การสังเคราะห์เสียงขลุ่ยไทยโดยใช้ SINUSOIDAL PLUS NOISE MODEL

THAI FLUTE SOUND SYNTHESIS USING SINUSOIDAL PLUS NOISE MODEL



รุ่งกานต์ ศิริเจริญไชย

RUNGKARN SIRICHAROENCHAI

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2544

ISBN 974-648-421-4

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การสังเคราะห์เสียงขลุ่ยไทยโดยใช้ SINUSOIDAL PLUS NOISE MODEL

THAI FLUTE SOUND SYNTHESIS USING SINUSOIDAL PLUS NOISE MODEL



รุ่งกานต์ ศิริเจริญไชย

RUNGKARN SIRICHAROENCHAI

เลขหมู่.....
เลขทะเบียน 40778
วัน, เดือน, ปี 26 พ.ย. 2544

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

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พ.ศ. 2544

ISBN 974-648-421-4

THAI FLUTE SOUND SYNTHESIS
USING SINUSOIDAL PLUS NOISE MODEL

RUNGKARN SIRICHAROENCHAI

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR THE DEGREE OF

MASTER OF ENGINEERING IN INFORMATION TECHNOLOGY

SCHOOL OF GRADUATE STUDIES

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2001

ISBN 974-648-421-4

COPYRIGHT 2001
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
SCHOOL OF GRADUATE STUDIES
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKARBANG

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ การสังเคราะห์ เสียงขลุ่ยไทยโดยใช้ Sinusoidal Plus Noise Model
THAI FLUTE SOUND SYNTHESIS USING SINUSOIDAL PLUS NOISE MODEL

ชื่อนักศึกษา นางสาวรุ่งกานต์ ศิริเจริญไชย

รหัสประจำตัว 40067017

ปริญญา วิทยาศาสตรมหาบัณฑิต

สาขาวิชา เทคโนโลยีสารสนเทศ

อาจารย์ผู้ควบคุมวิทยานิพนธ์ ดร.นพพร โชติกกำธร

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
ดร.นพพร	โชติกกำธร	
รศ.ดร.รัตติกร	วรากูลศิริพันธ์	
รศ.ดร.วิเชียร	เปรมชัยสวัสดิ์	
ดร.อาริต	ธรรมโน	
อาจารย์อัครินทร์	คุณกิตติ	

วัน/เดือนปี ที่สอบ 10 กันยายน 2544 เวลา 10.00 น. เป็นต้นไป

สถานที่สอบ ณ คณะเทคโนโลยีสารสนเทศ สจล.

บัณฑิตวิทยาลัยรับรองแล้ว

(รศ.ดร.บุญวัฒน์ อัตชู)

คณบดีบัณฑิตวิทยาลัย

วันที่ 9 **เดือน** กันยายน **พ.ศ.** 2544

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การสังเคราะห์เสียงขลุ่ยไทยโดยใช้ Sinusoidal Plus Noise Model
นักศึกษา	นางสาว รุ่งกวนต์ ศิริเจริญไชย
รหัสประจำตัว	40067017
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
พ.ศ.	2544
อาจารย์ผู้ควบคุมวิทยานิพนธ์	ดร.นพพร โชติกกำธร

บทคัดย่อ

วิทยานิพนธ์ฉบับนี้นำเสนอผลการศึกษาวิจัยการประยุกต์ใช้แนวทางการวิเคราะห์และสังเคราะห์เสียงเชิงสเปกตรัมกับเครื่องดนตรีขลุ่ยไทย โดยอาศัยแบบจำลองการสังเคราะห์เสียงดนตรีแบบ Sinusoidal Plus Noise Model ในงานวิจัยได้ทำการวิเคราะห์เสียงขลุ่ย (ในช่วงภาวะเสถียรและช่วงที่มีการเปลี่ยนแปลง) เพื่อหาค่าพารามิเตอร์สำหรับใช้ในการสังเคราะห์เสียงด้วยแบบจำลองดังกล่าว เช่น ค่าความถี่พื้นฐานและฮาร์โมนิก ค่าแอมปลิจูดและค่าสัมประสิทธิ์ตัวกรองสัญญาณ โดยจากการวิเคราะห์สเปกตรัมของเสียงขลุ่ยพบว่า ประกอบด้วยความถี่พื้นฐานและความถี่ฮาร์โมนิกจำนวน 10 ค่าความถี่ และจากการวัดคุณภาพของเสียงที่สังเคราะห์ได้ด้วยการฟังโดยอาศัยการให้คะแนนจากผู้ที่มีพื้นฐานทางด้านดนตรีไทย พบว่าจำนวนสัมประสิทธิ์ของตัวกรองสัญญาณ (ซึ่งที่ใช้ในกระบวนการสังเคราะห์) อย่างน้อย 15 จำนวน ให้เสียงสังเคราะห์ที่ยังคงลักษณะเสียงขลุ่ย โดยพบว่าการใช้เออาร์โมเดลในการสังเคราะห์เสียงขลุ่ยในส่วนของสัญญาณรบกวนได้ผลที่ดีกว่าโมเดลที่ใช้ในงานวิจัยอื่น และค่าพารามิเตอร์ที่ได้จากการวิเคราะห์ได้นำมาใช้ในการพัฒนาโปรแกรมบนเครื่องคอมพิวเตอร์เพื่อการสังเคราะห์เสียงในแบบเรียลไทม์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	Thai Flute Sound Synthesis Using Sinusoidal Plus Noise Model
Student	Miss Rungkarn Siricharoenchai
Student ID.	40067017
Degree	Master of Science
Programme	Information Technology
Year	2001
Thesis Advisor	Dr. Nopporn Chotikakamthorn

ABSTRACT

This thesis presents the results from research study on the application of spectral analysis and synthesis technique to the sound of Thai flute. In this study the sinusoidal plus noise model was applied. We analyzed recorded sound of Thai flute (during stationary and non-stationary parts) to estimate the values of parameters required for sound synthesis. Such parameters include frequency, amplitude and filter coefficients. From the analysis, we found that the sound of Thai flute is generally composed of 10 principal harmonics. From the sound quality listening test, participated by Thai musical-trained listeners, it has been found that at least 15 filter coefficients are needed in generating the noise part, to obtain the synthesized sound which can be recognized as a sound of Thai flute. In addition, the use of AR model for synthesizing the noise part was found to yield superior result as compared to the MA model used in other research works. Finally, with necessary parameters obtained from the analysis stage, Thai flute sound synthesizer has been successfully implemented on a personal computer using C++ and DirectX.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยดี เพราะได้รับความเมตตาจากอาจารย์ผู้ควบคุมวิทยานิพนธ์ ดร. นพพร โชติกกำธร ซึ่งได้ให้คำปรึกษาในการแก้ปัญหาทุก ๆ เรื่องและให้คำชี้แนะรวมทั้งแนวทางในการทำวิจัยตลอดมา ข้าพเจ้ารู้สึกซาบซึ้งในความกรุณาของท่านและขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้

นอกจากนั้นข้าพเจ้าขอขอบคุณและเอ่ยนามถึงบุคคลต่าง ๆ ที่ได้มีส่วนร่วมในการสร้างงานวิจัยชิ้นนี้ให้เสร็จสมบูรณ์ขึ้นมาได้

ขอกราบขอบพระคุณบุพการี ผู้ให้สติปัญญา ความคิดอ่าน และคอยให้กำลังใจข้าพเจ้าตลอดมา

ขอขอบพระคุณ อาจารย์ธนิสร์ ศรีกลิ่นดี ที่กรุณาเล่นขลุ่ยเพียงออ เพื่อการบันทึกเป็นข้อมูลในการวิจัยและนางสาวสุวรรณี กิ่งเพชร ที่ให้การบันทึกเสียงขลุ่ยประกอบการศึกษาลักษณะสัญญาณเสียงในเบื้องต้น

ขอขอบพระคุณอาจารย์สุบิน จันทร์แก้ว อาจารย์เสนาะ หลวงสุนทร อาจารย์สุวิทย์ แก้วกระมล อาจารย์ฉลาก โพธิ์สามต้น อาจารย์ถุงเงิน ทองโด อาจารย์ประมวล อรรถชีพ อาจารย์อุทัย แก้วละเอียด และอาจารย์บุญส่ง ธรรมวณิชย์ อาจารย์ประจำภาควิชาศิลปนิเทศ คณะมนุษยศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ที่กรุณาเสียสละเวลาในการประเมินคุณภาพของเสียงสังเคราะห์ในเบื้องต้น พร้อมทั้งเอื้อเฟื้อสถานที่เพื่อทำการวิจัย

ขอขอบพระคุณ ผศ.วิเชียร อ่อนละมูล ผศ.จันทร์พร นวลแพ่ง อาจารย์อดุล วงศ์แก้ว และอาจารย์สมศักดิ์ ท้าวสบาย อาจารย์ประจำภาควิชาดนตรี สถาบันราชภัฏจันทรเกษม ได้ให้ความช่วยเหลือข้าพเจ้าในการจัดหาน้องนักศึกษาเข้าร่วมการประเมินคุณภาพเสียงและเอื้อเฟื้อสถานที่ในการทำการประเมินอีกด้วย และขอขอบคุณน้องนักศึกษาปี 2 3 และ 4 ทั้งเอกดนตรีไทยและสากลในภาคการศึกษา 2543 สถาบันราชภัฏจันทรเกษม ที่ให้ความร่วมมือในการทำแบบประเมินคุณภาพเสียงเป็นอย่างดี

ขอขอบพระคุณครูเอ๋ ที่ได้ให้ความรู้เกี่ยวกับขลุ่ยไทยมากขึ้น

ขอขอบพระคุณอาจารย์ประจำคณะเทคโนโลยีสารสนเทศทุกท่านที่ได้อบรมสั่งสอนข้าพเจ้าให้มีความรู้เพิ่มพูน รวมทั้งขอขอบคุณคณะเทคโนโลยีสารสนเทศที่สนับสนุนเครื่องมือและอุปกรณ์ที่ใช้ในการทดลอง ตลอดจนห้องปฏิบัติการ Multimedia and Virtual Research Lab ที่ได้เอื้อเฟื้อสถานที่ตลอดการทำวิจัยนี้

ขอขอบคุณเจ้าหน้าที่คณะเทคโนโลยีสารสนเทศที่ได้ให้ความอนุเคราะห์ในความสะดวกต่าง ๆ ในการทำงานตั้งแต่ต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น ไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารนี้ทุกครั้งที่มีการนำไปใช้

ขอขอบคุณพี่ ๆ เพื่อน ๆ นักศึกษาทุกคนที่ได้ให้ความช่วยเหลือ ให้คำแนะนำและกำลังใจข้าพเจ้ามาตลอด

สุดท้ายนี้ขอขอบพระคุณบัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่ได้ให้ทุนสนับสนุนการทำวิจัยในครั้งนี้

คุณค่าและประโยชน์อันพึงมีจากวิทยานิพนธ์ฉบับนี้ ข้าพเจ้าขอมอบแด่มีพระคุณทุกท่าน

รุ่งกานต์ ศิริเจริญไชย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....I

บทคัดย่อภาษาอังกฤษ..... II

กิตติกรรมประกาศ..... III

สารบัญ..... V

สารบัญตาราง..... VIII

สารบัญรูป..... IX

บทที่ 1 บทนำ..... 1

1.1 ความเป็นมาและความสำคัญของปัญหา..... 1

1.2 วัตถุประสงค์ของการศึกษา..... 2

1.3 ทฤษฎีหรือแนวคิดที่ใช้ในการวิจัย..... 2

1.4 ขอบเขตของการวิจัย..... 3

1.5 ขั้นตอนการศึกษา..... 3

1.6 ประโยชน์ที่คาดว่าจะได้รับ..... 4

1.7 โครงสร้างของวิทยานิพนธ์..... 4

บทที่ 2 หลักการสังเคราะห์เสียงดนตรี..... 6

2.1 ความเป็นมาของการสังเคราะห์เสียงด้วยระบบดิจิตอล..... 6

2.2 วิธีการสังเคราะห์เสียงดนตรี..... 7

2.2.1 การสังเคราะห์เสียงโดยใช้วิธีการแซมปลิงหรือเวปเทเบิล..... 7

2.2.2 การสังเคราะห์เสียงด้วยแบบจำลองเชิงกายภาพ..... 7

2.2.3 การสังเคราะห์เสียงด้วยแบบจำลองเชิงสเปกตรัม..... 7

2.2.3.1 Sinusoidal Model..... 7

2.2.3.2 Sinusoidal Plus Noise Model..... 8

2.2.4 เปรียบเทียบแนวทางการสังเคราะห์เสียงดนตรี..... 10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.3 หลักการสังเคราะห์เสียงด้วยแบบจำลอง Sinusoidal Plus Noise..... 10

2.3.1ขั้นตอนการวิเคราะห์สัญญาณเสียง..... 11

2.3.1.1 ขั้นตอนการวิเคราะห์สัญญาณไซน์..... 12

2.3.1.2 การวิเคราะห์ในส่วนของสัญญาณรบกวน.....13

2.3.1.2.1 การวิเคราะห์ค่าสัมประสิทธิ์ตัวกรองสัญญาณสำหรับ MA โมเดล.... 17

2.3.1.2.2 การวิเคราะห์ค่าสัมประสิทธิ์ตัวกรองสัญญาณสำหรับ ARโมเดล..... 18

2.3.2 ขั้นตอนการสังเคราะห์เสียง.....20

บทที่ 3 ข้อพิจารณาในการสังเคราะห์เสียงแบบ Sinusoidal Plus Noise Model.....22

3.1 การพิจารณาชนิดวินโดว์ที่เหมาะสม.....22

3.2 การเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์ให้กับสัญญาณเสียง.....26

บทที่ 4 การวิเคราะห์และการสังเคราะห์เสียงขลุ่ยไทยด้วยแบบจำลอง Sinusoidal Plus Noise..... 35

4.1 การวิเคราะห์ในส่วนของสัญญาณไซน์..... 35

4.1.1 การวิเคราะห์สัญญาณเสียงในช่วงภาวะเสถียร..... 35

4.1.2 การวิเคราะห์สัญญาณเสียงในช่วงภาวะที่ไม่เสถียร.....37

4.2 การวิเคราะห์ในส่วนของสัญญาณรบกวน.....40

4.2.1 การวิเคราะห์ค่าสัมประสิทธิ์ตัวกรองสัญญาณสำหรับ MA โมเดล..... 40

4.2.2 การวิเคราะห์ค่าสัมประสิทธิ์ตัวกรองสัญญาณสำหรับ MA โมเดล..... 42

4.2.3 การเปรียบเทียบความเหมาะสมของการใช้ MA และ AR โมเดล สำหรับ
การสังเคราะห์ค่าคุณสมบัติทางสถิติของสัญญาณรบกวนที่เหมาะสม.....42

4.2.3.1 การประเมินแบบออปเจคทีฟ..... 43

4.2.3.2 การประเมินแบบซับเจคทีฟ.....44

4.3 การเปรียบเทียบแนวทางการสังเคราะห์เสียงแบบ Sinusoidal Model
และ Sinusoidal Plus Noise Model.....45

4.3.1 จำนวนพารามิเตอร์ที่ใช้ในการสังเคราะห์ต่อคุณภาพเสียง.....45

4.3.2 ความซับซ้อนในการคำนวณ.....47

4.4 การประเมินระดับคุณภาพของเสียง..... 47

เอกสารนี้เป็นเ... ไม่ว่ากรณีใดๆ ... และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

บทที่ 5 การพัฒนาโปรแกรมเพื่อการสังเคราะห์เสียงขลุ่ยไทยแบบเรียลไทม์.....51

5.1 แนวคิดในการพัฒนาโปรแกรม.....51

5.2 ความต้องการของระบบในการพัฒนาโปรแกรม.....51

5.3 หลักการทำงานของโปรแกรม.....52

5.4 การใช้งานโปรแกรม.....54

บทที่ 6 สรุปผลการวิจัยและข้อเสนอแนะ.....57

6.1 สรุปผลการวิจัย.....57

6.2 บทวิเคราะห์.....58

6.3 ข้อเสนอแนะและแนวทางในการทำวิจัยต่อ.....59

เอกสารอ้างอิง.....60

ภาคผนวก.....62

ภาคผนวก ก.....63

ภาคผนวก ข.....67

ภาคผนวก ค ขลุ่ยไทย.....69

ภาคผนวก ง สเปกตรัมของโน้ตขลุ่ยไทย.....73

ภาคผนวก จ ลักษณะของสัญญาณรบกวน.....78

ภาคผนวก ฉ ลักษณะของสัญญาณเสียงสังเคราะห์.....81

ภาคผนวก ช บทความที่ได้รับการตีพิมพ์.....86

ประวัติผู้เขียน.....105

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ หน้า

2.1 คุณสมบัติของการสังเคราะห์เสียงในแต่ละวิธี.....10

3.1 ค่าความคลาดเคลื่อนเฉลี่ยของค่าความถี่และค่าแอมปลิจูดของสัญญาณไซน์ลำดับที่หนึ่ง จากการใช้วินโดว์ชนิดต่าง..... 25

3.2 ค่าความคลาดเคลื่อนเฉลี่ยของค่าความถี่และค่าแอมปลิจูดของสัญญาณไซน์ลำดับที่สอง จากการใช้วินโดว์ชนิดต่าง ๆ.....26

4.1 เปอร์เซ็นต์ความคลาดเคลื่อนเฉลี่ยของค่าความถี่..... 36

4.2 เปอร์เซ็นต์ความคลาดเคลื่อนเฉลี่ยของค่าแอมปลิจูด.....36

4.3 พารามิเตอร์ของสัญญาณไซน์ลำดับที่ต่าง ๆ ในแต่ละตัวโน้ต..... 33

4.4 ค่าสัมประสิทธิ์ตัวกรองสัญญาณที่ได้จากการวิเคราะห์ด้วยเอ็มเอโมเดล.....41

4.5 ค่าสัมประสิทธิ์ตัวกรองสัญญาณที่ได้จากการวิเคราะห์ด้วยเออาร์โมเดล..... 42

4.6 ค่าความผิดพลาดเฉลี่ยของค่าสหสัมพันธ์ของสัญญาณรบกวนสังเคราะห์ที่ได้ จากทั้งสองโมเดล ตั้งแต่ time lag ที่ 0-501.....44

4.7 ระดับคุณภาพของสัญญาณรบกวนสังเคราะห์.....44

4.8 การเปรียบเทียบระดับคุณภาพเสียงสังเคราะห์จากวิธีการสังเคราะห์เสียง แบบ Sinusoidal Model และ Sinusoidal Plus Noise Model.....47

4.9 ผลการประเมินคุณภาพของเสียง จากการใช้ค่าสัมประสิทธิ์ตัวกรองสัญญาณ ที่จำนวนต่าง ๆ.....49

ก.1 ค่าความคลาดเคลื่อนในแต่ละช่วงความถี่.....66

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่ หน้า

2.1 การแยกองค์ประกอบของสัญญาณเสียงออกเป็นสองส่วน.....10

2.2 ไดอะแกรมขั้นตอนการสังเคราะห์เสียง..... 11

2.3 ไดอะแกรมขั้นตอนการวิเคราะห์สัญญาณเสียงด้วยวิธี Sinusoidal Plus Noise.....11

2.4 สัญญาณเสียงคูณกับวินโดว์ฟังก์ชัน.....14

2.5 โครงสร้างของเอ็มเอโมเดล.....15

2.6 โครงสร้างของเออาร์โมเดล.....15

2.7 ตัวอย่างสเปกตรัมของสัญญาณ..... 16

3.1 ลักษณะวินโดว์ฟังก์ชันประเภทต่าง ๆ..... 23

3.2 ค่าความคลาดเคลื่อนของค่าความถี่ที่ได้จากการใช้วินโดว์ทุกประเภท..... 27

3.3 ค่าความคลาดเคลื่อนของค่าแอมปลิจูดจากการใช้วินโดว์แต่ละประเภท.....28

3.4 การเปรียบเทียบสเปกตรัมที่ได้จากการใช้วินโดว์ประเภทต่างๆ.....29

3.5 เปอร์เซ็นต์ความผิดพลาดของค่าความถี่ เมื่อเพิ่มจำนวนข้มูลที่มีค่าศูนย์
ตั้งแต่ 1000 จนถึง 66000 ค่าโดยเพิ่มทีละ 10000 ค่า.....31

3.6 เปอร์เซ็นต์ความผิดพลาดของค่าแอมปลิจูด เมื่อเพิ่มจำนวนข้อมูลที่มีค่าศูนย์
ตั้งแต่ 1000 จนถึง 66000 ค่า โดยเพิ่มทีละ 10000 ค่า..... 32

3.7 ค่าความคลาดเคลื่อนของค่าความถี่ที่ได้จากการเพิ่มจำนวนศูนย์
เป็นจำนวนเท่าของสัญญาณ.....33

3.8 ค่าความคลาดเคลื่อนของค่าแอมปลิจูดที่ได้จากการเพิ่มจำนวนศูนย์
เป็นจำนวนเท่าของสัญญาณ.....34

4.1 ลักษณะสัญญาณในช่วงที่มีการเปลี่ยนแปลง.....39

4.2 ตัวอย่างการเปรียบเทียบสเปกตรัมของโน้ตต้นแบบกับที่สังเคราะห์ขึ้นใหม่..... 41

4.3 เพาเวอร์สเปกตรัมของสัญญาณรบกวนของโน้ตโด.....42

4.4 การเปรียบเทียบค่าสหสัมพันธ์ของสัญญาณต้นแบบกับสัญญาณรบกวนสังเคราะห์
ในเอ็มเอโมเดลและเออาร์โมเดลซึ่งใช้จำนวนสัมประสิทธิ์ตัวกรองสัญญาณเท่ากัน
101 จำนวน..... 43

4.5 สัมประสิทธิ์ตัวกรองสัญญาณที่ได้จากการวิเคราะห์แบบเออาร์โมเดล..... 46

4.6 การเปรียบเทียบสัญญาณเสียงต้นแบบ (ก) และสัญญาณเสียงสังเคราะห์ (ข) ของโน้ตโด..... 50

ก.1 ค่าความคลาดเคลื่อนของค่าความถี่ในแต่ละช่วงความถี่..... 64

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่ หน้า

ก.2 ค่าความคลาดเคลื่อนของค่าแอมปลิจูดในแต่ละช่วงความถี่.....65
ค.1 ลักษณะขลุ่ยหลีบ.....69
ค.2 ลักษณะขลุ่ยเพียงออ.....70
ค.3 ลักษณะขลุ่ยอู้.....70
ค.4 ลักษณะคลื่นสัญญาณเสียงขลุ่ยเพียงออตั้งแต่โน้ตโดจนถึงโน้ตซอลสูง.....71
ง.1 การเปรียบเทียบสเปกตรัมส่วนที่เป็น attack (ซ้าย) และส่วนที่เป็น decay (ขวา) ของสัญญาณเสียงต้นแบบกับที่สังเคราะห์ขึ้นใหม่.....73
จ.1 ลักษณะสัญญาณรบกวนที่ได้จากขั้นตอนการวิเคราะห์สัญญาณรบกวน.....78
ฉ.1 ลักษณะสัญญาณเสียงสังเคราะห์ ตั้งแต่โน้ตโดจนถึงโน้ตซอลสูง.....81

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในการสังเคราะห์เสียงด้วยคอมพิวเตอร์ จำเป็นต้องมีแบบจำลองที่ดีเพื่อที่จะได้เสียงที่มีคุณภาพดีและมีความซับซ้อนในการคำนวณน้อย. ซึ่งแนวทางในการสังเคราะห์เสียงนั้นมีหลายวิธี เช่น วิธีการแซมปลิงซึ่งเป็นวิธีการเก็บเสียงในรูปของคลื่นเสียง (waveform) โดยเก็บค่าแซมเปิลของเสียง ทำให้วิธีดังกล่าวมีความซับซ้อนในการคำนวณต่ำและให้เสียงสังเคราะห์ที่มีคุณภาพดี (ในแง่ความเป็นธรรมชาติ) แต่มีขีดความสามารถในการดัดแปลงแก้ไขเสียงได้จำกัด เช่น การทำ Vibrato[1] และ Audio morphing[2] ไม่สามารถกระทำได้ เนื่องจากเทคนิคดังกล่าวอาศัยหลักการเปลี่ยนแปลงค่าในเชิงความถี่ ซึ่งวิธีการแซมปลิงไม่สามารถจะแก้ไของค์ประกอบทางด้านความถี่ของเสียงที่บันทึกไว้ได้

นอกจากวิธีการแซมปลิงแล้ว เทคนิคการสังเคราะห์เสียงใหม่ ๆ ได้ถูกพัฒนาขึ้นในปัจจุบันมีอาทิเช่น การสังเคราะห์เสียงเชิงกายภาพ วิธีการดังกล่าวอาศัยลักษณะปรากฏการณ์ทางกายภาพของการเกิดเสียงของเครื่องดนตรีเพื่อการสังเคราะห์เสียง [4] [14] โดยพารามิเตอร์ต่าง ๆ ที่ใช้ในการสังเคราะห์เสียงเป็นค่าคุณสมบัติทางกายภาพของเครื่องดนตรีแต่ละชนิด วิธีการดังกล่าวให้เสียงสังเคราะห์ที่มีความเป็นธรรมชาติ แต่อย่างไรก็ดีการสังเคราะห์เสียงแนวทางนี้มีข้อจำกัดเรื่องความยืดหยุ่นในการดัดแปลงเสียง และมีความยุ่งยากเมื่อนำไปประยุกต์ใช้กับเสียงของเครื่องดนตรีแต่ละชนิด เนื่องจากในการสังเคราะห์เครื่องดนตรีแต่ละประเภทจำเป็นต้องทราบแบบจำลองเฉพาะของเครื่องดนตรีชนิดนั้น ๆ

วิธีการสังเคราะห์เสียงเชิงสเปกตรัมเป็นวิธีการสังเคราะห์เสียงอีกวิธีหนึ่งที่ได้รับการพัฒนาขึ้นเมื่อไม่นานมานี้ [9][10] วิธีการดังกล่าวมีข้อดีคือสามารถสังเคราะห์เสียงได้หลายประเภทไม่จำกัดเฉพาะเสียงเครื่องดนตรี และมีความยืดหยุ่นสูงกว่าในการดัดแปลงหรือสร้างเสียงใหม่เมื่อเปรียบเทียบกับวิธีการสังเคราะห์เสียงเชิงกายภาพและการสังเคราะห์เสียงด้วยการแซมปลิง

งานวิจัยที่นำเสนออาศัยแบบจำลองเชิงสเปกตรัมในการวิเคราะห์และสังเคราะห์เสียงเครื่องดนตรีขลุ่ยไทย เนื่องจากว่าแบบจำลองการสังเคราะห์เสียงเชิงสเปกตรัมมีความยืดหยุ่นสูงสามารถนำมาประยุกต์ใช้กับเครื่องดนตรีหลากหลายประเภทรวมทั้งขลุ่ยไทยได้ และสามารถเปลี่ยนแปลงค่าเพื่อการสร้างเสียงที่แปลกใหม่ได้ ในขณะที่การสังเคราะห์เสียงเชิงกายภาพต้องอาศัยผลการศึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

---

[1] Vibrato คือ การเปลี่ยนความถี่หลักของเสียงแต่เปลี่ยนในระดับที่น้อย

[2] Audio morphing คือ การผสมเสียงเข้าด้วยกันโดยเป็นการผสมในเชิงความถี่

วิเคราะห์แบบจำลองเชิงฟิสิกส์ของเสียงขลุ่ยไทยในการสร้างแบบจำลองเฉพาะ ซึ่งจากการศึกษายังไม่พบว่ามีผู้ใดทำการวิเคราะห์แบบจำลองเชิงฟิสิกส์สำหรับเสียงดนตรีขลุ่ยไทย

1.2 วัตถุประสงค์ของการศึกษา

งานวิจัยนี้มุ่งศึกษาการสังเคราะห์เสียงเชิงสเปกตรัมเพื่อนำมาประยุกต์ใช้ในการสังเคราะห์เสียงขลุ่ยไทย ซึ่งในการสังเคราะห์เสียงจำเป็นต้องหาส่วนประกอบเชิงความถี่ของขลุ่ยไทย โดยเป้าหมายของงานวิจัยมีดังนี้

1) ประยุกต์ใช้วิธีการสังเคราะห์เสียง Sinusoidal Plus Noise สำหรับการสังเคราะห์เสียงขลุ่ยไทย โดยทำการวิเคราะห์หาค่าพารามิเตอร์ต่าง ๆ เพื่อให้สามารถนำมาสร้างครึ่งเสียงในเครื่องดนตรีขลุ่ยไทย

2) ศึกษาเปรียบเทียบความเหมาะสมในการสังเคราะห์เสียงสัญญาณรบกวนด้วยเอ็มเอโมเดลซึ่งใช้อยู่ในงานวิจัยเดิมและเออาร์โมเดลที่งานวิจัยนี้เสนอ

3) นำเสนอการประยุกต์ใช้วิธีการวิเคราะห์โดยวิธีการสังเคราะห์ (Analysis By Synthesis) สำหรับการวิเคราะห์เสียงในช่วงที่ไม่เสถียร

4) ศึกษาวินโดว์ฟังก์ชันที่เหมาะสมสำหรับการวิเคราะห์สัญญาณเสียงขลุ่ยไทย

5) ศึกษาแนวทางการเพิ่มจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์สำหรับการวิเคราะห์เชิงความถี่ (Spectral Analysis)

6) ศึกษาเปรียบเทียบแนวทางการสังเคราะห์เสียง Sinusoidal Model และ Sinusoidal Plus Noise Model ทั้งในเรื่องของความซับซ้อนในการคำนวณ ตลอดจนจำนวนพารามิเตอร์ที่เหมาะสมในการสังเคราะห์เสียง

7) พัฒนาโปรแกรมสังเคราะห์เสียงแบบเรียลไทม์บนเครื่องพีซี

1.3 ทฤษฎีหรือแนวคิดที่ใช้ในการวิจัย

ทฤษฎีของฟูเรียร์กล่าวไว้ว่า สัญญาณใด ๆ ประกอบไปด้วยสัญญาณไซน์ที่ความถี่ และแอมปลิจูดต่าง ๆ ที่เหมาะสมรวมกัน ซึ่งการแยกองค์ประกอบของสัญญาณเสียงออกเป็นองค์ประกอบย่อย ๆ ที่อยู่ในรูปแบบของฟังก์ชันไซน์ที่ความถี่ต่าง ๆ นั้นสามารถกระทำได้โดยการวิเคราะห์สัญญาณเสียงเชิงสเปกตรัม และการสังเคราะห์สัญญาณดังกล่าวขึ้นใหม่สามารถกระทำได้โดยการสร้างสัญญาณไซน์ที่เป็นองค์ประกอบของสัญญาณนั้นรวมกัน ซึ่งการสังเคราะห์เสียงเชิงสเปกตรัมตั้งอยู่บนพื้นฐานของทฤษฎีดังกล่าว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารนี้ทุกครั้งที่มีการนำไปใช้

การสังเคราะห์เสียงด้วยแบบจำลอง Sinusoidal Plus Noise เป็นแบบจำลองการสังเคราะห์เสียงเชิงสเปกตรัมประเภทหนึ่งที่มีความเหมาะสมในการประยุกต์ใช้กับเสียงดนตรีหรือเสียงพูด เนื่องจากว่าเสียงเหล่านี้มีส่วนของสัญญาณรบกวน (noise:ในที่นี้มิได้หมายถึงเสียงหรือสัญญาณที่

ไม่พึงปรารถนา) อยู่ด้วย ซึ่งเสียงรบกวนเมื่อทำการวิเคราะห์ในเชิงสเปกตรัมพบว่าประกอบไปด้วยองค์ประกอบเชิงความถี่ที่มีค่าแอมปลิจูดต่ำเป็นจำนวนมาก การสังเคราะห์เสียงภายใต้โมเดลนี้อยู่บนพื้นฐานการสังเคราะห์เสียงดนตรีด้วยผลรวมขององค์ประกอบของสัญญาณ 2 ส่วน [7] [11] คือ ส่วนของสัญญาณคลื่นไซน์ซึ่งเป็นองค์ประกอบของค่าความถี่ที่มีค่าแอมปลิจูดสูงและส่วนที่เป็นสัญญาณรบกวนซึ่งไม่สามารถจำลองได้ดีด้วยฟังก์ชันไซน์ หากแต่สามารถอธิบายได้ด้วยคุณสมบัติทางสถิติ โดยการสังเคราะห์เสียงด้วยวิธีการดังกล่าวจะได้จากการหาผลรวมขององค์ประกอบสัญญาณไซน์ที่ความถี่ต่าง ๆ รวมกับส่วนของสัญญาณรบกวน ซึ่งสามารถสังเคราะห์ได้โดยการกรองสัญญาณ white noise ด้วยฟิลเตอร์ที่เหมาะสม

1.4 ขอบเขตของการวิจัย

ในงานวิจัยนี้จะทำการวิเคราะห์หาส่วนประกอบเชิงความถี่ของเสียงโน้ตแต่ละตัวตั้งแต่เสียงโดต่ำถึงเสียงซอลสูงรวม 12 เสียงของขลุ่ยเพียงออและขลุ่ยไทยประเภทอื่น ๆ เพื่อหาค่าพารามิเตอร์ที่จะนำมาใช้ในการสังเคราะห์เสียง และทำการทดลองศึกษาความสัมพันธ์ระหว่างจำนวนพารามิเตอร์ที่ใช้กับคุณภาพของเสียงที่ได้ ตลอดจนศึกษาวิธีการลดความซับซ้อนในการประมวลผล และทำการพัฒนาโปรแกรมสังเคราะห์เสียงบนเครื่องพีซี

1.5 ขั้นตอนการศึกษา

การวิจัยจะทำการวิเคราะห์และสังเคราะห์เสียงขลุ่ยไทย ซึ่งมีขั้นตอนของการศึกษาโดยสรุปดังนี้

1) รวบรวมข้อมูลเสียงดนตรีขลุ่ยไทย โดยทำการบันทึกเสียงในทุกตัวโน้ต

2) ทำการวิเคราะห์เสียงขลุ่ยเพียงอยในทุกตัวโน้ต โดยใช้โปรแกรม MATLAB® ช่วยในการวิเคราะห์เชิงสเปกตรัมรวมทั้งการสังเคราะห์เสียงขึ้นใหม่

3) ศึกษาผลของชนิดวินโดว์ ขนาดวินโดว์และการเพิ่มจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์ ในขั้นตอนฟาสฟูเรียร์ทรานฟอร์มต่อค่าความถูกต้องของค่าพารามิเตอร์ต่าง ๆ มากน้อยเพียงใด

4) ศึกษาเปรียบเทียบความเหมาะสมของการใช้เอ็มเอโมเดล (Moving Average Model) และเออาร์โมเดล (Autoregressive Model) โมเดล สำหรับการสังเคราะห์ค่าคุณสมบัติทางสถิติของสัญญาณรบกวนที่เหมาะสม

5) ทำการทดลองหาจำนวนพารามิเตอร์ที่เหมาะสมในการสังเคราะห์เสียงโดยคำนึงถึงคุณภาพของเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารนี้ทุกครั้งที่มีการนำไปใช้

6) เปรียบเทียบแนวทางการสังเคราะห์เสียงแบบ Sinusoidal Model และ Sinusoidal Plus Noise Model ในแง่ของคุณภาพของเสียงและความซับซ้อนในการสังเคราะห์เสียงดนตรี

7) พัฒนาโปรแกรมการสังเคราะห์เสียงบนเครื่องพีซี โดยโปรแกรมทำการอ่านค่าพารามิเตอร์ต่าง ๆ ที่เก็บอยู่ในไฟล์เพื่อนำมาใช้ในการสังเคราะห์เสียงในเรียลไทม์

1.6 ประโยชน์ที่คาดว่าจะได้รับ

ในการสังเคราะห์เสียงด้วยวิธีการ Sinusoid Plus Noise Model นอกจากจะให้มีเสียงที่ใกล้เคียงเสียงของเครื่องดนตรีจริง ยังให้ผลดีในด้านต่าง ๆ ด้วย ยกตัวอย่างเช่น

1) สามารถแก้ไขหรือปรับแต่งเสียง ซึ่งในเครื่องดนตรีจริงไม่สามารถทำได้ เช่นการผสมเสียงเครื่องดนตรี 2 ชนิดเข้าด้วยกัน (morphing) หรือการพัฒนารูปแบบเสียงใหม่ ๆของดนตรี

2) ประโยชน์ซึ่งสืบเนื่องมาจากการศึกษาองค์ประกอบในเชิงความถี่ของเครื่องดนตรีไทย สามารถนำไปใช้ในการออกแบบเครื่องดนตรีจริง และการเปรียบเทียบคุณภาพของขลุ่ยได้

3) ลดปริมาณข้อมูล ให้ง่ายต่อการจัดเก็บและการส่งผ่านข้อมูล ในกรณีของมิดีไฟล์ (MIDI) จะเป็นชุดคำสั่งให้สังเคราะห์เสียงขลุ่ยไทย โดยมีอัลกอลิทึมในการสร้างเสียงขลุ่ยไทยซึ่งจะใช้ข้อมูลของค่าความถี่และค่าแอมปลิจูดจำนวนไม่มาก ดังนั้นจะเห็นได้ว่าปริมาณข้อมูลที่ใช้ในการสังเคราะห์เสียงด้วยแนวทางนี้มีจำนวนน้อยมากเมื่อเทียบกับการจัดเก็บเป็นลำดับของเสียง ในการจัดเก็บเป็นลำดับของเสียงจะเปลืองเนื้อที่ในการจัดเก็บมาก เสียงที่มีคุณภาพดีมากเท่าไรก็ต้องใช้แซมปลิงเรท (sampling rate) ที่ค่าสูง ๆ เมื่อแซมปลิงเรทที่ค่ามาก ๆ ก็ต้องใช้เนื้อที่ในการจัดเก็บมาก อย่างเช่น ไฟล์เวปจะจัดเก็บค่าเสียงจริง ทำให้ขนาดของไฟล์ใหญ่เมื่อเทียบกับมิดี้ไฟล์ ซึ่งจัดเก็บฟอร์แมตของคำสั่งเพื่อให้ซินทีไซน์เซอร์สร้างเสียงที่เหมือนจริง

1.7 โครงสร้างของวิทยานิพนธ์

ในงานวิทยานิพนธ์นี้ได้แบ่งเนื้อหาออกเป็น 6 บทต่อไปนี้

บทที่ 1 จะกล่าวถึงที่มาของงานวิจัยและวัตถุประสงค์ของงานวิจัย ตลอดจนทฤษฎีและหลักการที่เกี่ยวข้อง

บทที่ 2 เป็นเนื้อหาเกี่ยวกับหลักการสังเคราะห์เสียงดนตรีที่มีอยู่ในปัจจุบัน พร้อมทั้งกล่าวถึงแนวทางสังเคราะห์เสียงด้วยแบบจำลอง Sinusoidal Plus Noise โดยแยกขั้นตอนออกเป็นการวิเคราะห์และการสังเคราะห์เสียง

บทที่ 3 ข้อพิจารณาในการสังเคราะห์เสียงด้วยแบบจำลอง Sinusoidal Plus Noise โดยจะกล่าวถึงการเลือกใช้ชนิดวินโดว์ที่เหมาะสม การเพิ่มจำนวนจุดข้อมูลที่มีค่าศูนย์เข้ากับสัญญาณเพื่อให้ได้ค่าพารามิเตอร์ที่ถูกต้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4 กล่าวถึงการประยุกต์ใช้แนวทางการสังเคราะห์เสียงด้วยแบบจำลอง Sinusoidal Plus Noise กับเสียงขลุ่ยไทย โดยในรายละเอียดจะเป็นขั้นตอนการเก็บข้อมูลค่าความถี่ แอมปลิจูด และค่าสัมประสิทธิ์ตัวกรองสัญญาณ นอกจากนั้นยังทำการเปรียบเทียบโมเดลการสังเคราะห์สัญญาณ

รบกวนในแบบเอ็มเอและเออาร์โมเดลเพื่อการพิจารณาโมเดลที่เหมาะสม ตลอดจนทำการประเมินคุณภาพของเสียงสังเคราะห์

บทที่ 5 รายละเอียดในการพัฒนาโปรแกรมเพื่อการสังเคราะห์เสียงขลุ่ยไทย โดยจะเป็นหลักการทำงานของแอพพลิเคชั่น และรายละเอียดของระบบที่ใช้ในการพัฒนาโปรแกรมและทดสอบโปรแกรม

บทที่ 6 บทสรุปและข้อเสนอแนะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
หลักการสังเคราะห์เสียงดนตรี

2.1 ความเป็นมาของการสังเคราะห์เสียงด้วยระบบดิจิตอล

Digital Sound Synthesis คือการสังเคราะห์เสียงด้วยอุปกรณ์ดิจิตอล (digital synthesizer) ซึ่งเทคนิคนี้มีพื้นฐานมาจากการสังเคราะห์เสียงแบบอนาลอก (analog synthesis) การสังเคราะห์เสียงด้วยระบบดิจิตอลนี้จำแนกตามลักษณะอุปกรณ์ที่ใช้ในการสังเคราะห์เสียงได้ 2 ประเภท

1) Hardware Based Sound Synthesis เป็นการสังเคราะห์เสียงโดยใช้อุปกรณ์ที่ออกแบบเฉพาะสำหรับการสังเคราะห์เสียง

2) Software Based Sound Synthesis คือการสังเคราะห์เสียงโดยใช้เครื่องคอมพิวเตอร์สร้างเสียงขึ้นใหม่ ซึ่งการสังเคราะห์เสียงในลักษณะนี้สามารถทำได้โดยการเขียนโปรแกรมให้คอมพิวเตอร์สร้างเสียงตามที่ต้องการ ตัวอย่างเช่น Csound หรือ MATLAB® หรืออาจใช้โปรแกรมสำเร็จรูปในการสังเคราะห์เสียงก็ได้ เช่น Audio Architecture หรือ Grain Wave

ในการสังเคราะห์เสียงดนตรีโดยใช้คอมพิวเตอร์นี้ จำเป็นต้องคำนึงถึงปัจจัยในด้านต่าง ๆ ที่เกี่ยวข้องกับการสังเคราะห์เสียง เพื่อที่จะทำให้ได้วิธีการสังเคราะห์เสียงที่เหมาะสมกับความต้องการ [9] โดยมีรายละเอียดดังนี้

1) คุณภาพของเสียง เสียงสังเคราะห์ต้องมีคุณภาพดี ซึ่งหมายถึงเสียงที่มีความเป็นธรรมชาติ

2) ความยืดหยุ่น มีความสามารถที่จะแก้ไขดัดแปลงเสียงได้โดยเปลี่ยนแปลงค่าพารามิเตอร์ที่ใช้อยู่ ซึ่งในกรณีนี้วิธีการแซมปลิงมีความยืดหยุ่นน้อย แต่หากเป็นการสังเคราะห์เสียงแบบเอฟเอ็ม (FM synthesis) จะมีความยืดหยุ่นมากกว่า

3) การใช้งานได้อย่างกว้างขวาง สามารถนำไปใช้ในการสังเคราะห์เสียงได้หลายประเภท ไม่จำกัดเฉพาะเสียงใดเสียงหนึ่ง

4) ความซับซ้อนในการคำนวณ เป็นตัวกำหนดความต้องการทรัพยากรซึ่งได้จากการนับจำนวนคำสั่งที่ใช้ในการสังเคราะห์สัญญาณเสียงใน 1 แซมเปิล

ในการเลือกประยุกต์ใช้วิธีการสังเคราะห์เสียงต่าง ๆ ควรให้ความสำคัญกับปัจจัยที่กล่าวมาข้างต้น นอกจากนี้ยังต้องคำนึงถึงความเหมาะสมกับลักษณะงาน เพื่อให้แนวทางการสังเคราะห์เสียงดนตรีที่ถูกนำไปใช้จะทำให้เกิดประโยชน์สูงสุดได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 วิธีการสังเคราะห์เสียงดนตรี

แนวทางการสังเคราะห์เสียงที่สำคัญในปัจจุบันจำแนกได้ดังนี้

2.2.1 การสังเคราะห์เสียงโดยใช้วิธีการแซมปลิงหรือเวปเทเบิล (Sampling/ Wavetable Synthesis)

ซินทีไซเซอร์ (synthesizer) ที่ใช้วิธีการแซมปลิง จะเก็บบันทึกเสียงดนตรีจากเสียงดนตรีจริงในแต่ละตัวโน้ตไว้โดยจัดเก็บในรูปแบบของคลื่นเสียง ดังนั้นเมื่อต้องการสังเคราะห์เสียง ข้อมูลเหล่านี้จะถูกดึงออกมาแปลงเป็นเสียงดนตรี วิธีการดังกล่าวจึงมีความซับซ้อนในการคำนวณน้อย แต่อย่างไรก็ดีซินทีไซเซอร์ชนิดนี้ขาดความยืดหยุ่นในบางประการ เช่น การปรับเพิ่ม/ลดองค์ประกอบเชิงความถี่ไม่สามารถกระทำได้

2.2.2 การสังเคราะห์เสียงด้วยแบบจำลองเชิงกายภาพ (Physical Synthesis) [4] [14]

การสังเคราะห์เสียงด้วยแบบจำลองเชิงกายภาพ เป็นเทคนิคการสังเคราะห์เสียงโดยวิเคราะห์ลักษณะทางกายภาพของเครื่องดนตรีชนิดนั้น ๆ ว่าให้เสียงได้อย่างไร โดยจะต้องคำนึงถึงปัจจัยต่าง ๆ ที่เป็นองค์ประกอบทางกายภาพของเครื่องดนตรีซึ่งมีผลต่อการเกิดเสียง แล้วนำมาสร้างแบบจำลองทางคณิตศาสตร์เพื่อเลียนแบบการสร้างเสียงของเครื่องดนตรีชนิดนั้น ๆ การสังเคราะห์เสียงด้วยวิธีนี้ต้องอาศัยข้อมูลเชิงฟิสิกส์ของเสียงสำหรับเครื่องดนตรีแต่ละชนิด ทำให้โมเดลที่ใช้กับเครื่องดนตรีชนิดหนึ่งไม่สามารถนำมาใช้กับเสียงดนตรีชนิดอื่นได้ แต่อย่างไรก็ดีเสียงสังเคราะห์ที่ได้จะมีความเป็นธรรมชาติเนื่องจากค่าพารามิเตอร์ที่ใช้ในการสังเคราะห์เสียงนั้นเป็นค่าคุณสมบัติทางกายภาพของเครื่องดนตรีจริง

2.2.3 การสังเคราะห์เสียงด้วยแบบจำลองเชิงสเปกตรัม (Spectral Synthesis)

การสังเคราะห์เสียงเชิงสเปกตรัมได้อาศัยทฤษฎีของฟูเรียร์ที่กล่าวว่า ฟังก์ชันใด ๆ สามารถเขียนได้ในรูปของผลรวมของฟังก์ชันไซน์หรือโคไซน์ โดยมีเงื่อนไขว่าจะต้องมีจำนวนองค์ประกอบของความถี่ (frequency component) ที่มากพอจึงจะให้ค่าที่ใกล้เคียงกับฟังก์ชันนั้น ๆ

เทคนิคการสร้างเสียงแบบจำลองเชิงสเปกตรัมสามารถลดข้อจำกัดของการสังเคราะห์เสียงด้วยวิธีการแซมปลิงในแง่ความยืดหยุ่นได้ เช่นการปรับแต่งหรือสร้างเสียงสังเคราะห์ใหม่จากเสียงเดิม โดยแบบจำลองที่จัดอยู่ในแนวทางดังกล่าวมีอยู่ 2 วิธี คือ [10]

2.2.3.1 Sinusoidal Model [7]

เทคนิคการสร้างเสียงด้วยแบบจำลอง Sinusoidal Model ได้โมเดลสัญญาณเสียงเป็นผลรวมของสัญญาณไซน์ที่ความถี่ต่าง ๆ ตามสมการที่ 2.1 ดังนั้นการใช้แนวทางนี้ในการสังเคราะห์เสียงจะเป็นการลดข้อจำกัดของวิธีการแซมปลิงที่มีอยู่ นอกจากนี้แล้ววิธีการดังกล่าวสามารถแก้ไขดัด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แปลงเสียงได้ง่าย เช่น การทำเสียงสั้น/ยาว หรือแม้แต่การสร้างเสียงแปลกใหม่โดยเปลี่ยนแปลงค่าในเชิงความถี่ก็สามารถทำได้

$$s(n) = \sum_{r=1}^{R} A_r(n)\cos\left(\theta_r(n)\right) \tag{2.1}$$

โดยที่ $s(n)$ คือสัญญาณที่ได้

$A_r(n)$ คือค่าแอมปลิจูดของสัญญาณลำดับที่ r

$\theta_r(n)$ คือค่าความถี่เชิงมุมและค่าเฟสของสัญญาณไซน์ลำดับที่ r

n คือลำดับที่ของจุดข้อมูล

R คือจำนวนสัญญาณไซน์ทั้งหมดที่ใช้ในการสังเคราะห์เสียง

2.2.3.2 Sinusoidal Plus Noise Model [7] [11]

ในกรณีเสียงที่มีส่วนประกอบของความถี่ในช่วงกว้าง (อันเป็นลักษณะของเสียงดนตรีส่วนใหญ่) การสังเคราะห์เสียงด้วยวิธี Sinusoidal Model จำเป็นต้องใช้จำนวนองค์ประกอบของสัญญาณความถี่เป็นจำนวนมากเพื่อให้ได้เสียงที่ใกล้เคียงกับเสียงจริง ทำให้ต้องอาศัยค่าพารามิเตอร์ต่าง ๆ เป็นจำนวนมากในการสังเคราะห์เป็นผลให้มีความซับซ้อนในการคำนวณมากขึ้น การสังเคราะห์เสียงแบบ Sinusoidal Plus Noise Model ถูกพัฒนาขึ้นเพื่อลดข้อจำกัดดังกล่าว [10] โดยแยกการจำลองสัญญาณเสียงออกเป็นองค์ประกอบของความถี่ที่มีค่าแอมปลิจูดสูงและองค์ประกอบของความถี่ที่มีค่าแอมปลิจูดต่ำ (ในที่นี้เรียกว่าสัญญาณรบกวน) ดังแสดงในรูปที่ 2.1 ซึ่งสามารถเขียนให้อยู่ในรูปของสมการทางคณิตศาสตร์ได้ดังนี้

$$s(n) = \sum_{r=1}^{R} A_r(n)\cos(\theta_r(n)) + e(n) \tag{2.2}$$

โดยที่ $s(n)$ คือสัญญาณเสียง (input signal)

