A PROTOTYPE OF SPEECH RECOGNITION SYSTEM FOR THAI LANGUAGE WITH SPEAKER INDEPENDENCE

THANYARAT PRUTPAPOP

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE
(COMPUTER SCIENCE)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2008

COPYRIGHT OF MAHIDOL UNIVERSITY

Copyright by Mahidol University

ระบบต้นแบบการรู้จำเสียงพูดภาษาไทยโดยไม่ขึ้นกับผู้พูด (A PROTOTYPE OF SPEECH RECOGNITION SYSTEM FOR THAI LANGUAGE WITH SPEAKER INDEPENDENCE)

ธัญรัตน์ พฤฒปภพ 4737241 SCCS/M

วท.ม. (วิทยาการคอมพิวเตอร์)

คณะกรรมการควบคุมวิทยานิพนธ์ : ศุภชัย ตั้งวงศ์ศานต์, Ph.D., สุกัญญา พงษ์สุภาพ, Ph.D., ชมทิพ พรพนมชัย, Ph.D.

บทคัดย่อ

งานวิจัยนี้เป็นกา รพัฒนาต้นแบบของระบบรู้จำเสียงพูดภาษาไทยโดยไม่ขึ้นกับผู้พูด ซึ่งการพัฒนาตัวต้นแบบนั้นใช้หลักการของ Hidden Markov Model ในการพัฒนาเพื่อให้ได้ Acoustic model ที่มีประสิทธิภาพและมีอัตราการรู้จำสูง โดยงานวิจัยนี้ได้นำ HTK (Hidden Markov Toolkit) ซึ่งเป็นเครื่องมือที่ใช้ในการสร้างระบบรู้จำเสียงพูด มาช่วยในการสร้าง acoustic model การ training และ การ recognition โดยทำการประยุกต์ใช้ HTK ให้เหมาะสมกับเสียงภาษาไทย โดยใช้หลักการของ International Phonetic Alphabet (I.P.A) และทำการศึกษา Syntactic และ Semantic ในภาษาไทยที่มีโครงสร้างไม่ซับซ้อนมากนัก เพื่อนำไปใช้ในการเพิ่มอัตราความถูกต้องของระบบ

การพัฒนาตัวต้นแบบใช้บุคคลจำนวน 20 คน แบ่งเป็นชาย 10 คน หญิง 10 คน สำหรับการ Training ใช้ชาย หญิงอิสระที่ไม่อยู่ในกลุ่ม Training จำนวน 6 คน สำหรับการทดสอบระบบ โดยมีคำศัพท์ในระบบทั้งหมด 128 คำ และนำคำศัพท์ดังกล่าวมาสร้างเป็นบทสนทนาโต้ตอบกันระหว่างชายหญิงในการทดลอง ผลงานการวิจัยในระดับ Acoustic มีผลเบื้องต้นดังนี้ 55.21% สำหรับ Top 1 และ 57.39% สำหรับ Top 3 ผลจากการใช้ Syntactic Rule ส่งผลให้ accuracy เพิ่มขึ้นเป็น 66.52% สำหรับ Top 1 และ 73.04% สำหรับ Top 3 และผลจากการใช้ Semantic Rule อย่างง่าย ปรับปรุงแล้วส่งผลให้อัตราความถูกต้องเพิ่มขึ้นเป็น 76.56% สำหรับ Top 1 และ 86.52% สำหรับ Top 3

สรุประบบต้นแบบของระบบรู้จำเสียงพูดภาษาไทยโดยไม่ขึ้นกับผู้พูดที่ได้ออกแบบและพัฒนาขึ้นสามารถบรรลุตามวัตถุประสงค์ที่ตั้งไว้

108 หน้า

Copyright by Mahidol University

A PROTOTYPE OF SPEECH RECOGNITION SYSTEM FOR THAI LANGUAGE WITH SPEAKER INDEPENDENCE

THANYARAT PRUTPAPOP 4737241 SCCS/M

M.Sc.(COMPUTER SCIENCE)

THESIS ADVISORS: SUPACHAI TANGWONGSAN, Ph.D., SUKANYA PHONGSUPHAP, Ph.D., CHOMTIP PORNPANOMCHAI, Ph.D.

ABSTRACT

This research presents the development of a prototype of speech recognition for the Thai language with the class of speaker independence. The main concept is based on the Hidden Markov Model (HMM) on the acoustic level. The system makes use of a tool known as Hidden Markov toolkit, which is the instrument employed to generate the model in both steps of training and recognition. However, as the tool was originally designed for English language, in order to make it applicable, Thai speech syllables had to be first mapped into standard codes of International Phonetic Alphabet (IPA) accordingly. The next steps were the application of syntactic and semantic rules in the prototype in order to improve the accuracy rate of recognition.

In the experiment, 20 people, 10 males and 10 females were invited, and a set of 128 simple vocabulary words commonly formed at the Primary 4 level were used for training. Then 6 new people were invited to conduct the testing by speaking several sets of conversation in meaningful sentences based on those 128 vocabularies. The results showed that on the acoustic level, only 55.21% (Top 1) and 57.39% (Top 3) are obtained with no surprise. However, after applying the syntactic rules for further processing, the result shows a better improvement as 66.52% (Top 1) and 73.04% (Top 3) were obtained. Finally, with the semantic rules for filtration in the process, the results showed a significant improvement as 75.65% (Top 1) and 86.52% (Top 3) were obtained respectively.

In conclusion, the developed prototype has met the set objectives accordingly and satisfactorily.

KEY WORDS: SPEECH RECOGNITION/ HIDDEN MARKOV MODEL

108 pp.

Copyright by Mahidol University