



การประมวลเสียงพูด โดยการประมาณเชิงเส้น
VOICE PROCESSING BY LINEAR PREDICTION

โดย

นายสุนทร อรอินทร์ 34108439

นายอัฐ เครือฟัก 34109506

อาจารย์ที่ปรึกษา

รศ. ดร. กอบชัย เดชหาญ

ดร. ไกรสิน ส่งวัฒนา

ปริญญาโท สำหรับปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2537

ปริญญาานิพนธ์ปีการศึกษา 2537

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การประมวลเสียงพูดโดยการประมาณเชิงเส้น

ผู้จัดทำ

1. นายสุนทร อรอินทร์ เลขประจำตัว 34108439
2. นายอัฐ เครือฟัก เลขประจำตัว 34109506



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การประมวลเสียงพูดโดยการประมาณเชิงเส้น
VOICE PROCESSING BY LINEAR PREDICTION

โดย นายสุนทร อรอินทร์
นายอัฐ เครือฟัก

อาจารย์ที่ปรึกษา รศ. ดร. กอบชัย เดชหาญ
ดร. ไกรสิน ส่องวัฒนา

บทคัดย่อ

โครงการนี้เป็น การนำเสนอวิธีการประมวลสัญญาณเสียงพูดโดยการประมาณเชิงเส้น โดยใช้วิธีออโตคอร์เรเลชันมาคำนวณหาค่าพารามิเตอร์ของการประมาณเชิงเส้น และหาความถี่พิทช์ของเสียงพูดด้วยวิธี SIFT (Simplified Inverse Filter Tracking) จากนั้น ทำการวิเคราะห์ และสังเคราะห์ กับตัวอย่างสัญญาณเสียงพูด

Abstract

This project presents a method of linear prediction as applied to voice processing. The technique uses the autocorrelation method to calculate parameters of linear predictor. The pitch frequency of voice signal is found by using the simplified inverse filter tracking (SIFT) algorithm. The results are used in the analysis and synthesis of simple speech signal.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทที่ 1 บทนำ	1
บทที่ 2 ระบบการพูดและเสียงพูดของมนุษย์	3
บทที่ 3 แบบจำลองระบบกำเนิดเสียงพูด	9
บทที่ 4 หลักการประมวลเชิงเส้น	11
บทที่ 5 การคำนวณค่าสัมประสิทธิ์ของ LPC และอัตราการขยาย	16
บทที่ 6 เกณฑ์การตัดสินใจเสียงว้อยซ์ และเสียงอันว้อยซ์ และการหาคาบเวลาพิทช์ของเสียงพูดด้วยวิธี SIFT	21
บทที่ 7 ผลการทดลอง	25
บทที่ 8 สรุป	34
บทแทรกที่ 1 โปรแกรมที่ใช้ในการคำนวณ	36
บทแทรกที่ 2 ตัวอย่างการวิเคราะห์และสังเคราะห์สัญญาณเสียง	48
บทแทรกที่ 3 การออกแบบวงจรกรองเชิงเลข	71
เอกสารอ้างอิง	76
กิตติกรรมประกาศ	78

บทที่ 1

บทนำ

ในยุคโลกาภิวัตน์นี้ นับได้ว่าเป็นยุคแห่งการสื่อสารโดยแท้ เนื่องจากไม่ว่าจะอยู่ ณ มุมใดในโลกมนุษย์ย่อมต้องการติดต่อสื่อสารด้วยกันทั้งสิ้น ไม่ว่าจะเป็น ทางด้านธุรกิจการค้า หรือการติดต่อกับญาติสนิทมิตรสหาย ซึ่งมีได้หลายทาง แต่การติดต่อสื่อสารที่สะดวกและรวดเร็วที่สุด เห็นจะได้แก่ การสื่อสารโทรคมนาคม ด้วยเหตุนี้เองจึงทำให้การสื่อสารโทรคมนาคมเป็นที่นิยมกันอย่างแพร่หลายจนอาจจะกล่าวได้ว่า เป็นส่วนหนึ่งของชีวิตประจำวันก็ว่าได้ ซึ่งการโทรคมนาคมที่ใช้บ่อยและเห็นชัดที่สุดก็คือ โทรศัพท์ ซึ่งเป็นการติดต่อสื่อสารทางเสียงพูด ฉะนั้นการที่จะประหยัดคู่สายโทรศัพท์ ย่อมก่อประโยชน์แก่การสื่อสารเป็นอันมาก

การประมวลเชิงเส้นสามารถนำไปประยุกต์ใช้กับการสื่อสารทางเสียงได้หลายด้านเช่น การเก็บข้อมูลเสียง (voice mail) หรือการจดจำเสียงพูด (speech recognition) เนื่องจากวิธีนี้เป็นวิธีที่สามารถบีบอัด(compress) สัญญาณเสียงพูด ซึ่งมีมานานแล้วแต่ไม่เป็นที่แพร่หลายเนื่องจากต้องอาศัยการคำนวณที่ยุ่งยากและซับซ้อน จวบจนปัจจุบันพัฒนาการทางด้านสารกึ่งตัวนำ(semiconductor) ได้ก้าวล้ำจนสามารถสร้างเครื่องมือคำนวณที่มีประสิทธิภาพ และมีความเร็วสูง อีกทั้งยังมีการพัฒนาทางด้านการประมวลสัญญาณเชิงเลข(digital signal processing) ทำให้การประมวลสัญญาณเชิงเส้นมีความสำคัญมากยิ่งขึ้น

วัตถุประสงค์ในการทำปฏิญานิพนธ์ เรื่องการประมวลสัญญาณเชิงเส้นของเสียงพูด มีดังนี้

1. ศึกษาขั้นตอนการประมวลเชิงเส้นของสัญญาณเสียงพูด โดยวิธี ออโตคอร์รัเลชัน เพื่อนำมาใช้ในการทดลองการเลียนเสียงพูดของมนุษย์
2. สร้างแบบจำลองของระบบการกำเนิดเสียงพูด(vocal system) อย่างง่าย พร้อมทั้งวิเคราะห์หาค่าพารามิเตอร์ของระบบที่สร้างขึ้น
3. เปรียบเทียบสัญญาณที่สร้างขึ้นกับสัญญาณจริง ว่ามีความเชื่อถือได้มากน้อยเพียงไร
4. เปรียบเทียบประสิทธิภาพของการประมวลเชิงเส้น เมื่อเพิ่มอันดับในการประมวล

เนื่องจากปฏิญานิพนธ์นี้ เป็นการเริ่มต้นศึกษาหลักการประมวลเชิงเส้น ดังนั้น จึงเป็นเพียงการเขียนโปรแกรมเพื่อจำลองส่วนต่างๆ ในการประมวล และบันทึกผลออกมาในรูปของกราฟ โดยใช้ดิิจิตอลคอมพิวเตอร์ และ โปรแกรมที่ใช้ในการเขียนชื่อว่า 'MATLAB' ขั้นตอนของปฏิญานิพนธ์แบ่งออกเป็น 8 บทดังนี้

บทที่ 2 จะกล่าวถึงระบบการพูดของมนุษย์ ในด้านองค์ประกอบทางสรีระ และหน้าที่ของอวัยวะต่างๆ ของมนุษย์ในการเปล่งเสียงพูด เพื่อให้เกิดแนวคิดอย่างกว้างๆ เกี่ยวกับแบบจำลองระบบกำเนิดเสียงพูด ซึ่งจะได้อีกกล่าวถึงต่อไป

บทที่ 3 จะกล่าวถึงแบบจำลองระบบกำเนิดเสียงพูด ซึ่งจะจำลองลักษณะทางกายภาพของอวัยวะที่ใช้ในการกำเนิดเสียงพูด ออกมาเป็นบล็อกไดอะแกรมอย่างง่าย

บทที่ 4 เป็นทฤษฎีพื้นฐานของการประมาณเชิงเส้น ซึ่งอาศัยหลักการ กำลังสองน้อยที่สุด (least square estimate) จนกระทั่งได้สูตรในการคำนวณค่าสัมประสิทธิ์ต่างๆ

บทที่ 5 จะเป็นขั้นตอนในการคำนวณค่าสัมประสิทธิ์ของการประมาณ และอัตราการขยาย ซึ่งจะกล่าวถึงการเตรียมสัญญาณก่อนการวิเคราะห์จนถึงขั้นการคำนวณจริง

บทที่ 6 เป็นขั้นตอนการหาคาบเวลาพิชต์ด้วยขั้นตอนวิธี SIFT ซึ่งอาศัยหลักการในการหาสัญญาณความผิดพลาด โดยลดอันดับในการประมาณเชิงเส้น นอกจากนี้ยังได้เสนอเกณฑ์ ในการตัดสินใจเสียงว้อยซ์ และเสียงอันว้อยซ์อีกด้วย

บทที่ 7 เป็นขั้นในการทดลองและวัดผล ซึ่งในที่นี้จะนำเสนอในรูปแบบของกราฟเป็นส่วนใหญ่ เนื่องจากการนำเสนอที่ชัดเจนพอที่จะเปรียบเทียบให้สามารถเข้าใจได้

บทที่ 8 บทนี้จะสรุปเกี่ยวกับการทดลองทั้งหมดที่ได้ทำมา

ในตอนท้ายของปฏิญานีพนธ์นี้ ได้รวบรวมโปรแกรมต่างๆ ที่ใช้ซึ่งเขียนโดยโปรแกรม MATLAB ในการคำนวณ และ ผลการทดลองของตัวอย่างสัญญาณเสียงจำนวนหนึ่ง

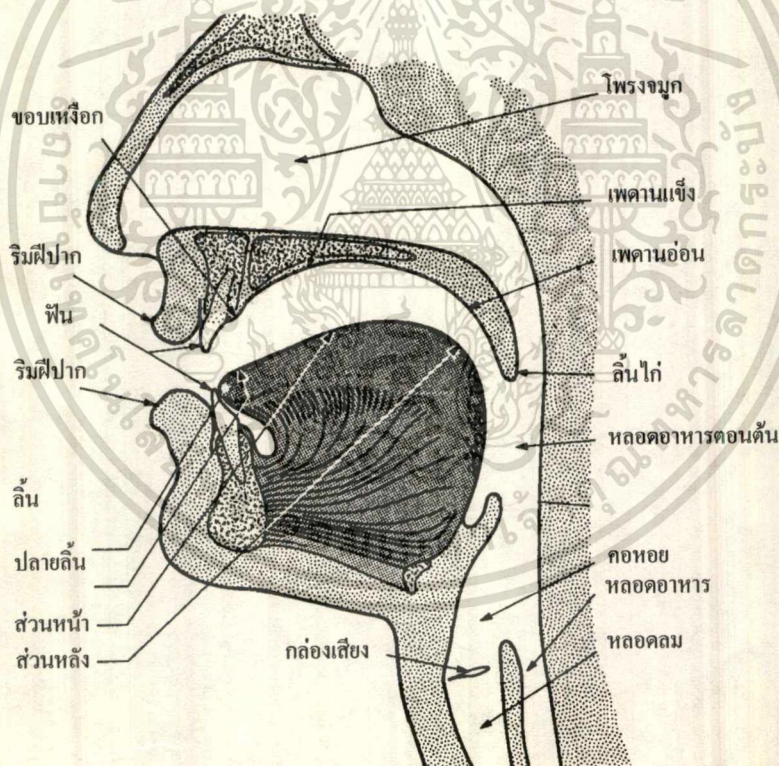
บทที่ 2

ระบบการพูดและเสียงพูดของมนุษย์

จากการศึกษาด้านกายวิภาคศาสตร์ของมนุษย์ (human anatomy) วิชาที่ว่าด้วยเสียงของภาษา (phonetics) และศาสตร์ทางด้านเสียง (acoustics) ช่วยให้เราเข้าใจขั้นตอนการทำงานร่วมกันของอวัยวะต่างๆ ในการเปล่งเสียงพูด ตลอดจนลักษณะทางกายภาพของเสียงพูด เพื่อนำมาวิเคราะห์และสร้างแบบจำลองเลียนแบบเสียงพูดของมนุษย์

2.1 กายวิภาคของระบบการเปล่งเสียงของมนุษย์

การทำให้เกิดเสียงเป็นหน้าที่หนึ่งของระบบหายใจ การออกเสียงหรือการพูดของมนุษย์แต่ละครั้ง จะต้องมีการทำงานร่วมกันของอวัยวะต่างๆ ของร่างกายดังรูปที่ 2.1 อันประกอบด้วย



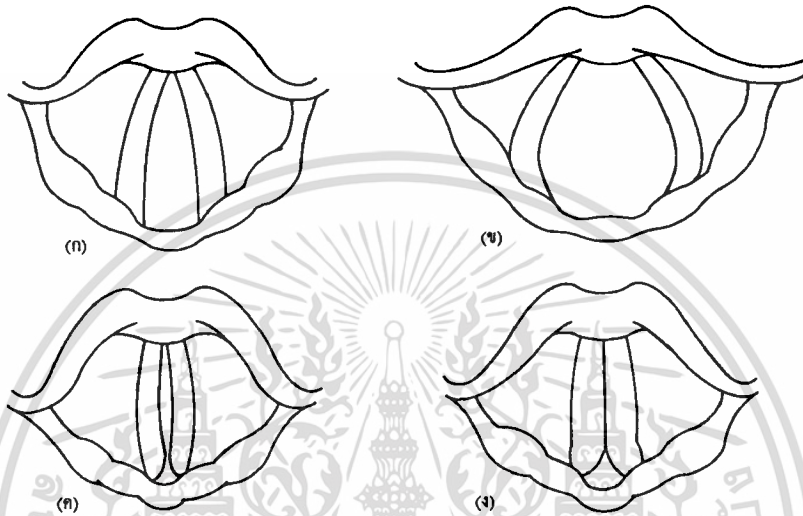
รูปที่ 2.1 ภาพตัดขวางแสดงอวัยวะในระบบการพูดของมนุษย์

2.1.1 ปอดและกระบังลม ทำหน้าที่สำคัญในการหายใจ และเป็นต้นกำเนิดการไหลของอากาศ
ในกระบวนการผลิตเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.2 หลอดลม ทำหน้าที่นำอากาศจากปอดผ่านกล่องเสียง และเป็นอวัยวะที่อยู่ด้านหน้าของหลอดอาหาร

2.1.3 กล่องเสียง เป็นอวัยวะพิเศษที่ทำหน้าที่เป็นทางเดินอากาศเวลาหายใจ และเป็นตัวผลิตพัลส์ (pulse) ของอากาศขณะเปล่งเสียง ซึ่งประกอบด้วยเส้นเสียง (vocal cords) และช่องสายเสียง (glottis) รูปร่างของกล่องเสียงและเส้นเสียงในลักษณะต่างๆ แสดงในรูปที่ 2.2



รูปที่ 2.2 รูปร่างกล่องเสียงขณะ (ก) หายใจปกติ (ข) หายใจเข้าลึกๆ (ค) กำลังส่งเสียง (ง) ส่งเสียงกระซิบหรือเสียงแผ่ว

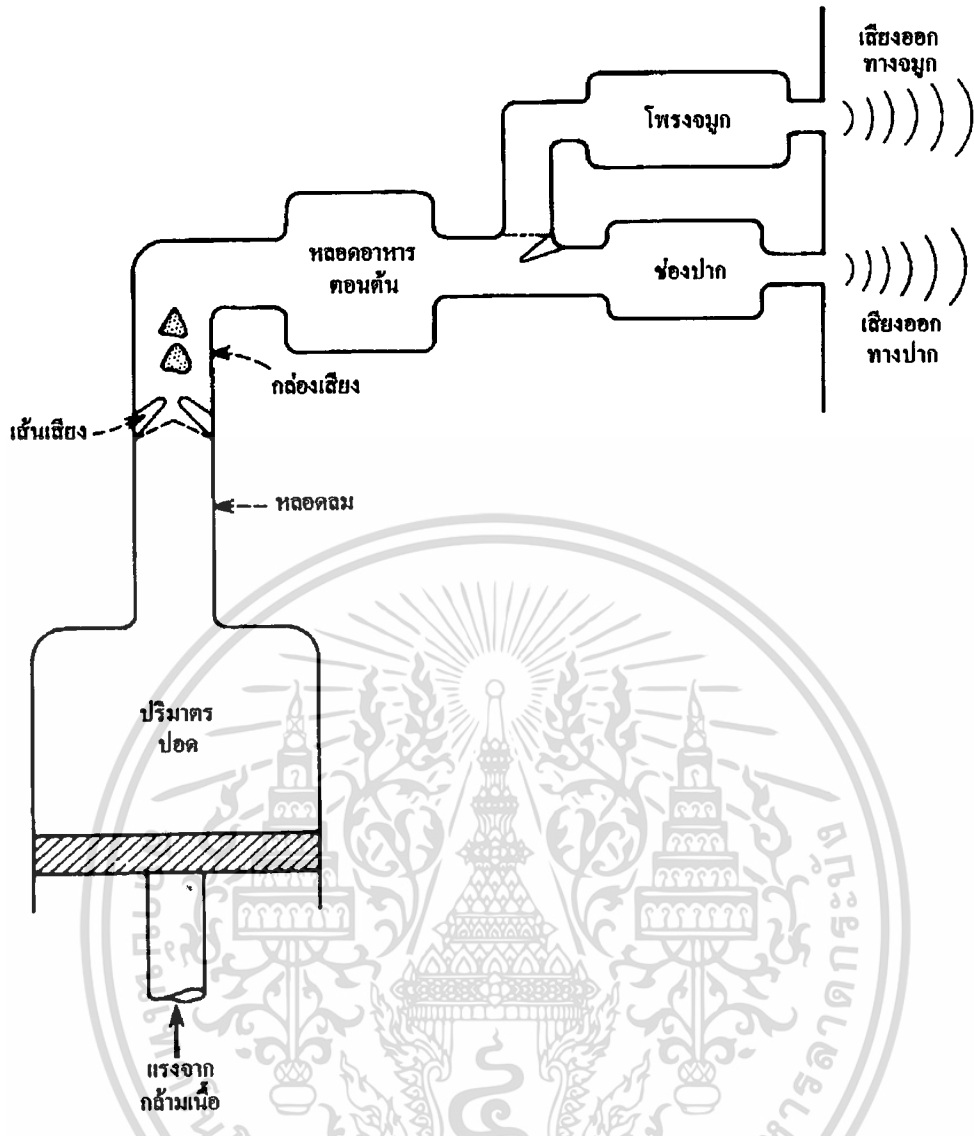
2.1.4 ช่องปากและส่วนของหลอดอาหารตอนต้น อวัยวะกลุ่มนี้อยู่ต่อจากกล่องเสียงอาจเรียกว่าอวัยวะกำทอนเสียง (vocal tract) ทำหน้าที่กำทอนเสียง โดยให้กำทอนทั้งเสียงที่เกิดจากกล่องเสียงและเสียงที่เกิดภายในช่องปาก ขนาดของอวัยวะกำทอนเสียงขึ้นอยู่กับตำแหน่งของลิ้น ริมฝีปาก ขากรรไกร และเพดานอ่อน และเปลี่ยนแปลงไปตามการออกเสียง

2.1.5 โพรงจมูก เริ่มจากเพดานอ่อนจนถึงรูจมูกทั้งสอง ทำหน้าที่กำทอนเสียงร่วมกับช่องปาก เมื่อมีการเปล่งเสียงที่ออกทางจมูก (nasal sounds) เช่นเสียง /ม/, /น/ และ /ง/ เป็นต้น

2.2 กระบวนการผลิตเสียงพูด

จากระบบเสียงพูดของมนุษย์ สามารถแสดงเป็นแผนภาพของระบบกำเนิดเสียงดังรูปที่ 2.3 ซึ่งเราสามารถจำแนกกลไกการสร้างเสียงพูดของมนุษย์ได้ 3 แบบ [13] ดังนี้

2.2.1 อากาศไหลจากปอดจะถูกมอดูเลต (modulate) โดยการสั่นของเส้นเสียงทำให้เกิดคลื่นเสียงลักษณะคล้ายพัลส์ที่มีคาบเวลาแบบควอไซ (quasi-periodic pulse-like excitation)



รูปที่ 2.3 แผนภาพระบบเสียงพูดของมนุษย์

2.2.2 อากาศที่ไหลจากปอดถูกทำให้ปั่นป่วน ด้วยการบังคับให้ไหลผ่านช่องแคบอันเกิดจากการบีบตัวของอวัยวะในช่องปาก ทำให้เกิดเสียงลักษณะคล้ายเสียงรบกวน (noise-like excitation)

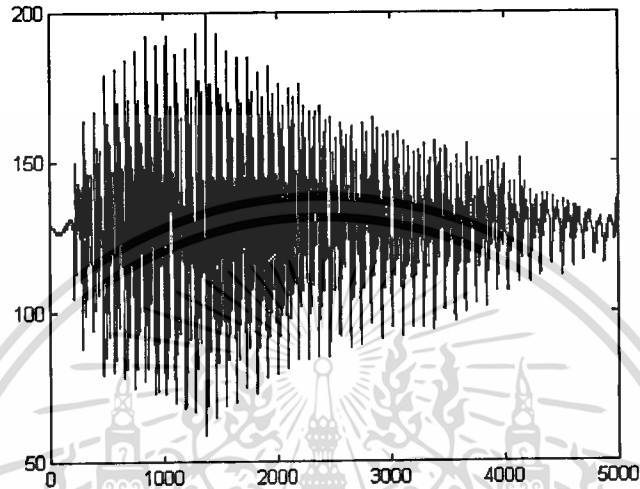
2.2.3 อากาศที่ไหลถูกกัก และเกิดแรงดันอยู่ภายในส่วนของช่องปากที่ปิด จากนั้นจึงปล่อยให้ อากาศที่มีแรงดัน พุ่งออกไปอย่างรวดเร็ว ทำให้เกิดการกระตุ้นเป็นเสียงในช่วงเริ่มต้น (transient excitation)

2.3 เสียงพูดของมนุษย์

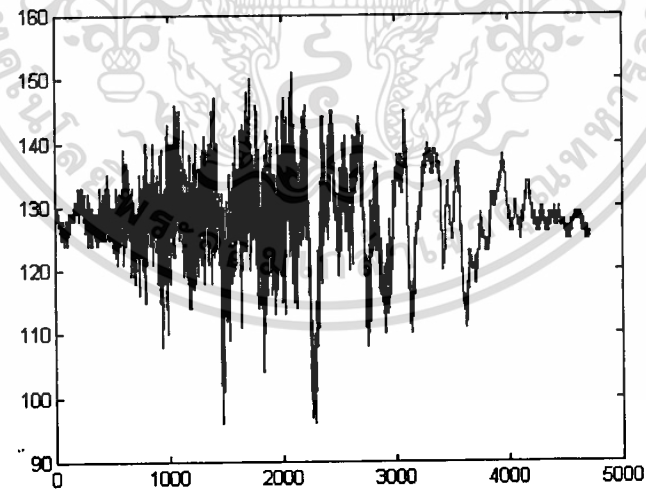
เสียงพูด เป็นคลื่นตามยาว (longitudinal wave) เกิดจากการสั่นของอนุภาคตัวกลาง นั่นคือ อากาศ และทิศทางการสั่นของอนุภาคจะอยู่ในทิศเดียวกันกับทิศทางของการเคลื่อนที่ คลื่นเสียงเป็น คลื่นที่เปลี่ยนแปลงไปตามเวลา

เสียงพูดแบ่งได้เป็น 2 ชนิดตามการกำเนิดเสียงหรือโหมด (mode) การกระตุ้น คือ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.1 เสียงวอยซ์หรือเสียงโฆษะ (voiced sounds) เกิดจากการบังคับอากาศให้ผ่านช่องสายเสียง ทำให้มีการเปลี่ยนแปลงความตึงหย่อนของเส้นเสียง โดยเส้นเสียงจะสั่นและเกิดเป็นพัลส์ (pulse) ของอากาศไปกระตุ้นอวัยวะกำทอนเกิดเป็นเสียงวอยซ์ ตัวอย่างเสียงวอยซ์ได้แก่ เสียงสระ เสียงพยัญชนะที่ต้องออกเสียงจากลำคอ (voiced consonants) เช่นเสียง /a/ ดังรูปที่ 2.4(ก)



รูปที่ 2.4(ก) รูปคลื่นของสัญญาณเสียง /a/ ซึ่งเป็นเสียงวอยซ์



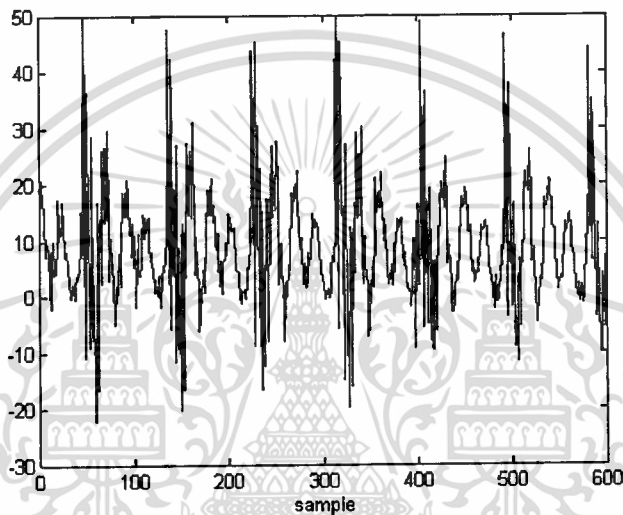
รูปที่ 2.4(ข) รูปคลื่นของสัญญาณเสียง /sh/ ซึ่งเป็นเสียงอันวอยซ์

2.3.2 เสียงอันวอยซ์หรือเสียงอโฆษะ (voiceless หรือ unvoiced sounds) เป็นเสียงที่เกิดในช่องปากหรือโพรงจมูก โดยอวัยวะภายในช่องปาก ริมฝีปาก ขวางการไหลของอากาศให้ผ่านได้เป็นช่องเล็กๆ อากาศจึงไหลผ่านอย่างรวดเร็วและปั่นป่วนจนกระทั่งสร้างเป็นเสียงรบกวนช่วงความถี่กว้าง เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาติให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(broad-spectrum noise) ตัวอย่างเสียงอันวอยซ์ได้แก่ เสียงพยัญชนะที่ไม่ได้เกิดจากลำคอ (voiceless consonants) เช่นเสียง /sh/ ซึ่งแสดงดังรูปที่ 2.4(ข)

2.4 คาบเวลาพิทช์ (pitch period)

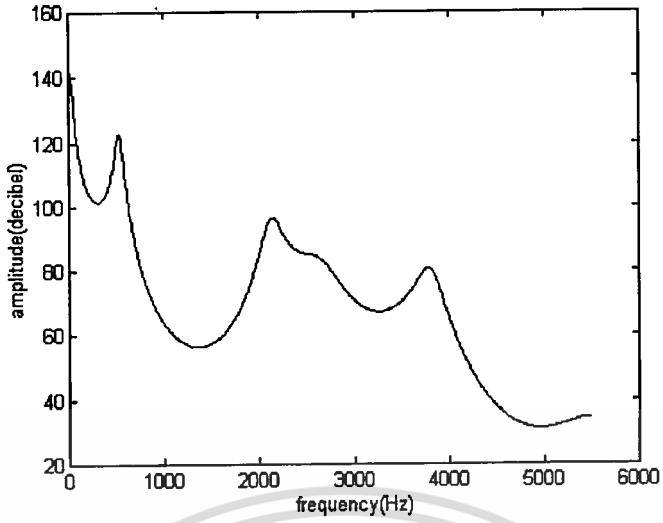
คาบเวลาพิทช์ คือคาบเวลาขบวนพัลส์ของอากาศที่เส้นเสียง ซึ่งเป็นอินพุต (input) ของการเปล่งเสียงวอยซ์ ส่วนกลับของคาบเวลาพิทช์เรียกว่าความถี่พิทช์ (pitch frequency) คาบเวลาพิทช์ของเสียง /a/ แสดงดังรูปที่ 2.5



รูปที่ 2.5 รูปคลื่นและคาบเวลาพิทช์ของเสียง /a/

2.5 ความถี่ฟอร์แมนท์ (formant frequency)

ความถี่ฟอร์แมนท์ คือความถี่กำทอน (resonance frequency) ของกลุ่มอวัยวะกำทอนเสียง ความถี่ฟอร์แมนท์ของเสียงจะมีค่าเท่าใดขึ้นอยู่กับ การบังคับขนาดและรูปทรงของอวัยวะส่วนนี้ อันเกิดเป็นเสียงความถี่ที่แตกต่างกันนั้นคือเสียงพูดต่างๆ นั่นเอง ความถี่ฟอร์แมนท์สำหรับเสียงพูดคำหนึ่งๆ อาจมีได้หลายค่า ดังรูปที่ 2.6 เป็นสเปกตรัมของเสียง /a/ ซึ่งมีความถี่ฟอร์แมนท์ที่ 1, 2, 3 และ 4 อยู่ ณ ตำแหน่งความถี่ ที่ 550, 2150, 2600 และ 3750 เฮิรตซ์ตามลำดับ



รูปที่ 2.6 สเปกตรัมของเสียง /a/

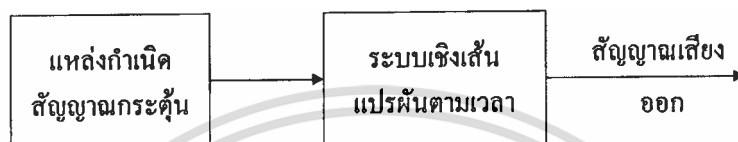


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



แบบจำลองระบบกำเนิดเสียงพูด

ในบทที่ผ่านมาเราได้ทราบการทำงานของระบบเสียงพูดของมนุษย์ จากแผนภาพรูปที่ 2.2 เราสามารถแสดงบล็อกไดอะแกรมจำลองระบบกำเนิดเสียงเบื้องต้นดังรูปที่ 3.1 จากรูปมีการแยกภาคแหล่งกำเนิดสัญญาณกระตุ้นออกจากส่วนกำเนิดเสียงซึ่งแทนด้วยระบบเชิงเส้นแปรผันตามเวลา



รูปที่ 3.1 บล็อกไดอะแกรมจำลองระบบกำเนิดเสียงเบื้องต้น

จากรูปที่ 3.1 แหล่งกำเนิดสัญญาณกระตุ้นทำหน้าที่แทนการทำงานของปอดและกล่องเสียง ส่วนนี้จะผลิตขบวนการพัลส์ที่มีความยาวเวลาที่เปลี่ยนแปลงเสียงวอยซ์ และให้กำเนิดเสียงซึ่งคล้ายเสียงรบกวนขณะเปล่งเสียงอันวอยซ์

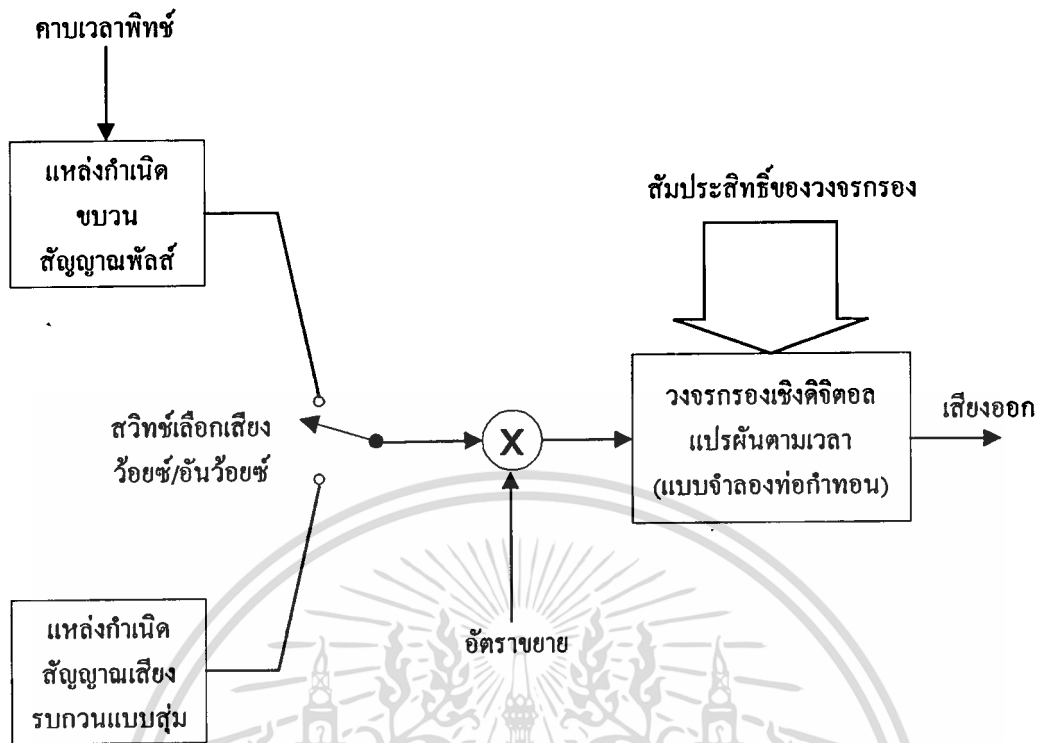
ส่วนที่สองเป็นท่อนำทอนเสียง จะแทนการทำงานของช่องปากและโพรงจมูก ทำหน้าที่เสมือนตัวกรองสัญญาณ (filter) ที่ยอมให้ความถี่ฟอร์แมนต์ผ่านได้ ซึ่งสามารถแทนด้วยระบบเชิงเส้นแปรผันตามเวลา (time-varying linear system)