$A_r(n)$ และ $\theta_r(n)$ คือค่าแอมปลิจูดและค่าเฟสของสัญญาณไซน์ลำดับที่ r

$e(n)$ คือส่วนของสัญญาณรบกวนซึ่งสามารถคำนวณหาได้จากสมการต่อไปนี้

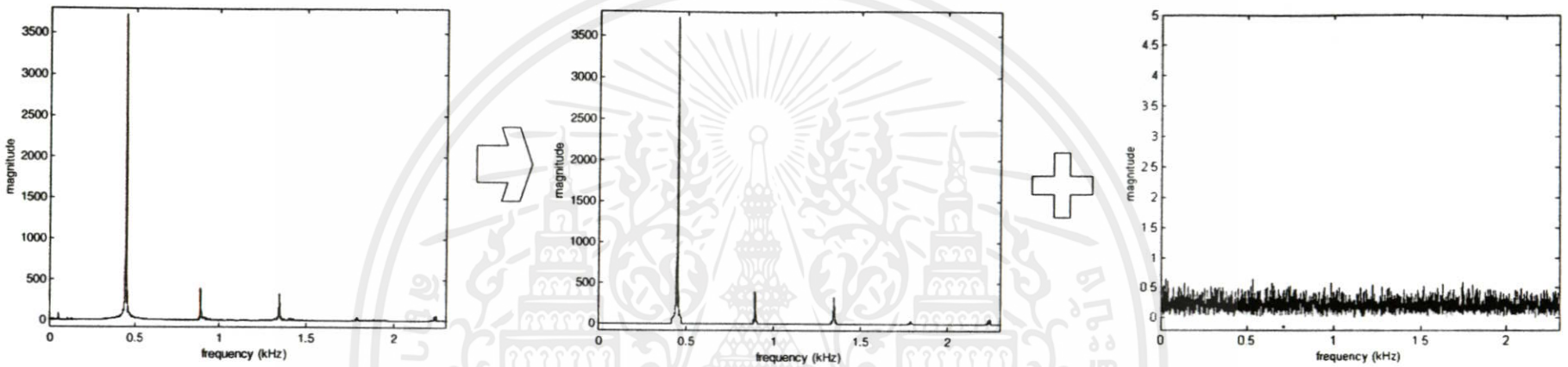
$$e(n) = \sum_{m=0}^{N-1} b(n,m)x(n-m) \tag{2.3}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ $x(n)$ คือสัญญาณ white noise

$b(n,m)$ คือค่าสัมประสิทธิ์ตัวกรองสัญญาณ (filter coefficient) ลำดับที่ m ณ เวลา n

N คือจำนวนสัมประสิทธิ์ตัวกรองสัญญาณที่ใช้



(ก) (ข) (ค)

รูปที่ 2.1 แสดงตัวอย่างการแยกองค์ประกอบของสัญญาณเสียงออกเป็นสองส่วน

ก. สเปกตรัมของสัญญาณเสียงที่ทำการวิเคราะห์

ข. สเปกตรัมของส่วนที่เป็นสัญญาณคลื่นไซน์

ค. สเปกตรัมของส่วนที่เป็นสัญญาณเสียงรบกวน (noise)

โดยทั่วไปเสียงพูดหรือเสียงดนตรีจะมีเสียงเหล่านี้เป็นองค์ประกอบอยู่ด้วย ดังนั้นการนำเทคนิคนี้มาประยุกต์ใช้ในการสังเคราะห์เสียงดนตรีจะช่วยลดความซับซ้อนในการคำนวณ อีกทั้งจำนวนพารามิเตอร์ที่ใช้ในการจัดเก็บก็ลดลงอีกด้วย เนื่องจากว่าการสังเคราะห์เสียงไม่จำเป็นต้องใช้องค์ประกอบของสัญญาณในทุกความถี่ โดยในการสังเคราะห์เสียงในลักษณะนี้จะกระทำโดยใช้องค์ประกอบของสัญญาณ ณ ความถี่ที่มีแอมปลิจูดชัดเจนเท่านั้น แล้วนำมารวมกับสัญญาณเสียงรบกวนที่สร้างขึ้นให้มีคุณสมบัติเชิงสถิติที่เหมาะสม (อันจะได้กล่าวถึงต่อไป) ประกอบกันเป็นเสียงสังเคราะห์ที่ต้องการ

2.2.4 เปรียบเทียบแนวทางการสังเคราะห์เสียงดนตรี

วิธีการสังเคราะห์เสียงแต่ละวิธีการที่กล่าวมาแล้วข้างต้นนั้นมีข้อดีและข้อเสียที่แตกต่างกัน โดยสามารถสรุปได้ดังตารางที่ 2.1 [3] [12]

ตารางที่ 2.1 แสดงคุณสมบัติของการสังเคราะห์เสียงในแต่ละวิธี

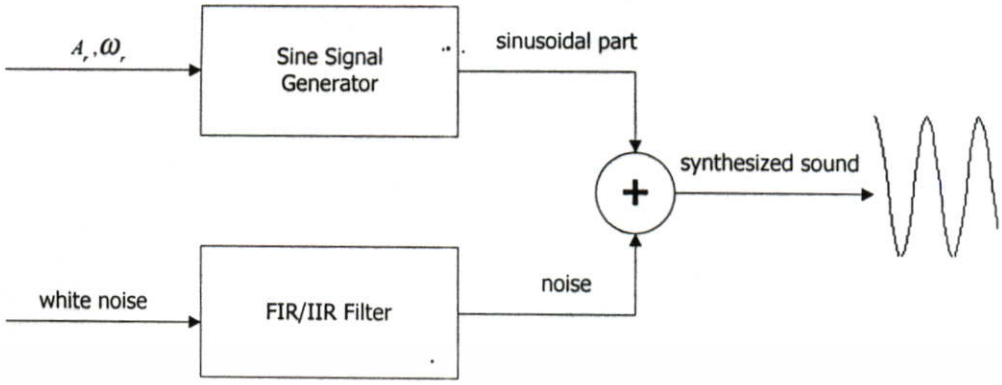
วิธีการสังเคราะห์	แซมปลิง (sample playback)	สเปกตรัมโมเดล (spectral model)	ฟิสิกคอลโมเดล (physical model)
คุณภาพ (Quality)	+	+	+
ความยืดหยุ่น (Flexibility)	-	+	0
การใช้งานได้อย่างกว้างขวาง (Generality)	+	+	-
เวลาที่ใช้ในการคำนวณ (Computing time)	+	- -	-
ความง่ายในการโปรแกรม (Programming ease)	+	-	- -
ปริมาณข้อมูล (Amount of data)	-	0	+
การควบคุมแบบอินเทอแอคทีฟ (Interactive control)	-	+	+

+ ให้ความหมายในเชิงบวก - ให้ความหมายในเชิงลบ 0 ให้ความหมายเป็นกลาง

2.3 หลักการสังเคราะห์เสียงด้วยแบบจำลอง Sinusoidal Plus Noise

เนื่องจากว่าการสังเคราะห์เสียงแนวทางนี้ได้โมเดลสัญญาณออกเป็นสองส่วนคือส่วนสัญญาณไซน์และสัญญาณรบกวน ซึ่งในการสังเคราะห์เสียงจะต้องสร้างส่วนของสัญญาณไซน์โดยใช้วิธีการสังเคราะห์แบบแอดดิทีฟ (additive synthesis) และสร้างส่วนของสัญญาณรบกวนโดยทำการสร้าง white noise ผ่านฟิลเตอร์จะได้สัญญาณรบกวน แล้วนำไปรวมกับสัญญาณไซน์จะได้สัญญาณเสียงสังเคราะห์ที่ต้องการ (ดูรูปที่ 2.2 ประกอบ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

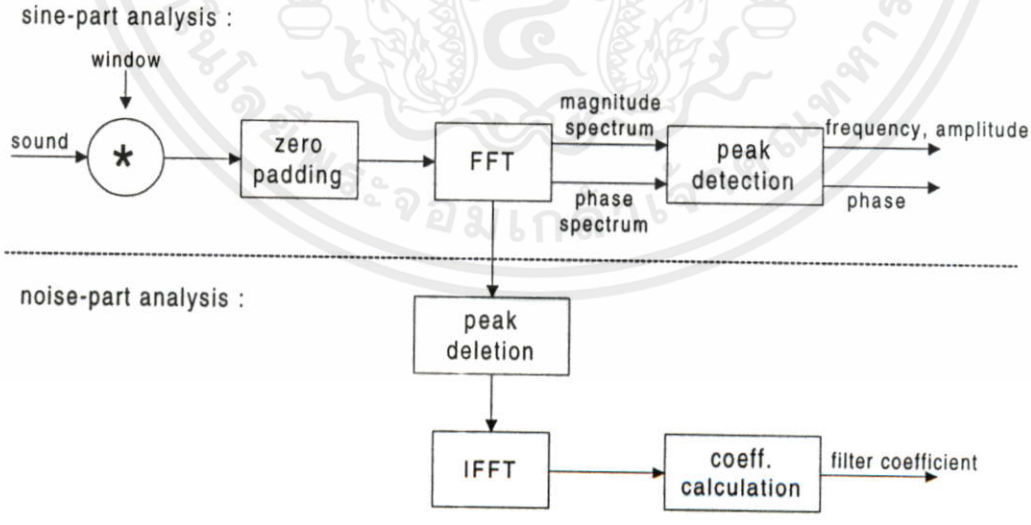


รูปที่ 2.2 แสดงไดอะแกรมขั้นตอนการสังเคราะห์เสียง

จากรูปที่ 2.2 จะเห็นได้ว่าในการสังเคราะห์เสียงจำเป็นต้องอาศัยค่าพารามิเตอร์ต่าง ๆ ทั้งในส่วนของสัญญาณไซน์และสัญญาณรบกวน โดยส่วนของสัญญาณไซน์ต้องการค่าความถี่และค่าแอมปลิจูด และส่วนของสัญญาณรบกวนต้องการสัมประสิทธิ์ตัวกรองสัญญาณ ซึ่งค่าพารามิเตอร์เหล่านี้ได้จากการวิเคราะห์สัญญาณเสียงต้นแบบ โดยรายละเอียดของการวิเคราะห์และสังเคราะห์เสียงด้วยวิธีการสังเคราะห์เสียงแบบ Sinusoidal Plus Noise มีดังนี้

2.3.1 ขั้นตอนของการวิเคราะห์สัญญาณเสียง

ในขั้นตอนนี้จะแยกการวิเคราะห์ออกเป็นสองส่วนคือ ส่วนของการวิเคราะห์สัญญาณไซน์และส่วนของสัญญาณรบกวน ดังแสดงรายละเอียดในรูปที่ 2.3



รูปที่ 2.3 แสดงไดอะแกรมขั้นตอนการวิเคราะห์สัญญาณเสียงด้วยวิธี Sinusoidal Plus Noise

เอกสารนี้ใช้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.1.1 ขั้นตอนการวิเคราะห์สัญญาณไซน์

การวิเคราะห์สเปกตรัมของสัญญาณเสียงนี้จะใช้วิธี STFT (Short Time Fourier Transform) หรือวิธีการแบ่งสัญญาณเสียงออกเป็นช่วงย่อย ๆ ซึ่งจะทำให้ได้ค่าสเปกตรัมที่มีการเปลี่ยนแปลงไปตามเวลา [11] แต่การได้มาซึ่งค่าความถี่และแอมปลิจูดที่ถูกต้องจะขึ้นอยู่กับปัจจัย 3 ประการคือ ชนิดของวินโดว์ที่นำมาคูณกับสัญญาณเสียงจริง ขนาดของวินโดว์ของสัญญาณ และการเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์เข้ากับสัญญาณที่วิเคราะห์ โดยมีรายละเอียดดังนี้

1) ชนิดของวินโดว์

การเลือกใช้ชนิดของวินโดว์นั้นมีสิ่งที่ต้องพิจารณามีอยู่ 2 ประการ [8] คือ ความกว้างของเมนโลบ (mainlobe) และ ไซด์โลบ (sidelobe) ที่สูงที่สุด ซึ่งในทางอุดมคตินั้น เมนโลบควรมีขนาดที่แคบ จึงจะทำให้ได้ค่าความถี่ที่มีความถูกต้องมากขึ้น และไซด์โลบควรมีค่าต่ำ แต่โดยทั่วไปแล้วไม่มีวินโดว์ชนิดใดที่มีลักษณะดังกล่าว ดังนั้นในการเลือกใช้ชนิดของวินโดว์จะต้องเลือกพิจารณาระหว่างวินโดว์ที่มีเมนโลบที่กว้างแต่มีไซด์โลบที่ต่ำ หรือมีเมนโลบที่แคบแต่มีไซด์โลบที่สูง ทั้งนี้ขึ้นอยู่กับลักษณะสเปกตรัมสัญญาณที่นำมาวิเคราะห์ด้วย โดยในการวิเคราะห์เชิงสเปกตรัมเพื่อให้ได้ค่าที่ถูกต้องมากขึ้นจะต้องทำการคูณสัญญาณที่นำมาวิเคราะห์ด้วยวินโดว์ฟังก์ชันในโดเมนของเวลาก่อนที่จะทำการวิเคราะห์เชิงสเปกตรัม

2) ขนาดของวินโดว์

ขนาดของวินโดว์ที่เหมาะสมจะมีผลทำให้การหาค่าความถี่หลักและความถี่ฮาร์โมนิกมีความถูกต้องมากขึ้น ซึ่งขนาดของวินโดว์เริ่มต้นที่เหมาะสมนั้นมีรูปแบบตามสมการต่อไปนี้ [5]

$$M \geq \frac{B \cdot f_s}{f_o} \tag{2.4}$$

โดยที่ M คือขนาดของวินโดว์ มีหน่วยเป็นแซมเปิล

B คือเมนโลบของวินโดว์ที่ใช้ มีหน่วยเป็นแซมเปิล

f_s คือความถี่แซมเปิล

f_o คือความถี่พื้นฐาน (fundamental frequency)

ทั้งนี้ค่า B ขึ้นอยู่กับวินโดว์ฟังก์ชันที่นำมาคูณกับสัญญาณเสียงก่อนการวิเคราะห์เชิงสเปกตรัม

3) การเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์เข้ากับสัญญาณที่วิเคราะห์ (zero padding) [5] [8]

เนื่องจากว่าการคำนวณหาสเปกตรัมของสัญญาณใช้วิธีการดีสครีตฟูเรียร์ทรานฟอร์ม ซึ่งเป็นการประมาณค่าความถี่ตามจำนวนแซมเปิลของสัญญาณที่นำมาวิเคราะห์ โดยจะประมาณค่าแต่ละจุดข้อมูลแซมเปิลด้วยระยะห่างที่เท่ากัน ดังนั้นหากใช้ขนาดของสัญญาณที่มีปริมาณมากขึ้น จะทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ให้มีค่าความถี่ที่ทำการประมาณเพิ่มขึ้น อันเป็นผลให้ได้ค่าความถี่ที่ความละเอียดมากขึ้น นั่นคือได้ค่าสเปกตรัมที่ถูกต้องมากขึ้น แต่ในความเป็นจริงแล้วไม่มีสัญญาณเสียงใด ๆ ที่มีความยาวมากพอสำหรับการวิเคราะห์เชิงสเปกตรัมที่ให้ผลได้ถูกต้อง ด้วยข้อจำกัดดังกล่าวจึงต้องทำการเพิ่มจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์ให้กับสัญญาณเสียงเพื่อเป็นการเพิ่มค่าความละเอียดของค่าความถี่ (frequency resolution) ให้ได้ลักษณะสเปกตรัมที่มีความถูกต้องมากขึ้น โดยการเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์ควรจะต้องอยู่ภายใต้เงื่อนไขดังต่อไปนี้ [8]

3.1) N ควรมีขนาดเท่ากับ 2 เท่าของขนาดวินโดว์ M เป็นอย่างน้อย

3.2) N/M ที่ค่ามากจะทำให้ค่าพารามิเตอร์ถูกต้องมากขึ้น

สำหรับการเพิ่มจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์จะทำการเพิ่มที่ช่วงไหนของสัญญาณเสียงก็ได้ แต่ในงานวิจัยนี้จะทำการเพิ่มจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์ต่อท้ายสัญญาณเสียงที่นำมาวิเคราะห์เพื่อความสะดวกในการนำสัญญาณดังกล่าวมาวิเคราะห์เชิงสเปกตรัมต่อไป

ขั้นตอนการวิเคราะห์ในส่วนของสัญญาณไซน์ [5] มีรายละเอียดดังนี้

1) แบ่งสัญญาณเสียงออกเป็นเฟรมย่อย ๆ โดยกำหนดให้มีขนาดเท่ากับ M แซมเปิล (M มีค่าเป็นไปตามสมการที่ 2.4)

2) นำสัญญาณเสียงในทุก ๆ เฟรมที่ได้จากขั้นตอนที่ 1 มาคูณกับวินโดว์ฟังก์ชันที่มีขนาดเท่ากันโดยให้ (ดูรูปที่ 2.4 ประกอบ)

$$s_w(n) = s(n) \cdot w(n) \qquad t = 0,\ldots, M-1 \tag{2.5}$$

โดยที่ $s_w(n)$ คือสัญญาณที่คูณกับวินโดว์

$s(n)$ คือสัญญาณขนาด M แซมเปิล

$w(n)$ คือวินโดว์ฟังก์ชัน

3) เพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์ให้กับสัญญาณที่คูณกับวินโดว์ฟังก์ชันแล้ว และทำการคำนวณค่าสเปกตรัมของสัญญาณ เช่น โดยอาศัยอัลกอลิทึม FFT

4) ตรวจหาจุดที่มีค่าความชันสูง (peak) จากสเปกตรัมที่ได้ในแต่ละเฟรม (ดูรายละเอียดดูเอกสารอ้างอิง [5]) จากนั้นคำนวณหาค่าความถี่และแอมปลิจูดที่จุดนั้น ๆ สำหรับใช้ในกระบวนการสังเคราะห์เสียง

2.3.1.2 การวิเคราะห์ในส่วนของสัญญาณรบกวน

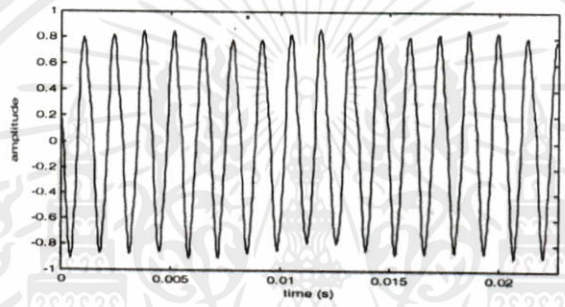
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัญญาณรบกวนที่กล่าวถึงในงานวิจัยนี้ เป็นลักษณะของสัญญาณที่ไม่สามารถระบุค่าได้เมื่อเวลาเปลี่ยนไป แต่สามารถที่จะบ่งบอกด้วยคุณสมบัติทางสถิติ โดยคุณสมบัติทางสถิติได้จากการวิเคราะห์ค่าสัมประสิทธิ์ตัวกรองสัญญาณ ดังนั้นการฟิลเตอร์สัญญาณ white noise ด้วยค่า

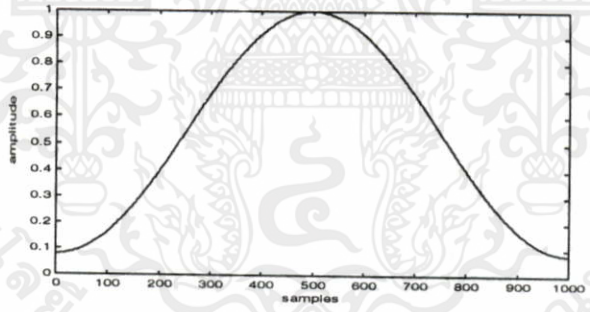
สัมประสิทธิ์ตัวกรองสัญญาณเหล่านี้จะสามารถสร้างสัญญาณรบกวนที่มีคุณสมบัติทางสถิติเหมือนกันได้ ซึ่งโมเดลของสัญญาณรบกวนมีอยู่ 2 วิธีคือ เอ็มเอโมเดลซึ่งเป็นโมเดลที่ใช้ในงานวิจัยที่มีอยู่ในปัจจุบันและเออาร์โมเดลซึ่งเป็นโมเดลที่งานวิจัยนี้นำมาศึกษาเปรียบเทียบ

1) เอ็มเอโมเดล (Moving Average Model)

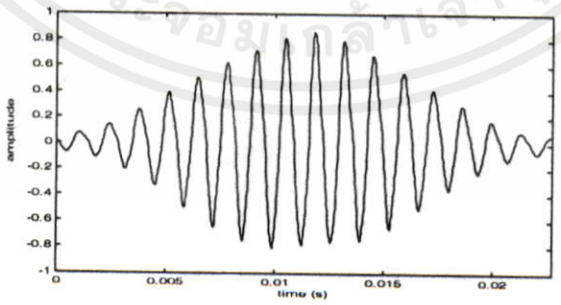
เอ็มเอโมเดล เป็นแบบจำลองสัญญาณรบกวนที่ใช้ FIR Filter ซึ่งมีโครงสร้างตามรูปที่ 2.5 โดยสัญญาณเอาต์พุต $y(n)$ คือสัญญาณรบกวนที่เกิดจากการคอนโวลูชัน (convolution) สัญญาณ white noise (สัญญาณอินพุต $x(n)$) ด้วยค่าสัมประสิทธิ์ตัวกรองสัญญาณ ($b(i)$) โดยสามารถเขียนเป็นสมการทางคณิตศาสตร์ดังสมการ 2.6 [13]



(ก)



(ข)



(ค)

รูปที่ 2.4 แสดงสัญญาณเสียงคูณกับวินโดว์ฟังก์ชัน
ก.สัญญาณเสียงที่ต้องการวิเคราะห์
ข.แฮมมิงวินโดว์
ค.สัญญาณเสียงที่คูณกับแฮมมิงวินโดว์

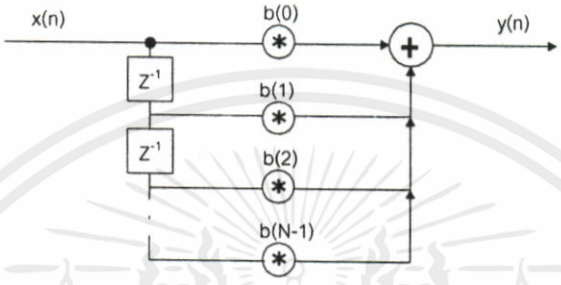
เอกสารนี้ใช้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$y(n)=\sum_{i=0}^{N-1} x(n-i)b(i) \tag{2.6}$$

โดยที่ $y(n)$ คือสัญญาณรบกวน

$x(n)$ คือสัญญาณ white noise

$b(i)$ คือสัมประสิทธิ์ตัวกรองสัญญาณซึ่งมีจำนวน N ค่า



รูปที่ 2.5 แสดงโครงสร้างของ เอ็มเอโมเดล

2) เออาร์โมเดล (Autoregressive Model)

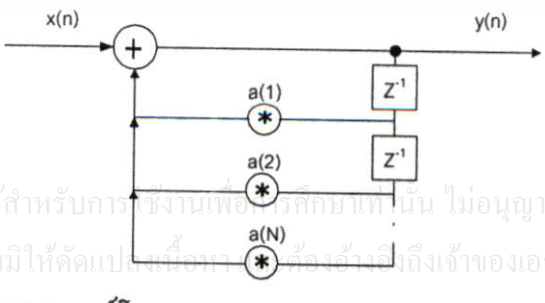
เออาร์โมเดลเป็นแบบจำลองสัญญาณรบกวนที่ใช้ IIR Filter ซึ่งมีโครงสร้างดังรูปที่ 2.6 โดยสัญญาณเอาต์พุต $y(n)$ คือสัญญาณรบกวนที่เกิดจากการกรองสัญญาณ white noise (สัญญาณอินพุต $x(n)$) ด้วยฟิลเตอร์แบบ IIR ที่มีสัมประสิทธิ์ตัวกรองสัญญาณ ($a(i)$) โดยสามารถเขียนเป็นสมการทางคณิตศาสตร์ได้ดังนี้ [13]

$$y(n)=\sum_{i=1}^{N} a(i)y(n-i)+x(n) \tag{2.7}$$

โดยที่ $y(n)$ คือสัญญาณรบกวน

$x(n)$ คือสัญญาณ white noise

$a(i)$ คือสัมประสิทธิ์ตัวกรองสัญญาณซึ่งมีจำนวน N ค่า

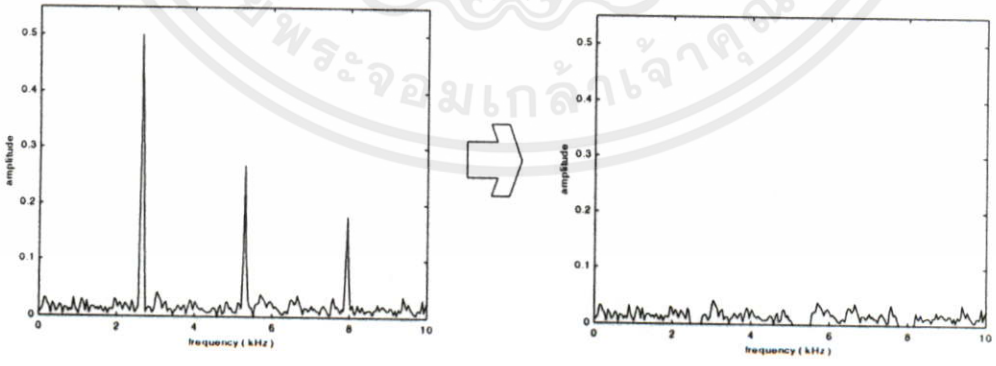


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.6 แสดงโครงสร้างของเออาร์โมเดล

วิธีการคำนวณหาสัญญาณรบกวนจากสัญญาณเสียงโดยการลบส่วนของสัญญาณไซน์ออกจากสัญญาณเสียงนั้น สามารถทำได้ทั้งในโดเมนของเวลา (time domain) และโดเมนของความถี่ (frequency domain) ในกรณีโดเมนของเวลานั้นสามารถกระทำได้โดยนำสัญญาณเสียงจริงลบส่วนของสัญญาณไซน์หรือไซน์นูซอยด์ที่สังเคราะห์ขึ้นใหม่ [11] แต่ในงานวิจัยนี้ได้นำเสนอวิธีการคำนวณหาสัญญาณรบกวนในโดเมนของความถี่ เนื่องจากว่าค่าพารามิเตอร์ต่าง ๆ (ค่าความถี่ และค่าแอมปลิจูด) ที่ได้จากการวิเคราะห์เชิงสเปกตรัมไม่ใช่ค่าที่ถูกต้องจริง ๆ เมื่อนำมาคำนวณหาสัญญาณรบกวนในโดเมนของเวลาตามที่งานวิจัยเดิมเสนอมา จะทำให้ได้ส่วนของสัญญาณรบกวนที่ยังคงมีส่วนของสัญญาณไซน์อยู่ สำหรับการคำนวณสัญญาณรบกวนในโดเมนของความถี่ที่ใช้ในงานวิจัยนี้สามารถทำได้โดยกำหนดค่าสเปกตรัมที่มีค่าความชันสูงให้มีค่าเป็นศูนย์เพื่อแยกส่วนของสัญญาณไซน์ออกจากสัญญาณรบกวน ซึ่งอาจทำให้มีการตัดส่วนของสัญญาณรบกวนที่อยู่บริเวณที่มีค่าความชันสูงไปด้วย และส่งผลให้ได้ส่วนของสัญญาณรบกวนที่ไม่ใช่สัญญาณรบกวนจริง อย่างไรก็ดีสัญญาณรบกวนที่ได้จากวิธีการข้างต้นน่าจะมีความใกล้เคียงสัญญาณรบกวนจริง เนื่องจากว่าส่วนของสเปกตรัมที่ทำการกำหนดค่าเท่ากับศูนย์เป็นเพียงส่วนน้อยเมื่อเทียบกับขนาดของสเปกตรัมทั้งหมด โดยรายละเอียดของการคำนวณหาสัญญาณรบกวนด้วยวิธีดังกล่าวมีดังนี้

1) กำหนดค่าในช่วงของสเปกตรัมที่มีค่าความชันสูงให้มีค่าเป็นศูนย์ (ดูรูปที่ 2.7 ประกอบ) จะได้สเปกตรัมในส่วนของสัญญาณรบกวน
2) สร้างสัญญาณเสียงของส่วนที่เป็นสัญญาณรบกวน โดยการหาอินเวอร์สฟูเรียร์ทรานฟอร์ม (Inverse Fourier Transform) ของสเปกตรัมที่ได้ตัดส่วนที่เป็นจุดสูงสุด (peak) ออกไปแล้ว
3) จากนั้นนำสัญญาณรบกวนในโดเมนของเวลาตัดจำนวนข้อมูลที่มีค่าเป็นศูนย์ออกเท่ากับจำนวนข้อมูลศูนย์ที่ทำการเพิ่มเข้ากับสัญญาณในส่วนของการวิเคราะห์สัญญาณไซน์



(ก) (ข)

รูปที่ 2.7 แสดงตัวอย่างสเปกตรัมของสัญญาณ
ก. สเปกตรัมของสัญญาณเสียงที่ต้องการวิเคราะห์
ข. สเปกตรัมของสัญญาณเสียงที่ตัดส่วนของสัญญาณไซน์ออกแล้ว

เอกสารนี้ใช้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

4) นำสัญญาณที่ทำการตัดข้อมูลศูนย์แล้วซึ่งจะมีขนาดของสัญญาณเท่ากับขนาดสัญญาณต้นแบบที่นำมาวิเคราะห์มาหารด้วยวินโดว์ฟังก์ชันที่ขนาดเท่ากันแบบจุดต่อจุด (วินโดว์ฟังก์ชันต้องเป็นวินโดว์ฟังก์ชันชนิดเดียวกับที่ใช้ในการวิเคราะห์ส่วนของสัญญาณไซน์)

5) จะได้สัญญาณรบกวนที่สามารถนำมาวิเคราะห์หาสัมประสิทธิ์ตัวกรองสัญญาณ (filter coefficient) ในแต่ละโมเดลต่อไป

2.3.1.2.1 การวิเคราะห์ค่าสัมประสิทธิ์ตัวกรองสัญญาณสำหรับเอ็มเอโมเดล

ขั้นตอนการวิเคราะห์สัมประสิทธิ์ตัวกรองสัญญาณสำหรับเอ็มเอโมเดลมีดังนี้

1) เริ่มจากคำนวณหาค่าเพาเวอร์สเปกตรัมของสัญญาณรบกวน ซึ่งมีขั้นตอนดังนี้

1.1) แบ่งสัญญาณเสียงออกเป็นวินโดว์ที่มีขนาดเท่ากับ M แซมเปิล โดยให้แต่ละวินโดว์เหลื่อมกัน (overlap) $M/2$ แซมเปิล

1.2) คำนวณหาฟูเรียร์ทรานฟอร์มในแต่ละวินโดว์

1.3) นำค่าฟูเรียร์ทรานฟอร์มทั้งหมดมาเฉลี่ยด้วยจำนวนวินโดว์ดังสมการต่อไปนี้

$$P(\omega)=\frac{1}{K}\sum_{i=1}^{K}\left|F_i(\omega)\right|^2 \qquad (2.8)$$

โดยที่ $P(\omega)$ คือเพาเวอร์สเปกตรัม

K คือจำนวนวินโดว์ทั้งหมด

$F_i(\omega)$ คือฟูเรียร์ทรานฟอร์มลำดับที่ i

2) พิจารณาหาค่าสัมประสิทธิ์ตัวกรองสัญญาณในสมการที่ 2.8

$$P(\omega)=\{H(\omega)\cdot H^{*}(\omega)\}\cdot\sigma^2 \qquad (2.9)$$

โดยที่ $H(\omega)$ คือ ฟูเรียร์ทรานฟอร์มของสัมประสิทธิ์ตัวกรองสัญญาณ

เนื่องจาก white noise ที่ทำการสร้างขึ้นมาเพื่อผ่านฟิลเตอร์นั้นเป็นเลขสุ่ม (random number) ที่มีคุณสมบัติการกระจายแบบปกติ (normal distribution) ซึ่งมีค่า σ^2 เท่ากับ 1 จะได้

$$P(\omega) = H(\omega)\cdot H^{*}(\omega) \qquad (2.10)$$

จากสมการที่ 2.10 มีค่า $H(\omega)$ หลายค่าที่ทำให้ $P(\omega)$ เท่ากับผลคูณของ $H(\omega)$ กับ $H^{*}(\omega)$ แต่ในที่นี้จะเลือกค่า $H(\omega)$ ดังสมการต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$H(\omega)=\sqrt{P(\omega)} \qquad (2.11)$$

40778

จากสเปกตรัมของสัมประสิทธิ์ตัวกรองสัญญาณ สามารถหาค่าสัมประสิทธิ์ตัวกรองสัญญาณโดยการทำอินเวอร์สฟูเรียร์ทรานฟอร์มดังสมการ 2.12

$$b(n)=\frac{1}{2\pi}\int_{-\pi}^{\pi}H(\omega)e^{j\omega_n}d\omega \tag{2.12}$$

โดยที่ $b(n)$ คือค่าสัมประสิทธิ์ตัวกรองสัญญาณ

2.3.1.2.2 การวิเคราะห์ค่าสัมประสิทธิ์ตัวกรองสัญญาณสำหรับเออาร์โมเดล

จากสมการที่ 2.7 $y(n)$ เป็นสัญญาณรบกวนต้นแบบที่นำมาวิเคราะห์ $x(n)$ คือสัญญาณ white noise และ a_i คือค่าสัมประสิทธิ์ตัวกรองสัญญาณที่ต้องการ จากสมการดังกล่าวนำมาคูณด้วย $y(n-1)$ แล้วคำนวณค่าความคาดหวัง (Expectation) จะได้สมการดังนี้

$$y(n)=\sum_{i=0}^{N}a_i y(n-i)+x(n)$$

$$E\{y(n)y(n-1)\} = E\left\{\left(\sum_{i=1}^{N}a_i y(n-i)+x(n)\right)y(n-1)\right\}$$

$$=E\left\{\left(\sum_{i=1}^{N}a_i y(n-i)\right)y(n-1)\right\}+E\{x(n)y(n-1)\} \tag{2.13}$$

จากสมการ 2.13 $x(n)$และ $y(n-1)$ ไม่มีความสัมพันธ์กัน ซึ่งทำให้ $E\{x(n)y(n-1)\}$[1] มีค่าเท่ากับศูนย์ จะได้สมการเขียนในรูปของเมตริกได้ดังนี้

$$E\{y(n)y(n-1)\} =\left[E\{y(n-1)y(n-1)\}\ \ldots\ E\{y(n-1)y(n-N)\}\right]\begin{bmatrix}a_1\\a_2\\\vdots\\a_N\end{bmatrix} \tag{2.14}$$

แต่เนื่องจาก $C_{yy}(m)\equiv E\{y(n)y(n-m)\}$ สำหรับค่า n ใด ๆ

ซึ่ง
$$C_{yy}(m)=\sum_{i=1}^{N}a_i C_{yy}(m-i) \tag{2.15}$$

โดยที่ C_{yy} คือสหความสัมพันธ์ (correlation)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[1] $E\{\ \}$คือ expectation function

m คือค่า time lag

เมื่อนำสมการที่ 2.14 มาเขียนในรูปของค่าสหสัมพันธ์ได้ดังนี้

$$C_1 = \begin{bmatrix} C_0 & C_1 & C_2 & C_3 & \cdots & C_{N-4} & C_{N-3} & C_{N-2} & C_{N-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} \tag{2.16}$$

จากสมการ 2.16 ไม่สามารถนำมาแก้สมการหาค่าสัมประสิทธิ์ตัวกรองสัญญาณ a_N ได้ จึงต้องทำการหาค่าสหสัมพันธ์ตั้งแต่ time lag ที่ 1 จนถึงที่ N (โดยที่ N คือจำนวนสัมประสิทธิ์ตัวกรองสัญญาณที่ต้องการ) ดังแสดงในสมการที่ 2.17

$$\begin{bmatrix} C_{yy}(1) \\ C_{yy}(2) \\ \vdots \\ C_{yy}(N) \end{bmatrix} = \begin{bmatrix} C_{yy}(0) & C_{yy}(1) & \cdots & C_{yy}(N-1) \\ C_{yy}(1) & C_{yy}(0) & \cdots & C_{yy}(N-2) \\ \vdots & \vdots & \vdots & \vdots \\ C_{yy}(N-1) & C_{yy}(N-2) & \cdots & C_{yy}(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} \tag{2.17}$$

กำหนดให้ $$B = \begin{bmatrix} C_{yy}(0) & C_{yy}(1) & \cdots & C_{yy}(N-1) \\ C_{yy}(1) & C_{yy}(0) & \cdots & C_{yy}(N-2) \\ \vdots & \vdots & \vdots & \vdots \\ C_{yy}(N-1) & C_{yy}(N-2) & \cdots & C_{yy}(0) \end{bmatrix} \tag{2.18}$$

คูณด้วย B^{-1} ทั้งสองฝั่งสมการที่ 2.17 จะสามารถคำนวณหาค่าสัมประสิทธิ์ตัวกรองสัญญาณ a_i

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} = B^{-1} \begin{bmatrix} C_{yy}(1) \\ C_{yy}(2) \\ \vdots \\ C_{yy}(N) \end{bmatrix} \tag{2.19}$$

ขั้นตอนการคำนวณหาค่าสัมประสิทธิ์ตัวกรองสัญญาณ

1) สัญญาณรบกวนที่ได้นำมาคำนวณค่าสหสัมพันธ์ ณ time lag ตั้งแต่ 0 ถึง N โดยที่ N เป็นจำนวนค่าสัมประสิทธิ์ตัวกรองสัญญาณที่ต้องการ

2) นำค่าที่ได้มาคำนวณค่าสัมประสิทธิ์ตัวกรองสัญญาณตามสมการที่ 2.19

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.2 ขั้นตอนการสังเคราะห์เสียง

ขั้นตอนการสังเคราะห์เสียงแยกพิจารณาเป็น 2 ส่วน คือส่วนของการสร้างสัญญาณไซน์และส่วนของสัญญาณรบกวน ซึ่งมีรายละเอียดดังนี้

1) การสร้างสัญญาณคลื่นไซน์ $s(n)$ นั้นจะใช้ค่าพารามิเตอร์ที่ได้จากขั้นตอนการวิเคราะห์ในส่วนของสัญญาณไซน์ เนื่องจากค่าเฟสไม่มีผลต่อการสังเคราะห์เสียงจึงมิได้นำค่าเฟสมาใช้ในการสังเคราะห์ตามสมการต่อไปนี้

$$s(n)=\sum_{r=1}^{R} A_r(n)\cos\left(\omega_r(n)\right) \tag{2.20}$$

โดยที่ $s(n)$ คือสัญญาณเสียง (input signal)

$A_r(n)$ คือค่าแอมปลิจูดของสัญญาณไซน์ลำดับที่ r

$\omega_r(n)$ คือค่าความถี่เชิงมุมของสัญญาณไซน์ลำดับที่ r ซึ่งมีหน่วยเป็นเรเดียน

2) การสร้างสัญญาณรบกวนนั้นจะใช้สัมประสิทธิ์ตัวกรองสัญญาณที่ได้จากการวิเคราะห์ในส่วนของสัญญาณรบกวน โดยการสังเคราะห์เสียงสัญญาณรบกวนด้วยวิธีเอ็มเอโมเดลและการสังเคราะห์สัญญาณรบกวนด้วยวิธีเออาร์โมเดล จะเป็นไปตามสมการที่ 2.19 และ 2.20 ตามลำดับ

$$e(n)=\sum_{i=0}^{N-1} w(n-i)b_i \tag{2.21}$$

โดยที่ b_i คือสัมประสิทธิ์ตัวกรองสัญญาณ (filter coefficient)

$w(n)$ คือสัญญาณ white noise

N คือจำนวนสัมประสิทธิ์ตัวกรองสัญญาณ

$$e(n)=\sum_{i=1}^{N} a_i e(n-i)+w(n) \tag{2.22}$$

โดยที่ a_i คือสัมประสิทธิ์ตัวกรองสัญญาณ (filter coefficient)

$w(n)$ คือสัญญาณ white noise

N คือจำนวนสัมประสิทธิ์ตัวกรองสัญญาณ

ในส่วนของการสร้างสัญญาณ white noise จะเป็นสร้างค่าตัวเลขสุ่มที่มีการกระจายแบบปกติ ซึ่งในงานวิจัยนี้ได้ใช้ฟังก์ชัน randn ใน MATLAB® ในการสร้างสัญญาณ white noise และสำหรับรายละเอียดในการสร้างส่วนของสัญญาณ white noise สามารถศึกษาเพิ่มเติมได้จากเอกสารอ้างอิง [6]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) จากนั้นนำสัญญาณที่สร้างขึ้นใน 2 ขั้นตอนข้างต้นมารวมกันแล้วคูณด้วยเอนวิลอปของสัญญาณต้นแบบ จะได้สัญญาณเสียงสังเคราะห์ที่ต้องการ ดังแสดงในรูปที่ 2.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

ข้อพิจารณาในการสังเคราะห์เสียงแบบ Sinusoidal Plus Noise Model

เนื่องจากการสังเคราะห์เสียงด้วยแบบจำลอง Sinusoidal Plus Noise สามารถนำประยุกต์ใช้ในการสังเคราะห์เสียงได้หลายประเภทและมีความยืดหยุ่นสูงในการดัดแปลงหรือสร้างสัญญาณเสียงขึ้นมาใหม่ได้ ดังนั้นงานวิจัยนี้จึงนำเอาแนวทางการสังเคราะห์เสียงดังกล่าวมาประยุกต์ใช้ในการศึกษาการวิเคราะห์และสังเคราะห์เสียงขลุ่ยไทย โดยจะทำการวิเคราะห์หาส่วนประกอบเชิงความถี่เพื่อหาค่าพารามิเตอร์ที่จะนำมาใช้ในการสังเคราะห์เสียง แต่เนื่องจากว่าขั้นตอนในการวิเคราะห์สัญญาณเสียงขลุ่ยไทย จำเป็นต้องศึกษาแนวทางการวิเคราะห์ข้อมูลที่ให้ค่าองค์ประกอบเชิงความถี่ใกล้เคียงความจริงมากที่สุด ตามที่ได้กล่าวไว้ในบทที่ 2 ค่าองค์ประกอบเชิงความถี่ที่ได้จากการวิเคราะห์จะมีความถูกต้องมากน้อยเพียงใดขึ้นอยู่กับชนิดของวินโดว์ที่นำมาคูณกับสัญญาณเสียง ขนาดของวินโดว์ และการเพิ่มจำนวนข้อมูลที่มีค่าเป็นศูนย์เข้ากับสัญญาณเสียง โดยจะกล่าวถึงรายละเอียดในส่วนถัดไป

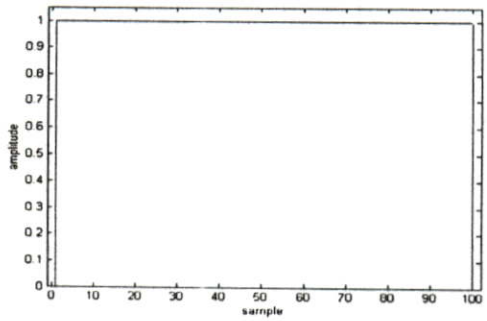
3.1 การพิจารณาชนิดวินโดว์ที่เหมาะสม

ชนิดของวินโดว์ที่นำมาคูณกับสัญญาณเสียง จะมีผลต่อความถูกต้องของค่าองค์ประกอบเชิงความถี่ที่วัดได้จากสเปกตรัมของสัญญาณเสียง ซึ่งค่าความถูกต้องจากการใช้วินโดว์แต่ละชนิดยังขึ้นอยู่กับระยะห่างระหว่างสองค่าความถี่ใดๆ (Δ) ของสัญญาณเสียงที่นำมาวิเคราะห์อีกด้วย จากการวิเคราะห์สัญญาณเสียงขลุ่ยไทย พบว่าสเปกตรัมในส่วนของสัญญาณไซน์ประกอบไปด้วยความถี่ที่เป็นฮาร์โมนิกกันซึ่งมีระยะห่างกันพอสมควร ดังนั้นในการทดลองหาค่าความคลาดเคลื่อนเฉลี่ยจากการใช้วินโดว์แต่ละชนิดจึงต้องทำการหาค่าความคลาดเคลื่อนเฉลี่ยที่ค่า Δ ต่าง ๆ เนื่องจากว่าค่าระยะห่างระหว่างความถี่ที่แท้จริงของสัญญาณเสียงขลุ่ยไทยเป็นเท่าใดนั้นยังไม่แน่ชัด โดยวินโดว์ที่เหมาะสมสำหรับการวิเคราะห์เชิงสเปกตรัมของสัญญาณเสียงขลุ่ยไทยนั้นต้องให้ค่าความคลาดเคลื่อนเฉลี่ยน้อยที่สุด ทั้งนี้ในการทดลองได้ทำการเปรียบเทียบวินโดว์ประเภทต่าง ๆ ดังนี้ เรคแทงคิวลาร์วินโดว์ (rectangular window) แฮมมิงวินโดว์ (hamming window) แฮนนิงวินโดว์ (hanning window) และแบล็คแมนวินโดว์ (blackman window) ซึ่งลักษณะวินโดว์แต่ละประเภทแสดงในรูปที่ 3.1

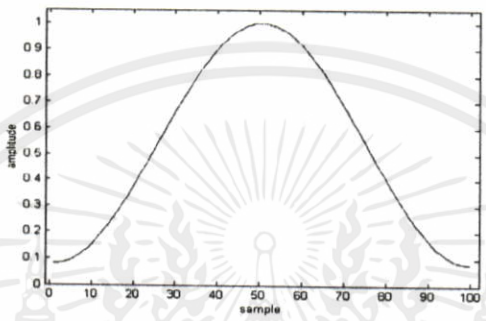
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.1 การทดลอง

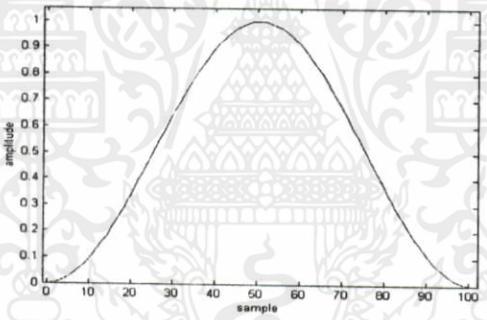
ในการทดลองนี้ได้สร้างลักษณะสัญญาณเสียงที่มีองค์ประกอบเชิงความถี่ใกล้เคียงเสียงตัวโน้ตโดของขลุ่ยเพียงออ เพื่อทำการเปรียบเทียบค่าความคลาดเคลื่อนเฉลี่ยจากการใช้วินโดว์แต่ละชนิด (เหตุที่ใช้โน้ตโดเพราะว่าในการวิเคราะห์ความถี่ยิ่งใกล้กันเท่าไรมีโอกาสที่จะวิเคราะห์ผิด



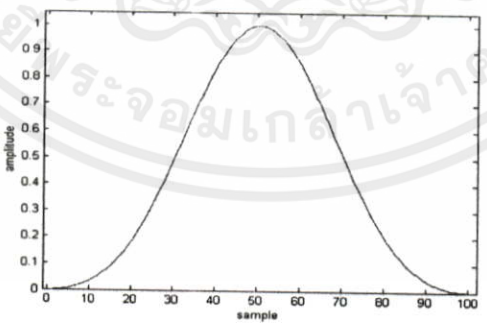
(ก)



(ข)



(ค)



(ง)

รูปที่ 3.1 แสดงลักษณะวินโดว์ฟังก์ชันประเภทต่าง ๆ

ก. เรคแทงคิวลาร์วินโดว์

ข. แฮมมิงวินโดว์

ค. แฮนนิงวินโดว์

ง. แบล็คแมนวินโดว์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พลาดได้ และโน้ตโดมีระยะห่างระหว่างความถี่สองความถี่ที่ใกล้กันที่สุดเมื่อเทียบกับโน้ตอื่น ๆ ในเสียงขลุ่ย ซึ่งถือว่าเป็นเวิร์สเคส (worst case) และหากสามารถประมาณค่าได้จากกรณีดังกล่าวทำให้สามารถนำไปใช้กับโน้ตตัวอื่น ๆ ได้) โดยการวัดค่าความคลาดเคลื่อนเฉลี่ยจะเฉลี่ยเฉพาะช่วงความถี่หนึ่ง มิได้ทำการทดลองในทุกช่วงความถี่ เนื่องจากว่าค่าความคลาดเคลื่อนเฉลี่ยที่ได้จากการเปลี่ยนแปลงค่าความถี่ที่มีค่าอยู่ในช่วง BIN[1] ใด ๆ สามารถนำมาเป็นตัวแทนของค่าความคลาดเคลื่อนเฉลี่ยที่พิจารณาจากทุก ๆ ความถี่ได้ (รายละเอียดแสดงในภาคผนวก ก) ซึ่งในการทดลองจะทำการหาค่าความคลาดเคลื่อนเฉลี่ยจากการใช้ค่า Δ ที่ต่างกัน โดยสร้างสัญญาณที่ประกอบไปด้วยสัญญาณไซน์ที่ 2 ค่าความถี่ ณ ค่า Δ ต่าง ๆ และทำการเปลี่ยนแปลงทั้ง 2 ค่าความถี่โดยการเพิ่มค่าความถี่ในปริมาณที่เท่าๆ กันให้อยู่ในช่วง 1 BIN เพื่อนำมาเฉลี่ยค่าความคลาดเคลื่อน ซึ่งมีรายละเอียดดังนี้

1) สร้างสัญญาณขนาด 10000 แซมเปิล ซึ่งเกิดจากการรวมกันของสัญญาณไซน์ 2 ค่าความถี่ ซึ่งแต่ละสัญญาณเป็นไปตามสมการที่ 3.1

$$s(n)=\sum_{r=1}^{2} A_r \cos\left(2\pi f_r n\right)+e(n) \qquad (3.1)$$

โดยที่ $s(n)$ คือสัญญาณเสียง

$A_1=1,\ A_2=0.0831$

$f_1=0.0102,\ f_2=f_1+\Delta$ (กำหนด Δ เท่ากับ 0.0098 0.01 0.0102 0.0104 0.0106 ซึ่งเป็นการกำหนดระยะห่างของค่าความถี่ที่ค่าต่าง ๆ)

$e(n)$ คือสัญญาณรบกวน (ในที่นี้ใช้ฟังก์ชัน randn ในโปรแกรม MATLAB® สังเคราะห์ขึ้นมา) โดยที่ $n=0,\ldots,9999$

2) ในการหาค่าเฉลี่ยที่ค่า Δ เท่ากับ 0.0098 จะทำการสร้างสัญญาณตามสมการ 3.1 โดยกำหนดค่าความถี่ f_r (โดยที่ r เท่ากับ 1 และ 2) ให้เปลี่ยนแปลงในช่วง 1 BIN (ทั้งนี้ความถี่ทั้งสอง f_1 และ f_2 จะทำการเพิ่มหรือลดด้วยจำนวนค่าที่เท่ากันเสมอ แต่ค่าแอมปลิจูดจะคงค่าเดิมเสมอ) คือค่าความถี่จะมีการเปลี่ยนแปลงโดยอยู่ในช่วง $f_r-(1/20000)$ และ $f_r+(1/20000)$ ซึ่งจะเพิ่มค่าความถี่ทีละ 1/100000 ดังนั้นสัญญาณที่ได้ทั้งหมดจะมีค่าเท่ากับ 11 สัญญาณ

3) นำสัญญาณแต่ละสัญญาณมาคูณกับวินโดว์ที่มีขนาดเท่ากัน ทั้งนี้วินโดว์ประเภทต่าง ๆ ได้ใช้ฟังก์ชันในโปรแกรม Matlab® สร้างขึ้นมาให้มีขนาดข้อมูลเท่ากับสัญญาณไซน์ที่สร้างขึ้นในที่นี้คือ 10000 แซมเปิล

4)ใช้ฟาสฟูเรียร์ทรานฟอร์มหรือ FFT อัลกอลิทึมในการวิเคราะห์สเปกตรัมของสัญญาณ เพื่อคำนวณหาค่าความถี่และแอมปลิจูด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[1] BIN คือ ระยะห่างระหว่างจุดข้อมูลที่ทำการประมาณค่าในการหาดีสครีตฟูเรียร์ทรานฟอร์ม หรือ 1/จำนวนข้อมูลทั้งหมดที่ใช้ในการหาดีสครีตฟูเรียร์ทรานฟอร์ม

5)หาค่าความผิดพลาดเฉลี่ยของความถี่และแอมปลิจูดที่ได้โดยทำการหาค่ามีนสแคว์เออร์เรอร์ (MSE) ซึ่งสมการในการคำนวณนั้นดูรายละเอียดได้ที่ภาคผนวก ข

6)ทำการคำนวณค่าความคลาดเคลื่อนเฉลี่ยที่ค่า Δ ต่าง ๆ ดังนี้ 0.01 0.0102 0.0104 0.0106 โดยทำตามขั้นตอนที่ 1-5 เพียงแต่เปลี่ยนเฉพาะค่า Δ ที่ค่าต่าง ๆ

ในการหาค่าคลาดเคลื่อนเฉลี่ยจากการใช้วินโดว์แต่ละชนิด (เรคแทงคิวลาร์วินโดว์ แฮมมิงวินโดว์ แฮนนิงวินโดว์ และแบล็คแมนวินโดว์) จะต้องทำการสร้างสัญญาณเสียงตามขั้นตอนที่ได้กล่าวไว้ในข้อ 1-6 โดยเปลี่ยนวินโดว์ฟังก์ชันที่นำมาคูณในข้อ 3

3.1.2 ผลการทดลอง

จากการทดลองพบว่าโดยเฉลี่ยแล้ว การใช้วินโดว์ชนิดต่าง ๆ จะให้ค่าความคลาดเคลื่อนของความถี่เท่ากัน และค่าความคลาดเคลื่อนของแอมปลิจูดใกล้เคียงกัน (ค่าความคลาดเคลื่อนดังกล่าวได้จากการหาค่าสัมบูรณ์ของผลต่างระหว่างค่าจริงที่ใช้ในการสังเคราะห์และค่าที่ได้จากการวิเคราะห์สัญญาณที่สังเคราะห์ขึ้น) ซึ่งแสดงดังรูปที่ 3.2 และ 3.3 ตามลำดับ

การหาค่าเฉลี่ยของค่าความคลาดเคลื่อนจากการใช้วินโดว์แต่ละชนิดจะกระทำโดยการเฉลี่ยค่าความคลาดเคลื่อนในแต่ละค่าความถี่ที่เปลี่ยนแปลงในช่วงความถี่หนึ่ง ๆ ซึ่งพบว่าในการใช้วินโดว์ต่างชนิดกันให้ค่าความคลาดเคลื่อนเฉลี่ยของค่าความถี่เท่ากัน แม้ในค่า Δ ที่ต่างกัน และค่าความคลาดเคลื่อนของค่าแอมปลิจูดเฉลี่ยมีค่าใกล้เคียงกัน และในแต่ละวินโดว์จะให้ค่าความคลาดเคลื่อนเท่ากันในทุกค่า Δ แสดงดังตารางที่ 3.1 และตารางที่ 3.2

ตารางที่ 3.1 แสดงค่าความคลาดเคลื่อนเฉลี่ยของค่าความถี่และค่าแอมปลิจูดของสัญญาณไซน์ลำดับที่หนึ่ง จากการใช้วินโดว์ชนิดต่าง ๆ

ค่าความคลาดเคลื่อนเฉลี่ย		ระยะห่างระหว่างสองค่าความถี่ (Δ)				
		0.0098	0.01	0.0102	0.0104	0.0106
แรคแทงคิวลาร์วินโดว์	ค่าความถี่	9.22E-06	9.22E-06	9.22E-06	9.22E-06	9.22E-06
	ค่าแอมปลิจูด	0.0547	0.0547	0.0547	0.0547	0.0547
แฮมมิงวินโดว์	ค่าความถี่	9.22E-06	9.22E-06	9.22E-06	9.22E-06	9.22E-06
	ค่าแอมปลิจูด	0.0272	0.0272	0.0272	0.0272	0.0272
แฮนนิงวินโดว์	ค่าความถี่	9.22E-06	9.22E-06	9.22E-06	9.22E-06	9.22E-06
	ค่าแอมปลิจูด	0.0225	0.0225	0.0225	0.0225	0.0225
แบล็คแมนวินโดว์	ค่าความถี่	9.22E-06	9.22E-06	9.22E-06	9.22E-06	9.22E-06
	ค่าแอมปลิจูด	0.0176	0.0176	0.0176	0.0176	0.0176

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.2 แสดงค่าความคลาดเคลื่อนเฉลี่ยของค่าความถี่และค่าแอมปลิจูดของสัญญาณไซน์ลำดับที่สอง จากการใช้วินโดว์ชนิดต่าง ๆ

ค่าความคลาดเคลื่อนเฉลี่ย		ระยะห่างระหว่างสองค่าความถี่ (Δ)				
		0.0098	0.01	0.0102	0.0104	0.0106
แรคแทงคิวลาร์วินโดว์	ค่าความถี่	9.22E-06	9.22E-06	9.22E-06	9.22E-06	9.22E-06
	ค่าแอมปลิจูด	0.0041	0.0041	0.0041	0.0041	0.0041
แฮมมิงวินโดว์	ค่าความถี่	9.22E-06	9.22E-06	9.22E-06	9.22E-06	9.22E-06
	ค่าแอมปลิจูด	0.0021	0.0021	0.0021	0.0021	0.0021
แฮนนิงวินโดว์	ค่าความถี่	9.22E-06	9.22E-06	9.22E-06	9.22E-06	9.22E-06
	ค่าแอมปลิจูด	0.0018	0.0018	0.0018	0.0018	0.0018
แบล็คแมนวินโดว์	ค่าความถี่	9.22E-06	9.22E-06	9.22E-06	9.22E-06	9.22E-06
	ค่าแอมปลิจูด	0.0014	0.0014	0.0014	0.0014	0.0014

3.1.3 สรุปผลการทดลอง

เมื่อพิจารณาจากการใช้วินโดว์ทั้ง 4 ชนิด พบว่าระยะห่างระหว่างสองค่าความถี่ไม่มีผลต่อค่าความถูกต้องของค่าความถี่และค่าแอมปลิจูด ซึ่งวินโดว์ทั้ง 4 ชนิดนี้จะให้ค่าความคลาดเคลื่อนเฉลี่ยของความถี่เท่ากัน แต่สำหรับค่าความคลาดเคลื่อนเฉลี่ยของแอมปลิจูด แบล็คแมนวินโดว์จะให้ค่าความคลาดเคลื่อนเฉลี่ยน้อยที่สุด ดังนั้นในการวิเคราะห์สัญญาณเสียงขลุ่ยไทยเพื่อการหาค่าพารามิเตอร์ต่าง ๆ ที่จะกล่าวถึงต่อไปจะใช้แบล็คแมนวินโดว์ในการวิเคราะห์

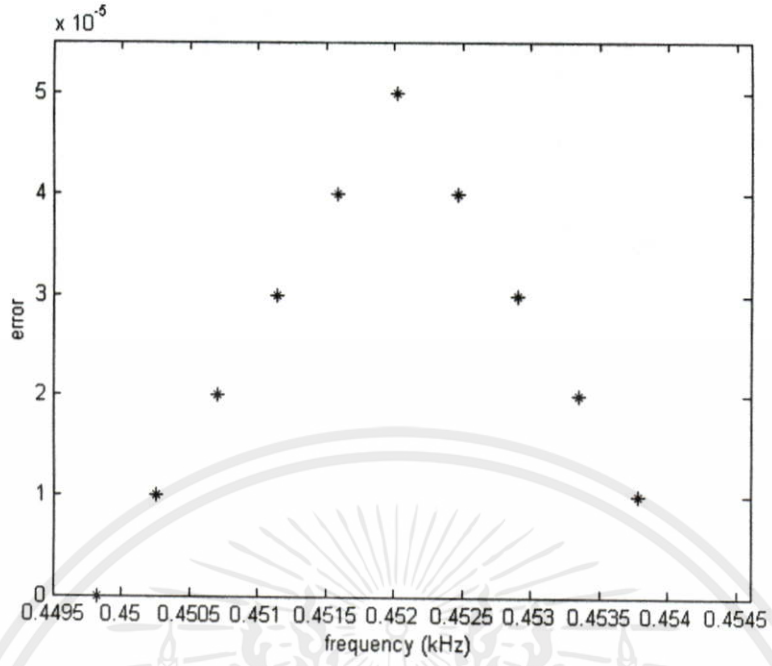
นอกจากนี้จากการศึกษาลักษณะสเปกตรัม พบว่าแฮมมิงวินโดว์ แฮนนิงวินโดว์ และแบล็คแมนวินโดว์จะให้สเปกตรัมที่แยกส่วนของสัญญาณไซน์ชัดเจนมากกว่าแรคแทงคิวลาร์วินโดว์ ซึ่งจะเป็นประโยชน์ในการวิเคราะห์หาส่วนของสัญญาณรบกวนได้ดี (พิจารณารูปที่ 3.4)

จากผลการทดลองในตารางที่ 3.1 และ 3.2 พบว่าแบล็คแมนวินโดว์มีความเหมาะสมในการนำมาใช้เพื่อการวิเคราะห์สัญญาณเสียงขลุ่ยไทยมากกว่าวินโดว์ประเภทอื่น เนื่องจากการใช้วินโดว์ชนิดนี้โดยเฉลี่ยจะให้ค่าพารามิเตอร์ที่ถูกต้องสำหรับสัญญาณไซน์ที่มีช่วงห่างของสัญญาณดังที่พบในเสียงขลุ่ยไทย

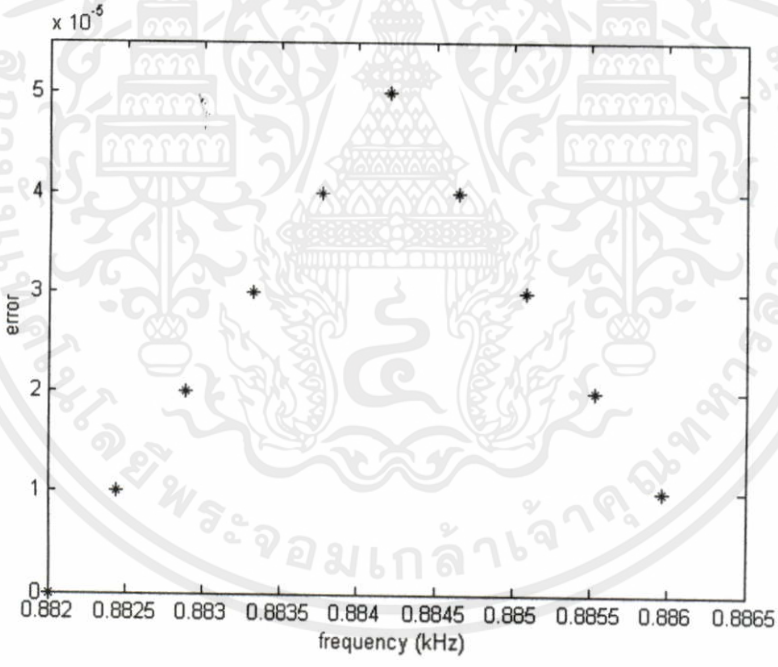
3.2 การเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์ให้กับสัญญาณเสียง

การวิเคราะห์หาองค์ประกอบเชิงความถี่ของสัญญาณเสียงเพื่อให้ได้ค่าที่ถูกต้องมากขึ้นนั้น จำเป็นต้องเพิ่มจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์เข้าไปกับสัญญาณ แต่จากการทดลองพบว่าการเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์อาจจะทำให้มีความคลาดเคลื่อนมากขึ้นได้ ทั้งนี้ขึ้นอยู่กับจำนวนจุดของข้อมูลที่มีค่าเป็นศูนย์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการใช้



(ก)



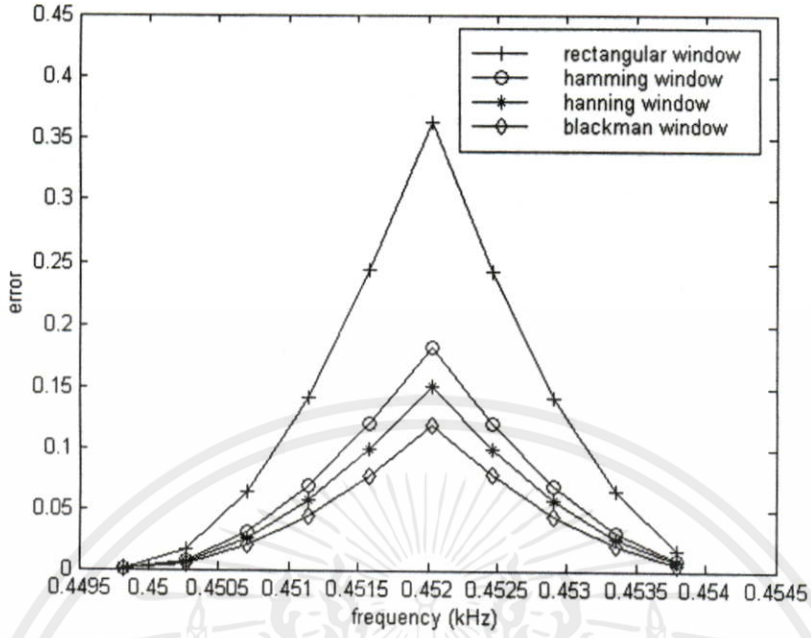
(ข)

รูปที่ 3.2 แสดงค่าความคลาดเคลื่อนของค่าความถี่ที่ได้จากการใช้วินโดว์ในทุกประเภท

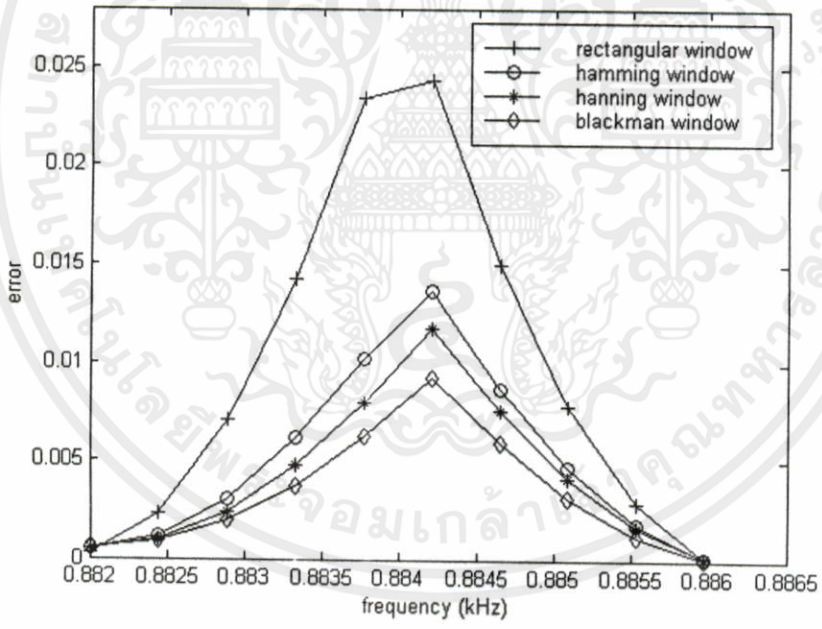
ก. ค่าความคลาดเคลื่อนของค่าความถี่ที่ฟังก์ชันไซน์ลำดับที่หนึ่ง ณ $\Delta = 0.0098$

ข. ค่าความคลาดเคลื่อนของค่าความถี่ที่ฟังก์ชันไซน์ลำดับที่สอง ณ $\Delta = 0.0098$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)



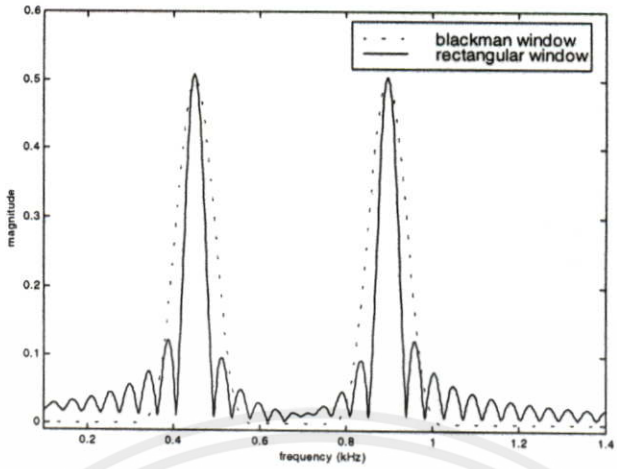
(ข)

รูปที่ 3.3 แสดงค่าความคลาดเคลื่อนของค่าแอมปลิจูดจากการใช้วินโดว์แต่ละประเภท

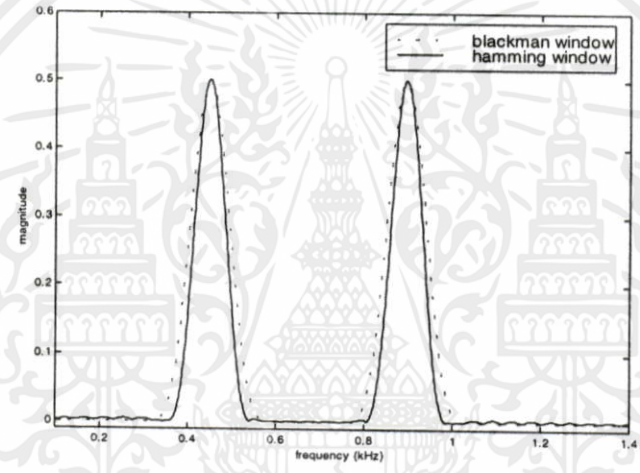
ก. ค่าความคลาดเคลื่อนของค่าแอมปลิจูดที่ฟังก์ชันไซน์ลำดับที่หนึ่ง ณ $\Delta = 0.0098$

ข. ค่าความคลาดเคลื่อนของค่าแอมปลิจูดที่ฟังก์ชันไซน์ลำดับที่สอง ณ $\Delta = 0.0098$

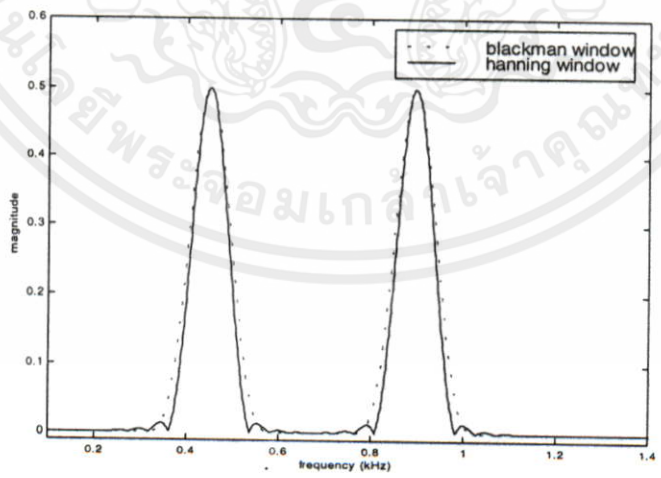
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)



(ข)



(ค)

รูปที่ 3.4 แสดงการเปรียบเทียบสเปกตรัมที่ได้จากการใช้วินโดว์ประเภทต่าง ๆ

ก. เปรียบเทียบสเปกตรัมที่ได้จากการใช้แบล็คแมนวินโดว์และเรคแทงคิวลาร์วินโดว์

ข. เปรียบเทียบสเปกตรัมที่ได้จากการใช้แบล็คแมนวินโดว์และแฮมมิงวินโดว์

ค. เปรียบเทียบสเปกตรัมที่ได้จากการใช้แบล็คแมนวินโดว์และแฮนนิงวินโดว์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.1 การทดลอง

สร้างสัญญาณเสียงซึ่งประกอบด้วยสัญญาณไซน์ที่ความถี่ต่าง ๆ โดยกำหนดให้มีค่าใกล้เคียงกับสัญญาณเสียงโดของขลุ่ยเพียงออ จากนั้นทำการเพิ่มจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์เข้าไปในสัญญาณ ซึ่งมีรายละเอียดดังนี้

1) สร้างสัญญาณตามสมการ 3.1 โดยที่ $A_1 = 1$, $A_2 = 0.0831$ และ $f_1 = 0.0102$, $f_2 = 0.0203$ และ $e(n)$ ใช้ฟังก์ชัน randn ใน MATLAB® สังเคราะห์ขึ้นมา โดยที่ $n = 0,\ldots,999$

2) นำสัญญาณที่ได้คูณกับแบล็คแมนวินโดว์ในโดเมนของเวลาด้วยขนาดจำนวนข้อมูล 1000 แซมเปิลเท่ากัน

3) สัญญาณที่ทำการวินโดว์แล้ว นำมาเพิ่มจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์ตั้งแต่ 1,000 ถึง 66,000 ค่า

4) ใช้ฟาสฟูเรียร์ทรานฟอร์มอัลกอลิทึมในการวิเคราะห์สเปกตรัมของสัญญาณ เพื่อคำนวณหาค่าความถี่และแอมปลิจูด

5) หาค่าความผิดพลาดของความถี่และแอมปลิจูดที่ได้โดยทำการหาค่าสัมบูรณ์ของผลต่างระหว่างค่าที่ใช้ในการสังเคราะห์สัญญาณขึ้นมากับค่าที่ได้จากการวิเคราะห์สัญญาณเสียงสังเคราะห์ดังกล่าว

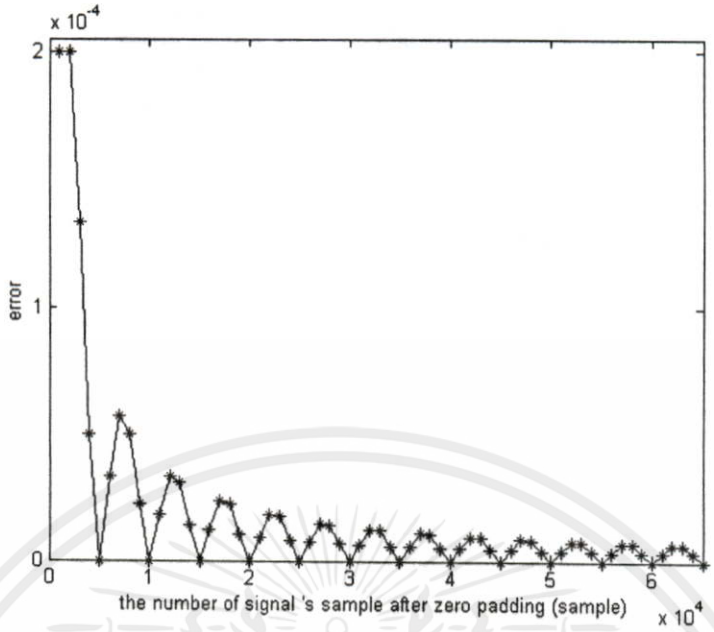
3.2.2 ผลการทดลอง

พิจารณารูปที่ 3.5 และ 3.6 จะเห็นว่า การเพิ่มจำนวนข้อมูลที่มีค่าเป็นศูนย์เข้าไปในสัญญาณเสียงนั้นมีผลต่อความถูกต้องของค่าความถี่และค่าแอมปลิจูดของสัญญาณเสียง นั้นคือ เมื่อเพิ่มจำนวนข้อมูลที่มีค่าเป็นศูนย์มากขึ้นก็จะทำให้ค่าความถี่และค่าแอมปลิจูดมีแนวโน้มที่จะมีความถูกต้องมากขึ้นเช่นกัน

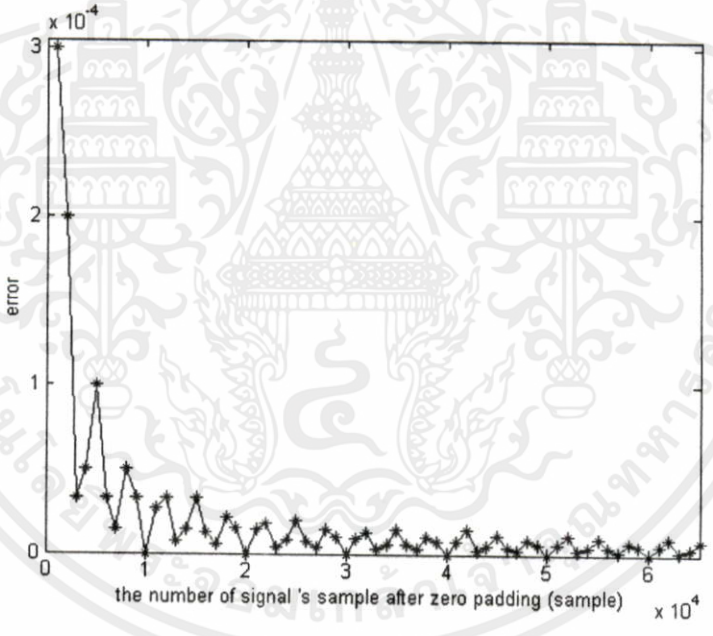
3.2.3 วิเคราะห์ผลการทดลอง

วิธีการคำนวณหาสเปกตรัมของสัญญาณดิจิตอล หรือที่เรียกว่า ดีสครีตฟูเรียร์ทรานฟอร์ม (Discrete Fourier Transform) เป็นการคำนวณหาองค์ประกอบของค่าความถี่ของสัญญาณดีสครีตในโดเมนความถี่ สัญญาณเสียงที่มีขนาด N แซมเปิล จะมีองค์ประกอบของค่าความถี่เชิงมุมตั้งแต่ 0 ถึง 2π เรเดียน โดยแต่ละค่าความถี่จะมีระยะห่างกัน $2\pi / N$ เรเดียน ซึ่งมีความเป็นไปได้ที่ค่าความถี่ในแต่ละแซมเปิลจะคลาดเคลื่อนไป π / N เรเดียน เนื่องจากค่าความถี่จริงอาจอยู่ในช่วงระยะห่างของค่าความถี่สองค่าใด ๆ จำนวนข้อมูล N จึงมีผลต่อความถูกต้องของค่าองค์ประกอบของสัญญาณไซน์ หากจำนวนข้อมูลเพิ่มเป็น $2N, 4N, \ldots, 2^n N$ จะทำให้ค่าความคลาดเคลื่อนลดลงเป็น $\pi / 2N, \pi / 4N, \ldots, \pi / 2^n N$ แต่ในความเป็นจริงข้อมูลแซมเปิลของสัญญาณไม่สามารถเพิ่มขึ้นได้ ดังนั้นการลดความคลาดเคลื่อนจึงกระทำโดยการเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์ให้กับสัญญาณแทนอย่างไรก็ดี การเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์อาจไม่ได้เพิ่มความถูกต้องของการวัดค่าสเปกตรัม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)



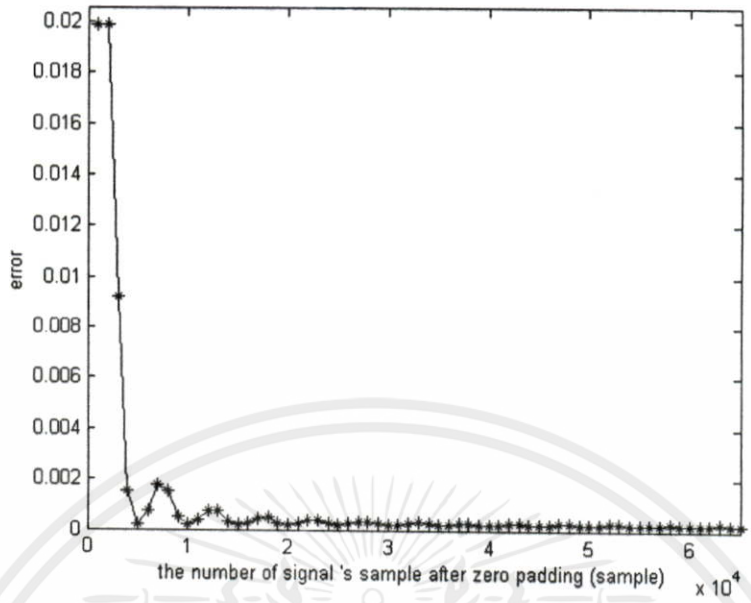
(ข)

รูปที่ 3.5 แสดงค่าความผิดพลาดของค่าความถี่ เมื่อเพิ่มจำนวนข้อมูลที่มีค่าศูนย์ตั้งแต่ 1000 จนถึง 66000 ค่าโดยเพิ่มทีละ 10000 ค่า

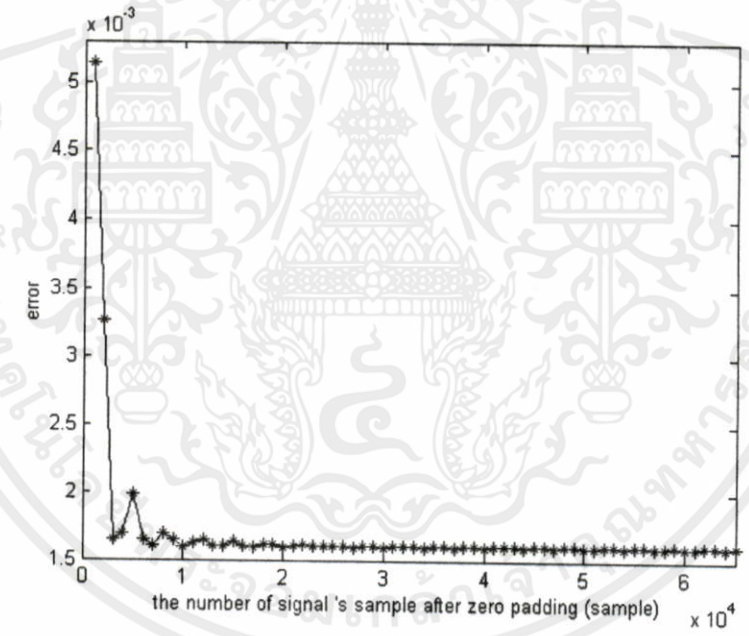
ก. ค่าความผิดพลาดของค่าความถี่ที่ฟังก์ชันไซน์ลำดับที่หนึ่ง

ข. ค่าความผิดพลาดของค่าความถี่ที่ฟังก์ชันไซน์ลำดับที่สอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)



(ข)

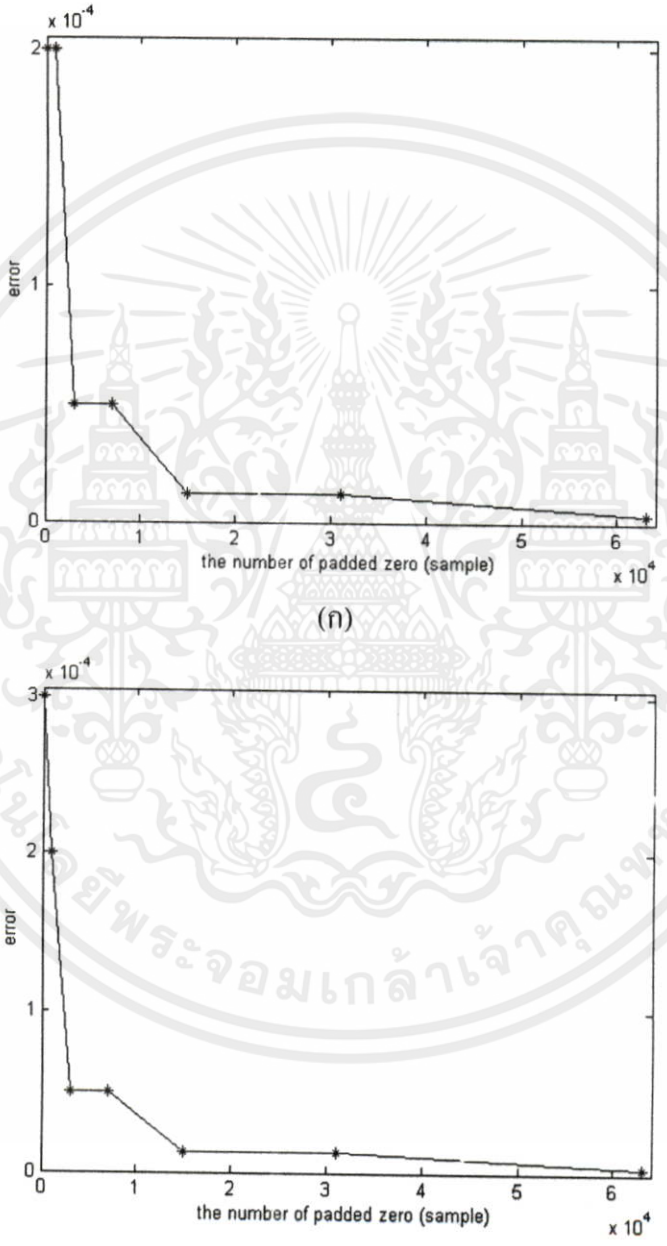
รูปที่ 3.6 แสดงค่าความผิดพลาดของค่าแอมปลิจูด เมื่อเพิ่มจำนวนข้อมูลที่มีค่าศูนย์ตั้งแต่ 1000 จนถึง 66000 ค่าโดยเพิ่มทีละ 10000 ค่า

ก. ค่าความผิดพลาดของค่าแอมปลิจูดที่ฟังก์ชันไซน์ลำดับที่หนึ่ง

ข. ค่าความผิดพลาดของค่าแอมปลิจูดที่ฟังก์ชันไซน์ลำดับที่สอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เสมอไป (ดูรูป 3.5 และ 3.6) จากการวิเคราะห์ข้างต้นพบว่า ค่าความคลาดเคลื่อนจะลดลง เมื่อขนาดของสัญญาณที่เพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์มีค่าเป็นจำนวนเท่าของขนาดสัญญาณเดิม ดังแสดงในรูปที่ 3.7 และ 3.8 ดังนั้นเพื่อให้การเพิ่มจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์มีผลทำให้ค่าความคลาดเคลื่อนลดลง (หรือไม่เพิ่มขึ้น) เสมอ ในที่นี้จึงได้เสนอแนวทางการเพิ่มจุดข้อมูลที่เป็นศูนย์ตามสมการที่ 3.1



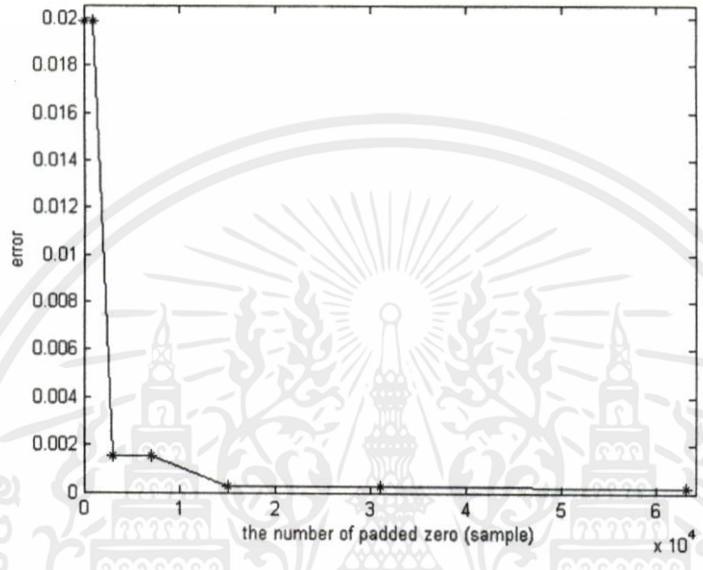
รูปที่ 3.7 แสดงค่าความคลาดเคลื่อนที่ได้จากการเพิ่มจำนวนศูนย์เป็นจำนวนเท่าของสัญญาณ
ก. ค่าความคลาดเคลื่อนของค่าความถี่ที่ฟังก์ชันไซน์ลำดับที่หนึ่ง
ข. ค่าความคลาดเคลื่อนของค่าความถี่ที่ฟังก์ชันไซน์ลำดับที่สอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเอกสารทุกครั้งที่มีการนำไปใช้

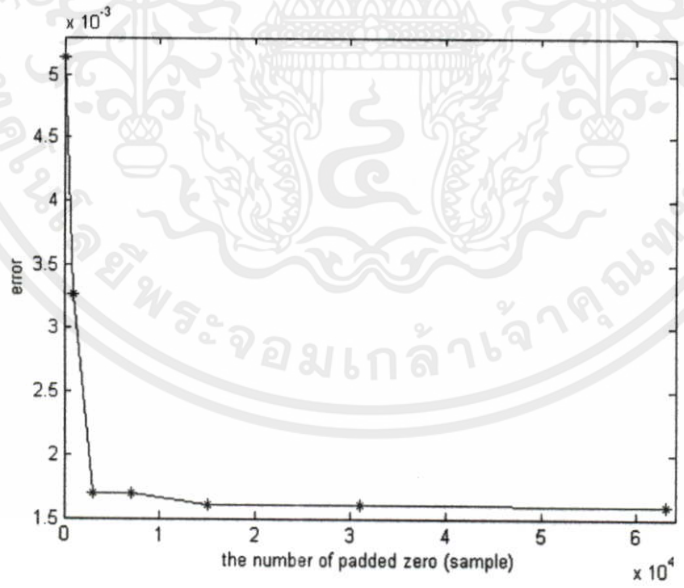
$$T = N(2^k - 1) \tag{3.1}$$

โดยที่ T คือจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์ที่เพิ่มเข้ากับสัญญาณ

N คือจำนวนแซมเปิลของสัญญาณที่พิจารณา $k = 0,1,2,3,...;$



(ก)



(ข)

รูปที่ 3.8 แสดงค่าความคลาดเคลื่อนที่ได้จากการเพิ่มจำนวนศูนย์เป็นจำนวนเท่าของสัญญาณ
ก. ค่าความคลาดเคลื่อนของค่าแอมปลิจูดที่ฟังก์ชันไซน์ลำดับที่หนึ่ง
ค. ค่าความคลาดเคลื่อนของค่าแอมปลิจูดที่ฟังก์ชันไซน์ลำดับที่สอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
การวิเคราะห์และการสังเคราะห์เสียงขลุ่ยไทยด้วยแบบจำลอง Sinusoidal Plus Noise

ขั้นตอนในการเก็บรวบรวมข้อมูลเสียงดนตรีขลุ่ยไทยเพื่อทำการวิเคราะห์นั้น ได้ทำการบันทึกเสียงขลุ่ยเพียงออตั้งแต่เสียงโดต่ำจนถึงเสียงซอลสูงรวมทั้งหมด 12 ตัวโน้ต นอกจากนี้ยังทำการจัดเก็บเสียงโน้ตของขลุ่ยไทยชนิดอื่น อย่างเช่น ขลุ่ยหลิบ ขลุ่ยอู้ และขลุ่ยเพียงออเทียบสากล (คุณสมบัติขลุ่ยไทยชนิดต่าง ๆ พิจารณาภาคผนวก ค) ซึ่งเสียงโน้ตทั้งหมดนี้จัดเก็บในรูปของเวปไฟล์ที่อัตราแซมปลิง 44100 เฮิรตซ์ และใช้การจัดเก็บแบบ 16 บิตต่อหนึ่งแซมเปิล โดยโน้ตแต่ละตัวจะทำการบันทึก 5 ครั้ง มีความยาวประมาณ 1.5 วินาที (ลักษณะสัญญาณเสียงโน้ตต่าง ๆ พิจารณาภาคผนวก ค)

4.1 การวิเคราะห์ในส่วนของสัญญาณไซน์

เนื่องจากว่าลักษณะของสัญญาณเสียงขลุ่ยเพียงออที่นำมาวิเคราะห์ประกอบไปด้วย 2 ส่วน คือส่วนสัญญาณเสียงที่มีการเปลี่ยนแปลง (attack และ decay) และส่วนของสัญญาณเสียงที่อยู่ในภาวะเสถียร (sustain) ดังนั้นในการวิเคราะห์จึงต้องแยกการวิเคราะห์สัญญาณเสียงเป็น 2 ส่วน ซึ่งมีรายละเอียดดังนี้

4.1.1 การวิเคราะห์สัญญาณเสียงในช่วงภาวะเสถียร

ในการวิเคราะห์สัญญาณเสียงที่อยู่ในภาวะเสถียร ได้ตั้งสมมติฐานว่าในช่วงเวลาที่เปลี่ยนไปสัญญาณเสียงน่าจะมีองค์ประกอบเชิงความถี่ที่เหมือนกัน ดังนั้นเพื่อเป็นการพิสูจน์ข้อสมมติฐานดังกล่าว จึงได้ทำการทดลองแบ่งสัญญาณออกเป็นวินโดว์ย่อย ๆ เพื่อที่จะหาค่าองค์ประกอบเชิงความถี่ของแต่ละวินโดว์ย่อยและนำมาเปรียบเทียบกัน ทั้งนี้ขนาดของวินโดว์ของสัญญาณจะมีค่าเป็น 250 500 1000 2000 และ 4000 แซมเปิล ในการทดลองนี้เลือกสัญญาณเสียงของโน้ตลา โดยจะนำสัญญาณที่แบ่งเป็นวินโดว์ย่อย ๆ มาคูณกับแบล็คแมนวินโดว์ จากนั้นทำการเพิ่มจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์เข้ากับสัญญาณในแต่ละวินโดว์ให้มีขนาด 640000 แซมเปิล แล้วนำมาคำนวณค่าองค์ประกอบเชิงความถี่ในแต่ละวินโดว์ ซึ่งการหาเปอร์เซ็นต์ความคลาดเคลื่อนเฉลี่ยของสองวินโดว์ที่ติดกันใด ๆ ได้จากการ

เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของผลงานทุกครั้งที่มีการใช้

1) พิจารณาหาเปอร์เซ็นต์ความผิดพลาดของค่าความถี่และค่าแอมปลิจูดโดยหาค่าสัมบูรณ์ของผลต่างระหว่างค่าที่ได้วินโดว์ที่ติดกันใด ๆ แล้วหารด้วยค่าที่ได้จากวินโดว์แรกในสองวินโดว์ที่นำมาหาผลต่าง

2) เฉลี่ยเปอร์เซ็นต์ความผิดพลาดทั้งหมดด้วยจำนวนวินโดว์ย่อยลบหนึ่ง

จากการทดลองพบว่าในแต่ละวินโดว์จะมีองค์ประกอบเชิงความถี่ใกล้เคียงกัน และเปอร์เซ็นต์ความคลาดเคลื่อนของค่าองค์ประกอบเชิงความถี่ระหว่าง 2 วินโดว์ใด ๆ ที่ติดกันเมื่อเทียบกับค่าองค์ประกอบเชิงความถี่ของวินโดว์แรกอยู่ในระดับที่ยอมรับได้ (ซึ่งแสดงรายละเอียดดังตารางที่ 4.1 และ 4.2) แม้ว่าค่าเปอร์เซ็นต์ความคลาดเคลื่อนเฉลี่ยของแอมปลิจูด ณ ความถี่ฮาร์โมนิกต่าง ๆ (ที่มิใช่ความถี่พื้นฐาน) จะมีค่าค่อนข้างสูง แต่เนื่องจากค่าแอมปลิจูดที่คำนวณได้นี้มีค่าน้อยมาก จึงทำให้ค่าความคลาดเคลื่อนที่เกิดขึ้นจึงไม่ได้มีผลต่อสัญญาณเสียงที่เปลี่ยนแปลงไปตามเวลาแต่อย่างใด ดังนั้นสามารถสรุปข้อสันนิษฐานได้ว่าในช่วงเวลาที่เปลี่ยนไปสัญญาณเสียงมีองค์ประกอบเชิงความถี่ที่เหมือนกัน

นอกจากนี้ขนาดของวินโดว์ที่มีความยาวมาก เมื่อนำมาวิเคราะห์เชิงสเปกตรัมจะให้สเปกตรัมของสัญญาณได้ชัดเจนมากที่สุด ดังนั้นในการวิเคราะห์หาองค์ประกอบเชิงความถี่ของเสียงขลุ่ยไทยของโน้ตต่าง ๆ จึงเลือกใช้ขนาดวินโดว์ที่มีขนาดสูงสุดจากข้อมูลที่บันทึก (ทั้งนี้ขนาดวินโดว์ของแต่ละตัวโน้ตขึ้นอยู่กับการบันทึกแต่ละครั้ง) เพื่อนำมาวิเคราะห์หาค่าพารามิเตอร์ โดยขั้นตอนการวิเคราะห์เชิงสเปกตรัมในทุก ๆ ตัวโน้ตมีดังนี้

ตารางที่ 4.1 แสดงเปอร์เซ็นต์ความคลาดเคลื่อนเฉลี่ยของค่าความถี่

เปอร์เซ็นต์ความคลาดเคลื่อนเฉลี่ย	ขนาดวินโดว์ (แซมเปิล)				
	250	500	1000	2000	4000
สัญญาณไซน์ลำดับที่ 1	0.1194	0.1012	0.1109	0.1534	0.1357
สัญญาณไซน์ลำดับที่ 2	0.9599	0.6322	0.1965	0.1361	0.1253
สัญญาณไซน์ลำดับที่ 3	0.755	0.3586	0.1671	0.1221	0.1114
สัญญาณไซน์ลำดับที่ 4	1.1957	0.4335	0.1675	0.1324	0.1149
สัญญาณไซน์ลำดับที่ 5	0.4831	0.2263	0.1726	0.1267	0.1086

ตารางที่ 4.2 แสดงเปอร์เซ็นต์ความคลาดเคลื่อนเฉลี่ยของค่าแอมปลิจูด

เปอร์เซ็นต์ความคลาดเคลื่อนเฉลี่ย	ขนาดวินโดว์ (แซมเปิล)				
	250	500	1000	2000	4000
สัญญาณไซน์ลำดับที่ 1	1.0051	1.4151	1.4065	1.9603	2.2799
สัญญาณไซน์ลำดับที่ 2	12.8877	13.3478	8.8689	6.2237	8.8583
สัญญาณไซน์ลำดับที่ 3	15.8475	8.1919	6.1625	4.8889	3.9149
สัญญาณไซน์ลำดับที่ 4	22.7577	23.2111	18.4023	21.2484	11.0695
สัญญาณไซน์ลำดับที่ 5	12.8688	14.8631	15.4838	10.0166	3.4267

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) นำสัญญาณเสียงขลุ่ยที่มีขนาดสูงสุดซึ่งอยู่ในช่วงภาวะเสถียร (N แซมเปิล) มาคูณกับแบล็คแมนวินโดว์

2) เพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์ให้กับสัญญาณที่ทำการคูณกับแบล็คแมนวินโดว์แล้ว เป็นจำนวน $N(2^k-1)$ แซมเปิล ในที่นี้เลือกค่า $k=5$ เนื่องจากว่าค่า k ที่มากกว่านี้ไม่ได้ทำให้ค่าองค์ประกอบเชิงความถี่ที่คำนวณได้มีความถูกต้องมากขึ้น อันเนื่องมาจากว่าขนาดของสัญญาณเองมีความยาวไม่ต่ำกว่า 50000 แซมเปิลซึ่งถือว่าเป็นขนาดที่ยาวพอที่จะทำการวิเคราะห์เชิงสเปกตรัมได้ค่าที่ใกล้เคียงความถูกต้องอยู่แล้ว

3) คำนวณค่าสเปกตรัมของสัญญาณโดยใช้อัลกอลิทึมฟาสฟูเรียร์ทรานฟอร์ม

4) จากสเปกตรัมที่ได้ ใช้อัลกอลิธึมตรวจหาจุดที่มีค่าความชันสูงเพื่อคำนวณค่าความถี่ และค่าแอมปลิจูด

ในการวิเคราะห์องค์ประกอบเชิงความถี่ของแต่ละตัวโน้ต จะทำการวิเคราะห์ในทุกตัวอย่างข้อมูลทั้งหมด 5 ตัวอย่าง จากนั้นนำค่าที่ได้จากการวิเคราะห์ทั้ง 5 ตัวอย่างมาเฉลี่ยค่าองค์ประกอบเชิงความถี่ของโน้ตแต่ละตัว สำหรับค่าความถี่ได้จากการประมาณค่าความถี่เฉลี่ยอีกทีด้วยวิธี curve fitting (รายละเอียดแสดงในภาคผนวก ข) เพื่อให้ได้ค่าความถี่ฮาร์โมนิกต่าง ๆ เป็นจำนวนเท่าของค่าความถี่พื้นฐานพอดี ซึ่งจะทำให้สัญญาณสังเคราะห์ที่ได้ยังคงลักษณะรูปคลื่นเหมือนสัญญาณต้นแบบ ดังแสดงในตารางที่ 4.3 (อนึ่งจากการทดลองเปลี่ยนแปลงค่าเฟสในการสังเคราะห์เสียงพบว่าค่าเฟสไม่มีผลในการสังเคราะห์เสียงแต่อย่างใด จึงมิได้นำค่าเฟสที่วิเคราะห์ได้มาใช้ในการสังเคราะห์)

4.1.2 การวิเคราะห์สัญญาณเสียงในช่วงภาวะที่ไม่เสถียร

เนื่องจากสัญญาณเสียงในช่วงที่มีการเปลี่ยนแปลง (ไม่เสถียร เช่น เสียงในช่วงต้นและช่วงสิ้นสุดของแต่ละโน้ต) เป็นสัญญาณที่มีความยาวไม่มาก ดังแสดงในรูปที่ 4.1 ซึ่งในการวิเคราะห์เชิงสเปกตรัมโดยวิธีการฟูเรียร์ทรานฟอร์มนั้นจะต้องใช้ข้อมูลจำนวนมากพอและสัญญาณที่นำมาวิเคราะห์ควรอยู่ในภาวะเสถียรตลอดช่วงของข้อมูลที่ทำการวิเคราะห์จึงจะได้ค่าความถี่และค่าแอมปลิจูดที่ถูกต้อง แต่เนื่องจากสัญญาณในช่วงนี้มีช่วงที่คงที่ที่แคบมาก จึงยากแก่การคำนวณค่าพารามิตอร์ต่าง ๆ ได้ถูกต้อง ด้วยเหตุนี้จึงทำให้การวิเคราะห์สัญญาณเสียงในช่วงดังกล่าวไม่อาจกระทำได้ อย่างไรก็ดีเนื่องจากเสียงในช่วงภาวะเสถียรและไม่เสถียรเกิดจากกลไกทางฟิสิกส์ที่ใกล้เคียงกัน ผู้วิจัยจึงได้ตั้งสมมติฐานว่าสัญญาณเสียงของโน้ตแต่ละตัวน่าจะประกอบไปด้วยความถี่ที่ค่าต่าง ๆ เหมือนกัน ไม่ว่าจะเป็นช่วงภาวะเสถียรและไม่เสถียร หากแต่มีความแตกต่างกันตรงที่ในช่วงที่ไม่เสถียรค่าแอมปลิจูดมีการเปลี่ยนแปลงตลอดเวลา ดังนั้นเพื่อเป็นการตรวจสอบสมมติฐานดังกล่าว จึงได้ทำการวิเคราะห์โดยอาศัยหลักการสังเคราะห์สัญญาณเสียง (Analysis by Synthesis) โดยใช้ค่าพารามิเตอร์ต่าง ๆ ที่ได้จากการวิเคราะห์ในช่วงภาวะเสถียรข้างต้นมาทำการสังเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 แสดงค่าพารามิเตอร์ของสัญญาณไซน์ลำดับที่ต่าง ๆ ในแต่ละตัวโน้ต

โน้ต	สัญญาณไซน์ลำดับที่ / พารามิเตอร์	1	2	3	4	5	6	7	8	9	10
โด	ความถี่ (เฮิรตซ์)	447.9	895.7	1344	1792	2239	2687	3135	3583	4031	4479
	แอมปลิจูด	1	0.086	0.066	0.01	0.011	0.004	0.005	0.006	6E-04	4E-04
เร	ความถี่ (เฮิรตซ์)	497.2	994.4	1492	1989	2486	2983	3480	3978	4475	4972
	แอมปลิจูด	1	0.277	0.369	0.075	0.022	0.005	0.004	0.002	9E-04	9E-04
มี	ความถี่ (เฮิรตซ์)	549.3	1099	1648	2197	2746	3296	3845	4394	4944	5493
	แอมปลิจูด	1	0.174	0.151	0.008	0.007	0.004	0.002	0.001	7E-04	4E-04
ฟา	ความถี่ (เฮิรตซ์)	596.6	1193	1790	2386	2983	3580	4176	4773	5370	5966
	แอมปลิจูด	1	0.168	0.042	0.003	0.003	0.004	0.001	2E-04	4E-04	2E-04
ซอล	ความถี่ (เฮิรตซ์)	665.8	1332	1997	2663	3329	3995	4660	5326	5992	6658
	แอมปลิจูด	1	0.198	0.054	0.002	0.007	0.002	7E-04	4E-04	3E-04	1E-04
ลา	ความถี่ (เฮิรตซ์)	743.2	1486	2230	2973	3716	4459	5202	5945	6689	7432
	แอมปลิจูด	1	0.067	0.027	0.004	0.009	9E-04	8E-04	3E-04	2E-04	1E-04
ที	ความถี่ (เฮิรตซ์)	836.2	1672	2509	3345	4181	5017	5853	6689	7525	8362
	แอมปลิจูด	1	0.086	0.011	0.004	0.007	4E-04	6E-04	3E-04	2E-04	1E-04
โดสูง	ความถี่ (เฮิรตซ์)	917.3	1835	2752	3669	4587	5504	6421	7338	8256	9173
	แอมปลิจูด	1	0.115	0.015	0.008	0.002	7E-04	4E-04	1E-04	1E-04	1E-04
เรสูง	ความถี่ (เฮิรตซ์)	1013	2027	3040	4054	5067	6080	7094	8107	9121	10134
	แอมปลิจูด	1	0.034	0.006	0.002	0.002	3E-04	6E-04	1E-04	1E-04	1E-04
มีสูง	ความถี่ (เฮิรตซ์)	1120	2240	3359	4479	5599	6719	7839	8958	1008	1120
	แอมปลิจูด	1	0.007	0.015	0.002	0.002	6E-04	3E-04	1E-04	1E-04	1E-04
ฟาสูง	ความถี่ (เฮิรตซ์)	1210	2420	3631	4841	6051	7261	8471	9682	1089	1210
	แอมปลิจูด	1	0.007	0.031	7E-04	0.002	3E-04	2E-04	1E-04	1E-04	1E-04
ซอลสูง	ความถี่ (เฮิรตซ์)	1327	2653	3980	5306	6633	7960	9286	10613	11940	13266
	แอมปลิจูด	1	0.002	0.008	7E-04	0.001	2E-04	1E-04	1E-04	1E-04	1E-04

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

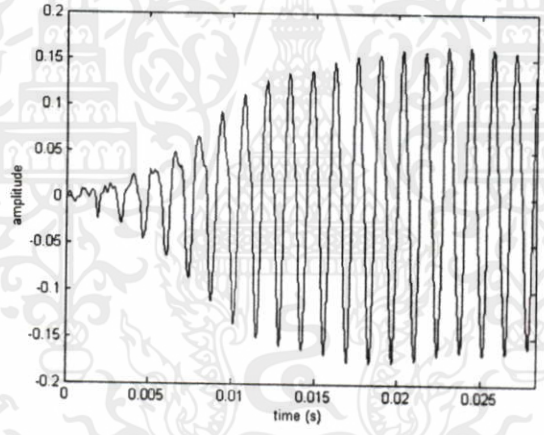
เสียงขึ้น จากนั้นนำสัญญาณเสียงที่สังเคราะห์มาคูณกับเอลเวลลอป (envelop) ของสัญญาณเสียงจริงในช่วงที่มีการเปลี่ยนแปลงและทำการเปรียบเทียบสเปกตรัมของทั้งสองสัญญาณ โดยมีรายละเอียดในการทดลองดังนี้

1) แบ่งสัญญาณเสียงขลุ่ยต้นแบบในช่วงภาวะที่ไม่เสถียร (ทั้งในช่วง attack และ decay) ออกเป็นวินโดว์ย่อย ๆ โดยกำหนดให้มีขนาดเท่ากับ 500 แซมเปิล จากนั้นนำสัญญาณเสียงในแต่ละวินโดว์มาคูณกับแบล็คแมนวินโดว์ และคำนวณหาค่าสเปกตรัมของสัญญาณในทุกวินโดว์โดยใช้ฟาสฟูเรียร์ทรานฟอร์ม

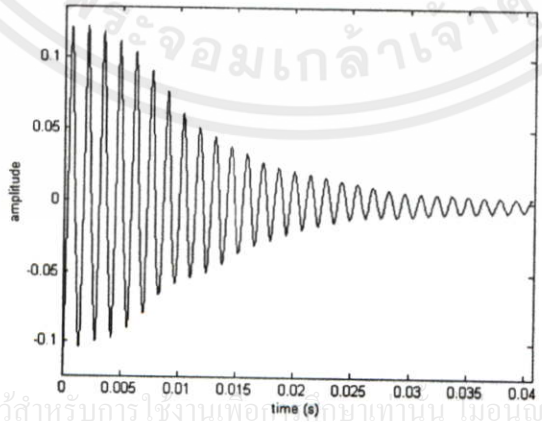
2) นำสัญญาณเสียงสังเคราะห์ซึ่งได้จากการคำนวณตามสมการที่ 2.18 มาคูณกับเอนเวลลอปของสัญญาณเสียงต้นแบบ (เนื่องจากสเปกตรัมของสัญญาณรบกวนมีค่าต่ำมากเมื่อเทียบส่วนสเปกตรัมของสัญญาณไซน์ จึงมิได้นำมาพิจารณาเพื่อการสังเคราะห์เสียง) จากนั้นกระทำในลักษณะเดียวกันกับสัญญาณเสียงต้นแบบในขั้นตอนที่ 1

3) ทำการเปรียบเทียบสเปกตรัมของทั้งสองสัญญาณเสียงในแต่ละวินโดว์

4) ในกรณีที่สเปกตรัมของทั้งสองสัญญาณมีค่าแอมปลิจูดไม่เท่ากัน จะต้องทำการปรับเปลี่ยนค่าแอมปลิจูดที่ใช้ในการสังเคราะห์เสียงใหม่จนกว่าจะได้ลักษณะของสเปกตรัมที่ใกล้เคียงกัน



(ก)



(ข)

รูปที่ 4.1 แสดงลักษณะสัญญาณในช่วงที่มีการเปลี่ยนแปลง

ก. สัญญาณเสียงในช่วง attack

ข. สัญญาณเสียงในช่วง decay

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อพิจารณาสเปกตรัมของเสียงสังเคราะห์ที่ได้เปรียบเทียบกับสัญญาณเสียงจริงในช่วงภาวะที่ไม่เสถียรพบว่าลักษณะสเปกตรัมให้ค่าความถี่ที่เหมือนกัน มีเพียงค่าแอมปลิจูดที่มีค่าไม่เท่ากัน แต่เมื่อทำการปรับเปลี่ยนค่าแอมปลิจูดแล้วทำการสังเคราะห์เสียงใหม่ จะได้สเปกตรัมของสองสัญญาณเสียงที่มีลักษณะใกล้เคียงกันทั้งความถี่และแอมปลิจูด ดังแสดงในรูปที่ 4.2 ซึ่งเป็นตัวอย่างสเปกตรัมเพียงวินโดว์เดียว (รายละเอียดของโน้ตอื่นแสดงในภาคผนวก ง)

จากการวิเคราะห์ข้างต้น สามารถสรุปได้ว่าสัญญาณเสียงที่มีการเปลี่ยนแปลงหรือสัญญาณเสียงที่อยู่ในภาวะที่เสถียร จะมีส่วนประกอบของความถี่ที่เหมือนกัน ซึ่งจะมีเพียงเอนเวลลอปเท่านั้นที่มีการเปลี่ยนแปลง ดังนั้นในการสังเคราะห์สัญญาณเสียงในช่วงเวลาที่มีการเปลี่ยนแปลงจึงสามารถใช้ค่าพารามิเตอร์ชุดเดียวกับสัญญาณเสียงที่อยู่ในภาวะเสถียรได้ โดยจะปรับเปลี่ยนเฉพาะค่าแอมปลิจูดที่แปรตามเวลาเท่านั้น

4.2 การวิเคราะห์ในส่วนของสัญญาณรบกวน

การวิเคราะห์ในส่วนของสัญญาณรบกวนของแต่ละตัวโน้ตจะกระทำโดยการกำหนดส่วนของสเปกตรัมที่มีค่าความชันสูงให้มีค่าเท่ากับศูนย์ ซึ่งจะได้สเปกตรัมในส่วนของสัญญาณรบกวน จากนั้นทำการหาอินเวอร์สฟูเรียร์ทรานฟอร์มเพื่อให้ได้สัญญาณเสียงของส่วนที่เป็นสัญญาณรบกวน (ลักษณะสัญญาณรบกวนของแต่ละตัวโน้ตแสดงในภาคผนวก จ) สัญญาณรบกวนที่ได้นี้สามารถนำมาวิเคราะห์หาสัมประสิทธิ์ตัวกรองสัญญาณสำหรับเอ็มเอโมเดลและเออาร์โมเดลซึ่งมีรายละเอียดดังนี้

4.2.1 การวิเคราะห์ค่าสัมประสิทธิ์ตัวกรองสัญญาณสำหรับเอ็มเอโมเดล

การวิเคราะห์ค่าสัมประสิทธิ์ตัวกรองสัญญาณรบกวนสำหรับเอ็มเอโมเดลจะกระทำกับโน้ตโดเท่านั้น โดยเริ่มจากคำนวณเพาเวอร์สเปกตรัมของสัญญาณรบกวนซึ่งมีขั้นตอนดังนี้

1) แบ่งสัญญาณรบกวนในโดเมนของเวลาออกเป็นวินโดว์ย่อยโดยกำหนดให้มีขนาดเท่ากับ 1000 แซมเปิล โดยให้แต่ละวินโดว์เลื่อมกัน 500 แซมเปิล

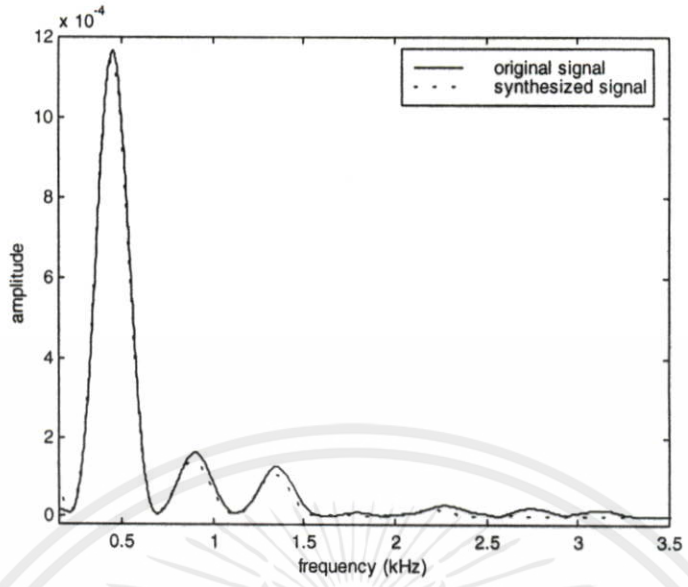
2) คำนวณฟูเรียร์ทรานฟอร์มในแต่ละวินโดว์โดยใช้อัลกอลิทึมฟาสฟูเรียร์ทรานฟอร์ม

3) นำค่าฟูเรียร์ทรานฟอร์มทั้งหมดมาเฉลี่ยด้วยจำนวนวินโดว์ตามสมการที่ 2.8 ซึ่งจะได้เพาเวอร์สเปกตรัมของสัญญาณรบกวนดังตัวอย่างในรูปที่ 4.3

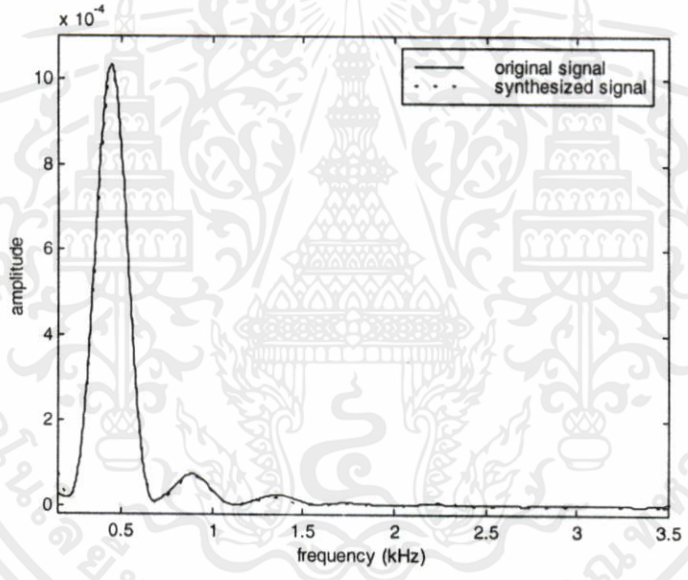
4) นำเพาเวอร์สเปกตรัมที่ได้มาหาอินเวอร์สฟูเรียร์ทรานฟอร์มของฟังก์ชัน $H(\omega)$ ตามสมการที่ 2.12 ซึ่งจะได้ค่าสัมประสิทธิ์ตัวกรองสัญญาณดังแสดงเป็นตัวอย่างตามตารางที่ 4.4

ค่าสัมประสิทธิ์ตัวกรองสัญญาณที่ได้จากคำนวณ จะนำไปใช้ในการเปรียบเทียบความเหมาะสมในการใช้โมเดลสัญญาณรบกวนในหัวข้อถัดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)



(ข)

รูปที่ 4.2 แสดงตัวอย่างการเปรียบเทียบสเปกตรัมของโน้ตโดต้นแบบกับที่สังเคราะห์ขึ้นใหม่

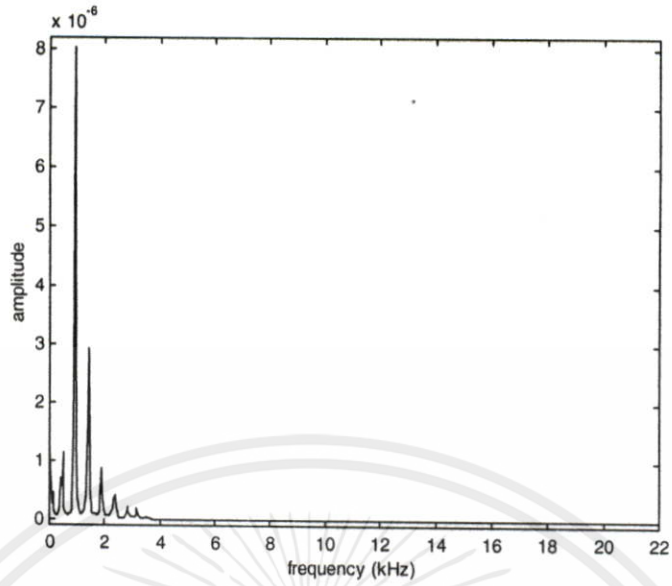
ก. ส่วนที่เป็น attack

ข. ส่วนที่เป็น decay

ตารางที่ 4.4 แสดงค่าสัมประสิทธิ์ตัวกรองสัญญาณที่ได้จากการวิเคราะห์ด้วยเอ็มเอโมเดล

สปส.ตัวกรอง / สัญญาณ / โน้ต	1	2	3	4	5	6	7	8	9	10	11
โด	0.003	0.004	0.006	0.007	0.01	0.014	0.01	0.009	0.007	0.006	0.004

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.3 แสดงเพาเวอร์สเปกตรัมของสัญญาณรบกวนของโน้ตโด

4.2.2 การวิเคราะห์ค่าสัมประสิทธิ์ตัวกรองสัญญาณสำหรับเออาร์โมเดล

ในการวิเคราะห์แบบเออาร์โมเดลจะทำการคำนวณค่าสหสัมพันธ์ของสัญญาณรบกวนของโน้ตโด ณ time lag ต่าง ๆ ตั้งแต่ 0 ถึง จำนวนสัมประสิทธิ์ตัวกรองสัญญาณที่ต้องการ แล้วนำมาคำนวณค่าสัมประสิทธิ์ตัวกรองสัญญาณตามสมการที่ 2.17 ซึ่งค่าสัมประสิทธิ์ตัวกรองสัญญาณดังกล่าวแสดงดังตารางที่ 4.5

ตารางที่ 4.5 แสดงสัมประสิทธิ์ตัวกรองสัญญาณที่ได้จากการวิเคราะห์ด้วยเออาร์โมเดล

สปส.ตัวกรองสัญญาณ / โน้ต	1	2	3	4	5	6	7	8	9	10	11
โด	1	-1.083	-0.228	0.104	0.178	0.139	0.052	-0.013	-0.032	-0.044	-0.023

4.2.3 การเปรียบเทียบความเหมาะสมของการใช้เอ็มเอและเออาร์โมเดลสำหรับการสังเคราะห์ค่าคุณสมบัติทางสถิติของสัญญาณรบกวนที่เหมาะสม

เนื่องจากเอ็มเอโมเดลและเออาร์โมเดลเป็นโมเดลที่สามารถนำมาใช้ในการสังเคราะห์ในสัญญาณรบกวน จึงได้ทำการเปรียบเทียบโมเดลทั้งสองเพื่อเลือกแนวทางที่เหมาะสมที่ใช้จำนวนสัมประสิทธิ์ตัวกรองสัญญาณน้อย ในขณะที่ให้คุณภาพเสียงที่เท่ากัน โดยการเปรียบเทียบจะทำการประเมินทั้งด้านออปเจคทีฟ (objective assessment) และ ซับเจคทีฟ (subjective assessment) โดยจะเลือกใช้สัญญาณรบกวนที่ได้จากโน้ตโด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

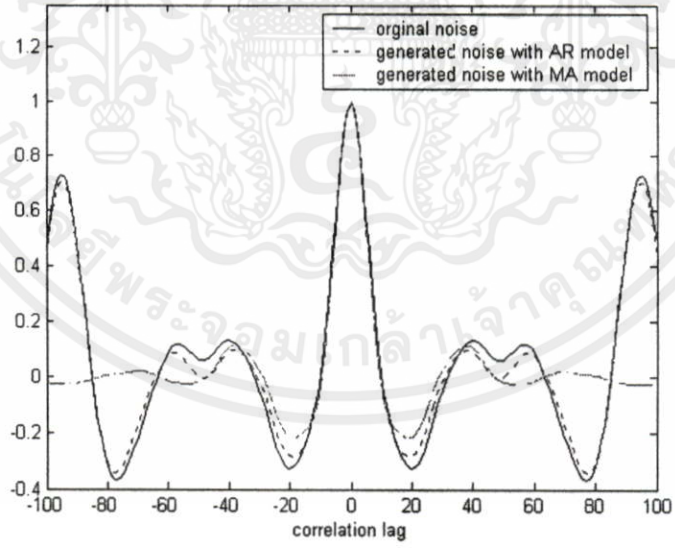
4.2.3.1 การประเมินแบบออปเจคทีฟ

การเปรียบเทียบเอ็มเอและเออาร์โมเดลด้วยการประเมินแบบออปเจคทีฟ จะเป็นการคำนวณค่าความผิดพลาดของค่าสหสัมพันธ์ของสัญญาณรบกวนที่สังเคราะห์จากทั้ง 2 โมเดล โดยจะนำมาเปรียบเทียบกับค่าสหสัมพันธ์ของสัญญาณรบกวนต้นแบบ ซึ่งรายละเอียดมีดังต่อไปนี้

1) สังเคราะห์สัญญาณรบกวนในแต่ละโมเดล โดยใช้จำนวนสัมประสิทธิ์ตัวกรองสัญญาณตั้งแต่ 51 101 151 201 301 และ 501 ค่า ตามที่คำนวณได้จากหัวข้อ 4.2.1 และ 4.2.2

2) นำสัญญาณที่ได้จากขั้นตอนที่ 1 มาหาค่าสหสัมพันธ์เปรียบเทียบกับค่าสหสัมพันธ์ของสัญญาณรบกวนต้นแบบ เพื่อหาค่าความผิดพลาดเฉลี่ยในแต่ละโมเดล โดยค่าความผิดพลาดเฉลี่ยของค่าสหสัมพันธ์ได้จากการหาค่ามีนสแคว์เออร์เรอร์ (MSE ดูรายละเอียดที่ภาคผนวก ข)

จากการเปรียบเทียบทั้งสองโมเดลโดยพิจารณารูปที่ 4.4 พบว่าค่าสหสัมพันธ์ของสัญญาณรบกวนที่ได้จากเออาร์โมเดลนั้นมีความใกล้เคียงสัญญาณรบกวนต้นแบบใกล้เคียงกว่าเอ็มเอโมเดล และเมื่อทำการคำนวณค่าความผิดพลาดเฉลี่ยในแต่ละโมเดลดังแสดงในตารางที่ 4.6 ซึ่งเห็นได้ว่าเออาร์โมเดลให้ผลค่าความผิดพลาดเฉลี่ยน้อยกว่าจากการใช้จำนวนสัมประสิทธิ์ที่เท่ากัน นอกจากนี้ยังพบว่าในระดับค่าเฉลี่ยที่ใกล้เคียงกัน จำนวนสัมประสิทธิ์ตัวกรองสัญญาณที่ได้จาก เอ็มเอโมเดล ต้องใช้มากกว่าถึง 3 เท่า



รูปที่ 4.4 แสดงการเปรียบเทียบค่าสหสัมพันธ์ของสัญญาณต้นแบบกับสัญญาณรบกวนสังเคราะห์ในเอ็มเอโมเดลและเออาร์โมเดลซึ่งใช้จำนวนสัมประสิทธิ์ตัวกรองสัญญาณเท่ากัน 101 จำนวน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.6 แสดงค่าความผิดพลาดเฉลี่ยของค่าสหสัมพันธ์ของสัญญาณรบกวนสังเคราะห์ที่ได้จากทั้งสองโมเดล ตั้งแต่ time lag ที่ 0-501

จำนวนสัมประสิทธิ์ตัวกรอง / วิธีการ / สัญญาณ	51	101	151	201	301	501
เอ็มเอโมเดล	0.0541	0.0537	0.0502	0.0245	0.0194	0.007
เออาร์โมเดล	0.0472	0.0025	0.0021	0.0021	0.0019	0.0025

4.2.3.2 การประเมินแบบซับเจคทีฟ

การประเมินผลแบบซับเจคทีฟ จะเป็นการประเมินผลจากการฟัง ซึ่งผู้ที่เข้ามาทดสอบเป็นผู้ที่มีความเชี่ยวชาญทางด้านดนตรีทั้งสิ้นจำนวน 22 คน ในการวัดผลครั้งนี้ ได้ให้ผู้เข้ารับการทดสอบเปรียบเทียบเสียงสังเคราะห์จากทั้งสองแนวทาง โดยต้องฟังเสียงต้นแบบก่อน แล้วพิจารณาว่าเสียงใดคล้ายคลึงเสียงสัญญาณรบกวนต้นแบบมากกว่า จากนั้นนำข้อมูลที่ได้มาเก็บเป็นคะแนนดังแสดงตามตารางที่ 4.7 ซึ่งรายละเอียดของการเตรียมข้อมูลเพื่อการทดสอบครั้งนี้มีดังนี้

1) สังเคราะห์เสียงสัญญาณรบกวนโดยใช้สัมประสิทธิ์ตัวกรองสัญญาณของเออาร์โมเดล จำนวน 100 ค่า

2) สังเคราะห์เสียงสัญญาณรบกวนโดยใช้สัมประสิทธิ์ตัวกรองสัญญาณของเอ็มเอโมเดล จำนวน 100 200 500 800 และ 1000 ค่า

3) ทำการเปรียบเทียบเสียงสังเคราะห์กับสัญญาณต้นแบบทั้งหมด 5 ครั้ง ซึ่งในแต่ละครั้ง จะให้ฟังเสียงต้นแบบ เสียงสังเคราะห์แบบเออาร์โมเดลที่สัมประสิทธิ์ตัวกรองสัญญาณ100 ค่า และเสียงสังเคราะห์แบบเอ็มเอโมเดล(แต่ละครั้งจะใช้เสียงสังเคราะห์แบบเอ็มเอโมเดลที่ค่าสัมประสิทธิ์ตัวกรองสัญญาณที่ต่างกัน) จากนั้นทำการเปรียบเทียบว่าเสียงใดมีความใกล้เคียงเสียงต้นแบบมากกว่า โดยหลักการให้คะแนนมีดังนี้

- เสียงสังเคราะห์ใดที่มีความใกล้เคียงเสียงต้นแบบมากกว่า จะให้คะแนนเท่ากับ 3
- เสียงสังเคราะห์ใดที่มีความใกล้เคียงเสียงต้นแบบน้อยกว่า จะให้คะแนนเท่ากับ 1
- เสียงสังเคราะห์ทั้งสองใกล้เคียงเสียงต้นแบบเท่ากัน จะให้คะแนนเท่ากับ 2 ทั้งคู่

ตารางที่ 4.7 แสดงระดับคุณภาพของสัญญาณรบกวนสังเคราะห์

วิธีการสังเคราะห์สัญญาณรบกวน	เออาร์โมเดล	เอ็มเอโมเดล				
จำนวนสัมประสิทธิ์ตัวกรองสัญญาณ	100	100	200	500	800	1000
ระดับคุณภาพเฉลี่ย (1-3)	2.5	1.1	1.1	1.2	1.9	1.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการทดลองพบว่า แนวทางการสังเคราะห์ในส่วนของสัญญาณรบกวนด้วยเออาร์โมเดล จะให้สัญญาณเสียงสังเคราะห์ที่ใกล้เคียงเสียงต้นแบบมากกว่า ดังนั้นขั้นตอนการวิเคราะห์ค่าสัมประสิทธิ์ตัวกรองสัญญาณในส่วนของสัญญาณรบกวน จึงใช้วิธีการเออาร์โมเดลโดยทำการคำนวณหาจากสัญญาณรบกวนของแต่ละตัวโน้ตทั้ง 5 ตัวอย่างสัญญาณ แล้วนำมาเฉลี่ยดังแสดงในรูปที่ 4.5

4.3 การเปรียบเทียบแนวทางการสังเคราะห์เสียงแบบ Sinusoidal Model และ Sinusoidal Plus Noise Model

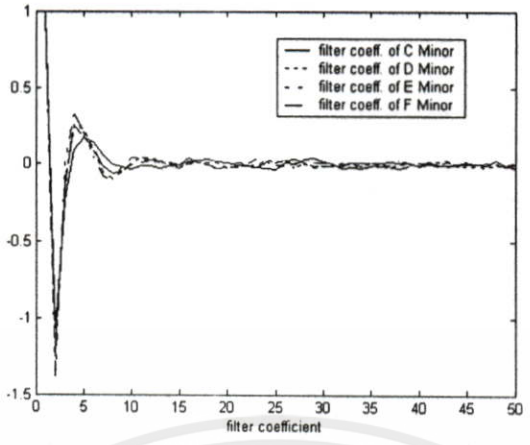
ในการเปรียบเทียบแนวทางการสังเคราะห์เสียงแบบ Sinusoidal Model และ Sinusoidal Plus Noise Model ได้ทำการทดลองเพื่อศึกษาถึงความซับซ้อนในการคำนวณของแต่ละแนวทางโดยคำนึงถึงคุณภาพของเสียง ซึ่งในการทดลองจะทำการเปรียบเทียบจำนวนพารามิเตอร์ที่ใช้ในการสังเคราะห์เสียงกับความซับซ้อนในการคำนวณ

4.3.1 จำนวนพารามิเตอร์ที่ใช้ในการสังเคราะห์ต่อคุณภาพเสียง

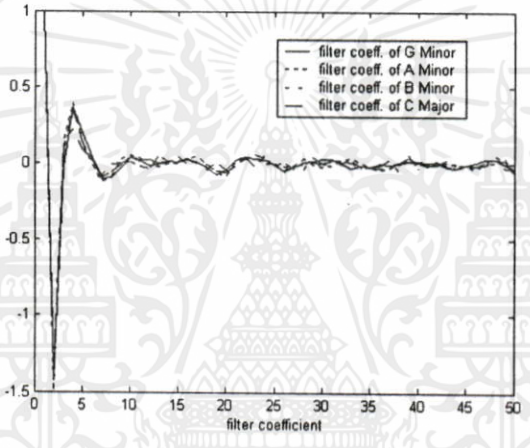
ในการทดลองนี้จะเปรียบเทียบจำนวนพารามิเตอร์ที่ใช้ในการสังเคราะห์เสียงโน้ตโด โดยที่การสังเคราะห์แบบ Sinusoidal Model จะสร้างสัญญาณเสียงที่ประกอบด้วยจำนวนสัญญาณไซน์ตั้งแต่ 100 500 1000 5000 10000 สัญญาณ เพื่อนำมาเปรียบเทียบกับแนวทางการสังเคราะห์เสียงแบบ Sinusoidal Plus Noise Model ที่ใช้จำนวนพารามิเตอร์ในการสร้างส่วนของสัญญาณไซน์จำนวน 20 ค่า และในส่วนของสัญญาณรบกวน 100 ค่า โดยในขั้นตอนของการประเมินนั้น ได้ให้ผู้ฟังฟังเสียงโน้ตโดต้นแบบก่อน จากนั้นให้ฟังเสียงสังเคราะห์จากทั้งสองแนวทางเปรียบเทียบกันว่าเสียงใดใกล้เคียงเสียงต้นแบบมากกว่า โดยเสียงสังเคราะห์ที่ได้จากแนวทางการสังเคราะห์เสียงแบบ Sinusoidal Plus Noise จะนำไปเปรียบเทียบกับเสียงสังเคราะห์ที่ได้จากแนวทางการสังเคราะห์เสียงแบบ Sinusoidal Model ทุกเสียง (ทั้ง 5 สัญญาณเสียง) ซึ่งสามารถสรุปผลการให้คะแนนของผู้ฟังจำนวน 22 คนได้ในตารางที่ 4.8 (รายละเอียดการให้คะแนนเหมือนกับหัวข้อ 4.2.3.2)

แม้ว่าวิธีการสังเคราะห์เสียง Sinusoidal Model จะใช้จำนวนซินนูซอยน์ถึง 10000 จำนวน แต่คุณภาพเสียงที่ได้ก็ยังไม่สามารถนำมาเทียบกับวิธีการสังเคราะห์เสียงแบบ Sinusoidal Plus Noise ซึ่งมีส่วนโมเดลของสัญญาณรบกวนที่ทำให้จำนวนพารามิเตอร์ในการสังเคราะห์เสียงนั้นลดลง อีกทั้งให้คุณภาพเสียงที่ใกล้เคียงเสียงต้นแบบ

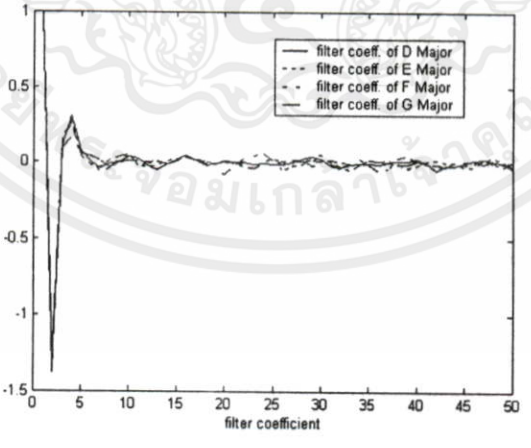
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)



(ข)



(ค)

รูปที่ 4.5 แสดงค่าสัมประสิทธิ์ตัวกรองสัญญาณที่ได้จากวิเคราะห์แบบเออาร์โมเดล
ก. ค่าสัมประสิทธิ์ตัวกรองสัญญาณของโน้ตโด จนถึง โน้ตฟา
ข. ค่าสัมประสิทธิ์ตัวกรองสัญญาณของโน้ตซอล จนถึง โน้ตโดสูง
ค. ค่าสัมประสิทธิ์ตัวกรองสัญญาณของโน้ตเรสูง จนถึง โน้ตซอลสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.8 แสดงการเปรียบเทียบระดับคุณภาพเสียงสังเคราะห์จากวิธีการสังเคราะห์เสียงแบบ Sinusoidal Model และ Sinusoidal Plus Noise Model

วิธีการสังเคราะห์เสียง	Sinusoidal Plus Noise Model		Sinusoidal Model				
จำนวนพารามิเตอร์	ซินนูซอยด์	สปส.ตัวกรองสัญญาณ	ซินนูซอยด์				
	10	100	100	500	1000	5000	10000
ระดับคุณภาพเฉลี่ย (1-3)	2.7		1.05	1.22	1.3	1.4	2.1

4.3.2 ความซับซ้อนในการคำนวณ

การวัดความซับซ้อนในการคำนวณจะทำการวัดจำนวนครั้งของการคำนวณในการสร้างเสียงสังเคราะห์ของแต่ละวิธี (โดยจะกำหนดให้การคำนวณในลักษณะ บวก ลบ คูณ และหาร จะถูกนับการคำนวณเป็นหนึ่งเท่ากัน) จากการสังเคราะห์เสียงในแต่ละแซมเปิลที่อัตราการแซมปลิงที่ 44100 เฮิรตซ์ พบว่า

- Sinusoidal Model ซึ่งใช้จำนวนส่วนประกอบเชิงความถี่ 10000 ค่าความถี่ ให้การคำนวณเท่ากับ 40,000 ครั้ง/แซมเปิล (flops/sample)
- Sinusoidal Plus Noise Model ซึ่งใช้จำนวนส่วนประกอบเชิงความถี่ 10 ค่าความถี่ และใช้จำนวนสัมประสิทธิ์ตัวกรองสัญญาณในส่วนของสัญญาณรบกวนอีก 100 ค่า ให้การคำนวณเท่ากับ 241 ครั้ง/แซมเปิล (flops/sample)

จะเห็นได้ว่ามีความซับซ้อนในการคำนวณจากการใช้แนวทางการสังเคราะห์เสียงโดยวิธี Sinusoidal Model มากกว่า Sinusoidal Plus Noise Model ประมาณ 166 เท่า ในระดับของคุณภาพเสียงที่ใกล้เคียงกัน ดังนั้นการประยุกต์ใช้แนวทางดังกล่าวเพื่อสร้างแอพพลิเคชั่นของการสังเคราะห์เสียงในลักษณะเรียลไทม์จึงไม่เหมาะสม

4.4 การประเมินระดับคุณภาพของเสียง

การประเมินระดับคุณภาพของเสียงสังเคราะห์ที่ใช้แนวทางการสังเคราะห์เสียงแบบ Sinusoidal Plus Noise Model จะเน้นพิจารณาที่จำนวนพารามิเตอร์ที่ใช้ในส่วนของการสร้างสัญญาณรบกวนเป็นหลัก เนื่องจากว่าคุณภาพของเสียงสังเคราะห์ขึ้นอยู่กับจำนวนพารามิเตอร์ในส่วนนี้ด้วย แต่อย่างไรก็ตาม การใช้พารามิเตอร์จำนวนมากอาจให้เสียงที่มีคุณภาพไม่แตกต่างจากการใช้จำนวนพารามิเตอร์ที่น้อย ดังนั้นในการประเมินระดับคุณภาพเสียง จะทำการหาจำนวนพารามิเตอร์ที่น้อยที่สุดที่ยังคงให้เสียงสังเคราะห์ที่มีคุณภาพใกล้เคียงกับเสียงต้นแบบ โดยขั้นตอนในการเตรียมข้อมูลเพื่อทำการประเมินผลมีดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) สร้างสัญญาณไซน์ที่ความถี่และแอมปลิจูดต่าง ๆ (แล้วแต่ชนิดของตัวโน้ต) ตามสมการที่ 2.18 โดยกำหนดให้ n มีค่าเท่ากับจำนวนแซมเปิลของสัญญาณเสียงต้นแบบ

2) สร้างสัญญาณรบกวนโดยใช้จำนวนสัมประสิทธิ์ตัวกรองสัญญาณตั้งแต่ 2 15 50 100 ตามสมการที่ 2.20 โดยกำหนดให้ n มีค่าเท่ากับจำนวนแซมเปิลของสัญญาณเสียงต้นแบบ (ในการวัดคุณภาพเสียงครั้งนี้ได้ใช้สัมประสิทธิ์ตัวกรองสัญญาณ 2 จำนวนเป็นตัวแทนสัมประสิทธิ์ตัวกรองสัญญาณที่อยู่ระหว่าง 2-10 จำนวนเพื่อการเปรียบเทียบระดับคุณภาพเสียง เนื่องจากว่าในการทดลองเบื้องต้นจำนวนสัมประสิทธิ์ตัวกรองสัญญาณระหว่าง 2-10 ค่า ไม่ได้ทำให้เสียงสังเคราะห์มีความใกล้เคียงเสียงต้นแบบ)

3) นำสองส่วนมารวมกันตามสมการที่ 2.2 แล้วคูณด้วยเอนเวลลอปของสัญญาณเสียงต้นแบบ

4) ทำการสร้างเสียงสังเคราะห์ตามขั้นตอน 1-3 ในทุก ๆ ตัวโน้ต

จากนั้นนำเสียงสังเคราะห์ทั้งหมดที่ได้ประเมินคุณภาพของเสียงจากการฟังโดยให้ผู้ที่มีความเชี่ยวชาญทางด้านดนตรีไทยจำนวน 30 ท่าน ซึ่งเป็นนักศึกษาปี 3 และปี 4 เอกดนตรีไทยและสากลในสาขาเครื่องเป่าเป็นส่วนใหญ่ โดยในการประเมินคุณภาพของเสียงแต่ละตัวโน้ตนั้นจะให้เปรียบเทียบเสียงโน้ตทั้งหมด 6 ครั้ง (4 เสียงสังเคราะห์เวียนเปรียบเทียบกันทีละคู่) ในแต่ละครั้งจะให้ฟังเสียง 2 คู่ โดยคู่แรกคือเสียงต้นแบบกับเสียงสังเคราะห์ที่ 1 คู่ที่ 2 คือเสียงต้นแบบกับเสียงสังเคราะห์ที่ 2 จากนั้นเปรียบเทียบความเหมือนระหว่าง 2 คู่เสียงนี้ ซึ่งมีหลักการพิจารณาดังนี้

- คู่เสียงใด มีเสียงที่ใกล้เคียงกันมากกว่าเมื่อเทียบกับอีกคู่ จะให้คะแนนเท่ากับ 3
- คู่ที่มีความเหมือนน้อยกว่าจะให้คะแนนเท่ากับ 1
- ทั้งสองคู่มีระดับความเหมือนไม่แตกต่าง จะให้คะแนนเท่ากับ 2 ทั้งสองคู่

จากนั้นนำคะแนนทั้งหมดมาหาค่าเฉลี่ยแล้วปรับสเกลระดับคะแนนให้อยู่ในช่วง 1-5 โดยระดับคะแนนที่ 5 เป็นคะแนนเสียงสังเคราะห์ที่ได้คะแนนสูงสุด ดังแสดงในตารางที่ 4.9

เหตุที่ผู้วิจัยไม่ได้ทำการกำหนดระดับคะแนนให้อยู่ในช่วง 1-5 ในตอนแรกของการประเมินเนื่องจากว่าเสียงสังเคราะห์ที่ได้มีความใกล้เคียงกันมาจนผู้ที่ทำการประเมินคุณภาพของเสียงไม่สามารถบอกได้ว่าเสียงสังเคราะห์นั้นอยู่ที่ระดับคะแนนที่เท่าไร ดังนั้นผู้วิจัยจึงใช้วิธีการเปรียบเทียบความเหมือนของเสียงสังเคราะห์แทนดังได้กล่าวมาข้างต้น จากนั้นจึงนำมาปรับสเกลให้อยู่ในช่วง 1-5

ในการประเมินผลการฟัง ผู้เข้ารับการทดสอบครั้งนี้กล่าวโดยรวมว่าทุกเสียงสังเคราะห์ยังคงลักษณะเสียงขลุ่ยไทยทั้งสิ้น แต่สิ่งที่แตกต่างกันคือเสียงของลม (ซึ่งก็คือส่วนของสัญญาณรบกวน) โดยจะมากน้อยขึ้นอยู่กับผู้เป่า เสียงสังเคราะห์แต่ละเสียงอาจเป็นการเป่าคนละครั้งหรือเป็นคนละคนเป่าก็ได้ ดังนั้นสิ่งที่ทำให้เสียงสังเคราะห์ไม่ได้มีความเหมือนเสียงต้นแบบ อาจเปรียบได้กับการได้ยินเสียงขลุ่ยที่เป่าคนละครั้ง ทั้งนี้การสร้างสัญญาณรบกวนแต่ละครั้งก็เทียบได้กับลมที่เป่าออกไปแต่ละครั้ง โดยจากการฟังเสียงโน้ตต้นแบบที่ได้จากการบันทึกหลายครั้ง ผู้เข้ารับการทดสอบได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารนี้ทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.9 แสดงผลการประเมินคุณภาพของเสียง จากการใช้ค่าสัมประสิทธิ์ตัวกรองสัญญาณที่จำนวนต่าง ๆ

ระดับคุณภาพเสียง	จำนวนสัมประสิทธิ์ตัวกรองสัญญาณ			
โน้ต	100	50	15	2
โด	5	4.4	5	3.3
เร	4.8	5	4.5	2.3
มี	4.8	5	4.8	3.4
ฟา	4.9	5	4.6	3.5
ซอล	4.9	5	4.5	2.6
ลา	4.6	4.4	4.6	2.9
ที	5	4.6	4.6	3.5
โดสูง	4.8	4.8	5	3.5
เรสูง	4.7	4.6	5	4.4
มีสูง	4.7	4.6	5	4.5
ฟาสูง	4.9	4.9	5	5
ซอลสูง	4.7	4.6	5	4.1
ค่าเฉลี่ยจากทุกตัวโน้ต	4.8	4.7	4.8	3.5

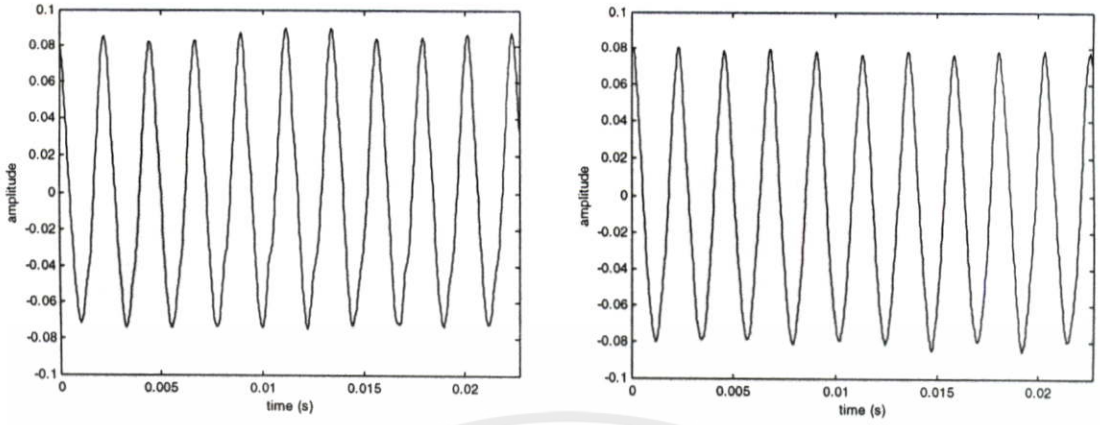
กล่าวว่าเสียงดังกล่าวไม่เหมือนกัน (เหมือนในที่นี้ มีความหมายว่า เสียงสองเสียงไม่มีความแตกต่างกันเลย เหมือนกับได้ยินเสียงเดียวกันสองครั้ง)

จากผลการพิจารณาคุณภาพของเสียง พบว่าจำนวนสัมประสิทธิ์ตัวกรองสัญญาณที่ใช้ในการสังเคราะห์ในส่วนของสัญญาณรบกวนเท่ากับ 15 จำนวน ให้ระดับคุณภาพเสียงโดยเฉลี่ยทุกตัวโน้ตอยู่ที่ 4.8 ซึ่งหากทำการเพิ่มจำนวนสัมประสิทธิ์ตัวกรองสัญญาณมากกว่า 15 จำนวน ระดับคุณภาพเสียงที่ได้ไม่ได้มีความแตกต่างกันอย่างมีนัยสำคัญ อย่างไรก็ดี ในการประเมินคุณภาพเสียงโดยเฉพาะในโน้ตเสียงสูง ผู้รับการทดสอบส่วนใหญ่กล่าวว่าโน้ตเสียงสูง จะมีเสียงลมไม่ชัดเจนเมื่อเทียบกับโน้ตเสียงต่ำ แต่จะมีระดับเสียงพิท (pitch) ชัดเจนมากกว่า จึงยากแก่การตัดสินได้ว่าเสียงสังเคราะห์ใดคล้ายเสียงต้นแบบมากกว่ากัน ซึ่งทำให้ระดับคะแนนที่ได้จากการประเมินในบางตัวโน้ต จะมีค่าสูงจากการใช้จำนวนสัมประสิทธิ์ตัวกรองสัญญาณเพียง 2 ค่า

นอกจากนี้ลักษณะของสัญญาณเสียงสังเคราะห์ยังคล้ายคลึงกับสัญญาณต้นแบบ ดังแสดงตัวอย่างของโน้ตโดในรูปที่ 4.6 (ในรายละเอียดของลักษณะของสัญญาณเสียงสังเคราะห์ทั้งหมดตั้งแต่โน้ตโด ถึง โน้ตซอลสูงแสดงในภาคผนวก ฉ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการประยุกต์ใช้วิธีการสังเคราะห์เสียงแบบ Sinusoidal Plus Noise Model กับเสียงเครื่องดนตรีขลุ่ยไทย ทำให้ทราบว่าเสียงขลุ่ยไทยประกอบด้วยความถี่พื้นฐานและความถี่ฮาร์โมนิก



(ก) (ข)

รูปที่ 4.6 แสดงการเปรียบเทียบสัญญาณเสียงต้นแบบ (ก) และสัญญาณเสียงสังเคราะห์ (ข) ของโน้ตโด

จำนวน 10 ความถี่ (ในบางตัวโน้ตมีมากกว่า 10 ค่าความถี่ แต่เนื่องจากว่าค่าแอมปลิจูดที่ความถี่เหล่านี้มีค่าน้อยมากเมื่อเทียบกับค่าความถี่พื้นฐาน ซึ่งไม่มีผลต่อการสร้างเสียงสังเคราะห์ จึงไม่ได้แสดงมา ณ ที่นี่) และในส่วนของการสังเคราะห์สัญญาณโดยคำนึงถึงความซับซ้อนในการคำนวณและคุณภาพของเสียงสังเคราะห์ที่ได้ ต้องใช้ค่าสัมประสิทธิ์ตัวกรองสัญญาณจำนวน 15 ค่าขึ้นไป (เมื่อรวมกับส่วนของสัญญาณคลื่นไซน์) จึงจะได้เสียงสังเคราะห์ที่ยังคงลักษณะเสียงขลุ่ยไทย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
การพัฒนาโปรแกรมเพื่อการสังเคราะห์เสียงขลุ่ยไทยแบบเรียลไทม์

ปัจจุบันได้มีการพัฒนาซอฟต์แวร์ซินทีไซเซอร์เพื่อการสังเคราะห์เสียงขลุ่ยไทยขึ้นมา โดยอาศัยหลักการของเวปเทเบิลในการสังเคราะห์ แม้ว่าจะให้เสียงที่เหมือนจริงแต่ขาดความยืดหยุ่นในการปรับแต่งเสียง สำหรับงานวิจัยนี้ได้ทำการพัฒนาซอฟต์แวร์ซินทีไซเซอร์สำหรับการสังเคราะห์เสียงขลุ่ยไทยเช่นกัน โดยอาศัยโมเดลการสังเคราะห์เสียงแบบ Sinusoidal Plus Noise ซึ่งไม่เหมือนกับซินทีไซเซอร์ที่มีอยู่ในปัจจุบัน โมเดลดังกล่าวจะให้เสียงขลุ่ยสังเคราะห์ที่มีความเป็นธรรมชาติ เนื่องจากโมเดลของสัญญาณรบกวน (ซึ่งเป็นโมเดลของลมที่เป่าออกมาในแต่ละครั้ง) จะสร้างสัญญาณที่มีการเปลี่ยนแปลงตามคุณสมบัติทางสถิติ (Stochastic signal) ซึ่งทำให้เสียงสังเคราะห์ที่ได้มีความเป็นธรรมชาติ

5.1 แนวคิดในการพัฒนาโปรแกรม

ในการพัฒนาซอฟต์แวร์ซินทีไซเซอร์ จะเป็นการสังเคราะห์เสียงแบบ Sinusoidal Plus Noise เมื่อผู้ใช้กดคีย์บอร์ด ซินทีไซเซอร์จะทำการสังเคราะห์และเล่นเสียงโน้ตดนตรีของขลุ่ยไทยจนกระทั่งผู้ใช้ปล่อยคีย์บอร์ด ซึ่งจะเหมือนกับการใช้ซินทีไซเซอร์ที่มีอยู่ในปัจจุบัน นอกจากนี้แล้วซอฟต์แวร์ซินทีไซเซอร์นี้ ยังสามารถเพิ่มชนิดของเครื่องดนตรีอื่นเข้าไปได้ โดยต้องมีลักษณะคุณสมบัติที่อยู่ภายใต้โมเดลการสังเคราะห์เสียงแบบ Sinusoidal Plus Noise ทั้งนี้เพื่อประโยชน์ในการใช้ซอฟต์แวร์ สำหรับการเล่นเครื่องดนตรีชนิดนั้นได้อีกด้วย

สำหรับขั้นตอนการเล่นเสียงโน้ตดนตรีออกมานั้น จำเป็นต้องติดต่อกับการ์ดเสียงเพื่อส่งข้อมูลที่ได้สังเคราะห์ขึ้นมาไปยังบัฟเฟอร์ของการ์ดเสียง โดยในการติดต่อกับการ์ดเสียงได้ใช้ DirectX ทำให้การติดต่อกับการ์ดเสียงง่ายขึ้น และในขณะเดียวกันผู้วิจัยเองไม่จำเป็นต้องสร้างแอพพลิเคชันให้รองรับการ์ดเสียงทุกรุ่น เนื่องจากว่าแอพพลิเคชันดังกล่าวสามารถเรียกใช้คุณสมบัติหรือคุณลักษณะของฮาร์ดแวร์ได้โดยผ่าน DirectX แม้รันภายใต้สภาพแวดล้อมที่ต่างกัน (ในที่นี้อาจเป็นการใช้การ์ดเสียงต่างรุ่นกัน)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ความต้องการของระบบในการพัฒนาโปรแกรม

ในการพัฒนาซอฟต์แวร์ซินทีไซเซอร์

- คุณสมบัติทางด้านฮาร์ดแวร์

หน่วยประมวลผลกลาง Intel Pentium II 350

ขนาดหน่วยความจำ 64 MB

ความจุฮาร์ดดิส 6.5 GB

การ์ดเสียง Yamaha OPL3-Sax WDM

- ระบบปฏิบัติการ

วินโดว์ 98

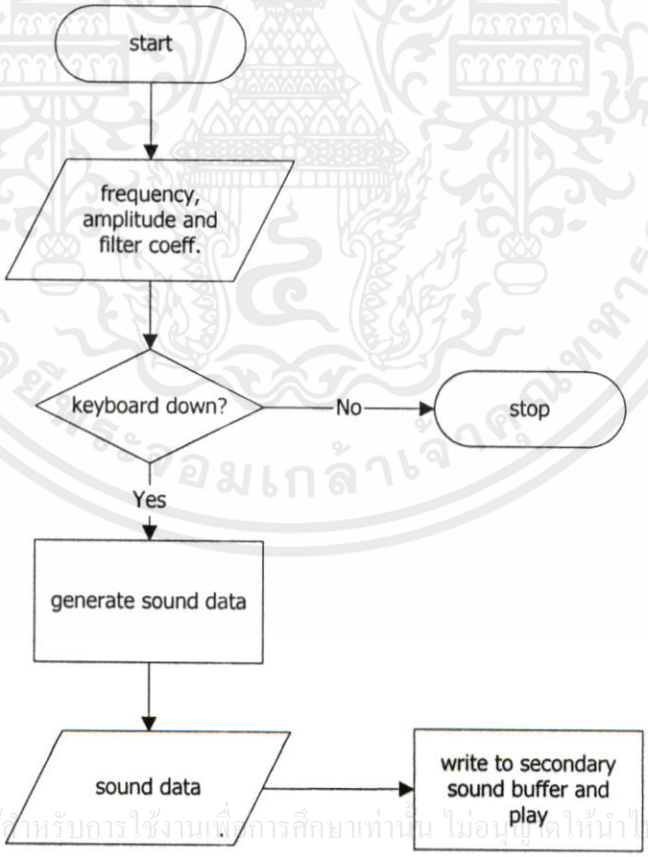
- เครื่องมือที่ใช้ในการพัฒนา

DirectX 8.0

Visucl C++ 6.0

5.3 หลักการทำงานของโปรแกรม

หลักการทำงานโดยทั่วไปจะคล้ายกับฮาร์ดแวร์ซินทีไซเซอร์ คือเมื่อผู้ใช้กดคีย์บอร์ดหรือเมาส์ โปรแกรมจะทำการเล่นเสียงดนตรี และเมื่อผู้ใช้ยกนิ้วออกจากคีย์บอร์ดหรือเมาส์ เสียงดนตรีจะหยุดเล่น โดยลักษณะไดอะแกรมการทำงานแสดงดังรูปที่ 5.1



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

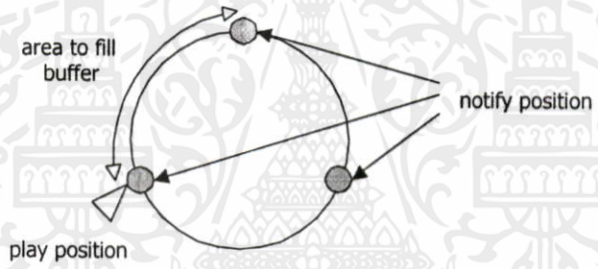
รูปที่ 5.1 แสดงลักษณะไดอะแกรมการทำงานของโปรแกรมการสังเคราะห์เสียง

ขั้นตอนการทำงานของโปรแกรมมีค่าแอมปลิจูด ค่าความถี่ และค่าสัมประสิทธิ์ตัวกรองสัญญาณ (ที่ได้จากการวิเคราะห์สเปกตรัมของสัญญาณเสียง) เป็นค่าอินพุต ส่วนเอาต์พุตได้จากการนำค่าพารามิเตอร์เหล่านี้มาคำนวณภายใต้โมเดล Sinusoidal Plus Noise ซึ่งในรายละเอียดแสดงดังนี้

1.อ่านค่าแอมปลิจูด ค่าความถี่ และสัมประสิทธิ์ตัวกรองสัญญาณจากไฟล์ข้อมูล

2.มีการรันโปรเซส 2 โปรเซสเพื่อรอรับคำสั่งให้ทำงาน

- โปรเซสแรกจะทำการคำนวณค่าเอาต์พุตจำนวนหนึ่ง ทันทีที่มีอีเวนท์ (Event) ส่งมา
- โปรเซสที่สอง เป็นโปรเซสที่เขียนข้อมูลลงในเซ็กเกิลดารี บัฟเฟอร์ (Secondary buffer) ของการ์ดเสียง เมื่อตำแหน่งที่เล่นเสียงดนตรี (Play position) ในบัฟเฟอร์มาถึง ณ ตำแหน่งที่ทำการมาร์คไว้ (Notify position) จะส่งอีเวนท์ไปให้โปรเซสเขียนข้อมูลลงบัฟเฟอร์ เนื่องจากบัฟเฟอร์ดังกล่าวเป็นเซอร์เคิลบัฟเฟอร์ (Circle buffer) ในการเขียนข้อมูลจะเขียนข้อมูลตั้งแต่ตำแหน่งมาร์คก่อนหน้าตำแหน่งมาร์คปัจจุบัน จนถึงตำแหน่งมาร์คปัจจุบัน ดังแสดงในรูปที่ 5.2



รูปที่ 5.2 แสดงเซ็กเกิลดารีบัฟเฟอร์ของการ์ดเสียง

3. ทันทีที่กดคีย์บอร์ดลง จะทำการกำหนดการทำงานของสองไทม์เมอร์ (Timer)

- ไทม์เมอร์แรก (Periodic Timer)จะส่งอีเวนท์ทุก 0.1 วินาทีไปให้โปรเซสแรกทำการคำนวณค่าเอาต์พุตจำนวน 2205 แซมเปิล เนื่องจากว่าอัตราการเล่นเสียงดนตรีของโปรแกรมนี้อยู่ที่ 22050 แซมเปิล/วินาที ซึ่งแสดงว่า ภายใน 1 วินาทีโมเดลการคำนวณควรจะคำนวณค่าข้อมูลทั้งหมด 22050 แซมเปิล แต่เนื่องจากภายใต้สิ่งแวดล้อมที่ได้ทำการรันโปรแกรมนี้จะให้ข้อมูลมากถึง 40000 แซมเปิลต่อวินาทีซึ่งเกินความต้องการ ดังนั้นในกระบวนการคำนวณเพื่อที่จะได้จำนวนเอาต์พุตตามต้องการจึงจำเป็นต้องกำหนดให้ทุก 0.1 วินาที คำนวณค่าข้อมูลจำนวน 2205 แซมเปิล
- ไทม์เมอร์ที่สอง (Oneshot Timer) เป็นไทม์เมอร์ที่กำหนดให้ทำการเขียนข้อมูลลงบัฟเฟอร์และทำการเล่น หลังจากที่กดคีย์บอร์ดไปแล้ว 0.05 วินาที เนื่องจากว่าการเขียนข้อมูลลงบัฟเฟอร์ ในครั้งแรกนั้นต้องเขียนลงให้เต็มบัฟเฟอร์ (ซึ่งโปรแกรมนี้ กำหนดให้มีขนาด 22050*0.05*y

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งไม่อนุญาตให้นำไปใช้

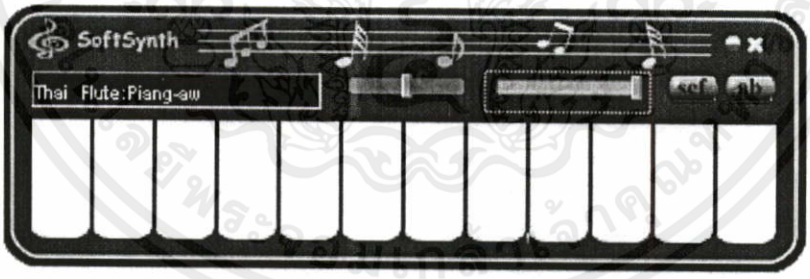
ไบต์ โดยที่ y เป็นจำนวนไบต์ใน 1 แซมเปิล) ดังนั้นทำให้มีความล่าช้าในการเล่นไปประมาณ 0.05 วินาที

4. เมื่อปล่อยคีย์บอร์ดขึ้น จะได้จำนวนแซมเปิลของข้อมูลทั้งหมดตั้งแต่เริ่มกดคีย์บอร์ดจนกระทั่งปล่อยคีย์บอร์ด จากนั้นจะทำการหยุดการทำงานของไทม์เมอร์แรก (Kill timer) เพื่อสิ้นสุดขั้นตอนการคำนวณค่าเอาต์พุต สำหรับจำนวนแซมเปิลที่ได้จะนำมาใช้ในการหยุดกระบวนการเล่นเสียงในบัฟเฟอร์ โดยจะทำการเปรียบเทียบจำนวนข้อมูลทั้งหมดที่เล่นไปแล้วกับจำนวนข้อมูลจริงที่ต้องเล่น โดยเมื่อจำนวนข้อมูลทั้งหมดที่เล่นไปแล้วมากกว่าหรือเท่ากับจำนวนข้อมูลจริงที่ต้องเล่นก็จะสั่งหยุดเล่นทันที

นอกจากนี้แอพพลิเคชันนี้ยังมีฟังก์ชันพิเศษที่สามารถเพิ่มชนิดเครื่องดนตรีเข้ามา แต่ต้องมีคุณลักษณะของค่าพารามิเตอร์ต่างๆ อยู่ภายใต้โมเดล Sinusoidal Plus Noise และผู้ใช้สามารถเลือกเล่นเครื่องดนตรีชนิดนั้นได้อีกด้วย

5.4 การใช้งานโปรแกรม

โปรแกรมนี้มีลักษณะอินเทอร์เฟซคล้ายกับคีย์บอร์ด ทำให้การใช้งานง่าย โดยผู้ใช้สามารถกดที่คีย์บอร์ดในแถวที่สอง (q ,w,e,r,t,y,u,I,o,p,[,]) ซึ่งเริ่มจาก q คือโน้ตโด ไปจนถึง] คือโน้ตซอลสูง และในกรณีที่ผู้ใช้เมาส์ สามารถเลื่อนเมาส์ไปที่รูปคีย์บอร์ดและกดปุ่มเมาส์ซ้าย ดังแสดงในรูปที่ 5.3



รูปที่ 5.3 แสดงอินเทอร์เฟซของซอฟต์แวร์ซินทีไซเซอร์

สำหรับการเพิ่มหรือลดชนิดเครื่องดนตรีสามารถกระทำได้โดยการกดปุ่ม  (ดูรูปที่ 5.4 ประกอบ) ในการเพิ่มเครื่องดนตรีจะเป็นการอ่านค่าพารามิเตอร์จากเท็กไฟล์ โดยไฟล์ดังกล่าวต้องมีค่าพารามิเตอร์ต่างๆ ดังนี้

แถวที่ 1 จะเป็นชื่อเครื่องดนตรี โดยอยู่ในเครื่องหมาย " " เช่น "Thai Flute:Piang-aw"

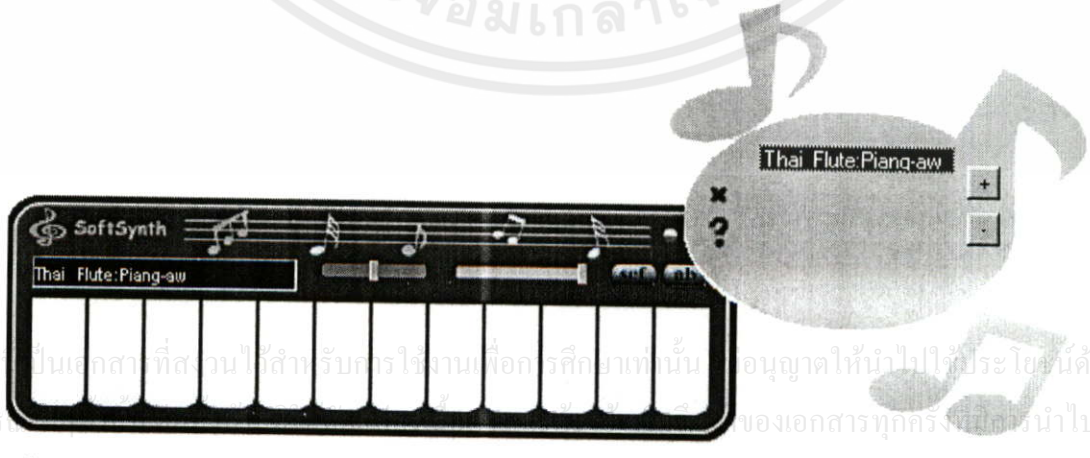
แถวที่ 2 จะเป็นค่าความถี่ทั้งหมด 8 ค่าของโน้ตโด

แถวที่ 3 จะเป็นค่าแอมปลิจูดทั้งหมด 8 ค่าของโน้ตโด

แถวที่ 4 จะเป็นค่าความถี่ทั้งหมด 8 ค่าของโน้ตเร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แถวที่ 5 จะเป็นค่าแอมปลิจูดทั้งหมด 8 ค่าของโน้ตเร
แถวที่ 6 จะเป็นค่าความถี่ทั้งหมด 8 ค่าของโน้ตมี
แถวที่ 7 จะเป็นค่าแอมปลิจูดทั้งหมด 8 ค่าของโน้ตมี
แถวที่ 8 จะเป็นค่าความถี่ทั้งหมด 8 ค่าของโน้ตฟา
แถวที่ 9 จะเป็นค่าแอมปลิจูดทั้งหมด 8 ค่าของโน้ตฟา
แถวที่ 10 จะเป็นค่าความถี่ทั้งหมด 8 ค่าของโน้ตซอล
แถวที่ 11 จะเป็นค่าแอมปลิจูดทั้งหมด 8 ค่าของโน้ตซอล
แถวที่ 12 จะเป็นค่าความถี่ทั้งหมด 8 ค่าของโน้ตลา
แถวที่ 13 จะเป็นค่าแอมปลิจูดทั้งหมด 8 ค่าของโน้ตลา
แถวที่ 14 จะเป็นค่าความถี่ทั้งหมด 8 ค่าของโน้ตที
แถวที่ 15 จะเป็นค่าแอมปลิจูดทั้งหมด 8 ค่าของโน้ตที
แถวที่ 16 จะเป็นค่าความถี่ทั้งหมด 8 ค่าของโน้ตโดสูง
แถวที่ 17 จะเป็นค่าแอมปลิจูดทั้งหมด 8 ค่าของโน้ตโดสูง
แถวที่ 18 จะเป็นค่าความถี่ทั้งหมด 8 ค่าของโน้ตเรสูง
แถวที่ 19 จะเป็นค่าแอมปลิจูดทั้งหมด 8 ค่าของโน้ตเรสูง
แถวที่ 20 จะเป็นค่าความถี่ทั้งหมด 8 ค่าของโน้ตมีสูง
แถวที่ 21 จะเป็นค่าแอมปลิจูดทั้งหมด 8 ค่าของโน้ตสูง
แถวที่ 22 จะเป็นค่าความถี่ทั้งหมด 8 ค่าของโน้ตฟาสูง
แถวที่ 23 จะเป็นค่าแอมปลิจูดทั้งหมด 8 ค่าของโน้ตฟาสูง
แถวที่ 24 จะเป็นค่าความถี่ทั้งหมด 8 ค่าของโน้ตซอลสูง
แถวที่ 25 จะเป็นค่าแอมปลิจูดทั้งหมด 8 ค่าของโน้ตซอลสูง
แถวที่ 26 จะเป็นค่าสัมประสิทธิ์ตัวกรองสัญญาณทั้งหมด 14 ค่า



รูปที่ 5.4 แสดงอินเทอร์เฟซของซอฟต์แวร์ซินทีไซเซอร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น ห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แอพพลิเคชันนี้นอกจากสามารถสังเคราะห์เสียงขลุ่ยไทยได้แล้วยังสามารถทำการโหลดไฟล์เพื่อใช้ในการเล่นเครื่องดนตรีประเภทอื่น ซึ่งไฟล์เหล่านี้ จะต้องจัดเก็บค่าพารามิเตอร์ต่าง ๆ (ค่าความถี่ ค่าแอมปลิจูด และค่าสัมประสิทธิ์ตัวกรองสัญญาณ) ให้อยู่ในฟอร์แมตที่กำหนด โดยค่าพารามิเตอร์ต่าง ๆ จะมีค่าแตกต่างกันไปตามลักษณะเสียงของเครื่องดนตรีที่ได้จากการวิเคราะห์เชิงสเปกตรัมและการวิเคราะห์สัญญาณรบกวน นอกจากนั้นแล้วแอพพลิเคชันยังสามารถทำการปรับเพิ่มหรือลดระดับความดังของเสียง และปรับความดังของลำโพงซ้าย/ขวาได้อีกด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6
สรุปผลการวิจัยและข้อเสนอแนะ

6.1 สรุปผลการวิจัย

งานวิจัยนี้ได้นำเสนอแนวทางการสังเคราะห์เสียงสำหรับเสียงเครื่องดนตรีขลุ่ยไทย โดยได้ประยุกต์ใช้โมเดลการสังเคราะห์เสียง Sinusoidal Plus Noise ซึ่งวิธีการดังกล่าวได้โมเดลสัญญาณเสียงออกเป็น 2 ส่วนคือ ส่วนของซินนูซอยด์ที่อาศัยทฤษฎีของฟูเรียร์ และส่วนของสัญญาณรบกวนที่โมเดลได้ด้วยคุณสมบัติทางสถิติ โดยในขั้นตอนการวิเคราะห์สเปกตรัมเสียงขลุ่ยไทย ผู้วิจัยได้นำเสนอแนวทางเพิ่มเติมในขบวนการวิเคราะห์หาค่าพารามิเตอร์ต่าง ๆ เพื่อให้ได้ค่าที่ใกล้เคียงความจริงที่สุด จากผลการทดลองในบทที่ 3 พบว่า แบล็คแมนวินโดว์เมื่อนำมาคูณกับสัญญาณเสียงก่อนทำการวิเคราะห์เชิงสเปกตรัม ให้ค่าความถูกต้องของค่าความถี่และค่าแอมปลิจูดมากกว่าวินโดว์ชนิดอื่น นอกจากนี้วินโดว์ชนิดนี้ยังให้สเปกตรัมที่แยกส่วนของสัญญาณไซน์ได้ชัดเจน อันเป็นประโยชน์ในการวิเคราะห์ส่วนของสัญญาณรบกวน สำหรับกรณีการเพิ่มจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์เข้ากับสัญญาณเสียงนั้นควรเพิ่มโดยให้ได้สัญญาณที่มีขนาดเป็นจำนวนเท่าของสัญญาณเดิม จึงจะทำให้ค่าความถี่และค่าแอมปลิจูดมีความถูกต้องมากขึ้น โดยแนวทางดังกล่าวนี้ได้นำมาใช้ในการวิเคราะห์ในส่วนของซินนูซอยด์ ซึ่งการวิเคราะห์ได้แบ่งเป็น 2 ส่วนคือส่วนสัญญาณเสียงที่มีการเปลี่ยนแปลง และส่วนสัญญาณเสียงที่อยู่ในภาวะเสถียร จากการวิเคราะห์ส่วนของสัญญาณเสียงที่มีการเปลี่ยนแปลงด้วยวิธีการสังเคราะห์สัญญาณเสียงพบว่า ลักษณะสเปกตรัมมีความใกล้เคียงกันกับสเปกตรัมของสัญญาณเสียงที่อยู่ในภาวะเสถียร ดังนั้นในการคำนวณหาค่าพารามิเตอร์เพื่อการสร้างสัญญาณไซน์นี้ จึงได้ใช้ส่วนสัญญาณเสียงที่อยู่ในภาวะเสถียรในการวิเคราะห์เป็นหลัก และในการวิเคราะห์ในส่วนของสัญญาณรบกวนพบว่า ค่าสัมประสิทธิ์ตัวกรองสัญญาณที่ได้จากการใช้เออาร์โมเดลจะให้ลักษณะสัญญาณใกล้เคียงกับสัญญาณรบกวนต้นแบบมากกว่าเอ็มเอโมเดลที่ใช้ในงานวิจัยด้านการสังเคราะห์เสียงแบบ Sinusoidal Plus Noise เดิม

สำหรับในกระบวนการสังเคราะห์เสียงเพื่อให้ได้เสียงที่มีคุณภาพและมีความซับซ้อนในการคำนวณน้อย ผู้วิจัยได้พยายามลดจำนวนสัมประสิทธิ์ตัวกรองสัญญาณลงจนถึง 15 จำนวนในการสร้างส่วนของสัญญาณรบกวนเพื่อนำมารวมกับส่วนของสัญญาณไซน์ที่มีองค์ประกอบของค่าความถี่ 10 จำนวน และจากผลการประเมินคุณภาพเสียงจากผู้เชี่ยวชาญทางด้านดนตรีพบว่าเสียงสังเคราะห์ที่ได้ก็ยังคงให้เสียงที่ใกล้เคียงเสียงต้นแบบ

เอกสารนี้ใช้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.2 บทวิเคราะห์

จากการศึกษาแนวทางการสังเคราะห์เสียง Sinusoidal Plus Noise เพื่อประยุกต์ใช้กับเสียงของเครื่องดนตรีขลุ่ยไทย พบว่าแนวทางนี้มีข้อดีเหนือแนวทางการสังเคราะห์เสียงประเภทอื่นๆ ในหลายประเด็น โดยเฉพาะในเรื่องความยืดหยุ่นในการปรับแต่งเสียง

6.2.1 ความยืดหยุ่นในการสร้างเสียงแปลกใหม่

แนวทางการสังเคราะห์เสียงดังกล่าวมีความยืดหยุ่นในการปรับสร้างเสียงแปลกใหม่ หรือแม้แต่การทำออดิโอมอร์ฟฟิงก็สามารถกระทำได้ง่ายกว่าแนวทางอื่น โดยทั่วไปแล้วซินทีไซเซอร์จะใช้วิธีการแซมปลิงหรือที่เรียกว่าเวปเทเบิลไม่สามารถสังเคราะห์เสียงที่แปลกออกไปได้ แต่สำหรับแนวทางการสังเคราะห์เสียง Sinusoidal Plus Noise สามารถกระทำได้โดยการเปลี่ยนแปลงองค์ประกอบเชิงความถี่ ในที่นี้ได้เสนอวิธีการดัดแปลงเสียงไว้ดังนี้

- ปรับเปลี่ยนค่าแอมปลิจูดในบางค่าความถี่หรือทุกค่าความถี่
- ใช้ค่าความถี่เพียงบางค่าร่วมกับการปรับเปลี่ยนค่าแอมปลิจูดเดิม
- ใช้ค่าความถี่เพียงบางค่าแต่ยังคงค่าแอมปลิจูดเดิม

นอกจากนี้อาจสร้างเสียงสังเคราะห์โดยการผสมผสานทางด้านความถี่ของเครื่องดนตรีสองชนิดใด ๆ ซึ่งสามารถทำได้ดังนี้

- ใช้ค่าเฉลี่ยค่าความถี่ของเสียงดนตรีทั้งสองพร้อมทั้งค่าแอมปลิจูดของเสียงใดเสียงหนึ่ง
- ใช้ค่าเฉลี่ยของค่าแอมปลิจูดร่วมกับค่าความถี่ของเสียงดนตรีชนิดใดชนิดหนึ่งเป็นหลัก
- ใช้ค่าเฉลี่ยของค่าแอมปลิจูดและค่าความถี่เพื่อใช้ในการสังเคราะห์เสียง

จากการทดลองพบว่า เสียงสังเคราะห์ที่เกิดจากการผสมผสานระหว่างขลุ่ยเพียงออกับขลุ่ยหลิบให้เสียงสังเคราะห์ที่ยังคงลักษณะเสียงดนตรีซึ่งไม่ใช่เสียงระดับเสียงที่มีในขลุ่ยเพียงออหรือขลุ่ยหลิบ โดยเสียงที่ได้น่าจะเป็นเสียงประเภทครึ่งเสียงมากกว่า

ข้อดีอีกประการหนึ่งของการสังเคราะห์เสียงด้วยวิธีการดังกล่าว คือสามารถสร้างเสียงสังเคราะห์ที่มีระดับเสียงสูงต่ำของตัวโน้ตเท่ากับเสียงดนตรีสากล ซึ่งกระทำได้โดยปรับค่าความถี่ให้มีค่าเท่ากับความถี่มาตราฐาน ตลอดจนสามารถสร้างเสียงสังเคราะห์ที่เป็นครึ่งเสียง ทำให้มีคุณสมบัติพิเศษในการเล่นเสียงดนตรีขลุ่ยไทยร่วมกับดนตรีสากลได้

6.2.2 ความเป็นธรรมชาติของเสียง

ให้เสียงที่มีความเป็นธรรมชาติของเสียง เนื่องจากโมเดลของสัญญาณรบกวนจะสร้างสัญญาณที่มีการเปลี่ยนแปลงไปตามคุณสมบัติทางสถิติที่ได้จากการวิเคราะห์ ทำให้สัญญาณเสียงที่สังเคราะห์ในแต่ละครั้งไม่เหมือนกัน หากเปรียบเทียบกับขลุ่ย ก็เปรียบได้กับเป็นคนละคนเป่าหรือขลุ่ยคนละเลา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.2.3 การใช้งานได้อย่างกว้างขวาง

แนวทางดังกล่าวมีรูปแบบโมเดลที่สามารถประยุกต์ใช้ได้กับเครื่องดนตรีอื่น ๆ ได้ เพราะเสียงดนตรีหรือว่าเสียงพูดจะมีองค์ประกอบของสัญญาณไซน์ที่แตกต่างกัน เพียงแต่เปลี่ยนองค์ประกอบของสัญญาณก็จะสามารถสังเคราะห์เสียงที่ต้องการได้

6.2.4 จำนวนพารามิเตอร์ที่ใช้

จำนวนพารามิเตอร์ที่ใช้ในการสังเคราะห์เสียงน้อย มีเพียงค่าความถี่ ค่าแอมปลิจูดและค่าสัมประสิทธิ์ตัวกรองสัญญาณ จากการทดลองในบทที่ 4 ในการสังเคราะห์เสียงด้วยแนวทางดังกล่าวใช้จำนวนพารามิเตอร์เพียง 1 ใน 100 ของจำนวนพารามิเตอร์ทั้งหมดที่แนวทางการสังเคราะห์เสียงแบบ Sinusoidal ใช้ ซึ่งก็ให้เสียงที่มีคุณภาพใกล้เคียงกัน

6.3 ข้อเสนอแนะและแนวทางในการทำวิจัยต่อ

จากการวิเคราะห์และสังเคราะห์เสียงขลุ่ยไทยด้วยวิธีการสังเคราะห์เสียง Sinusoidal Plus Noise สามารถพัฒนาเป็นซอฟต์แวร์ซินทีไซเซอร์ที่ใช้สังเคราะห์เสียงแบบเรียลไทม์ ได้ นอกจากนี้โมเดลดังกล่าวสามารถนำมาพัฒนาเป็นซินทีไซเซอร์ที่สามารถสร้างเสียงแปลกใหม่ หรือแม้แต่การทำออดิโอมอร์ฟฟิงได้

เนื่องจากว่าโมเดลดังกล่าวสามารถนำมาประยุกต์ใช้ในการสังเคราะห์เสียงเครื่องดนตรีได้หลายชนิด จึงน่าจะมีการนำวิธีการดังกล่าวมาใช้กับเครื่องดนตรีไทยในวงมโหรีเพื่อการวิเคราะห์ค่าพารามิเตอร์ที่จำเป็นในการสังเคราะห์ และนำมาสร้างซอฟต์แวร์ซินทีไซเซอร์ที่สามารถเล่นเพลงมิดีไทยเดิมได้ โดยข้อมูลที่จัดเก็บมีเพียงค่าพารามิเตอร์ของเครื่องดนตรีแต่ละชนิดและโมเดลในการสร้างเสียงสังเคราะห์ก็ใช้ร่วมกัน ทำให้ขนาดของซอฟต์แวร์มีขนาดเล็กสะดวกในการเผยแพร่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] ประเสริฐ คชายุทธเดช. *โน้ตดนตรีไทย เล่ม 1 ขลุ่ย*. หน้า 6.

[2] เสริม หนูจ้อย. *ทฤษฎีประกอบการเรียนวิชาศิลปศึกษาแขนงวิชาดนตรีศึกษา*. กรุงเทพมหานคร: โรงพิมพ์อักษรสยามการพิมพ์. หน้า 2-4.

[3] JanBen, H., *Synthesis Methods*, [Online]. Available:http://www.neuroinformatik.ruhr-uni-bochum.de/ini/PEOPLE/heja/sy-prog/node3.html

[4] Nomer, D., *Musical Physical Modeling*, 1998. [Online]. Available: http://web800.com/music music/vl/physmodl.htm.

[5] Peevers, A.W., *A Real Time 3D Signal Analysis/Synthesis Tool Based on the Short Time Fourier Transfor*, 1999. [Online]. Available: http://cnmat.CNMAT.Berkeley.EDU/ ~alan/MS-html/MSv2_ToC.html.

[6] Press, H. W. et. al., *Numerical Recipes in C : The Art of Scientific Computing*, 2nd ed., New York, USA: Cambridge University Press, 1992, pp. 288-290.

[7] Rodet, X., *Musical Sound Signal Analysis/ Synthesis: Sinusoidal+Residual and Elementary Waveform Models*, Paris, France, 1997.

[8] Serra, X. and J. Smith., *Spectral Modeling Synthesis: A Sound Analysis/Synthesis System based on a Deterministic plus Stochastic Decomposition, Computer Music Journal* 14 (4):12—24, 1990.

[9] Serra, X., *Current Perspectives in the Digital Synthesis of Musical Sounds*, Pompeau Fabra University, 1997. [Online]. Available:http://www.iua.upf.es/~xserra/articles/current.

[10] Serra, X., *Integrating Complementary Spectral Models in the Design of a Musical Synthesizer, The Proceeding of the International Computer Music Conference* , 1997.

[11] Serra, X., *Musical Sound Modeling With Sinusoidal Plus Noise, Musical Signal Processing*. Swets& Zeitlinger Publishers,1997.

[12] Smith, J.O., *Viewpoints on the History of Digital Synthesis., Proceedings of the 1991 International Computer Music Conference*, Montreal, San Fracisco, 1-10 Oct. 1991.

[13] Steiglit, K., *A Digital Signal Processing Primer with Applications to Digital Audio and Computer Music,* Menlo Park, CA : Addison-Wesley, 1996.

[14] Tero Tolonen, *Model-Based Analysis and Resynthesis of Acoustic Guitar Tones*, Master's thesis. Report no. 46, Helsinki University of Technology, Department of Electrical and

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Communications Engineering, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland,pp 33-35 , Jan. 1998. [Online] Available: http://bekesy.hut.fi/~ttolonen/thesis.html

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

การเปรียบเทียบชนิดวินโดว์ในเรื่องความถูกต้องของค่าสเปกตรัม สามารถพิจารณาจากความคลาดเคลื่อนเฉลี่ยของสัญญาณเสียงที่สร้างขึ้นจากการใช้วินโดว์แต่ละชนิด แต่เนื่องจากว่าการคำนวณหาสเปกตรัมของสัญญาณเสียงในงานวิจัยนี้ได้ใช้วิธีการดีสครีตฟูเรียร์ทรานฟอร์ม ซึ่งเป็นการประมาณค่าความถี่ในแต่ละจุดข้อมูลแซมเปิลด้วยระยะห่างที่เท่ากัน จึงมีความเป็นไปได้ที่ว่า ค่าความถี่ที่คำนวณได้ จะมีความคลาดเคลื่อนจากความเป็นจริงได้ ซึ่งจะมากหรือน้อยนั้นขึ้นอยู่กับตำแหน่งจุดข้อมูลที่ทำการประมาณค่าอยู่ ณ ตำแหน่งที่ใกล้ค่าความถี่จริงมากน้อยเพียงไร ซึ่งค่าความคลาดเคลื่อนของค่าความถี่ต่าง ๆ ที่มิได้อยู่ ณ ตำแหน่งที่มีการประมาณค่าไว้ จะมีค่าเพิ่มขึ้นเรื่อย ๆ เมื่อค่าความถี่จริงอยู่ ณ ตำแหน่งที่ห่างจากจุดที่มีการประมาณค่าไว้มากขึ้น จนถึง ณ ตำแหน่งที่ค่าความถี่จริงอยู่ตรงจุดกึ่งกลางระหว่าง 2 จุดที่มีการประมาณค่าไว้ จะมีค่าความคลาดเคลื่อนสูงสุด และจะลดลงเรื่อย ๆ จนมีค่าเป็นศูนย์เมื่อค่าความถี่จริงอยู่ ณ ตำแหน่งที่ประมาณค่าไว้พอดี ซึ่งค่าความคลาดเคลื่อนจะเป็นลักษณะอย่างนี้ตลอดทุกช่วงความถี่ ดังนั้นการหาความคลาดเคลื่อนเฉลี่ยไม่จำเป็นต้องหาตลอดทุกย่านความถี่ หากเพียงแต่เฉลี่ยค่าความคลาดเคลื่อนของความถี่ที่มีค่าอยู่ในช่วงระหว่างจุดข้อมูลที่มีการประมาณค่า (BIN) ก็สามารถนำมาเป็นตัวแทนค่าความคลาดเคลื่อนเฉลี่ยได้ ดังแสดงตามการทดลองข้างล่างนี้

1. การทดลอง

การทดลองนี้จะสร้างสัญญาณเสียงที่เกิดจากสัญญาณไซน์ 2 ความถี่รวมกันซึ่งมีขนาด 50000 แซมเปิล จากนั้นทำการเพิ่มค่าความถี่ทั้งสองให้มีค่าอยู่ในช่วง 1 BIN ซึ่งจะได้ค่าความถี่อยู่ในช่วง f และ f+ 1.0e-005 (โดยที่ f คือค่าความถี่เดิม) เพื่อพิจารณาค่าความคลาดเคลื่อนเฉลี่ยในช่วงความถี่นี้ และเปรียบเทียบค่าที่ได้กับช่วงความถี่อื่นโดยสร้างสัญญาณเสียงโดยให้ความถี่ทั้งสองเลื่อนไปจากค่าความถี่เดิมเท่ากับ +2.0e-005 และ +4.0e-005 แล้วทำการปรับค่าความถี่ทั้งสองตามที่ได้กล่าวข้างต้น ซึ่งในรายละเอียดแสดงดังต่อไปนี้

1) สร้างสัญญาณ $s3$ ซึ่งมีทั้งหมด 3 ช่วงความถี่

$s3_1 = s1_1 + s2_1$; $s1_1 = \cos(2\pi*(0:49999)*f_1)$ $s2_1 = (0.08)*\cos(2\pi*(0:49999)*f_2)$

โดยที่ f_1 อยู่ในช่วง 0.0102 ถึง 0.010218 โดยจะเพิ่มค่าทีละ 1/500000

f_2 อยู่ในช่วง 0.0203 ถึง 0.020318 โดยจะเพิ่มค่าทีละ 1/500000

$s3_2 = s1_2 + s2_2$; $s1_2 = \cos(2\pi*(0:49999)*f_1)$ $s2_2 = (0.08)*\cos(2\pi*(0:49999)*f_2)$ โดยที่ f_1 อยู่ในช่วง 0.01022 ถึง 0.010238 โดยจะเพิ่มค่าทีละ 1/500000

f_2 อยู่ในช่วง 0.02032 ถึง 0.020338 โดยจะเพิ่มค่าทีละ 1/500000

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเอกสารทุกครั้งที่มีการนำไปใช้

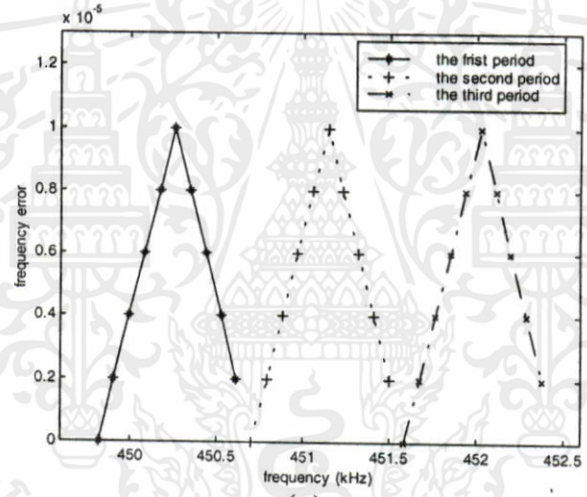
$s3_2 = s1_2 + s2_2$; $s1_2 = \cos(2\pi*(0:49999)*f_1)$ $s2_2 = (0.08)*\cos(2\pi*(0:49999)*f_2)$ โดยที่ f_1 อยู่ในช่วง 0.01024 ถึง 0.010258 โดยจะเพิ่มค่าทีละ 1/500000

f_2 อยู่ในช่วง 0.02034 ถึง 0.020358 โดยจะเพิ่มค่าทีละ 1/500000

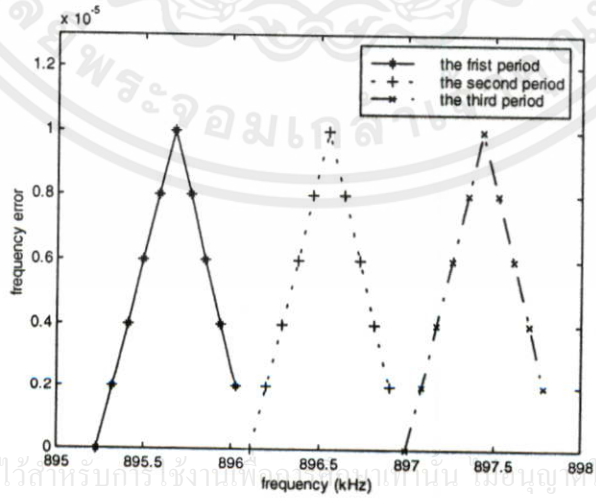
2) คำนวณหาฟูเรียร์ทรานฟอร์มของสัญญาณ ใช้ในการวิเคราะห์สเปกตรัมของสัญญาณ เพื่อคำนวณค่าความคลาดเคลื่อนของความถี่และแอมปลิจูดที่ได้ในแต่ละช่วงความถี่

2 ผลการทดลอง

การคำนวณหาค่าความคลาดเคลื่อนของค่าความถี่และค่าแอมปลิจูดในแต่ละช่วงความถี่ พบว่าค่าความคลาดเคลื่อนจะมีลักษณะเพิ่มมากขึ้นเรื่อย ๆ จนมีค่ามากที่สุดเมื่อค่าความถี่จริงของสัญญาณที่สร้างขึ้นมีค่าอยู่ ณ กึ่งกลางระหว่างจุดข้อมูลที่ทำการประมาณค่าไว้และจะลดลงไปเรื่อย ๆ แสดงดังรูปที่ ก.1 และ ก.2 ตามลำดับ



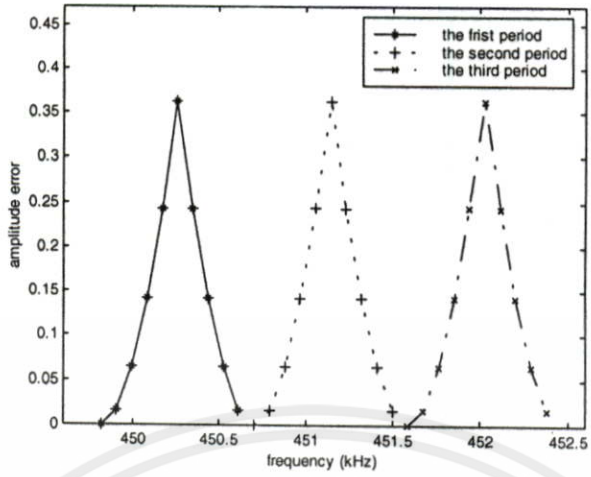
(ก)



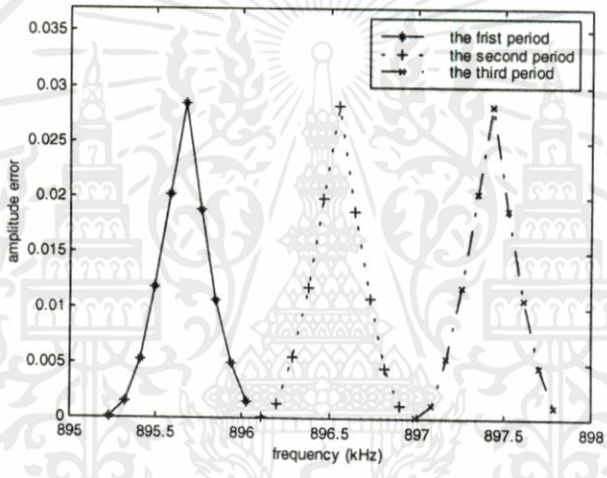
(ข)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ก.1 แสดงค่าความคลาดเคลื่อนของค่าความถี่ในแต่ละช่วงความถี่
ก.ความคลาดเคลื่อนของค่าความถี่ที่ฟังก์ชันไซน์ลำดับที่หนึ่ง
ข.ความคลาดเคลื่อนของค่าความถี่ที่ฟังก์ชันไซน์ลำดับที่สอง



(ก)



(ข)

รูปที่ ก.2 แสดงค่าความคลาดเคลื่อนของค่าแอมปลิจูดในแต่ละช่วงความถี่
ก.ความคลาดเคลื่อนของค่าแอมปลิจูดที่ฟังก์ชันไซน์ลำดับที่หนึ่ง
ข.ความคลาดเคลื่อนของค่าแอมปลิจูดที่ฟังก์ชันไซน์ลำดับที่สอง

เมื่อทำการเฉลี่ยค่าความคลาดเคลื่อนในแต่ละช่วงความถี่โดยพิจารณาจากการคำนวณค่ารูทมีนสแคว์เออร์เรอร์ (RMSE โดยรายละเอียดของสมการแสดงในภาคผนวก ข) พบว่า ค่าเฉลี่ยที่ได้มีค่าเท่ากันแสดงดังตารางที่ ก.1

3 สรุปผลการทดลอง

การคำนวณหาค่าความคลาดเคลื่อนเฉลี่ยไม่ว่าจะเป็นช่วงความถี่ใด ๆ จะให้ค่าที่เท่ากัน เนื่องมาจากว่าการประมาณค่าความถี่และแอมปลิจูดด้วยวิธีการดีสครีตฟูเรียร์ทรานฟอร์ม จะทำให้ได้ค่าความคลาดเคลื่อนที่เท่ากันในทุกช่วงความถี่ หากนำมาสร้างกราฟจะได้กราฟที่มีลักษณะเหมือนกันทุกช่วงความถี่ ดังแสดงในรูปที่ ก.1 และ ก.2 ดังนั้นการคำนวณค่าความคลาดเคลื่อนเฉลี่ย จึงไม่จำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารนี้ทุกครั้งที่มีการนำไปใช้

เป็นต้องคำนวณหาค่าเฉลี่ยตลอดทุกช่วงความถี่ เพียงแต่ทำการคำนวณหาค่าความคลาดเคลื่อนของความถี่ที่มีค่าอยู่ ณ ตำแหน่งระหว่างสองจุดข้อมูลใด ๆ (ภายใน BIN ใด ๆ) มาเฉลี่ยเพื่อเป็นตัวแทนของความคลาดเคลื่อนของค่าความถี่ทั้งหมดได้

ตารางที่ ก.1 แสดงค่าความคลาดเคลื่อนเฉลี่ยในแต่ละช่วงความถี่

ค่าความคลาดเคลื่อนเฉลี่ย	ช่วงความถี่แรก	ช่วงความถี่ที่สอง	ช่วงความถี่ที่สาม
ค่าความถี่ในสัญญาณไซน์แรก	1.8439e-006	1.8439e-006	1.8439e-006
ค่าความถี่ในสัญญาณไซน์ที่สอง	1.8439e-006	1.8439e-006	1.8439e-006
ค่าแอมปลิจูดในสัญญาณไซน์แรก	0.0547	0.0547	0.0547
ค่าแอมปลิจูดในสัญญาณไซน์ที่สอง	0.0043	0.0043	0.0043

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข

1.Error ค่าความคลาดเคลื่อนหาได้จาก

$$E = \left| I_i - J_i \right| \tag{ข.1}$$

โดยที่ $I_i - J_i$ คือค่าความคลาดเคลื่อน

2.Mean Square Error เป็นการคำนวณค่าความคลาดเคลื่อนเฉลี่ย โดยมีสมการดังนี้

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(I_i - J_i\right)^2 \tag{ข.2}$$

3.Root Mean Square Error เป็นการคำนวณค่าความคลาดเคลื่อนเฉลี่ยอย่างละเอียด โดยมีสมการดังนี้

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(I_i - J_i\right)^2} \tag{ข.3}$$

4.Curve Fitting

เนื่องจากเราต้องการให้ค่าความถี่ฮาร์โมนิกต่าง ๆ มีค่าเป็นจำนวนเท่าของความถี่พื้นฐานดังสมการนี้

$$fn_i = ax_i \tag{ข.4}$$

โดยที่ fn_i คือค่าความถี่

x_i คือค่าตั้งแต่ 1,2,3 ... (ลำดับฮาร์โมนิกต่าง ๆ)

ซึ่งในการทำ curve fitting เพื่อให้ได้ค่าความถี่ลำดับต่าง ๆ ตามสมการ ข.4 ควรให้ค่าเออร์เรอร์น้อยที่สุด ในที่นี้ก็คือผลรวมของผลต่างระหว่างค่าความถี่เดิมกับค่าความถี่ใหม่ทั้งหมดยกกำลังสองควรมีค่าน้อยที่สุด

$$e_i = fn_i - fo_i \tag{ข.5}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\sum_{i=1}^{N}(e_i^{\,2})=\sum_{i=1}^{N}(fn_i-fo_i)^2$$

$$=\sum_{i=1}^{N}(ax_i-fo_i)^2$$

$$=\sum_{i=1}^{N}(a^2x_i^{\,2}-2fo_iax_i+fo_i^{\,2})$$

$$=a^2\sum_{i=1}^{N}x_i^{\,2}-2a\sum_{i=1}^{N}(fo_ix_i)+\sum_{i=1}^{N}fo_i^{\,2} \qquad \text{(ข.6)}$$

จากนั้นเมื่อทำการหาค่าอนุพันธ์ลำดับที่ 1 ของ $\sum_{i=1}^{N}(e_i^{\,2})$ แล้วกำหนดให้มีค่าเท่ากับศูนย์

$$\frac{df(a)}{da}=0$$

$$2a\sum_{i=1}^{N}x_i^{\,2}-2\sum_{i=1}^{N}(fo_ix_i)=0 \qquad \text{(ข.7)}$$

จะได้ค่า a ที่เป็นค่ากลางดังสมการ

$$a=\frac{\sum_{i=1}^{N}(fo_ix_i)}{\sum_{i=1}^{N}x_i^{\,2}} \qquad \text{(ข.8)}$$

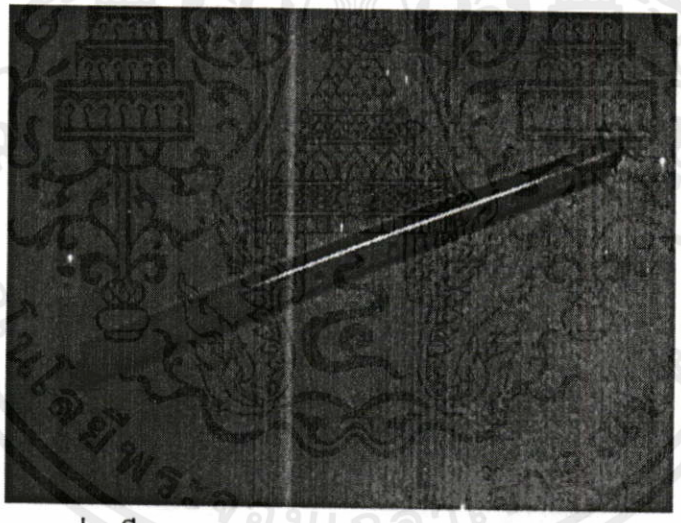
นำค่า a ไปแทนในสมการ ข.4 จะได้ค่าความถี่ต่าง ๆ ที่เป็นจำนวนเท่าของความถี่พื้นฐาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค
ขลุ่ยไทย

ขลุ่ยเป็นเครื่องดนตรีประเภทเครื่องเป่า ซึ่งใช้บรรเลงในวงดนตรีไทย ความไพเราะของขลุ่ยจะขึ้นอยู่กับวัสดุที่ใช้ ซึ่งหากเป็นไม้เนื้อแข็ง อย่างเช่น ไม้ไผ่ ไม้ชิงชัน แก่นไม้มะขาม เสียงขลุ่ยจะมีความนุ่มนวลมากกว่าขลุ่ยที่ทำจากเอสล่อนหรือพลาสติก โดยทั่วไปแล้วขลุ่ยที่นิยมใช้กันมากที่สุดคือขลุ่ยเพียงออเนื่องจากมีระดับเสียงปานกลาง ซึ่งหากแบ่งตามระดับช่วงเสียงแล้วขลุ่ยจำแนกได้ 3 ประเภท คือ [1] [2]

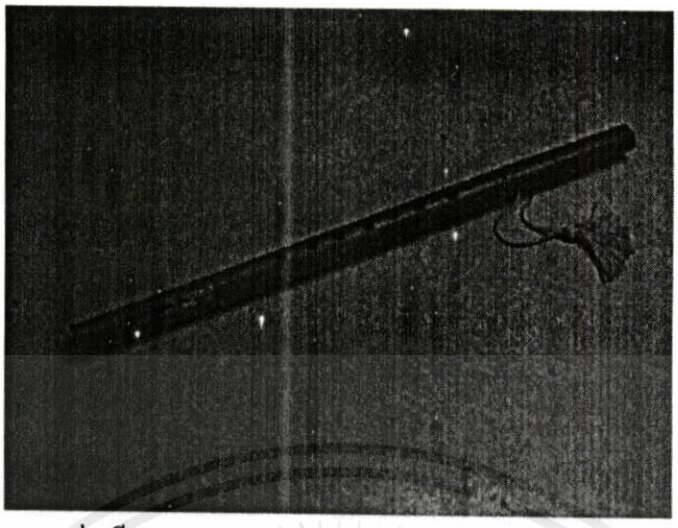
1.ขลุ่ยหลีบ เป็นขลุ่ยขนาดเล็ก ลักษณะจะยาวประมาณ 35 เซ็นติเมตร กว้างประมาณ 2 เซ็นติเมตร (ดูรูปที่ ค.1) มีเสียงสูงเล็กแหลม นิยมใช้บรรเลงในวงเครื่องสายเครื่องคู่ และวงเครื่องสายปี่ชวา



รูปที่ ค.1 แสดงลักษณะขลุ่ยหลีบ

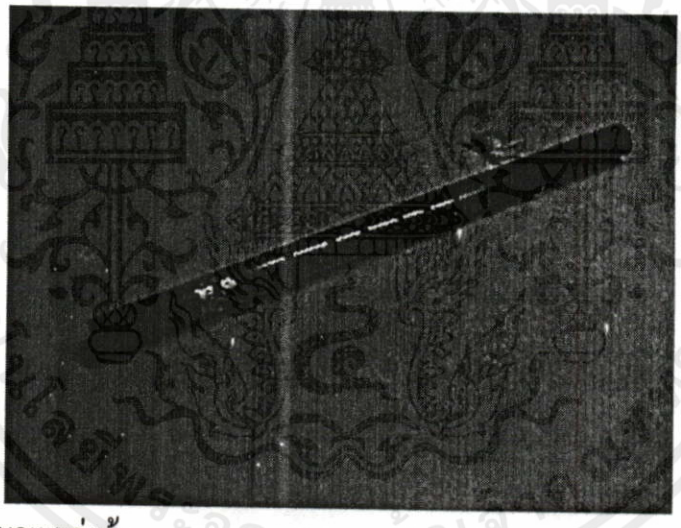
2.ขลุ่ยเพียงออ เป็นขลุ่ยขนาดกลาง ยาวประมาณ 45 เซ็นติเมตร กว้างประมาณ 2.5 เซ็นติเมตร (ดูรูปที่ ค.2) ให้เสียงปานกลาง นิยมใช้ในวงเครื่องสายวงเล็ก วงปี่พาทย์และวงมโหรี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ค.2 แสดงลักษณะขลุ่ยเพียงออ

3.ขลุ่ยอู้ เป็นขลุ่ยขนาดใหญ่ยาวประมาณ 60 เซ็นติเมตร กว้างประมาณ 4 เซ็นติเมตร (ดูรูปที่ ค.3) มีเสียงต่ำ นิยมใช้ในวงมโหรีเครื่องใหญ่

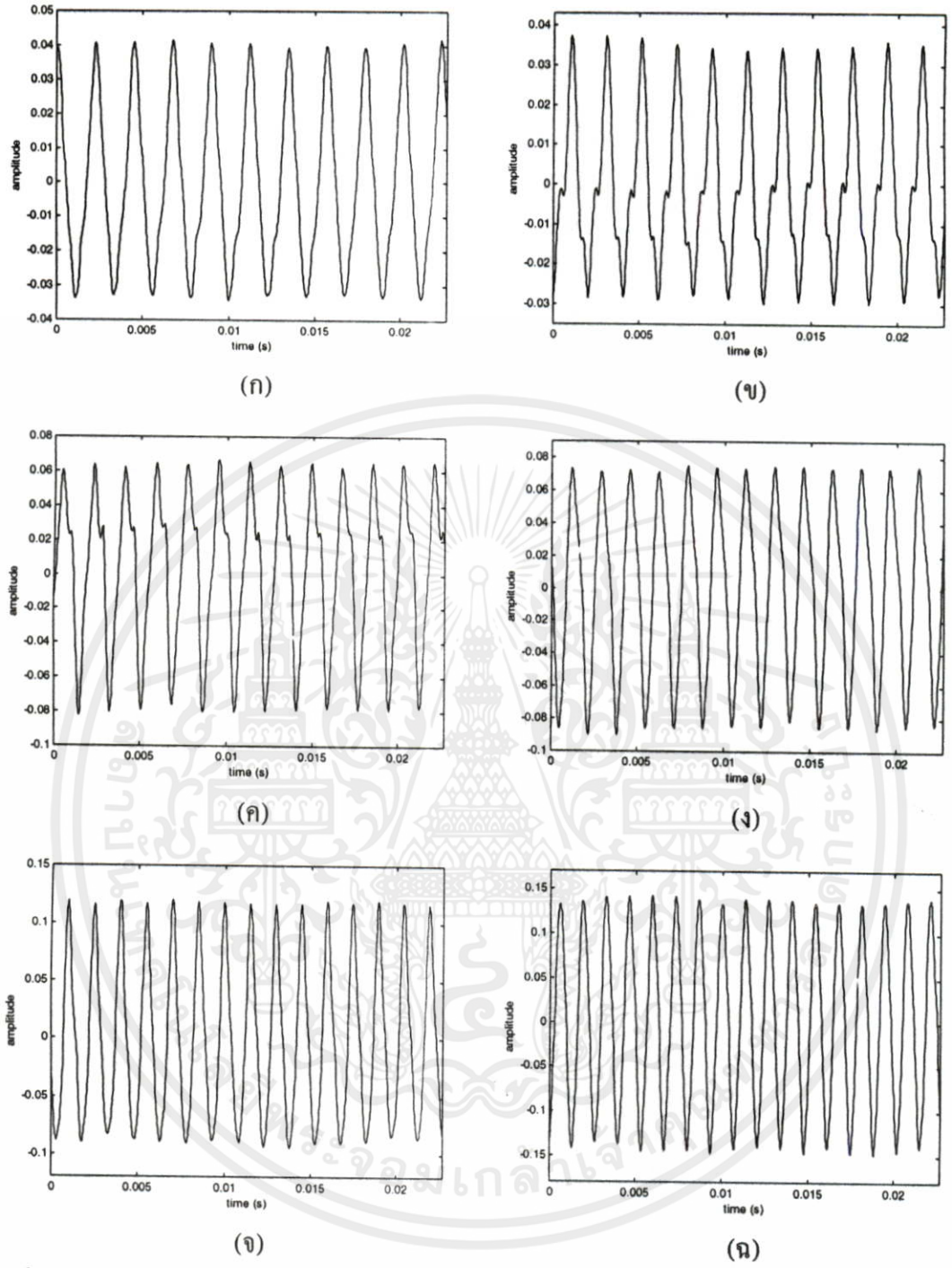


รูปที่ ค.3 แสดงลักษณะขลุ่ยอู้

ลักษณะสัญญาณเสียงขลุ่ยไทย

ลักษณะสัญญาณเสียงขลุ่ยไทยของขลุ่ยเพียงออที่ได้ทำการบันทึกไว้เพื่อเป็นข้อมูลในการวิเคราะห์เชิงสเปกตรัมแสดงดังรูปข้างล่างต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ค.4 แสดงลักษณะคลื่นสัญญาณเสียงขลุ่ยเพียงออตั้งแต่โน้ตโดจนถึงโน้ตซอลสูง

ก. โน้ตโด

ข. โน้ตเร

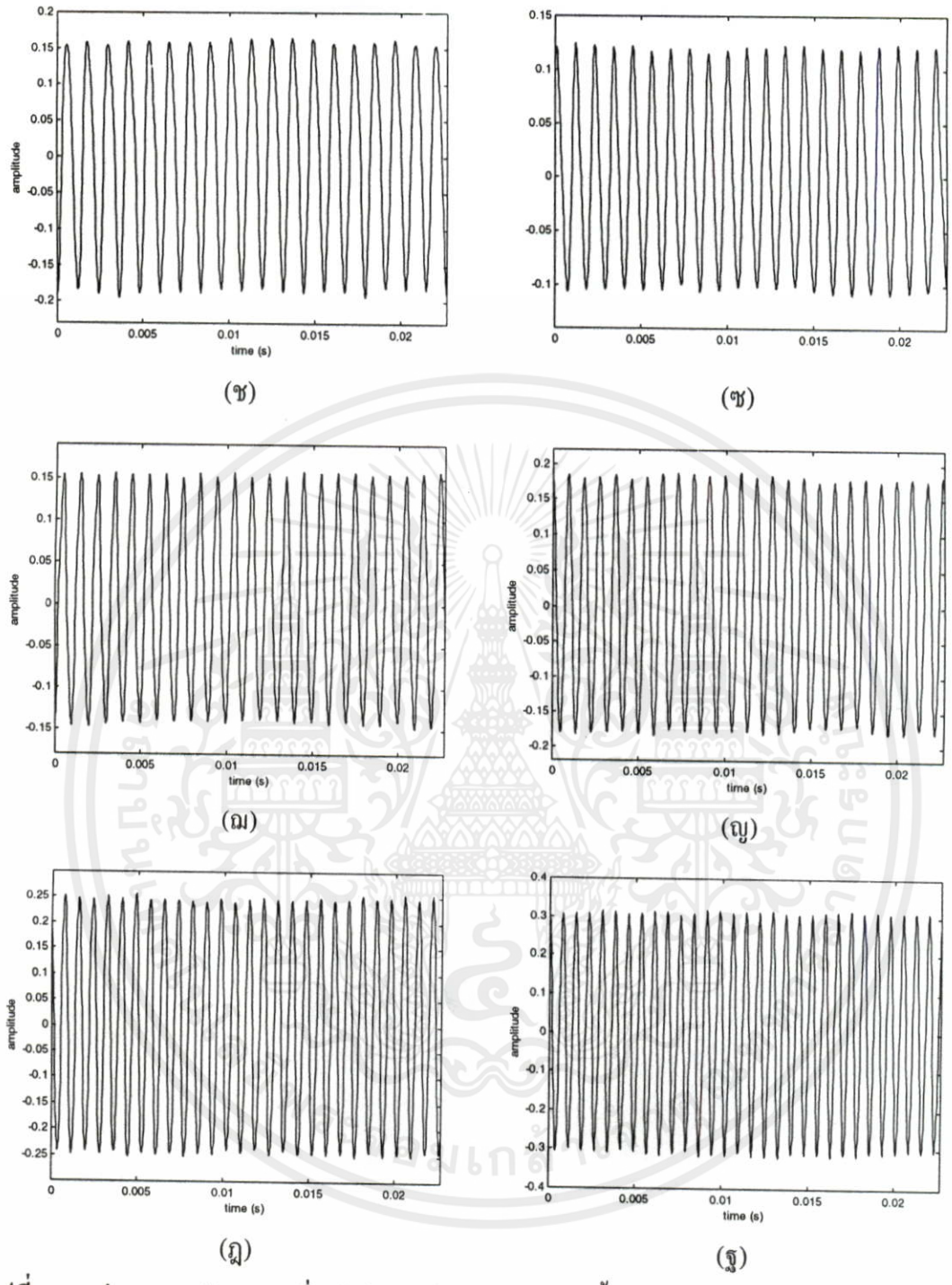
ค. โน้ตมี

ง. โน้ตฟา

จ. โน้ตซอล

ฉ. โน้ตลา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น ทั้งนี้ห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ ค.4 (ต่อ) แสดงลักษณะคลื่นสัญญาณเสียงขลุ่ยเพียงออตั้งแต่โน้ตโดจนถึงโน้ตซอลสูง

ช.โน้ตที

ซ.โน้ตโดสูง

ฌ.โน้ตเรสูง

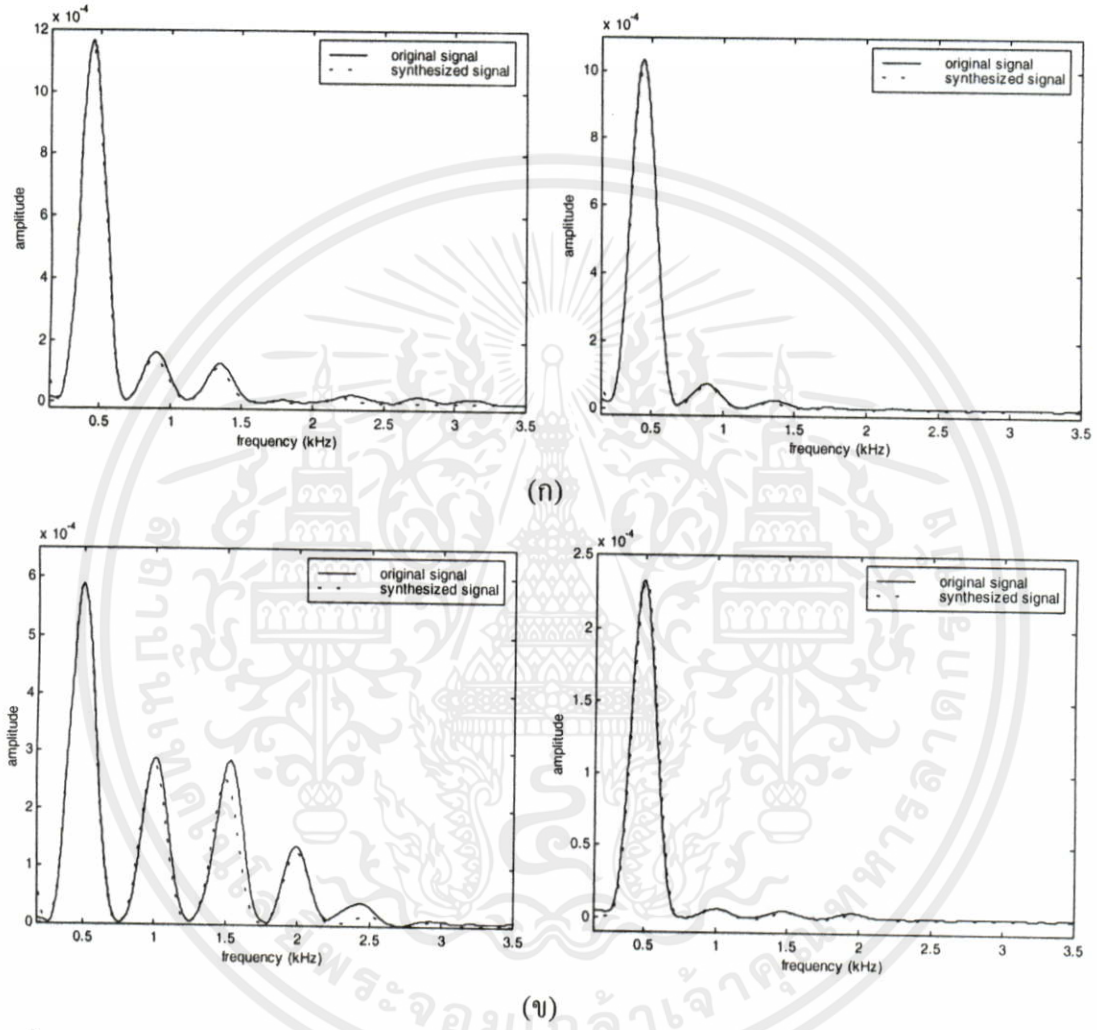
ญ.โน้ตมีสูง

ฎ.โน้ตฟาสูง

ฐ.โน้ตซอลสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ง
สเปกตรัมของโน้ตขลุ่ยไทย

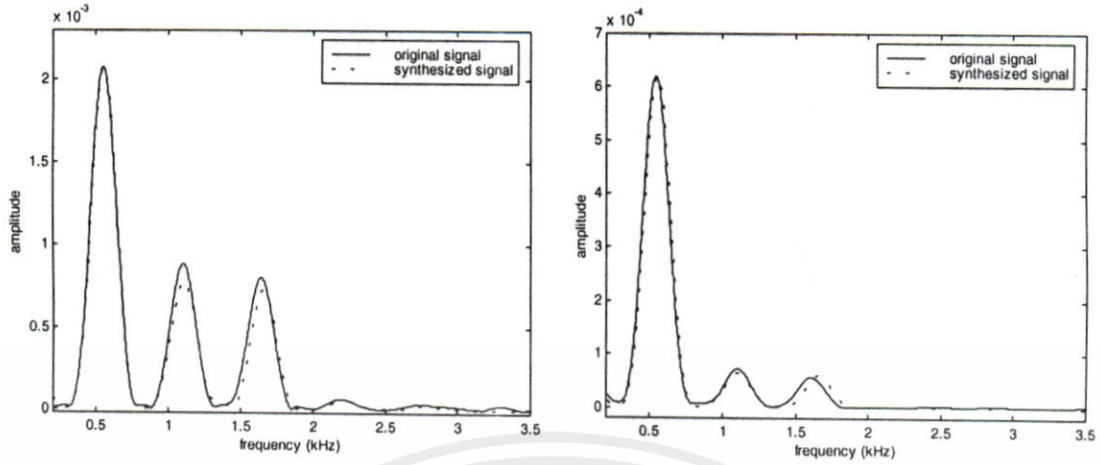


รูปที่ ง.1 แสดงการเปรียบเทียบสเปกตรัมส่วนที่เป็น attack (ซ้าย) และส่วนที่เป็น decay (ขวา) ของสัญญาณเสียงต้นแบบกับที่สังเคราะห์ขึ้นใหม่

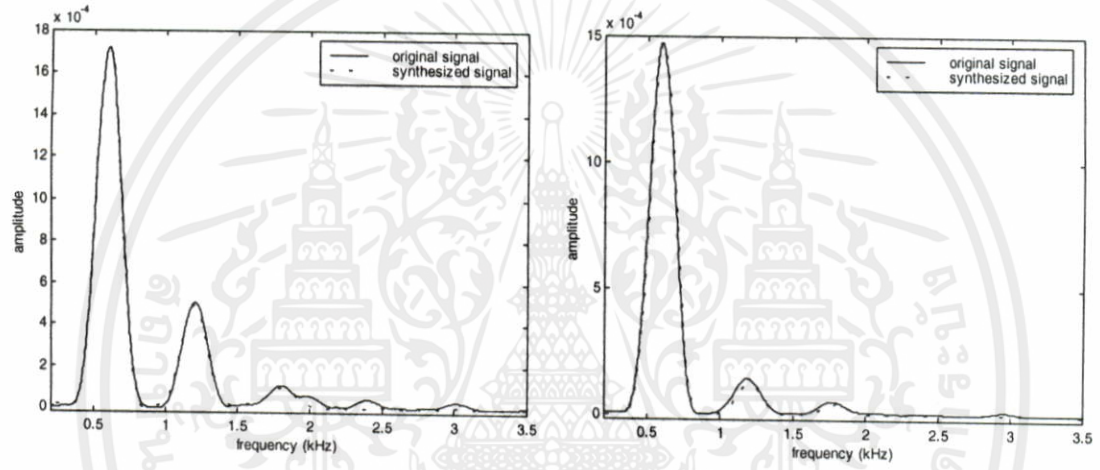
ก. โน้ตโด

ข. โน้ตเร

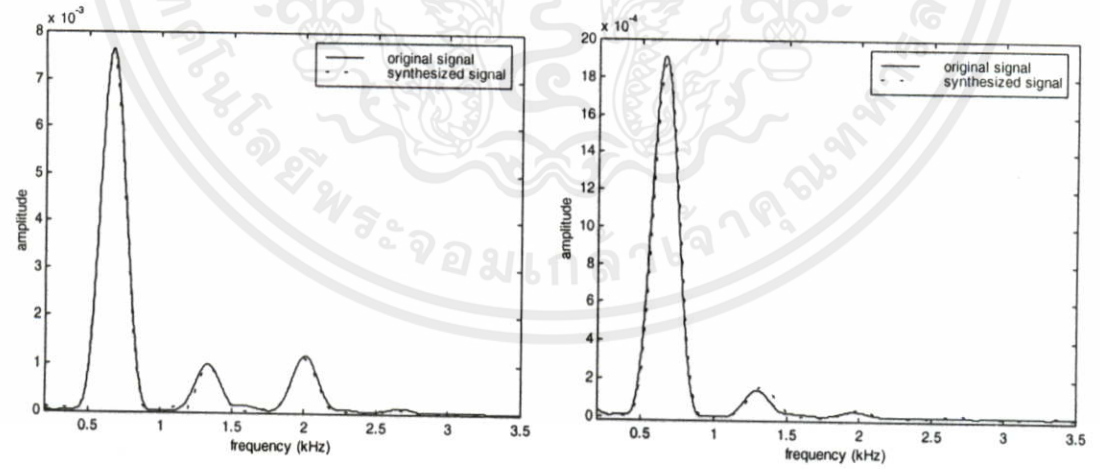
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ค)



(ง)



(จ)

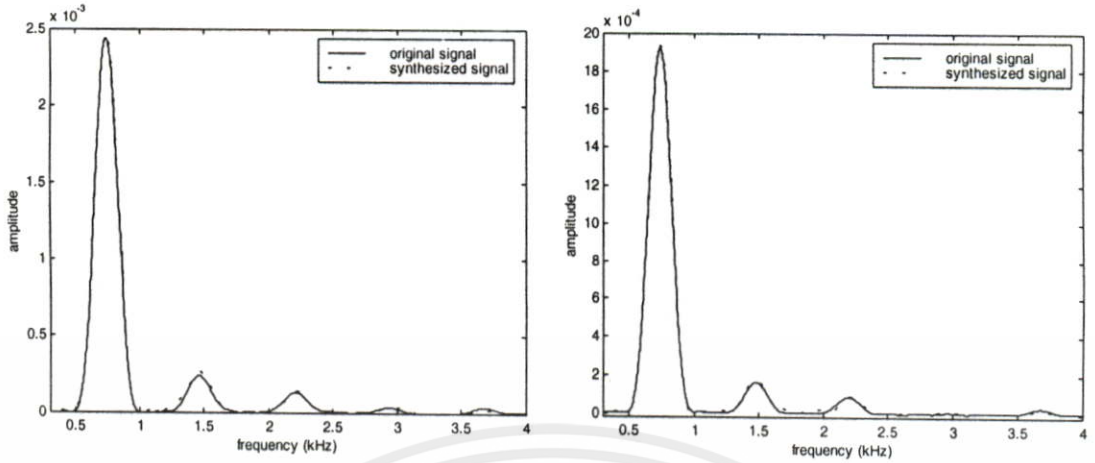
รูปที่ ง.1 (ต่อ) แสดงการเปรียบเทียบสเปกตรัมส่วนที่เป็น attack (ซ้าย) และส่วนที่เป็น decay (ขวา) ของสัญญาณเสียงต้นแบบกับที่สังเคราะห์ขึ้นใหม่

ค. โน้ตมี

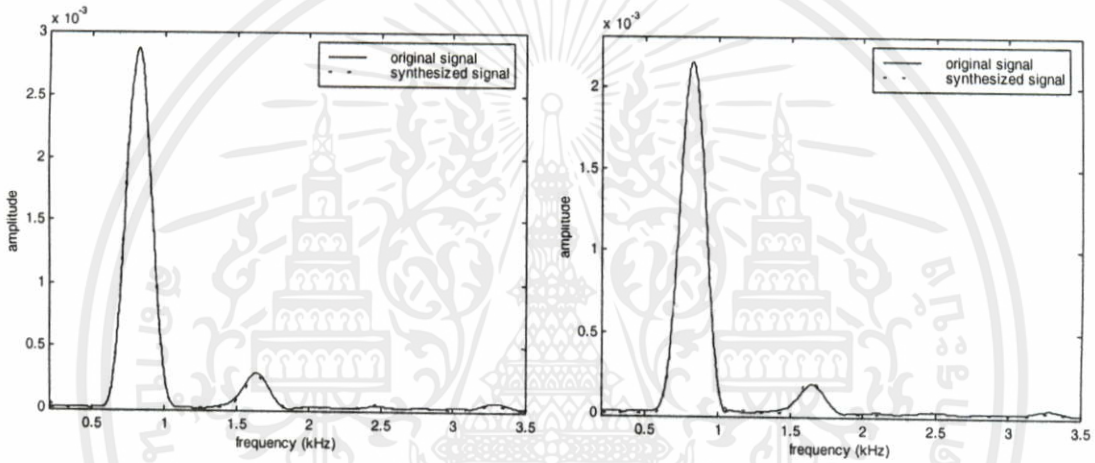
ง. โน้ตฟา

จ. โน้ตซอล

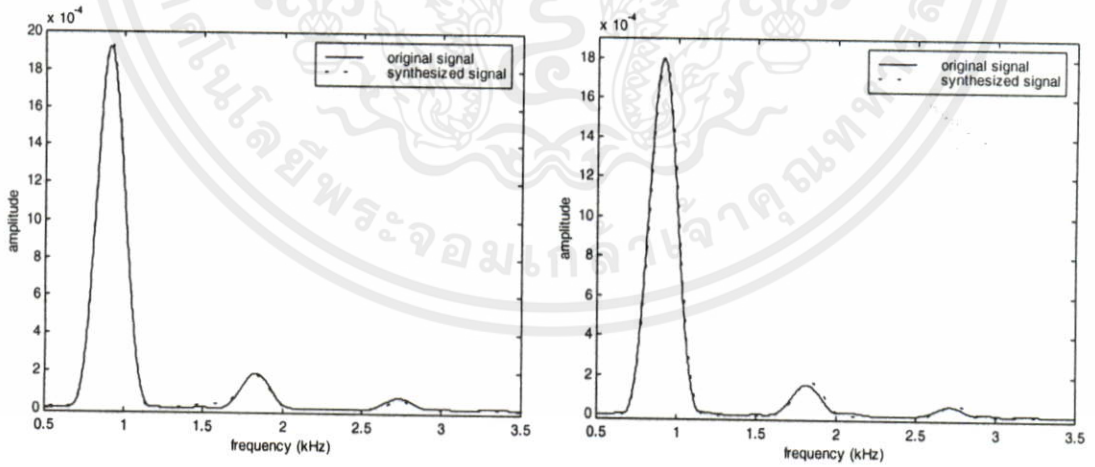
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ฉ)



(ช)



(ซ)

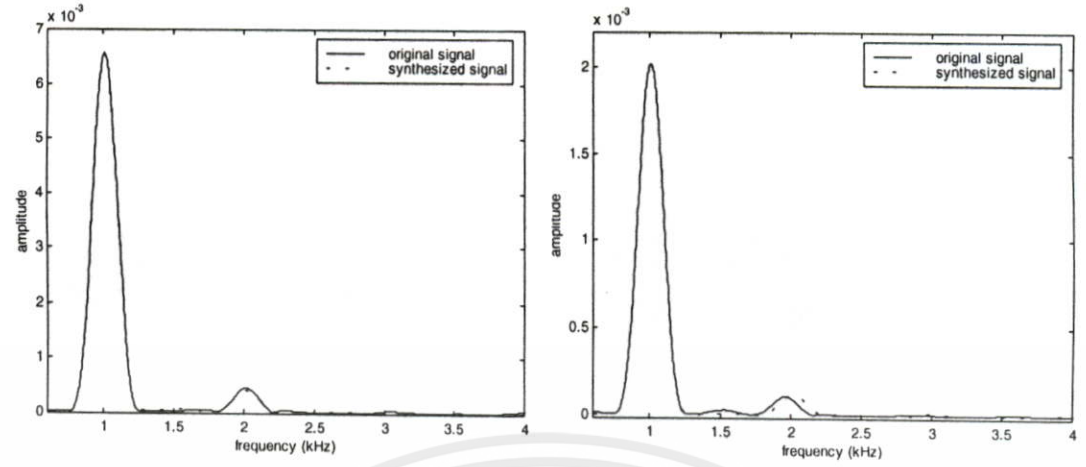
รูปที่ ง.1 (ต่อ) แสดงการเปรียบเทียบสเปกตรัมส่วนที่เป็น attack (ซ้าย) และส่วนที่เป็น decay (ขวา) ของสัญญาณเสียงต้นแบบกับที่สังเคราะห์ขึ้นใหม่

ฉ. โน้ตลา

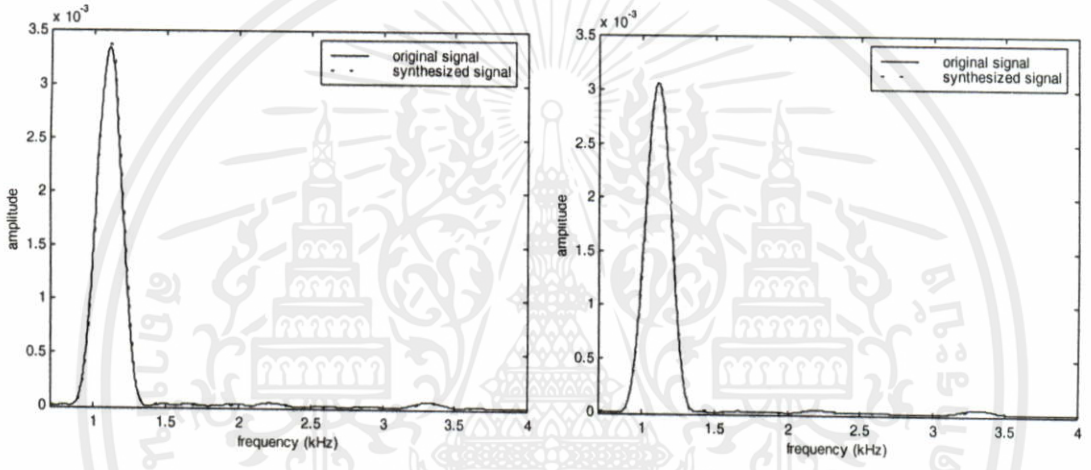
ช. โน้ตที

ซ. โน้ตโดสูง

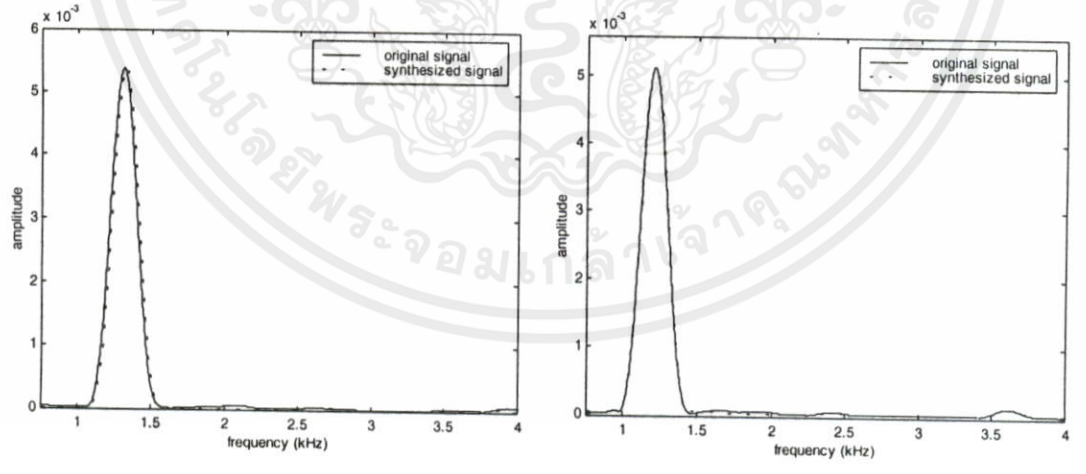
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น ห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ฌ)



(ญ)



(ฎ)

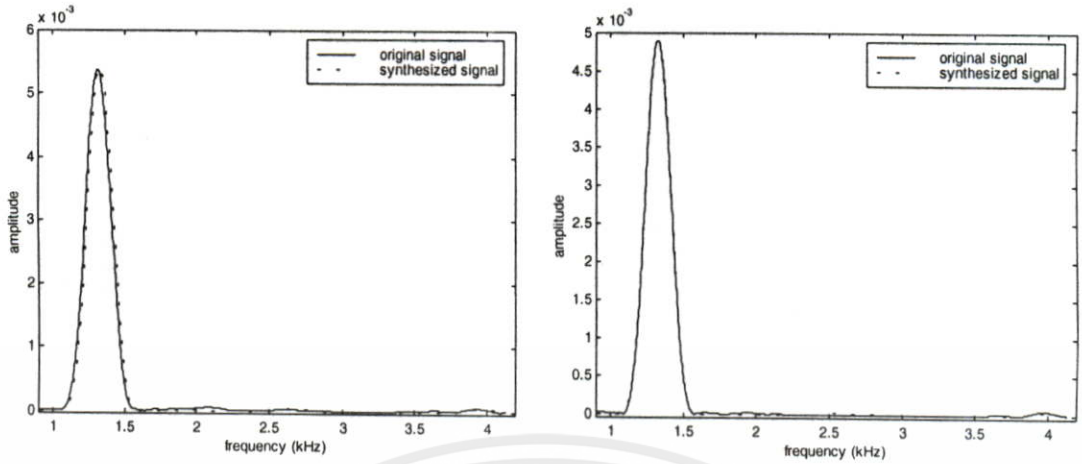
รูปที่ ง.1 (ต่อ) แสดงการเปรียบเทียบสเปกตรัมส่วนที่เป็น attack (ซ้าย) และส่วนที่เป็น decay (ขวา) ของสัญญาณเสียงต้นแบบกับที่สังเคราะห์ขึ้นใหม่

ฌ. โน้ตเรสูง

ญ. โน้ตมีสูง

ฎ. โน้ตฟาสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ฐ)

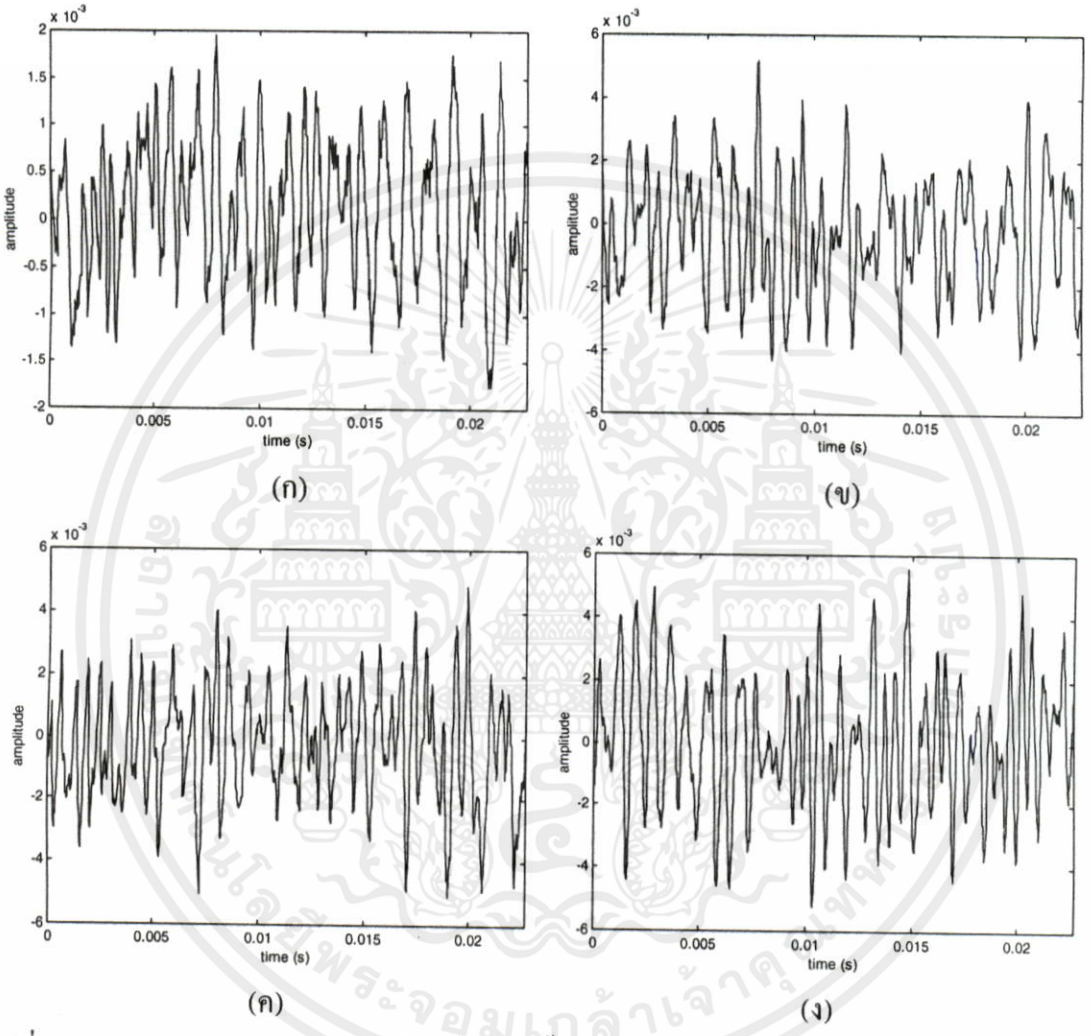
รูปที่ ง.1 (ต่อ) แสดงการเปรียบเทียบสเปกตรัมส่วนที่เป็น attack (ซ้าย) และส่วนที่เป็น decay (ขวา) ของสัญญาณเสียงต้นแบบกับที่สังเคราะห์ขึ้นใหม่

ฐ. โน้ตซอลสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก จ

ลักษณะของสัญญาณรบกวน



รูปที่ จ.1 แสดงลักษณะสัญญาณรบกวนที่ได้จากขั้นตอนการวิเคราะห์สัญญาณรบกวน

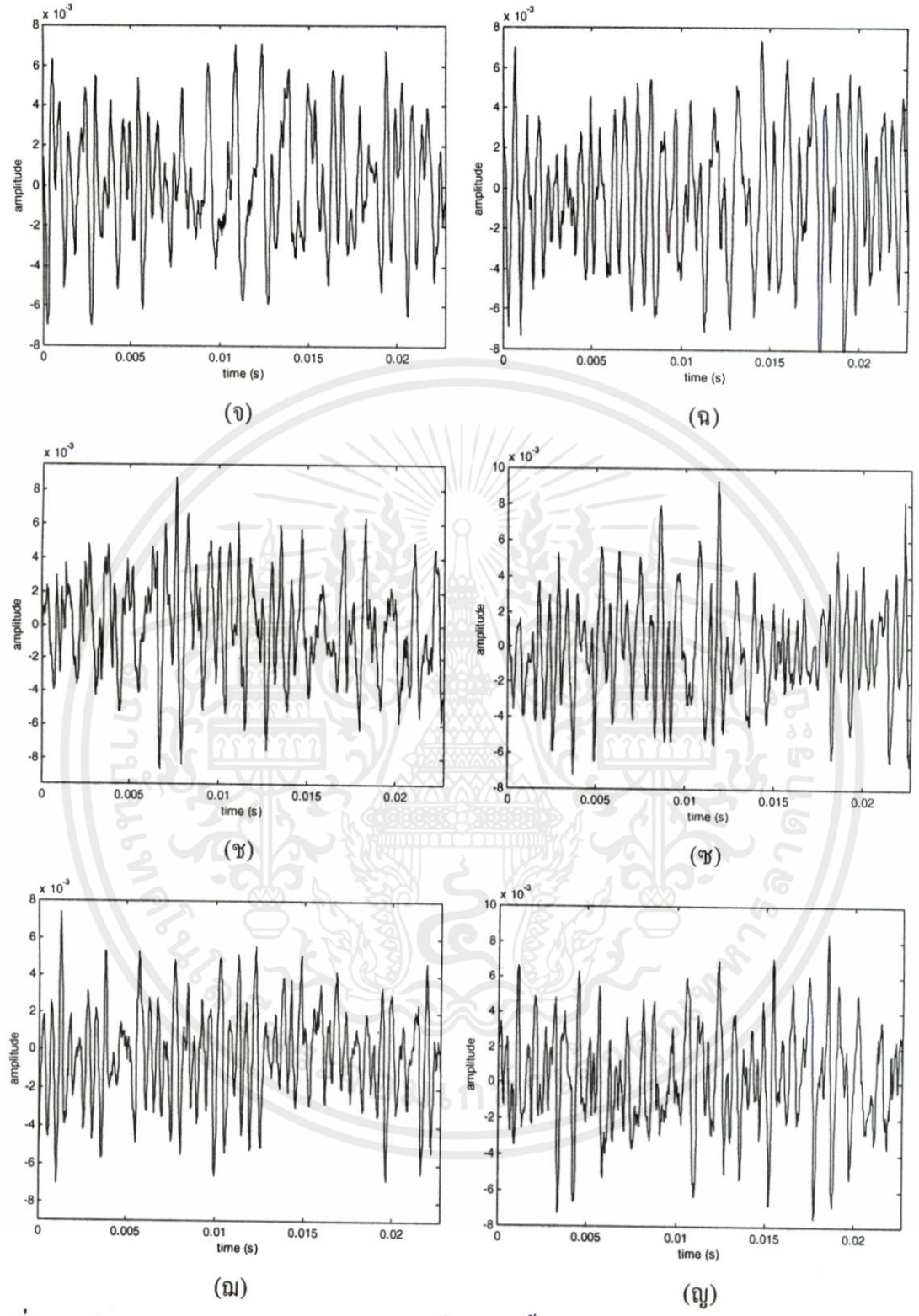
ก. โน้ตโด

ข. โน้ตเร

ค. โน้ตมี

ง. โน้ตฟา

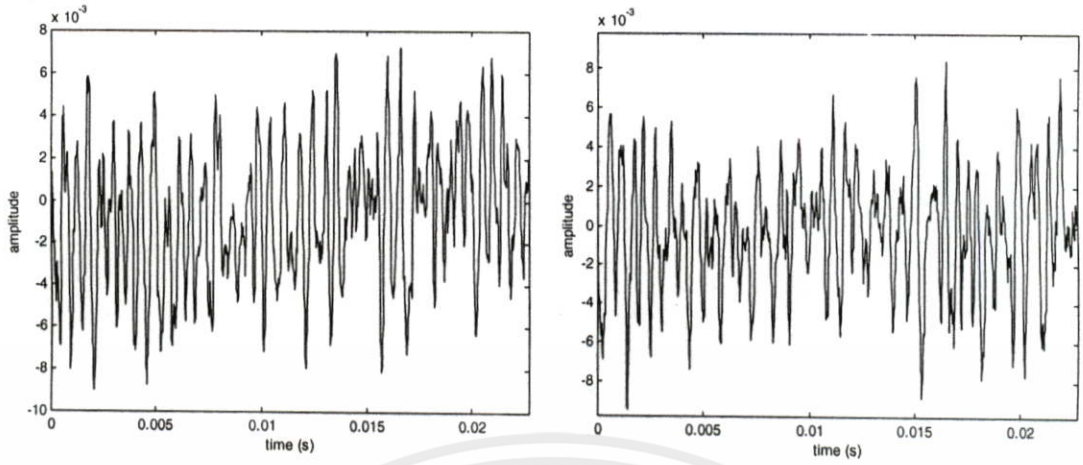
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ จ.1 (ต่อ) แสดงลักษณะสัญญาณรบกวนที่ได้จากขั้นตอนการวิเคราะห์สัญญาณรบกวน

จ. โน้ตซอล
ฉ. โน้ตลา
ช. โน้ตที
ซ. โน้ตโดสูง
ฌ. โน้ตเรสูง
ญ. โน้ตมีสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น และห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ฎ) (ฐ)

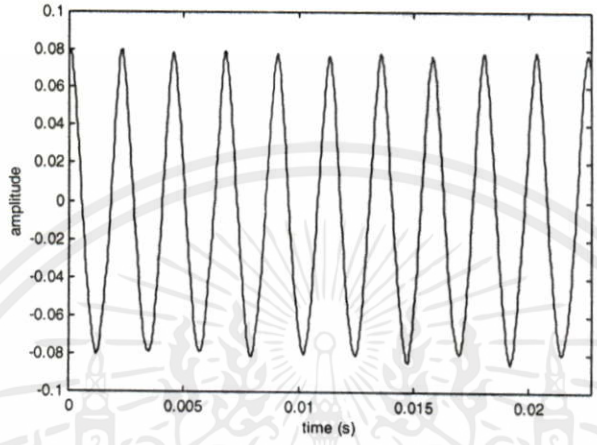
รูปที่ จ.1 (ต่อ) แสดงลักษณะสัญญาณรบกวนที่ได้จากขั้นตอนการวิเคราะห์สัญญาณรบกวน

ฎ. โน้ตฟาสูง

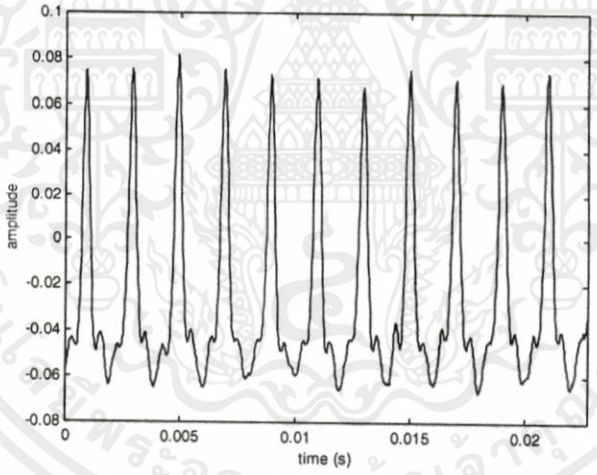
ฐ. โน้ตซอลสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ฉ
ลักษณะของสัญญาณเสียงสังเคราะห์



(ก)



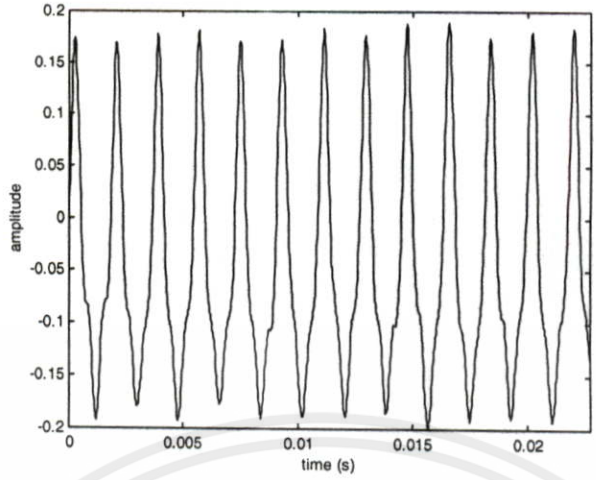
(ข)

รูปที่ ฉ.1 แสดงลักษณะสัญญาณเสียงสังเคราะห์

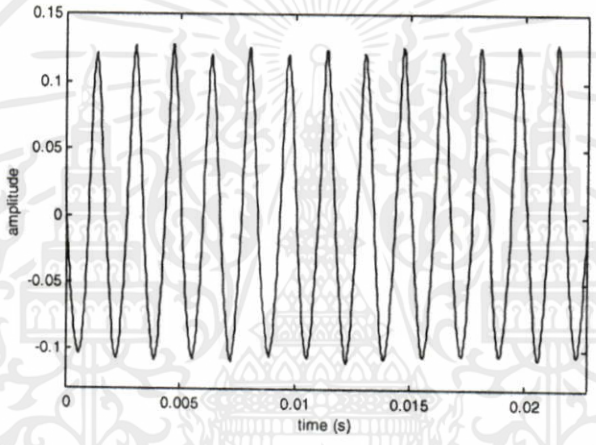
ก. โน้ตโด

ข. โน้ตเร

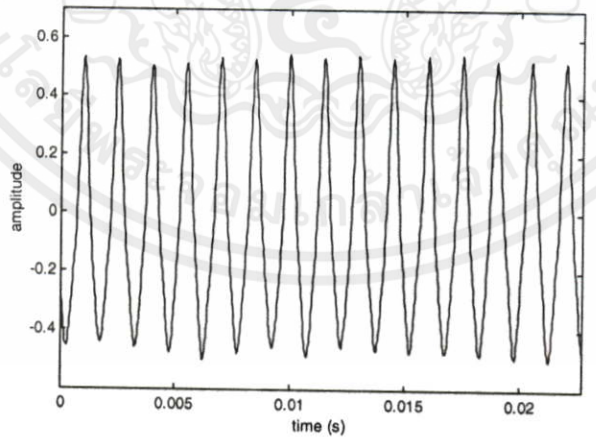
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ค)



(ง)



(จ)

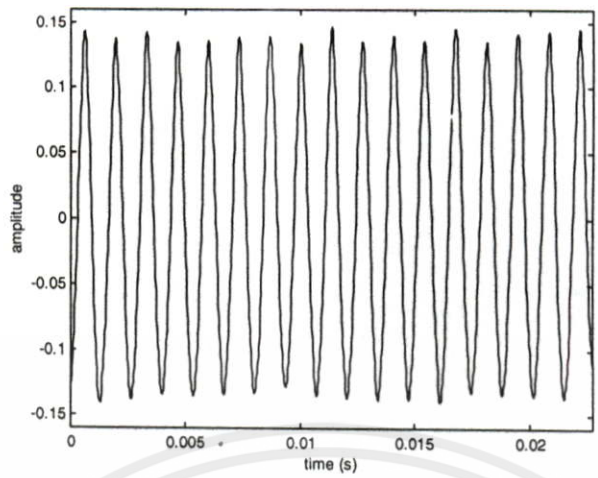
รูปที่ ฉ.1 (ต่อ) แสดงลักษณะสัญญาณเสียงสังเคราะห์

ค. โน้ตมี

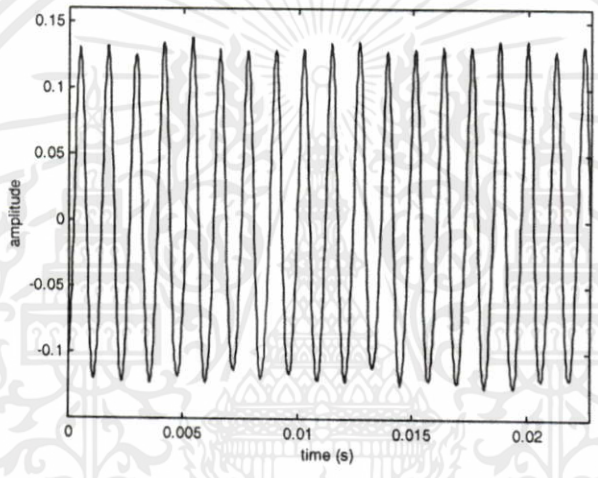
ง. โน้ตฟา

จ. โน้ตซอล

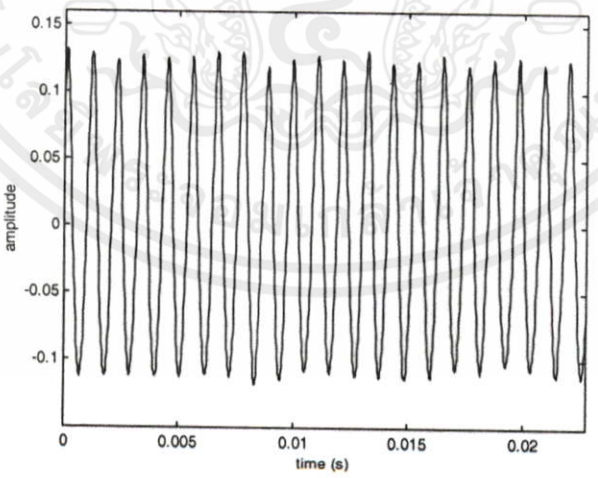
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ฉ)



(ช)



(ซ)

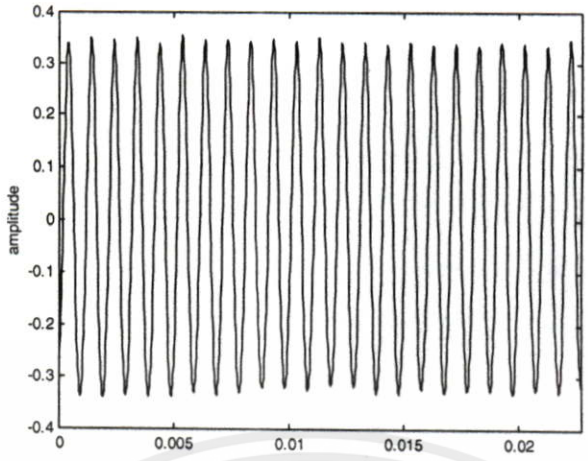
รูปที่ ฉ.1 (ต่อ) แสดงลักษณะสัญญาณเสียงสังเคราะห์

ฉ. โน้ตลา

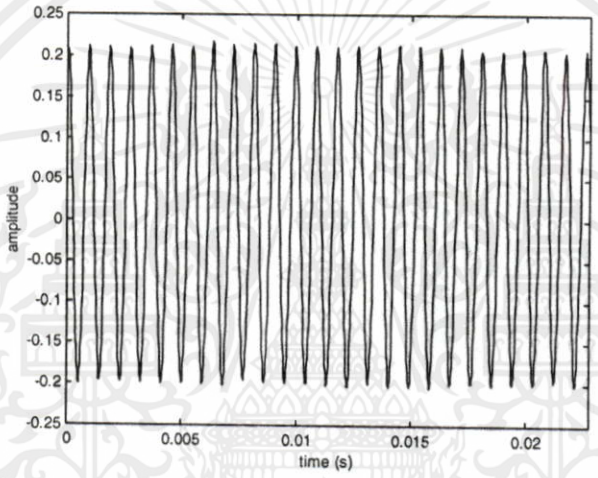
ช. โน้ตที

ซ. โน้ตโดสูง

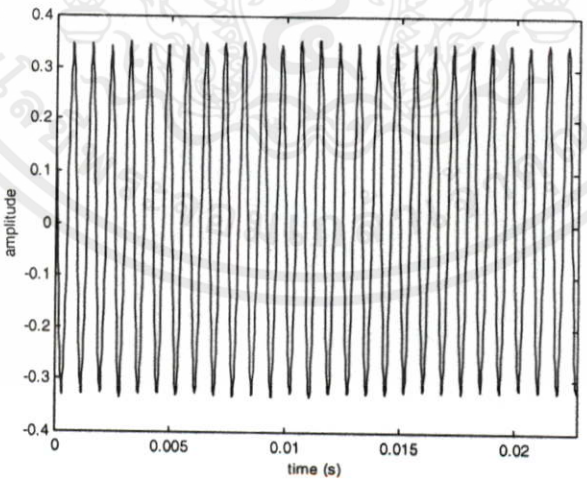
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น และห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ฌ)



(ญ)



(ฎ)

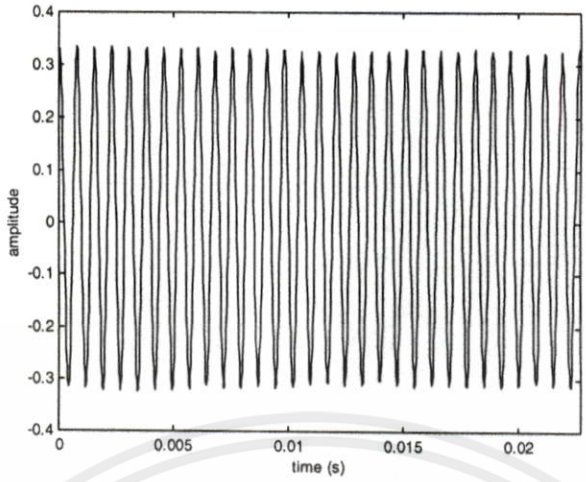
รูปที่ ฉ.1 (ต่อ) แสดงลักษณะสัญญาณเสียงสังเคราะห์

ฌ. โน้ตเรสูง

ญ. โน้ตมีสูง

ฎ. โน้ตฟาสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ฐ)

รูปที่ ฉ.1 (ต่อ) แสดงลักษณะสัญญาณเสียงสังเคราะห์

ฐ. โน้ตซอลสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ฉ
บทความที่ได้รับการตีพิมพ์

1. รุ่งกานต์ ศิริเจริญไชย และ นพพร โชติกกำธร. "การวิเคราะห์และการสังเคราะห์เสียงขลุ่ยไทยโดยใช้โมเดลแบบ Sinusoidal Plus Noise." สาร NECTEC, ปีที่ 7 ฉบับที่ 36 กันยายน-ตุลาคม 2543. หน้า 30-38.
2. Siricharoenchai, R. and N. Chotikakamthorn. "Thai Flute Sound Synthesis Using Sinusoidal Plus Noise Model." *The 1999 National Computer Science and Engineering Conference*, Bangkok, Thailand.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของผลงานนี้

บทความทั่วไป

โดย: รุ่งกานต์ ศิริเจริญไชย
นพพร โชติกกำธร

การวิเคราะห์และการสังเคราะห์เสียงขลุ่ยไทย โดยใช้โมเดลแบบ Sinusoidal Plus Noise



ABSTRACT - This paper presents the results from research study on the application of spectral analysis and synthesis technique for the sound of Thai flute. In this study the sinusoidal plus noise model was applied. We analyzed recorded sound of Thai flute (both the steady and transient state) to estimate the values of parameters required for sound synthesis. Such parameters include frequency, amplitude and filter coefficients. From the analysis, we found that the sound of Thai flute is generally composed of 5-12 principal harmonics, depending on the pitch of sound analyzed. From the sound quality listening test, participated by Thai musical-trained listeners, it has been found that at least 3 filter coefficients are needed in generating the noise part, to obtain the synthesized sound which can be recognized as a sound of Thai flute.

KEY WORDS--sound synthesis, sinusoidal plus noise model, spectral estimation, Thai flute

เอกสารนี้เป็นเอกสารของสำนักวิทยบริการฯ ใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น ทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทคัดย่อ

บทความนี้นำเสนอผลการศึกษาวิจัยการประยุกต์ใช้แนวทางการวิเคราะห์และสังเคราะห์เสียงเชิงสเปกตรัมกับเครื่องดนตรีขลุ่ยไทย โดยอาศัยแบบจำลองการสังเคราะห์เสียงดนตรีแบบ Sinusoidal Plus Noise Model ในงานวิจัยได้ทำการวิเคราะห์เสียงขลุ่ย (ในช่วงภาวะเสถียรและช่วงที่มีการเปลี่ยนแปลง) เพื่อหาค่าพารามิเตอร์สำหรับใช้ในการสังเคราะห์เสียงด้วยแบบจำลองดังกล่าว เช่น ค่าความถี่พื้นฐานและฮาร์โมนิก ค่าแอมปลิจูดและค่าสัมประสิทธิ์ตัวกรองสัญญาณ โดยจากการวิเคราะห์สเปกตรัมของเสียงขลุ่ยพบว่า ประกอบด้วยความถี่พื้นฐานและความถี่ฮาร์โมนิกจำนวน 5-12 ค่าความถี่ ซึ่งขึ้นอยู่กับว่าเป็นเสียงโน้ตตัวใด และจากการวัดคุณภาพของเสียงที่สังเคราะห์ได้ด้วยการฟัง โดยอาศัยการให้คะแนนจากผู้ที่มีพื้นฐานทางด้านดนตรีไทย พบว่าจำนวนสัมประสิทธิ์ของตัวกรองสัญญาณ (ซึ่งที่ใช้ในกระบวนการสังเคราะห์) อย่างน้อย 3 จำนวน ให้เสียงสังเคราะห์ที่ยังคงลักษณะเสียงขลุ่ย

คำสำคัญ – การสังเคราะห์เสียง แบบจำลองการสังเคราะห์เสียงซินนูซอยดัลพลัสนอยซ์ การประมาณค่าเชิงสเปกตรัม ขลุ่ยไทย

บทนำ

การสังเคราะห์เสียงดนตรีด้วยคอมพิวเตอร์นั้นควรเลือกใช้แบบจำลองที่ดีที่เหมาะสมกับการประยุกต์ใช้ ซึ่งในปัจจุบันมีวิธีการสังเคราะห์เสียงอยู่หลายวิธี เช่น วิธีการแซมปลิง ซึ่งเป็นวิธีการจัดเก็บเสียงในรูปของคลื่นเสียง ทำให้เสียงสังเคราะห์ที่ได้จากวิธีการดังกล่าวมีคุณภาพดี (ในแง่ของความเป็นธรรมชาติ) แต่มีขีดความสามารถในการดัดแปลงแก้ไขเสียงได้จำกัด

นอกจากวิธีการดังกล่าวแล้ว เทคนิคการสังเคราะห์เสียงใหม่ๆ ได้ถูกพัฒนาขึ้นในปัจจุบันมีอาทิ เช่น การสังเคราะห์เสียงเชิงกายภาพ วิธีการดังกล่าวอาศัยการจำลองลักษณะปรากฏการณ์ทางกายภาพของการเกิดเสียงเพื่อการสังเคราะห์ (1) โดยพารามิเตอร์ต่างๆ ที่ใช้ในการสังเคราะห์เสียงเป็นค่าคุณสมบัติทางกายภาพของเครื่องดนตรีแต่ละชนิด ทำให้ได้เสียงสังเคราะห์ที่มีความเป็นธรรมชาติ แต่มีความยุ่งยากเมื่อนำไปประยุกต์ใช้กับเสียงของเครื่องแต่ละชนิด เนื่องจากในการสังเคราะห์เสียงเครื่องดนตรีแต่ละประเภทจำเป็นต้องทราบแบบจำลองของเครื่องดนตรีชนิดนั้น

วิธีการสังเคราะห์เสียงเชิงสเปกตรัมเป็นวิธีการสังเคราะห์เสียงอีกวิธีหนึ่งที่ได้รับการพัฒนาขึ้นเมื่อไม่นานมานี้ (2)(3) วิธีการดังกล่าวมีข้อดีคือสามารถสังเคราะห์เสียงได้หลายประเภทไม่จำกัดเฉพาะเสียงเครื่องดนตรี และมีความยืดหยุ่นสูงกว่าในการดัดแปลงหรือสร้างเสียงใหม่เมื่อเปรียบเทียบกับวิธีการสังเคราะห์เสียงเชิงกายภาพและการสังเคราะห์เสียงด้วยการแซมปลิง จากงานวิจัยที่ผ่านมา ได้นำวิธีการสังเคราะห์เสียงดังกล่าวมาสร้างซอฟต์แวร์ซินทีไซเซอร์ (Software Synthesizer) เพื่อการสังเคราะห์เสียงเครื่องดนตรี (4) แต่สำหรับเครื่องดนตรีไทยยังไม่ปรากฏ งานวิจัยที่นำเสนอในบทความนี้จึงอาศัยแบบจำลองเชิงสเปกตรัมในการวิเคราะห์และสังเคราะห์เสียงเครื่องดนตรีขลุ่ยไทย ในการวิจัยผู้วิจัยได้ทำการวิเคราะห์หาส่วนประกอบเชิงความถี่ของเสียงขลุ่ยไทยเพื่อหาค่าพารามิเตอร์สำหรับใช้ในการสังเคราะห์เสียง นอกจากนี้ยังนำเสนอเทคนิคการสังเคราะห์เพื่อการวิเคราะห์ (Analysis by Synthesis) ในการวิเคราะห์ช่วงสัญญาณที่ไม่เสถียรตลอดจนทำการศึกษาถึงคุณภาพของเสียงและความซับซ้อนในการคำนวณด้วยวิธีการสังเคราะห์เสียงตามแนวทางดังกล่าว

การสังเคราะห์เสียงด้วยแบบจำลองเชิงสเปกตรัม (Spectral Modeling)

แนวทางการสังเคราะห์เสียงเชิงสเปกตรัมเริ่มจากการวิเคราะห์สเปกตรัมของเสียงเพื่อหาค่าพารามิเตอร์สำหรับใช้ในการสังเคราะห์เสียง โดยค่าพารามิเตอร์ต่างๆ เช่น ค่าความถี่และค่าแอมปลิจูดที่ได้จากการวิเคราะห์ดังกล่าวจะถูกนำใช้ในขั้นตอนของการสังเคราะห์เสียงต่อไป

การสังเคราะห์เสียงเชิงสเปกตรัมอาจแบ่งย่อยออกได้เป็น 2 แนวทาง (5) ดังต่อไปนี้

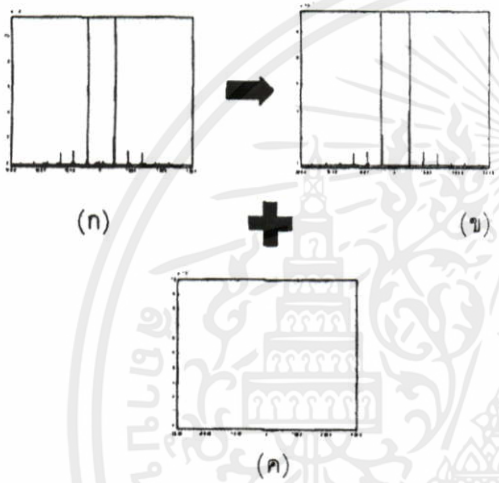
1) Sinusoidal Model เป็นวิธีการที่จำลองเสียงดนตรีด้วยผลรวมของสัญญาณไซน์ที่ความถี่ต่างๆ ดังแสดงในสมการต่อไปนี้

$$s(n) = \sum_{i=1}^{} A_i \cos(\omega_i n + \theta_i) \quad (1)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ $s(n)$ คือสัญญาณที่ได้ A_r คือค่าแอมปลิจูดของสัญญาณลำดับที่ r ω_r คือค่าความถี่เชิงมุมของสัญญาณไซน์ลำดับที่ r θ_r คือค่าเฟสของสัญญาณไซน์ลำดับที่ r n คือลำดับที่ของจุดข้อมูล และ R คือจำนวนสัญญาณไซน์ที่พิจารณาทั้งหมด

2) Sinusoidal Plus Noise Model การสังเคราะห์เสียงภายใต้โมเดล Sinusoidal Plus Noise Model อยู่บนพื้นฐานของการจำลองเสียงดนตรีด้วยผลรวมขององค์ประกอบสัญญาณ 2 ส่วน (6)(7) คือส่วนที่เป็นสัญญาณคลื่นไซน์และส่วนที่เป็นสัญญาณรบกวน (Noise:ในที่นี้หมายถึง เสียงที่มีลักษณะของสัญญาณที่ไม่สามารถจำลองได้ด้วย



รูปที่ 1 แสดงตัวอย่างการแยกองค์ประกอบของสัญญาณเสียงออกเป็นสองส่วน
ก. สเปกตรัมของสัญญาณเสียงที่ทำการวิเคราะห์
ข. สเปกตรัมของส่วนที่เป็นสัญญาณคลื่นไซน์
ค. สเปกตรัมของส่วนที่เป็นสัญญาณเสียงรบกวน (noise)

ฟังก์ชันไซน์ หากแต่สามารถอธิบายได้ด้วยคุณสมบัติทางสถิติ มิได้หมายถึงเสียงหรือสัญญาณที่ไม่พึงปรารถนา) ดังแสดงในรูปที่ 1

วิธีการสังเคราะห์เสียงของแนวทางนี้ประกอบไปด้วย 2 ส่วน ดังรูปที่ 2 คือ ส่วนของสัญญาณคลื่นไซน์ซึ่งขั้นตอนในการสร้างสัญญาณส่วนนี้ จะกระทำในลักษณะเดียวกันกับวิธีการ Sinusoidal Model ตามสมการที่ (1) และส่วนของสัญญาณรบกวนซึ่งสามารถสังเคราะห์ได้โดยการกรองสัญญาณ white noise ด้วยฟิลเตอร์ที่เหมาะสม เสียงสังเคราะห์ที่ได้ด้วยโมเดลดังกล่าวแสดงดังสมการต่อไปนี้

$$s(n) = \sum_{r=1}^{R} A_r(n)\cos(\omega_r n) + e(n) \qquad (2)$$

โดยที่ $s(n)$ คือสัญญาณที่ได้ (output signal) $A_r(n)$ คือค่าแอมปลิจูดของสัญญาณไซน์ลำดับที่ r $\omega_r n$ คือค่าเฟสของสัญญาณไซน์ลำดับที่ r R คือจำนวนสัญญาณไซน์ที่พิจารณาทั้งหมด $e(n)$ คือส่วนของสัญญาณรบกวนซึ่งสามารถคำนวณหาได้จากสมการต่อไปนี้

$$e(n) = \sum_{k=0}^{M-1} h(k)w(n-k) \qquad (3)$$

โดยที่ $h(k)$ คือสัมประสิทธิ์ตัวกรองสัญญาณ (filter coefficient) $w(n)$ คือสัญญาณ white noise และ M คือจำนวนสัมประสิทธิ์ตัวกรองสัญญาณ

จากสมการข้างต้นจะเห็นว่าพารามิเตอร์ที่ต้องใช้ในการสังเคราะห์ คือ ค่าแอมพลิจูด ค่าความถี่ และค่าสัมประสิทธิ์ตัวกรองสัญญาณ ซึ่งค่าพารามิเตอร์เหล่านี้สามารถหาได้จากการวิเคราะห์เสียงดนตรีจริง

การสังเคราะห์เสียงโดยใช้แบบจำลอง Sinusoidal Plus Noise

การสังเคราะห์เสียงด้วยแนวทางนี้มี 2 ขั้นตอนหลัก คือ ขั้นตอนการวิเคราะห์สัญญาณเสียงหาค่าพารามิเตอร์ต่างๆ เพื่อนำไปใช้สร้างเสียงสังเคราะห์ ในขั้นตอนของการสังเคราะห์ ในการวิเคราะห์สัญญาณเสียง แบ่งออกเป็นสองส่วนคือส่วนของสัญญาณไซน์และในส่วนของสัญญาณรบกวน โดยมีรายละเอียดดังนี้

การวิเคราะห์ในส่วนของสัญญาณคลื่นไซน์

ในส่วนของการวิเคราะห์สเปกตรัมของสัญญาณเพื่อให้ได้ค่าความถี่และแอมพลิจูดที่ถูกต้อง ชนิดของวินโดว์ที่นำมาคูณกับสัญญาณเสียงจริง ขนาดของวินโดว์ของสัญญาณ และการเพิ่มจำนวนศูนย์เข้ากับสัญญาณจะให้ค่าพารามิเตอร์ที่มีความถูกต้องต่างกัน

1) ชนิดของวินโดว์

เนื่องจากชนิดของวินโดว์ที่นำมาคูณกับสัญญาณจริงจะให้ผลค่าความถูกต้องของความถี่และแอมพลิจูดแตกต่างกัน ซึ่งในที่นี้เป็นการประยุกต์ใช้กับสัญญาณเสียงขลุ่ยไทยซึ่งมีช่วงห่างระหว่างความถี่ฮาร์โมนิกพอสมควร ดังนั้นจึงเลือกใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

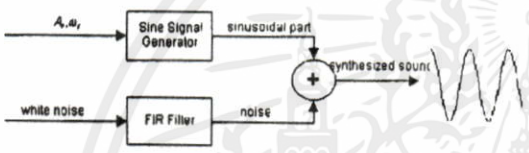
แฮมมิ่งวินโดว์ เนื่องจากให้ค่าที่มีความถูกต้องมากกว่าเรคแทงคิวลาร์วินโดว์ นอกจากนี้แฮมมิ่งวินโดว์ทำให้การพิจารณาหาค่าพารามิเตอร์ในส่วนของสัญญาณรบกวนมีความถูกต้องมากขึ้นอีกด้วย

2) ขนาดของวินโดว์

ขนาดของวินโดว์ที่เหมาะสมมีผลทำให้การหาค่าความถี่หลักและความถี่ฮาร์โมนิกมีความถูกต้องมากขึ้น ซึ่งขนาดของวินโดว์เริ่มต้นที่เหมาะสมเป็นไปตามสมการต่อไปนี้ (8)

$$M \geq \frac{B \cdot f_s}{f_o} \qquad (4)$$

โดยที่ M คือขนาดของวินโดว์ B คือ mainlobe ของวินโดว์ที่ใช้ f_s คือความถี่แซมเปิล f_o คือความถี่พื้นฐาน (fundamental frequency)



รูปที่ 2 แสดงไดอะแกรมขั้นตอนการสังเคราะห์เสียง

3) การเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์เข้ากับสัญญาณที่วิเคราะห์ (Zero padding)

โดยทั่วไปการเพิ่มจุดข้อมูล มีผลทำให้การวิเคราะห์มีความถูกต้องมากขึ้น แต่อย่างไรก็ดี จากการทดลองนี้เราพบว่าการเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์อาจทำให้มีความคลาดเคลื่อนมากขึ้นได้ ทั้งนี้ขึ้นอยู่กับจำนวนจุดของข้อมูลที่มีค่าเป็นศูนย์ จากการวิเคราะห์พบว่า หากเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์ตามสมการต่อไปนี้จะทำให้ค่าพารามิเตอร์ (เช่น ค่าความถี่และแอมพลิจูด) มีความถูกต้องมากขึ้นตามจำนวนจุดที่เพิ่มขึ้น

$$T=N.(2^k-1) \qquad (5)$$

โดยที่ T คือจำนวนจุดข้อมูลที่มีค่าเป็นศูนย์ที่เพิ่มเข้ากับสัญญาณ N คือจำนวนแซมเปิลของสัญญาณที่พิจารณา k = 0,1,2,3,... ,

ขั้นตอนการวิเคราะห์ในส่วนของสัญญาณไซน์ มีรายละเอียดดังนี้

1. นำส่วนของสัญญาณที่ต้องการวิเคราะห์จำนวน M แซมเปิล (โดยจำนวนM เป็นไปตามสมการที่ 4) คูณกับวินโดว์ฟังก์ชัน (ในที่นี้เลือกใช้แฮมมิ่งวินโดว์)

$$s_w(n)=s(n).\ w(n) \quad ,n = 0,...,M-1 \qquad (6)$$

โดยที่ $s_w(n)$ คือสัญญาณที่คูณกับแฮมมิ่งวินโดว์ $s(n)$ คือสัญญาณขนาด M แซมเปิล $w(n)$ คือแฮมมิ่ง วินโดว์ซึ่งหาได้จากสมการต่อไปนี้ (9)

$$w(n) = 0.54 - 0.46 \cdot \cos\left(\frac{2\pi n}{D-1}\right), 0 \leq n < D \qquad (7)$$

โดยที่ D คือขนาดของแฮมมิ่งวินโดว์

2. เพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์ให้กับสัญญาณตามสมการที่ (5)

3. คำนวณค่าสเปกตรัมของสัญญาณ เช่น โดยอาศัยอัลกอริทึม FFT

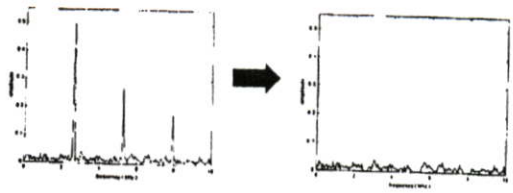
4. จากสเปกตรัมที่ได้ ทำการตรวจหาจุดที่มีค่าความชันสูง (Peak) เพื่อคำนวณค่าความถี่ แอมพลิจูด ที่จุดนั้นๆ สำหรับใช้ในกระบวนการสังเคราะห์เสียง

การวิเคราะห์ในส่วนของสัญญาณรบกวน

ในการคำนวณหาส่วนของสัญญาณรบกวนโดยการลบส่วนของสัญญาณไซน์ออกจากสัญญาณเสียง สามารถทำได้ทั้งในโดเมนของเวลา (time domain) และโดเมนของความถี่ (frequency domain) กรณีโดเมนของเวลา ทำได้โดยนำสัญญาณเสียงจริงลบส่วนของสัญญาณไซน์หรือไซน์นูซอยด์ที่สังเคราะห์ขึ้นใหม่ โดยในทางปฏิบัติทำได้ยาก เนื่องจากว่าสัญญาณที่ได้จากการลบอาจมีส่วนของสัญญาณไซน์เหลืออยู่ ซึ่งเกิดขึ้นได้ในกรณีที่ค่าพารามิเตอร์ที่ใช้ในการสังเคราะห์ไม่ใช่ค่าที่ถูกต้องจริงๆ ดังนั้นในงานวิจัยนี้จึงพิจารณาที่โดเมนของความถี่ ซึ่งจะให้ส่วนของสัญญาณรบกวนที่ถูกต้องมากกว่า โดยการวิเคราะห์ในส่วนของสัญญาณรบกวนมีขั้นตอนดังนี้

1. จากสเปกตรัมของสัญญาณที่ได้จากขั้นตอนการวิเคราะห์ในส่วนของสัญญาณคลื่นไซน์ ทำการหาสเปกตรัมของส่วนสัญญาณรบกวนโดยกำหนดค่าในช่วงของสเปกตรัมที่มีค่าความชันสูง (ตามที่ตรวจหาจากขั้นตอนที่ 4 ในการวิเคราะห์ข้างต้น) ให้มีค่าเป็นศูนย์ ดังรูปที่ 3 โดยเริ่มจาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก) (ข)

รูปที่ 3 แสดงตัวอย่างสเปกตรัมของสัญญาณ
ก. สเปกตรัมของสัญญาณเสียงที่ต้องการวิเคราะห์
ข. สเปกตรัมของสัญญาณเสียงที่ตัดส่วนของสัญญาณไซน์ออกแล้ว

ค่าต่ำสุดที่เป็นจุดเริ่มของความชันจนถึงค่าต่ำสุดซึ่งเป็นจุดสิ้นสุดของค่าความชัน

2. สร้างสัญญาณเสียงของส่วนที่เป็นสัญญาณรบกวน โดยการหาอินเวร์สฟูเรียร์ทรานฟอร์ม (Inverse Fourier Transform) ของสเปกตรัมที่ได้ในข้อ 1

3. คำนวณเพาเวอร์สเปกตรัมของสัญญาณรบ-กวนโดยเริ่มจาก

3.1. แบ่งสัญญาณออกเป็นวินโดว์ขนาด M แซมเปิล โดยให้แต่ละวินโดว์เหลื่อมกัน (overlap) M/2 แซมเปิล

3.2. คำนวณหาฟูเรียร์ทรานฟอร์มในแต่ละวินโดว์

3.3. นำค่าฟูเรียร์ทรานฟอร์มทั้งหมดมาเฉลี่ยด้วยจำนวนวินโดว์ดังสมการต่อไปนี้

$$P(\omega) = \frac{1}{K}\sum_{i=1}^{K}|F_i(\omega)|^2 \qquad (8)$$

โดยที่ คือเพาเวอร์สเปกตรัม K คือจำนวนวินโดว์ทั้งหมด คือฟูเรียร์ทรานฟอร์มลำดับที่ i

4. หาค่าสัมประสิทธิ์ของตัวกรองสัญญาณ h(n) ซึ่ง โดยการทำอินเวอร์สฟูเรียร์ทราน- ฟอร์มของฟังก์ชัน โดยที่

$$H(\omega) = \sqrt{P(\omega)} \qquad (9)$$

การสังเคราะห์เสียง

หลังจากที่ได้วิเคราะห์สัญญาณเสียงทั้ง 2 ส่วน จะได้ค่าพารามิเตอร์ที่สามารถนำมาใช้ในการสังเคราะห์เสียง ซึ่งมีรายละเอียดดังนี้

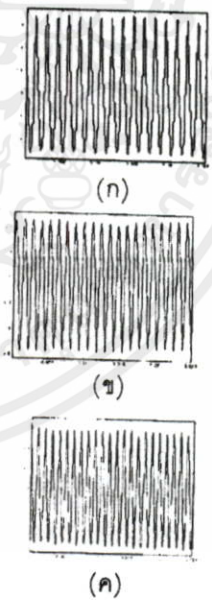
1. สร้างสัญญาณคลื่นไซน์ตามสมการที่ (1) โดยใช้ค่าพารามิเตอร์ที่หาได้จากขั้นตอนการวิเคราะห์ในส่วนของสัญญาณไซน์

2. สร้างสัญญาณรบกวนตามสมการที่(3) โดยใช้สัมประสิทธิ์ตัวกรองสัญญาณที่ได้จากการวิเคราะห์ในส่วนของสัญญารบกวน

3. นำสัญญาณที่สร้างขึ้นใน 2 ขั้นตอนข้างต้นมารวมกันเป็นสัญญาณเสียงสังเคราะห์ที่ต้องการ

ผลการทดลอง

ในการทดลองสังเคราะห์เสียงขลุ่ยผู้วิจัยได้ทำการบันทึกเสียงขลุ่ยเพียงออเก็บเป็นไฟล์เวฟ ที่อัตราการแซมปลิง 44100 เฮิรตซ์ และใช้การจัดเก็บแบบ 16 บิตต่อหนึ่งแซมเปิล โดยในบทความนี้จะยกตัวอย่างการพิจารณาเสียงโน้ตตัวซอล ที โดสูง (สำหรับโน้ตตัวอื่นๆ ดูเอกสารอ้างอิง (10)) ซึ่งมีความยาวประมาณ 1.5 วินาที โดยแต่ละตัวโน้ตใช้จำนวน 5 สัญญาณคลื่นเสียง (จำนวนข้อมูลทั้งหมด 5 ชุด จากการบันทึก 5 ครั้ง) ตัวอย่างของสัญญาณเสียงแสดงดังรูปที่ 4 หลังจากที่ได้นำสัญญาณดังกล่าวมาวิเคราะห์เพื่อหาส่วนประกอบของสัญญาณไซน์ที่ฮาร์โมนิกต่างๆ และคำนวณหาค่าสัมประสิทธิ์ตัวกรองสัญญาณปรากฏว่าพารามิเตอร์ในส่วนของสัญญาณไซน์และสัญญาณรบกวนของ 5 สัญญาณมีค่าใกล้เคียงกัน ดังนั้นจึงเลือกแสดงเพียงสัญญาณเดียว



(ก)

(ข)

(ค)

รูปที่ 4 แสดงคลื่นเสียง
ก. โน้ตตัวซอล
ข. โน้ตตัวมี
ค. โน้ตตัวโดสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนการวิเคราะห์

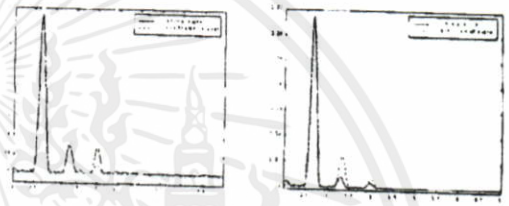
เริ่มจากการวิเคราะห์ในส่วนของสัญญาณไซน์ก่อน จะได้สเปกตรัมของสัญญาณซึ่งนำมาใช้ในการวิเคราะห์ส่วนของสัญญาณรบกวนต่อไป โดยมีรายละเอียดดังนี้

4.1.1. การวิเคราะห์ในส่วนของสัญญาณไซน์

เนื่องจากสัญญาณเสียงโน้ตทุกตัวประกอบไปด้วย 2 ส่วน คือ ส่วนสัญญาณเสียงที่มีการเปลี่ยนแปลง (Attack และ Decay) และส่วนของสัญญาณเสียงที่อยู่ในภาวะเสถียร (Sustain) ดังนั้นในการวิเคราะห์จึงต้องแยกสัญญาณเสียงเป็น 2 ส่วน โดยต้องแบ่งสัญญาณออกเป็นวินโดว์ย่อยๆ เพื่อวิเคราะห์ดูการเปลี่ยนแปลงของสเปกตรัมในแต่ละวินโดว์

1. วิเคราะห์สัญญาณเสียงที่อยู่ในภาวะเสถียร โดยกำหนดขนาดของวินโดว์เริ่มต้น 250 500 1000 2000 และ 40000 แซมเปิล ปรากฏว่าสเปกตรัมของสัญญาณที่ได้ในแต่ละวินโดว์ย่อยมีค่าใกล้เคียงกันมาก จากการทดลองพบว่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของความถี่และแอมพลิจูดมีค่าเท่ากับ 0.64% และ 2.1% ตามลำดับ ซึ่งอยู่ในระดับที่ยอมรับได้ นอกจากนี้ขนาดวินโดว์ที่มีความยาวมากจะให้สเปกตรัมของสัญญาณได้ชัดเจนมากที่สุด ในการทดลองนี้จึงเลือกใช้วินโดว์ที่มีขนาดสูงสุดจากข้อมูลที่บันทึก (50000 แซมเปิล)เพื่อใช้ในการวิเคราะห์หาค่าพารามิเตอร์ต่างๆ ซึ่งแสดงดังตารางที่ 1 ค่าความถี่ที่แสดงดังตารางข้างต้นได้จากการใช้วิธีประมาณค่า Curve Fitting จากค่าที่วัดได้เพื่อให้ได้ค่าความถี่ฮาร์โมนิกต่างๆเป็นจำนวนเท่าของค่าความถี่พื้นฐานพอดี ซึ่งจะทำให้สัญญาณสังเคราะห์ที่ได้ยังคงลักษณะรูปคลื่นเหมือนสัญญาณต้นแบบ (อนึ่งจากการทดลองเปลี่ยนแปลงค่าเฟสในการสังเคราะห์เสียงพบว่า ค่าเฟสไม่มีผลในการสังเคราะห์เสียงแต่อย่างใด จึงมิได้นำค่าเฟสที่วิเคราะห์ได้มาใช้ในการสังเคราะห์)

2. วิเคราะห์สัญญาณเสียงในช่วงที่มีการเปลี่ยนแปลง (ไม่เสถียร เช่น เสียงในช่วงต้นและช่วงสิ้นสุดของแต่ละโน้ต) เนื่องจากสัญญาณเสียงช่วงนี้เป็นสัญญาณที่มีความยาวไม่มาก ดังนั้นการวิเคราะห์สเปกตรัมของสัญญาณ จำเป็นต้องแบ่งสัญญาณออกเป็นวินโดว์ย่อยๆ ที่มีขนาดเล็ก อันทำให้ความละเอียด (Frequency Resolution) ในการวัดความถี่ลดลง แต่จากสมมติฐานที่ว่าสัญญาณเสียงของโน้ตแต่ละตัว น่าจะประกอบไปด้วยความถี่ที่ค่าต่างๆเหมือนกัน ไม่ว่าจะเป็นช่วงภาวะเสถียร หรือช่วงที่แอมพลิจูดของสัญญาณเสียงมีการเปลี่ยนแปลง จึงได้ทำการวิเคราะห์สัญญาณเสียงด้วยวิธีการสังเคราะห์สัญญาณเสียง (Analysis by Synthesis) โดยอาศัยค่าพารามิเตอร์ต่างๆ ที่ได้จากการวิเคราะห์ในช่วงเสถียรข้างต้น มาทำการสังเคราะห์เสียงขึ้น จากนั้นนำสัญญาณที่สังเคราะห์มาคูณกับเอนเวลลอป (Envelop) ของสัญญาณเสียงจริงในช่วงที่มีการเปลี่ยนแปลง จากการพิจารณาสเปกตรัมของเสียงสังเคราะห์ที่ได้เปรียบเทียบกับสเปกตรัมของสัญญาณเสียงจริงในช่วงดังกล่าว พบว่ามีความใกล้เคียงกัน ดังแสดงในรูปที่ 5



(ก) (ข)

รูปที่ 5 แสดงการเปรียบเทียบสเปกตรัมของโน้ตตัวซอลต้นแบบกับที่สังเคราะห์ขึ้นใหม่
ก.ส่วนที่เป็น attack ข. ส่วนที่เป็น decay

จากการวิเคราะห์ข้างต้น สรุปได้ว่าสัญญาณเสียงที่มีการเปลี่ยนแปลงหรืออยู่ในภาวะเสถียร จะมีส่วนประกอบของความถี่ที่เหมือนกัน เฉพาะเอนเวลลอปเท่านั้นที่มีการเปลี่ยนแปลง ดังนั้นในการสังเคราะห์เสียงในช่วงดังกล่าวจึงสามารถใช้พารามิเตอร์ชุดเดียวกับในภาวะเสถียรได้

4.1.2. การวิเคราะห์ในส่วนของสัญญาณรบกวน

หลังจากทำการวิเคราะห์ส่วนของสัญญาณคลื่นไซน์แล้ว การวิเคราะห์ในส่วนของสัญญาณรบกวนกระทำโดยกำหนดส่วนของสเปกตรัมที่มีค่าความชันสูงให้มีค่าเท่ากับศูนย์ จะได้สเปกตรัมในส่วนของสัญญาณรบกวน ในการวิเคราะห์ในส่วนนี้ได้แบ่งสัญญาณออกเป็นวินโดว์ย่อยๆ เพื่อหาสเปกตรัมเฉลี่ยของสัญญาณรบกวนซึ่งมีขนาด 500 แซมเปิล และแต่ละวินโดว์เหลื่อมกัน 250 แซมเปิล แสดงดังรูปที่ 6 จากนั้นนำสเปกตรัมที่ได้มาหาอินเวอร์สฟูเรียร์ทรานฟอร์มของฟังก์ชัน ตามสมการที่ 9 จะได้ค่าสัมประสิทธิ์ตัวกรองสัญญาณแสดงดังตารางที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการศึกษาเพื่อประโยชน์ทางการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โน้ต	สัญญาณไซน์ลำดับที่ / พารามิเตอร์	1	2	3	4	5	6	7	8	9	10
เสียงซอล	ค่าความถี่ (เฮิรตซ์)	666	1332	1998	2664	3330	3996	4662	5328	5994	6660
	ค่าแอมพลิจูด	1	0.1899	0.0515	0.0013	0.0027	0.0012	0.0014	0.0001	0.0002	0.0001
เสียงที	ค่าความถี่ (เฮิรตซ์)	836.4	1672.8	2509.2	3345.6	4181.9	5018.3	5854.7	6691.1	-	-
	ค่าแอมพลิจูด	1	0.951	0.0219	0.0013	0.0122	0.0005	0.0003	0.0004	-	-
เสียงโดสูง	ค่าความถี่ (เฮิรตซ์)	911.2	1822.3	2733.5	3644.8	4550	54604	6372.2	-	-	-
	ค่าแอมพลิจูด	1	0.1617	0.0169	0.0053	0.003	0.0003	0.0003	-	-	-

ตารางที่ 1 แสดงค่าพารามิเตอร์ของสัญญาณไซน์ที่ลำดับต่าง ๆ และค่าสัมประสิทธิ์ตัวกรองสัญญาณ

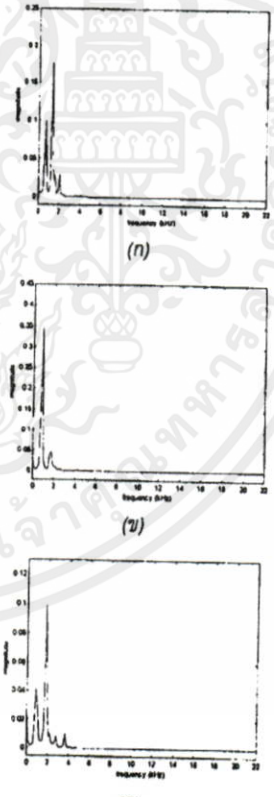
ขั้นตอนการสังเคราะห์เสียง

การสังเคราะห์เสียงแบ่งออกเป็นสองส่วนคือ ส่วนของการสร้างเสียงสัญญาณไซน์และส่วนของการสร้างเสียงที่มีลักษณะของสัญญาณรบกวน ในส่วนแรกทำได้โดยนำค่าพารามิเตอร์ที่ได้จากการวิเคราะห์ในส่วนของสัญญาณไซน์ มาสร้างสัญญาณส่วนของไซน์นูซอยด์ตามสมการที่ (1) แต่จากการพิจารณาจากสมการดังกล่าว จะพบว่ามีความซ้ำซ้อนในการคำนวณ เนื่องจากในการสร้างสัญญาณไซน์ที่ความถี่ต่างๆ ที่เป็นฮาร์โมนิกของความถี่พื้นฐาน จะมีค่า $\cos(\omega n)$ บางค่าที่เท่ากัน เช่น $\cos(\omega n)$ มีค่าเท่ากับ $\cos(\omega_2(n/2))$ หรือ $\cos(\omega_2 n)$ มีค่าเท่ากับ $\cos(\omega_4(n/2))$ ทั้งนี้ $\omega_2 = 2\omega$ และ $\omega_4 = 2\omega_2$ ตามลำดับ ดังนั้นการลดความซ้ำซ้อนโดยอาศัยความจริงดังกล่าวจะเป็นประโยชน์สำหรับงานวิจัยในอนาคตในการสร้างแอพพลิเคชันแบบ real-time ซึ่งมีความเป็นไปได้ที่จะลดความซ้ำซ้อนได้ถึง 17 เปอร์เซ็นต์

ในส่วนของสัญญาณรบกวน ใช้สมการที่ (3) สังเคราะห์สัญญาณรบกวน ทั้งนี้ได้ลดจำนวนสัมประสิทธิ์ตัวกรองสัญญาณลงเรื่อยๆ โดยกำหนดจำนวนขั้นต่ำที่ใช้ 3 ค่า และนำทั้งสองส่วนมารวมกันตามสมการที่ (2) จากนั้นทำการคูณด้วยเอนเวลลอป ในส่วนของ Attack และ Decay จะได้สัญญาณเสียงที่มีลักษณะคล้ายคลึงกับสัญญาณเสียงต้นแบบดังแสดงในรูปที่ 7

การประเมินระดับคุณภาพของเสียง

ในการวัดคุณภาพของเสียงกระทำโดยใช้ผู้ที่มีความเชี่ยวชาญทางด้านดนตรีไทยจำนวน 10 ท่าน ฟังเสียงสังเคราะห์เทียบกับเสียงต้นแบบทีละคลื่นเสียง โดยในครั้งแรกให้ฟังเสียงต้นแบบและทุกคลื่นที่สังเคราะห์ขึ้นก่อนสำหรับการให้คะแนนแบ่งออกเป็น5 ระดับ (โดยที่ 1 คือระดับคุณภาพเสียงต่ำ จนถึง 5



รูปที่ 6 ค่าเฉลี่ยสเปกตรัมของส่วนสัญญาณรบกวน
ก. ค่าเฉลี่ยสเปกตรัมของสัญญาณรบกวนในโน้ตซอล
ข. ค่าเฉลี่ยสเปกตรัมของสัญญาณรบกวนในโน้ตที
ค. ค่าเฉลี่ยสเปกตรัมของสัญญาณรบกวนในโน้ตโด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จำนวนสัมประสิทธิ์ตัวกรองสัญญาณ / โน้ต	1	2	3	4	5	6	7	8	9	10	11
เสียงซอล	0.0071	0.0096	0.0123	0.0149	0.0173	0.0219	0.0173	0.0149	0.0123	0.0096	0.0071
เสียงที	0.0086	0.0112	0.0140	0.0172	0.0203	0.0267	0.0203	0.0172	0.0140	0.0112	0.0086
เสียงโดสูง	0.0037	0.0066	0.0099	0.0133	0.0163	0.0216	0.0163	0.0133	0.0099	0.0066	0.0037

ตารางที่ 2 แสดงค่าสัมประสิทธิ์ตัวกรองสัญญาณ

ระดับคุณภาพเสียงดีมาก) ทั้งนี้การประเมินนั้นขึ้นอยู่กับประสบการณ์และความพึงพอใจของผู้เข้ามาทดสอบด้วย ซึ่งจากผลของการพิจารณาคุณภาพของเสียงพบว่าจำนวนสัมประสิทธิ์ตัวกรองสัญญาณในช่วง 3-99 ค่า ให้เสียงสังเคราะห์ที่มีคุณภาพอยู่ในระดับคะแนนที่ 4

ข้อดีของการสังเคราะห์เสียงเชิงสเปกตรัม

แนวทางการสังเคราะห์ดังกล่าวมีข้อได้เปรียบเหนือแนวทางการสังเคราะห์เสียงอื่น โดยเฉพาะในเรื่องของความยืดหยุ่นในการสร้างเสียงแปลกใหม่ซินทีไซน์เซอร์โดยทั่วไปจะใช้วิธีการแซมปลิง หรือที่เรียกว่าเวปเทเปิล (Wavetable) ซึ่งไม่สามารถสังเคราะห์เสียงที่แปลกออกไปได้เช่นการทำเทคนิคออดิโอมอร์ฟิง (Audio Morphing) แต่ในกรณีของการสังเคราะห์เสียงเชิงสเปกตรัม สามารถกระทำได้โดยการเปลี่ยนแปลงองค์ประกอบเชิงความถี่ ในที่นี้ได้เสนอวิธีการดัดแปลงเสียงไว้ดังนี้

1. ปรับเปลี่ยนค่าแอมพลิจูดในบางค่าความถี่หรือทุกค่าความถี่
2. ใช้ค่าความถี่เพียงบางค่าร่วมกับการปรับเปลี่ยนค่าแอมพลิจูดเดิม
3. ใช้ค่าความถี่เพียงบางค่าแต่ยังคงค่าแอมพลิจูดเดิม

จากวิธีการดังกล่าวหรือสร้างเสียงสังเคราะห์จากการผสมผสานทางด้านความถี่ของเสียงดนตรีทั้งสองชนิด ซึ่งสามารถกระทำได้ดังนี้

1. ใช้ค่าเฉลี่ยค่าความถี่ของเสียงดนตรีทั้งสองพร้อมทั้งค่าแอมพลิจูดของเสียงใดเสียงหนึ่ง
2. ใช้ค่าเฉลี่ยของค่าแอมพลิจูดร่วมกับค่าความถี่ของเสียงดนตรีชนิดใดชนิดหนึ่งเป็นหลัก
3. ใช้ค่าเฉลี่ยของค่าแอมพลิจูดและค่าความถี่เพื่อใช้ในการสังเคราะห์เสียง

จากการทดลองพบว่า เสียงสังเคราะห์ที่เกิดจากการผสมผสานระหว่างขลุ่ยเพียงออกับขลุ่ยหลิบ โดยมีผู้เข้ารับการทดสอบซึ่งทักษะทางด้านดนตรี ให้ความเห็นว่า เสียงสังเคราะห์ที่ได้ยังคงลักษณะเสียงดนตรีแต่ไม่ใช่เสียงขลุ่ยเพียงออหรือขลุ่ยหลิบ

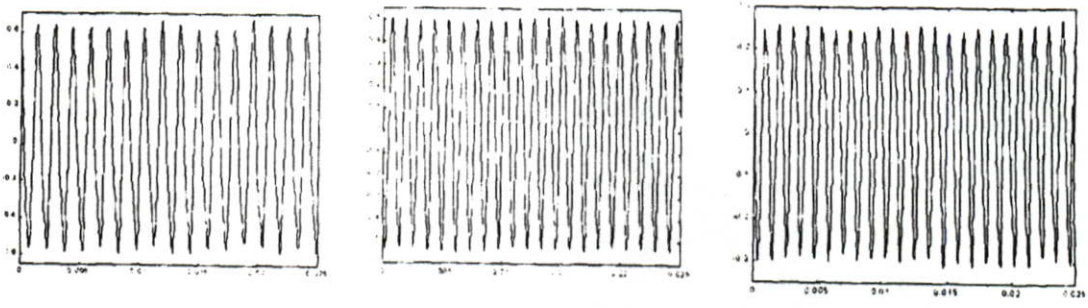
ข้อดีอีกประการหนึ่งของการสังเคราะห์เสียงด้วยวิธีการดังกล่าว คือสามารถสร้างเสียงสังเคราะห์ที่มีระดับเสียงสูงต่ำของตัวโน้ตเท่ากับเสียงดนตรีสากล ซึ่งกระทำได้โดยปรับค่าความถี่ให้มีค่าเท่ากับความถี่มาตรฐานสากล ตลอดจนสามารถสร้างเสียงสังเคราะห์ที่เป็นครึ่งเสียง ทำให้มีคุณสมบัติพิเศษในการเล่นเสียงดนตรีขลุ่ยไทยร่วมกับดนตรีสากลได้

นอกจากนี้แล้ว เมื่อเปรียบเทียบกับวิธีการแซมปลิงแล้ว วิธีการนี้มีข้อดีในแง่ของปริมาณข้อมูลที่ใช้ในการสังเคราะห์เสียงน้อยกว่า แต่ให้ระดับคุณภาพเสียงที่เท่ากันอีกด้วย

บทสรุป

บทความนี้ได้นำเสนอเทคนิคการสังเคราะห์เสียงด้วยวิธีการ Sinusoidal plus noise model ซึ่งการสังเคราะห์เสียงแนวทางนี้มีข้อดีคือเป็นโมเดลที่สามารถใช้ได้กับเครื่องดนตรีทุกประเภท เพียงแต่เปลี่ยนค่าพารามิเตอร์ที่ได้จากการวิเคราะห์สเปกตรัมของสัญญาณเสียงแต่ละชนิด จากการวิเคราะห์พบว่าเสียงขลุ่ยไทยประกอบด้วยความถี่พื้นฐานและความถี่ฮาร์โมนิกจำนวน 5-12 ความถี่ และในส่วนของการสังเคราะห์สัญญาณโดยคำนึงถึงความซับซ้อนในการคำนวณ พบว่าต้องใช้ค่าสัมประสิทธิ์ตัวกรองสัญญาณจำนวน 3 ค่าขึ้นไป (เมื่อนำมารวม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก) (ข) (ค)

รูปที่ 7 แสดงลักษณะคลื่นสัญญาณเสียงสังเคราะห์ ก.โน้ตตัวซอล ข.โน้ตตัวที ค.โน้ตตัวโดสูง

กับส่วนของสัญญาณคลื่นไซน์) จึงจะได้เสียงสังเคราะห์ที่ยังคงลักษณะเสียงขลุ่ย จากการศึกษาพบว่าโมเดลดังกล่าวมีความเหมาะสมที่จะใช้เสียงขลุ่ยไทย เนื่องจากใช้จำนวนพารามิเตอร์เพียงเล็กน้อยในการแทนเสียงเครื่องดนตรีดังกล่าว นอกจากนี้แล้วยังมีความเป็นไปได้ในการประยุกต์ใช้โมเดลดังกล่าวเพื่อการวิเคราะห์เปรียบเทียบคุณภาพของเสียงดนตรีจากเครื่องดนตรีจริงได้อีกด้วย

เอกสารอ้างอิง

[1] Nomer, D. "Musical Physical Modeling", [Online]. Available: http://web800.com/music/vl/physmodl.htm. 1998

[2] Serra, X. "Current Perspectives in the Digital Synthesis of Musical Sounds", Pompeau Fabra University. [Online]. Available:http://www.iua.upf.es/~xserra/articles/current. 1997.

[3] Smith, J.O. "Viewpoints on the History of Digital Synthesis", Proceedings of the 1991 International Computer Music Conference. Montreal, San Fracisco, 1-10 Oct. 1991.

[4] Music Technology Group. "Spectral Modeling Synthesis", [Online]. Available: http://www.iua.upf.es/~sms/. 1999

[5] Serra, X. "Integrating Complementary Spectral Models in the Design of a Musical Synthesizer", The Proceeding of the International Computer Music Conference , 1997.

[6] Rodet, X. "Musical Sound Signal Analysis/ Synthesis: Sinusoidal+Residual and Elementary Waveform Models", Paris, France, 1997.

[7] Serra, X. "Musical Sound Modeling With Sinusoidal Plus Noise.", Musical Signal Processing. Swets& Zeitlinger Publishers,1997.

[8] Peevers, A.W. "A Real Time 3D Signal Analysis/ Synthesis Tool Based on the Short Time Fourier Transform", [Online]. Available: http://cnmat.CNMAT.Berkeley.EDU/~alan/MS-html/MSv2_ToC.html. 1999

[9] Steiglit, K. A Digital Signal Processing Primer with Applications to Digital Audio and Computer Music.Menlo Park, CA : Addison-Wesley, 1996.

[10]Siricharoenchai, R. and N. Chotikakamthorn. "Thai Flute Sound Synthesis Using Sinusoidal Plus Noise Model." The 1999 National Computer Science and Engineering Conference, Bangkok, Thailand.

กิตติกรรมประกาศ

ขอขอบพระคุณ อ. ธนิสร์ ศรีกลิ่นดี ที่กรุณาเล่นขลุ่ยเพียงออเพื่อการบันทึกเก็บเป็นข้อมูลในการทำวิจัย และนางสาวสุวรรณี กิ่งเพชร ที่ให้การบันทึกเสียงขลุ่ยประกอบการศึกษาลักษณะสัญญาณเสียงในเบื้องต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thai Flute Sound Synthesis Using Sinusoidal Plus Noise Model

Rungkarn Siricharoenchai and Nopporn Chotikakamthorn
Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang,
Bangkok 10520, Thailand
E-mail: s0067017@kmitl.ac.th , nopporn@it.kmitl.ac.th

Abstract: This paper presents research result on Thai flute musical sound synthesis based on spectral synthesis approach, using a sinusoidal plus noise model. We analyzed recorded sound of Thai flute, only steady-state part, to estimate the values of parameters required for sound synthesis. Such parameters include frequency ,amplitude and filter coefficients. From the analysis, we found that the sound of Thai flute is composed of 5-12 principal harmonic, depending on the pitch of sound analyzed. From the sound quality listening test by general listeners, it has been found that 50 filter coefficients in generating noise part of the synthesized sound gives synthesized sound which is indifferent from the original one.
Key words:Sound synthesis, Sinusoidal plus noise model, Spectral estimation, Thai flute

1.บทนำ

แนวทางในการสังเคราะห์เสียงดนตรีด้วยคอมพิวเตอร์มีหลายวิธี เช่น วิธีการที่อาศัยการจัดเก็บเสียงในรูปของคลื่นเสียง(waveform) หรือที่เรียกกันว่าวิธีการแซมปลิง (sampling) วิธีการสังเคราะห์เสียงดังกล่าวให้เสียงที่มีคุณภาพดีแต่การดัดแปลงแก้ไขเสียงดนตรีทำได้ไม่มาก นอกจากวิธีการดังกล่าวแล้ว ในปัจจุบันได้มีการพัฒนาเทคนิคการสังเคราะห์เสียงใหม่ๆ หลายวิธี เช่น การสังเคราะห์เสียงเชิงกายภาพ ซึ่งอาศัยลักษณะปรากฏการณ์ทางกายภาพของการเกิดเสียงเครื่องดนตรีชนิดนั้นๆมาเป็นแบบจำลองในการสังเคราะห์เสียง[1] พารามิเตอร์ที่ใช้ในการควบคุมจึงเหมือนของเครื่องดนตรีจริงและเสียงสังเคราะห์จะมีความเป็นธรรมชาติเหมือนกับเสียงที่ได้จากเครื่องดนตรีจริง แต่อย่างไรก็ดีการสังเคราะห์เสียงแนวทางนี้มีข้อจำกัดคือ ในการนำไปประยุกต์ใช้กับเสียงเครื่องดนตรีแต่ละชนิด จำเป็นต้องสร้างแบบจำลองเฉพาะสำหรับเครื่องดนตรีนั้นๆ และยังมีข้อจำกัดในด้านความยืดหยุ่นในการดัดแปลงเสียง นอกเหนือจากการสังเคราะห์เสียงเชิงกายภาพแล้ว เทคนิคการสังเคราะห์เสียงอีกวิธีหนึ่งที่ได้รับการพัฒนาขึ้นมาเมื่อไม่นานมานี้ คือ วิธีการสังเคราะห์เสียงเชิงสเปกตรัม ซึ่งมีข้อดีคือสามารถสังเคราะห์เสียงได้หลายประเภทไม่จำกัดเฉพาะเสียงเครื่องดนตรี[6],[7] และมีความยืดหยุ่นสูงกว่าในการดัดแปลงหรือสร้างเสียงได้ นอกจากนี้โมเดลในการสังเคราะห์เสียงสามารถนำไปประยุกต์ใช้กับเสียงดนตรีใดๆก็ได้ งานวิจัยที่เสนอในบทความนี้จึงอาศัยแนวทางการสังเคราะห์เชิงสเปกตรัมในการวิเคราะห์และสังเคราะห์เสียงเครื่องดนตรีจุ่ยไทย โดยในการวิจัยผู้วิจัยได้ทำการวิเคราะห์ส่วนประกอบเชิงความถี่ของเสียงขลุ่ยไทยเพื่อหาค่าพารามิเตอร์สำหรับใช้ในการสังเคราะห์เสียง ตลอดจนได้ทำการทดลองทดสอบคุณภาพของเสียงสังเคราะห์ที่ได้

2.การสังเคราะห์เสียงด้วยแบบจำลองเชิงสเปกตรัม (Spectral Modeling)

แนวทางการสังเคราะห์เสียงเชิงสเปกตรัมเริ่มด้วยการวิเคราะห์สเปกตรัมของเสียงเพื่อหาค่าพารามิเตอร์สำหรับใช้ในของการสังเคราะห์เสียง โดยค่าพารามิเตอร์ต่างๆ เช่น ค่าความถี่ แอมปลิจูดและเฟส จะถูกใช้ในการสร้างสัญญาณเสียงสังเคราะห์ แนวทางการสังเคราะห์เสียงเชิงสเปกตรัมที่สำคัญมี 2 แนวทาง [5] ดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1. Sinusoidal Model

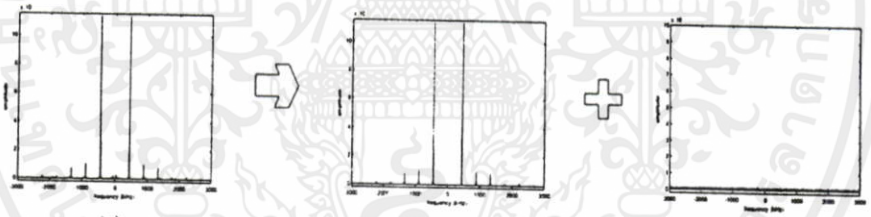
รูปแบบการจำลองแบบ Siinusoidal เป็นวิธีการที่จำลองเสียงดนตรีด้วยผลรวมของสัญญาณไซน์ที่ความถี่ต่างๆ ดังแสดงในสมการต่อไปนี้

$$s(n) = \sum_{r=1}^{R} A_r \cos(\omega_r n + \theta_r) \qquad (1)$$

โดยที่ $s(n)$ คือสัญญาณที่ได้ A_r คือค่าแอมปลิจูดของสัญญาณลำดับที่ r ω_r คือค่าความถี่เชิงมุมของสัญญาณไซน์สำดับที่ r θ_r คือค่าเฟสของสัญญาณไซน์ลำดับที่ r n คือลำดับที่ของจุดข้อมูล R คือจำนวนสัญญาณไซน์ที่พิจารณาทั้งหมด

2.2 Sinusoidal Plus Noise Model

การสังเคราะห์เสียงภายใต้โมเดล Sinusoidal Plus Noise Model อยู่บนพื้นฐานของการจำลองเสียงดนตรีด้วยผลรวมขององค์ประกอบสัญญาณ 2 ส่วน [3],[4] คือ ส่วนที่เป็นสัญญาณคลื่นไซน์ และส่วนที่เป็นสัญญาณรบกวน (Noise)[1] ดังแสดงในรูปที่ 1



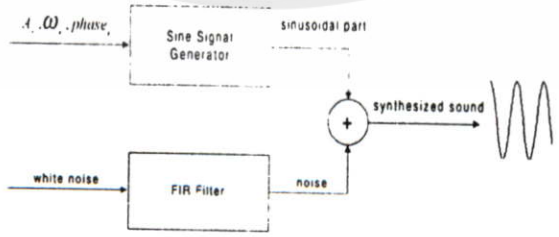
(ก) (ข) (ค)

รูปที่ 1 ตัวอย่างแสดงการแยกองค์ประกอบของสัญญาณเสียงออกเป็นสองส่วน
ก. สเปกตรัมของสัญญาณเสียงที่ทำการวิเคราะห์ ข. สเปกตรัมของส่วนที่เป็นสัญญาณคลื่นไซน์
ค. สเปกตรัมของส่วนที่เป็นสัญญาณเสียงรบกวน (noise)

วิธีการสังเคราะห์เสียงในแนวทางนี้ประกอบไปด้วย 2 ส่วน (ดูรูปที่ 2 ประกอบ) คือ ส่วนของสัญญาณคลื่นไซน์ ซึ่งขั้นตอนในการสร้างสัญญาณส่วนนี้ จะกระทำในลักษณะเดียวกันกับวิธีการ Sinusoidal Model ตามสมการที่ (1) และส่วนของสัญญาณรบกวน ซึ่งสามารถสังเคราะห์ได้โดยการกรองสัญญาณ white noise ด้วยฟิลเตอร์ที่เหมาะสม ด้วยแบบจำลองดังกล่าวเสียงดนตรีประกอบด้วยส่วนต่างๆ ดังสมการต่อไปนี้

$$s(n) = \sum_{r=1}^{R} A_r \cos(\omega_r n + \theta_r) + e(n) \qquad (2)$$

โดยที่ $s(n)$ คือสัญญาณที่ได้ (output signal) $A_r(n)$ คือค่าแอมปลิจูดของสัญญาณไซน์ลำดับที่ r ω_r คือค่าความถี่เชิงมุมของสัญญาณไซน์สำดับที่ r $\theta_r(n)$ คือค่าเฟสของสัญญาณไซน์ลำดับที่ r R คือจำนวนสัญญาณไซน์ที่พิจารณาทั้งหมด $e(n)$ คือ ส่วนของสัญญาณรบกวน ซึ่งสามารถคำนวณหาได้จากสมการต่อไปนี้



รูปที่ 2 ไดอะแกรมแสดงขั้นตอนการสังเคราะห์เสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[1] สัญญาณรบกวนในที่นี้หมายถึง เสียงที่มีลักษณะของสัญญาณที่ไม่สามารถจำลองได้ด้วยฟังก์ชันไซน์ หากแต่สามารถอธิบายได้ด้วยคุณสมบัติทางสถิติ มิได้หมายถึง เสียงหรือสัญญาณที่ไม่พึงปรารถนา

$$e(n)=\sum_{k=0}^{K-1}h(k)w(n-k) \qquad (3)$$

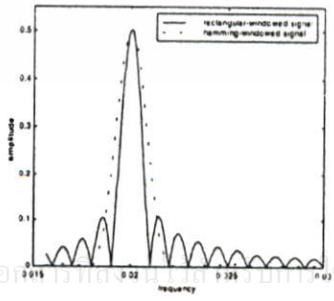
โดยที่ $h(k)$ คือสัมประสิทธิ์ตัวกรองสัญญาณ (filter coefficient) $w(n)$ คือสัญญาณ white noise และ K คือจำนวนสัมประสิทธิ์ตัวกรองสัญญาณ

จากสมการข้างต้นพบว่ามีค่าพารามิเตอร์ที่ต้องใช้ในการสังเคราะห์ คือ ค่าแอมปลิจูด ค่าความถี่ ค่าเฟส และค่าสัมประสิทธิ์ตัวกรองสัญญาณ ซึ่งค่าพารามิเตอร์เหล่านี้สามารถหาได้จากการวิเคราะห์เสียงดนตรีที่เป็นสัญญาณจริง โดยขั้นตอนการวิเคราะห์มีดังต่อไปนี้

2.2.1.การวิเคราะห์ในส่วนของสัญญาณคลื่นไซน์

ในการวิเคราะห์เชิงสเปกตรัม เพื่อให้ได้ค่าความถี่และแอมปลิจูดที่ถูกต้อง ชนิดของวินโดว์ ขนาดของวินโดว์ และการเพิ่มค่าศูนย์ (zero padding) มีความสำคัญ

- ชนิดของวินโดว์ มีผลต่อการให้ค่าความถี่หลักและความถี่ฮาร์โมนิกที่ถูกต้องมากขึ้น เนื่องจากโดยลักษณะสเปกตรัมของสัญญาณเสียงขลุ่ยไทยนั้นมีช่วงห่างระหว่างความถี่พอสมควร ฉะนั้นการใช้
- แฮมมิ่งวินโดว์ (hamming window) น่าจะให้ความถูกต้องมากกว่าเรคแทงคิวลาร์วินโดว์ (rectangular window) ซึ่งผลการทดลองเป็นไปตามสมมติฐานดังกล่าว นอกจากนี้ความถูกต้องในการหาค่าความถี่หลักยังทำให้การวิเคราะห์หาค่าพารามิเตอร์ในส่วนของสัญญาณรบกวนมีความถูกต้องมากขึ้นอีกด้วย ตัวอย่างสเปกตรัมที่ได้จากการใช้วินโดว์ทั้ง 2 ประเภทแสดงในรูปที่ 3



รูปที่ 3 เปรียบเทียบสเปกตรัมของสัญญาณที่คูณกับ rectangular window และ hamming window

- ขนาดของวินโดว์ หากใช้ขนาดเหมาะสม จะทำให้ค่าความถี่หลักและความถี่ฮาร์โมนิกชัดเจนมากขึ้น ขนาดของวินโดว์เริ่มต้นควรเป็นไปตามสมการข้างล่างนี้ [2]

$$M \geq \frac{B \cdot f_s}{f_o} \qquad (4)$$

โดยที่ M คือขนาดของวินโดว์ B คือ mainlobe ของวินโดว์ที่ใช้ f_s คือความถี่แซมเปิล f_o คือความถี่พื้นฐาน (fundamental frequency)

- การเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์ เข้ากับสัญญาณเสียง (Zero padding) โดยทั่วไปมีผลทำให้การวิเคราะห์มีความถูกต้องมากขึ้น แต่อย่างไรก็ดี จากการทดลองนี้เราพบว่าการเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์ อาจทำให้มีความคลาดเคลื่อนมากขึ้นได้ ทั้งนี้ขึ้นอยู่กับจำนวนจุดของข้อมูลที่มีค่าเป็นศูนย์ที่เพิ่มเข้ามา จากการวิเคราะห์และผลการทดลองพบว่า หากเพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์ตามสมการต่อไปนี้

$$s=2^t(N+x) \qquad (5)$$

โดยที่ s คือจำนวนแซมเปิลที่เป็นไปได้ที่ทำให้ค่าที่คำนวณได้มีความถูกต้องมากขึ้น N คือจำนวนแซมเปิลของสัญญาณที่พิจารณา x คือ จำนวนค่าศูนย์ที่เพิ่มให้กับสัญญาณ k คือ 0,1,2,3,... ,จะทำให้ค่าพารามิเตอร์ (เช่น ค่าความถี่ และแอมปลิจูด) มีความถูกต้องมากขึ้นตามจำนวนจุดที่เพิ่มขึ้น

ขั้นตอนการวิเคราะห์ในส่วนของสัญญาณไซน์ มีรายละเอียดดังนี้

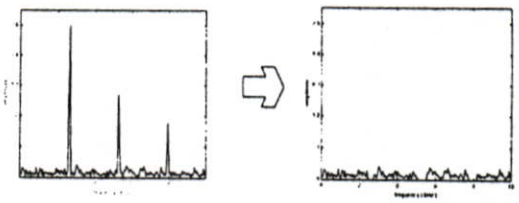
1.นำส่วนของสัญญาณที่ต้องการวิเคราะห์จำนวน M แซมเปิล (โดยจำนวน M ต้องเป็นไปตามสมการ 4) คูณกับแฮมมิ่งวินโดว์

$$s_w(n)=s(n).\,w(n) \quad , n = 0,...,M-1 \qquad (6)$$

โดยที่ $s_w(n)$ คือสัญญาณที่คูณกับแฮมมิ่งวินโดว์ $s(n)$ คือสัญญาณขนาด M แซมเปิล $w(n)$ คือแฮมมิ่งวินโดว์ที่มีขนาด M เท่ากัน

2.เพิ่มจุดข้อมูลที่มีค่าเป็นศูนย์ให้กับสัญญาณตามสมการที่ (5)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก) (ข)

รูปที่ 4 ตัวอย่างแสดงสเปกตรัมของสัญญาณ ก. สเปกตรัมของสัญญาณเสียงที่ต้องการวิเคราะห์ ข. สเปกตรัมของสัญญาณเสียงที่ตัดส่วนของสัญญาณไซน์ออกแล้ว

3.คำนวณค่าสเปกตรัมของสัญญาณ เช่นโดยอาศัยอัลกอลิธึ่ม FFT

4.จากสเปกตรัมที่ได้ ทำการตรวจหาจุดที่มีค่าความชันสูง (Peak) เพื่อคำนวณค่าความถี่ แอมปลิจูดและเฟส ที่จุดนั้นๆสำหรับใช้ในกระบวนการสังเคราะห์เสียง

2.2.2.การวิเคราะห์ในส่วนของสัญญาณรบกวน

การวิเคราะห์ในส่วนของสัญญาณรบกวนมีขั้นตอนดังนี้

1.จากสเปกตรัมของสัญญาณที่วิเคราะห์ได้จากขั้นแรก เราจะได้สเปกตรัมของส่วนสัญญาณรบกวนโดยกำหนดค่าในช่วงของสเปกตรัมที่มีค่าความชันสูงตามที่ตรวจหาจากขั้นตอนที่ 4 ในการวิเคราะห์ส่วนของสัญญาณไซน์ ให้มีค่าเป็นศูนย์ (ดูรูปที่ 4 ประกอบ)

2.สร้างสัญญาณเสียงของส่วนที่เป็นสัญญาณรบกวน โดยการหาอินเวอร์สฟูเรียร์ทรานฟอร์ม (Inverse Fourier Transform) ของสเปกตรัมที่ได้ในข้อ 1

3.คำนวณหาค่าสหสัมพันธ์ (correlation) ของสัญญาณรบกวน ที่ได้จากข้อ 2 ดังสมการต่อไปนี้

$$c_{xx}(l)=\frac{1}{N}\sum_{n=1}^{N-1}x(n)x(n-1) \qquad (7)$$

โดยที่ $c_{xx}(l)$ คือค่าสหสัมพันธ์ และ l คือค่า time lag

4.หาค่าสัมประสิทธิ์ของตัวกรองสัญญาณ $h(n)$ ซึ่ง n=0...,n-1 โดยการทำอินเวอร์สฟูเรียร์ทรานฟอร์มของฟังก์ชัน $H(\omega)$ โดยที่

$$H(\omega)=\sqrt{C(\omega)} \qquad (8)$$

$C(\omega)$ ในสมการที่ (8) หาได้จากการทำฟูเรียร์ทรานฟอร์ม ฟังชันก์สหสัมพันธ์ $c_{xx}(l)$ โดยค่า l ที่พิจารณาจะมีค่าตั้งแต่ $-L$ ถึง L โดยเลือกค่า L ที่ทำให้ $c_{xx}(l)$, $l > L$, มีค่าใกล้ศูนย์

2.2.3.การสังเคราะห์เสียง

หลังจากที่ได้วิเคราะห์สัญญาณเสียงทั้ง 2 ส่วน (ดังแสดงในรายละเอียด 2.2.1 และ 2.2.2) จะได้ค่าพารามิเตอร์ที่สามารถนำมาใช้ในขั้นตอนการสังเคราะห์เสียง ซึ่งมีรายละเอียดดังต่อไปนี้

1.สร้างสัญญาณคลื่นไซน์ตามสมการที่ (1) โดยใช้ค่าพารามิเตอร์ที่หาได้จากขั้นตอนการวิเคราะห์ในส่วนของสัญญาณไซน์

2.สร้างสัญญาณรบกวนตามสมการที่(3)โดยใช้สัมประสิทธิ์ตัวกรองสัญญาณที่ได้จากการวิเคราะห์ในส่วนของสัญญาณรบกวน

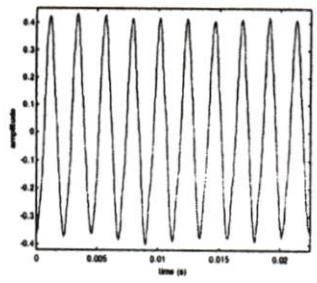
3.นำสัญญาณที่สร้างขึ้นใน 2 ขั้นตอนข้างต้นมารวมกันเป็นสัญญาณเสียงสังเคราะห์ที่ต้องการ

3.ผลการทดลอง

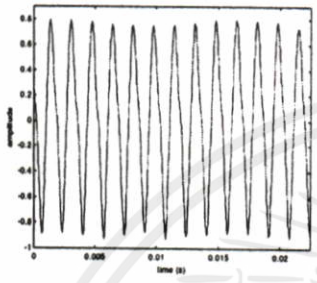
ในการทดลองสังเคราะห์เสียงขลุ่ย ผู้วิจัยได้ทำการบันทึกเสียงขลุ่ยเพียงออเก็บเป็นไฟล์เวป (wave files) ที่อัตราแซมปลิง (sampling rate) 44100 เฮิร์ท และใช้การจัดเก็บแบบ 16 บิตต่อหนึ่งแซมเปิล โดยในบทความนี้จะยกตัวอย่างการพิจารณาเสียงโน้ตตัวโด ฟา ลา ในช่วงภาวะเสถียรซึ่งมีความยาวประมาณ 1.5 วินาที ซึ่งแต่ละตัวโน้ตใช้จำนวน 5 สัญญาณคลื่นเสียง[2] ตัวอย่างของสัญญาณเสียงแสดงดังรูปที่ 5 หลังจากที่ได้นำสัญญาณดังกล่าวมาวิเคราะห์ เพื่อหาส่วนประกอบของสัญญาณไซน์

[2]จำนวนข้อมูล 5 ชุดโดยทำการบันทึก 5 ครั้ง

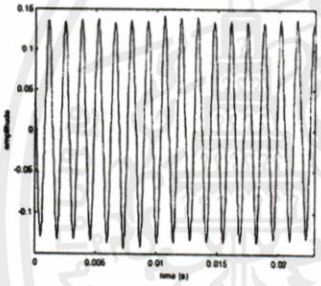
เอกสารนี้ถูกจัดทำขึ้นเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)



(ข)



(ค)

รูปที่ 5 คลื่นเสียงของโน้ต
ก. โน้ตตัวโด ข. โน้ตตัวฟา ค. โน้ตตัวลา

ที่ฮาร์โมนิกต่างๆและคำนวณหาค่าสัมประสิทธิ์ตัวกรองสัญญาณ ปรากฏว่าพารามิเตอร์ในส่วนของสัญญาณไซน์และสัญญาณรบกวนของ 5 สัญญาณมีค่าใกล้เคียงกัน ดังนั้นจึงเลือกแสดงเพียงสัญญาณเดียว

3.1.ขั้นตอนการวิเคราะห์

เริ่มจากการวิเคราะห์ในส่วนของสัญญาณไซน์ก่อน เราจะได้สเปกตรัมของสัญญาณซึ่งนำมาใช้ในการวิเคราะห์ส่วนของสัญญาณรบกวนต่อไป โดยมีรายละเอียดดังนี้

(1) การวิเคราะห์ในส่วนของสัญญาณไซน์

แบ่งสัญญาณออกเป็นวินโดว์ย่อยๆ เพื่อวิเคราะห์ดูการเปลี่ยนแปลงของสเปกตรัมในแต่ละวินโดว์ โดยกำหนดขนาดของวินโดว์เริ่มต้น 250 500 1000 2000 และ 4000 แซมเปิล ปรากฏว่าสเปกตรัมของสัญญาณที่ได้ในแต่ละวินโดว์ย่อยมีค่าใกล้เคียงกันมาก จากการวัดเปอร์เซ็นต์ความคลาดเคลื่อนของความถี่และแอมปลิจูดในแต่ละวินโดว์ พบว่าค่าเฉลี่ยอยู่ในระดับที่ยอมรับได้ เนื่องจากขนาดวินโดว์ที่มีความยาวมากจะให้สเปกตรัมของสัญญาณได้ชัดเจนมากที่สุด ในการทดลองนี้จึงเลือกใช้วินโดว์ที่มีขนาดเท่ากับ 50000 แซมเปิลในการวิเคราะห์เพื่อหาค่าพารามิเตอร์ต่างๆซึ่งแสดงดังตารางที่ 1 (อนึ่งจากการทดลองเปลี่ยนแปลงค่าเฟสในการสังเคราะห์เสียง พบว่า ค่าเฟสไม่มีผลต่อการสังเคราะห์เสียง จึงมิได้นำค่าเฟสที่วิเคราะห์ได้มาใช้ในการสังเคราะห์)

(2) การวิเคราะห์ในส่วนของสัญญาณรบกวน

หลังจากการทำการวิเคราะห์ส่วนของสัญญาณคลื่นไซน์แล้ว การวิเคราะห์ในส่วนของสัญญาณรบกวนกระทำโดยกำหนดส่วนของสเปกตรัมที่มีค่าความชันสูงให้มีค่าเท่ากับศูนย์ จะได้สเปกตรัมในส่วนของสัญญาณรบกวน ในการวิเคราะห์ส่วนนี้ ได้แบ่งสัญญาณออกเป็นเฟรมย่อยๆ เพื่อพิจารณาดูความเปลี่ยนแปลงในคุณสมบัติทางสถิติในแต่ละช่วงของสัญญาณ จากรูปที่ 6 พบว่า ค่าสหสัมพันธ์มีความคล้ายคลึงกันตลอดช่วงที่ศึกษา ดังนั้น ในการคำนวณค่าสัมประสิทธิ์ตัวกรองสัญญาณ จึงได้ใช้จำนวนแซมเปิลทั้งหมดของสัญญาณ (50000 แซมเปิล) เพื่อให้ได้ค่าทางสถิติที่ถูกต้องที่สุด

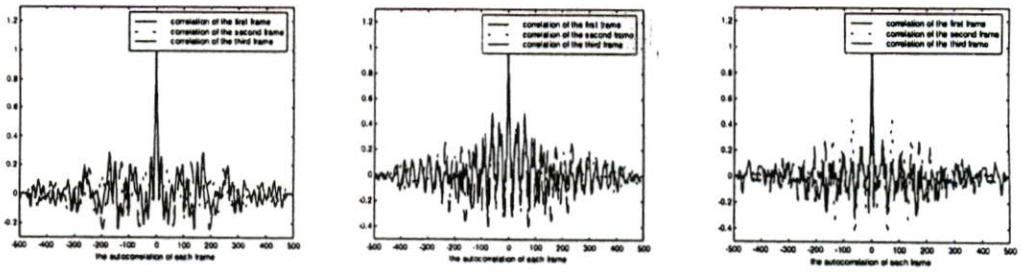
3.2.ขั้นตอนการสังเคราะห์เสียง

นำค่าพารามิเตอร์ที่ได้จากการวิเคราะห์ มาสร้างสัญญาณเสียงในส่วนของสัญญาณไซน์ตามสมการที่ (1) และส่วนของสัญญาณรบกวนตามสมการที่ (3) สังเคราะห์สัญญาณรบกวน ทั้งนี้ได้เลือกจำนวนสัมประสิทธิ์ตัวกรองสัญญาณที่ 25 51 75 และ 129 จากนั้นนำทั้งสองส่วนมารวมกันตามสมการที่ (2) เป็นสัญญาณเสียงสังเคราะห์ที่ต้องการ ดังตัวอย่างแสดงในรูปที่ 7

3.3.การประเมินระดับคุณภาพของเสียง

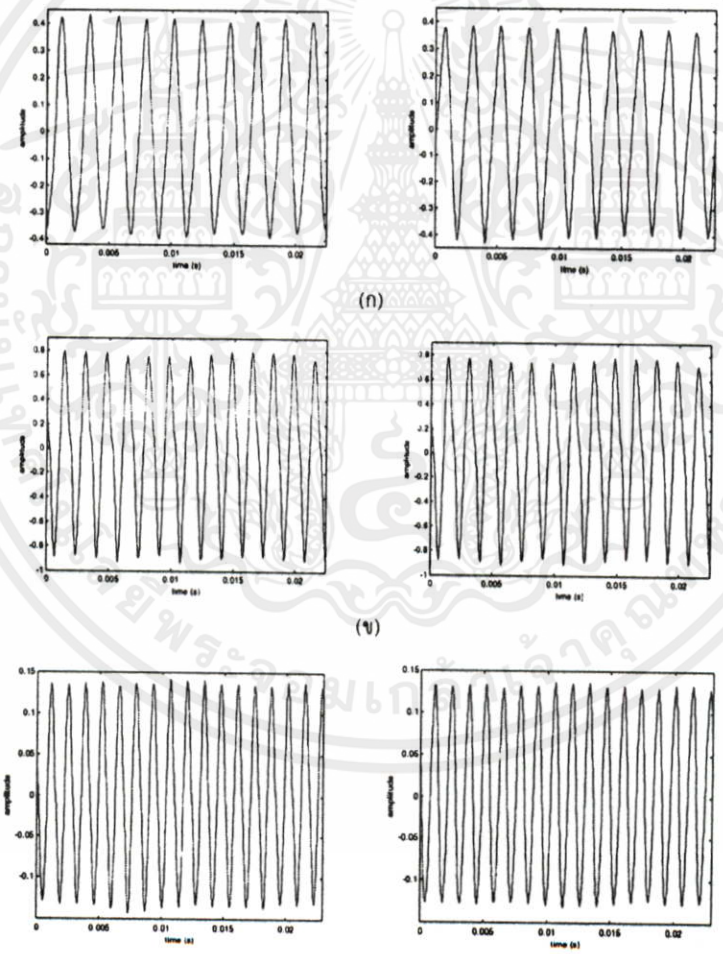
ในการวัดคุณภาพของเสียงกระทำโดยให้ผู้เข้ามาทดสอบซึ่งเป็นบุคคลทั่วไป ฟังเสียงสังเคราะห์เทียบกับเสียงต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก) (ข) (ค)

รูปที่ 6 ค่าสหสัมพันธ์ในแต่ละช่วงของสัญญาณรบกวน
ก. ค่าสหสัมพันธ์ของสัญญาณรบกวนในโน้ตโด ข.ค่าสหสัมพันธ์ของสัญญาณรบกวนในโน้ตฟา
ค. ค่าสหสัมพันธ์ของสัญญาณรบกวนในโน้ตลา



(ก)

(ข)

(ค)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 7 แสดงการเปรียบเทียบสัญญาณเสียงต้นแบบ (ซ้าย) สัญญาณเสียงสังเคราะห์ (ขวา)
ก. โน้ตตัวโด ข. โน้ตตัวฟา ค. โน้ตตัวลา

ตารางที่ 1 ค่าพารามิเตอร์ที่สัญญาณไซน์ที่ลำดับต่างๆ

โน้ต	สัญญาณไซน์ลำดับที่ / พารามิเตอร์	1	2	3	4	5	6	7	8	9	10	11	12
เสียงโด	ค่าความถี่ (เฮิร์ท)	445.6	891.9	1338	1783.6	2230.1	2676.6	3123.2	3572.1	4015.7	4465.3	4967.4	5366.1
	ค่าแอมปลิจูด	1	0.0965	0.0591	0.0096	0.0115	0.0047	0.0044	0.0065	0.0006	0.0006	0.0004	0.0003
เสียงฟา	ค่าความถี่ (เฮิร์ท)	599.5	1200	1796.9	2395.3	2999	3598.6	4192.6	4783.4	5394.1	5985.8	6606.6	-
	ค่าแอมปลิจูด	1	0.1891	0.0443	0.0041	0.0034	0.004	0.0015	0.0004	0.0003	0.0002	0.0002	-
เสียงลา	ค่าความถี่ (เฮิร์ท)	738.7	1477.4	2215.6	2955.1	3694.3	4432.9	5170.7	5894.4	6676.6	7420.3	-	-
	ค่าแอมปลิจูด	1	0.0643	0.0294	0.0054	0.0097	0.001	0.0006	0.0003	0.0003	0.0002	-	-

แบบทีละคลื่นเสียง โดยในครั้งแรกจะฟังเสียงต้นแบบและทุกคลื่นเสียงที่สังเคราะห์ขึ้นก่อน สำหรับการให้คะแนนนั้นแบ่งออกเป็น 5 ระดับ (โดยที่ 1 คือระดับคุณภาพเสียงต่ำ จนถึง 5 ระดับคุณภาพเสียงดีมาก) จากผลของการพิจารณาคุณภาพของเสียง พบว่าที่จำนวนสัมประสิทธิ์ตัวกรองสัญญาณเท่ากับ 51 ค่า ให้เสียงที่มีคุณภาพอยู่ในระดับคะแนนที่ 4.5

4.บทสรุป

บทความนี้ได้นำเสนอเทคนิคการสังเคราะห์เสียงด้วยวิธีการ sinusoidal plus noise ซึ่งการสังเคราะห์เสียงแนวทางนี้มีข้อดีเหนือกว่าเทคนิคอื่น เนื่องจากเป็นโมเดลที่สามารถใช้ได้ทั่วไป เพียงแต่เปลี่ยนค่าพารามิเตอร์ที่ได้จากการวิเคราะห์สเปกตรัมของสัญญาณเสียงแต่ละชนิด ซึ่งสามารถนำมาประยุกต์ใช้ได้กับเสียงดนตรีขลุ่ยไทยได้ โดยจากการวิเคราะห์มีจำนวนความถี่พื้นฐานและความถี่ฮาร์โมนิกที่ใช้ในการสังเคราะห์สัญญาณคลื่นไซน์ 5-12 ค่าความถี่ ส่วนในการสังเคราะห์สัญญาณรบกวน พบว่าต้องใช้ค่าสัมประสิทธิ์ตัวกรองสัญญาณจำนวนสูงกว่า 50 ค่าขึ้นไป (ซึ่งเมื่อนำมารวมกับส่วนของสัญญาณคลื่นไซน์) จึงจะได้เสียงสังเคราะห์ที่ไม่สามารถแยกความแตกต่างจากเสียงขลุ่ยจริง สำหรับงานวิจัยในอนาคต จะประยุกต์ใช้เทคนิคดังกล่าวข้างต้นในส่วนของสัญญาณที่มีการเปลี่ยนแปลง (attack และ decay) รวมทั้งศึกษาความเป็นไปได้ในการลดความซับซ้อนในการคำนวณด้วย

5.เอกสารอ้างอิง

[1] Nomer, D. "Musical Physical Modeling", [Online]. Available: http://web800.com/music/vl/physmodl.htm. 1998

[2] Peevers, A.W. "A Real Time 3D Signal Analysis/Synthesis Tool Based on the Short Time Fourier Transform", [Online]. Available: http://cnmat.CNMAT.Berkeley.EDU/~alan/MS-html/MSv2_ToC.html 1999

[3] Rodet, X. "Musical Sound Signal Analysis/Synthesis: Sinusoidal+Residual and Elementary Waveform Models", Paris, France, 1997.

[4] Serra, X. "Musical Sound Modeling With Sinusoidal Plus Noise.", *Musical Signal Processing*. Swets& Zeitlinger Publishers,1997.

[5] Serra, X. "Integrating Complementary Spectral Models in the Design of a Musical Synthesizer", *The Proceeding of the International Computer Music Conference* , 1997.

[6] Serra, X. "Current Perspectives in the Digital Synthesis of Musical Sounds", *Pompeau Fabra University*. [Online]. Available: http://www.iua.upf.es/~xserra/articles/current. 1997.

เอกสารนี้ ... ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ... และต้องอ้างอิงถึงเจ้าของลิขสิทธิ์ ... ไปใช้

[7] Smith, J.O. "Viewpoints on the History of Digital Synthesis", *Proceedings of the 1991 International Computer Music Conference*, Montreal, San Fracisco, 1-10 Oct. 1991.

6.กิตติกรรมประกาศ

ขอขอบพระคุณ อ. ธนิสร์ ศรีกลิ่นดี ที่กรุณาเล่นขลุ่ยเพียงออ เพื่อการบันทึกเก็บเป็นข้อมูลในการทำวิจัย และนางสาวสุวรรณี กิ่งเพชร ที่ให้การบันทึกเสียงขลุ่ยประกอบการศึกษาลักษณะสัญญาณเสียงในเบื้องต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อผู้เขียน	นางสาวรุ่งกานต์ ศิริเจริญไชย
วันเดือนปีเกิด	22 มกราคม 2518
สถานที่เกิด	จังหวัดสุรินทร์
วุฒิการศึกษาระดับปริญญาตรี	ศิลปศาสตรบัณฑิต
สถานที่สำเร็จการศึกษา	มหาวิทยาลัยธรรมศาสตร์
ปีที่สำเร็จการศึกษา	2540
ทุนการศึกษาที่เคยได้รับ	สิงหาคม 2543 ได้รับทุนงบประมาณประจำปี 2543 มิถุนายน 2544 ได้รับทุนสนับสนุนการทำวิทยานิพนธ์ประจำปี 2544

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้