



## บทที่ 1

### บทนำ

#### 1.1 ความเบื้องต้น

ปัจจุบันการศึกษาทางด้านการรู้จำเสียงพูด (Voice/Speech Recognition) มีการศึกษากันมากขึ้นและมีความก้าวหน้าไปอย่างมาก ทั้งนี้เนื่องมาจากมีผลงานวิจัยออกมามากมายในต่างประเทศโดยตีพิมพ์ผ่านวารสารต่างประเทศต่าง ๆ อยู่ตลอด อีกทั้งความสามารถในการประมวลผลมีประสิทธิภาพสูงขึ้นกว่าแต่ก่อนมาก นอกจากนี้แล้วเครื่องคอมพิวเตอร์ก็ได้มีราคาแพงเช่นในอดีต รวมทั้งการที่มีบริษัทต่าง ๆ หันมาพัฒนาอุปกรณ์ต่อพ่วงประเภทการ์ดเสียงที่มีประสิทธิภาพสูงในราคาถูกลง เช่น การ์ดซาวด์บลาสเตอร์ เป็นต้น จึงทำให้การศึกษาในด้านการรู้จำเสียงพูดเป็นไปได้สะดวกขึ้นกว่าแต่ก่อน โดยการศึกษาส่วนใหญ่จะมุ่งไปที่การนำไปใช้งานต่าง ๆ ได้จริง ซึ่งการรู้จำเสียงพูดนี้ยึดหลักของความถูกต้อง (accuracy) และความเป็น real time เป็นหลัก

ในการศึกษาเกี่ยวกับการรู้จำเสียงพูดนั้นมีการศึกษาในรูปแบบต่าง ๆ มากมาย ในที่นี้จะแบ่ง 2 แบบ คือ การระบุตัวผู้พูด (Speaker Identification) (Dante and Sarma, 1979) และการรู้จำเสียงพูดของผู้พูด (Speaker Recognition) ซึ่งในส่วนของ การรู้จำเสียงพูดของผู้พูดนั้นยังสามารถแบ่งออกได้อีก 4 แบบ ด้วยกันคือ

ก. การรู้จำคำเดี่ยวโดยขึ้นต่อผู้พูด (Speaker Dependent Isolated Word Recognition)

ข. การรู้จำคำเดี่ยวโดยไม่ขึ้นต่อผู้พูด (Speaker Independent Isolated Word Recognition)

ค. การรู้จำคำต่อเนื่องโดยขึ้นต่อผู้พูด (Speaker Dependent Continuous Words Recognition)

ง. การรู้จำคำต่อเนื่องโดยไม่ขึ้นต่อผู้พูด (Speaker Independent Continuous Words Recognition)

ในวิทยานิพนธ์ฉบับนี้ จะมุ่งศึกษาเรื่องของระบบการรู้จำเสียงพูดของกลุ่มตัวเลขภาษาไทยแบบไม่ขึ้นต่อผู้พูด โดยจะเน้นผลความถูกต้องของเสียงพูดของตัวเลขไทยคำเดียว

โดด ๆ เป็นหลัก ในการศึกษาจะใช้ดิสครีตฮาร์ตลีย์ทรานส์ฟอร์ม (Discrete Hartley Transform, DHT) (สุนิสา จันทวิกุล, 2536) และ ไดนามิกโทมวาร์ปิง (Furui, 1989) ช่วยในการวิเคราะห์

## 1.2 วัตถุประสงค์ของงานวิจัย

- ก. เพื่อศึกษาถึงวิธีการรู้จำเสียงพูดที่เป็นกลุ่มคำเฉพาะตัวเลขไทยแบบไม่ขึ้นกับผู้พูด
- ข. เพื่อประยุกต์เอาทฤษฎีดิสครีตฮาร์ตลีย์ทรานส์ฟอร์มมาใช้ในการรู้จำเสียงพูด
- ค. เพื่อจัดทำโปรแกรมใช้งานสำหรับรู้จำเสียงพูดตัวเลขไทยโดยไม่ขึ้นกับผู้พูด
- ง. เพื่อเป็นแนวทางในการศึกษาวิธีการรู้จำเสียงพูดภาษาไทยที่ไม่ได้จำกัดแต่เพียงตัวเลขเท่านั้น

## 1.3 ขอบเขตของงานวิจัย

- ก. ฮาร์ดแวร์ ใช้เครื่อง IBM PC/AT หรือ Compatible และ Sound Card
- ข. ภาษา ใช้ภาษาซีในการพัฒนาโปรแกรม
- ค. อินพุต ได้จากการบันทึกผ่าน Sound Card เป็นเสียงของตัวเลขภาษาไทย 0 ถึง 9 รวมทั้งคำว่า "สิบ", "เอ็ด", "ยี่", "ร้อย", "พัน", "หมื่น", "แสน", "ล้าน" โดยให้ข้อมูลขนาด 8 บิต
- ง. วิธีการ ใช้ฮาร์ตลีย์ทรานส์ฟอร์มและไดนามิกโทมวาร์ปิง

## 1.4 ขั้นตอนและวิธีการ

- ก. ศึกษากระบวนการรู้จำเสียงพูดในลักษณะต่าง ๆ
- ข. กำหนดขอบเขตวิทยานิพนธ์
- ค. เขียนโปรแกรมวิเคราะห์
- ง. เก็บข้อมูลเสียงเพื่อใช้ในการวิเคราะห์
- จ. การทดสอบและแก้ไขปรับปรุงโปรแกรมใช้งานให้เหมาะสม
- ฉ. สรุปผล

### 1.5 เป้าหมายของงานวิจัย

- ก. โปรแกรมใช้งานสำหรับการรู้จำเสียงพูดตัวเลขไทยคำหนึ่ง ๆ แบบต่างบุคคล
- ข. สามารถทำการรู้จำเสียงตัวเลขไทยคำหนึ่ง ๆ ได้โดยเน้นที่คำเดียวโดด ๆ ได้จากกลุ่มทดสอบจำนวน 20 คน
- ค. ความถูกต้องในการรู้จำเสียงพูดคำเดียวโดด ๆ อย่างน้อย 70 % โดยใช้ชุดเปรียบเทียบเพียงชุดเดียว