8.1 แบบจำลองระบบกำเนิดสัญญาณเสียงพูดเชิงเลข

จากแบบจำลองรูปที่ 3.1 ประกอบกับการใช้เทคนิคการประมวลผลสัญญาณเชิงเลข (digital signal processing) สามารถแสดงเป็นแบบจำลองระบบกำเนิดเสียงพูดเชิงเลขดังรูปที่ 3.2

3.1.1 แหล่งกำเนิดสัญญาณกระตุ้น แบ่งเป็น 2 ชนิดตามประเภทเสียง โดยชนิดแรกเป็นแหล่งกำเนิดเสียงวอยซ์ ทำหน้าที่ผลิตขบวนการสัญญาณพัลส์ด้วยคาบเวลาคงที่ และชนิดที่สองเป็นแหล่งกำเนิดสัญญาณเสียงรบกวนแบบสุ่ม ในกรณีที่เกิดเสียงอันวอยซ์

3.1.2 แบบจำลองอวัยวะนำทอนเสียง เป็นวงจรกรองเชิงดิจิตอลซึ่งเป็นชนิดที่มีแต่โพลทั้งหมด (all-pole digital filter) หรือ วงจรกรองเชิงเลขแบบป้อนกลับ (recursive digital filter) สำหรับการวิเคราะห์แบบจำลองท่อนำทอนจะกล่าวในเรื่องการวิเคราะห์ตัวแปรแบบเสียง



รูปที่ 3.2 แบบจำลองระบบสร้างสัญญาณเสียงพูด

3.2 พารามิเตอร์ (parameter) ของระบบกำเนิดเสียงพูด

การสร้างระบบเสียงพูดจำเป็นต้องวิเคราะห์หาค่าตัวแปรหรือพารามิเตอร์ที่ต้องใช้ตามแบบจำลองและระบบประมวลสัญญาณเสียง ซึ่งประกอบด้วย

- 3.2.1 ดัชนีหรือพารามิเตอร์บอกชนิดเสียงพูดว่าเป็นเสียงว้อยซ์หรืออันว้อยซ์
- 3.2.2 คาบเวลาพิทซ์ของขบวนการพัลส์ในกรณีเสียงว้อยซ์
- 3.2.3 อัตราการขยาย (gain) ของวงจรรองหรือค่ารากกำลังสองเฉลี่ย (rms value) ของเสียง
- 3.2.4 พารามิเตอร์ของแบบจำลองทอค่าทอนหรือสัมประสิทธิ์ของวงจรรอง
- 3.2.5 อัตราการสุ่มตัวอย่างสัญญาณเสียงหรืออัตราเร็วข้อมูล

บทที่ 4

หลักการประมาณเชิงเส้น

เนื่องจากสัญญาณเสียงมีพารามิเตอร์ที่สำคัญต่างๆ ดังที่กล่าวไว้ในบทที่ 2 และ 3 การวิเคราะห์การเข้ารหัสโดยการประมาณเชิงเส้น (linear predictive coding : LPC) จึงถูกใช้ในการหาค่าพารามิเตอร์พื้นฐานต่างๆ ที่กล่าวไว้ในหัวข้อ 3.2 เพื่อใช้ในการเข้ารหัสสำหรับส่งสัญญาณเสียงด้วยความเร็วบีทการย่อข้อมูล และการจดจำเสียง ด้วยเหตุที่ว่าวิธีนี้เป็นการวิเคราะห์พารามิเตอร์ต่างๆ ที่ให้ความแม่นยำสูง และสามารถย่อข้อมูลได้อย่างมีประสิทธิภาพ

หลักการพื้นฐานของการประมาณเชิงเส้นคือ การประมาณค่าสัญญาณจากผลรวมเชิงเส้นของสัญญาณก่อนหน้า

สมมติว่าสัญญาณเดิมเป็น $s(n)$ การประมาณค่าสัญญาณเป็น $s'(n)$ ดังนั้นเราสามารถอธิบายการประมาณเชิงเส้นได้ด้วยสมการต่อไปนี้

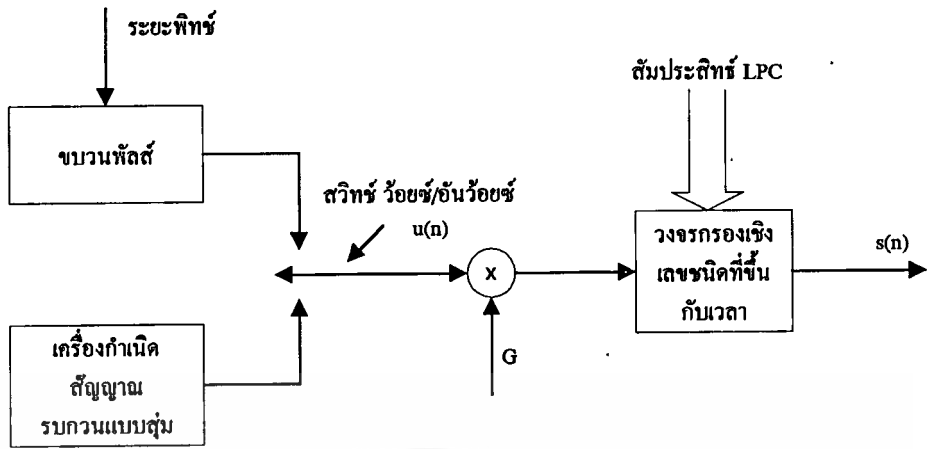
$$s'(n) = \sum_{k=1}^p \alpha_k s(n-k) \quad (4.1)$$

เมื่อ α_k เป็นค่าคงที่ เรียกวิธีการนี้ว่าการประมาณเชิงเส้นอันดับ p โดยมีเงื่อนไขว่า ค่า α_k ที่ใช้ในการประมาณจะต้องทำให้ผลรวมของกำลังสองของความคลาดเคลื่อน $\{s(n) - s'(n)\}^2$ มีค่าน้อยที่สุด นั่นคือ $\sum e^2(n) = \sum \{s(n) - s'(n)\}^2$ มีค่าต่ำที่สุด ซึ่งการประมาณเชิงเส้นมีอยู่หลายวิธี ได้แก่ [13]

- 4.1 วิธีโควาเรียนซ์ (Covariance Method)
- 4.2 วิธีออโตคอร์รีเลชัน (Autocorrelation Method)
- 4.3 วิธีแลตทิซ (Lattice Method)
- 4.4 วิธีวงจรรองผกผัน (Inverse Filter Method)
- 4.5 วิธีการประมาณสเปกตรัม (Spectrum Estimation Method)

และอื่นๆ อีกหลายวิธี แต่ในที่นี้จะเลือกใช้วิธีออโตคอร์รีเลชัน หรือ วิธีอັดคัสัมพันธ์

จากหลักการพื้นฐานของการประมาณเชิงเส้น และรูปที่ 3.2 ซึ่งเป็นแบบจำลองระบบสร้างสัญญาณเสียง เราสามารถเขียน บล็อกไดอะแกรมการทำการประมาณเชิงเส้น มาสร้างสัญญาณเสียงพูดได้ดังรูปที่ 4.1



รูปที่ 4.1 บล็อกโคโตะแกรมแสดงโมเดลการสร้างสัญญาณเสียงพูดอย่างง่าย

จากรูปที่ 4.1 สามารถเขียนสมการได้เป็น

$$s(n) = G \cdot u(n) + \sum_{k=1}^p a_k s(n-k) \quad (4.2)$$

การประมาณเชิงเส้นโดยใช้สัมประสิทธิ์ $\{\alpha_k\}$ คือ

$$s'(n) = \sum_{k=1}^p \alpha_k s(n-k) \quad (4.3)$$

ดังนั้น ความคลาดเคลื่อน คือ

$$\begin{aligned} e(n) &= s(n) - s'(n) \\ &= s(n) - \sum_{k=1}^p \alpha_k s(n-k) \end{aligned} \quad (4.4)$$

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

$$\begin{aligned} A(z) &= \frac{E(z)}{S(z)} \\ &= 1 - \sum_{k=1}^p \alpha_k z^{-k} \end{aligned} \quad (4.5)$$

จากสมการ (4.2)-(4.4) จะเห็นได้ว่า ถ้า $\{\alpha_k\} = \{a_k\}$ แล้ว

$$e(n) = G \cdot u(n) \quad (4.6)$$

ในการประมวลสัญญาณจะมีการแบ่งสัญญาณออกเป็นช่วงสั้นๆ (windowing) ช่วงละประมาณ 20 - 40 มิลลิวินาที ซึ่งจะได้มีการอธิบายต่อไป ดังนั้น ค่าผสมรวมของกำลังสองของความคลาดเคลื่อน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{aligned}
E_n &= \sum_m e_n^2(m) \\
&= \sum_m [s_n(m) - s'_n(m)]^2 \\
&= \sum_m \left[s_n(m) - \sum_{k=1}^p \alpha_k s_n(m-k) \right]^2
\end{aligned} \tag{4.7}$$

โดยที่ n คือ ช่วงที่ n ของสัญญาณที่ใช้คำนวณ เพราะฉะนั้นเพื่อให้ได้ค่า E_n ต่ำที่สุดจะต้องมีเงื่อนไขว่า

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

จากสมการ (4.7)

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

$\partial E_n / \partial \alpha_i = 0$ ก็ต่อเมื่อ

$$\sum_{k=1}^p \alpha_k \sum_m s_n(m-k)s_n(m-i) = \sum_m s_n(m)s_n(m-i) \quad \text{เมื่อ } i = 1, 2, \dots, p \tag{4.8}$$

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

$$\sum_{k=1}^p \alpha_k \phi_n(i,j) = \phi_n(i,0) \tag{4.9}$$

โดยสมการ (4.7)-(4.8) จะได้ว่า

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

และจาก $\phi_n(i,j) = \sum_m s_n(m-k)s_n(m-i)$

$$E_n = \phi_n(0,0) - \sum_{k=1}^p \alpha_k \phi_n(0,k) \tag{4.10}$$

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

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

$$\text{ให้ } R_n(k) = \sum_{m=0}^{N-1-k} s_n(m)s_n(m+k) \tag{4.11}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\therefore \phi_n(i, j) = R_n(|i - k|) \quad \text{เมื่อ } i = 1, 2, \dots, p; \quad k = 0, 1, 2, \dots, p$$

จากสมการ (4.9) จะได้ว่า

$$\sum_{k=1}^p \alpha_k R_n(|i - k|) = R_n(i) \quad \text{เมื่อ } i = 1, 2, \dots, p \quad (4.12)$$

และจากสมการ (4.10)

$$E_n = R_n(0) - \sum_{k=1}^p \alpha_k R_n(k) \quad (4.13)$$

จากสมการ (4.12) เขียนให้อยู่ในรูปเมตริกซ์ได้เป็น

$$\begin{bmatrix} R_n(0) & R_n(1) & \cdots & R_n(p-1) \\ R_n(1) & R_n(0) & \cdots & R_n(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_n(p-1) & R_n(p-2) & \cdots & R_n(0) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix} = \begin{bmatrix} R_n(1) \\ R_n(2) \\ \vdots \\ R_n(p) \end{bmatrix} \quad (4.14)$$

หรือ

$$R_n \cdot \alpha = r_n \quad (4.15)$$

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

จากสมการ (4.2) และ (4.4)

$$G \cdot u(n) = s(n) - \sum_{k=1}^p a_k s(n-k)$$

$$e(n) = s(n) - \sum_{k=1}^p \alpha_k s(n-k)$$

เมื่อ $a_k = \alpha_k$ จะได้ว่า

$$e(n) = G \cdot u(n)$$

$$\begin{aligned} \therefore E_n &= \sum_{m=0}^{N-1} e^2(m) = G \sum_{m=0}^{N-1} u^2(m) \\ &= R_n(0) - \sum_{k=1}^p \alpha_k R_n(k) \end{aligned} \quad (4.16)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากสมการ (4.15) เราสามารถหา α โดยตรงดังนี้

$$\alpha = R_n^{-1} \cdot r_n \quad (4.17)$$

และจากสมการ (4.16) เราสามารถหาค่า G โดยตรงจาก

$$G = \frac{R_n(0) - \sum_{k=1}^p \alpha_k R_n(k)}{\sum_{m=0}^{N-1} u^2(m)} \quad (4.18)$$

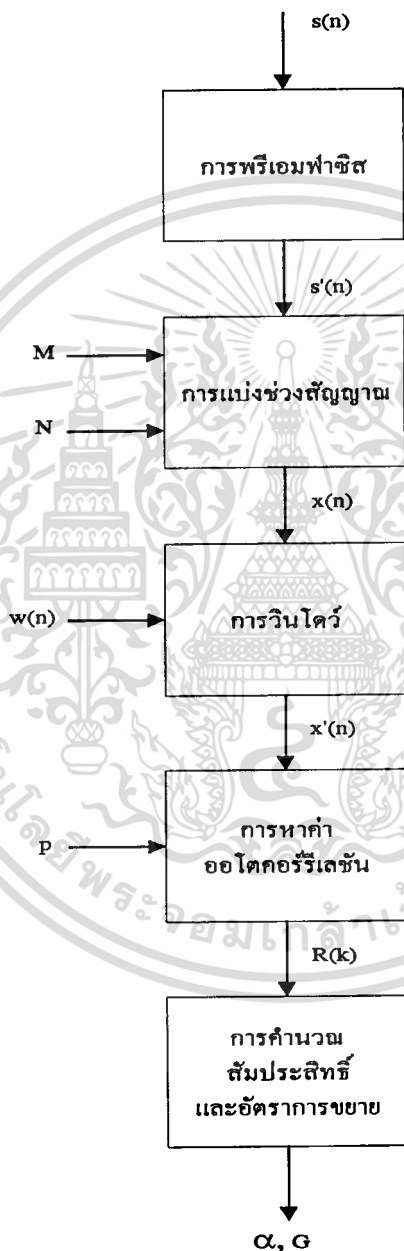
การคำนวณหาค่าพารามิเตอร์ต่างๆ อันได้แก่ ค่าสัมประสิทธิ์ของ LPC ค่าอัตราขยาย และคาบเวลาพิทช์ จะกล่าวในบทที่ 5 และบทที่ 6



บทที่ 5

การคำนวณค่าสัมประสิทธิ์ของ LPC และอัตราการขยาย

ในการหาค่าสัมประสิทธิ์ LPC (α) และอัตราการขยาย (G) เราจำเป็นต้องมีการเตรียมข้อมูลในการวิเคราะห์ก่อน ซึ่งขั้นตอนในการเตรียมสัญญาณเข้าเป็นดังรูป 5.1



รูปที่ 5.1 แสดงขั้นตอนการเตรียมสัญญาณในการวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การวิเคราะห์สามารถแบ่งเป็นขั้นตอนต่างๆ ดังนี้ [15]

- 5.1 การพรีเอมฟาซิส (preemphasis)
- 5.2 การแบ่งช่วงสัญญาณ (frame blocking)
- 5.3 การวินโดว์ (windowing)
- 5.4 การหาอโตคอรีเลชัน (autocorrelation analysis)
- 5.5 การหาสัมประสิทธิ์ α และค่าอัตราขยาย G

5.1 การพรีเอมฟาซิส

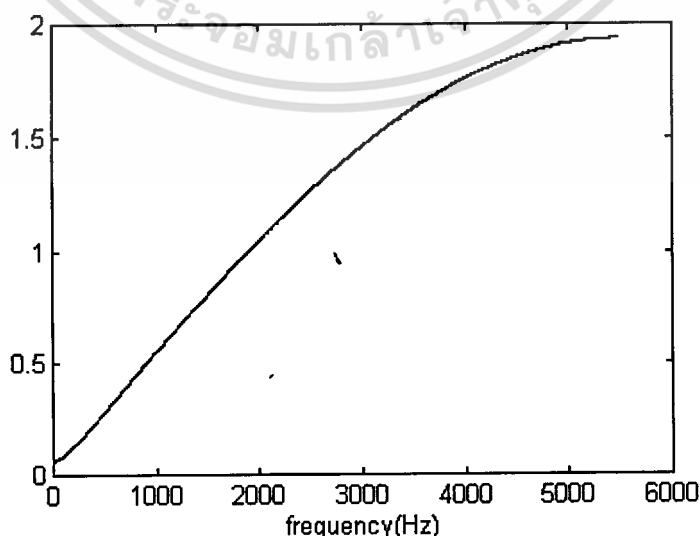
เนื่องจากสัญญาณเสียงพูดของมนุษย์ จะมีองค์ประกอบส่วนใหญ่อยู่บริเวณความถี่ต่ำ เมื่อเทียบกับแถบความถี่ปฏิบัติงาน (bandwidth) ไม่เกิน 5 กิโลเฮิร์ตซ์ ดังนั้น เพื่อให้อัตราส่วนสัญญาณเสียงต่อสัญญาณรบกวน (signal to noise ratio : SNR) มีค่าค่อนข้างคงที่ตลอดช่วงความถี่ปฏิบัติงานนี้ เราจึงต้องมีการพรีเอมฟาซิส โดยเน้นความถี่สูงให้มีขนาดสูงขึ้น นั่นคือ การพรีเอมฟาซิสก็คือการกรองสัญญาณด้วยวงจรกรองความถี่สูงผ่าน (high pass filter) ซึ่งมักนิยมใช้วงจรกรองอันดับหนึ่ง มีฟังก์ชันถ่ายโอนเป็น

$$H(z) = 1 - a \cdot z^{-1} \quad \text{โดยที่ } 0.9 < a < 1.0$$

เมื่อเทียบกับรูป 5.1 เราจะได้ว่า

$$s'(n) = s(n) - a \cdot s(n-1)$$

ยิ่งค่า a ใกล้ 1 เท่าใด ความถี่สูงก็จะยิ่งถูกขยายมากขึ้นเท่านั้น ค่า a ที่นิยมสำหรับใช้ในการหาพารามิเตอร์ของ LPC คือ $15/16 = 0.9375$ [14] เมื่อนำฟังก์ชันถ่ายโอนมาพล็อตกราฟของขนาดเทียบกับความถี่จะได้ดังรูปที่ 5.2

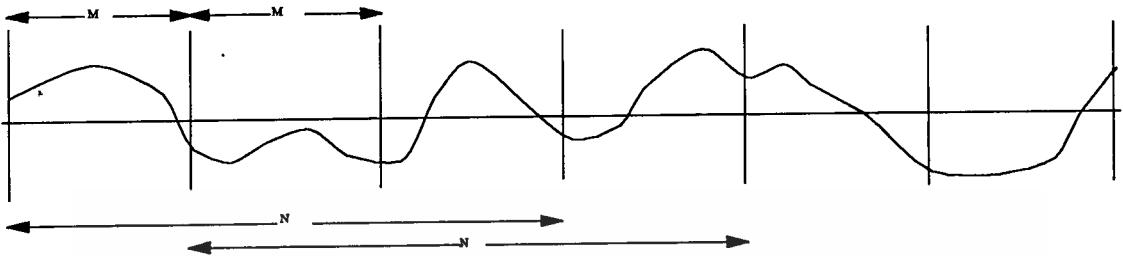


รูปที่ 5.2 แสดงขนาดสเปกตรัมของฟังก์ชันถ่ายโอนของการพรีเอมฟาซิส

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 การแบ่งช่วงสัญญาณ

สัญญาณที่ผ่านการพรีเอมฟาซิสแล้ว ($s'(n)$) จะถูกตัดแบ่งออกเป็นช่วงๆ หรือ เฟรม ช่วงละ N ตัวอย่างสัญญาณ การวิเคราะห์จะวิเคราะห์ทีละช่วงของแต่ละ N ตัวอย่างสัญญาณ ดังรูปที่ 5.3



รูปที่ 5.3 แสดงการแบ่งช่วงของสัญญาณที่ใช้ในการวิเคราะห์

โดยช่วงในการวิเคราะห์แต่ละช่วงจะถูกเลื่อนไปเป็นระยะ M ช่วงสัญญาณ จะเห็นได้ว่า ถ้าค่า M โตกว่าค่า N ในการเลื่อนของช่วงในการวิเคราะห์จะทำให้บางสัญญาณ ไม่ถูกใช้ในการวิเคราะห์ ก็จะเป็นการสูญเสียส่วนหนึ่งทำให้ผลที่ได้ไม่ถูกต้องเท่าที่ควร ถ้าค่า M เล็กกว่า N ก็จะทำให้ตัวอย่างสัญญาณ ทุกตัวถูกนำมาวิเคราะห์ ยิ่งค่า M เล็กเท่าใด ความแม่นยำในการวิเคราะห์ก็จะยิ่งสูงขึ้นเท่านั้น แต่ก็จะทำให้การคำนวณช้าลง

หลักการกำหนดขนาดของวินโดว์

5.2.1 วินโดว์จะต้องสั้นพอที่ คุณสมบัติของเสียงที่เราสนใจจะวิเคราะห์นั้น ยังไม่เปลี่ยนแปลงในวินโดว์

5.2.2 วินโดว์จะต้องยาวพอที่ จำนวนตัวอย่างสัญญาณในวินโดว์ สามารถนำมาคำนวณหาคุณสมบัติที่ต้องการได้

5.2.3 วินโดว์ที่ติดกัน ไม่ควรจะสั้นจนกระโดดข้ามข้อมูลบางส่วนไป แต่ควรเลื่อนวินโดว์ให้ น้อยกว่าขนาดของเฟรม

เนื่องจาก เราใช้ความถี่ในการสุ่มสัญญาณ 11.025 กิโลเฮิร์ตซ์ ในการวิเคราะห์นี้ เราเลือกใช้ค่า $N = 300$ และ ค่า $M = 100$ นั่นคือ ช่วงในการวิเคราะห์คือ 27.21 มิลลิวินาที และระยะในการเลื่อนเฟรมประมาณ 9 มิลลิวินาที

5.3 การวินโดว์

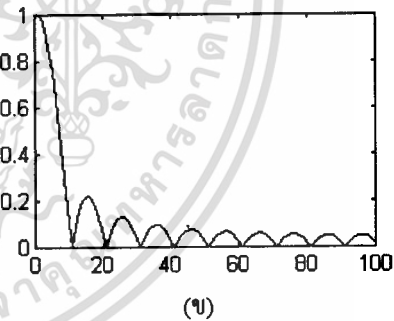
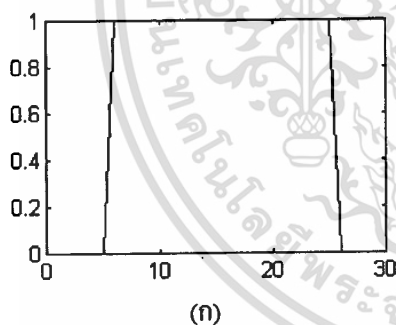
พิจารณาช่วงสัญญาณ N ตัวอย่างสัญญาณของช่วงใดๆ ที่ตัดมาวิเคราะห์ (รูปที่ 5.4) จะเห็นว่าที่ขอบของเฟรมที่ตัดมานี้มีความไม่ต่อเนื่องของสัญญาณ ถ้ามองในโดเมนความถี่ ก็จะมีความถี่สูงเกิดขึ้น ดังนั้นเพื่อที่จะลดองค์ประกอบทางความถี่ที่สูงเหล่านี้ เราจะคูณด้วยฟังก์ชันวินโดว์เพื่อลดความไม่ต่อเนื่องของสัญญาณที่ขอบ และไม่ทำให้สเปกตรัมของสัญญาณในช่วงความถี่ต่ำเปลี่ยนแปลงไปมากนัก [1], [6] ในที่นี่จะใช้ฟังก์ชันวินโดว์แฮมมิง (Hamming window function) [6] ซึ่งนิยามโดยสมการดังนี้คือ

$$w(n) = 0.54 - 0.46\cos(2\pi n/(N-1)) \quad \text{เมื่อ } n = 0, 1, 2, \dots, N-1 \quad (5.1)$$

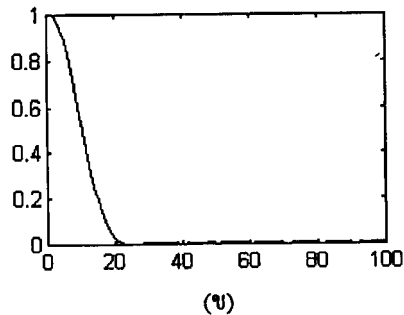
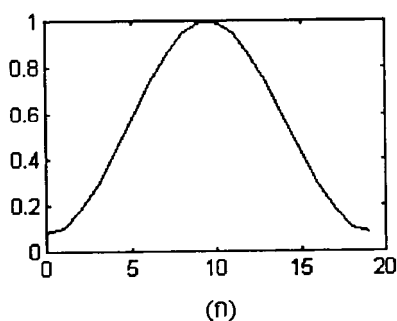
รูปที่ 5.5 และ 5.6 แสดงองค์ประกอบทางเวลาและทางความถี่ของฟังก์ชันวินโดว์แบบสี่เหลี่ยม และแฮมมิง ตามลำดับ จะเห็นว่าสเปกตรัมของวินโดว์แฮมมิงมีริพเพิล (ripple) น้อยกว่าของวินโดว์สี่เหลี่ยม



รูปที่ 5.4 แสดงส่วนของสัญญาณที่ตัดมาวิเคราะห์



รูปที่ 5.5 แสดงวินโดว์แบบสี่เหลี่ยม (ก) ในโดเมนเวลา (ข) ในโดเมนความถี่



รูปที่ 5.6 แสดงวินโดว์แบบแฮมมิง (ก) ในโดเมนเวลา (ข) ในโดเมนความถี่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้น จากรูปที่ 5.1 เราจะได้ว่า

$$x'(n) = x(n) \cdot w(n) \quad (5.2)$$

5.4 การคำนวณออกโศคอร์รีเลชัน

จากสมการ (4.11)

$$R_n(k) = \sum_{m=0}^{N-1-k} s_n(m)s_n(m+k) \quad \text{เมื่อ } k = 0, 1, 2, \dots, p \quad (5.3)$$

ดังนั้น จากสมการ (5.2)-(5.3) จะได้ว่า

$$R_n(k) = \sum_{m=0}^{N-1-k} x'(m)x'(m+k) \quad (5.4)$$

5.5 การหาสัมประสิทธิ์ α และอัตราขยาย G

เมื่อได้ค่า $R_n(0), R_n(1), R_n(2), \dots, R_n(p)$ แล้ว ก็สามารถหาค่า α และค่า G ได้จากสมการ (4.17) และ (4.18) ตามลำดับ นั่นคือ

$$\alpha = R_n^{-1} \cdot r_n \quad (5.5)$$

$$G = \frac{R_n(0) - \sum_{k=1}^p \alpha_k R_n(k)}{\sum_{m=0}^{N-1} u^2(m)} \quad (5.6)$$

บทที่ 6

เกณฑ์การตัดสินเสียงว้อยซ์และเสียงอันว้อยซ์ และการหาคาบเวลาพิทซ์ของเสียงพูดด้วยวิธี SIFT

คาบเวลาพิทซ์ เป็นพารามิเตอร์แสดงความเป็นคาบของเสียงว้อยซ์ในช่วงเวลาที่พิจารณา อีกทั้งยังเป็นพารามิเตอร์ตัวหนึ่งของการบีบอัดสัญญาณเสียง และการสร้างขบวนการสัญญาณพัลส์เพื่อสังเคราะห์เสียงพูดตามบล็อกโคอะแกรมรูปที่ 4.1 โดยในบทนี้ได้เสนอวิธีหาคาบเวลาพิทซ์ของเสียงพูดด้วยวิธี SIFT (simple inverse filtering tracking) [13], [14]

การหาคาบเวลาพิทซ์ด้วยวิธี SIFT แสดงเป็นบล็อกโคอะแกรมดังรูปที่ 6.1 ซึ่งการทำงานในแต่ละขั้นตอนอธิบายได้ดังนี้

6.1 ลำดับสัญญาณเสียงพูดที่ถูกสุ่มสัญญาณด้วยความถี่ 11.025 กิโลเฮิร์ตซ์ จากการสังเกตค่าความถี่พิทซ์ของเสียงมนุษย์ซึ่งมักมีค่าต่ำกว่า 900 เฮิร์ตซ์ ฉะนั้นจึงนำสัญญาณเสียงผ่านเข้าวงจรกรองความถี่ต่ำที่มีค่าความถี่คัทออฟ 900 เฮิร์ตซ์ (บพแทรกที่ 3) และสัญญาณเสียง $s(n)$ ที่ออกจากวงจรกรองแสดงดังรูปที่ 6.2

6.2 ลดจำนวนตัวอย่างลำดับสัญญาณเป็นอัตราส่วน 6 ต่อ 1 ด้วยการเลือกเก็บค่าของลำดับสัญญาณมา 1 ตัวอย่างจากทุกๆ 6 ตัวอย่าง การลดจำนวนตัวอย่างของลำดับสัญญาณลงทำให้ลดเวลาในการคำนวณ เสียงที่ผ่านการลดจำนวนตัวอย่างแสดงดังรูปที่ 6.3

6.3 นำสัญญาณที่ได้มาผ่านการเตรียมสัญญาณเพื่อการวิเคราะห์ ด้วยการทำพรีเอมฟาซิส แบ่งช่องสัญญาณเป็นเฟรม ขนาด $300/11.025 = 27.21$ มิลลิวินาที (50 ตัวอย่าง) โดยเลื่อนเฟรมสัญญาณครั้งละ $100/11.025 = 9.07$ มิลลิวินาที (16 ตัวอย่าง) และผ่านวินโดว์ ซึ่งได้อธิบายไว้แล้วในบทที่ 5

6.4 นำสัญญาณที่ผ่านการเตรียมการวิเคราะห์แล้วนั้นมาคำนวณค่าสัมประสิทธิ์ การประมาณสัญญาณเสียงโดยใช้สัมประสิทธิ์จำนวน 4 ตัว ทีละเฟรม เพื่อใช้เป็นพารามิเตอร์ในการประมาณสัญญาณเสียงเชิงเส้นอันดับ 4 ซึ่งค่าสัมประสิทธิ์ทั้ง 4 ตัว หาจากสมการ (5.5)

6.5 หาสัญญาณคลาดเคลื่อนหรือสัญญาณผิดพลาด (error signal) $e(n)$ ระหว่างสัญญาณเสียงพูดจริง $s(n)$ กับสัญญาณจากการประมาณเชิงเส้นอันดับ 4 ซึ่งแทนด้วย $s'(n)$ และแสดงเป็นสมการจะได้

