

หัวข้อวิทยานิพนธ์	การพัฒนาคุณลักษณะสำคัญของสัมประสิทธิ์เซปสตรัมบนสเกลเมล สำหรับระบบการรู้จำเสียงพูด
นักศึกษา	นางสุจิตรา จินะวงษ์
รหัสประจำตัว	034820402015 -3
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา	วิศวกรรมไฟฟ้า แขนงวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม
ปีการศึกษา	2552
อาจารย์ผู้ควบคุมวิทยานิพนธ์	ดร. จักรี ศรีนนท์ฉัตร

บทคัดย่อ

ระบบโครงข่ายประสาทเทียม (Artificial Neural Networks) ได้มีการประยุกต์ใช้ในงานวิจัยทางด้านการรู้จำเสียงพูด ทั้งนี้ระบบโครงข่ายประสาทเทียมมีความจำเป็นที่จะต้องเรียนรู้และจดจำคุณลักษณะของสัญญาณเสียงพูด โดยการเลือกใช้พารามิเตอร์ของเสียงคำพูด การเลือกจำนวนโนดเพื่อปรับโครงสร้างของโครงข่ายประสาทเทียมให้มีความเหมาะสมกับสัญญาณนั้นๆ ทั้งนี้จากการเลือกใช้พารามิเตอร์ของเสียงคำพูดที่ไม่เหมาะสม ทำให้จำนวนโนดในชั้นอินพุตของโครงข่ายประสาทเทียมมีจำนวนมาก ส่งผลให้การคำนวณและการฝึกฝนโครงข่ายประสาทเทียมใช้เวลานานเกินไป

วิทยานิพนธ์นี้นำเสนอการลดจำนวนคุณลักษณะสำคัญของสัมประสิทธิ์เซปสตรัมบนสเกลเมล (Mel Frequency Cepstral Coefficient : MFCC) สำหรับระบบการรู้จำเสียงพูด เพื่อใช้ในการฝึกฝนโครงข่ายประสาทเทียม การดึงคุณลักษณะเด่นของสัญญาณเสียงพูด จะใช้วิธีการหาค่าสัมประสิทธิ์เซปสตรัมบนสเกลเมล โดยแบ่ง เป็น 4, 5, 6, 7 และ 10 อันดับ ในส่วนของการรู้จำ จะใช้ระบบโครงข่ายประสาทเทียม ประเภทเพอเซปตรอนหลายชั้น (Multilayer perceptron : MLP) และ มีการเรียนรู้แบบแพร่กลับ (Backpropagation) สัญญาณเสียง ทางด้านอินพุต ได้มาจากผู้ชาย 100 คน และผู้หญิง 100 คน พุคคนละ 2 รอบ จะได้ข้อมูลเสียงพูดละ 400 เสียง นำมาฝึกฝนและทดสอบโครงข่ายประสาทเทียม

ผลการทดลองพบว่า อัตราความถูกต้องของการรู้จำโดยใช้ MFCC 4 คือ 87.1% (ผู้หญิง) ใช้เวลาในการฝึกฝนคือ 11 นาที และ 80.2 % (ผู้ชาย) ใช้เวลาในการฝึกฝนคือ 12 นาที ส่วน MFCC 10 ให้อัตราการรู้จำที่ 90.9% (ผู้หญิง) และ 85.3% (ผู้ชาย) ใช้เวลาในการฝึกฝนเท่ากันคือ 22 นาที ซึ่ง MFCC4 ใช้เวลาน้อยกว่า MFCC10 เกือบ 2 เท่า

คำสำคัญ : โครงข่ายประสาทเทียม , สัมประสิทธิ์เซปสตรัมบนสเกลเมล, การรู้จำเสียงพูด

Thesis Title : IMPROVEMENT MFCC FEATURE FOR SPEECH
RECOGNITION SYSTEM

Student Name : Mrs. Suchitra Jeenawong

Student ID : 034820402015 -3

Degree Award : Master of Engineering

Study Program : Electrical Engineering
(Electronic and Telecommunication Engineering)

Academic Year: 2009

Thesis Advisor/s : Dr. Jakkree Srinonchat

ABSTRACT

An artificial neural network (ANNs) has been recently applied in the speech recognition research area. ANNs always requires the training and recognition process using the parameters of speech signal, number of neurons. However, the non-appropriate number of parameters for training ANNs is directly affected to the time processing and training process which will take long time to process.

This thesis presents the reducing speech feature based on Mel Frequency Cepstral Coefficient to train ANNs for speech recognition system. There are different number of MFCC coefficients such as 4, 5, 6, 7 and 10 to train the ANNs. The Multilayer perceptron and backpropagation technique of ANNs are applied for recognition process. Speech signal from 200 peoples (100 men and 100 women) are selected to be input signal which each speaker speaks two times. Therefore there are 400 input signal to train the ANNs.

The results shown that, in the term of MFCC4, the system provides the accuracy ratio 87.1% using 11 minutes for training and 80.2% using 12 minutes for training for female and male speech signal respectively. In the term of MFCC10, the system provides the accuracy ratio 90.9% and 85.3% for female and male speech signal respectively and both cases use 22 minutes. It can notice that MFCC 4 using time to training less than MFCC 10 approximately 2 times.

Keywords : Artificial neural networks , Mel frequency cepstral coefficient, Speech recognition

กิตติกรรมประกาศ

วิทยานิพนธ์นี้สำเร็จลุล่วงไปได้ ด้วยความช่วยเหลือเป็นอย่างดีของ ดร.จักรี ศรีนนท์ฉัตร ซึ่งเป็นอาจารย์ผู้ควบคุมวิทยานิพนธ์ และสมาชิกในห้องปฏิบัติการและวิจัยทางด้านการประมวลผลสัญญาณทุกท่านสำหรับคำแนะนำและข้อคิดเห็นต่างๆ ทั้งในส่วนของการทำงานวิจัยและในส่วนของการทำรูปเล่มวิทยานิพนธ์ ขอขอบพระคุณคณาจารย์ทุกท่านที่ได้อบรมสั่งสอนและประสิทธิ์ประสาทความรู้วิชาการต่างๆ ให้ตลอดระยะเวลาที่ผ่านมา

ขอขอบคุณนักศึกษาระดับปริญญาโท สาขาวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม คณะวิศวกรรมศาสตร์ นักศึกษาคณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี สำหรับข้อมูลเสียงทั้งหมดที่นำมาใช้ในงานวิจัยและวิทยานิพนธ์นี้

ขอขอบคุณมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรีที่ให้การสนับสนุนทางการเงิน ทำยที่สุคนธ์ขอขอบคุณบิดามารดา ญาติพี่น้อง สามีและลูกสาวสุดที่รักสำหรับกำลังใจที่มีให้เสมอมา

สุจิตรา จินะวงษ์

30 มีนาคม 2553

