

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การบีบอัดสัญญาณเสียง

VOICE COMPRESSION



เลขหมู่.....
เลขทะเบียน..... 82010
วัน,เดือน,ปี..... 4 ก.ค. 2551

b. 119 13420
i.

ปริญญาานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
ภาควิชาวิศวกรรมระบบควบคุม สาขาวิชาวิศวกรรมระบบควบคุม
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2550

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญาโท ปีการศึกษา 2550

สาขาวิชาวิศวกรรมศาสตรระบบควบคุม

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การบีบอัดสัญญาณเสียง

Voice Compression

ผู้จัดทำ

1. นาย ธนากร เสงี่ยมิษฐ์ 47010317
2. นางสาว สุชาดา พิษพันธุ์ 47010852



..... อาจารย์ที่ปรึกษา

(รศ.ดร. วันชัย รุ่งรุจา)

..... อาจารย์ที่ปรึกษา

(อ. ชวัลชัย คำศรี)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การบีบอัดสัญญาณเสียง

นาย ธนากร เสงี่ยมชัย 47010317
นางสาว สุชาดา พิษพันธ์ 47010852
อาจารย์ที่ปรึกษา
รศ. ดร. วันชัย ธีรวัชร
อาจารย์ รัชชัย คำศรี
ปีการศึกษา 2550

บทคัดย่อ

ปริญญาานิพนธ์ฉบับนี้ เป็นการศึกษาการบีบอัดสัญญาณเสียง ซึ่งในปัจจุบันความสำคัญด้านการเก็บข้อมูลเริ่มเข้ามามีบทบาทมากยิ่งขึ้น จึงต้องมีการพัฒนาเทคโนโลยีในการสื่อสาร จึงทำให้เกิดเทคโนโลยีการบีบอัดสัญญาณเสียงขึ้นมา เพื่อพัฒนาการเก็บข้อมูลให้มีประสิทธิภาพมากขึ้น โครงการงานนี้จึงมีจุดประสงค์เพื่อที่จะทำการวิจัยด้านการบีบอัดข้อมูลเสียง เพื่อจัดการข้อมูลให้มีขนาดเล็กลง และยังคงคุณภาพของข้อมูลไว้อย่างครบถ้วนโดยใช้ทฤษฎีตามมาตรฐาน SPEEX และใช้ไมโครคอนโทรลเลอร์ dsPIC30F6014 ขนาด 16 บิต เป็นตัวประมวลผล

การทดสอบสมรรถนะของระบบทำโดยการบันทึกเสียงสนทนา 2 สัญญาณเสียง โดยใช้โปรแกรม wavepad แล้วพล็อตสัญญาณเสียงด้วยโปรแกรม “matlab” เปรียบเทียบสัญญาณก่อนและหลังการบีบอัดสัญญาณเสียงโดยใช้ความถี่ออสซิลเลเตอร์ 6.144MHz เปรียบเทียบปริจิสเตอร์ควม $N1 : M1$ เป็น 2:13 กับ 1:6 ซึ่ง จากผลการทดสอบพบว่า หลังทำการบีบอัดสัญญาณแล้ว เสียงยังคงมีคุณภาพในระดับที่ยอมรับได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

VOICE COMPRESSION

Mr. Thanakorn Hengmeechai 47010317

Ms. Suchada Puechpan 47010852

Adviser

Asst. Prof. Vanchai Riewruja

Mr. Thawatchai Kamsri

Academic Year 2007

ABSTRACT

In this thesis, a speech signal processing for speech compression is studied. Technology of the communication through the computer network, transfer signal (voice and data), has an important role in the world today. Smaller size of signal became more important in communication. The scope of the study in this thesis is to research for theory of speech compression in order to reduce data size with completion. This thesis is based on SPEEX codec and used a microcontroller of 16 bits as CPU.

The efficiency of the speech compression is evaluated based on 2 speech signals using wavepad program. Then, plot the signal before and after compression using "matlab" by setting frequency as 6.144 MHz compared with register N1 : M1 as 2:13 and 1:6. As the result, after compression speech signal have a satisfied quality.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี ด้วยคำแนะนำ คำปรึกษาและคอยดูแลจากหลาย ๆ ฝ่ายด้วยกัน โดยเฉพาะอาจารย์ที่ปรึกษาที่คอยให้ความเอาใจใส่ แนะนำและช่วยเหลือเสมอมา ซึ่งก็คือ อาจารย์รัชชัช คำศรี และขอขอบพระคุณหัวหน้าภาควิชาซึ่งก็คือ รศ.ดร.วันชัย ธีรรัฐจา

นอกเหนือจากนี้ต้องขอขอบพระคุณอาจารย์ทุกท่านในภาควิชาวิศวกรรมระบบควบคุม คณะวิศวกรรมศาสตร์สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังเป็นอย่างยิ่งที่ได้ช่วยประสิทธิ์ประสาทวิชาความรู้ให้แก่คณะผู้จัดทำ

ขอขอบคุณภาควิชาวิศวกรรมระบบควบคุมที่ได้จัดเตรียมสิ่งอำนวยความสะดวก เช่น อินเทอร์เน็ตความเร็วสูง บอร์ดทดลอง dsPicdem เงินทุน ซึ่งช่วยให้การวิจัย การค้นคว้าหาความรู้ต่าง ๆ และพัฒนาโปรแกรมเป็นไปได้อย่างสะดวกและรวดเร็ว

สุดท้ายต้องขอขอบคุณบุคคลที่สำคัญที่สุดในชีวิตที่ทำให้ข้าพเจ้ามีวันนี้ นั่นคือ บิดา มารดาและบุคคลในครอบครัว อันเป็นที่เคารพรัก ซึ่งได้เลี้ยงดู คอยสั่งสอนข้าพเจ้ามาเป็นอย่างดี พร้อมให้โอกาสในการศึกษาอย่างเต็มที่ และยังให้กำลังใจ ความรักเสมอมา ข้าพเจ้าขอกราบขอบพระคุณมา ณ ที่นี้ด้วย

นาย ธนากร เสงี่ยมิชัย
น.ส. สุชาดา พิษพันธ์

สารบัญ

	หน้า
บทคัดย่อ	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญรูปภาพ	VII
สารบัญตาราง	IX
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์ของ โครงการงาน	2
1.3 ขอบเขตของ โครงการงาน	2
1.4 ขั้นตอนของ โครงการงาน	2
1.5 ประโยชน์ที่คาดว่าจะได้รับจากโครงการงาน	2
1.6 รายละเอียดของปริญญานิพนธ์	3
บทที่ 2 ทฤษฎีและความรู้ที่เกี่ยวข้อง	4
2.1 ทฤษฎีการสร้างเสียงพูด	4
2.1.1 อวัยวะที่ใช้ในการออกเสียงพูด	4
2.1.2 การเกิดเสียง	5
2.2 ทฤษฎีการประมวลผลสัญญาณดิจิทัล	6
2.2.1 การประมวลผลข้อมูลโดยการแปลงฟูเรียร์แบบเร็ว	6
2.2.2 ทฤษฎีการสุ่มตัวอย่าง	11
2.2.3 การแบ่งช่วงสัญญาณ	12
2.2.4 การเข้ารหัสของสัญญาณเสียงแบบพัลส์โค้ดมอดูเลชัน	13
2.2.5 การสังเคราะห์เสียงเชิงคณิตศาสตร์	18
2.2.6 เฟสล็อกกลูป	21
2.3 สถาปัตยกรรมไมโครคอนโทรลเลอร์ dsPIC30F6014	28
2.3.1 คุณสมบัติโดยรวมของ dsPIC30F6014	28
2.3.2 โมดูลสำหรับเชื่อมต่ออุปกรณ์อนุกรม แบบ SPI	32
2.3.3 โมดูลสำหรับเชื่อมต่ออุปกรณ์อนุกรม แบบ DCI	34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.4 ทฤษฎีและความรู้ที่เกี่ยวข้องของการบีบอัดข้อมูลเสียงมาตรฐาน SPEEX	37
2.4.1 ความรู้เบื้องต้นเกี่ยวกับ SPEEX	37
2.4.2 Speex และ CELP	38
2.4.3 การนำ Speex ไปประยุกต์ใช้	42
บทที่ 3 หลักการออกแบบ	44
3.1 หน่วยเข้าและถอดรหัสเสียง	45
3.1.1 คุณสมบัติของ Si3000	45
3.1.2 การเชื่อมต่อแบบดิจิทัล	46
3.1.3 การเลือกความถี่สัญญาณนาฬิกา	49
3.1.4 การกำหนดค่ารีจิสเตอร์สำหรับตัวกำเนิดสัญญาณนาฬิกา	50
3.2 หน่วยแสดงผล	51
3.3 หน่วยประมวลผลกลาง	52
3.3.1 การเตรียมสัญญาณในการวิเคราะห์	52
3.3.2 การวิเคราะห์สัญญาณขั้นที่หนึ่ง	55
3.3.3 การวิเคราะห์สัญญาณขั้นที่สอง	57
3.3.4 การวิเคราะห์สัญญาณขั้นที่สาม	58
3.3.5 การเวทค่าพารามิเตอร์	59
3.3.2 การจัดระดับเวกเตอร์	59
บทที่ 4 การทดลอง	61
4.1 วัตถุประสงค์ของการทดลอง	61
4.2 วิธีการทดลอง	62
4.3 ผลการทดลอง	63
4.3 สรุปผลการทดลอง	67
บทที่ 5 บทวิจารณ์และสรุป	68
5.1 สรุปผลการทดลอง	68

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
5.2 ปัญหาที่พบและแนวทางแก้ไข	69
5.3 ข้อเสนอแนะและแนวทางในการค้นคว้าพัฒนา	69
ภาคผนวก ก ไอซีเข้าและถอดรหัสเสียง S13000	70
เอกสารอ้างอิง	81



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ

รูปที่	หน้า
2.1 การเปรียบเทียบจำนวนครั้งของการคำนวณระหว่าง DFT และ FFT	7
2.2 ค่าเอกซ์โพเนนเชียลของ W (แฟกเตอร์การหมุน)	8
2.3 ความสัมพันธ์ระหว่างโดเมนเวลาและโดเมนความถี่	9
2.4 แสดงฟังก์ชันหน้าต่างในแบบต่างๆ	11
2.5 ระบบการประมวลผลสัญญาณแอนะล็อกโดยใช้การประมวลผลสัญญาณดิจิทัล	11
2.6 แสดงการแบ่งช่วงสัญญาณ	12
2.7 กระบวนการเข้ารหัสและถอดรหัสของระบบพีซีเอ็ม	13
2.8 การชักตัวอย่างสัญญาณเสียง	14
2.9 สัญญาณแอนะล็อกดั้งเดิมก่อนทำการจัดระดับสัญญาณ	15
2.10 สัญญาณดิจิทัลหลังขั้นตอนทำการจัดระดับสัญญาณ	15
2.11 การแทรกขั้นตอนของการอัดสัญญาณและการยัดสัญญาณลงในระบบพีซีเอ็ม	16
2.12 คุณลักษณะการคอมเพรสชันของไดโอด	17
2.13 แบบอย่างคุณลักษณะของการคอมเพรสชัน	18
2.14 แสดงการสังเคราะห์สัญญาณเสียง	18
2.15 บล็อกไดอะแกรมของการแปลงเสียง	20
2.16 แผนผังของเฟสล็อกกลูป	21
2.17 แผนผังของหน่วยสังเคราะห์ความถี่	23
2.18 คุณลักษณะ ช่วงพลวัต (dynamic range) ในการเปลี่ยนความถี่เฟสล็อกกลูป	25
2.19 ลักษณะสมบัติของส่วนเปรียบเทียบเฟสชนิดแอนะล็อก	27
2.20 ลักษณะสมบัติ ส่วนเปรียบเทียบเฟสชนิดดิจิทัล	27
2.21 ลักษณะสมบัติของ VCO	27
2.22 บล็อกไดอะแกรม ของ โมดูล SPI	33
2.23 การเชื่อมต่อ โมดูล SPI ระหว่างอุปกรณ์มาสเตอร์ และอุปกรณ์สเลฟ	34
2.24 บล็อกไดอะแกรม ของ โมดูล DCI	35
2.25 แสดงตัวอย่างการเชื่อมต่อกับอุปกรณ์เข้าและถอดรหัส	36
2.26 ไทมิงไดอะแกรมของการทำงานแบบ สเตเฟรมซิงค์ ชนิดหลายช่องทาง	36
2.27 การถอดรหัสแบบทั่วไปของ CELP	38
2.28 การวิเคราะห์เฟรมในระบบวงเปิด	40

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ (ต่อ)

รูปที่	หน้า
2.29 การวิเคราะห์เชิงสังเคราะห์แบบวงปีคอปทิมิซเซชันของเฟรมช้อย	40
3.1 บล็อกไดอะแกรมของไอซี Si3000	45
3.2 ไทมิงไดอะแกรมของการเชื่อมต่อแบบอนุกรม	46
3.3 การร้องขอเฟรมทุดิยุมิง	48
3.4 การเขียนข้อมูลลงเฟรมทุดิยุมิง	48
3.5 การอ่านข้อมูลจากเฟรมทุดิยุมิง	49
3.6 ระบบการสร้างสัญญาณนาฬิกาด้วยเฟสล็อกกลูป	49
3.7 บล็อกไดอะแกรมขั้นตอนการบีบอัดสัญญาณเสียง	52
3.8 การหาจุดเริ่มต้นและจุดสุดท้ายของสัญญาณเสียง โดยการเปรียบเทียบพลังงาน	54
3.9 การหาเฟรมเริ่มต้นและเฟรมสุดท้ายของสัญญาณเสียง โดยการหาอัตราการตัดศูนย์	56
3.10 ขั้นตอนการนอร์มอลไลซ์สัญญาณ	57
4.1 เสียงสนทนาที่ 1 ก่อนเข้าสู่ระบบ	63
4.2 เสียงสนทนาที่ 1 หลังเข้าสู่ระบบ	63
4.3 เสียงสนทนาที่ 2 ก่อนเข้าสู่ระบบ	64
4.4 เสียงสนทนาที่ 2 หลังเข้าสู่ระบบ	64
4.5 เสียงสนทนาที่ 1 ก่อนเข้าสู่ระบบ	65
4.6 เสียงสนทนาที่ 1 หลังเข้าสู่ระบบ	65
4.7 เสียงสนทนาที่ 2 ก่อนเข้าสู่ระบบ	66
4.8 เสียงสนทนาที่ 2 หลังเข้าสู่ระบบ	66

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 แสดงคุณสมบัติของแต่ละฟังก์ชันหน้าต่าง1	11
2.2 ความสัมพันธ์ระหว่างรูปแบบทางกายภาพกับรูปแบบทางคณิตศาสตร์ของสัญญาณเสียง	19
2.3 ย่านสัญญาณความถี่นาฬิกาที่สามารถใช้ร่วมกับวงจรเฟสล็อกภายใน	32
3.1 คุณลักษณะการสวิตช์ซิงของการเชื่อมต่อแบบอนุกรม	46
3.2 โหมดการเชื่อมต่อแบบอนุกรม	47
3.3 ตัวอย่างค่าความถี่จาก MCLK ในการสร้างความถี่สุ่มตัวอย่าง 8 kHz	50
4.1 ความถี่ออสซิลเลเตอร์ และคาร์ริสเตอร์ควมคุมที่ใช้ในการเปลี่ยนค่าเฟสล็อก	61



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของปัญหา

ในปัจจุบันนี้เทคโนโลยีต่างๆ ได้เข้ามามีบทบาทในชีวิตประจำวันของมนุษย์มากขึ้น จึงทำให้เกิดการศึกษา ค้นคว้า วิจัยเพื่อค้นหาเทคโนโลยีใหม่ๆ มาใช้เพื่ออำนวยความสะดวก สบายต่อมนุษย์ โดยเฉพาะเทคโนโลยีที่เกี่ยวกับการสื่อสารในปัจจุบัน ซึ่งมนุษย์ได้สังเกตเห็นว่าการสื่อสารนั้นเป็นสิ่งที่สำคัญและจำเป็นต่อมนุษย์เป็นอย่างมาก เนื่องจากมนุษย์เป็นสัตว์สังคมที่ต้องมีการติดต่อสื่อสารกันอยู่ตลอดเวลา โดยเฉพาะการพูดเป็นการสื่อสารที่ทำได้ง่าย ทำให้เทคโนโลยีการสื่อสารได้ถูกพัฒนาไปอย่างกว้างขวาง

การเข้าใจของสิ่งที่มนุษย์ได้ยินนั้น เป็นกระบวนการที่ซับซ้อน เกินที่จะจำลองเป็นรูปแบบหนึ่งๆ ได้ ถึงแม้ว่าเราจะรู้สึกได้ว่า เราเข้าใจสิ่งที่เราได้ยินได้ง่าย แต่ในคอมพิวเตอร์จำเป็นต้องอาศัยการคำนวณเป็นจำนวนมหาศาลเพื่อวิเคราะห์คำพูดเพียงหนึ่งพยางค์ โดยอาศัยทฤษฎี ความรู้ด้านต่างๆ มาประยุกต์ใช้ในการสร้างระบบที่สามารถแยกแยะสิ่งที่เราพูดให้ได้

เทคโนโลยีทางการบีบอัดสัญญาณเสียง เป็นเทคโนโลยีการสื่อสาร ซึ่งได้มีการพัฒนาและ ทำวิจัยต่อๆ กันมาเป็นระยะเวลาอันแล้ว เทคโนโลยีนี้ทำให้การป้อนข้อมูลนั้นสามารถใส่เสียงเป็นอินพุตได้ แทนที่จะใช้การคีย์ข้อมูลผ่านคีย์บอร์ดเท่านั้น ซึ่งเป็นสิ่งที่ทุกคนคาดหวังว่าในอนาคตจำเป็นต้องมีเทคโนโลยีเช่นนี้เกิดขึ้น โดยหลักการของการบีบอัดเสียงเป็นการวิเคราะห์ข้อมูลเสียงที่พูดออกไป นำข้อมูลเก็บไว้ในหน่วยความจำซึ่งสามารถเรียกฟังได้ตามต้องการ

โดยในโครงการนี้จะพัฒนาระบบการบีบอัดเสียงพูด ด้วยการใช้บอร์ด dsPicdem ซึ่งมี dsPIC30F6014 เป็นตัวประมวลผล เนื่องจากผู้จัดทำมีความเห็นว่า ในอดีตที่ผ่านมาการพัฒนาระบบรู้จำเสียงจะใช้คอมพิวเตอร์พัฒนาเป็นส่วนใหญ่ จึงทดลองนำบอร์ด ds30F6014 มาทำการบีบอัดเสียงว่าจะมีประสิทธิภาพมากน้อยเพียงใด สามารถนำมาประยุกต์ใช้ในชีวิตประจำวันได้หรือไม่ เช่น ใช้เสียงในการกดเบอร์โทรศัพท์, การสั่งให้เปิด/ปิดไฟ, การสั่งงานในคอมพิวเตอร์หรืออุปกรณ์อิเล็กทรอนิกส์ด้วยเสียง อีกทั้งยังสามารถนำไปเป็นแนวทางในการวิเคราะห์คำพูดแบบต่อเนื่อง หรือใช้ในการวิเคราะห์ภาษาอื่นได้อีกด้วย

1.2 วัตถุประสงค์ของโครงการ

1. เพื่อศึกษาคุณลักษณะของเสียง
2. เพื่อศึกษาทฤษฎีการบีบอัดสัญญาณเสียงแบบต่างๆ
3. เพื่อศึกษาการประยุกต์ใช้มาตรฐาน SPEEX กับบอร์ดทดลอง ในกรณีที่มีทรัพยากรจำกัด
4. เพื่อนำการบีบอัดสัญญาณเสียงไปใช้กับโปรแกรมประยุกต์ต่างๆ

1.3 ขอบเขตของโครงการ

โครงการนี้เน้นไปในเรื่องของการศึกษาทฤษฎีการบีบอัดสัญญาณเสียงโดยใช้มาตรฐาน SPEEX ซึ่งจะทำให้การทดลองโดยใช้บอร์ด dsPIC30F6014 เป็นบอร์ดทดลอง โดยเราจะศึกษาทฤษฎีต่างๆ ในการบีบอัดเสียง และเน้นด้านมาตรฐาน SPEEX เป็นหลัก โดยเป้าหมายโครงการนี้เป็น การย่อส่วนจากคอมพิวเตอร์มาใช้บอร์ดทดลองแทน ซึ่งจะเน้นศึกษาด้านการนำมาตรฐาน SPEEX มาประยุกต์ใช้กับบอร์ดทดลองเพื่อจะได้นำไปพัฒนาต่อไป

1.4 ขั้นตอนของโครงการ

1. ศึกษาธรรมชาติของเสียงพูด และทฤษฎีเบื้องต้นเกี่ยวกับการบีบอัดสัญญาณ
2. ศึกษาทฤษฎีหลักของมาตรฐาน SPEEX ที่ใช้ในการบีบอัดสัญญาณเสียง
3. ศึกษาและพัฒนาระบบการบีบอัดเสียงบนบอร์ด dsPIC30F6014
4. เขียน โปรแกรมจำลองการทำงานด้วยเสียงบน MPLAB IDE v7.60 โดยรับข้อมูลจากบอร์ด
5. วิเคราะห์ผลการทดลองที่ได้ โดยวัดผลจากกราฟที่พล็อตจากโปรแกรม matlab และ เสียงที่ออกจากลำโพงว่ามีข้อมูลสมบูรณ์หรือไม่
6. สรุปผลของโครงการ และหาแนวทางในการพัฒนาโครงการต่อไป

1.5 ประโยชน์ที่คาดว่าจะได้รับจากโครงการ

1. สามารถนำการบีบอัดเสียงพูดที่มีลักษณะไม่ต่อเนื่องไปประยุกต์ใช้กับโปรแกรมประยุกต์บางอย่างได้ เช่น การใช้เสียงในการสั่งงานในคอมพิวเตอร์ด้วยเสียง นอกจากนี้ยังสามารถนำไปช่วยอำนวยความสะดวกให้กับคนพิการได้อีกด้วย
2. สามารถนำเอาทฤษฎีต่างๆ ที่เกี่ยวข้องมาใช้ในการเก็บข้อมูลเสียงไว้ ยามที่ต้องการเรียกฟังเมื่อไรก็สามารถทำได้
3. สามารถนำเอาผลจากโครงการนี้ไปพัฒนางานวิจัยอื่นๆ ที่เกี่ยวข้องกันกับโครงการนี้ได้ เช่น ใช้เป็นแนวทางในการทดลองการบีบอัดเสียงเพื่อลดเนื้อที่ในการเก็บข้อมูลได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.6 รายละเอียดของปริญญาานิพนธ์

บทที่ 1 จะกล่าวถึง ความสำคัญและที่มาของปัญหา วัตถุประสงค์ของโครงการขอบเขตของโครงการ ขั้นตอนของโครงการ และประโยชน์ที่คาดว่าจะได้รับจากโครงการ

บทที่ 2 จะกล่าวถึง ทฤษฎีต่าง ๆ ที่เกี่ยวข้องกับการบีบอัดเสียง เช่น ทฤษฎีของเสียง ทฤษฎีการประมวลผลสัญญาณดิจิทัล คุณสมบัติต่างๆของ dsPIC30F6014 และมาตรฐานการบีบอัดสัญญาณเสียง SPEEX

บทที่ 3 จะกล่าวถึง การออกแบบและการสร้าง คุณสมบัติบอร์ด dsPIC30F6014 ว่าเราจะสามารถทำการบีบอัดเสียงโดยนำอัลกอริทึมของมาตรฐาน SPEEX มาประยุกต์ใช้กับบอร์ด dsPIC30F6014 เป็นตัวประมวลผลได้อย่างไร

บทที่ 4 จะกล่าวถึง การทดลองและผลการทดลอง เพื่อดูว่าประสิทธิภาพของการบีบอัดเสียงโดยใช้มาตรฐาน SPEEX กับบอร์ด dsPIC30F6014 เป็นตัวประมวลผลมีความสมบูรณ์หรือความถูกต้องของสัญญาณเสียงมากน้อยเพียงใด

บทที่ 5 จะกล่าวถึงสรุป ว่ามาตรฐาน SPEEX เมื่อนำมาใช้กับบอร์ด dsPIC30F6014 มีความสามารถในการบีบอัดเสียงได้หรือไม่ หรือมีความสามารถมากน้อยเพียงใด ปัญหาที่พบในระหว่างการทำโครงการนี้และข้อเสนอแนะต่างๆเพื่อนำไปพัฒนาต่อไป

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

ในการทำกระบวนการการรู้จำเสียงนั้น เราจำเป็นต้องศึกษาทฤษฎีต่างๆที่เกี่ยวข้อง เพื่อที่จะสามารถดำเนินโครงการได้อย่างถูกต้อง เช่น ศึกษาเรื่องเสียงของคนเราว่ามีคุณสมบัติอย่างไร ศึกษาว่าเราจะสามารถนำเสียงมาประมวลผลหรือหาค่าตัวแทนของเสียงได้อย่างไร เราจะสามารถรู้ได้อย่างไรว่าเสียงที่เราพูดนั้นเป็นเสียงใด มีวิธีการเปรียบเทียบอย่างไร ซึ่งเป็นคำถามที่เราจะต้องหาคำตอบต่อไป

2.1 ทฤษฎีการสร้างเสียงพูด

การพูดของมนุษย์มีใช้อาการที่เกิดที่ปากอย่างเดียวเท่านั้น หากเริ่มจากลมหายใจเข้าของมนุษย์ที่นำลมเข้าสู่ปอดจากนั้นจะใช้ลมจากปอด ซึ่งก็คือลมหายใจออก มาทำให้เกิดเสียงพูด โดยลมจะถูกบังคับให้ผ่านอวัยวะต่าง ๆ ที่สำคัญ เราสามารถจำเสียงและวิธีการพูดของคน ๆ หนึ่งได้ เนื่องจากเมื่อพิจารณาเสียงที่พูดออกมาในแต่ละครั้งอาจจะสังเกตลักษณะที่แตกต่างกันได้ เสียงพูดที่จะอธิบายด้วยหลักเกณฑ์ทางวิทยาศาสตร์แม้ว่าในภาษาหนึ่ง ๆ จะมีเสียงต่างกันมากบ้างน้อยบ้าง แต่ละเสียงก็สามารถนำมาพิจารณาและอธิบายให้รู้ลักษณะการออกเสียงและตำแหน่งที่เกิดเสียงได้ คำอธิบายนี้จะทำให้เข้าใจลักษณะเสียงทุกเสียง วิชาที่ว่าด้วยเสียงพูดเรียกว่า วิชาสัทศาสตร์ (Phonetics)

2.1.1 อวัยวะที่ใช้ในการออกเสียงพูด

อวัยวะที่ใช้ในการออกเสียงมีอยู่หลายส่วน แต่ละส่วนสามารถทำให้เสียงพูดแตกต่างกันไปได้ อวัยวะที่ใช้ในการเปล่งเสียงพูดแบ่งออกเป็น 2 ประเภท คือ

ประเภท 1 อวัยวะที่ใช้ในการทำอาหาร (Articulator) คืออวัยวะที่เคลื่อนไหวเพื่อผลัดลมไปยังส่วนต่าง ๆ อวัยวะที่สำคัญ คือ ลิ้น ซึ่งเป็นส่วนที่เคลื่อนไหวได้มากที่สุด

ประเภท 2 ที่เป็นตำแหน่งที่เกิดเสียงต่าง ๆ (Point Of Articulator) คือตำแหน่งที่เกิดเสียงต่าง ๆ เช่น ริมฝีปาก ฟัน เพดาน ส่วนต่าง ๆ เป็นต้น

อวัยวะส่วนที่มีหน้าที่ในการออกเสียงโดยตรงมีดังนี้

1. ริมฝีปาก เป็นอวัยวะส่วนที่เคลื่อนไหวได้มากและทำให้เสียงแตกต่างกันได้มาก เราอาจบังคับปากอยู่ชิด กัน ห่างกัน ขึ้นออกมา หรือห่อกลม เป็นต้น ลักษณะริมฝีปากแบบต่าง ๆ นี้ล้วนมีอิทธิพลต่อการออกเสียงและทำให้เสียงแตกต่างกันไปทั้งสิ้น

2. ฟัน เป็นอวัยวะที่ทำให้เกิดเสียงหลายชนิด เช่น เมื่อฟันกดลงบนริมฝีปากล่าง หรือฟันล่าง ลมที่ผ่านออกมาโดยแรงจะลอดช่องที่พอดฟันได้ออกมา ทำให้เกิดเป็นเสียงชนิดที่เรียกว่าเสียงเสียดแทรก เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. ปุ่มเหงือก เป็นส่วนนูนออกมาอยู่หลังฟันด้านบน ถ้าเอาลิ้นแตะจะรู้ว่าเป็นคลื่น
 4. เพดานแข็ง หรือเพดานปาก คือส่วนเฉพาะที่โค้งเป็นกระดูกแข็ง
 5. เพดานอ่อน คือส่วนที่อยู่ต่อจากเพดานแข็งไปข้างใน มีลักษณะเป็นกระดูกอ่อนที่ขยับขึ้น-ลงได้ เวลาหายใจเพดานอ่อนและลิ้นไก่ซึ่งอยู่ปลายเพดานอ่อนลดระดับลงมา เปิดช่องว่างให้ลมออกทางจมูกเวลาพูดส่วนใหญ่ปลายเพดานอ่อนและลิ้นไก่อจะถูกยกขึ้นไปจรดหลังคอก นอกจากเวลาออกเสียงนาสิกเท่านั้นที่เพดานอ่อนจะลดระดับลงมา เพื่อให้ลมออกทางจม
 6. ลิ้นไก่ เป็นก้อนเนื้อเล็ก ๆ อยู่ต่อปลายเพดานอ่อนตรงกลางปาก อวัยวะส่วนนี้สั้นรัว
 7. ลิ้น เป็นส่วนที่เคลื่อนไหวมากที่สุดในการออกเสียงพูด จึงต้องแบ่งออกเป็น 3 ส่วน ตามหน้าที่ในการออกเสียง คือ
 - 7.1) ปลายลิ้น คือส่วนปลายลิ้นซึ่งสามารถยกขึ้นไปแตะกับอวัยวะส่วนต่าง ๆ ในปากตอนบนได้โดยง่าย
 - 7.2) หน้าลิ้น คือส่วนที่อยู่ตรงข้ามกับเพดานแข็ง ถ้าวางลิ้นราบกับปาก
 - 7.3) หลังลิ้น ถ้าวางลิ้นราบกับปาก ลิ้นส่วนนี้จะอยู่ตรงข้ามกับเพดานอ่อน
 8. แผ่นเนื้อปากหลอดลม เป็นก้อนเนื้อเล็ก ๆ คล้ายลิ้นไก่ อยู่ต่อโคนลิ้นลงไปในคอ มีหน้าที่ปิดช่องลมในขณะที่รับประทานอาหาร และเปิดช่องลมเมื่อพูด
 9. กรวยคอ หมายถึง โพรงคอที่อยู่ถัดจากช่องปากลงไปจนถึงเส้นเสียง
 10. เส้นเสียง หรือสายเสียง เป็นอวัยวะสำคัญที่ทำให้เกิดเสียง เส้นเสียงมีลักษณะเป็นกล้ามเนื้อ 2 แผ่นปิดขวางอยู่บริเวณปากช่องหลอดลมจากด้านหลังมาด้านหน้า ระหว่างเส้นเสียงจะมีช่องว่างซึ่งเป็นช่องผ่านให้ลมไปถึงปอดและออกมาจากปอดได้ ช่องว่างนี้เรียกว่า ช่องว่างระหว่างเส้นเสียง (Glottis) เส้นเสียงทั้งสองสามารถดึงให้ห่างออกหรือดึงเข้าหากันได้ เส้นเสียงนี้เป็นส่วนสำคัญที่ทำให้เกิดเสียงพูดขึ้น
 11. ช่องจมูก หมายถึง โพรงในช่องจมูกซึ่งอยู่เหนือลิ้นไก่ขึ้นไป เป็นช่องที่ลมซึ่งผ่านเส้นเสียงขึ้นมาจะผ่านออกไปทางจมูกได้เมื่อเวลาหายใจและเวลาออกเสียงนาสิก ในเวลาที่พูดเสียงอื่นลิ้นไก่อจะถูกยกขึ้นไปปิดช่องจมูกเพื่อให้ลมออกมาทางปาก
- เสียงที่เกิดขึ้นนั้นไม่ว่าจะเป็นเสียงประเภทใด จะผ่านไปตามทางเดินของเสียง โดยจะเริ่มตั้งแต่ช่องว่างระหว่างเส้นเสียงถึงริมฝีปาก ในทางวิทยาศาสตร์สรุปได้ว่าทางเดินของเสียงคือ ท่อนำเสียงที่มีรูปร่างไม่แน่นอน
12. เส้นเสียงปลอม เป็นอวัยวะ ที่มีลักษณะเหมือนเส้นเสียงแต่อยู่เหนือเส้นเสียงขึ้นไป เส้นเสียงปลอมนี้เข้าใจกันว่าจะดึงเข้าหากันเมื่อเวลาพูดเสียงกระซิบ

2.1.2 การเกิดของเสียง

การเกิดของเสียงแบ่งออกเป็น 3 ขั้นตอนคือ

ขั้นตอนที่ 1 จุดเริ่มต้น เป็นขั้นตอนที่ลมเริ่มถูกขับออกจากปอด ผ่านเข้าไปสู่ขั้นตอนที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 2 การคัดแปลงลมที่เสี้ยนเสียง อวัยวะที่ใช้ในขั้นตอนนี้คือส่วนที่ต่อจากปอดขึ้นมาถึงกล่องเสียงและที่กล่องเสียง เส้นเสียงจะทำหน้าที่เป็นลิ้นปิดเปิดทำให้เกิดเสียง 2 ชนิดคือ

1. เสียงก้อง หรือเสียงโหมยะ (Voice) คือเสียงที่เกิดจากเส้นเสียงปิดกั้นลมไว้ ลมที่ผ่านออกมาจะเพิ่มแรงดันมากขึ้นจนเส้นเสียงปิด-เปิดสลับกันไป เป็นผลให้สัญญาณเสียงที่ได้ (Speech Waveform) มีลักษณะเป็นคาบ (Quasi-periodic) ซึ่งสามารถเรียกความถี่ในการปิด-เปิดของเส้นเสียงว่า ความถี่มูลฐาน (Fundamental Frequency : F_0)

2. เสียงไม่ก้อง หรือเสียงอโหมยะ (Unvoice หรือ Voiceless) คือเสียงที่เกิดในขณะที่เส้นเสียงเปิดอากาศที่ไหลผ่านอย่างรวดเร็วทำให้เกิดเสียงที่เป็นเสียงของสัญญาณรบกวน (Noise) ซึ่งไม่เป็นคาบ

ขั้นตอนที่ 3 การเปลี่ยนแปลงลักษณะเส้นเสียง อวัยวะที่ใช้คือส่วนที่ต่อจากกล่องเสียงจนถึงริมฝีปาก โดยลมที่ผ่านออกจากกล่องเสียงจะทำให้เกิดเสียงในลักษณะต่าง ๆ ซึ่งจะเกิดจากการเปลี่ยนแปลงของอวัยวะที่กล่าวไว้ในหัวข้อที่แล้ว

2.2 ทฤษฎีการประมวลผลสัญญาณดิจิทัล

ความรู้เกี่ยวกับการประมวลผลสัญญาณดิจิทัลนั้น นับเป็นส่วนสำคัญที่จำเป็นจะต้องศึกษาก่อน เพื่อเป็นพื้นฐานของการออกแบบระบบการกำจัดเสียงรบกวน โดยต้องทำความเข้าใจเกี่ยวกับหลักการประมวลผลสัญญาณดิจิทัลเบื้องต้น ซึ่งประกอบไปด้วยหัวข้อต่างๆ ดังนี้

2.2.1 การประมวลผลข้อมูลโดยการแปลงฟูเรียร์แบบเร็ว

1. การแปลงฟูเรียร์

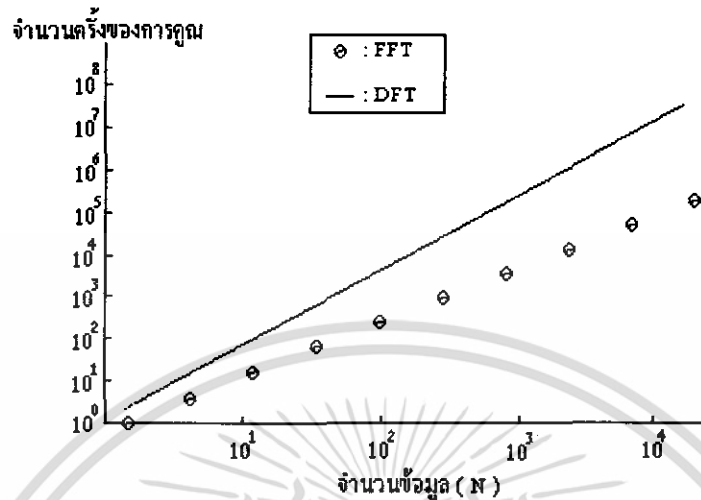
ข้อมูลส่วนใหญ่จะเป็นเหตุการณ์ที่เปลี่ยนแปลงไปตามเวลา ซึ่งข้อมูลที่แสดงได้เหล่านี้จะสามารถแสดงได้โดยใช้การประกอบกันของคลื่นฟังก์ชันตรีโกณมิติที่มีความถี่ต่างกัน ซึ่งฟังก์ชันที่ได้จากแกนของเวลาสามารถแสดงได้ด้วยฟังก์ชันที่แสดงบนแกนของความถี่ การแปลงแบบนี้เรียกว่า “การแปลงฟูเรียร์” ความสัมพันธ์ระหว่างฟังก์ชันในแกนของเวลา $F(t)$ และ ฟังก์ชันในแกนของความถี่ $F(\omega)$ ได้ดังสมการที่ (2.1) และ (2.2) ดังนี้

$$F(\omega) = \int_{-\infty}^{+\infty} F(t)e^{-j\omega t} dt \quad (2.1)$$

ถ้าทำการแปลงฟูเรียร์กับสัญญาณดิจิทัลจะเรียกการแปลงนั้นว่า การแปลงฟูเรียร์แบบไม่ต่อเนื่อง (Discrete Fourier Transform : DFT) ซึ่งจะแทนการอินทิเกรตจากอนันต์ถึงลบอนันต์ ด้วยการหาค่าผลรวมที่จำกัด (finite summation) การแปลงฟูเรียร์นี้ไม่ต่อเนื่องทั้งโดเมนเชิงเวลา และโดเมนเชิงความถี่ นอกจากนี้ยังมีการพัฒนาการแปลงฟูเรียร์ให้เร็วขึ้น หรือที่เรียกว่า “การแปลงฟู

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เรียกรวดเร็ว(Fast Fourier Transform : FFT)” ซึ่งเป็นอัลกอริทึมที่ข่นเวลาในการคำนวณ โดยลดจำนวนครั้งการคำนวณสำหรับแปลงฟูเรียร์แบบไม่ต่อเนื่อง ให้กับข้อมูลที่มีขนาดใหญ่ ดังแสดงในรูปที่ 2.1



รูปที่ 2.1 การเปรียบเทียบจำนวนครั้งของการคำนวณระหว่าง DFT และ FFT

2. การแปลงฟูเรียร์แบบเร็ว

สมมติให้ $F(t)$ เป็นสัญญาณต่อเนื่อง การแปลงฟูเรียร์แบบไม่ต่อเนื่อง จะแสดงดังสมการที่

(2.2) ดังนี้

$$C_k = \frac{1}{N} \sum_{t=1}^{N-1} f(t) \frac{e^{-j2\pi kt}}{N} \quad (2.2)$$

เมื่อ: C_k คือ ผลรวมการแปลงฟูเรียร์แบบไม่ต่อเนื่อง ณ จุดสัญญาณที่ k
 k คือ จุดของข้อมูลตั้งแต่ 0 ถึง N

วิธีคำนวณการแปลงฟูเรียร์แบบไม่ต่อเนื่อง จำเป็นต้องคำนวณผลบวกทั้งหมด N^2 ครั้ง ซึ่งมีขนาดใหญ่่มาก ดังนั้นวิธีคำนวณการแปลงฟูเรียร์แบบเร็วจึงพัฒนาขึ้นเพื่อลดจำนวนครั้งของการคูณ โดยทำการแปลงการคูณนั้นให้เป็นการบวกด้วยการรวมพจน์ก่อนเข้าคูณ ดังสมการที่ (2.3)

กำหนดให้

$$W = e^{-j(2\pi/N)} = \cos\left(\frac{2\pi}{N}\right) - j \sin\left(\frac{2\pi}{N}\right) \quad (2.3)$$

เมื่อ: W คือ แฟกเตอร์การหมุน

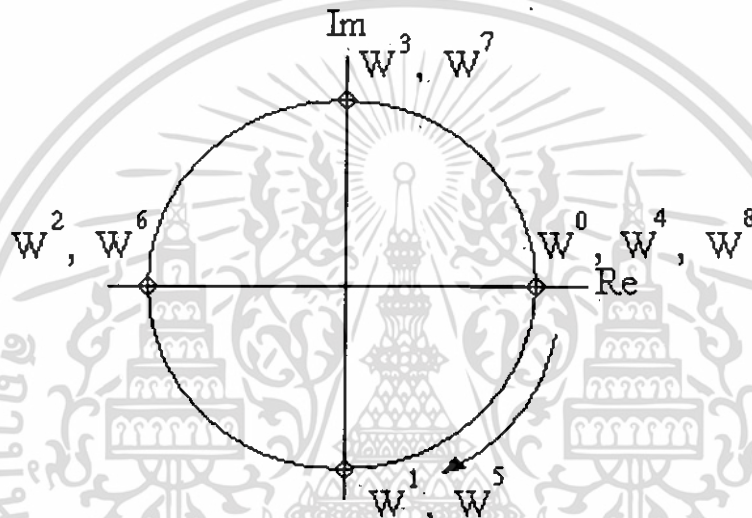
สมการที่(2.2) จะถูกแปลงเป็นสมการที่(2.4) ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$C_k = \frac{1}{N} \sum_{i=1}^{N-1} f(i) * W^{ki} \quad \text{เมื่อ } n = ki \quad (2.4)$$

ในสมการนี้เรียกว่า W^{ki} ว่า “แฟกเตอร์การหมุนเฟส (phase-rotation factor)”

แนวคิดพื้นฐานการแปลงฟูเรียร์แบบเร็วจะใช้คุณสมบัติ 2 อย่างของ W^N เช่น รูปที่ 2.2 แสดงการแบ่งวงกลมหนึ่งหน่วย (unit circle) ที่มีจุดศูนย์กลางอยู่ที่จุดกำเนิด (origin) ของระนาบเชิงซ้อน (complex plane) ออกเป็น N ส่วนเท่า ๆ กัน เมื่อมีการเพิ่ม N ก็จะหมุนไปทิศลบ ดังนั้นจากคุณสมบัติการสมมาตร ค่าของ W^6, W^2 จะเท่ากับค่าของ $-W^0$ และจากคุณสมบัติความเป็นรอบ W^4, W^8 จะเท่ากับค่าของ W^0 จึงสามารถแบ่งการคูณได้ดังสมการที่ (2.4)



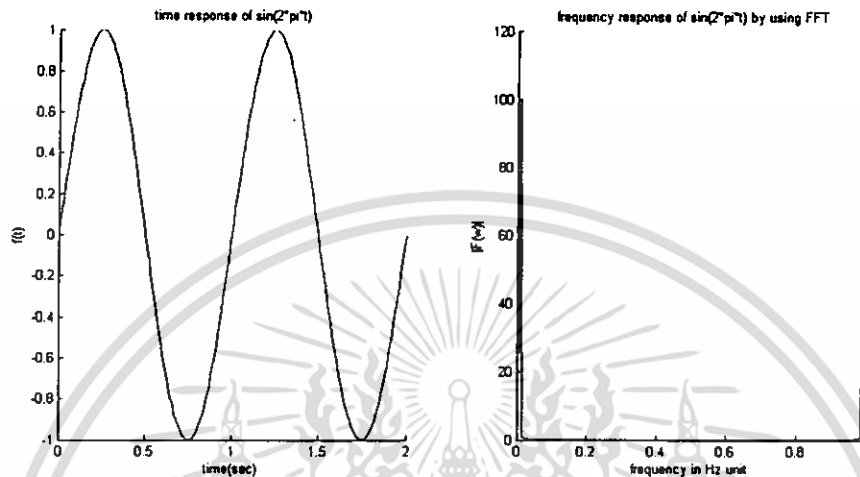
รูปที่ 2.2 ค่าเอกซ์โพเนนเชียลของ W (แฟกเตอร์การหมุน)

3. การประมวลผลแบบใช้วินโดว์

การแปลงฟูเรียร์แบบเร็วจะประมวลผลข้อมูลตัวเลขที่ส่งมาเป็นส่วนๆ จากชุดของข้อมูลที่มีทั้งหมด การตัดข้อมูลหรือคลื่นออกมาสั้นๆ แบบนี้เรียกว่า “การใช้วินโดว์ (หน้าต่างเวลา)” แต่เดิมการแปลงฟูเรียร์แบบเร็วถูกนิยามอยู่บนข้อมูลที่มีความยาวไม่จำกัด (วิธีการ DFT ก็เช่นเดียวกัน) วิธีการแปลงฟูเรียร์แบบเร็วจะตัดคลื่นออกเป็นวินโดว์ และสมมุติว่าคลื่นนั้นเป็นคลื่นที่เกิดซ้ำกันไปเรื่อยๆ จากนั้นจะทำการวิเคราะห์สัญญาณ ถ้าความยาวของข้อมูลที่วิเคราะห์นั้นเป็นจำนวนเต็มเท่าของรอบความถี่แต่ละอัน คลื่นที่แปลงฟูเรียร์แบบเร็ว นั้นจะเหมือนกับคลื่นอินพุตที่เข้ามา และจะได้ ไลน์สเปกตรัม (line spectrum) อันเดียวดังรูปที่ 2.3 แต่ถ้าความยาวของข้อมูลที่วิเคราะห์นั้นไม่เป็นจำนวนเต็มเท่าของรอบความถี่ (ปลายเริ่มกับปลายปิดไม่เชื่อมต่อกัน) ในกรณีนี้จะต้องจัดการกับคลื่นที่บิดเบี้ยวไป (distorted wave) โดยทฤษฎีแล้วเราไม่สามารถรวมพลังงานไว้ที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สเปกตรัมนั้นแต่จะเกิดการขยายไปข้างๆหรือที่เรียกว่า “ไซด์โลบ (side lobe)” การรั่วของพลังงาน เช่นนี้เรียกว่า “ค่าผิดพลาดจากการรั่ว (leakage error)” ตามทฤษฎีแล้ว เพื่อให้ได้มาซึ่งสเปกตรัม อันเดียวจำเป็นต้องใช้ข้อมูลที่ยาวไม่จำกัดแต่โดยทั่วไปแล้ว การแปลงฟูเรียร์แบบเร็วจะประมวลผล สัญญาณข้อมูลในช่วงที่จำกัด ดังนั้นจึงเกิดค่าผิดพลาดจากการรั่ว



รูปที่ 2.3 ความสัมพันธ์ระหว่าง โดเมนเวลาและ โดเมนความถี่

วิธีหนึ่งในการป้องกันค่าผิดพลาดจากการรั่วที่พบโดยทั่วไปคือ “วิธีประมวลผลด้วยวินโดว์ (window processing)” โดยจะใช้ฟังก์ชันรูปภูเขาที่ปลาย 2 ข้างของเฟรมเป็นศูนย์ ทำการฟิลเตอร์ สัญญาณ ซึ่งจะทำให้ปลายเริ่มกับปลายจบเชื่อมต่อกันฟังก์ชันแบบนี้เรียกว่า “ฟังก์ชันหน้าต่าง (window function)” การใช้ฟังก์ชันหน้าต่างในการทำซิงโครไนซ์สัญญาณที่วิเคราะห์เรียกว่า “การประมวลผลแบบใช้วินโดว์” ผลลัพธ์ที่ได้จะมีรูปร่างของสเปกตรัมใกล้เคียงกับไลน์สเปกตรัม

ส่วนของทฤษฎีจะใช้ฟังก์ชันหน้าต่างในการควบคุมค่าของ แถบความถี่เปลี่ยน ,ขนาดของ ริปเปิ้ล ซึ่งมีอยู่หลายหน้าต่างเช่น หน้าต่างสี่เหลี่ยม , หน้าต่างแฮมมิง , หน้าต่างฮานนิง , Barlett , Keiser ในแต่ละแบบจะมีผลที่ออกมาไม่เหมือนกัน เช่นการลดริปเปิ้ลแถบผ่าน และริปเปิ้ลแถบ หยุด ที่ไม่เท่ากัน คุณสมบัติของฟังก์ชันหน้าต่างในแบบต่างๆ แสดงได้ดังสมการที่ (2.5) ถึง (2.11) และแสดงรูปของหน้าต่างๆในรูปที่ 2.4

หน้าต่างสี่เหลี่ยม

$$W(n) = \begin{cases} 1 & \text{เมื่อ } -\omega < \omega < \omega_c \\ 0 & \text{นอกจากนั้น} \end{cases} \quad (2.5)$$

หน้าต่างฮานนิง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$W(n) = 0.5 \left[(1) - \cos \left(2 \cdot \pi \cdot \frac{n}{N-1} \right) \right] \quad \text{เมื่อ } 0 < n < N-1 \quad (2.6)$$

หน้าต่างแฮมมิง

$$W(n) = 0.54 - 0.46 \cos \left(2 \cdot \pi \cdot \frac{n}{N-1} \right) \quad \text{เมื่อ } 0 < n < N-1 \quad (2.7)$$

หน้าต่างแบร์เรต

$$W(n) = 1 - 2 \cdot \frac{\left| n - \left(\frac{N-1}{2} \right) \right|}{N-1} \quad \text{เมื่อ } 0 < n < N-1 \quad (2.8)$$

หน้าต่างแบล็กแมน

$$W(n) = 0.42 + 0.5 \cos \left(2 \cdot \pi \cdot \frac{n}{N-1} \right) + 0.08 \cos \left(4 \cdot \pi \cdot \frac{n}{N-1} \right) \quad \text{เมื่อ } 0 < n < N-1 \quad (2.9)$$

หน้าต่างโคเชอร์

$$W(n) = \frac{I_0(\beta) \cdot \sqrt{1 - \left[\frac{\left(\frac{n}{N-1} \right)^2}{\left(\frac{N-1}{2} \right)^2} \right]}}{I_0(\beta)} \quad \text{เมื่อ } 0 < n < N-1 \quad (2.10)$$

= 0 นอกจากนั้น

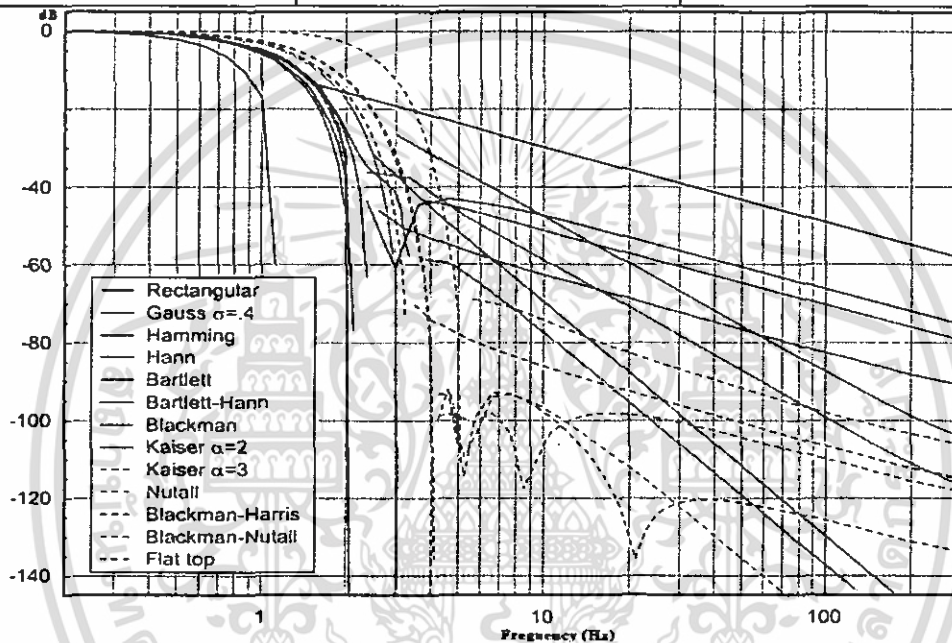
เมื่อ β คือพารามิเตอร์ในการควบคุม แถบความถี่เปลี่ยนและ ริบเบิลแถบผ่าน , ริบเบิลแถบหยุด ดังสมการที่ 2.11

$$I_0(x) = 1 + \sum_{k=1}^L \left[\left(\frac{1}{k!} \right) \cdot \left(\frac{x}{2} \right)^k \right]^2 \quad (2.11)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.1 แสดงคุณสมบัติของแต่ละฟังก์ชันหน้าต่าง

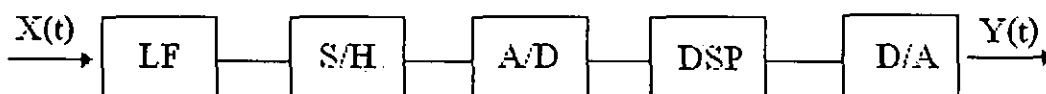
ฟังก์ชันหน้าต่าง	แถบความถี่เปลี่ยน	การลดทอนของแถบหยุด (dB)
สี่เหลี่ยม	$1.8\pi/N$	20.9
แบร์เร็ด	$6.1\pi/N$	25
ฮานนิง	$6.2\pi/N$	43.9
แฮมมิง	$6.6\pi/N$	54.5
แบล็กแมน	$11\pi/N$	75.3
ไคเซอร์	ปรับค่าได้	ปรับค่าได้



รูปที่ 2.4 แสดงฟังก์ชันหน้าต่างในแบบต่างๆ

2.2.2 ทฤษฎีการสุ่มตัวอย่าง

อาจกล่าวได้ว่าสัญญาณในธรรมชาติส่วนมาก เช่น สัญญาณเสียง สัญญาณการสั่นสะเทือนของพื้นโลก คลื่นหัวใจ หรือการแปรค่าไปของอุณหภูมิ เหล่านี้เป็นไปในลักษณะที่มีความสัมพันธ์กับเวลาที่เปลี่ยนไป หรือกล่าวได้ว่าเป็นสัญญาณแอนะล็อก การนำสัญญาณเหล่านี้ไปประมวลผลในลักษณะของดิจิทัล จะต้องใช้ระบบการประมวลผลดังรูปที่ 2.5



รูปที่ 2.5 ระบบการประมวลผลสัญญาณแอนะล็อกโดยใช้การประมวลผลสัญญาณดิจิทัล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งตามรูป วงจร S/H เป็นวงจรสุ่มและคงค่าสัญญาณไว้ เพื่อให้วงจร A/D ทำการแปลงสัญญาณจากแอนะล็อกเป็นสัญญาณดิจิทัล

ในการแปลงสัญญาณแอนะล็อกเป็นสัญญาณดิจิทัล ความถี่ในการสุ่มตัวอย่างนั้นจะต้องมากพอที่จะไม่ทำให้สัญญาณสูญเสียข้อมูลที่สำคัญออกไป

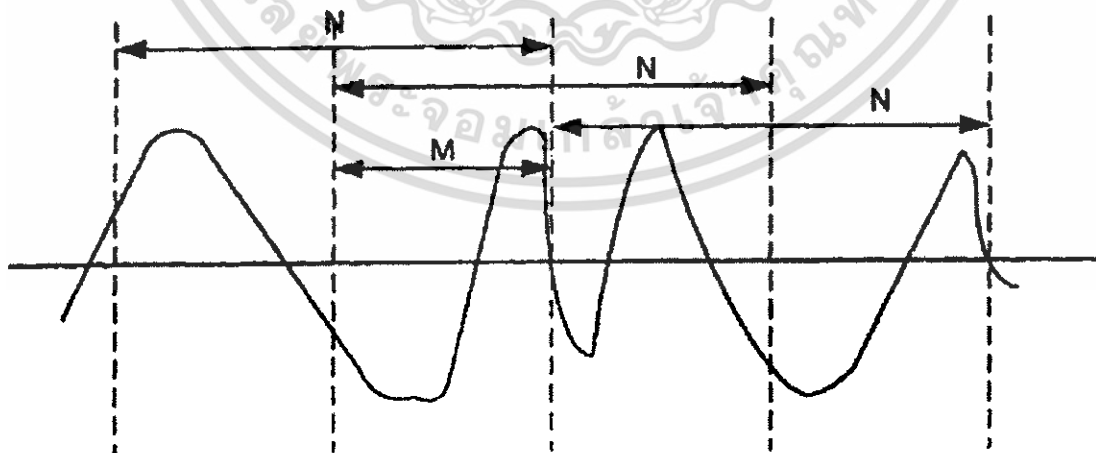
ทฤษฎีการสุ่มตัวอย่าง (sampling theory) ของ Shannon กล่าวไว้ว่า ถ้าหากมีสัญญาณแอนะล็อก $x(t)$ ที่ทำการแปลงฟูรีเยร์ หรือเพาเวอร์ สเปกตรัม (power spectrum) ของสัญญาณมีแถบความถี่ปฏิบัติงานเท่ากับ f_0 แล้วเราจะสามารถทำการสุ่มโดยที่สัญญาณนั้นไม่สูญเสียเนื้อหาที่สำคัญก็ต่อเมื่อความถี่ในการสุ่ม f_s มีค่ามากกว่าสองเท่าของความถี่ f_0 ดังสมการที่ (2.12)

$$f_s \geq 2f_0 \quad (2.12)$$

โดยทั่วไปเราอาจสุ่มด้วยความถี่ $f_s = 2f_0$ พอดี ค่าความถี่นี้เรียกว่าความถี่ในควิสต์ (Nyquist frequency) แต่ในทางปฏิบัติเพื่อหลีกเลี่ยงการไม่เป็นเชิงเส้นของการสุ่ม เราจึงมักใช้ความถี่ในการสุ่มที่มากกว่าความถี่ในควิสต์ และความถี่ในการสุ่มที่ใช้นั้นขึ้นอยู่กับลักษณะงาน ไม่กำหนดขอบเขตที่แน่นอน ดังนั้นการสุ่มสัญญาณที่จะได้ผลถูกต้องนั้น จำเป็นต้องรู้แถบความถี่ในการปฏิบัติงานของสัญญาณ ดังนั้นในการประมวลผลเราอาจใช้ตัวกรองความถี่ต่ำ (low pass filter) เป็นตัวกำหนดแถบความถี่ของสัญญาณก่อนที่จะเข้าสู่วงจร S/H

2.2.3 การแบ่งช่วงสัญญาณ

ข้อมูลเสียงที่อยู่ในโดเมนเวลา จะถูกแบ่งออกเป็นช่วงสัญญาณ (Block) ซึ่งแต่ละช่วงสัญญาณจะประกอบด้วยสัญญาณเสียง N สัญญาณ และในแต่ละช่วงสัญญาณจะถูกวิเคราะห์โดยการเลื่อนสัญญาณไปครั้งละ M สัญญาณ ดังแสดงในรูปที่ 2.6



รูปที่ 2.6 แสดงการแบ่งช่วงสัญญาณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

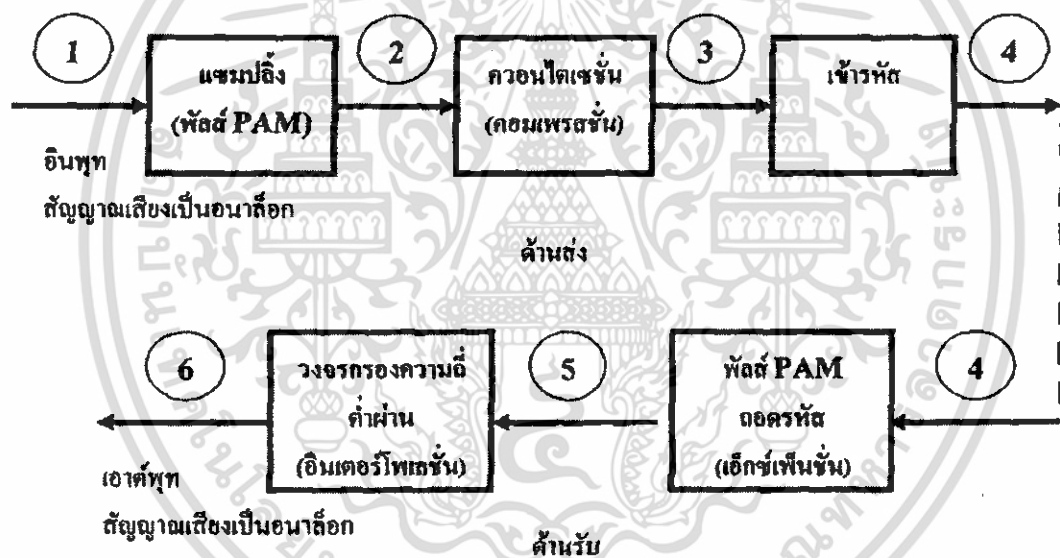
จะเห็นว่าถ้าค่า M มีค่าน้อยกว่าค่า N มากเท่าใด จะทำให้การวิเคราะห์สัญญาณมีความแม่นยำมากเท่านั้น และถ้าค่า M มีค่ามากกว่า N แล้วจะทำให้สัญญาณบางส่วนไม่ถูกใช้ในการวิเคราะห์ จะเกิดความผิดพลาดในการวิเคราะห์สัญญาณในส่วนต่อไป

การกำหนดขนาดของช่วงสัญญาณมีเงื่อนไขในการเลือกดังนี้

1. ค่า M จะต้องสั้นพอที่ทำให้คุณสมบัติของเสียงไม่เปลี่ยนแปลง
2. ค่า N จะต้องยาวพอที่จำนวนของตัวอย่างมีเพียงพอต่อการหาสัมประสิทธิ์
3. การเลื่อนการวิเคราะห์ (ค่า M) ต้องไม่ข้ามข้อมูล

2.2.4 การเข้ารหัสของสัญญาณเสียงแบบพัลส์โค้ดมอดูเลชัน

วิธีการเข้ารหัสสัญญาณเสียงในระบบสื่อสารแบบดิจิทัลที่ได้รับการพัฒนา และนำมาใช้อย่างกว้างขวางก็คือ พัลส์โค้ดมอดูเลชัน (Pulse Code Modulation : PCM) ซึ่งเป็นวิธีการหนึ่งในการเปลี่ยนสัญญาณเสียงจากแอนะล็อกให้เป็นดิจิทัล โดยจะประกอบด้วยส่วนประกอบต่างๆ ดังนี้



รูปที่ 2.7 กระบวนการเข้ารหัสและถอดรหัสของระบบพีซีเอ็ม

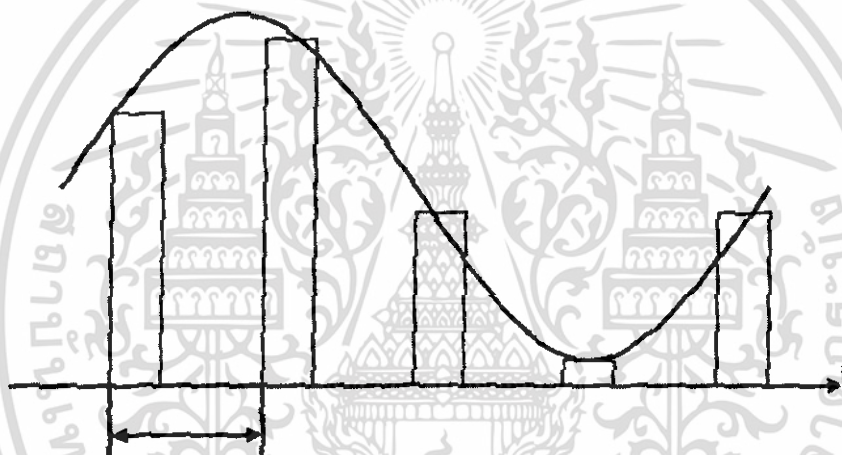
จากรูปที่ 2.7 แสดงขั้นตอนการประมวลผลสัญญาณ เพื่อให้ได้สัญญาณพีซีเอ็ม กล่าวอย่างกว้างๆก็คือ ระบบพีซีเอ็มเป็นการจัดการสัญญาณพัลส์แอมพลิจูดมอดูเลชัน (Pulse Amplitude Modulation : PAM) โดยการนำสัญญาณพีเอเอ็มไปทำการเข้ารหัส (coding) เป็นสัญญาณดิจิทัล หลังจากนั้นจึงนำสัญญาณดิจิทัลที่ได้นั้นส่งผ่านระบบต่อไปและทางด้านรับก็จะทำการถอดรหัส (decoding) เป็นสัญญาณพีเอเอ็ม แล้วนำไปตีมอดูเลตเพื่อให้ได้สัญญาณเดิมกลับมา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการส่งโดยระบบพีซีเอ็มนั้นต้องส่งด้วยจำนวนบิต 8 บิต แต่ยังมีการเข้ารหัสอีกแบบหนึ่ง ที่ทำการส่งด้วยรหัสสั้นกว่า 8 บิต คือการเข้ารหัสสัญญาณเสียงแบบอแด็ปทีฟวีดิฟเฟอร์เรนเชียลพัลส์โค้ดมอดูเลชัน ส่งผลให้การและถอดรหัสสัญญาณมีความเร็วมากขึ้น แต่ประสิทธิภาพเท่าเดิม

กระบวนการเข้าและถอดรหัสของระบบพีซีเอ็มสามารถอธิบายองค์ประกอบแต่ละส่วนของระบบได้ดังนี้

1. การซ้กตัวอย่าง



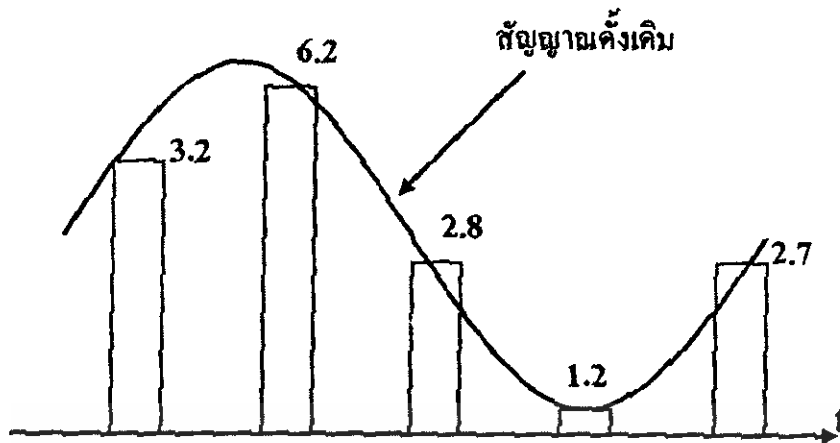
รูปที่ 2.8 การซ้กตัวอย่างสัญญาณเสียง

การซ้กตัวอย่าง (sampling) คือการทำสัญญาณซึ่งมีค่าต่อเนื่องให้เป็นแบบดิสครีต ในช่วงเวลาที่เท่าๆกัน ดังรูปที่ 2.8 และจากทฤษฎีการซ้กตัวอย่าง ถ้าเก็บตัวอย่าง (sample) ด้วยอัตรา 2 เท่าหรือมากกว่าความถี่สูงสุดของสัญญาณแอนะล็อกแล้วจะทำให้สัญญาณเดิมกลับคืนมาได้ เนื่องจากสัญญาณเสียงสนทนาที่พบถูกจำกัดให้มีความถี่ระหว่าง 300 - 3400 Hz ดังนั้นอัตราการซ้กตัวอย่างต่ำสุดจะต้องเท่ากับ 6.8 kHz สำหรับในทางปฏิบัติจะใช้ 8 kHz คือการซ้กตัวอย่างทุกๆ 125 us

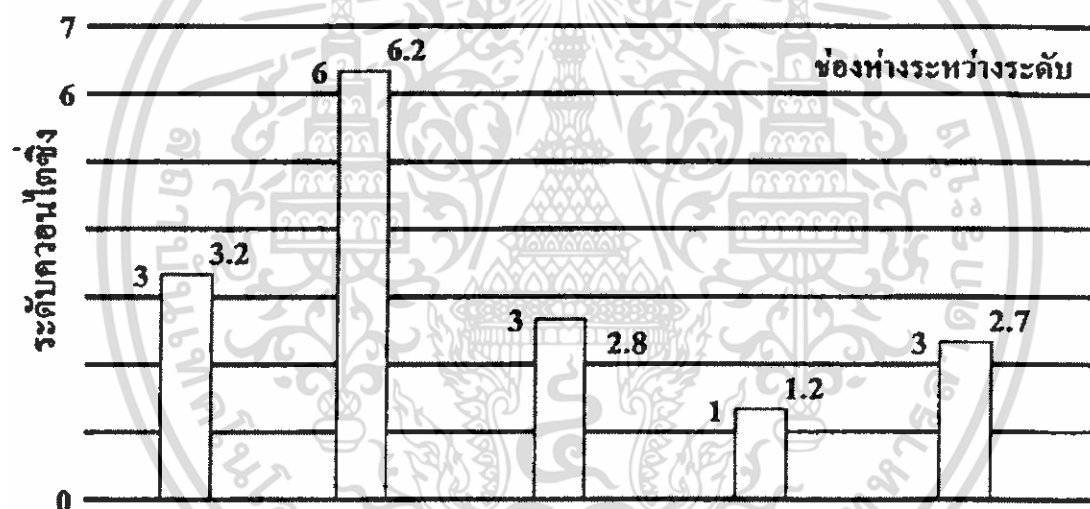
2. การจัดระดับ

ขบวนพัลส์พีเอเอ็มที่ผ่านการซ้กตัวอย่างมาแล้ว ยังถือว่าเป็นสัญญาณชนิดแอนะล็อกอยู่ คือจะมีแอมพลิจูดที่เปลี่ยนแปลงอย่างต่อเนื่องไปกับเวลาที่เป็นช่วง ๆ การจัดระดับคือกระบวนการที่เปลี่ยนแปลงแอมพลิจูดของพัลส์พีเอเอ็มเหล่านั้นให้เป็นตัวเลขแบบดิสครีต ดังแสดงในรูปที่ 2.9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.9 สัญญาณแอนะล็อกคั้งเดิมก่อนทำการจัดระดับสัญญาณ



รูปที่ 2.10 สัญญาณดิจิตอลหลังขณะทำการจัดระดับสัญญาณ

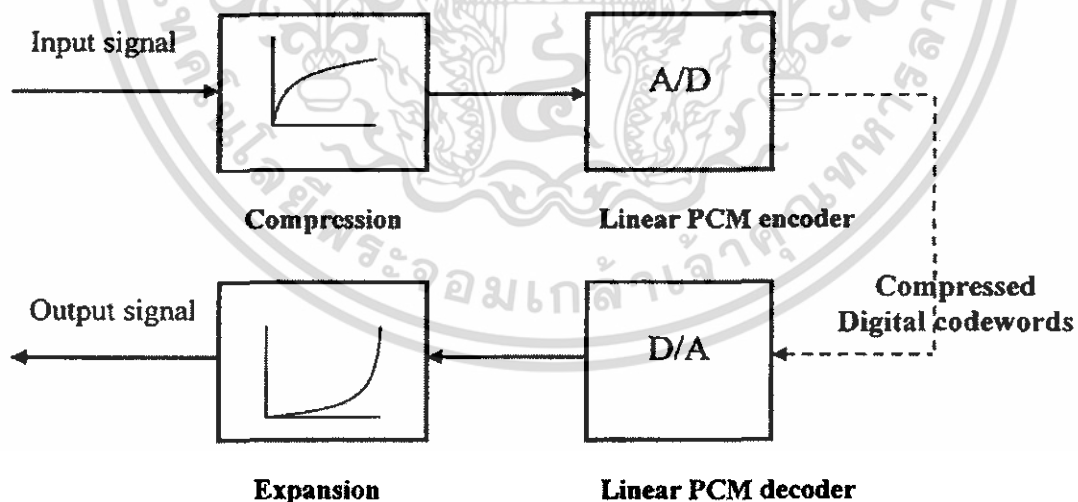
จากรูปที่ 2.10 แอมพลิจูดของการชักตัวอย่างทุกตัวของพีเอเอ็มจะถูกจัดให้เป็นระดับ ซึ่งเรียกว่า ระดับการควอนไทซ์ (quantize level) โดยมีระยะห่างระหว่างระดับข้างเคียง เรียกว่า ควอนไทซ์ซึ่งอินเทอร์วัล (quantizing interval) เท่ากัน กรณีนี้เรียกว่าการจัดแบบยูนิฟอร์ม (uniform quantizing) ขนาดของแอมพลิจูดทุกตัวจะแสดงด้วยระดับการควอนไทซ์ที่ใกล้เคียงที่สุด เช่น ขนาดของแอมพลิจูดที่ $t = t_1$ คือ 3.2 จะจัดให้เป็นระดับ 3 หรือค่าแอมพลิจูดที่ $t = t_2$ มีขนาด 6.2 จะจัดให้เป็น 6 เป็นต้น จะเห็นได้ว่าสัญญาณพีเอเอ็มที่ถูกจัดระดับแล้วนี้จะเป็นเพียงค่าโดยประมาณของสัญญาณแอนะล็อกเท่านั้น ดังนั้นส่วนเกินและส่วนขาดจากการจัดระดับจึงเป็นค่าผิดพลาดระหว่างเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัญญาณเดิมและค่าที่จัดระดับ ซึ่งค่าผิดพลาดนี้เรียกว่า ควอนไตซิงนอยส์ (quantizing noise) หรือ ค่าความผิดพลาดจากการควอนไตซิง (quantizing distortion)

จากหลักการที่กล่าวมานี้ ในทางปฏิบัติจะไม่สามารถหลีกเลี่ยงควอนไตซิงนอยส์ได้ แต่เพื่อรักษาคุณภาพเสียงในการสนทนาให้ดี จึงจำเป็นต้องทำการควอนไตซิงนอยส์นี้ให้น้อยลง ในเบื้องต้นคือการลดควอนไตซิงนอยส์อินเทอร์วัลให้แคบลงก็สามารถลดควอนไตซิงนอยส์ได้

3. คอมแพนดิง

ตามที่ได้กล่าวมาแล้วว่าเราไม่สามารถหลีกเลี่ยงควอนไตซิงนอยส์ได้ แต่เราสามารถทำให้มันลดลงได้โดยการควอนไตซิงนอยส์อินเทอร์วัล หรือการเพิ่มจำนวนระดับนั่นเอง แต่เมื่อจำนวนระดับเพิ่มขึ้นแล้วจำนวนบิตที่ใช้ก็เพิ่มขึ้นด้วย จึงจำเป็นต้องใช้ความเร็วในการส่งสัญญาณดิจิทัลสูงขึ้น ตามปกติควอนไตซิงนอยส์จะเกิดขึ้นสม่ำเสมอในทุกอินเทอร์วัลโดยไม่เกี่ยวข้องกับแอมพลิจูดของสัญญาณเดิม ในกรณีที่สัญญาณมีระดับสูง คุณภาพการเข้ารหัสของสัญญาณเสียงจะดีกว่าของสัญญาณซึ่งมีระดับต่ำ ดังนั้นจึงจำเป็นต้องพิจารณาควอนไตซิงนอยส์ในบริเวณที่สัญญาณมีระดับต่ำ การจัดแบบนอนยูนิฟอร์ม (non-uniform quantizing) คือบริเวณที่สัญญาณมีแอมพลิจูดต่ำ จะควอนไตซิงอินเทอร์วัลแคบๆ และในทางตรงกันข้าม บริเวณที่สัญญาณแอมพลิจูดสูงจะใช้ ควอนไตซิงนอยส์อินเทอร์วัลกว้างๆ ซึ่งการทำให้เป็นแบบนอนยูนิฟอร์มนั้นจะใช้หลักการคอมแพนดิงเข้าช่วย ดังรูปที่ 2.11



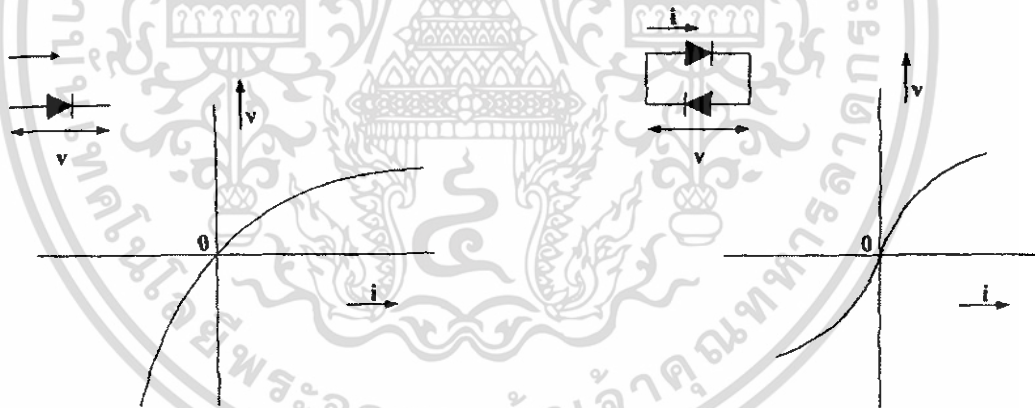
รูปที่ 2.11 การแทรกขั้นตอนของการอัดสัญญาณและการยัดสัญญาณลงในระบบพีซีเอ็ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

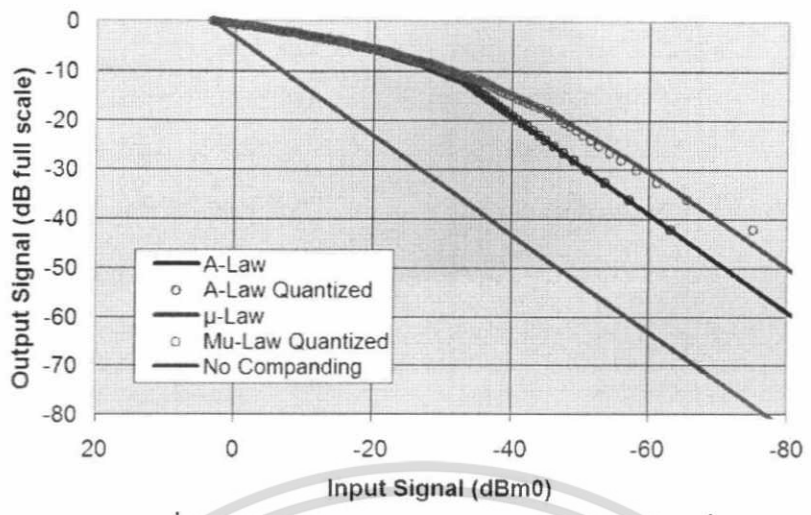
การทำการคอมแพนดิงเมื่อนำมาใช้ในระบบพีซีเอ็มจะทำให้ประสิทธิภาพในการใช้จำนวนบิตในการเข้ารหัสสูงขึ้น กรณีที่จัดระดับแบบยูนิฟอร์มนั้นจะต้องใช้ประมาณ 2,000 ระดับ จึงจะสามารถรักษาคุณภาพของเสียงให้ดี ในการเข้ารหัสต้องใช้ถึง 11 บิตต่อหนึ่งแซมเปิ้ล แต่ถ้าใช้แบบนอนยูนิ-ฟอร์มแล้วจะใช้เพียง 7 บิต ซึ่งจะมีระดับเพียง 128 ระดับเท่านั้น ก็เพียงพอที่จะทำให้อัตราส่วนของสัญญาณต่อการควอนไทซ์ซึ่งน้อยสก็ใกล้เคียงกับการจัดระดับแบบยูนิฟอร์ม CCITT กำหนดให้ใช้ 8 บิต ต่อแซมเปิ้ล 1 ตัว และระดับการควอนไทซ์เท่ากับ 256 ระดับ ก็จะเป็นการรับรองว่าเสียงพูดจะมีคุณภาพดี

ลักษณะของการคอมเพรสเซอร์นั้นจะแสดงเป็นแบบลอการิทึม รูปแบบโดยทั่วไปจะใช้คุณสมบัติของ V-I ของไดโอด ตามรูปที่ 2.11 กรณีที่ใช้คอมเพรสเซอร์จะมีกระแส I เป็นอินพุท โวลเตจ V เป็นเอาต์พุท สำหรับกรณีที่ใช้เอกซ์แพนเดอร์จะมีโวลเตจเป็นอินพุท และกระแสเป็นเอาต์พุท

คุณลักษณะของการคอมเพรสเซอร์ที่ใช้สำหรับการเข้ารหัสสัญญาณเสียงในปัจจุบันคือ μ -law ซึ่งในทวีปอเมริกาเหนือและญี่ปุ่นเป็นหลัก โดยที่ค่า μ ที่ใช้คือ 255 สำหรับในประเทศไทย ยุโรปและประเทศไทยใช้ A -law ซึ่งโดยหลักการจะเหมือนกับกับกฎ μ -law แต่รายละเอียดจะแตกต่างกัน โดยค่า A ที่ใช้คือ 87.6 คุณลักษณะทั้งสองนี้แสดงไว้ในรูปที่ 2.12 เฉพาะกรณี $\mu = 255$ และ $A = 87.6$



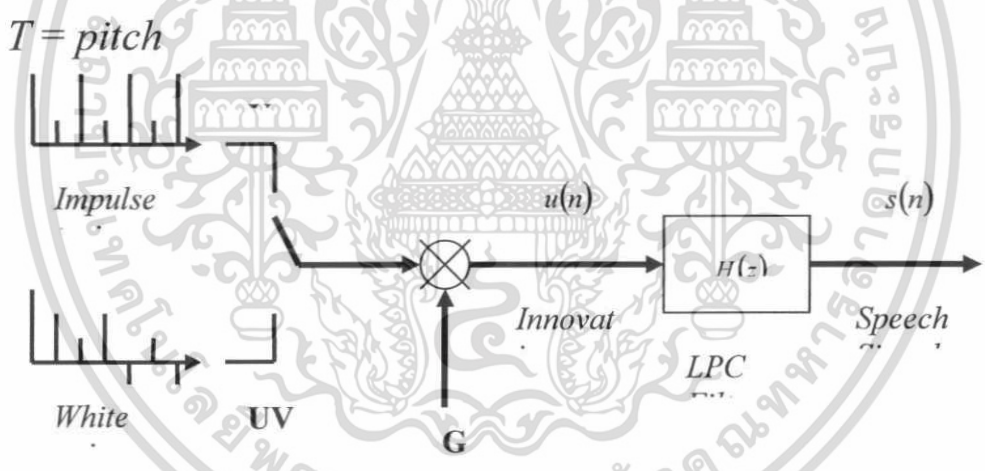
รูปที่ 2.12 คุณลักษณะการคอมเพรสชันของไดโอด



รูปที่ 2.13 แบบอย่างคุณลักษณะของการคอมเพรสชัน

กรณีที่มี $\mu = 100$ จะใช้คอมเพรสเซอร์ตามรูปที่ 2.11 แต่กรณีที่มี $\mu = 255$ และ $A = 87.6$ จะใช้วงจรคอมเพรสเซอร์ที่มีคุณลักษณะเป็นเส้นตรง โดยแยกออกเป็นส่วนๆ เรียกว่าเซกเมนต์ (segment)

2.2.5 การสังเคราะห์เสียงเชิงคณิตศาสตร์



รูปที่ 2.14 แสดงการสังเคราะห์สัญญาณเสียง

จากรูปที่ 2.14 แสดงแบบจำลองทางคณิตศาสตร์สำหรับการสังเคราะห์เสียง โดยอธิบายได้ดังนี้

- สัญญาณเสียงที่สังเคราะห์ขึ้นได้จากการนำสัญญาณเสียงที่มีลักษณะเป็นคาบ (voiced) และสัญญาณเสียงที่ไม่เป็นคาบ (unvoiced) มาทำการมอดูเลชัน และผ่านวงจรกรอง
- เปรียบเทียบความสัมพันธ์ระหว่างรูปแบบทางกายภาพของเสียงกับรูปแบบทางคณิตศาสตร์ได้ดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.2 ความสัมพันธ์ระหว่างรูปแบบทางกายภาพกับรูปแบบทางคณิตศาสตร์ของสัญญาณเสียง

รูปแบบทางกายภาพ	รูปแบบทางคณิตศาสตร์
ลม (Air)	$u(n)$ (Innovations)
การสั่นของเส้นเสียง (Vocal Cord Vibration)	V (Voiced)
คาบการสั่นของเส้นเสียง (Vocal Cord Vibration Period)	T (Pitch period)
เสียงที่ไม่มีการสั่นของเส้นเสียง (Fricatives and Plosives)	UV (Unvoiced)
ปริมาณของลมที่ออกจากปอด (Air Volume)	G (Gain)

1. การหาสัมประสิทธิ์ของวงจรรอง

แนวคิดของการหาสัมประสิทธิ์ของวงจรรอง LP คือการประมาณให้ตัวอย่างของเสียงที่เวลาใดๆ $s(n)$ สามารถประมาณค่าได้จากการทำความเป็นไปได้เชิงเส้น (linear combination) ของตัวอย่างสัญญาณเสียงอินพุตก่อนหน้าจำนวน p ตัวอย่าง ดังสมการที่ (2.13)

$$s(n) \approx a_1 s(n-1) + a_2 s(n-2) + \dots + a_p s(n-p) \quad (2.13)$$

ซึ่งสัมประสิทธิ์ของ a_1, a_2, \dots, a_p นั้นถือว่ามีค่าคงที่ในช่วงเฟรมที่ทำการวิเคราะห์ เราสามารถแปลงสมการที่ (2.13) จากการประมาณให้เป็นเท่ากับได้โดยเพิ่มสัญญาณกระตุ้นเข้าปอดดังสมการที่ (2.14)

$$s(n) = \sum_{i=1}^p a_i s(n-i) + Gu(n) \quad (2.14)$$

โดยที่ $u(n)$ เป็นสัญญาณกระตุ้นที่ถูกทำให้เป็นหนึ่งหน่วย (normalize) แล้ว เมื่อ G เป็นอัตราขยายของสัญญาณกระตุ้น ซึ่งหากพิจารณาสมการดังกล่าวบนโดเมนของ Z จะได้ความสัมพันธ์ดังนี้

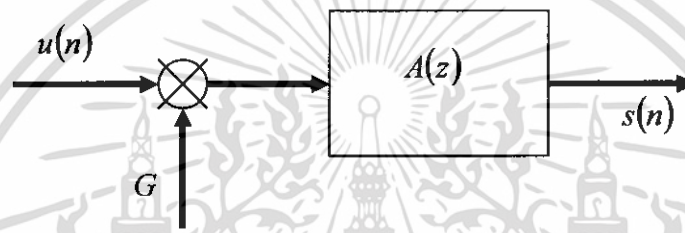
$$S(z) = \sum_{i=1}^p a_i z^{-i} S(z) + GU(z) \quad (2.15)$$

ทำให้เราได้ฟังก์ชันถ่ายโอนเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$H(z) = \frac{S(z)}{GU(z)} = \frac{1}{1 - \sum_{i=1}^p a_i z^{-i}} = \frac{1}{A(z)} \quad (2.16)$$

ความหมายของสมการนี้แสดงได้ด้วยรูปที่ 2.15 ซึ่งแหล่งกำเนิดของสัญญาณกระตุ้นปล่อยสัญญาณกระตุ้น $u(n)$ ออกมา สัญญาณจะถูกขยายด้วยค่า G และกลายเป็นสัญญาณอินพุตที่ป้อนให้กับวงจรกรอง เพื่อสร้างสัญญาณเสียง $s(n)$ ออกมา จากความรู้ในปัจจุบันเราพบว่าสัญญาณกระตุ้นของเสียงพูดนั้นเป็นแบบพัลส์เทรน (pulse train) หรือสัญญาณเสียงที่เป็นคาบ (voiced sound) หรือเป็นสัญญาณแบบสุ่ม (unvoiced sound) ส่วนวงจรกรองนั้นเป็นการทำงานที่ผสมผสานกันของอวัยวะต่างๆที่ใช้ออกเสียง (ช่องคอ, ทางเดินอากาศ, ช่องปาก และลิ้น)



รูปที่ 2.15 บล็อกไดอะแกรมของการแปลงเสียง

การประมาณความเป็นไปได้เชิงเส้นของสัญญาณเสียงในอดีต โดยนิยามค่าของสัญญาณที่ประมาณจะแสดงด้วยสมการที่ (2.17) ดังนี้

$$\tilde{s}(n) = \sum_{i=1}^p a_i s(n-i) \quad (2.17)$$

ค่าความผิดพลาดในการประมาณ $e(n)$ แสดงได้โดยสมการที่ (2.18)

$$e(n) = s(n) - \tilde{s}(n) = s(n) - \sum_{k=1}^p a_k s(n-k) \quad (2.18)$$

โดยที่มีค่าความผิดพลาดของฟังก์ชันถ่ายโอน (error transfer function) ดังสมการที่ (2.19)

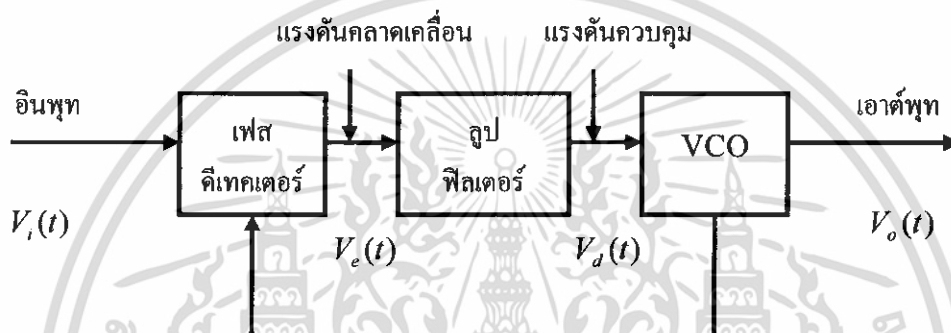
$$A(z) = \frac{E(z)}{S(z)} = 1 - \sum_{k=1}^p a_k z^{-k} \quad (2.19)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นเมื่อให้ $s(n)$ เป็นสัญญาณจริง ที่ถูกสร้างขึ้นมาจากระบบเชิงเส้นตามสมการที่ (2.14) จะพบว่า $e(n)$ จะมีค่าเท่ากับสัญญาณกระตุ้นที่ถูกขยายแล้ว $Gu(n)$

2.2.6 เฟสล็อกกลุ๊ป

เฟสล็อกกลุ๊ปเป็นระบบป้อนกลับ(feedback system) ที่บังคับให้วงจรออสซิลเลเตอร์มีความถี่หรือเฟสเปลี่ยนแปลงไปตามความถี่ หรือเฟสของสัญญาณอ้างอิงภายนอก เฟสล็อกกลุ๊ปประกอบด้วยภาคสำคัญ 3 ภาค คือ ภาคเทียบเฟสหรือเฟสดีเทคเตอร์ ภาคฟิลเตอร์ และภาค VCO (Voltage Controlled Oscillator) ดังรูปที่ 2.16



รูปที่ 2.16 แผนผังของเฟสล็อกกลุ๊ป

สมมุติว่าสัญญาณความถี่อ้างอิงภายนอกเป็นสัญญาณรายคาบมาที่อินพุต ภาคเทียบเฟสทำหน้าที่เปรียบเทียบเฟสระหว่างสัญญาณอ้างอิงกับสัญญาณจาก VCO เอาต์พุตที่ได้จากภาคเฟสดีเทคเตอร์จะเป็นแรงดันที่มีแอมพลิจูดเป็นสัดส่วนกับผลต่างในเฟสของสัญญาณทั้งสอง ที่ทำการเปรียบเทียบแรงดันผลต่างนี้ป้อนไปให้วงจรฟิลเตอร์ซึ่งเป็นฟิลเตอร์ชนิดกรองความถี่ต่ำผ่านเอาต์พุตเฉพาะความถี่ต่างๆที่ต้องการ เพื่อส่งไปควบคุมการออสซิลเลเตอร์ของ VCO ต่อไป

เมื่อลู๊ปอยู่ในสถานะล็อก (lock) ความถี่ของ VCO จะเท่ากับความถี่ของสัญญาณอินพุตพอดี อาจจะมีเฟสแตกต่างกันไป แต่ค่าเฟสที่แตกต่างจะมีค่าคงที่ (constant phase difference) ในกรณีที่มีเฟสไม่ตรงกันภาคเฟสดีเทคเตอร์จะจ่ายแรงดันคลาดเคลื่อน (error voltage) ไปควบคุมการทำงานของ VCO เพื่อมิให้เฟสคลาดเคลื่อนจนกว่าจะเข้าสู่สถานะล็อก เอาต์พุตของ VCO จึงมีแอมพลิจูดคงที่เสมอ แต่ความถี่จะเปลี่ยนแปลงตามความถี่ของสัญญาณอินพุต

จากที่ได้กล่าวมาแล้ว เราอาจจะอธิบายการทำงานในลักษณะสมการทางไฟฟ้าได้อีกดังนี้ คือ เมื่อไม่มีสัญญาณเข้า ($V_i = 0$) แรงดัน V_d และ V_i (ที่ใช้ควบคุม VCO) จะเป็นศูนย์ VCO จะกำเนิดสัญญาณความถี่หนึ่งเรียกว่า ความถี่อิสระ (free running frequency) f_{∞} เมื่อใส่สัญญาณเข้าที่ V_i ที่วงจรกรองความถี่ต่ำผ่าน แรงดัน V_d ที่ใช้ควบคุม VCO จะยังคงเป็นศูนย์ และ VCO ยังคงทำงานที่ความถี่ f_{∞} ถ้า f_i ไม่ต่างจาก f_{∞} มากนัก นั่นคือ $|f_i - f_{\infty}| \leq \Delta f_p$ ($2\Delta f_p$ มีชื่อว่าช่วงดึง

เข้าหรือ pull-in range) สัญญาณ V_c จะมีลักษณะไม่สมมาตร กล่าวคือมีองค์ประกอบไฟตรงไม่เท่ากับ 0 V_d จะค่อยๆเพิ่มค่าขึ้นและหลังจากช่วงเวลาหนึ่งเรียกว่า เวลาดึงเข้า (pull-in time) ความถี่ f_d ที่ค่อยๆปรับไปยังค่า f_i จะต่างจาก f_i เพียง $|f_i - f_\infty| = \Delta f_p$ ($2\Delta f_p$ มีชื่อว่าช่วงจับหรือ capture range) ถึงตอนนี้ f_o จะแปรเข้าสู่ค่า f_i อย่างรวดเร็วจับพลาตัน เรียกว่าเกิดการล็อก V_d ก็จะเปลี่ยนแปลงจับพลาตันเช่นกัน ทั้งนี้สำหรับ VCO โดยทั่วไปความถี่ f_o เป็นปฏิภาคกับการบ่ายเบนจากความถี่อิสระ ดังสมการที่(2.20) นั่นคือ

$$f_o - f_\infty = k_o V_d k_o V_d \quad (2.20)$$

โดยที่ค่า k_o เรียกว่าความไวของ VCO มีหน่วยเป็น $\frac{rad/s}{V}$

การได้มาซึ่งการล็อกอาจใช้วิธีได้ดังนี้

1. ณ ขณะหนึ่ง ความถี่ของ VCO ต่างจากความถี่ขาเข้า f_i น้อยกว่า Δf_c จะเกิดการล็อกโดยจับพลาตันเรียกว่า เกิดการจับ (capture)
2. ณ ขณะหนึ่ง ความถี่ของ VCO ต่างจากความถี่ขาเข้า f_i มากกว่า Δf_c แต่ไม่น้อยกว่า Δf_p ความถี่ของ VCO จะค่อยๆเลื่อนเข้าหาความถี่ f_i เรียกว่า เกิดการดึงเข้า (pull-In)
3. นอกช่วงดึงเข้าหรือในกรณีที่การดึงเข้าใช้เวลานานเกินไป อาจเพิ่มวงจรเพื่อควบคุมให้ความถี่ของ VCO กว้างไปเพื่อเสาะหาความถี่ของสัญญาณ หรือ ถ้าเสียรบกวนน้อย ก็อาจเพิ่มแถบความถี่ของรอบ (tracking) ความถี่ของ VCO จะสามารถตามรอยความถี่ f_i ได้ครบโดที่ $|f_i - f_\infty| = 2\Delta f_H$ ($2\Delta f_H$ มีชื่อว่าช่วงคงไว้หรือช่วงล็อก)

เราสามารถนำเฟสล็อกกลูไปใช้ในการสังเคราะห์ หรือผลิตความถี่ที่มีความเที่ยงตรง และมีเสถียรภาพเทียบเท่าสัญญาณอ้างอิงได้ วงจรนี้เรียกว่า วงจรสังเคราะห์ความถี่ ระบบสังเคราะห์ความถี่จะช่วยให้เราสามารถสังเคราะห์สัญญาณเอาต์พุตจาก VCO ให้มีความถี่ตามต้องการได้หลายความถี่โดยมีความเที่ยงตรงและเสถียรสูงเทียบเท่าคริสตอลออสซิลเลเตอร์

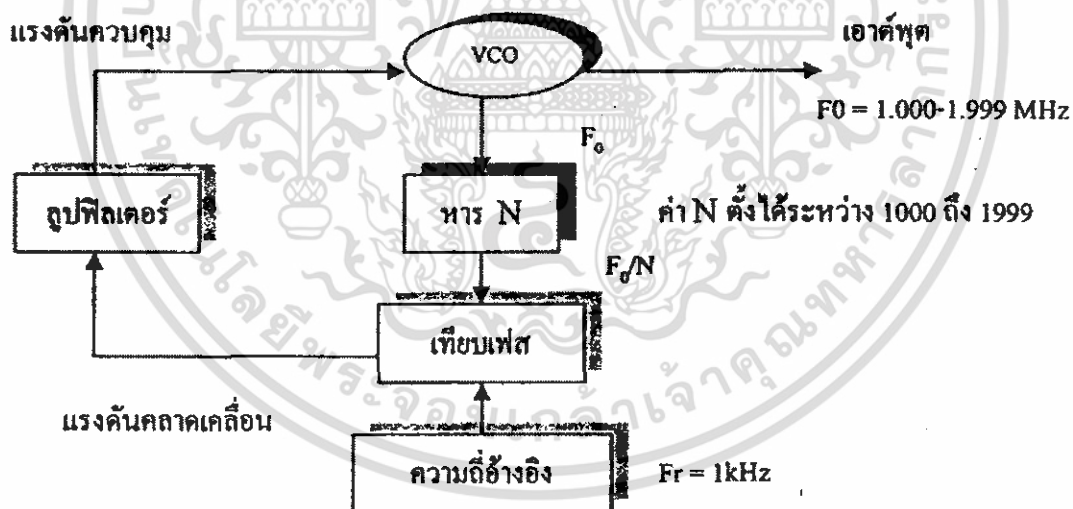
1. คุณสมบัติของวงจรสังเคราะห์ความถี่

นอกจากวงจรสังเคราะห์ความถี่จะต้องมีคุณสมบัติเกี่ยวกับช่วงความถี่ (frequency range) ที่ต้องผลิตและเรโซลูชันระหว่างขั้นแล้ว คุณสมบัติอื่นๆของวงจรสังเคราะห์ความถี่ก็มีความสำคัญเช่นกัน

โดยปกติวงจรสังเคราะห์ความถี่จะสามารถกำเนิดสัญญาณเพียงสัญญาณเดี่ยวแต่เลือกความถี่ได้หลายค่า ในช่วงความถี่ใช้งาน และมีความละเอียดของความถี่ขึ้นอยู่กับเรโซลูชัน ในกรณีที่เราเปลี่ยนหลายค่า ในช่วงความถี่ใช้งาน และมีรายละเอียดของความถี่ขึ้นอยู่กับเรโซลูชัน ในกรณีที่เราเปลี่ยนความถี่จากค่าหนึ่งไปยังอีกค่าหนึ่ง วงจรสังเคราะห์ความถี่จะต้องเปลี่ยนตามได้เร็วทันที กล่าวอีกอย่างหนึ่งคือ ล็อกความถี่ได้ในเวลาอันรวดเร็ว นั่นคือ ช่วงเวลาล็อก (lock-up time) สั้น คุณสมบัติการล็อกความถี่ใหม่ได้รวดเร็วนั้นมีความจำเป็นอย่างยิ่ง

วงจรสังเคราะห์ความถี่ที่ดี จะต้องสามารถผลิตสัญญาณความถี่เดี่ยว โดยปราศจากความถี่แปลกปลอมต่างๆ คุณสมบัตินี้เรียกว่า ความบริสุทธิ์ของสเปกตรัม (spectrum purity) นั่นคือความถี่ฮาร์โมนิกและสปีวเรียดต่างๆนั้น จะต้องถูกกำจัดออกให้เหลือน้อยที่สุด นอกจากนี้ข้อจำกัดจากวงจรออสซิลเลเตอร์จะทำให้วงจรสังเคราะห์ความถี่ไม่บริสุทธิ์ ไม่ใช่เพียงความถี่เดี่ยว ในช่วงใกล้เคียงกับความถี่ที่ต้องการ นอยส์ดังกล่าวนี้เรียกว่า เฟส นอยส์ (phase noise)

ความเที่ยงตรง (accuracy) และเสถียรภาพ (stability) ทางความถี่ของวงจรสังเคราะห์ความถี่ขึ้นอยู่กับสัญญาณอ้างอิง โดยทั่วไปสัญญาณอ้างอิงมักจะเป็นวงจรออสซิลเลเตอร์ชนิดใช้แร่บัตั้วความถี่ ฉะนั้นวงจรสังเคราะห์ความถี่จะมีเสถียรภาพ และความเที่ยงตรงเทียบเท่ากับคริสตอลออสซิลเลเตอร์



รูปที่ 2.17 แผนผังของหน่วยสังเคราะห์ความถี่

2. การใช้เฟสล็อกคูลในการสังเคราะห์ความถี่

แม้ว่าระบบสังเคราะห์ความถี่จะมีความซับซ้อนเพียงใด เมื่อพิจารณาถึงลงไปแล้วจะพบว่า มีเฟสล็อกคูลเป็นหัวใจในการสังเคราะห์ความถี่เสมอ รูปที่ 2.17 เป็นตัวอย่างของระบบสังเคราะห์ความถี่อย่างง่าย ประกอบด้วย 5 ภาค คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. VCO เป็นออสซิลเลเตอร์กำเนิดสัญญาณเอาต์พุทของระบบสังเคราะห์ความถี่
2. ภาควิทยาศาสตร์อิเล็กทรอนิกส์คริสตอลจากออสซิลเลเตอร์หรือสัญญาณอื่นๆ
3. ภาคเทียบเฟส
4. ภาคความถี่อ้างอิง
5. ภาคฟิลเตอร์ซึ่งทำหน้าที่กรองเอาเฉพาะความถี่ค่าไปใช้

ภาคเฟสดีเทกเตอร์

จากรูปจะเห็นว่าสัญญาณอินพุทของภาคเทียบเฟสจาก 2 แหล่งคือ

1. จาก VCO มีความถี่เท่ากับ F_o / N
2. จากสัญญาณอ้างอิงมีความถี่เท่ากับ F_o

เอาต์พุทที่ได้จากการเปรียบเทียบเฟสก็คือ ผลต่างระหว่างสัญญาณ F/N กับ F ซึ่งจะกรองเอาเฉพาะความถี่ค่าเท่านั้น เพื่อบังคับการออสซิลเลตของวงจร VCO ให้ทำการปรับแก้ความถี่ (หรือเฟส) ให้ตรงจนกว่าความถี่ของสัญญาณทั้งสองจะเท่ากันในสภาวะล็อกความถี่ของ VCO เพื่อผ่านวงจรหาร จะเท่ากับความถี่อ้างอิง นั่นคือ

$$F = N \times F_o \quad (2.21)$$

ภาคออสซิลเลเตอร์แบบควบคุมความถี่ด้วยระดับแรงดัน (V_{cc})

วงจรวจร VCO นับว่ามีความสำคัญมากในระบบเฟสล็อกลูป เพราะจะเป็นตัวกำหนดเสถียรภาพเชิงความถี่ จะต้องมีคุณสมบัติคือ ไปด้วยนี้คือ

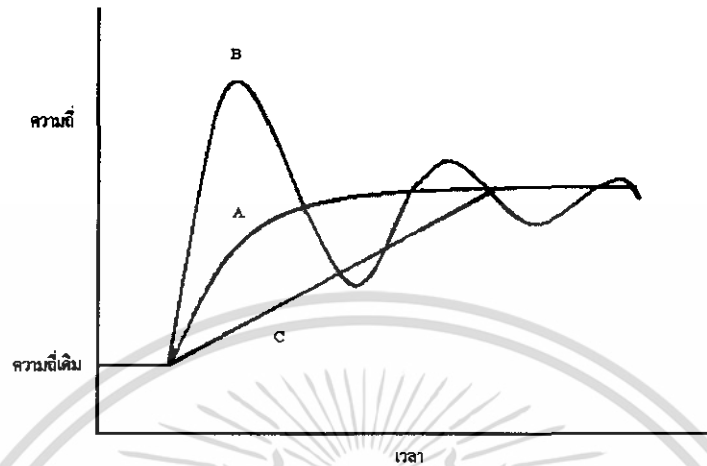
1. การแปลแรงดันเป็นความถี่มีลักษณะเป็นเชิงเส้น
2. มีเสถียรภาพเชิงความถี่
3. ทำงานได้ที่ความถี่สูง
4. ช่วงตามรอยกว้าง
5. ปรับคลื่นได้ง่าย

คุณสมบัติเหล่านี้มักไม่พบใน VCO ที่เป็นตัวเดียวกัน ดังนั้นเราต้องเลือกชนิดของ VCO ตามคุณสมบัติที่เราต้องการ

ภาคฟิลเตอร์

เป็นวงจรมีหน้าที่กรองความถี่ค่าผ่านทำหน้าที่กรองเอาเฉพาะความถี่ค่ามาควบคุมความถี่ของ VCO และในที่นี้เราใช้ฟิลเตอร์ประเภทพาสซีฟ ฟิลเตอร์นี้เป็นตัวกำหนดคุณสมบัติการเปลี่ยนแปลงความถี่ก่อนเข้าสู่สภาวะล็อกที่เรียกว่า คุณสมบัติชั่วคราว ถ้าเลือกอัตราขยายและค่าคงตัวของฟิลเตอร์ไม่เหมาะสม ความถี่ของเฟสล็อกลูปจะไม่ล็อกและจะเปลี่ยนไปเปลี่ยนมา ดังนั้น ค่าคงตัวของฟิลเตอร์จะต้องไม่มากเกินไป เพื่อว่าทุกครั้งที่เปลี่ยนความถี่เฟสล็อกลูปจะล็อกได้ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เร็ว โดยไม่มีการสละหรือใช้เวลาเปลี่ยนความถี่อย่างรวดเร็ว แต่ค่าคงตัวเวลาก็ไม่ควรจะน้อยเกินไปจนกระทั่งความถี่สั่นหรือไม่นิ่ง



รูปที่ 2.18 คุณลักษณะ ช่วงพลวัต (dynamic range) ในการเปลี่ยนความถี่เฟสล็อกกลูป

รูปที่ 2.18 ซึ่งแสดงการเปลี่ยนความถี่ของ VCO จะเห็นว่าเส้นทางการเปลี่ยนแปลงแรงดันมี 3 เส้นทาง เส้นทาง A เป็นเส้นทางหน่วงวิกฤต (critically damped) ใช้เวลาในการเปลี่ยนสู่ความถี่ใหม่ น้อยที่สุด เส้นทาง B เรียกว่า เส้นทางหน่วงขาด (undamped) มีการสละ (หรือ ออสซิลเลต) เนื่องจากจากโอเวอร์ชูต เส้นทาง C เป็นเส้นทางหน่วงเกิน (overdamped) ไม่มีโอเวอร์ชูตแต่เวลาที่ใช้ในการเข้าสู่ความถี่ใหม่จะช้า เส้นทาง A เป็นเส้นทางที่ดีที่สุดในการออกแบบค่าคงตัวเวลาของลูปฟิลเตอร์ เพราะใช้เวลาเปลี่ยนความถี่เร็ว และไม่มีโอเวอร์ชูต

ภาคหาร N

ภาคหาร N (programmable divider) เวลาที่ใช้ในการล็อกความถี่เมื่อมีค่าน้อยที่สุดจะไม่เท่ากับเมื่อมีค่ามากที่สุด วงจรหารเกิดจากวงจรนับฐานสิบ (decade counter) หลายชุดมาต่อรวมกับเกตต่างๆ เพื่อให้สามารถเลือกสั่งให้วงจรนับทำหน้าที่หารความถี่ได้ตามตัวเลขที่ตั้งไว้ วงจรหาร N นี้เป็นตัวรับคำสั่งเกี่ยวกับความถี่ไปควบคุม VCO เพื่อให้กำเนิดสัญญาณตามที่ต้องการ ตัว N จะเป็นตัวที่กำหนดย่านความถี่และจำนวนช่องความถี่

ลักษณะการป้อนข้อมูล N ให้กับวงจรหาร N แบบขนาน (parallel) คือข้อมูลแต่ละบิตจะป้อนเข้าพร้อมๆกัน ส่วนลักษณะการป้อนข้อมูล N ให้กับวงจรหาร N แบบอนุกรมแต่ละบิตจะป้อนเข้าไม่พร้อมกัน วงจรหารประเภทนี้มีความซับซ้อนมากกว่า เพราะต้องมีสัญญาณนาฬิกา มีวงจรแลตซ์ ในการป้อนข้อมูลวงจรหาร N ประเภทนี้จะควบคุมการทำงานด้วยไมโครคอมพิวเตอร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัญหาสำคัญของซินธิไซเซอร์อีกอย่างหนึ่งก็คือ วงจรหาร N (หรือวงจรหารที่โปรแกรมได้) ไม่สามารถทำงานที่ความถี่สูงกว่า 25 MHz ได้ ฉะนั้นเราจึงต้องหาทางลดทอนความถี่ที่ป้อนแก่ วงจรหาร N ลง เพื่อให้วงจรล็อกของวงจรหาร N ทำงานได้ วิธีต่างๆที่นิยมใช้ได้แก่ ใช้ความถี่จาก ออสซิลเลเตอร์พิเศษ (บางครั้งเรียกว่า ออสซิลเลเตอร์ PLL) มามิกซ์กับ VCO ให้ความถี่ลดลงก่อนที่ จะป้อนให้แก่วงจรหาร

ภาคกำเนิดความถี่อ้างอิง

ภาคกำเนิดความถี่อ้างอิงนิยมใช้คริสตอลออสซิลเลเตอร์ ความถี่ออสซิลเลเตอร์อ้างอิงนี้ เป็นตัวกำหนดร โขลูชันและเสถียรภาพของความถี่อ้างอิงที่ดี จึงทำให้สามารถสังเคราะห์ความถี่ที่มี เสถียรภาพที่ดีด้วย

3. การคีมอดูเลตสัญญาณแอฟเอ็มด้วยเฟสล็อกกลูป

เฟสล็อกกลูปเป็นระบบป้อนกลับ ประกอบด้วยเครื่องกำเนิดสัญญาณที่มีความถี่เปลี่ยนแปลง ตามระดับแรงดันขาเข้า (VCO) และส่วนเปรียบเทียบเฟส (phase comparator)

สัญญาณเข้าจะเปรียบเทียบเฟสกับสัญญาณอ้างอิง (reference signal) ที่กำเนิดจาก VCO ส่วนเปรียบเทียบเฟสจะให้แรงดันขาออกที่มีระดับขึ้นกับความแตกต่างระหว่างเฟสของสัญญาณเข้า ทั้งสอง เมื่อผ่านเครื่องกรองความถี่ต่ำผ่านสัญญาณขาออกบางส่วนจะถูกป้อนกลับเพื่อควบคุมการ กำเนิดความถี่ของ VCO จนกระทั่งความถี่ทั้งสองสัญญาณเท่ากัน แรงดันที่ป้อนกลับจะคงที่และทำให้ ความถี่ทั้งสองเกาะกันไป โดยมีเฟสต่างกันที่ค่าคงที่ความถี่กลาง (center frequency หรือ free-running frequency)

ต่อไปนี้จะอธิบายส่วนประกอบที่สำคัญของเฟสล็อกกลูป

ส่วนเปรียบเทียบเฟส มีอยู่ 2 ชนิดคือ ชนิดแอนะล็อก และชนิดดิจิทัล ชนิดแอนะล็อก สร้างขึ้นโดยใช้ตัวผสม (mixer) หรือตัวคูณ (multiplier) ดังนี้

$$V_1(t) = A \sin(\omega_1 t + \theta_1) \quad (2.22)$$

$$V_2(t) = A \sin(\omega_2 t + \theta_2) \quad (2.23)$$

โดย $V_1(t)$ และ $V_2(t)$ เป็นสัญญาณขาเข้า ซึ่งจะได้สัญญาณขาออก $V_o(t)$ ของการคูณดังนี้

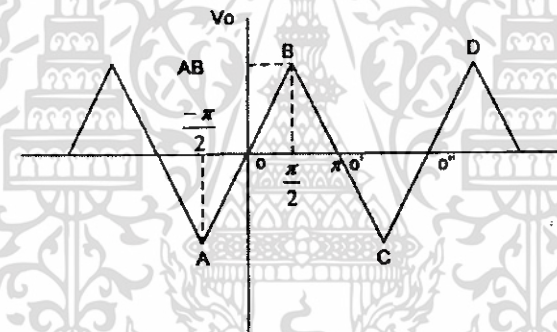
$$V_o(t) = \frac{1}{2} AB \cos[(\omega_1 + \omega_2)t + (\theta_1 + \theta_2)] + \frac{1}{2} AB \cos[(\omega_1 - \omega_2)t + (\theta_1 - \theta_2)] \quad (2.24)$$

เมื่อผ่านเครื่องกรองความถี่ต่ำผ่านก็จะได้สัญญาณขาออก ที่ขึ้นอยู่กับความแตกต่างของเฟสซึ่งมีลักษณะสมบัติ ดังรูปที่ 2.19



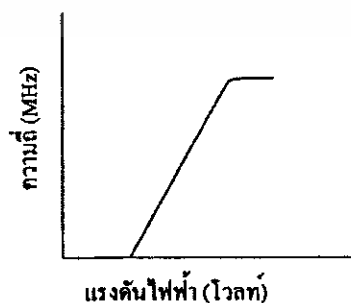
รูปที่ 2.19 ลักษณะสมบัติของส่วนเปรียบเทียบเฟสชนิดแอนะล็อก

ถ้าสำหรับอุปกรณ์เปรียบเทียบชนิดดิจิทัลนั้น สร้างขึ้นได้หลายวิธี เช่น ใช้ เอ็คคลูซีฟ-ออร์ หรืออุปกรณ์ชนิดมีความจำ เช่น ใช้ ฟลิป-ฟลอป ซึ่งจะมีลักษณะสมบัติเช่นเดียวกับแบบแอนะล็อก แต่จะต่างกันที่เป็นเส้นตรง ดังรูปที่ 2.20



รูปที่ 2.20 ลักษณะสมบัติ ส่วนเปรียบเทียบเฟสชนิดดิจิทัล

VCO สามารถสร้างขึ้นได้หลายวิธี ไดโอดคาปาซิเตนซ์ชนิดปรับค่าแรงดัน (voltage variable capacitance diode) ซึ่งจะใช้กันดีในชื่อ วาแรคเตอร์ (varactor) หรือ มัลติไวเบรเตอร์ (multivibrator) ชนิดต่างๆ ที่สามารถควบคุมความถี่ได้ด้วยแรงดัน รูปที่ 2.21 เป็นลักษณะสมบัติทั่วไปของ VCO



รูปที่ 2.21 ลักษณะสมบัติของ VCO

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมมุติให้สัญญาณเฟรมเข้าที่ ส่วนเปรียบเทียบเฟส คือ

$$A \cos[\omega_c t + \psi(t)]$$

ให้ความไวทางความถี่ (Frequency Sensitivity) ของ VCO คือ K_0 rad/sec คือค่าของความแตกต่างหรือการเปลี่ยนแปลงของความถี่เชิงมุมขณะใดขณะหนึ่ง ω_1 ที่เกิดจากการเปลี่ยนแปลงแรงดัน V ที่ขาเข้าของ VCO ดังสมการที่ (2.25) นั่นคือ

$$K_0 = \frac{d\omega_1}{dv} \quad (2.25)$$

สัญญาณนี้จะกระตุ้นให้เกิดสัญญาณจากเครื่องสั่นสะท้อนความถี่ (Oscillator) ใน VCO ดังสมการที่ (2.26)

$$B \cos[\omega_c t + K_0 \int_{-\infty}^t V_0(\lambda) d\lambda] \quad (2.26)$$

2.3 สถาปัตยกรรมไมโครคอนโทรลเลอร์ dsPIC30F6014

สำหรับโครงการนี้ได้เลือกใช้ไมโครคอนโทรลเลอร์ dsPIC30F6014 ในการประมวลผลสัญญาณเสียง dsPIC คือ ไมโครคอนโทรลเลอร์ 16 บิต และไมโครคอนโทรลเลอร์นี้ ถูกจัดให้อยู่ในอนุกรม DSC (Digital Signal Controller) ซึ่งเป็นไมโครคอนโทรลเลอร์ที่ได้รับการออกแบบมาเป็นพิเศษเพื่องานประมวลผลสัญญาณดิจิทัลสำหรับสร้างระบบควบคุมอัตโนมัติที่มีความสามารถสูง และมีโมดูลพิเศษต่างๆ ที่ช่วยในการประมวลผลทางด้านเสียง

2.3.1 คุณสมบัติโดยรวมของ dsPIC30F6014

ในหัวข้อนี้จะขออธิบายถึงคุณสมบัติและความสามารถต่างๆ ภายในไมโครคอนโทรลเลอร์ รวมถึง โมดูลพิเศษที่ใช้สำหรับเชื่อมต่อกับอุปกรณ์ภายนอกในการทำโครงการนี้

1. คุณสมบัติของซีพียู

- เป็นไมโครคอนโทรลเลอร์ที่ใช้ ซีพียู แบบ RISC (Reduced Instruction Set Computer : RISC) ถูกออกแบบมาให้มีคำสั่งในการทำงานจำนวนน้อย ทำให้โครงสร้างภายในไมโครคอนโทรลเลอร์ สามารถประมวลผลได้อย่างมีประสิทธิภาพ อีกทั้งคำสั่งในการทำงานไม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซับซ้อน ทำให้สามารถทำงานได้รวดเร็ว สามารถใช้กับสัญญาณนาฬิกาความถี่สูงมากๆ ได้ การทำงานถูกออกแบบมาให้ใช้รีจิสเตอร์ภายใน จึงจำเป็นต้องมีการถ่ายเทข้อมูลจากหน่วยความจำไปไว้ในรีจิสเตอร์ก่อน

- ออกแบบด้วย สถาปัตยกรรมแบบฮาร์วาร์ดที่ถูกปรับปรุงใหม่ (modified harvard architecture) เป็นโพรเซสเซอร์ที่สร้างขึ้นมาให้เหมาะกับการประมวลผลสัญญาณแบบเรียลไทม์ มีการแยกบัสระหว่างบัสข้อมูล กับบัสคำสั่งโปรแกรม เพื่อหลีกเลี่ยงการแย่งไปป์ไลน์ (pipeline) ที่เป็นสาเหตุให้ระบบช้าลง โดยปกติแล้วสถาปัตยกรรมฮาร์วาร์ดเป็นโครงสร้างของ DSP (Digital Signal Processor : DSP)
- คำสั่งการคำนวณและบวก (Multiply and Accumulation command : MAC) มีการทำงานแบบไปป์ไลน์ ทำให้ไม่เกิดการว่างการทำงานเนื่องจากต้องรอผลของส่วนอื่น จึงทำให้การคำนวณการแปลงฟูเรียร์แบบเร็วทำได้เร็วขึ้น
- มี 84 คำสั่งมาตรฐาน สามารถรองรับรูปแบบการอ้างแอดเดรสได้อย่างอิสระ
- ชุดคำสั่งมีขนาด 24 บิต สามารถประมวลผลข้อมูลได้ 16 บิต
- มีขนาดพื้นที่หน่วยความจำโปรแกรมแบบแฟลช ความจุ 144 Kbytes ลบและเขียนใหม่ได้ไม่น้อยกว่า 10,000 ครั้ง สามารถโปรแกรมหน่วยความจำโปรแกรมได้โดยทางซอฟต์แวร์ บรรจุชุดคำสั่งได้จริง 48 K words
- มีขนาดพื้นที่หน่วยความจำข้อมูลทั้งหมด มีขนาด 64 Kbytes (32K words) ภายในจะมีพื้นที่หน่วยความจำข้อมูลแรม 8 Kbytes แบ่งออกเป็น 2 ส่วนได้แก่ หน่วยความจำข้อมูลแรม X และ Y แต่ละส่วนมีขนาด 16 บิต ความจุ 4 Kbytes ในแต่ละส่วนจะมีตัวกำหนดแอดเดรส (Address Generation Unit : AGU) แยกออกจากกัน คำสั่งคูณและบวกของชุดคำสั่งการประมวลผลสัญญาณดิจิทัล (DSP instructions) จาก 2 แหล่งสามารถทำงานผ่าน ตัวกำหนดแอดเดรส X และ Y
- มีหน่วยความจำข้อมูลแรม 8 Kbytes
- ความเร็วในการทำงานสูงถึง 30 ล้านคำสั่งต่อวินาที (Million Instruction Per Second : MIPS)
- มีรีจิสเตอร์ W จัดอยู่ในรูปแบบอะเรย์ (working register array) มีขนาด 16 บิต จำนวน 16 ตัว เป็นรีจิสเตอร์หลักที่ใช้ในการทำงาน เนื่องจากโครงสร้างเป็นอะเรย์จึงสามารถรองรับทั้งข้อมูล, ค่าแอดเดรส หรือค่าของรีจิสเตอร์ใดๆ ที่ต้องนำมาประมวลผล รีจิสเตอร์ W ตัวที่ 0-14 (W0 – W14)ใช้ในการประมวลผลหลัก ส่วนรีจิสเตอร์ W ตัวที่ 15 (W15) ใช้ทำงานร่วมกับตัวชี้สแต็ก (software stack pointer) ในการทำงานของโปรแกรมย่อย และบริการอินเตอร์รัปต์ (interrupt service)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- รองรับแหล่งกำเนิดสัญญาณอินเทอร์รับต์ได้สูงสุด 41 แหล่ง โดยสามารถกำหนดระดับความสำคัญ (selectable priority levels) ได้ 8 ระดับ
- รับความถี่สัญญาณนาฬิกาจากภายนอกได้ ตั้งแต่ 4 MHz - 10 MHz ในกรณีใช้งานร่วมกับวงจรเฟสล็อกภายใน เลือกปรับ ตัวคูณได้ 3 ระดับคือ 4, 8 หรือ 16 เท่า

2. คุณสมบัติด้านการประมวลผลสัญญาณดิจิทัล

- มีวงจรเฟตช์ข้อมูลคู่ (Dual data fetch) ทำให้สามารถประมวลผลข้อมูลได้อย่างรวดเร็ว
- สามารถเข้าถึงหน่วยคำจำข้อมูลได้ 2 รูปแบบพิเศษ คือ แบบวนรอบ และ แบบกลับบิต
 - การเข้าถึงหน่วยความจำข้อมูล แบบวนรอบ (modulo addressing) สามารถใช้ได้ในพื้นที่หน่วยความจำข้อมูล X และ Y วิธีการนี้ใช้เพื่อลดการทำงานของแบบวนรอบ (loop program) สำหรับ DSP algorithms เป็นหลัก
 - การเข้าถึงหน่วยความจำข้อมูล แบบกลับบิต (bit-reversed addressing) สามารถใช้ได้ใน X AGU เท่านั้น เพื่อการเข้าถึงข้อมูลอย่างมีประสิทธิภาพ โดยการจัดเรียงข้อมูล เพื่อเข้าสู่ radix-2 FFT algorithms ทำได้ง่ายขึ้น
- มีแอมพลิจูด (accumulator) ขนาด 40 บิต 2 ตัว ใช้เก็บข้อมูลสำหรับประมวลผลทางคณิตศาสตร์
 - เมื่อประมวลผลเชิงเศษส่วน สามารถเก็บข้อมูลได้ตั้งแต่ -256.0 ถึง $256 \cdot 2^{-31}$
 - เมื่อประมวลผลเลขจำนวนเต็ม สามารถเก็บข้อมูลได้ตั้งแต่ $-549,755,813,888$ ถึง $549,755,813,887$
- มีหน่วยประมวลผลทางคณิตศาสตร์และลอจิก (Arithmetic Logic Unit : ALU) ขนาด 40บิต
- มีหน่วยประมวลผลด้านการคูณเลขจำนวนเต็มและเศษส่วน 17×17 บิต ในรูปของฮาร์ดแวร์ รองรับการคำนวณค่าตัวเลขทั้งที่คิดและไม่คิดเครื่องหมาย โดยบิตที่ 16 จะเป็นบิตเครื่องหมาย สำหรับการให้ผลลัพธ์สามารถให้ผลได้ทั้งเลขจำนวนเต็มขนาด 32 บิต และ เลขทศนิยม ใช้เวลาในการทำงานเพียง 1 รอบการทำงาน (Instruction Cycle Period)
 - เมื่อทำงานในโหมดเลขจำนวนเต็ม จะให้ผลลัพธ์เป็น ค่า two's complement แบบคิดเครื่องหมาย ค่าที่ได้จะอยู่ในช่วง -2^{N-1} to $2^{N-1} - 1$
 - สำหรับเลขจำนวนเต็มขนาด 16 บิต จะมีค่าอยู่ระหว่าง -32768 ($0x8000$) ถึง 32767 ($0x7FFF$)
 - สำหรับเลขจำนวนเต็มขนาด 32 บิต จะมีค่าอยู่ระหว่าง $-2,147,483,648$ ถึง $2,147,483,645$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- เมื่อทำงานในโหมดเลขเศษส่วน จะให้ผลลัพธ์เป็น ค่า two's complement ของเลขเศษส่วนแบบคิตรีงหมาย ค่าที่ได้จะอยู่ในช่วง -1.0 ถึง $(1 - 2^{1-N})$
- สำหรับเลขทศนิยมขนาด 16 บิต จะมีค่าอยู่ระหว่าง -1.0 ถึง 0.999969482 และมีค่าความละเอียด 3.01518×10^{-5}
- ชุดคำสั่งการประมวลผลสัญญาณดิจิทัลทั้งหมดใช้เวลาในการทำงานเพียง 1 รอบการทำงาน
- ตัวเลื่อนข้อมูลสองทิศทางแบบบาร์เรล (barrel shifter) ขนาด 40 บิต สามารถเลื่อนไปทางซ้ายหรือขวาได้ภายใน 1 รอบการทำงาน เมื่อได้รับข้อมูลที่กำหนดเครื่องหมาย โดยถ้าเป็นค่าบวก จะเลื่อนข้อมูลไปทางขวา ถ้าเป็นค่าลบ จะเลื่อนข้อมูลไปทางซ้าย ให้ผลลัพธ์ขนาด 40 บิต

3. คุณสมบัติด้านการเชื่อมต่อกับอุปกรณ์ภายนอก

- สามารถจ่ายกระแสออกทางขาพอร์ต ได้ 25 mA ทั้งแบบกระแสซิงค์และซอร์ส
- ไทมเมอร์/เคาน์เตอร์ ขนาด 16 บิต จำนวน 5 ตัว สามารถต่อใช้งานร่วมกันเป็น ไทมเมอร์ 32 บิต ได้
- มีส่วนเชื่อมต่อกับอุปกรณ์อนุกรมในแบบ SPI (Serial Peripheral Interface : SPI)
- มีส่วนเชื่อมต่อกับอุปกรณ์เข้ารหัสและถอดรหัสเสียง ในแบบ DCI (Data Converter Interface : DCI)

4. คุณสมบัติพิเศษด้านอื่นๆ

- มีพาวเวอร์อนรีเซต (POR), พาวเวอร์อัพไทมเมอร์ (PWRT) และ ออสซิลเลเตอร์สตาร์ทอัพไทมเมอร์ (OST)
- มีวอตช์ด็อกไทมเมอร์แบบโปรแกรมทางซอฟต์แวร์ได้ (Programmable Watchdog Timer : WDT) ทำหน้าที่รีเซตไมโครคอนโทรลเลอร์ในกรณีที่เกิดการทำงานทางซอฟต์แวร์ผิดพลาดจนไม่สามารถทำงานอื่นต่อได้
- รองรับการโปรแกรมในวงจรแบบอนุกรม (In-Circuit Serial Programming : ICSP)

5. การรองรับการหารเลข

รองรับการหารเลขอย่างเต็มรูปแบบ โดยไม่ต้องสร้างโปรแกรมย่อยขึ้นมาจัดการ แบ่งออกเป็น 5 รูปแบบ ดังนี้

1. การหารเลขเศษส่วนหรือทศนิยม 16 บิตแบบคิตรีงหมาย
2. การหารเลขจำนวนเต็มที่มีตัวตั้ง 32 บิต ด้วยตัวหาร 16 บิตแบบคิตรีงหมาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. การหารเลขจำนวนเต็มที่มีตัวตั้ง 32 บิต ด้วยตัวหาร 16 บิตแบบไม่คิดเครื่องหมาย
4. การหารเลขจำนวนเต็มที่มีตัวตั้ง 16 บิต ด้วยตัวหาร 16 บิตแบบคิดเครื่องหมาย
5. การหารเลขจำนวนเต็มที่มีตัวตั้ง 16 บิต ด้วยตัวหาร 16 บิตแบบไม่คิดเครื่องหมาย

6. ระบบสัญญาณนาฬิกา

- สามารถเลือกแหล่งกำเนิดสัญญาณนาฬิกาจากภายในหรือภายนอกได้
- มีวงจรเฟสล็อกอยู่ภายในเพื่อเพิ่มความถี่ในการทำงานของสัญญาณนาฬิกา ดังแสดงในตารางที่ 2.3
- มีส่วนตรวจสอบความผิดพลาดของสัญญาณนาฬิกาเพื่อเปลี่ยนแหล่งกำเนิดสัญญาณนาฬิกาได้

ตารางที่ 2.3 ย่านสัญญาณความถี่นาฬิกาที่สามารถใช้ร่วมกับวงจรเฟสล็อกภายใน

ความถี่ขาเข้า	ตัวคูณเฟสล็อก	ความถี่ขาออก
4 MHz -10 MHz	x4	16 MHz -40 MHz
4 MHz -10 MHz	x8	32 MHz -80 MHz
4 MHz -7.5 MHz	x16	64 MHz -120 MHz

2.3.2 โมดูลสำหรับเชื่อมต่ออุปกรณ์อนุกรม แบบ SPI

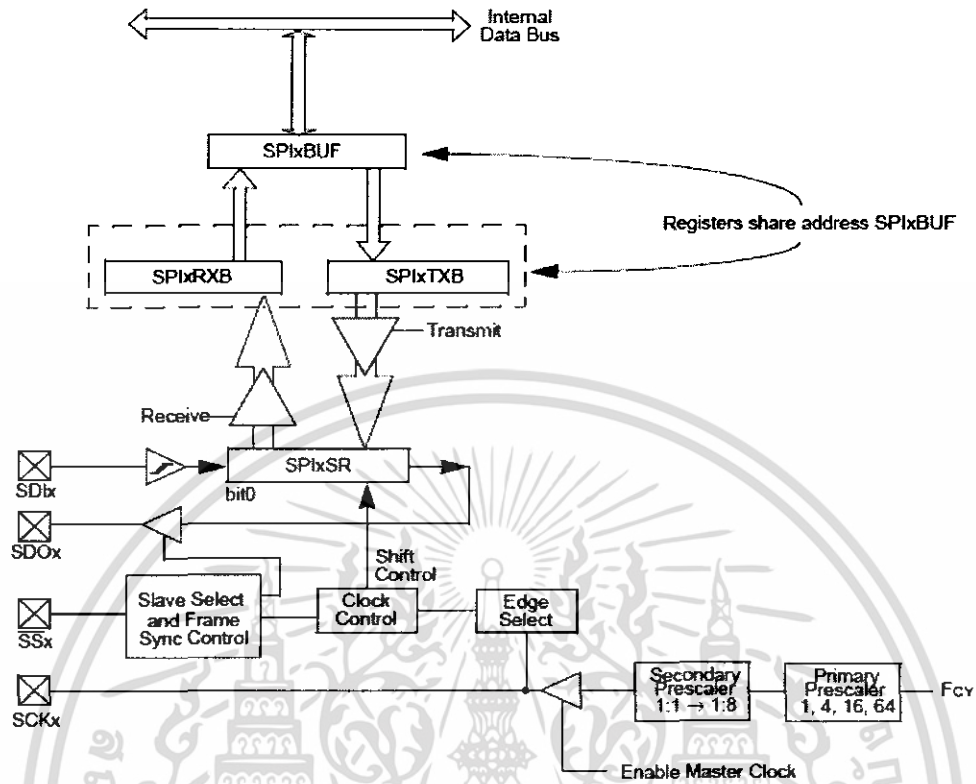
การติดต่อแบบ SPI เป็นการติดต่ออนุกรมแบบเป็นจังหวะ (synchronous serial interface) อย่างหนึ่ง ช่วยเพิ่มความสะดวกในการติดต่อกับอุปกรณ์ภายนอก ที่มีระบบการติดต่อแบบ SPI เหมือนกัน เช่น หน่วยความจำภายนอก EEPROM, ไอซีวัดอุณหภูมิ, อุปกรณ์แปลงสัญญาณแอนะล็อกเป็นดิจิทัล, อุปกรณ์แปลงสัญญาณดิจิทัลเป็นแอนะล็อก หรือ ไมโครคอนโทรลเลอร์ตัวอื่น สำหรับในโปรเจกต์นี้จะใช้สำหรับติดต่อกับ ไอซีขับหน้าจอแสดงผล LCD

โมดูล SPI มีขาสำหรับต่อใช้งาน 4 ขา ซึ่งใช้ร่วมกับขาพอร์ตปกติด้วย ประกอบด้วย

1. SDIx Pin (Serial Data Input) เป็นขารับข้อมูลอนุกรม
2. SDOx Pin (Serial Data Output) เป็นขาส่งข้อมูลอนุกรม
3. SCKx Pin (Shift Clock input or output) เป็นขาสัญญาณนาฬิกาอนุกรม
4. SSx Pin (active low slave select or frame synchronization I/O pulse) เป็นขาเลือกอุปกรณ์สเลฟ ทำงานที่ลอจิก “0” หรือใช้เป็นขารับส่งสัญญาณพัลส์สำหรับกำหนดจังหวะของเฟรมข้อมูลอนุกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แต่อย่างไรก็ตาม ปกติแล้วจะใช้เพียง 3 ขาหลักคือ SDI, SDO และ SCK ขึ้นอยู่กับข้อกำหนดโหมดการทำงาน ในรูปที่ 2.23 แสดงบล็อกไดอะแกรมการทำงานของ โมดูล SPI ใน dsPIC

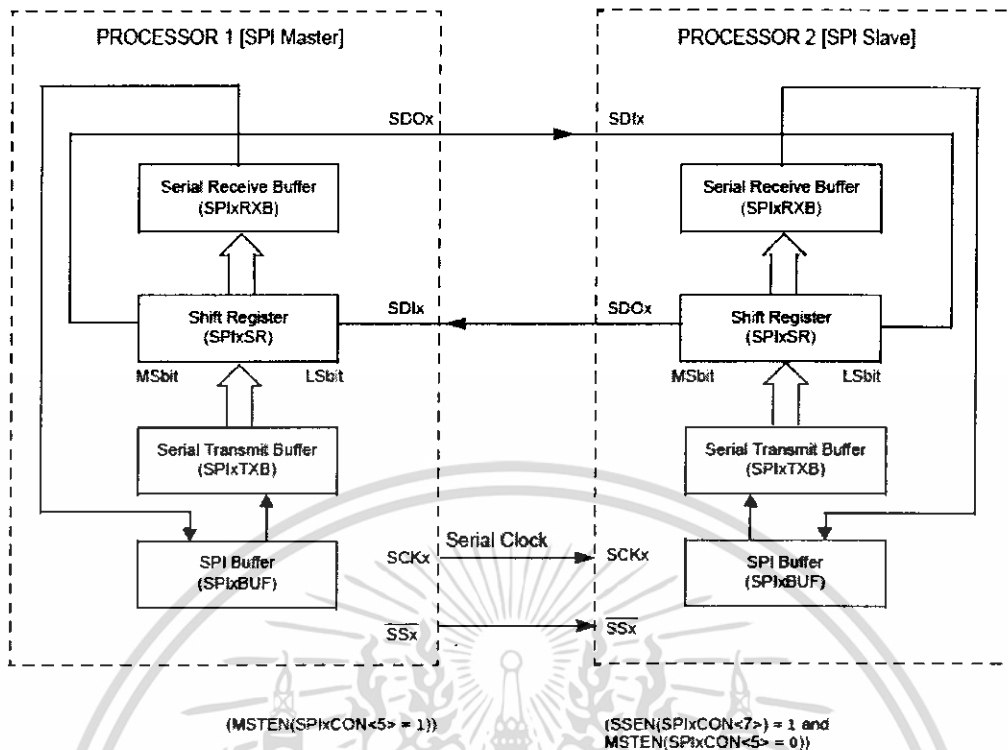


Note: The SPIxTXB and SPIxRXB registers are memory mapped to the SPIxBUF register.

รูปที่ 2.22 บล็อกไดอะแกรม ของ โมดูล SPI

สำหรับโครงการนี้ได้กำหนดให้โมดูล SPI ทำงานในโหมด มาสเตอร์ และมีรูปแบบการเชื่อมต่อระหว่างไมโครคอนโทรเลอร์กับไอซีเข้าและถอดรหัสเสียง ดังรูปที่ 2.24

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



- Note 1: Using the \overline{SSx} pin in Slave mode of operation is optional.
 Note 2: User must write transmit data to/read received data from SP1xBUF. The SP1xTXB and SP1xRXB registers are memory mapped to SP1xBUF.

รูปที่ 2.23 การเชื่อมต่อโมดูล SPI ระหว่างอุปกรณ์มาสเตอร์ และอุปกรณ์สเลฟ

2.3.3 โมดูลสำหรับเชื่อมต่ออุปกรณ์อนุกรม แบบ DCI

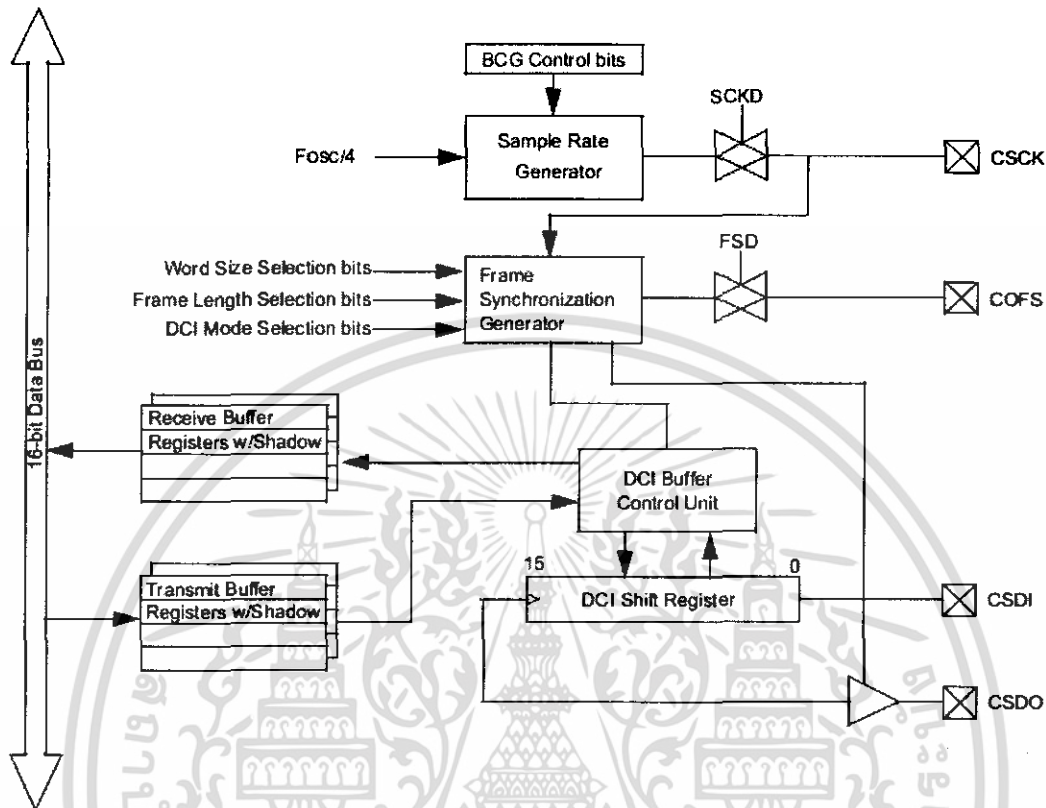
โมดูล DCI ช่วยให้การเชื่อมต่อกับอุปกรณ์ภายนอกทำได้ง่ายขึ้น อาทิเช่น อุปกรณ์เข้าและถอดรหัสเสียง (audio codecs) หรือ อุปกรณ์แปลงสัญญาณแอนะล็อกเป็นดิจิทัล, อุปกรณ์แปลงสัญญาณดิจิทัลเป็นแอนะล็อก สำหรับโปรเจกต์นี้จะใช้สำหรับติดต่อกับ ไอซี S13000

1. คุณสมบัติทั่วไปของ โมดูล DCI

- สามารถกำหนดขนาดเวิร์ดข้อมูล (data word size) ได้สูงสุดที่ขนาด 16 บิต ด้วยวิธีการโปรแกรม
- รองรับการทำงานได้สูงสุด 16 ไทม์สล็อต (time slot) ใน เฟรมข้อมูล (data frame) และรองรับขนาดเฟรม (frame size) ได้สูงสุด 256 บิต โดยมีบิตควบคุมอยู่ในทุกๆ ไทม์สล็อตในเฟรมข้อมูล ซึ่งจะคอยพิจารณาว่าเมื่อใด DCI จะรับหรือส่งข้อมูล
- มีแหล่งที่พักระหว่างข้อมูล (data buffer) สามารถเก็บข้อมูลได้ถึง 4 เวิร์ดข้อมูล โดยไม่แย่งทรัพยากรของ CPU

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในรูปที่ 2.25 จะแสดงถึงลักษณะการทำงานของส่วนประกอบต่างๆภายในโมดูล DCI ซึ่งแสดงถึงการทำงานรับส่งข้อมูลระหว่าง โมดูล DCI กับ บัสข้อมูล (data bus)



รูปที่ 2.24 บล็อกไดอะแกรม ของ โมดูล DCI

2. ขาสำหรับต่อใช้งาน โมดูล DCI

โมดูล DCI มีขาสำหรับต่อใช้งาน 4 ขา ประกอบด้วย

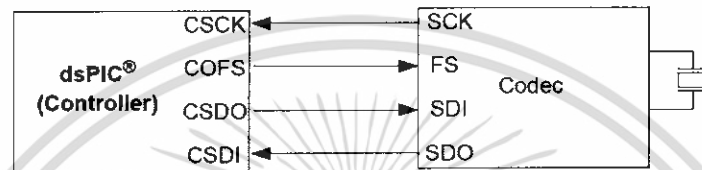
1. CSCK Pin (Serial Clock input or output) ทำหน้าที่ เชื่อมต่อสัญญาณนาฬิกาเข้าด้วยกัน สามารถกำหนดให้ไมโครคอนโทรลเลอร์เป็น ตัวรับสัญญาณนาฬิกาจากอุปกรณ์ภายนอกหรือเป็นตัวส่งสัญญาณนาฬิกา ผ่านทางบิตควบคุม
2. CSDO Pin (Serial Data Output) ถูกกำหนดให้เป็นขาส่งข้อมูลออก เมื่อโมดูล DCI ถูกเอนเอเบิล (enable) สามารถกำหนดว่าจะส่งข้อมูลออกหรือไม่ ผ่านทางบิตควบคุม
3. CSDI Pin (Serial Data Input) ถูกกำหนดให้เป็นขารับข้อมูลเข้า เมื่อโมดูล DCI ถูกเอนเอเบิล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. COFS Pin (Frame Synchronization) ทำหน้าที่ ซิงโครไนซ์ (synchronize) การถ่ายโอนข้อมูล ซึ่งเกิดขึ้นบน CSDO และ CSDI สามารถกำหนดได้ว่าจะรับหรือส่งพัลส์ (frame synchronization pulse) ผ่านทางบิตควบคุม

ในรูปที่ 2.26 นี้ ไมโครคอนโทรลเลอร์เป็นตัวส่งพัลส์ไปยังไอซีเข้าและถอดรหัสเสียงจึงจัดว่า โมดูล DCI ทำงานในโหมดมาสเตอร์

- ในกรณีที่ไมโครคอนโทรลเลอร์สร้างพัลส์ ส่งให้อุปกรณ์ภายนอกโมดูล DCI จะเป็น มาสเตอร์
- ในกรณีที่ไมโครคอนโทรลเลอร์รับพัลส์ จากอุปกรณ์ภายนอก โมดูล DCI จะเป็น สเลฟ



Codec Generates SCK

ข้อสังเกต : วงจรออสซิลเลเตอร์ของอุปกรณ์เข้าและถอดรหัส ทำหน้าที่ สร้างสัญญาณ SCK

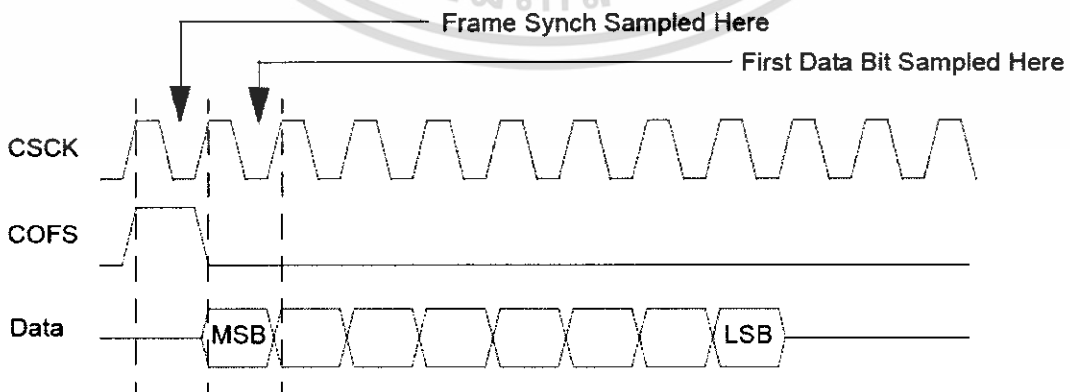
รูปที่ 2.25 แสดงตัวอย่างการเชื่อมต่อกับอุปกรณ์เข้าและถอดรหัส

3. ลักษณะการเชื่อมต่อ โมดูล DCI

โมดูล DCI มีลักษณะการเชื่อมต่ออยู่ 3 รูปแบบใหญ่ๆดังนี้

1. แบบสื่อสารอนุกรม (framed synchronous serial Transfer (single or multi-Channel))
2. แบบระบบบัส I²S (Inter-IC Sound interface)
3. แบบ AC-Link compliant mode

สำหรับในโครงการนี้ กำหนดให้ การสื่อสารอนุกรม DCI ทำงานแบบ สเลฟเฟรมซิงค์ ชนิดหลายช่องทาง (Slave Frame Sync & Multi-Channel mode) โดยรูปที่ 2.27 แสดงถึงจังหวะการส่งสัญญาณนาฬิกาเพื่อทำงานซิงโครไนซ์กับเฟรมข้อมูล



รูปที่ 2.26 Timing Diagram ของการทำงานแบบ สเลฟเฟรมซิงค์ ชนิดหลายช่องทาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับในบทที่ 2 นี้ ได้กล่าวถึง ทฤษฎีที่เกี่ยวข้องกับการประมวลผลสัญญาณดิจิทัล และความรู้ทางด้านไมโครคอนโทรลเลอร์ที่เกี่ยวข้องในการเขียนโปรแกรม รวมถึงโมดูลพิเศษต่างๆ เพื่อไปใช้ในการทำโครงการนี้ ในบทที่ 3 นั้นจะเป็นการอธิบายถึงการนำความรู้และทฤษฎีในบทที่ 2 มาใช้ในการออกแบบโครงการนี้

2.4 ทฤษฎีและความรู้ที่เกี่ยวข้องการบีบอัดข้อมูลเสียงมาตรฐาน SPEEX

Speex เริ่มมีขึ้นในปี 2002 เป็นการบีบอัดเสียงที่เป็นโอเพ่นซอร์ส Speex เป็นอัลกอริธึมแบบ Code Excited Linear Prediction (CELP) ไม่เหมือนกับการบีบอัดของ Vorbis ที่เคยมี เพราะ Speex ดีสำหรับเสียงในการสื่อสารที่มีดีเลย์ต่ำบนแพ็คเกจที่เชื่อถือไม่ได้ ปรินซิปลินี้จะนำเสนอข้อมูลที่น่าสนใจของ Speex เทคโนโลยีที่เกี่ยวข้องกันและความสามารถในการนำไปใช้งานบนแพลตฟอร์มต่างๆ การพัฒนาล่าสุดของ Speex เช่น ฟิกซ์พอยท์ พอร์ท, การกำจัดเสียงก้องบนเสียงเพลง และการกำจัดสัญญาณรบกวน

2.4.1 ความรู้เบื้องต้นเกี่ยวกับ SPEEX

Speex codec ดีสำหรับเสียงพูดและถูกออกแบบมาให้เหมาะสำหรับการสื่อสารภายในที่มีดีเลย์บนแพ็คเกจที่มีปัญหา

Speex เป็นอัลกอริธึม CELP ที่มีชื่อเสียง แต่เพราะการจำกัดสิทธิบัตร เทคนิคบางอย่างที่เหมือนกับการใช้โค้ดทางพีชคณิต (ACELP) นั้นใช้งานไม่ได้ อย่างไรก็ตาม Speex สามารถที่จะเปรียบเทียบคุณภาพกับเจ้าของ codec ที่บิตเรทเดียวกันได้ ในขณะที่การซัพพอร์ตคุณลักษณะของมันนั้นจะไม่พบในมาตรฐานการบีบอัดเสียงอื่นๆ คุณลักษณะเหล่านี้ได้รวมถึงบิตเรทที่เปลี่ยนแปลงได้ซึ่งทำให้การใช้งานบิตดีขึ้น เมื่อมีการเข้ารหัส (encoding) และการเอนเบคขณะเข้ารหัส จะมีวิธีง่ายในการอินเทอร์เฟซไว้ด์แบนด์ (16 kHz) แชนแนลด้วยเลก้าซีนาร์โรแบนด์ (8 kHz) ของโทรศัพท์

Speex ค่อยๆเปลี่ยนแปลงจนกลายเป็นเครื่องมือสมบูรณ์แบบ รวมทั้งอัลกอริธึมในการควบคุมสัญญาณรบกวน, เสียงก้องของเพลง และการควบคุมการเกิดการกระตุ่นในช่วงหนึ่งนั้น Speex จะกลายเป็นพอร์ทของสถาปัตยกรรมที่ไม่มีฟลอยติ่งพอยท์ยูนิต(floating-point unit)และจะมีการยอมให้ Speex ใช้งานได้บนอุปกรณ์เอนเบคที่มีอยู่แล้วบนฟิกซ์พอยท์ CPU หรือ DSP

Speex เกิดขึ้นเพราะเป็นโอเพ่นซอร์ส และฟรีจากเจ้าของซอฟต์แวร์ทั้งหลาย ในเวลานั้นตัวเลือกสำหรับการบีบอัดเสียงมีแค่ G.711 codec (รู้จักกันในชื่อ μ -law และ A-law) เท่านั้น Vorbis codec ก็มีไว้สำหรับการบีบอัดอডিโอทั่วไปด้วย แต่ไม่เหมาะกับการใช้กับเสียงพูดมากนัก

2.4.2 Speex และ CELP

Speex ใช้พื้นฐานของ CELP ที่ทำงานแบบ Code Excited Linear Prediction ซึ่งเทคนิคของ CELP มีอยู่ 3 อย่างคือ

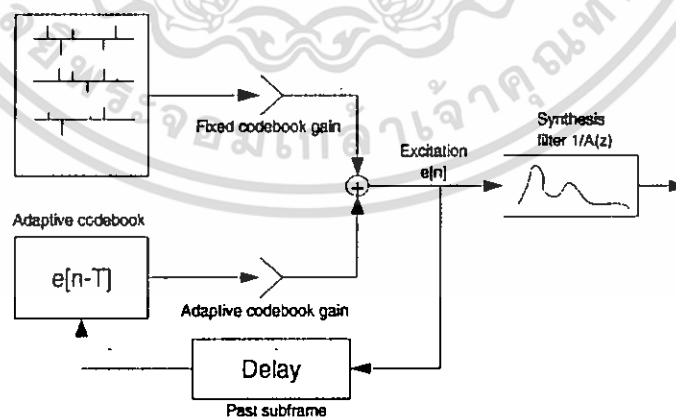
- 1) การใช้โมเดลของการทำนายเชิงเส้น (LP) กับ โมเดลของไวคอดแตรีก
- 2) การใช้อะแดปทีฟและพิกซ์โค้ดบุ๊กที่ปรับอินพุท (การกระตุ้น) ของ LP โมเดล
- 3) การค้นหาแบบวงปิดใน “Perceptually weighted domain”

1. Source-Filter Speech Model

ก่อนจะเข้าเนื้อหาของ Speex และ CELP เราขอแนะนำ Source-Filter Model ของโมเดลของเสียงพูดนี้ สมมุติว่าไวคอด ฮอร์คเป็นที่มาของเสียงแบบ spectrally flat (“สัญญาณกระตุ้น”) และ ไวคอดแตรีกทำงานเหมือนตัวกรอง ของรูปร่างที่เป็นสเปกตรัมของเสียงพูดมากมาย ขณะที่ยังประมาณการอยู่โมเดลของการบีบอัดเสียงใช้กันอย่างกว้างขวางเพราะความง่ายของมัน ประโยชน์ของมันยังเป็นเหตุผลว่าทำไมการบีบอัดเสียงส่วนมาก (รวมทั้ง Speex ด้วย) ทำงานได้เช่นบนสัญญาณเพลง

เสียงที่แตกต่างสามารถเห็นได้โดยการกระตุ้นของมัน (แหล่งกำเนิด) และ spectrally shape (ตัวกรอง) การสั่นเป็นจังหวะ(เช่น สระ) มีสัญญาณกระตุ้นเป็นระยะๆ และสามารถประมาณได้โดยขบวนการอิมพัลส์บน time domain หรือโดยการที่ความเร็วปกตินั้นสอดคล้องกับ frequency domain อีกอย่างคือความขาก (เช่น เสียง “s”, “sh”, “f”) เป็นสัญญาณกระตุ้นที่เรียกว่า “voice fricative” (เช่น “z” และ “v”) มีสัญญาณกระตุ้นที่สอดคล้องกันและส่วนของสัญญาณรบกวน

2. การถอดรหัส Speex



รูปที่ 2.27 การถอดรหัสแบบทั่วไปของ CELP

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก่อนจะสำรวจ กระบวนการเข้ารหัสที่ซับซ้อนของ Speex เราจะแนะนำตัวถอดรหัส Speex ก่อน เหตุผลคือความง่ายของตัวถอดรหัสสาร์โรแบนด์ รูปที่ 2.27 บรรยายเรื่องทั่วไปของตัวถอดรหัส CELP การกระตุ้นเกิดจากการรวมส่วนของอะแด็ปทีฟโค้ดบุ๊ก (aka pitch) และ พิกซ์โค้ดบุ๊ก (aka innovation)

$$e[n] = e_a[n] + e_f[n] \quad (2.27)$$

ซึ่ง $e_a[n]$ เป็นอะแด็ปทีฟโค้ดบุ๊ก คอนทริบิวชัน และ $e_f[n]$ เป็นพิกซ์โค้ดบุ๊กคอนทริบิวชัน

ตัวกรองที่ทำให้รูปร่างของการกระตุ้นมีโมเดลของโพลทั้งหมด (อิมพัลส์เรสพอนด์ที่เป็นอนันต์) อยู่ในรูป $1/A(z)$ ซึ่ง $A(z)$ เรียกว่า prediction filter ซึ่งเป็นตัวบอกการใช้การทำนายเชิงเส้น (Lenson-Durbin algorithm) ตัวกรองของโพลทั้งหมดถูกใช้เพราะว่ามันเป็นการแทนโวลคอลเทร็กที่ดี และเพราะว่ามันง่ายในการคำนวณ

ใน Speex เฟรมจะมีระยะ 20 ms ซึ่งมีนาร์โรวแบนด์ 160 แคมเบิ้ล ทุกๆเฟรมจะถูกแบ่งเป็น 4 เฟรมย่อย เฟรมละ 40 แคมเบิ้ล ซึ่งจะเข้ารหัสได้ถี่ขึ้น ในโหมคส่วนมากมีแค่ตัวกรองสังเคราะห์ และ global excitation gain (ดูข้างใต้) ที่เข้ารหัสบนพื้นฐานเฟรมย่อย (sub-frame)

Speex รวม 3 ข้อแตกต่างที่สำคัญ เมื่อเปรียบเทียบกับ CELP codec ล่าสุดอย่างแรก ขณะที่ CELP codec ตัวใหม่ล่าสุด ทำประโยชน์ด้านการประมาณค่าเฟรคชันแนลของพิทซ์บนแกนเดียว Speex ใช้เลขจำนวนเต็มเพื่อเข้ารหัสคาบเสียง แต่ใช้ตัวทำนาย 3 แท็บ (3 เกน) อะแด็ปทีฟโค้ดบุ๊กคอนทริบิวชัน สามารถแสดงได้ดังนี้

$$e_a[n] = g_0 e[n-T-1] + g_1 e[n-T] + g_2 e[n-T+1] \quad (2.28)$$

ซึ่ง g_0 , g_1 และ g_2 เป็นพิทซ์เกนและเป็นจอยท์ลี่ ควอนไทซ์ (VQ)

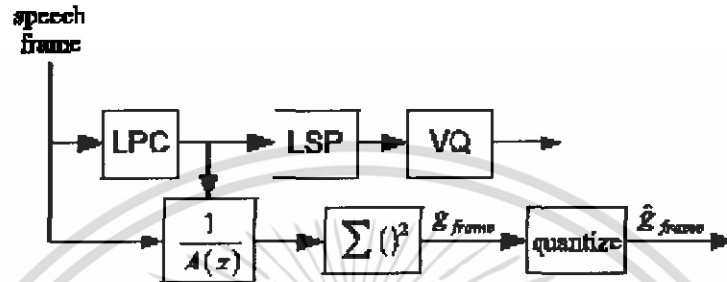
CELP codec มากมายในปัจจุบัน ใช้การทำนาย moving average (MA) เพื่อเข้ารหัสพิกซ์โค้ดบุ๊กที่กำหนดไว้ ซึ่งจะทำได้การโค้ดคิงที่ดีขึ้นนิดหน่อยในการนำสิ่งพิเศษที่ขึ้นอยู่กับเฟรมเก่า Speex ทำการเข้ารหัสพิกซ์โค้ดบุ๊กเกน เป็น global excitation gain บนเฟรมที่ดีเท่าเซตของเกนของเฟรมย่อย อย่างถูกต้อง

Speex ใช้การควอนไทซ์เฟรมย่อยในสัญญาณอิน โนเวชั่น (พิกซ์โค้ดบุ๊ก) นี้ไม่ใช้การควอนไทซ์ที่มีประสิทธิภาพเท่าไร แต่เป็นสิ่งที่ดีที่สุดที่สามารถใช้สัทริบิตรที่ถูกจำกัดได้ ทุกๆ เฟรมย่อยจะถูกแบ่งเป็นเวกเตอร์ย่อย ในย่าน 5-20 แคมเบิ้ล ทุกเวกเตอร์ย่อยถูกเลือกจากบิตเรทที่ขึ้นกับโค้ดบุ๊กและ เวกเตอร์ย่อย ทั้งหมดถูกเชื่อมโยงกัน ดังตัวอย่างโหมค 3.95 kbps ใช้เวกเตอร์ย่อยขนาด

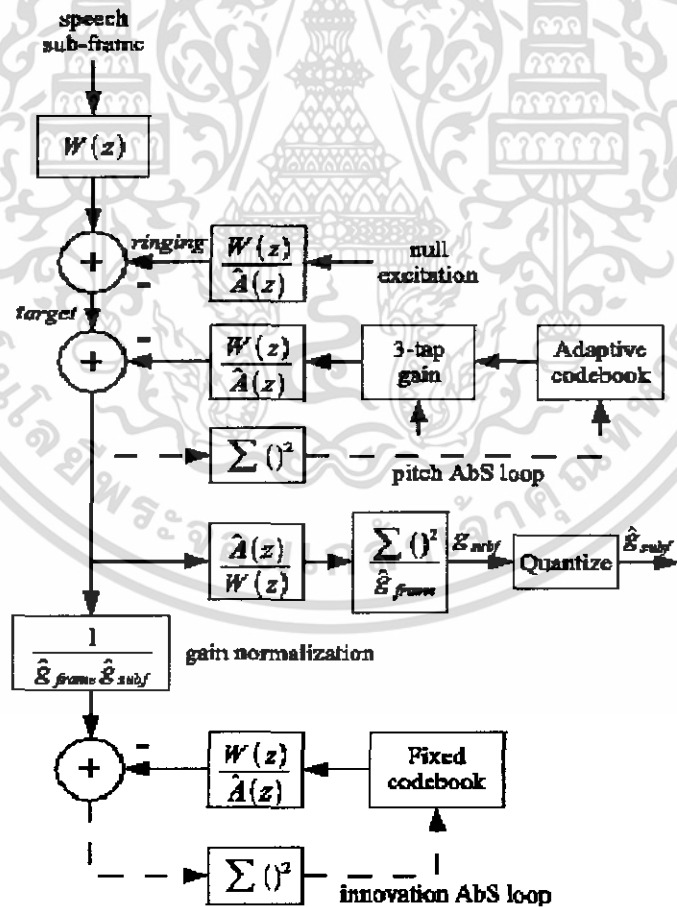
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

20 แชมเปิ้ล 32 เอนทรี ในโค้ดบุ๊ก (5 บิต) หมายความว่า การเปลี่ยนแปลงคือการเข้ารหัส 10 บิต ต่อหนึ่งเฟรมย่อยหรือ 2000 bps อีกตัวอย่างหนึ่ง โหมด 18.2 kbps ใช้เวกเตอร์ย่อยขนาด 5 แชมเปิ้ล 256 เอนทรีในโค้ดบุ๊ก (8 บิต) ดังนั้น การเปลี่ยนแปลงจะใช้ 64 บิตต่อเฟรมย่อย หรือ 12800 bps

3. การเข้ารหัส Speex



รูปที่ 2.28 การวิเคราะห์เฟรมในระบบวงเปิด



รูปที่ 2.29 การวิเคราะห์เชิงสังเคราะห์แบบวงปิดของพหุนามเชิงซ้อนของเฟรมย่อย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลักสำคัญเบื้องหลัง CELP เรียกว่า การวิเคราะห์โดยการสังเคราะห์ (AbS) หมายความว่า การเข้ารหัส (วิเคราะห์) ทำงานโดยการอพยพโพลาไรซ์เพื่อเซพซอล สัญญาณที่ถูกถอดรหัส ในวงปิดตาม ทฤษฎี CELP สตรีมที่ดีที่สุด เกิดจากการพยายามรวมบิตทั้งหมดที่เป็นไปได้ แล้วเลือกมาแค่หนึ่ง ที่ จะให้การถอดรหัสสัญญาณที่เสียงที่ดีที่สุด ซึ่งไม่สามารถเป็นไปได้ เพราะเหตุผล 2 ข้อ คือ: ความต้องการที่สับสนบนฮาร์ดแวร์ ที่มีอยู่ทั้งหลาย และการเลือกเสียงที่ดีที่สุดเพื่อเป็นตัวเลือกการ ฟังของมนุษย์

เพื่อการเข้ารหัสแบบเรียลไทม์ โดยคำนวณแหล่งกำเนิดที่มีการจำกัดขอบเขตการมอง CELP ถูกยุบให้เล็กลง จึงจัดการได้ง่ายขึ้น ทำให้การหาเพอร์เซพซอลเวทคิงฟังก์ชันทำได้ง่าย ใน กรณียของ Speex ออพติไมซ์เซชันจะมี 4 ขั้นตอน คือ

- 1) การวิเคราะห์การทำนายเชิงเส้นใช้ตัดสินตัวกรองสังเคราะห์ที่เปลี่ยนเป็นสัมประสิทธิ์ Line Spectral Pair (LSP) และ เวกเตอร์ควอนไทซ์
- 2) อะแด็ปทีฟโค้ดบุ๊กเอนทรี และ เกน ร่วมกันค้นหาการรวมกันของพิตช์เอนทรีที่ดีที่สุด โดยใช้ AbS
- 3) พิกส์โค้ดบุ๊กเอนทรีเป็นลักษณะ “วงเปิด” แต่ในสัญญาณกระตุ้นเท่านั้น
- 4) พิกส์โค้ดบุ๊กมีไว้สำหรับเอนทรีที่ดีที่สุดในการใช้ AbSD

ออพติไมซ์เซชันของขั้นตอนที่ 2) และ 4) ทำงานโดยใช้ชื่อ “เพอร์เซพซอลเวทโดเมน” หมายความว่า เราพยายามลดความแตกต่างที่สังเกตได้ของต้นฉบับ เพื่อที่จะทำอย่างนั้น ได้เวทคิงฟิลเตอร์จึงถูก ประยุกต์บนสัญญาณอินพุตดังนี้ :

$$W(z) = \frac{A(z/\gamma_1)}{A(z/\gamma_2)} \quad (2.29)$$

$A(z)$ เป็นตัวกรองสังเคราะห์และ γ_1 และ γ_2 ควบคุมรูปร่างของตัวกรองโดยย้ายโพลและซีโ โร ไปที่ตรงกลางวงกลม 1 หน่วย ของ z-transform (Speex ใช้ $\gamma_1 = 0.9$ และ $\gamma_2 = 0.6$) ฟิลเตอร์เป็น การประมาณค่าอย่างหยาบมาก ๆ ในสิ่งที่รู้จักกันดีในชื่อ “masking curve” ในการบีบอัดเสียง เช่น Vorbis ผลกระทบทั้งหมดของเวทคิงฟิลเตอร์ $W(z)$ คือตัวเข้ารหัสที่จะยอมให้สัญญาณรบกวนเข้า มามากขึ้นที่ความถี่ที่มีระดับกำลังสูง และมีสัญญาณรบกวนน้อยลงที่ความถี่ที่มีระดับกำลังต่ำ

4. ไวด์แบนด์ CELP (SB-CELP)

โหมดไวด์แบนด์ใน Speex ต่างจากไวด์แบนด์ในการบีบอัดเสียงอื่นๆ มากที่สุด เพราะมัน แยกสัญญาณเป็น 2 เฟรมย่อย (SB-CELP) แบนด์ที่ต่ำกว่าถูกเข้ารหัสโดยใช้การเข้ารหัสสนาร์โรว์ แบนด์ซึ่งจะได้เปรียบเพราะทำให้มันง่ายในการอินเตอร์โอบอเรท กับระบบ “เลกาซี” (เช่น PSTN) ที่ทำงานเป็นนาร์โรว์แบนด์ ส่วนแบนด์ที่สูงกว่าจะถูกเข้ารหัสโดยใช้ตัวเข้ารหัส CELP ที่คล้ายกับ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวเข้ารหัสสาร์โร้แบนด์ที่มีความแตกต่างกันตรงที่ใช้แค่ฟิกส์โค้ดบุ๊ก ไม่ใช่อะแด็ปทีฟโค้ดบุ๊ก ข้อมูลของแบนด์สูงถูกเข้ารหัส หลังจากนาร์โร้แบนด์ (แบนด์ต่ำ) ข้อมูลส่วนสุดท้ายไว้ค้แบนด์เฟรมที่ให้นาร์โร้แบนด์เฟรมที่สมบูรณ์ถูกทิ้งไปนั่นเอง ตัวถอดรหัสไว้ค้แบนด์สามารถที่จะพิสูจน์ว่าข้อมูลแบนด์สูงหายไปและสามารถถอดรหัสนาร์โร้แบนด์เฟรมได้

2.4.3 การนำ Speex ไปประยุกต์ใช้

Speex ต่างจากสปีชโคเดคอื่นๆมาก เพราะมันซัพพอร์ตในหลายๆแซมปลิงเรท และหลายๆ บิตเรท ในทุกๆ แซมปลิงเรทสิ่งหนึ่งที่ต้องตัดสินใจก่อนเลือกโคเดค (หรือในกรณีนี้โหมคไหนเป็นของโคเดค) คือ:

- แบนด์วิธที่มีเป็นเท่าไร
- คุณภาพที่ต้องการเป็นเท่าไร
- คีเลย์ที่ต้องการเป็นเท่าไร

เมื่อออกแบบ ลักษณะอย่างแรกก็ต้องพิจารณาความจริงว่า การรวม IP, UDP และ RTP ก่อให้เกิดการส่งโอเวอเฮด 40 ไบต์ ต่อแพ็กเก็ต เมื่อการส่ง 1 แพ็กเก็ตในทุกๆ 20 ms ถูกทำไปแล้วเป็นปกติ โอเวอเฮดจะเท่ากับ 16 kbps! พิจารณาว่าเห็นได้ชัด ในเหตุการณ์ที่เกิดมากที่สุด คือ มีการเร่งนิตหน้อยในการใช้บิตเรทที่ต่ำมากๆ

1. การเลือกโหมคที่ถูกต้อง

ตัวเลือกแรกที่ต้องทำเมื่อใช้ Speex คือ การใช้แซมปลิงเรทสำหรับแอปพลิเคชันใหม่นั้น ต้องการใช้ไว้ค้แบนด์คือโอไอ (16 kHz แซมปลิงเรท) เท่าที่เป็นไปได้ ไว้ค้แบนด์ของเสียงพูดชัดเจนกว่า เสียงของนาร์โร้แบนด์และเข้าใจง่ายกว่า ในหลายๆกรณี มันดีที่จะซัพพอร์ตนาร์โร้แบนด์ด้วย ทั้งทางตรงหรือผ่าน ตัวเข้ารหัส/ตัวถอดรหัส ของไว้ค้แบนด์ ในบริบทนี้ นาร์โร้แบนด์มีประโยชน์กับเหตุผลที่สอดคล้องกันทั้งคู่ และเพราะที่บิตเรทต่ำๆ (<10 mbps) บางทีคุณภาพก็ดีกว่าไว้ค้แบนด์

เมื่อแซมปลิงเรทถูกเลือกแล้ว ส่วนหนึ่งคือการตัดสินใจในการเลือกบิตเรทตามความต้องการท้วๆไป คือใช้บิตเรทคงที่ (CBR) เพราะความจุของฟิกส์เซนแนลลือกอย่างหนึ่ง VBR มีประสิทธิภาพดีกว่าในด้านการบีบอัดไฟล์ เพราะ บิตสามารถแยกไว้ในที่ที่ต้องการได้ เมื่อคำนวณเป้าหมายของบิตเรทเป็นสิ่งสำคัญที่จะพิจารณาโอเวอเฮด เมื่อใช้ RTP และ 1 เฟรมต่อแพ็กเก็ต คือ 16 kbps สำหรับการเข้ารหัสไฟล์ใน Ogg โอเวอเฮดจะประมาณ 1 ไบต์ ต่อ เฟรม ดังนั้นจึงไม่เป็นปัญหาเลย สำหรับทุกๆบิตเรทต่ำๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ปัญหาทั่วไปของ layer 8

มีปัญหาที่เกิดขึ้นบ่อยเมื่อมีการซัพพอร์ตใน Speex มีบางอย่างที่ทำให้โปรแกรม เกิดความผิดพลาด ส่วนใหญ่เป็นการส่งหรือการรับที่ผิด ลักษณะที่สำคัญที่สุดของสัญญาณอินพุต คือ แอมพลิจูด Speex สามารถเข้ารหัสสัญญาณมากมายในย่าน $[-35768, 32767]$ อย่างไรก็ตาม มันทึที่สุดที่จะทำให้เกิดประโยชน์จากย่านฟูลไดนามิกโดยไม่ใกล้กับจุดที่เกิดการแซทเทอเรทมากเกินไป เพราะอาจเป็นสาเหตุให้เสียงพูดที่ถูกถอดรหัสนั้นเกิดการแซทเทอเรท แม้ว่าเสียงพูดที่ถูกเข้ารหัสจะไม่ใช่เป็นอย่างนั้นก็ตาม ซึ่งมันจะดีที่สุดในที่ระดับสัญญาณอินพุต ดังนั้น ค่าสูงสุด ของแอมพลิจูด จะอยู่ระหว่าง 5000 ถึง 20000

ลักษณะที่สำคัญอื่นๆของสัญญาณคือความถี่ที่มี DC offset ควรจะหลีกเลี่ยงค่าทั้งหมด เพราะมันเป็นสาเหตุให้ตัวเข้ารหัสของ Speex ผลิตไฟล์เสียงที่แย่ (DC ไม่จำเป็นสำหรับออดิโอ) DC offset บางทีอาจเป็นสาเหตุจากซาวด์การ์ดที่มีราคาถูก หมายความว่าสัญญาณจะไม่อยู่ตรงกลางรอบๆซีโร่ (zero-mean) กรณีที่ง่ายมากๆ DC offset สามารถย้ายอย่างง่ายดายโดยใช้ notch filter ในรูป:

$$n(z) = \frac{1 - z^{-1}}{1 - \alpha z^{-1}} \quad (2.30)$$

ซึ่งค่าของ $\alpha = 0.98$ ซึ่งเหมาะทั้งตัวกรองนาร์โรแบนด์และ ไวด์แบนด์ ขณะที่เห็นผลกระทบได้น้อยลง บางทีความถี่ต่ำๆ จำนวนมาก อาจเป็นสาเหตุของการเข้ารหัสเฟรมย่อยของ Speex ได้ ตามกฎทั่วไป นาร์โรแบนด์ออดิโอควรจะเป็น high-pass filter ที่ 300 Hz และ ไวด์แบนด์ฟิลเตอร์ควรเป็น ฟิลเตอร์ที่ 50 Hz มันเป็นหนึ่งในลักษณะแผนงานของ Speex เพื่อทำให้เกิดตัวกรองโดยอัตโนมัติ แต่มันยังไม่ได้ถูกสร้างขึ้นเลย

บทที่ 3

หลักการออกแบบ

การออกแบบการบีบอัดสัญญาณเสียง ในเบื้องต้นนั้นต้องเข้าใจหลักการทำงานในภาพรวมของระบบก่อนว่ามีหลักการทำงานอย่างไร แล้วจึงพิจารณาส่วนย่อยต่างๆ อันประกอบไปด้วยหน่วยประมวลผลกลาง, หน่วยเข้าและถอดรหัสเสียง และหน่วยแสดงผล โดยหน่วยย่อยทั้งสามนี้จะทำงานร่วมกัน ภายใต้ระบบการทำงานที่ถูกออกแบบไว้ ท้ายที่สุดจึงเป็นการนำเสนอโครงสร้างการทำงานในส่วนของโปรแกรม ซึ่งอธิบายโครงสร้างได้ตามรูปที่ 3.1 อันประกอบด้วยส่วนประกอบต่างๆ ดังนี้

1. หน่วยเข้าและถอดรหัสเสียง

ทำหน้าที่สองอย่างคือ เมื่อรับสัญญาณเสียงแอนะล็อกจากไมโครโฟนไอซี Si3000 จะแปลงสัญญาณแอนะล็อกเป็นสัญญาณดิจิทัล และนำสัญญาณดิจิทัลนั้นส่งไปประมวลผลในรูปแบบเฟรมข้อมูล ด้วยการเชื่อมต่อแบบอนุกรม (DCI) ในทางกลับกันเมื่อไมโครคอนโทรลเลอร์บีบอัดสัญญาณเสร็จแล้ว ไอซีSi3000จะแปลงสัญญาณดิจิทัลเป็นสัญญาณแอนะล็อกและส่งเสียงแอนะล็อกนั้นทางลำโพง หน่วยเข้าและถอดรหัสเสียงสามารถแสดงได้ดังรูปที่ 3.1

2. หน่วยแสดงผล

ในโครงการนี้ผู้จัดทำได้ใช้จอแสดงผลในการแสดงค่าฟังก์ชันของการบีบอัดสัญญาณเสียงในแต่ละโหมคการทำงานของโปรแกรม ดังรูปที่3.1 และเนื่องจากผู้จัดทำใช้บอร์ดทดลอง dsPICDEM 1.1 ในการพัฒนาโปรแกรม ซึ่งในบอร์ดนี้มีส่วนแสดงผลอยู่แล้วอันประกอบด้วยจอแสดงผลแบบผลึกเหลว (Liquid Crystal Display : LCD) และหลอดไฟ ในส่วนของจอแสดงผลเป็นแบบคอตเมตริกซ์ (dot-matrix)122x32 ขนาด 4 บรรทัด 20 ตัวอักษร ใช้ไอซี PIC18F242 LCD เป็นคอนโทรลเลอร์ ติดต่อกับ LCD ในโหมค 8 บิต และ dsPIC30F6014 ซึ่งเป็นไมโครคอนโทรลเลอร์ตัวหลักจะทำการเชื่อมต่อกับไอซี PIC18F242 LCD ผ่านการสื่อสารแบบอนุกรม SPI โดยผู้จัดทำกำหนดให้ dsPIC30F6014 เป็นมาสเตอร์ และไอซี PIC18F242 LCD เป็นสลาฟ

3. หน่วยประมวลผลกลาง

เป็นหน่วยที่มีความสำคัญมากที่สุด ในโครงการนี้ใช้ไมโครคอนโทรลเลอร์ dsPIC30F6014 ขนาด 32 บิตเป็นตัวประมวลผล ภายในได้สร้างระบบการบีบอัดสัญญาณเสียงโดยใช้ภาษาซี ร่วมกับภาษาแอสเซมบลีในการพัฒนาโปรแกรม และใช้ MPLAB C30 เป็นตัวคอมไพเลอร์ ดังรูปที่ 3.1(C) ซึ่งบล็อกไดอะแกรมการทำงานของระบบบีบอัดสัญญาณจะได้กล่าวต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

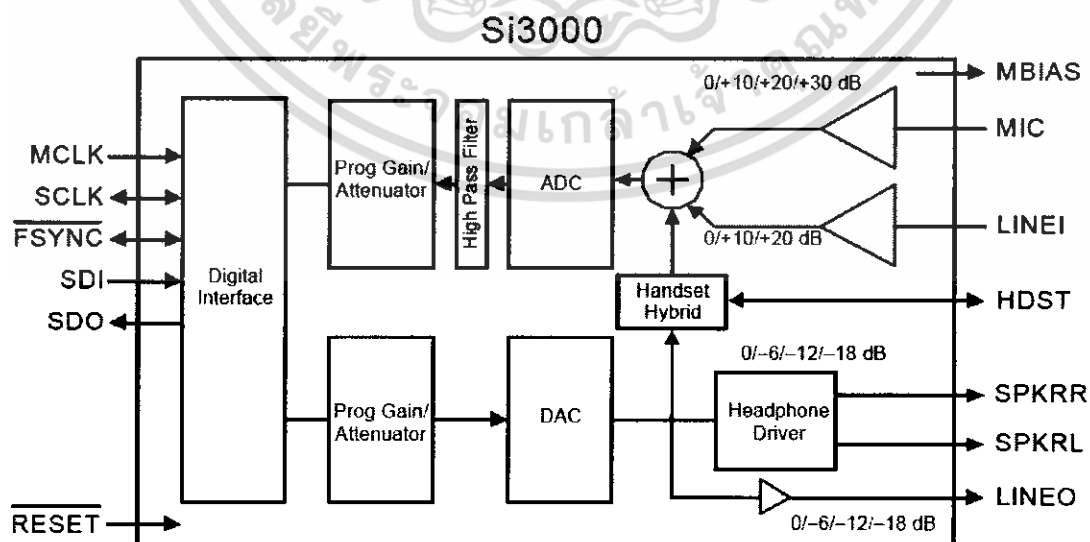
3.1 หน่วยเข้าและถอดรหัสเสียง

ในโครงการนี้จะใช้หน่วยเข้าและถอดรหัสเสียง Si3000 ซึ่งเป็นไอซีสำเร็จรูปที่มีฟังก์ชันการทำงานกับเสียงหลากหลาย และที่สำคัญคือไอซี Si3000 สามารถรองรับการประมวลผลแบบดิจิทัลซึ่งสอดคล้องกับความต้องการของโครงการ สำหรับคุณสมบัติของไอซี Si3000 จะอธิบายได้ดังต่อไปนี้

3.1.1 คุณสมบัติของ Si3000

- แปลงสัญญาณแอนะล็อกเป็นสัญญาณดิจิทัลและแปลงสัญญาณดิจิทัลเป็นสัญญาณแอนะล็อกขนาด 10 บิต
- สามารถปรับเพิ่ม,ลดสัญญาณเสียงในช่วง -34.5 ถึง 12 dB โดยการปรับค่าทางโปรแกรม
- สามารถรับสัญญาณแอนะล็อกได้ 3 แชนแนล
- เชื่อมต่อแบบอนุกรมกับไมโครคอนโทรลเลอร์
- ขยายสัญญาณเสียงอินพุต ได้สูงสุด 30 dB
- รองรับความถี่สุ่มตัวอย่างได้ 4 ถึง 12 kHz

การทำงานของไอซี Si3000 เริ่มจากรับเสียงแอนะล็อกจากไมโครโฟน หรือจากสายออดิโอ (audio) โดยตรง จากนั้นสัญญาณจะถูกขยายในช่วง 0 ถึง 30 เดซิเบล สัญญาณแอนะล็อกนั้นจะถูกแปลงเป็นสัญญาณดิจิทัล แล้วผ่านตัวกรองความถี่สูง และโปรแกรมเกนตามลำดับ ข้อมูลดิจิทัลนั้นจะติดต่อสื่อสารกับหน่วยประมวลผลผ่านการเชื่อมต่อแบบอนุกรม DCI ในทางกลับกันข้อมูลดิจิทัลจะถูกส่งออกมาที่ไอซี Si3000 ผ่าน โปรแกรมเกนเข้าที่พุด และทำการแปลงสัญญาณดิจิทัลเป็นสัญญาณแอนะล็อกกลับ และส่งสัญญาณนั้นออกทางสายออดิโอ หรือลำโพงซึ่งในที่นี้ขึ้นอยู่กับ การกำหนดค่าทางโปรแกรม การทำงานของไอซี Si3000 อธิบายได้ดังรูปที่ 3.1



รูปที่ 3.1 บล็อกไดอะแกรมของไอซี Si3000

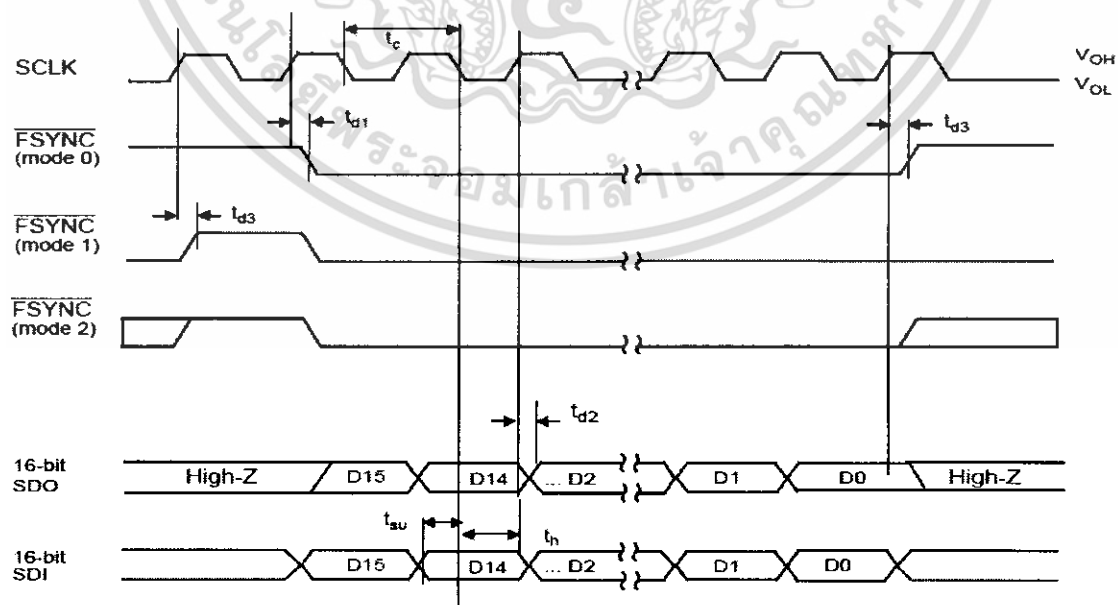
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.2 การเชื่อมต่อแบบดิจิทัล

การเชื่อมต่อแบบดิจิทัลแบ่งออกเป็น 3 โหมด ในแต่ละโหมดจะมีช่วงจังหวะการรับส่งข้อมูลในเฟรมที่ขึ้นกับจังหวะสัญญาณนาฬิกา SCLK ดังรูปที่ 3.2 โดยมีคุณลักษณะการสวิตซ์ชิงของการเชื่อมต่อแบบอนุกรมดังตารางที่ 3.1

ตารางที่ 3.1 คุณลักษณะการสวิตซ์ชิงของการเชื่อมต่อแบบอนุกรม

Parameter	Symbol	Min	Type	Max	Unit
Cycle Time, SCLK	t_c	354	1/256 Fs	-	ns
SCLK Duty Cycle	t_{dy}	-	50	-	%
Delay Time, SCLK \uparrow to FSYNC \downarrow	t_{d1}	-	-	10	ns
Delay Time, SCLK \uparrow to SDO Valid	t_{d2}	-	-	20	ns
Delay Time, SCLK \uparrow to FSYNC \uparrow	t_{d3}	-	-	10	ns
Setup Time, SDI, before SCLK \downarrow	t_{su}	25	-	-	ns
Hold Time, SDI, after SCLK \downarrow	t_h	20	-	-	ns
Setup Time, FSYNC (mode 2) before MCLK \downarrow	t_{su}	25	-	-	ns
Hold Time, FSYNC (mode 2) after MCLK \downarrow	t_h	20	-	-	ns



รูปที่ 3.2 ไทมิงไดอะแกรมของการเชื่อมต่อแบบอนุกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เลือกโหมดการรับส่งข้อมูล โดยใช้ตัวต้านทาน 50 k Ω พัลส์/ควาน์ ที่ขา SDO และ SDI ตามตารางที่ 3.2 สำหรับในโครงการนี้จะกำหนดให้รับส่งข้อมูลแบบอนุกรมในโหมดที่ 1 โดยให้ ไอซี Si3000 เป็นมาสเตอร์

ตารางที่ 3.2 โหมดการเชื่อมต่อแบบอนุกรม

โหมด	SCLK*	SDO*	รายละเอียด
0	0	0	\overline{FSYNC} เฟรมข้อมูล
1	0	1	\overline{FSYNC} เป็นสัญญาณขึ้นบันไดเริ่มเฟรมข้อมูล
2	1	0	สเตฟโหมด
3	1	0	ไม่ใช้งาน

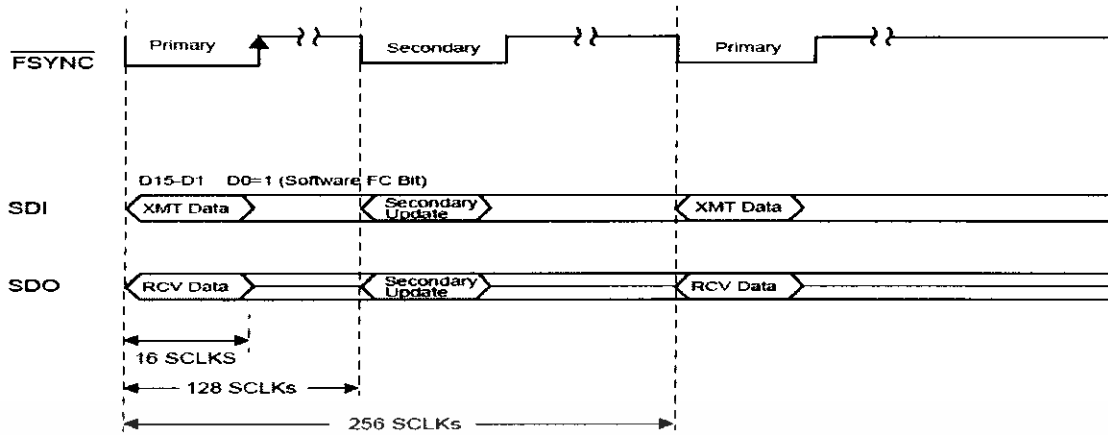
*เมื่อ : พัลส์/พลาควาน์

เมื่อกำหนดให้ไอซี Si3000 เป็นมาสเตอร์ (master clock - MCLK) จะเป็นอินพุท และ (serial data clock : SCLK) จะเป็นเอาต์พุท และ SCLK จะทำงานที่ 256 บิต ต่อเฟรม เมื่อเฟรมเรท (frame rate) มีค่าเท่ากับความเร็วสุ่มตัวอย่าง (F_s) โดยการส่งและรับข้อมูลจะแบ่งออกเป็น 2 เฟรม คือ

- เฟรมปฐมภูมิ จะใช้สำหรับเก็บจุดตัวอย่างข้อมูลสัญญาณเสียงดิจิตอล
- เฟรมทุติยภูมิ จะใช้ติดต่อกับรีจิสเตอร์ ภายในของไอซี Si3000 และจะเกิดขึ้นเฉพาะเวลาถูกร้องขอเท่านั้น

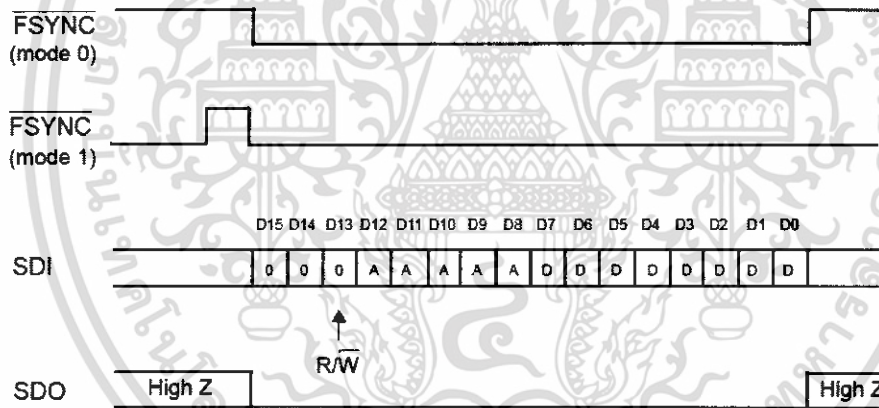
การส่งข้อมูลในเฟรมปฐมภูมิ (จาก dsPIC ไปยังไอซี Si3000) จะใช้ LSB เป็นบิตร้องขอเฟรมทุติยภูมิ ดังนั้นข้อมูลเสียงที่ส่งมาจะมีเพียง 15 บิตเท่านั้น และเมื่อเฟรมทุติยภูมิ ไม่ปรากฏก็จะไม่เกิดการส่งข้อมูลในไทม์สล็อตนี้

การรับข้อมูลในเฟรมปฐมภูมิ (จากไอซี Si3000 ไปยัง dsPIC) ในโหมด 0 หรือ 1 จะรับทางขา SDO พร้อมกับข้อมูลเสียง เมื่อ dsPIC ได้รับเฟรมปฐมภูมิ แล้วจะตรวจสอบ LSB ว่าจะรับข้อมูลจาก ไอซี Si3000 ต่อ หรือส่งเฟรมทุติยภูมิออกไป สำหรับการส่งและรับข้อมูลในเฟรมทุติยภูมิ จะส่งและรับข้อมูลแบบ 16 บิต ดังรูปที่ 3.3



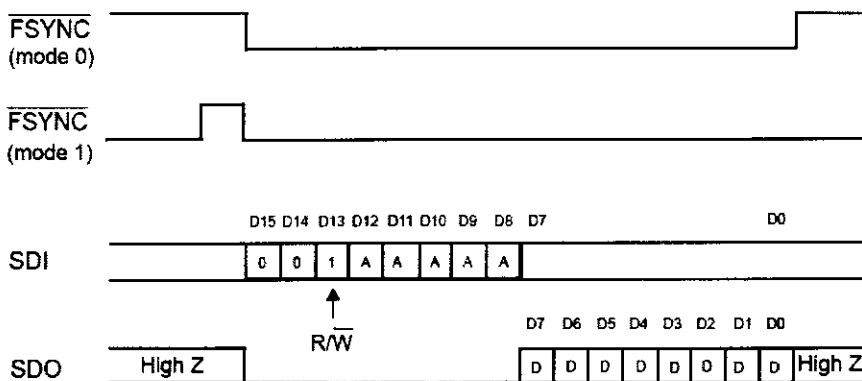
รูปที่ 3.3 การร้องขอเฟรมทุดิยภูมิ

ระหว่างการเขียนข้อมูลบิต R/W จะถูกเคลียร์ ตามด้วย แอคเตสของรีจิสเตอร์ที่เขียนขนาด 5 บิตแล้วตามด้วย ข้อมูลขนาด 8 บิต และจะมีรีจิสเตอร์เพียงตัวเดียวเท่านั้นที่จะถูกอ่านหรือเขียนในเฟรมทุดิยภูมิดังรูปที่ 3.4



รูปที่ 3.4 การเขียนข้อมูลลงเฟรมทุดิยภูมิ

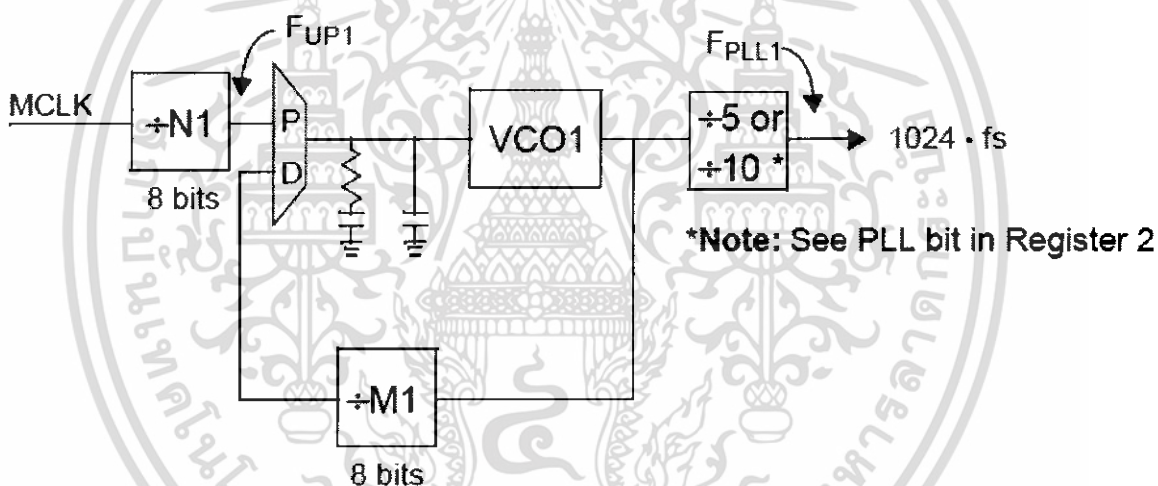
ระหว่างการอ่านข้อมูล บิต R/W จะถูกกำหนดให้เป็น 1 ตามด้วย แอคเตสของรีจิสเตอร์ที่อ่านขนาด 5 บิต ส่วนข้อมูลขนาด 8 บิต ที่เหลือจะใช้กับ รีจิสเตอร์ควบคุม (control register) ซึ่งข้อมูลจะถูกเขียนผ่านขา SDO ดังรูปที่ 3.5



รูปที่ 3.5 การอ่านข้อมูลจากเฟรมทศนิยม

3.1.3 การเลือกความถี่สัญญาณนาฬิกา

ไอซี Si3000 สามารถสร้างความถี่สุ่มตัวอย่าง โดยใช้ข้อสซิงเลเตอร์จากภายนอก (MCLK) เป็นตัวกำหนด



รูปที่ 3.6 ระบบการสร้างสัญญาณนาฬิกาด้วยเฟสล็อกคูล

สัญญาณนาฬิกาที่ใช้ในการสุ่มตัวอย่างจะถูกสร้างมาจากเฟสล็อกคูล ซึ่งจากรูปที่ 3.6 F_{PLL1} จะเป็นตัวกำหนดค่าความถี่สุ่มตัวอย่าง สถาปัตยกรรมระบบนี้เป็นแบบวงปิด ที่ใช้รีจิสเตอร์ 2 ตัว เป็นตัวกำหนดความถี่สุ่มตัวอย่าง ซึ่งใช้วิธีกำหนดค่าทางโปรแกรม นอกจากนั้นความถี่จาก MCLK ยังสามารถรองรับในช่วงความถี่ 1 MHz ถึง 60 MHz โดยตัวกำเนิดสัญญาณนาฬิกา (clock generator) จะสร้างความถี่สัญญาณนาฬิกา ที่ $1024 * F_s$ เมื่อ F_s คือความถี่สุ่มตัวอย่างที่ต้องการ

3.1.4 การกำหนดค่ารีจิสเตอร์สำหรับตัวกำเนิดสัญญาณพิก้า

ประกอบด้วยรีจิสเตอร์ 2 ตัวได้แก่

Register 3 - N1 divider, 8 bits

Register 4 - M1 divider, 8 bits

เมื่อ N1 และ M1 เป็นรีจิสเตอร์ ขนาด 8 บิต แบบไม่คิดเครื่องหมาย

เราสามารถคำนวณค่า N1 และ ค่า M1 ได้ตามสมการความสัมพันธ์ระหว่างความถี่จาก MCLK และความถี่สุ่มตัวอย่าง ดังสมการที่ (3.21)

$$\frac{M1}{N1} = \frac{5 * 1024 * Fs}{MCLK} \quad (3.21)$$

เมื่อ: Fs คือ ความถี่สุ่มตัวอย่าง

ตารางที่ 3.3 ตัวอย่างค่าความถี่จาก MCLK ในการสร้างความถี่สุ่มตัวอย่าง 8 kHz

ความถี่ออสซิลเลเตอร์ MCLK (MHz)	N1	M1
1.8432	9	200
4.000	25	256
4.0960	1	10
5.2800	33	256
5.7600	9	64
6.1440	3	20
8.1920	1	5
9.2160	9	40
10.0800	63	256
10.5600	33	128
11.0592	27	100
12.288	3	10
14.7456	9	25
16.000	25	64

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านธุรกิจ

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลำดับการเขียนรีจิสเตอร์ มีความสำคัญมากสำหรับความถี่ที่ได้จากระบบนี้ N1 จะต้องถูกเขียนก่อน M1 เสมอ

ข้อสังเกต ค่า N1 และ ค่า M1 ที่ถูกเขียนจริงจะต้องน้อยกว่าการคำนวณอยู่ 1 เสมอ

ขั้นตอนสุดท้ายของการออกแบบ คือการปรับแก้อัตราความถี่ (update rate) ตามสมการที่ (3.22) และ (3.23)

$$\text{PLL lock time} < 1 \text{ ms} \quad (3.22)$$

$$F_{up} = \frac{F_{MCLK}}{N1} \geq 144 \text{ kHz} \quad (3.23)$$

เมื่อ: F_{up} คือ ความถี่ที่เข้ามาเพื่อทำการปรับแก้ ดังแสดงในรูปที่ 3.6

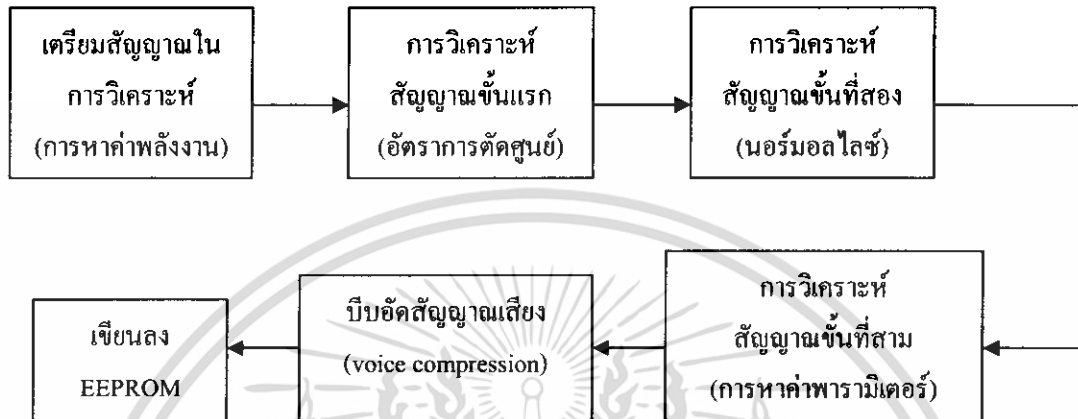
เมื่อทราบความถี่จาก MCLK และความถี่สุ่มตัวอย่างที่ออกแบบ จะทำให้สามารถคำนวณค่ารีจิสเตอร์ M1 และ N1 ได้ สิ่งที่ต้องพิจารณาอีกสิ่งหนึ่ง คือช่วงของแต่ละรีจิสเตอร์ ควรเป็นช่วงที่น้อยที่สุดของการปรับแก้อัตราความถี่ของเฟสล็อกคูล์ นั้นหมายความว่า M1 และ N1 นั้นต้องเป็นอัตราส่วนอย่างต่ำ นอกจากนั้น ค่าที่คำนวณออกมาต้องลบออก “1” ในเวลาเขียนข้อมูลลงรีจิสเตอร์เสมอ เนื่องจากการเก็บค่าในรีจิสเตอร์ จะเริ่มตั้งแต่ 0

3.2 หน่วยแสดงผล

ในโครงการนี้ผู้จัดทำได้ใช้จอแสดงผลในการแสดงค่าฟังก์ชันของการบีบอัดสัญญาณเสียงในแต่ละโหมดการทำงานของโปรแกรม ดังรูปที่ 3.1(B) และเนื่องจากผู้จัดทำใช้บอร์ดทดลอง dsPICDEM 1.1 ในการพัฒนาโปรแกรม ซึ่งในบอร์ดนี้มีส่วนแสดงผลอยู่แล้วอันประกอบด้วยจอแสดงผลแบบผลึกเหลว (Liquid Crystal Display : LCD) และหลอดไฟ ในส่วนของจอแสดงผลเป็นแบบคอตเมตริกซ์ (dot-matrix) 122x32 ขนาด 4 บรรทัด 20 ตัวอักษร ใช้ไอซี PIC18F242 LCD เป็นคอนโทรลเลอร์ ติดต่อกับ LCD ในโหมด 8 บิต และ dsPIC30F6014 ซึ่งเป็นไมโครคอนโทรลเลอร์ตัวหลักจะทำการเชื่อมต่อกับไอซี PIC18F242 LCD ผ่านการสื่อสารแบบอนุกรม SPI โดยผู้จัดทำกำหนดให้ dsPIC30F6014 เป็นมาสเตอร์ และไอซี PIC18F242 LCD เป็นสเลฟ ดังที่ได้กล่าวถึงลักษณะการเชื่อมต่อมาแล้วในบทที่ 2

3.3 หน่วยประมวลผลกลาง

ในที่นี้ผู้จัดทำจะแยกอธิบายถึงกระบวนการบีบอัดสัญญาณเสียง และขั้นตอนการทำงานของโปรแกรม (flow chart) โดยรวมเพื่อให้ง่ายต่อการเข้าใจดังนี้



รูปที่ 3.7 บล็อกโคอะแกรมขั้นตอนการบีบอัดสัญญาณเสียง

จากบล็อกโคอะแกรมขั้นตอนการบีบอัดสัญญาณเสียง เราสามารถอธิบายเป็นส่วนๆ ได้ดังนี้

3.3.1 การเตรียมสัญญาณในการวิเคราะห์

เป็นขั้นตอนการเตรียมสัญญาณ เพื่อนำไปวิเคราะห์สัญญาณในขั้นต่อไป โดยในขั้นตอนนี้เป็นการหาขอบเขตของค่าซึ่งเป็นส่วนที่แท้จริงที่จะนำไปใช้ต่อในส่วนอื่นๆ

การวิเคราะห์หาค่าพลังงาน

ใช้หลักการวิเคราะห์สัญญาณในโดเมนเวลาเข้ามาช่วย สำหรับขั้นตอนนี้มีดังต่อไปนี้

1. คำนวณค่าพลังงานของสัญญาณในช่วงเวลาที่ถูกบันทึกทั้งหมด โดยการคำนวณหาค่าพลังงานของสัญญาณซึ่งจะทำการคำนวณเป็นเฟรม กำหนดให้เฟรมหนึ่งมีจำนวนแซมเปิล 160 แซมเปิล และทำการคำนวณเป็นเฟรมต่อไปเรื่อยๆจนครบทุกเฟรม
2. พิจารณาค่าพลังงานสูงสุดของสัญญาณ (E_{max}) และพิจารณาว่า ค่าพลังงานสูงสุดที่หาได้นั้นอยู่ที่เฟรมใด แล้วสมมติให้เป็นเฟรมที่ m
3. ทำการกำหนดพลังงานอ้างอิง (E_{ref}) โดยพลังงานอ้างอิงสามารถคำนวณได้จาก α เท่าของพลังงานสูงสุด ($\alpha * E_{max}$) ซึ่งในที่นี้เลือกใช้ค่า $\alpha = 0.1$ จากการที่หาค่าพลังงานสูงสุดจะทำให้ทราบว่า เฟรมเริ่มต้นและเฟรมสุดท้ายของค่าจะอยู่นอกช่วงที่เราหา $\alpha * E_{max}$ ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

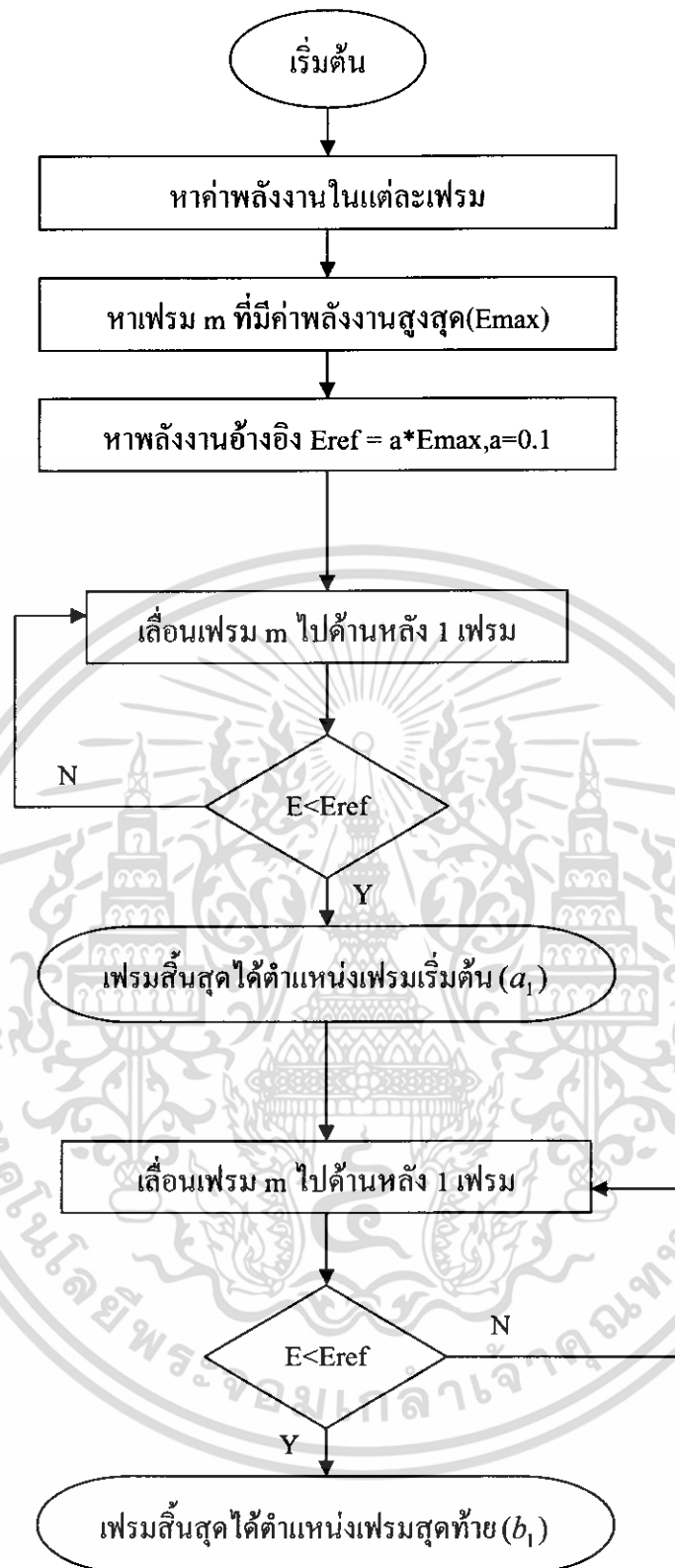
4. ทำการหาเฟรมเริ่มต้นของค่า โดยจะเริ่มพิจารณาที่ค่าพลังงานสูงสุดเป็นเฟรมแรก (เฟรมที่ m) ซึ่งนำค่าพลังงานของสัญญาณ เฟรมที่ m ไปเปรียบเทียบกับพลังงานอ้างอิง แล้วเลื่อนเฟรมที่จะเปรียบเทียบกับพลังงานอ้างอิงไปด้านหลังเรื่อยๆ (เฟรมที่ $m-1$, $m-2$, ..., 2 , 1) จนกว่าจะพบเฟรมที่มีค่าพลังงานน้อยกว่าค่าพลังงานอ้างอิงเป็นเฟรมแรก ก็จะได้เฟรมเริ่มต้นของสัญญาณ (a_1)

5. การหาเฟรมสุดท้ายของค่า จะทำการพิจารณาเช่นเดียวกับการหาเฟรมเริ่มต้นของค่า คือ เริ่มเปรียบเทียบจากเฟรมที่มีพลังงานสูงสุด (เฟรมที่ m) เช่นกัน แต่การเลื่อนเฟรมจะเลื่อนไปทางด้านหน้าแทน (เฟรมที่ $m+1$, $m+2$, ...) จนกว่าจะพบเฟรมที่มีค่าพลังงานน้อยกว่าค่าพลังงานอ้างอิง ก็จะได้จุดสุดท้ายของสัญญาณ (b_1)

สามารถหาเฟรมเริ่มต้นและเฟรมสุดท้าย ได้ดังรูปที่ 3.8



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.8 การหาจุดเริ่มต้นและจุดสุดท้ายของสัญญาณเสียงโดยการเปรียบเทียบพลังงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.2 การวิเคราะห์สัญญาณขั้นที่หนึ่ง

เป็นขั้นตอนการวิเคราะห์สัญญาณ เพื่อให้เป็นการแน่ใจว่า การหาขอบเขตของค่าในขั้นตอนการหาค่าพลังงานนั้นถูกต้อง ทำให้การการบีบอัดสัญญาณเสียงมีความแม่นยำมากขึ้น

การวิเคราะห์หาค่าอัตราการตัดศูนย์

เนื่องจากการใช้การวิเคราะห์โดยการหาค่าพลังงานของเสียงอาจเกิดการผิดพลาด จึงนำเอาการวิเคราะห์หาค่าอัตราการตัดศูนย์มาร่วมพิจารณา โดยมีวิธีการดังนี้

1. คำนวณหาค่าอัตราการตัดศูนย์ในช่วงเวลาที่ถูกรับที่ทั้งหมด โดยการคำนวณหาค่าอัตราการตัดศูนย์จะคำนวณเป็นเฟรม กำหนดให้แต่ละเฟรมมีตัวอย่าง 160 ตัวอย่าง และทำการคำนวณจนครบช่วงเวลาของเสียงที่ถูกรับที่ทั้งหมด

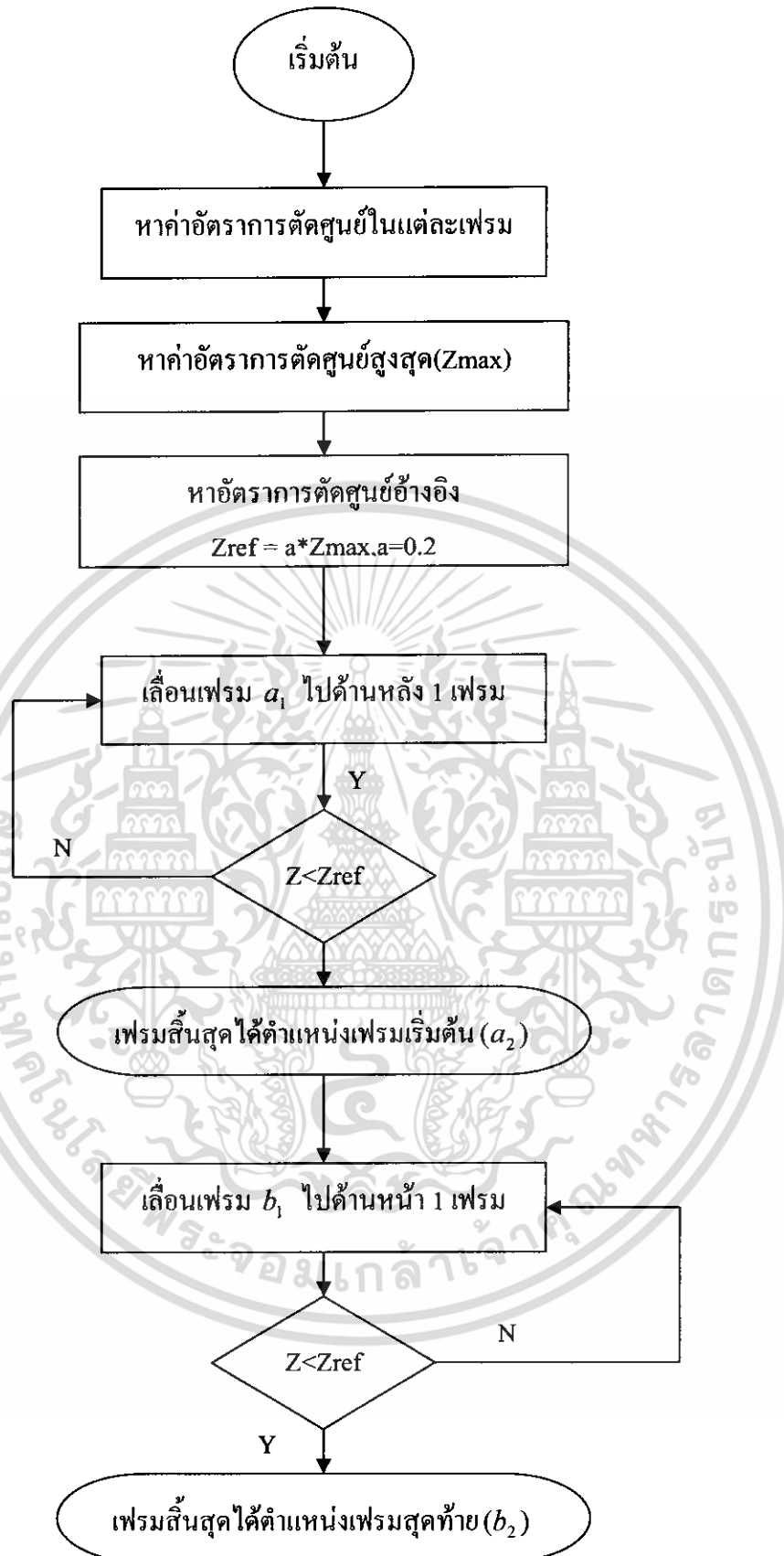
2. กำหนดระดับสูงสุด (Z_{max}) ให้เป็นเฟรมที่ p

3. กำหนดระดับอ้างอิง (Z_{ref}) มีค่าเป็น a เท่าของอัตราการตัดศูนย์สูงสุด ($a * Z_{ref}$) โดยค่า a จะเท่ากับ 0.1 เช่นกัน

4. จากจุดเริ่มต้น (a_1) ที่หาได้จากขั้นตอนที่ 1 จะเป็นเฟรมแรกที่พิจารณาโดยจะทำการเปรียบเทียบอัตราการตัดศูนย์ของเฟรมดังกล่าวกับอัตราการตัดศูนย์อ้างอิง ถ้ามีค่ามากกว่า เฟรมเริ่มต้นจะถูกเปลี่ยนตำแหน่ง โดยจะเลื่อนไปยังเฟรมแรกที่มีค่าน้อยกว่าอัตราการตัดศูนย์อ้างอิง (a_2) และจะถือว่าจุดนี้เป็นเฟรมเริ่มต้นที่แท้จริง ซึ่งการเลื่อนเฟรมจะเลื่อนไปด้านหลัง (เฟรมที่ a_{-1}, a_{-2}, \dots) แต่ในกรณีที่ a_1 น้อยกว่าอัตราการตัดศูนย์อ้างอิง จะถือว่า a_1 ที่หาได้จากขั้นตอนที่ 1 เป็นเฟรมเริ่มต้นที่ถูกต้องอยู่แล้ว

5. การหาเฟรมสุดท้าย จากเฟรมสุดท้าย (b_2) ที่ได้จากขั้นตอนที่ 1 จะพิจารณาเหมือนข้อ 3 คือจะเป็นเฟรมแรกที่จะเปรียบเทียบอัตราการตัดศูนย์อ้างอิง ถ้ามีค่ามากกว่า เฟรมสุดท้ายจะถูกเปลี่ยนตำแหน่ง โดยจะเลื่อนเฟรมจะเลื่อนไปด้านหน้า (เฟรมที่ b_{+1}, b_{+2}, \dots) จนกว่าจะเจอเฟรมแรกที่มีค่าน้อยกว่าอัตราการตัดศูนย์อ้างอิง (b_2) และจะถือว่าเฟรมนี้เป็นเฟรมสุดท้ายที่แท้จริง แต่ถ้าอัตราการตัดศูนย์ที่ b_1 น้อยกว่าอัตราการตัดศูนย์อ้างอิงก็กล่าวได้ว่า b_1 เป็นเฟรมสุดท้ายที่ถูกต้องอยู่แล้วเช่นกัน

จากขั้นตอนทั้งหมดที่กล่าวมานี้สามารถดำเนินการได้ดังรูปที่ 3.9



รูปที่ 3.9 การหาเฟรมเริ่มต้นและเฟรมสุดท้ายของสัญญาณเสียง โดยการหาอัตราการตัดศูนย์

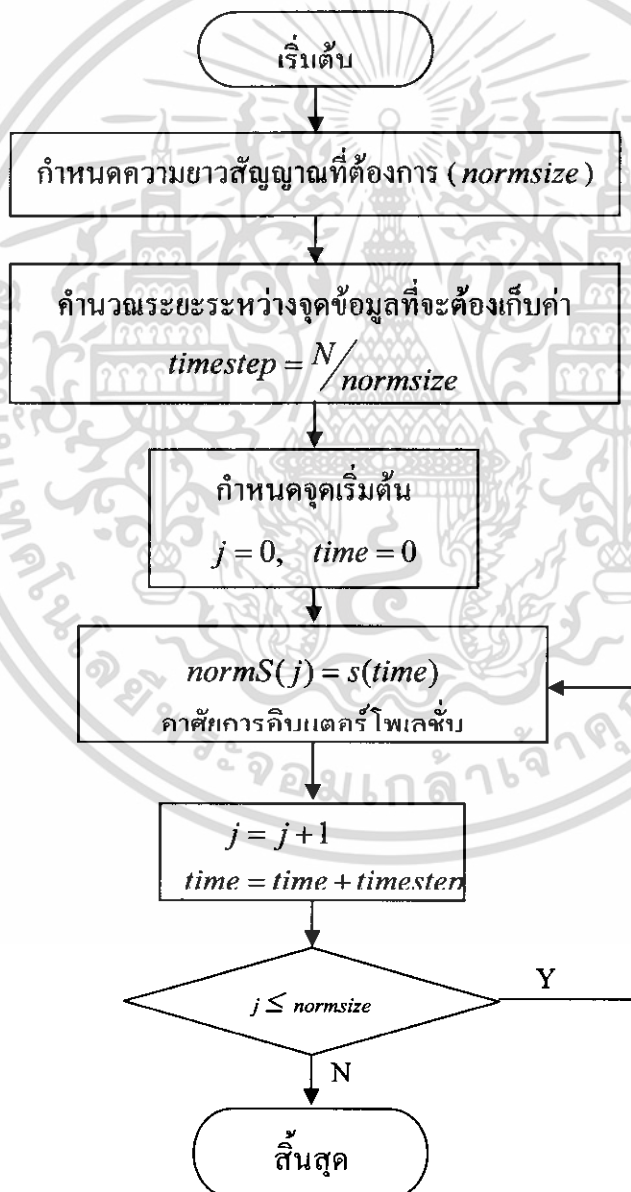
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.3 การวิเคราะห์สัญญาณขั้นที่สอง

เนื่องจากสัญญาณเสียงพูดแต่ละคำมีความยาวไม่เท่ากันจึงต้องมีการนอร์มอลไลซ์สัญญาณเสียงพูดให้ยาวเท่ากัน

การนอร์มอลไลซ์

การนอร์มอลไลซ์ทำขึ้นเพื่อให้สัญญาณที่สนใจมีความยาวเท่าๆ กันโดยหลักการเปลี่ยนคาบการสุ่ม และใช้หลักการอินเตอร์โพลชัน(Interpolation) ในการประมาณค่าหากจุดที่ทำการสุ่มใหม่นั้นไม่ตรงกับจุดซึ่งมีข้อมูลอยู่ ขั้นตอนการนอร์มอลไลซ์สัญญาณเสียงพูดแสดงดังรูปที่ 3.10 เมื่อ สัญญาณอินพุตคือ $s(n)$ โดยมีความยาว N และเอาต์พุตคือสัญญาณ $normS(n)$



รูปที่ 3.10 ขั้นตอนการนอร์มอลไลซ์สัญญาณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.2 การวิเคราะห์สัญญาณขั้นที่สาม

เป็นการหาค่าพารามิเตอร์แบ่งออกเป็น

การหาค่าพรีเอมฟาซิส

เป็นการนำสัญญาณที่ได้จากขั้นตอนที่สองมาผ่านระบบซึ่งมีฟังก์ชันถ่ายโอน

$$H(z) = 1 - az^{-1} \text{ เมื่อ } 0.9 \leq a \leq 1.0$$

โดยในโครงการนี้เลือกใช้ $a = 15/16 = 0.937$

การแบ่งช่วงสัญญาณ

จากสัญญาณที่ผ่านการพรีเอมฟาซิสจะแบ่งพิจารณาเป็นเฟรมโดยแต่ละเฟรมมีขนาด N จุดข้อมูลและมีข้อมูลเหลื่อมกับเฟรมก่อนหน้า M จุดข้อมูล โดยต้องให้ค่า M มีค่าน้อยกว่า N การคำนวณค่า n เฟรมที่ l สามารถทำได้โดยอาศัยสมการ

$$x_l(n) = s'(Ml + n) \text{ เมื่อ } n = 0, 1, \dots, N-1$$

ค่าที่ใช้ในการทดลองคือ $M = 20$ และ $N = 160$ และในการทดลองนี้ทำการกำหนดให้มีการสุ่มตัวอย่างของสัญญาณด้วยความถี่ 8,000 เฮิรตซ์ ดังนั้นช่วงเวลาในการวิเคราะห์ของแต่ละเฟรมคือ $160/8000 = 20ms$ และเวลาที่เหลื่อมกันในแต่ละเฟรมคือ $20/8000 = 2.5ms$

และเพื่อลดความไม่ต่อเนื่องของเฟรม จะนำสัญญาณที่แบ่งเป็นช่วงนั้นผ่านเข้าช่องแคบแฮมมิงซึ่งมีสมการเป็น

$$w(n) = 0.54 - 0.46 \cos\left(2\pi \frac{n}{N}\right), \quad 0 \leq n \leq N-1$$

การวิเคราะห์คุณลักษณะของเสียง

ในการวิเคราะห์ใช้การหาค่าสัมประสิทธิ์ซีอีแอลทีเพราะเป็นวิธีที่ทำให้ข้อมูลมีขนาดเล็ก แต่ให้ผลการทดลองในระดับที่ยอมรับได้ และเป็นหลักของ Speex เช่นกัน

3.3.5 การเวทค่าพารามิเตอร์

ทำการเวท หรือถ่วงน้ำหนักสัมประสิทธิ์เซปสตรีมที่ได้ ช่วงลำดับต้นๆและท้ายๆของเฟรมที่นำมาวิเคราะห์จะเกิดความคลาดเคลื่อนมากกว่าบริเวณส่วนอื่น เพราะฉะนั้น เราจะทำการลดความคลาดเคลื่อนเหล่านั้นด้วยฟังก์ชันเวทดัง (ดูได้จากบทที่ 2) จะได้พารามิเตอร์สุดท้ายคือ

$$C_m' = C_m * W_m$$

3.3.6 การจัดระดับเวกเตอร์

การจัดระดับเวกเตอร์เป็นเทคนิคการเข้ารหัสเทคนิคหนึ่งที่มีคุณภาพ มันจะทำการเข้ารหัสเวกเตอร์ พารามิเตอร์ของคลื่น ที่แทนสเปกตรัมของเสียง ซึ่งเป็นเวกเตอร์อินพุตที่ตัวเลขเป็นจำนวนเต็ม โดยนำไปเปรียบเทียบกับโค้ดบุคที่ได้สร้างเก็บไว้ก่อนแล้ว โค้ดบุคที่ถูกเลือกนั้นจะเป็นตัวใกล้เคียงกับเวกเตอร์อินพุตผลที่ได้จากเวกเตอร์ควอนไทเซชัน ที่แสดงออกมาขึ้นอยู่กับเซตของโค้ดบุคที่สร้างเก็บไว้ก่อนแล้วนั่นเอง ขั้นตอนการจัดระดับเวกเตอร์ มี 2 ขั้นตอน คือ การสร้างโค้ดบุค และการเปรียบเทียบ

1. การสร้างโค้ดบุค

สร้างโค้ดบุคขึ้นมาเป็นตัวแทนสเปกตรัมของเสียงที่สังเกตได้ เพื่อลดจำนวนข้อมูลของพารามิเตอร์ที่ได้มาจากขั้นตอนการวิเคราะห์สัญญาณขั้นที่ 3 การสร้างโค้ดบุคมีขั้นตอนดังนี้

นำเทรนนึงเซตมาใช้ในการสร้างโค้ดบุค

ขนาดโค้ดบุคของการควอนไทซ์ แบบเวกเตอร์ คือ $M = 2^b$ เวกเตอร์ และ เพื่อที่จะหาเซตของ M โค้ดบุคที่ดีที่สุด จำนวนเวกเตอร์อินพุตจะต้องมากกว่าโค้ดบุคหลายๆ

การสุ่มค่าเริ่มต้น

การสุ่มค่าเริ่มต้น เป็นวิธีหนึ่งในการออกแบบโค้ดบุค ซึ่งคือการเลือกค่าเริ่มต้นของโค้ดบุค เรียกโค้ดบุคที่ได้จากการสุ่มค่าเริ่มต้นนี้ว่า แรนดอมโค้ดบุค (Random Codebook) ถึงแม้วิธีนี้จะไม่ใช่วิธีที่ค้นหาค่าโค้ดบุคที่ได้จากการสุ่มก็ให้ผลเป็นที่ยอมรับ

การหาความคลาดเคลื่อน

การหาความคลาดเคลื่อนเป็นส่วนที่จำเป็นและเป็นประโยชน์ต่อการออกแบบโค้ดบุค สมการทางพีชคณิตที่ใช้ในการหาระยะทางมีหลายวิธี แต่ที่นำมาใช้คือ การหาความคลาดเคลื่อนกำลังสองรวม (Total Square Error) ซึ่งเป็นวิธีที่ง่ายและเร็ว ถ้าสัญญาณมี P มิติ สามารถหาระยะห่างระหว่างสัญญาณอินพุต (x) กับเวกเตอร์โค้ด (y) โดยสมการ

$$d(v_1, v_2) = \|v_1 - v_2\| = \sum_{i=0}^{k-1} (x_i - y_i)^2$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การจัดกลุ่มและการหาจุดศูนย์กลางของกลุ่ม

การจัดกลุ่มเป็นการจัดเวกเตอร์อินพุตเข้าไปตามกลุ่มต่างๆ ของเรนคอมโด้คบุก โดยพิจารณาระยะทางหรือความคลาดเคลื่อนน้อยที่สุด ของเวกเตอร์อินพุต x กับเวกเตอร์โด้คบุก y ซึ่งเป็นโด้คบุก จากนั้นจะทำการหาค่าเฉลี่ยของแต่ละกลุ่ม เพื่อเป็นค่ากลางของกลุ่มนั้นๆจะได้

$$\bar{Y} = \frac{1}{L} \sum_{i=1}^L x_i$$

\bar{Y} เป็นจุดศูนย์กลางซึ่งเป็นเวกเตอร์ที่อยู่ตรงกลางของ $\{x_i\}_{i=1}^L$ ซึ่งแต่ละมิติจะไม่ขึ้นแก่กัน หมายความว่าแต่ละ y_k เป็นค่ากลางของ $\{x_i\}_{i=1}^L$ ทำ 2 ขั้นตอนนี้ซ้ำ จะเกิดการลู่เข้า โดยความคลาดเคลื่อนรวมจะต่ำกว่าค่าหนึ่งๆ ซึ่งค่าความคลาดเคลื่อนรวมจะลดลงทุกครั้งที่มีการคำนวณซ้ำใหม่ จึงขึ้นกับค่าที่กำหนดว่าจะต้องการให้ความคลาดเคลื่อนรวมน้อยกว่าค่ากลางดังกล่าวของแต่ละกลุ่ม จะถูกเก็บเป็นเวกเตอร์โด้คบุกจะได้ว่า y เป็นควอนโด้คบุกของค่า x

โดย $q(\cdot)$ เป็นคอเปอร์เรเตอร์ของการควอนโด้คบุก y ถูกเรียกว่าเอาท์พุทเวกเตอร์ของค่า x โดย y เป็นค่าโด้คบุกหนึ่งใน $y_{ii} = [y_{i1} y_{i2} \dots y_{ip}]$

$$Y = \{y_i; 1 \leq i \leq M\} \quad \text{โดย} \quad y_{ii} = [y_{i1} y_{i2} \dots y_{ip}]$$

Y เป็นเซตของโด้คบุก M เป็นขนาดของโด้คบุกและ $\{y_i\}$ เป็นเซตของเวกเตอร์โด้คบุก อาจเรียกว่าโด้คบุกอ้างอิงและ M อาจเรียกว่าจำนวนระดับชั้น จะทำการแบ่งเวกเตอร์ x ไปใน M เซล $\{C_i; 1 \leq i \leq M\}$ เมื่อ x อยู่ในเซล C_i

$$q(x) = y_i \quad \text{ถ้า} \quad x \in C_i$$

2. การเปรียบเทียบ

เป็นขั้นตอนการเปรียบเทียบพารามิเตอร์กับโด้คบุกอ้างอิงที่จัดทำไว้แล้ว ผลลัพธ์ที่ได้จะเป็นเวกเตอร์โด้คบุกที่มีจำนวนข้อมูลลดลงจากขั้นตอนการวิเคราะห์ขั้นที่ 3 โดยเป็นการวัดระยะห่างระหว่างเวกเตอร์โด้คบุกกับพารามิเตอร์ที่ต้องการทดสอบ ด้วยวิธีการหาความคลาดเคลื่อนกำลัง 2 รวม (Total Square Error)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

การทดลอง

ในบทนี้จะกล่าวถึงการทดลองและผลการทดลอง ซึ่งในโครงการนี้ ผู้จัดทำได้เก็บตัวอย่างเสียงสนทนาสองเสียง จากนั้นทำการวิเคราะห์ด้วยเสียงสนทนานั้นด้วยโปรแกรม “matlab” ซึ่งเป็นโปรแกรมที่นิยมใช้กันอย่างแพร่หลาย เพื่อเพื่อผลตอบสนองเชิงความถี่ และผลตอบสนองเชิงเวลาของเสียงสนทนาทั้งสอง ผู้จัดทำได้ทำการศึกษาเปรียบเทียบตามวัตถุประสงค์ดังนี้

4.1 วัตถุประสงค์ของการทดลอง

1. ศึกษาและเปรียบเทียบผลของการเปลี่ยนค่าเฟสล็อกกลุ่ม ซึ่งใช้ในการเชื่อมต่อแบบ DCI ระหว่าง ไอซีเข้าและถอครหัสเสียง Si3000 กับ ไมโครคอนโทรลเลอร์ โดยเลือกใช้ออสซิลเลเตอร์จากภายนอก และกำหนดค่ารีจิสเตอร์ควบคุมจากการคำนวณในบทที่ 3 ดังแสดงในตารางที่ 4.1

ตารางที่ 4.1 ความถี่ออสซิลเลเตอร์ และค่ารีจิสเตอร์ควบคุมที่ใช้ในการเปลี่ยนค่าเฟสล็อกกลุ่ม

ความถี่ออสซิลเลเตอร์ (MHz)	รีจิสเตอร์ควบคุม N1	รีจิสเตอร์ควบคุม M1
6.144	2	13
6.144	1	6

2. ศึกษาและเปรียบเทียบผลตอบสนองเชิงความถี่ และผลตอบสนองเชิงเวลาของเสียงสนทนา ระหว่างเสียงสนทนาหลังเข้าสู่ระบบที่ถูกบันทึกโดยตรง และเสียงสนทนาที่ถูกบันทึกไว้หลังผ่านระบบที่มีการบีบอัดสัญญาณเสียง

ผู้จัดทำได้ทำการออกแบบการทดลองออกเป็นสองตอน เพื่อศึกษาผลของการเปลี่ยนค่าเฟสล็อกกลุ่มต่อผลตอบสนองเชิงความถี่ และผลตอบสนองเชิงเวลาของเสียงสนทนา เมื่อใช้ความถี่สุ่มตัวอย่างที่ 8 KHz ดังนี้

ตอนที่ 1 .เสียงสนทนาที่ 1 ใช้ความถี่ออสซิลเลเตอร์ 6.144MHz รีจิสเตอร์ควบคุม N1 : M1 คือ 1 : 6 และ 2 : 13

ตอนที่ 2 .เสียงสนทนาที่ 2 ใช้ความถี่ออสซิลเลเตอร์ 6.144MHz รีจิสเตอร์ควบคุม N1 : M1 คือ 1 : 6 และ 2 : 13

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

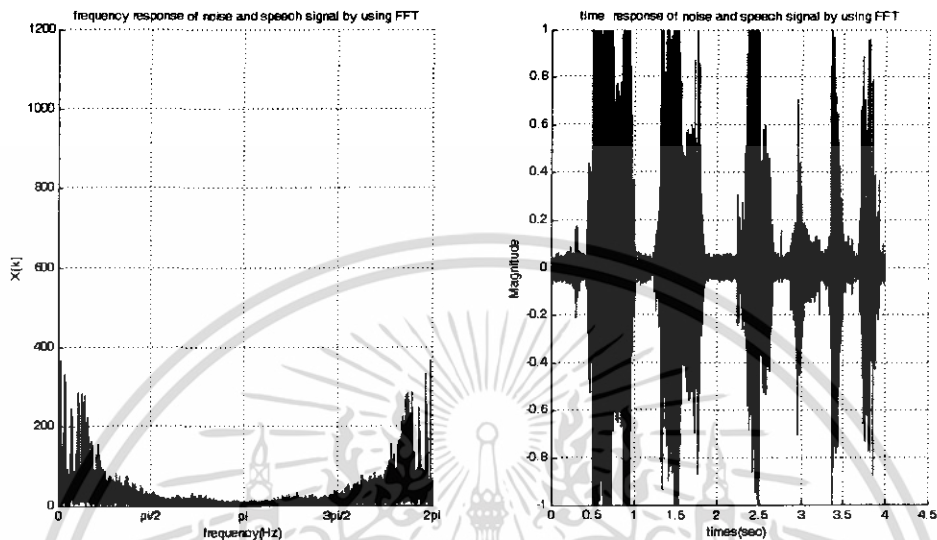
4.2 วิธีการทดลอง

1. บันทึกเสียงสนทนา 2 สัญญาณเสียง โดยใช้โปรแกรม wavepad บันทึกเป็นไฟล์เสียง PCM ออกดิโอแซมเปิ้ล 16 บิต ความถี่สุ่มตัวอย่าง 8 kHz
2. พล็อตสัญญาณเสียงก่อนการบีบอัดสัญญาณเสียงแบบผลตอบสนองเชิงความถี่ และผลตอบสนองเชิงเวลาของเสียงสนทนา ด้วยโปรแกรม “matlab”
3. เปิดเสียงสนทนาที่บันทึกไว้ในข้อ 1 จากนั้นต่อสายออกดิโอของเครื่องเล่นเสียงเข้ากับบอร์ดทดลอง (mic in : J16)
4. ทำการบีบอัดสัญญาณเสียง โดยใช้ความถี่ออสซิลเลเตอร์ 6.144MHz รีจิสเตอร์ควบคุม N1 : M1 คือ 1:6
5. บันทึกเสียงสนทนาด้วยโปรแกรม wavepad เช่นเดียวกับข้อ 1
6. พล็อตสัญญาณเสียงหลังการบีบอัดสัญญาณเสียงแบบผลตอบสนองเชิงความถี่ และผลตอบสนองเชิงเวลาของเสียงสนทนา ด้วยโปรแกรม “matlab”
7. เปลี่ยนรีจิสเตอร์ควบคุม N1 : M1 เป็น 2:13 จากนั้นทำการทดลองในข้อ 3,4,5,6 อีกครั้ง

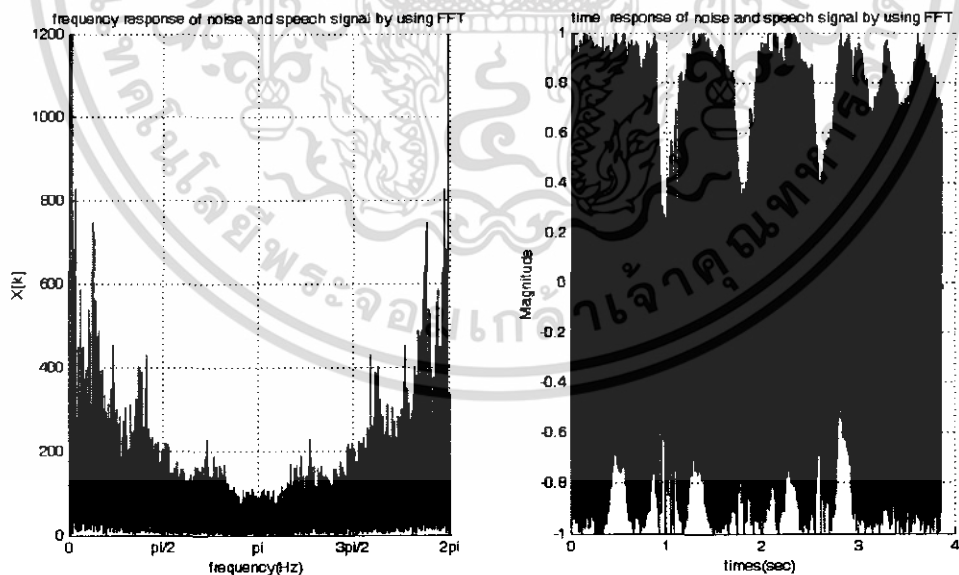
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 ผลการทดลอง

ตอนที่ 1 ศึกษาผลตอบสนองเชิงความถี่ และผลตอบสนองเชิงเวลาของเสียงสนทนาที่ 1 เมื่อ กำหนดให้ความถี่ออสซิลเลเตอร์ 6.144 MHz รัจิสเตอร์ควมคุม N1:M1 คือ 1:6



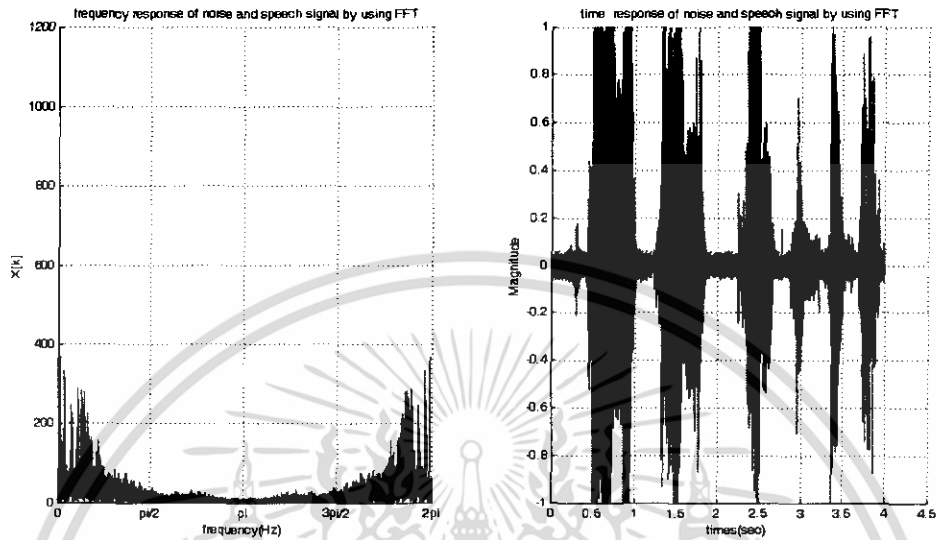
รูปที่ 4.1 เสียงสนทนาที่ 1 ก่อนเข้าสู่ระบบ



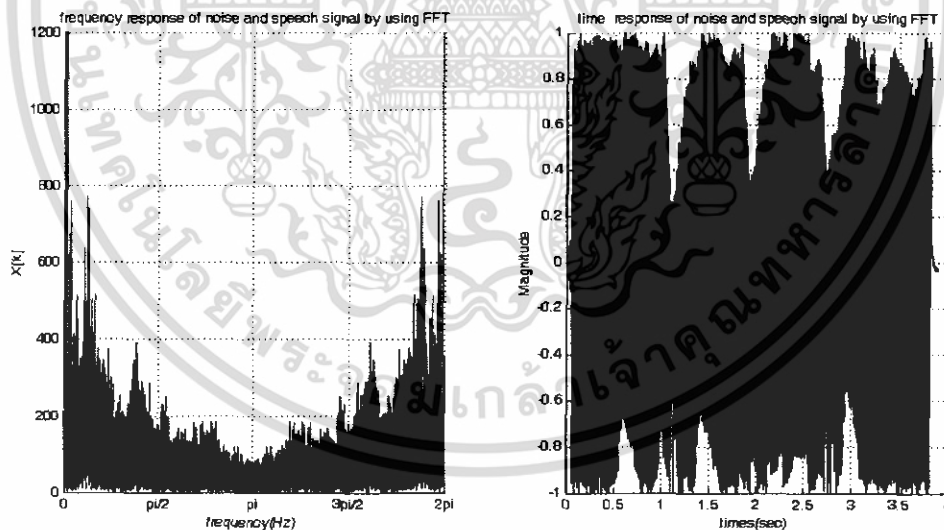
รูปที่ 4.2 เสียงสนทนาที่ 1 หลังเข้าสู่ระบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตอนที่ 1(ต่อ) ศึกษาผลตอบสนองเชิงความถี่ และผลตอบสนองเชิงเวลาของเสียงสนทนาที่ 1 เมื่อ กำหนดให้ความถี่ออสซิลเลเตอร์ 6.144 MHz รัจิสเตอร์ควบคุม N1:M1 คือ 2:13



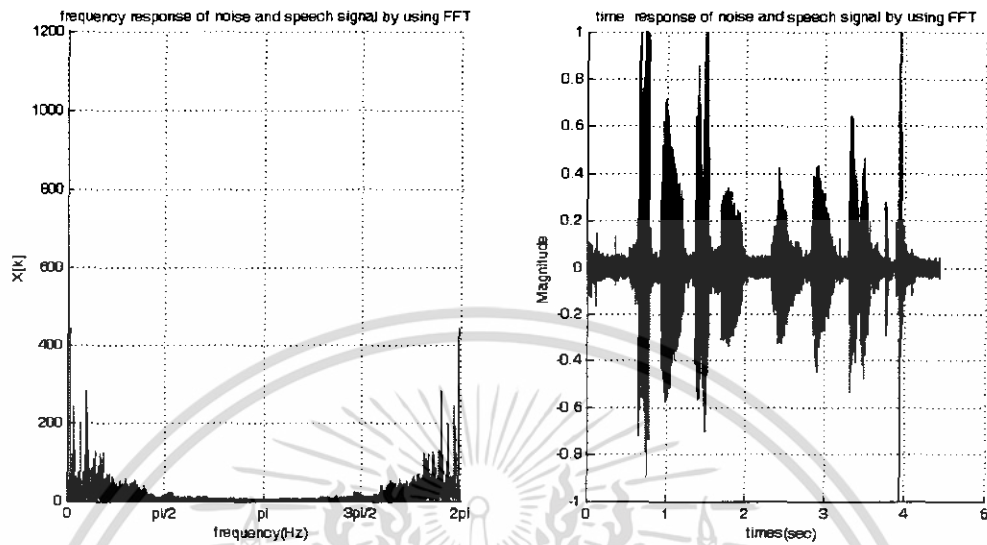
รูปที่ 4.3 เสียงสนทนาที่ 1 ก่อนเข้าสู่ระบบ



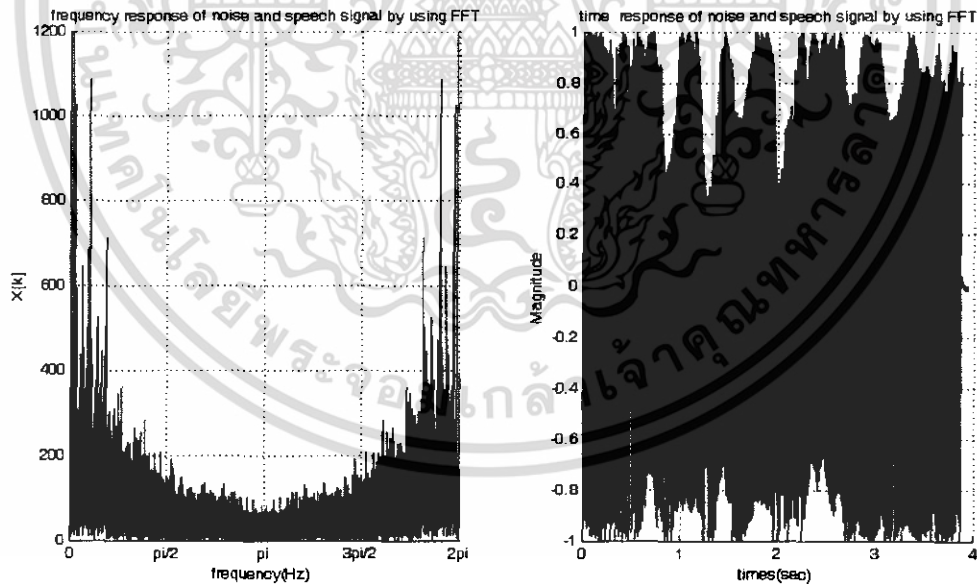
รูปที่ 4.4 เสียงสนทนาที่ 1 หลังเข้าสู่ระบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตอนที่ 2 ศึกษาผลตอบสนองเชิงความถี่ และผลตอบสนองเชิงเวลาของเสียงสนทนาที่ 2 เมื่อ กำหนดให้ความถี่ออสซิลเลเตอร์ 6.144 MHz รัจิสเตอร์ควมคุม N1:M1 คือ 1:6



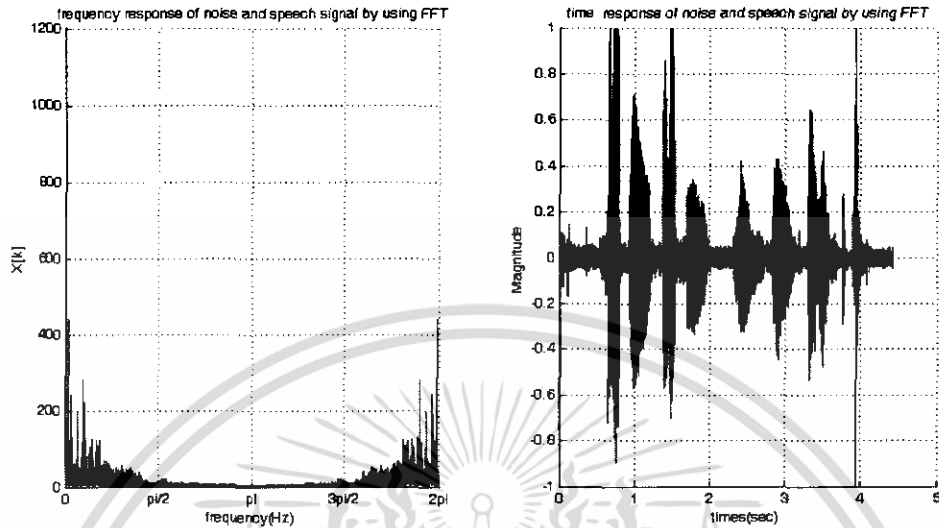
รูปที่ 4.5 เสียงสนทนาที่ 2 ก่อนเข้าสู่ระบบ



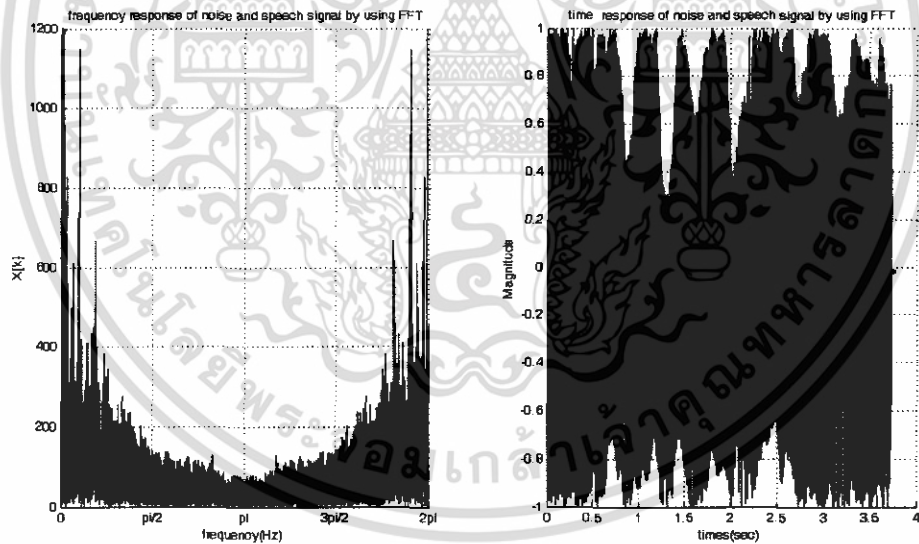
รูปที่ 4.6 เสียงสนทนาที่ 2 หลังเข้าสู่ระบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตอนที่ 2(ต่อ) ศึกษาผลตอบสนองเชิงความถี่ และผลตอบสนองเชิงเวลาของเสียงสนทนาที่ 2 เมื่อ กำหนดให้ความถี่ออสซิลเลเตอร์ 6.144 MHz รีจิสเตอร์ควบคุม N1:M1 คือ 2:13



รูปที่ 4.7 เสียงสนทนาที่ 2 ก่อนเข้าสู่ระบบ



รูปที่ 4.8 เสียงสนทนาที่ 2 หลังเข้าสู่ระบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 สรุปผลการทดลอง

1. ศึกษาผลการเปลี่ยนค่าเฟสล็กอกลูประหว่างไอซีเข้าและถอดรหัสเสียง กับไมโครคอนโทรลเลอร์ โดยเลือกใช้ออสซิลเลเตอร์จากภายนอก และมีการกำหนดค่ารีจิสเตอร์ควบคุมจากการคำนวณ

เมื่อเปลี่ยนค่ารีจิสเตอร์ควบคุม N1:M1 จากเดิม 1:6 เป็น 2:13 พบว่าไอซี Si3000 จะใช้เวลาส่งเฟรมปฐมภูมิไปยังไมโครคอนโทรลเลอร์นานกว่า นั่นคือรีจิสเตอร์ควบคุม 1:6 จะใช้เวลาในการสร้างความถี่สุ่มตัวอย่างนานกว่า ซึ่งสอดคล้องกับการเขียนค่าเฟสล็กอกลูใน Si3000 ที่ค่าจริงจากการคำนวณจะต้องน้อยกว่าค่าเดิมหนึ่งค่าเวลาเขียนโปรแกรม แต่ในทางกลับกันเมื่อ Si3000 สามารถสร้างความถี่สุ่มตัวอย่างได้สมบูรณ์แล้วจะสรุปผลตอบสนองของสัญญาณได้ดังนี้

ผลตอบสนองเชิงความถี่ของเสียงสนทนาที่ใช้รีจิสเตอร์ควบคุม 2:13 พบว่าขนาด $|H(j\omega)|$ จะมีขนาดสูงกว่า เสียงสนทนาที่ใช้รีจิสเตอร์ควบคุม 1:6 ดังแสดงในรูปที่ 4.2 กับ 4.4 เป็นต้น

ผลตอบสนองเชิงเวลาของเสียงสนทนาที่ใช้รีจิสเตอร์ควบคุม 1:6 พบว่าสามารถบีบอัดสัญญาณเสียง และขยายขนาดของเสียงสนทนาได้ดีกว่าการใช้รีจิสเตอร์ควบคุม 2:13 ที่จะเห็นได้ชัดว่าสัญญาณจะละเอียดกว่าดังแสดงในรูปที่ 4.6 กับ 4.8 เป็นต้น

2. ศึกษาเปรียบเทียบผลตอบสนองเชิงความถี่ และผลตอบสนองเชิงเวลาของเสียงสนทนา ระหว่างเสียงสนทนาที่ก่อนเข้าสู่ระบบที่ถูกบันทึกโดยตรง และเสียงสนทนาที่ถูกบันทึกไว้หลังผ่านระบบการบีบอัดสัญญาณเสียง

ผลตอบสนองเชิงความถี่ของเสียงสนทนา ที่ถูกบันทึกโดยตรงจะมีขนาดของโลบหลักสูง แต่ภายหลังเมื่อเสียงสนทนาเข้าสู่ระบบการบีบอัดสัญญาณเสียง โลบหลักจะมีขนาดสูงขึ้นและจะมีขนาดของโลบข้างที่กว้างขึ้น นั่นคือจะมีเสียงสัญญาณรบกวนเพิ่มเข้ามา ดังแสดงในรูปที่ 4.1 กับ 4.2, รูปที่ 4.3 กับ 4.4, รูปที่ 4.5 กับ 4.6, และรูปที่ 4.7 กับ 4.8 เป็นต้น

จากการทดลองทำให้ทราบว่า การบีบอัดสัญญาณเสียงโดยใช้ค่าเฟสล็กอกลู 1:6 และ 2:13 ให้ผลของเสียงที่มีคุณภาพใกล้เคียงกันมาก โดยที่ค่าเฟสล็กอกลู 1:6 จะให้สัญญาณเสียงที่ครบถ้วนกว่าที่ค่าเฟสล็กอกลู 2:13 เนื่องจากมีการแซมปลิงน้อยกว่านั่นเอง และเมื่อพิจารณาเสียงก่อนเข้าสู่ระบบเปรียบเทียบกับเสียงหลังเข้าสู่ระบบแล้ว พบว่าเสียงที่เข้าสู่ระบบจะมีสัญญาณรบกวนอยู่ด้วย ซึ่งเมื่อพิจารณาด้านคุณภาพแล้วจะเห็นว่า คุณภาพยังอยู่ในเกณฑ์ที่สามารถยอมรับได้ แสดงว่ามาตรฐาน Speex ให้ผลสัญญาณเสียงที่ยอมรับได้ในระดับที่น่าพอใจ ดังนั้น ถ้ามีการกำจัดสัญญาณรบกวนออกไปและพัฒนาโปรแกรมให้ดีขึ้นก็จะทำให้คุณภาพเสียงดีขึ้นกว่าเดิม ซึ่งจะอยู่ในขั้นตอนของการพัฒนาต่อไป

บทที่ 5

บทวิจารณ์และสรุป

5.1 สรุปผลการทดลอง

ปฏิญานิพนธ์ฉบับนี้นำเสนอ การบีบอัดสัญญาณเสียงโดยใช้มาตรฐาน Speex ซึ่งขั้นตอนการทดลองประกอบด้วยหน่วยย่อยสามส่วนคือ หน่วยประมวลผลกลาง หน่วยเข้ารหัสและถอดรหัสเสียง และหน่วยแสดงผล ซึ่งในโครงการนี้เราได้เลือกหน่วยประมวลผลกลางโดยใช้ไมโครคอนโทรลเลอร์ dsPIC30F6014 ขนาด 16 บิต เพื่อใช้ในการประมวลผลสัญญาณเสียงแทนเครื่องคอมพิวเตอร์ ซึ่ง dsPIC30F6014 นั้นมีโมดูลพิเศษต่างๆ ที่ใช้งานได้ง่าย และทันสมัย คุณสมบัติในหลายๆด้านนั้น dsPIC30F6014 สามารถทำงานได้ดีกว่าไมโครคอนโทรลเลอร์ตระกูลอื่นๆ อีกทั้งยังมีความสามารถรองรับการประมวลผลสัญญาณเสียงได้เป็นอย่างดี

เมื่อทำการทดลองกับสัญญาณเสียงแล้ว ผู้จัดทำได้พล็อตผลตอบสนองเชิงความถี่ และผลตอบสนองเชิงเวลาของสัญญาณเสียงสนทนาออกมา เปรียบเทียบผลของการเปลี่ยนค่าเฟสล็อกคูล์ ซึ่งใช้ในการเชื่อมต่อแบบ DCI ระหว่าง ไอซีเข้าและถอดรหัสเสียง Si3000 กับไมโครคอนโทรลเลอร์ โดยเลือกใช้ออสซิลเลเตอร์จากภายนอก และกำหนดค่ารีจิสเตอร์ควบคุมจากการคำนวณแล้ว ผลตอบสนองเชิงเวลาของเสียงสนทนาที่ใช้รีจิสเตอร์ควบคุม 1:6 พบว่าสามารถบีบอัดสัญญาณเสียง และขยายขนาดของเสียงสนทนาได้ไม่ต่างกับการใช้รีจิสเตอร์ควบคุม 2:13 และเมื่อเปรียบเทียบผลตอบสนองเชิงความถี่ และผลตอบสนองเชิงเวลาของเสียงสนทนา ระหว่างเสียงสนทนาหลังเข้าสู่ระบบที่ถูกบันทึกโดยตรง และเสียงสนทนาที่ถูกบันทึกไว้หลังผ่านระบบที่มีการบีบอัดสัญญาณเสียง พบว่าเสียงสนทนาหลังเข้าสู่ระบบ ผลตอบสนองเชิงความถี่ของเสียงสนทนา ที่ถูกบันทึกโดยตรงจะมีขนาดของโพลหลักสูงแต่ภายหลังเมื่อเสียงสนทนาเข้าสู่ระบบการบีบอัดสัญญาณเสียง โพลหลักจะมีขนาดลดลงแต่จะมีขนาดของโพลข้างที่กว้างขึ้น นั่นคือจะมีเสียงสัญญาณรบกวนแทรกเข้ามาด้วย ซึ่งจากการทดลอง เมื่อพิจารณาด้านคุณภาพของเสียงแล้วจะเห็นว่า คุณภาพยังอยู่ในเกณฑ์ที่สามารถยอมรับได้ แสดงให้เห็นว่ามาตรฐาน Speex เมื่อนำมาใช้ในการบีบอัดสัญญาณเสียงแล้วสามารถให้ผลสัญญาณเสียงที่ยอมรับได้ในระดับที่น่าพอใจ ดังนั้น ถ้ามีการกำจัดสัญญาณรบกวนออกไปและพัฒนาโปรแกรมให้ดีขึ้นก็จะทำให้คุณภาพเสียงดีขึ้นกว่าเดิม ซึ่งจะอยู่ในขั้นตอนของการพัฒนาต่อไป

5.2 ปัญหาที่พบและแนวทางแก้ไข

1. เสียงสนทนาที่ผ่านระบบการบีบอัดแล้วจะมีเสียงรบกวนแฝงอยู่ และยังคงมีเสียงก้อง เสียงสั้นอยู่บ้าง ซึ่งยังยอมรับได้ในระดับหนึ่ง แต่ยังไม่ดีพอสำหรับการบีบอัดเสียงที่สมบูรณ์แบบ จึงจำเป็นต้องศึกษาข้อดีอยู่แล้วปรับปรุงเพิ่ม
2. การพัฒนาด้านอุปกรณ์โดยใช้บอร์ดทดลองอาจไม่เสถียรเท่าการใช้คอมพิวเตอร์ แต่เป็นโครงการที่ขอส่วนการทดลองลงมาจึงเหมาะแก่การพัฒนาเริ่มต้นมากกว่าระบบใหญ่
3. การใช้บอร์ดทดลองเมื่อต่อใช้งานร่วมกับไมค์และลำโพงแล้ว ทำให้เกิดสัญญาณรบกวนเพิ่มขึ้น

5.3 ข้อเสนอแนะและแนวทางในการค้นคว้าพัฒนา

เพื่อประสิทธิภาพในการบีบอัดสัญญาณเสียงให้มีประสิทธิภาพมากยิ่งขึ้น ควรมีการวิจัยด้านการกำจัดเสียงรบกวน, เสียงก้อง หรือ การสั้นให้น้อยลง ซึ่งจากมาตรฐานการบีบอัดที่ใช้ นั้นถือว่ายอมรับได้ แต่ยังไม่ดีพอ ถ้ามีการพัฒนาอุปกรณ์ให้ดีขึ้น ก็จะสามารถพัฒนาเทคโนโลยีด้านการบีบอัดเสียงได้ดียิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

- [1] ชีรพงษ์ เจนวิริยะกุล, นรินทร์ทิพย์ เซาว์ทอง, ชญาดา ชันบุญโส, “การเข้ารหัสเสียงโดยวิธี เอ ดี พี ซี เอ็ม”, ปรินญาณิพนธ์วิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมโทรคมนาคม คณะวิศวกรรมศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2547.
- [2] ธนพล ตั้งชูพงศ์, ธนสิทธิ์ ชวลิตธารา, “การเข้ารหัสสัญญาณเสียงแบบ Real Time สำหรับมาตรฐาน ITU-T G.729 บน DSP chip TMS 320C6713”, ปรินญาณิพนธ์วิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2547.
- [3] เจ จันทรสุภฤกษ์, ชีรดา จันทรสมิ, “การเข้ามืบอัดเสียง”, ปรินญาณิพนธ์วิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมระบบควบคุม คณะวิศวกรรมศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2544.
- [4] ชัยยุทธ อัสวเมธีรุ่งโรจน์, เอกณรงค์ ใจยงค์, “การกำจัดเสียงรบกวนด้วยไมโครคอนโทรลเลอร์ dsPIC30F6014”, ปรินญาณิพนธ์วิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมระบบควบคุม คณะวิศวกรรมศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2549.
- [5] คำรัส โสมภีร์, ฤทัย รุ่งสีทอง, “การตระหนักรู้เสียงพูด”, ปรินญาณิพนธ์วิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมระบบควบคุม คณะวิศวกรรมศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2548.
- [6] บุญกิต นิลประพัฒน์, พัลลภ กลิ่นขจร, ภาคภูมิ ไพโรจน์, “การรู้จำเสียง”, ปรินญาณิพนธ์วิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมระบบควบคุม คณะวิศวกรรมศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2548.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาคผนวก

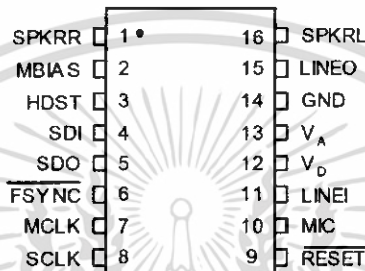
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

ไอซีเข้าและถอดรหัสเสียง Si3000

ไอซีเข้าและถอดรหัสเสียง Si3000 เป็นอุปกรณ์แปลงสัญญาณดิจิทัลเป็นแอนะล็อก และแอนะล็อกเป็นดิจิทัล 16 บิต มีรายละเอียดต่างๆดังนี้

1. Pin Descriptions: Si3000



Pin #	Pin Name	Description
1	SPKRR	Speaker Right Output. Analog output capable of driving a 60 Ω load.
2	MBIAS	Microphone bias output.
3	HDST	Handset Input/Output. Handset analog input/output.
4	SDI	Serial Port Data In. Serial communication and control data that is generated by the Si3000 to the system DSP.
5	SDO	Serial Port Data Out. Serial communication data that is provided by the Si3000 to the system DSP.
6	FSYNC	Frame Sync Output. Data framing signal that is used to indicate the start and stop of a communication data frame.
7	MCLK	Master Clock Input. High speed master clock input. Generally supplied by the system crystal clock or DSP.
8	SCLK	Serial Port Bit Clock Input/Output. Controls the serial data on SDO and latches the data on SDI. This pin is an input in slave mode and an output in master mode.
9	RESET	Reset. An active low input that is used to reset all control registers to a defined initialized state. Also used to bring the Si3000 out of sleep mode.
10	MIC	MIC Input. Microphone level or line level input. This input contains selectable gain of 0, 10, 20, or 30 dB with a full scale input level of 1 V_{RMS} .
11	LINEI	Line Input. Line level input with selectable gain of 0, 10, or 20 dB. The full scale input level is 1 V_{RMS} .
12	V_D	Digital Supply Voltage. Provides the digital supply voltage to the Si3000. Nominally either 5 or 3.3 V.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Pin #	Pin Name	Description
14	GND	Ground. Connects to the system digital ground.
15	LINEO	Line Output. Line level analog output with a 1 V _{RMS} full scale output level.
16	SPKRL	Speaker Left Output. Analog output capable of driving a 60 Ω load.

2. Control Registers

Note: Any register not listed here is reserved and should not be written. Any register bit labelled reserved should be written to zero during writes to the register. Register 0 can be read (always returns 0) and written safely.

Table 1. Register Summary

Register	Name	Bit 7	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0
1	Control 1	SR			SPD	LPD	HPD	MPD	CPD
2	Control 2				HPFD	PLL	DL1	DL2	
3	PLL1 Divide N1	Divider N1							
4	PLL1 Multiply M1	Multiplier M1							
5	RX Gain Control 1	LIG	LIM		MCG	MCM	HIM	IIR	
6	ADC Volume Control				RXG			LOM	HOM
7	DAC Volume Control				TXG			SLM	SRM
8	Status Report	SLSC	SRSC	LOSC					
9	Analog Attenuation					LOT		SOT	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Register 1. Control 1

Bit	D7	D6	D5	D4	D3	D2	D1	D0
Name	SR			SPD	LPD	HPD	MPD	CPD
Type	R/W			R/W	R/W	R/W	R/W	R/W

Reset settings = 0000_0000

Bit	Name	Function
7	SR	Software Reset. 1 = Sets all registers to their reset value. 0 = Enables chip for normal operation. Note: Bit will automatically clear after being set.
6:5	Reserved	Read returns zero.
4	SPD	Speaker Drive Power Down. 1 = Normal operation 0 = Power down left and right speaker drive.
3	LPD	Line Drive Power Down. 1 = Normal operation 0 = Power down line driver.
2	HPD	Handset Drive Power Down. 1 = Normal operation 0 = Power down handset driver.
1	MPD	MIC Bias Power Down. 1 = Power down MIC bias buffer. 0 = Normal operation
0	CPD	Chip Power Down. 1 = Puts Si3000 into power down mode. 0 = Normal operation

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Register 2. Control 2

Bit	D7	D6	D5	D4	D3	D2	D1	D0
Name				HPFD	PLL	DL1	DL2	
Type	R/W R/W R/W R/W							

Reset Settings = 0000_0000

Bit	Name	Function
7:5	Reserved	Read returns zero.
4	HPFD	High Pass Filter (HPF) Disable. 1 = HPF disabled 0 = HPF enabled
3	PLL	PLL Divide by 10. 1 = Sets final stage of PLL to divide by 10. 0 = Sets final stage of PLL to divide by 5.
2	DL1	Digital Loopback. 1 = Enables digital loopback (DAC analog out → ADC analog in). 0 = Normal operation
1	DL2	Digital Loopback. 1 = Enables digital loopback (DAC one bit → ADC one bit). 0 = Normal operation
0	Reserved	Read returns zero.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Register 3. PLL1 Divide N1

Bit	D7	D6	D5	D4	D3	D2	D1	D0
Name	Divider N1							
Type	R/W							

Reset settings = 0000_0000

Bit	Name	Function
7:0	N1	N1. Contains the (value – 1) for determining the output frequency on PLL.

Register 4. PLL1 Multiply M1

Bit	D7	D6	D5	D4	D3	D2	D1	D0
Name	Multiplier M1							
Type	R/W							

Reset settings = 0000_0000

Bit	Name	Function
7:0	M1	M1. Contains the (value – 1) for determining the output frequency on PLL.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Register 5. RX Gain Control 1

Bit	D7	D6	D5	D4	D3	D2	D1	D0
Name	LIG		LIM	MCG		MCM	HIM	IIR
Type	R/W		R/W	R/W		R/W	R/W	R/W

Reset settings = 0100_0111

Bit	Name	Function
7:6	LIG	Line in Gain. 11 = 20 dB gain 10 = 10 dB gain 01 = 0 dB gain 00 = Reserved
5	LIM	Line in Mute. 1 = Line input muted 0 = Line input goes to mixer
4:3	MCG	MIC Input Gain. 11 = 30 dB gain 10 = 20 dB gain 01 = 10 dB gain 00 = 0 dB gain
2	MCM	MIC Input Mute. 1 = Mute MIC input 0 = MIC input goes into mixer.
1	HIM	Handset Input Mute. 1 = Mute handset input 0 = Handset input goes into mixer.
0	IIR	IIR Enable. 1 = Enables IIR filter 0 = Enables FIR filter

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Register 6. ADC Volume Control

Bit	D7	D6	D5	D4	D3	D2	D1	D0	
Name		RXG					LOM	HOM	
Type		R/W					R/W	R/W	

Reset settings = 0101_1100

Bit	Name	Function
7	Reserved	Read returns zero.
6:2	RXG	RX PGA Gain Control. 11111 = 12 dB 10111 = 0 dB 00000 = -34.5 dB LSB = 1.5 dB
1	LOM	Line Out Mute. 0 = Mute 1 = Active
0	HOM	Handset Out Mute. 0 = Mute 1 = Active

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Register 7. DAC Volume Control

Bit	D7	D6	D5	D4	D3	D2	D1	D0
Name		TXG					SLM	SRM
Type	R/W					R/W	R/W	

Reset settings = 0101_1100

Bit	Name	Function
7	Reserved	Read returns zero.
6:2	TXG	TX PGA Gain Control. 11111 = 12 dB 10111 = 0 dB 00000 = -34.5 dB LSB = 1.5 dB
1	SLM	SPKR_L Mute. 0 = Mute 1 = Active
0	SRM	SPKR_R Mute. 0 = Mute 1 = Active

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Register 8. Status Report

Bit	D7	D6	D5	D4	D3	D2	D1	D0
Name	SLSC	SRSC	LOSC					
Type	R	R	R					

Reset settings = 0000_0000

Bit	Name	Function
7	SLSC	SPK_L Short Circuit. 1 = Indicate short circuit status is detected at left speaker. 0 = Normal mode
6	SRSC	SPK_R Short Circuit. 1 = Indicate short circuit status is detected at right speaker. 0 = Normal mode
5	LOSC	Line Out Short Circuit. 1 = Indicate short circuit status is detected at line out. 0 = Normal mode
4:0	Reserved	Read returns zero.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Register 9. Analog Attenuation

Bit	D7	D6	D5	D4	D3	D2	D1	D0
Name					LOT		SOT	
Type					R/W		R/W	

Reset settings = 0000_0000

Bit	Name	Type
7:4	Reserved	Read returns zero.
3:2	LOT	Line Out Attenuation. 11 = -18 dB analog attenuation on Line Output. 10 = -12 dB analog attenuation on Line Output. 01 = -6 dB analog attenuation on Line Output. 00 = 0 dB analog attenuation on Line Output.
2:0	SOT	Speaker Out Attenuation. 11 = -18 dB analog attenuation on Speaker Output. 10 = -12 dB analog attenuation on Speaker Output. 01 = -6 dB analog attenuation on Speaker Output. 00 = 0 dB analog attenuation on Speaker Output.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้