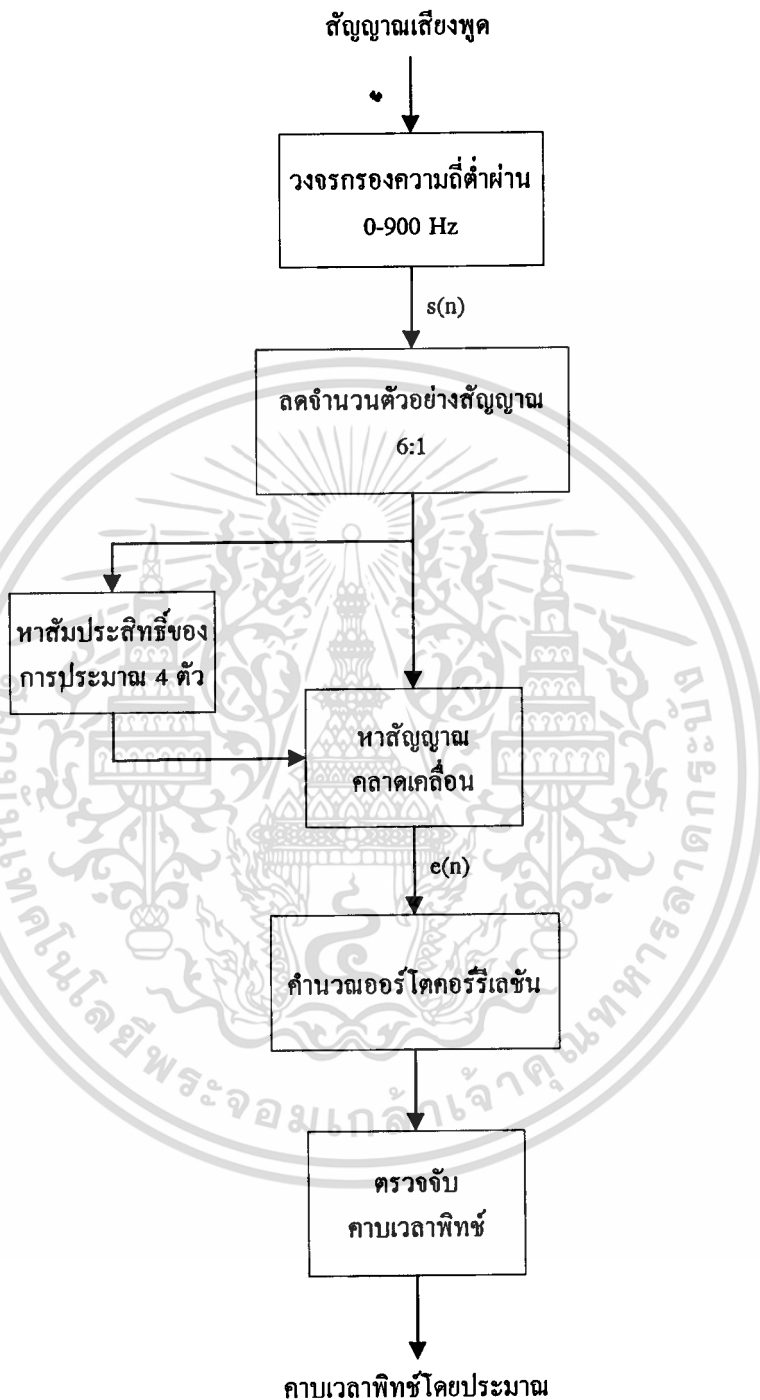
$$s'(n) = \sum_{k=1}^4 \alpha_k s(n-k) \quad (6.1)$$

จะได้สัญญาณคลาดเคลื่อนคือ

$$e(n) = s(n) - s'(n) \quad (6.2)$$

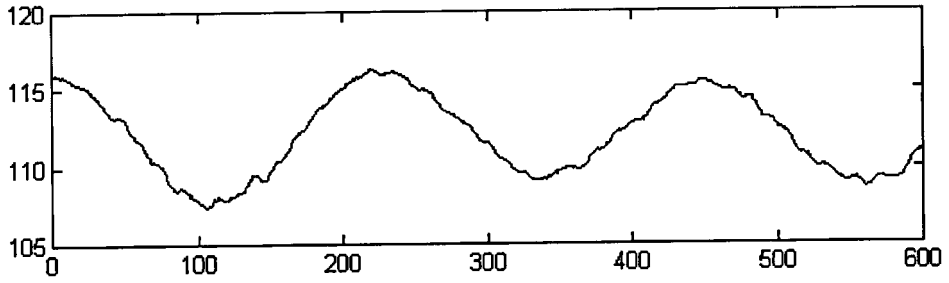
บล็อกโคอะแกรมการหาสัญญาณคลาดเคลื่อนแสดงดังรูปที่ 6.4 และรูปคลื่นของสัญญาณคลาดเคลื่อนแสดงดังรูปที่ 6.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

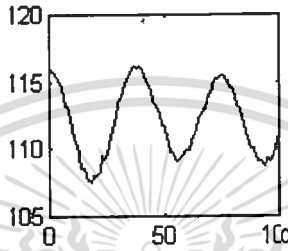


รูปที่ 6.1 บล็อกไดอะแกรมของวิธี SIFT ในการหาคาบเวลาพิทช์

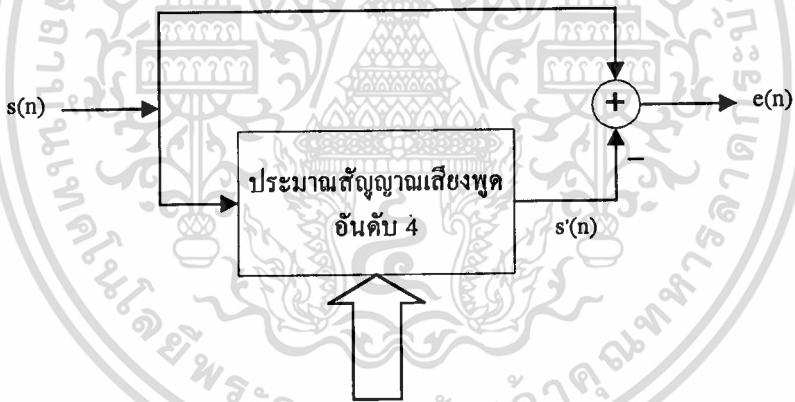
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.2 รูปคลื่นสัญญาณเสียงก่อนผ่านเข้าขั้นตอนการลดจำนวนตัวอย่าง

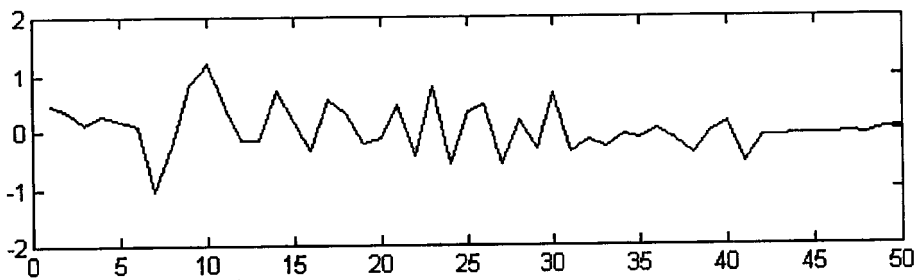


รูปที่ 6.3 รูปคลื่นสัญญาณเสียงที่ผ่านขั้นตอนการลดจำนวนตัวอย่างแล้ว



สัมประสิทธิ์การประมาณเชิงเส้น 4 ตัว

รูปที่ 6.4 บล็อกโคอะแกรมการหาสัญญาณคลาดเคลื่อน



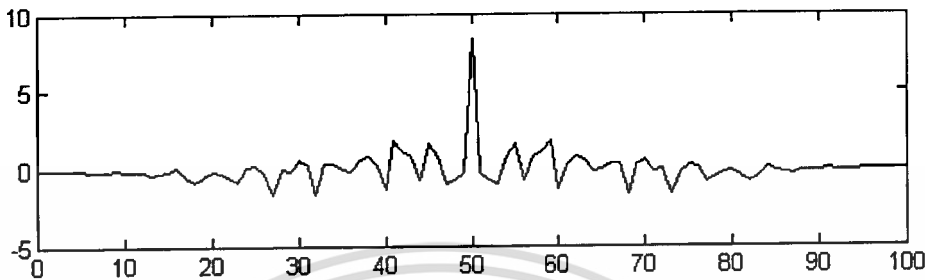
รูปที่ 6.5 รูปคลื่นสัญญาณคลาดเคลื่อน e(n)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.6 จำนวนหาอโตคอร์รีเลชันของสัญญาณคลาดเคลื่อนจากสมการ

$$R(k) = \sum e(n)+e(n+k) \quad \text{เมื่อ } k = 1, 2, 3, 4 \quad (6.3)$$

ภาพสัญญาณที่ผ่านการทำอโตคอร์รีเลชันแสดงดังรูปที่ 6.6



รูปที่ 6.6 ผลจากการทำอโตคอร์รีเลชันของสัญญาณคลาดเคลื่อน

6.7 คาบเวลาพิทช์ก็คือค่า 6 เท่าของระยะในแกนเวลาระหว่างตำแหน่งตรงกลาง กับตำแหน่งของค่าอโตคอร์รีเลชันสูงสุดค่าแรก

6.8 เกณฑ์ในการตัดสินว่าเป็นเสียงวอยซ์ หรือเสียงอันวอยซ์ จะดูที่อัตราส่วนระหว่างค่าอโตคอร์รีเลชันสูงสุดค่าแรก กับค่าอโตคอร์รีเลชันที่ตำแหน่งตรงกลาง ถ้าอัตราส่วนนี้ มากกว่า 0.378 [14] ให้ถือว่าเป็นเสียงวอยซ์ นอกเหนือจากนั้นแล้วให้ถือว่าเป็นเสียงอันวอยซ์

จากกระบวนการข้างต้นสามารถเขียนโพลิวาร์ทของการคำนวณหาคาบเวลาพิทช์ด้วยวิธี SIFT และโปรแกรมคำนวณหาคาบเวลาพิทช์ด้วยวิธี SIFT อยู่ในบทแทรกที่ 1

บทที่ 7

ผลการทดลอง

ในบทนี้ ได้รวมผลการทดลอง ในการหาค่าพารามิเตอร์ต่างๆ (สัมประสิทธิ์ของวงจรรองและอัตราขยาย) โดยวิธีออคเตอร์รีเลชัน ในการประมาณเชิงเส้นรวมถึงพล็อตสเปกตรัมของสัญญาณเสียงเปรียบเทียบกับสเปกตรัมของวงจรรองการประมาณเชิงเส้น

7.1 ผลในการวิเคราะห์พารามิเตอร์ของ LPC

จากการทดลองโดยการนำเสียงว้อยซ์และอันว้อยซ์ มาหาพารามิเตอร์ต่างๆ ของ LPC ได้ผลดังตารางที่ 7.1

ตารางที่ 7.1 แสดงค่าพารามิเตอร์ของเฟรมหนึ่งเฟรมของสัญญาณเสียง
เสียงที่ 1 แทนเสียงอันว้อยซ์
เสียงที่ 2 แทนเสียงว้อยซ์

พารามิเตอร์	เสียงที่ 1	เสียงที่ 2
G	24.656531184	54.323564026
a(1)	-1.247326144	0.860407461
a(2)	-1.041388796	-1.232998471
a(3)	-0.337743688	1.488384402
a(4)	-0.405478347	-1.342893828
a(5)	-0.201609377	1.572002474
a(6)	-0.367430317	-1.268723123
a(7)	-0.118282064	0.892472479
a(8)	-0.043897573	-0.933865803
a(9)	0.188637772	0.178204565
a(10)	-0.228553726	-0.341942969
a(11)	-0.320994684	0.034636535
a(12)	-0.212322522	0.171222710
pitch	-	84

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7.2 การสร้างสัญญาณเสียงกลับคืนจากสัมประสิทธิ์ที่ได้

ขั้นตอนต่อไป เราจะสร้างสัญญาณเสียงโดยอาศัย บล็อกไดอะแกรมที่ 4.1 และ ตารางที่ 7.1 ซึ่งผลการทดลองอยู่ในรูปที่ 7.1-7.4 ดังนี้

รูปที่ 7.1-7.2 แสดงการเปรียบเทียบสัญญาณเสียงก่อนและหลังการวิเคราะห์ LPC

รูปที่ 7.1 เสียงอันว้อยซ์

รูปที่ 7.2 เสียงว้อยซ์

รูปที่ 7.3-7.4 แสดงการเปรียบเทียบ สเปกตรัมของสัญญาณเสียงกับสเปกตรัมที่ได้จากสัมประสิทธิ์ LPC

รูปที่ 7.3 เสียงอันว้อยซ์

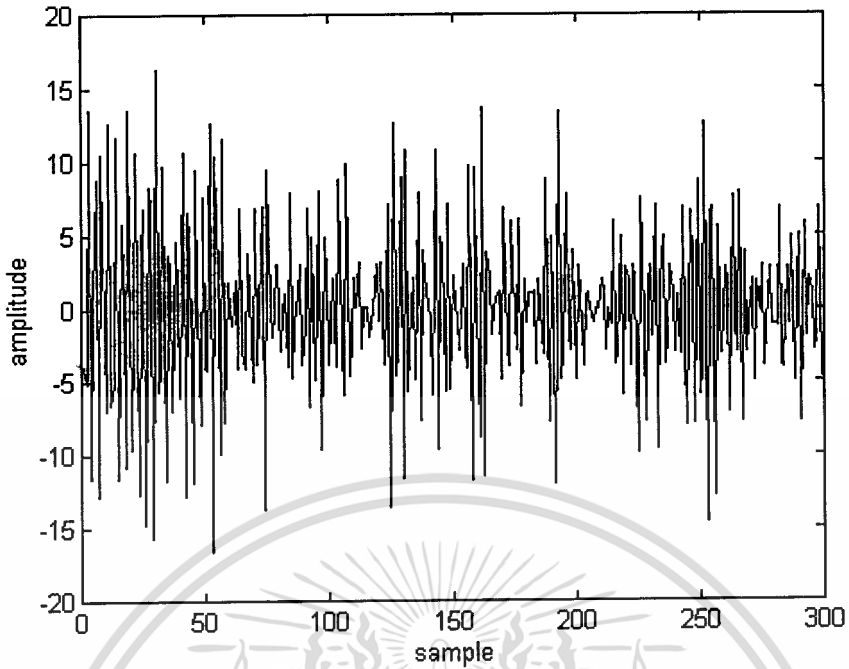
รูปที่ 7.4 เสียงว้อยซ์

จากรูปที่ 7.1(ก) สัญญาณมีลักษณะค่อนข้างสุ่ม ไม่เป็นคาบ คล้ายกับสัญญาณรบกวนแบบสุ่มหรือแรนด้อมนอยส์ เนื่องจากเป็นเสียงอันว้อยซ์ เมื่อสร้างเสียงจากสัมประสิทธิ์ในตารางที่ 7.1 ตามบล็อกไดอะแกรมที่ 4.1 โดยป้อนสัญญาณเข้าเป็นสัญญาณรบกวนแบบสุ่มจะได้สัญญาณออกดังรูปที่ 7.1(ข) ส่วนในรูป 7.2(ก) ซึ่งเป็นเสียงว้อยซ์ จะสามารถสร้างได้โดยใช้สัญญาณเข้าเป็นขบวนพัลส์ที่มีระยะห่างระหว่างพัลส์เท่ากับระยะพิทช์ จะได้สัญญาณออกดังรูปที่ 7.2(ข) จะเห็นได้ว่าสัญญาณในรูปที่ 7.1(ข) และ 7.2(ข) ที่สร้างได้มีความคล้ายคลึงกับรูปที่ 7.1(ก) และ 7.1(ข)

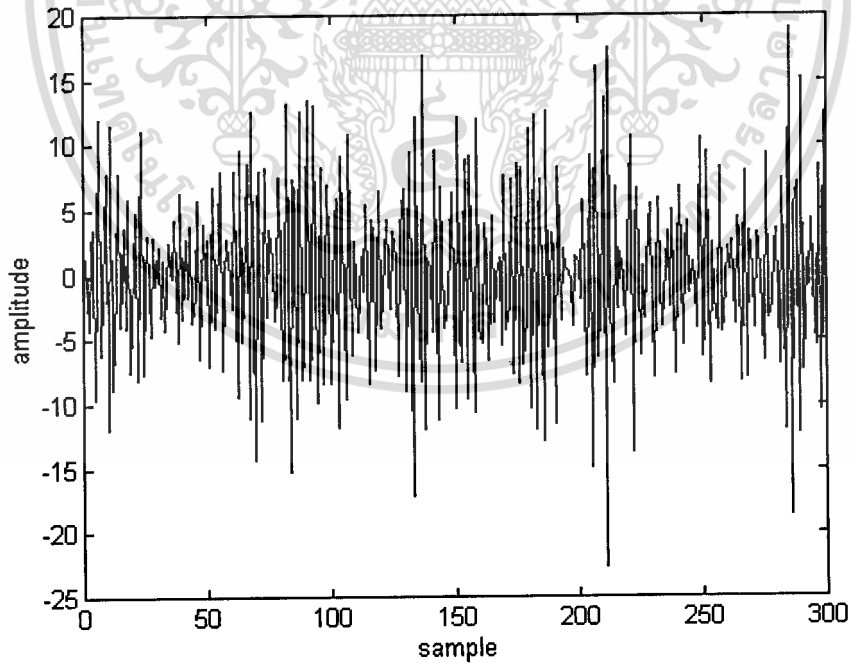
ส่วนรูปที่ 7.3 และ 7.4 เป็นการเปรียบเทียบสเปกตรัมของสัญญาณเสียงกับสเปกตรัมที่คำนวณจากสัมประสิทธิ์ LPC ของเสียงอันว้อยซ์ และเสียงว้อยซ์ ตามลำดับ จะเห็นว่า

1. สเปกตรัมของเสียงว้อยซ์ จะมียอด (peak) ที่ความถี่ต่ำ (น้อยกว่า 1000 เฮิรตซ์) ดังรูปที่ 7.4 ประมาณ 650 เฮิรตซ์ ซึ่งถือว่าเป็นความถี่ฟอร์แมนท์ที่หนึ่ง และยังมียอดถัดๆ ไปเรียกว่าความถี่ฟอร์แมนท์ที่ 2, 3, ... ตามลำดับ ส่วนเสียงอันว้อยซ์ยอดของความถี่ส่วนใหญ่จะมีค่าสูงที่ความถี่สูง (มากกว่า 1000 เฮิรตซ์)

2. สเปกตรัมที่ได้จากสัมประสิทธิ์ LPC จะเป็นค่าเฉลี่ยของสเปกตรัมเสียงเดิมทั้งเสียงว้อยซ์และเสียงอันว้อยซ์ ซึ่งก็เป็นการประมาณสเปกตรัมของเสียงนั่นเอง

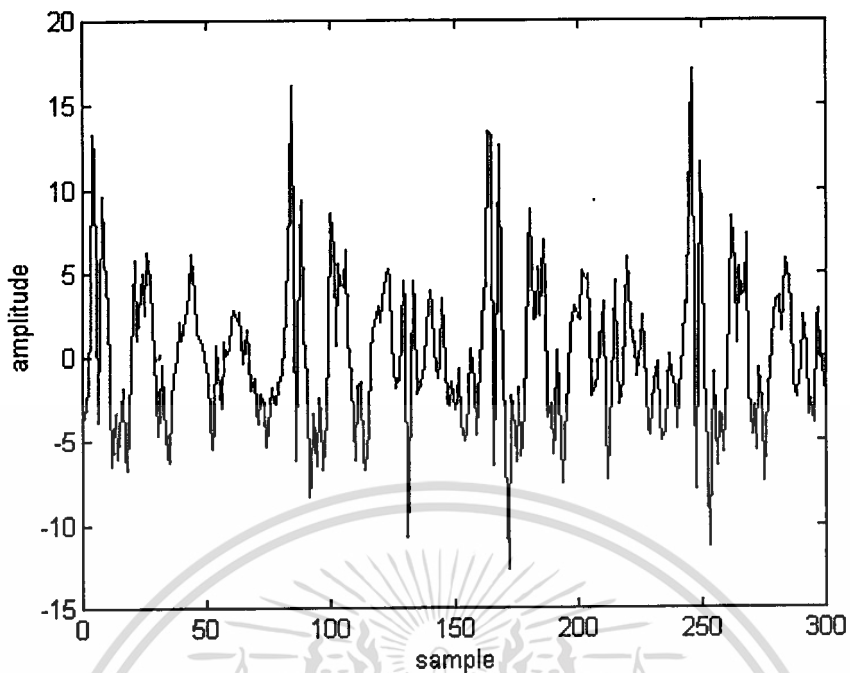


รูปที่ 7.1(ก) สัญญาณเสียงอินวอยซ์

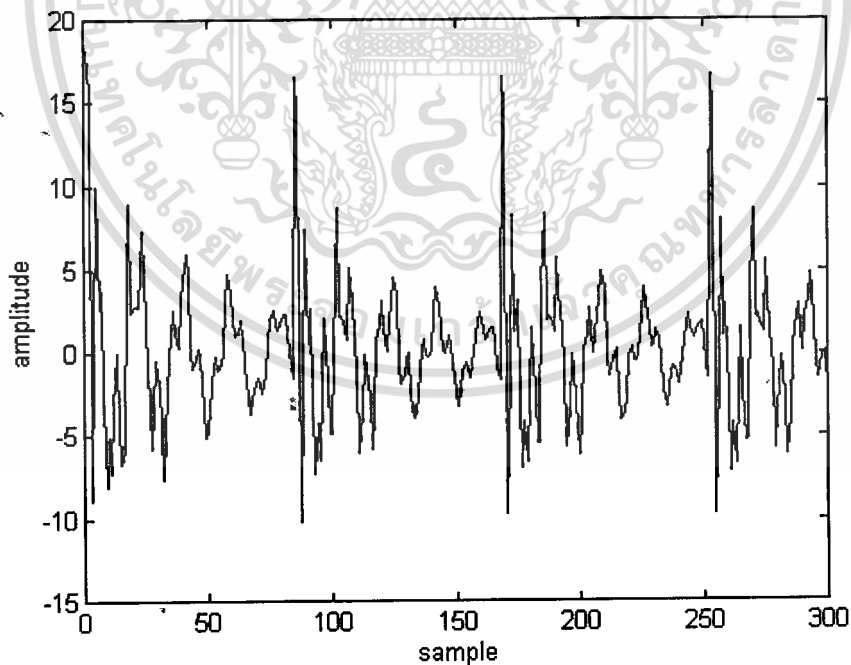


รูปที่ 7.1(ข) สัญญาณเสียงที่สร้างจากสัมประสิทธิ์ LPC

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

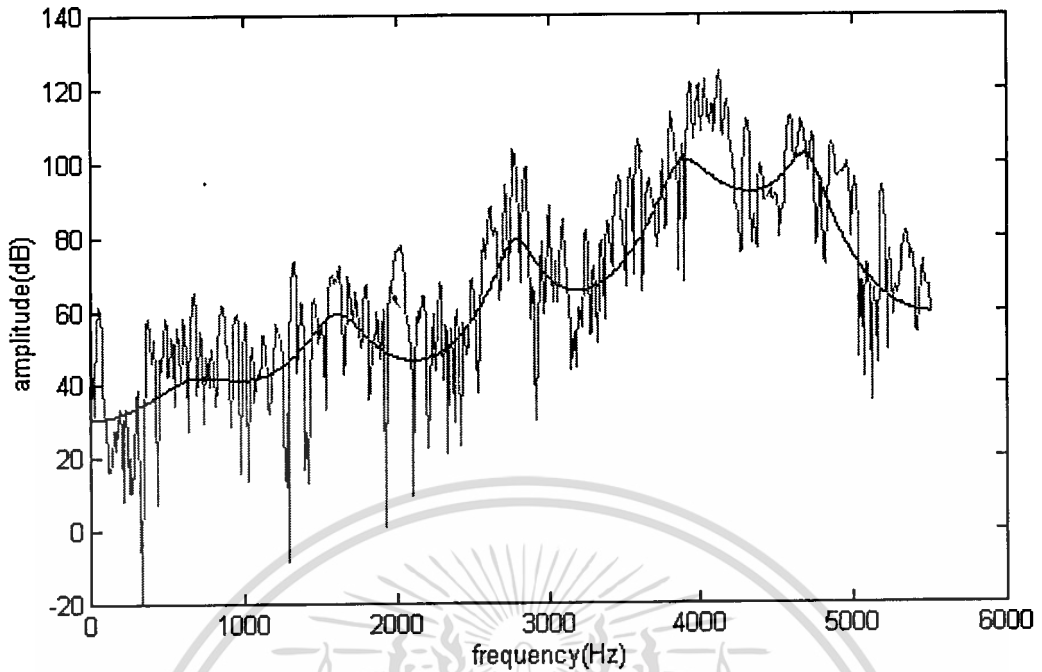


รูปที่ 7.2(ก) สัญญาณเสียงว้อยซ์

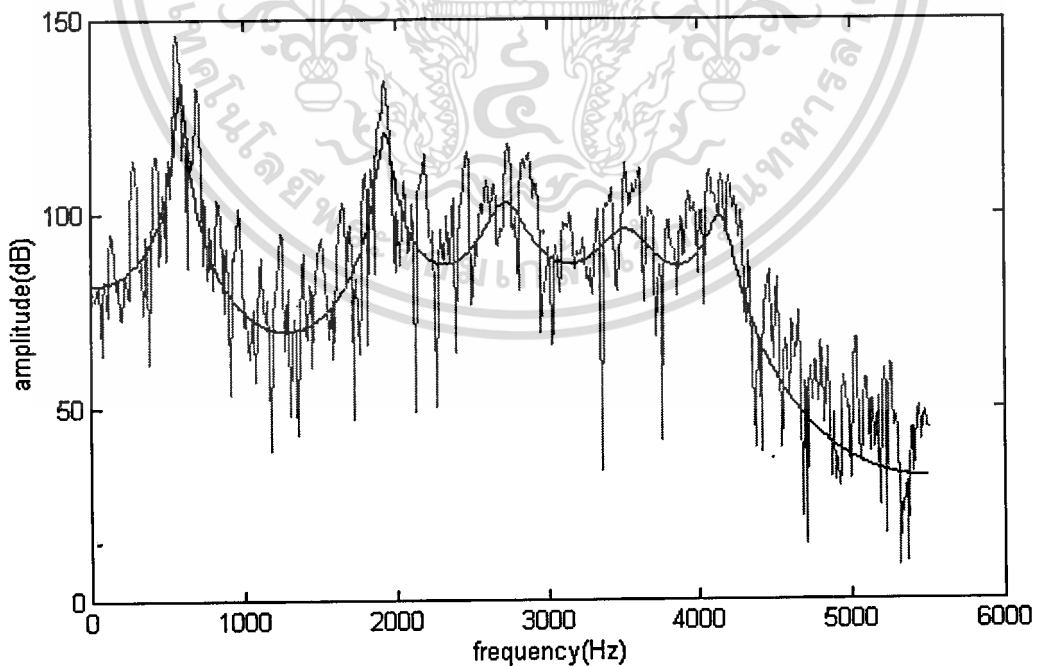


รูปที่ 7.2(ข) สัญญาณเสียงที่สร้างจากสัมประสิทธิ์ LPC

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 7.3 สเปกตรัมของสัญญาณเสียงอันว้อยซ์เปรียบเทียบกับสเปกตรัมที่ได้จากสัมประสิทธิ์ LPC



รูปที่ 7.4 สเปกตรัมของสัญญาณเสียงว้อยซ์เปรียบเทียบกับสเปกตรัมที่ได้จากสัมประสิทธิ์ LPC

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7.3 สเปกตรัมและสัญญาณความผิดพลาด เมื่อค่า p เปลี่ยนไป

ทดลองเปลี่ยนค่า p เป็น 4, 8, 12, 16 และ 20 แล้วเปรียบเทียบสเปกตรัม และสัญญาณความผิดพลาดที่ได้

รูปที่ 7.5 เปรียบเทียบสเปกตรัมที่ได้จากสัมประสิทธิ์ LPC กับสเปกตรัมเสียงเดิม

รูปที่ 7.5(ก) เมื่อ $p = 4$

รูปที่ 7.5(ข) เมื่อ $p = 8$

รูปที่ 7.5(ค) เมื่อ $p = 12$

รูปที่ 7.5(ง) เมื่อ $p = 16$

รูปที่ 7.5(จ) เมื่อ $p = 20$

รูปที่ 7.6 แสดงสัญญาณความคลาดเคลื่อน

รูปที่ 7.6(ก) เมื่อ $p = 4$

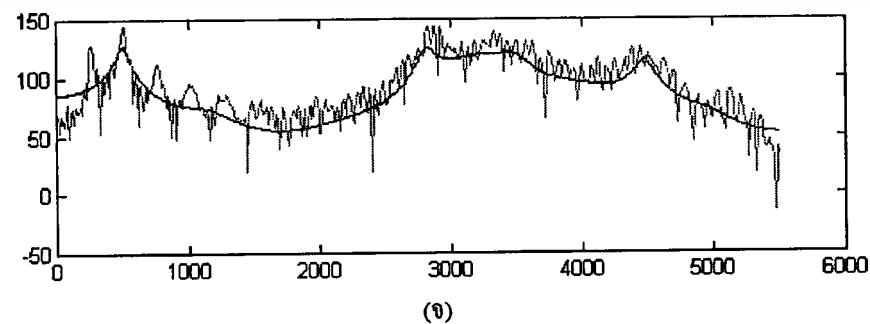
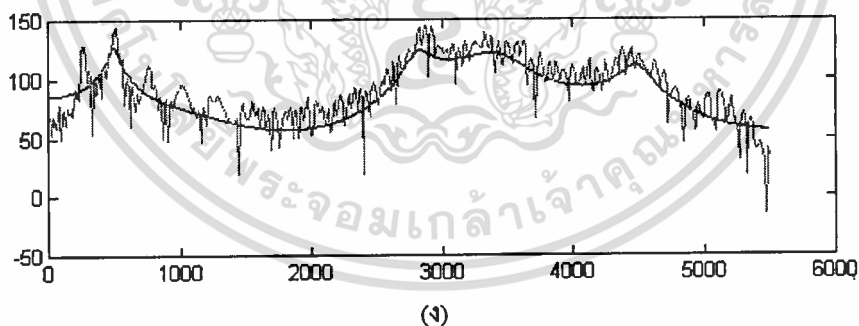
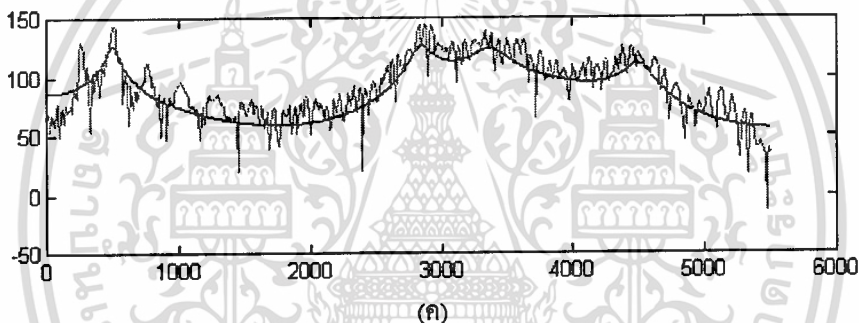
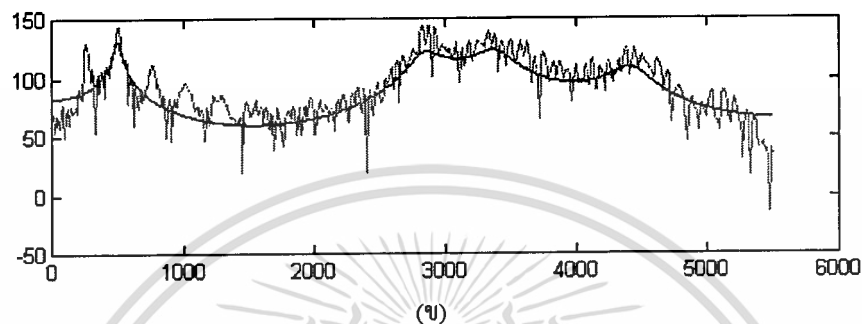
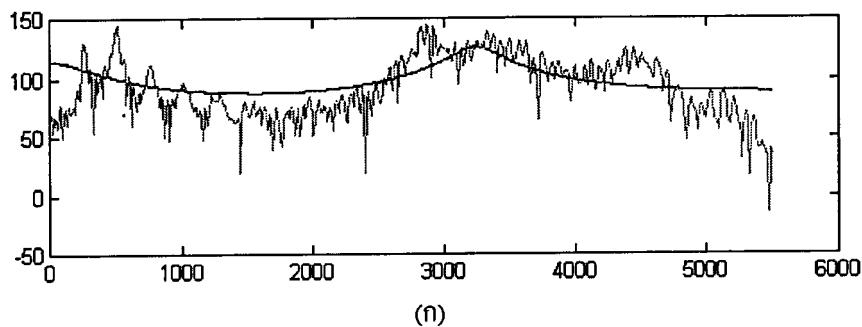
รูปที่ 7.6(ข) เมื่อ $p = 8$

รูปที่ 7.6(ค) เมื่อ $p = 12$

รูปที่ 7.6(ง) เมื่อ $p = 16$

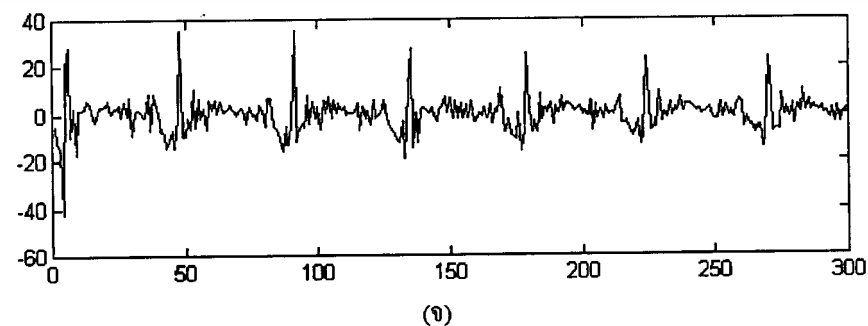
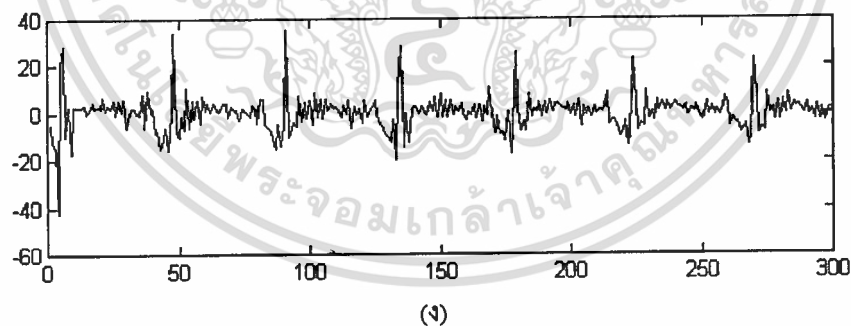
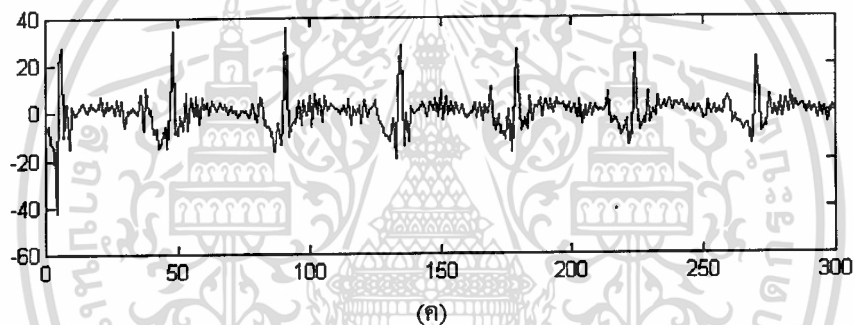
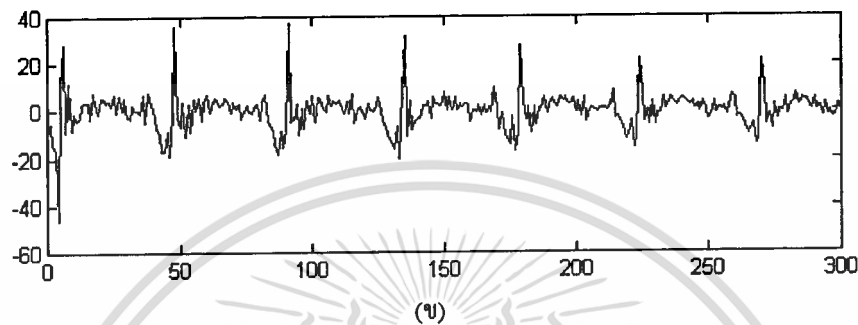
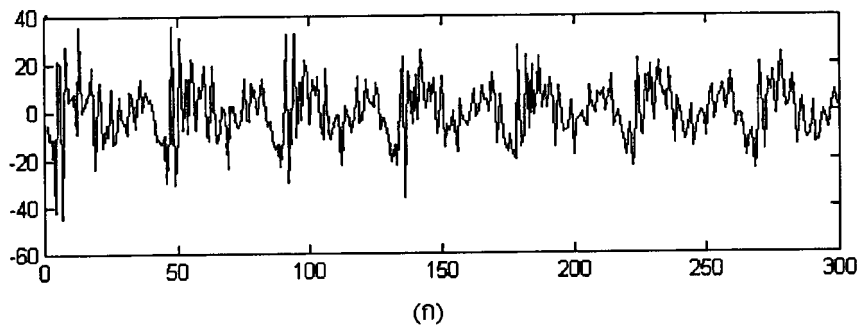
รูปที่ 7.6(จ) เมื่อ $p = 20$

เมื่อเปลี่ยนค่า p เป็น 4, 8, 12, 16 และ 20 จะพบว่า การประมาณสเปกตรัมจะมีความใกล้เคียงมากยิ่งขึ้น เมื่อค่า p ยิ่งสูงขึ้น ดังจะสังเกตได้ในรูปที่ 7.5 ส่วนในรูปที่ 7.6 เป็นการแสดงสัญญาณความผิดพลาดที่เกิดจากการประมาณเชิงเส้น พบว่า สัญญาณความผิดพลาดจะมีขนาดลดลงเมื่อค่า p สูงขึ้น อย่างไรก็ตาม ถึงแม้การเพิ่มค่า p จะทำให้การประมาณมีความใกล้เคียงมากยิ่งขึ้น แต่ก็ทำให้การคำนวณมีความยุ่งยาก และใช้เวลายาวนานขึ้น ดังนั้นเพื่อให้ได้ประสิทธิภาพสูงที่สุด จึงต้องมีการเลือกค่า p ที่เหมาะสม



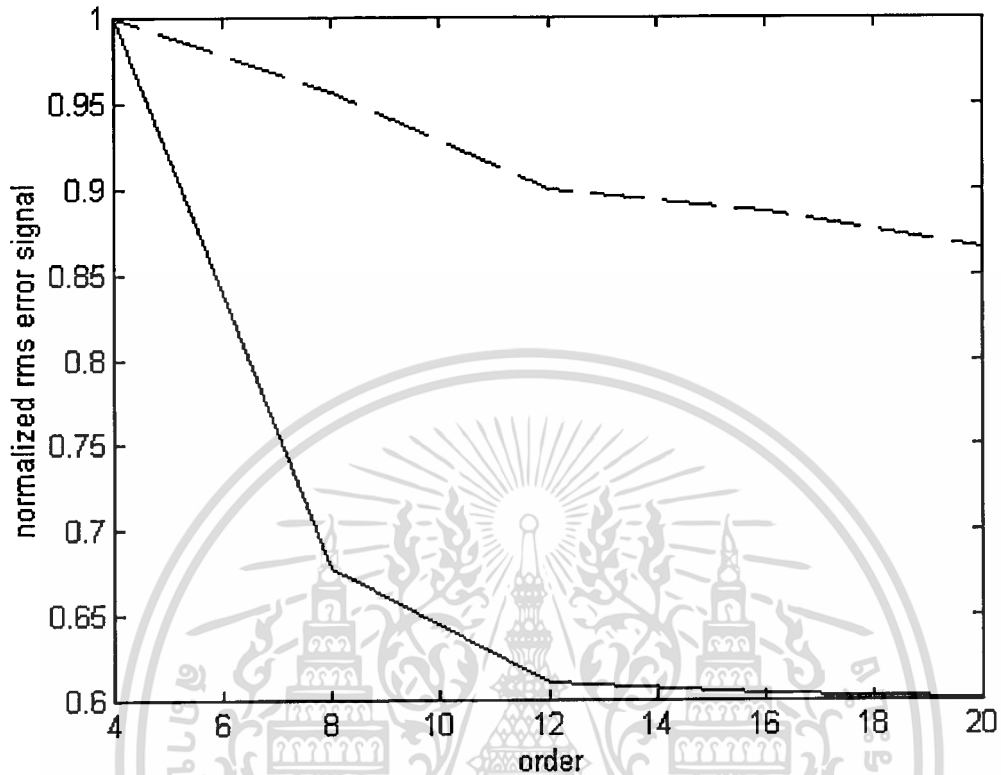
รูปที่ 7.5 สเปกตรัมของเสียงเปรียบเทียบกับสเปกตรัมที่ได้จากสัมประสิทธิ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ภายในห้องเรียนเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 7.6 สัญญาณความคลาดเคลื่อน (ก) $p=4$ (ข) $p=8$ (ค) $p=12$ (ง) $p=16$ (จ) $p=20$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 7.7 กราฟสัญญาณความผิดพลาดที่ผ่านการนอร์มอลไลซ์
เส้นประ แทน เสียงอันวอยซ์
เส้นทึบ แทน เสียงวอยซ์

ถ้าเรานำค่าพารามิเตอร์ต่างๆ ที่ได้มาเข้ารหัสแทนสัญญาณเสียงที่ละเฟรมซึ่งประกอบด้วย [14]

อัตราขยาย G	1 ไบต์
สัมประสิทธิ์ LPC	12 ไบต์
ระยะพิทช์รวมทั้งรหัสของเสียงวอยซ์ และอันวอยซ์	1 ไบต์

จะเห็นว่าเราสามารถลดอัตราในการส่งข้อมูลได้ในอัตรา $100/(1+12+1) = 7.1$

บทที่ 8

สรุป

การประมาณเชิงเส้นคือการประมาณสัญญาณเสียง จากผลรวมเชิงเส้นของสัญญาณเสียงในอดีต โดยอาศัยเกณฑ์กำลังสองของสัญญาณความคลาดเคลื่อนมีค่าต่ำสุดในการหาสัมประสิทธิ์การประมาณเชิงเส้น

จากการทดลองโดยใช้สัญญาณเสียงที่สุ่มด้วยความถี่ 11.025 กิโลเฮิร์ตซ์ ความยาว ประมาณ 0.5 ถึง 1 วินาที โดยแบ่งเป็นเฟรมๆ ละ 300 ตัวอย่างสัญญาณ และเลื่อนเฟรมทีละ 100 ตัวอย่างสัญญาณ และใช้การประมาณเชิงเส้นอันดับที่ 12 เราสามารถสรุปได้ดังนี้

1. เราสามารถหาความถี่พิทช์ได้โดยใช้โปรแกรม SIFT แต่ค่าความถี่พิทช์ที่ได้ยังมีความคลาดเคลื่อนอยู่บ้างเนื่องจากความคลาดเคลื่อนในการคำนวณ ของโปรแกรม จึงได้มีการปรับปรุงโดยใช้อาศัยวงจรกรองมัธยฐานซึ่งอธิบายไว้ในบทแทรกที่ 1

2. จากสัมประสิทธิ์ที่ได้ เราสามารถคำนวณหาสเปคตรัม และพบว่าสเปคตรัมที่ได้มีความสอดคล้องกันกับสเปคตรัมของสัญญาณเสียงที่นำมาใช้ในการทดลองทั้งเสียงวอยซ์และเสียงอันวอยซ์ สำหรับเสียงวอยซ์นั้น จะมีขอดความถี่ที่เด่นในช่วงความถี่ที่ต่ำกว่า 1 กิโลเฮิร์ตซ์ เรียกว่า ความถี่ฟอร์แมนท์ที่หนึ่ง นอกจากนี้ยังมีความถี่ฟอร์แมนท์ที่ 2, 3,... ส่วนเสียงอันวอยซ์ ที่ความถี่ต่ำ แอมพลิจูดของสเปคตรัมจะมีค่าต่ำกว่า แอมพลิจูดที่ความถี่สูง

3. เมื่อเปรียบเทียบสเปคตรัมโดยเปลี่ยนค่าจำนวนสัมประสิทธิ์ของ LPC (p) แล้วพบว่า เมื่อใช้จำนวนสัมประสิทธิ์ในการประมาณมากขึ้น การประมาณสเปคตรัมจะมีความใกล้เคียงสูงขึ้น

4. เมื่อนำพารามิเตอร์มาสร้างสัญญาณเสียงกลับคืน โดยใช้ระยะพิทช์จากโปรแกรม SIFT (ในบทแทรกที่ 1) ในการสร้างขบวนพัลส์ สำหรับสัญญาณเสียงวอยซ์ และ สัญญาณรบกวนแบบสุ่ม สำหรับสัญญาณเสียงอันวอยซ์ สามารถสร้างเสียงขึ้นมาในระดับที่พอจะจำแนกได้ว่าเป็นคำว่าอะไร แต่ยังไม่มีความชัดเจนในเรื่องสำเนียงอยู่มาก ถึงแม้ว่าจะได้มีการปรับปรุงระยะพิทช์โดยใช้วงจรกรองมัธยฐานแล้วก็ตาม(ในการสร้างสัญญาณเสียงนั้น ได้ใช้วิธีการอินเตอร์โพลเทท(interpolate) เพื่อให้ได้ความถูกต้องแม่นยำยิ่งขึ้น)

5. อัตราในการบีบสัญญาณที่คำนวณได้ ประมาณ 7 เท่า

แนวทางการพัฒนาศึกษาต่อไป

- ขั้นตอนต่างๆ ในการคำนวณในโปรแกรมต่างๆ ไม่ว่าจะเป็น โปรแกรม MAIN , LPC หรือ SIFT ก็ตามยังเป็นเพียงโปรแกรมที่เขียนเพื่อเป็นการจำลองการใช้งานจริงเท่านั้น แต่ยังมีได้คำนึงถึงการออปติไมซ์(optimization) ขั้นตอนวิธีแต่อย่างใด ดังนั้นการคำนวณจึงใช้เวลานานพอสมควรสำหรับเสียงความยาว 0.5 วินาที ที่สุ่มสัญญาณด้วยความถี่ 11.025 กิโลเฮิรตซ์ และใช้ค่า p เท่ากับ 12 จะใช้เวลาในการคำนวณสัมประสิทธิ์ ประมาณไม่ถึง 1 นาที เมื่อใช้คอมพิวเตอร์รุ่น 486 DX2-66 ในการประมวลผล

- พัฒนาขั้นตอนวิธีที่ได้ สร้างเป็นวงจร(hardware) และนำไปใช้ในการบีบอัดสัญญาณเสียง เพื่อประโยชน์ในการติดต่อสื่อสาร

- เพราะวาระยะพิทซ์ที่ได้ยังมีความคลาดเคลื่อนอยู่มาก ถึงแม้จะมีการปรับปรุงโดยใช้การกรองแบบมัลติแชนแนลก็ตาม ดังนั้นจึงเป็นเรื่องน่าสนใจที่จะศึกษาค้นคว้าในการปรับปรุงระยะพิทซ์ให้ดีขึ้น เพราะจากผลการทดลองสร้างเสียงขึ้นนั้น เสียงที่ได้ยังพอที่จะจำแนกออกได้ว่าเป็นคำว่าอะไร แต่ยังมี ความเพี้ยนในเรื่องสำเนียง ถ้ามีการปรับปรุงระยะพิทซ์ให้ดีขึ้นแล้วเสียงที่สร้างน่าจะมีคุณภาพดีขึ้น

- นำหลักการการประมาณเชิงเส้นไปใช้ในการจดจำเสียงพูด

- เพิ่มข้อมูลสัญญาณเข้า จากคำเป็นประโยค

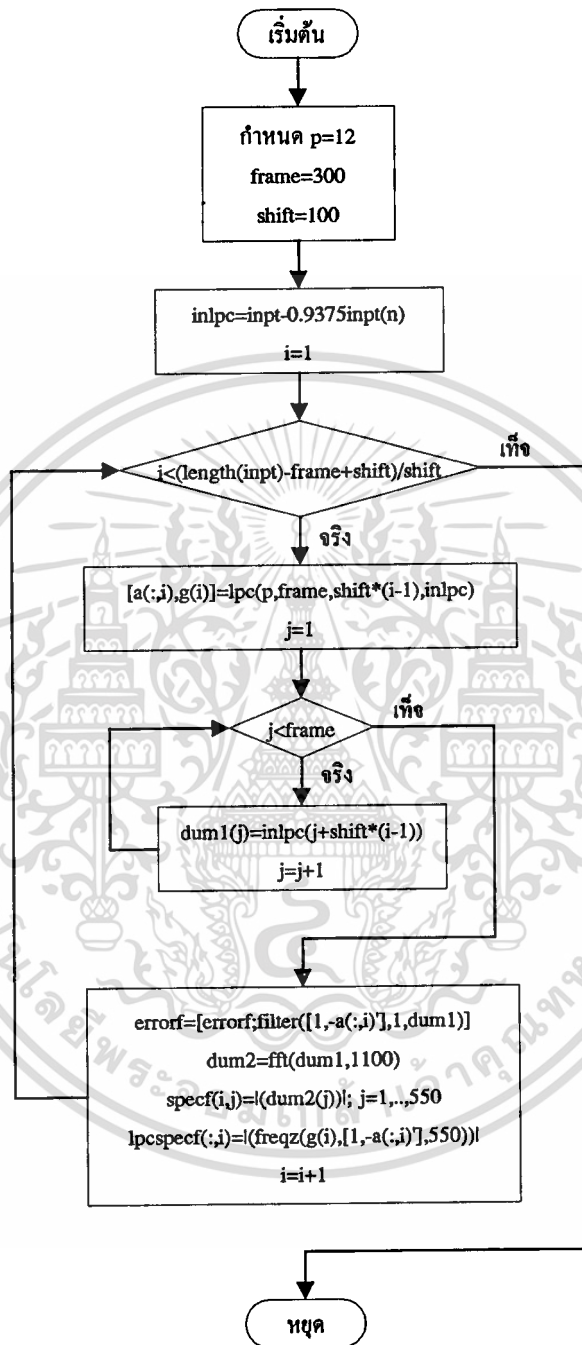
บทแทรกที่ 1

โปรแกรมที่ใช้ในการคำนวณ

ในบทนี้ได้รวบรวมโปรแกรมต่างๆ ที่ใช้ในการคำนวณและวิเคราะห์ซึ่งในที่นี่จะใช้ไวยากรณ์ของโปรแกรม MATLAB โดยแบ่งออกเป็น

- โปรแกรมหลัก ชื่อ MAIN
- โปรแกรมการคำนวณสัมประสิทธิ์และอัตราขยาย ชื่อ LPC
- โปรแกรมการหาคาบเวลาพิทช์ ชื่อ SIFT
- โปรแกรมการสร้างสัญญาณเสียง ชื่อ RECONSTRUCT
- โปรแกรมวงจรกรองมีรยฐาน ชื่อ MEDIANS





รูปที่ ๓.๑.๑ โฟลวชาร์ทของโปรแกรม MAIN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

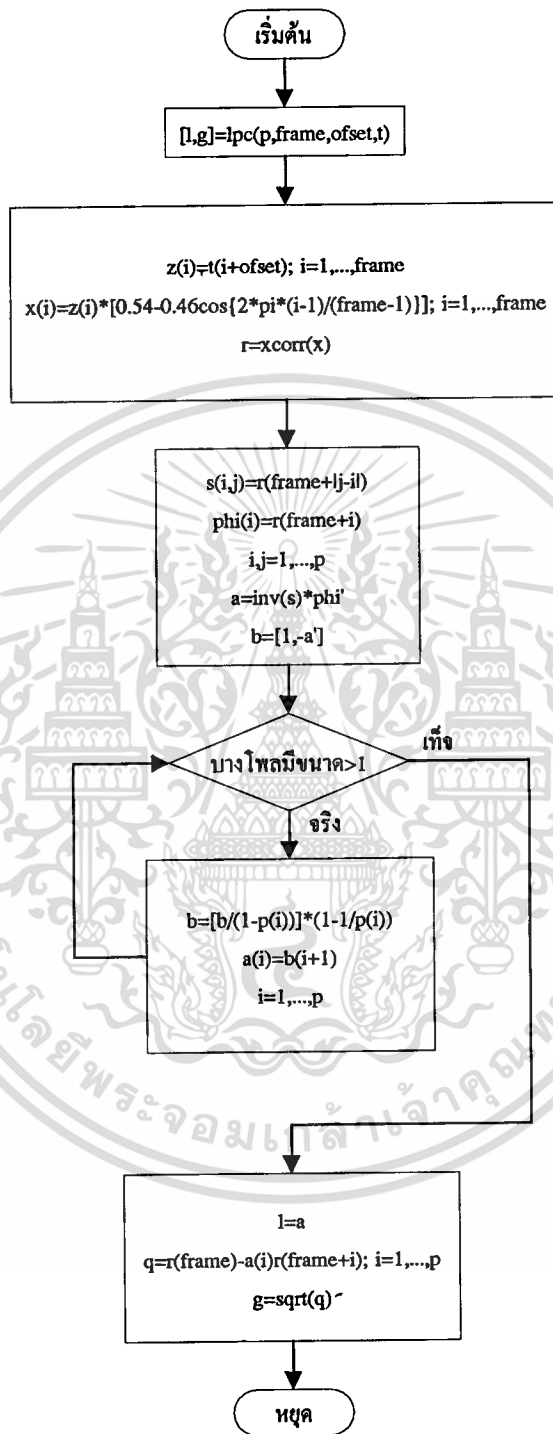
โปรแกรม MAIN

```

clear;
% โหลดไฟล์ .wav ที่จะใช้ในการวิเคราะห์ เข้ามาเก็บไว้ในตัวแปรชื่อ inpt
[inpt,fp,form]=loadwave('c:\matlab\toolbox\work\a_noi.wav');
% เซตค่าคงที่ต่างๆ ดังนี้
errorf=[];
p=12; % จำนวนสัมประสิทธิ์ของวงจรกรอง
frame=300; % ความยาวของเฟรม
shift=100; % ระยะในการเลื่อนเฟรม
% นำ input ชื่อ inpt มาผ่านการเอมฟาซิส แล้วเก็บไว้ในตัวแปร inlpc
inlpc=filter([1,-0.9375],[1],inpt);
% การคำนวณพารามิเตอร์ต่างๆ ในการประมาณ เก็บไว้ในตัวแปร a(j,i) และ g(i)
for i=1:(length(inpt)-frame+shift)/shift;
    [a(:,i),g(i)]=lpc(p,frame,shift*(i-1),inlpc);
    for j=1:frame;
        dum1(j)=inlpc(j+shift*(i-1));
    end
% นำสัมประสิทธิ์ที่ได้มาหาความคลาดเคลื่อนกับสัญญาณเสียงพูด
    errorf=[errorf;filter([1,-a(:,i)'],1,dum1)];
% คำนวณสเปกตรัมของสัญญาณเสียง
    dum2=fft(dum1,1100);
    for j=1:550;
        specf(i,j)=abs(dum2(j));
    end
% คำนวณสเปกตรัมจากสัมประสิทธิ์ที่หาได้
    lpcspecf(:,i)=abs(freqz(g(i),[1,-a(:,i)],550));
end
lpcspecf=lpcspecf';
% เซตโดเมนความถี่ จาก 0 ถึง 5500 เฮิรตซ์
for i=1:550;
    f(i)=10*i;
end

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ท1.2 โพลซาร์ทของโปรแกรม LPC

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรม LPC

```

% ตั้งชื่อ ฟังก์ชัน lpc
function [l,g] = lpc(p,frame,offset,t)

format long;
% คำนวณ input ที่จะใช้ในการคำนวณ เมื่อกำหนดขนาด(frame) และจุดเริ่มต้น(offset) ของข้อมูล
for i=1:frame;
    z(i)=t(i+offset);
end
% นำสัญญาณที่ได้มาผ่านการวินโดว์เพื่อเตรียมวิเคราะห์ต่อไป
for i=1:frame;
    x(i)=z(i)*(0.54-0.46*cos(2*pi*(i-1)/(frame-1)));
end
% หาอโตคอร์รีเลชันและเมตริกซ์ของอโตคอร์รีเลชัน
r=xcorr(x);
for i=1:p,
    for j=i:p,
        s(i,j)=r(frame+j-i);
        s(j,i)=s(i,j);
    end
end
phi(i)=r(frame+i);
end
a=inv(s)*phi';
b=[1,-a'];
% ปรับปรุงเสถียรภาพของวงจรกรองที่ได้ โดยแทนโพลของวงจรกรองที่มีขนาดเกินหนึ่งด้วยคอนจูเกตเชิงซ้อน
c=roots([1,-a']);
for i=1:p,
    if abs(c(i))>1
        b=deconv(b,[1,-c(i)]);
        b=conv(b,[1,-1/c(i)]);
    end
end
end

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

a(i)=-b(i+1);

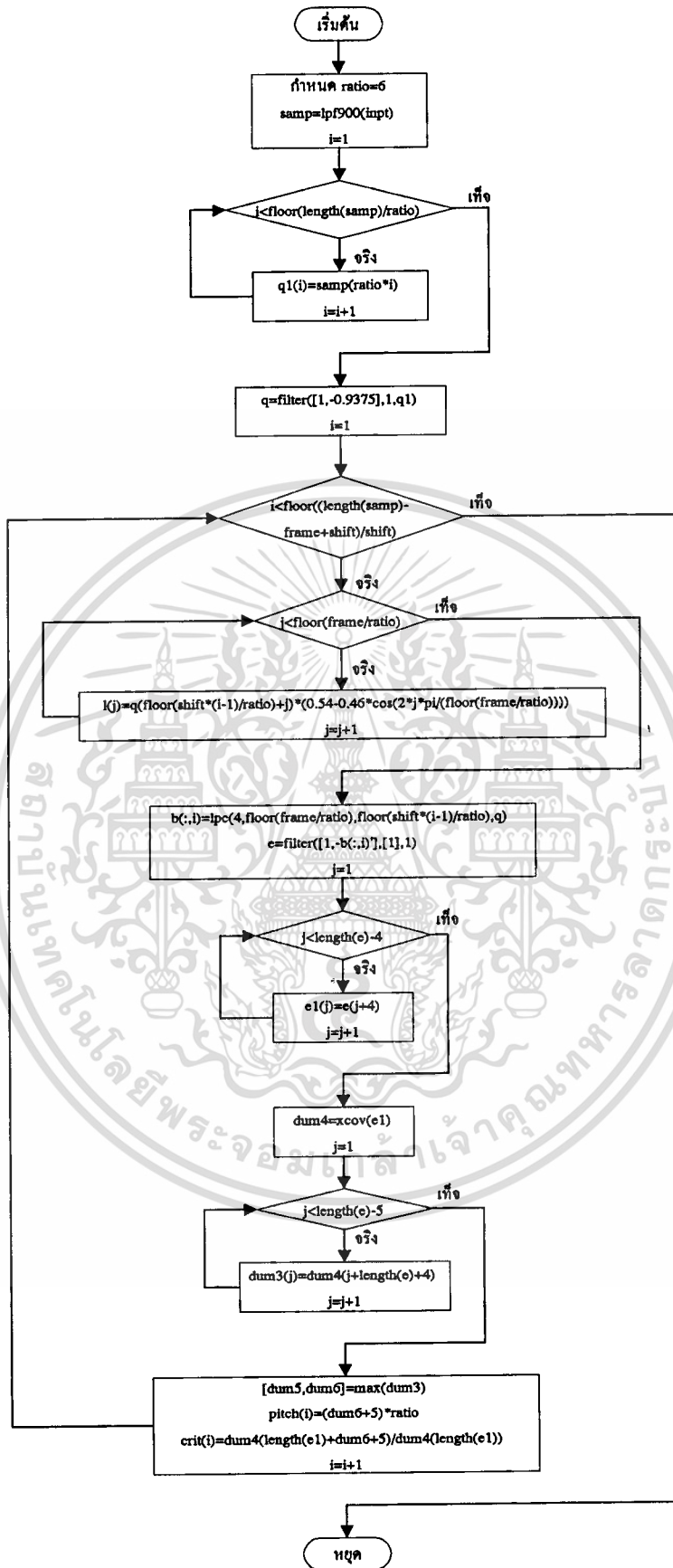
end

l=a;
% คำนวณอัตราขยาย G
q=r(frame);
for i=1:p,
    q=q-a(i)*r(frame+i);
end
g=sqrt(q);

```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ๓1.3 ฟลวชาร์ทของโปรแกรม SIFT

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

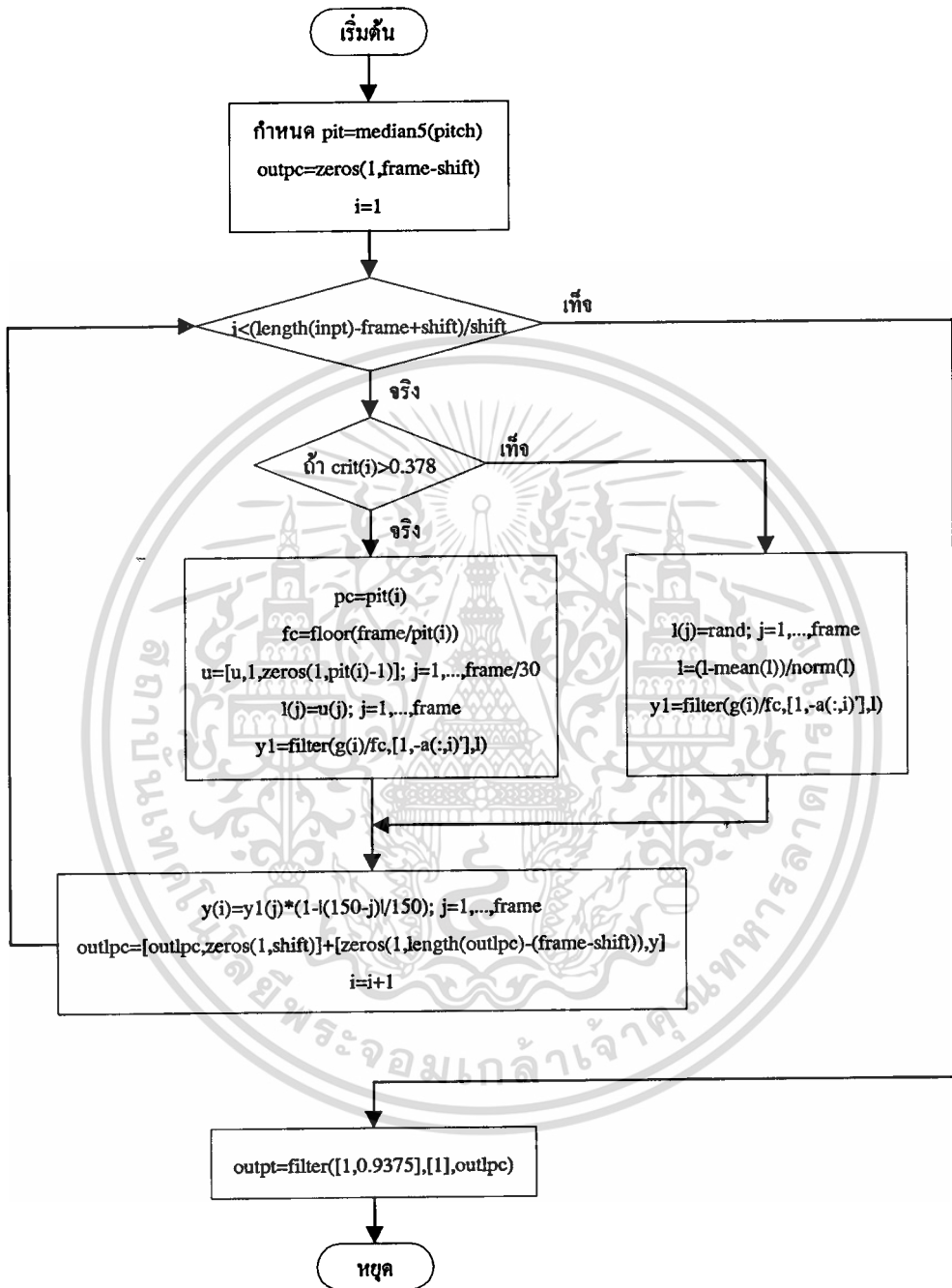
โปรแกรม SIFT

```

% กำหนดอัตราส่วนในการสุ่มสัญญาณ 6:1
ratio=6;
% เตรียม input เพื่อใช้ในการหาระยะพิทช์ โดยผ่านวงจรกรองความถี่ต่ำผ่าน 900 เฮิรตซ์ และทำการสุ่มสัญญาณ
% ด้วยอัตราส่วน 6 ต่อ 1
samp=lpf900(inpt);
for i=1:floor(length(samp)/ratio);
    q1(i)=samp(ratio*i);
end
% ผ่านการพรีเอมฟาซิส
q=filter([1,-0.9375],1,q1);
for i=1:floor((length(samp)-frame+shift)/shift);
% แบ่งเฟรมของ input และ การวินโดว์ด้วยวินโดว์แฮมมิง
    l=[];
    for j=1:floor(frame/ratio);
        l(j)=q(floor(shift*(i-1)/ratio)+j)*(0.54-0.46*cos(2*j*pi/(floor(frame/ratio))));
    end
% คำนวณสัมประสิทธิ์รีจิงกรองอันดับ 4
    b(:,i)=lpc(4,floor(frame/ratio),floor(shift*(i-1)/ratio),q);
% หาสัญญาณความคลาดเคลื่อนจากสัมประสิทธิ์ที่ได้
    e=filter([1,-b(:,i)],[1],l);
    for j=1:length(e)-4;
        e1(j)=e(j+4);
    end
% หาออโตคอรรีเลชันเพื่อใช้ในการวิเคราะห์ระยะพิทช์
    dum4=xcov(e1);
    for j=1:length(e1)-5;
        dum3(j)=dum4(j+length(e1)+4);
    end
    [dum5,dum6]=max(dum3);
    pitch(i)=(dum6+5)*ratio;
    crit(i)=dum4(length(e1)+dum6+5)/dum4(length(e1))
end

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ๓1.4 โฟลวชาร์ทของโปรแกรม RECONSTRUCT

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

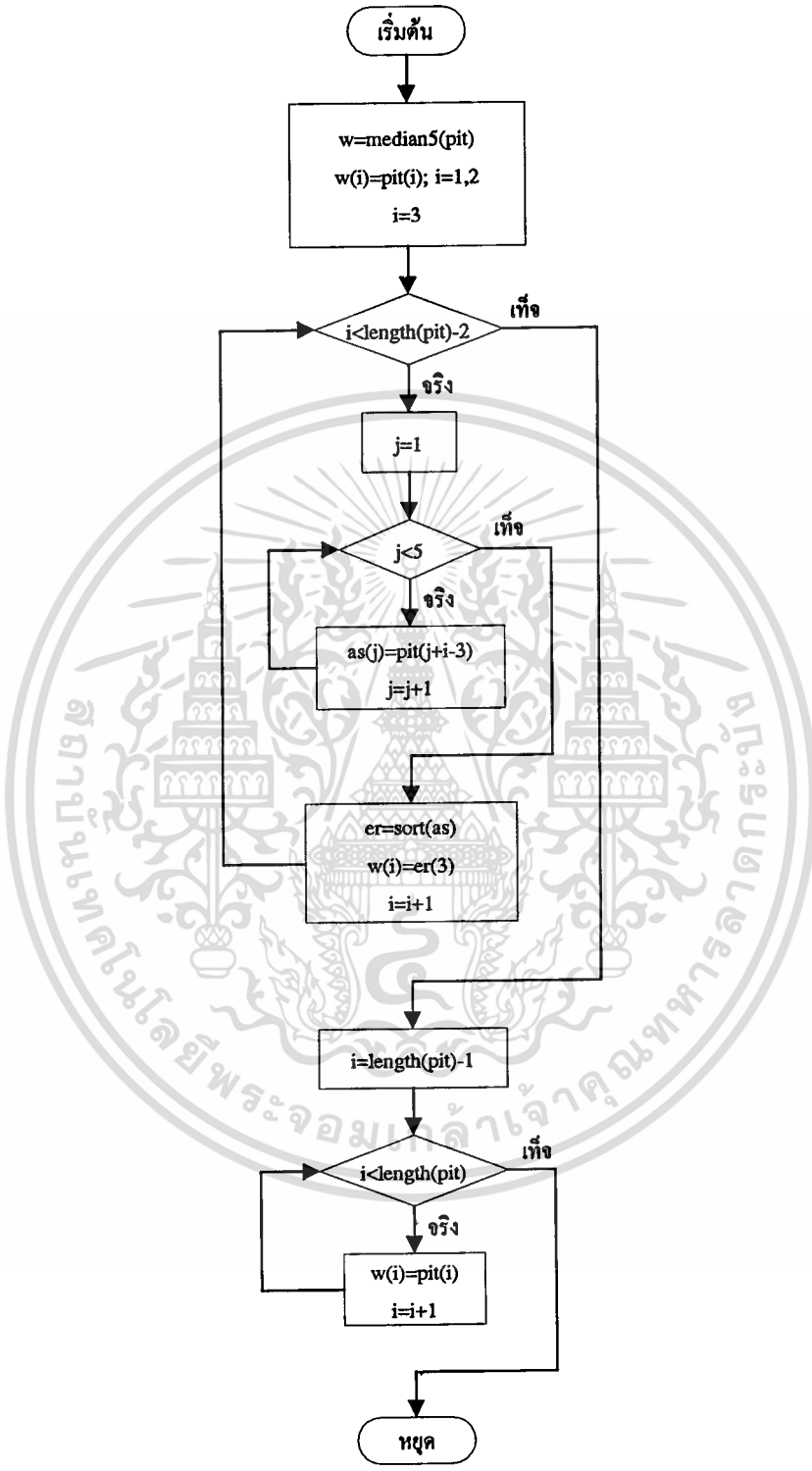
โปรแกรม RECOSTRUCT

```

pit=median5(pitch);
outlpc=zeros(1,frame-shift);
for i=1:(length(inpt)-frame+shift)/shift;
% เชค crit ถ้าโตกว่า 0.38 เป็นเสียงวอยซ์ ถ้าไม่เป็นเสียงอันวอยซ์
    if crit(i)>0.378
        pc=pit(i);
        fc=floor(frame/pit(i));
        u=[];
% เตรียมสัญญาณเข้าเป็นขบวนพัลส์ สำหรับสร้างเสียงวอยซ์
        for j=1:frame/30;
            u=[u,1,zeros(1,pit(i)-1)];
        end
        for j=1:frame;
            l(j)=u(j);
        end
        y1=filter(g(i)/fc,[1,-a(:,i)],l);
    else
% เตรียมสัญญาณเข้าเป็นแรนด้อมนอยส์สำหรับสร้างเสียงอันวอยซ์
        for j=1:frame;
            l(j)=rand;
        end
        l=(l-mean(l))/norm(l);
        y1=filter(g(i),[1,-a(:,i)],l);
    end
% การอินเตอร์โพลเลท
    for j=1:frame;
        y(j)=y1(j)*(1-abs(150-j)/150);
    end
    outlpc=[outlpc,zeros(1,shift)]+[zeros(1,length(outlpc)-(frame-shift)),y];
end
% การดีเอมฟาซิส
outpt=filter([1,0.9375],[1],outlpc);

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ๓1.5 ไพลวาร์ทของโปรแกรม MEDIAN5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

MEDIAN5

```

% ตั้งชื่อฟังก์ชัน
function w=median5(pit)

for i=1:2;
w(i)=pit(i);
end
for i=3:length(pit)-2;
    for j=1:5;
        as(j)=pit(j+i-3);
    end

