

รายงานวิจัย

ระบบรู้จำเสียงพูดตัวเลขภาษาไทยแบบต่อเนื่องโดยใช้ แบบจำลองยิตเดนมาร์คอฟ

An Application of Cepstral/Temporal Features for Connected
Thai Numeral Recognition with Hidden Markov Models

โดย

มนตรี กาญจนะเดชะ
กรดา เขมะพันธุ์มนัส
รุ่งศิริ ธีรวิธานนท์

ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์

งานวิจัยนี้ได้รับงบประมาณสนับสนุนจากเงินรายได้คณะวิศวกรรมศาสตร์
มหาวิทยาลัยสงขลานครินทร์

สิงหาคม 2545

ชื่อโครงการ	ระบบรู้จำเสียงพูดตัวเลขภาษาไทยแบบต่อเนื่องโดยใช้แบบจำลองฮิดเดนมาร์คอฟ An Application of Cepstral/Temporal Features for Connected Thai Numeral Recognition with Hidden Markov Models
ผู้วิจัย	นายมนตรี กาญจนะเดชะ นางสาวกรดา เขมะพันธ์มณี นางสาวรุ่งศิริ ธีรวิทยานนท์
หน่วยงาน	ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์

บทคัดย่อ

โครงการวิจัยนี้เป็นการพัฒนาระบบรู้จำเสียงพูดตัวเลขภาษาไทยแบบต่อเนื่องโดยใช้แบบจำลองฮิดเดนมาร์คอฟร่วมกับลักษณะเด่นของเสียงซึ่งอยู่ในรูปของสัมประสิทธิ์เคปสตรอลความถี่เมล ขั้นตอนแรกของการวิจัยคือการเก็บตัวอย่างเสียงจากผู้พูดต่างๆ โดยเสียงพูดที่บันทึกได้จะถูกนำไปกำกับหน่วยเสียงเพื่อกำหนดขอบเขตของคำพูดแต่ละคำ ขั้นตอนต่อไปเป็นการสร้างตัวรู้จำเสียงพูดโดยใช้แบบจำลองฮิดเดนมาร์คอฟด้วยโปรแกรม HTK จากนั้นได้ทำการทดสอบความแม่นยำของตัวรู้จำ และปรับค่าพารามิเตอร์ต่างๆ เพื่อให้ผลการรู้จำดีที่สุด ขั้นตอนสุดท้ายคือการออกแบบและสร้างโปรแกรมตัวรู้จำเสียงพูดที่สามารถทำงานได้ในเวลาจริงภายใต้ระบบปฏิบัติการ Windows จากผลการทดลองสรุปได้ว่าระบบรู้จำเสียงพูดตัวเลขภาษาไทยที่พัฒนาขึ้นมีความแม่นยำสูง และสามารถปรับปรุงและนำไปในงานอื่นๆ ได้

Project Title	An Application of Cepstral/Temporal Features for Connected Thai Numeral Recognition with Hidden Markov Models
Investigators	Dr. Montri Karnjanadecha Miss Korada Kemaphanmanus Miss Rungsiri Thirawitayanon
Department	Department of Computer Engineering, Faculty of Engineering, Prince of Songkla University

Abstract

This report presents a development of a real-time continuous Thai digit recognition system using Hidden Markov Model with Mel-frequency cepstral coefficients. The first phase of this work concerned collection of speech data from various speakers. All recorded speech was transcribed in order to specify boundaries of each word. Next phase was to build and train a continuous speech recognizer using the available speech data. The recognizer was built using Hidden Markov Model Toolkit. The performance of the recognizer was evaluated and model parameters were tuned to achieve best accuracy. After a good set of speech models was obtained, software was designed and implemented for the recognizer to run in real-time under Windows operating systems. Experimental results indicated that the system yields high accuracy and it could be modified to fit other applications.