บทที่ 2
วรรณกรรมปริทัศน์และทฤษฎีที่เกี่ยวข้อง

2.1 วรรณกรรมปริทัศน์

2.1.1 งานวิจัยที่เกี่ยวข้องทางด้านการวิเคราะห์เสียงดนตรีไทย

การวิเคราะห์เสียงดนตรีไทยด้วยการประมวลผลสัญญาณดิจิตอลมาประยุกต์ใช้ได้เริ่มขึ้นอย่างเป็นรูปธรรมในปี พ.ศ. 2545 โดย สราวุฒิ สุจิตจร (2545) ได้ทำการวิเคราะห์ระบบเสียงดนตรีไทย โดยเทคนิคที่ใช้ในการวิเคราะห์ ได้แก่ เทคนิคการแปลงฟูริเยร์เต็มหน่วย (Discrete Fourier Transform-DFT) เทคนิคการแปลงฟูริเยร์ช่วงเวลาสั้น (Short-Time Fourier Transform-STFT) และเทคนิคออโตรีเกรสซีฟโมเดล (Auto Regressive Model-AR) ซึ่งหลังจากการใช้เทคนิคดังที่กล่าวมาแล้วในการวิเคราะห์พบว่าเทคนิค DFT และเทคนิค STFT ให้ผลการวิเคราะห์ที่สมจริงและใกล้เคียงกันมาก ส่วนเทคนิค AR ไม่เหมาะที่จะใช้วิเคราะห์เสียงดนตรีไทย เพราะให้ผลการวิเคราะห์ที่เป็นไปไม่ได้ เช่น ความถี่เสียงมี สูงกว่าเสียงฟา ซึ่งในความเป็นจริงแล้วความถี่เสียงฟาจะต้องสูงกว่าเสียงมี เป็นต้น

ถัดมาในปี พ.ศ. 2547 ฤดีรัตน์ ชินเวชกิจวานิชย์ และคณะ (2547) ได้ทำการวิเคราะห์ระบบเสียงดนตรีไทยโดยเป็นการวิจัยต่อเนื่องมาจาก สราวุฒิ สุจิตจร (2545) และได้เพิ่มเทคนิคการวิเคราะห์สัญญาณเข้ามาอีกหนึ่งกรรมวิธีคือ การกระจายเชิงโหมด (Modal Distribution-MD) ซึ่งผลการวิเคราะห์ที่ได้พบว่าเทคนิค MD ให้ผลการวิเคราะห์ที่ดีที่สุด เนื่องจากแสดงการไล่ระดับเสียงตามความสูงต่ำของเสียงได้ถูกต้อง และใน 1 ทบเสียง เช่น เสียงโดสูงมีความถี่ใกล้เคียงกับสองเท่าของเสียงโดต่ำ ซึ่งเป็นไปตามหลักทฤษฎีดนตรีทั่วไป

ถัดมาในปี พ.ศ. 2549 กิตติพงษ์ มีสวาสดิ์ และคณะ (2549) ได้ทำการประเมินระบบเสียงของขลุ่ยเพียงออ โดยกลุ่มตัวอย่างของขลุ่ยประกอบด้วยขลุ่ยไม้จากแหล่งต่าง ๆ กันรวม 8 เลา ทำการวิเคราะห์ความถี่แต่ละระดับเสียงของขลุ่ยแต่ละเลาด้วยการแปลงฟูริเยร์ในช่วงเวลาสั้น และนำผลที่ได้มาวิเคราะห์ข้อมูลทางสถิติ พบว่าค่าเบี่ยงเบนมาตรฐานความถี่ของขลุ่ยทั้ง 8 เลามีค่าไม่เกิน 20 Hz ค่าเบี่ยงเบนมาตรฐานระยะพิตช์มีค่าไม่เกิน 25 เซนต์ และพบว่าค่าเฉลี่ยสัมบูรณ์ของขลุ่ยมีความแตกต่างกัน แต่ระยะพิตช์เฉลี่ยของขลุ่ยทั้ง 8 เลาสอดคล้องกับรูปแบบระยะพิตช์ของเสียงที่ได้จากการศึกษาก่อนหน้านี้ (สราวุฒิ สุจิตจร, 2545; ฤดีรัตน์ ชินเวชกิจวานิชย์, 2547)

2.1.2 งานวิจัยที่เกี่ยวข้องทางด้านการสังเคราะห์เสียงดนตรีไทย

ในปี พ.ศ. 2544 รุ่งกานต์ ศิริเจริญไชย (2544) ได้ทำการสังเคราะห์เสียงขลุ่ยเพียงออโดยใช้แบบจำลองสัญญาณไซน์ผสมสัญญาณรบกวน (Sinusoidal Plus Noise Model) ซึ่งงานวิจัยนี้ได้ทำการวิเคราะห์หาค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในการสังเคราะห์ โดยการทดลองจะสร้างแบบจำลองสัญญาณเสียงออกเป็น 2 ส่วน คือ ส่วนที่เป็นสัญญาณไซนูซอยด์ (Sinusoid)

