

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

ในบทนี้ จะกล่าวถึงผลสรุปการวิจัย และข้อเสนอแนะในการพัฒนางานวิจัยต่อไปในอนาคต

5.1 สรุปผลการวิจัย

จากผลการวิจัย ซึ่งแบ่งวิธีการรู้จำตามลักษณะสำคัญของเสียงที่ใช้ในการรู้จำ สามารถสรุปผลได้ดังนี้

1. ขั้นตอนประมวลผลก่อนการรู้จำที่ใช้ มีเพียงการคำนวณพลังงาน และคาบการสั้นของเสียง เพื่อใช้ในการหาจุดสิ้นสุดของเสียง ซึ่งเป็นวิธีที่แม่นยำกว่าการใช้การคำนวณพลังงานเพียงอย่างเดียว

2. จำนวนชุดรหัสที่ใช้ในการรู้จำคำศัพท์เพียงคำเดียวและคำศัพท์ตัวเลขสามารถใช้ได้ทั้งแบบ 128 ชุดรหัส และ 256 ชุดรหัส

3. จำนวนสถานะของแบบจำลองฮิดเดน มาร์คอฟสามารถเลือกใช้เพียง 3 สถานะได้ แต่ถ้าเพิ่มจำนวนคำศัพท์และเพิ่มชุดคำศัพท์ที่มีจำนวนพยางค์มากขึ้น จำนวนสถานะต้องเพิ่มขึ้นด้วย

4. จำนวนเสียงพูดเพื่อฝึกฝนควรมากกว่า 20 คน ในการวิจัยใช้จำนวนเสียงพูดเพื่อฝึกฝน 40 คนให้อัตราการรู้จำที่สูงเพิ่มขึ้นมาก

5. ในการรู้จำเสียงพูดของชุดเสียงพูดเพื่อฝึกฝนจะให้้อัตราการรู้จำสูงมาก ในทุกวิธี ดังนั้นสามารถนำวิธีการรู้จำโดยอาศัยลักษณะสำคัญข้างต้น ใช้งานในงานประยุกต์การรู้จำเสียงพูดของชุดฝึกฝนที่ต้องการการคำนวณไม่มากนักได้

6. การคำนวณค่า Isf ไม่ควรคำนวณค่าโดยประมาณ โดยการละ L2 และ L3 เพราะจำทำให้อัตราการรู้จำลดลงอย่างมาก

7. การคำนวณค่า Isf โดยผ่านตัวกรองมัธยฐานไม่ช่วยให้การรู้จำดีขึ้นอย่างชัดเจน จึงไม่จำเป็นต้องนำมาคำนวณ

8. จากผลการทดลองวิธีที่ดีที่สุดในการรู้จำมี 2 ทางเลือก คือ ทางเลือกที่หนึ่ง ถ้าต้องการการรู้จำเสียงพูดที่มีการคำนวณต่ำที่สุด แต่ยังให้ผลการรู้จำที่ดี แต่อาจไม่ดีเลิศ กับระบบรู้จำที่ไม่มีคำศัพท์มากนัก ควรเลือกวิธีที่ 4 ให้ดัชนีของชุดรหัส G.729 โดยตรง ทางเลือกที่สอง ถ้าต้องการระบบรู้จำที่มีความผิดพลาดต่ำ ซึ่งต้องใช้ในการคำนวณเพิ่มมากขึ้น ควรเลือกวิธีที่ 1 ใช้ LSF โดยตรงเป็นลักษณะสำคัญ

5.2 ข้อเสนอแนะ

1. ในขั้นตอนการหาจุดสิ้นสุดของเสียงพูด สามารถใช้การคำนวณคาบการสั่นของเสียงรวมผนวกกับพลังงานได้ แต่ในมาตรฐานการเข้ารหัส G.729 Annex B จะให้ค่าบิตเพิ่มเติมจากมาตรฐาน G.729 แบบปกติคือ บิตที่บ่งบอกว่าเสียงช่วงนั้นเป็นเสียงเงียบหรือมีเสียงปกติ เพื่อไม่ต้องทำการเข้ารหัสในขณะที่เสียงเงียบ โดยทำการคำนวณจากพลังงาน , จำนวนการตัดผ่านค่าศูนย์ (zero crossing rate), และการเปลี่ยนแปลงของ Isf (line spectral frequency) ซึ่งมีขนาด 1 บิต สามารถใช้บิตนี้หาจุดสิ้นสุดของเสียงพูดได้ โดยไม่ต้องคำนวณใด ๆ ทั้งสิ้นเป็นการลดการคำนวณได้อย่างมาก

2. ถ้าต้องการอัตราการรู้จำที่สูงขึ้นและสามารถรู้จำคำศัพท์ได้หลากหลายขึ้น อาจต้องเพิ่มการคำนวณ ให้ซับซ้อนขึ้น โดยเฉพาะการรู้จำคำพูดภาษาไทยจะมีเสียงสูงต่ำของวรรณยุกต์เกี่ยวข้องด้วย อาจใช้คาบการสั่นของเสียงช่วยแยกเสียงเป็นกลุ่ม ๆ ก่อน จึงทำการรู้จำ

3. ในงานวิจัยทำการเปรียบเทียบจำนวนเสียงพูดเพื่อฝึกฝนจาก 20 คนเป็น 40 คนซึ่งให้ผลการรู้จำที่สูงขึ้นมาก ดังนั้นถ้าเพิ่มจำนวนเสียงพูดมากขึ้นอีก อาจทำให้อัตราการรู้จำสูงขึ้นอีก

4. ในการเลือกเสียงพูดที่ใช้ในการฝึกฝน ควรเป็นเสียงพูดที่ออกเสียงถูกต้อง เพื่อให้ระบบรู้จำคำที่ถูกต้อง จะมีผลต่ออัตราการรู้จำมาก