



6.1 สรุป

ระบบการรับรู้เสียงพูดที่ทำการวิจัยนี้ ได้ทำการทดสอบโดยการคัดเลือกคำที่ใช้ทดสอบออกเป็น 2 กลุ่ม คือในกลุ่มที่ 1 จะคัดเลือกโดยใช้หลักเกณฑ์การออกเสียงทางด้านภาษาศาสตร์ จำนวน 15 คำ และกลุ่มที่ 2 ได้จากการคัดเลือกมาจากคำที่ใช้มากที่สุดในชีวิตประจำวัน จำนวน 10 คำ แบบอ้างอิงจำนวน 10 แบบในแต่ละคำจะถูกคัดเลือกมาจากนักศึกษาชายทั้งหมด 30 คน การทดสอบทำโดยนักศึกษาชายอีก 5 คน โดยให้แต่ละคนพูดคนละ 10 ครั้ง ผลการทดสอบปรากฏว่า คำในกลุ่มที่ 1 มีอัตราความถูกต้องในการรับรู้ร้อยละ 85.8 และคำในกลุ่มที่ 2 มีอัตราความถูกต้องในการรับรู้ร้อยละ 94.3

จากผลการทดสอบสามารถสรุปได้ดังนี้

ก. การคัดเลือกคำที่ใช้ในการรับรู้ ในการคัดเลือกคำที่ใช้ในการรับรู้ จะมีผลโดยตรงต่ออัตราความถูกต้องในการรับรู้ เนื่องจากสำหรับกลุ่มคำที่มีการออกเสียงคล้ายคลึงกัน จะมีอัตราความถูกต้องในการรับรู้ต่ำกว่ากลุ่มคำที่มีการออกเสียงแตกต่างกันอย่างชัดเจน ดังจะเห็นได้จากผลการทดลอง คำที่ใช้ในทดสอบในกลุ่มที่ 1 จะมีอัตราความถูกต้องในการรับรู้ต่ำกว่าคำในกลุ่มที่ 2

ข. การคัดเลือกแบบอ้างอิง ในการคัดเลือกแบบอ้างอิงที่ใช้ในการเปรียบเทียบนั้น จำเป็นต้องมีจำนวน และการกระจายมากเพียงพอ เพราะเมื่อทำการคัดเลือก จะได้แบบอ้างอิงที่สามารถครอบคลุมความแตกต่างระหว่างผู้พูด ได้มากที่สุด ซึ่งจะช่วยให้ผลการรับรู้มีอัตราความถูกต้องสูงขึ้นด้วย

ค. การรับรู้ การรับรู้เป็นการคำนวณหาค่าระยะทางระหว่างแบบทดสอบกับแบบอ้างอิง พร้อมทั้งทำการตัดสินใจหาผลลัพธ์ ซึ่งสามารถแบ่งการประมวลผลออกเป็น

1. การเปรียบเทียบแบบไดนามิกโปรแกรมมิ่ง เป็นวิธีในการหาค่าความแตกต่างของรูปแบบสองรูปแบบ ที่มีความสามารถในการปรับตัวตามแนวแกนเวลา การคำนวณระยะทางของรูปแบบทั้งสองจะ ได้จากการเปรียบเทียบค่าสัมประสิทธิ์การสะท้อนของรูปแบบทั้งสอง

2. กฎการตัดสินใจ ในการตัดสินใจของระบบรับรู้เสียงพูดแบบต่างบุคคล ไม่สามารถใช้ค่าระยะทางที่น้อยที่สุดระหว่างแบบอ้างอิงกับแบบทดสอบมาตัดสินใจแทน เนื่องจากความแปรปรวนอันเกิดจากผู้พูดต่าง ๆ กันอาจทำให้แบบอ้างอิงที่มีระยะทางน้อยที่สุด เป็นผลลัพธ์ที่ไม่ถูกต้องได้ จึงได้นำกฎ K-nearest Neighbour (KNN) มาใช้โดยหาแบบอ้างอิงที่มีระยะทางน้อยที่สุดของแต่ละคำจำนวน K แบบ แล้วนำค่าเฉลี่ยของระยะทางมาใช้ในการตัดสินใจ ซึ่งทำให้อัตราความถูกต้องในการรับรู้สูงขึ้น

3. การกำหนดค่าความแตกต่างที่ยอมรับได้ (Reject Value) เพื่อใช้ลดความผิดพลาดในการรับรู้ อันเนื่องมาจากผลลัพธ์ที่ได้มาจากการตัดสินใจมีค่าใกล้เคียงกันมาก ในกรณีที่ระยะถึงผลลัพธ์สามารถทำให้เกิดความผิดพลาดได้ง่าย ในการกำหนดค่านี้จำเป็นต้องอาศัยการทดลองเพื่อหาค่าที่เหมาะสมโดยยึดเอาอัตราความถูกต้องเป็นหลัก

4. การทำ Post Processing จะช่วยเพิ่มความถูกต้องในการรับรู้ สำหรับผลการรับรู้ที่มีความสับสนเกิดขึ้นบ่อย เช่นจากการทดลองคำว่า "ปาก" กับ "มาก" เป็นต้น จึงทำการหาลักษณะ (Feature) เพื่อจำแนกความแตกต่างของคำทั้งสองนี้ ในที่นี้ใช้ค่า Zero Crossing เป็นลักษณะที่ใช้ในการจำแนก

6.2 ข้อเสนอแนะ

ก. ทำการแยกกลุ่มย่อย (Classify) ของข้อมูล ในระบบการรับรู้เสียงพูดที่พัฒนาขึ้น จะทำการเปรียบเทียบแบบทดสอบกับแบบอ้างอิงทั้งหมด ถ้ามีการหาลักษณะ (Feature) ของคำทั้งหมด แล้วนำมาทำการแยกกลุ่มย่อยก่อน แบบอ้างอิงที่เหลือในการเปรียบเทียบจะลดลง อัตราความถูกต้องในการรับรู้จะสูงขึ้น


ข. ปรับปรุงความยืดหยุ่นในการคัดเลือกแบบอ้างอิง ในระบบการคัดเลือกแบบอ้างอิงที่พัฒนาขึ้น จะทำการกำหนดจำนวนแบบอ้างอิงในแต่ละคำไว้แน่นอน ซึ่งอาจจะมากหรือไม่เพียงพอที่จะเป็นตัวแทนของแบบอ้างอิงทั้งหมดได้ ดังนั้นการนำเทคนิคการคัดเลือกแบบอ้างอิง โดยการไม่กำหนดจำนวนแบบอ้างอิง แต่กำหนดค่าความแตกต่างของระยะทางที่ยอมรับได้ จะทำให้ได้แบบอ้างอิงที่ดีขึ้น

ค. ในการเก็บเสียงเพื่อนำไปสร้างเป็นแบบอ้างอิง หรือแบบทดสอบ จำเป็นต้องมีเครื่องมือที่มีความแม่นยำสูง และมีสถานที่ที่เหมาะสม เพื่อลดผลของสภาพแวดล้อมที่มีต่อความถูกต้องในการรับรู้

6.3 ข้อเสนอแนะในการประยุกต์ใช้งาน

ก. ใช้เชื่อมต่อกับโปรแกรมประยุกต์ ในการเชื่อมต่อระบบการรับรู้เสียงพูดกับโปรแกรมประยุกต์ต่าง ๆ จะช่วยให้โปรแกรมประยุกต์ใช้ประโยชน์จากระบบการรับรู้เสียงพูดได้เต็มที่

ข. ใช้พัฒนาเป็นอุปกรณ์ใช้งานที่สามารถรับคำสั่งเป็นเสียงพูดได้ เช่น โทรศัพท์อัตโนมัติ เครื่องถ่ายเอกสาร เป็นต้น



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย