บทที่ 2

ทฤษฎีและหลักการ

ในบทนี้เป็นการกล่าวถึงทฤษฎีและหลักการพื้นฐานสำหรับการทำวิทยานิพนธ์ โดยเนื้อหาที่จะกล่าวถึงในบทนี้ประกอบด้วย หัวข้อ 2.1 ทฤษฎีดนตรีเบื้องต้น ซึ่งจะได้อธิบายหลักการพื้นฐานและคุณสมบัติของเสียงดนตรี, ลักษณะดนตรีไทยและดนตรีสากล หัวข้อ 2.2 ระนาดเอก กล่าวถึงลักษณะและองค์ประกอบของระนาดเอก หัวข้อ 2.3 การประมวลผลสัญญาณดิจิตอล อธิบายถึงหลักการพื้นฐานของการประมวลผลสัญญาณเสียง หลักการและลักษณะของตัวกรองความถี่ หัวข้อ 2.4 กล่าวถึงโครงข่ายประสาทเทียม โดยได้มุ่งเน้นไปที่โครงข่ายประสาทเทียมแบบ Feed Forward Neural Network หัวข้อที่ 2.5 โครงสร้างของไฟล์มีดี (MIDI) จะได้อธิบายถึงลักษณะการจัดเก็บข้อมูลของไฟล์มีดีและสุดท้ายคือหัวข้อที่ 2.6 สรุปเนื้อหาของบทนี้

2.1 ทฤษฎีดนตรีเบื้องต้น

หลักการพื้นฐานของและคุณสมบัติของเสียงดนตรีที่จะกล่าวถึงในหัวข้อนี้จะได้นำเสนอเกี่ยวกับพื้นฐานคุณสมบัติของเสียง, ตัวโน้ตและจังหวะ ซึ่งข้อมูลเหล่านี้จะเป็นพื้นฐานนำไปสู่ความเข้าใจในหลักการของดนตรีเบื้องต้น

2.1.1 คุณสมบัติของเสียง

เสียง เกิดขึ้นจากการสั่นไหวของวัตถุ เมื่อวัตถุใดเกิดการสั่นไหวขึ้นจะทำให้เกิดการสั่นสะเทือนในชั้นบรรยากาศ ทำให้เกิดคลื่นเสียงส่งต่อไปยังหูของคนเรา โดยปกติแล้วหูของคนเราจะได้ยินเสียงที่มีจำนวนรอบต่อวินาทีในการสั่นเทือนหรือความถี่ของคลื่นเสียงอยู่ในช่วง 20–20,000 เฮิรตซ์ [20]

คลื่นเสียงที่เกิดขึ้นมีลักษณะแตกต่างกันไปขึ้นอยู่กับ ลักษณะของแหล่งกำเนิดเสียง, จำนวนรอบต่อวินาทีของการสั่นสะเทือน ซึ่งเราสามารถจำแนกคุณสมบัติพื้นฐานของเสียงได้ดังต่อไปนี้

2.1.1.1 ระดับเสียง (Pitch)

ระดับเสียง คือ เสียงสูงและเสียงต่ำ เราสามารถเปรียบเทียบเสียงสูงเสียงต่ำได้โดยการฟัง ถ้าหากเสียงทั้งสองมีระดับเสียงต่างกันมากก็สามารถแยกแยะได้ง่ายแต่ถ้าเสียงมีระดับเสียง

ต่างกันน้อยก็จะทำให้พิจารณาแยะแยะลำบาก

จำนวนรอบต่อวินาทีของการสั่นสะเทือนหรือความถี่การสั่นของแหล่งกำเนิดเสียงเป็นตัวกำหนดระดับเสียง ถ้าหากแหล่งกำเนิดเสียงสั่นเร็วก็จะให้ระดับเสียงสูงแต่ถ้าหากแหล่งกำเนิดเสียงสั่นช้าก็จะให้ระดับเสียงต่ำ หรืออาจกล่าวได้ว่าวัตถุที่สั่นด้วยความถี่สูงกว่าจะให้ระดับเสียงที่สูงกว่านั่นเอง โดยถ้าหากวัตถุมีการสั่นด้วยความถี่เพิ่มมากขึ้นเป็น 2 เท่าของความถี่เดิม ระดับที่ได้จะสูงขึ้น 1 ช่วงคู่แปด (Octave)

2.1.1.2 ความเข้มเสียง (Volume)

ความเข้มเสียง คือ เสียงดังเสียงเบา ความเข้มเสียงเกิดจากแรงสั่นสะเทือนของวัตถุที่เป็นแหล่งกำเนิดเสียง ถ้าหากแรงสั่นสะเทือนของวัตถุมีมากก็จะทำให้ความเข้มเสียงเพิ่มมากขึ้นไปด้วย ขนาดของความเข้มเสียงวัดได้จากความสูงของคลื่นเสียง (Amplitude) นอกจากแรงสั่นสะเทือนจะมีผลความเข้มเสียงแล้วนั้น ระยะทางจากแหล่งกำเนิดเสียงจนถึงปลายทางก็มีผลด้วยเช่นกัน คือ ถ้าหากแหล่งกำเนิดเสียงอยู่ไกลก็จะทำให้ความเข้มเสียงลดลงหรือเสียงเบาลงนั่นเอง

2.1.1.3 สีสันเสียง (Timbre)

สีสันเสียง คือ คลื่นเสียงที่มีรูปร่างลักษณะที่แตกต่างกัน โดยเสียงที่มีแหล่งกำเนิดเสียงที่แตกต่างกัน หรือ เครื่องดนตรีที่แตกต่างกันจะให้ลักษณะคลื่นเสียงที่แตกต่าง ตัวอย่างเช่น เสียงเปียโน และเสียงกีตาร์ ที่บรรเลงเพลงระดับเสียงเดียวกัน ผู้ฟังจะสามารถรับรู้ได้ว่าเสียงทั้งสองแตกต่างกัน

พาร์เชียล (Partials) คือ กลุ่มของระดับเสียงจำนวนหนึ่งที่ประกอบกันเป็นเสียงที่เราได้ยิน โดยจำนวนของพาร์เชียล, ความเข้มเสียงของพาร์เชียลและการกระจายตัวของพาร์เชียลทั้งหมดนี้จะเป็นตัวกำหนดสีสันเสียง ตัวอย่างเช่น เสียงที่มีจำนวนพาร์เชียลน้อยจะมีเสียงใสกว่าเสียงที่มีจำนวนพาร์เชียลมาก สำหรับเสียงที่ไม่จำนวนพาร์เชียลเลยจะทำให้เสียงนั้นฟังแล้วไม่ไพเราะ เป็นต้น

2.1.1.4 คุณภาพเสียง (Tone Quality)

คุณภาพเสียง คือ ลักษณะของเสียงดีมากดีน้อย ซึ่งคุณสมบัตินี้จะเป็นคุณสมบัติเฉพาะตัวของเครื่องดนตรีแต่ละชิ้น เช่น โน้ตดนตรีตัวเดียวกันที่บรรเลงจากเปียโน 2 ตัว จะให้คุณภาพเสียงที่แตกต่างกัน ตัวแรกอาจจะให้เสียงทีฟังแล้วนุ่มและใสกว่า เป็นต้น

2.1.1.5 ความยาวเสียง (Duration)

ความยาวเสียง คือ ระยะเวลาของเสียง เสียงที่เกิดขึ้นจะต้องมีระยะเวลาของเสียง ไม่ว่าจะเป็นเสียงยาวหรือเสียงสั้น เสียงดนตรีที่เกิดขึ้นนั้นจะมีระยะเวลาเข้ามาเกี่ยวข้องเพื่อเป็นการกำหนดจังหวะของเพลง

2.1.2 ตัวโน้ตและจังหวะ (Note and Rhythm)

ตัวโน้ตและจังหวะเป็นหัวใจสำคัญในภาษาดนตรี [21] เนื่องจากในการบรรเลงแต่ละเพลงนั้น นักดนตรีจะต้องบรรเลงเพลงตามโน้ตและจังหวะที่ได้กำหนดไว้ สำหรับในหัวข้อนี้จะได้กล่าวอธิบายถึงเรื่องตัวโน้ตและจังหวะของตัวโน้ตดังต่อไปนี้

2.1.2.1 ชื่อตัวโน้ต

ตัวโน้ตมีอยู่ทั้งสิ้น 7 ตัวด้วยกัน โดยเมื่อเราเรียกชื่อตัวโน้ตตามลำดับครบหมดแล้วสามารถที่จะใช้ชื่อตัวโน้ตวนรอบต่อไปได้ โดยนำชื่อตัวโน้ตตัวที่ 1 มาเรียงต่อจากชื่อโน้ตตัวที่ 7 สำหรับระบบในการเรียกชื่อตัวโน้ตที่นิยมใช้กันนั้นมีอยู่ 2 ระบบคือ ระบบโซ-ฟา (So-Fa System) และระบบตัวอักษร (Letter System)

ระบบโซ-ฟา (So-Fa System) เป็นระบบการเรียกชื่อตัวโน้ตที่สามารถเรียงตามลำดับจากโน้ตที่มีเสียงต่ำไปยังโน้ตที่มีเสียงสูงและเรียงลำดับจากโน้ตเสียงสูงไปยังโน้ตเสียงต่ำได้ โดยเรียกชื่อโน้ตดังนี้ คือ โด (Do) เร (Re) มี (Mi) ฟา (Fa) โซ (So) ลา (La) ที (Ti) โด ในกรณีเรียงจากโน้ตเสียงต่ำไปโน้ตเสียงสูงและในทางกลับกันจะเรียกชื่อโน้ต โด ที ลา โซ ฟา มี เร โด ในกรณีที่เรียกโน้ตจากเสียงสูงไปโน้ตเสียงต่ำ

