



## บทที่ 1

### บทนำ

ในปัจจุบันเทคโนโลยีต่างๆ มีการพัฒนาอย่างรวดเร็ว รวมถึงเทคโนโลยีด้านคอมพิวเตอร์ ซึ่งจะถูกนำไปใช้ร่วมกับเทคโนโลยีด้านอื่นๆ อีกมากมาย เช่น การสื่อสาร, ระบบควบคุม ฯลฯ คอมพิวเตอร์จึงถูกพัฒนาให้มีขีดความสามารถมากขึ้นเรื่อยๆ เพื่อทำงานต่างๆ แทนมนุษย์ได้ การรู้จักก็เป็นการพัฒนาความสามารถอีกด้านของคอมพิวเตอร์ โดยอาศัยเทคนิคของ Digital Signal เพื่อให้สามารถติดต่อกับคอมพิวเตอร์ได้คล่องตัวยิ่งขึ้น

ในอดีตมีการวิเคราะห์ว่าการสอนให้คอมพิวเตอร์สามารถรู้จำเสียงพูดของมนุษย์นั้นทำไม่ได้ เนื่องจากการพูดของมนุษย์มีความซับซ้อน และมีความแตกต่างกันในแต่ละบุคคล ความเร็ว รวมถึงเสียงสูงเสียงต่ำ แต่ก็มี การวิจัยเพื่อให้คอมพิวเตอร์สามารถรู้จำเสียงพูดเรื่อยๆ จนในปัจจุบันการรู้จำเสียงพูดสามารถจะใช้งานได้ มีการนำไปใช้กับอุปกรณ์ต่างๆ เช่น ATM (Automated Teller Machine) และโทรศัพท์มือถือ ถึงแม้ว่าคอมพิวเตอร์จะไม่สามารถรู้จำเสียงได้เท่ามนุษย์ แต่เราสามารถสอนให้คอมพิวเตอร์รู้จำได้ในขีดจำกัดระดับหนึ่ง โดยมีการกำหนดขอบเขตของการรู้จำ เช่น จำกัดคำที่จะสามารถรู้จำได้ มีการคาดคะเนกันว่าในอนาคต อุปกรณ์เครื่องใช้เกือบทุกประเภทจะสามารถสั่งงานได้โดยเสียงพูด

งานวิจัยเกี่ยวกับเสียงพูดของมนุษย์ ได้มีการวิจัยหลายๆ ด้าน นับตั้งแต่การวิเคราะห์เสียง (speech analysis) การสังเคราะห์เสียง (speech synthesis) การรู้จำคำพูด (speech recognition) และการรู้จำเสียง (voice recognition) ในส่วนของ speech recognition นั้น เป็นงานวิจัยที่มีการวิจัยกันอย่างแพร่หลาย มีการศึกษาทั้งแบบขึ้นกับผู้พูด (dependent) และไม่ขึ้นกับผู้พูด (Independent) ในส่วนของการศึกษาระณีไม่ขึ้นกับผู้พูด (Independent) มีการศึกษาทั้งแบบคำโดด (Isolated word) (Rabiner and Levinson, 1981), (Lamel et al., 1981) และ คำต่อเนื่อง (continuous word) (Kai-fu and Hon, 1989)

งานวิจัยที่วิจัยเกี่ยวกับการรู้จำเสียงภาษาไทย ได้มีการวิจัยโดยใช้กรรมวิธีต่างๆ เช่น การวิจัยการรู้จำเสียงพูดคำเดี่ยวต่างบุคคล (ไพศาล ธรรมโพธิ์ทอง 2532) เป็นงานวิจัยที่ใช้เทคนิคของไดนามิกโปรแกรมมิ่ง และใช้ K-Nearest Neighbour (KNN) เป็นกฎการตัดสินใจ และงานวิจัยการรู้จำเสียงพูดภาษาไทยในประโยค ต่อเนื่อง โดยใช้หน่วยเสียงพยางค์เป็นหน่วยเสียงในการเปรียบเทียบ (ทวี ปทุมทาน 2529) โดยใช้เทคนิคการประมาณพินธะเชิงเส้น (Linear Prediction) และไดนามิกโปรแกรมมิ่ง

ในงานวิจัยฉบับนี้ได้นำวิธีการ Hidden Markov Model (HMM) มาใช้วิเคราะห์คำโดดแบบไม่ขึ้นกับผู้พูด มีการจำกัดขอบเขตของการวิจัยโดยกำหนดให้รู้จำเสียงตัวเลขภาษาไทย 0-9 ใช้บุคคลในการหาโมเดล 20 คนที่เป็นชาย(เพื่อทดสอบความเป็นไปได้กับกลุ่มย่อยก่อน จึงมีการจำกัดต้นแบบรู้จำให้เป็นชาย เพราะชายและหญิงมีเสียงที่แตกต่างกันมาก) HMM เหมาะสำหรับการวิเคราะห์ระบบที่เป็น stochastic รวมทั้งมีความรวดเร็วในการคำนวณ ต้องการเนื้อที่ในการเก็บน้อย ซึ่งมีผลทำให้จำนวนคำที่สามารถรู้จำได้มีจำนวนมากด้วย แต่มีข้อเสียคือต้องการรูปแบบเสียงจำนวนมากในการหาโมเดล(model) จึงมีงานวิจัยต่างๆ ที่พยายามปรับปรุงความถูกต้องในการรู้จำที่มีรูปแบบของเสียงในการหาโมเดลจำนวนจำกัด (Koo and UN, 1990), (Kim and Lee, 1991)



งานวิจัยนี้เป็นงานวิจัยเพื่อศึกษาว่าวิธีการของ HMM สามารถนำมาใช้ได้ดีในการรู้จำเสียง ภาษาไทยหรือไม่ ซึ่งจะเป็นประโยชน์ในการพัฒนาการรู้จำโดยใช้วิธี HMM กับเสียงภาษาไทยคำโดดอื่น ๆ และเสียงภาษาไทยแบบคำต่อเนื่องต่อไป ในวิทยานิพนธ์แบ่งออกเป็น 5 บทดังนี้

บทที่ 1 บทนำ ที่ได้กล่าวมาข้างต้น

บทที่ 2 รูปแบบการรู้จำเสียง

เป็นการกล่าวถึงทฤษฎีการทำ feature ของเสียงพูด รวมถึงวิธีการที่จะนำมาวิเคราะห์เสียงพูด

บทที่ 3 ขั้นตอนการทดลอง

จะกล่าวถึง algorithm ที่ถูกนำมาใช้ในการทดลอง โดยอาศัยทฤษฎีที่ได้กล่าวมาในบทที่ 3

บทที่ 4 ผลการทดลอง

เป็นการรวบรวมผลการทดลองที่ได้ทดลองทั้งหมด โดยอาศัยทฤษฎีจากบทที่ 3 และขั้นตอนเป็นไปตามบทที่ 3

บทที่ 5 สรุปผลการทดลองและวิจารณ์

สรุปเกี่ยวกับผลการทดลองทั้งหมดและข้อเสนอแนะสำหรับผู้ที่จะทำการวิจัยและพัฒนาการรู้จำเสียงโดยวิธี HMM ต่อไป

ภาคผนวก

ภาคผนวก ก. เป็นการทดลองเพื่อหาพารามิเตอร์ที่เหมาะสม ที่จะนำไปใช้กับวิธีการในขั้นตอนต่าง ๆ

ภาคผนวก ข. เป็นการยกตัวอย่างผลการทดลองที่มีผู้ได้ทำวิจัยไว้ ในการรู้จำเสียงพูดตัวเลขภาษาอังกฤษด้วยวิธีต่าง ๆ

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย