บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

1. เมื่อฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วนกับระบบการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูดที่นำเสนอในวิทยานิพนธ์นี้สามารถสรุปได้ดังนี้ อันดับของสัมประสิทธิ์การประมาณพันธะเชิงเส้นที่เหมาะสมเท่ากับ 10 อันดับของสัมประสิทธิ์เซปสตรอลที่เหมาะสมเท่ากับ 12 และอันดับของสัมประสิทธิ์เซปสตรอลบนความถี่เมลที่เหมาะสมเท่ากับ 14 และเมื่อฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการหาคาดหวังที่มากที่สุด อันดับของสัมประสิทธิ์การประมาณพันธะเชิงเส้นที่เหมาะสมเท่ากับ 12 ส่วนอันดับของสัมประสิทธิ์เซปสตรอลและสัมประสิทธิ์เซปสตรอลบนความถี่เมลที่เหมาะสมเท่ากับ 14

2. ลักษณะสำคัญที่ให้อัตราการบ่งชี้ผู้พูดดีที่สุดคือสัมประสิทธิ์เซปสตรอลบนความถี่เมล จากการเปรียบเทียบกับระบบที่ใช้สัมประสิทธิ์การประมาณพันธะเชิงเส้นและสัมประสิทธิ์เซปสตรอลเป็นลักษณะสำคัญเมื่อฝึกฝนต้นแบบอ้างอิงชุดรหัสด้วยขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วน และขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุด

3. ขนาดของต้นแบบอ้างอิงชุดรหัสที่เพิ่มขึ้นโดยรวมทำให้อัตราการบ่งชี้ผู้พูดของระบบมีค่าสูงขึ้น แต่มีบางกรณีที่อัตราการบ่งชี้ผู้พูดมีค่าลดลงเมื่อเพิ่มขนาดของต้นแบบอ้างอิงชุดรหัสให้มากขึ้น เนื่องจากจำนวนเวกเตอร์รหัสที่เพิ่มขึ้นในต้นแบบอ้างอิงชุดรหัสมีผลทำให้ประสิทธิภาพในการควอนไทซ์แบบเวกเตอร์ลดลง ดังนั้นการแทนข้อมูลเสียงพูดด้วยต้นแบบอ้างอิงชุดรหัสเกิดความผิดพลาดมากขึ้นและเป็นสาเหตุทำให้การบ่งชี้ผู้พูดผิดมากขึ้น ทั้งนี้ประสิทธิภาพของการควอนไทซ์แบบเวกเตอร์ขึ้นอยู่กับปริภูมิของข้อมูล ขนาดของต้นแบบอ้างอิงชุดรหัส ขั้นตอนวิธีการสร้างและฝึกฝนต้นแบบอ้างอิงชุดรหัสเป็นสำคัญ

4. การเพิ่มประสิทธิภาพของระบบการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูดสามารถกระทำได้โดยเพิ่มจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟให้มากเพียงพอ เพราะว่าแบบจำลองฮิดเดนมาร์คอฟเป็นแบบจำลองที่เหมาะสมกับสัญญาณเสียงพูดที่มีความยาวมาก เนื่องจากว่ามีการรับข้อมูลขาเข้าเป็นแบบลำดับของข้อมูลที่เรียกว่าลำดับของค่าสังเกตและถ้าลำดับของค่าสังเกตนี้มีจำนวนมากสามารถเพิ่มความแตกต่างของข้อมูลเสียงพูดของผู้พูดแต่ละคนได้ โดยใช้สถานะของแบบจำลองที่มาจำนวนมากขึ้นสามารถแสดงความแตกต่างได้มากขึ้นเช่นกัน

5. การเลือกบทคำพูดที่เหมาะสมกับระบบการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูดนั้นสามารถเพิ่มประสิทธิภาพให้แก่ระบบด้วย จากผลการวิจัยเห็นได้ว่าบทคำพูดที่มีอัตราการบ่งชี้ผู้พูดมากที่สุดในทุกๆ กรณีที่พิจารณาในงานวิจัยนี้คือประโยคที่ 3 ซึ่งเป็นบทคำพูดที่ประกอบขึ้นด้วยตัวเลขดังที่กล่าวไว้ในบทที่ 3 โดยมีการเรียงลำดับตามเสียงขึ้นและตกสลับกันทำให้เกิดความแตกต่างในการออกเสียงของผู้พูดแต่ละคนมาก และเนื่องจากเป็นประโยคของตัวเลขทำให้มีจังหวะการออกเสียงของผู้พูดแต่ละคนแตกต่างกัน ในขณะที่

ประโยคที่ 1 และ 2 ถูกกำหนดขึ้นเพื่อสื่อความหมายในการประยุกต์ใช้งานกับระบบรักษาความปลอดภัย โดยประโยคที่ 1 และ 2 มีอัตราการบ่งชี้ผู้พูดใกล้เคียงกัน แต่ประโยคที่ 1 มีอัตราการบ่งชี้ผู้พูดโดยรวมมากกว่าประโยคที่ 2 เล็กน้อยดังแสดงในตารางที่ 4.15 เพราะว่าประโยคที่ 1 เป็นประโยคที่มีจำนวนพยางค์มากที่สุด (6 พยางค์) ทำให้การออกเสียงของผู้พูดแต่ละคนมีโอกาสแตกต่างกันมากกว่าประโยคที่ 2 (4 พยางค์) แม้ว่าประโยคที่ 3 มีพยางค์น้อยกว่าประโยคที่ 1 แต่จากผลการวิจัยแสดงให้เห็นว่าการเลือกบทคำพูดที่เหมาะสมมีอิทธิพลมากกว่าจำนวนพยางค์ของบทคำพูด

6. การเปรียบเทียบขั้นตอนวิธีการฝึกฝนต้นแบบอ้างอิงชุดรหัสในการทดลองสามารถสรุปตามผลการวิจัยได้ว่าขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุดจะมีประสิทธิภาพดีกว่าขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วนเมื่อมีขนาดของต้นแบบอ้างอิงชุดรหัสมีค่าน้อยๆ แต่ถ้าเพิ่มขนาดของต้นแบบอ้างอิงชุดรหัสให้มากเพียงพอแล้วขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วนจะมีประสิทธิภาพดีกว่า และถ้าคำนึงถึงเวลาที่ใช้ในการฝึกฝนต้นแบบอ้างอิงชุดรหัสของทั้งสองขั้นตอนวิธีการพบว่าขั้นตอนวิธีการแบ่งเฉลี่ย K ส่วนใช้เวลาในการประมวลผลน้อยกว่าขั้นตอนวิธีการหาค่าคาดหวังที่มากที่สุดในทุกๆ กรณีดังแสดงในตารางที่ 5.1 เนื่องจากว่าขั้นตอนวิธีการหาค่าคาดหวังมีการคำนวณแมตริกซ์ความแปรปรวนที่มีขนาดตามมิติในปริภูมิของสัมประสิทธิ์ที่ใช้เป็นลักษณะสำคัญในทุกๆ รอบของการปรับเวกเตอร์รหัสและเมตริกซ์ความแปรปรวนจึงต้องใช้เวลาในการประมวลผลนานมาก

ตารางที่ 5.1 เวลาที่ใช้ในการประมวลผลจากการทดลองปรับเปลี่ยนค่าพารามิเตอร์ของระบบ

กระบวนการ	เวลาที่ใช้ในการฝึกฝน	เวลาที่ใช้ในการทดสอบ
สกัดลักษณะสำคัญ	LPC = CEP = MFCC	
อันดับของลักษณะสำคัญ	16>14>12>10 (ทุกกรณีใช้เวลาไม่ถึงวินาที)	
ขนาดของต้นแบบอ้างอิงชุดรหัส	250>200>150>100>50 เวลาขึ้นอยู่ขั้นตอนวิธีการฝึกฝน	250>200>150>100>50 เวลาขึ้นอยู่ขั้นตอนวิธีการฝึกฝน
ขั้นตอนวิธีการฝึกฝนต้นแบบอ้างอิงชุดรหัส	K-means ใช้เวลาประมาณ 10-30 นาที EM ใช้เวลาประมาณ 1 - 5 ชั่วโมง	K-means ใช้เวลาน้อยกว่า EM
จำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟ	30>25>20>15>10>5	30>25>20>15>10>5

7. ถึงแม้ว่าการเพิ่มขนาดของต้นแบบอ้างอิงชุดรหัสและจำนวนสถานะของแบบจำลองฮิดเดนมาร์คอฟเป็นการเพิ่มอัตราการบ่งชี้ผู้พูดก็ตาม แต่ก็เป็นการเพิ่มเวลาที่ใช้ในการประมวลผลด้วย โดยเฉพาะการเพิ่มจำนวนสถานะของแบบจำลองต้องใช้เวลาในการฝึกฝนและทดสอบระบบมากขึ้นตามจำนวนสถานะที่เพิ่มมากขึ้น ในขณะที่การเพิ่มขนาดของต้นแบบอ้างอิงชุดรหัสเป็นการเพิ่มเวลาในช่วงการฝึกฝนเท่านั้นแต่

เวลาในช่วงทดสอบระบบเพิ่มขึ้นเล็กน้อยมาก เมื่อเปรียบเทียบเวลาที่ใช้ในการเพิ่มขนาดของต้นแบบอ้างอิงชุดรหัสและจำนวนสถานะของแบบจำลองพบว่าเวลาที่ใช้ในการประมวลผลจากจำนวนสถานะที่เพิ่มขึ้นใช้เวลามากกว่าเวลาที่ใช้ในการประมวลผลของระบบจากขนาดที่เพิ่มขึ้นของต้นแบบอ้างอิงชุดรหัสในช่วงเวลาฝึกฝนและทดสอบระบบดังแสดงในตารางที่ 5.2

8. เทคนิคการฝึกฝนต้นแบบอ้างอิงชุดรหัสที่ใช้ในระบบนี้เป็นต้นแบบอ้างอิงชุดรหัสแบบชุดเดี่ยวทำให้ใช้เวลาในการประมวลผลน้อยกว่าต้นแบบอ้างอิงชุดรหัสหลายชุด แต่ก็มีข้อเสียคือการเพิ่มจำนวนผู้พูดเข้าไปในฐานข้อมูลของระบบจะต้องฝึกฝนชุดรหัสใหม่ทุกครั้ง ในขณะที่ต้นแบบอ้างอิงชุดรหัสหลายชุดจะมีการฝึกฝนเฉพาะข้อมูลเสียงพูดของผู้พูดใหม่เท่านั้น

ตารางที่ 5.2 เปรียบเทียบเวลาที่ใช้ในการประมวลผลตามกระบวนการต่างๆ ของระบบ

กระบวนการ	เวลาที่ใช้ในการฝึกฝน	เวลาที่ใช้ในการทดสอบ
ประมวลผลสัญญาณเบื้องต้นและสกัดลักษณะสำคัญ	น้อยมาก (ใช้เวลาไม่ถึงวินาทีต่อหนึ่งเสียง)	
การควอนไทซ์แบบเวกเตอร์	K-means ใช้เวลาน้อยมากในทุกๆ ขนาดของต้นแบบอ้างอิงชุดรหัสที่ปรับเปลี่ยน ส่วน EM ขึ้นอยู่กับขนาดของต้นแบบอ้างอิงชุดรหัสและใช้เวลามากกว่า K-means	
แบบจำลองฮิดเดนมาร์คอฟ	ประมาณ 15-30 นาทีต่อแบบจำลอง ขึ้นกับจำนวนสถานะของแบบจำลอง	ประมาณ 0.1-3 นาทีในการทดสอบ 60 เสียงกับ 12 แบบจำลอง

หมายเหตุ ในการประมวลผลระบบการบ่งชี้ผู้พูดแบบขึ้นกับบทคำพูดนี้ใช้คอมพิวเตอร์ที่มีหน่วยประมวลผลกลางเป็น Pentium II ความถี่ 450 MHz

5.2 ข้อเสนอแนะ

1. ควรเพิ่มขนาดของฐานข้อมูลที่ใช้ในระบบการบ่งชี้ผู้พูด ซึ่งได้แก่ข้อมูลเสียงพูดในชุดฝึกฝนและทดสอบรวมถึงจำนวนของผู้พูดในระบบด้วย อีกทั้งการเก็บตัวอย่างข้อมูลเสียงพูดของผู้พูดแต่ละบุคคลควรกระทำหลายครั้งที่เวลาแตกต่างกันและมากเพียงพอ เพื่อที่สามารถบ่งบอกเอกลักษณ์ของแต่ละบุคคลได้โดยพยายามเก็บความแปรปรวนของผู้พูดให้ได้มากที่สุด

2. ควรใช้เทคนิคในการฝึกฝนต้นแบบอ้างอิงชุดรหัสหลายชุดเพื่อง่ายต่อการเพิ่มจำนวนของผู้พูดในระบบและอาจจะเป็นการเพิ่มประสิทธิภาพในการควอนไทซ์แบบเวกเตอร์ก็เป็นได้

3. ควรมีการวิเคราะห์และใช้ลักษณะสำคัญแทนสัญญาณเสียงพูดหลายๆ ลักษณะสำคัญ เพื่อแสดงถึงรูปลักษณ์ต่างๆ ของแต่ละบุคคล เช่น ค่าพลังงาน ความถี่มูลฐาน และความเร็วในการพูด เป็นต้น และควรมีการจำแนกเพศของผู้พูดก่อนเพื่อเป็นการลดความผิดพลาดในการบ่งชี้ผิดเป็นเพศตรงข้าม