

การเปรียบเทียบประสิทธิภาพการบีบอัดเสียงพูดโดยใช้เทคนิคเวฟเลต

Comparison Efficiency of Speech Compression using Wavelet Technique

วีระยุทธ คุณรัตนศิริ¹ และจักรี ศรีนนท์จักร²

¹สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ

7/1 ถนนพหลโยธิน ต.สวนใหญ่ อ.เมือง จ.นนทบุรี 11000 E-mail: weerayuth.k@rmutsb.ac.th

²ห้องปฏิบัติการและวิจัยทางด้านการประมวลผลสัญญาณ

ภาควิชาวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

39 ถนนรังสิต-นครนายก ต.คลองหก อ.ธัญบุรี จ.ปทุมธานี 12110 E-mail: jakkree.s@en.rmutt.ac.th

บทคัดย่อ

การบีบอัดสัญญาณเสียงมีจุดประสงค์ที่จะลดขนาดของสัญญาณเสียงและรักษาคุณภาพของสัญญาณเสียงให้มีลักษณะเหมือนกับเสียงต้นฉบับมากที่สุด ดังนั้นการเลือกใช้เทคนิคที่เหมาะสมในการบีบอัดสัญญาณเสียงจึงมีความสำคัญเป็นอย่างยิ่ง บทความนี้นำเสนอการเปรียบเทียบประสิทธิภาพของการบีบอัดสัญญาณเสียงโดยใช้เวฟเลต ทั้งนี้เวฟเลตมีอยู่ด้วยกันหลายชนิดและหลายระดับ (Level) ในการบีบอัดสัญญาณ ในบทความนี้ได้เลือกเวฟเลต 3 ชนิด ได้แก่ Daubechies, Symlet, Coiflets เพื่อนำมาเปรียบเทียบประสิทธิภาพของเวฟเลตทั้งสามชนิดๆ ละ 5 ระดับ ซึ่งงานวิจัยได้วิเคราะห์สัญญาณเสียงทั้งโดเมนเวลา และโดเมนทางความถี่ ซึ่งกระบวนการแยกองค์ประกอบของเวฟเลต นั้นมีการแยกออกเป็นองค์ประกอบของความถี่ต่ำและความถี่สูง โดยงานวิจัยนี้เลือกใช้เฉพาะความถี่ต่ำ หลังจากที่ได้สัญญาณที่ถูกบีบอัดแล้ว จะทำการแปลงสัญญาณที่บีบอัดคืนกลับเป็นสัญญาณเสียง เพื่อคำนวณหาค่าการสูญเสียของสัญญาณโดยวัดจากค่า SNR, PSNR และอัตราการบีบอัดสัญญาณเสียง จากการทดลองเห็นได้ว่า เวฟเลตชนิด Daubechies3 ในระดับที่ 2 มีค่า SNR และค่า PSNR ที่ดีที่สุดคือเท่ากับ 51.2578 dB ส่วนชนิด Daubechies1 และ Symlet1 ได้ค่าอัตราการบีบอัดเท่ากับ 3.9978

คำสำคัญ: สัญญาณเสียงพูด, การวิเคราะห์เวฟเลต, สเปกตรัมเชิงความถี่

Abstract

The propose of speech compression is to reduce the speech signal to low bit rate while maintain the quality of speech signal as same as the original. Therefore the technique that uses to compress the signal is so important. This article presents a comparison performance of wavelet technique for speech compression. The Daubechies, Symlet, and Coiflets wavelet technique is selected to perform in this experiment. These techniques have compressed each speech signal into five levels in condition of

time domain and frequency domain analysis. However, the wavelet technique is separated signal to be low and high frequencies section. This experiment selected only the low frequency to process. Speech signal is then decompressed and measured the quality of speech using the Signal to Noise Ratio (SNR), Peak Signal to Noise Ratio (PSNR). Also the compression rate is calculated. The results show that Daubechies 3 level 2 provides 51.2578 dB of SNR and PSNR. Also the Symlet1 and Daubechies1 give 3.9978 of compression ratio. Keywords: Speech Signal, Wavelet Analysis, Frequency Spectrum

1. บทนำ

ในการนำสัญญาณเสียงไปประยุกต์ใช้ในการรู้จำ เพื่อเพิ่มความรวดเร็วในการประมวลผล จำเป็นอย่างยิ่งที่จะต้องลดขนาดของสัญญาณเสียงที่เป็นสัญญาณอินพุตลง โดยเวฟเลตเป็นหนึ่งในเทคนิคที่ได้รับความนิยม แต่ปัจจุบันมีเวฟเลตหลายชนิด [1] โดยในเวฟเลตแต่ละชนิด ยังมีหลายระดับ ซึ่งทำให้ได้คุณภาพของสัญญาณเสียงที่แตกต่างกัน

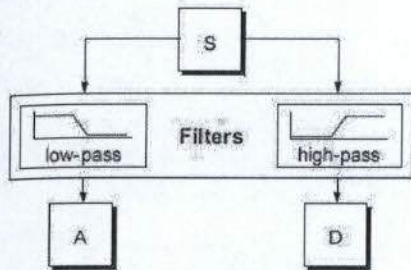
งานวิจัยนี้นำเสนอการเปรียบเทียบประสิทธิภาพของเวฟเลต 3 ชนิด ชนิดละ 5 ระดับ ตั้งแต่ระดับที่ 1-5 เพื่อให้สามารถเลือกเวฟเลตได้เหมาะสมกับการประมวลสัญญาณเสียง นอกจากนั้นยังแสดงให้เห็นคุณลักษณะของสัญญาณเสียงในโดเมนเวลา เพื่อแสดงถึงการสูญเสียของแอมพลิจูดหรือพลังงาน และโดเมนของความถี่เพื่อแสดงถึงการสูญเสียเชิงความถี่ของสัญญาณเสียง

2. หลักการ

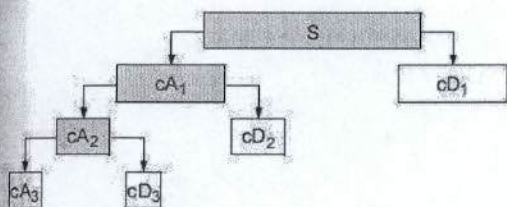
2.1 การบีบอัดสัญญาณเสียงโดยเวฟเลต

การแปลงเวฟเลตแบบไม่ต่อเนื่อง (Discrete Wavelet Transform : DWT) ประกอบด้วยขั้นตอนหลัก 2 ขั้นตอนคือ การแยกองค์ประกอบ (Wavelet Decomposition) ซึ่งจะทำให้การแยกออกเป็นค่าสัมประสิทธิ์ 2 ตัว ได้แก่ Approximate Coefficient ก็คือองค์ประกอบที่ได้จากความถี่ต่ำ และ Detail Coefficient คือองค์ประกอบที่ได้จาก

ความถี่สูง โดยการบีบอัดข้อมูลของเวฟเล็ตในแต่ละระดับ จะทำการลดจำนวนข้อมูลลงครึ่งหนึ่ง เนื่องจากการแยกออกเป็น 2 ส่วน คือ Approximate และ Detail อย่างละ 50% ดังนั้นในทุกๆระดับของการบีบอัดสัญญาณโดยเวฟเล็ตที่เพิ่มสูงขึ้น จำนวนข้อมูลจะลดลงครึ่งหนึ่งในทุกๆ ระดับ อีกขั้นตอนหนึ่งก็คือการคืนกลับสัญญาณเสียง (Wavelet Reconstruction) เป็นขั้นตอนที่นำสัญญาณที่ได้จากการบีบอัดแปลงกลับเป็นสัญญาณเสียง ดังรูปที่ 1 และรูปที่ 2



รูปที่ 1 การแยกองค์ประกอบโดยเวฟเล็ต



รูปที่ 2 สัมประสิทธิ์ที่ได้จากการแยกองค์ประกอบโดยเวฟเล็ต

ในปัจจุบันมีการพัฒนาเวฟเล็ตออกมาหลายชนิด โดยแบ่งแยกตามลักษณะงานต่างๆ ซึ่งในแต่ละชนิดของเวฟเล็ตจะมีคุณลักษณะที่แตกต่างกัน ซึ่งสามารถแยกออกได้ดังตารางที่ 1

ตารางที่ 1 คุณลักษณะของเวฟเล็ตชนิดต่าง ๆ [1]

	Compact Support	Symmetry	Orthogonal	Bi-orthogonal
Haar	Y	Y	Y	Y
Db	Y	N	Y	Y
Coif	Y	N	Y	Y
Sym	Y	N	Y	Y
Mayer	N	Y	Y	Y
Bior	Y	N	N	Y

2.2 การหาประสิทธิภาพของการบีบอัดโดยใช้เวฟเล็ต

ในการหาประสิทธิภาพของการบีบอัดสัญญาณเสียงด้วยวิธีเวฟเล็ต สามารถทดสอบหาประสิทธิภาพได้หลายวิธี ในงานด้านการประมวลผลสัญญาณเสียงนิยมใช้การหาค่าผิดพลาด SNR และ PSNR

ซึ่งเป็นการเปรียบเทียบข้อผิดพลาดระหว่างสัญญาณอินพุตเทียบกับสัญญาณเอาต์พุต ซึ่งจะทำให้ทราบถึงคุณภาพของสัญญาณเสียงที่ลดลงจากการบีบอัดสัญญาณเสียง นอกจากนี้ยังสามารถหาอัตราการบีบอัดของสัญญาณเสียงที่ได้จากกระบวนการเวฟเล็ต ดังสมการต่อไปนี้

1. Signal to Noise Ratio : SNR

$$SNR = 10 \log_{10} \left| \frac{\sigma_x^2}{\sigma_e^2} \right| \quad (1)$$

2. Peak Signal to Noise Ratio : PSNR

$$PSNR = 10 \log_{10} \frac{NX^2}{\|x-r\|^2} \quad (2)$$

โดย N คือ ความยาวของสัญญาณเสียงที่ถูกคืนกลับ

X คือ ค่าเฉลี่ยกำลังสองสูงสุด

$\|x-r\|^2$ คือ ความแตกต่างของค่าพลังงานระหว่าง

สัญญาณต้นฉบับและสัญญาณที่ถูกคืนกลับ

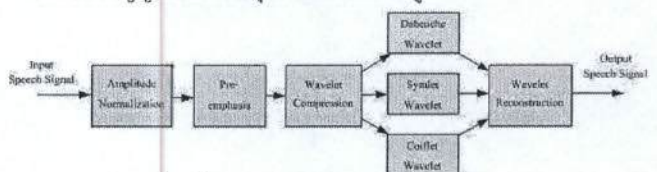
3. อัตราการบีบอัดสัญญาณ (Compression Ratio)

$$C = \frac{\text{Length}(x(n))}{\text{Length}(cWC)} \quad (3)$$

โดย cWC คือ ความยาวของข้อมูลสัญญาณเสียงที่ถูกบีบอัดแล้ว

3. ผลการทดลองการบีบอัดสัญญาณเสียงโดยวิธีเวฟเล็ต

ในการทดลองนี้ทำการบันทึกเสียงแบบโมโน ที่ความถี่ 8 กิโลเฮิรท์ซ โดยมีขนาดข้อมูลละ 8 บิต ค่าตัวอย่างที่ใช้ในการทดสอบคือ "ศูนย์" และใช้โปรแกรม MATLAB ในการประมวลผลและบีบอัดสัญญาณเสียง โดยก่อนเข้าสู่กระบวนการบีบอัดสัญญาณเสียงจะทำการปรับค่าสัญญาณเสียงอินพุตให้เหมาะสม ดังรูปที่ 1

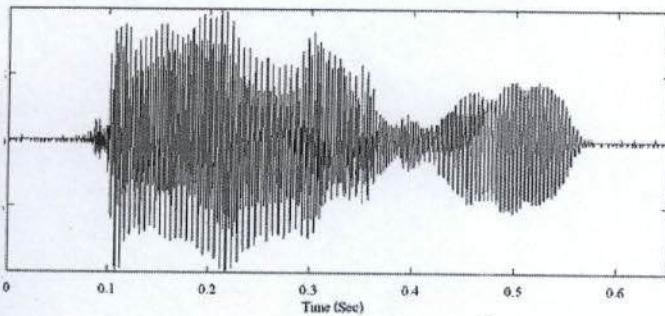


รูปที่ 3 ขั้นตอนการบีบอัดเสียงโดยเวฟเล็ต

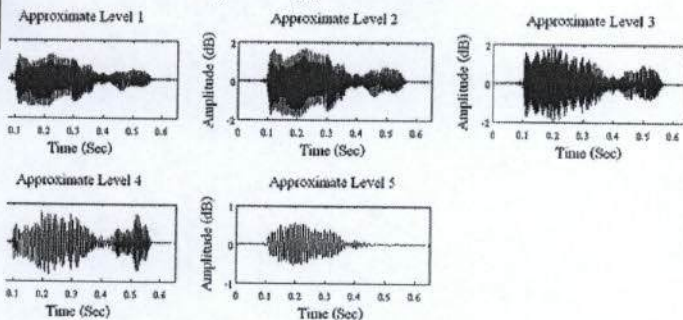
ในขั้นตอนของการประมวลผลสัญญาณเสียง สามารถทำการวิเคราะห์ได้ทั้งโดเมนเชิงเวลา (Time Domain) ซึ่งจะให้รายละเอียดของค่าพลังงานในช่วงเวลาต่าง ๆ และโดเมนเชิงความถี่ (Frequency Domain) ซึ่งแสดงองค์ประกอบของความถี่ โดยในการบีบอัดเสียงจะมีการสูญเสียสัญญาณเชิงความถี่เกิดขึ้น ในการทดลองเลือกใช้เวฟเล็ต 3 ชนิด คือ Daubechies, Symlet, Coiflets โดยในแต่ละชนิดทำการวิเคราะห์ตั้งแต่ระดับ 1 ถึงระดับที่ 5 เมื่อได้สัญญาณที่บีบอัดแล้ว ก็จะนำเข้าสู่กระบวนการหาประสิทธิภาพของสัญญาณ โดยใช้ค่า SNR, PSNR และอัตราการบีบอัดสัญญาณ

3.1 การเปรียบเทียบประสิทธิภาพเชิงเวลา

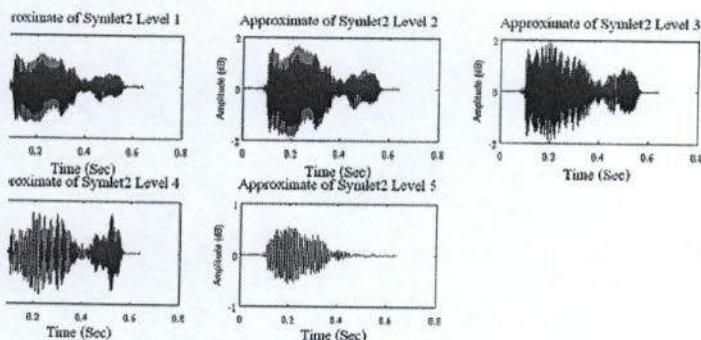
จากการทดลองพบว่าหลังจากทำการบีบอัดสัญญาณเสียง แล้วทำการแปลงกลับเป็นสัญญาณเสียงต้นฉบับ พบว่าค่าแอมพลิจูด หรือค่าพลังงานที่ได้จากสัญญาณเสียงที่แปลงกลับในโดเมนเชิงเวลา มีค่าลดต่ำลง อันเนื่องจากการแยกองค์ประกอบของความถี่ต่ำและความถี่สูงในกระบวนการของเวฟเลต ซึ่งทำให้จำนวนข้อมูลลดลง ส่งผลให้เวฟเลตตั้งแต่ระดับที่สี่เป็นต้นไปของเวฟเลตทั้งสามชนิด มีลักษณะของสัญญาณที่ผิดเพี้ยนไปจากสัญญาณต้นฉบับเป็นอย่างมาก ดังแสดงรูปที่ 4 และสัญญาณคืนกลับในโดเมนของเวลาของเวฟเลตทั้ง 3 ชนิด ดังรูปที่ 5 ถึงรูปที่ 7



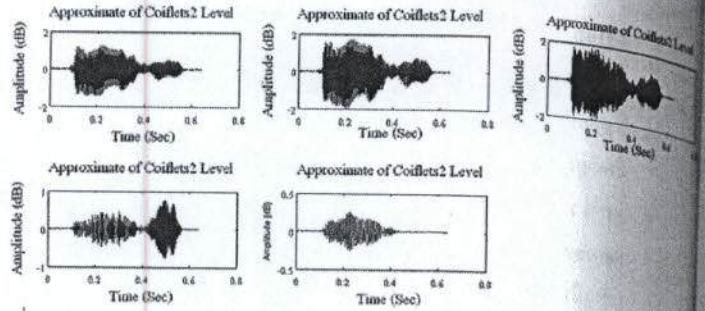
รูปที่ 4 สัญญาณเสียงคืนฉบับ



รูปที่ 5 สัญญาณเสียงที่ได้จากการคืนกลับสัญญาณของเวฟเลตชนิด Daubechies2 ระดับที่ 1 ถึง 5 ในโดเมนเชิงเวลา



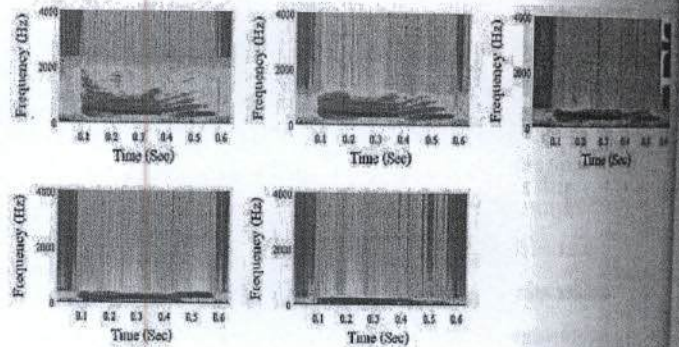
รูปที่ 6 สัญญาณเสียงที่ได้จากการคืนกลับสัญญาณของเวฟเลตชนิด Symlet2 ระดับที่ 1 ถึง 5 ในโดเมนเชิงเวลา



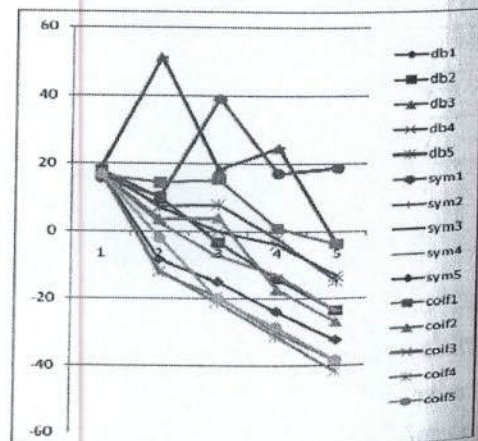
รูปที่ 7 สัญญาณเสียงที่ได้จากการคืนกลับสัญญาณของเวฟเลตชนิด Coiflet2 ระดับที่ 1 ถึง 5 ในโดเมนเชิงเวลา

3.2 การเปรียบเทียบประสิทธิภาพเชิงความถี่

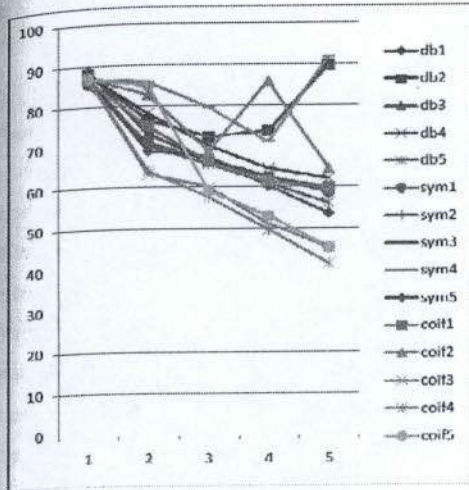
จากกระบวนการแปลงกลับเป็นสัญญาณเสียง ในการวิเคราะห์เชิงความถี่ พบว่าในเวฟเลตระดับที่สูงขึ้น องค์ประกอบของสัญญาณเชิงความถี่เริ่มลดต่ำลง โดยสังเกตที่เวฟเลตระดับที่ 4 ของทุกชนิด พบว่ามีการสูญเสียเชิงความถี่ เมื่อเปรียบเทียบกับระดับที่ 1-3 ส่งผลให้ไม่เหมาะสมที่จะนำไปวิเคราะห์ในกระบวนการถัดไปได้ ดังรูปที่ 8



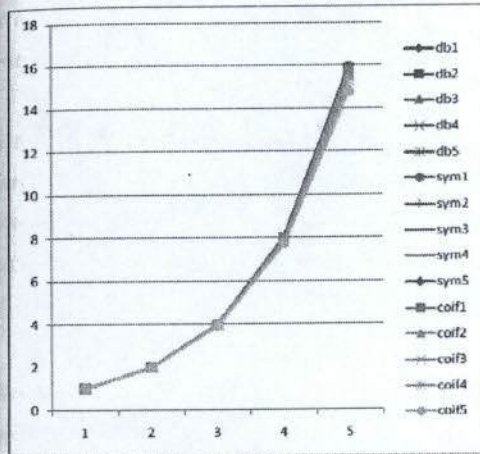
รูปที่ 8 สัญญาณเสียงที่ได้จากการคืนกลับสัญญาณของเวฟเลตชนิด Daubechies2 ระดับที่ 1 ถึง 5 ในโดเมนเชิงความถี่



รูปที่ 9 ค่า SNR ที่ได้จากการบีบอัดสัญญาณเสียง



รูปที่ 10 ค่า PSNR ที่ได้จากการบีบอัดสัญญาณเสียง



รูปที่ 11 อัตราการบีบอัดสัญญาณเสียงต้นฉบับเปรียบเทียบกับสัญญาณเสียงที่แปลงกลับ

4. สรุป

งานวิจัยนี้นำเสนอการเปรียบเทียบประสิทธิภาพของการบีบอัดเสียงโดยใช้ดีสครีตเวฟเลต 3 ชนิดคือ Daubechies, Symlet, Coiflets โดยแต่ละชนิด ๆ ละ 5 ระดับ พบว่าในการบีบอัดเสียงในตั้งแต่ระดับที่ 4 ขึ้นไปของเวฟเลตทุกชนิด มีการผิดเพี้ยนของสัญญาณเป็นอย่างมากเมื่อเปรียบเทียบกับสัญญาณเสียงอินพุต ทั้งในโดเมนเชิงเวลาและโดเมนเชิงความถี่ ซึ่งทำให้สัญญาณเสียงที่ได้ในระดับดังกล่าวไม่เหมาะสมต่อการนำมาใช้งาน ในงานวิจัยนี้เลือกทดสอบการบีบอัดสัญญาณเสียงในระดับที่ 1 ถึง 3 จากการหาประสิทธิภาพของการบีบอัดสัญญาณเสียง พบว่า เวฟเลตชนิด Daubechies3 ในระดับที่ 2 มีค่า SNR และ PSNR เท่ากับ 51.2578 dB ซึ่งเป็นค่าที่ดีที่สุดเมื่อเปรียบเทียบกับเวฟเลตชนิดอื่น ๆ ส่วนอัตราการบีบอัดสัญญาณเสียงที่ดีที่สุด ได้แก่ เวฟเลตชนิด Daubechies1 และ Symlet1

ในระดับที่ 3 โดยมีค่าเท่ากับ 3.9978 เท่าเมื่อเปรียบเทียบกับสัญญาณต้นฉบับ

5. เอกสารอ้างอิง

- [1] Yan, L., G. Liu, "Selection of The Best Wavelet Base for Speech Signal", Intelligent Multimedia, Video and Speech Processing, 2004.
- [2] Rufiner, H. L. and J. Goddard C, "A method of wavelet selection in phoneme recognition", Circuits and Systems, 1997.
- [3] Najih, A. M. M. A., A. R. bin Ramli, "Speech compression using discrete wavelet transform", Telecommunication Technology, 2003. NCTT 2003 Proceedings.
- [4] Hosny, N. M., S. H. El-Ramly, "Novel techniques for speech compression using wavelet transform", Microelectronics, 1999. ICM '99.
- [5] Junejo, N., N. Ahmed, "Speech and Image Compression Using Discrete Wavelet Transform". Advances in Wired and Wireless Communication, 2005.
- [6] Agbinya, J. I., "Discrete wavelet transform techniques in speech processing", TENCON '96.
- [8] J.Srinonchat, S.Danaher, J.I.H.Allen, "Double Clustering Algorithm applied to Speaker Dependent Information", SPECOM'2004.



นายวีระยุทธ คุณรัตนศิริ จบการศึกษาระดับปริญญาตรี วิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม พ.ศ. 2550 จากสถาบันเทคโนโลยีปทุมวัน ปัจจุบันกำลังศึกษาในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี งานวิจัยที่สนใจ คือ การประมวลผลสัญญาณเสียง



จักรี ศรีนนท์ถักร จบการศึกษาระดับปริญญาตรีวิศวกรรมไฟฟ้า แขนงวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม ในปี พ.ศ. 2538 จากสถาบันเทคโนโลยีราชมงคลธัญบุรี และจบการศึกษาระดับปริญญาโทและเอก จาก

Northumbria University, UK ในปี พ.ศ.2544 และพ.ศ.2548 ตามลำดับในสาขาวิศวกรรมไฟฟ้า สาขาเอกวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม ปัจจุบันดำรงตำแหน่งอาจารย์ผู้สอนที่ภาควิชาวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี งานวิจัยที่สนใจ คือ การประมวลผลสัญญาณ การบีบอัดสัญญาณเสียงและภาพ โครงสร้างจำลองระบบประสาทเทียม