ระบบตัวอักษร (Letter System) สามารถเรียกชื่อตัวโน้ตเรียงลำดับได้เช่นเดียวกับระบบโซ-ฟา (So-Fa System) แต่ใช้ชื่อในการเรียกแตกต่างกันดังต่อไปนี้คือ เอ (A) บี (B) ซี (C) ดี (D) อี (E) เอฟ (F) จี (G) เอ ในกรณีเรียงจากโน้ตเสียงต่ำไปโน้ตเสียงสูงและในทางกลับกันจะเรียกชื่อโน้ต เอ จี เอฟ อี ดี ซี บี เอ ในกรณีที่เรียกโน้ตจากเสียงสูงไปโน้ตเสียงต่ำ

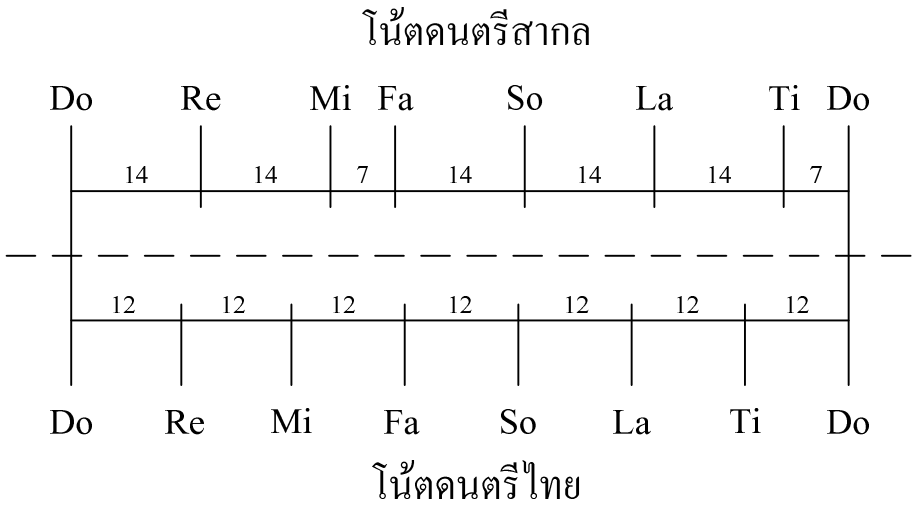
2.1.2.2 ช่วงคู่แปด (Octave)

เนื่องจากโน้ตดนตรีที่ใช้มีจำนวนมากแต่ชื่อตัวโน้ตที่มีนั้นมีเพียง 7 ตัว ดังนั้นในการเรียกชื่อโน้ตดนตรีจึงมีการวนรอบซ้ำตามลำดับชื่อตัวโน้ต สำหรับตัวโน้ตที่มีชื่อซ้ำกันแต่อยู่ต่างระดับเสียงกัน แสดงว่ามีระยะห่างของระดับเสียงเป็นช่วงคู่แปด ซึ่งอาจจะห่างกันมากกว่านั้นก็ได้

ขึ้นอยู่กับระดับเสียงของตัวโน้ต ทั้งนี้เราอาจเรียกได้ว่าโน้ตทั้งสองอยู่ต่างออกเทฟ (Octave) กัน

2.1.2.3 การแบ่งระยะในช่วงคู่แปด

สำหรับในดนตรีไทย ในการแบ่งช่วงเสียงนั้นจะแตกต่างกันกับดนตรีสากล คือ ในดนตรีไทยจะแบ่งช่วงเสียงเท่ากันหมดคือมี 7 เสียงต่างจากดนตรีสากลที่มีเสียงเต็ม 5 เสียงและมีครึ่งเสียงอยู่ 2 เสียงคือเสียงมี (Mi) และเสียงที (Ti) ดังแสดงในรูปที่ 2.1



รูปที่ 2.1 การแบ่งช่วงเสียงของดนตรีสากลและดนตรีไทย [22]

สำหรับในดนตรีสากล เราสามารถแบ่งระยะช่วงคู่แปดออกเป็นส่วนย่อยได้เป็น 12 ช่วงครึ่งเสียง (Semitone) ตัวอย่างเช่น ระยะระหว่างโน้ตมีและโน้ตฟาคือ 1 ช่วงครึ่งเสียงและระยะระหว่างโน้ตฟาและโน้ตโซคือ 2 ช่วงครึ่งเสียงหรือ 1 ช่วงเสียงเต็ม โดยเครื่องหมายที่ใช้ในการกำหนดระดับเสียงที่เป็นครึ่งเสียงนั้นจะใช้เครื่องหมายดังนี้คือ

เครื่องหมายชาร์ป (Sharp, ♯) ใช้ในการแปลงระดับเสียงของตัวโน้ตให้สูงขึ้นครึ่งเสียง

เครื่องหมายแฟลต (Flat, ♭) ใช้ในการแปลงระดับเสียงของตัวโน้ตให้ต่ำลงครึ่งเสียง

2.1.2.4 ค่าของตัวโน้ต

ชื่อตัวโน้ตใช้ในการระบุระดับเสียง แต่เนื่องจากในทางดนตรีนั้นในแต่ละระดับเสียงที่บรรเลงออกมาจะต้องมีความเกี่ยวข้องกับระยะเวลาด้วย หรือจะกล่าวได้ว่า โน้ตแต่ะตัวจะต้องมีตัวกำหนดระยะเวลาของโน้ต

สำหรับชนิดของตัวโน้ตที่แตกต่างกันจะเป็นตัวกำหนดระยะเวลาของโน้ตที่แตกต่างกัน ดังต่อไปนี้คือ

1. โน้ตตัวกลม (Whole note)
2. โน้ตตัวขาว (Half note)
3. โน้ตตัวดำ (Quarter note)
4. โน้ตเขบ็ตหนึ่งชั้น (Eighth note)
5. โน้ตเขบ็ตสองชั้น (Sixteenth note)
6. โน้ตเขบ็ตสามชั้น (Thirty-Second note)

สำหรับค่าระยะเวลาของตัวโน้ตที่กล่าวถึงจะเป็นค่าที่มีความสัมพันธ์กัน นั่นคือ โน้ตตัวกลมจะมีค่าระยะเวลาเป็น 2 เท่าของโน้ตตัวขาว, โน้ตตัวขาวจะมีค่าระยะเวลาเป็น 2 เท่าของโน้ตตัวดำ, โน้ตตัวดำมีค่าระยะเวลาเป็น 2 เท่าของโน้ตเขบ็ตหนึ่งชั้น ค่าระยะเวลานี้จะเป็นไปตามลำดับชั้น

2.1.2.5 กลุ่มเคาะ (Beat)

กลุ่มเคาะ คือ การเคาะเน้นที่มีจังหวะสม่ำเสมอ หรืออาจเรียกได้ว่าเป็นหน่วยเวลาพื้นฐาน

2.1.2.6 อัตราความเร็ว (Tempo)

อัตราความเร็ว เป็นการกำหนดความช้าเร็วของบทเพลง โดยจะกำหนดให้กับเครื่องเคาะจังหวะ ซึ่งการกำหนดอัตราเร็วจะส่งผลต่อค่าตัวโน้ต ดังตัวอย่างต่อไปนี้

ตัวอย่างที่ 1. ถ้ากำหนดอัตราเร็วเท่ากับ 60 บีต (Beat) ต่อนาที หมายถึงมีจำนวนการเคาะให้จังหวะทั้งหมด 60 ครั้งต่อนาทีหรืออาจกล่าวได้ว่ามีการเคาะจังหวะทุก 1 วินาที ดังนั้น ถ้ากำหนดให้โน้ตตัวดำ มีค่าระยะเท่ากับ 1 จังหวะหรือ 1 บีต แสดงว่าโน้ตตัวดำมีค่าระยะเวลาเท่ากับ 1 วินาทีซึ่งจะส่งผลให้โน้ตตัวขาวมีค่าเท่ากับ 2 วินาที, โน้ตตัวกลมเท่ากับ 4 วินาที ตามลำดับค่าของตัวโน้ต

ตัวอย่างที่ 2. ถ้ากำหนดอัตราเร็วเท่ากับ 120 บีตต่อนาที หมายถึงมีจำนวนการเคาะให้จังหวะทั้งหมด 120 ครั้งต่อนาทีหรืออาจกล่าวได้ว่ามีการเคาะจังหวะทุก 0.5 วินาที ดังนั้น ถ้ากำหนดให้โน้ตตัวดำ มีค่าระยะเท่ากับ 1 จังหวะหรือ 1 บีต แสดงว่าโน้ตตัวดำมีค่าระยะเวลาเท่ากับ 0.5 วินาทีซึ่งจะส่งผลให้โน้ตตัวขาวมีค่าเท่ากับ 1 วินาที, โน้ตตัวกลมเท่ากับ 2 วินาที ตามลำดับค่าของตัวโน้ต

2.1.3 ดนตรีไทย

สำหรับการศึกษาดนตรีไทยนั้นพบว่ามีข้อมูลพื้นฐานบางอย่างที่แตกต่างจากดนตรีสากล ดังตัวอย่างบางส่วนที่ได้นำเสนอไปใน หัวข้อ 2.1.2.3 สำหรับในหัวข้อนี้จะได้นำเสนอพื้นฐานและองค์ประกอบของดนตรีไทยเพิ่มเติมดังต่อไปนี้

2.1.3.1 เครื่องดนตรีไทย

เครื่องดนตรีไทยมีอยู่หลายประเภท โดยในที่นี้จะแบ่งแยกตามวิธีการที่ทำให้เกิดเสียงหรือจังหวะของ ซึ่งมีอยู่ 4 วิธีด้วยกันคือ

1. เครื่องดีด คือ อุปกรณ์ที่ใช้มือหรือสิ่งใดสิ่งหนึ่งดีดสายแล้วมีเสียงดังเกิดขึ้น ตัวอย่างเครื่องดนตรีประเภทเครื่องดีด เช่น จะเข้, กระจับปี่, พิณน้ำเต้า
2. เครื่องสี คือ อุปกรณ์ที่ใช้เส้นหางม้าหลาย ๆเส้นรวมกันแล้วสีให้มีเสียงดังเกิดขึ้น ตัวอย่างเครื่องดนตรีประเภทเครื่องสี เช่น ซอด้วง, ซออู้, ซอสามสาย
3. เครื่องตี คือ อุปกรณ์ที่ใช้มือหรือไม้ตีแล้วมีเสียงดังเกิดขึ้น ตัวอย่างเครื่องดนตรีประเภทเครื่องตี เช่น กรับ, ฆ้อง, ฉิ่ง, กลอง, ตะโพน, ระนาด
4. เครื่องเป่า คือ อุปกรณ์ที่ปากเป่าลมเข้าไปแล้วมีเสียงดังเกิดขึ้น ตัวอย่างเครื่องดนตรีประเภทเป่า เช่น ขลุ่ยกรวด, ขลุ่ยเพียงออ, ปี่ไฉน, ปี่ชวา

2.1.3.2 โน้ตดนตรีไทย

สำหรับชื่อโน้ตในดนตรีไทยเรียกชื่อเหมือนกับโน้ตดนตรีสากล มีการแบ่งช่วงคู่แปดเช่นเดียวกันกับโน้ตสากลคือ เรียกชื่อโน้ตวนรอบตามลำดับตัวโน้ต แต่แตกต่างกันตรงลักษณะการแบ่งช่วงตามที่ได้กล่าวไว้แล้วในหัวข้อ 2.1.2.3 โดยปกติทั่วไปแล้วในการเขียนสัญลักษณ์ตัวโน้ตไทยจะเขียนตามความนิยมของครูเพลงคือ เขียนตัวย่อภาษาไทยแทนสัญลักษณ์ตัวโน้ตคือ ด, ร, ม, ฟ, ซ, ล, ท ตามลำดับ แต่ถ้าหากโน้ตเพลงอยู่ในช่วงออกเทฟที่สูงกว่าจะใช้การเขียนจุดไว้ด้านบนตัวโน้ต และ ถ้าหากโน้ตเพลงอยู่ในช่วงออกเทฟที่ต่ำกว่าจะใช้การเขียนจุดไว้ด้านล่างตัวโน้ตและโดยปกติการเขียนโน้ตเพลงของดนตรีไทยนั้นจะใช้การเขียนโน้ตเพลงโดยกำหนดให้แต่ละบรรทัดแบ่งออกเป็น 8 ห้องแต่ละห้องมีโน้ตขนาด 1 จังหวะอยู่ภายใน 4 ตัวโน้ต ทั้งนี้โน้ต 1 จังหวะที่กำหนดจะมีระยะเวลาเท่าใดนั้นขึ้นอยู่กับการกับจังหวะของอุปกรณ์เครื่องดนตรีที่ให้จังหวะ เช่น กรับและฉิ่ง ตัวอย่างโน้ตเพลงดนตรีไทยแสดงดังรูปที่ 2.2 ตัวอย่างโน้ตเพลงดนตรีไทย

ท่อนที่ 1

ซ	ล	ด	ล̇	ซ	ฟ	ล̰	ล	ซ	ม	ล	ล̇	ซ	ร	ร	ด̇

ท่อนที่ 2

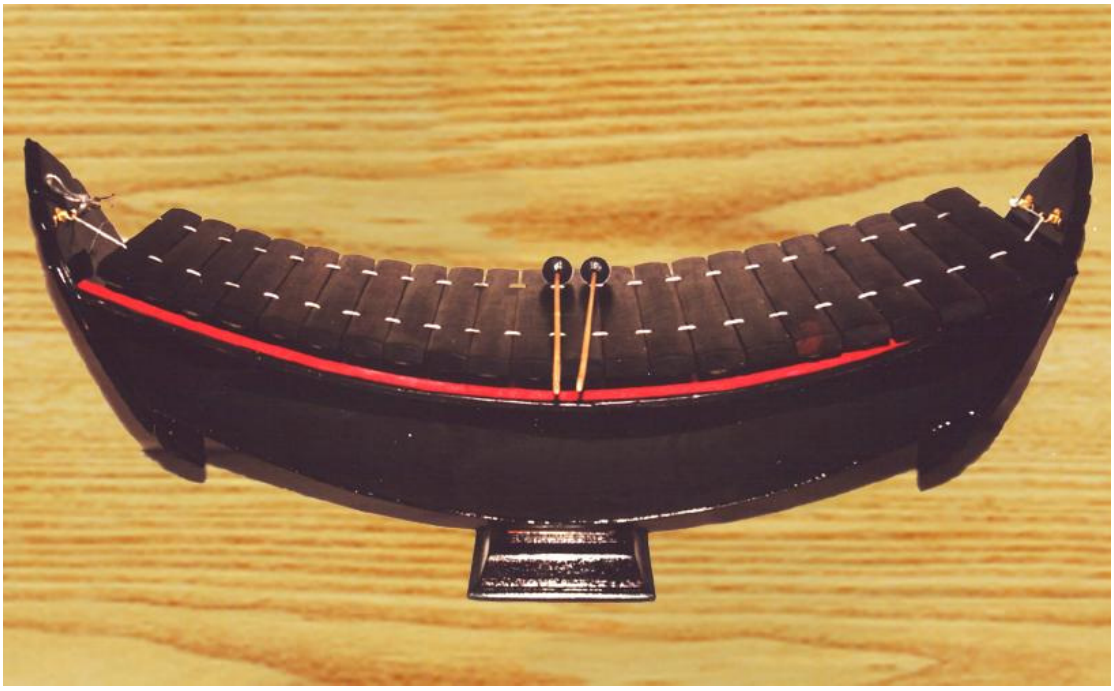
ด	ร	ร̇	ม	ซ	ฟ	ม	ร	ซ	ม	ซ̰	ล̇	ซ	ด	ร	ม

รูปที่ 2.2 ตัวอย่างโน้ตเพลงดนตรีไทย

2.2 ระนาดเอก

2.2.1 ลักษณะของระนาดเอก

ระนาดเอกเป็นเครื่องดนตรีไทยประเภทตีชนิดหนึ่ง [23][24] ที่วิวัฒนาการมาจากกรับ แต่เดิมใช้กรับสองอันตีเป็นจังหวะ ต่อมาได้มีการนำเอากรับหลาย ๆ อันวางเรียงต่อกันลงไปและแก้ไขให้มีขนาดลดหลั่นกัน จากนั้นจึงใช้เชือกร้อยไม้กรับขนาดต่าง ๆ กันนั้นให้ติดกันและขึงไว้บนรางและใช้ไม้ตีเพื่อให้เกิดเสียง ในการปรับระดับเสียงได้มีการนำเอาตะกั่วผสมกับขี้ผึ้งมาถ่วงเสียงโดยติดไว้ที่หัวท้ายของไม้กรับนั้น ให้เกิดเสียงไพเราะยิ่งขึ้น สำหรับไม้กรับที่ประดิษฐ์ขึ้นมานั้นเรียกว่า “ลูกระนาด” และเมื่อนำลูกระนาดมาผูกติดกันเป็นแผ่นเดียวกันว่า “ผืนระนาด”



รูปที่ 2.3 ระนาดเอก

2.2.2 องค์ประกอบของระนาดเอก

องค์ประกอบของระนาดเอกดังแสดงในรูปที่ 2.3 มีองค์ประกอบหลักอยู่ 3 ส่วนด้วยกัน คือ

2.2.2.1 ผืนระนาดเอก

ผืนระนาดเอกคือลูกระนาดมาผูกติดกันเป็นแผ่นเดียวกัน ลูกระนาดนี้ทำด้วยไม้ไผ่บง หรือไม้แก่น เช่นไม้ชิงชัน ไม้มะหาด ไม้พะยุงก็ได้ โดยนำมาเหลาให้ได้ตามขนาดที่ต้องการแล้วนำเอาตะกั่วผสมกับขี้ผึ้งมาถ่วงเสียง ระนาดเอกในปัจจุบันมีจำนวนลูกระนาดทั้งหมด 21 ลูก โดยมีขนาดลดหลั่นกันลงไปตามลำดับจากซ้ายมือไปทางขวามือของผู้บรรเลงระนาด

2.2.2.2 รางระนาดเอก

รางระนาดเอกทำด้วยไม้เนื้อแข็งเช่น ไม้สัก ไม้มะริด ไม้มะเกลือ ไม้ประดู่ เป็นต้น รูปร่างคล้ายเรือสมัยโบราณ คือตอนกลางป่องเป็นกระพุ้งเล็กน้อยเพื่อให้เสียงก้องกังวาน ตอนหัวและท้ายโค้งงอนขึ้นมีฝาประกบ 2 แผ่น บนขอบรางด้านบนทั้งสองข้างของฝาประกบจะใช้หวายเส้นซึ่งมีผ้าพันโดยรอบ ติดเป็นแนวยาวไว้ตลอด เพื่อรองรับผืนระนาดขณะที่ลดลงจากตะขอเกี่ยวป้องกันบริเวณใต้ท้องของผืนระนาดไม่ให้เป็นรอยได้ง่าย

2.2.2.3 ไม้ตีระนาดเอก

ไม้ตีระนาดเอกทำด้วยไม้ไผ่เหลาเป็นท่อนกลมเล็ก ๆ 2 อัน โดยทั่วไปแล้วไม้ตีระนาดเอกจะมี 2 แบบด้วยกันคือ แบบที่ 1 คือหัวไม้ตีทำด้วย ด้ายพันด้วยผ้าชุบรัก ลักษณะเป็นปื้นกลม เวลาตีมีเสียงดังแข็งกร้าวเรียกว่า “ไม้แข็ง” ส่วนแบบที่ 2 คือหัวไม้ตีทำด้วยผ้าติดตะกั่วเล็กน้อย ทาแป้งเปียกแล้วพันด้วยด้ายสีเส้นเล็ก ๆโดยรอบอย่างสวยงาม เวลาตีมีเสียงทุ้มนุ่มนวลเรียกว่า “ไม้นวม”

2.3 การประมวลผลสัญญาณดิจิตอล (Digital Signal Processing)

สำหรับในหัวข้อนี้จะได้กล่างอธิบายถึงขั้นตอนและหลักการพื้นฐานของการประมวลสัญญาณเสียง ร่วมถึงหลักการของตัวกรองสัญญาณความถี่ประเภทต่าง ๆ

2.3.1 ขั้นตอนและหลักการพื้นฐานของการประมวลสัญญาณเสียง

การประมวลผลสัญญาณดิจิตอล (Digital Signal Processing: DSP) [25][26] เป็นการประมวลผลสัญญาณเชิงตัวเลขที่มีความความแม่นยำในการคำนวณและประสิทธิภาพดีกว่าการประมวลผลสัญญาณในรูปแบบระบบสัญญาณอนาล็อก

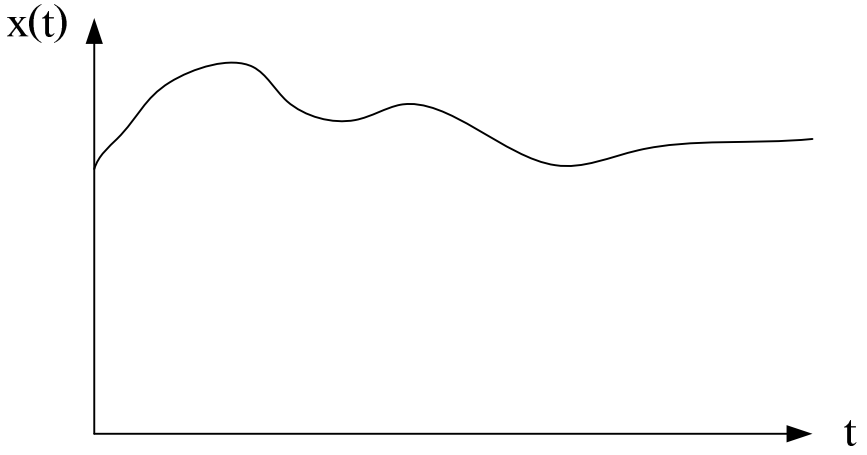
2.3.1.1 สัญญาณ (Signal)

สัญญาณส่วนใหญ่ในทางปฏิบัตินั้นเป็นสัญญาณเวลา ดังนั้นจึงสามารถแบ่งสัญญาณตามลักษณะจำเพาะของตัวแปรในเชิงเวลาและค่าของสัญญาณได้ดังนี้คือ

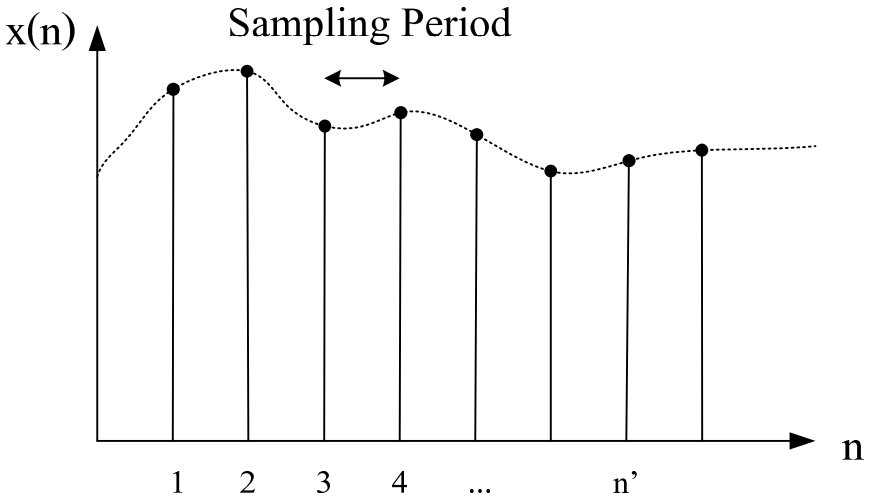
1. สัญญาณเชิงเวลาต่อเนื่อง (Continuous Data Signal) คือ สัญญาณที่มีรูปคลื่นของสัญญาณแปรค่าไปอย่างต่อเนื่องกับพิสัยเวลาแต่แอมพลิจูด (Amplitude) ของสัญญาณไม่ได้เจาะจงว่าจะต้องแปรไปอย่างต่อเนื่อง สำหรับสัญญาณอนาล็อกนั้นจัดเป็นสัญญาณเชิงเวลาต่อเนื่องด้วยเช่นกัน เพียงแต่มีแอมพลิจูดของสัญญาณมีการเปลี่ยนแปลงอย่างต่อเนื่องไปด้วย
2. สัญญาณเชิงเวลาเต็มหน่วย (Discrete Time Signal) คือ สัญญาณที่ค่าของฟังก์ชัน (Function) กำหนดเฉพาะเซตของเวลาที่แน่นอนอันหนึ่งเท่านั้นซึ่งยังสามารถแบ่งย่อยออกได้เป็น 2 แบบคือ
   - 2.1 สัญญาณเชิงข้อมูลเต็มหน่วยเป็นสัญญาณที่มีค่าของแอมพลิจูดเท่ากันทุกประการกับค่าที่อยู่ในสัญญาณอนาล็อกต้นแบบ
   - 2.2 สัญญาณเชิงเลข (Digital Signal) เป็นสัญญาณที่มีค่าแอมพลิจูดเป็นค่าที่แน่นอนคือ 0 และ 1

2.3.1.2 การสุ่มสัญญาณ (Signal Sampling)

สัญญาณที่เราพิจารณาคือสัญญาณดิจิตอล แต่เนื่องจากโดยทั่วไปแล้วนั้นสัญญาณที่ได้จากแหล่งกำเนิด เช่น สัญญาณเสียง จะอยู่ในรูปของสัญญาณอนาล็อก ดังนั้นจึงจำเป็นต้องมีการเปลี่ยนสัญญาณอนาล็อกให้เป็นสัญญาณเชิงเวลาเต็มหน่วย (Analog to Digital Converter: (A/D)) ก่อน และเพื่อเป็นการป้องกันการสูญเสียข้อมูลจึงต้องมีการสุ่มตัวอย่าง (Sampling) ข้อมูล โดยกระบวนการแปลงนี้สามารถอธิบายได้ดังรูปที่ 2.4 และรูปที่ 2.5



รูปที่ 2.4 สัญญาณเชิงเวลาต่อเนื่อง



รูปที่ 2.5 สัญญาณเชิงเวลาเต็มหน่วย

สำหรับหลักการของการสุ่มข้อมูลนั้นเพื่อป้องกันไม่ให้เกิดการสูญเสียสัญญาณที่สำคัญ เราสามารถสุ่มตัวอย่างข้อมูลได้โดยการใช้ความถี่ในการสุ่ม (Sampling Rate) ที่มีค่ามากกว่าหรือเท่ากับ 2 เท่าของความถี่สูงสุดที่มีอยู่ในสัญญาณข้อมูลดังสมการที่ (2-1)

$$f_s \geq 2f_0 \qquad (2\text{-}1)$$

เมื่อกำหนดให้

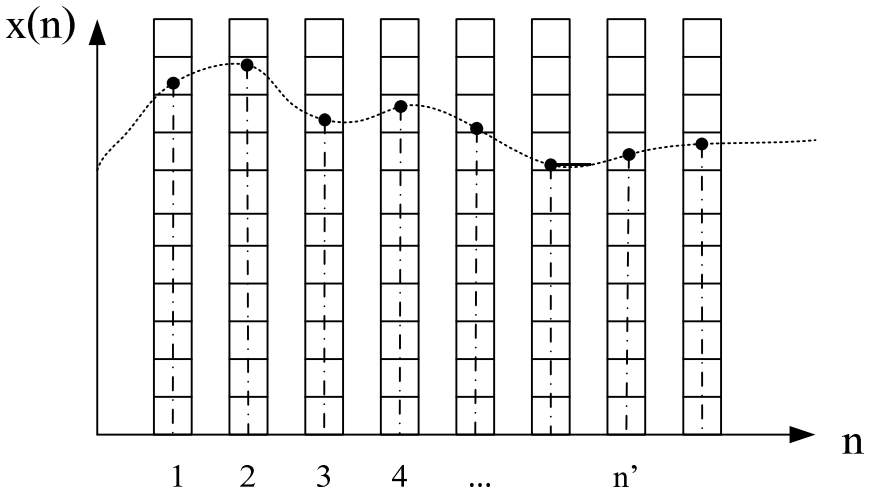
f_s คือ ค่าความถี่ในการสุ่มตัวอย่าง

f_0 คือ ค่าความถี่สูงสุดที่มีอยู่ในสัญญาณข้อมูล

จากสมการที่ (2-1) ถ้าหากค่าความถี่ในการสุ่มตัวอย่างมีค่าเท่ากับ $2f_0$ จะเรียกค่าความถี่นี้ว่า ความถี่ไนควิสต์ (Nyquist Frequency) และเรียกคาบเวลา $\frac{1}{2f_0}$ ว่าเป็นช่วงเวลาการสุ่มของไนควิสต์ ถ้าหากค่าความถี่ในการสุ่มสัญญาณมีค่าน้อยกว่า $2f_0$ จะทำให้มีความผิดพลาดของสัญญาณ (Aliasing) เกิดขึ้นเมื่อนำค่าแปลงจากสัญญาณเชิงตัวเลขกลับไปเป็นสัญญาณอนาล็อก (Digtal to Analog Convert: (D/A)) ทำให้ไม่สามารถนำสัญญาณอนาล็อกเดิมกลับมาได้

2.3.2 การควอนไตซ์ (Quantization)

เนื่องจากคอมพิวเตอร์เก็บข้อมูลในรูปของหน่วยข้อมูลที่เป็นบิต (Bit) ซึ่งทำให้มีความยาวจำกัดในการเก็บข้อมูลดังแสดงในรูปที่ 2.6 ดังนั้นขนาดข้อมูลที่ได้จากการสุ่มข้อมูลในแต่ละจุดจะจัดเก็บในหน่วยข้อมูล ขึ้นอยู่กับความละเอียดของข้อมูลที่ต้องการ ถ้าต้องการความละเอียดของข้อมูลสูงก็ใช้จำนวนบิตในการจัดเก็บข้อมูลสูง แต่ถ้าต้องการความละเอียดต่ำก็ใช้จำนวนบิตในการจัดเก็บน้อย



รูปที่ 2.6 การควอนไตซ์

2.3.3 การคำนวณหาค่าพลังงาน (Energy Detection)

การคำนวณหาค่าพลังงานของเสียงสามารถทำได้โดยการพิจารณาค่าผลรวมของแต่ละจุด (Sample) ที่อยู่ในช่วงสัญญาณเสียงที่ต้องการ ทั้งนี้เนื่องจากค่าในแต่ละจุดนั้นอาจจะมีค่าเป็นได้ทั้งบวกและลบ ดังนั้นก่อนการหาผลรวมจะต้องทำให้เป็นบวกทั้งหมดเสียก่อน ดังแสดงในสมการที่ (2-2)

$$E = \sum_{i=1}^{N} | s(i) | \qquad (2-2)$$

โดยที่

$s(i)$ คือ สัญญาณเสียงลำดับที่ i ในช่วงที่ต้องการหาค่าพลังงาน

N คือ จำนวน Sample ทั้งหมดในช่วงที่ต้องการหาค่าพลังงาน

E คือ ค่าพลังงานรวมทั้งหมด

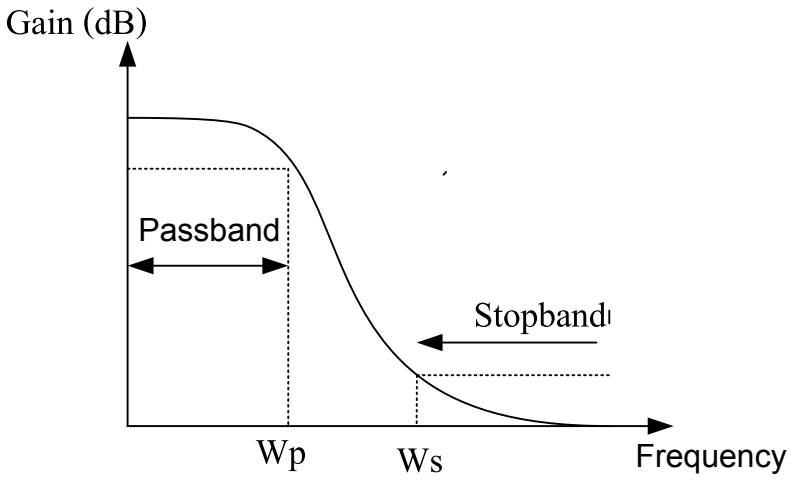
2.3.4 ตัวกรอง(Filter)

ตัวกรอง คือ สิ่งที่ยอมให้สัญญาณความถี่ที่ต้องการผ่านและกำจัดสัญญาณที่มีความถี่ที่ไม่ต้องการออก ตัวกรองส่วนมากจะใช้ในการกำจัดสัญญาณรบกวน (Noise) และลดความผิดเพี้ยนของสัญญาณ เนื่องจากตัวกรองมีวิธีการออกแบบที่ไม่ยากแต่สามารถใช้งานได้อย่างมีประสิทธิภาพ ดังนั้นตัวกรองจึงถูกนำมาประยุกต์มาใช้ในการประมวลผลสัญญาณ

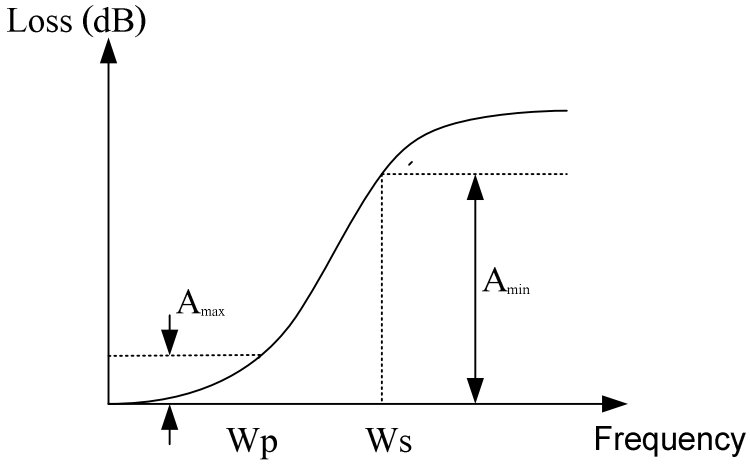
ตัวกรองสามารถแบ่งได้ 2 ระบบตามสัญญาณที่ป้อนเข้ามาคือ ระบบตัวกรองอนาล็อก (Analog Filter) และระบบตัวกรองเชิงตัวเลข หรือระบบตัวกรองดิจิตอล (Digital Filter) นอกจากนี้เราอาจจะแบ่งตัวกรองตามลักษณะฟังก์ชันถ่ายโอน (Transfer Function) ได้ 4 แบบ คือ

2.3.4.1 ตัวกรองผ่านความถี่ต่ำ (Lowpass Filter: LPF)

ตัวกรองผ่านความถี่ต่ำมีคุณสมบัติพื้นฐานคือ ผ่านความถี่ต่ำและบั่นทอนความถี่สูง นอกจากนี้ตัวกรองความถี่ต่ำจะต้องผ่านสัญญาณ DC จนถึง cutoff frequency, Wp ซึ่งอยู่ในช่วง Passband จะสูญเสียไม่เกิน A_{max} dB และความถี่ที่สูงกว่า Ws จนถึงอนันต์จะสูญเสียอย่างต่ำ A_{min} dB ดังแสดงในรูปที่ 2.7 และรูปที่ 2.8



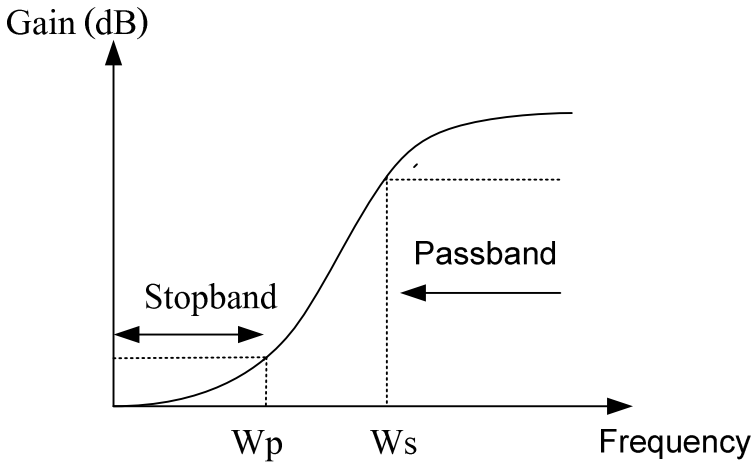
รูปที่ 2.7 ผลตอบสนองของตัวกรองผ่านความถี่ต่ำ



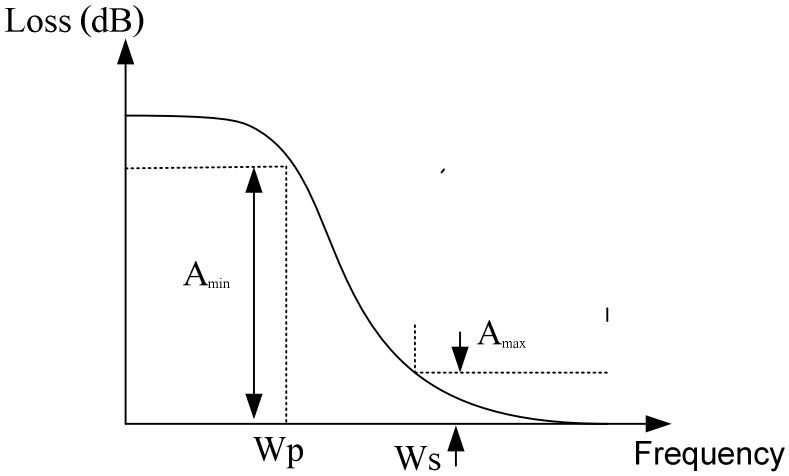
รูปที่ 2.8 ผลการสูญเสียของตัวกรองผ่านความถี่ต่ำ

2.3.4.2 ตัวกรองผ่านความถี่สูง (Highpass Filter: HPF)

ตัวกรองผ่านความถี่สูงจะผ่านสัญญาณที่มีความถี่สูงกว่า Wp โดยจะมีการสูญเสียไม่เกิน A_{max} และความถี่ DC ถึง Ws จะมีการสูญเสียอย่างน้อย A_{min} ดังแสดงในรูปที่ 2.9 และ รูปที่ 2.10



