

## บทที่ 5

### สรุปผลการวิจัย

#### ผลการวิจัย

การแบ่งเสียงพูดเป็นเซกเมนต์ประกอบไปด้วยขั้นตอนย่อยสองขั้นตอน คือ ขั้นตอนการตรวจหาขอบเขตของหน่วยเสียงและขั้นตอนการสร้างกราฟของเซกเมนต์ วิทยานิพนธ์นี้นำเสนอวิธีการแบ่งเสียงพูดเป็นเซกเมนต์แบบซัพพอร์ตเวกเตอร์แมชชีน โดยใช้สารสนเทศสวนศาสตร์เพื่อนำไปใช้กับระบบรู้จำเสียงพูดแบบอาศัยเซกเมนต์ โดยเปรียบเทียบประสิทธิภาพกับวิธีการแบ่งเสียงพูดเป็นเซกเมนต์แบบอาศัยเครื่องรู้จำเสียงพูดที่ใช้อยู่เดิม ด้วยความหวังที่ว่าวิธีการแบ่งเสียงพูดเป็นเซกเมนต์ที่เสนอนี้จะสามารถทำงานได้รวดเร็วกว่าและสร้างกราฟของเซกเมนต์ที่มีคุณภาพสูงกว่าได้ ในหัวข้อนี้จะนำเสนอผลการวิจัยซึ่งแบ่งออกเป็นสองส่วนคือ การตรวจหาขอบเขตของหน่วยเสียง และการสร้างกราฟของเซกเมนต์ซึ่งมีรายละเอียดดังนี้

#### 1. การตรวจหาขอบเขตของหน่วยเสียง

กระบวนการตรวจหาขอบเขตของหน่วยเสียงตามวิธีที่เสนอนี้ เริ่มต้นจากการจำแนกหน่วยเสียงออกเป็นประเภทตามลักษณะการออกเสียง โดยอาศัยซัพพอร์ตเวกเตอร์แมชชีนมาช่วยแบ่งแยกสัญลักษณ์ลักษณะการออกเสียงแต่ละแบบ ซึ่งการแบ่งแยกนี้จะใช้ลักษณะสำคัญของเสียงพูดที่ได้จากการใช้ความรู้และสารสนเทศทางสวนศาสตร์ หลังจากนั้นจึงนำผลการจำแนกลักษณะการออกเสียงมาแยกเอาข้อมูลที่เป็นขอบเขตของหน่วยเสียงออกมา แล้วคำนวณค่าคะแนนความมั่นใจให้กับขอบเขตของหน่วยเสียงที่หามาได้ โดยวิธีการนี้มีจุดเด่นสามประการคือ

1. ในการตรวจหาขอบเขตของหน่วยเสียงด้วยวิธีการนี้ การแบ่งแยกสัญลักษณ์ลักษณะการออกเสียงแต่ละแบบจะสามารถใช้ลักษณะสำคัญที่แตกต่างกันได้ ทำให้สามารถเลือกใช้ลักษณะสำคัญที่เหมาะสมกับการจำแนกสัญลักษณ์ลักษณะสำคัญแต่ละประเภทได้อย่างอิสระ
2. การตรวจหาขอบเขตของหน่วยเสียงด้วยวิธีการนี้ สามารถเพิ่มเติมลักษณะสำคัญที่น่าจะเป็นประโยชน์ต่อการแบ่งแยกสัญลักษณ์แต่ละแบบได้โดยง่าย ทำให้สามารถเพิ่มประสิทธิภาพการแบ่งแยกสัญลักษณ์ลักษณะการออกเสียงแต่ละแบบให้มีความแม่นยำมากขึ้นได้ หากอาศัยสารสนเทศสวนศาสตร์มาวิเคราะห์เลือกใช้ลักษณะสำคัญเพิ่ม
3. การตรวจหาขอบเขตของหน่วยเสียงด้วยวิธีการนี้มีการคิดคะแนนให้กับขอบเขตของหน่วยเสียงที่หามาได้ ค่าคะแนนนี้จะแทนระดับความมั่นใจของการที่ขอบเขตของ

หน่วยเสียงจะตรงกันกับขอบเขตของหน่วยเสียงอ้างอิง ซึ่งสามารถนำไปใช้ประโยชน์ต่อในขั้นตอนการรู้จำเสียงพูดเพื่อเพิ่มประสิทธิภาพการรู้จำเสียงพูดต่อไปได้

และจากการทดลองตรวจหาขอบเขตของหน่วยเสียง โดยเมื่อเปรียบเทียบวิธีการตรวจหาขอบเขตของหน่วยเสียงที่เสนอกับวิธีการตรวจหาขอบเขตของหน่วยเสียงแบบอาศัยเครื่องรู้จำเสียงพูด พบว่าวิธีการตรวจหาขอบเขตของหน่วยเสียงที่เสนอนี้มีประสิทธิภาพดีกว่าทั้งในทางด้านคุณภาพของขอบเขตของหน่วยเสียงที่ตรวจหามาได้และความเร็วในการทำงาน โดยที่ระดับความคลาดเคลื่อนยินยอม 20 มิลลิวินาที ขอบเขตของหน่วยเสียงที่ได้จากวิธีการที่เสนอนี้มีเปอร์เซ็นต์ความแม่นยำและความครอบคลุมสูงกว่าขอบเขตของหน่วยเสียงที่ได้จากวิธีการตรวจหาขอบเขตของหน่วยเสียงแบบอาศัยเครื่องรู้จำเสียงพูดที่ทำงานได้เร็วพอๆกัน อยู่ถึง 8.3% และ 5.1% ตามลำดับ และเมื่อเปรียบเทียบกับวิธีการตรวจหาขอบเขตของหน่วยเสียงแบบอาศัยเครื่องรู้จำเสียงพูดที่ยอมเสียเวลาทำงานนานๆเพื่อให้ได้คำตอบที่ครอบคลุมมากขึ้น พบว่าวิธีการตรวจหาขอบเขตของหน่วยเสียงที่เสนอนี้ยังสามารถตรวจหาขอบเขตของหน่วยเสียงได้ความแม่นยำและความครอบคลุมสูงกว่าถึง 9.3% และ 3.52% ตามลำดับ

## 2. การสร้างกราฟของเซกเมนต์

วิทยานิพนธ์นี้ได้เสนอวิธีการสร้างกราฟของเซกเมนต์สามวิธีได้แก่ วิธีการสร้างกราฟของเซกเมนต์แบบเชื่อมต่อทุกขอบเขตของหน่วยเสียงซึ่งเป็นวิธีดั้งเดิมที่ใช้ในการแบ่งเสียงพูดเป็นเซกเมนต์แบบอาศัยเครื่องรู้จำเสียงพูดที่นำมาใช้เปรียบเทียบกับวิธีการสร้างกราฟอีกสองแบบที่เหลือ วิธีการสร้างกราฟของเซกเมนต์แบบอาศัยการเปลี่ยนแปลงสเปกตรัม และวิธีการสร้างกราฟแบบหลายระดับ

วิธีการสร้างกราฟของเซกเมนต์แบบเชื่อมต่อทุกขอบเขตของหน่วยเสียงมีข้อดีอยู่ที่ การสร้างกราฟแบบนี้จะเชื่อมต่อขอบเขตของหน่วยเสียงครบทุก ทำให้กราฟที่ได้มีความครอบคลุมสูงกว่าวิธีอื่นทั้งหมด แต่มีข้อเสียสำคัญคือกราฟนั้นจะมีขนาดใหญ่มาก ทำให้ขั้นตอนการรู้จำเสียงพูดซึ่งนำกราฟนี้ไปใช้จะต้องเสียเวลาในการทำงานมากตามไปด้วย ส่วนวิธีการสร้างกราฟของเซกเมนต์แบบอาศัยการเปลี่ยนแปลงสเปกตรัมมีข้อดีตรงที่สามารถลดขนาดของกราฟได้เป็นจำนวนมากอีกทั้งยังใช้เวลาในการทำงานน้อยมากเมื่อเทียบกับเวลาที่ใช้ในขั้นตอนการตรวจหาขอบเขตของหน่วยเสียง แต่มีข้อเสียคือกราฟที่ได้จะมีความครอบคลุมลดลงมาก ในระหว่างที่วิธีการสร้างกราฟแบบหลายระดับมีข้อได้เปรียบตรงที่สามารถลดขนาดของกราฟได้มากกว่า อีกทั้งยังรักษาระดับความครอบคลุมไว้ได้มากกว่ากราฟที่ได้จากวิธีการสร้างกราฟของเซกเมนต์แบบอาศัยการเปลี่ยนแปลงสเปกตรัม

