< EN - 126 > (69)

ET4_11

การวิเคราะห์และเปรียบเทียบเทคนิคการสกัดคุณลักษณะของสัญญาณที่เป็นเสียงและเสียงรบกวน

Analysis and Comparison of Voiced and Unvoiced Classification Techniques

สุภาธิณี กรสิงห์[1] เกรียงศักดิ์ พัฒนบุรี[2] เฉลิมเกียรติ สุดาชา[3] และ จักรี ศรีนนท์ฉัตร[4]
สาขาวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
39 ม.1 ต.คลองหก อ.ธัญบุรี จ.ปทุมธานี 12110 โทรศัพท์: 02-549-3588 E-mail: top123354@hotmail.com

บทคัดย่อ

บทความนี้นำเสนอการวิเคราะห์และเปรียบเทียบเทคนิคการสกัดคุณลักษณะของสัญญาณที่เป็นเสียงและเสียงรบกวน ทั้งนี้การสกัดหรือแยกแยะสัญญาณที่เป็นเสียงและเสียงรบกวนนั้นมีประโยชน์และจำเป็นอย่างยิ่งสำหรับการประมวลผลสัญญาณเสียง ระบบการรู้จำเสียงพูด การบีบอัดสัญญาณเสียง และสังเคราะห์เสียง ในบทความนี้ได้ทำการทดลองกับเสียงของผู้พูดทั้งชายและหญิง รวมถึงภาษาที่ใช้นั้นมีทั้งภาษาอังกฤษและภาษาไทย เทคนิคที่นำใช้ในงานวิจัยนี้อาศัยคุณสมบัติทางการกระจายน้ำหนักของพลังงานเสียง (Weight distribution) และการกระจายของความถี่ (Frequency distribution) จากผลการทดลองพบว่าวิธีที่พบค่า error น้อยที่สุดคือ วิธีที่ 5 ส่วนวิธีที่มีค่า error มากที่สุดคือวิธีที่ 1

คำสำคัญ: การประมวลผลสัญญาณเสียง, การแยกแยะสัญญาณเสียงและเสียงรบกวน, การกระจายของความถี่ในสัญญาณเสียง

Abstract

This article describes an analysis and comparison of voiced and unvoiced classification techniques. Voiced and unvoiced extraction techniques are wildly useful and necessary for speech processing, speech recognition, speech compression and speech synthesis. This article uses male and female speech within Thai and English language as the input speech. These techniques are used in this research based on the weight distribution and frequency distribution. The results show that the best and worst performance is the first and fifth technique respectively.

Keywords: speech signal processing, voiced and unvoiced classification, frequency distribution in speech signal