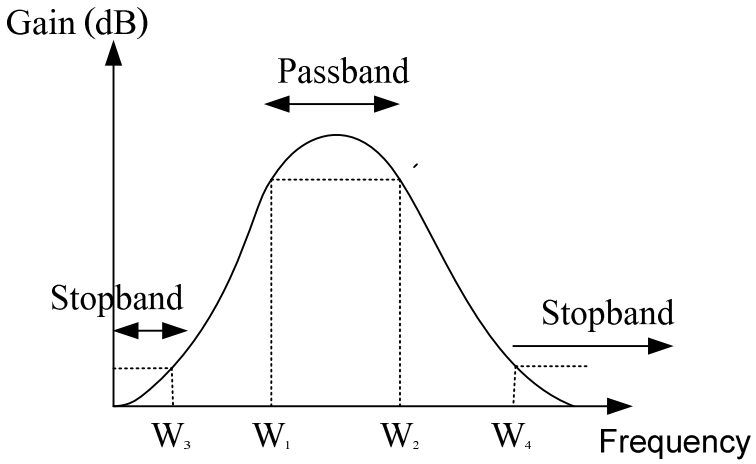
รูปที่ 2.9 ผลตอบสนองของตัวกรองผ่านความถี่สูง



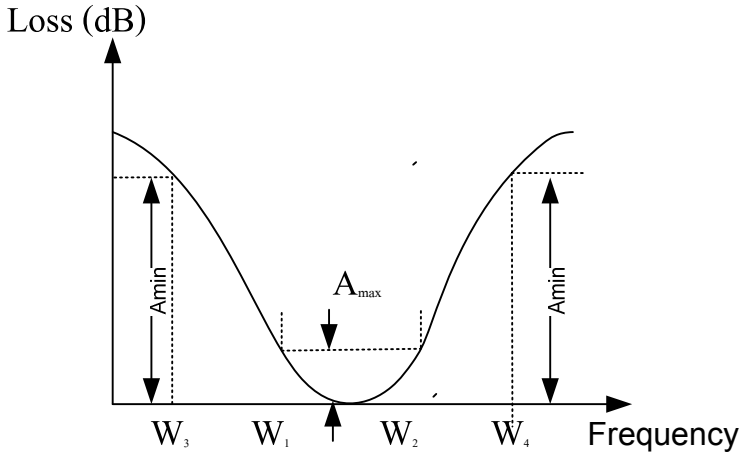
รูปที่ 2.10 ผลการสูญเสียของตัวกรองผ่านความถี่สูง

2.3.4.3 ตัวกรองผ่านแถบความถี่ (Bandpass Filter: BPF)

ตัวกรองผ่านแถบความถี่จะผ่านช่วงความถี่ระหว่าง W_1 ถึงความถี่ W_2 ด้วยการสูญเสียไม่เกิน A_{max} และความถี่ที่น้อยกว่า W_3 และความถี่ที่มากกว่า W_4 จะมีการสูญเสียอย่างน้อย A_{min} ดังแสดงในรูปที่ 2.11 และ รูปที่ 2.12



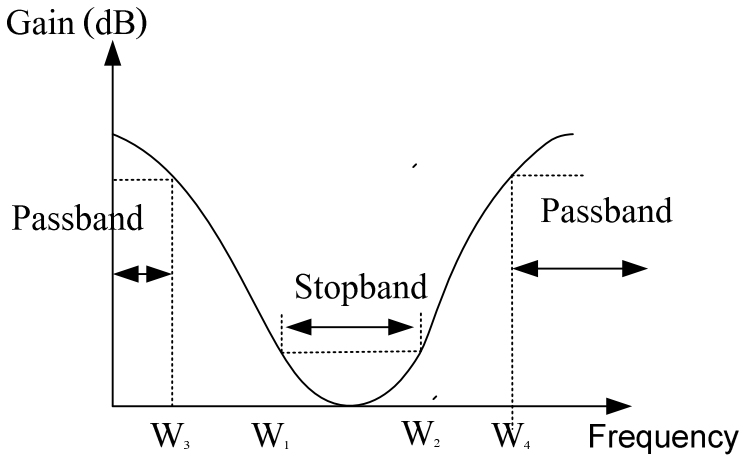
รูปที่ 2.11 ผลตอบสนองของตัวกรองผ่านแถบความถี่



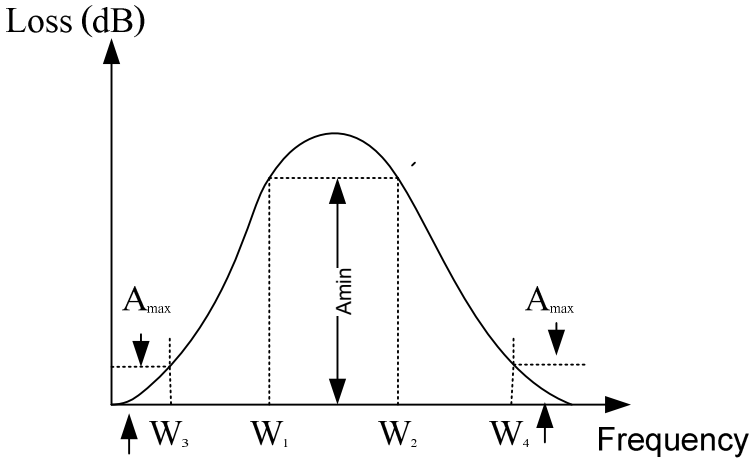
รูปที่ 2.12 ผลการสูญเสียของตัวกรองผ่านแถบความถี่

2.3.4.4 ตัวกรองหยุดแถบความถี่ (Bandstop Filter: BSF)

ตัวกรองหยุดแถบความถี่จะผ่านความถี่ที่ต่ำกว่า W_3 และความถี่ที่สูงกว่า W_4 ด้วยการสูญเสียไม่เกิน A_{max} และช่วงความถี่ระหว่าง W_1 ถึง W_2 จะมีการสูญเสียอย่างน้อย A_{min} ดังแสดงในรูปที่ 2.13 และ รูปที่ 2.14



รูปที่ 2.13 ผลตอบสนองของตัวกรองหยุดแถบความถี่



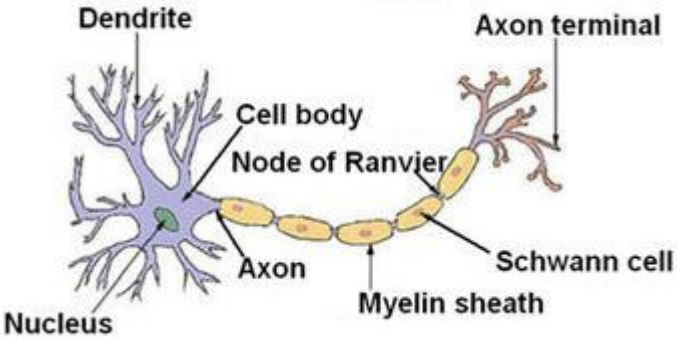
รูปที่ 2.14 ผลการสูญเสียของตัวกรองหยุดแถบความถี่

2.4 โครงข่ายประสาทเทียม

2.4.1 หลักการและแนวคิดของโครงข่ายประสาทเทียม

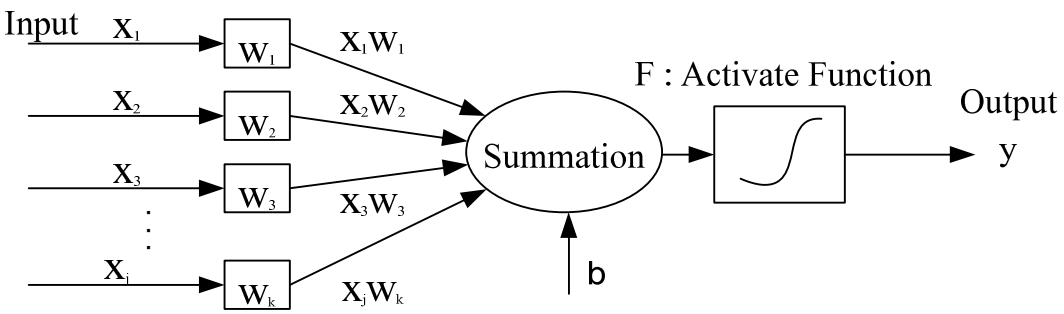
โครงข่ายประสาทเทียม (Artificial Neural Network) [27][28] คือแบบจำลองทางคณิตศาสตร์ จำลองการทำงานของเครือข่ายประสาทในสมองมนุษย์ ใช้ในการประมวลผลสารสนเทศด้วยการคำนวณแบบคอนเนคชันนิสต์ (Connectionist) ซึ่งมีความสามารถในการเรียนรู้และจดจำรูปแบบ (Pattern Recognition) และการอุปมานความรู้ (Knowledge Deduction) เช่นเดียวกับความสามารถที่มีในสมองของมนุษย์

แนวคิดของโครงข่ายประสาทเทียมนี้ได้มาจากการศึกษาข่ายงานไฟฟ้าชีวภาพ (Bioelectric Network) ในสมอง ซึ่งประกอบด้วย เซลล์ประสาท หรือ “นิวรอน” (Neurons) และ จุดประสานประสาท (Synapses) แต่ละเซลล์ประสาทประกอบด้วยปลายในการรับกระแสประสาท เรียกว่า “เดนไดรท์” (Dendrite) ซึ่งเป็น input และปลายในการส่งกระแสประสาทเรียกว่า “แอคซอน” (Axon) ซึ่งเป็นเหมือน output ของเซลล์ ดังแสดงในรูปที่ 2.15 เซลล์เหล่านี้ทำงานด้วยปฏิกิริยาไฟฟ้าเคมี เมื่อมีการกระตุ้นด้วยสิ่งเร้าภายนอกหรือกระตุ้นด้วยเซลล์ด้วยกัน กระแสประสาทจะวิ่งผ่านเดนไดรท์เข้าสู่นิวเคลียสซึ่งจะเป็นตัวตัดสินว่าต้องกระตุ้นเซลล์อื่น ๆ ต่อหรือไม่ ถ้ากระแสประสาทแรงพอ นิวเคลียสก็จะกระตุ้นเซลล์อื่น ๆ ต่อไปผ่านทางแอคซอน



รูปที่ 2.15 โครงสร้างของเซลล์ประสาท

หลักการของโครงข่ายประสาทเทียมที่ใช้ในระบบคอมพิวเตอร์นั้น เซลล์ประสาท (Neurons) จะประกอบด้วยส่วน input และ output เช่นเดียวกัน โดยจำลองให้ input แต่ละอันมีตัวกำหนดน้ำหนัก (Weight) ของ input โดยเซลล์ประสาทแต่ละหน่วยจะมีค่าอ้างอิง (Threshold) เพื่อเป็นตัวกำหนดว่าค่าน้ำหนักรวมของ input ต้องมากเพียงพอที่จะสามารถส่งผลลัพธ์ไปยังเซลล์ประสาทตัวอื่นได้หรือไม่ เมื่อนำเซลล์ประสาทแต่ละหน่วยมาต่อและทำงานร่วมกัน การทำงานนี้ในทางตรรกแล้วก็จะเหมือนกับปฏิกิริยาเคมีที่เกิดในสมอง เพียงแต่ในระบบคอมพิวเตอร์ทุกอย่างเป็นตัวเลขเท่านั้นเองดังแสดงในรูปที่ 2.16



รูปที่ 2.16 โครงสร้างแบบจำลองเซลล์ประสาทในระบบคอมพิวเตอร์

จากรูปที่ 2.16 จะได้สมการค่าผลลัพธ์ของหน่วยประสาทเทียมได้สมการที่ (2-3) คือ

$$y = F(x_1 w_1 + x_2 w_2 + x_3 w_3 + ... + x_j w_k + b) \qquad (2\text{-}3)$$

โดยที่

b คือค่าเอนเอียงของหน่วยประสาทเทียม

F คือฟังก์ชันการแปลงถ่ายทอดข้อมูล (Activate Function)

y คือผลลัพธ์ของหน่วยประสาทเทียม

$x_1, x_2, x_3, ..., x_j$ คือค่าข้อมูลป้อนเข้าสู่หน่วยประสาทเทียม

$w_1, w_2, w_3, ..., w_k$ คือค่าถ่วงน้ำหนักของข้อมูลเข้าหน่วยประสาทเทียม

หรืออาจจะเขียนในรูปของเมตริกซ์ดังสมการที่ (2-4) คือ

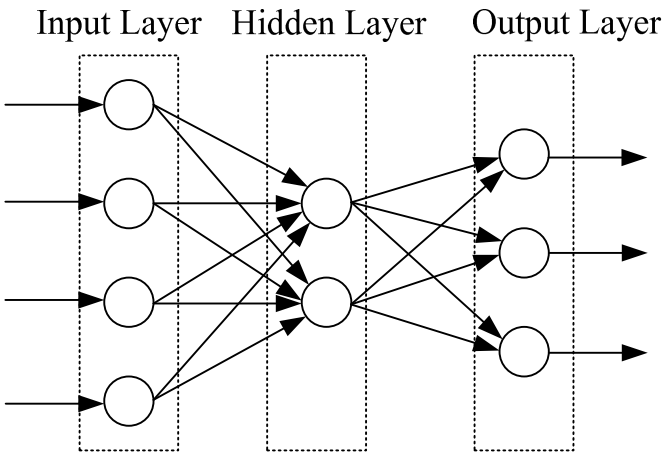
$$y = F(Wx + b) \qquad (2\text{-}4)$$

เมื่อกำหนดให้ Wx คือเมตริกซ์ของผลคูณระหว่างค่าถ่วงน้ำหนักและข้อมูลที่ป้อนเข้าสู่โครงข่ายประสาทเทียมในแต่ละอินพุต

สำหรับการเรียนรู้ของหน่วยประสาทเทียมนั้นทำได้โดยการป้อนค่าตัวอย่างข้อมูลเพื่อให้หน่วยประสาทเทียมทำการปรับค่าถ่วงน้ำหนัก (Weight) แต่ละตัวให้เหมาะสมเพื่อให้ได้ผลลัพธ์ที่ให้ค่าใกล้เคียงกับค่าผลลัพธ์ที่ต้องการมากที่สุด

2.4.2 โครงข่ายประสาทเทียมแบบ Feed Forward Network (Multilayer Perceptron)

ในการทำงานเราสามารถนำหน่วยประสาทเทียมแต่ละหน่วยมาต่อเชื่อมกันเป็นโครงข่ายประสาทเทียมเพื่อให้ทำงานร่วมกันได้ โดยในการนำมาต่อเชื่อมกันนั้น สามารถเชื่อมต่อหน่วยเซลล์ประสาทได้เป็นหลาย ๆ ชั้น (MutiLayer) ในแต่ละชั้นจะประกอบไปด้วยเซลล์ประสาทเทียมซึ่งอาจจะมีจำนวนเซลล์ประสาทเทียมไม่เท่ากันก็ได้และรวมถึงชนิดของฟังก์ชันการแปลงถ่ายทอดข้อมูลของในแต่ละชั้นอาจจะไม่เหมือนกันก็ได้เช่นกัน ดังแสดงในรูปที่ 2.17



รูปที่ 2.17 โครงข่ายประสาทเทียมแบบหลายชั้น

หน่วยประสาทเทียมในชั้นที่รับข้อมูลอินพุตเข้ามานั้นจะเรียกว่า Input Layer และชั้นที่ให้ผลลัพธ์ออกมาเรียกว่าชั้น Output Layer ส่วนชั้นที่อยู่ระหว่างกลางเรียกว่าชั้น Hidden Layer ซึ่งในการทำงานจริงนั้นชั้น Hidden layer อาจจะมีมากกว่า 1 ชั้นก็ได้ ทั้งนี้ขึ้นอยู่กับความซับซ้อนของงาน

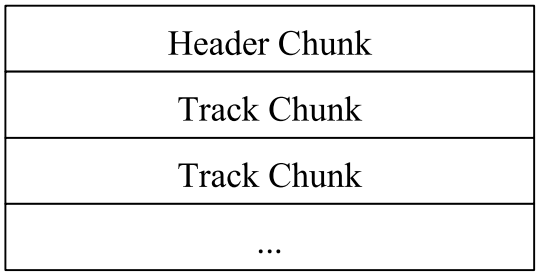
จากรูปที่ 2.17 สังเกตุได้ว่าข้อมูลที่ป้อนเข้าไปนั้นจะส่งผ่านจากเซลล์ประสาทเรียงไปในทิศทางเดียวตามลำดับชั้น คือ ส่งจากชั้น Input Layer ต่อไปยังชั้น Hidden Layer และได้ผลลัพธ์ออกมาที่ชั้น Output Layer โดยไม่มีการป้อนข้อมูลย้อนกลับและแม้แต่เซลล์ประสาทที่ต่ออยู่ในชั้นเดียวกันก็ไม่มีการเชื่อมต่อถึงกัน เราเรียกการเชื่อมต่อโครงข่ายประสาทเทียมแบบนี้ว่า Feed-Forward Network

2.5 โครงสร้างของไฟล์มีดี (MIDI Structure)

ไฟล์มีดี (MIDI : Music Instrument Digital Interface) เป็นมาตรฐานของอุตสาหกรรมดนตรีแบบอิเล็กทรอนิกส์ ซึ่งจะใช้สำหรับการส่งและแลกเปลี่ยนสัญญาณเสียงในรูปแบบที่อุปกรณ์อิเล็กทรอนิกส์สามารถใช้งานได้ โดยจะเป็นเทคโนโลยีที่เปรียบเสมือนการเก็บโน้ตเพลง เนื่องจากข้อมูลแบบ MIDI จะเป็นคำสั่งในการสังเคราะห์เสียงแทนที่จะเป็นเสียงเพลงจริง ๆ และจะใช้อุปกรณ์ ซินธิไซเซอร์ (Sythesizer) ในการรับคำสั่งจากข้อมูล MIDI ทำให้สามารถแก้ไขหรือปรับแต่งเพลงได้ทีละตัวโน้ต รวมทั้งสามารถปรับแต่งจังหวะได้โดยไม่กระทบกระเทือนถึงระดับเสียงของตัวโน้ต เนื่องจากรูปไฟล์เสียงในรูปแบบของมีดีนั้นไม่ได้เป็นการเก็บข้อมูลทั้งหมดของเสียงในแต่ละจุด (Sample) แต่จะเก็บในรูปแบบของสัญลักษณ์ของตัวโน้ตและข้อมูลจำเป็นอื่น ๆ ตามรูปแบบมาตรฐาน ที่ได้กำหนดไว้ใน [29]

Chunk คือ กลุ่มของข้อมูล ไฟล์มีดี อาจประกอบไปด้วย chunk ได้หลาย ๆกลุ่มซึ่งแต่ละ

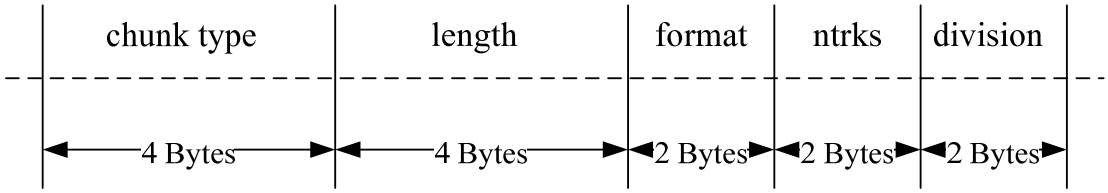
กลุ่มจะมีขนาดแตกต่างกันไป โดยในแต่ละ chunk จะเริ่มต้นด้วยข้อมูลขนาด 4 ไบต์ที่จัดอยู่ในรูปของรหัสแอสกี (ASCII) ซึ่งรหัสนี้จะเป็นตัวบอกชนิดของ chunk สำหรับโครงสร้างของไฟล์มีดีนั้นจะประกอบไปด้วย 2 ส่วนคือ Header chunk กับ Track chunk ดังแสดงในรูปที่ 2.18



รูปที่ 2.18 โครงสร้างของไฟล์มีดี (MIDI Structure)

2.5.1 Header Chunk

Header chunk จะเป็นส่วนเริ่มต้นของไฟล์มีดีซึ่งจะใช้ในการระบุข้อมูลพื้นฐานบางอย่างเกี่ยวกับข้อมูลในไฟล์ โดยรูปแบบที่สมบูรณ์ของ Header chunk แสดงดังในรูปที่ 2.19



รูปที่ 2.19 โครงสร้างของ Header Chunk

จากโครงสร้างของ Header Chunk ในรูปที่ 2.19 จะเห็นได้ว่าไฟล์มีดีจะมี Header Chunk ขนาด 14 ไบต์ ซึ่งมีรายละเอียดของฟิลด์ต่างๆ ดังนี้

- Chunk type เป็น การระบุชนิดของ chunk โดยจะแสดงในรูปแบบของตัวอักษร ASCII 4 ตัวคือ ‘MThd’
- Length เป็นขนาดข้อมูล 32 บิตใช้ในการระบุจำนวนข้อมูลที่ตามหลังมา โดยปกติแล้วจะเป็นเลข 6 ซึ่งหมายความว่ามีข้อมูลขนาด 6 ไบต์ตามมา
- Format เป็นการระบุรูปแบบของไฟล์ มีขนาด 2 ไบต์ รูปแบบของไฟล์ MIDI ซึ่งมีด้วยกัน 3 รูปแบบดังนี้

  รูปแบบที่ 0 คือ 00 00 ซึ่งหมายถึง ไฟล์รูปแบบที่ 0 จะมีได้ 16 channels, 1 track

  รูปแบบที่ 1 คือ 00 01 ซึ่งหมายถึง ไฟล์รูปแบบที่ 1 จะมีได้ 1 channel, มีได้หลาย tracks และสามารถเล่นเสียงแต่ละ track ได้พร้อมกัน

รูปแบบที่ 2 คือ 00 02 ซึ่งหมายถึง ไฟล์รูปแบบที่ 2 สามารถมีได้หลาย tracks แต่ต้องเล่นเสียงตามลำดับ tracks

- ntrks ใช้ในการระบุจำนวนของ tracks ในไฟล์ ดังนั้นถ้าหากเป็นMIDI ที่เป็นรูปแบบที่ 0 จะมีจำนวน track เท่ากับ 1 นั่นเอง
- Division ใช้ในการระบุจำนวนตัวหาร ควอร์เตอร์โน้ต (Quarter Note, โน้ตตัวดำ) ซึ่งจะใช้ในคำนวณเวลาของตัวโน้ต ในการระบุความหมายของ delta-times มีได้สองรูปแบบดังแสดงในคือ

  รูปแบบแรก บิตที่ 15 มีค่าเป็น 0 ส่วน 14 บิตที่เหลือแสดงจำนวน ticks ต่อหนึ่งโน้ตตัวดำ

  รูปแบบที่สอง บิตที่ 15 มีค่าเป็น1 ส่วน 14 บิตที่เหลือจะประกอบไปด้วยสองส่วน คือ SMPTE และ MIDI Time Code

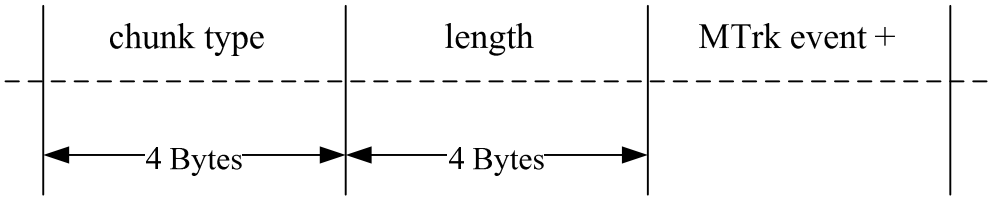
|15 |14 8 | 7 0|

0	ticks per quarter note	
1	negative SMPTE format	Ticks per frame

รูปที่ 2.20 การระบุความหมายของ Delta-times

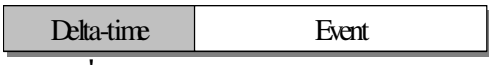
2.5.2 Track Chunk

Track chunk จะบรรจุข้อมูล MIDI ต่อหนึ่ง track ซึ่งแต่ละ track จะจัดเก็บข้อมูลอยู่ในลำดับของ MIDI events โดยรูปแบบของ Track chunk แสดงดังรูปที่ 2.21



รูปที่ 2.21 โครงสร้างของ Track chunk

สำหรับเครื่องหมาย '+' หมายถึงสามารถมีได้อย่างน้อยตั้งแต่ 1 ขึ้นไปและสำหรับรูปแบบของ MTrk event จะประกอบไปด้วย 2 ส่วนคือส่วนของ delta-time และ event ดังแสดงในรูปที่ 2.22



รูปที่ 2.22 รูปแบบของ Mtrk event

จากโครงสร้างของ Track Chunk ในรูปที่ 2.21และรูปที่ 2.22 จะเห็นได้ว่าไฟล์มีดีจะมี Track Chunk ซึ่งมีรายละเอียดของฟิลด์ต่าง ๆ ดังนี้

- Chunk type เป็น การระบุชนิดของ chunk โดยจะแสดงในรูปแบบของตัวอักษร ASCII 4 ตัวคือ 'MTrk'
- Length เป็นขนาดข้อมูล 32 บิตใช้ในการระบุจำนวนข้อมูลที่อยู่ในแต่ละ track
- delta-time ใช้สำหรับกำหนดค่าจำนวนเวลาก่อนที่ event ถัดมาจะเกิดขึ้น ตัวอย่างเช่น ถ้าต้องการให้ event แรกเกิดขึ้นที่จุดเริ่มต้นของ track หรือ ให้ event ที่สองเกิดขึ้นตามมาทันที ให้กำหนดค่า delta-time เท่ากับ 0

  เนื่องจากต้องการให้ไฟล์มีดีมีขนาดเล็กที่สุดเท่าที่จะเป็นไปได้ ดังนั้นในการจัดเก็บข้อมูลของ delta-time จึงอาจจะเป็นไปได้ตั้งแต่ 1, 2, 3 หรือ 4 ไบต์ก็ได้ ซึ่งจะเป็นการจัดอยู่ในรูปแบบของ Variable-Length Quantity โดยบิตสูงสุดของแต่ละไบต์จะถูกกำหนดให้เป็น 1 ยกเว้นไบต์สุดท้าย (ไบต์ 0) เพราะฉะนั้นช่วงเวลาที่เป็นไปได้ในแต่ละไบต์คือ 0-127 (00-7F) แต่ถ้าช่วงเวลาเป็น 128 ก็จะต้องใช้ข้อมูล 2 ไบต์ในการจัดเก็บนั่นคือ '81' '00' บิตสูงของไบต์แรกถูกกำหนดให้เป็น 1 เพื่อบอกให้ทราบว่าไม่ใช่ไบต์สุดท้าย ดังตัวอย่าง เช่น

  ฐาน 10→ ฐานสอง-------→ Variable-Length Quantity

  127--→01111111-----→7F

  128--→10000000-----→10000001 00000000 ('81' '00')

- event ใช้ในการควบคุมการทำงานซึ่งประกอบไปด้วย 3 ส่วนดังนี้คือ

  <event> = <MIDI event> | <sysex event>|<meta-event>

- MIDI event ใช้ในการแสดงเหตุการณ์ที่ต้องการ เช่นการเล่นโน้ต (Note on), การหยุดเล่นโน้ต (Note off) เป็นต้น รวมถึงหมายเลขโน้ต (Note number) ของ MIDI channel ที่เกี่ยวข้องด้วย MIDI channel คือ ชนิดของเสียงเครื่องดนตรี โดยระบบ MIDI สามารถส่งสัญญาณได้พร้อมกัน 16 channels นั่นคือ เราสามารถเล่นเสียงดนตรีได้พร้อมกันถึง 16 ชนิด โดยปกติแล้วนั้น channel ที่ 1 – 9 และ channel 11-16 จะใช้สำหรับเสียงเครื่องดนตรีที่เป็นทำนองเพลง (melody) ทั่วไป ส่วน channel ที่ 10 ใช้สำหรับเสียงกลอง อย่างไรก็ตามอาจจะกำหนด channel ที่แตกต่างไปได้

- sysex event ใช้ในการกำหนดข้อความหรือคำสั่งพิเศษบางอย่าง
- meta-event ใช้ในการระบุข้อมูลที่เป็น non-MIDI เช่น การปิด track

2.6 สรุป

ในบทนี้ได้กล่าวถึงหลักการพื้นฐานและคุณสมบัติของเสียงดนตรี, ลักษณะดนตรีไทยและดนตรีสากล โดยได้มุ่งเน้นไปที่ ระนาดเอก ซึ่งจะเห็นได้ว่าเครื่องดนตรีไทยมีความแตกต่างจากเครื่องดนตรีสากลในหลาย ๆด้าน นอกจากนี้ยังได้กล่าวอธิบายทฤษฎีที่ใช้ในการประมวลผลสัญญาณดิจิตอลและหลักการของตัวกรองความถี่ประเภทต่าง ๆ ซึ่งข้อมูลเหล่านี้จะได้นำไปใช้ในการพัฒนาระบบรู้จำโน้ตเสียงของระนาดเอก

สำหรับโครงข่ายประสาทเทียมที่ได้กล่าวอธิบายนั้นจะได้นำไปประยุกต์ใช้ในการตัดสินใจระบุค่าตัวโน้ตและจัดแสดงผลลัพธ์ให้ออกมาอยู่ในรูปแบบของไฟล์มิดี ซึ่งเนื้อหาเกี่ยวกับการออกแบบระบบสำหรับระบุตัวโน้ตจากเสียงระนาดเอกจะได้กล่าวในบทที่ 3 และผลลัพธ์จากการทดลองจะได้กล่าวไว้ในบทที่ 4 ตามลำดับ