โดยใช้กรรมวิธีการสังเคราะห์เสียงแบบแอดดิทีฟ (Additive Synthesis) และส่วนที่เป็นสัญญาณรบกวน (Noise) โดยกระบวนการสังเคราะห์เสียงนั้น ส่วนของการสร้างสัญญาณไซนูซอยด์ต้องใช้ความถี่มูลฐานและความถี่ฮาร์มอนิกส์จำนวน 10 ค่าความถี่ และในส่วนของการสร้างสัญญาณรบกวนต้องใช้จำนวนสัมประสิทธิ์ของตัวกรองสัญญาณอย่างน้อย 15 จำนวน และจากผลการทดลองพบว่าแบบจำลองเออาร์ (AR Model) จะให้ลักษณะสัญญาณรบกวนสังเคราะห์ใกล้เคียงกับสัญญาณรบกวนต้นฉบับมากกว่าแบบจำลองเอ็มเอ (MA Model) ซึ่งผลการทดลองแสดงให้เห็นว่าเสียงสังเคราะห์ที่ได้จะยังคงให้เสียงใกล้เคียงกับเสียงต้นฉบับ สำหรับข้อดีของการสังเคราะห์เสียงด้วยกรรมวิธีนี้คือ มีความยืดหยุ่นในการในการปรับแก้พารามิเตอร์ต่าง ๆ ที่ใช้ในการสังเคราะห์ แต่อย่างไรก็ตามกรรมวิธีนี้ไม่เหมาะสมในการนำไปประยุกต์สร้างเป็นเครื่องสังเคราะห์เสียงดนตรี เนื่องจากต้องการทรัพยากรในการคำนวณมาก

ต่อมาในปี พ.ศ. 2550 กิตติพงษ์ มีสวาสดิ์ และภาคภูมิ โชคชัยรุ่งโรจน์ (2550) ได้ทำการวิเคราะห์และสังเคราะห์เสียงเครื่องดนตรีไทยประเภทพิณพื้นบ้าน โดยใช้กรรมวิธีจำลองสภาพทางกายภาพ โดยสร้างแบบจำลองของส่วนประกอบของเครื่องดนตรีไทยจากสมการเชิงอนุพันธ์ย่อย (Partial Differential Equation) และแปลงรูปสมการนี้ให้อยู่ไม่ต่อเนื่อง (Discrete) โดยใช้กรรมวิธีผลต่างสืบเนื่องในแกนเวลา (Finite Difference Time Domain) ซึ่งผลการวิเคราะห์ที่ได้พบว่าระบบขั้นเสียงของพิณพื้นบ้านนั้นมีความใกล้เคียงกับระบบขั้นเสียงดนตรีสากล แต่มีความแตกต่างจากระบบขั้นเสียงของดนตรีไทย และในส่วนของการสังเคราะห์เสียงนั้นเสียงสังเคราะห์ที่ได้มีความแตกต่างจากเสียงพิณต้นฉบับแต่ยังคงลักษณะของเสียงพิณเอาไว้ได้ แต่อย่างไรก็ตามแนวทางนี้ใช้ทรัพยากรในการคำนวณสูงและใช้เวลานานในการประมวลผลสัญญาณ

2.1.3 งานวิจัยที่เกี่ยวข้องทางด้านการสังเคราะห์เสียงดนตรีด้วยกรรมวิธี Subtractive Synthesis

ในปี ค.ศ. 1975 Makhoul (1975) ได้อธิบายถึงการนำการประมาณค่าเชิงเส้น (Linear Predictive Coding) มาใช้ในการวิเคราะห์สัญญาณไม่ต่อเนื่องทางเวลา (Discrete-Time Signals) ซึ่งในขณะนั้นได้มีการนำกรรมวิธีนี้มาประยุกต์ใช้ทางด้านนิวโรฟิสิกส์ (Neurophysics) ในการตรวจวัดคลื่นไฟฟ้าสมอง (EEG Signals) เพื่อหาระดับของคลื่นสมองเพื่อใช้ในการวินิจฉัยโรค สำหรับทางด้านจีโอฟิสิกส์ (Geophysics) จะใช้ในการวิเคราะห์หาแหล่งน้ำมันจากสัญญาณที่ได้จากเครื่องวัดแผ่นดินไหว (Seismograph) สำหรับทางด้านการสื่อสาร (Communication) จะใช้ในการวิเคราะห์และสังเคราะห์เสียงพูด โดยทำการวิเคราะห์หาพารามิเตอร์ต่าง ๆ จากสัญญาณเสียงพูด จากนั้นเครื่องส่งจะทำการส่งข้อมูลเฉพาะค่าพารามิเตอร์เหล่านั้นไปยังเครื่องรับเพื่อทำการสังเคราะห์เสียงพูดกลับคืนมา โดยวิธีการนี้จะช่วยลดจำนวนข้อมูลที่จะใช้ส่งทำให้การสื่อสารมีความรวดเร็วมากยิ่งขึ้นเพราะไม่ต้องส่งข้อมูลไปทั้งหมด

งานวิจัยนี้ได้นำการประมาณค่าเชิงเส้นมาใช้ออกแบบแบบจำลองตัวกำทอนสำหรับขลุ่ยเพียงออซึ่งเป็นส่วนหนึ่งของการสังเคราะห์เสียงด้วยกรรมวิธี Subtractive Synthesis โดยจะได้นำเสนอต่อไปในหัวข้อทฤษฎีที่เกี่ยวข้อง

ถัดมาในปี ค.ศ. 2004 Sturdivant (2004) ได้ทำการสังเคราะห์เสียงบีแฟลทคลาริเน็ท (B^b Clarinet) โดยใช้กรรมวิธี Subtractive Synthesis ซึ่งงานวิจัยนี้ได้ทำการออกแบบตัวกรองสัญญาณ 2 ชนิด ที่ใช้ในการสังเคราะห์เสียง คือ Autoregression Autocorrelation Liner Predictors และ Least Squares Filter ซึ่งจากผลการทดลองที่ได้พบว่าตัวกรองสัญญาณทั้งสองชนิดนี้มีความแตกต่างกันน้อยมาก เมื่อนำมาสร้างเป็นสัญญาณเสียงสังเคราะห์ ซึ่งมนุษย์ไม่สามารถแยกแยะความแตกต่างของเสียงได้ แต่อย่างไรก็ตามในด้านของการประมวลผลสัญญาณตัวกรองสัญญาณชนิด Autoregression Autocorrelation Linear Predictors ให้ผลดีที่ดีกว่าในด้านความต้องการในการทรัพยากรในการคำนวณที่ต่ำกว่า

2.1.4 งานวิจัยที่เกี่ยวข้องทางด้านการเข้ารหัสแบนด์ย่อยโดยใช้การประมาณค่าเชิงเส้น

ในปี พ.ศ. 2552 ธัญลักษณ์ ดิษฐ์สูงเนิน (2552) ได้นำเทคนิคการเข้ารหัสสัญญาณเสียงพูดโดยใช้การประมาณค่าเชิงเส้นผนวกเข้ากับเทคนิคการเข้ารหัสแบนด์ย่อยซึ่งนิยมใช้กับสัญญาณเสียงที่เป็นออดิโอ (Audio) มาใช้ในการเข้ารหัสสัญญาณเสียงพูด โดยเรียกเทคนิคใหม่นี้ว่า การประมาณค่าเชิงเส้นแบนด์ย่อย ซึ่งเทคนิคใหม่นี้ได้นำไปเข้ารหัสสัญญาณเสียงพูดและทำการเปรียบเทียบกับเทคนิคแบบปัจจุบันอีก 3 กรรมวิธีด้วยกัน ได้แก่ การประมาณค่าเชิงเส้นแบบดั้งเดิม การประมาณค่าเชิงเส้นแบบ 2 และ 4 แบนด์ย่อยที่ใช้การแบ่งแบนด์ย่อยโดยการแปลงเวฟเลต ซึ่งจากผลการทดลองพบว่าเทคนิคใหม่นี้สามารถสังเคราะห์สัญญาณเสียงพูดที่มีคุณภาพดีที่สุด และใช้อัตราบิตที่ต่ำกว่าอีก 3 กรรมวิธีที่ได้กล่าวมาข้างต้น

ในหัวข้อต่อไปจะนำเสนอเกี่ยวกับ เทคนิคการวิเคราะห์สัญญาณเสียงขลุ่ยเพียงออด้วยฟังก์ชันความหนาแน่นสเปกตรัมกำลัง เทคนิคการประมาณค่าเชิงเส้นแบบดั้งเดิมและเทคนิคการประมาณค่าเชิงเส้นแบนด์ย่อยสำหรับนำไปออกแบบแบบจำลองตัวกำทอนสำหรับขลุ่ยเพียงออโดยละเอียด กรรมวิธีการสังเคราะห์เสียงด้วยกรรมวิธี Subtractive Synthesis และแบบจำลองขอบของสัญญาณเสียง เป็นลำดับสุดท้าย

2.2 ทฤษฎีที่เกี่ยวข้อง

สำหรับทฤษฎีที่ใช้ในงานวิจัยนี้จะประกอบไปด้วย 5 ส่วนหลักด้วยกันได้แก่ การวิเคราะห์สัญญาณโดยใช้กรรมวิธีของเวลช์ การประมาณค่าเชิงเส้น การประมาณค่าเชิงเส้นแบนด์ย่อย การสังเคราะห์เสียงด้วยกรรมวิธี Subtractive Synthesis และแบบจำลองขอบของสัญญาณเสียงตามลำดับ

2.2.1 การวิเคราะห์สัญญาณด้วยฟังก์ชันความหนาแน่นสเปกตรัมกำลังด้วยกรรมวิธีของเวลช์

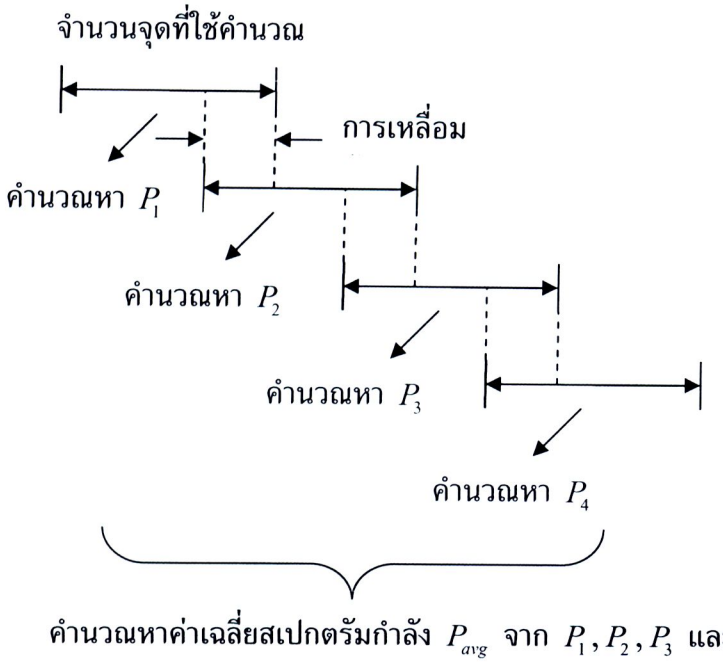
กรรมวิธีที่ใช้ในการประมาณค่าสเปกตรัมมีอยู่ด้วยกันหลายวิธี ซึ่งหนึ่งในกรรมวิธีที่เป็นที่นิยมและใช้กันอย่างกว้างขวางก็คือ Average Modified Periodogram (Proakis, Manolakis, 1996) หรือเรียกสั้น ๆ ว่า กรรมวิธีของเวลช์ (Welch's Method) สำหรับการคำนวณหาค่าเฉลี่ยสเปกตรัมกำลัง P_{avg} สามารถแสดงได้ดังสมการที่ (1)

$$P_{avg}(f) = \frac{1}{K}\sum_{i=1}^{K} P_{xi}(f) \qquad (1)$$

โดยที่ K คือ จำนวนบล็อกที่ใช้ในการแบ่งสัญญาณก่อนทำการวิเคราะห์

P_{xi} คือ ค่าสเปกตรัมกำลังในแต่ละบล็อก

ซึ่งหลักการของกรรมวิธีนี้คือลำดับแรกนำสัญญาณที่จะนำมาวิเคราะห์มาแบ่งเป็นบล็อก จากนั้นนำแต่ละบล็อกมาคูณด้วยฟังก์ชันหน้าต่าง (Window Function) และคำนวณหาค่าสเปกตรัมกำลังในแต่ละบล็อกด้วยกรรมวิธีการแปลงฟูเรียร์แบบรวดเร็ว (Fast Fourier Transform-FFT) จากนั้นทำการเลื่อนจุดให้มีการเหลื่อมล้ำ (Overlap) ตามที่กำหนดและทำการคำนวณซ้ำตามวิธีการที่กล่าวมาจนครบทุกบล็อก และขั้นตอนสุดท้ายจึงนำค่าสเปกตรัมกำลังในแต่ละบล็อกมาคำนวณเพื่อหาค่าเฉลี่ย โดยงานวิจัยนี้ใช้ฟังก์ชัน Pwelch ซึ่งเป็นฟังก์ชันสำเร็จรูปในโปรแกรม Matlab® ที่ใช้การคำนวณหาค่าเฉลี่ยสเปกตรัมกำลังด้วยกรรมวิธีของเวลช์ ซึ่งพารามิเตอร์ต่าง ๆ ที่ใช้ในการวิเคราะห์จะอยู่ในหัวข้อที่ 3.3 โดยการจัดแบ่งบล็อกของสัญญาณเพื่อหาค่าเฉลี่ยสเปกตรัมกำลังตามกรรมวิธีของเวลช์ดังภาพที่ 2.1



ภาพที่ 2.1 การจัดแบ่งบล็อกของสัญญาณเพื่อหาค่าเฉลี่ยสเปกตรัมกำลังตามกรรมวิธีของเวลช์

2.2.2 การประมาณค่าเชิงเส้น (Linear Predictive Coding-LPC)

การประมาณค่าเชิงเส้นเป็นกรรมวิธีที่นิยมใช้ในการเข้ารหัสการประมาณค่าสัญญาณเสียงพูด (Chu, 2003) ซึ่งในส่วนของงานวิจัยนี้จะเป็นการประยุกต์ในอีกแนวทางหนึ่งคือนำมาใช้ในการจำลองตัวกำทอนสำหรับเสียงขลุ่ยเพียงออ โดยกำหนดให้สัญญาณที่ได้จากการประมาณค่าแทนด้วย $\hat{x}(n)$ ซึ่งได้จากผลรวมของสัญญาณเสียงก่อนหน้า $x(n-i)$ คูณกับค่าสัมประสิทธิ์การประมาณค่าเชิงเส้น a_i ซึ่งกรรมวิธีการคำนวณหาค่าค่าสัมประสิทธิ์การประมาณค่าเชิงเส้นแสดงดังสมการที่ (2) ถึง (8)

$$\hat{x}(n) = -\sum_{i=1}^{p} a_i x(n-i) \tag{2}$$

โดยที่ p คือ อันดับของการประมาณค่าเชิงเส้น
จากสมการที่ (2) สามารถหาค่าความผิดพลาดจากการประมาณได้ตามสมการที่ (3)

$$e(n) = x(n) - \hat{x}(n) \tag{3}$$

$$e(n) = x(n) + \sum_{i=1}^{p} a_i x(n-i) \tag{4}$$

โดยที่ $x(n)$ คือ สัญญาณเสียงที่นำมาทำการประมาณค่าเชิงเส้น

ในการประมาณค่าเชิงเส้น สัญญาณที่ได้จากการประมาณขึ้นมานั้นจะต้องมีลักษณะใกล้เคียงกับสัญญาณต้นฉบับมากที่สุด ดังนั้นจึงต้องทำการประมาณให้มีค่าผิดพลาดกำลังสองให้น้อยที่สุด โดยค่าผิดพลาดกำลังสองแสดงได้ดังสมการที่ (5)

$$\varepsilon = e^2(n) = \left[x(n) + \sum_{i=1}^{p} a_i x(n-i) \right]^2 \tag{5}$$

จากสมการที่ (5) สามารถคำนวณหาค่า a_i ได้จากเงื่อนไขค่าผิดพลาดกำลังสองน้อยสุดก็คือ $\frac{d\varepsilon}{da_i} = 0$ ซึ่งผลลัพธ์ที่ได้แสดงได้ดังสมการที่ 6

$$\sum_{i=1}^{p} a_i R(m-i) = -R(m); m = 1, 2, ..., p \tag{6}$$

จากสมการที่ (6) สามารถจัดให้อยู่ในรูปของเมตริกซ์ได้ดังสมการที่ (7)

$$\begin{bmatrix} R(0) & R(1) & \dots & R(p-1) \\ R(1) & R(0) & \dots & R(p-2) \\ \dots & \dots & \dots & \dots \\ R(p-1) & R(p-2) & \dots & R(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_3 \end{bmatrix} = - \begin{bmatrix} R(1) \\ R(2) \\ \dots \\ R(p) \end{bmatrix} \tag{7}$$

หรือเขียนเป็นสัญลักษณ์ได้ดังสัมการที่ (8)

$$R\vec{a} = -\vec{r} \tag{8}$$

โดยที่ R คือ ออโตคอร์รีเลชั่นเมตริกซ์

$\vec{r}$ คือ ครอสคอร์รีเลชั่นเวกเตอร์

$\vec{a}$ คือ เวกเตอร์ของค่าสัมประสิทธิ์การประมาณค่าเชิงเส้น

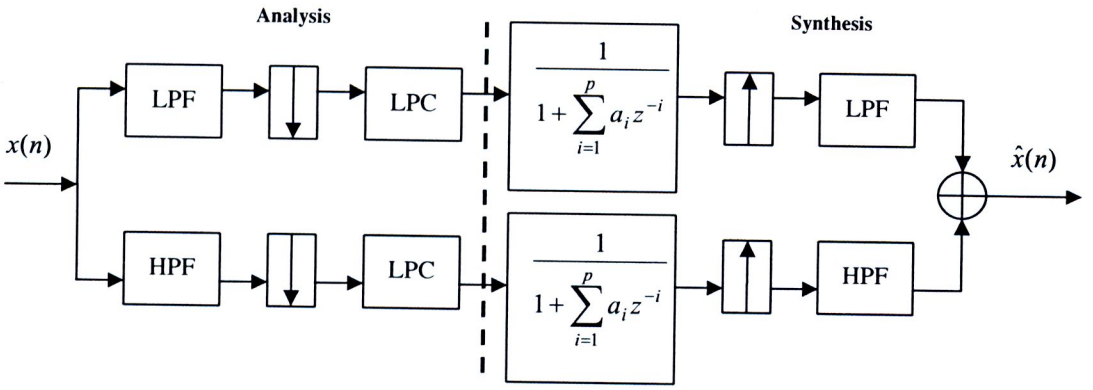
จากสมการที่ (7) สามารถแก้สมการเพื่อหาค่า a_i โดยอาศัยอัลกอริทึมของ Levinson และ Durbin (Makhoul, 1975) และในส่วนแบบจำลองของสัญญาณที่ได้จากการประมาณสามารถแสดงได้ดังสมการที่ (9)

$$H(z) = \frac{1}{1 + \sum_{i=1}^{p} a_i z^{-i}} \tag{9}$$

จากสมการที่ (9) แบบจำลองนี้จะนำไปใช้ออกแบบแบบจำลองตัวกำทอนสำหรับขลุ่ยเพียงออโดยใช้การประมาณค่าเชิงเส้นแบบดั้งเดิม และการประมาณค่าเชิงเส้นแบนด์ย่อยในลำดับต่อไป

2.2.3 การประมาณค่าเชิงเส้นแบนด์ย่อย (Sub-Band Linear Predictive Coding-SBLPC)

การประมาณค่าเชิงเส้นแบนด์ย่อยเป็นเทคนิคที่นำเสนอโดย ธัญญลักษณ์ ดิษฐ์สูงเนิน (2552) ซึ่งกรรมวิธีนี้เป็นเทคนิคการเข้ารหัสสัญญาณเสียงพูดที่ปรับปรุงจากการนำการประมาณค่าเชิงเส้นแบบดั้งเดิมผนวกเข้ากับการเข้ารหัสแบนด์ย่อยได้เป็นการเข้ารหัสสัญญาณเสียงพูดแบบใหม่ขึ้นมา โดยสามารถแสดงกระบวนการทำงานของกรรมวิธีนี้ได้ดังภาพที่ 2.2



ภาพที่ 2.2 กระบวนการการประมาณค่าเชิงเส้นแบนด์ย่อย

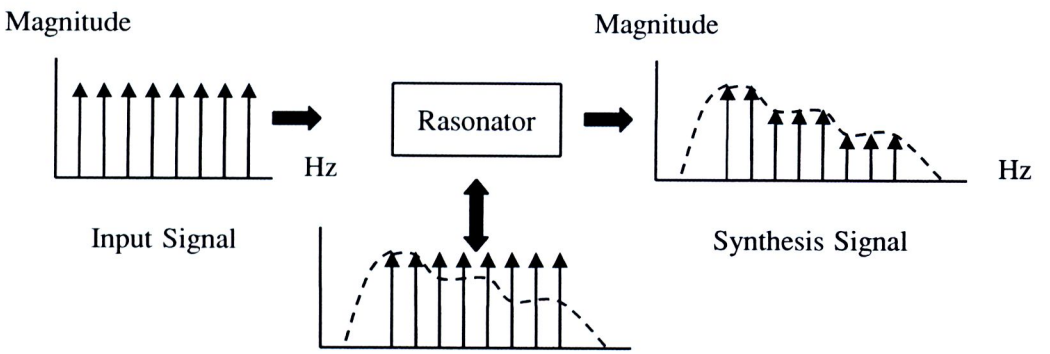
จากภาพที่ 2.2 ในลำดับแรกจะเป็นขั้นตอนการวิเคราะห์สัญญาณเสียงโดยกำหนดให้สัญญาณเสียง $x(n)$ เป็นสัญญาณที่จะนำมาประมาณค่า จากนั้นนำสัญญาณมาผ่านตัวกรองความถี่ต่ำผ่านและตัวกรองความถี่สูงผ่าน เพื่อแยกสัญญาณเป็นสองแบนด์ย่อยคือแบนด์ความถี่ต่ำและแบนด์ความถี่สูงตามลำดับ ถัดมานำสัญญาณทั้งสองแบนด์ที่ได้มาทำการลดอัตราสุ่มลง จากนั้นทำการคำนวณหาค่าสัมประสิทธิ์การประมาณค่าเชิงเส้นเพื่อใช้ในการสังเคราะห์เสียงต่อไป

ขั้นตอนต่อมาเมื่อต้องการสังเคราะห์เสียงให้นำสัมประสิทธิ์การประมาณค่าเชิงเส้นที่ได้จากขั้นตอนการวิเคราะห์มาสร้างแบบจำลองสัญญาณตามสมการที่ (9) จากนั้นทำการเพิ่มอัตราสุ่มทั้งแบนด์ความถี่ต่ำและแบนด์ความถี่สูงให้เท่ากับอัตราสุ่มสัญญาณก่อนทำการวิเคราะห์ และนำสัญญาณสังเคราะห์แบนด์ความถี่ต่ำมาผ่านตัวกรองความถี่ต่ำผ่านและนำสัญญาณสังเคราะห์

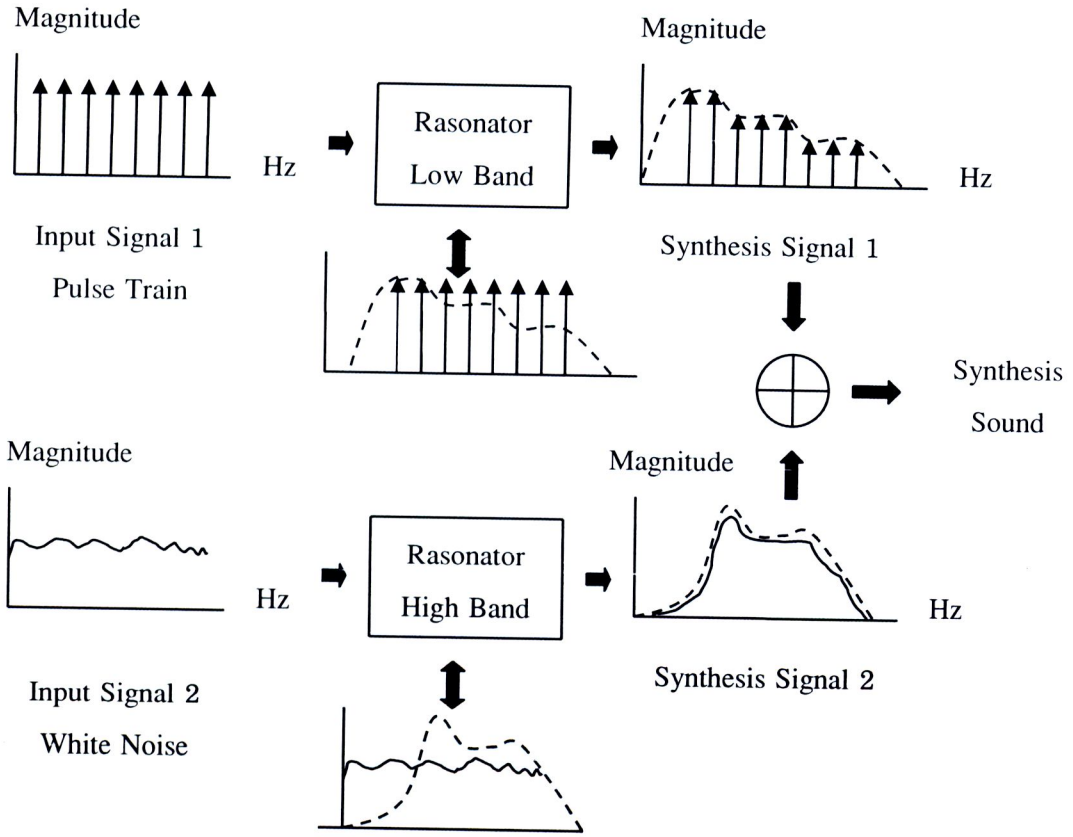
แบนด์ความถี่สูงมาผ่านตัวกรองความถี่สูงผ่าน ซึ่งเมื่อได้สัญญาณสังเคราะห์ทั้งสองแบนด์แล้วในขั้นตอนสุดท้ายให้นำสัญญาณสังเคราะห์ทั้งสองสัญญาณมารวมกันเพื่อสร้างเป็นเสียงสังเคราะห์ที่มีความสมบูรณ์ขึ้นมา

2.2.4 การสังเคราะห์เสียงด้วยกรรมวิธี Subtractive Synthesis

การสังเคราะห์เสียงด้วยกรรมวิธี Subtractive Synthesis (Cook, 2004) กรรมวิธีนี้ในลำดับแรกจะต้องทำการวิเคราะห์สัญญาณเสียงต้นฉบับเพื่อหาสัมประสิทธิ์การประมาณค่าเชิงเส้นเพื่อนำมาใช้สร้างแบบจำลองของสัญญาณ ซึ่งในวิจัยนี้ได้นำมาใช้สร้างแบบจำลองตัวกำทอนสำหรับขลุ่ยเพียงออ โดยหน้าที่หลักของตัวกำทอนจะทำหน้าที่ตัดลดองค์ประกอบเชิงความถี่ที่ไม่ต้องการออกไป หรือกล่าวอีกนัยหนึ่งก็คือลบองค์ประกอบเชิงความถี่ที่ไม่ต้องการจากสัญญาณขาเข้า จากนั้นเมื่อต้องการสร้างเสียงสังเคราะห์ ขั้นแรกให้สร้างสัญญาณขาเข้า (Input Signal) ที่มีองค์ประกอบเชิงความถี่ทุก ๆ ความถี่ ซึ่งงานวิจัยนี้ได้เลือกใช้สัญญาณสองชนิดด้วยกัน ได้แก่ สัญญาณ Pulse Train และ White Noise โดยสัญญาณ Pulse Train จะใช้ป้อนในแบบจำลองสัญญาณแบนด์ความถี่ต่ำ และ White Noise จะใช้ป้อนในแบบจำลองสัญญาณแบนด์ความถี่สูง ซึ่งกรรมวิธีการสังเคราะห์เสียงด้วยกรรมวิธี Subtractive Synthesis ที่จะนำเสนอในงานวิจัยนี้จะแตกต่างจากการสังเคราะห์เสียงด้วยกรรม Subtractive Synthesis โดยทั่วไป (ตัวกำทอนจะสร้างโดยใช้การประมาณค่าเชิงเส้นแบบดั้งเดิม) ที่จะเลือกใช้สัญญาณขาเข้าชนิดใดชนิดหนึ่งเพียงสัญญาณเดียวเท่านั้น โดยสามารถแสดงขั้นตอนของกรรมวิธีการสังเคราะห์เสียงด้วยกรรมวิธี Subtractive Synthesis โดยทั่วไปได้ดังภาพที่ 2.3 และกรรมวิธีที่นำเสนอในงานวิจัยนี้แสดงดังรูปที่ 2.4



ภาพที่ 2.3 กระบวนการสังเคราะห์เสียงด้วยกรรมวิธี Subtractive Synthesis ทั่วไป



ภาพที่ 2.4 กระบวนการสังเคราะห์เสียงด้วยกรรมวิธี Subtractive Synthesis ที่นำเสนอ

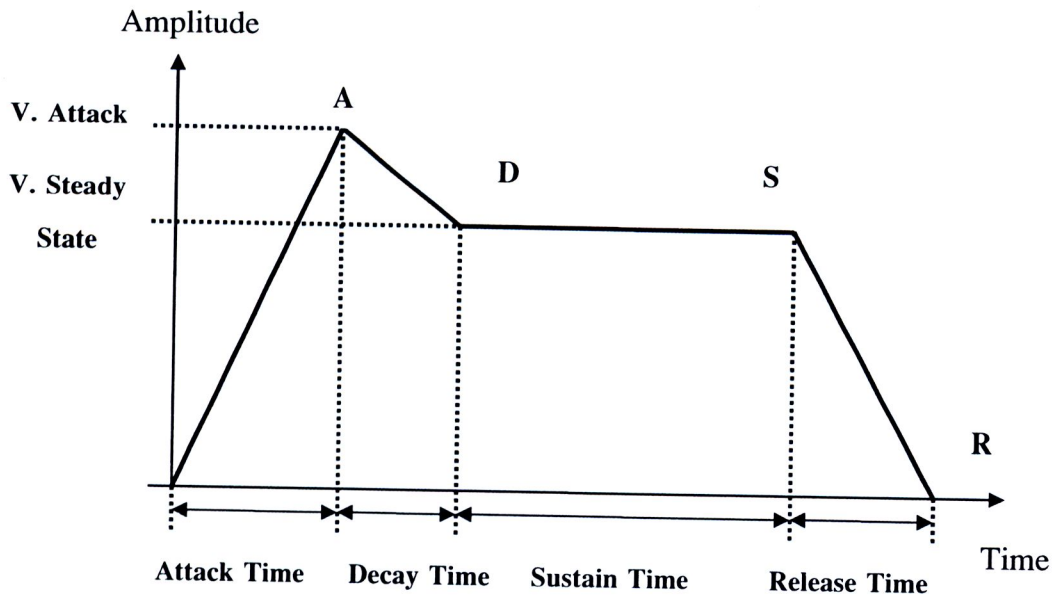
สำหรับขั้นตอนกรรมวิธีการสังเคราะห์อย่างละเอียดนั้นจะนำเสนอในบทที่ 3 กรรมวิธีทดลองเป็นลำดับต่อไป

สำนักงานคณะกรรมการวิจัยแห่งชาติ
ห้องสมุดงานวิจัย
วันที่ 12 มิ.ย. 2556
เลขทะเบียน 208859
เลขเรียกหนังสือ

2.2.5 แบบจำลองขอบของสัญญาณเสียง

เป้าหมายหลักของการสังเคราะห์เสียงด้วยกรรมวิธี Subtractive Synthesis นั้น อยู่ที่การจำลององค์ประกอบเชิงความถี่ของสัญญาณเสียงต้นฉบับที่จำเป็นให้ดีที่สุดโดยใช้ตัวกำทอน แต่นอกเหนือจากองค์ประกอบเชิงความถี่ของสัญญาณเสียงแล้ว องค์ประกอบทางเวลาก็มีความสำคัญไม่แพ้กัน ซึ่งเราสามารถกำหนดองค์ประกอบทางเวลาของสัญญาณเสียง (Envelope Model) ได้ ซึ่งปัจจุบันแบบจำลองที่นิยมใช้มากที่สุดในอุตสาหกรรมการสังเคราะห์เสียงดนตรีก็คือ แบบจำลองขอบของสัญญาณ (ADSR Model) (Dodge, 1997)

ขอบของสัญญาณเป็นกรอบที่แสดงถึงระดับพลังงานในแต่ละช่วงของสัญญาณเสียงนั้น ๆ ซึ่งมีรูปร่างเฉพาะในแต่ละเสียง โดยขอบของสัญญาณนั้นมีความสำคัญในกระบวนการสังเคราะห์เสียงเป็นอย่างมาก เนื่องจากทำหน้าที่ในการกำหนดจุดเริ่มต้นและจุดสิ้นสุดของเสียง อีกทั้งยังเป็นคุณลักษณะเด่นที่จะบ่งบอกว่าเสียงที่สังเคราะห์นั้นเป็นเสียงชนิดใด โดยดูได้จากรูปร่างขอบของสัญญาณนั่นเอง ซึ่งแบบจำลองขอบของสัญญาณเสียง (วีระ ทองไพบูลย์, 2549) แสดงได้ดังภาพที่ 2.5



ภาพที่ 2.5 แบบจำลองขอบของสัญญาณเสียง

งานวิจัยนี้จะกล่าวถึงขอบของสัญญาณเสียงในกรณีเสียงดนตรีที่เป็นเครื่องเป่าเท่านั้น โดยที่ Attack Time คือ ช่วงที่ผู้เล่นดนตรีทำการปฏิสัมพันธ์กับเครื่องดนตรี หรือกล่าวอีกนัยหนึ่งคือช่วงลมที่เป่าจากผู้เล่นมากระทบเครื่องดนตรีในระยะเวลาสั้น ๆ ก่อนที่ระดับพลังงานของเสียงดนตรีจะถูกลดระดับเข้าสู่สถานะอยู่ตัวหรือเรียกว่าช่วง Decay Time โดยในส่วนของ Sustain Time นั่นขึ้นอยู่กับช่วงเวลาที่ผู้เล่นทำการบรรเลงเครื่องดนตรี และในลำดับสุดท้ายเมื่อผู้เล่นหยุด

ทำการบรรเลงระดับพลังงานของเสียงจะยังคงอยู่ชั่วขณะก่อนจะลดสู่ศูนย์ ซึ่งเรียกช่วงเวลานี้ว่า Release Time