รายงานผลการประดิษฐ์
ทุนอุดหนุนโครงการสิ่งประดิษฐ์

เรื่อง

ระบบนำเข้าเสียงพูด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จฬ
วศ 15
007011

โดย

วีระ ริ้วพิทักษ์

2535

ชื่อโครงการ ระบบนำเข้าเสียงพูด

ชื่อผู้ดำเนินงาน ผู้ช่วยศาสตราจารย์ ดร.วีระ ริ้วพิทักษ์

เดือนและปีที่ทำวิจัยเสร็จ มกราคม 2535

บทคัดย่อ

ระบบนำเข้าเสียงพูด เป็นระบบที่ทำให้เครื่องคอมพิวเตอร์รับรู้เสียงพูดภาษาไทยเป็นวลีสั้น ๆ ของแต่ละบุคคล แล้วนำผลการรับรู้ไปค้นหารายละเอียดที่ต้องการจากฐานข้อมูลภาษาไทยที่ได้เตรียมไว้แล้ว รายละเอียดที่ค้นได้จะแสดงบนจอภาพเป็นภาษาไทย การทำงานในขั้นแรกจะต้องให้เครื่องเรียนรู้โดยเสียงพูดที่รับเข้าทางไมโครโฟน จะผ่านการวิเคราะห์โดยใช้เทคนิคของ LPC (Linear Predictive Coding) ด้วยบอร์ดไมโครโปรเซสเซอร์ความเร็วสูง สร้างเป็นแบบอ้างอิง ในขั้นต่อมาจะเป็นการรับรู้เสียงพูด โดยนำเสียงพูดมาวิเคราะห์แล้วนำไปเปรียบเทียบกับแบบอ้างอิง เมื่อได้แบบอ้างอิงที่ใกล้เคียงที่สุด จะนำแบบดังกล่าวไปเป็นตัวค้นหารายละเอียดในฐานข้อมูลต่อไป

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

I 12847495

Project Title: Voice Input System

Name of Investigator: Weera RIEWPITUK, D.Ing.

Finished Date: January 1992

Abstract

Voice Input System is the system which enables computer to recognize Thai language in short phases for each person. Its recognition will be searched for detailed data needed from Thai database which has already prepared. The detailed data found will be displayed in Thai on the screen. At the first step, the computer has to recognize voice input through microphone. Voice inut will be analyzed by Linear Prediction Coding technique using high speed micro processor board set up as reference patterns. The next step, voice recognition will be analyzed and the result will be compared with reference in order to receive the most similar reference pattern. This reference will be used to search for further detailed data in database.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ระบบนำเข้าเสียงพูด

Voice Input System

ดร. วีระ ริ้วพิทักษ์

ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

1. บทนำ

1.1 ความเป็นมาของปัญหา ระบบคอมพิวเตอร์ที่มีใช้กันอยู่ในปัจจุบัน มีหน่วยนำเข้าข้อมูลหลักก็คือ แป้นพิมพ์ การพัฒนาในเทคโนโลยีส่วนนี้เป็นไปได้ช้ามากเมื่อเทียบกับในส่วนอื่นของระบบคอมพิวเตอร์ ไม่ว่าโปรเซสเซอร์ หน่วยความจำ ซอร์ฟแวร์ ได้พัฒนาไปไกลมากแล้ว แต่การนำเข้า ก็ยังใช้การป้อนข้อมูลด้วยแป้นพิมพ์เหมือนเดิม ความสามารถในการติดต่อกับระบบคอมพิวเตอร์ (Man-machine communication) ที่เป็นอยู่ในปัจจุบันค่อนข้างจะจำกัด จึงได้มีความพยายามที่จะเพิ่มขีดความสามารถของการติดต่อกับระบบคอมพิวเตอร์ กลุ่มแรกได้มีการพัฒนาด้านการอ่านของระบบ ได้แก่ บาร์โค๊ด (Bar Code) เครื่องอ่านลายมือเขียน (วัลนพ ต้นฤดี, 2533) และเครื่องอ่านตัวอักษรด้วยแสง (Optical Character Reader) (HIRUNVANICHAKORN, P., et al. 1984) กลุ่มที่สอง ได้มีการพัฒนาทางด้านการฟังและพูดภาษาไทยของระบบ การพูดเป็นการพัฒนาวิธีการสร้างเสียงพูด (Synthesis) มีการพัฒนาการสร้างโดยใช้หน่วยครึ่งพยางค์ (Demisyllable)(SALAVARI C. and IMAI S.,1983) และการสังเคราะห์เสียงวรรณยุกต์ (SALAVARI C.,IMAI S.,1984) ส่วนด้านการฟังหรือการรับรู้เสียงพูดภาษาไทย มีการพัฒนาระบบ ที่นำเอาวรรณยุกต์ไปช่วยในการรับรู้ (Recognition) สำหรับการรับรู้คำพูดแบบคำโดย บุคคลเดียว (RIEWPITUK W. ,1985)

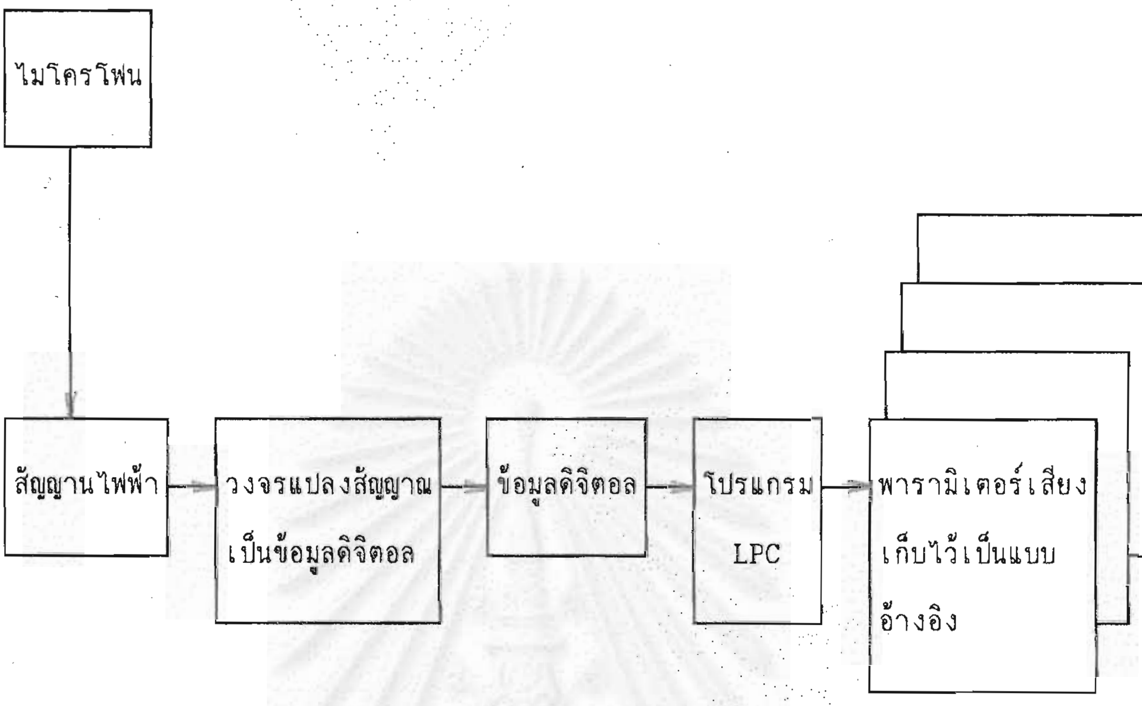
1.2 <u>คุณสมบัติของระบบ</u> ระบบนำเข้าเสียงพูดที่พัฒนาขึ้นนี้ จะรับเสียงพูดภาษาไทยเฉพาะบุคคลเป็นคำสั้นๆ แล้วนำคำดังกล่าวไปค้นหารายละเอียดในฐานข้อมูลภาษาไทยที่ได้จัดเตรียมไว้แล้ว เมื่อค้นหาพบจะแสดงรายละเอียดบนจอภาพ

1.3 <u>ประโยชน์ของระบบ</u> ระบบนำเข้าเสียงพูด เป็นระบบที่ทำให้เครื่องคอมพิวเตอร์รับรู้เสียงพูดของมนุษย์เป็นการเพิ่มหน่วยนำเข้า (Input Unit) ของเครื่องคอมพิวเตอร์อีกรูปแบบหนึ่งการนำเข้าของข้อมูลในรูปเสียงพูด จะมีอัตราการนำเข้าสูงกว่าการนำเข้าด้วยแป้นพิมพ์ที่นิยมใช้กันอยู่โดยทั่วไป อีกทั้งยังไม่ต้องการฝึกฝนเป็นพิเศษ เพราะทุกคนพูดเป็นอยู่แล้ว นอกจากนั้นยังทำให้มือว่างที่จะใช้งานอื่นได้อีกด้วย ระบบที่พัฒนาขึ้น นำไปใช้งานกับการสอบถามข้อมูล งานควบคุมอุปกรณ์ งานตรวจวัดค่าจากมาตร เป็นต้น

2. <u>การทำงานของระบบนำเข้าเสียงพูด</u>

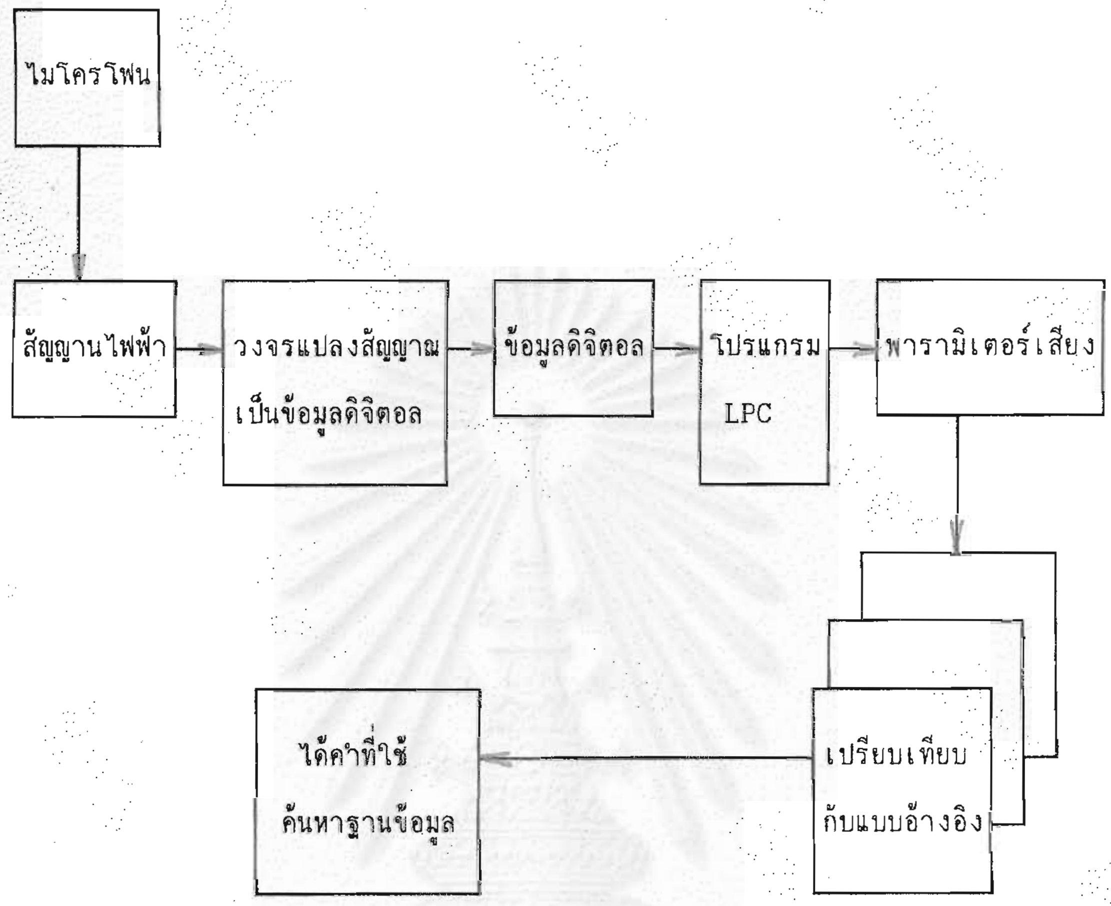
การทำงานของระบบนำเข้าเสียงพูดประกอบด้วยขั้นตอนต่างๆ ได้แก่ ขั้นตอนการเรียนรู้เสียงพูด ซึ่งจะเป็นการสร้างแบบอ้างอิงเสียงพูดเก็บไว้ในเครื่องคอมพิวเตอร์ ขั้นตอนการรับรู้เสียงพูด จะเป็นการนำเอาเสียงที่พูดเข้าไปเปรียบเทียบกับแบบอ้างอิง เพื่อหาว่าใกล้เคียงกับแบบอ้างอิงแบบใดมากที่สุด ขั้นการสร้างฐานข้อมูล โดยจะสร้างฐานข้อมูลภาษาไทย ที่มีเขตข้อมูลสำหรับใช้ในการค้นหา (Key field) เป็นภาษาไทย และขั้นสุดท้ายจะเป็นการนำเสียงพูดไปค้นหาฐานข้อมูล

2.1 <u>ขั้นการเรียนรู้เสียงพูด</u> เสียงพูดแต่ละชุดของเสียงพูดที่ผ่านไมโครโฟน จะเปลี่ยนเป็นสัญญาณไฟฟ้าแบบแอนนาลอก แล้วจึงเปลี่ยนรูปเป็นสัญญาณดิจิตอล เก็บไว้ในหน่วยความจำ ข้อมูลที่เก็บในหน่วยความจำของแต่ละชุดของเสียงพูดจะผ่านวิธีการวิเคราะห์ โดยเทคนิคของ LPC (Linear Predictive Coding) ได้เป็นพารามิเตอร์ของ LPC ดังรูปที่ 1 เสียงทุกเสียงที่ต้องการให้เครื่องรับรู้ จะต้องผ่านการเรียนรู้ก่อนเครื่องจะทำการเก็บเสียงทุกเสียงที่เรียนรู้ไว้เป็นแบบอ้างอิง



รูปที่ 1 การเรียนรู้เสียงพูด

2.2 <u>ขั้นการรับรู้เสียงพูด</u> จะเป็นการรับเสียงพูดเข้ามาเปรียบเทียบกับแบบอ้างอิง แบบอ้างอิงที่ใกล้เคียงที่สุด ถือว่าเป็นเสียงนั้น การทำงานอยู่ในรูปที่ 2 เสียงพูดจะผ่านขั้นตอนการวิเคราะห์ในทันทีทันใด แล้วนำไปเปรียบเทียบกับแบบอ้างอิงในหน่วยความจำ ขั้นตอนนี้เป็นขั้นตอนที่ยากที่สุด เนื่องจากว่าเสียงพูดแต่ละครั้งในคำเดียวกัน จะไม่เหมือนกันทีเดียว มักจะสั้นหรือยาวไปบ้าง วิธีในการเปรียบเทียบมีด้วยกันหลายวิธี แต่วิธีหนึ่งที่นิยมใช้กันมาก ได้แก่ วิธีการเปรียบเทียบโดยใช้ ไดนามิคโปรแกรมมิ่ง วิธีนี้จะช่วยยืดหยุ่นได้ในแกนเวลา การพูดเร็วหรือช้าไปบ้างก็ยังได้ผลดีอยู่ ปัญหาของไดนามิกโปรแกรมมิ่ง ก็คือ ใช้เวลาในการคำนวนและเปรียบเทียบสูงมาก เพื่อให้การตอบสนองได้เร็ว จำเป็นต้องใช้ไมโครโปรเซสเซอร์ความเร็วสูง มาทำงานนี้โดยเฉพาะ ในระบบนี้ใช้บอร์ดทีมีไมโครโปรเซสเซอร์ความเร็วสูง ของ Texas Instrument ในตระกูล TMS 320 ช่วยในการคำนวณ



รูปที่ 2 การรับรู้เสียงพูด

2.3 <u>ขั้นการสร้างฐานข้อมูล</u> ฐานข้อมูลที่สร้างเป็นภาษาไทย เพื่อให้สะดวกในการนำไปใช้งานจึงเลือกไดรเวอร์ภาษาไทยที่ไม่ขึ้นกับฮาร์ดแวร์ ได้แก่ VTHAI ที่พัฒนาขั้นที่จุฬาลงกรณ์มหาวิทยาลัย และเลือกฐานข้อมูลที่เป็นที่นิยม มีประสิทธิภาพสูง ใช้งานร่วมกับภาษาไทยได้เป็นอย่างดี กับมีความสามารถในเรื่องโปรแกรมในตัวเอง ฐานข้อมูลดังกล่าวได้แก่ FOXPRO การกำหนดโครงสร้างของฐานข้อมูล จะกำหนดเป็น 2 ส่วน ส่วนแรกจะเรียกว่า ข้อมูลที่เป็นเงื่อนไขในการค้นหา ส่วนที่สอง จะเป็นรายละเอียดของข้อมูลที่ต้องการทราบ

การควบคุม การค้นหา การแสดงผล ทำได้ด้วยความสามารถของฐานข้อมูลเองกับความสามารถทางด้านโปรแกรมที่มากับฐานข้อมูล โปรแกรมจะมีหน้าที่คอยรับข้อมูลที่เป็นเงื่อนไขในการค้นหามาจาก โปรแกรมรับรู้เสียงพูด ตัวอย่างโปรแกรมโครงสร้างและข้อมูลอยู่ในภาคผนวก ก.

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.4 <u>การนำเสียงพูดไปค้นหาฐานข้อมูล</u> ส่วนนี้จะเป็นการเชื่อมโยงส่วนการรับรู้เสียงพูด กับฐานข้อมูลที่ได้เตรียมไว้แล้ว โปรแกรมส่วนนี้จะเขียนขึ้นด้วยภาษาที่เรียกว่าภาษาเสียงพูด (Language of speech) ซึ่งจะเป็นการกำหนดประโยคต่างๆ รวมทั้งไวยากรณ์ ขึ้นมาแล้วนำข้อมูลที่กำหนดเก็บไว้ในไฟล์ หลังจากนั้นจะทำการแปล (Compile) ให้เป็นโปรแกรมที่วิ่งได้ แล้วนำไปจัดไว้เป็นโปรแกรมที่วิ่งในหน่วยความจำแบบฝังตัว (Resident program) ตัวอย่างโปรแกรมอยู่ในภาคผนวก ข. โปรแกรมนี้จะเชื่อมโยงผลจากการรับรู้เสียงพูด กับคำที่เป็นเงื่อนไขในการค้นหา แล้วส่งต่อคำที่เป็นเงื่อนไขไปให้โปรแกรมฐานข้อมูลอีกทีหนึ่ง

3. <u>ส่วนประกอบของระบบ</u>

การทำงานของระบบดังที่กล่าวมาแล้ว จะต้องประกอบด้วยส่วนที่เป็นฮาร์ดแวร์ ได้แก่ตัวเครื่องไมโครคอมพิวเตอร์ และบอร์ดคำนวณความเร็วสูง ส่วนซอร์ฟแวร์นั้น จะเป็นส่วนทำงานตามหน้าที่ดังกล่าว

3.1 <u>ฮาร์ดแวร์</u>

ฮาร์ดแวร์หลัก ได้แก่ เครื่องไมโครคอมพิวเตอร์ ระดับ IBM XT หรือสูงกว่า มีหน่วยความจำ 640 KByte พร้อมฮาร์ดดิส ขนาดความจุ 40 MByte เป็นอย่างน้อย

บอร์ดคำนวณความเร็วสูง ที่มีไมโครโปรเซสเซอร์ TMS 32010 เป็นตัวคำนวณ เสียบบนสล๊อตของไมโครคอมพิวเตอร์ ซึ่งบอร์ดนี้จะรับเสียงพูดจากไมโครโฟน เปลี่ยนเป็นข้อมูล และทำกระบวนการวิเคราะห์ เปรียบเทียบ บอร์ดจะทำงานอยู่ภายใต้การควบคุมของไมโครคอมพิวเตอร์

ไมโครโฟน ที่มีคุณสมบัติรับเสียงพูดด้านหน้าเท่านั้น เพื่อจะได้ไม่เก็บเสียงสิ่งแวดล้อม

3.2 <u>ซอฟต์แวร์</u>

โปรแกรมควบคุมระบบ ได้แก่ DOS Version 3.0 ขึ้นไป พร้อมไดรเวอร์ภาษาไทย VITHAI ซึ่งจะทำให้การนำเข้าและแสดงผลเป็นภาษาไทยที่ไม่ขึ้นกับฮาร์แวร์

โปรแกรมฐานข้อมูลที่ใช้คือ FOXPRO Version 2.0 ที่เลือกโปรแกรมฐานข้อมูลดังกล่าว เพราะมีประสิทธิภาพสูง เขียนโปรแกรมได้ด้วย และทำงานเข้ากับภาษาไทย VTHAI ได้เป็นอย่างดี

โปรแกรมภาษาเสียงพูด ซึ่งเป็นยูทิลิตี้ ของบริษัท TI ที่จะช่วยการสร้างโมดูลในการเชื่อมโยงการรับข้อมูลจากไมโครโฟนกับฐานข้อมูล และเป็นโปรแกรมที่วิ่งอยู่ด้วยกับโปรแกรมฐานข้อมูลในหน่วยความจำ

4. <u>ลักษณะการใช้งาน</u>

ก่อนการใช้งาน จะต้องจัดเตรียมฐานข้อมูลให้เรียบร้อยเสียก่อน โดยแยกข้อมูลในแต่ละเร็คคอร์ดออกเป็นสองส่วน ส่วนแรกเป็นข้อมูลที่ใช้เป็นเงื่อนไขในการค้นหา ส่วนที่สองเป็นข้อมูลที่ต้องการค้นหาเพื่อแสดงบนจอภาพ ฐานข้อมูลที่ใช้เป็นตัวอย่าง ได้แก่ฐานข้อมูล อัตราแลกเปลี่ยนเงินตราต่างประเทศ กับเงินบาทของไทย ส่วนที่ใช้ในการค้นหา คือชื่อของเงิน เช่น "ดอลลาร์สหรัฐ" "เยน" เป็นต้น ส่วนรายละเอียดที่ได้จากการค้นหาได้แก่ อัตราซื้อขาย เมื่อเทียบกับเงินบาทไทย

ต่อมาจะต้องป้อนข้อมูลให้เครื่องเรียนรู้ โดยจะต้องพูดคำที่ใช้เป็นเงื่อนไขในการค้นหาทุกคำ ได้แก่ คำว่า "ดอลลาร์สหรัฐ" "เยน" เป็นต้น คำพูดนี้จะต้องตรงกับในฐานข้อมูลที่ได้เตรียมไว้แล้ว การป้อนข้อมูลในเสียงพูดในขั้นนี้ ต้องให้ความระมัดระวังเป็นอย่างมาก เพราะจะเก็บไว้เป็นแบบเปรียบเทียบ ควรพูดให้ชัดถ้อยชัดคำ ไม่เร็วหรือช้าเกินไป

เมื่อเตรียมฐานข้อมูล และสอนเครื่องให้รู้จักเสียงครบทุกเสียงแล้ว ก็จะเป็นการนำไปใช้งาน การใช้งานก็จะเป็นการถามเครื่องโดยพูดคำที่ต้องการทราบเข้าที่ไมโครโฟน เครื่องจะค้นหาข้อมูลจากฐานข้อมูล แล้วแสดงรายละเอียดบนจอภาพ การทำงานเป็นขั้นตอนอยู่ในภาคผนวก ค.

5. <u>สรุป</u> ระบบนำเข้าเสียงพูดเป็นการนำเอาเทคโนโลยีเสียงพูด ประสานกับฐานข้อมูล และการแสดงผลภาษาไทย เพื่อนำไปประยุกต์ใช้กับการเรียกหารายละเอียดของข้อมูลที่ต้องการโดยใช้เสียงเป็นเงื่อนไขในการค้นหา จำนวนคำที่เรียกค้นได้ประมาณ 50

คำ ซึ่งเพียงพอสำหรับงานหลายประเภท และเป็นตัวอย่างของการนำเอาระบบนำเข้าเสียงพูดไปใช้งาน แนวทางในการค้นคว้า วิจัย เพื่อให้ใช้งานได้กว้างขวาง จะต้องเพิ่มจำนวนคำให้มากขึ้น และเพิ่มความถูกต้อง ระบบที่อยู่ในปัจจุบันรับจำนวนคำได้ตั้งแต่ 50 คำ จนถึงมากกว่า 10,000 คำ มีความถูกต้องมากกว่า 95 % (LEE K, et al. 1990) จำนวนคำที่มากขึ้นนั้นขึ้นอยู่กับขนาดหน่วยความจำ และความเร็วของการคำนวณซึ่งจะต้องให้ผลตอบสนองในเวลาที่ต้องการได้ ส่วนความถูกต้องจะอยู่ที่พารามิเตอร์ของเสียง วิธีการเปรียบเทียบ ขณะนี้มีงานวิจัยที่นำเอาเทคโนโลยีของนิวรอลเน็ตเวิร์ค มาช่วยในการเปรียบซึ่งให้ผลลัพธ์ได้มากกว่า 98 % (WAIBEL A. and HAMPSHIRE, 1989) การนำนิวรอลเน็ทเวิร์คมาช่วยในงานรับรู้เสียงพูด จะช่วยในการแก้ปัญหาในเรื่องของพารามิเตอร์ที่ยังหาความหมายไม่ได้ แต่นิวรอลเน็ตเวิร์คจะเรียนรู้ได้เอง

เอกสารอ้างอิง

1. RIEWPITUK W., Thai isolated word recognition, Doctor of Engineering Thesis, Paris XI, 1985.

2. ___________, TI-speech application toolkit guide, Texas Instrument, 1989.

3. ___________, FoxPro version 2.0 manual, Fox software, 1991.

4. SARAVARI C., IMAI S., A demisyllable approach to speech systhesis of Thai - A tone language, J. Acoust. Soc. Jpn.(E)4,2 (1983).

5. SARAVARI C., IMAI S., An automatic procedure for extracting demisyllables from isolated mono syllabic source words for use in speech synthesis-by-rule of Thai, J. Acoust. Soc. Jpn. (E) 4,2,(1983).

5. HIRANVANICHAKORN P., et al., Recognition of Thai character recognition, Trans. IECE Japan, August 1984.

6. วัลนพ ตันฤดี, ระบบการรับรู้ลายมือเขียนภาษาไทย, วิทยานิพนธ์ ปริญญามหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย 2532.

7. WAIBEL A. and HAMPSHIRE J., Building block for speech, Byte, August 1989.

8. LEE K. et al., The spoken word, Byte, July 1990.

สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก.

โปรแกรมและโครงสร้างข้อมูล

โปรแกรมค้นหาฐานข้อมูล เขียนด้วย Foxpro

โปรแกรม MONEY1.PRG

```
SET ECHO OFF
SET ESCAPE OFF
TEST =.T.

SELECT 1
if ( file("moneyndx") )
   erase moneyndx.ntx
endif
USE MONEY
INDEX ON NAME TO MONEYNDX

DO WHILE TEST

     CLEAR
     MUNUCHOICE = SPACE(30)
     @ 05,10 SAY "อัตราแลกเปลี่ยนเงินตรา"
     @ 10,10 SAY "1 ดอลลาร์สหรัฐ"
     @ 11,10 SAY "2 ดอลลาร์สิงคโปร์"
     @ 12,10 SAY "3 ดอลลาร์สหรัฐ"
     @ 13,10 SAY "4 ดอลลาร์ฮ่องกง"
     @ 14,10 SAY "5 ฟรังก์สวิส"
```

```
    @ 15,10 SAY "6 ปอนด์สเตอร์ลิง"
    @ 16,10 SAY "7 มาร์กเยอรมัน"
    @ 17,10 SAY "8 ริมินบี (หยวนจีน)"
    @ 18,10 SAY "9 เยน (ต่อ 100 เยน)"
    @ 19,10 SAY "0 เลิกการทำงาน"
    @ 21,10 SAY "เลือก" "GET MENUCHOICE
    READ

IF MENUCHOICE = '0' .OR. MENUCHOICE = "เลิกการทำงาน"
    TEST = .F.
ELSE
    DO CASE
        CASE MENUCHOICE = '1'
            LOCATE ALL FOR NAME = "ดอลลาร์สหรัฐ"
        CASE MENUCHOICE = "ดอลลาร์สหรัฐ"
            LOCATE ALL FOR NAME = "ดอลลาร์สหรัฐ"
        CASE MENUCHOICE = '2'
            LOCATE ALL FOR NAME = "ดอลลาร์สิงคโปร์"
        CASE MENUCHOICE = "ดอลลาร์สิงคโปร์"
            LOCATE ALL FOR NAME = "ดอลลาร์สิงคโปร์"
        CASE MENUCHOICE = '3'
            LOCATE ALL FOR NAME = "ดอลลาร์มาเลเซีย"
        CASE MENUCHOICE = "ดอลลาร์มาเลเซีย"
            LOCATE ALL FOR NAME = "ดอลลาร์มาเลเซีย"
        CASE MENUCHOICE = '4'
            LOCATE ALL FOR NAME = "ดอลลาร์ฮ่องกง"
        CASE MENUCHOICE = "ดอลลาร์ฮ่องกง"
```

064622

```
            LOCATE ALL FOR NAME = "ดอลลาร์ฮ่องกง"
        CASE MENUCHOICE = '5'
            LOCATE ALL FOR NAME = "ฟรังก์สวิส"
        CASE MENUCHOICE = "ฟรังก์สวิส"
            LOCATE ALL FOR NAME = "ฟรังก์สวิส"
        CASE MENUCHOICE = '6'
            LOCATE ALL FOR NAME = "ปอนด์สเตอร์ลิง"
        CASE MENUCHOICE = "ปอนด์สเตอร์ลิง"
            LOCATE ALL FOR NAME = "ปอนด์สเตอร์ลิง"
        CASE MENUCHOICE = '7'
            LOCATE ALL FOR NAME = "มาร์กเยอรมัน"
        CASE MENUCHOICE = "มาร์กเยอรมัน"
            LOCATE ALL FOR NAME = "มาร์กเยอรมัน"
        CASE MENUCHOICE = '8'
            LOCATE ALL FOR NAME = "รีมินบี (หยวนจีน)"
        CASE MENUCHOICE = "รีมินบี (หยวนจีน)"
            LOCATE ALL FOR NAME = "รีมินบี (หยวนจีน)"
        CASE MENUCHOICE = '9'
            LOCATE ALL FOR NAME = "เยน"
        CASE MENUCHOICE = "เยน"
            LOCATE ALL FOR NAME = "เยน"
    ENDCASE
    DO MENEY2
  ENDIF
ENDDO
```

โปแกรม MONEY2.PRG

```
@ 05,10 SAY "      "
@ 05,10 SAY NAME
@ 10,10 SAY "ซื้อ" + STR(BUY,8,2)
@ 12,10 SAY "ขาย" + STR(SALE,8,2)
RETURN
```

โครงสร้างข้อมูล MONEY.DBF

```
Structure for database: D:money.dbf
Number of data records:
Date of last update   : 01/13/92
Field     Field Name     Type        Width    Dec
    1     NAME           Character      30
    2     BUY            Numeric         8      2
    3     SALE           Numeric         8      2
** Total **                             47
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ข.

โปรแกรมภาษาเสียงพูด

```
LANGUAGE NEWDOS:

WORDS

DOS_COMMAND  = clear scren     "CLS^OD", /*CLS and press ENTER*/

               1               "1^OD",

               2               "2^OD",

               3               "3^OD",

               4               "4^OD",

               5               "5^OD",

               6               "6^OD",

               7               "7^OD",

               8               "8^OD",

               9               "9^OD",

               0               "0^OD",

               usa             "ดอลลาร์สหรัฐ^OD",

               singapore       "ดอลลาร์สิงคโปร์^OD",

               malasia         "ดอลลาร์มาเลเซีย^OD",

               hongkong        "ดอลลาร์ฮ่องกง^OD"

               swiss           "ฟรังก์สวิส^OD"

               sterling        "ปอนด์สเตอร์ลิง^OD"

               mark            "มาร์กเยอรมัน^OD"

               reeminbee       "รีมินบี^OD"

               yen             "เยน^OD"

               exit            "เลิกการทำงาน^OD"

               ENTER           "^OD";  /* press enter     */

               backspace       "^08",  /* press backspace */
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
          stop          "^13",  /* <ctrl> S        */
          cancel        "^03",  /* <ctrl> C        */
GRAMMAR NEWDOS:
        IF DOS_COMMAND THEN DONE;
DONE:
        GOTO NEWDOS;
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ค.

การใช้งานระบบนำเข้าเสียงพูด

การเรียนรู้เสียงพูด

การเรียนรู้เสียงพูด เริ่มจากการใช้งานโปรแกรมภาษาเสียงพูด จะทำโดยการเรียกใช้โปรแกรมที่ชื่อ LR ซึ่งเมื่อเรียกโปรแกรม LR แล้ว จะมีหน้าจอดังรูปที่ ค.1 และโปรแกรมจะถามว่า Enter language pathname : เพื่อให้ป้อนชื่อไฟล์ที่จะทำการเรียนรู้ ในที่นี้ให้ชื่อว่า MONEY

หลังจากนั้นก็จะมีเมนูขึ้นมาดังรูปที่ ค.2 ซึ่งผู้ใช้จะทำการแปลโปรแกรมเสียก่อน โดยใช้ การเลือกหัวข้อ Compile Language แล้วจึงเลือกหัวข้อ Enroll Words ส่วนหัวข้อ Test Recognition ใช้ทดสอบว่าการเรียนรู้ได้ถูกต้องหรือไม่ ส่วนการแก้ไขเงื่อนไขต่างๆ ทำได้โดยเลือกเมนู Change Configuration เมื่อกำหนดทุกอย่างเรียบร้อย ก็เริ่มการทำงานโดยเลือกหัวข้อStart Recognition โปรแกรมจะฝังตัวในหน่วยความจำวิ่งเป็นแบบแบ็คกราวด์ รายละเอียดของแต่ละหัวข้อพอสังเขปมีดังนี้

Compile Language

จะเป็นการแปลไฟล์โปรแกรมภาษาเสียงพูด ที่จะทำการรับรู้ว่าถูกต้องตามหลักไวยากรณ์หรือไม่ โปรแกรมภาษาเสียงพูดอยู่ในภาคผนวก ข.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Enroll Words

จะเป็นการทำการเรียนรู้เสียงของผู้ใช้ ซึ่งก็คือการสร้างแบบอ้างอิงของเสียงขึ้นมา โดยการที่ผู้ใช้จะต้องเลือกคำที่จะทำการ Enroll จากไฟล์ ดังรูปที่ ค.3 เมื่อเลือกได้แล้วผู้ใช้จะต้องออกเสียงของคำนั้นตามต้องการ เมื่อเรียบร้อยแล้ว จะมีความยาวของเสียงแสดง ดังรูปที่ ค.4

Test Recognition

จะเป็นการทดสอบการรับรู้ว่าเมื่อออกเสียงตามที่ได้ทำการ Enroll แล้ว ผลที่ได้

I 12847495

จะเป็นอย่างไร

Change Configuration

จะเป็นการให้ผู้ใช้สามารถเปลี่ยนแปลงค่าต่าง ๆ ในการทำรับรู้ เช่น

- Set Enrollment Parameters
- Set Recognition Parameters
- Assign Board Number
- Choose I/O Device
- Modify Input Gain
- Modify Output Volume

ดังรูปที่ ค. 5

Start Recognitiony

จะเป็นการเริ่มการใช้งาน Language Recognition

<u>การรับรู้เสียงพูด</u>

ก่อนอื่นต้องเรียกโปรแกรม FOXPRO ก่อน แล้วจึงเรียกโปรแกรม MONEY.PRG เป็นโปรแกรมที่ใช้สอบถามข้อมูลจากฐานข้อมูล อัตราแลกเปลี่ยนเงินตรา (MONEY.DBF) เกี่ยวกับอัตราแลกเปลี่ยนเงินตราต่างประเทศ

โปรแกรมจะมีชื่อของเงินตราต่างประเทศสกุลต่าง ๆ ขึ้นมาให้ ดังนี้

1 ดอลลาร์สหรัฐ

2 ดอลลาร์สิงคโปร์

3 ดอลลาร์มาเลเซีย

4 ดอลลาร์ฮ่องกง

5 ฟรังก์สวิส

6 ปอนด์สเตอร์ลิง

7 มาร์กเยอรมัน

จุฬาลงกรณ์มหาวิทยาลัย

8 รีมินบี (หยวนจีน)

9 เยน (ต่อ 100 เยน)

0 เลิกการทำงาน

และจะถามผู้ใช้ว่าจะเลือกสอบถามถึงอัตราแลกเปลี่ยนเงินสกุลใด ซึ่งผู้ใช้ต้องใช้เสียง พูดว่าจะเลือกเงินสกุลใด โดยผู้ใช้อาจพูดเฉพาะเลขที่นำหน้า เช่น "1" หรือ "2" ก็ได้ หรือผู้ใช้ป้อนเป็นประโยคว่า "ดอลลาร์สหรัฐ" หรือ "ดอลลาร์สิงคโปร์" หรือ "เยน" ซึ่งเสียงนี้จะไปตรวจสอบกับไฟล์ MONEY.DBF ว่าตรงกับตัวแปรใด และจะนำข้อมูลของตัวแปรนั้นมาทำการพิมพ์แทนการพิมพ์ของผู้ใช้

เมื่อป้อนข้อมูลเรียบร้อยแล้วโปรแกรมก็จะนำข้อมูลที่ได้ไปทำการประมวลผล หรือค้นหาข้อมูลที่ต้องการ และนำข้อมูล หรืออัตราแลกเปลี่ยนเงินที่ค้นหาได้มาแสดงต่อไป ซึ่งในการแสดงผลนั้น จะทำโดยเรียกใช้โปรแกรมย่อย ที่ชื่อ MONEY2.PRG

โดยผลบนจอภาพที่ได้จะเป็นดังนี้

ดอลลาร์สหรัฐ

ซื้อ : 25.12

ขาย : 25.72

TI-SPEECH LANGUAGE RECOGNITION UTILITY

Enter language pathname : MONEY

Version 2.2

Copyright (c) 1986, Texas Instruments Incorporated.

All Rights reserved

TI-SPEECH is a trademark of Texas Instruments Incorporated.

จุฬาลงกรณ์มหาวิทยาลัย

Quit (ESC)

รูปที่ ค.1

TI-SPEECH LANGUAGE RECOGNITION UTILITY

Select a Command

Commands

Compile Language

Generate Test Sentences

Input Test Sentences

Enroll Words

Update Words

Test Recognition

Change Recognition

Start Recognition

Access Editor

Language pathname : MONEY

Template pathname : MONEY

Board Number : 0 Mike/Speaker Input Gain : 4 Output Volume : 9

Quit (ESC)

รูปที่ ค.2

TI-SPEECH LANGUAGE RECOGNITION UTILITY : Enroll Word

Enter Template pathname: MONEY

Words	Words Spellings	Length (ms)
USA		
HONGKONG		
MARK		
YEN		
ENTER		

Main Menu (ESC)

รูปที่ ค.3

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

TI-SPEECH LANGUAGE RECOGNITION UTILITY : Enroll Word

Please Say : USA

ENROLL WORDS IN A SENTENCES.	Words Spellings	Length (ms)
	USA	15

Main Menu (ESC)

รูปที่ ค.4

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

TI-SPEECH LANGUAGE RECOGNITION UTILITY : Change Configuration

Select a Command

Command	Recognition Parameter	
Set Enrollment Parameters	Minimum utterance detection threshol	256
Set Recognition Parameters	Inter utterance silence gap	25
Assign Board Number	Sentence accuracy threashold	-100
Choose I/O Device	Word accuracy threshold	3000
Modify Input Gain	Intermediate result flag	1
Modify Output Volume	Hypothesis timeout theshold	32639

Language pathname : MONEY

Template pathname : MONEY

Board Number : 0 Mike/Speaker Input Gain : 4 Outupt Volume : 9

Nochange (ESC)

รูปที่ ค. 5