

การลดความซับซ้อนสำหรับระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสตรีมมิ่ง
โดยใช้โครงสร้างการรองรับแถบย่อยขึ้นกับการประมวลผลสัญญาณในโดเมนความถี่



นายบุญชัย กฤตยานันต์

สถาบันวิทยบริการ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า


คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2548

ISBN 974-53-2040-4

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

COMPLEXITY REDUCTION FOR STEREOPHONIC ACOUSTIC ECHO CANCELLATION
USING SUBBAND FILTERING STRUCTURE BASED ON FREQUENCY-DOMAIN PROCESSING



Mr.Boonchai Kritayanun

สถาบันวิทยบริการ
A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Electrical Engineering
Department of Electrical Engineering

Faculty of Engineering
Chulalongkorn University

Academic Year 2005

ISBN 974-53-2040-4

บุญชัย กฤตยานันต์ : การลดความซับซ้อนสำหรับระบบการตัดสัญญาณเสียงสะท้อนในการ
 สัมมนาแบบสเตอริโอโดยใช้โครงสร้างการกรองแถบย่อยขึ้นกับการประมวลผลสัญญาณในโดเมน
 ความถี่ (COMPLEXITY REDUCTION FOR STEREOPHONIC ACOUSTIC ECHO
 CANCELLATION USING SUBBAND FILTERING STRUCTURE BASED ON FREQUENCY-
 DOMAIN PROCESSING) อ. ที่ปรึกษา : ผศ.ดร.นิศาชล ตั้งเสงี่ยมวิสัย, 96 หน้า. ISBN 974-53-
 2040-4.

วิทยานิพนธ์นี้นำเสนอโครงสร้างไฮบริดที่ใช้การกรองแถบย่อยร่วมกับการประมวลผลในโดเมน
 ความถี่เพื่อปรับปรุงสมรรถนะของระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอ (SAEC) ใน
 ด้านอัตราการรู้ของระบบที่เร็วขึ้น ในขณะที่โครงสร้างที่นำเสนอมีความซับซ้อนทางการคำนวณที่ต่ำกว่าของ
 ระบบ SAEC แบบปกติ เทคนิคที่นำเสนอได้เลือกใช้ขั้นตอนวิธี AP2 สำหรับการประมวลผลสัญญาณในแถบ
 ความถี่ต่ำ ในแถบความถี่สูงจะเลือกใช้ขั้นตอนวิธี F-NLMS แบบหนึ่งช่องสัญญาณที่มีการประมวลผลใน
 โดเมนความถี่เพื่อลดความซับซ้อนทางการคำนวณของระบบ SAEC โดยรวมให้ต่ำกว่าของระบบ SAEC
 แบบปกติ ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error) ในแต่ละแถบความถี่จะถูกใช้เป็น
 เกณฑ์ในการตัดสินใจการพิจารณาเลือกใช้ระหว่างขั้นตอนวิธี AP2 และขั้นตอนวิธี F-NLMS

สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา.....วิศวกรรมไฟฟ้า.....
 สาขาวิชา.....วิศวกรรมไฟฟ้า.....
 ปีการศึกษา 2548

ลายมือชื่อนิสิต.....
 ลายมือชื่ออาจารย์ที่ปรึกษา.....

4670697621 : MAJOR ELECTRICAL ENGINEERING

KEY WORD: STEREOPHONIC ACOUSTIC ECHO CANCELLATION / SUBBAND FILTERING / FAST AFFINE PROJECTION

BOONCHAI KRITTAYANUN : COMPLEXITY REDUCTION FOR STEREOPHONIC ACOUSTIC ECHO CANCELLATION USING SUBBAND FILTERING STRUCTURE BASED ON FREQUENCY-DOMAIN PROCESSING. THESIS ADVISOR : ASST.PROF.NISACHON TANGSANGIUMVISAI, Ph.D., 96 pp. ISBN 974-53-2040-4.

In this thesis, a hybrid structure employing subband and frequency-domain processing is proposed. The proposed technique aims to improve the performance of Stereophonic Acoustic Echo Cancellation (SAEC) systems, in terms of convergent rate, while keeping the computational complexity lower than that of the conventional fullband approaches. The proposed technique employs the two-channel version of Affine Projection (AP2) algorithm for processing the signal in low-frequency region. In the high-frequency region, the frequency-domain version of mono-channel Normalized Least Mean Square (F-NLMS) algorithm is chosen in this thesis for reducing the computational complexity of the SAEC system, as compared to the conventional one. The Mean Square Error (MSE) in each subband is used as a criterion for algorithm selection between AP2 and F-NLMS.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

DepartmentElectrical Engineering
Field of study....Electrical Engineering
Academic year....2005

Student's signature.....Boonchai Kritayanun.....
Advisor's signature.....Nisachon Tangsangiumvisai.....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี กระผมใคร่ขอกราบขอบพระคุณอย่างสูงสำหรับความช่วยเหลืออย่างดียิ่งของ ผู้ช่วยศาสตราจารย์ ดร.นิศาชล ตั้งเสงี่ยมวิสัย อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งท่านได้ให้คำแนะนำและข้อคิดเห็นต่าง ๆ พร้อมทั้งแรงกระตุ้นและแรงบันดาลใจในการทำวิจัยมาด้วยดีตลอดมา

ขอขอบคุณอาจารย์ทุก ๆ ท่านที่ได้ให้ความรู้ในการศึกษาในการนำไปใช้เกี่ยวกับงานวิจัยนี้

ขอขอบคุณโครงการวิจัยร่วมเสริมสร้างความเชื่อมโยงระหว่างภาควิชาวิศวกรรมไฟฟ้าและภาคเอกชนทางด้านการวิจัยและพัฒนา ในการใช้ห้องบันทึกเสียงเกี่ยวกับการทำงานวิจัยนี้

ขอขอบคุณ รุ่นพี่ รุ่นน้อง เพื่อน ๆ และคนรอบตัวผู้วิจัยทุก ๆ คน ในห้องปฏิบัติการวิจัย การประมวลผลสัญญาณดิจิทัลทุก ๆ คน สำหรับความช่วยเหลือและกำลังใจในการทำวิจัยตลอดมา ขอขอบคุณที่ทำให้วันและเวลาของผู้วิจัยผ่านไปอย่างมีความหมาย

สุดท้ายขอกราบขอบพระคุณบิดามารดาและครอบครัว ตลอดจนญาติ ๆ ทุกคนที่เป็นกำลังใจและให้การสนับสนุนแก่ผู้วิจัยมาโดยตลอดจนสำเร็จการศึกษา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ	ช
สารบัญตาราง.....	ฅ
สารบัญภาพ	ญ
บัญชีคำศัพท์.....	ท

บทที่

1. บทนำ.....	1
1.1 ความสำคัญของปัญหา	1
1.2 วัตถุประสงค์.....	3
1.3 เป้าหมายและขอบเขตงานวิจัย	4
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	4
1.5 ขั้นตอนและวิธีการดำเนินการ	4
2. หลักการและขั้นตอนวิธีสำคัญ.....	6
2.1 หลักการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี Steepest-Descent ²	7
2.2 หลักการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี Least Mean Square ²	12
2.3 หลักการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี Normalized Least Mean Square ²	13
2.4 หลักการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี Fast Normalized Least Mean Square ²	14
2.5 หลักการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี Affine Projection ²	18
2.6 หลักการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี Fast Least Square ²	21
2.7 ระบบการตัดสัญญาณเสียงสะท้อนด้วยโครงสร้างการกรองแถบย่อย	22
2.7.1 ระบบหลายอัตรา (Multirate System)	23
2.7.1.1 การทำเดซิเมชัน (Decimation)	23
2.7.1.2 การทำอินเทอร์โพลชัน (Interpolation).....	27
2.7.2 การกรองแถบย่อย (Subband Filtering)	30
2.7.3 การประยุกต์ใช้การกรองแถบย่อยในโดเมนเวลาและโดเมนความถี่	31

2.8 การเปรียบเทียบความซับซ้อนทางการคำนวณ (Computational Complexity)	33
3 โครงสร้างไฮบริดระหว่างขั้นตอนวิธีแบบโมโนและขั้นตอนวิธีแบบสเตริโอ	36
3.1 โครงสร้างไฮบริดระหว่างขั้นตอนวิธีแบบโมโนและขั้นตอนวิธีแบบสเตริโอ	36
3.2 ความซับซ้อนทางการคำนวณของโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS ใน [18].....	38
3.3 โครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอในวิทยานิพนธ์นี้	40
3.4 ความซับซ้อนทางการคำนวณของโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอในวิทยานิพนธ์นี้	43
4 ผลการจำลองแบบและการวิเคราะห์ผล	47
4.1 แบบจำลองของระบบ SAEC	47
4.2 ผลการจำลองแบบระบบ SAEC ที่มีการประมวลผลสัญญาณในโดเมนเวลา	49
4.3 ผลการจำลองแบบระบบ SAEC ที่มีการประมวลผลสัญญาณในโดเมนความถี่	51
4.4 ผลการจำลองแบบระบบ SAEC ที่มีการประยุกต์ใช้โครงสร้างการกรองแถบย่อย ร่วมกับการประมวลผลสัญญาณในโดเมนเวลาและโดเมนความถี่.....	52
4.5 การเปรียบเทียบความซับซ้อนทางการคำนวณ (Computational Complexity)	61
4.5.1 ขั้นตอนวิธี NLMS2	61
4.5.2 ขั้นตอนวิธี F-NLMS2	61
4.5.3 ขั้นตอนวิธี Subband NLMS2	63
4.5.4 Subband F-NLMS2	64
4.6 ผลการจำลองแบบโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18] เปรียบเทียบกับ โครงสร้างไฮบริดที่นำเสนอ	66
5 สรุปผลการวิจัยและข้อเสนอแนะ	76
5.1 สรุปผลการวิจัย	76
5.2 ข้อเสนอแนะสำหรับงานวิจัยในอนาคต	78
รายการอ้างอิง.....	79
ภาคผนวก.....	81
บทความที่ได้รับการเผยแพร่	82
ประวัติผู้เขียนวิทยานิพนธ์.....	95

สารบัญตาราง

ตาราง	หน้า
ตารางที่ 2.1 ความซับซ้อนทางการคำนวณของขั้นตอนวิธี NLMS ช่องสัญญาณที่ j	34
ตารางที่ 2.2 ความซับซ้อนทางการคำนวณของขั้นตอนวิธี F-NLMS ช่องสัญญาณที่ j	34
ตารางที่ 2.3 ความซับซ้อนทางการคำนวณของขั้นตอนวิธี AP ช่องสัญญาณที่ j	34
ตารางที่ 2.4 ความซับซ้อนทางการคำนวณของขั้นตอนวิธี FLS ช่องสัญญาณที่ j	35
ตารางที่ 4.1 ความซับซ้อนทางการคำนวณของขั้นตอนวิธีในโดเมนเวลา	51
ตารางที่ 4.2 การเปรียบเทียบความซับซ้อนทางการคำนวณของแต่ละขั้นตอนวิธีเทียบกับขั้นตอนวิธี NLMS2 เมื่อกำหนดให้ $N_{\text{tap}} = 64$ ค่า	65
ตารางที่ 4.3 การเปรียบเทียบข้อดีข้อเสียของแต่ละขั้นตอนวิธี	66
ตารางที่ 4.4 ความซับซ้อนทางการคำนวณของแต่ละขั้นตอนวิธีที่ศึกษาในวิทยานิพนธ์นี้	75

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญญภาพ

ภาพประกอบ	หน้า
รูปที่ 1.1 แบบจำลองระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตริโอ (SAEC) ...	2
รูปที่ 2.1 ระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตริโอ	6
รูปที่ 2.2 โครงสร้างการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี Steepest-Descent	8
รูปที่ 2.3 โครงสร้างการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี LMS2 ที่ประมวลผลในโดเมนเวลา	12
รูปที่ 2.4 โครงสร้างการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี F-NLMS2 ด้วยวิธีเก็บข้อมูลซ้อนเหลื่อม 50 เปอร์เซ็นต์	18
รูปที่ 2.5 การทำเดซิเมชัน	23
รูปที่ 2.6 ผลตอบสนองทางความถี่รูปแบบต่างๆ ของคลังวงจรรอง	24
รูปที่ 2.7 การลดอัตราการซ้กตัวอย่าง	25
รูปที่ 2.8 ตัวอย่างการลดอัตราการซ้กตัวอย่างด้วยตัวประกอบเดซิเมชัน $D = 2$	27
รูปที่ 2.9 การทำอินเทอร์โพลชัน	27
รูปที่ 2.10 การเพิ่มอัตราการซ้กตัวอย่าง	28
รูปที่ 2.11 ตัวอย่างการเพิ่มอัตราการซ้กตัวอย่างด้วยตัวประกอบอินเทอร์โพลชัน $I = 2$	29
รูปที่ 2.12 คลังวงจรรองควอแดรเจอร์-มิลเลอร์ สองช่องสัญญาณ ($M = 2$)	29
รูปที่ 2.13 การวิเคราะห์ด้วยคลังวงจรรองและการสังเคราะห์ด้วยคลังวงจรรอง	30
รูปที่ 2.14 ตัวอย่างโครงสร้างการกรองแถบย่อย M แถบความถี่ย่อย	31
รูปที่ 2.15 การประยุกต์ใช้การกรองแถบย่อยกับระบบ SAEC ในโดเมนเวลาด้วยขั้นตอนวิธี NLMS2	32
รูปที่ 2.16 การประยุกต์ใช้การกรองแถบย่อยกับระบบ SAEC ในโดเมนความถี่ด้วยขั้นตอนวิธี F-NLMS2	33
รูปที่ 3.1 บล็อกไดอะแกรมแสดงโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS ใน [18]	37
รูปที่ 3.2 ผลตอบสนองเชิงความถี่ของสัญญาณเสียงเข้าหนึ่งช่องสัญญาณที่ใช้ในวิทยานิพนธ์ฉบับนี้ เมื่อความถี่การซ้กตัวอย่างเท่ากับ 16 KHz	40
รูปที่ 3.3 ผลตอบสนองเชิงความถี่ของวงจรรองวิเคราะห์ที่มีโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอในวิทยานิพนธ์นี้	42
รูปที่ 3.4 บล็อกไดอะแกรมแสดงโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอในวิทยานิพนธ์นี้	43

รูปที่ 4.1 ลักษณะวิถีสีสะท้อนทางเสียงจำนวน 2,048 ค่าจากลำโพงตัวที่ 1 ไปที่ไมโครโฟนตัวที่ 1 ในห้องรับ (\mathbf{h}_1) และจากลำโพงตัวที่ 2 ไปที่ไมโครโฟนตัวที่ 1 ในห้องรับ (\mathbf{h}_2) ตามลำดับ 48

รูปที่ 4.2 สัญญาณเข้าของระบบที่เป็นสัญญาณเสียงพูดจากลำโพงตัวที่ 1 และตัวที่ 2 ตามลำดับ 48

รูปที่ 4.3 ค่า WEVN เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี NLMS2: $\mu = 0.5$ 50

รูปที่ 4.4 ค่า WEVN เปรียบเทียบสมรรถนะระหว่าง (1) ขั้นตอนวิธี NLMS2: $\mu = 0.5$ กับ (2) ขั้นตอนวิธี FLS2: $E_a(0) = E_b(0) = 1200, W = 1$ กับ (3) - (5) ขั้นตอนวิธี AP2: $\mu = 0.5$, $p=2, 3$ และ 5 ตามลำดับ 50

รูปที่ 4.5 ค่า WEVN เมื่อประมวลผลสัญญาณด้วย (1) ขั้นตอนวิธี NLMS2 ($\mu = 0.5$) เปรียบเทียบกับ (2) ขั้นตอนวิธี F-NLMS2 ($\mu = 0.3$) 52

รูปที่ 4.6 ความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูด 53

รูปที่ 4.7 ผลตอบสนองอิมพัลส์ของคลังวงจรกรองวิเคราะห์เอกรูปเมื่อ $M = 2$ 54

รูปที่ 4.8 ค่า WEVN เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี Subband NLMS2 : $M = 2, \mu = 0.04$ 54

รูปที่ 4.9 ค่า WEVN เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี Subband F-NLMS2 : $M = 2, \mu = 0.022$ 55

รูปที่ 4.10 ผลตอบสนองอิมพัลส์ของคลังวงจรกรองวิเคราะห์เอกรูปเมื่อ $M = 4$ 56

รูปที่ 4.11 ค่า WEVN เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี Subband NLMS2 : $M = 4, \mu = 0.004$ 56

รูปที่ 4.12 ค่า WEVN เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี Subband F-NLMS2 : $M = 4, \mu = 0.018$ 57

รูปที่ 4.13 ผลตอบสนองอิมพัลส์ของคลังวงจรกรองวิเคราะห์เอกรูปเมื่อ $M = 8$ 58

รูปที่ 4.14 ค่า WEVN แถบความถี่ย่อยที่ 0-3 เมื่อประมวลผลสัญญาณด้วย ขั้นตอนวิธี Subband NLMS2: $M = 8, \mu = 0.002$ 58

รูปที่ 4.15 ค่า WEVN แถบความถี่ย่อยที่ 4-7 เมื่อประมวลผลสัญญาณด้วย ขั้นตอนวิธี Subband NLMS2: $M = 8, \mu = 0.002$ 59

รูปที่ 4.16 กราฟแสดงค่า WEVN แถบความถี่ย่อยที่ 0-3 เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี Subband F-NLMS2: $M = 8, \mu = 0.01$ 59

รูปที่ 4.17 กราฟแสดงค่า WEVN แถบความถี่ย่อยที่ 4-7 เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี Subband F-NLMS2: $M = 8, \mu = 0.01$ 60

รูปที่ 4.18 การแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18]	67
รูปที่ 4.19 ค่า ERLE ของโครงสร้างไฮบริดระหว่าง FLS2 (0-1 kHz): $E_a(0) = E_b(0) = 10, W = 0.999$ และ NLMS (1-8 kHz): $\mu = 0.03$	67
รูปที่ 4.20 การแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอบนแบบที่หนึ่ง	68
รูปที่ 4.21 ค่า ERLE เปรียบเทียบสมรรถนะระหว่าง (1) โครงสร้างไฮบริดระหว่าง FLS2 (0-1 kHz): $E_a(0) = E_b(0) = 10, W = 0.999$ และ NLMS (1-8 kHz): $\mu = 0.03$ กับ (2) โครงสร้างไฮบริดที่นำเสนองระหว่าง AP2 (0-1 kHz): $p=2, \mu = 0.3$ และ F-NLMS (1-8 kHz): $\mu = 0.02$	68
รูปที่ 4.22 ค่า MSE ด้วยขั้นตอนวิธี NLMS2 ($M = 8$)	69
รูปที่ 4.23 การแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอบนแบบที่สอง	70
รูปที่ 4.24 ค่า ERLE เปรียบเทียบสมรรถนะระหว่าง (1) โครงสร้างไฮบริดระหว่าง FLS2 (0-1 kHz): $E_a(0) = E_b(0) = 10, W = 0.999$ และ NLMS (1-8 kHz): $\mu = 0.03$ กับ (2) โครงสร้างไฮบริดที่นำเสนองระหว่าง AP2 (0-4 kHz): $p=2, \mu = 0.3$ และ F-NLMS (4-8 kHz): $\mu = 0.02$	70
รูปที่ 4.25 การแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอบนแบบที่สาม	71
รูปที่ 4.26 ค่า ERLE เปรียบเทียบสมรรถนะระหว่าง (1) โครงสร้างไฮบริดระหว่าง FLS2 (0-1 kHz): $E_a(0) = E_b(0) = 10, W = 0.999$ และ NLMS (1-8 kHz): $\mu = 0.03$ กับ (2) โครงสร้างไฮบริดที่นำเสนองระหว่าง AP2 (0-4 kHz): $p=2, \mu = 0.3$ และ F-NLMS (4-8 kHz): $\mu = 0.02$ กับ (3) โครงสร้างไฮบริดที่นำเสนองระหว่าง AP2 (0-2 kHz): $p=2, \mu = 0.3$ และ F-NLMS (2-8 kHz): $\mu = 0.02$	71
รูปที่ 4.27 การแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอบนแบบที่สี่	72
รูปที่ 4.28 ค่า ERLE เปรียบเทียบสมรรถนะระหว่าง (1) โครงสร้างไฮบริดระหว่าง FLS2 (0-1 kHz): $E_a(0) = E_b(0) = 10, W = 0.999$ และ NLMS (1-8 kHz): $\mu = 0.03$ กับ (2) โครงสร้างไฮบริดที่นำเสนองระหว่าง AP2 (0-4 kHz): $p=2, \mu = 0.3$ และ F-NLMS	

(4-8 kHz): $\mu = 0.02$ กับ (3) โครงสร้างไฮบริดที่นำเสนองระหว่าง AP2 (0-1 kHz): $p=2$,
 $\mu = 0.3$ และ F-NLMS2 (1-4 kHz): $\mu = 0.02$ $\mu = 0.3$ และ F-NLMS2 (4-8 kHz):
 $\mu = 0.02$ 73

รูปที่ 4.29 ค่า ERLE เปรียบเทียบสมรรถนะระหว่าง (1) โครงสร้างไฮบริดระหว่าง FLS2 (0-1
 kHz): $E_d(0) = E_b(0) = 10$, $W = 0.999$ และ NLMS (1-8 kHz): $\mu = 0.03$ กับ
 (2) โครงสร้างไฮบริดที่นำเสนองระหว่าง AP2 (0-4 kHz): $p=2$, $\mu = 0.3$ และ F-NLMS
 (4-8 kHz): $\mu = 0.02$ กับ (3) โครงสร้างไฮบริดที่นำเสนองระหว่าง AP2 (0-1 kHz): $p=2$,
 $\mu = 0.3$ และ F-NLMS2 (1-4 kHz): $\mu = 0.02$ $\mu = 0.3$ และ F-NLMS2 (4-8 kHz):
 $\mu = 0.02$ กับ (4) ขั้นตอนวิธี NLMS2: $\mu = 0.5$ 74



สถาบันวิทยบริการ
 จุฬาลงกรณ์มหาวิทยาลัย

ดัชนีคำศัพท์

Abrupt Change	การเปลี่ยนแปลงแบบฉับพลัน
Acoustic Echo	สัญญาณเสียงสะท้อน
Acoustic Echo Path	วิถีสะท้อนทางเสียง
Adaptive Filter Updating	การปรับปรุงวงจรรองแบบปรับตัว
Adaptation Gain Updating	การปรับปรุงอัตราขยายแบบปรับตัว
Adaptive Filter	วงจรรองแบบปรับตัว
Aliasing Effect	ผลความผิดเพี้ยนภาพ
Analysis Filter Bank	คลังวงจรรองวิเคราะห์
A Priori Error	เวกเตอร์ความผิดพลาดค่าก่อนหน้า
Autocorrelation Matrix	เมทริกซ์อัตโนมัติสัมพันธ์
Background Noise	สัญญาณรบกวนพื้นหลัง
Backward Prediction Coefficients	สัมประสิทธิ์การทำนายย้อนกลับ
Block Processing	การประมวลผลแบบบล็อก
Computational Complexity	ความซับซ้อนทางการคำนวณ
Convergence Rate	อัตราการลู่เข้า
Convolution Sum	ผลบวกการประสาน
Cost Function	ฟังก์ชันต้นทุน
Coupling	การเชื่อมต่อ
Cross-correlation Vector	เวกเตอร์สหสัมพันธ์ไขว้
Cut-off Frequency	ความถี่ตัดออก
Decimation	เดซิเมชัน
Diagonal Matrix	เมทริกซ์ทแยงมุม
Discrete Fourier Transform	การแปลงฟูรีเยร์ไม่ต่อเนื่อง
Discrete Signal	สัญญาณวิยุต
Distortion Function	ฟังก์ชันผิดเพี้ยน
Diverge	ลู่ออก
Division	การหาร
Downsampling	การลดการชักตัวอย่าง
Eigenvalues	ค่าลักษณะเฉพาะ

Eigenvalue Spread	การกระจายของค่าลักษณะเฉพาะ
Eigenvectors	เวกเตอร์ลักษณะเฉพาะ
Even Distribution	การแจกแจงแบบคู่
Expectation	ค่าคาดหวัง
Exponential Decay	การลดระดับแบบเลขชี้กำลัง
Fast Fourier Transform	การแปลงฟูริเยร์แบบเร็ว
Filter Bank	คลังวงจรรอง
Finite Impulse Response	ผลตอบสนองอิมพัลส์แบบจำกัด
Forgetting Factor	ตัวประกอบลืม
Forward Prediction Coefficients	สัมประสิทธิ์การทำนายล่วงหน้า
Frequency Bin	บินความถี่
Frequency Response	ผลตอบสนองทางความถี่
Fullband Structure	โครงสร้างแถบเต็ม
Half Wave Rectifier	ตัวทำกระแสตรงแบบครึ่งคลื่น
High-pass	สูงผ่าน
Homogeneous Difference Equation	สมการผลต่างเอกพันธ์
Hybrid Structure	โครงสร้างไฮบริด
Identity Matrix	เมทริกซ์เอกลักษณ์
Implementation	การนำไปปฏิบัติได้
Inherent Delay	เวลาประวิงในตัว
Instantaneous Value	ค่า ณ ขณะใดขณะหนึ่ง
Integer	จำนวนเต็ม
Interpolation	อินเทอร์โพลेशन
Iteration	การวนซ้ำ
Jointly Stationary	สเตชันนารีร่วม
Leakage	การรั่วไหล
Linear Time-invariant	เชิงเส้นและไม่แปรเปลี่ยนตามเวลา
Linear Time-variant System	เชิงเส้นและแปรเปลี่ยนตามเวลา
Low-pass	ต่ำผ่าน
Marginally-overlapping	ซ้อนทับกันพอดี
Mean Square Error	ค่าคลาดเคลื่อนกำลังสองเฉลี่ย

Memory	หน่วยความจำ
Minimum	น้อยที่สุด
Misadjustment	ค่าผิดพลาดในการปรับแก้
Multirate System	ระบบหลายอัตรา
Non-flat Spectrum	สเปกตรัมไม่เรียบ
Non-linear Transformation	การแปลงไม่เชิงเส้น
Non-overlapping	ไม่ซ้อนเหลื่อมกัน
Normal Random Signal	สัญญาณสุ่มปรกติ
One-way Transmission	การส่งผ่านทางเดียว
Optimal	เหมาะสมที่สุด
Order	อันดับ
Orthogonal Set	เซตตั้งฉาก
Overlapping	ซ้อนเหลื่อมกัน
Overlap-save Method	วิธีเก็บข้อมูลซ้อนเหลื่อม
Parallel Processing	การประมวลผลแบบขนาน
Posterior Error	เวกเตอร์ความผิดพลาดค่าหลัง
Post-filter	โพลีฟิวเตอร์
Power Spectrum Density	ความหนาแน่นของสเปกตรัมกำลัง
Pre-filter	พรีฟิวเตอร์
Projection Order	อันดับของการฉาย
Pseudo-inverse Matrix	เมตริกซ์ผกผันเทียม
Psychoacoustic	จิตวิทยาทางเสียง
Quadrature Mirror Filter Bank	คลังวงจรรองควอดเรเจอร์-มิลเลอร์
Real Additions and Subtractions	การบวกและลบจำนวนจริง
Real Multiplications	การคูณจำนวนจริง
Real-time	เวลาจริง
Receiving Room	ห้องรับ
Regularization Matrix	เมตริกซ์เรกกูลาไรเซชัน
Reverberation Time	เวลาสะท้อนกลับ
Ripple	ระลอก
Scaling	การสเกล

Sideband	แถบความถี่ข้าง
Similarity Transformation	การแปลงแบบคล้าย
Slowly Time-varying	การแปลงตามเวลาแบบช้า
Steady-state	สถานะคงตัว
Step-size	ค่าช่วงก้าว
Stereophonic Acoustic Echo Cancellation	ระบบการตัดสัญญาณเสียงสะท้อนในการ สัมมนาแบบสเตอริโอ
Subband Filtering Structure	โครงสร้างการกรองแถบย่อย
Synthesis Filter Bank	คลังวงจรรองสังเคราะห์
Tap-weight Vector	เวกเตอร์สัมประสิทธิ์น้ำหนัก
Teleconferencing System	ระบบสัมมนาทางไกล
Tradeoff	ภาวะถ่วงดุล, การชดเชยข้อดีข้อเสีย
Transfer Function	ฟังก์ชันถ่ายโอน
Transversal Filter	วงจรรองแบบเส้นตัดขวาง
Transmitting Room	ห้องส่ง
Under-determined	ไม่สามารถกำหนดได้
Uniform Filter Bank	คลังวงจรรองเอกรูป
Unitary	ยูนิแทรี
Unit Circle	วงกลมหนึ่งหน่วย
Update Equation	สมการปรับให้เป็นปัจจุบัน
Upsampling	การเพิ่มการซีกตัวอย่างข้อมูล
Variance	ความแปรปรวน
Weight-error Vector	เวกเตอร์สัมประสิทธิ์น้ำหนักผิดพลาด
Weighting Factor	ตัวประกอบถ่วงน้ำหนัก
Z-transformation	การแปลง Z

บทที่ 1

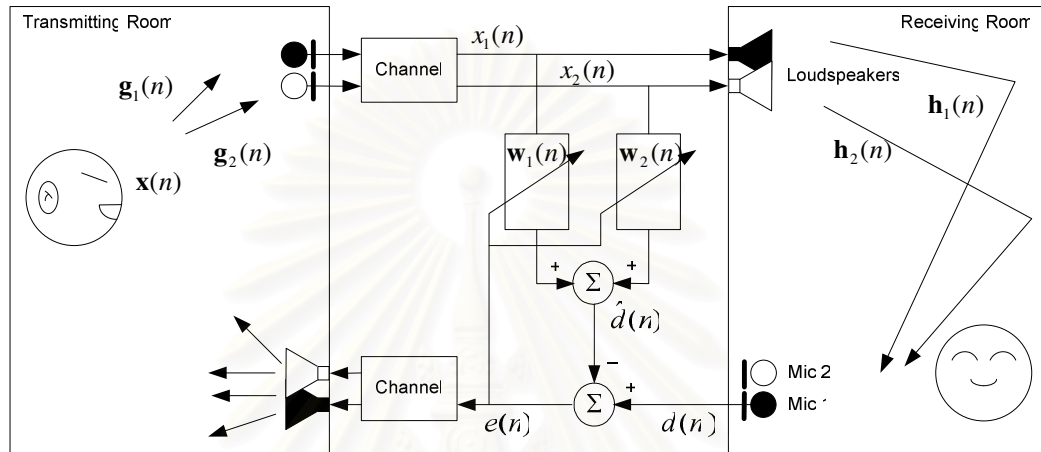
บทนำ

1.1 ความสำคัญของปัญหา

สัญญาณเสียงสะท้อน (Acoustic Echo) เป็นปัญหาที่พบในระบบการสื่อสารทางเสียงหลายประเภท เช่น การสื่อสารทางโทรศัพท์เคลื่อนที่แบบที่ใช้อุปกรณ์ Hand-free หรือในระบบสัมมนาทางไกล (Teleconferencing System) เป็นต้น ซึ่งเกิดจากวิถีสะท้อนทางเสียง (Acoustic Echo Paths, AEP) ระหว่างลำโพงกับไมโครโฟน [1] สำหรับระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอ (Stereophonic Acoustic Echo Cancellation, SAEC) จะประยุกต์ใช้วงจรกรองแบบปรับตัวชนิดผลตอบสนองอิมพัลส์แบบจำกัด (Finite Impulse Response, FIR) ในการจำลองวิถีสะท้อนทางเสียง [2] ในปัจจุบัน การสัมมนาแบบสเตอริโอซึ่งเป็นการสื่อสารทางเสียงที่มีสองช่องสัญญาณหรือมากกว่าสองช่องสัญญาณ จะได้รับความนิยมมากกว่าการสัมมนาทางเสียงแบบช่องสัญญาณเดียว เนื่องจากสามารถรองรับการใช้งานของผู้ร่วมสัมมนาในห้องสัมมนาได้อย่างทั่วถึง เมื่อเปรียบเทียบกับการสัมมนาแบบช่องสัญญาณเดียวที่ผู้พูดจะต้องพูดผ่านไมโครโฟนเพียงตัวเดียว อีกทั้ง การจัดวางตำแหน่งของไมโครโฟนที่ห่างกันในห้องส่ง จะทำให้การรับรู้ทางเสียงจากลำโพงในห้องรับสมจริงมากกว่าการสัมมนาทางเสียงแบบช่องสัญญาณเดียว

เมื่อพิจารณาการสัมมนาแบบสเตอริโอซึ่งประกอบด้วยลำโพงและไมโครโฟนทั้งทางด้านห้องส่ง (Transmitting Room) และห้องรับ (Receiving Room) ฝั่งละ 2 ชุด ดังแสดงในรูปที่ 1.1 ผู้พูดในห้องส่งทำการพูด สัญญาณเสียงพูด $x(n)$ จะถูกส่งผ่านช่องสัญญาณไปยังห้องรับ และเกิดการเชื่อมต่อ (Coupling) ระหว่างสัญญาณเสียงที่ถูกขยายขนาดจากไมโครโฟนและลำโพงที่อยู่ใกล้กันในห้องรับ เรียกว่าวิถีสะท้อนทางเสียง ซึ่งเป็นผลจากการสะท้อนของสัญญาณเสียงพูดกับบริเวณโดยรอบ ส่งผลให้ผู้พูดในห้องส่งได้ยินเสียงของตนเองและรบกวนการสนทนา จึงมีการนำวงจรกรองแบบปรับตัว $w(n)$ มาประยุกต์ใช้ในระบบการตัดสัญญาณเสียงสะท้อนที่ไม่ต้องการระหว่างการสนทนา โดยวงจรกรองแบบปรับตัวจะทำการประมาณวิถีสะท้อนทางเสียงให้ได้ใกล้เคียงมากที่สุด เมื่อพิจารณาการตัดสัญญาณเสียงสะท้อนที่เกิดจากวิถีสะท้อนทางเสียงระหว่างลำโพงทั้ง 2 ตัวและไมโครโฟนตัวที่ 1 (Mic 1) ผลต่างระหว่างสัญญาณเสียงสะท้อนจากไมโครโฟนตัวที่ 1 ในห้องรับ $d(n)$ และสัญญาณประมาณของเสียงสะท้อน $\hat{d}(n)$ ที่เป็นสัญญาณออกของวงจรกรองแบบปรับตัว เรียกว่าสัญญาณผิดพลาด $e(n)$ ซึ่งจะถูกส่งกลับไปยังห้องส่ง ดังนั้นเมื่อสัญญาณผิดพลาดมีค่าเข้าใกล้ศูนย์ สัญญาณเสียงสะท้อนจะสามารถถูกตัด

นอกจากการสัมมนาแบบสเตอริโอได้ ทั้งนี้ การทำงานดังกล่าวสามารถนำไปใช้งานในการแก้ปัญหาสัญญาณเสียงสะท้อนที่เกิดจากวิถีสะท้อนทางเสียงระหว่างลำโพงทั้งสองกับไมโครโฟนอีกตัวหนึ่งได้ด้วย (Mic 2) สำหรับวิทยานิพนธ์ฉบับนี้จะพิจารณาเฉพาะการตัดสัญญาณเสียงสะท้อนที่ไมโครโฟนเพียงตัวเดียวเท่านั้น



รูปที่ 1.1 แบบจำลองระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอ (SAEC)

สำหรับระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอ การเลือกจำนวนสัมประสิทธิ์ของวงจรรองแบบปรับตัว (หรือเรียกว่าความยาวของวงจรรองแบบปรับตัว) ในการประมาณวิถีสะท้อนทางเสียงจะพิจารณาให้มีความเหมาะสม ถ้าจำนวนสัมประสิทธิ์ของวงจรรองแบบปรับตัวที่เลือกใช้มีค่าน้อยเกินไป จะทำให้เกิดความผิดพลาดอย่างมากในการประมาณวิถีสะท้อนทางเสียง แต่ถ้าเลือกจำนวนสัมประสิทธิ์ของวงจรรองแบบปรับตัวมากเกินไป ความซับซ้อนทางการคำนวณของระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอจะมากขึ้นตามไปด้วย ส่งผลให้อัตราการลู่เข้าของวงจรรองแบบปรับตัวในระบบการตัดสัญญาณเสียงสะท้อนช้าลง ซึ่งจะไม่เหมาะสมกับการประยุกต์ใช้กับงานเวลาจริง (Real-time) ทั้งนี้จำนวนสัมประสิทธิ์ของวงจรรองแบบปรับตัวจะขึ้นอยู่กับการประมาณค่าของเวลาสะท้อนกลับ (Reverberation Time) ของสัญญาณเสียงในแต่ละสภาพแวดล้อม เช่น เวลาสะท้อนกลับของสภาพแวดล้อมภายในรถยนต์จะอยู่ที่ 50-100 ms หรือสำหรับสภาพแวดล้อมในการสัมมนาแบบสเตอริโอที่เป็นสำนักงานจะมีเวลาสะท้อนกลับที่มากกว่า คือ ประมาณ 300-500 ms [3] ดังนั้นจำนวนสัมประสิทธิ์ของวงจรรองแบบปรับตัวควรมีค่าประมาณ 800-1,600 ค่าในภายในรถยนต์ และ 4,800-8,000 ค่าในกรณีภายในสำนักงาน เมื่อความถี่ซีกตัวอย่างของระบบเท่ากับ 16 kHz เป็นต้น จะเห็นว่า จำนวนสัมประสิทธิ์ของวงจรรองแบบปรับตัวที่เหมาะสมสำหรับระบบการตัด

สัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอต้องมีจำนวนมาก ซึ่งทำให้ความซับซ้อนทางการคำนวณของระบบการตัดสัญญาณเสียงสะท้อนสูงขึ้น ดังนั้นวิทยานิพนธ์ฉบับนี้จะศึกษาและพัฒนากระบวนการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอ เพื่อปรับปรุงสมรรถนะของการตัดสัญญาณเสียงสะท้อน โดยลดความซับซ้อนทางการคำนวณ และเพิ่มอัตราการสุ่มเข้าของวงจรกรองแบบปรับตัวที่ใช้กับระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอ

วิธีที่ได้รับความนิยมอย่างแพร่หลายในการลดความซับซ้อนทางการคำนวณสำหรับระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอ คือ การประยุกต์ใช้โครงสร้างการกรองแถบย่อย (Subband Filtering Structure) ร่วมกับระบบการตัดสัญญาณเสียงสะท้อนแบบปกติ [4] - [6] โครงสร้างการกรองแถบย่อยเป็นการแบ่งข้อมูลของสัญญาณที่ต้องการประมวลผลออกเป็นความถี่ย่อยหลายแถบความถี่ เพื่อให้สามารถประมวลผลสัญญาณในแต่ละแถบความถี่ย่อยพร้อมกันได้ เรียกว่าการประมวลผลแบบขนาน (Parallel Processing) เมื่อใช้กระบวนการเดซิเมชัน (Decimation Process) ในระบบที่มีโครงสร้างการกรองแถบย่อย เพื่อลดอัตราการสุ่มเข้าตัวอย่างของระบบ จะทำให้จำนวนสัมประสิทธิ์ของวงจรกรองแบบปรับตัวในแต่ละแถบความถี่ย่อยมีค่าลดลงไปจากระบบเดิมที่ไม่ใช้โครงสร้างการกรองแถบย่อย จึงสามารถลดความซับซ้อนทางการคำนวณของระบบเดิมลงได้ นอกจากนี้โครงสร้างการกรองแถบย่อยยังเหมาะสมกับการประมวลผลสัญญาณเสียงพูด เนื่องจากคุณลักษณะของสัญญาณเสียงพูดที่มีสเปกตรัมของสัญญาณไม่เรียบ (Non-flat Spectrum) ดังนั้น เมื่อแบ่งสเปกตรัมข้อมูลของสัญญาณแบ่งออกเป็นหลายแถบความถี่ย่อย จะทำให้ค่าการกระจายของลักษณะเฉพาะ (Eigenvalue Spread) ของเมทริกซ์อัตโนมัติสหสัมพันธ์ (Autocorrelation Matrix) ของสัญญาณเสียงพูดในแต่ละแถบความถี่ย่อยมีการกระจายตัวน้อยกว่าค่าดังกล่าวที่มีโครงสร้างแถบเต็ม (Fullband Structure) ตามปกติ ทำให้วงจรกรองแบบปรับตัวในแต่ละแถบความถี่ย่อยมีอัตราการสุ่มเข้าที่เร็วขึ้น นอกจากนี้ อีกวิธีหนึ่งที่สามารถลดความซับซ้อนทางการคำนวณของระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอได้ คือ การประมวลผลสัญญาณในโดเมนความถี่ [7] - [10] โดยใช้หลักการของการประมวลผลแบบบล็อก (Block Processing) และการแปลงฟูริเยร์ไม่ต่อเนื่อง (Discrete Fourier Transform) เพื่อทำการปรับปรุงสัมประสิทธิ์ของวงจรกรองแบบปรับตัวในโดเมนความถี่ จึงสามารถลดความซับซ้อนทางการคำนวณของระบบการตัดสัญญาณเสียงสะท้อนได้

1.2 วัตถุประสงค์

ศึกษาและพัฒนากระบวนการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอ ซึ่งใช้วงจรกรองแบบปรับตัวในการจำลองวิถีสะท้อนทางเสียง โดยอาศัยโครงสร้างการกรองแถบย่อย และการประมวลผลสัญญาณในโดเมนความถี่ เพื่อปรับปรุงสมรรถนะของระบบการตัด

สัญญาณเสียงสะท้อน และเพื่อลดความซับซ้อนทางการคำนวณ นอกจากนี้ ยังศึกษาถึงความ เป็นได้ในการเพิ่มอัตราการรู้เข้าของวงจรแบบปรับตัวในระบบการตัดสัญญาณเสียงสะท้อน นี้ด้วย

1.3 เป้าหมายและขอบเขตของการวิจัย

วิทยานิพนธ์ฉบับนี้ศึกษาโครงสร้างการกรองแถบย่อย และการประมวลผลสัญญาณใน โดเมนความถี่สำหรับระบบการตัดสัญญาณเสียงสะท้อนในการสम्मมนาแบบสเตอริโอ เพื่อลดความ ซ้ำซ้อนทางการคำนวณเมื่อเปรียบเทียบกับระบบการตัดสัญญาณเสียงสะท้อนแบบปกติที่ทำงาน ในโดเมนเวลา และเพื่อศึกษาความเป็นไปได้ในการปรับปรุงสมรรถนะการทำงานของระบบการตัด สัญญาณเสียงสะท้อนในด้านอัตราการรู้เข้าของวงจรแบบปรับตัวให้เร็วกว่าระบบดังกล่าวที่ ทำงานในโดเมนเวลา นอกจากนี้ความถูกต้องแม่นยำในการจำลองวิธีสะท้อนทางเสียงของวงจร กรองแบบปรับตัวของระบบการตัดสัญญาณเสียงสะท้อนในการสम्मมนาแบบสเตอริโอ จะถูกสังเกต จากค่าผิดพลาดในการปรับแก้ (Misadjustment) ให้มีค่าต่ำกว่าของระบบการตัดสัญญาณเสียง สะท้อนในการสम्मมนาแบบสเตอริโอแบบปกติ

1.4 ประโยชน์ที่คาดว่าจะได้รับ

แนวทางในการพัฒนาการแก้ปัญหาหาระบบการตัดสัญญาณเสียงสะท้อนในการสम्मมนา แบบสเตอริโอที่มีสมรรถนะสูง อัตราการรู้เข้าเร็วและความซับซ้อนทางการคำนวณต่ำ

1.5 ขั้นตอนและวิธีการดำเนินการ

1. ศึกษาถึงที่มาของปัญหาและแนวทางการแก้ปัญหาของระบบการตัดสัญญาณเสียง สะท้อนในการสम्मมนาแบบสเตอริโอ
2. ศึกษาวิธีการตัดสัญญาณเสียงสะท้อน ที่ทำการประมวลผลในโดเมนเวลาและโดเมน ความถี่
3. ประยุกต์ใช้โครงสร้างการกรองแถบย่อยเพื่อปรับปรุงสมรรถนะของระบบการตัด สัญญาณเสียงสะท้อน
4. เขียนโปรแกรมทดสอบการจำลองการทำงานของระบบการตัดสัญญาณเสียงสะท้อนใน การสम्मมนาแบบสเตอริโอในสิ่งแวดล้อมต่างๆ

5. วิเคราะห์และประเมินผลสมรรถนะในการทำงานของระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสตรีโอและหาค่าความซับซ้อนทางการคำนวณของระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสตรีโอที่มีโครงสร้างต่างๆ กัน
6. ศึกษาผลการทำงานของระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสตรีโอที่มีโครงสร้างการกรองแถบย่อย
7. ศึกษาผลการทำงานของระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสตรีโอที่มีโครงสร้างไฮบริด ที่นำเสนอขึ้นในวิทยานิพนธ์นี้เปรียบเทียบกับวิธีต่างๆ ที่ศึกษาในวิทยานิพนธ์นี้
8. สรุปผลงานวิจัยและจัดทำรูปเล่มวิทยานิพนธ์ฉบับสมบูรณ์

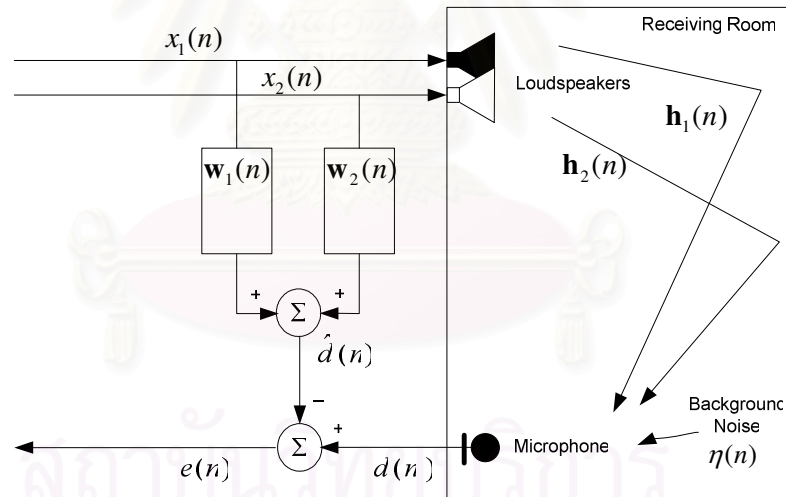


สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

หลักการและขั้นตอนวิธีสำคัญ

ในบทนี้จะกล่าวถึงหลักการทำงานของวงจรกรองแบบปรับตัว (Adaptive Filter) ในระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอ และขั้นตอนวิธีต่างๆ ทั้งในโดเมนเวลาและโดเมนความถี่ที่ใช้กับวงจรกรองแบบปรับตัว โดยขั้นตอนวิธีในโดเมนเวลาที่สนใจศึกษาในวิทยานิพนธ์นี้ ได้แก่ ขั้นตอนวิธี Steepest-descent, LMS, NLMS [2], AP [11], [12] และ FLS [2], [13], [14] ส่วนขั้นตอนวิธีในโดเมนความถี่ที่สนใจศึกษา คือ ขั้นตอนวิธี F-NLMS [2], [15] และยังสามารถศึกษาการประยุกต์ใช้โครงสร้างการกรองแถบย่อย (Subband Filtering Structure) ร่วมกับการประมวลผลสัญญาณทั้งในโดเมนเวลาและโดเมนความถี่ นอกจากนี้ จะทำการเปรียบเทียบความซับซ้อนทางการคำนวณ (Computational Complexity) ของแต่ละขั้นตอนวิธีที่ศึกษาในวิทยานิพนธ์นี้ไว้ด้วย



รูปที่ 2.1 ระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอ

หลักการงานของแต่ละขั้นตอนวิธีที่ใช้สำหรับระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอ คือ การปรับปรุงสัมประสิทธิ์ของวงจรกรองแบบปรับตัวของระบบดังกล่าวให้มีความใกล้เคียงกับวิธีสะท้อนทางเสียงในห้องรับให้มากที่สุด ด้วยอัตราการสุ่มค่าตอบของระบบที่รวดเร็วเพียงพอในการทำให้สัญญาณผิดพลาดที่ถูกส่งกลับไปยังห้องส่งมีค่าน้อยที่สุด โดยที่สัญญาณผิดพลาดเป็นความแตกต่างระหว่างสัญญาณจากไมโครโฟนในห้องรับกับสัญญาณ

ออกที่ได้จากวงจรรองแบบปรับตัว ดังรูปที่ 2.1 เมื่อสัญญาณผิดพลาดมีค่าน้อยที่สุด แสดงว่าสัญญาณเสียงสะท้อนถูกตัดได้อย่างมีประสิทธิภาพ นอกจากนี้ ความซับซ้อนทางการคำนวณของขั้นตอนวิธีที่จะเลือกใช้สำหรับระบบการตัดสัญญาณเสียงสะท้อนนั้นจะต้องมีค่าต่ำ เหมาะสมแก่การนำไปปฏิบัติได้ (Implementation) รายละเอียดการทำงานของแต่ละขั้นตอนวิธีในการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวจะได้กล่าวถึงในหัวข้อย่อยต่อไป

ทั้งนี้ ขั้นตอนวิธีต่างๆ ที่ใช้ในการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวสองช่องสัญญาณ เช่น ขั้นตอนวิธี Steepest-descent สองช่องสัญญาณ จะขอเรียกว่าขั้นตอนวิธี Steepest-descent2 สำหรับขั้นตอนวิธี Steepest-descent ช่องสัญญาณเดียว จะเรียกว่าขั้นตอนวิธี Steepest-descent เป็นต้น

2.1 หลักการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี Steepest-Descent2

เมื่อพิจารณาระบบการตัดสัญญาณเสียงสะท้อนแบบสเตรโที่ที่ใช้วงจรรองแบบปรับตัวที่มีโครงสร้างการทำงานแบบเส้นตัดขวาง (Transversal Filter) ดังแสดงในรูปที่ 2.2 เวกเตอร์สัญญาณเข้าของระบบ ณ เวลา n ที่มีจำนวนสัมประสิทธิ์ L ตัว สามารถจำลองได้เป็น $\mathbf{x}_j(n)$ เมื่อ $j = 1, 2$ ดังนี้

$$\mathbf{x}_j(n) = [x_j(n), x_j(n-1), \dots, x_j(n-L+1)]^T \quad (2.1)$$

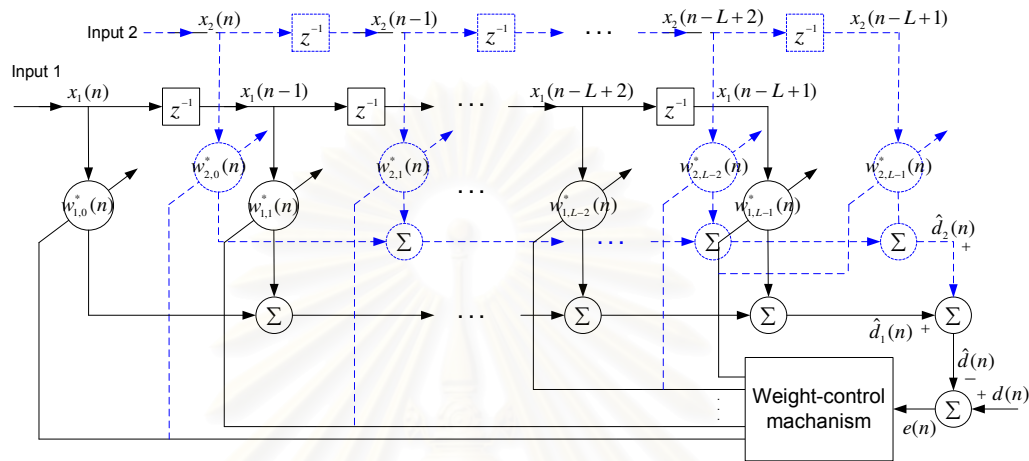
เมื่อ $\langle \cdot \rangle^T$ คือ เมทริกซ์สลับเปลี่ยน และเวกเตอร์สัมประสิทธิ์ของวงจรรองแบบปรับตัว ณ เวลา n ที่มีขนาด L เท่ากันเป็น

$$\mathbf{w}_j(n) = [w_{j,0}(n), w_{j,1}(n), \dots, w_{j,L-1}(n)]^T \quad (2.2)$$

เมื่อ $j = 1, 2$

หลักการสำคัญของขั้นตอนวิธี Steepest Descent2 คือ การพยายามลดค่าคลาดเคลื่อนกำลังสองเฉลี่ยให้มีค่าน้อยที่สุด โดยเริ่มการทำงานจากการกำหนดค่าเริ่มต้นให้แก่สัมประสิทธิ์ของวงจรรองแบบปรับตัวหรือ เวกเตอร์สัมประสิทธิ์น้ำหนัก (Tap-weight Vector) โดยค่าสัมประสิทธิ์น้ำหนักจะถูกปรับปรุงทุกๆ จำนวนรอบหรือจำนวนการวนซ้ำ (Iteration) ของการคำนวณ ค่าของเวกเตอร์สัมประสิทธิ์น้ำหนักที่ได้จากการวนซ้ำในรอบสุดท้ายจะถูกรวบรวมที่เรียกว่า Wiener

Solution [2] ซึ่งเป็นคำตอบของค่าสัมประสิทธิ์ของระบบไม่ทราบค่าที่เหมาะสมที่สุด (Optimal) ซึ่งในที่นี้ คือ ค่าวิถีสะท้อนทางเสียงในห้องรับ โดยที่อัตราการเรียนรู้เข้าสู่คำตอบของสัมประสิทธิ์ของวงจรกรองแบบปรับตัวจะแปรผันตรงกับขนาดของค่าช่วงก้าว (Step-size)



รูปที่ 2.2 โครงสร้างการทำงานของวงจรกรองแบบปรับตัวด้วยขั้นตอนวิธี Steepest-Descent2

จากรูปที่ 2.2 ระบบการตัดสัญญาณเสียงสะท้อนแบบสเตริโอต้องการประมาณสัญญาณออกของวงจรกรองแบบปรับตัว $\hat{d}(n)$ ให้มีความใกล้เคียงกับสัญญาณไมโครโฟนในห้องรับ $d(n)$ ให้มากที่สุด ดังนั้นสัญญาณความผิดพลาดของการประมาณ $e(n)$ หาได้จากความแตกต่างระหว่างสัญญาณออกของวงจรกรองแบบปรับตัว $\hat{d}(n)$ กับสัญญาณไมโครโฟน $d(n)$

$$e(n) = d(n) - \hat{d}(n) \quad (2.3)$$

$$= d(n) - \mathbf{w}_1^H(n)\mathbf{x}_1(n) - \mathbf{w}_2^H(n)\mathbf{x}_2(n) \quad (2.4)$$

โดยที่ $\mathbf{w}_j^H(n)\mathbf{x}_j(n)$ เป็นผลบวกการประสาน (Convolution Sum) ระหว่างเวกเตอร์สัมประสิทธิ์ของวงจรกรองแบบปรับตัวกับเวกเตอร์สัญญาณเข้าของแต่ละช่องสัญญาณ

ถ้าเวกเตอร์สัญญาณเข้า $\mathbf{x}_j(n)$ มีความสัมพันธ์แบบสเตชันนารีร่วม (Jointly Stationary) กับสัญญาณไมโครโฟน $d(n)$ ดังนั้นค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error) หรือฟังก์ชันต้นทุน (Cost Function) จะเป็นดังสมการที่ (2.5)

$$\mathbf{J}(n) = E\{e^2(n)\} \quad (2.5)$$

เมื่อ $E\{\cdot\}$ เป็นค่าคาดหวัง (Expectation) ของสัญญาณผิดพลาด กำหนดให้ σ_d^2 เป็นความแปรปรวน (Variance) ของสัญญาณไมโครโฟน $d(n)$ เวกเตอร์ \mathbf{p}_j เป็นเวกเตอร์สหสัมพันธ์ไขว้ (Cross-correlation Vector) ระหว่างเวกเตอร์สัญญาณเข้า $\mathbf{x}_j(n)$ กับสัญญาณไมโครโฟน $d(n)$

$$\mathbf{p}_j = E\{\mathbf{x}_j(n)d^*(n)\} \quad (2.6)$$

และเมทริกซ์ \mathbf{R}_j เป็นเมทริกซ์อัตสหสัมพันธ์ (Autocorrelation Matrix) ของเวกเตอร์สัญญาณเข้า $\mathbf{x}_j(n)$ แต่ละตัวเป็น

$$\mathbf{R}_j = E\{\mathbf{x}_j(n)\mathbf{x}_j^H(n)\} \quad (2.7)$$

แทนค่า $e(n)$ จากสมการที่ (2.4) ลงในสมการ (2.5) จะได้ฟังก์ชันต้นทุนเป็น

$$\mathbf{J}(n) = \sigma_d^2 - \sum_{j=1}^2 (\mathbf{w}_j^H(n)\mathbf{p}_j + \mathbf{p}_j^H\mathbf{w}_j(n) - \mathbf{w}_j^H(n)\mathbf{R}_j\mathbf{w}_j(n)) \quad (2.8)$$

ในการหาค่า $\mathbf{w}_j(n)$ ที่เหมาะสมที่สุดหรือเมื่อ $\mathbf{J}(n)$ มีค่าน้อยที่สุด (Minimum) สามารถทำได้โดยการหาค่าเวกเตอร์เกรเดียนท์ของ $\mathbf{J}(n)$ ซึ่งจะได้เป็น

$$\nabla(\mathbf{J}(n)) = \sum_{j=1}^2 (-2\mathbf{p}_j + 2\mathbf{R}_j\mathbf{w}_j(n)) \quad (2.9)$$

สมการปรับปรุงค่าสัมประสิทธิ์น้ำหนักของวงจรกรองแบบปรับตัว $\mathbf{w}_j(n)$ (Update Equation) หาได้จากการคำนวณแบบวนซ้ำดังต่อไปนี้

$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + \frac{1}{2}\mu[-\nabla_j(\mathbf{J}(n))] \quad (2.10)$$

เมื่อ $\nabla_j(\mathbf{J}(n)) = -2\mathbf{p}_j + 2\mathbf{R}_j\mathbf{w}_j(n)$ และ μ เป็นค่าบวกขนาดเล็กที่ใช้ทำการสเกล (Scaling) การปรับปรุงค่าสัมประสิทธิ์ของ $\mathbf{w}_j(n)$ หรือจะเรียกว่า ค่าช่วงก้าว (Step-size) เมื่อแทนค่า $\nabla_j(\mathbf{J}(n))$ ในสมการที่ (2.10) จะได้ว่า

$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + \mu[\mathbf{p}_j - \mathbf{R}_j\mathbf{w}_j(n)] \quad (2.11)$$

ขั้นตอนวิธี Steepest-Descent2 ใช้การป้อนกลับในการควบคุมการทำงานของวงจรรองแบบปรับตัว ดังนั้น เสถียรภาพของระบบที่ใช้ขั้นตอนวิธีนี้จะขึ้นอยู่กับพารามิเตอร์สองตัวคือ ค่าช่วงก้าว μ และเมทริกซ์อัตราสลับพันธุ \mathbf{R} [2] ดังนั้น จึงต้องหาเงื่อนไขของค่าช่วงก้าวและเมทริกซ์อัตราสลับพันธุที่ทำให้ระบบเสถียร ด้วยการนิยามเวกเตอร์สัมประสิทธิ์น้ำหนักผิดพลาด (Weight-error Vector) ที่เวลา n ดังต่อไปนี้

$$\mathbf{c}_j(n) = \mathbf{w}_j(n) - \mathbf{w}_{j,o} \quad (2.12)$$

โดย $\mathbf{w}_{j,o}$ เป็นค่าที่เหมาะสมที่สุดของเวกเตอร์สัมประสิทธิ์ของวงจรรองแต่ละตัว ซึ่งนิยามมาจาก สมการ Wiener-Hopf [2] โดยมีความสัมพันธ์กับเมทริกซ์ \mathbf{R}_j และเวกเตอร์ \mathbf{p}_j คือ

$$\mathbf{R}_j \mathbf{w}_{j,o} = \mathbf{p}_j \quad (2.13)$$

เมื่อแทนค่าเวกเตอร์ \mathbf{p}_j จากสมการที่ (2.13) ลงในสมการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวแต่ละช่องสัญญาณลงในสมการที่ (2.11) จะสามารถหาค่าเวกเตอร์สัมประสิทธิ์น้ำหนักผิดพลาดได้ใหม่เป็น

$$\mathbf{c}_j(n+1) = (\mathbf{I} - \mu \mathbf{R}_j) \mathbf{c}_j(n) \quad (2.14)$$

โดย \mathbf{I} เป็นเมทริกซ์เอกลักษณ์ (Identity Matrix) เมื่อใช้การแปลงแบบคล้าย (Similarity Transformation) ที่มีคุณสมบัติยูนิแทรี (Unitary) จะสามารถเขียนเมทริกซ์อัตราสลับพันธุ \mathbf{R}_j ได้ใหม่ในรูปของ

$$\mathbf{R}_j = \mathbf{Q}_j \mathbf{\Lambda}_j \mathbf{Q}_j^H \quad (2.15)$$

โดยเมทริกซ์ \mathbf{Q}_j ถูกเรียกว่าเมทริกซ์ยูนิแทรี (Unitary Matrix) มีค่าในแนวคอลัมน์เป็นเซตตั้งฉาก (Orthogonal Set) ของเวกเตอร์ลักษณะเฉพาะ (Eigenvectors) จากค่าลักษณะเฉพาะ (Eigenvalues) จำนวน L ตัวของเมทริกซ์ \mathbf{R}_j นิยามด้วย $\lambda_{j,1}, \lambda_{j,2}, \dots, \lambda_{j,L}$ และเมทริกซ์ $\mathbf{\Lambda}_j$ คือเมทริกซ์ทแยงมุม (Diagonal Matrix) ที่มีสมาชิกในแนวทแยงมุมเป็นค่าลักษณะเฉพาะตั้งแต่ $\lambda_{j,1}, \lambda_{j,2}, \dots, \lambda_{j,L}$ ตามลำดับ ดังนั้น สมการที่ (2.14) จึงเขียนใหม่ได้ดังนี้

$$\mathbf{c}_j(n+1) = (\mathbf{I} - \mu \mathbf{Q}_j \Lambda_j \mathbf{Q}_j^H) \mathbf{c}_j(n) \quad (2.16)$$

จากคุณสมบัติของเมทริกซ์ยูนิแทรี \mathbf{Q} ที่ \mathbf{Q}^H เท่ากับ \mathbf{Q}^{-1} จึงสามารถคูณ \mathbf{Q}^H ทั้งสองข้างของสมการ (2.16) ได้เป็น

$$\mathbf{Q}_j^H \mathbf{c}_j(n+1) = (\mathbf{I} - \mu \Lambda_j) \mathbf{Q}_j^H \mathbf{c}_j(n) \quad (2.17)$$

โดยกำหนดให้ $\mathbf{v}_j(n) = \mathbf{Q}_j^H \mathbf{c}_j(n)$ ดังนั้น $\mathbf{v}_j(n) = \mathbf{Q}_j^H [\mathbf{w}_j(n) - \mathbf{w}_{j,o}]$ ด้วย ซึ่งสามารถหาความสัมพันธ์ของ $\mathbf{v}_j(n)$ และ $\mathbf{v}_j(n+1)$ ได้เป็น

$$\mathbf{v}_j(n+1) = (\mathbf{I} - \mu \Lambda_j) \mathbf{v}_j(n) \quad (2.18)$$

สำหรับค่าลักษณะเฉพาะตัวที่ k ($\lambda_{j,k}$) ใดๆของแต่ละช่องสัญญาณของขั้นตอนวิธี Steepest-Descent สามารถเขียนสมการที่ (2.18) ได้เป็น

$$v_{j,k}(n+1) = (1 - \mu \lambda_{j,k}) v_{j,k}(n) ; k = 1, 2, \dots, L \quad (2.19)$$

จะเห็นว่าสมการที่ (2.19) เป็นสมการผลต่างเอกพันธ์ (Homogeneous Difference Equation) อันดับหนึ่ง โดยกำหนดให้ $v_{j,k}(n)$ มีค่าเริ่มต้นเป็น $v_{j,k}(0)$ จะได้ว่า

$$v_{j,k}(n) = (1 - \mu \lambda_{j,k})^n v_{j,k}(0) ; k = 1, 2, \dots, L \quad (2.20)$$

เนื่องจากค่าลักษณะเฉพาะของเมทริกซ์อัตราส่วนสัมพันธ์ \mathbf{R}_j เป็นจำนวนจริงบวก ผลตอบ $v_{j,k}(n)$ ของสมการที่ (2.20) จะไม่ลู่ออก (Diverge) ก็ต่อเมื่อ

$$-1 < 1 - \mu \lambda_{j,k} < 1 ; \forall k \quad (2.21)$$

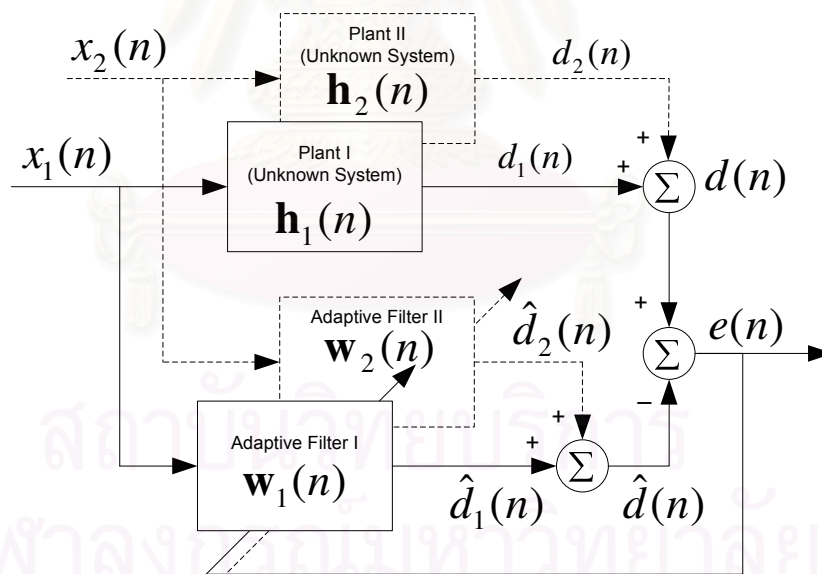
ดังนั้น เพื่อรักษาความเสถียรภาพของระบบ ค่าช่วงก้าว μ ควรอยู่ในช่วง

$$0 < \mu < \frac{2}{\lambda_{j,\max}} \quad (2.22)$$

เมื่อ $\lambda_{j,\max}$ เป็นค่าลักษณะเฉพาะของเมทริกซ์อัตราสหสัมพันธ์ \mathbf{R}_j ที่มีค่ามากที่สุด ดังนั้น ค่าช่วงก้าว μ และค่าลักษณะเฉพาะ $\lambda_{j,\max}$ ของเมทริกซ์อัตราสหสัมพันธ์ \mathbf{R}_j จะมีผลโดยตรงต่ออัตราการลู่เข้าสู่สภาวะอยู่ตัว (Convergence Rate) ของวงจรรองแบบปรับตัว ทั้งนี้อัตราการลู่เข้าของวงจรรองแบบปรับตัวจะแปรผันตรงกับค่าช่วงก้าวที่เลือกใช้ ถ้าเลือกใช้ค่าช่วงก้าวที่มีค่ามากจะทำให้วงจรรองแบบปรับตัวลู่เข้าสู่ค่าตอบได้เร็ว หากแต่ค่าช่วงก้าวแปรผกผันกับค่าผิดพลาดในการปรับแก้ (Misadjustment) กล่าวคือ การเลือกใช้ค่าช่วงก้าวที่มีค่ามากจะทำให้ค่าผิดพลาดในการปรับแก้มีค่าสูงด้วย ดังนั้น การเลือกใช้ค่าช่วงก้าวที่เหมาะสม จึงต้องมีการชดเชยข้อดีข้อเสีย (Trade-off) ระหว่างอัตราการลู่เข้า และค่าผิดพลาดในการปรับแก้

2.2 หลักการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี Least Mean Square2

การทำงานของวงจรรองแบบปรับตัวสำหรับระบบการตัดสัญญาณเสียงสะท้อนแบบสเตอริโอในระบบการสื่อสารทางเสียงที่มีสองช่องสัญญาณ เมื่อมีการควบคุมด้วยขั้นตอนวิธี Least Mean Square2 (LMS2) ที่ประมวลผลในโดเมนเวลา สามารถแสดงโดยใช้บล็อกไดอะแกรมดังรูปที่ 2.3



รูปที่ 2.3 โครงสร้างการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี LMS2 ที่ประมวลผลในโดเมนเวลา

ขั้นตอนวิธี LMS2 เป็นขั้นตอนวิธีที่มีความซับซ้อนทางการคำนวณที่ต่ำ ถึงแม้ว่าจะมีอัตราการลู่เข้าที่ค่อนข้างช้า เมื่อสัญญาณเข้าของระบบการตัดสัญญาณเสียงสะท้อนแบบสเตอริโอเป็น

สัญญาณเสียงพูด ทั้งนี้ขั้นตอนวิธี LMS2 ได้รับการพัฒนามาจากขั้นตอนวิธี Steepest Descent2 แต่ใช้ค่าประมาณของเวกเตอร์เกรเดียนท์ $\hat{\nabla}(\mathbf{J}(n)) = \sum_{j=1}^2 (-2\hat{\mathbf{p}}_j + 2\hat{\mathbf{R}}_j \mathbf{w}_j(n))$ แทน กล่าวคือ การเลือกใช้ค่า ณ ขณะใดขณะหนึ่ง (Instantaneous Value) แทนการคำนวณหาเมทริกซ์อัตรา สหสัมพันธ์ \mathbf{R}_j และการคำนวณหาเวกเตอร์สหสัมพันธ์ไขว้ \mathbf{p}_j ดังนั้น $\hat{\mathbf{R}}_j = \mathbf{x}_j(n)\mathbf{x}_j^H(n)$ และ $\hat{\mathbf{p}}_j = \mathbf{x}_j(n)d^*(n)$ จึงทำให้สมการปรับปรุงค่าสัมประสิทธิ์น้ำหนักของวงจรรองแบบปรับตัว $\mathbf{w}_j(n)$ ในสมการที่ (2.10) ของขั้นตอนวิธี Steepest-descent2 กลายเป็น

$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + \mu \mathbf{x}_j(n)e(n) \quad (2.23)$$

เมื่อสัญญาณผิดพลาดหาได้จาก

$$e(n) = d(n) - \hat{d}(n) \quad (2.24)$$

และสัญญาณออกของวงจรรองแบบปรับตัวหาได้จาก

$$\hat{d}(n) = \mathbf{w}_1^H(n)\mathbf{x}_1(n) + \mathbf{w}_2^H(n)\mathbf{x}_2(n) \quad (2.25)$$

ทั้งนี้ เมื่อเวกเตอร์สัญญาณเข้า $\mathbf{x}_j(n)$ และเวกเตอร์สัมประสิทธิ์ของวงจรรองแบบปรับตัว $\mathbf{w}_j(n)$ แต่ละตัวมีขนาดยาว L เท่ากัน

2.3 หลักการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี Normalized Least Mean Square2

ค่าประมาณของเวกเตอร์เกรเดียนท์ $\mu \mathbf{x}_j(n)e(n)$ เมื่อ $j = 1, 2$ ในสมการปรับปรุง สัมประสิทธิ์ของวงจรรองแบบปรับตัวสองช่องสัญญาณด้วยขั้นตอนวิธี LMS2 จะเป็นสัดส่วน โดยตรงกับเวกเตอร์สัญญาณเสียงเข้า $\mathbf{x}_j(n)$ ของระบบการตัดสัญญาณเสียงสะท้อนแบบ สเตอริโอ เมื่อสัญญาณเสียงเข้ามีขนาดใหญ่มากจะทำให้ค่าสัญญาณผิดพลาด $e(n)$ ในการ ประมาณค่าเวกเตอร์เกรเดียนท์มีค่าเพิ่มขึ้นด้วย สำหรับการปรับปรุงค่าสัมประสิทธิ์ของวงจรรอง ด้วยขั้นตอนวิธี Normalized Least Mean Square2 (NLMS2) นั้น จัดการกับปัญหานี้โดยการทำ นอร์มัลไลซ์ค่าประมาณเวกเตอร์เกรเดียนท์ด้วยขนาดของเวกเตอร์สัญญาณเข้า $\|\mathbf{x}_j(n)\|^2$ แต่ละ

ตัวเมื่ $\|\cdot\|$ คือ Euclidean Norm ดังนั้น สมการปรับปรุงค่าสัมประสิทธิ์ของวงจรรองแบบปรับตัว $\mathbf{w}_j(n)$ ด้วยขั้นตอนวิธี NLMS หาได้จาก

$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + \frac{\mu \mathbf{x}_j(n)e(n)}{\varepsilon + \|\mathbf{x}_j(n)\|^2} \quad (2.26)$$

เมื่อ ε คือ ค่าคงที่เล็ก ๆ เพื่อป้องกันการลู่ออกของวงจรรองแบบปรับตัวกรณีที่มีสัญญาณเข้ามีขนาดน้อย [2]

การเลือกใช้ค่าช่วงก้าว μ ที่น้อยเกินไปจะทำให้ค่าผิดพลาดในการปรับแก้ลดลง แต่ส่งผลให้อัตราการลู่เข้าของวงจรรองแบบปรับตัวช้าลงด้วย ในทางตรงกันข้าม การเลือกใช้ค่าช่วงก้าวที่มากเกินไปจะทำให้อัตราการลู่เข้าเร็วแต่ค่าผิดพลาดในการปรับแก้จะมากขึ้น ในกรณีของขั้นตอนวิธี NLMS2 ค่าช่วงก้าวที่ใช้ในการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวจะเปลี่ยนแปลงตามเวลา โดยเขียนได้เป็น

$$\mu_j(n) = \frac{\mu}{\varepsilon + \|\mathbf{x}_j(n)\|^2} \quad (2.27)$$

ซึ่งแตกต่างจากการใช้ค่าช่วงก้าวคงที่ในขั้นตอนวิธี LMS2 ทั้งนี้ ค่า $\mu_j(n)$ ที่เปลี่ยนแปลงตามเวลาในขั้นตอนวิธี NLMS2 จะทำให้ชดเชยข้อดีข้อเสียระหว่างอัตราการลู่เข้าและค่าผิดพลาดในการปรับแก้ จึงทำให้ระบบการตัดสัญญาณเสียงสะท้อนที่ใช้ขั้นตอนวิธี NLMS2 จะมีอัตราการลู่เข้าที่เหมาะสมกว่าขั้นตอนวิธี LMS2 เมื่อสัญญาณเข้าเป็นสัญญาณเสียงพูด ดังนั้น ในวิทยานิพนธ์ฉบับนี้จึงเลือกใช้ขั้นตอนวิธี NLMS2 ในการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัว เป็นขั้นตอนวิธีมาตรฐานสำหรับเปรียบเทียบกับขั้นตอนวิธีอื่นๆ ต่อไป

2.4 หลักการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี Fast Normalized Least Mean Square2

การปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวตามขั้นตอนวิธี NLMS2 เพื่อตัดสัญญาณเสียงสะท้อน มีการนำเสนอทั้งในโดเมนเวลาและโดเมนความถี่อย่างต่อเนื่อง [3], [6] – [10] สำหรับการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวในโดเมนเวลาในการใช้งานของระบบการตัดสัญญาณเสียงสะท้อนแบบช่องสัญญาณเดี่ยวและระบบการตัดสัญญาณเสียงสะท้อนแบบสเตอริโอ นั้น ระบบต้องการจำนวนสัมประสิทธิ์ของวงจรรองจำนวนมากในการ

ประมาณวิธีสะท้อนทางเสียงให้มีความถูกต้องและมีประสิทธิภาพ จึงทำให้การประมวลผลของ วงจรกรองแบบปรับตัวในโดเมนเวลาที่มีความซับซ้อนทางการคำนวณที่สูง ถึงแม้ว่าจะเลือกใช้ ขั้นตอนวิธีที่มีความซับซ้อนทางการคำนวณที่ต่ำ เช่น ขั้นตอนวิธี NLMS2 เป็นต้น การลดความ ซับซ้อนทางการคำนวณนี้สามารถทำได้โดยการปรับปรุงสัมประสิทธิ์ของวงจรกรองแบบปรับตัวใน โดเมนความถี่ [7]-[10] เช่น ขั้นตอนวิธี Block Least Mean Square2 (BLMS2) [7] ที่ประมวลผล แบบบล็อก และขั้นตอนวิธี Fast Normalized Least Mean Square2 (F-NLMS2) [2] เป็นต้น

เมื่อกำหนดให้ $\mathbf{w}_j(k)$ เป็นเวกเตอร์สัมประสิทธิ์ของวงจรกรองแบบปรับตัวความยาว L ของช่องสัญญาณ $j = 1, 2$ ณ บล็อกที่ k จะได้ว่า

$$\mathbf{w}_j(k) = [w_{j,0}(k), w_{j,1}(k), \dots, w_{j,L-1}(k)]^T, \quad k = 0, 1, \dots \quad (2.28)$$

สำหรับเวลา n ของการคำนวณในโดเมนเวลาตามปกติและบล็อกที่ k หรือรอบการคำนวณแบบ บล็อกมีความสัมพันธ์กัน ดังนี้

$$n = kL + r, \quad r = 0, 1, \dots, L - 1 \quad (2.29)$$

เมื่อการคำนวณแบบบล็อกนี้ ทำการแบ่งข้อมูลของสัญญาณออกเป็นบล็อกละ L เพื่อทำการ ประมวลผลในครั้งหนึ่งๆ ดังนั้น เวกเตอร์สัญญาณเข้าแบบสเตรโไอของแต่ละช่องสัญญาณ ณ เวลา n เป็น

$$\mathbf{x}_j(n) = [x_j(n), x_j(n-1), \dots, x_j(n-L+1)]^T \quad (2.30)$$

ซึ่งมีความยาวเช่นเดียวกับความยาวของจำนวนสัมประสิทธิ์ของวงจรกรองแบบปรับตัว สัญญาณออกของวงจรกรองแบบปรับตัวจากแต่ละช่องสัญญาณ ณ เวลา n หาได้จาก

$$\hat{d}_j(n) = \mathbf{w}_j^T(k) \mathbf{x}_j(n) \quad (2.31)$$

ซึ่งสามารถเขียนได้เป็น

$$\hat{d}_j(kL + r) = \mathbf{w}_j^T(k) \mathbf{x}_j(kL + r), \quad r = 0, 1, \dots, L - 1 \quad (2.32)$$

เมื่อกำหนดให้สัญญาณเข้าของไมโครโฟนในห้องรับเป็น $d(n)$ เวกเตอร์สัญญาณผิดพลาดจึงหาได้จาก

$$e(kL+r) = d(kL+r) - \sum_{j=1}^2 \hat{d}_j(kL+r), \quad r = 0,1,\dots,L-1 \quad (2.33)$$

อย่างไรก็ตาม จำนวนสัมประสิทธิ์ของวงจรรองแบบปรับตัวจะถูกจำกัดให้มีค่าคงที่ตลอดการประมวลผลในบล็อกหนึ่งๆ ซึ่งเป็นส่วนที่มีความแตกต่างอย่างชัดเจนเมื่อเปรียบเทียบกับการประมวลผลในโดเมนเวลาตามปกติ ที่สัมประสิทธิ์ของวงจรรองแบบปรับตัวจะถูกปรับปรุงค่าทุกๆ ครั้งที่มีข้อมูลใหม่ของสัญญาณเข้า ดังนั้น สมการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวเมื่อใช้ขั้นตอนวิธี LMS2 และทำการประมวลผลแบบบล็อก หาได้จาก

$$\mathbf{w}_j(k+1) = \mathbf{w}_j(k) + \mu \sum_{r=0}^{L-1} \mathbf{x}_j^H(kL+r)e(kL+r) \quad (2.34)$$

เมื่อ μ เป็นค่าช่วงก้าว โดยใช้ผลรวมของสัญญาณผิดพลาดทั้งหมดในหนึ่งบล็อกในการปรับปรุงสัมประสิทธิ์ของ $\mathbf{w}_j(k)$

ทั้งนี้ การประมวลผลของผลบวกคอนโวลูชันในการประมวลผลแบบบล็อก สามารถทำให้มีประสิทธิภาพสูงขึ้น โดยใช้ขั้นตอนวิธีการแปลงฟูริเยร์ไม่ต่อเนื่อง (Discrete Fourier Transform, DFT) ดังนั้น จึงสามารถพูดได้ว่าเป็นการประมวลผลสัญญาณในโดเมนความถี่ของขั้นตอนวิธี LMS2 ที่มีการประมวลผลแบบบล็อก โดยเรียกขั้นตอนวิธีนี้ว่าขั้นตอนวิธี LMS2 แบบเร็ว (Fast LMS2, F-LMS2) [7], [17]

ขั้นตอนวิธี F-LMS2 นี้ สามารถถูกปรับปรุงสัมประสิทธิ์ได้โดยการทำให้ค่าช่วงก้าว μ เป็นฟังก์ชันของค่าพลังงานของสัญญาณเข้า [19] กล่าวคือ

$$\mu_r = \frac{\alpha}{P_{r,j}}, \quad r = 0,1,\dots,L-1 \text{ และ } j = 1,2 \quad (2.35)$$

เมื่อ α เป็นค่าคงที่ และค่าประมาณของพลังงานของสัญญาณเข้าในแต่ละบินความถี่ (Frequency Bin) หาได้จาก

$$P_{r,j}(k) = \gamma P_{r,j}(k-1) + (1-\gamma) |X_{i,j}(k)|^2, \quad i = 0,1,\dots,2L-1 \quad (2.36)$$

เมื่อ γ คือ ตัวประกอบลืม (Forgetting Factor) ซึ่งเป็นค่าคงที่ จะได้ว่าเมทริกซ์ทแยงมุม (Diagonal Matrix) ของค่าช่วงก้าวจะเป็น

$$\boldsymbol{\mu}(k) = \alpha \mathbf{D}_j(k) \quad (2.37)$$

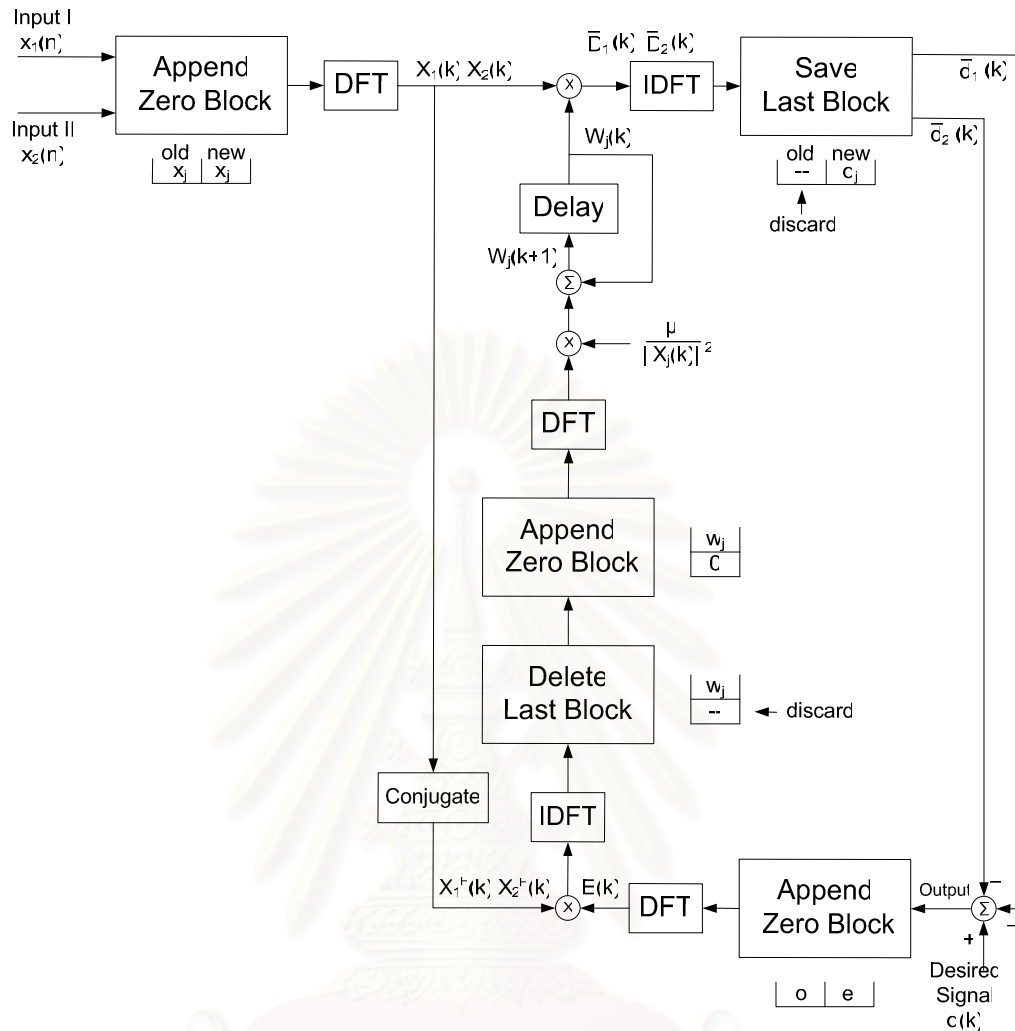
เมื่อ $\mathbf{D}_j(k)$ เป็นเมทริกซ์ทแยงมุมที่หาได้จาก

$$\mathbf{D}_j(k) = \text{diag}[\mathbf{P}_{0,j}^{-1}(k) \quad \mathbf{P}_{1,j}^{-1}(k) \quad \dots \quad \mathbf{P}_{2L-1,j}^{-1}(k)] \quad (2.38)$$

ดังนั้น สมการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวในโดเมนความถี่ด้วยขั้นตอนวิธี F-NLMS2 ซึ่งหาได้จากวิธีเก็บข้อมูลซ้อนเหลื่อม (Overlap-save Method) 50 เปอร์เซ็นต์ [15] คือ ความยาวของเวกเตอร์การแปลงจะเป็นสองเท่าของเวกเตอร์ในโดเมนเวลา (N -point DFT = $2L$) จะเป็นดังต่อไปนี้

$$\mathbf{W}_j(k+1) = \mathbf{W}_j(k) + \frac{\boldsymbol{\mu}(k)}{\varepsilon + \|\mathbf{X}_j(k)\|^2} \mathbf{X}_j^H(k) \mathbf{E}(k), \quad k = 0, 1, \dots \quad (2.39)$$

เมื่อ $\mathbf{X}_j^H(k) \mathbf{E}(k)$ เป็นการคูณแบบแถวลำดับ (Array Multiply) โดยที่ $\mathbf{X}(k)$ คือ เวกเตอร์ในโดเมนความถี่ของสัญญาณเข้า ณ บล็อก k หาได้จากการ DFT $\begin{bmatrix} \mathbf{x}(k) \\ \mathbf{x}(k-1) \end{bmatrix}$, $\mathbf{E}(k)$ คือ เวกเตอร์ในโดเมนความถี่ของสัญญาณผิดพลาด ณ บล็อก k หาได้จากการ DFT $\begin{bmatrix} \mathbf{0} \\ \mathbf{e}(k) \end{bmatrix}$ และ $\mathbf{W}_j(k)$ เป็นสัมประสิทธิ์ในโดเมนความถี่จากการแปลงฟูริเยร์ของสัมประสิทธิ์ของวงจรรองแบบปรับตัว $\mathbf{w}_j(k)$ กล่าวคือ $\mathbf{W}_j(k) = \text{DFT} \begin{bmatrix} \mathbf{w}_j(k) \\ \mathbf{0} \end{bmatrix}$ อย่างไรก็ตาม การประมวลผลแบบบล็อกนั้นจะเกิดเวลาประวิงในตัว (Inherent Delay) เนื่องจากจะต้องรอข้อมูลใหม่จำนวน L ตัวทุกครั้งก่อนการประมวลผลหนึ่งบล็อก ซึ่งเวลาประวิงที่มากเกินไปอาจไม่เหมาะสมแก่การประยุกต์ใช้กับงานในเวลาจริง (Real-time) ดังนั้น โครงสร้างการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี F-NLMS2 ด้วยวิธีเก็บข้อมูลซ้อนเหลื่อม 50 เปอร์เซ็นต์สามารถแสดงได้ดังรูปที่ 2.4



รูปที่ 2.4 โครงสร้างการทำงานของวงจรกรองแบบปรับตัวด้วยขั้นตอนวิธี F-NLMS2 ด้วยวิธีเก็บข้อมูลซ้อนเหลื่อม 50 เปอร์เซ็นต์

2.5 หลักการทำงานของวงจรกรองแบบปรับตัวด้วยขั้นตอนวิธี Affine Projection2

ขั้นตอนวิธี Affine Projection2 (AP2) เป็นขั้นตอนวิธีที่มีอัตราการลู่เข้าและความซับซ้อนทางการคำนวณอยู่ระหว่างขั้นตอนวิธี NLMS2 และขั้นตอนวิธี RLS2 [2] กล่าวคือขั้นตอนวิธี AP2 จะมีอัตราการลู่เข้าที่แปรผันตรงกับอันดับ (Order, p) ของขั้นตอนวิธี กล่าวคือ ถ้าอันดับเท่ากับ 1 ขั้นตอนวิธี AP2 จะมีลักษณะเดียวกับขั้นตอนวิธี NLMS2 ถ้าเพิ่มอันดับสูงขึ้นจะทำให้อัตราการลู่เข้าเร็วขึ้น แต่ความซับซ้อนทางการคำนวณก็จะเพิ่มขึ้นตามจำนวนอันดับที่เพิ่มขึ้นเช่นกัน ขั้นตอนวิธี AP2 มีสมการที่ใช้ในการคำนวณดังต่อไปนี้

เมื่อกำหนดให้เวกเตอร์ของสัญญาณไมโครโฟนในห้องรับขนาด $p \times 1$ เป็น

$$\mathbf{d}(n) = [d(n) \quad d(n-1) \quad \dots \quad d(n-p+1)]^T \quad (2.40)$$

และเมทริกซ์ข้อมูลของสัญญาณเข้าแต่ละช่องสัญญาณ $i = 1, 2$ ขนาด $L \times p$ เป็น

$$\mathbf{X}_i(n) = [\mathbf{x}_i(n) \quad \mathbf{x}_i(n-1) \quad \dots \quad \mathbf{x}_i(n-p+1)] \quad (2.41)$$

$$\mathbf{X}_i(n) = \begin{bmatrix} x_i(n) & x_i(n-1) & \dots & x_i(n-p+1) \\ x_i(n-1) & x_i(n-2) & \dots & x_i(n-p) \\ \vdots & \vdots & \ddots & \vdots \\ x_i(n-L+1) & x_i(n-L) & \dots & x_i(n-p-L+2) \end{bmatrix} \quad (2.42)$$

โดยที่เวกเตอร์ความผิดพลาดค่าก่อนหน้า (A Priori Error) ขนาด $p \times 1$ เป็น

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}_1^T(n) \mathbf{w}_1(n-1) - \mathbf{X}_2^T(n) \mathbf{w}_2(n-1) \quad (2.43)$$

และเวกเตอร์ความผิดพลาดค่าหลัง (Posterior Error) ขนาด $p \times 1$ จะเป็น

$$\xi(n) = \mathbf{d}(n) - \mathbf{X}_1^T(n) \mathbf{w}_1(n) - \mathbf{X}_2^T(n) \mathbf{w}_2(n) \quad (2.44)$$

การปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัว ต้องการทำให้ค่าผิดพลาดของเวกเตอร์ความผิดพลาดค่าหลัง $\xi(n) = \mathbf{0}$ ดังนั้น เมื่อหาค่าผลต่างของสมการที่ (2.43) และสมการที่ (2.44) จะได้เป็น

$$\mathbf{e}(n) = \sum_{i=1}^2 \mathbf{X}_i^T(n) (\mathbf{w}_i(n) - \mathbf{w}_i(n-1)) \quad (2.45)$$

เมื่อ

$$\Delta \mathbf{w}_i(n) \equiv \mathbf{w}_i(n) - \mathbf{w}_i(n-1) \quad (2.46)$$

หรือเขียนได้เป็น

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + \Delta\mathbf{w}_i(n) \quad (2.47)$$

ถ้า $p < L$ สมการที่ (2.42) จะเป็นสมการที่มีจำนวนตัวแปรมากกว่าจำนวนสมการ (Under-determined) ทำให้ไม่สามารถแก้สมการได้โดยตรง และคำตอบที่เหมาะสมที่สุดของ $\Delta\mathbf{w}_i(n)$ สามารถหาได้จาก

$$\Delta\mathbf{w}_i(n) = [\mathbf{X}_i^T(n)]^{\ast-1} \mathbf{e}(n) \quad (2.48)$$

เมื่อ $[\mathbf{X}_i^T(n)]^{\ast-1}$ คือ เมตริกซ์ผกผันเทียม (Pseudo-inverse Matrix) ของเมตริกซ์ $\mathbf{X}_i^T(n)$ เขียนได้เป็น

$$[\mathbf{X}_i^T(n)]^{\ast-1} = \mathbf{X}_i(n) [\mathbf{X}_i^T(n) \mathbf{X}_i(n)]^{-1} \quad (2.49)$$

เมื่อแทนค่า $[\mathbf{X}_i^T(n)]^{\ast-1}$ จากสมการที่ (2.49) ลงในสมการที่ (2.48) จะได้

$$\Delta\mathbf{w}_i(n) = \mathbf{X}_i(n) [\mathbf{X}_i^T(n) \mathbf{X}_i(n)]^{-1} \mathbf{e}(n) \quad (2.50)$$

ซึ่งจะเห็นได้ว่า ถ้าค่า $[\mathbf{X}_i^T(n) \mathbf{X}_i(n)]^{-1}$ มีการกระจายค่าลักษณะเฉพาะ (Eigenvalue Spread) ที่มีค่ามาก จะทำให้เกิดปัญหาในการหาค่าเมตริกซ์ผกผัน จึงต้องมีการเพิ่มเทอมเมตริกซ์เรกกูลาไรเซชัน (Regularization Matrix, $\delta\mathbf{I}$) เพื่อป้องกันการหารด้วยค่าน้อยมากๆ ดังนั้น สมการที่ (2.50) จึงกลายเป็น

$$\Delta\mathbf{w}_i(n) = \mathbf{X}_i(n) [\mathbf{X}_i^T(n) \mathbf{X}_i(n) + \delta\mathbf{I}]^{-1} \mathbf{e}(n) \quad (2.51)$$

เมื่อกำหนดให้

$$\boldsymbol{\varepsilon}(n) = [\mathbf{X}_i^T(n) \mathbf{X}_i(n) + \delta\mathbf{I}]^{-1} \mathbf{e}(n) \quad (2.52)$$

ดังนั้น สมการปรับปรุงค่าสัมประสิทธิ์ของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี AP2 ในสมการที่ (2.47) สามารถเขียนใหม่ได้เป็น

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + \mu \mathbf{X}_i(n)\mathbf{e}(n) \quad (2.53)$$

2.6 หลักการทำงานของวงจรรองแบบปรับตัวด้วยขั้นตอนวิธี Fast Least Square2

ถึงแม้ว่าขั้นตอนวิธี Fast Least Square2 (FLS2) จะมีอัตราการลู่เข้าที่เร็วใกล้เคียงกับขั้นตอนวิธี Recursive Least Squares2 (RLS2) [2] แต่ก็มีข้อจำกัดข้อหนึ่งคือมีความซับซ้อนทางการคำนวณที่สูงเมื่อเปรียบเทียบกับขั้นตอนวิธี LMS2 และ NLMS2 หลักการปรับปรุงสัมประสิทธิ์ด้วยขั้นตอนวิธี FLS2 สามารถแบ่งได้เป็น 2 ส่วน คือ การปรับปรุงสัมประสิทธิ์ของอัตราขยายแบบปรับตัว (Adaptation Gain Updating) และการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัว (Adaptive Filter Updating) โดยมีรายละเอียดดังต่อไปนี้

กำหนดตัวแปรเริ่มต้น ณ เวลา n โดยที่ $i = 1, 2$ ได้แก่

สัมประสิทธิ์ของวงจรรองแบบปรับตัว	$\mathbf{w}_i(n)$
สัมประสิทธิ์การทำนายล่วงหน้า (Forward Prediction Coefficients)	$\mathbf{A}_i(n)$
สัมประสิทธิ์การทำนายย้อนกลับ (Backward Prediction Coefficients)	$\mathbf{B}_i(n)$
เวกเตอร์สัญญาณเข้า	$\mathbf{X}_i(n)$
การทำนายพลังงานของค่าคลาดเคลื่อนล่วงหน้าและย้อนกลับ	$E_{i,a}(n), E_{i,b}(n)$
การทำนายอัตราส่วนของค่าคลาดเคลื่อนล่วงหน้าและย้อนกลับ	$\alpha_i(n)$
ตัวประกอบถ่วงน้ำหนัก (Weighting Factor)	\mathbf{W}_i

และตัวแปรเริ่มต้น ณ เวลา $n+1$ ได้แก่

เวกเตอร์สัญญาณเข้า	$\mathbf{x}_i(n+1)$
สัญญาณไมโครโฟนในห้องรับ	$d(n+1)$

ดังนั้น สมการการปรับปรุงสัมประสิทธิ์ของอัตราขยายแบบปรับตัว เป็นดังต่อไปนี้

$$e_{i,a}(n+1) = x_i(n+1) - \mathbf{A}_i^T(n)\mathbf{X}_i(n) \quad (2.54)$$

$$\mathbf{A}_i(n+1) = \mathbf{A}_i(n) + \mathbf{G}_i(n)e_{i,a}(n+1)/\alpha_i(n) \quad (2.55)$$

$$\mathbf{E}_{i,a}(n+1) = (\mathbf{E}_{i,a}(n) + e_{i,a}(n+1)e_{i,a}(n+1)/\alpha_i(n))\mathbf{W}_i \quad (2.56)$$

$$\mathbf{G}_{i,1}(n+1) = \begin{bmatrix} 0 \\ \mathbf{G}_i(n) \end{bmatrix} + \frac{e_{i,a}(n+1)}{\mathbf{E}_{i,a}(n+1)} \begin{bmatrix} 1 \\ -\mathbf{A}_i(n+1) \end{bmatrix} = \begin{bmatrix} \mathbf{M}_i(n+1) \\ \mathbf{m}_i(n+1) \end{bmatrix} \quad (2.57)$$

$$e_{i,b}(n+1) = x_i(n+1-L) - \mathbf{B}_i^T(n)\mathbf{X}_i(n+1) \quad (2.58)$$

$$\mathbf{G}_i(n+1) = \mathbf{M}_i(n+1) + \mathbf{m}_i(n+1)\mathbf{B}_i(n) \quad (2.59)$$

$$\alpha_{i,1}(n+1) = \alpha_i(n) + e_{i,a}(n+1)e_{i,a}(n+1)/\mathbf{E}_{i,a}(n+1) \quad (2.60)$$

$$\alpha_i(n+1) = \alpha_{i,1}(n+1) - \mathbf{m}_i(n+1)e_{i,b}(n+1) \quad (2.61)$$

$$\mathbf{E}_{i,b}(n+1) = (\mathbf{E}_{i,b}(n) + e_{i,b}(n+1)e_{i,b}(n+1)/\alpha_i(n+1))\mathbf{W}_i \quad (2.62)$$

$$\mathbf{B}_i(n+1) = \mathbf{B}_i(n) + \mathbf{G}_i(n+1)e_{i,b}(n+1)/\alpha_i(n+1) \quad (2.63)$$

และสมการการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัว คือ

$$e_i(n+1) = d(n+1) - \mathbf{w}_i^T(n)\mathbf{X}_i(n+1) \quad (2.64)$$

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \mathbf{G}_i(n+1)e_i(n+1) \quad (2.65)$$

2.7 ระบบการตัดสัญญาณเสียงสะท้อนด้วยโครงสร้างการกรองแถบย่อย

โครงสร้างการกรองแถบย่อย (Subband Filtering Structure) ได้รับความนิยมนำไปใช้อย่างแพร่หลายในการประมวลผลสัญญาณดิจิทัล เพื่อปรับปรุงสมรรถนะของการประมวลผลสัญญาณทั้งในโดเมนเวลาและโดเมนความถี่ ทั้งนี้ ระบบที่มีโครงสร้างการกรองแถบย่อยจะทำการประมวลผลโดยใช้หลักการของระบบหลายอัตรา (Multirate System) เพื่อลดปริมาณข้อมูลที่ใช้ในการประมวลผล และใช้คลังวงจรรอง (Filter Bank) ในการแบ่งวงจรรองอันดับสูงออกเป็นวงจรรองอันดับต่ำหลายวงจรรองที่ช่วงความถี่ต่างๆกัน ข้อดีของการประมวลผลด้วยโครงสร้างการกรองแถบย่อย คือ การเพิ่มอัตราการใช้ของระบบ และการลด

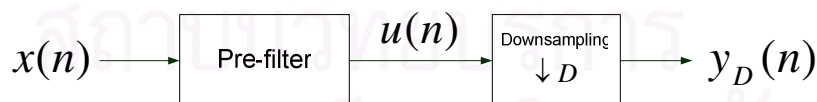
ความซับซ้อนทางการคำนวณ เนื่องจากจำนวนสัมประสิทธิ์ของวงจรรองในแต่ละแถบความถี่ย่อยมีจำนวนลดลงเมื่อเปรียบเทียบกับระบบปกติที่มีการประมวลผลในโดเมนเวลา [6]-[8] นอกจากนี้ ถึงแม้ว่าระบบที่มีโครงสร้างการกรองแถบย่อยจะทำการประมวลผลสัญญาณแบบบล็อก หากแต่ถ้าโครงสร้างการกรองแถบย่อยอาศัยการประมวลผลสัญญาณผ่านวงจรรองอันดับต่ำหลายวงจรรองพร้อมกัน จึงช่วยลดเวลาประวิงที่เกิดขึ้นกับการประมวลผลสัญญาณแบบบล็อก [5]

2.7.1 ระบบหลายอัตรา (Multirate System)

ลักษณะเฉพาะที่สำคัญของการประมวลผลสัญญาณโดยใช้โครงสร้างการกรองแถบย่อย คือ การใช้หลักการของระบบหลายอัตรา [5] ทั้งนี้ การประมวลผลแบบหลายอัตราเป็นการแบ่งข้อมูลออกเป็นแถบความถี่ย่อยหลายแถบความถี่ เพื่อให้สามารถประมวลผลข้อมูลในแต่ละแถบความถี่ย่อยพร้อมกันได้ นอกจากนี้ระบบหลายอัตรายังสามารถเพิ่มหรือลดความถี่ในการซิกตัวอย่างของสัญญาณในแต่ละแถบความถี่ย่อย เพื่อการประมวลผลสัญญาณที่มีอัตราการซิกตัวอย่างแตกต่างไปจากสัญญาณเข้าของระบบเดิม ดังนั้น ประโยชน์ของระบบหลายอัตราคือการประมวลผลที่รวดเร็ว เนื่องจากสามารถประมวลผลข้อมูลในแต่ละแถบความถี่ย่อยพร้อมกัน และลดความซับซ้อนในการคำนวณของแต่ละแถบความถี่ย่อยได้ เนื่องจากอันดับที่ต่ำลงของวงจรรองแต่ละตัวในคลังวงจรรอง จึงสามารถนำไปใช้ในทางปฏิบัติได้อย่างมีประสิทธิภาพ

องค์ประกอบที่สำคัญของระบบหลายอัตรา ได้แก่ การทำเดซิเมชัน (Decimation) และการทำอินเทอร์โพลชัน (Interpolation) ดังรายละเอียดในหัวข้อย่อยต่อไปนี้

2.7.1.1 การทำเดซิเมชัน (Decimation)



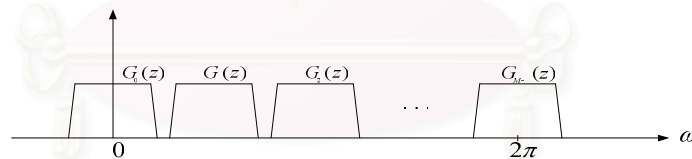
รูปที่ 2.5 การทำเดซิเมชัน

การทำเดซิเมชันคือ การวิเคราะห์สัญญาณผ่านพรีฟิวเตอร์ (Pre-filter) เพื่อให้ได้ข้อมูลในแถบความถี่ส่วนที่ต้องการ จากนั้นจึงทำการลดการซิกตัวอย่าง (Downsampling) เพื่อลดปริมาณข้อมูลที่ใช้ในการประมวลผล ดังรูปที่ 2.5 โดยที่การออกแบบผลตอบสนองของทางความถี่ (Frequency Response) ของพรีฟิวเตอร์ สามารถแบ่งได้เป็นสามประเภทด้วยกัน [4] คือ

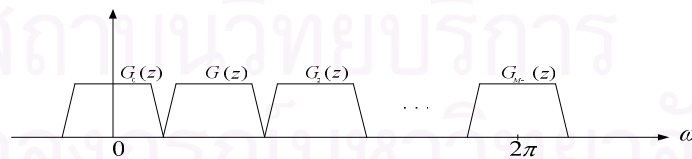
1. วงจรกรองที่ผลตอบสนองทางความถี่ไม่ซ้อนเหลื่อมกัน (Non-overlapping Frequency Response) โดยความถี่ตัดออก (Cut-off Frequency, ω_c) ของวงจรกรองแต่ละตัวมีค่าเป็น $0 < \omega_c < \pi/M$ เมื่อ M เป็นจำนวนแถบความถี่ย่อย ดังแสดงในรูปที่ 2.6 (ก) วงจรกรองจะทำการกรองข้อมูลในแถบความถี่ย่อยใดๆ โดยไม่มีข้อมูลของแถบความถี่รอบข้างติดมาด้วย แต่ข้อมูลในช่วงความถี่ระหว่างวงจรกรองที่ติดกันจะไม่ถูกนำมาพิจารณาด้วย

2. วงจรกรองที่ขอบของผลตอบสนองทางความถี่ที่ติดกันซ้อนทับกันพอดี (Marginally-overlapping Frequency Response) ความถี่ตัดออกของวงจรกรองประเภทนี้มีค่า $\omega_c = \pi/M$ เมื่อแบ่งข้อมูลออกเป็น M แถบความถี่ย่อย จะทำให้ข้อมูลในทุกช่วงความถี่ถูกแบ่งมาส่วนละเท่าๆ กันเพื่อพิจารณาพร้อมกันในแต่ละแถบความถี่ย่อย ดังแสดงในรูปที่ 2.6 (ข) วงจรกรองรูปแบบนี้นำไปใช้ในการส่งข้อมูลผ่านช่องสัญญาณโดยไม่ถูกรบกวนจากช่องสัญญาณรอบข้าง นอกจากนี้ ในแต่ละแถบความถี่ย่อยจะมีการทำเดซิเมชันด้วยการลดอัตราการซีกตัวอย่างด้วยตัวประกอบเดซิเมชัน D ที่มีค่ามากที่สุดที่ไม่ทำให้เกิดผลความผิดเพี้ยนภาพ (Aliasing Effect)

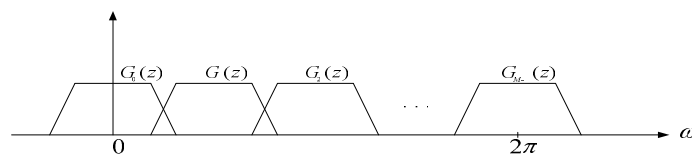
3. วงจรกรองที่ผลตอบสนองทางความถี่ซ้อนเหลื่อมกัน (Overlapping Frequency Response) คือ ความถี่ตัดออกของวงจรกรองมีค่า $\omega_c > \pi/M$ ดังแสดงในรูปที่ 2.6 (ค) จึงทำให้มีข้อมูลของแถบความถี่รอบข้างติดมาประมวลผลในแถบความถี่ใกล้เคียงด้วย วงจรกรองรูปแบบนี้สามารถประยุกต์ใช้ในบางระบบ เช่น การเข้ารหัสสัญญาณดิจิทัล ที่ต้องการข้อมูลของแถบความถี่รอบข้างสำหรับการสังเคราะห์สัญญาณออก [4]



(ก) Non-overlapping Frequency Response



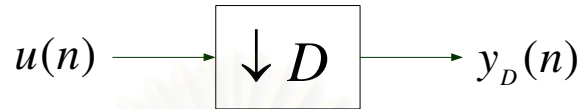
(ข) Marginally-overlapping Frequency Response



(ค) Overlapping Frequency Response

รูปที่ 2.6 ผลตอบสนองทางความถี่รูปแบบต่างๆ ของคลังวงจรกรอง

โดยทั่วไปการออกแบบพีวีพีแตรในการทำเดซิเมชันเพื่อให้ขอบของผลตอบสนองทางความถี่ที่ติดกันซ้อนทับกันพอดี (ดังหัวข้อที่ 2) ทำได้ยาก ส่วนมากจะเกิดการซ้อนเหลื่อมกันของผลตอบสนองทางความถี่ เนื่องจากระลอก (Ripple) ของแถบความถี่ข้าง (Sideband) ทำให้ข้อมูลเกิดการรั่วไหล (Leakage) ไปยังแถบความถี่ข้างได้



รูปที่ 2.7 การลดอัตราการชักตัวอย่าง

สำหรับการลดการชักตัวอย่างข้อมูลในแต่ละแถบความถี่ย่อย ด้วยตัวประกอบเดซิเมชัน D (Decimation Factor) จะเป็นไปตามรูปที่ 2.7 สัญญาณออกของขั้นตอนวิธีนี้เรียกว่าสัญญาณวิยุต (Discrete Signal) $y_D(n)$ ซึ่งมีความสัมพันธ์กับสัญญาณเข้า $u(n)$ ในแต่ละแถบความถี่ย่อย ดังนี้

$$y_D(n) = u(Dn) \quad \text{เมื่อ } n = 0, 1, 2, \dots \quad (2.66)$$

เมื่อทำการแปลง Z (Z-transformation) กับสัญญาณ $y_D(n)$ จะได้ว่า

$$Y_D(z) = \sum_{n=-\infty}^{\infty} u(Dn)z^{-n} \quad (2.67)$$

สัญญาณ $u(Dn)$ เป็นสัญญาณวิยุตเนื่องจาก n เป็นจำนวนเต็ม (Integer) ใดๆ จึงสามารถเขียนแทนด้วย $u_{\text{int}}(Dn)$ [16] ดังนั้นสมการที่ (2.67) จะเขียนได้เป็น

$$Y_D(z) = \sum_{n=-\infty}^{\infty} u_{\text{int}}(Dn)z^{-n} \quad (2.68)$$

สัญญาณ $u_{\text{int}}(n)$ จะมีค่าเท่ากับ $u(n)$ เมื่อ n เป็นจำนวนเท่าของ D และจะมีค่าเป็นศูนย์ เมื่อ n เป็นค่าอื่นที่ไม่เป็นจำนวนเท่าของ D จึงสามารถแสดงความสัมพันธ์ระหว่างสัญญาณ $u_{\text{int}}(n)$ กับ $u(n)$ ได้ดังสมการที่ (2.69)

$$u_{\text{int}}(n) = c(n)u(n) \quad (2.69)$$

โดยที่

$$c(n) = \begin{cases} 1, & n = 0, \pm D, \pm 2D, \dots \\ 0, & \text{otherwise} \end{cases} \quad (2.70)$$

เมื่อให้ $W_D = e^{\frac{-j2\pi}{D}}$ เป็นการแจกแจงแบบคู่ (Even Distribution) บนวงกลมหนึ่งหน่วย (Unit Circle) ตัวแปร $c(n)$ สามารถเขียนได้เป็น

$$c(n) = \frac{1}{D} \sum_{k=0}^{D-1} W_D^{-kn} \quad (2.71)$$

ดังนั้นเมื่อทำการแปลง z กับสมการที่ (2.69) และแทนค่า $c(n)$ จากสมการที่ (2.71) จะได้ว่า

$$U_{\text{int}}(z) = \sum_{n=-\infty}^{\infty} c(n)u(n)z^{-n} \quad (2.72)$$

$$= \frac{1}{D} \sum_{n=-\infty}^{\infty} \left(\sum_{k=0}^{D-1} W_D^{-kn} \right) u(n)z^{-n} \quad (2.73)$$

$$= \frac{1}{D} \sum_{k=0}^{D-1} U(zW_D^k) \quad (2.74)$$

เมื่อกำหนดให้ $Dn = k$ และใช้ความสัมพันธ์ในสมการที่ (2.68) และ (2.72) สัญญาณที่ผ่านการลดอัตราการซีกตัวอย่างในโดเมน z สามารถเขียนได้เป็น

$$Y_D(z) = \sum_{k=-\infty}^{\infty} u_{\text{int}}(k)z^{-k/D} \quad (2.75)$$

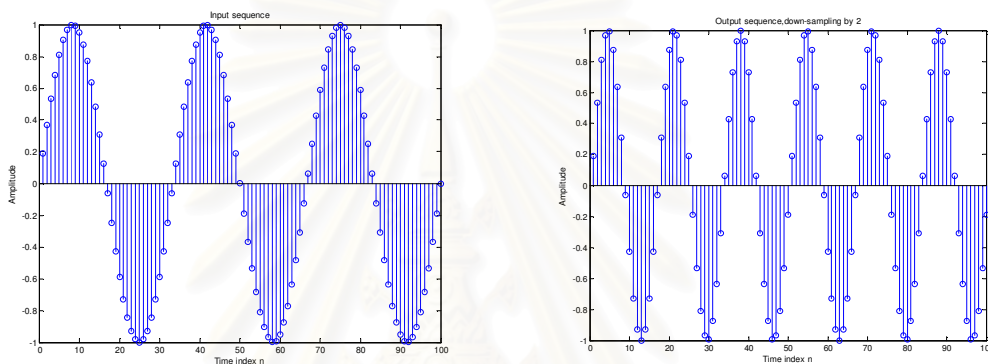
$$= U_{\text{int}}(z^{1/D}) \quad (2.76)$$

และเมื่อนำสมการที่ (2.74) แทนค่าลงในสมการที่ (2.76) จะได้ว่า

$$Y_D(z) = \frac{1}{D} \sum_{k=0}^{D-1} U\left(z^{1/D} W_D^k\right) \tag{2.77}$$

สำหรับสัญญาณคงตัวใดๆ สามารถแทนค่า z ด้วย $e^{j\omega}$ ดังนั้นผลตอบสนองทางความถี่ (Frequency Response) ของสัญญาณที่ถูกลดอัตราการซีกตัวอย่าง $Y_D(z)$ หาได้จาก

$$Y_D(e^{j\omega}) = \frac{1}{D} \sum_{k=0}^{D-1} U\left(e^{j(\omega-2\pi k)/D}\right) \tag{2.78}$$



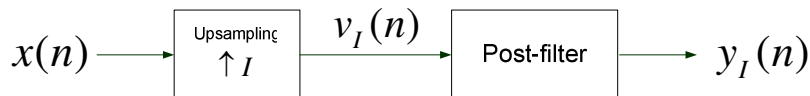
(ก) สัญญาณเข้า $u(n)$

(ข) สัญญาณออก $y_D(n)$

รูปที่ 2.8 ตัวอย่างการลดอัตราการซีกตัวอย่างด้วยตัวประกอบเดซิเมชัน $D = 2$

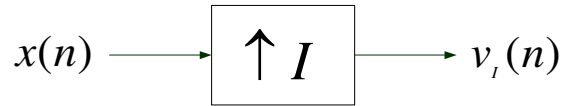
พิจารณารูปที่ 2.8 (ก) และ 2.8 (ข) พบว่าสัญญาณออกที่ได้จากการลดอัตราการซีกตัวอย่างสัญญาณเข้าของระบบด้วยตัวประกอบเดซิเมชัน D นั้น จะทำการซีกตัวอย่างด้วยอัตราที่ลดลง D เท่า โดยที่คุณลักษณะของสัญญาณในทางเวลายังคงแสดงถึงฟังก์ชันของสัญญาณเดิมอยู่

2.7.1.2 การทำอินเทอร์โพลชัน (Interpolation)



รูปที่ 2.9 การทำอินเทอร์โพลชัน

การทำอินเทอร์โพลชัน คือ การเพิ่มการซีกตัวอย่างข้อมูล (Upsampling) ทั้งนี้ ส่วนมากมักตามด้วยขั้นตอนการสังเคราะห์ข้อมูลด้วยโพลฟิลเตอร์ (Post-filter) เพื่อกำจัดผลความผิดเพี้ยนภาพเมื่อผลตอบสนองทางความถี่เกิดการซ้อนเหลื่อมกัน ดังรูปที่ 2.9



รูปที่ 2.10 การเพิ่มอัตราการซีกตัวอย่าง

เมื่อพิจารณาส่วนของการเพิ่มอัตราการซีกตัวอย่างด้วยตัวประกอบอินเทอร์โพลชัน I (Interpolation Factor) ดังแสดงในรูปที่ 2.10 ความสัมพันธ์ของสัญญาณออก $v_I(n)$ กับสัญญาณเข้า $x(n)$ เป็นไปดังสมการ

$$v_I(n) = \begin{cases} x(n/I), & n = 0, \pm I, \pm 2I, \dots \\ 0, & \text{otherwise} \end{cases} \quad (2.79)$$

ซึ่งหมายความว่า ระหว่างสัญญาณที่ถูกซีกตัวอย่างสองตัวอย่างที่อยู่ติดกันจะมีการเติมศูนย์จำนวน $I-1$ ตัวอย่าง เมื่อทำการแปลง Z กับสมการที่ (2.79) จะได้ว่า

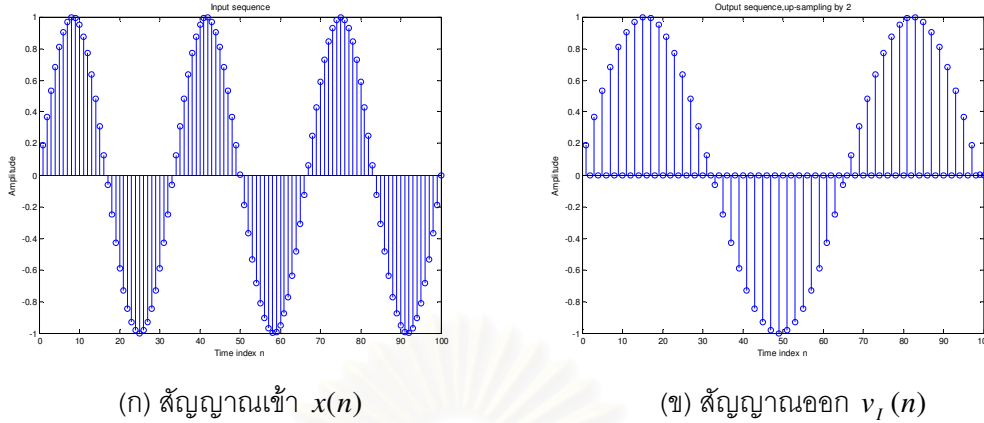
$$V_I(z) = \sum_{n=-\infty}^{\infty} x(n/I)z^{-n} \quad (2.80)$$

กำหนดให้ $n/I = m$ สมการที่ (2.80) สามารถเขียนได้เป็น

$$V_I(z) = \sum_{m=-\infty}^{\infty} x(m)z^{-mI} \quad (2.81)$$

$$= X(z^I) \quad (2.82)$$

ตัวอย่างการเพิ่มอัตราการซีกตัวอย่างด้วยตัวประกอบอินเทอร์โพลชัน $I = 2$ ดังแสดงในรูปที่ 2.11 พบว่าสัญญาณออกที่ได้จากการเพิ่มอัตราการซีกตัวอย่างสัญญาณเข้าของระบบด้วยตัวประกอบอินเทอร์โพลชัน $I = 2$ นั้น จะมีการเติมศูนย์จำนวน $I-1$ ตัวอย่างระหว่างสัญญาณที่ถูกซีกตัวอย่างสองตัวอย่างที่อยู่ติดกัน โดยที่คุณลักษณะของสัญญาณในทางเวลายังคงแสดงถึงฟังก์ชันของสัญญาณเดิมอยู่

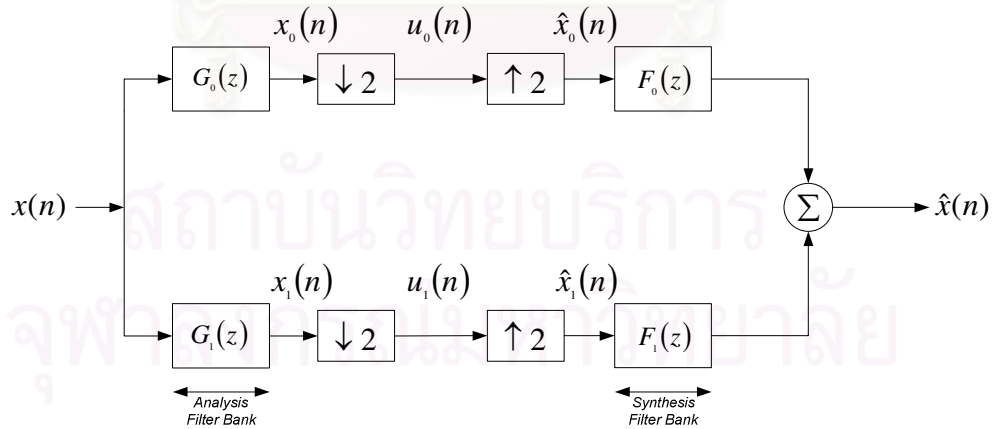


(ก) สัญญาณเข้า $x(n)$

(ข) สัญญาณออก $v_1(n)$

รูปที่ 2.11 ตัวอย่างการเพิ่มอัตราการซ้กตัวอย่างด้วยตัวประกอบอินเทอร์โพลชัน $I = 2$

สำหรับการกำจัดผลความผิดเพี้ยนภาพเมื่อผลตอบสนองทางความถี่เกิดการซ้อนเหลื่อมกันของพรีฟิวเตอร์ สามารถทำได้โดยการออกแบบโพสฟิวเตอร์ (Post-filter หรือเรียกว่า Anti-aliasing Filter) ในการทำอินเทอร์โพลชัน ได้แสดงดังตัวอย่างในการออกแบบคลังวงจรรองควอดเรเจอร์-มิลเลอร์ (Quadrature Mirror Filter Bank) ที่ปรับปรุงมาจากคลังวงจรรองแบบสองช่องแถบความถี่ย่อย ($M = 2$) ในกระบวนการนี้ให้ $G_0(z)$ และ $G_1(z)$ แทนพรีฟิวเตอร์ต่ำผ่าน (Low-pass Pre-filter) และสูงผ่าน (High-pass Pre-filter) ตามลำดับ ส่วน $F_0(z)$ และ $F_1(z)$ แทนโพสฟิวเตอร์ต่ำผ่านและสูงผ่านตามลำดับ ดังรูปที่ 2.12



รูปที่ 2.12 คลังวงจรรองควอดเรเจอร์-มิลเลอร์ สองช่องสัญญาณ ($M = 2$)

จากรูปที่ 2.12 สัญญาณสังเคราะห์ในโดเมน Z สามารถหาได้จาก

$$\hat{X}(z) = T(z)X(z) + A(z)X(-z) \quad (2.83)$$

เมื่อฟังก์ชันบิดเบี้ยว (Distortion Function) นิยามโดย

$$T(z) = \frac{1}{2}[G_0(z)F_0(z) + G_1(z)F_1(-z)] \quad (2.84)$$

และฟังก์ชันถ่ายโอน (Transfer Function) ซึ่งมีผลความบิดเบี้ยวภาพของสัญญาณ $X(-z)$ เป็น

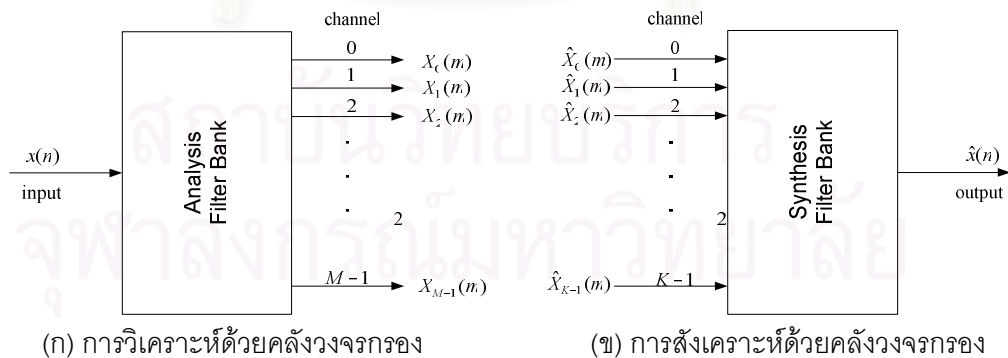
$$A(z) = \frac{1}{2}[G_0(-z)F_0(z) + G_1(-z)F_1(z)] \quad (2.85)$$

เมื่อต้องการให้ไม่มีความบิดเบี้ยวภาพของสัญญาณ สามารถออกแบบโพลีโนเมียล $F(z)$ ได้โดยการกำหนดให้ $A(z) = 0$ ดังนั้น

$$F_0(z) = G_1(-z) \quad \text{และ} \quad F_1(z) = -G_0(-z) \quad (2.86)$$

ด้วยการเลือกโพลีโนเมียล ดังสมการ (2.86) จะสามารถสังเคราะห์สัญญาณออก ที่มีส่วนประกอบเหมือนกับสัญญาณเข้า จึงสามารถกำจัดความบิดเบี้ยวภาพของสัญญาณได้

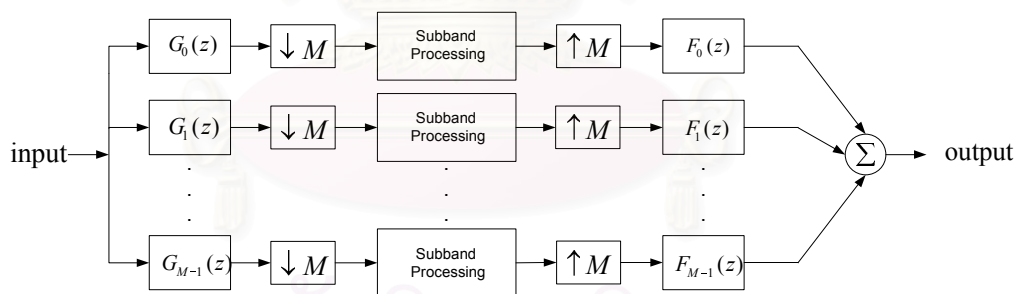
2.7.2 การกรองแถบย่อย (Subband Filtering)



รูปที่ 2.13 การวิเคราะห์ด้วยคลังวงจรรองและการสังเคราะห์ด้วยคลังวงจรรอง

การวิเคราะห์ด้วยคลังวงจรรอง (Analysis Filter Bank) เป็นการทำให้สัญญาณโดยที่คลังวงจรรองวิเคราะห์จะแบ่งข้อมูลในโดเมนความถี่ออกเป็นแถบข้อมูลย่อย M แถบความถี่ดังรูปที่ 2.13 (ก) โดยใช้พีวีเตอร์และลดการชักตัวอย่างข้อมูลในแต่ละแถบความถี่ย่อย จากนั้นจึงนำข้อมูลแต่ละแถบย่อยไปประมวลผล เนื่องจากสามารถประมวลผลข้อมูลในแต่ละความถี่พร้อมกันจึงประมวลผลได้รวดเร็วยิ่งขึ้น และลดความซับซ้อนของจำนวนสัมประสิทธิ์ของวงจรรอง ส่วนการสังเคราะห์ด้วยคลังวงจรรอง (Synthesis Filter Bank) เป็นการสังเคราะห์สัญญาณในแต่ละแถบความถี่ย่อยที่รับเข้ามา โดยการเพิ่มการชักตัวอย่างข้อมูลที่รับเข้ามาแล้วจึงใช้พีวีเตอร์ทำการสังเคราะห์ข้อมูลให้กลับเป็นข้อมูลรวมที่มีความถี่เท่าเดิม ดังรูปที่ 2.13 (ข) สำหรับรูปแบบของคลังวงจรรองจะแบ่งออกได้เป็นสองประเภทด้วยกัน คือ คลังวงจรรองเอกกรุป (Uniform Filter Bank) และคลังวงจรรองไม่เอกกรุป (Non-uniform Filter Bank) โดยที่คลังวงจรรองเอกกรุปนั้นได้แสดงตัวอย่างดังรูปที่ 2.6 ซึ่งผลตอบสนองทางความถี่ของวงจรรองในแต่ละแถบความถี่ย่อยมีค่าเท่ากัน ดังนั้น ข้อมูลสัญญาณเข้าจะถูกแบ่งเป็น M แถบความถี่ย่อยที่เท่ากัน สำหรับคลังวงจรรองไม่เอกกรุปนั้น ผลตอบสนองทางความถี่ของวงจรรองในแต่ละแถบความถี่ย่อยจะมีค่าไม่เท่ากัน ขึ้นอยู่กับการออกแบบคลังวงจรรอง

2.7.3 การประยุกต์ใช้การกรองแถบย่อยในโดเมนเวลาและโดเมนความถี่



รูปที่ 2.14 ตัวอย่างโครงสร้างการกรองแถบย่อย M แถบความถี่ย่อย

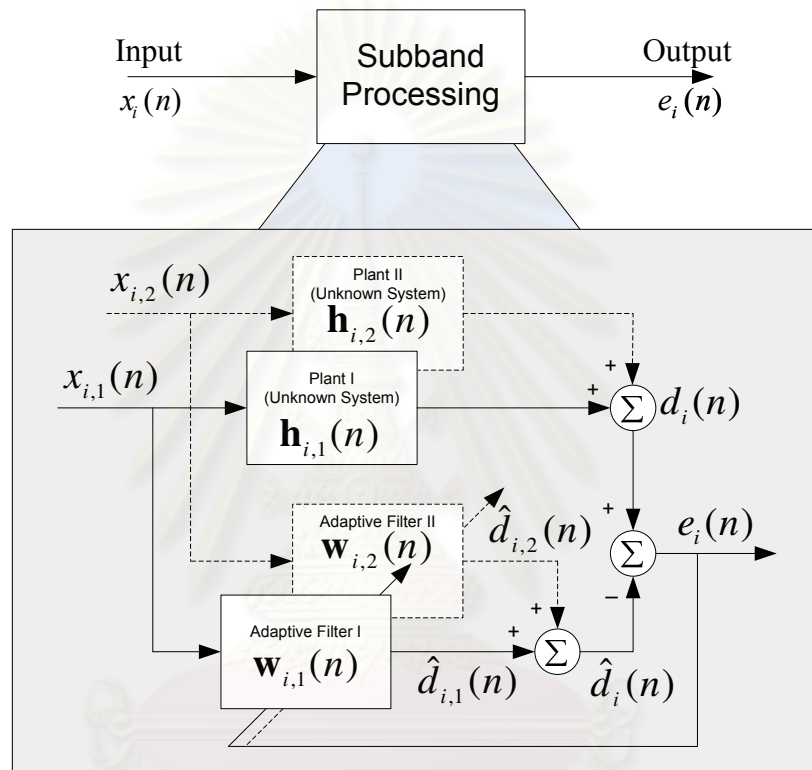
รูปที่ 2.14 แสดงตัวอย่างโครงสร้างการกรองแถบย่อย M แถบความถี่ย่อย โดยที่

$G_i(z)$ เมื่อ $i = 0, 1, \dots, M-1$ แทนคลังวงจรรองวิเคราะห์ M แถบความถี่

และ

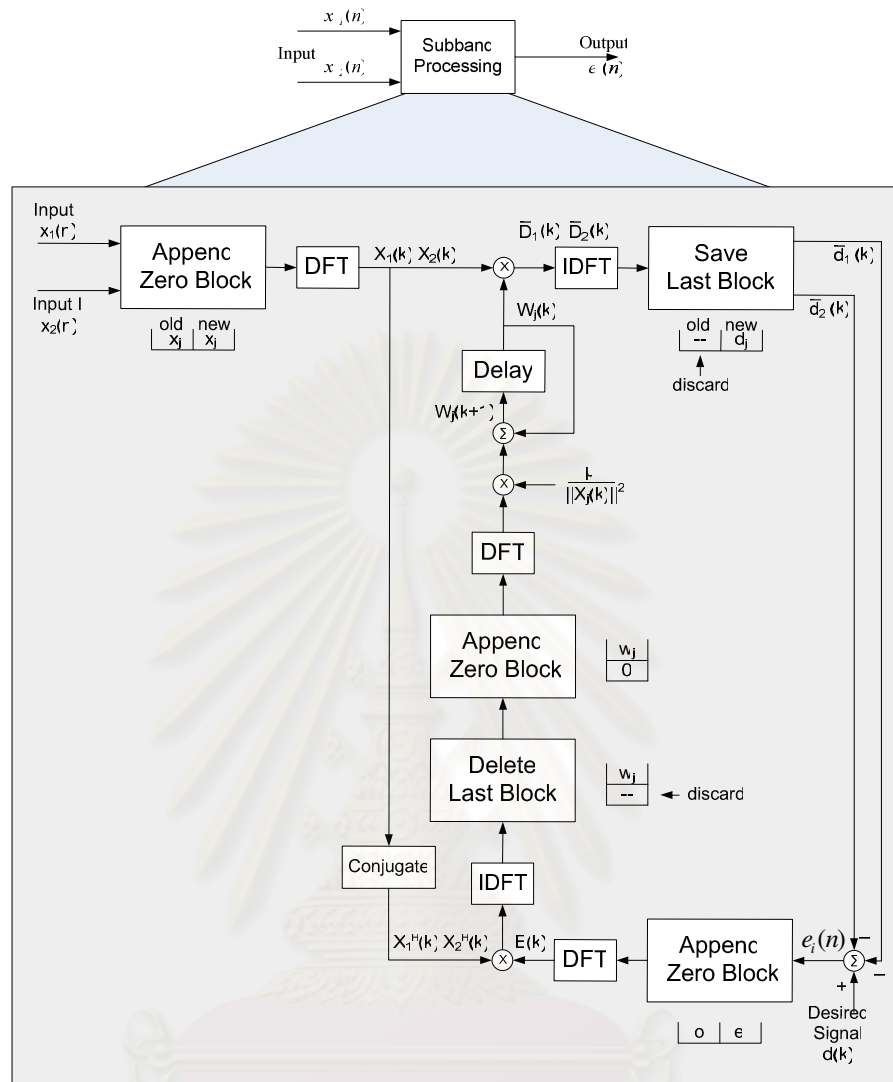
$F_i(z)$ เมื่อ $i = 0, 1, \dots, M-1$ แทนคลังวงจรรองสังเคราะห์ M แถบความถี่

หลังจากทำการวิเคราะห์ด้วยคลังวงจรกรองวิเคราะห์กับสัญญาณที่ต้องการประมวลผล จะได้ข้อมูลในแต่ละแถบความถี่ย่อยจำนวน M แถบความถี่ดังรูปที่ 2.14 ในหัวข้อนี้จะเป็นการ แสดงตัวอย่างการประยุกต์ใช้ระบบการตัดสัญญาณเสียงสะท้อนในโดเมนเวลาและในโดเมน ความถี่ด้วยขั้นตอนวิธี NLMS2 ในแต่ละแถบความถี่ย่อย โดยที่ระบบการตัดสัญญาณเสียง สะท้อนในโดเมนเวลา ได้แสดงดังรูปที่ 2.15 โดยที่ $i = 1, 2, \dots, M$ เป็นจำนวนแถบความถี่ย่อย



รูปที่ 2.15 การประยุกต์ใช้การกรองแถบย่อยกับระบบ SAEC ในโดเมนเวลา ด้วยขั้นตอนวิธี NLMS2

สำหรับการประยุกต์ใช้ระบบการตัดสัญญาณเสียงสะท้อนในโดเมนความถี่ด้วยขั้นตอนวิธี F-NLMS2 ของแต่ละแถบความถี่ย่อย ได้แสดงดังรูปที่ 2.16 โดยที่ $i = 1, 2, \dots, M$ เป็นจำนวนแถบ ความถี่ย่อย



รูปที่ 2.16 การประยุกต์ใช้การกรองแถบย่อยกับระบบ SAEC ในโดเมนความถี่ ด้วยขั้นตอนวิธี F-NLMS2

2.8 การเปรียบเทียบความซับซ้อนทางการคำนวณ (Computational Complexity)

หัวข้อนี้เป็นการเปรียบเทียบความซับซ้อนทางการคำนวณของแต่ละขั้นตอนวิธีต่อการอบการคำนวณ (Iteration) โดยจะพิจารณาในพจน์ของการบวกและลบจำนวนจริง (Real Additions and Subtractions, RASs) การคูณจำนวนจริง (Real Multiplications, RMPs) และการหาร (Division, div) ทั้งหมดนี้ต่อหนึ่งรอบการวนซ้ำ หรืออาจกล่าวได้ว่าต่อข้อมูลของสัญญาณเข้าหนึ่งตัวอย่าง (Per Input Sample) ทั้งนี้ จะพิจารณาที่การประมวลผลสัญญาณในช่องสัญญาณหนึ่งๆ โดย

กำหนดให้การหารหนึ่งครั้ง (1 div) ประมาณด้วยการคูณเท่ากับ 16 การคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่าง [14], L คือ ความยาวสัมประสิทธิ์ของวงจรรองแบบปรับตัว สำหรับความซับซ้อนทางการคำนวณของการแปลงฟูริเยร์ไม่ต่อเนื่องจะคิดเป็นแบบ Radix-2 ในโดเมนความถี่ [15] ดังนั้น แต่ละ N -point Real DFT (และ IDFT) ต้องการการคำนวณ $N \log_2 N$ RMPs เมื่อ $N = 2L$ อันดับของขั้นตอนวิธี AP = p และ $\text{Inv}(\mathbf{M}_{p \times p})$ คือ เมทริกซ์ผกผันขนาด $p \times p$

ตารางที่ 2.1 ความซับซ้อนทางการคำนวณของขั้นตอนวิธี NLMS ช่องสัญญาณที่ j

Adaptive Filter :	\times / \div	$+ / -$
$e(n) = d(n) - \mathbf{w}^T(n)\mathbf{x}(n)$	L	L
$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu \mathbf{x}(n)e(n)}{\varepsilon + \ \mathbf{x}(n)\ ^2}$	$L+3+1\text{div}$	$L+3$

ตารางที่ 2.2 ความซับซ้อนทางการคำนวณของขั้นตอนวิธี F-NLMS ช่องสัญญาณที่ j

Adaptive Filter :	\times / \div	$+ / -$
1 DFT/ (1 IDFT)	$10 \log_2 L + 10$	$10 - 5/L$
$\mathbf{E}(k) = \mathbf{D}(k) - \mathbf{W}(k)\mathbf{X}(k)$	8	8
$\mathbf{W}(k+1) = \mathbf{W}(k) + \frac{\mathbf{M}(k)}{\varepsilon + \ \mathbf{X}(k)\ ^2} \mathbf{X}^H(k)\mathbf{E}(k)$	$24+2\text{div}$	8

ตารางที่ 2.3 ความซับซ้อนทางการคำนวณของขั้นตอนวิธี AP ช่องสัญญาณที่ j

Adaptive Filter :	\times / \div	$+ / -$
$e(n) = d(n) - \mathbf{X}^T(n)\mathbf{w}(n)$	pL	pL
$\boldsymbol{\varepsilon}(n) = [\mathbf{X}^T(n)\mathbf{X}(n) + \delta \mathbf{I}]^{-1} e(n)$	$p^2(L+1) + \text{Inv}(\mathbf{M}_{p \times p})$	p^2L
$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \mathbf{X}(n)\boldsymbol{\varepsilon}(n)$	$pL+1$	pL

ตารางที่ 2.4 ความซับซ้อนทางการคำนวณของขั้นตอนวิธี FLS ช่องสัญญาณที่ j

available at time: n	adaptive filter: $\mathbf{w}(n)$, forward prediction: $\mathbf{A}(n)$, backward prediction: $\mathbf{B}(n)$, data vector: $\mathbf{X}(n)$, adaptation gain: $\mathbf{G}(n)$, prediction error energy: $E_a(n)$, forgetting factor: W	
new data at time: $n+1$	input signal: $x(n+1)$, microphone signal: $d(n+1)$	
Adaptation Gain Updating:		
	\times / \div	$+ / -$
$e_a(n+1) = x(n+1) - \mathbf{A}^T(n)\mathbf{X}(n)$	L	L
$\mathbf{A}(n+1) = \mathbf{A}(n) + \mathbf{G}(n)e_a(n+1)/\alpha(n)$	$L+1\text{div}$	L
$E_a(n+1) = (E_a(n) + e_a(n+1)e_a(n+1)/\alpha(n))W$	$2+1\text{div}$	1
$\mathbf{G}_1(n+1) = \begin{bmatrix} 0 \\ \mathbf{G}(n) \end{bmatrix} + \frac{e_a(n+1)}{E_a(n+1)} \begin{bmatrix} 1 \\ -\mathbf{A}(n+1) \end{bmatrix} = \begin{bmatrix} \mathbf{M}(n+1) \\ \mathbf{m}(n+1) \end{bmatrix}$	$L+1+1\text{div}$	$L+1$
$e_b(n+1) = x(n+1-L) - \mathbf{B}^T(n)\mathbf{X}(n+1)$	L	L
$\mathbf{G}(n+1) = \mathbf{M}(n+1) + \mathbf{m}(n+1)\mathbf{B}(n)$	L	L
$\alpha_1(n+1) = \alpha(n) + e_a(n+1)e_a(n+1)/E_a(n+1)$	$1+1\text{div}$	1
$\alpha(n+1) = \alpha_1(n+1) - \mathbf{m}(n+1)e_b(n+1)$	1	1
$E_b(n+1) = (E_b(n) + e_b(n+1)e_b(n+1)/\alpha(n+1))W$	$2+1\text{div}$	1
$\mathbf{B}(n+1) = \mathbf{B}(n) + \mathbf{G}(n+1)e_b(n+1)/\alpha(n+1)$	$L+1\text{div}$	L
Adaptive Filter:		
$e(n+1) = d(n+1) - \mathbf{w}^T(n)\mathbf{X}(n+1)$	L	L
$\mathbf{w}(n+1) = \mathbf{w}(n) + \mathbf{G}(n+1)e(n+1)$	L	$L+1$

ในวิทยานิพนธ์ฉบับนี้จะพิจารณาถึงข้อดีข้อเสียของแต่ละขั้นตอนวิธีในด้านประสิทธิภาพของการทำงานและความซับซ้อนทางการคำนวณ และเลือกใช้ขั้นตอนวิธีมากกว่าหนึ่งขั้นตอนวิธีขึ้นไปมาผสมผสานกันในการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวเพื่อเพิ่มประสิทธิภาพ โดยมีหลักการเลือกใช้แต่ละขั้นตอนวิธี คือ ต้องเป็นขั้นตอนวิธีที่ประยุกต์ใช้ได้ง่าย และมีความซับซ้อนทางการคำนวณค่อนข้างต่ำ ดังนั้น ในบทที่ 4 จะศึกษาการนำโครงสร้างการกรองแถบย่อยมาใช้งานร่วมกับการประมวลผลในโดเมนความถี่ เพื่อเพิ่มอัตราการสุ่มเข้าและลดความซับซ้อนทางการคำนวณของระบบการตัดสัญญาณเสียงสะท้อนในโดเมนเวลา

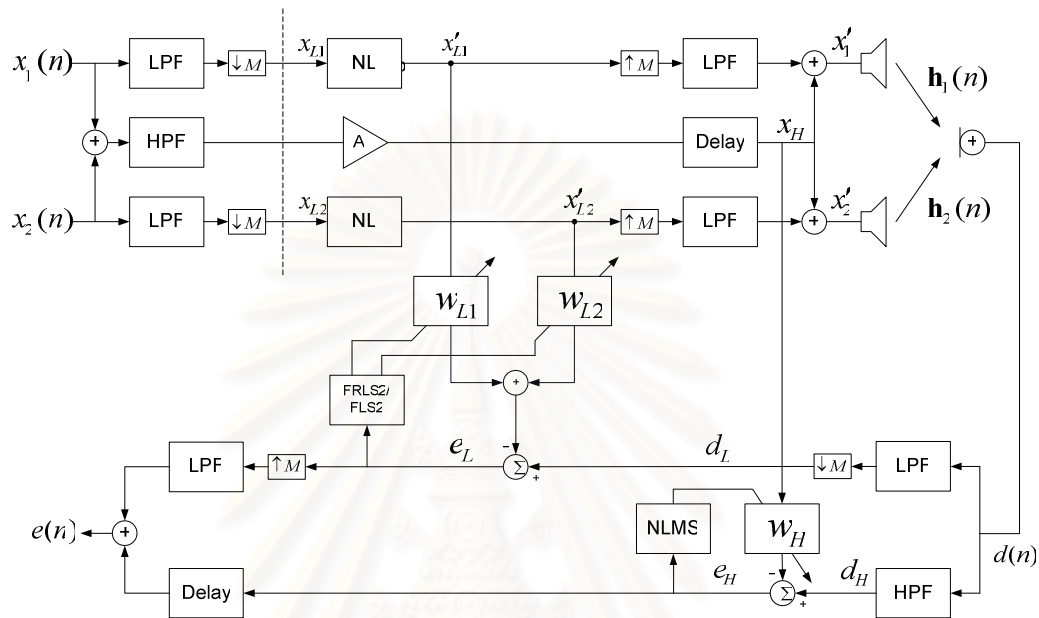
โครงสร้างไฮบริดระหว่างขั้นตอนวิธีแบบโมโนและขั้นตอนวิธีแบบสเตริโอ

ในบทที่ 3 นี้จะกล่าวถึงหลักการทำงานของโครงสร้างไฮบริด (Hybrid Structure) ระหว่างขั้นตอนวิธีแบบโมโนและขั้นตอนวิธีแบบสเตริโอ สำหรับระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตริโอใน [18] และนำเสนอโครงสร้างไฮบริดใหม่เพื่อปรับปรุงสมรรถนะการตัดสัญญาณเสียงสะท้อนในด้านอัตราการลู่เข้าของวงจรรองแบบปรับตัวให้เร็วยิ่งขึ้นกว่าระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตริโอใน [18] โดยที่ความซับซ้อนทางการคำนวณยังต่ำกว่าของระบบดังกล่าว นอกจากนี้ จะเปรียบเทียบสมรรถนะของระบบการตัดสัญญาณเสียงสะท้อนที่มีโครงสร้างไฮบริดทั้งสองแบบกับระบบปกติที่มีการทำงานในโดเมนเวลาและโดเมนความถี่

3.1 โครงสร้างไฮบริดระหว่างขั้นตอนวิธีแบบโมโนและขั้นตอนวิธีแบบสเตริโอ

โครงสร้างไฮบริดระหว่างขั้นตอนวิธีแบบโมโนและขั้นตอนวิธีแบบสเตริโอเป็นการประยุกต์ใช้โครงสร้างการกรองแถบย่อยร่วมกับคลังวงจรรองไม่เอกรูปในระบบการตัดสัญญาณเสียงสะท้อนที่มีหนึ่งช่องสัญญาณ (AEC) และระบบการตัดสัญญาณเสียงสะท้อนแบบสเตริโอ (SAEC) มีหลักการทำงานเริ่มจากการใช้คลังวงจรรองไม่เอกรูปทำการแบ่งข้อมูลของสัญญาณเข้าในแต่ละช่องสัญญาณออกเป็นหลายแถบความถี่ย่อย แล้วจึงประมวลผลสัญญาณในแต่ละแถบความถี่ย่อยด้วยวงจรรองแบบปรับตัวที่มีขั้นตอนวิธีที่เหมาะสมกับลักษณะของสัญญาณในแต่ละแถบความถี่ย่อยนั้นๆ การเลือกใช้ขั้นตอนวิธีเพื่อการประมวลผลสัญญาณในแต่ละแถบความถี่ย่อยจะพิจารณาถึงเหมาะสมกับการประมวลผลสัญญาณเสียงพูด ทั้งนี้ จากการศึกษาคุณลักษณะของสัญญาณเสียงพูด พบว่า ความหนาแน่นของสเปกตรัมกำลัง (Power Spectrum Density, PSD) ของสัญญาณเสียงในช่วงความถี่ต่ำ (300-3400 Hz) จะมีค่ามากกว่าความหนาแน่นของสเปกตรัมกำลังในช่วงความถี่ที่สูงขึ้น และจากหลักการของจิตวิทยาทางเสียง (Psychoacoustic) ซึ่งพลังงานของสัญญาณเสียงเชิงสเตริโอจะอยู่ในช่วงความถี่ที่ต่ำกว่า 1 kHz ดังนั้น การเลือกใช้ขั้นตอนวิธีที่ให้อัตราการลู่เข้าสู่ค่าตอบเร็ว เช่น FLS, RLS เป็นต้น เพื่อการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวที่ใช้กับสัญญาณข้อมูลในแต่ละแถบความถี่ต่ำ จะสามารถเพิ่มอัตราการลู่เข้าของระบบการตัดสัญญาณเสียงสะท้อน หากแต่ว่าความซับซ้อนทางการคำนวณจะมากกว่าการใช้ขั้นตอนวิธีที่มีความซับซ้อนทางการคำนวณต่ำ เช่น ขั้นตอนวิธี NLMS เป็นต้น สำหรับการประมวลผลสัญญาณในแต่ละแถบความถี่ที่สูงขึ้นซึ่งมีข้อมูลของ

สัญญาณเสียงอยู่น้อยกว่า จะเลือกใช้ขั้นตอนวิธีที่มีความซับซ้อนทางการคำนวณที่ต่ำลงเพื่อลดความซับซ้อนทางการคำนวณของระบบโดยรวม แต่จะทำให้อัตราการลู่เข้าของการประมวลผลในแถบความถี่เหล่านี้ลดต่ำลงด้วย



รูปที่ 3.1 บล็อกไดอะแกรมแสดงโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS ใน [18]

โครงสร้างไฮบริดระหว่าง FRLS2 และ NLMS ใน [18] ได้อ้างอิงจากหลักการของจิตวิทยาทางเสียง ดังนั้น บทความดังกล่าวจึงได้นำขั้นตอนวิธี Fast Recursive Least Squares (FRLS2) [2] ที่ให้อัตราการลู่เข้าสู่ค่าตอบที่เร็วมาใช้ปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวในแถบความถี่ตั้งแต่ 0-900 Hz ของทั้งสองช่องสัญญาณ เมื่อความถี่การซักรั่วอย่างของสัญญาณเสียงเข้าที่ถี่ใน [18] เท่ากับ 16 kHz นอกจากนี้ การใช้ตัวประกอบเดซิเมชัน M ยังสามารถลดความซับซ้อนทางการคำนวณของแถบความถี่ย่อยได้ด้วย ทั้งนี้ ในบทความที่ [18] ได้ใช้การแปลงไม่เชิงเส้น (Non-linear Transformation, NL) ในการลดสหสัมพันธ์ (Correlation) ของสัญญาณเสียงเชิงสเตอริโอ เพื่อปรับปรุงการลู่เข้าสู่ค่าตอบของวงจรรองแบบปรับตัวในระบบการตัดสัญญาณเสียงสะท้อนแบบสเตอริโอให้ลู่เข้าสู่ค่าตอบอย่างถูกต้องมากขึ้น ส่วนข้อมูลของสัญญาณเสียงในแถบความถี่ตั้งแต่ 900-8000 Hz ซึ่งมีพลังงานของสัญญาณเสียงเชิงสเตอริโออยู่น้อยกว่าในช่วงความถี่ 0-900 Hz จะทำการประมวลผลแบบช่องสัญญาณเดียว โดยทำการรวมสัญญาณทั้งสองช่องสัญญาณเป็นช่องสัญญาณเดียวและใช้ระบบการตัดสัญญาณเสียงสะท้อนแบบช่องสัญญาณเดียวในการตัดสัญญาณเสียงสะท้อนแทนระบบการตัดสัญญาณเสียงสะท้อนแบบสเตอริโอ จะสามารถลดความซับซ้อนทางการคำนวณของระบบโดยรวมได้ โดยใน [18] ได้

เลือกใช้ขั้นตอนวิธี NLMS ในการประมวลผลสัญญาณในช่วงความถี่สูงดังกล่าว โดยที่บล็อกไดอะแกรมของโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS ใน [18] สามารถแสดงได้ ดังรูปที่ 3.1

3.2 ความซับซ้อนทางการคำนวณของโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS ใน [18]

ความซับซ้อนทางการคำนวณของโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS ใน [18] สามารถแบ่งได้เป็น 2 ส่วน คือ ความซับซ้อนทางการคำนวณจากการประมวลผลสัญญาณในแถบความถี่ต่ำ (0-900 Hz) ด้วยขั้นตอนวิธี FRLS2 ในวิทยานิพนธ์นี้จะเลือกใช้ขั้นตอนวิธี FLS2 [15] ในการประมวลผลสัญญาณแทนขั้นตอนวิธี FRLS2 สำหรับส่วนที่สอง คือ ความซับซ้อนทางการคำนวณจากการประมวลผลสัญญาณในแถบความถี่สูง (900-8000 kHz) ด้วยขั้นตอนวิธี NLMS

กำหนดให้ M คือ ตัวประกอบเดซิเมตร
 L คือ จำนวนสัมประสิทธิ์ของวงจรกรองแบบปรับตัว
 N_{tap} คือ จำนวนสัมประสิทธิ์ของวงจรกรองวิเคราะห์และวงจรกรองสังเคราะห์

โดยที่การคำนวณความซับซ้อนของแต่ละขั้นตอนวิธีจะคำนวณจากค่าการคูณจำนวนจริง (Real Multiplications, RMPs) และค่าการบวกและลบจำนวนจริง (Real Additions and Subtractions, RASs) ต่อสัญญาณเข้าหนึ่งตัวอย่าง

การกรองแถบย่อยในแถบความถี่ต่ำ (0-900 Hz) จะต้องการ $4N_{\text{tap}}$ RMPs และ $4(N_{\text{tap}} - 1)$ RASs สำหรับการวิเคราะห์และสังเคราะห์สัญญาณเข้าแต่ละตัว ($x_i(n)$, $i=1,2$) ส่วนการวิเคราะห์สัญญาณอ้างอิง $d(n)$ และการสังเคราะห์สัญญาณผิดพลาด $e(n)$ จะต้องการอย่างละ N_{tap} RMPs และ $N_{\text{tap}} - 1$ RASs ดังนั้น การคำนวณสำหรับวิเคราะห์และสังเคราะห์สัญญาณที่ต้องการรวมทั้งสิ้นคิดเป็น $6N_{\text{tap}}$ RMPs และ $6(N_{\text{tap}} - 1)$ RASs ต่อสัญญาณเข้าหนึ่งตัวอย่าง

สำหรับขั้นตอนวิธี FLS2 ที่ใช้ประมวลผลในแถบความถี่นี้จะได้จากสมการดังตารางที่ 2.4 ในบทที่ 2 โดยคิดเป็นสองช่องสัญญาณและจำนวนสัมประสิทธิ์ในแถบความถี่ย่อยจะลดลงด้วยตัวประกอบเดซิเมตรเหลือเพียง $\frac{L}{M}$ ดังนั้น ความซับซ้อนทางการคำนวณของขั้นตอนวิธี FLS2 ในแถบความถี่ย่อยจะเป็น $2 \times \left[8 \left(\frac{L}{M} \right) + 5 + 6 \text{div} \right]$ RMPs และ $2 \times \left[8 \left(\frac{L}{M} \right) + 6 \right]$ RASs ต่อสัญญาณเข้าหนึ่งตัวอย่าง

สำหรับการแปลง NL ที่ใช้ในโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS ใน [18] ที่ถูกจำลองขึ้นในวิทยานิพนธ์นี้ จะใช้ตัวทำกระแสตรงแบบครึ่งคลื่น (Half Wave Rectifier) ซึ่งมีความซับซ้อนทางการคำนวณเท่ากับ 2 RASs ต่อสัญญาณเข้าหนึ่งตัวอย่าง

ดังนั้น ความซับซ้อนทางการคำนวณทั้งหมดเมื่อใช้ขั้นตอนวิธี FLS2 ในแถบความถี่ต่ำ เมื่อพิจารณาการคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างจะคิดเป็น

$$6N_{\text{tap}} + 2 \times \left[8 \left(\frac{L}{M} \right) + 5 + 6\text{div} \right] \quad (3.1)$$

และการบวกจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างจะคิดเป็น

$$6(N_{\text{tap}} - 1) + 2 \times \left[8 \left(\frac{L}{M} \right) + 6 \right] + 2 \quad (3.2)$$

สำหรับขั้นตอนวิธี NLMS ในแถบความถี่สูงนั้น จะต้องการการวิเคราะห์สัญญาณเสียงเข้า และสัญญาณอ้างอิงเท่ากับ $2N_{\text{tap}}$ RMPs และ $2(N_{\text{tap}} - 1)$ RASs ต่อสัญญาณเข้าหนึ่งตัวอย่าง และความซับซ้อนทางการคำนวณด้วยขั้นตอนวิธี NLMS ในแถบความถี่สูงนั้นจะได้จากสมการดังตารางที่ 2.1 ในบทที่ 2 โดยคิดในกรณีหนึ่งของสัญญาณ ดังนั้น ความซับซ้อนทางการคำนวณทั้งหมดเมื่อใช้ขั้นตอนวิธี NLMS ในแถบความถี่สูง เมื่อพิจารณาการคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างจะคิดเป็น

$$2N_{\text{tap}} + [2L + 3 + 1\text{div}] \quad (3.3)$$

และการบวกจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างจะคิดเป็น

$$2(N_{\text{tap}} - 1) + [2L + 3] \quad (3.4)$$

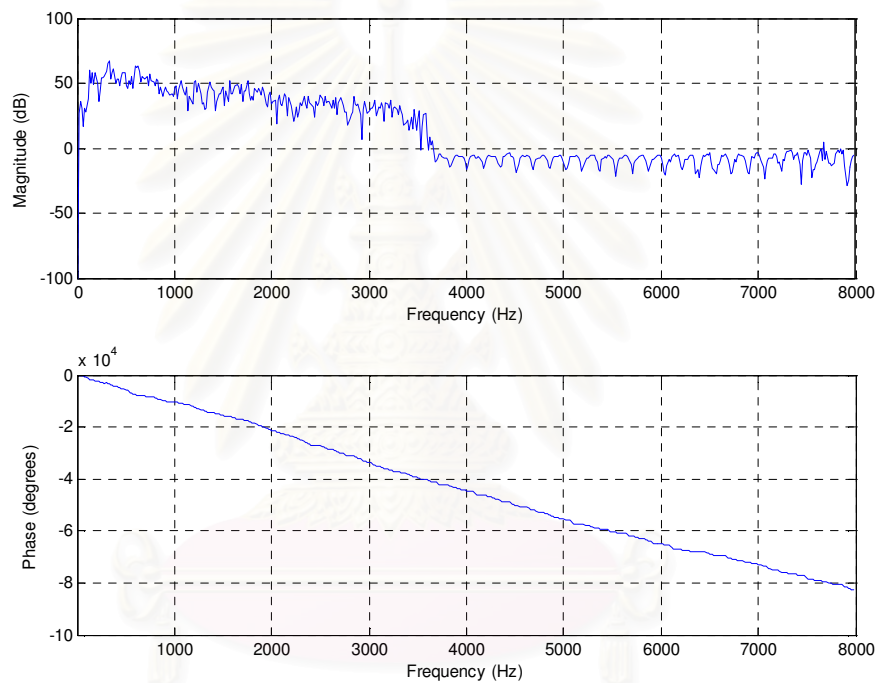
ดังนั้น ความซับซ้อนทางการคำนวณรวมทั้งหมดของโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS ใน [18] หาได้จากผลรวมของการประมวลผลสัญญาณในแถบความถี่ต่ำและแถบความถี่สูง คิดเป็นการคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างได้

$$8N_{\text{tap}} + 2L + 16\left(\frac{L}{M}\right) + 13 + 13\text{div} \quad (3.5)$$

และการบวกจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างได้เป็น

$$8(N_{\text{tap}} - 1) + 2L + 16\left(\frac{L}{M}\right) + 17 \quad (3.6)$$

3.3 โครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอในวิทยานิพนธ์นี้



รูปที่ 3.2 ผลตอบสนองเชิงความถี่ของสัญญาณเสียงเข้าหนึ่งช่องสัญญาณ

ที่ใช้ในวิทยานิพนธ์ฉบับนี้ เมื่อความถี่การซัดตัวอย่างเท่ากับ 16 kHz

เมื่อทำการศึกษาผลตอบสนองเชิงความถี่ของสัญญาณเสียงเข้าหนึ่งช่องสัญญาณที่ใช้ในวิทยานิพนธ์ฉบับนี้ เมื่อความถี่การซัดตัวอย่างเท่ากับ 16 kHz พบว่า ความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดในช่วง 0-1 kHz จะมีขนาดมากกว่าในช่วงความถี่อื่นๆ และมีขนาดลดลงมนี้น้อยมากในช่วงความถี่ที่สูงขึ้นจนถึง 4 kHz ดังแสดงในรูปที่ 3.2 ดังนั้น ในวิทยานิพนธ์ฉบับนี้จึงนำเสนอโครงสร้างไฮบริดแบบใหม่เพื่อใช้ในระบบการตัดสัญญาณเสียงสะท้อนแบบสเตอริโอ โดยเลือกใช้ขั้นตอนวิธีที่เหมาะสมกับความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดในแต่ละแถบความถี่ย่อย เพื่อเพิ่มสมรรถนะของระบบการตัดสัญญาณเสียง

สะท้อนแบบสเตริโอให้มีอัตราการลู่เข้าสู่คำตอบที่รวดเร็วยิ่งขึ้น โดยที่ความซับซ้อนทางการคำนวณของระบบที่นำเสนอจะต้องต่ำลง เมื่อเปรียบเทียบกับระบบการตัดสัญญาณเสียงสะท้อนแบบสเตริโอแบบปกติที่ทำงานในโดเมนเวลา และที่ทำงานในโดเมนความถี่ รวมทั้งเมื่อเปรียบเทียบกับระบบการตัดสัญญาณเสียงสะท้อนแบบสเตริโอ ที่มีโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS ใน [18] ด้วย ทั้งนี้ หลักการในการแบ่งแถบความถี่ย่อย และขั้นตอนวิธีที่แตกต่างกับโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS ใน [18] จะถูกนำเสนอขึ้นในหัวข้อถัดไป

โครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอในวิทยานิพนธ์ฉบับนี้จะทำการประมวลผลสัญญาณเสียงพูด ซึ่งเป็นสัญญาณเข้าของระบบการตัดสัญญาณเสียงสะท้อนแบบสเตริโอ โดยใช้วงจรกรองวิเคราะห้ในการแบ่งความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดออกเป็น 3 ส่วนด้วยกัน ได้แก่ ส่วนที่หนึ่ง คือ แถบความถี่ตั้งแต่ 0-1 KHz ซึ่งเป็นแถบความถี่ที่มีความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดมากที่สุด ส่วนที่สอง คือ แถบความถี่ตั้งแต่ 1-4 KHz ถึงแม้ว่าความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดในแถบความถี่นี้จะมีค่าน้อยกว่าในส่วนที่หนึ่งแต่ก็ยังมีค่ามากอยู่ และส่วนที่สาม คือ แถบความถี่ตั้งแต่ 4-8 KHz ในแถบความถี่นี้จะมีความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดอยู่น้อยมาก โดยที่ผลตอบสนองเชิงความถี่ของวงจรกรองวิเคราะห้ที่มีโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ได้แสดงดังรูปที่ 3.3 เมื่อกำหนดให้ $G_i(n)$ เมื่อ $i=0, \dots, M-1$ และ $G_H(n)$ เป็นวงจรกรองวิเคราะห้ในแต่ละแถบความถี่ย่อย สำหรับรายละเอียดของการออกแบบวงจรกรองวิเคราะห้จะเป็นดังต่อไปนี้

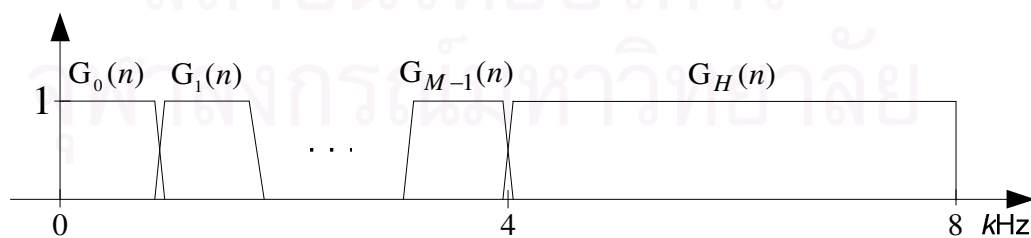
ส่วนที่หนึ่ง คือ แถบความถี่ตั้งแต่ 0-1 KHz ซึ่งเป็นแถบความถี่ที่มีความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดมากที่สุดเมื่อพิจารณาจากค่าผลตอบสนองเชิงความถี่ของสัญญาณเสียงเข้าดังรูปที่ 3.2 ดังนั้น ในแถบความถี่ย่อยนี้จึงจำเป็นต้องใช้ขั้นตอนวิธีที่ให้อัตราการลู่เข้าสู่คำตอบเร็วที่สุดในการประมวลผลสัญญาณ โดยจะเลือกใช้ขั้นตอนวิธี AP2 เนื่องจากมีอัตราการลู่เข้าสู่คำตอบที่เร็วกว่าขั้นตอนวิธีอื่นที่พิจารณาในวิทยานิพนธ์นี้ อีกทั้งขั้นตอนวิธี AP2 ยังมีข้อได้เปรียบกว่าขั้นตอนวิธีอื่น คือ สามารถปรับเพิ่มอัตราการลู่เข้าสู่คำตอบให้เร็วขึ้นโดยการเพิ่มอันดับของการฉาย (Projection Order, p) ให้มากขึ้น และสามารถปรับลดอันดับของการฉาย ในกรณีที่ต้องการลดความซับซ้อนทางการคำนวณของระบบการตัดสัญญาณเสียงสะท้อนแบบสเตริโอ

ส่วนที่สอง คือ แถบความถี่ตั้งแต่ 1-4 KHz ถึงแม้ว่าความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดในแถบความถี่นี้จะมีค่าน้อยกว่าในส่วนที่หนึ่งแต่ก็ยังมีค่ามาก ดังนั้น ในช่วงแถบความถี่ย่อยนี้ยังสามารถปรับปรุงอัตราการลู่เข้าสู่ของวงจรกรองแบบปรับตัวให้เร็วกว่าของโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS ใน [18] ที่ประมวลผลในแถบความถี่ย่อยนี้แบบ

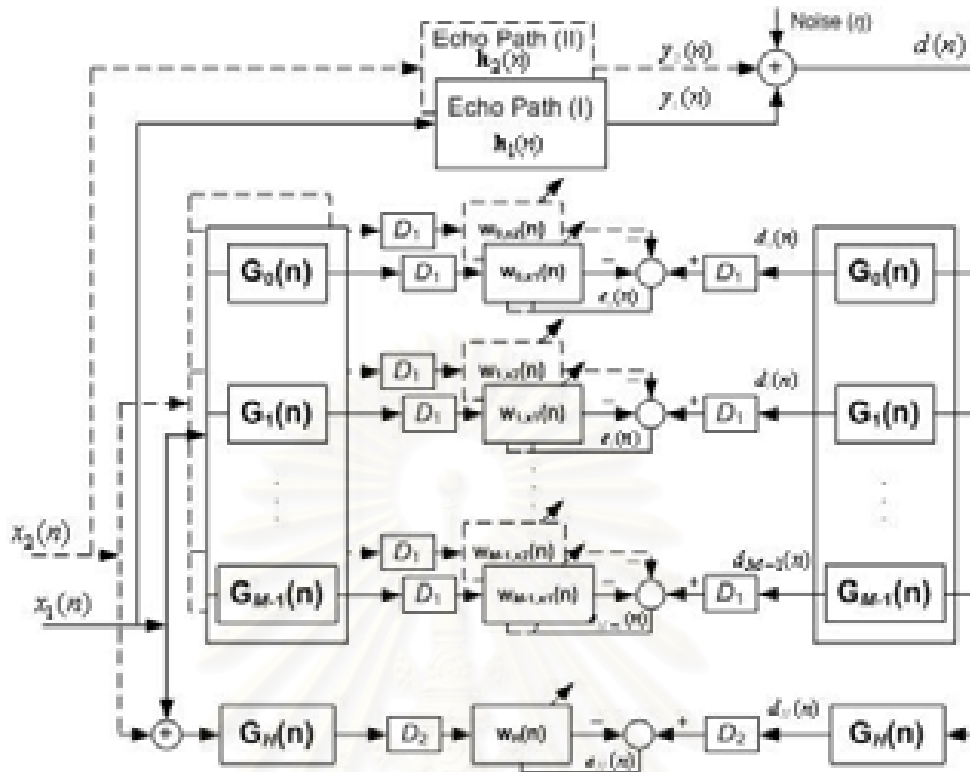
ช่องสัญญาณเดียว วิทยานิพนธ์นี้จึงเลือกใช้ขั้นตอนวิธีแบบสองช่องสัญญาณในการประมวลผลสัญญาณในแถบความถี่ย่อยนี้ เพื่อเพิ่มอัตราการสุ่มเข้าของวงจรรองแบบปรับตัวให้เร็วยิ่งขึ้น โดยจะพิจารณาเลือกใช้ขั้นตอนวิธี F-NLMS2 ในการประมวลผล อีกทั้ง ขั้นตอนวิธี F-NLMS2 ซึ่งประมวลผลสัญญาณในโดเมนความถี่จะมีความซับซ้อนทางการคำนวณต่ำกว่าขั้นตอนวิธี NLMS2 ที่ประมวลผลในโดเมนเวลา โดยการแบ่งแถบความถี่ในส่วนที่สองนี้ สามารถแบ่งแถบความถี่ย่อยได้มากกว่าหนึ่งแถบความถี่ ดังตารางที่ 2.1 และ 2.2 ในบทที่ 2 โดยใช้คลังวงจรรองวิเคราะห์ $G_i(n)$ เมื่อ $i=1, \dots, M-1$ ดังนั้น ความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดในแถบความถี่ตั้งแต่ 1-4 KHz จะสามารถแบ่งได้เท่ากับ $M-1$ แถบความถี่ย่อย โดยที่การเลือกจำนวนแถบความถี่ย่อย M จะได้ทำการทดลองในบทที่ 4 ต่อไป เพื่อให้ได้สมรรถนะของการประมวลผลของโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่ดีที่สุด และมีความซับซ้อนทางการคำนวณที่ต่ำ

ส่วนที่สาม คือ แถบความถี่ตั้งแต่ 4-8 KHz ในแถบความถี่นี้ จะมีความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดอยู่น้อยมาก จึงสามารถรวมสเปกตรัมกำลังของทั้งสองช่องสัญญาณเป็นช่องสัญญาณเดียว แล้วประมวลผลด้วยขั้นตอนวิธีเพียงหนึ่งช่องสัญญาณเพื่อลดความซับซ้อนทางการคำนวณของระบบโดยรวม โดยไม่ทำให้เกิดความผิดพลาดในการประมาณวิถีสะท้อนมากนัก ขั้นตอนวิธีที่เลือกใช้ในแถบความถี่นี้ คือ ขั้นตอนวิธี F-NLMS เนื่องจากมีความซับซ้อนทางการคำนวณที่ต่ำกว่าขั้นตอนวิธี NLMS

ดังนั้น ด้วยการแบ่งสเปกตรัมของสัญญาณเสียงพูดออกเป็นสามส่วนด้วยคลังวงจรรองวิเคราะห์ และประมวลผลสัญญาณในแต่ละแถบความถี่ย่อยดังกล่าว บล็อกไดอะแกรมแสดงโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอ สามารถแสดงได้ดังรูปที่ 3.4 โดยที่ D_1 และ D_2 เป็นตัวประกอบเดซิเบลขึ้น ทั้งนี้ โดยทั่วไปแล้ว D_1 จะมีค่าเท่ากับ $2M$ และ D_2 จะมีค่าเท่ากับ 2



รูปที่ 3.3 ผลตอบสนองเชิงความถี่ของวงจรรองวิเคราะห์ที่มีโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอในวิทยานิพนธ์นี้



รูปที่ 3.4 บล็อกไดอะแกรมแสดงโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอในวิทยานิพนธ์นี้

3.4 ความซับซ้อนทางการคำนวณของโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอในวิทยานิพนธ์นี้

ความซับซ้อนทางการคำนวณของโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอสามารถแบ่งการคำนวณออกได้เป็น 3 ส่วนด้วยกัน โดยกำหนดให้ตัวประกอบเดซิเมชันที่ทำให้เกิดอัตราการซีกตัวอย่างแบบวิกฤต (Critical Sampling Rate) ดังนั้น $D_1 = 2M$ และ $D_2 = 2$

ส่วนที่หนึ่ง คือ แถบความถี่ตั้งแต่ 1-4 kHz การกรองแถบย่อยจะต้องการการคำนวณสำหรับการวิเคราะห์สัญญาณเข้าแต่ละตัว ($x_i(n), i = 1, 2$) การวิเคราะห์สัญญาณอ้างอิง $d(n)$ และการสังเคราะห์สัญญาณผิดพลาด $e(n)$ คิดเป็น $4N_{tap}$ RMPs และ $4(N_{tap} - 1)$ RASs ต่อสัญญาณเข้าหนึ่งตัวอย่าง

สำหรับขั้นตอนวิธี AP2 ที่ใช้ประมวลผลในแถบความถี่นี้จะได้จากสมการดังตารางที่ 2.3 ในบทที่ 2 โดยความซับซ้อนทางการคำนวณจะคิดเป็นสองช่องสัญญาณและจำนวนสัมประสิทธิ์ในแถบความถี่ย่อยจะลดลงด้วยตัวประกอบเดซิเมชันเหลือเพียง $\frac{L}{2M}$ ดังนั้น ความซับซ้อนทางการคำนวณของขั้นตอนวิธี AP2 ในแถบความถี่ย่อยนี้จะเท่ากับ

$2 \times \left[p^2 \left(\frac{L}{2M} \right) + 2p \left(\frac{L}{2M} \right) + p^2 + 1 + \text{Inv}(M_{\text{pxp}}) \right]$ RMPs และ $2 \times \left[p^2 \left(\frac{L}{2M} \right) + 2p \left(\frac{L}{2M} \right) \right]$ RASs ต่อสัญญาณเข้าหนึ่งตัวอย่าง

ดังนั้น ความซับซ้อนทางการคำนวณทั้งหมดเมื่อใช้ขั้นตอนวิธี AP2 ในแถบความถี่นี้ เมื่อพิจารณาการคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างจะคิดเป็น

$$4N_{\text{tap}} + \left[p^2 \left(\frac{L}{M} \right) + 2p \left(\frac{L}{M} \right) + 2p^2 + 2 + 2\text{Inv}(M_{\text{pxp}}) \right] \quad (3.7)$$

และการบวกจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างคิดเป็น

$$4(N_{\text{tap}} - 1) + \left[p^2 \left(\frac{L}{M} \right) + 2p \left(\frac{L}{M} \right) \right] \quad (3.8)$$

ส่วนที่สอง คือ แถบความถี่ตั้งแต่ 1-4 kHz ในส่วนที่สองนี้จะทำการแบ่งความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงออกเป็นแถบความถี่ย่อยตั้งแต่แถบความถี่ที่ 1 ถึง $M-1$ รวมเป็นจำนวน $M-1$ แถบความถี่ย่อย โดยที่ในแต่ละแถบความถี่ย่อยจะต้องการ $4N_{\text{tap}}$ RMPs และ $4(N_{\text{tap}} - 1)$ RASs ต่อสัญญาณเข้าหนึ่งตัวอย่าง สำหรับการวิเคราะห์สัญญาณเข้าแต่ละตัว ($x_i(n)$, $i=1,2$) การวิเคราะห์สัญญาณอ้างอิง $d(n)$ และการสังเคราะห์สัญญาณผิดพลาด $e(n)$

สำหรับขั้นตอนวิธี F-NLMS2 ที่ใช้ประมวลผลในแถบความถี่นี้จะได้จากสมการดังตารางที่ 2.2 ในบทที่ 2 ดังนั้น ความซับซ้อนทางการคำนวณของขั้นตอนวิธี F-NLMS2 ในแต่ละแถบความถี่ย่อยจะเท่ากับ $20 \log_2 \left(\frac{L}{2M} \right) + 148$ RMPs และ $20 \left(\frac{L}{2M} \right) - 2$ RASs ต่อสัญญาณเข้าหนึ่งตัวอย่าง ดังนั้น ความซับซ้อนทางการคำนวณทั้งหมดเมื่อใช้ขั้นตอนวิธี F-NLMS2 ในแถบความถี่ย่อยนี้ เมื่อพิจารณาการคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างจะคิดเป็น

$$(M-1) \left[4N_{\text{tap}} + 20 \log_2 \left(\frac{L}{2M} \right) + 148 \right] \quad (3.9)$$

และการบวกจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างคิดเป็น

$$(M-1) \left[4(N_{\text{tap}} - 1) + 20 \left(\frac{L}{2M} \right) - 2 \right] \quad (3.10)$$

ส่วนที่สาม คือ แถบความถี่ตั้งแต่ 4-8 kHz ในส่วนนี้จะต้องการ $3N_{\text{tap}}$ RMPs และ $3(N_{\text{tap}} - 1)$ RASs ต่อสัญญาณเข้าหนึ่งตัวอย่าง สำหรับการวิเคราะห์สัญญาณเข้าหนึ่งช่องสัญญาณ การวิเคราะห์สัญญาณอ้างอิง $d(n)$ และการสังเคราะห์สัญญาณผิดพลาด $e(n)$

สำหรับขั้นตอนวิธี F-NLMS ที่ใช้ประมวลผลในแถบความถี่นี้จะได้จากสมการดังตารางที่ 2.2 ในบทที่ 2 เช่นกัน โดยจะคำนวณความซับซ้อนในการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวเพียงหนึ่งช่องสัญญาณ ดังนั้น ความซับซ้อนทางการคำนวณของขั้นตอนวิธี F-NLMS ในแถบความถี่ย่อยนี้จะเท่ากับ $10\log_2\left(\frac{L}{2M}\right) + 74$ RMPs และ $10\left(\frac{L}{2M}\right) - 1$ RASs ต่อสัญญาณเข้าหนึ่งตัวอย่าง ดังนั้น ความซับซ้อนทางการคำนวณทั้งหมดเมื่อใช้ขั้นตอนวิธี F-NLMS2 ในแถบความถี่นี้ เมื่อพิจารณาการคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างจะคิดเป็น

$$3N_{\text{tap}} + 10\log_2\left(\frac{L}{2M}\right) + 74 \quad (3.11)$$

และการบวกจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างคิดเป็น

$$3(N_{\text{tap}} - 1) + 10\left(\frac{L}{2M}\right) - 1 \quad (3.12)$$

ดังนั้น ความซับซ้อนทางการคำนวณรวมทั้งหมดของโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอ หากจากผลรวมของการประมวลผลสัญญาณทั้งสามส่วน คิดเป็นการคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างได้เท่ากับ

$$(4M + 3)N_{\text{tap}} + \left[p^2 + 2p\left(\frac{L}{M}\right) + 2p^2 + 2 + 2\text{Inv}(M_{\text{pxp}}) \right] + \left[(20M - 10)\log_2\left(\frac{L}{2M}\right) + (148M - 74) \right] \quad (3.13)$$

และความซับซ้อนทางการคำนวณรวมทั้งหมดในเชิงการบวกจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่าง จะเท่ากับ

$$(4M - 3)(N_{\text{tap}} - 1) + \left[p^2\left(\frac{L}{M}\right) + 2p\left(\frac{L}{M}\right) \right] + \left[(20M - 10)\left(\frac{L}{2M}\right) - 2M + 1 \right] \quad (3.14)$$

ทั้งนี้ การเปรียบเทียบความซับซ้อนทางการคำนวณและอัตราการใช้ค่าตอบของวงจรแบบปรับตัวเมื่อใช้โครงสร้างไฮบริดระหว่าง FRLS2 และ NLMS ใน [18] กับโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอในวิทยานิพนธ์นี้ เมื่อแทนค่าพารามิเตอร์ต่างๆ แล้ว จะได้นำเสนอในบทที่ 4 ต่อไป



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4

ผลการจำลองแบบและการวิเคราะห์ผล

บทที่ 4 นี้จะเป็นการจำลองแบบระบบ SAEC เพื่อเปรียบเทียบสมรรถนะของระบบการตัดสัญญาณเสียงสะท้อนทั้งด้านอัตราการรู้เข้าของวงจรกรองแบบปรับได้และความซับซ้อนทางการคำนวณ โดยขั้นตอนวิธีในโดเมนเวลาที่จะใช้ในการเปรียบเทียบสมรรถนะของระบบ ได้แก่ ขั้นตอนวิธี NLMS2, FLS2 และ AP2 สำหรับการเปรียบเทียบสมรรถนะของระบบระหว่างขั้นตอนวิธีในโดเมนเวลาและขั้นตอนวิธีในโดเมนความถี่ จะศึกษาประสิทธิภาพการทำงานของระบบเมื่อใช้ขั้นตอนวิธี NLMS2 และ F-NLMS2 จากนั้นจึงทำการจำลองแบบระบบที่มีโครงสร้างการกรองแถบย่อยร่วมกับการประมวลผลสัญญาณทั้งในโดเมนเวลาและโดเมนความถี่เพื่อเปรียบเทียบสมรรถนะของระบบเมื่อทำการประมวลผลสัญญาณในโดเมนเวลาตามปกติ หัวข้อสุดท้าย จะเป็น การเปรียบเทียบสมรรถนะของระบบการตัดสัญญาณเสียงสะท้อนที่มีโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18] กับโครงสร้างไฮบริดที่นำเสนอขึ้นในวิทยานิพนธ์นี้

4.1 แบบจำลองของระบบ SAEC

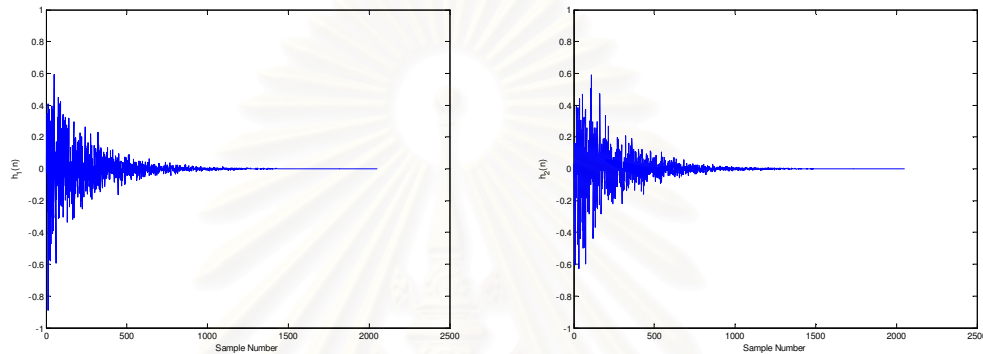
หัวข้อนี้จะแสดงค่าพารามิเตอร์ต่างๆ ของระบบ SAEC และสมการของตัววัดสมรรถนะของระบบ เริ่มด้วยการจำลองลักษณะวิธีสะท้อนทางเสียงจากลำโพงตัวที่ 1 และตัวที่ 2 ($\mathbf{h}_1(n)$ และ $\mathbf{h}_2(n)$) ไปที่ไมโครโฟนตัวที่ 1 ในห้องรับ โดยวิธีสะท้อนทางเสียง $\mathbf{h}_1(n)$ และ $\mathbf{h}_2(n)$ จะจำลองได้จากสัญญาณสุ่มปรกติ (Normal Random Signal) ที่มีการลดระดับแบบเลขชี้กำลัง (Exponential Decay) ให้มีจำนวนสัมประสิทธิ์ของวิธีสะท้อนทางเสียง (L) เท่ากับ 2,048 ค่า และสมมติให้วิธีสะท้อนทางเสียงเป็นเชิงเส้นและไม่แปรเปลี่ยนตามเวลา (Linear Time-invariant, LTI) กล่าวคือ $\mathbf{h}_1(n) = \mathbf{h}_1$ และ $\mathbf{h}_2(n) = \mathbf{h}_2$ ดังรูปที่ 4.1 จากนั้น กำหนดให้จำนวนสัมประสิทธิ์ของวงจรกรองแบบปรับได้ ($\mathbf{w}_1(n)$ และ $\mathbf{w}_2(n)$) มีจำนวน 2,048 ค่า เช่นกัน

สำหรับสัญญาณเข้าของระบบ ($\mathbf{x}_1(n)$ และ $\mathbf{x}_2(n)$) จะเป็นสัญญาณเสียงพูดแบบสเตอริโอที่มีอัตราการซีกตัวอย่าง $f_s = 16$ kHz และถูกนอร์มัลไลซ์ให้มีค่าเฉลี่ยเท่ากับศูนย์และความแปรปรวนเท่ากับหนึ่ง ดังแสดงในรูปที่ 4.2 และกำหนดให้สัญญาณรบกวนพื้นหลัง (Background Noise) ในห้องรับมีขนาด 30 dB SNR (Signal to Noise Ratio)

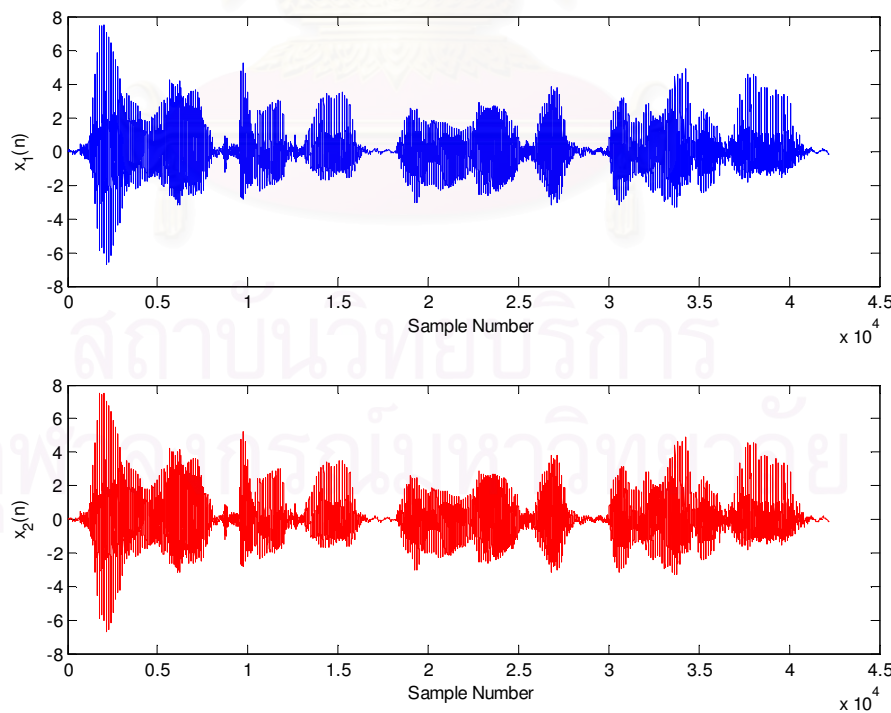
การประเมินสมรรถนะการทำงานของวงจรกรองแบบปรับได้เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธีในโดเมนเวลา ขั้นตอนวิธีในโดเมนความถี่ และเมื่อใช้โครงสร้างการกรองแถบย่อย

ร่วมกับการประมวลผลสัญญาณทั้งในโดเมนเวลาและโดเมนความถี่ จะได้จากการวัดค่า Weight Error Vector Norm (WEVN) ดังสมการที่ (4.1)

$$\text{WEVN}(n) = 10 \times \log_{10} \frac{\|\mathbf{h}_1 - \mathbf{w}_1(n)\|^2 + \|\mathbf{h}_2 - \mathbf{w}_2(n)\|^2}{\|\mathbf{h}_1\|^2 + \|\mathbf{h}_2\|^2} \quad (4.1)$$



รูปที่ 4.1 ลักษณะวิธีสะท้อนทางเสียงจำนวน 2,048 ค่าจากลำโพงตัวที่ 1 ไปที่ไมโครโฟนตัวที่ 1 ในห้องรับ (\mathbf{h}_1) และจากลำโพงตัวที่ 2 ไปที่ไมโครโฟนตัวที่ 1 ในห้องรับ (\mathbf{h}_2) ตามลำดับ



รูปที่ 4.2 สัญญาณเข้าของระบบที่เป็นสัญญาณเสียงพูดจากลำโพงตัวที่ 1 และตัวที่ 2 ตามลำดับ

จะเห็นว่า ถ้าวงจรรองแบบปรับตัวสามารถจำลองวิธีสะท้อนทางเสียงได้ใกล้เคียงมากขึ้น จะทำให้ค่า WEVN มีค่าลดต่ำลง ดังนั้น ระบบการตัดสัญญาณเสียงสะท้อนจึงทำงานได้อย่างมีประสิทธิภาพมากขึ้น

สำหรับการประเมินสมรรถนะการทำงานของโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18] กับโครงสร้างไฮบริดที่น่าเสนอ จะพิจารณาด้วยค่า Echo Return Loss Enhancement (ERLE) แทนค่า WEVN เนื่องจาก สัญญาณผิดพลาดของแต่ละแถบความถี่ย่อยสามารถนำมาสังเคราะห์เป็นสัญญาณผิดพลาด $e(n)$ ได้ โดยค่า ERLE จะทำการเปรียบเทียบความสัมพันธ์ระหว่างสัญญาณเข้าของไมโครโฟนในห้องรับ (สัญญาณเสียงสะท้อน) และสัญญาณออกของระบบ (สัญญาณผิดพลาดที่ถูกส่งกลับไปยังห้องส่ง) ดังสมการที่ (4.2)

$$ERLE(n) = 10 \times \log_{10} \left(\frac{\sum_{i=0}^{N_j} d^2(n-i)}{\sum_{i=0}^{N_j} e^2(n-i)} \right) \quad (4.2)$$

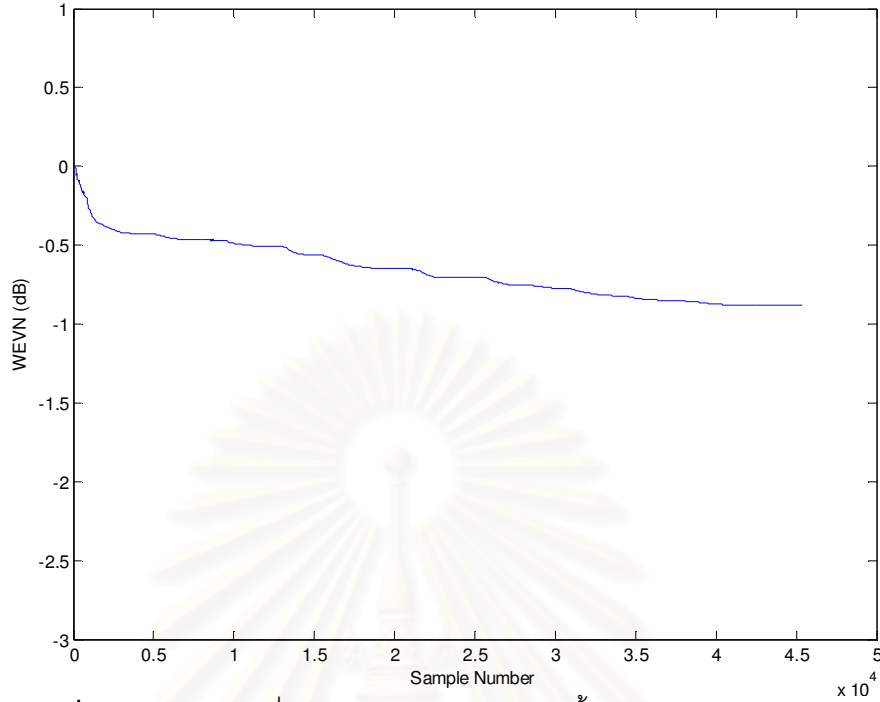
จะเห็นว่า ถ้าสัญญาณผิดพลาดที่ถูกส่งกลับไปยังห้องส่งมีขนาดลดลงเมื่อเทียบกับสัญญาณเสียงสะท้อนที่ไมโครโฟนในห้องรับ แสดงว่าสัญญาณเสียงสะท้อนที่ถูกส่งกลับไปยังห้องส่งมีขนาดลดลง หรือหมายความว่าระบบสามารถตัดสัญญาณเสียงสะท้อนได้อย่างมีประสิทธิภาพมากขึ้น โดยสังเกตจากค่า ERLE ที่มีค่าเพิ่มมากขึ้น

4.2 ผลการจำลองแบบระบบ SAEC ที่มีการประมวลผลสัญญาณในโดเมนเวลา

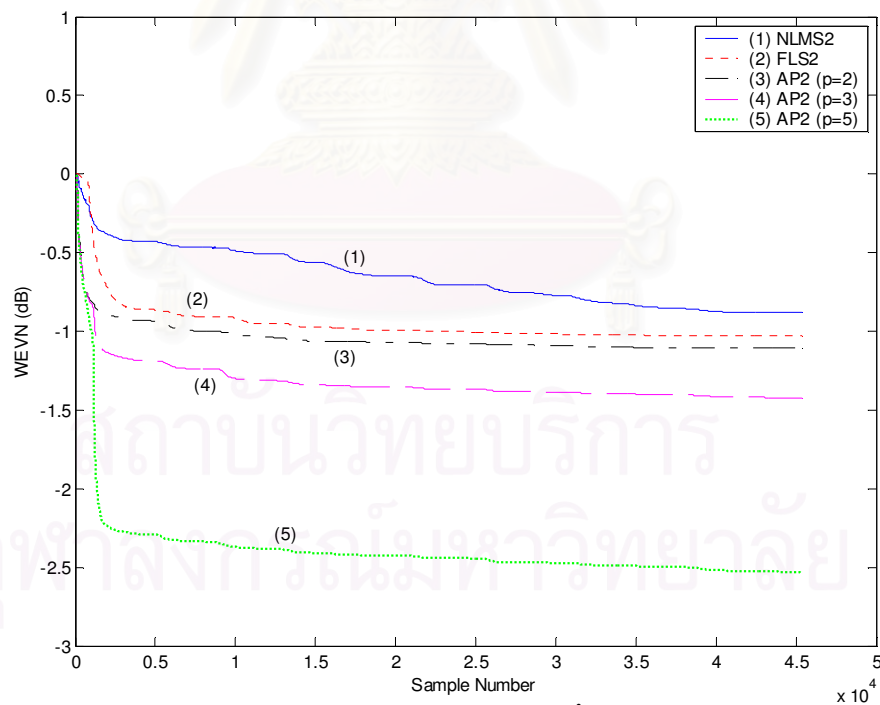
การจำลองระบบ SAEC ที่มีการประมวลผลสัญญาณในโดเมนเวลาเพื่อเปรียบเทียบสมรรถนะการตัดสัญญาณเสียงสะท้อน จะเลือกใช้ขั้นตอนวิธี NLMS2, FLS2 และ AP2 ในการประมวลผลสัญญาณ โดยจะเปรียบเทียบสมรรถนะของระบบทางด้านอัตราการลู่เข้าของวงจรรองแบบปรับตัวและความซับซ้อนทางการคำนวณ

ผลการจำลองแบบด้วยขั้นตอนวิธี NLMS2 ดังบล็อกไดอะแกรมรูปที่ 2.3 จากบทที่ 2 โดยเลือกค่าช่วงก้าว μ ที่ทำให้วงจรรองแบบปรับตัวมีอัตราการลู่เข้าที่เร็วที่สุด สามารถสังเกตค่า WEVN ดังแสดงในรูปที่ 4.3

จากนั้น จึงทำการจำลองแบบขั้นตอนวิธี FLS2 และขั้นตอนวิธี AP2 ที่อันดับ $p=2, 3$ และ 5 ตามลำดับ โดยการปรับเลือกพารามิเตอร์ที่ทำให้วงจรรองแบบปรับตัวของแต่ละขั้นตอนวิธีดังกล่าวมีอัตราการลู่เข้าที่เร็วที่สุด จะได้ค่า WEVN ดังแสดงในรูปที่ 4.4



รูปที่ 4.3 ค่า WEVN เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี NLMS2: $\mu = 0.5$



รูปที่ 4.4 ค่า WEVN เปรียบเทียบสมรรถนะระหว่าง (1) ขั้นตอนวิธี NLMS2: $\mu = 0.5$ กับ

(2) ขั้นตอนวิธี FLS2: $E_a(0) = E_b(0) = 1200$, $W = 1$ กับ

(3) - (5) ขั้นตอนวิธี AP2: $\mu = 0.5$, $p=2, 3$ และ 5 ตามลำดับ

ตารางที่ 4.1 ความซับซ้อนทางการคำนวณของขั้นตอนวิธีในโดเมนเวลา

Adaptive Filter Algorithm :	\times / \div	$+ / -$
a. NLMS2	8230	8198
b. FLS2	36904	36867
c. AP2 (p = 2)	32822	32768
d. AP2 (p = 3)	61570	61440

จากรูปที่ 4.4 แสดงให้เห็นว่าขั้นตอนวิธี FLS2 มีอัตราการลู่เข้าที่เร็วกว่าขั้นตอนวิธี NLMS2 แต่ความซับซ้อนทางการคำนวณของขั้นตอนวิธี FLS2 จะสูงกว่าขั้นตอนวิธี NLMS2 ดังตารางที่ 4.1 (b) และเมื่อเปรียบเทียบสมรรถนะของระบบกับขั้นตอนวิธี AP2 พบว่า ขั้นตอนวิธี AP2 อันดับ p=2 ให้อัตราการลู่เข้าที่เร็วกว่าขั้นตอนวิธี FLS2 และเมื่อเพิ่มอันดับ p ให้สูงขึ้น (p=3 และ p=5 ตามลำดับ) ขั้นตอนวิธี AP2 จะให้อัตราการลู่เข้าที่เร็วยิ่งขึ้น แต่ความซับซ้อนทางการคำนวณก็จะมีสูงขึ้น ดังตารางที่ 4.1 (c) และ 4.1 (d) ตามลำดับ โดยที่ความซับซ้อนทางการคำนวณของขั้นตอนวิธี AP2 เมื่อเลือกใช้อันดับ p=3 และ p=5 จะสูงกว่าขั้นตอนวิธีอื่นมากเกินไประดับนี้ ในวิทยานิพนธ์นี้จะเลือกพิจารณาการประมวลผลสัญญาณด้วยขั้นตอนวิธี AP2 อันดับ p=2 เพื่อเปรียบเทียบกับผลการประมวลผลสัญญาณด้วยขั้นตอนวิธีอื่นต่อไป

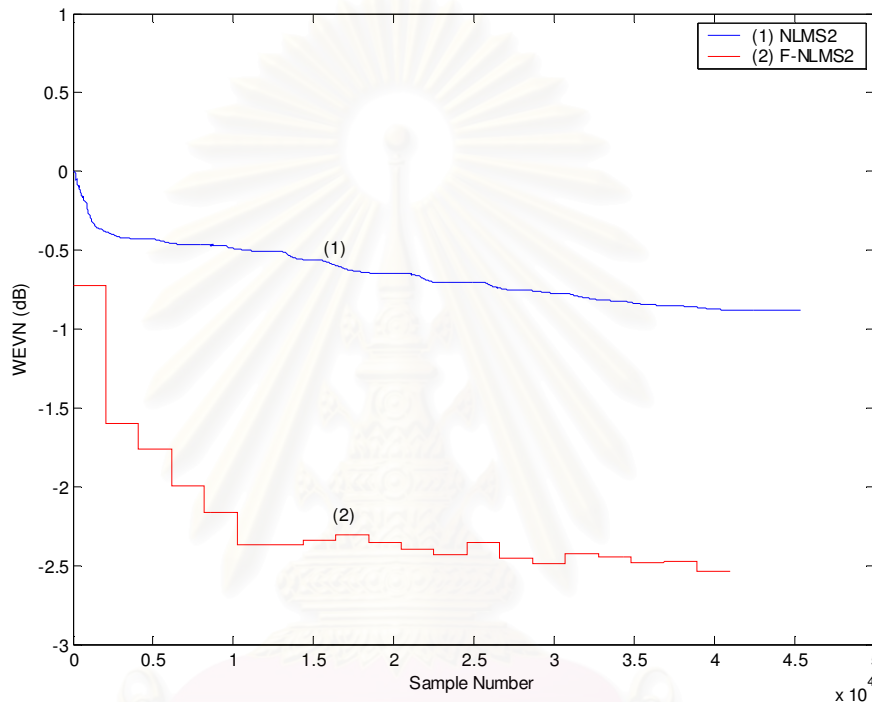
4.3 ผลการจำลองแบบระบบ SAEC ที่มีการประมวลผลสัญญาณในโดเมนความถี่

สำหรับการจำลองระบบ SAEC ที่มีการประมวลผลสัญญาณในโดเมนความถี่ จะเลือกใช้ขั้นตอนวิธี F-NLMS2 เพื่อเปรียบเทียบกับอัตราการลู่เข้าของระบบกับขั้นตอนวิธี NLMS2 ที่ประมวลผลในโดเมนเวลา โดยเลือกค่าช่วงก้าว μ ที่ทำให้ระบบมีอัตราการลู่เข้าที่เร็วที่สุด

ผลการจำลองแบบระบบ SAEC ด้วยขั้นตอนวิธี F-NLMS2 ดังบล็อกไดอะแกรมรูปที่ 2.4 ในบทที่ 2 โดยสังเกตจากค่า WEVN เปรียบเทียบกับขั้นตอนวิธี NLMS2 สามารถแสดงได้ดังรูปที่ 4.5

จากรูปที่ 4.5 จะเห็นว่าระบบ SAEC ที่มีการประมวลผลสัญญาณในโดเมนความถี่ด้วยขั้นตอนวิธี F-NLMS2 จะให้อัตราการลู่เข้าที่เร็วกว่าการประมวลผลในโดเมนเวลาด้วยขั้นตอนวิธี NLMS2 อีกทั้ง ค่าผิดพลาดในการปรับแก้ที่สถานะคงตัว (Steady-state) จะมีค่าน้อยกว่า แสดงว่าการประมวลผลสัญญาณในโดเมนความถี่ให้สมรรถนะที่ดีกว่า แต่การประมวลผลสัญญาณในโดเมนความถี่นั้นจะเกิดเวลาประวิงขึ้น เนื่องจากต้องรอข้อมูลบล็อกใหม่ของสัญญาณเข้า ซึ่งไม่

เป็นที่ต้องการในการใช้งานบางประเภทที่ทำงานในเวลาจริง เช่น ระบบการสัมมนาแบบสตรีมมิ่ง ดังนั้น จึงควรจำกัดเวลาประวิงที่เกิดขึ้นจากการประมวลผลให้มีค่าต่ำกว่ามาตรฐานที่กำหนดด้วย เช่น มาตรฐานของ International Telecommunication Union (ITU-T G.114) ระบุไว้ว่าเวลาประวิงของการส่งข้อมูลผ่านช่องสัญญาณทางเดียว (One-way Transmission) ไม่ควรเกิน 400 ms [20]

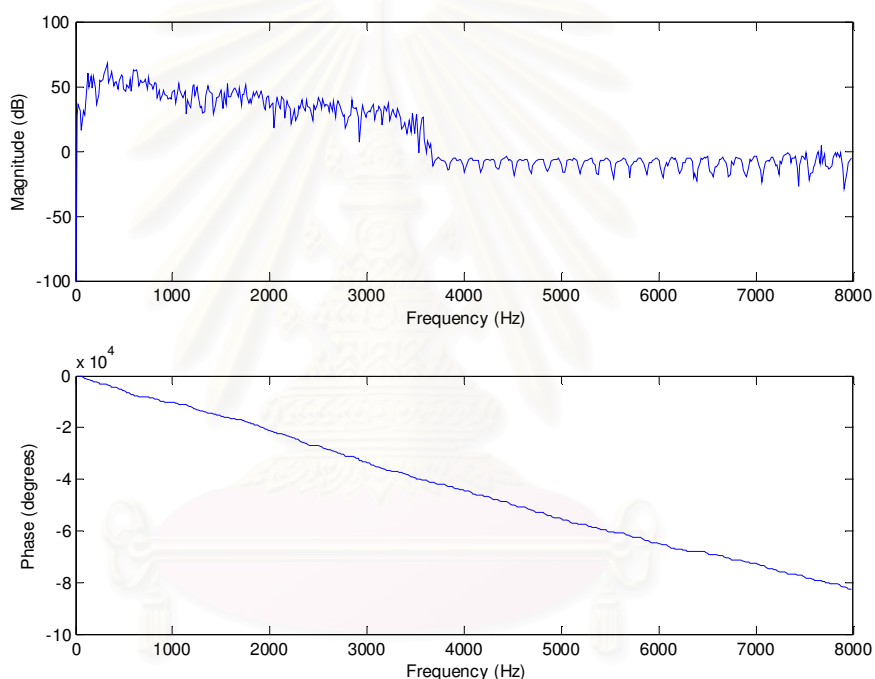


รูปที่ 4.5 ค่า WEVN เมื่อประมวลผลสัญญาณด้วย (1) ขั้นตอนวิธี NLMS2 ($\mu = 0.5$)
เปรียบเทียบกับ (2) ขั้นตอนวิธี F-NLMS2 ($\mu = 0.3$)

4.4 ผลการจำลองแบบระบบ SAEC ที่มีการประยุกต์ใช้โครงสร้างการกรองแถบย่อยร่วมกับการประมวลผลสัญญาณในโดเมนเวลาและโดเมนความถี่

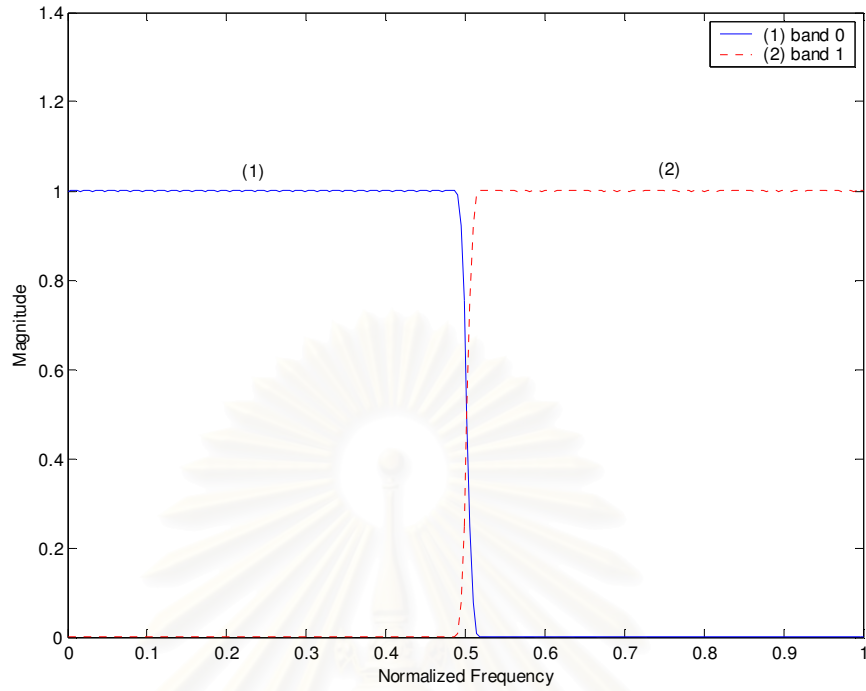
ในหัวข้อนี้จะศึกษาการประยุกต์ใช้โครงสร้างการกรองแถบย่อยร่วมกับการประมวลผลสัญญาณในโดเมนเวลาและโดเมนความถี่กับระบบ SAEC โดยเลือกใช้ขั้นตอนวิธี NLMS2 สำหรับการประมวลผลสัญญาณในโดเมนเวลาและขั้นตอนวิธี F-NLMS2 สำหรับการประมวลผลสัญญาณในโดเมนความถี่ และเมื่อทำการประยุกต์ใช้โครงสร้างการกรองแถบย่อยร่วมกับขั้นตอนวิธี NLMS2 และ F-NLMS2 จะเรียกว่า Subband NLMS2 และ Subband F-NLMS2 ตามลำดับ

จากการวิเคราะห์ผลตอบสนองอิมพัลส์ของสัญญาณเสียงพูดจากรูปที่ 4.2 ด้วยฟังก์ชัน “Freqz” จากโปรแกรม MATLAB จะได้ความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูด ดังแสดงในรูปที่ 4.6 พบว่า ความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดจะอยู่ในช่วงความถี่ต่ำ (0-4 kHz) มากกว่าในช่วงความถี่ที่สูงขึ้น เมื่อทำการแบ่งสเปกตรัมกำลังของสัญญาณเสียงพูดออกเป็น M แถบความถี่ย่อย ด้วยคลังวงจรกรองวิเคราะห์เอกรูป โดยออกแบบคลังวงจรกรองวิเคราะห์เอกรูปด้วยฟังก์ชัน “Fir1” จากโปรแกรม MATLAB ที่มีสัมประสิทธิ์ของผลตอบสนองอิมพัลส์แบบจำกัดจำนวน 64 ค่า ดังรูปที่ 4.7 จากนั้นทำการประมวลผลสัญญาณในแต่ละแถบความถี่ย่อยด้วยขั้นตอนวิธี NLMS2 โดยแบ่งเป็นกรณีต่างๆ ดังต่อไปนี้

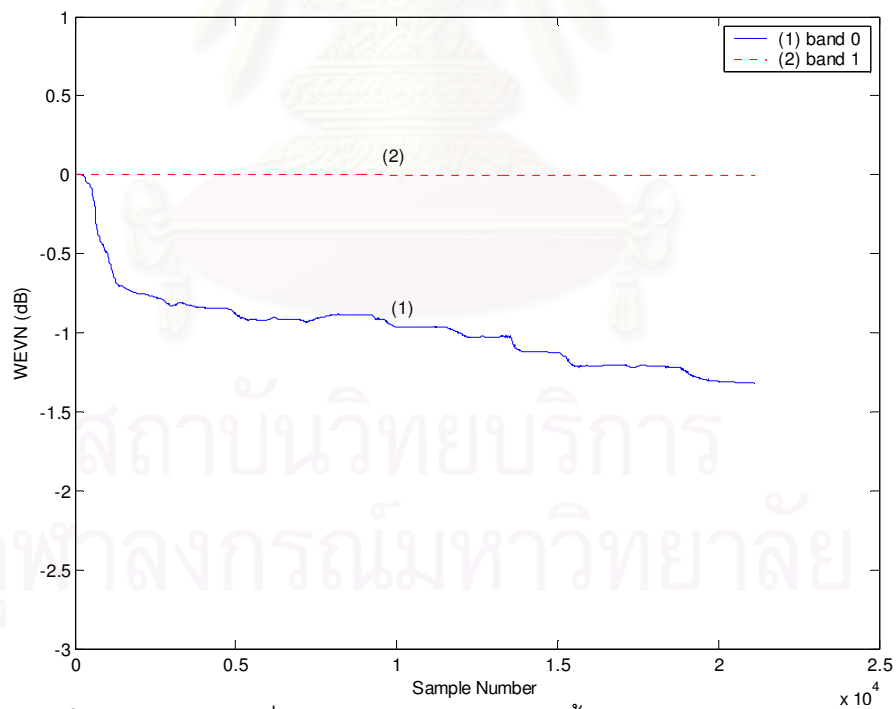


รูปที่ 4.6 ความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูด

สถาบันวิจัยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

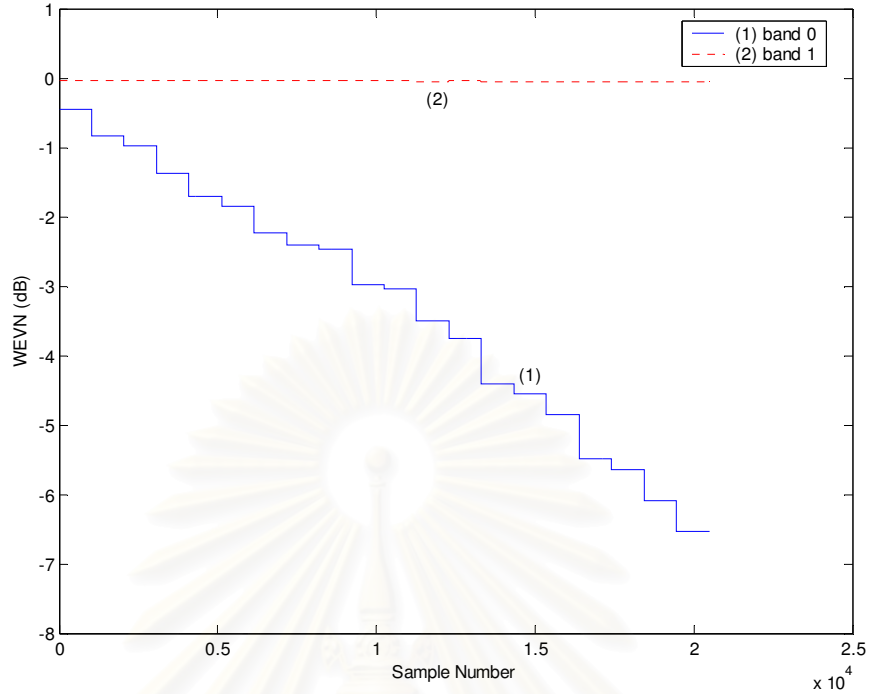


รูปที่ 4.7 ผลตอบสนองของอิมพัลส์ของคลังวงจรรองวิเคราะห์เอกรูปเมื่อ $M = 2$



รูปที่ 4.8 ค่า WEVN เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี Subband NLMS2

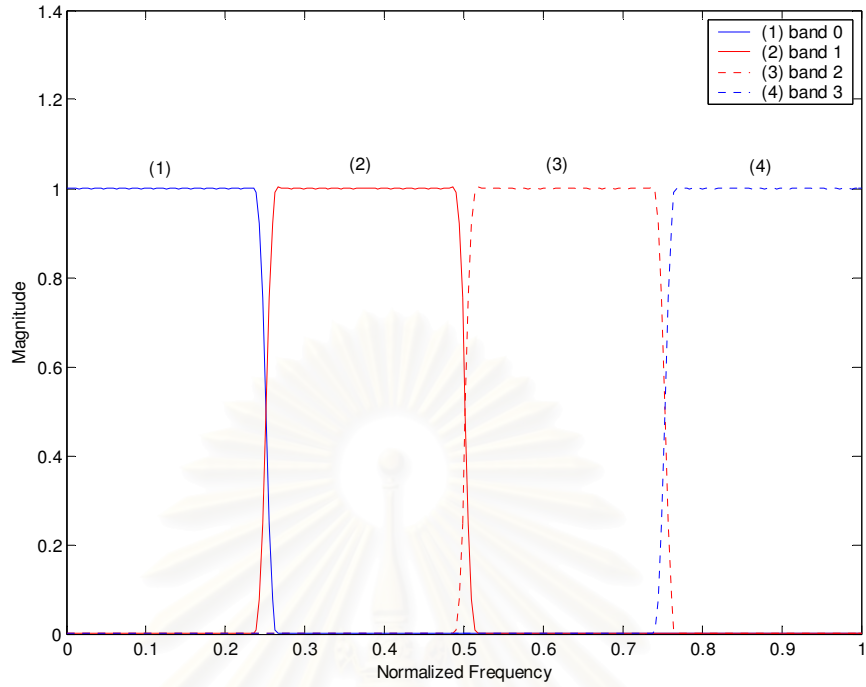
$$: M = 2, \mu = 0.04$$



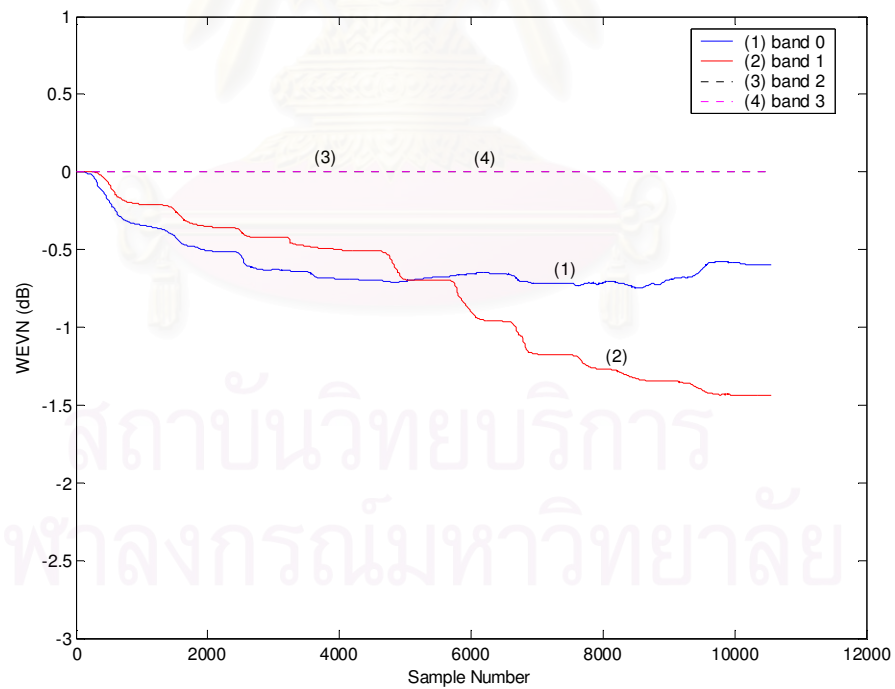
รูปที่ 4.9 ค่า WEVN เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี Subband F-NLMS2

$$: M = 2, \mu = 0.022$$

เมื่อแบ่งความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงด้วยคลังวงจรของวิเคราะห์ $M = 2$ ดังรูปที่ 4.7 พบว่า เมื่อเลือกใช้ค่าช่วงก้าวเป็น 0.04 ทั้งสองแถบความถี่ย่อย การประมวลผลในแถบความถี่ย่อยที่ 1 จะให้ค่าอัตราการลู่เข้าที่เร็วกว่าในแถบความถี่ที่ 2 เนื่องจากความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดจะอยู่ในช่วงความถี่ต่ำมากกว่า ในช่วงความถี่สูง อีกทั้ง ค่าผิดพลาดในการปรับแก้ของการประมวลผลด้วยการกรองแถบย่อยจะน้อยกว่าการประมวลผลด้วยแถบเต็ม ดังแสดงในรูปที่ 4.8 และ 4.9 ซึ่งพบว่า การประมวลผลในโดเมนเวลา (เมื่อใช้ขั้นตอนวิธี Subband NLMS2) จะให้สมรรถนะที่ต่ำกว่าการประมวลผลในโดเมนความถี่ (เมื่อใช้ขั้นตอนวิธี Subband F-NLMS2) โดยสังเกตผ่านทางกราฟแสดงค่า WEVN ในช่วงสถานะอยู่ตัวที่ลดลงได้ถึงประมาณ 1.3 dB เมื่อใช้ขั้นตอนวิธี Subband NLMS2 สำหรับค่า WEVN เมื่อใช้ขั้นตอนวิธี Subband F-NLMS2 ในแถบความถี่ที่ 1 สามารถลดลงได้ถึงประมาณ -6 dB ซึ่งลดลงได้มากกว่าในแถบความถี่ที่ 2 เพราะมีความหนาแน่นของสเปกตรัมกำลังอยู่ต่ำกว่า ในช่วงความถี่ต่ำ จึงสรุปได้ว่า การประมวลผลสัญญาณของระบบ SAEC ด้วยโครงสร้างการกรองแถบย่อยในโดเมนความถี่จะให้อัตราการลู่เข้าที่เร็วกว่าการประมวลผลสัญญาณด้วยโครงสร้างการกรองแถบย่อยในโดเมนเวลา

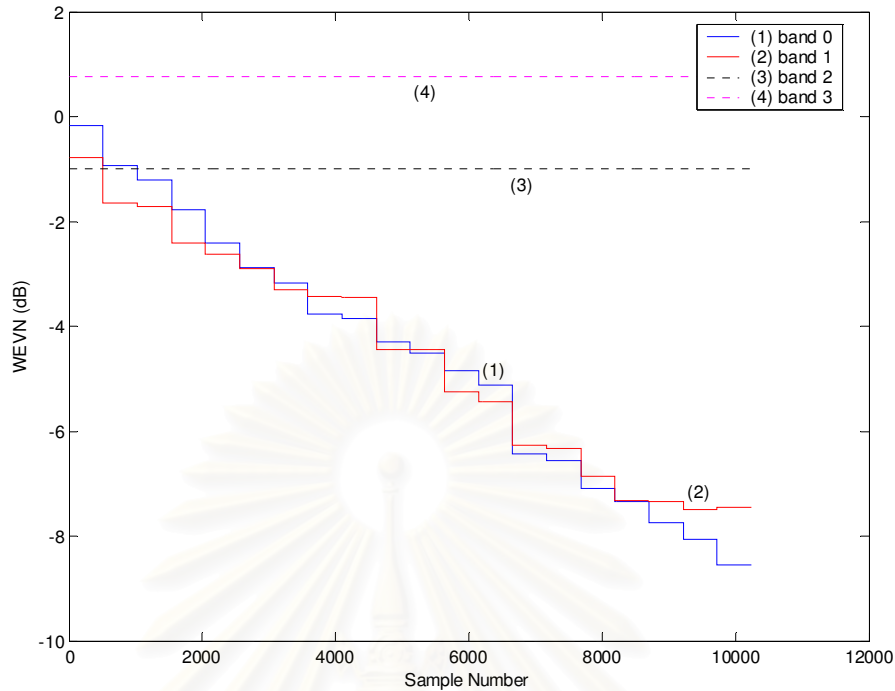


รูปที่ 4.10 ผลตอบสนองอิมพัลส์ของคลังวงจรกรองวิเคราะห์เอกรูปเมื่อ $M = 4$



รูปที่ 4.11 ค่า WEVN เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี Subband NLMS2

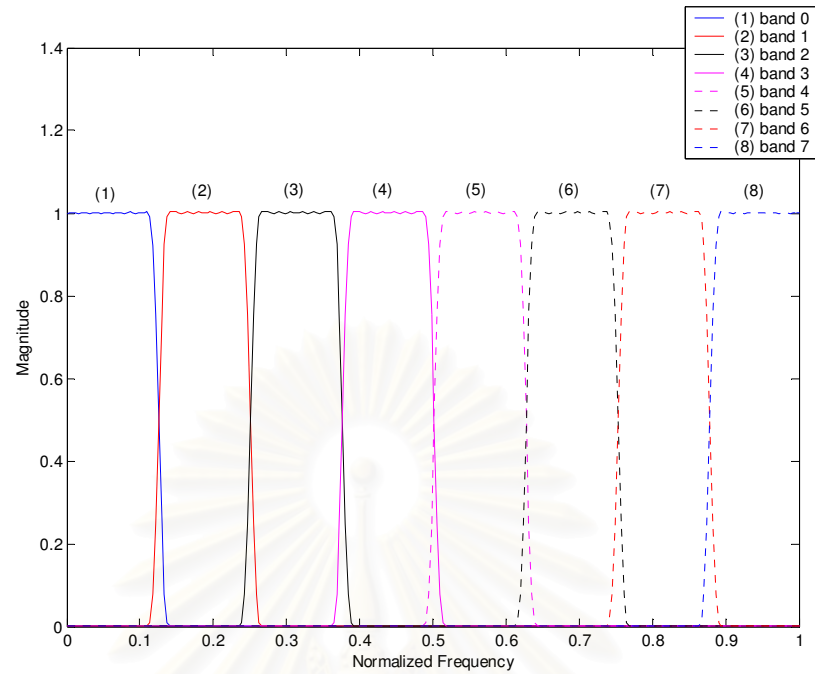
$$: M = 4, \mu = 0.004$$



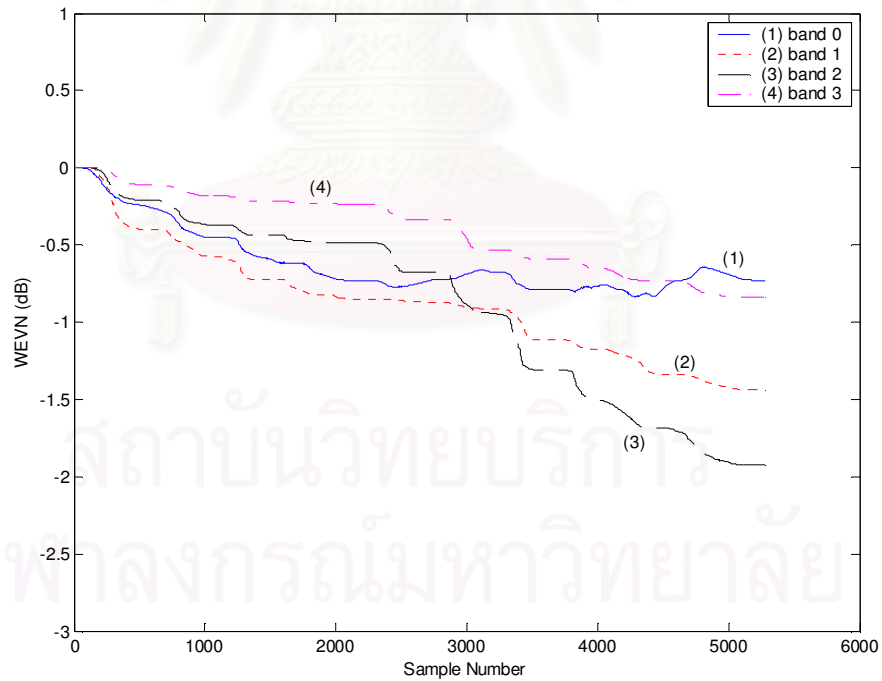
รูปที่ 4.12 ค่า WEVN เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี Subband F-NLMS2

$$: M = 4, \mu = 0.018$$

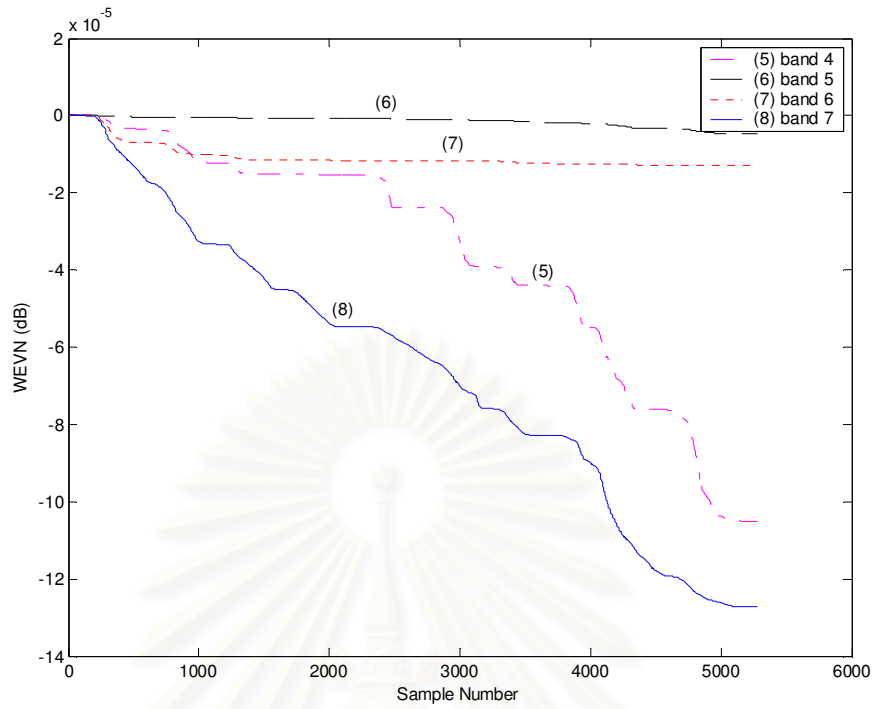
เมื่อแบ่งข้อมูลของสัญญาณเสียงด้วยคลังวงจรรองวิเคราะห์ $M = 4$ ดังรูปที่ 4.10 พบว่าการใช้จำนวนแถบความถี่ย่อยของคลังวงจรรองวิเคราะห์ที่มากขึ้นในการประมวลผล จะส่งผลให้ระบบ SAEC มีอัตราการลู่เข้าที่เร็วขึ้นด้วย โดยเฉพาะอย่างยิ่งในช่วงแถบความถี่ที่มีความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงอยู่มาก คือ ในแถบความถี่ย่อยที่ 1 และ 2 ทั้งนี้ การประมวลผลสัญญาณในโดเมนเวลาให้ค่า WEVN ในช่วงสถานะอยู่ตัวประมาณ -1.5 dB ดังรูปที่ 4.11 และการประมวลผลในโดเมนความถี่ให้ค่า WEVN ที่ลดลงได้ต่ำที่สุดในแถบความถี่ที่ 1 และ 2 คือลดลงได้ถึงประมาณ -8 dB ดังรูปที่ 4.12



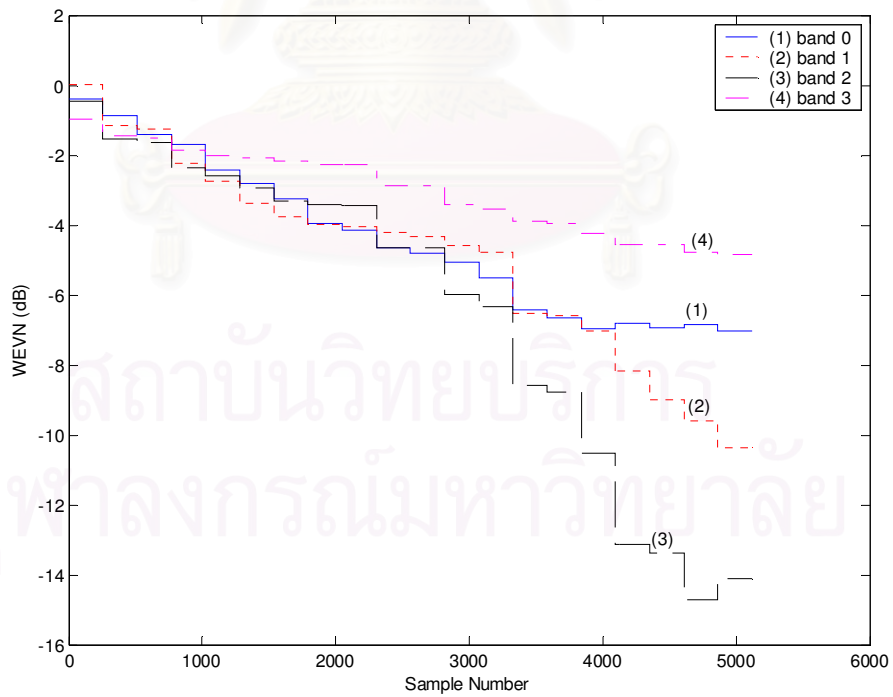
รูปที่ 4.13 ผลตอบสนองอิมพัลส์ของคลังวงจรกรองวิเคราะห์เอกรูปเมื่อ $M = 8$



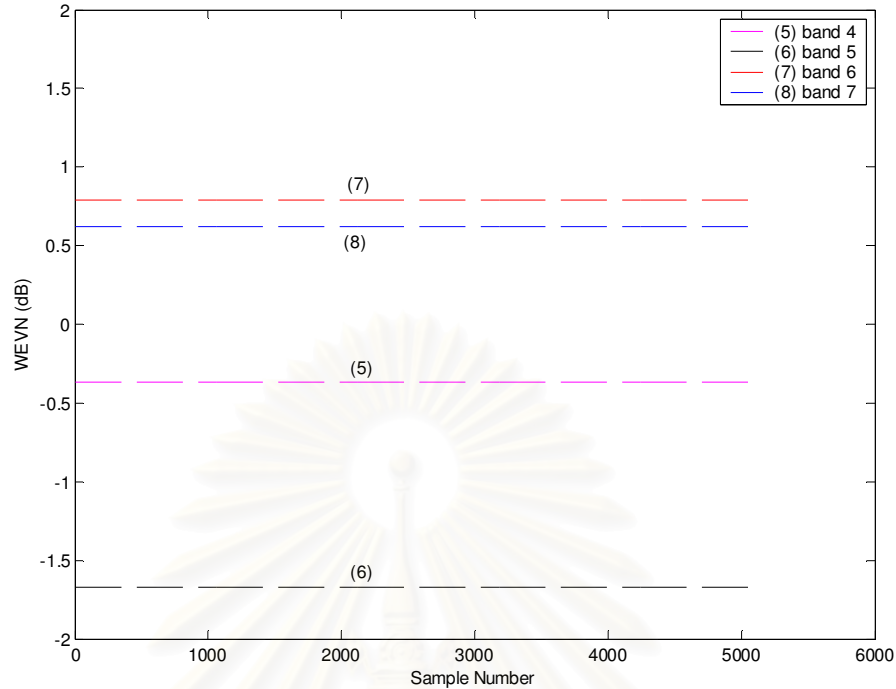
รูปที่ 4.14 ค่า WEVN แถบความถี่ย่อยที่ 0-3 เมื่อประมวลผลสัญญาณด้วย
ขั้นตอนวิธี Subband NLMS2: $M = 8, \mu = 0.002$



รูปที่ 4.15 ค่า WEVN แถบความถี่ย่อยที่ 4-7 เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี Subband NLMS2: $M = 8, \mu = 0.002$



รูปที่ 4.16 กราฟแสดงค่า WEVN แถบความถี่ย่อยที่ 0-3 เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี Subband F-NLMS2: $M = 8, \mu = 0.01$



รูปที่ 4.17 กราฟแสดงค่า WEVN แถบความถี่ย่อยที่ 4-7 เมื่อประมวลผลสัญญาณด้วยขั้นตอนวิธี

Subband F-NLMS2: $M = 8, \mu = 0.01$

เมื่อแบ่งความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงด้วยคลังวงจรรองวิเคราะห์ $M = 8$ ดังรูปที่ 4.13 พบว่าการใช้คลังวงจรรองวิเคราะห์ $M = 8$ ในการประมวลผลสัญญาณ จะทำให้ระบบ SAEC มีค่าอัตราการลู่เข้าที่เร็วกว่าการใช้คลังวงจรรองวิเคราะห์ด้วย $M = 4$ และ $M = 2$ ตามลำดับ สำหรับการประมวลผลในโดเมนเวลาให้ค่า WEVN ในช่วงสถานะอยู่ตัวที่ลดลงได้ต่ำที่สุดประมาณ -2 dB ในแถบความถี่ที่ 2 ดังรูปที่ 4.14 และการประมวลผลในโดเมนความถี่ให้ค่า WEVN ที่ลดลงได้ต่ำที่สุดในแถบความถี่ที่ 3 คือ ลดลงได้ถึงประมาณ -14 dB ดังรูปที่ 4.16 จึงสรุปได้ว่า การทำงานของวงจรรองแบบปรับตัวในระบบ SAEC ด้วยโครงสร้างการกรองแถบย่อยในโดเมนความถี่จะมีสมรรถนะที่ดีกว่าการประมวลผลสัญญาณด้วยโครงสร้างการกรองแถบย่อยในโดเมนเวลา อีกทั้ง การเพิ่มจำนวนแถบความถี่ย่อยให้มากขึ้น จะเพิ่มอัตราการลู่เข้าแก่การประมวลผลสัญญาณในแถบความถี่ย่อยที่มีสเปกตรัมกำลังของสัญญาณเสียงพุดอยู่มากมีอัตราการลู่เข้าที่เร็วขึ้น

4.5 การเปรียบเทียบความซับซ้อนทางการคำนวณ (Computational Complexity)

หัวข้อนี้เป็นการเปรียบเทียบความซับซ้อนทางการคำนวณของระบบ SAEC ด้วยขั้นตอนวิธี NLMS2 ที่ประมวลผลสัญญาณในโดเมนเวลาและ F-NLMS2 ที่ประมวลผลสัญญาณในโดเมนความถี่ สำหรับโครงสร้างการกรองแบบย่อ โดยเปรียบเทียบด้วยค่าการคูณจำนวนจริง (Real Multiplications, RMPs) และค่าการบวกและลบจำนวนจริง (Real Additions and Subtractions, RASs) ต่อสัญญาณเข้าหนึ่งตัวอย่าง เมื่อการหารหนึ่งครั้งมีความซับซ้อนทางการคำนวณคิดเป็น 16 RMPs [14] , M คือจำนวนแถบความถี่ย่อย และ L คือ จำนวนสัมประสิทธิ์ของวงจรรองแบบปรับตัว

4.5.1 ขั้นตอนวิธี NLMS2

ขั้นตอนวิธี NLMS2 ต้องการการคำนวณ

- สำหรับการหาค่าสัญญาณผิดพลาดของแต่ละช่องสัญญาณ ($e_i(n)$ เมื่อ $i = 1, 2$) ต้องการ L RMPs และ L RASs
- สำหรับการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวแต่ละตัว ($w_i(n)$ เมื่อ $i = 1, 2$) ต้องการ $L + 3$ RMPs กับ 1 div และ $L + 3$ RASs

ดังนั้น ขั้นตอนวิธี NLMS2 ต้องการการคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างคิดเป็น

$$4L + 38 \quad (4.3)$$

และการบวกและลบจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างคิดเป็น

$$4L + 6 \quad (4.4)$$

4.5.2 ขั้นตอนวิธี F-NLMS2

เมื่อกำหนดให้สัญญาณเข้าเป็นจำนวนจริงและความซับซ้อนทางการคำนวณของการแปลงฟูริเยร์ไม่ต่อเนื่องจะคิดเป็นแบบ Radix-2 ในโดเมนความถี่ [15] ดังนั้น แต่ละ N -point Real DFT (และ IDFT) ต้องการการคำนวณ $N \log_2 N$ RMPs และ $N - 1$ RASs เมื่อ $N = 2L$ และหนึ่งการคูณจำนวนเชิงซ้อนเท่ากับ 4 RMPs ในขณะที่การคูณจำนวนจริงกับจำนวนเชิงซ้อนเท่ากับ 2 RMPs

เมื่อพิจารณารูปที่ 2.4 ในบทที่ 2

- การประมวลผลสัญญาณหนึ่งบล็อกต้องการการคำนวณ 5 DFT/IDFTs เท่ากับ $5N \log_2 N$ RMPs และ $5(N-1)$ RASs [14]

ขั้นตอนวิธี F-NLMS2 ต้องการการคำนวณ

- สำหรับการคำนวณเวกเตอร์สัญญาณออกของวงจรของแบบปรับตัวแต่ละตัวต้องการ $4N$ RMPs
- สำหรับการคำนวณเวกเตอร์สัญญาณผิดพลาดต้องการ N RASs
- สำหรับการคำนวณเวกเตอร์สหสัมพันธ์ไขว้ (Cross-correlation) ของสัญญาณเข้าและสัญญาณผิดพลาดต้องการ $4N$ RMPs
- สำหรับการคูณเวกเตอร์ค่าช่วงก้าวต้องการ $2N$ RMPs
- และสำหรับหาค่าพลังงานเฉลี่ยของสัญญาณเข้าทุกบล็อกการคำนวณต้องการ $22N$ RMPs และ N RASs

ดังนั้น ขั้นตอนวิธี FNLMS2 ต้องการการคูณจำนวนจริงสำหรับแต่ละบล็อกการประมวลผลเท่ากับ

$$2(5N \log_2 N) + 2(32N) \quad (4.5)$$

และการบวกและลบจำนวนจริงสำหรับแต่ละบล็อกการประมวลผลเท่ากับ

$$5(N-1) + 2N \quad (4.6)$$

เมื่อความยาวของสัมประสิทธิ์วงจรของแบบปรับตัวเท่ากับ L การคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างจะได้เป็น

$$20 \log_2 L + 148 \quad (4.7)$$

และการบวกและลบจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างคิดเป็น

$$14 - \frac{5}{L} \quad (4.8)$$

4.5.3 ขั้นตอนวิธี Subband NLMS2

การคำนวณความซับซ้อนของขั้นตอนวิธี NLMS2 สำหรับการกรองแถบย่อย M แถบ ความถี่ย่อยในโดเมนเวลา และความยาวสัมประสิทธิ์ของวงจรรองวิเคราะห์และวงจรรองสังเคราะห์เท่ากับ N_{tap} จะต้องการการคำนวณ

- สำหรับการวิเคราะห์สัญญาณเข้าแต่ละตัว ($\mathbf{x}_i(n)$, $i = 1, 2$) และสัญญาณเข้าของไมโครโฟนในห้องรับ $d(n)$ ต้องการ $3MN_{\text{tap}}$ RMPs และ $3M(N_{\text{tap}} - 1)$ RASs
- สำหรับการสังเคราะห์สัญญาณผิดพลาด $e(n)$ ต้องการ MN_{tap} RMPs และ $M(N_{\text{tap}} - 1)$ RASs

ดังนั้น จึงต้องการการคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างสำหรับวิเคราะห์และสังเคราะห์สัญญาณได้เป็น

$$4MN_{\text{tap}} \quad (4.9)$$

และการบวกและลบจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างคิดเป็น

$$4M(N_{\text{tap}} - 1) \quad (4.10)$$

กำหนดให้ L_s เป็นความยาวสัมประสิทธิ์ของวงจรรองในแต่ละแถบความถี่ย่อย โดยที่ $L_s = \frac{L}{M}$ ดังนั้น การคำนวณขั้นตอนวิธี NLMS2 ในแต่ละแถบความถี่ย่อยจากสมการที่ (4.3) และสมการที่ (4.4) การคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างคิดเป็น

$$4\left(\frac{L}{M}\right) + 38 \quad (4.11)$$

และการบวกและลบจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างคิดเป็น

$$4\left(\frac{L}{M}\right) + 6 \quad (4.12)$$

ดังนั้น การคูณจำนวนจริงทั้งหมดต่อสัญญาณเข้าหนึ่งตัวอย่างจากสมการที่ (4.9) และ (4.11) จะเป็น

$$4\left(\frac{L}{M}\right) + 38 + 4MN_{\text{tap}} \quad (4.13)$$

และการบวกและลบจำนวนจริงทั้งหมดต่อสัญญาณเข้าหนึ่งตัวอย่างจากสมการที่ (4.10) และ (4.12) คิดเป็น

$$4\left(\frac{L}{M}\right) + 6 + 4M(N_{\text{tap}} - 1) \quad (4.14)$$

4.5.4 Subband F-NLMS2

ขั้นตอนวิธี F-NLMS2 ในโดเมนความถี่สำหรับการกรองแถบย่อย M แถบความถี่ย่อย จะต้องการการคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างสำหรับวิเคราะห์และสังเคราะห์สัญญาณ ดังสมการที่ (4.9) คือ

$$4MN_{\text{tap}} \quad (4.15)$$

และการบวกและลบจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างดังสมการที่ (4.10) คือ

$$4M(N_{\text{tap}} - 1) \quad (4.16)$$

สำหรับการคูณจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างในแต่ละแถบความถี่ย่อยจะเป็นดังสมการที่ (4.7) จะเท่ากับ

$$20 \log_2 \left(\frac{L}{M} \right) + 148 \quad (4.17)$$

และการบวกและลบจำนวนจริงต่อสัญญาณเข้าหนึ่งตัวอย่างดังสมการที่ (4.8) คือ

$$14 - \frac{5M}{L} \quad (4.18)$$

ดังนั้น การคูณจำนวนจริงทั้งหมดต่อสัญญาณเข้าหนึ่งตัวอย่างจากสมการที่ (4.15) และ (4.17) จะเป็น

$$20 \log_2 \left(\frac{L}{M} \right) + 148 + 4MN_{\text{tap}} \quad (4.19)$$

และการบวกและลบจำนวนจริงทั้งหมดต่อสัญญาณเข้าหนึ่งตัวอย่างจากสมการที่ (4.16) และ (4.18) จะเป็น

$$14 - \frac{5M}{L} + 4M(N_{\text{tap}} - 1) \quad (4.20)$$

ตารางที่ 4.2 การเปรียบเทียบความซับซ้อนทางการคำนวณของแต่ละขั้นตอนวิธีเทียบกับขั้นตอนวิธี NLMS2 เมื่อกำหนดให้ $N_{\text{tap}} = 64$ ค่า

Parameter		Complexity of each Algorithm / Complexity of NLMS2					
L	M	F-NLMS2		Subband NLMS2		Subband F-NLMS2	
		×/÷	+/-	×/÷	+/-	×/÷	+/-
512	1	0.1572	0.0068	-	-	-	-
	2	-	-	0.7546	0.7468	0.3931	0.2522
	4	-	-	0.7546	0.7429	0.6194	0.4975
	8	-	-	1.1227	1.1091	1.1103	0.9883
1024	1	0.0842	0.0034	-	-	-	-
	2	-	-	0.6284	0.6236	0.2032	0.1263
	4	-	-	0.5046	0.4968	0.3222	0.2491
	8	-	-	0.6284	0.4929	0.5651	0.4949
2048	1	0.0447	0.0017	-	-	-	-
	2	-	-	0.5645	0.5618	0.1045	0.0632
	4	-	-	0.3779	0.3735	0.1643	0.1247
	8	-	-	0.3779	0.3716	0.2863	0.2476
	16	-	-	0.5645	0.5550	0.5327	0.4935

ตารางที่ 4.2 แสดงการเปรียบเทียบความซับซ้อนทางการคำนวณของแต่ละขั้นตอนวิธีเทียบกับขั้นตอนวิธี NLMS จะเห็นว่าเมื่อความยาวสัมประสิทธิ์ของวงจรรอง (L) มีค่าเพิ่มขึ้น ความซับซ้อนในการคำนวณของแต่ละขั้นตอนวิธีจะมีค่าลดลงตามลำดับ โดยที่การประมวลผลในโดเมนความถี่จะมีความซับซ้อนในการคำนวณน้อยกว่าการประมวลผลในโดเมนเวลา และการประมวลผลโดยใช้การกรองแถบย่อยจะมีความซับซ้อนในการคำนวณน้อยกว่าการประมวลผลในแถบเต็ม และเมื่อเพิ่มจำนวนแถบความถี่ย่อยในการประมวลผล ความซับซ้อนในการคำนวณจะยิ่งมากขึ้น และเมื่อพิจารณาการแบ่งความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงออกเป็น $M = 16$ แถบความถี่ย่อย จะให้ค่าความซับซ้อนในการคำนวณที่มากกว่ากรณีที่ $M < 16$ แถบความถี่ย่อย ดังนั้น สำหรับโครงสร้างไฮบริดที่นำเสนอในวิทยานิพนธ์ฉบับนี้จึงเลือกศึกษาการปรับปรุงสมรรถนะของระบบการตัดสัญญาณเสียงสะท้อนในการสัมมนาแบบสเตอริโอโดยใช้การกรองแถบย่อยเป็น $M \leq 8$ แถบความถี่ย่อยเท่านั้น ดังนั้น จึงสามารถสรุปประสิทธิภาพการทำงานของแต่ละขั้นตอนวิธี ในด้านของความซับซ้อนทางการคำนวณ (Computational Complexity) อัตราการลู่เข้า (Convergence Rate) ของวงจรรองแบบปรับตัว และเวลาประวิงในตัว (Inherent Delay) ที่เกิดขึ้น ดังแสดงในตารางที่ 4.3

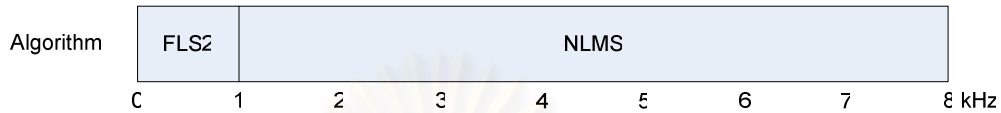
ตารางที่ 4.3 การเปรียบเทียบข้อดีข้อเสียของแต่ละขั้นตอนวิธี

	NLMS2	F-NLMS2	Subband NLMS2	Subband F-NLMS2
Computational Complexity	High	Lowest	Medium	Low
Convergence Rate	Slow	Fast	Medium	Fastest
Inherent Delay	No	High	No	Low

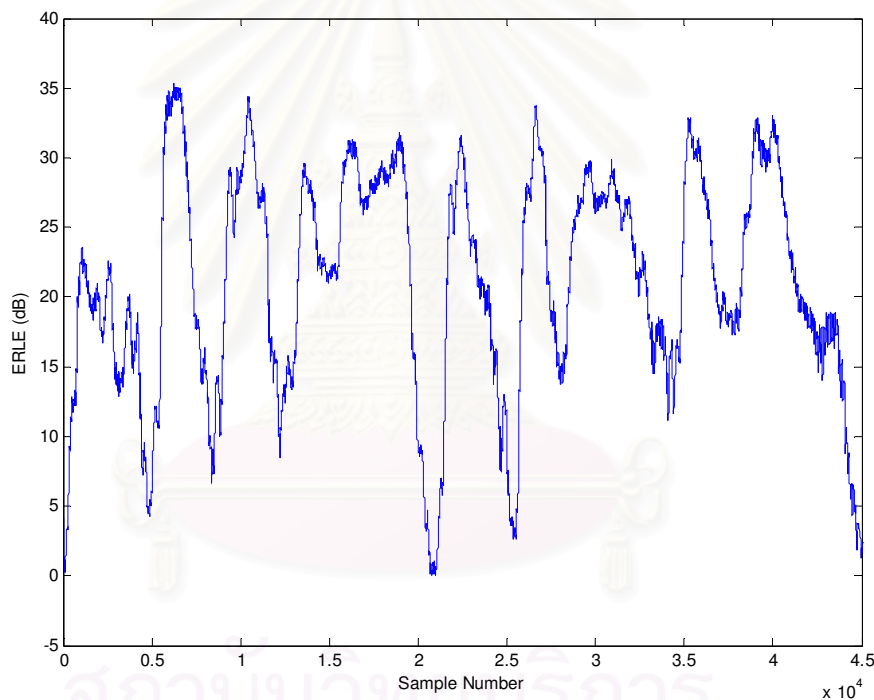
4.6 ผลการจำลองแบบโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18] เปรียบเทียบกับโครงสร้างไฮบริดที่นำเสนอ

หัวข้อนี้จะทำการเปรียบเทียบอัตราการลู่เข้าของวงจรรองแบบปรับตัวของระบบ SAEC ที่มีโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18] กับระบบ SAEC ที่มีโครงสร้างไฮบริดที่นำเสนอในวิทยานิพนธ์นี้ โดยผลการจำลองแบบของโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS

[18] ที่แถบความถี่ย่อยแรก (0-1 kHz) จะประมวลผลด้วยขั้นตอนวิธี FLS2 สำหรับแถบความถี่ที่สูงขึ้น (1-8 kHz) จะรวมสัญญาณทั้งสองช่องสัญญาณเป็นช่องสัญญาณเดี่ยวและทำการประมวลผลด้วยขั้นตอนวิธี NLMS ดังรูปที่ 4.18 สามารถสังเกตค่า ERLE ดังแสดงในรูปที่ 4.19



รูปที่ 4.18 การแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18]



รูปที่ 4.19 ค่า ERLE ของโครงสร้างไฮบริดระหว่าง FLS2 (0-1 kHz):

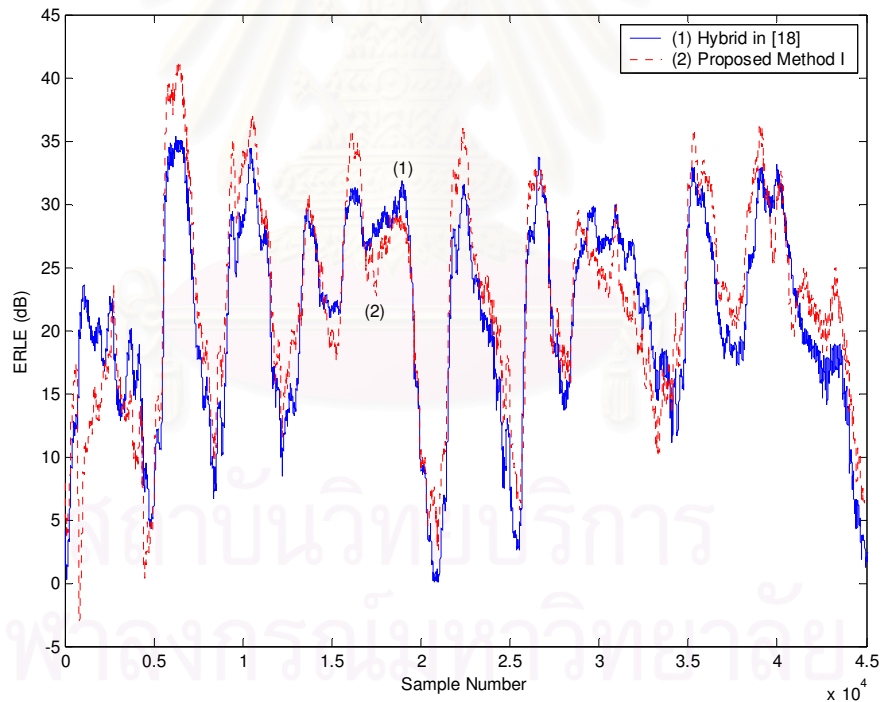
$$E_a(0) = E_b(0) = 10, W = 0.999 \text{ และ NLMS (1-8 kHz): } \mu = 0.03$$

จากหัวข้อที่ 4.2 และ 4.3 แสดงให้เห็นว่าขั้นตอนวิธี AP2 อันดับ $p=2$ ให้อัตราการลู่เข้าของวงจรรองแบบปรับตัวที่เร็วกว่าขั้นตอนวิธี FLS2 และขั้นตอนวิธี F-NLMS2 ให้อัตราการลู่เข้าที่เร็วกว่าขั้นตอนวิธี NLMS2 อีกทั้ง ความซับซ้อนทางการคำนวณของขั้นตอนวิธี F-NLMS2 ยังต่ำกว่าขั้นตอนวิธี NLMS2 ดังนั้น โครงสร้างไฮบริดที่น่าเสนอจะจำลองแบบการใช้ขั้นตอนวิธี AP2 อันดับ $p=2$ ในการปรับปรุงสัมประสิทธิ์ของวงจรรองในแถบความถี่ต่ำ (0-1 kHz) แทนขั้นตอนวิธี

FLS2 สำหรับการปรับปรุงสัมประสิทธิ์ของวงจรรองในแถบความถี่สูง (1-8 kHz) จะใช้ขั้นตอนวิธี F-NLMS แทนขั้นตอนวิธี NLMS โดยการแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอรหว่าง AP2 (0-1 kHz): $p=2$, $\mu = 0.3$ และ F-NLMS (1-8 kHz): $\mu = 0.02$ เมื่อแถบความถี่ย่อยแรกประมวลผลด้วยขั้นตอนวิธี AP2 สำหรับแถบความถี่ที่สูงขึ้นจะรวมเป็นช่องสัญญาณเดียวและทำการประมวลผลด้วยขั้นตอนวิธี F-NLMS ดังสรุปในรูปที่ 4.20 และค่า ERLE ของระบบดังกล่าวสามารถแสดงได้ดังรูปที่ 4.21

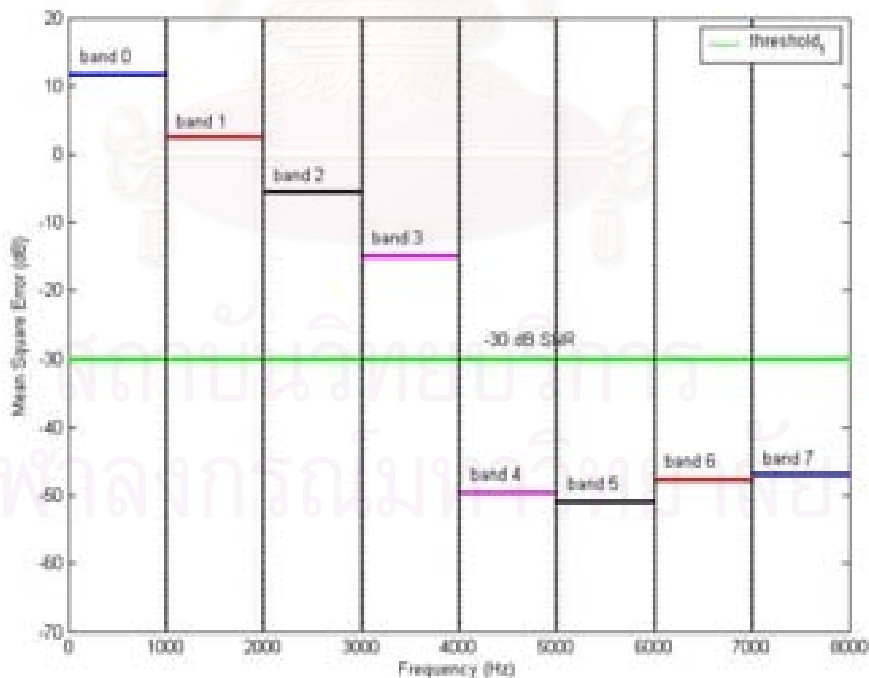


รูปที่ 4.20 การแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอแบบที่หนึ่ง



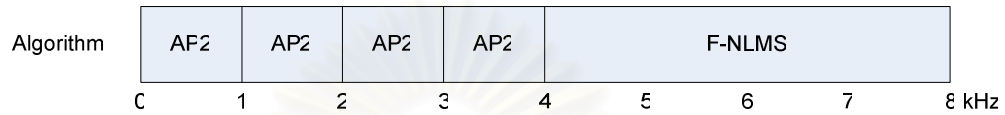
รูปที่ 4.21 ค่า ERLE เปรียบเทียบสมรรถนะระหว่าง (1) โครงสร้างไฮบริดระหว่าง FLS2 (0-1 kHz): $E_a(0) = E_b(0) = 10$, $W = 0.999$ และ NLMS (1-8 kHz): $\mu = 0.03$ กับ (2) โครงสร้างไฮบริดที่นำเสนอรหว่าง AP2 (0-1 kHz): $p=2$, $\mu = 0.3$ และ F-NLMS (1-8 kHz): $\mu = 0.02$

จะเห็นว่าโครงสร้างไฮบริดที่นำเสนอในรูปแบบที่ 4.21 ให้ค่าสมรรถนะ ERLE ใกล้เคียงกับโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18] และมีความซับซ้อนทางการคำนวณที่ต่ำกว่าดังตารางที่ 4.4 (e) และ 4.4 (f) และเมื่อพิจารณาความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดในรูปแบบที่ 4.6 จะพบว่า แถบความถี่ย่อยตั้งแต่ 1-4 kHz ยังมีค่าความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดอยู่มาก ดังนั้น โครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18] ที่ใช้ขั้นตอนวิธี NLMS ในการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวในแถบความถี่นี้ ยังไม่สามารถตัดสัญญาณเสียงสะท้อนได้ดีเพียงพอ โดยสามารถพิจารณารายละเอียดได้ดังรูปที่ 4.22 ที่ทำการแบ่งแถบความถี่ย่อยของสัญญาณเสียงพูดออกเป็น $M = 8$ แถบความถี่ย่อย จากนั้นจึงใช้ขั้นตอนวิธี NLMS2 ในการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวในแต่ละแถบความถี่ย่อย พบว่าแถบความถี่ย่อยที่ 1 (band 1: 1-2 kHz), ความถี่ย่อยที่ 2 (band 2: 2-3 kHz) และความถี่ย่อยที่ 3 (band 3: 3-4 kHz) ซึ่งเป็นแถบความถี่ตั้งแต่ 1-3 kHz มีค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error, MSE) มากกว่าแถบความถี่ย่อยที่ 4 ถึง 7 และมากกว่าสัญญาณรบกวนพื้นหลัง -30 dB SNR ดังนั้น ในแถบความถี่ตั้งแต่ 1-3 kHz จึงควรใช้ขั้นตอนวิธีที่ให้อัตราการลู่เข้าเร็วยิ่งกว่าขั้นตอนวิธี NLMS2

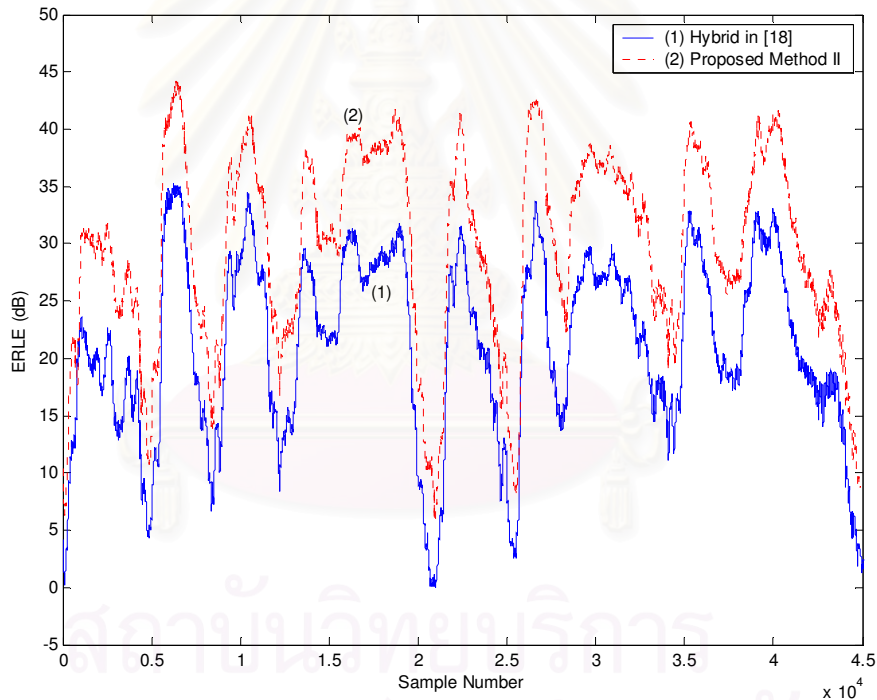


รูปที่ 4.22 ค่า MSE ด้วยขั้นตอนวิธี NLMS2 ($M = 8$)

ในการจำลองแบบต่อไป จึงได้นำเสนอการใช้ขั้นตอนวิธี AP2 ในการปรับปรุงสัมประสิทธิ์ของวงจรแบบปรับตัวในแถบความถี่ย่อยที่ 1, 2 และ 3 ตามลำดับ โดยการแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอใหม่จะเป็นดังแสดงในรูปที่ 4.23 ค่า ERLE ที่สังเกตได้แสดงในรูปที่ 4.24



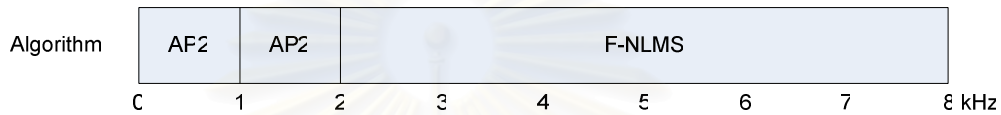
รูปที่ 4.23 การแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอแบบที่สอง



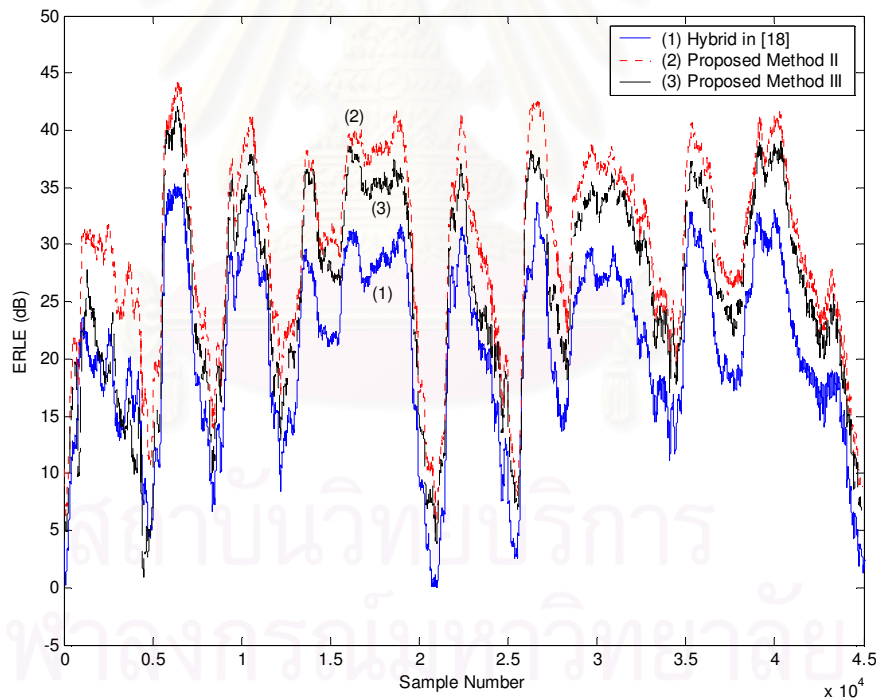
รูปที่ 4.24 ค่า ERLE เปรียบเทียบสมรรถนะระหว่าง (1) โครงสร้างไฮบริดระหว่าง FLS2 (0-1 kHz): $E_a(0) = E_b(0) = 10$, $W = 0.999$ และ NLMS (1-8 kHz): $\mu = 0.03$ กับ (2) โครงสร้างไฮบริดที่นำเสนอระหว่าง AP2 (0-4 kHz): $p=2$, $\mu = 0.3$ และ F-NLMS (4-8 kHz): $\mu = 0.02$

จากรูปที่ 4.24 พบว่าโครงสร้างไฮบริดที่นำเสนอมีอัตราการลู่เข้าสู่คำตอบของวงจรแบบปรับตัวเร็วกว่าโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18] แต่ความซับซ้อนทางการ

คำนวณจะมีค่าสูงชันมากดังตารางที่ 4.4 (g) เนื่องจากใช้ขั้นตอนวิธี AP2 ในการประมวลผล สัญญาณแถบความถี่ย่อยที่ 0 – 3 ดังนั้น การจำลองแบบต่อไปจึงเลือกใช้ขั้นตอนวิธี AP2 กับแถบความถี่ย่อยที่ 0 (0-1 kHz) และแถบความถี่ย่อยที่ 1 (1-2 kHz) เท่านั้น ส่วนในแถบความถี่ตั้งแต่ 2-8 kHz จะใช้ขั้นตอนวิธี F-NLMS โดยการแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วย โครงสร้างไฮบริดที่นำเสนอจะถูกปรับปรุงใหม่เป็นดังแสดงในรูปที่ 4.25 และค่า ERLE ที่สังเกต ได้แสดงในรูปที่ 4.26



รูปที่ 4.25 การแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอ แบบที่สาม



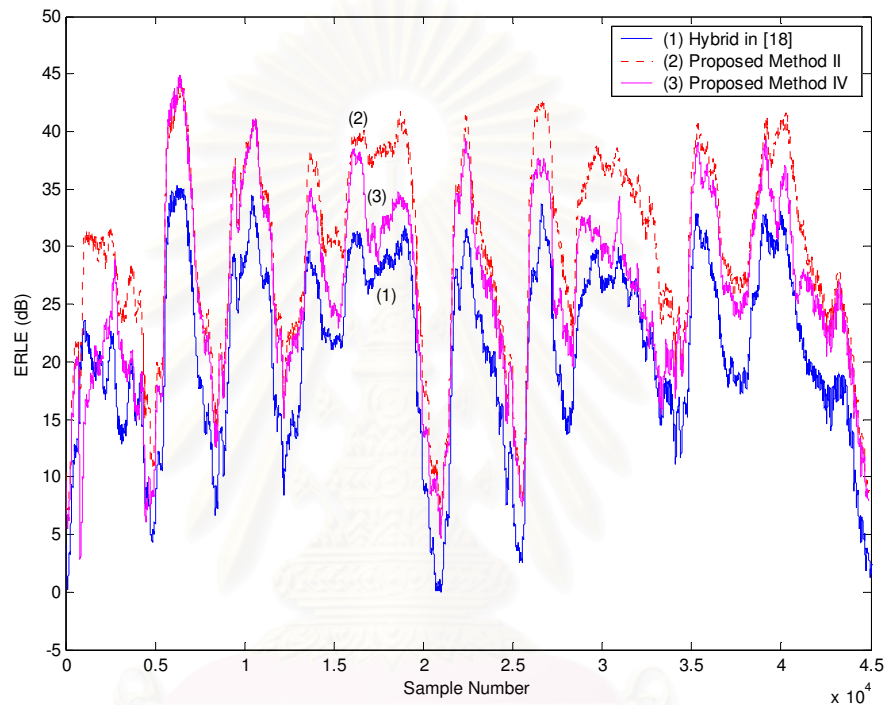
รูปที่ 4.26 ค่า ERLE เปรียบเทียบสมรรถนะระหว่าง (1) โครงสร้างไฮบริดระหว่าง FLS2 (0-1 kHz): $E_a(0) = E_b(0) = 10$, $W = 0.999$ และ NLMS (1-8 kHz): $\mu = 0.03$ กับ (2) โครงสร้างไฮบริดที่นำเสนอระหว่าง AP2 (0-4 kHz): $p=2$, $\mu = 0.3$ และ F-NLMS (4-8 kHz): $\mu = 0.02$ กับ (3) โครงสร้างไฮบริดที่นำเสนอระหว่าง AP2 (0-2 kHz): $p=2$, $\mu = 0.3$ และ F-NLMS (2-8 kHz): $\mu = 0.02$

จากรูปที่ 4.26 เมื่อเลือกใช้ขั้นตอนวิธี AP2 กับแถบความถี่ย่อยที่ 1 และ 2 จะทำให้ค่า ERLE ที่ลดลงจากโครงสร้างไฮบริดที่นำเสนอแบบที่สอง แต่ค่าสมรรถนะก็ยังมากกว่าโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18] และเมื่อพิจารณาค่าความซับซ้อนทางการคำนวณดังตารางที่ 4.4 (h) พบว่าความซับซ้อนทางการคำนวณของวิธีนี้จะลดต่ำลงมากเมื่อเทียบกับโครงสร้างไฮบริดที่นำเสนอแบบที่สองจนใกล้เคียงกับโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18] แต่การเลือกใช้ขั้นตอนวิธี AP2 กับแถบความถี่ย่อยที่ 1 และ 2 เท่านั้น จะทำให้สัญญาณผิดพลาดที่เกิดในแถบความถี่ย่อยที่ 3 และ 4 มีเพิ่มค่ามากขึ้น ซึ่งไม่เหมาะสมที่จะประยุกต์ใช้กับสัญญาณเสียงพูดในแถบความถี่ย่อยที่ 3 และ 4 ที่ค่าความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดยังมีค่ามาก ดังนั้น ในการจำลองแบบต่อไปจะเลือกใช้ขั้นตอนวิธี F-NLMS2 ในการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัว เนื่องจากความซับซ้อนทางการคำนวณของขั้นตอนวิธี F-NLMS2 น้อยกว่าขั้นตอนวิธี AP2 มาก จึงสามารถใช้ขั้นตอนวิธี F-NLMS2 กับแถบความถี่ย่อยที่ 1-3 โดยที่ความซับซ้อนทางการคำนวณของโครงสร้างไฮบริดที่นำเสนอใหม่ในแบบที่สี่นี้มีค่าลดต่ำกว่าโครงสร้างไฮบริดใหม่ในแบบที่สองและแบบที่สามที่นำเสนอก่อนหน้านี้ ทั้งนี้ ในแถบความถี่ย่อยที่ 0 จะยังเลือกใช้ขั้นตอนวิธี AP2 เช่นเดิม เนื่องจากความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดในแถบความถี่นี้มีค่ามากที่สุด จึงควรใช้ขั้นตอนวิธีที่ทำให้การปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวมีอัตราการลู่เข้าสู่ค่าตอบเร็วที่สุด

จากรูปที่ 4.28 โครงสร้างไฮบริดที่นำเสนอระหว่าง AP2 (0-1 kHz): $p=2$, $\mu = 0.3$ และ F-NLMS2 (1-4 kHz): $\mu = 0.02$ $\mu = 0.3$ และ F-NLMS2 (4-8 kHz): $\mu = 0.02$ เป็นโครงสร้างไฮบริดที่ให้อัตราการลู่เข้าสู่ค่าตอบได้เร็วกว่าแบบอื่นๆ ที่ศึกษาในวิทยานิพนธ์นี้ อีกทั้ง ความซับซ้อนทางการคำนวณของโครงสร้างไฮบริดที่นำเสนอนี้ยังต่ำมาก ดังนั้น โครงสร้างไฮบริดแบบที่สี่ที่นำเสนอจึงเป็นอีกทางเลือกหนึ่งสำหรับการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวที่นำเสนอขึ้นในวิทยานิพนธ์ฉบับนี้

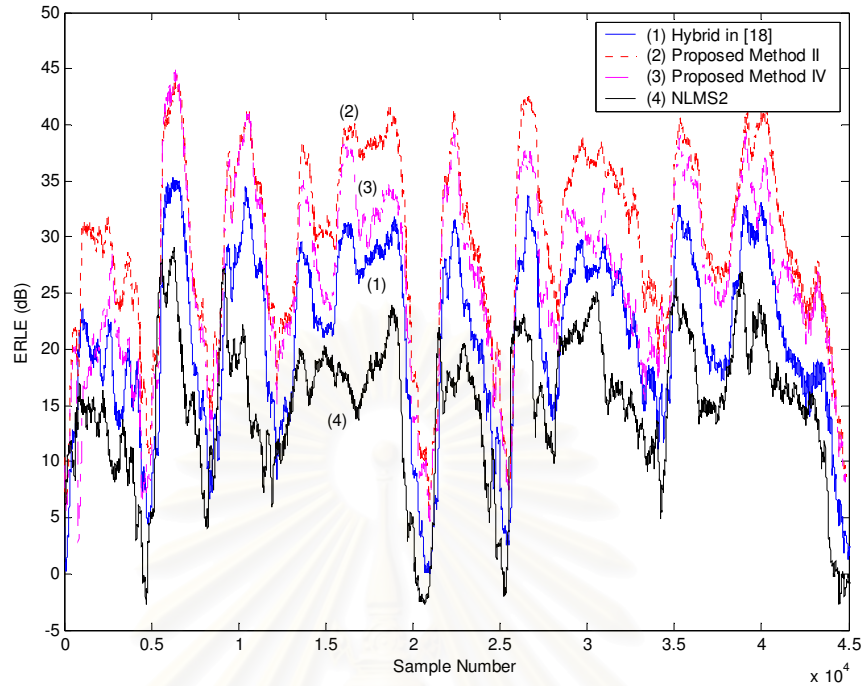
Algorithm	AF2	F-NLMS2	F-NLMS2	F-NLMS2	F-NLMS				
	C	1	2	3	4	5	6	7	8 kHz

รูปที่ 4.27 การแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอแบบที่สี่



รูปที่ 4.28 ค่า ERLE เปรียบเทียบสมรรถนะระหว่าง (1) โครงสร้างไฮบริดระหว่าง FLS2 (0-1 kHz): $E_a(0) = E_b(0) = 10$, $W = 0.999$ และ NLMS (1-8 kHz): $\mu = 0.03$ กับ (2) โครงสร้างไฮบริดที่นำเสนองระหว่าง AP2 (0-4 kHz): $p=2$, $\mu = 0.3$ และ F-NLMS (4-8 kHz): $\mu = 0.02$ กับ (3) โครงสร้างไฮบริดที่นำเสนองระหว่าง AP2 (0-1 kHz): $p=2$, $\mu = 0.3$ และ F-NLMS2 (1-4 kHz): $\mu = 0.02$ $\mu = 0.3$ และ F-NLMS2 (4-8 kHz): $\mu = 0.02$

จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 4.29 ค่า ERLE เปรียบเทียบสมรรถนะระหว่าง (1) โครงสร้างไฮบริดระหว่าง FLS2 (0-1 kHz): $E_a(0) = E_b(0) = 10$, $W = 0.999$ และ NLMS (1-8 kHz): $\mu = 0.03$ กับ (2) โครงสร้างไฮบริดที่นำเสนองระหว่าง AP2 (0-4 kHz): $p=2$, $\mu = 0.3$ และ F-NLMS (4-8 kHz): $\mu = 0.02$ กับ (3) โครงสร้างไฮบริดที่นำเสนองระหว่าง AP2 (0-1 kHz): $p=2$, $\mu = 0.3$ และ F-NLMS2 (1-4 kHz): $\mu = 0.02$ $\mu = 0.3$ และ F-NLMS2 (4-8 kHz): $\mu = 0.02$ กับ (4) ขั้นตอนวิธี NLMS2: $\mu = 0.5$

รูปที่ 4.29 แสดงค่าสมรรถนะ ERLE ของโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18] และโครงสร้างไฮบริดแบบที่สองและแบบที่สี่ที่นำเสนอขึ้นในวิทยานิพนธ์นี้ เพื่อเปรียบเทียบกับขั้นตอนวิธี NLMS2 จากการศึกษพบว่า โครงสร้างไฮบริดแบบที่สองจะให้อัตราการลู่เข้าที่เร็วที่สุด โครงสร้างไฮบริดแบบที่สี่จะมีอัตราการลู่เข้าอยู่ระหว่างโครงสร้างไฮบริดแบบที่สองและโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS [18] โดยที่โครงสร้างไฮบริดทุกรูปแบบที่นำเสนอขึ้นในวิทยานิพนธ์นี้จะมีอัตราการลู่เข้าที่เร็วกว่าเมื่อเปรียบเทียบกับขั้นตอนวิธี NLMS2

เมื่อกำหนดให้ $\text{Inv}(M_{2 \times 2})$ มีความซับซ้อนทางการคำนวณเท่ากับ 22 RMPs และ 1 ADSs, และ $\text{Inv}(M_{3 \times 3})$ เท่ากับ 55 RMPs และ 14 ADSs. [21] ค่าความซับซ้อนทางการคำนวณของแต่ละขั้นตอนวิธีสามารถสรุปได้ดังตารางที่ 4.4

ตารางที่ 4.4 ความซับซ้อนทางการคำนวณของแต่ละขั้นตอนวิธีที่ศึกษาในวิทยานิพนธ์นี้

Adaptive Filter Algorithm :	\times / \div	$+ / -$
a. FLS2	36904	36867
b. AP2 (p = 2)	32822	32768
c. AP2 (p = 3)	61570	61440
d. NLMS2	8230	8198
e. [18] : FLS2 (0 – 1 kHz) + NLMS (1 – 8 kHz)	9275	9221
f. Proposed I: AP2 (p = 2) (0 – 1 kHz) + F-NLMS (1 – 8 kHz)	4836	4614
g. Proposed II: AP2 (p = 2) (0 – 4 kHz) + F-NLMS (4 – 8 kHz)	18054	17658
h. Proposed III: AP2 (p = 2) (0 – 2 kHz) + F-NLMS (2 – 8 kHz)	9242	8962
i. Proposed IV: AP2 (p = 2) (0 – 1 kHz) + F-NLMS2 (1 – 4 kHz) + F-NLMS (4 – 8 kHz)	6518	5454

จากตารางที่ 4.4 พบว่าโครงสร้างไฮบริดแบบที่หนึ่ง (ตารางที่ 4.4 f) จะมีความซับซ้อนทางการคำนวณต่ำที่สุด เมื่อพิจารณาจากค่าการคูณจำนวนจริง (RMPs) และค่าการบวกและลบจำนวนจริง (RASs) และโครงสร้างไฮบริดที่มีความซับซ้อนทางการคำนวณมากที่สุด คือ โครงสร้างไฮบริดแบบที่สอง (ตารางที่ 4.4 g) สำหรับโครงสร้างไฮบริดแบบที่สี่ (ตารางที่ 4.4 i) จะมีความซับซ้อนทางการคำนวณลดต่ำกว่าโครงสร้างไฮบริดทั้งสองแบบที่กล่าวมา จะเห็นว่า การประมวลผลสัญญาณด้วยโครงสร้างไฮบริดจะมีความซับซ้อนทางการคำนวณโดยรวมต่ำกว่าการประมวลผลสัญญาณด้วยขั้นตอนวิธีที่มีโครงสร้างแถมเต็มที่ประมวลผลด้วยขั้นตอนวิธีเพียงขั้นตอนเดียว

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

วิทยานิพนธ์นี้ได้ศึกษาการใช้โครงสร้างการกรองแถบย่อยร่วมกับการประมวลผลสัญญาณในโดเมนความถี่ เพื่อปรับปรุงสมรรถนะการทำงานของระบบการตัดสัญญาณเสียงสะท้อนในการสนทนาแบบสเตริโอ ในด้านการลดความซับซ้อนทางการคำนวณและการเพิ่มอัตราการสุ่มเข้าของวงจรกรองแบบปรับตัวให้เร็วกว่าระบบการตัดสัญญาณเสียงสะท้อนแบบปกติที่ทำงานในโดเมนเวลา การเลือกใช้ขั้นตอนวิธีในการประมวลผลสัญญาณในแต่ละแถบความถี่ย่อยจะพิจารณาถึงประสิทธิภาพการทำงาน ข้อดี ข้อเสียของแต่ละขั้นตอนวิธีในด้านอัตราการสุ่มเข้าและความซับซ้อนทางการคำนวณ นอกจากนี้ การรวมกันของขั้นตอนวิธีต่างๆ ได้อย่างถูกต้องและนำเสนอในการปรับปรุงสัมประสิทธิ์ของวงจรกรองแบบปรับตัว เพื่อให้ได้ขั้นตอนวิธีรวมที่มีประสิทธิภาพสูงสุด เหมาะสมแก่การใช้งานในระบบการตัดสัญญาณเสียงสะท้อน ทั้งนี้ ความถูกต้องแม่นยำในการจำลองวิธีสะท้อนทางเสียงของวงจรกรองแบบปรับตัวจะถูกสังเกตจากค่าผิดพลาดในการปรับแก้ผ่านทางค่า WEVN และค่า ERLE ให้มีค่าต่ำกว่าของระบบการตัดสัญญาณเสียงสะท้อนในการสนทนาแบบสเตริโอแบบปกติ

จากการศึกษาที่เกี่ยวข้อง พบว่า โครงสร้างไฮบริดระหว่างขั้นตอนวิธีแบบโมโนและขั้นตอนวิธีแบบสเตริโอที่ประยุกต์ใช้โครงสร้างการกรองแถบย่อยร่วมกับคลังวงจรกรองไม่เอกรูปเป็นโครงสร้างแบบหนึ่งที่สามารถนำมาใช้งานในระบบการตัดสัญญาณเสียงสะท้อนแบบสเตริโอได้อย่างดี เนื่องจากสามารถลดความซับซ้อนทางการคำนวณของระบบโดยรวมเมื่อเปรียบเทียบกับ การเลือกใช้ขั้นตอนวิธีบางขั้นตอนวิธีที่ทำให้อัตราการสุ่มเข้าของวงจรกรองแบบปรับตัวเพิ่มสูงขึ้น โดยที่โครงสร้างไฮบริดระหว่าง FRLS2 และ NLMS ใน [18] ได้อ้างอิงจากหลักการของจิตวิทยาทางเสียงที่ว่า ความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดเชิงสเตริโอในช่วงความถี่ต่ำกว่า 1 kHz จะมีค่ามากกว่าความหนาแน่นของสเปกตรัมกำลังในช่วงความถี่ที่สูงขึ้น ดังนั้น บทความดังกล่าวจึงได้เลือกใช้ขั้นตอนวิธี FRLS2 ที่ให้อัตราการสุ่มเข้าสู่ค่าตอบที่เร็วในการปรับปรุงสัมประสิทธิ์ของวงจรกรองแบบปรับตัวเมื่อทำการประมวลผลสัญญาณในแถบความถี่ตั้งแต่ 0-900 Hz สำหรับแต่ละช่องสัญญาณ เมื่อความถี่การซัดตัวอย่างของสัญญาณเสียงเข้าที่ใช้ใน [18] เท่ากับ 16 kHz นอกจากนี้ การใช้ตัวประกอบเดซิเมชัน (M) ยังสามารถลดความซับซ้อนทางการคำนวณของการประมวลผลสัญญาณในแต่ละแถบความถี่ย่อยได้ด้วย ทั้งนี้ ในบทความที่ [18] ได้ใช้การแปลงไม่เชิงเส้นในการลดสหสัมพันธ์ของสัญญาณเสียงเชิงสเตริโอ เพื่อปรับปรุงการ

ลู่เข้าสู่ค่าตอบของวงจรรองแบบปรับตัวในระบบการตัดสัญญาณเสียงสะท้อนแบบสเตอริโอให้ลู่เข้าสู่ค่าตอบอย่างถูกต้องมากขึ้น หากแต่ว่าการแปลงไม่เชิงเส้นดังกล่าว อาจส่งผลให้เกิดการลดทอนคุณภาพของสัญญาณเสียงได้

สำหรับข้อมูลของสัญญาณเสียงในแถบความถี่ตั้งแต่ 900-8000 Hz ซึ่งมีพลังงานของสัญญาณเสียงเชิงสเตอริโออยู่น้อยกว่าในช่วงความถี่ 0-900 Hz ผู้แต่งใน [18] เลือกที่จะทำการประมวลผลแบบช่องสัญญาณเดียว โดยทำการรวมสัญญาณทั้งสองช่องสัญญาณเป็นช่องสัญญาณเดียวและใช้ระบบการตัดสัญญาณเสียงสะท้อนแบบช่องสัญญาณเดียวในการตัดสัญญาณเสียงสะท้อนแทนระบบการตัดสัญญาณเสียงสะท้อนแบบสเตอริโอ เพื่อลดความซับซ้อนทางการคำนวณของระบบโดยรวม สำหรับขั้นตอนวิธีที่ [18] เลือกใช้ในการประมวลผลสัญญาณในช่วงความถี่สูงดังกล่าว คือ ขั้นตอนวิธี NLMS

อย่างไรก็ตาม จากการศึกษาคุณลักษณะของสัญญาณเสียงพูด พบว่า ในช่วงแถบความถี่ย่อยระหว่าง 1-4 kHz นั้น ความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงยังมีค่ามากกว่าความหนาแน่นของสเปกตรัมกำลังในช่วงความถี่ที่สูงกว่า 4 kHz ดังนั้น จึงควรทำการปรับปรุงประสิทธิภาพของการประมวลผลในช่วงแถบความถี่ย่อยดังกล่าว วิทยานิพนธ์นี้จึงนำเสนอโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS เพื่อใช้ในระบบการตัดสัญญาณเสียงสะท้อนแบบสเตอริโอ โดยจะเลือกใช้ขั้นตอนวิธี AP2 ในการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวในแถบความถี่ที่ 0-1 kHz เนื่องจากขั้นตอนวิธีนี้สามารถให้อัตราการลู่เข้าสู่ค่าตอบเร็วขึ้นได้เมื่ออันดับของการฉายเพิ่มมากขึ้น และสามารถปรับลดอันดับของการฉายในกรณีที่ต้องการลดความซับซ้อนทางการคำนวณ สำหรับช่วงแถบความถี่ย่อยที่สูงขึ้น ได้เลือกใช้การผสมผสานกันระหว่างขั้นตอนวิธี F-NLMS2 และ F-NLMS ที่ทำการประมวลผลสัญญาณในโดเมนความถี่ เพื่อให้ระบบการตัดสัญญาณเสียงสะท้อนโดยรวมสามารถลดความซับซ้อนทางการคำนวณได้ ทั้งนี้ จะไม่นำวิธีการลดสหสัมพันธ์ของสัญญาณเสียงมาใช้กับระบบการตัดสัญญาณเสียงสะท้อนดังกล่าว เพื่อไม่ให้มีผลกระทบต่อคุณภาพของสัญญาณเสียง

จากการจำลองแบบ พบว่า โครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอขึ้นในวิทยานิพนธ์นี้จะให้อัตราการลู่เข้าที่เร็วกว่าโครงสร้างไฮบริดระหว่าง FLS2 และ NLMS ใน [18] และมีค่าความซับซ้อนทางการคำนวณที่ต่ำกว่าด้วย นอกจากนี้ เมื่อเปรียบเทียบกับขั้นตอนวิธี NLMS2 ซึ่งเป็นขั้นตอนวิธีพื้นฐานสำหรับระบบการตัดสัญญาณเสียงสะท้อนในการสลับแบบสเตอริโอแบบปกติ โครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอจะให้สมรรถนะที่ดีกว่าในด้านการเพิ่มอัตราการลู่เข้าของวงจรรองแบบปรับตัวให้เร็วขึ้น รวมทั้งยังสามารถลดความซับซ้อนทางการคำนวณได้ต่ำกว่าขั้นตอนวิธี NLMS2 อีกด้วย

5.2 ข้อเสนอแนะสำหรับงานวิจัยในอนาคต

งานที่ควรได้รับการศึกษาหรือพัฒนาต่อไปในอนาคต คือ

1. ศึกษาถึงประสิทธิภาพการทำงานของระบบการตัดสัญญาณเสียงสะท้อนด้วยโครงสร้างไฮบริดที่นำเสนอในสภาพแวดล้อมที่วิถีสะท้อนทางเสียงเป็นระบบเชิงเส้นและแปรเปลี่ยนตามเวลา (Linear Time-variant System) เช่น การเปลี่ยนแปลงแบบฉับพลันของวิถีสะท้อนทางเสียง (Abrupt Change of the Acoustic Echo Path) หรือการแปลงตามเวลาแบบช้าของวิถีสะท้อนทางเสียง (Slowly Time-varying of the Acoustic Echo Path) เป็นต้น
2. ศึกษาถึงความเป็นไปได้ในการนำไปสร้างจริง (Real-time Implementation) ของระบบการตัดสัญญาณเสียงสะท้อนด้วยโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอ โดยเฉพาะอย่างยิ่งด้านความซับซ้อนทางการคำนวณ รวมถึงหน่วยความจำ (Memory) ที่จำเป็นในระบบ

รายการอ้างอิง

1. M. M. Sondhi, D. R. Morgan and J. L. Hall. Stereophonic Acoustic Echo Cancellation- An Overview of the Fundamental Problem. IEEE Signal Processing Letters. 2, 8 (August 1995) : 148-151.
2. S. Haykin. Adaptive Filter Theory. Englewood Cliffs. NJ. : Prentice Hall, 1996.
3. G. Schmidt. Acoustic Echo and Noise Control for Low-Cost Processor. DSP World Spring Desing Conference. pp. 1-28. April, 2000.
4. R. E. Crochiere and L. R. Rabiner. Multirate Digital Signal Processing. Englewood Cliffs. NJ. : Prentice Hall, 1983.
5. N. J. Fliege. Multirate Digital Signal Processing. NY : John Wiley & Sons, 1994.
6. F. Amano, H. P. Meana and A. de Luca. A Multirate Acoustic Echo Canceled Structure. IEEE Trans. Communications. 43, 7 (July 1995) : 2172-2176.
7. G. A. Clark, S. K. Mitra and S. R. Parker. Block Implementation of Adaptive Digital Filters. IEEE Trans. Circuits and Systems. 28, 6 (June 1981) : 584-592.
8. K. Eneman and M. Moonen. Hybrid Subband/Frequency-Domain Adaptive Systems. Signal Processing. 8, 1 (January 2001) : 117-136.
9. T. Ogunfunmi and A. Peterson. Alternative Implementations for the Frequency-domain LMS Adaptive Filter. IEEE Trans. Acoust. Speech and Signal Processing. 3, 1 (April 1990) : 1441-1444.
10. A. O. Ogunfunmi and A. M. Peterson. On the Implementation of the Frequency-domain LMS Adaptive Filter. IEEE Trans. Circuits and Systems II. 39, 5 (May 1992) : 318-322.
11. S. L. Gay and S. Tavathia. The Fast Affine Projection Algorithm. in Proc. IEEE ICASSP'95. pp. 3023-3026. May, 1995.
12. F. Amand, J. Benesty, A. Gilloire and Y. Grenier. A Fast Two-channel Affine Projection Algorithm for Stereophonic Acoustic Echo Cancellation. in Proc. IEEE ICASSP'96. pp. 949-952. May, 1996.
13. M. Bellanger. Adaptive Digital Filtering and Signal Analysis. NY : Marcel Dekker Inc, 1987.

14. G. V. Moustakidis and S. Theodoridis. Fast Newton Transversal Filters – A new class of adaptive estimation algorithms”. IEEE Trans. Signal Processing. 39 (October 1991) : 2184-2193.
15. J. J. Shynk. Frequency-domain and Multirate Adaptive Filtering. IEEE Signal Processing Magazine. 9, 1 (January 1992) : 14-37.
16. N. Tangsangiumvisai. 2102797 Statistical Digital Signal Processing. (Lecture node). 2nd semester, 2003.
17. E. R. Ferrara. Fast Implementation of LMS Adaptive Filters. IEEE Trans. Acoust. Speech and Signal Processing. 28, 4 (August 1980) : 474-475.
18. J. Benesty, D. R. Morgan and M. M. Sondhi. A Hybrid Mono/Stereo Acoustic Echo Canceler. IEEE Trans. Speech and Audio Processing. 6, 5 (September 1998) : 468-475.
19. E.R. Ferrara. Freq-domain Adaptive Filtering in Adaptive Filters. Englewood Cliffs. NJ. : Prentice Hall, 1985.
20. ITU Telecommunication Standardization Sector. ITU-T Recommendation G.114 : One-way Transmission Time. available at <http://www.itu.int/ITU-T/>.
21. D. Miyasaki. Technique : Inverse Matrix. available at <http://www.cvl.iis.u-tokyo.ac.jp/~miyazaki/tech/teche23.html>.



ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทความทางวิชาการที่ได้รับการเผยแพร่

1. B. Kritayanun and N. Tangsangiumvisai, "On Performance Improvement of Stereophonic Acoustic Echo Cancellation through the use of Hybrid Structure," in the Proceedings of IEEE International Midwest Symposium on Circuits and Systems (MWSCAS'05, Cincinnati, Ohio, USA), 7 – 10 August 2005.
2. B. Kritayanun and N. Tangsangiumvisai, "Complexity Reduction of SAEC Systems using a Hybrid Mono/Stereo Structure," in the Proceedings of National Electrical Engineering Conference (EECON'05, Phuket, Thailand), 20 – 21 October 2005.
3. B. Kritayanun and N. Tangsangiumvisai, "Improved Performance for Stereophonic Acoustic Echo Cancellation using Subband Structure," in the Proceedings of IEEE International Region 10 Technical Conference on Analog and Digital Techniques in Electrical Engineering (TENCON'05, Melbourne, Australia), 22 – 24 November 2005.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

On Performance Improvement of Stereophonic Acoustic Echo Cancellation through the use of Hybrid Structure

B. Krittayanun and N. Tangsangiumvisai

Digital Signal Processing Research Laboratory, Department of Electrical Engineering,
Faculty of Engineering, Chulalongkorn University, Bangkok, 10330, Thailand.
e-mail: Nisachon.T@eng.chula.ac.th

Abstract

This paper proposes an adaptive filtering algorithm for Stereophonic Acoustic Echo Cancellation (SAEC) in order to obtain computational complexity reduction and convergence rate improvement. The hybrid mono/stereo structure between Fast Affine Projection (FAP) and Frequency-domain Normalized Least Mean Square (F-NLMS) algorithms is introduced, based on subband processing. Simulation results on recorded speech signals suggest that convergence rate can be improved at low computational cost.

INTRODUCTION

Most of adaptive filtering algorithms for Stereophonic Acoustic Echo Cancellation (SAEC) systems are processing in time domain, such as the two-channel Least Mean Square (LMS2), two-channel Recursive Least Squares (RLS2) algorithms [1] – [3], etc. The tradeoff between convergence rate and computational complexity between LMS2 and RLS2 is achieved through a number of fast versions of RLS-type (FRLS) algorithms [4] – [6]. These algorithms, however, encounter the instability problem. Subband and frequency-domain implementations [7] – [9] are alternative approaches for system complexity reduction. Via the use of the decimation process, the sampling rate in each frequency band can be reduced. The number of adaptive filter taps in each subband is D times smaller than the fullband ones when D represents a decimation factor. Thus, lower computational complexity than the conventional SAEC systems can be achieved. In contrast to the fullband system, the eigenvalue spread of the autocorrelation input matrix, \mathbf{R} , in each subband can be reduced by processing the input signal into M subbands. Hence, faster convergence rate can potentially be obtained. In addition, block-based adaptation of the frequency-approach usually employs fast Fourier Transform (FFT) algorithm for computational efficiency.

A hybrid structure employing subband and frequency-domain processing is proposed in this paper. The proposed technique is based on the investigation of the hybrid mono/stereo structure in [10] for further improvement on the performance of the SAEC system, in terms of convergence rate, while keeping the computational complexity lower than the conventional fullband approach. For fast convergence speed, the proposed technique employs the two-channel version of the Fast Affine Projection (FAP2) algorithm [11]. In high-frequency regions, the frequency-domain version of mono-channel Normalized Least Mean

Square (F-NLMS) algorithm is chosen for lowering the computational complexity of the SAEC system, as compared to the conventional one. The Mean Square Error (MSE) in each subband is used as a criterion for algorithm selection between FAP2 and F-NLMS.

THE HYBRID STRUCTURE FOR SAEC SYSTEM

A. Hybrid Mono/Stereo AEC in [10]

Based on the psychoacoustical principle that stereo effect is due mostly to sound energy below about 1 kHz, two AEC systems are differently applied to low-frequency and high-frequency signals. The FRLS2 algorithm [5] is employed in the low-frequency subbands (100-900 Hz) in order to obtain fast convergence rate, when the sampling frequency (f_s) is 16 kHz. In addition, a decimation factor, D , is applied to each channel for further complexity reduction of the SAEC system. The sum of the stereo signals is highpass filtered (900-8000 Hz) to obtain mono-channel signal since little amount of speech signal energy is contained in the high-frequency region. The mono-channel NLMS algorithm is employed for complexity reduction, as compared to the conventional SAEC system that normally employs two-channel algorithms. Note that, the non-linear transformation (NL) is included in each channel in order to decorrelate partially the low-frequency stereo signals [10].

B. The Proposed Hybrid Structure for SAEC

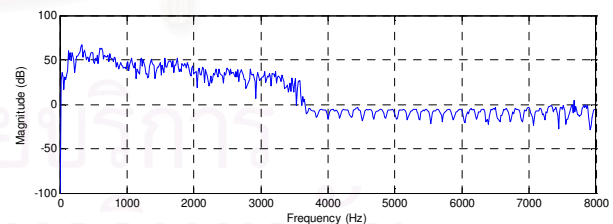


Figure 1. Frequency response of a speech signal $x_1(n)$ (sampling frequency is $f_s = 16$ kHz).

By considering at the magnitude response of one channel of stereo speech signals with $f_s = 16$ kHz, as shown in Fig. 1, its significant energy is contained in frequencies below 2 kHz, and gradually decreases when the frequency reaches beyond 4 kHz. According to speech energy in each subband, adaptive filtering algorithms can be chosen to be employed in each subband differently. Algorithms with fast convergence rate should be applied to the particular subbands which contain the most significant amount of spectral

density of speech inputs. On the other hand, algorithms with slow convergence rate, but with lower computational complexity, can be employed to those subbands with less spectral density of speech signal in order to reduce the computational complexity of the overall system.

The criterion for selection of fast or slow convergence rate adaptive filtering algorithms to be processed in each subband is suggested in this paper to be the average version of the MSE in each subband. For example, with the choice of $M = 8$ number of subbands, the MSE plots for each subband are illustrated in Figure 2, when all subbands employ the two-channel NLMS (NLMS2) algorithm. It can be seen that, the MSE levels in band 0 up to band 3 are higher than those in band 4 to band 7. The threshold level of MSE criterion is chosen to be equal to -30 dB, whose magnitude is the output Signal-to-Noise Ratio (SNR) level of the added background noise at ‘Mic1’ in all simulations. For this particular case, the first 4 subbands will employ adaptive algorithms with fast convergence rate since the MSE values are greater than the selected threshold level. On the other hand, the adaptive algorithms with slower convergence rate but with lower complexity than the one chosen for the low-frequency subbands are operated in the last 4 bands.

In the proposed hybrid structure, the algorithm with fast convergence rate is chosen to be the two-channel version of the Fast Affine Projection (FAP2) [11]. The faster speed of convergence of FAP2 can be obtained by increasing the projection order, p . However, very high projection order results in large computational complexity of the SAEC system. For the low complexity adaptive algorithm employed in the high-frequency region, the mono-channel version of frequency-domain Normalized Least Mean Square (F-NLMS) algorithm is suggested, instead of the mono-channel NLMS algorithm as chosen in the hybrid structure in [10]. Hence,

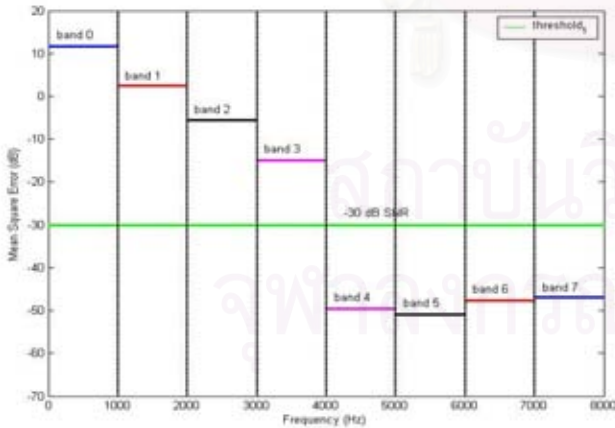


Figure 2. The MSE values of each subband ($M = 8$).

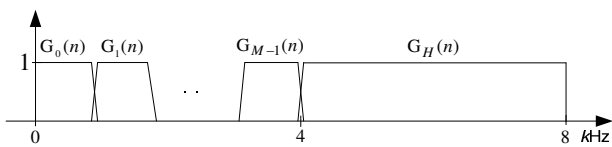


Figure 3. Frequency response of the proposed AFB.

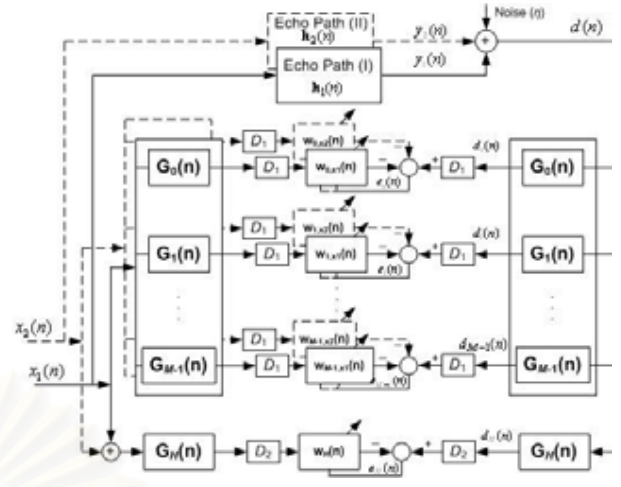


Figure 4. The proposed hybrid structure for SAEC.

the proposed structure for SAEC can achieve faster convergence rate than the fullband one. Moreover, the computational complexity of the proposed hybrid structure should be lower than the subband SAEC system that employs two-channel algorithms in all subbands.

The Analysis Filter Bank (AFB) with slightly overlapping frequency responses, as depicted in Figure 3, is employed for splitting the input signals in each channel, $x_i(n)$ for $i = 1, 2$. Additionally, another AFB is needed for the microphone signal in the receiving room, $d(n)$. With the chosen threshold level, the frequency responses of AFBs for low-frequency ($0 - 4$ kHz) signals, $G_i(z)$, for $i = 0, 1, \dots, M - 1$, are as depicted in Figure 3. Note that $G_0(z)$ is the prototype filter of AFB and other filters are obtained via the frequency-shifting method [12]. The decimation factor in the first 4 subbands is $D_1 = 2M$ where $M = 4$ subbands. The frequency response $G_H(z)$ for the signals with higher frequencies than 4 kHz is a half-band highpass filter. In this final subband, the decimation factor is $D_2 = 2$. The block diagram of the proposed system is given in Figure 4.

COMPUTATIONAL COMPLEXITY ANALYSIS

In this section, the computational complexity in mono-channel version of each algorithm, in terms of the number of real multiplications per input sample (RMPs) and number of additions and subtractions, is given in Table 1 - 4. The adaptive filter length is chosen to be L whereas N_{tap} denotes the length of the prototype filter of AFB. The diagonal matrix of the step-size parameter of F-NLMS is represented by $\mathbf{M}(k)$ and the block size of F-NLMS is $N = 2L$. Note that, in this paper, the two-channel Fast Least Squares (FLS2) [4] is used for the implementation of the hybrid structure in [10].

EXPERIMENTAL RESULTS

Speech signals, sampled at 16 kHz, were normalized to have zero mean and unity variance, as depicted in Figure 5,

and used as the input signals $x_1(n)$ and $x_2(n)$ of the stereo teleconferencing system. The acoustic echo signals due to AEPs in the receiving room that were assumed to be time-invariant, were obtained as the convolution sum between the input signals and the L -tap impulse responses $\mathbf{h}_1(n)$ and $\mathbf{h}_2(n)$. The 30 dB SNR of background noise, that was assumed to be of zero-mean and uncorrelated with both input signals, was added to the microphone signal in the receiving room. The length of adaptive filter was chosen to be $L = 2048$ for both AEC and SAEC parts. Note that, the adaptive filter length in each of the first 4 subbands and the final subband was equal to D_1 and D_2 respectively, where $N_{\text{tap}} = 64$.

For comparison, the performance of adaptive filters was evaluated in terms of Echo Return Loss Enhancement (ERLE), given by

$$\text{ERLE}(n) = 10 \times \log_{10} \left(\frac{\sum_{i=0}^{N_j} d^2(n-i)}{\sum_{i=0}^{N_j} e^2(n-i)} \right) \quad (1)$$

First, the performance of the SAEC system with fullband structure employing FAP2 with $p = 2$ and FLS2 are investigated. It is shown in Figure 6 that FAP2 obtains superior performance than FLS2, both in terms of faster convergence rate and lower computational complexity. (see Table 5).

The proposed hybrid structure is subsequently investigated, when $f_s = 16$ kHz. The hybrid structure in [10] employs FLS2 in the first subband (0 – 1 kHz) and the mono-channel NLMS in the other subbands (1 – 8 kHz). The MSE value in each subband is calculated and compared with the threshold level of 30 dB. By employing FAP2 in the first subband of the proposed structure, it is found that all MSE values in the first 4 subbands are greater than the selected threshold. Hence, FAP2 is employed in the first 4 subbands. For this case, the overall system obtains much faster convergence rate than the conventional SAEC system, employing fullband NLMS2, as demonstrated in Figure 7. The computational complexity is, however, greater than the conventional system. As compared to the hybrid structure in [10], similar performance in terms of convergence rate can be obtained, but the proposed technique gives larger complexity. Therefore, as a compromise between complexity and convergence speed performance of the system, FAP2 is only employed in the first subband whereas the two-channel F-NLMS (F-NLMS2) algorithm is used in the other subbands that MSE criterion is met, despite its inferior convergence rate. The ERLE performance is similar to that in [10], and is faster than the conventional system employing NLMS2, as shown in Figure 8. The computational complexity of the proposed structure in this case is the lowest among other investigated technique. The computational complexity of each algorithm for all cases is given in Table 5, when 1 division is equivalent to 16 RMPs [8], and $\text{Inv}(M_{2 \times 2})$ equals to 22 RMPs and 1 addition.

Table 1. Fast Least Squares (FLS)

available at time: n	adaptive filter: $\mathbf{w}(n)$, forward prediction: $\mathbf{A}(n)$, backward prediction: $\mathbf{B}(n)$, data vector: $\mathbf{X}(n)$, adaptation gain: $\mathbf{G}(n)$, prediction error energy: $E_a(n)$, forgetting factor: W	
new data at time: n+1	input signal: $x(n+1)$, reference signal: $d(n+1)$	

Adaptation Gain Updating:	\times/\div	$+/-$
$e_a(n+1) = x(n+1) - \mathbf{A}^T(n)\mathbf{X}(n)$	L	L
$\mathbf{A}(n+1) = \mathbf{A}(n) + \mathbf{G}(n)e_a(n+1)/\alpha(n)$	$L+1\text{div}$	L
$E_a(n+1) = (E_a(n) + e_a(n+1)e_a(n+1)/\alpha(n))W$	$2+1\text{div}$	1
$\mathbf{G}_1(n+1) = \begin{bmatrix} 0 \\ \mathbf{G}(n) \end{bmatrix} + \frac{e_a(n+1)}{E_a(n+1)} \begin{bmatrix} 1 \\ -\mathbf{A}(n+1) \end{bmatrix} = \begin{bmatrix} \mathbf{M}(n+1) \\ m(n+1) \end{bmatrix}$	$L+1+1\text{div}$	$L+1$
$e_b(n+1) = x(n+1-L) - \mathbf{B}^T(n)\mathbf{X}(n+1)$	L	L
$\mathbf{G}(n+1) = \mathbf{M}(n+1) + m(n+1)\mathbf{B}(n)$	L	L
$\alpha_1(n+1) = \alpha(n) + e_a(n+1)e_a(n+1)/E_a(n+1)$	$1+1\text{div}$	1
$\alpha(n+1) = \alpha_1(n+1) - m(n+1)e_b(n+1)$	1	1
$E_b(n+1) = (E_b(n) + e_b(n+1)e_b(n+1)/\alpha(n+1))W$	$2+1\text{div}$	1
$\mathbf{B}(n+1) = \mathbf{B}(n) + \mathbf{G}(n+1)e_b(n+1)/\alpha(n+1)$	$L+1\text{div}$	L

Adaptive Filter:	\times/\div	$+/-$
$e(n+1) = d(n+1) - \mathbf{w}^T(n)\mathbf{X}(n+1)$	L	L
$\mathbf{w}(n+1) = \mathbf{w}(n) + \mathbf{G}(n+1)e(n+1)$	L	$L+1$

Table 2. Normalized Least Mean Square (NLMS)

Adaptive Filter :	\times/\div	$+/-$
$e(n) = d(n) - \mathbf{w}^T(n)\mathbf{x}(n)$	L	L
$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\varepsilon + \ \mathbf{x}(n)\ ^2} \mathbf{x}(n)e(n)$	$L+3+1\text{div}$	$L+3$

Table 3. Fast Affine Projection (FAP)

Adaptive Filter :	\times/\div	$+/-$
$e(n) = d(n) - \mathbf{X}^T(n)\mathbf{w}(n)$	pL	pL
$\phi(n) = [\mathbf{X}^T(n)\mathbf{X}(n) + \delta\mathbf{I}]^{-1} e(n)$	$p^2(L+1) + \text{Inv}(M_{pp})$	p^2L
$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\varepsilon + \ \mathbf{X}(n)\ ^2} \mathbf{X}(n)\phi(n)$	$pL+1$	pL

Table 4. Fast Normalized Least Mean Square (F-NLMS)

Adaptive Filter :	\times/\div	$+/-$
FFT/IFFT	$5N \log_2 N$	$5(N-1)$
$\mathbf{E}(k) = \mathbf{D}(k) - \mathbf{W}(k)\mathbf{X}(k)$	$4N$	N
$\mathbf{W}(k+1) = \mathbf{W}(k) + \frac{\mathbf{M}(k)}{\varepsilon + \ \mathbf{X}(k)\ ^2} \mathbf{X}^H(k)\mathbf{E}(k)$	$N(12+1\text{div})$	N

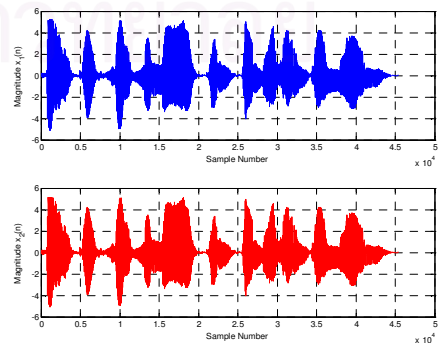


Figure 5. Speech input signals.

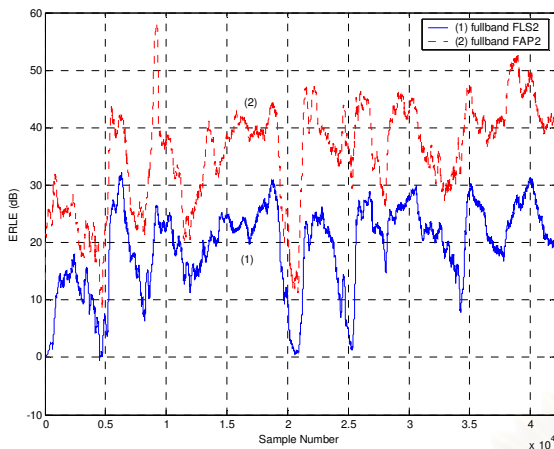


Figure 6. ERLE performance of fullband SAEC system. (1) FLS2 ($E_a(0)=1300$, $W=1$) and (2) FAP2 ($p=2$, $\mu=0.5$)

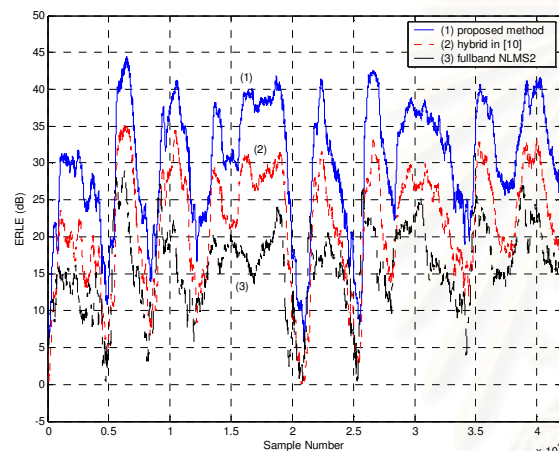


Figure 7. Comparison of ERLE performance: (1) proposed method (FAP2 : $p=2$, $\mu=0.2$, F-NLMS: $\mu=0.02$) (2) hybrid one in [10] (3) fullband SAEC (NLMS2: $\mu=0.3$)

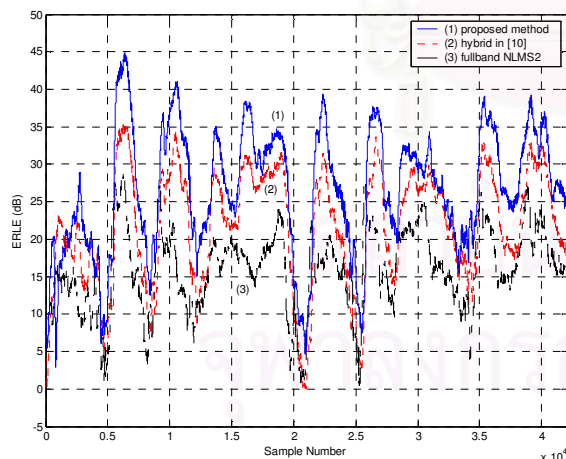


Fig. 8: Comparison of ERLE performance: (1) proposed method (FAP2 : $p=2$, $\mu=0.2$, F-NLMS2 and F-NLMS: $\mu=0.02$) (2) hybrid one in [10] (3) fullband SAEC (NLMS2: $\mu=0.3$)

CONCLUSIONS

A hybrid structure between FAP2 and F-NLMS algorithms, based on subband processing, has been proposed for SAEC. Simulation results based on speech signals indicate im-

proved performance of the proposed technique in terms of convergence rate and computational complexity reduction, as compared to the fullband structure. Without any signal decorrelation technique, the stereo perception of the loudspeaker signals in the receiving room is not degraded.

Table 5. Computational complexity of each algorithm

Adaptive Filter Algorithm :	\times/\div	$+/-$
a. Fullband FLS2	36904	36867
b. Fullband FAP2 ($p=2$)	32822	32768
c. Fullband NLMS2	8230	8198
d. [10] : FLS2 (0 – 1 kHz) + NLMS (1 – 8 kHz)	8551	8452
e. Proposed: FAP2 ($p=2$) (0 – 4 kHz) + F-NLMS (4 – 8 kHz)	18054	17658
f. Proposed: FAP2 ($p=2$) (0 – 1 kHz) + F-NLMS2 (1 – 4 kHz) + F-NLMS (4 – 8 kHz)	6518	5454

ACKNOWLEDGEMENTS

This work has been supported in part by the Cooperation Project between the Department of Electrical Engineering and Private Sector for Research and Development, Chulalongkorn University, Thailand.

REFERENCES

- [1] M. M Sondhi, et. al., "Stereo Acoustic Echo Cancellation -- An Overview of the Fundamental Problem," *IEEE Signal Processing Letters*, vol. 2, no.8, pp.148-151, Aug. 1995.
- [2] Haykin .S., *Adaptive Filter Theory*, 1996, Prentice Hall .
- [3] G. Schmidt, "Acoustic Echo and Noise Control for Low-Cost Processor," *DSP World Spring Design Conference*, pp. 1–28, Apr. 2000.
- [4] M. Bellanger, "Adaptive Digital Filtering and Signal Analysis," Marcel Dekker Inc, 1987.
- [5] J. Benesty, et. al., "Adaptive filtering algorithms for stereophonic acoustic echo cancellation," in *Proc. IEEE ICASSP'95*, pp. 3099-3102, May 1995.
- [6] G. V. Moustakidis, and S. Theodoridis, "Fast Newton Transversal Filters – A new class of adaptive estimation algorithms," *IEEE Trans. Signal Processing*, vol. 39, pp. 2184-2193, Oct. 1991.
- [7] A. Gilloire, and M. Vetterli, "Adaptive Filtering in Subbands with Critical Sampling: Analysis, Experiments, and Application to Acoustic Echo Cancellation," *IEEE Trans. Acoust. Speech and Signal Processing*, vol. 40, no. 8, pp. 1862-1875, Aug. 1992.
- [8] P. P. Vaidyanathan, "Multirate Systems and Filter Banks," Prentice Hall, 1993.
- [9] J. J. Shynk, "Frequency-domain and Multirate Adaptive Filtering," *IEEE Signal Processing Magazine*, vol. 9, no. 1, pp.14-37, Jan. 1992.
- [10] J. Benesty, et. al., "A Hybrid Mono/Stereo Acoustic Echo Canceller," *IEEE Trans. Speech and Audio Processing*, vol. 6, no. 5, pp. 468-475, Sep. 1998.
- [11] F. Amand, et. al., "A fast two-channel affine projection algorithm for stereophonic acoustic echo cancellation," in *Proc. IEEE ICASSP'96*, pp. 949-952, May, 1996.
- [12] C. Cowan and P. Grant, "Adaptive Filter", Prentice Hall, 1985.

Complexity Reduction of SAEC Systems using a Hybrid Mono/Stereo Structure

B. Krittanun and N. Tangsangiumvisai

Digital Signal Processing Research Laboratory, Department of Electrical Engineering,
Faculty of Engineering, Chulalongkorn University, Bangkok, 10330, Thailand.
Phone 0-2218-6491-2, 0-22186497, Fax 0-2251-8991, E-mail: Nisachon.T@eng.chula.ac.th

Abstract

This paper proposes an adaptive filtering algorithm for Stereophonic Acoustic Echo Cancellation (SAEC) in order to obtain computational complexity reduction and convergence rate improvement. The hybrid mono/stereo structure between Fast Affine Projection (FAP) and Frequency-domain Normalized Least Mean Square (F-NLMS) algorithms is introduced, based on subband processing. Simulation results on recorded speech signals suggest that convergence rate can be improved at low computational complexity.

Keywords: stereophonic acoustic echo cancellation, subband filtering, fast affine projection

1. General Information

Most of adaptive filtering algorithms for Stereophonic Acoustic Echo Cancellation (SAEC) systems are processing in time domain, such as the two-channel Least Mean Square (LMS2), two-channel Recursive Least Squares (RLS2) algorithms [1] – [3], etc. The tradeoff between convergence rate and computational complexity between LMS2 and RLS2 is achieved through a number of fast versions of RLS-type (FRLS) algorithms [4] – [6]. These algorithms, however, encounter the instability problem. Subband and frequency-domain implementations [7] – [9] are alternative approaches for system complexity reduction. Via the use of the decimation process, the sampling rate in each frequency band can be reduced. The number of adaptive filter taps in each subband is D times smaller than the fullband ones when D represents a decimation factor. Thus, lower computational complexity than the conventional SAEC systems can be achieved. In contrast to the fullband system, the eigenvalue spread of the autocorrelation input matrix, \mathbf{R} , in each subband can be reduced by processing the input signal into M subbands. Hence, faster convergence rate can potentially be obtained. In addition, block-based adaptation of the frequency-approach usually employs fast Fourier Transform (FFT) algorithm for computational efficiency.

A hybrid structure employing subband and frequency-domain processing is proposed in this paper. The proposed technique is based on the investigation of the hybrid mono/stereo structure in [10] for further improvement on the performance of the SAEC system, in terms of convergence rate, while keeping the computational complexity lower than the conventional fullband approach. For fast convergence speed, the proposed technique employs the two-channel version of the Fast Affine Projection (FAP2) algorithm [11]. In high-frequency regions, the

frequency-domain version of mono-channel Normalized Least Mean Square (F-NLMS) algorithm is chosen for lowering the computational complexity of the SAEC system, as compared to the conventional one. The Mean Square Error (MSE) in each subband is used as a criterion for algorithm selection between FAP2 and F-NLMS.

2. The Hybrid Structure for SAEC System

2.1 Hybrid Mono/Stereo AEC in [10]

Based on the psychoacoustical principle that stereo effect is due mostly to sound energy below about 1 kHz, two AEC systems are differently applied to low-frequency and high-frequency signals. The FRLS2 algorithm [5] is employed in the low-frequency subbands (100-900 Hz) in order to obtain fast convergence rate, when the sampling frequency (f_s) is 16 kHz. In addition, a decimation factor, D , is applied to each channel for further complexity reduction of the SAEC system. The sum of the stereo signals is highpass filtered (900-8000 Hz) to obtain mono-channel signal since little amount of speech signal energy is contained in the high-frequency region. The mono-channel NLMS algorithm is employed for complexity reduction, as compared to the conventional SAEC system that normally employs two-channel algorithms. Note that, the non-linear transformation (NL) is included in each channel in order to decorrelate partially the low-frequency stereo signals [10].

2.2 The Proposed Hybrid Structure for SAEC

By considering at the magnitude response of one channel of stereo speech signals with $f_s = 16$ kHz, as shown in Fig. 1, its significant energy is contained in frequencies below 2 kHz, and gradually decreases when the frequency reaches beyond 4 kHz. According to speech energy in each subband, adaptive filtering algorithms can be chosen to be employed in each subband differently. Algorithms with fast convergence rate should be applied to the particular subbands which contain the most significant amount of spectral density of speech inputs. On the other hand, algorithms with slow convergence rate, but with lower computational complexity, can be employed to those subbands with less spectral density of speech signal in order to reduce the computational complexity of the overall system.

The criterion for selection of fast or slow convergence rate adaptive filtering algorithms to be processed in each subband is suggested in this paper to be the average version of the MSE in each subband. For example, with the choice of $M = 8$ number of subbands, the MSE

plots for each subbands are illustrated in Fig. 2, when all subbands employ the two-channel NLMS (NLMS2) algorithm. It can be seen that, the MSE levels in band 0 up to band 3 are higher than those in band 4 to band 7. The threshold level of MSE criterion is chosen to be equal to -30 dB, whose magnitude is the output Signal-to-Noise Ratio (SNR) level of the added background noise at ‘Mic1’ in all simulations. For this particular case, the first 4 subbands will employ adaptive algorithms with fast convergence rate since the MSE values are greater than the selected threshold level. On the other hand, the adaptive algorithms with slower convergence rate but with lower complexity than the one chosen for the low-frequency subbands are operated in the last 4 bands.

In the proposed hybrid structure, the algorithm with fast convergence rate is chosen to be the two-channel version of the Fast Affine Projection (FAP2) [11]. The faster speed of convergence of FAP2 can be obtained by increasing the projection order, p . However, very high projection order results in large computational complexity of the SAEC system. For the low complexity adaptive algorithm employed in the high-frequency region, the mono-channel version of frequency-domain Normalized Least Mean Square (F-NLMS) algorithm is suggested, instead of the mono-channel NLMS algorithm as chosen in the hybrid structure in [10]. Hence, the proposed structure for SAEC can achieve faster convergence rate than the fullband one. Moreover, the computational complexity of the proposed hybrid structure should be lower than the subband SAEC system that employs two-channel algorithms in all subbands.

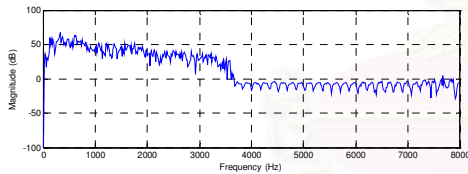


Fig. 1: Frequency response of a speech signal $x_1(n)$ (sampling frequency is $f_s = 16$ kHz).

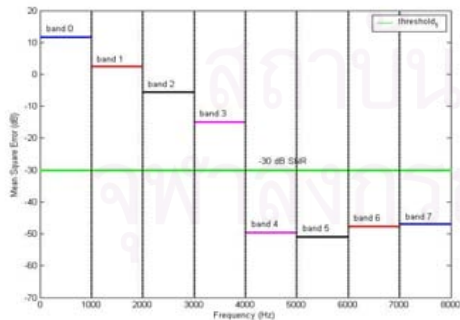


Fig. 2: The MSE values of each subband ($M = 8$).

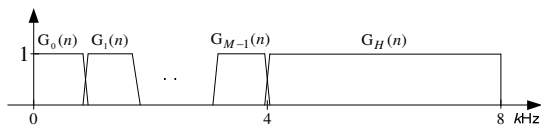


Fig. 3: Frequency response of the proposed AFB.

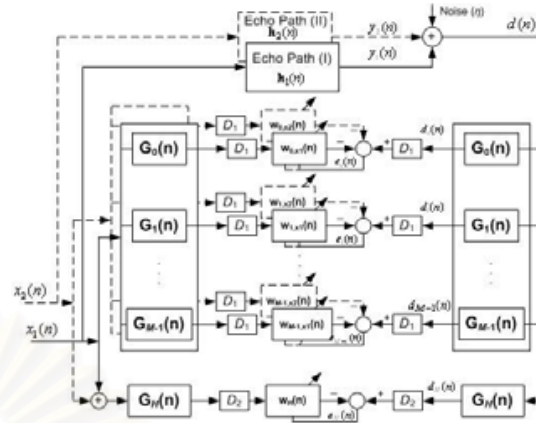


Fig. 4: The proposed hybrid structure for SAEC.

The Analysis Filter Bank (AFB) with slightly overlapping frequency responses, as depicted in Fig. 3, is employed for splitting the input signals in each channel, $x_i(n)$ for $i = 1, 2$. Additionally, another AFB is needed for the microphone signal in the receiving room, $d(n)$. With the chosen threshold level, the frequency responses of AFBs for low-frequency ($0 - 4$ kHz) signals, $G_i(z)$, for $i = 0, 1, \dots, M-1$, are as depicted in Fig. 3. Note that $G_0(z)$ is the prototype filter of AFB and other filters are obtained via the frequency-shifting method [12]. The decimation factor is $D_1 = 2M$ where $M = 4$ subbands. The frequency response $G_H(z)$ for the signals with higher frequencies than 4 kHz is a half-band highpass filter. In this final subband, the decimation factor is $D_2 = 2$. The block diagram of the proposed system is given in Fig. 4.

3. Computational Complexity Analysis

In this section, the computational complexity in mono-channel version of each algorithm, in terms of the number of real multiplications per input sample (RMPs) and number of additions and subtractions, is given in the Table 1 - 4. The adaptive filter length is chosen to be L whereas N_{tap} denotes the length of the prototype filter of AFB. The diagonal matrix of the step-size parameter of F-NLMS is represented by $M(k)$ and the block size of F-NLMS is $N = 2L$. Note that, in this paper, the two-channel Fast Least Squares (FLS2) [4] is used for the implementation of the hybrid structure in [10].

4. Experimental Results

Speech signals, sampled at 16 kHz, were normalized to have zero mean and unity variance, as depicted in Fig. 5, and used as the input signals $x_1(n)$ and $x_2(n)$ of the stereo teleconferencing system. The acoustic echo signals due to AEPs in the receiving room that were assumed to be time-invariant, were obtained as the convolution sum between the input signals and the L -tap impulse responses $\mathbf{h}_1(n)$ and $\mathbf{h}_2(n)$. The 30 dB SNR of background

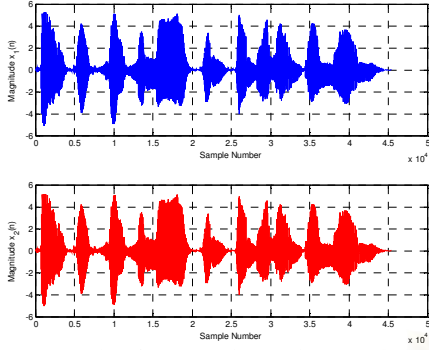


Fig. 5: Speech input signals.

noise, that was assumed to be of zero-mean and uncorrelated with both input signals, was added to the microphone signal in the receiving room. The length of adaptive filter was chosen to be $L = 2048$ for both AEC and SAEC parts. Note that, the adaptive filter length in each of the first 4 subbands and the final subband was equal to D_1 and D_2 respectively, where $N_{\text{tap}} = 64$.

For comparison, the performance of adaptive filters was evaluated in terms of Echo Return Loss Enhancement (ERLE), given by

$$\text{ERLE}(n) = 10 \times \log_{10} \left(\frac{\sum_{i=0}^{N_1} d^2(n-i)}{\sum_{i=0}^{N_1} e^2(n-i)} \right) \quad (1)$$

The proposed hybrid structure is subsequently investigated, when $f_s = 16$ kHz. The hybrid structure in [10] employs FLS2 in the first subband (0 – 1 kHz) and the mono-channel NLMS in the other subbands (1 – 8 kHz). The MSE value in each subband is calculated and compared with the threshold level of 30 dB. By employing FAP2 in the first subband of the proposed structure, it is found that all MSE values in the first 4 subbands are greater than the selected threshold. Hence, FAP2 is employed in the first 4 subbands. For this case, the overall system obtains much faster convergence rate than the conventional SAEC system, employing fullband NLMS2, as demonstrated in Fig. 6. The computational complexity is, however, greater than the conventional system. As compared to the hybrid structure in [10], similar performance in terms of convergence rate can be obtained, but the proposed technique gives larger complexity. Therefore, as a compromise between complexity and convergence speed performance of the system, FAP2 is only employed in the first subband whereas the two-channel F-NLMS (F-NLMS2) algorithm is used in the other subbands that MSE criterion is met, despite its inferior convergence rate. The ERLE performance is similar to that in [10], and is faster than the conventional system employing NLMS2, as shown in Fig. 7. The computational complexity of the proposed structure in this case is the lowest among other investigated technique. The computational complexity of each algorithm for all cases is given in Table 5, when 1 division is equivalent to 16 RMPs [8], and $\text{Inv}(M_{2 \times 2})$ equals to 22 RMPs and 1 addition.

Table 1: Fast Least Squares (FLS)

available at time: n	adaptive filter: $\mathbf{w}(n)$, forward prediction: $\mathbf{A}(n)$, backward prediction: $\mathbf{B}(n)$, data vector: $\mathbf{X}(n)$, adaptation gain: $\mathbf{G}(n)$, prediction error energy: $E_a(n)$, forgetting factor: W	
new data at time: n+1	input signal: $x(n+1)$, reference signal: $d(n+1)$	
Adaptation Gain Updating:	\times/\div	$+/-$
$e_a(n+1) = x(n+1) - \mathbf{A}^T(n)\mathbf{X}(n)$	L	L
$\mathbf{A}(n+1) = \mathbf{A}(n) + \mathbf{G}(n)e_a(n+1)/\alpha(n)$	$L+1\text{div}$	L
$E_a(n+1) = (E_a(n) + e_a(n+1)e_a(n+1)/\alpha(n))W$	$2+1\text{div}$	1
$\mathbf{G}_1(n+1) = \begin{bmatrix} 0 \\ \mathbf{G}(n) \end{bmatrix} + \frac{e_a(n+1)}{E_a(n+1)} \begin{bmatrix} 1 \\ -\mathbf{A}(n+1) \end{bmatrix} = \begin{bmatrix} \mathbf{M}(n+1) \\ \mathbf{m}(n+1) \end{bmatrix}$	$L+1+1\text{div}$	$L+1$
$e_b(n+1) = x(n+1-L) - \mathbf{B}^T(n)\mathbf{X}(n+1)$	L	L
$\mathbf{G}(n+1) = \mathbf{M}(n+1) + \mathbf{m}(n+1)\mathbf{B}(n)$	L	L
$\alpha_1(n+1) = \alpha(n) + e_a(n+1)e_a(n+1)/E_a(n+1)$	$1+1\text{div}$	1
$\alpha(n+1) = \alpha_1(n+1) - \mathbf{m}(n+1)e_b(n+1)$	1	1
$E_b(n+1) = (E_b(n) + e_b(n+1)e_b(n+1)/\alpha(n+1))W$	$2+1\text{div}$	1
$\mathbf{B}(n+1) = \mathbf{B}(n) + \mathbf{G}(n+1)e_b(n+1)/\alpha(n+1)$	$L+1\text{div}$	L
Adaptive Filter:		
$e(n+1) = d(n+1) - \mathbf{w}^T(n)\mathbf{X}(n+1)$	L	L
$\mathbf{w}(n+1) = \mathbf{w}(n) + \mathbf{G}(n+1)e(n+1)$	L	$L+1$

Table 2: Normalized Least Mean Square (NLMS)

Adaptive Filter :	\times/\div	$+/-$
$e(n) = d(n) - \mathbf{w}^T(n)\mathbf{X}(n)$	L	L
$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\varepsilon + \ \mathbf{X}(n)\ ^2} \mathbf{X}(n)e(n)$	$L+3+1\text{div}$	$L+3$

Table 3: Fast Affine Projection (FAP)

Adaptive Filter :	\times/\div	$+/-$
$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\mathbf{w}(n)$	pL	pL
$\phi(n) = \left[\mathbf{X}^T(n)\mathbf{X}(n) + \delta \mathbf{I} \right]^{-1} \mathbf{e}(n)$	$p^2(L+1) + \text{Inv}(M_{p \times p})$	p^2L
$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\varepsilon + \ \mathbf{X}(n)\ ^2} \mathbf{X}(n)\phi(n)$	$pL+1$	pL

Table 4: Fast Normalized Least Mean Square (F-NLMS)

Adaptive Filter :	\times/\div	$+/-$
FFT/IFFT	$5N \log_2 N$	$5(N-1)$
$\mathbf{E}(k) = \mathbf{D}(k) - \mathbf{W}(k)\mathbf{X}(k)$	$4N$	N
$\mathbf{W}(k+1) = \mathbf{W}(k) + \frac{\mathbf{M}(k)}{\varepsilon + \ \mathbf{X}(k)\ ^2} \mathbf{X}^H(k)\mathbf{E}(k)$	$N(12+1\text{div})$	N

Table 5: Computational Complexity of each algorithm

Adaptive Filter Algorithm :	\times/\div	$+/-$
a. Fullband FLS2	36904	36867
b. Fullband FAP2 ($p = 2$)	32822	32768
c. Fullband NLMS2	8230	8198
d. [10] : FLS2 (0 – 1 kHz) + NLMS (1 – 8 kHz)	8551	8452
e. Proposed: FAP2 ($p = 2$) (0 – 4 kHz) + F-NLMS (4 – 8 kHz)	18054	17658
f. Proposed: FAP2 ($p = 2$) (0 – 1 kHz) + F-NLMS2 (1 – 4 kHz) + F-NLMS (4 – 8 kHz)	6518	5454

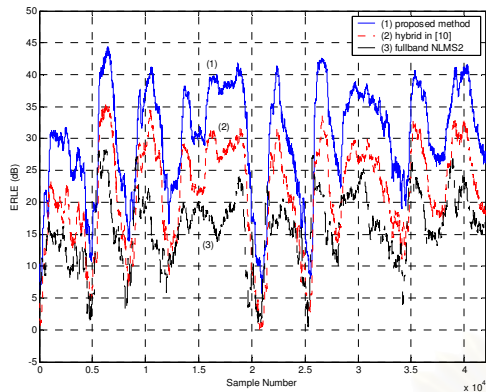


Fig. 6: Comparison of ERLE performance: (1) proposed method (FAP2 : $p = 2$, $\mu = 0.2$, F-NLMS: $\mu = 0.02$) (2) hybrid one in [10] (3) fullband SAEC (NLMS2: $\mu = 0.3$)

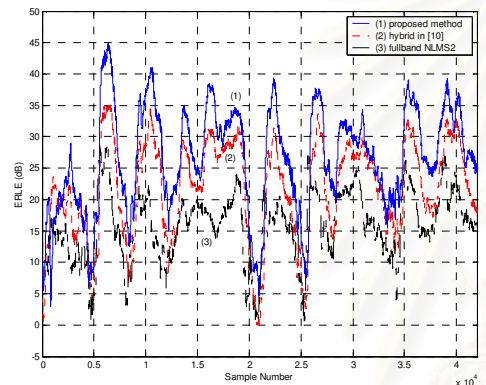


Fig. 7: Comparison of ERLE performance: (1) proposed method (FAP2 : $p = 2$, $\mu = 0.2$, F-NLMS2 and F-NLMS: $\mu = 0.02$) (2) hybrid one in [10] (3) fullband SAEC (NLMS2: $\mu = 0.3$)

5. Conclusions

A hybrid structure between FAP2 and F-NLMS algorithms, based on subband processing, has been proposed for SAEC. Simulation results based on speech signals indicate improved performance of the proposed technique in terms of convergence rate and computational complexity reduction, as compared to the fullband structure. Without any signal decorrelation technique, the stereo perception of the loudspeaker signals in the receiving room is not degraded.

6. Acknowledgements

This work has been supported in part by the Cooperation Project between the Department of Electrical Engineering and Private Sector for Research and Development, Chulalongkorn University, Thailand.

References

[1] M. M Sondhi, et. al., "Stereo Acoustic Echo Cancellation -- An Overview of the Fundamental Prob-

lem," IEEE Signal Processing Letters, vol. 2, no.8, pp.148-151, Aug. 1995.

[2] S. Haykin, "Adaptive Filter Theory," Prentice Hall, 1996.

[3] G. Schmidt, "Acoustic Echo and Noise Control for Low-Cost Processor," DSP World Spring Design Conference, pp. 1-28, Apr. 2000.

[4] M. Bellanger, "Adaptive Digital Filtering and Signal Analysis," Marcel Dekker Inc, 1987.

[5] J. Benesty, et. al., "Adaptive filtering algorithms for stereophonic acoustic echo cancellation," in Proc. IEEE ICASSP'95, pp. 3099-3102, May 1995.

[6] G. V. Moustakidis, and S. Theodoridis, "Fast Newton Transversal Filters - A new class of adaptive estimation algorithms," IEEE Trans. Signal Processing, vol. 39, pp. 2184-2193, Oct. 1991.

[7] A. Gilloire, and M. Vetterli, "Adaptive Filtering in Subbands with Critical Sampling: Analysis, Experiments, and Application to Acoustic Echo Cancellation," IEEE Trans. Acoust. Speech and Signal Processing, vol. 40, no. 8, pp. 1862-1875, Aug. 1992.

[8] P. P. Vaidyanathan, "Multirate Systems and Filter Banks," Prentice Hall, 1993.

[9] J. J. Shynk, "Frequency-domain and Multirate Adaptive Filtering," IEEE Signal Processing Magazine, vol. 9, no. 1, pp.14-37, Jan. 1992.

[10] J. Benesty, et. al., "A Hybrid Mono/Stereo Acoustic Echo Canceller," IEEE Trans. Speech and Audio Processing, vol. 6, no. 5, pp. 468-475, Sep. 1998.

[11] F. Amand, et. al., "A fast two-channel affine projection algorithm for stereophonic acoustic echo cancellation," in Proc. IEEE ICASSP'96, pp. 949-952, May, 1996.

[12] C. Cowan and P. Grant, "Adaptive Filter", Prentice Hall, 1985.



Nisachon Tangsangiumvisai received the M.Eng. degree in Electrical and Electronic Engineering and PhD. degree in Signal Processing from Imperial College, London, U.K. in 1997 and 2001, respectively. She currently serves as a lecturer at the Department of Electrical Engineering, Chulalongkorn University, Bangkok, Thailand. Her research interests include Adaptive Filters and Digital Signal Processing in Telecommunications.



Boonchai Krittanun received the B.Eng. degree in Electrical Engineering from Chulalongkorn University, Bangkok, Thailand in 2002. He is currently a master student in the department of Electrical Engineering, Chulalongkorn University, Bangkok, Thailand. His research interest is on Stereophonic Acoustic Echo Cancellation.

Improved Performance for Stereophonic Acoustic Echo Cancellation using Subband Structure

Boonchai Krittanun

Department of Electrical Engineering
Chulalongkorn University
Bangkok, Thailand
Boonchai.K@student.chula.ac.th

Nisachon Tangsangiumvisai

Department of Electrical Engineering
Chulalongkorn University
Bangkok, Thailand
Nisachon.T@chula.ac.th

Abstract—An adaptive filtering approach is proposed in this paper for Stereophonic Acoustic Echo Cancellation (SAEC) in order to obtain convergence rate improvement and computational complexity reduction. The hybrid mono/stereo structure between Fast Affine Projection (FAP) and Frequency-domain Normalized Least Mean Square (F-NLMS) algorithms is introduced, based on subband adaptive filtering. Simulation results on speech signals suggest that enhanced performance in terms of rate of convergence can be achieved at low computational cost.

I. INTRODUCTION

In two-channel voice communication systems, such as stereo teleconferencing systems, which provide better localization of talkers to audiences in the receiving room than the mono-channel systems, Stereophonic Acoustic Echo Cancellation (SAEC) is generally needed. In an SAEC system, as illustrated in Figure 1., four Finite Impulse Response (FIR) adaptive filters are normally employed to identify four impulse responses of unknown acoustic echo paths (AEPs) [1]. In this paper, the elimination of acoustic echo due to two AEPs between a pair of loudspeakers and one microphone (Mic1) in the receiving room are considered.

Having the highly correlated input signals, such as speech signals, the SAEC system usually obtains slow convergence rate, especially when employing the two-channel Least Mean Square (LMS2) algorithm. This is due to the fact that the LMS2 algorithm is sensitive to the condition number of the autocorrelation input matrix, \mathbf{R} [2]. Although the two-channel Recursive Least Squares (RLS2) algorithm is well-known to obtain very fast initial convergence rate in a stationary and low-noise environment, due to its convergence rate being essentially independent of the condition number of \mathbf{R} , its high computational complexity is, however, unsuitable for low-cost and real-time implementation [3]. Several fast versions of RLS-type algorithms have been proposed with lower complexity than RLS2 [4] – [6]. However, most of these algorithms might encounter the instability and slow convergence rate. Alternatively, an approach to reduce the computational complexity of the SAEC system is via the use of subband adaptive filtering [7], [8]. Due to the parallel processing of the input signals into small frequency subbands together with the decimation process, tap weights of adaptive

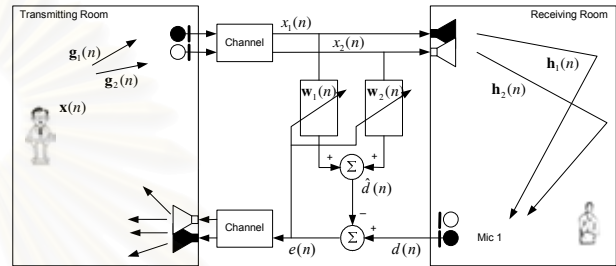


Figure 1. A diagram for Stereophonic Acoustic Echo Cancellation.

filters in each subband can be reduced, thus complexity reduction is obtained. Moreover, by dividing the input signals into small subbands, the eigenvalue spread of \mathbf{R} in each subband can be reduced, as compared to that of the fullband signals. Hence, faster convergence rate of the subband SAEC system than the conventional one is possible. Similarly, frequency-domain adaptive filtering, based on block processing, also offers complexity reduction, as compared to time-domain implementation. This is due to the Fast Fourier Transform (FFT) algorithm being normally employed for computational efficiency [9].

Unlike the Acoustic Echo Cancellation (AEC), strong correlation between the transmitted stereo signals within the SAEC system, however, prevents the two adaptive filters from identifying the two AEPs in the receiving room uniquely and correctly. Several techniques have been proposed to partially decorrelate the stereo signals so that the misalignment of the adaptive filters can be reduced. However, the distortion of stereo perception is another severe concern to degrade the quality of stereo signals. One example of the stereo decorrelation techniques is non-linear(NL)transformation [10].

Based on the observation that the stereo effect is due mostly to speech energy below about 1 kHz, the hybrid mono/stereo structure for SAEC is proposed in [11] to improve the performance of SAEC system, with lower computational complexity than the fullband system. From the advantages of the Fast Affine Projection (FAP) algorithm [12] that convergence rate can be improved by increasing the projection order, it is therefore proposed in this paper another hybrid structure for SAEC, based on subband adaptive

filtering. The low-frequency speech signals employ the two-channel FAP (FAP2) algorithm for fast rate of convergence. In higher frequency-region, many combinations of AEC and SAEC employing the frequency-domain Normalized Least Mean Square (F-NLMS) algorithms are investigated in order to obtain the best performance in terms of convergence rate, with the compromise of low complexity. Without any signal decorrelation technique, our proposed structure yields better performance in terms of convergence rate than the hybrid structure in [11]. Thus, the stereo perception of the received signals in the receiving room is not degraded.

This paper is organized as follows. Section II described the proposed structure for SAEC system, as compared to the hybrid mono/stereo one in [11]. Details of computational complexities of each algorithm involved in this paper are given in Section III, followed by simulation results in Section IV. Finally, the conclusions are given in Section V.

II. THE PROPOSED SAEC SYSTEM

First, the hybrid mono/stereo AEC [11] is summarized in section A. The proposed structure for SAEC is then presented in section B.

A. Hybrid mono/stereo AEC in [11]

Due to the existence of most of the energy in speech signals at frequencies below 1 kHz, two AEC systems are differently applied to low-frequency and high-frequency signals. With the sampling frequency (f_s) of 16 kHz, the stereo signals are lowpass filtered (100-900 Hz) and a fullband SAEC system is employed. The adaptive filtering algorithm chosen for the SAEC part is one of the fast versions of two-channel RLS (FRLS2) [5] in order to obtain fast convergence rate and hence, the acoustic echo signal can be eliminated efficiently. In addition, a decimation factor, D , is applied to each channel of the stereo signals for further complexity reduction of the SAEC system.

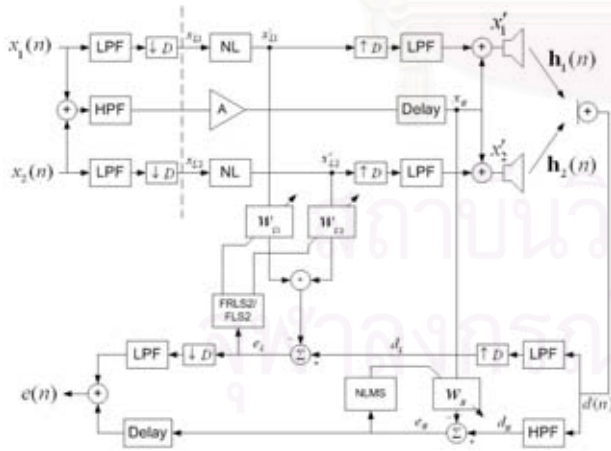


Figure 2. Hybrid Mono/Stereo AEC in [11]

For the high-frequency stereo signals, which contain less speech energy than the low-frequency ones, the mono-channel

signal is obtained from the highpass filtered (900-8000 Hz) of the sum of both stereo signals. The Normalized LMS (NLMS) algorithm is employed as the conventional AEC part to decrease the complexity of the overall system. The block diagram of the hybrid mono/stereo AEC system is shown in Figure 2. Note that, the NL transformation is added to stereo signals in each low-frequency channel for stereo signal decorrelation.

B. The proposed hybrid structure for SAEC

In this section, the main features of the proposed hybrid structure for SAEC are described. By considering at the magnitude response of a speech signal of $f_s = 16$ kHz in Figure 3, it can be seen that its significant energy is contained in the low frequencies below 1 kHz, and gradually decreases to 0 dB when the frequency reaches around 4 kHz. Hence, by dividing the input signals into small frequency subbands and processing each subband differently, e.g. using different adaptive filtering algorithms, the convergence rate can be improved while reducing the complexity of the overall system. Hence, three analysis filter banks (AFB) are required; two AFBs for the input signals in each channel, $x_i(n)$ for $i = 1, 2$, and one AFB for the microphone signal in the receiving room, $d(n)$. The AFBs for low-frequency (0 – 4 kHz) signals have their frequency response $G_i(z)$, for $i = 0, 1, \dots, M-1$, where M is the number of subbands, as depicted in Figure 4. Note that $G_0(z)$ is the prototype filter of AFB and other filters are obtained via the frequency-shifting method [13]. The final subband, which contains the signals in frequencies higher than 4 kHz, is obtained using a half-band highpass filter with frequency response $G_H(z)$. In order to fit the most significant part of spectral density of speech input signal into the first subband, the low-frequency signals are divided into $M = 4$ subbands. Thus, band 0, band 1, ..., band 4 are sometimes used to refer to these subbands. Since, band 1 (1 – 2 kHz) still contains significant energy of speech signals, less number of subbands, e.g. $M = 2$, is also investigated in Section IV to see if the performance of the SAEC system can be further obtained.

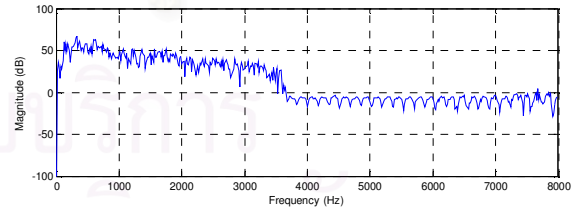


Figure 3. Frequency response of a speech signal $x_1(n)$ ($f_s = 16$ kHz).

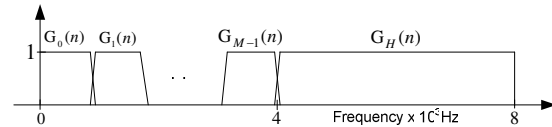


Figure 4. Frequency response of the proposed AFB.

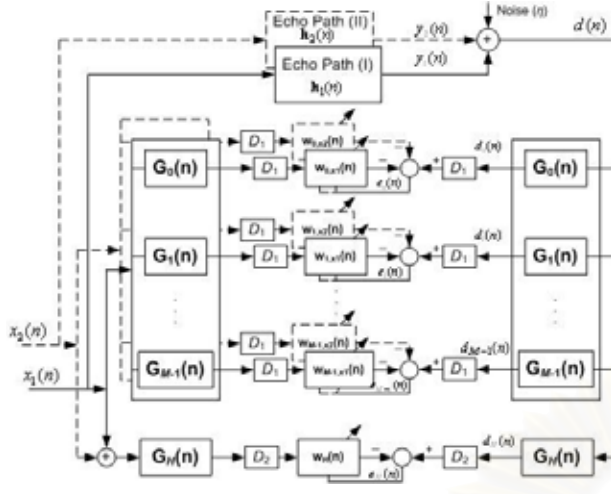


Figure 5. The proposed hybrid structure for SAEC

The adaptive filtering algorithm with very fast convergence rate is suggested to be employed for frequency subbands with high-dense speech spectral content. Thus, the two-channel FAP (FAP2) algorithm is employed in band 0. The faster speed of convergence of FAP2 can be obtained by increasing the projection order, p . However, the higher the order p , the more computational complexity of the SAEC system. As a compromise between complexity and convergence speed performance of the system, those $M-1$ subbands below 4 kHz employ the two-channel F-NLMS (F-NLMS2) algorithm. This is due to the fact that F-NLMS2 has lower computational complexity than that of FAP2 despite its inferior convergence rate. The decimation factor in the first M subbands is $D_1 = 2M$.

The stereo signals in the final subband of frequencies higher than 4 kHz are combined together since they contain negligible amount of speech spectra. The mono-channel F-NLMS algorithm is employed in this subband. This yields further complexity reduction of the overall system. In this final subband, the decimation factor is $D_2 = 2$. The block diagram of the proposed system is given in Figure 5.

III. ADAPTIVE FILTER ALGORITHMS AND COMPUTATIONAL COMPLEXITIES

In this section, the computational complexity of each algorithm in mono-channel version, in terms of the number of real multiplications per input sample (RMPs) and number of additions and subtractions, is given in the following tables. The adaptive filter length is chosen to be L whereas N_{tap} denotes the length of prototype filter of AFB. The diagonal matrix of the step-size of F-NLMS is represented by $\mathbf{M}(k)$ and the block size of the F-NLMS algorithm equals to $N = 2L$. Note that, in this paper, the two-channel Fast Least Squares (FLS2) [4] is used for the implementation of the hybrid structure in [11] instead of FRLS2.

TABLE I
FAST LEAST SQUARES (FLS)

available at time: n	adaptive filter: $\mathbf{w}(n)$, forward prediction: $\mathbf{A}(n)$, backward prediction: $\mathbf{B}(n)$, data vector: $\mathbf{X}(n)$, adaptation gain: $\mathbf{G}(n)$, prediction error energy: $E_a(n)$, forgetting factor: W		
new data at time: $n+1$	input signal: $x(n+1)$, reference signal: $d(n+1)$		
Adaptation Gain Updating:		\times/\div	$+/-$
$e_a(n+1) = x(n+1) - \mathbf{A}^T(n)\mathbf{X}(n)$		L	L
$\mathbf{A}(n+1) = \mathbf{A}(n) + \mathbf{G}(n)e_a(n+1)$		L	L
$e_a(n+1) = x(n+1) - \mathbf{A}^T(n+1)\mathbf{X}(n)$		L	L
$E_a(n+1) = W E_a(n) + e_a(n+1)e_a(n+1)$		2	1
$\mathbf{G}_1(n+1) = \begin{bmatrix} 0 \\ \mathbf{G}(n) \end{bmatrix} + \frac{\varepsilon_a(n+1)}{E_a(n+1)} \begin{bmatrix} 1 \\ -\mathbf{A}(n+1) \end{bmatrix} = \begin{bmatrix} \mathbf{M}(n+1) \\ \mathbf{m}(n+1) \end{bmatrix}$		$L+1+\text{div}$	$L+1$
$e_b(n+1) = x(n+1-L) - \mathbf{B}^T(n)\mathbf{X}(n+1)$		L	L
$\mathbf{G}(n+1) = \frac{1}{1-m(n+1)e_b(n+1)}(\mathbf{M}(n+1)+m(n+1)\mathbf{B}(n))$		$L+1+\text{div}$	$L+1$
$\mathbf{B}(n+1) = \mathbf{B}(n) + \mathbf{G}(n+1)e_b(n+1)$		L	L
Adaptive Filter:			
$e(n+1) = d(n+1) - \mathbf{w}^T(n)\mathbf{X}(n+1)$		L	L
$\mathbf{w}(n+1) = \mathbf{w}(n) + \mathbf{G}(n+1)e(n+1)$		L	$L+1$

TABLE II
Normalized Least Mean Square (NLMS)

Adaptive Filter :	\times/\div	$+/-$
$e(n) = d(n) - \mathbf{w}^T(n)\mathbf{x}(n)$	L	L
$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\varepsilon + \ \mathbf{x}(n)\ ^2} \mathbf{x}(n)e(n)$	$L+3+\text{div}$	$L+3$

TABLE III
Fast Affine Projection (FAP)

Adaptive Filter :	\times/\div	$+/-$
$e(n) = d(n) - \mathbf{X}^T(n)\mathbf{w}(n)$	pL	pL
$\phi(n) = [\mathbf{X}^T(n)\mathbf{X}(n) + \delta \mathbf{I}]^{-1} e(n)$	$p^3(L+1) + \text{Inv}(M_{\text{pp}})$	p^2L
$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\varepsilon + \ \mathbf{X}(n)\ ^2} \mathbf{X}(n)\phi(n)$	$pL+1$	pL

TABLE IV
Fast Normalized Least Mean Square (F-NLMS)

Adaptive Filter :	\times/\div	$+/-$
FFT/IFFT	$5N \log_2 N$	$5(N-1)$
$\mathbf{E}(k) = \mathbf{D}(k) - \mathbf{W}(k)\mathbf{X}(k)$	$4N$	N
$\mathbf{W}(k+1) = \mathbf{W}(k) + \frac{\mathbf{M}(k)}{\varepsilon + \ \mathbf{X}(k)\ ^2} \mathbf{X}^H(k)\mathbf{E}(k)$	$N(12+1\text{div})$	N

IV. EXPERIMENTAL RESULTS

Speech signals, sampled at 16 kHz, were normalized to have zero mean and unity variance, and used as the input signals $x_1(n)$ and $x_2(n)$ of the stereo teleconferencing system, as depicted in Figure 6. The acoustic echo signals due to AEPs $\mathbf{h}_1(n)$ and $\mathbf{h}_2(n)$ in the receiving room that were assumed to be time-invariant, was obtained as the convolution sum between the input signals and the L -tap $\mathbf{h}_1(n)$ and $\mathbf{h}_2(n)$. The background noise that was assumed to be of zero-mean and uncorrelated with $x_1(n)$ and $x_2(n)$ was added to the microphone signal in the receiving room, $d(n)$, with the output Signal-to-Noise Ratio (SNR) of 30 dB. The length of adaptive filter was chosen to be $L = 2048$ for both AEC and SAEC parts. Note that, the adaptive filter length in the first M subbands and the final subband was equal to L/D_1 and L/D_2 respectively, where $N_{\text{tap}} = 64$.

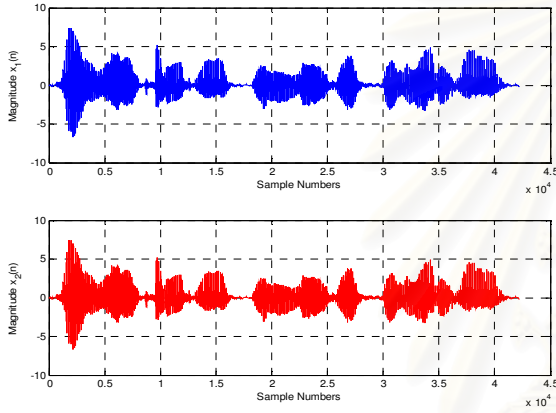


Figure 6. Speech Input Signals.

For comparison, the performance of adaptive filters was evaluated in terms of Echo Return Loss Enhancement (ERLE), given by

$$\text{ERLE}(n) = 10 \log_{10} \left(\frac{\sum_{i=0}^{N_j} d^2(n-i)}{\sum_{i=0}^{N_j} e^2(n-i)} \right) \quad (1)$$

where N_j was a small block to calculate the average.

First, by considering at the fullband structure of SAEC, it was found that FAP2 with $p = 2$, obtained faster convergence rate than FLS2, as shown in Figure 7. Moreover, the computational complexity of FAP2 was less than that of FLS2 (see Table V).

Then, for all cases of the proposed technique, F-NLMS was employed for the AEC part in the final subband. Due to the existence of speech energy in the interval of 1 – 3 kHz, i.e. band 1 to band 3, three cases were observed for different combinations of AEC and SAEC employing F-NLMS and F-NLMS2, respectively. Case I, $M = 4$ and the decimation factor was $D_1 = 8$, FAP2 with $p = 2$ was employed in band 0 while F-NLMS2 was employed in each subband from band 1 to band 3.

Case II, $M = 2$ and $D_1 = 4$, band 0 employed FAP2 and band 1 employed F-NLMS2. Case III was similar to Case II apart from that the projection order was increased to be $p = 3$. These cases were compared with the hybrid structure in [11] with $D=8$, where FLS2 was employed in band 0, and the mono-channel NLMS was used in the high-frequency band (1 – 8 kHz). (see Table V:c) Moreover, the hybrid structure with larger low-frequency region was investigated when $D=4$. (see Table V:d)

For Case I, the proposed structure yields faster convergence rate than the hybrid one in [11], as shown in Figure 8. In addition, the computational complexity of the proposed method is lower than that in [11]. (see Table V:c and V:e) Again, for Case II, improvement in terms of convergence rate and computational complexity of the proposed structure is obtained, as compared to that in [11]. (Table V:d and V:f). This is illustrated in Figure 9. In addition, the convergence performance of Case II is better than Case I due to the use of FAP2 in longer duration (0 - 2 kHz) than in Case I (0 - 1 kHz).

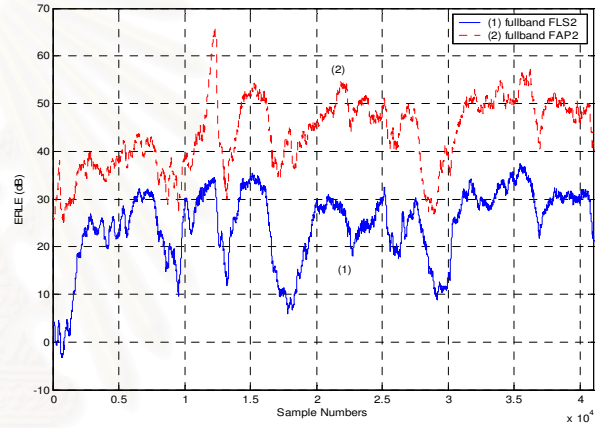


Figure 7. ERLE performance of FLS2 ($E_a(0) = 1200, W = 1$) and FAP2 ($p = 2, \mu = 0.5$). (fullband structure)

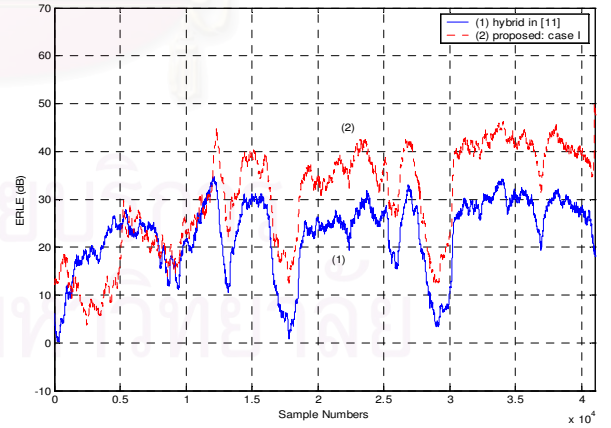


Figure 8. ERLE performance of the proposed technique (case I) (FAP2 : $p = 2, \mu = 0.3$, F-NLMS2 and F-NLMS : $\mu = 0.01$) as compared with that in [11] (FLS2 (0-1 kHz): $E_a(0) = 10, W = 0.999$, NLMS: $\mu = 0.03$).

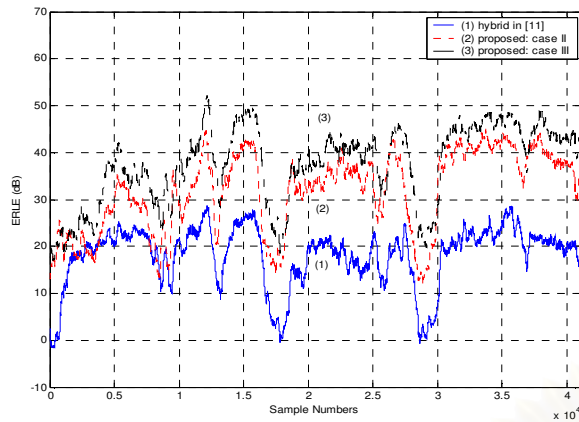


Figure 9. ERLE performance of the proposed technique (case II, FAP2 : $p = 2$, $\mu = 0.3$, F-NLMS2 and F-NLMS : $\mu = 0.01$, case III, FAP2 : $p = 3$, $\mu = 0.3$, F-NLMS2 and F-NLMS : $\mu = 0.01$) as compared with that in [11] (FLS2 (0-2 kHz): $E_a(0) = 10$, $W = 0.999$, NLMS: $\mu = 0.03$).

Case III demonstrates that the speed of convergence of the proposed structure can be further improved by increasing the projection order of FAP2. This, however, results in higher computational complexity than the $p = 2$ case, as shown in Table V:g. The computational complexity of each algorithm for all cases is given in Table V, when 1 division is equivalent to 16 RMPs [11], $\text{Inv}(M_{2 \times 2})$ equals to 22 RMPs and 1 addition, and $\text{Inv}(M_{3 \times 3})$ equals to 55 RMPs and 14 additions.

V. CONCLUSIONS

A hybrid structure between FAP2 and F-NLMS algorithms has been proposed in this paper for SAEC. By employing the subband structure of the SAEC system, computational complexity can be reduced, as compared to the fullband structure. A suitable choice of projection order for FAP2 is necessary to obtain convergence rate improvement while keeping the computational complexity low, as compared to that in [11]. Simulation results based on speech signals indicate improved performance of the proposed technique in terms of convergence rate via ERLE, as compared to that of the hybrid structure in [11]. Moreover, the computational complexity of the proposed structure is less than that in [11].

ACKNOWLEDGMENT

This work has been supported in part by the Cooperation Project between the Department of Electrical Engineering and Private Sector for Research and Development, Chulalongkorn University, Thailand.

TABLE V
Computational Complexity of each Algorithm

Adaptive Filter Algorithm :	\times/\div	$+/-$
a. Fullband FLS2	36904	36867
b. Fullband FAP2 ($p = 2$)	32822	32768
c. [11] : FLS2 (0 – 1 kHz) + NLMS (1 – 8 kHz)	9275	9221
d. [11] : FLS2 (0 – 2 kHz) + NLMS (2 – 8 kHz)	13883	13829
e. Case I : FAP2 ($p = 2$) + F-NLMS2 ($M = 4$) + F-NLMS (4 – 8 kHz)	6518	5454
f. Case II : FAP2 ($p = 2$) + F-NLMS2 ($M = 2$) + F-NLMS (4 – 8 kHz)	9516	8976
g. Case III : FAP2 ($p = 3$) + F-NLMS2 ($M = 2$) + F-NLMS (4 – 8 kHz)	16760	16158

REFERENCES

- [1] M. M. Sondhi, et. al., "Stereophonic Acoustic Echo Cancellation -- An Overview of the Fundamental Problem," *IEEE Signal Processing Letters*, vol. 2, no.8, pp.148-151, Aug. 1995.
- [2] Haykin .S.,*Adaptive Filter Theory*," 1996,Prentice Hall .
- [3] G. Schmidt, "Acoustic Echo and Noise Control for Low-Cost Processor," *DSP World Spring Design Conference*, pp. 1-28, Apr. 2000.
- [4] M. Bellanger, "Adaptive Digital Filtering and Signal Analysis," Marcel Dekker Inc, 1987.
- [5] J. Benesty, F. Amand, A. Gilloire, and Y. Grenier, "Adaptive filtering algorithms for stereophonic acoustic echo cancellation," in *Proc. IEEE ICASSP'95*, pp. 3099-3102, May 1995.
- [6] G. V. Moustakidis, and S. Theodoridis, "Fast Newton Transversal Filters – A new class of adaptive estimation algorithms," *IEEE Trans. Signal Processing*, vol. 39, pp. 2184-2193, Oct. 1991.
- [7] A. Gilloire, and M. Vetterli, "Adaptive Filtering in Subbands with Critical Sampling: Analysis, Experiments, and Application to Acoustic Echo Cancellation," *IEEE Trans. Acoust. Speech and Signal Processing*, vol. 40, no. 8, pp. 1862-1875, Aug. 1992.
- [8] A. Nakagawa, et. al., "Subband acoustic echo canceller using two different analysis filters and 8th order projection algorithm," in *Proc. IWAENC'97*, vol. 1, , pp. 140-143, Sep. 1997.
- [9] J. J. Shynk, "Frequency-domain and Multirate Adaptive Filtering," *IEEE Signal Processing Magazine*, vol. 9, no. 1, pp.14-37, Jan. 1992.
- [10] J. Benesty, D. R. Morgan and M. M. Sondhi, "A Better Understanding and an Improved Solution to the Specific Problems of Stereophonic Acoustic Echo Cancellation", *IEEE Trans. on Speech and Audio Processing*, vol. 6, no. 2, pp. 156-165, Mar. 1998.
- [11] J. Benesty, D. R. Morgan, and M. M. Sondhi, "A Hybrid Mono/Stereo Acoustic Echo Canceller," *IEEE Trans. Speech and Audio Processing*, vol. 6, no. 5, pp. 468-475, Sep. 1998.
- [12] F. Amand, J. Benesty, A. Gilloire and Y. Grenier "A fast two-channel affine projection algorithm for stereophonic acoustic echo cancellation," in *Proc. IEEE ICASSP'96*, pp. 949-952, May, 1996.
- [13] C. Cowan and P. Grant, "Adaptive Filter", Prentice Hall, 1985.

ประวัติผู้เขียนวิทยานิพนธ์

นายบุญชัย กฤตยานันต์ เกิดวันที่ 25 มิถุนายน พ.ศ. 2525 ที่จังหวัดกรุงเทพมหานคร เข้าศึกษาในหลักสูตรวิศวกรรมศาสตรบัณฑิต คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยในปีการศึกษา 2542 สำเร็จการศึกษาวิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยในปีการศึกษา 2545 เข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยในปีการศึกษา 2546



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย