



**NONLINEAR PREDICTION WITH
NEURAL NETWORK FOR SPEECH CODING**

WATCHARIN MONPENGPINIT



อธิปัทนาร

จาก

บัณฑิตวิทยาลัย มหาวิทยาลัยมหิดล

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

2001

ISBN 974-04-0171-6

COPYRIGHT OF MAHIDOL UNIVERSITY

TH

W 324 n

2001

4237697SCCS/M : MAJOR : COMPUTER SCIENCE ; M.Sc. (COMPUTER SCIENCE)

KEY WORD : DIGITAL SPEECH CODING, NON-LINEAR PREDICTION, NEURAL NETWORK

WATCHARIN MONPENGPINIT : NONLINEAR PREDICTION WITH NEURAL NETWORK FOR SPEECH CODING. THESIS ADVISORS : DAMRAS WONGSAWANG, Ph.D., SUKANYA PHONGSUPHAP, Ph.D. 115p. ISBN 974-40-0171-6

Digital speech coding is a method to transform analog speech waveform into digital speech waveform. One recent effective approach utilizes a neural network model as a digital speech coder. Neural network is a computer model that emulates human brain functions. Its architecture consists of three layers: input, hidden and output layer. Radial basis function network (RBFN) is a perfect feed forward neural network model that is applied as a nonlinear prediction speech coding algorithm. However, it is a noisy speech coder that produced some distortion of speech waveform.

This thesis proposed the use of a speech coding algorithm called Feedback Radial Basis Function Network (FRBFN). It used feedback information from output layer of the network as an additional input of the hidden layer. The methods of mathematical simulation and prototype of speech coding simulation were employed to test and analyze the results of this proposed algorithm comparing with the existing RBFN model.

The results showed that FRBFN can improve the prediction performance of radial basis function network with small increasing processing unit of the hidden layer of the neural network model. It also produced less distortion speech waveform than the existing speech coder using RBFN algorithm. The details of the FRBFN was described and analyzed. The results of its experiments were presented and compared among other speech coders. Finally, further development for future works are suggested.

4237697SCCS/M: สาขาวิชา : วิทยาการคอมพิวเตอร์; วท.ม. (วิทยาการคอมพิวเตอร์)

วัชรินทร์ มนพ่งพินิจ : การแปลงเสียงพูดจากสัญญาณอนาลอกเป็นสัญญาณดิจิทัล ด้วยวิธีการคาดหมายแบบไม่ใช้เส้นตรง โดยใช้เครือข่ายสมองเทียม (NONLINEAR PREDICTION WITH NEURAL NETWORK FOR SPEECH CODING). คณะกรรมการควบคุมวิทยานิพนธ์ : คำรัส วงศ์สว่าง, Ph.D., สุกัญญา พงษ์สุภาพ, Ph.D. 115 หน้า. ISBN 974-04-0171-6

การแปลงเสียงพูดจากสัญญาณอนาลอกให้เป็นสัญญาณดิจิทัล คือการแปลงสัญญาณที่มีความต่อเนื่องทางด้านเวลาและลูกคลื่นของสัญญาณ ให้เป็นสัญญาณแบบไม่ต่อเนื่อง ซึ่งเรียกว่าสัญญาณดิจิทัล วิธีการแปลงสัญญาณดังกล่าวมีหลากหลายรูปแบบ แต่ในปัจจุบันวิธีการหนึ่งที่ได้รับการยอมรับ ว่ามีประสิทธิภาพสูงคือการใช้เครือข่ายสมองเทียมมาช่วยในการแปลงรูปแบบของสัญญาณ เครือข่ายสมองเทียมคือการสร้าง Computer model เพื่อใช้จำลอง Function ในการทำงานของสมองมนุษย์ ซึ่งสถาปัตยกรรมของเครือข่ายสมองเทียม ประกอบไปด้วยสามชั้น คือ ชั้นล่างสุดเรียกว่า Input layer, ชั้นถัดไปเรียกว่า Hidden layer และชั้นสุดท้ายเรียกว่า Output layer เครือข่ายสมองเทียมมี Model ที่ดี model หนึ่ง เรียกว่า Radial basis function network (RBFN) ถูกนำมาใช้เป็นวิธีการแปลงสัญญาณเสียง โดยที่ทิศทางการไหลของ Information เป็นไปในทิศทางเดียวไม่มีการย้อนกลับ เราจึงเรียก RBFN ว่าเป็น Feed forward neural network model และถึงแม้ว่า RBFN จะเป็น Model ที่ดี แต่ก็ยังมีข้อผิดพลาดในการแปลงสัญญาณเสียงพูดอยู่

งานวิจัยนี้ได้นำเสนอวิธีการที่มีชื่อเรียกว่า Feedback Radial Basis Function Network (FRBFN) ซึ่งเป็นเครือข่ายสมองเทียม ที่มีทิศทางการไหลของ Information กลับจาก Output layer มาสู่ Hidden layer โดยที่ Information ที่ไหลกลับมานี้ เป็นเสมือน Input พิเศษที่ป้อนเข้าสู่ Hidden layer ในงานวิจัยนี้ได้ทำการทดลอง และวิเคราะห์ผลของการทดลองโดยการจำลองการทำงานทางด้านคณิตศาสตร์ (Mathematical simulation) และการพัฒนาแบบจำลองสำหรับการแปลงสัญญาณอนาลอกของเสียงพูดให้เป็นสัญญาณดิจิทัล (Prototype of speech coding simulation) ผลของการทดลองแสดงให้เห็นว่าข้อผิดพลาดในการแปลงสัญญาณเสียงเป็นสัญญาณดิจิทัลลดลง ในขณะที่จำนวนของ process unit เพิ่มขึ้นเพียงเล็กน้อย เมื่อเทียบกับการแปลงสัญญาณเสียงเป็นสัญญาณดิจิทัล โดยใช้ RBFN model และได้แนะนำจนถึงวิธีการที่จะทำการพัฒนางานวิจัยนี้ต่อไปในอนาคต