

รายงานการวิจัยพัฒนาและวิศวกรรมฉบับสมบูรณ์
รหัสโครงการ NT-B-06-4C-18-411

ระยะที่ 1

โครงการวิจัย

การพัฒนาระบบพูดแทนพิมพ์ภาษาไทย

Development of a Thai Speech Dictation System

คณะผู้วิจัย

ผศ.ดร.มนตรี กาญจนเดชะ

ผศ.กฤษณ์ ไทยะวานิชย์

นายวรพต ชูกำเนิด

นางปฏิมากร จันทร์พริ้ม

โครงการวิจัยนี้ได้รับทุนอุดหนุนการวิจัย พัฒนา และวิศวกรรม
จากศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ
สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ
ประจำปีงบประมาณ 2545

บทคัดย่อ

เป้าหมายของโครงการวิจัยนี้คือการพัฒนาต้นแบบของโปรแกรมพูดแทนพิมพ์ โดยการพัฒนาจะเริ่มต้นจากการบันทึกเสียงพูดจากผู้พูดต่างๆ แล้วจึงนำไปรวมกับข้อมูลเสียงที่ได้จากการทำงานของกลุ่มวิจัยอื่น เพื่อจัดทำเป็นฐานข้อมูลเสียงภาษาไทยที่มีมาตรฐาน และสามารถนำไปแจกจ่ายให้กับนักวิจัยที่สนใจ หลังจากพัฒนาฐานข้อมูลเสียงเสร็จเรียบร้อยแล้ว ขั้นตอนต่อไปจะเป็นการนำฐานข้อมูลเสียงนี้มาใช้ในการพัฒนาต้นแบบของระบบรู้จำเสียงพูดภาษาไทย แบบคำต่อเนื่องและไม่ขึ้นกับผู้พูด โดยจะใช้ซอฟต์แวร์ HTK เป็นหลักเพื่อการสร้างตัวรู้จำเสียงพูด ในขั้นตอนสุดท้ายจะเป็นการพัฒนาซอฟต์แวร์ต้นแบบของระบบพูดแทนพิมพ์ภาษาไทย โดยจะเน้นการวิจัยไปที่การพัฒนาซอฟต์แวร์ที่มีหน้าที่การทำงานเช่นเดียวกับซอฟต์แวร์ HTK แต่มีความเหมาะสมกว่าในเชิงของการนำไปใช้ในโปรแกรมประยุกต์ อย่างไรก็ตามรายงานนี้เป็นรายงานฉบับสมบูรณ์ของโครงการวิจัยในระยะที่ 1 ซึ่งมีเป้าหมายเพื่อรวบรวมเสียงพูดจากผู้พูดจำนวน 100 คน เพื่อสร้างฐานข้อมูลเสียงพูดสำหรับการวิจัยในระยะต่อไป

Abstract

This project proposes development of a working prototype of speech dictation software. The project will begin with recording of speech from various speakers. The speech data will be transcribed, verified and combined with the data obtained from other research groups. Thus a standard Thai speech corpus will be created and then distributed to interested researchers. Once a speech corpus is obtained, a prototype of a speaker independent, continuous Thai speech recognition system will be developed. Freely available speech recognition software, HTK, will be used in this phase to build a complete recognizer. Use of HTK will enormously reduce the development time of the project and will evaluate the feasibility of developing a high-performance Thai speech recognition system. The last step will be to write speech recognition software that has similar functionality to HTK and then utilize it in the speech dictation system prototype which is the final goal of the project. This report describes the first phase of the project which concerns a recording of speech sample from 100 speakers. The goal is to build a speech corpus which will be used in the next phases of our research project.