% เรียงข้อมูล
    er=sort(as);
%เลือกเอาค่าตรงกลาง
    w(i)=er(3);
end
for i=length(pit)-1:length(pit);
    w(i)=pit(i);
end

```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทแทรกที่ 2

ตัวอย่างของการวิเคราะห์และสังเคราะห์สัญญาณเสียง

ในบทนี้ได้รวมผลการทดลองนอกเหนือจากที่แสดงไว้ในบทที่ 7 ในการหาค่าพารามิเตอร์ต่างๆ (สัมประสิทธิ์ของวงจรรองและอัตรากรขยาย) โดยวิธีออคโตออร์รีเลขัน ในการประมาณเชิงเส้นรวมถึงพล็อตสเปกตรัมของสัญญาณเสียงเปรียบเทียบกับสเปกตรัมของวงจรรองการประมาณเชิงเส้น โดยแบ่งเป็น 4 ตัวอย่างคือ

1. เสียงอันว้อยซ์ของผู้ชาย
2. เสียงว้อยซ์ของผู้ชาย
3. เสียงอันว้อยซ์ของผู้หญิง
4. เสียงว้อยซ์ของผู้หญิง

ตารางที่ ท2.1 แสดงค่าพารามิเตอร์ของเฟรมหนึ่งเฟรมของสัญญาณเสียง

เสียงที่ 1 แทนเสียงอันว้อยซ์ของผู้ชาย

เสียงที่ 2 แทนเสียงว้อยซ์ของผู้ชาย

เสียงที่ 3 แทนเสียงอันว้อยซ์ของผู้หญิง

เสียงที่ 4 แทนเสียงว้อยซ์ของผู้หญิง

พารามิเตอร์	เสียงที่ 1	เสียงที่ 2	เสียงที่ 3	เสียงที่ 4
G	24.656531184	54.323564026	12.454005918	78.435264926
a(1)	-1.247326144	0.860407461	-0.561921225	-0.170430741
a(2)	-1.041388796	-1.232998471	-0.316828746	-0.413555535
a(3)	-0.337743688	1.488384402	-0.123101049	1.210919148
a(4)	-0.405478347	-1.342893828	-0.020392531	0.373740056
a(5)	-0.201609377	1.572002474	0.279909104	0.511149733
a(6)	-0.367430317	-1.268723123	-0.050644561	-1.053406567
a(7)	-0.118282064	0.892472479	0.111221020	-0.234837452
a(8)	-0.043897573	-0.933865803	0.013879693	-0.574064647
a(9)	0.188637772	0.178204565	0.107368452	0.478540058
a(10)	-0.228553726	-0.341942969	0.104216395	-0.054685122
a(11)	-0.320994684	0.034636535	0.235676561	0.058926182
a(12)	-0.212322522	0.171222710	0.144186450	-0.169664346

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์และสงวนสิทธิ์ในเนื้อหาทั้งหมด ไม่อนุญาตให้ทำซ้ำโดยไม่ได้รับอนุญาตจากสำนักพิมพ์

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ท2.1 แสดงรายละเอียดของเสียงอันว้อยซ์ของผู้ชาย

รูปที่ ท2.2 แสดงรายละเอียดของเสียงว้อยซ์ของผู้ชาย

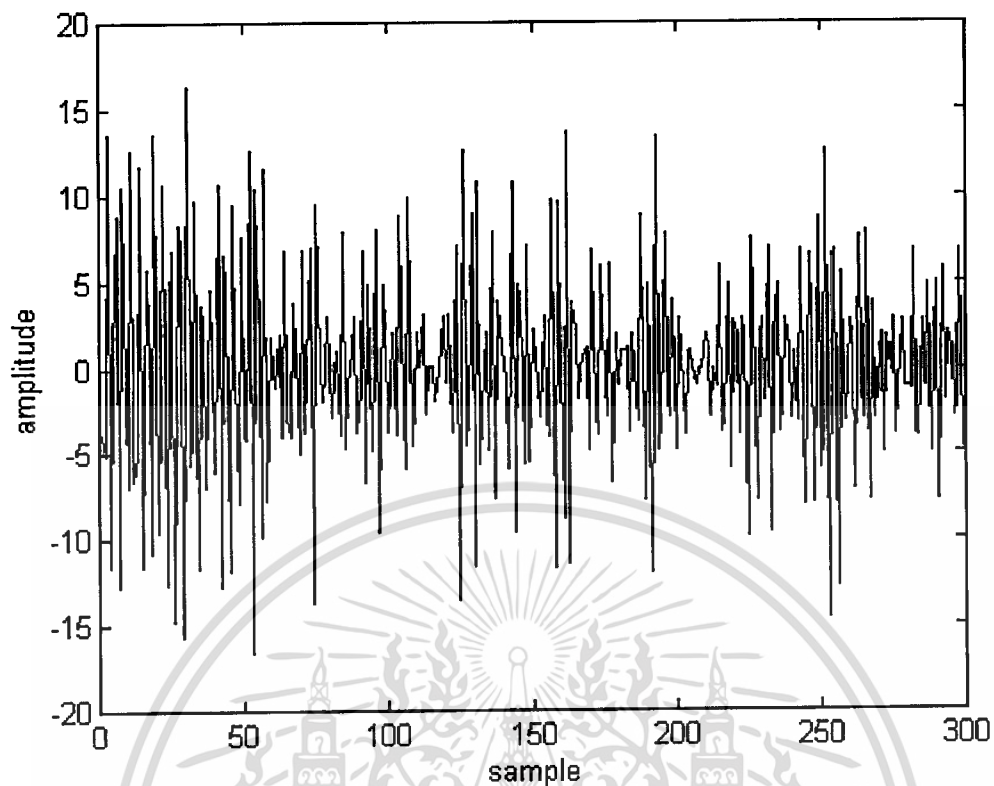
รูปที่ ท2.3 แสดงรายละเอียดของเสียงอันว้อยซ์ของผู้หญิง

รูปที่ ท2.4 แสดงรายละเอียดของเสียงว้อยซ์ของผู้หญิง

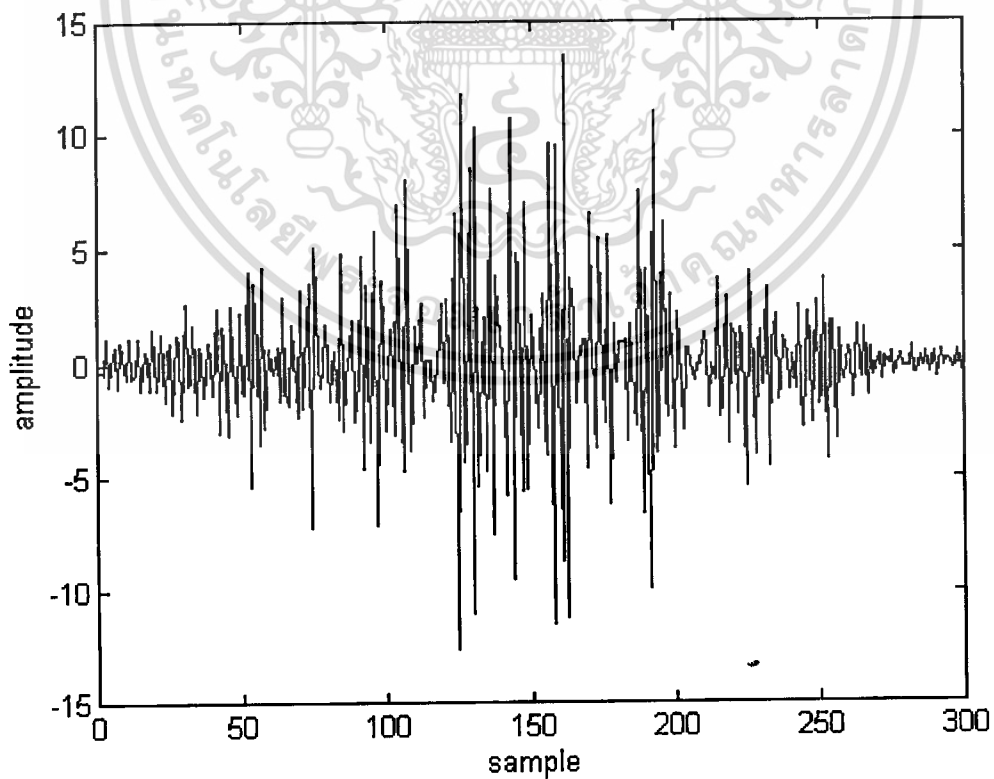
โดยแบ่งแยกย่อยดังนี้

- (ก) แสดงสัญญาณเสียงช่วงหนึ่งเฟรม
- (ข) แสดงสัญญาณเสียงที่ผ่านการวินโดว์แฮมมิง
- (ค) แสดงสเปกตรัมของสัญญาณ ในรูป (ข)
- (ง) แสดงสเปกตรัมของ LPC
- (จ) แสดงสัญญาณความคลาดเคลื่อน
- (ฉ) แสดงสเปกตรัมของสัญญาณความคลาดเคลื่อน



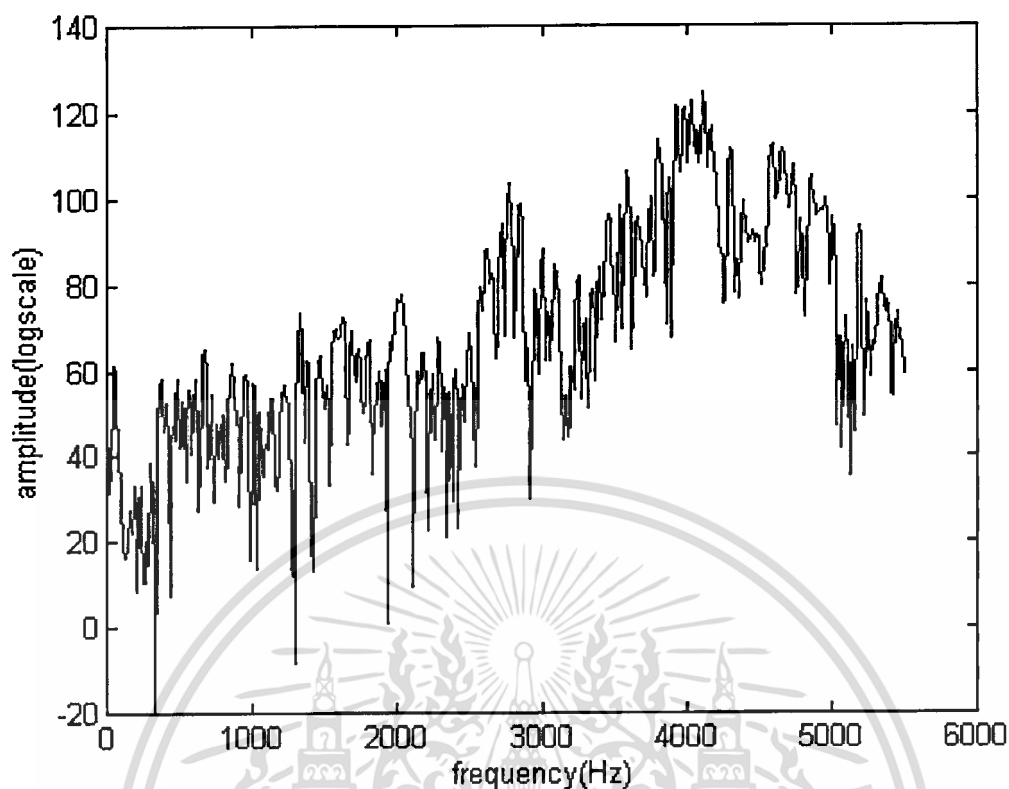


รูปที่ ท2.1(ก) รูปคลื่นของเสียงอันว้อยซ์ของผู้ชาย

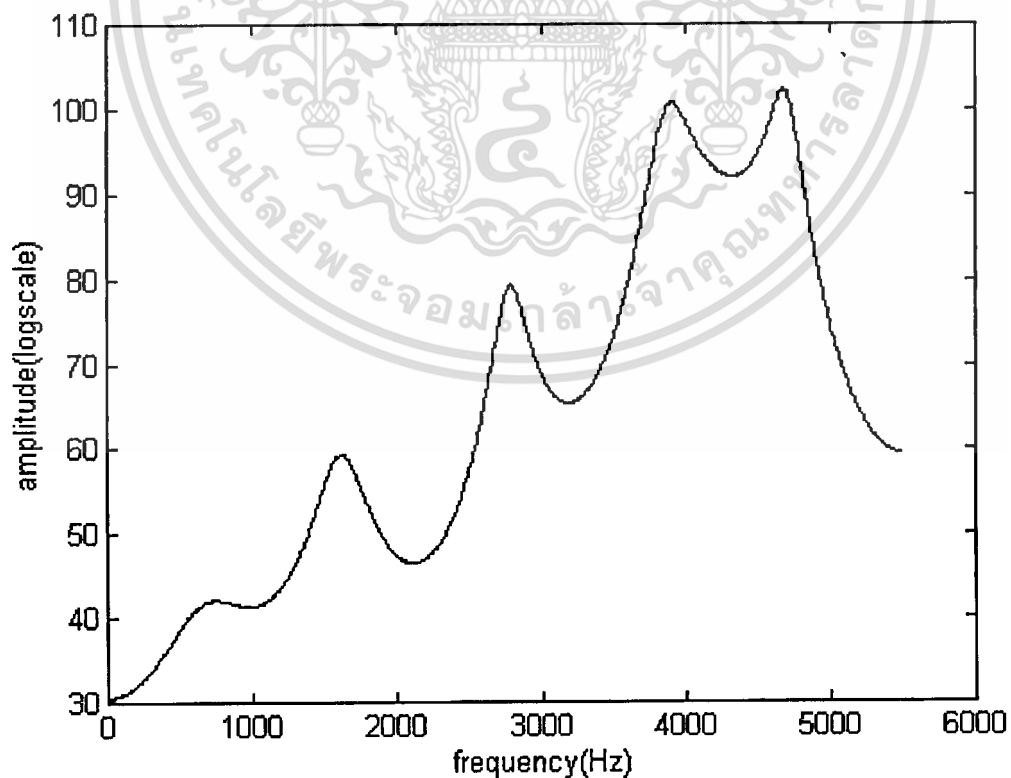


รูปที่ ท2.1(ข) รูปคลื่นที่ผ่านการวินโดว์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

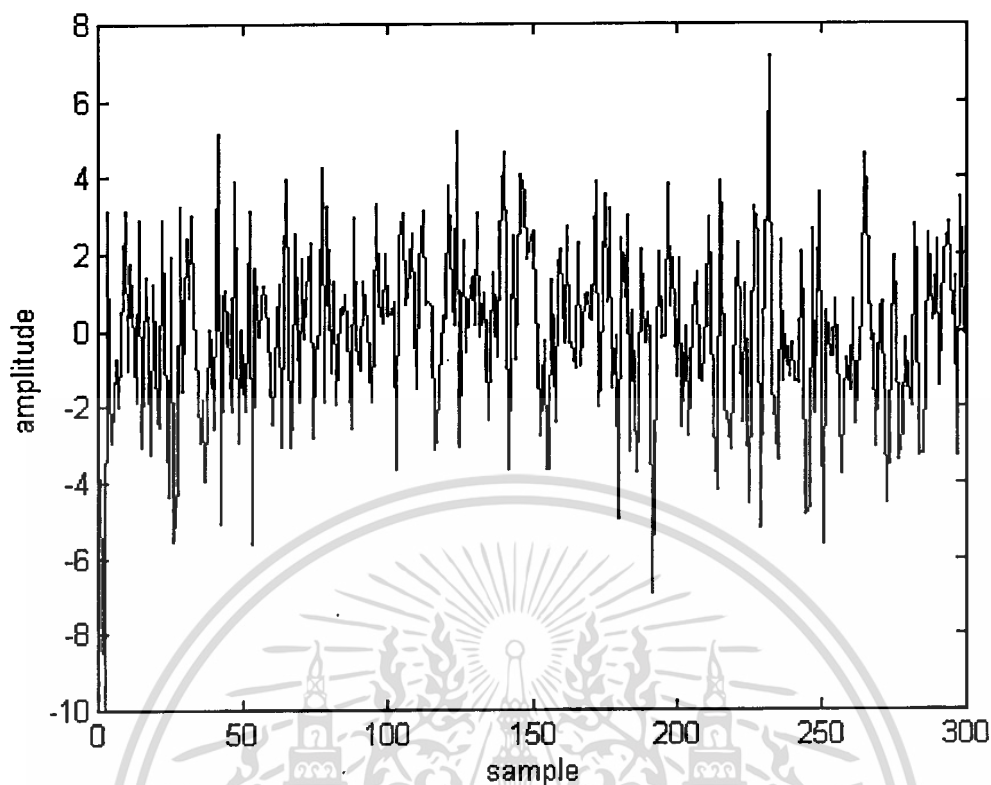


รูปที่ 2.1(ค) สเปกตรัมของเสียง

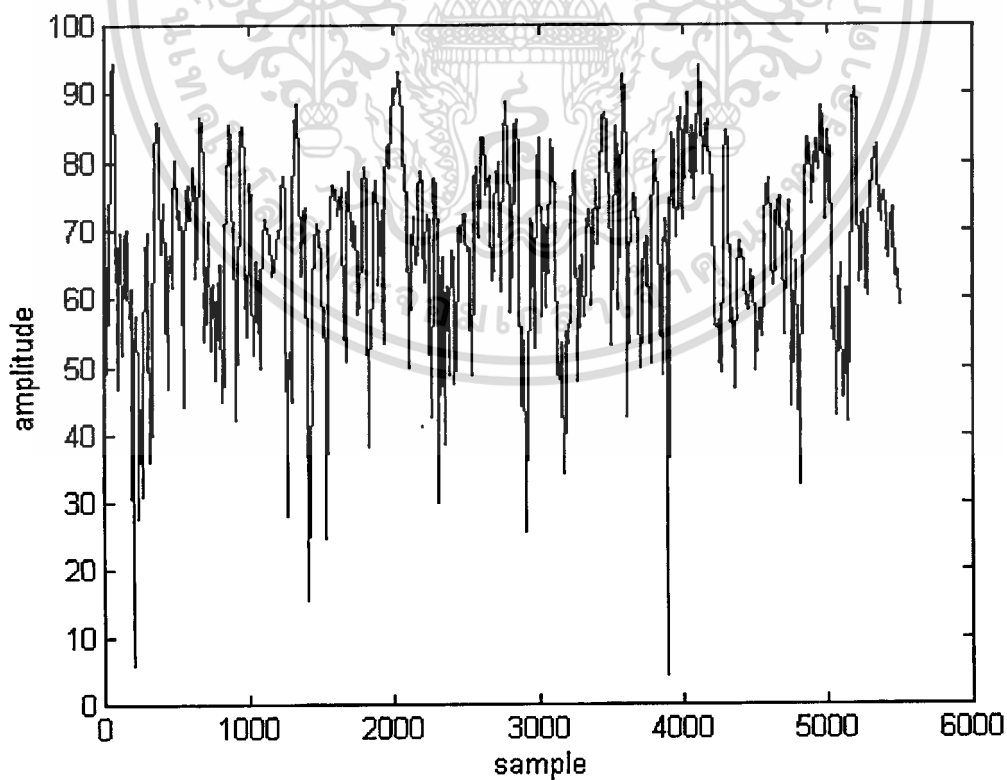


รูปที่ 2.1(ง) สเปกตรัมที่ได้จากสัมประสิทธิ์ของวงจรกรอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

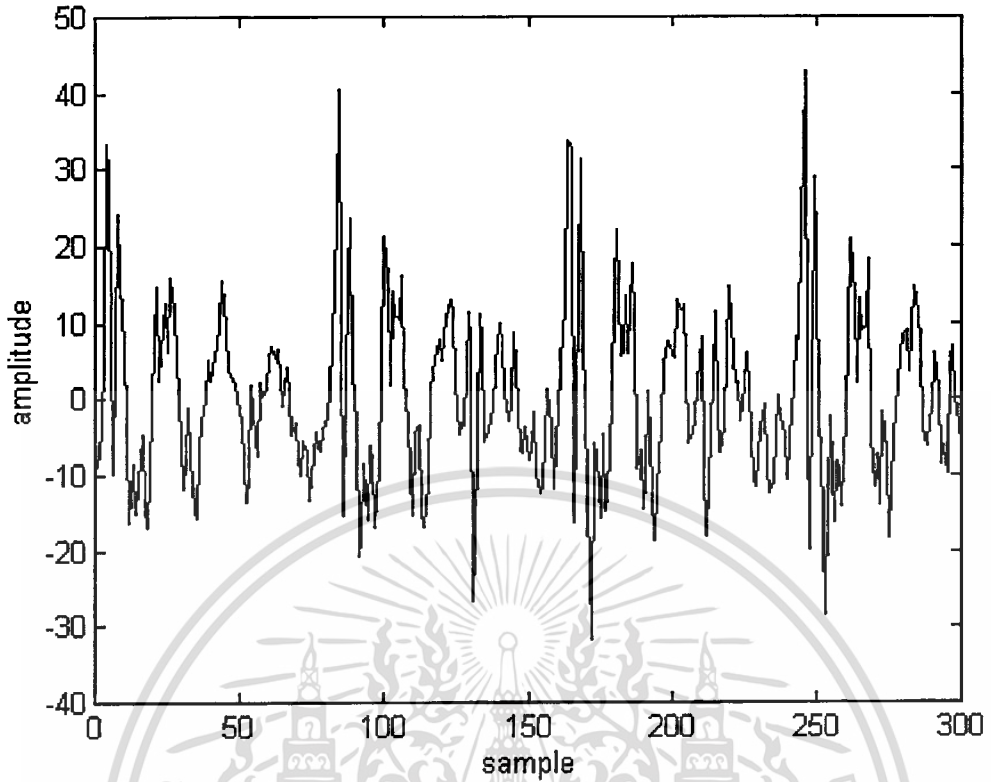


รูปที่ ๓2.1(จ) สัญญาณความคลาดเคลื่อน

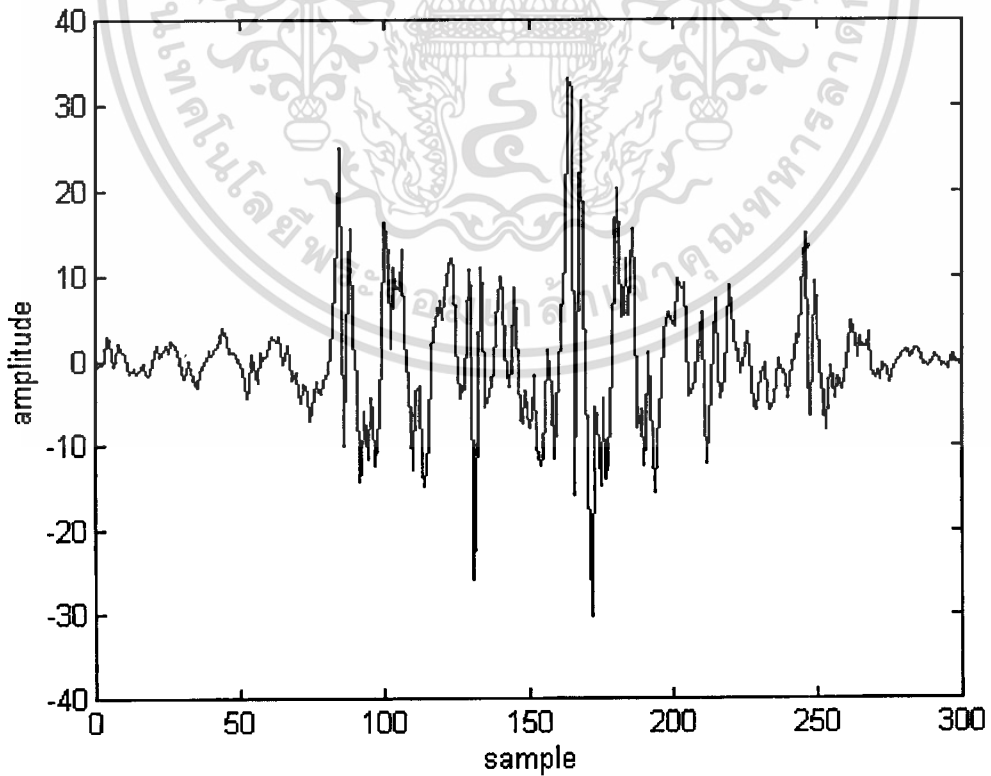


รูปที่ ๓2.1(ข) สเปกตรัมของสัญญาณความคลาดเคลื่อน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

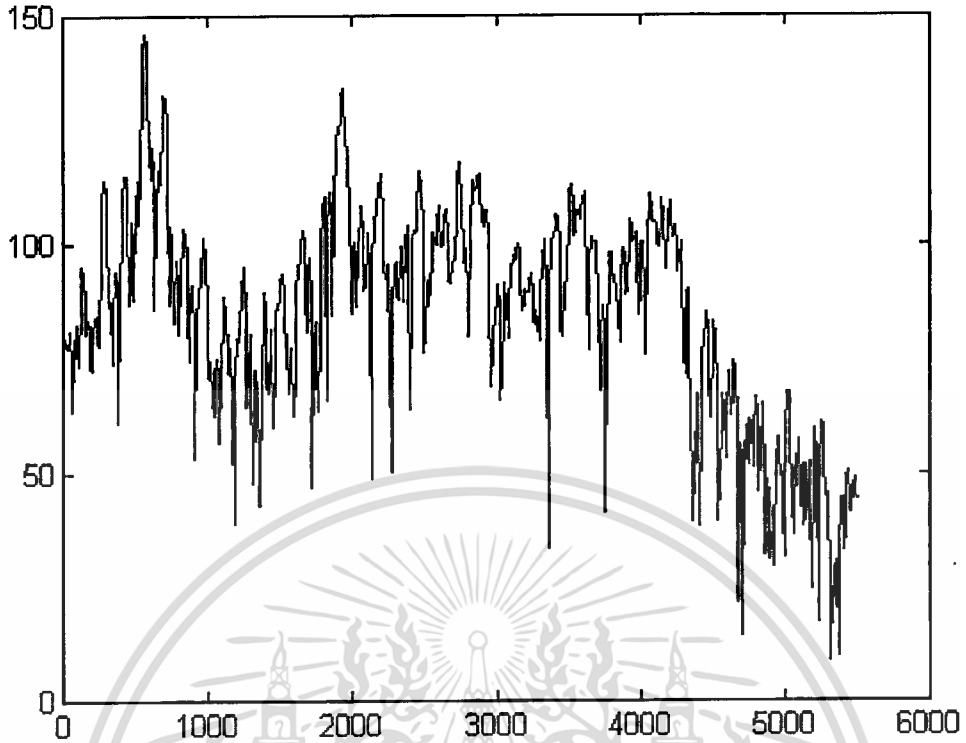


รูปที่ ๓.๒.๒(ก) รูปคลื่นของเสียงว้อยซ์ของผู้ชาย

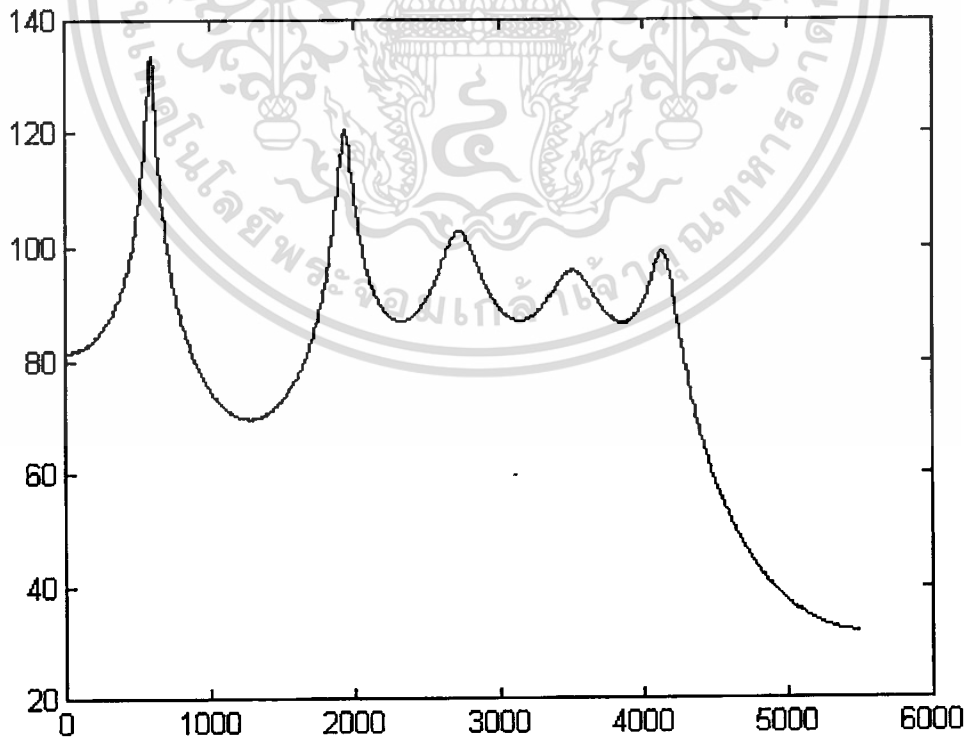


รูปที่ ๓.๒.๒(ข) รูปคลื่นที่ผ่านการวินโดว์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

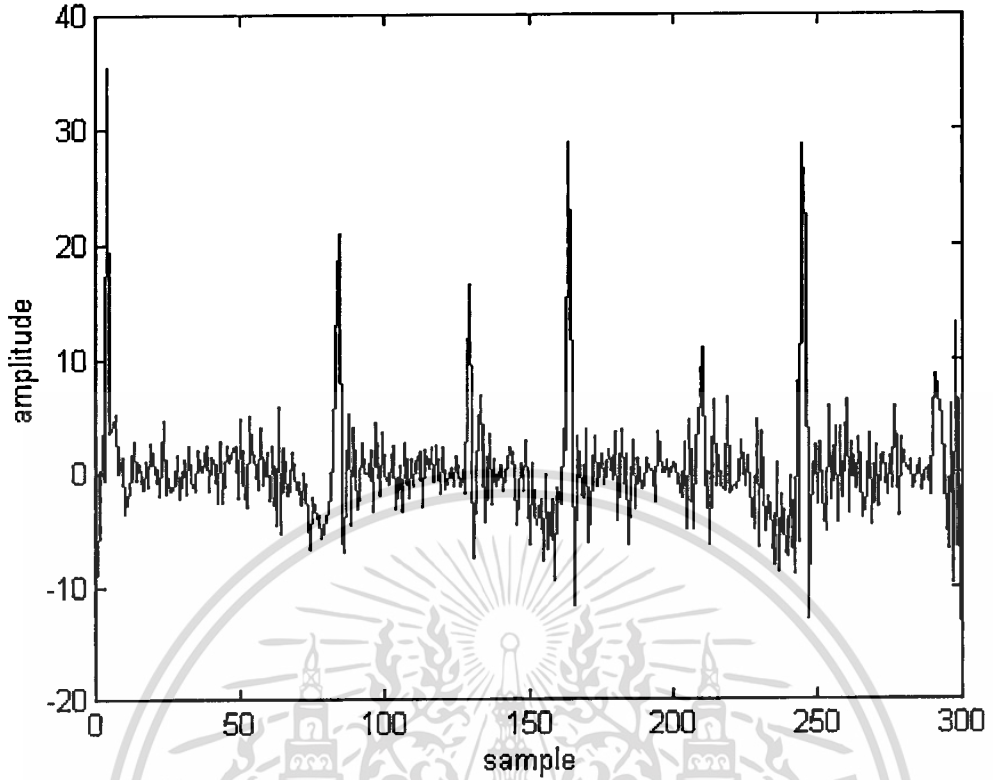


รูปที่ ๓.๒.๒(ค) สเปกตรัมของเสียง

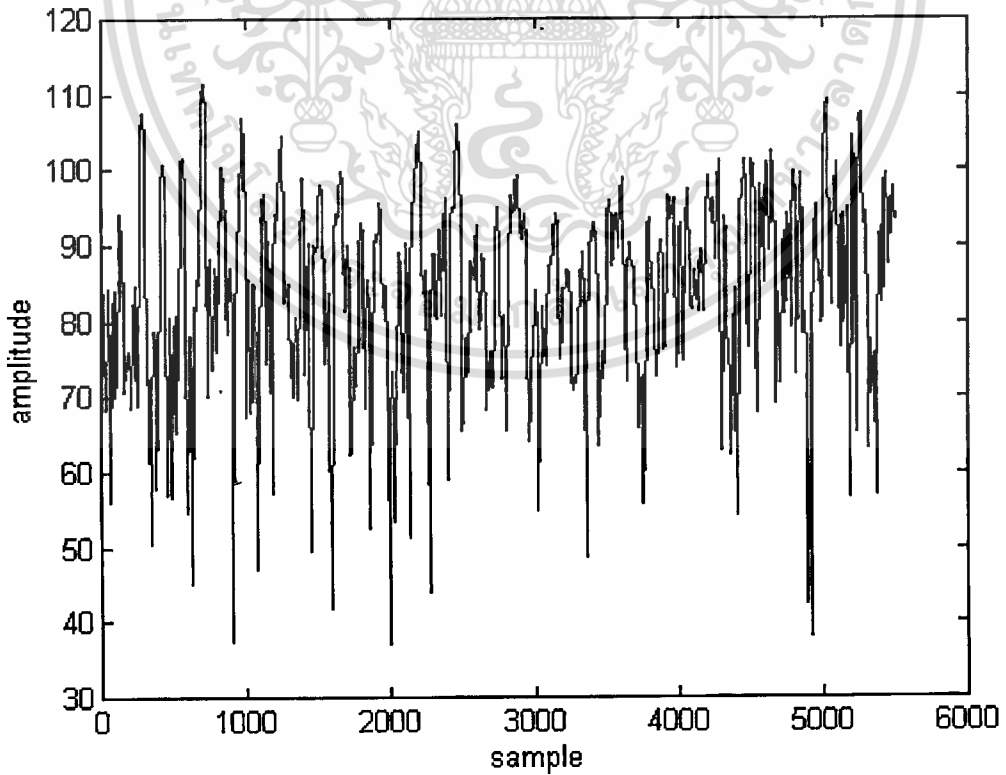


รูปที่ ๓.๒.๒(ง) สเปกตรัมที่ได้จากสัมประสิทธิ์ของวงจรรอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

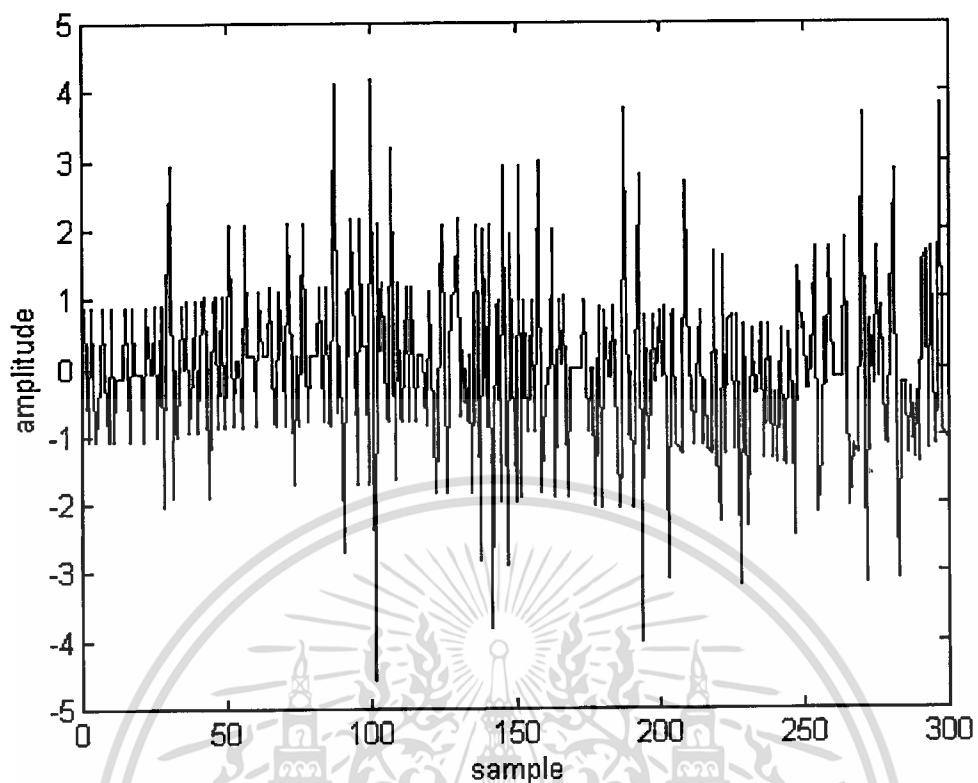


รูปที่ 2.2(จ) สัญญาณความคลาดเคลื่อน

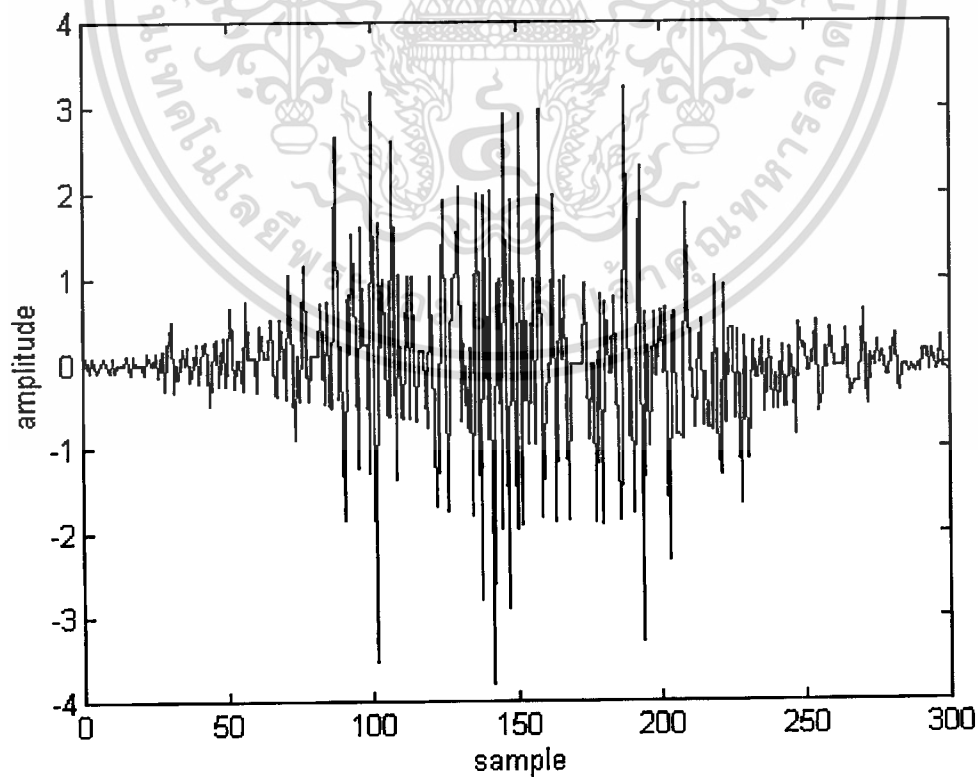


รูปที่ 2.2(ข) สเปกตรัมของสัญญาณความคลาดเคลื่อน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

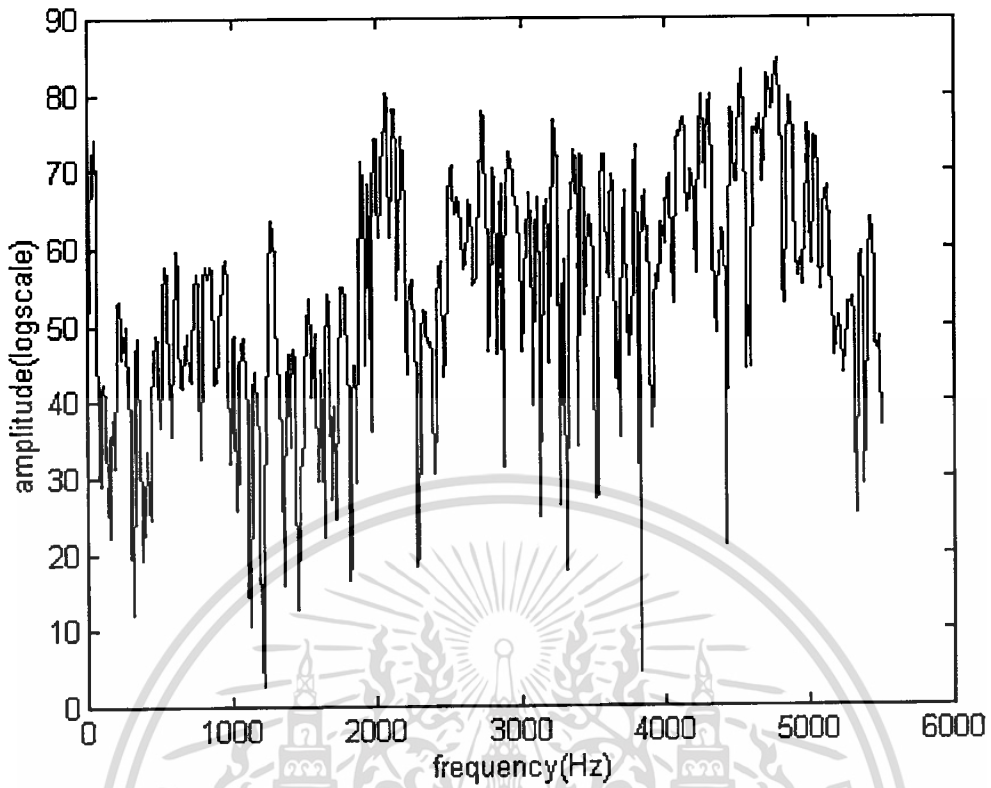


รูปที่ ๓.๓(ก) รูปคลื่นของเสียงอันวอยซ์ของผู้หญิง

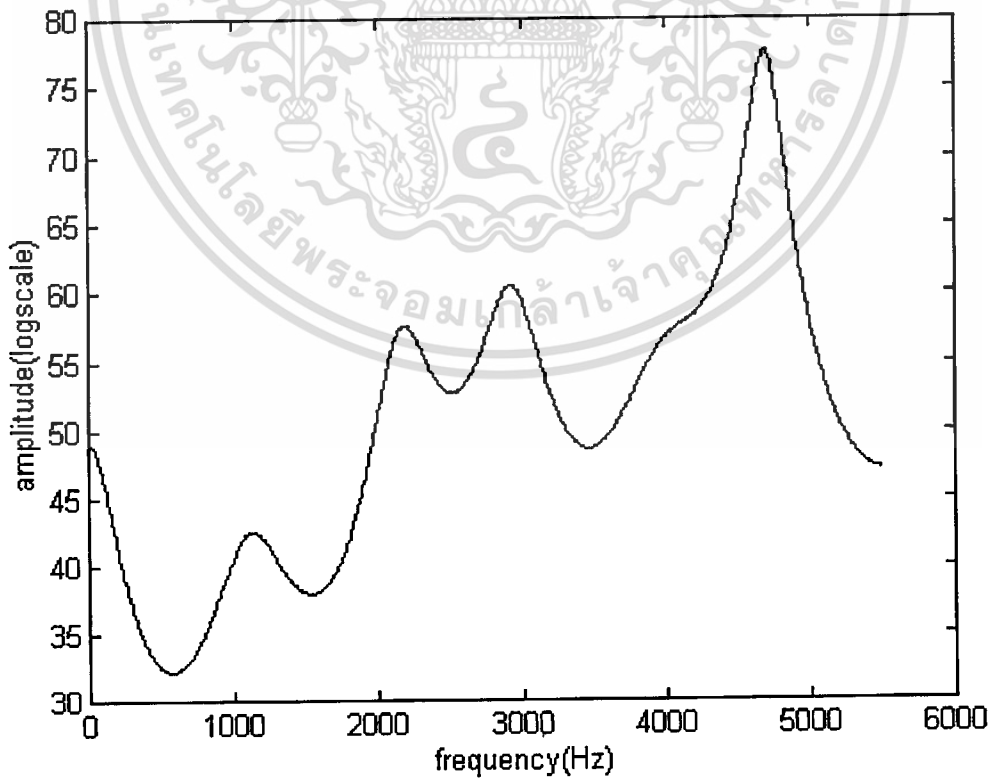


รูปที่ ๓.๓(ข) รูปคลื่นที่ผ่านการวินโดว์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

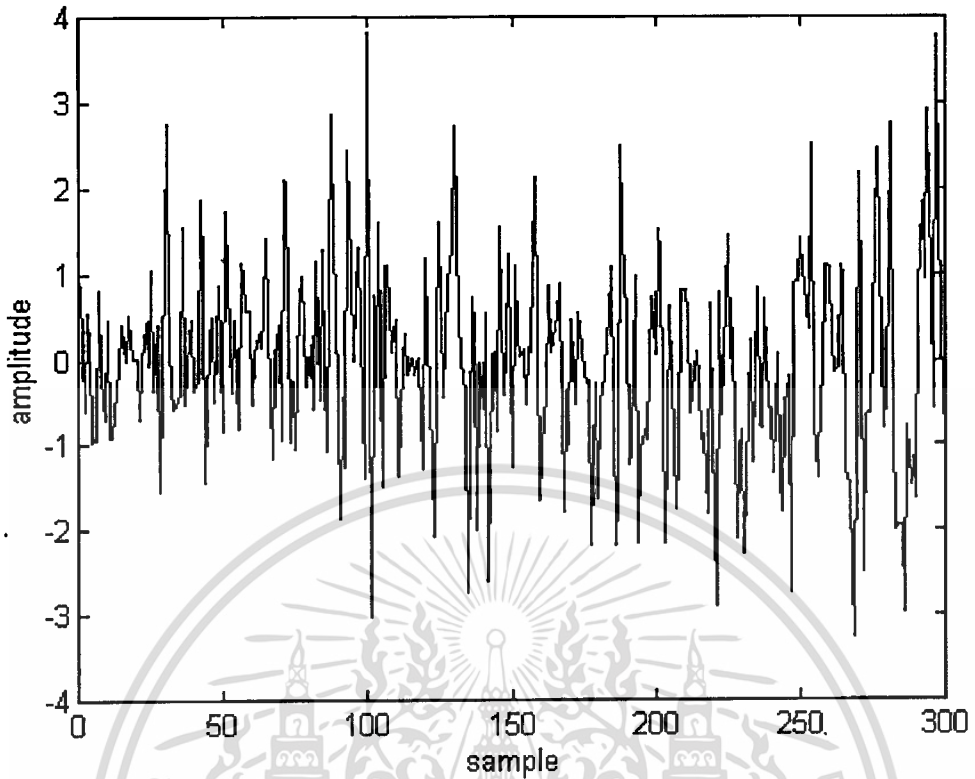


รูปที่ ๓.๓(ค) สเปกตรัมของเสียง

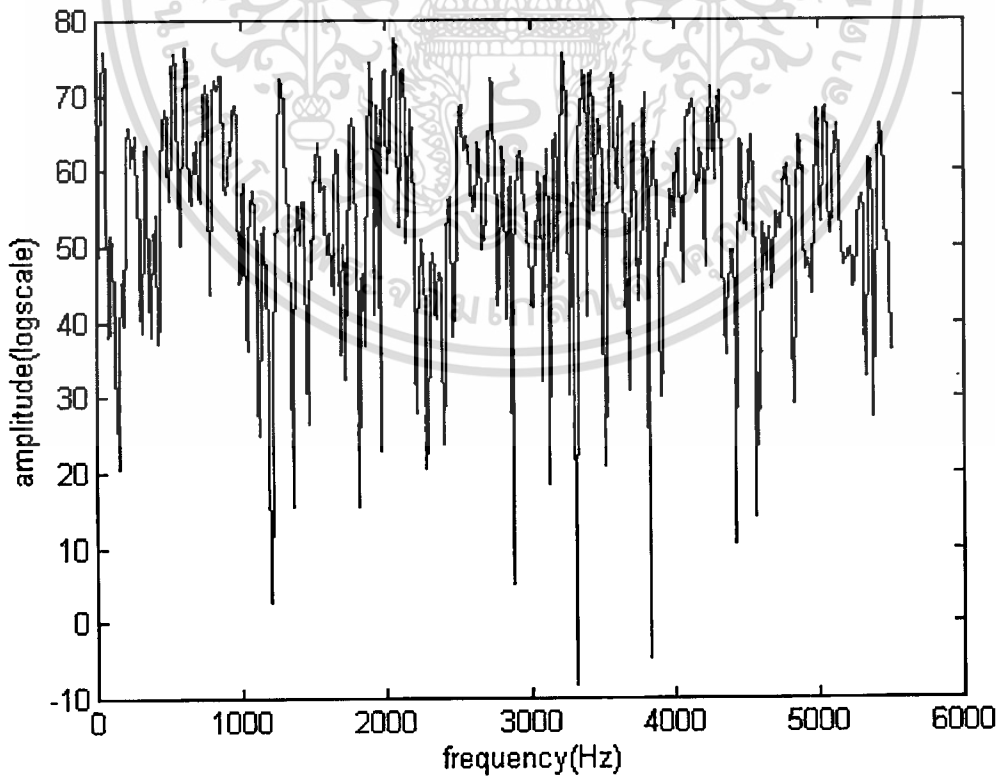


รูปที่ ๓.๓(ง) สเปกตรัมที่ได้จากสัมประสิทธิ์ของวงจกรอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

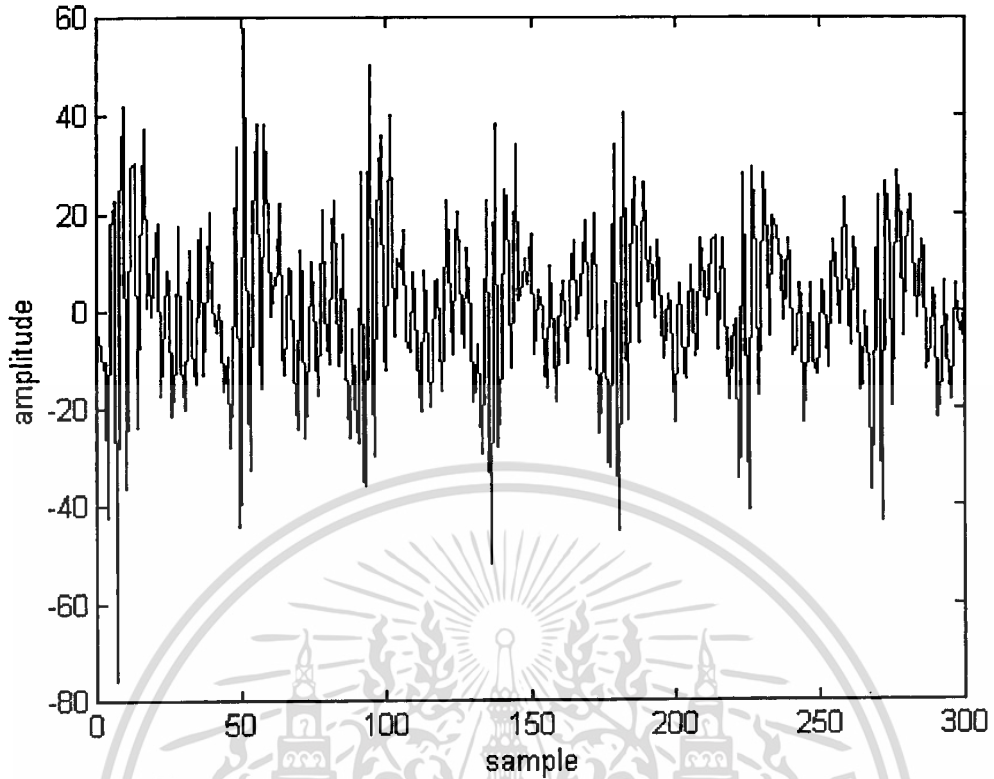


รูปที่ 2.3(จ) สัญญาณความคลาดเคลื่อน

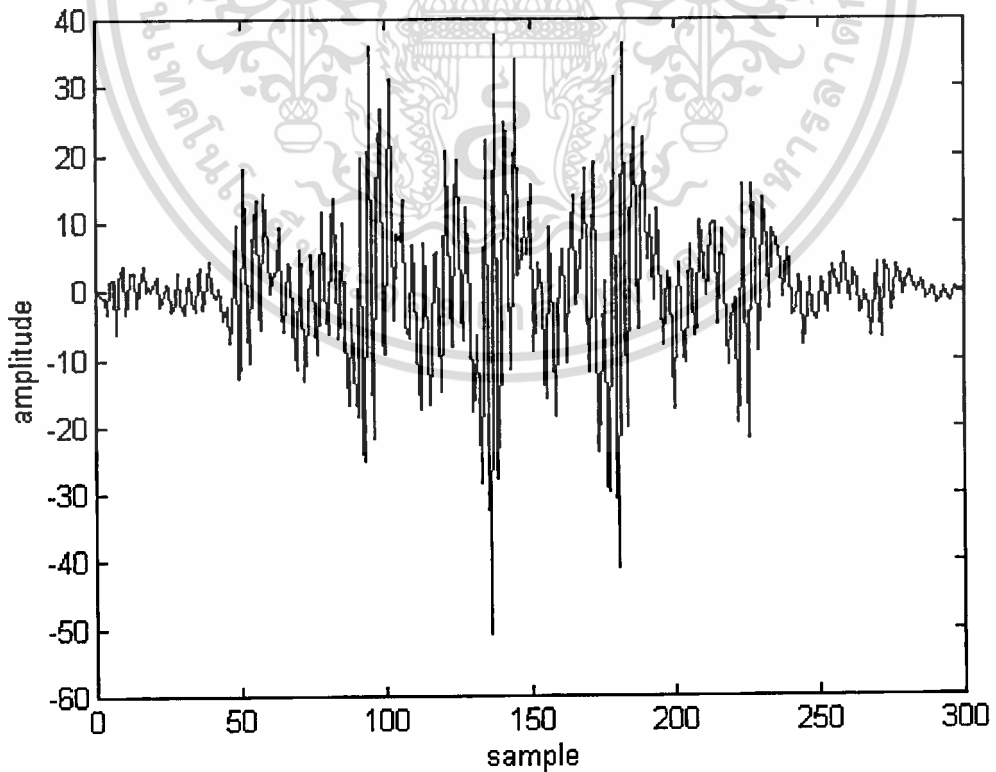


รูปที่ 2.3(ข) สเปกตรัมของสัญญาณความคลาดเคลื่อน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

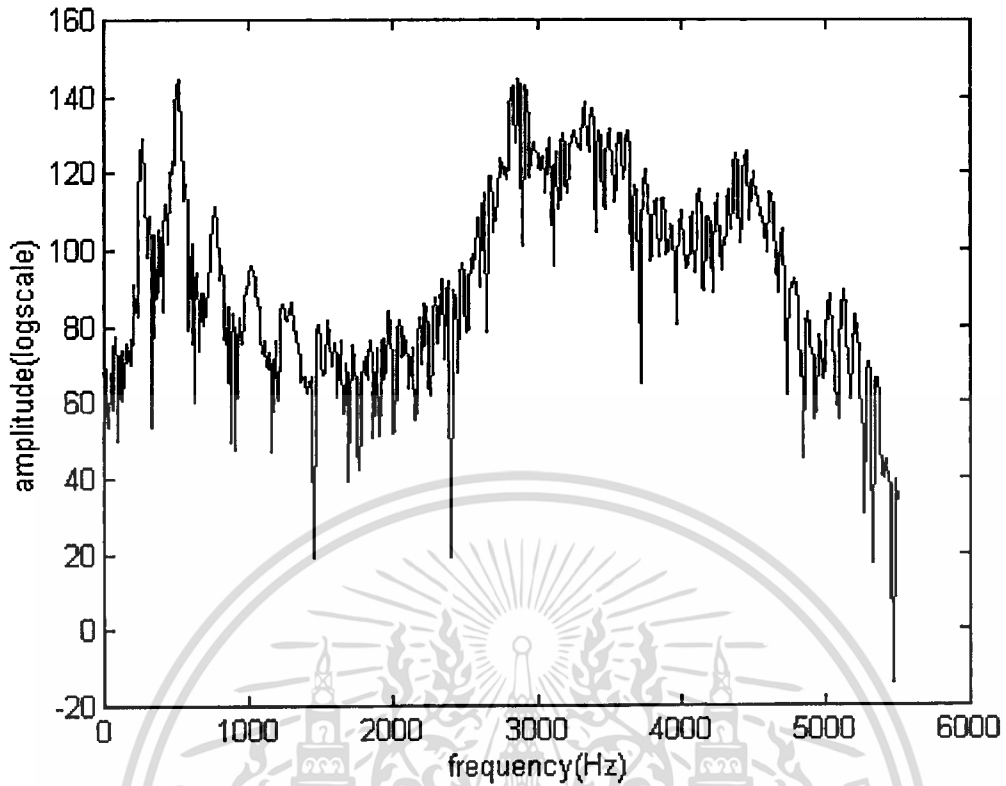


รูปที่ ท2.4(ก) รูปคลื่นของเสียงวอยซ์ของผู้หญิง

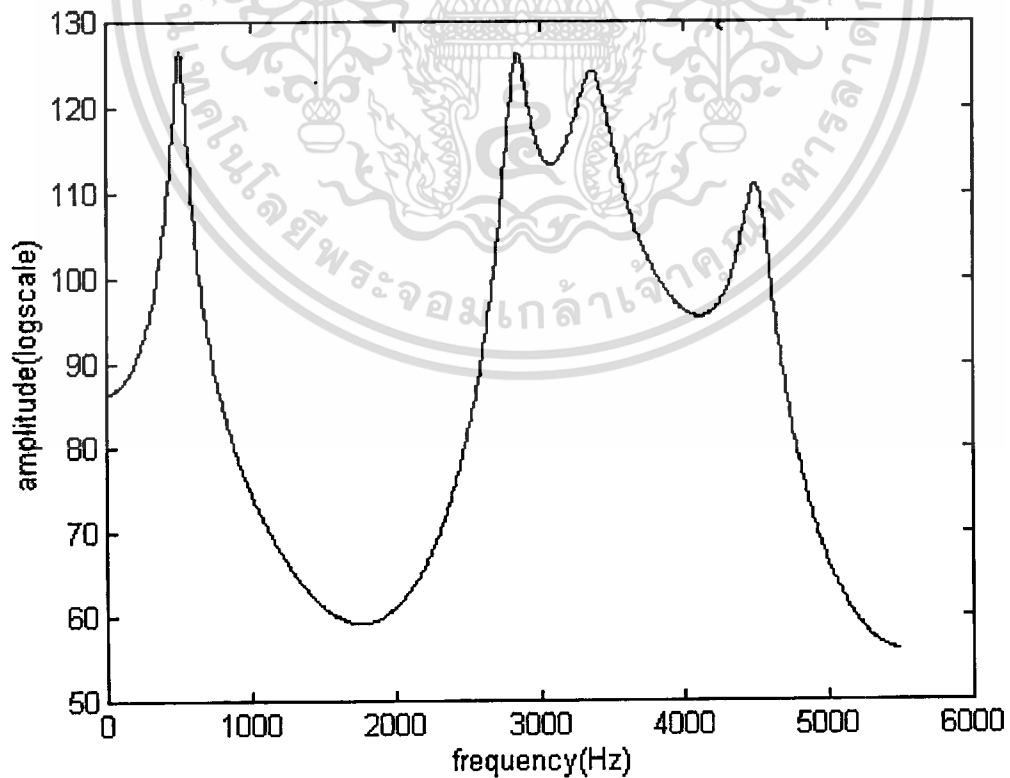


รูปที่ ท2.4(ข) รูปคลื่นที่ผ่านการวินโดว์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

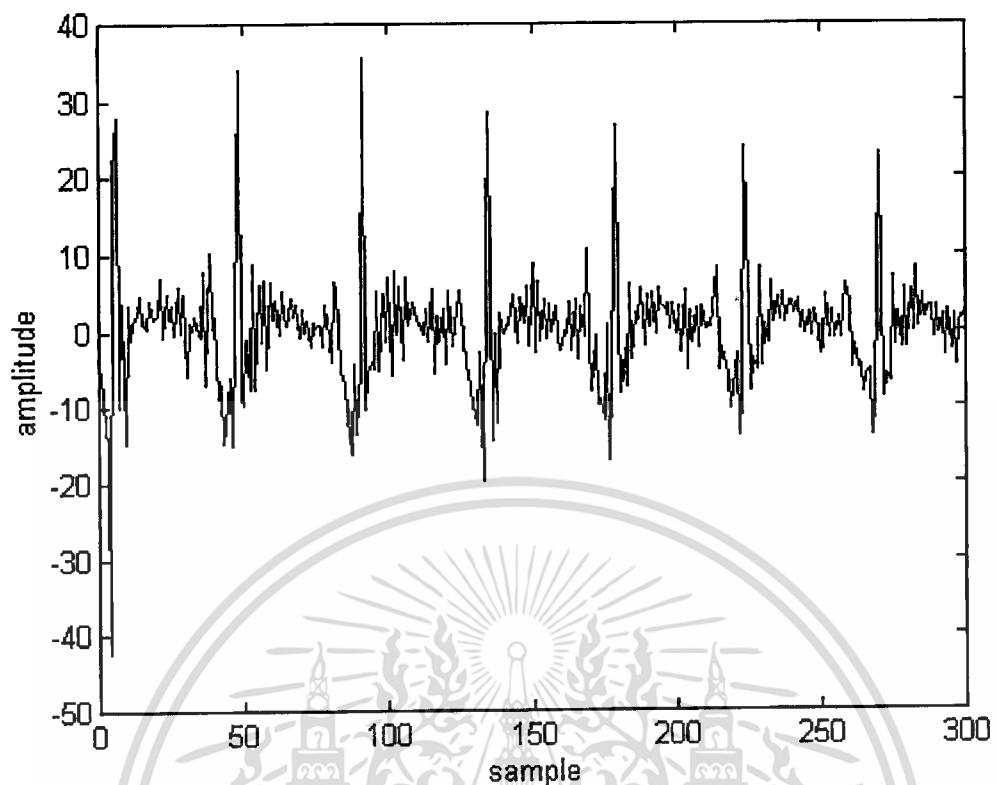


รูปที่ ๓2.4(ค) สเปกตรัมของเสียง

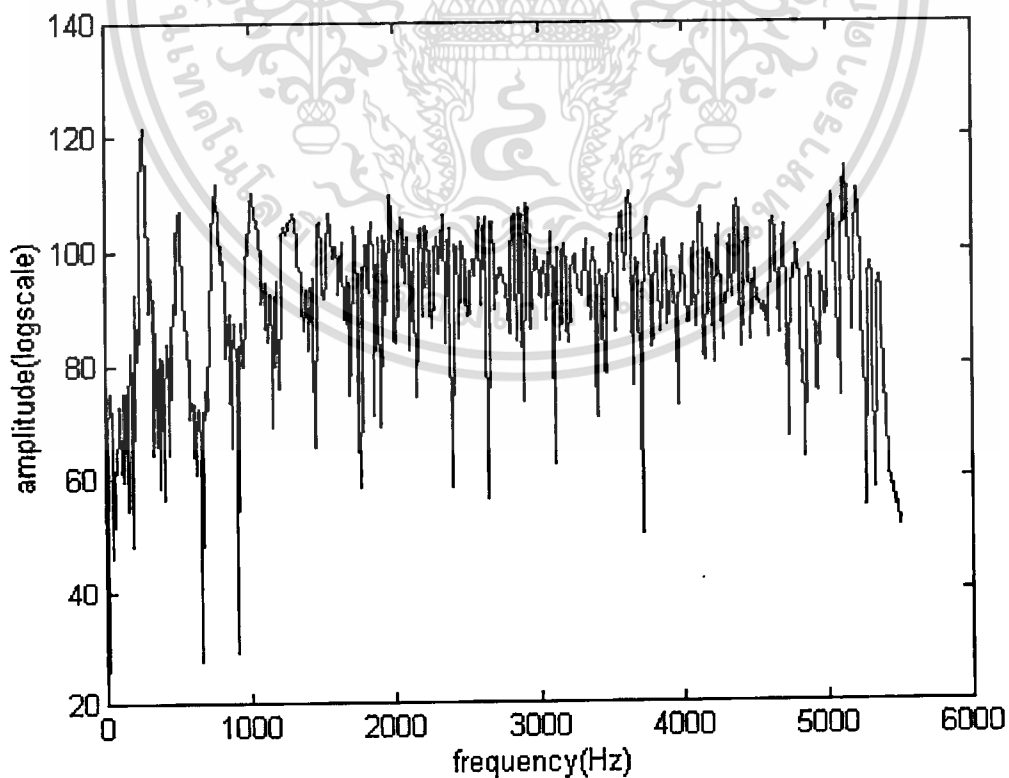


รูปที่ ๓2.4(ง) สเปกตรัมที่ได้จากสัมประสิทธิ์ของวงจรรอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.4(จ) สัญญาณความคลาดเคลื่อน



รูปที่ 2.4(ข) สเปกตรัมของสัญญาณความคลาดเคลื่อน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เปรียบเทียบสเปกตรัมเมื่อเปลี่ยนค่า p

ทดลองเปลี่ยนค่า p เป็นค่า 4, 8, 12, 16 และ 20 แล้วเปรียบเทียบสเปกตรัม และสัญญาณความคลาดเคลื่อน เสียงที่ใช้ในที่นี้คือ เสียงวอยซ์และเสียงอันวอยซ์ของผู้ชายและผู้หญิง ดังรูปที่ ท2.5-ท2.12

รูปที่ ท2.5 แสดงสเปกตรัมของเสียงอันวอยซ์ของผู้ชาย โดยรูป (ก), (ข), (ค), (ง) และ (จ) แทน $p = 4, 8, 12, 16$ และ 20 ตามลำดับ

รูปที่ ท2.6 แสดงสัญญาณความผิดพลาดในการประมาณของเสียงอันวอยซ์ของผู้ชาย โดยรูป (ก), (ข), (ค), (ง) และ (จ) แทน $p = 4, 8, 12, 16$ และ 20 ตามลำดับ

รูปที่ ท2.7 แสดงสเปกตรัมของเสียงวอยซ์ของผู้ชาย โดยรูป (ก), (ข), (ค), (ง) และ (จ) แทน $p = 4, 8, 12, 16$ และ 20 ตามลำดับ

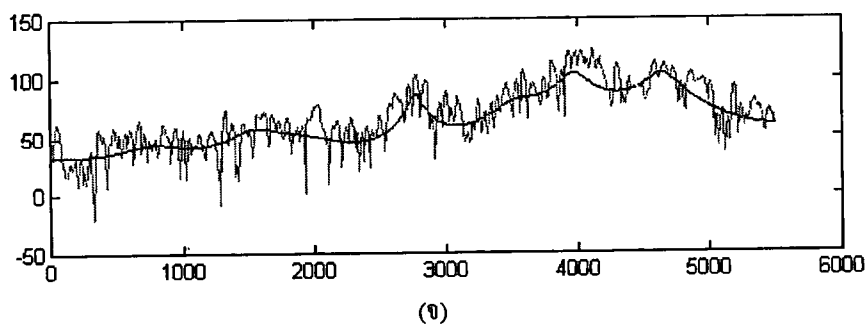
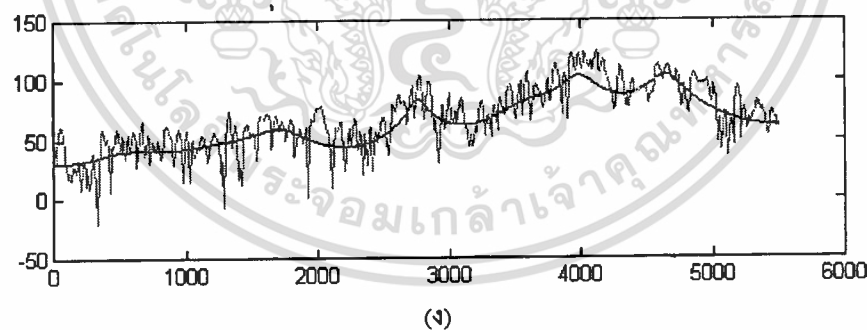
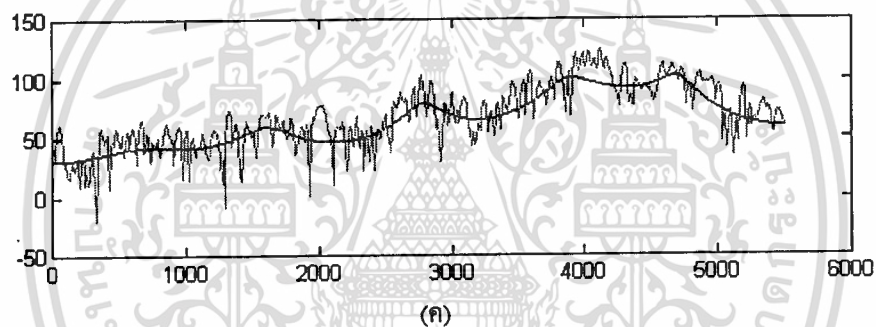
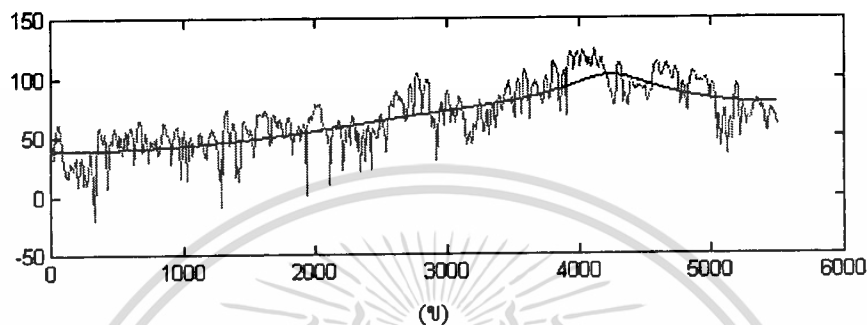
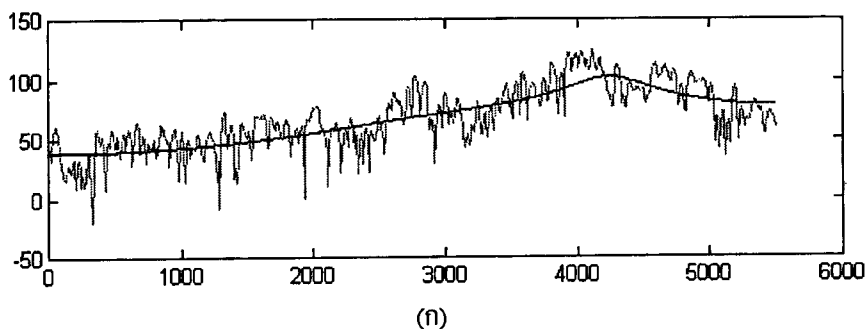
รูปที่ ท2.8 แสดงสัญญาณความผิดพลาดในการประมาณของเสียงวอยซ์ของผู้ชาย โดยรูป (ก), (ข), (ค), (ง) และ (จ) แทน $p = 4, 8, 12, 16$ และ 20 ตามลำดับ

รูปที่ ท2.9 แสดงสเปกตรัมของเสียงอันวอยซ์ของผู้หญิง โดยรูป (ก), (ข), (ค), (ง) และ (จ) แทน $p = 4, 8, 12, 16$ และ 20 ตามลำดับ

รูปที่ ท2.10 แสดงสัญญาณความผิดพลาดในการประมาณของเสียงอันวอยซ์ของผู้หญิง โดยรูป (ก), (ข), (ค), (ง) และ (จ) แทน $p = 4, 8, 12, 16$ และ 20 ตามลำดับ

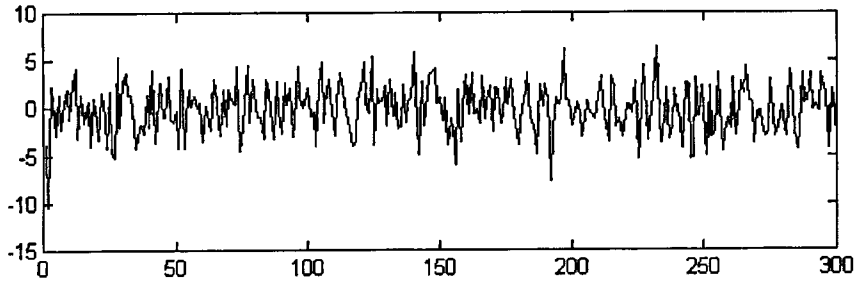
รูปที่ ท2.11 แสดงสเปกตรัมของเสียงวอยซ์ของผู้หญิง โดยรูป (ก), (ข), (ค), (ง) และ (จ) แทน $p = 4, 8, 12, 16$ และ 20 ตามลำดับ

รูปที่ ท2.12 แสดงสัญญาณความผิดพลาดในการประมาณของเสียงวอยซ์ของผู้หญิง โดยรูป (ก), (ข), (ค), (ง) และ (จ) แทน $p = 4, 8, 12, 16$ และ 20 ตามลำดับ

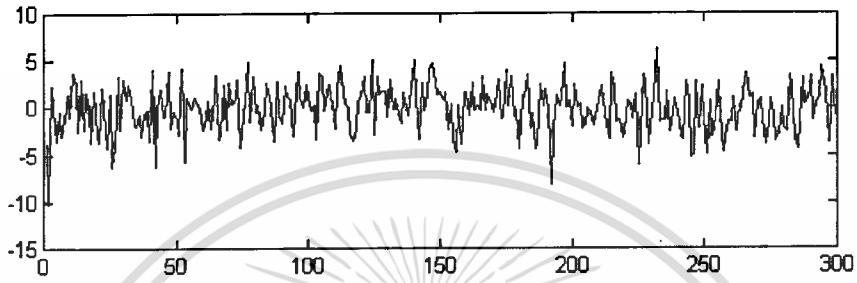


รูปที่ 2.5 สเปกตรัมของเสียงอันว้อซซ์ของผู้ชายเปรียบเทียบกับสเปกตรัมที่ได้จากสัมประสิทธิ์

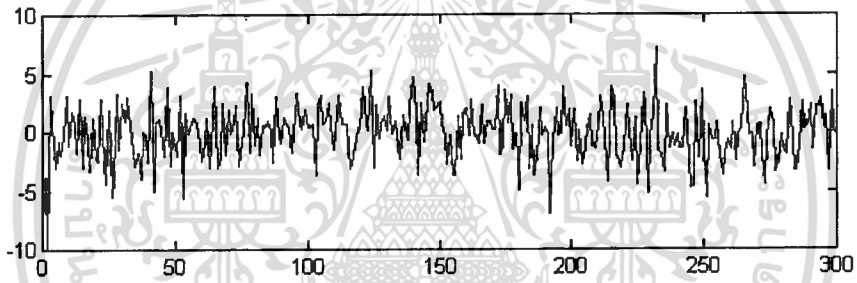
ของวงจรรอง (ก) $p=4$ (ข) $p=8$ (ค) $p=12$ (ง) $p=16$ (จ) $p=20$
 เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ของ บริษัท สยามอินเตอร์คอมมูนิเคชั่น จำกัด (มหาชน) ผู้จัดทำให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



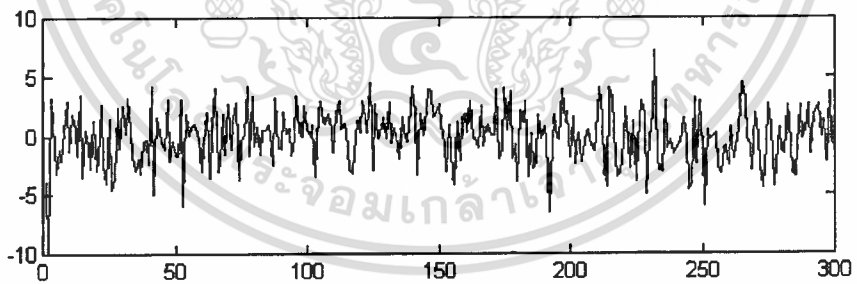
(ก)



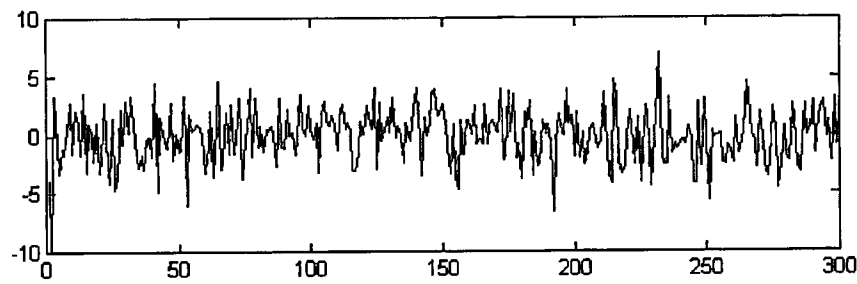
(ข)



(ค)



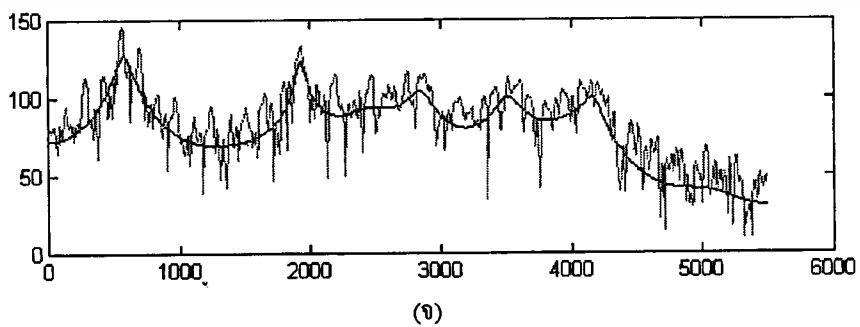
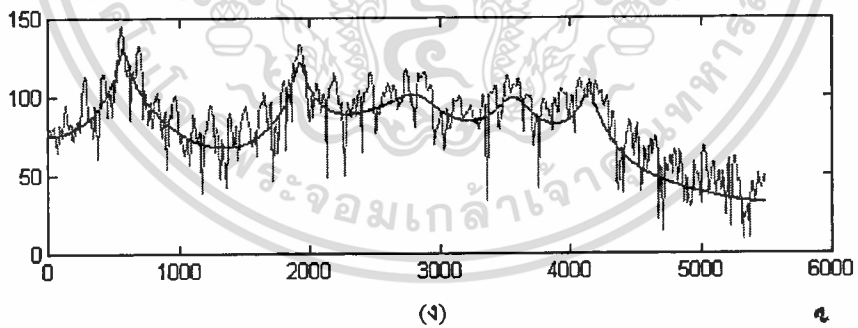
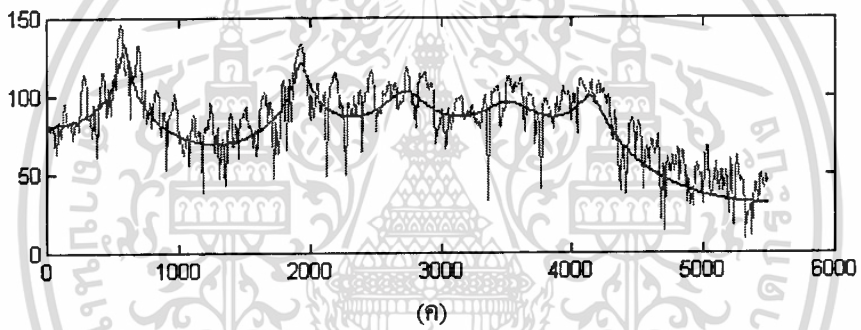
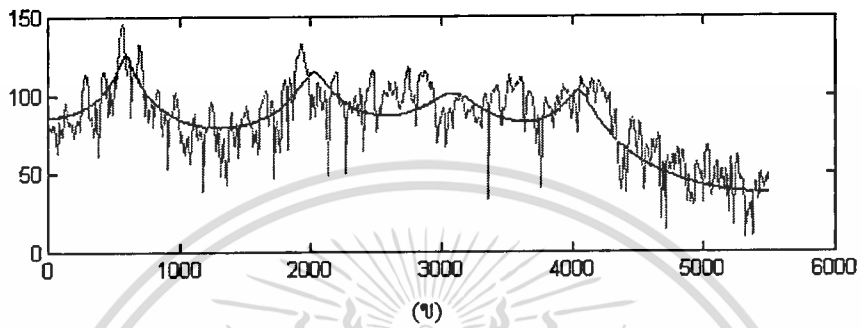
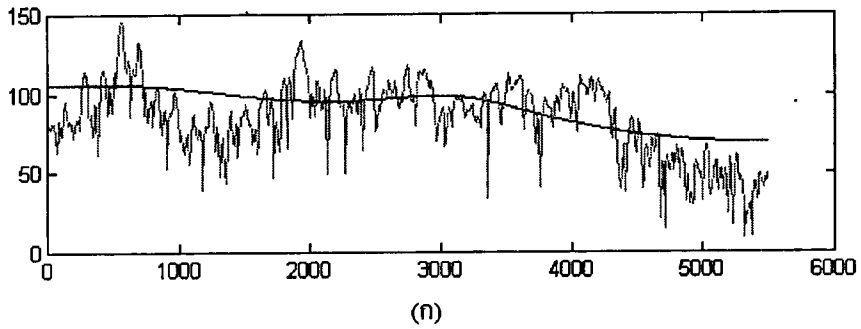
(ง)



(จ)

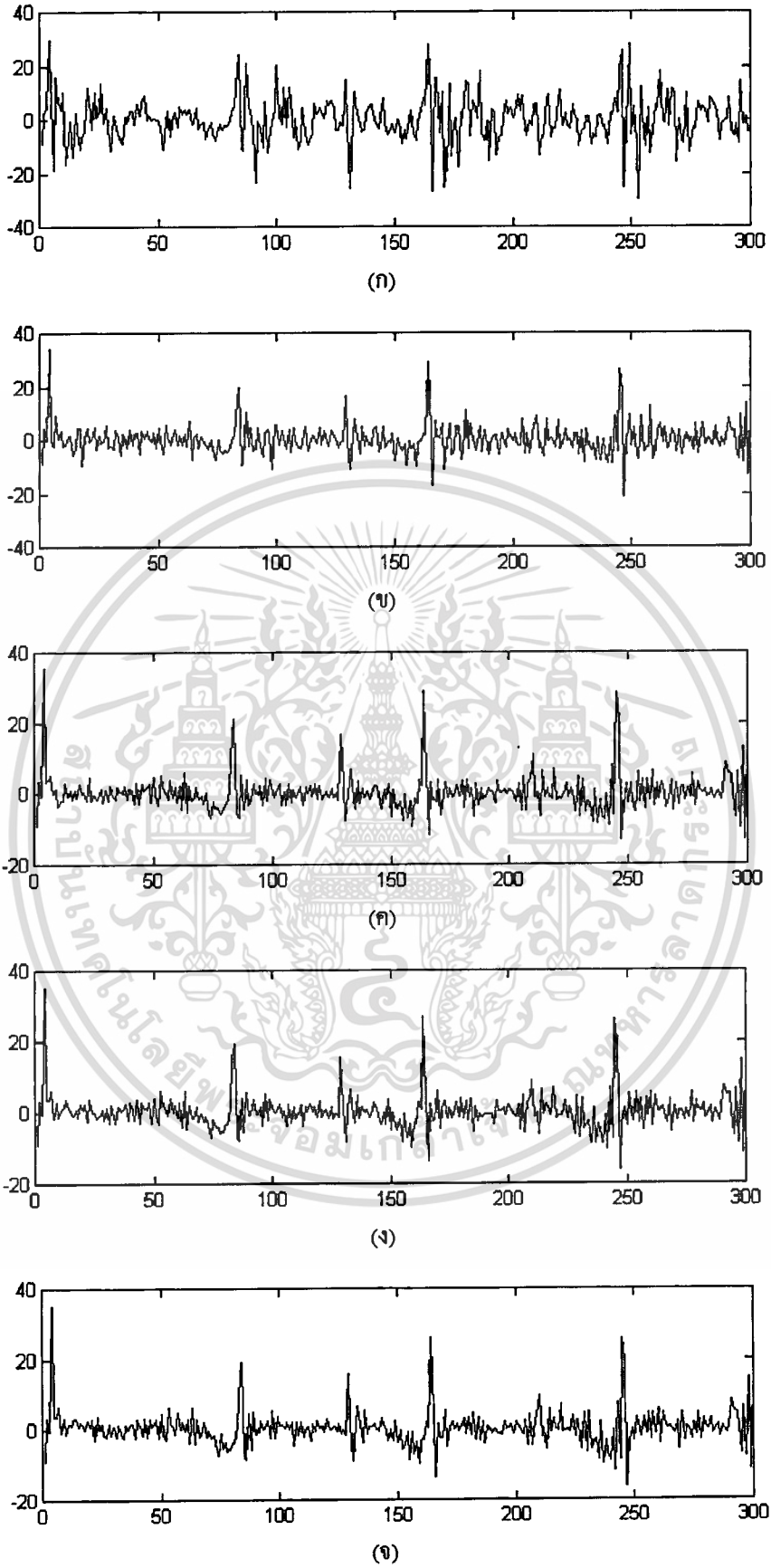
รูปที่ 2.6 สัญญาณความคลาดเคลื่อนของเสียงอันว้อยซ์ผู้ชาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับ (ก) $p=4$ (ข) $p=8$ (ค) $p=12$ (ง) $p=16$ (จ) $p=20$ ให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



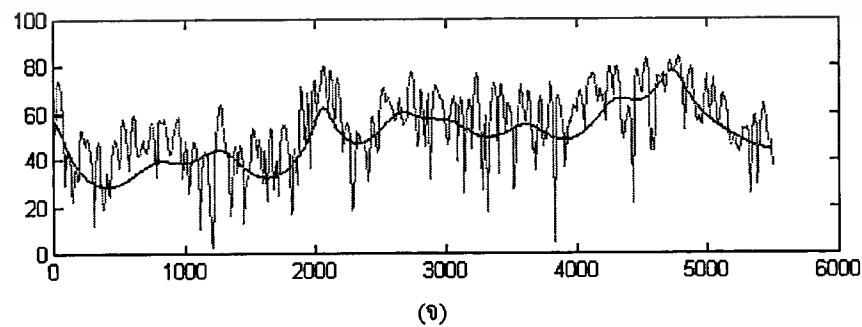
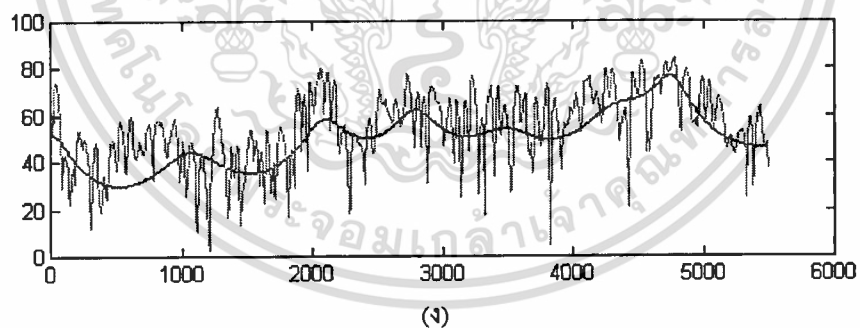
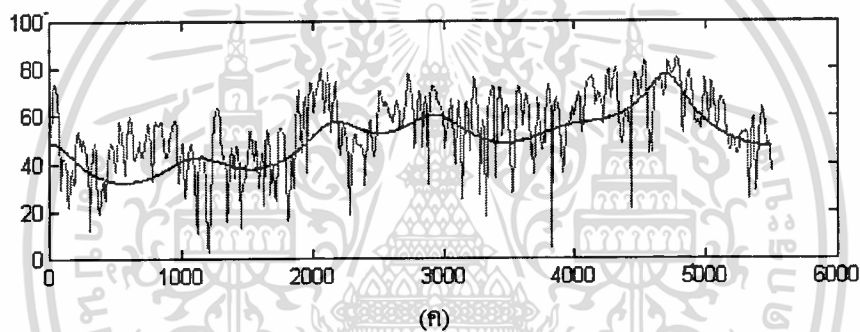
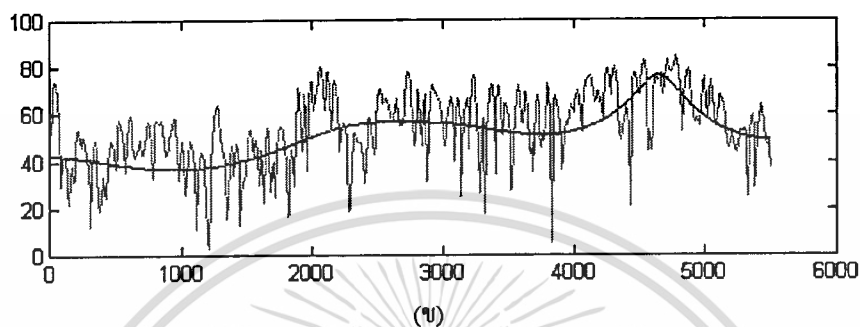
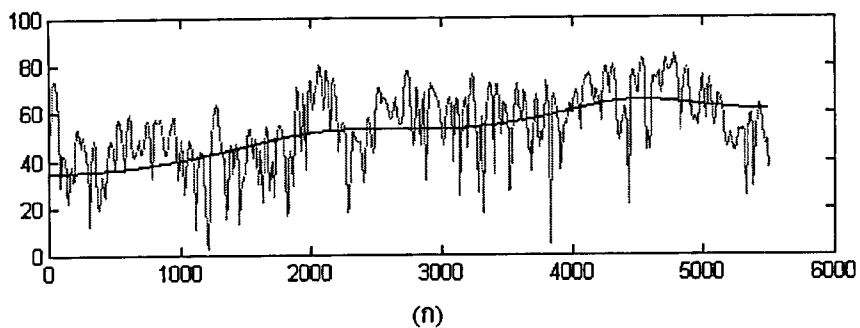
รูปที่ 2.7 สเปกตรัมของเสียงรบกวนของผู้ขายเปรียบเทียบกับสเปกตรัมที่ได้จากสัมประสิทธิ์

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ของโรงเรียนเพื่อใช้ในการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.8 สัญญาณความคลาดเคลื่อนของเสียงวีย์ซซ์ผู้ชาย

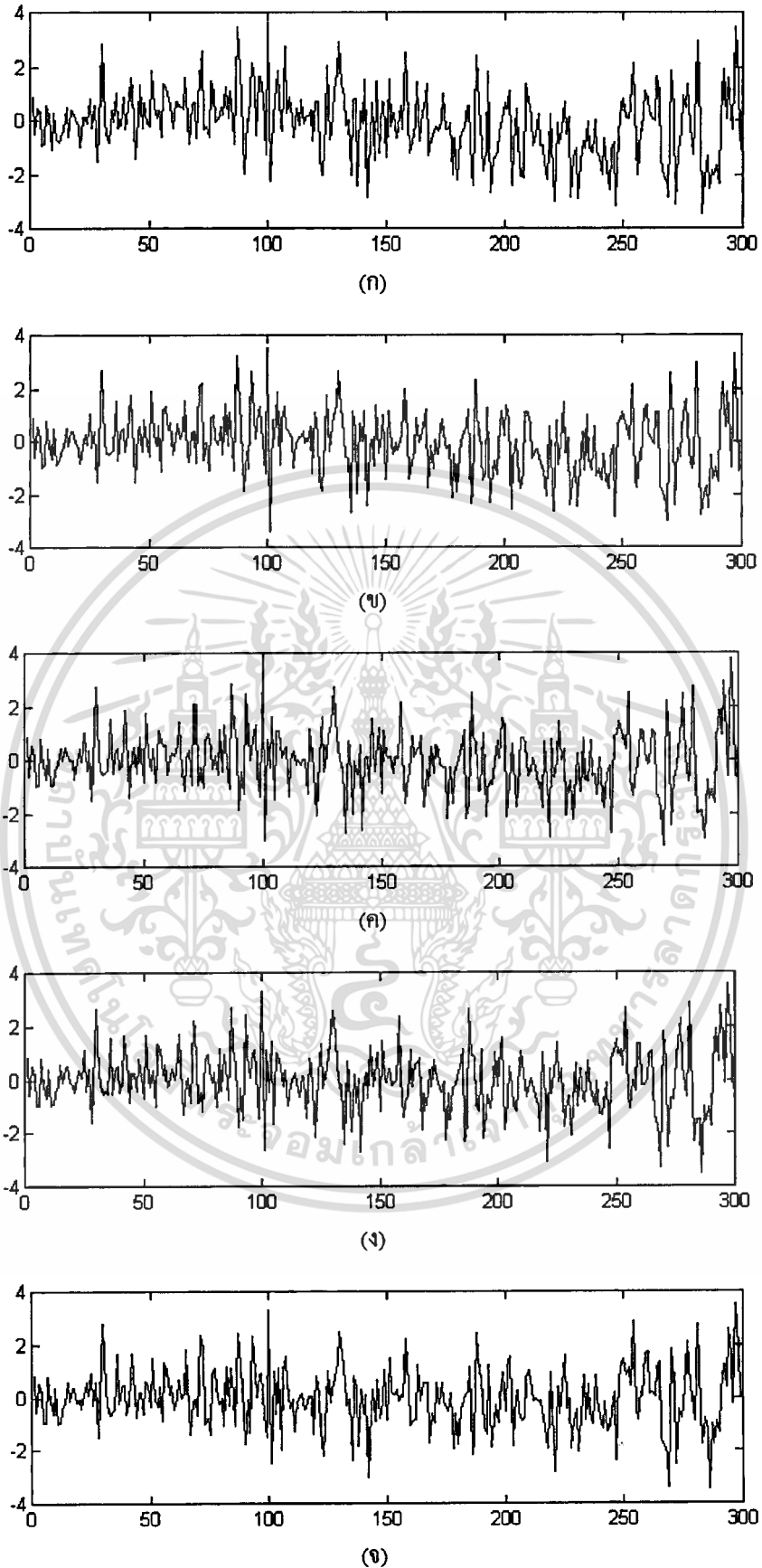
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับก (ก) $p=4$ (ข) $p=8$ (ค) $p=12$ (ง) $p=16$ (จ) $p=20$ นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ท.2.9 สเปกตรัมของเสียงอันวอยซ์ของผู้หญิงเปรียบเทียบกับสเปกตรัมที่ได้จากสัมประสิทธิ์

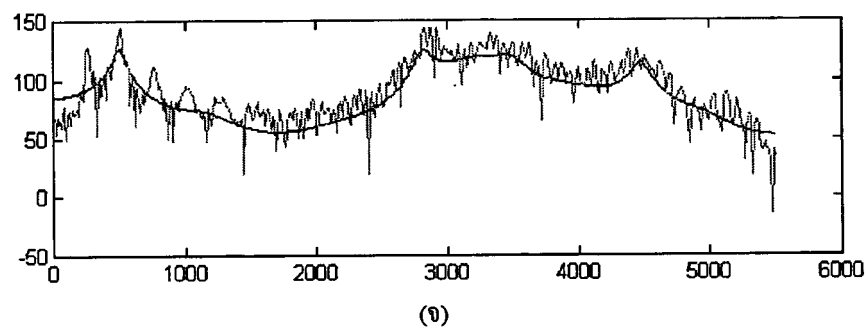
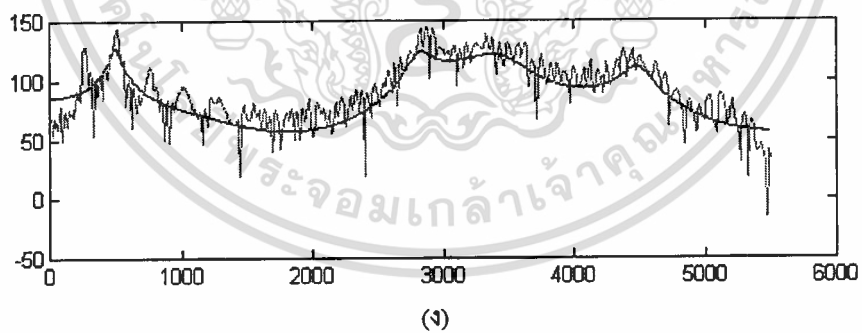
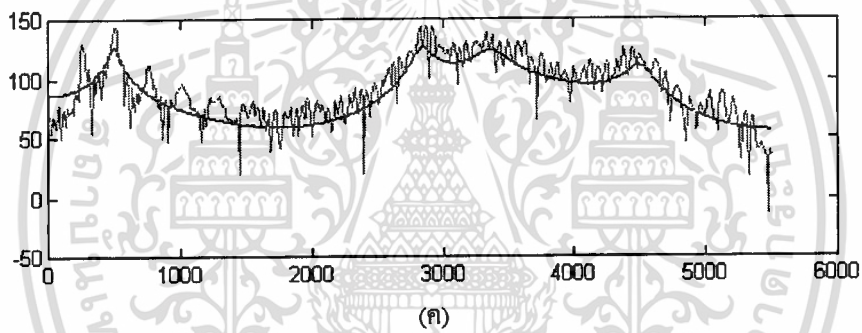
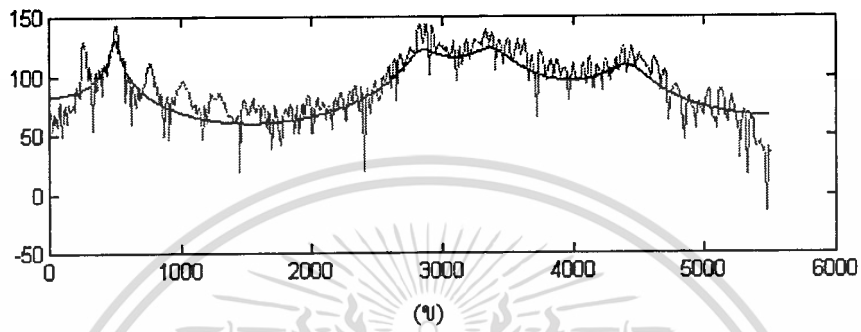
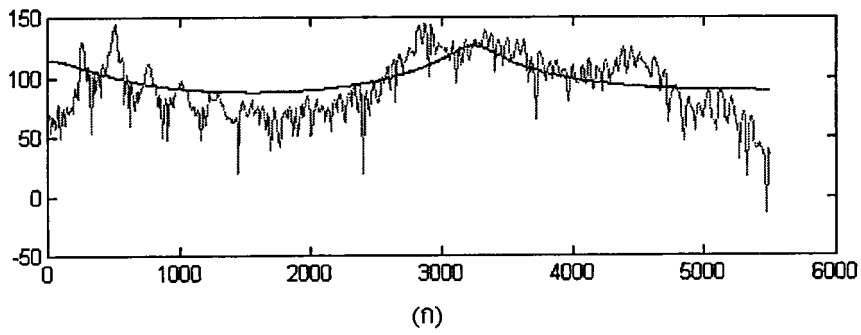
ของวงจรรอง (ก) $p=4$ (ข) $p=8$ (ค) $p=12$ (ง) $p=16$ (จ) $p=20$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



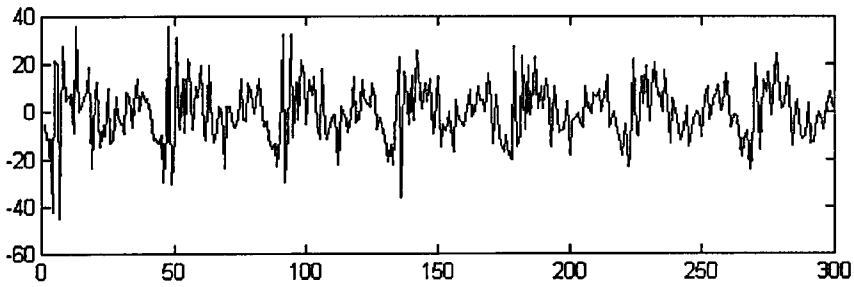
รูปที่ 2.10 สัญญาณความคลาดเคลื่อนของเสียงอันว้อยซ์ผู้หญิง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับ (ก) $p=4$ (ข) $p=8$ (ค) $p=12$ (ง) $p=16$ (จ) $p=20$ ให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

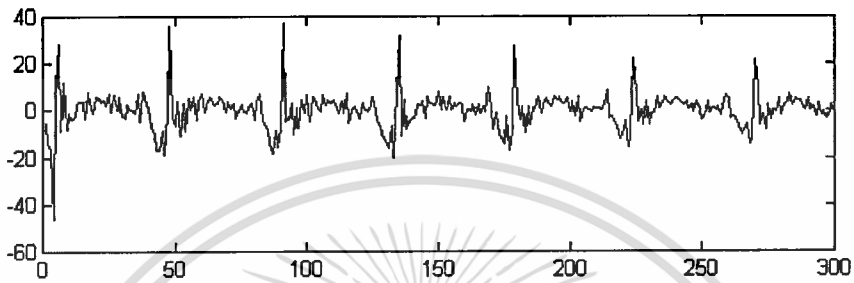


รูปที่ ๓2.11 สเปกตรัมของเสียงวอยซ์ของผู้หญิงเปรียบเทียบกับสเปกตรัมที่ได้จากสัมประสิทธิ์

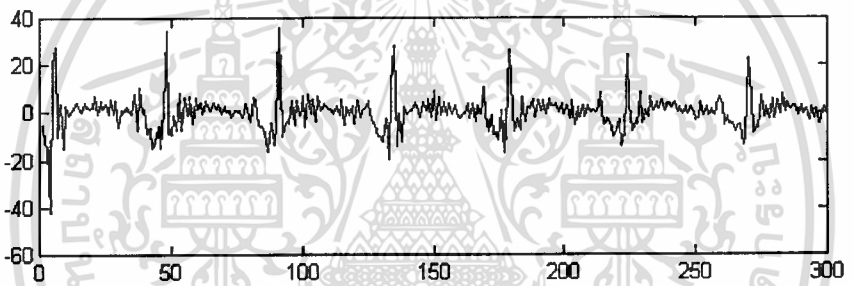
เอกสารนี้เป็นเอกสารที่สงวนของวังจรรย์รอง (ก) $p=4$ (ข) $p=8$ (ค) $p=12$ (ง) $p=16$ (จ) $p=20$ ห้านำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



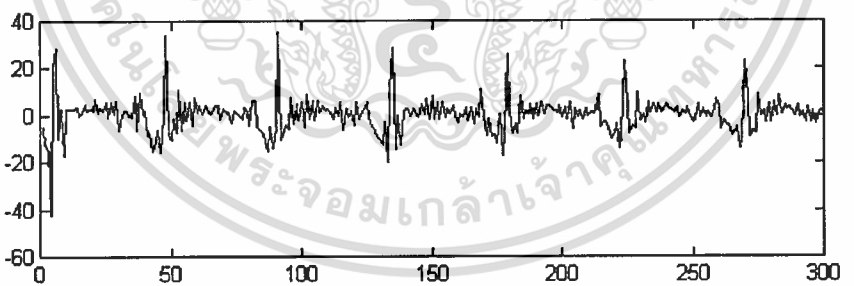
(ก)



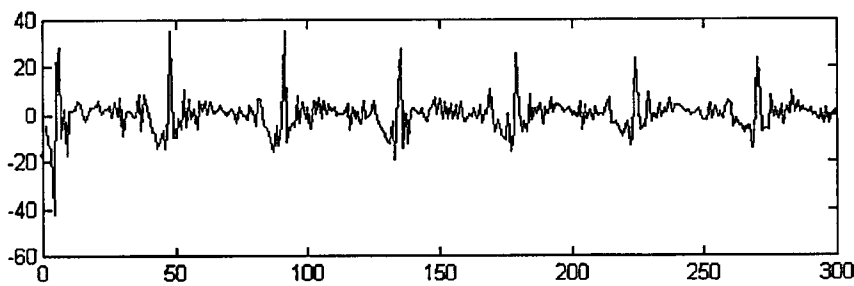
(ข)



(ค)



(ง)



(จ)

รูปที่ ๓.๑๒ สัญญาณความคลาดเคลื่อนของเสียงรบกวนผู้หญิง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกา (ก) $p=4$ (ข) $p=8$ (ค) $p=12$ (ง) $p=16$ (จ) $p=20$ นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทแทรกที่ 3

การออกแบบวงจรกรองเชิงเลข

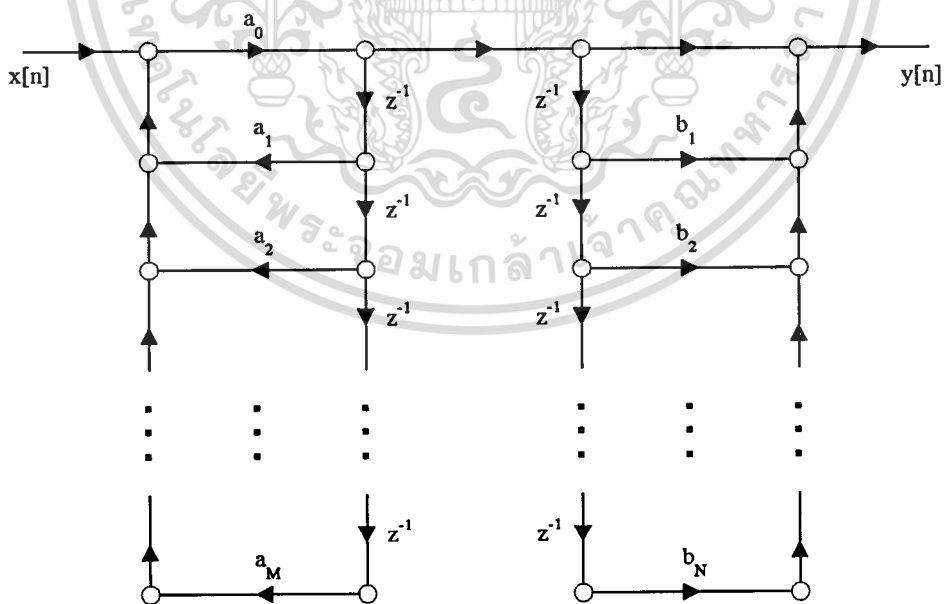
วงจรกรองเชิงเลข (digital filter) จัดเป็นระบบการประมวลสัญญาณเชิงเลขแบบหนึ่งที่เราสามารถแสดงความสัมพันธ์ระหว่างลำดับสัญญาณเข้า $x(n)$ และลำดับสัญญาณออก $y(n)$ ได้ในรูปของสมการผลต่างสลับเนื่อง (difference equation) อันดับที N ดังนี้

$$y(n) = \sum_{j=0}^M a_j x(n-j) - \sum_{k=1}^N b_k y(n-k) \quad (ท3.1)$$

เมื่อ a_j และ b_k เป็นสัมประสิทธิ์ของสมการผลต่างสลับเนื่อง ใช้การแปลง Z (Z-transform) ดังนั้นฟังก์ชันถ่ายโอนของระบบ (transfer function) ซึ่งใช้สัญลักษณ์ $H(z)$ คือ

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{j=0}^M a_j z^{-j}}{1 + \sum_{k=1}^N b_k z^{-k}} \quad (ท3.2)$$

และมีแผนผังการไหลของสัญญาณ (signal flow graph) ดังรูปที่ ท3.1



รูปที่ 3.1 แผนผังการไหลของสัญญาณของวงจรกรองเชิงเลข

ในโครงการนี้ต้องใช้วงจรกรองสองวงจร คือวงจรกรองความถี่ต่ำผ่าน 4 กิโลเฮิร์ตซ์ ใช้กรองสัญญาณเสียงที่จะนำมาวิเคราะห์การประมาณเชิงเส้น และวงจรกรองความถี่ต่ำผ่าน 900 เฮิร์ตซ์ ใช้ในวงจรตรวจจับคาบเวลาพิทช์ วงจรกรองทั้งสองคำนวณได้ดังนี้

ท3.1 วงจรกรองเชิงเลขแบบป้อนกลับชนิดความถี่ต่ำผ่าน (low pass recursive digital filter) อันดับ 4 โดยมีคุณสมบัติทางความถี่ดังนี้

ท3.1.1 ที่ความถี่ 0-4 กิโลเฮิร์ตซ์ มีอัตราขยายเท่ากับ 1 ค่าคลาดเคลื่อนไม่เกิน 0.01

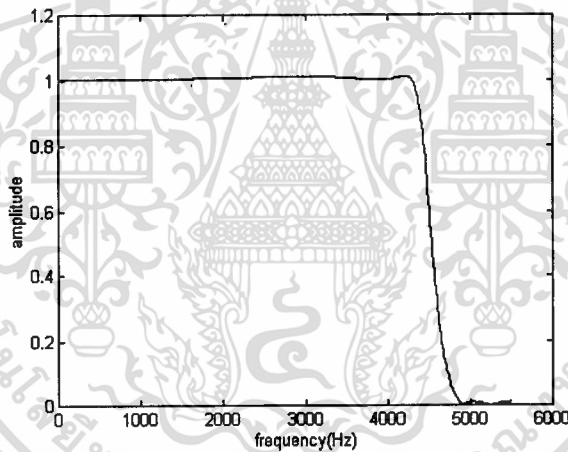
ท3.1.2 ตั้งแต่ความถี่ 5 กิโลเฮิร์ตซ์ ขึ้นไป มีอัตราขยายไม่เกิน 0.01

ท3.1.3 ความถี่ในการสุ่มสัญญาณ (sampling frequency) 11.025 กิโลเฮิร์ตซ์

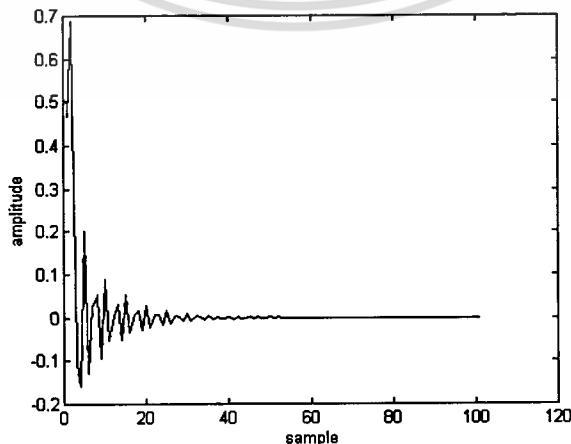
จากคุณสมบัติดังกล่าว สามารถออกแบบวงจรกรองชนิดอิลิปติก (elliptic) [9] จะได้ฟังก์ชันถ่ายโอนเป็น

$$H(z) = (0.467 + 1.802z^{-1} + 2.67z^{-2} + 1.802z^{-3} + 0.467z^{-4}) / (1 + 2.386z^{-1} + 10.27z^{-2} + 1.156z^{-3} + 0.231z^{-4}) \quad (ท3.3)$$

สามารถวาดกราฟแสดงขนาดของ $H(z)$ เทียบกับแกนความถี่ ได้ดังนี้



รูปที่ 3.2(ก) กราฟค่าสัมบูรณ์ของฟังก์ชันถ่ายโอนในสมการ (ท3.3)



รูปที่ 3.2(ข) กราฟผลตอบสนองของอิมพัลส์ของวงจรกรอง

เอกสารนี้เป็นเอกสารที่สงวนไว้เพื่อใช้ในการเรียนการสอนเท่านั้น เมื่อผู้ดูแลเนื้อหาเห็นว่าไม่เหมาะสมหรือไม่ควรเผยแพร่ หรือหากมีข้อผิดพลาดประการใด ขออภัยและขอสงวนสิทธิ์ในเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากสมการ (ท3.3) เราจะได้สัมประสิทธิ์ของสมการผลต่างสืบเนื่องคือ

$$a_0 = 0.46674455$$

$$a_1 = 1.80176901$$

$$a_2 = 2.67129358$$

$$a_3 = 1.80176901$$

$$a_4 = 0.46674455$$

$$b_1 = 2.3860733$$

$$b_2 = 2.4349378$$

$$b_3 = 1.1564779$$

$$b_4 = 0.23083172$$

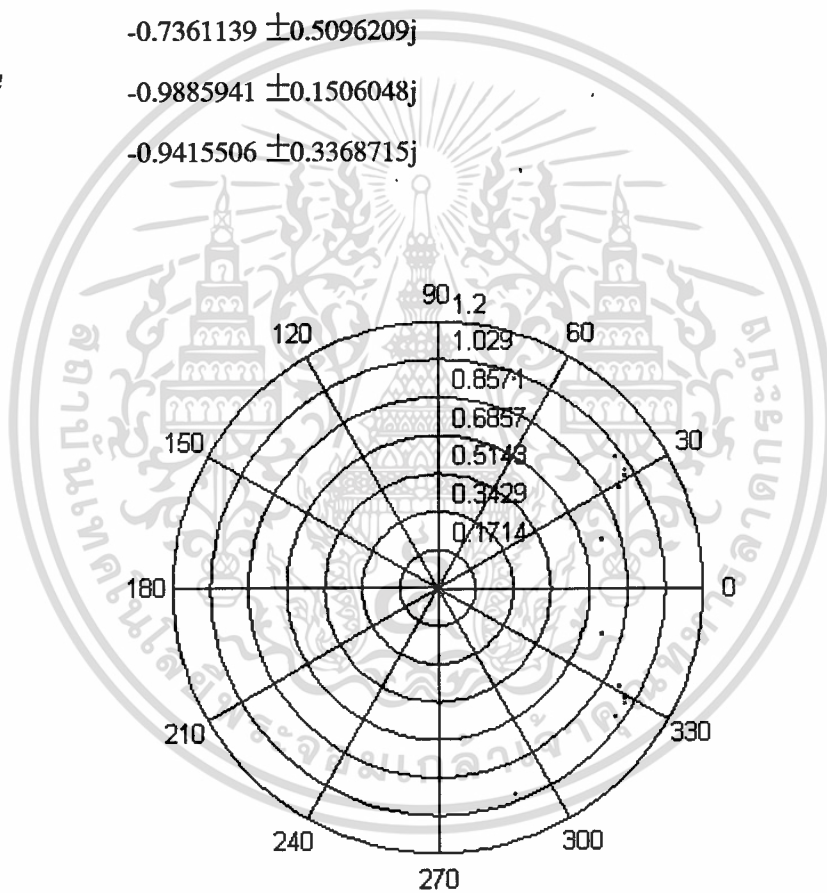
และ จะได้ค่าโพล (pole) และซีโร (zero) ของวงจรรองคือ

ค่าโพล $-0.4569228 \pm 0.2814134j$

$$-0.7361139 \pm 0.5096209j$$

ค่าซีโร $-0.9885941 \pm 0.1506048j$

$$-0.9415506 \pm 0.3368715j$$



รูปที่ ๓.3 กราฟแสดงตำแหน่งค่าโพลและซีโรของวงจรรองตามสมการ (ท3.3)

ซึ่งเขียนสมการผลต่างสืบเนื่องได้เป็น

$$y(n) = 0.4667x(n)+1.8017x(n-1)+2.6712x(n-2)+1.8017x(n-3)+0.4667x(n-4)$$

$$-2.3860y(n-1)-2.4349y(n-2)-1.1564y(n-3)-0.23083y(n-4)$$

(ท3.4)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ท3.2 วงจรกรองเชิงเลขแบบป้อนกลับชนิดความถี่ต่ำผ่าน อันดับ 6 โดยมีคุณสมบัติทางความถี่ ดังนี้

ท3.2.1 ที่ความถี่ 0-900 เฮิรตซ์ มีอัตราขยายเท่ากับ 1 ค่าคลาดเคลื่อนไม่เกิน 0.04

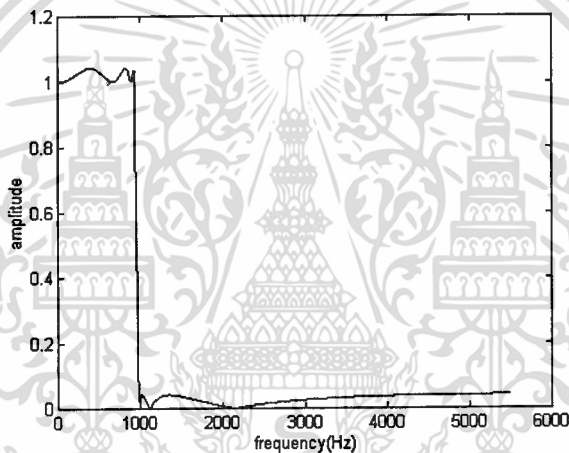
ท3.2.2 ตั้งแต่ความถี่ 1 กิโลเฮิรตซ์ ขึ้นไป มีอัตราขยายไม่เกิน 0.04

ท3.2.3 ความถี่ในการรุ่มสัญญาณ 11.025 กิโลเฮิรตซ์

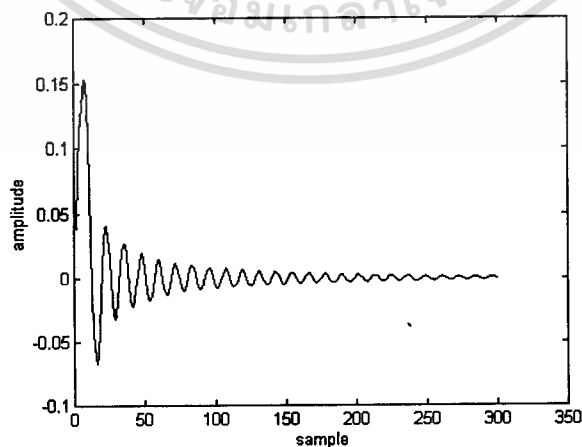
จากคุณสมบัติดังกล่าว สามารถออกแบบวงจรกรองชนิดอิลิปติก [9] จะได้ฟังก์ชันถ่ายโอนเป็น

$$H(z) = (0.05 - 0.2z^{-1} + 0.4z^{-2} - 0.5z^{-3} + 0.4z^{-4} - 0.2z^{-5} + 0.05z^{-6}) / (1 - 4.8z^{-1} + 10.3z^{-2} - 12z^{-3} + 8z^{-4} - 3.2z^{-5} + 0.5z^{-6}) \quad (ท3.5)$$

สามารถวาดกราฟแสดงขนาดของ $H(z)$ เทียบกับแกนความถี่ ได้ดังนี้



รูปที่ 3.4(ก) กราฟแสดงค่าสัมบูรณ์ของฟังก์ชันถ่ายโอนในสมการ (ท3.5)



รูปที่ 3.4(ข) กราฟผลตอบสนองอิมพัลส์ของวงจรกรอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากสมการ (ท3.5) เราจะได้สัมประสิทธิ์ของสมการผลต่างสืบเนื่องคือ

$$a_0 = 0.046468830$$

$$a_1 = -0.18644581$$

$$a_2 = 0.37448168$$

$$a_3 = -0.46197630$$

$$a_4 = 0.37448168$$

$$a_5 = -0.18644581$$

$$a_6 = 0.046468830$$

$$b_1 = -4.8453796$$

$$b_2 = 10.275010$$

$$b_3 = -12.101068$$

$$b_4 = 8.3213027$$

$$b_5 = -3.1621918$$

$$b_6 = 0.51935979$$

และจะได้ค่าโพล (pole) และซีโร (zero) ของวงจรวงคือ

ค่าโพล $0.7465326 \pm 0.2169899j$

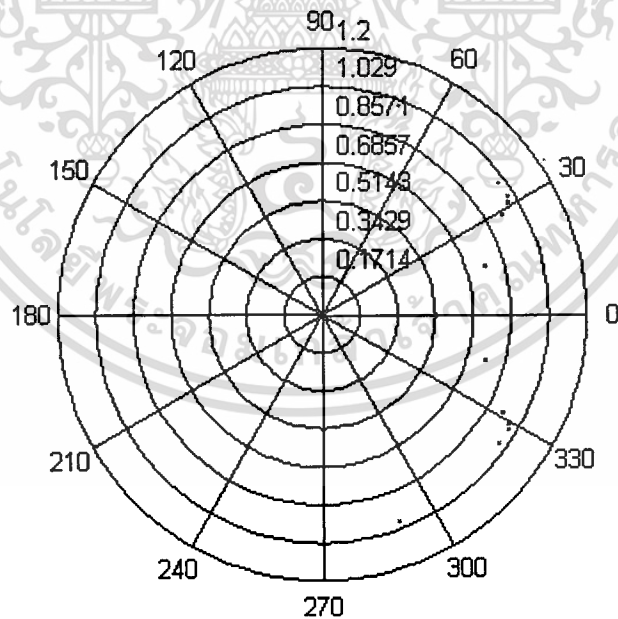
$$0.8232332 \pm 0.4469322j$$

$$0.8529941 \pm 0.5018250j$$

ค่าซีโร $0.8088087 \pm 0.5880718j$

$$0.3494310 \pm 0.9369621j$$

$$0.8478987 \pm 0.5301583j$$



รูปที่ ๓.5 กราฟแสดงตำแหน่งค่าโพลและซีโรของวงจรวงตามสมการ (ท3.5)

ซึ่งเขียนสมการผลต่างสืบเนื่องได้เป็น

$$y(n) = 0.046x(n) - 0.186x(n-1) + 0.374x(n-2) - 0.461x(n-3) + 0.374x(n-4) - 0.186x(n-5) + 0.046x(n-6)$$

เอกสารนี้เป็นเอกสาร +4.845y(n-1) - 10.275y(n-2) + 12.101y(n-3) - 8.321y(n-4) + 3.162y(n-5) - 0.519y(n-6) ดัง (ท3.6) ำ

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

- [1] วัลลภ สุระกำพลธร, *การประมวลผลสัญญาณเชิงเลข*, กรุงเทพฯ : บริษัท ไคนาพรีนซ์ จำกัด, 2533.
- [2] สุเชียร เกียรติสุนทร, *การประมาณพัลลภเชิงเส้นเสียงพูด*, วิทยานิพนธ์ปริญญาโท, บัณฑิต วิทยาลัยสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, พ.ศ. 2525.
- [3] J. Makhoul, *Linear Prediction : A Tutorial Review*, Proc. IEEE, vol. 63, no. 4, pp. 561-579, April 1975.
- [4] B. S. Atal, *Predictive Coding of Speech at Low Bit Rates*, IEEE Trans. Communications, vol. COM-30, no. 4, April, 1982.
- [5] B. S. Atal and S. L. Hanauer, *Speech Analysis and Synthesis by Linear Prediction of the Speech Wave*, J. Acoust. Soc. Amer. 50, pp. 637-655, April, 1971.
- [6] A. V. Oppenheim and R. W. Schaffer, *Digital Signal Processing*, New York : John Wiley & Sons, 1976.
- [7] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing : Principles, Algorithms, and Applications*, New York : Macmillan, 1988.
- [8] J. S. Lim and A. V. Oppenheim, *Advance Topics in Signal Processing*, New Jersey : Prentice-Hall, 1988.
- [9] R. Kuc, *Introduction to Digital Signal Processing*, Singapore : McGraw-Hill, 1982.
- [10] N. B. Jones, *Digital Signal Processing*, London : Peter Peregrinus, 1982.
- [11] S. J. Orfanidis, *Optimum Signal Processing : an Introduction*, Singapore : McGraw-Hill, 1990.
- [12] D. O' Shaughnessy, *Speech Communication : Human and Machine*, USA : Addison-Wesley, 1987.
- [13] L. R. Rabiner and R. W. Schaffer, *Digital Processing of Speech Signals*, New Jersey : Prentice-Hall, 1978.
- [14] P. E. Papamichalis, *Practical Approaches to Speech Coding*, New Jersey : Prentice-Hall, 1987.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [15] L. Rabiner and Biing-Hwang Juang, *Fundamentals of Speech Recognition*, New Jersey : Prentice-Hall, 1993.
- [16] The MathWorks, Inc., *Matlab : High-Performance Numeric Computation and Visualization Software*, Mass : The MathWorks, Inc., 1992.
- [17] A. Stolz, *The Sound Blaster Book*, Michigan : Abacus, 1993.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ปริญญาานิพนธ์ฉบับนี้ มีอาจสำเร็จลุล่วงได้ถ้าปราศจากซึ่งคุณพ่อ คุณแม่ ท่านผู้ให้กำเนิด คอยเลี้ยงดูเราและคอยให้กำลังใจเราตลอดมา ลูกขอกราบขอบพระคุณมา ณ โอกาสนี้ ขอบขอบพระคุณ อาจารย์ กอบชัย เศษหาญ และอาจารย์ ไกรสิน ส่วงวัฒนา ที่ให้คำปรึกษาเมื่อเกิดปัญหา ตลอดจนแหล่งข้อมูลข่าวสารที่จำเป็น เอกสารทางวิชาการต่างๆ รวมทั้งการตรวจทานต้นฉบับและการพิมพ์ปริญญาานิพนธ์ฉบับนี้ ขอบขอบพระคุณ อาจารย์ โมไนย ไกรฤกษ์ สำหรับหนังสือคู่มือโปรแกรม MATLAB



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้