

ระบบควบคุมคอมพิวเตอร์ด้วยเสียงพูดภาษาไทยโดยใช้
เทคนิค การวิเคราะห์สเปกตรัมและโครงข่ายประสาทเทียม



นายพงษ์ศักดิ์ ชูงาม

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2544

ISBN 974-03-1639-5

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

I 20325137

28 เม.ย. 2547

A COMPUTER CONTROLLED SYSTEM BY THAI SPEECH USING
SPECTRUM ANALYSIS AND AN ARTIFICIAL NEURAL NETWORK

Mr. Pongsak Choo-Ngam

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science

Department of Computer Engineering

Faculty of Engineering


Chulalongkorn University

Academic year 2001

ISBN 974-03-1639-5

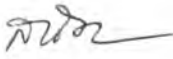
หัวข้อวิทยานิพนธ์ ระบบควบคุมคอมพิวเตอร์ด้วยเสียงพูดภาษาไทย โดยใช้เทคนิคการ
วิเคราะห์สเปกตรัมและโครงข่ายประสาทเทียม
โดย นาย พงษ์ศักดิ์ ชูงาม
สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร.สาธิต วงศ์ประทีป

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วน
หนึ่งของการศึกษาตามหลักสูตรปริญญาโทบริหารธุรกิจ

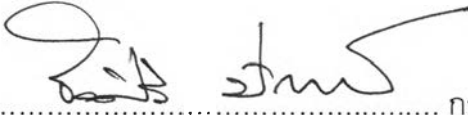
 คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สมศักดิ์ ปัญญาแก้ว)

คณะกรรมการสอบวิทยานิพนธ์

 ประธานกรรมการ
(อาจารย์ ดร.อาทิตย์ ทองทักษ์)

 อาจารย์ที่ปรึกษา
(ผู้ช่วยศาสตราจารย์ ดร.สาธิต วงศ์ประทีป)

 กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล)

 กรรมการ
(อาจารย์ ชัยศิริ ปันนิตานนท์)

พงษ์ศักดิ์ ชูงาม : ระบบควบคุมคอมพิวเตอร์ด้วยเสียงพูดภาษาไทยโดยใช้ เทคนิค การวิเคราะห์
สเปกตรัม และโครงข่ายประสาทเทียม (A COMPUTER CONTROLLED SYSTEM BY THAI
SPEECH USING SPECTRUM ANALYSIS AND AN ARTIFICIAL NEURAL NETWORK)

อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร. สราจิต วงศ์ประทีป , 64 หน้า. ISBN 974-03-1639-5

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อพัฒนาวิธีการรู้จำเสียงพูด โดยการวิเคราะห์เชิงความถี่ เพื่อหาลักษณะ
เด่นของเสียงพูดในรูปแบบของแถบความถี่ และ นิวรอลเน็ตเวิร์กแบบแบ็กพรอพาคชัน โดยใช้แถบความถี่เป็น
ข้อมูลอินพุตสำหรับ นิวรอลเน็ตเวิร์ก และพัฒนาโปรแกรมต้นแบบ เพื่อแสดงการทำงานของระบบจริง

ชุดข้อมูลเสียงที่ใช้ทดสอบ ประกอบด้วยเสียง 50 เสียง โดยกำหนดเพื่อแทนคำสั่งหรือปุ่มบนแป้นกด
เมื่อโปรแกรมได้รับเสียง โปรแกรมจะกำหนดจุดเริ่มต้นของเสียง และคำนวณหาแถบความถี่ของเสียง แถบ
ความถี่ จะเป็นข้อมูลรับเข้าของ นิวรอลเน็ตเวิร์ก เพื่อหาวิธีแบบที่เข้ากันได้ กับข้อมูลที่มีการสอนไว้

ผลจากการทดลอง ระบบสามารถรู้จำเสียงถูกต้อง 87.7 เปอร์เซ็นต์ พบปัญหาของระบบอยู่ที่ระบบรับ
สัญญาณเสียง การคำนวณแถบความถี่ เป็นารคำนวณเป็นแบบช่วงเวลา ดังนั้นในบางครั้งกรอบของข้อมูลรับ
เข้า ไม่สามารถครอบคลุมสัญญาณเสียง แถบความถี่จะผิดพลาด ถ้าสัญญาณเสียงไม่สมบูรณ์ โปรแกรมตัวอย่าง
เป็นต้นแบบของการพัฒนา การรู้จำคำพูดแบบต่อเนื่อง

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต
สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่ออาจารย์ที่ปรึกษา
ปีการศึกษา 2544

4171461521 : MAJOR COMPUTER SCIENCE

SPEECH RECOGNITION / NEURAL NETWORK / FAST FOURIER TRANSFORM / FFT

PONGSAK CHOO-NGAM : A COMPUTER CONTROLLED SYSTEM BY THAI SPEECH
USING SPECTRUM ANALYSIS AND AN ARTIFICIAL NEURAL NETWORK


THESIS ADVISOR : ASSISTANT PROF. SARTID VONGPRADHIP, Ph.D. 64 pp.

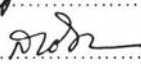
ISBN 974-03-1639-5

The purpose of this research is to develop a speech recognition algorithm using frequency domain analysis for specify pattern of spectrum and back propagation neural network . Results of Spectrum analysis are feeded to neural network. An example program is developed to show the process of algorithm in real system.

A set of 50 speeches is used and these speeches are window commands or key when program receive speech. The program find a starting point of speech and calculate frequency spectrum of the speech. Frequency spectrum is input pattern for neural network and the results of neural network are matched with the training pattern.

From The results, the system can recognize speeches with 87.7 % Correction. It is found that a problem is in the input signal system. Calculation of short time spectrum can not cover speeches signal. A spectrum of frequency is lost if speeches signal are not completed. An example program is modeled to develop a continuous speech recognition.

Department/Program Computer Engineering Student's signature 

Field of study Computer Science Advisor's signature 

กิตติกรรมประกาศ

ข้าพเจ้าขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร. สราจิต วงศ์ประทีป อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาสละเวลาให้คำปรึกษา และคำแนะนำต่าง ๆ ในการทำวิทยานิพนธ์ฉบับนี้ ข้าพเจ้าขอขอบคุณ ดร. บุญเสริม กิจศิริกุล ที่ได้ให้คำแนะนำและคำวิจารณ์ที่เป็นประโยชน์เกี่ยวกับการทำวิทยานิพนธ์ฉบับนี้

ข้าพเจ้าขอขอบพระคุณ นายวุฒิพงษ์ พรสุขจันทร์ ที่ได้ให้ข้อมูลที่เป็นประโยชน์ ในการทำวิทยานิพนธ์ ขอขอบคุณที่ทำงานทุกแห่งที่ให้โอกาส ข้าพเจ้า และน้องๆ เพื่อน ๆ ทุกคน ที่ให้กำลังใจและช่วยเหลือ สุดท้ายข้าพเจ้าขอกราบขอบพระคุณ คุณแม่ที่เป็นกำลังใจให้ข้าพเจ้าเสมอ รักคุณแม่เสมอ

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ	ฉ
สารบัญ	ช
บทที่	
1. บทนำ	1
1.1 แนวเหตุผล	1
1.2 วัตถุประสงค์	1
1.3 ขอบเขต	2
1.4 ประโยชน์	2
2. หลักการรู้จำเสียงพูด	3
2.1 รูปแบบการพูดในระบบรู้จำเสียงพูด	3
2.2 พื้นฐานของระบบรู้จำเสียง	4
2.3 โครงสร้างระบบรู้จำเสียงพูด	5
2.4 การแปลงสัญญาณอนาล็อกเป็นสัญญาณดิจิทัล	6
2.5 การจัดรูปแบบของสัญญาณก่อนการประมวลผล	7
2.6 วิธีการวิเคราะห์คุณลักษณะของเสียงพูด	9
2.7 การหาความคล้ายคลึงกันของรูปแบบ	30
3. ระบบวิเคราะห์และรู้จำเสียงพูด	35
3.1 โครงสร้างระบบ	35
3.2 การทำงานของระบบรู้จำเสียงพูด	36
3.3 โปรแกรม วิเคราะห์และรู้จำเสียงพูด	40
3.4 การใช้งานโปรแกรมวิเคราะห์และรู้จำเสียงพูด	43
4. การทดลองและผลการทดลอง	45
4.1 อุปกรณ์และองค์ประกอบในการทดลอง	45
4.2 ข้อมูลที่ใช้ในการทดลอง และวิธีทดลอง	45
4.3 วันที่ผลการทดลอง	47
5. สรุปผลการทดลองและข้อเสนอแนะ	59
5.1 สรุปผลการทดลอง	59
5.2 ปัญหาและอุปสรรค	59
5.2 ข้อเสนอแนะ	60
รายการอ้างอิง	61
ภาคผนวก	63
ประวัติผู้เขียนวิทยานิพนธ์	64