โดยในการทดลองเปรียบเทียบประสิทธิภาพวิธีการสร้างกราฟของเซกเมนต์แบบอาศัยการเปลี่ยนแปลงสเปกตรัมกับวิธีการสร้างกราฟของเซกเมนต์แบบจับคู่ทุกขอบเขตของหน่วยเสียง พบว่า กราฟที่ได้จากวิธีการสร้างกราฟของเซกเมนต์แบบอาศัยการเปลี่ยนแปลงสเปกตรัมมีขนาดเล็กกว่า 7 – 8 เท่าโดยประมาณ แต่มีความครอบคลุมน้อยกว่ากราฟที่ได้จากการสร้างกราฟของเซกเมนต์แบบจับคู่ทุกขอบเขตของหน่วยเสียงถึง 23.96% (จาก 88.98% เป็น 65.02%) และในการทดลองเปรียบเทียบประสิทธิภาพวิธีการสร้างกราฟของเซกเมนต์แบบหลายระดับกับวิธีการสร้างกราฟของเซกเมนต์แบบจับคู่ทุกขอบเขตของหน่วยเสียงพบว่า กราฟที่ได้จากวิธีการสร้างกราฟของเซกเมนต์แบบหลายระดับมีขนาดเล็กกว่าประมาณ 14 เท่าซึ่งเล็กกว่ากราฟที่ได้จากวิธีการสร้างกราฟของเซกเมนต์แบบอาศัยการเปลี่ยนแปลงสเปกตรัม ในขณะที่มีความครอบคลุมน้อยกว่ากราฟที่ได้จากการสร้างกราฟของเซกเมนต์แบบจับคู่ทุกขอบเขตของหน่วยเสียงประมาณ 11.58% (จาก 88.98% เป็น 77.40%) ซึ่งลดลงไปน้อยกว่ากราฟที่ได้จากวิธีการสร้างกราฟของเซกเมนต์แบบอาศัยการเปลี่ยนแปลงสเปกตรัม ดังนั้นเมื่อพิจารณาความจำเป็นของระบบรู้จำเสียงพูดแบบอาศัยเซกเมนต์ที่จะต้องทำงานได้อย่างรวดเร็ว วิธีการสร้างกราฟของเซกเมนต์แบบหลายระดับจึงมีความเหมาะสมที่สุด

โดยสรุปแล้ววิทยานิพนธ์นี้เสนอวิธีการแบ่งเสียงพูดเป็นเซกเมนต์สำหรับการรู้จำเสียงพูดแบบอาศัยเซกเมนต์แบบใหม่ ที่นำเอาสารสนเทศสวณศาสตร์และซอฟต์แวร์เวกเตอร์แมชชีนมาประยุกต์ใช้เพื่อให้ประสิทธิภาพทั้งในด้านคุณภาพของผลลัพธ์ และความเร็วในการทำงานดีกว่าวิธีการแบ่งเสียงพูดเป็นเซกเมนต์แบบอาศัยเครื่องรู้จำเสียงพูด ดังนั้นวิธีการแบ่งเสียงพูดเป็นเซกเมนต์ที่เสนอนี้ จึงสามารถนำไปใช้งานในระบบรู้จำเสียงพูดแบบอาศัยเซกเมนต์ได้เป็นอย่างดี และส่งผลให้การรู้จำเสียงพูดมีความถูกต้องมากยิ่งขึ้น นอกจากนี้วิธีการแบ่งเสียงพูดที่เสนอยังอาจนำไปประยุกต์ใช้ในการเตรียมข้อมูลเสียงสำหรับระบบสังเคราะห์เสียงพูด หรือนำไปเตรียมข้อมูลฝึกฝนสำหรับเครื่องรู้จำเสียงพูดอื่นๆ ได้

## ข้อเสนอแนะ

แม้ว่าวิธีการแบ่งเสียงพูดเป็นเซกเมนต์โดยใช้สารสนเทศสวณศาสตร์ที่เสนอไว้ในวิทยานิพนธ์นี้ จะมีประสิทธิภาพดีกว่าทั้งในด้านคุณภาพของผลลัพธ์และความเร็วในการทำงานเมื่อเปรียบเทียบกับวิธีการแบ่งเสียงพูดเป็นเซกเมนต์แบบอาศัยเครื่องรู้จำเสียงพูด แต่ประสิทธิภาพที่ได้นั้น ยังคงห่างจากประสิทธิภาพของวิธีการแบ่งเสียงพูดเป็นเซกเมนต์ด้วยคนอยู่ ดังนั้นงานวิจัยนี้จึงยังสามารถพัฒนาได้อีกในหลายทาง

การเพิ่มขึ้นตอนการกรองเซกเมนต์ที่มีขนาดเล็กเกินกว่าจะเป็นหน่วยเสียงได้ออกไป ก็เป็นหนทางหนึ่งในการเพิ่มประสิทธิภาพของการแบ่งเสียงพูดเป็นเซกเมนต์ โดยอาจเลือกกรองเฉพาะเซกเมนต์ที่มีระยะห่างระหว่างขอบเขตด้านซ้ายและด้านขวาซึ่งสั้นกว่า 10 มิลลิวินาที เพื่อให้กราฟของเซกเมนต์ที่ได้มีขนาดเล็กลง ทำให้กราฟมีคุณภาพมากยิ่งขึ้น

นอกจากนี้อาจทำการทดลองใช้ซอฟต์แวร์เวกเตอร์แมชชีนที่ใช้ฟังก์ชันเคอร์เนลแบบอื่นๆ นอกเหนือไปจากซอฟต์แวร์เวกเตอร์แมชชีนที่ใช้ฟังก์ชันเคอร์เนลแบบเชิงเส้นและแบบพหุนาม เช่น ซอฟต์แวร์เวกเตอร์แมชชีนที่ใช้ฟังก์ชันเคอร์เนลแบบอาร์บีเอฟ มาใช้จำแนกเสียงพูดตามลักษณะการออกเสียง ซึ่งอาจจะสามารถแบ่งแยกสัทลักษณะลักษณะการออกเสียงได้แม่นยำมากยิ่งขึ้น ส่งผลให้ได้ขอบเขตของหน่วยเสียงที่มีคุณภาพเพิ่มมากขึ้นตามไปด้วย

และส่วนสำคัญอีกประการหนึ่งของการปรับปรุงประสิทธิภาพการแบ่งเสียงพูดเป็นเซกเมนต์ด้วยวิธีการนี้คือ การใช้ลักษณะสำคัญเพิ่มเติมโดยใช้สารสนเทศสวณศาสตร์มาประกอบการพิจารณาสำหรับนำไปใช้ในการแบ่งแยกสัทลักษณะลักษณะการออกเสียงแต่ละแบบให้เหมาะสม เพื่อช่วยให้สามารถตรวจหาขอบเขตของหน่วยเสียงได้แม่นยำมากยิ่งขึ้น ทำให้ได้กราฟของเซกเมนต์ที่น่าขอบเขตของหน่วยเสียงนี้ไปใช้มีคุณภาพเพิ่มมากขึ้นด้วย