

## โครงการวิจัยย่อยลำดับที่ 8

### เรื่อง การปรับปรุงคุณภาพการส่งสัญญาณเสียงบนเครือข่าย X (เฟสที่สอง)

#### Achieving high subjective voice quality in Voice over "X"-networks (Phase II)

#### 1. ผู้รับผิดชอบโครงการงาน ผศ. ดร. นิสาชล ตั้งเสงี่ยมวิสัย

ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ถนนพญาไท เขตปทุมวัน กรุงเทพฯ 10330

โทรศัพท์ 02-2186909 โทรสาร 02-2186912

e-mail : [Nisachon.T@chula.ac.th](mailto:Nisachon.T@chula.ac.th)

บทที่ 1

บทนำ

#### 1.1 ความสำคัญและที่มาของโครงการงาน

ในระบบการสื่อสารทางเสียงนั้น คุณภาพของการส่งสัญญาณเสียงบนเครือข่ายใด ๆ ถูกกำหนดด้วยความชัดเจนของสัญญาณเสียง และความหน่วงเวลา (Delay) เป็นปัจจัยหลัก โดยความหน่วงเวลาถูกระบุว่าเป็นเวลาที่เกิดขึ้นสำหรับผู้รับก่อนได้ยินเสียงของผู้พูด อย่างไรก็ตามสัญญาณเสียงสะท้อน (Echo) เกิดขึ้น โดยสามารถถูกแบ่งเป็นประเภท Acoustic Echo และประเภท Electrical Echo ขึ้นอยู่กับลักษณะการเกิดเสียงสะท้อนที่แตกต่างกัน ผู้พูดจะได้ยินเสียงสะท้อนของตนเมื่อความหน่วงเวลาหนึ่งรอบ (Round-trip Delay) ที่เกิดขึ้นในระบบการสื่อสารทางเสียงมีค่ามากกว่าประมาณ 50 – 150 มิลลิวินาที ทั้งนี้ขึ้นอยู่กับความดังของสัญญาณเสียงที่สะท้อนกลับมายังผู้พูด และค่าความหน่วงเวลาหนึ่งรอบก่อนที่ผู้พูดจะได้ยินเสียงสะท้อนของตนก็มีค่าแตกต่างกันไปตามเพศ วัย และความสามารถในการได้ยินของผู้พูด

ดังนั้น ปัจจัยหนึ่งของการปรับปรุงคุณภาพสัญญาณเสียงที่ส่งออกไป คือการกำจัดสัญญาณเสียงสะท้อนที่เกิดขึ้นในระบบการสื่อสารทางเสียงนั้นๆ ในโครงการวิจัยนี้ ได้พิจารณาถึงการกำจัดสัญญาณเสียงสะท้อนประเภท Acoustic Echo เท่านั้น (Acoustic Echo Cancellation, AEC) [1] ซึ่งได้ศึกษาแล้วในสองเฟสแรกของโครงการวิจัย “การกำจัดสัญญาณสะท้อนทางเสียงเมื่อมีสถานการณ์ Double-Talk ในระบบการสนทนาแบบสเตอริโอ” (1 มิถุนายน 2545 – 30 พฤษภาคม 2547) โดยทำการศึกษาการกำจัดสัญญาณสะท้อนทางเสียงในระบบการสนทนาแบบโมโน (ช่องสัญญาณเดียว) และแบบสเตอริโอ (สองช่องสัญญาณ) [1, 2, 8] ซึ่งพบว่าสามารถนำระบบ AEC ไปประยุกต์ใช้กับการกำจัดสัญญาณเสียงสะท้อนในระบบการสนทนาทางเสียงแบบสเตอริโอ (Stereophonic Acoustic Echo Cancellation, SAEC) [1, 2] ได้ หากแต่พบว่า ปัญหาสหสัมพันธ์ข้าม (Cross-correlation) ระหว่างสัญญาณเสียงแบบสเตอริโอในระบบ SAEC

จำเป็นต้องได้รับการแก้ไข สำหรับวิธีลดสหสัมพันธ์ข้าม (Decorrelation Techniques) ระหว่างสัญญาณในแต่ละช่องสัญญาณของสัญญาณเสียงแบบสเตอริโอ ได้แก่ การใช้วงจรมองไม่เชิงเส้น (Nonlinear Transformation) [2, 6] การใช้วงจรรองทุกความถี่แบบเปลี่ยนแปลงตามเวลา (Higher-Order Time-Varying Allpass Filters, HO-TV-APF) [3, 4, 7] และ การใช้สัญญาณรบกวนแบบปรับตัวได้ [10] เป็นต้น ดังรายละเอียดในรายงานฉบับสมบูรณ์ในสองเฟสแรกของโครงการวิจัย สำหรับเฟสที่สาม เป็นการศึกษาถึง “การปรับปรุงคุณภาพการส่งสัญญาณเสียงบนเครือข่าย X” (1 มิถุนายน 2547 – 30 พฤษภาคม 2548) เพื่อนำระบบ AEC และ SAEC มาใช้ในการกำจัดสัญญาณเสียงสะท้อนออกจากระบบการสื่อสารบนเครือข่าย “X” อาทิ เช่น VoIP VoFR VoATM เป็นต้น โดยได้ทำการศึกษาถึงการเข้ารหัสสัญญาณเสียงที่ทำการกำจัดเสียงสะท้อนแล้ว ด้วยตัวเข้ารหัสสัญญาณเสียง (Speech Coder) เช่น G.729 Codec เป็นต้น หรือตัวเข้ารหัสอื่นที่มีความเหมาะสมต่อคุณสมบัติของเครือข่าย “X” นั้น ๆ เพื่อรักษาไว้ซึ่งคุณภาพของสัญญาณเสียงในระบบ ภายใต้ความหน่วงเวลาที่ต่ำ สำหรับระบบการสื่อสารทางเสียงที่น่าสนใจระบบหนึ่งในปัจจุบัน คือ การสื่อสารทางเสียงบนเครือข่ายอินเทอร์เน็ต (Voice Over IP, VoIP) ดังนั้น มาตรฐานการเข้ารหัสสัญญาณเสียงพูดที่น่าสนใจเลือกศึกษาในเฟสที่สามของโครงการวิจัย คือ G.723.1 และ G.729 [11, 12]

ปัจจัยสำคัญอีกปัจจัยหนึ่งในการพิจารณาถึงคุณภาพของการส่งสัญญาณเสียง คือ ความชัดเจนของสัญญาณเสียง ดังนั้น ในเฟสที่สี่ของโครงการวิจัยนี้ จะศึกษาถึงเทคนิคต่างๆ ที่ใช้ในการกำจัดสัญญาณรบกวน [xx] และนำเทคนิคการกำจัดสัญญาณรบกวนที่เหมาะสมมาประยุกต์ใช้งานร่วมกับระบบกำจัดสัญญาณเสียงสะท้อน เพื่อพัฒนาประสิทธิภาพการกำจัดสัญญาณเสียงสะท้อนให้เพิ่มขึ้น และทำให้คุณภาพของสัญญาณเสียงที่ส่งออกไปในระบบมีความชัดเจนมากขึ้น ปราศจากเสียงสะท้อนที่หลงเหลืออยู่ หรือสัญญาณเสียงรบกวนใดๆ โดยจะเน้นไปที่ระบบ AEC เพียงอย่างเดียวก่อน สำหรับประสิทธิภาพการทำงานของ การกำจัดเสียงสะท้อนในโครงการนี้ จะอ้างอิงถึงมาตรฐาน ITU Standard G.165

นอกจากนี้ การทดสอบประสิทธิภาพการทำงานของ การกำจัดเสียงสะท้อนในเวลาจริง (Real-time) จะได้ทดลองสร้างจริง (Implementation) ระบบ AEC ลงบนบอร์ด FPGA (FPGA Board) รุ่น DS-KIT-2VP7FF672-EURO โดยขั้นตอนวิธี (Algorithm) ที่เลือกใช้ควบคุมการทำงานของวงจรรองแบบปรับตัว คือ ขั้นตอนวิธี Least Mean Square (LMS) และ / หรือขั้นตอนวิธี Normalised Least Mean Square (NLMS) เนื่องจากความซับซ้อนในการคำนวณที่ต่ำ [5] โดยในเฟสที่สี่นี้ จะพัฒนาบอร์ด FPGA ต่อจากเฟสที่สาม และเน้นถึงโครงสร้างที่ให้ความซับซ้อนในการสร้างจริงต่ำ

## 1.2 วัตถุประสงค์ของโครงการวิจัย

- 1.2.1 ศึกษา วิเคราะห์ และพัฒนาระบบ AEC/SAEC เพื่อกำจัดสัญญาณเสียงสะท้อนในการสื่อสารทางเสียงบนช่องสัญญาณต่างๆ ให้มีประสิทธิภาพการทำงานใกล้เคียงกับมาตรฐาน ITU Standard G.165

- 1.2.2 ศึกษาถึงความเป็นไปได้ในการกำจัดหรือลดผลกระทบของสัญญาณรบกวน เพื่อเพิ่มคุณภาพสัญญาณเสียงที่จะส่งไปบนเครือข่ายหนึ่งๆ โดยเน้นทางด้านความคมชัดของสัญญาณเสียงพูดเป็นพิเศษ
- 1.2.3 อาศัย FPGA Board ที่ออกแบบไว้ในเฟสก่อนหน้านี เพื่อทำการทดสอบในเวลาจริงได้

### 1.3 ขอบเขตของโครงการวิจัย

ในโครงการวิจัยนี้ จะเน้นไปที่การศึกษาความเป็นไปได้ในการกำจัดหรือลดผลกระทบของสัญญาณรบกวน เพื่อเพิ่มคุณภาพสัญญาณเสียงที่จะส่งไปบนเครือข่ายใดๆ โดยเน้นทางด้านความคมชัดของสัญญาณเสียงพูดเป็นพิเศษ ทั้งนี้ระบบ AEC ที่พัฒนาขึ้นจะต้องสามารถกำจัดสัญญาณเสียงสะท้อนในการสื่อสารทางเสียงบนช่องสัญญาณใดๆ ให้มีประสิทธิภาพการทำงานใกล้เคียงกับมาตรฐาน ITU Standard G.165 โดยพิจารณาประสิทธิภาพการทำงานของระบบที่ศึกษาผ่านการจำลองระบบบนคอมพิวเตอร์ (Simulation) นอกจากนี้ จะได้ทำการสร้างจริงระบบ AEC ลงบนบอร์ด FPGA เพื่อศึกษาประสิทธิภาพการทำงานจริงของระบบ AEC ในเวลาจริง

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

## บทที่ 2

### วิธีการดำเนินการวิจัย

ในเฟสที่สี่นี้ แผนงานวิจัยถูกจัดไว้เป็นระยะเวลา 1 ปี โดยจะถูกดำเนินการดังแสดงในตารางที่ 2-1 และมีการจ้างผู้ช่วยวิจัยจำนวน 3 คน (กำลังศึกษาในระดับปริญญาโท) ตลอดโครงการวิจัย

ตารางที่ 2-1 แผนการดำเนินงานของโครงการนี้ตลอดระยะเวลา 12 เดือน

กิจกรรม	ระยะเวลาปฏิบัติการ					
	เดือน	เดือน	เดือน	เดือน	เดือน	เดือน
	1-2	3-4	5-6	7-8	9-10	11-12
2.1 ศึกษาเทคนิคการกำจัดสัญญาณเสียงรบกวนแบบต่างๆ เพื่อหาเทคนิคที่เหมาะสมนำมาใช้ร่วมกับระบบ AEC	■	■				
2.2 จำลองระบบบนคอมพิวเตอร์ เทคนิคการกำจัดสัญญาณเสียงรบกวนที่คัดเลือกแล้วจากระยะที่ 1		■	■			
2.3 ทำการทดสอบระบบ AEC ที่พัฒนาขึ้นด้วยการสร้างจริงบนบอร์ด FPGA			■	■	■	
2.4 ปรับแต่งพารามิเตอร์ต่าง ๆ และเขียนรายงานสรุปผล					■	■

## 2.1 ศึกษาเทคนิคการกำจัดสัญญาณเสียงรบกวนแบบต่างๆ เพื่อหาเทคนิคที่เหมาะสมนำมาใช้ร่วมกับระบบ AEC

ในช่วง 4 เดือนแรกของโครงการ จะทำการศึกษา วิเคราะห์ ถึงเทคนิคต่างๆ ที่ใช้สำหรับกำจัดสัญญาณรบกวน เพื่อนำมาใช้งานร่วมกับระบบ AEC โดยใช้ฐานข้อมูลที่มีอยู่ทั่วไป เช่น เอกสารทางวิชาการ วารสารทางวิชาการ หรือรายงานจาก ห้องสมุด และอินเทอร์เน็ต เป็นต้น

## 2.2 จำลองระบบบนคอมพิวเตอร์เทคนิคการกำจัดสัญญาณเสียงรบกวนที่คัดเลือกแล้วจากระยะที่ 1

พิจารณาคัดเลือกเทคนิคการกำจัดสัญญาณเสียงรบกวนที่ศึกษาในระยะที่ 1 เพื่อทำการจำลองระบบบนคอมพิวเตอร์ ผ่านทางโปรแกรม MATLAB โดยทำการจำลองระบบการกำจัดสัญญาณเสียงรบกวนเพียงลำพัง กับการจำลองระบบการกำจัดสัญญาณเสียงรบกวนร่วมกับระบบ AEC

สำหรับสัญญาณเสียงที่ใช้ในการจำลองระบบบนคอมพิวเตอร์ เป็นสัญญาณเสียงพูดจากฐานข้อมูลที่มีอยู่แล้วในเฟสก่อนหน้า ซึ่งได้ทำการบันทึกเสียง ในสถานะแวดล้อมที่ไม่มีสัญญาณเสียงรบกวน ส่วนการเพิ่มสัญญาณเสียงรบกวนให้กับสัญญาณเสียงเข้าของระบบ AEC เพื่อทำการจำลองสถานะแวดล้อมต่างๆ ในบริเวณโดยรอบของผู้ใช้ให้สมจริงนั้น ทำโดยการเพิ่มสัญญาณรบกวนขาว (Additive White Gaussian Noise) เข้าไปในการจำลองระบบบนคอมพิวเตอร์ นอกจากนี้ อาจมีการบันทึกเสียงเพิ่มเติมสำหรับสัญญาณเสียงพูดในสถานะแวดล้อมที่จัดให้มีสัญญาณเสียงรบกวนด้วย

## 2.3 ทำการทดสอบระบบ AEC ที่พัฒนาขึ้นด้วยการสร้างจริงบนบอร์ด FPGA

ทดสอบการใช้งานของระบบ AEC ที่ทำงานร่วมกับเทคนิคการกำจัดสัญญาณเสียงรบกวนด้วยบอร์ด FPGA เพื่อประเมินผลประสิทธิภาพการกำจัดเสียงสะท้อนในเวลาจริง

## 2.4 ปรับแต่งพารามิเตอร์ต่าง ๆ และเขียนรายงานสรุปผล

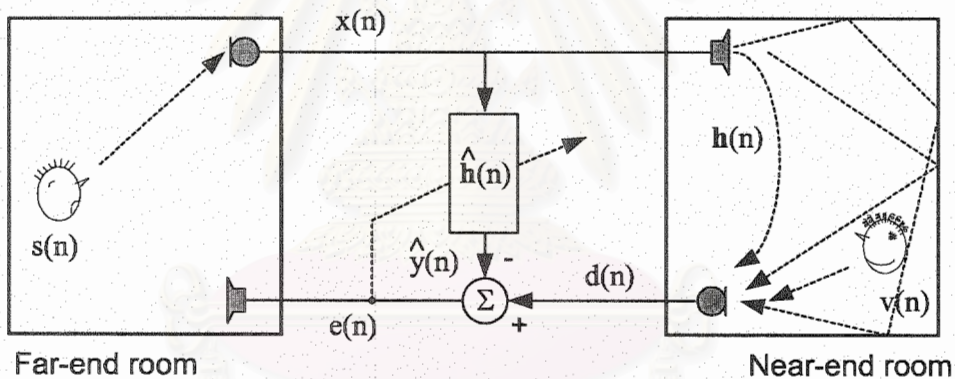
สรุปแนวทางในการปรับปรุงระบบที่นำเสนอในโครงการวิจัยนี้ เพื่อเพิ่มประสิทธิภาพการทำงาน เช่น ปรับปรุงอัตราการลู่เข้าให้เร็วขึ้น หรือลดความซับซ้อนในการคำนวณ (Computational Complexity) ของระบบโดยรวม และเขียนรายงานฉบับสมบูรณ์ของโครงการวิจัย

### บทที่ 3

#### ผลการดำเนินการ

การศึกษาประสิทธิภาพการทำงานของระบบกำจัดเสียงสะท้อนแบบ AEC ในโครงการนี้แบ่งออกเป็น 4 ส่วน คือ (1) ทบทวนถึงระบบ AEC ซึ่งได้ศึกษาแล้วในเฟสก่อนหน้าของโครงการ (2) ผลการศึกษาของระบบ SAEC (3) ศึกษาเทคนิคการกำจัดสัญญาณเสียงรบกวนแบบต่างๆ เพื่อหาเทคนิคที่เหมาะสมนำมาใช้งานร่วมกับระบบ AEC อันได้แก่ การกำจัดสัญญาณรบกวนแบบปรับตัว (Adaptive Noise Cancellation, ANC) และเทคนิคการลบออกของสเปกตรัม (Spectral Subtraction, SS) (4) การใช้งานร่วมกันของระบบ AEC กับเทคนิค ANC (5) การจำลองระบบบนคอมพิวเตอร์ของการทำงานร่วมกันของระบบ AEC กับเทคนิค ANC และ (6) การทดสอบระบบ AEC ที่พัฒนาขึ้นด้วยการสร้างจริงบนบอร์ด FPGA รายละเอียดจะได้กล่าวถึงในหัวข้อย่อต่อไปนี้

#### 3.1 ผลการศึกษาระบบ AEC



รูปที่ 1 การกำจัดเสียงสะท้อนในระบบการสนทนาแบบโมโน (AEC)

เมื่อพิจารณาทางฝั่งห้องรับ (Receiving Room) หรือเรียกว่าห้องใกล้ (Near-end Room) สัญญาณเสียงสะท้อนแบบ Acoustic Echo เกิดขึ้นจากการเชื่อมต่อทางเสียง (Acoustic Coupling) ระหว่างลำโพงกับไมโครโฟนภายในห้องใกล้ ดังแสดงในรูปที่ 1 ทั้งนี้ การกำจัดสัญญาณเสียงสะท้อนที่เกิดขึ้นภายในห้องใกล้นี้ (ระบบ AEC) สามารถทำได้โดยใช้วงจรกรองแบบปรับตัว (Adaptive Filter) เพื่อทำการจำลองวิถีสะท้อนทางเสียง (Acoustic Echo Path) และหักออกจากสัญญาณไมโครโฟนก่อนที่จะถูกส่งกลับไปยังคู่สนทนาในห้องส่ง (Transmitting Room) หรือเรียกว่าห้องไกล (Far-end Room) สำหรับการกำจัดสัญญาณเสียงสะท้อนที่เกิดขึ้นภายในห้องไกล ก็สามารถใช้อัตรากรองแบบปรับตัวอีกตัวหนึ่ง เพื่อทำการจำลองวิถีสะท้อนทางเสียงทางฝั่งห้องไกลได้เช่นกัน

เมื่อสัญญาณขาเข้าของระบบ AEC เป็นสัญญาณเสียงพูดจากห้องใกล้  $x(n)$  สัญญาณไมโครโฟนที่ห้องใกล้  $d(n)$  ประกอบด้วยเสียงสะท้อนที่เกิดจากวิถีสะท้อนทางเสียงในห้องใกล้  $h(n)$  ซึ่งถูกจำลองขึ้นจากการคอนโวลูชันสัญญาณขาเข้า  $x(n)$  กับวิถีสะท้อนทางเสียง  $h(n)$  ดังสมการ

$$d(n) = h^T(n)x(n) \quad (1)$$

โดยที่  $h(n) = [h_1(n), h_2(n), \dots, h_L(n)]^T$  เป็นเวกเตอร์สัมประสิทธิ์ของวิถีสะท้อนทางเสียงในห้องใกล้ที่มีขนาด  $L$  และ  $x(n) = [x(n), x(n-1), \dots, x(n-L+1)]^T$  เป็นเวกเตอร์สัญญาณเสียงขาเข้าที่มีขนาด  $L$  เช่นกัน โดยทั่วไปแล้ว จะจำลองให้สัญญาณไมโครโฟน  $d(n)$  มีสัญญาณรบกวนเบื้องหลัง (Background Noise) ภายในห้องใกล้รวมอยู่ด้วย เพื่อความสมจริงมากขึ้น สำหรับสัญญาณรบกวนเบื้องหลังนี้ สามารถจำลองได้โดยใช้สัญญาณรบกวนขาว (White Gaussian Noise)

สำหรับสัญญาณผิดพลาด  $e(n)$  จากการประมาณสัญญาณเสียงสะท้อนของระบบ AEC หาได้จากผลต่างของสัญญาณเสียงสะท้อนในห้องใกล้และค่าประมาณของสัญญาณเสียงสะท้อนที่เป็นสัญญาณออกของวงจรกรองแบบปรับตัว  $\hat{h}(n)$  ที่ทำการจำลองวิถีสะท้อนทางเสียงในห้องใกล้และมีจำนวนสัมประสิทธิ์เท่ากับ  $L$  เช่นกัน

$$e(n) = d(n) - \hat{h}^T(n)x(n) \quad (2)$$

โดยสัญญาณผิดพลาด  $e(n)$  เป็นสัญญาณขาออกของระบบ AEC และจะถูกส่งต่อไปยังห้องใกล้ ดังนั้นเมื่อวงจรกรองแบบปรับตัว  $\hat{h}(n)$  สามารถจำลองวิถีสะท้อนทางเสียงได้อย่างใกล้เคียง เสียงสะท้อนก็สามารถถูกกำจัดไปได้ในที่สุด

การปรับปรุงค่าสัมประสิทธิ์ของวงจรกรองแบบปรับตัวในโครงงานวิจัยนี้จะใช้ระเบียบวิธี Normalised Least Mean Square (NLMS) ดังสมการที่ (3)

$$\hat{h}(n+1) = \hat{h}(n) + \mu \frac{x(n)e(n)}{\varepsilon + \|x(n)\|^2} \quad (3)$$

เมื่อ  $\square$  คือค่าช่วงก้าว (step-size) ของระเบียบวิธี NLMS ที่ใช้ในการปรับการทำงานของวงจรกรองแบบปรับตัว และ  $\varepsilon$  เป็นค่าคงที่เล็กๆ เพื่อป้องกันการลู่ออกของวงจรกรองแบบปรับตัวเมื่อขนาดของสัญญาณเสียงเข้ามีขนาดเล็กมากในช่วงหนึ่งๆ

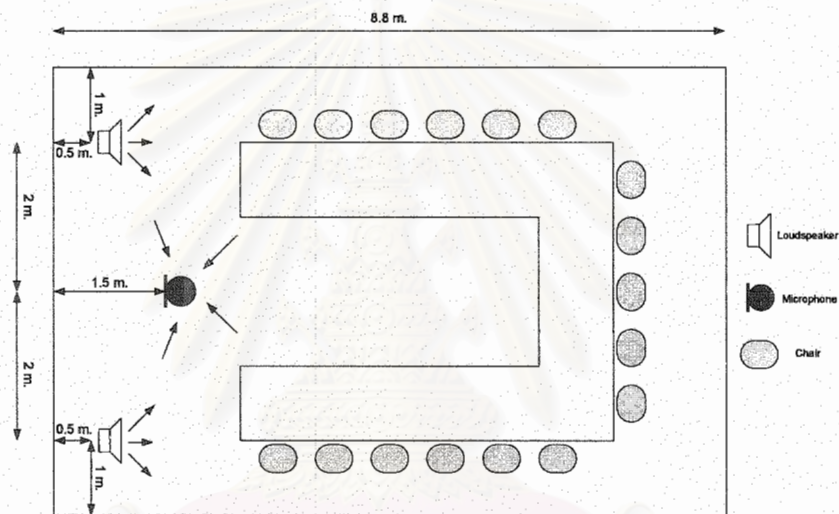
การศึกษามาตรฐาน ITU Standard G.165 สำหรับอ้างอิง และเปรียบเทียบประสิทธิภาพการทำงานของ การกำจัดสัญญาณเสียงสะท้อน สามารถดูรายละเอียดได้ในภาคผนวก ข

### 3.2 ผลการศึกษาระบบ SAEC

ในหัวข้อนี้จะกล่าวถึงการศึกษาวิธีสะท้อนทางเสียงในห้องสัมมนาทางไกลแบบสเตอริโอ และ โครงสร้าง Hybrid ที่ใช้กับระบบ SAEC

#### 3.2.1 การหาวิธีสะท้อนทางเสียงในห้องสัมมนาทางไกลแบบสเตอริโอ

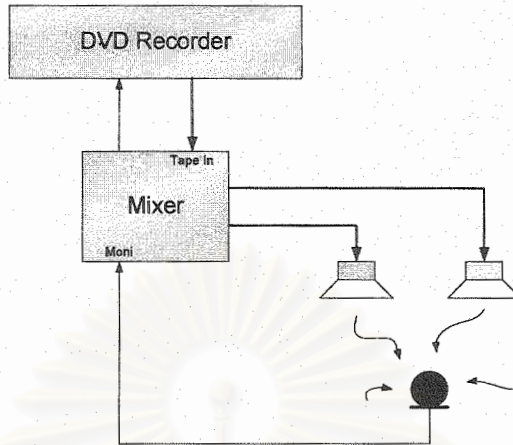
ได้ทำการจำลองสถานการณ์ของระบบการสัมมนาแบบสเตอริโอในห้องรับ (Receiving Room) และ วัดสัญญาณเสียงสะท้อนที่เข้าสู่ไมโครโฟน เมื่อสิ่งแวดล้อมภายในห้องรับเป็นดังรูปที่ 2



รูปที่ 2 แบบจำลองสถานการณ์ของระบบการสัมมนาแบบสเตอริโอในห้องรับ

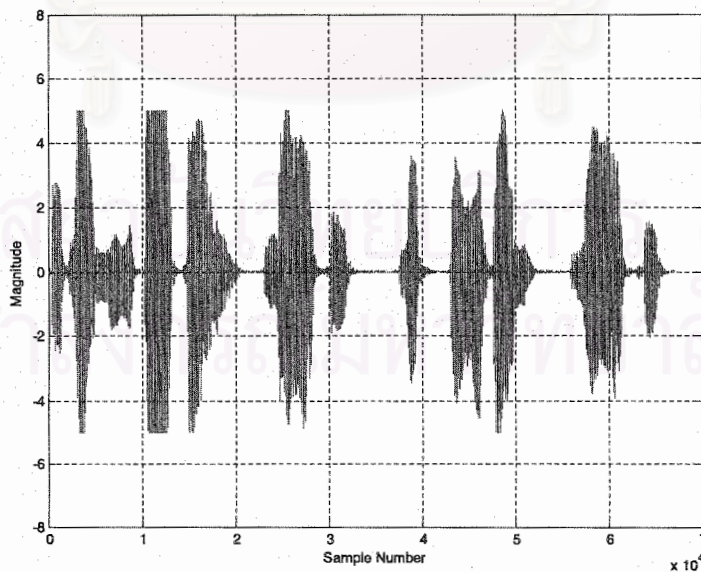
รูปที่ 2 แสดงรูปห้องรับขนาด กว้าง x ยาว x สูง = 4 m. x 8.8 m. 2.65 m. กำหนดให้ลำโพงและ ไมโครโฟนอยู่สูงจากพื้นห้อง = 1.5 m.



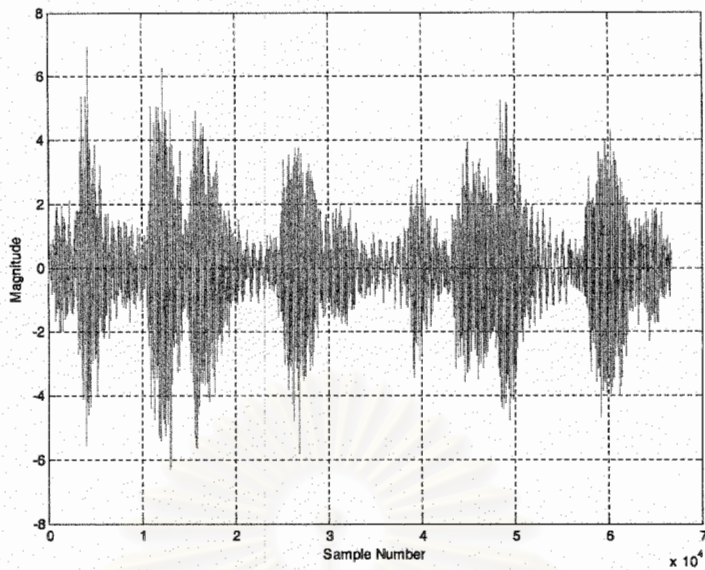


รูปที่ 3 รูปแสดงการจัดชุดทดลองการจำลองวิถีสะท้อนทางเสียง

การทดลองในหัวข้อนี้จะจัดชุดทดลองดังรูปที่ 3 โดยเริ่มด้วยการจำลองการสัญญาณเสียงพูดดังรูปที่ 4 ให้เล่นด้วยเครื่อง DVD Recorder สัญญาณเสียงพูดนี้จะถูกส่งผ่านเข้าเครื่อง Mixer ที่มีอัตราขยาย Tape In = 2 dB เพื่อให้สัญญาณเสียงพูดออกจากลำโพงทั้งสองตัวในห้องรับ และเกิดการสะท้อนกับสิ่งแวดล้อมภายในห้องรับ สัญญาณเสียงพูดที่เกิดการสะท้อนขึ้นนี้จะเป็นสัญญาณเข้าของไมโครโฟน เมื่อกำหนดให้ Moni ของ Mixer ที่ต่อมาจากไมโครโฟนมีอัตราขยาย = 5 dB จากนั้น Mixer จะส่งผ่านสัญญาณนี้ไปยัง DVD recorder เพื่อทำการบันทึกต่อไป โดยการบันทึกสัญญาณเข้าของไมโครโฟนจาก DVD recorder จะได้ผลการบันทึก ดังรูปที่ 5

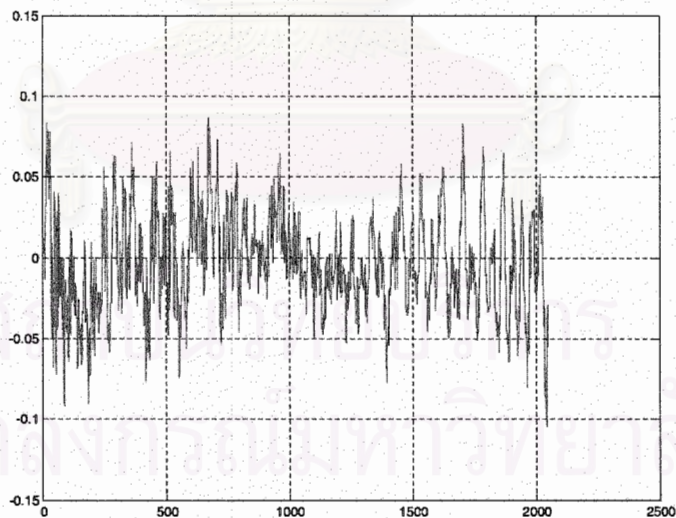


รูปที่ 4 สัญญาณเสียงเข้าของระบบจากลำโพงตัวที่ 1

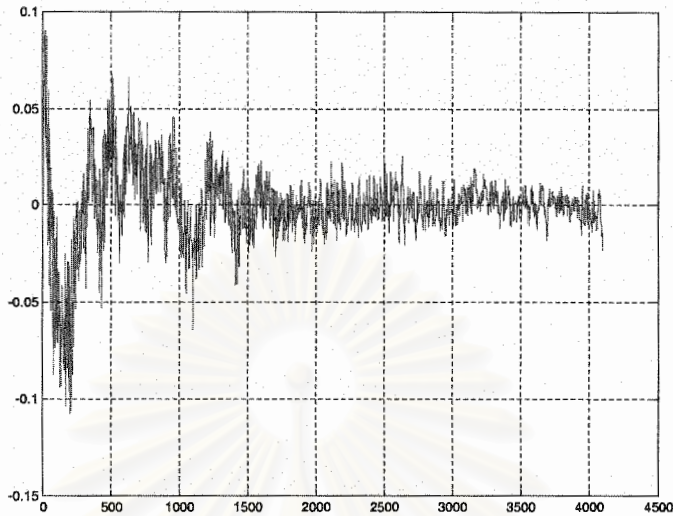


รูปที่ 5 สัญญาณไมโครโฟนในห้องรับ

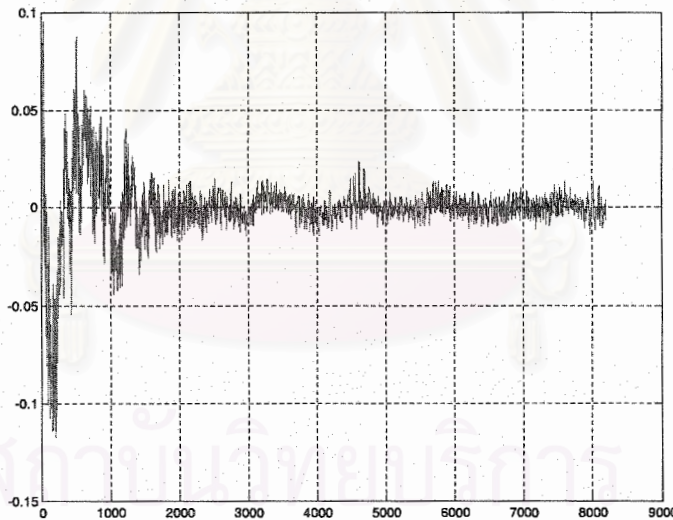
เมื่อได้สัญญาณเข้าของระบบและสัญญาณเข้าของไมโครโฟน จะใช้ขั้นตอนวิธี NLMS ในการปรับปรุงสัมประสิทธิ์ของวงจรกรองแบบปรับตัว เพื่อจำลองวิถีสะท้อนทางเสียงในห้องรับ โดยผลการจำลองวิถีสะท้อนทางเสียงเมื่อกำหนดให้จำนวนสัมประสิทธิ์ของวงจรกรองแบบปรับตัวเท่ากับ 1024, 2048 และ 4096 ค่า จะได้ผลการจำลองดังต่อไปนี้



รูปที่ 6 แบบจำลองวิถีสะท้อนทางเสียงเมื่อใช้วงจรกรองแบบปรับตัวจำนวน 1024 ค่า และ  $\mu_{NLMS} = 0.5$



รูปที่ 7 แบบจำลองวิธีสะท้อนทางเสียงเมื่อใช้วงจรกรองแบบปรับตัวจำนวน 2048 ค่า  
และ  $\mu_{NLMS} = 0.5$



รูปที่ 8 แบบจำลองวิธีสะท้อนทางเสียงเมื่อใช้วงจรกรองแบบปรับตัวจำนวน 4096 ค่า  
และ  $\mu_{NLMS} = 0.5$

จากการทดลอง พบว่า คุณลักษณะของวิธีสะท้อนทางเสียงจะมีค่ามากในช่วงเริ่มต้น และมีค่าลดลงเรื่อยๆ เมื่อค่าสัมประสิทธิ์มีค่ามากขึ้น นอกจากนี้ การเลือกจำนวนสัมประสิทธิ์ของวงจรกรองแบบปรับตัวตั้งแต่ 4096 ค่า ขึ้นไป จะสามารถจำลองวิธีสะท้อนทางเสียงที่ค่าวิธีสะท้อนช่วงท้ายมีลักษณะถูกลงได้ใกล้เคียงมากขึ้น แต่จากการจำลองวิธีสะท้อนด้วยเครื่อง DVD Recorder และ Mixer ซึ่งอุปกรณ์ทั้งสองตัวนี้จะมี Analog/Digital Converter และ Amplifier ภายในตัวอุปกรณ์ ซึ่งทำให้ขนาดของวิธีสะท้อนที่จำลองได้

คลาดเคลื่อน อีกทั้ง อัตราการซัดตัวอย่างของสัญญาณเข้าของระบบจากลำโพงและสัญญาณเข้าของไมโครโฟนมีค่าไม่เท่ากัน ทำให้การประมวลผลด้วยขั้นตอนวิธี NLMS อาจเกิดความคลาดเคลื่อนขึ้นได้

การหาค่าวิธีสะท้อนทางเสียงเมื่อสัญญาณเข้าและสัญญาณอ้างอิงถูกแปลงไปสู่โดเมนความถี่ (FFT)

กำหนดให้สัญญาณเข้าของระบบ SAEC คือ  $x(n)$  และสัญญาณเข้าสู่ไมโครโฟน คือ  $d(n)$  เมื่อวิธีสะท้อนทางเสียงในห้องรับ คือ  $h(n)$  ดังนั้น จะได้ความสัมพันธ์ระหว่าง  $x(n)$  กับ  $d(n)$  ดังนี้

$$d(n) = h(n) * x(n) \quad (4)$$

เมื่อทำการแปลงสมการที่ (4) ไปสู่โดเมนความถี่ด้วยฟังก์ชัน FFT จะได้ว่า

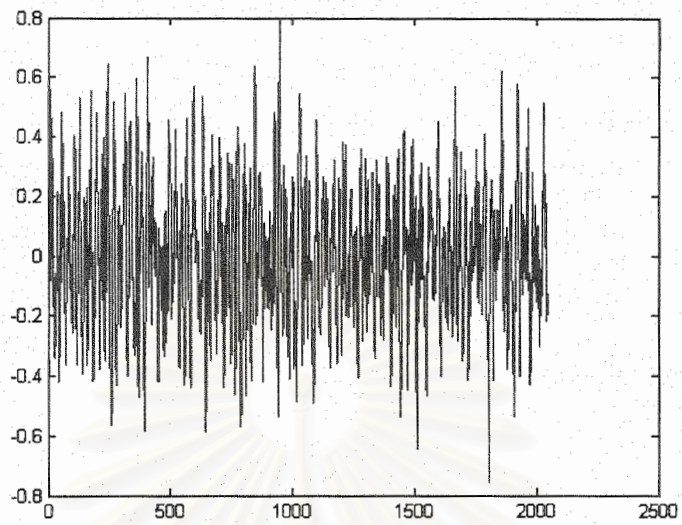
$$D(s) = H(s)X(s) \quad (5)$$

$$H(s) = \frac{D(s)}{X(s)} \quad (6)$$

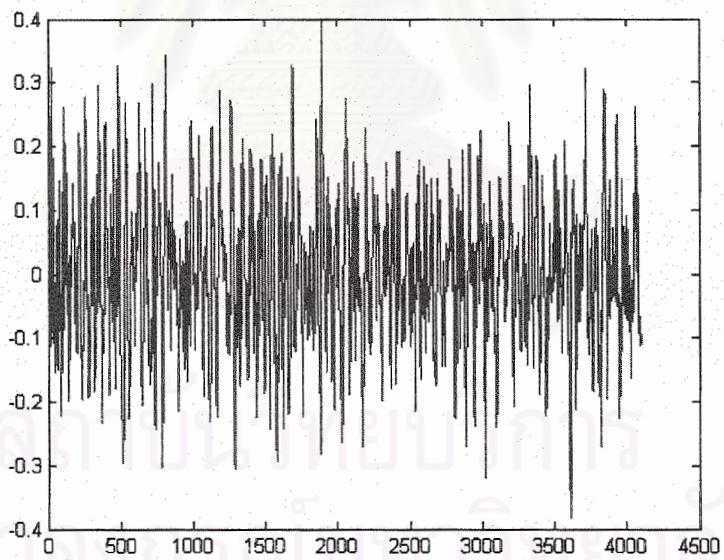
ดังนั้น จึงสามารถหาค่าวิธีสะท้อนทางเสียงได้จากการแปลงผกผันสมการที่ (6) ดังนี้

$$h(n) = \text{IFFT}\{H(s)\} = \text{IFFT}\left\{\frac{D(s)}{X(s)}\right\} \quad (7)$$

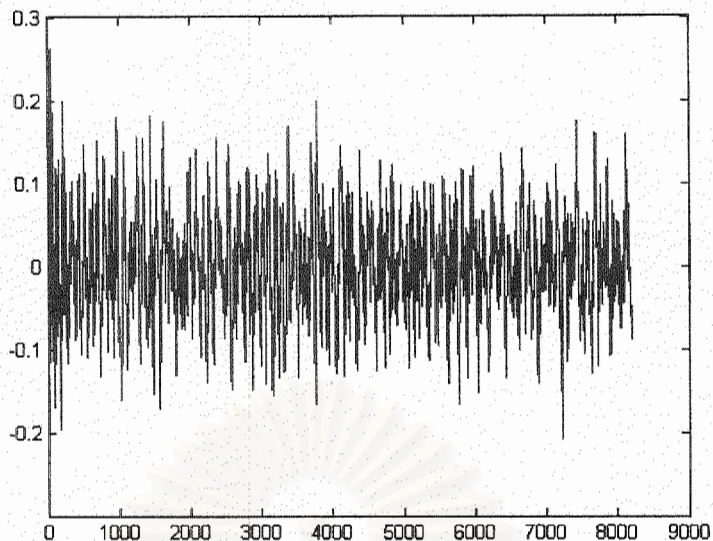
การหาค่าวิธีสะท้อนทางเสียง  $h(n)$  เมื่อสัญญาณเข้าของระบบ  $x(n)$  เป็นดังรูปที่ 4 และสัญญาณไมโครโฟน  $d(n)$  เป็นดังรูปที่ 5 จะได้ค่าลักษณะของวิธีสะท้อนทางเสียงที่มีค่าสัมประสิทธิ์จำนวน 2048, 4096, 8192 ค่า ดังนี้



รูปที่ 9 แบบจำลองวิถีสะท้อนทางเสียงจำนวน 2048 ค่า ด้วยการแปลง FFT



รูปที่ 10 แบบจำลองวิถีสะท้อนทางเสียงจำนวน 4096 ค่า ด้วยการแปลง FFT



รูปที่ 11 แบบจำลองวิถีสะท้อนทางเสียงจำนวน 8192 ค่า ด้วยการแปลง FFT

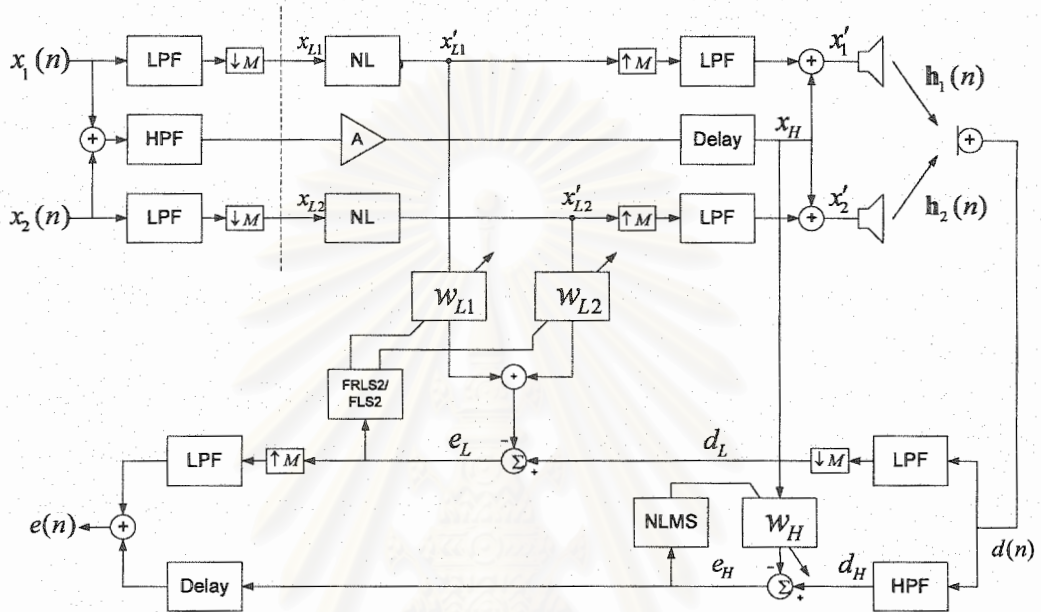
จากผลการทดลอง พบว่า ลักษณะของวิถีสะท้อนทางเสียงไม่ได้มีค่าคู่ลง เนื่องจากการแปลงสัญญาณเข้าและสัญญาณไมโครโฟนไปสู่โดเมนความถี่ จะไม่สามารถหาค่าวิถีสะท้อน ณ ขณะใดขณะหนึ่งได้ ดังนั้น วิถีสะท้อนทางเสียงที่ได้จากการแปลงฟูรีเยร์ผกผันที่ได้ จึงเป็นผลรวมของค่าวิถีสะท้อนทางเสียงทุกช่วงเวลา

### 3.2.2 โครงสร้าง Hybrid ที่ใช้กับระบบ SAEC

โครงสร้างไฮบริดเป็นการประยุกต์ใช้การกรองแบบย่อย (Subband Filtering) เพื่อแบ่งข้อมูลออกเป็นแถบความถี่ย่อยหลายแถบความถี่ จากนั้นจึงทำการการประมวลผลสัญญาณในแต่ละแถบความถี่ย่อย ด้วยการเลือกขั้นตอนวิธีที่เหมาะสมกับลักษณะของข้อมูลในแถบความถี่ย่อยนั้น

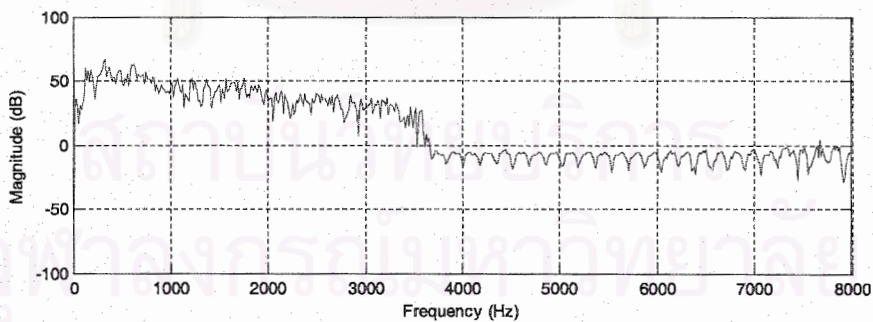
โครงสร้างไฮบริดที่ได้ศึกษาจากบทความ [27] ได้อ้างอิงจากหลักการจิตวิทยาทางเสียง ว่าพลังงานของสัญญาณเสียงเชิงสเตรียโอจะอยู่ในช่วงความถี่ที่ต่ำกว่า 1 kHz ดังนั้น ในบทความนี้จึงเลือกใช้ขั้นตอนวิธีที่ให้อัตราการลู่เข้าสู่ค่าตอบเร็ว (ขั้นตอนวิธี FRLS สองช่องสัญญาณ, FRLS2 [28]) ในการปรับปรุงสัมประสิทธิ์ของวงจรกรองแบบปรับตัวในช่วงความถี่ตั้งแต่ 0-1 kHz สำหรับการประมวลผลสัญญาณในแถบความถี่ที่สูงขึ้นซึ่งมีข้อมูลของสัญญาณเสียงอยู่น้อยกว่า จะทำการรวมสัญญาณสเตริโอเป็นช่องสัญญาณเดียวใช้ขั้นตอนวิธีที่ให้อัตราการลู่เข้าสู่ค่าตอบที่น้อยกว่า (ขั้นตอนวิธี NLMS) เพื่อลดความซับซ้อนทางการคำนวณของระบบโดยรวม โดยที่บล็อกโคอะแกรมของโครงสร้างไฮบริดในบทความที่ [27] สามารถแสดงได้ ดังรูปที่ 12

ทั้งนี้ ในบทความที่ [27] ได้ใช้การแปลงไม่เชิงเส้น (Non-linear Transformation, NL) ในการลดสหสัมพันธ์ (Correlation) ของสัญญาณเสียงเชิงสเตรโอในแถบความถี่ต่ำ เพื่อปรับปรุงการเข้าสู่ผู้ฟังของวงจรแบบปรับตัวในระบบ SAEC ให้ผู้เข้าสู่ผู้ฟังอย่างถูกต้องมากขึ้น



รูปที่ 12 บล็อกไดอะแกรมแสดงโครงสร้างไฮบริดในบทความที่ [27]

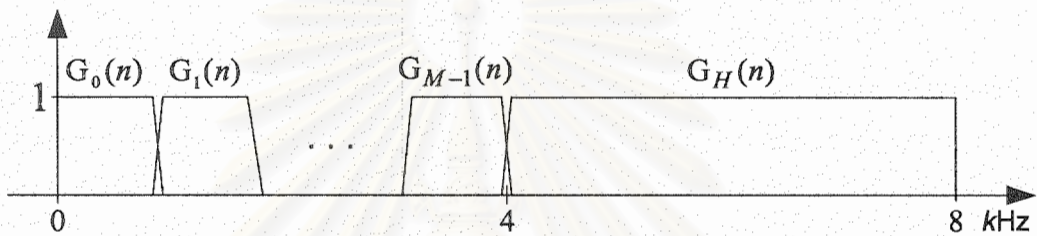
### โครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS



รูปที่ 13 ผลตอบสนองเชิงความถี่ของสัญญาณเสียงเข้าหนึ่งช่องสัญญาณ เมื่อความถี่การซัดตัวอย่างเท่ากับ 16 kHz

จากการศึกษาผลตอบสนองเชิงความถี่ของสัญญาณเสียงเข้าหนึ่งช่องสัญญาณ เมื่อความถี่การซัดตัวอย่างเท่ากับ 16 kHz พบว่า ความหนาแน่นของสเปกตรัมกำลัง (Power Spectral Density, PSD) ของสัญญาณเสียงพูดจะมีค่ามากในช่วง 0-4 kHz โดยมีค่ามากที่สุดในช่วง 0-1 kHz และค่า PSD จะมีขนาด

ลดลงจนมีค่าน้อยมากในช่วงความถี่ที่สูงขึ้นจนถึง 4 kHz ดังแสดงในรูปที่ 13 ดังนั้น ในหัวข้อนี้จึงได้นำเสนอโครงสร้างไฮบริดแบบใหม่เพื่อใช้ในระบบ SAEC โดยเลือกใช้ขั้นตอนวิธีที่เหมาะสมกับค่า PSD ของสัญญาณเสียงพูดในแต่ละแถบความถี่ย่อย โดยใช้วงจรกรองวิเคราะห้ในการแบ่งความหนาแน่นของสเปกตรัมกำลังของสัญญาณเสียงพูดออกเป็น 3 ส่วนด้วยกัน ได้แก่ ส่วนที่หนึ่ง คือ แถบความถี่ตั้งแต่ 0-1 kHz ซึ่งเป็นแถบความถี่ที่มีค่า PSD ของสัญญาณเสียงพูดมากที่สุด ส่วนที่สอง คือ แถบความถี่ตั้งแต่ 1-4 kHz ถึงแม้ว่าค่า PSD ของสัญญาณเสียงพูดในแถบความถี่นี้จะมีค่าน้อยกว่าในส่วนที่หนึ่งแต่ก็ยังมีค่ามากอยู่ และส่วนที่สาม คือ แถบความถี่ตั้งแต่ 4-8 kHz ในแถบความถี่นี้จะมีค่า PSD ของสัญญาณเสียงพูดอยู่น้อยมาก โดยที่ผลตอบสนองเชิงความถี่ของวงจรกรองวิเคราะห้ที่มีโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ได้แสดงดังรูปที่ 14 เมื่อกำหนดให้  $G_i(n)$  เมื่อ  $i = 0, \dots, M-1$  และ  $G_H(n)$  เป็นวงจรกรองวิเคราะห้ในแต่ละแถบความถี่ย่อย สำหรับรายละเอียดของการออกแบบวงจรกรองวิเคราะห้จะเป็นดังต่อไปนี้



รูปที่ 14 ผลตอบสนองเชิงความถี่ของวงจรกรองวิเคราะห้ที่มีโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS

ส่วนที่หนึ่ง คือ แถบความถี่ตั้งแต่ 0-1 kHz ซึ่งเป็นแถบความถี่ที่มีค่า PSD ของสัญญาณเสียงพูดมากที่สุดเมื่อพิจารณาจากค่าผลตอบสนองเชิงความถี่ของสัญญาณเสียงเข้าดังรูปที่ 12 ดังนั้น ในแถบความถี่ย่อยนี้จึงจำเป็นต้องใช้ขั้นตอนวิธีที่ให้อัตราการลู่เข้าสู่ค่าตอบเร็วที่สุดในการประมวลผลสัญญาณ โดยจะเลือกใช้ขั้นตอนวิธี AP2 [29] เนื่องจากมีอัตราการลู่เข้าสู่ค่าตอบที่เร็ว อีกทั้ง ขั้นตอนวิธี AP2 ยังมีข้อได้เปรียบกว่าขั้นตอนวิธีอื่น คือ สามารถปรับเพิ่มอัตราการลู่เข้าสู่ค่าตอบให้เร็วยิ่ง โดยการเมื่ออันดับของการฉาย (Projection Order,  $p$ ) ให้มากขึ้น และสามารถปรับลดอันดับของการฉาย ในกรณีที่ต้องการลดความซับซ้อนทางการคำนวณของระบบ SAEC

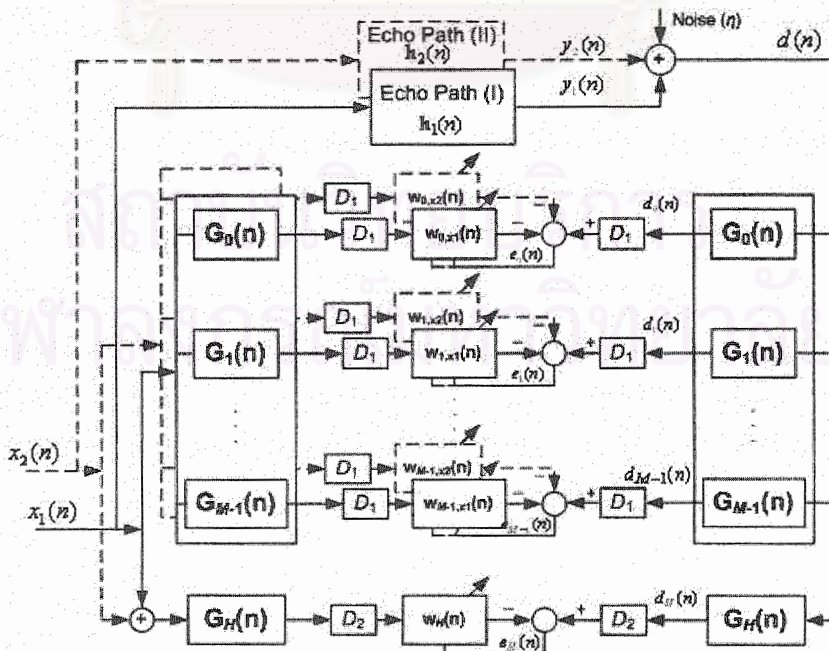
ส่วนที่สอง คือ แถบความถี่ตั้งแต่ 1-4 kHz ถึงแม้ว่าค่า PSD ของสัญญาณเสียงพูดในแถบความถี่นี้จะมีค่าน้อยกว่าในส่วนที่หนึ่งแต่ก็ยังมีค่ามาก ดังนั้น ในช่วงแถบความถี่ย่อยนี้ยังสามารถปรับปรุงอัตราการลู่เข้าสู่ของวงจรกรองแบบปรับตัวให้เร็วกว่าของโครงสร้างไฮบริดในบทความที่ [1] ที่ประมวลผลในแถบความถี่ย่อยนี้แบบช่องสัญญาณเดียว ในหัวข้อนี้ จึงเลือกใช้ขั้นตอนวิธีแบบสองช่องสัญญาณในการประมวลผลสัญญาณในแถบความถี่ย่อยนี้ เพื่อเพิ่มอัตราการลู่เข้าสู่ของวงจรกรอง



แบบปรับตัวให้เร็วยิ่งขึ้น โดยจะพิจารณาเลือกใช้ขั้นตอนวิธี F-NLMS2 [30] ในการประมวลผล อีกทั้ง ขั้นตอนวิธี F-NLMS2 ซึ่งประมวลผลสัญญาณ ใน โดเมนความถี่จะมีความซับซ้อนทางการคำนวณต่ำกว่าขั้นตอนวิธี NLMS2 ที่ประมวลผลในโดเมนเวลา โดยการแบ่งแถบความถี่ในส่วนนี้ สามารถแบ่งแถบความถี่ย่อยได้มากกว่าหนึ่งแถบความถี่โดยใช้คลังวงจรรองวิเคราะห์  $G_i(n)$  เมื่อ  $i=1, \dots, M-1$  ดังนั้นค่า PSD ของสัญญาณเสียงพูดในแถบความถี่ตั้งแต่ 1-4 kHz จะสามารถแบ่งได้เท่ากับ  $M-1$  แถบความถี่ย่อย เพื่อให้ได้สมรรถนะของการประมวลผลของโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่ดีที่สุด และมีความซับซ้อนทางการคำนวณที่ต่ำ

ส่วนที่สาม คือ แถบความถี่ตั้งแต่ 4-8 kHz ในแถบความถี่นี้จะมีระหว่าง AP2 และ F-NLMS ของสัญญาณเสียงพูดอยู่น้อยมาก จึงสามารถรวมสเปกตรัมกำลังของทั้งสองช่องสัญญาณเป็นช่องสัญญาณเดี่ยวแล้วประมวลผลด้วยขั้นตอนวิธีเพียงหนึ่งช่องสัญญาณเพื่อลดความซับซ้อนทางการคำนวณของระบบโดยรวม โดยไม่ทำให้เกิดความผิดพลาดในการประมาณวิถีสะท้อนมากนัก ขั้นตอนวิธีที่เลือกใช้ในแถบความถี่นี้ คือ ขั้นตอนวิธี F-NLMS เนื่องจากมีความซับซ้อนทางการคำนวณที่ต่ำกว่าขั้นตอนวิธี NLMS

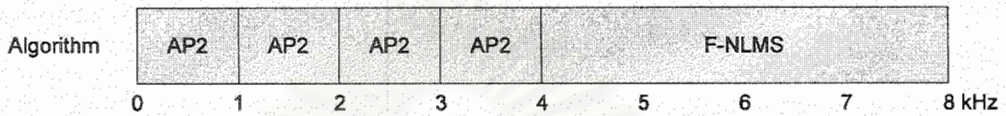
ดังนั้น ด้วยการแบ่งสเปกตรัมของสัญญาณเสียงพูดออกเป็นสามส่วนด้วยคลังวงจรรองวิเคราะห์ และประมวลผลสัญญาณในแต่ละแถบความถี่ย่อยดังกล่าว บล็อกโคอะแกรมแสดงโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS ที่นำเสนอ สามารถแสดงได้ดังรูปที่ 14 โดยที่  $D_1$  และ  $D_2$  เป็นตัวประกอบเดซิเมชัน ทั้งนี้ โดยทั่วไปแล้ว  $D_1$  จะมีค่าเท่ากับ  $2M$  และ  $D_2$  จะมีค่าเท่ากับ 2



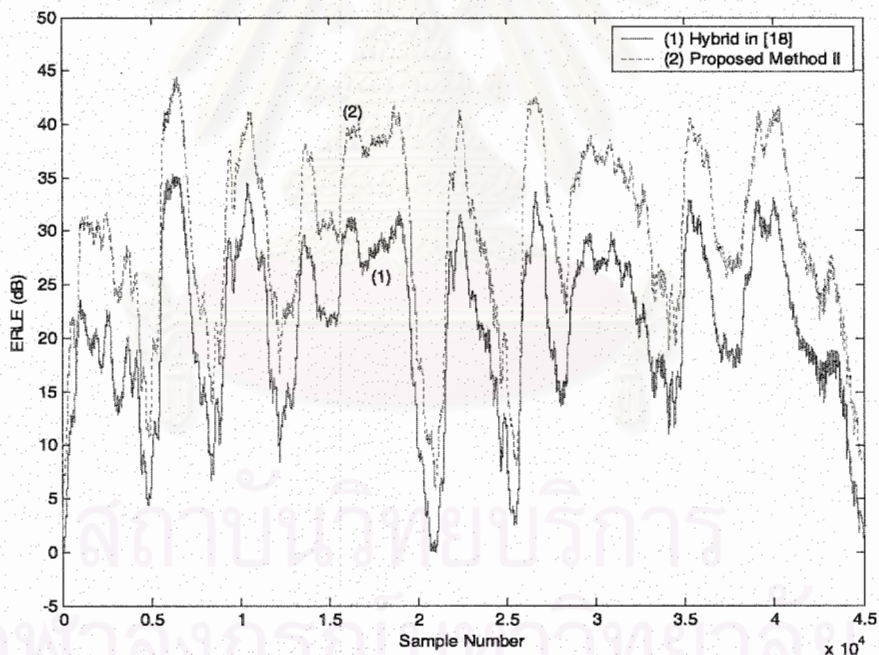
รูปที่ 15 บล็อกโคอะแกรมแสดงโครงสร้างไฮบริดระหว่าง AP2 และ F-NLMS

ผลการจำลองแบบระบบที่มีโครงสร้างไฮบริด

ในการจำลองแบบต่อไปนี้จะได้นำเสนอการใช้ขั้นตอนวิธี AP2 ในการปรับปรุงสัมประสิทธิ์ของวงจรกรองแบบปรับตัวในแถบความถี่ย่อยที่ 0, 1, 2 และ 3 ตามลำดับ โดยการแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอจะเป็นดังแสดงในรูปที่ 16 ค่า ERLE ที่สังเกตได้แสดงในรูปที่ 17



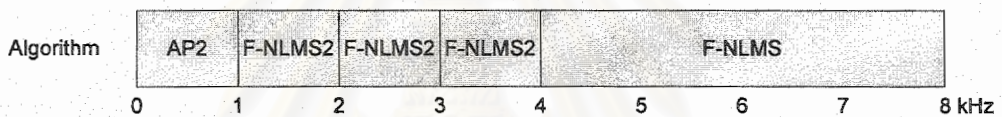
รูปที่ 16 การแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอ



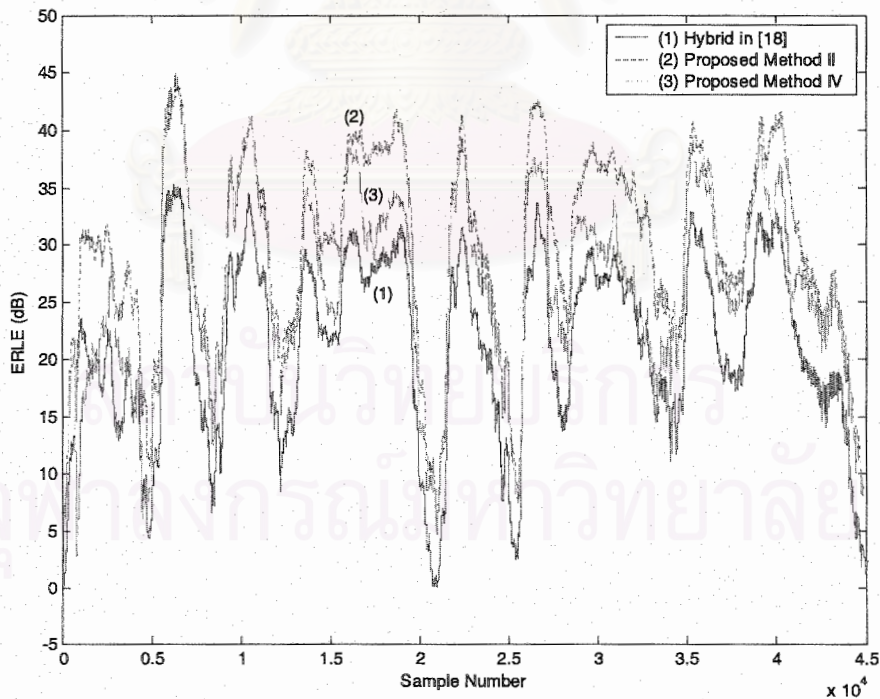
รูปที่ 17 ค่า ERLE เปรียบเทียบสมรรถนะระหว่าง (1) โครงสร้างไฮบริดในบทความที่ [27] (2) โครงสร้างไฮบริดที่นำเสนอระหว่าง AP2 (0-4 kHz):  $p=2$ ,  $\mu = 0.3$  และ F-NLMS (4-8 kHz):  $\mu = 0.02$

จากรูปที่ 17 พบว่าโครงสร้างไฮบริดที่นำเสนอ มีอัตราการเข้าสู่ค่าคอบของวงจรรองแบบปรับตัวเร็วกว่าโครงสร้างในบทความที่ [27] แต่ความซับซ้อนทางการคำนวณจะมีค่าสูงขึ้นมากดังตารางที่ 1 (b) เนื่องจากการเลือกใช้ขั้นตอนวิธี AP2 ในการประมวลผลสัญญาณแถบความถี่ย่อยที่ 0-3

ต่อไปจะนำเสนอโครงสร้างไฮบริดที่ลดความซับซ้อนทางการคำนวณให้ต่ำลง โดยมีโครงสร้างดังรูปที่ 18 เมื่อแถบความถี่ย่อยที่ 1, 2 และ 3 ประมวลผลด้วยขั้นตอนวิธี F-NLMS2 ที่มีความซับซ้อนทางการคำนวณที่ต่ำกว่าขั้นตอนวิธี FAP2 จากรูปที่ 19 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):  $\mu=0.02$  เป็นโครงสร้างไฮบริดที่ให้อัตราการเข้าสู่ค่าคอบได้เร็วขึ้น อีกทั้ง ความซับซ้อนทางการคำนวณของโครงสร้างไฮบริดที่นำเสนอนี้ยังต่ำมาก ดังนั้น โครงสร้างไฮบริดแบบที่สี่ที่นำเสนอนี้จึงเป็นอีกทางเลือกหนึ่งสำหรับใช้ในการปรับปรุงสัมประสิทธิ์ของวงจรรองแบบปรับตัวที่นำเสนอขึ้นในโครงการวิจัยนี้



รูปที่ 18 การแบ่งแถบความถี่ย่อยสำหรับการประมวลผลด้วยโครงสร้างไฮบริดที่นำเสนอ



รูปที่ 19 ค่า ERLE เปรียบเทียบสมรรถนะระหว่าง (1) โครงสร้างไฮบริดในบทความที่ [27]

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

เมื่อกำหนดให้  $Inv(M_{2 \times 2})$  มีความซับซ้อนทางการคำนวณเท่ากับ 22 RMPs และ 1 ADSs, และ  $Inv(M_{3 \times 3})$  เท่ากับ 55 RMPs และ 14 ADSs. [5] ค่าความซับซ้อนทางการคำนวณของแต่ละขั้นตอนวิธีสามารถสรุปได้ดังตารางที่ 1

ตารางที่ 1 ความซับซ้อนทางการคำนวณของแต่ละขั้นตอนวิธีที่ศึกษาในโครงการวิจัยนี้

Adaptive Filter Algorithm :	$\times / \div$	$+ / -$
a. FLS2 (0 – 1 kHz) + NLMS (1 – 8 kHz)	9275	9221
b. Proposed II: AP2 (p = 2) (0 – 4 kHz) + F-NLMS (4 – 8 kHz)	18054	17658
c. Proposed IV: AP2 (p = 2) (0 – 1 kHz) + F-NLMS2 (1 – 4 kHz) + F-NLMS (4 – 8 kHz)	6518	5454

จากตารางที่ 1 พบว่าโครงสร้างไฮบริดแบบ c จะมีความซับซ้อนทางการคำนวณต่ำที่สุด เมื่อพิจารณาจากค่าการคูณจำนวนจริง (RMPs) และค่าการบวกและลบจำนวนจริง (RASs) และมีค่าต่ำกว่าโครงสร้างไฮบริดในบทความที่ [27]

### 3.3 ผลการศึกษาเทคนิคการกำจัดสัญญาณเสียงรบกวน

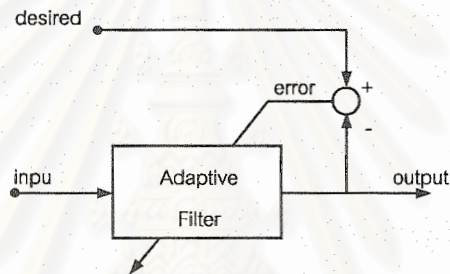
ในหัวข้อย่อนี้ จะกล่าวถึงผลการศึกษาเทคนิคต่างๆ ที่ใช้ในการกำจัดสัญญาณเสียงรบกวน ได้แก่ การกำจัดสัญญาณรบกวนแบบปรับตัว (Adaptive Noise Cancellation, ANC) และเทคนิคการลบออกของสเปกตรัม (Spectral Subtraction, SS)

#### 3.3.1 ระบบ ANC

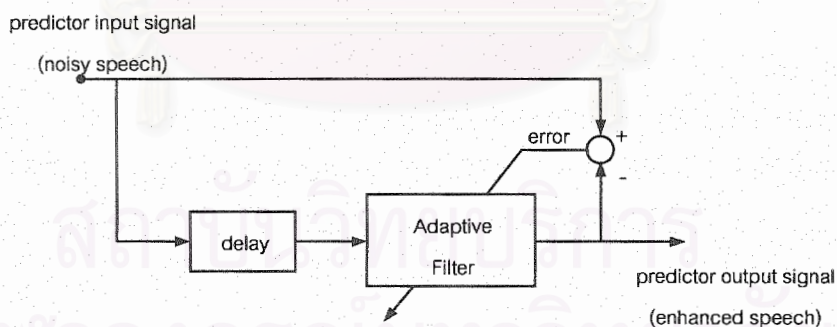
ความสามารถในการเข้าใจในคำพูดของมนุษย์ถูกจำกัดเมื่อมีสัญญาณรบกวนพื้นหลังเข้ามารบกวน สำหรับผู้ที่มีความสามารถในการได้ยินในระดับปกติสามารถเข้าใจในเสียงที่ถูกรบกวนได้ เมื่อระดับกำลังของสัญญาณรบกวนมีค่าไม่เกินระดับกำลังของสัญญาณเสียงพูด แต่สำหรับผู้ที่มีความสามารถในการได้ยินต่ำกว่าปกติแล้ว สัญญาณรบกวนจะมีผลกระทบต่อความเข้าใจในเสียงพูดเป็นอย่างมาก จุดหมายของระบบ ANC ก็เพื่อดึงเอาสัญญาณเสียงพูดออกมาจากสัญญาณเสียงพูดที่ถูกรบกวน [31] ซึ่งในหัวข้อนี้จะแบ่งระบบ ANC ออกเป็นสองประเภทได้แก่ ระบบ ANC แบบหนึ่งไมโครโฟน และ ระบบ ANC แบบสองไมโครโฟน

### 3.3.1.1 ระบบ ANC แบบหนึ่งไมโครโฟน

หลักการทำงานของระบบ ANC แบบหนึ่งไมโครโฟนคือการอาศัยประโยชน์จากความแตกต่างของสัญญาณเสียงพูดและสัญญาณรบกวนแบบเพิ่มเติม (Additive Noise) สัญญาณเสียงพูดเป็นสัญญาณที่สามารถทำนายได้ในช่วงสั้นๆ ช่วงหนึ่ง แต่สัญญาณรบกวนจะไม่สามารถทำนายได้ (อันที่จริงแล้วสัญญาณรบกวนจะสามารถทำนายได้ยิ่งน้อยหากยังมีช่วงความถี่กว้าง และโดยทั่วไปแล้วสัญญาณรบกวนจะมีช่วงความถี่ซึ่งกว้างมาก ทำให้ไม่สามารถทำนายได้) วงจรกรองแบบปรับตัวถูกใช้เป็นเครื่องมือในการทำนายแบบปรับตัว เพื่อใช้ในการแยกสัญญาณเสียงพูดออกจากสัญญาณรบกวน ตัวทำนายดังกล่าวถูกสร้างเพื่อปรับตัวอย่างรวดเร็วให้ทันกับการเป็นสัญญาณจดนิ่ง (Stationary Signal) ที่เป็นช่วงๆ ของสัญญาณเสียงพูด [31]



รูปที่ 20 วงจรกรองแบบปรับตัวที่ใช้ในระบบ ANC แบบหนึ่งไมโครโฟน

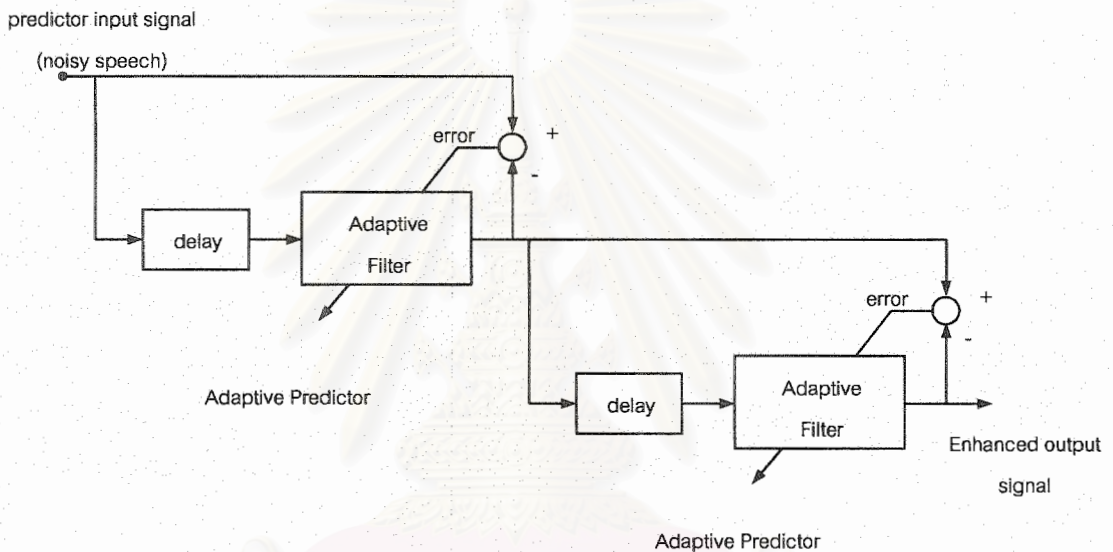


รูปที่ 21 แผนภาพของระบบ ANC แบบหนึ่งไมโครโฟน

ระบบ ANC แบบหนึ่งไมโครโฟนดังรูปที่ 21 สามารถทำนายสัญญาณรายคาบได้อย่างสมบูรณ์ เนื่องจากสัญญาณรายคาบนั้นมีคุณสมบัติเป็นสัญญาณจุดนิ่งตลอดทุกเวลา ในทางตรงกันข้ามสัญญาณเสียงพูด ไม่สามารถถูกทำนายได้อย่างสมบูรณ์ เนื่องจากมีคุณสมบัติการเป็นสัญญาณจุดนิ่งเพียงบางช่วงเวลาเท่านั้น ดังนั้นสัญญาณเสียงพูดจะสามารถถูกทำนายได้บนช่วงเวลาสั้นๆ ช่วงหนึ่ง ไม่สมบูรณ์แต่ก็เป็นการประมาณที่ดี กล่าวคือสามารถทำนายได้อย่างถูกต้องไปในอนาคตที่ไม่ไกล ซึ่งผิดกับสัญญาณรายคาบที่สามารถทำนายไปยังอนาคตที่ไกลเท่าใดก็ได้ ในทางตรงกันข้ามกับสัญญาณรายคาบและสัญญาณเสียงพูด

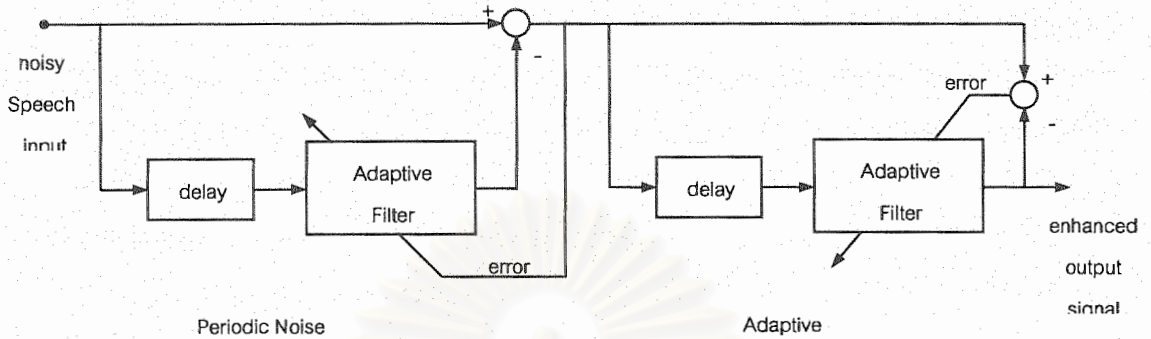
สัญญาณรบกวนที่มีช่วงความถี่กว้างจะไม่สามารถถูกทำนายได้ สามารถกล่าวโดยประมาณได้ว่าจำนวนครั้งที่สามารถทำนายได้ในอนาคตจะเท่ากับส่วนกลับของช่วงกว้างของแถบความถี่ของสัญญาณรบกวน ดังนั้นถ้าจะใช้ตัวทำนายในการแยกแยะสัญญาณเสียงพูดกับสัญญาณรบกวนจะต้องทำการเลือกค่าจำนวนในอนาคตที่ต้องการจะทำนายค่าหนึ่งซึ่งทำให้ตัวทำนายสามารถทำนายสัญญาณเสียงพูดได้แต่ไม่สามารถทำนายสัญญาณรบกวนได้ กล่าวคือเป็นค่าในอนาคตที่ไกลเกินไปสำหรับสัญญาณรบกวนในการทำนายแต่พอที่จะสามารถทำนายสัญญาณเสียงพูดได้ในช่วงสั้นๆ ค่าของจำนวนในอนาคตสามารถเลือกได้จากค่าน่วงเวลา (delay) ก่อนเข้าวงจรกรองแบบปรับตัวนั่นเอง

นอกจากนี้ยังสามารถทำการเพิ่มประสิทธิภาพของระบบจัดสัญญาณรบกวนแบบปรับตัวโดยใช้ไมโครโฟนหนึ่งตัวในรูปที่ 21 ได้โดยการต่อเพิ่มระบบดังกล่าวด้วยระบบเดิมดังรูปที่ 22



รูปที่ 22 แผนภาพของระบบจัดสัญญาณรบกวนแบบปรับตัวในรูปที่ 21 ต่อเพิ่มเข้าด้วยกัน

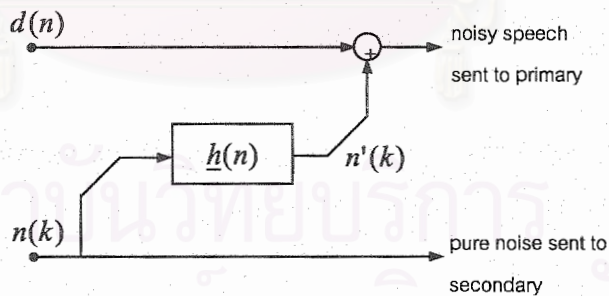
ข้อเสียของระบบที่กล่าวมานี้คือ สัญญาณรบกวนที่มีลักษณะเป็นรายคาบ เช่น เสียงเครื่องปรับอากาศ จะสามารถถูกทำนายและผ่านมายังสัญญาณขาออกได้ ระบบดังรูปที่ 23 เป็นการออกแบบเพื่อป้องกันสัญญาณรบกวนแบบรายคาบ โดยประกอบด้วยตัวจัดสัญญาณรบกวนแบบรายคาบ (Periodic Noise Canceller) ต่อเพิ่มกับตัวทำนายแบบปรับตัวในรูปที่ 21 หลักการทำงานของตัวจัดสัญญาณรบกวนแบบรายคาบเป็นหลักเดียวกับตัวทำนายแบบปรับตัว แต่ใช้ในทางที่ต่างออกไป เพื่อเป็นการป้องกันการรบกวนสัญญาณเสียงพูดทิ้งไปด้วย ค่าหน่วงเวลาที่เลือกใช้ในตัวจัดสัญญาณรบกวนแบบรายคาบจะต้องมากพอเพื่อวงจรกรองแบบปรับตัวจะไม่สามารถทำนายสัญญาณเสียงพูดได้ (มากกว่าช่วงสั้นๆ ที่สัญญาณเสียงพูดยังคงเป็น สัญญาณแบบจุดนิ่ง)



รูปที่ 23 แผนภาพของระบบขจัดสัญญาณรบกวนแบบรายคาบต่อเพิ่มด้วยระบบในรูปที่ 21

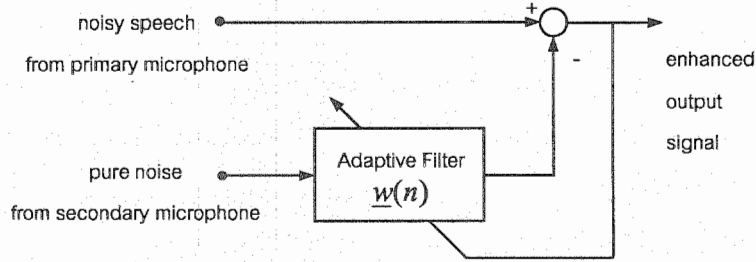
### 3.3.1.2 ระบบ ANC แบบสองไมโครโฟน

ระบบ ANC แบบสองไมโครโฟนสองตัวถูกริเริ่มขึ้นในช่วงปี 1960 โดยถูกมองเป็นขั้นตอนวิธีการปรับปรุงโดยใช้วงจรกรองแบบปรับตัวที่มีพื้นฐานคือใช้มีสัญญาณขาเข้าหลัก (Primary Input Source) และสัญญาณอ้างอิงรอง (Secondary Reference Source) สัญญาณขาเข้าหลักเป็นสัญญาณที่ประกอบด้วยสัญญาณที่ต้องการ (desired signal) รวมกับสัญญาณรบกวน  $n'(k)$  ส่วนสัญญาณอ้างอิงรองถูกสมมติว่ามีเพียงสัญญาณรบกวน  $n(k)$  ซึ่งมีสหสัมพันธ์กับสัญญาณรบกวน  $n'(k)$  แต่ไม่มีสหสัมพันธ์กับสัญญาณที่ต้องการ [32]



รูปที่ 24 แสดงให้เห็นถึงความสัมพันธ์ของสัญญาณขาเข้าหลัก (Primary Input Source) และสัญญาณอ้างอิงรอง (Secondary Reference Source)

แผนภาพแสดงระบบ ANC แบบสองไมโครโฟนขั้นพื้นฐานถูกแสดงไว้ในรูปที่ 25 โดยใช้วงจรกรองแบบปรับตัวซึ่งใช้ขั้นตอนวิธีแบบค่าเฉลี่ยกำลังสองน้อยที่สุด (Normalized Least Mean Square Algorithm, NLMS) ในการประมาณสัญญาณรบกวนแบบเพิ่มเติม แล้วจึงนำไปทำการหักล้างออกจากสัญญาณขาเข้าหลัก



รูปที่ 25 แผนภาพของระบบ ANC แบบสองไมโครโฟน

ถึงแม้ว่าระบบขจัดสัญญาณรบกวนแบบปรับตัวโดยใช้ไมโครโฟนสองตัวจะให้อัตราสัญญาณต่อสัญญาณรบกวนที่ต่ำ คุณภาพของระบบอาจถูกทำให้ลดลงได้โดยหากสัญญาณขาเข้าหลักมีองค์ประกอบของสัญญาณรบกวนที่ไม่มีสหสัมพันธ์กับสัญญาณอ้างอิงรอง ซึ่งสัญญาณรบกวนที่ไม่มีสหสัมพันธ์กับสัญญาณอ้างอิงรองนี้ จะไม่สามารถถูกขจัดทิ้งไปได้จึงส่งผ่านไปยังสัญญาณทางขาออก

ปัญหาอีกประการที่เกิดขึ้นกับระบบขจัดสัญญาณรบกวนแบบปรับตัวโดยใช้ไมโครโฟนสองตัวคือ หากสัญญาณอ้างอิงรองมีสหสัมพันธ์กับสัญญาณเสียงพูดบางส่วนในสัญญาณขาเข้าหลักซึ่งถูกเรียกว่า cross talk ส่งผลให้วงจรกรองแบบปรับตัวขจัดเสียงพูดบางส่วนนั้นไป ทำให้เกิดการบิดเพี้ยนของสัญญาณขึ้น [33]

### 3.3.2 ระบบ SS

ระบบ SS เป็นระบบขจัดสัญญาณรบกวนแบบหนึ่งไมโครโฟนประเภทหนึ่งซึ่งมีหลักการ คือ เรียนรู้ลักษณะทางความถี่ของสัญญาณรบกวนและใช้ลักษณะทางความถี่ของสัญญาณรบกวนนั้นมาทำการสร้างสัญญาณรบกวนโดยประมาณขึ้นมาเพื่อที่จะนำไปหักล้างกับสัญญาณรบกวนที่ปะปนมากับสัญญาณที่ต้องการ การเรียนรู้ลักษณะทางความถี่ของสัญญาณรบกวน ทำได้โดยทำการตรวจจับช่วงเงียบจากสัญญาณที่รับมา โดยดูจากค่าพลังงานโดยประมาณในช่วงนั้น หากมีค่าสูงกว่าค่าที่ตั้งไว้จะถือว่าเป็นช่วงที่ไม่เงียบ เมื่อตรวจจับช่วงที่ถือว่าเป็นช่วงเงียบได้แล้ว ก็จะทำการประมาณลักษณะทางความถี่ของสัญญาณรบกวนโดยถือว่าสัญญาณที่รับมาในช่วงเงียบนี้คือสัญญาณรบกวนเท่านั้น

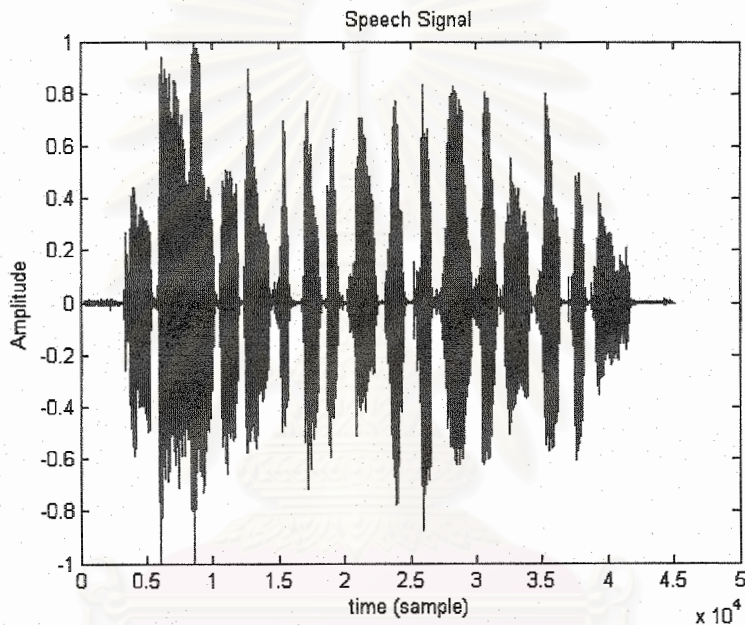
### 3.4 การใช้งานร่วมกันของระบบ AEC กับเทคนิค ANC

เนื่องจากเหตุผลที่ว่าระบบขจัดสัญญาณเสียงสะท้อนจะทำงานได้เร็วลงเมื่ออยู่ในสภาพแวดล้อมที่มีสัญญาณรบกวน ดังนั้นจึงมีความพยายามที่จะนำระบบขจัดสัญญาณรบกวนมาช่วยเพิ่มประสิทธิภาพของระบบขจัดสัญญาณเสียงสะท้อน

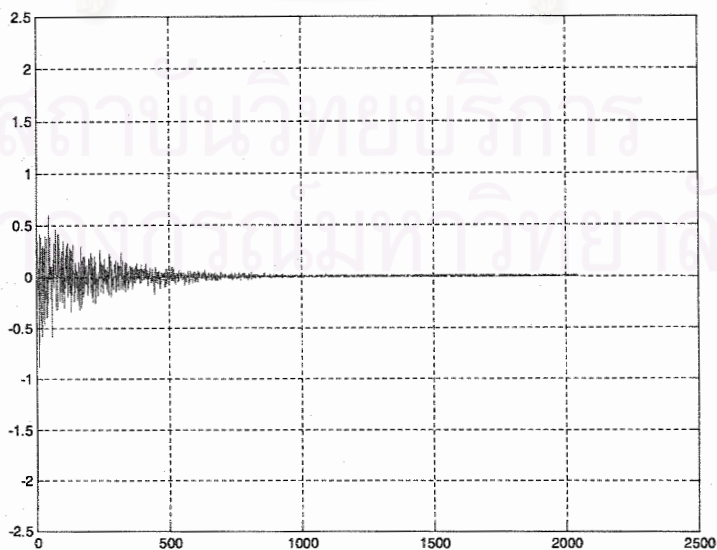


### 3.5 ผลการจำลองระบบบนคอมพิวเตอร์ของการใช้งานร่วมกันของระบบ AEC กับเทคนิค ANC

ในการจำลองระบบบนคอมพิวเตอร์ด้วยโปรแกรม MATLAB สัญญาณเสียงขาเข้าของระบบ AEC เป็นสัญญาณเสียงพูด (Speech Signal) มีความถี่ซีกตัวอย่าง (Sampling Frequency,  $f_s$ ) เป็น 16 kHz ดังแสดงในรูปที่ 26 ในส่วนของแบบจำลองที่ใช้สำหรับวิธีสะท้อนทางเสียงในห้องใกล้ จะอ้างอิงถึงวิธีสะท้อนทางเสียงที่วิเคราะห์ได้จากการบันทึกสัญญาณเสียงสะท้อนจริง ภายในห้องสัมมนาทางไกลแบบสเตอริโอ แต่จะเลือกมาใช้เพียงวิธีสะท้อนทางเสียงในห้องใกล้วิธีเดียว จากสองวิธีสะท้อนที่หาได้ จึงทำการจำลองดังแสดงได้ในรูปที่ 27



รูปที่ 26 สัญญาณเสียงพูดเป็นสัญญาณขาเข้าของระบบ AEC (SNR = 4.2 dB)



รูปที่ 27 วิธีสะท้อนทางเสียงในห้องรับ ( $L = 2048$ )

การศึกษาเปรียบเทียบประสิทธิภาพการทำงานของวงจรกรองแบบปรับตัวของระบบ AEC จะแสดงผ่านทางค่า Weight Error Vector Norm (WEVN) ดังสมการ

$$\text{WEVN}(n) = 10 \times \log_{10} \left( \frac{\|\mathbf{h}(n) - \hat{\mathbf{h}}(n)\|^2}{\|\mathbf{h}(n)\|^2} \right) \quad (8)$$

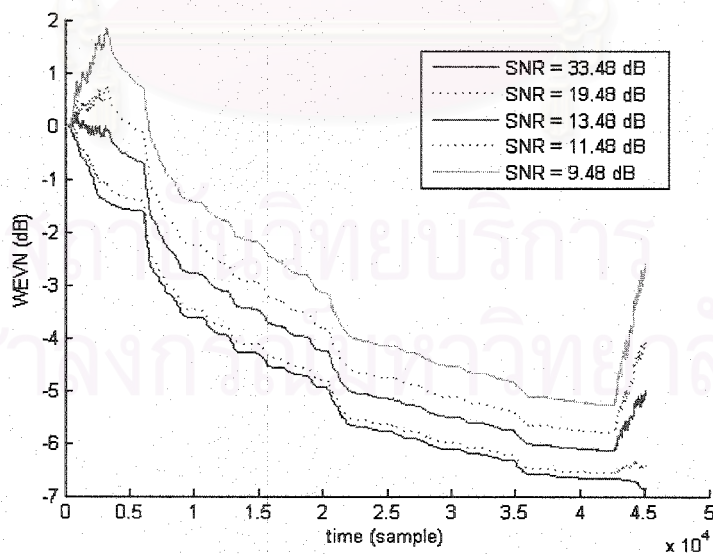
สำหรับการวัดสมรรถนะของวงจรกรองแบบปรับตัวในระบบ AEC ที่ไม่ทราบค่าวิถีสะท้อนทางเสียง สามารถเลือกใช้ค่า Echo Return Loss Enhancement (ERLE) ดังสมการ

$$\text{ERLE}(n) = 10 \times \log_{10} \left( \frac{\sum_{i=0}^{N-1} d(n-i)^2}{\sum_{i=0}^{N-1} e(n-i)^2} \right) \quad (9)$$

เมื่อ N คือหน้าต่างในการหาค่าเฉลี่ยของข้อมูล

ในหัวข้อย่อนี้ จะได้แบ่งการจำลองระบบบนคอมพิวเตอร์ ออกเป็น 3 กรณีของ (1) การใช้งานระบบ AEC (2) การใช้งานระบบ ANC และ (3) การใช้งานร่วมกันของระบบ AEC กับระบบ ANC ตามลำดับ

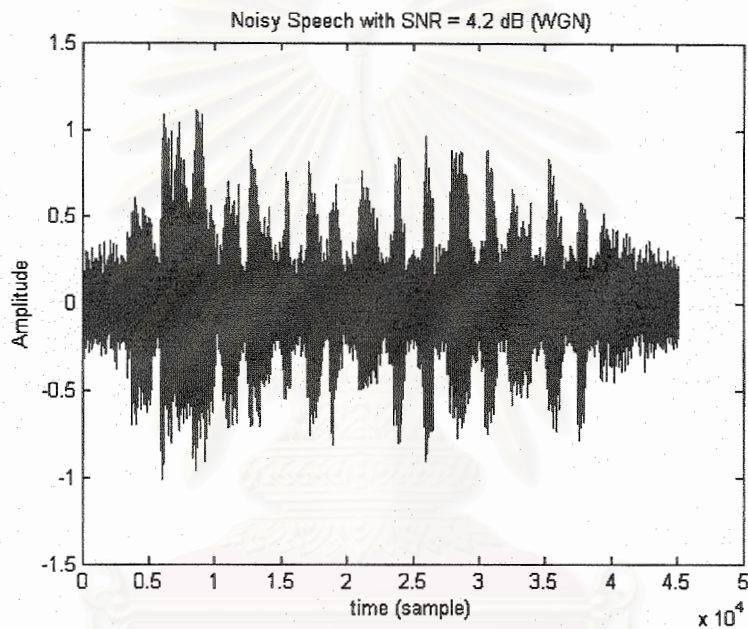
### 3.5.1 การใช้งานระบบ AEC



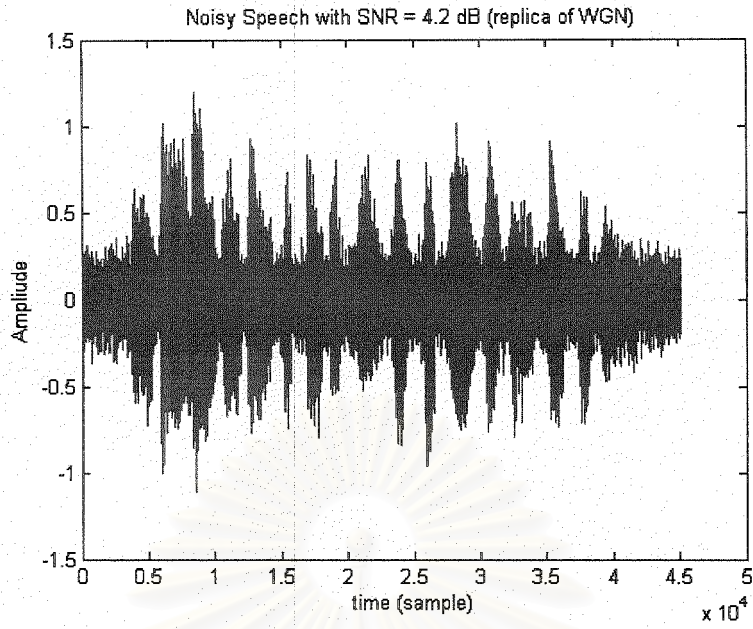
รูปที่ 28 ค่า WEVN ของระบบ AEC ในสภาพแวดล้อมที่มีสัญญาณรบกวนในระดับต่างๆ

ระบบ AEC ดังที่ได้กล่าวถึงในหัวข้อ 3.4 จะทำงานได้เลวลงในสภาพแวดล้อมที่มีสัญญาณรบกวน ทั้งนี้เนื่องจากสัญญาณรบกวนทำให้ความเป็นเชิงเส้นของระบบหายไป ซึ่งมีสัญญาณรบกวนเพิ่มมากขึ้นเท่าใด ระบบ AEC จะยิ่งทำงานได้เลวลงดังรูปที่ 28 โดยในที่นี้ใช้การวัดผลการทำงาน โดยแสดงเป็นค่า WEVN ซึ่งบ่งชี้ถึงความใกล้เคียงของระบบจำลองของระบบ AEC กับวิธีสะท้อนทางเสียงในห้องสัมมนา

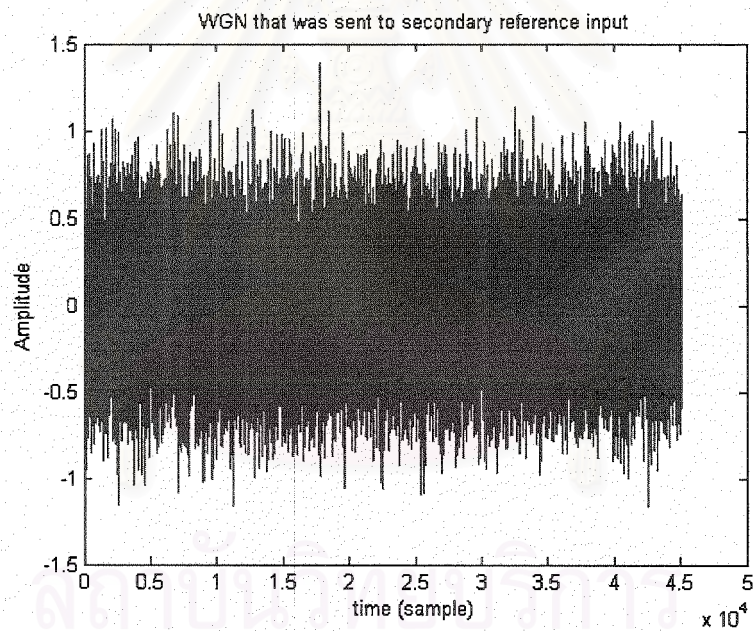
### 3.5.2 การใช้งานระบบ ANC



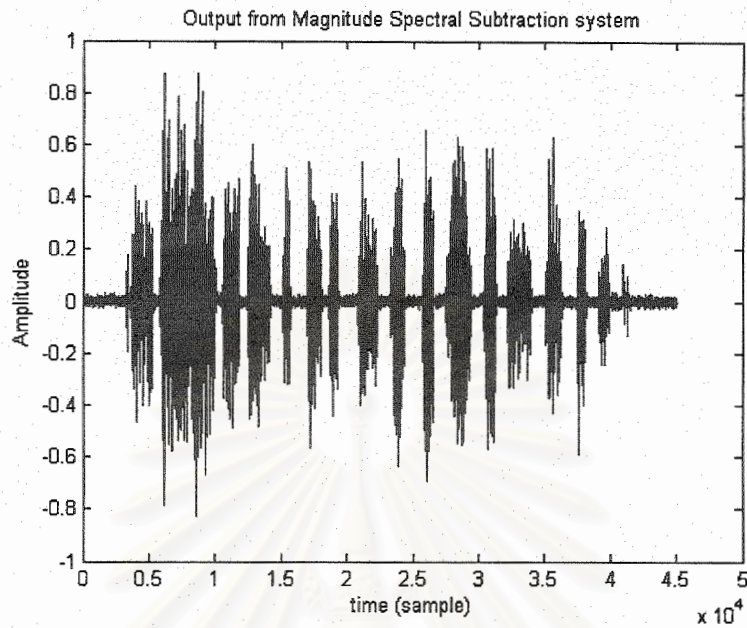
รูปที่ 29 สัญญาณเสียงพูดที่ถูกรบกวนด้วยสัญญาณรบกวนแบบ White Gaussian Noise ซึ่งใช้เป็นสัญญาณขาเข้าของระบบ ANC แบบสองไมโครโฟน และ ระบบ SS โดยมีค่า SNR = 4.2 dB



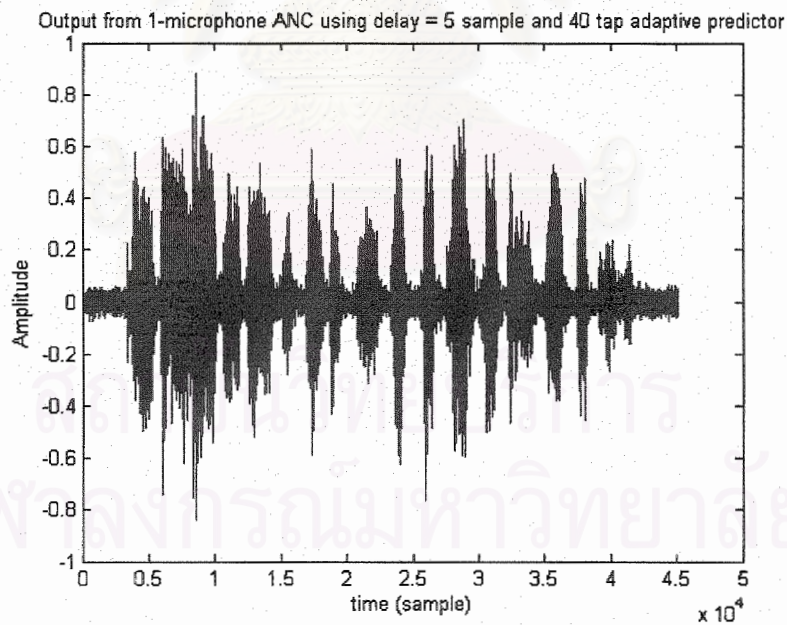
รูปที่ 30 สัญญาณขาเข้าหลักของระบบ ANC แบบสองไมโครโฟน โดยมีค่า SNR = 4.2 dB



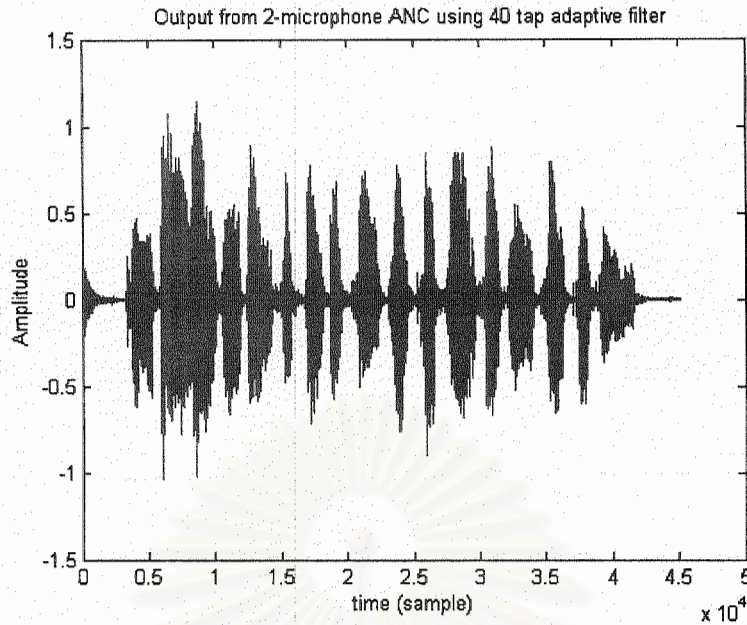
รูปที่ 31 สัญญาณอ้างอิงรองของระบบ ANC แบบสองไมโครโฟน



รูปที่ 32 สัญญาณขาออกของระบบ SS เมื่อสัญญาณเข้าเป็นดังรูปที่ 10 SNR ของสัญญาณนี้ คือ 23.346 dB



รูปที่ 33 สัญญาณขาออกของระบบ ANC แบบหนึ่งไมโครโฟน ที่ใช้ตัวหน่วงเวลาจำนวน 5 sample และใช้วงจรแบบปรับตัวซึ่งมีค่าสัมประสิทธิ์จำนวน 40 ตัว เมื่อสัญญาณเข้าเป็นดังรูปที่ 10 SNR = 16.424 dB

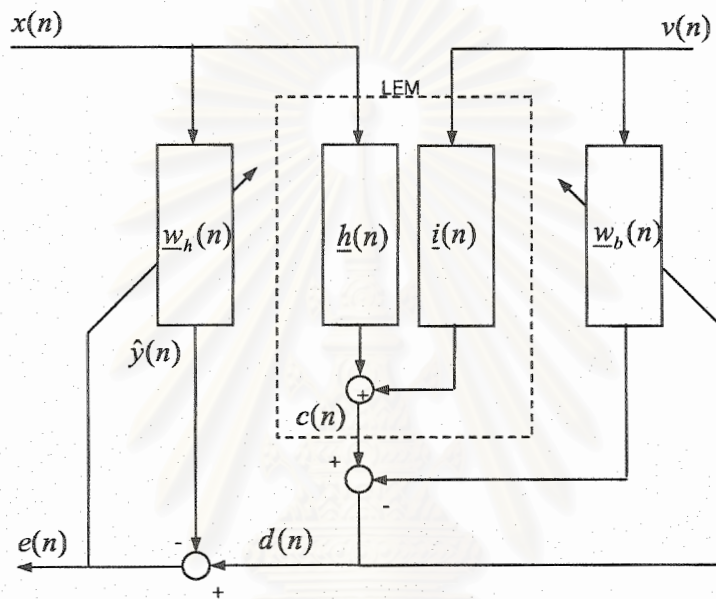


รูปที่ 34 แสดงสัญญาณขาออกของระบบ ANC แบบสองไมโครโฟน โดยจำนวนสัมประสิทธิ์ของวงจรกรองแบบปรับตัวคือ 40 ตัว เมื่อสัญญาณเข้าที่เข้าหลักและอ้างอิงเป็นดังรูปที่ 35 และรูปที่ 36 ตามลำดับ SNR = 26.968 dB

จากผลการทดลองพบว่าระบบ ANC แบบสองไมโครโฟน ให้ผลการขจัดสัญญาณรบกวนที่ดีที่สุด แต่เช่นกันย่อมต้องแลกมาซึ่งความซับซ้อนของระบบที่มากที่สุดตามไปด้วย ในขณะที่ระบบ ANC แบบหนึ่งไมโครโฟน มีความซับซ้อนน้อยกว่าและให้ผลที่ดีเช่นกัน และเมื่อเพิ่มความซับซ้อนของระบบ ANC แบบหนึ่งไมโครโฟน โดยเปลี่ยนไปใช้ตามแบบในรูปที่ 21 หรือ 22 จะให้ผลซึ่งดียิ่งขึ้นไปอีก (แต่ความซับซ้อนของระบบก็จะเพิ่มขึ้นเช่นกัน) ส่วนระบบ SS แม้จะมีความซับซ้อนของระบบน้อยที่สุดในสามระบบ ให้ผลการลดลงของสัญญาณรบกวนซึ่งออกมาในรูปของอัตราสัญญาณต่อสัญญาณรบกวนที่ดีเทียบเท่ากับระบบ ANC แบบสองไมโครโฟน แต่หากทำการวัดผลแบบอิงผู้ฟัง (Subjective Test) แล้ว จะให้ผลที่ค่อนข้างเลว เนื่องจากระบบ SS เป็นเพียงการประมาณสัญญาณรบกวน ดังนั้นจะให้ผลเพียงภาพรวมของสัญญาณรบกวนที่ต่ำลง กล่าวคือทำให้ค่าเฉลี่ยของสัญญาณรบกวนมีค่าต่ำลง แต่สิ่งที่ตามมาคือทำให้ค่าแปรปรวนของสัญญาณรบกวนเพิ่มมากขึ้น ทำให้เมื่อทำการฟังสัญญาณจากระบบดังกล่าวจะได้ยินเสียงสัญญาณรบกวนอย่างชัดเจน และนอกจากนี้ระบบ SS จะมีส่วนที่มีพฤติกรรมไม่เชิงเส้นอยู่ในระบบด้วย ทำให้สัญญาณเสียงพูดที่ได้มีการบิดเบี้ยวและทำให้ความสามารถในการฟังรู้เรื่องของสัญญาณเสียงพูดลดลง ผลของความบิดเบี้ยวจากระบบหักล้างทางความถี่เนื่องจากความแปรปรวนที่สูงขึ้นถูกเรียกว่า สัญญาณรบกวนแบบดนตรี (Musical Noise) [33] งานที่เหมาะสมกับระบบหักล้างทางความถี่จึงควรเป็นงานที่ต้องการความซับซ้อนที่ต่ำและไม่ได้มีเป้าหมายให้มนุษย์เป็นผู้รับสาร ใน [33] ได้นำระบบ SS ในการขจัดสัญญาณรบกวนเพื่อช่วยเพิ่มประสิทธิภาพของระบบรู้จำเสียงพูดที่ใช้ในรถยนต์ เนื่องจากระบบดังกล่าวผู้รับสารเป็นคอมพิวเตอร์

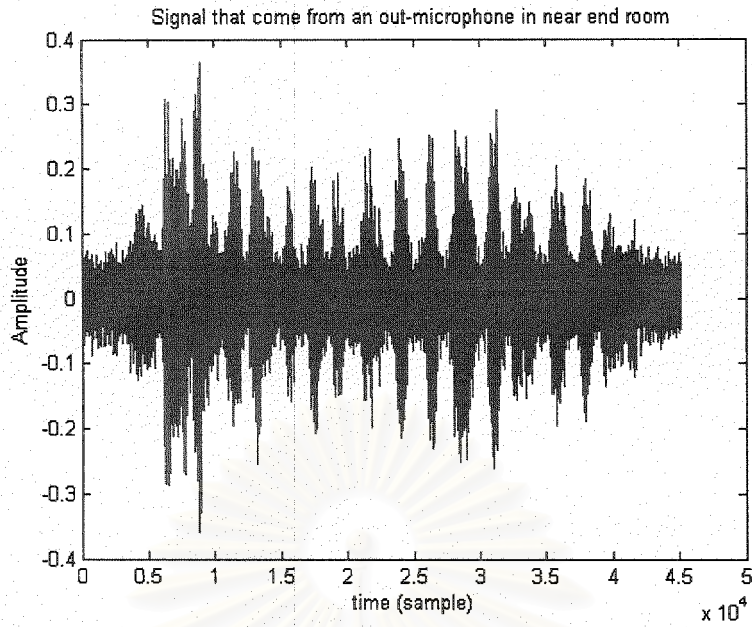
### 3.5.3 การใช้งานร่วมกันของระบบ AEC กับระบบ ANC

โครงการวิจัยนี้นำเสนอการศึกษาการรวมกันของระบบขจัดเสียงสะท้อนในระบบสัมมนาทางไกลกับระบบ ANC แบบสองไมโครโฟน เนื่องจากเป็นระบบที่ผู้รับสารเป็นมนุษย์ และต้องการผลการลดสัญญาณรบกวนที่ดีที่สุดเพื่อช่วยประสิทธิภาพในการจำลองระบบของระบบขจัดสัญญาณเสียงสะท้อนให้ดียิ่งขึ้น ระบบที่นำเสนอเป็นดังรูปที่ 3.4.3.1

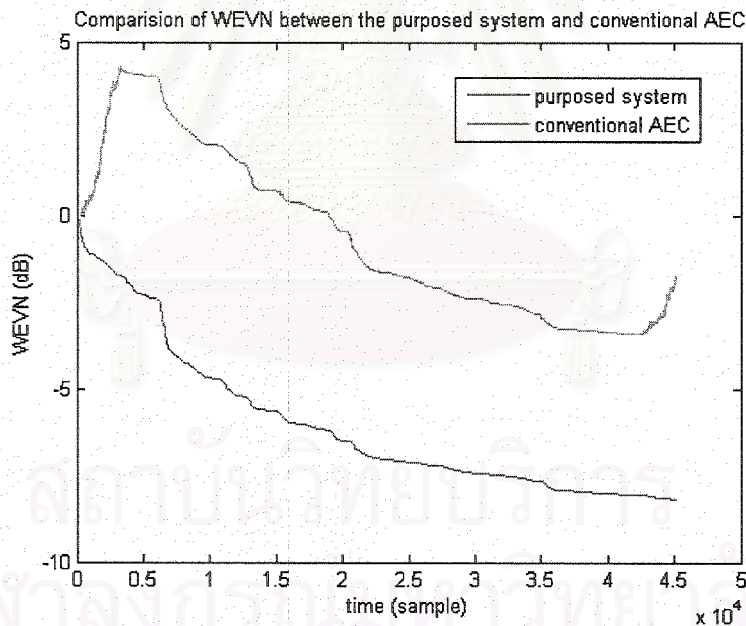


รูปที่ 35 แผนภาพระบบ AEC โดยเพิ่มเติมด้วยระบบ ANC แบบสองไมโครโฟน

แนวคิดของระบบที่นำเสนอเริ่มต้นมาจากความเข้าใจที่ว่า วงจรกรองแบบปรับตัวสามารถทำงานทางด้านการแสดงเอกลักษณ์ของระบบ (System Identification) ได้ดีก็ต่อเมื่อระบบเป็นเชิงเส้น ทั้งนี้เนื่องมาจากระเบียบวิธีต่างๆ ที่ทำการพัฒนาขึ้นมาเป็นระเบียบวิธีซึ่งมีสมมติฐานอยู่บนหัวข้อที่ว่าระบบที่ต้องการแสดงเอกลักษณ์นั้นเป็นระบบเชิงเส้นนั่นเอง เมื่อพิจารณาถึงห้องโถงไกลของระบบขจัดสัญญาณเสียงสะท้อนที่มีสัญญาณรบกวนร่วมอยู่ด้วย จะเห็นว่าระบบดังกล่าวนี้จะไม่มีคุณสมบัติความเป็นเชิงเส้น อาจพิจารณาได้ง่ายๆ ได้จากกรณีซึ่งแม้ไม่มีสัญญาณเข้าใดๆ แต่สัญญาณขาออกจะยังคงมีอยู่ (สัญญาณรบกวนเพียงอย่างเดียว) ความไม่เป็นเชิงเส้นนี้เองที่เป็นเหตุให้การปรับตัวเพื่อแสดงเอกลักษณ์ของระบบของวงจรกรองแบบปรับตัวเลวลง และเลวลงไปอีกหากระดับของสัญญาณรบกวนเพิ่มมากยิ่งขึ้น ดังนั้นหากต้องการแก้ไขปัญหาก็เกิดขึ้นต้องทำให้ระบบดังกล่าวเป็นระบบเชิงเส้นโดยนำสัญญาณรบกวนมาเป็นสัญญาณเข้าอีกตัวหนึ่งของวงจรกรองแบบปรับตัวด้วย



รูปที่ 36 สัญญาณจากไมโครโฟนขาออกของห้องใกล้ ( $c(n)$  ในรูปที่ 16) SNR = 5.75 dB เพื่อนำไปทำการทดลองกับระบบที่นำเสนอและระบบขจัดสัญญาณเสียงสะท้อนทั่วไปต่อไป



รูปที่ 37 แสดงผลการเปรียบเทียบค่า Weight Error Vector Norm (WEVN) ระหว่างระบบที่นำเสนอกับระบบขจัดสัญญาณเสียงสะท้อนทั่วไป ในสภาพแวดล้อมที่มีระดับสัญญาณรบกวน SNR = 5.75 dB



### 3.6 ผลการทดสอบระบบ AEC ที่พัฒนาขึ้นด้วยการสร้างจริงบนบอร์ด FPGA

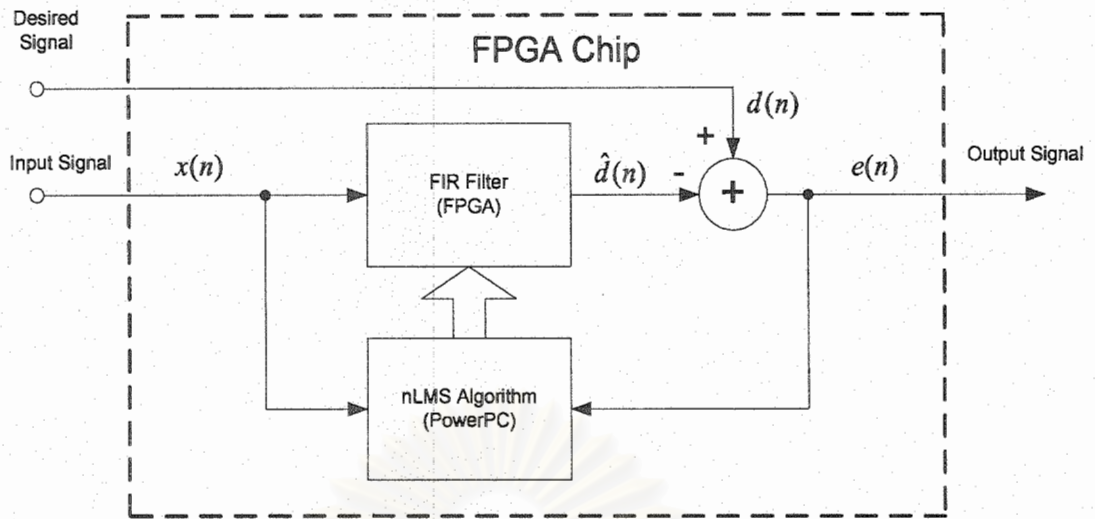
ในเฟสที่สามของโครงการ บอร์ด FPGA รุ่น DS-KIT-2VP7FF672-EURO ถูกเลือกใช้เพื่อนำมาสร้างวงจรกรองแบบปรับตัวของระบบ AEC ในรูปแบบของฮาร์ดแวร์ (Hardware) ที่มีโครงสร้าง

แบบผลตอบสนองอิมพัลส์จำกัด (Finite Impulse Response, FIR) จำนวน 128 แทป (Tap, จำนวนสัมประสิทธิ์ของวงจรกรองแบบปรับตัว) โดยใช้ขั้นตอนวิธี NLMS ควบคุมการทำงานของวงจรกรองแบบปรับตัว

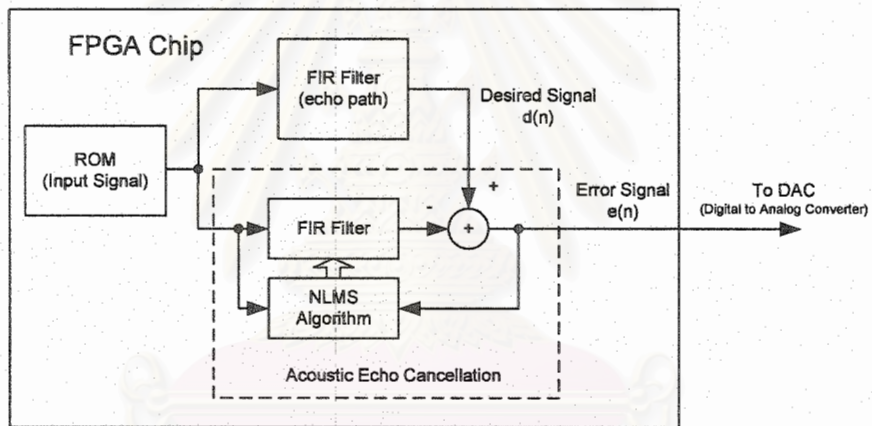
โครงสร้างภายในบอร์ด FPGA นี้ประกอบไปด้วย 2 ส่วนหลัก คือ

- ชิป FPGA (FPGA Chip) ซึ่งเป็นชิปของบริษัท Xilinx ตระกูล Virtex-II Pro รุ่น XC2VP7 ในการออกแบบจะใช้ชิป FPGA สร้างเป็นฮาร์ดแวร์ของวงจรกรองแบบปรับตัวที่มีโครงสร้างแบบ FIR โดยโครงสร้างวงจรกรองแบบปรับตัวภายในชิป FPGA ดังแสดงในรูปที่ 10
- โปรแกรม PowerPC<sup>®</sup> ซึ่งเป็น RISC Processor ของ IBM<sup>®</sup>  
โปรแกรม PowerPC<sup>®</sup> จะใช้ทำการคำนวณค่าสัมประสิทธิ์ของวงจรกรองแบบปรับตัวและส่งค่าที่ได้ไปทำการปรับค่าสัมประสิทธิ์ของวงจรกรองแบบปรับตัว ทั้งนี้โปรแกรม PowerPC<sup>®</sup> จะช่วยลดทรัพยากร (จำนวนเกต - Gate) ที่ใช้ในการสร้างวงจรกรองแบบปรับตัวลงได้ และนอกจากนี้ยังทำให้เกิดความสะดวกในกรณีที่ต้องการเปลี่ยนแปลงโปรแกรมในการปรับค่าสัมประสิทธิ์ ของวงจรกรองแบบปรับตัว

ในระยะแรก การทดสอบวงจรของระบบ AEC ได้ทำการสร้าง ROM มาทำการเก็บข้อมูลสัญญาณเข้า ซึ่งเป็นข้อมูลเสียงพูด เพื่อป้อนให้แก่วงจร ส่วนที่เอาต์พุตของวงจร ได้ทำการต่อวงจร DAC (Digital to Analog Converter) เพื่อวัดสัญญาณความผิดพลาด (Error Signal) ของระบบ AEC โดยบล็อกไดอะแกรมแสดงภาพรวมทั้งหมดของระบบ AEC ถูกแสดงดังในรูปที่ 39 สำหรับโปรแกรมคำสั่งการทำงานของระบบ AEC ถูกเขียน โดยใช้ภาษา VHDL ลงบนชิป FPGA

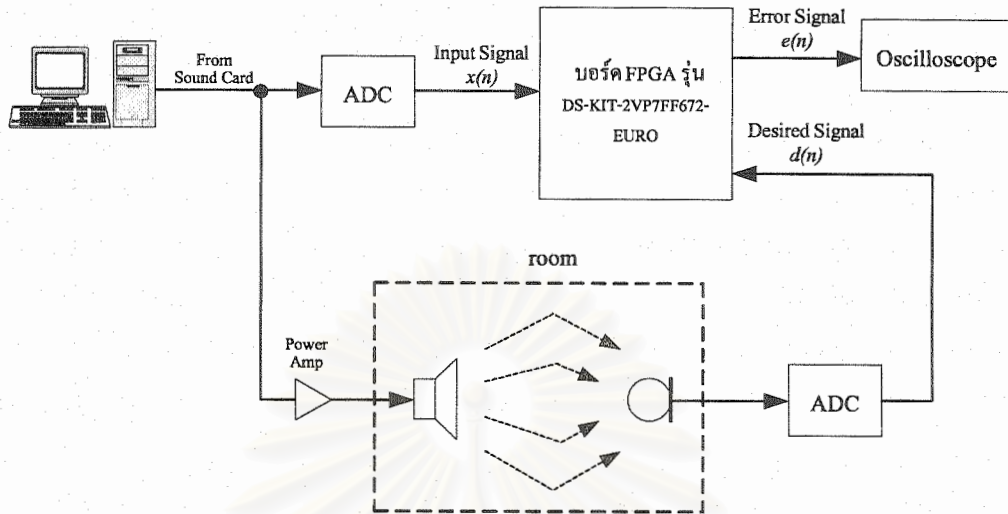


รูปที่ 38 โครงสร้างวงจรกรองแบบปรับตัวภายในชิป FPGA



รูปที่ 39 บล็อกไดอะแกรมที่ใช้ในระยะแรกสำหรับการทดสอบวงจรของระบบ AEC

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 40 บล็อกไดอะแกรมแสดงภาพรวมทั้งหมดของระบบ AEC

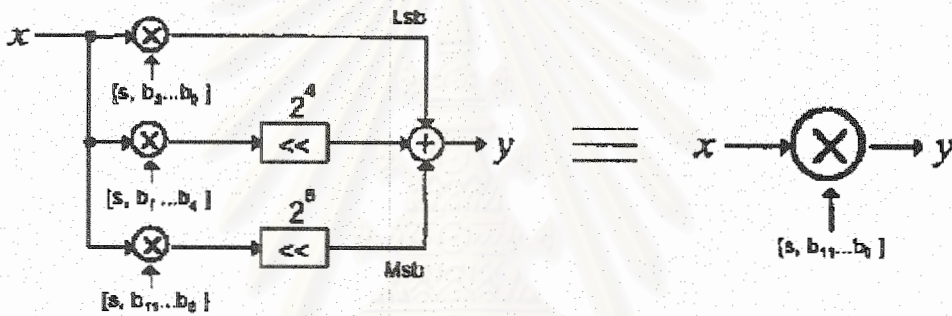
ในระยะต่อมา ผลการทดสอบในเวลาจริง ทำการป้อนข้อมูลสัญญาณเข้าจากภายนอกที่ได้มาจาก Sound Card ของคอมพิวเตอร์ ซึ่งเป็นสัญญาณแอนะล็อก (Analog) นำมาผ่านวงจร ADC (Analog to Digital Converter) เพื่อแปลงให้เป็นสัญญาณดิจิทัล (Digital) ส่งให้บอร์ด FPGA ใช้ทำการทดสอบ (แทนการป้อนข้อมูลโดย ROM ในระยะแรก) สำหรับวิธีสะท้อนทางเสียงระหว่างลำโพงกับไมโครโฟนในห้องโถงนั้น ได้เลือกใช้วงจรกรองขนาด 128 แท็ป ที่สร้างขึ้นมาในชิป FPGA ดังนั้น บล็อกไดอะแกรมที่ใช้ทำการทดสอบวงจรของระบบ AEC จึงกลายเป็นดังแสดงในรูปที่ 40

สำหรับการทดสอบการใช้งาน ADC ร่วมกับ DAC นั้น ได้ทดลองนำ DAC มาต่อกับ ADC โดยตรง และทำการป้อนสัญญาณแอนะล็อกเข้าที่อินพุตของ ADC จากนั้นวัดสัญญาณแอนะล็อกที่เอาต์พุตของ DAC ผลปรากฏว่า ข้อมูลสัญญาณออกของวงจรที่ความถี่สูงๆ มีลักษณะสเปกตรัมที่ใกล้เคียงกับสเปกตรัมของสัญญาณที่อินพุต แต่ที่ย่านความถี่ต่ำๆ ไม่มีสเปกตรัมของสัญญาณขาออกของวงจรเลย ทั้งนี้เนื่องจาก วงจร ADC มีตัวเก็บประจุ (Capacitor) ซึ่งทำหน้าที่เชื่อมต่อสัญญาณ (Coupling) และตัวต้านทาน (Resistor) ภายในวงจร โดย ADC ทำตัวเสมือนเป็นวงจรกรองผ่านสูง (High-Pass Filter) ที่มีความถี่ตัด (Cutoff Frequency) ประมาณ 3 kHz ต่อยู่ที่อินพุตของวงจร ADC จึงกรองข้อมูลในย่านความถี่ต่ำ รวมทั้งในย่านความถี่ของสัญญาณเสียงทิ้งไป จึงต้องหาวิธีแก้ไขปัญหานี้ด้วย

ในการสร้างวงจรดิจิทัล โดยเฉพาะอย่างยิ่ง การสร้างตัวกรองดิจิทัลนั้น ปัญหาที่สำคัญก็คือ การสร้างตัวคูณ (Multiplier) ที่ต้องใช้พื้นที่หรือทรัพยากร (Resource) ในการสร้างค่อนข้างมากเมื่อเทียบกับวงจรดิจิทัลอื่นๆ จึงได้ศึกษาเปรียบเทียบเทคนิคการสร้างตัวคูณแบบโครงสร้างตัวประมวลผลเลขคณิตกระจาย (Distributed Arithmetic, DA) กับ เทคนิคการสร้างตัวคูณแบบโครงสร้างตัวประมวลผลแบบแอก

คิวมูล (Multiply and Accumulation, MAC) เพื่อหาความเป็นไปได้ในการลดพื้นที่ของโครงสร้างวงจร FIR จากการศึกษพบว่า

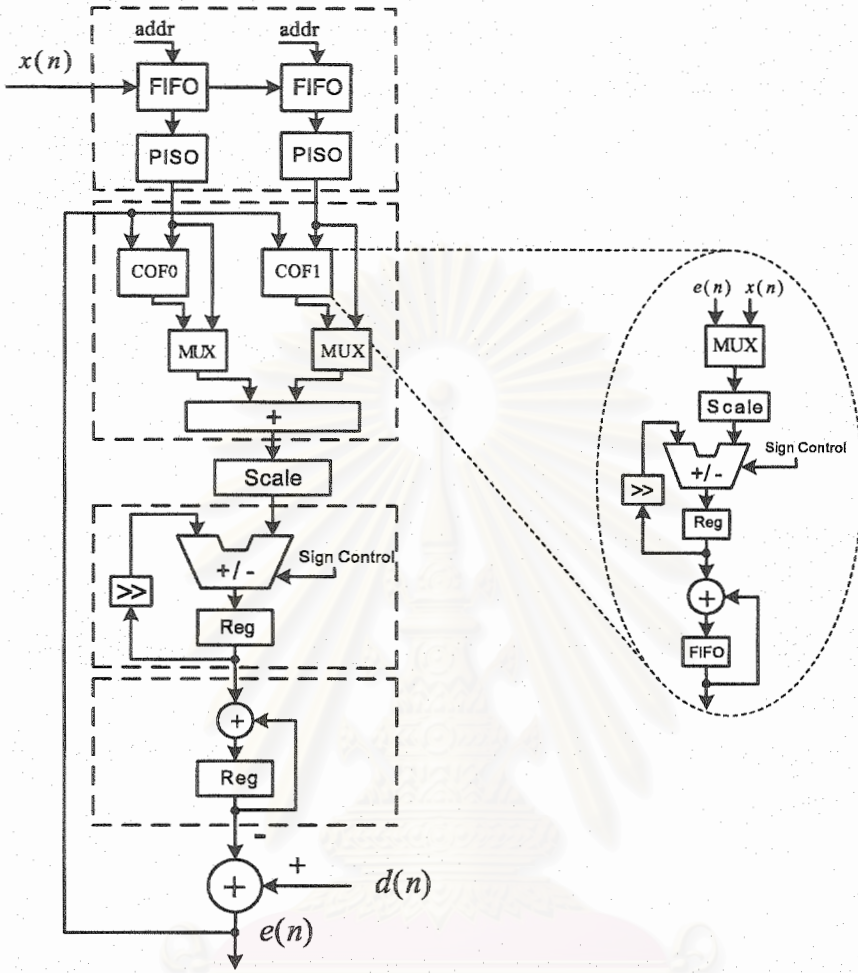
- เทคนิค DA เป็นการแยกการคูณออกเป็นแถวย่อย และทำการคูณในแต่ละแถวด้วยจำนวนบิตที่ต่ำลง ก่อนที่จะนำมารวมกันตามสัดส่วนการคูณที่ควรจะเป็นเพื่อสร้างสัญญาณออกที่ต้องการ ดังแสดงในรูปที่ 41 จึงมีข้อดี คือ จะทำให้จำนวนเกตที่ต้องใช้ในการสร้างตัวคูณลดลง เนื่องจากเป็นการกระจายข้อมูลที่จะนำมาคูณ ทำให้วงจรกินไฟน้อยลง อีกทั้งยังจะมีความเร็วเพิ่มขึ้นด้วย ซึ่งจะเอื้ออำนวยให้สามารถสร้างวงจรกรองแบบปรับตัวของระบบ AEC ที่มีจำนวนแท็ปสูงขึ้นถึง 1024 - 2048 แท็ป ตามที่ต้องการ ได้ด้วยบอร์ด Xilinx ที่มีอยู่
- เทคนิค MAC ช่วยในการลดจำนวนตัวคูณที่ต้องใช้ในวงจรลงได้ เนื่องจากการประมวลผลในลักษณะการวนรอบคำนวณ ซึ่งจะมีลักษณะเป็นการนำทรัพยากรกลับมาใช้ (Reuse) ในการประมวลผลครั้งใหม่



รูปที่ 41 โครงสร้าง DA

ดังนั้น ในการสร้างวงจรกรองแบบปรับตัวของระบบ AEC เมื่ออาศัยเทคนิคการประมวลผลแบบ MAC ร่วมกับเทคนิคแบบ DA จะสามารถช่วยลดจำนวนทรัพยากรในการสร้างวงจรลงได้มาก โดยโครงสร้างแบบ DA ร่วมกับ MAC แสดงดังในรูปที่ 42

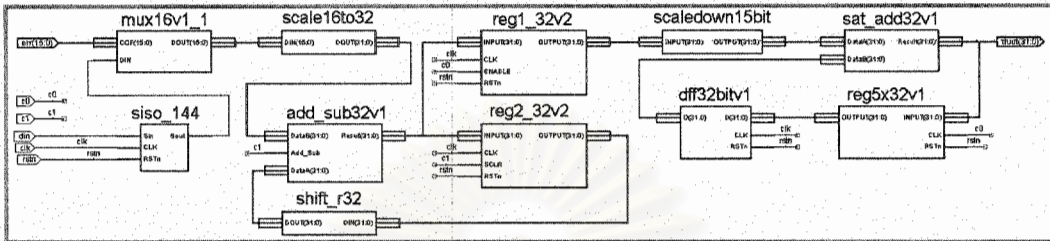
สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 42 ตัวกำจัดเสียงสะท้อนโดยใช้โครงสร้างแบบ DA ร่วมกับ MAC

การทำงานเริ่มจาก วงจร FIFO (First-in First-out) ทำหน้าที่เลื่อนข้อมูลหรือเปรียบเสมือนเคาะในตู้กรงทั่วไปๆ หลังจากข้อมูลในแต่ละชุดถูกเลื่อนแล้ววงจร PISO (Parallel-in Serial-out) จะทำหน้าที่ดึงข้อมูลในแต่ละตำแหน่งของ FIFO ออกมาและเลื่อนออกมาทีละบิต โดยเลื่อนบิตที่มีนัยสำคัญค่าสุดท้ายออกมา ก่อน จากนั้นส่งข้อมูลที่ละบิตไปยังวงจร COF เพื่อคูณกับข้อมูลสัญญาณค่าผิดพลาด  $e(n)$  ได้เป็นค่าสัมประสิทธิ์แต่ละแท็บของวงจรกรองแบบปรับตัว โดยค่าสัมประสิทธิ์เหล่านี้จะถูกเก็บไว้ใน FIFO ที่อยู่ในวงจร COF เพื่อนำมาบวกกับค่าสัมประสิทธิ์ที่แท็บเดิมในเวลาถัดไป วงจร PISO นอกจากจะส่งข้อมูลอินพุตไปยังวงจร COF แล้วยังส่งข้อมูลไปยังวงจร MUX ซึ่งเปรียบได้กับวงจรแอนด์ เนื่องจากข้อมูลเลื่อนเข้ามาทีละบิตเพื่อคูณกับค่าสัมประสิทธิ์ของวงจรกรองแบบปรับตัว จึงเสมือนเป็นการแอนด์กันระหว่างข้อมูลอินพุตที่เข้ามาทีละบิต (มีค่าเป็น 0 หรือ 1 เท่านั้น) กับค่าสัมประสิทธิ์ของวงจรกรองแบบปรับตัว เมื่อแอนด์กันเสร็จแล้วจะนำผลที่ได้มารวมกันและทำการสเกล (Scale) ค่า โดยการสเกลค่าเป็นการเติมศูนย์เข้าที่ตำแหน่ง LSB (Least Significant Bit) จึงเสมือนเป็นการเลื่อนข้อมูลทั้งหมดไปอยู่ทางซ้ายมือสุด

โครงสร้างภายในของวงจร COF ดังแสดงในรูปที่ 43 สร้างโดยใช้ภาษา VHDL โดยวงจร mux16v1\_1 เปรียบได้กับวงจร MUX วงจร scale16to32 เปรียบได้กับ Scale วงจรแอกคูมูเลต (Accumulate) ซึ่งประกอบด้วย reg1\_32v2, reg2\_32v2, add\_sub32v1, shift\_r32 และ scaledown15bit เปรียบได้กับวงจรบวก/ลบ, วงจร Reg และวงจรเลื่อนขวา (>>) วงจร sat\_add32v1, reg5x32v1 และ dff32bitv1 เปรียบได้กับวงจรบวกและ FIFO ในรูปที่ 42



รูปที่ 43 โครงสร้างภายในของ COF

ในส่วนของวงจรแอกคูมูเลเตอร์ ซึ่งประกอบไปด้วยวงจรที่ทำหน้าที่เป็นทั้งวงจรบวกและวงจรถบในตัว โดยมีสัญญาณควบคุมการบวก/ลบ วงจรดังกล่าวนี้จะทำงานเป็นวงจรถบเพียงครั้งเดียวเท่านั้นในแต่ละชุดของข้อมูลอินพุต 1 คำ คือเมื่อข้อมูลอินพุตตำแหน่ง MSB (Most Significant Bit) เข้ามา ทั้งนี้เนื่องจากระบบตัวเลขที่ใช้ในการประมวลผลเป็นแบบคอมพลีเมนต์ที่ 2 (2's Complement) ซึ่งตำแหน่ง MSB จะเป็นบิตแสดงเครื่องหมาย (Sign Bit) ส่วนวงจรรีจิสเตอร์ (REG) จะทำหน้าที่พักข้อมูลและเป็นกันชน (Buffer) เพราะข้อมูลที่เอาต์พุตจะถูกป้อนกลับมายังวงจรบวก/ลบอีกครั้งโดยผ่านทางวงจรเลื่อนขวา แต่เนื่องจากทั้งวงจรเลื่อนขวาและวงจรถบ/ลบ เป็นวงจรแบบคอมบินเนชัน (Combination Circuit) ซึ่งไม่มีหน่วยความจำ (Memory) ในตัว ทำให้ต้องมีการเพิ่มวงจรที่ทำหน้าที่เป็นหน่วยความจำ เพื่อเก็บข้อมูลก่อนทำการประมวลผลของข้อมูลแต่ละชุดในวงจร เมื่อประมวลผลเสร็จแล้วจะส่งผลลัพธ์มายังวงจรในส่วนของวงจรถบสะสม เพื่อทำการบวกแบบสะสม โดยวงจรในส่วนนี้เทียบได้กับการประมวลผลแบบ MAC โดยจะบวกสะสมเป็นจำนวนรอบเท่ากับจำนวนตำแหน่งของข้อมูลภายในวงจร FIFO จึงจะทำให้ได้ข้อมูลเอาต์พุตเพื่อนำไปลบกับข้อมูลของสัญญาณที่ต้องการ  $d(n)$  ต่อไป

ในการสร้างวงจรโดยใช้ภาษา VHDL ได้เริ่มสร้างวงจรขนาด 32 แท็บก่อนเพื่อให้ทราบถึงปัญหาต่างๆ ก่อนที่จะสร้างวงจรที่มีจำนวนแท็บมากขึ้น ซึ่งโครงสร้างของวงจรถบแบบปรับตัวที่ใช้เทคนิค DA ร่วมกับ MAC นี้สามารถเพิ่มจำนวนแท็บให้มากขึ้นในอนาคต โดยไม่ต้องเริ่มต้นสร้างวงจรใหม่อีกครั้ง การเพิ่มจำนวนแท็บสามารถทำได้ 2 วิธีคือ การเพิ่มจำนวนรอบในการประมวลผล (ในกรณีที่ต้องการประหยัดทรัพยากรในการสร้างวงจร) หรือ เพิ่มขนาดของวงจรในการประมวลผลแต่ยังคงใช้จำนวนรอบในการประมวลผลเท่าเดิม (ในกรณีที่ต้องการประสิทธิภาพทางด้านความเร็วในการประมวลผล)

## บทที่ 4

### สรุปผลการทดลองและข้อเสนอแนะ

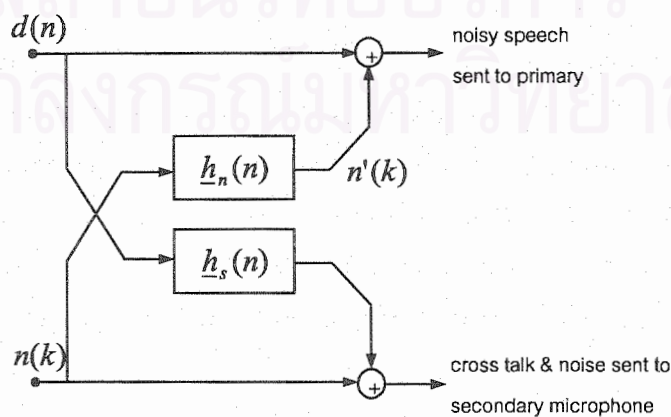
#### 4.1 สรุปผลการทดลอง

การกำจัดสัญญาณเสียงสะท้อนในระบบ AEC ในสถานการณ์ ST ที่ศึกษาในโครงการวิจัยนี้ สามารถถูกเพิ่มประสิทธิภาพการทำงานได้ โดยการประยุกต์ใช้งานระบบ AEC ร่วมกันกับระบบ ANC เพื่อกำจัดหรือลดผลของสัญญาณรบกวน นอกจากนี้ สัญญาณเสียงสะท้อนที่เหลืออยู่จากการกำจัดในขั้นแรก โดยระบบ AEC สามารถถูกกำจัดทิ้งได้เพิ่มเติมเมื่อระบบ ANC ถูกนำมาใช้งานโดยทำการต่อรวมเข้าไปในระบบ AEC ดังรูปที่ 35

เมื่อเปรียบเทียบประสิทธิภาพการทำงานของระบบที่นำเสนอ พบว่าให้ผลใกล้เคียงกับมาตรฐาน ITU Standard G.165 นอกจากนี้ คุณภาพของสัญญาณเสียงที่ส่งออกไปในระบบมีความชัดเจนมากขึ้น ดังเห็นได้จากผลการทดลองในหัวข้อที่ 3.5

#### 4.2 ข้อเสนอแนะและส่วนที่จะดำเนินการต่อไป

- สำหรับในเฟสต่อไปของโครงการวิจัยนี้ ต้องการศึกษาความเป็นไปได้ในการพัฒนาการกำจัดสัญญาณเสียงสะท้อนของระบบ AEC บนบอร์ด FPGA เพื่อทดสอบในเวลาจริง ที่มีเทคนิคการใช้งานร่วมกันของระบบ AEC กับระบบ ANC เพื่อเพิ่มประสิทธิภาพในการกำจัดสัญญาณรบกวนหรือความเป็นไปได้ในการพัฒนาระบบ AEC บนบอร์ด FPGA ที่มีโครงสร้างแบบความถี่ย่อย (Subband Structure) ที่ได้พัฒนาขึ้นในเฟสที่สาม เพื่อเพิ่มประสิทธิภาพการทำงานของระบบที่มีความซับซ้อนในการคำนวณต่ำกว่าระบบ AEC แบบปกติที่ทำงานในโดเมนเวลา
- พัฒนาระบบ ANC โดยคำนึงถึงความเหมือนจริงของระบบในทางปฏิบัติมากยิ่งขึ้น จากรูปที่ 24 จะเห็นว่าในทางปฏิบัติจริงสัญญาณรบกวนที่เราจัดหามา จะได้รับผลกระทบจากสัญญาณเสียงพูด ด้วยการรบกวนดังกล่าวถูกเรียกว่า cross talk ดังรูปที่ 44



รูปที่ 44

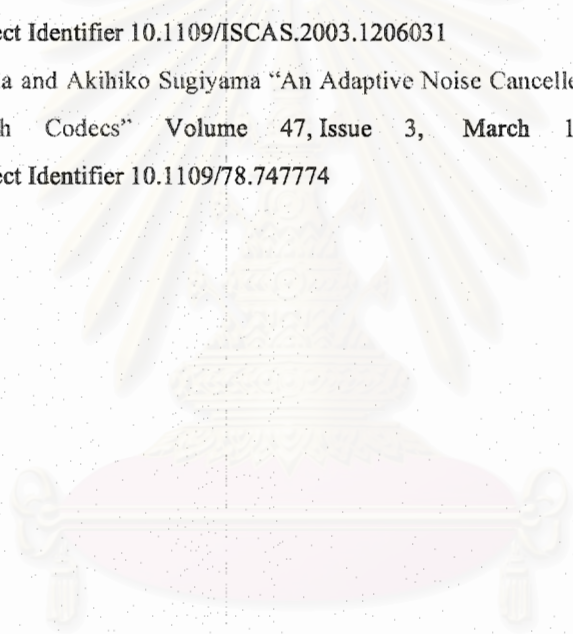
## เอกสารอ้างอิง

1. C. Breining, *et. al.*, "Acoustic Echo Control – An Application of Very-High-Order Adaptive Filters", IEEE Signal Processing Magazine, pp. 42 – 69, July 1999.
2. M. M. Sondhi, D. R. Morgan and J. L. Hall, "Stereophonic Acoustic Echo Cancellation – An Overview of the Fundamental Problem", IEEE Signal Processing Letters, vol. 2, no. 8, pp. 148-151, 1995.
3. J. Benesty, D. R. Morgan and M. M. Sondhi, "A better understanding and an improved solution to the problems of stereophonic acoustic echo cancellation", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'97, Munich, Germany), vol. 1, pp. 303-306, 1997.
4. N. Tangsangiumvisai, J.A. Chambers and A. G. Constantinides, "Higher-order time-varying allpass filters for signal decorrelation in stereophonic acoustic echo cancellation", IEE Electronics Letters, vol. 35, no. 1, pp. 88-90, 1999.
5. N. Forsyth, N. Tangsangiumvisai, P. A. Naylor and J. A. Chambers, "A subband signal decorrelation scheme for stereophonic acoustic echo cancellation using higher-order time-varying allpass filters", International Workshop on Acoustic Echo and Noise Control (IWAENC'99, Pennsylvania, USA.), pp. 24-27, 1999.
6. S. Haykin, "Adaptive Filter Theory", Prentice Hall, Englewood Cliffs, NJ, 3rd Edition, 1996.
7. J. Benesty, F. Amand, A. Gilloire and Y. Grenier, "Adaptive Filtering Algorithms for Stereophonic Acoustic Echo Cancellation", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'95, Detroit, USA.), vol. 5, pp. 3099-3102, 1995.
8. N. Tangsangiumvisai and J. A. Chambers, "A fast converging and stable adaptive filtering algorithm for stereophonic acoustic echo cancellation", European Signal Processing Conference (EUSIPCO'2000, Tampere, Finland), vol. 2, 2000.
9. A. Gilloire, E. Moulines, D. Slock and P. Duhamel, "State of the Art in Acoustic Echo Cancellation" in "Digital Signal Processing in Telecommunications", Springer-Verlag, NY, pp. 45-91, 1996.
10. N. Ithisoponkul and N. Tangsangiumvisai, "A Double-Talk Detection Technique using the Correlation Behaviour of the Gradient Vector for Acoustic Echo Cancellation", Proceedings of IEEE International Midwest Symposium on Circuits and Systems (MWSCAS'04, Hiroshima, Japan), July 2004.
11. P. Surin, N. Tangsangiumvisai, and S. Aramvith, "An Adaptive Noise Decorrelation Technique for Stereophonic Acoustic Echo Cancellation", to be published in the Proceedings of annual technical conference of IEEE region 10 (TENCON'04, Chiang Mai, Thailand), November 2004.



12. ITU Recommendation G.723.1
13. ITU Recommendation G.729
14. S. Gudvangen and S. J. Flockton, "Modelling of acoustic transfer functions for echo cancellers", IEE Proceedings of Vision, Image and Signal Processing, vol. 142, no. 1, 1995.
15. A. P. Liavas and P. A. Regalia, "Acoustic echo cancellation : Do IIR models offer better modeling than their FIR counterparts?" IEEE Transactions on Signal Processing, vol. 46, no. 9, 1998.
16. A. Gilloire and M. Vetterli, "Adaptive Filtering in Subbands with Critical Sampling : Analysis, Experiments and Application to Acoustic Echo Cancellation", IEEE Transactions on Signal Processing, vol. 40, no. 8, pp. 1862-1875, 1992.
17. J. J. Syhnk, "Frequency-domain and multirate adaptive filtering", IEEE Signal Processing Magazine, vol. 9, pp. 14-37, 1992.
18. S. Shimauchi and S. Makino, "Stereo projection echo canceller with true echo path estimation", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'95 Detroit, USA.), vol. 5, pp. 3059-3062, 1995.
19. P. A. Naylor, O. Tanrikulu and A. G. Constantinides, "Subband Adaptive Filtering for Acoustic Echo Cancellation Using Allpass Polyphase IIR Filter-banks", IEEE Transactions on Speech and Audio Processing, vol. 6, no. 2, pp. 143-155, 1998.
20. D. R. Morgan and J. C. Thi, "A delayless subband adaptive filter architecture", IEEE Transaction on Signal Processing, vol. 43, no. 8, pp. 1819-1830, 1995.
21. N. Tangsangiumvisai and J. A. Chambers, "A fullband-subband automatic switching scheme for stereophonic acoustic echo cancellation", IEE Electronics Letters, vol. 36, no. 25, pp. 2109-2111, 2000.
22. S. Minami and T. Kawasaki, "A Double Talk Detection Method for an Echo Canceller", IEEE International Conference on Communications (ICC), pp. 1492-1497, 1985.
23. H. Ye and B. Wu, "A New Double-Talk Detection Algorithm Based on the Orthogonality Theorem", IEEE Transactions on Communications, vol. 39, no. 11, pp. 1542-1545, 1991.
24. C. Carlemalm, F. Gustaffson and B. Wahlberg, "On the problem of detection and discrimination of double talk and change in the echo path", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'96, Atlanta, USA.), vol. 5, pp. 2742-2745, 1996.
25. C. Carlemalm and A. Logothetis, "On detection of double talk and changes in the echo path using a markov modelulated channel model", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'97, Munich, Germany), vol. 5, pp. 3869-3872, 1997.
26. S. J. Pushparajah and J. A. Chambers, "A smater method for Acoustic Echo Cancellation in the presence of Double Talk", International Workshop on Acoustic Echo and Noise Control (IWAENC'99, Pennsylvania, USA.), pp. 184-186, 1999.

27. J. Benesty, D. R. Morgan and M. M. Sondhi, "A Hybrid Mono/Stereo Acoustic Echo Canceled", IEEE Trans. on Speech and Audio Processing, vol. 6, no. 5, pp. 4680-475, September 1998.
28. G. V. Moustakidis and S. Theodoridis, "Fast Newton Transversal Filters – A new class of adaptive estimation algorithms", IEEE Trans. on Signal Processing, vol. 39, pp. 2184-2193, October 1991.
29. S. L. Gay and S. Tavathia, "The Fast Affine Projection Algorithm", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'95), pp. 3023-3026, May 1995.
30. E. R. Ferrara, "Fast Implementation of LMS Adaptive Filters", IEEE Trans. on Acoust. Speech and Signal Processing, vol. 28, no. 4, pp. 474-475, August 1980.
31. Widrow Bernard (US) "Speech Enhancement in the presence of background noise"
32. M. Gabrea "Two Microphones Speech Enhancement System Based On a Double Affine Projection Algorithm" Volume 2, 25-28 May 2003 Page(s):II-544 - II-547 vol.2 Digital Object Identifier 10.1109/ISCAS.2003.1206031
33. Shigeji Ikeda and Akihiko Sugiyama "An Adaptive Noise Canceller with Low Signal Distortion for Speech Codecs" Volume 47, Issue 3, March 1999 Page(s):665 - 674 Digital Object Identifier 10.1109/78.747774



สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก

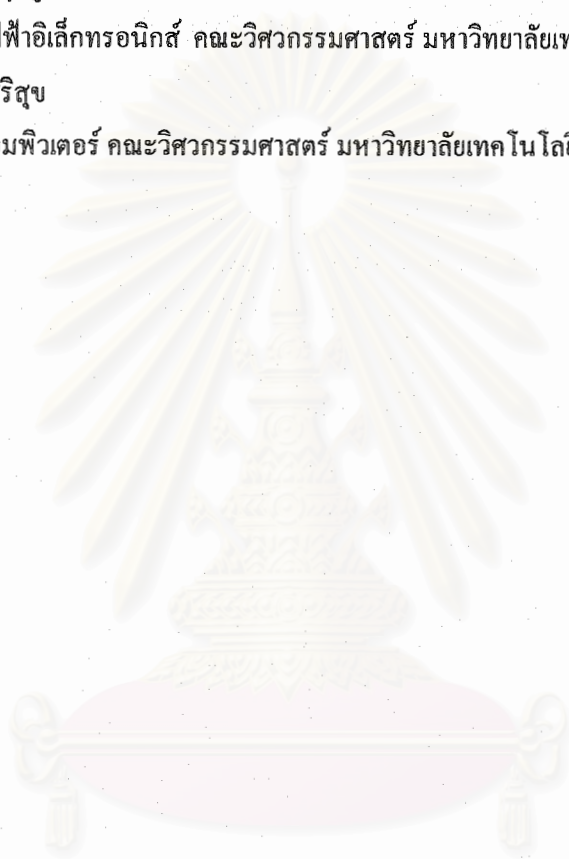
รายชื่อผู้ทรงคุณวุฒิพร้อมสถานที่การทำงานที่สามารถประเมินโครงการน้อย

• ดร. พีระพล ยวภูมิตานนท์

ภาควิศวกรรมไฟฟ้าอิเล็กทรอนิกส์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีมหานคร

• ดร. เผ่าศักดิ์ ศิริสุข

ภาควิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีมหานคร



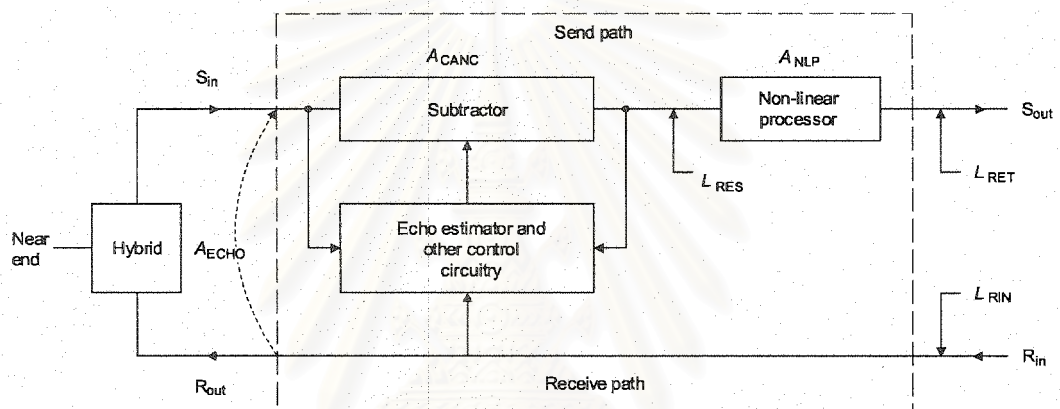
สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

ITU-T G.165

Recommendation G.165

Recommendation G.165 เป็นเอกสารมาตรฐานสำหรับอ้างอิงในการออกแบบหรือพัฒนา Echo Cancellers โดยในรายงานฉบับนี้จะกล่าวโดยสังเขปเกี่ยวกับตัวชี้วัดมาตรฐานและค่าเกณฑ์ในการจำกัดขีดความสามารถในการทำงานของ Echo Cancellers

1. นิยามตัวชี้วัดที่เกี่ยวข้องกับ Echo Canceller ตามมาตรฐาน Recommendation G.165 มีดังนี้



รูปที่ 1

echo loss ( $A_{ECHO}$ )

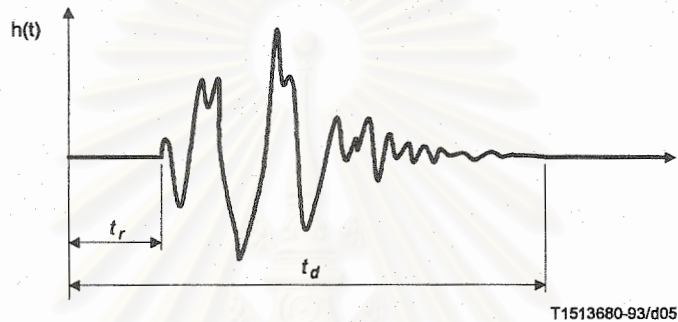
การลดทอนของสัญญาณจาก receive-out port ( $R_{out}$ ) ถึง send-in port ( $S_{in}$ ) ของ echo canceller ระหว่างการส่ง

pure delay ( $t_r$ )

เวลาประวิงจาก  $R_{out}$  port ถึง  $S_{in}$  port เป็นเวลาประวิงที่เกิดในวิธีการสะท้อนห้องใกล้ (near end)

(near-end) echo path or end delay ( $t_d$ )

เป็นผลสุดท้ายของการรบกวนกันของ pure delay แต่ละ path ซึ่งแสดงให้เห็นเป็นภาพได้ดังรูปที่ 2 คำนิยามนี้แสดงให้เห็นเพียง echo path เดียว ถ้าเป็นหลาย path จะทำให้  $t_d$  ของ path ทั้งหมด กลายเป็น echo path หรือ end delay ที่มีค่ามากที่สุด ใน path ต่างๆ เหล่านั้น



รูปที่ 2

**cancellation ( $A_{CANC}$ )**

การลดทอนของสัญญาณเสียงสะท้อนเมื่อผ่านวิถีส่ง echo canceller นิยามนี้รวมถึงกระบวนการไม่เชิงเส้นที่เกิดขึ้นที่ฝั่งออกของ canceller

**residual echo level ( $L_{RES}$ )**

ระดับของสัญญาณเสียงสะท้อนคงเหลือที่ send-out port ของ echo canceller หลังจากการกำจัดที่ไม่สมบูรณ์ของวงจรการ สัมพันธ์กับค่า receive-in signal  $L_{Rin}$  โดย

$$L_{RES} = L_{Rin} - A_{ECHO} - A_{CANC}$$

ไม่รวมกระบวนการไม่เชิงเส้นด้วย

**nonlinear processor (NLP)**

อุปกรณ์จะมีการนิยามระดับ suppression threshold เมื่อ:

- สัญญาณที่ตรวจได้มีระดับต่ำกว่า threshold จะถูกระงับ
  - สัญญาณที่ตรวจได้มีระดับสูง threshold จะถูกส่งผ่านโดยไม่มีการลดทอน
- หมายเหตุ

- 1 ความเที่ยงตรงของกระบวนการไม่เชิงเส้นการตรวจสอบและอัลกอริทึมที่

2 ตัวอย่างของกระบวนการไม่เชิงเส้น คือ analogue centre clipper เมื่อทุกสัญญาณที่ค่าต่ำกว่า threshold จะถูกปรับเป็นค่าเล็กๆ บางค่า

**nonlinear processing loss ( $A_{NLP}$ )**

การลดทอนของ residual echo level ลงเนื่องจากการผ่านเข้าไปใน nonlinear processor

หมายเหตุ – ถึงแม้ว่าการลดทอนของสัญญาณเนื่องจาก nonlinear processor ไม่อาจวัดได้ในหน่วย dB แต่การให้ความสำคัญกับการลดทอนนี้ควรถูกคำนึงถึงเสมอ

**returned echo level ( $L_{RET}$ )**

ระดับของสัญญาณที่ send-out port ของ echo canceller ซึ่งจะวนกลับมาที่ผู้พูด จะรวมผลของการลดทอนจากกระบวนการไม่เชิงเส้น  $L_{RET}$  สัมพันธ์กับ  $L_{Rin}$  โดย

$$L_{RET} = L_{Rin} - (A_{ECHO} + A_{CANC} + A_{NLP})$$

กรณีที่ไม่มีการวนกลับมาที่ผู้พูด  $L_{RES} = L_{RET}$

**combined loss ( $A_{COM}$ )**

ผลรวมของ echo loss, cancellation loss และ nonlinear processing loss (ถ้ามี). นิยามกับ  $L_{Rin}$  โดยที่

$$L_{RET} = L_{Rin} - A_{COM}, \text{ where } A_{COM} = A_{ECHO} + A_{CANC} + A_{NLP}$$

**convergence time**

เวลาที่ระบบจำลองใช้ในการปรับตัว เริ่มนับตั้งแต่ให้สัญญาณเข้าเข้าไปในระบบ จนกระทั่งสัญญาณออกจากระบบจำลองมีค่าเข้าใกล้สัญญาณจากระบบจริงเกินค่าที่กำหนดค่าหนึ่ง

## 2. คุณลักษณะสมบัติของ Echo Cancellor

Recommendation G.165 ใช้สำหรับออกแบบ echo cancellers ที่สมมุติให้การกำจัดเสียงสะท้อนกระทำที่ภาคส่งเท่านั้น แล้วจึงส่งสัญญาณไปภาครับ

เสียงสะท้อนใน 2-wire หรือวงจรโทรศัพท์ที่รวม 2- และ 4-wire เกิดจากการไม่สอดคล้องกันของความต้านทานเชิงซ้อน จึงใช้ echo canceller เพื่อลดเสียงสะท้อนให้อยู่ในเกณฑ์ที่รับได้

เสียงสะท้อนที่เกิดขึ้นใน send-in port ของ echo canceller คือการลดทอนและเวลาประวิงที่ย้อนกลับมาจากห้องไกล เช่น เสียงสะท้อนของเสียงเข้าที่ถูกเปลี่ยนแปลงด้วยวิธีการสะท้อน ส่วนมากวิธีสะท้อนจะวิเคราะห์เป็นผลตอบอิมพัลส์ (ดังรูปที่ 2). ผลตอบของวิธีสะท้อนแสดงในรูป pure delay  $t_r$ , เนื่องจากเวลา

ประวิงภายในของการส่งและการขยายตัวของสัญญาณ ผลรวมของเวลาประวิงวิธีสะท้อนคือ  $t_d$  ค่าของเวลาประวิงและการขยายตัวจะแปรตามคุณสมบัติของวิธีสะท้อน เช่น อาจแปรตามความแตกต่างขอโครงข่ายของแต่ละประเทศ

จะทำการสมมติว่า echo path จะเป็นการเปลี่ยนแปลงแบบเชิงเส้นเท่านั้น และ echo path ไม่ได้เป็นระบบที่เปลี่ยนแปลงตลอดเวลา ประสิทธิภาพของ echo canceller จะขึ้นอยู่กับความเป็นเชิงเส้นของ echo path ระหว่าง  $R_{out}$  ไปยัง  $S_{in}$

วิธีสะท้อนอาจเป็น ได้ทั้งสายแอนาล็อกและดิจิทัล ระดับสัญญาณที่ถูกตัดยอดของสายดิจิทัลได้นิยามไว้ใน Recommendation G.711 ที่ระดับยอดคลื่น 3.1 dBm0 sine wave การประยุกต์ใช้ระดับตัดยอดจะกระทำก่อนที่สัญญาณจะส่งเข้า  $R_{in}$  เพื่อไม่ให้สมรรถนะของการกำจัดเสียงสะท้อนลดลงมาก สำหรับสัญญาณระดับสูง

การสูญเสียของวิธีสะท้อนในหน่วย dB (หัวข้อ 2.2) คือการสูญเสียที่น้อยที่สุดจาก  $R_{out}$  ถึง  $S_{in}$  ของ echo canceller จะเท่ากับผลต่างของระดับทั้งสองช่องทางบวก 6 dB การออกแบบ Echo cancellers ใน Recommendation นี้จะให้การลดทอนวิธีสะท้อน (AECHO) เท่ากับ 6 dB หรือมากกว่านั้น สำหรับ (AECHO) ที่น้อยกว่า 6 dB Echo cancellers จะยังทำงานได้แต่ให้ประสิทธิภาพที่ลดลง

Echo canceller ต้องสามารถจำลอง echo path impulse response ได้ โดยที่หลายๆ กรณีใช้ sampling rate = 8000 Hz ตาม Nyquist rate ดังนั้นจำนวน sample ของ echo estimator จะต้องเหมาะสมสำหรับ path ที่ทำการประมาณ เนื่องจากหากใช้ น้อยเกินไปก็จะทำให้ไม่สามารถประมาณ path ได้อย่างถูกต้อง และหากใช้มากเกินไปจะทำให้เกิดสัญญาณรบกวนขึ้นเนื่องจากส่วนหางของ impulse response จะไม่เป็น 0

เมื่อสัญญาณเสียงวนกลับถูกส่งไปพร้อมกับสัญญาณ double talk ตัว echo canceller อาจแปลความหมายสัญญาณทั้งหมดเป็นเสียงสะท้อน และพยายามกำจัด ซึ่งจะเป็นการลดคุณภาพของสัญญาณ echo cancellation ไม่ได้แค่กำจัดสัญญาณเสียงสะท้อนเท่านั้น แต่ยังทำการลดทอนสัญญาณเสียง double talk ด้วย ดังนั้นในหัวข้อนี้จึงเสนอ 2 วิธีเพื่อแก้ไขปัญหานี้ วิธีแรกคือการใช้อัลกอริทึม ซึ่งทำให้การปรับปรุงสัมประสิทธิ์ช้าลงในช่วงที่เกิด double talk อีกวิธีคือการใช้ double talk detector ซึ่งคล้ายกับการใช้ echo suppressors

ดังนั้น echo cancellers ต้องมีหลักพื้นฐานดังต่อไปนี้:

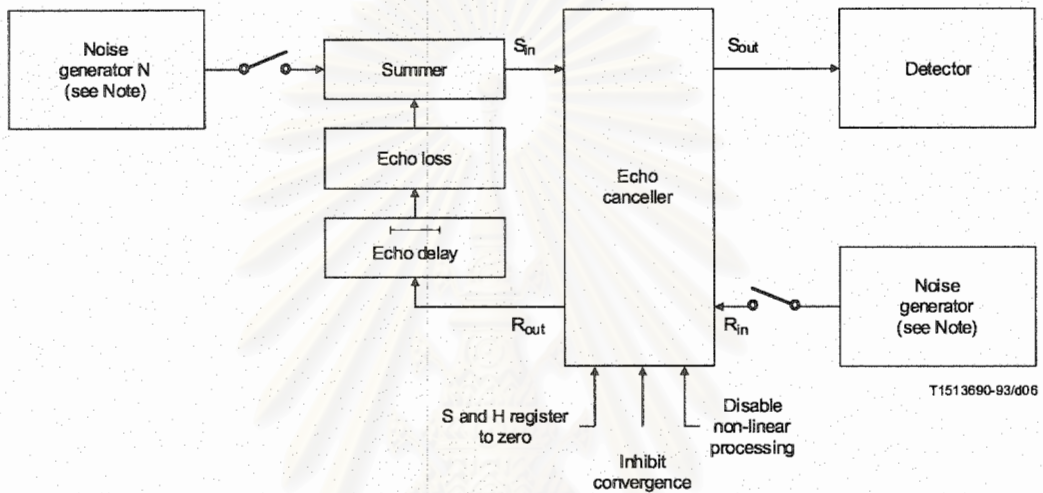
- 1) การลู่เข้าที่รวดเร็ว
- 2) ให้ค่าเสียงสะท้อนป้อนกลับที่น้อยระหว่าง single talk

3) อัตราการลู่ออกซ้ำช่วงที่เกิด double talk

3. Echo canceller performance

ข้อกำหนดทางสมรรถนะของ echo cancellers จะรวมกระบวนการไม่เชิงเส้นด้วย (Annex A สำหรับ echo cancellers ซึ่งไม่รวมกระบวนการไม่เชิงเส้น)

Recommendation G.165 ประกอบด้วยการทดสอบหลายๆ การทดสอบเพื่อกำหนดสมรรถนะของ echo canceller โดยมีแบบจำลองของระบบทดสอบดังรูปที่ 3



NOTE – The requirements in 3.4.2 are based on the use of band-limited white noise (300-3400 Hz) as the test signal. Noise shaped in accordance with Recommendation G.227 may also be used. However, the applicability of the requirements in 3.4.2 requires confirmation and is under study.

The use of alternative test signals more representative of real speech and possible changes in test procedures and requirements are also under study.

รูปที่ 3

วัตถุประสงค์หลักของ echo canceller คือการควบคุมเสียงสะท้อนของสัญญาณเสียงพูด ด้วยการสังเคราะห์ผลตอบอิมพัลส์วิถีการสะท้อน เพื่อนำไปประมาณค่าเสียงสะท้อนแล้วลบออกจากวงจรเสียงสะท้อนจริง การสังเคราะห์ต้องใช้ได้กับสัญญาณเสียงเข้า แต่เนื่องจากการนิยามสัญญาณเสียงที่ใช้ทดลองทำได้ยาก ดังนั้นการทดลองนี้จึงใช้ band-limited noise เป็นสัญญาณทดสอบแทน และควรออกแบบกระบวนการไม่เชิงเส้นใน echo canceller ให้มีค่าน้อยที่สุด และหลีกเลี่ยงผลของการตัดยอดคลื่นในกรณี double-talk และ background noise contrast (ดัง test 9 described later in this Recommendation)



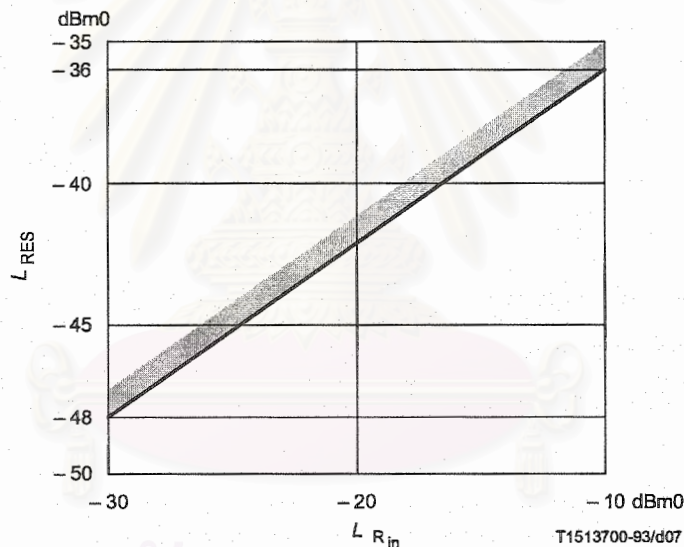
### Steady state residual and returned echo level test

การทดสอบนี้เพื่อยืนยันว่าการกำจัดเสียงสะท้อนที่จุด steady state (ACANC) จะคงเหลือค่าเสียงสะท้อนน้อยเพียงพอที่จะใช้กระบวนการไม่เชิงเส้น

กำหนดค่าเริ่มต้นของ H register เป็นศูนย์ และรับสัญญาณเข้านานพอที่เสียงสะท้อนคงเหลือจะเข้าสู่ steady state

### Requirement (provisional)

ด้วยค่าเริ่มต้นของ H register เป็นศูนย์ การบวนการไม่เชิงเส้นจะตัดการทำงานถ้าสัญญาณรับเข้า  $L_{Rin} \geq -30$  dBm0 และ  $\leq 0$  dBm0 และทุกค่า echo loss  $\geq 6$  dB และเวลาประวิงของวิธีสะท้อน  $t_d \leq \Delta$  ms<sup>4)</sup> ระดับเสียงสะท้อนคงเหลือต้องน้อยกว่าหรือเท่ากับระดับเสียงสะท้อนดังรูปที่ 4

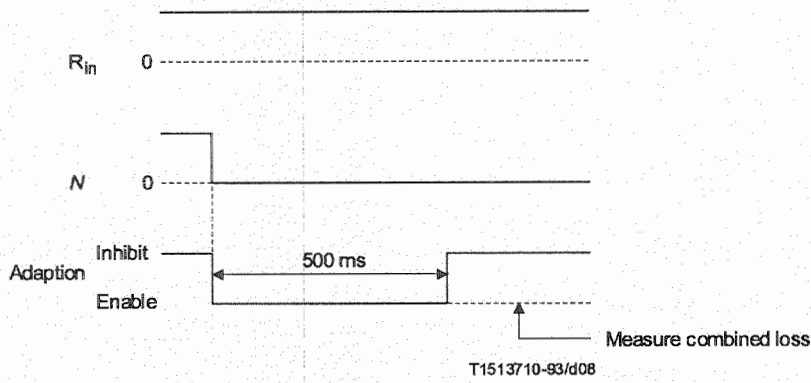


รูปที่ 4

### Convergence test

การทดสอบนี้เพื่อยืนยันว่า echo canceller จะเข้าสู่ค่าตอบได้เร็วพอทุกๆระดับสัญญาณเข้าและวิธีสะท้อน และระดับเสียงสะท้อนคงเหลือจะน้อยเพียงพอ กำหนดให้ H register ที่ค่าเริ่มต้นเป็นศูนย์และปรับปรุงตามรอบกาคำนวณ double talk detector ตั้งอยู่ในสถานะทำงาน ให้สัญญาณเข้าทาง  $S_{in}$  และ  $R_{in}$  สัญญาณที่  $S_{in}$  จะถูกกำจัดและทำการปรับปรุงสัมประสิทธิ์ของ cancellers ทันที ระดับการปรับปรุงวัดด้วยระดับเสียงสะท้อนคงเหลือ จะขึ้นอยู่กับคุณสมบัติการเข้าสู่ของ echo canceller และ double talk detector

4) Different echo cancellers may be designed to work satisfactorily for different echo path delays depending on their application in various networks. Thus  $\Delta$ , whenever it appears in this Recommendation, represents the echo path delay,  $t_d$ , for which the echo canceller is designed.



กระบวนการทดสอบคือ เริ่มต้นให้ H register เป็น 0 และไม่อนุญาตให้ทำการปรับค้ว สัญญาณ N ถูกใส่เข้าไปที่ระดับ  $-10 \text{ dBm0}$  และสัญญาณใดๆ ถูกใส่เข้าที่  $R_{in}$  จากนั้น N จะถูกตัดทิ้ง ในขณะที่เดียวกับที่ อนุญาตให้ระบบปรับค้วได้ หลังจากนั้น 500 ms จึงทำการห้ามปรับค้วอีกครั้ง

**Requirement**

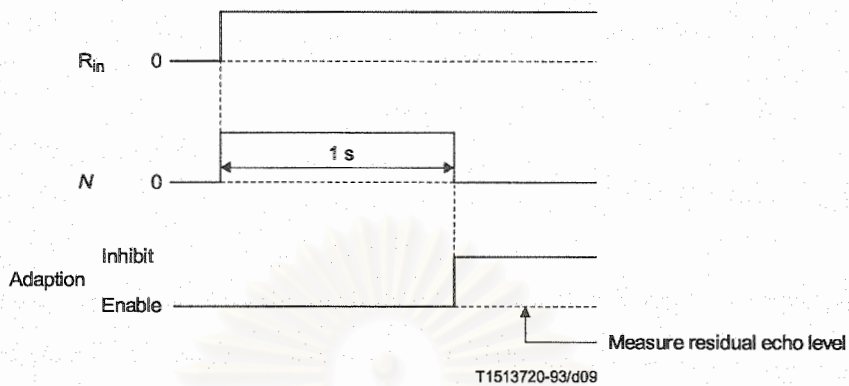
ด้วยค่าเริ่มต้นของ H register เป็นศูนย์ ทุกค่า  $L_{Rin} \geq -30 \text{ dBm0}$  และ  $\leq 0 \text{ dBm0}$  เมื่อเวลาผ่านไป 500 ms และทุกค่าการสูญเสียเสียงสะท้อน  $\geq 6 \text{ dB}$  และเวลาประวิงของวิถีสะท้อน  $t_d \leq \Delta \text{ ms}$  ผลรวมของการสูญเสีย ( $A_{COM} = A_{ECHO} + A_{CANC} + A_{NLP}$ ) ต้อง  $\geq 27 \text{ dB}$ .

**Performance under conditions of double talk**

การทดสอบนี้เป็นการทดสอบสมรรถนะของ canceller ภายใต้สภาวะ double talk ว่าสามารถป้องกันหรือชะลอการปรับปรุงสัมประสิทธิ์ เพื่อหลีกเลี่ยงการกำจัดสัญญาณที่มากเกินไป

1. การทดสอบนี้เพื่อยืนยันว่า double talk detection จะไม่ตัดสินใจลดการปรับปรุงสัมประสิทธิ์เข้ามีค่าต่ำ และทำให้ double talk detector หยุดการปรับปรุงสัมประสิทธิ์

กำหนดให้ H register มีค่าเริ่มต้นเป็นศูนย์ สำหรับเวลาประวิงของเสียงสะท้อนและการสูญเสียเสียงสะท้อนบางค่า ให้สัญญาณเข้าทาง  $R_{in}$  (ดังรูป 9) และสัญญาณรบกวนที่มีระดับต่ำเพียงพอที่จะไม่รบกวนการลู่เข้าของ echo canceller เข้าทาง  $S_{in}$  สัญญาณนี้จะไม่ทำให้ double talk detector ทำงาน หลังจากเวลาผ่านไป 1 s ให้หยุดการปรับปรุงสัมประสิทธิ์ และวัดค่าเสียงสะท้อนคงเหลือกระบวนการไม่เชิงเส้นควรจะ *disabled*

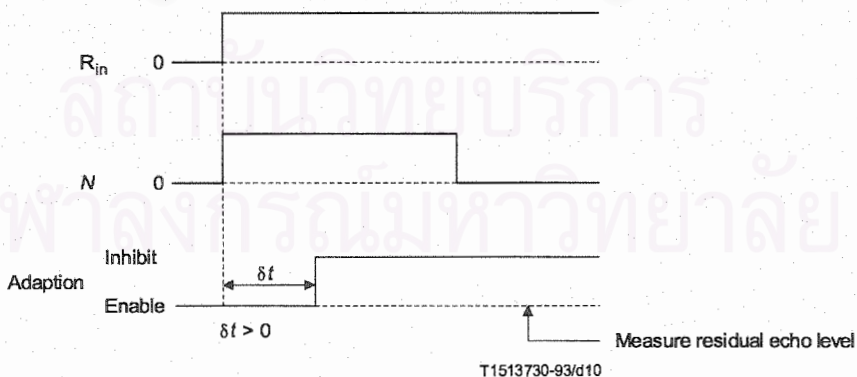


**Requirement**

ด้วยค่าเริ่มต้นของ H register เป็นศูนย์ สำหรับทุกค่า  $L_{Rin} \geq -25$  dBm0 และ  $\leq -10$  dBm0,  $N = L_{Rin} - 15$  dB,  $AECHO \geq \Delta$  6 dB และเวลาประวิงของวิถีสะท้อน  $t_d \leq \Delta$  ms การลู่เข้าควรเกิดขึ้นภายใน 1.0 s และ  $L_{RES}$  ต้องน้อยกว่า  $\leq N$

- การทดสอบนี้เพื่อให้มั่นใจว่า double talk detector จะทำงานได้เร็วเพียงพอที่จะป้องกันการลู่ออกกระหว่างที่เกิด double talking

กำหนดให้ตอนแรกระบบลู่เข้าสู่ค่าตอบ ให้สัญญาณเข้าทาง  $R_{in}$  ทันที (ดังรูป 10) และสัญญาณ  $N$  เข้าทาง  $S_{in}$  ซึ่งมีระดับน้อยกว่า  $R_{in}$  และทำให้ double talk detector ทำงาน เมื่อเวลา arbitrary time ผ่านไป  $\delta t > 0$  หยุดการทำงานของ canceller และวัดค่าเสียงสะท้อนคงเหลือ กระบวนการไม่เชิงเส้นควรจะ disabled



**Requirement**

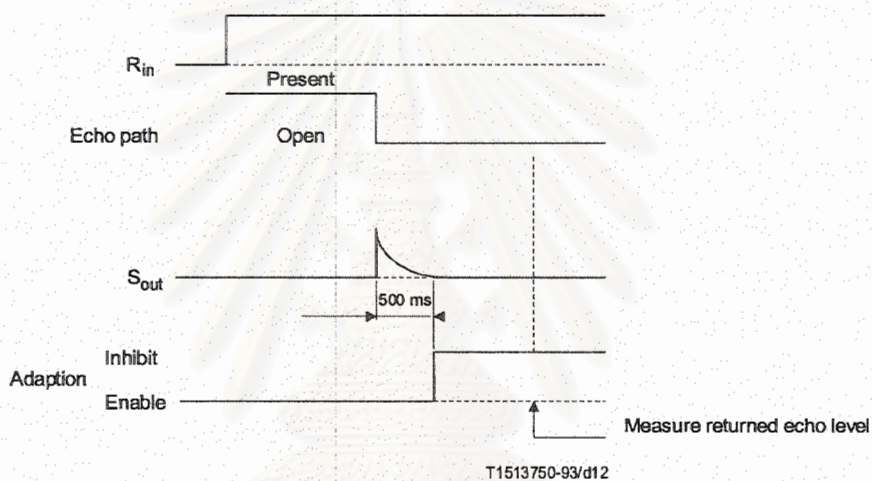
กำหนดให้ตอนแรกระบบลู่เข้าสู่ค่าตอบ สำหรับทุกค่า  $L_{Rin} \geq -30$  dBm0 และ  $\leq -10$  dBm0 และทุกค่า  $N \geq L_{Rin}$  การสูญเสียเสียงสะท้อน  $\geq 6$  dB และเวลาประวิงวิถีสะท้อน  $t_d \leq \Delta$  ms ระดับเสียงสะท้อน

คงเหลือเมื่อใส่สัญญาณเข้า  $L_{Rin}$  และ  $N$  ทุกคาบเวลาไม่ควรเพิ่มขึ้นเกิน 10 dB จากข้อกำหนด steady state test

#### Infinite return loss convergence test

การทดสอบนี้เพื่อให้มั่นใจว่า echo canceller มีค่าเฉลี่ยบางค่าที่จะป้องกันการเกิดเสียงสะท้อนที่ไม่ต้องการ โดยที่ H register จะเก็บรูปแบบวิธีการสะท้อนจากรอบการคำนวณก่อนหน้า หรือรอบการคำนวณปัจจุบัน และวิธีสะท้อนถูกเปิด (circuit echo vanishes) ในขณะที่สัญญาณเข้าสู่  $R_{in}$ .

กำหนดให้ตอนแรกระบบถูกรับเข้าสู่คำตอบ ค่าวิธีสะท้อนของ echo canceller ถูกปรับจนในขณะที่สัญญาณกำลังเข้า  $R_{in}$  หลังจากนั้น 500 ms ค่าสัญญาณการสะท้อนที่ออกทาง  $S_{out}$  ควรเป็นตาม (รูปที่ 12). กระบวนการไม่เชิงเส้นควร be disabled



#### Requirement (provisional)

กำหนดให้ตอนแรกระบบถูกรับเข้าสู่คำตอบ สำหรับทุกค่าการสูญเสียเสียงสะท้อน  $\geq 6$  dB และทุกค่า  $L_{Rin} \geq -30$  dBm0 และ  $\leq -10$  dBm0 ระดับเสียงสะท้อนคงเหลือที่  $S_{out}$  หลังจากทีวิธีสะท้อนถูกรับจนเป็นเวลา 500 ms ควรจะ  $\leq \Delta -37$  dBm0

#### สรุป

- Echo canceller ควรลดสัญญาณระหว่าง  $L_{Rin}$  กับ  $L_{RES}$  ให้ได้ประมาณ 20 dB (steady state condition)
- ช่วงเวลาที่ใช้ในการปรับตัวคือต้องเข้าสู่เงื่อนไข steady state condition ภายในเวลาไม่เกิน 500 ms
- ต้องมีอัตราการส่งออกที่ช้าเมื่อมีสัญญาณ double talk และ ณ ช่วงสั้นๆ ค่าหนึ่ง  $L_{RES}$  พลังงานขาออกไม่ควรสูงกว่า steady state condition เกิน 10 dB