

การบีบอัดข้อมูลเสียง
SPEECH COMPRESSION



โดย
นางสาวเจ จันท์สุภฤกษ์
นางสาวธีรดา จันท์สมิ



4

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
สาขาวิชาวิศวกรรมโทรคมนาคม
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2543

๖๖๖
๖๖๖๖
๖๖๖

เลขหน้.....
เลขทะเบียน 42136
วัน, เดือน, ปี 1 3 พ.ค. 2545

b.....
i.....

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ใดๆ
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การบีบอัดข้อมูลเสียง
SPEECH COMPRESSION

โดย

นางสาวเจ จันทร์ศุภฤกษ์ 40010143

นางสาวธีรดา จันทร์ศมี 40010324

อาจารย์ที่ปรึกษา

ดร.สุทธิชัย นพนาถิพงษ์

อาจารย์กฤษฎณ์ วงรุจิระ

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
สาขาวิชาวิศวกรรมโทรคมนาคม
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2543

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญาโทปีการศึกษา 2543

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การบีบอัดข้อมูลเสียง

SPEECH COMPRESSION

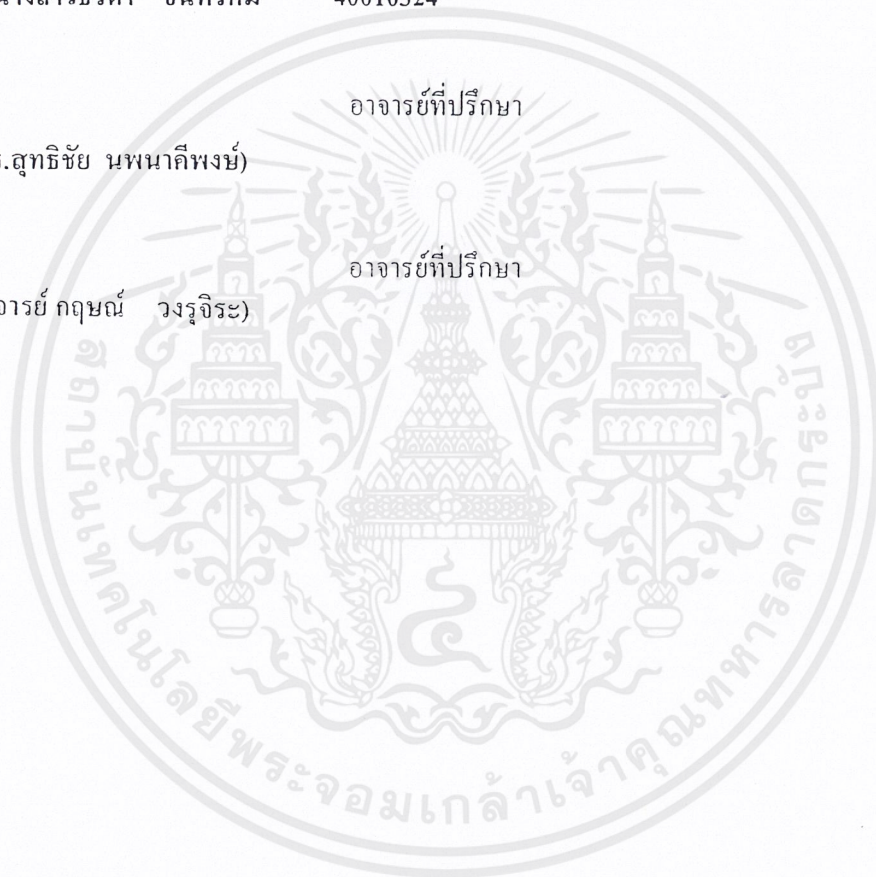
ผู้จัดทำ

1. นางสาวเจ จันทร์ศุภฤกษ์ 40010143

2. นางสาวธีรดา จันทร์ศมี 40010324

อาจารย์ที่ปรึกษา
(ดร.สุทธิชัย นพนาถิพงษ์)

อาจารย์ที่ปรึกษา
(อาจารย์ กฤษณ์ วงจริระ)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การบีบอัดข้อมูลเสียง

SPEECH COMPRESSION

โดย นางสาวเจ จันทรสุภฤกษ์ 40010143

นางสาวธีรดา จันทรศมี 40010324

อาจารย์ที่ปรึกษา ดร. สุทธิชัย นพนาถิพงษ์

อาจารย์ กฤษณ์ วงจรจิระ

บทคัดย่อ

โครงการชิ้นนี้เป็นการศึกษาการบีบอัดข้อมูลเสียง ซึ่งในปัจจุบันนี้การติดต่อสื่อสารโทรคมนาคมมีความก้าวหน้าไปอย่างมาก โดยเฉพาะการสื่อสารผ่านเครือข่ายคอมพิวเตอร์ ขนาดของข้อมูลเริ่มเข้ามามีบทบาทในการสื่อสารในปัจจุบันมากยิ่งขึ้น จึงต้องมีการพัฒนาเทคโนโลยีในการสื่อสารซึ่งทำให้เกิดประโยชน์สูงสุดในการส่งข้อมูลต่างๆ โดยพัฒนาข้อมูลข่าวสารต่างๆ ให้มีขนาดเล็กลง โดยยังคงคุณภาพของข้อมูลไว้อย่างครบถ้วน

ABSTRACT

The progression of Communication Technology especially the communication through the computer network, transfer signal (voice and data), has an important role in the world today. The smaller size of signal became more important in communication. So, this project is the study of Speech Compression to reduce the size of voice data file for supporting the development of nowadays communication. This compression could make the most benefit in transferring voice data file, and almost maintain the quality of original data.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

| | หน้า |
|---|------|
| บทที่ 1 บทนำ | 1 |
| บทที่ 2 ทฤษฎีและหลักการ | 2 |
| 2.1 ลักษณะของเสียงพูด | 2 |
| 2.2 อวัยวะที่ใช้ในการออกเสียง | 2 |
| 2.3 การเกิดของเสียง | 4 |
| 2.4 ทางเดินเสียง | 4 |
| 2.5 การเข้ารหัสเสียงเกี่ยวข้องกับการสุ่มข้อมูลและการควอนไทซ์ขนาด | 4 |
| 2.6 การเข้ารหัสเสียงแบบต่างๆ | 5 |
| 2.6.1 พัลส์โค้ดมอดูเลชัน | 5 |
| 2.6.2 แอลพีซี10 | 6 |
| 2.6.3 โค้ดเอ็กซ์ไซต์ลิเนียร์พรีดิกทีฟ | 6 |
| 2.6.4 โลว์ดีเลย์ซีอีแอลพี | 7 |
| 2.6.5 เวกเตอร์ซิมเอกซ์ไซต์ลิเนียร์พรีดิกทีฟ | 8 |
| 2.7 การกระตุ้นเสียงและการทำสเปกตรัลให้เรียบ | 8 |
| 2.8 การเข้ารหัสเสียงแบบช่องสัญญาณการกระตุ้นเสียง | 10 |
| 2.9 การกระตุ้นเสียงและการกระตุ้นค่าผิดพลาดของการเข้ารหัสเสียงแบบแอลพีซี | 11 |
| 2.10 การเข้ารหัสเสียงโดยวิธีการทำนาย | 13 |
| 2.11 การเข้ารหัสเสียงแบบการทำนายอะแด็ปทีฟ | 15 |
| 2.12 การเข้ารหัสแบบค้อย | 17 |
| 2.13 การเข้ารหัสเสียงแบบมัลติพัลส์แอลพีซี | 18 |
| 2.14 การบีบข้อมูลเสียงพูดด้วยวิธีการทำนายเชิงเส้นด้วยการกระตุ้นด้วยรหัส | 20 |
| 2.14.1 การเปลี่ยนแปลงให้เป็นซีอีแอลพี | 22 |
| 2.14.2 อนุกรมของโค้ดบุ๊กที่ไม่เป็นเกาส์เซียน | 22 |
| 2.14.3 โลว์ดีเลย์ซีอีแอลพี | 23 |
| 2.15 การลดเวลาการค้นหาในการเข้ารหัสแบบซีอีแอลพี | 24 |
| 2.15.1 ฟิวเจอร์ซิมพลิฟิเคชัน | 25 |
| 2.15.2 การเพิ่มความเร็วการค้นหา | 26 |
| 2.15.3 การค้นหาโค้ดบุ๊กแบบมัลติเรโซลูชัน | 28 |
| 2.15.4 การทำลายอนุกรมพาเซี่ยล | 29 |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

| | หน้า |
|--|-----------|
| 2.15.5 ไม้ค้ำยันแบบโครงเหล็ก | 29 |
| 2.15.6 อะแดปทีฟไม้ค้ำยัน | 30 |
| 2.15.7 ไม้ค้ำยันแบบลิเนียร์คอมบิเนชัน | 31 |
| 2.16 การวิเคราะห์สัญญาณเสียงในช่วงเวลาสั้นๆ | 31 |
| 2.17 รูปแบบของช่องแคบ | 32 |
| 2.18 การหาค่าพลังงานของสัญญาณ | 33 |
| 2.19 การวิเคราะห์เสียงเบื้องต้น | 33 |
| 2.20 การเตรียมสัญญาณเสียงในการวิเคราะห์ มีขั้นตอนดังต่อไปนี้ | 33 |
| 2.20.1 การพรีเอมฟาสีส | 33 |
| 2.20.2 การแบ่งช่วงสัญญาณ | 34 |
| 2.21 การวิเคราะห์หาคุณลักษณะเสียง | 35 |
| 2.22 การคำนวณหาออคโตคอร์รีเลชัน | 36 |
| 2.23 คุณสมบัติของฟังก์ชันออคโตคอร์รีเลชัน | 36 |
| 2.24 การหาพารามิเตอร์แอลพีซี | 39 |
| 2.25 การเปลี่ยนพารามิเตอร์แอลพีซี เป็นสัมประสิทธิ์เชปสตรัม | 40 |
| 2.26 การหาค่าพารามิเตอร์ | 40 |
| 2.27 การจัดระดับแวกเตอร์ | 40 |
| บทที่ 3 การคำนวณและการสร้าง | 44 |
| 3.1 หลักการของซีอีแอลพี | 45 |
| 3.2 สัญญาณอินพุตและสัญญาณเอาต์พุต | 46 |
| 3.3 การวิเคราะห์แอลเอสพี | 46 |
| 3.3.1 การวิเคราะห์แอลพีซี | 47 |
| 3.3.2 การแปลงพารามิเตอร์แอลพีซีเป็นพารามิเตอร์แอลเอสพี | 47 |
| 3.3.3 การทำให้แอลเอสพีมีค่าคงที่ | 50 |
| 3.4 การหาค่าพารามิเตอร์ของ ไม้ค้ำยัน | 50 |
| 3.4.1 ปัญหาในการค้นหา | 50 |
| 3.4.2 การแบ่งเฟรมเป็นเฟรมย่อย | 51 |
| 3.4.3 การรวมฟิลเตอร์เพอร์เซปชวล และสเปคตรัม | 52 |
| 3.4.4 การค้นหาแบบอนุกรม | 54 |
| 3.4.5 การค้นหาพร้อมกันอย่างมีประสิทธิภาพ | 55 |
| 3.4.6 วิธีการที่รวดเร็วสำหรับ ไม้ค้ำยันแบบพิเศษ | 56 |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

| | หน้า |
|---|------|
| 3.5 สังเคราะห์เสียง | 58 |
| 3.5.1 กระบวนการการสังเคราะห์ | 58 |
| 3.5.2 การแปลงแอลเอสพีเป็นแอลพีซี | 58 |
| บทที่ 4 การทดลองและผลการทดลอง | |
| การทดลองที่ 4.1 การทดสอบโปรแกรมการติดต่อกับผู้ใช้และส่วนแสดงผล การทำงานในส่วน ต่างๆ | 61 |
| การทดลองที่ 4.2 ตรวจสอบผลการบีบอัดโดยการดูรูปของสัญญาณที่ถอดรหัสได้ เปรียบเทียบกับสัญญาณอินพุตที่รับเข้ามาทั้งในแกนเวลา และแกนความถี่ | 66 |
| การทดลองที่ 4.3 ตารางแสดงค่าขนาดของข้อมูลเมื่อผ่านการเข้ารหัสแบบต่างๆ | 74 |
| การทดลองที่ 4.4 ตรวจสอบค่าสัมประสิทธิ์ไลน์สเปคตรัมแพร์ (แอลเอสพี) | 78 |
| การทดลองที่ 4.5 ตรวจสอบค่าสัมประสิทธิ์ไลน์สเปคตรัมแพร์ (แอลเอสพี) หลังผ่านการ เข้ารหัสแล้ว | 82 |
| การทดลองที่ 4.6 ตรวจสอบค่าความผิดพลาด | 90 |
| การทดลองที่ 4.7 ตารางแสดงค่าขนาดของข้อมูลเมื่อผ่านการเข้ารหัสแบบต่างๆ | 93 |
| บทที่ 5 บทวิจารณ์และบทสรุป | 95 |
| 5.1 สรุปผลการทดลอง | 95 |
| 5.2 บทวิจารณ์และการพัฒนา | 95 |
| ภาคผนวก | |
| เอกสารอ้างอิง | |

สารบัญรูปภาพ

| | หน้า |
|--|------|
| รูปที่ 2.1 ตัวเข้ารหัสของซีอีแอลพี แบบปรกติ | 7 |
| รูปที่ 2.2 ตัวถอดรหัสของวีเอสอีแอลพี | 8 |
| รูปที่ 2.3 การทำสเปคตรัมให้ราบเรียบของสัญญาณเบสแบนด์เพื่อทำการผลิตสัญญาณ กระตุ้นตัวเข้ารหัสเสียง | 9 |
| รูปที่ 2.4 วงจรซิกแซก | 9 |
| รูปที่ 2.5 ขั้นตอนการทำงานของการทำสเปคตรัมให้ราบเรียบ | 10 |
| รูปที่ 2.6 การเข้ารหัสเสียงโดยการกระตุ้นและแสดงตัวทำการวิเคราะห์ (ส่วนบน) และ ส่วนสังเคราะห์ (ด้านล่าง) | 10 |
| รูปที่ 2.7 แสดงการเลือกวิธีของการทำสเปคตรัมให้ราบเรียบ โดยการใช้ คอแวนซ์แชมปลิ่ง | 12 |
| รูปที่ 2.8 แสดงบล็อกโคอะแกรมอย่างง่ายของวีอีแอลพี และอาร์อีแอลพี | 12 |
| รูปที่ 2.9 ดิฟเฟอเรนเชียลพัลส์โค้ดมอดูเลชัน | 14 |
| รูปที่ 2.10 อะแดปทีฟดิฟเฟอเรนเชียลพัลส์โค้ดมอดูเลชัน | 14 |
| รูปที่ 2.11 การมอดูเลชันเคลดต้าต่อเนื่องความชัน | 15 |
| รูปที่ 2.12 แนวความคิดแบบ การประมาณเชิงเส้นลูคิเมนเทลี | 16 |
| รูปที่ 2.13 ฟรีดิกเตอร์และควอนไทเซอร์ในรูป | 16 |
| รูปที่ 2.14 แสดงแนวความคิดอะแดปทีฟฟรีดิกทีฟโค้ดดิ้ง (adaptive predictive coding :APC) | 17 |
| รูปที่ 2.15 แสดงผลการทดลองเพื่อที่จะหาค่าความเหมาะสมของระดับของสัญญาณรบกวน (noise) ที่อยู่ใน คริติคัลแบนด์เราจะพบว่าสัญญาณรบกวน ควรมีค่าต่ำกว่าสัญญาณอยู่ 18 – 28 dB ก่อนที่จะ ผ่านการมาสค์ (mask) | 18 |
| รูปที่ 2.16 แสดง QMF แบบ 4 ช่องสัญญาณ (channel) | 18 |
| รูปที่ 2.17 แอลพีซีหลายรูปคลื่น | 19 |
| รูปที่ 2.18 ตัวสังเคราะห์ซีอีแอลพี | 20 |
| รูปที่ 2.19 ตัวเข้ารหัสสัญญาณซีอีแอลพี พื้นฐาน | 21 |
| รูปที่ 2.20 การเข้ารหัสโลว์ดีเลย์ซีอีแอลพี | 23 |
| รูปที่ 2.21 ตัวเข้ารหัสแบบซีอีแอลพี หลังจากทำการลดระดับครั้งแรก | 25 |
| รูปที่ 2.22 อธิบายรูปที่ 2.21 | 25 |
| รูปที่ 2.23 ตัวเข้ารหัสแบบซีอีแอลพี ที่มีการนำหน่วยความจำออก | 26 |
| รูปที่ 2.24 โครงสร้างต้นไม้ของเคลดต้าบิท | 29 |
| รูปที่ 2.25 แสดงขั้นตอนการเตรียมสัญญาณในการวิเคราะห์ | 33 |
| รูปที่ 2.26 วงจรกรองความถี่สูงผ่าน | 34 |

สารบัญรูปภาพ (ต่อ)

| | หน้า |
|--|------|
| รูปที่ 2.27 การแบ่งช่วงของสัญญาณ | 34 |
| รูปที่ 2.28 แสดงวินโดว์แบบแฮมมิง | 35 |
| รูปที่ 2.29 แสดงขั้นตอนในการวิเคราะห์หาคุณลักษณะของเสียง | 36 |
| รูปที่ 2.30 บล็อกไดอะแกรมแสดงโมเดลการสร้างสัญญาณเสียงพูดอย่างง่าย | 37 |
| รูปที่ 2.31 แสดงการกระจายเฟรมของเสียงพูดแต่ละจุดแทนเฟรมของเสียง | 41 |
| รูปที่ 2.32 การรวมกลุ่มของเฟรมเสียงเพื่อสร้างโค้ด X | 41 |
| รูปที่ 2.33 บล็อกไดอะแกรมของเวกเตอร์ควอนไทซ์ | 41 |
| รูปที่ 2.34 ขั้นตอนของเวกเตอร์ควอนไทเซชัน | 42 |
| รูปที่ 3.1 บล็อกไดอะแกรมรวมของการเข้ารหัส | 44 |
| รูปที่ 3.2 บล็อกไดอะแกรมการวิเคราะห์ซีอีแอลพี | 46 |
| รูปที่ 3.3 อุปกรณ์วิเคราะห์ซีอีแอลพี | 54 |
| รูปที่ 4.1 แสดงหน้าจอการเริ่มต้น การทำงาน โดยให้เลือกเข้ารหัสหรือถอดรหัส | 61 |
| รูปที่ 4.2 แสดงหน้าจอการทำงานถ้าเลือกการเข้ารหัส โดยจะให้เลือกเสียงที่ต้องการเข้ารหัส | 62 |
| รูปที่ 4.3 แสดงหน้าจอการทำงานถ้าเลือกการเข้ารหัส โดยจะให้พิมพ์ชื่อไฟล์ที่ต้องการเก็บเสียงนั้น | 62 |
| รูปที่ 4.4 แสดงหน้าจอขณะกำลังเริ่มทำการทำงาน | 63 |
| รูปที่ 4.5 แสดงหน้าจอขณะกำลังทำการทำงาน โดยจะมีการแสดงสถานะการทำงานในแต่ละเฟรม | 63 |
| รูปที่ 4.6 แสดงหน้าจอเสร็จสิ้นการเข้ารหัส โดยจะให้เลือกทำการทำงานต่อหรือออกจากโปรแกรม | 64 |
| รูปที่ 4.7 แสดงหน้าจอการทำงานถ้าเลือกการถอดรหัส โดยจะให้เลือกไฟล์(*.celp)ที่ต้องการถอดรหัส | 64 |
| รูปที่ 4.8 แสดงหน้าจอการทำงานถ้าเลือกการถอดรหัส โดยจะให้เลือกพิมพ์ชื่อไฟล์ที่ต้องการเก็บเสียงนั้น | 65 |
| รูปที่ 4.9 แสดงหน้าจอเมื่อเสร็จสิ้นการถอดรหัส โดยจะให้เลือกทำการทดสอบเสียงด้วยการพล็อตหรือ ทดลองโดยการฟังเสียง หรือออกจากโปรแกรม | 65 |
| รูปที่ 4.10 รูปแสดงการเปรียบเทียบสัญญาณเอาท์พุทหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุท โดยแสดงในแกนเวลา ของสัญญาณทดสอบ (Five.wav) ขนาด 9 กิโลไบต์ | 66 |
| รูปที่ 4.11 รูปแสดงการเปรียบเทียบสัญญาณเอาท์พุทหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุท โดยแสดงในแกนความถี่ ของสัญญาณทดสอบ (Five.wav) ขนาด 9 กิโลไบต์ | 67 |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ (ต่อ)

หน้า

| | | |
|-------------|---|----|
| รูปที่ 4.12 | รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุตหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุต โดยแสดงในแกนเวลา ของสัญญาณทดสอบ (สวีสวี_อิน.wav) ขนาด 27 กิโลไบต์ | 68 |
| รูปที่ 4.13 | รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุตหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุต โดยแสดงในแกนความถี่ ของสัญญาณทดสอบ (สวีสวี_อิน.wav) ขนาด 27 กิโลไบต์ | 69 |
| รูปที่ 4.14 | รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุตหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุต โดยแสดงในแกนเวลา ของสัญญาณทดสอบ (โน้ต.wav) ขนาด 47 กิโลไบต์ | 70 |
| รูปที่ 4.15 | รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุตหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุต โดยแสดงในแกนความถี่ ของสัญญาณทดสอบ (โน้ต.wav) ขนาด 47 กิโลไบต์ | 71 |
| รูปที่ 4.16 | รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุตหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุต โดยแสดงในแกนเวลา ของสัญญาณทดสอบ (เจ.wav) ขนาด 49 กิโลไบต์ | 72 |
| รูปที่ 4.17 | รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุตหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุต โดยแสดงในแกนความถี่ ของสัญญาณทดสอบ (เจ.wav) ขนาด 49 กิโลไบต์ | 73 |
| รูปที่ 4.18 | แสดงค่าสัมประสิทธิ์การเข้ารหัสแบบประมาณเชิงเส้นของเสียงทดสอบ (Five.wav) ขนาด 9 กิโลไบต์ | 74 |
| รูปที่ 4.19 | แสดงค่าสัมประสิทธิ์การเข้ารหัสแบบประมาณเชิงเส้นของเสียงทดสอบ (สวีสวี_อิน.wav) ขนาด 26 กิโลไบต์ | 75 |
| รูปที่ 4.20 | แสดงค่าสัมประสิทธิ์การเข้ารหัสแบบประมาณเชิงเส้นของเสียงทดสอบ (โน้ต.wav) ขนาด 47 กิโลไบต์ | 76 |
| รูปที่ 4.21 | แสดงค่าสัมประสิทธิ์การเข้ารหัสแบบประมาณเชิงเส้นของเสียงทดสอบ (เจ.wav) ขนาด 49 กิโลไบต์ | 77 |
| รูปที่ 4.22 | แสดงค่าสัมประสิทธิ์แอลเอสพีของเสียงทดสอบ (Five.wav) ขนาด 9 กิโลไบต์ | 78 |
| รูปที่ 4.23 | แสดงค่าสัมประสิทธิ์แอลเอสพีของเสียงทดสอบ (สวีสวี_อิน.wav) ขนาด 26 กิโลไบต์ | 79 |
| รูปที่ 4.24 | แสดงค่าสัมประสิทธิ์แอลเอสพีของเสียงทดสอบ (โน้ต.wav) ขนาด 47 กิโลไบต์ | 80 |
| รูปที่ 4.25 | แสดงค่าสัมประสิทธิ์แอลเอสพีของเสียงทดสอบ (เจ.wav) ขนาด 49 กิโลไบต์ | 81 |
| รูปที่ 4.26 | แสดงค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (Five.wav) ขนาด 9 กิโลไบต์ | 82 |

สารบัญรูปภาพ (ต่อ)

หน้า

| | |
|---|----|
| รูปที่ 4.27 แสดงค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (สวัสดี_อื่น.wav) ขนาด 26 กิโลไบต์ | 83 |
| รูปที่ 4.28 แสดงค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (โน้ต.wav) ขนาด 47 กิโลไบต์ | 84 |
| รูปที่ 4.29 แสดงค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (เจ.wav) ขนาด 49 กิโลไบต์ | 85 |
| รูปที่ 4.30 แสดงการประมาณค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (Five.wav) ขนาด 9 กิโลไบต์ | 86 |
| รูปที่ 4.31 แสดงการประมาณค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (สวัสดี_อื่น.wav) ขนาด 26 กิโลไบต์ | 87 |
| รูปที่ 4.32 แสดงการประมาณค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (โน้ต.wav) ขนาด 47 กิโลไบต์ | 88 |
| รูปที่ 4.33 แสดงการประมาณค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (เจ.wav) ขนาด 49 กิโลไบต์ | 89 |
| รูปที่ 4.34 แสดงค่ามโนสแควร์เออเรอร์ (Mean-square error) ของสัญญาณเสียงทดสอบ (Five.wav) | 90 |
| รูปที่ 4.35 แสดงค่ามโนสแควร์เออเรอร์ (Mean-square error) ของสัญญาณเสียงทดสอบ (สวัสดี_อื่น.wav) | 91 |
| รูปที่ 4.36 แสดงค่ามโนสแควร์เออเรอร์ (Mean-square error) ของสัญญาณเสียงทดสอบ (โน้ต.wav) | 91 |
| รูปที่ 4.37 แสดงค่ามโนสแควร์เออเรอร์ (Mean-square error) ของสัญญาณเสียงทดสอบ (เจ.wav) | 92 |

สารบัญตาราง

| | หน้า |
|--|------|
| ตารางที่ 4.1 แสดงการเปรียบเทียบขนาดของข้อมูลเมื่อทำการเข้ารหัสวิธีต่าง ๆ | 93 |
| ตารางที่ 4.2 แสดงค่าอัตราการบีบอัดข้อมูลเมื่อทำการเข้ารหัสวิธีต่าง ๆ | 94 |



บทที่ 1

บทนำ

การสื่อสารในปัจจุบันนั้น ได้มีการพัฒนาไปอย่างรวดเร็ว มีการนำเอาเทคโนโลยีต่างๆมาใช้ เพื่อเพิ่มประสิทธิภาพให้กับการสื่อสาร ยกตัวอย่าง เช่น การสื่อสารแบบเคลื่อนที่ในระบบดิจิทัลซึ่งมีข้อดีในด้านความทนทานต่อสัญญาณรบกวนซึ่งจะทำให้คุณภาพเสียงที่ดีขึ้น และในด้านของการใช้งานการติดต่อสื่อสารผ่านอินเทอร์เน็ต ซึ่งในการสื่อสารเหล่านี้ มักจะพบปัญหาคือในเรื่องของขนาดของข้อมูลเสียง ซึ่งโดยปกติจะมีช่วงความถี่อยู่ที่ 200-3400 เฮิรตซ์ ใช้อัตราการแปลงข้อมูล 8000 ตัวอย่างต่อวินาที รหัสที่ใช้สำหรับแทนตัวอย่างแต่ละตัวอย่างมีขนาด 8 บิต จึงทำให้ได้อัตราของข้อมูลเป็นอัตรา 64 กิโลบิตต่อวินาที ถ้าต้องการทำการเก็บเสียงไว้ในช่วงเวลาหนึ่งนาที่ที่จะต้องมีการใช้หน่วยความจำถึง 480 กิโลไบต์ และยังถ้าต้องการคุณภาพเสียงที่ดียิ่งขึ้น ก็จะต้องมีการใช้จำนวนบิตต่อตัวอย่างของเสียง และอัตราการแปลงข้อมูลเสียงที่สูงขึ้น ไปอีก จะทำให้ปริมาณของข้อมูลผลลัพธ์ที่ได้สูงมาก จึงจำเป็นที่จะต้องมีการควบคุมขนาดของข้อมูลอย่างมีประสิทธิภาพ เพื่อให้เกิดประโยชน์สูงสุดในการเก็บบันทึกข้อมูล และประโยชน์ในการใช้ช่องสัญญาณการสื่อสารได้อย่างคุ้มค่าที่สุด และทำให้สามารถส่งข้อมูลได้จำนวนมากขึ้นในสายนำสัญญาณที่มีแบนด์วิธจำกัด ดังนั้นเนื่องจากการที่มีความต้องการการสื่อสารที่เพิ่มขึ้นอย่างรวดเร็ว จึงได้มีการศึกษาในเรื่องของเสียงอย่างแพร่หลาย อาทิเช่น คุณลักษณะต่างๆของเสียง ความถี่ และแอมพลิจูด เพื่อนำไปพัฒนาขั้นตอนในการบีบอัดข้อมูล และใช้ในการส่งข้อมูลเสียงต่อไป

ในปริญญานิพนธ์นี้ ได้มีการศึกษา และทดลองการบีบอัดข้อมูลเสียง (Speech compression) ซึ่งเป็นการเพิ่มประสิทธิภาพให้กับการส่งข้อมูลเสียง โดยมีข้อจำกัดที่ว่า ข้อมูลที่ผ่านการบีบอัด เมื่อถึงเครื่องรับแล้วจะได้ข้อมูลที่มีความผิดพลาดอยู่ในระดับที่ยอมรับได้ของสัญญาณเสียงพูด โดยจะศึกษาในการเข้ารหัสสัญญาณ ด้วยวิธีการต่างๆ ซึ่งเราสามารถพัฒนาวิธีการเข้ารหัสสัญญาณ เพื่อให้ได้ข้อมูลที่มีขนาดเล็ก แต่มีประสิทธิภาพในการเก็บรายละเอียดของข้อมูลที่สำคัญไว้ โดยในปริญญานิพนธ์นี้ได้เลือกใช้วิธีการบีบอัดแบบการกระตุ้นเชิงเส้นกับโค้ดบุ๊ก (Code Excited Linear Prediction : CELP) ในการทำการบีบอัดข้อมูลเสียง ซึ่งเป็นวิธีการที่สามารถบีบอัดข้อมูลได้ค่อนข้างมาก และมีความผิดพลาดที่ยอมรับในระดับของสัญญาณเสียงพูด ซึ่งเป็นมาตรฐานที่พัฒนาขึ้นมาโดยใช้หลักการเข้ารหัสแบบการเข้ารหัสเสียง (Vocoder) หาสัญญาณการประมาณเชิงเส้น และทำการกระตุ้นสัญญาณการประมาณเชิงเส้นที่ได้กับโค้ดบุ๊กที่กำหนดไว้และเก็บค่าเป็นดัชนีของโค้ดบุ๊กนั้น

บทที่ 2

ทฤษฎีหรือหลักการ

ในปัจจุบันนี้การสื่อสารแบบใช้สายโดยใช้เส้นใยแก้วนำแสง (optical fiber) ที่มีขนาดของแถบความถี่สูงเริ่มมีราคาลดลงและมีการใช้งานกันมากขึ้น แต่อย่างไรก็ตามในการสื่อสารแบบไร้สายเช่นสายโทรศัพท์เคลื่อนที่แบบเซลลูลาร์ (cellular) หรือการสื่อสารดาวเทียมก็ยังคงมีความจำเป็นในการส่งวนการใช้แถบความถี่อยู่ สัญญาณเสียงพูดที่ใช้กันในปัจจุบันส่วนใหญ่จะอยู่ในรูปแบบของสัญญาณดิจิทัล เพื่อให้ได้ทำการประมวลผล เก็บรักษาหรือทำการส่งได้โดยใช้โปรแกรมคอมพิวเตอร์ การเก็บข้อมูลแบบดิจิทัลทำให้เกิดอัตราของข้อมูลที่สูง ซึ่งต้องการแถบความถี่ในการส่งหรือเก็บที่สูงเช่นกัน การเข้ารหัสเสียงพูด (Speech coding) หรือการบีบอัดเสียงพูด (Speech Compression) คือวิธีการหารูปแบบของข้อมูลดิจิทัลที่กะทัดรัดเพื่อใช้แสดงแทนสัญญาณเสียงพูดเพื่อนำวัตถุประสงค์ในการส่งและเก็บอย่างมีประสิทธิภาพ หลักการสำคัญคือการนำเสนอเสียงพูดโดยใช้จำนวนบิตที่น้อยที่สุดในขณะที่ยังคงรักษาคุณภาพการรับฟังของเสียงนั้นเอาไว้ได้

2.1 ลักษณะของเสียงพูด

คนเราเปล่งเสียงพูดด้วยอวัยวะที่ใช้ในการออกเสียง ทำเสียงตามที่มิในระบบภาษาของตน แม้ว่าคนที่อยู่ในสังคมเดียวกันจะใช้ภาษาเดียวกันแต่ถ้าพิจารณาเสียงที่เปล่งออกมาแล้วจริงๆ แล้วแต่ละครั้ง ก็อาจจะสังเกตเห็นลักษณะที่ต่างกันได้ เราจึงสามารถจำเสียง จำวิธีพูดของคนที่เราคุ้นเคยได้ เสียงพูดที่จะอธิบายด้วยหลักเกณฑ์ทางวิทยาศาสตร์แม้ว่าในภาษาหนึ่งๆ จะมีเสียงต่างกันมากบ้างน้อยบ้าง แต่ละเสียงก็สามารถนำมาพิจารณาและอธิบายให้รู้ลักษณะการออกเสียงและตำแหน่งที่เกิดเสียงได้ ถ้าอธิบายนี้จะทำให้เข้าใจลักษณะเสียงทุกเสียง วิชาที่ว่าด้วยเสียงพูดเรียกว่า วิชาสัทศาสตร์ (Phonetics)

ในการศึกษาเสียงพูดแบ่งได้เป็น 2 ลักษณะ คือ

สรีรศาสตร์ (Articulatory) เป็นการศึกษาเสียงพูดจากอวัยวะและการเคลื่อนไหวอวัยวะที่ทำให้เกิดเสียงพูด การอธิบายนี้ก็จะอธิบายโดยอาศัยลักษณะและอาการเคลื่อนไหวของอวัยวะที่เกี่ยวข้องในการเปล่งเสียงพูดนั้น

กลศาสตร์ (Acoustic Phonetics) เป็นการศึกษาเสียงพูดจากลักษณะคลื่นเสียงที่ผู้พูดเปล่งออกมาแล้ว และผู้ฟังได้ยินว่ามีลักษณะทางกลศาสตร์อย่างไร การศึกษาตามแนวนี้ต้องอาศัยความรู้ทางฟิสิกส์และคณิตศาสตร์ ช่วยอธิบายลักษณะของการเกิดคลื่นเสียง

2.2 อวัยวะที่ใช้ในการออกเสียง

อวัยวะที่ใช้ในการออกเสียงมีอยู่หลายส่วน แต่ละส่วนสามารถทำให้เสียงพูดแตกต่างกันไปได้ อวัยวะเหล่านี้ มีปากและส่วนต่างๆ ในปาก ช่องคอ กล่องเสียง ช่องว่างในปาก และช่องว่างในจมูก อวัยวะที่ใช้ในการเปล่งเสียงพูด แบ่งออกเป็น 2 ประเภท คือ

อวัยวะที่ใช้ในการทำอาหาร(Articulator) คืออวัยวะที่เคลื่อนไหว เพื่อผลิตลมไปยังส่วนต่างๆ อวัยวะที่สำคัญคือ ลิ้น ซึ่งเป็นส่วนที่เคลื่อนไหวมากที่สุด

อวัยวะซึ่งเป็นตำแหน่งซึ่งเกิดเสียงต่างๆ(Point Of Articulator) คือ ตำแหน่งที่เกิดเสียงต่างๆ เช่น ริมฝีปาก ฟัน เพดาน ส่วนต่างๆ เป็นต้น

อวัยวะส่วนที่มีหน้าที่ในการออกเสียง โดยตรงมี ดังนี้

(1) ริมฝีปาก เป็นอวัยวะที่เคลื่อนไหวได้มาก และทำให้เสียงแตกต่างกันได้มาก เราอาจบังคับให้ริมฝีปากอยู่ชิดติดกัน ห่างกัน ขึ้นออกมา เป็นต้นก็ได้ ลักษณะริมฝีปากแบบต่างๆล้วนมีอิทธิพลต่อการออกเสียงและทำให้เสียงแตกต่างกันไปทั้งสิ้น

(2) ฟัน เป็นอวัยวะที่ทำให้เกิดเสียงหลายชนิด เช่น เมื่อกดลงบนริมฝีปากล่าง ลมที่ผ่านออกมาโดยแรงจะลอดช่องออกมาซึ่งทำให้เสียงเกิดได้

(3) ปุ่มเหงือก เป็นส่วนนูนออกมา อยู่หลังฟันด้านบนและจะรู้สึกว่าเป็นคลื่นปุ่มเหงือก เป็นบริเวณที่ทำให้เกิดเสียงปุ่มเหงือก

(4) เพดานแข็งหรือเพดานอ่อน คือส่วนที่เป็นกระดูกแข็ง

(5) เพดานอ่อน คือส่วนที่ต่อเพดานแข็งไปข้างใน มีลักษณะเป็นกระดูกอ่อนที่ขยับได้ในเวลาที่หายใจ เพดานอ่อนและลิ้นไก่ ซึ่งอยู่ปลายเพดานอ่อนลดระดับลงมาเปิดช่องว่างให้ลมออกทางจมูก เวลาพูดส่วนใหญ่ ปลายเพดานอ่อนและลิ้นไก่อจะถูกยกขึ้นไปจรดหลังคอก้นอกจากเวลาออกเสียง นานิกเท่านั้น

(6) ลิ้นไก่ เป็นก้อนเนื้อเล็กๆอยู่ปลายเพดานอ่อนตกกลางปากลิ้นร้วได้

(7) ลิ้น ลิ้นเป็นส่วนที่เคลื่อนไหวมากที่สุดในการออกเสียง ซึ่งสามารถแบ่งได้ 3 ส่วน

-ปลายลิ้น คือส่วนของปลายลิ้นซึ่งสามารถยกขึ้นไปแตะกับส่วนต่างๆได้

-หน้าลิ้น คือส่วนที่อยู่ตรงข้ามกับเพดานแข็ง ถ้าวางลิ้นราบกับปาก เช่นเดียวกับเวลาที่ไม่ได้พูด

-หลังลิ้น ถ้าวางลิ้นราบกับปาก ส่วนนี้จะอยู่ตรงข้ามกับเพดานอ่อน

(8) แผ่นเนื้อปากหลอดลม เป็นเนื้อก้อนเล็กๆคล้ายลิ้นไก่อยู่ตรงโคนลิ้น ลงไปในคอ มี หน้าที่ปิดช่องลม ในขณะที่รับประทานอาหาร และเปิดช่องลมเมื่อพูด

(9) ช่องคอ อยู่ถัดจากช่องปากไปจนถึงเส้นเสียง

(10) เส้นเสียง เป็นอวัยวะที่สำคัญที่ทำให้เกิดเสียง เส้นเสียงมีลักษณะเป็นกล้ามเนื้อ แผ่น ภายในกล่องเสียงปิดขวางอยู่ปากช่องหลอดลม จากด้านหลังมาด้านหน้า ระหว่างเส้นเสียงจะมีช่องว่างซึ่งเป็นช่องผ่านไทม์ส ไปถึงปอด และออกมาจากปอดได้ ช่องนี้เรียกว่า ช่องว่างระหว่างเส้นเสียง

(11) ช่องจมูก หมายถึง โพรงในช่องจมูก ซึ่งอยู่เหนือลิ้นไก่ขึ้นไปเป็นช่องที่ลมซึ่งผ่าน เส้นเสียงขึ้นมาจะปานออกไปทางจมูกได้เมื่อเวลาหายใจและเวลาออกเสียงนาสิก

เสียงที่เกิดขึ้นนั้นไม่ว่าจะเป็นเสียงประเภทใดจะผ่านไปยังทางเดินของเสียง โดยจะเริ่มตั้งแต่ช่องว่างระหว่าง เส้นเสียงถึงริมฝีปาก ในทางวิทยาศาสตร์สรุปได้ว่า ทางเดินของเสียง คือท่อนำเสียงที่มีรูปร่างไม่แน่นอน

2.3 การเกิดของเสียง (Speech Production)

การเกิดของเสียงแบ่งออกเป็น 3 ขั้นตอน คือ

ขั้นตอนที่ 1 จุดเริ่มต้น เป็นขั้นตอนที่ลมเริ่มถูกขับออกจากปอด ผ่านเข้าไปสู่ขั้นตอนที่ 2

ขั้นตอนที่ 2 การดัดแปลงลมที่เส้นเสียงอวัยวะ ที่ใช้ในขั้นตอนนี้คือส่วนที่ต่อจากปอดขึ้นมาถึงกล่องเสียง และที่ กล่องเสียง เส้นเสียงจะทำหน้าที่เป็นลิ้นปิดเปิดทำให้เกิดเสียง 2 ชนิด คือ

1. เสียงก้อง (Voiced) เกิดจากเส้นเสียงปิดกั้นลมไว้ ลมที่ผ่านมาจะเพิ่มแรงดันมากขึ้นจนเส้นเสียงปิดเปิดสลับกันไป ทำให้เกิดเสียงก้องขึ้นมา ซึ่งสามารถเรียกความถี่ในการปิดเปิดเส้นเสียงว่า “ความถี่มูลฐาน” มีลักษณะเป็นรายคาบ(periodic)และมีค่าคาบของเสียง(pitch period)เป็นส่วนประกอบสำคัญ ตัวอย่างของเสียงก้องได้แก่ เสียงสระต่างๆ เสียงพยัญชนะ บ,ก ซึ่งเปล่งออกมาทางปาก เสียงพยัญชนะ น,ม,ง ซึ่งเปล่งออกมาทางจมูกที่เรียกว่าเสียงนาสิก (Nasal sound)

2. เสียงไม่ก้อง (Unvoiced) เสียงชนิดนี้เส้นเสียงจะเปิดตลอดเวลาที่ลมผ่าน ลมจึงผ่านออกมาได้สะดวกทำให้เกิดเสียงไม่ก้องขึ้น ไม่เป็นรายคาบ แต่จะมีลักษณะเป็นสัญญาณของเสียงรบกวน (noise) ตัวอย่างของเสียงไม่ก้องได้แก่ เสียงพยัญชนะ ฟ,ต,ซ

ขั้นตอนที่ 3 การเปลี่ยนแปลงลักษณะเส้นเสียง อวัยวะที่ใช้คือส่วนที่ต่อจากกล่องเสียงจนถึงริมฝีปาก โดยลมที่ผ่านออกจากกล่องเสียงจะทำให้เกิดเสียงในลักษณะต่างๆ ซึ่งจะเกิดจากการเปลี่ยนแปลงอวัยวะต่างๆ ที่กล่าวเอาไว้ในหัวข้อที่แล้ว

2.4 ทางเดินเสียง

ทางเดินเสียง คือ ช่องที่เสียงจะเดินทางผ่านจากช่องกำเนิดเสียงริมฝีปาก แสดงถึงส่วนประกอบของอวัยวะผลิตเสียง ส่วนของทางเดินเสียงคือช่องระหว่างลิ้นไก่ (valum) เพดานแข็ง (hard palate) กับลิ้น (tonque) โดยเสียงทั้งเสียงก้องและไม่ก้องก็ต้องผ่านช่องทางนี้ทั้งลิ้น ส่วนในกรณีของเสียงนาสิกเสียงจากต้นกำเนิดจะต้องผ่านทางโพรงจมูก (nasal cavity) ควบคู่ไปกับผ่านทางปาก

2.5 การเข้ารหัสเสียงเกี่ยวข้องกับการสุ่มข้อมูล (Sampling) และการควอนไทซ์ขนาด (Amplitude quantization)

ขนาดอัตราสุ่มข้อมูลของเสียงพูดมักจะคงที่ประมาณ 2 เท่าของขนาดแถบความถี่ของสัญญาณเสียงแบบอนาล็อก ดังนั้นการควอนไทซ์จะเป็นวิธีการสำคัญในการกำหนดรูปแบบการเข้ารหัสเสียงพูด ซึ่งแบ่งได้เป็น 2 แบบใหญ่ๆ คือ

1. การควอนไทซ์โดยตรงหรือแบบนอนพาราเมตริก คือ การนำสัญญาณไปนารีมาแสดงแทนสัญญาณเสียงโดยตรง

2. การควอนไทซ์แบบพาราเมตริก คือการนำสัญญาณไปนารีมาแสดงแทนโมเดลของเสียงหรือ/และพารามิเตอร์ทางความถี่ของเสียง

โดยวิธีการควอนไทซ์ที่สำคัญจะกล่าวถึงในหัวข้อถัดไปเสียงพูดโดยปกติของมนุษย์มีความถี่ไม่

เกิน 4 กิโลเฮิร์ตและจะถูกสุ่มข้อมูลที่อัตรา 8 กิโลเฮิร์ตซ์ การเข้ารหัสเสียงโดยการควอนไทซ์โดยตรงที่ง่ายเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่สุดคือ พัลส์โค้ดมอดูเลชัน(Pulse-Code Modulation : PCM) เสียงพูดที่เข้ารหัสที่อัตรา 64 กิโลบิตต่อวินาทีที่ใช้ลอคาลิทิมีพีซีเอ็ม (logarithm PCM) ถูกเรียกว่าการเข้ารหัสเสียงแบบไม่บีบอัดข้อมูล (noncompressed) ใช้เป็นการเข้ารหัสแบบอ้างอิงเทียบกับการเข้ารหัสแบบอื่นๆ ซึ่งอาจจะแบ่งการเข้ารหัสเสียงพูดเป็นพวกได้ตามอัตราข้อมูลได้ดังนี้

1. การเข้ารหัสอัตราข้อมูลสูง ใช้อัตราข้อมูลมากกว่า 16 กิโลบิตต่อวินาที
2. การเข้ารหัสอัตราข้อมูลขนาดกลาง ใช้อัตราข้อมูลอยู่ในช่วง 8-16 กิโลบิตต่อวินาที
3. การเข้ารหัสอัตราข้อมูลต่ำ ใช้อัตราข้อมูลอยู่ในช่วง 2.4-8 กิโลบิตต่อวินาที
4. การเข้ารหัสอัตราข้อมูลต่ำมาก ใช้อัตราข้อมูลอยู่ในช่วง 2.4 กิโลบิตต่อวินาที

การเข้ารหัสเสียงพูดที่อัตราตั้งแต่ขนาดกลางลงมานั้นต้องใช้ขบวนการวิเคราะห์-สังเคราะห์ (analysis-synthesis) ในขั้นของการวิเคราะห์จะหาชุดของพารามิเตอร์ที่ใช้แทนสัญญาณเสียงที่ถูกเข้ารหัสได้อย่างมีประสิทธิภาพ และในขั้นของการสังเคราะห์ค่าพารามิเตอร์เหล่านี้จะถูกถอดรหัสและสร้างเสียงพูดกลับมา การวิเคราะห์อาจจะเป็นได้ทั้งแบบวงปิด (closed loop) หรือแบบวงเปิด (open loop) ในแบบวงปิดค่าพารามิเตอร์จะถูกค้นหาจากค่าความแตกต่างระหว่างเสียงต้นฉบับกับเสียงที่ถูกสร้างขึ้นมา นั่นคือใน ส่วนวงปิดจะต้องมีส่วนสังเคราะห์อยู่ภายใน กระบวนการแบบหลังนี้จะเรียกว่าการวิเคราะห์จากการสังเคราะห์ (analysis-by-synthesis)

การเข้ารหัสเสียงพูดแบบพารามิเตอร์อาจเรียกได้อีกอย่างว่า “การเข้ารหัสเสียงตามลักษณะของเสียงพูดหรือโวลโคดเดอร์ (Speech-specific coders หรือ Voice coders หรือ Vocoders)” เป็นการเข้ารหัสที่เน้นในเรื่องของคุณภาพในการรับฟังของเสียงพูดโดยไม่จำเป็นต้องได้สัญญาณที่เหมือนเดิมทุกประการ โวลโคดเดอร์สามารถทำงานที่อัตราข้อมูลต่ำมากได้โดยให้คุณภาพของเสียงที่ระดับเสียงสังเคราะห์ โดยที่ในอัตราข้อมูลที่สูงขึ้นก็จะให้คุณภาพของเสียงที่ดีขึ้น

2.6 การเข้ารหัสเสียงแบบต่างๆ

การเข้ารหัสเสียงแบ่งได้เป็น 2 พวกใหญ่ๆ คือการเข้ารหัสเสียงตามรูปคลื่น (Waveform Coding) และการเข้ารหัสเสียงตามลักษณะของเสียงพูดหรือโวลโคดเดอร์ การเข้ารหัสเสียงตามรูปคลื่นที่สำคัญที่จะยกตัวอย่างในที่นี้ได้แก่ พีซีเอ็ม, ดีพีซีเอ็ม และ เอดีพีซีเอ็ม ส่วนโวลโคดเดอร์ที่จะพูดถึงได้แก่ แอลพีซี 10, ซีอีแอลพี, แอลดี ซีอีแอลพี และ วีเอสอีแอลพี ซึ่งจะพูดถึงหลักการทำงานโดยรวมดังนี้

2.6.1) พัลส์โค้ดมอดูเลชัน (Pulse Code Modulation : PCM) คือการเข้ารหัสเสียงแบบแรกๆ เป็นตัวอย่างที่สำคัญของการเข้ารหัสเสียงแบบ นอนพารามิเตอร์ เริ่มจากการสุ่มข้อมูลเสียงด้วยอัตรา 8 กิโลเฮิรตซ์ แล้วนำข้อมูลแต่ละตัวไปผ่านการควอนไทซ์เป็นข้อมูลดิจิทัลขนาด 8 บิตโดยใช้ตารางของการควอนไทซ์แบบลอคาลิทิมี ทำให้ได้เป็นนอนยูนิฟอร์ม (non-uniform PCM) ที่มีอยู่ 2 แบบคือ แอลอว์ (A-law) หรือ ไมโครลอว์ (μ -law) พีซีเอ็มที่อัตราข้อมูล 64 กิโลบิตต่อวินาที มีอัตราส่วนของสัญญาณต่อเสียงรบกวน (signal-to-noise ratio) ประมาณ 83 เดซิเบล เป็นการเข้ารหัสเสียงที่ใช้อ้างอิงเพื่อเปรียบเทียบ

กับการเข้ารหัสเสียงแบบอื่นๆ และได้เป็นมาตรฐานของการเข้ารหัสเสียง G.711 ที่กำหนดโดย CCITT ในปีทศวรรษที่ 1960

ต่อมามีการดัดแปลงข้อมูลที่จะถูกควอนไทซ์ โดยแทนที่จะใช้ตัวข้อมูลโดยตรงก็จะใช้ผลต่างของข้อมูลตัวปัจจุบันกับข้อมูลที่ได้ออกจากการทำนายพันระเชิงเส้น นั่นคือจะทำการควอนไทซ์

$$e(n) = s(n) - \hat{s}(n) \quad (2.1)$$

โดยที่ $s(n)$ คือข้อมูลเสียงตัวปัจจุบันและ

$$\hat{s}(n) = \sum_{i=1}^M \hat{a}(i)s(n-i) \quad (2.2)$$

คือเสียงที่ได้จากการทำนายพันระเชิงเส้นโดยที่ $\hat{a}(i)$ เป็นค่าที่มีเก็บไว้ทั้งในเครื่องเข้ารหัสและเครื่องถอดรหัสเป็นวิธีการที่ใช้ในดิฟเฟอเรนเชียลพีซีเอ็ม (Differential PCM : DPCM) และเมื่อมีการพัฒนาการปรับเปลี่ยนขนาดได้ทั้งค่าของขนาดขั้นในการควอนไทซ์และค่าของ $\hat{a}(i)$ จะได้เป็นอะแดพทีฟดีพีซีเอ็ม (Adaptive DPCM : ADPCM) ซึ่งมีการใช้อัตราเพียง 4 บิตต่อหนึ่งข้อมูล 32 กิโลบิตต่อวินาที เอดีพีซีเอ็ม มีคุณภาพของเสียงที่ดีมากและเป็นมาตรฐานของการเข้ารหัสเสียง G.721 ที่กำหนดโดย CCITT ในปี 1988

ต่อมาวิธีรูปแบบการเข้ารหัสเสียงแบบต่างๆ เกิดขึ้นมากมาย ซึ่งคงไม่สามารถกล่าวถึงในที่นี้ได้หมด ที่จะพูดต่อไปนี้เป็นเรื่องโวลโคเดอร์ซึ่งใช้หลักการกำเนิดเสียงที่เลียนแบบการกำเนิดเสียงพูดของมนุษย์ โดยเน้นที่เรื่องของโวลโคเดอร์แบบการทำนายพันระเชิงเส้น (Linear Predictive vocoder หรือ LPC) เนื่องจากเป็นเทคนิคที่มีการพัฒนาอย่างกว้างขวางที่สุดในช่วง 2 ทศวรรษหลังนี้

2.6.2) แอลพีซีสิบ (LPC-10) เป็นการเข้ารหัสรุ่นแรกที่ใช้เทคนิคของแอลพีซี โดยการส่งลำดับของตัวกระตุ้นจะใช้ตัวกระตุ้นแบบขบวนของพัลส์ที่มีช่วงห่างกันเท่ากับคาบของเสียง (Pitch-pulse excitation) สำหรับเสียงก้องและใช้สัญญาณรบกวนสำหรับเสียงไม่ก้อง ซึ่งเป็นการกำหนดค่าลำดับของตัวกระตุ้นที่พยายามทำให้คุณภาพของเสียงที่ได้อยู่ในระดับเสียงสังเคราะห์ มีการคำนวณค่าสัมประสิทธิ์ของการทำนายจำนวน 10 ตัวจึงมีชื่อว่า แอลพีซี 10

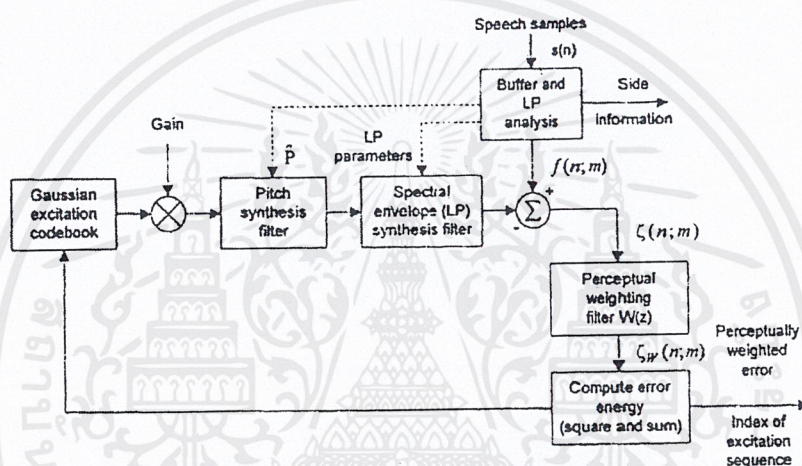
คุณภาพของเสียงที่ได้อยู่ในระดับเสียงสังเคราะห์ (synthetic quality) เริ่มใช้ในของเล่นเพื่อการศึกษาศึกษา (Speak and Spell) ของบริษัทเท็กซัสอินสตรูเมนต์ ในปี 1970s และเป็นมาตรฐานของการเข้ารหัสเสียง (Federal Standard FS1015) ในปี 1984 การส่งข้อมูลระหว่างตัวเข้ารหัสและตัวถอดรหัสของแอลพีซี-10 จะส่งค่าสัมประสิทธิ์ของการทำนายและลำดับของตัวกระตุ้นไป ถ้าเป็นเสียงก้องก็จะส่งเฉพาะความสูงของพัลส์และคาบของเสียง ถ้าเป็นเสียงไม่ก้องก็จะส่งเฉพาะค่าส่งไปให้ทางด้านรับกำเนิดสัญญาณรบกวนเอาเองทำให้จำนวนข้อมูลที่ส่งต่ำ ทำงานได้ที่อัตราข้อมูลเท่ากับ 2400 บิตต่อวินาที

2.6.3) โค้ดเอ็กซิตีลีนีร์พรีดิกทีฟ (Code-Excited Linear Prediction : CELP) คือการเข้ารหัส

เสียงแบบแอลพีซี ที่ใช้การเข้ารหัสเวกเตอร์ $e(n)$ ด้วยรหัสที่มีเก็บอยู่ในชุดรหัสจำนวนจำกัด โดยการเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทำงานจะทำกับสัญญาณเสียงที่เป็นเวกเตอร์ ให้คุณภาพของเสียงที่ดีโดยใช้อัตราข้อมูลที่ต่ำซีอีแอลพี เป็นเครื่องเข้ารหัสที่มีการวิเคราะห์โดยการวิเคราะห์หั่นคือที่ด้านเข้ารหัสจะมีการสังเคราะห์เสียงตัวปัจจุบัน ขึ้นมาจากข้อมูลที่มีอยู่ แล้วนำไปเปรียบเทียบกับเสียงจริงที่เข้ามา ข้อแตกต่างที่ได้จะเป็นสัญญาณความผิดพลาด (Error signal) ที่ผ่านการถ่วงคุณภาพของการรับฟังแล้ว ซึ่งจะถูกนำไปผ่านการควอนไทซ์แบบเวกเตอร์เพื่อหาดัชนีของลำดับของตัวกระตุ้นที่เหมาะสมที่สุดแล้วส่งดัชนีนั้นไปยังด้านรับ นอกจากนี้แล้ว ยังต้องมีการส่งข้อมูลข้างเคียง (Side information) ไปตามช่องสัญญาณด้วย ข้อมูลเหล่านี้ได้แก่ สัมประสิทธิ์ของตัวทำนายและค่าคาบของเสียง

ซีอีแอลพี อาจจะเรียกได้ว่าเป็นตัวเข้ารหัสแบบไฮบริด (Hybrid Coder) เพราะว่ามีคุณสมบัติของทั้งการเข้ารหัสเสียงตามรูปคลื่นและการเข้ารหัสเสียงตามลักษณะของเสียงพูดรวมไว้ด้วยกัน



รูปที่ 2.1 ตัวเข้ารหัสของซีอีแอลพีแบบปรกติ

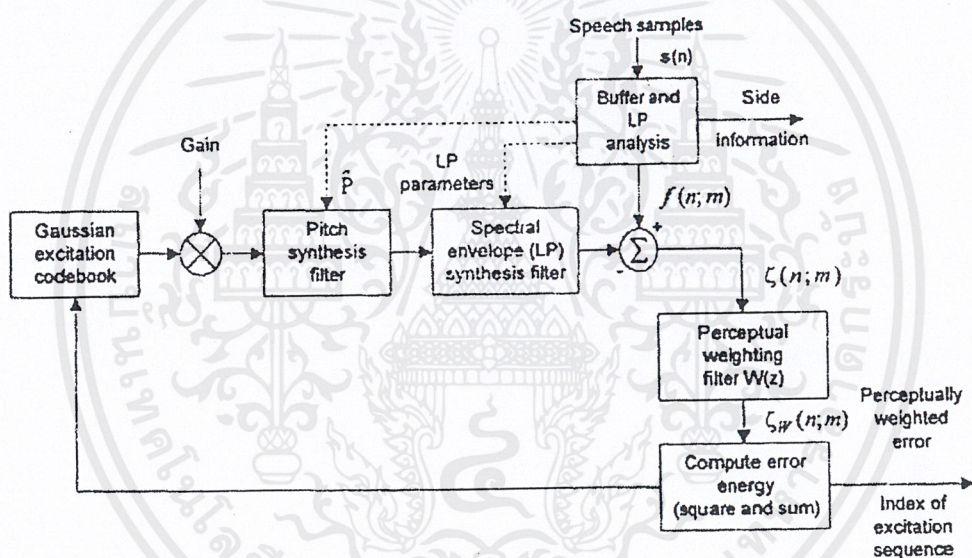
ในการควอนไทซ์แบบเวกเตอร์ใช้ชุดรหัสขนาด 10 บิตมีลำดับของตัวกระตุ้นได้ทั้งหมด 1024 แบบต่อเวกเตอร์ของตัวกระตุ้น (Excitation vector) ขนาด 40 ตัวอย่าง ซึ่งพบว่าเพียงพอสำหรับให้คุณภาพเสียงที่ดีซีอีแอลพี ที่ทำงานที่อัตราข้อมูล 16,000 บิตต่อวินาทีจะให้คุณภาพเสียงระดับเครือข่าย ซึ่งดีกว่าระบบอื่นที่ทำงานในอัตราเดียวกันแต่จะมีความล่าช้าทางเวลาที่มากกว่าเนื่องจากการเข้ารหัสที่ต้องใช้การวิเคราะห์ข้อมูลจำนวนมาก การทำงานเป็นลักษณะเฟรมต่อเฟรม ซึ่งขนาดเฟรมที่ใช้จะเป็นเวลาประมาณ 40 ถึง 60 มิลลิวินาที แต่ยังมีวิธีที่จะลดความล่าช้าทางเวลาลงไปอีก ดังในข้อต่อไป

2.6.4) โลดดีเลย์ซีอีแอลพี (Low-Delay CELP: LD-CELP) ใช้อัตราข้อมูล 16 กิโลบิตต่อวินาที โดยให้คุณภาพของเสียงที่ดีเท่ากับหรือดีกว่าของเอดีพีซีเอ็ม แต่มีความต้องการการคำนวณที่สูงมาก ได้เป็นมาตรฐานของการเข้ารหัสเสียง G.728 ที่กำหนดโดย CCITT ในปี 1992 การทำงานโดยรวมเป็นการพัฒนามาจากซีอีแอลพี แบบปกติแต่สามารถลดความล่าช้าทางเดียว (one-way delay) ลงมาเหลือน้อยกว่า 2 มิลลิวินาที โดยการปรับเปลี่ยนค่าแบบย้อนหลัง (backward adaptation) ของสัมประสิทธิ์ตัวทำนาย พันระเชิงเส้น อัตราขยายและสัมประสิทธิ์การท่วงคุณภาพการรับฟัง และยังใช้ขนาดของเวกเตอร์ของตัวกระตุ้นขนาดเล็กเพียง 5 ตัวอย่างเท่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6.5) เวกเตอร์ซัมเอ็กซิตไลเนียร์พรีดิคทีฟ (Vector Sum Excited Linear Prediction : VSELP) ให้คุณภาพเสียงระดับการสื่อสาร โดยใช้อัตราข้อมูล 8 กิโลบิตต่อวินาที เป็นมาตรฐานของการเข้ารหัสเสียงของระบบโทรศัพท์เคลื่อนที่ดิจิทัลแบบเซลล์ลาร์ที่ใช้ในทวีปอเมริกาเหนือ มีลักษณะของซุกรหัสที่แตกต่างไปจากซีอีแอลพี คือ มีซุกรหัส 3 ชุด ซุกรหัส 1 และซุกรหัส 2 แต่ละซุกรหัสจะประกอบด้วย 128 คำรหัสที่เกิดจากการรวมกันแบบเชิงเส้นของเบซิส(basis) 7 ตัว แทนที่จะกำหนดมาเป็น 128 คำรหัสที่เป็นอิสระต่อกันและพื้นฐานเหล่านี้จะมีการปรับค่าเพื่อลดค่าความผิดพลาดถ่วงน้ำหนักทั้งหมด (total perceptually weighted error) ให้น้อยที่สุด

ฟิลเตอร์สเตทระยะยาว (long-term filter state) เป็นซุกรหัสที่ประกอบด้วย 128 คำรหัสเช่นกัน โดยเป็นซุกรหัสแบบปรับเปลี่ยนได้ผลลัพธ์จากแหล่งกำเนิดตัวกระตุ้น (Excitation source) ทั้ง 3 นี้จะถูกคูณด้วยอัตราขยายซึ่งแยกอิสระกันของแต่ละตัวแล้วนำมารวมกันเพื่อผ่านตัวกรองในส่วนในตัวเข้ารหัสก็จะคล้ายกันกับตัวเข้ารหัสของซีอีแอลพี ทั่วไป

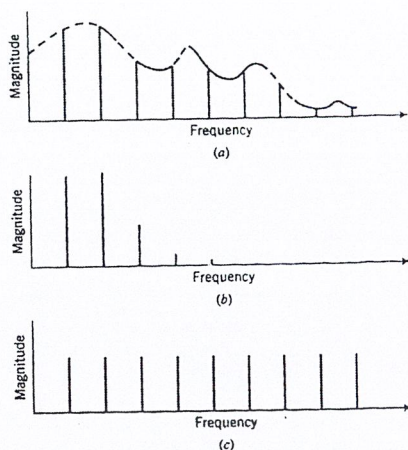


รูปที่ 2.2 ตัวถอดรหัสของ VSELP

2.7 การกระตุ้นเสียงและการทำสเปกตรัมให้เรียบ (Voice Excitation and spectral flattening)

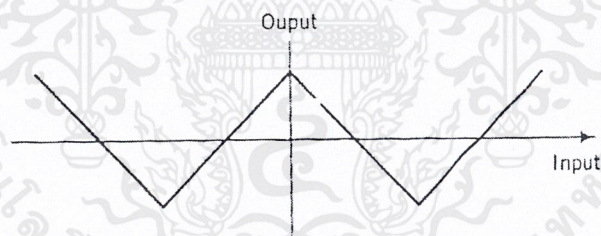
เกิดจากความพยายามที่จะปรับปรุงคุณภาพของอุปกรณ์เข้ารหัสเสียง (Vocoder) ซึ่งมาจากความเข้าใจที่ว่า การที่หน้าที่การกระตุ้น (Excitation function) ไม่มีประสิทธิภาพจะส่งผลให้เกิดการลดทอนคุณภาพของเสียง โดยที่จะกล่าวถึงเป็นสิ่งแรกคือ การพัฒนา ตัวเข้ารหัสแบบกระตุ้นเสียง (Voice-excited Vocoder) โดยแสดงไว้ในรูปที่ 2.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



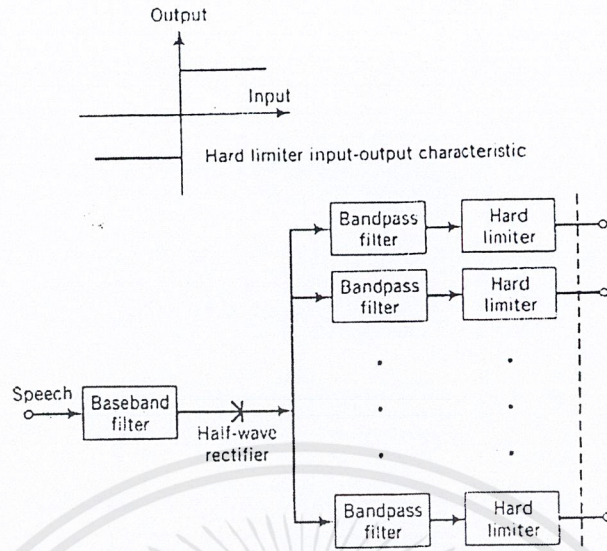
รูปที่ 2.3 การทำสเปกตรัมให้ราบเรียบของสัญญาณเบสแบนด์เพื่อทำการผลิตสัญญาณกระตุ้นตัวเข้ารหัสเสียง

จากรูปที่ 2.3 แสดง สเปกตรัมระยะสั้น (short-time spectrum) ของสัญญาณเสียง โดยเป็นขั้นตอนที่ก่อให้เกิดสเปกตรัม ดังรูปที่ 2.3a ซึ่งเป็นรูปของสเปกตรัมต้นแบบที่ถูกทำให้ราบเรียบ (Spectral flattening) แล้วจึงเป็นสัญญาณกระตุ้น (Excitation signal) ที่เหมาะสมได้ โดยในรูปที่ 2.3a จะเป็นสัญญาณสเปกตรัมต้นแบบ และได้มีการนำสัญญาณที่ผ่านวงจรกรองความถี่ต่ำ ดังรูปที่ 2.3b มาผ่านวงจรซิกแซก (zig-zag network) ดังรูปที่ 2.4 จะเกิดสัญญาณที่มีการลดทอนหลายสัญญาณ



รูปที่ 2.4 วงจรซิกแซก

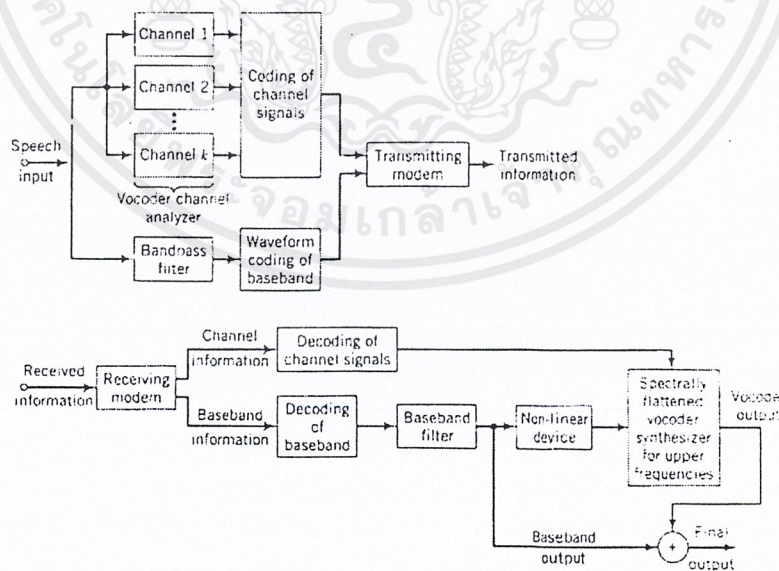
สัญญาณเสียง (speech) หลังจากผ่านวงจรกรองความถี่ต่ำ จะได้สัญญาณเบสแบนด์ และเมื่อนำไปผ่านวงจรลดทอน (distortion network) เช่น วงจรเรียงกระแสแบบครึ่งคลื่น (half-wave rectifier) จะได้ฮาร์โมนิกด้านบน (upper harmonics) และเมื่อนำสัญญาณที่ได้ไปผ่านฟิลเตอร์แบงก์ (filter bank) จะเป็นการขยายสเปกตรัมในส่วนที่สนใจทั้งหมด และเอาที่พู่ที่ได้จากวงจรกรองแถบความถี่ผ่านจะถูกจำกัดได้ยาก ดังรูปที่ 2.5



รูปที่ 2.5 ขั้นตอนการทำงานของการทำงานการทำสเปกตรัมให้ราบเรียบ

2.8 การเข้ารหัสเสียงแบบช่องสัญญาณการกระตุ้นเสียง (Voice-Excited Channel Vocoder)

วิธีการทำให้สเปกตรัมราบเรียบ (spectrum flattening) ไม่ใช่วิธีการตามอุดมคตินัก เช่น ถ้ามี ฮาร์โมนิกเกิดขึ้น 2 ฮาร์โมนิกในฟิลเตอร์ตัวเดียวกัน ฮาร์โมนิกที่มีความแรงน้อยกว่าจะถูกกด (suppress) โดยฮาร์โมนิกที่มีความแรงมากกว่า อย่างไรก็ตามแนวคิดนี้ได้ถูกนำมาใช้เป็นส่วนหนึ่งของการเข้ารหัสเสียงโดยการกระตุ้น (voice-excited channel vocoder) ดังรูปที่ 2.9



รูปที่ 2.6 การเข้ารหัสเสียงโดยการกระตุ้น (Voice-excited channel vocoder) และแสดงตัวทำการวิเคราะห์ (ส่วนบน) และ ส่วนสังเคราะห์ (ด้านล่าง)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยทั่วไปแบนด์วิธของสัญญาณเบสแบนด์จะมีค่าอยู่ระหว่าง 600-900 เฮิรตซ์ ในสัญญาณที่มีแบนด์วิธสูงกว่า จะมีคุณภาพดีกว่า แต่ก็จะมีอัตราการส่ง (transmission rate) สูงกว่าเช่นกัน

สัญญาณเบสแบนด์จะถูกส่งโดยเทคนิคการเข้ารหัสรูปคลื่น (waveform) มาตรฐาน และที่เครื่องรับสัญญาณเบสแบนด์จะมีการถอดรหัส และทำสเปคตรัมให้ราบเรียบโดยมีการกระตุ้นในส่วนของตัวเข้ารหัสเสียง (vocoder) สำหรับความถี่ที่สูงกว่าสัญญาณเบสแบนด์ และยังผ่านการเข้ารหัสเสียงพูด (speech) เพื่อที่จะได้เอาที่พูดออกมา

รูปที่ 2.5 แสดงเบสแบนด์ฟิลเตอร์ และส่วนของฟิลเตอร์แบงก์การวิเคราะห์ (synthesis filter bank) ที่เครื่องรับ อุปกรณ์ไม่เป็นเชิงเส้น (non linear) เป็นอุปกรณ์ที่จำกัดได้ยาก และในการที่จะกำจัดฮาร์โมนิกที่เกิดจากตัวจำกัด (limiter) จะมีการใช้ฟิลเตอร์แบงก์ (filter bank) พิเศษต่อกับ ตัวจำกัด และช่องสัญญาณรับของการเข้ารหัสสัญญาณเสียง (vocoder) เกิดขึ้นที่บริเวณส่วนบนของรูปที่ 2.6 และจะนำไปผ่านมอดูเลเตอร์ที่อยู่ระหว่าง ตัวจำกัด (limiter) กับ ฟิลเตอร์แบงก์ (filter bank)

2.9 การกระตุ้นเสียงและการกระตุ้นค่าความผิดพลาดของการเข้ารหัสเสียงแบบแอลพีซี (Voice-Excited and Error-Signal_Excited LPC Vocoders)

หลักการของการกระตุ้นเสียงของการเข้ารหัสเสียงแอลพีซี (voice-excited LPC vocoder) เหมือนกับการเข้ารหัสเสียงช่องสัญญาณ (channel vocoder) โดยการทำให้สเปคตรัมให้ราบเรียบจะเกิดขึ้นและผ่านการอินทิเกรตเพื่อเข้าสู่ส่วนของ การเข้ารหัสเสียงแอลพีซี (LPC vocoder)

ในระบบแอลพีซี จะมีการเพิ่มเติมส่วนประกอบเข้าไป คือเศษเหลือ (residual) หรือ สัญญาณผิดพลาด (error signal)

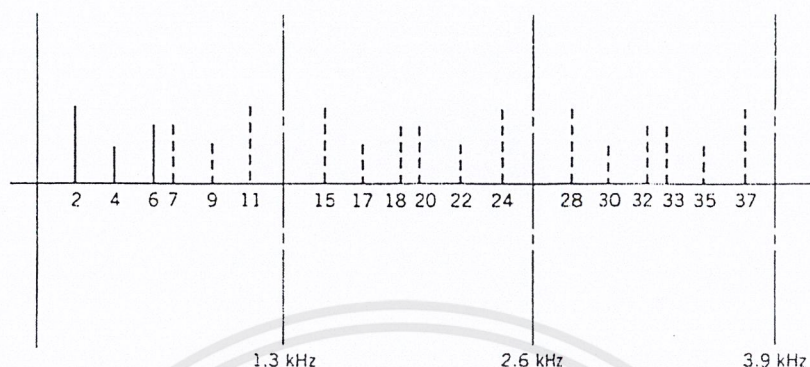
การกระตุ้นการสังเคราะห์แอลพีซี (LPC analysis) ด้วยสัญญาณผิดพลาดที่สมบูรณ์แบบ จะเป็นการทำให้เกิดอินพุตเสียงพูดอีกครั้งหนึ่ง (reproduce) แต่ปัญหาที่คือการที่จะหารูปแบบที่เหมาะสมสำหรับการบีบอัดสัญญาณผิดพลาดที่จะสามารถปรับตัวเพื่อทำหน้าที่เป็นตัวกระตุ้นที่ดีได้

มีการค้นคว้าของการกระตุ้นเศษเหลือของการเข้ารหัสเสียงแบบแอลพีซี (residual excited LPC vocoder) ที่สามารถทำงานได้ที่ความเร็ว 9600 บิตต่อวินาที และ การกระตุ้นสัญญาณพื้นฐานของแอลพีซี (voice baseband-excited LPC) โดยมีการใช้สัญญาณผิดพลาด หรือสัญญาณเสียงต้นแบบ (original speech) แต่ก็ยังคงมีปัญหาเช่นกันคือ ที่ส่วนทำการวิเคราะห์ (analyzer) สัญญาณจะต้องผ่านตัวกรองความถี่ต่ำผ่านก่อน เพื่อที่จะรักษาแบนด์วิธของการส่ง

การลดทอนแบบไม่เป็นเส้นตรง และการทำให้สเปคตรัมให้ราบเรียบ จะเกิดขึ้นที่ส่วนของการสังเคราะห์ (synthesizer) โดยใช้วิธีการนำเศษเหลือ (residual) มาผ่านอุปกรณ์ลดทอนแบบไม่เป็นเส้นตรง เพื่อที่จะนำความถี่สูงกลับมา ซึ่งวิธีนี้เป็นวิธีที่ไม่ค่อยน่าพอใจนัก เนื่องจากองศาของการเบี่ยงเบนจากการทำให้สเปคตรัมให้ราบเรียบในแบบที่ต้องการนั้น ขึ้นอยู่กับการขยายในอัตราสูงของการเปลี่ยนแปลงทางเวลาของสเปคตรัมของสัญญาณเสียงพูด (speech) และได้มีการเสนอทางเลือกสำหรับวิธีใหม่ ดังอธิบายในรูปที่ 2.7 จะเห็นว่าสัญญาณเบสแบนด์จะถูกจำลองโดยการทำการแซมปลิง (down sampling) ที่เหมาะสม

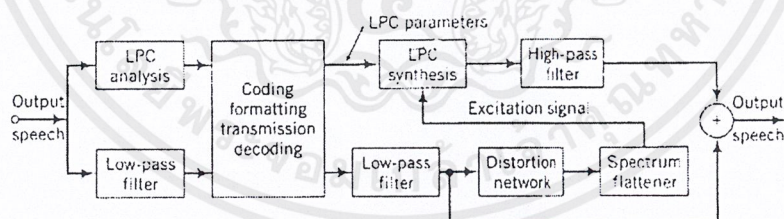
ซึ่งในแต่ละแบบตามอุดมคติ จะมีการเบี่ยงเบน 2 แบบ คือ สัญญาณที่ถูกจำลองจะประกอบไปด้วยเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขอบของสเปกตรัล (spectral envelope) ที่เหมือนกับสัญญาณเบสแบนด์ และองค์ประกอบของฮาร์โมนิกของสัญญาณจำลองจะ ไม่เป็นจำนวนเต็มเท่าของความถี่พื้นฐาน ยกเว้นในบางกรณีเท่านั้น

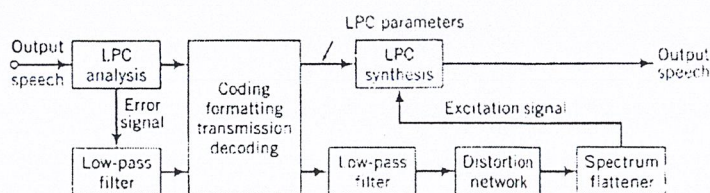


รูปที่ 2.7 แสดงการเลือกวิธีของการทำสเปกตรัมให้ราบเรียบโดยการใช้อัตราแซมปลิง

และยังมีบางรูปแบบใช้แนวความคิดสัญญาณความผิดพลาดของแอลพีซี (LPC error-signal) เพื่อที่จะแสดงการทำสเปกตรัมให้ราบเรียบ ซึ่งก็คือวิธี การกระตุ้นเสียงการประมาณเชิงเส้น (VELP :voice-excited linear prediction) ซึ่งจะตรงกันข้ามกับวิธีการกระตุ้นเศษเหลือการประมาณเชิงเส้น (REL P : residual-excited linear prediction) และส่วนของ การวิเคราะห์ แอลพีซี จะถูกแสดงในรูปของสัญญาณที่รับได้ การเบี่ยงเบน และสัญญาณเบสแบนด์ที่ก่อให้เกิดสัญญาณผิดพลาด ตัวใหม่ และถ้าเราให้การวิเคราะห์แบบนี้มีความถูกต้องของสเปกตรัม เราจะได้ว่าสัญญาณผิดพลาดที่ได้นั้น จะมีสเปกตรัมที่ราบเรียบ และจะเป็นประโยชน์แก่การทำการกระตุ้น



(a) VELP (Voice Excited Linear Prediction)



(b) RELP (Residual-Excited Linear Prediction)

รูปที่ 2.8 แสดงบล็อกไดอะแกรมอย่างง่ายของวีอีแอลพี และ อาร์อีแอลพี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งในวิธีแอลพี นั้นการกระตุ้นจะได้รับโดยตรงจากสัญญาณที่นับได้ ซึ่งก็คือ เสียงพูดที่ผ่านตัวกรองความถี่ต่ำผ่านแล้วและสัญญาณเบสแบนด์จะสามารถถูกรวมเข้าสู่เสียงที่ถูกเข้ารหัสช่วงสูงผ่าน (high-pass vocoded speech) และในวิธีอาร์อีแอลพี นั้นการประมวลผลสัญญาณความผิดพลาด จะถูกทำเฉพาะในส่วนการกระตุ้นเท่านั้น

ได้มีการเปรียบเทียบระหว่างระบบแบ่งแยกช่องสัญญาณ (split-band system) กับเสียงหรือระบบการกระตุ้นสัญญาณความผิดพลาด (voice or error-signal-excited system) เป็นมิลลิวินาที โดยในแบบระบบแบ่งแยกช่องสัญญาณ (split band system) เป็นมิลลิวินาที นั้นสัญญาณเบสแบนด์จะถูกส่งไป เพื่อเป็นการแทนที่การลดทอนของสัญญาณที่เครื่องรับเพื่อทำการสร้างองค์ประกอบความถี่ที่สูงกว่า โดยในระบบนี้จะก่อให้เกิดประโยชน์จากการตรวจจับพิทช์ (pitch detector) อย่างชัดเจนที่ความถี่ พื้นฐาน (Baseband) จะมีการเข้ารหัสเวฟฟอร์ม (waveform) ในขณะที่ ความถี่สูงกว่าจะถูกเข้ารหัสตัวต่อตัว โดยหลาย ๆ วิธี ส่วนข้อได้เปรียบของวิธีการกระตุ้นเสียงหรือเศษเหลือ (voice or residual-excited) คือ

- พิทช์ (pitch) และการตรวจจับสัญญาณเสียงไม่มีความจำเป็น ดังนั้นจะเป็นการประหยัดด้านฮาร์ดแวร์ (Hardware) และความจุช่องสัญญาณ
 - คุณภาพในสภาพแวดล้อมที่มีสิ่งรบกวนได้ถูกปรับปรุงเนื่องจากไม่มีพารามิเตอร์พิทช์ (pitch parameter) ที่จะถูกประเมิน หรือการตัดสินใจคุณภาพเสียงเกิดขึ้น
- ดังนั้นจะเห็นว่า ระบบการแยกแถบความถี่ (split-band) นั้นมีความต้องการด้านฮาร์ดแวร์ เป็นพิเศษ หรือบางทีอาจเป็นซอฟต์แวร์ด้วยการทดสอบ ไชโครอะคูสติก (psycho acoustic) แสดงให้เห็นว่า ที่ความถี่สูงกว่า 1500 เฮิร์ตซ์ความผิดพลาดของพิทช์ จะไม่มีผลใด ๆ

2.10 การเข้ารหัสเสียงด้วยวิธีการทำนาย (Waveform coding with predictive Methods)

โดยปกติแล้ว เนื่องจากสเปกตรัมของสปีคไม่ได้เป็นผลมาจากคอร์รีเลชัน (correlation) ของเวฟฟอร์มแซมเปิล (waveform sample) ดังนั้นการที่จะสามารถใช้อัตราบิต (bit rate) ที่น้อย ๆ จะสามารถเป็นไปได้โดยใช้วิธีเดลต้ามอดูเลชัน (Delta modulation)

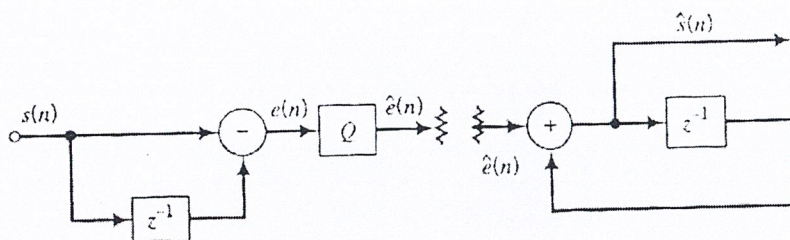
เราจะสามารถสรุปได้ว่า มีบิตจำนวนหนึ่งจะถูกใช้ในการเข้ารหัสความแตกต่างระหว่าง สองแซมเปิลที่ติดกันมากกว่าที่จะต้องใช้ไปในการควอนไทเซชันแซมเปิลโดยตรง โดยสามารถสรุปเป็นสูตรได้ดังนี้

$$e(n) = s(n) - s(n-1) \quad (2.3)$$

$$s(n) - \text{signal} \quad (2.4)$$

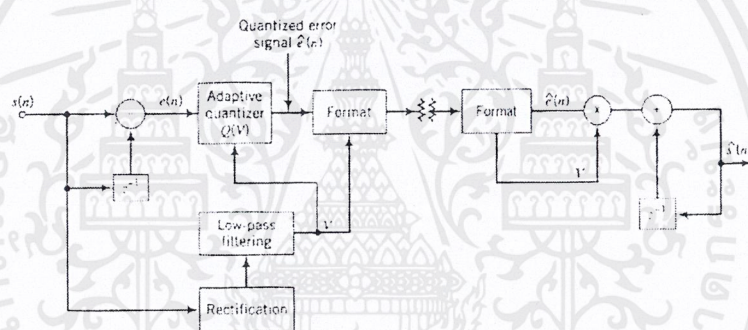
$$\hat{e}(n) - \text{quantize signal of } e(n) \quad (2.5)$$

ระบบในรูป 2.9 เป็นระบบคิฟเฟอเรนเชียลพัลส์โค้ดมอดูเลชัน (differential pulse code modulation : DPCM) และมีการแสดงการเข้ารหัสสัญญาณและถอดรหัสสัญญาณ



รูปที่ 2.9 ระบบคิฟเฟอเรนเชียลพัลส์โค้ดมอดูเลชัน

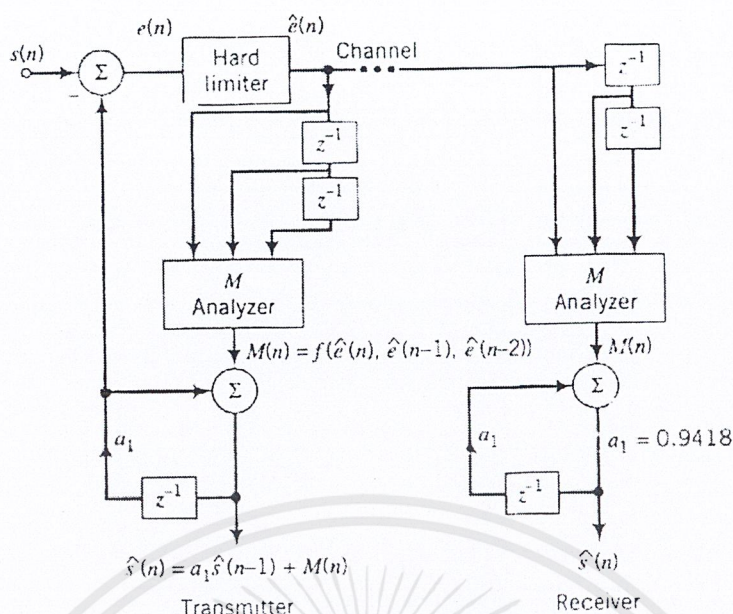
และมีวิธีการที่เรียกว่าเทคนิคอะแดปทีฟควอนไทซ์เซชัน (adaptive quantization technique) ที่สามารถลดจำนวนบิตได้ ดังแสดงในรูปที่ 2.10



รูปที่ 2.10 ระบบอะแดปทีฟคิฟเฟอเรนเชียลพัลส์โค้ดมอดูเลชัน

ความยากของระบบในรูปที่ 2.10 คือ ปัญหาความชันมากเกินไป (slope-overload problem) ซึ่งถ้าตัวอินทิเกรต (integrator) มีการกำหนดค่าขนาดของขั้น (step size) ผลตอบสนองจะมีการเพิ่มขึ้นของสัญญาณอย่างรวดเร็ว เป็นผลทำให้สัญญาณที่ถูกเข้ารหัสมีค่าตกลง

และเพื่อที่จะแก้ปัญหานี้ จึงได้มีการเสนอวิธีให้ใช้ควอนไทเซอร์ ที่สามารถปรับค่าสเต็ปไซส์ ได้ ดังรูปที่ 2.11 ซึ่งเรียกว่าซีวีเอสดี (Continuously variable slope delta Modulation : CVSD)



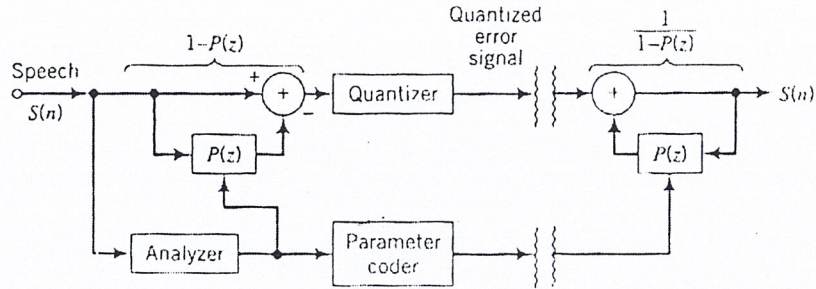
รูปที่ 2.11 เกล็ดต้ามอดูเลชันหลายความชันแบบต่อเนื่อง (Continuously variable slope delta modulation)

ในขั้นแรก การสังเคราะห์สัญญาณเสียงจะเกิดขึ้นที่ภาคส่ง ดังนั้นเอาท์พุทของมันจะสามารถแสดงผลได้อย่างต่อเนื่อง และสามารถทำการเปรียบเทียบกับสัญญาณอินพุทได้ ดังนั้นสัญญาณผิดพลาด ก็คือค่าผลต่างระหว่างสัญญาณอินพุท กับ สัญญาณที่สังเคราะห์ขึ้นมา

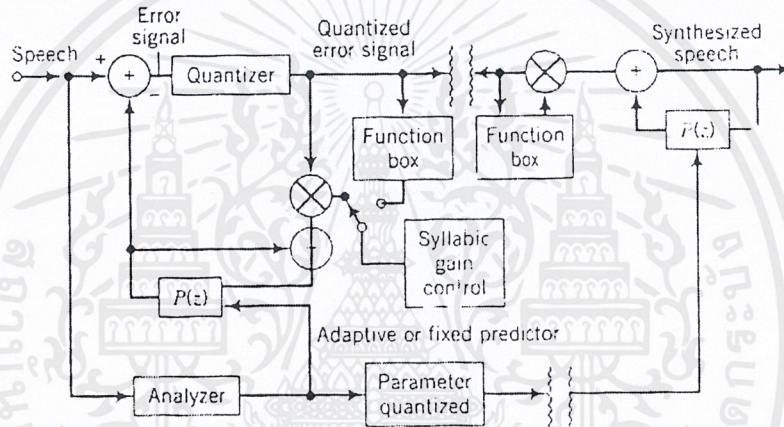
จากรูปที่ 2.11 ตัวทำการวิเคราะห์ M จะทำหน้าที่เปลี่ยนควอนไทซ์สเต็ป ถ้ามีสัญญาณผิดพลาด 3 สัญญาณติดกันที่มีสัญลักษณ์เดียวกันสเต็ปไซส์ จะถูกเพิ่มค่าขึ้น ถ้าเป็นกรณีอื่น ๆ สเต็ปไซส์ จะถูกลดค่าลง การเพิ่มขึ้นหรือลดลงของสเต็ปไซส์ นี้้อาจจะมีค่าไม่เท่ากัน

2.11 การเข้ารหัสเสียงแบบการทำนายอะแด็ปทีฟ (Adaptive predictive coding of speech)

จากรูปที่ 2.12 แสดงถึงพื้นฐานของรูปแบบของการประมาณเชิงเส้น (linear prediction model) $1-p(z)$ จะแทนการกรองเชิงเส้น (linear filtering) ของสปีชแชนเนล และตัวทำการวิเคราะห์ (analyzer) จะทำการเลือกพารามิเตอร์ของ $|1-p(z)|$ ตามวิธีการ (algorithm) ที่จะใช้



รูปที่ 2.12 แนวความคิดแบบ ลูติเมนทาร์ลิเนียร์พรีดิกชัน (Rudimentary linear prediction concept)

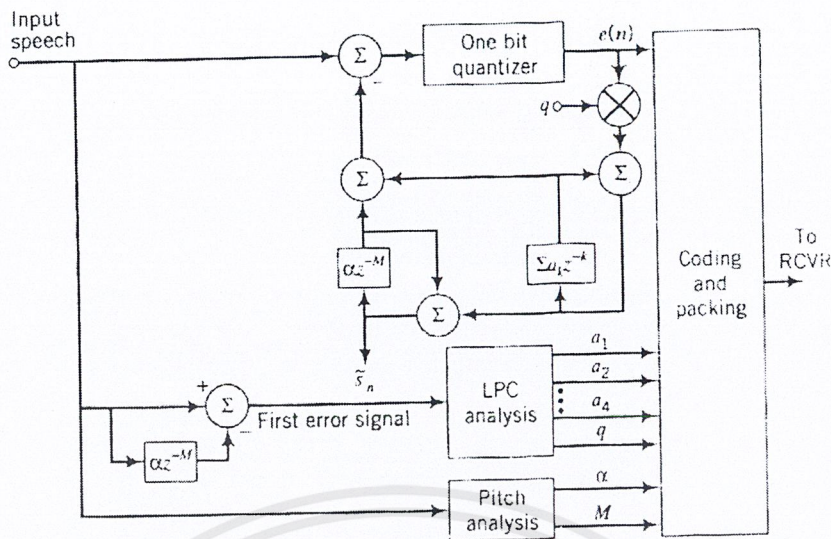


รูปที่ 2.13 พรีดิกเตอร์และควอนไทเซอร์ในลูป

โดยทั้งสัญญาณผิดพลาด และพารามิเตอร์ จะถูกควอนไทซ์ และถูกส่งพร้อมกัน ตัวอินเวอร์สฟิลเตอร์ (inverse filter) ที่ส่วนทำการสังเคราะห์ (synthesizer) จะถูกกระตุ้นจากสัญญาณผิดพลาดที่ได้รับ เพื่อทำการผลิตสัญญาณเอาท์พุท

ตัวกรองสัญญาณที่เกิดจากการสังเคราะห์จะมีอยู่ในอะนาลิเซอร์ฟีดแบคลูป (analyzer feedback loop) และแมกนิจูด (magnitude) ของสัญญาณผิดพลาด ที่ถูกควอนไทซ์ แล้วจะถูกปรับค่าเพื่อให้มีความผิดพลาดน้อยลง ซึ่งควบคุมอัตราการขยาย (gain control) ที่อยู่ในลูปย้อนกลับ (feed back loop) สามารถปรับค่าสำหรับแต่ละสปีชเซมเปิ้ล (ดังเช่นในซีวีเอสดี :CVSD)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.14 แสดงแนวความคิดของ อะแดพทีฟพรีดิกทีฟโค้ดดิ้ง (adaptive predictive coding :APC)

ในระบบนี้ การทำนายเสียงมีพื้นฐานอยู่บนสองอย่างคือ (ก) วิธีการวัดค่าของคลื่นเสียง 1 คาบพื้นฐาน (ข) การวิเคราะห์แอลพีซี ซึ่งได้แก่สัญญาณผิดพลาดซึ่งเป็นผลมาจาก (ก)

โดยคาบพื้นฐานได้มาจากการคำนวณของตัวจับสัญญาณพิทช์ (pitch detector) และ α_1 ได้มาจากการคำนวณค่าที่น้อยที่สุดของค่าความผิดพลาดเฉลี่ยกำลังสอง (Mean-Square-error) ของสัญญาณผิดพลาดตัวแรก พารามิเตอร์ เช่น a_1, a_2, a_3 และ a_4 ได้จากการคำนวณในขั้นตอนของการวิเคราะห์แอลพีซีของ 1 ตัวแรก

$e(n)$ คือ สัญญาณผิดพลาด ที่ถูกส่ง จะคำนวณดังที่เห็นในภาพ โดยพารามิเตอร์ q ถูกควบคุมที่อัตราของพยางค์เสียงและทรานสเฟอร์ฟังก์ชัน (transfer function) ระหว่าง สัญญาณ $e(n)$ กับ $\hat{s}(n)$ จะถูกมองในรูปของ ตัวกรองการสังเคราะห์ที่เหมาะสม

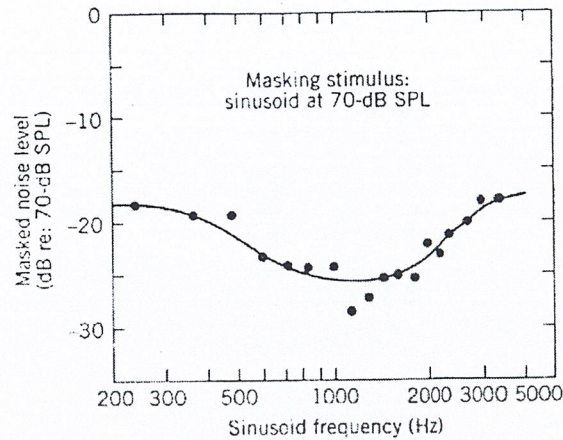
เอพีซีอัลกอริทึม ถูกจัดการเพื่อที่จะลด อัตราข้อมูลผิดพลาด โดยใช้บิต จำนวนน้อย ๆ โดยที่ระบบเอพีซี ทำงานได้ดีที่อัตรา 8 กิโลบิตต่อวินาที

2.12 การเข้ารหัสแบนด์ย่อย (Subband coding)

การเข้ารหัสเวฟฟอร์มต่างจากการเข้ารหัสเสียง (vocoding) เนื่องจากไม่สามารถให้ประโยชน์ในการผลิตสัญญาณเสียงได้อย่างชัดเจน อย่างไรก็ตามความเหมาะสมของสัญญาณเสียงและระบบการรับฟังของมนุษย์สามารถปรับปรุงให้ดีขึ้นโดยการออกแบบระบบการส่งสัญญาณเสียงที่ผ่านการเข้ารหัสแล้ว ดังตัวอย่างเช่น ในการคุยโทรศัพท์จะมีช่วงของการสนทนาที่ 50 เปอร์เซ็นต์ ของช่วงเวลาทั้งหมด และสัญญาณแถบความถี่แคบสามารถทำการควอนไทซ์สัญญาณรบกวนได้ ถ้าสเปกตรัมของสัญญาณรบกวนอยู่ในย่านของความถี่สัญญาณ แนวความคิดนี้นำไปสู่การเข้ารหัสแบนด์ย่อย (subband coding) ซึ่งในขั้นแรกเสียง จะถูกนำไปผ่านฟิลเตอร์แบงก์ (filter bank) หลังจากนั้นแต่ละย่านความถี่จะถูกควอนไทซ์แบบ

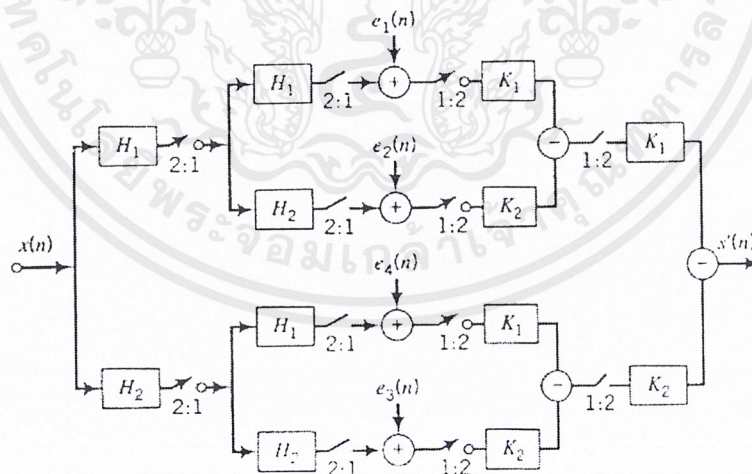
แยกกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.15 แสดงผลการทดลองเพื่อที่จะหาค่าความเหมาะสมของระดับของสัญญาณรบกวน (noise) ที่อยู่ใน คริติคัลแบนด์ (critical band) เราจะพบว่าสัญญาณรบกวน ควรมีค่าต่ำกว่าสัญญาณอยู่ 18–28 dB ก่อนที่จะผ่านการมาส์ค (mask)

ได้มีการเสนอฟิลเตอร์แบบควอดรูเชอร์มิลเลอร์ (quadrature mirror : QMFs) ซึ่งมีคุณสมบัติดังนี้ ภายใต้สภาวะที่แน่นอน การเกิดเอเลียสซิ่ง (aliasing) จะมีผลมาจากการซ้อนทับกันของย่านความถี่ ซึ่งสามารถยกเลิกได้ และ คิวเอ็มเอฟ (QMF) ที่จะแยกแบนด์วิธ ทั้งหมดเป็น 2 ย่านที่สมมาตรกัน โดยฟิลเตอร์ r ตัวบนจะสามารถรับจากฟิลเตอร์ตัวล่างได้พอดี โดยการแทนที่ของสัมประสิทธิ์ที่เป็นเลขคู่



รูปที่ 2.16 แสดงคิวเอ็มเอฟแบบ 4 ช่องสัญญาณ (channel)

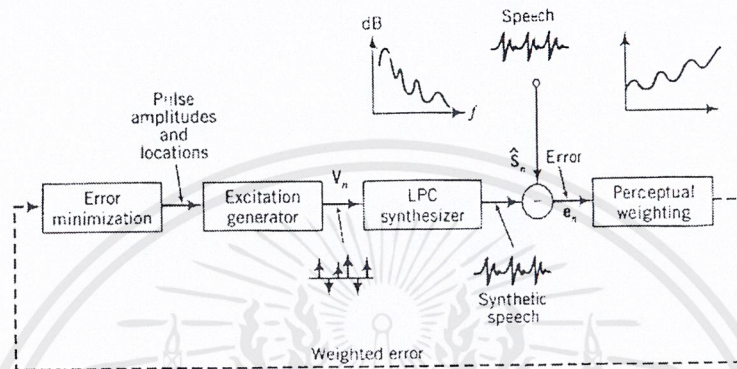
2.13 การเข้ารหัสเสียงแอลพีซีแบบหลายพัลส์ (Multipulse LPC Vocoder)

ในการเข้ารหัสเสียงแบบหลายพัลส์นั้นเศษเหลือจะถูกเสียระบบโดยผลรวมของอิมพัลส์ (impulse) หลาย ๆ ระดับซึ่งมักมีลักษณะเป็นหลายแอมพลิจูดและตำแหน่ง ดังนั้นการที่จะส่งสัมประสิทธิ์แอลพีซี ไปนั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แอมพลิจูดและตำแหน่งของการกระตุ้นอิมพัลส์ จะต้องถูกส่งไปด้วย ปัญหาจึงอยู่ที่ว่า ทำอย่างไรจึงจะสามารถระบุตำแหน่งและความสูงของอิมพัลส์ได้

ซึ่งเราจะสามารถทำได้โดยวิธีการวิเคราะห์โดยการสังเคราะห์ (analysis by synthesis) ของการเลือกอิมพัลส์ ซึ่งจะทำให้ค่าผลต่างระหว่างสัญญาณเสียงจริงกับสัญญาณเสียงที่เครื่องรับสร้างขึ้นใหม่มีค่าน้อยที่สุด ดังแสดงในรูปที่ 2.17



รูปที่ 2.17 มัลติพัลส์แอลพีซี (Multipulse LPC)

โดยอิมพัลส์จะถูกกรองสัญญาณผ่านตัวกรองของการสังเคราะห์แอลพีซี ที่สามารถสร้างสัญญาณที่ถูกเข้ารหัสขึ้นมาใหม่ซึ่งถูกนำไปลบออกจากสัญญาณจริง และนำไปผ่านเพอร์เซปชวลเวทติ้งฟิลเตอร์ (perceptual weighting filter) โดยเพอร์เซปชวลเวทติ้งฟิลเตอร์ ได้จากสมการ

$$w(z) = \frac{A(z)}{A(z/c)} \quad (2.6)$$

ซึ่ง c คือ ค่าคงที่ที่ทำหน้าที่ควบคุมจำนวนของเวทติ้ง
มีค่าเท่ากับ 1 : ไม่มีการเวทติ้ง (no weighting)
และ 0.7 – 0.9 โดยปกติ

$A(z)$ คือ ค่าพรีดิคเตอร์ฟิลเตอร์ (prediction filter)

เพอร์เซปชวลเวทติ้งฟิลเตอร์จะถือว่าการได้ยินของมนุษย์ที่มีผลต่อการกระทำระหว่างการกระตุ้นซีควเนต์ (sequence) กับแอลพีซี พารามิเตอร์ ต่อจากนั้น ค่าความผิดพลาดของเพอร์เซปชวลเวทติ้ง (perceptual weighted) จะถูกทำให้มีค่าน้อยที่สุดโดยการใช้ขบวนการที่มีประสิทธิภาพเพื่อเลือกค่าตำแหน่งอิมพัลส์ที่ดีที่สุด และ แอมพลิจูด

ถึงแม้ว่าการเข้ารหัสแอลพีซีแบบหลายพัลส์ ที่ประกอบด้วยการสังเคราะห์พิทช์ (pitch synthesis) จะค่อนข้างมีความสมบูรณ์และให้การสังเคราะห์เสียงที่มีคุณภาพสูง และยังมี การปรับปรุงคุณภาพของมันอยู่หลายทาง ซึ่งหนึ่งในนั้นก็คือซีอีแอลพี ซึ่งปรับปรุงมาจากการสังเกตการกระตุ้นลำดับซึ่งมีความคล้ายคลึงกันอย่างมาก จาก 1 คาบของพิทช์ไปยังคาบต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยการเพิ่มส่วนฟิลเตอร์สังเคราะห์พิทช์ ซึ่งมีสมการดังนี้

$$\theta_p(z) = \frac{c}{1 - bz^{-T}} \quad (2.7)$$

c : scale factor

b : ตัวควบคุมจำนวนของการทำนายค่า ($0 < b < 1$)

T : คาบของพิทช์

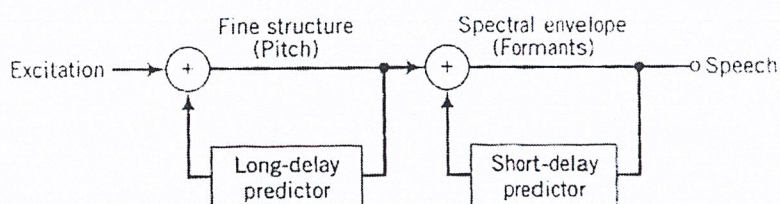
การเพิ่มส่วนฟิลเตอร์สังเคราะห์พิทช์เพื่อที่จะให้ผลต่างของสัญญาณกระตุ้น กับระดับสเกลเวอร์ชัน (scaled version) ของตัวก่อนหน้าเท่านั้นที่จะถูกส่งไป

2.14 การบีบข้อมูลเสียงพูดด้วยวิธีการทำนายเชิงเส้นด้วยการกระตุ้นด้วยรหัส (Code Excited Linear Predictive Coding ; CELP)

แม้ว่าการทำการทำนายเชิงเส้นหลายลูกคลื่น (multipulse LPC) จะมีการสังเคราะห์ระดับเสียงสูงต่ำ (pitch synthesis) จะสามารถปรับปรุงคุณลักษณะของการทำนายเชิงเส้นแบบเก่า แต่ด้วยการปรับเทคนิคการทำการเข้ารหัสแบบเวกเตอร์ควอนไทเซชัน นำมาใช้กับปัญหาการเข้ารหัสสัญญาณกระตุ้น (excitation signal) จะทำให้มีคุณลักษณะที่ดีขึ้นและนำไปสู่ระดับของสัญญาณรบกวนที่สามารถได้ยินได้ (audible distortion) ที่ต่ำกว่ารวมถึงความกว้างของช่องสัญญาณที่ต้องการจะมัน้อยลง โดยการเข้ารหัสแบบนี้จะเรียกว่า การบีบข้อมูลเสียงพูดด้วยวิธีการทำนายเชิงเส้นด้วยการกระตุ้นด้วยรหัส (Code Excited Linear Predictive Coding : CELP)

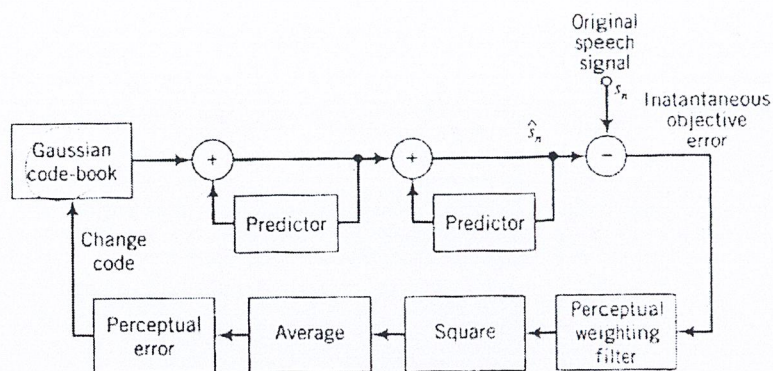
การเข้ารหัสแบบซีอีแอลพี พื้นฐาน ถูกแนะนำขึ้นมาโดยชอเฟอร์ (Schroder) และ เอเทล (Atal) โดยกล่าวว่า การวิเคราะห์ข้อมูลเสียงในการเข้ารหัสแบบซีอีแอลพี จะเหมือนกับวิธีการที่ใช้กับการเข้ารหัสแบบการทำนายการเปลี่ยนแปลง (Adaptive Predictive Coder : APC)

ในการเข้ารหัสแบบซีอีแอลพี ความแตกต่างของการกระตุ้นจากเฟรมหนึ่งไปอีกเฟรมหนึ่งจะไม่เป็นเป็นลักษณะพิเศษเจาะจงในอิมพัลส์ (impulse) บางตัว แต่จะเป็นตัวเลขที่เป็นค่าคงที่ในอนุกรมของโค้ดบิตตัวหนึ่ง ดังนั้นเพื่อจะทำการสังเคราะห์สัญญาณใหม่ อุปกรณ์สังเคราะห์ของซีอีแอลพี ดังรูป 2.18 จะสามารถนำมาใช้ได้ ในอุปกรณ์การเข้ารหัสซีอีแอลพี พื้นฐาน อนุกรมของโค้ดบิตจะถูกพิจารณาก่อน มีค่าเฉลี่ยเป็นศูนย์ มีความแตกต่างค่าเดียว และใช้การสุ่มตัวเลขแบบ เกาส์เซียน (Gaussian)



รูปที่ 2.18 ตัวสังเคราะห์ซีอีแอลพี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.19 ตัวเข้ารหัสสัญญาณซีอีแอลพี พื้นฐาน

ในการเข้ารหัสจะเลือกใช้โค้ดบุ้คแบบเกาส์เซียน (Gaussian) เพราะจากการตรวจสอบคุณลักษณะของข้อมูลเสียงทำให้ทราบว่าอนุกรมการกระตุ้นในการเข้ารหัสสัญญาณแบบแอลพีซีด้วยตัวกรองสังเคราะห์สังเคราะห์พิทช์ (pitch synthesis filter) โดยธรรมชาติแล้วจะเป็นแบบเกาส์เซียน (Gaussian) ซึ่งโค้ดบุ้คแบบอื่น ๆ นอกเหนือจากนี้จะอธิบายในภายหลัง

การเข้ารหัสแบบซีอีแอลพี นี้จะทำงานหนักในการเลือกอนุกรมของโค้ดบุ้ค หรือ อนุกรมการกระตุ้นที่จะทำให้การลดทอนที่สังเกตได้มีค่าน้อยที่สุด โดยจะสามารถทำได้ด้วยวิธีการวิเคราะห์โดยการสังเคราะห์ดังรูป 2.19 ซึ่งที่อยู่ของคำรหัส (code word) สำหรับแต่ละเฟรมซึ่งลดค่าความผิดพลาดที่ยอมรับได้แล้วจะถูกส่งไปยังตัวรับในรูปแบบของพารามิเตอร์มาตรฐานของแอลพีซี

การทำงานของซีอีแอลพี แบบแรกเริ่มหรือมาตรฐานที่เป็นพื้นฐาน จะทำการค้นหาตามโค้ดบุ้คทั้งหมดซึ่งค่อนข้างจะเป็นการเปลืองการทำงาน โดยโค้ดบุ้คทั้งหมดจะมีจำนวน 5,121,024 ตัว ซึ่งแต่ละตัวจะประกอบด้วย 40 ตัวอย่าง ซึ่งจะเป็นช่วงเวลา 5 มิลลิวินาที สำหรับอัตราการสุ่มตัวอย่างที่อัตรา 8 กิโลเฮิร์ตซ์ โดยที่แต่ละอนุกรมในโค้ดบุ้คจะถูกกรองด้วยการสังเคราะห์พิทช์ (pitch synthesis) และฟิลเตอร์การสังเคราะห์เชิงเส้น (LP synthesis filter) และลบจากเสียงจริงๆ จากนั้นจะถูกกรองโดยฟิลเตอร์เพอเซพชวลเวกต์ดิง (Perceptual weighting filter) ซึ่งจะเป็นการคำนวณ จำนวนมากโดยวิธีการลดเวลาการเข้ารหัสจะพิจารณาได้ในตอนต่อไป

อัตราส่วนของอัตราการขยายสำหรับแต่ละอนุกรมการกระตุ้นสามารถจัดไว้ว่าเหมือนกับแอมพลิจูดของอิมพัลส์ (Impulse Amplitude) ในมัลติพัลส์แอลพีซี (multipulse LPC) โดยพารามิเตอร์ของการทำนายเชิงเส้นจะถูกเข้ารหัสโดยใช้การเข้ารหัสแบบเวกเตอร์ควอนไทซ์เซชัน (Vector quantization) เหมือนที่ได้อธิบายไปแล้วในข้อ 2.13

เมื่อคาบของพิทช์ (pitch) และอนุกรมการกระตุ้นเปลี่ยนแปลงรวดเร็วยิ่งขึ้นมากกว่าสะเปก Spectra characteristic ของสัญญาณเสียง คัดนี้ของโค้ดบุ้ค อัตราการขยายและการประมาณค่าพิทช์ (Pitch) จะถูกส่งบ่อยขึ้นมากกว่าสัมประสิทธิ์ของการประมาณเชิงเส้น

มาตรฐาน FS1016 (Federal Standard 1016) กำหนด 4800 บิตต่อวินาที ของการเข้ารหัสซีอีแอลพี

ไว้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. พารามิเตอร์ของการประมาณเชิงเส้นสลิปตัวจะถูกปรับปรุงทุกๆ 30 มิลลิวินาที โดยใช้ 34 บิทหรือค่า 1133.3 บิตต่อวินาที
2. การประมาณค่าพิทช์ (pitch) ทั้งคาบเวลาและการประมาณการย้อนกลับของอัตราขยายระยะยาว (long-term feedback gain) จะทำการส่งเป็นจำนวนสี่ครั้งทุกๆช่วงเวลา 7.5 มิลลิวินาทีโดยอัตราขยายย้อนกลับ จะถูกเข้ารหัสเป็นตัวเลข 8 และ 16 บิท เมื่อมันถูกเข้ารหัสเป็นตัวเลข 8 บิท คาบเวลาของพิทช์ จะถูกส่งไป ในระหว่างคาบ 8 บิทนั้น ค่าความแตกต่าง 6 บิทจะถูกส่งออกไปจากกรณีที่พิทช์ไม่เปลี่ยนแปลง อย่างมากหรืออย่างรวดเร็วบ่อยนักจึงเพียงพอโดยค่าพิทช์ ประมาณรวม 1600 บิตต่อวินาที
3. ชุดท้ายตัวแสดงการกระตุ้น ไม้ค้อน จะถูกส่งทุกๆ 7.5 มิลลิวินาทีโดยใช้ครรชน 9 บิทเพื่อเจาะจงว่าต้องการใช้เวกเตอร์ใดในจำนวน 512 เวกเตอร์ใน ไม้ค้อน และอัตราขยาย 5 บิท เพื่อเจาะจงขนาด ซึ่งรวมแล้วจะต้องการ 1866.67 บิตต่อวินาที เมื่อเพิ่ม 200 บิตต่อวินาที ของการซิงโครไนซ์ (synchronize) และ การแก้ไขความผิดพลาดล่วงหน้าจะเป็น 4800 บิตต่อวินาที

2.14.1) การเปลี่ยนแปลงของซีอีแอลที

เหมือนกับเทคนิคใหม่ๆทั่วไปคือการเข้ารหัสเสียงแบบซีอีแอลที พื้นฐานจะมีการเปลี่ยนแปลงและมีการพัฒนาวิธีการ โดยจะถูกแบ่งเป็นสองประเภทคือการปรับปรุงที่มีจุดมุ่งหมายโดยเฉพาะเจาะจงที่จะลดการคำนวณที่เกี่ยวข้องในการค้นหาไม้ค้อน ซึ่งจะแสดงในส่วนที่ 2.15 ส่วนเป้าหมายของการปรับปรุงอื่นจะอยู่ที่การเพิ่มความแม่นยำหรือการปรับปรุงอื่นๆซึ่งจะนำเสนอต่อไป

2.14.2) อนุกรมของไม้ค้อนที่ไม่เป็นเกาส์เซียน (Non-Gaussian codebook sequence)

แม้ว่าสัญญาณรบกวนแบบเกาส์เซียน จะสามารถแทนการเปลี่ยนแปลงการกระตุ้นได้แต่มันจะไม่ดีนักในการเป็นตัวอย่างการเริ่มต้นของเสียงจากเสียงเงียบหรือเสียงส่วนที่เป็นเสียงไม่ก้อง (unvoiced) เมื่อแอลทีซี หลายลูกคลื่นจะสามารถเจาะจงตำแหน่งในอนุกรมการกระตุ้นและเพิ่มคาบอย่างรวดเร็ว ในคาบที่ต้องการในอิมพัลส์ที่ต้องการในเสียงก้อง (voice) ไม้ค้อน แบบเกาส์เซียนอย่างเดียวไม่เพียงพอจะไม่สามารถเจาะจงมากนักในทางเลือกของมัน โดยสามารถแก้ไขได้โดยรวมอนุกรมของบางการกระตุ้นที่แน่นอน (deterministic excitation) ใน ไม้ค้อน

ไม้ค้อนที่ปรับปรุงแล้วจะถูกแบ่งออกเป็น 4 ส่วนแสดงระดับของการกระตุ้นที่แตกต่างกัน 4 ระดับ ในระดับแรกจะใช้ครึ่งหนึ่งของ 512 ตำแหน่งไม้ค้อนเป็นอนุกรมการกระตุ้นสุ่มมาตรฐาน (standard stochastic excitation sequence) อีกระดับจะพิจารณาสัญญาณพัลส์เดี่ยวๆ รวมด้วย เพื่อสามารถใช้ได้กับการเริ่มต้นของเสียงที่รวดเร็ว ระดับที่สามแสดงค่าของสัญญาณพัลส์ที่คาบของพิทช์ปัจจุบัน ซึ่งประมาณในส่วนของการวิเคราะห์เชิงเส้นและใช้ในตัวกรองสัญญาณของตัวสังเคราะห์พิทช์ ซึ่งจะถูกรวมเพื่อการเริ่มของเสียงที่ค่อยๆเป็นไปทีละขั้น ในระดับที่สี่จะเกี่ยวข้องกับสัญญาณพัลส์ จากช่องสายเสียง (Glottal) สำหรับรายละเอียดการกระตุ้นมากขึ้นที่คาบของพิทช์ที่ต่ำกว่า

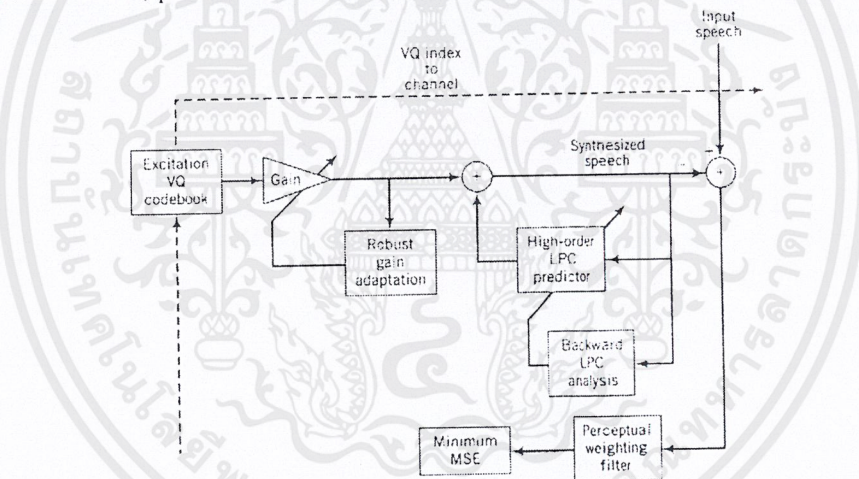
วิธีการอื่นๆของการแก้ปัญหาแบบซีอีแอลที นี้จะใช้ฟังก์ชันไซน์และโคไซน์เป็นส่วนหนึ่งของไม้ค้อน การใช้อนุกรมของไม้ค้อนแบบพิจารณาได้ (deterministic) ไม่เพียงปรับปรุงสมรรถภาพของการเข้ารหัสแต่จะสามารถลดการคำนวณที่ต้องการลงได้ด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.14.3) โลติเลย์ซีอีแอลพี (Low-Delay CELP)