1. คำนำ

การประมวลผลของสัญญาณเสียง [1] นั้นเป็นส่วนหนึ่งของการประยุกต์ใช้การประมวลผลสัญญาณทางด้านดิจิตอลเข้ากับสัญญาณเสียง โดยธรรมชาตินั้นสัญญาณเสียงจะเป็นสัญญาณในลักษณะที่เป็นสัญญาณต่อเนื่อง (Continues signal) แต่มีรูปแบบที่ไม่ซ้ำกัน (non-stationary signal) [2] ทั้งนี้สัญญาณถูกนำทำการวิจัยและประยุกต์ใช้ในหลายๆด้านเช่นการบีบอัดสัญญาณเสียง (speech compression) [3][4] การรู้จำสัญญาณเสียง (speech recognition) การแยกแยะผู้พูด (speaker identification) รวมถึงการสังเคราะห์เสียง (speech synthesis) เป็นต้น ทั้งนี้ในสัญญาณเสียงจะประกอบด้วยสัญญาณที่เป็นเสียง (voiced signal) และสัญญาณที่ไม่ใช่เสียง (unvoiced signal) หรือเสียงรบกวน ซึ่งในการแยกแยะคุณลักษณะของสัญญาณเสียงนั้นการดำเนินการวิจัยในหลายๆวิธี เช่น ในงานวิจัย [5] ได้นำเสนอการสกัดคุณลักษณะของสัญญาณที่เป็นเสียงและเสียงรบกวนรวมถึงการนำสัญญาณเหล่านั้นมาทำการสังเคราะห์ขึ้นใหม่โดยอาศัยหลักการของ Gaussian Mixer โดยนำไปประยุกต์ใช้กับการบีบอัดสัญญาณเสียงซึ่งจากผลการวิจัยพบว่าสามารถนำไปเพิ่มประสิทธิภาพของการบีบอัดสัญญาณเสียงแบบ MELP ได้ ในงานวิจัยของ [6] ได้นำเสนอการรู้จำอารมณ์ของคำพูดโดยอาศัยการแยกแยะสัญญาณเสียงและเสียงรบกวนแบบการแยกคุณลักษณะเฉพาะของคำพูดเพื่อนำไปใช้ประโยชน์ระบบการโต้ตอบของหุ่นยนต์จากผลการวิจัยพบว่าการแยกแยะนั้นมีผลต่อความแม่นยำในการบอกถึงความรู้สึกจากคำพูดนั้นๆ ในงานวิจัย [7] ได้นำเสนอเทคนิคที่ชื่อว่า Instantaneous Frequency Amplitude Spectrum (IFAS) ในการแยกแยะสัญญาณเสียงและเสียงรบกวน ซึ่งจากผลการทดลองพบว่าเทคนิคนี้สามารถแยกแยะสัญญาณเสียงของผู้ชายได้ดีกว่าเสียงของผู้หญิงประมาณ 5% และในงานวิจัย [8] ได้ประยุกต์ใช้เทคนิคเวฟเล็ตในการแยกแยะสัญญาณเสียงและเสียงรบกวน ผลการวิจัยพบว่าเทคนิคที่ใช้ในการวิจัยนี้สามารถแยกแยะสัญญาณเสียงและเสียงรบกวนของผู้ชายได้ดีกว่าเสียงของผู้หญิงอันเนื่องมาจากความถี่เสียงของผู้ชายที่มีความถี่ต่ำกว่าเสียงผู้หญิง

ในบทความนี้นำเสนอการวิเคราะห์และเปรียบเทียบการสกัดคุณลักษณะของสัญญาณที่เป็นเสียงและสัญญาณรบกวนโดยวิธีต่างๆซึ่งจะอธิบายถึงทฤษฎีไว้ในส่วนที่ 2 ของบทความ การทดลองจะถูกนำเสนอในส่วนที่ 3 ผลการทดลองและสรุปผลการทดลองจะถูกกล่าวถึงในลำดับต่อๆไป

2. การแยกแยะสัญญาณเสียงและสัญญาณรบกวน

เสียงพูด (Speech) คือเสียงที่มนุษย์เปล่งออกมาเพื่อใช้ติดต่อสื่อสารกันซึ่งเสียงนั้นจะประกอบไปด้วย สัญญาณเสียงและสัญญาณรบกวนหรือสัญญาณที่ไม่ใช่เสียง ซึ่งเทคนิคที่ใช้ในการแยกแยะ

คุณลักษณะของสัญญาณเสียงในบทความนี้มีด้วยกัน 5 วิธีโดยขึ้นอยู่กับค่าพลังงานช่วงสั้น (Short Time Energy) เพื่อแยกสัญญาณเสียงออกจากสัญญาณที่ไม่ใช่เสียง

2.1 การตัดสินใจวิธีที่ 1

วิธีนี้เป็นการตัดสินใจเพื่อหา Voice และ Unvoiced โดยสมมุติค่าการตัดสินใจ (Threshold) ของค่าพลังงานที่สูงสุด (Max Energy) ในแต่ละเฟรมโดยเทียบค่าเฉลี่ยของแต่ละเฟรม ดังสมการที่ (1) และ (2)

$$E_n = \frac{1}{w}\sum_{n=1}^{w} x(n) \quad (1)$$

$$Th = max_m \times ratio \quad (2)$$

เมื่อ E_n คือ ค่าเฉลี่ยของพลังงานทั้งหมดใน 1 เฟรม

x คือ การสุ่ม (Sampling)

n คือ จำนวนครั้งของการสุ่ม

m คือ เฟรม (Frame)

w คือ จำนวนการสุ่มทั้งหมดใน 1 เฟรม

max_m คือ ค่าพลังงานสูงสุดของสัญญาณเสียงในแต่ละเฟรม

Th คือ ค่าการตัดสินใจ (Threshold)

2.2 การตัดสินใจวิธีที่ 2

วิธีนี้จะคล้ายกับวิธีที่ 1ซึ่งแตกต่างกันที่ค่า max ของวิธีนี้คือค่าพลังงานสูงสุดของสัญญาณเสียงทั้งหมดส่วนวิธีที่ 1 จะเป็นค่าพลังงานสุดสุดในแต่ละเฟรม

2.3 การตัดสินใจวิธีที่ 3

วิธีนี้จะเป็นการตัดสินใจระหว่าง เสียง (Voice) และ ไม่ใช่เสียง (Unvoiced) หรือสัญญาณรบกวน (Noise) โดยแยกแยะด้วยค่าการตัดสินใจ (Threshold) จากค่าเฉลี่ยของผลรวมของทุกเฟรมเทียบกับค่าเฉลี่ยของแต่ละเฟรม ดังสมการที่ (1), (3) และ (4)

$$E_n = \frac{1}{w}\sum_{n=1}^{w} x(n)$$

$$E_a = \frac{1}{m}\sum_{i=1}^{n} E_n(i) \quad (3)$$

$$Th = E_a \times ratio \quad (4)$$

เมื่อ E_a คือค่าเฉลี่ยของผลรวมของทุกเฟรม

2.4 การตัดสินใจวิธีที่ 4

วิธีนี้จะเป็นการตัดสินใจระหว่าง เสียง (Voice) และ ไม่ใช่เสียง (Unvoiced) หรือสัญญาณรบกวน (Noise) โดยแยกแยะด้วยค่าการตัดสินใจ (Threshold) จากค่าเฉลี่ยของจากผลรวมของทุกเฟรมที่สูงที่สุดเทียบกับค่าเฉลี่ยของแต่ละเฟรม ดังสมการที่ (1), (3) และ (5)

$$E_n = \frac{1}{w}\sum_{n=1}^{w} x(n) \text{ และ } E_a = \frac{1}{m}\sum_{i=1}^{n} E_n(i)$$

$$Th = \max(E_a) \times ratio \quad (5)$$

โดยแยกระหว่าง Voice กับ Unvoicedได้ดังสมการที่

$$\text{Result} = \begin{cases} E_n > th : Vf \\ E_n < th : Uf \end{cases}$$

โดยที่ Vf คือ เสียง และ Uf คือ ไม่ใช่เสียง

2.5 การตัดสินใจวิธีที่ 5

วิธีเป็นการ หา Voice และ Unvoiced โดยสมมุติค่าการตัดสินใจ (Threshold) ของค่าพลังงานสูงสุดในสัญญาณเสียง โดยเปรียบเทียบทุก Sampling จะได้ค่าที่เป็น voiced และ Unvoiced จากนั้นนำค่าที่ได้ไปเปรียบเทียบกับ rate% ของ sampling ทั้งหมด ดังสมการที่ (2), (6) (7) และ (8)

$$Th = max \times ratio$$

$$S = \begin{cases} 1, & Th \geq x(n); voiced \\ 0, & Th < x(n); unvoiced \end{cases} \quad (6)$$

$$E_S = \frac{1}{w}\sum_{n=1}^{w} S(n) \quad (7)$$

$$Th_s = \begin{cases} E_s \geq ratio_s & ; Vf \\ E_s < ratio_s & ; Uf \end{cases} \quad (8)$$

เมื่อ E_s คือ ค่าเฉลี่ยของพลังงานใหม่ทั้งหมดใน 1 เฟรม

S คือ ค่าใหม่ของการสุ่ม (Sampling)

$ratio_s$ คือ ค่าเปอร์เซ็นต์ที่ใช้เปรียบแต่ละเฟรม

Th_s คือ ค่าตัดสินใจโดยเทียบกับค่า $ratio_s$

3.ขั้นตอนการทดลอง

สัญญาณอินพุตที่ใช้ในการวิจัยนี้เป็นสัญญาณเสียง8 บิต ที่อัตราสุ่ม 8 KHz จากนั้นทำการปรับแต่งสัญญาณเสียงเบื้องต้นให้เหมาะสมในการวิเคราะห์ โดยทำการ Normalization ซึ่งขั้นตอนนี้ทำการปรับระดับของสัญญาณให้อยู่ในช่วง -1 ถึง 1 เพื่อปรับให้เสียงอยู่ในระดับเดียวกันก่อนนำไปประมวลผล ดังสมการที่ (9) และ (10)

$$c = 1 + (1 - max(abs(x(i))) \quad (9)$$

$$x(i) = x(i) * c \quad (10)$$

โดยที่ $x(i)$ คือค่าสัญญาณเสียงอินพุต

c คืออัตราการขยายสัญญาณ

จากนั้นทำการแบ่งสัญญาณเสียงออกเป็นเฟรมย่อยๆก่อนนำเสียงพูดมาวิเคราะห์ที่ละส่วนจะทำให้วิเคราะห์ง่ายขึ้นและลดความซับซ้อนในการประมวลผล โดยขนาดของหน้าต่าง (Window) มีค่าอยู่ในช่วงที่เหมาะสมที่ใช้ในการทดลองเลือกที่ 200 จากนั้นสัญญาณเสียงจะถูกนำไปวิเคราะห์ในกระบวนการตัดสินใจในหัวข้อการแยกแยะสัญญาณเสียงและสัญญาณรบกวน

4. ผลการทดลอง

4.1 วิเคราะห์ผลการทดลองการตัดสินใจวิธีที่ 1

ค่าการตัดสินใจที่ 10% ของค่าพลังงานสูงสุดของสัญญาณเสียงในแต่ละเฟรมจะได้ voice 58 เฟรม และ unvoiced 167 เฟรม จะเห็นว่าผลลัพธ์ที่ได้แตกต่างจากต้นฉบับมาก ดังนั้นผลลัพธ์ที่

ได้จะไม่ค่อยมีความต่อเนื่องฟังแล้วจะเป็นเสียงสะดุด ส่วน 3% , 2% และ 1% ของค่าพลังงานสูงสุดของสัญญาณเสียงในแต่ละเฟรม จะได้ voice 90, 103 และ 124 เฟรม ส่วน unvoiced 135, 122 และ 101 เฟรม ตามลำดับ ซึ่งผลลัพธ์ที่ได้แตกต่างจากต้นฉบับเพียงเล็กน้อย ดังนั้นเสียงใหม่ที่ได้จะมีความต่อเนื่องมากขึ้นแต่ฟังแล้วมีความคล้ายเสียงต้นฉบับ โดยเฉพาะที่ 1% จะมีรูปสัญญาณที่คล้ายกับต้นฉบับมากที่สุดจากการสุ่มค่าในการทดลองครั้งนี้ ดังตารางที่ 1

ตารางที่ 1 การแยกแยะ Voice และ Unvoiced ด้วยวิธีที่ 1

Threshold (%)	*Voice Frames*	*Unvoiced Frames*
10	58	167
5	78	147
4	81	144
3	90	135
2	103	122
1	124	101

4.2 วิเคราะห์ผลการทดลองการตัดสินใจวิธีที่ 2

ค่าการตัดสินใจที่ 10% ของค่าพลังงานสูงสุดของสัญญาณเสียงในแต่ละเฟรม จะได้ voice 36 เฟรม และ unvoiced 189 เฟรม จะเห็นว่าผลลัพธ์ที่ได้แตกต่างจากเสียงต้นฉบับมาก ดังนั้นเสียงใหม่ที่ได้จะไม่ค่อยมีความต่อเนื่องฟังแล้วจะเป็นเสียงสะดุด ส่วนที่ 3% ของค่าพลังงานสูงสุดของสัญญาณเสียงในแต่ละเฟรมจะได้ voice 60, 64 เฟรม และ unvoiced 165, 161 เฟรม ตามลำดับ จากเสียงที่ได้ใหม่จะเห็นว่าผลลัพธ์ที่ได้แตกต่างจากต้นฉบับแต่ดีกว่าที่ Threshold ที่ 10% ดังนั้นจะสัญญาณเสียงที่ได้จะมีความต่อเนื่องมากขึ้น ดังตารางที่ 2

ตารางที่ 2 การแยกแยะ Voice และ Unvoiced ด้วยวิธีที่ 2

Threshold (%)	*Voice Frames*	*Unvoiced Frames*
10	36	189
5	60	165
4	64	161
3	71	154
2	85	140
1	110	115

4.3 วิเคราะห์ผลการทดลองการตัดสินใจวิธีที่ 3

ค่าการตัดสินใจ ที่ 100% ของค่าเฉลี่ยของผลรวมของทุกเฟรมเทียบกับค่าเฉลี่ยของแต่ละเฟรมจะได้ voice 64 เฟรม และ unvoiced 161 เฟรม จะเห็นว่าสัญญาณเสียงใหม่ที่ได้แตกต่างจากต้นฉบับมากไม่ค่อยมีความต่อเนื่องฟังแล้วจะเป็นเสียงสะดุด ส่วนที่ 50% จะได้ voice 86 เฟรม และ unvoiced 139 เฟรม ดีกว่าค่าการตัดสินใจที่ 100% และดีที่สุดในการใช้การแยกแยะสัญญาณเสียงและสัญญาณรบกวน และที่ค่าการตัดสินใจ ที่ 15% จะได้ voice 123 เฟรม และ unvoiced 102 เฟรม แต่ไม่สามารถใช้แยกแยะสัญญาณเสียงและสัญญาณรบกวนได้ ดังตารางที่ 3

ตารางที่ 3 การแยกแยะ Voice และ Unvoiced ด้วยวิธีที่ 3

Threshold (%)	*Voice Frames*	*Unvoiced Frames*
100	64	161
50	86	139
30	106	119
25	110	115
20	115	110
15	123	102

4.4 วิเคราะห์ผลการทดลองการตัดสินใจวิธีที่ 4

ค่าการตัดสินใจที่ 100% ค่าเฉลี่ยของจากผลรวมของทุกเฟรมที่สูงที่สุดเทียบกับค่าเฉลี่ยของแต่ละเฟรมจะได้ voice 1 เฟรม และ unvoiced 224 เฟรม จะเห็นว่าสัญญาณเสียงใหม่ที่ได้แตกต่างจากต้นฉบับมาก ส่วนที่ 8% จะได้ voice 86 เฟรม และ unvoiced 139 เฟรม ดีกว่าที่ค่าการตัดสินใจที่ 100% และดีที่สุดในการใช้การแยกแยะสัญญาณเสียงและสัญญาณรบกวน และค่าการตัดสินใจที่ 2% จะได้ voice 125 เฟรม และ unvoiced 100 เฟรม แต่ไม่สามารถใช้แยกแยะสัญญาณเสียงและสัญญาณรบกวนได้ ดังตารางที่ 4

ตารางที่ 4 การแยกแยะ Voice และ Unvoiced ด้วยวิธีที่ 4

Threshold (%)	*Voice Frames*	*Unvoiced Frames*
100	1	224
10	83	142
8	85	140
5	106	119
3	118	107
2	125	100

4.5 วิเคราะห์ผลการทดลองการตัดสินใจวิธีที่ 5

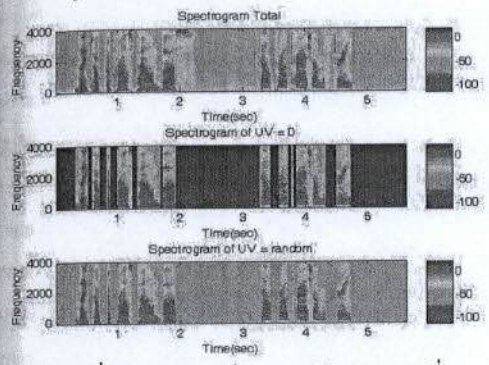
ค่าการตัดสินใจ (Threshold) ที่ 6% ลดลงมาจนถึงค่าการตัดสินใจที่ 1% และค่าใหม่ของการสุ่ม (Sampling) ที่ 70 ลดลงมาจนถึงค่าใหม่ของการสุ่มที่ 50 ได้ดังนี้ จากการวิเคราะห์โดยศึกษาผลลัพธ์จะสังเกตว่าถ้าค่าการตัดสินใจมาก เสียงที่ได้จะสะดุดมากจนถึงฟังแล้วไม่สามารถจับใจความในคำหรือประโยคนั้นได้ ในการทดลองนี้ ค่าที่สามารถยอมรับได้ในคุณภาพของเสียงที่ได้ใกล้เคียงต้นฉบับจะอยู่ที่ค่าการตัดสินใจ ที่ 1% และค่าใหม่ของการสุ่ม ที่ 70, 60, 50 นอกจากนี้แล้วคุณภาพของเสียงที่ได้ค่อนข้างต่ำ

ตารางที่ 5 การแยกแยะ Voice และ Unvoiced ด้วยวิธีที่ 5

Threshold (%)	*Sampling (%)*					
	90		*70*		*50*	
	Voice Frames	*Unvoiced Frames*	*Voice Frames*	*Unvoiced Frames*	*Voice Frames*	*Unvoiced Frames*
6	ไม่มี	225	22	203	48	177
5	ไม่มี	226	31	194	53	172
4	3	222	43	182	60	165
3	14	211	52	173	63	162
2	27	198	61	164	79	146
1	50	175	81	144	105	120

4.6 วิเคราะห์ผลการทดลองในลักษณะของความถี่

เมื่อผ่านกระบวนการตัดสินใจในการแยกแยะสัญญาณเสียงและสัญญาณรบกวน นำผลการทดลองมาวิเคราะห์การสูญเสียในเชิงความถี่ ทำการทดลองแทนค่าด้วยสัญญาณที่มีค่าเท่ากับศูนย์และสัญญาณสุ่มเข้าไปแทนที่สัญญาณรบกวน จากการทดลองพบว่าเมื่อแทนค่าสัญญาณรบกวนด้วยค่าศูนย์ จะเกิดความสูญเสียในเชิงความถี่และมีลักษณะผิดเพี้ยนจากสัญญาณต้นฉบับเกิดขึ้นอย่างมาก ส่วนการแทนค่าด้วยสัญญาณสุ่มเข้าแทนค่าสัญญาณสุ่มจะเกิดความสูญเสียในเชิงความถี่และมีลักษณะผิดเพี้ยนจากสัญญาณต้นฉบับแต่ดีกว่าการแทนค่าสัญญาณรบกวนด้วยค่าศูนย์ ดังแสดงในรูปที่ 1



รูปที่ 1 การวิเคราะห์สัญญาณเสียงทางความถี่

5. สรุป

บทความนี้เป็นการเสนอแนวทางในการลดสัญญาณที่ไม่เป็นเสียงหรือสัญญาณรบกวน ก่อนที่จะนำเข้าสู่กระบวนการอื่นๆต่อไป จากการทดลองทั้งหมด 5 วิธี จะได้เฟรมที่เป็น voice และ unvoiced ที่ใกล้เคียงกันยกเว้นวิธีที่ 1 เนื่องจากเป็นการคำนวณที่ใช้ค่าการตัดสินใจเทียบกับค่าพลังงานสูงสุดของแต่ละเฟรมจึงไม่สามารถแยกแยะได้ว่าเฟรมใดคือ voice เฟรมใดคือ unvoiced สามารถลดสัญญาณที่ไม่เป็นเสียงหรือสัญญาณรบกวนได้ดีขึ้นกว่าวิธีอื่นได้มากที่สุด

เอกสารอ้างอิง

[1] J.R. Deller, J.G.Proakis and J.H.Hansen, "Discrete-Time Processing of Speech Signals", 2000

[2] จักรกฤษ อ่อนชื่นจิตร, "การวิเคราะห์แนวทางเดินเสียงพูดในรูปแบบของสัมประสิทธิ์คู่เส้นสเปกตรัมร่วมกับ Double Clustering", วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์, มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, 2551.

[3] อรอนงค์ วิริยานุรักษ์นคร "การพัฒนาเทคนิคเวคเตอร์ควอนไทซ์เซชั่นสำหรับการบีบอัดสัญญาณเสียง", วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์, มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, 2553.

[4] สุทธี ทับทองดี, "การพัฒนาเทคนิคสำหรับการบีบอัดสัญญาณเสียงพูดภาษาไทยในมาตรฐาน LPC-10", วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์, มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, 2553.

[5] W. Xuan, D. Xiaoyan, C. Huijuan and T. Kun, "Voiced/ Unvoiced classification recovery in the speech decoder based on GMM", 9th International Conference on Signal Processing, 2008, pp. 546 – 548

[6] E.H.Kim, K.H.Hyun, S.H.Kim, and Y.K.Kwak, "Speech emotion recognition separately from voiced and unvoiced sound for emotional interaction robot", International Conference on Control Automation and Systems, 2008, pp. 2014 - 2019

[7] D.Arifianto, "Dual Parameters for Voiced-Unvoiced Speech Signal Determination", International Conference on Acoustics Speech and Signal Processing, 2007, pp. IV-749 - IV-752

[8] R.s.Cai, Y.T.Zhu, Y.M.Guo, "Wavelet-Based Multi-Feature Voiced/Unvoiced Speech Classification Algorithm", Conference on Wireless Mobile and Sensor Networks, 2007, pp. 897 – 900

ประวัติผู้เขียนบทความ

สุภาธิณี กรสิงห์ : สำเร็จการศึกษาระดับปริญญาตรี ด้านวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม ในปี พ.ศ. 2552 จากมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ปัจจุบันกำลังศึกษาระดับปริญญาโทด้านวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม

เกรียงศักดิ์ พัฒนบุรี : สำเร็จการศึกษาระดับปริญญาตรีด้านวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม ในปี พ.ศ. 2553 จากมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ปัจจุบันกำลังศึกษาระดับปริญญาโทด้านวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม

เฉลิมเกียรติ สุตาชา : สำเร็จการศึกษาระดับปริญญาตรีด้านวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม ในปี พ.ศ. 2553 จากมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ปัจจุบันกำลังศึกษาระดับปริญญาโทด้านวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม

จักรี ศรีนนท์ฉัตร สำเร็จการศึกษระดับปริญญาเอก จาก Northumbria Universisty, UK. ในปี พ.ศ. 2548 ในสาขาวิศวกรรมไฟฟ้า สาขาเอกวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม ปัจจุบันดำรงตำแหน่งอาจารย์ผู้สอนที่ภาควิชาอิเล็กทรอนิกส์และโทรคมนาคมมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี