การพัฒนาเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณรบกวนแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ

IMPROVEMENT VOICED AND UNVOICED CLASSIFICATION TECHNIQUE BASED ON REAL TIME PROCESSING USING FPGA BOARD

เฉลิมเกียรติ สุตาชา

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์
มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ปีการศึกษา 2555

ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

การพัฒนาเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณรบกวนแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ

เฉลิมเกียรติ สุตาชา

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์
มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ปีการศึกษา 2555

ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

หัวข้อวิทยานิพนธ์	การพัฒนาเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณรบกวนแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ
ชื่อ - นามสกุล	นายเฉลิมเกียรติ สุตาชา
สาขาวิชา	วิศวกรรมไฟฟ้า
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์จักรี ศรีนนท์ฉัตร, Ph.D.
ปีการศึกษา	2555

บทคัดย่อ

ปัจจุบันงานวิจัยในการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณรบกวนหรือสัญญาณเสียงที่ไม่ใช่เสียงพูดแบบฐานเวลาจริงมีความต้องการการประมวลผลที่ความเร็วสูง ดังนั้นวิทยานิพนธ์นี้ได้นำเสนอการพัฒนาเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ (Field Programmable Gate Array: FPGA) และยังได้พัฒนาเทคนิคที่สามารถทำงานทั้งบนซอฟต์แวร์และฮาร์ดแวร์

บอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro ถูกใช้ในการทดลองนี้สำหรับการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดแบบฐานเวลาจริงด้วยเทคนิคเส้นตัดสินใจ ซึ่งได้ศึกษาและทดลองหาค่าเส้นตัดสินใจที่เหมาะสมบนหลักการของ Signal to Noise Ratio (SNR) และ Normalized Root Mean Squared Error (NRMSE) จากนั้นได้ออกแบบระบบและสังเคราะห์เป็นวงจรลอจิกติดตั้งบนบอร์ดเอฟพีจีเอ ทั้งนี้การทดลองได้ใช้สัญญาณเสียงทั้งหมด 80 เสียง โดยแบ่งเป็นกลุ่ม ดังนี้ เสียงพูดของผู้หญิงและผู้ชายที่มีความยาว 5 วินาที และ 60 วินาที อย่างละ 20 เสียง

ผลทดลองพบว่าค่าเส้นตัดสินใจที่เหมาะสมคือร้อยละ 1 มีค่าอัตราบิบอัดข้อมูล 1.970 เท่า เทคนิคการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์มีค่า SNR เท่ากับ 20.024 dB และค่า NRMSE เท่ากับ 0.113 ส่วนเทคนิคการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่มมีค่า SNR เท่ากับ 19.398 dB ค่า NRMSE เท่ากับ 0.105 ดังนั้นเทคนิคการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์ให้ผลลัพธ์ที่ดีกว่า เมื่อออกแบบระบบและติดตั้งบนบอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro พบว่าเทคนิคดังกล่าวสามารถทำการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดแบบฐานเวลาจริงได้ โดยที่ใช้เวลาในการประมวลผล 25 มิลลิวินาที บนระบบที่มีอัตราสุ่มสัญญาณ 8 กิโลเฮิรตซ์ที่สัญญาณเสียงขนาด 18 บิต

คำสำคัญ : สัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด เอฟพีจีเอ ระบบแบบฐานเวลาจริง

Thesis Title	Improvement Voiced and Unvoiced Classification Technique Based on Real Time Processing Using FPGA Board
Name - Surname	Mr. Chalermkiat Sutacha
Program	Electrical Engineering
Thesis Advisor	Assistant Professor Jakkree Srinonchat, Ph.D.
Academic Year	2012

ABSTRACT

Recently, the research of voiced and noise or unvoiced classification based on real time system has a requirement of high speed processing. This thesis presents an improvement voiced and unvoiced classification technique based on real time processing using FPGA board and also develops technique which can be operate in software and hardware.

Xilinx Virtex-II Pro FPGA board is used in this experiment for voiced and unvoiced classification based on real time system with threshold value technique. The appropriate threshold value is studied and investigated based on Signal to Noise Ratio (SNR) and Normalized Root Mean Squared Error (NRMSE). At this point, the design and synthesis logic system is implemented on FPGA board. In the experiments, there are 80 speech signals which are used as input data. These signals can be categorized into groups that consist of male and female speech signal with the length of 5 and 60 seconds respectively.

The results show that the appropriate threshold value is 1%. The compression ratio is 1.970. The replacement of unvoiced signal with zero padding technique provides 20.024 dB and 0.113 for SNR and NRMSE respectively. The replacement of unvoiced signal with random signal provides 19.398 dB and 0.105 for SNR and NRMSE respectively. The replacement of unvoiced signal with zero padding technique is better than the replacement of unvoiced signal with random signal. In the implement of Xilinx Virtex-II Pro shows that those techniques can be classify voiced and unvoiced based on real time system which used 25 ms for processing process with sampling rate as 8 kHz. Each sampling uses 18 bits.

Keywords: voiced and unvoiced, FPGA, real time system

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลืออย่างดียิ่งของ ผู้ช่วยศาสตราจารย์ ดร. จักรี ศรีนนท์ฉัตร อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งได้ให้คำแนะนำข้อคิดเห็นและสนับสนุนในการทำวิจัยมาด้วยดีตลอด ผู้วิจัยจึงขอกราบขอบพระคุณมา ณ ที่นี้

กราบขอบพระคุณ คณะกรรมการสอบวิทยานิพนธ์ ดร. อำนวย เรืองวารี ดร. ณฐภัทร พันธ์คง และผู้ทรงคุณวุฒิ ผู้ช่วยศาสตราจารย์ ดร. คมกฤตย์ ชมสุวรรณ ที่ให้คำแนะนำ ความรู้และประสบการณ์ที่เป็นแนวทางในการทำวิทยานิพนธ์นี้ให้สำเร็จลุล่วงด้วยดี และเป็นประโยชน์ในทำการวิจัยครั้งต่อไป

ขอขอบคุณพี่ๆ เพื่อนๆ ที่ห้องปฏิบัติการและวิจัยทางด้านการประมวลผลสัญญาณ ที่ได้ช่วยเหลือเกี่ยวกับข้อมูล รวมถึงคำแนะนำต่างๆ ตลอดเวลาที่การทำวิจัยอย่างดียิ่ง

สุดท้ายนี้ผู้วิจัยขอกราบขอบพระคุณ บิดามารดาที่ให้การสนับสนุนแก่ผู้วิจัยเสมอมาจนสำเร็จการศึกษา

เฉลิมเกียรติ สุตาชา

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ค

บทคัดย่อภาษาอังกฤษ ..... ง

กิตติกรรมประกาศ ..... จ

สารบัญ ..... ฉ

สารบัญตาราง ..... ซ

สารบัญภาพ ..... ฌ

คำอธิบายสัญลักษณ์และคำย่อ ..... ฎ

บทที่

1 บทนำ 1

1.1 ความเป็นมาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์ของการศึกษา 2

1.3 ขอบเขตของการศึกษา 3

1.4 ข้อจำกัดของการศึกษา 3

1.5 ประโยชน์ที่คาดว่าจะได้รับ 3

2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 4

2.1 เอฟพีจีเอ 4

2.2 ภาษาวีเอชดีแอล 6

2.3 คุณลักษณะเสียงพูดของมนุษย์ 9

2.4 การประมวลผลสัญญาณเสียงพูด 10

2.5 อัตราการบีบอัดสัญญาณ 12

2.6 การวัดความค่าผิดพลาด 13

2.7 การวัดค่าประสิทธิภาพของสัญญาณ 13

2.8 ระบบแบบฐานเวลาจริง 14

2.9 งานวิจัยที่เกี่ยวข้อง 14

3 วิธีการดำเนินงานวิจัย 19

3.1 ขั้นตอนการทดลอง 19

สารบัญ (ต่อ)

บทที่ หน้า

3.2 การประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอ 20

3.3 เทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด 21

3.4 การคัดเลือกค่าเส้นการตัดสินใจ 22

3.5 การออกแบบและการทำงานบนเอฟพีจีเอ 24

4 ผลการทดลอง 37

4.1 การเปรียบเทียบประสิทธิภาพและค่าผิดพลาดของการบีบอัดสัญญาณเสียงพูด 37

4.2 ผลการออกแบบการทำงานบนชิปเอฟพีจีเอ 43

4.3 ผลการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด 47

5 สรุปผลการทดลอง 52

5.1 สรุปผลงานวิจัย 52

5.2 ข้อเสนอแนะ 55

รายการอ้างอิง 56

ภาคผนวก 59

ภาคผนวก ก อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ที่ค่าเส้นตัดสินใจที่ 0.5 - 7% 60

ภาคผนวก ข ตัวอย่างการบีบอัดสัญญาณเสียงพูด 72

ภาคผนวก ค ผลงานตีพิมพ์เผยแพร่ 85

ประวัติผู้เขียน 118

สารบัญตาราง

ตารางที่ หน้า

3.1 รายละเอียดของชิป LM4550 AC97 Audio CODEC ที่ต่ออยู่กับขาเอฟพีจี 26

3.2 รายละเอียดของ Push Button และ DIP Switch ที่ต่ออยู่กับขาเอฟพีจี 27

3.3 คำอธิบายหน้าที่ของรีจิสเตอร์ 29

3.4 การเชื่อมต่อของ AC97 Audio CODEC 29

4.1 การวัดค่าประสิทธิภาพของสัญญาณเสียงพูดที่ค่าเส้นตัดสินใจที่ 1% 43

5.1 การวัดค่าประสิทธิภาพของสัญญาณเสียงพูดที่ค่าเส้นตัดสินใจที่ 0.5 - 7% 53

5.2 การใช้ทรัพยากรภายในชิปเอฟพีจีเอ 54

สารบัญภาพ

ภาพที่ หน้า

2.1 ชิปเอฟพีจีเอ Xilinx Virtex-II Pro XC2VP30 5
2.2 ขั้นตอนการออกแบบจากบนลงล่าง 9
2.3 ลักษณะสัญญาณเสียงก้องหรือเสียงโฆษะ 10
2.4 ลักษณะสัญญาณเสียงไม่ก้องหรือเสียงอโฆษะ 10
2.5 การแบ่งสัญญาณเสียงออกเป็นเฟรมย่อย 11
3.1 ขั้นตอนการทำงานของงานวิจัย 19
3.2 ไดอะแกรมของระบบแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด ... 20
3.3 การแยกแยะสัญญาณเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยเส้นตัดสินใจ 21
3.4 ผลการแยกแยะสัญญาณเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยเส้นตัดสินใจ..... 22
3.5 ตัวอย่างสัญญาณเสียงพูด ที่เวลา 5 วินาที 23
3.6 ตัวอย่างสัญญาณเสียงพูด ที่เวลา 60 วินาที 23
3.7 บอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro XC2VP30 25
3.8 ไดอะแกรมของการเชื่อมต่อกับบอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro XC2VP30..... 26
3.9 บล็อกไดอะแกรมการจัดวางอุปกรณ์ชิป LM4550 AC97 Audio CODEC 27
3.10 บล็อกไดอะแกรมการจัดวางอุปกรณ์ชิปขยายกำลังเสียง 150 mW..... 28
3.11 การอินเตอร์เฟส LM4550 AC97 Audio CODEC กับเอฟพีจีเอ 28
3.12 การเก็บข้อมูลสัญญาณเสียงออกเป็นเฟรมย่อยด้วยหน่วยความจำ 30
3.13 ผังงานการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด..... 32
3.14 แผนภาพไดอะแกรมขั้นตอนการแยกแยะสัญญาณเสียงพูด..... 33
3.15 การสังเคราะห์วงจรลอจิกบนบนชิปเอฟพีจีเอจากซอฟแวร์ Xilinx ISE 34
3.16 การแยกแยะสัญญาณเสียงพูดในกระบวนการรับส่งข้อมูล 35
3.17 รูปแบบการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดจากการคำนวณบนเอฟพีจีเอ..... 36
3.18 การทำงานบนบอร์ดเอฟพีจีเอ Virtex-II Pro (XC2VP30)..... 36
4.1 ค่า SNR สัญญาณเสียงพูดผู้หญิงและผู้ชาย 5 วินาที ที่ค่าเส้นตัดสินใจที่ 0.5 - 7%..... 38
4.2 ค่า NRMSE สัญญาณเสียงพูดผู้หญิงและผู้ชาย 5 วินาที ที่ค่าเส้นตัดสินใจที่ 0.5 - 7%... 38
4.3 ค่า SNR สัญญาณเสียงพูดและผู้ชาย 60 วินาที ที่ค่าเส้นตัดสินใจที่ 0.5 - 7% 39
4.4 ค่า NRMSE สัญญาณเสียงพูดผู้หญิงและผู้ชาย 60 วินาที ที่ค่าเส้นตัดสินใจที่ 0.5 - 7%. 39

สารบัญภาพ (ต่อ)

ภาพที่ หน้า

4.5 ค่า SNR สัญญาณเสียงพูด 5 และ 60 วินาที ที่ค่าเส้นตัดสินใจที่ 0.5 - 7% 40

4.6 ค่า NRMSE สัญญาณเสียงพูด 5 และ 60 วินาที ที่ค่าเส้นตัดสินใจที่ 0.5 - 7% 40

4.7 อัตราบีบอัดข้อมูลสัญญาณเสียงพูด 5 วินาที ที่เส้นตัดสินใจที่ 0.5 - 7% 41

4.8 อัตราบีบอัดข้อมูลสัญญาณเสียงพูด 60 วินาที ที่เส้นตัดสินใจที่ 0.5 - 7% 42

4.9 อัตราบีบอัดข้อมูลสัญญาณเสียงพูดที่เส้นตัดสินใจที่ 0.5 - 7% 42

4.10 ผลจำลองการประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอจากซอฟแวร์ Xilinx ISE..... 44

4.11 ผลจำลองการส่งข้อมูลด้วยสัญญาณศูนย์บนชิปเอฟพีจีเอ 44

4.12 ผลจำลองการส่งข้อมูลด้วยสัญญาณสุ่มบนชิปเอฟพีจีเอ 44

4.13 ผลจำลองการแยกแยะสัญญาณเสียงพูดบนชิปเอฟพีจีเอ เมื่อ N เท่ากับ 199 หรือ 399 .. 45

4.14 ผลจำลองการแยกแยะสัญญาณเสียงพูดบนชิปเอฟพีจีเอเมื่อ pointer เท่ากับ 199 46

4.15 ผลการส่งข้อมูลบนชิปเอฟพีจีเอจากสัญญาณ AC97_SDATA_OUT 46

4.16 การบีบอัดสัญญาณเสียงพูด 5 วินาที ด้วยโปรแกรม MATLAB ในโดเมนเวลา..... 47

4.17 การบีบอัดสัญญาณเสียงพูด 5 วินาที บนบอร์ดเอฟพีจีเอในโดเมนเวลา..... 48

4.18 การบีบอัดสัญญาณเสียงพูด 60 วินาที ด้วยโปรแกรม MATLAB ในโดเมนเวลา..... 48

4.19 การบีบอัดสัญญาณเสียงพูด 60 วินาที บนบอร์ดเอฟพีจีเอในโดเมนเวลา..... 49

4.20 การบีบอัดสัญญาณเสียงพูด 5 วินาที ด้วยโปรแกรม MATLAB ในสเปกโตรแกรม 50

4.21 การบีบอัดสัญญาณเสียงพูด 5 วินาที บนบอร์ดเอฟพีจีเอในสเปกโตรแกรม 50

4.22 การบีบอัดสัญญาณเสียงพูด 60 วินาที ด้วยโปรแกรม MATLAB ในสเปกโตรแกรม ... 51

4.23 การบีบอัดสัญญาณเสียงพูด 60 วินาที บนบอร์ดเอฟพีจีเอในสเปกโตรแกรม 51

5.1 สัญญาณรบกวนภายในระบบ 54

คำอธิบายสัญลักษณ์และคำย่อ

EEPROM	Electrically Erasable Programmable Read Only Memory
FPGA	Field Programmable Gate Array
HDL	Hardware Description Language
NRMSE	Normalized Root Mean Squared Error
PCM	Pulse Code Modulation
SRAM	Static Read Only Memory
SNR	Signal to Noise Ratio
V/UV	Voiced and Unvoiced
VHDL	Very High Speed Integrated Circuit Hardware Description Language

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

เอฟพีจีเอนั้นได้จัดเป็นอุปกรณ์สารกึ่งตัวนำที่มีโครงสร้างภายในประกอบด้วยส่วนของวงจรลอจิกที่สามารถชนิดโปรแกรมกำหนดการทำงานได้ จึงสามารถออกแบบโครงสร้างภายในของชิปเอฟพีจีเอให้มีหน้าที่การทำงานเหมือนลอจิกเกตพื้นฐาน เช่น AND OR NOT หรือรวมกันเพื่อการทำงานที่มีความซับซ้อนมากขึ้น เช่น ฟังก์ชันเข้ารหัส ฟังก์ชันถอดรหัสหรือฟังก์ชันทางคณิตศาสตร์ เป็นต้น และสามารถสร้างบล็อกของหน่วยความจำจากพื้นที่ของสารกึ่งตัวนำสร้างเป็นหน่วยความจำภายในได้ การออกแบบเอฟพีจีเอนั้นได้ใช้ภาษาเอชดีแอล (Hardware Description Language: HDL) เช่น การออกแบบการประมวลผลดิจิตอล (Digital Signal Processing) เป็นต้น

ในการออกแบบวงจรดิจิตอลอิเล็กทรอนิกส์ที่มีเอฟพีจีเออยู่บนแผงวงจรด้วยนั้น จะช่วยให้ผู้ออกแบบสามารถลดขนาดของแผงวงจรและอุปกรณ์ลดลงได้ รวมทั้งสามารถออกแบบได้รวดเร็ว เนื่องจากวงจรที่ออกแบบนั้นสามารถจำลองการทำงาน (Simulation) ผ่านทางโปรแกรมคอมพิวเตอร์ออกแบบโปรแกรมการทำงาน เช่น ภาษาวีเอชดีแอล (Very-High-Speed Integrated Circuit Hardware Description Language: VHDL) เป็นต้น เมื่อทดสอบจากการจำลองการทำงานจนผลที่ได้ถูกต้อง ก็จะทำการบรรจุวงจรที่ออกแบบไว้บนเอฟพีจีเอสามารถนำไปใช้งานได้ทันทีและสามารถปรับเปลี่ยนหรือแก้ไขการทำงานของวงจรได้ภายหลังได้ อย่างไรก็ตามข้อกำหนด (Configuration) ของเอฟพีจีเอจะหายไปหลังจากปิดไฟเลี้ยง ดังนั้นจะต้องมีหน่วยความจำภายนอกเพื่อรักษาข้อกำหนดของเอฟพีจีเอไว้ ซึ่งจะมีกระบวนการอ่านข้อกำหนดนั้นโดยอัตโนมัติหลังจากได้รับไฟเลี้ยง

สัญญาณเสียงพูดนั้นเป็นสัญญาณสนใจในเรื่องของประมวลผลสัญญาณแบบฐานเวลาจริง (Real Time) เพื่อที่จะพัฒนาต่อในเรื่องของเทคโนโลยีต่างๆ และการนำมาประยุกต์ใช้ในงานทางการสื่อสารอีกเป็นจำนวนมาก การประมวลผลของสัญญาณเสียงเป็นการประมวลผลสัญญาณทางด้านดิจิตอลเข้ากับสัญญาณเสียงแบบฐานเวลาจริงโดย ซึ่งการประมวลผลดังกล่าวอาจส่งผลต่อการเกิดการหน่วงเวลาเกิดขึ้น การดำเนินการแบบฐานเวลาจริงโดยนั้นต้องอาศัยตัวประมวลผลที่มีความสามารถค่อนข้างสูง โดยเทคโนโลยีในปัจจุบันเอฟพีจีเอเป็นหนึ่งในทางเลือกที่มีการนำมาประยุกต์ใช้ในงานวิจัยด้านการประมวลผลสัญญาณเสียง

ผู้วิจัยได้ศึกษาและออกแบบระบบการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณรบกวนหรือสัญญาณเสียงที่ไม่ใช่เสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอโดยการใช้เส้นตัดสินใจ แต่

ทั้งนี้การประมวลผลของการแยกแยะสัญญาณเสียงพบว่ามีปัญหาในการประมวลผลแบบฐานเวลาจริง โดยและเสียงพูดในส่วนระงับการส่งสัญญาณเสียงที่ไม่ใช่เสียงพูดหรือเสียงรบกวน เพราะความเปลี่ยนแปลงของสัญญาณเสียงนั้นเกิดขึ้นอย่างรวดเร็วและคาดเดาล่วงหน้าได้ยาก รวมทั้งความถูกต้องในการแยกแยะเสียงพูดแบบฐานเวลาจริงโดยนั้นต้องอาศัยตัวประมวลผลที่มีความสามารถค่อนข้างสูง ดังนั้นผู้วิจัยจึงเลือกใช้บอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro จะมีชิป XC2VP30 เป็นหน่วยประมวลผลกลาง โดย XC2VP30 มีเกททั้งหมด 30,816 ลอจิกเซลล์ ถูกนำมาใช้ในการวิจัยครั้งนี้ มีอุปกรณ์อำนวยความสะดวกด้วยอุปกรณ์อินพุตและเอาต์พุต เช่น ชิป LM4550 AC97 Audio CODEC เป็นหน่วยควบคุมสัญญาณเสียง สามารถต่อร่วมกับหน่วยความจำภายนอกได้ เป็นต้น อีกหนึ่งปัญหาที่พบจากการระงับการส่งสัญญาณเสียงที่ไม่ใช่เสียงพูดหรือเสียงรบกวนคือ ผู้ใช้จะได้ยินเสียงที่เงียบมากจนรู้สึกเหมือนกับว่าโดนสายตัด จึงได้ศึกษาการส่งสัญญาณเสียงที่ไม่ใช่เสียงพูดหรือเสียงรบกวนด้วยการส่งสัญญาณศูนย์และสัญญาณสุ่มจากการส่งสัญญาณเสียงทั้งสัญญาณเพื่อศึกษาผลที่เกิดขึ้นโดยทั้งบนซอฟต์แวร์และฮาร์ดแวร์

ผู้วิจัยได้ใช้ประโยชน์จากการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด เนื่องจากสัญญาณเสียงพูดมีลักษณะเป็นกลุ่มข้อมูลเสียงและกลุ่มข้อมูลที่ไม่เป็นเสียง ทำการส่งเฉพาะกลุ่มข้อมูลที่เป็นข้อมูลเสียงเท่านั้น เพื่อทำให้ลดขนาดของข้อมูลในการส่งข้อมูลลง เพื่อเพิ่มความคุ้มค่าในการส่งข้อมูลภายในเครือข่าย เพราะการสนทนาการพูดกับการฟังมักไม่เกิดขึ้นพร้อมกัน การแยกแยะเสียงพูดและเสียงที่ไม่ใช่เสียงพูดที่มีความแม่นยำสูงนั้น ส่งผลให้สามารถลดการสิ้นเปลืองแถบความกว้างของสัญญาณและลดความคับคั่งของการใช้ทรัพยากรในเครือข่าย ดังนั้นงานวิจัยนี้จึงนำเสนอการพัฒนาเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณรบกวนแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ

1.2 วัตถุประสงค์ของการศึกษา

1.2.1 ศึกษาเทคนิคสำหรับการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด

1.2.2 พัฒนาการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดบนพื้นฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ

1.2.3 วิเคราะห์คุณภาพของสัญญาณเสียงพูดหลังจากแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด

1.3 ขอบเขตของการศึกษา

1.3.1 คัดเลือกเส้นตัดสินใจที่เหมาะสมสำหรับการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด

1.3.2 ออกแบบเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดให้สามารถลดขนาดของสัญญาณเสียงโดยมีอัตราบีบอัดข้อมูลไม่น้อยกว่า 1.5

1.3.3 วิเคราะห์คุณภาพของสัญญาณเสียงที่แยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดโดยใช้เส้นตัดสินใจได้ด้วยวิธีการ SNR และ NRMSE

1.3.4 สามารถออกแบบระบบแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ

1.4 ข้อจำกัดของการศึกษา

1.4.1 เสียงที่นำวิจัยต้องเป็นสัญญาณเสียงพูด โดยแบ่งเป็นเสียงเป็น 2 ส่วน ส่วนที่ 1 จะบันทึกจะแบ่งเป็น 4 กลุ่ม คือ เสียงผู้ชาย 5 วินาที 20 เสียง เสียงผู้หญิง 5 วินาที 20 เสียงผู้ชาย 60 วินาที 20 เสียง และ เสียงผู้หญิง 60 วินาที 20 เสียง ส่วนที่ 2 ใช้ในการประมวลผลบนหน่วยประมวลผลบนเอฟพีจีเอ จะรับเสียงจากไมโครโฟนและคอมพิวเตอร์

1.4.2 ประมวลผลโดยใช้โปรแกรม MATLAB และ Xilinx ISE

1.4.3 หน่วยประมวลผลจะใช้เอฟพีจีเอตระกูล Xilinx รุ่น Virtex-II Pro XC2VP30

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 พัฒนางานวิจัยที่เกี่ยวข้องกับการบีบอัดสัญญาณเสียงพูดแบบฐานเวลาจริงโดยโดยใช้บอร์ดเอฟพีจีเอ โดยใช้เทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด

1.5.2 เป็นแนวทางในการพัฒนาการบีบอัดสัญญาณเสียงเพื่อลดขนาดของข้อมูลและสามารถนำไปประยุกต์ใช้ในระบบการสื่อสาร เช่น การบีบอัดข้อมูลเสียงพูดจากสัญญาณเสียงที่มีกว่า 2 แหล่งกำเนิดพร้อมกัน

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การพัฒนาเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณรบกวนหรือสัญญาณเสียงที่ไม่ใช่เสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอเป็นการประยุกต์เอฟพีจีเอใช้ในงานวิจัยด้านการประมวลผลสัญญาณ โดยต้องศึกษาทฤษฎีที่เกี่ยวข้องดังต่อไปนี้คือ เอฟพีจีเอ ภาษาวีเอชดีแอล คุณลักษณะเสียงพูดของมนุษย์ การประมวลผลสัญญาณเสียงพูด อัตราการบีบอัดสัญญาณ และการวัดความค่าผิดพลาดและการวัดค่าประสิทธิภาพของสัญญาณ นอกจากนี้ยังกล่าวถึงงานวิจัยที่เกี่ยวข้องกับการพัฒนาเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ

2.1 เอฟพีจีเอ

เอฟพีจีเอเป็นอุปกรณ์สารกึ่งตัวนำที่มีโครงสร้างภายในประกอบด้วยส่วนของวงจรลอจิกที่สามารถชนิดโปรแกรมกำหนดการทำงานได้ จึงสามารถออกแบบโครงสร้างภายในของเอฟพีจีเอให้มีหน้าที่การทำงานเหมือนลอจิกเกตพื้นฐาน เช่น AND OR NOT หรือใช้รวมกันเพื่อทำงานที่มีความซับซ้อนมากขึ้น เช่น ฟังก์ชันการเข้ารหัส ฟังก์ชันถอดรหัสหรือฟังก์ชันทางคณิตศาสตร์ เป็นต้น และสามารถสร้างบล็อกของหน่วยความจำจากพื้นที่ของสารกึ่งตัวนำสร้างเป็นหน่วยความจำภายในได้ การโปรแกรมซึ่งก็จะมีอยู่ด้วยกัน 2 แบบคือ การโปรแกรมโดยการทำให้เกิดการเปลี่ยนแปลงทางกายภาพของตัวชิปและการโปรแกรมโดยการใช้หน่วยความจำอธิบายได้ดังนี้

1) การโปรแกรมโดยการทำให้เกิดการเปลี่ยนแปลงทางกายภาพนั้นจะแบ่งได้เป็น 2 วิธีคือ FUSE เป็นวิธีการโปรแกรมที่สามารถทำได้เพียงครั้งเดียว ซึ่งหลังจากที่โปรแกรมแล้วจุดเชื่อมต่อจะขาดจากกัน และ ANTI FUSE เป็นวิธีการโปรแกรมที่คล้ายกับแบบ FUSE แต่ต่างกันที่หลังจากทำการโปรแกรมแล้วจุดเชื่อมต่อจะเชื่อมต่อถึงกัน

2) การโปรแกรมโดยการใช้หน่วยความจำนั้นจะแบ่งได้เป็น 2 วิธีคือ Electrically Erasable Programmable Read Only Memory (EEPROM) Based FPGA การโปรแกรมแบบนี้มักเรียกว่า CPLD ซึ่งเทคโนโลยีที่ใช้จะเหมือนกับ EEPROM ทำให้มีความจำของเกทต่ำ โดยทั่วไปจะน้อยกว่า 20,000 เกท แต่ข้อดีของ EEPROM บนพื้นฐานเอฟพีจีเอก็เป็นคือสามารถเก็บข้อมูลโปรแกรมลงไปได้โดยไม่จำเป็นต้องมีไฟเลี้ยง และในการโปรแกรมจะใช้ทรานซิสเตอร์1 ตัวต่อ 1 บิต ซึ่งการโปรแกรมสามารถทำได้ประมาณ 10,000 ครั้ง และอีกวิธีคือ Static Read Only Memory (SRAM) Based FPGA การที่

โปรแกรมแบบนี้จะใช้เทคโนโลยีในการโปรแกรมเหมือนกับ SRAM ทำให้สามารถโปรแกรมซ้ำได้ไม่จำกัดจำนวนครั้ง นอกจากนี้ยังมีความจุของเกทในระดับปานกลางถึงสูงมาก ประมาณ 10,000 - 1,000,000 เกท ซึ่งข้อดีของ SRAM บนพื้นฐานเอฟพีจีเอนั้นก็คือใช้เวลาในการโปรแกรมน้อย ซึ่งการโปรแกรมทำได้ง่ายเทียบได้กับการเขียน SRAM ทั่วไป และเหมาะสำหรับการออกแบบวงจรที่มีแบบวงจรที่มีความสลับซับซ้อน ส่วนข้อเสียคือ ไม่สามารถเก็บโปรแกรมในภาวะที่ไม่มีไฟเลี้ยงได้ ดังนั้นเอฟพีจีเอชนิดนี้จึงมักใช้ควบคู่กับ ROM เพื่อเก็บโปรแกรมและทำการโหลดโปรแกรมลงในตัวชิปในขณะที่เริ่มต้นใช้งาน

ปัจจัยที่ทำให้การออกแบบเอฟพีจีเอทำได้ง่ายและสะดวกรวดเร็วเพราะว่าผู้ออกแบบไม่จำเป็นต้องทราบถึงโครงสร้างภายในของตัวชิป เพียงแต่มีความรู้เกี่ยวกับขั้นตอนการออกแบบลอจิกก็เพียงพอแล้วต่างกับการใช้ไมโครโปรเซสเซอร์ซึ่งจำเป็นต้องศึกษาโครงสร้างภายในรวมถึงภาษาแอสเซมบลี (Assembly) ของไมโครโปรเซสเซอร์นั้นด้วยและรวมถึงการออกแบบเอฟพีจีเอนั้นได้ใช้ภาษาวีเอชดีแอลซึ่งเป็นวิธีการที่มีความยืดหยุ่นสูงสามารถทำได้รวดเร็วและจำเป็นต้องทราบถึงลักษณะของวงจรที่ต้องการว่าจะเชื่อมต่อกันอย่างไร เพียงแต่กำหนดลักษณะการทำงานให้มันจากนั้นตัวซอฟต์แวร์จะทำ การสังเคราะห์และเพิ่มประสิทธิภาพให้ทั้งหมด นอกจากนี้จากภาษาที่ใช้ยังเป็นมาตรฐานเดียวกันสามารถใช้ได้กับชิปทุกตัวและทุกบริษัท ในส่วนการโปรแกรมนั้นสามารถทำการโปรแกรมข้อมูลผ่านสายดาวน์โหลดทางพอร์ตของคอมพิวเตอร์ก็สามารถโปรแกรมตัวชิปขณะที่อยู่ในระบบได้โดยไม่จำเป็นต้องถอดโปรแกรมข้างนอกซึ่งทำให้ง่ายในการแก้ไขและพัฒนา

ภาพที่ 2.1 ชิปเอฟพีจีเอ Xilinx Virtex-II Pro XC2VP30

2.1.1 การโปรแกรมอุปกรณ์เอฟพีจีเอ

การดาวน์โหลดนี้ก่อนอื่นนั้นต้องทำการแปลงแบบวงจรรวมที่ได้ให้เป็นข้อมูลวงจร ซึ่งอยู่ในรูปของบิตสตรีม ก่อนแล้วจึงดาวน์โหลดลงไปเพื่อให้อุปกรณ์เอฟพีจีเอมีฟังก์ชันการทำงานตามโมเดลที่ผู้ออกแบบต้องการ ซึ่งในขั้นตอนนี้จะใช้วิธีที่แตกต่างกันออกไป ซึ่งโดยอุปกรณ์เอฟพีจีเอของแต่ละบริษัทผู้ผลิตที่ใช้คือ ในกรณีที่เป็นอุปกรณ์เอฟพีจีเอชนิดที่ต้องโปรแกรมโดยใช้วิธี SRAM นั้น ในการใช้งานของผู้ออกแบบจะต้องเก็บข้อมูลวงจรไว้ในหน่วยความจำประเภท EPROM หรือ Serial PROM ด้วยเพื่อจะใช้งานสะดวกขึ้น เพราะว่าในการใช้งานโมเดลครั้งต่อไปไม่ต้องดาวน์โหลดข้อมูลวงจรจากเครื่องคอมพิวเตอร์อีก เพราะมีข้อมูลวงจรเก็บอยู่ในหน่วยความจำอยู่ แล้วแต่กรณีที่อุปกรณ์เอฟพีจีเอเป็นชนิดที่โปรแกรมโดยใช้วิธี EPROM หรือ Anti-FUSE นั้นก็ไม่จำเป็นที่จะต้องมีหน่วยความจำสำหรับเก็บข้อมูลวงจร เพราะว่าอุปกรณ์เอฟพีจีเอชนิดนี้เมื่อดาวน์โหลดข้อมูลวงจรลงไปข้อมูลที่ดาวน์โหลดลงไปก็ยังคงอยู่ในอุปกรณ์เอฟพีจีเอและครั้งต่อไปก็ใช้งานโมเดลที่ออกแบบไว้

2.1.2 เครื่องมือสำหรับการออกแบบเอฟพีจีเอ

ส่วนสำคัญที่ใช้ในการทำเอฟพีจีเอคือซอฟต์แวร์ที่ใช้ตั้งแต่เขียนโค้ดอธิบายฮาร์ดแวร์จนกระทั่งเมื่อทำการดาวน์โหลดลงในอุปกรณ์เอฟพีจีเอซึ่งในซอฟต์แวร์ที่ใช้นี้ต้องเป็นซอฟต์แวร์ที่ทำงานต่อเนื่องกันได้ สำหรับซอฟต์แวร์ที่ใช้ทำการจำลองการทำงานของวงจรนั้น ต้องสามารถใช้งานต่อเนื่องกับซอฟต์แวร์ที่ใช้ทั้งระบบเพราะโมเดลที่ได้จากวิธีการทำขั้นตอนต่างๆ ด้วยซอฟต์แวร์ ต่างๆ ต้องเอามาจำลองการทำงานได้ และในการจำลองการทำงานของวงจรที่ใช้นั้นควรใช้ซอฟต์แวร์ตัวเดียวกันตลอดทั้งระบบเพื่อจะได้เปรียบเทียบผลได้ง่าย ในอดีตซอฟต์แวร์ส่วนใหญ่จะใช้งานอยู่บนคอมพิวเตอร์สมรรถนะสูงอย่างเวิร์คสเตชั่น โดยในปัจจุบันมีการพัฒนาซอฟต์แวร์ที่ใช้บนพีซีเพิ่มขึ้นซึ่งเราสามารถลดค่าใช้จ่ายในด้านอุปกรณ์คอมพิวเตอร์ตัวอย่างโปรแกรมที่ใช้ เช่น Xilinx ISE 10.1 เป็นต้น

2.2 ภาษาวีเอชดีแอล

ปัจจุบันในการออกแบบระบบดิจิตอลมีความซับซ้อนและขนาดของระบบเพิ่มทุกขณะส่งผลให้มีการนำคอมพิวเตอร์เพื่อช่วยในการออกแบบระบบดิจิตอลมากขึ้น อีกทั้งอุปกรณ์และวิธีออกแบบใหม่ๆได้พัฒนาขึ้นเพื่อช่วยอำนวยความสะดวกให้นักออกแบบมากขึ้น การออกแบบโดยใช้การวาดวงจร (Capture Schematic) ผู้ออกแบบจะต้องมีทักษะสูงในการออกแบบ และต้องใช้เวลามากในการออกแบบ จำลองการทำงาน และตลอดจนถึงการแก้ไขความถูกต้องของระบบ (Debugging)

สำหรับภาษาบรรยายฮาร์ดแวร์เอชดีแอล ก็เป็นเครื่องมือที่ได้รับการพัฒนามาอย่างต่อเนื่องในการปรับปรุงขบวนการออกแบบระบบดิจิตอลให้เป็นไปอย่างมีประสิทธิภาพ

2.2.1 ประวัติความเป็นมาของภาษาวีเอชดีแอล

ภาษาวีเอชดีแอลเป็นภาษาบรรยายฮาร์ดแวร์ประเภทหนึ่งและเป็นภาษาโปรแกรมระดับสูงที่ใช้ในการออกแบบฮาร์ดแวร์ในระบบดิจิตอล ตัวภาษาสามารถบรรยายพฤติกรรมการทำงานในรูปแบบลำดับขั้น (Hierarchy) และสามารถเขียนได้หลายรูปแบบ ดังนั้นภาษาวีเอชดีแอลจึงน่าสนใจที่จะศึกษาและนำไปใช้งานเป็นอย่างยิ่ง สำหรับมาตรฐานของภาษาที่ใช้บรรยายพฤติกรรมวงจรหรือฮาร์ดแวร์ สามารถสรุปได้ดังนี้

1) ต้องเป็นภาษาที่นำไปเขียนในรูปแบบระบบดิจิตอลและมีคุณสมบัติที่สามารถเข้าใจได้ทั้งมนุษย์และเครื่องคอมพิวเตอร์โดยไม่มีการแปลหรือเปลี่ยนแปลงอีก

2) ต้องเป็นภาษาที่เขียนขึ้นสำหรับใช้จำลองการทำงานของวงจร

3) ต้องการภาษาที่เป็นมาตรฐานหรือภาษากลางที่ทำให้สามารถเผยแพร่ผลงานการออกแบบกันภายในกลุ่มนักออกแบบด้วยกันได้

2.2.2 การออกแบบระบบดิจิตอล

ในการออกแบบระบบดิจิตอล จะเริ่มต้นจากการกำหนดหน้าที่การทำงานของระบบที่ต้องการซึ่งโดยส่วนใหญ่แล้วระบบดิจิตอลประกอบด้วย 2 ส่วนคือ ส่วนที่เป็นระบบฮาร์ดแวร์และส่วนที่เป็นซอฟต์แวร์โดยทั่วไปแล้วการออกแบบทั้ง 2 ส่วนดังกล่าวนี้จะต้องทำไปพร้อมๆกัน หลังจากนั้นจะเป็นขั้นตอนการกำหนดหน้าที่การทำงานของระบบฮาร์ดแวร์และซอฟต์แวร์ โดยในระหว่างการออกแบบนั้นผู้ออกแบบจะต้องมีการทดสอบการทำงานร่วมกันเพื่อทำให้สามารถแก้ไขความผิดพลาดในการออกแบบที่เกิดขึ้น

2.2.3 การออกแบบจากบนลงล่าง

ในการพัฒนาวงจรรวมดิจิตอลขนาดใหญ่ที่มีความซับซ้อน ผู้ออกแบบมักจะมองการออกแบบในรูปของบล็อกไดอะแกรมก่อนที่ทำวิเคราะห์ให้ลึกถึงรายละเอียดต่อไป ภาษาวีเอชดีแอลนั้นอนุญาตให้อธิบายและวิเคราะห์การ ทำงานของแต่ละบล็อก รวมถึงการปรับปรุงการทำงานจากผลที่วิเคราะห์เพื่อให้ได้การทำงานตามที่ต้องการ นอกจากนี้ยังสามารถเพิ่มเติมในรายละเอียดในแต่ละขั้นตอนได้ ซึ่งหลักการนี้สอดคล้องกับหลักการออกแบบจากบนลงล่าง

จากภาพที่ 2.1 แสดงถึงขั้นตอนของการออกแบบจากบนลงล่าง การออกแบบจากบนลงล่างในแต่ละขั้นตอนมีดังนี้

1) สร้างข้อกำหนดของความต้องการ และวิเคราะห์ระบบ เพื่อหาแนวความคิดและหลักการในการแก้ปัญหา

2) เขียนรูปแบบของระบบที่ต้องการออกแบบโดยภาษาวีเอชดีแอลหรือภาษาเอชดีแอลอื่นๆ สำหรับบรรยายพฤติกรรมการ ทำงาน พร้อมทั้งจำลองการทำงาน เพื่อเปรียบเทียบและตรวจสอบความถูกต้องกับข้อกำหนด

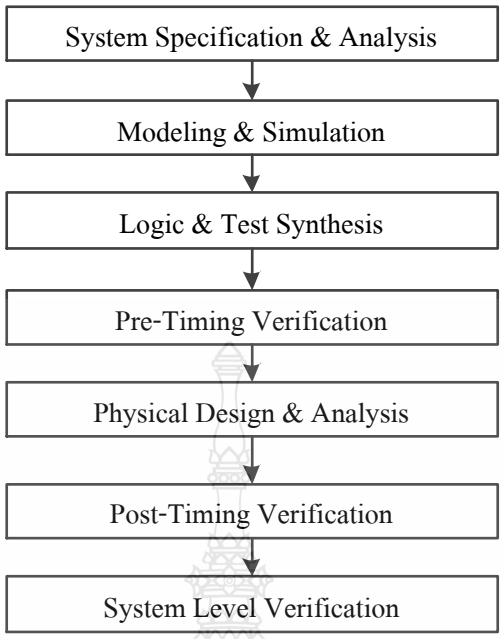
3) หลังจากที่ได้หลักการขั้นต้นพร้อมแนวความคิดที่ผ่านการตรวจสอบแล้วหลักการนี้จะถูกเพิ่มเติมในรายละเอียดลงมา เป็นลำดับขั้นที่สอง จนกระทั้งอยู่ในระดับที่จะนำไปผลิตวงจริง หรือสังเคราะห์ในขั้นตอนนี้เองเทคโนโลยีที่จะมารองรับ วงจรออกแบบจะถูกกำหนดขึ้น และระบบช่วยการออกแบบจะสังเคราะห์วงจรที่ได้จากรูปแบบที่เขียนขึ้นให้อยู่ในรูปของ วงจรที่ประกอบด้วยอุปกรณ์อิเล็กทรอนิกส์ หรือวงจรในระดับเกท และการเชื่อมต่อระหว่างกันของอุปกรณ์เหล่านั้นหรือไม่ก็อยู่ในรูปของระดับเกท (Gate-Level Netlist) ที่สามารถนำไปผลิตในอุปกรณ์อื่นได้

4) หลังจากการสังเคราะห์วงจรให้อยู่ในระดับเกทแล้ว ข้อมูลนี้จะใช้สำหรับจำลองการทำงานในเรื่อง ความถูกต้องของฟังก์ชัน พร้อมกับนำข้อมูลที่เกี่ยวข้องกับเวลาเข้ามาประกอบการพิจารณาด้วย ซึ่งตามปกติแล้วอุปกรณ์ ทางอิเล็กทรอนิกส์ทุกชิ้นจะมีเวลาหน่วงของการแพร่กระจาย (Propagation Delay Time) เสมอ ถึงแม้ว่าจะมันเป็นเวลาที่น้อยมากในระดับนาโนวินาทีก็ตาม แต่ถ้าภายในวงจรหนึ่งประกอบด้วยเกทของฟังก์ชันต่างๆ จำนวน 10,000 เกท ขึ้นไป เวลาดังกล่าวนี้จะสะสมกันมากขึ้น จนอาจทำให้การทำงานของวงจรรวมทั้งหมดผิดพลาดไป หรือไม่สามารถทำ งานในย่านความถี่สัญญาณนาฬิกาที่สูงได้

5) ผลิตเป็นวงจรโดยที่นำข้อมูลที่ได้จากการสังเคราะห์มาผลิต ซึ่งอาจจะอยู่ในรูปของแผงวงจรไฟฟ้า ที่ประกอบด้วยอุปกรณ์หลายๆ ชิ้นหรืออยู่ในรูปของวงจรรวม ASIC

6) ทำการตรวจสอบการทำงานและตัวแปรทางด้านเวลาทั้งหมด เพื่อความถูกต้องของวงจรเป็นครั้งสุดท้ายก่อนที่นำไปรวมเข้ากับอุปกรณ์อื่นๆ ให้เป็นระบบดิจิตอล เนื่องจากในขั้นตอนนี้วงจรที่ได้ออกแบบ จะประกอบด้วยจุดต่อทางอินพุทและเอาท์พุท ซึ่งเป็นจุดต่อสำหรับการรับและส่งสัญญาณกับภายนอก

7) นำวงจรที่ออกแบบไว้ประกอบเข้ากับอุปกรณ์อื่นๆ ให้เป็นระบบที่สมบูรณ์ แล้วทำการทดสอบการทำงานทั้งระบบร่วมกับอุปกรณ์อื่นๆ อีกครั้งเพื่อควบคุมคุณภาพของผลิตภัณฑ์



ภาพที่ 2.2 ขั้นตอนการออกแบบจากบนลงล่าง

2.3 คุณลักษณะเสียงพูดของมนุษย์

เสียงพูดจะมีลักษณะเป็นคลื่นตามยาว (Longitudinal Wave) ซึ่งเกิดจากการสั่นของอนุภาคตัวกลาง นั่นคือ อากาศ และทิศทางการสั่นของอนุภาคจะอยู่ในทิศเดียวกันกับทิศทางของการเคลื่อนที่คลื่นเสียงเป็นคลื่นที่เปลี่ยนแปลงไปตามเวลา เสียงพูดแบ่งออกได้เป็น 2 ชนิดตามการกำเนิดเสียงหรือการกระตุ้น

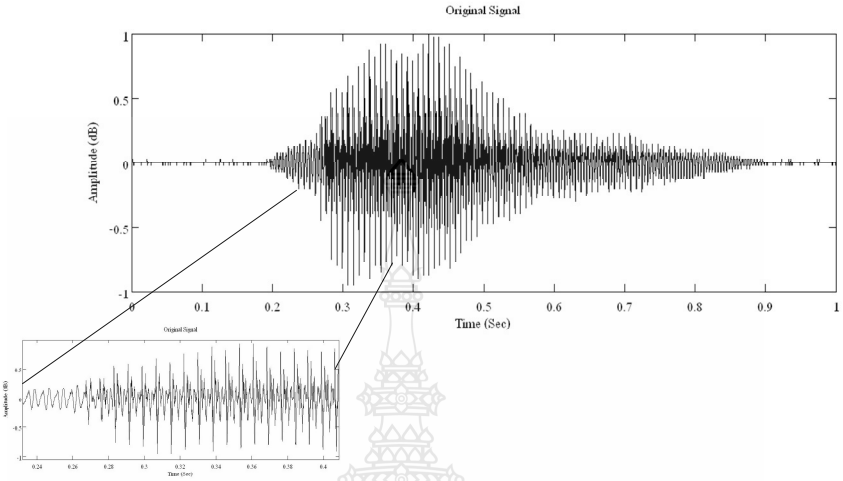
2.3.1 เสียงก้องหรือเสียงโฆษะ

เสียงก้องหรือเสียงโฆษะ (Voiced) เกิดจากการบังคับอากาศให้ผ่านช่องสายเสียง ทำให้มีการเปลี่ยนแปลงความตึงหย่อนของเส้นเสียง โดยเส้นเสียงจะสั่นและเกิดเป็นพัลส์ (Pulse) ของอากาศไปกระตุ้นอวัยวะกำทอนเกิดเป็นเสียงก้อง ตัวอย่างเสียงก้องได้แก่ เสียงสระ เสียงพยัญชนะ ที่ต้องออกเสียงจากลำคอ (Voiced Consonants) จะแสดงในภาพที่ 2.3

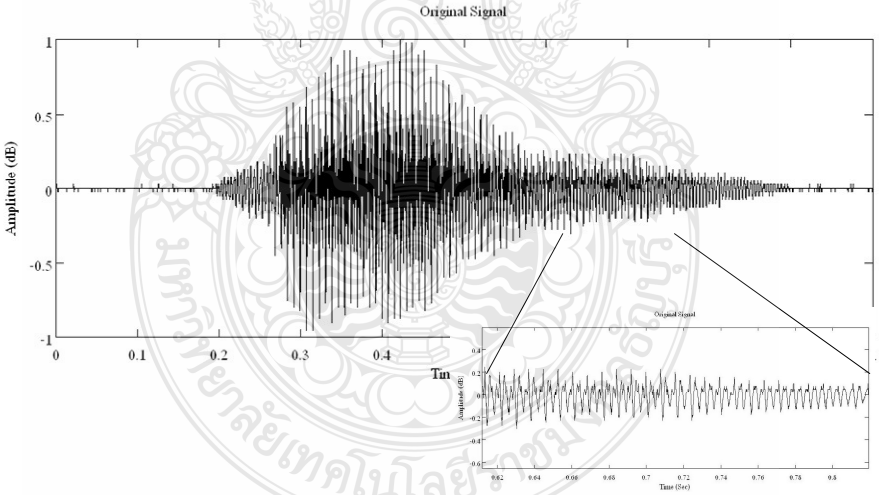
2.3.2 เสียงไม่ก้องหรืออโฆษะ

เสียงไม่ก้องหรืออโฆษะ (Unvoiced หรือ Voiceless) เป็นเสียงที่ไม่เกิดจากการสั่นของเส้นเสียง แต่เกิดในช่องปากหรือโพรงจมูก โดยอวัยวะภายในช่องปาก ริมฝีปาก ซึ่งอาจขัดขวางการไหลของอากาศให้ผ่านได้เป็นช่องเล็กๆ ทำให้อากาศมีการไหลผ่านอย่างรวดเร็วและปั่นป่วนจนกระทั่ง

สามารถสร้างเป็นเสียงรบกวนช่วงความถี่กว้าง (Broad Spectrum Noise) ตัวอย่างเสียงไม่ก้องได้แก่ เสียงพยัญชนะที่ไม่ได้เกิดจากลำคอ (Voiceless Consonants) จะแสดงในภาพที่ 2.4



ภาพที่ 2.3 ลักษณะสัญญาณเสียงก้องหรือเสียงโฆษะ

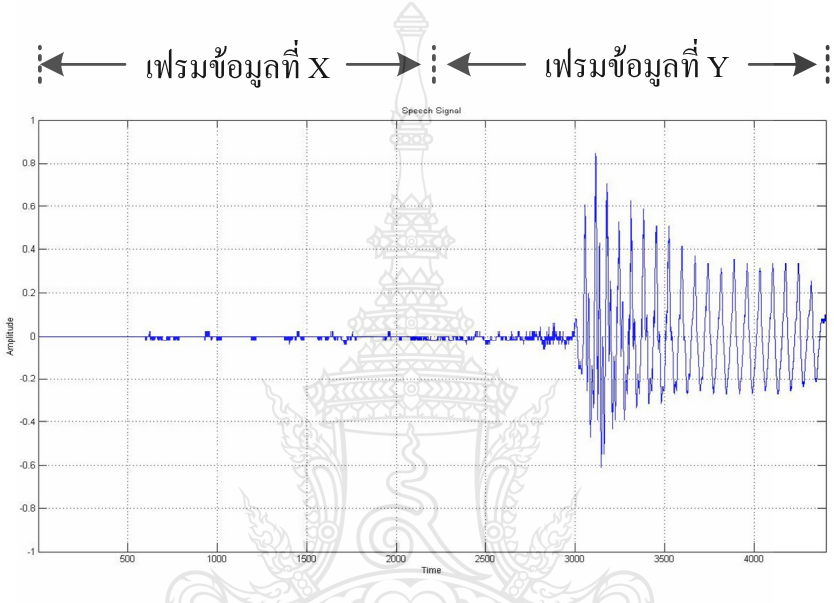


ภาพที่ 2.4 ลักษณะสัญญาณเสียงไม่ก้องหรือเสียงอโฆษะ

2.4 การประมวลผลสัญญาณเสียงพูด

โดยธรรมชาติของสัญญาณเสียงพูดจะไม่เสถียรและเปลี่ยนแปลงตามเวลา (Non-Stationary) ดังนั้นเมื่อต้องการนำสัญญาณเสียงพูดมาประมวลผลสัญญาณดิจิตอล จึงจำเป็นต้องแบ่ง

สัญญาณออกเป็นช่วงสั้นๆ เพื่อให้สัญญาณเสียงมีความเสถียรและไม่มีการเปลี่ยนแปลงตามเวลา (Stationary) จากนั้นจึงนำสัญญาณไปประมวลผลต่อไปได้กรอบเสียงพูด (Speech Frame) ควรมีความยาวประมาณ 10 - 30 มิลลิวินาที (ms) จะทำให้สัญญาณเสียงพูด ในแต่ละกรอบเป็นสัญญาณที่มีความเสถียรและไม่เปลี่ยนแปลงตามเวลา กรรมวิธีการประมวลสัญญาณเสียงพูดเบื้องต้นจะปรับสัญญาณเสียงพูดให้มีความเหมาะสมในการนำสัญญาณไปวิเคราะห์ในรูปแบบต่างๆ ต่อไป



ภาพที่ 2.5 การแบ่งสัญญาณเสียงออกเป็นเฟรมย่อย

2.4.1 การแยกแยะสัญญาณที่เป็นเสียงและเสียงรบกวนหรือสัญญาณเสียงที่ไม่ใช่เสียงพูด

1) วิธีการใช้ค่าแอมปลิจูด (Amplitude) เริ่มต้นจากการหาค่าสัมบูรณ์ของแอมปลิจูดของสัญญาณเสียงแต่ละจุดที่ n ใดๆ ดังสมการที่ 2.1

$$E(n) = |s(n)| \tag{2.1}$$

โดยได้ใช้แนวความคิดที่ว่าส่วนที่เป็นสัญญาณเสียงจะมีค่าที่สูงกว่าส่วนที่ไม่ใช่สัญญาณเสียง ข้อดีของการใช้ค่าแอมปลิจูดคือคำนวณง่ายและใช้เวลาน้อย แต่มีข้อเสียคือมีผิดพลาดได้ถ้ามีสัญญาณรบกวนที่มีค่าแอมปลิจูดสูง

2) วิธีการใช้ค่าพลังงาน (Energy) ในการหาค่าพลังงานของสัญญาณเสียง โดยทั่วไปจะใช้เทคนิคการแบ่งสัญญาณเสียงออกเป็นส่วนย่อย (Frame) จากนั้นจึงนำสัญญาณเสียงแต่ละส่วนย่อยมาหาค่าพลังงาน (Short-Time Energy) โดยได้ใช้แนวความคิดที่ว่าส่วนที่เป็นสัญญาณเสียงจะเป็นส่วนที่มีระดับพลังงานที่สูงกว่าส่วนที่ไม่ใช่สัญญาณเสียง

3) วิธีการหาสเปกโตรแกรม (Spectrogram) จะเป็นการแสดงแผนภาพในแกนความถี่และแกนเวลาโดยจะใช้กลักการแปลงฟูเรียร์เพื่อใช้ในการสร้างแผนภาพสเปกโตรแกรมซึ่งจะทำให้ทราบว่าสัญญาณเสียงที่เวลาต่างๆมีความถี่เป็นอย่างไรและเมื่อพิจารณาในแนวแกนแซดจะทำให้ทราบถึงค่าระดับพลังงานของสัญญาณเสียงที่ความถี่นั้นๆซึ่งจะบ่งบอกเป็นค่าสีต่างๆบนแผนภาพ

4) วิธีอัตราการตัดผ่านศูนย์ (Zero-Crossing Rate) ใช้หลักการรูปสัญญาณเสียงมีการตัดกับแกนเวลา ซึ่งบริเวณที่เป็นสัญญาณเสียงจะมีอัตราการตัดกับแกนเวลาน้อยกว่าบริเวณที่ไม่ใช่สัญญาณเสียงและเนื่องจากบริเวณของสัญญาณเสียงส่วนใหญ่จะมีค่าพลังงานอยู่ในช่วงความถี่ต่ำและบริเวณที่ไม่ใช่สัญญาณเสียงจะมีพลังงานอยู่ในช่วงความถี่สูงประกอบกับวิธีอัตราการตัดผ่านศูนย์มีความสัมพันธ์โดยตรงกับความถี่ของสัญญาณเสียงโดยสมการแสดงได้ดังสมการที่ 2.2 และ 2.3

$$Sign\left(s\left(n\right)\right)=\begin{cases}1 & ;s\left(n\right)>0\\ 0 & ;Others\end{cases} \tag{2.2}$$

$$Z=\frac{1}{2N}\sum_{n=1}^{N}\left|Sign\left(s\left(n\right)\right)-Sign\left(s\left(n-1\right)\right)\right| \tag{2.3}$$

เทคนิคในการวิเคราะห์สัญญาณเสียงพูดในการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดและแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดในการวิจัยได้ประยุกต์เทคนิคที่กล่าวมาจนได้เป็นเทคนิคการใช้เส้นตัดสินใจ รวมทั้งการใช้คุณสมบัติ การเก็บข้อมูลสัญญาณเสียงออกเป็นเฟรมย่อยด้วยหน่วยความจำเพื่อให้กระบวนการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดสามารถทำงานแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ โดยในขั้นตอนของวิธีดังกล่าวได้ถูกอธิบายไว้ในบทที่ 3

2.5 อัตราการบีบอัดสัญญาณ

อัตราการบีบอัดสัญญาณ (Compression Ratio) คือ อัตราส่วนระหว่างสัญญาณอินพุตที่ได้เปรียบเทียบกับสัญญาณเอาต์พุต เนื่องจากสัญญาณเสียงที่ใช้ในวิทยานิพนธ์ฉบับนี้มีขนาดความยาวที่

แตกต่างกัน ดังนั้นจึงมีการนำเอาค่าของเฟรมที่บอกสถานะที่แสดงว่าเป็นเฟรมเสียงพูด หรือไม่เป็นเฟรมของเสียงพูดมาร่วมในการคำนวณหาค่าอัตราการบีบอัดสัญญาณ ซึ่งหาได้ดังสมการที่ 2.4

$$\text{Compression Ratio} = \frac{\text{x(n)}}{\text{y(n)}} \tag{2.4}$$

โดยที่ $x(n)$ คือ สัญญาณต้นฉบับ

$y(n)$ คือ สัญญาณที่ส่ง

2.6 การวัดความค่าผิดพลาด

การวัดค่าความผิดพลาด (Error) ของสัญญาณจะใช้การวัดค่าเชิงวัตถุ เนื่องจากเป็นวิธีการที่ทำได้ง่ายไม่ยุ่งยากใช้เวลาน้อย การหาค่าความผิดพลาดระหว่างสัญญาณก่อนบีบอัดและสัญญาณที่ผ่านการบีบอัดแล้ว โดยในงานวิจัยนี้ได้ใช้การวัดความค่าผิดพลาดด้วยวิธี Normalized Root Mean Squared Error (NRMSE) ซึ่งเป็นวิธีที่ใช้ในการคำนวณหาประสิทธิภาพของการแทนค่าข้อมูลที่ขาดหายไป ซึ่งหาได้ดังสมการที่ 2.5

$$NRMSE = \sqrt{\frac{(p(n)-r(n))^2}{(p(n)-\mu p(n))^2}} \tag{2.5}$$

โดยที่ $p(n)$ คือ สัญญาณต้นฉบับ

$r(n)$ คือ สัญญาณที่สร้างขึ้นมาใหม่

$\mu p(n)$ คือ ค่าเฉลี่ยของสัญญาณต้นฉบับ

2.7 การวัดค่าประสิทธิภาพของสัญญาณ

การหาค่าประสิทธิภาพของสัญญาณจะใช้การวัดค่าเชิงวัตถุ ซึ่งสามารถวัดคุณภาพของอัลกอริทึมในการบีบอัดข้อมูลได้ โดยการหาค่าประสิทธิภาพของสัญญาณนั้นได้ใช้เทคนิค Signal to Noise Ratio (SNR) เป็นการพิจารณาหาค่าอัตราส่วนสัญญาณเสียงต้นฉบับต่อสัญญาณรบกวนแล้วทำการเปรียบเทียบค่าระหว่างสัญญาณเสียงต้นฉบับกับสัญญาณเสียงบีบอัด โดยที่การหาค่าประสิทธิภาพของสัญญาณ SNR ซึ่งหาได้ดังสมการที่ 2.6

$$SNR = 10\log_{10}\left|\frac{\sigma_x^2}{\sigma_e^2}\right| \quad (2.6)$$

โดยที่ σ_x^2 คือ ค่าเฉลี่ยกำลังสองของสัญญาณ

σ_e^2 คือ ค่าเฉลี่ยระหว่างของความแตกต่างระหว่างสัญญาณต้นฉบับเทียบกับสัญญาณคืนกลับ

2.8 ระบบแบบฐานเวลาจริง

ระบบแบบฐานเวลาจริง (Real Time Systems) คือระบบที่สามารถให้ผลการตอบสนองจากระบบอย่างทันทีทันใดเมื่อได้รับอินพุตเข้าไป ในทางอุดมคติระบบแบบฐานเวลาจริงนั้นจะเป็นระบบที่ไม่เสียเวลาในการประมวลผลหรืออาจจะกล่าวได้ว่าเวลาที่ใช้ในการประมวลนั้นเป็นศูนย์ แต่ในทางปฏิบัติการทำงานแบบฐานเวลาจริงมีข้อจำกัดจึงไม่สามารถทำงานได้ตามอุดมคติ โดยทำได้เพียงการลดเวลาการประมวลผลให้น้อยที่สุดจนไม่สามารถเห็นความแตกต่างของช่วงเวลาที่ป้อนอินพุตเข้าไปและได้รับเอาต์พุตออกมา ซึ่งในส่วนของเวลาที่เกิดความแตกต่างนี้เรียกว่าเวลาตอบสนอง (Response Time) ผู้ใช้งานทั่วไปต้องการเวลาตอบสนองให้น้อยที่สุดเพื่อประสิทธิภาพของระบบที่ดี แต่เนื่องการทำงานบนระบบแบบฐานเวลาจริงจะมีเวลาในการทำงานที่ถูกจำกัด ดังนั้นถ้าระบบให้ผลลัพธ์ได้ไม่ทันระบบก็จะล้มเหลว ในการจำแนกระบบแบบฐานเวลาจริงจะแบ่งเป็น 2 ระบบดังนี้

2.8.1 ระบบแบบฐานเวลาจริงตายตัว (Hard Real Time Systems) เป็นระบบที่รับรองว่าจะได้รับการตอบสนองภายในเวลาที่กำหนด ไม่สามารถหยุดรอไม่ได้และต้องทำงานเสร็จตรงตามเวลา ระบบนี้จะต้องมีการจัดการข้อมูล ซึ่งในการเก็บข้อมูลจะเก็บในหน่วยความจำช่วงสั้น (Short-Term Memory) หรือหน่วยความจำประเภทรอม

2.8.3 ระบบแบบฐานเวลาจริงยืดหยุ่น (Soft Real Time Systems) เป็นระบบที่มีความเข้มงวดน้อยกว่าระบบแบบฐานเวลาจริงตายตัวเพราะสามารถรองรับการประมวลผลตามความสำคัญของงาน โดยสามารถประมวลผลงานที่สำคัญมากกว่าให้ประมวลผลก่อนได้

2.9 งานวิจัยที่เกี่ยวข้อง

งานวิจัยทางด้านการประยุกต์ใช้งานและการประมวลผลสัญญาณเสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอนั้น ได้มีการพัฒนามาอย่างหลากหลายในหลายแนวทางและมีความแตกต่างกันออกไปตามวัตถุประสงค์ของการวิจัยนั้นๆ ในการวิจัยครั้งนี้จึงได้คัดเลือกงานวิจัยที่เกี่ยวข้องและมีความใกล้เคียงกับงานวิจัยที่กำลังศึกษา ดังตัวอย่างต่อไปนี้

2.9.1 ในงานวิจัย [1] นำเสนอในหัวข้อ “Enhancement Pattern Analysis Technique for Voiced/Unvoiced Classification” งานวิจัยนี้ได้เสนอ 4 เทคนิคในการวิเคราะห์การแยกแยะเสียงโฆษะหรือเสียงอโฆษะบนพื้นฐานของสัมประสิทธิ์คู่เส้นสเปกตรัม โดยเทคนิคทั้งหมดจะถูกเปรียบเทียบประสิทธิภาพกับเทคนิคฉันทลักษณ์ งานวิจัยนี้ได้นำสัญญาณเสียงพูด 10 นาทีมาเข้ากระบวนเทคนิคการตัดสินใจทั้ง 4 เทคนิคเพื่อแยกแยะเสียงโฆษะหรือเสียงอโฆษะ ผลการทดลองพบว่าเทคนิคที่ 1 เป็นเทคนิคมีประสิทธิภาพที่ดีในการหาคุณภาพการแยกแยะเสียงโฆษะและเสียงอโฆษะที่ 89.29% ที่จำนวนเสียงโฆษะเท่ากับ 19,356 เฟรม

2.9.2 ในงานวิจัย [2] นำเสนอในหัวข้อ “Time-Compression of Speech in Information Talks Using Spectral Entropy” ได้นำเสนอวิธีแปลกใหม่สำหรับการบีบอัดข้อมูลทางเวลาสำหรับเสียงพูดบนพื้นฐานการวัดข้อมูลของเอนโทรปี บทความได้แสดงให้เห็นประยุกต์ใช้ของเอนโทรปีสเปกตรัมของคุณสมบัติในการระบุกลุ่มที่สำคัญของเสียงพูด ด้วยการแยกประมวลผลที่ละส่วน จากนั้นจึงทำการคำนวณผลเพื่อใช้ค่าตัดสินใจเพื่อระบุกลุ่มที่สำคัญของเสียงพูดซึ่งช่วยให้การบีบอัดข้อมูลทางเวลาสำหรับเสียงพูดมีประสิทธิภาพ จากหลักการสามารถนำไปพัฒนาและประยุกต์ใช้การบีบอัดข้อมูลทางเวลาสำหรับเสียงพูดบนฐานเวลาจริงโดยการปรับค่าตัดสินใจ

2.9.3 ในงานวิจัย [3] นำเสนอในหัวข้อ “การเปรียบเทียบประสิทธิภาพการรับส่งข้อมูลร่วมกับการบีบอัดสัญญาณเสียงโดยการแปลงเวฟเล็ต” ได้ทำการเปรียบเทียบประสิทธิภาพการของสัญญาณเสียงพูดที่มีการบีบอัดทั้งสัญญาณกับการส่งเฉพาะเสียงพูดที่บีบอัดเท่านั้น เนื่องจากสัญญาณเสียงพูดมีลักษณะเป็นกลุ่มข้อมูลเสียงและกลุ่มข้อมูลที่ไม่เป็นเสียง ทำการส่งที่บีบอัดเฉพาะกลุ่มข้อมูลที่เป็นข้อมูลเสียงเท่านั้น เพื่อทำให้ลดขนาดของข้อมูลในการส่งข้อมูลลง การส่งสัญญาณเฉพาะสัญญาณเสียงพูดมีอัตราการบีบอัดที่ 4.64 เท่าของสัญญาณเสียงพูดต้นฉบับ โดยที่มีค่า SNR = 18.62 dB และค่า PSNR = 32.7 dB ซึ่งใกล้เคียงกับสัญญาณเสียงพูดต้นฉบับ

2.9.4 ในงานวิจัย [4] นำเสนอในหัวข้อ “User Voice Idenfication on FPGA” ได้ศึกษาและพัฒนาอัลกอริทึมระบุเสียงผู้พูดใช้หน่วยประมวลผลสัญญาณดิจิตอลด้วยความก้าวหน้าด้านเทคโนโลยีซิลิกอนของเอฟพีจีเอ ซึ่งอัลกอริทึมนั้นพัฒนาบนพื้นฐานการใช้งานซอฟต์แวร์ โดยจะปรับปรุงสัญญาณเสียงและใช้ประมวลผลโดยใช้หน้าต่างแบบฮัมมิ่ง ก่อให้เกิดการสกัดลักษณะด้วยเทคนิค LPC ซึ่งตัวอัลกอริทึมจะดำเนินการบนพื้นฐานของเอฟพีจีเอ

2.9.5 ในงานวิจัย [5] นำเสนอในหัวข้อ “A Voice Activity Detection System Based on FPGA” ได้พัฒนาระบบการตรวจหากิจกรรมเสียงโดยใช้หน่วยประมวลผลเอฟพีจีเอ ระดับของเสียงและอัตราส่วนสัญญาณต่อสัญญาณรบกวนควอนไทล์เป็นพารามิเตอร์ของอัลกอริทึมระบบนี้ ตัว

อัลกอริทึมจะดำเนินการในฮาร์ดแวร์อยู่บนเอฟพีจีเอ Xilinx XC5VLX330 มีชิป AD1836A เป็นตัวควบคุมสัญญาณเสียง ซึ่งระบบได้ใช้หน่วยความจำแบบดูอัลพอร์ตภายในเอฟพีจีเอ เป็นบัฟเฟอร์แบบขนานจะถูกเก็บไว้เพื่อเพิ่มความเร็วในการประมวลผล ระบบมีอัตราสุ่มสัญญาณ 16 กิโลเฮิรตซ์และใช้เสียงขนาด 16 บิต การคำนวณตรวจหากิจกรรมเสียงใช้เวลา 15 มิลลิวินาที จากการวิจัยนี้จะเป็นแนวไปประยุกต์ใช้ในหลายๆ ด้าน เช่น การเข้ารหัสเสียงพูดการรู้จำเสียงพูด เป็นต้น

2.9.6 ในงานวิจัย [6] นำเสนอในหัวข้อ “An FPGA-based Voice Signal Preprocessor for The Real-Time Cross-Correlation” ได้พัฒนาระบบการประมวลผลสัญญาณที่เป็นเสียงสำหรับการครอสคอรรีเรชั่นบนฐานเวลาจริงซึ่งสามารถคำนวณการครอสคอรรีเรชั่นของของสัญญาณเสียงจากจำนวนของไมโครโฟนโดยใช้หน่วยประมวลผลเอฟพีจีเอ แต่ละการครอสคอรรีเรชั่น คำนวณโดยใช้สัญญาณเสียง 200 มิลลิวินาที ด้วยการเลื่อนหน้าต่างทีละ 10 มิลลิวินาที ระบบมีอัตราสุ่มสัญญาณ 48 กิโลเฮิรตซ์และใช้เสียงขนาด 16 บิต ตัวอัลกอริทึมจะดำเนินการในฮาร์ดแวร์อยู่บนเอฟพีจีเอ Xilinx XC4VLX200 มีชิป AD1836A เป็นตัวควบคุมสัญญาณเสียง ซึ่งสามารถแบบฐานเวลาจริง

2.9.7 ในงานวิจัย [7] นำเสนอในหัวข้อ “FPGA Implementation of Real Time Acoustic Noise Suppression by Spectral Subtraction using Dynamic Moving Average Method” ได้พัฒนาระบบโดยใช้หน่วยประมวลผลเอฟพีจีเอของการกำจัดเสียงรบกวนบนฐานเวลาจริง โดยการลดเชิงสเปคตรัลใช้วิธีค่าเฉลี่ยเคลื่อนที่แบบไดนามิก การเพิ่มตัวเลือกที่ไม่ซ้ำกันของค่าเฉลี่ยเคลื่อนที่แบบไดนามิก การระยะปรับประมาณการและรับมือกับการเปลี่ยนแปลงในระดับเสียง อัตราส่วนสัญญาณต่อสัญญาณรบกวนถูกทดสอบที่สภาวะแวดล้อมที่แตกต่างกันและใช้การปรับปรุงอัตราส่วนสัญญาณต่อสัญญาณรบกวนเป็นตัวชี้วัดความน่าเชื่อถือ ตัวอัลกอริทึมจะดำเนินการในฮาร์ดแวร์อยู่บนเอฟพีจีเอ Cyclone II EP2C35F672C6 มีชิป WM8731 เป็นตัวควบคุมสัญญาณเสียง จากการวิจัยนี้จะเป็นแนวทางการการนำชิปเอฟพีจีเอไปประยุกต์ใช้ในหลายๆ ด้าน เช่น สตรีมมิ่งสัญญาณเสียงพูดและสามารถนำมาใช้ในเชิงอุตสาหกรรม การเชื่อมแบบระบบบัสและระบบโทรศัพท์มือถือ เป็นต้น

2.9.8 ในงานวิจัย [8] นำเสนอในหัวข้อ “Proposed FPGA Hardware Architecture for High Frame Rate (>>100 fps) Face Detection Using Feature Cascade Classifiers” ได้ออกแบบเอฟพีจีเอสำหรับการตรวจจับใบหน้าที่อัตราเฟรมสูงมากกว่า 100 เฟรมต่อวินาทีโดยใช้การจัดประเภทตามลำดับชั้น เทคนิคเอดาบูทในวิธีฝึกฝนการตรวจจับใบหน้าซึ่งให้ผลรวดเร็วกว่าวิธีอื่นโดยคัดแยกด้วยการประมวลผลสองชุดแบบอนุกรม ระบบสามารถคำนวณแต่ละหน้าต่างย่อยในหนึ่งสัญญาณนาฬิกา โดยเก็บข้อมูลในหน่วยความจำภายในชิปเอฟพีจีเอซึ่งจะลูกจำกัดความสามารถตามสัญญาณ

นาฬิกาบนบอร์ดเอฟพีจีเอ ดังนั้นระบบสามารถทำงานได้บนเอฟพีจีเอ Xilinx Vertex-II Pro ผลการทดลองแสดงให้เห็นว่าอัตราเฟรมบรรลุ 143 เฟรมต่อวินาที สำหรับขนาดของภาพ 640 x 480 พิกเซล

2.9.9 ในงานวิจัย [9] นำเสนอในหัวข้อ “External Memory System Optimization for FPGA Based Implementation of Speech Signal Processing” ได้ทำการศึกษาการดำเนินการเพิ่มประสิทธิภาพระบบหน่วยความจำภายนอกสำหรับเอฟพีจีเอบนพื้นฐานของการประมวลผลสัญญาณเสียง ซึ่งในปัจจุบันเอฟพีจีเอเป็นหนึ่งในทางเลือกที่มีการนำมาประยุกต์ใช้ในระบบการประมวลผลบนพื้นฐานเวลาจริงสัญญาณ เพราะว่าในประมวลผลบนพื้นฐานเวลาจริงสัญญาณนั้นต้องการความเร็วสูง และความแม่นยำสูง จำเป็นต้องใช้หน่วยความจำสำหรับเก็บข้อมูลและการคำนวณ ผลการทดลองบนหน่วยความจำภายนอกชนิด SDRAM ขนาด 128 เมกกะไบต์บนแผงวงจรชนิด 2 เลเยอร์พบว่าสามารถทำงานได้บนเอฟพีจี Xilinx โมดูล

จากงานวิจัยที่กล่าวมาข้างต้นการประมวลผลของสัญญาณเสียงเป็นการประยุกต์การประมวลผลสัญญาณทางด้านดิจิตอลเข้ากับสัญญาณเสียง โดยในงานวิจัยที่ [1-3] นั้นได้กล่าวถึงการประมวลสัญญาณเสียงแต่พื้นฐานการออกแบบและการทำงานจะอยู่บนซอฟต์แวร์ เนื่องจากผู้วิจัยได้มีแนวคิดที่พัฒนาการประมวลผลสัญญาณเสียงจากพื้นฐานการออกแบบและการทำงานบนซอฟต์แวร์ไปสู่ฮาร์ดแวร์และให้สามารถประมวลผลบนฐานเวลาจริง ในงานวิจัยที่ [4-7] ปัจจุบันเอฟพีจีเอนิยมนำมาประยุกต์ใช้ในงานวิจัยด้านการประมวลสัญญาณเสียงโดยใช้บอร์ดเอฟพีจีเอซึ่งได้ทำการวิจัยและประยุกต์ใช้ในหลายๆด้าน เช่น การบีบอัดสัญญาณเสียง การรู้จำสัญญาณเสียง เป็นต้น ทั้งนี้ในสัญญาณเสียงมีลักษณะเป็นกลุ่มข้อมูลจะประกอบด้วยสัญญาณที่เป็นเสียงและสัญญาณที่ไม่ใช่เสียงพูดหรือเสียงรบกวน ในการวิเคราะห์สัญญาณเสียงพูดหลากหลายเทคนิค เช่น วิธีการใช้ค่าแอมปลิจูด วิธีการใช้ค่าพลังงาน เป็นต้น แต่พบว่าวิธีดังกล่าวอาจจะมีข้อจำกัดเมื่อทำการประมวลผลบนพื้นฐานเวลาจริง เช่น ต้องรู้จุดเริ่มและจุดสิ้นสุดของทั้งสัญญาณ เป็นต้น จึงต้องศึกษาหลักการที่เหมาะสม ในงานวิจัยที่ [8-9] นั้นได้มีการกล่าวถึงการจัดเก็บข้อมูลในหน่วยความจำ ซึ่งเป็นหัวใจในการประมวลผลสัญญาณบนพื้นฐานเวลาจริง ในการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดบนฐานเวลาจริงนั้นเทคนิคที่นำมาวิเคราะห์สัญญาณต้องลดข้อจำกัดการประมวลผลบนพื้นฐานเวลาจริงให้น้อยที่สุดเพื่อให้ออกแบบและการทำงานบนซอฟต์แวร์ได้

ดังนั้นในงานวิจัยนี้ได้นำเสนอการพัฒนาเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ ซึ่งแตกต่างกับงานวิจัยที่ [4-7] ในส่วนของการจัดการหน่วยความจำภายในให้สามารถรองรับการทำงานบนฐานเวลาจริงและการใช้เทคนิคเส้นตัดสินใจในการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด ได้ศึกษาค่า

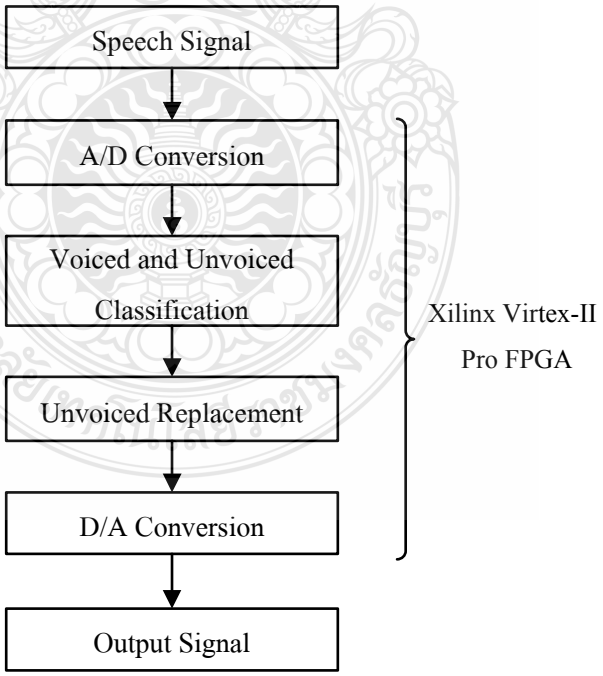
เส้นตัดสินใจที่เหมาะด้วยโปรแกรม MATLAB จากค่า SNR และ NRMSE จากนั้นได้ออกแบบระบบด้วยภาษาวีเฮชดีแอลและสังเคราะห์เป็นวงจรลอจิกติดตั้งไปที่บอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro ซึ่งข้อดีของเทคนิคดังกล่าวสามารถทำการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดบนฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ ใช้เวลาประมวลผลเพียง 25 มิลลิวินาที โดยที่สัญญาณเสียงนั้นมีขนาด 18 บิต บนระบบที่มีอัตราสุ่มสัญญาณเสียง 8 กิโลเฮิรตซ์

บทที่ 3
วิธีการดำเนินงานวิจัย

เนื้อหาในส่วนนี้นำเสนอการออกแบบระบบแยกแยะสัญญาณที่เป็นเสียงและสัญญาณรบกวนหรือสัญญาณเสียงที่ไม่ใช่เสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอซึ่งขั้นตอนต่างๆ นั้นจะประกอบด้วย ขั้นตอนการทดลองการประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอ เทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด การคัดเลือกค่าเส้นการตัดสิน การออกแบบและการทำงานบนเอฟพีจีเอ

3.1 ขั้นตอนการทดลอง

จากข้อมูลที่ศึกษาบอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro ที่ใช้ในการวิจัยครั้งนี้ได้ถูกออกแบบด้วยภาษาวีเฮชดีแอล ออกแบบระบบแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยการประมวลผลสัญญาณแบบฐานเวลาจริงและสังเคราะห์เป็นวงจรลอจิกไปที่ชิปเอฟพีจีเอ ซึ่งขั้นตอนการทำงานของระบบดังกล่าวจะแสดงในภาพที่ 3.1



ภาพที่ 3.1 ขั้นตอนการทำงานของงานวิจัย

3.2 การประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอ

การประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอมีขั้นตอนการทำงานของโปรแกรมจะแสดงในภาพที่ 3.2 สามารถแบ่งเป็นขั้นตอนได้ดังนี้

1) รับข้อมูลเสียงภายนอกเข้ามา โดยได้ออกแบบให้สามารถเลือกข้อมูลเสียงได้จากไมโครโฟนและคอมพิวเตอร์

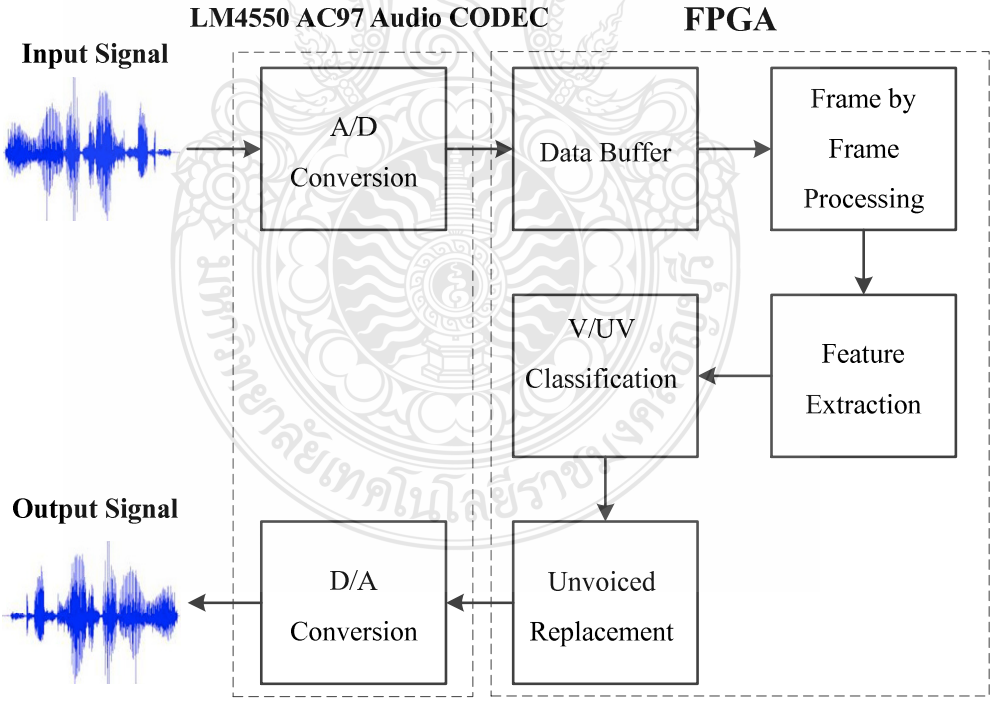
2) บันทึกข้อมูลเสียงลงหน่วยความจำถูกเก็บตามแอดเดรสของหน่วยความจำ กำหนดด้วยขนาดของหน้าต่าง (Window)

3) สกัดคุณลักษณะเด่นของสัญญาณ

4) ประมวลผลเพื่อแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดจากข้อมูลเฟรมเสียงการสกัดคุณลักษณะเด่นของสัญญาณ

5) ส่งข้อมูลเฉพาะส่วนสัญญาณเสียง ในส่วนสัญญาณเสียงที่ไม่ใช่เสียงพูดจะถูกแทนที่ด้วยข้อมูลจากการคำนวณของเอฟพีจีเอไปยังระบบ

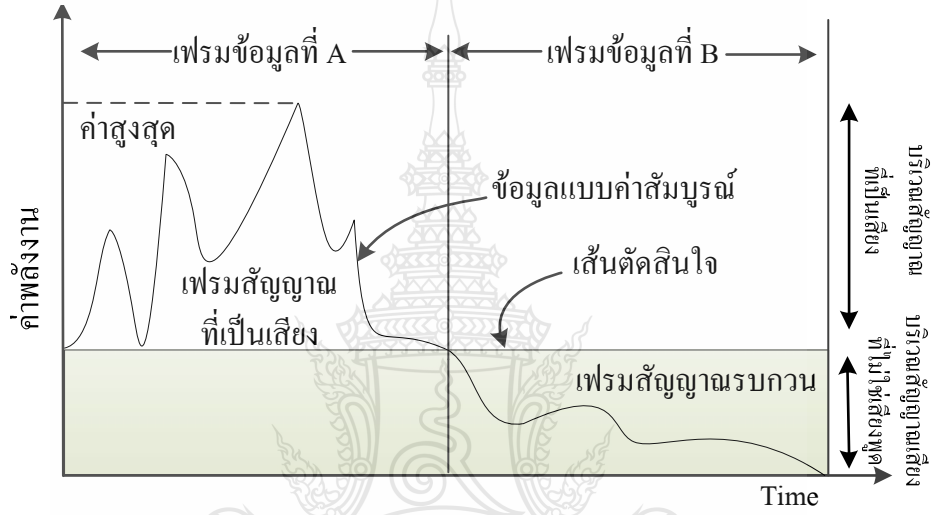
6) คืนกลับข้อมูลสัญญาณเสียงไปยังลำโพง



ภาพที่ 3.2 ไดอะแกรมของระบบแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด

3.3 เทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด

การแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดโดยใช้บอร์ดเอฟพีจีเอ โดยพัฒนาเทคนิคจากข้อดีของเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด [10] โดยพบว่าเทคนิคที่ดีที่สุดคือวิธีที่ 5 เพราะว่าที่พบค่าความผิดพลาดน้อยที่สุด สามารถนำไปปรับปรุงและออกแบบด้วยภาษาวีเฮชดีแอลและสังเคราะห์เป็นวงจรลอจิกไปที่ชิปเอฟพีจีเอ ซึ่งมีขั้นตอนการดำเนินการดังนี้



ภาพที่ 3.3 การแยกแยะสัญญาณเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยเส้นตัดสินใจ [11]

ในการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยการใช้ค่าตัดสินใจจะแสดงในภาพที่ 3.3 โดยที่พลังงานสูงสุดมีค่าเท่ากับ 1 เมื่อรับข้อมูลสัญญาณเสียงมายังเส้นการตัดสินใจทำการเปรียบเทียบทุกค่าในเฟรมข้อมูล เพื่อลดขั้นตอนการประมวลผลจึงไม่คำนึงถึงเครื่องหมาย แล้วทำการนับจำนวนค่าที่ผ่านเส้นการตัดสินใจ เพื่อนำไปเปรียบเทียบค่าเปอร์เซ็นต์เพื่อใช้เปรียบเทียบค่าที่ผ่านเส้นการตัดสินใจ โดยกำหนดไว้ที่ค่า 25 เปอร์เซ็นต์ เพื่อแยกแยะเฟรมสัญญาณเสียงและเฟรมสัญญาณเสียงที่ไม่ใช่เสียงพูด ดังสมการที่ 3.1 - 3.4

$$Th = Ratio1 \tag{3.1}$$

$$S = \begin{cases} 1, & Th \le x(n); voiced \\ 0, & Th > x(n); unvoiced \end{cases} \tag{3.2}$$

$$E_S = \sum_{n=1}^{w} S(n) \tag{3.3}$$

$$\mathrm{Re} = \begin{cases} E_s \geq Ratio2 & ;Voiced\ frame \\ E_s < Ratio2 & ;Unvoiced\ frame \end{cases} \tag{3.4}$$

เมื่อ Th คือ ค่าเส้นตัดสินใจ

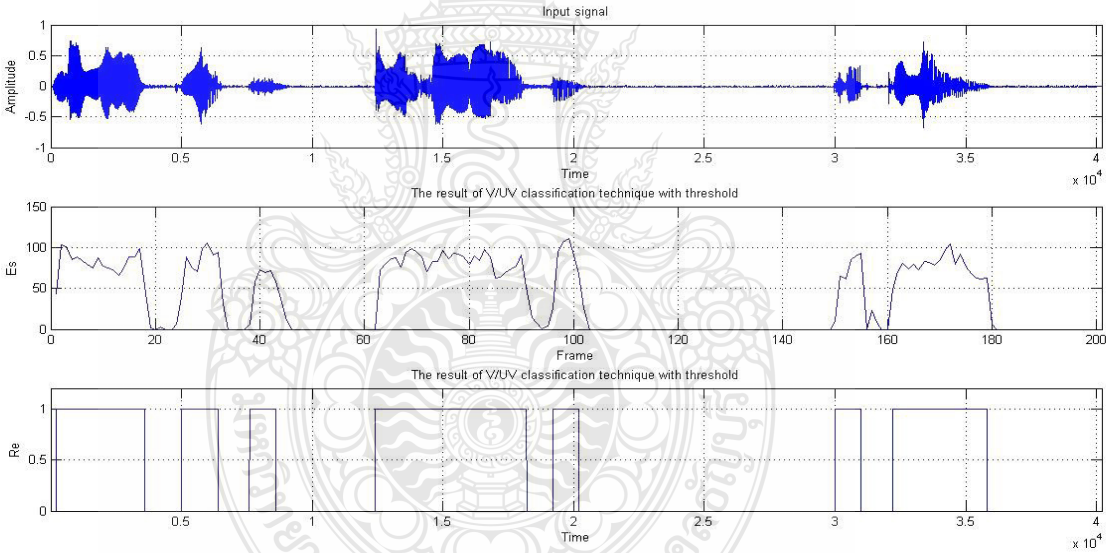
$Ratio1$ คือ ค่าเปอร์เซ็นต์ในการปรับค่าเส้นการตัดสินใจ

S คือ จำนวนค่าที่ผ่านเส้นการตัดสินใจ

E_S คือ ผลรวมจำนวนค่าที่ผ่านเส้นการตัดสินใจ

$Ratio2$ คือ ค่าเปอร์เซ็นต์เพื่อใช้เปรียบเทียบค่าที่ผ่านเส้นการตัดสินใจในเฟรม

Re คือ ผลการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด



ภาพที่ 3.4 ผลการแยกแยะสัญญาณเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยเส้นตัดสินใจ

3.4 การคัดเลือกค่าเส้นการตัดสินใจ

กระบวนการการคัดเลือกค่าเส้นการตัดสินใจ เพื่อนำไปใช้ในการแยกแยะสัญญาณเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด เพื่อเลือกค่าเส้นตัดสินใจที่เหมาะสมที่สุดสำหรับงานวิจัยนี้ โดยค่าตัดสินใจจะถูกเลือกใช้ในการแยกแยะสัญญาณเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยการใช้เส้นการตัดสินใจ มีขั้นตอนการดำเนินการดังนี้

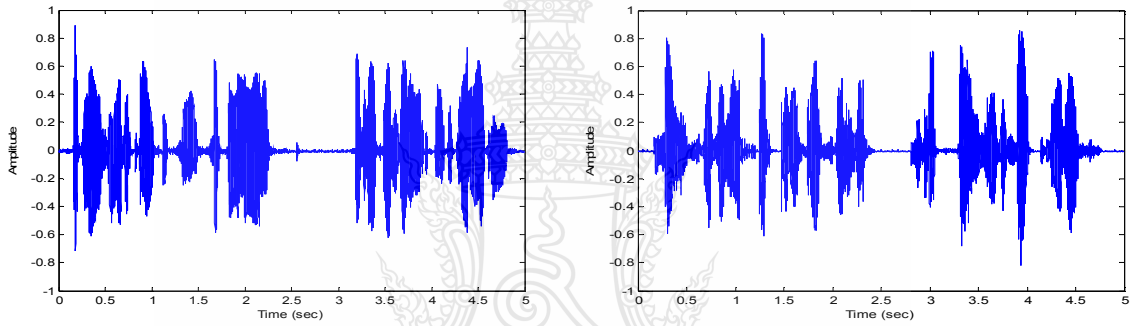
3.4.1 เตรียมสัญญาณเสียงพูด

สัญญาณเสียงพูดที่ใช้สำหรับงานวิจัยนี้มีทั้งหมด 80 เสียง ซึ่งเป็นสัญญาณเสียงภาษาอังกฤษและแต่ละสัญญาณเสียงเป็นประโยคคำพูดที่แตกต่างกัน โดยสามารถแบ่งกลุ่มตามเวลาและแยกเป็นสัญญาณเสียงของผู้ชายและผู้หญิงได้ดังนี้

1) เสียงผู้ชายที่เวลา 5 วินาที และ 60 วินาที อย่างละ 20 เสียง

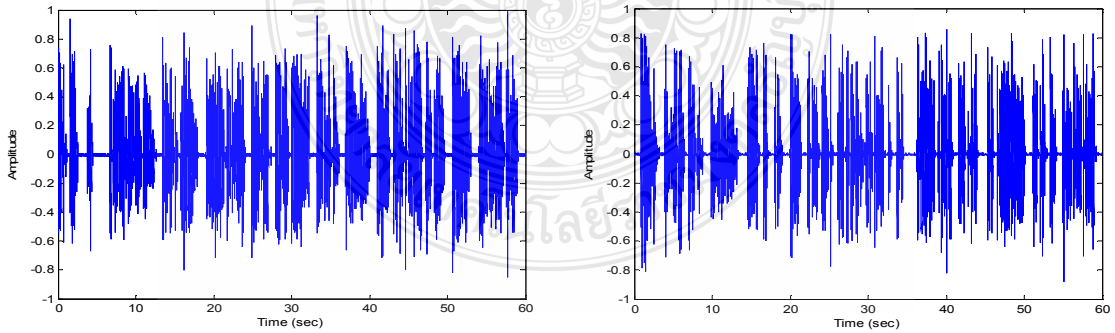
2) เสียงผู้หญิงที่เวลา 5 วินาที และ 60 วินาที อย่างละ 20 เสียง

3) สัญญาณเสียงที่นำมาทดลอง มีอัตราการสุ่มตัวอย่าง (Sampling Rate) อยู่ที่ 8 กิโลเฮิรตซ์ โดยแต่ละข้อมูลมีขนาด 8 บิต ใช้ช่องสัญญาณเดียว (Mono) โดยรูปแบบจะเป็นมาตรฐานของระบบพีซีเอ็ม (Pulse-Code Modulation: PCM)



(ก) เสียงพูดของผู้หญิง (ข) เสียงพูดของผู้ชาย

ภาพที่ 3.5 ตัวอย่างสัญญาณเสียงพูด ที่เวลา 5 วินาที



(ก) เสียงพูดของผู้หญิง (ข) เสียงพูดของผู้ชาย

ภาพที่ 3.6 ตัวอย่างสัญญาณเสียงพูด ที่เวลา 60 วินาที

3.4.2 การวัดค่าผิดพลาด

งานวิจัยนี้เลือกใช้เทคนิค NRMSE ในการหาความค่าผิดพลาดระหว่างสัญญาณเสียงก่อนบีบอัดสัญญาณเสียงกับสัญญาณเสียงที่ผ่านการบีบอัดแล้ว ซึ่งเทคนิค NRMSE นั้นเป็นวิธีที่ใช้ในการคำนวณหาประสิทธิภาพของการแทนค่าข้อมูลที่ขาดหายไป โดยรายละเอียดของสมการได้ถูกกล่าวไว้ในบทที่ 2 สำหรับการวัดค่าความผิดพลาดนั้น ถ้าค่าของ NRMSE ที่มีค่าต่ำจะบ่งบอกถึงคุณภาพของการบีบอัดที่ดี

3.4.3 การวัดค่าประสิทธิภาพของสัญญาณ

สำหรับการหาค่าประสิทธิภาพของสัญญาณเสียงก่อนบีบอัดและสัญญาณเสียงที่ผ่านการบีบอัดแล้ว โดยงานวิจัยนี้ใช้เทคนิคในการหาค่าประสิทธิภาพ คือเทคนิค SNR นั้นเป็นเทคนิคในการพิจารณาหาค่าอัตราส่วนสัญญาณเสียงต้นฉบับต่อสัญญาณรบกวนมีหน่วยเป็นเดซิเบล (dB) โดยในรายละเอียดของสมการได้ถูกกล่าวไว้ในบทที่ 2 สำหรับการวัดค่าประสิทธิภาพของสัญญาณเสียง ถ้าค่าของ SNR มีค่าเดซิเบลที่สูง จะเป็นการบ่งบอกคุณภาพในการบีบอัดที่มีประสิทธิภาพที่ดี

3.5 การออกแบบและการทำงานบนเอฟพีจีเอ

งานวิจัยนี้ได้เลือกใช้บอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro เพราะว่าสามารถเป็นได้ทั้งบอร์ดทดลองและบอร์ดพัฒนาระบบจะแสดงในภาพที่ 3.7 นั้นมีชิป XC2VP30 เป็นหน่วยประมวลผลกลาง โดย XC2VP30 มีเกททั้งหมด 30,816 ลอจิกเซลล์ ถูกนำมาใช้ในการวิจัยครั้งนี้ มีอุปกรณ์อำนวยความสะดวกด้วยอุปกรณ์ทางอินพุตและเอาต์พุต เช่น มีชิป LM4550 AC97 Audio CODEC เป็นตัวควบคุมสัญญาณเสียง สามารถต่อรวมกับหน่วยความจำภายนอกได้ เป็นต้น เพื่อให้ผู้ทดลองได้เรียนรู้การออกแบบวงจรดิจิตอลตั้งแต่วงจรขั้นพื้นฐานจนไปถึงขั้นนำไปพัฒนาออกแบบสร้างวงจรขนาดใหญ่ได้ด้วยตัวเอง ออกแบบระบบแยกแยะสัญญาณเสียงพูดด้วยการประมวลผลสัญญาณแบบฐานเวลาจริงด้วยภาษาวีเฮชดีแอลและสังเคราะห์เป็นวงจรลอจิกไปที่ชิปเอฟพีจีเอต่อไป

3.4.1 คุณสมบัติทั่วไปของบอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro XC2VP30

บอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro XC2VP30 จะประกอบไปด้วยอุปกรณ์ต่างๆ จะแสดงในภาพที่ 3.5 ยกตัวอย่าง เช่น

1) Virtex-II Pro XC2VP30 มีทรัพยากร 30,816 ลอจิกเซลล์ 136 18-bit Multipliers 2,448 กิโลไบต์ of Block RAM และ 2 PowerPC Processors

2) DDR SDRAM จำนวน 1 ช่องและสามารถรองรับหน่วยความจำได้สูงสุด 2 Gbytes

3) 10/100 Ethernet จำนวน 1 พอร์ท

4) USB2 จำนวน 1 พอร์ท

5) Compact Flash Card จำนวน 1 ช่อง

6) XSGA Video จำนวน 1 พอร์ท

7) LM4550 AC97 Audio CODEC

8) SATA 3Gbps จำนวน 3 ช่อง

9) PS/2 จำนวน 1 พอร์ท

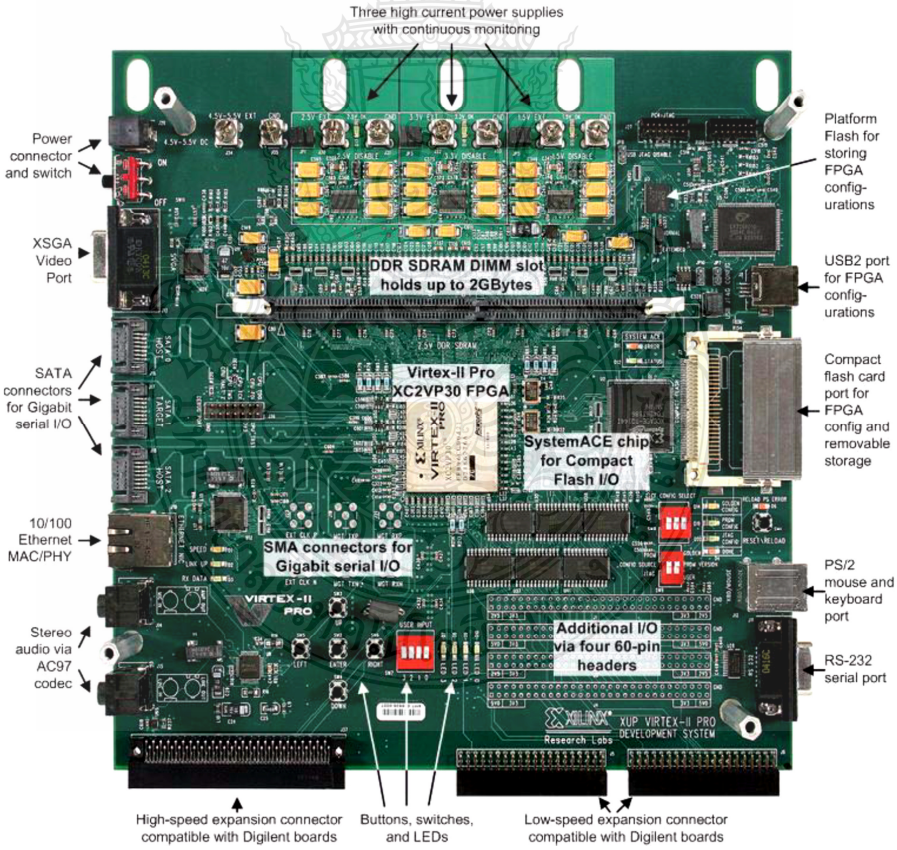
10)RS-232 จำนวน 1 พอร์ท

11)Push Button Switch จำนวน 5 ตัว

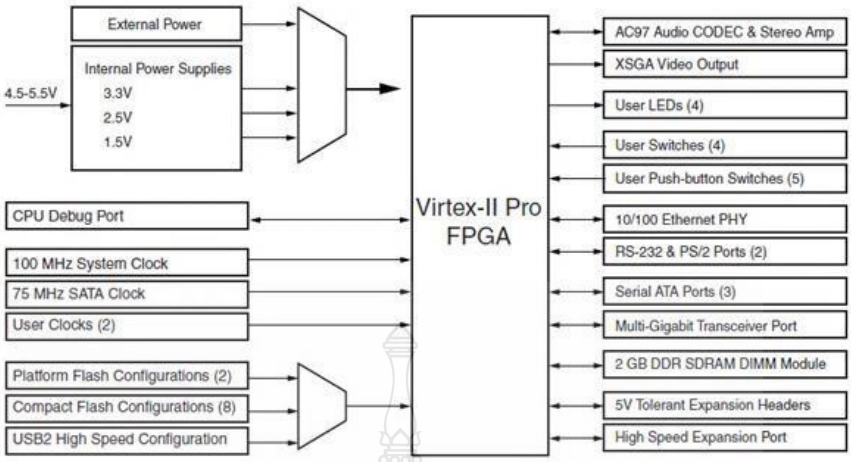
12)DIP Switch 4 บิต

13)คอนเนคเตอร์ ชนิด High Speed และ ชนิด Low Speed จำนวน 1 พอร์ท

14)100 MHz Oscillator



ภาพที่ 3.7 บอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro XC2VP30



ภาพที่ 3.8 ไดอะแกรมของการเชื่อมต่อกับบอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro XC2VP30

3.4.2 ระบบเสียงในบอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro XC2VP30

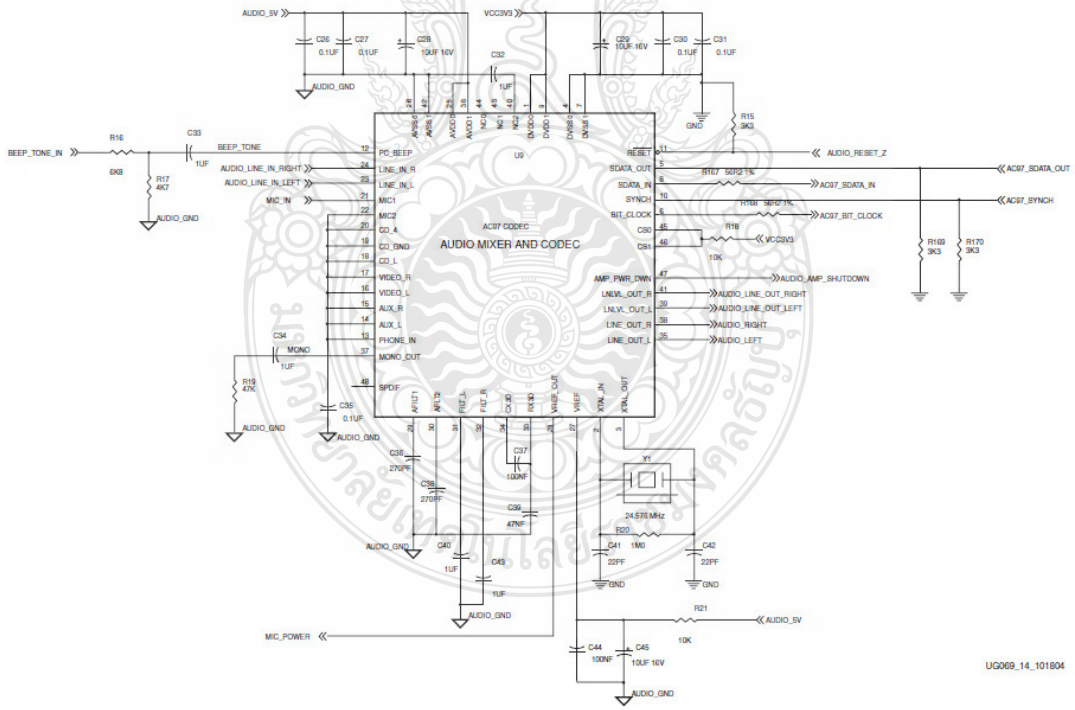
บอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro XC2VP30 มีชิป LM4550 AC97 Audio CODEC เป็นตัวควบคุมสัญญาณเสียงและชิปขยายกำลังเสียงขนาด 150 มิลลิวัตต์ ในส่วนบล็อกไดอะแกรมการจัดวางอุปกรณ์มีรายละเอียดการจัดวางจะแสดงดังภาพที่ 3.9 และ 3.10 โดยมีรายละเอียดของขาอุปกรณ์ที่ต่ออยู่กับขาเอฟพีจีจะแสดงดังตารางที่ 3.1 และ 3.2

ตารางที่ 3.1 รายละเอียดของชิป LM4550 AC97 Audio CODEC ที่ต่ออยู่กับขาเอฟพีจี

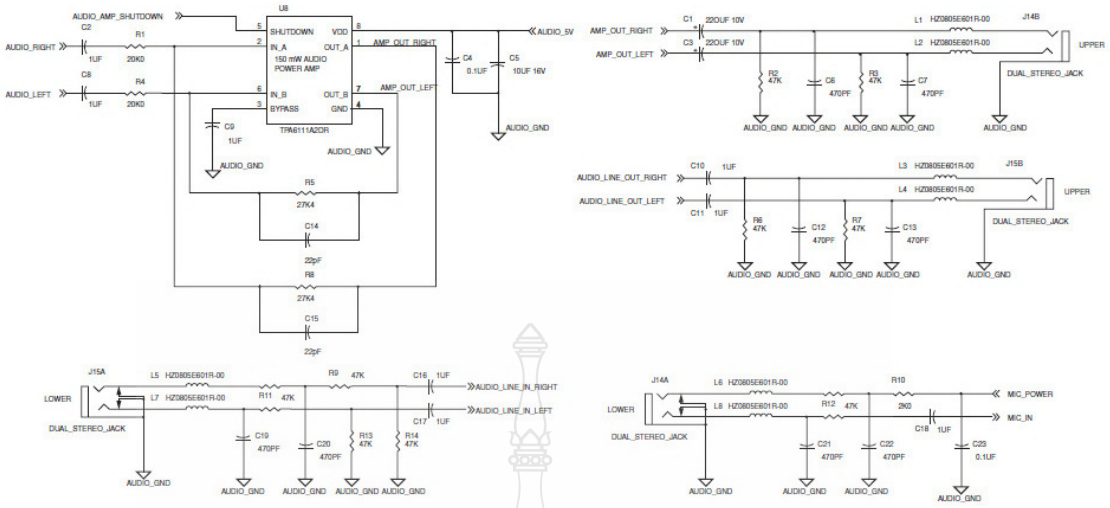
Signal	Direction	FPGA Pin	I/O Type	Drive	Slew
AC97_SDATA_OUT	O	E8	LVTTL	8 mA	SLOW
AC97_SDATA_IN	I	E9	LVTTL	-	-
AC97_SYNCH	O	F7	LVTTL	8 mA	SLOW
AC97_BIT_CLOCK	I	F8	LVTTL	-	-
AUDIO_RESET_Z	O	E6	LVTTL	8 mA	SLOW
BEEP_TONE_IN	O	E7	LVTTL	8 mA	SLOW

ตารางที่ 3.2 รายละเอียดของ Push Button และ DIP Switch ที่ต่ออยู่กับขาเอฟพีจี

Signal	Direction	FPGA Pin	I/O Type	Drive	Slew
SW_0	I	AC11	LVCMOS25	-	-
SW_1	I	AD11	LVCMOS25	-	-
SW_2	I	AF8	LVCMOS25	-	-
SW_3	I	AF9	LVCMOS25	-	-
PB_ENTER	I	AG5	LVTTL	-	-
PB_UP	I	AH4	LVTTL	-	-
PB_DOWN	I	AG3	LVTTL	-	-
PB_LEFT	I	AH1	LVTTL	-	-
PB_RIGHT	I	AH2	LVTTL	-	-



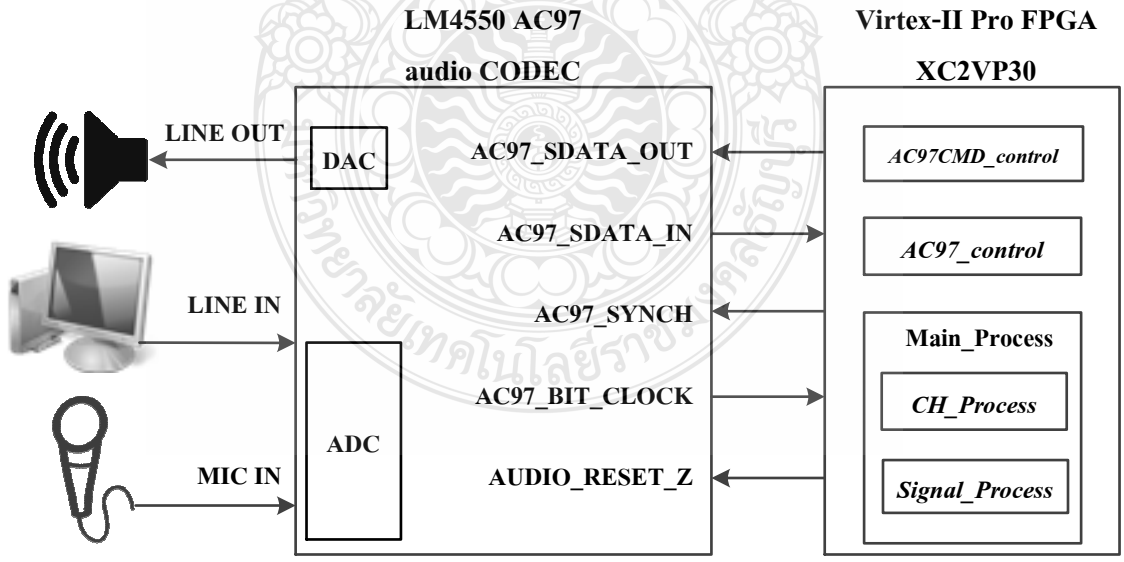
ภาพที่ 3.9 บล็อกไดอะแกรมการจัดวางอุปกรณ์ชิป LM4550 AC97 Audio CODEC



ภาพที่ 3.10 บล็อกไดอะแกรมการจัดวางอุปกรณ์ชิปขยายกำลังเสียง 150 mW

3.4.3 การออกแบบและการอินเตอร์เฟส LM4550 AC97 Audio CODEC กับเอฟพีจีเอ

การออกแบบและการทำงานบนเอฟพีจีเอในส่วนการอินเตอร์เฟส LM4550 AC97 Audio CODEC กับเอฟพีจีเอ จะแสดงในภาพที่ 3.11



ภาพที่ 3.11 การอินเตอร์เฟส LM4550 AC97 Audio CODEC กับเอฟพีจีเอ

ตารางที่ 3.3 คำอธิบายหน้าที่ของรีจิสเตอร์

รีจิสเตอร์	คำอธิบาย
Reset Register (00h)	ทำการเปลี่ยนแปลงค่ารีจิสเตอร์ทั้งหมดให้เป็นค่าเริ่มต้น
Master Volume Register (02h)	เปลี่ยนระดับเสียงหรือปิดเสียงของ LINE_OUT ในช่วง 0 dB ถึง 46.5dB ที่ค่าระดับละ 1.5 dB
Headphone Volume Register (04h)	เปลี่ยนระดับเสียงหรือปิดเสียงของ HP_OUT ในช่วง 0 dB ถึง 46.5dB ที่ค่าระดับละ 1.5 dB
Mic Volume Register (0Eh)	ปรับระดับเสียงไมโครโฟน
Record Select Register (1Ah)	เลือกแหล่งที่มาของเสียง
Record Gain Register (1Ch)	ควบคุมอัตราการขยายสัญญาณ ในช่วง 0dB-22.5dB ที่ค่าระดับละ 1.5 dB
Sample Rate Control Registers (2Ch,32h)	ตั้งค่าอัตราการสุ่มข้อมูลของสัญญาณ ต่ำสุดที่ 4KHz และสูงสุดที่ 48KHz

ตารางที่ 3.4 การเชื่อมต่อของ AC97 Audio CODEC

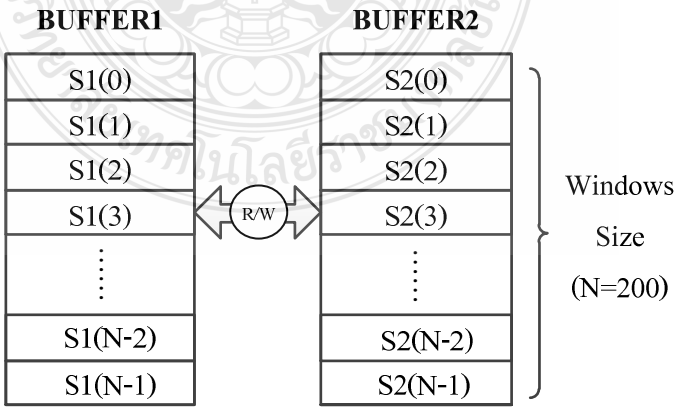
สัญญาณ	คำอธิบาย
AC97_SDATA_OUT	รับข้อมูลจากตัวควบคุมไปที่ AC97 Audio CODEC โดยมีทั้งเฟรมข้อมูลเสียง PCM และเฟรมข้อมูลที่ใช้ควบคุมการทำงานของ DACs.
AC97_SDATA_IN	ส่งข้อมูลจาก AC97 Audio CODEC ไปยังตัวควบคุม โดยมีทั้งเฟรมสถานะและเฟรมข้อมูลเสียง PCM จาก ADCs.
AC97_SYNCH	เป็นสัญญาณที่ใช้กำหนดขอบเขต AC Link Frames โดยอ้างอิงขอบขาขึ้นของสัญญาณ
AC97_BIT_CLOCK	รับสัญญาณนาฬิกา AC Link ความถี่ 12.288MHz
AUDIO_RESET_Z	ควบคุมรีจิสเตอร์ของระบบ ภายในเงื่อนไขค่าเริ่มต้น

เริ่มต้นรับข้อมูลเสียงจากภายนอกเข้ามา โดยได้ออกแบบให้สามารถเลือกข้อมูลเสียงได้จากไมโครโฟนและคอมพิวเตอร์และเริ่มต้นระบบใหม่ได้ สัญญาณจะส่งมายังผ่านการแปลงสัญญาณอนาล็อกเป็นสัญญาณดิจิตอลขนาด 18 บิต ผ่านทางขาสัญญาณ AC97_SDATA_IN ส่งต่อไปยังหน่วย

ประมวลผล เพื่อทำหน้าที่บันทึกเสียงลงหน่วยความจำที่ละเฟรมเสียงข้อมูล ประมวลผลเพื่อแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด แล้วส่งข้อมูลดิจิตอลจากตัวควบคุมผ่านทางขาสัญญาณ AC97_SDATA_OUT จะทำการคืนกลับเป็นสัญญาณอนาล็อกไปยังลำโพงเพื่อแปลงสัญญาณไฟฟ้าให้เป็นเสียง โดยมีสัญญาณจาก AC97_SYNCH และ AC97_BIT_CLOCK เพื่อช่วยควบคุมจังหวะการทำงานของชิป LM4550 AC97 Audio CODEC ร่วมกับโปรแกรมออกแบบการควบบนบอร์ดเอฟพีจีเอ คำอธิบายหน้าที่ของรีจิสเตอร์จะแสดงดังตารางที่ 3.3 และการเชื่อมต่อของอุปกรณ์จะแสดงดังตารางที่ 3.4

3.4.4 การออกแบบและการทำงานบนเอฟพีจีเอ

การออกแบบและการทำงานบนเอฟพีจีเอในส่วนการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด จากเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดจะออกแบบระบบแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยการประมวลผลสัญญาณแบบฐานเวลาจริง จะนำสัญญาณเสียงแอนะล็อกมาจัดเก็บในรูปแบบที่เป็นดิจิตอลเพื่อการนำไปการประมวลผลสัญญาณเสียงโดยกรรมวิธีทางดิจิตอล เราจำเป็นต้องมีการสุ่มเอาค่าจากสัญญาณแอนะล็อกนำมาจัดเก็บเป็นค่าดิจิตอล [12] จะได้สัญญาณที่เรียกว่าสัญญาณไม่ต่อเนื่องทางเวลา (Discrete-Time Signal) [13] โดยส่วนคุณภาพของสัญญาณไม่ต่อเนื่องทางเวลากำหนดด้วยอัตราการสุ่มสัญญาณ แต่การจะประมวลผลสัญญาณแบบฐานเวลาจริงนั้น จะถูกแบ่งออกเป็นเฟรมตามความเหมาะสมกับงานที่ใช้การประมวลผลการแบ่งสัญญาณเสียงออกเป็นเฟรมย่อย สัญญาณเสียงเสียงจะถูกเก็บเป็นเฟรม ซึ่งขนาดหน้าต่างที่เหมาะสมในการทดลองเลือกที่ 200 ตำแหน่ง เพื่อที่จะสร้างพื้นที่ใช้จัดเก็บบัฟเฟอร์ข้อมูล 2 เฟรมจะแสดงดังภาพที่ 3.12



ภาพที่ 3.12 การเก็บข้อมูลสัญญาณเสียงออกเป็นเฟรมย่อยด้วยหน่วยความจำ

โดยเฟรมที่หนึ่งจะใช้ประมวลผลในส่วนการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดและเฟรมที่สองจะใช้ในการคืนกลับสัญญาณเสียง โดยจะเมื่อเฟรมที่หนึ่งประมวลผลเสร็จแล้วเฟรมที่สองจะใช้ประมวลผลในส่วนการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดและเฟรมที่หนึ่งจะใช้ในการคืนกลับสัญญาณเสียงต่อไป ซึ่งจะช่วยให้เอฟพีจีเอสามารถประมวลผลสัญญาณเสียงแบบฐานเวลาจริงเพื่อการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดได้ [14]

การออกแบบขั้นตอนการทำงานของระบบการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณรบกวนแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ จะทำให้ทราบแนวทางการออกแบบโปรแกรมบนเอฟพีจีเอให้กระทำตามขั้นตอนได้ทุกขั้นตอนและมีประสิทธิภาพและลดขั้นตอนการผิดพลาดได้ ซึ่งขั้นตอนการทำงานจะแสดงในภาพที่ 3.13 จะมีขั้นตอนดังนี้

1) ตรวจสอบสถานะของการทำงานด้วยสัญญาณ Reset เมื่อสัญญาณ Reset มีค่าเท่ากับ 0 ระบบจะทำการคืนค่ากลับไปสู่ค่าเริ่มต้น แต่เมื่อสัญญาณ Reset นั้นมีค่าเท่ากับ 1 จะทำงานในขั้นตอนต่อไป

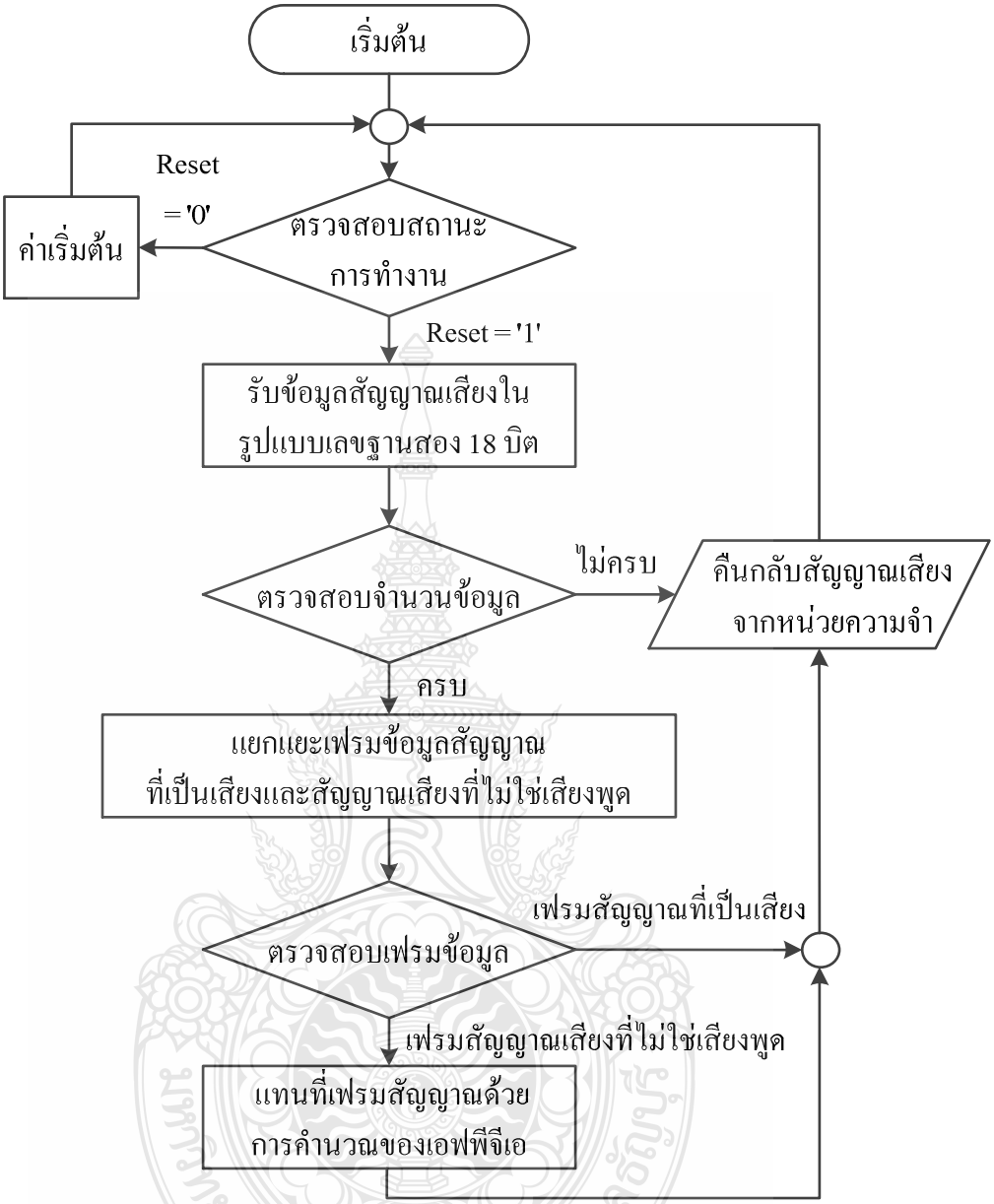
2) รับข้อมูลสัญญาณเสียงเป็นข้อมูลดิจิตอลในรูปแบบเลขฐานสอง 18 บิต ระบบสามารถเลือกรับข้อมูลเสียงได้จากไมโครโฟนและคอมพิวเตอร์

3) ตรวจสอบจำนวนข้อมูลภายในหน่วยความจำ ถ้าจำนวนข้อมูลไม่ครบตามขนาดหน้าต่างที่ 200 ตำแหน่ง จะคืนกลับสัญญาณเสียงจากหน่วยความจำก่อนที่จะกลับไปขั้นตอนเริ่มต้น แต่ถ้าจำนวนข้อมูลครบตามขนาดหน้าต่างที่ 200 ตำแหน่งจะทำงานในขั้นตอนต่อไป

4) ประมวลผลเพื่อแยกแยะเฟรมข้อมูลสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด

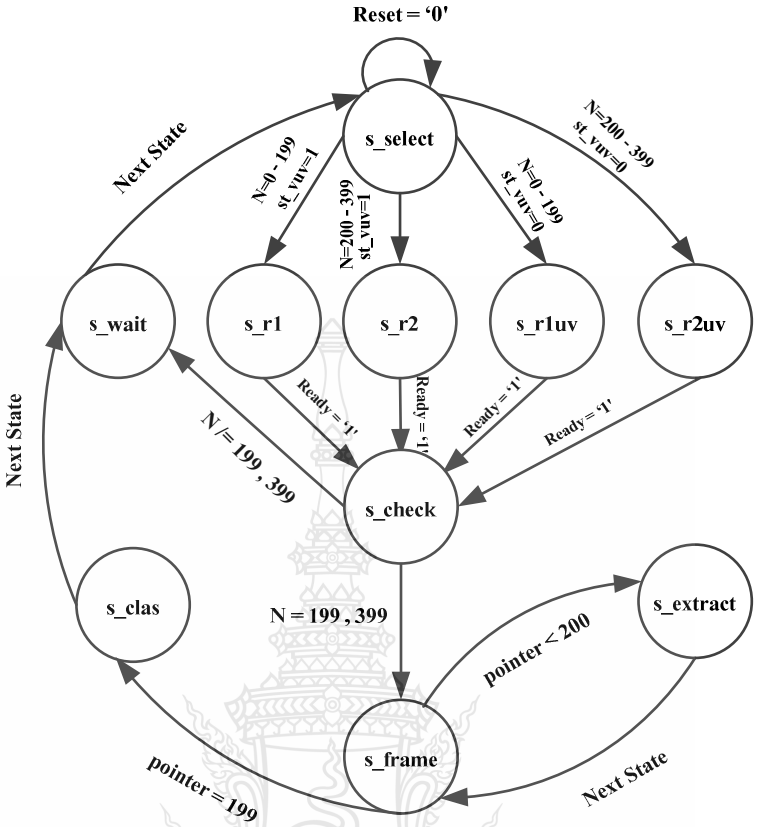
5) ตรวจสอบเฟรมข้อมูล ถ้าเป็นเฟรมข้อมูลสัญญาณที่เป็นเสียงจะคืนกลับสัญญาณเสียงจากหน่วยความจำก่อนที่จะกลับไปขั้นตอนเริ่มต้น แต่ถ้าเป็นเฟรมสัญญาณเสียงที่ไม่ใช่เสียงพูดจะทำงานในขั้นตอนต่อไป

6) แทนที่เฟรมสัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยการคำนวณของเอฟพีจีเอ ก่อนที่จะคืนกลับสัญญาณเสียงจากหน่วยความจำก่อนที่จะกลับไปขั้นตอนเริ่มต้น



ภาพที่ 3.13 ผังงานการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด

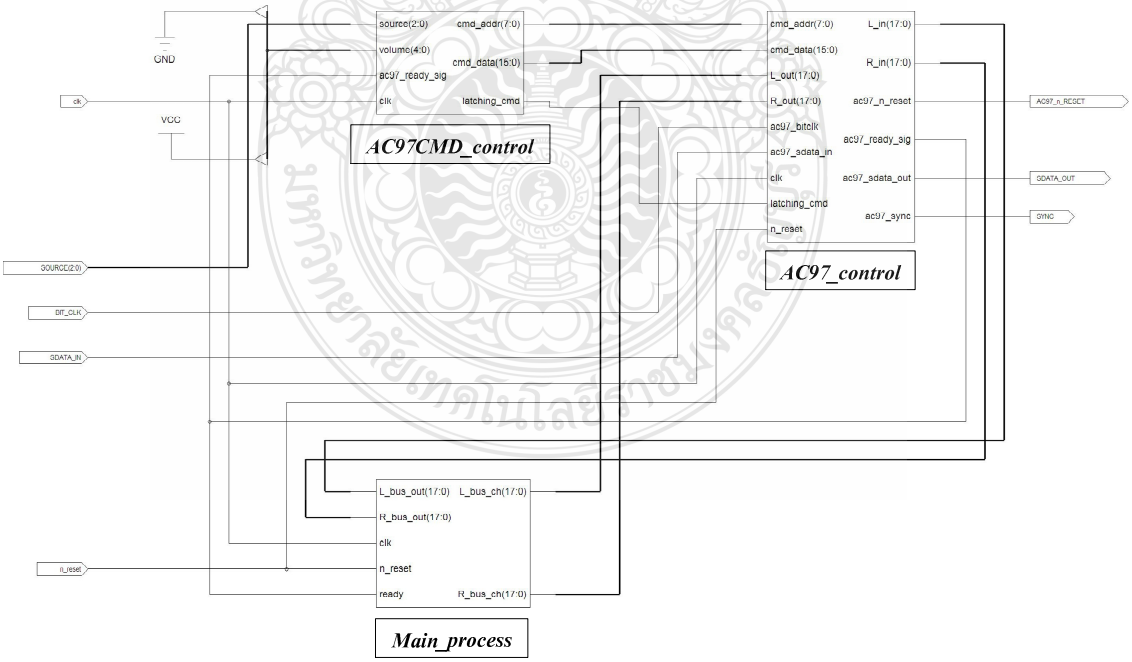
การออกแบบหน่วยควบคุมระบบการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดบนฐานเวลาจริงโดยหน่วยประมวลผลเอฟพีจีเอได้บรรยายด้วยแผนภาพไดอะแกรม ซึ่งขั้นตอนการทำงานของระบบดังกล่าวจะแสดงในภาพที่ 3.14



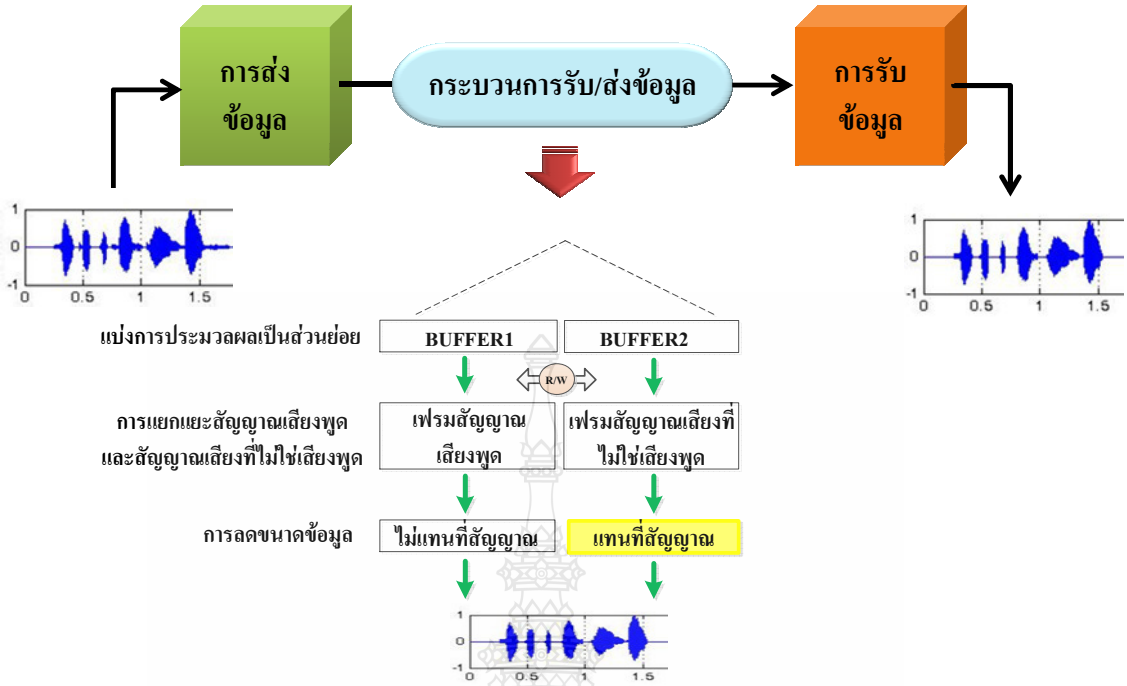
ภาพที่ 3.14 แผนภาพไดอะแกรมขั้นตอนการแยกแยะสัญญาณเสียงพูด

โดยเริ่มที่สถานะ s_select ถ้าสัญญาณ Reset มีค่าเท่ากับ 0 ค่าตัวแปรจะเข้าสู่สภาวะเริ่มต้นและมีสถานะเป็น s_select แต่ถ้า Reset มีค่าเท่ากับ 1 สถานะจะเปลี่ยนไปเป็นสถานะอื่นขึ้นอยู่กับสัญญาณ st_vuv และ N เป็นตัวกำหนดการทำงานของโปรแกรมและหน่วยความจำ โดยได้ออกแบบการทำงานไว้ 4 กรณี ในกรณีที่ 1 เมื่อสัญญาณ st_vuv มีค่าเท่ากับ 1 และ N มีค่าช่วง 0 ถึง 199 สถานะจะเปลี่ยนไปเป็น s_r1 ทำบันทึกข้อมูลสัญญาณเสียงในหน่วยความจำตัวที่ 1 และทำการอ่านข้อมูลจากหน่วยความจำตัวที่ 2 ไปยังระบบเพื่อคืนกลับข้อมูลสัญญาณเสียงต่อไป ในกรณีที่ 2 เมื่อสัญญาณ st_vuv มีค่าเท่ากับ 1 และ N มีค่าช่วง 200 ถึง 399 สถานะจะเปลี่ยนไปเป็น s_r2 ทำบันทึกข้อมูลสัญญาณเสียงในหน่วยความจำตัวที่ 2 และทำการคืนกลับข้อมูลสัญญาณเสียงจากหน่วยความจำตัวที่ 1 ไปยังระบบเพื่อคืนกลับข้อมูลสัญญาณเสียงต่อไป กรณีที่ 3 เมื่อสัญญาณ st_vuv มีค่าเท่ากับ 0 และ N มีค่าช่วง 0 ถึง 199 สถานะจะเปลี่ยนไปเป็น s_r1uv ทำบันทึกข้อมูลสัญญาณเสียงในหน่วยความจำตัวที่ 1 และทำการส่งข้อมูลจากการแทนที่ด้วยข้อมูลจากการคำนวณของเอฟพีจีเอไปยังระบบเพื่อคืน

กลับข้อมูลสัญญาณเสียงต่อไปและกรณีที่ 4 เมื่อสัญญาณ st_vuv มีค่าเท่ากับ 0 และ N มีค่าช่วง 200 ถึง 399 สถานะจะเปลี่ยนไปเป็น s_r1uv ทำบันทึกข้อมูลสัญญาณเสียงในหน่วยความจำตัวที่ 1 และทำการส่งข้อมูลจากการแทนที่ด้วยข้อมูลจากการคำนวณของเอฟพีจีเอไปยังระบบเพื่อคืนกลับข้อมูลสัญญาณเสียงต่อไป เมื่อสัญญาณ Ready มีค่าเท่ากับ 1 ค่าสถานะจะเปลี่ยนไปเป็น s_check ถ้า N มีค่าไม่เท่ากับ 199 และ 399 ค่าสถานะจะเปลี่ยนไปเป็น s_wait จากนั้นค่าสถานะจะเปลี่ยนไปเป็นสถานะเริ่มต้นที่ s_select แต่ถ้า N นั้นมีค่าเท่ากับ 199 และ 399 ค่าสถานะจะเปลี่ยนไปเป็น s_frams ซึ่งการประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอเสียงจะถูกเก็บเป็นเฟรมซึ่งขนาดของหน้าต่าง มีค่าอยู่ในช่วงที่เหมาะสมที่ใช้ในการทดลอง โดยเลือกที่ 200 จะกำหนดจำนวน pointer ของหน่วยความจำ เมื่อเข้าสู่สถานะ s_frams ค่าสถานะจะเปลี่ยนไปเป็นสถานะ s_extract ทำการสกัดคุณลักษณะเด่นของสัญญาณแล้วเก็บข้อมูลเฟรมเสียงนั้นๆ เมื่อ pointer นั้นมีค่าไม่เท่ากับ 199 สถานะจะกลับไปที่ s_frams แต่ถ้า pointer มีค่าเท่ากับ 199 สถานะจะเปลี่ยนไปเป็น s_clas เพื่อแยกแยะเสียงโฆษะและเสียงอโฆษะโดยอาศัยข้อมูลของเฟรมเสียงที่ได้จากการสกัดคุณลักษณะเด่นของสัญญาณ เสร็จแล้วสถานะจะเปลี่ยนไปเป็น s_wait จากนั้นค่าสถานะจะเปลี่ยนไปเป็นสถานะเริ่มต้นที่ s_select เสร็จแล้วจะทำการทดสอบการออกแบบบนชิปเอฟพีจีเอก่อนที่จะสังเคราะห์เป็นวงจรลอจิก ซึ่งได้ออกแบบทั้งหมด 3 ส่วน คือ AC97CMD_control AC97_control และ Main_process ดังภาพที่ 3.15

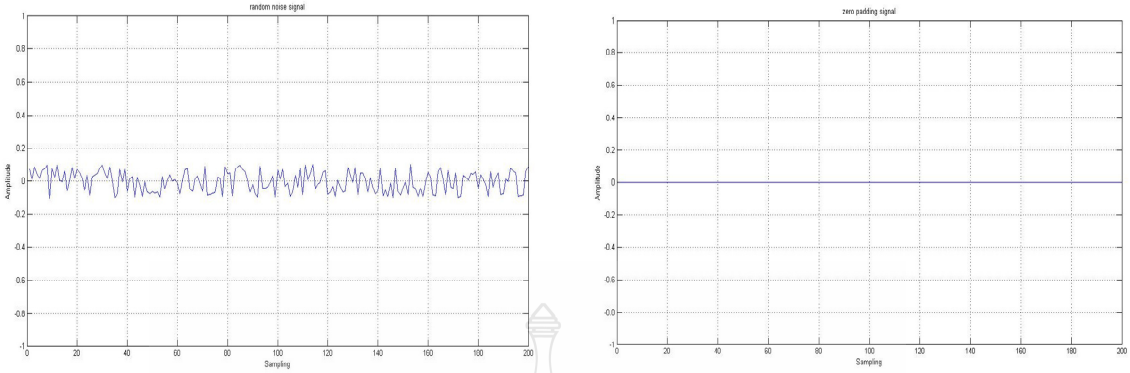


ภาพที่ 3.15 การสังเคราะห์วงจรลอจิกบนบนชิปเอฟพีจีเอจากซอฟแวร์ Xilinx ISE



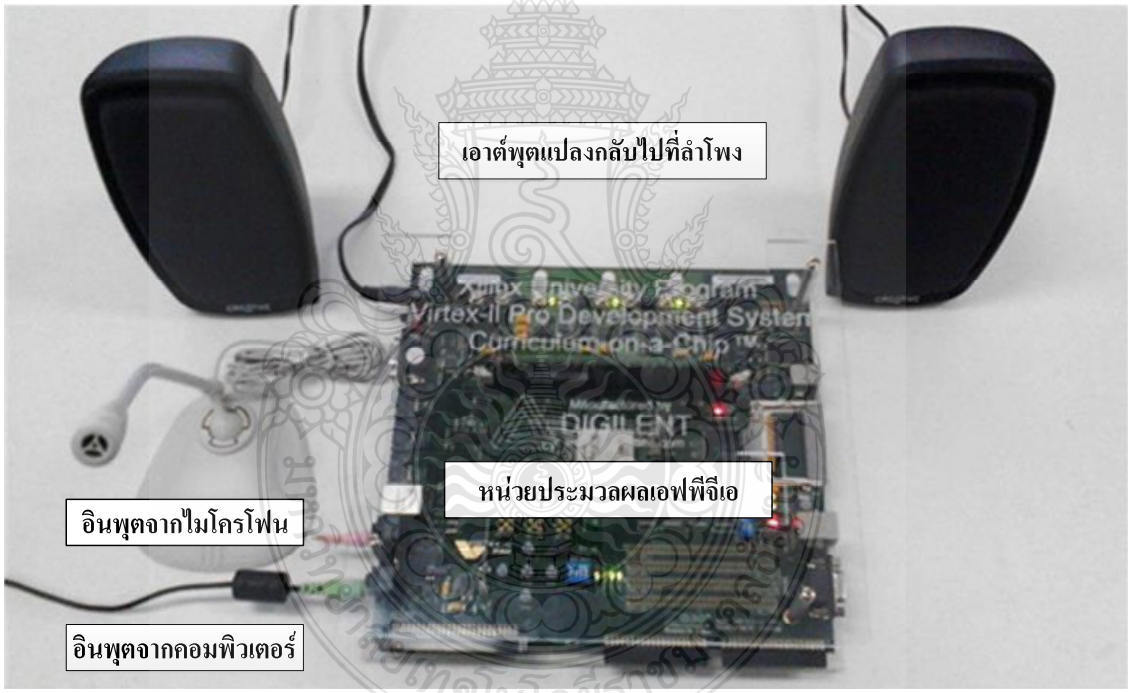
ภาพที่ 3.16 การแยกแยะสัญญาณเสียงพูดในกระบวนการรับส่งข้อมูล

จากภาพที่ 3.16 ได้แสดงกระบวนการรับส่งข้อมูลซึ่งผู้วิจัยได้ใช้ประโยชน์จากการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด [15-16] เนื่องจากสัญญาณเสียงพูดนั้นมีลักษณะเป็นกลุ่มข้อมูลเสียงและกลุ่มข้อมูลที่ไม่เป็นเสียงจะทำการส่งเฉพาะกลุ่มข้อมูลที่เป็นข้อมูลเสียงเท่านั้น เพื่อทำการลดขนาดของข้อมูลในการส่งข้อมูล [17-18] ซึ่งในการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดจากการคำนวณบนเอฟพีจีเอ จากภาพที่ 3.17 ผู้วิจัยได้ศึกษาการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูด 2 รูปแบบ คือ แทนที่ด้วยศูนย์และสัญญาณสุ่ม ดังภาพที่ เมื่อได้ออกแบบและทดสอบระบบดังกล่าวเสร็จแล้วจึงติดตั้งโปรแกรมลงบนบอร์ดเอฟพีจีเอเพื่อทำการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดในกระบวนการส่งข้อมูลแบบฐานเวลาจริง โดยทางด้านอินพุตโดยได้ออกแบบให้รับสัญญาณเสียงจากไมโครโฟนและคอมพิวเตอร์ ทางด้านเอาท์พุตได้ออกแบบให้ส่งสัญญาณเสียงไปยังลำโพง ทั้งด้านอินพุตและเอาต์พุตได้เชื่อมต่ออยู่บนหน่วยประมวลผลเอฟพีจีเอดังภาพที่ 3.18



(ก) สัญญาณสุ่ม (ข) สัญญาณศูนย์

ภาพที่ 3.17 รูปแบบการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดจากการคำนวณบนเอฟพีจีเอ



ภาพที่ 3.18 การทำงานบนบอร์ดเอฟพีจีเอ Virtex-II Pro (XC2VP30)

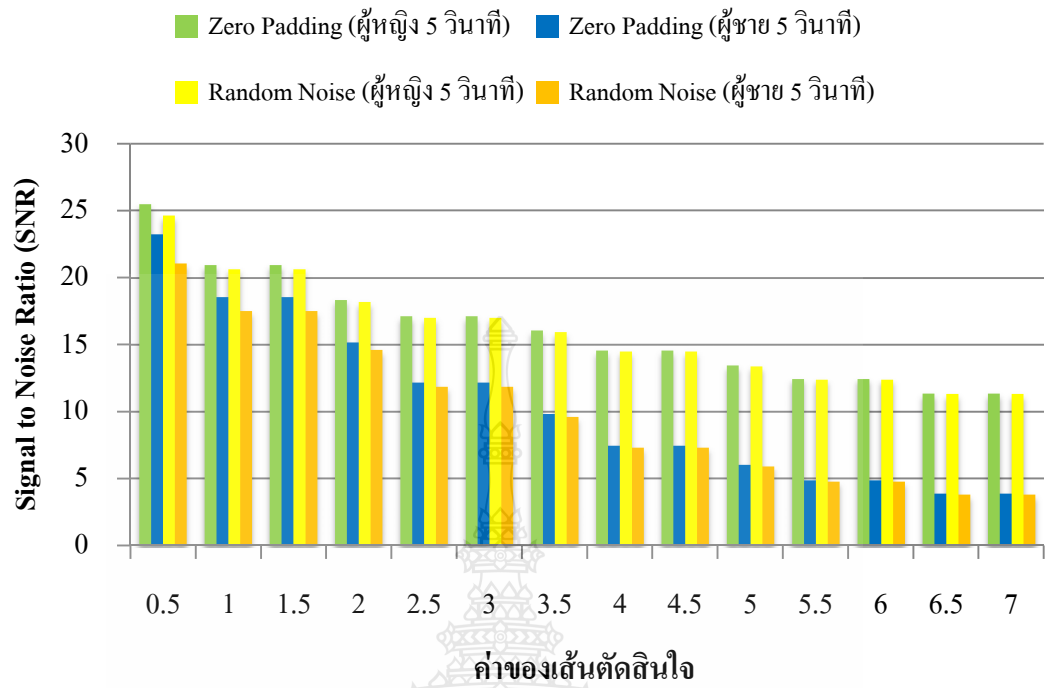
บทที่ 4
ผลการทดลอง

บทนี้เป็นนำเสนอผลการออกแบบระบบแยกแยะสัญญาณที่เป็นเสียงและเสียงรบกวนหรือสัญญาณเสียงที่ไม่ใช่เสียงพูดบนฐานเวลาจริงโดยหน่วยประมวลผลเอฟพีจีเอ ในขั้นตอนแรกจะเป็นการเปรียบเทียบประสิทธิภาพและค่าผิดพลาดของการบีบอัดสัญญาณเสียงพูดเพื่อหาค่าเส้นตัดสินใจที่เหมาะสม ขั้นตอนที่สองจะเป็นผลการ ออกแบบระบบประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอบนฐานเวลาจริงและขั้นตอนสุดท้ายจะนำสัญญาณเสียงพูดผู้ชายและผู้หญิงมาทำการทดสอบ ซึ่งผลการดำเนินการวิจัยมีดังต่อไปนี้

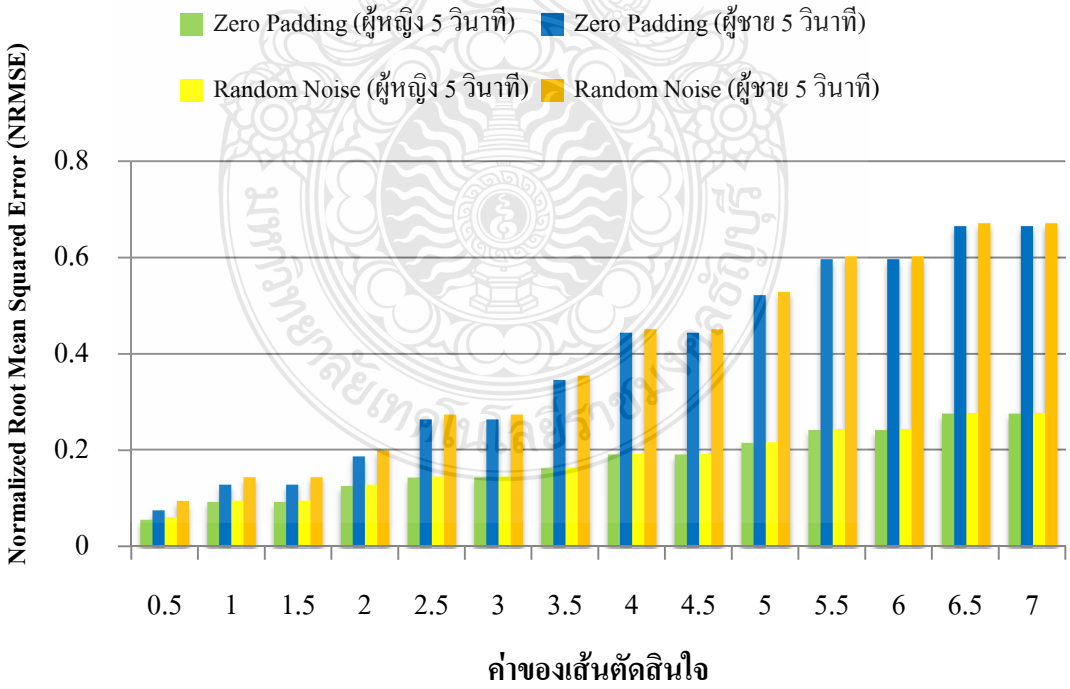
4.1 การเปรียบเทียบประสิทธิภาพและค่าผิดพลาดของการบีบอัดสัญญาณเสียงพูด

ในงานวิจัยนี้ได้ใช้เสียงพูดของผู้หญิงและผู้ชาย มีความยาว 5 วินาที และ 60 วินาที ทั้งหมด 80 สัญญาณ ในการหาค่าเส้นตัดสินใจที่เหมาะสม ในงานวิจัยนำเสนอการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์และการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่ม [19] โดยใช้โปรแกรม MATLAB [20] และนำไปออกแบบระบบเพื่อใช้งานบนฐานเวลาจริงโดยหน่วยประมวลผลเอฟพีจีเอในการวิเคราะห์สัญญาณและเปรียบเทียบประสิทธิภาพของการบีบอัดสัญญาณเสียงพูด ซึ่งผลจะแสดงดังภาพที่ 4.1 - 4.6

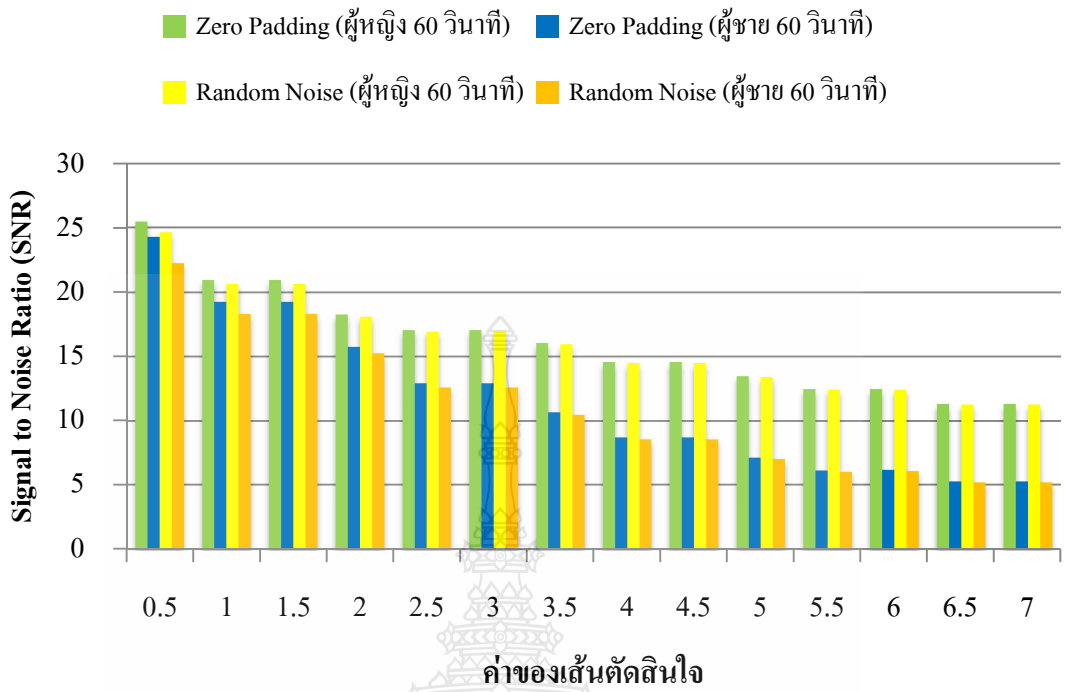
จากภาพที่ 4.1 - 4.4 เป็นการแสดงผลในการวัดค่าประสิทธิภาพของสัญญาณเสียงพูดผู้หญิงและผู้ชาย จากการหาค่า SNR และ ค่า NRMSE ของเสียงพูดที่ผ่านการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดที่ค่าเส้นตัดสินใจที่ 0.5 - 7% กับสัญญาณเสียงพูดผู้หญิงและผู้ชายที่เวลา 5 วินาที และ 60 วินาที จะสังเกตว่าค่า SNR ของเสียงพูดผู้หญิงนั้นมีค่าประสิทธิภาพของการบีบอัดสัญญาณเสียงพูดที่ดีกว่าเสียงพูดผู้ชายและแนวโน้มของค่าประสิทธิภาพของการบีบอัดสัญญาณเสียงพูดจะขึ้นอยู่กับค่าเส้นตัดสินใจ เมื่อค่าเส้นตัดสินใจมีค่าเพิ่มขึ้นพบว่าค่าประสิทธิภาพของการบีบอัดสัญญาณเสียงพูดจะลดลง ในส่วนค่า NRMSE จะสังเกตเห็นว่าของเสียงพูดผู้หญิงนั้นมีค่าความผิดพลาดที่น้อยกว่าเสียงพูดผู้ชายและแนวโน้มของค่าความค่าผิดพลาดของการบีบอัดสัญญาณเสียงพูดจะขึ้นอยู่กับค่าเส้นตัดสินใจ เมื่อค่าเส้นตัดสินใจมีค่าเพิ่มขึ้นความค่าผิดพลาดของการบีบอัดสัญญาณเสียงพูดจะสูงขึ้น ซึ่งเหตุผลเกิดจากการเพิ่มค่าตัดสินใจมากขึ้นทำให้สัญญาณเสียงพูดถูกระบุเป็นส่วนสัญญาณเสียงที่ไม่ใช่เสียงพูดมากขึ้น



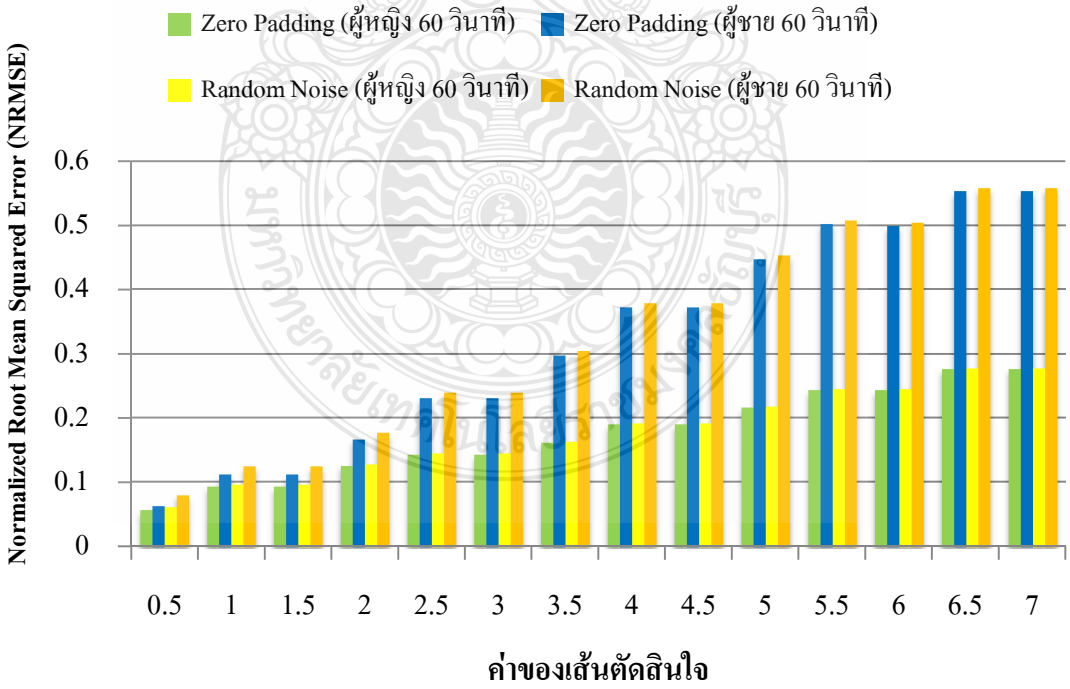
ภาพที่ 4.1 ค่า SNR สัญญาณเสียงพูดผู้หญิงและผู้ชาย 5 วินาที ที่ค่าเส้นตัดสินใจที่ 0.5 - 7%



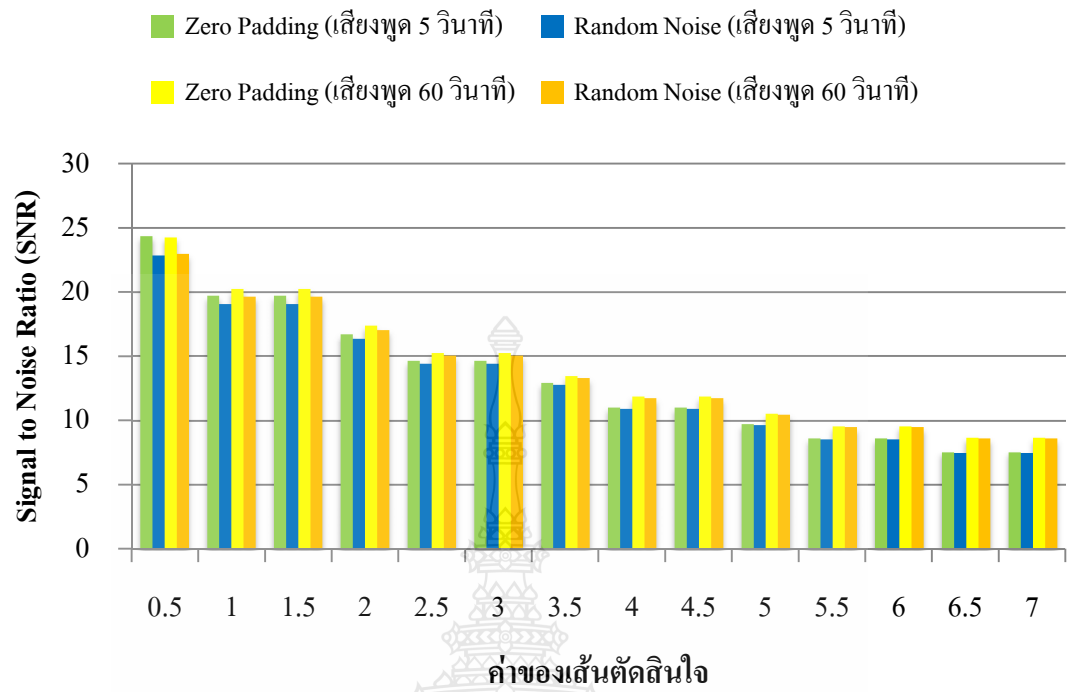
ภาพที่ 4.2 ค่า NRMSE สัญญาณเสียงพูดผู้หญิงและผู้ชาย 5 วินาที ที่ค่าเส้นตัดสินใจที่ 0.5 - 7%



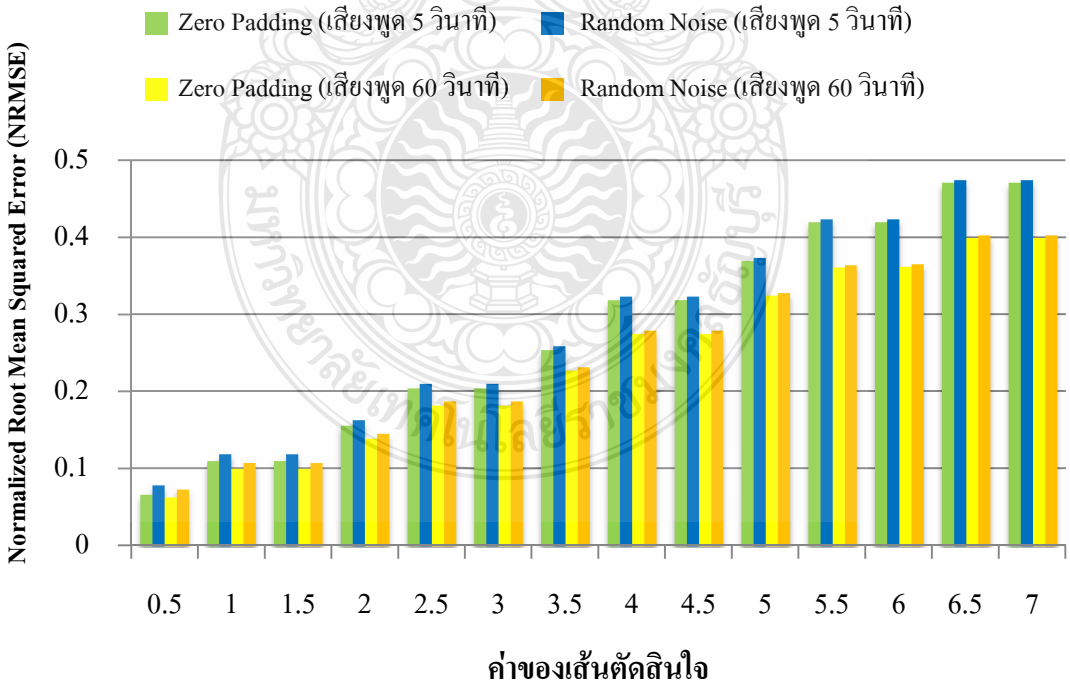
ภาพที่ 4.3 ค่า SNR สัญญาณเสียงพูดและผู้ชาย 60 วินาที ที่ค่าเส้นตัดสินใจที่ 0.5 - 7%



ภาพที่ 4.4 ค่า NRMSE สัญญาณเสียงพูดผู้หญิงและผู้ชาย 60 วินาที ที่ค่าเส้นตัดสินใจที่ 0.5 - 7%

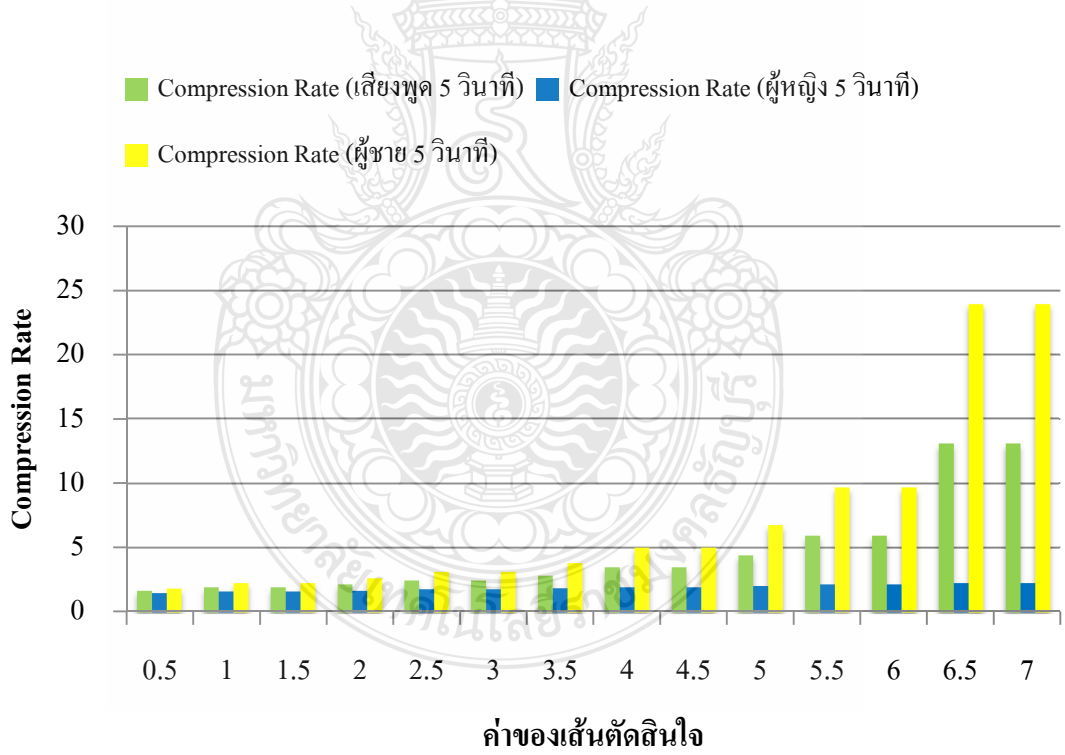


ภาพที่ 4.5 ค่า SNR สัญญาณเสียงพูด 5 และ 60 วินาที ที่ค่าเส้นตัดสินใจที่ 0.5 - 7%

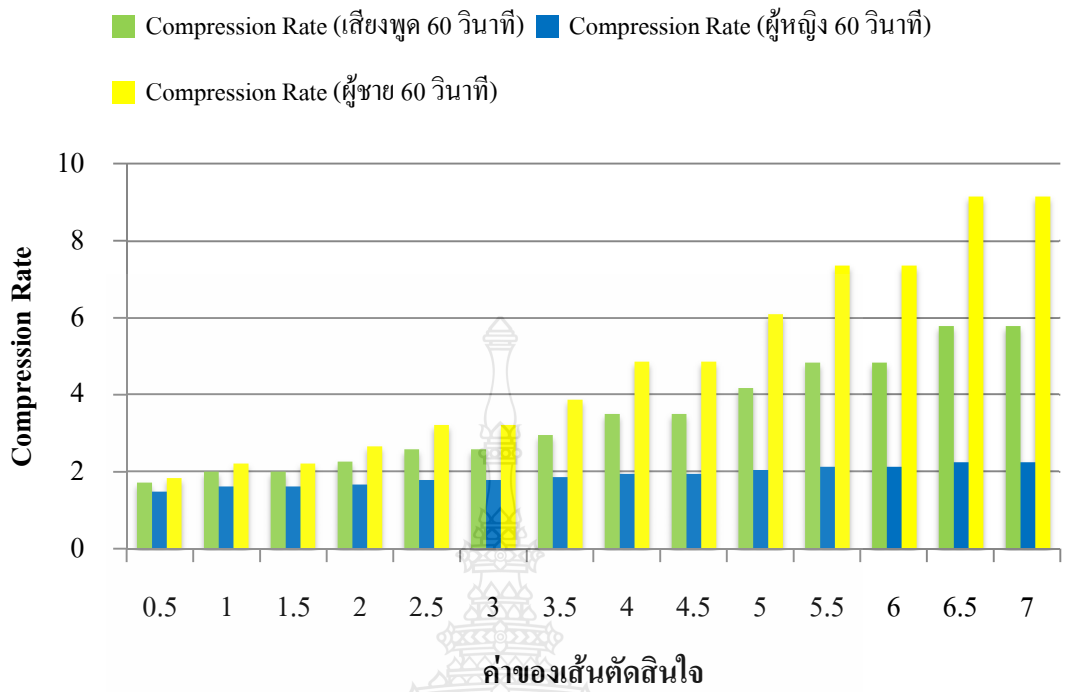


ภาพที่ 4.6 ค่า NRMSE สัญญาณเสียงพูด 5 และ 60 วินาที ที่ค่าเส้นตัดสินใจที่ 0.5 - 7%

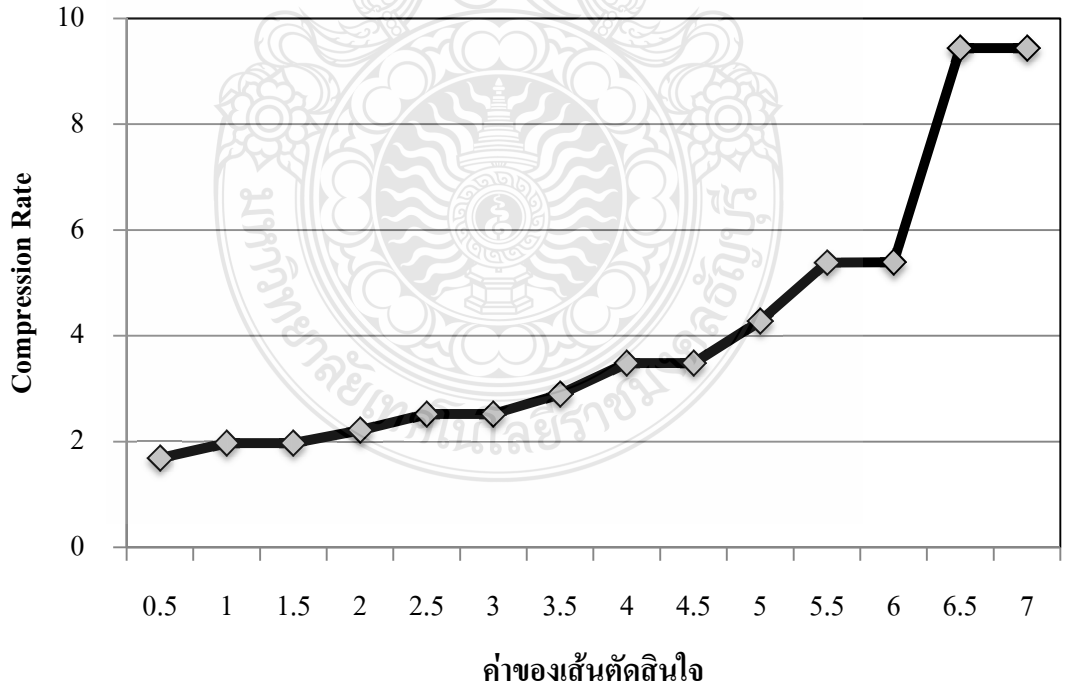
จากภาพที่ 4.5 - 4.6 นั้นเป็นการแสดงผลการวัดค่าประสิทธิภาพของสัญญาณเสียงพูดของแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์และการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่ม ได้จากการหาค่า SNR และ NRMSE ของเสียงพูดที่ผ่านการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดที่ค่าเส้นตัดสินใจที่ 0.5 - 7% กับสัญญาณเสียงพูดที่เวลา 5 วินาที และ 60 วินาที จะสังเกตเห็นได้ว่าค่า SNR ของการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์นั้น จะมีค่าประสิทธิภาพของการบีบอัดสัญญาณเสียงพูดที่ดีกว่าการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่มเล็กน้อยและแนวโน้มของค่าประสิทธิภาพของการบีบอัดสัญญาณเสียงพูดจะขึ้นอยู่กับค่าเส้นตัดสินใจ เมื่อค่าเส้นตัดสินใจมีค่าเพิ่มขึ้นพบว่าค่าประสิทธิภาพของการบีบอัดสัญญาณเสียงพูดจะลดลง ในส่วนของค่า NRMSE จะสังเกตว่าแนวโน้มของค่าความค่าผิดพลาดของการบีบอัดสัญญาณเสียงพูดจะขึ้นอยู่กับค่าเส้นตัดสินใจ เมื่อค่าเส้นตัดสินใจมีค่าเพิ่มขึ้นความค่าผิดพลาดของการบีบอัดสัญญาณเสียงพูดจะสูงขึ้น ซึ่งเหตุผลเกิดจากการเพิ่มค่าตัดสินใจมากขึ้นทำให้สัญญาณเสียงพูดถูกระบุเป็นส่วนสัญญาณเสียงที่ไม่ใช่เสียงพูดมากขึ้น



ภาพที่ 4.7 อัตราบีบอัดข้อมูลสัญญาณเสียงพูด 5 วินาที ที่เส้นตัดสินใจที่ 0.5 - 7%



ภาพที่ 4.8 อัตราบีบอัดข้อมูลสัญญาณเสียงพูด 60 วินาที ที่เส้นตัดสินใจที่ 0.5 - 7%



ภาพที่ 4.9 อัตราบีบอัดข้อมูลสัญญาณเสียงพูดที่เส้นตัดสินใจที่ 0.5 - 7%

จากภาพที่ 4.7 - 4.9 เป็นการแสดงผลการวัดค่าประสิทธิภาพของสัญญาณเสียงพูดจากการหาค่าอัตราบีบอัดข้อมูลเสียงพูดที่ผ่านการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดที่ค่าเส้นตัดสินใจที่ 0.5 - 7% กับสัญญาณเสียงพูด 5 วินาที และ 60 วินาที จะสังเกตเห็นได้ว่าอัตราบีบอัดข้อมูลสัญญาณของเสียงพูดจะขึ้นอยู่กับค่าเส้นตัดสินใจ เมื่อค่าเส้นตัดสินใจมีค่าเพิ่มขึ้นพบว่าอัตราบีบอัดข้อมูลสัญญาณจะสูงขึ้นตามค่าเส้นตัดสินใจ ซึ่งเหตุผลเกิดจากการเพิ่มค่าตัดสินใจมากขึ้นทำให้สัญญาณเสียงพูดถูกระบุเป็นส่วนสัญญาณเสียงที่ไม่ใช่เสียงพูดมากขึ้น จากในงานวิจัยที่ [10] พบว่าค่าเส้นตัดสินใจที่ 1% และการสุ่มสัญญาณที่ 50% นั้นจะมีค่าที่เหมาะสมในการใช้แยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด เมื่อได้ทำการพัฒนาเทคนิคให้เหมาะสมในการประมวลผลแบบฐานเวลาจริงนั้นได้กล่าวไว้ในบทที่ 3 และทำการทดลองจะได้ผลดังในตารางที่ 4.1 และในส่วนผลลัพธ์อื่นๆ นั้นจะแสดงไว้อยู่ในภาคผนวก ก

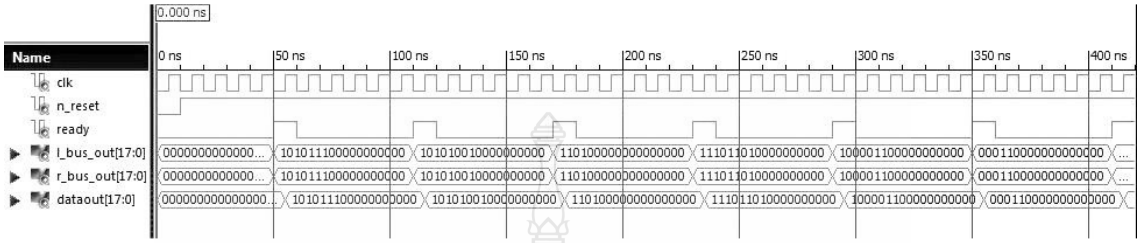
ตารางที่ 4.1 การวัดค่าประสิทธิภาพของสัญญาณเสียงพูดที่ค่าเส้นตัดสินใจที่ 1%

อัตราบีบอัดข้อมูล	SNR		NRMSE	
	การแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์	การแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่ม	การแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์	การแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่ม
1.970	20.024	19.398	0.105	0.113

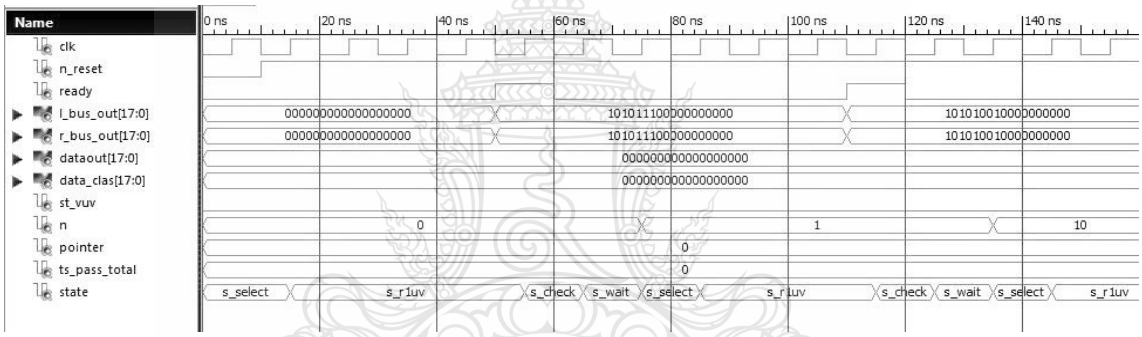
4.2 ผลการออกแบบการทำงานบนชิปเอฟพีจีเอ

จากภาพที่ 4.10 เป็นการแสดงผลการประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอด้วยการจำลองการทำงานลงบนชิปเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเฮชดีแอลด้วยซอฟแวร์ Xilinx ISE โดยที่ระบบจะเป็นการประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอบนฐานเวลาจริงโดยไม่มีการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด จากภาพที่ 3.14 แผนภาพถูกนำมาเปลี่ยนเป็นลอจิกเกท โดยที่ระบบจะเป็นการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดบนชิปเอฟพีจีเอบนฐานเวลาจริงและแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยการคำนวณของเอฟพีจีเอ จากภาพที่ 4.11 เป็นการแสดงผลด้วยการจำลองการทำงานลงบนชิปเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเฮชดีแอลด้วยซอฟแวร์ Xilinx ISE จะเห็นได้ว่าสัญญาณ st_vuv มีค่าเท่ากับ 0 และ N มีค่าช่วง 0 - 199 สถานะจะเปลี่ยนจาก s_select ไปเป็น

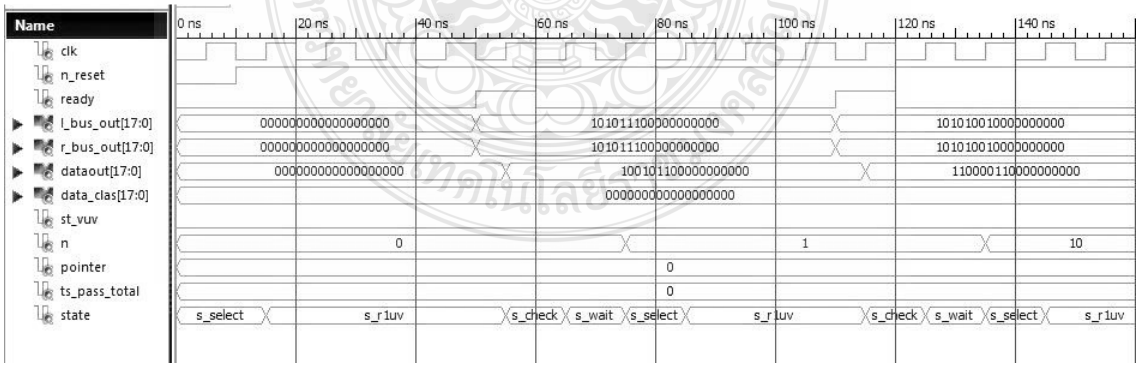
s_r1uv ซึ่งจะทำบันทึกข้อมูลสัญญาณเสียงในหน่วยความจำตัวที่ 1 และจะทำการส่งข้อมูลด้วยสัญญาณศูนย์ไปยังระบบเพื่อคืนกลับข้อมูลสัญญาณเสียง จากนั้นสถานะจะเปลี่ยนไปเป็น s_check, s_wait และ s_select เพื่อการประมวลในครั้งต่อไป



ภาพที่ 4.10 ผลจำลองการประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอจากซอฟแวร์ Xilinx ISE

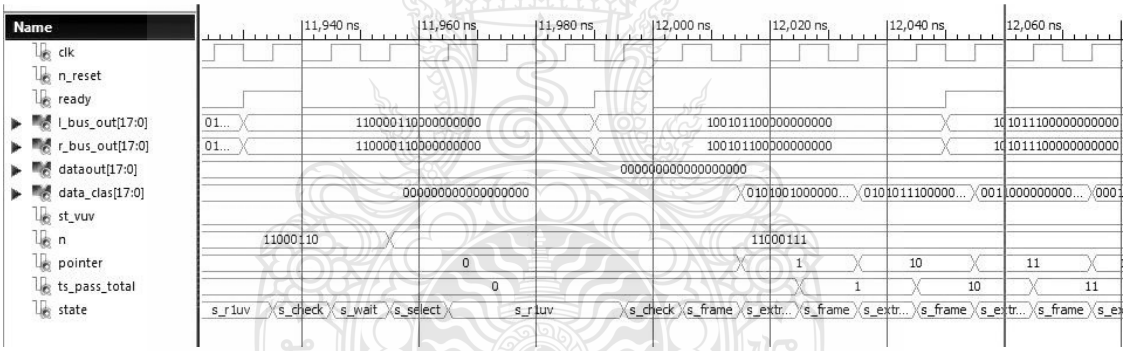


ภาพที่ 4.11 ผลจำลองการส่งข้อมูลด้วยสัญญาณศูนย์บนชิปเอฟพีจีเอ



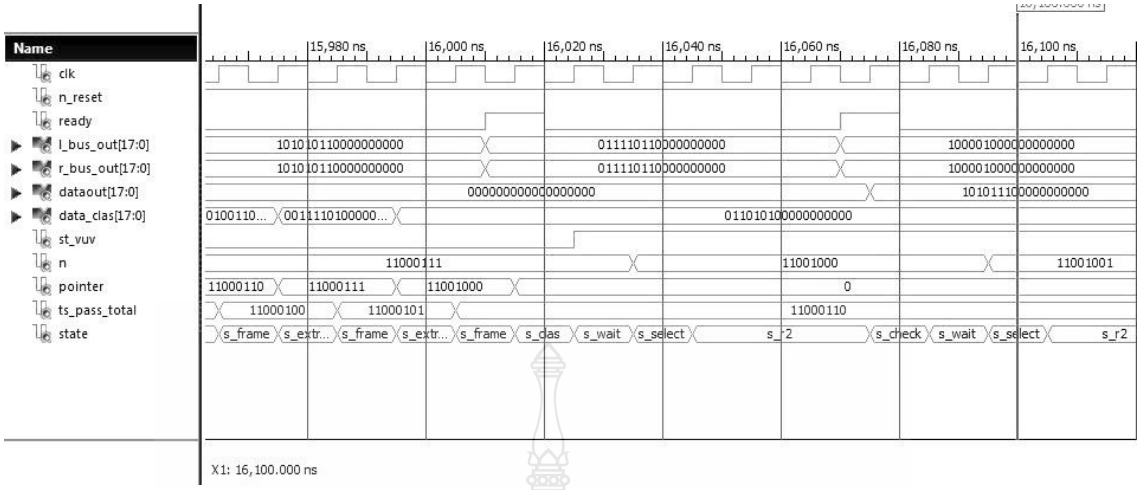
ภาพที่ 4.12 ผลจำลองการส่งข้อมูลด้วยสัญญาณสุ่มบนชิปเอฟพีจีเอ

จากภาพที่ 4.12 จะเป็นการแสดงผลการประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอด้วยการจำลองการทำงานลงบนชิปเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเฮชดีแอลด้วยซอฟแวร์ Xilinx ISE จะได้ว่าสัญญาณ st_vuv มีค่าเท่ากับ 0 และ N มีค่าช่วง 0 - 199 สถานะจะเปลี่ยนจาก s_select ไปเป็น s_r1uv ซึ่งจะทำบันทึกข้อมูลสัญญาณเสียงในหน่วยความจำตัวที่ 1 และทำการส่งข้อมูลสัญญาณสุ่มไปยังระบบเพื่อคืนกลับข้อมูลสัญญาณเสียง จากนั้นสถานะจะเปลี่ยนไปเป็น s_check, s_wait และ s_select เพื่อการประมวลในครั้งต่อไป จากภาพที่ 4.13 เป็นการแสดงผลการประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอด้วยการจำลองการทำงานลงบนชิปเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเฮชดีแอลด้วยซอฟแวร์ Xilinx ISE จะได้ว่าสัญญาณ N มีค่าเท่ากับ 199 และ 399 ค่าสถานะจะเปลี่ยนไปเป็น s_frams การประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอเสียงจะถูกจัดเก็บเป็นเฟรม ก่อนที่ค่าสถานะจะเปลี่ยนไปเป็นสถานะ s_extract จะทำการสกัดคุณลักษณะเด่นของสัญญาณแล้วเก็บข้อมูลเฟรมเสียงนั้นๆ เมื่อสัญญาณ pointer มีค่าไม่เท่ากับ 199 สถานะจะกลับไปที่ s_frams จนกว่าสัญญาณ pointer มีค่าเท่ากับ 199



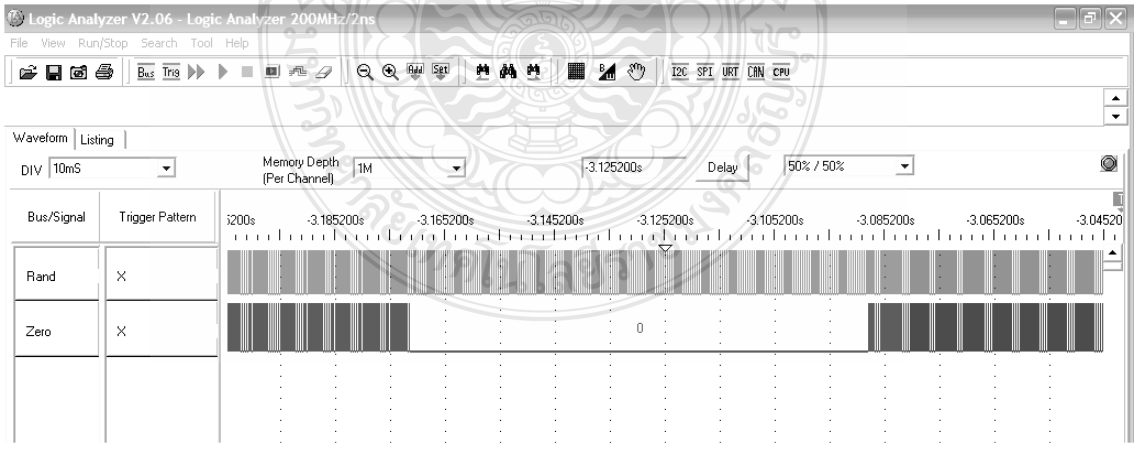
ภาพที่ 4.13 ผลจำลองการแยกแยะสัญญาณเสียงพูดบนชิปเอฟพีจีเอ เมื่อ N เท่ากับ 199 หรือ 399

จากภาพที่ 4.14 เป็นการแสดงผลการประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอด้วยผลการจำลองการทำงานลงบนชิปเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเฮชดีแอลด้วยซอฟแวร์ Xilinx ISE เมื่อสัญญาณ pointer มีค่าเท่ากับ 199 สถานะจะเปลี่ยนจาก s_frams ไปเป็น s_clas เพื่อการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดโดยอาศัยข้อมูลของเฟรมเสียงที่ได้จากการสกัดคุณลักษณะเด่นของสัญญาณ เสร็จแล้วสถานะจะเปลี่ยนไปเป็น s_wait, s_select เพื่อทำการประมวลในครั้งต่อไป ซึ่งผลการจำลองการทำงานลงบนชิปเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเฮชดีแอลด้วยซอฟแวร์ Xilinx ISE



ภาพที่ 4.14 ผลจำลองการแยกแยะสัญญาณเสียงพูดบนชิปเอฟพีจีเอ เมื่อ pointer เท่ากับ 199

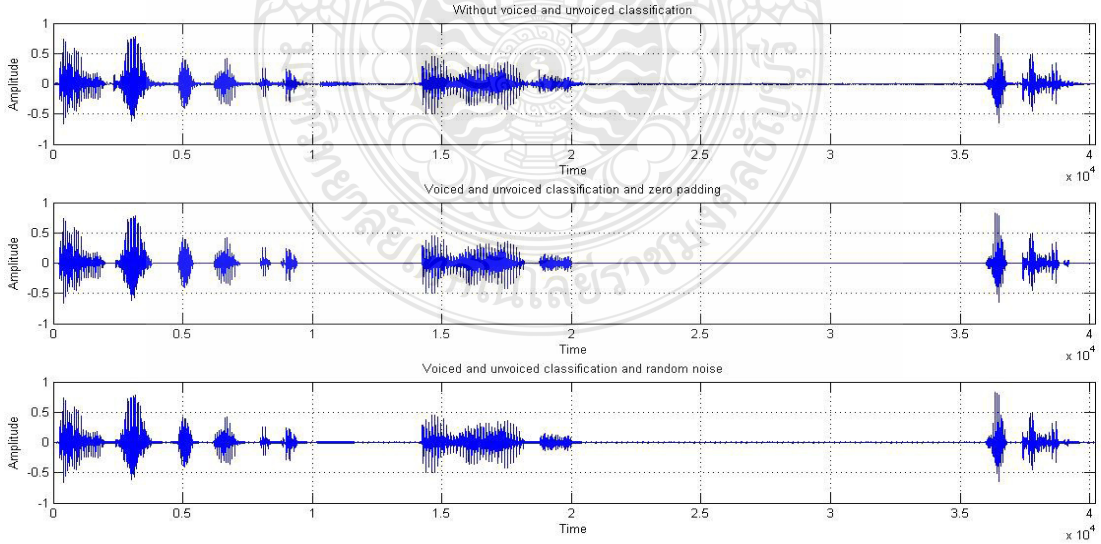
จากผลการจำลองการทำงานบนชิปเอฟพีจีเอพบว่าสามารถทำงานตามคำสั่งที่ได้ออกแบบไว้ หลังจากนั้นซอฟแวร์ Xilinx ISE จะช่วยสังเคราะห์เป็นวงจรลอจิกที่ตัวชิปเอฟพีจีเอและทำการติดตั้งบนบอร์ดเอฟพีจีเอ เพื่อควบคุมการทำงานระหว่างชิปเอฟพีจีเอและชิป LM4550 AC97 Audio CODEC ทำการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ โดยได้แทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยการคำนวณของเอฟพีจีเอ ซึ่งได้วัดการทำงานดังกล่าวจากสัญญาณ AC97_SDATA_OUT จะแสดงผลดังภาพที่ 4.15



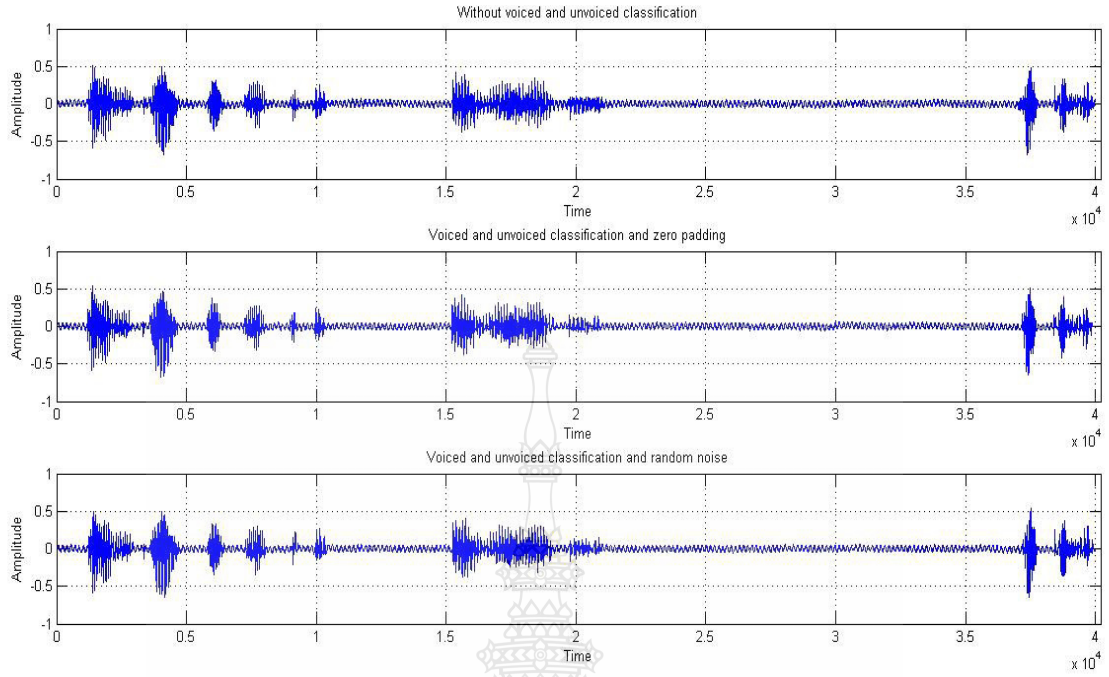
ภาพที่ 4.15 ผลการส่งข้อมูลบนชิปเอฟพีจีเอจากสัญญาณ AC97_SDATA_OUT

4.3 ผลการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด

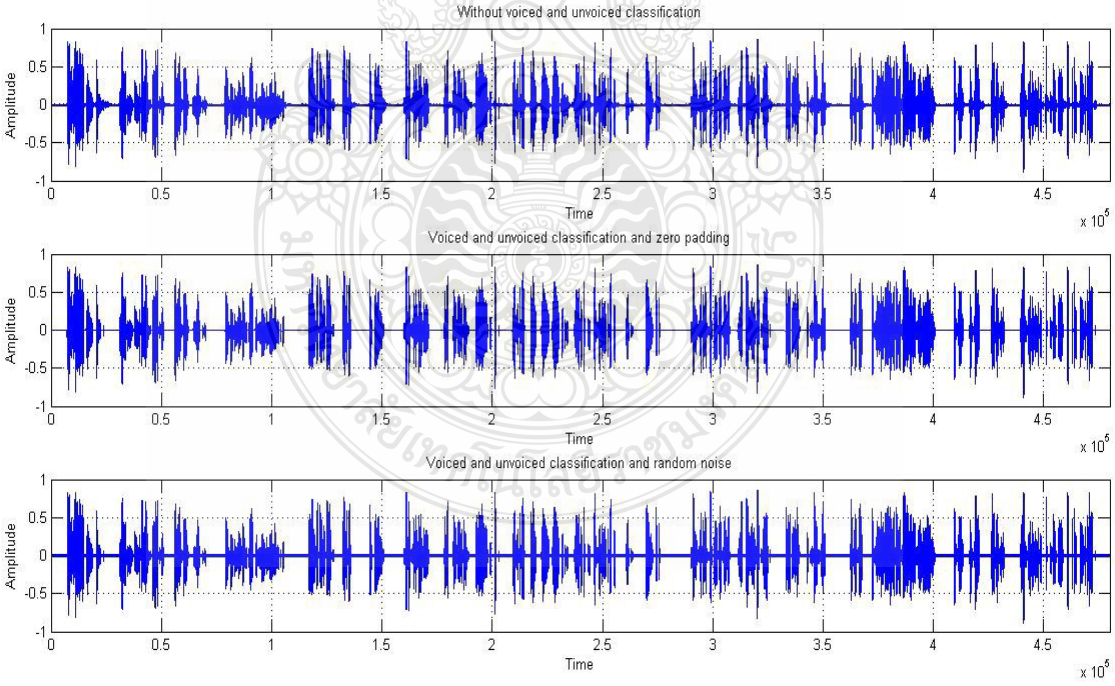
ผลการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดจากการแทนที่ที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์และการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่มนั้น ในงานวิจัยนั้นได้ทดสอบ 2 ส่วน คือ การทดสอบด้วยโปรแกรม MATLAB และบอร์ดเอฟพีจีเอ ซึ่งได้ทำการวิเคราะห์สัญญาณ 2 รูปแบบ ในรูปแบบแรกจะวิเคราะห์สัญญาณในโดเมนเวลา รูปแบบที่สองจะวิเคราะห์สัญญาณในโดเมนสเปกโตรแกรมทำให้สามารถวิเคราะห์ลักษณะต่างๆ ของสัญญาณเสียงทั้งทางเวลา หรือความถี่ หลังจากนั้นจึงจะคาดเดารูปร่างของทางเสียง (Vocal Tract) ตลอดจนแหล่งกำเนิดสัญญาณ ในตำแหน่งเวลาที่สอดคล้องกับลักษณะทางเวลาหรือทางความถี่ของสัญญาณเสียง ซึ่งจากภาพที่ 4.16 - 4.19 นั้นเป็นการเปรียบเทียบการบีบอัดสัญญาณเสียงพูดระหว่างสัญญาณเสียงที่ไม่ผ่านการแยกแยะส่วนสัญญาณเสียงกับสัญญาณเสียงที่ไม่ใช่เสียงพูดกับการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์และการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่มในโดเมนเวลา พบว่าการบีบอัดสัญญาณเสียงพูดด้วยโปรแกรม MATLAB ในส่วนสัญญาณเสียงที่แทนสัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่มจะมีความคล้ายสัญญาณสัญญาณเสียงต้นฉบับมากกว่าแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์ แต่ในการบีบอัดสัญญาณเสียงพูดบนบอร์ดเอฟพีจีเอ พบว่าการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์และการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่มในโดเมนเวลาไม่แตกต่างกัน เนื่องจากการประมวลผลบนเวลาจริงบนบอร์ดเอฟพีจีเอเกิดสัญญาณรบกวนภายในระบบขึ้น



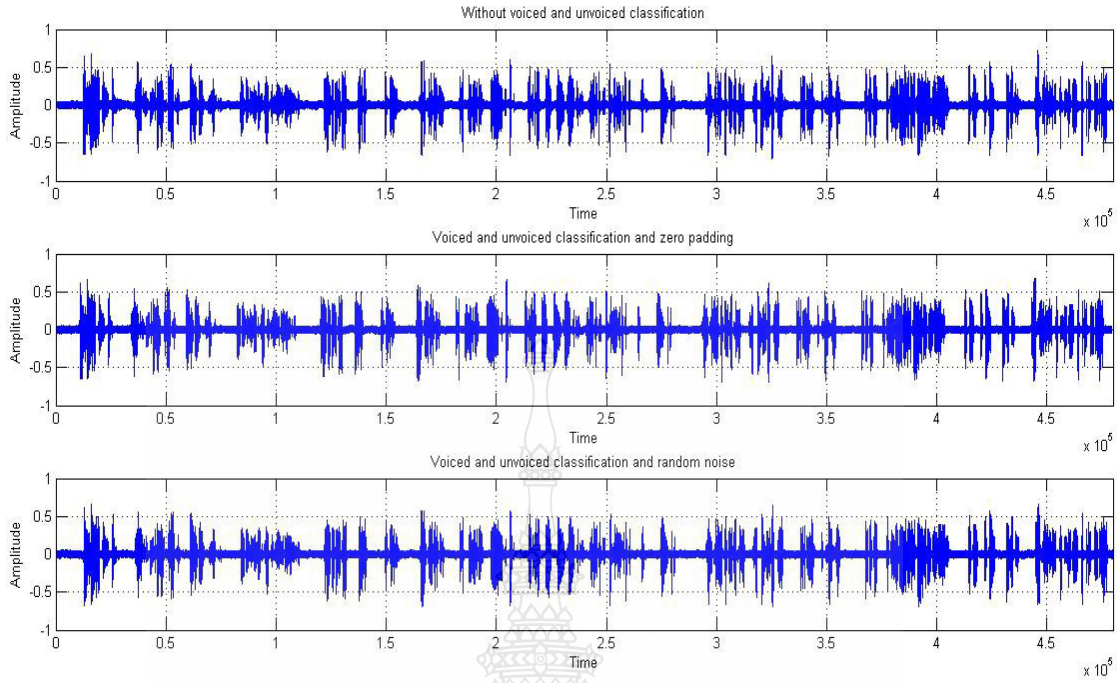
ภาพที่ 4.16 การบีบอัดสัญญาณเสียงพูด 5 วินาที ด้วยโปรแกรม MATLAB ในโดเมนเวลา



ภาพที่ 4.17 การบีบอัดสัญญาณเสียงพูด 5 วินาที บนบอร์ดเอฟพีจีเอในโดเมนเวลา

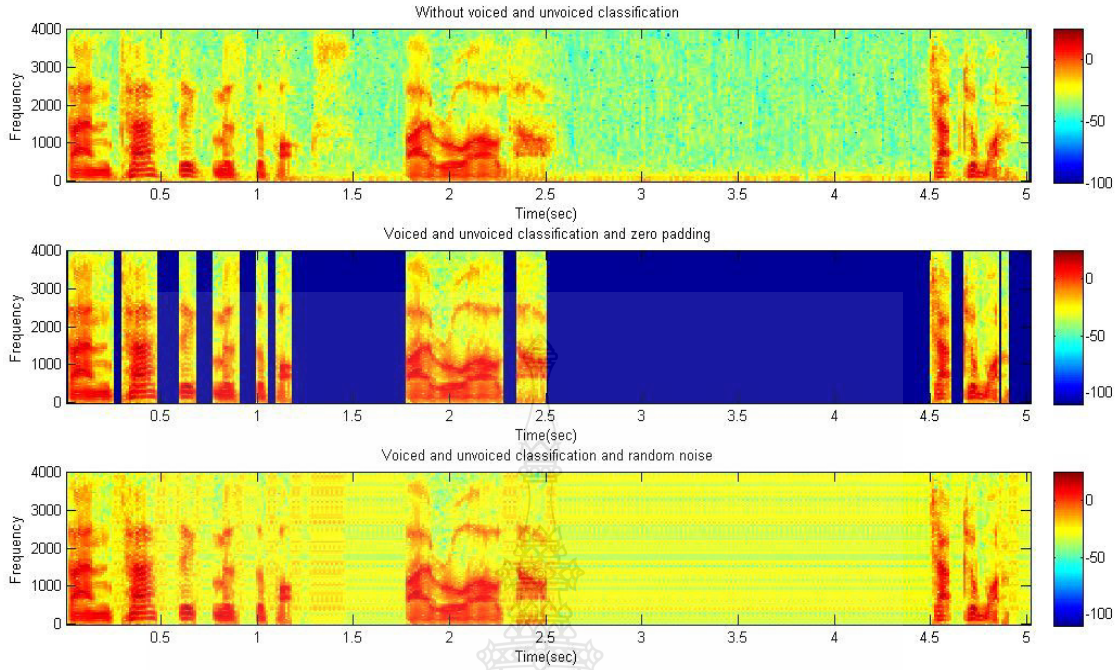


ภาพที่ 4.18 การบีบอัดสัญญาณเสียงพูด 60 วินาที ด้วยโปรแกรม MATLAB ในโดเมนเวลา

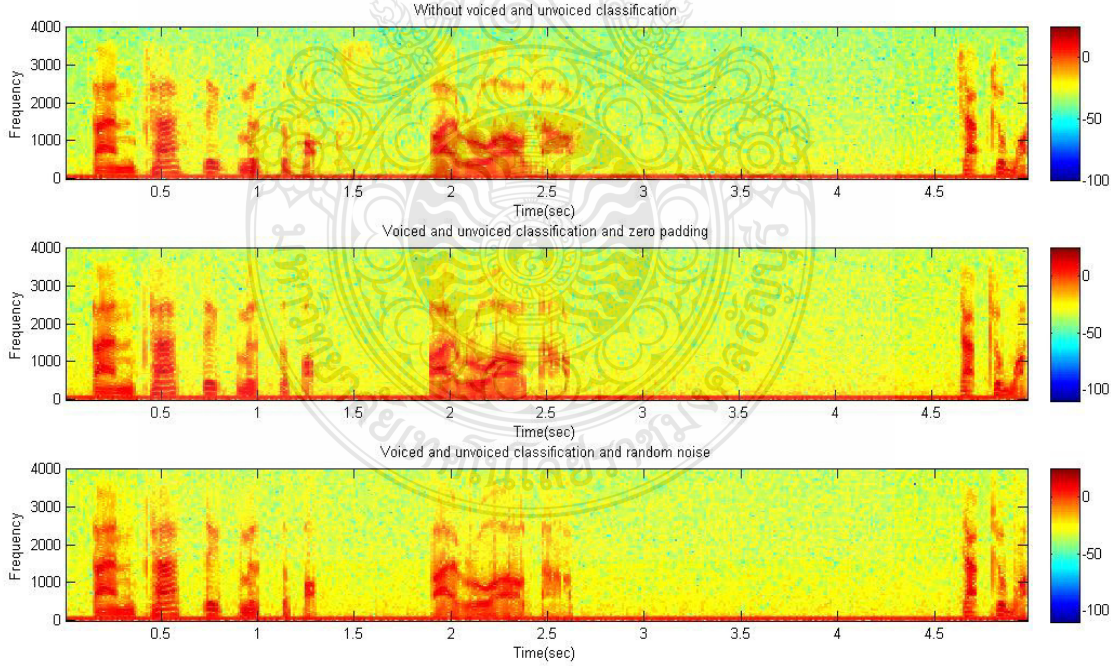


ภาพที่ 4.19 การบีบอัดสัญญาณเสียงพูด 60 วินาที บนบอร์ดเอฟพีจีเอในโดเมนเวลา

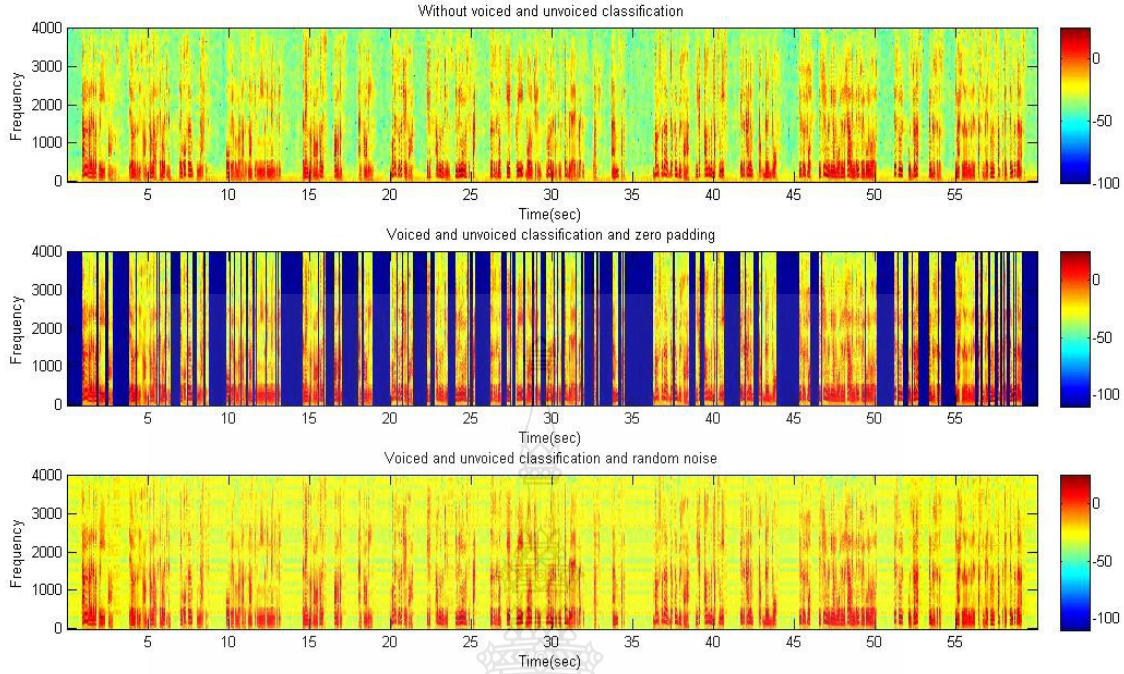
จากภาพที่ 4.20 - 4.23 นั้นเป็นการเปรียบเทียบในส่วนการบีบอัดสัญญาณเสียงพูดระหว่างสัญญาณเสียงที่ไม่ผ่านการแยกแยะส่วนสัญญาณเสียงกับสัญญาณเสียงที่ไม่ใช่เสียงพูดกับการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์และการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่มในโดเมนสเปกโตรแกรม พบว่าการบีบอัดสัญญาณเสียงพูดด้วยโปรแกรม MATLAB ซึ่งสัญญาณเสียงที่แทนสัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่มจะมีความคล้ายสัญญาณสัญญาณเสียงต้นฉบับมากกว่าแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์ แต่ในการบีบอัดสัญญาณเสียงพูดบนบอร์ดเอฟพีจีเอพบว่าการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์และการแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่มในโดเมนสเปกโตรแกรมไม่แตกต่างกัน เนื่องจากการประมวลผลบนเวลาจริงบนบอร์ดเอฟพีจีเอเกิดสัญญาณรบกวนภายในระบบขึ้น ในส่วนผลลัพธ์อื่นๆ จะแสดงไว้ในภาคผนวก ข



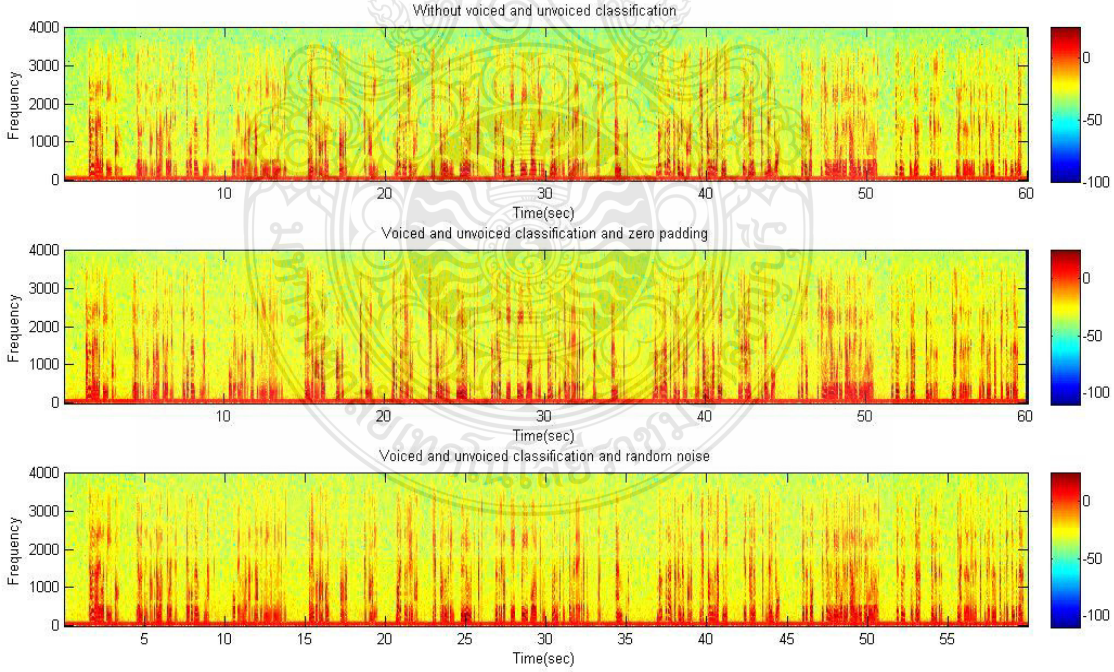
ภาพที่ 4.20 การบีบอัดสัญญาณเสียงพูด 5 วินาที ด้วยโปรแกรม MATLAB ในสเปกโตรแกรม



ภาพที่ 4.21 การบีบอัดสัญญาณเสียงพูด 5 วินาที บนบอร์ดเอฟพีจีเอในสเปกโตรแกรม



ภาพที่ 4.22 การบีบอัดสัญญาณเสียงพูด 60 วินาที ด้วยโปรแกรม MATLAB ในสเปกโตรแกรม



ภาพที่ 4.23 การบีบอัดสัญญาณเสียงพูด 60 วินาที บนบอร์ดเอฟพีจีเอในสเปกโตรแกรม

บทที่ 5
สรุปผลการทดลอง

ในบทนี้จะกล่าวถึงสรุปการพัฒนาเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณรบกวนหรือสัญญาณเสียงที่ไม่ใช่เสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ ข้อเสนอแนะสำหรับแนวทางการพัฒนางานวิจัยต่อไป

5.1 สรุปผลงานวิจัย

วิทยานิพนธ์ฉบับนี้เป็นการพัฒนาเทคนิคการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณรบกวนแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ โดยปกติแล้วในกระบวนการส่งสัญญาณเสียงที่มีการบีบอัดสัญญาณเสียงนั้น เพื่อต้องการลดขนาดของสัญญาณเสียงพูดจากสัญญาณเสียงต้นฉบับ และต้องการให้สัญญาณเสียงพูดมีคุณสมบัติใกล้เคียงกับสัญญาณเสียงพูดต้นฉบับมากที่สุดหรือขึ้นอยู่กับการนำเสียงนั้นๆ ไปประยุกต์ใช้งาน ซึ่งสัญญาณเสียงพูดที่นำมาวิเคราะห์ในงานวิจัยนี้ได้แก่ เสียงของผู้หญิงและผู้ชายที่มีความยาว 5 วินาที และ 60 วินาที ได้เปรียบเทียบค่าประสิทธิภาพสัญญาณโดยใช้หลักการ SNR และ NRMSE ในงานวิจัยได้นำเสนอการหาค่าเส้นตัดสินใจที่เหมาะสมกับการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดและแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์และสัญญาณสุ่ม ที่ค่าเส้นตัดสินใจ 0.5 – 7.0% ด้วยโปรแกรม MATLAB ดังตารางที่ 5.1 จากในงานวิจัยที่ [10] พบว่าค่าเส้นตัดสินใจที่ 1% และการสุ่มสัญญาณที่ 50% นั้นจะมีค่าที่เหมาะสมในการใช้แยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด เมื่อได้ทำการพัฒนาเทคนิคให้เหมาะสมในการประมวลผลแบบฐานเวลาจริงนั้น พบว่าจากค่าเส้นตัดสินใจได้เลือกใช้ที่ 1% เพราะสามารถลดขนาดของสัญญาณเสียงโดยมีอัตราบีบอัดข้อมูลไม่น้อยกว่า 1.5 ตามที่ขอบเขตกำหนดไว้ ซึ่งมีค่าอัตราบีบอัดข้อมูล 1.970 มีค่า SNR แทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์เท่ากับ 20.024 dB และแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่มเท่ากับ 19.398 dB และค่า NRMSE SNR แทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์เท่ากับ 0.113 และแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่มเท่ากับ 0.105 ซึ่งมีค่าประสิทธิภาพและค่าผิดพลาดของการบีบอัดสัญญาณเสียงพูดที่เหมาะสม หลังจากนั้นได้นำไปออกแบบระบบบีบอัดสัญญาณเสียงพูดบนพื้นฐานเวลาจริงบนบอร์ดเอฟพีจีเอด้วยภาษาวีเฮชดีแอล ก่อนที่จะสังเคราะห์เป็นวงจรลอจิกไปที่ชิปเอฟพีจีเอ ซึ่งชิปเอฟพีจีเอที่ถูกโปรแกรมจะควบคุมการทำงาน LM4550 AC97 Audio CODEC ที่ทำการติดตั้งบนบอร์ดเอฟพีจีเอ Xilinx Virtex-II Pro XC2VP30 ทำให้สามารถการประมวลผลสัญญาณแบบฐานเวลาจริงได้

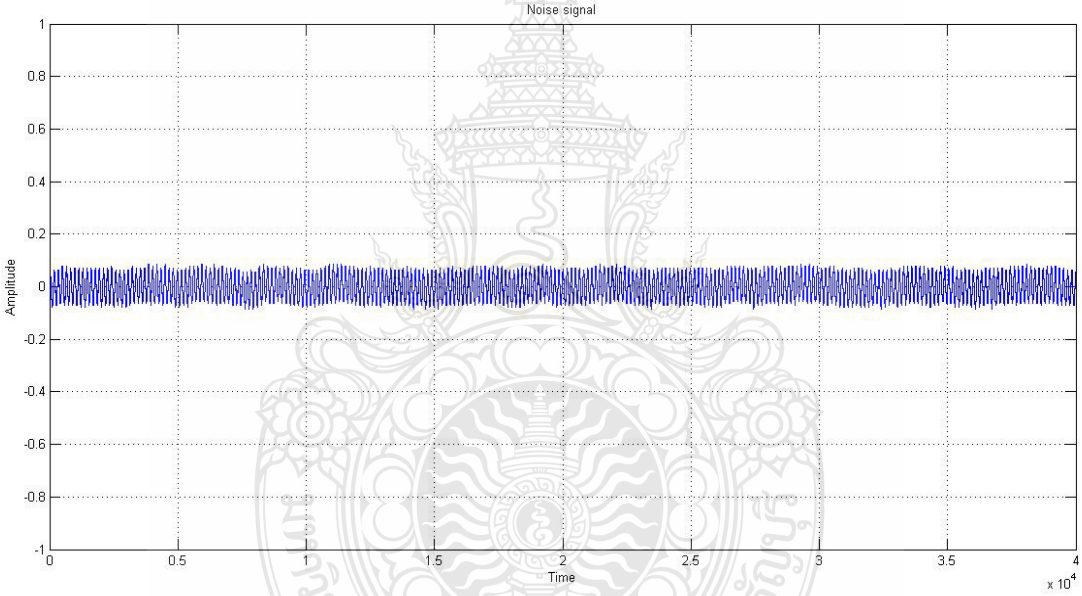
ตารางที่ 5.1 การวัดค่าประสิทธิภาพของสัญญาณเสียงพูดที่ค่าเส้นตัดสินใจที่ 0.5 - 7%

ค่าเส้นตัดสินใจ (%)	อัตราบีบอัดข้อมูล	SNR		NRMSE	
		แทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์	แทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่ม	แทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์	แทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่ม
0.5	1.691	24.321	22.939	0.064	0.075
1.0	1.970	20.024	19.398	0.105	0.113
1.5	1.970	20.024	19.398	0.105	0.113
2.0	2.212	17.104	16.756	0.1471	0.154
2.5	2.526	14.959	14.736	0.1933	0.198
3.0	2.526	14.959	14.736	0.193	0.198
3.5	2.898	13.214	13.059	0.240	0.245
4.0	3.488	11.480	11.369	0.296	0.300
4.5	3.488	11.480	11.369	0.296	0.300
5.0	4.281	10.191	10.105	0.346	0.350
5.5	5.388	9.154	9.085	0.390	0.394
6.0	5.390	9.153	9.084	0.390	0.394
6.5	9.438	8.186	8.128	0.435	0.438
7.0	9.438	8.186	8.128	0.435	0.438

งานวิจัยได้ศึกษาระบบประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอบนฐานเวลาจริง 3 ระบบ โดยที่ระบบแรกจะประมวลผลสัญญาณเสียงบนชิปเอฟพีจีเอบนฐานเวลาจริงที่ไม่ได้จัดการสัญญาณเสียงที่ไม่ใช่เสียงพูดระบบที่สองจะออกแบบระบบแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดบนฐานเวลาจริงโดยประมวลผลบนเอฟพีจีเอโดยแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยศูนย์และระบบสามจะออกแบบระบบแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดบนฐานเวลาจริงโดยประมวลผลบนเอฟพีจีเอและแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดด้วยสัญญาณสุ่มจากการคำนวณบนเอฟพีจีเอ ออกแบบระบบด้วยโปรแกรม Xilinx ISE ซึ่งได้ผลจากการออกแบบระบบ 3 ระบบนั้นจะใช้ทรัพยากรภายในชิปเอฟพีจีเอดังตารางที่ 5.2

ตารางที่ 5.2 การใช้ทรัพยากรภายในชิปเอฟพีจีเอ

Name	Without V/UV Classification		With V/UV Classification and Zero Padding		With V/UV Classification and Random Noise	
	Used/Available	Util.	Used/Available	Util.	Used/Available	Util.
Slice Flip Flops	166/27392	0. 6%	276/27392	1.0 %	301/27392	1.1%
occupied slice	126/13,696	0. 9%	1,056/13,696	7.7 %	984/13,696	7.2%
4 input LUT	137 /27,392	0. 5%	1,854 /27,392	6.8 %	1,646/27,392	6.0%
bonded	10 /556	1.8%	10 /556	1.8%	10 /556	1.8%
BUFMUXs	2 /16	12.5%	2 /16	12.5%	2 /16	12.5%



ภาพที่ 5.1 สัญญาณรบกวนภายในระบบ

จากตารางที่ 5.2 จะสังเกตว่าการใช้ทรัพยากรภายในชิปเอฟพีจีเอเมื่อสังเคราะห์เป็นวงจรลอจิกไปที่ชิปเอฟพีจีเอนั้นมีแนวโน้มในการใช้ทรัพยากรภายในชิปเอฟพีจีเอจะมากขึ้นเมื่อได้เพิ่มโปรแกรมเพิ่มในส่วนการแยกแยะสัญญาณที่เป็นเสียงและเสียงรบกวนและการแทนที่สัญญาณรบกวนนั้น พบว่าไม่เกินกว่าทรัพยากรที่มีในชิปเอฟพีจีเอและโดยระบบที่ออกแบบสามารถทำงานบนฐานเวลาจริงได้โดยใช้เวลาในการประมวลผลเพียง 25 มิลลิวินาที ซึ่งระบบมีอัตราสุ่มสัญญาณ 8 กิโลเฮิรตซ์และใช้เสียงขนาด 18 บิต ซึ่งไม่ส่งผลต่อการเกิดการหน่วงเวลา แต่หลังจากการแทนที่

สัญญาณรบกวนด้วยสัญญาณศูนย์และสัญญาณสุ่มบนฐานเวลาบนบอร์ดเอฟพีจีเอพบว่าสัญญาณรบกวนขึ้นภายในระบบที่เกิดขึ้นดังภาพที่ 5.1 ทำให้ระหว่างแทนที่สัญญาณรบกวนด้วยสัญญาณศูนย์และสัญญาณสุ่มแบบเวลาจริงไม่เกิดความแตกต่างกันเมื่อเทียบกับบนซอฟต์แวร์

5.2 ข้อเสนอแนะ

5.2.1 เทคนิคในการวิเคราะห์สัญญาณเสียงพูดในการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดและแทนที่สัญญาณสัญญาณเสียงที่ไม่ใช่เสียงพูดนั้นมีหลายเทคนิค ในการวิจัยได้เลือกเทคนิคเส้นตัดสินใจเพื่อใช้กับการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูด ดังนั้นควรเปรียบเทียบเทคนิคอื่นๆ เพื่อเปรียบเทียบกับผลการวิจัยแต่เทคนิคนั้นต้องเหมาะสมกับการประผลบนพื้นฐานเวลาจริง

5.2.2 ควรมีการปรับปรุงสัญญาณเสียงพูดหลังจากทำการบีบอัดสัญญาณ เช่น การใช้ตัวกรองคาลมาน (Kalman Filter) เพื่อลดสัญญาณรบกวนที่เกิดขึ้นจากการประมวลผลบนฐานเวลาจริง

5.2.3 ควรนำการการแยกแยะสัญญาณที่เป็นเสียงและสัญญาณเสียงที่ไม่ใช่เสียงพูดและแทนที่สัญญาณเสียงที่ไม่ใช่เสียงพูดบนพื้นฐานเวลาจริงบนบอร์ดเอฟพีจีเอไปประยุกต์ใช้ในการประมวลผลที่ต้องอาศัยหน่วยประมวลผลที่มีความสามารถสูง เช่น การบีบอัดข้อมูลเสียงพูดจากสัญญาณเสียงที่มีกว่า 2 แหล่งกำเนิดพร้อมกันหรือการบีบอัดข้อมูลที่มีทั้งภาพและเสียงพร้อมกัน เช่น การประชุมทางไกลผ่านจอภาพ (Video Conference)

รายการอ้างอิง

[1] Kreangsak Pattanaburi and Jakkree Srinonchat, “Enhancement Pattern Analysis Technique for Voiced/Unvoiced Classification,” **IEEE-International Symposium on Computer Consumer and Control (IS3C)**, 2012. pp. 389-392.

[2] Kushki, A. and Plataniotis, K.N., “Time-Compression of Speech in Information Talks Using Spectral Entropy,” **Eight International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS'07)**, 2007. pp. 80.

[3] วีระยุทธ คุณรัตนสิริ และ จักรี ศรีนนท์ฉัตร, “การเปรียบเทียบประสิทธิภาพการรับส่งข้อมูลร่วมกับการบีบอัดสัญญาณเสียงโดยการแปลงเวฟเล็ต,” **การประชุมวิชาการ งานวิจัยและพัฒนาเชิงประยุกต์ ครั้งที่ 3 (ECTI- CARD)**, 2554. หน้า 382-385.

[4] Xu, J. Ariyaeeinia, A. and Sotudeh, R., “User Voice Idenfication on Fpga,” **Perspectives in Pervasive Computing,** 2005. pp. 75-83.

[5] Junhee Jung, Seunghun Jin, Dongkyun Kim, Hyung Soon Kim, Jong Suk Choi and Jae Wook Jeon, “A Voice Activity Detection System Based on FPGA,” **International Conference on Control Automation and Systems,** 2010. pp 2304-2308.

[6] Jung Uk Cho, Dae Ro Lee, Jong Hyun Park, Hyung Soon Kim, Chang Hoon Lee Jong Suk Choi and Jae Wook Jeon, “An FPGA-Based Voice Signal Preprocessor for The Real-Time Cross-Correlation,” **International Conference on Control Automation and Systems**, 2007. pp. 793-797.

[7] Upal Mahbub, Tauhidur Rahman and A. B. M. H. Rashid, “FPGA Implementation of Real Time Acoustic Noise Suppression by Spectral Subtraction using Dynamic Moving Average Method,” **IEEE Symposium on Industrial Electronics and Applications (ISIEA 2009)**, 2009. pp. 365-370.

[8] Hung-Chih Lai, Marios Savvides and Tsuhan Chen, “Proposed FPGA Hardware Architecture for High Frame Rate (>>100 fps) Face Detection Using Feature Cascade Classifiers,” **IEEE International Conference on Biometrics Theory Applications and Systems 2007 (BTAS 2007)**, 2007. pp. 1-6.

[9] Jakkree Srinonchat, “External Memory System Optimization for FPGA Based Implementation of Speech Signal Processing,” **Proceedings of the World Congress on Engineering Vol. II 2012 (WCE 2012)**, 2012.

[10] สุภาธิณี กรสิงห์, เกรียงศักดิ์ พัฒนบุรี, เฉลิมเกียรติ สุตาชา และ จักรี ศรีนนท์ฉัตร, “การวิเคราะห์และเปรียบเทียบเทคนิคการสกัดคุณลักษณะของสัญญาณที่เป็นเสียงและเสียงรบกวน,” **การประชุมวิชาการเครือข่ายวิศวกรรมไฟฟ้า มหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 3**, 2554. หน้า 335-338.

[11] Chalermkiat Sutacha and Jakkree Srinonchat, “Improvement Voiced and Unvoiced Classification Technique Based on Real Time Processing Using FPGA Board,” **KMITL Science and Technology Journal (KMITL-STJ)**, Vol. 12, No. 2, 2013.

[12] เฉลิมเกียรติ สุตาชา และ จักรี ศรีนนท์ฉัตร, “การพัฒนาเทคนิคการประมวลผลสัญญาณเสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ,” **การประชุมวิชาการเครือข่ายวิศวกรรมไฟฟ้า มหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 4,** 2555. หน้า 585-588.

[13] เฉลิมเกียรติ สุตาชา, อรรถพล เพียรผลดีสกุล และ จักรี ศรีนนท์ฉัตร, “การประยุกต์ใช้เอฟพีจีเอสำหรับการบีบอัดข้อมูล,” **การประชุมวิชาการมหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 3**, 2553.

[14] เฉลิมเกียรติ สุตาชา และ จักรี ศรีนนท์ฉัตร, “การพัฒนาเทคนิคการแยกแยะเสียงโฆษะและเสียงอโฆษะแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ,” **การประชุมเครือข่ายวิชาการบัณฑิตศึกษาแห่งชาติ ครั้งที่ 1 (GS-NETT 2012)**, 2555. หน้า 34.

[15] Kreangsak Pattanaburi and Jakkree Srinonchat, “Devlopment Voiced-Unvoiced Classification Technique Based On LSP Coefficients**,” Electical Engineering Network 2012 of Rajamangala University of Technology (EENET 2012)**, 2012. pp. 541-543.

[16] Supatinee Kornsing and Jakkree Srinonchat, “Enhancement Speech Compression Technique Using Modern Wavelet Tranforms,” **International Symposium on Computer Consumer and Control**, 2012. pp. 393-396.

[17] Noman Junejo, Ahmed, N. Unar, M.A. Rajput ans A.Q.K., “Speech and Image Compression Using Discrete Wavelet Transform,” **IEEE/Sarnoff Symposium on Advances in Wired and Wireless Communication**, 2005. pp. 45-48.

[18] Song Lin, Dimitrios Gunopulos, Vana Kalogeraki, Stefano Lonardi, “A Data Compression Technique for Sensor Networks with Dynamic Bandwidths Allocation,” **IEEE-Proceedings of the 12th International Symposium on Temporal Representation and Reasoning (TIME’05)**, 2005. pp. 186-188.

[19] เฉลิมเกียรติ สุตาชา และ จักรี ศรีนนท์ฉัตร, “การบีบอัดสัญญาณเสียงโดยใช้เทคนิคซีโร่แพดดิ้งบนพื้นฐานเวลาจริงบนบอร์ดเอฟพีจีเอ,” **การประชุมวิชาการเครือข่ายวิศวกรรมไฟฟ้า มหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 5**, 2556.

[20] รศ.ดร. มนัส สังวรศิลป์ และ วรรัตน์ ภัทรอมรกุล, **คู่มือการใช้งาน MATLAB ฉบับสมบูรณ์**. พิมพ์ครั้งที่ 1. กรุงเทพฯ: สำนักพิมพ์อินโฟเพรส, 2543.

ภาคผนวก

ภาคผนวก ก

อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ที่ค่าเส้นตัดสินใจที่ 0.5 - 7%

ตารางที่ ก.1 อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ของสัญญาณเสียงผู้หญิงที่เวลา 5 วินาที

ค่าเส้นตัดสินใจ (%)	อัตราบีบอัดข้อมูล	SNR		NRMSE	
		แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม	แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม
0.5	1.484	25.465	24.644	0.056	0.061
1.0	1.615	20.983	20.684	0.092	0.095
1.5	1.615	20.983	20.684	0.092	0.095
2.0	1.668	18.341	18.176	0.126	0.128
2.5	1.782	17.143	17.013	0.143	0.145
3.0	1.782	17.143	17.013	0.143	0.145
3.5	1.854	16.077	15.973	0.161	0.163
4.0	1.943	14.601	14.523	0.190	0.191
4.5	1.943	14.601	14.523	0.190	0.191
5.0	2.040	13.490	13.429	0.216	0.217
5.5	2.146	12.505	12.454	0.243	0.244
6.0	2.146	12.505	12.454	0.243	0.244
6.5	2.268	11.359	11.317	0.276	0.278
7.0	2.268	11.359	11.317	0.276	0.278

ตารางที่ ก.2 อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ของสัญญาณเสียงผู้ชายที่เวลา 5 วินาที

ค่าเส้นตัดสินใจ (%)	อัตราบีบอัดข้อมูล	SNR		NRMSE	
		แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม	แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม
0.5	1.844	23.261	21.107	0.075	0.095
1.0	2.262	18.539	17.515	0.128	0.143
1.5	2.262	18.539	17.515	0.128	0.143
2.0	2.623	15.202	14.640	0.186	0.198
2.5	3.127	12.175	11.849	0.264	0.274
3.0	3.127	12.175	11.849	0.264	0.274
3.5	3.811	9.838	9.619	0.346	0.355
4.0	4.994	7.513	7.369	0.446	0.453
4.5	4.994	7.513	7.369	0.446	0.453
5.0	6.752	6.096	5.983	0.523	0.530
5.5	9.711	4.867	4.778	0.596	0.603
6.0	9.711	4.867	4.778	0.596	0.603
6.5	23.896	3.895	3.820	0.665	0.670
7.0	23.896	3.895	3.820	0.665	0.670

ตารางที่ ก.3 อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ของสัญญาณเสียงพูดที่เวลา 5 วินาที

ค่าเส้นตัดสินใจ (%)	อัตราบีบอัดข้อมูล	SNR		NRMSE	
		แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม	แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม
0.5	1.66	24.36	22.88	0.07	0.08
1.0	1.94	19.76	19.10	0.11	0.12
1.5	1.94	19.76	19.10	0.11	0.12
2.0	2.15	16.77	16.41	0.16	0.16
2.5	2.45	14.66	14.43	0.20	0.21
3.0	2.45	14.66	14.43	0.20	0.21
3.5	2.83	12.96	12.80	0.25	0.26
4.0	3.47	11.06	10.95	0.32	0.32
4.5	3.47	11.06	10.95	0.32	0.32
5.0	4.40	9.79	9.71	0.37	0.37
5.5	5.93	8.69	8.62	0.42	0.42
6.0	5.93	8.69	8.62	0.42	0.42
6.5	13.08	7.63	7.57	0.47	0.47
7.0	13.08	7.63	7.57	0.47	0.47

ตารางที่ ก.4 อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ของสัญญาณเสียงผู้หญิงที่เวลา 60 วินาที

ค่าเส้นตัดสินใจ (%)	อัตราบีบอัดข้อมูล	SNR		NRMSE	
		แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม	แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม
0.5	1.484	25.465	24.644	0.056	0.061
1.0	1.615	20.983	20.684	0.092	0.095
1.5	1.615	20.983	20.684	0.092	0.095
2.0	1.668	18.341	18.176	0.126	0.128
2.5	1.782	17.143	17.013	0.143	0.145
3.0	1.782	17.143	17.013	0.143	0.145
3.5	1.854	16.077	15.973	0.161	0.163
4.0	1.943	14.601	14.523	0.190	0.191
4.5	1.943	14.601	14.523	0.190	0.191
5.0	2.040	13.490	13.429	0.216	0.217
5.5	2.146	12.505	12.454	0.243	0.244
6.0	2.146	12.505	12.454	0.243	0.244
6.5	2.268	11.359	11.317	0.276	0.278
7.0	2.268	11.359	11.317	0.276	0.278

ตารางที่ ก.5 อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ของสัญญาณเสียงผู้ชายที่เวลา 60 วินาที

ค่าเส้นตัดสินใจ (%)	อัตราบีบอัดข้อมูล	SNR		NRMSE	
		แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม	แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม
0.5	1.829	24.290	22.282	0.063	0.079
1.0	2.234	19.310	18.387	0.111	0.123
1.5	2.234	19.310	18.387	0.111	0.123
2.0	2.676	15.775	15.263	0.166	0.177
2.5	3.223	12.946	12.635	0.230	0.239
3.0	3.223	12.946	12.635	0.230	0.239
3.5	3.872	10.715	10.515	0.297	0.304
4.0	4.876	8.753	8.609	0.372	0.378
4.5	4.876	8.753	8.609	0.372	0.378
5.0	6.107	7.124	7.014	0.448	0.453
5.5	7.384	6.111	6.025	0.502	0.507
6.0	7.389	6.168	6.080	0.499	0.504
6.5	9.158	5.266	5.194	0.553	0.557
7.0	9.158	5.266	5.194	0.553	0.557

ตารางที่ ก.6 อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ของสัญญาณเสียงพูดที่เวลา 60 วินาที

ค่าเส้นตัดสินใจ (%)	อัตราบีบอัดข้อมูล	SNR		NRMSE	
		แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม	แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม
0.5	1.718	24.279	23.002	0.062	0.073
1.0	2.001	20.288	19.696	0.099	0.107
1.5	2.001	20.288	19.696	0.099	0.107
2.0	2.278	17.436	17.097	0.139	0.145
2.5	2.597	15.258	15.042	0.182	0.187
3.0	2.597	15.258	15.042	0.182	0.187
3.5	2.963	13.470	13.323	0.226	0.231
4.0	3.507	11.902	11.792	0.275	0.279
4.5	3.507	11.902	11.792	0.275	0.279
5.0	4.166	10.589	10.503	0.323	0.327
5.5	4.846	9.623	9.554	0.361	0.364
6.0	4.852	9.621	9.551	0.362	0.365
6.5	5.793	8.745	8.686	0.399	0.402
7.0	5.793	8.745	8.686	0.399	0.402

ตารางที่ ก.7 อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ของสัญญาณเสียงพูดที่เวลา 5 และ 60 วินาที

ค่าเส้นตัดสินใจ (%)	อัตราบีบอัดข้อมูล	SNR		NRMSE	
		แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม	แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม
0.5	1.691	24.321	22.939	0.064	0.075
1.0	1.970	20.024	19.398	0.105	0.113
1.5	1.970	20.024	19.398	0.105	0.113
2.0	2.212	17.104	16.753	0.147	0.154
2.5	2.526	14.959	14.736	0.193	0.198
3.0	2.526	14.959	14.736	0.193	0.198
3.5	2.898	13.214	13.059	0.240	0.245
4.0	3.488	11.479	11.369	0.296	0.300
4.5	3.488	11.479	11.369	0.296	0.300
5.0	4.281	10.191	10.105	0.346	0.350
5.5	5.388	9.154	9.085	0.390	0.394
6.0	5.390	9.153	9.084	0.391	0.394
6.5	9.438	8.186	8.127	0.435	0.438
7.0	9.438	8.186	8.127	0.435	0.438

ตารางที่ ก.8 อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ของสัญญาณเสียงผู้หญิงที่เวลา 5 วินาที ณ เส้นตัดสินใจที่ 1%

ชื่อไฟล์	อัตราบีบอัดข้อมูล	SNR		NRMSE	
		แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม	แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม
speech1	2.209	24.249	23.152	0.061	0.070
speech2	1.349	23.174	22.546	0.069	0.075
speech3	1.457	27.739	26.618	0.041	0.047
speech4	1.297	27.006	25.910	0.045	0.051
speech5	1.575	24.624	23.973	0.059	0.063
speech6	1.471	27.614	26.624	0.042	0.047
speech7	1.460	27.136	26.024	0.044	0.050
speech8	1.748	20.259	19.853	0.097	0.102
speech9	1.703	20.505	20.204	0.094	0.098
speech10	1.274	27.272	26.436	0.043	0.048
speech11	1.600	21.224	20.853	0.087	0.091
speech12	1.429	23.191	22.683	0.069	0.073
speech13	1.274	28.719	27.638	0.037	0.042
speech14	1.457	29.818	28.180	0.032	0.039
speech15	1.515	27.063	26.019	0.044	0.050
speech16	1.358	27.099	26.049	0.044	0.050
speech17	1.467	30.080	28.706	0.031	0.037
speech18	1.211	23.561	23.319	0.066	0.068
speech19	1.558	24.393	23.848	0.060	0.064
speech20	1.266	24.567	24.252	0.059	0.061
ค่าเฉลี่ย	1.484	25.465	24.644	0.056	0.061

ตารางที่ ก.9 อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ของสัญญาณเสียงผู้ชายที่เวลา 5 วินาที ณ เส้นตัดสินใจที่ 1%

ชื่อไฟล์	อัตราบีบอัดข้อมูล	SNR		NRMSE	
		แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม	แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม
speech21	3.000	19.494	17.859	0.106	0.128
speech22	3.242	20.233	19.144	0.097	0.110
speech23	1.933	15.554	14.942	0.167	0.179
speech24	2.273	19.589	18.589	0.105	0.118
speech25	1.587	23.733	22.698	0.065	0.073
speech26	1.763	21.643	20.531	0.083	0.094
speech27	1.763	18.355	17.451	0.121	0.134
speech28	1.478	19.051	18.641	0.112	0.117
speech29	3.092	11.389	10.553	0.270	0.297
speech30	2.116	14.268	13.553	0.194	0.210
speech31	2.247	13.729	12.991	0.206	0.224
speech32	2.422	21.570	20.047	0.083	0.100
speech33	1.613	19.430	18.919	0.107	0.113
speech34	3.226	16.497	15.480	0.150	0.168
speech35	1.852	19.827	18.720	0.102	0.116
speech36	1.811	22.228	20.634	0.077	0.093
speech37	3.571	17.489	16.139	0.134	0.156
speech38	2.151	22.701	21.563	0.073	0.084
speech39	1.570	20.495	19.370	0.095	0.108
speech40	2.532	13.502	12.485	0.212	0.238
ค่าเฉลี่ย	2.262	18.539	17.515	0.128	0.143

ตารางที่ ก.10 อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ของสัญญาณเสียงผู้หญิงที่เวลา 60 วินาที ณ เส้นตัดสินใจที่ 1%

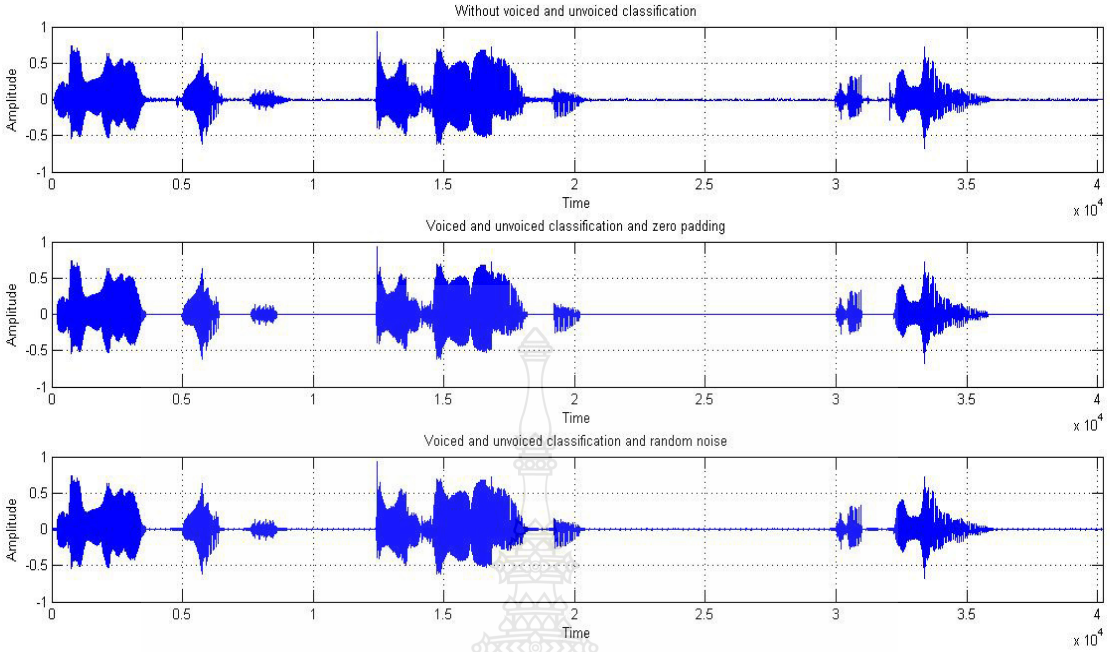
ชื่อไฟล์	อัตราบีบอัดข้อมูล	SNR		NRMSE	
		แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม	แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม
speech41	1.844	20.054	19.789	0.099	0.103
speech42	1.638	21.162	20.899	0.088	0.090
speech43	1.724	20.505	20.274	0.094	0.097
speech44	1.690	20.664	20.435	0.093	0.095
speech45	1.688	21.715	21.423	0.082	0.085
speech46	1.669	21.368	21.127	0.085	0.088
speech47	1.690	20.987	20.751	0.089	0.092
speech48	2.037	20.135	19.892	0.099	0.101
speech49	1.564	23.264	22.941	0.069	0.071
speech50	1.790	20.538	20.303	0.094	0.097
speech51	1.782	20.995	20.727	0.089	0.092
speech52	1.944	18.984	18.787	0.112	0.115
speech53	1.780	22.831	22.496	0.072	0.075
speech54	1.826	19.962	19.761	0.101	0.103
speech55	2.056	20.754	20.464	0.092	0.095
speech56	1.808	22.073	21.798	0.079	0.081
speech57	1.874	22.240	21.914	0.077	0.080
speech58	1.624	21.967	21.756	0.080	0.082
speech59	1.649	23.853	23.551	0.064	0.066
speech60	1.704	21.242	21.002	0.087	0.089
ค่าเฉลี่ย	1.769	21.265	21.004	0.087	0.090

ตารางที่ ก.11 อัตราบีบอัดข้อมูล ค่า SNR และค่า NRMSE ของสัญญาณเสียงผู้ขายที่เวลา 60 วินาที ณ เส้นตัดสินใจที่ 1%

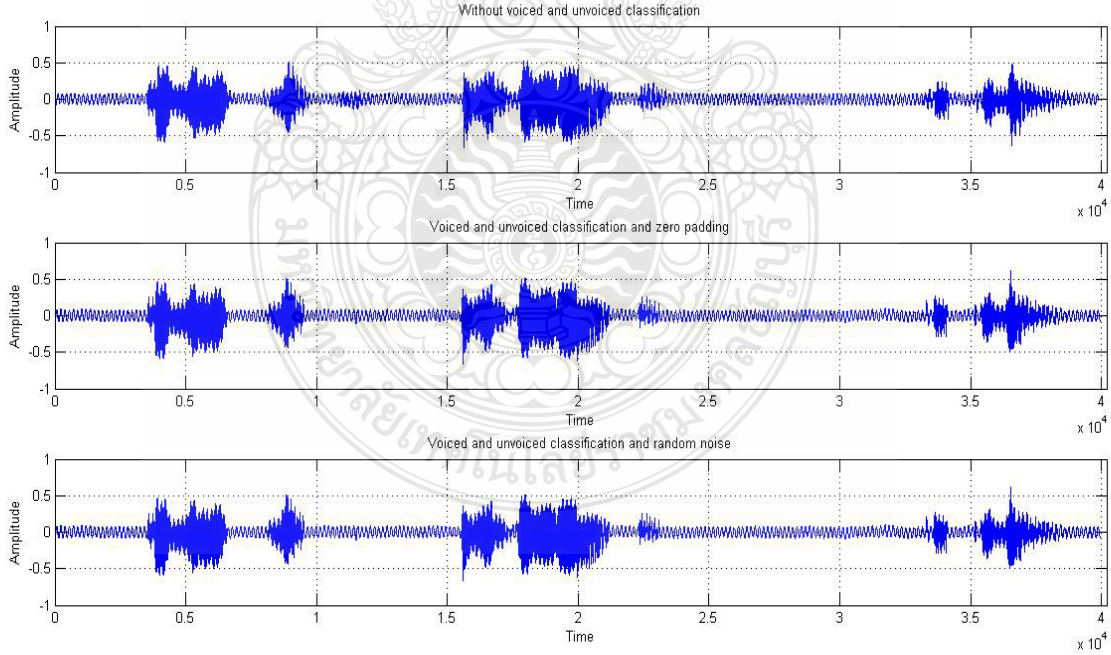
ชื่อไฟล์	อัตราบีบอัดข้อมูล	SNR		NRMSE	
		แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม	แทนที่สัญญาณรบกวนด้วยศูนย์	แทนที่สัญญาณรบกวนด้วยสัญญาณสุ่ม
speech61	1.983	20.715	19.876	0.092	0.101
speech62	2.680	17.779	16.816	0.129	0.144
speech63	1.889	19.077	18.322	0.111	0.121
speech64	2.344	18.389	17.527	0.120	0.133
speech65	2.254	15.919	15.113	0.160	0.176
speech66	1.818	21.080	20.277	0.088	0.097
speech67	2.186	19.206	18.410	0.110	0.120
speech68	2.556	17.801	16.762	0.129	0.145
speech69	2.621	16.245	15.178	0.154	0.174
speech70	2.195	20.684	19.707	0.092	0.103
speech71	2.587	17.775	16.334	0.129	0.153
speech72	2.187	21.107	20.203	0.088	0.098
speech73	2.003	21.238	20.311	0.087	0.097
speech74	2.261	20.728	19.815	0.092	0.102
speech75	2.021	19.859	18.815	0.102	0.115
speech76	1.915	21.205	20.337	0.087	0.096
speech77	2.418	17.211	16.393	0.138	0.152
speech78	1.739	22.989	22.180	0.071	0.078
speech79	2.601	18.566	17.619	0.118	0.132
speech80	2.418	18.636	17.751	0.117	0.130
ค่าเฉลี่ย	2.234	19.310	18.387	0.111	0.123

ภาคผนวก ข

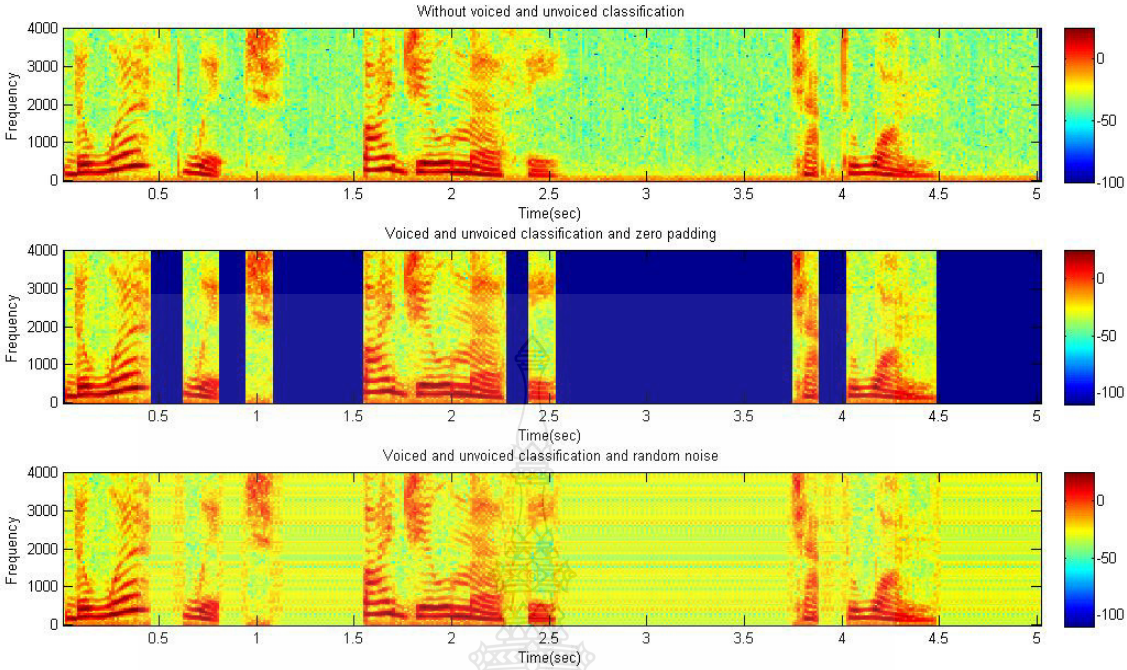
ตัวอย่างการบีบอัดสัญญาณเสียงพูด



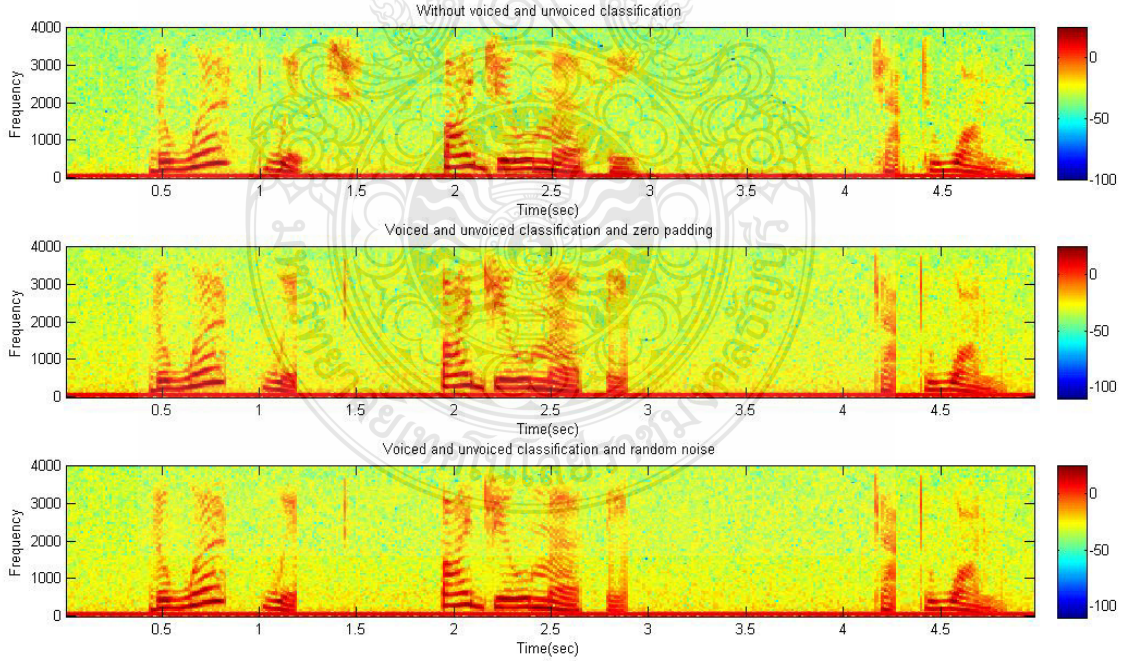
ภาพที่ ข.1 การบีบอัดสัญญาณเสียงพูดในโดเมนเวลา 5 วินาทีด้วยโปรแกรม MATLAB เสียงที่ 1



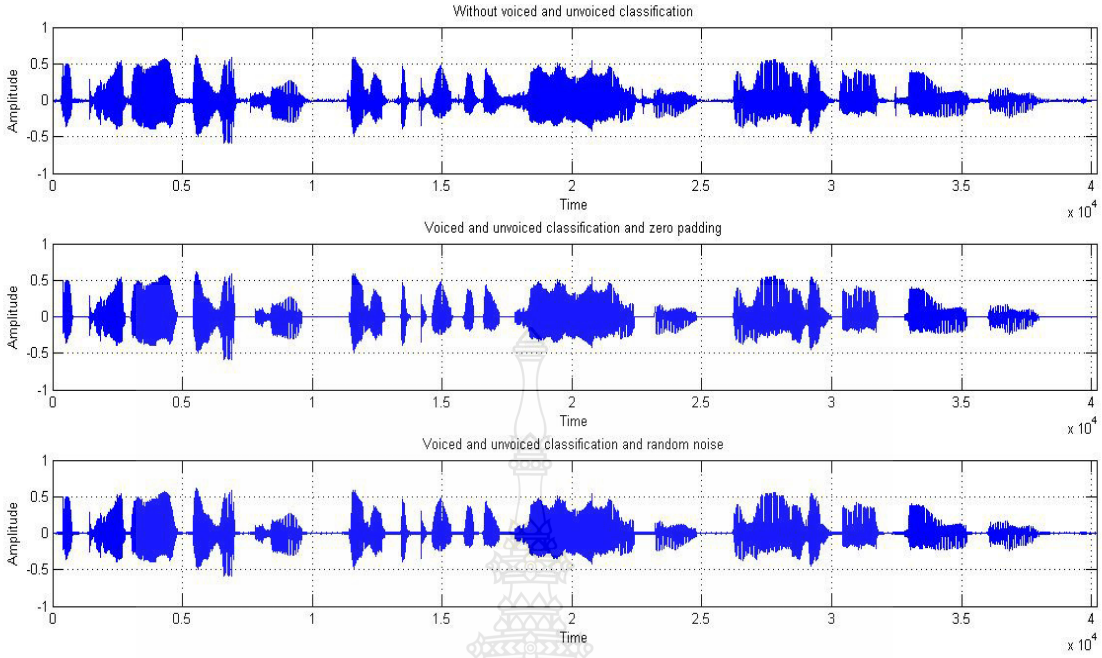
ภาพที่ ข.2 การบีบอัดสัญญาณเสียงพูดในโดเมนเวลา 5 วินาทีบนบอร์ดเอฟพีจีเอเสียงที่ 1



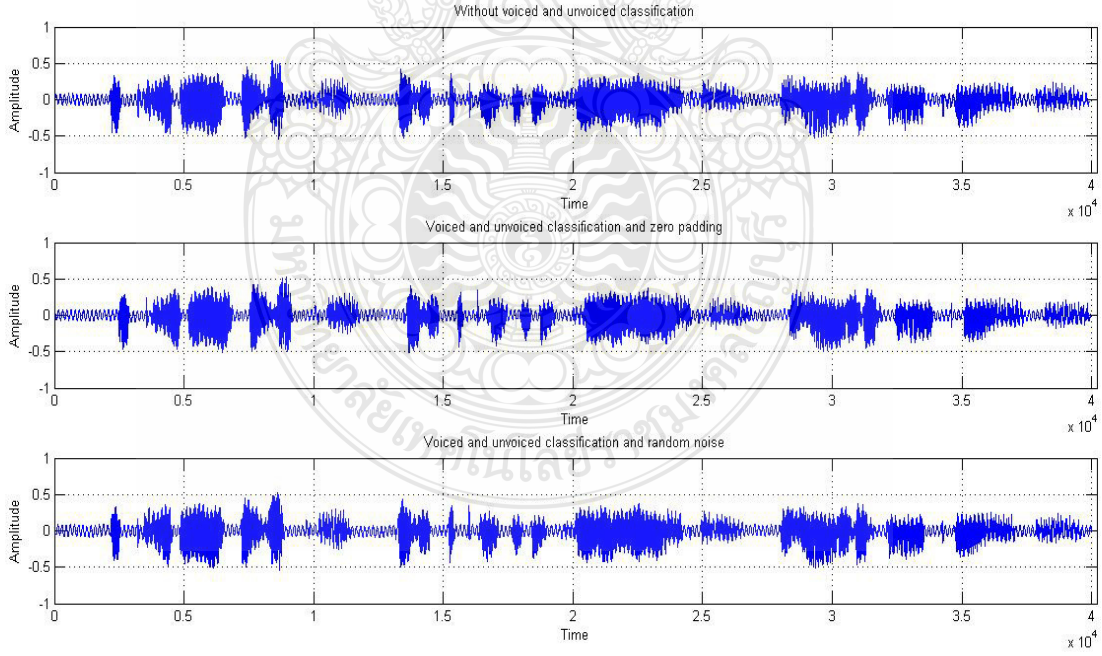
ภาพที่ ข.3 การบีบอัดสัญญาณเสียงพูดในสเปกโตรแกรม 5 วินาทีด้วยโปรแกรม MATLAB เสียงที่ 1



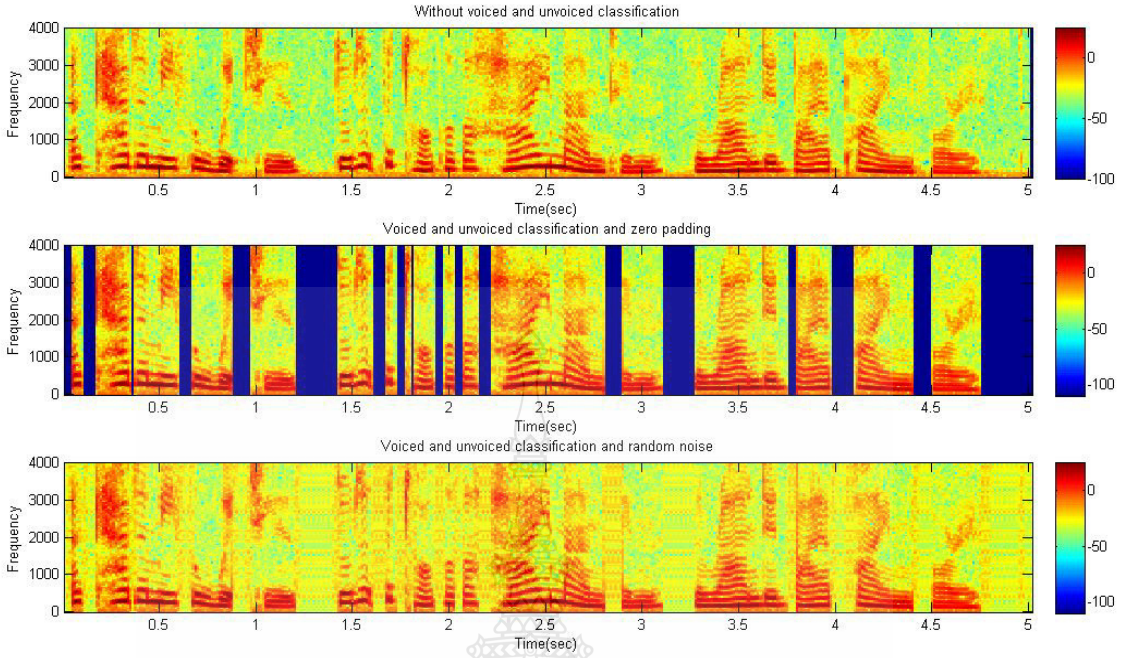
ภาพที่ ข.4 การบีบอัดสัญญาณเสียงพูดในสเปกโตรแกรม 5 วินาทีบนบอร์ดเอฟพีจีเอเสียงที่ 1



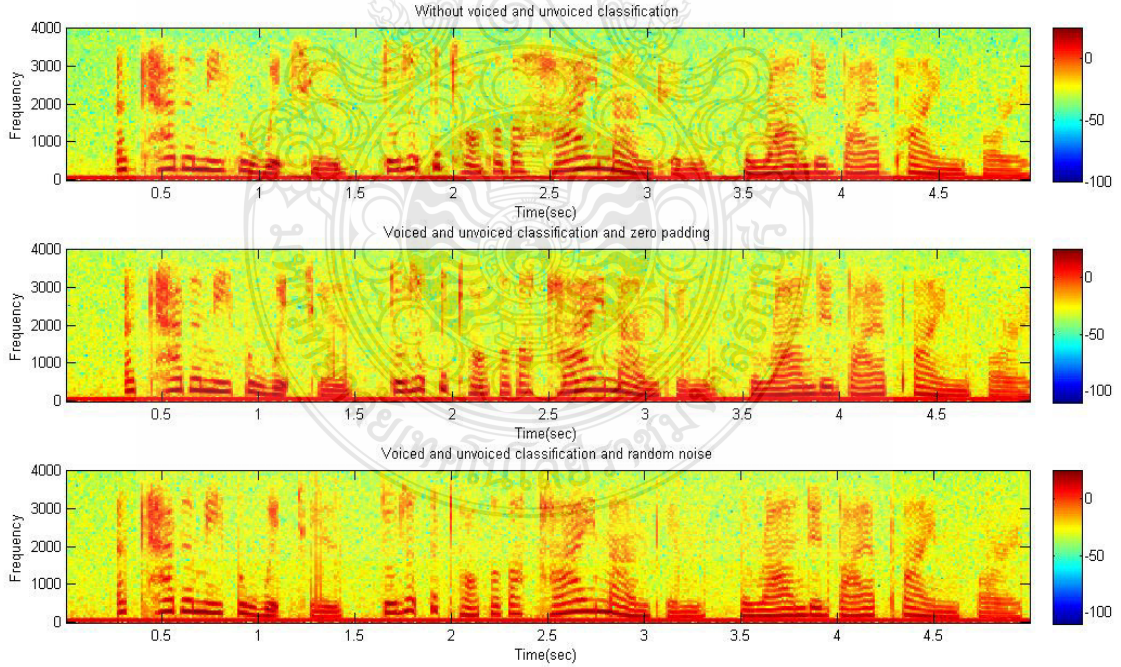
ภาพที่ ข.5 การบีบอัดสัญญาณเสียงพูดในโดเมนเวลา 5 วินาทีด้วยโปรแกรม MATLAB เสียงที่ 2



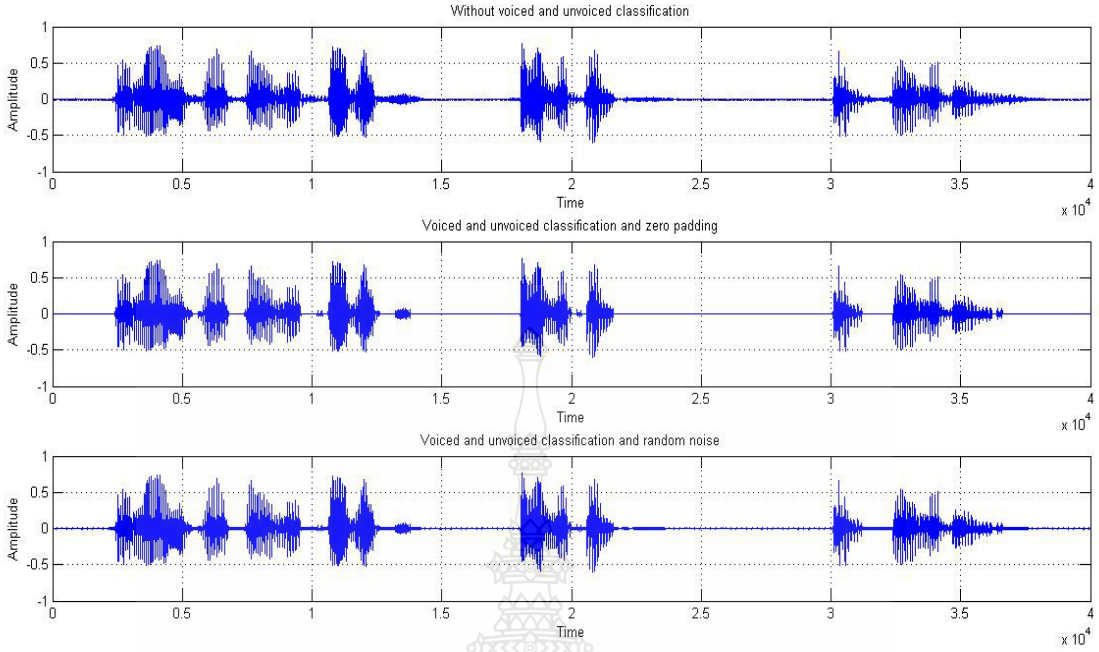
ภาพที่ ข.6 การบีบอัดสัญญาณเสียงพูดในโดเมนเวลา 5 วินาทีบนบอร์ดเอฟพีจีเอเสียงที่ 2



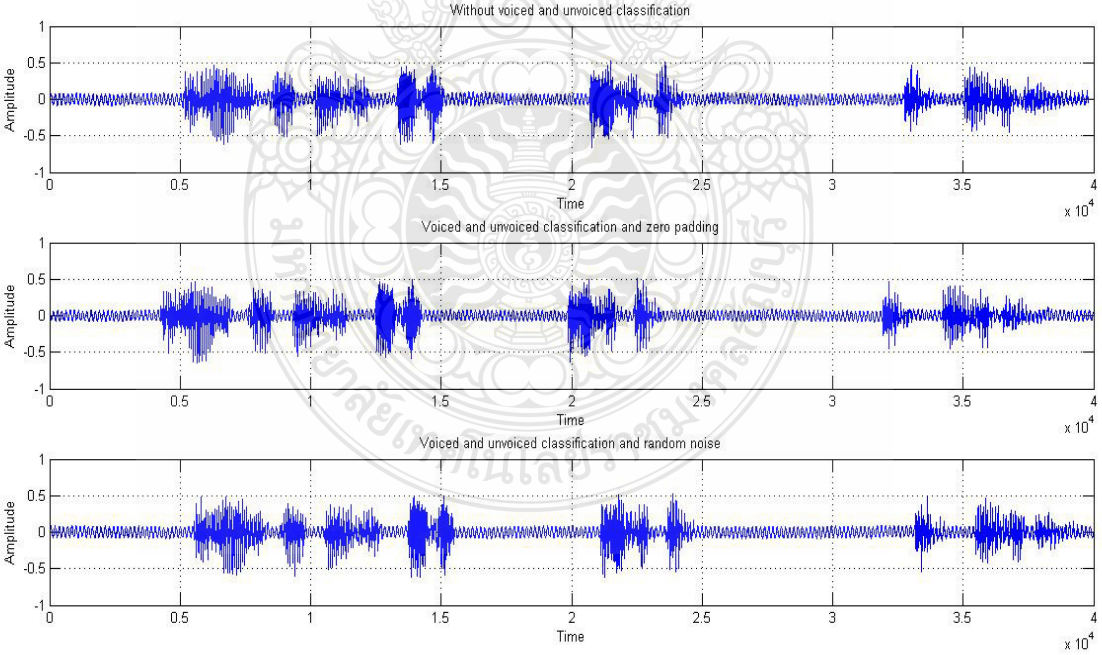
ภาพที่ ข.7 การบีบอัดสัญญาณเสียงพูดในสเปกโตรแกรม 5 วินาทีด้วยโปรแกรม MATLAB เสียงที่ 2



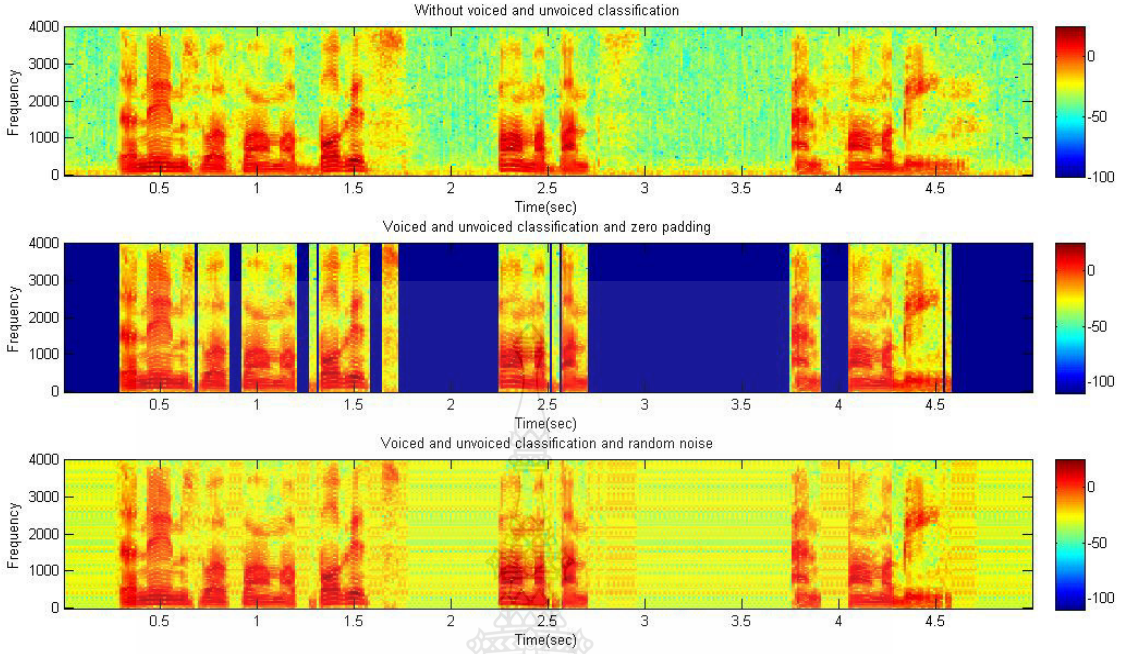
ภาพที่ ข.8 การบีบอัดสัญญาณเสียงพูดในสเปกโตรแกรม 5 วินาทีบนบอร์ดเอฟพีจีเอเสียงที่ 2



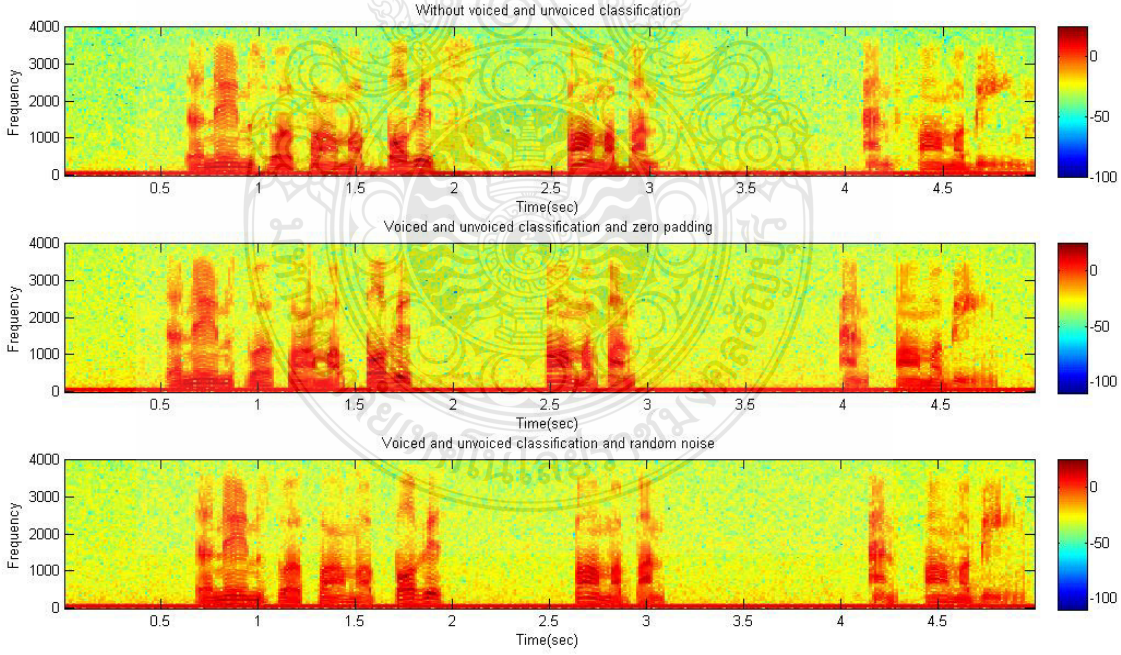
ภาพที่ ข.9 การบีบอัดสัญญาณเสียงพูดในโดเมนเวลา 5 วินาทีด้วยโปรแกรม MATLAB เสียงที่ 3



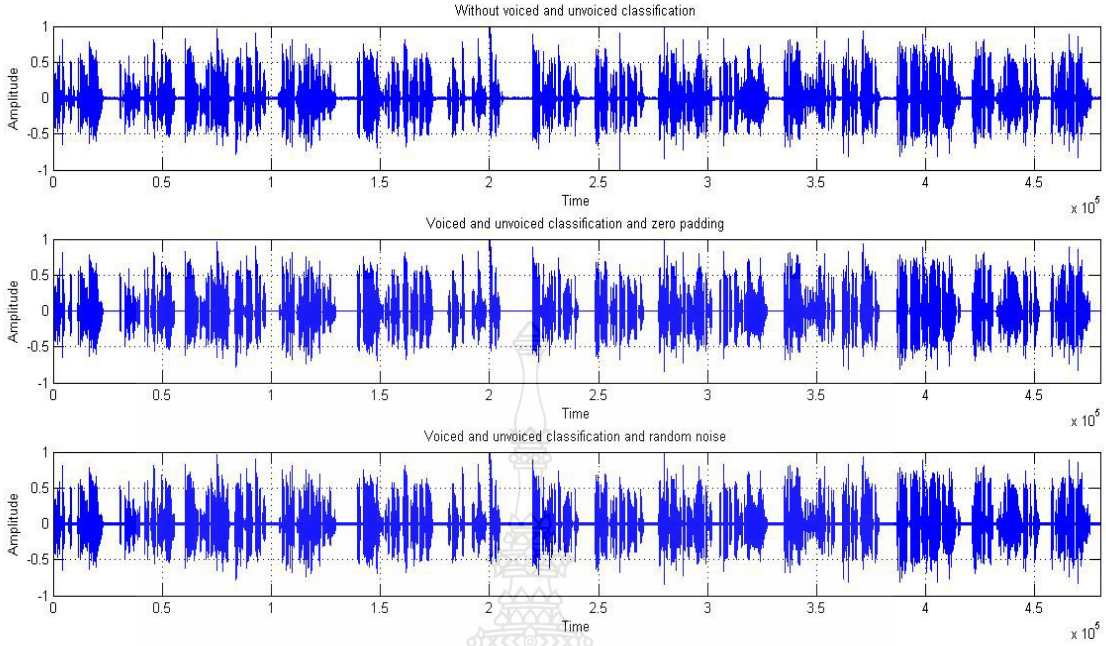
ภาพที่ ข.10 การบีบอัดสัญญาณเสียงพูดในโดเมนเวลา 5 วินาทีบนบอร์ดเอฟพีจีเอเสียงที่ 3



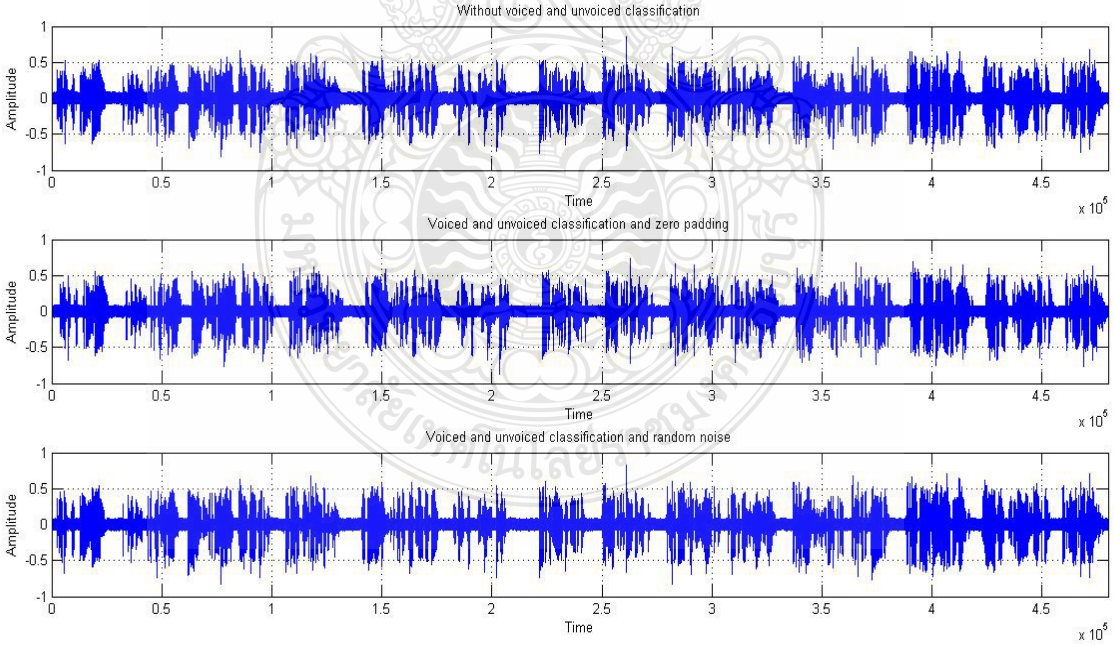
ภาพที่ ข.11 การบีบอัดสัญญาณเสียงพูดในสเปกโตรแกรม 5 วินาทีด้วยโปรแกรม MATLAB เสียงที่ 3



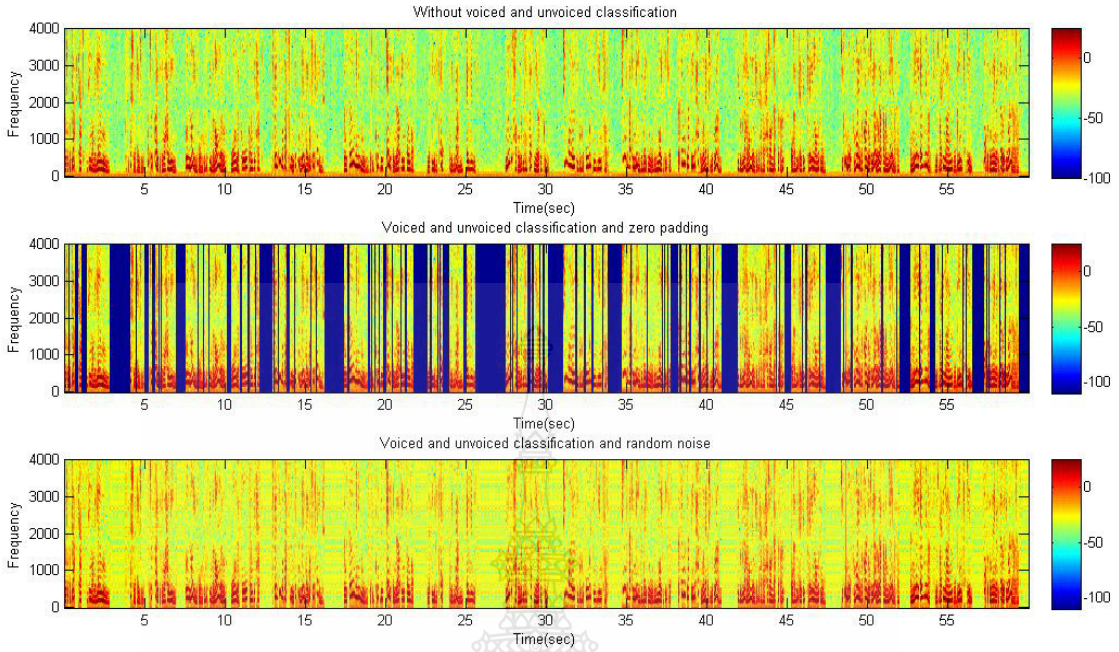
ภาพที่ ข.12 การบีบอัดสัญญาณเสียงพูดในสเปกโตรแกรม 5 วินาทีบนบอร์ดเอฟพีจีเอเสียงที่ 3



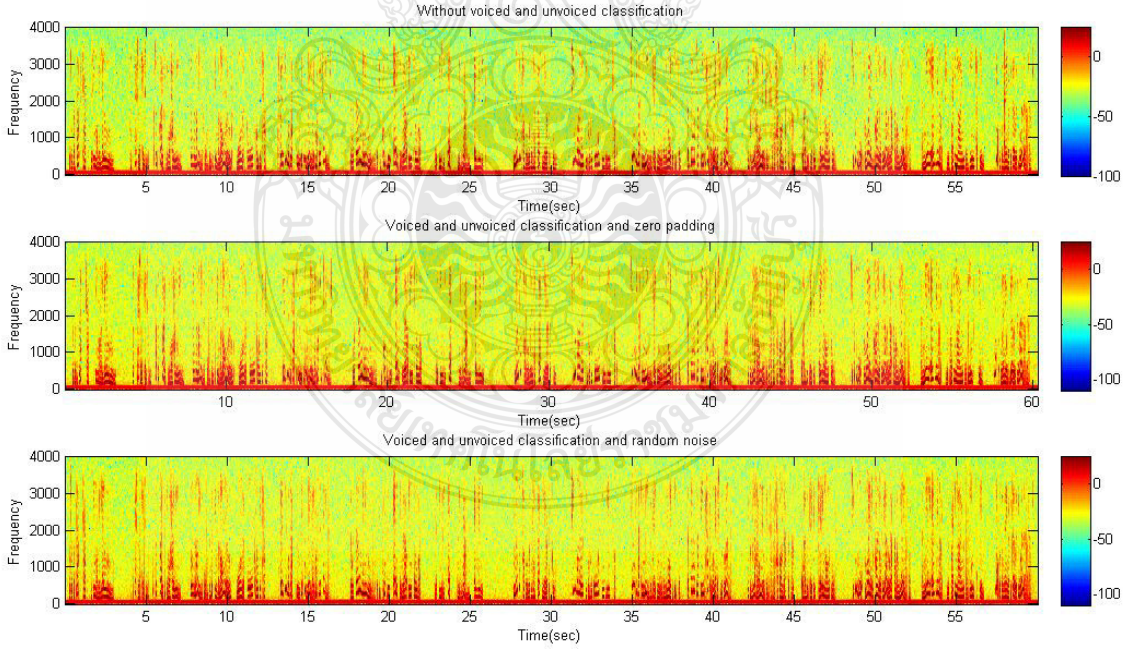
ภาพที่ ข.13 การบีบอัดสัญญาณเสียงพูดในโดเมนเวลา 60 วินาทีด้วยโปรแกรม MATLAB เสียงที่ 4



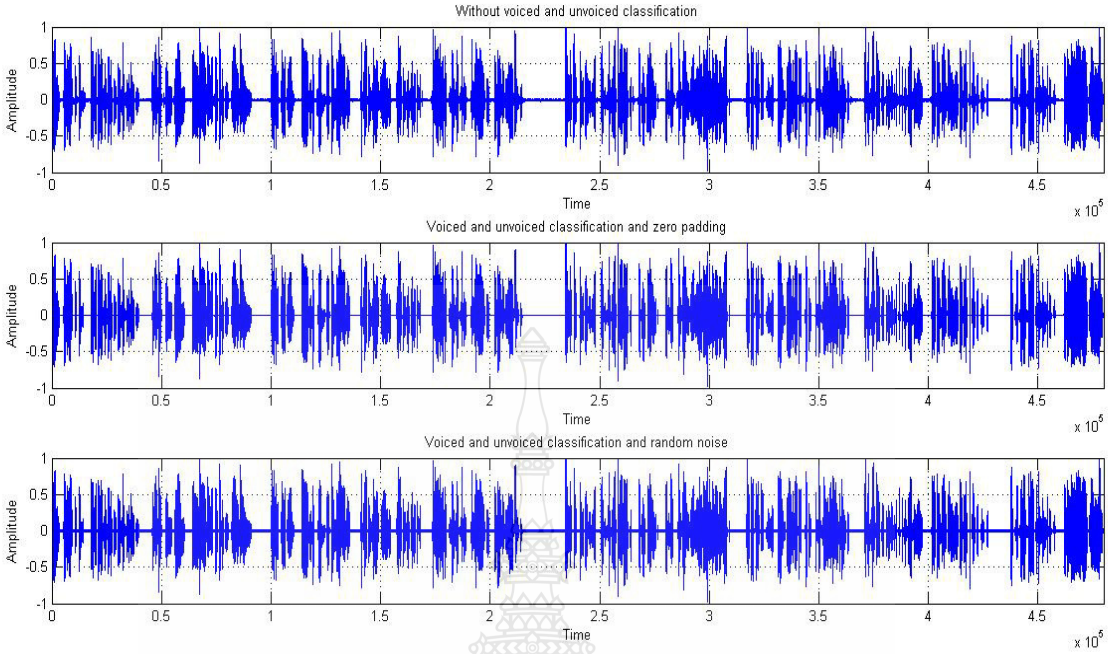
ภาพที่ ข.14 การบีบอัดสัญญาณเสียงพูดในโดเมนเวลา 60 วินาทีบนบอร์ดเอฟพีจีเอเสียงที่ 4



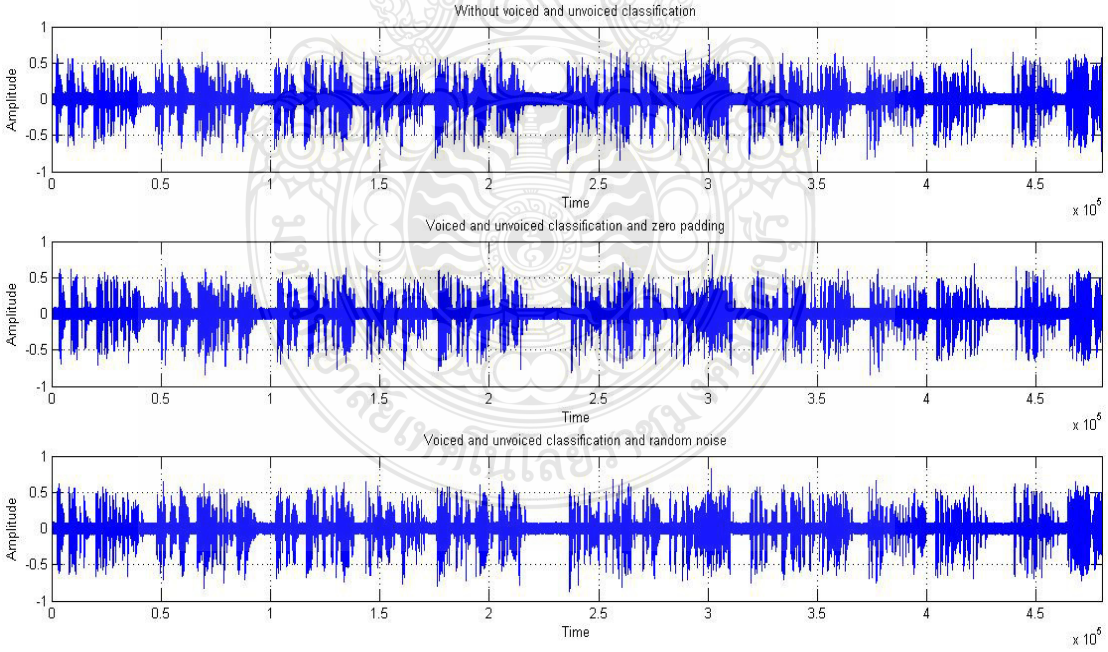
ภาพที่ ข.15 การบีบอัดสัญญาณเสียงพูดในสเปกโตรแกรม 60 วินาทีด้วยโปรแกรม MATLAB เสียงที่ 4



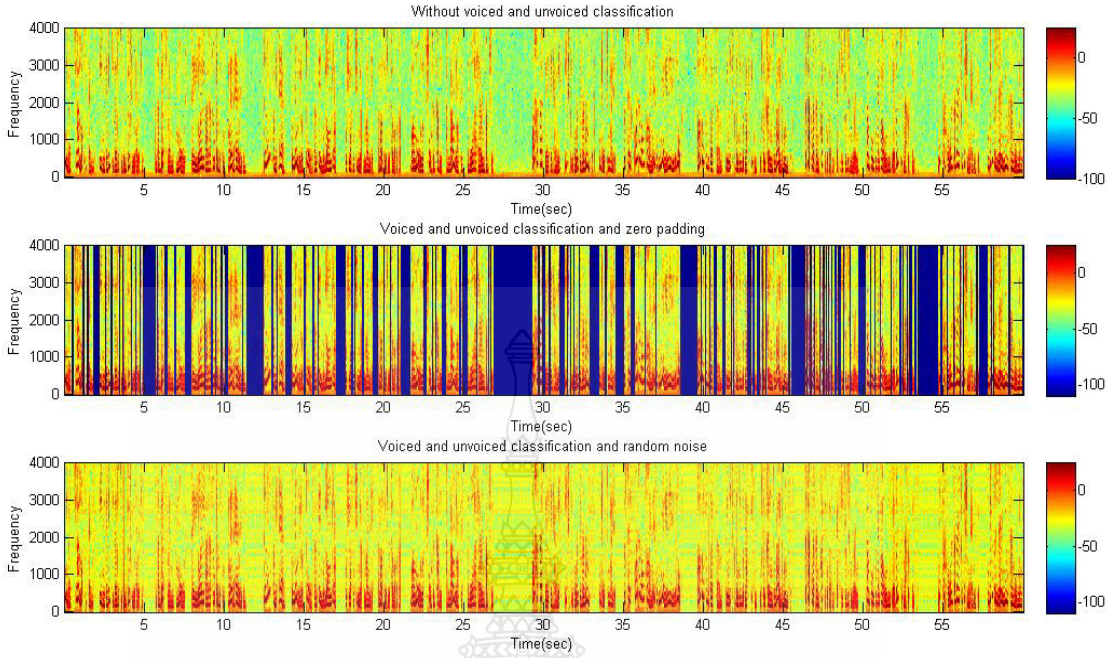
ภาพที่ ข.16 การบีบอัดสัญญาณเสียงพูดในสเปกโตรแกรม 60 วินาทีบนบอร์ดเอฟพีจีเอเสียงที่ 4



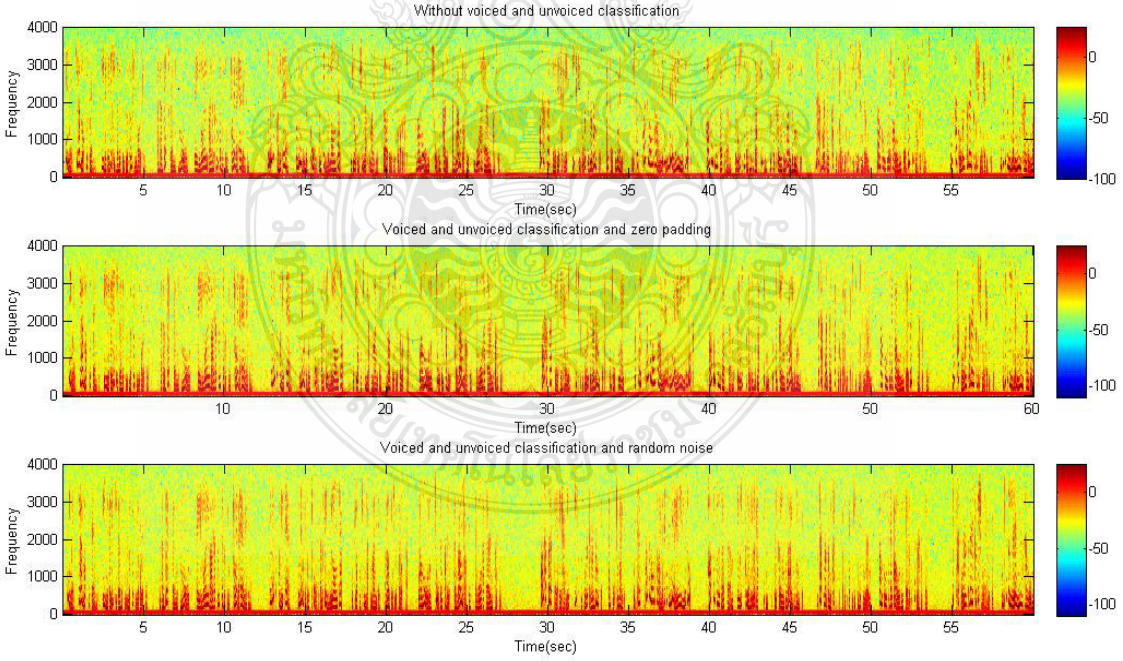
ภาพที่ ข.17 การบีบอัดสัญญาณเสียงพูดในโดเมนเวลา 60 วินาทีด้วยโปรแกรม MATLAB เสียงที่ 5



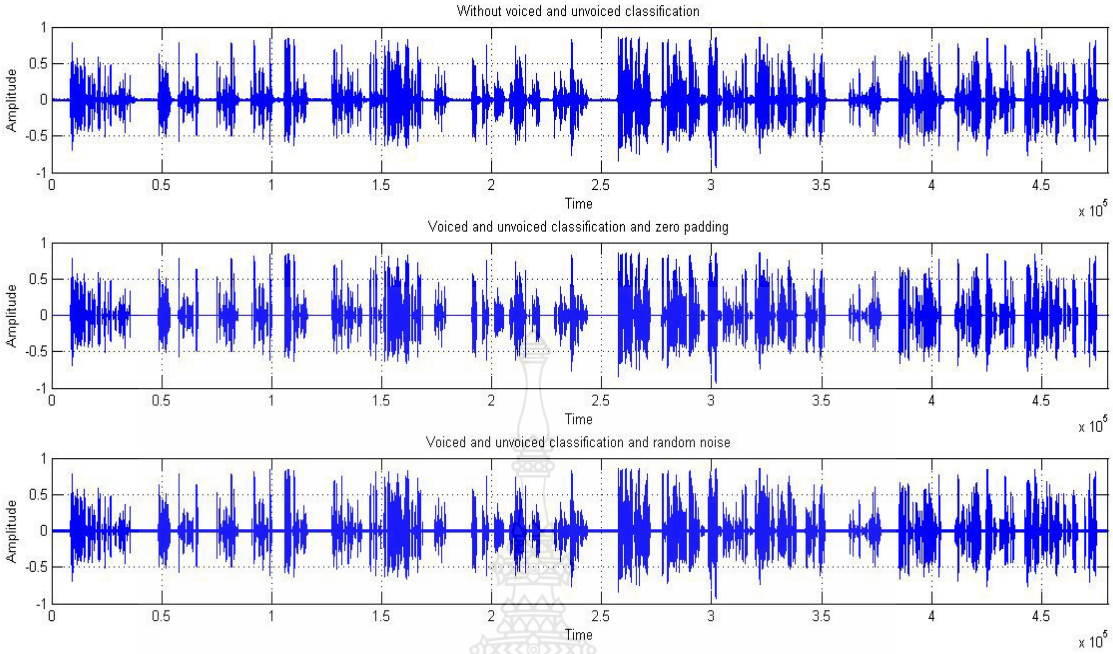
ภาพที่ ข.18 การบีบอัดสัญญาณเสียงพูดในโดเมนเวลา 60 วินาทีบนบอร์ดเอฟพีจีเอเสียงที่ 5



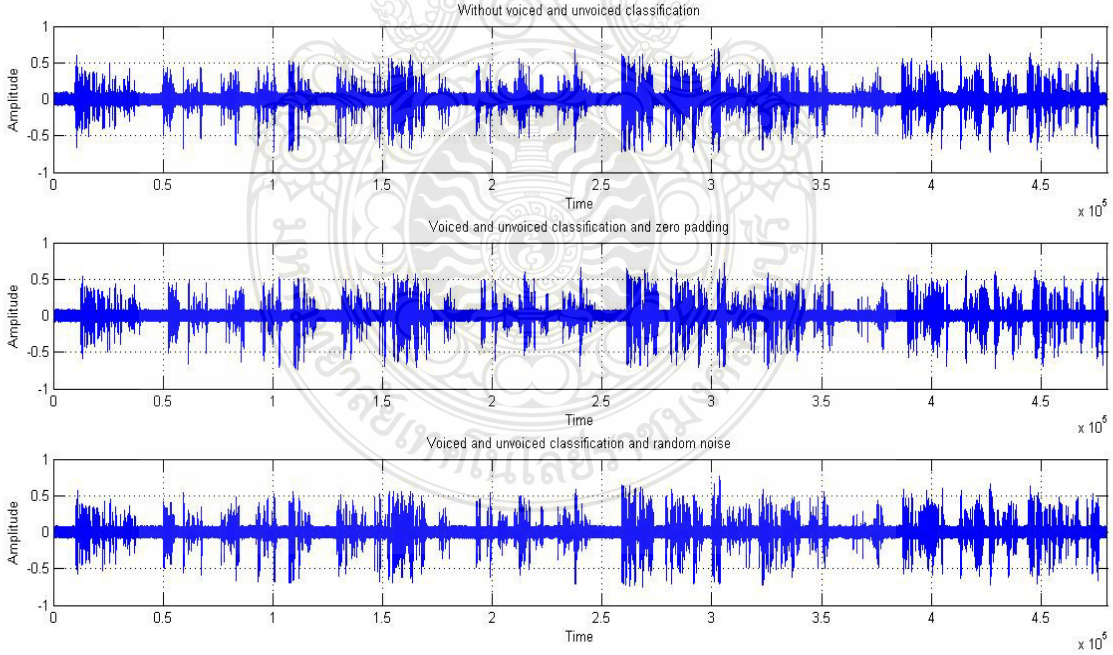
ภาพที่ ข.19 การบีบอัดสัญญาณเสียงพูดในสเปกโตรแกรม 60 วินาทีด้วยโปรแกรม MATLAB เสียงที่ 5



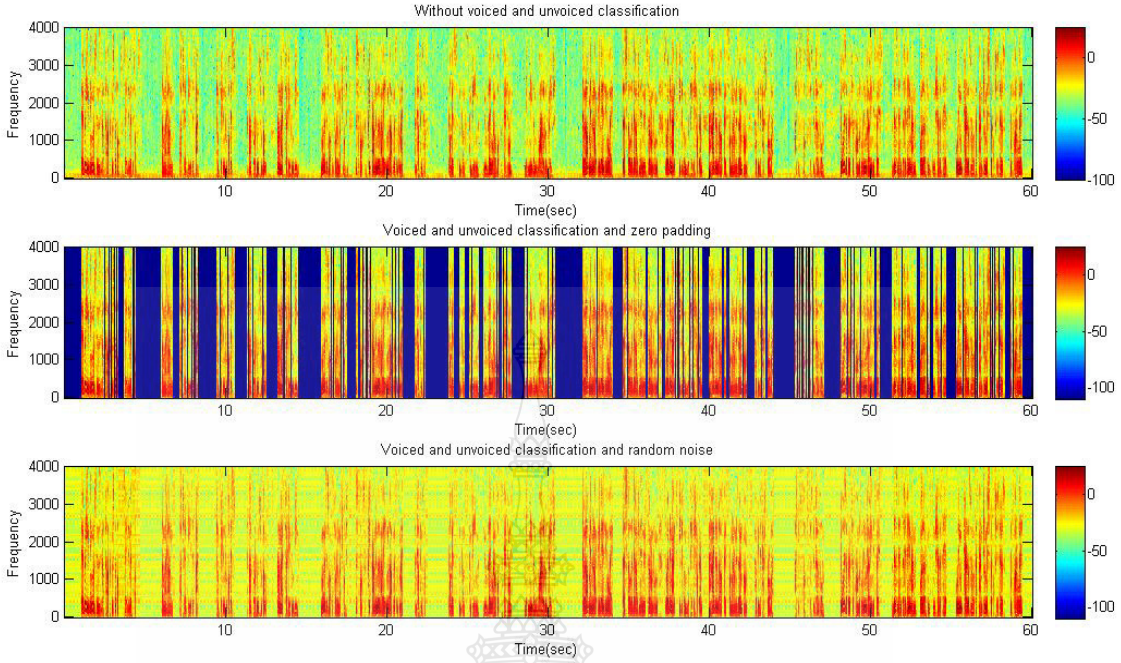
ภาพที่ ข.20 การบีบอัดสัญญาณเสียงพูดในสเปกโตรแกรม 60 วินาทีบนบอร์ดเอฟพีจีเอเสียงที่ 5



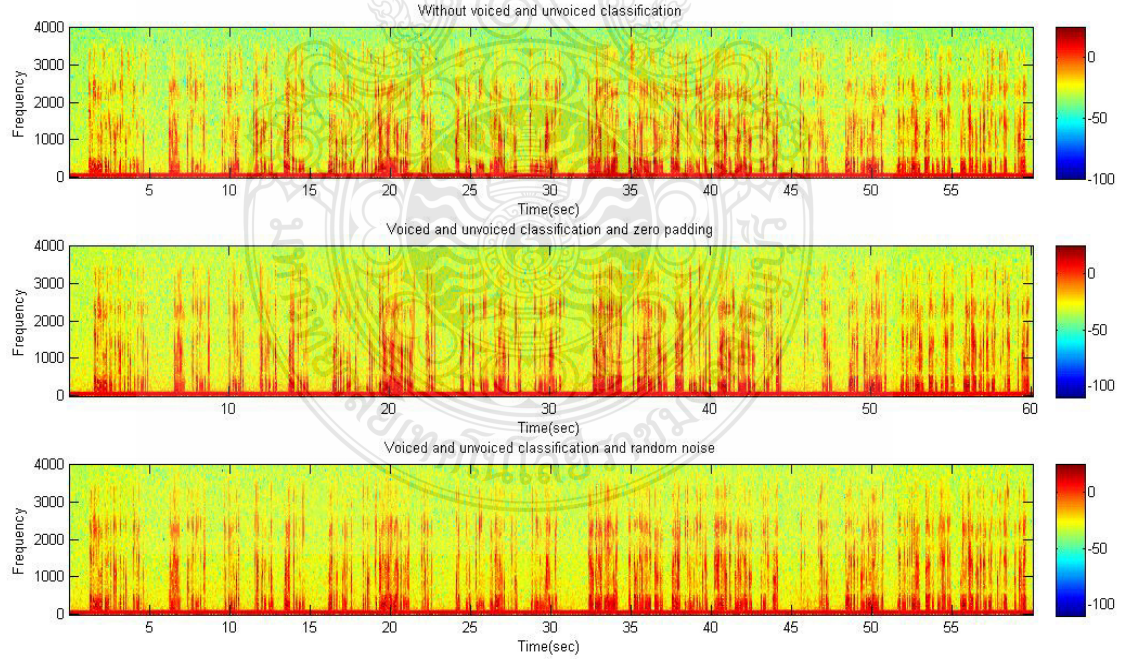
ภาพที่ ข.21 การบีบอัดสัญญาณเสียงพูดในโดเมนเวลา 60 วินาทีด้วยโปรแกรม MATLAB เสียงที่ 6



ภาพที่ ข.22 การบีบอัดสัญญาณเสียงพูดในโดเมนเวลา 60 วินาทีบนบอร์ดเอฟพีจีเอเสียงที่ 6



ภาพที่ ข.23 การบีบอัดสัญญาณเสียงพูดในสเปกโตรแกรม 60 วินาทีด้วยโปรแกรม MATLAB เสียงที่ 6



ภาพที่ ข.24 การบีบอัดสัญญาณเสียงพูดในสเปกโตรแกรม 60 วินาทีบนบอร์ดเอฟพีจีเอเสียงที่ 6

ภาคผนวก ค

ผลงานตีพิมพ์เผยแพร่

Electrical Engineering Network 2012

of Rajamangala University of Technology (EENET 2012)

CONFERENCE TOPICS

GROUP 1 (PE)
Power Electronics, Electric Machines, Motor Control and Drive, Measurement, Control and Robotics.

GROUP 2 (PW)
Power System, Transmission and Distribution, High Voltage and Electrical Energy, Generating Systems.

GROUP 3 (RE)
Renewable Energy, Energy Saving Technologies, Industry Specific Energy Conversion and Conditioning Technologies, Materials for Energy and Environment.

GROUP 4 (TE)
Telecommunication, Electronics, Information and Communication Technologies, Antennas, Microwave Theory and Techniques.

GROUP 5 (CP)
Computer Technologies and Network, Computer Graphics, Machine Learning and Human-Computer Interaction.

GROUP 6 (GN)
Education in Electrical Engineering, Simulation Software and Design tools, Related Topics in Electrical Engineering.

EENET 2012

GRAND PARADISE HOTEL
Nong Khai, THAILAND
April 3-5, 2012

รายชื่อผู้พิจารณาบทความการประชุมเครือข่ายวิชาการวิศวกรรมไฟฟ้า
มหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 4

รศ. ดร.วีระเชษฐ์ ขันเงิน	
รศ. ดร.วิจิตร กิณเรศ	
รศ. ดร.มนตรี ศิริปรัชญานันท์	
รศ. ดร.เวคิน ปิยะรัตน์	
รศ. ดร.เดชา พวงดาวเรือง	
รศ. ดร.วิบูลย์ ชื่นแขก	
รศ. ดร. เสถียร ธัญญศรีรัตน์	
ผศ. ดร.แนบบุญ หุนเจริญ	
ผศ. ดร.อภินันท์ อุรโสภณ	
ผศ. ดร.วรวัฒน์ เสงี่ยมวิบูล	
ผศ. ดร.อาทิตย์ โสตรโยม	
ผศ. ดร.เผด็จ เผ่าละออ	
ผศ. ดร.กองพล อารีรักษ์	
ดร.ยุทธนา ขำสุวรรณ์	
ดร.นิวัตร์ อังควิศิษฐพันธ์	
รศ.ดร.โกศล โอฬารไพโรจน์	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาภาคพายัพ เชียงใหม่
ดร.อุเทน คำน่าน	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาภาคพายัพ เชียงใหม่
นายณรงค์ นันทกุศล	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาภาคพายัพ เชียงใหม่
ผศ.กฤษดา ยิ่งขยัน	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาภาคพายัพ เชียงใหม่
นายเอกทัศน์ พฤษวรรณ	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาภาคพายัพ เชียงใหม่
ดร.จัตตุฤทธิ์ ทองปรอน	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาภาคพายัพ เชียงใหม่
ผศ.ชาญชัย เดชธรรมรงค์	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาภาคพายัพ เชียงใหม่
ดร.จักรกฤษณ์ เคลือบวัง	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เขตพื้นที่ตาก
นายสมนึก เครือสอน	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เขตพื้นที่ตาก
นายทัศนะ ถมทอง	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เขตพื้นที่ตาก
นายณรงค์ฤทธิ์ พิมพ์คำวงศ์	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เขตพื้นที่ตาก
ผศ.อภิศักดิ์ ขันแก้วหล้า	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เขตพื้นที่ตาก

ผศ.สุรสิทธิ์ แสนทอน	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เขตพื้นที่ตาก
นายเอกลักษณ์ สุมนพันธุ์	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เขตพื้นที่ตาก
นายปรีชา มหาไม้	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เขตพื้นที่ตาก
นายก่อเกียรติ อ๊อดทรัพย์	มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เขตพื้นที่ตาก
รศ.ดร.ธวัช เกิดชื่น	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
ดร.พินิจ ศรีธร	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
ผศ.ประเสริฐ เผื่อนหมื่นไวย	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
ผศ.พันธ์พงศ์ อภิชาตกุล	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
ผศ.สุทธินันท์ ต้นโพธิ์	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
ผศ.ศิริชัย ลาภาสระน้อย	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
ผศ.กฤตวิทย์ บัวใหญ่	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
ผศ.วุฒิชัย สง่างาม	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
นายกิตติวุฒิ จีนนะบุตร	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
นายรุ่งเพชร ก่องนอก	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
นายเอกจิต คุ้มวงศ์	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
นางอุษา คงเมือง	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
นายชิติสรรค์ วิชิโต	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
ดร.วรรณรีย์ วงศ์ไตรรัตน์	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา
นายบุญช่วย เจริญผล	มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพจ.กรุงเทพฯ
นายวุฒิวัฒน์ คงรัตนประเสริฐ	มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพจ.กรุงเทพฯ
นายชูศักยฐ กมลขันติธร	มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพจ.กรุงเทพฯ
นายวินัย เมธาวิทิต	มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพจ.กรุงเทพฯ
นายชาญฤทธิ์ ธาราสันติสุข	มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพจ.กรุงเทพฯ
นายประหยัด กองสุข	มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก จ.จันทบุรี
นายภัทรพงศ์ อัญชันภาติ	มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก จ.จันทบุรี
นายสมพล โคศรี	มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก จ.จันทบุรี
นายทัศพันธ์ สุวรรณทัต	มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก จ.จันทบุรี
นายจตุรงค์ จตุรเชิดชัยสกุล	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
นายพูนศรี วรรณการ	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
ดร.ณัฐพงศ์ พันธุนะ	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ

ดร.นัฐโชติ รักไทยเจริญชีพ	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
นายนิติพันธ์ คุณประเสริฐ	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
นายธนารัตน์ ตันมณีประเสริฐ	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
นายศุภวุฒิ เนตรโพธิ์แก้ว	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
นายทง ลานธารทอง	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
นายสมเกียรติ ทองแก้ว	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
ผศ.พิชญ ดาราพงษ์	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
นายมนัส บุญเทียรทอง	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
นายพนา ดุสิตากร	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
ผศ.กิจจา ลักษณ์อำนวยการ	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
ผศ.โกศล นิธิโสภา	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
ผศ.จรินทร์ จุลวานิช	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
นายนิลมิต นิลาศ	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
นายเกรียงไกร เหลืองอำพล	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
นายวณพันธ์ วัยวุฒิ	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
นายณัฏฐ์วรินทร์ ทองรักษ์	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
ผศ.ศรีศักดิ์ น้อยไร่ภูมิ	มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร จ.กรุงเทพฯ
ผศ.ดร.ประมุข อุณหเลขกะ	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.นนทบุรี
นางสาวพัชรนันท์ ศรีธนาอุทัยกร	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.นนทบุรี
ผศ.สรายุธ ทองกุลภัทร์	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.นนทบุรี
ผศ.วารุณี ศรีสงคราม	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.นนทบุรี
ดร.ยุทธนา กันทะพะเยา	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.นนทบุรี
ผศ.เฉลิมพล เรืองพัฒนาวิวัฒน์	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.นนทบุรี
ผศ.กระจ่าง พิทักษ์วงศ์วิทยา	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.นนทบุรี
ดร.ศรีสุดา ไชยทองสุก	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.นนทบุรี
รศ.นภัทร วัจนเทพินทร์	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.นนทบุรี
รศ.สมพันธ์ อำพาวัน	มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย จ.สงขลา
ดร.ชูวงศ์ วัฒนศักดิ์ภูบาล	มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย จ.สงขลา
ผศ.วิสุทธิ์ พงศ์พฤกษธาตุ	มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย จ.สงขลา
ดร.สุริยา แก้วอาษา	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร

ดร.นิธิโรจน์ พรสุวรรณเจริญ	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
ดร.เมธา ทัศคร	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
ดร.นำพน พิพัฒน์ไพบูลย์	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
ผศ.วิชัย ครองกิจศิริ	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
นายเอกวิทย์ หายักวงษ์	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
นายวีระ ธันยาภิรักษ์	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
นายรัก สกุลพงศ์	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
นายนครินทร์ ศรีปัญญา	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
นายกฤตยา สมสัย	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
นายจงเจริญ คุ้มบุญ	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
นายกฤษฎา บุญมีวิเศษ	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
นายจีระพงศ์ ศรีวิชัย	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
นายวีระชัย จรบุรมย์	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสกลนคร
นายปฏิวัติ บุญมา	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.สุพรรณบุรี
นายเฉลียว เกตุแก้ว	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.สุพรรณบุรี
ผศ.ดร.ปรีชา สาคะรังค์	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.สุพรรณบุรี
ดร.ภัควัฒน์ จันทร์ตรี	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.สุพรรณบุรี
นายไพบูลย์ เกียรติสุขคณธร	มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ จ.สุพรรณบุรี
รศ.พันธ์ พิริยะวรรธน์	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตขอนแก่น
ผศ.ประวิช เปรียบเหมือน	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตขอนแก่น
ผศ.ดร.ศักดิ์ระวี ระวีกุล	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตขอนแก่น
ผศ.ประยงค์ เสาร์แก้ว	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตขอนแก่น
ผศ.จักรวัฒน์ บุตรบุญชู	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตขอนแก่น
ดร.สุระ ตันดี	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตขอนแก่น
ดร.ณรงค์ สีหาจ่อง	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตขอนแก่น
ผศ.ดร.สมชัย หิรัญวโรดม	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ผศ.ดร.วันชัย ทรัพย์สิงห์	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ผศ.วิชัย ผดุงศิลป์	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ผศ.ศิริชัย แดงเอม	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ผศ.ณัฐวุฒิ โสมะเกษตรินทร์	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

ดร.ฉัตรชัย	ศุภพิทักษ์สกุล	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ดร.บุญยัง	ปลั่งกลาง	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ดร.สุรินทร์	แหงมงาม	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ดร.กฤษณ์ชนม์	ภูมิกิตติพชญ์	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ดร.ณฐภัทร	พันธ์คง	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
นายพินิจ	จิตจริง	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
นายสมชาย	เบียนสูงเนิน	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
นายพร้อมศักดิ์อภิรติกุล		มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
นายนิติพงศ์	ปานกลาง	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
นายธีระพล	เหมือนขาว	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
นายณัฐพล	หาอุปละ	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ผศ.จินตนา	นาคะสุวรรณ	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ผศ.ปราชญ์	ดาบบัง	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ดร.จักรี ศรีนนท์ฉัตร		มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ดร.อำนวย	เรืองอำนวย	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ดร.ไพฑูรย์	รักเหลือ	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
นายวิโรจน์	พิราจเนนชัย	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
นายพงษ์ศักดิ์ อำภา		มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
นายบุญยิ่ง	นบนอบ	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
นายอภิรดา	นามแสง	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ดร.สุทินัน	พรอนุรักษ์	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ดร.วินัย วิชัยพาณิชย์		มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ดร.กิตติวัณณ์ นิ่มเกิดผล		มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ผศ.อิฐอารัญ	ปิติมล	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
นายณัชติพงศ์	อูทอง	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

TE21	Current Controlled Voltage-mode Universal Filter with High Input Impedance	553
TE22	การวิเคราะห์ประสิทธิภาพการแฮนด์ออฟแบบมัลติลิงค์สำหรับเครือข่าย NEMO และ 3G	557
TE23	สายอากาศโมโนโพลรูปตัววีที่มีการปรับจูนระนาบกราวด์สำหรับประยุกต์ใช้งานย่านความถี่กว้าง	561
TE24	วงจรกรองแถบความถี่ผ่านแบบขั้นบันไดลำดับที่ 2 โดยใช้ CCTA	565
TE25	วงจรกรองผ่านแถบแบบไม่สมมาตรโดยรูปแบบสายป้อนสัญญาณอินเตอร์ดิจิตอล	569
TE26	High Output-impedance Current-mode Universal Filter Using CCCIIs	573
TE27	Research of Adaptive Channel Estimation in MIMO-OFDM Systems	577
TE28	วงจรกำลังสองโหมดกระแสที่ควบคุมได้ด้วยวิธีทางอิเล็กทรอนิกส์และอิสระทางอุณหภูมิโดยใช้ MO-CFTA เพียงหนึ่งตัว	581
TE29	การพัฒนาเทคนิคการประมวลผลสัญญาณเสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ	585
TE30	การวัดการแพร่กระจายของสนามแม่เหล็กจากอาร์เอฟไอดีความถี่ 13.56MHz โดยใช้ตัวตรวจจับแบบอาร์เรย์	589
TE31	สายอากาศแพตช์รูปก้ามปูความถี่แถบคู่ที่ป้อนด้วยสายนำสัญญาณระนาบร่วม	593
TE32	โทรศัพท์พื้นฐานโดยใช้เทคโนโลยี RFID	597
TE33	วงจรกำเนิดสัญญาณฐานเวลานาฬิกาดิจิตอลความเที่ยงตรงสูง	601
TE34	Current-mode Universal Filter with Independent Control of Pole frequency and Quality factor	606
TE35	Current-mode Multiphase Sinusoidal Oscillators Using CFTAs	610
TE36	วงจรกรองผ่านแถบความถี่ไมโครสตริปรูปสี่เหลี่ยมแบบมีการคัปปลิงสำหรับใช้งานย่านความถี่แถบกว้าง	614
TE38	วงจรตรวจรู้ภาพโดยใช้คลื่นไมโครเวฟ	622

การประชุมวิชาการเครือข่ายวิศวกรรมไฟฟ้า มหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 4 วันที่ 3-5 เมษายน 2555

การพัฒนาเทคนิคการประมวลผลสัญญาณเสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ

Improvement Speech Signal Real Time Processing Techniques Using FPGA Board

เฉลิมเกียรติ สุดาชา และจักรี ศรีนนท์ฉัตร
สาขาวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
39 ม.1 ต.คลองหก อ.ธัญบุรี จ.ปทุมธานี 12110 โทรศัพท์: 02-549-3588 E-mail: pickup2530@hotmail.com

บทคัดย่อ

บทความนี้นำเสนอการพัฒนาเทคนิคสำหรับการประมวลผลสัญญาณเสียงพูดบนบอร์ดเอฟพีจีเอ (Field Programmable Gate Arrays: FPGA) บนฐานเวลาจริง สัญญาณเสียงพูดถูกคำนวนในรูปแบบตัวเลขทศนิยมแบบ Fixed-Point สัญญาณเสียงพูดอินพุตจะถูกประมวลผลที่ 2^n บิตโดยอาศัยคุณสมบัติเฉพาะตัวของหน่วยความจำและหน่วยประมวลผลบนบอร์ด งานวิจัยนี้การออกแบบจะใช้ภาษาเอชดีแอล (VHSIC Hardware Description Language : VHDL) บนบอร์ด FPGA Discovery - III XC3S200F4 ผลการทดลองพบว่าเทคนิคนี้สามารถทำงานได้บนระบบฐานเวลาจริงได้

คำสำคัญ: การประมวลผลสัญญาณเสียงพูด, เอฟพีจีเอ, การประมวลผลบนฐานเวลาจริง

Abstract

This article presents the development of speech processing technique in Field Programmable Gate Arrays (FPGA) board based on real time system. Speech signal is computed in the Fixed-Point format. The input signal is processed in 2^n bits which used the characteristic of ram and central processing unit on broad for implementation. This research uses the VHSIC Hardware Description Language to design and operate the FPGA Discovery - III XC3S200F4 broad. The results show that this technique can be operated in real time system

Keywords: speech processing, FPGA, real time processing system

1. บทนำ

การประมวลผลของสัญญาณเสียง [1] นั้นเป็นส่วนหนึ่งของการประยุกต์ใช้การประมวลผลสัญญาณทางด้านดิจิตอลเข้ากับสัญญาณเสียง โดยที่เสียงพูด (Speech) คือเสียงที่มนุษย์เปล่งออกมาเพื่อใช้ติดต่อสื่อสารกันซึ่งเสียงนั้นจะประกอบไปด้วยสัญญาณเสียงและสัญญาณรบกวนหรือสัญญาณที่ไม่ใช่เสียง ซึ่งในธรรมชาติสัญญาณเสียงจะเป็นสัญญาณในลักษณะที่เป็นสัญญาณต่อเนื่อง (Continues Signal) แต่มีรูปแบบที่ไม่ซ้ำกัน (Non-Stationary Signal) [2] และมีคุณสมบัติต่างกัน เช่น สมบัติทางการกระจายน้ำหนักของพลังงานเสียง (Weight Distribution) และการกระจายของความถี่ (Frequency Distribution) [3] เป็นต้น ทั้งนี้สัญญาณถูกนำมาทำการวิจัยและประยุกต์ใช้ในหลายด้าน เช่น การบีบอัดสัญญาณเสียง (Speech Compression) [4, 5] การรู้จำสัญญาณเสียง (Speech Recognition) การแยกแยะผู้พูด (Speaker Identification) รวมถึงการสังเคราะห์เสียง (Speech Synthesis) เป็นต้น เทคโนโลยีเอฟพีจีเอเป็นอุปกรณ์สารกึ่งตัวนำชนิดโปรแกรมได้ โดยนำไปโปรแกรมวงจรที่ได้จำลองการทำงานลงบนชิพเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเฮชดีแอลลงบนอุปกรณ์เอฟพีจีเอ ซึ่งบอร์ดเอฟพีจีเอเป็นทางเลือกที่มีราคาไม่แพง ความถี่สัญญาณนาฬิกาสูง มีฟังก์ชันการทำงานตามที่ผู้ออกแบบต้องการ นำไปประยุกต์ใช้งานพร้อมยังสามารถปรับปรุงและแก้ไขชุดคำสั่งได้ง่าย [6]

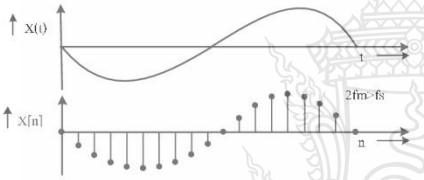
ในบทความนี้นำเสนอการพัฒนาเทคนิคการประมวลผลสำหรับสัญญาณเสียงพูดบนบอร์ดเอฟพีจีเอบนฐานเวลาจริง การประมวลผลสัญญาณเสียงพูดรูปแบบตัวเลขทศนิยมแบบ Fixed-Point ทำการทดสอบโดยจำลองการทำงานลงบนชิพเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเฮชดีแอล ทำการทดสอบผลบนคอมพิวเตอร์ พร้อมทั้งสามารถสังเคราะห์เป็นวงจรลอจิคไปที่ตัวชิพเอฟพีจีเอบนบอร์ด FPGA Discovery - III XC3S200F4

2. การประมวลผลสัญญาณเสียงบนชิพเอฟพีจีเอ

รูปที่ 1 บล็อกไดอะแกรมการทำงานของโปรแกรมบนชิพเอฟพีจีเอ

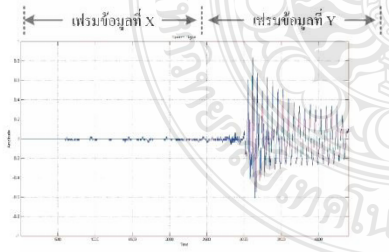
การประมวลผลสัญญาณเสียงพูดบนบอร์ดเอฟพีจีเอนั้นเป็นการออกแบบการทำงานชิพเอฟพีจีเอสำหรับประมวลผลสัญญาณเสียงพูดในรูปแบบที่เป็นข้อมูลดิจิตอลด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเฮชดีแอลลงไปเพื่อให้ตัวอุปกรณ์เอฟพีจีเอ จากรูปที่ 1 บล็อกไดอะแกรมการทำงานของโปรแกรมบนชิพเอฟพีจีเอ ในขั้นตอนแรกจะรับข้อมูลสัญญาณเสียงมาเก็บไว้ที่หน่วยความจำภายในเอฟพีจีประมวลผลในการแบ่งข้อมูลสัญญาณเสียงตามที่ออกแบบและเมื่อทำการแบ่งข้อมูลเสร็จแล้ว จะทำการประมวลผลคืนข้อมูลสัญญาณเสียงพูด

การนำสัญญาณเสียงแอนะลอกมาจัดเก็บในรูปแบบที่เป็นดิจิตอลเพื่อการนำไปการประมวลผลสัญญาณเสียงโดยกรรมวิธีทางดิจิตอล เราจำเป็นต้องมีการสุ่มเอาค่าจากสัญญาณแอนะลอกมาจัดเก็บเป็นดิจิตอล จะได้สัญญาณที่เรียกว่าสัญญาณไม่ต่อเนื่องทางเวลา (Discrete-Time Signal) และคุณภาพของสัญญาณไม่ต่อเนื่องทางเวลากำหนดด้วยอัตราการสุ่มสัญญาณ (Sampling Rate) ดังรูปที่ 2 แต่ในทางปฏิบัติการสุ่มสัญญาณจะต้องมีค่ามากกว่าหรือเท่ากับสองเท่าของความถี่ที่ทำการสุ่มสัญญาณ



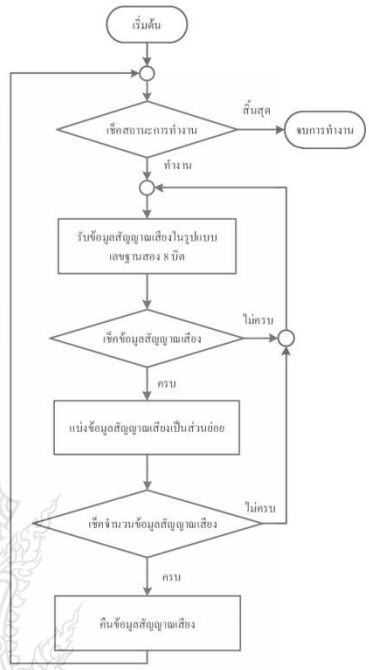
รูปที่ 2 การนำสัญญาณเสียงแอนะลอกมาจัดเก็บในรูปแบบดิจิตอล

ในการประมวลผลสัญญาณเสียงบนชิพเอฟพีจีเอเสียงจะถูกแบ่งออกเป็นเฟรมตามความเหมาะสมกับงานที่ใช้การประมวลผลการแบ่งสัญญาณเสียงออกเป็นเฟรมย่อยดังรูปที่ 3 ก่อนนำมาวิเคราะห์ที่ละส่วนจะทำให้วิเคราะห์ง่ายขึ้นและลดความซับซ้อนในการประมวลผลสัญญาณ



รูปที่ 3 การแบ่งสัญญาณเสียงออกเป็นเฟรมย่อย

2.1 การออกแบบโปรแกรมและการทำงานบนเอฟพีจีเอ

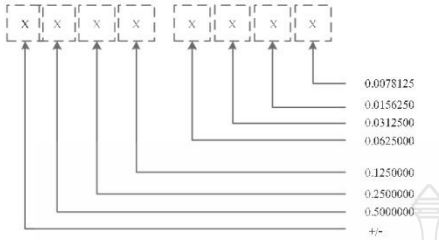


รูปที่ 4 ขั้นตอนการทำงานของโปรแกรมที่ออกแบบบนชิพเอฟพีจีเอ

จากรูปที่ 4 เป็นการแสดงให้เห็นถึงขั้นตอนการประมวลผลสัญญาณเสียงพูดบนบอร์ดเอฟพีจีเอ โดยออกแบบการจำลองการทำงานลงบนชิพเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเฮชดีแอลไปที่ในตัวชิพเอฟพีจีเอด้วยโปรแกรม Xilinx ISE

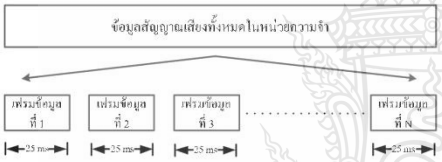
3. การทดลอง

สัญญาณเสียงพูดที่ใช้ในการวิจัยนี้เป็นสัญญาณเสียง 8 บิต โดยจะใช้รูปแบบการกำหนดการเก็บข้อมูลสัญญาณเสียงพูดแบบตัวเลขทศนิยมแบบ Fixed-Point ดังรูปที่ 5 โดยที่อัตราสุ่มสัญญาณอยู่ที่ 8 kHz และมีอัตราบิต (Bit Rate) อยู่ที่ 64 kbps ทำการแบ่งสัญญาณเสียงออกเป็นเฟรมย่อย ก่อนนำมาวิเคราะห์ที่ละส่วนจะทำให้วิเคราะห์ง่ายขึ้นและลดความซับซ้อนในการประมวลผลสัญญาณ



รูปที่ 5 รูปแบบการกำหนดข้อมูลสัญญาณเสียงพูดแบบตัวเลขทศนิยมแบบ Fixed-Point

ขนาดของหน้าต่าง (Window) มีค่าอยู่ในช่วงที่เหมาะสมที่ใช้ในการทดลองเลือกที่ 200 จะเท่ากับเฟรมละ 25 ms ดังรูปที่ 6 จากนั้นสัญญาณเสียงจะจำลองการทำงานลงบนชิพเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเสชดีแอล ในตัวชิพเอฟพีจีเอด้วยโปรแกรม Xilinx ISE ต่อไป



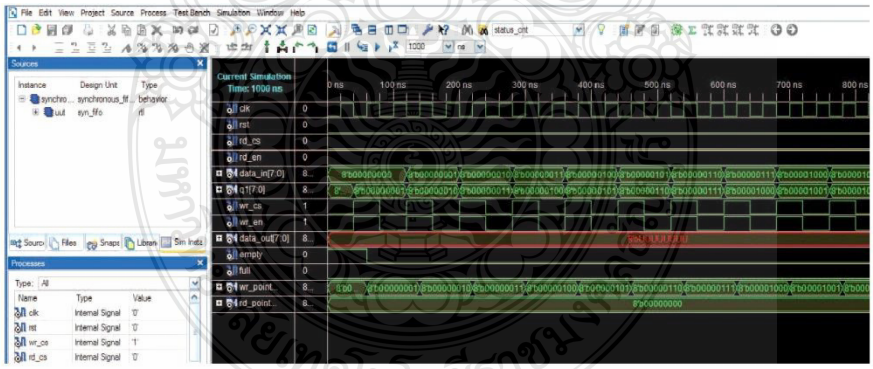
รูปที่ 6 การแบ่งสัญญาณเสียงออกเป็นเฟรมจากข้อมูลในหน่วยความจำ

4. ผลการทดลอง

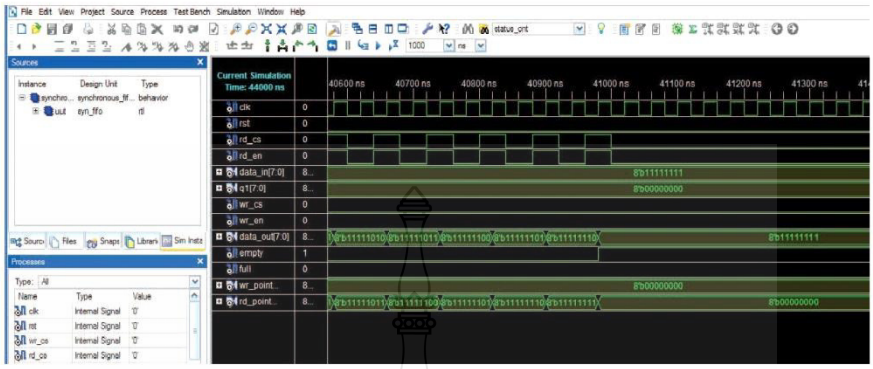
งานวิจัยนี้ได้ศึกษาและพัฒนาเทคนิคการประมวลผลสัญญาณเสียงพูดบนบอร์ดเอฟพีจีเอ ซึ่งเป็นการพัฒนาเทคนิคในการประมวลผลของสัญญาณบนฐานเวลาจริงมาประยุกต์ใช้จะออกแบบโปรแกรมการจัดการสัญญาณเสียง ซึ่งสัญญาณเสียงพูดอินพุตจะถูกประมวลผลที่ 2^8 บิต มีรูปแบบการกำหนดการเก็บข้อมูลสัญญาณเสียงพูดแบบตัวเลขทศนิยมแบบ Fixed-Point แบ่งเสียงเป็นเฟรมที่ 200 (25 ms) แล้วทำการจำลองการทำงานลงบนชิพเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเสชดีแอลไปที่ในตัวชิพเอฟพีจีเอด้วยโปรแกรม Xilinx ISE ในขั้นตอนแรกจะรับข้อมูลสัญญาณเสียงมาเก็บไว้ที่หน่วยความจำภายในเอฟพีจี โดยเก็บข้อมูลสัญญาณเสียงพูดแบบตัวเลขทศนิยมแบบ Fixed-Point ที่ 2^8 บิต และประมวลผลในการแบ่งข้อมูลสัญญาณเสียงพูดดังรูปที่ 7 และเมื่อทำการแบ่งข้อมูลเสร็จแล้ว จะทำการประมวลคืนข้อมูลสัญญาณเสียงพูด ดังรูปที่ 8

5. สรุป

การพัฒนาเทคนิคสำหรับการประมวลผลสัญญาณเสียงพูดบนบอร์ดเอฟพีจีเอบนฐานเวลาจริงเป็นการประยุกต์ใช้ชิพชิพเอฟพีจีเอในการประมวลผลสัญญาณ จากการจำลองการทำงานลงบนชิพเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเสชดีแอลไปที่ในตัวชิพเอฟพีจีเอด้วยโปรแกรม Xilinx ISE นั้น ผลการทดลองพบว่าเทคนิคนี้สามารถทำงานได้บนระบบฐานเวลาจริงได้ จากการวิจัยนี้จะเป็นแนวทางการการนำชิพชิพเอฟพีจีเอไปประยุกต์ใช้ในหลายๆ ด้าน เช่น การบีบอัดสัญญาณเสียง การรู้จำสัญญาณเสียง การแยกแยะผู้พูด รวมถึงการสังเคราะห์เสียงต่อไป



รูปที่ 7 ผลจำลองการประมวลผลสัญญาณเสียงพูดด้วยชิพเอฟพีจีเอจากโปรแกรม Xilinx ISE ก.



รูปที่ 8 ผลจำลองการประมวลผลสัญญาณเสียงพูดด้วยชิพเอฟพีจีเอจากโปรแกรม Xilinx ISE ข.

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] สุภาธิณี กรสิงห์, เกรียงศักดิ์ พัฒนบุรี, เฉลิมเกียรติ สุดาชา และ จักรี ศรีนนท์ฉัตร "การวิเคราะห์และเปรียบเทียบเทคนิคการสกัดคุณลักษณะของสัญญาณที่เป็นเสียงและเสียงรบกวน" การประชุมวิชาการเครือข่ายวิศวกรรมไฟฟ้า มหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 4, 2554

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

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

[6] เฉลิมเกียรติ สุดาชา, อรรถพล เพียรผลดีสกุล และ จักรี ศรีนนท์ฉัตร "การประยุกต์ใช้เอฟพีจีเอสำหรับการบีบอัดข้อมูล" การประชุมวิชาการมหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 3, 2553

ประวัติผู้เขียนบทความ

เฉลิมเกียรติ สุดาชา : สำเร็จการศึกษาระดับปริญญาตรี ด้านวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม ในปี พ.ศ. 2553 จากมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ปัจจุบันกำลังศึกษาระดับปริญญาโทด้านวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม

จักรี ศรีนนท์ฉัตร สำเร็จการศึกษระดับปริญญาเอกจาก Northumbria University, UK. ในปี พ.ศ. 2548 ในสาขาวิศวกรรมไฟฟ้า สาขาเอกวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม ปัจจุบันดำรงตำแหน่งอาจารย์ผู้สอนที่ภาควิชาอิเล็กทรอนิกส์และโทรคมนาคมมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

การประชุมเครือข่ายวิชาการบัณฑิตศึกษาแห่งชาติ ครั้งที่ 1

1st Conference on Graduate Student Network of Thailand

(GS-NETT 2012)

"บัณฑิตศึกษาไทย ใฝ่วิจัยสหวิทยาการ"

18 ธันวาคม 2555

ณ ศูนย์ประชุมธรรมศาสตร์ รังสิต มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต

สำนักงานบริหารการวิจัย มหาวิทยาลัยธรรมศาสตร์

รายชื่อคณะกรรมการจัดการประชุม
การประชุมเครือข่ายวิชาการบัณฑิตศึกษาแห่งชาติ ครั้งที่ 1

คณะกรรมการที่ปรึกษา

อธิการบดีมหาวิทยาลัยธรรมศาสตร์
รองอธิการบดีฝ่ายบริหาร ศูนย์ลำปาง
รองอธิการบดีฝ่ายบริหาร ศูนย์รังสิต
รองอธิการบดีฝ่ายวิชาการ

คณะกรรมการฝ่ายวิชาการ

รองอธิการบดีฝ่ายบริหาร ศูนย์ลำปาง		ประธานกรรมการ
ผู้ช่วยอธิการบดีฝ่ายบริการวิชาการและวิจัย		รองประธานกรรมการ
ศาสตราจารย์ ดร.อัญญา ขันธวิทย์	มหาวิทยาลัยธรรมศาสตร์	กรรมการ
ศาสตราจารย์ ดร.พิณทิพย์ ทวยเจริญ	ข้าราชการบำนาญ	กรรมการ
ศาสตราจารย์ ดร.ตีรณ พงศ์มงพัฒน์	จุฬาลงกรณ์มหาวิทยาลัย	กรรมการ
ศาสตราจารย์ ดร. ธเนศ อาภรณ์สุวรรณ	มหาวิทยาลัยธรรมศาสตร์	กรรมการ
ศาสตราจารย์ ดร. ศิริลักษณ์ โรจนกิจอำนวย	มหาวิทยาลัยธรรมศาสตร์	กรรมการ
ศาสตราจารย์ ดร.ปรีดา วิบูลย์สวัสดิ์	ที่ปรึกษาสำนักงานกองทุนสนับสนุนการวิจัย	กรรมการ
ศาสตราจารย์ ดร.สำเริง จักรใจ	ม.เทคโนโลยีพระจอมเกล้าธนบุรี	กรรมการ
ศาสตราจารย์ ดร.สมนึก ตั้งเติมสิริกุล	สถาบันเทคโนโลยีนานาชาติสิรินธร	กรรมการ
ศาสตราจารย์ นพ.สุรศักดิ์ ฐานีพานิชสกุล	จุฬาลงกรณ์มหาวิทยาลัย	กรรมการ
ศาสตราจารย์ ทพ. ดร. สิทธิชัย ขุนทองแก้ว	มหาวิทยาลัยธรรมศาสตร์	กรรมการ
ศาสตราจารย์ พิเศษ ดร.กาญจนา เงารังษี	มหาวิทยาลัยนเรศวร	กรรมการ
ศาสตราจารย์ ดร. เกศรา ณ บางช้าง	มหาวิทยาลัยธรรมศาสตร์	กรรมการ
รองศาสตราจารย์ ดร.ตรีทศ เหล่าศิริหงษ์ทอง	มหาวิทยาลัยธรรมศาสตร์	กรรมการ
รองศาสตราจารย์ ดร. ทศพล เกียรติเจริญผล	มหาวิทยาลัยศรีนครินทรวิโรฒ	กรรมการ
รองศาสตราจารย์ ดร. สมชาย นำประเสริฐชัย	มหาวิทยาลัยเกษตรศาสตร์	กรรมการ
รองศาสตราจารย์ ดร. ศากุน บุญอิต	มหาวิทยาลัยธรรมศาสตร์	กรรมการ
ผู้ช่วยศาสตราจารย์ ดร.ทิพวรรณ ปิ่นวณิชย์กุล	ม.เทคโนโลยีพระจอมเกล้าธนบุรี	กรรมการ
ผู้ช่วยศาสตราจารย์ ดร.ปรารถนา ปุณณกิติเกษม	มหาวิทยาลัยมหิดล	กรรมการ
ผู้ช่วยศาสตราจารย์ ดร. สิทธิพร พิมพ์สกุล	สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง	กรรมการ
ผู้ช่วยศาสตราจารย์ ดร. รัญชนา สินธวาลัย	มหาวิทยาลัยสงขลานครินทร์	กรรมการ
อาจารย์ ดร. จักรวาล คุณะดิลก	มหาวิทยาลัยบูรพา	กรรมการ
อาจารย์ ดร. ธันยวัต สมใจทวีพร	สถาบันการจัดการปัญญาภิวัฒน์	กรรมการ
ผู้อำนวยการสำนักงานบริหารการวิจัย		กรรมการและเลขานุการ
นางบุญถม วชิราธาดา		ผู้ช่วยเลขานุการ
นางสาวเฟื่องกาญจน์ สารัมภานนท์		ผู้ช่วยเลขานุการ

คณะทำงานดำเนินงานจัดประชุม

รองอธิการบดีฝ่ายบริหาร ศูนย์ลำปาง	ที่ปรึกษา
ผู้ช่วยอธิการบดีฝ่ายบริการวิชาการและวิจัย	ประธานคณะทำงาน
รองคณบดีฝ่ายวิชาการ คณะรัฐศาสตร์	คณะทำงาน
รองคณบดีฝ่ายวิชาการ คณะสาธารณสุขศาสตร์	คณะทำงาน
รองคณบดีฝ่ายบัณฑิตศึกษา คณะวิทยาศาสตร์และเทคโนโลยี	คณะทำงาน
รองคณบดีฝ่ายบัณฑิตศึกษา คณะทันตแพทยศาสตร์	คณะทำงาน
รองคณบดีฝ่ายวางแผนและวิจัย วิทยาลัยนวัตกรรม	คณะทำงาน
ผู้ช่วยคณบดีฝ่ายบัณฑิตศึกษาและวิจัย คณะศิลปศาสตร์	คณะทำงาน
ผู้ช่วยคณบดีฝ่ายบัณฑิตศึกษา คณะวิทยาศาสตร์และเทคโนโลยี	คณะทำงาน
ผู้อำนวยการโครงการปริญญาโท คณะวารสารศาสตร์และสื่อสารมวลชนผู้อำนวยการโครงการบัณฑิตศึกษา	คณะทำงาน
สาขาเทคนิคการแพทย์ คณะสหเวชศาสตร์ผู้ช่วยผู้อำนวยการโครงการบัณฑิตศึกษา สาขาชีวเวชศาสตร์ คณะสหเวชศาสตร์	คณะทำงาน
ผู้อำนวยการกองกิจการนักศึกษา	คณะทำงาน
หัวหน้างานประชาสัมพันธ์	คณะทำงาน
ผู้อำนวยการสำนักงานบริหารการวิจัย	เลขานุการ
หัวหน้างานส่งเสริมและเผยแพร่งานวิจัย	ผู้ช่วยเลขานุการ
นายทศพร สมใจ	ผู้ช่วยเลขานุการ

คณะกรรมการอำนวยการจัดประชุมวิชาการ

นางสาวอัจฉรา จินดารัตน์	ประธานกรรมการ
นางสาวทัศนีย์ ดาวเรือง	กรรมการ
นางสาวเบญจวรรณ ประจวบลาภ	กรรมการ
นางวิลาวัลย์ กางการ	กรรมการ
นายอุทัย ขันธมาลัย	กรรมการ
นางสาวจินตนา เที่ยงตรง	กรรมการ
นางสาวเฟื่องกาญจน์ สารัมภานนท์	กรรมการ
นางสาวศิตกมล วงษ์สมบูรณ์	กรรมการ
นางสาวสุปราณี ปราชญ์รัตนะกวี	กรรมการ
นางสาวปิยนาถ พุทธรัตน์	กรรมการ
นางสาวศิริรัตน์ จำปาสด	กรรมการ
นายรติ สมันตรัฐ	กรรมการ
นายวัฒนา แกล้วทนงค์	กรรมการ
นายวิระศักดิ์ คงแก้ว	กรรมการและที่ปรึกษาด้านไอที
นางบุญถม วชิราธาดา	กรรมการและเลขานุการ
นายเถกิงศักดิ์ ไชยา	ผู้ช่วยเลขานุการ
นายทศพร สมใจ	ผู้ช่วยเลขานุการ

รายชื่อผู้ทรงคุณวุฒิพิจารณาบทความ

ศ.ดร.สมชาติ ฉันทศิริวรรณ	มหาวิทยาลัยธรรมศาสตร์
ศ.ดร.สมนึก ตั้งเติมสิริกุล	สถาบันเทคโนโลยีนานาชาติสิรินธร
ศ.ทพ. ดร. สิทธิชัย ขุนทองแก้ว	มหาวิทยาลัยธรรมศาสตร์
ศ.กิตติคุณ ดร.วริทธิ์ อึ้งภากรณ์	จุฬาลงกรณ์มหาวิทยาลัย
ศ.พิเศษ ดร.กาญจนา เงารังษี	มหาวิทยาลัยนเรศวร
ศ.ดร. ธนัญชัย ลีภักดิ์ปรีดา	สถาบันเทคโนโลยีนานาชาติสิรินธร
รศ. ดร. ไชยณรงค์ จักรธรานนท์	มหาวิทยาลัยธรรมศาสตร์
รศ. ดร. ดุลยโชติ ชลศึกษ์	มหาวิทยาลัยธรรมศาสตร์
รศ.ดร.สักกมน เทพหัสดิน ณ อยุธยา	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
รศ. ดร. บุษบา พฤกษาพันธุ์รัตน์	มหาวิทยาลัยธรรมศาสตร์
รศ.ดร. ธวัชชัย อ่อนจันทร์	สถาบันเทคโนโลยีนานาชาติสิรินธร
รศ.ดร.แคทลียา ปัทมพรหม	มหาวิทยาลัยธรรมศาสตร์
รศ.ดร. มรรยาท รุจิวิชชญ์	มหาวิทยาลัยธรรมศาสตร์
รศ.ทพญ.ดร.ศิริวรรณ สืบนุการณ์	มหาวิทยาลัยธรรมศาสตร์
รศ.ดร.จารุวัตร เจริญสุข	สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
รศ.ดร.ไพศาล นาผล	มหาวิทยาลัยศรีนครินทรวิโรฒ
รศ.ดร.ชินนะพงษ์ บำรุงทรัพย์	มหาวิทยาลัยธรรมศาสตร์
รศ.ดร.อรุณพร อิฐรัตน์	มหาวิทยาลัยธรรมศาสตร์
รศ.ดร.มงคล มงคลวงศ์โรจน์	สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
รศ.ภก.ดร.บุญชู ศรีตุลารักษ์	จุฬาลงกรณ์มหาวิทยาลัย
รศ.พัชรีวัลย์ ปั้นเหน่งเพ็ชร	มหาวิทยาลัยขอนแก่น
รศ.ดร.ตรีทศ เหล่าศิริหงษ์ทอง	มหาวิทยาลัยธรรมศาสตร์
รศ.ดร. ศากุน บุญอิต	มหาวิทยาลัยธรรมศาสตร์
รศ.พญ. อรพรรณ โพชนุกูล	มหาวิทยาลัยธรรมศาสตร์
ผศ. ดร. บรรยงค์ รุ่งเรืองด้วยบุญ	มหาวิทยาลัยธรรมศาสตร์
ผศ.ดร.ทวีศักดิ์ กิจกาญจนารัตน์	มหาวิทยาลัยธรรมศาสตร์
ผศ.ดร. อภิวัฒน์ มุตตามระ	มหาวิทยาลัยธรรมศาสตร์
ผศ.ดร.วาทิต ภักดี	มหาวิทยาลัยธรรมศาสตร์
ผศ.ดร. จิระพล กลิ่นบุญ	มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ
ผศ.ดร.บัณฑิต กฤตาคม	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน
ผศ.สารดา จารุพันธ์	มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก วิทยาเขตอุเทนถวาย
ผศ.ดร.นพพร ลีปรีชานนท์	มหาวิทยาลัยธรรมศาสตร์
ผศ.ดร.วารุณี กลิ่นไกล	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ผศ. พิศิษฐ์ แสง-ชูโต	มหาวิทยาลัยรามคำแหง
ผศ. ประสิทธิ์ สันติกาญจน์	มหาวิทยาลัยกรุงเทพ
ผศ.ดร.ประภาส ปิ่นตบแต่ง	จุฬาลงกรณ์มหาวิทยาลัย
ผศ.ยุบลวรรณ ตั้นเธียรรัตน์	มหาวิทยาลัยกรุงเทพ
ผศ.มณฑล ใจกุศล	สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ผศ.ญาณวุฒิ สุพิชญางกูร	มหาวิทยาลัยรังสิต
ผศ.ดร.พรระพีพัฒน์ ภาสบุตร	มหาวิทยาลัยธรรมศาสตร์
ผศ.ดร. พรพรรณ วีระปรียากูร	มหาวิทยาลัยธรรมศาสตร์
ผศ. กนกพันธรน์ โลกุตรวงศ์	สถาบันการจัดการปัญญาภิวัฒน์
ดร.นฤมล ทับจุมพล	จุฬาลงกรณ์มหาวิทยาลัย
ดร.วราภรณ์ กลิ่นบุญ	มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์
ดร.ณัฏฐ์ มากุล	มหาวิทยาลัยราชภัฏพระนคร
ดร.รัฐศักดิ์ พรหมมาศ	มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์
ทพญ.ดร.กมลพรรณ ภักดี	มหาวิทยาลัยธรรมศาสตร์
ดร.ธันยวัต สมใจทวีพร	สถาบันการจัดการปัญญาภิวัฒน์
ดร.ธีรพจน์ เวศพันธุ์	มหาวิทยาลัยอีสเทิร์นเอเชีย
ดร.ธีรเวทย์ ประมวญรัฐการ	ผู้ประเมินภายนอก
ดร.ภาสกร อินทุมาร	มหาวิทยาลัยศิลปากร
ดร.ธาตรี ใต้ฟ้าพูล	จุฬาลงกรณ์มหาวิทยาลัย
ดร.สมศักดิ์ วงษ์ประดับไชย	มหาวิทยาลัยธรรมศาสตร์
ดร.วีรชัย โรยนรินทร์	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ดร. กริช เจียมจิโรจน์	มหาวิทยาลัยธรรมศาสตร์
ดร. สัปปินันทน์ เอกอำพน	มหาวิทยาลัยธรรมศาสตร์
ดร.นพรัตน์ พฤกษ์ทวีศักดิ์	มหาวิทยาลัยธรรมศาสตร์
ดร.สโรชา เจริญวัย	มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ดร. อรรถพร วิเศษสินธุ์	มหาวิทยาลัยเกษตรศาสตร์
ดร.วรรณี เอกศิลป์	มหาวิทยาลัยรังสิต
ดร.อมรรัตน์ แก้วประดับ	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
ดร. สมชาย ศรียาบ	มหาวิทยาลัยเชียงใหม่
อาจารย์เสกสรรค์ สุทธิสงค์	มหาวิทยาลัยอีสเทิร์นเอเชีย
อาจารย์วัชระ ลอยสมุทร	มหาวิทยาลัยรังสิต
อาจารย์จตุพร สถากุลเจริญ	มหาวิทยาลัยรังสิต

Sci-Tech 033	การใช้ประโยชน์ตัวพารามิเตอร์ Q และ R ของตัวกรองคาลมานสำหรับลดสัญญาณรบกวนแบบไวท์เกาส์เซียนในการบีบอัดสัญญาณเสียง *เจษฎ์ รัศมียงทอง และ จักรี ศรีนนท์ฉัตร*	33
Sci-Tech 034	การประยุกต์ใช้การวิเคราะห์โครงข่ายในระบบสารสนเทศภูมิศาสตร์เพื่อแนะนำเส้นทางการท่องเที่ยวในจังหวัดสุพรรณบุรี *ธัญญรัตน์ ไชยคราม*	34

Sci-Health : วิทยาศาสตร์สุขภาพ

Sci-Health 001	บทบาทของ p38 ในการเจริญและทำให้เกิดหลอดเลือดและน้ำเหลืองของเซลล์มะเร็งบริเวณศีรษะและคอ *Kantima Leelahavanichkul, Panomwat Amornphimoltham, Alfredo A. Molinolo, John R. Basile, Sittichai Koontongkaew, J. Silvio Gutkind*	35
Sci-Health 002	Characteristics of Clinacanthus nutans Extraction from Thailand and Indonesia (Preliminary Study) *Moehamad Orliando Roeslan, Thaweephol Dechatiwongse Na Ayudhya, Sittichai Koontongkaew*	37
Sci-Health 003	ผลของโปรแกรมการผ่อนคลายกล้ามเนื้อแบบก้าวหน้าร่วมกับกลุ่มจิตบำบัดแบบประคับประคองต่อระดับความซึมเศร้าของผู้สูงอายุที่ป่วยด้วยโรคเรื้อรัง *เกสร มุ้ยจีน ชมชื่น สมประเสริฐ และมรรยาท รุจิวิชญ์*	38
Sci-Health 004	การเสริมสร้างความรู้สึกมีคุณค่าในตนเองของผู้ป่วยวัยรุ่นเสพติดแอมเฟตามีน *ธนาธิป หอมหวล,มรรยาท รุจิวิชญ์ และชมชื่น สมประเสริฐ*	40
Sci-Health 005	ผลของโปรแกรมไบโอฟีดแบคแบบควบคุมการทำงานของคลื่นประสาทอัลฟ่าร่วมกับการฝึกสร้างจินตนาการต่อพฤติกรรมการแสดงออกของเด็กสมาธิสั้น *อรวรรณ จันทร์มณี มรรยาท รุจิวิชญ์ และชมชื่น สมประเสริฐ*	42
Sci-Health 006	ผลของโปรแกรมไบโอฟีดแบคร่วมกับการสร้างจินตนาการต่อระดับพฤติกรรมก้าวร้าวของผู้เสพสารแอมแฟตามีน *ตรีนุช ราษฎร์ดุษดี มรรยาท รุจิวิชญ์ และเพ็ญพักตร์ อุทิศ*	44
Sci-Health 007	Inhibitory Effect on Nitric Oxide Production and Antioxidant Activity of Thai Medicinal Plants in Zingiberaceae Family for Osteoarthritis Treatment *Areeya Anuthakoengkun, Sunita Makchuchit, Arunporn Itharat*	46
Sci-Health 008	In vitro cytotoxic, antioxidant and antimicrobial activities of Derris scandens *Nuntika Prommee, Ormanee Prajuabjinda , Arunporn Itharat*	47

Sci-Tech 034

การประชุมเครือข่ายวิชาการบัณฑิตศึกษาแห่งชาติครั้งที่ 1
18 ธันวาคม 2555 มหาวิทยาลัยธรรมศาสตร์

การพัฒนาเทคนิคการแยกแยะเสียงโฆษะและเสียงอโฆษะแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ

Development Voiced and Unvoiced Classification Technique based on Real Time Processing Using FPGA Board

เฉลิมเกียรติ สุตาชา[*] จักรี ศรีนนท์ฉัตร

[1, 2] ภาควิชาวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม คณะวิศวกรรมศาสตร์
มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
E-mail: {pickup2530, jakkree_s }@hotmail.com

บทคัดย่อ

บทความนี้นำเสนอการพัฒนาเทคนิคการแยกแยะเสียงโฆษะ (Voiced) และเสียงอโฆษะ(Unvoiced) บนพื้นฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ ซึ่งบอร์ด Virtex-II Pro นั้นมี Xlinx ชิป (XC2VP30) เป็นหน่วยประมวลผลกลางถูกนำมาใช้ในการวิจัยครั้งนี้ โดย XC2VP30 มีเกททั้งหมด 30,816 ลอจิกเซลล์และสามารถต่อร่วมกับหน่วยความจำภายนอกได้ จากผลการทดลองพบว่าระบบที่ออกแบบสามารถทำงานบนฐานเวลาจริงได้โดยใช้เวลาในการประมวลผลเพียง 4.17-50 ms ซึ่งไม่ส่งผลต่อการเกิดการหน่วงเวลาเมื่อทำงานพัฒนาระบบบนฐานเวลาจริง อีกทั้งสัญญาณเสียงที่ได้ยังคงรักษาคุณภาพไว้ใกล้เคียงกับสัญญาณต้นฉบับ ซึ่งในจุดนี้เองชิป XC2VP30 สามารถพัฒนาและประยุกต์ใช้ในการบีบอัดสัญญาณเสียงและการจดจำสัญญาณเสียง

Abstract

This article presents the development voiced and unvoiced classification technique based on real time processing using FPGA board. The Virtex-II Pro board which consists of XC2VP30 chip as central processor unit is used in this research. The XC2VP30 chip consists of 30,816 logic cells and also it can operate with external memory. The experiment results show that this system can be function on real time system. It used only 4.17-50 ms for time processing which does not effect to delay time processing on real time system. Moreover, the quality of the output speech signal is still similar to the original speech signal. This is the major point that the XC2VP30 chip can be develop to use in speech compression and speech recognition.

Keywords: Voiced and unvoiced classification, Xilinx XC2VP30, Real time processing system

1. บทนำ

งานวิจัยทางด้านการแยกแยะเสียงโฆษะและเสียงอโฆษะเป็นที่น่าสนใจในงานทางด้านการประมวลผลสัญญาณเสียง [1] ทั้งนี้ได้อาศัยคุณสมบัติทางการกระจายน้ำหนักของพลังงานเสียง (Weight distribution) และการกระจายของความถี่ (Frequency distribution) การแยกแยะสัญญาณเสียงและสัญญาณรบกวน และใน [1] ได้นำเสนอการวิเคราะห์และเปรียบเทียบเทคนิคการสกัดคุณลักษณะของสัญญาณที่เป็นเสียงและเสียงรบกวน ซึ่งผลการทดลองแสดงให้เห็นว่าวิธีที่พบค่าความคาดเคลื่อน (Error) น้อยที่สุดคือวิธีที่ 5 ทั้งนี้การดำเนินการวิจัยดังกล่าวยังไม่ได้

ดำเนินการบนฐานเวลาจริงซึ่งอาจส่งผลต่อการเกิดการหน่วงเวลา (Delay) ในการดำเนินการบนฐานเวลาจริงนั้นต้องอาศัยตัวประมวลผลที่มีความสามารถค่อนข้างสูง โดยเทคโนโลยีในปัจจุบันเอฟพีจีเอเป็นหนึ่งในทางเลือกที่มีการนำมาประยุกต์ใช้ในงานวิจัยด้านการประมวลผลสัญญาณเสียง เทคโนโลยีเอฟพีจีเอเป็นอุปกรณ์สารกึ่งตัวนำชนิดโปรแกรมได้ด้วยการเขียนบรรยายพฤติกรรมและจำลองการทำงานลงด้วยภาษาวีเฮชดีแอล ซึ่งบอร์ดเอฟพีจีเอมีความถี่สัญญาณนาฬิกาสูง นำไปประยุกต์ใช้งานพร้อมยังสามารถปรับปรุงและแก้ไขชุดคำสั่งได้ง่าย [2] ในระบบการเตรียมข้อมูลสัญญาณเสียงแบบฐานเวลาจริงบนพื้นฐานเอฟพีจีเอโดยได้อาศัยเทคนิค Cross-correlation ของสัญญาณเสียงจากไมโครโฟน ใช้เสียงขนาด 16 บิต เฟรมละ 200 ms ด้วยการเลื่อนหน้าต่าง 10 ms ด้วยอัตราสุ่มสัญญาณ 48 KHz แบบฐานเวลาจริง [3] สำหรับเทคโนโลยีในการผลิตวงจรรวมที่มีระบบอยู่บนชิปโดยการออกแบบโมดูลอินเตอร์เฟสแบบ Wishbone Bus ในการติดต่อสื่อสารและสำหรับการประมวลผลสัญญาณเสียงโดยการใช้ AC97 Controller Core เทคโนโลยีในการผลิตวงจรรวมที่มีระบบอยู่บนชิปสามารถดำเนินการบนชิปเอฟพีจีเอ [4] ในการพัฒนาอัลกอริทึมการจำแนกเสียงโดยสกัดคุณลักษณะเด่นของสัญญาณบนเอฟพีจีเอด้วยเทคโนโลยีในการผลิตวงจรรวมที่มีระบบอยู่บนชิป [5] การประมวลผลสัญญาณเสียงพูดบนบอร์ดเอฟพีจีเอสัญญาณเสียงพูดถูกคำนวณในรูปแบบตัวเลขทศนิยมแบบ Fixed-Point สัญญาณเสียงพูดอินพุตจะถูกประมวลผลที่ 2^8 บิตโดยอาศัยคุณสมบัติเฉพาะตัวของหน่วยความจำและหน่วยประมวลผลบนบอร์ดเอฟพีจีเอ Discovery - III XC3S200F4 ทำงานได้บนระบบฐานเวลาจริงได้ [6]

บทความนี้นำเสนอแนวทางการพัฒนาเทคนิคการแยกแยะเสียงโฆษะและเสียงอโฆษะแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ โดยพัฒนาเทคนิคจากข้อดีของเทคนิคการแยกแยะเสียงโฆษะและเสียงอโฆษะใน [1] มาปรับปรุงและออกแบบด้วยภาษาวีเฮชดีแอล ประมวลผลสัญญาณแบบฐานเวลาจริง สังเคราะห์เป็นวงจรลอจิกไปที่ชิพเอฟพีจีเอติดตั้งบนบอร์ดเอฟพีจีเอ Virtex-II Pro (XC2VP30) มีชิป LM4550 AC97 Audio CODEC เป็นตัวควบคุมสัญญาณเสียง ทำหน้าที่บันทึกข้อมูลสัญญาณเสียงลงหน่วยความจำตามขนาดหน้าต่าง ประมวลผล แยกแยะเสียงโฆษะและเสียงอโฆษะ

2. การประมวลผลและแยกแยะสัญญาณเสียง

2.1 การประมวลผลสัญญาณเสียงบนชิพเอฟพีจีเอ

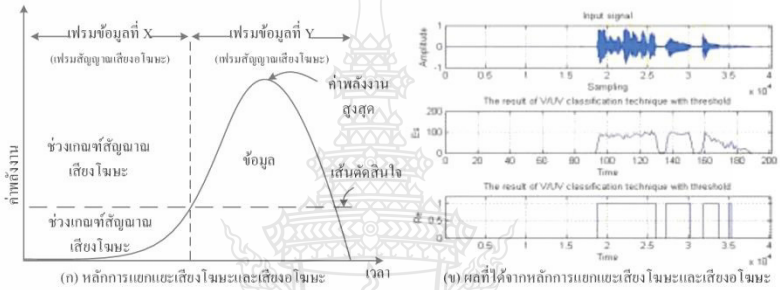
การพัฒนาเทคนิคสำหรับการแยกแยะเสียงโฆษะและเสียงอโฆษะบนพื้นฐานเอฟพีจีเอ ขั้นตอนการทำงานของโปรแกรมสามารถแบ่งเป็นขั้นตอนได้ดังนี้

1. รับข้อมูลเสียงจากภายนอกเข้ามา โดยได้ออกแบบให้ระบบสามารถเลือกข้อมูลเสียงได้จากไมโครโฟนและคอมพิวเตอร์
2. บันทึกข้อมูลเสียงลงหน่วยความจำถูกเก็บตามแอดเดรสของหน่วยความจำด้วยการกำหนดขนาดของหน้าต่าง (Window)
3. สกัดคุณลักษณะเด่นของสัญญาณจากข้อมูลในหน่วยความจำ
4. ประมวลผลเพื่อแยกแยะเสียงโฆษะและเสียงอโฆษะจากการสกัดคุณลักษณะเด่นของสัญญาณ
5. คืนกลับข้อมูลสัญญาณเสียงแยกแยะเสียงโฆษะและเสียงอโฆษะไปยังลำโพง

2.1 การแยกแยะเสียงโฆษะและเสียงอโฆษะ

เสียงโฆษะหรือเสียงก้องเกิดจากการบังคับอากาศให้ผ่านช่องสายเสียง ทำให้มีการเปลี่ยนแปลงความตึงหย่อนของเส้นเสียง โดยเส้นเสียงจะสั่นและเกิดเป็นพัลส์ของอากาศไปกระตุ้นอวัยวะกำทอนเกิดเป็นเสียงก้อง ส่วนเสียงอโฆษะหรือเสียงไม่ก้องเป็นเสียงที่ไม่เกิดจากการสั่นของเส้นเสียง แต่เกิดในช่องปากหรือโพรงจมูก โดยอวัยวะภายในช่องปาก ริมฝีปาก ซึ่งอาจขัดขวางการไหลของอากาศ ทำให้อากาศมีการไหลผ่านอย่างรวดเร็วจนกระทั่งสร้างเป็นเสียง

รบกวน ในการแยกแยะเสียงโฆษะและเสียงอโฆษะด้วยการใช้ค่าตัดสินใจ (Threshold) แสดงในรูปที่ 1 ได้แสดงหลักแยกแยะเสียงโฆษะและเสียงอโฆษะและผลที่ได้ ซึ่งได้กำหนดไว้ที่ค่าเปอร์เซ็นต์ในการปรับค่าเส้นการตัดสินใจไว้ที่ 3 เปอร์เซ็นต์ของค่าพลังงานสูงสุด ซึ่งมีประสิทธิ์ภาพที่ดีในการแยกแยะเสียงโฆษะและเสียงอโฆษะ โดยที่พลังงานสูงสุดมีค่าเท่ากับ 1 เมื่อรับข้อมูลสัญญาณเสียงมายังเส้นการตัดสินใจทำการเปรียบเทียบทุกค่าในเฟรมข้อมูล เพื่อลดขั้นตอนการประมวลผลจึงไม่คำนึงถึงเครื่องหมาย แล้วทำการนับจำนวนค่าที่ผ่านเส้นการตัดสินใจ เพื่อนำไปเปรียบเทียบค่าเปอร์เซ็นต์เพื่อใช้เปรียบเทียบค่าที่ผ่านเส้นการตัดสินใจ กำหนดไว้ที่ค่า 25 เปอร์เซ็นต์ เพื่อแยกแยะเสียงโฆษะและเสียงอโฆษะดังสมการที่ (1), (2) , (3) และ (4)



รูปที่ 1 การแยกแยะเสียงโฆษะและเสียงอโฆษะด้วยการใช้เส้นการตัดสินใจ

$$Th = Ratio1 \tag{1}$$

$$S = \begin{cases} 1, & x(n) \geq Th; voiced \\ 0, & x(n) < Th; unvoiced \end{cases} \tag{2}$$

$$E_S = \sum_{n=1}^{w} S(n) \tag{3}$$

$$\mathrm{Re} = \begin{cases} E_s \geq ratio2 & ;Voiced\ frame \\ E_s < ratio2 & ;Unvoiced\ frame \end{cases} \tag{4}$$

เมื่อ Th คือ ค่าเส้นการตัดสินใจ

$Ratio1$ คือ ค่าเปอร์เซ็นต์ในการปรับค่าเส้นการตัดสินใจ

S คือ จำนวนค่าที่ผ่านเส้นการตัดสินใจ

E_S คือ ผลรวมจำนวนค่าที่ผ่านเส้นการตัดสินใจ

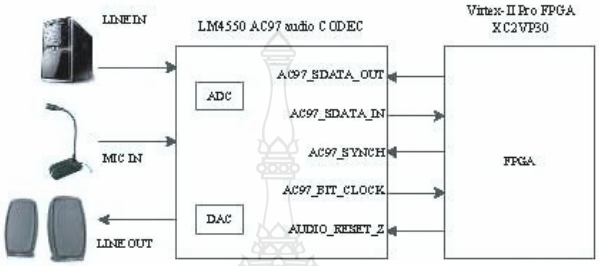
$Ratio2$ คือ ค่าเปอร์เซ็นต์เพื่อใช้เปรียบเทียบค่าที่ผ่านเส้นการตัดสินใจในเฟรม

Re คือ ผลการแยกแยะเสียงโฆษะและเสียงอโฆษะ

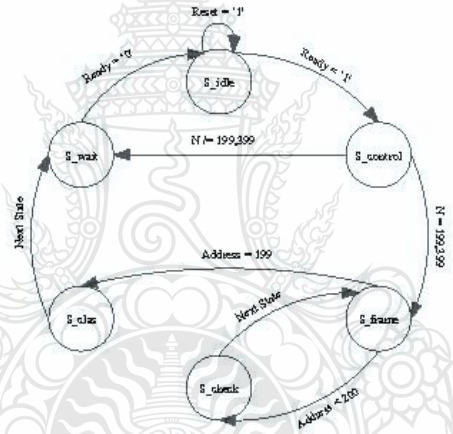
3. การออกแบบและการทำงานบนเอฟพีจีเอ

บทความนี้ได้เลือกใช้บอร์ดเอฟพีจีเอ Virtex-II Pro (XC2VP30) ที่มีชิป LM4550 AC97 Audio CODEC เป็นตัวควบคุมสัญญาณเสียง เป็นการออกแบบและการทำงานในส่วนการอินเตอร์เฟส LM4550 AC97 Audio

CODEC กับเอฟพีจีเอ แสดงในรูปที่ 2 ทั้งนี้การรับข้อมูลเสียงจากภายนอกเข้ามา ระบบออกแบบให้สามารถเลือกข้อมูลเสียงได้จากไมโครโฟนและคอมพิวเตอร์



รูปที่ 2 การอินเตอร์เฟส LM4550 AC97 Audio CODEC กับเอฟพีจีเอ



รูปที่ 3 แผนภาพไดอะแกรมของขั้นตอนการแยกแยะเสียงโฆษะและเสียงอโฆษะ

การออกแบบและการทำงานบนเอฟพีจีเอในการแยกแยะเสียงโฆษะและเสียงอโฆษะแสดงในรูปที่ 3 การออกแบบหน่วยควบคุมนี้ได้บรรยายด้วยแผนภาพไดอะแกรม โดยเริ่มที่สถานะ S_idle ถ้าสัญญาณ Reset มีค่าเท่ากับ 1 ค่าตัวแปรจะเข้าสู่สภาวะเริ่มต้นและมีสถานะเป็น S_idle แต่ถ้า Reset มีค่าเท่ากับ 0 สถานะจะเปลี่ยนไปเป็น S_control ขึ้นอยู่กับสัญญาณ Ready ถ้าเป็น 0 จะไม่เปลี่ยนสถานะรอจนเป็น 1 จะเข้าสู่สถานะ S_control ทำบันทึกข้อมูลสัญญาณเสียงให้ครบแอดเดรสของหน่วยความจำ ถ้า N มีค่าไม่เท่ากับ 199 และ 399 จะส่งข้อมูลที่อยู่ในหน่วยความจำตามแอดเดรส เพื่อทำการคืนกลับข้อมูลสัญญาณเสียงและสถานะจะเปลี่ยนไปเป็น S_wait เพื่อรอสัญญาณ Ready มีค่าเท่ากับ 0 สถานะจะเปลี่ยนไปเป็น S_idle แต่ถ้า N มีค่าเท่ากับ 199 และ 399 สถานะจะเปลี่ยนไปเป็น S_frams การประมวลผลสัญญาณเสียงบนชิพเอฟพีจีเอเสียงจะถูกเก็บเป็นเฟรมซึ่งขนาดของหน้าต่าง มีค่าอยู่ในช่วงที่เหมาะสมที่ใช้ในการทดลองเลือกที่ 200 [7] ใช้กำหนดจำนวนแอดเดรสของหน่วยความจำ เมื่อเข้าสู่

สถานะ S_frams จะทำการสกัดคุณลักษณะเด่นของสัญญาณแล้วเก็บข้อมูลเฟรมเสียงนั้นๆ เมื่อแอดเดรสมีค่าเท่ากับ 199 สถานะจะเปลี่ยนไปเป็น S_clas เพื่อแยกแยะเสียงโฆษะและเสียงอโฆษะโดยอาศัยข้อมูลของเฟรมเสียงที่ได้จากการสกัดคุณลักษณะเด่นของสัญญาณ เสร็จแล้วสถานะจะเปลี่ยนไปเป็น S_wait เพื่อรอสัญญาณ Ready มีค่าเท่ากับ 0 จะนำไปสู่สถานะ S_idle ต่อไป

4. ผลการทดลอง

จากรูปที่ 3 ไดอะแกรมจะถูกนำมาเปลี่ยนเป็นลอจิกเกทในส่วนของ Classification ซึ่งถูกออกแบบบน Xilinx XC2VP30 รวมถึงวงจรในส่วนของ AC97CMD_control และ AC97_control ซึ่งทำหน้าที่ควบคุมและอินเตอร์เฟสกับ LM4550 AC97 Audio CODEC ด้วย ในส่วนการอินเตอร์เฟส การควบคุมจังหวะการทำงาน การแปลงสัญญาณอนาล็อกเป็นสัญญาณดิจิตอล การแยกแยะเสียงโฆษะและเสียงอโฆษะ การแปลงสัญญาณดิจิตอลเป็นสัญญาณอนาล็อกด้วยภาษาวีเฮชดีแอล การออกแบบด้วยภาษาวีเฮชดีแอลนั้นเพื่อสร้างเป็นตัวประมวลผลสำหรับทำงานบนอุปกรณ์ลอจิกที่มีความสามารถในการปรับเปลี่ยนได้ แล้วทำการจำลองการทำงานลงบนชิพเอฟพีจีเอด้วยการเขียนบรรยายพฤติกรรมการทำงานด้วยภาษาวีเฮชดีแอลด้วยซอฟแวร์ Xilinx ISE

รูปที่ 4 ผลการเสียงโฆษะและเสียงอโฆษะจากซอฟแวร์ Xilinx ISE เมื่อ N มีค่าเท่ากับ 199

รูปที่ 5 ผลการแยกแยะเสียงโฆษะและเสียงอโฆษะจากซอฟแวร์ Xilinx ISE เมื่อแอดเดรสมีค่าเท่ากับ 199

ผลการจำลองเมื่อเข้าสู่สถานะ S_control, S_frame, S_check S_clas, S_wait และ S_idle เมื่อ N มีค่าเท่ากับ 199 ดังรูปที่ 4 จะแสดงการเปลี่ยนสถานะจาก S_control ที่ใช้ควบคุมการเก็บข้อมูลเสียงบนหน่วยความจำไปยัง S_frame และ S_check เพื่อเข้าสู่การสกัดคุณลักษณะเด่นของสัญญาณและเมื่อแอดเดรสมีค่าเท่ากับ 199 ดังรูปที่ 5 จะแสดงการเปลี่ยนสถานะจาก S_frame ไปยังสถานะ S_check เพื่อนำค่าที่ได้จากการสกัดคุณลักษณะเด่นของสัญญาณเข้าสู่ขั้นตอนการแยกแยะเสียงโฆษะและเสียงอโฆษะต่อไป จากนั้นซอฟแวร์ Xilinx ISE ช่วยสังเคราะห์

เป็นวงจรลอจิตที่ตัวชิพเอฟพีจีเอ ติดตั้งบนบอร์ดเอฟพีจีเอ Virtex-II Pro (XC2VP30) ดังรูปที่ 6 โดยการแยกแยะเสียงโฆษะและเสียงอโฆษะ ขั้นตอนการประมวลผลทั้งหมดกระทำบนบอร์ดเอฟพีจีเอ พบว่าเทคนิคสำหรับการแยกแยะเสียงโฆษะและเสียงอโฆษะบนพื้นฐานเอฟพีจีเอ สามารถทำได้บนระบบฐานเวลาจริง เนื่องจากต้องเก็บข้อมูลเสียงเพื่อนำไปใช้ในการประมวลผล ทำให้เกิดการหน่วงเวลาจากการประมวลผล ตั้งแต่ 4.17-50 ms ขึ้นอยู่กับอัตราการสุ่มสัญญาณสามารถปรับค่าได้ที่ 4 KHz จนถึง 48 KHz

รูปที่ 6 การพัฒนาเทคนิคสำหรับการแยกแยะเสียงโฆษะและเสียงอโฆษะบนบอร์ดเอฟพีจีเอ Virtex-II Pro (XC2VP30)

5. สรุป

บทความนี้ได้เสนอการพัฒนาเทคนิคการแยกแยะเสียงโฆษะและเสียงอโฆษะแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอด้วยการใช้เส้นการตัดสินใจ โดยการออกแบบระบบนี้ได้ใช้ Virtex-II Pro (XC2VP30) มีชิป LM4550 AC97 Audio CODEC เป็นตัวควบคุมสัญญาณเสียง ซึ่งในการดำเนินการบนฐานเวลาจริงนั้นต้องอาศัยตัวประมวลผลที่มีความสามารถค่อนข้างสูง ผลการทดลองแสดงให้เห็นว่าบอร์ดเอฟพีจีเอสามารถแยกแยะเสียงโฆษะและเสียงอโฆษะบนระบบฐานเวลาจริงได้ โดยใช้เวลาในการประมวลผล 4.17-50 ms ใช้ Slice Flip Flops 191 ตัว Occupied Slice 301 ตัวและใช้ BUFMUXs 2 ตัว อีกทั้งข้อดีของ XC2VP30 ที่ประกอบอยู่บน Virtex-II Pro สามารถต่อเพิ่มหน่วยความจำจากภายนอก ได้ช่วยสนับการคำนวณและทำงานบนฐานเวลาจริงให้สมบูรณ์ ดังนั้นเทคนิคที่ได้นำเสนอในบทความนี้สามารถประยุกต์ใช้ได้กับการประมวลผลสัญญาณเสียงทั้งระบบการบีบอัดสัญญาณเสียงและการจดจำคุณลักษณะของสัญญาณเสียงบนฐานเวลาจริงได้

เอกสารอ้างอิง

[1] สุภาธิณี กรสิงห์, เกรียงศักดิ์ พัฒนบุรี, เฉลิมเกียรติ สุตาชา และ จักรี ศรีนนท์ฉัตร "การวิเคราะห์และเปรียบเทียบเทคนิคการสกัดคุณลักษณะของสัญญาณที่เป็นเสียงและเสียงรบกวน" การประชุมวิชาการเครือข่ายวิศวกรรมไฟฟ้า มหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 4, นนทบุรี, 2554, หน้า 335-338

[2] เฉลิมเกียรติ สุตาชา, อรรถพล เพียรผลดีสกุล และ จักรี ศรีนนท์ฉัตร "การประยุกต์ใช้เอฟพีจีเอสำหรับการบีบอัดข้อมูล" การประชุมวิชาการมหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 3, กรุงเทพมหานคร, 2553

[3] Jung Uk Cho, Dae Ro Lee, Jong Hyun Park, Hyung Soon Kim, Chang Hoon Lee, Jong Suk Choi and Jae Wook Jeon "An FPGA-based voice signal preprocessor for the real-time cross-correlation" International Conference on Control, Automation and Systems 2007, Seoul, South Korea, 2007, pp. 793-97.

[4] Ayas Kanta Swain and Kamala Kanta Mahapatra, "Low Cost System on Chip Design for Audio Processing" Proceedings of the International Multi Conference of Engineers and Computer Scientists 2010, Hong Kong, 2010

[5] J Xu, A Ariyaeeinia and R Sotudeh" USER VOICE IDENFICATION ON FPGA" Perspectives in Pervasive Computing, London, England, 2005, pp. 75-83.

[6] เฉลิมเกียรติ สุตาชา และ จักรี ศรีนนท์ฉัตร การพัฒนาเทคนิคการประมวลผลสัญญาณเสียงพูดแบบฐานเวลาจริงโดยใช้บอร์ดเอฟพีจีเอ" การประชุมวิชาการเครือข่ายวิศวกรรมไฟฟ้า มหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 5, หนองคาย, 2555, หน้า 585-588

[7] Kreangsak Pattanaburi and Jakkree Srinonchat, "Enhancement Pattern Analysis Technique for Voiced/Unvoiced Classification," IEEE-International symposium on computer, consumer and control, Taichung, Taiwan, 2012, pp. 389-392

Conference Guide & Abstract Book
iEast
2012
International Conference on Engineering,
Applied Sciences, and Technology
November 21 - 24, 2012
The Swissôtel Le Concorde
Bangkok Thailand
Organized by:
Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang, Thailand
ECTI Association
IEEE THAILAND SECTION
JSST
ICROS
Institute of Control, Robotics and Systems
NECTEC
ptt Group
Thaioil Group

Advisory Committee

Prabhas Chongstitvatana, ECTI Association

Wirash Kanchanapibul, IEEE Thailand Section

Satoshi Tanaka, JSST, Japan

Nam-Suk Hur, ICROS, Korea

Boonsanong Ratanasoontragul, KMITL, Thailand

Peerawut Suwanjan, KMITL, Thailand

Sakchai Choochote, KMITL, Thailand

Dusanee Thanaboripat, KMITL, Thailand

Chanboon Sathitwiriyawong, KMITL, Thailand

Wanna Tungjaroenchai, KMITL, Thailand

Jiti Nukeaw, KMITL, Thailand

Apinunt Thanachayanont, KMITL, Thailand

Anantawat Kunakorn, KMITL, Thailand

Jirasek Trimetsoontorn, KMITL, Thailand

Organizing Committee

General Chair

Suchatvee Suwansawat, KMITL, Thailand

General Co-Chair

Somsak Choomchuay, KMITL, Thailand

General Secretary

Taworn Benjanarasuth, KMITL, Thailand

Technical Program Chair

Chuwong Phongcharoenpanich, KMITL, Thailand

Technical Program Co-Chairs

Apisak Popan, KMITL, Thailand

Pronsawan Assawasaengrat, KMITL, Thailand

Tanimnun Jeanaksorn, KMITL, Thailand

Nopporn Chotikakamthorn, KMITL, Thailand

Arthit Petchsasithon, KMITL, Thailand

Publication Chairs

Don Isarakorn, KMITL, Thailand

Sorawat Chivapreecha, KMITL, Thailand

Publicity Chairs

Tanawan Pinnarat, KMITL, Thailand

Jongkol Ngamwiwit, KMITL, Thailand

Wipoo Sriseubsai, KMITL, Thailand

Information Chairs

Pitak Thumwarin, KMITL, Thailand

Pikulkaew Tangtisanon, KMITL, Thailand

Special Session Chairs

Pitikhate Sooraksa, KMITL, Thailand

Sathaporn Promwong, KMITL, Thailand

Local Arrangement Chairs

Ruchira Taprap, KMITL, Thailand

Noppadol Maneerat, KMITL, Thailand

Maradee Phongpipatpong, KMITL, Thailand

Finance Chairs

Amnach Khawne, KMITL, Thailand

Wiboon Prompanich, KMITL, Thailand

International Relation Committee

Hisayuki Aoyama, University of Electro-Communications, Japan

Fernando Resende, University of Washington, USA

Danick Briand, Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland

Tian-Bo Deng, Toho University, Japan

Hiroaki Kikuchi, Tokai University, Japan

Takenobu Matsuura, Tokai University, Japan

Sutheera	Puntheeranurak	Pitak	Thumwarin
Arnon	Sakonkanapong	Sekson	Timakul
Adisorn	Sawetwiwat	Thanit	Trisuwannawat
Kingkarn	Sookhanaphibarn	Satean	Tunyasrirut
Pitikhate	Sooraksa	Santi	Wattananusorn
Wipoo	Sriseubsai	Pitchanun	Wongsiritorn
Pikulkaew	Tangtisanon	Winadda	Wongwiriyapan
Ruchira	Taprap	Chunquan	Xu
Sontisuk	Terachaichayut	Chumpol	Yuangyai
Porjai	Thamakorn		

Contents

Conference Organization	I
Sponsorships	II
Message from Honorary Chairs	VI
Message from General Chair	VIII
Message from Technical Program Chair	IX
Advisory Committee	X
Organizing Committee	XI
International Relation Committee	XIII
List of Reviewers	XIV
Local Information	XVI
Travel Information	XVII
Additional Information	XIX
Map	XX
Floor Plan	XXI
Conference Program	XXIII
Keynote Speech	XXVII
Papers' Abstracts	1
Author Index	122

International Conference on Engineering, Applied Sciences, and Technology
November 21 – 24, 2012, Swissôtel Le Concorde, Bangkok, Thailand

This paper will divert to KMITL Science and Technology Journal.

PID : 00075

Improvement Voiced And Unvoiced Classification Technique Based on Real Time Processing Using FPGA Board

Chalermkiat Sutacha
Rajamangala University of Technology Thanyaburi, Thailand
Jakkree Srinonchat
Rajamangala University of Technology Thanyaburi, Thailand

Keywords:
voiced and unvoiced classification.Xilinx XC2VP30.real time processing system

Abstract:

This article presents an improvement voiced and unvoiced (V/UV) classification technique based on real time processing using Field Programmable Gate Array (FPGA) board. The Virtex-II Pro board which consists of XC2VP30 chip as central processor unit is used in this research. The XC2VP30 chip consists of 30,816 logic cells and also it can operate with external memory. The experiment results show that this system can be function on real time system. It used only 4.17-30 ms for time processing which does not effect to delay time process on real time system. Moreover, the quality of the output speech signal is still similar to the original speech signal. This is the major point that the XC2VP30 chip can be develop to use in speech compression and speech recognition.

ประวัติผู้เขียน

ชื่อ – นามสกุล	นายเฉลิมเกียรติ สุตาชา
วัน เดือน ปีเกิด	11 ธันวาคม 2530
ที่อยู่	47/1 หมู่ 2 ตำบล หัวดอน อำเภอ เขื่องใน จังหวัด อุบลราชธานี 34150
การศึกษา	
2553 – ปัจจุบัน	กำลังศึกษาปริญญาโท คณะวิศวกรรมศาสตร์ สาขาวิชา วิศวกรรมไฟฟ้า แขนง วิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
2549 – 2553	ปริญญาตรี คณะวิศวกรรมศาสตร์ สาขาวิชา วิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
2546 – 2549	ประกาศนียบัตรวิชาชีพ สาขาวิชาไฟฟ้าและอิเล็กทรอนิกส์ วิทยาลัยเทคนิคอุบลราชธานี