ในทอริบายนี้ การแปลงแบบซีอีแอลพี จะมีบัฟเฟอร์ ไปกำกับความเร็ว 20 มิลลิวินาที และดังนั้นจะสามารถหน่วงเวลาเสียงพูด 20-40 มิลลิวินาที ซึ่งสามารถเป็นปัญหาของการใช้กับสายโทรศัพท์ถ้าไม่มีการใช้ระบบการตัดเสียงสะท้อน สำหรับการปรับปรุง ซึ่งการหน่วงเวลาถูกนำมาพิจารณา โดยวิธีการลดความล่าช้า ของซีอีแอลพี สามารถเป็นไปได้ การหน่วงเวลาทางเดียว ของลำดับที่ 2 มิลลิวินาที จะประสบความสำเร็จ วิธีการโลติเลย์ซีอีแอลพี จะมีความสามารถลดลงเป็นผลมาจากอัตราบิตที่มากขึ้นปกติต้องการ 16000 บิตต่อวินาที สำหรับคุณภาพทั้งหมดของสัญญาณเสียงพูด

การลดค่าความล่าช้าจะทำได้โดยใช้ช่องที่เล็กกว่า คือประมาณ 5 แซมเปิ้ล หรือ 0.625 มิลลิวินาที และส่วนเพิ่มเติมของอะแด็ปเทชันเพื่อทำนายคุณลักษณะของสัญญาณจะมีแนวโน้มอย่างไร ดังรูปที่ 2.20 การทำนายและอัตราขยายอะแด็ปเทชันถ่ายทอดจะมีส่วนการสังเคราะห์ มองที่ส่วนซึ่งเพิ่งผ่านไปของอนุกรมการกระตุ้น และลำดับของการกระตุ้น (excitation sequence) ที่ถูกกระตุ้น และจากการปรับปรุงตัวสร้างสัญญาณลิเนียร์พรีดิกทีฟสำรอง ซึ่งสร้างการประมาณใหม่ของสัมประสิทธิ์เกน (gain coefficient) และสเปกตรัมฟิลเตอร์ (spectral filter coefficient)



รูปที่ 2.20 การเข้ารหัสโลติเลย์ซีอีแอลพี (low-delay CELP coder)

ตัวเข้ารหัสจะทำโดยคำนวณการสังเคราะห์โดยประมาณของเกน และสัมประสิทธิ์ฟิลเตอร์แบบสเปกตรัม (spectral filter coefficient) และมันชดเชยโดยปรับลำดับของการกระตุ้น เพื่อสร้างเอาท์พุท ที่ออกแบบไว้ ดังนั้นข้อมูลเดียวที่ส่งข้ามช่องสัญญาณ คือ ดัชนีของโค้ดบุ๊กของการกระตุ้น (excitation codebook) ข้อดีคือเพียงแค่อุปกรณ์ปัจจุบันจะถูกใช้ การอะแด็ปเทชัน ทั้งหมดเกิดขึ้นบนพื้นฐานของอนุกรมการกระตุ้น ซึ่งจะถูกลดรหัสโดยตัวสังเคราะห์อย่างสิ้นมากหลังจากพวกมันได้ ในการเข้ารหัส ดังนั้น จะไม่มีเฟรมยาวๆ (long-term) เช่น 20 มิลลิวินาทีเกิดขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตั้งแต่การทำนายพิทช์จะต้องการอดีตอย่างน้อยที่ความยาวของคาบของพิทช์ มันจะทำให้เกิดคีย์ และดังนั้นไม่สามารถใช้ได้ ด้วยเหตุผลนี้ลำดับของฟิลเตอร์ลิเนียร์พรีดิคทีฟจะต้องเพิ่มขึ้นถึงประมาณ 50 เท่า

ซึ่งได้กล่าวไปก่อนหน้านี้แล้ว อนุกรมการกระตุ้นจะลดลงอย่างรุนแรงในระยะ 5 แคมเปิล ดัชนีของโค้ดบุ๊ก ยังคงเป็น 10 บิต แต่ 3 บิต จะถูกระบุไว้สำหรับการควบคุมแกน และ 7 บิตสำหรับรูปแบบการกระตุ้นลูกลับ เนื่องจาก 128 รูปแบบจะเป็นไปได้ อนุกรมการลูของเกาส์เซียน (Gaussian) ธรรมดาไม่สามารถให้การเลือกที่เพียงพอได้ ดังนั้นจะทำการเป็นเป็นกระบวนการพิจารณาแบบโคลสลูป (close-loop optimization) ซึ่งเลือกและพิจารณาในส่วนของเวคเตอร์ในโค้ดบุ๊กโดยใช้ข้อมูลตรวจสอบที่แท้จริงเพื่อลดการลดทอนที่คาดไว้

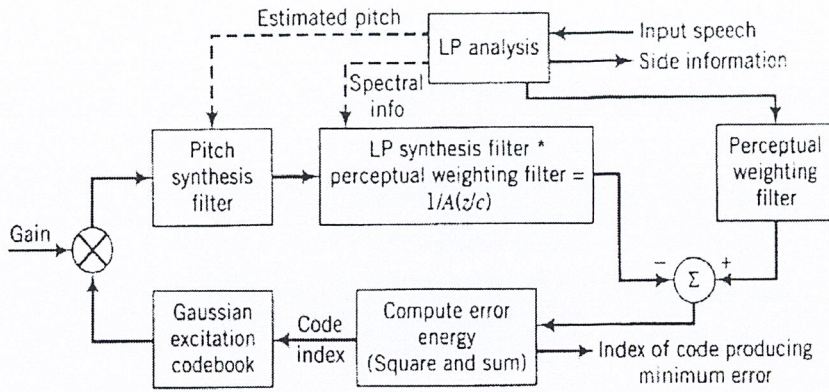
ความคิดพลาดของช่องสัญญาณจะเป็นสิ่งสำคัญในการเข้ารหัสเสียง และจะถูกจัดการด้วยการกระทำที่หน้าสนใจในอุปกรณ์นี้ ดัชนีของ โค้ดบุ๊ก จะถูกเข้ารหัสด้วยรหัสเกรย์ (gray code) เลขฐาน 2 ดังเช่นรหัสซึ่งแตกต่างด้วย 1 บิต ซึ่งส่วนมากเหมือนกัน (gray code คือ code เลขฐาน 2 ซึ่งอนุกรมของบิตแตกต่าง โดยบิตเพียงบิตเดียว ตัวอย่างเช่น รหัสเกรย์ แบบ 3 บิตจะเป็น $0=000$, $1=001$, $2=011$, $3=010$, $4=110$, $5=111$, $6=101$, $7=100$ ดังนั้นการเกิดความคิดพลาดของบิต จะถูกจัดการได้ดียิ่งขึ้น

อุปกรณ์อื่นๆของโลคัลเลย์ซีอีแอลพี (low-delay CELP) ก็ได้ถูกคิดค้นขึ้นโดยเฉพาะ อุปกรณ์ซึ่งใช้อะแดปทีฟโค้ดบุ๊ก (adaptive code book) (ในขั้นตอนที่ 2.15.6)

2.15 การลดเวลาการค้นหาในการเข้ารหัสแบบซีอีแอลพี (Reducing codebook search time in CELP)

ในตอนเริ่มแรกนั้น ซีอีแอลพี ห่างไกลจากการดำเนินการแบบในเวลาจริง (real time) โดยอุปกรณ์เริ่มแรก จะต้องการ 125 วินาที ของ คอมพิวเตอร์เครวัน (Cray-1) เวลาของหน่วยประมวลผลกลาง เพื่อสร้างสัญญาณเสียง 1 วินาที นอกจากการเปลี่ยนแปลงต่างๆไปในความเร็วหน่วยประมวลผลกลาง ยังมีการเสนอวิธีการในเวลาจริง อื่นๆอีก

การจำกัดขึ้นในซีอีแอลพี จะค้นหาผ่านพจนานุกรมโค้ดเวิร์ด (code-word) เพื่อหาอนุกรมการกระตุ้น ซึ่งลดค่าการลดทอนที่สังเกตได้ลดลง จากวิธีการที่แสดงจะต้องทำการส่งผ่านอนุกรมในโค้ดบุ๊กผ่านออลโพลไฟในทอิมพัลส์ฟิลเตอร์ (all pole finite Impulse (IIL) filter) 3 ตัว คือตัวฟิลเตอร์การสังเคราะห์พิทช์ (pitch synthesis filter) ฟิลเตอร์การสังเคราะห์ลิเนียร์พรีดิคทีฟ (LP synthesis filter) และ เพอร์เซปชวลเวตติ้งฟิลเตอร์ (perceptual weighting filter) ดังรูปที่ 2.7 มันจึงจำเป็นที่จะต้องคำนวณความคิดพลาดที่เกิดขึ้นร่วมกันของอนุกรม การค้นหาลักษณะนี้ค่อนข้างจำเป็นแม้แต่กับอุปกรณ์ดิจิทัลซิกแนลโปรเซสซิ่ง (DSP) ที่มีความเร็วสูงที่สุด หลายๆวิธีการของปัญหานี้ถูกนำเสนอขึ้น ซึ่งบางอย่างจะถูกพิจารณาตอนนี้



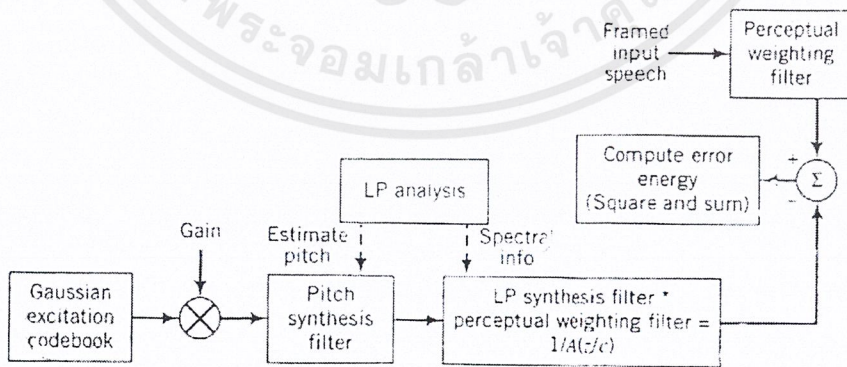
รูปที่ 2.21 ตัวเข้ารหัสแบบซีอีแอลพี หลังจากทำการลดระดับครั้งแรก

2.15.1) ฟิเตอร์ซิมพลิฟิเคชัน (Filter Simplification)

ด้วยการตรวจสอบการจัดการของฟิเตอร์ของซีอีแอลพี อย่างระมัดระวัง มันเป็นไปได้ที่จะตัดแปลงโครงสร้างปกติเพื่อลด IIR (filter operation) จาก 3 เป็น 1 โดยไม่มีหน่วยความจำ ทารอธิบายอย่างย่อๆ และผลจะถูกกล่าวถึงต่อไปในที่นี่ สำหรับรายละเอียดจะต้องอ่านจาก

การเปลี่ยนแปลงครั้งแรกเพื่อเคลื่อนเวดจ์ฟิเตอร์ (weighting filter) นี้ยอมรับได้จากเอาท์พุทของผลรวม (Σ) ยังอินพุทของมันจะเท่ากันเมื่อฟิเตอร์เป็นลิเนียร์เวดจ์ฟิเตอร์ (linear weighting filter) ที่ยอมรับได้ตามรูปแบบ $\frac{A(z)}{A(z/c)}$ ซึ่งเมื่อวางหลังจากฟิเตอร์สังเคราะห์พิทช์ (pitch synthesis filter) ซึ่งมีรูปแบบ $\frac{1}{A(z)}$ ทั้งสองจะถูกรวมกันเพื่อสร้างฟิเตอร์เดี่ยวรูปแบบ $\frac{1}{A(z/c)}$ ตามรูป

2.21 และ 2.22



รูปที่ 2.22 อธิบายรูปที่ 2.21

ต่อมาตัวกรองสังเคราะห์พิทช์ (pitch synthesis filter) สามารถเคลื่อนจากจุดของการค้นหา โค้ด บิ๊ก ซึ่งฟิเตอร์สังเคราะห์พิทช์ มีรูปแบบ $B(z) = 1-bz^{-T}$ ซึ่ง b คือเกนแฟกเตอร์ (gain factor) และ T คือ

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ ห้ามนำไปเผยแพร่โดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หา G ด้วยการกำหนดอนุพันธ์ของ E เพื่อเปรียบเทียบกับ G เป็น 0

$$0 = \frac{\partial E}{\partial G} = -2 \sum_n x_n g_n + 2G \sum_n g_n^2 \quad (2.9)$$

$$G = \frac{\sum x_n g_n}{\sum g_n^2} \quad (2.10)$$

ดังนั้น

$$E = \sum_n x_n^2 - 2 \frac{\sum x_n g_n}{\sum g_n^2} \sum x_n g_n + \left(\frac{\sum x_n g_n}{\sum g_n^2} \right)^2 \sum g_n^2 \quad (2.11)$$

$$E = \sum_n x_n^2 - \frac{(\sum x_n g_n)^2}{\sum g_n^2} \quad (2.12)$$

อย่างไรก็ตาม $\sum x_n^2$ ขึ้นอยู่กับอินพุตและ $\sum g_n^2$ คือสเกลแฟกเตอร์ขึ้นอยู่กับแอสพีซี สเตปครัลฟิลเตอร์และโค้ดเวิร์ค ดังนั้นเราจะคำนวณเบื้องต้น $\alpha_k = \sqrt{\sum g_n^2}$ และเลือกค่า g_n ที่ดีที่สุดเพื่อเพิ่ม $(\sum x_n g_n) / \alpha_k$

ดังนั้นผลผลิต ภัยคุกคามภายในของฟิลเตอร์โค้ดเวิร์ค และอนุกรม T/P ทุกระดับแบ่งโดยปัจจัยตรงข้ามที่เหมาะสมกับพลังงานในอนุกรมโค้ดบุ๊ก ซึ่งจะถูกทำให้เพิ่มขึ้น สเกลแฟกเตอร์นี้ไม่ต้อง อย่างไรก็ตามจะต้องคำนวณเป็นเวลาจริง ตั้งแต่สัมประสิทธิ์แอสพีซี สเตปครัล เป็นควอนไทซ์เวกเตอร์ และอนุกรมโค้ดบุ๊กถูกควอนไทซ์ การตรวจสอบในตารางเป็นวิธีการหนึ่งที่จะทำให้ได้มาซึ่งค่าที่ถูกต้องสำหรับดัชนีของโค้ดบุ๊กที่ให้มาและชุดสัมประสิทธิ์แอสพีซี วิธีการนี้ยังต้องการแต่ละโค้ดเวิร์คให้ถูกรองโดย $\frac{1}{[A(z/c)]}$

ก่อนที่อนุกรมภายในของมันจะถูกคำนวณ อย่างไรก็ตามจะสามารถลดการคำนวณอินเนอร์โพรดักต์ของโค้ดเวิร์คที่ไม่ได้ผ่านการกรอง ด้วยเวอร์ชันที่ผ่านการกรองของอนุกรมของอินพุตดังต่อไปนี้

จุดมุ่งหมายเพื่อเพิ่ม $P = \sum x_n g_n$ เมื่อ x_n คืออนุกรมอินพุต และ g_n คือผลตอบสนองของอินเวอร์สฟิลเตอร์ซึ่งไม่มีความจำ $\frac{1}{[A(z/c)]}$ ไปโค้ดเวิร์คปัจจุบัน c_n ในที่นี้ g_n จะสามารถเขียนเป็นคอนโวลูชันของ c และอิมพัลส์เรสponse f_n ของ

$$g_n = \sum_i c_i f_{n-i} \quad (2.13)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้น

$$P = \sum_{n=0}^{N-1} \left(X_n \sum_{i=0}^{N-1} C_i f_{n-1} \right) = \sum_{i=0}^{N-1} \left(C_i \sum_{n=0}^{N-1} X_n f_{n-1} \right) \quad (2.14)$$

หรือ

$$P = \sum_{i=0}^{N-1} C_i d_i$$

$$d_i = \sum_{n=0}^{N-1} X_n f_{n-1} \quad (2.15)$$

ดังนั้นเราจะคำนวณเพียงอินเนอร์โพรดักต์ของอินพุทของแบกเวิร์ดฟิลเตอร์ α ด้วยเวกเตอร์โค้ดเวิร์ดที่ไม่ผ่านฟิลเตอร์ c_i ดังนั้น การค้นหาโค้ดบุ๊กจะมีประสิทธิภาพลดอินเนอร์คอตโพรดักต์ของแต่ละโค้ดเวิร์ดด้วยเวกเตอร์หนึ่ง ซึ่งจะคำนวณเพียงเฟรมละครั้ง ซึ่งการคำนวณนี้จะเป็นไปได้ยากกว่าการกรองทั้งหมด ต่อ การเปรียบเทียบโค้ดเวิร์ด

2.15.3) การค้นหาโค้ดบุ๊กแบบมัลติโวลูชัน (Multiresolution Codebook Search)

การกำหนดระยะทางการคำนวณซึ่งลำดับโค้ดบุ๊กเข้ากันได้ดีที่สุด มีหลายวิธีที่จะค้นหาได้มีประสิทธิภาพมากกว่าวิธีการค้นหาทั้งหมด การค้นหาในโค้ดบุ๊กหลายชั้น (multi stage) นำมาซึ่งการดำเนินการเบื้องต้น, การคำนวณเข้มข้นน้อยลง และดังนั้นความถูกต้องของการหาค้นหา ในฐานข้อมูลเมื่อจำกัดจำนวนของอนุกรมสมาชิกที่เป็นไปได้ลง จากขนาดทั้งหมดของฐานข้อมูลไปเป็นส่วนหนึ่งของซบเซตที่เหมาะสมมากขึ้น ซึ่งจะค้นหาซ้ำและแบบที่ถูกต้อง วิธีนี้ยังถูกใช้อย่างกว้างขวางในการทำอื่น ๆ เช่น อิมเมจโพรเซสซิ่ง

ในเวอร์ชันหนึ่งของการค้นหาโค้ดบุ๊กหลายชั้น เบื้องต้นระยะเปรียบเทียบจะแสดงที่การตัดสินใจชั่วคราวที่ตั้งต้นโดยโลพาสฟิลเตอร์และ ทำการเดซิเมต (decimating) ด้วย 5 และ 1024 การเข้าโค้ดบุ๊กและอนุกรมเอ็กไซเตชันที่ออกแบบไว้ คะแนนจากการเปรียบเทียบการลดเหล่านี้จะถูกใช้เฉพาะสมาชิกที่ดีที่สุด 70 ตัว ซึ่งเปรียบเทียบที่การตัดสินใจแบบเต็ม

ผู้เขียนรายงานว่าความเร็วโดยแฟกเตอร์ของ g จะส่งเร็วด้วยพารามิเตอร์ โดยแฟกเตอร์เดซิเมตชัน (Decimation factor) และขนาดซบเซตเล็กลงถูกเลือกโดยการลอกเลียนแบบ ดังนั้น 99 เปอร์เซ็นต์ ของเวลา การค้นหาที่น้อยลงจะไปทำลายอนุกรมที่ถูกต้อง ใช้เริ่มแรก 2 ระยะที่เดซิเมชัน (decimation) ของ 5 และ 2 อนุญาตให้เพิ่มความเร็วด้วยแฟกเตอร์ของ 13 ด้วยความถูกต้องยังคงเป็น 99 เปอร์เซ็นต์

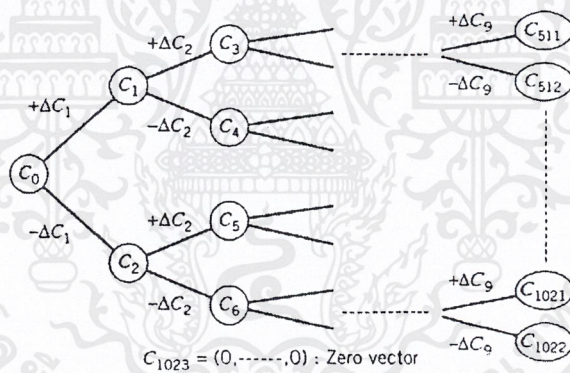
การลดลงของค่าโค้ดบุ๊กสามารถคำนวณได้โดยไม่ต้องพึ่งเครื่องคอมพิวเตอร์หลักจะอิสระจากสัญญาณเสียงพูด ผลนี้จะเพิ่มการเก็บที่ต้องการเล็กน้อย แต่มันจะไปปรับปรุงการคำนวณเวลาการชดเชย

2.15.4) การทำลายอนุกรมพาเซี่ยล (Partial Sequence Elimination)

อีกวิธีการหนึ่งของการค้นหาแบบมัลติสแตจจะไม่เปรียบเทียบกับอนุกรมโค้ดบู้กทั้งหมดที่ถูกลดลง แต่เปรียบเทียบแค่ส่วนของอนุกรมที่ผลการตัดสินใจทั้งหมดในวิธีนี้ k แซมเปิ้ลแรกของการออกแบบอนุกรมการกระตุ้นจะเปรียบเทียบกับ k แซมเปิ้ลแรกของแต่ละเวกเตอร์โค้ดบู้ก ซึ่งจะใช้วิธีนี้ในการลดความกว้างของระดับที่ 2 ซึ่งเป็นการค้นหาเพิ่มความยาวโดยตัวเลขที่ระบุคงที่บางค่าของโค้ดบู้กส่วนใหญ่ นี่คือการรายงานเพื่อลดจำนวนของตัวคูณ และตัวกระทำที่เป็นเลขคี่ (odd operation) 75 เปอร์เซ็นต์

2.15.5) โค้ดบู้กเดลต้าแบบรูปต้นไม้

ฐานข้อมูลทรี-สตรีคเจอร์มักจะใช้กับการค้นหาที่มีประสิทธิภาพ ดังเช่น พ่อ แม่ และลูก ๆ จะต้องมีความสัมพันธ์ในความหมายบางอย่าง ดังนั้น ถ้าค่าในคอลเวเนชันนอล ลินีเยร์ โค้ดบู้ก (conventional linear codebook) จะถูกกำหนดเป็น โหนดของทรีจะไม่มีประโยชน์ในการคำนวณ อนุกรมการสุ่มแบบเกาส์เซียน (Gaussian) จะไม่ปรากฏตัวของมันเป็นลำดับ แต่โดยการใช้ชุดขอบเขตของอนุกรมเรนคอมเพื่อสร้างรากฐานของชุดที่ใหญ่ขึ้น ชุดที่ใหญ่กว่าจะสามารถเข้ารหัสใน ทรี ซึ่งความสูงเป็นสัดส่วนกับจำนวนเวกเตอร์เบื้องต้นสามารถทำได้วิธี ทรี-สตรีคเจอร์ เดลต้า โค้ดบู้ก



รูปที่ 2.24 โครงสร้างต้นไม้ของเดลต้าโค้ดบู้ก

ทรี สตรีคเจอร์ เดลต้า โค้ดบู้ก สร้างสัดส่วนขึ้นจากจำนวนที่จำกัด L ของเวกเตอร์เริ่มแรกก่อนทำการคำนวณ Δc_0 ผ่าน Δc_{L-1} (ดังรูป 2.24) ลำดับของโค้ดบู้กสามารถหาได้โดยเริ่มต้นด้วยโค้ดเวกเตอร์ที่เรารู้ (root) และเพิ่มหรือลด เดลต้าโค้ดบู้กเวกเตอร์จะถูกพบระหว่างการสำรวจเส้นทางไปยังโหนดที่ต้องการ (ซึ่ง $c_0 = \Delta c_0$) ตัวอย่างเช่น c_5 คือ $\Delta c_0 - \Delta c_1 + \Delta c_2$ ในกรณีนี้เวกเตอร์ $2^L - 1$ ในโค้ดบู้กจะถูกสร้างจากเดลต้าโค้ดบู้กเวกเตอร์ $L, \Delta c_0 \dots \Delta c_{L-1}$

ทรี-สตรีคเจอร์ เดลต้า โค้ดบู้ก ไม่เพียงแต่ลดหน่วยความจำที่ต้องการแต่ยังลดการคำนวณที่ต้องการอีกด้วย โดยอันแรกจะลดเป็นส่วนของ 100 ซึ่งอันหลังจะลดด้วยส่วนของ 70 เมื่อเทียบกับโค้ดบู้กแบบธรรมดา จากการเพิ่มขึ้นโดยธรรมชาติของ ทรี-สตรีคเจอร์ค่าของโค้ดบู้กที่เหมาะสมโดยตรงที่สุดจะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แสดงใน โค้ดบุ๊ก เดลต้า Δ มากกว่าใน โค้ดบุ๊กทั้งหมดซึ่งจะลดจำนวนของการกรองและคอร์รัเลชัน (หรือ คออดโพรคัตของเวกเตอร์ในคำศัพท์ที่กลับกัน) เพื่อจะเป็นเชิงเส้นกับ L ก่อนหน้าที่ค่าพลังงานของ อนุกรมของโค้ดบุ๊กนี้ถูกกรองจะต้องถูกคำนวณ และใช้เพื่อกลับระดับภายในโพรคัต ซึ่งจะถูกคำนวณไว้ พอเหมาะกับกำลังสองของ L อย่างไรก็ตาม จะถูกทำลายถ้าตารางลูก-อัพถูกใช้ดังที่พบก่อนหน้านี้ในส่วน ที่ 2.15.2

ทางเลือกของรหัสเวกเตอร์ของ เดลต้า จะถูกคำนวณโดยไม่ต้องพึ่งเครื่องคอมพิวเตอร์หลักโดยการวิเคราะห์สัญญาณข้อมูลเสียงเป็นจำนวนมาก และเลือกเวกเตอร์ตั้งฉากหลัก ซึ่งทำให้เกิดการลดทอน ตั้งสุดเมื่อใช้รหัสตรวจสอบข้อมูล การทำซ้ำวิธีการดังกล่าวจะทำให้ เดลต้าโค้ดเวกเตอร์ปรับปรุงขึ้น

จะต้องถูกบันทึกไว้ว่า จากวิธีการแบบทริสตรัคเจอร์ของโค้ดบุ๊กและความจริงว่าตัวแสดงโค้ดบุ๊ก ที่มีอยู่จะอยู่ทั้งส่วนของไปไม้ของทริ และในโหนดภายในบางส่วนของเดลต้าโค้ดเวกเตอร์จะถูกใช้บ่อยถ้า ส่วนอื่น โดยเฉพาะเดลต้าโค้ดเวกเตอร์ที่ตัวชี้ต่ำจะถูกใช้มากกว่าที่ตัวชี้สูง ๆ (hing index) กรณีที่ต่ำสุดคือ Δ_{c_0} ซึ่งเป็นองค์ประกอบของเวกเตอร์ทั้งหมดในโค้ดบุ๊กและ $\Delta_{c_{L-1}}$ ซึ่งเป็นเพียงองค์ประกอบของครึ่ง หนึ่งของเวกเตอร์โค้ดบุ๊ก ดังนั้นการจัดลำดับของเดลต้าโค้ดบุ๊กเวกเตอร์จะสำคัญมาก เดลต้าเวกเตอร์ซึ่ง สอดคล้องมากกว่ากับการแจกแจงอนุกรมการกระตุ้นปัจจุบัน ควรจะจัดอยู่ใกล้รากของต้นไม้ และเดลต้า เวกเตอร์ซึ่งเพียงแค่มัผัสกับอนุกรมการกระตุ้นควรจัดอยู่ในระดับที่ต่ำกว่า ซึ่งจะกระตุ้นให้เกิดเทคนิค เรียกว่า การเรียง เดลต้า เวกเตอร์ (Delta vector sorting) ซึ่งลำดับของเดลต้าเวกเตอร์เป็นกุญแจในการจัด บรรจุสเปกตรัมของอนุกรมการกระตุ้นที่ออกแบบไว้ เดลต้าเวกเตอร์จะถูกกรองผ่านแอลพีซี สเปกตรัม เชปปีงฟิลเตอร์ และถูกจัดเรียงโดยน้ำหนักพลังงาน (weighted energy) เดลต้าเวกเตอร์ซึ่งมีค่าพลังงานที่ มากกว่าจะสามารถใช้ได้อย่างมีประสิทธิภาพที่สุดเพื่อสร้างอนุกรมการกระตุ้นที่ออกแบบไว้ และจะถูกจัด ความใกล้ชิดที่สุดกับรากของทริ ดังนั้นสัมประสิทธิ์การกรองแอลพีซี (LPC spectral shaping filter) จะเป็นทั้ง ตัวเข้ารหัส และตัววิเคราะห์ ไม่ต้องส่งข้อมูลเพิ่มเติม

2.15.6) อดแอฟทีฟโค้ดบุ๊ก (Adaptive Codebook)

ในวิธีการที่ผ่านมาของซีอีแอลพี การใช้โค้ดบุ๊กจะทำการคำนวณก่อนและยังคงค่าคงที่จนกระทั่ง ทำการเข้ารหัสและถอดรหัสออกเรียบร้อยแล้วซึ่งการทำลักษณะนี้จะไม่ตรงตามคุณลักษณะที่ว่า สัมประสิทธิ์ของสัญญาณเสียงเปลี่ยนแปลงตามเวลา ดังนั้นจะเป็นประโยชน์ที่จะมีโค้ดบุ๊กซึ่งเปลี่ยนแปลง ตามสัญญาณเสียง วิธีการนี้ถูกใช้ในวิธีการใหม่ ๆ ของซีอีแอลพี หลักการนี้จะถูกอธิบายสำหรับการ กระตุ้นแบบ เชลฟ์-เอกซ์แทรกท์ (self-extracted) ในรูปแบบธรรมดาส่วนมากของอดแอฟทีฟโค้ดบุ๊กมี บรรจุโค้ดบุ๊กจากการสังเคราะห์สัญญาณที่ถูกกระตุ้นที่ผ่าน ๆ มา ดังนั้นคิดว่าโค้ดบุ๊กไม่ใช่ชุดของ k คิสทริทซึ่งอนุกรม N แซมเปิ้ลโค้ดเวกเตอร์ซึ่งถูกเลือกมาแต่ยังบันทึกประวัติเชิงเส้นของการกระตุ้น kN ที่ ผ่านมา (โดยปกติ $kN+N-1$) การเลือกโค้ดเวกเตอร์เกี่ยวข้องกับการพิจารณาอดีตของ kN แซมเปิ้ล เมื่อ เลือกแซมเปิ้ลของ N ต่อเนื่องสำหรับใช้อีกครั้งเป็นโค้ดเวกเตอร์ปัจจุบัน ตั้งแต่ดัชนีของโค้ดปัจจุบันจาก 0 ถึง $kN-1$ จะต้องการ $\log_2(k)+\log_2(N)$ บิต ดังนั้น kN จะถูกเลือกเป็นปกติประมาณ 128 ซึ่งด้วย N ยังคง ประมาณ 40 ดังนั้นโค้ดบุ๊กทั้งหมดจะค่อนข้างเล็ก โดย k จะไม่มีความสำคัญโดยโค้ดเวิร์ดจะเป็นทุกส่วน เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของอนุกรมต้องเนื่องหนึ่ง ดัชนีของโค้ดบางครั้งจะเรียกว่า แล็ก (lag) หรือแทนด้วย L ตั้งแต่มันแสดงจำนวนของเวลาที่กลับไปนำอนุกรมการกระตุ้นกลับคืนมาสู่สภาพเดิม

อะแดปทีฟโค้ดบุ๊กก็มีประโยชน์ทางการคำนวณโดยไม่ต้องการเพอร์เซปชวลเวตติงฟิลเตอร์ จากการที่อนุกรมการกระตุ้นได้ถูกย่อมาจากอนุกรมการกระตุ้นเมื่อเร็ว ๆ นี้ มันจะเป็นค่าเวตติงที่ถูกต้องที่ยอมรับได้ ตั้งแต่การกระตุ้นครั้งที่แล้วเป็น แม้ว่าข้อสรุปนี้จะดูคววน แต่ข้อพิสูจน์และการทำตามจะมีบางเครื่องมือกรองและวัดลำดับอนุกรมการกระตุ้นก่อนเพิ่มมันเข้าไปในโค้ดบุ๊ก

อะแดปทีฟโค้ดบุ๊กจะดีสำหรับอนุกรมการกระตุ้นเสียงแบบสเตตีสเตท (steady-state) แต่จะไม่เหมาะสมนักสำหรับการเริ่มต้นเสียง (voicing onset) หรือเสียงสั้นในคำคอ เพื่อแก้ไขข้อบกพร่องดังกล่าว การพิจารณาสแตติกโค้ด (static code) ก่อนจะถูกรวมอยู่ในอีกส่วนหนึ่งของโค้ดบุ๊กมันควรจะบันทึกไว้ว่าอนุกรมเหล่านี้สามารถเป็นทางผ่านย่อยของฟิลเตอร์เพอเซปชวลเวตติงด้วยเช่นกัน เมื่อมีการสุ่มมากขึ้นในธรรมชาติและจึงท้วงว่าการทำเสียงต่าง ๆ ให้เป็นรูปตามต้องการจะไม่จำเป็น

2.15.7) โค้ดบุ๊กแบบลิเนียร์คอมบิเนชัน (Linear Combination Codebook)

แม้ว่าการสร้างโค้ดบุ๊กจากการรวมเชิงเส้น เวกเตอร์พื้นฐานจะกล่าวไปแล้วในการประยุกต์กับการเข้ารหัสโค้ดเวิร์คแบบทรีในตอนที่ 2.15.5 และอีกเครื่องมือหนึ่งที่จะอธิบายต่อไปนี้จะแตกต่างกันอย่างสิ้นเชิง เพื่อประกันการกระทำเฉพาะรายการจัดโค้ดบุ๊กที่อธิบายนี้จะใช้ในระบบซึ่งอธิบายในส่วนถัดไป

แทนที่จะใช้โครงสร้างแบบทรีเพื่อรวมเวกเตอร์เบื้องต้น, จะสามารถใช้เลขฐาน 2 แทนเพื่อควบคุมวิธีการรวมกันของเวกเตอร์พื้นฐานพิจารณา k เวกเตอร์พื้นฐานใช้รูปแบบ 2^k เข้าเอนทรีโค้ดบุ๊ก ใน 2^k เอนทรีโค้ดบุ๊กนี้ แต่ละโค้ดบุ๊กเอนทรีจะเป็นการรวมเชิงเส้นของเวกเตอร์พื้นฐาน ซึ่งเวกเตอร์พื้นฐานจะจัดขนาดเป็น -1 หรือ 1 เท่านั้น ดังนั้นจะง่ายต่อการดู วิธีการที่ 2^k ตัวจะถูกสร้างจาก k บิตโค้ดบุ๊กอินเดกซ์ (k -bit codebook index)

2.16 การวิเคราะห์สัญญาณเสียงในช่วงเวลาสั้นๆ

เนื่องจากว่าสัญญาณเสียงเป็นสัญญาณที่แปรตามเวลา มีการแปรเปลี่ยนที่ไม่แน่นอน เช่น ในขณะที่พูดซ้ำๆ รูปร่างของโพรงเสียง รวมทั้งลักษณะรูปแบบของการกระตุ้นอาจจะไม่มีการเปลี่ยนแปลง ในช่วงเวลาที่นานที่สุดประมาณ 200 มิลลิวินาที แต่ในขณะที่พูดอย่างรวดเร็ว อาจจะมีช่วงที่ไม่มีการเปลี่ยนแปลง สั้นมาก คือประมาณ 80 มิลลิวินาที ก็ได้

ดังนั้นเทคนิคในการวิเคราะห์เสียงพูดส่วนใหญ่แล้วจะสมมติให้สัญญาณเสียงมีคุณสมบัติที่เปลี่ยนแปลงสัมพันธ์กับเวลาอย่างเชื่องช้า นั่นก็คือ เราจะต้องแบ่งทำการวิเคราะห์หาพารามิเตอร์ของสัญญาณเสียงที่พูดในช่วงเวลาสั้นๆ เหมือนมองผ่านช่องแคบๆ ที่เรียกว่า ซอร์ทไทม์วินโดว์ (short-time window) เมื่อเทียบตามเวลาที่เสียงอาจจะมีการเปลี่ยนแปลงได้ เพื่อจะได้มองเห็นเหมือนกับว่า เราหาพารามิเตอร์นั้นๆ ได้มาจากสัญญาณเสียงที่อยู่ภายในช่องแคบๆ และมีความเสถียรภายในช่วงเวลาสั้นๆ

เทคนิคส่วนใหญ่จะกำหนดให้พารามิเตอร์ได้มาจากค่าเฉลี่ยจากการวิเคราะห์ภายในช่วงเวลาแคบๆ นั้น สำหรับกรณีที่ต้องพิจารณาพารามิเตอร์ต่างๆ ที่มีการเปลี่ยนแปลง ก็จะมีการแบ่งสัญญาณเสียงเอกสารนี้เป็นเอกสารที่ส่งวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ออกเป็นช่องหลายๆช่อง หรืออาจจะเรียก กรอบการวิเคราะห์ (analysis frame) ดังนั้นพารามิเตอร์ต่างๆจะสามารถหาได้ทันทีเพียงพอที่จะติดตามการเปลี่ยนแปลงของสัญญาณ สำหรับในช่วงที่สัญญาณมีการเปลี่ยนแปลงช้า อาจกำหนดให้ช่องแคบมีขนาดใหญ่ประมาณ 100 มิลลิวินาที แต่ในทางตรงข้ามถ้าสัญญาณมีการเปลี่ยนแปลงเร็ว ก็ต้องใช้ช่องแคบที่มีขนาดเล็กมากๆ ประมาณ 5-10 มิลลิวินาที เพื่อป้องกันการสูญหายของรายละเอียดของสัญญาณถัดไป

2.17 รูปแบบของช่องแคบ (windows)

การกำหนดขนาดของช่องแคบที่ใช้ขึ้นอยู่กับ

(1) ช่องจะต้องสั้นพอ ที่จะให้คุณสมบัติของเสียงที่กำลังพิจารณาไม่มีการเปลี่ยนแปลง อย่างมีนัยสำคัญในช่องแคบนั้น

(2) ช่องแคบจะต้องยาวพอที่จะทำให้การจัดเตรียมตัวอย่างของเสียง เพื่อจะนำไปคำนวณหาพารามิเตอร์ให้ได้ตามต้องการ อย่างเช่น ในกรณีที่มีสัญญาณรบกวนเข้ามาแทรกอยู่ในบางช่วงในสัญญาณเสียงด้วย ถ้าเราเลือกใช้ขนาดช่องแคบที่มีขนาดใหญ่กว่า เมื่อทำการหาค่าพารามิเตอร์โดยเฉลี่ย กฎจะทำให้ส่วนประกอบของสัญญาณรบกวนถูกตัดทิ้ง หรือมองข้ามไป

(3) ช่องแคบที่เหมาะสม ไม่ควรสั้นเกินกว่าช่วงหนึ่งคาบของสัญญาณเสียง ในช่วงที่กำลังวิเคราะห์ เงื่อนไขนี้จะมีผลต่อค่าเฟรมเรท (frame rate) ซึ่งก็คือจำนวนครั้งต่อวินาทีที่ทำการวิเคราะห์สัญญาณเสียง โดยการขยับช่องแคบไปเป็นคาบๆตามแกนเวลา ตามปกติ เฟรมเรทจะมีค่าประมาณ 2 เท่าของส่วนกลับของขนาดช่องแคบ นั่นก็คือช่องแคบถัดๆกัน ไปจะมีการซ้อนทับกัน 50 เปอร์เซ็นต์

การนำฟังก์ชันของช่องแคบที่มีช่วงขนาดจำกัด $w(n)$ มาคูณเข้ากับสัญญาณ $s(n)$ จะทำให้ได้กลุ่มตัวอย่างของเสียงพูดที่ถูก ที่ถูกกำหนดน้ำหนักให้แปรไปตามรูปร่างของช่องแคบรูปแบบของช่องแคบที่ง่ายที่สุดคือ กรอบแบบสี่เหลี่ยม (rectangular window) ซึ่งมีนิยามดังนี้

$$w(n) = \begin{cases} 1, & n = 0, 1, \dots, N-1 \\ 0, & n \text{ อื่นๆ} \end{cases}$$

ในสมการนี้คือ การกำหนดช่องของการวิเคราะห์ให้มีจำนวนตัวอย่าง N ตัวอย่าง

รูปแบบของช่องแคบในลักษณะนี้ก็มีฟังก์ชันของช่องแคบหลายลักษณะด้วยกัน ตัวอย่างเช่น แบล็คแมน (Blackman), บาร์เลต (Barlett), แฮมมิง (Hamming), แฮนนิง (Hanning) เป็นต้น โดยฟังก์ชันของช่องแคบที่นิยมใช้กันมาก ในการวิเคราะห์สัญญาณเสียงก็คือฟังก์ชันแฮมมิง (Hamming) ซึ่งมีรูปร่างตามลักษณะของ (cosine pulse) มีนิยามดังนี้

$$w(n) = \begin{cases} 0.54 - 0.46 \cos(2\pi n)/(N-1), & n=0, 1, \dots, N-1 \\ 0, & n \text{ อื่นๆ} \end{cases}$$

2.18 การหาค่าพลังงานของสัญญาณเสียง

พลังงานของสัญญาณ เป็นตัวแทนอันหนึ่งที่เรามักจะนำมาใช้ในการวิเคราะห์ลักษณะต่างๆของสัญญาณต่างๆไป โดยพลังงานของสัญญาณ $s(n)$ ใดๆ ที่แปรตามเวลาสามารถนิยามได้ว่า

$$Energy = \sum_{n=-\infty}^{\infty} s^2$$

แต่สำหรับสัญญาณเสียงซึ่งเป็นสัญญาณที่แปรเปลี่ยนอยู่ตลอด ไม่มีเสถียรภาพตามเวลา เราจะต้องแบ่งสัญญาณออกมาพิจารณาเป็นช่วงเล็กๆ ตามแกนเวลา เรียกว่า แบ่งออกเป็นเฟรมเช่น เฟรมละประมาณ 10-30 วินาทีหรือเฟรมละ 100 ตัวอย่างเป็นต้น ดังนั้นก็สามารถหาพลังงานของเสียงในแต่ละเฟรมได้

$$Energy_l(m) = \sum_{n=0}^{N-1} s^2(n)$$

โดยที่ l แทนลำดับของเฟรมข้อมูลเสียง , $l = 0, 1, 2, 3, \dots, L$

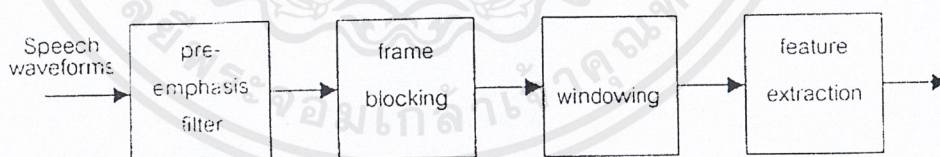
N จำนวนข้อมูลเสียงในแต่ละเฟรม

การวัดค่าพลังงาน ดังในสมการ นั้นมีข้อจำกัดตรงที่ว่า มันจะมีความไวต่อสัญญาณที่มีขนาดใหญ่ เนื่องจากเราใช้วิธียกกำลังสองค่าของสัญญาณอินพุต ดังนั้นการแก้ปัญหาอย่างหนึ่งก็คือการวัดพลังงานของเสียงโดยใช้สมการดังนี้

$$Energy_l(m) = \sum_{n=0}^{N-1} |s(n)|$$

2.19 การวิเคราะห์เสียงเบื้องต้น (Speech Pre-Processing)

ในการวิเคราะห์สัญญาณเสียง เราจำเป็นต้องมีการเตรียมข้อมูลการวิเคราะห์ก่อน ซึ่งขั้นตอนในการเตรียมสัญญาณเข้าเป็นดังรูปที่ 2.25



รูปที่ 2.25 แสดงขั้นตอนการเตรียมสัญญาณในการวิเคราะห์

2.20 การเตรียมสัญญาณเสียงในการวิเคราะห์ มีขั้นตอนดังต่อไปนี้

2.20.1 การพรีเอมฟาสีส (Preemphasis)

เนื่องจากสัญญาณเสียงพูดของมนุษย์ จะมีองค์ประกอบส่วนใหญ่อยู่บริเวณความถี่ต่ำ เมื่อเทียบกับความถี่ปฏิบัติงานไม่เกิน 5 กิโลเฮิรตซ์ ดังนั้น เพื่อให้อัตราส่วนของสัญญาณต่อสัญญาณเสียงรบกวน (SNR) มีค่าคงที่ตลอดช่วงความถี่ปฏิบัติงานจึงทำการพรีเอมฟาสีส โดยเน้นให้ความถี่สูงมีขนาดสูงขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

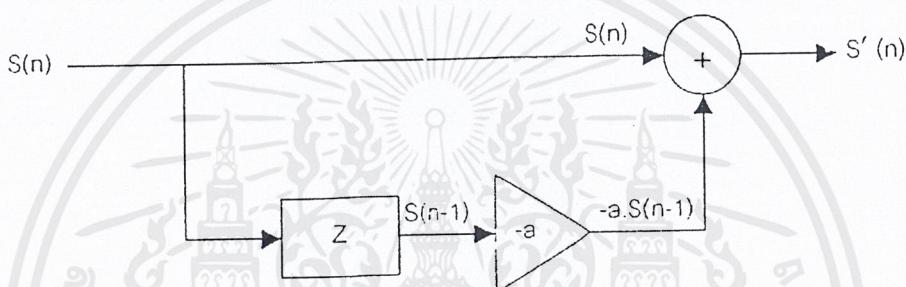
การพรีอิมฟาสซิส คือ การกรองสัญญาณด้วยวงจรกรองความถี่สูงผ่าน (High Pass Filter) ซึ่งนิยมใช้วงจรอันดับหนึ่ง ซึ่งตัวกรองเชิงเลขแบบง่ายที่สุด มีรูปแบบสมการดังนี้

$$y(n) = x(n) - ax(n-1) \quad (2.16)$$

มีฟังก์ชันถ่ายโอนเป็น

$$H(z) = 1 - az^{-1} \quad ; \quad 0.9 < a < 1.0 \quad (2.17)$$

สมมติสัญญาณเดิมเป็น $S(n)$ เมื่อผ่านวงจรกรองแล้วจะเป็น $S'(n)$



รูปที่ 2.26 วงจรกรองความถี่สูงผ่าน

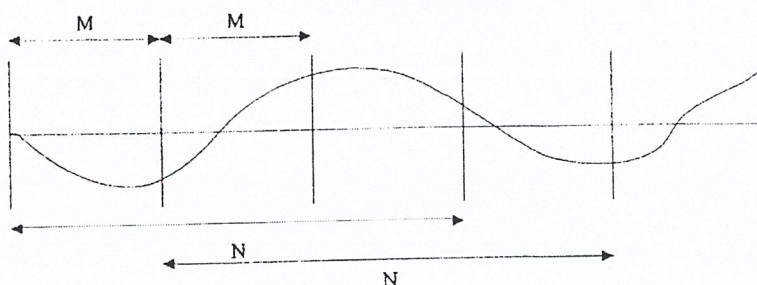
จะได้ว่า

$$S'(n) = S(n) - a.S(n-1) \quad (2.18)$$

ยิ่งค่า a เข้าใกล้ 1 เท่าใด ความถี่สูงจะถูกขยายขึ้นมากเท่านั้น ค่า a ที่นิยมสำหรับใช้ในการหาพารามิเตอร์ของแอลพีซี คือค่า $15/16 = 0.9375$

2.20.2) การแบ่งช่วงสัญญาณ (Frame Blocking)

สัญญาณที่ผ่านการพรีอิมฟาสซิสแล้ว ถูกตัดมาวิเคราะห์ทีละเฟรม เฟรมละ N ตัวอย่างสัญญาณในการวิเคราะห์ทีละช่วงของ N ตัวอย่าง สัญญาณนั้นจะวิเคราะห์ โดยเลื่อนเป็นระยะ M ช่วงสัญญาณ จนหมดสัญญาณเสียงที่นำมาวิเคราะห์ ดังรูปที่ 2.6 ซึ่งระยะ N ที่เหมาะสมจะต้องมีค่ามากกว่า M จะทำให้วิเคราะห์สัญญาณได้แม่นยำ แต่ถ้าค่า N น้อยเกินไปการวิเคราะห์จะล่าช้า

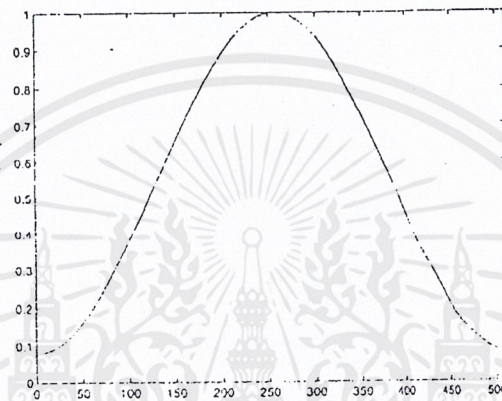


รูปที่ 2.27 การแบ่งช่วงของสัญญาณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไมอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เนื่องจากที่เราตัดสัญญาณเสียงมาวิเคราะห์ทีละเฟรมทำให้ที่ขอบเฟรมเกิดความไม่ต่อเนื่องของสัญญาณ ถ้ามองในโดเมนความถี่สูง ดังนั้นเพื่อจะลดองค์ประกอบทางความถี่สูงเหล่านี้ เราจะต้องคูณด้วยฟังก์ชันวินโดว์เพื่อลดความไม่ต่อเนื่องของสัญญาณที่ขอบและไม่ทำให้สเปกตรัมของสัญญาณในช่วงความถี่ต่ำเปลี่ยนแปลงไปมากนัก ในที่นี้จะใช้ฟังก์ชันวินโดว์แฮมมิง (Hamming window function) ซึ่งนิยามโดยสมการ

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right) \quad \text{เมื่อ } n = 0, 1, 2, \dots, N-1 \quad (2.19)$$



รูปที่ 2.28 แสดงวินโดว์แบบแฮมมิง

ในการวิเคราะห์เสียงโดยใช้ฟังก์ชันวินโดว์ จะพบว่าสัญญาณที่ผ่านการกรองโดยใช้วินโดว์นั้น จะมีการแกว่งขึ้นลงมากขึ้นกับช่วงเวลาของวินโดว์ (ความกว้างของวินโดว์) คือ ถ้าช่วงของวินโดว์สั้น จะมีการแกว่งขึ้นลงอย่างรวดเร็ว และถ้าช่วงของการวินโดว์ยาวจะมีการแกว่งขึ้นลงอย่างช้าๆ ดังนั้นการเลือกช่วงเวลาของการวินโดว์ ต้องให้อยู่ในช่วงที่เหมาะสม คือไม่ให้เอาที่พูของสัญญาณไม่แกว่งช้าหรือเร็วจนเกินไปอยู่ในช่วงระหว่าง 10 – 30 มิลลิวินาที

เมื่อคูณกับฟังก์ชันแล้วจะได้

$$x'(n) = w(n).x(n) \quad (2.20)$$

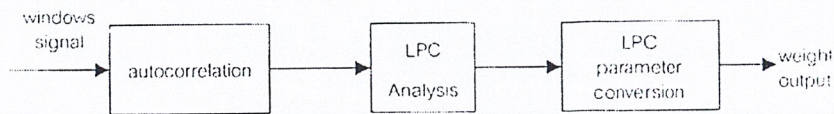
2.21 การวิเคราะห์หาคุณลักษณะเสียง (Feature extraction)

การวิเคราะห์หาคุณลักษณะของเสียง เป็นเทคนิคในการแปลงข้อมูลที่มีอยู่มากมาย ให้เป็นส่วนเล็กๆ ซึ่งส่วนเล็กๆ นี้จะแสดงคุณสมบัติของคลื่นเสียงนั้นๆ โดยจะใช้วิธีการป้อนสัญญาณผ่านวงจรกรอง (Filter) และการประเมินเชิงเส้น (Linear Predictive Coding)

หลักการพื้นฐานของการประเมินเชิงเส้น คือ การประเมินเสียงจากผลรวมเชิงเส้นของสัญญาณเสียงในอดีต โดยอาศัยหลักเกณฑ์กำลังสองของสัญญาณความคลาดเคลื่อน มีค่าต่ำสุดในการหาสัมประสิทธิ์การประเมินเชิงเส้น กล่าวคือหลังจากการทำพีริออดฟาซิสและวินโดว์ครบทุกตัวในหนึ่งเฟรมแล้ว จะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทำการคำนวณออโตคอร์รีเลชัน คือที่จะทำการหาค่าสัมประสิทธิ์แอลพีซี และอัตราขยาย G มีขั้นตอนแสดงดังรูปที่ 2.29



รูปที่ 2.29 แสดงขั้นตอนในการวิเคราะห์หาคุณลักษณะของเสียง

2.22 การคำนวณหาออโตคอร์รีเลชัน (Autocorrelation)

สมมติว่าสัญญาณเดิมเป็น $S(n)$ การประมาณ ค่าสัญญาณเป็น $S'(n)$ ดังนั้นสามารถอธิบายการประมาณเชิงเส้นด้วยสมการดังนี้

$$S'(n) = \sum_{k=1}^p \alpha_k S(n-k) \quad (2.21)$$

เมื่อ α_k เป็นค่าคงที่ เรียกวิธีการประเมินเชิงเส้นอันดับที่ p โดยมีเงื่อนไขว่า ค่า α_k ที่ใช้ในการประมาณที่จะต้องทำให้ผลรวมของกำลังสองของความคลาดเคลื่อน $\{S(n) - S'(n)\}^2$ มีค่าน้อยที่สุด ซึ่งใช้ในการประเมินเชิงเส้น (Autocorrelation Method) หรือ วิธีออตตัมพันธ์

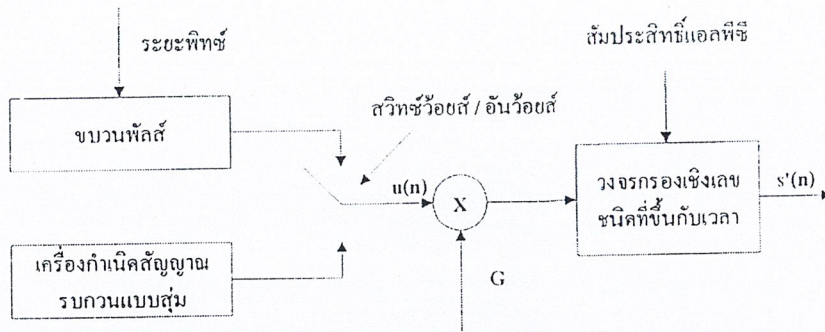
การคำนวณออโตคอร์รีเลชัน เป็นวิธีการหาสัมประสิทธิ์แอลพีซี โดยฟังก์ชันออโตคอร์รีเลชัน ซึ่งเป็นการเปรียบเทียบสัญญาณกับสัญญาณของตัวเองที่ถูกเปลี่ยนไปตามแกนเวลา ที่ใช้วิธีนี้เนื่องจากเป็นการคำนวณที่มีการแก้สมการที่น้อยกว่าวิธีอื่นๆ และมีความแน่นอนในด้านเสถียรภาพ อีกทั้งมีการเก็บข้อมูลที่น้อยกว่า

2.23 คุณสมบัติของฟังก์ชันออโตคอร์รีเลชัน

1. เป็นฟังก์ชันคู่ $R(k) = R(-k)$
2. จะเป็นค่าสูงสุดเมื่อเปรียบเทียบสัญญาณของตัวเองที่ตำแหน่งทางแกนเวลาเดียวกัน คือ

$$R(0) = \max$$

จากหลักการพื้นฐานของการประมาณเชิงเส้นและแบบจำลองระบบการสร้างสัญญาณเสียง สามารถเขียนบล็อกไดอะแกรมการทำประมาณเชิงเส้นมาสร้างสัญญาณเสียงพูดได้ดังรูป



รูปที่ 2.30 บล็อกไดอะแกรมแสดงโมเดลการสร้างสัญญาณเสียงพูดอย่างง่าย

จากรูปที่ 2.8 สามารถเขียนสมการได้เป็น

$$S(n) = G \times u(n) + \sum_{k=1}^p \alpha_k S(n-k) \quad (2.22)$$

การประมาณเชิงเส้นโดยการใช้สัมประสิทธิ์ $\{\alpha_k\}$ คือ

$$S'(n) = \sum_{k=1}^p \alpha_k S(n-k) \quad (2.23)$$

ดังนั้นความคลาดเคลื่อน คือ

$$e(n) = S(n) - S'(n) \\ e(n) = S(n) - \sum_{k=1}^p \alpha_k S(n-k) \quad (2.24)$$

ฟังก์ชันถ่ายโอนระหว่าง $e(n)$ และ $S(n)$ คือ

$$A(z) = E(z) / S(z) \\ = 1 - \sum_{k=1}^p \alpha_k z^{-k} \quad (2.25)$$

จากสมการที่ (2.22) และ (2.24) จะเห็นได้ว่าถ้า $\{\alpha_k\} = \{a_k\}$ แล้ว

$$e(n) = G \times u(n) \quad (2.26)$$

ดังนั้น ค่าผลรวมของกำลังสองของความคลาดเคลื่อน

$$E_n = \sum_m e_n^2(m) \\ E_n = \sum_m [S(m) - S'(m)]^2 \quad (2.27)$$

โดยที่ n คือช่วงที่ n ของสัญญาณที่ใช้คำนวณ เพราะฉะนั้นเพื่อให้ได้ค่า E_n ต่ำที่สุดจะต้องมีเงื่อนไขว่า

$$\frac{\partial E_n}{\partial \alpha_i} = 0 \quad \text{เมื่อ } i = 1, 2, 3, \dots, p$$

จากสมการ (2.27)

$$\frac{\partial E_n}{\partial \alpha_i} = -2s(m-i) \sum_m \left[s_n(m) - \sum_{k=1}^p \alpha_k s_n(m-k) \right] \quad \text{เมื่อ } i = 1, 2, 3, \dots, p \\ = -2 \left[\sum_m s(m)s(m-i) - \sum_{k=1}^p \sum_m \alpha_k s(m-k)s(m-i) \right]$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\frac{\partial E_n}{\partial \alpha_i} = 0 \quad \text{ก็ต่อเมื่อ}$$

$$\sum_{k=1}^p \alpha_k \sum_m s_n(m-k)s_n(m-i) = \sum_{m-i} s_n(m)s_n(m-i) \quad (2.28)$$

เมื่อ $i = 1, 2, 3, \dots, p$

ถ้าเรากำหนดให้ $\phi_n(i, j) = \sum s_n(m-k)s_n(m-i)$ เพราะฉะนั้น

$$\sum_{k=1}^p \alpha_k \phi_n(i, j) = \phi_n(i, 0) \quad (2.29)$$

โดยสมการ (2.27) – (2.28) จะได้ว่า

$$E_n = \sum_m s_n^2(m) - \sum_{k=1}^p \alpha_k \sum_m s_n(m)s_n(m-k)$$

และจาก

$$\begin{aligned} \phi_n(i, j) &= \sum s_n(m-k)s_n(m-i) \\ E_n &= \phi_n(0, 0) - \sum_{k=1}^p \alpha_k \phi_n(0, k) \end{aligned} \quad (2.30)$$

สมมติว่าใน 1 เฟรม ของสัญญาณ ที่ตัดมามีจำนวน N ตัวอย่าง คือ $s_n(0), s_n(1), s_n(2), \dots, s_n(N-1)$ ในที่นี้เราให้ $s_n(m) = 0$ เมื่อ $m < 0$ หรือ $m > N-1$ เพราะฉะนั้น

$$\begin{aligned} \phi_n(i, j) &= \sum_m s_n(m-k)s_n(m-i) \\ &= \sum_{m=0}^{N-1-(j-k)} s_n(m)s_n(m+i-k) \quad 0 \leq k \leq p, 1 \leq i \leq p \end{aligned}$$

$$\text{ให้} \quad R_n(k) = \sum_m^{N-1-k} s_n(m)s_n(m+k) \quad \text{เมื่อ } k = 0, 1, 2, \dots, p \quad (2.31)$$

ดังนั้น จากสมการ (2.20) และ (2.21) จะได้ว่า

$$R_n(k) = \sum_{m=0}^{N-1-k} x'(m)x'(m+k) \quad (2.32)$$

จากสมการที่ (2.29) จะได้ว่า

$$\sum_{k=1}^p \alpha_k R_n(|i-k|) = R_n(i) \quad \text{เมื่อ } i = 1, 2, 3, \dots, p$$

สามารถเขียนให้อยู่ในรูปเมตริกได้เป็น

$$\begin{bmatrix} R_n(0) & R_n(1) & \cdots & R_n(p-1) \\ R_n(1) & R_n(0) & \cdots & R_n(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_n(p-1) & R_n(p-2) & \cdots & R_n(0) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix} = \begin{bmatrix} R_n(1) \\ R_n(2) \\ \vdots \\ R_n(p) \end{bmatrix}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$R_n \cdot \alpha = r_n \quad (2.33)$$

เมื่อ

$$R_n = \begin{bmatrix} R_n(0) & R_n(1) & \cdots & R_n(p-1) \\ R_n(1) & R_n(0) & \cdots & R_n(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_n(p-1) & R_n(p-2) & \cdots & R_n(0) \end{bmatrix}, \quad \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix}, \quad r_n = \begin{bmatrix} R_n(1) \\ R_n(2) \\ \vdots \\ R_n(p) \end{bmatrix}$$

2.24 การหาพารามิเตอร์ แอลพีซี

พารามิเตอร์ แอลพีซี ได้แก่ สัมประสิทธิ์ α และอัตราขยาย G เมื่อได้ค่า $R_n(0), R_n(1), R_n(2), \dots, R_n(p)$ จากสมการ (2.18) แล้วก็สามารถหาค่า α นั่นคือ

$$\alpha = R_n^{-1} \cdot r_n$$

และจากสมการ (2.22) และ (2.24) จะได้ว่า

$$e(n) = G \cdot u(n) \\ \therefore E_n = \sum_{m=0}^{N-1} e^2(m) = G \sum_{m=0}^{N-1} u^2(n) \quad (2.34)$$

จากสมการ (2.30) จะได้ว่า

$$E_n = \phi_n(0,0) - \sum_{k=1}^p \alpha_k \phi_n(0,k) \\ = R_n(0) - \sum_{k=1}^p \alpha_k R_n(k) \quad (2.35)$$

และจากสมการ (2.34) เราสามารถหาค่า G ได้โดยตรงจาก

$$G^2 = \frac{R_n(0) - \sum_{k=1}^p \alpha_k R_n(k)}{\sum_{m=0}^{N-1} u^2(m)} \quad (2.36)$$

ขั้นตอนนี้จะได้ว่าวิเคราะห์สัญญาณไปหนึ่งเฟรมแล้ว โดยในหนึ่งเฟรมนี้จะได้ค่าสัมประสิทธิ์ทั้งหมด 12 ค่า ($P = 12$); $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_p$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.25 การเปลี่ยนพารามิเตอร์แอลพีซี เป็นสัมประสิทธิ์เซปสตรีม

หลังจากการหาสัมประสิทธิ์แอลพีซี และ อัตราการขยายใน 1 เฟรมแล้ว จะเปลี่ยนให้เป็นสัมประสิทธิ์เซปสตรีม เนื่องจากการรู้จำเสียงพูดนั้น สัมประสิทธิ์เซปสตรีมเป็นพารามิเตอร์ที่มีลักษณะน่าเชื่อถือได้ดีกว่าสัมประสิทธิ์แอลพีซี ทั้งยังมีความใกล้ชิดกับการรับรู้เสียงตามความรู้สึกของมนุษย์โดยแท้จริง สัมประสิทธิ์เซปสตรีมสามารถหาได้โดยตรงจากสัมประสิทธิ์แอลพีซี ดังนี้

$$C_0 = \ln G$$

เป็นสัมประสิทธิ์ตัวแรกซึ่งเป็นแกน

$$Q \approx \frac{3}{2} p$$

โดย $p = 10$ ดังนั้น $Q = 12$ ไม่รวมแกน (C_0) จะได้ว่า สัมประสิทธิ์เซปสตรีมใน 1 เฟรม = 12 ตัว

$$C_m = a_m + \sum_{k=1}^{m-1} \left(\frac{k}{m} \right) C_k a_{m-k} \quad ; \quad 1 \leq m \leq p$$

$$C_m = \sum_{k=1}^{m-1} \left(\frac{k}{m} \right) C_k a_{m-k} \quad ; \quad m > p$$

2.26 การเวทค่าพารามิเตอร์ (Parameter Weighting)

เนื่องจากสัมประสิทธิ์เซปสตรีมที่ได้นั้น ช่วงลำดับต้น ๆ และลำดับท้าย ๆ ของเฟรมที่นำมาวิเคราะห์จะเกิดความคลาดเคลื่อนมากกว่าบริเวณส่วนอื่น เพราะฉะนั้นจึงทำการถ่วงน้ำหนักเพื่อลดค่าความคลาดเคลื่อนดังกล่าวนี้ ด้วยฟังก์ชันเวทดัง ดังนี้คือ

$$w_m = \left[1 + \frac{Q}{2} \sin \left(\frac{\pi m}{Q} \right) \right] \quad ; \quad 1 \leq m \leq Q$$

จะได้พารามิเตอร์สุดท้ายคือ

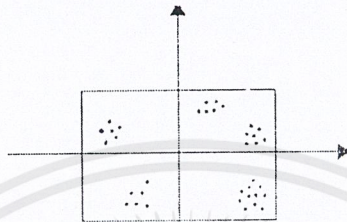
$$C'_m = C_m \cdot w_m \quad (2.37)$$

จากนั้นก็พิจารณาให้ครบทุกเฟรมของข้อมูล เมื่อพิจารณาเรียบร้อยแล้วก็จะนำไปจัดกลุ่มเสียง และสร้างแบบจำลองเสียงเพื่อใช้ในการเปรียบเทียบต่อไป

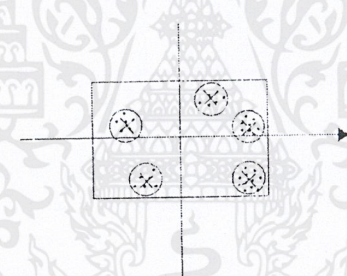
2.27 การจัดระดับเวกเตอร์ (Vector Quantization)

เวกเตอร์ควอนไทเซชัน เป็นวิธีการลดไคเมนชัน (Dimension) หรือจำนวนข้อมูล เวกเตอร์อินพุท หรือเซตเทรนนิ่ง หรือ พารามิเตอร์ที่ได้จากขั้นแอลพีซีจะถูกเลือกมากลุ่มหนึ่งซึ่งใช้เป็นตัวแทนของข้อมูลจำนวนหนึ่งหรือเรียกว่าการค้นหาโค้ดบุ๊ก อินพุทที่เข้ามาจะถูกเปรียบเทียบกับโค้ดบุ๊ก ที่มีอยู่ โดยพิจารณาว่าอินพุทที่เข้ามานั้นห่างจากโค้ดบุ๊ก ไคน้อยที่สุด อินพุทดังกล่าวจะถูกแทนด้วยเวกเตอร์โค้ด (index) นั้น อินพุททุกตัวเป็นสมาชิกของเวกเตอร์โค้ดใด ๆ จะถูกนำมาหาจุดศูนย์กลางร่วมใหม่ และนำเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จุดศูนย์กลางที่ได้นี้ไปทำการหาความคลาดเคลื่อนกับสมาชิกทุกตัว ถ้าค่าความคลาดเคลื่อนที่ได้มีค่ามากกว่าค่าที่กำหนดไว้ค่าหนึ่งหรือค่าที่ยอมรับได้ ก็จะนำศูนย์กลางใหม่นั้นไปเป็นโค้ตบุ๊ก แทน และจะทำการจัดกลุ่มอินพุตเข้ากับโค้ตบุ๊ก ใหม่ที่ได้ และหาค่าความคลาดเคลื่อนอีกครั้งทำอย่างนี้ซ้ำ ๆ จนกระทั่งค่าความคลาดเคลื่อนมีค่าน้อยถึงค่าที่ยอมรับได้ ก็จะถือว่าโค้ตบุ๊ก ที่ดีที่สุดจะเป็นตัวแทนของอินพุตทั้งหมด จะสังเกตได้ว่าทุกครั้งที่มีการหาโค้ตบุ๊ก ใหม่ นั้น ค่าความคลาดเคลื่อนที่ได้จะมีค่าลดลงทุกครั้งด้วย

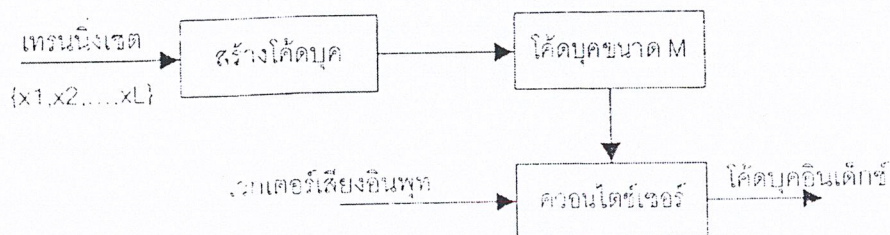


รูปที่ 2.31 แสดงการกระจายเฟรมของเสียงพูดแต่ละจุดแทนเฟรมของเสียง



รูปที่ 2.32 การรวมกลุ่มของเฟรมเสียงเพื่อสร้างโค้ด X

ตัวอย่างเวกเตอร์ควอนไทเซชัน สมมติให้เวกเตอร์แต่ละตัวมี 2 มิติ และทำการหาโค้ตบุ๊ก ขนาด 8 เวกเตอร์ ทั้งหมดจะถูกจัดเข้ากลุ่มของโค้ตบุ๊ก ต่าง ๆ แล้วทำการหาจุดศูนย์กลางใหม่ โดยการเฉลี่ยค่าเวกเตอร์สมาชิกทุกตัวที่อยู่ในกลุ่มเดียวกัน ผลที่ได้คือ โค้ตบุ๊ก 8 ตัว เป็นตัวแทนของเวกเตอร์ทั้งหมด



รูปที่ 2.33 บล็อกไดอะแกรมของเวกเตอร์ควอนไทซ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำงานของควอนไทซ์แบบเวกเตอร์ แบ่งเป็น ขั้นตอน ดังนี้

การสร้างโค้ดบุ๊ก (Codebook) โดยวิธีเคมีน (K-means) จะเริ่มจากขั้นตอนการประมาณเชิงเส้นของเสียงตัวอย่างจำนวนมาก จะได้เทรนนึงเซต ซึ่งประกอบด้วยเวกเตอร์สเปคตรัมจำนวน L เฟรม ; $x = \{x_i; 1 \leq i \leq L\}$ เฟรมละ P มิติ ; $x = [x_1, x_2, \dots, x_p]$ แล้วนำข้อมูลที่ได้นำมาทำการสร้างกลุ่มของแบบอ้างอิง

ในระบบการรับรู้เสียงพูดแบบต่าง ๆ จะใช้อ้างอิงจำนวนหนึ่งจากผู้พูดจำนวนมาก เพื่อที่จะได้ครอบคลุมความแปรปรวนแบบต่าง ๆ ที่เกิดระหว่างผู้พูดแต่ละคน เนื่องจากถ้าใช้แบบอ้างอิงจำนวนมาก เวลาที่ใช้ในการตอบสนองจะมาก เนื้อที่หน่วยความจำสำรองที่ใช้เก็บแบบอ้างอิงจะเพิ่ม และเมื่อเพิ่มแบบอ้างอิงไประดับหนึ่งความถูกต้องในการรับรู้ก็จะเริ่มคงที่ ดังนั้นการจัดกลุ่มแบบอ้างอิงใหม่เพื่อให้ได้แบบอ้างอิงที่เหมาะสม และสามารถใช้เป็นตัวแทนของแบบอ้างอิงที่มีอยู่ทั้งหมดได้ อัลกอริทึมที่ใช้ได้แก่ วิธีการแบบเคมีน (K-means Algorithm) ขั้นตอนการสร้างโค้ดบุ๊กมีดังนี้

1) นำเทรนนึงเซตมาใช้ในการสร้างโค้ดบุ๊ก

ขนาดโค้ดบุ๊กของการควอนไทซ์ แบบเวกเตอร์ คือ $M = 2^b$ เวกเตอร์ และเพื่อที่จะหาเซตของ M โค้ดบุ๊กที่ดีที่สุด จำนวนเวกเตอร์อินพุตจะต้องมากกว่าขนาดโค้ดบุ๊กมาก ๆ

2) การสุ่มค่าเริ่มต้น

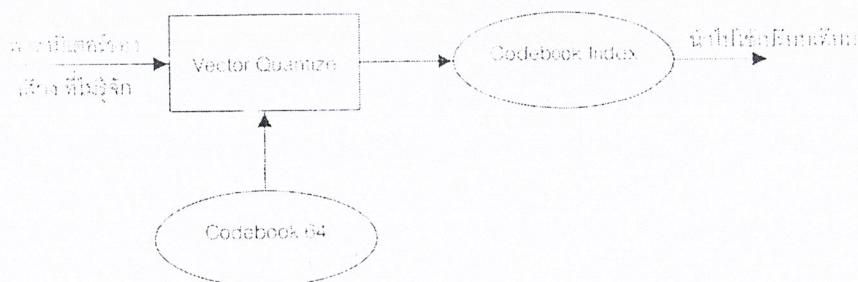
การสุ่มค่าเริ่มต้น เป็นวิธีหนึ่งในการออกแบบโค้ดบุ๊ก ซึ่งคือ การเลือกค่าเริ่มต้นของโค้ดบุ๊ก เรียกโค้ดบุ๊กที่ได้จากการสุ่มค่าเริ่มต้นนี้ว่า แรนดอมโค้ดบุ๊ก (Random Codebook) ถึงแม้วิธีนี้จะไม่ใช่วิธีที่ดีนัก แต่โค้ดบุ๊กที่ได้จากการสุ่มก็ได้ผลเป็นที่ยอมรับ

3) การหาความคลาดเคลื่อน

การหาความคลาดเคลื่อน เป็นส่วนที่จำเป็นและเป็นประโยชน์ต่อการออกแบบโค้ดบุ๊ก สมการทางพีชคณิตที่ใช้ในการหาระยะทางมีหลายวิธี แต่ที่นำมาใช้คือ การหาความคลาดเคลื่อนกำลังสองรวม (Total square error) ซึ่งมีวิธีที่ง่ายและรวดเร็ว

ถ้าสมการมี P มิติ สามารถหาระยะห่างระหว่างอินพุต (x) กับเวกเตอร์โค้ด (y) โดยสมการ

$$d(v_1, v_2) = \|v_1 - v_2\|^2 = \sum_{i=0}^{k-1} (x_i - y_i)^2 \quad (2.39)$$



รูปที่ 2.34 ขั้นตอนของเวกเตอร์ควอนไทซ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4) การจัดกลุ่ม (Classification) และการหาจุดศูนย์กลางของกลุ่ม (center cluster)

การจัดกลุ่มเป็นการแบ่งเวกเตอร์อินพุตเข้าไปตามกลุ่มต่าง ๆ ของแรนดอมโกลด์บुक โดยพิจารณาระยะทางหรือความคลาดเคลื่อนน้อยที่สุดของแต่ละเวกเตอร์อินพุต x กับเวกเตอร์โกลด์บुक y ซึ่งเป็นโกลด์บुकจากนั้นจะทำการหาค่าเฉลี่ยของแต่ละกลุ่ม เพื่อเป็นค่ากลางของกลุ่มนั้น ๆ จะได้

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

\bar{Y} เป็นจุดศูนย์กลางซึ่งเป็นเวกเตอร์ที่อยู่ตรงกลางของ $\{x_i\}_{i=1}^L$ ซึ่งแต่ละมิติจะไม่ขึ้นแก่กัน หมายความว่าแต่ละ y_k เป็นค่ากลางของ $\{x_i\}_{i=1}^L$ ทำ 2 ขั้นตอนซ้ำ จะเกิดการลู่เข้า (convergent) โดยความคลาดเคลื่อนรวมจะต่ำกว่าค่าหนึ่ง ๆ ซึ่งค่าความคลาดเคลื่อนรวมจะลดลงทุกครั้งที่มีการคำนวณซ้ำใหม่ จึงขึ้นกับค่าที่กำหนดว่าต้องการให้ความคลาดเคลื่อนรวมน้อยกว่า ค่ากลางดังกล่าวของแต่ละกลุ่มจะถูกเก็บเป็นเวกเตอร์โกลด์บุกจะได้ว่า y เป็นควอนไทซ์ของค่า x

โดย $q(\cdot)$ เป็นโอเปอเรเตอร์ของควอนไทซ์ y ถูกเรียกว่าเอาท์พุทเวกเตอร์ของค่า x โดย y เป็นค่าใดค่าหนึ่งใน $Y = \{y_i, 1 \leq i \leq M\}$ โดย $y_i = [y_{i1} \ y_{i2} \ \dots \ y_{ip}]$ Y เป็นเซตของโกลด์บุก

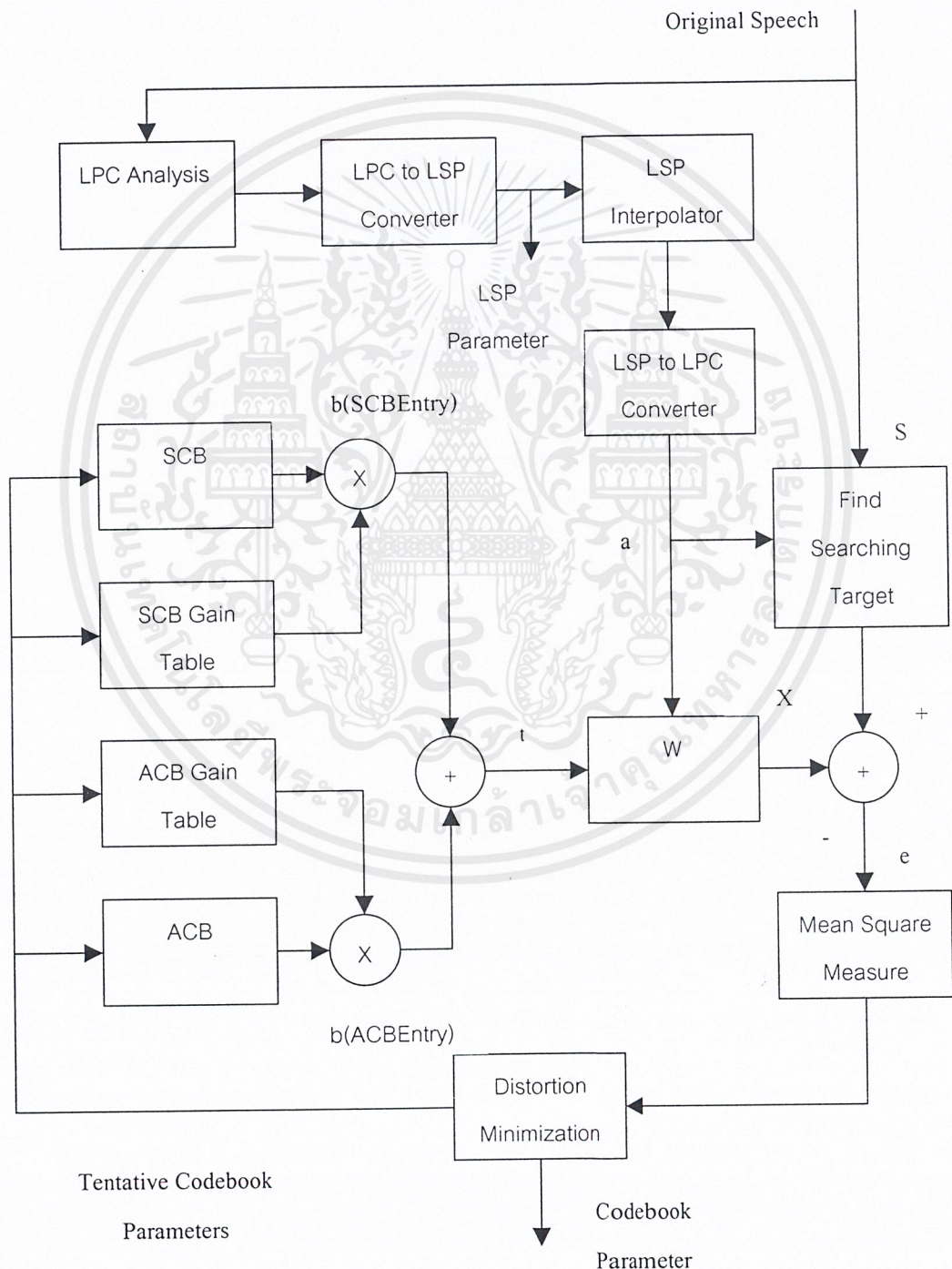
M เป็นขนาดของโกลด์บุก และ $\{y_i\}$ เป็นเซตของเวกเตอร์โกลด์บุก อาจเรียกว่าโกลด์บุกอ้างอิง และ M อาจเรียกว่าจำนวนระดับขั้น จะทำการแบ่งเวกเตอร์ X ไปใน M เซล $\{C_i; 1 \leq i \leq M\}$ เมื่อ X อยู่ในเซล C_i

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

บทที่ 3

การคำนวณและการออกแบบ

เทคนิคของการบีบอัดข้อมูลเสียง (speech compression) โดยทั่วไปนั้น มีพื้นฐานอยู่บน 2 การทำงานที่แตกต่างกันคือ การกำจัดข้อมูลที่มีมากเกินไป (redundancy) และการกำจัดข้อมูลที่ไม่มีความจำเป็น (irrelevancy) โดยในการกำจัดข้อมูลที่มีมากเกินไปจะใช้วิธีการทำนาย (prediction) และ/หรือ การแปลง (transform) เพื่อลดจำนวนของข้อมูลที่มีมากเกินไป ซึ่งเป็นการลดจำนวนอัตราบิต (bit rate)



รูปที่ 3.1 บล็อกไดอะแกรมรวมของการเข้ารหัส

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนวิธีที่ใช้ในการลดข้อมูลที่ไม่มีความจำเป็น จะทำโดยใช้วิธีการลดอัตราบิตผ่านการควอนไทเซชันของค่าประกอบของเวลาจากการหาค่าผิดพลาดในการทำนาย หรือ การแปลงค่าสัมประสิทธิ์

3.1 หลักการของซีอีแอลพี (CELP)

. ตัวการทำนาย (predictor) 2 ตัว ซึ่งได้แก่ พิตช์ (pitch) และ สเปกตรัลพรีดิกเตอร์ (spectral predictor) เป็นตัวที่ทำการกำจัดข้อมูลที่มาจนเกินไป ซึ่งเป็นผลมาจากการคอร์รีเลชัน (correlation) ของสปีช แซมเปิ้ล (speech sample) ทั้งเทอมระยะยาวและสั้น (long term and short term)

. โคลลสลับ (close loop) โดยการใช้ เพอร์เซปชวลเวกเตอร์ควอนไทเซอร์ (perceptual vector quantizer) เช่น โค้ดบุ๊ก (codebook) เพื่อที่จะทำการลดข้อมูลที่ไม่มีความจำเป็นจากองค์ประกอบทางเวลาของการทำนายความผิดพลาด โดยโค้ดบุ๊กจะเก็บสัญญาณแบบสุ่ม (random signal) หรือ สโตเคสติก (stochastic) เป็นต้นแบบของสัญญาณกระตุ้น สำหรับ พรีดิกเตอร์ ฟิลเตอร์ (predictor filter) ทั้ง 2 ตัว และ เพอร์เซปชวล เวตติ้ง ฟิลเตอร์ (perceptual weighting filter) ทำหน้าที่วัดค่าการสะท้อนเป็น มิน-สแควร์ เออริอร์ (mean square error) ให้เป็นความผิดพลาดที่ยอมรับได้

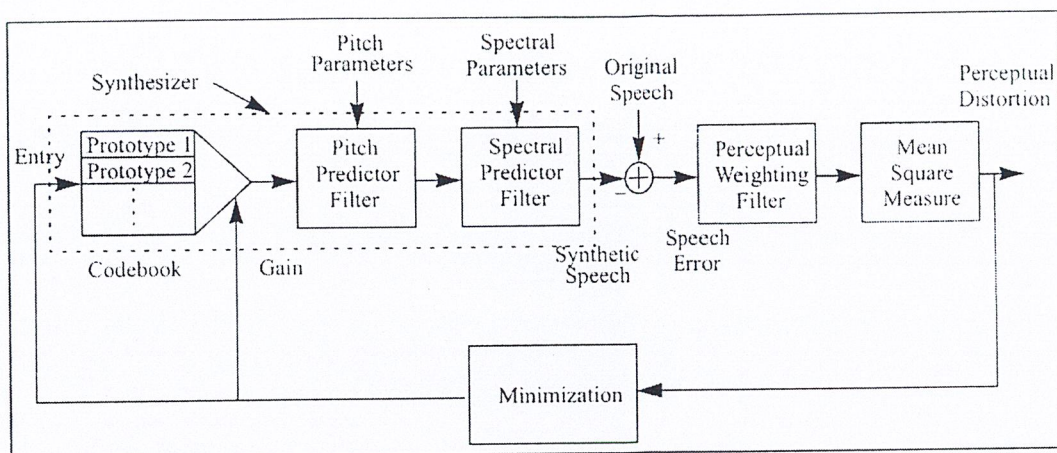
วิธีการบีบอัดข้อมูลเสียงแบบซีอีแอลพี ประกอบด้วย

- . กลุ่มของค่าพารามิเตอร์ สเปกตรัล พรีดิกเตอร์
- . กลุ่มของค่าพารามิเตอร์ พิตช์ พรีดิกเตอร์
- . โค้ดบุ๊กพารามิเตอร์ทั้ง เอนทรี (entry) และเกน (gain)

. วิธีซีอีแอลพี สามารถส่งข้อมูลและเก็บข้อมูลได้ที่อัตราต่ำกว่าหรือเท่ากับ 4.8 กิโลบิตต่อวินาที อัลกอริธึมของการบีบอัดข้อมูลเสียงเริ่มจากการหาค่าพารามิเตอร์การทำนายและทำการหาโค้ดบุ๊กพารามิเตอร์ที่ทำการตอบสนองกับต้นแบบของการกระตุ้นที่สามารถทำให้ค่าความผิดพลาดมีค่าน้อยที่สุด และดีคอมเพรสเซอร์ (decompressor) ของซีอีแอลพี จะใช้เพอร์เซปชวลโค้ดบุ๊กพารามิเตอร์ เพื่อทำการผลิตสัญญาณกระตุ้น (excitation signal) ที่ทำหน้าที่กระตุ้นฟิลเตอร์แบบพิตช์กับสเปกตรัลที่ต่อแคสเคด (cascade) โดยผลที่ได้คือสปีชที่ผ่านการดีคอมเพรสแล้ว

ในแต่ละขั้นตอน แต่ละต้นแบบจะต้องมีการผ่านฟิลเตอร์ 3 ตัวคือ พิตช์ฟิลเตอร์, สเปกตรัลฟิลเตอร์ และเพอร์เซปชวลฟิลเตอร์ และจะนำมาผ่านขั้นตอนการหาค่ามินสแควร์

สเปกตรัลพารามิเตอร์ จะถูกนำมาควอนไทซ์จนได้เป็นแอลเอสพีส่วนพิตช์ฟิลเตอร์เป็นโค้ดบุ๊กชนิดหนึ่งที่เรียกว่า ACB และ โค้ดบุ๊กสำหรับสัญญาณที่เป็นสัญญาณสุ่ม (random signal) จะเรียกว่า SCB



รูปที่ 3.2 บล็อกไดอะแกรม (Block Diagram) การวิเคราะห์ซีอีแอลพี ประกอบไปด้วย

3.2 สัญญาณอินพุต(Input) และ สัญญาณเอาต์พุต (Output)

ในทางปฏิบัติซีอีแอลพี เป็นวิธีการเข้ารหัสแบบแบ่งเป็นบล็อก ซึ่งแต่ละเฟรมประกอบไปด้วย PCM 240 แซมเปิ้ล ซึ่งมีขนาดของข้อมูลทั้งหมด 1.92 กิโลบิต เมื่อทำการบีบอัดแล้วจะมีขนาดของข้อมูล 144 บิต

พารามิเตอร์ซีอีแอลพี ประกอบด้วย

- . พารามิเตอร์ไลน์สเปกตรัมแพร์ (line spectrum pair :LSP)
- . พารามิเตอร์โค้ดบุ๊กอะแดปทีฟ (adaptive codebook :ACB)
- . พารามิเตอร์โค้ดบุ๊กสโตคัสติก (stochastic codebook :SCB)

โดยทั้งพารามิเตอร์แอลเอสพี และอะแดปทีฟโค้ดบุ๊ก (ACB) และ สโตคัสติกโค้ดบุ๊ก (SCB) เป็นการเรียงลำดับของตารางการควอนไทซ์และโค้ดบุ๊กซึ่งเรียกว่าอะแดปทีฟโค้ดบุ๊ก (ACB) กับตารางอัตราขยายของอะแดปทีฟโค้ดบุ๊ก (ACB) และ สโตคัสติกโค้ดบุ๊ก (SCB) กับตารางอัตราขยายของสโตคัสติกโค้ดบุ๊ก (SCB) ซึ่งทั้งหมดต้องการ 138 บิต ซึ่งจะใช้ 6 บิต สำหรับการแก้ไขส่วนผิดพลาด (error correction) การซิงโครไนซ์ (synchronization) และการขยายในอนาคต (Future Expansion) ซึ่งซีอีแอลพี ที่ทำการบีบอัดเรียกว่าการวิเคราะห์ (analyzer) ประกอบไปด้วย การวิเคราะห์แอลเอสพี เพื่อที่จะได้พารามิเตอร์แอลเอสพี การหาโค้ดบุ๊กทั้งแบบโค้ดบุ๊กอะแดปทีฟ (adaptive codebook) และ โค้ดบุ๊กสโตคัสติก (stochastic codebook) ส่วนตัวการแปลงกลับของซีอีแอลพี (CELP decompressor) เรียกว่าส่วนสังเคราะห์ (synthesizer)

3.3 การวิเคราะห์แอลเอสพี (LSP Analysis)

ประกอบไปด้วย 3 ขั้นตอน คือ

- 1.) การหาค่าลิเนียร์พรีดิกทีฟโค้ดดิ้ง (แอลพีซี) (linear predictive coding (LPC)) จากพีซีเอ็ม แซมเปิ้ล (PCM sample) เพื่อเป็นการแทนข้อมูลของสเปกตรัล
- 2.) ทำการแปลงพารามิเตอร์แอลพีซีเป็นพารามิเตอร์แอลเอสพีเพื่อเป็นการแทนค่า

สัมประสิทธิ์

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์การใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 3.) การทำให้ค่าพารามิเตอร์แอลเอสพี ที่ได้มีค่าที่คงที่ (stability)

3.3.1) การวิเคราะห์แอลพีซี (LPC)

เป้าหมายของการวิเคราะห์แอลพีซี คือการหาค่าพารามิเตอร์ของแอลพีซี ซึ่งก็คือ a_i (เมื่อรวมสัมประสิทธิ์ทุกตัวคือ a) มีการตอบสนองกับสเปกตรัมฟิลเตอร์ซึ่งเป็นรูปแบบของ vocal tract ของมนุษย์ ซึ่งรูปแบบที่นิยมใช้มากที่สุดคือ 10 ออร์เดอร์ ดิจิตอลฟิลเตอร์แบบออลโพล (10-order Digital filter all pole) ($H(z)$) ประกอบไปด้วยสัมประสิทธิ์ a_i 10 ตัว

$$H(z) = \frac{1}{1 + \sum_{i=1}^{10} a_i z^{-i}} = \frac{1}{A(z)} \quad (3.1)$$

โดยกำหนดให้อินพุทของฟิลเตอร์เป็นสัญญาณแบบ zero-mean (t) และเอาท์พุทของฟิลเตอร์คือ \hat{s} ใน z โดเมนจะได้ว่า

$$\hat{S}(z) = H(z)T(z) \quad (3.2)$$

การวิเคราะห์แอลพีซี จะทำการหาค่า $\|s - \hat{s}\|$ ที่มีค่าน้อยที่สุดโดย a_i ของเวกเตอร์ a คือพารามิเตอร์แอลพีซี ซึ่งได้จากสมการ

$$0 = \sum_{i=0}^{10} a_i r_{i-j} \quad a_0 = 1 \quad j = 1, \dots, 10 \quad (3.3)$$

โดย r_i คือ ออโตคอร์รีเลชัน ซึ่งมีสมการดังนี้

$$r_i = \sum_{n=i}^{N-1} s[n]s[n-i] \quad i = 0, \dots, 10 \quad (3.4)$$

3.3.2) การแปลงพารามิเตอร์แอลพีซีเป็นพารามิเตอร์แอลเอสพี

ระบบจะมีการควอนไทซ์ค่า a โดยใช้วิธีการวิเคราะห์แอลเอสพี เนื่องจาก a_i มีค่าเป็นจำนวนจริง 10 ค่า ซึ่งต้องมีการใช้บิตจำนวนมากในการแทน ($10 \times 16 = 160$ บิต) การหาพารามิเตอร์แอลเอสพี จะมีการใช้บิตจำนวนน้อยเพียง 34 บิต เนื่องจากเป็นจำนวนเต็ม 10 ค่า มีขอบเขตตั้งแต่ 0 - 8 หรือ 0 - 16 มีการตอบสนองกับเอนทรี (entry) ที่เหมาะสมของตารางค่าแอลเอสพี

การแสดงการแปลงเราจะให้ a แทนค่า z_i ซึ่งก็คือค่าซีโร่ (zero) ของสมการโพลีโนเมียล $p(z)$ และ $q(z)$ มีความสัมพันธ์ดังนี้

$$\begin{aligned}
 A(z) &= \frac{1}{2}(p(z) + q(z)) \\
 p(z) &= A(z) + z^{-11}A(z^{-1}) \\
 q(z) &= A(z) - z^{-11}A(z^{-1})
 \end{aligned}
 \tag{3.5}$$

โดยที่สมการโพลีโนเมียล $p(z)$ และ $q(z)$ แทนคิจิตอลฟิลเตอร์ $H(z)$ ซึ่งทั้ง $p(z)$ และ $q(z)$ จะมีซีโร่อย่างละ 11 ตัวซึ่งหมายถึง a

โดย z_i สามารถแทน ω_i ได้โดย

$$\omega_j = \arg(z_j); \quad i = 0, \dots, 9 \tag{3.6}$$

โดย $\arg(\cdot)$ คือ ค่าอาร์กิวเมนต์ของตัวแปรเชิงซ้อนเมื่อ $z = 1$ และ $z = -1$ จะเป็นค่าซีโร่ของ $p(z)$ และ $q(z)$ เสมอ ตามลำดับ ดังนั้นซีโร่ที่เหลือ 20 ค่า ของทั้ง $p(z)$ และ $q(z)$ โดยจะมีลักษณะสมมาตรบนแกนจริง (real axis) และอยู่ในวงกลมหนึ่งหน่วย (unit circle) ในระนาบ z

ดังนั้นซีโร่ 10 ค่าที่อยู่ต่ำกว่าแกนจริงจะไม่มีค่าเป็นจริงเหลือเพียงซีโร่ 10 ค่าที่อยู่เหนือแกนจริง จึงมีค่า ω_j จะประกอบไปด้วยตำแหน่งคู่และคี่โดยเป็นตำแหน่งของ $p(z)$ และ $q(z)$ สลับกันไป 10 ค่า เราจึงสรุปได้ว่า ω_j 10 ค่านี้อาจสามารถสร้างซีโร่ของทั้ง $p(z)$ และ $q(z)$ ขึ้นมาใหม่ซึ่งแทนค่า a

เราสามารถได้ค่า ω_j จากการควอนไทซ์ อย่างไรก็ตามเราสามารถควอนไทซ์ค่า a_j ได้โดยตรงโดยค่าเปลี่ยนแปลงของ a_j มีค่ากว้าง (มีค่าที่มีความสำคัญของ a_j อยู่มาก) ซึ่งต้องใช้ควอนไทซ์ สเต็ป จำนวนมากเพื่อลดการผิดพลาดที่เกิดจากการควอนไทซ์ (quantization error) หรือแต่ละ ω_i จะมีขอบเขตจำกัดเนื่องจากขอบเขตของ ω_i ไม่เกี่ยวข้องกับช่วงย่อย s_j ในช่วงจำนวนจริง $0 - \pi$

$$\begin{aligned}
 \omega_j &\in S_j; & 0 \leq \cup S_j < \pi \\
 i \neq j &\Rightarrow (S_j \cap S_i = \emptyset); & j = 0, \dots, 9
 \end{aligned}
 \tag{3.7}$$

ดังนั้นการใช้ควอนไทซ์สเต็ป จำนวนน้อยสำหรับการหาค่า ω_j จะทำให้ได้ค่าผิดพลาดจากการควอนไทซ์โดยจะใช้ตารางแอลเอสพี ในการควอนไทซ์ ω_j สำหรับแต่ละค่าของ ω_j จะสร้างเซตของค่าควอนไทซ์ 8 ค่า หรือ 16 ค่า ถ้าเป็นตำแหน่งที่ 2 ถึง 5 โดยค่าตารางแอลเอสพี เป็นค่าตารางแอลเอสพี เป็นค่า $[j, i]$ โดย j มี 10 ค่าตั้งแต่ 0 ถึง 9 และค่า i มีค่าตั้งแต่ 0 ถึง 7 หรือ 0 ถึง 15 โดยเราจะทำการหา ω_j ที่ใกล้เคียงที่สุดในรายการของ j แล้วเราจะให้ค่า i เป็นค่า LSP_j และทำเช่นเดียวกันกับค่าทั้งหมดของ j เราจะให้ค่า LSP_j แทน ค่า ω_j ที่ถูกควอนไทซ์ซึ่ง LSP_j จะเป็นพารามิเตอร์แอลเอสพี ซึ่งจะแทน a

วิธีการนี้จะมีประสิทธิภาพมากเนื่องจากมีความต้องการเพียง 34 บิตสำหรับค่า a แต่ละตัวโดยจะเป็น $(3 + 4 + 4 + 4 + 4 + 3 + 3 + 3 + 3 + 3)$

โดยมีอีกวิธีที่เร็วกว่าคือวิธีการตรวจสอบค่าซีโร่ ครอสซิง (zero crossing) ของคู่โพลีโนเมียลคู่ใหม่

เอกสารนี้ $\tilde{p}(x)$ และ $\tilde{q}(x)$ ซึ่งมีความสัมพันธ์กับค่า $p(z)$ และ $q(z)$ ดังนี้ ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$x_j = \cos \omega_j \quad (3.8)$$

โดย $\tilde{p}(x)$ และ $\tilde{q}(x)$ ได้จาก

$$\begin{aligned} \tilde{p}(x) &= \sum_{i=0}^5 b_i x^i \\ \tilde{q}(x) &= \sum_{i=0}^5 c_i x^i \end{aligned} \quad (3.9)$$

โดยค่าสัมประสิทธิ์ b และ c คือ

$$\begin{aligned} b_5 &= 32 \\ b_4 &= 16p_1 \\ b_3 &= 8(p_2 - 5) \\ b_2 &= 4(p_3 - 4p_1) \\ b_1 &= 2(p_4 - 3p_2 + 5) \\ b_0 &= p_5 - 2p_3 + 2p_1 \end{aligned} \quad (3.10)$$

$$\begin{aligned} c_5 &= 32 \\ c_4 &= 16q_1 \\ c_3 &= 8(q_2 - 5) \\ c_2 &= 4(q_3 - 4q_1) \\ c_1 &= 2(q_4 - 3q_2 + 5) \\ c_0 &= q_5 - 2q_3 + 2q_1 \end{aligned} \quad (3.11)$$

โดย p_i และ q_i คือค่าสัมประสิทธิ์ของ $p(z)$ และ $q(z)$ ตามลำดับโดย i อ้างอิงถึง เทอมของโพลีโนเมียลประกอบด้วย z_i และ p_0 และ q_0 จะมีค่าเป็นศูนย์เสมอ

โดยค่า a ที่ได้ตามสมการที่ 7 และ p_i กับ q_i ที่เหลือจะได้อาจจากการวนรูปหาค่าจาก 1 ถึง 5 ของสมการ

$$\begin{aligned} p_i &= a_i - a_{11-i} p_{i-1} \\ q_i &= a_i - a_{11-i} q_{i-1} \end{aligned} \quad (3.12)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยวิธีนี้ชี้ให้เห็นว่าค่า x แต่ละค่าสัมพันธ์กับค่าซีโร่ (zero) ของ $p(z)$ หรือ $q(z)$ เนื่องจาก $p(x)$ หรือ $q(x)$ มีค่าเป็นศูนย์ตามลำดับ ดังนั้นจะมีการประยุกต์ค่า x ให้สอดคล้องตอบสนองกับ ω ในตาราง แอลเอสพี และค่าโพลีโนเมียล $\tilde{p}(x)$ และ $\tilde{q}(x)$ และการหาค่าซีโร่ครอสซิง (zero crossing)

3.3.3) การทำให้แอลเอสพีมีค่าคงที่

การแปลงแอลเอสพี เป็นแอลเอสพี นั้นจะมีความเปลี่ยนแปลงสภาพได้มากเนื่องจาก $H(z)$ เป็นรีเคอร์ซีฟฟิลเตอร์ (Recursive Filter) การลดทอนค่า a อาจจะทำให้โพล (pole) เคลื่อนออกจากวงกลมหนึ่งหน่วย (unit circle) ได้ง่าย เป็นผลทำให้ฟิลเตอร์ $H(z)$ มีค่าไม่แน่นอน และการแปลงแอลเอสพี ยังเกิดการลดทอนเนื่องจากการควอนไทซ์ (Quantization error)

ดังนั้น ถ้าค่าของ ω_j มีค่าเพิ่มขึ้นเพียงอย่างเดียวจาก 0 ถึง π ดิจิตอลฟิลเตอร์ $H(z)$ จะมีความคงที่อย่างแน่นอน

ดังนั้นก่อนที่จะทำการส่งค่า LSPj จะต้องทำการหาลำดับของค่า ω_j ที่ตอบสนองกับค่า LSPj ถ้าลำดับของค่าไม่เพิ่มขึ้นเพียงอย่างเดียวแต่มีค่าที่สลับกันไปมาจะถือว่าเป็นการทำลายความคงที่ของดิจิตอลฟิลเตอร์ดังนั้นระบบจะทำการแทนที่ค่าแอลเอสพี ด้วยค่า LSPj ที่มีความคงที่ของเฟรมที่ผ่านมา

และในบางครั้ง การกำหนดค่าช่วงของการควอนไทซ์จะทำให้เกิดปัญหากับความคงที่ จะเกิดขึ้นเมื่อค่า ω_j ที่ติดกันมีค่าใกล้เคียงกันมาก ดังนั้นตารางจะไม่สามารถแยกแยะออกจากกันได้ หรือในกรณีที่ค่า ω_j มีค่าไม่อยู่ในความครอบคลุมของตาราง ค่าแอลเอสพี ที่ได้จะไม่เที่ยงตรง และไม่มีค่าแน่นอนวิธีแก้ไขปัญหานี้คือ การขยายแบนด์วิดธ์

$$c_i = a_i \gamma^i \quad (3.13)$$

โดย γ คือค่าองค์ประกอบการขยาย (Expanding factor) ปกติมีค่า 0.994 และ i เป็นค่าดัชนี โดยมีค่าตั้งแต่ 1 ถึง 10

3.4 การหาค่าพารามิเตอร์ของโค้ดบุ๊ก (Codebook Parameter Searching)

3.4.1) ปัญหาในการค้นหา

ในการหาค่าพารามิเตอร์ของโค้ดบุ๊กจะต้องทำการลดค่าความผิดพลาดลงให้เหลือค่าความผิดพลาดลงให้เหลือค่าความผิดพลาดที่รับรู้ได้ (Perceptual distortion) น้อยที่สุดโดย

$$\|e\|^2 = \|s - \hat{s}\|_W^2 = \|P_W |s - \hat{s}|\|^2 \quad (3.14)$$

เมื่อค่า $\| \cdot \|$ แสดงถึงค่าระดับเฉลี่ยหรือขนาดของเวกเตอร์และ P_W แทนค่าเวกต์ริงฟิลเตอร์ที่ยอมรับได้ (perceptual weighting filter) แสดงได้โดย

$$P_W(z) = \frac{H\left(\frac{z}{\gamma}\right)}{H(z)} \quad 0 \leq \gamma \leq 1 \quad (3.15)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จะแทน r ด้วย 0.8 ซึ่งสมการของ $P_W(Z)$ จะทำให้สมการที่ 2 มีมาตรฐานในการวัดได้มากกว่าการวัดจากรูปแบบที่เหมือนกับรูปคลื่นโดยตรง

ค่าพารามิเตอร์ในโค้ดบुक จะมีผลกระทบต่อค่าความคลาดเคลื่อนในสมการที่ 14 โดยการกระตุ้น t และ \hat{s} โดยโค้ดบुक จะประกอบด้วย ดันแบบ หรือ โค้ดเวิร์ดบี (Code words b) ซึ่งจะเป็นขบวนของอาร์เรย์(Array) ของอิมพัลส์ $b[n]$ ซึ่งแต่ละโค้ดบुकจะถูกชี้โดยตัวเข้าโค้ดบुक เรียกว่า ซีบี เอนทรี (CB Entry : Codebook Entry) ซึ่งสำหรับแต่ละโค้ดบुक จะมีตารางของอัตราขยายเกนเทเบิล (Gain Table) ซึ่งประกอบด้วยตัวประกอบซึ่งเป็นอัตราขยายซึ่งเป็นเลขจำนวนจริง แต่ละค่าของอัตราขยายจะชี้โดยค่าการเข้าตารางของอัตราขยายเรียกว่า เกนเอนทรี (Gain Entry) ดังนั้นสำหรับ เอซีบี (ACB) และ เอสซีบี (SCB) จะมีการเข้ารหัสในโค้ดบुकแบบเปลี่ยนแปลง (ACB Entry) และการเข้ารหัสในโค้ดบुकแบบสุ่มค่า (SCB Entry) ตามลำดับ เมื่อการเข้ารหัสตารางอัตราขยาย เอซีบี และ เอสซีบี เกนเอนทรี ตามลำดับ กลุ่มของการเรียงลำดับเหล่านี้จะสร้าง t โดยขึ้นอยู่กับ

$$t = b_{(a)}(ACBEntry)g_{(a)}(ACBGainEntry) + b_{(s)}(SCBEntry)g_{(s)}(SCBGainEntry) \quad (3.16)$$

ค่า $b(a)(ACB Entry)$ และ $b(s)(SCB Entry)$ จะเป็น ACB และ SCB ที่ไปยังที่โค้ดเวิร์ด ด้วย ACB Entry และ SCB Entry ตามลำดับ เมื่อ $g(a)(ACB Gain Entry)$ และ $g(s)(SCB Gain Entry)$ จะเป็นค่าปัจจัยของอัตราขยาย เอซีบี และเอสซีบี ซึ่งจะทำการชี้โดย เกนเอนทรี และเอสซีบีเกนเอนทรี ตามลำดับ สำหรับค่า s ที่กำหนดให้และเมตริกซ์ของโค้ดบुक t จะสร้าง \hat{s} และ e จากสมการ (2) และสมการ (14) ตามลำดับ

ปัญหาของการค้นหาสำหรับเป้าหมายการค้นหา s ในการหาค่า เอซีบี เอนทรี, เอสซีบีเอนทรี, เอซีบีเกนเอนทรี และ เอสซีบีเกนเอนทรี จะขึ้นอยู่กับค่า e ที่น้อยที่สุดจากสมการที่ (16)

เพื่อแก้ปัญหาการค้นหาจะมีวิธีการมากมาย แต่อย่างไรก็ตามเราไม่สามารถใช้วิธีการทั้งหมดรวมกันได้ ดังนั้นในที่นี้เราจะใช้วิธีการค้นหาอย่างรวดเร็ว (ตามมาตรฐาน FS 1016) และยังคงพิจารณาถึงความต่อเนื่องในวิธีการด้วย

3.4.2) การแบ่งเฟรมเป็นเฟรมย่อย (Breaking the frame into Subframes)

วิธีการหนึ่งที่ลดการคำนวณในการค้นหา คือ การลดขนาดของโค้ดบुकเช่น ลดจำนวนของดันแบบในโค้ดบुक แต่วิธีการนี้จะเป็นการเพิ่มความผิดพลาดของการควอนไทซ์เวกเตอร์ เพื่อลดความผิดพลาดของการควอนไทซ์นี้ หนึ่งคือการลดความยาว เช่น ลมิตซ์ของดันแบบ อย่างไรก็ตาม วิธีการนี้จะเพิ่มจำนวนบิตที่ต้องการเพราะเราจะต้องการค่าดันแบบที่มากขึ้นเพื่อแทนในส่วนของเมตริกซ์ t วิธีการของ FS 1016 จะแก้ปัญหานี้โดยใช้ค่าความยาวเริ่มต้น 60 แซมเปิ้ล ซึ่งหมายความว่าจะมีการแยกเป้าหมายการค้นหาจากหนึ่งเฟรม แยกเป็นสี่เฟรมของ s ในเฟรมย่อย และวิธีการค้นหา 4 วิธีการ ในการค้นหาและเข้ารหัสเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของเฟรมหนึ่ง ๆ เป็นผลทำให้มีตัวการเข้ารหัสเป็นกลุ่ม 4 กลุ่ม ทำให้ เอสซีบี มีการลดขนาดลดลงจนเป็น 512 เพื่อรักษาคุณภาพของเสียงเดิม

พิจารณาได้ว่าโค้ดบุ้ค เอสซีบี ซึ่งปกติจะแทนหนึ่งอะแดปทีฟแท็บ (Adaptive Tap) และพิทช์ฟิลเตอร์แบบโพลทั้งหมด แต่ขนาดของมันจะไม่ได้ทำการพิจารณาด้วยค่าเหล่านี้ โดยขนาดของเอสซีบีพิจารณาช่วงของความถี่พิทช์ที่มันสามารถครอบคลุมได้ สำหรับการกระตุ้น $x[n]$ ฟิลเตอร์จะสร้าง

$$y[n] = gy[n-d] + x[n] \quad (3.17)$$

ด้วย g ซึ่งเป็นสัมประสิทธิ์ของฟิลเตอร์ ซึ่งเท่ากับอัตราการขยายเอสซีบี และ d จะเป็นตำแหน่งแท็บ ซึ่งเป็นเสมือนเอนทรีของ เอสซีบี การเปลี่ยนค่า d จะมีการเปลี่ยนความถี่พิทช์ (ในหน่วยเฮิรตซ์) ขึ้นอยู่กับสมการ

$$\text{Pitch Frequency} = \frac{\text{Sampling Frequency}}{d} \quad (3.18)$$

FS 1016 จะครอบคลุมความถี่พิทช์ระหว่าง 54 เฮิรตซ์ ถึง 400 เฮิรตซ์ ต้องการค่า d ระหว่าง 20 ถึง 147 ดังนั้นจะใช้ขนาดของ ACB เป็น 128 มาตรฐาน FS 1016 โดยปกติจะให้ขนาด 256 เพื่อปรับปรุงความละเอียดของพิทช์ในย่านความถี่สูง ซึ่งส่วนมากเป็นเสียงของผู้หญิง โดยจากสมการ (18) ซึ่งแสดงว่าความละเอียดของพิทช์ในย่านความถี่สูงจะมีค่าหยาบ ดังนั้นจึงมีการเพิ่มค่าเอนทรีใน ACB เพื่อปรับปรุงคุณภาพของส่วนความถี่สูง แต่เพื่อลดการคำนวณเราจะตัดส่วนประกอบนี้ออกไป

การค้นหาเฟรมย่อยจะสามารถทำให้การเปลี่ยนแปลงสัมประสิทธิ์แอลพีซี เรียบขึ้น โดยการประมาณค่า ดังนั้นในแต่ละเฟรมย่อย $i = 1, \dots, 4$ วิธีการจะใช้ค่า $H(z)$ ที่แตกต่างกันจากการประมาณพารามิเตอร์ของแอลเอสพี โดย

$$\omega_j = \frac{9-2i}{8} \text{Previous} \omega_j + \frac{2i-1}{8} \text{Present} \omega_j \quad (3.19)$$

ดังนั้นระบบจะต้องเก็บพารามิเตอร์แอลเอสพี จากเฟรมที่ผ่านมาเสมอ

3.4.3) การรวมฟิลเตอร์เพอร์เซปชวลและสเปกตรัล (Combining Perceptual and Spectral Filters)

จากสมการที่(14) ของการลดทอน แต่ละชนิดจะต้องผ่านฟิลเตอร์แอลพีซี และฟิลเตอร์ถ่วงน้ำหนัก โดยในแกนของ z จะมีเวกเตอร์การสูญเสียเป็น

$$\begin{aligned} E(z) &= P_w(z) \{ S(z) - \hat{S}(z) \} \\ &= P_w(z) S(z) - P_w(z) H(z) T(z) \\ &= Y(z) - W(z) T(z) \\ &= Y(z) - X(z) \end{aligned} \quad (3.20)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ

$$Y(z) = P_w(z)S(z) \quad (3.21)$$

$$W(z) = P_w(z)H(z) = \frac{H\left(\frac{z}{\gamma}\right)}{H(z)}H(z) = H\left(\frac{z}{\gamma}\right) \quad (3.22)$$

$$X(z) = W(z)T(z) \quad (3.23)$$

และจะได้ค่า

$$\|e\|^2 = \|y - x\|^2 \quad (3.24)$$

จะมีปัญหาเล็กน้อยในวิธีนี้ ถ้าเรากำหนดสมการ (23) โดยใช้เวกเตอร์และเมทริกซ์ ในรูปของเมทริกซ์ โดยฟิลเตอร์ $w(z)$ จะประมาณเมทริกซ์ w แสดงโดย

$$W = \begin{bmatrix} w[0] & 0 & \dots & 0 \\ w[1] & w[0] & \dots & 0 \\ \dots & \dots & \dots & \dots \\ w[59] & w[58] & \dots & w[0] \end{bmatrix} \quad (3.25)$$

เมื่อ $w(i)$ เป็นอิมพัลส์เรสponse (impulse response) ของ $w(z)$ เมื่อ

$$x = Wt \quad (3.26)$$

โดยผลการค้นหาจะมีคุณภาพดีเมื่อตัวส่งเคราะห์ใช้ $H(z)$ ในรูปแบบเมทริกซ์ โดยให้ z ที่ตัวส่งเคราะห์ เช่น $Z[n]$ เป็นเอาต์พุทของ $H(z)$ เมื่อ อินพุทของมันเป็นซีโร สำหรับ เฟรมย่อยทั้งหมด โดยปกติค่า z จะเป็นนอนซีโร (non zero) ขึ้นกับความจูนอนซีโรของ delay element ของ $H(z)$ ซึ่งเป็นผลมาจากการกระตุ้นครั้งที่แล้ว ดังนั้นเสียงจริงจากตัวส่งเคราะห์เป็น

$$\hat{s} = Ht + z \quad (3.27)$$

อุปกรณ์วิเคราะห์จะต้อง ทำการชดเชยจากวิธีการซึ่งเราทำการลดลงไป ดังสมการ (14) แต่ยังคงใช้รวมฟิลเตอร์ w ด้วยสมการ $x = Wt$ จากสมการ $\hat{s} = Ht + z$ เรามี

$$P_w \hat{s} = P_w Ht + P_w z = x + P_w z \quad (3.28)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใช้ผลลัพธ์จากสมการ เราจะได้

$$\begin{aligned} \|c\|^2 &= \|P_w s - x - P_w z\|^2 \\ &= \|P_w (s - z) - x\|^2 \\ &= \|\tilde{y} - x\|^2 \end{aligned} \tag{3.29}$$

ปัจจุบัน \tilde{y} เป็นเป้าหมายการค้นหาใหม่แสดงโดย

$$\tilde{y} = P_w (s - z) \tag{3.30}$$

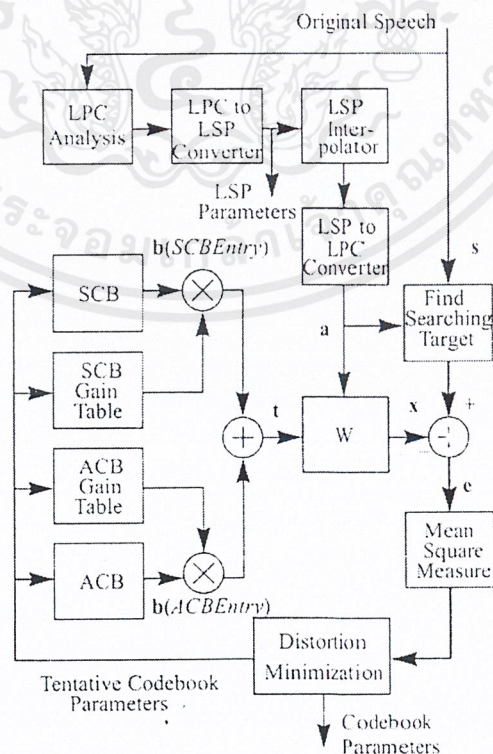
ให้ (CBEntry, GainEntry) เป็นเวกเตอร์ของค่าความผิดพลาดที่ตรวจสอบได้ เกี่ยวข้องกับโค้ดบุ๊ก เอนทรี CBEntry และค่าเอนทรีอัตราขยาย (GainEntry) ลดลงอย่างชัดเจน

$$\begin{aligned} \|c(\text{CBEntry}, \text{GainEntry})\|^2 &= \\ \|\tilde{y} - x(\text{CBEntry}, \text{GainEntry})\|^2 \end{aligned} \tag{3.31}$$

จะเท่ากับการลดสมการ (20) ด้วย z ได้ถูกนำมาพิจารณาด้วย โดยรูป 2 จะเป็นโครงสร้างใหม่

3.4.4) การค้นหาแบบอนุกรม (Serial Search)

เพื่อลดการคำนวณ วิธีการจะค้นหาแบบอนุกรมกับพารามิเตอร์ของ ACB ก่อนพารามิเตอร์ของ SCB โดยระบบจะใช้ 512 และ 128 เอนทรี สำหรับ SCB และ ACB ตามลำดับ และ 16 เอนทรีสำหรับตารางของอัตราขยายแต่ละตาราง ถ้าวิธีการต้องค้นหาโค้ดบุ๊กทั้งหมดในเวลาเดียวกันจะต้องทำการค้นหาใน $512 \times 128 \times 16 \times 16 = 16,777,216$ เอนทรี โดยถ้าทำการค้นหาแบบอนุกรม จะทำการบน $512 \times 16 + 128 \times 16 = 10,240$ เท่านั้น



รูปที่ 3.3 อุปกรณ์วิเคราะห์หีสี่อีแอลพี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นอุปกรณ์ค้นหา ACB และ SCB จะแตกต่างกันในเป้าหมายในการค้นหา โดยเป้าหมายในการค้นหาของ ACB เป็น \tilde{y} ดังที่แสดงในสมการ (30) ค่าที่ได้จาก ACB พารามิเตอร์เพียงตัวเดียวสามารถสร้าง x จากสมการ (26) แต่จะทำให้เกิดค่า $\|e\|_2$ มีค่าสูงโดยพารามิเตอร์ของ SCB ต้องสร้างสัญญาณเติมลงในช่องว่าง ระหว่าง \tilde{y} และ x ดังนั้น $\tilde{y} - wt$ จะกลายเป็นเป้าหมายของการค้นหา SCB เมื่อค่า t ได้จากสมการ (16) ใช้ ACB พารามิเตอร์ที่รับมาใหม่ โดยไม่มีพารามิเตอร์ของ SCB

3.4.5) การค้นหาร่วมกันอย่างมีประสิทธิภาพ (Joint Optimization Search)

วิธีการค้นหาร่วมกันอย่างมีประสิทธิภาพนี้ จะค้นหาโค้ดบุ๊กและเอนทรีของอัตราการขยายนี้ในกระบวนการเดียวกัน จะลดจำนวนของกระบวนการต้นแบบ ในการลด $\|e\|_2$ (CBEntry, GainEntry) โดยระบบจะต้องทำการค้นหาสมาชิกทุกตัวของ CBEntry และ GainEntry อย่างไรก็ตามจอยท์ออปติไมซ์เซชัน (Joint Optimization) จะกำหนดค่าเอนทรีที่มีประสิทธิภาพสำหรับแต่ละ CBEntry ดังนั้นวิธีการจะมีประสิทธิภาพในการค้นหา CBEntry เท่านั้น แทนที่จะต้องทำการค้นหา ใน 10,240 เอนทรี วิธีการนี้จะต้องการการค้นหาบน $512+128 = 640$ เอนทรี ผลของการคำนวณนี้จะทำการบันทึกลำดับในรูปของแมกนิจูด วิธีการพื้นฐานเป็นดังนี้

- 1) สำหรับทุกๆ โค้ดบุ๊กเอนทรีซึ่งเรียกว่า CBEntry จำนวน V ซึ่งบางครั้งเรียกว่า x ที่ถูกนอร์มอลไลซ์ เช่น x ที่ได้มาจากยูนิตเกน (Unit Gain) ตามสมการ

$$v = Wb[\text{CBEntry}] \quad (3.32)$$

ในที่นี้ $b[\text{CBEntry}]$ จะเป็นต้นแบบในโค้ดบุ๊กถูกชี้โดย CBEntry โดยวิธีการนี้จะเรียกว่าการคอนโวลูชัน (convolution)

- 2) สำหรับทุกๆ CBEntry จะคำนวณ GainEntry ร่วมกับ CBEntry เช่นถ้า g เป็นอัตราการขยายซึ่งแบ่งเป็น t อย่างชัดเจน

$$x = gv \quad (3.33)$$

วิธีการหนึ่งเพื่อลดสมการ (31) คือการเพิ่มค่าจุดยอดในเทอมของผลผลิตภายใน (inner-product)

$$\text{Peak} = \langle \tilde{y}, x \rangle - \langle x, x \rangle = g \langle \tilde{y}, v \rangle - g^2 \langle v, v \rangle \quad (3.34)$$

เพื่อค่า g ที่ดีที่สุดเพื่อเพิ่มสมการ (34) จะได้สมการ (35) ด้วยค่า g และรากของมัน รากที่ให้ค่า g ที่ดีที่สุดจะเป็น

$$g = \frac{\langle \tilde{y}, v \rangle}{\langle v, v \rangle} \quad (3.35)$$

นอกจากนี้ ค่า GainEntry เป็นดัชนี ซึ่งค่าในตารางของอัตราขยาย เป็นค่าที่ใกล้เคียงค่า g มากที่สุด

- 3) สำหรับทุกๆ CBEEntry การคำนวณยังคงใช้ค่าสูงสุดใช้ในสมการ (33)
- 4) หา CBEEntry ที่มีระยะทางใกล้ที่สุด ซึ่งคือตัวที่มีค่าพิกสูงสุด ค่า CBEEntry และ GainEntry จะกลายเป็นค่าพารามิเตอร์ในโค้ดบู้คที่ต้องการ

3.4.6) วิธีการที่รวดเร็วสำหรับโค้ดบู้คแบบพิเศษ

วิธีการค้นหาจะใช้วิธีการคอนโวลูชันแบบรวดเร็ว สำหรับการคอนโวลูชันในสมการ (32) โดยตัดคุณสมบัติการซ้อนทับกันของโค้ดบู้คแต่ละชั้น ผลลัพธ์ของการคอนโวลูชันเอนทรีจะสามารถใช้ในการคำนวณคอนโวลูชันของเอนทรีถัดไป ดังนั้นการออกแบบ SCB โดยให้ชั้นเริ่มต้นทุกองค์ประกอบของโค้ดบู้คมาจากอาร์เรย์ γ มี 1082 องค์ประกอบ สมมติว่าองค์ประกอบเบื้องต้นซึ่งโดย CBEEntry (เช่น b (CBEEntry)) จะเป็น $b_{CBEEntry}[i]$ ด้วยค่า $i = 0, \dots, 59$ และให้ค่าองค์ประกอบชั้นเป็น

$$b_{CBEEntry} = [i] = r[2(511 - CBEEntry) + i] \quad (3.36)$$

จะสามารถยืนยันได้ว่าต้นแบบจะมีการซ้อนทับกัน ตัวอย่างเช่น องค์ประกอบส่วนมากของต้นแบบจะเป็นองค์ประกอบของต้นแบบอื่นที่ใกล้เคียงกันด้วย

ด้วย SCB พิเศษนี้จะสามารถหาค่า $v_{CBEEntry}[i]$ ใช้สมการ ดังต่อไปนี้

$$b_{CBEEntry} [i] = \sum_{j=0}^{59} w[j - i] b_{CBEEntry} [j] \quad (3.37)$$

เพื่อทำให้ง่ายขึ้นจะกำหนดค่า $u(CBEEntry, i, j)$ เป็น

$$u(CBEEntry, i, j) = w[j - i] b_{CBEEntry} [j] = w[j - i] r[2(511 - CBEEntry) + j] \quad (3.38)$$

ซึ่งจะได้

$$V_{CBEntry} = \sum_{j=0}^1 u(CBEntry, i, j) + \sum_{j=2}^{61} u(CBEntry, i, j) - \sum_{j=60}^{61} u(CBEntry, i, j) \quad (3.39)$$

เราจะได้อ่า จะสามารถยืนยันได้อย่างง่ายดายว่าเทอมกลางจะเป็น $V_{CBEntry-1}[i]$ เพราะคุณสมบัติการซ้อนทับกัน ด้วยเหตุผลนี้ปัจจุบันแทนที่จะแสดง 60 เทอมของการคูณและการคำนวณ (Multiple and accumulate : MAC) ตามสมการ (37) วิธีการนี้จะทำการคำนวณ $V_{CBEntry-1}[i]$ ใน 4 เทอมเท่านั้นเพื่อรับค่าหัวและท้าย และใช้การคำนวณครั้งก่อนหน้าค่านวณ $V_{CBEntry-1}[i]$ ในเทอมกลาง จะสามารถลดขนาดการคำนวณลงได้จนถึง 15 เท่า

เรายังสามารถหลีกเลี่ยงการคำนวณในเทอมท้ายถ้าเราสามารถหาค่าอาร์เรย์ยาว V' และ อาร์เรย์สั้น V'' ซึ่งมีความยาว 1082 และ 60 ตามลำดับซึ่งแสดงในวิธีการต่อไปนี้

1. เริ่มต้นด้วยการคำนวณ $V_0[i]$ ใช้วิธีการเดิมโดยใช้สมการ (37) เป็นจุดเริ่มต้นของการทำ และทำการเก็บค่าผลลัพธ์ลงในเมตริกซ์ V' ซึ่งว่างอยู่ดังสมการ

$$v'[I] = v_0[I] \quad I = 0, \dots, 59 \quad (3.40)$$

2. คำนวณจุดยอดและลำดับเอนทรีของอัตราขยายและเก็บค่าจุดยอดที่ดีที่สุดและเอนทรีของอัตราขยายที่ดีที่สุดตามลำดับ และเก็บค่าเอนทรีของโค้ดบู้กลงในเอนทรีของโค้ดบู้ที่ดีที่สุด ซึ่งในกรณีนี้เป็น 0 ดังนั้นสำหรับทุก ๆ โค้ดบู้เอนทรี = 1, ..., 511
3. คำนวณเทอมหัว 60 เทอม และเก็บใน V''
4. เปลี่ยนค่าอาร์เรย์ V' ขึ้นอยู่กับสมการ

$$v'[i + 2CBEntry] \leftarrow V'[i + 2CBEntry] + V'' \quad (3.41)$$

5. คำนวณจุดยอดและเอนทรีของอัตราขยาย โดยจะได้ค่า V จากค่า V' ดังสมการ

$$v'[i] = V'[i + 2CBEntry]; \quad I = 0, \dots, 59 \quad (3.42)$$

6. เปรียบเทียบจุดยอดด้วยค่าตัวเลขค่าสุดยอดที่ดีที่สุด โดยเริ่มต้นเป็น 0 ถ้าจุดยอดปัจจุบันมีขนาดใหญ่กว่าจุดยอดที่ดีที่สุด จะต้องกำหนดค่าจุดยอดนั้นแทนจุดยอดที่ดีที่สุดเดิม และเก็บค่า $CBEntry$ และ $GainEntry$ ด้วยค่า $CBEntry$ และ $GainEntry$ ที่ดีที่สุดตามลำดับ

เมื่อทำตามขั้นตอนดังกล่าวแล้ว สำหรับทุก ๆ เอนทรี ค่าพารามิเตอร์ที่ต้องการจะมาอยู่ในค่า $CBEntry$ และ $GainEntry$ ที่ดีที่สุด

ดังนั้นการลดการคำนวณในวิธี FS 1016 ตัว SCB จะใช้ $b_{CBEntry}[i]$ ที่ไม่เพียงแต่ซ้อนทับกัน แต่จะมีค่าน้อยมากโดยค่าส่วนใหญ่เป็น 0 ถึง 77 เปอร์เซ็นต์ และประกอบด้วยเลขเพียง 3 ตัว คือ -1, 0, 1 เท่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นก่อนการคำนวณ เทอมหัวในขั้นที่ 3 ด้านบน จะทำการตรวจสอบว่า $bCBEntry[j]$ เป็นศูนย์หรือไม่ โดย 60 การคำนวณของเทอมที่ใช้ $bCBEntry$ นี้ในขั้นตอนที่ 3 จะถูกข้ามไปถึง 77 เปอร์เซนต์

และด้วย $bCBEntry[j]$ ซึ่งประกอบด้วยค่าเพียง 3 ค่าเท่านั้นในการคำนวณการคูณในเทอมของหัว จะไม่จำเป็นอีกต่อไป เพราะจะคูณด้วย 1 และ -1 ซึ่งเท่ากับเป็นการเปลี่ยนเครื่องหมายเท่านั้น

เราสามารถลดการคำนวณของการหาพารามิเตอร์ของโค้ดบู้คแบบอะแดปทีฟได้โดยการพิจารณาคุณลักษณะของเสียงของมนุษย์ โดยค่าพิทช์จะไม่มีเปลี่ยนแปลงอย่างทันทีทันใดในเวลาที่ 2 เฟรมย่อย หรือ 15 มิลลิวินาที ซึ่งหมายความว่า ความแตกต่างระหว่างโค้ดบู้คเอนทรีที่ถูกเลือกมาในเฟรมย่อยที่ติด ๆ กันจะมีค่าน้อยกว่า 64 เอนทรี ดังนั้นเราสามารถใช้การเข้ารหัสเซลล์ที่เข้ารหัสเฉพาะส่วนของความแตกต่างของเอนทรีเท่านั้น และทุก ๆ การเข้ารหัสจะต้องการจุดอ้างอิง โดยตามมาตรฐาน FS 1016 จะใช้โค้ดบู้คแบบอะแดปทีฟในเฟรมย่อยที่เป็นคี่ในการอ้างอิง และใช้การเข้ารหัสแบบเซลล์ในเฟรมย่อยที่เป็นเฟรมคู่ ยกตัวอย่างเช่น เอนทรีในเฟรมย่อยที่ 2 และเฟรมย่อยที่ 4 จะแทนด้วยค่าความแตกต่างระหว่างเอนทรีจริงและเอนทรีของเฟรมย่อยที่แล้ว

วิธีการนี้จะลดการคำนวณได้เพราะ การค้นหาในเฟรมคู่จะทำบนกลุ่มย่อยของโค้ดบู้คแบบอะแดปทีฟเท่านั้น (64 เอนทรีแทนที่ของเดิม 128 เอนทรี) โดยวิธีการนี้จะลดจำนวนบิตที่ใช้ลงด้วย เมื่อจำนวนบิตที่แทนความแตกต่างของค่าจะมีการใช้บิตในการแสดงค่าน้อยกว่าการแทนเอนทรีจริง ๆ

3.5 สังเคราะห์เสียง (Speech Synthesis)

3.5.1) กระบวนการการสังเคราะห์ (Synthesis Process)

อุปกรณ์สังเคราะห์สัญญาณซีอีแอลพี จะสร้างเมตริกซ์ s ซึ่งแทน 240 แซมเปิลจากกลุ่มของพารามิเตอร์ซีอีแอลพี โดยตัวสังเคราะห์จะต้องสร้างฟิลเตอร์ $H(z)$ ใช้ประมาณค่าพารามิเตอร์แอลเอสพี และคำนวณพัลส์การกระตุ้น t สำหรับเฟรมย่อยหนึ่งใช้พารามิเตอร์ในโค้ดบู้คตามสมการที่ 17 และทำการใส่ค่าอิมพัลส์การกระตุ้น t ไปในฟิลเตอร์ $H(z)$ เพื่อสังเคราะห์ 60 องค์กรประกอบของเสียง \hat{s} ใช้สมการ (2) และทำตามขั้นตอนที่ 3 ป้อยขึ้นจนครบ 240 องค์กรประกอบ \hat{s} ซึ่งได้อธิบายไปแล้วก่อนหน้านี้และทำการแปลงแอลเอสพี เป็นพารามิเตอร์แอลพีซี ดังหัวข้อต่อไปนี้

3.5.2) การแปลงแอลเอสพีเป็นแอลพีซี

ต้องการสร้างเมตริกซ์ a จากการประมาณค่า LSPs ω_j โดยจะแสดง Xp และ Xq ดังนี้

$$\begin{aligned} Xp_i &= \omega_{2i} \\ Xq_i &= \omega_{2i+1} \end{aligned} \quad (3.43)$$

และทำการแปลงดังขั้นตอนต่อไปนี้

1. สร้างอาร์เรย์ b กลับมาดังสมการ (10) ตามสมการต่อไปนี้

$$\begin{aligned}
 b_5 &= 32 \\
 b_4 &= -b_5 \sum_{i=0}^4 xp_{i+1} \\
 b_3 &= b_5 \sum_{i=1}^4 \sum_{j=i+1}^5 xp_i xp_j \\
 b_2 &= -b_5 \sum_{i=0}^3 \sum_{j=i+1}^4 \sum_{n=j+1}^5 xp_i xp_j xp_n \\
 b_1 &= b_5 \sum_{i=1}^2 \sum_{j=i+1}^3 \sum_{m=j+1}^4 \sum_{n=m+1}^5 xp_i xp_j xp_m xp_n \\
 b_0 &= -b_5 (xp_1 xp_2 xp_3 xp_4 xp_5)
 \end{aligned} \tag{3.44}$$

2. สร้างพารามิเตอร์ $p(z)$ กลับมาตามสมการดังต่อไปนี้

$$\begin{aligned}
 p_1 &= \frac{b_4}{16} \\
 p_2 &= \frac{b_3 + 40}{8} \\
 p_3 &= \frac{b_2 + 16p_1}{4} \\
 p_4 &= \frac{b_1 + 6p_2 - 10}{2} \\
 p_5 &= b_0 + 2p_3 - 2p_1
 \end{aligned} \tag{3.45}$$

3. สร้างอาร์เรย์ c กลับมาดังสมการ (11) ตามสมการต่อไปนี้

$$\begin{aligned}
 c_5 &= 32 \\
 c_4 &= -c_5 \sum_{i=0}^4 xq_{i+1} \\
 c_3 &= c_5 \sum_{i=1}^4 \sum_{j=i+1}^5 xq_i xq_j \\
 c_2 &= -c_5 \sum_{i=0}^3 \sum_{j=i+1}^4 \sum_{n=j+1}^5 xq_i xq_j xq_n \\
 c_1 &= c_5 \sum_{i=1}^2 \sum_{j=i+1}^3 \sum_{m=j+1}^4 \sum_{n=m+1}^5 xq_i xq_j xq_m xq_n \\
 c_0 &= -c_5 (xq_1 xq_2 xq_3 xq_4 xq_5)
 \end{aligned} \tag{3.46}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. จะได้ค่าชุดของพารามิเตอร์ $q(z)$ ตามสมการต่อไปนี้

$$\begin{aligned}
 q_1 &= \frac{c_4}{16} \\
 q_2 &= \frac{c_3 + 40}{8} \\
 q_3 &= \frac{c_2 + 16q_1}{4} \\
 q_4 &= \frac{c_1 + 6q_2 - 10}{2} \\
 q_5 &= c_0 + 2q_3 - 2q_1
 \end{aligned} \tag{3.47}$$

5. และขั้นตอนสุดท้ายใช้ p และ q สร้าง a โดยส่วนกลับของสมการ (12) ดังต่อไปนี้

$$\begin{aligned}
 a_0 &= 1 \\
 p_0 &= q_0 = 1 \\
 a_i &= \frac{p_{i-1} + p_i + q_i - q_{i-1}}{2}; \quad i = 1, \dots, 5 \\
 a_{11-i} &= \frac{p_{i-1} + p_i + q_i - q_i}{2}; \quad i = 1, \dots, 5
 \end{aligned} \tag{3.48}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

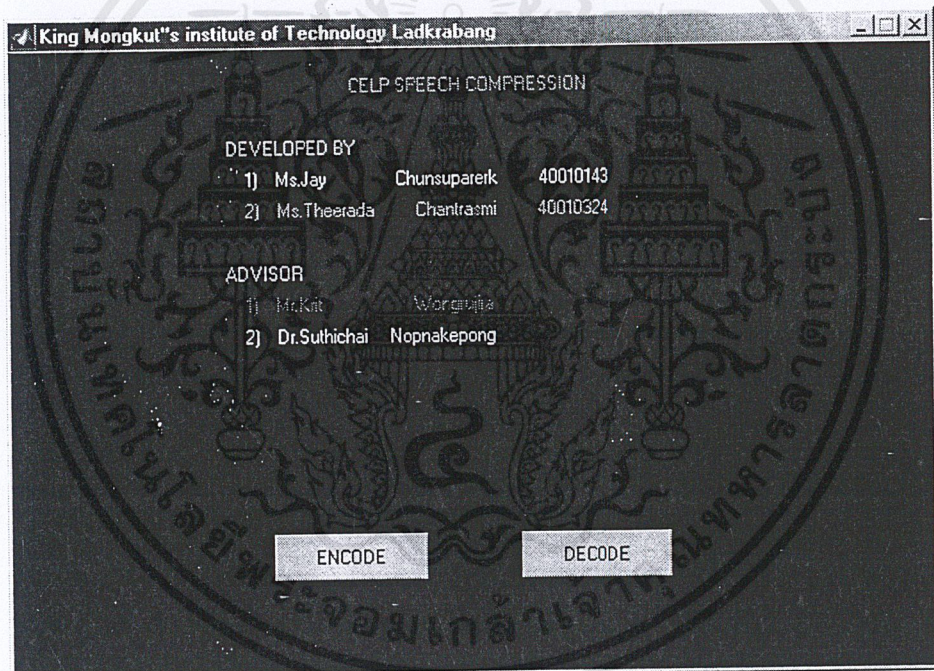
การทดลองและผลการทดลอง

การทดลองที่ 4.1 ทดสอบโปรแกรมการติดต่อกับผู้ใช้และส่วนแสดงผลการทำงานในส่วนต่างๆ จุดประสงค์การทดลอง เพื่อทดสอบโปรแกรมส่วนที่ใช้ในการติดต่อกับผู้ใช้เพื่อทำการรับไฟล์เสียงจากทางผู้ใช้ และให้ผู้ใช้ทำการเลือกปลายทาง รวมถึงส่วนแสดงผลการทำงานในขณะนั้นของโปรแกรมการแปลง CELP อีกด้วย

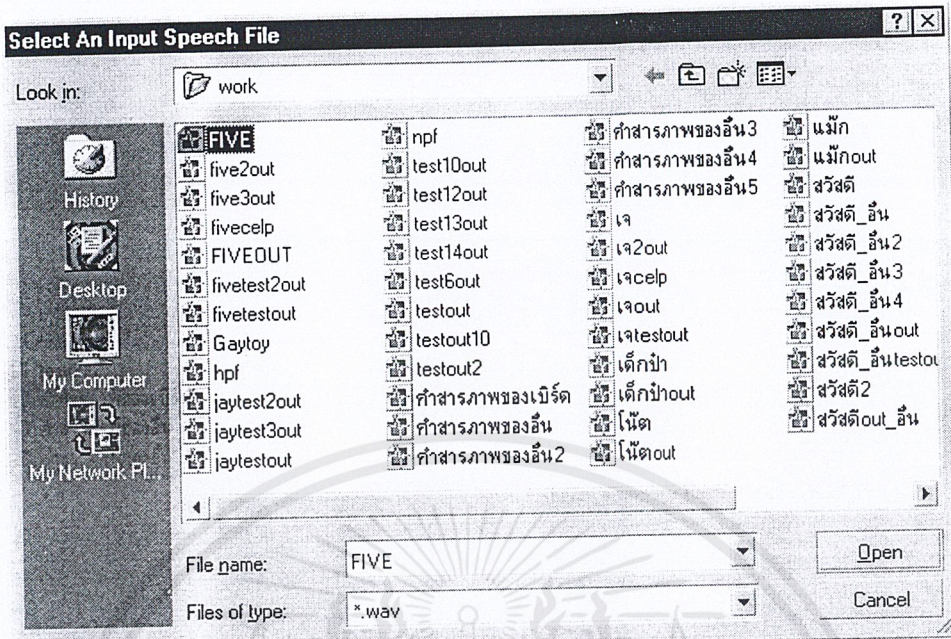
ขั้นตอนการทดลอง

1. เรียกโปรแกรม MATLAB
2. เรียกโปรแกรมการติดต่อกับผู้ใช้ซึ่งเขียนไว้เป็นฟังก์ชัน (guiinit(screen))

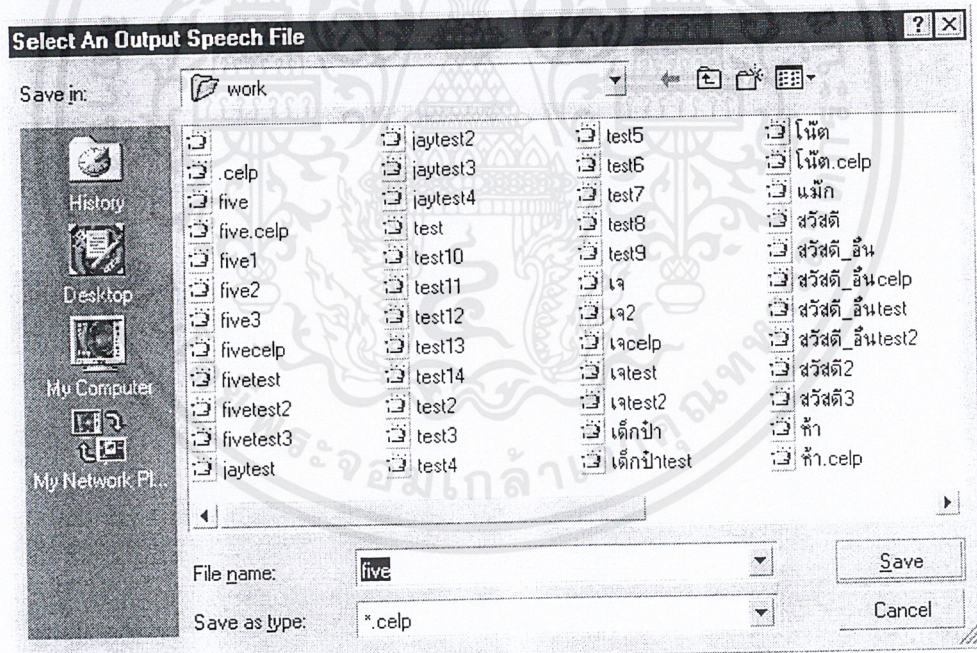
ผลการทดลอง



รูปที่ 4.1 แสดงหน้าจอการเริ่มต้นการทำงานโดยให้เลือกเข้ารหัสหรือถอดรหัส

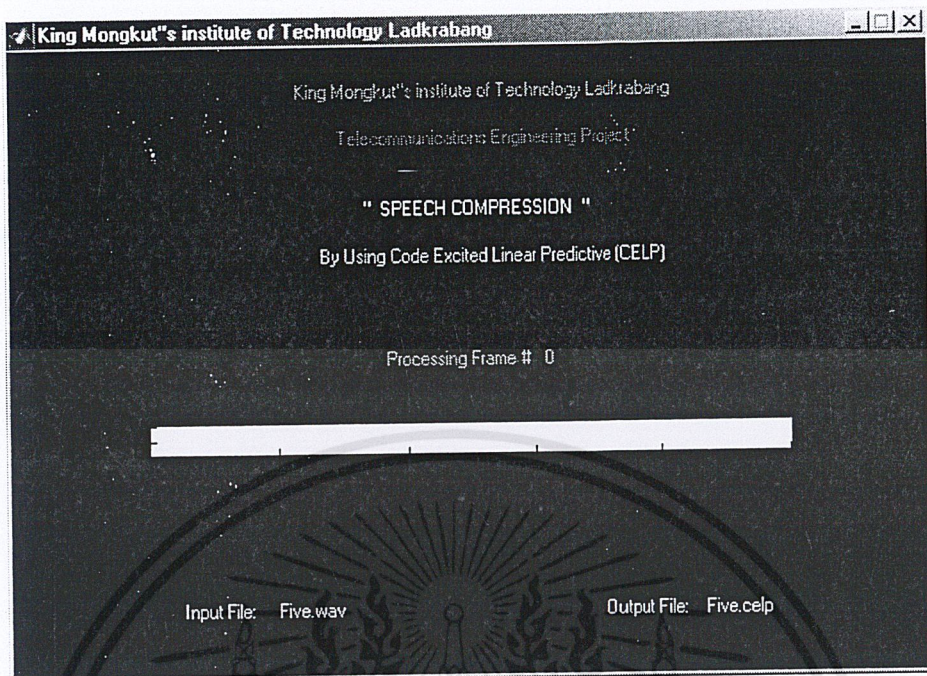


รูปที่ 4.2 แสดงหน้าจอการทำงานถ้าเลือกการเข้ารหัส โดยจะให้เลือกเสียงที่ต้องการเข้ารหัส

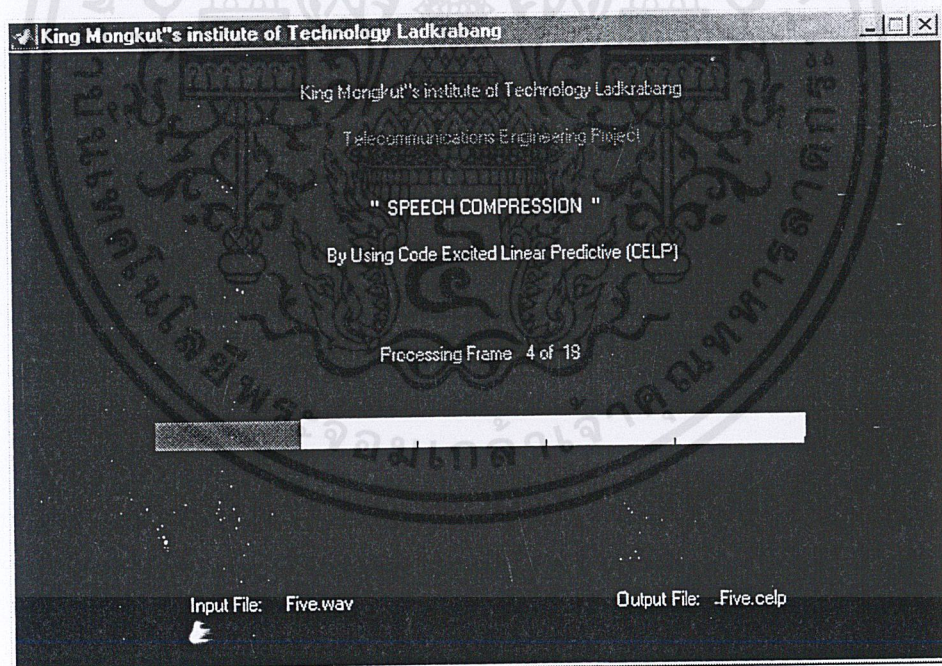


รูปที่ 4.3 แสดงหน้าจอการทำงานถ้าเลือกการเข้ารหัส โดยจะให้พิมพ์ชื่อไฟล์ที่ต้องการเก็บเสียงนั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

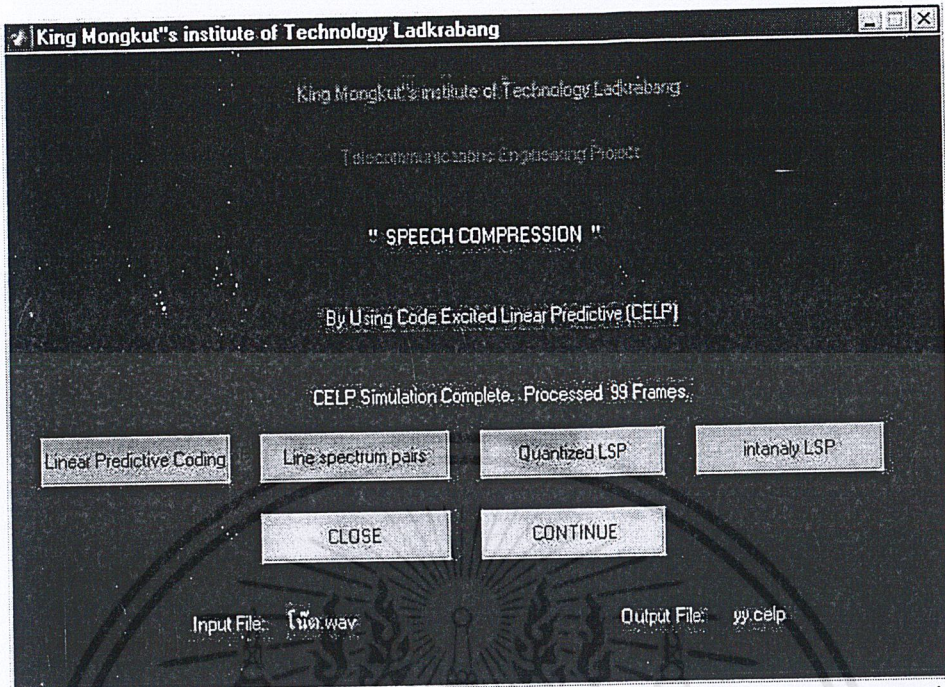


รูปที่ 4.4 แสดงหน้าจอขณะกำลังเริ่มทำการทำงาน

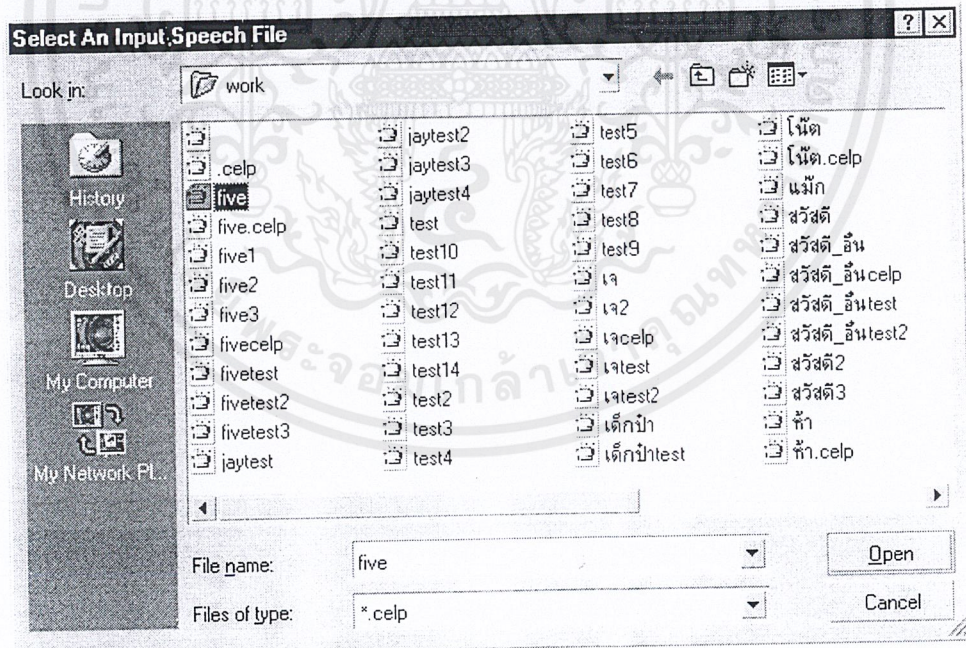


รูปที่ 4.5 แสดงหน้าจอขณะกำลังทำการทำงาน โดยจะมีการแสดงสถานะการทำงานในแต่ละเฟรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

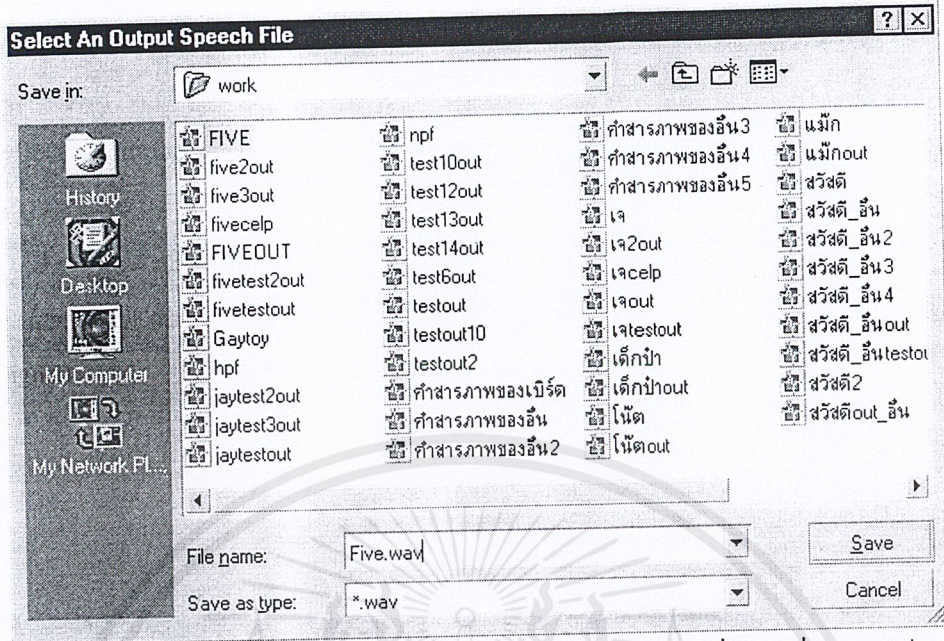


รูปที่ 4.6 แสดงหน้าจอเมื่อเสร็จสิ้นการเข้ารหัสโดยจะให้เลือกทำการทำงานต่อ ทดสอบสัญญาณที่ได้ในขั้นตอนต่างๆ หรือ ออกจากโปรแกรม

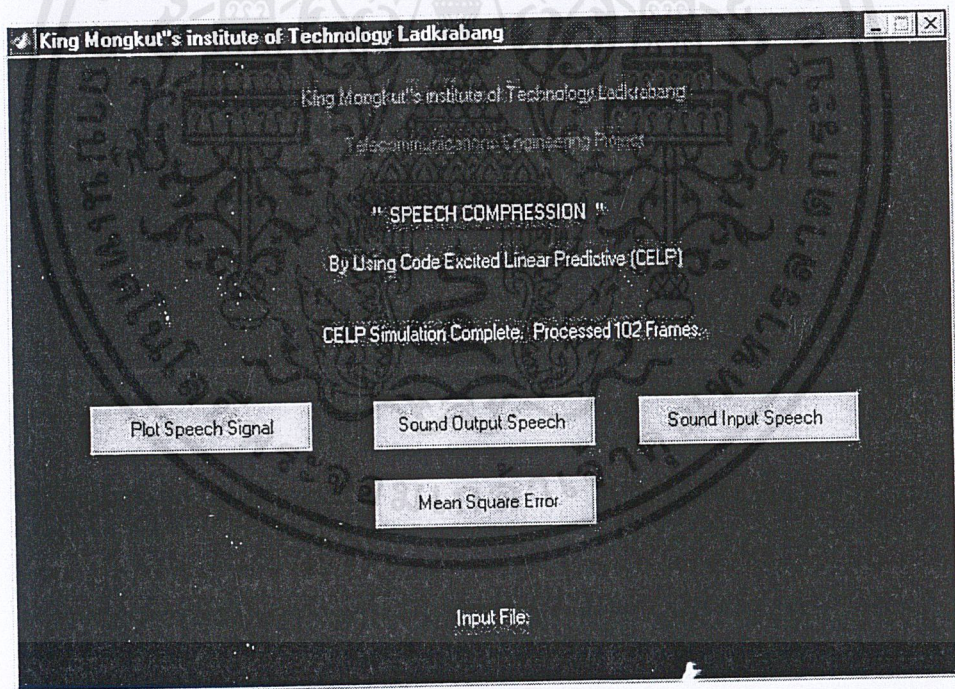


รูปที่ 4.7 แสดงหน้าจอการการทำงานถ้าเลือกการถอดรหัส โดยจะให้เลือกไฟล์(.celp)ที่ต้องการถอดรหัส

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.8 แสดงหน้าจอการทำงานถ้าเลือกการถอดรหัส โดยจะให้พิมพ์ชื่อไฟล์ที่ต้องการเก็บเสียงนั้น



รูปที่ 4.9 แสดงหน้าจอเมื่อเสร็จสิ้นการถอดรหัสโดยจะให้เลือกทำการทดสอบเสียงด้วยการพล็อต หรือทดลองโดยการฟังเสียง หรือออกจากโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทดลองที่ 4.2 ตรวจสอบผลการบีบอัดโดยการดูรูปของสัญญาณที่ถอดรหัสได้เปรียบเทียบกับสัญญาณอินพุทที่รับเข้ามาทั้งในแกนเวลาและแกนความถี่

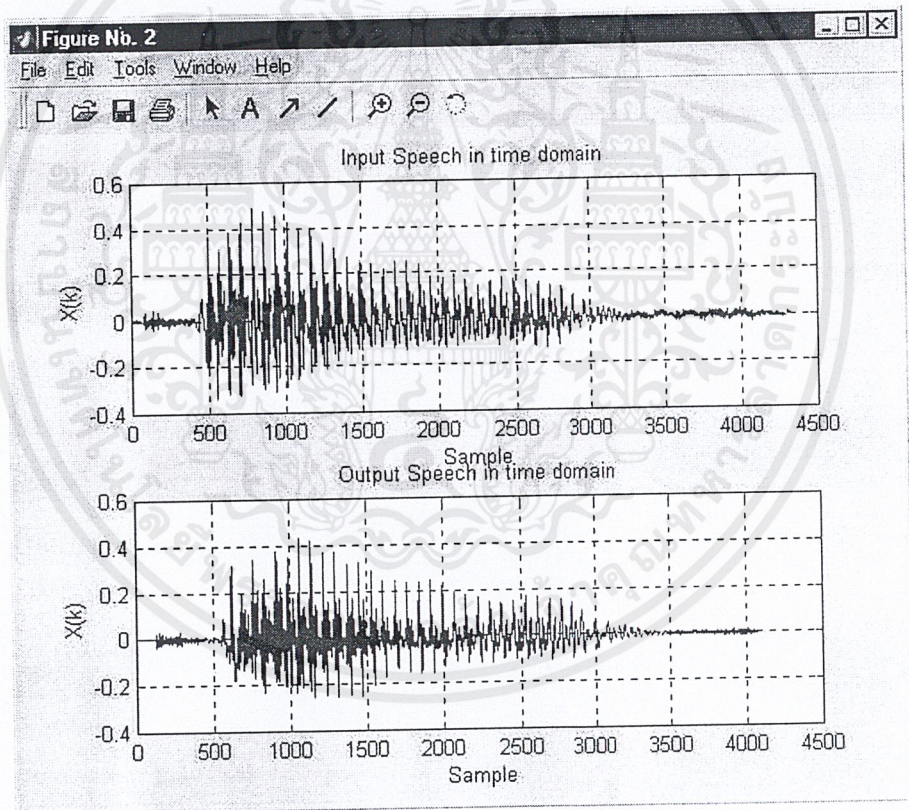
จุดประสงค์การทดลอง เพื่อตรวจสอบคุณภาพของรูปสัญญาณว่าแตกต่างจากต้นแบบไปมากน้อยเพียงไร

ขั้นตอนการทดลอง 1. ทำการรับข้อมูลเสียงทดสอบเข้ามา และใช้โปรแกรมทำการแปลงให้เป็นไฟล์ที่ถูก

บีบอัด (.celp) และทำการแปลงกลับเป็นไฟล์เสียงเหมือนเดิม

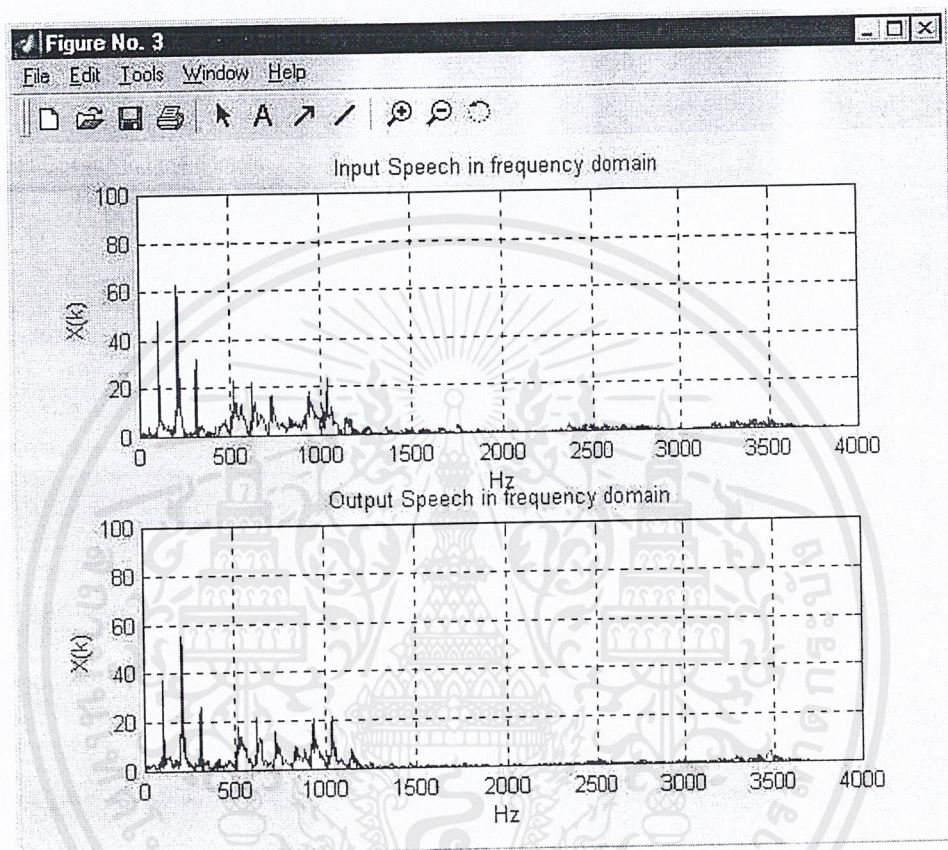
2. ทำการพล็อตดูรูปของสัญญาณจากหน้าจอ โดยคลิกที่คำสั่งให้พล็อต (Plot output speech signal)

ผลการทดลอง



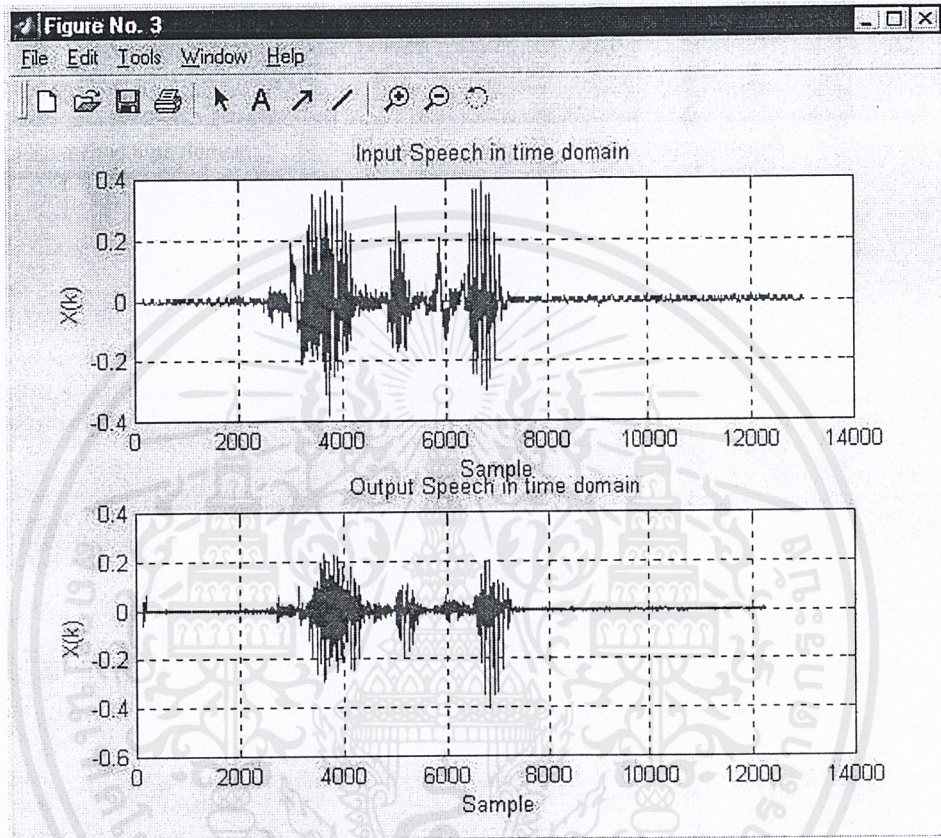
รูปที่ 4.10 รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุทหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุท โดยแสดงในแกนเวลา ของสัญญาณทดสอบ (Five.wav) ขนาด 9 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



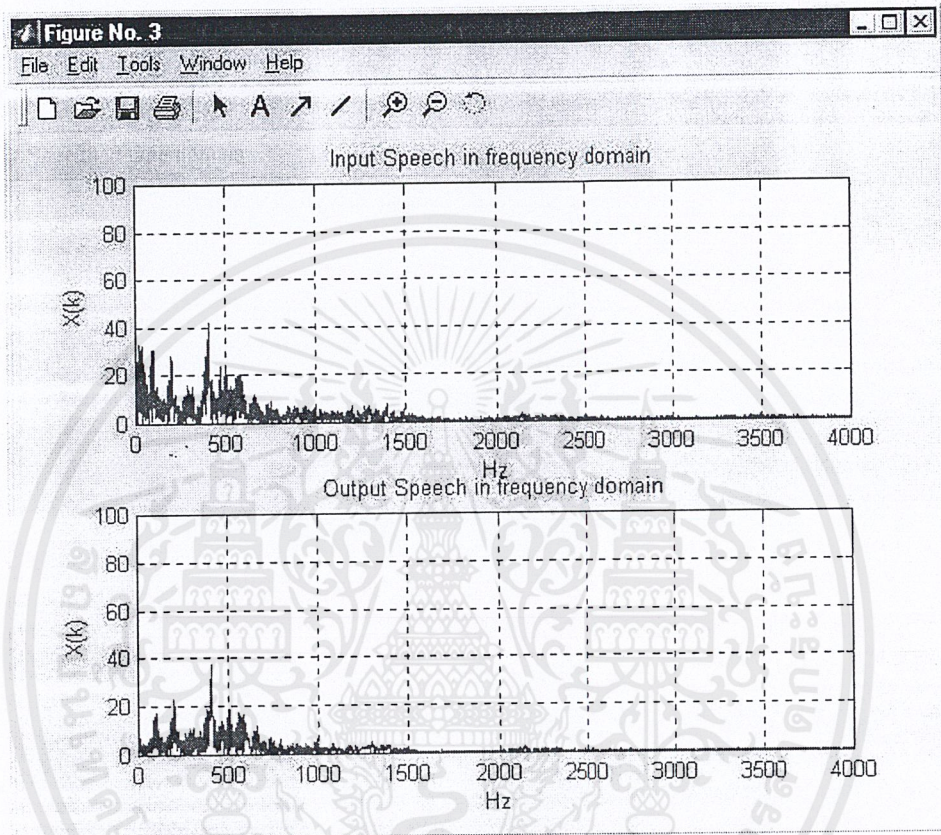
รูปที่ 4.11 รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุตหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุต โดยแสดงในแกนความถี่ ของสัญญาณทดสอบ (Five.wav) ขนาด 9 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



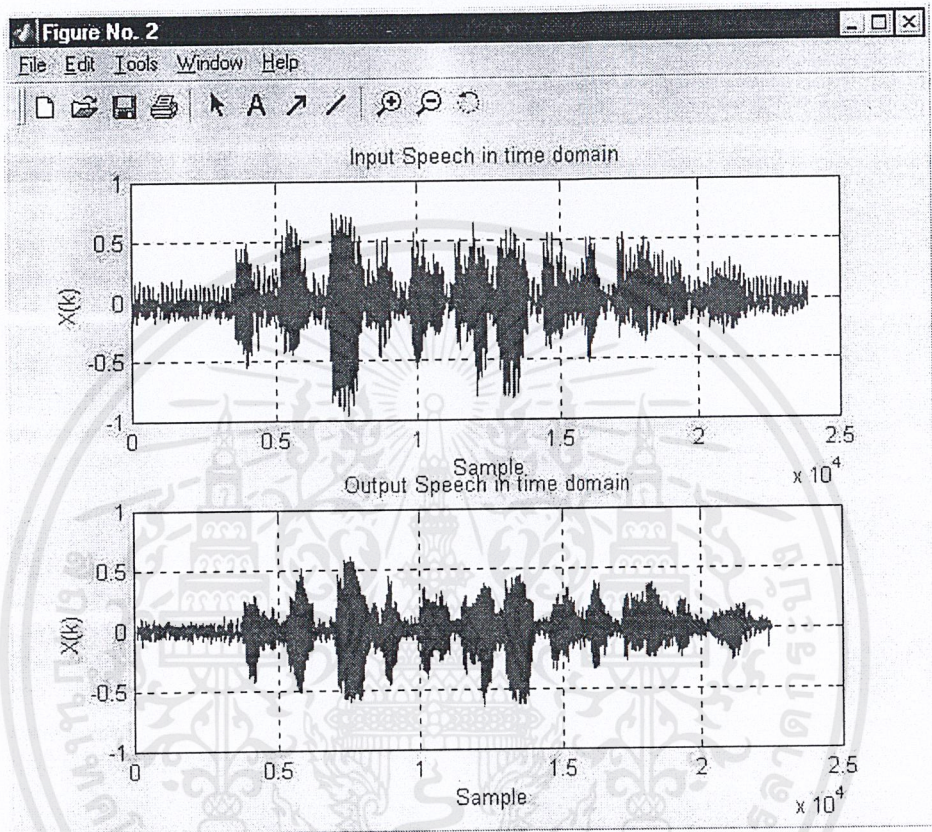
รูปที่ 4.12 รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุทหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุท โดยแสดงในแกนเวลา ของสัญญาณทดสอบ (สวัสดี_อิน.wav) ขนาด 26 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



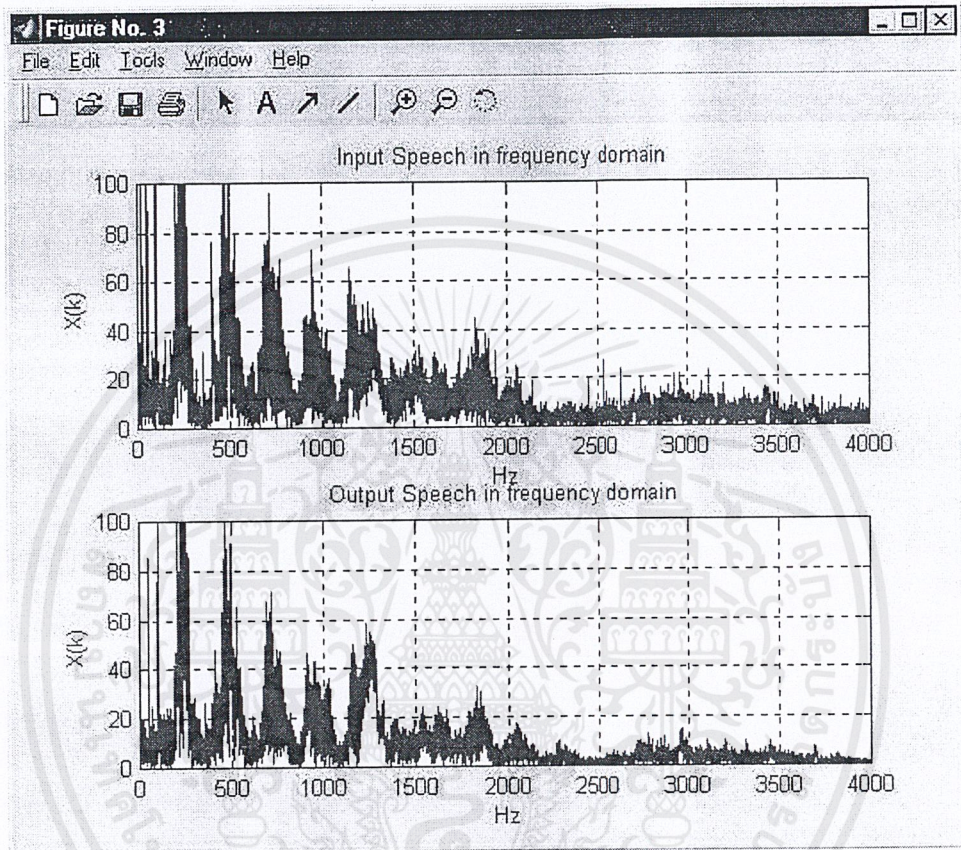
รูปที่ 4.13 รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุตหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุต โดยแสดงในแกนความถี่ ของสัญญาณทดสอบ (สวัสดี_อิน.wav) ขนาด 26 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



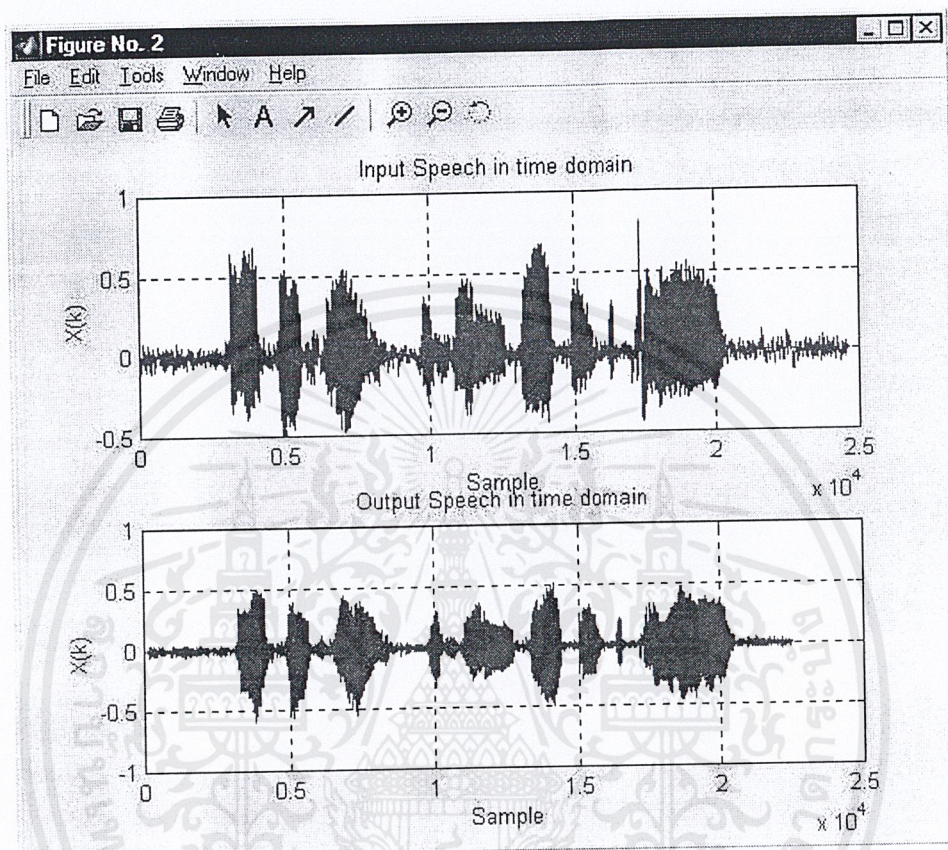
รูปที่ 4.14 รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุทหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุท โดยแสดงในแกนเวลา ของสัญญาณทดสอบ (ไฟล์.wav) ขนาด 47 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



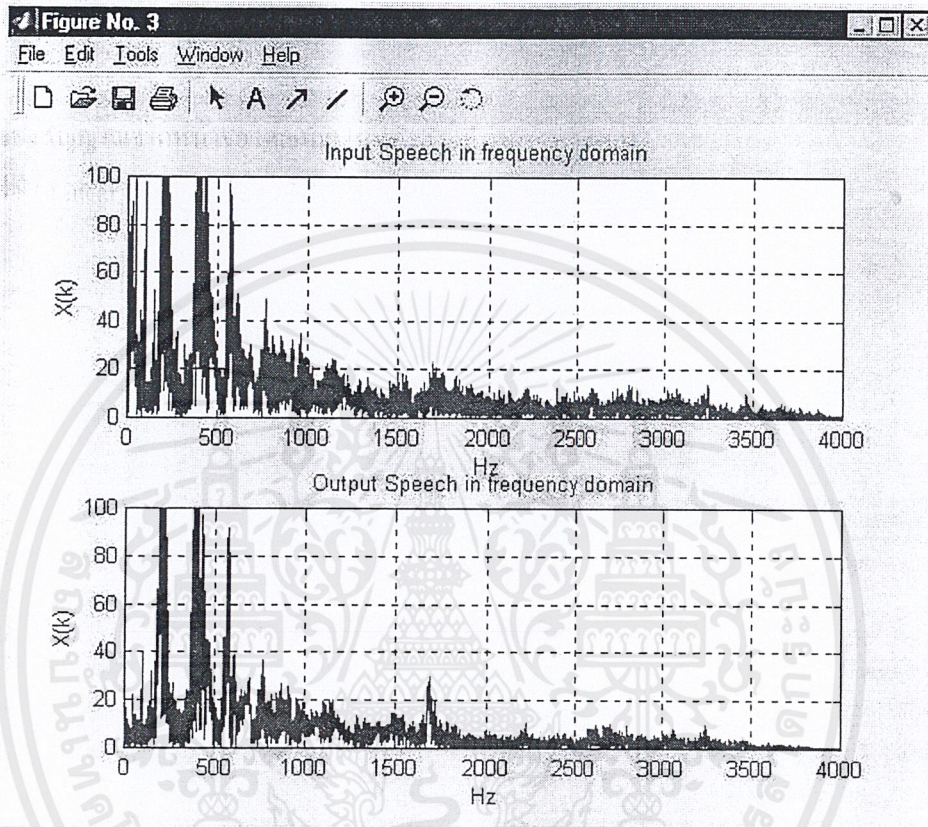
รูปที่ 4.15 รูปแสดงการเปรียบเทียบสัญญาณเอาท์พุทหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุท โดยแสดงในแกนความถี่ ของสัญญาณทดสอบ (โน้ต.wav) ขนาด 47 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.16 รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุทหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุท โดยแสดงในแกนเวลา ของสัญญาณทดสอบ (เจ.wav) ขนาด 49 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



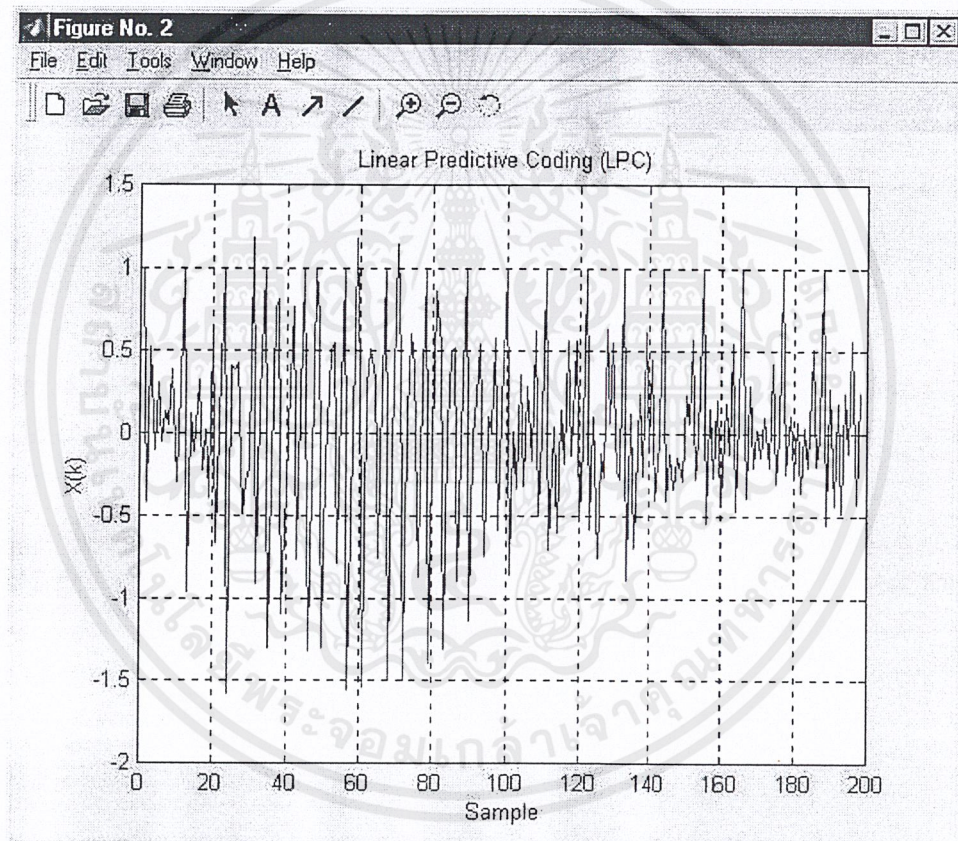
รูปที่ 4.17 รูปแสดงการเปรียบเทียบสัญญาณเอาต์พุตหลังจากการบีบอัด และคลายข้อมูลออกมาเปรียบเทียบกับอินพุท โดยแสดงในแกนความถี่ ของสัญญาณทดสอบ (เจ.wav) ขนาด 49 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทดลองที่ 4.3 ตรวจสอบค่าสัมประสิทธิ์การเข้ารหัสแบบประมาณเชิงเส้น (แอลพีซี)

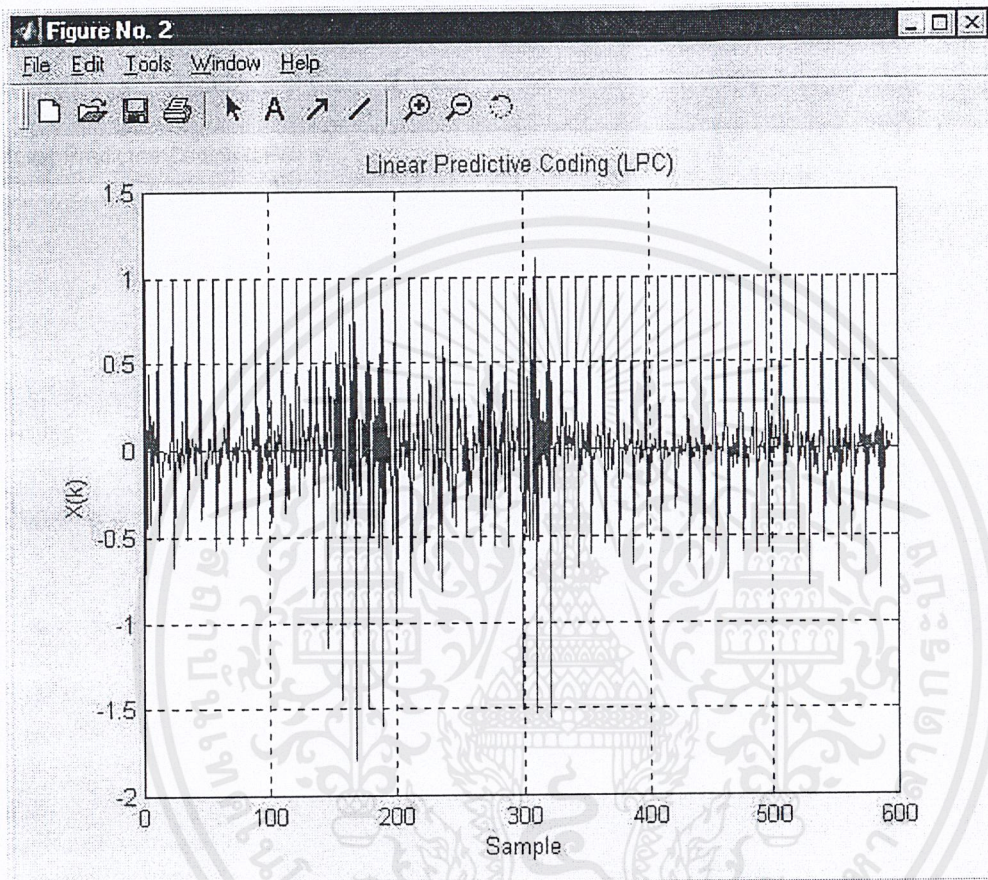
จุดประสงค์การทดลอง เพื่อตรวจสอบค่าสัมประสิทธิ์การเข้ารหัสแบบประมาณเชิงเส้น (แอลพีซี) ของสัญญาณเสียงไฟล์ต่างๆที่รับเข้ามา เพื่อแสดงผลการดำเนินงาน และ ตรวจสอบความถูกต้องโดยการดูรูปของสัญญาณ

- ขั้นตอนการทดลอง
1. ทำการรับข้อมูลเสียงทดสอบเข้ามา และ ใช้โปรแกรมทำการแปลงให้เป็นสัญญาณการประมาณเชิงเส้น โดยทำที่ละเฟรมและเก็บค่าผลรวม
 2. ทำการพล็อตดูรูปของสัญญาณจากหน้าจอโดยคลิกไปที่คำสั่งให้ทำการพล็อตค่าสัมประสิทธิ์แอลพีซี (Linear Predictive Coding)



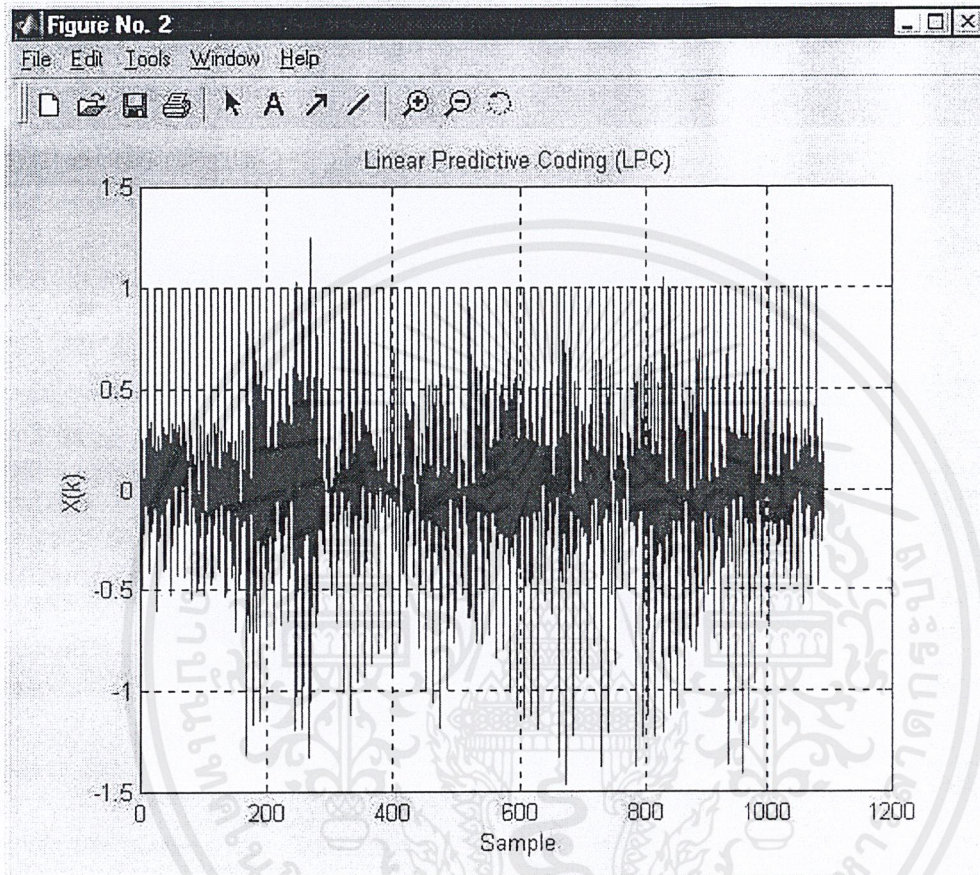
รูปที่ 4.18 แสดงค่าสัมประสิทธิ์การเข้ารหัสแบบประมาณเชิงเส้นของเสียงทดสอบ (Five.wav) ขนาด 9 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



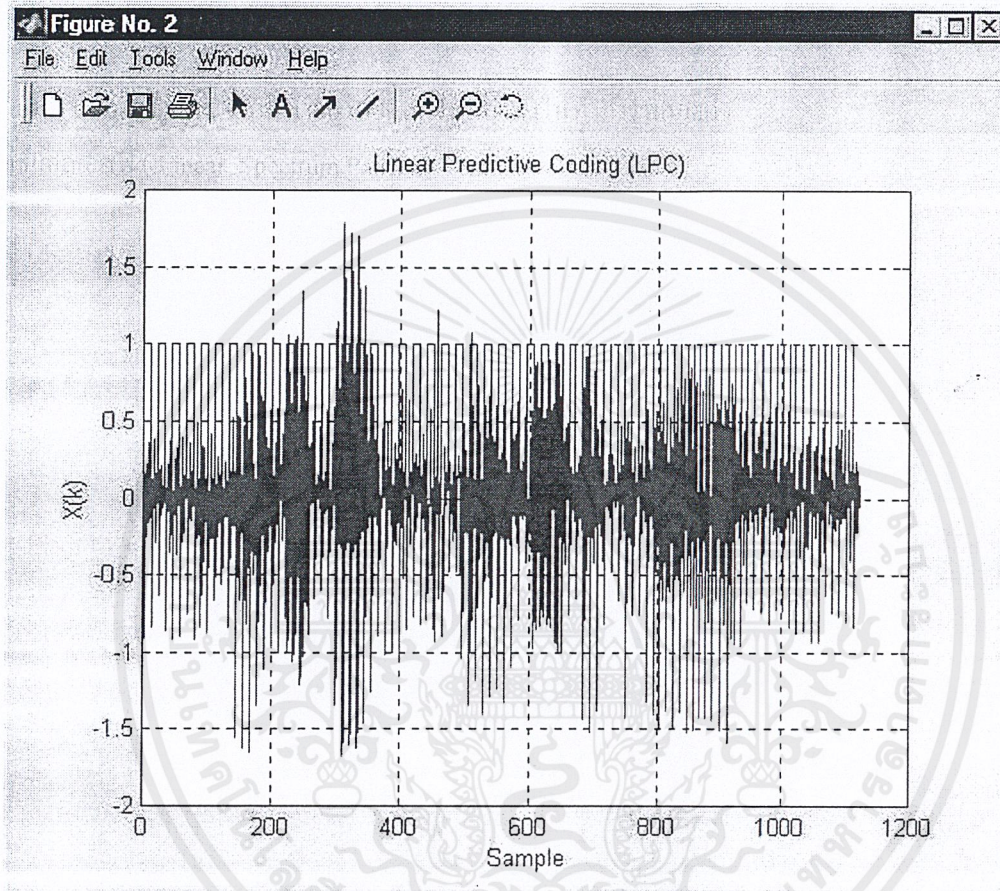
รูปที่ 4.19 แสดงค่าสัมประสิทธิ์การเข้ารหัสแบบประมาณเชิงเส้นของเสียงทดสอบ (สวัสดี_อัน.wav)
ขนาด 26 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.20 แสดงค่าสัมประสิทธิ์การเข้ารหัสแบบประมาณเชิงเส้นของเสียงทดสอบ (ไฟล์.wav)
ขนาด 47 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.21 แสดงค่าสัมประสิทธิ์การเข้ารหัสแบบประมาณเชิงเส้นของเสียงทดสอบ (เจ.wav)
ขนาด 49 กิโลไบต์

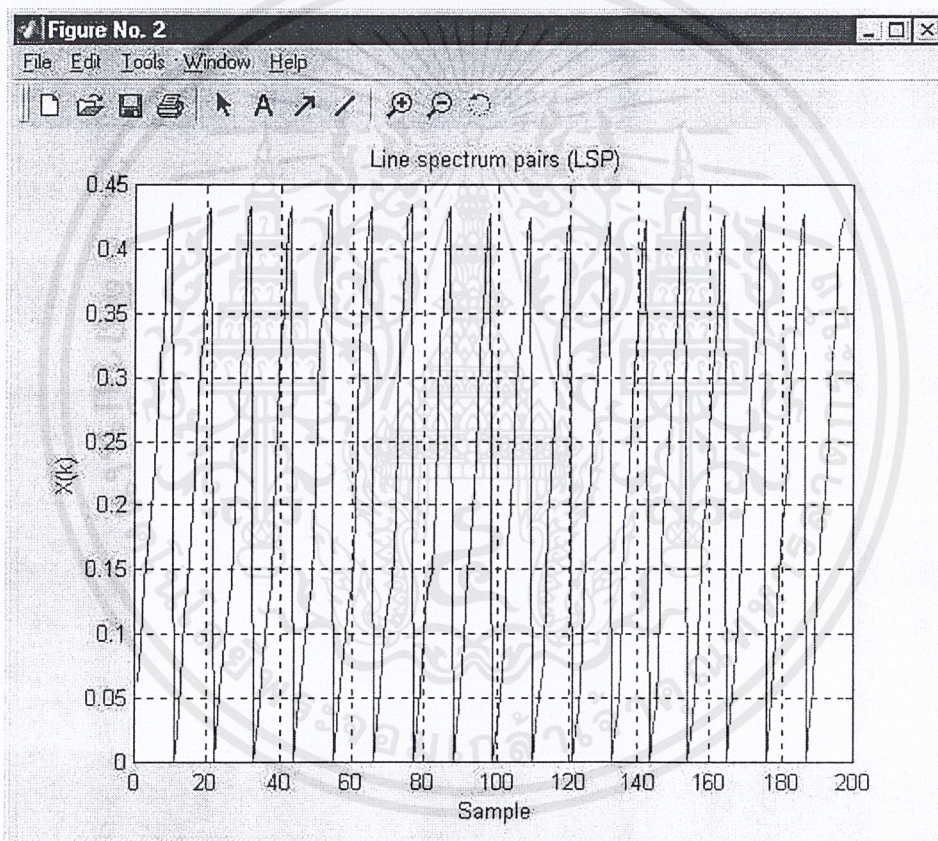
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทดลองที่ 4.4 ตรวจสอบค่าสัมประสิทธิ์ไลน์สเปกตรัมแพร์ (แอลเอสพี)

จุดประสงค์การทดลอง เพื่อตรวจสอบค่าสัมประสิทธิ์ไลน์สเปกตรัมแพร์ (แอลเอสพี) ของสัญญาณเสียงไฟล์ต่างๆที่รับเข้ามา เพื่อแสดงผลการดำเนินงาน และ ตรวจสอบความถูกต้องโดยการดูรูปของสัญญาณ

ขั้นตอนการทดลอง

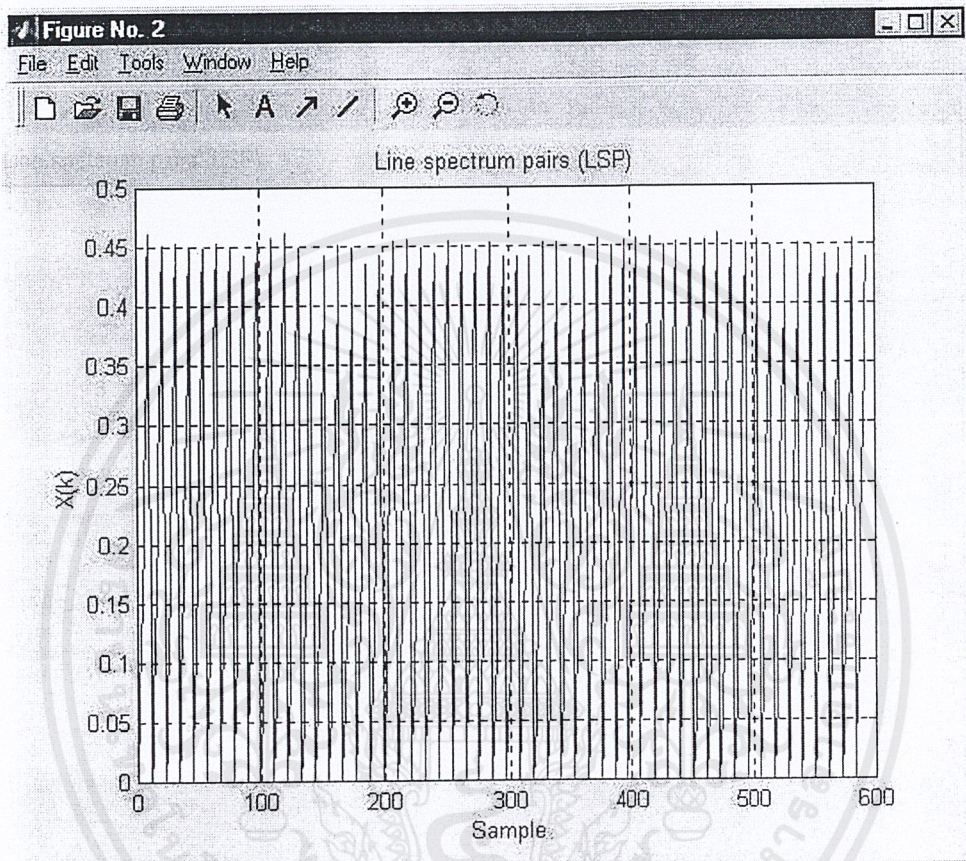
1. ทำการรับข้อมูลเสียงทดสอบเข้ามา และใช้โปรแกรมทำการแปลงให้เป็นสัญญาณการประมาณเชิงเส้น โดยทำที่ละเฟรมและเก็บค่าผลรวม
2. ทำการพล็อตดูรูปของสัญญาณจากหน้าจอ โดยคลิกไปที่คำสั่งให้ทำการพล็อตค่าสัมประสิทธิ์แอลเอสพี (Linear Spectrum Pair)



รูปที่ 4.22 แสดงค่าสัมประสิทธิ์แอลเอสพีของเสียงทดสอบ (Five.wav)

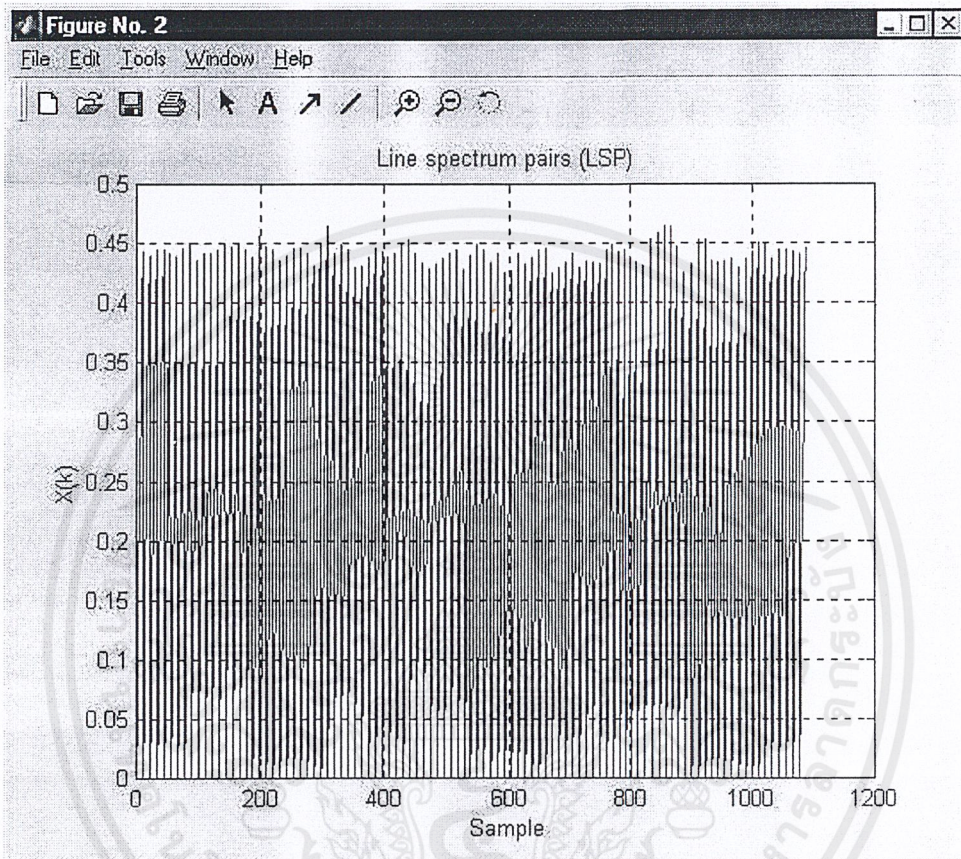
ขนาด 9 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



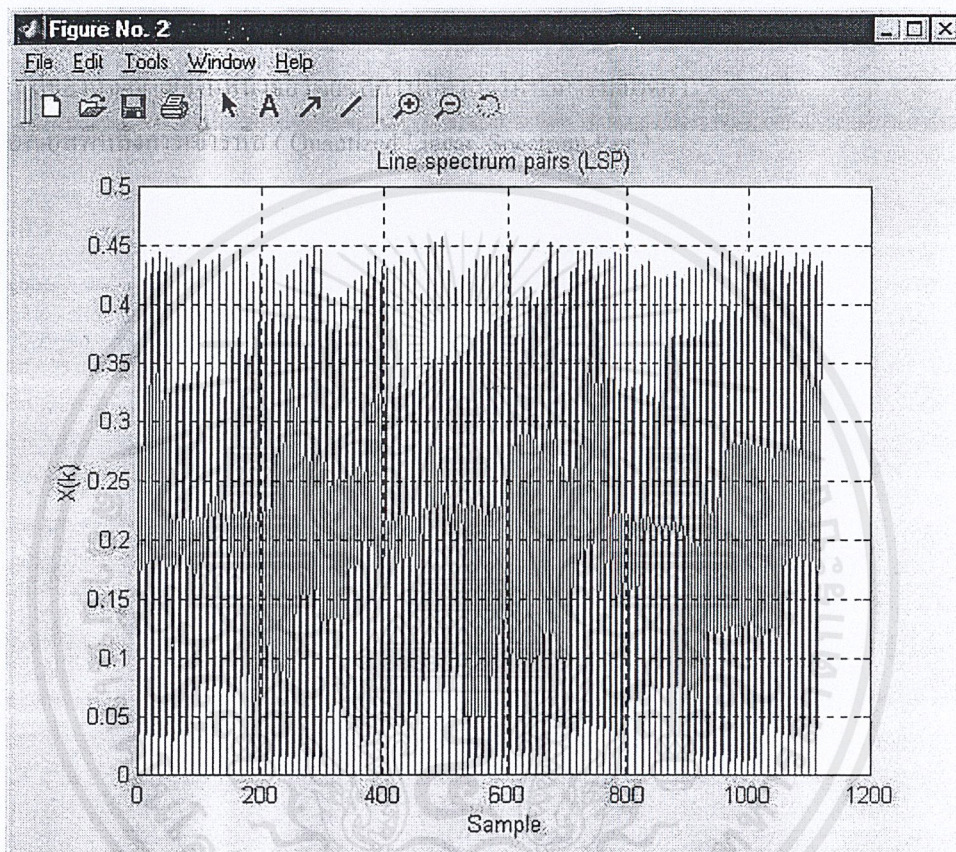
รูปที่ 4.23 แสดงค่าสัมประสิทธิ์เอลเอสพีของเสียงทดสอบ (สวัสดี_อิน.wav)
ขนาด 26 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.24 แสดงค่าสัมประสิทธิ์แอลเอสพีของเสียงทดสอบ (โน้ต.wav)
ขนาด 47 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



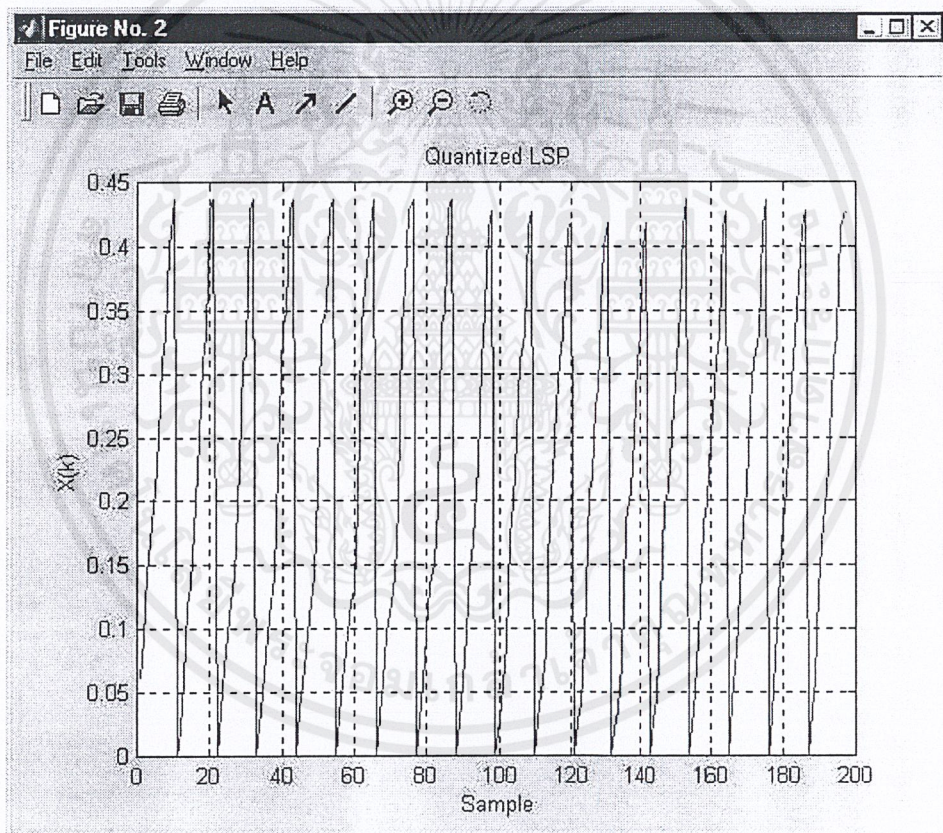
รูปที่ 4.25 แสดงค่าสัมประสิทธิ์แอลเอสพีของเสียงทดสอบ (เจ.wav)

ขนาด 49 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

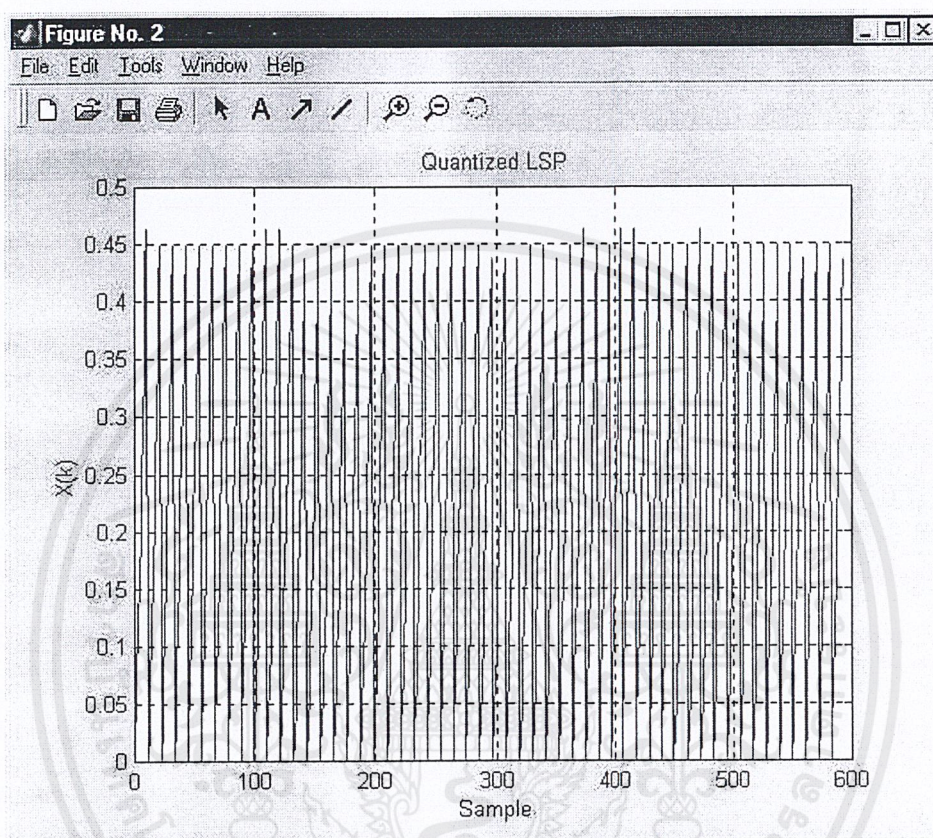
การทดลองที่ 4.5 ตรวจสอบค่าสัมประสิทธิ์โหนดสเปกตรัมแพร์ (แอลเอสพี) หลังผ่านการเข้ารหัสแล้ว จุดประสงค์การทดลอง เพื่อตรวจสอบค่าสัมประสิทธิ์โหนดสเปกตรัมแพร์ (แอลเอสพี) หลังจากการเข้ารหัสกับ ตาราง ของสัญญาณเสียงไฟล์ต่างๆที่รับเข้ามา เพื่อแสดงผลการดำเนินงาน และ ตรวจสอบความคลาดเคลื่อน โดยการดูรูปของสัญญาณ

- ขั้นตอนการทดลอง
1. ทำการรับข้อมูลเสียงทดสอบเข้ามา และใช้โปรแกรมทำการแปลงให้เป็นสัญญาณการประมาณเชิงเส้น และทำการเข้ารหัสโดยทำทีละเฟรมและเก็บค่าผลรวม
 2. ทำการพล็อตดูรูปของสัญญาณจากหน้าจอ โดยคลิกไปที่คำสั่งให้ทำการพล็อตค่าสัมประสิทธิ์แอลเอสพีหลังการเข้ารหัส (Quantized Linear Spectrum Pair)



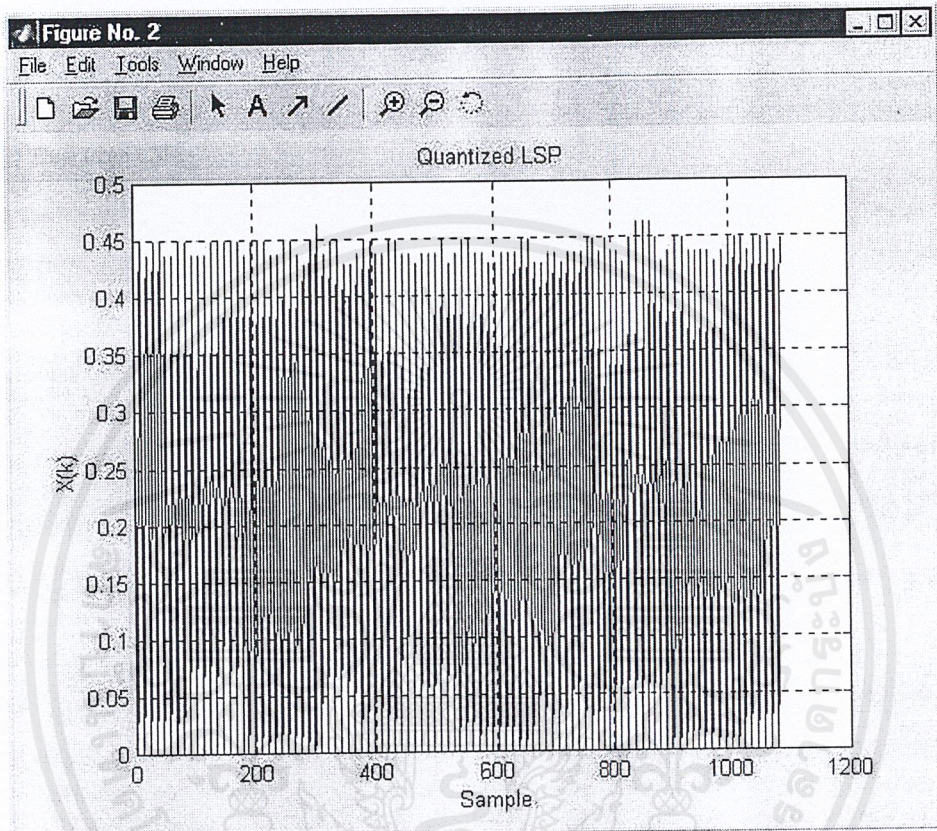
รูปที่ 4.26 แสดงค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (Five.wav) ขนาด 9 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



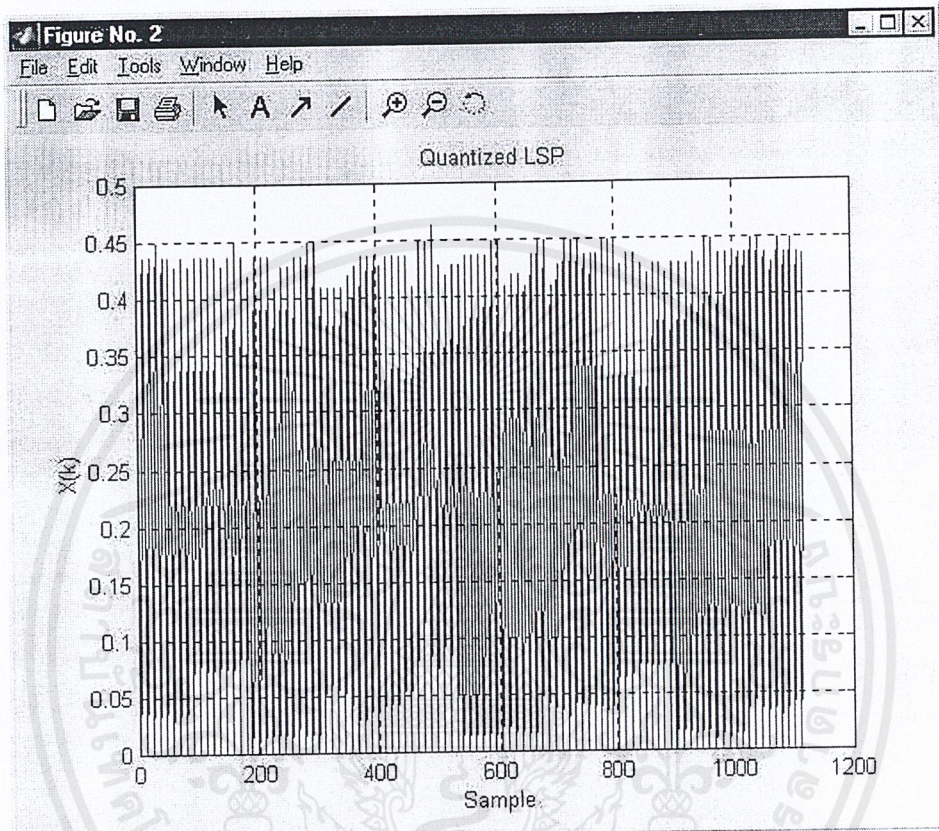
รูปที่ 4.27 แสดงค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (สวัสดี_อัน.wav) ขนาด 26 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



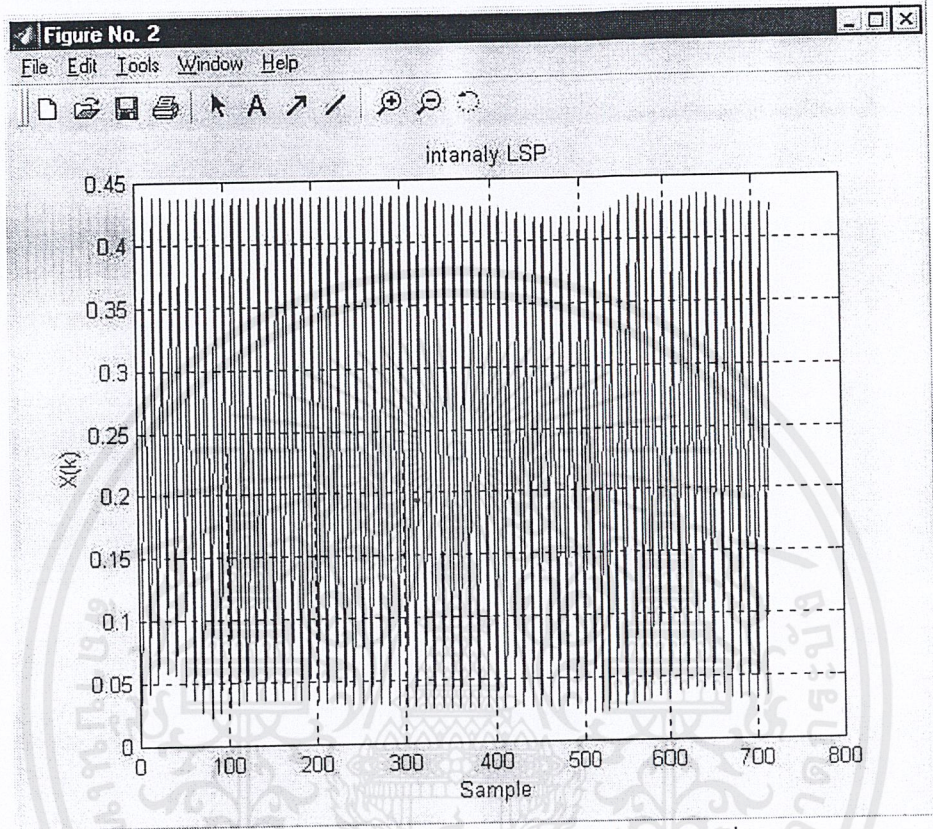
รูปที่ 4.28 แสดงค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (โน้ต.wav)
ขนาด 47 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



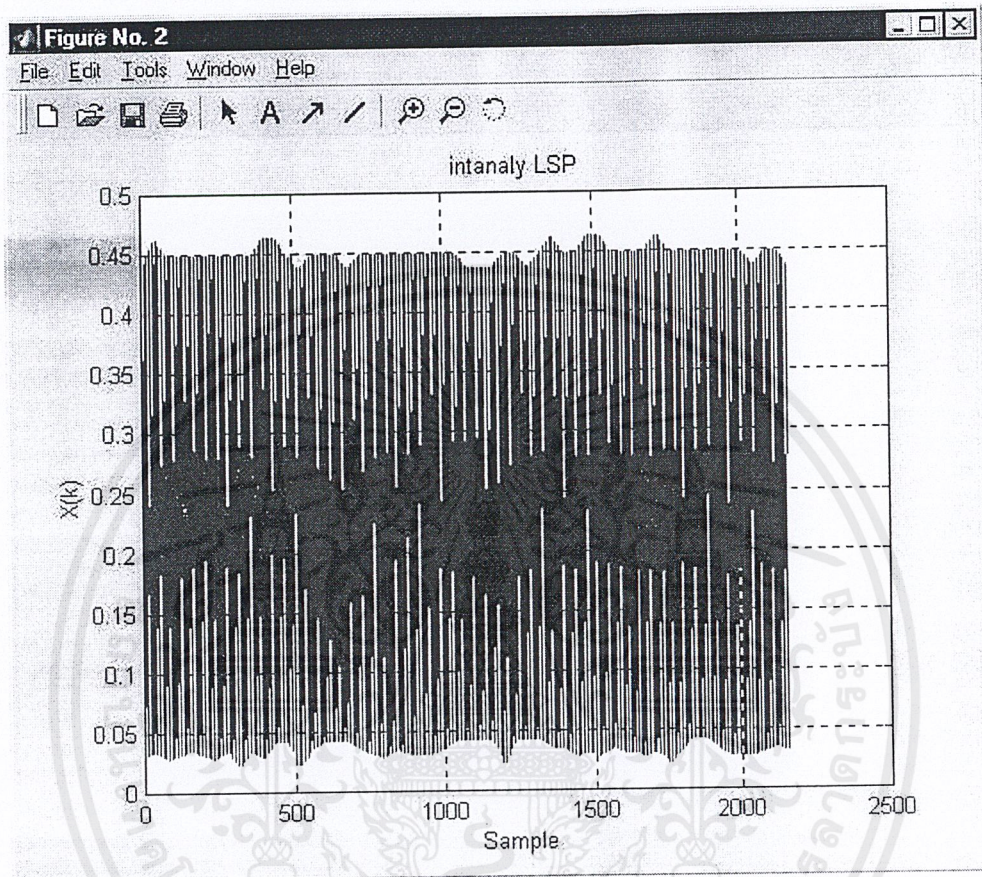
รูปที่ 4.29 แสดงค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (เจ.wav)
ขนาด 49 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



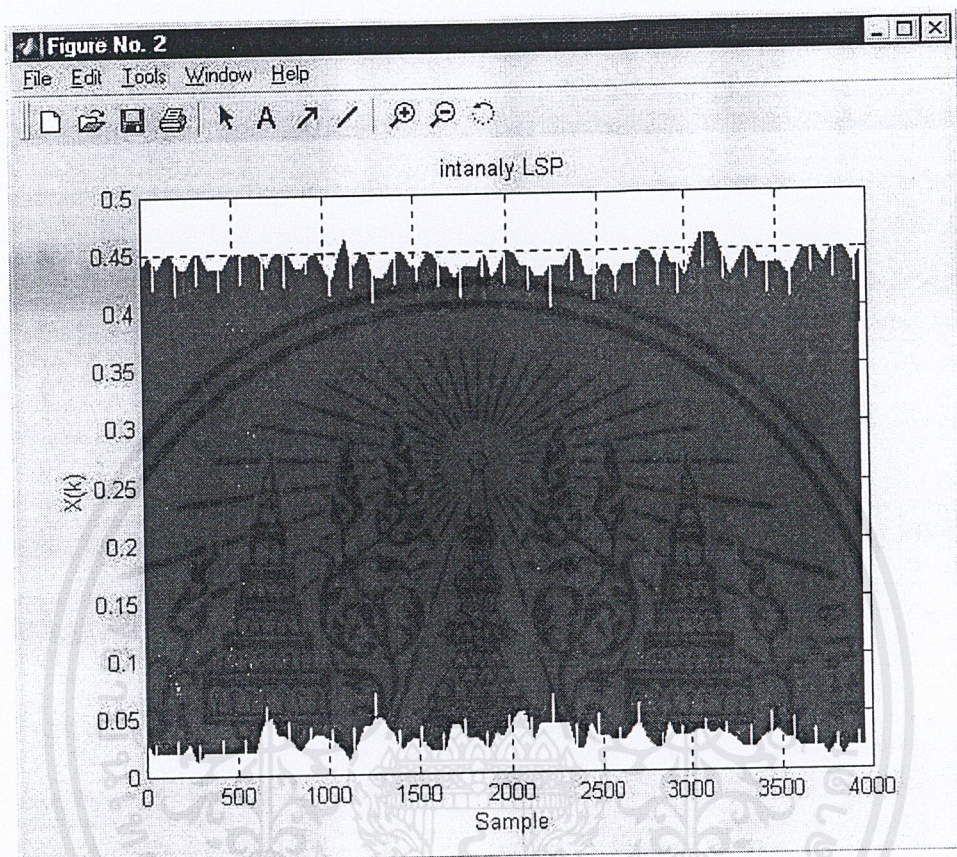
รูปที่ 4.30 แสดงการประมาณค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (Five.wav) ขนาด 9 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



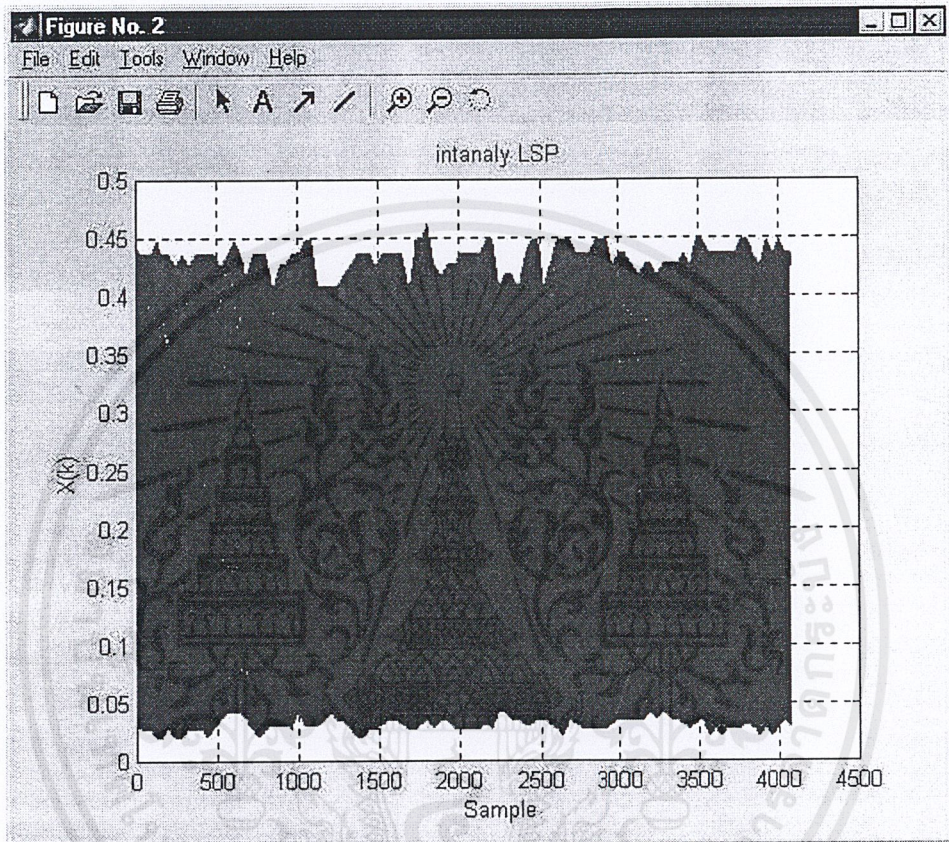
รูปที่ 4.31 แสดงการประมาณค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (สวัสดี_อิน.wav) ขนาด 26 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.32 แสดงการประมาณค่าสัมประสิทธิ์แอลเอสพีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (ไฟล์.wav) ขนาด 47 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



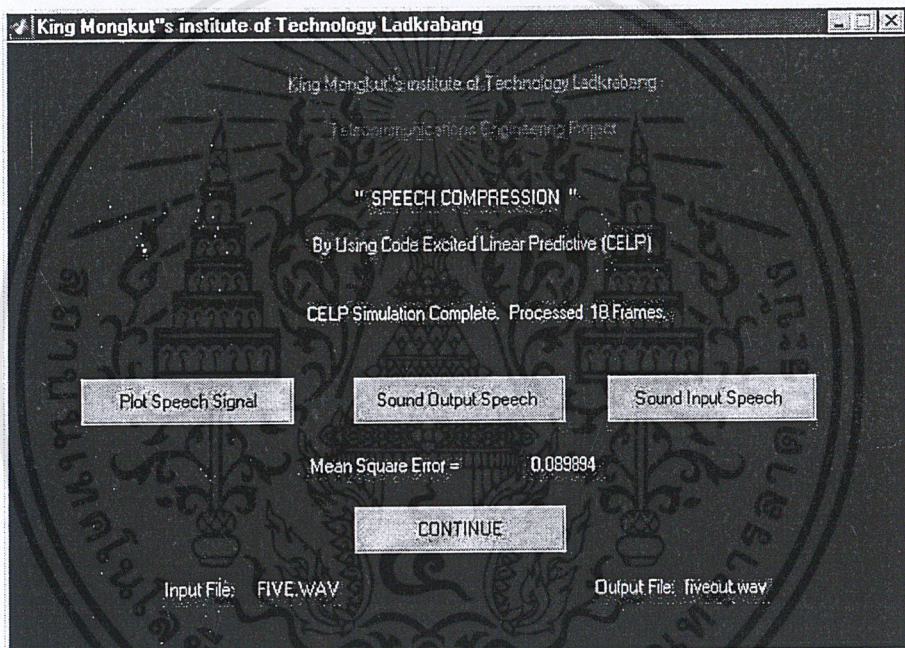
รูปที่ 4.33 แสดงการประมาณค่าสัมประสิทธิ์แอลเอสทีหลังผ่านการควอนไทเซชันของเสียงทดสอบ (เจ.wav)
ขนาด 49 กิโลไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทดลองที่ 4.6 ตรวจสอบค่าความผิดพลาด (Mean Square Error)

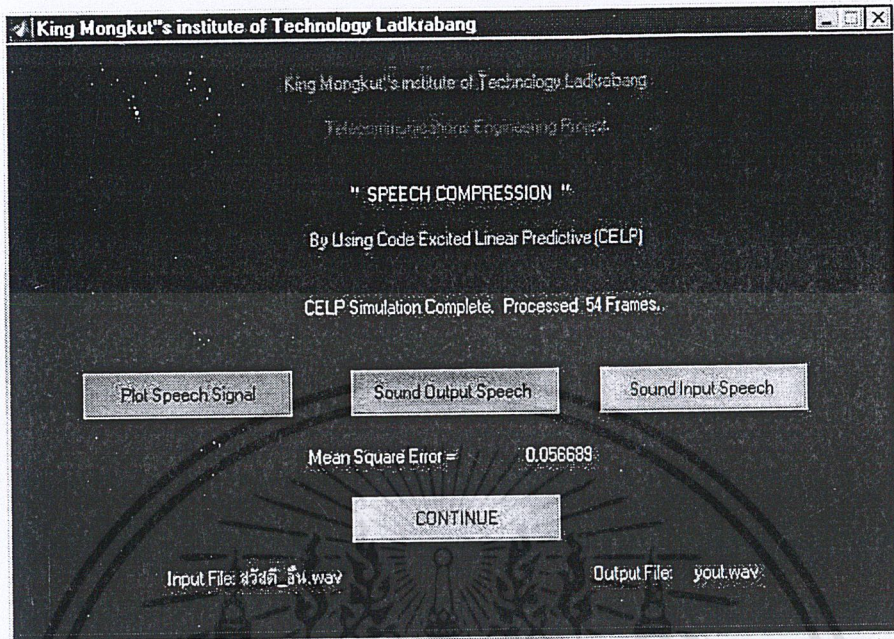
จุดประสงค์การทดลอง เพื่อตรวจสอบค่าความผิดพลาด ของสัญญาณเสียงไฟล์ต่างๆหลังจากการทำการแปลงสัญญาณแบบซีอีแอลพี และทำการแปลงกลับเป็นสัญญาณเดิมเรียบร้อยแล้ว

- ขั้นตอนการทดลอง
1. ทำการรับข้อมูลเสียงทดสอบเข้ามา และใช้โปรแกรมทำการเข้ารหัสและทำการถอดรหัส
 2. ตรวจสอบหาค่าความผิดพลาดของสัญญาณที่ออกมาเทียบกับสัญญาณอินพุต

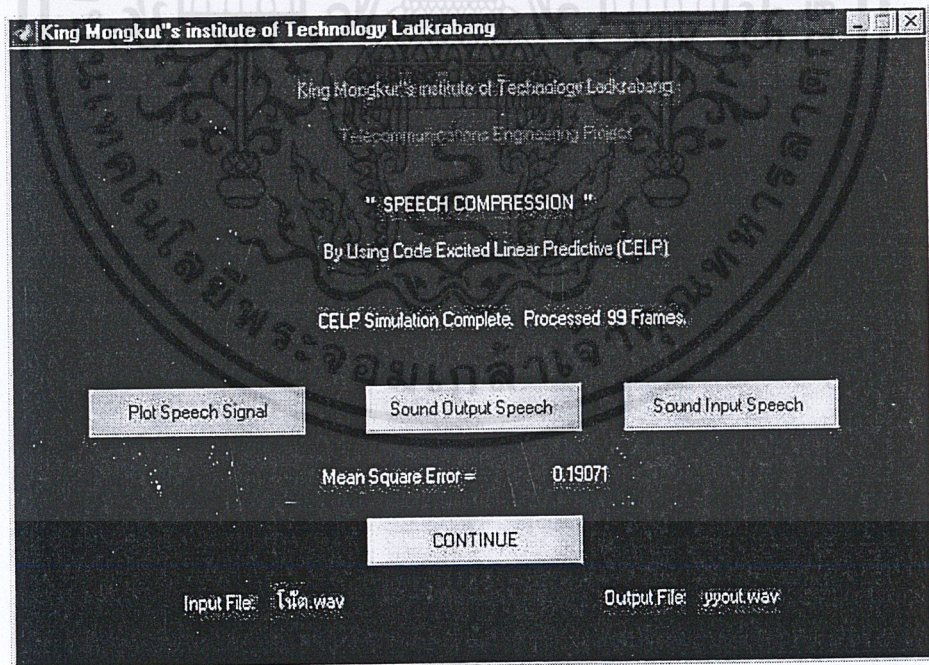


รูปที่ 4.34 แสดงค่ามินสแควร์เออเรอร์ (Mean-square error) ของสัญญาณเสียงทดสอบ (Five.wav)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

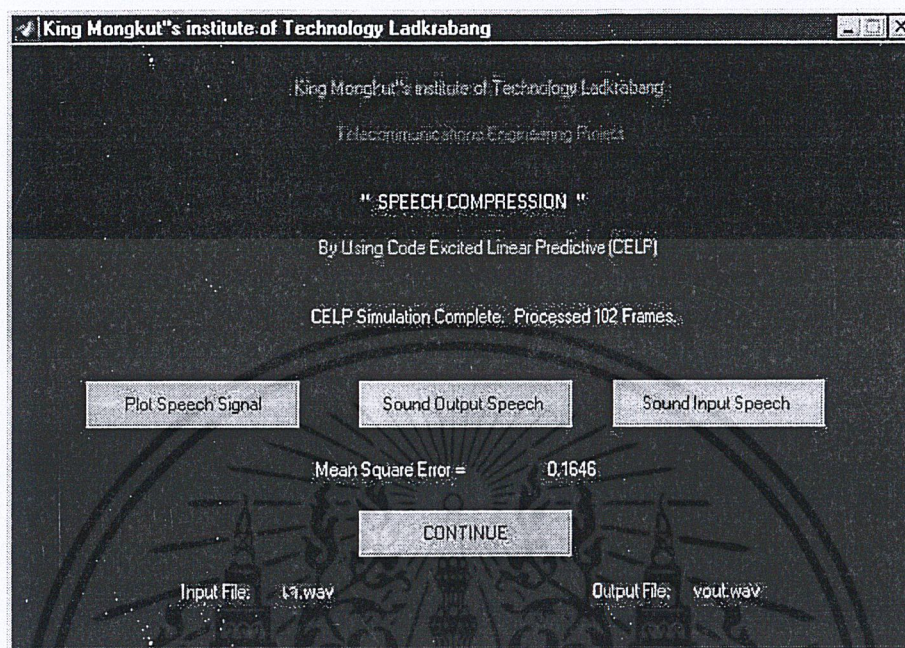


รูปที่ 4.35 แสดงค่ามีนสแควร์เอเรอร์ (Mean-square error) ของสัญญาณเสียงทดสอบ (สวัสดิ์_อิน.wav)



รูปที่ 4.36 แสดงค่ามีนสแควร์เอเรอร์ (Mean-square error) ของสัญญาณเสียงทดสอบ (โน้ต.wav)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.37 แสดงค่ามีนสแควร์เอเรอร์ (Mean-square error) ของสัญญาณเสียงทดสอบ (เจ.wav)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทดลองที่ 4.7 ตารางแสดงค่าขนาดของข้อมูลเมื่อผ่านการเข้ารหัสแบบต่างๆ

จุดประสงค์การทดลอง เพื่อทำการเปรียบเทียบขนาดไฟล์ของข้อมูลอินพุท และข้อมูลที่ทำการบีบอัดแล้ว
หลายๆตัว เปรียบกันและเปรียบเทียบกับค่าการเข้ารหัสโดยวิธีการการเข้ารหัสแบบอื่นๆ

- ขั้นตอนการทดลอง
1. เก็บข้อมูลเสียงเป็นไฟล์ (.wav) โดยทำการใช้ไฟล์เสียงตัวอย่างในการทดสอบ
ไฟล์เสียงเป็นเสียงพูดขนาดต่างๆ
 2. ทำการเข้ารหัสและถอดรหัสและทำการแสดงผลในตาราง

ผลการทดลอง

| ชื่อไฟล์ | ขนาดเดิม | ขนาดหลังจากการบีบอัด | | |
|-------------|----------|----------------------|-----------|---------|
| | (KB) | แบบ CELP | แบบ ADPCM | แบบ MP3 |
| ห้า | 9 | 1 | 2.17 | 1.25 |
| Hello | 18 | 2 | 4.24 | 2.49 |
| สวัสดี_อัน3 | 23 | 2 | 5.75 | 3.03 |
| สวัสดี | 25 | 2 | 6.20 | 3.23 |
| สวัสดี2 | 25 | 2 | 6.20 | 3.23 |
| สวัสดี3 | 25 | 2 | 6.20 | 3.23 |
| สวัสดี_อัน | 26 | 3 | 7.18 | 3.75 |
| สวัสดี_อัน4 | 28 | 3 | 7.00 | 3.69 |
| แม็ก | 40 | 4 | 9.98 | 5.10 |
| สวัสดี_อัน2 | 46 | 4 | 11.51 | 6.06 |
| ไนต์ | 47 | 4 | 11.98 | 6.16 |
| เจ | 49 | 4 | 12.34 | 6.32 |
| อัน4 | 49 | 4 | 12.26 | 6.46 |
| อัน5 | 138 | 11 | 34.52 | 18.19 |
| อัน3 | 156 | 13 | 39.02 | 20.56 |
| อัน2 | 250 | 20 | 62.54 | 32.95 |
| อัน | 308 | 25 | 77.05 | 40.59 |
| เบิร์ด | 346 | 28 | 86.55 | 45.60 |

ตารางที่ 4.1 แสดงการเปรียบเทียบขนาดของข้อมูลเมื่อทำการเข้ารหัสวิธีต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

| ชื่อไฟล์ | อัตราการใช้บิต (N:1) | | |
|-------------|-------------------------|-------|------|
| | CELP | ADPCM | MP3 |
| ห้า | 9.00 | 4.15 | 7.20 |
| Hello | 9.00 | 4.25 | 7.23 |
| สวัสดี_อัน3 | 11.50 | 4.00 | 7.59 |
| สวัสดี | 12.50 | 4.03 | 7.74 |
| สวัสดี2 | 12.50 | 4.03 | 7.74 |
| สวัสดี3 | 12.50 | 4.03 | 7.74 |
| สวัสดี_อัน | 8.67 | 3.62 | 6.93 |
| สวัสดี_อัน4 | 9.33 | 4.00 | 7.59 |
| แม็ก | 10.00 | 4.01 | 7.84 |
| สวัสดี_อัน2 | 11.50 | 4.00 | 7.59 |
| โน้ต | 11.75 | 3.92 | 7.63 |
| เจ | 12.25 | 3.97 | 7.75 |
| อัน4 | 12.25 | 4.00 | 7.59 |
| อัน5 | 12.55 | 4.00 | 7.59 |
| อัน3 | 12.00 | 4.00 | 7.59 |
| อัน2 | 12.50 | 4.00 | 7.59 |
| อัน | 12.32 | 4.00 | 7.59 |
| เบิร์ด | 12.36 | 4.00 | 7.59 |

ตารางที่ 4.2 แสดงค่าอัตราการใช้บิตข้อมูลเมื่อทำการเข้ารหัสวิธีต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5 บทวิจารณ์และบทสรุป

5.1 สรุปผลการทดลอง

การทดลองนี้เป็นการศึกษาการบีบอัดข้อมูลเสียง เพื่อลดขนาดของข้อมูลเสียง เพื่อประโยชน์ในการเพิ่มความสามารถในการส่งสัญญาณเสียงออกไป หลังจากที่ได้นำเสียงมาทำการวิเคราะห์ โดยเริ่มจากการเก็บข้อมูลเสียง และทำการเพิ่มอัตราส่วนของเสียงต่อสัญญาณรบกวน โดยนำไปทำพีเอฟเอซิสหรือผ่านการเข้าตัวกรองแบบให้ความถี่สูงผ่าน จะพบว่า จะได้เสียงที่มีอัตราส่วนของสัญญาณเสียงต่อสัญญาณรบกวนที่ดีขึ้น โดยจะลดองค์ประกอบของเสียงที่มีความถี่ต่ำลงได้ และนำเสียงที่ได้มานี้ไปทำการหาค่า เฟรมบล็อคล็อก และคูณด้วยค่าวินโดว์ และนำสัญญาณที่ได้ไปหาค่าสัมประสิทธิ์แอลพีซี โดยกำหนดลำดับของซีอีแอลพี เป็นลำดับที่ 10 (order 10) เพื่อใช้ในการคำนวณหาค่านิพจน์ของโค้ตบุ๊ก และค่าอัตราการขยายเพื่อบันทึกไว้เป็นข้อมูลหรืออยู่ในรูปของไฟล์ที่เป็นไฟล์ที่ผ่านกระบวนการซีอีแอลพี ซึ่งเก็บในรูปแบบของไฟล์ตัวอักษร (text file) เมื่อต้องการแปลงกลับให้อยู่ในรูปไฟล์เสียงจะต้องทำการอ่านค่าในแต่ละบรรทัดของฟรมนั้นๆและนำมาทำการสร้างสัญญาณเสียงกลับมาซึ่งจะได้เสียงที่มีคุณภาพในระดับที่ยอมรับได้ในระดับหนึ่งสำหรับเสียงพูด

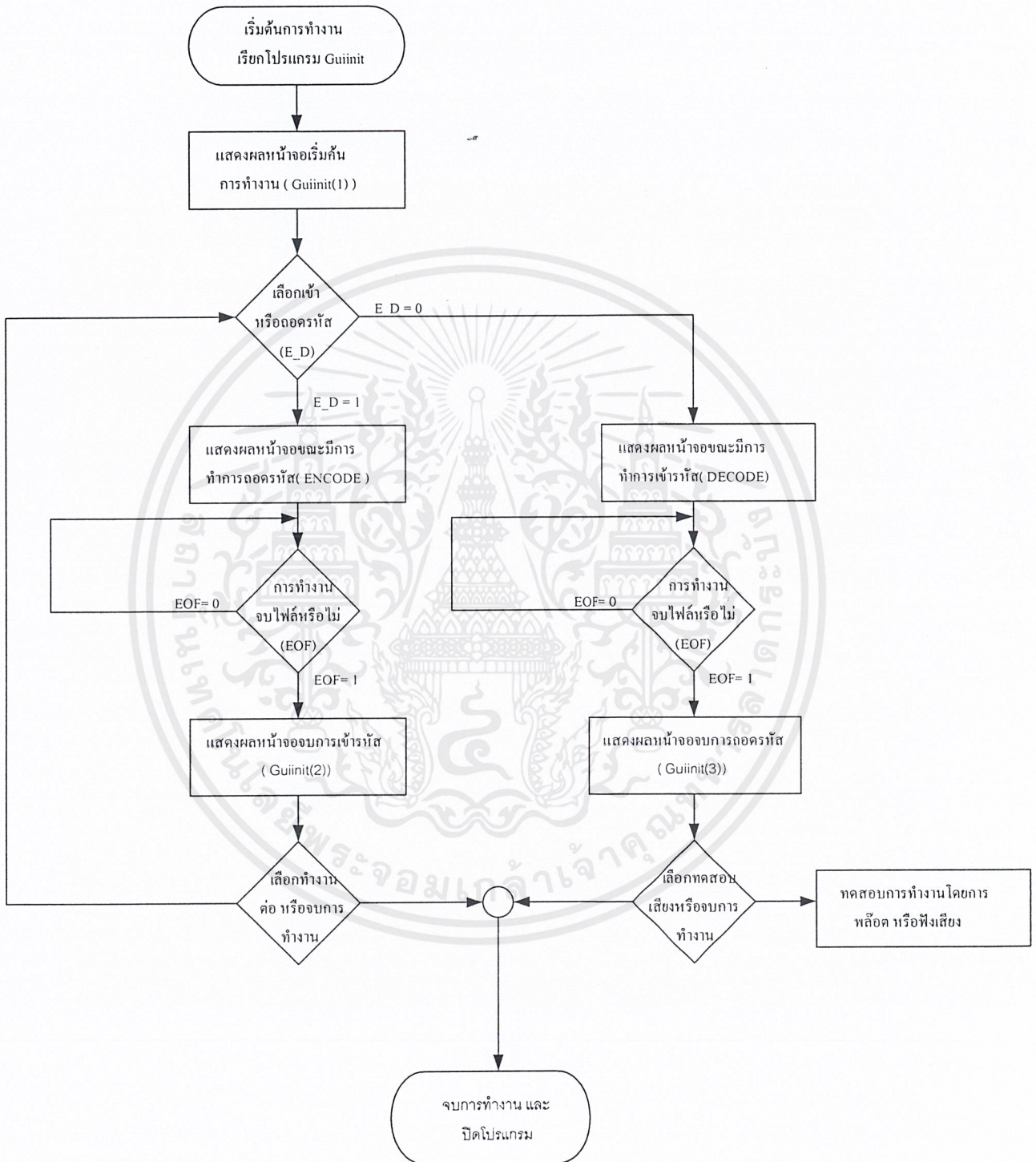
5.2 บทวิจารณ์และการพัฒนา

ปัญหาจากการจากการศึกษาปริญญาโทในครั้ง นี้ จะพบว่า มีปัญหาเนื่องจากการประมวลผลล่าช้าของโปรแกรม MATLAB ต้องใช้เวลานานในการทำการบีบอัด และในการคลายข้อมูลคั้งนั้นไม่สามารถใช้ในการส่งสัญญาณแบบตามเวลาจริงได้และในการค้นหาโค้ตบุ๊กจะมีการเกิดความผิดพลาดในการคำนวณและการเขียนโปรแกรมขึ้นทำให้ผลที่ได้ไม่ดีเท่าที่ควรจะต้องมีการพัฒนาการเขียนในส่วนของการค้นหาโค้ตบุ๊กให้ลดค่าความผิดพลาดลง หรืออาจใช้หลักในการหาโค้ตบุ๊กแบบอื่นๆที่ได้ถูกพัฒนาขึ้นมาในการแก้ไขปัญหาต่างๆ เช่น อาจใช้วิธีการหาโค้ตบุ๊กแบบลดความล่าช้า(Low-Delay CELP) หรือวิธีการผลรวมของเวกเตอร์ย่อย (Vector-Sum Excited Linear Predictive) และในการทดสอบยังมีปัญหาจากสภาพแวดล้อมขณะทำการบันทึกเสียง ซึ่งถ้าสภาพแวดล้อมไม่ดี ผลจากการบันทึกเสียงจะออกมาในลักษณะที่มีเสียงรบกวนมาก ทำให้เกิดเสียงที่มีคุณภาพต่ำด้วย



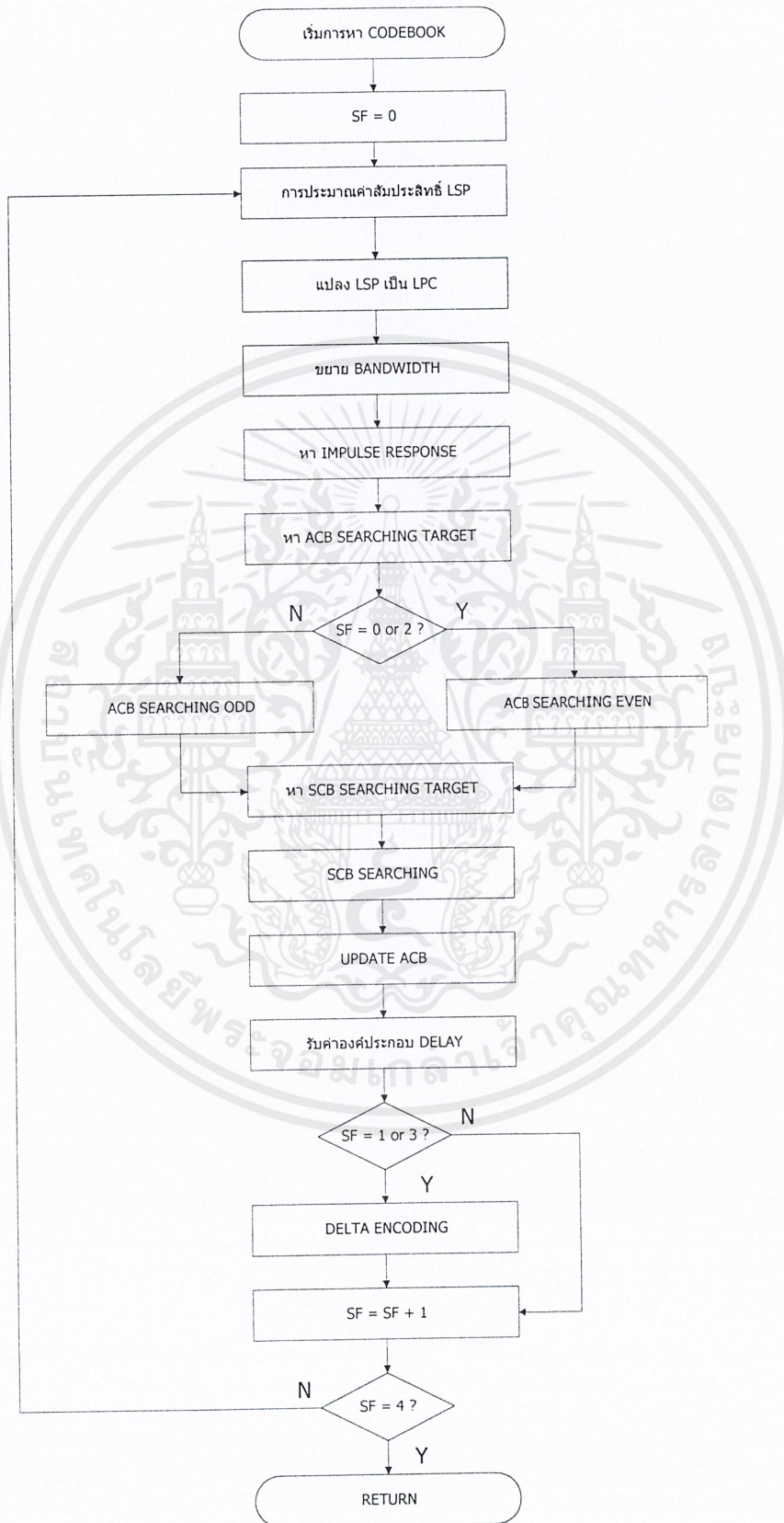
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โฟลว์ชาร์ต(Flow Chart) โปรแกรมส่วนแสดงผลการทำงาน (Guiinit)



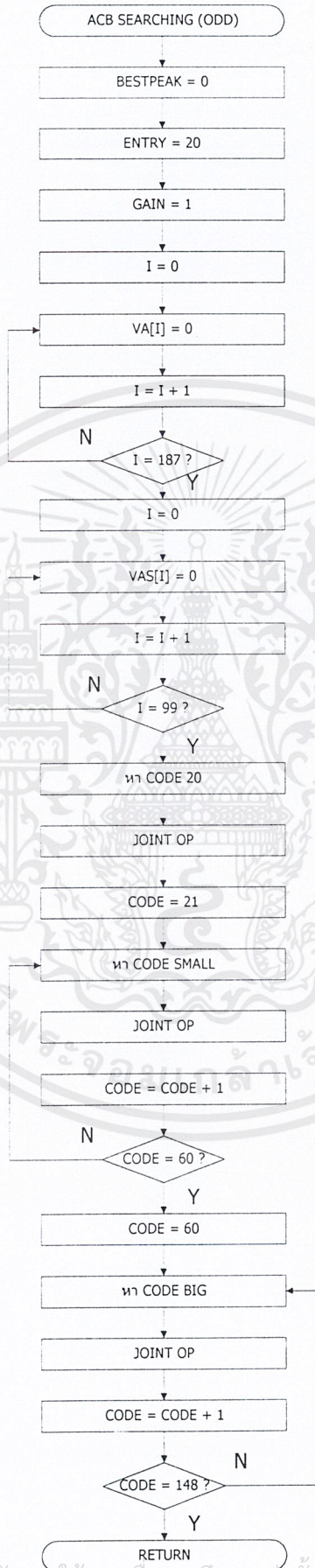
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โฟลว์ชาร์ท (Flowchart) แสดงขั้นตอนการหาโค้ดบุ้ค



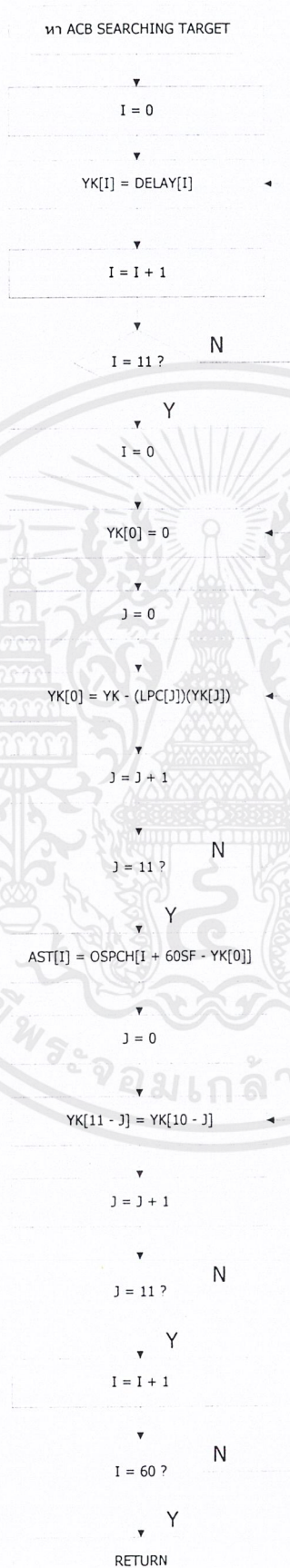
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โฟลว์ชาร์ท (Flowchart) แสดงขั้นตอนการหาโค้ดบุ๊กแบบอะแด็ปทีฟของเฟรมลี



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

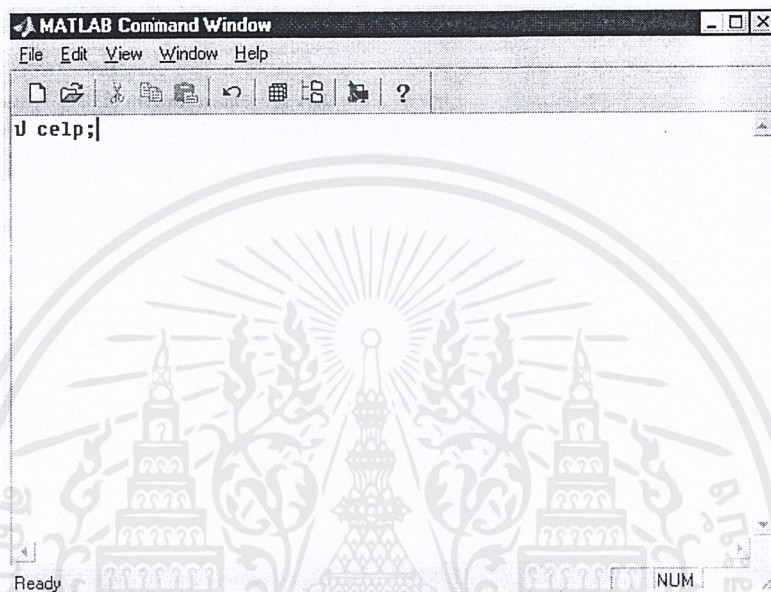
โฟลว์ชาร์ท (Flowchart) ใ้ค้ดบู้กแบบอะแต่ปทีฟ



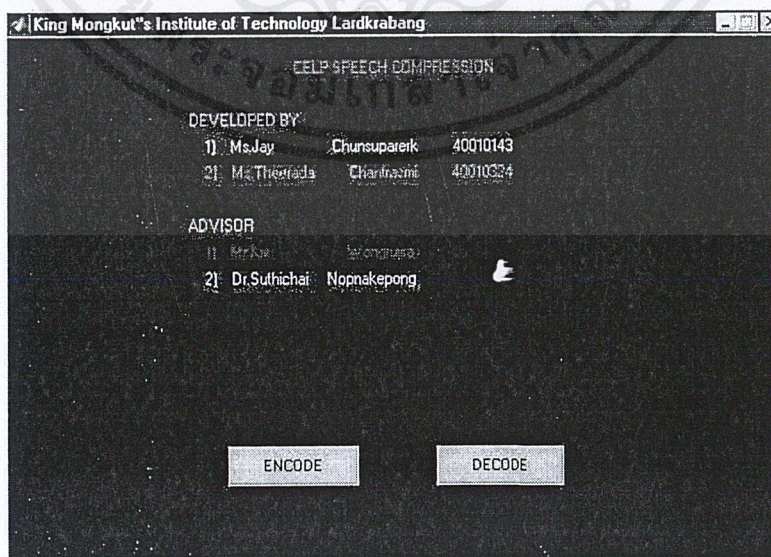
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีการใช้โปรแกรมการเข้ารหัสเสียง

การเริ่มต้นการเข้ารหัสเสียงซึ่งเขียนโปรแกรมไว้โดยใช้โปรแกรมเมทแล็บ (MATLAB) จะเริ่มต้นด้วยการเรียกโปรแกรมเมทแล็บ (MATLAB) ขึ้นมาและทำการเรียกไปยังไฟล์ชื่อ CELP โดยการพิมพ์คำว่า CELP ลงในหน้าจอการทำงานดังรูป

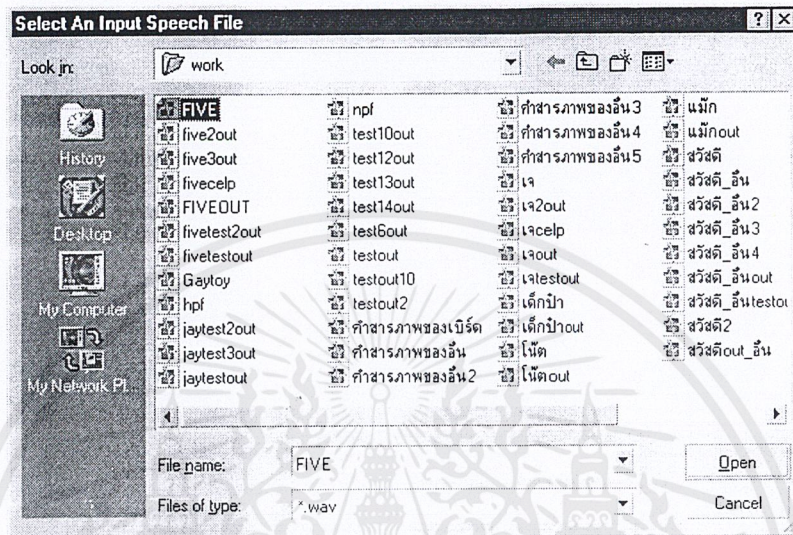


หลังจากนั้น โปรแกรมจะทำการเรียกไปยังโปรแกรมที่ทำการเขียนไว้และแสดงหน้าจอการทำงาน โดยที่การเริ่มต้นจะเริ่มต้นที่มีการให้ผู้ใช้เลือกว่าต้องการทำการเข้ารหัสหรือทำการถอดรหัสโดยหน้าจอการทำงานจะแสดงได้ดังรูป

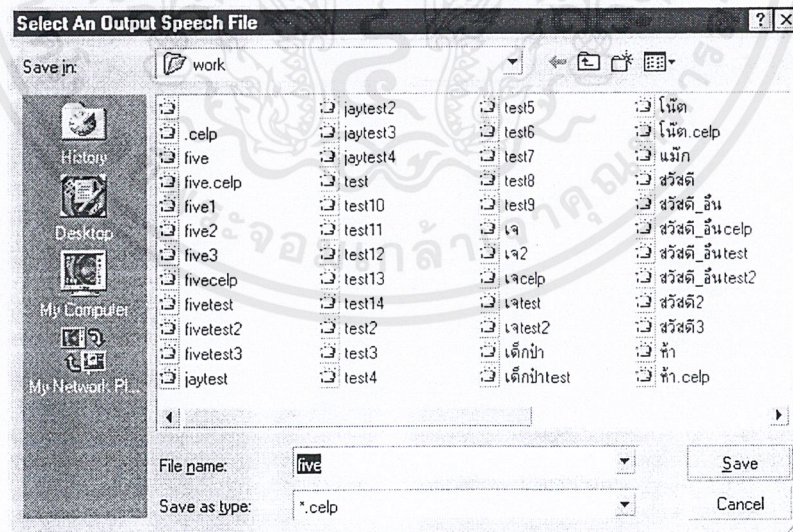


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยเมื่อผู้ใช้ต้องการเข้ารหัสสัญญาณเสียงจะต้องเลือกไปยังปุ่ม ENCODE โปรแกรมก็จะทำเข้ารหัส โดยเริ่มต้นแสดงหน้าจอการทำงานให้ผู้ใช้ทำการเลือกไฟล์เสียงที่ต้องการเข้ารหัสเสียง โดยกำหนดให้ผู้ใช้เลือกไฟล์ที่เป็นไฟล์เสียง (*.wav) เป็นอินพุตในการทำงาน และเลือกปุ่มเปิด (OPEN) ดังรูป



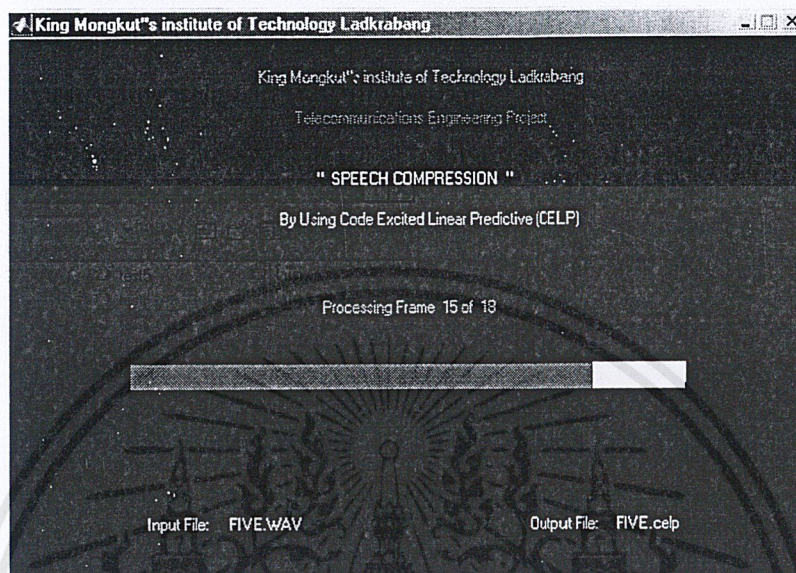
หลังจากนั้นจะให้ผู้ใช้ทำการกำหนดชื่อไฟล์ที่ต้องการจะให้ป็นเอาท์พุทของการทำงาน โดยให้เป็นไฟล์ตัวอักษรและกำหนดให้เป็นไฟล์ (*.celp) ดังรูป



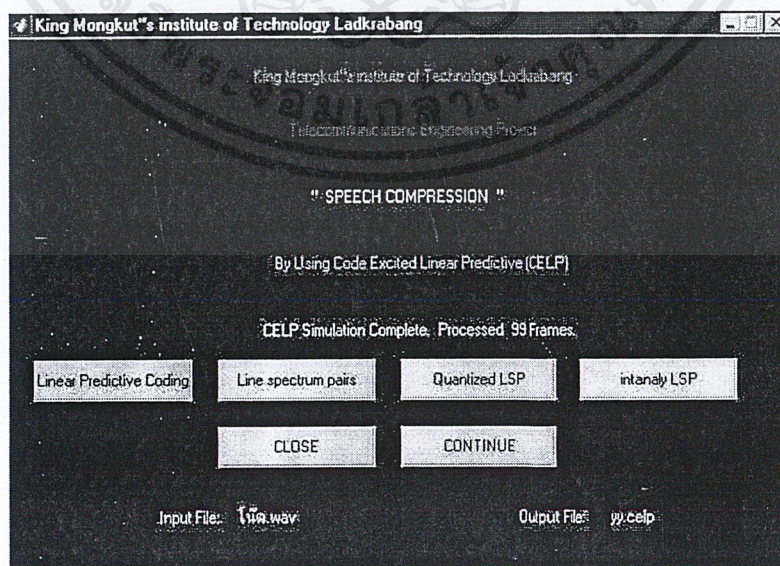
หลังจากทำการเลือกไฟล์อินพุทและเอาท์พุทเรียบร้อยแล้วโปรแกรมจะเริ่มการทำงาน โดยในการทำงานนั้นจะทำการทำงานเป็นแบบทำทีละเฟรม โดยเมื่อทำแต่ละเฟรมจะมีการแสดงผลการทำงานเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลักษณะของกราฟรูปแท่งแสดงสถานะการทำงานในแต่ละเฟรม และแสดงจำนวนเฟรมทั้งหมดที่จะต้องมีการทำงาน และชื่อไฟล์อินพุต และเอาต์พุตของการทำงานอยู่ในหน้าจอแสดงสถานะดังรูป

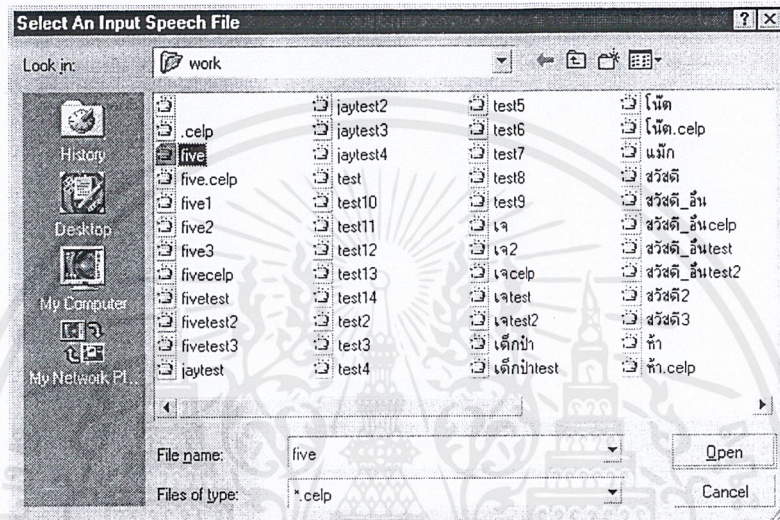


โดยในการทำงานแต่ละเสียงอินพุตจะมีเฟรมการทำงาน และระยะเวลาในการทำงานไม่เท่ากันขึ้นอยู่กับขนาดของเสียงที่รับเข้ามา เมื่อทำการเข้ารหัสเสร็จเรียบร้อยแล้วโปรแกรมจะมีการให้เลือกรการทำงานต่อไป โดยจะสามารถเลือกให้เริ่มการทำงานในขั้นตอนแรกใหม่โดยจะกลับไปหน้าจอการทำงานซึ่งเป็นการเลือกการเข้ารหัสหรือการถอดรหัสโดยเลือกที่ปุ่ม CONTINUE หรือสามารถเลือกให้หยุดการทำงานโดยการเลือกปุ่ม CLOSE และยังสามารถตรวจสอบสถานะการทำงานในขั้นตอนต่างๆ ได้ ดังรูป

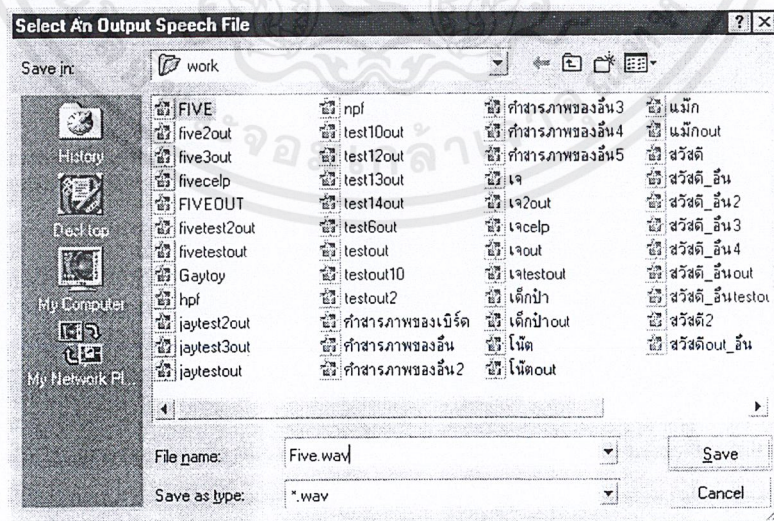


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยเมื่อทำการเลือกขั้นตอนที่ต้องการทดสอบโดยเลือกไปยังปุ่มที่ต้องการแล้วจะมีการแสดงผลออกมาเป็นการพล็อตกราฟแสดงผลขั้นตอนนี้ ๆ แต่ถ้าต้องการทำการถอดรหัสเมื่อทำการเลือกการถอดรหัสโดยเลือกที่ปุ่ม DECODE ในหน้าจอแสดงการทำงานแรกจะเริ่มทำการถอดรหัสข้อมูลเสียงโดยเริ่มต้นให้มีการให้เลือกสัญญาณที่ผ่านการบีบอัดมาแล้วในรูปของไฟล์(*.celp)มาเป็นอินพุตของการทำงานดังรูป

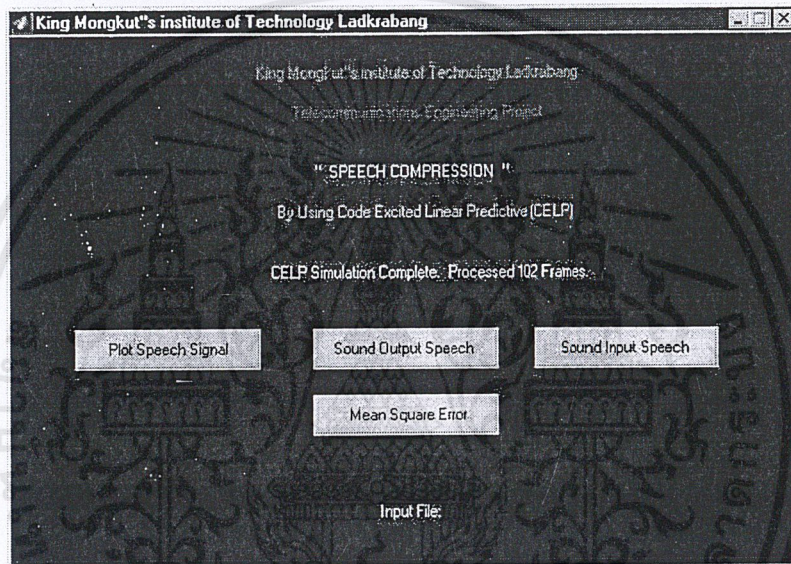


และหลังจากนั้นจะให้มีการเลือกกำหนดไฟล์เอาต์พุตที่ต้องการให้เป็นเอาต์พุตของการทำงาน โดยให้กำหนดเป็นไฟล์เสียง(*.wav) ดังรูป

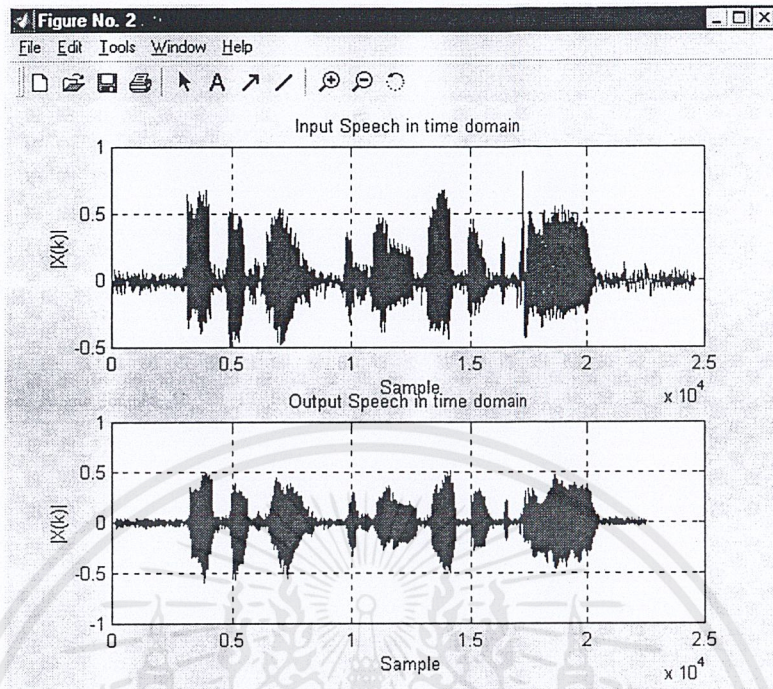


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

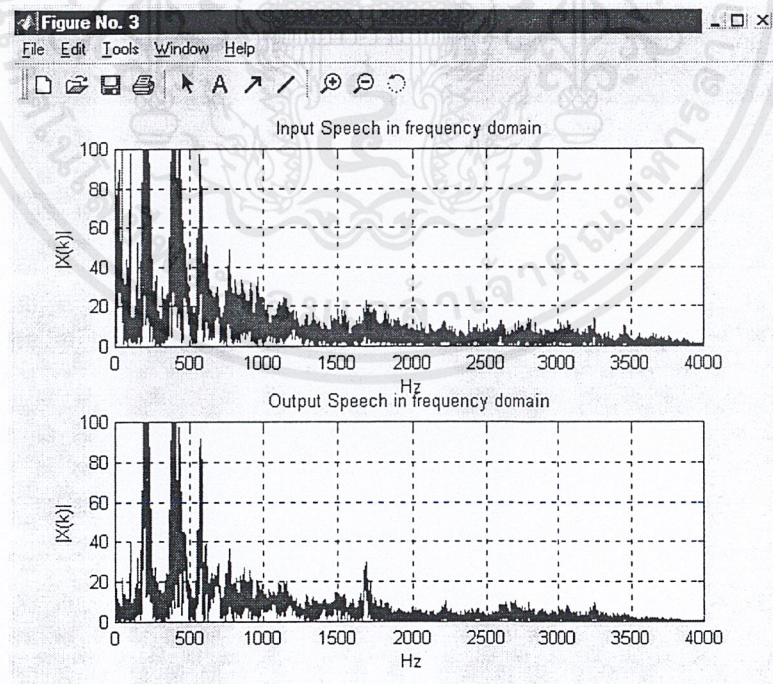
โดยหลังจากทำการถอดรหัสเรียบร้อยแล้ว โปรแกรมจะมีการให้เลือกรับเลือกการทำการตรวจสอบเสียงที่ทำการถอดรหัสออกมาได้โดยมีให้เลือกรับเลือกการตรวจสอบโดยการพล็อตดูรูปคลื่นของสัญญาณในแกนเวลาและแกนความถี่เปรียบเทียบกับกับสัญญาณเอาท์พุท ด้วยการเลือกปุ่ม Plot Speech Signal และมีการเลือกรับเลือกตรวจสอบโดยการฟังเสียงเอาท์พุทโดยการเลือกปุ่ม Sound Output Speech และสามารถทำการเปรียบเทียบฟังเสียงอินพุท โดยการเลือกที่ปุ่ม Sound Input Speech ซึ่งจะเป็นการทดสอบฟังเสียงอินพุทเดิมของสัญญาณเพื่อทำการเปรียบเทียบกับเอาท์พุทที่ได้มา หรือเลือกทำการออกจากการทำงานและทำการปิดโปรแกรม โดยการเลือกที่ปุ่ม CLOSE โดยมีหน้าจอแสดงการให้เลือกรับเลือกคำสั่งต่างๆ ดังรูป



จะสามารถเลือกรับเลือกการดูรูปคลื่นของสัญญาณเอาท์พุทและทำการดูค่าความผิดพลาดของสัญญาณได้โดยเมื่อเลือกปุ่ม Plot Speech Signal จะมีการทำการพล็อตกราฟเสียงที่เป็นเอาท์พุทหลังจากการทำการบีบอัดและทำการคลายข้อมูลออกมาเรียบร้อยแล้ว โดยโปรแกรมจะทำการพล็อตเปรียบเทียบกับสัญญาณอินพุทที่รับเข้ามาตอนเริ่มแรกในรูปแบบของแกนเวลาดังรูป

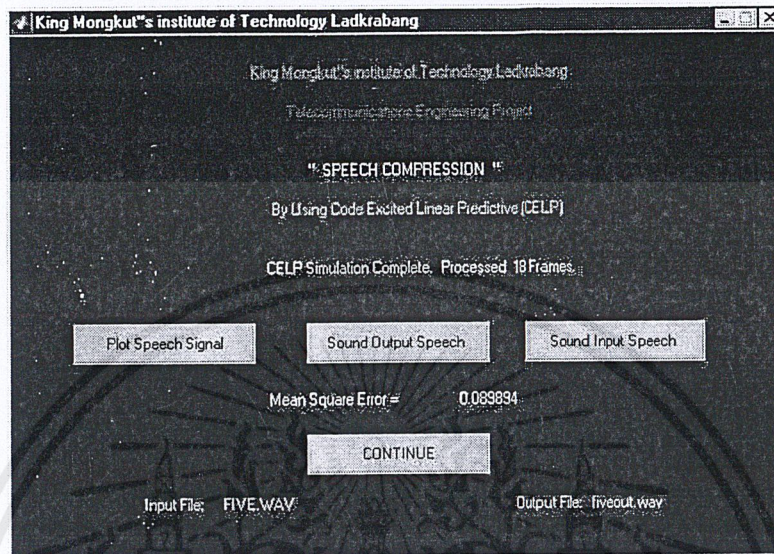


และจะทำการแปลงค่าสัญญาณให้อยู่ในรูปแกนความถี่ และแสดงเปรียบเทียบสัญญาณกับสัญญาณอินพุท ที่เข้ามาโดยแปลงให้อยู่ในแกนของความถี่เช่นเดียวกัน และจะแสดงผลเปรียบเทียบดังรูป



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และเมื่อทำการเลือกปุ่มแสดงค่าความผิดพลาด ก็จะมีการแสดงค่าความผิดพลาดในการทำงานออกมา
ทางหน้าจอแสดงผลดังรูป



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

%+++++
%      function guinit      +
%+++++
function guinit( screen )
global guiState guiF1 guiSline FrameCnt guiOpt guiColors guiIfile
guiOfile guiNPFfile guiHPFfile fnInput fnNPFfile fnHPFile fnPFile E_D
plotPara plotParal ANALYSIS
if strcmp( guiColors, 'white' ) == 1
    cFIG = 'w';
else
    cFIG = 'k';
end
PCOL = 'k';
ss = get( 0, 'ScreenSize' );
xorg = ( ss(3) - 600 ) / 2;
yorg = ( ss(4) - 400 ) / 2;
F1POS = [ xorg yorg 600 400 ];
guiDepth = get( 0, 'ScreenDepth' );
NameStr = 'King Mongkut"s institute of Technology Ladkrabang';
WSstr = ' ';
EndStr = sprintf( 'CELP Simulation Complete.  Processed %3d Frames.',
FrameCnt );
FileStr = sprintf( 'Input File: %12s%12sOutput File: %12s', guiIfile,
WSstr, guiOfile );
StatStr = sprintf( 'Processing Frame #%4d', FrameCnt );
L1STR = NameStr;
L2STR = 'Telecommunications Engineering Project';
X1STR = '" SPEECH COMPRESSION "' ;
X2STR = ' By Using Code Excited Linear Predictive (CELP)';
PSTR = 'Percent Complete (%)';

% 0 start screen
if screen == 0
    cr;

% 1 processing screen
elseif screen == 1

    guiF1 = figure( 'Position', F1POS, 'number', 'off', 'name', ...
                    NameStr, 'resize', 'off', 'Menubar', 'none',
'color', PCOL );
    clf;
    hh=uicontrol('style','text','string','King Mongkut"s institute of
Technology Ladkrabang',...
                'position', [110,330,400,50]);
    set(hh,'backg',get(gcf,'color'),'foreg','r');
    hh=uicontrol('style','text','string','Telecommunications
Engineering Project',...
                'position', [110,300,400,50]);
    set(hh,'backg',get(gcf,'color'),'foreg','b');
    hh=uicontrol('style','text','string,'" SPEECH COMPRESSION "',...
                'position', [80,285,450,20]);
    set(hh,'backg',get(gcf,'color'),'foreg','y');
    hh=uicontrol('style','text','string',' By Using Code Excited Linear
Predictive (CELP)',...
                'position', [90,255,450,20]);
    set(hh,'backg',get(gcf,'color'),'foreg','g');

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

    hh=icontrol('style','text','string',FileStr,'position',
[5,30,600,20]);
    set(hh,'backg',get(gcf,'color'),'foreg','c');
    guiSline=icontrol('style','text','string',StatStr,'position',
[100,190,400,20]);
    set(guiSline,'backg',get(gcf,'color'),'foreg','m');

    guiPax = axes('position',[0.15,0.35,0.7,0.05]);
    x=[ 0,0,0,0 ];
    y=[ 0,1,1,0];
    fill(x,y,'b');
    axis([0,100,0,1.0]);
    set(guiPax,'box','off');
    xlabel(PSTR);
    hold on
    drawnow;

% 2 compression output screen
elseif screen == 2
    close;
    guiF1 = figure( 'Position', FlPOS, 'number', 'off', 'name',
NameStr, ...
                    'resize', 'off', 'Menubar', 'none', 'color', PCOL
);
    clf;
    hh=icontrol('style','text','string','King Mongkut's institute of
Technology Ladkrabang',...
                'position', [110,330,400,50]);
    set(hh,'backg',get(gcf,'color'),'foreg','r');
    hh=icontrol('style','text','string','Telecommunications
Engineering Project',...
                'position', [110,290,400,50]);
    set(hh,'backg',get(gcf,'color'),'foreg','b');
    hh=icontrol('style','text','string','" SPEECH COMPRESSION "',...
                'position', [80,270,450,20]);
    set(hh,'backg',get(gcf,'color'),'foreg','y');
    hh=icontrol('style','text','string',' By Using Code Excited Linear
Predictive (CELP)',...
                'position', [90,220,450,20]);
    set(hh,'backg',get(gcf,'color'),'foreg','g');

    hh=icontrol('style','text','string',FileStr,'position',
[5,30,600,20]);
    set(hh,'backg',get(gcf,'color'),'foreg','c');
    hh=icontrol('style','text','string',EndStr,'position',
[115,170,400,20]);
    set(hh,'backg',get(gcf,'color'),'foreg',[1,1,1]-get(gcf,'color'));

    OKButton = uicontrol('Style','Pushbutton','Position',
[200,100,90,30],'Callback',...
                        'close;', 'String', 'CLOSE');
    CONButton = uicontrol('Style','Pushbutton','Position',
[310,100,90,30],'Callback',...
                      'close;init;', 'String', 'CONTINUE');

% 3 decompression output screen
elseif screen== 3

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

close;
guiFl = figure( 'Position', FlPOS, 'number', 'off', 'name',
NameStr, ...
                'resize', 'off', 'Menubar', 'none', 'color', PCOL
);
clf;
hh=uicontrol('style','text','string','King Mongkut"s institute of
Technology Ladkrabang',...
            'position', [110,330,400,50]);
set(hh,'backg',get(gcf,'color'),'foreg','r');
hh=uicontrol('style','text','string','Telecommunications
Engineering Project',...
            'position', [110,300,400,50]);
set(hh,'backg',get(gcf,'color'),'foreg','b');
hh=uicontrol('style','text','string','" SPEECH COMPRESSION "',...
            'position', [80,285,450,20]);
set(hh,'backg',get(gcf,'color'),'foreg','y');
hh=uicontrol('style','text','string',' By Using Code Excited Linear
Predictive (CELP)',...
            'position', [90,255,450,20]);
set(hh,'backg',get(gcf,'color'),'foreg','g');

hh=uicontrol('style','text','string',FileStr,'position',
[5,30,600,20]);
set(hh,'backg',get(gcf,'color'),'foreg','c');
hh=uicontrol('style','text','string',EndStr,'position',
[120,210,400,20]);
set(hh,'backg',get(gcf,'color'),'foreg',[1,1,1]-get(gcf,'color'));

PLOTButton = uicontrol('Style','Pushbutton','Position',
[50,150,140,30],'Callback',...
    'note(1);','String','Plot Speech Signal');
SOUNDButton = uicontrol('Style','Pushbutton','Position',
[230,150,140,30],'Callback',...
    'note(2);','String','Sound Output Speech ');
SOUNDButton = uicontrol('Style','Pushbutton','Position',
[400,150,140,30],'Callback',...
    'note(3);','String','Sound Input Speech ');
OKButton = uicontrol('Style','Pushbutton','Position',
[260,80,90,30],'Callback',...
    'close;','String','CLOSE');

end

%+++++++function cr+++++++
function cr
ss = get( 0, 'ScreenSize' );
xorg = ( ss(3) - 600 ) / 2;
yorg = ( ss(4) - 400 ) / 2;
FlPOS = [ xorg yorg 600 400 ];
TOP = 365; W = 600; H = 20; COL1 = 70; COL2 = 80; COL3 = 90; COL4 =
170;
C1P = [COL1+150 TOP W H];
C1aP= [COL1+70 TOP-2*H W H];
C2P = [COL1+70 TOP-3*H W H];
C3P = [COL1+70 TOP-4*H W H];
C4P = [COL1+70 TOP-6*H W H];

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

C5P = [COL1+70 TOP-7*H W H];
C6P = [COL1+70 TOP-8*H W H];
NOCPOS = [COL1+260,60,100,30];
OKCPOS = [COL1+100,60,100,30];

NameStr = 'King Mongkut"s Institute of Technology Lardkrabang ';
C1S = 'CELP SPEECH COMPRESSION';
ClaS= 'DEVELOPED BY';
C2S = '    1)    Ms.Jay                Chunsuparerk        40010143';
C3S = '    2)    Ms.Theerada            Chantrasmi        40010324';
C4S = 'ADVISOR ';
C5S = '    1)    Mr.Krit                    Wongrujira ';
C6S = '    2)    Dr.Suthichai            Nopnakepong';

guiF1 = figure( 'Position', FlPOS, 'number', 'off', 'name', NameStr,
...
                'resize', 'off', 'Menubar', 'none','color','k' );
clf;
guitext( C1P, C1S, 'k', 'r' );
guitext( ClaP, ClaS, 'k', 'g' );
guitext( C2P, C2S, 'k', 'w' );
guitext( C3P, C3S, 'k', 'm' );
guitext( C4P, C4S, 'k', 'c' );
guitext( C5P, C5S, 'k', 'b' );
guitext( C6P, C6S, 'k', 'y' );

OKButton =
uicontrol('Style','Pushbutton','Position',OKCPOS,'Callback',...
          'close; ENCD;', 'String','ENCODE');
NOButton =
uicontrol('Style','Pushbutton','Position',NOCPOS,'Callback',...
          'close; DEC;', 'String','DECODE');
%++++++function guitext++++++
function guitext( pos, str, bc, fc )
uicontrol('style','text','position', pos, ...
          'string', str, 'backg',bc, 'foreg',fc,'horiz','left' );

%++++++function guistat++++++
function guistat
global guiSline FrameCnt guiFrames;

% DISPLAY NEW STATUS
if FrameCnt<=guiFrames
StatStr = sprintf( 'Processing Frame %3d of %3d', FrameCnt, guiFrames
);
set( guiSline, 'string', StatStr );

% COMPUTE PERCENT COMPLETE AND UPDATE BAR GRAPH
Pcomp = (FrameCnt/guiFrames) * 100;
px = [0,0,Pcomp,Pcomp];
py = [0,1,1,0];
patch(px,py,'r');
end
if ( rem(FrameCnt, 3) == 0 ) | ( FrameCnt == guiFrames )
    drawnow;
end

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

%+++++++ function diskio ++++++++
function [ outbuf, status ] = diskio( mode, fileptr, input, length )
global EOF READ WRITE
if mode == READ
    [ outbuf, samples ] = fread( fileptr, length, 'short' );
    if samples ~= length
        status = EOF;
    else
        status = samples;
    end
else
    samples = fwrite( fileptr, input, 'short');
    if samples ~= length
        error( '**** disk write error ****\n' );
    end
    outbuf = input;
    [status q]=size(samples);
end

%+++++++ function celpexec ++++++++
%
%+++++++

global EOF LFRAME READ guiIfile guiOfile guiFrames CELPpath plotPara
plotParal LFRAME
jay=0;
WAVHDRSIZE = 44;
CELPhdr = '';
% OPEN INPUT AND OUTPUT FILES
if E_D ==0
    if E_D==0
        [ file, path ] = uigetfile( '*.wav', 'Select An Input Speech
        File' );
        plotPara=file;

% OPEN FILE AND COMPUTE INPUT SPEECH FRAMES (TRAP CANCEL EXCEPTION)
if file ~= 0

% UPDATE SEARCH PATH FOR INPUT FILES
[ fpi, emsg ] = fopen( [path,file], 'r' );
fHdrOffset = WAVHDRSIZE;
guiIfile = file;

% READ WAVE FILE HEADER
rID = fread( fpi, 4, 'char' );
rLen = fread( fpi, 1, 'long' );
wID = fread( fpi, 4, 'char' );
ChunkID = fread( fpi, 4, 'char' );
fLen = fread( fpi, 1, 'long' );
wFormatTag = fread( fpi, 1, 'short' );
nChannels = fread( fpi, 1, 'short' );
nSamplesPerSec = fread( fpi, 1, 'long' );
nAvgBytesPerSec = fread( fpi, 1, 'long' );
nBlockAlign = fread( fpi, 1, 'short' );
FormatSpecific = fread( fpi, fLen-14, 'char' );
dID = fread( fpi, 4, 'char' );

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

dLen = fread( fpi, 1, 'long' );

% VERIFY VALID .WAV FILE HAS BEEN SELECTED
WORDSIZE = nBlockAlign;
FSAMP = nSamplesPerSec;
c = zeros( 9, 1 );
c(1) = strcmp( setstr( rID ), 'RIFF' );
c(2) = strcmp( setstr( wId ), 'WAVE' );
c(3) = strcmp( setstr( ChunkID ), 'fmt ' );
c(4) = strcmp( setstr( dId ), 'data' );
c(5) = (wFormatTag == 1);
c(6) = (nChannels == 1);
c(7) = (nAvgBytesPerSec == (FSAMP * WORDSIZE));
c(8) = WORDSIZE == 2;
c(9) = FSAMP == 8000;
else
    error( 'CELP Simulation aborted!' );
end
if fpi == -1
    fclose( 'all' );
    error( emsg );
else
    % DETERMINE FRAME COUNT
    fseek( fpi, 0, 'eof' );
    guiFrames = fix( ( ftell(fpi) - fHdrOffset ) / (WORDSIZE*LFRAME)
);
    fseek( fpi, fHdrOffset, 'bof' );
end
fp=fpi;
npf=0;;
hpf=0;
end

% OPEN OUTPUT FILE
if E_D==0
    [ file, path ] = uinputfile( '*.celp', 'Select An Output Speech
File' );
    cd(CELPpath);
    guiOfile=strcat(file, '.celp');
else
    [ file, path ] = uinputfile( '*.wav', 'Select An Output Speech
File');
    plotParal=file;
    cd(CELPpath);
end
if E_D==1
% OPEN OUTPUT FILES AND GENERATE .WAV HEADERS
if file ~= 0
    % UPDATE SEARCH PATH FOR INPUT FILES
    [ fpo, emsg ] = fopen( [path,file], 'w' );
    [ fpnpf, emsg ] = fopen( [path,'npf.wav'], 'w' );
    [ fphpf, emsg ] = fopen( [path,'hpf.wav'], 'w' );
    fHdrOffset = WAVHDRSIZE;
    guiOfile = file;

    % CREATE .WAV HEADERS

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

if ~wavhdr( fpo, guiFrames, LFRAME, CELPhdr )
    error( 'Unable to create output .WAV file!' );
end
if ~wavhdr( fpnpf, guiFrames, LFRAME, CELPhdr )
    error( 'Unable to create non-postfiltered .WAV file!' );
end
if ~wavhdr( fphpf, guiFrames, LFRAME, CELPhdr )
    error( 'Unable to create high-pass filtered .WAV file!' );
end
else
    error( 'CELP Simulation aborted!' );
end

% TRAP FILE OPEN EXCEPTIONS
if fpo == -1
    fclose( 'all' );
    error( emsg );
end
fp=fpo;
hpf=fphpf;
npf=fpnpf;
end
fpi= fp;fpnpf=npf;fphpf=hpf;
else
    [ file, path ] = uigetfile( '*.celp', 'Select An Input Speech File'
);
    fid=fopen(file);
    c_line=0;
    while 1
        line=fgetl(fid);
        if ~isstr(line) break,end
        c_line=c_line+1;
    end
    fclose(fid);
    guiFrames=c_line;
    [fpo,fpnpf,fphpf] = setup(E_D);
end
% INITIALIZE STATUS OR TUTORIAL WINDOW, DEPENDING UPON GUI STATE
guiinit(1);
while 1 == 1
    % READ AN INPUT FRAME
    if E_D==0
        [ iarf, status ] = diskio( READ, fpi, [], LFRAME );
        if ( status ~= EOF )
            FrameCnt = FrameCnt + 1;
            pointer = 1;
            global guiSline
            if guiState == 0
                if FrameCnt<=guiFrames
                    StatStr = sprintf( 'Processing Frame %3d of %3d',
                    FrameCnt, guiFrames );
                    set( guiSline, 'string', StatStr );
                    Pcomp = (FrameCnt/guiFrames) * 100;
                    px = [0,0,Pcomp,Pcomp];
                    py = [0,1,1,0];
                    patch(px,py,'r');
                end
            end
        end
    end
end

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

        end
        if ( rem(FrameCnt, 3) == 0 ) | ( FrameCnt == guiFrames )
            drawnow;
        end
    end
    celpanal
else
    fclose( 'all' );
    guinit( 2 );
    break;
end
else
    fid=fopen(file);
    for a=1:c_line
        line=fgetl(fid);
        if a==c_line
            jay=1;
        end
        FrameCnt = FrameCnt + 1;
        pointer = 1;
        global guiSline
        if guiState == 0
            if FrameCnt<=guiFrames
                StatStr = sprintf( 'Processing Frame %3d of %3d',
                    FrameCnt, guiFrames );
                set( guiSline, 'string', StatStr );
                Pcomp = (FrameCnt/guiFrames) * 100;
                px = [0,0,Pcomp,Pcomp];
                py = [0,1,1,0];
                patch(px,py,'r');
            end
            if ( rem(FrameCnt, 3) == 0 ) | ( FrameCnt == guiFrames )
                drawnow;
            end
        end
    end
    celpsyn
    if jay==1
        fclose(fid);
        guinit(3);
        break;
    end
end
break;
end
end

```

```

%+++++
%          function celpanal          +
%+++++
global FrameCnt guiState SubFrameCnt guiFDM2 guisdM2 guidm2
snew = min( [ (iarf .* scale)'; maxv ] )';
snew = max( [ snew'; minv ] )';

% RUN HIGHPASS FILTER
[ snew, dhpfl ] = filter( bhpf, ahpf, snew, dhpfl );

ssub( 1:ll/2 ) = sold( (ll/2)+1:ll );

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

ssub( (11/2)+1:11 ) = snew( 1:11/2 );
sold = snew;

% DO LPC SPECTRAL ANALYSIS (OPEN LOOP)
si=snew;aay=11;omega=0.994127;
global MAXNO TRUE FALSE FrameCnt
unstable = FALSE;
atemp = zeros( MAXNO+1, 1);
atemp(1) = 1.0;

s = si .* hamw;
c = zeros( 10, 1 );
c0 = sum( s(1:aay) .* s(1:aay) );
for i = 1:10
    c(i) = sum( s(i+1:aay) .* s(1:aay-i) );
end

if c0 < 0.0
    unstable = TRUE;
end
global MAXNO
rc = zeros( MAXNO, 1 );
tmp = zeros( MAXNO, 1 );

if c0 <= 0.00
    rc( 1:10) = 0.00;
    return
end
alpha = c0;
atemp(2) = -c(1) / c0;
rc(1) = atemp(2);
beta = c(1);

% RECURSION
for i = 1:10-1
    alpha = alpha + ( beta * rc(i) );
    beta = c( i+1 );
    beta = beta + sum( c( 1:i ) .* atemp( i+1:-1:2 ) );
    rc( i+1 ) = -beta / alpha;
    tmp( 1:i ) = rc( i+1 ) .* atemp( i+1:-1:2 );
    atemp( 2:i+1 ) = atemp( 2:i+1 ) + tmp( 1:i );
    atemp( i+2 ) = rc( i+1 );
end

pc=atemp;
pcexp( 1:10+1, 1 ) = ( pc( 1:10+1 ) .* ( 0.994127 .^ (0:10)' ) );
a=pcexp;

% MATCH RCs to PCs
rc = pctorc( a,10 );
if any( abs(rc) > 1.0 )
    unstable = TRUE;
end
if unstable == TRUE
    printf( 'autohf: unstable lpc analysis at frame %d\n', FrameCnt );
    a( 2:10+1 ) = 0.00;
    rc( 1:10 ) = 0.00;

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

end
    fcn=a;rcn=rc;
% GENERATE LSP
[ newfreq, lspflag ] = pctolsp2( fcn, no );
if lspflag == TRUE
    fprintf( 'celp: Bad "new" lsp at frame: %d\n', FrameCnt );
end
unqfreq = newfreq;
global MAXNO lspQ
FSCALE = 8000.00;
findex = zeros( MAXNO, 1 );
newfreq = FSCALE * newfreq;
levels = ( 2 .^ sbits ) - 1;
for i = 1:no
    dist = abs( newfreq(i) - lspQ(i,1:levels(i)+1) );
    [ low, findex(i) ] = min( dist );

    if i > 1
        if lspQ( i,findex(i) ) <= lspQ( i-1,findex(i-1) )
            errorup = abs( newfreq(i) - lspQ(i,min(findex(i)+1,levels
                (i) )) ) + ...
                abs( newfreq(i-1) - lspQ(i-1, findex(i-1)) );
            errorndn = abs( newfreq(i) - lspQ(i,findex(i)) ) + ...
                abs( newfreq(i-1) - lspQ(i-1,max(findex(i-1)-
                1,0)) );
            if errorup < errorndn
                findex(i) = min( findex(i)+1, levels(i) );
                while ( lspQ(i,findex(i)) < lspQ(i-1,findex(i-1)) )
                    findex(i) = min( findex(i)+1, levels(i) );
                end
            elseif i == 1
                findex(i-1) = max( findex(i-1)-1, 0 );
            elseif lspQ( i-1, max(findex(i-1)-1, 0) ) > lspQ( i-2,
                findex(i-2) )
                findex(i-1) = max( findex(i-1)-1, 0 );
            else
                findex(i) = min( findex(i)+1, levels(i) );
                while lspQ( i, findex(i) ) < lspQ( i-1, findex(i-1) )
                    findex(i) = min( findex(i)+1, levels(i) );
                end
            end
        end
    end
end
end
end
findex = findex - 1;
Q = reshape( lspQ',160,1);
newfreq( 1:no ) = Q( ((0:no-1)*16)+findex(1:no)+1 ) / FSCALE;
% CONVERT LSPs TO AUTOCORRELATION
unqpc = lsptopc( unqfreq, no );
newpc = lsptopc( newfreq, no );
unqrc = pctorc( unqpc, no );
newrc = pctorc( newpc, no );
unqac = rctoac( unqrc, no );
newac = rctoac( newrc, no );

% FIND DISTANCES BETWEEN THE AUTOCORRELATION SEQUENCES

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

[ dm, sumdm, iframedm ] = dist( no, 4*no, unqac, newac, sumdm2,
framedm2 );
dm2=dm;sumdm2=sumdm;framedm2=iframedm;
guidm2 = dm2; guisd2 = sumdm2; guiFDM2 = framedm2;
for i = 1:length(sbits)
    last = pointer+sbits(i)-1;
    stream( last:-1:pointer ) = ( int2bin( findex(i), sbits(i) ) )';
    pointer = last + 1;
end
lsp = intanaly( newfreq, nn, no );

% SEARCH ADAPTIVE AND STOCHASTIC CODEBOOKS FOR EACH SUBFRAME
for subframe = 1:nn
    fci = lsptopc( lsp(subframe,:), no );
    fc = fci;
    nseg = nseg + 1;

    sfl = ((subframe-1)*1)+1; sfh = sfl+1-1;
    v(sfl:sfh) = csub( ssub(sfl:sfh), 1, 1p );
    if rem( subframe, 2 ) ~= 0
        [ stream, pointer ] = ...
        packtau( tauptr-minptr, pbits(1), pencode, stream, pointer );
    else
        [ stream, pointer ] = ...
        cpack( tauptr-minptr, pbits(2), pointer, stream );
    end

    % PITCH GAIN INDEX
    [ stream, pointer ] = cpack( pindex, pbits(3), pointer, stream );

    % STOCHASTIC CODEWORD INDEX
    [ stream, pointer ] = cpack( cbindex-1, cbbits, pointer, stream );

    % STOCHASTIC GAIN INDEX
    [ stream, pointer ] = cpack( gindex, cbgbits, pointer, stream );
end
if protect == TRUE
    codeword( 1:CODELENGTH2 ) = stream( bitprotect( 1:CODELENGTH2 ) );
    [ paritybit, codeword ] = encodham( CODELENGTH1, CODELENGTH2,
    hmatrix, codeword );
    % PACK FUTURE BIT
    [ stream, pointer ] = cpack( 0, 1, pointer, stream );
    % PACK PARITY BITS
    [ stream, pointer ] = cpack( codeword( CODELENGTH2+1:CODELENGTH1 ),
    ...[1,1,1,1], pointer, stream );
    % TOGGLE AND PACK THE SYNC BIT
    sync = ~sync;
    [ stream, pointer ] = cpack( sync, 1, pointer, stream );
end
savestream = stream;
pstream = stream( bitperm );
line = puthex( STREAMBITS, pstream );
fprintf(guiOfile , '%s\n', line );

%+++++++Pctolsp2 ++++++
function [ freq, lspflag ] = pctolsp2( a, m )
global TRUE FALSE MAXNO lastfreq

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

EPS = 1.00e-6;
N = 128;
NB = 15;
MAXORD = 24;
p = zeros( MAXORD, 1 );
q = zeros( MAXORD, 1 );
freq = zeros( MAXNO+1, 1 );
mp = m + 1;
mh = fix ( m/2 );

p( 1:mh ) = a( 2:mh+1 ) + a( m+1:-1:m-mh+2 );
q( 1:mh ) = a( 2:mh+1 ) - a( m+1:-1:m-mh+2 );

% COMPUTE P AT F=0
fl = 0.0;
pxl = 1.0 + sum( p(1:mh) );

% SEARCH FOR ZEROS OF P
nf = 1;
i = 1;
while i <= N
    mb = 0;
    fr = i * ( 0.5 / N );
    pxr = cos( mp * pi * fr );
    jc = mp - ( 2 * ( 1:mh )' );
    ang = pi * fr * jc;
    pxr = pxr + sum( cos(ang) .* p(1:mh) );
    tpxr = pxr;
    tfr = fr;
    if ( pxl * pxr ) <= 0.00
        mb = mb + 1;
        fm = fl + ( fr - fl ) / ( pxl - pxr ) * pxl;
        pxm = cos( mp * pi * fm );
        jc = mp - ( (1:mh)' * 2 );
        ang = pi * fm * jc;
        pxm = pxm + sum( cos(ang) .* p(1:mh) );
        if ( pxm * pxl ) > 0.00
            pxl = pxm;
            fl = fm;
        else
            pxr = pxm;
            fr = fm;
        end
    while (abs(pxm) > EPS) & ( mb < 4 )
        mb = mb + 1;
        fm = fl + ( fr - fl ) / ( pxl - pxr ) * pxl;
        pxm = cos( mp * pi * fm );
        jc = mp - ( (1:mh)' * 2 );
        ang = pi * fm * jc;
        pxm = pxm + sum( cos(ang) .* p(1:mh) );
        if ( pxm * pxl ) > 0.00
            pxl = pxm;
            fl = fm;
        else
            pxr = pxm;
            fr = fm;
        end
    end
    i = i + 1;
end

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

        end
    end
    if ( (pxl-pxr) * pxl ) == 0
        freq( 1:m ) = ( 1:m ) * 0.04545;
        fprintf( 'pctolsp2: default lsps used, avoiding /0\n' );
        return
    end
    freq(nf) = fl + (fr-fl) / (pxl-pxr) * pxl;
    nf = nf + 2;
    if nf > m-1
        break
    end
end
end
pxl = tpxr;
fl = tfr;
i = i + 1;
end

% SEARCH FOR ZEROS OF Q
freq(m+1) = 0.5;
fl = freq(1);
qxl = sin( pi * mp * fl );
jc = mp - ( (1:mh)' * 2 );
ang = pi * fl * jc;
qxl = qxl + sum( sin(ang) .* q(1:mh) );

i = 2;
while i < mp
    mb = 0;
    fr = freq(i+1);
    qxr = sin( mp * pi * fr );
    jc = mp - ( (1:mh)' * 2 );
    ang = pi * fr * jc;
    qxr = qxr + sum( sin(ang) .* q(1:mh) );
    tqxl = qxl;
    tfr = fr;
    tqxr = qxr;
    mb = mb + 1;
    fm = ( fl + fr ) * 0.5;
    qxm = sin( mp * pi * fm );
    jc = mp - ( (1:mh)' * 2 );
    ang = pi * fm * jc;
    qxm = qxm + sum( sin(ang) .* q(1:mh) );
    if ( qxm * qxl ) > 0.00
        qxl = qxm;
        fl = fm;
    else
        qxr = qxm;
        fr = fm;
    end
    while ( abs(qxm) > EPS*tqxl ) & ( mb < NB )
        mb = mb + 1;
        fm = ( fl + fr ) * 0.5;
        qxm = sin( mp * pi * fm );
        jc = mp - ( (1:mh)' * 2 );
        ang = pi * fm * jc;
        qxm = qxm + sum( sin(ang) .* q(1:mh) );

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

        if ( qxm * qxl ) > 0.00
            qxl = qxm;
            fl = fm;
        else
            qxr = qxm;
            fr = fm;
        end
    end
end
if ( ( qxl - qxr ) * qxl ) == 0
    freq(1:m) = lastfreq(1:m);
    fprintf( 'pctolsp2: last lsp used, avoiding /0\n' );
    return
end
freq(i) = fl + ( fr - fl ) / ( qxl - qxr ) * qxl;
qxl = tqxr;
fl = tfr;
i = i + 2;
end
freq = freq(1:m);
lspflag = FALSE;
if ( any(freq==0.00) | any(freq==0.5) )
    lspflag = TRUE;
end
for i = 2:m
    if freq(i) < freq(i-1)
        lspflag = TRUE;
        fprintf( 'pctolsp2: non-monotonic lsp\n' );
        tempfreq = freq(i);
        freq(i) = freq(i-1);
        freq(i-1) = tempfreq;
    end
end
for i = 2:m
    if freq(i) < freq(i-1)
        fprintf( 'pctolsp2: Reset to previous lsp values\n' );
        freq(1:m) = lastfreq;
        break;
    end
end
lastfreq = freq(1:m);

%+++++ Lsp34 +++++
function [ freq, findex ] = lsp34( freq, no, bits )
global MAXNO lspQ
FSCALE = 8000.00;
findex = zeros( MAXNO, 1 );
freq = FSCALE * freq;
levels = ( 2 .^ bits ) - 1;
for i = 1:no
    dist = abs( freq(i) - lspQ(i,1:levels(i)+1) );
    [ low, findex(i) ] = min( dist );
    if i > 1
        if lspQ( i,findex(i) ) <= lspQ( i-1,findex(i-1) )
            errorup = abs( freq(i) - lspQ(i,min(findex(i)+1,levels(i)) ) ) + ...
                abs( freq(i-1) - lspQ(i-1, findex(i-1)) );
            errordn = abs( freq(i) - lspQ(i,findex(i)) ) + ...

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้


```

%+++++++function cbsearch+++++++
function v = cbsearch( l, v )
global cbindex StochCB ncsize mxsw gindex cbgbits cbgtype TRUE FALSE
MAXNCSIZE
LEN = 30;
gain = zeros( MAXNCSIZE, 1 );
err = zeros( MAXNCSIZE, 1 );
codeword = ( 2 * MAXNCSIZE ) - 2;
cbindex = 1;
[ gain(1), err(1) ] = cgain( StochCB(codeword+1:codeword+1), 1, TRUE,
LEN );
emax = err(1);
codeword = codeword - 2;
for i = 1:ncsize-1
    [ gain(i+1), err(i+1) ] = cgain( StochCB(codeword+1:codeword+1), 1,
FALSE, LEN );
    codeword = codeword - 2;
    if err(i+1) >= emax
        emax = err(i+1);
        cbindex = i + 1;
    end
end
codeword = 2 * (MAXNCSIZE - cbindex);
[ gain(cbindex), err(cbindex) ] = cgain(
StochCB(codeword+1:codeword+1), 1, TRUE, LEN );
if mxsw == 1
    gain(cbindex) = mexcite3( gain(cbindex) );
end
if strcmp( cbgtype, 'none' ) == 0
    if cbgbits == 5
        [ gindex, gain(cbindex) ] = gaincode( gain(cbindex) );
    else
        fprintf( 'cbsearch: not quantizing cbgain\n' );
    end
end
v = ( gain(cbindex) * StochCB(codeword+1:codeword+1) );

%+++++++function csub+++++++
function v = csub( s, l, lp )
global MAXNO MAXPA e0 d1a d1b d2a d2b d3a d3b d4a d4b mxsw nseg;
e0 = zeros( l, 1 );
[ d1a, d2a, d3a, d4a ] = confg( s, l, d1a, d2a, d3a, d4a, 0, 1, 1, 1 );
d2a = d2b;
d3a = d3b;
d4a = d4b;
impulse(1);
if mxsw == 1
    mexcitel( 1 );
end
if lp == 1
    psearch( 1 );
elseif rem( nseg, 2 ) == 1
    psearch( lp );
end
e0 = zeros( l, 1 );
[ d1a, d2a, d3a, d4a ] = confg( s, l, d1a, d2a, d3a, d4a, 1, 1, 1, 1 );

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

if mxsw == 1
    mexcite2( 1 );
end
% STOCHASTIC CODEBOOK SEARCH
v = cbsearch( 1, 0 );
e0 = v;
[ d1b, d2b, d3b, d4b ] = config( s, 1, d1b, d2b, d3b, d4b, 1, 1, 1, 1 );
d1a = d1b;
d2a = d2b;
d3a = d3b;
d4a = d4b;

%+++++++ config ++++++++
function [ d1, d2, d3, d4 ] = ...
    config( s, 1, d1, d2, d3, d4, isw1, isw2, isw3, isw4 )

global MAXNO idb no bb e0 fc gamma2
fctemp = zeros( MAXNO+1, 1 );
% COMPUTE INITIAL STATE, 1/P(z)
if isw1 ~= 0
    [ e0, d1 ] = pitchvq( e0, 1, d1, idb, bb, 'long' );
end
% COMPUTE INITIAL STATE, 1/A(z)
if isw2 ~= 0
    [ d2, e0 ] = polefilt( fc, no, d2, e0, 1 );
end
e0 = s - e0;
% COMPUTE INITIAL STATE, A(z)
if isw3 ~= 0
    [ d3, e0 ] = zerofilt( fc, no, d3, e0, 1 );
end
% COMPUTE INITIAL STATE, 1/A(z/gamma)
if isw4 ~= 0
    fctemp = bwexp( gamma2, fc, no );
    [ d4, e0 ] = polefilt( fctemp, no, d4, e0, 1 );
end

%+++++++function polefilt+++++++
function [ z, xy ] = polefilt( a, n, z, xy, len )
if a(1) ~= 1.0
    fprintf( 'polefilt: bad coefficients\n' );
end
for t = 1:len
    z(1) = xy(t) - sum( z(2:n+1) .* a(2:n+1) );
    z(2:n+1) = z(1:n);
    xy(t) = z(1);
end

%+++++++function zerofilt ++++++++
function [ z, xy ] = zerofilt( b, n, z, xy, len )

for t = 1:len
    z(1) = xy(t);
    ar = sum( z(2:n+1) .* b(2:n+1) );
    z(2:n+1) = z(1:n);
    xy(t) = ar + ( z(1) * b(1) );
end

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

%+++++
%      function celpsyn      +
%+++++
global EOF LFRAME READ WRITE guiSin  guiSout  guiRCs  guiState  guidm
guisdm  guiFDM  avgSnrNPF  avgSnrPF  avgSnrHPF  guiExcit  SnrNPF  SnrPF
SnrHPF

STREAMBITS=144;
pointer=145;
pstream = gethex( 144, line );
stream( bitperm ) = pstream;
[ stream, errors, total ] = biterror( ber, mask, stream, STREAMBITS,
errors, total );
if protect == TRUE
    if SimType == SYNTHESIS
        pointer = 139;
    else
        pointer = pointer - PARITYLENGTH - 2;
    end
    [ codeword( CODELENGTH2+1:CODELENGTH1 ), pointer ] = ...
unpack( stream, [1,1,1,1], pointer );
codeword( 1:CODELENGTH2 ) = stream( bitprotect );
codeword( 11 ) = 0;
[ codeword, twoerror, syndrome ] = ...
decodham( CODELENGTH1, hmatrix, syndrometable, paritybit, codeword
);

twoerror = FALSE;
if syndrome ~= 0
    syndrome = 1;
end
syndavg = ( 1.0 - (1.0/SYNDRUN) ) * syndavg + ( 1.0 / SYNDRUN )
* syndrome;
stream( bitprotect ) = codeword( 1:CODELENGTH2 );
if twoerror == TRUE
    fprintf( 'celpexec: two errors have occured in frame %d\n',
FrameCnt );
end
end
pointer = 0;
[ findex, pointer ] = unpack( stream, sbits, pointer );
newfreq = lspdecod( findex, no );

lsp = intsynth( newfreq, nn, no, twoerror, syndavg );

bitpointer = pointer;
[ taus, bitpointer ] = dcoatau( pbits(1), pbits(2), bitsum1, bitsum2,
...
bitpointer, nn, stream, pdecode,
pdtabi );
[ pgs, bitpointer ] = dcodpg( pbits(3), bitsum1, bitsum2, bitpointer,
nn, stream );
...
[ cbi, bitpointer ] = dcodcbi( cbbits, bitsum1, bitsum2, bitpointer,
nn, stream );
...
[ cbg, bitpointer ] = dcodcbg( cbgbits, bitsum1, bitsum2, bitpointer,
nn, stream );
...

if SimType == ANALYSIS

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

nseg = nseg - nn;
end

for subframe = 1:nn

nseg = nseg + 1;
vl = ((subframe-1)*l)+1; vh = vl+l-1;

cbindex = cbi(subframe);
decodedgain = cbg(subframe);

if protect == TRUE
    decodedgain = smcbg( decodedgain, twoerror, syndavg, cbg,
        subframe );
end

vdecoded(vl:vh) = vdecode( decodedgain, l, cbindex );

if protect == TRUE
    taus(subframe) = smt( taus(subframe), twoerror, syndavg, taus
        (3), subframe );
end
bb(1) = taus(subframe);
bb(3) = pgs(subframe);

if protect == TRUE
    bb(3) = smpg( bb(3), twoerror, syndavg, pgs, subframe );
end

[ vdecoded(vl:vh), dps ] = pitchvq( vdecoded(vl:vh), l, dps, idb,
    bb, 'long' );
guiExcit(vl:vh) = vdecoded(vl:vh);

fci = lsptopc( lsp(subframe,:), no );
[ dss, vdecoded(vl:vh) ] = polefilt( fci, no, dss, vdecoded(vl:vh),
    l );

vdecoded(vl:vh) = vdecoded(vl:vh) * descale;
npf(vl:vh) = min( [ (vdecoded(vl:vh))'; maxv(vl:vh) ] )';
npf(vl:vh) = round( max( [ (npf(vl:vh))'; minv(vl:vh) ] )' );

[ junkbuf, status ] = diskio( WRITE, fpnpf, npf(vl:vh), l );
if SimType == ANALYSIS

    [ sumsnr, framesnr, snrflag, SnrNPF ] = segsnr( ssub(vl:vh),
npf(vl:vh), ...
        sumsnr, framesnr );

    if ( snrflag == TRUE )
        [ dm, sumdm, framedm ] = disto( ssub(vl:vh), npf(vl:vh),
hamws, ...
            l, no, sumdm, framedm );

        guidm = dm;
        guisd = sumdm;
        guiFDM = framedm;
        avgSnrNPF = 10*log10(sumsnr/framesnr);
    end
end

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

end

[ vdecoded(vl:vh), ip, op, dp1, dp2, dp3 ] = ...
postfilt( vdecoded(vl:vh), l, ALPHA, BETA, ip, op, dp1, dp2, dp3,
fci, no );

while clip( vdecoded( vl:vh ) ) == TRUE
    vdecoded( vl:vh ) = 0.05 * vdecoded( vl:vh );
    fprintf( 'celpsyn: Clipping detected @ frame %d\n', FrameCnt );
end

pf(vl:vh) = min( [ (vdecoded(vl:vh))'; maxv(vl:vh) ] )';
pf(vl:vh) = round( max( [ (pf(vl:vh))'; minv(vl:vh) ] )' );

[ junkbuf, status ] = diskio( WRITE, fpo, pf(vl:vh), l );

[ dhpfl0, vdecoded(vl:vh) ] = zerofilt( ahpfo, 2, dhpfl0, vdecoded
(vl:vh), l );
[ dhpfl2o, vdecoded(vl:vh) ] = polefilt( bhpfo, 2, dhpfl2o, vdecoded
(vl:vh), l );

hppf(vl:vh) = min( [ (vdecoded(vl:vh))'; maxv(vl:vh) ] )';
hppf(vl:vh) = round( max( [ (hppf(vl:vh))'; minv(vl:vh) ] )' );

[ junkbuf, status ] = diskio( WRITE, fphpf, hppf(vl:vh), l );

if SimType == ANALYSIS
    [ sumsnrp, framesnrp, snrflagp, dummy1 ] = segsnr( ssub(vl:vh),
sumsnrp, framesnrp );
    [ sumsnrh, framesnrh, snrflagh, dummy1 ] = segsnr( ssub(vl:vh),
sumsnrh, framesnrh );
    if snrflagp == TRUE
        avgSnrPF = 10*log10(sumsnrp/framesnrp);
    end
    if snrflagh == TRUE
        avgSnrHPF = 10*log10(sumsnrh/framesnrh);
    end
end
end

if SimType == ANALYSIS

    [ dummy1, dummy2, dummy3, SnrNPF ] = segsnr( ssub, npf, 0, 0 );
    [ dummy1, dummy2, dummy3, SnrPF ] = segsnr( ssub, pf, 0, 0 );
    [ dummy1, dummy2, dummy3, SnrHPF ] = segsnr( ssub, hppf, 0, 0 );
end

if guiState == 1
    guiRCs = fci;
    guiSin = ssub;
    guiSout = [ npf, pf, hppf ];
    gui;

    break;
end

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

%+++++++ vdecode ++++++++

function vdecoded = vdecode( decodedgain, l, cbindex )
global StochCB FrameCnt MAXNCSIZE
vdecoded = zeros( l, 1 );
codeword = ( 2 * ( MAXNCSIZE - cbindex ) ) + 1;
if codeword < 1
    fprintf( 'vdecode: cbindex > MAXNCSIZE @ frame %d\n', FrameCnt );
    codeword = 1;
end
vdecoded = StochCB( codeword:codeword+1-1 ) * decodedgain;

%+++++++pitchvq+++++++
function [ rar, buf ] = pitchvq( rar, idim, buf, idimb, b, type )
global MAXLP
k = idimb - idim;
start = k + 1;
m = fix( b(1) );
frac = b(1) - m;
buf2 = zeros( MAXLP, 1 );
buf( 1:k ) = buf( (1+idim):(k+idim) );
if abs(frac) < 1.0e-4
    if ( idimb-m ) >= k+1
        if m < 30
            buf( (k+1):(k+m) ) = buf( (k-m+1):k );
            buf( (k+m+1):(k+m+m) ) = buf( (k+1):(k+m) );
            buf( (k+m+m+1):idimb ) = buf( (k+m+1):(idimb-m) );
        else
            buf( (k+1):(k+m) ) = buf( (k-m+1):k );
            buf( (k+m+1):idimb ) = buf( (k+1):(idimb-m) );
        end
    else
        buf( (k+1):idimb ) = buf( (k+1-m):(idimb-m) );
    end
end
end

if abs(frac) > 1.0e-4
    if strcmp( type, 'long' ) == 1
        [ buf, buf2 ] = ldelay( buf, start, idim, frac, m );
    else
        [ buf, buf2 ] = delay( buf, start, idim, frac, m );
    end
    buf( (k+1):(k+idim) ) = buf2( 1:idim );
end

imin = k+1; imax = k+idim;
rar( 1:idim ) = rar( 1:idim ) + ( b(3) * buf( imin:imax ) );
buf( imin:imax ) = rar( 1:idim );

%+++++++ wavhdr ++++++++
function status = wavhdr( fpo, frames, frameSize, fmt )
WAVHDRSIZE = 44;
c = zeros(14,1);
dLen = frames * frameSize * 2;
rLen = dLen + WAVHDRSIZE - 8;
c(1) = fwrite( fpo, 'RIFF', 'char' ); % rId

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

c(2) = fwrite( fpo, rLen, 'long' );           % rLen
c(3) = fwrite( fpo, 'WAVE', 'char' );        % wId
c(4) = fwrite( fpo, 'fmt ', 'char' );        % fId
c(5) = fwrite( fpo, 16+length(fmt), 'long');  % fLen = 16 +
fmtSpecific field
c(6) = fwrite( fpo, 1, 'short' );            % wfmtTag (1=PCM)
c(7) = fwrite( fpo, 1, 'short' );            % nchan
c(8) = fwrite( fpo, 8000, 'long' );           % nSampSec
c(9) = fwrite( fpo, 16000, 'long' );          % nAvgBytesSec
c(10) = fwrite( fpo, 2, 'short' );            % nBlkAlign
c(11) = fwrite( fpo, 16, 'short' );           % fmtSpecific field -
bits/sample
c(12) = fwrite( fpo, fmt, 'char' );           % fmtSpecific field -
bits/sample
c(13) = fwrite( fpo, 'data', 'char' );        % dId
c(14) = fwrite( fpo, dLen, 'long' );          % dLen

```

```

if sum(c) ~= 25 + length(fmt)
    status = 0;
else
    status = 1;
end

```

```

%+++++++ DCOECBG ++++++
function [ cbg, bitpointer ] = dcoecbg( cbgbits, bitsum1, bitsum2, ...
    bitpointer, nn, stream )

```

```

global gainlog5
cbg = zeros( nn, 1 );
pointer = bitpointer;
for i = 1:nn
    [ index, pointer ] = unpack( stream, cbgbits, pointer );
    if cbgbits == 5
        cbg(i) = gainlog5( index+1 );
    else
        fprintf( 'gaindecode: Unquantized stochastic codebook
gain.\n' );
    end
    if any( [ 1 3 5 ] == i )
        pointer = pointer + bitsum2 - cbgbits;
    elseif any( [ 2 4 6 ] == i )
        pointer = pointer + bitsum1 - cbgbits;
    else
        fprintf( 'dcoecbg: Error in decoding stochastic codebook
gain\n' );
    end
end
end

```

```

bitpointer = bitpointer + cbgbits;

```

```

%+++++++ DCODCBI ++++++
function [ cbi, bitpointer ] = dcodcbi( cbbits, bitsum1, bitsum2, ...
    bitpointer, nn, stream )

```

```

cbi = zeros( nn, 1 );
pointer = bitpointer;

```

```

for i = 1:nn

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

[ cbi(i), pointer ] = unpack( stream, cbbits, pointer );
if any( [ 1 3 5 ] == i )
    pointer = pointer + bitsum2 - cbbits;
elseif any( [ 2 4 6 ] == i )
    pointer = pointer + bitsum1 - cbbits;
else
    fprintf( 'dcodcbi: Error in decoding code book indicies\n' );
end
end
cbi = cbi + 1;
bitpointer = bitpointer + cbbits;

%+++++++ DCOSPG ++++++
function [ pgs, bitpointer ] = dcospg( pgbits, bitsum1, bitsum2, ...
    bitpointer, nn, stream )

global pitch2max5
pgs = zeros( nn, 1 );
pointer = bitpointer;
for i = 1:nn
    [ index, pointer ] = unpack( stream, pgbits, pointer );
    pgs(i) = pitch2max5( index+1 );
    if ( i == 1 ) | ( i == 3 ) | ( i == 5 )
        pointer = pointer + bitsum2 - pgbits;
    elseif ( i == 2 ) | ( i == 4 ) | ( i == 6 )
        pointer = pointer + bitsum1 - pgbits;
    else
        fprintf( 'dcodpg: Error in decoding pitch gain\n' );
    end
end
bitpointer = bitpointer + pgbits;

%+++++++ DCODTAU ++++++
function [ taus, bitpointer ] = dcodtau( taubits, taudelta, bitsum1,
... bitsum2, bitpointer, nn, ... stream, pddecode, pdtabi )
global plevel1 plevel2 lptrDT pdelay
taus = zeros( nn, 1 );
pointer = bitpointer;
for i = 1:nn
    if rem( i,2 ) ~= 0
        [ tptr, pointer ] = unpack( stream, taubits, pointer );
        taus(i) = pddecode( tptr+1 );
        pointer = pointer + bitsum1 - taubits;
    else
        [ tptr, pointer ] = unpack( stream, taudelta, pointer );
        pointer = pointer + bitsum2 - taudelta;
        mnptr = lptrDT - ( (plevel2/2) - 1 );
        mxptr = lptrDT + (plevel2/2);
        if mnptr < 0
            mnptr = 0;
        end
        if mxptr > ( plevel1-1 )
            mnptr = plevel1 - plevel2;
        end
        taus(i) = pdelay( tptr + mnptr + 1 );
    end
    lptrDT = pdtabi( tptr+1 );
end
end

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

bitpointer = bitpointer + taubits;

%+++++++ DCODHAM ++++++
function [ codeword, twoerror, synflag ] = decodham( codelength1,
hmatrix, ... syndrometable, ... paritybit, codeword )
global TRUE FALSE
parityflag = FALSE;
twoerror = FALSE;
errorflag = 0;
if parityflag == TRUE
    synflag = rem( sum( codeword(1:codelength1) ), 2 );
    errorflag = errorflag + ( paritybit ~= synflag );
end
synflag = 0;
for i = 1:codelength1
    if codeword(i) ~= 0
        synflag = bitxor( synflag, hmatrix(i), 16 );
    end
end
if synflag ~= 0
    if ( errorflag ~= 1 ) & ( parityflag == TRUE )
        twoerror = TRUE;
        return
    end
    j = syndrometable( synflag );
    codeword( j ) = ~codeword( j );
end
end

```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

%+++++++ Bitperm ++++++

124, 91, 2, 86, 103, 56, 114, 26, 71, 109, 40, 11,
76, 129, 63, 32, 46, 134, 18, 87, 140, 49, 133, 6,
96, 119, 60, 105, 23, 74, 94, 48, 120, 15, 110, 53,
62, 126, 29, 108, 47, 84, 38, 21, 104, 75, 61, 65,
3, 97, 141, 88, 137, 27, 95, 81, 115, 125, 10, 68,
41, 82, 106, 33, 44, 132, 90, 59, 19, 54, 4, 99,
39, 111, 24, 80, 100, 67, 13, 57, 123, 30, 93, 52,
83, 22, 128, 70, 117, 1, 136, 28, 77, 142, 64, 8,
45, 113, 34, 130, 73, 37, 17, 139, 122, 85, 5, 118,
50, 69, 107, 12, 78, 42, 43, 138, 16, 116, 51, 101,
25, 127, 36, 72, 112, 9, 55, 89, 79, 31, 121, 58,
66, 143, 20, 98, 131, 35, 7, 135, 102, 14, 92, 144

%+++++++ PDDECODE ++++++

54.66667,28.75000,27.25000,27.75000,110.00000,57.66667,55.33334,
56.33334,55.00000,28.50000,27.00000,27.50000,111.00000,57.33334,
55.66667,56.00000,109.00000,29.25000,54.33334,28.25000,114.00000,
58.33334,112.00000,56.66667,108.00000,29.00000,54.00000,28.00000,
115.00000,58.00000,113.00000,57.00000,73.66667,76.00000,74.66667,
76.33334,81.00000,77.66667,80.00000,77.00000,74.00000,75.66667,
75.00000,76.66667,79.66667,78.00000,79.33334,77.33334,37.00000,
38.00000,37.33334,38.33334,40.33334,78.66667,40.00000,39.33334,
74.33334,75.33334,37.66667,38.66667,79.00000,78.33334,39.66667,
39.00000,88.00000,89.00000,20.00000,123.00000,86.00000,87.00000,
20.33334,20.66667,83.00000,82.00000,40.66667,41.00000,84.00000,
85.00000,41.33334,41.66667,44.00000,44.33334,133.00000,134.00000,
43.33334,43.66667,21.33334,21.00000,22.00000,21.66667,128.00000,
127.00000,42.66667,43.00000,42.33334,42.00000,63.33334,31.50000,
130.00000,129.00000,63.66667,31.75000,135.00000,136.00000,66.00000,
33.00000,67.33334,34.00000,65.33334,32.50000,67.66667,68.00000,
64.33334,32.25000,65.00000,64.66667,64.00000,32.00000,132.00000,
131.00000,66.33334,33.25000,66.66667,33.50000,65.66667,32.75000,
67.00000,33.75000,52.33334,102.00000,53.33334,106.00000,25.00000,
49.66667,26.50000,99.00000,51.33334,50.66667,53.00000,105.00000,
25.33334,50.33334,26.25000,101.00000,52.00000,103.00000,53.66667,
107.00000,24.66667,49.33334,26.75000,98.00000,51.66667,51.00000,
52.66667,104.00000,25.66667,50.00000,26.00000,100.00000,48.33334,
97.00000,45.33334,91.00000,48.00000,96.00000,45.00000,90.00000,
24.33334,48.66667,44.66667,139.00000,24.00000,49.00000,22.33334,
147.00000,46.66667,94.00000,46.00000,93.00000,47.33334,95.00000,
45.66667,92.00000,23.33334,47.00000,23.00000,46.33334,23.66667,
47.66667,22.66667,143.00000,34.66667,34.33334,35.33334,35.00000,
68.33334,68.66667,70.33334,69.66667,69.33334,69.00000,70.66667,
70.00000,137.00000,138.00000,141.00000,140.00000,36.66667,36.33334,
35.66667,36.00000,73.33334,72.66667,71.00000,72.00000,73.00000,
72.33334,71.33334,71.66667,146.00000,145.00000,142.00000,144.00000,
62.33334,30.75000,61.66667,30.25000,62.00000,30.50000,61.33334,
30.00000,121.00000,119.00000,60.33334,59.33334,61.00000,60.00000,
60.66667,59.66667,63.00000,31.00000,124.00000,29.50000,62.66667,
31.25000,125.00000,29.75000,126.00000,116.00000,117.00000,58.66667,
122.00000,118.00000,120.00000,59.00000

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

%+++++++ PDECODE ++++++

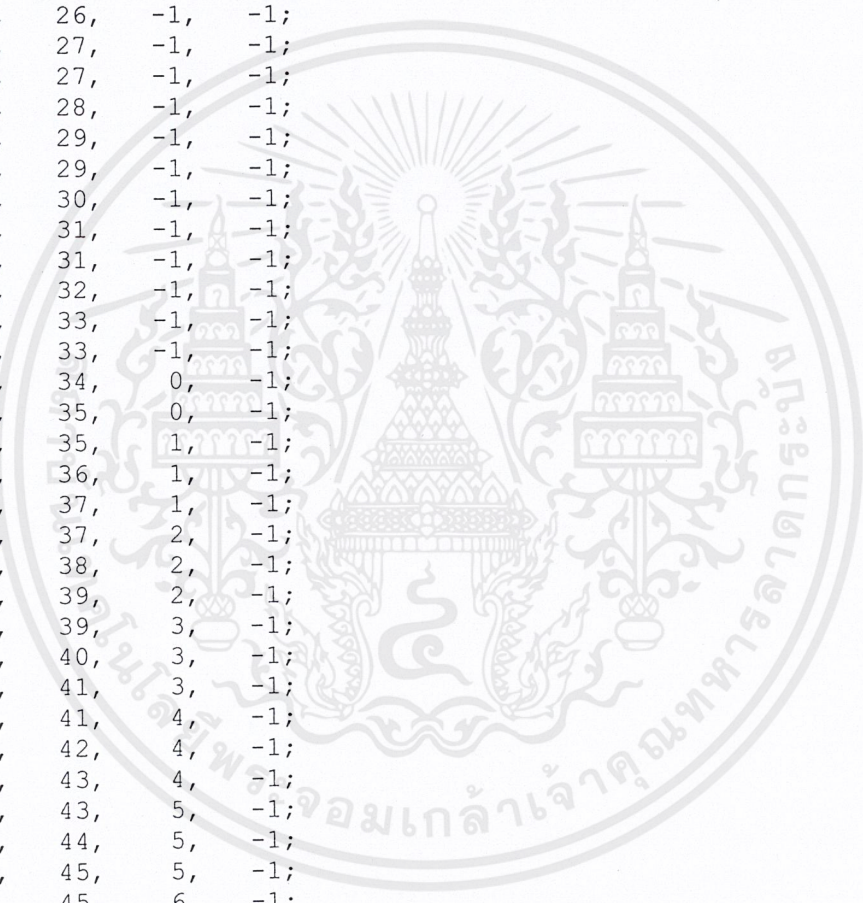
66, 70, 71, 87, 86, 89, 88, 174, 190, 186, 184, 188, 172,
168, 148, 132, 140, 156, 158, 142, 134, 150, 10, 2, 11, 3,
27, 19, 9, 1, 25, 17, 243, 247, 231, 227, 229, 225, 241,
245, 97, 101, 117, 113, 109, 125, 105, 121, 123, 127, 107, 193,
192, 195, 194, 210, 211, 209, 208, 48, 50, 58, 49, 51, 59,
63, 55, 62, 54, 52, 74, 75, 78, 79, 95, 94, 92, 93,
84, 85, 80, 81, 170, 166, 162, 182, 178, 187, 176, 185, 180,
189, 164, 160, 169, 173, 149, 133, 157, 141, 137, 153, 136, 152,
144, 128, 154, 138, 130, 146, 26, 18, 0, 8, 6, 14, 15,
7, 23, 31, 13, 5, 29, 21, 251, 255, 235, 239, 237, 234,
238, 236, 230, 226, 228, 224, 244, 240, 96, 100, 116, 112, 115,
114, 108, 124, 104, 120, 122, 126, 106, 110, 111, 196, 197, 201,
200, 199, 203, 198, 202, 214, 218, 219, 215, 217, 213, 216, 212,
32, 40, 56, 34, 42, 57, 41, 33, 35, 43, 39, 47, 37,
45, 61, 53, 60, 46, 44, 38, 36, 73, 72, 76, 77, 68,
69, 64, 65, 167, 163, 183, 179, 177, 181, 165, 161, 151, 135,
159, 143, 129, 145, 155, 139, 131, 147, 24, 16, 4, 12, 22,
30, 20, 28, 249, 250, 253, 233, 254, 232, 252, 67, 242, 246,
248, 91, 90, 99, 98, 119, 118, 82, 83, 102, 103, 204, 205,
171, 207, 206, 222, 191, 223, 221, 220, 175

%+++++++ PDELAY ++++++

20.00000, 20.33334, 20.66667, 21.00000, 21.33334, 21.66667, 22.00000,
22.33334, 22.66667, 23.00000, 23.33334, 23.66667, 24.00000, 24.33334,
24.66667, 25.00000, 25.33334, 25.66667, 26.00000, 26.25000, 26.50000,
26.75000, 27.00000, 27.25000, 27.50000, 27.75000, 28.00000, 28.25000,
28.50000, 28.75000, 29.00000, 29.25000, 29.50000, 29.75000, 30.00000,
30.25000, 30.50000, 30.75000, 31.00000, 31.25000, 31.50000, 31.75000,
32.00000, 32.25000, 32.50000, 32.75000, 33.00000, 33.25000, 33.50000,
33.75000, 34.00000, 34.33334, 34.66667, 35.00000, 35.33334, 35.66667,
36.00000, 36.33334, 36.66667, 37.00000, 37.33334, 37.66667, 38.00000,
38.33334, 38.66667, 39.00000, 39.33334, 39.66667, 40.00000, 40.33334,
40.66667, 41.00000, 41.33334, 41.66667, 42.00000, 42.33334, 42.66667,
43.00000, 43.33334, 43.66667, 44.00000, 44.33334, 44.66667, 45.00000,
45.33334, 45.66667, 46.00000, 46.33334, 46.66667, 47.00000, 47.33334,
47.66667, 48.00000, 48.33334, 48.66667, 49.00000, 49.33334, 49.66667,
50.00000, 50.33334, 50.66667, 51.00000, 51.33334, 51.66667, 52.00000,
52.33334, 52.66667, 53.00000, 53.33334, 53.66667, 54.00000, 54.33334,
54.66667, 55.00000, 55.33334, 55.66667, 56.00000, 56.33334, 56.66667,
57.00000, 57.33334, 57.66667, 58.00000, 58.33334, 58.66667, 59.00000,
59.33334, 59.66667, 60.00000, 60.33334, 60.66667, 61.00000, 61.33334,
61.66667, 62.00000, 62.33334, 62.66667, 63.00000, 63.33334, 63.66667,
64.00000, 64.33334, 64.66667, 65.00000, 65.33334, 65.66667, 66.00000,
66.33334, 66.66667, 67.00000, 67.33334, 67.66667, 68.00000, 68.33334,
68.66667, 69.00000, 69.33334, 69.66667, 70.00000, 70.33334, 70.66667,
71.00000, 71.33334, 71.66667, 72.00000, 72.33334, 72.66667, 73.00000,
73.33334, 73.66667, 74.00000, 74.33334, 74.66667, 75.00000, 75.33334,
75.66667, 76.00000, 76.33334, 76.66667, 77.00000, 77.33334, 77.66667,
78.00000, 78.33334, 78.66667, 79.00000, 79.33334, 79.66667, 80.00000,
81.00000, 82.00000, 83.00000, 84.00000, 85.00000, 86.00000, 87.00000,
88.00000, 89.00000, 90.00000, 91.00000, 92.00000, 93.00000, 94.00000,
95.00000, 96.00000, 97.00000, 98.00000, 99.00000, 100.00000, 101.00000,
102.00000, 103.00000, 104.00000, 105.00000, 106.00000, 107.00000,

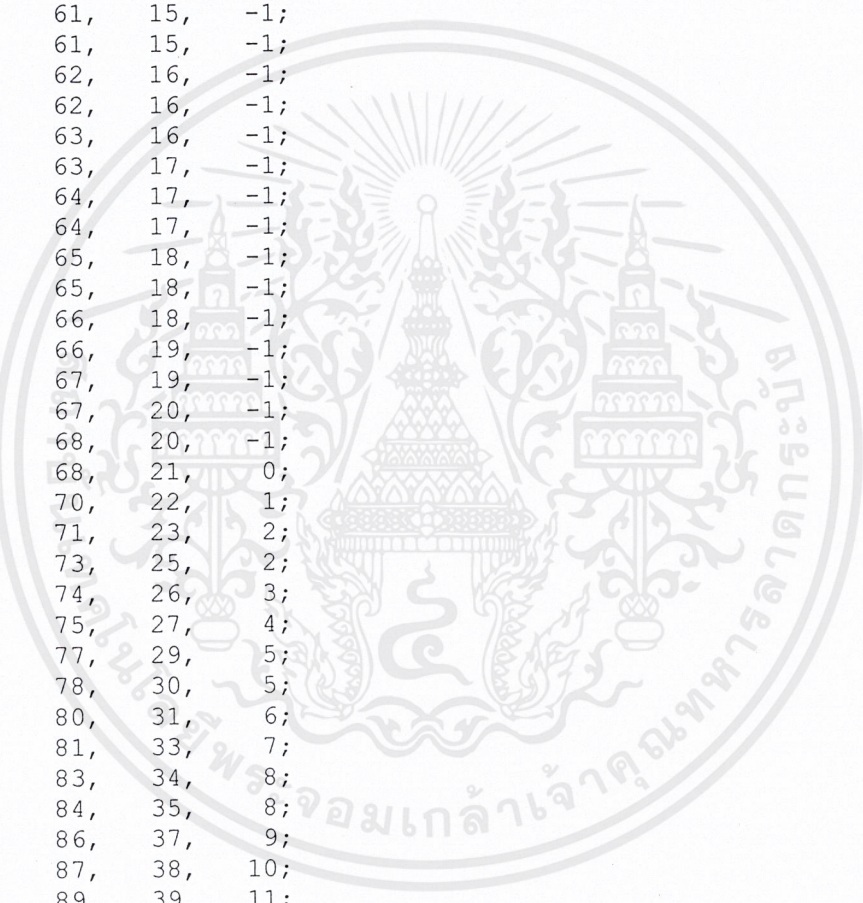
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

| | | | |
|----|-----|-----|-----|
| 1, | 18, | -1, | -1; |
| 1, | 18, | -1, | -1; |
| 1, | 19, | -1, | -1; |
| 1, | 19, | -1, | -1; |
| 1, | 20, | -1, | -1; |
| 1, | 21, | -1, | -1; |
| 1, | 21, | -1, | -1; |
| 1, | 22, | -1, | -1; |
| 1, | 23, | -1, | -1; |
| 1, | 23, | -1, | -1; |
| 1, | 24, | -1, | -1; |
| 1, | 25, | -1, | -1; |
| 1, | 25, | -1, | -1; |
| 1, | 26, | -1, | -1; |
| 1, | 27, | -1, | -1; |
| 1, | 27, | -1, | -1; |
| 1, | 28, | -1, | -1; |
| 1, | 29, | -1, | -1; |
| 1, | 29, | -1, | -1; |
| 1, | 30, | -1, | -1; |
| 1, | 31, | -1, | -1; |
| 1, | 31, | -1, | -1; |
| 1, | 32, | -1, | -1; |
| 1, | 33, | -1, | -1; |
| 1, | 33, | -1, | -1; |
| 2, | 34, | 0, | -1; |
| 2, | 35, | 0, | -1; |
| 2, | 35, | 1, | -1; |
| 2, | 36, | 1, | -1; |
| 2, | 37, | 1, | -1; |
| 2, | 37, | 2, | -1; |
| 2, | 38, | 2, | -1; |
| 2, | 39, | 2, | -1; |
| 2, | 39, | 3, | -1; |
| 2, | 40, | 3, | -1; |
| 2, | 41, | 3, | -1; |
| 2, | 41, | 4, | -1; |
| 2, | 42, | 4, | -1; |
| 2, | 43, | 4, | -1; |
| 2, | 43, | 5, | -1; |
| 2, | 44, | 5, | -1; |
| 2, | 45, | 5, | -1; |
| 2, | 45, | 6, | -1; |
| 2, | 46, | 6, | -1; |
| 2, | 47, | 6, | -1; |
| 2, | 47, | 7, | -1; |
| 2, | 48, | 7, | -1; |
| 2, | 49, | 7, | -1; |
| 2, | 49, | 8, | -1; |
| 2, | 50, | 8, | -1; |
| 2, | 51, | 8, | -1; |
| 2, | 51, | 9, | -1; |
| 2, | 52, | 9, | -1; |
| 2, | 52, | 9, | -1; |
| 2, | 53, | 10, | -1; |
| 2, | 53, | 10, | -1; |
| 2, | 54, | 10, | -1; |



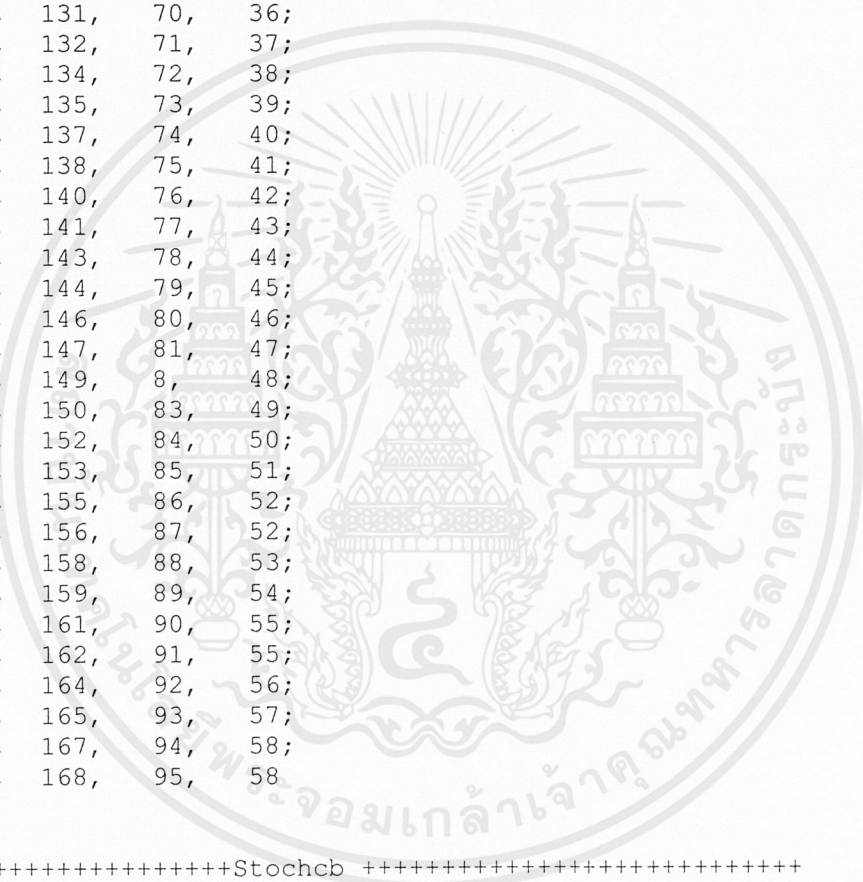
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

| | | | |
|----|------|-----|-----|
| 2, | 54, | 11, | -1; |
| 2, | 55, | 11, | -1; |
| 2, | 55, | 11, | -1; |
| 2, | 56, | 12, | -1; |
| 2, | 56, | 12, | -1; |
| 2, | 57, | 12, | -1; |
| 2, | 57, | 13, | -1; |
| 2, | 58, | 13, | -1; |
| 2, | 58, | 13, | -1; |
| 2, | 59, | 14, | -1; |
| 2, | 59, | 14, | -1; |
| 2, | 60, | 14, | -1; |
| 2, | 60, | 15, | -1; |
| 2, | 61, | 15, | -1; |
| 2, | 61, | 15, | -1; |
| 2, | 62, | 16, | -1; |
| 2, | 62, | 16, | -1; |
| 2, | 63, | 16, | -1; |
| 2, | 63, | 17, | -1; |
| 2, | 64, | 17, | -1; |
| 2, | 64, | 17, | -1; |
| 2, | 65, | 18, | -1; |
| 2, | 65, | 18, | -1; |
| 2, | 66, | 18, | -1; |
| 2, | 66, | 19, | -1; |
| 2, | 67, | 19, | -1; |
| 2, | 67, | 20, | -1; |
| 2, | 68, | 20, | -1; |
| 3, | 68, | 21, | 0; |
| 3, | 70, | 22, | 1; |
| 3, | 71, | 23, | 2; |
| 3, | 73, | 25, | 2; |
| 3, | 74, | 26, | 3; |
| 3, | 75, | 27, | 4; |
| 3, | 77, | 29, | 5; |
| 3, | 78, | 30, | 5; |
| 3, | 80, | 31, | 6; |
| 3, | 81, | 33, | 7; |
| 3, | 83, | 34, | 8; |
| 3, | 84, | 35, | 8; |
| 3, | 86, | 37, | 9; |
| 3, | 87, | 38, | 10; |
| 3, | 89, | 39, | 11; |
| 3, | 90, | 41, | 11; |
| 3, | 92, | 42, | 12; |
| 3, | 93, | 43, | 13; |
| 3, | 95, | 45, | 14; |
| 3, | 96, | 46, | 14; |
| 3, | 98, | 47, | 15; |
| 3, | 99, | 49, | 16; |
| 3, | 101, | 50, | 17; |
| 3, | 102, | 51, | 17; |
| 3, | 104, | 52, | 18; |
| 3, | 105, | 53, | 19; |
| 3, | 107, | 54, | 20; |
| 3, | 108, | 55, | 21; |
| 3, | 110, | 56, | 22; |



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

| | | | |
|----|------|-----|-----|
| 3, | 111, | 57, | 23; |
| 3, | 113, | 58, | 24; |
| 3, | 114, | 59, | 25; |
| 3, | 116, | 60, | 26; |
| 3, | 117, | 61, | 27; |
| 3, | 119, | 62, | 28; |
| 3, | 120, | 63, | 29; |
| 3, | 122, | 64, | 30; |
| 3, | 123, | 65, | 31; |
| 3, | 125, | 66, | 32; |
| 3, | 126, | 67, | 33; |
| 3, | 128, | 68, | 34; |
| 3, | 129, | 69, | 35; |
| 3, | 131, | 70, | 36; |
| 3, | 132, | 71, | 37; |
| 3, | 134, | 72, | 38; |
| 3, | 135, | 73, | 39; |
| 3, | 137, | 74, | 40; |
| 3, | 138, | 75, | 41; |
| 3, | 140, | 76, | 42; |
| 3, | 141, | 77, | 43; |
| 3, | 143, | 78, | 44; |
| 3, | 144, | 79, | 45; |
| 3, | 146, | 80, | 46; |
| 3, | 147, | 81, | 47; |
| 3, | 149, | 8, | 48; |
| 3, | 150, | 83, | 49; |
| 3, | 152, | 84, | 50; |
| 3, | 153, | 85, | 51; |
| 3, | 155, | 86, | 52; |
| 3, | 156, | 87, | 52; |
| 3, | 158, | 88, | 53; |
| 3, | 159, | 89, | 54; |
| 3, | 161, | 90, | 55; |
| 3, | 162, | 91, | 55; |
| 3, | 164, | 92, | 56; |
| 3, | 165, | 93, | 57; |
| 3, | 167, | 94, | 58; |
| 3, | 168, | 95, | 58; |



§++++++Stochcb++++++§

| | | | | | | | | | |
|------|------|------|------|------|-----|------|------|------|------|
| 0., | 1., | 0., | 0., | 0., | 1., | 0., | 0., | 0., | 0., |
| 0., | 0., | 1., | 0., | 1., | 0., | 0., | 0., | 0., | -1., |
| 0., | 0., | -1., | 0., | -1., | 0., | 1., | 0., | 0., | 0., |
| 0., | 0., | 0., | 0., | -1., | 0., | -1., | -1., | 0., | 0., |
| -1., | 0., | -1., | 0., | -1., | 0., | -1., | 0., | 0., | 0., |
| 0., | 0., | 0., | 0., | -1., | 0., | 0., | 0., | 1., | 0., |
| 0., | 0., | 0., | -1., | 0., | 0., | 0., | 0., | 1., | 0., |
| 0., | 0., | 0., | 0., | 0., | 0., | 0., | 1., | -1., | 0., |
| 1., | 0., | 1., | 0., | -1., | 1., | 0., | 0., | 0., | 0., |
| 0., | -1., | 0., | 0., | 1., | 0., | 0., | 0., | -1., | 0., |
| 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., |
| 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., |
| 0., | -1., | -1., | 0., | 0., | 0., | 0., | 0., | 0., | 0., |
| 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | -1., | -1., |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

| | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0. | 1. | -1. | 0. | 0. | -1. | -1. | 0. | 0. | 0. |
| 0. | 1. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. |
| 0. | 0. | 0. | -1. | 0. | -1. | -1. | 0. | 0. | -1. |
| 0. | 1. | 0. | 0. | 0. | 1. | 0. | 0. | 0. | 0. |
| 0. | -1. | 0. | 0. | 0. | -1. | 0. | 0. | 0. | 0. |
| 0. | 0. | -1. | 0. | 1. | 0. | 0. | 0. | 0. | 0. |
| 0. | 0. | -1. | 1. | 0. | 0. | 0. | -1. | 1. | 0. |
| 0. | 1. | 0. | 0. | -1. | 0. | 0. | 0. | 0. | 0. |
| 0. | -1. | 0. | 0. | 1. | -1. | 0. | 0. | 1. | 0. |
| 0. | 0. | 0. | 0. | 0. | 1. | 1. | 1. | 0. | 0. |
| 1. | 1. | 0. | 0. | 0. | -1. | 0. | 0. | 0. | -1. |
| 0. | 0. | -1. | 1. | 0. | 0. | 0. | 0. | 0. | -1. |
| -1. | 0. | 0. | 0. | 0. | -1. | 0. | 0. | 0. | 0. |
| 0. | -1. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. |
| 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. |
| 0. | 1. | 1. | 0. | -1. | -1. | 0. | 0. | 1. | 0. |
| 0. | 0. | 0. | 0. | 0. | -1. | 0. | 0. | -1. | 1. |
| 0. | 0. | 0. | 0. | 0. | 0. | -1. | 0. | 0. | -1. |
| 0. | 0. | 0. | 0. | 0. | -1. | 0. | 0. | -1. | 0. |
| 0. | 0. | 0. | 0. | -1. | -1. | -1. | 0. | -1. | -1. |
| 0. | 1. | 0. | 0. | 0. | 0. | 1. | 0. | 0. | 0. |
| 1. | 0. | 0. | 0. | 0. | 1. | -1. | 0. | 0. | 0. |
| 0. | 0. | 0. | 1. | -1. | 0. | 0. | -1. | 0. | -1. |
| 0. | 0. | -1. | 0. | 0. | 1. | 0. | 0. | 0. | 0. |
| -1. | -1. | 0. | -1. | 0. | 0. | 0. | 0. | 0. | 0. |
| 0. | 0. | 0. | 1. | 0. | 0. | -1. | 0. | 0. | 0. |
| -1. | 0. | 1. | 0. | 0. | -1. | 0. | 0. | 0. | -1. |
| 0. | 0. | 0. | -1. | 0. | 0. | 0. | 0. | -1. | 0. |
| 0. | -1. | -1. | 0. | 0. | 0. | 1. | 0. | -1. | 0. |
| 0. | 0. | 0. | 0. | 1. | 0. | 0. | 0. | 0. | 0. |
| 0. | 0. | 1. | 1. | 0. | 0. | 0. | 1. | 1. | 0. |
| 0. | 0. | -1. | 0. | 0. | 0. | 0. | 0. | -1. | 1. |
| 1. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 1. |
| 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. |
| 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | -1. | 0. |
| 1. | 1. | 0. | 0. | 0. | 0. | 1. | 0. | 0. | 1. |
| 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. |
| 0. | 0. | 0. | 0. | 1. | 0. | 0. | 1. | 0. | -1. |
| -1. | 1. | 0. | 0. | 0. | 1. | 0. | 1. | 0. | 1. |
| 0. | 0. | -1. | 0. | 0. | 0. | 0. | 0. | 0. | 0. |
| 0. | -1. | 0. | -1. | 0. | 0. | 0. | -1. | 0. | 0. |
| -1. | 0. | 1. | 1. | 0. | 0. | 0. | 0. | 0. | 0. |
| 0. | 0. | 0. | 0. | 0. | 0. | 0. | -1. | -1. | 0. |
| 0. | 0. | 0. | 1. | 0. | 0. | 0. | 0. | 0. | 0. |
| -1. | 0. | -1. | -1. | 0. | 0. | -1. | 0. | -1. | 0. |
| 0. | 0. | 0. | 0. | 1. | 0. | 1. | 0. | 0. | 0. |
| 0. | 0. | 0. | 1. | -1. | 0. | 1. | 0. | 0. | 0. |
| 0. | -1. | 0. | 0. | 0. | 0. | 0. | 0. | 1. | 0. |
| 0. | -1. | -1. | 0. | 0. | 0. | 0. | 0. | 0. | 1. |
| 0. | 0. | 0. | 0. | 1. | -1. | 0. | 1. | 0. | 1. |
| 0. | 1. | 0. | 0. | 0. | 1. | 0. | 0. | 0. | 0. |
| -1. | 0. | 0. | 0. | -1. | 0. | 0. | 0. | 0. | 0. |
| 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. |
| 0. | 0. | 0. | 0. | 0. | 0. | 0. | -1. | 0. | 0. |
| 0. | 0. | 0. | 0. | 0. | 0. | 1. | 0. | 0. | -1. |
| 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. | 0. |
| 1. | 0. | 0. | 1. | 0. | 0. | 1. | 0. | 0. | 0. |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

| | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|------|
| 0., | 0., | 0., | -1., | 1., | 0., | 1., | 0., | 0., | -1., |
| 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 1., |
| -1., | 0., | 0., | 0., | -1., | 0., | 0., | 0., | 0., | -1., |
| 0., | 0., | 0., | 0., | 1., | -1., | 0., | -1., | 1., | 0., |
| 0., | 0., | 0., | -1., | 0., | 0., | 0., | 1., | 0., | 0., |
| 0., | 0., | 0., | -1., | 0., | 0., | 0., | 0., | 0., | 0., |
| -1., | 0., | -1., | 1., | 1., | -1., | 0., | 0., | 0., | 1., |
| 0., | 1., | 0., | 0., | 0., | -1., | 0., | 0., | 0., | 0., |
| 0., | 0., | 0., | 0., | 0., | 0., | 1., | 0., | 1., | 0., |
| 0., | 0., | 0., | 0., | -1., | 0., | 0., | 0., | 1., | 0., |
| 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., |
| 0., | 0., | -1., | 0., | 0., | 0., | 1., | 0., | 0., | 0., |
| 0., | -1., | 0., | 1., | -1., | 0., | 1., | 0., | 0., | 0., |
| 0., | 0., | -1., | 0., | 0., | 0., | 1., | 0., | -1., | 1., |
| -1., | 0., | 0., | 1., | 0., | 0., | 0., | 0., | -1., | 0., |
| 0., | 0., | -1., | 0., | 0., | 0., | 0., | -1., | 0., | 0., |
| 0., | 0., | 1., | 0., | 0., | 0., | -1., | 0., | 0., | 0., |
| 0., | 0., | 0., | 1., | 0., | 0., | 0., | 0., | 0., | 0., |
| 1., | -1., | 0., | 0., | 1., | 0., | 0., | 0., | 0., | -1., |
| 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., |
| -1., | 0., | 0., | 0., | 1., | 0., | 0., | 1., | 0., | 0., |
| 0., | 0., | 1., | 0., | 0., | 1., | 0., | 0., | -1., | 0., |
| 0., | 0., | 1., | 0., | 0., | 0., | 0., | 0., | 0., | 0., |
| 1., | 0., | 0., | -1., | 0., | 0., | -1., | 0., | 0., | 0., |
| 1., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., |
| 0., | 0., | 0., | 1., | 0., | 0., | 0., | 0., | 0., | 1., |
| 0., | 0., | -1., | 1., | -1., | 0., | 1., | 0., | 0., | 0., |
| 0., | 0., | 0., | 0., | 1., | -1., | 0., | 0., | 0., | 0., |
| 0., | 0., | 0., | 0., | -1., | 1., | 0., | 0., | 0., | 0., |
| 0., | 0., | 0., | 0., | 0., | 0., | 0., | 1., | 0., | 0., |
| 0., | 0., | 0., | -1., | 0., | 0., | 0., | 0., | -1., | 0., |
| 0., | 0., | -1., | 0., | -1., | 0., | 0., | 0., | 0., | 0., |
| 0., | -1., | 0., | 0., | -1., | 0., | 0., | 0., | 0., | 0., |
| 0., | 0., | 0., | 1., | -1., | 0., | 0., | 0., | 0., | -1., |
| 0., | 1., | 0., | 0., | 1., | 0., | 0., | 0., | 0., | -1., |
| 0., | 0., | 0., | 0., | 1., | 0., | 0., | 0., | 0., | -1., |
| 0., | 0., | -1., | 0., | 0., | 0., | 0., | 0., | 0., | 1., |
| 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0., | 0. |

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ปริญญาบัตรชั้นนี้ สำเร็จลุล่วงได้ด้วยดี ด้วยความช่วยเหลือจากบุคคลหลายๆท่าน ทั้งในด้านความรู้ ความช่วยเหลือ เครื่องมือและอุปกรณ์ต่างๆ ตลอดจนถึงการสั่งสอน ให้คำแนะนำ อีกทั้งยังคอยเป็นกำลังใจในปริญญาบัตรชั้นนี้ ซึ่งต้องกล่าวถึงบุคคลเหล่านี้

บิดา มารดา ที่คอยเป็นกำลังใจและสนับสนุนในเรื่องต่างๆ ในการทำปริญญาบัตรชั้นนี้

อาจารย์กฤษณ์ วงจรูระ และ ดร.สุทธิชัย นพาคีพงษ์ อาจารย์ที่ปรึกษาในการทำปริญญาบัตร ที่คอยให้คำแนะนำ ความช่วยเหลือ ให้คำปรึกษาในเรื่องต่างๆ และยังให้ความสนับสนุนในทุกๆด้าน

คุณเศรษฐกร กาเมือง, คุณพงศ์ศิริ จุฑาจรัส, คุณพดุมพิงษ์ เลิศชัยมงคล ที่คอยให้คำแนะนำต่างๆ ให้คำปรึกษาในเรื่องข้อมูลที่เป็นประโยชน์อย่างสูงแก่ปริญญาบัตรชั้นนี้

คุณวฤช คุณสุวรรณ, คุณจรรุวรรณ อังสนากุล, คุณต่อลาภ วินิจมงคลสิน ที่ช่วยให้รายงานฉบับนี้ สำเร็จลุล่วงไปได้ด้วยดี

คุณวริทธิ สันธิเดช, คุณอรนุช เลิศสุวรรณกิจ, คุณฤทธิษฐ์ หิตะวัฒน์กุล, คุณวรุศม์ บุญธรรม, คุณอังคณา โชติเวชกุล, คุณสงคราม สุนทรานู, คุณอังศุมารินทร์ สุพรรณรัตน์, คุณชนินทร์ มหารักษ์ ที่คอยเป็นกำลังใจให้เสมอมา นับตั้งแต่เริ่มปริญญาบัตรชั้นนี้

คณะผู้จัดทำ

1 เมษายน 2544

เอกสารอ้างอิง

- [1] Panos E.Papamichalis, Texas Instruments, Inc. Rice University, "Practical Approaches to Speech Coding":Prentice-Hal , Inc. 1987
- [2] Lawrence Rabiner and Biing-Hwang Jung, "Fundamental of Speech Recognition , New Jersey : Prentice Hall , 1993
- [3] Speech Analysis and Synthesis by linear prediction of the speech wave, B.S. ATAL and SUZANNE L.HANAUER, Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey 07974
- [4] Codebook-excited linear predictive (CELP) coding can reduce the bit rate required for digital telephony by an order of magnitude, Roger Dettmer
- [5] A new LPC Interpolation Technique for CELP Coders, Mei yong
- [6] Neural Code-Excited Linear Prediction for Low Power Speech Compression, SriGouri Kamarsu and H.C. Card, Department of Electrical and Computer Engineering University of Manitoba, Winnipeg, MBR3T 5V6
- [7] Analysis and Quantization Procedures for a Real-Time Implementation of a 4.8 kb/s CELP coder, F.J. Casajus-Quiros, L.A. Henandez-Gomez* and C.Garcia-Mateo
- [8] A 4.8 kbps Homomorphic Vocoder Using Analysis-by-Synthesis Excitation Analysis, Jae H. Chung and Ronald W. Schafer, Georgia Institute of Technology School of Electrical Engineering Atlanta, Georgia 30332, U.S.A.
- [9] Fundamental of Digital Signal Processing, Lonnie C. Ludeman, New Mexico State University, John Weley & Sons
- [10] Fast CELP Algorithm and Implementation for Speech Compression, Department of Electrical and Computer Engineering and Telecommunication Research Laboratories, University of Manitoba
- [11] อภิชาติ เอกผจญสวัสดิ์ และ อานนท์ ดารามาลัย, การจดจำเสียงพูด, วิทยานิพนธ์ ปีการศึกษา 2541, ภาควิชาวิศวกรรมโทรคมนาคม, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง