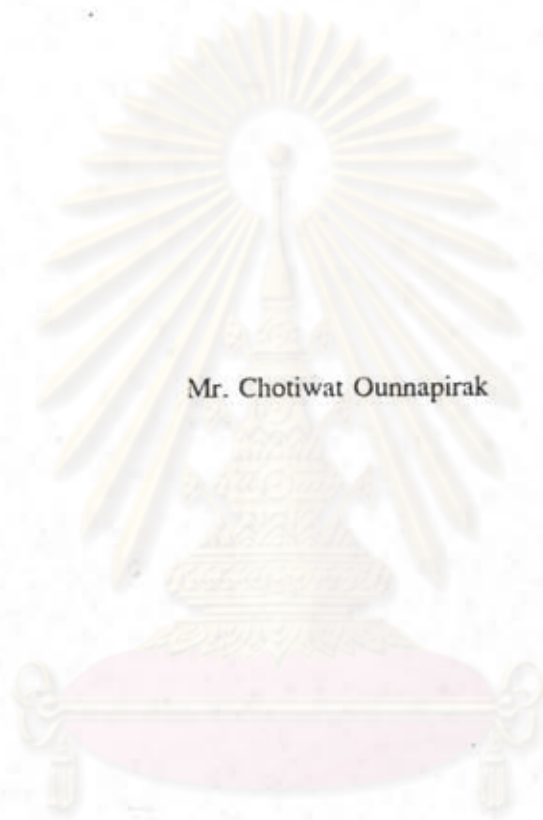




Speech Compression Using Wavelet Transform and LPC Vector Quantization



Mr. Chotiwat Ounnampirak

ศูนย์วิทยทรัพยากร

จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Engineering

Department of Electrical Engineering

Graduate School

Chulalongkorn University

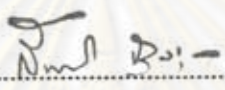
1995

ISBN 974-632-415-2

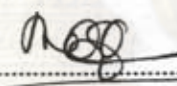
หัวข้อวิทยานิพนธ์ การบิบบย่อเสียงพูดโดยใช้การแปลงเวฟเลตและแอลพีซีเวกเตอร์ควอนไทเซชัน  
โดย นาย โชติวัฒน์ อุณนาภิรักษ์  
ภาควิชา วิศวกรรมไฟฟ้า  
อาจารย์ที่ปรึกษา รองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล


---

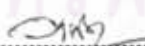
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัยฉบับนี้เป็นส่วนหนึ่ง  
ของการศึกษาตามหลักสูตรปริญญาโทบัณฑิต

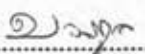
  
..... คณบดีบัณฑิตวิทยาลัย  
( รองศาสตราจารย์ ดร. สันติ อุงสุวรรณ )

คณะกรรมการสอบวิทยานิพนธ์

  
..... ประธานกรรมการ  
( รองศาสตราจารย์ ดร. ณรงค์ อยู่ถนอม )

  
..... อาจารย์ที่ปรึกษา  
( รองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล )

  
..... กรรมการ  
( อ. ดร. วาতিต เบนจพลงกุล )

  
..... กรรมการ  
( ดร. บวรกุล จิตต์ประเสริฐ )

พิมพ์ต้นฉบับบทคัดย่อวิทยานิพนธ์ภายในกรอบสี่เหลี่ยมนี้เพียงแผ่นเดียว

โชติวัฒน์ อุณาภิรักษ์ : การบีบย่อเสียงพูดโดยใช้การแปลงเวฟเลตและแอลพีซีแวลูแวลูแวลู  
ควอนไทเซชัน ( SPEECH COMPRESSION USING WAVELET TRANSFORM AND LPC  
VECTOR QUANTIZATION ) อ. ที่ปรึกษา : รศ.ดร.สมชาย จิตะพันธ์กุล 48 หน้า.  
ISBN 974-632-415-2

วิทยานิพนธ์นี้จะนำเสนอวิธีการใหม่ในการบีบย่อเสียงพูด ซึ่งสามารถเพิ่มอัตราการบีบย่อของวิธี CELP โดยที่ยังคงคุณภาพเสียงที่ดีไว้ โดยจะใช้เวฟเลตแตกแยกมาบีบย่อสัญญาณเศษเหลือ (residual signal) ที่ได้จากการบีบย่อเสียงพูดโดยวิธี CELP (Code-Excited Linear Prediction) แล้วตัดส่วนที่มีความสำคัญน้อยออกไปโดยพิจารณาเลือก level ที่ดีที่สุด

ปัจจัยที่มีผลต่ออัตราการบีบย่อและคุณภาพเสียงในการบีบย่อเสียงพูดโดยวิธีที่เสนอในวิทยานิพนธ์นี้ได้แก่ order ของ mother wavelet ที่ใช้, ขนาดของ LPC VQ codebook, ขนาดของ excitation VQ codebook, และค่า threshold ที่ใช้ในการพิจารณาคัดองค์ประกอบของสัญญาณที่มีความสำคัญน้อย

ระบบการบีบย่อเสียงพูดที่สร้างขึ้นสามารถให้ bit-rate ประมาณ 9,600 bps ที่คุณภาพเสียงที่ดี โดยสามารถเพิ่มอัตราการบีบย่อของการบีบย่อเสียงพูดโดยวิธี CELP ได้มากกว่า 1.5 เท่า โดยใช้เงื่อนไขการบีบย่อที่เหมาะสม

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา..... วิศวกรรมไฟฟ้า  
สาขาวิชา..... DSP  
ปีการศึกษา..... 2537

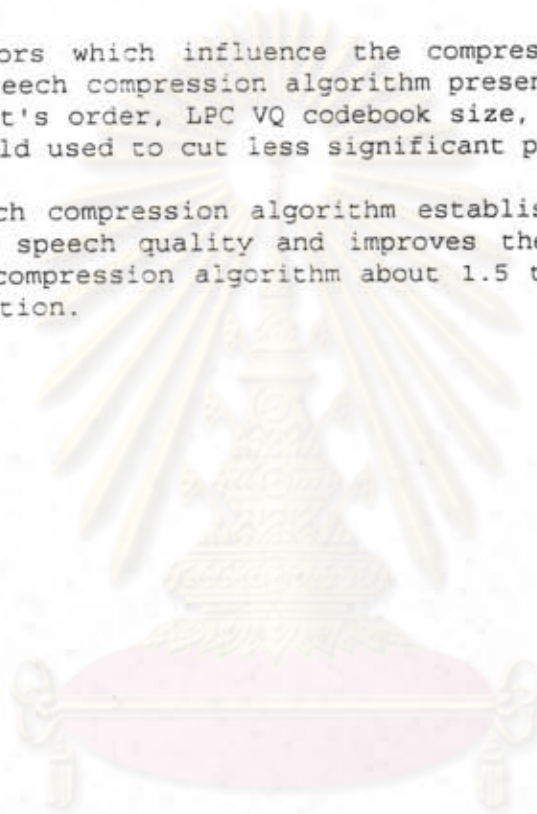
ลายมือชื่อนิสิต.....  
ลายมือชื่ออาจารย์ที่ปรึกษา.....  
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

## C515474: MAJOR ELECTRICAL ENGINEERING  
KEY WORD: THAI SPEECH COMPRESSION / WAVELET / LPC  
CHOTIWAT OUNNAPIRAK : SPEECH COMPRESSION USING WAVELET  
TRANSFORM AND LPC VECTOR QUANTIZATION. THESIS ADVISOR :  
ASSO. PROF. SOMCHAI JITAPUNKUL 48 pp. ISBN 974-632-415-2

In this thesis, we introduce a new algorithm to speech compression which allows us to increase the compression ratio of CELP while keeping good speech quality. By applying wavelet packet transform to residual signal obtained from CELP(Code-Excited Linear Predictive Coding), we cut less significant parts of transform signal using best level criteria.

The factors which influence the compression ratio and speech quality in the speech compression algorithm presented in this thesis are the mother wavelet's order, LPC VQ codebook size, excitation VQ codebook size, and threshold used to cut less significant parts of signal.

The speech compression algorithm established gives a 9,600 bps bit-rate at good speech quality and improves the compression ratio of the CELP speech compression algorithm about 1.5 times under appropriate compression condition.



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา.....วิศวกรรมไฟฟ้า

สาขาวิชา.....DSP

ปีการศึกษา.....2527

ลายมือชื่อนิสิต.....

ลายมือชื่ออาจารย์ที่ปรึกษา.....

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

### กิตติกรรมประกาศ

ขอขอบพระคุณ รศ. คร. สมชาย จิตะพันธ์กุล อาจารย์ที่ปรึกษาที่กรุณาให้แนวคิดในการทำวิทยานิพนธ์ ให้ความอนุเคราะห์เอกสารอ้างอิงจำนวนมาก และสละเวลาตรวจแก้วิทยานิพนธ์ฉบับนี้จนเสร็จลุล่วง

ขอขอบพระคุณ อ. สุวิทย์ นาคพีระบุท สำหรับเอกสารอ้างอิงจำนวนมาก

ขอขอบคุณ พี่วัชรพันธุ์ สำหรับเอกสารอ้างอิงจำนวนมาก

ขอขอบคุณ มนชัย สำหรับคำแนะนำในการทำวิทยานิพนธ์

ขอขอบคุณ ลำไย สำหรับน้ำใจ และความช่วยเหลือในการทำวิทยานิพนธ์

และสุดท้ายนี้ ขอขอบพระคุณ พ่อ แม่ ปู่ และทุกคนในครอบครัว ที่ให้โอกาส และความเข้าใจ จนทำวิทยานิพนธ์นี้สำเร็จ



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฅ
สารบัญภาพ.....	ฉุ
บัญชีคำย่อ.....	ฎ
บทที่ 1 บทนำ.....	1
ความสำคัญของวิทยานิพนธ์.....	1
จุดประสงค์ของวิทยานิพนธ์.....	2
ขอบเขตของวิทยานิพนธ์.....	2
วิธีการดำเนินงาน.....	2
ประโยชน์ที่คาดว่าจะได้รับจากวิทยานิพนธ์.....	3
บทที่ 2 เวฟเลตและการแปลงเวฟเลต.....	4
ความเป็นมาของเวฟเลต.....	4
การแปลงเวฟเลต.....	4
การแปลงเวฟเลตแพกเกต.....	9
บทที่ 3 การบีบย่อเสียงพูดโดยวิธี CELP.....	13
การบีบย่อเสียงพูด.....	13
หลักการบีบย่อเสียงพูดโดยวิธี CELP.....	16
บทที่ 4 การบีบย่อสัญญาณเสียงพูดด้วยเวฟเลต.....	19
การบีบย่อสัญญาณเสียงพูดโดยใช้เวฟเลต.....	19
บทที่ 5 การทดลองบีบย่อสัญญาณเสียงพูดด้วยเวฟเลต.....	28
อุปกรณ์และวิธีการเก็บข้อมูล.....	28
การออกแบบสร้าง codebook.....	29

การวัดคุณภาพของสัญญาณเสียงพูด.....	31
การทดลองบีบอัดสัญญาณเสียงพูดภายใต้เงื่อนไขต่างๆ.....	31
การทดลองบีบอัดสัญญาณเสียงพูดทั่วไป.....	35
เปรียบเทียบอัตราการบีบอัดเมื่อใช้การแปลงเวฟเลตแพคเกตและไม่ใช้.....	36
บทที่ 6 บทสรุปและข้อเสนอแนะ.....	38
บทสรุป.....	38
ปัญหาที่เกิดขึ้น.....	39
ข้อเสนอแนะ.....	39
รายการอ้างอิง.....	41
ภาคผนวก parameter ของ Daubechies wavelet.....	46
ประวัติผู้เขียน.....	47



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



## สารบัญภาพ

	หน้า
รูปที่ 2-1 แสดงผลตอบเชิงความถี่ของ STFT และ WT.....	7
รูปที่ 2-2 แผนผังการทำ Discrete Wavelet Transform ถึง level 3.....	9
รูปที่ 2-3 ตัวอย่างการเลือก wavelet packets โดยเงื่อนไข best level.....	11
รูปที่ 2-4 ตัวอย่างการเลือก wavelet packets โดยเงื่อนไข best basis.....	11
รูปที่ 3-1 แสดง block diagram ของตัวเข้ารหัสแบบ CELP.....	19
รูปที่ 3-2 แสดง block diagram ของตัวเข้ารหัสแบบ LD-CELP.....	19
รูปที่ 4-1 แสดง block diagram ของตัวเข้ารหัส (encoder).....	19
รูปที่ 4-2 แสดง block diagram ของตัวถอดรหัส (decoder).....	20
รูปที่ 4-3 กราฟแสดงความสัมพันธ์ระหว่าง LPC Order และ Residual Power.....	21
รูปที่ 4-4 แสดงสัญญาณต่างๆในการเข้ารหัสและถอดรหัสใน 1 frame (20ms).....	26


  
 ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย

## สารบัญตาราง

	หน้า
ตารางที่ 4-1 แสดงความสัมพันธ์ระหว่าง LPC Order และ Residual Power.....	21
ตารางที่ 4-2 แสดงการทำ scalar quantization ของ pitch gain (g).....	23
ตารางที่ 4-3 แสดงการทำ scalar quantization ของ gain (G).....	24
ตารางที่ 5-1 แสดงระดับการให้คะแนนของ MOS (Mean Opinion Score).....	31
ตารางที่ 5-2 แสดงผลของการบีบอัดสัญญาณเสียงพูดโดยใช้ขนาดของ LPC codebook ต่าง ๆ กัน.....	32
ตารางที่ 5-3 แสดงผลการบีบอัดสัญญาณเสียงโดยใช้ขนาดของ excitation codebook ต่าง ๆ กัน.....	33
ตารางที่ 5-4 แสดงผลการบีบอัดสัญญาณเสียงโดยใช้ order ของ mother wavelet ต่าง ๆ กัน.....	34
ตารางที่ 5-5 แสดงผลการบีบอัดสัญญาณเสียงโดยใช้ค่า threshold ต่าง ๆ กัน.....	34
ตารางที่ 5-6 แสดงผลการบีบอัดสัญญาณเสียงใน training sequence และนอก training sequence.....	36

## บัญชีคำย่อ

ACELP	Adaptive Code-Excited Linear Prediction
CELP	Code-Excited Linear Prediction
DWT	Discrete Wavelet Transform
ITU	International Telecom Union
LD-CELP	Low Delay Code-Excited Linear Prediction
LPC	Linear Predictive Coding
LTP	Long-Term Prediction
VQ	Vector Quantization
WT	Wavelet Transform
WPKT	Wavelet Packet Transform

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย