเทียม ระบบได้รับการออกแบบให้ทำการฝึกฝนเพื่อหาต้นแบบอ้างอิงและทำการทดสอบอัตราการรู้จำหน่วยเสียงสระเดี่ยว 9 หน่วยเสียง นอกจากนี้ยังสร้างแบบจำลองช่วงเวลาของเสียงสระเพื่อจำแนกสระเสียงสั้น-ยาวสำหรับรู้จำหน่วยเสียงสระเดี่ยวทั้งหมดภายหลังจากการใช้โครงข่ายประสาทเทียมในการรู้จำหน่วยเสียงสระเดี่ยว 9 หน่วยเสียง

2. ข้อมูลเสียงพูดแบ่งเป็น 2 ชุด ได้แก่ ข้อมูลเสียงพูดชุดฝึกฝนประกอบด้วยเสียงพูดสองครั้งจำนวน 30 คนและข้อมูลเสียงพูดชุดทดสอบประกอบด้วยเสียงพูดสองครั้งของผู้พูดที่ไม่ได้อยู่ในชุดฝึกฝนจำนวน 20 คน

3. ในขั้นตอนกรรมวิธีหาจุดสิ้นสุดเสียงพูดและพยางค์เพื่อหาขอบเขตของหน่วยเสียงสระที่ต้องการรู้จำจะอาศัยเส้นระดับพลังงานในการวิเคราะห์จุดสิ้นสุดเสียงพูดและพยางค์ โดยการกำหนดระดับจุดเริ่มเปลี่ยนที่มีค่าร้อยละ 10 ของค่าพลังงานสูงสุดของเสียงพูด นอกจากการใช้เส้นระดับพลังงานในการวิเคราะห์จุดสิ้นสุดเสียงพูดและพยางค์แล้ว ในงานวิจัยนี้ยังได้ใช้ความถี่มูลฐานเพื่อให้ได้ความแม่นยำในการหาขอบเขตของหน่วยเสียงสระในกรณีที่เสียงพยัญชนะต้นเป็นเสียงพ่นลมซึ่งจะมีค่าพลังงานมากกว่าระดับจุดเริ่มเปลี่ยน

4. ในขั้นตอนการสกัดค่าลักษณะสำคัญจะใช้ค่าลักษณะสำคัญได้แก่ สัมประสิทธิ์การประมาณพันธะเชิงเส้น สัมประสิทธิ์ cepstral ความถี่ฟอร์แมนท์และความเข้มของสเปกตรัม โดยใช้ค่าลักษณะสำคัญจำนวน 10 กรอบข้อมูลเสียงพูด

5. ในขั้นตอนการสร้างต้นแบบของหน่วยเสียงสระจะใช้โครงข่ายประสาทเทียมสำหรับฝึกฝนค่าลักษณะสำคัญเพื่อให้ได้ต้นแบบของหน่วยเสียงสระ ในงานวิจัยนี้ได้ทดลองใช้โครงข่ายประสาทเทียมสร้างต้นแบบของหน่วยเสียงสระหน่วยเสียงสระเดี่ยวทั้งหมด 9 หน่วยเสียงและทดลองใช้โครงข่ายประสาทเทียมในการแบ่งกลุ่มหน่วยเสียงสระและใช้โครงข่ายประสาทเทียมย่อยในการสร้างต้นแบบของหน่วยเสียงสระที่ได้จากการแบ่งกลุ่ม

6. ในงานวิจัยนี้ได้ทำการเปลี่ยนแปลงค่าพารามิเตอร์ของระบบเพื่อศึกษาถึงผลกระทบที่มีต่ออัตราการรู้จำของระบบ ค่าพารามิเตอร์ของระบบที่ทำการเปลี่ยนแปลงได้แก่ จำนวนเสียงผู้พูดฝึกฝน ค่าลักษณะสำคัญ และจำนวนโครงข่ายประสาทเทียม

7. โครงข่ายประสาทเทียมที่ใช้ในงานวิจัยนี้เป็นโครงข่ายประสาทเทียมแบบ MLP ที่มีการฝึกฝนด้วยขั้นตอนวิธีการส่งค่าย้อนกลับ (Backpropagation Algorithm) แบบ Pattern Mode มีจำนวนชั้นซ่อนตัวหนึ่งชั้น จำนวนโนดในชั้นซ่อนตัวหาได้จากการทดลองซึ่งพบว่าจำนวนโนดในชั้นซ่อนตัว 150 โนดมีความเหมาะ

สมที่สุดคือให้อัตราการรู้จำสูงที่สุดสำหรับค่าลักษณะสำคัญเกือบทุกค่า จำนวนโนดที่มากกว่านี้ (200 โนด และ 250 โนด) จะให้อัตราการรู้จำที่สูงขึ้นเพียงเล็กน้อยเท่านั้นหรือลดลงในบางกรณี

8. จำนวนผู้พูดฝึกฝนมีผลต่ออัตราการรู้จำ จำนวนผู้พูดที่เพิ่มขึ้นทำให้ระบบมีอัตราการรู้จำที่สูงขึ้น เนื่องจากระบบสามารถสร้างต้นแบบของหน่วยเสียงให้มีความครอบคลุมความแปรผันของเสียงพูดได้ดีขึ้น

9. อัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบของระบบที่มีการแบ่งกลุ่มเสียงสระตามความถี่ฟอร์แมนท์ที่ 1 และความถี่ฟอร์แมนท์ที่ 2 มีค่าเท่ากับร้อยละ 90.34 และร้อยละ 88.22 ตามลำดับ ซึ่งมีค่าสูงกว่าอัตราการรู้จำของระบบที่ใช้โครงข่ายประสาทเทียมเดียวในการรู้จำหน่วยเสียงสระที่มีค่าเท่ากับร้อยละ 85.92

10. ผลการทดสอบอัตราการรู้จำแบบไม่ขึ้นกับผู้พูดของข้อมูลทดสอบมีค่าสูงสุดร้อยละ 90.34 เมื่อใช้ค่าลักษณะสำคัญเป็นสัมประสิทธิ์ cepstral 16 อันดับที่คำนวณจากสัมประสิทธิ์การประมาณพันธะเชิงเส้น 16 อันดับและใช้จำนวนผู้พูดฝึกฝน 30 คน โดยใช้โครงข่ายประสาทเทียมในการแบ่งกลุ่มหน่วยเสียงสระตามความถี่ฟอร์แมนท์ที่ 1 และใช้โครงข่ายประสาทเทียมย่อยในการรู้จำหน่วยเสียงสระที่ได้จากการแบ่งกลุ่ม

11. การใช้โครงข่ายประสาทเทียมในการแบ่งกลุ่มหน่วยเสียงสระตามความถี่ฟอร์แมนท์ และใช้โครงข่ายประสาทเทียมย่อยในการรู้จำหน่วยเสียงสระที่ได้จากการแบ่งกลุ่มจะใช้เวลาในการฝึกฝนและเวลาในการรู้จำที่มากกว่าการใช้โครงข่ายประสาทเทียมโครงข่ายเดียวในการรู้จำหน่วยเสียงสระทั้งหมด ในการฝึกฝนระบบที่มีการใช้โครงข่ายประสาทเทียมย่อยจะต้องทำการฝึกฝนโครงข่ายประสาทเทียมทั้งหมด 4 โครงข่าย โดยแบ่งเป็นโครงข่ายสำหรับแบ่งกลุ่มเสียงสระ 1 โครงข่ายและโครงข่ายรู้จำหน่วยเสียงสระอีก 3 โครงข่าย ส่วนในขั้นตอนการรู้จำจะต้องผ่านโครงข่ายประสาทเทียม 2 ครั้งคือครั้งแรกจะต้องผ่านโครงข่ายสำหรับแบ่งกลุ่มเสียงสระและจะต้องผ่านโครงข่ายรู้จำหน่วยเสียงสระกลุ่มย่อยอีกครั้งหนึ่ง

12. ในงานวิจัยนี้ทำการสร้างแบบจำลองช่วงเวลาความยาวของสระเสียงสั้นและสระเสียงยาวให้มีการแจกแจงแบบเกาส์ (Gaussian Distribution) เปรียบเทียบแบบจำลองกับแผนภาพฮิสโตแกรมการแจกแจงช่วงเวลาของเสียงสระพบว่าแบบจำลองทางสถิติช่วงเวลาความยาวของสระเสียงสั้น-ยาวเป็นที่น่าเสนอเป็นแบบจำลองที่มีความสอดคล้องกับการแจกแจงของช่วงเวลาความยาวของสระเสียงสั้น-ยาว

5.2 ข้อเสนอแนะ

1. ควรทดลองใช้ค่าลักษณะสำคัญอื่น เช่น Perceptual Linear Prediction (PLP) Mel Frequency Cepstrum Coefficient (MFCC) หรือค่าลักษณะสำคัญที่มากกว่า 1 ค่าเช่น สัมประสิทธิ์การประมาณพันธะเชิงเส้นและสัมประสิทธิ์ cepstral หรือค่าลักษณะสำคัญและค่าอนุพันธ์เช่น สัมประสิทธิ์ cepstral และค่าอนุพันธ์ของสัมประสิทธิ์ cepstral (cepstral derivative) อาจทำให้อัตราการรู้จำสูงขึ้น

2. การใช้โครงข่ายประสาทเทียมในการแบ่งกลุ่มหน่วยเสียงสระตามความถี่ฟอร์แมนท์ และใช้โครงข่ายประสาทเทียมย่อยในการรู้จำหน่วยเสียงสระที่ได้จากการแบ่งกลุ่ม จำเป็นที่จะต้องปรับปรุงการแบ่งกลุ่มหน่วยเสียงสระให้มีประสิทธิภาพสูงที่สุดเพื่อทำให้อัตราการรู้จำโดยรวมของระบบมีค่าสูง

3. การสร้างแบบจำลองช่วงเวลาของเสียงสระเพื่อจำแนกสระเสียงสั้น-ยาวให้มีการแจกแจงเป็นแบบเกาส์ไม่เพียงพอที่จะจำแนกสระเสียงสั้น-ยาวได้อย่างมีประสิทธิภาพ จำเป็นต้องอาศัยความรู้ทางด้านภาษาศาสตร์มาช่วยวิเคราะห์เช่น เสียงวรรณยุกต์ หน่วยเสียงพยัญชนะข้างเคียง เป็นต้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย