สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การพัฒนาโปรแกรมสำหรับเล่นดนตรี ตามแผ่นโน้ตดนตรีที่ผู้ฟังป้อนให้

MOZART: MUSIC PLAYER FROM PROVIDE NOTE PAPER

T055082

นาย ณัฐพล ปัญญาไชย
นาย สุรเดช ปาลี

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2546

เลขหมู่.....
เลขทะเบียน 55082
วัน,เดือน,ปี - 8 มิ.ย. 2548

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

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

การพัฒนาโปรแกรมสำหรับเล่นดนตรี ตามแผ่นโน้ตดนตรีที่ผู้ฟังป้อนให้

MOZART: MUSIC PLAYER FROM PROVIDE NOTE PAPER

โดย

นาย ณัฐพล ปัญญาไชย

นาย สุรเดช ปาลี

อาจารย์ที่ปรึกษา

ดร. อรฉัตร จิตต์โสภักตร์

อ. เจริญ วงษ์ชุ่มเย็น

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2546

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์ปีการศึกษา 2546

ภาควิชา วิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การพัฒนาโปรแกรมสำหรับเล่นดนตรี ตามแผ่นโน้ตดนตรีที่ผู้ฟังป้อนให้

MOZART: MUSIC PLAYER FROM PROVIDE NOTE PAPER

คณะผู้จัดทำ นาย ณัฐพล ปัญญาไชย รหัสประจำตัว 44015323

นาย สุรเดช ปาลี รหัสประจำตัว 44015360

อรฉัตร จิตต์ ..... อาจารย์ที่ปรึกษา

(ดร. อรฉัตร จิตต์โสภักตร์)

..... อาจารย์ที่ปรึกษา

(อ. เจริญ วงษ์ชุ่มเย็น)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การพัฒนาโปรแกรมสำหรับเล่นดนตรี ตามแผ่นโน้ตดนตรีที่ผู้ฟังป้อนให้

นาย ณัฐพล ปัญญาไชย 44015323
นาย สุรเดช ปาลี 44015360
ดร. อรฉัตร จิตต์โสภักตร์ อาจารย์ที่ปรึกษา
อ. เจริญ วงษ์ชุ่มเย็น อาจารย์ที่ปรึกษา
ปีการศึกษา 2546

บทคัดย่อ

วิวัฒนาการของเทคโนโลยีต่างๆล้วนมีบทบาทที่สำคัญต่อชีวิตประจำวันของมนุษย์เรา ทุกอย่างย่อมมีการพัฒนาให้มีความสะดวกสบายในการใช้งานมากขึ้น แน่นอน ดนตรี ก็เป็นหนึ่งในหลายๆสิ่งเหล่านั้นที่มีการพัฒนามาอย่างต่อเนื่อง ซึ่งเมื่อเรากล่าวถึงดนตรีแล้ว ย่อมหลีกเลี่ยงไม่ได้ที่จะพูดถึงโน้ตดนตรี อันเป็นส่วนสำคัญที่จะทำให้เราสามารถสร้างสรรค์เสียงเพลง หรือแทนเสียงเพลงเหล่านั้นออกมาเป็นสัญลักษณ์ และนี่เองที่เป็นเหตุผลสำคัญของโครงงานในครั้งนี้ ที่จะได้นำเอาเทคนิคและวิธีการต่างๆมาประยุกต์ใช้ เพื่อเป็นทางเลือกหนึ่งในการให้ความสะดวกสบายเกี่ยวกับเสียงดนตรี

โครงงานนี้เป็นการพัฒนาโปรแกรมสำหรับประมวลผลภาพโน้ตดนตรีที่รับเข้ามา แล้วแปลงให้อยู่ในรูปของไฟล์มิดี้ โดยการทำงานของโปรแกรมโดยรวมนั้นจะแบ่งออกเป็น 3 ส่วนหลักๆคือ ส่วนการทำพรีโพรเซสซิ่ง, ส่วนการประมวลผลและรู้จำรูปแบบ และสุดท้ายส่วนสร้างไฟล์มิดี้ ซึ่งการทำงานนั้นจะเริ่มจากนำภาพโน้ตดนตรีมาผ่านกระบวนการทำพรีโพรเซสซิ่ง ต่อมาภาพจะถูกทำให้บรรทัด5เส้นหายไปเพื่อให้เหลือแต่โน้ตและสัญลักษณ์ที่จะทำการวิเคราะห์ จากนั้นภาพจะถูกส่งไปยังส่วนการประมวลผล และรู้จำรูปแบบ เพื่อวิเคราะห์ตัวโน้ตและสัญลักษณ์ต่างๆ สุดท้ายข้อมูลที่ได้นี้จะถูกนำไปสร้างเป็นไฟล์ มิดี้นั่นเอง นอกจากนี้จะมีการนำเสนอให้เห็นว่าเทคนิคต่างๆที่ได้ทำการศึกษาและทดลองมานั้นแต่ละแบบมีข้อดี หรือข้อเสียอย่างไร เหตุใดจึงเลือกเอาเทคนิคนั้นๆมาใช้ รวมทั้งผลการทดสอบโปรแกรมที่ได้ทำการพัฒนา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

MOZART: MUSIC PLAYER FROM PROVIDE NOTE PAPER

Mr. Nattapon Panyachai 44015323

Mr. Suradech Palee 44015360

Dr. Orachat Chitsobhuk Advisor

Mr. Charoen Vongchumyen Advisor

Academic Year 2003

ABSTRACT

Recently, Technology has become a part of life and it has been applied to satisfy human such as musical technology. Because musical note is still the standard for musicians, so we think of technology that applied by pattern recognition to analyze the structure of printed musical note. After that the computer can recognize musical note and play music from these data correctly.

The purpose of the project is to develop a program that transforms musical images received to midi files. The program is separated into 3 main parts that are preprocessing, processing and recognizing the patterns, and creating midi files. The first step is preprocessing musical images received. Then, the preprocessed images will be processed and recognized in order to analyze notes and symbols. Finally, midi files will be created from the processed information. Moreover, the project concludes with the advantages and disadvantages among other techniques, the reason why we choose those techniques and benchmark of the developed program.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ปริญญานิพนธ์ฉบับนี้สามารถสำเร็จลุล่วงได้เป็นอย่างดี ด้วยคำแนะนำ คำปรึกษา และได้รับการดูแลจากหลายๆฝ่ายด้วยกัน แน่นอนว่าแรงผลักดันสำคัญอันทำให้ปริญญานิพนธ์นี้สามารถสำเร็จลุล่วงได้นั่น ก็ด้วยความกรุณาจากท่านอาจารย์ที่ปรึกษา ดร. อรฉัตร จิตต์โสภักตร์ และ อ. เจริญ วงษ์ชุ่มเย็น ที่คอยประสิทธิประสาทวิชาความรู้ ให้ความเอาใจใส่ ห่วงใย และให้ความช่วยเหลือเสมอมา ซึ่งข้าพเจ้ารู้สึกทราบซึ้งในความอนุเคราะห์และขอขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณภาควิชาวิศวกรรมคอมพิวเตอร์ที่คอยจัดเตรียมสิ่งอำนวยความสะดวกต่างๆ ไม่ว่าจะเป็นห้องวิจัย อินเทอร์เน็ตความเร็วสูงสำหรับค้นคว้าหาข้อมูล หรือความอนุเคราะห์อื่นๆ เพื่อให้การดำเนินงานนั้นเป็นไปอย่างต่อเนื่อง ซึ่งท้ายที่สุดก็ได้เป็นปริญญานิพนธ์นี้ออกมา

ขอขอบคุณคุณรุ่นพี่ที่คอยให้คำปรึกษา เพื่อนๆที่คอยให้กำลังใจและเคียงข้างกันเสมอมา ไม่ว่าวันเวลาจะผ่านไปนานเท่าไรสิ่งหนึ่งที่ยังคงมีอยู่และจะยังคงมีไว้ตราบนานเท่านานคือมิตรภาพระหว่างเราและที่เห็นทีจะขาดเสียมิได้ ขอขอบคุณห้องฮาร์ดแวร์อันเปรียบเสมือนบ้านอีกหลังหนึ่งของเรา

สุดท้ายนี้ทุกสิ่งทุกอย่างที่กล่าวมาคงไม่อาจเกิดขึ้น หากข้าพเจ้าไม่ได้มายืนอยู่ตรงจุดนี้ ขอกราบขอบพระคุณบุพการีผู้ให้กำเนิด ที่คอยเลี้ยงดู อบรมสั่งสอน ที่สำคัญได้ให้โอกาสทางการศึกษาแก่ข้าพเจ้าอันทำให้มีวันนี้ วันที่เราประสบความสำเร็จ และความภาคภูมิใจเป็นของเราทุกคน

คุณค่าและประโยชน์อันพึงมีจากปริญญานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแด่ผู้มีพระคุณทุกท่าน

ณัฐพล ปัญญาไชย

สุรเดช ปาลี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้าที่
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VII
สารบัญภาพ	VIII
บทที่ 1 บทนำ	1
1.1 หลักการและเหตุผล	1
1.2 วัตถุประสงค์	1
1.3 ผลที่คาดว่าจะได้รับ	2
1.4 ขอบเขตของการพัฒนา	2
1.4.1 ข้อจำกัดของโปรแกรมเดิม	2
1.4.2 ขอบเขตของโปรแกรมที่ทำการพัฒนา	2
1.5 วิธีการดำเนินงาน	3
บทที่ 2 หลักการโน้ตดนตรีเบื้องต้น	4
2.1 บทนำ	4
2.2 ตัวโน้ต	4
2.3 ระดับเสียงและชื่อทางดนตรี	5
2.4 บรรทัด 5 เส้น	6
2.5 กุญแจ	6
2.6 เส้นกั้นห้องเพลง	7
2.7 เครื่องหมายกำหนดจังหวะ	8
2.8 ตัวหยุด	8
2.9 โน้ตประจุด	9
2.10 เครื่องหมายแปลงเสียง	9
2.10.1 การใช้เครื่องหมายชาร์ป	9
2.10.2 การใช้เครื่องหมายแฟลท	9
2.10.3 การใช้เครื่องหมายเนเจอรัล	10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

	หน้าที่
บทที่ 3 ไฟล์มิดี้	11
3.1 บทนำ	11
3.2 รูปแบบการส่งข้อมูลมิดี้	11
3.2.1 ข้อมูลแสดงแชลแนล	11
3.2.2 ข้อมูลระบบ	13
3.3 โครงสร้างไฟล์มิดี้	13
3.3.1 Header Chunk	14
3.3.2 Track Chunk	15
3.4 รูปแบบตัวอย่างไฟล์มิดี้	19
บทที่ 4 การทำพรีโพรเซสซิ่ง	21
4.1 บทนำ	21
4.2 การปรับภาพให้อยู่ในระนาบปกติ	21
4.2.1 Hough transform	21
4.2.2 ขั้นตอนการปรับระดับภาพ	23
4.3 การกำจัดสัญญาณรบกวน	25
4.3.1 Thresholding	25
4.3.2 Dilation	26
บทที่ 5 การเปรียบเทียบภาพ	29
5.1 บทนำ	29
5.2 การเปรียบเทียบค่าสหสัมพันธ์	29
5.3 ขั้นตอนการเปรียบเทียบภาพ	30
บทที่ 6 การออกแบบและการดำเนินงาน	32
6.1 บทนำ	32
6.2 ขั้นตอนการทำงานของโครงงาน	32
6.3 การแยกบรรทัด 5 เส้นออกจากภาพ	34
6.3.1 การหาตำแหน่งบรรทัด 5 เส้น	34
6.3.2 การลบบรรทัด 5 เส้น	35
6.3.3 การคืนภาพให้ตัวโน้ต	35

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

	หน้าที่
6.4 การหาขอบเขตตัวโน้ตและสัญลักษณ์	37
6.4.1 หลักการแยกองค์ประกอบ	37
6.4.2 Region Growing	38
6.4.3 ขั้นตอนการหาตำแหน่ง	39
6.5 การรู้จำรูปแบบโน้ตและสัญลักษณ์ทางดนตรี	41
6.5.1 กลุ่มที่มีความสูงต่ำกว่า 2 ช่องของบรรทัด 5 เส้น	42
6.5.1.1 กลุ่มที่มีความสูงภายใน 1 ช่อง	42
6.5.1.2 กลุ่มที่มีความสูงมากกว่า 1 ช่อง	45
6.5.2 กลุ่มที่มีความสูงเกินกว่า 2 ช่องแต่ไม่เกิน 3 ช่องของบรรทัด 5 เส้น	46
6.5.3 กลุ่มที่มีความสูงเกินกว่า 3 ช่องของบรรทัด 5 เส้น	46
6.5.3.1 กลุ่มที่มีเส้นตรงเพียงเส้นเดียว	46
6.5.3.1.1 กลุ่มที่มีเส้นตรงที่ยาวที่สุดอยู่ทางด้านซ้าย	47
6.5.3.1.2 กลุ่มทีมีเส้นตรงที่ยาวที่สุดอยู่ทางด้านขวา	49
6.5.3.1.3 กลุ่มที่มีเส้นตรงที่ยาวที่สุดอยู่ตรงกลาง	49
6.5.3.2 กลุ่มที่มีเส้นตรงมากกว่า 1 เส้น	49
6.5.4 การหาระดับเสียงสูงต่ำของตัวโน้ต	52
6.6 การแปลงข้อมูลให้อยู่ในรูปของไฟล์มิดี้	53
6.6.1 ข้อมูลที่ได้จากการรู้จำรูปแบบ	53
6.6.2 ขั้นตอนการสร้างไฟล์มิดี้	54
บทที่ 7 ผลการทดลอง	56
7.1 บทนำ	56
7.2 ประสิทธิภาพในการรู้จำ	56
บทที่ 8 บทสรุปและแนวทางในการพัฒนา	59
8.1 บทสรุป	59
8.2 แนวทางในการพัฒนาต่อ	60
ภาคผนวก ก	61
ภาคผนวก ข	62
บรรณานุกรม	67

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่		หน้าที่
2-1	แสดงประเภทของตัวโน้ต	4
2-2	แสดงประเภทของตัวหยุด	8
3-1	แสดงชุดข้อมูลของไฟล์มิดี้	11
3-2	แสดงชนิดและความหมายของข้อมูลเสียง	12
3-3	แสดงโครงสร้างของข้อมูลประเภท Chunk	13
3-4	แสดงโครงสร้างข้อมูลของไฟล์ MIDI	14
3-5	แสดงโครงสร้างของข้อมูลส่วน Header Chunk	14
3-6	แสดงโครงสร้างของข้อมูลส่วน Track Chunk	15
3-7	แสดงโครงสร้างของ MIDI Event	16
3-8	แสดงคำสั่งและความหมายต่างๆ ของ MIDI Event	17
3-9	แสดงคำสั่งและความหมายต่างๆของ Meta Event	18
4-1	แสดงรูปแบบตารางเก็บค่าผลลัพธ์ที่ได้จากการพิจารณา	23
5-1	แสดงการเปรียบเทียบภาพย่อยที่ต้องการหากับภาพหลักที่กำหนดไว้	31
6-1	แสดงโน้ตในกลุ่มที่มีความสูงภายใน 1 ช่องของบรรทัด 5 เส้น	43
6-2	แสดงสัญลักษณ์ในกลุ่มที่มีความสูงมากกว่า 1 ช่องแต่ไม่เกิน 2 ช่องของบรรทัด 5 เส้น	45
6-3	แสดงสัญลักษณ์ในกลุ่มที่มีความสูงเกินกว่า 2 ช่องแต่ไม่ถึง 3 ช่องของบรรทัด 5 เส้น	46
6-4	สัญลักษณ์ในกลุ่มที่มีความสูงเกินกว่า 3 ช่องของบรรทัด 5 เส้น	47
7-1	แสดงการทดสอบประสิทธิภาพการรู้จำของภาพในระนาบปกติ	57
7-2	แสดงการทดสอบประสิทธิภาพการรู้จำของภาพที่เอียงซ้าย	57
7-3	แสดงการทดสอบประสิทธิภาพการรู้จำของภาพที่เอียงขวา	58
ก-1	รหัสค่าเสียงตัวโน้ตเมื่อเทียบเป็นเลขฐานสิบที่ Octave ต่างๆ	61
ข-1	แสดงหมายเลขของเครื่องดนตรีแต่ละชนิด	62

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

หน้าที่

รูปที่

2-1 แสดงแผนภูมิกระจาย และการเทียบอัตราตัวโน้ต 5

2-2 แสดงระดับเสียง และชื่อทางดนตรี 5

2-3 แสดงระยะคู่แปด 5

2-4 แสดงบรรทัด 5 เส้น 6

2-5 แสดงการบันทึกตัวโน้ตลงบนบรรทัด 5 เส้น 6

2-6 แสดงลักษณะของกุญแจโน้ตดนตรี 6

2-7 แสดงโน้ตที่บันทึกลงบนกุญแจซอล และกุญแจฟา 7

2-8 แสดงเส้นกั้นห้องเพลง 7

2-9 แสดงสัญลักษณ์เครื่องหมายกำหนดจังหวะ 8

2-10 แสดงตัวอย่างการใช้ Dotted Note 9

2-11 แสดงเครื่องหมาย Sharp 9

2-12 แสดงเครื่องหมาย Flat 10

3-1 แสดงภาพโน้ตดนตรีที่จะทำการแปลงความให้อยู่ในรูปของไฟล์ MIDI 19

4-1 แสดงลักษณะของ Hough transform 22

4-2 แสดงรูปแบบของ Normal form ซึ่งสามารถแทนเส้นตรงได้ทุกๆมุม 23

4-3 แสดงการปรับภาพกรณีที่ภาพเอียงขวา 24

4-4 แสดงการปรับภาพกรณีที่ภาพเอียงซ้าย 24

4-5 แสดงลักษณะก่อนและหลังการปรับภาพเป็น 2 ระดับ 26

4-6 แสดงการพิจารณาพิกเซลที่อยู่รอบๆ จุด I(x , y) ซึ่งเป็นจุด ที่เรากำลังทำการพิจารณาอยู่ 26

4-7 แสดงเงื่อนไขทั้ง 4 ของ Binary Dilation 28

5-1 แสดงการเปรียบเทียบสหสัมพันธ์ระหว่างฟังก์ชั่น f(x , y) กับฟังก์ชั่น w(x , y) ณ ตำแหน่ง (x0 , y0) 30

5-2 แสดงภาพหลักของเครื่องหมาย ชาร์ป ที่สร้างไว้เป็นภาพหลัก 30

6-1 แสดงขั้นตอนการทำงานของโครงงาน 32

6-2 แสดงภาพโน้ตดนตรี และการสแกนภาพในแนวนอน 34

6-3 แสดงภาพตัวโน้ตที่ผ่านการลบบรรทัด 5 เส้น และเส้นน้อยออกไป 35

6-4 แสดงการลบบรรทัด 5 เส้น และการคืนส่วนของตัวโน้ตที่ถูกลบ 36

6-5 แสดงลักษณะการคืนภาพตัวโน้ต 36

6-6 แสดงภาพโน้ตที่ได้ผ่านการคืนค่าส่วนที่ขาดหายเสร็จเรียบร้อย 36

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ(ต่อ)

หน้าที่

รูปที่

6–7 แสดงการแยกองค์ประกอบ 37

6-8 แสดงการเชื่อมต่อกันของแต่ละพิกเซล 38

6-9 แสดงการหาตำแหน่งในขั้นตอนแรก 39

6-10 แสดงลำดับขั้นตอนการหาขอบเขต 40

6-11 แสดงตำแหน่งของตัวโน้ตที่ได้จากการหาขอบเขต 40

6-12 แสดงขอบเขตของโน้ตและเครื่องหมายทั้งหมด 41

6-13 แสดงตัวอย่างการจำแนกโน้ตและสัญลักษณ์ 42

6-14 แสดงการพิจารณาการเปลี่ยนแปลงค่าสีเพื่อระบุว่าเป็นโน้ตตัวกลม 43

6-15 แสดงสัญลักษณ์จุดซึ่งมีความกว้างและความสูงคล้ายสี่เหลี่ยมจัตุรัส 43

6-16 แสดงการพิจารณาตัวหยุดตัวกลมและตัวหยุดตัวขาวจากตำแหน่งที่ต่างกัน 44

6-17 สรุปวิธีการแยกตัวโน้ตและเครื่องหมาย ของกลุ่มที่มีความสูง ภายใน 1 ช่องของบรรทัด 5 เส้น 45

6-18 แสดงการแบ่งโน้ตออกเป็น 4 ส่วน 47

6-19 แสดงการวิเคราะห์เพื่อแยกแยะโน้ตตัวขาว และโน้ตตัวดำ 48

6-20 แสดงการวิเคราะห์เพื่อแยกแยะตัวเขบ็ด 1 ชั้น และตัวเขบ็ด 2 ชั้น 48

6-21 แสดงการเปลี่ยนแปลงค่าสีของโน้ตที่มีเส้นตรงอยู่ทางด้านขวา 49

6-22 แสดงการวิเคราะห์เพื่อแยกแยะตัวเขบ็ด 1 ชั้น และตัวเขบ็ด 2 ชั้น 49

6-23 แสดงความแตกต่างของการวาดโน้ตลงบนบรรทัด 5 เส้น 50

6-24 แสดงการแยกความแตกต่างระหว่างโน้ตที่มีหางชี้ขึ้นและชี้ลง 50

6-25 แสดงการพิจารณาตัวโน้ตที่มีหางชี้ขึ้น 50

6-26 แสดงการพิจารณาตัวโน้ตที่มีหางชี้ลง 51

6-27 แสดงการหาขอบเขตใหม่ของโน้ตกลุ่มที่มีเส้นตรงมากกว่า 1 เส้น 51

6-28 สรุปวิธีการแยกตัวโน้ตในกลุ่มที่มีความสูงเกินกว่า 3 ช่อง ของบรรทัด 5 เส้น 52

6-29 แสดงการหาระดับเสียงสูงต่ำของตัวโน้ต 52

6-30 แสดง ตัวอย่างโน้ตที่จะทำการสร้างไฟล์มิดี้ 53

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 หลักการและเหตุผล

เนื่องมาจากวิวัฒนาการทางเทคโนโลยีที่ก้าวหน้าอย่างต่อเนื่องจากอดีตจนถึงปัจจุบัน ทำให้เราได้เห็นถึงสิ่งต่างๆที่อยู่รอบตัวเราถูกนำมาพัฒนาปรับปรุงและเปลี่ยนแปลงให้ตอบสนองต่อความต้องการของมนุษย์และเพื่อให้สอดคล้องกับยุคสมัยไม่ว่าจะเป็น โทรศัพท์ ยานพาหนะ หรือที่เห็นได้เด่นชัดก็คือคอมพิวเตอร์ สิ่งต่างๆเหล่านี้ล้วนทำให้รูปแบบการดำเนินชีวิตของคนเราเปลี่ยนแปลงไป เช่นเดียวกับวิวัฒนาการของมนุษย์กับเสียงดนตรีที่อยู่ควบคู่กันจากอดีตจนถึงปัจจุบันและจะเป็นเช่นนี้เรื่อยไป เพราะว่ามนุษย์เราเป็นสิ่งมีชีวิต ที่มีความรู้สึกนึกคิด มีอารมณ์อ่อนไหว เช่นเดียวกับเสียงดนตรีที่มีความไพเราะและให้อารมณ์ในตัวของมันเอง วิวัฒนาการทางดนตรีนั้นได้รับการพัฒนามาเรื่อยๆอย่างต่อเนื่องเพื่อให้ตอบสนองต่อความต้องการมากที่สุด เมื่อพูดถึงดนตรีแล้วนั้น โน้ตดนตรีถือเป็นส่วนพื้นฐานของดนตรีเลยก็ว่าได้ อันเป็นส่วนสำคัญที่จะทำให้เราสามารถสร้างสรรค์เสียงเพลง หรือแทนเสียงเพลงเหล่านั้นออกมาให้อยู่ในรูปของสัญลักษณ์ ซึ่งไม่ว่าจะเป็นเครื่องดนตรีชนิดใดก็ล้วนแต่ใช้โน้ตดนตรีในการที่จะสร้างสรรค์เป็นบทเพลงขึ้นมา

ดังนั้นจากเหตุผลข้างต้น คณะผู้จัดทำโครงงานจึงมีแนวคิดที่จะทำการออกแบบและพัฒนาโปรแกรมที่สามารถแปลงภาพโน้ตดนตรีให้อยู่ในรูปของไฟล์มิดี้ เพื่อเป็นทางเลือกหนึ่งในการให้ความสะดวกสบายเกี่ยวกับเสียงดนตรี ซึ่งโครงงานนี้ได้มีผู้จัดทำก่อนหน้านี้แล้ว แต่ด้วยข้อจำกัดด้านการทำงานของโปรแกรมในหลายๆด้าน และประสิทธิภาพในการประมวลผลที่ไม่ดีเท่าที่ควร ทางผู้จัดทำจึงต้องการที่จะพัฒนาโปรแกรมขึ้นมาใหม่ โดยการนำเอาเทคนิคการประมวลผลต่างๆมาใช้เช่นการทำพรีโพรเซสซิ่งเป็นต้น เพื่อให้โปรแกรมสามารถตอบสนองต่อการใช้งานได้มากขึ้น และมีประสิทธิภาพในการประมวลผลที่ถูกต้อง

1.2 วัตถุประสงค์

1. เพื่อศึกษาทฤษฎีทางดนตรี ให้เข้าใจความหมายของโน้ตและสัญลักษณ์ต่างๆ อันจะทำให้โครงงานมีความถูกต้องและเป็นมาตรฐานสากล
2. เพื่อศึกษาทฤษฎีเกี่ยวกับโครงสร้างและการใช้งานของไฟล์มิดี้
3. เพื่อศึกษาหลักการและทฤษฎีการทำพรีโพรเซสซิ่ง
4. เพื่อศึกษาหลักการและทฤษฎีของการรู้จำรูปแบบ
5. เพื่อสร้างโปรแกรมแปลงจากภาพโน้ตดนตรีให้อยู่ในรูปของไฟล์มิดี้ ที่สามารถตอบสนองต่อการใช้งาน และมีศักยะภาพในการประมวลผลที่เพิ่มมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3 ผลที่คาดว่าจะได้รับ

1. ได้รับความรู้ ความเข้าใจเกี่ยวกับทฤษฎีของโน้ตดนตรี และไฟล์มิดี้
2. ได้รับความรู้ ความเข้าใจเกี่ยวกับการทำพรีโพรเซสซิ่ง
3. ได้รับความรู้ ความเข้าใจเกี่ยวกับเทคนิคการรู้จำรูปแบบ
4. โปรแกรมที่สามารถเปลี่ยนจากภาพโน้ตดนตรีให้อยู่ในรูปของไฟล์มิดี้ ซึ่งสามารถตอบสนองต่อการใช้งานได้ดี และมีประสิทธิภาพในการประมวลผล

1.4 ขอบเขตของการพัฒนา

โครงงานที่จัดทำขึ้นมานั้น จะมีคุณสมบัติของแต่ละส่วนคือ ส่วนแรกภาพที่รับเข้ามาจะทำการรับไฟล์ภาพมาจากการสแกนหรือไฟล์ภาพที่ได้จากโปรแกรมวาดโน้ต ส่วนที่สองการทำพรีโพรเซสซิ่งสามารถที่จะทำการลบ Noise ออกไปได้ และในกรณีที่ภาพเอียงจะสามารถปรับให้ภาพกลับอยู่ในระนาบปกติ ซึ่งภาพที่รับมานั้นสามารถเอียงได้ ± 45 องศา ส่วนที่สามการรู้จำรูปแบบสามารถที่จะตรวจสอบหาตำแหน่งบรรทัด 5 เส้น, วิเคราะห์ความหมายของตัวโน้ต และสัญลักษณ์ต่างๆจากภาพที่รับเข้ามา ส่วนสุดท้ายส่วนของไฟล์เอาท์พุตจะนำข้อมูลที่ได้จากการวิเคราะห์มาแปลงให้อยู่ในรูปของไฟล์มิดี้

สำหรับโครงงานที่จัดทำนั้นจะใช้ ภาษาจาวา และทำการพัฒนาบน Jbuilder Enterprise ซึ่งข้อจำกัดของโครงงานเดิม และขอบเขตของโปรแกรมที่ทำการพัฒนามีดังต่อไปนี้

1.4.1 ข้อจำกัดของโครงงานเดิม

- ภาพโน้ตที่ใช้ต้องอยู่ในรูปแบบ Bitmap (*.BMP)
- ภาพโน้ตที่ใช้ต้องอยู่ในรูปแบบ 8-bit Gray Scale
- ขนาดของภาพต้องไม่เกิน 1500*1500 พิกเซล
- หากภาพอินพุตที่รับเข้ามาไม่อยู่ในระนาบปกติ (ภาพเอียง) จะไม่สามารถทำการประมวลผลได้
- การกำจัดสัญญาณรบกวน (Noise) ขาดประสิทธิภาพ
- ไม่สามารถตรวจสอบชนิดของกุญแจได้
- ไม่สามารถตรวจสอบอัตราส่วนกำหนดจังหวะ
- ไม่สามารถตรวจสอบตัวเขบิ็ด 2 ชั้น, ตัวหยุดเขบิ็ด 2 ชั้นได้
- ไม่สามารถตรวจสอบตัวเขบิ็ด 1ชั้น, 2ชั้น และ 3 ชั้นที่ต่อกันหลายๆตัวได้
- ไฟล์มิดี้ที่ได้มีเสียงเครื่องดนตรีเพียงชิ้นเดียว

1.4.2 ขอบเขตของโปรแกรมที่ทำการพัฒนา

- เพิ่มความสามารถในการรับภาพ โดยภาพที่ใช้จะเป็นภาพพิมพ์ตัวโน้ตโดยรูปแบบของภาพที่สามารถนำมาใช้ได้แก่ BMP, JPG, GIF, PNG
- ภาพโน้ตที่ใช้สามารถอยู่ในรูปแบบ 8-bit Gray Scale, 256 Color(8-bit) หรือ Millions of Colors(24-bit)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ขนาดของภาพไม่เกิน 1500 * 1500 พิกเซล
- เพิ่มประสิทธิภาพในการปรับภาพที่เอียง ±45 องศา ให้กลับสู่ระนาบปกติได้
- เพิ่มประสิทธิภาพในการกำจัดสัญญาณรบกวน
- เพิ่มประสิทธิภาพในการตรวจสอบชนิดของกุญแจโดยสามารถตรวจสอบได้2 ชนิด คือ กุญแจซอล และกุญแจฟา
- เพิ่มประสิทธิภาพในการตรวจสอบอัตราส่วนกำหนดจังหวะ
- สามารถตรวจสอบโน้ตพื้นฐานต่างๆอันได้แก่ ตัวกลม, ตัวขาว, ตัวดำ, ตัวเขบ็ด1ชั้น
- สามารถทำการตรวจสอบเครื่องหมายต่างๆ อันได้แก่ ตัวหยุดตัวขาว, ตัวหยุดตัวดำ, ตัวหยุดเขบ็ด1ชั้น, โน้ตประจุด, เครื่องหมายแฟลท, เครื่องหมายชาร์ป, เครื่องหมายเนเจอรัล
- เพิ่มประสิทธิภาพในการตรวจสอบตัวเขบ็ด 2 ชั้น, 3 ชั้น และ ตัวหยุดเขบ็ด 2 ชั้น
- เพิ่มประสิทธิภาพในการตรวจสอบสอบตัวเขบ็ด 1ชั้น และ 2 ชั้นที่ต่อกันหลายๆตัว
- เพิ่มประสิทธิภาพในการกำหนดชนิดของเครื่องดนตรี โดยสามารถกำหนดได้ 3 ชนิด และไฟล์มิดี้สามารถเล่นประสานเสียงพร้อมกันได้

1.5 วิธีการดำเนินงาน

การดำเนินงานสำหรับโครงงานนี้นั้นจะแบ่งออกเป็น 3 ช่วงด้วยกันคือช่วงแรกจะเป็นช่วงของขั้นเตรียมการ ซึ่งในช่วงนี้จะทำการกำหนดเป้าหมายและขอบเขตของโครงงาน รวมทั้งทำการศึกษาเนื้อหาที่จำเป็นและเกี่ยวข้องทั้งหมดอันได้แก่ รายละเอียดเกี่ยวกับโน้ตดนตรี, ไฟล์มิดี้, การทำพรีโพรเซสซิ่งรูปแบบต่างๆ และเทคนิคการรู้จำรูปแบบ จากนั้นก็จะเข้าสู่ช่วงที่สองนั่นก็คือ ช่วงของขั้นดำเนินงาน ในช่วงนี้จะทำการออกแบบระบบ และพัฒนาโปรแกรมทั้งหมดของโครงงานโดยโปรแกรมได้ถูกแบ่งออกเป็น 2 ส่วนด้วยกันคือ ส่วนของการจัดการไฟล์ภาพที่รับเข้ามา และส่วนของการสร้างไฟล์มิดี้ เมื่อทำการพัฒนาทั้ง 2 ส่วนเสร็จสิ้นก็จะนำมารวมเข้าด้วยกันเพื่อให้สามารถทำงานร่วมกันได้ ซึ่งช่วงการดำเนินงานนี้ถือว่าเป็นช่วงที่มีความสำคัญเป็นอย่างมาก ดังนั้นจึงต้องใช้ความละเอียดและรอบคอบ อีกทั้งยังต้องหมั่นตรวจสอบอยู่เป็นระยะๆ ว่าการดำเนินงานที่ผ่านมานั้นตรงตามเป้าหมายและขอบเขตที่กำหนดหรือไม่ เพราะไม่เช่นนั้นอาจส่งผลให้ไม่เสร็จสิ้นทันเป้าหมายที่ได้วางไว้ ดังนั้นเมื่อทุกอย่างเป็นไปตามที่เราวางไว้ก็จะเข้าสู่ช่วงสุดท้ายนั้นก็คือ ช่วงของขั้นการประเมินผล ซึ่งแน่นอนว่าหากเรามีการตรวจสอบเป็นระยะๆในขั้นการดำเนินงานแล้ว โครงงานที่จัดทำย่อมเป็นไปตามสิ่งที่เราได้วางเป้าหมายไว้อย่างแน่นอน เราก็ตรวจสอบโปรแกรมขั้นสุดท้ายแล้วทำการสรุปผลการดำเนินงานทั้งหมดว่าที่ผ่านมานั่นเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
หลักการโน้ตดนตรีเบื้องต้น

2.1 บทนำ

โครงงานที่จัดทำในครั้งนี้จำเป็นอย่างยิ่งที่จะต้องทำการศึกษาเกี่ยวกับทฤษฎีทางดนตรี ให้เข้าใจความหมายของโน้ตและสัญลักษณ์ต่างๆ ทั้งนี้ก็เพื่อให้โครงงานที่จัดทำนั้น มีการประมวลผลและวิเคราะห์ภาพโน้ตดนตรีที่รับเข้ามาอย่างถูกต้องและเป็นมาตรฐานสากล ซึ่งการทำความเข้าใจนั้นอาจจำเป็นที่จะต้องใช้ระยะเวลาพอสมควร เนื่องจากว่าในส่วนของรายละเอียดค่อนข้างจะมีส่วนปลีกย่อยมากมาย ดังนั้นแล้วหากมีความตั้งใจจริงที่จะศึกษาโครงงาน ก็ควรจะมีความรู้ความเข้าใจในพื้นฐานทางดนตรีดังที่จะกล่าวต่อไปนี้

2.2 ตัวโน้ต

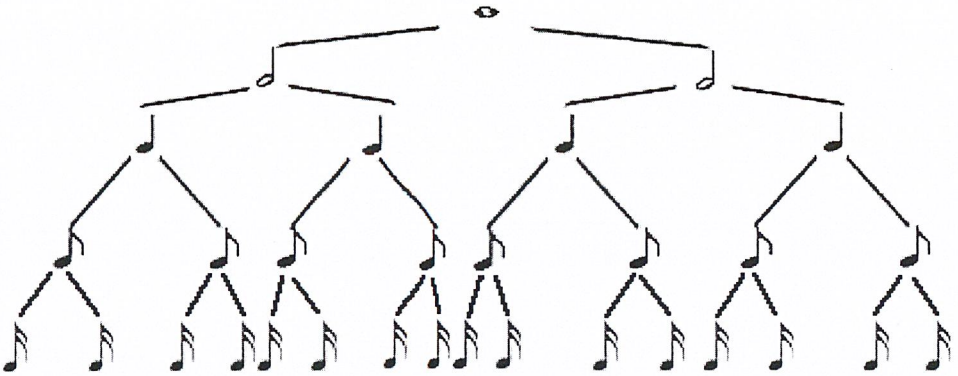
จะเป็นเครื่องหมาย หรือสัญลักษณ์ที่ใช้ในการบันทึกแทนเสียงดนตรี เพื่อให้ทราบว่าเสียงดนตรีนั้นจะความสั้นยาวของเสียงเป็นเช่นไร โดยสัญลักษณ์ของตัวโน้ตที่บอกความสั้นยาวของเสียงนั้นสามารถแสดงได้ตารางที่2-1

	โน้ตลักษณะนี้จะเรียกว่า “ตัวกลม” (Whole note) มี 4 จังหวะ
	โน้ตลักษณะนี้จะเรียกว่า “ตัวขาว” (Half note) มี 2 จังหวะ
	โน้ตลักษณะนี้จะเรียกว่า “ตัวดำ” (Quarter note) มี 1 จังหวะ
	โน้ตลักษณะนี้จะเรียกว่า “ตัวเขบ็ตหนึ่งชั้น” (Eighth note) มี ½ จังหวะ
	โน้ตลักษณะนี้จะเรียกว่า “ตัวเขบ็ตสองชั้น” (Sixteen note) มี¼ จังหวะ

ตารางที่ 2-1 แสดงประเภทของตัวโน้ต

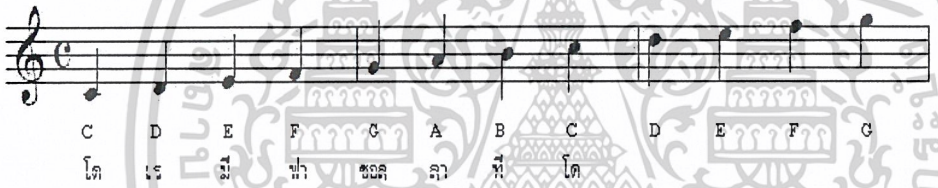
จากตารางจะเห็นได้ว่าค่าความยาวของเสียงตัวโน้ต โน้ตตัวกลมจะมีความยาวเสียงมากที่สุด และตัวอื่นๆคือ ตัวขาว ตัวดำ และตัวเขบ็ด1ชั้น จะสั้นลงตามลำดับ และเป็นสัดส่วนต่อกันดังแสดงในรูปที่ 2-1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2-1 แสดงแผนภูมิกระจาย และการเทียบอัตราตัวโน้ต

2.3 ระดับเสียง และชื่อทางดนตรี (Pitches and Musical names)

เสียงดนตรีนั้นนอกจากจะมีความยาวนานแล้ว จะต้องมีระดับเสียง สูง – ต่ำ ด้วยโดยจัดเรียงจากเสียงงต่ำไปหาสูง 7 เสียง ดังนี้ C D E F G A B โดยจัดเป็นชุด เรียงจากต่ำสุดไปหาสูงสุด

รูปที่ 2-2 แสดงระดับเสียง และชื่อทางดนตรี

เมื่อเวลาที่เราเรียกชื่อที่มีระดับเสียงต่ำหรือสูงกว่าที่กำหนดไว้ ก็ให้เรียกชื่อซ้ำอีกทบหนึ่งไปเรื่อยๆ จะเรียกเสียงที่มีชื่อเรียกเหมือนกัน แต่ระดับเสียงต่างกัน ว่ามีระยะคู่แปด หรือ Octave

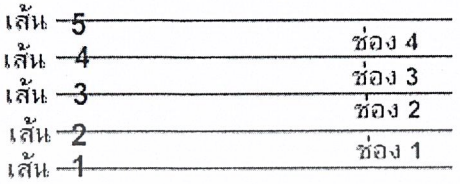
รูปที่ 2-3 แสดงระยะคู่แปด

จากรูปที่ 2-3 นั้นในการเทียบระดับเสียงสูงต่ำนั้น C D E F G A B ซึ่งมีจุดอยู่ด้านล่างจะมีเสียงที่ต่ำกว่า C D E F G A B และชุดสุดท้ายชุดที่มีจุดอยู่ด้านบนจะมีเสียงสูงกว่าในสองชุดแรกที่กล่าวมา C D E F G A B

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4 บรรทัด 5 เส้น (Staff)

บรรทัด 5 เส้นจะเป็นเส้นบรรทัดที่ใช้ในการบันทึกโน้ตดังรูปที่ 2-3(a) โดยเส้นด้านล่างใช้ในการบันทึกโน้ตเสียงต่ำ และเส้นที่สูงขึ้นมาใช้ในการบันทึกระดับเสียงโน้ตที่สูงขึ้นไปเป็นลำดับ การบันทึกระดับเสียงต่างๆในบรรทัด 5 เส้น เราสามารถทำได้ทั้งเขียนคาบเส้นดังรูปที่ 2-4(a) และเขียนในช่อง ดังรูปที่ 2-4(b)



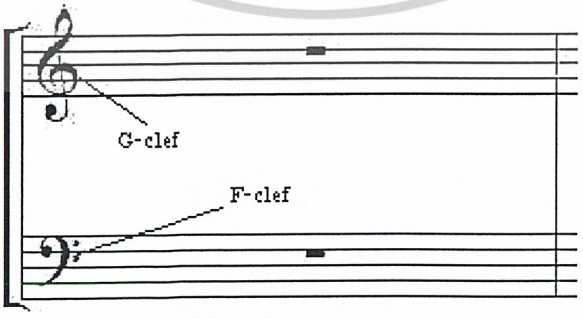
รูปที่ 2-4 แสดงบรรทัด 5 เส้น

(a) การบันทึกคาบเส้น (b) การบันในช่อง

รูปที่ 2-5 แสดงการบันทึกตัวโน้ตลงบนบรรทัด 5 เส้น

2.5 กุญแจ (Clef)

เป็นสัญลักษณ์ที่ใช้เป็นตัวกุญแจในการอ่านโน้ตต่างๆที่ถูกบันทึกลงบนบรรทัด 5 เส้น ถ้าไม่มีกุญแจบน บรรทัด 5 เส้น เราจะไม่สามารถอ่านระดับเสียงโน้ตออกมาได้ แต่จะทราบเพียงความแตกต่างของระดับเสียงของตัวโน้ตที่อยู่สูงกว่า หรือต่ำกว่ากุญแจที่เรานิยมใช้กันโดยทั่วไปมีอยู่ 2 ชนิดด้วยกันคือ กุญแจซอล (G Clef) และ กุญแจฟา (F clef) ดังแสดงในรูปที่ 2-6



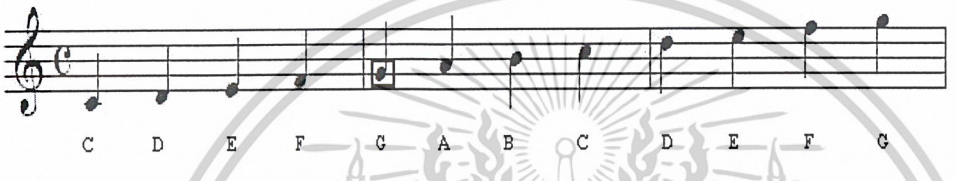
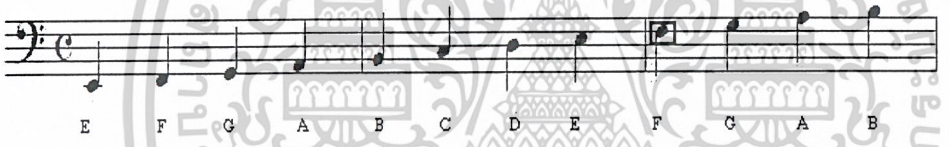
รูปที่ 2-6 แสดงลักษณะของกุญแจโน้ตดนตรี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.1 กุญแจซอล

เป็นกุญแจที่ใช้ในระดับเสียงปานกลาง ถึง ระดับเสียงสูงมากที่สุด เช่น Guitar เป็นต้น หัวของกุญแจจะคาบอยู่บนเส้นบรรทัดที่ 2 ของบรรทัด 5 เส้น ดังนั้นตัวโน้ตที่บรรทึกคาบเส้นเดียวกับหัวของกุญแจประจำหลักนี้ก็คือ “ซอล” (G) ตามชนิดและชื่อของกุญแจนั้น ดังแสดงในรูปที่ 2-7(a)

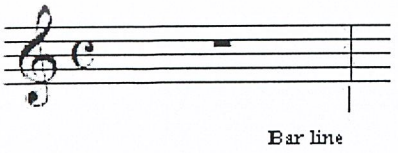
2.5.2 กุญแจฟา

เป็นกุญแจที่ใช้สำหรับเสียงต่ำ ถึง ระดับเสียงต่ำมาก เช่น Bass เป็นต้น หัวของกุญแจจะคาบอยู่บนเส้นบรรทัดที่ 4 ของบรรทัด 5 เส้น ดังนั่นตัวโน้ตที่บรรทึกคาบเส้นเดียวกับหัวของกุญแจประจำหลักนี้ก็คือ “ ฟา ” (F) ตามชนิดและชื่อของกุญแจนั้น ดังแสดงในรูปที่ 2-7(b)

(a) กุญแจซอล

(b) กุญแจฟา

รูปที่ 2-7 แสดงโน้ตที่บันทึกลงบนกุญแจซอล และกุญแจฟา

2.6 เส้นกั้นห้องเพลง (Bar Line)

เส้นกั้นห้องเพลง จะเป็นเส้นสำหรับแบ่งกั้นห้องเพลง ซึ่งเราจะทำการกั้นห้องเพลงเมื่อไหร่นั้นขึ้นอยู่กับTime signatureที่กำหนดอีกทีหนึ่ง



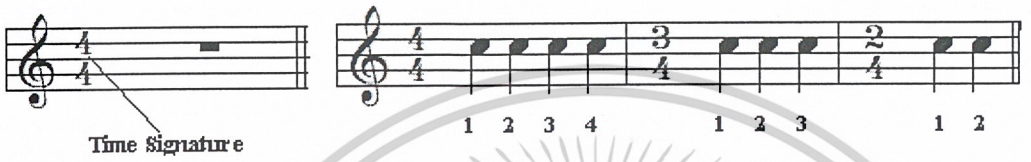
รูปที่ 2-8 แสดงเส้นกั้นห้องเพลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.7 เครื่องหมายกำหนดจังหวะ (Time Signature)

เครื่องหมายกำหนดจังหวะนั้นโดยปกติแล้ว เราจะพบเครื่องหมายนี้ได้หลังกุญแจ มีลักษณะการเขียนคล้ายตัวเลขในแบบเศษส่วน ดังแสดงในรูปที่ 2-9

ตอนเริ่มต้นของบทเพลงทุกเพลง จะมีเลข 2 ตัว ตัวหนึ่งอยู่ข้างบน อีกตัวหนึ่งอยู่ข้างล่าง ตัวเลขดังกล่าวคือ เครื่องหมายกำหนดจังหวะ ตัวเลขที่อยู่ด้านบนจะบอกจำนวนของจังหวะในแต่ละห้อง ส่วนตัวเลขที่อยู่ด้านล่าง จะบอกชนิดของตัวโน้ตที่มีค่าเท่ากับ 1 จังหวะ ยกตัวอย่างเช่น 4/4 หมายความว่า มี 4 จังหวะในแต่ละห้อง และเลข 4 ด้านล่างหมายความว่า โน้ตตัวดำ (QUARTER NOTE) มีค่าเท่ากับ 1 จังหวะ



รูปที่ 2-9 แสดงสัญลักษณ์เครื่องหมายกำหนดจังหวะ

2.8 ตัวหยุด (Rests)

ตัวหยุด เป็นเครื่องหมายที่ใช้เพื่อให้เงียบ หรือหยุดเสียงชั่วขณะ ซึ่งเสียงนั้นจะหยุดสั้นหรือนานเท่าใด ขึ้นอยู่กับ อัตราจังหวะและลักษณะของตัวหยุด ซึ่งลักษณะและอัตราตัวหยุดนั้น จะเทียบตรงกับตัวโน้ต แต่ตัวหยุดจะใช้ในทางตรงกันข้ามกับตัวโน้ตเท่านั้น ดังแสดงในตารางที่ 2-2

ตัวหยุด บันทึกไว้เพื่อให้เสียงเงียบเวลาบรรเลง

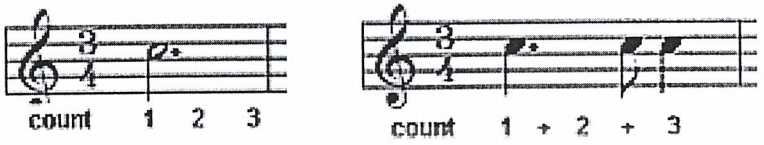
ตัวโน้ต บันทึกไว้เพื่อให้เกิดเสียงเวลาบรรเลง

	ตัวหยุดตัวกลม มีค่าหยุดเสียงเท่ากับโน้ตตัวกลม (หยุด 4 จังหวะ)
	ตัวหยุดตัวขาว มีค่าหยุดเสียงเท่ากับโน้ตตัวขาว (หยุด 2 จังหวะ)
	ตัวหยุดตัวดำ มีค่าหยุดเสียงเท่ากับโน้ตตัวดำ (หยุด 1 จังหวะ)
	ตัวหยุดตัวเขบ็ตหนึ่งชั้น มีค่าหยุดเสียงเท่ากับโน้ตตัวเขบ็ตหนึ่งชั้น (หยุด1/2 จังหวะ)

ตารางที่ 2-2 **แสดงประเภทของตัวหยุด**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.9 โน้ตประจุด (Dotted Note)

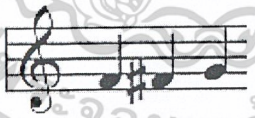
สัญลักษณ์โน้ตประจุดนี้ หากเราอ่านไปเจอในบทเพลงก็จะมีความหมายว่า ให้เราเพิ่มค่าโน้ตเข้าไปครึ่งหนึ่ง (1/2 จังหวะ) ของโน้ตที่มัน dotted อยู่ เช่น ถ้าเราเห็น half note dotted (ตัวขาวมี dotted) ก็มีความหมายว่า ให้เราทำการเพิ่มค่าโน้ตนี้ไป 1 จังหวะ เพราะว่าลำพังตัวขาวปกติจะมีค่า 2 จังหวะ ครึ่งหนึ่งของ 2 ก็คือ 1 ฉะนั้น เราจึงเพิ่มค่าโน้ตตัวขาวไปอีก 1 จังหวะ ก็จะมีค่าเป็น 3 จังหวะนั่นเอง

รูปที่ 2-10 แสดงตัวอย่างการใช้ Dotted Note

2.10 เครื่องหมายแปลงเสียง (Accidentals)

เครื่องหมายแปลงเสียง คือเครื่องหมายที่ใช้บันทึกขวางไว้หน้าตัวโน้ตที่ต้องการจะให้ระดับเสียงของตัวโน้ตนั้นๆ เปลี่ยนแปลงไปจากปกติ ไม่ว่าจะทำให้สูงขึ้นต่ำลง หรือกลับมาเท่าเดิมมีอยู่ 3 ชนิดด้วยกันคือ ช้าป, แฟล็ต และเนเจอรัล

2.10.1 การใช้เครื่องหมาย ชาร์ป (♯)

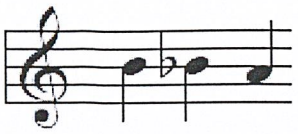
ถ้าต้องการจะให้เสียงเดิมเลื่อนสูงกว่าเดิมอีกครึ่งเสียง (1/2 จังหวะ) ให้ใช้เครื่องหมาย "#" บันทึกไว้หน้าตัวโน้ตที่ต้องการนั้น ในภาพ จะมีโน้ต F , F♯ ,G ซึ่ง F Sharp จะมีเสียงสูงกว่าF ครึ่งเสียง และมีเสียงต่ำกว่า G ครึ่งเสียง



รูปที่ 2-11 แสดงเครื่องหมาย Sharp

2.10.2 การใช้เครื่องหมาย แฟลท (♭)

ถ้าต้องการให้เสียงเดิมนั้นเลื่อนต่ำลงมาอีกครึ่งเสียง (1/2 จังหวะ) ให้ใช้เครื่องหมาย ♭ บันทึกไว้หน้าตัวโน้ตที่ต้องการนั้น ในภาพ จะมีโน้ต B, B♭, A ซึ่ง B Flat จะมีเสียงสูงกว่า B ครึ่งเสียง และมีเสียงต่ำกว่า A ครึ่งเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2-12 แสดงเครื่องหมาย Flat

2.10.3 การใช้เครื่องหมาย เนเจอรัล (♮)

ถ้าต้องการให้เสียงของตัวโน้ตที่ถูกเครื่องหมายแปลงเสียงบังคับไว้แต่แรกนั้นให้กลับสภาพตามเดิม ให้ใช้เครื่องหมาย ♮ บันทึกขวางหน้าตัวโน้ต จะมีผลทำให้เสียงของตัวโน้ตนั้นกลับสภาพปรกติ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
ไฟล์มิดี้

3.1 บทนำ

MIDI มาจากคำว่า Musical Instrument Digital Interface ซึ่งเป็นมาตรฐานในการสื่อสารข้อมูลแบบอนุกรม เพื่อใช้สำหรับสื่อสารระหว่างเครื่องดนตรีอิเล็กทรอนิกส์กัน หรือกับเครื่องคอมพิวเตอร์โดยข้อมูลที่ส่งไปนั้นเป็นข้อมูลแบบดิจิตอล ทำให้การแต่งเพลงหรือการบันทึกเสียงทำได้ง่ายขึ้นมาก สามารถนำเพลงมาแก้ไขได้สะดวกมากขึ้น

ไฟล์ MIDI ไม่ได้มีการเก็บเสียงดนตรีใดๆไว้เหมือนอย่างเทปเพลงหรือซีดีเพลง ข้อมูลทั้งหมดจะอยู่ในรูปของคำสั่งที่จะไปสั่งเครื่องดนตรีว่า ให้เปล่งเสียงโน้ตตัวใด(Note ON), ด้วยระดับความดังแค่ไหน (Velocity) และคำสั่งอื่นๆ ตามคุณสมบัติเฉพาะของเครื่องดนตรีแต่ละชนิด ด้วยเหตุที่เป็นไฟล์คำสั่งนี่เอง ทำให้มันมีขนาดที่เล็กมากๆ และจากความที่มันเป็นไฟล์คำสั่งแบบดิจิตอลนี่เอง นักคอมพิวเตอร์จึงสามารถนำข้อมูลดิจิตอลนี้มาพัฒนาด้วย จนในที่สุดทั้งคอมพิวเตอร์และเครื่องดนตรีก็สื่อสารกันได้อย่างสมบูรณ์โดยผ่านระบบ MIDIนี่เอง

3.2 รูปแบบของการส่งข้อมูลมิดี้

สำหรับการส่งข้อมูลในระบบ MIDI นั้นจะมีการส่งในลักษณะของ ชุดข้อมูล (Package) โดยที่ข้อมูล MIDI จะมีลักษณะดังนี้

MSB(Most Significant Bit) 1บิต	บิตของข้อมูล 7 บิต

ตารางที่ 3-1 แสดง ชุดข้อมูลของไฟล์ MIDI

- ถ้า MSB มีค่าเป็น 0 แสดงว่า ข้อมูลนี้เป็น ไบต์แสดงข้อมูล (Data Byte)
- ถ้า MSB มีค่าเป็น 1 แสดงว่า ข้อมูลนี้เป็น ไบต์แสดงสถานะ (Status Byte)

ทั้งนี้ในส่วนของไบต์แสดงสถานะนั้นจะสามารถแบ่งออกได้เป็น 2 ประเภทด้วยกันคือ ข้อมูลแสดงแชนแนล (Channel Message) และ ข้อมูลระบบ (System Message)

3.2.1 ข้อมูลแสดงแชนแนล

ไบต์แสดงสถานะประเภท ข้อมูลแสดงแชนแนลนั้นจะส่งข้อมูลตรงไปให้กับ แชนแนลใดแชนแนล หนึ่งในระบบเท่านั้น โดยที่ไบต์แสดงสถานะประเภทนี้ จะใช้บิตที่ 0 ถึงบิตที่ 3 เพื่อบอกถึง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แชนแนลที่ส่งไป ดังนั้นถ้าหากว่าอุปกรณ์ MIDI ที่เป็น Slave ตัวใดมีแชนแนลตรงกับ แชนแนลที่ส่งไป Slave ตัวนั้นก็จะตอบสนองต่อไบต์ข้อมูลที่ตามมา

ข้อมูลแสดงแชนแนล ยังสามารถแบ่งออกได้เป็น 2 ส่วนคือ ข้อมูลเสียง (Voice Message) และข้อมูลแสดงรูปแบบ (Mode Message)

3.2.1.1 ข้อมูลเสียง (Channel Voice Message)

โดยส่วนใหญ่แล้วข้อมูลแสดงแชนแนล จะเป็นประเภทข้อมูลเสียง ซึ่งหน้าที่ของข้อมูลเสียง คือ จะทำการสั่งให้ แชนแนลนั้นทำการเริ่มต้น หรือ หยุดเล่นโน้ตดนตรี เป็นต้น

ข้อมูลเสียงที่ใช้นั้นสามารถแสดงได้ดังตารางที่ 3-2

ข้อมูลเสียง	ความหมาย
Note on	ทำการเริ่มเล่นโน้ตดนตรี
Note off	ทำการหยุดเล่นโน้ตดนตรี
Control change	ควบคุมเกี่ยวกับตัวควบคุมเสียงบางประเภท เช่น ระดับเสียง (tone), ความดังของเสียง (volume)
Pitch Bend Change	เกี่ยวข้องกับการเอื้อนเสียงของแชนแนล
Program Change	ทำการเลือกเสียงของเครื่องดนตรีที่จะใช้ในแชนแนลนั้นๆ
Polyphonic Key	ข้อมูลชุดนี้จะถูกส่งหลังจากที่มีการเปลี่ยนแปลงความดัน เฉพาะ คีย์ ที่เราเล่นโน้ตดนตรีไปแล้ว
Channel Key Message	ข้อมูลชุดนี้จะเกี่ยวข้องกับการเปลี่ยนแปลง ความดัน ของ คีย์ เช่นเดียวกับ Polyphonic key แต่จะมีผลต่อเสียงทั้ง แชนแนล ไม่เฉพาะเจาะจง คีย์แบบ Polyphonic key

ตารางที่ 3-2 แสดงชนิดและความหมายของข้อมูลเสียง

3.2.1.2 ข้อมูลแสดงรูปแบบ (Channel Mode Message)

เป็นตัวควบคุมรูปแบบการทำงานของแต่ละ แชนแนล โดยจะทำการพิจารณาว่าจะให้แชนแนลนั้นจัดการกับข้อมูลเสียงที่ตามมาอย่างไร

ข้อมูลแสดงรูปแบบที่ใช้นั้น มีดังต่อไปนี้

1. Poly/Mono คือ การกำหนดให้อุปกรณ์ MIDI นั้นทำการเปล่งเสียงโน้ตดนตรีออกมาในแบบ Poly หรือ Mono โดยที่

- Mono การเปล่งเสียงนั้นจะสามารถ เล่นโน้ตดนตรีได้ทีละ 1 โน้ตเท่านั้น
- Poly การเปล่งเสียงนั้นจะสามารถเล่นโน้ตดนตรีได้ทีละหลายๆ โน้ตพร้อมกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. Omni On/Off เป็นการบังคับให้ แชนแนล นั้นตอบสนองเฉพาะข้อมูลที่ตรงกับ แชนแนล ของตน หรือตอบสนองในทุกๆ แชนแนล

- ถ้า Omni On จะเป็นการตอบสนองในทุกๆ แชนแนล
- ถ้า Omni Off จะเป็นการตอบสนองเฉพาะข้อมูลที่ตรงกับ แชนแนล ของตน

โดยปกติแล้ว ถ้านำอุปกรณ์ MIDI มาต่อกันให้เป็นระบบแล้ว รูปแบบที่จะนำมาใช้งานก็คือ Poly: Omni Off

3.2.2 ข้อมูลระบบ

ข้อมูลประเภทนี้จะทำการส่งออกไปโดยไม่เจาะจงเฉพาะแชนแนลใด แชนแนลหนึ่งในระบบ แต่ในทุกๆแชนแนลนั้นจะต้องตอบสนองต่อข้อมูลระบบนี้ ซึ่งข้อมูลระบบนั้นสามารถแบ่งได้เป็น

3.2.2.1 Common Message

ข้อมูลชุดนี้จะเกี่ยวข้องกับการเลือกเพลง การเลือกจังหวะ การเลือกตำแหน่งของเพลงที่จะเล่น และการบังคับให้อุปกรณ์ อะนาล็อก (Analog) ปรับแต่งระบบเสียงของตัวเอง (tune request)

3.2.2.2 System Real Time

ข้อมูลชุดนี้จะเกี่ยวข้องโดยตรงกับการสั่งเริ่มเล่นเพลง (Start), การหยุดเพลง (Stop) และข้อมูลที่สำคัญอื่นๆ เช่น ซิสเต็มคล็อก (System clock) จะใช้เกี่ยวกับระดับความเร็วของจังหวะเพลงว่าจะให้เร็วหรือช้า ยิ่งความถี่ในการส่งซิสเต็มคล็อกมาก เพลงก็จะยิ่งเล่นเร็วขึ้น

3.3 โครงสร้างไฟล์ MIDI

ไฟล์ MIDI จะถูกสร้างให้อยู่ในรูปแบบโครงสร้างที่เรียกว่า *"Chunk"* โดยในแต่ละ Chunk จะมีโครงสร้างภายในดังแสดงในตารางที่ 3-3

ชนิด	ความยาว	ข้อมูล
4 ไบต์	4 ไบต์	X ไบต์

ตารางที่ 3-3แสดงโครงสร้างของข้อมูลประเภท Chunk

- *ชนิดของ Chunk* เป็นส่วนที่ใช้บอกถึงประเภทของ Chunk นั้นๆ มีขนาดในการเก็บข้อมูลเท่ากับ 4 ไบต์ แบ่งได้เป็น 2 ประเภทคือ Header Chunk และ Track Chunk
- *ความยาวของ Chunk* เป็นส่วนที่ใช้บอกความยาวของข้อมูลที่ตามหลังมาว่า มีขนาดเท่าใด โดยมีขนาดในการเก็บข้อมูลเท่ากับ 4 ไบต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- *ข้อมูลของ Chunk* เป็นส่วนที่เก็บข้อมูลต่างๆ ของ Chunk โดยจะสามารถเก็บข้อมูลได้เท่าใดนั้น ก็ขึ้นอยู่กับค่าประกาศไว้ในส่วนของ ความยาวของChunk

รูปแบบของไฟล์ MIDI นั้นจะประกอบไปด้วยหนึ่ง Header Chunk ตามด้วย Track Chunk จำนวนตั้งแต่หนึ่ง Track Chunk ขึ้นไป

ไฟล์ MIDI	Chunk				
	ประเภท	ความยาว	ข้อมูล		
	MThd	6 ไบต์	*<format>*	*<tracks>*	*<division>*
	MTrk	*<length>*	*<Delta-Time><event>…*		
			.		
	MTrk	*<length>*	*<Delta-Time><event>…*		

ตารางที่ 3-4 แสดงโครงสร้างข้อมูลของไฟล์ MIDI

3.3.1 Header Chunk

Header Chunk จะเป็น ส่วนเริ่มต้นของไฟล์ MIDI เสมอและจะมีเพียง 1 Header Chunk เท่านั้น ซึ่งจะมีส่วนประกอบดังนี้

Header Chunk				
ประเภท	ความยาว	ข้อมูล		
4 ไบต์ แอสกี(ASCII)	4 ไบต์ (32 บิตไบนารี)	<length (6 ไบต์)>		
		16 บิต	16 บิต	16 บิต
MThd	<length>	<format>	<tracks>	<division>

ตารางที่ 3-5 แสดงโครงสร้างของข้อมูลส่วน Header Chunk

- *ส่วนบอกประเภทของ Chunk* เป็นตัวอักษร 4 ตัว ที่บอกว่า Chunk นี้เป็น Header Chunk โดยจะมีค่าเป็น “MThd” เสมอ ซึ่งจะมีขนาด 4 ไบต์ ซึ่งในรหัสแอสกี (ASCII) จะมีค่าเป็น ‘M = 4D’ ‘T = 54’ ‘h = 68’ ‘d = 64’

- *ความยาวของ Header Chunk* เป็นส่วนที่บอกความยาวของข้อมูลใน Header Chunk ว่ามีขนาดเท่าใด ซึ่งข้อมูลของ Header จะมีขนาดเป็น 6 ไบต์เสมอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- *ข้อมูลของ Header Chunk* ประกอบไปด้วย 3 ส่วนย่อยๆ คือ

  - ฟอร์แมต (Format) เป็นส่วนที่บอกชนิด แทร็ก (Track) ของ MIDI มีขนาดในการเก็บข้อมูล 2 ไบต์โดย ฟอร์แมตสามารถแบ่งได้เป็น 3 ชนิด คือ

    1. Single Track จะมีค่าในส่วนของ ฟอร์แมต เป็น 00 00
    2. Multiple Track แบบ Synchronous จะมีค่าในส่วนของ ฟอร์แมต เป็น 00 01
    3. Multiple Track แบบ Asynchronous จะมีค่าในส่วนของ ฟอร์แมต เป็น 00 02

  - แทร็ก (Track) เป็นส่วนที่ใช้บอกจำนวนของ Track Chunk ในไฟล์ MIDI มีขนาดในการเก็บข้อมูลเท่ากับ 2 ไบต์

  - Division เป็นส่วนที่บอกค่าเวลาของ Delta-Time ต่อโน้ตตัวดำ (Quarter note) มีขนาดในการเก็บข้อมูลเท่ากับ 2 ไบต์

3.3.2 Track Chunk

Track Chunk คือ ส่วนที่ใช้เก็บข้อมูลคำสั่งต่างๆ ของ ไฟล์ MIDI ในไฟล์ MIDI นั้น ประกอบไปด้วย Track Chunk จำนวนตั้งแต่ 1 Track Chunk ขึ้นไป ซึ่งรายละเอียดของส่วนประกอบของ Track Chunk นั้นมีดังนี้

Track Chunk		
ประเภท	**ความยาว**	**ข้อมูล**
4 ไบต์ แอสกี(ASCII)	4 ไบต์ (32 บิต ไบนารี)	<length ไบต์> (ข้อมูลชนิดไบนารี)
MTrk	<length>	<Delta-Time><event>…

ตารางที่ 3-6 แสดงโครงสร้างของข้อมูลส่วน Track Chunk

- *ส่วนบอกประเภทของ Chunk* เป็นตัวอักษร 4 ตัวที่บอกว่า Chunk นี้เป็น Track Chunk โดยจะมีค่าเป็น “MTrk” มีขนาด 4 ไบต์ ซึ่งในรหัสแอสกี จะมีค่าเป็น ‘M = 4d’ ‘T = 54’ ‘r =72’ ‘k=6B’

- *ความยาวของ Track Chunk* เป็นส่วนที่บอกความยาวของ ข้อมูลใน Track Chunk ว่ามีความยาวเท่าใด โดยมีขนาดในการเก็บข้อมูลเท่ากับ 4 ไบต์

- *ข้อมูลของ Track Chunk* จะประกอบขึ้นด้วย Track Event ตั้งแต่ 1 Track Event ขึ้นไป โดยที่แต่ละ Track Event จะประกอบไปด้วยส่วนย่อยๆ 2 ส่วน คือ

  1. Delta-Time เป็นส่วนที่ใช้สำหรับบอกค่าเวลาของเหตุการณ์นั้นๆ ซึ่งมีขนาดในการเก็บข้อมูลเท่ากับ 1-2 Bytes

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. Event เป็นส่วนของคำสั่งในแต่ละเหตุการณ์ที่สั่งให้ MIDI ทำงาน ขนาดของ Event นั้นจะขึ้นอยู่กับคำสั่งนั้นๆ Event สามารถแบ่งได้ดังนี้ คือ
   - MIDI Event คือ ส่วนที่บอกสถานะที่กำลังทำงาน (Running Status) ของเสียง และรูปแบบใน Channel ที่ต้องการ
   - Meta Event คือ ส่วนที่บอกข้อมูลระบบ จะไม่มีการกำหนด Channel และ ทุก Channel จะต้องตอบสนองต่อข้อมูลของระบบนี้

3.3.2.1 MIDI Event

MIDI Event นั้นสามารถแบ่งได้เป็น 2 ส่วนย่อย ๆ คือ

ส่วนของคำสั่ง	ส่วนของข้อมูล

ตารางที่ 3-7 แสดงโครงสร้างของ MIDI Event

- *ส่วนของคำสั่ง* จะมีขนาดเท่ากับ 1 ไบต์โดยมี 4 บิตแรก เป็นส่วนบอกคำสั่งที่จะใช้ในการทำงาน และ 4 บิตหลัง คือ ค่าของหมายเลขแชนแนล (Channel Number) ที่จะกำหนดให้ทำงาน
- *ส่วนของข้อมูล* จะเป็นส่วนบอกรายละเอียดของข้อมูลที่จะรับคำสั่งไปทำงาน เช่น หมายเลขโน้ต น้ำหนักในการเล่น (Velocity) เป็นต้น

คำสั่ง	ข้อมูล	รายละเอียด
8X	**NN VV** NN คือ หมายเลขโน้ต VV คือ น้ำหนักในการเล่น	ช่วงเลิกเล่นโน้ต (Note Off)
9X	**NN VV** NN คือ หมายเลขโน้ต VV คือ น้ำหนักในการเล่น	ช่วงเล่นโน้ต (Note On)
AX	**NN VV** NN คือ หมายเลขโน้ต VV คือ น้ำหนักในการเล่น	เพิ่มความดันของคีย์หลังจากที่กดคีย์นั้นไปแล้ว (Polyphonic Key Pressure)
BX	**CC VV** CC คือ หมายเลขตัวควบคุม VV คือ หมายเลขตัวควบคุมใหม่	เปลี่ยนตัวควบคุมที่ใช้ในการปรับแต่งเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

CX	**PP** PP คือ หมายเลขเครื่องดนตรี	ทำการกำหนดชนิดของเครื่องดนตรี
DX	**CC** CC คือ หมายเลขแชนแนล	กำหนดความดันเฉลี่ยของทุกๆ คีย์ของแชนแนล
EX	**BB TT** BB คือ Least Significant Bit TT คือ Most Significant Bit	การเอื้อนเสียง (Pitch Wheel Change)

ตารางที่ 3-8 แสดงคำสั่งและความหมายต่างๆ ของ MIDI Event

3.3.2.2 Meta Event

Meta Event จะเป็นข้อมูลจำพวก ชื่อแทร็ก, ข้อความเกี่ยวกับลิขสิทธิ์ ซึ่งโดยปกติแล้วนั้นจะไม่ถูกแสดงผลออกมา ในขณะที่คำสั่งต่างๆจะถูกประมวลผล แต่ก็ยังคงเป็นส่วนประกอบที่มีประโยชน์ต่อไฟล์ MIDIอยู่ รูปแบบทั่วๆไปของ Meta Events มีดังนี้

FF <Type> <Length> <Data>

<Type>

เป็น Single byte ใช้เพื่อบอกประเภทของ Meta event ซึ่งจะมีค่าตั้งแต่ 00 – 7F

<Length>

ใช้เพื่อบอกขนาดของข้อมูลที่ตามมา

<Data>

ข้อมูลต่างๆที่จะเก็บไว้ใน Meta Events ซึ่งจะมีขนาดตั้งแต่ 0 ไบต์ขึ้นไป

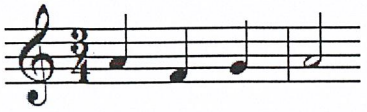
FF 00 02 SS SS	**Sequence Number**	
	ทำการตั้งลำดับหมายเลขของแทร็ก	
	SS SS	เป็นเลขฐานสองขนาด 16 บิต
FF 01 <length> <text>	**Text Event**	
	ข้อความเกี่ยวกับชื่อของแทร็ก	
	<length>	ค่าความยาว <text>
	<text>	ค่าแอสกีขนาดตามที่ประกาศไว้
FF 02 <length> <text>	**Copyright Notice**	
	ข้อความเกี่ยวกับลิขสิทธิ์	
FF 03 <length> <text>	**Track Name**	
	ข้อความบอกชนิดของ Track	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
55082

FF 04 <length> <text>	**Instrument Name**	
	ข้อความบอกชนิดของเครื่องดนตรีที่ใช้เล่น	
FF 05 <length> <text>	**Lyric**	
	ข้อความบอกทำนองของเพลง	
FF 06 <length> <text>	**Marker**	
	ข้อความระบุตำแหน่งที่จะเล่น	
FF 07 <length> <text>	**Cue Point**	
	ข้อความแนะนำเหตุการณ์ของเพลง	
FF 2F 00	**End of Track**	
	เป็นการบอกว่าจบ Track แล้ว	
FF 51 03 TT TT TT	**Set Tempo**	
	เป็นการกำหนดความเร็วของเพลง (Tempo) ถ้าไม่ได้ทำการตั้งไว้ค่ามาตรฐานจะเป็น 120 บีตต่อนาที ซึ่ง TT TT TT = 50 00 00	
	TT TT TT	ค่าจังหวะเป็นตัวเลขฐานสอง 24 บิต โดยมีหน่วยเป็น ไมโครวินาทีต่อโน้ตตัวดำ
FF 58 04 NN DD CC BB	**Time Signature**	
	ทำการกำหนดจังหวะของห้องเพลง	
	nn	ส่วนเศษ (Time signature numerator)
	dd	ตัวหาร (Time signature denominator)
	cc	MIDI Clocks ต่อ metronome tick
	bb	จำนวนของ 1/32 notes ต่อ 24 MIDI clocks (ค่ามาตรฐานคือ 8)
FF 59 02 SF MI	**Key Signature**	
	กำหนดคีย์ของเสียงเพลง 0 ใช้แทน C, จำนวนเต็มลบแทน เครื่องหมายแฟลตและ จำนวนเต็มบวกแทนเครื่องหมายชาร์ป	
	SF	จำนวนของ ชาร์ป หรือ แฟลต -7 = 7 แฟลต, 0 = C, +7 = 7 ชาร์ป
	MI	0 = Major key 1 = Minor key

ตารางที่ 3-9 แสดงคำสั่งและความหมายต่างๆของ Meta Event

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4 รูปแบบตัวอย่างไฟล์ MIDI

รูปที่ 3-1 แสดงภาพโน้ตดนตรีที่จะทำการแปลงความให้อยู่ในรูปของไฟล์ MIDI

เมื่อทำการแปลงให้อยู่ในรูปของไฟล์ MIDI จะได้ดังนี้ (แสดงเป็นเลขฐาน 16)

```
4D 54 68 64 00 00 00 06  00 01 00 02 01 80 4D
54 72 6B 00 00 00 10 00  FF 58 04 03 02 24 08
00 FF 51 03 07 A1 20 00  4D 54 72 6B 00 00 00
2B 00 C0 18 00 90 45 60  83 00 80 45 00 00 90
41 60 83 00 80 41 00 00  90 43 60 83 00 80 43
00 00 90 45 60 86 00 80  45 00 00 FF 2F 00
```

ข้อมูลของส่วน Header Chunk

4D 54 68 64	→	Chunk Type ในรหัส ASCII มีค่าเป็น MThd (Header Chunk)
00 00 00 06	→	ความยาวของ Header เท่ากับ 6 Bytes
00 01	→	ฟอร์แมตของ MIDI นี้คือ Format 1 (Multiple Track)
00 02	→	ไฟล์ MIDI นี้มีจำนวน 2 Track
01 80	→	กำหนดความยาวเสียงของโน้ตตัวดำเท่ากับ 384 (ฐาน 10)

ข้อมูลของส่วน Track Chunk แรก

4D 54 72 6B	→	Chunk Type ในรหัส ASCII มีค่าเป็น MTrk (Track Chunk)
00 00 00 10	→	ความยาวของ Track นี้เท่ากับ 16 Bytes
FF 58 04 03 02 24 08	→	ทำการกำหนดจังหวะของห้องเพลง
FF 51 03 07 A1 20	→	ทำการกำหนดความเร็วของการเล่นต่อโน้ตตัวดำ 1 ตัว

ข้อมูลของส่วน Track Chunk ที่สอง

4D 54 72 6B	→	Chunk Type ในรหัส ASCII มีค่าเป็น MTrk (Track Chunk)
00 00 00 2B	→	ความยาวข้อมูล Track นี้เท่ากับ 43 Bytes
C0 18	→	C0 คือ การตั้งเครื่องดนตรีที่ Channel 1
		18 คือ หมายเลขเครื่องดนตรี มีค่าเป็น 24 ในเลขฐาน 10
90 45 60	→	90 คือ ให้ Channel 1 ทำการเริ่มเล่นโน้ต
		45 หมายถึง โน้ตที่จะเล่นคือ A
		60 คือ ความหนักของเสียงที่เล่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

83 00	→	ค่าเวลาความยาวของเสียงของโน้ตตัวดำ
80 45	→	80 คือให้ Channel 1 ทำการหยุดเล่นโน้ต
		45 คือ โน้ตที่จะหยุดเล่นคือ A
90 41 60	→	90 คือ ให้ Channel 1 ทำการเริ่มเล่นโน้ต
		41 หมายถึง โน้ตที่จะเล่นคือ F
		60 คือ ความหนักของเสียงที่เล่น
83 00	→	ค่าเวลาความยาวของเสียงของโน้ตตัวดำ
80 41	→	80 คือให้ Channel 1 ทำการหยุดเล่นโน้ต
		41 คือ โน้ตที่จะหยุดเล่นคือ F
90 43 60	→	90 คือ ให้ Channel 1 ทำการเริ่มเล่นโน้ต
		43 หมายถึง โน้ตที่จะเล่นคือ G
		60 คือ ความหนักของเสียงที่เล่น
83 00	→	ค่าเวลาความยาวของเสียงของโน้ตตัวดำ
80 43	→	80 คือให้ Channel 1 ทำการหยุดเล่นโน้ต
		43 คือ โน้ตที่จะหยุดเล่นคือ G
90 45 60	→	90 คือ ให้ Channel 1 ทำการเริ่มเล่นโน้ต
		45 หมายถึง โน้ตที่จะเล่นคือ A
		60 คือ ความหนักของเสียงที่เล่น
86 00	→	ค่าเวลาความยาวของเสียงของโน้ตตัวขาว
80 45	→	80 คือให้ Channel 1 ทำการหยุดเล่นโน้ต
		45 คือ โน้ตที่จะหยุดเล่นคือ A
FF 2F 0	→	จบ Track

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
การทำพรีโพรเซสซิ่ง

4.1 บทนำ

ขั้นตอนการทำพรีโพรเซสซิ่ง (Pre-processing) นั้นถือว่ามีความสำคัญเป็นอย่างมาก ทั้งนี้เพราะว่าโปรแกรมที่จัดทำนั้นจะมีความถูกต้องในการรู้จำรูปแบบและสร้างไฟล์มิดี้ที่ได้มากเท่าใด ก็ขึ้นอยู่กับภาพที่จะส่งไปประมวลผล ซึ่งแน่นอนว่าหากภาพที่นำมาประมวลผลอยู่ในระนาบเอียง หรือมีสัญญาณรบกวน (noise) อยู่มาก การประมวลผลก็อาจเกิดความผิดพลาดขึ้นได้ ดังนั้นจึงต้องมีกระบวนการเพื่อปรับภาพให้กลับอยู่ในระนาบปกติ และกำจัดสัญญาณรบกวนออกไปก่อนที่จะนำไปใช้ในขั้นตอนอื่นๆ

ดังนั้นในบทนี้จะขอกล่าวถึงกระบวนการและขั้นตอนการทำพรีโพรเซสซิ่งที่ได้นำมาใช้ ว่ามีการทำงานและประสิทธิภาพเป็นเช่นใดบ้าง รวมทั้งเหตุใดจึงเลือกเอาเทคนิคนั้นๆมาใช้

4.2 การปรับภาพให้อยู่ในระนาบปกติ

ภาพที่จะนำมาประมวลผลนั้นจะต้องมีการตรวจก่อนว่าภาพนั้นอยู่ในระนาบที่ปกติหรือไม่ ทั้งนี้ก็เพราะหากภาพอยู่ในระนาบเอียง เมื่อส่งไปประมวลผลจะเกิดความผิดพลาดทั้งการตรวจสอบหาเส้นบรรทัดและส่วนของการรู้จำรูปแบบ ดังนั้นจึงต้องมีกระบวนการในการตรวจสอบภาพเพื่อปรับภาพที่เอียงให้อยู่ในระนาบปกติ ซึ่งเทคนิคที่ได้นำมาใช้นั้น จะใช้เทคนิคที่เรียกว่า Hough transform

4.2.1 Hough transform

Hough transform เป็นวิธีที่ได้รับการยอมรับเป็นมาตรฐานสำหรับการตรวจสอบภาพในทางวิชาการ ทางด้าน Computer Image Processing ซึ่งเทคนิคนี้มีความสามารถในการวิเคราะห์ภาพเพื่อตรวจสอบหาวัตถุ ไม่ว่าจะเป็นเส้นตรง, วงกลม, วงลี หรือรูปร่างที่ตรงกับแบบที่ต้องการ

สำหรับในโครงงานที่จัดทำจะนำเอา Hough transform มาตรวจสอบหาเส้นตรง (ในที่นี้ก็คือบรรทัด 5 เส้น) ซึ่งเราจะทำการตรวจสอบหาเส้นบรรทัดของภาพที่รับเข้ามาก็เพื่อจะหาว่าเส้นบรรทัดนั้นมีมุมเอียงไปจากระนาบปกติหรือไม่ ซึ่งถ้าหากพบว่ามีมุมเอียงเกิดขึ้นก็จะได้ทำการปรับให้อยู่ในระนาบปกตินั่นเอง

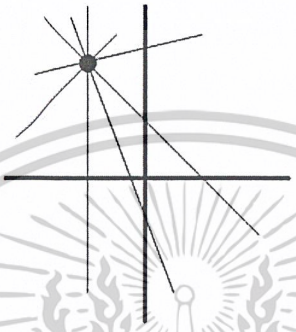
จุดภาพสีดำแต่ละจุดในภาพ สามารถที่จะมีเส้นตรงมากมายที่ลากผ่านจุดนั้นได้ โดยเส้นตรงแต่ละเส้นก็จะมีมุมที่แตกต่างกันไป ดังแสดงในรูปที่ 4-1(a) โดยลักษณะของแต่ละเส้นตรงนั้นสามารถแสดงออกมาได้ตามสมการความชันที่ 4.2.1-1

$$Y = mX + b \qquad \text{(สมการที่ 4.2.1-1)}$$

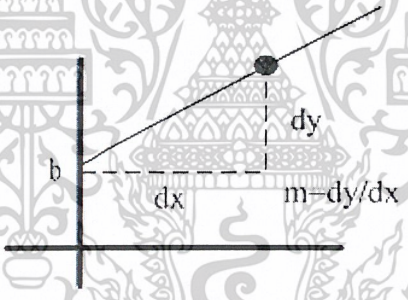
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ X, Y คือ ตำแหน่ง x และ y ของจุดภาพตามลำดับ ส่วน m แทนความชันของเส้นตรง และ b ก็คือจุดทีเส้นตรงลากผ่านแกน y แต่ในที่นี้หากสมการถูกแปลงในลักษณะที่เปลี่ยนไป คือให้ X และ Y เป็นค่าคงที่ ส่วน m และ b เป็นจุดตัดแกน x, y ก็จะสามารถแสดงออกมาอยู่ในรูปของอีกสมการหนึ่ง ดังแสดงในสมการที่ 4.2.1-2 ซึ่งจะเป็นสมการของเส้นตรงที่เป็นไปได้ทั้งหมดที่จะลากผ่านจุด x, y แสดงได้ดังรูปที่ 4-1(b)

$$b = -Xm + Y \qquad \text{(สมการที่ 4.2.1-2)}$$



(a) แสดงแต่ละจุดภาพสามารถมีเส้นตรงลากผ่านได้ไม่จำกัด



(b) แสดงจุดภาพที่อยู่บนเส้นตรงที่เราต้องการหา

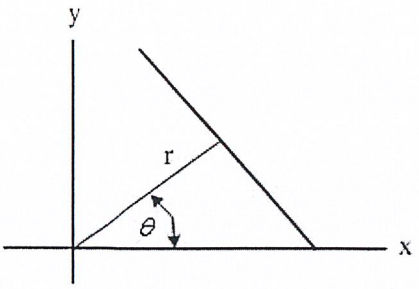
รูปที่ 4-1 แสดงลักษณะของ Hough transform

แต่ทั้งนี้ปัญหาหนึ่งที่เกิดจากการใช้สมการ $y = ax + b$ แทนแต่ละเส้นก็คือมุมที่เป็นไปได้จะมากมายจนนับไม่ถ้วน ซึ่งทางหนึ่งที่สามารถช่วยแก้ปัญหาได้ก็คือการเปลี่ยนมาใช้ในรูปของ normal form

$$r = x\cos\theta + y\sin\theta \qquad \text{(สมการที่ 4.2.1-3)}$$

โดยที่ r คือระยะทางจากจุด Origin ไปยังเส้นตรง, θ เป็นมุมเทียบกับแกน x ซึ่งเกิดจากการลากเส้นตรงตั้งฉากกับเส้นสมมุติมายังจุด Origin แสดงได้ดังรูปที่ 4-2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4-2 แสดงรูปแบบของ Normal form ซึ่งสามารถแทนเส้นตรงได้ทุกๆมุม

4.2.2 ขั้นตอนการปรับระดับภาพ

การปรับระดับภาพจะทำการพิจารณาจากจุดภาพที่เป็นสีดำทั้งหมด โดยจะทำการพิจารณาจากเส้นสมมุติซึ่ง ทำมุมตั้งฉากกับจุดกึ่งกลางของภาพโดยแต่ละมุมที่ทำการพิจารณาจะทำการเก็บค่าไว้ว่าเจอจุดสีดำทั้งหมดกี่จุด ดังนั้นมุม θ ที่เจอจุดสีดำมากสุดก็จะแสดงถึงมุมที่เอียงไปนั่นเองรูปแบบตารางสำหรับเก็บค่าที่ได้จากการพิจารณาแสดงได้ดังตารางที่ 4-1

θ / R	0°	1°	2°	3°	4°	----------	180°
1	x1	x2	x2	x3	x4		x5
⋮							
N							

ตารางที่ 4-1 แสดงรูปแบบตารางเก็บค่าผลลัพธ์ที่ได้จากการพิจารณา

จากตารางข้างต้นค่า x แต่ละค่าแทนค่าของจุดสีที่ได้จากการพิจารณาในแต่ละองศา ซึ่งหลังจากที่เราทราบมุมของภาพที่หักเหไปแล้ว เราก็จะทำการปรับภาพให้กลับคืนสู่ระนาบปกติ ซึ่งทั้งนี้ภาพที่เอียงสามารถเป็นไปได้ 2 กรณีคือ ภาพเอียงขวา และภาพเอียงซ้าย

จากรูปที่ 4-3 กรณีที่ภาพนั้นเอียงขวามุมที่เราจะต้องหมุนกลับนั้นสามารถพิจารณาได้ดังนี้

เมื่อ $\theta + X = 90^\circ$

$$X = 90^\circ - \theta$$

ตัวอย่าง สมมุติว่ามุม θ ที่หาได้นั้นเท่ากับ 45 องศา ก็จะได้ว่า

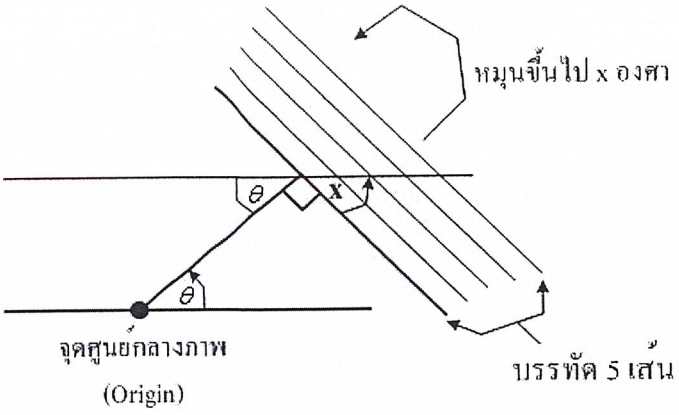
$$45 + X = 90^\circ$$

$$X = 90^\circ - 45^\circ$$

$$X = 45^\circ$$

ดังนั้นเราก็จะต้องหมุนขึ้นไปอีก 45 องศานั่นเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4-3 แสดงการปรับภาพกรณีที่ภาพเอียงขวา

ในกรณีที่ภาพเอียงซ้ายนั้นก็จะเป็นไปในทำนองเดียวกัน โดยจากรูปที่ 4-4 เป็นภาพที่เอียงซ้ายซึ่งมุมที่เราจะต้องหมุนกลับนั้นสามารถพิจารณาได้ดังนี้

เมื่อ $\theta + X = 90^\circ$

$$X = 90^\circ - \theta$$

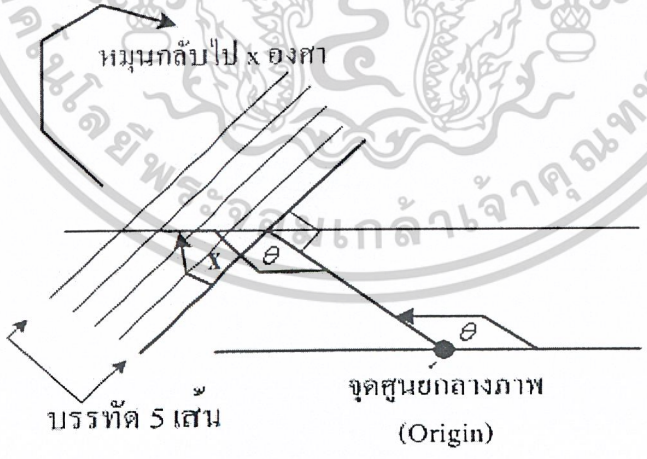
ตัวอย่าง สมมุติว่ามุม θ ที่หาได้นั้นเท่ากับ 120 องศา ก็จะได้ว่า

$$120 + X = 90^\circ$$

$$X = 90^\circ - 120^\circ$$

$$X = -30^\circ$$

ดังนั้นเราก็จะต้องหมุนย้อนกลับไปทั้งหมด -30 องศานั่นเอง



รูปที่ 4-4 แสดงการปรับภาพกรณีที่ภาพเอียงซ้าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 การกำจัดสัญญาณรบกวน

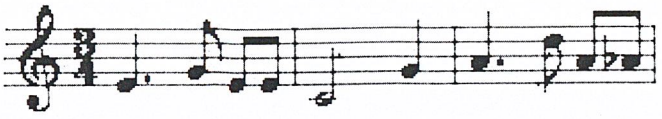
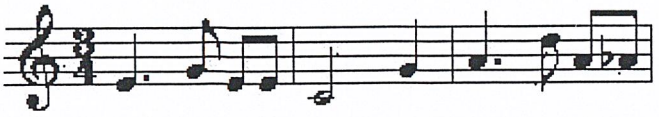
เทคนิคที่นำมาใช้ในการกำจัดสัญญาณรบกวนนั้น จะใช้เทคนิคที่เรียกว่าการทำ Thresholding ซึ่งเป็นการเปลี่ยนพิกเซลของภาพให้เหลือเพียงเพียง 2 ระดับ คือ 0 กับ 1 โดยจะทำการเปรียบเทียบค่าพิกเซลตำแหน่งนั้นๆ กับค่าคงที่ค่าหนึ่งที่เรียกว่า ค่าเทรชโฮล (Threshold Value) โดยถ้ามากกว่าจะทำการเปลี่ยนค่าพิกเซลที่ตำแหน่งนั้นให้เป็น 1 แต่หากว่าน้อยกว่าก็จะทำการเปลี่ยนค่าพิกเซลตำแหน่งนั้นให้เป็น 0 ประเด็นสำคัญของส่วนนี้จะอยู่ที่การเลือกใช้ค่าเทรชโฮลที่เหมาะสม โดยในโครงงานจะใช้ค่าเทรชโฮลเท่ากับ 238 ซึ่งได้มาจากการทดลองแล้วเปรียบเทียบผล เพื่อหาค่าเทรชโฮลที่เหมาะสมที่สุด

4.3.1 Thresholding

เมื่อทำการรับภาพมาจากการสแกน เพื่อความสะดวกและประสิทธิภาพในการวิเคราะห์ เราจึงต้องจัดการกับภาพที่ได้ให้มีความแตกต่างของค่าพิกเซลเพียงแค่ 2 ระดับ หรือ 2 สี คือ สีขาว ซึ่งเป็นสีของพื้นหลัง และ สีดำ ซึ่งเป็นสีของโน้ต/สัญลักษณ์ โดยจะทำการแปลงพิกเซลของภาพให้เป็นภาพไบนารี (Binary Image) ที่มีระดับความเข้มเพียง 2 ระดับคือ 0 กับ 1 โดยจุดภาพที่แทนด้วย 0 คือ จุดภาพที่มีสีดำ และจุดภาพที่แทนด้วย 1 คือจุดภาพที่มีสีขาว ดังนั้นภาพที่ได้จึงมีขนาดของข้อทูลที่เล็กทำให้การประมวลผลข้อมูลสามารถทำได้เร็วขึ้น

เทคนิคการทำเทรชโฮล (Thresholding Technique) เป็นกระบวนการที่ถูกนำใช้ในการสร้างภาพไบนารี เพื่อให้สามารถแยกภาพที่จะทำการพิจารณาออกจากพื้นหลังได้ง่าย โดยจะพิจารณาว่าพิกเซลใดควรจะเป็น สีขาว หรือ สีดำ หลักการนั้นจะทำการเปรียบเทียบค่าพิกเซลตำแหน่งนั้นๆ กับค่าคงที่ค่าหนึ่งที่เรียกว่า ค่าเทรชโฮล (Threshold Value) โดยถ้ามากกว่าจะทำการเปลี่ยนค่าพิกเซลที่ตำแหน่งนั้นให้เป็น 1 หรือเป็นสีขาว แต่หากว่าน้อยกว่าก็จะทำการเปลี่ยนค่าพิกเซลตำแหน่งนั้นให้เป็น 0 หรือเป็นสีดำ เทคนิคนี้นิยมใช้กันมากในกรณีที่ข้อมูลภาพมีความแตกต่างกันระหว่างวัตถุ (Object) และพื้นหลัง (Background)อย่างชัดเจน

ประเด็นสำคัญของส่วนทำ Thresholding นั้นจะอยู่ที่การเลือกค่า เทรชโฮล มาใช้เนื่องจากว่าค่าเทรชโฮลที่ไม่เหมาะสมย่อมส่งผลต่อความคมชัดของภาพและรายละเอียดของภาพบางส่วนที่อาจจะหายไป สำหรับวิธีคำนวณหาค่าเทรชโฮลนั้นมีอยู่ด้วยกันมีหลายวิธี แต่ละวิธีก็เหมาะสมกับลักษณะการทำงานที่แตกต่างกันไป ในโครงงานที่จัดทำนี้จะทำการหาค่าเทรชโฮลด้วยการกำหนดค่าล่วงหน้า (Pre-assigned Threshold Value) ซึ่งได้มาจากการทดลองแล้วเปรียบเทียบผลเพื่อหาค่าเทรชโฮลที่เหมาะสมที่สุด ภาพที่ 4-5(a) แสดงภาพโน้ตดนตรีก่อนทำการThresholding และหลังจากผ่านการทำ Thresholding จะได้ภาพที่มีความคมชัดขึ้นดังแสดงในรูปที่ 4-5(b)

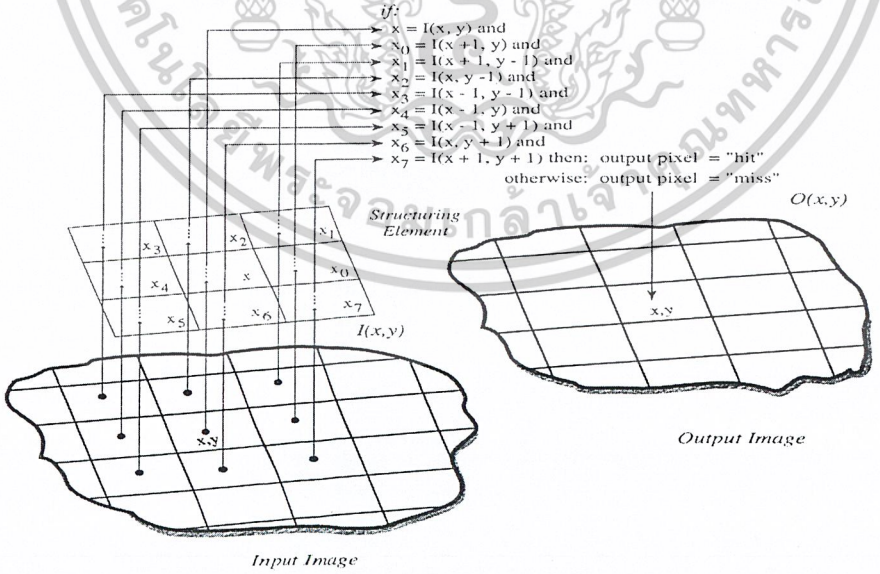
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(a) ก่อนการปรับภาพเป็น 2 ระดับ

(b) หลังการปรับภาพเป็น 2 ระดับ

รูปที่ 4-5 แสดงลักษณะก่อนและหลังการปรับภาพเป็น 2 ระดับ

4.3.2 Dilation

กระบวนการ Thresholding ในบางครั้งอาจไม่สามารถที่จะกำจัดสัญญาณรบกวนได้หมด ทั้งนี้เพราะว่าทุกพิกเซลไม่ได้เข้าอยู่ตามกลุ่มที่ต้องการ กล่าวคือ สัญญาณรบกวนในบางจุดอาจจะมีความเข้มที่น้อยกว่าค่าเทรชโฮล ทำให้เมื่อเข้าสู่ขั้นตอนการทำ Thresholding ก็จะนึกว่าสัญญาณรบกวนนั้นเป็นส่วนของเนื้อภาพก็จะทำการปรับให้เป็น ไบนารี 0 (สีดำ) ซึ่งแท้ที่จริงแล้วไม่ใช่ส่วนของภาพเลย ดังนั้นจึงต้องมีกระบวนการต่างๆเพิ่มเข้ามาเพื่อทำให้คุณภาพของภาพนั้นดียิ่งขึ้น

กระบวนการปรับภาพที่เพิ่มเข้ามาคือเทคนิคที่เรียกว่าการทำ Dilation ซึ่งมีหลักการคือจะทำการพิจารณาพิกเซลที่อยู่โดยรอบ ตำแหน่งพิกเซลที่เรากำลังทำการพิจารณา เพื่อที่จะกำหนดว่าตำแหน่งพิกเซลที่เรากำลังทำการพิจารณาอยู่นั้นควรจะถูกกำหนดให้เป็นค่า 0 หรือควรจะกำหนดให้เป็นค่า 1 วิธีการพิจารณาพิกเซลต่างๆที่อยู่โดยรอบตำแหน่งพิกเซลที่เรากำลังทำการพิจารณา แสดงได้ดังรูปที่ 4-6



รูปที่ 4-6 แสดงการพิจารณาพิกเซลที่อยู่รอบๆ จุด I(x , y) ซึ่งเป็นจุดที่เรากำลังทำการพิจารณาอยู่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากภาพข้างต้นนั้นจะเห็นว่าเอาท์พุตที่ได้จะมีอยู่ 2 กรณีคือ hit และ miss ซึ่งจะ hit ก็ต่อเมื่อค่า ณ พิกเซลที่เราสนใจและกำลังพิจารณาอยู่นั้น เหมือนกับพิกเซลรอบข้าง ส่วน miss นั้นก็เป็นในทางกลับกัน คือ พิกเซลที่เราสนใจและกำลังพิจารณาอยู่นั้น ไม่เหมือนกับพิกเซลรอบข้าง

การทำ Dilation นั้นจะให้พิกเซลที่เราสนใจนั้นแทนด้วยไบนารี 0 (เนื้อภาพซึ่งเป็นสีดำ)

0 0 0

0 0 0

0 0 0

โดยที่ O(x , y) = 0 (ดำ) สำหรับกรณี hit

= 1 (ขาว) สำหรับกรณี miss

หลักที่ใช้ในการตัดสินใจว่า ณ ตำแหน่งที่เรากำลังพิจารณาอยู่นั้นควรจะถูกกำหนดให้เป็นค่า 0 หรือควรจะถูกกำหนดให้เป็นค่า 1 นั้น จะมีอยู่ 4 เงื่อนไขด้วยกันคือ

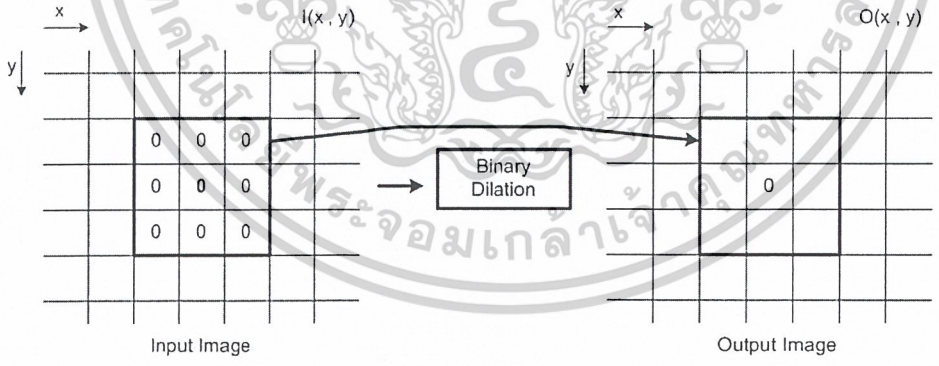
เมื่อพิกเซล อินพุต = 0, พิกเซลรอบข้างทั้งหมดเป็น 0 จะเป็น hit; O(x , y) = 0

เมื่อพิกเซล อินพุต = 1, พิกเซลรอบข้างทั้งหมด1 จะเป็น miss; O(x , y) = 1

เมื่อพิกเซล อินพุต = 1, พิกเซลรอบข้างมีทั้ง1และ 0 จะเป็น miss; O(x , y) = 1

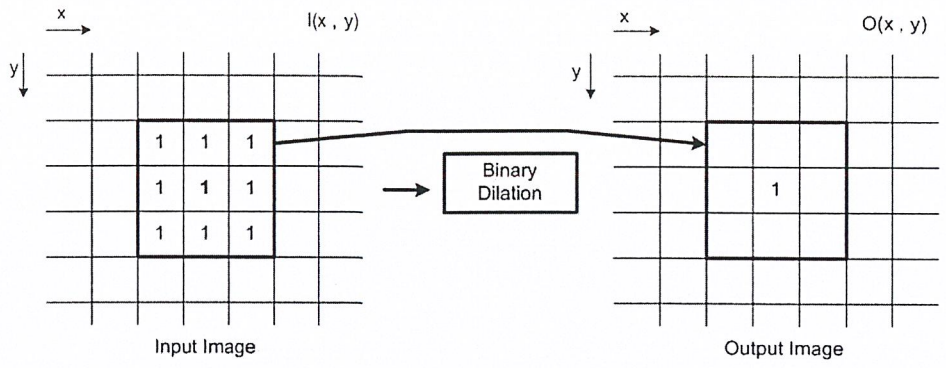
เมื่อพิกเซล อินพุต = 0, พิกเซลรอบข้างมีทั้ง1และ 0 จะเป็น miss; O(x , y) = 1

ทั้งนี้เพื่อให้สามารถทำความเข้าใจได้ง่ายขึ้น วิธีการทำ Dilation ทั้ง 4 เงื่อนไขนั้นสามารถแสดงได้ดังรูปที่ 4-7

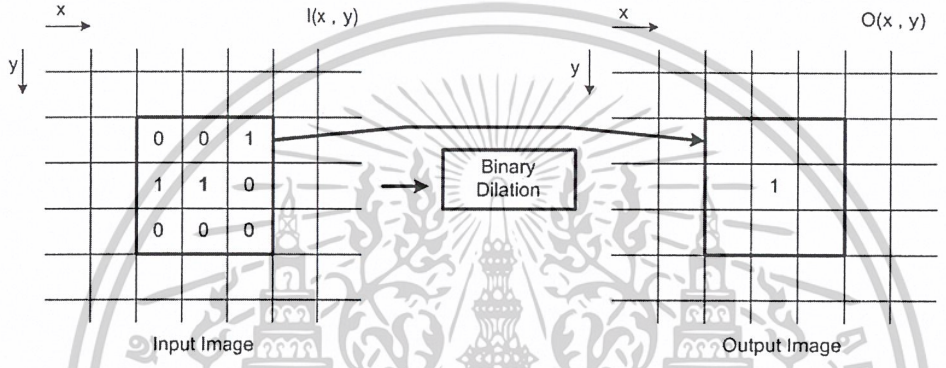


(a) แสดงเงื่อนไขที่ 1 พิกเซลอินพุต = 0, พิกเซลรอบข้างทั้งหมดเป็น 0

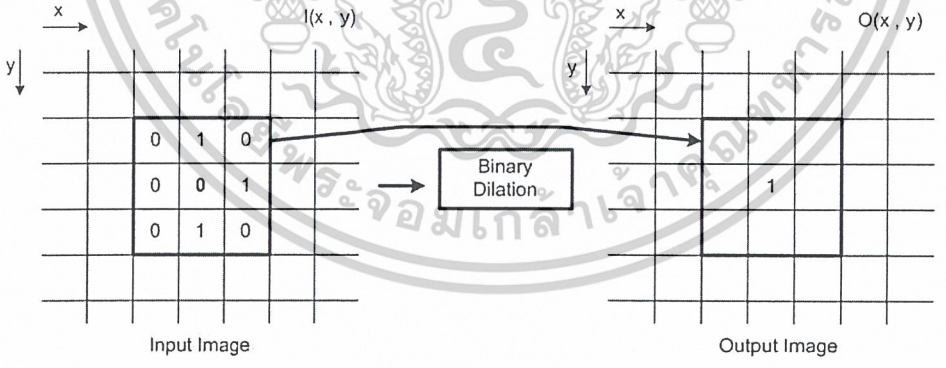
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(b) แสดงเงื่อนไขที่ 2 พิกเซลอินพุต = 1, พิกเซลรอบข้างทั้งหมดเป็น 1



(c) แสดงเงื่อนไขที่ 3 พิกเซลอินพุต = 1, พิกเซลรอบข้างมีทั้ง 1 และ 0



(d) แสดงเงื่อนไขที่ 4 พิกเซลอินพุต = 0, พิกเซลรอบข้างมีทั้ง 1 และ 0

รูปที่ 4-7 แสดงเงื่อนไขทั้ง 4 ของ Binary Dilation

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
การเปรียบเทียบภาพ

5.1 บทนำ

การที่จะทำให้เราสามารถแยกแยะได้ว่าภาพนั้นๆเป็นตัวโน้ตหรือเครื่องหมายใด เราอาจจะใช้วิธีการธรรมดาทั่วๆไปคือพิจารณาจากลักษณะเฉพาะของโน้ตหรือสัญลักษณ์นั้นๆเช่นการแบ่งเป็นสี่ส่วนแล้วดูว่ามีจุดสีดำอยู่ในส่วนใดบ้าง หรืออาจจะทำการพิจารณาจากการเปลี่ยนแปลงค่าสีบริเวณส่วนหัวตัวโน้ตก็ได้ แต่ทั้งนี้ในบางกรณีตัวโน้ตหรือเครื่องหมายบางตัวอาจไม่สามารถที่จะนำเอาลักษณะเฉพาะตัวมาใช้แยกแยะในการแยกแยะความแตกต่างได้ ดังนั้นจึงต้องหาวิธีการใหม่เพื่อให้สามารถแยกแยะและบอกได้ว่าเป็นโน้ตหรือสัญลักษณ์ใด

5.2 การเปรียบเทียบค่าสหสัมพันธ์ (Matching by Correlation)

การพิจารณาเพื่อเปรียบเทียบภาพนั้น ในที่นี้จะเราจะให้ภาพย่อยหรือภาพที่จะนำมาเปรียบเทียบแทนด้วยฟังก์ชั่น w(x , y) ที่มีขนาด J x K ส่วนภาพหลักหรือภาพต้นแบบที่เรากำหนดไว้จะถูกแทนด้วยฟังก์ชั่น f(x , y) ซึ่งมีขนาด M x N ซึ่งความสัมพันธ์ระหว่างทั้งสองฟังก์ชั่นนั้นค่า $J \leq M$ และ $K \leq N$ แต่ทั้งนี้ถึงแม้ว่าวิธีการหาค่าสหสัมพันธ์นั้นสามารถที่จะแสดงออกมาในรูปของเวคเตอร์ได้ แต่การนำเอาภาพย่อยไปเปรียบเทียบกับภาพหลักนั้นเป็นวิธีการที่ค่อนข้างได้รับความนิยมมากกว่า

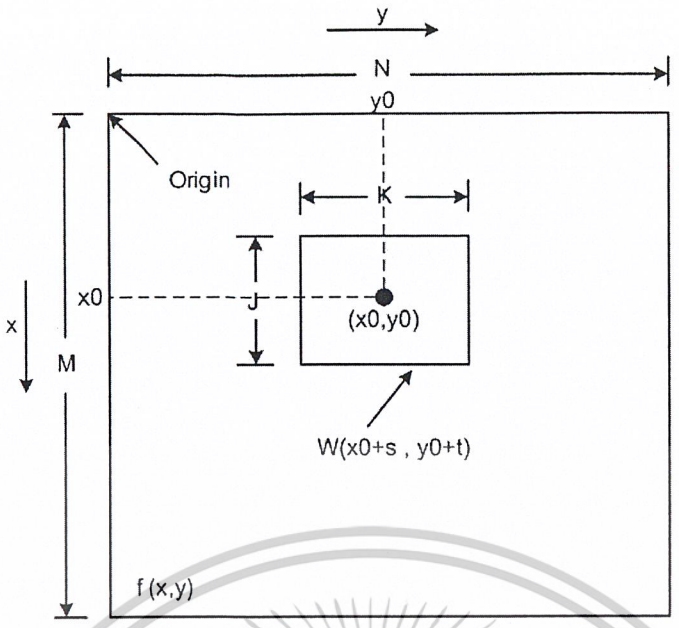
รูปแบบของวิธีการที่ง่ายที่สุด ในการที่จะหาค่าสหสัมพันธ์ระหว่างฟังก์ชั่น f(x , y) กับฟังก์ชั่น w(x , y) นั้นแสดงได้ดังสมการที่ 5.2-1

$$c(x,y) = \sum_{s} \sum_{t} f(s,t)w(x+s, y+t) \qquad (5.2\text{-}1)$$

จากสมการค่าของ x นั้นจะมีค่าตั้งแต่ 0, 1, 2,…., M-1 และในส่วนของค่า y นั้นจะมีค่าตั้งแต่ 0, 1, 2, …., N-1 ซึ่งผลรวมของสมการ จะเป็นค่าผลรวมของภาพทั้งหมดที่เกิดจากการซ้อนทับกัน ระหว่างฟังก์ชั่น w(x , y) กับฟังก์ชั่น f(x , y) ทั้งนี้การพิจารณาเพื่อตัดสินใจว่าภาพย่อยที่นำมาเปรียบเทียบนั้นตรงกับภาพหลักใดเราจะดูจากผลลัพธ์ที่ได้จากฟังก์ชั่น c(x , y) ซึ่งผลลัพธ์ที่มีค่ามากที่สุดจะทำให้เราทราบได้ว่าภาพย่อยนั้นตรงกับภาพหลักใดนั่นเอง

เพื่อให้เข้าใจได้ง่ายขึ้นเกี่ยวกับหลักในการเปรียบเทียบสหสัมพันธ์ระหว่างภาพหลักกับภาพย่อยนั้น สามารถที่จะแสดงออกมาได้ดังรูปที่ 5-1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5-1 แสดงการเปรียบเทียบสหสัมพันธ์ระหว่างฟังก์ชั่น f(x , y) กับ ฟังก์ชั่น w(x , y) ณ ตำแหน่ง (x0 , y0)

5.3 ขั้นตอนการเปรียบเทียบภาพ

ในการเปรียบเทียบภาพนั้น ในส่วนของภาพหลักที่เราจะใช้เป็นต้นแบบในการเปรียบเทียบ เราจะทำการสร้างไว้เป็นอาร์เรย์ขนาด 20 x 26 ตัวอย่างของภาพหลักที่สร้างไว้แสดงได้ดังรูปที่ 5-2

```
{ 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0 },
{ 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0 },
{ 0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,1,0,0,0 },
{ 0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,1,0,0,0 },
{ 0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,1,0,0,0 },
{ 0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,1,0,0,0 },
{ 0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,1,1,1,1 },
{ 0,0,0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1 },
{ 0,0,0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1 },
{ 0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1 },
{ 0,0,0,1,1,1,1,1,1,1,0,0,0,0,1,1,0,0,0 },
{ 0,0,0,1,1,1,1,1,1,1,0,0,0,0,1,1,0,0,0 },
{ 0,0,0,1,1,1,1,1,1,0,0,0,0,0,1,1,0,0,0 },
{ 0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,1,0,0,0 },
{ 0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,1,1,1,1 },
{ 0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,1,1,1,1 },
{ 0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1 },
{ 0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1 },
{ 0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1 },
{ 0,0,0,1,1,1,1,1,1,1,0,0,0,0,0,1,1,0,0,0 },
{ 0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,1,1,0,0,0 },
{ 0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,1,1,0,0,0 },
{ 0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,1,1,0,0,0 },
{ 0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,1,1,0,0,0 }
```

รูปที่ 5-2 แสดงภาพหลักของเครื่องหมาย ชาร์ป ที่สร้างไว้เป็นภาพหลัก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต่อมาในส่วนของภาพย่อยที่จะนำมาเปรียบเทียบนั้นก็จะได้มาจากการหาขอบเขตด้วยวิธีเซกเมนต์เทชั่น ซึ่งโน้ตหรือเครื่องหมายแต่ละตัว ก่อนที่จะนำมาเปรียบเทียบจะต้องทำการปรับขนาดให้เท่ากับภาพหลักก่อนทุกครั้ง ซึ่งหลังจากที่ทำการปรับขนาดแล้วก็จะนำเอาภาพย่อยนั้นไปเปรียบเทียบกับภาพหลักทั้งหมดที่เรามี โดยผลลัพธ์ที่ได้จากการเปรียบเทียบคู่ใดมีค่ามากที่สุดก็แสดงว่าภาพที่ต้องการหาก็คือภาพหลักของคู่นั้นนั่นเอง

ตัวอย่างผลลัพธ์ที่ได้จากฟังก์ชั่น c(x , y) ของการเปรียบเทียบภาพย่อยในที่นี้คือเครื่องหมายชาร์ปกับภาพหลักทั้งหมดที่เรามีแสดงได้ดังตารางที่ 5-1

ภาพย่อยที่ต้องการหา	ภาพหลักที่กำหนดไว้	ผลลัพธ์ที่ได้
เครื่องหมายชาร์ป (♯)	เครื่องหมายชาร์ป (♯)	0.37115384615348617
	เครื่องหมายเนเจอรัล (♮)	0.18461538461538463
	เครื่องหมายแฟลท (♭)	0.11346153846153846
	ตัวหยุดเขบ็ด 1 ชั้น (𝅘𝅥𝅯)	0.16153846153846155
	ตัวหยุดเขบ็ด 2 ชั้น (𝅘𝅥𝅰)	0.1076923076923077

ตารางที่ 5-1 แสดงการเปรียบเทียบภาพย่อยที่ต้องการหากับภาพหลักที่กำหนดไว้

ดังนั้นจากตารางข้างต้นจะเห็นได้ว่าภาพย่อยที่ต้องการหานั้นเมื่อนำไปเปรียบเทียบกับเครื่องหมายช้าปแล้วผลลัพธ์ที่ได้จะมีค่ามากที่สุดเมื่อเทียบกับคู่อื่นๆ จึงทำให้เราทราบได้ว่าเป็นเครื่องหมายชาร์ปนั่นเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

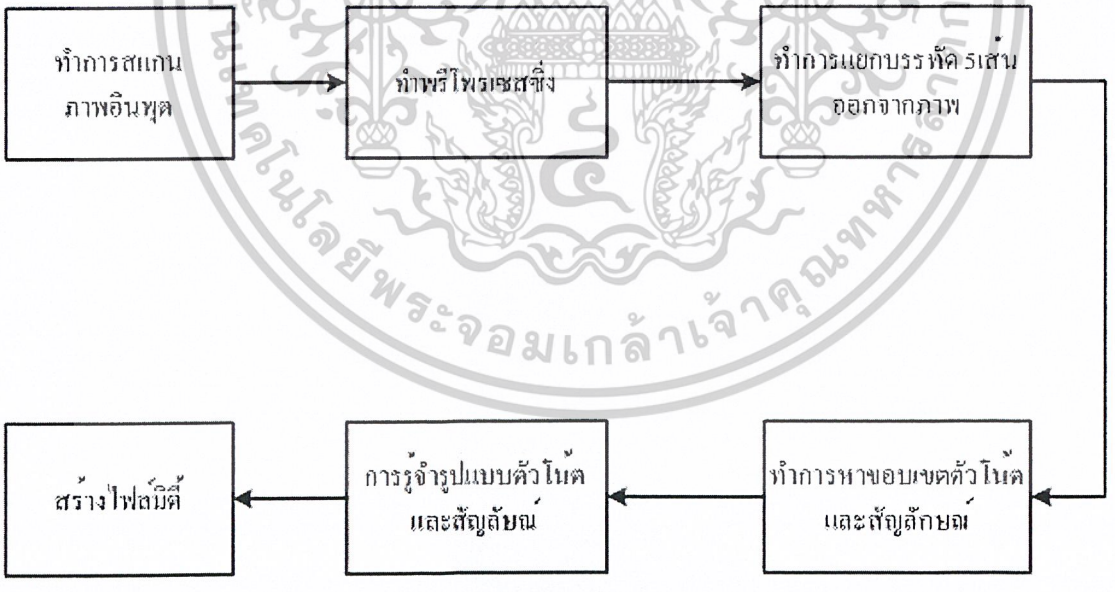
บทที่ 6
การออกแบบและการดำเนินงาน

6.1 บทนำ

เนื้อหาในบทนี้จะได้กล่าวถึงขั้นตอนการออกแบบโครงงานและกระบวนการในการรู้จำโน้ตดนตรีทั้งหมด ซึ่งภายหลังจากการออกแบบนั้นได้ทำการแบ่งการทำงานของโครงงานออกเป็นส่วนๆ โดยโครงงานจะเริ่มจากทำการรับภาพอินพุตเข้ามา แล้วนำภาพที่ได้มาผ่านกระบวนการพรีโพรเซสซิ่งเพื่อกำจัดสัญญาณรบกวนและปรับภาพที่เอียงให้อยู่ในระนาบปกติ จากนั้นจะทำการปรับภาพให้เป็น 2 ระดับเสร็จแล้วทำการตรวจสอบหาตำแหน่งของบรรทัด 5 เส้นเพื่อลบบรรทัด 5 เส้นออก ดังนั้นภาพที่ได้ก็จะเหลือเพียงแค่ตัวโน้ตและสัญลักษณ์เท่านั้น เราก็จะทำการตรวจสอบหาตำแหน่งตัวโน้ตและสัญลักษณ์แต่ละตัว ก่อนจะส่งเข้าสู่กระบวนการรู้จำรูปแบบต่อไป ดังนั้นเมื่อผ่านขั้นตอนต่างๆทั้งหมดเราก็จะทราบว่าเป็นตัวโน้ตหรือสัญลักษณ์ใด และมีความหมายว่าอย่างไร เราก็นำข้อมูลที่ได้ไปแปลงเป็นไฟล์มิดี้นั่นเอง

6.2 ขั้นตอนการทำงานของโครงงาน

จากที่ได้กล่าวมาข้างต้น สามารถนำขั้นตอนการทำงานโดยรวมของโครงงาน มาเขียนเป็นโฟลว์ชาร์ตได้ดังนี้



รูปที่ 6-1 แสดงขั้นตอนการทำงานของโครงงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1 การรับภาพอินพุต

การรับภาพภาพอินพุตที่จะนำมาใช้กับโปรแกรมนั้น จะทำการรับภาพโน้ตเข้ามาโดยผ่านการสแกนแล้วเก็บเป็นไฟล์ภาพ ซึ่งรูปแบบของภาพนั้น สามารถเป็นได้ 4 รูปแบบคือ BMP, JPG, GIF, PNG โดยจะเป็นภาพสีหรือขาวดำก็ได้และขนาดของภาพนั้นควรจะไม่เกิน 1500*1500 พิกเซล

2 การทำพรีโพรเซสซิ่ง

ภาพอินพุตที่รับเข้ามานั้นบางครั้งภาพอาจจะมี Noise อยู่หรือภาพอาจจะอยู่ในระนาบเอียงเราจึงจำเป็นต้องมีกระบวนการทำพรีโพรเซสซิ่ง เพื่อจัดการกับภาพอินพุตนั้นให้เหมาะสมก่อนนำไปวิเคราะห์ ซึ่งการทำพรีโพรเซสซิ่งนี้จะทำอยู่ 2 อย่างด้วยกันคือทำการลบ Noise และปรับภาพที่เอียงให้อยู่ในระนาบปกติ

3 การแยกบรรทัด 5 เส้นออกจากภาพ

ภาพโน้ตดนตรีที่เรารับเข้ามาจะถูกบันทึกอยู่บนบรรทัด 5 เส้น เพื่อบอกให้ทราบถึงระดับสูงต่ำของโน้ตแต่ละตัว การที่เราจะทำการแยกบรรทัด 5 เส้นออกนั้น เหตุผลก็เพื่อให้กระบวนการในการตรวจสอบหาพิกเซลของโน้ตและสัญลักษณ์ต่างๆนั้น สามารถทำได้ง่ายและมีความถูกต้องมากขึ้น แต่ก่อนที่จะมีการลบออกไปจะต้องทำการเก็บตำแหน่งของบรรทัด 5 เส้นทั้งหมดเอาไว้ด้วย เพื่อใช้ตรวจสอบหาระดับเสียงของโน้ตในภายหลัง

การหาตำแหน่งเส้นบรรทัด จะเริ่มโดยการหาขอบเขตของภาพที่แท้จริงก่อน จากนั้นจึงทำการสแกนพิกเซลของภาพในแนวนอนว่ามีเนื้อภาพเรียงต่อกันเป็นเส้นตรงเกินกว่าค่าที่กำหนดหรือไม่ ถ้าใช่ก็แสดงว่าเป็นเส้นบรรทัดก็ให้ทำการเก็บตำแหน่งไว้แล้วจึงลบออกไป ทั้งนี้ก็เพื่อให้เหลือเฉพาะตัวโน้ตและเครื่องหมายที่จะนำไปวิเคราะห์เท่านั้น

4 การหาขอบเขตตัวโน้ต และสัญลักษณ์ทางดนตรี

เมื่อภาพโน้ตดนตรีถูกลบบรรทัด 5 เส้นออกไป ก็จะเหลือเพียงแต่โน้ตและสัญลักษณ์เท่านั้น เราก็ทำการหาขอบเขตของโน้ตและสัญลักษณ์แต่ละตัวโดยใช้วิธีเซกเมนต์เทชั่น (Segmentation) ซึ่งหลักการของวิธีนี้จะทำการหาขอบเขตของภาพที่เป็นเนื้อเดียวกัน โดยการกระจายออกไปทุกทิศทางรอบพิกเซลจนกว่าจะสิ้นสุดเนื้อภาพ ก็จะได้ขอบเขตของโน้ตและสัญลักษณ์แต่ละตัวออกมา

5 การรู้จำรูปแบบโน้ต และสัญลักษณ์ทางดนตรี

หลักจากที่เราทราบตำแหน่งของโน้ตและสัญลักษณ์แต่ละตัวแล้ว เราก็นำไปวิเคราะห์เพื่อตรวจสอบว่าเป็นโน้ตหรือสัญลักษณ์ใด และมีความหมายว่าอย่างไร ก่อนจะแปลงเป็นไฟล์มิดี้ในลำดับต่อไปนั่นเอง ทั้งนี้การที่เราจะสามารถวิเคราะห์เพื่อแยกแยะความแตกต่างได้นั้นจะมีวิธีเฉพาะสำหรับโน้ตและสัญลักษณ์แต่ละตัวดังจะได้อธิบายให้เห็นในหัวข้อต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6 การแปลงข้อมูลให้อยู่ในรูปของไฟล์ MIDI

เมื่อเราสามารถทราบว่าเป็นโน้ตหรือสัญลักษณ์ใดแล้วเราก็ทำการนำข้อมูลทั้งหมดที่ได้ ไปแปลงให้อยู่ในรูปของไฟล์มิดี้ เพื่อให้สามารถเล่นเป็นเสียงเพลงได้ ซึ่งรูปแบบของไฟล์มิดีนั้นได้กล่าวไว้แล้วในเนื้อหาก่อนหน้านี้

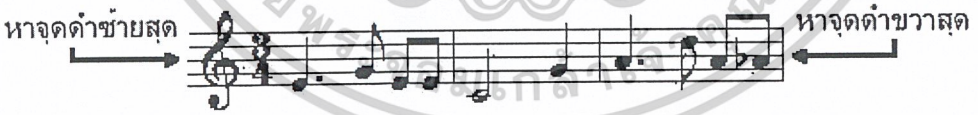
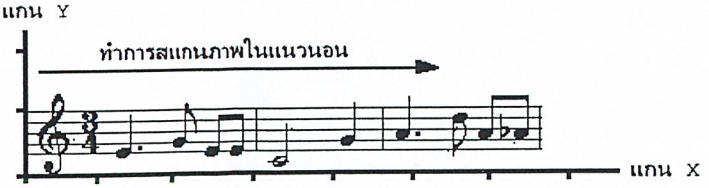
6.3 การแยกบรรทัด 5 เส้นออกจากภาพ

การแยกเส้นบรรทัดออกจากภาพ มีเหตุผลอยู่ 2 ประการคือ ประการแรกก็เพื่อทำการหาและเก็บตำแหน่งของเส้นบรรทัดไว้ใช้ในการหาระดับเสียงสูงต่ำของตัวโน้ต ประการที่สองก็เพื่อให้กระบวนการในการตรวจสอบหาแหน่งของตัวโน้ตและสัญลักษณ์ต่างๆนั้นสามารถทำได้สะดวกและง่ายขึ้น ซึ่งกระบวนการแยกเส้นบรรทัดนี้จะต้องทำ 3 ขั้นตอนคือ หาตำแหน่งบรรทัด 5 เส้น, ลบบรรทัด 5 เส้น และเติมส่วนที่ขาดหายเนื่องจากทำการลบเส้นบรรทัดออกไป

6.3.1 การหาตำแหน่งบรรทัด 5 เส้น

หลังจากที่ได้ทำพรีโพรเซสซิ่งแล้ว กระบวนการต่อจากนั้นเราจะทำการสแกนหาเส้นบรรทัด เพื่อจะได้เก็บตำแหน่งไว้แล้วแยกเส้นบรรทัดนั้นออกไป วิธีการหาตำแหน่งเส้นบรรทัดเราจะเริ่มต้นโดยการตรวจสอบหาจุดสีดำซ้ายสุดและขวาสุดของภาพก่อน เพื่อหาขอบเขตของภาพที่แท้จริง ทั้งนี้เพื่อป้องกันความผิดพลาดอันเนื่องมาจากภาพมีช่วงขอบที่กว้างทำให้การหาเส้นโดยเทียบกับขนาดของภาพนั้นไม่ได้ผล จากนั้นจึงทำการสแกนพิกเซลของภาพในแนวนอน หากพบว่าเป็นเป็นพิกเซลสีดำ (มีไบนารีเท่ากับ 0) และมีความยาวเกินกว่า 60 เปอร์เซ็นต์ของเนื้อภาพเราก็จะรู้ว่าเป็นเส้นบรรทัด ก็ให้ทำการเก็บตำแหน่งเส้นบรรทัดนั้นๆเอาไว้

รูป 6-2(a) จะเป็นการตรวจสอบหาเนื้อภาพที่แท้จริงโดยการหาจุดดำซ้ายสุด และขวาสุดของภาพ

รูป 6-2(b) แสดงการสแกนภาพในแนวนอน (แกน x) เพื่อหาตำแหน่งของบรรทัด 5 เส้น

(a) แสดงการหาขอบเขตของเนื้อภาพ

(b) แสดงการสแกนภาพในแนวนอนเพื่อหาบรรทัด 5 เส้น

รูปที่ 6-2 แสดงภาพโน้ตดนตรี และการสแกนภาพในแนวนอน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.3.2 การลบบรรทัด 5 เส้น

ในขั้นตอนก่อนหน้าทำให้เราทราบตำแหน่งของบรรทัด 5 เส้น ดังนั้นในส่วนนี้เราก็จะทำการลบเส้นบรรทัดนั้นออกไป

แต่ในส่วนของเส้นน้อยเราจะไม่สามารถใช้วิธีการดังกล่าวได้ เพราะว่าความยาวของเส้นน้อยไม่ถึง 60 เปอร์เซ็นต์ เราจึงนำเทคนิค Erosion มาใช้ในการลบส่วนของเส้นน้อยแทน

รูปที่ 6-3 แสดงภาพโน้ตดนตรีที่ผ่านการการลบบรรทัด 5 เส้น และเส้นน้อยออกไป แต่ทั้งนี้จะเห็นได้ว่าบางส่วนของตัวโน้ตและเครื่องหมายต่างๆนั้นได้หายไปด้วย ดังนั้นจึงต้องมีการเติมส่วนที่ขาดหาย ซึ่งจะได้กล่าวในหัวข้อต่อไป

รูปที่ 6-3 แสดงภาพตัวโน้ตที่ผ่านการลบบรรทัด 5 เส้น และเส้นน้อยออกไป

6.3.3 การคืนภาพให้ตัวโน้ต

ภายหลังจากที่ได้ทำการลบบรรทัด 5 เส้นและเส้นน้อยออกไป ส่งผลให้ภาพบางส่วนของโน้ตและเครื่องหมายต่างๆถูกลบออกไปด้วย ดังนั้นเราจึงต้องทำการคืนค่าให้กับส่วนที่ขาดหาย เพื่อให้ภาพที่จะนำไปวิเคราะห์นั้นสมบูรณ์คงเดิม ดังแสดงในรูปที่ 6-4

สำหรับการคืนค่าให้กับส่วนที่ขาดหายนั้น เราจะอ้างอิงกับตำแหน่งของบรรทัด 5 เส้นที่เก็บไว้ ซึ่งการคืนค่าเราจะตรวจสอบว่าอยู่ใน 3 เงื่อนไขที่เรากำหนดไว้หรือไม่ เงื่อนไขที่กล่าวถึงนั้นมีดังต่อไปนี้

- เงื่อนไขแรก ณ ตำแหน่งของเส้นบรรทัดนั้นๆ หากมีพิกเซลด้านบนและด้านล่างเป็นสีดำทั้งคู่ก็ให้คืนค่าสีดำเข้าไปตรงกลาง ดังแสดงในรูปที่ 6-5(a)
- เงื่อนไขที่สอง ณ ตำแหน่งของเส้นบรรทัดนั้นๆ หากมีพิกเซลด้านบนเป็นสีดำและด้านล่างเป็นสีขาวก็ให้ทำการคืนค่าสีดำเข้าไปตรงกลาง ดังแสดงในรูปที่ 6-5(b)
- เงื่อนไขที่สาม ณ ตำแหน่งของเส้นบรรทัดนั้นๆ หากมีพิกเซลด้านบนเป็นสีขาวและด้านล่างเป็นสีดำก็ให้ทำการคืนค่าสีดำเข้าไปตรงกลาง ดังแสดงในรูปที่ 6-5(c)

แต่หากว่า ณ ตำแหน่งของเส้นบรรทัดนั้นๆ มีพิกเซลด้านบนและด้านล่างเป็นสีขาวทั้งคู่ ซึ่งไม่ตรงกับเงื่อนไขทั้ง 3 ข้อที่เราตั้งไว้ เราก็จะไม่ทำการคืนค่าเข้าไป ดังแสดงในรูปที่ 6-5(d)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(a) การลบบรรทัด 5 เส้นออก *(b) การคืนส่วนของตัวโน้ต*

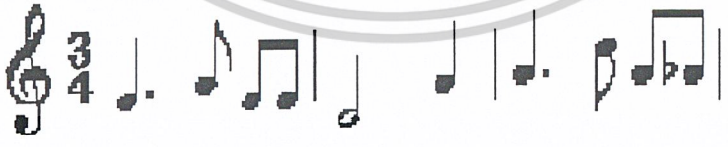
รูปที่ 6-4 แสดงการลบบรรทัด 5 เส้น และการคืนส่วนของตัวโน้ตที่ถูกลบ

(a) การคืนค่าที่พิกเซลบนและล่างเป็นสีดำ *(b) การคืนค่าที่พิกเซลด้านบนเป็นสีดำ*

(c) การคืนค่าที่พิกเซลด้านล่างเป็นสีดำ *(d) ลักษณะของรูปที่ไม่ต้องมีการคืนค่า*

รูปที่ 6-5 แสดงลักษณะการคืนภาพตัวโน้ต

ซึ่งภายหลังจากที่ได้ทำการคืนค่าส่วนที่ขาดหายเนื่องจากการลบบรรทัด 5 เส้นครบทุกตัวแล้วนั้น ก็จะได้ตัวโน้ตและเครื่องหมายต่างๆที่สมบูรณ์เหมือนเดิม ดังแสดงในรูปที่ 6-6



รูปที่ 6-6 แสดงภาพโน้ตที่ได้ผ่านการคืนค่าส่วนที่ขาดหายเสร็จเรียบร้อย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.4 การหาขอบเขตตัวโน้ตและสัญลักษณ์

กระบวนการทั้งหมดที่กล่าวมาตั้งแต่ต้นจนถึงตอนนี้ จะทำให้เราได้ภาพที่เหลือเพียงแค่ตัวโน้ตและเครื่องหมายต่างๆที่เกี่ยวข้อง อีกทั้งยังมีตำแหน่งของเส้นบรรทัดที่เราได้เก็บไว้ด้วย ซึ่งหลังจากนี้ไปก็จะเป็นส่วนที่สำคัญของโครงงานนั่นก็คือ ส่วนของการตรวจสอบหาขอบเขตตัวโน้ตและเครื่องหมายทั้งหมด และส่วนของกระบวนการรู้จำรูปแบบ

ในส่วนของการการหาขอบเขตตัวโน้ต เราจะใช้เทคนิคในการแยกข้อมูลภาพออกเป็นส่วนๆ หรือที่เรียกว่า Image Segmentation วิธีนี้จะทำการแยกองค์ประกอบต่างๆของรูปภาพออกจากกันตามลักษณะสำคัญที่เราทำการพิจารณา ขั้นตอนนี้ถือเป็นขั้นตอนที่สำคัญของการวิเคราะห์ภาพ ซึ่งข้อดีของการแยกข้อมูลภาพออกเป็นส่วนๆคือ

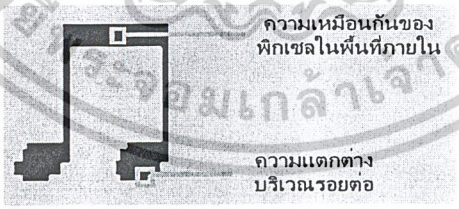
- ลดข้อมูลในภาพที่ไม่จำเป็นในการวิเคราะห์ลง
- จัดระเบียบข้อมูลในภาพให้เป็นกลุ่มได้ดีขึ้น
- แสดงข้อมูลในรูปที่เข้าใจได้ง่ายขึ้น

6.4.1 หลักการแยกองค์ประกอบ

ในการที่จะสามารถแยกแยะองค์ประกอบต่างๆของภาพออกมาได้นั้น หลักการสำคัญที่ถูกนำมาใช้ในการแยกองค์ประกอบจะมีอยู่ 2 หลักสำคัญๆคือ

1. การแยกองค์ประกอบตามความเหมือนกัน (Similarity) ของคุณสมบัติแต่ละพิกเซล ของรูปภาพภายในพื้นที่เดียวกัน
2. การแยกองค์ประกอบโดยดูจากความไม่ต่อเนื่อง (Discontinuity) ของคุณสมบัติของพิกเซลบริเวณรอยต่อระหว่างวัตถุในภาพกับฉากหลัง

หลักการแยกองค์ประกอบทั้ง 2 รูปแบบแสดงได้ดังรูปที่ 6-7



รูปที่ 6 –7 แสดงการแยกองค์ประกอบ

สำหรับเทคนิคการแยกข้อมูลภาพออกเป็นส่วนๆนั้นมีอยู่หลายวิธีด้วยกัน ซึ่งกลุ่มหลักๆที่มักถูกนำมาใช้นั้นก็อาทิเช่น

- Pixel Oriented Image Segmentation
- Region Oriented Image Segmentation
- Edge Oriented Image Segmentation

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับโครงงานที่ได้จัดทำนั้น ได้ทำการใช้เทคนิค Region Growing ซึ่งอยู่ในกลุ่มของ Region Oriented Image Segmentation ทั้งนี้เพราะว่าภาพหลังจากผ่านการลบเส้นบรรทัดออกจะเหลือเพียงแค่ตัวโน้ตและเครื่องหมาย ซึ่งแต่ละโน้ตและเครื่องหมายจะอยู่แยกจากกัน รวมทั้งความแตกต่างกับพื้นหลังก็สามารถแยกออกได้ชัดเจน ดังนั้นเมื่อนำวิธีนี้มาใช้ทำให้สามารถจัดการกับภาพได้ง่าย และได้ขอบเขตของโน้ตและเครื่องหมายที่ชัดเจน

กลุ่มของ Region Oriented Image Segmentation นั้นเป็นวิธีการแยกองค์ประกอบของภาพ โดยดูจากตำแหน่งของพิกเซลและความเหมือนกันของคุณสมบัติของพิกเซลภายในพื้นที่ โดยถ้าพิกเซลที่อยู่ติดกันและมีคุณสมบัติเหมือนกันจะถูกจัดให้เข้ากลุ่มเดียวกัน ซึ่งข้อดีของวิธีนี้จะทำให้เราได้พื้นที่ของภาพที่เราสนใจที่ต่อเนื่อง

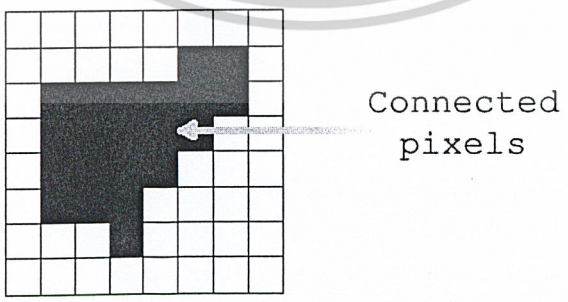
6.4.2 Region Growing

วิธีการ Region Growing จะนำพิกเซลข้างเคียงมาพิจารณาซึ่งจะทำการจัดกลุ่ม (Region) ของพิกเซลเหล่านี้เข้าไว้ด้วยกันโดยพิจารณาถึงความเข้มของพิกเซล (ค่าของพิกเซลมีค่าใกล้เคียงกัน)

ในการแยกองค์ประกอบของภาพ จะต้องมีการกำหนดกลุ่มที่ต้องการจะแบ่ง ***Ri*** โดยที่แต่ละกลุ่มจะต้องมีการกำหนดค่าความเข้มของพิกเซลเริ่มต้น $Si, i=1,.....N$ (ค่าเหล่านี้เราจะทำการกำหนดเอง) ซึ่งมีไว้เพื่อใช้สำหรับขยายตัวของกลุ่ม(growth) ในการขยายตัวของกลุ่มนี้จำเป็นจะต้องมีกฎเพื่อใช้เป็นวิธีการขยายตัวของกลุ่มรวมทั้งกฎของการตรวจสอบความเป็นอันหนึ่งอันเดียวกันของกลุ่ม (homogeneity) ของทุกระยะของการขยายตัว

ในแต่ละระยะของการขยายตัวของกลุ่ม $R_i^{(K)}, i = 1,....., N$ จะมีการตรวจสอบว่ามีพิกเซลที่ยังไม่ได้จัดกลุ่มหรือไม่ (เป็นพิกเซลทั้ง 8 ที่เชื่อมต่อ อยู่กับพิกเซลที่อยู่บริเวณขอบของกลุ่ม) และก่อนที่จะมีการกำหนดพิกเซลใดๆ (x) เข้าสู่กลุ่มใดกลุ่มหนึ่งจะต้องมีการตรวจสอบอีกว่ากลุ่มที่จะขยายนั้นยังคงมีความเป็นอันหนึ่งอันเดียวกันอีกหรือไม่ (Homogeneity) ดังสมการที่ 6.4.2-1

$$R(R_i^{(k)} \cup \{x\}) = TRUE \qquad \text{(สมการที่ 6.4.2-1)}$$



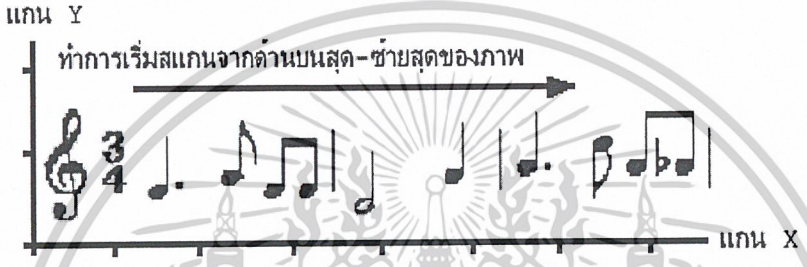
รูปที่ 6-8 แสดงการเชื่อมต่อกันของแต่ละพิกเซล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นประสิทธิภาพของการแยกองค์ประกอบของภาพวิธีนี้ จะขึ้นอยู่กับการเลือกค่าความเข้มของพิกเซลเริ่มต้นของในแต่ละกลุ่มที่เราได้กำหนดไว้ แต่ในกรณีของภาพที่เรานำมาใช้จะเป็นภาพที่ผ่านการปรับเป็น 2 ระดับแล้ว ดังนั้นจึงมีความแตกต่างระหว่างวัตถุกับพื้นหลังอย่างชัดเจน จึงทำให้การกำหนดค่าความเข้มเริ่มต้นนั้นสามารถทำได้ง่าย และได้การหาขอบเขตที่มีประสิทธิภาพ

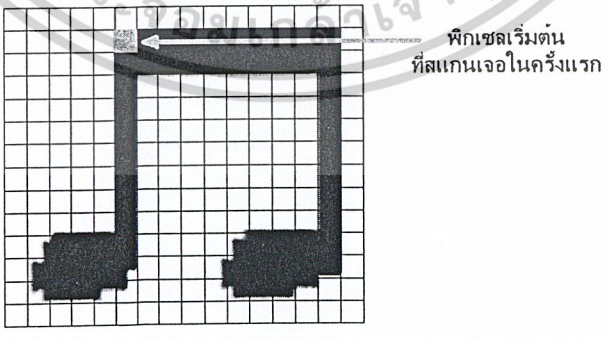
6.4.3 ขั้นตอนการหาตำแหน่ง

ในการหาตำแหน่งของตัวโน้ตและเครื่องหมายต่างๆทั้งหมด จะเริ่มโดยการสแกนจากด้านซ้ายสุดบนสุดไปจนถึงขวาสุดล่างสุดของภาพ เพื่อหาพิกเซลส่วนที่เป็นเนื้อภาพ (พิกเซลที่เป็นสีดำ) ดังแสดงในรูปที่ 6-9



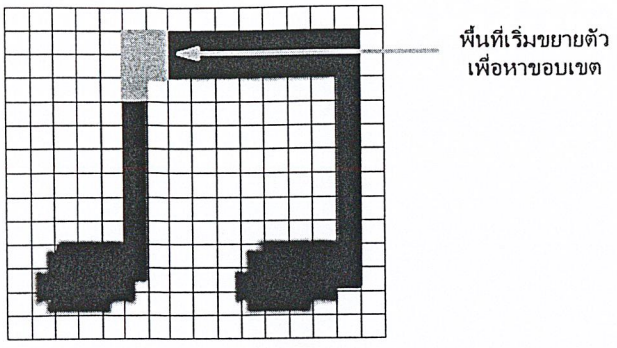
รูปที่ 6-9 แสดงการหาตำแหน่งในขั้นตอนแรก

เมื่อพบส่วนที่เป็นเนื้อภาพก็จะเริ่มทำการหาขอบเขตโดยใช้ วิธีการ Region Growing ตามที่กล่าวมา โดยจะเริ่มทำการค้นหาพิกเซลรอบข้างที่มีคุณสมบัติเหมือนกัน (มีค่าของพิกเซลใกล้เคียงกัน) ทำให้ขอบเขตของโน้ตหรือสัญลักษณ์นั้นๆ เริ่มโตขึ้นเรื่อยๆ ทั้งนี้การหาขอบเขตของโน้ตหรือสัญลักษณ์นั้นๆ จะสิ้นสุดลงเมื่อไม่สามารถที่จะขยายพื้นที่ออกไปได้มากว่านี้ เราก็จะได้ขอบเขตของโน้ตหรือสัญลักษณ์ตัวนั้นๆ จากนั้นเราก็เริ่มหาขอบเขตของตัวอื่นๆต่อไป ลำดับการหาขอบเขตแสดงได้ดังรูปที่ 6-10

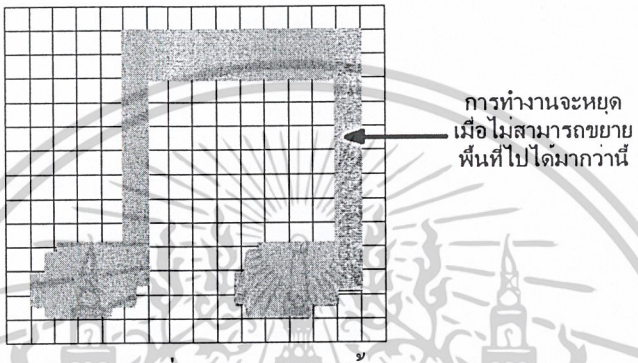


(a) พิกเซลแรกที่สแกนเจอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



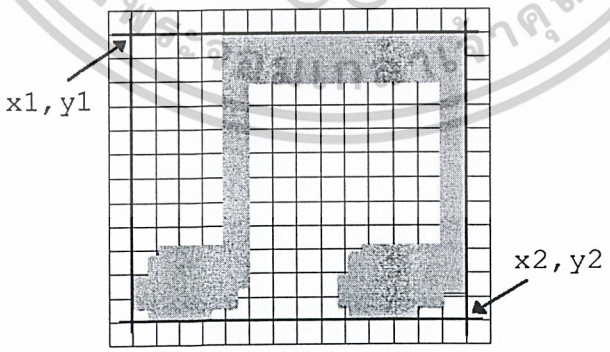
(b) ขอบเขตเริ่มขยายตัว



(c) เมื่อหาขอบเขตจนสิ้นสุด

รูปที่ 6-10 แสดงลำดับขั้นตอนการหาขอบเขต

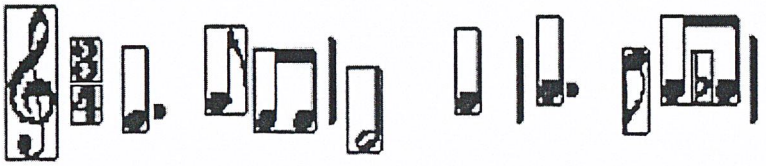
ภายหลังจากที่หาขอบเขตของโน้ตหรือเครื่องหมายนั้นๆเสร็จสิ้น เราก็จะทราบถึงตำแหน่งและขอบเขตของมัน โดยดูจากตำแหน่ง x ซ้ายสุด และ y บนสุด เทียบกับตำแหน่ง x ขวาสุด และ y ล่างสุด ก็จะได้เป็นขอบเขตออกมา ซึ่งเพื่อให้เข้าใจได้ง่ายขึ้นสามารถแสดงได้ในรูปที่ 6-11



รูปที่ 6-11 แสดงตำแหน่งของตัวโน้ตที่ได้จากการหาขอบเขต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในทุกๆครั้งที่เราหาขอบเขตของตัวโน้ตและสัญลักษณ์ได้ เราจะทำการเก็บตำแหน่งของแต่ละตัวเอาไว้ เพื่อจะได้นำโน้ตและเครื่องหมายนั้นๆไปผ่านกระบวนการรู้จำในขั้นตอนต่อไป โดยหลังจากที่ทำการหาขอบเขตของโน้ตและเครื่องหมายครบหมดทุกตัว ก็จะได้ขอบเขตแต่ละตัวดังแสดงในรูปที่ 6-12



รูปที่ 6-12 แสดงขอบเขตของโน้ตและเครื่องหมายทั้งหมด

6.5 การรู้จำรูปแบบตัวโน้ต และสัญลักษณ์ทางดนตรี

กระบวนการรู้จำรูปแบบเป็นส่วนที่มีความสำคัญต่อโครงงานในครั้งนี้เป็นอย่างมาก เพราะเป้าหมายหลักของเราคือการแปลงจากรูปภาพโน้ตดนตรีให้กลายเป็นไฟล์มิดี้เพื่อเล่นเป็นเสียงเพลงออกมา ดังนั้นหากกระบวนการรู้จำนี้ขาดประสิทธิภาพในการประมวลผล ขั้นตอนต่างๆที่ทำมาดีตั้งแต่ต้นก็ไม่มีความหมายใด

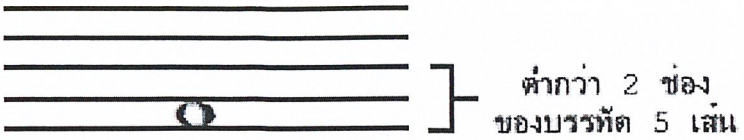
เมื่อในตอนนี้เราทราบขอบเขต, ตำแหน่งของตัวโน้ตและเครื่องหมายทั้งหมด เราก็จะนำแต่ละตัวนั้นมาทำการพิจารณา เพื่อวิเคราะห์ว่าเป็นตัวอะไรและในการใช้งานนั้นมีความหมายว่าอย่างไรบ้าง โดยรูปแบบการพิจารณาจะมีเงื่อนไขที่แตกต่างกันออกไป ตามลักษณะเฉพาะของโน้ตและเครื่องหมายแต่ละตัว

เงื่อนไขในการพิจารณาเพื่อแยกแยะตัวโน้ตและสัญลักษณ์นั้น ในโครงงานได้จำแนกออกเป็น 3 กลุ่มใหญ่ๆด้วยกันคือ

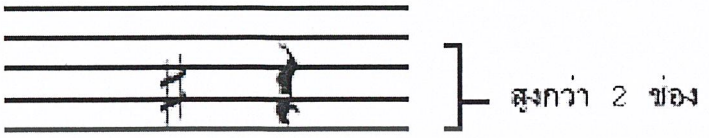
1. กลุ่มที่มีความสูงต่ำกว่า 2 ช่องบรรทัด 5 เส้น
2. กลุ่มที่มีความสูงเกินกว่า 2 ช่องแต่ไม่ถึง 3 ช่องของบรรทัด 5 เส้น
3. กลุ่มที่มีความสูงเกินกว่า 3 ช่องของบรรทัด 5 เส้น

(a) แสดงระยะของบรรทัด 5 เส้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(b) แสดงตัวอย่างกลุ่มที่มีความสูงต่ำกว่า 2 ช่องของบรรทัด 5 เส้น



(c) แสดงตัวอย่างกลุ่มที่มีความสูงเกิน 2 ช่องแต่ไม่ถึง 3 ช่องของบรรทัด 5 เส้น

(d) แสดงตัวอย่างกลุ่มที่มีความสูงเกินกว่า 3 ช่องของบรรทัด 5 เส้น

รูปที่ 6-13 แสดงตัวอย่างการจำแนกโน้ตและสัญลักษณ์

วิธีที่นำมาใช้วิเคราะห์ในกระบวนการรู้จำนั้น แต่ละกลุ่มก็จะมีวิธีที่แตกต่างกันออกไปตามลักษณะเฉพาะของตัวโน้ตและเครื่องหมายที่อยู่ในกลุ่มนั้นๆ วิธีที่ได้นำมาใช้ก็อาทิเช่น การพิจารณาจากการเปลี่ยนแปลงค่าสี, Matching by correlation เป็นต้น

6.5.1 กลุ่มที่มีความสูงต่ำกว่า 2 ช่องของบรรทัด 5 เส้น

เมื่อพิจารณาตัวโน้ตและเครื่องหมายต่างๆที่มีความสูงต่ำกว่า 2 ช่องของบรรทัด 5 เส้นสามารถแบ่งได้เป็น 2 กลุ่มด้วยกันคือ

- กลุ่มที่มีความสูงภายใน 1 ช่อง
- กลุ่มที่มีความสูงมากกว่า 1 ช่อง

6.5.1.1 กลุ่มทีมีความสูงภายใน 1 ช่อง

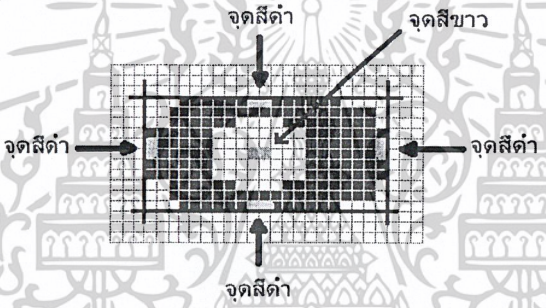
เมื่อพิจารณาถึงโน้ตและเครื่องหมายที่อยู่ในกลุ่มนี้จะมีอยู่ 4 ตัวด้วยกัน ดังที่แสดงในตารางที่ 6-1 ซึ่งการแยกแยะความแตกต่างภายในกลุ่มเพื่อชี้ชัดว่าเป็นโน้ตหรือเครื่องหมายใด จะใช้การพิจารณาจากลักษณะเฉพาะของแต่ละตัว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัญลักษณ์	ความหมาย
	โน้ตตัวกลม
	ตัวหยุดตัวกลม
	ตัวหยุดตัวขาว
	จุด

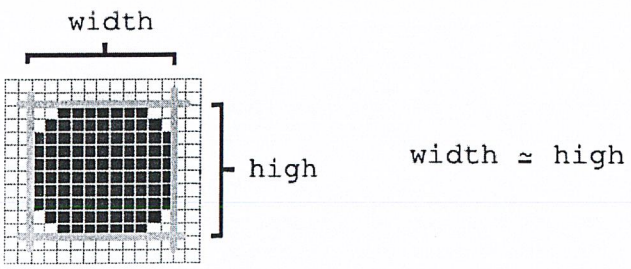
ตารางที่ 6-1 แสดงโน้ตในกลุ่มที่มีความสูงภายใน 1 ช่องของบรรทัด 5 เส้น

ในตอนเริ่มแรกเราจะทำการพิจารณาโดยดูจากการเปลี่ยนแปลงของค่าสีภายใน ซึ่งจะทำการตรวจเช็ค 5 จุดรอบขอบเขตของแต่ละตัวคือ บน ล่าง ซ้าย ขวา หากพบว่าเป็นสีดำทั้งหมดและตำแหน่งกึ่งกลางเป็นสีขาว เราก็จะสามารถแยกออกจากกลุ่มและระบุได้ว่าเป็นโน้ตตัวกลมนั่นเอง ทั้งนี้สัญลักษณ์ตัวอื่นๆจะไม่ตรงตามเงื่อนไขที่เราตรวจเช็ค การพิจารณาโน้ตตัวกลมแสดงได้ในรูปที่ 6-14



รูปที่ 6-14 แสดงการพิจารณาการเปลี่ยนแปลงค่าสีเพื่อระบุว่าเป็นโน้ตตัวกลม

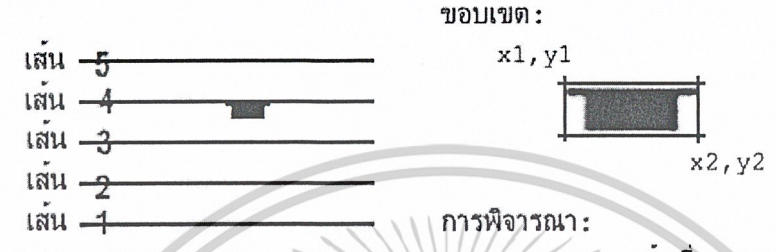
ขั้นตอนต่อไปเมื่อเราแยกโน้ตตัวกลมออกจากกลุ่มได้ก็จะเหลือเพียง ตัวหยุดตัวกลม ตัวหยุดตัวขาว และจุด ซึ่งเมื่อพิจารณาจากทั้ง 3 ตัวจะเห็นได้ว่า สัญลักษณ์จุดจะมีลักษณะที่เป็นสี่เหลี่ยมจัตุรัส คือขอบเขตจะมีความกว้างและความสูงที่ใกล้เคียงกัน ส่วนอีก2 ตัวคือตัวหยุดตัวกลมและตัวหยุดตัวขาวนั้นจะมีความกว้างที่มากกว่าความสูง ดังนั้นจากลักษณะเด่นนี้เราก็จะสามารถแยกสัญลักษณ์จุด ออกจากกลุ่มได้ การพิจารณาสัญลักษณ์จุดซึ่งมีความกว้างและความสูงที่ใกล้เคียงกันแสดงในรูปที่ 6-15



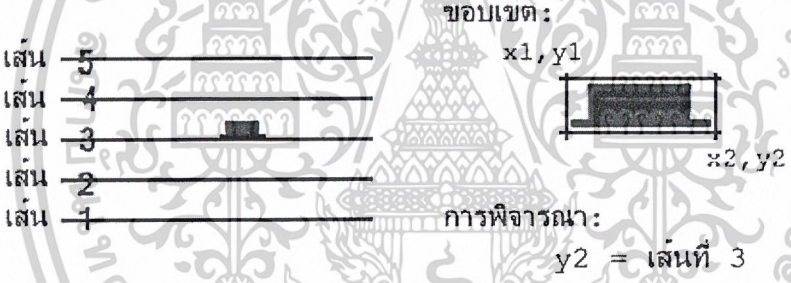
รูปที่ 6-15 แสดงสัญลักษณ์จุดซึ่งมีความกว้างและความสูงคล้ายสี่เหลี่ยมจัตุรัส

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในตอนนี้สัญลักษณ์ภายในกลุ่มก็จะเหลืออีกสองตัวคือ ตัวหยุดตัวกลมและตัวหยุดตัวขาว ซึ่งการพิจารณาเพื่อแยกแยะทั้ง 2 ตัวนี้ออกจากกันสามารถทำได้ง่ายมาก เพราะว่าทั้งสองตัวนั้นมีรูปแบบที่เฉพาะในการเขียนลงบนบรรทัด 5 เส้น คือ ตัวหยุดตัวกลมจะเขียนติดอยู่ล่างเส้นที่ 4 (พิจารณาจาก y1 = เส้นที่4) และตัวหยุดตัวขาวจะเขียนติดอยู่บนเส้นที่ 3 (พิจารณาจาก y2 = เส้นที่3) ดังนั้นในกลุ่มนี้ก็จะสามารถแยกแยะความแตกต่างเพื่อระบุว่าเป็นสัญลักษณ์อะไรได้ทั้งหมด การแยกระหว่างตัวหยุดตัวกลมและตัวหยุดตัวขาวแสดงในรูปที่ 6-16



(a) ตัวหยุดตัวกลม

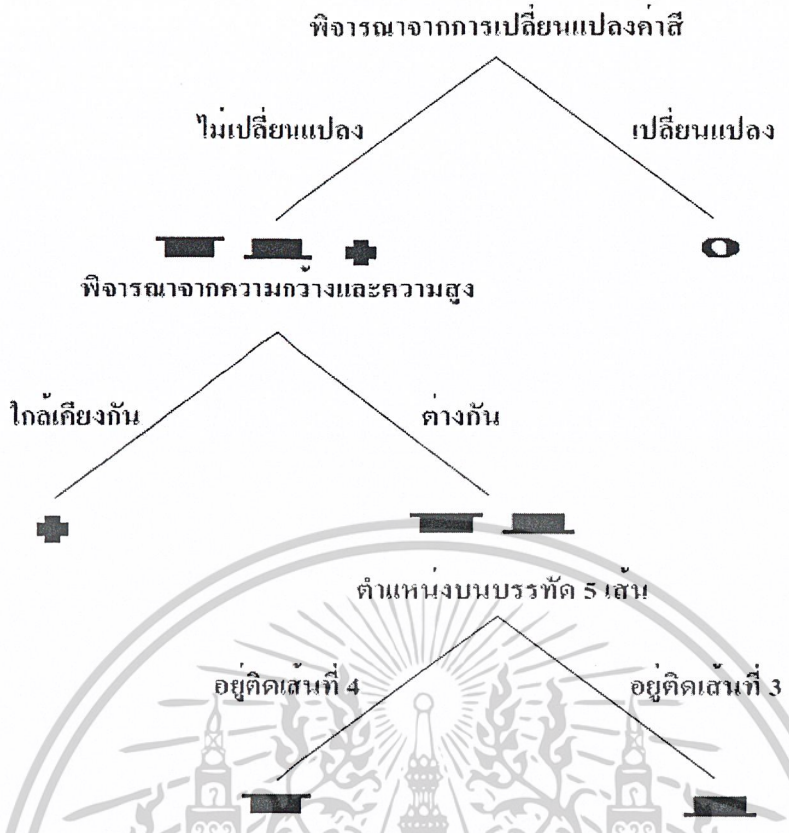


(b) ตัวหยุดตัวขาว

รูปที่ 6-16 แสดงการพิจารณาตัวหยุดตัวกลมและตัวหยุดตัวขาวจากตำแหน่งที่ต่างกัน

จากที่กล่าวมาทั้งหมดเราก็จะสามารถแยกแยะสัญลักษณ์ทั้ง 4 ตัวออกจากกัน โดยอาศัยลักษณะเฉพาะของแต่ละตัว ซึ่งสามารถสรุปถึงวิธีการที่นำมาใช้ในการพิจารณาได้ดังในรูปที่ 6-17

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6-17 สรุปวิธีการแยกตัวโน้ตและเครื่องหมาย ของกลุ่มที่มีความสูงภายใน 1 ช่องของบรรทัด 5 เส้น

6.5.1.2 กลุ่มที่มีความสูงมากกว่า 1 ช่อง

เมื่อพิจารณาในกลุ่มนี้จะมีอยู่ 5 ตัวด้วยกันดังแสดงในตารางที่ 6-2 ซึ่งจะสังเกตได้ว่าหากเราใช้การพิจารณาจากลักษณะเหมือนในกลุ่มแรก จะทำได้ยากเพราะว่าแต่ละตัวนั้นไม่ค่อยมีลักษณะเด่นที่สามารถนำมาใช้เพื่อแยกแยะแต่ละตัวออกจากกันได้ทั้งหมด ดังนั้นในกลุ่มนี้จะนำเอาเทคนิคของการเปรียบเทียบภาพ หรือเรียกอีกอย่างคือ Matching by correlation มาใช้

สัญลักษณ์	ความหมาย
𝄾	ตัวหยุดตัวเขบิ็ตหนึ่งชั้น
2 3 4 6	ตัวเลขที่ใช้ในอัตราส่วนกำหนดจังหวะ

ตารางที่ 6-2 แสดงสัญลักษณ์ในกลุ่มที่มีความสูงมากกว่า 1 ช่องแต่ไม่เกิน 2 ช่องของบรรทัด 5 เส้น

สำหรับขั้นตอนต่างๆในการวิเคราะห์โดยใช้การเปรียบเทียบรูปภาพนั้น ก็มีกระบวนการดังที่ได้กล่าวไปแล้วในส่วนของทฤษฎีการเปรียบเทียบรูปภาพในบทที่ 5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.5.2 กลุ่มที่มีความสูงเกินกว่า 2 ช่องแต่ไม่ถึง 3 ช่องของบรรทัด 5 เส้น

สัญลักษณ์ที่อยู่ในกลุ่มนี้จะมีอยู่ 5 ตัวด้วยกัน ดังแสดงในตารางที่ 6-3 และเช่นเดียวกันในกลุ่มนี้หากเราจะทำการพิจารณาจากลักษณะ จะทำได้ยากเพราะว่าสัญลักษณ์แต่ละตัวไม่มีลักษณะเด่นที่จะนำมาแยกแยะได้อย่างชัดเจน ดังนั้นเราก็จะนำเอาเทคนิคการเปรียบเทียบภาพมาใช้

สัญลักษณ์	ความหมาย
	ตัวหยุดตัวดำ
	ตัวหยุดเขบิ็ด 2 ชั้น
	เครื่องหมายชาร์ป
	เครื่องหมายแฟล็ต
	เครื่องหมายเนเจอรัล

ตารางที่ 6-3 แสดงสัญลักษณ์ในกลุ่มที่มีความสูงเกินกว่า 2 ช่องแต่ไม่ถึง 3 ช่องของบรรทัด 5 เส้น

สำหรับขั้นตอนต่างๆในการวิเคราะห์โดยใช้การเปรียบเทียบรูปภาพนั้น ก็มีกระบวนการดังที่ได้กล่าวไปแล้วในส่วนของทฤษฎีการเปรียบเทียบรูปภาพในบทที่ 5

6.5.3 กลุ่มที่มีความสูงเกินกว่า 3 ช่องของบรรทัด 5 เส้น

ในกลุ่มนี้จะใช้การพิจารณาจากลักษณะเฉพาะของแต่ละตัว ซึ่งลักษณะเฉพาะที่จะนำมาใช้ก็ได้แก่การพิจารณาจากส่วนของเส้นตรงที่ยาวที่สุด และการพิจารณาจากการเปลี่ยนแปลงค่าสี โดยเมื่อทำการพิจารณาจากเส้นตรงแล้วจะสามารถแบ่งออกได้เป็น 2 กลุ่มคือ

- กลุ่มที่มีเส้นตรงเพียงเส้นเดียว
- กลุ่มที่มีเส้นตรงมากกว่า 1 เส้น

6.5.3.1 กลุ่มที่มีเส้นตรงเพียงเส้นเดียว

จากตารางที่ 6-4 จะเห็นได้ว่าสัญลักษณ์แต่ละตัวนั้น เมื่อพิจารณาจากตำแหน่งของเส้นตรงที่ยาวสุดจะสามารถจำแนกออกเป็น 3 กลุ่มคือ

- กลุ่มที่มีเส้นตรงที่ยาวที่สุดอยู่ทางด้านซ้าย
- กลุ่มที่มีเส้นตรงที่ยาวที่สุดอยู่ทางด้านขวา
- กลุ่มที่มีเส้นตรงที่ยาวที่สุดอยู่ตรงกลาง

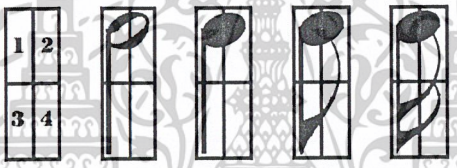
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัญลักษณ์			ความหมาย
เส้นตรงอยู่ด้านซ้าย	เส้นตรงอยู่ด้านขวา	เส้นตรงอยู่ตรงกลาง	
			โน้ตตัวขาว
			โน้ตตัวดำ
			โน้ตเขบิ็ด 1 ชั้น
			โน้ตเขบิ็ด 2 ชั้น

ตารางที่ 6-4 สัญลักษณ์ในกลุ่มที่มีความสูงเกินกว่า 3 ช่องของบรรทัด 5 เส้น

6.5.3.1.1 กลุ่มที่มีเส้นตรงที่ยาวที่สุดอยู่ทางด้านซ้าย

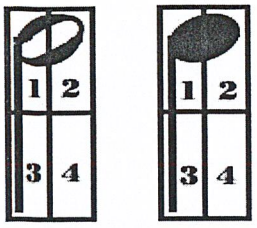
ในกลุ่มนี้ก็จะมีสัญลักษณ์ให้พิจารณาอยู่ 4 ตัว เราก็จะทำการแยกแยะความแตกต่างโดยการแบ่งสัญลักษณ์ทุกตัวออกเป็น 4 ส่วน ได้ดังรูปที่ 6-18



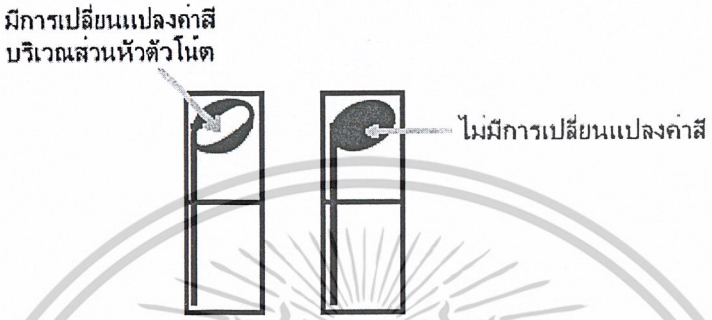
รูปที่ 6-18 แสดงการแบ่งโน้ตออกเป็น 4 ส่วน

หลังจากที่ทำการแบ่งออกเป็น 4 ส่วนจะเห็นได้ว่า โน้ตตัวขาวและโน้ตตัวดำจะมีจุดภาพที่เป็นสีดำอยู่ในส่วนที่ 1, 2 และ 3 เท่านั้น แสดงในรูปที่ 6-19(a) ดังนั้นเราจึงสามารถแยก 2 ตัวนี้ออกจากกลุ่ม จากกนั้นก็พิจารณาการเปลี่ยนแปลงค่าสีบริเวณส่วนหัวเพื่อแยกแยะความแตกต่างระหว่างโน้ตตัวขาวและตัวดำเป็นลำดับต่อไป การแยกแยะโน้ตตัวขาวและตัวดำแสดงดังรูปที่ 6-19(b)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



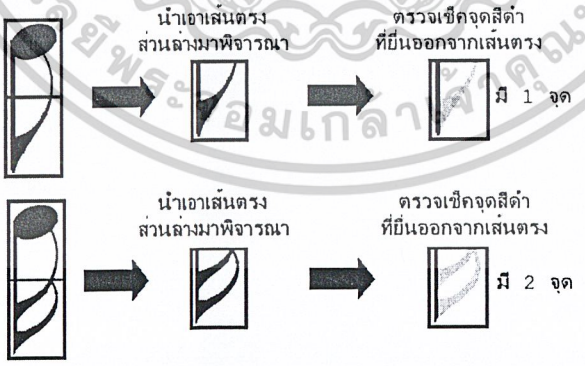
(a) โน้ตตัวขาวและตัวดำมีจุดภาพที่เป็นสีดำเฉพาะในส่วนที่ 1,2 และ3



(b) การพิจารณาการเปลี่ยนแปลงค่าสีบริเวณส่วนหัว

รูปที่ 6-19 แสดงการวิเคราะห์เพื่อแยกแยะโน้ตตัวขาว และโน้ตตัวดำ

เมื่อโน้ตตัวขาวและโน้ตตัวดำสามารถแยกแยะได้แล้ว ก็จะเหลืออีกสองตัวสำหรับกลุ่มนี้คือ ตัวเขบ็ด 1ชั้น และตัวเขบ็ด 2 ชั้น ซึ่งการแยกแยะทั้งสองตัวนี้จะเริ่มโดยการแบ่งครึ่งเส้นตรง จากนั้นนำเอาเส้นตรงส่วนล่างมาพิจารณาเพื่อดูว่ามีจุดสีดำที่ต่อออกมาจากเส้นตรงทั้งหมดกี่แห่ง ทั้งนี้เพื่อให้เข้าใจได้ง่ายขึ้น สามารถแสดงลำดับการวิเคราะห์ได้ดังรูปที่ 6-20

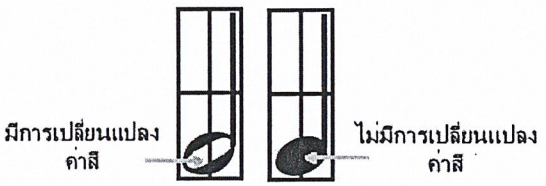


รูปที่ 6-20 แสดงการวิเคราะห์เพื่อแยกแยะตัวเขบ็ด 1 ชั้น และตัวเขบ็ด 2 ชั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.5.3.1.2 กลุ่มที่มีเส้นตรงที่ยาวที่สุดอยู่ทางด้านขวา

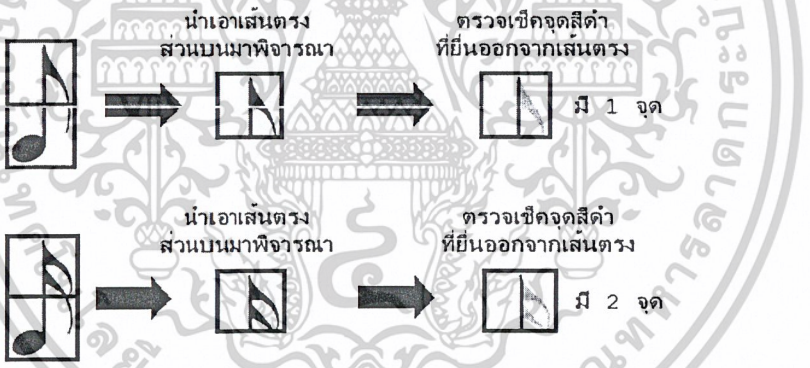
ในกลุ่มนี้ก็จะพิจารณาจากการเปลี่ยนแปลงค่าสีค่าบริเวณส่วนหัวของโน้ตเช่นเดียวกัน ซึ่งโน้ตตัวขาวนั้นจะมีการเปลี่ยนแปลงค่าสีบริเวณส่วนหัวตัวโน้ต ส่วนโน้ตตัวขาวจะไม่มีการเปลี่ยนแปลงค่าสี จึงทำให้สามารถรู้ได้ว่าเป็นโน้ตอะไร การพิจารณาแสดงในรูปที่ 6-21



รูปที่ 6-21 แสดงการเปลี่ยนแปลงค่าสีของโน้ตที่มีเส้นตรงอยู่ทางด้านขวา

6.5.3.1.3 กลุ่มที่มีเส้นตรงที่ยาวที่สุดอยู่ตรงกลาง

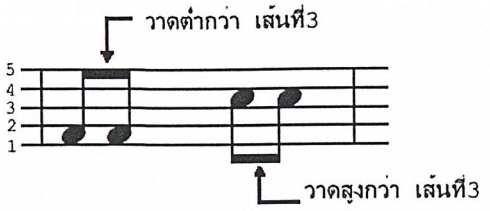
โน้ตกลุ่มนี้ก็จะมีอยู่สองตัวเช่นกัน การวิเคราะห์ก็จะทำการแบ่งครึ่งเส้นตรงออกเป็น 2 ส่วนแล้วนำเอาส่วนบนมาทำการพิจารณา เพื่อดูว่ามีจุดสีดำที่ต่อออกมาจากเส้นตรงทั้งหมดกี่แห่ง ทั้งนี้เพื่อให้เข้าใจได้ง่ายขึ้น สามารถแสดงลำดับการวิเคราะห์ได้ดังรูปที่ 6-22

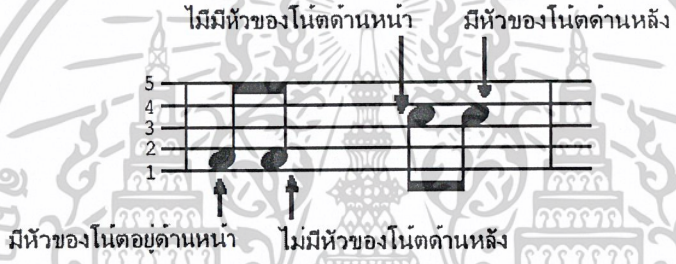


รูปที่ 6-22 แสดงการวิเคราะห์เพื่อแยกแยะตัวเขบ็ด 1 ชั้น และตัวเขบ็ด 2 ชั้น

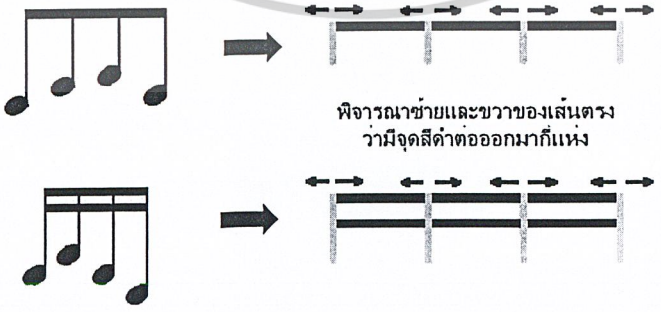
6.5.3.2 กลุ่มที่มีเส้นตรงมากกว่า 1 เส้น

กลุ่มนี้จะเกิดจาก การที่ตัวเขบ็ด 1 ชั้นหรือ 2 ชั้นอยู่ติดกันตั้งแต่ 2 ตัวขึ้นไปทำให้มีส่วนของเส้นตรงที่มากกว่า 1 เส้น การทีเราจะรู้ได้ว่าอยู่ติดกันกี่ตัวนั้นเราก็จะนับจากจำนวนของเส้นตรงที่มีจากนั้นเราก็จะทำการตรวจสอบอีกทีว่า เป็นตัวเขบ็ด 1 ชั้นหรือ 2ชั้น แต่ทั้งนี้ในการวาดโน้ตลงบนบรรทัด 5 เส้นจะมีเงื่อนไขอยู่ว่า หากโน้ตถูกวาดตั้งแต่เส้นที่ 3 ของบรรทัด 5 เส้นขึ้นไป หางของตัวโน้ตจะต้องชี้ลงด้านล่างแต่หากโน้ตถูกวาดไม่เกินเส้นที่ 3 เราก็จะวาดตามปกติคือมีหางชี้ขึ้นด้านบน ความแตกต่างของการวาดโน้ตแสดงได้ดังรูปที่ 6-23

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 6-23 แสดงความแตกต่างของการวาดโน้ตลงบนบรรทัด 5 เส้น

ดังนั้นจากความแตกต่างนี้เราจะทำการตรวจสอบก่อนว่าหางของโน้ตที่เราจะวิเคราะห์ชี้ขึ้นหรือลง ซึ่งหากเราพิจารณาที่เส้นซ้ายสุดและเส้นขวาสุดของแต่ละชุดเราจะสามารถแยกแยะได้ กล่าวคือโน้ตที่มีหางชี้ขึ้น เส้นซ้ายสุดจะมีหัวของโน้ตอยู่ด้านหน้า, เส้นขวาสุดจะ ไม่มีหัวของโน้ตอยู่ด้านหลัง ส่วนโน้ตที่มีหางชี้ลงเส้นซ้ายสุดจะไม่มีหัวของโน้ตอยู่ด้านหน้า, เส้นขวาสุดจะมีหัวของโน้ตตามมา ดังแสดงในรูปที่ 6-24

รูปที่ 6-24 แสดงการแยกความแตกต่างระหว่างโน้ตที่มีหางชี้ขึ้นและชี้ลง

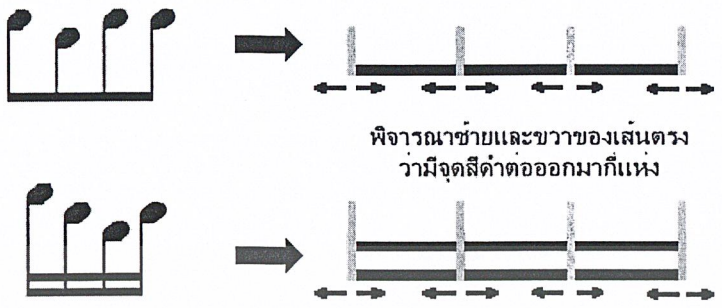
เมื่อในตอนนี้เราทราบได้ว่าโน้ตนั้นมีหางชี้ขึ้นหรือลงเราก็จะมาวิเคราะห์ต่อว่าเป็นตัวเขบ็ด 1 ชั้น หรือ 2 ชั้น โดยหากเป็นโน้ตที่มีหางชี้ขึ้นเราก็จะทำการแบ่งครึ่งเอาส่วนบนซึ่งเป็นหางมาพิจารณาว่าซ้ายและขวาของเส้นตรงแต่ละเส้นมีจุดสีดำต่อออกมาทั้งหมดกี่แห่งก็จะทราบได้ว่าเป็นเขบ็ด 1 ชั้นหรือ 2 ชั้น แสดงได้ดังรูปที่ 6-25



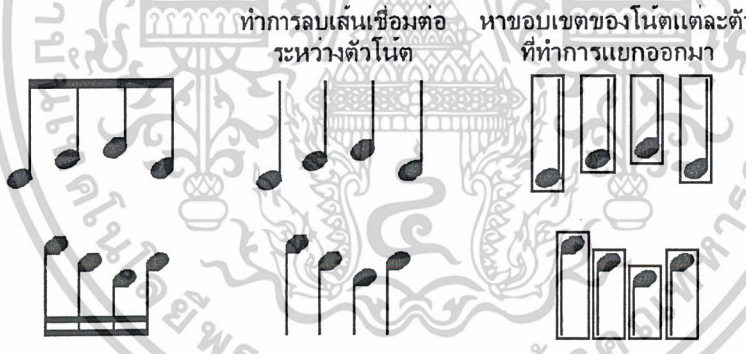
รูปที่ 6-25 แสดงการพิจารณาตัวโน้ตที่มีหางชี้ขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แต่หากเป็นโน้ตที่มีหางชี้ลงเราก็จะทำการแบ่งครึ่งเอาส่วนล่างซึ่งเป็นหางมาพิจารณาว่าซ้ายและขวาของเส้นตรงแต่ละเส้นมีจุดสีดำต่อออกมาทั้งหมดกี่แห่งเช่นกัน ก็จะทราบได้ว่าเป็นเขบ็ด 1 ชั้นหรือ 2 ชั้น แสดงได้ดังรูปที่ 6-26

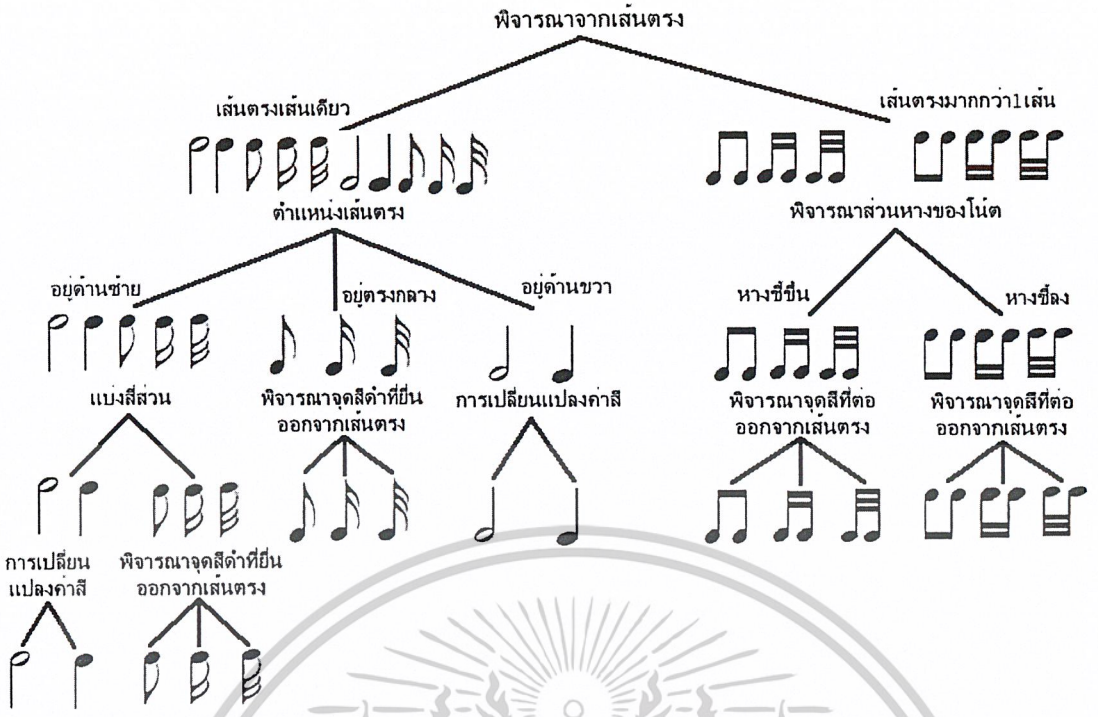


รูปที่ 6-26 แสดงการพิจารณาตัวโน้ตที่มีหางชี้ลง

ภายหลังจากที่ผ่านกระบวนการในการแยกแยะทั้งหมดของโน้ตในกลุ่มนี้เราก็จะทราบได้ว่าเป็นโน้ตอะไรและอยู่ติดกันทั้งหมดกี่ตัว ขั้นตอนสุดท้ายเราก็จะทำการลบเส้นเชื่อมต่อระหว่างโน้ตแต่ละตัวออก จากนั้นหาขอบเขตใหม่ของแต่ละตัวที่แยกออกมา เพื่อให้สามารถตรวจสอบหาระดับเสียงสูงต่ำสำหรับแต่ละตัวในขั้นตอนต่อไปได้ การหาขอบเขตใหม่แสดงในรูปที่ 6-27

รูปที่ 6-27 แสดงการหาขอบเขตใหม่ของโน้ตกลุ่มที่มีเส้นตรงมากกว่า 1 เส้น

ในตอนนี้เราก็สามารถแยกแยะโน้ตและเครื่องหมาย ในกลุ่มที่มีความสูงเกินกว่า 3 ช่องของบรรทัด 5 เส้นได้ทั้งหมด ซึ่งสมารถสรุปวิธีที่นำมาใช้เป็นลำดับๆ ดังแสดงในรูปที่ 6-28

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



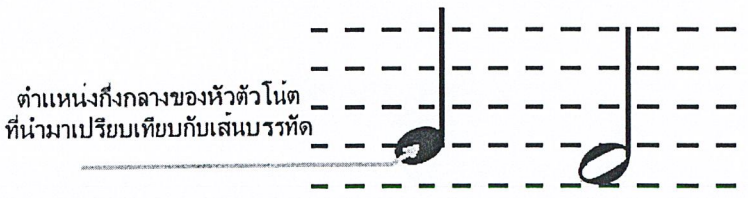
รูปที่ 6-28 สรุปวิธีการแยกตัวโน้ตในกลุ่มที่มีความสูงเกินกว่า 3 ช่องของบรรทัด 5 เส้น

6.5.4 การหาระดับเสียงสูงต่ำของตัวโน้ต

เมื่อเราสามารถแยกแยะโน้ตหรือเครื่องหมายได้ทั้งหมด และสามารถบอกได้ว่าแต่ละตัวมีความหมายว่าอย่างไร ขั้นตอนสุดท้ายของกระบวนการรู้จำเราก็จะทำการตรวจสอบหาระดับเสียงสูงต่ำของตัวโน้ตเพื่อนำไปใช้ในการสร้างไฟล์มิดี้

ทั้งนี้ในการหาระดับเสียงสูงต่ำของโน้ตแต่ละตัว เราจะพิจารณาโดยการนำเอาจุดกึ่งกลางส่วนหัวของตัวโน้ต มาเทียบกับตำแหน่งเส้นบรรทัดที่เราได้เก็บเอาไว้ เพื่อหาว่าวางอยู่ ณ ตำแหน่งใด ดังนั้นเราก็สามารถทราบได้ถึงระดับเสียงของโน้ตแต่ละตัว

การหาระดับเสียงสูงต่ำแสดงได้ดังรูปที่ 6-29 โดยเส้นประเป็นเส้นที่สมมุติขึ้นเพื่อแสดงตำแหน่งของเส้นบรรทัดที่เราได้เก็บไว้ อันจะทำให้เข้าใจการหาระดับเสียงสูงต่ำได้ง่ายขึ้น ดังนั้นโน้ตตัวดำตัวแรกก็จะมีระดับเสียงคือ ซอล (G) และตัวขาวมีระดับเสียงคือ ฟา (F)



รูปที่ 6-29 แสดงการหาระดับเสียงสูงต่ำของตัวโน้ต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.6 การแปลงข้อมูลให้อยู่ในรูปของไฟล์มิดี้

เมื่อภาพโน้ตที่เรานำมาประมวลผลผ่านขั้นตอนของการรู้จำรูปแบบ เราก็จะได้ข้อมูลที่จะนำไปสร้างเป็นไฟล์มิดี้ (*.MID) อันได้แก่ ชนิดของกุญแจ, อัตราส่วนกำหนดเวลา, ค่าของ Delta Time, ชนิดของตัวโน้ตหรือเครื่องหมายนั้นๆ และระดับเสียงสูง-ต่ำของโน้ต ซึ่งข้อมูลที่กล่าวมาทั้งหมดนั้นจะถูกเก็บอยู่ในอาร์เรย์ (Array)

สำหรับกระบวนการสร้างไฟล์มิดี้นั้น จะต้องทำอยู่ 3 ส่วนคือ

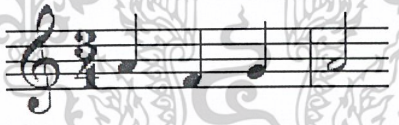
<Header Chunk> <Meta Event> <Midi Event>

ส่วนของ Header ที่เราจะต้องสร้างนั้นก็จะเป็นการกำหนดเกี่ยวกับรูปแบบของไฟล์มิดี้ที่จะสร้างว่าเป็น Single Track หรือ Multiple Track, จำนวนของแทร็กที่มี และความยาวเสียงของโน้ตตัวดำที่ใช้เป็นมาตรฐานอ้างอิงกับตัวอื่น

ส่วนของ Meta Event ในส่วนนี้จะเป็นการเก็บข้อมูลเกี่ยวกับอัตราส่วนกำหนดจังหวะ และความเร็วในการเล่นเพลง(Tempo)

ส่วนของ Midi Event ส่วนนี้ข้อมูลส่วนใหญ่ก็จะได้มาจากกระบวนการรู้จำรูปแบบ ข้อมูลที่จะเก็บก็มี ชนิดของเครื่องดนตรี(ผู้ใช้เลือกในโปรแกรม), ชื่อตัวโน้ตซึ่งแสดงถึงระดับเสียงสูงต่ำ, ค่า Delta Time ของโน้ตแต่ละตัว รวมไปถึงการกำหนดช่วงเล่นโน้ตและเลิกเล่นโน้ตอีกด้วย

6.6.1 ข้อมูลที่ได้จากการรู้จำรูปแบบ

ตัวอย่างภาพที่จะนำไปผ่านการรู้จำรูปแบบแสดงได้ดังรูปที่ 6-30 ซึ่งข้อมูลที่ได้จากการรู้จำรูปแบบมีดังต่อไปนี้

รูปที่ 6-30 แสดง ตัวอย่างโน้ตที่จะทำการสร้างไฟล์มิดี้

จากรูปที่ 6-31 เมื่อผ่านกระบวนการรู้จำแล้ว ข้อมูลต่างๆที่ได้ออกประมวลผลก็จะมีดังนี้

- ชนิดกุญแจ	กุญแจซอล
- อัตราส่วนกำหนดจังหวะ	เมื่อนำไปรวมกับชุดคำสั่งก็จะได้เป็น FF 58 04 **03 02** 24 08
- โน้ตตัวแรก	เป็นโน้ตตัวดำ, มีระดับเสียงเป็นเสียง ลา (A), มี Delta Time เท่ากับ 83 00h
- โน้ตตัวที่สอง	เป็นโน้ตตัวดำ, มีระดับเสียงเป็นเสียง ฟา (F), มี Delta Time เท่ากับ 83 00h
- โน้ตตัวที่สาม	เป็นโน้ตตัวดำ, มีระดับเสียงเป็นเสียง ซอล (G), มี Delta Time เท่ากับ 83 00h

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- โน้ตตัวที่สี่ เป็นโน้ตตัวขาว, มีระดับเสียงเป็นเสียง ลา(A),
มี Delta Time เท่ากับ 8C 00h

6.6.2 ขั้นตอนการสร้างไฟล์มิดี้

ตามที่กล่าวไปแล้วว่าการสร้างไฟล์มิดี้จะต้องทำอยู่ 3 ส่วนซึ่งสามารถอธิบายแต่ละส่วนได้ดังนี้

6.6.2.1 ส่วน Header

ส่วนนี้จะเริ่มโดยระบุประเภทของแทร็ก จากนั้นก็ตามด้วยความยาวของส่วน Header แล้วตามด้วยรูปแบบของแทร็ก และความยาวเสียงของโน้ตตัวดำที่จะใช้เป็นมาตรฐาน ก็จะได้ออกมาดังนี้

4D 54 68 64	→	Chunk Type ในรหัส ASCII มีค่าเป็น MThd (Header Chunk)
00 00 00 06	→	ความยาวของ Header เท่ากับ 6 Bytes
00 01	→	ฟอร์แมตของ MIDI นี้คือ Format 1 (Multiple Track)
00 02	→	ไฟล์ MIDI นี้มีจำนวน 2 Track
01 80	→	กำหนดความยาวเสียงของโน้ตตัวดำเท่ากับ 384 (ฐาน 10)

6.6.2.2 ส่วน Meta Event

ส่วนนี้ก็จะเริ่มด้วยการระบุประเภทของแทร็กเช่นกัน จากนั้นก็จะตามด้วยความยาวของแทร็กนี้ เสร็จแล้วก็จะใส่ข้อมูลในส่วนของอัตราส่วนกำหนดจังหวะ และกำหนดความเร็วในการเล่นเพลง ก็จะได้ออกมาดังนี้

4D 54 72 6B	→	Chunk Type ในรหัส ASCII มีค่าเป็น MTrk (Track Chunk)
00 00 00 10	→	ความยาวของ Track นี้เท่ากับ 16 Bytes
FF 58 04 03 02 24 08	→	ทำการกำหนดจังหวะของห้องเพลง
FF 51 03 07 A1 20	→	ทำการกำหนดความเร็วของการเล่นต่อโน้ตตัวดำ 1 ตัว

6.6.3 ส่วน Midi Event

เช่นเดียวกันก็จะเริ่มต้นด้วยการกำหนดประเภทของแทร็ก จากนั้นก็จะตามด้วยความยาวทั้งหมดของแทร็กนี้ ต่อมาก็จะทำการกำหนดชนิดของเสียงเครื่องดนตรี ซึ่งได้มาจากการเลือกในโปรแกรมก่อนที่จะทำการรู้จำรูปแบบ จากนั้นก็จะเริ่มทำการกำหนดช่วงเล่นโน้ต และหยุดเล่นโน้ตเป็นจังหวะๆตามค่า Delta Time ของโน้ตตัวนั้นๆ เมื่อทำจนครบทุกตัวแล้วก็จะจบด้วยแทร็กปิด ก็จะได้ออกมาดังนี้

4D 54 72 6B	→	Chunk Type ในรหัส ASCII มีค่าเป็น MTrk (Track Chunk)
00 00 00 2B	→	ความยาวข้อมูล Track นี้เท่ากับ 43 Bytes

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

C0 18	→	C0 คือ การตั้งเครื่องดนตรีที่ Channel 1
		18 คือ หมายเลขเครื่องดนตรี มีค่าเป็น 24 ในเลขฐาน 10
90 45 60	→	90 คือ ให้ Channel 1 ทำการเริ่มเล่นโน้ต
		45 หมายถึง โน้ตที่จะเล่นคือ A
		60 คือ ความหนักของเสียงที่เล่น
83 00	→	ค่าเวลาความยาวของเสียงของโน้ตตัวดำ
80 45	→	80 คือให้ Channel 1 ทำการหยุดเล่นโน้ต
		45 คือ โน้ตที่จะหยุดเล่นคือ A
90 41 60	→	90 คือ ให้ Channel 1 ทำการเริ่มเล่นโน้ต
		41 หมายถึง โน้ตที่จะเล่นคือ F
		60 คือ ความหนักของเสียงที่เล่น
83 00	→	ค่าเวลาความยาวของเสียงของโน้ตตัวดำ
80 41	→	80 คือให้ Channel 1 ทำการหยุดเล่นโน้ต
		41 คือ โน้ตที่จะหยุดเล่นคือ F
90 43 60	→	90 คือ ให้ Channel 1 ทำการเริ่มเล่นโน้ต
		43 หมายถึง โน้ตที่จะเล่นคือ G
		60 คือ ความหนักของเสียงที่เล่น
83 00	→	ค่าเวลาความยาวของเสียงของโน้ตตัวดำ
80 43	→	80 คือให้ Channel 1 ทำการหยุดเล่นโน้ต
		43 คือ โน้ตที่จะหยุดเล่นคือ G
90 45 60	→	90 คือ ให้ Channel 1 ทำการเริ่มเล่นโน้ต
		45 หมายถึง โน้ตที่จะเล่นคือ A
		60 คือ ความหนักของเสียงที่เล่น
86 00	→	ค่าเวลาความยาวของเสียงของโน้ตตัวขาว
80 45	→	80 คือให้ Channel 1 ทำการหยุดเล่นโน้ต
		45 คือ โน้ตที่จะหยุดเล่นคือ A
FF 2F 0	→	จบ Track

เมื่อทำการแปลงให้อยู่ในรูปของไฟล์ MIDI แล้วก็จะได้ออกมาดังนี้ (แสดงเป็นเลขฐาน 16)

```
4D 54 68 64 00 00 00 06  00 01 00 02 01 80 4D
54 72 6B 00 00 00 10 00  FF 58 04 03 02 24 08
00 FF 51 03 07 A1 20 00  4D 54 72 6B 00 00 00
2B 00 C0 18 00 90 45 60  83 00 80 45 00 00 90
41 60 83 00 80 41 00 00  90 43 60 83 00 80 43
00 00 90 45 60 86 00 80  45 00 00 FF 2F 00
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 7

ผลการทดลอง

7.1 บทนำ

ในส่วนของการทดลองนั้นจะเป็นการนำเอาภาพโน้ตและเครื่องหมายต่างๆ มาทำการทดสอบกับโปรแกรมในเงื่อนไขที่แตกต่างกันไปซึ่งมีอยู่ 3 กรณีคือ กรณีภาพที่รับเข้ามาอยู่ในระนาบปกติ, กรณีภาพที่รับมาอยู่ในระนาบเอียงซ้าย, กรณีภาพที่รับมาอยู่ในระนาบเอียงขวา ทั้งนี้กรณีภาพที่รับมาอยู่ในระนาบเอียงนั้นได้ทำการทดสอบภาพที่เอียงในระดับองศาต่างๆ เพื่อพิจารณาถึงความถูกต้องของโปรแกรมที่ได้ทำการพัฒนาว่ามีประสิทธิภาพในการรู้จำครอบคลุมในทุกเงื่อนไขหรือไม่

7.2 ประสิทธิภาพในการรู้จำ

จากที่ได้กล่าวไปแล้วว่าวิธีการที่ได้นำมาใช้เพื่อทดสอบประสิทธิภาพในการรู้จำนั้น ได้แบ่งออก 3 เงื่อนไข โดยแต่ละเงื่อนไขจะมีจำนวนสัญลักษณ์ที่นำมาทดสอบทั้งหมด100 ตัว

เงื่อนไขที่ 1 การทดสอบภาพสัญลักษณ์ต่างๆที่อยู่ในระนาบปกติ โดยกรณีนี้จะมีการเปรียบเทียบประสิทธิภาพกับโปรแกรมเดิมก่อนที่จะทำการพัฒนาด้วย

สัญลักษณ์ (100ตัว)	ภาพในระนาบปกติ	
	โปรแกรมเดิม	โปรแกรมพัฒนา
ตัวหยุดตัวกลม	100 %	100 %
ตัวหยุดตัวขาว	100 %	100 %
ตัวหยุดตัวดำ	74 %	100 %
ตัวหยุดเขบ็ด 1 ชั้น	100 %	100 %
ตัวหยุดเขบ็ด 2 ชั้น	-	100 %
เนเจอรัล	74 %	100 %
ชาร์ป	74 %	100 %
แฟลท	82 %	100 %
โน้ตประจุด	70 %	100 %
ตัวกลม	90 %	100 %
ตัวขาว	98 %	100 %
ตัวดำ	98 %	100 %
ตัวเขบ็ด 1 ชั้น	90 %	100 %

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวเขบิ๊ด 2 ชั้น	-	100 %
ตัวเขบิ๊ด 3 ชั้น	-	100 %

ตารางที่ 7-1 แสดงการทดสอบประสิทธิภาพการรู้จำของภาพในระนาบปกติ

เงื่อนไขที่ 2 การทดสอบภาพสัญลักษณ์ต่างๆที่อยู่ในระนาบเอียงซ้าย ในกรณีนี้ไม่ได้มีการเปรียบเทียบกับโปรแกรมเดิมก่อนทำการพัฒนา ทั้งนี้เพราะว่าโปรแกรมเดิมนั้นไม่สามารถทำการประมวลผลภาพที่อยู่ในระนาบเอียงได้นั่นเอง

สัญลักษณ์ (100ตัว)	ภาพในระนาบเอียงซ้าย (องศา)			
	5	10	20	45
ตัวหยุดตัวกลม	100 %	92 %	100 %	100 %
ตัวหยุดตัวขาว	100 %	95 %	100 %	98 %
ตัวหยุดตัวดำ	50 %	90 %	90 %	85 %
ตัวหยุดเขบิ๊ด 1 ชั้น	100 %	97 %	97 %	100 %
ตัวหยุดเขบิ๊ด 2 ชั้น	90 %	70 %	95 %	95 %
เนเจอรัล	60 %	30 %	20 %	20 %
ชาร์ป	98 %	95 %	95 %	96 %
แฟลท	95 %	94 %	94 %	90 %
โน้ตประจุด	60 %	65 %	60 %	70 %
ตัวกลม	60 %	85 %	60 %	60 %
ตัวขาว	98 %	75 %	80 %	75 %
ตัวดำ	85 %	84 %	90 %	80 %
ตัวเขบิ๊ด 1 ชั้น	85 %	85 %	90 %	85 %
ตัวเขบิ๊ด 2 ชั้น	60 %	60 %	65 %	70 %
ตัวเขบิ๊ด 3 ชั้น	55 %	60 %	60 %	70 %

ตารางที่ 7-2 แสดงการทดสอบประสิทธิภาพการรู้จำของภาพที่เอียงซ้าย

จากตารางที่7-2 จะเห็นได้ว่าประสิทธิภาพของการรู้จำจะขึ้นอยู่กับมุมที่เอียงไป ทั้งนี้เนื่องมาจากหากเราทำการปรับภาพที่เอียงไปให้กลับสู่ระนาบปกติ จะส่งผลให้เส้นบรรทัดนั้นมีขนาดที่ใหญ่ขึ้นการลบเส้นบรรทัดจึงทำได้ไม่หมด กระบวนการรู้จำจึงผิดพลาดตามไปด้วยนั่นเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เงื่อนไขที่ 3 การทดสอบภาพสัญลักษณ์ต่างๆที่อยู่ในระนาบเอียงขวา ในกรณีนี้ไม่ได้มีการเปรียบเทียบกับโปรแกรมเดิมก่อนทำการพัฒนา ทั้งนี้เพราะว่าโปรแกรมเดิมนั้นไม่สามารถทำการประมวลผลภาพที่อยู่ในระนาบเอียงได้นั่นเอง

สัญลักษณ์ (100ตัว)	ภาพในระนาบเอียงขวา (องศา)			
	5	10	20	45
ตัวหยุดตัวกลม	55 %	98 %	98 %	99 %
ตัวหยุดตัวขาว	40 %	95 %	95 %	95 %
ตัวหยุดตัวดำ	83 %	83 %	85 %	85 %
ตัวหยุดเขบ็ด 1 ชั้น	60 %	55 %	95 %	95 %
ตัวหยุดเขบ็ด 2 ชั้น	90 %	80 %	90 %	90 %
เนเจอรัล	20 %	25 %	20 %	20 %
ชาร์ป	75 %	75 %	85 %	98 %
แฟลท	75 %	80 %	75 %	72 %
โน้ตประจุด	60 %	60 %	60 %	60 %
ตัวกลม	50 %	80 %	80 %	80 %
ตัวขาว	75 %	80 %	75 %	90 %
ตัวดำ	80 %	80 %	90 %	95 %
ตัวเขบ็ด 1 ชั้น	85 %	80 %	95 %	95 %
ตัวเขบ็ด 2 ชั้น	60 %	60 %	65 %	70 %
ตัวเขบ็ด 3 ชั้น	60 %	60 %	60 %	70 %

ตารางที่ 7-3 แสดงการทดสอบประสิทธิภาพการรู้จำของภาพที่เอียงขวา

เช่นเดียวกันจากตารางที่7-3 จะเห็นได้ว่าประสิทธิภาพของการรู้จำจะขึ้นอยู่กับมุมที่เอียงไป ทั้งนี้เนื่องมาจากหากเราทำการปรับภาพที่เอียงไปให้กลับสู่ระนาบปกติ จะส่งผลให้เส้นบรรทัดนั้นมีขนาดที่ใหญ่ขึ้นการลบเส้นบรรทัดจึงทำได้ไม่หมด กระบวนการรู้จำจึงผิดพลาดตามไปด้วยนั่นเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 8
บทสรุปและแนวทางในการพัฒนา

8.1 บทสรุป

โปรแกรมที่ได้หลังจากทำการพัฒนาเมื่อพิจารณาจากการทำงานโดยรวมแล้ว ถือว่ามีประสิทธิภาพในการทำงานที่เพิ่มขึ้นจากเดิม ซึ่งสามารถสรุปออกมาได้ดังนี้

- โปรแกรมเดิมไม่มีการทำพรีโพรเซสซิ่ง แต่โปรแกรมที่ได้ทำการพัฒนาขึ้นมานั้นมีความสามารถในการทำพรีโพรเซสซิ่ง อันได้แก่การกำจัดสัญญาณรบกวน และการปรับภาพให้อยู่ในระนาบปกติ ส่งผลให้การประมวลผลมีประสิทธิภาพและการสร้างไฟล์มิดี้นั้นมีความถูกต้องที่เพิ่มมากขึ้น
- โปรแกรมมีความยืดหยุ่นเพิ่มขึ้นคือ สามารถที่จะรองรับฟอร์แมตของภาพได้หลากหลายกว่าเดิม,สามารถตรวจสอบหากุญแจ และอัตราส่วนกำหนดจังหวะได้ ทำให้ไฟล์มิดี้ที่ได้มีท่วงทำนองที่ไพเราะขึ้น
- โปรแกรมที่ทำการพัฒนา ส่วนของการรู้จำนั้นสามารถที่จะตรวจสอบโน้ตที่มีความซับซ้อนได้เพิ่มขึ้น คือสามารถที่จะทำการวิเคราะห์หาตัวเขบ็ด 1 ชั้น, 2 ชั้นและ 3 ชั้นที่เขียนติดกันได้สูงสุดถึง 5 ตัว

แม้จะมีประสิทธิภาพในการทำงานที่เพิ่มขึ้น แต่ทั้งนี้บางส่วนของโปรแกรมก็ยังนั้นคงมีข้อจำกัดในการทำงานหลงเหลืออยู่ ซึ่งสามารถจำแนกออกมาได้ดังนี้

1. ภาพที่ได้ภายหลังจากผ่านกระบวนการทำพรีโพรเซสซิ่งนั้น แม้ว่าจะสามารถปรับภาพให้กลับมาอยู่ในระนาบปกติ และสามารถลบ noise ออกไปได้ แต่ในบางครั้งหลังจากที่ทำการปรับภาพที่เอียงให้อยู่ในระนาบปกติ บริเวณขอบของเนื้อภาพ (บริเวณรอบๆจุดภาพที่เป็นสีดำ) จะเกิดการฟุ้งกระจายส่งผลให้ตัวเลขของอัตราส่วนกำหนดจังหวะซึ่งมีขนาดที่เล็ก เมื่อมีการฟุ้งกระจายเกิดขึ้นก็ไม่สามารถที่จะวิเคราะห์ได้ว่าเป็นเลขใด
2. บริเวณส่วนหางของตัวเขบ็ด 2 ชั้นที่ต่อกันหลายๆตัว หากระยะห่างของหางแต่ละชั้นอยู่ใกล้กันมากจะทำให้การตรวจสอบว่าเป็นตัวเขบ็ดกี่ขั้นนั้นเกิดความผิดพลาดได้
3. ไฟล์มิดี้ที่สร้างขึ้นมานั้นจะเป็นเพียงแค่การควบคุมให้เล่นโน้ตหรือหยุดเล่นโน้ตเท่านั้น จึงทำให้อาจฟังดูไม่ค่อยนุ่มนวลเท่าที่ควร
4. โปรแกรมไม่สามารถที่จะทำการตรวจสอบหาเส้นโยงได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8.2 แนวทางในการพัฒนาต่อ

ภายหลังจากที่ได้ทำการศึกษาเกี่ยวกับปัญหาต่างๆ ที่ยังคงหลงเหลืออยู่ในโปรแกรมแล้วนั้นก็สามารถสรุปแนวทางที่สามารถนำพัฒนาต่อเพื่อเพิ่มศักยะภาพของโปรแกรมให้ดียิ่งขึ้นดังนี้

1. ในส่วนของการทำพรีโพรเซสซิ่งนั้น ควรจะหาวิธีการที่สามารถลดการฟุ้งกระจายของขอบภาพที่เกิดจากการปรับภาพให้อยู่ในระนาบปกติ เพื่อให้กระบวนการรู้จำนั้นสามารถวิเคราะห์ได้ถูกต้องมากขึ้นนั่นเอง
2. เพิ่มประสิทธิภาพในการรู้จำให้หลากหลายยิ่งขึ้น เช่น สามารถที่จะตรวจสอบหาเส้นโยงได้
3. พัฒนารูปแบบของไฟล์เอาต์พุตให้หลากหลายขึ้นเช่น แปลงออกมาให้อยู่ในรูปของ Wave, MP3, พัฒนาให้มีการประมวลผลนั้นเป็นแบบ Real Time, พัฒนาโปรแกรมให้สามารถแปลงกลับจากไฟล์มิดี้ออกมาเป็นภาพโน้ตดนตรี เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก.

รหัสค่าเสียงของตัวโน้ต

Octave #	Note Numbers											
	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
-1	0	1	2	3	4	5	6	7	8	9	10	11
0	12	13	14	15	16	17	18	19	20	21	22	23
1	24	25	26	27	28	29	30	31	32	33	34	35
2	36	37	38	39	40	41	42	43	44	45	46	47
3	48	49	50	51	52	53	54	55	56	57	58	59
4	60	61	62	63	64	65	66	67	68	69	70	71
5	72	73	74	75	76	77	78	79	80	81	82	83
6	84	85	86	87	88	89	90	91	92	93	94	95
7	96	97	98	99	100	101	102	103	104	105	106	107
8	108	109	110	111	112	113	114	115	116	117	118	119
9	120	121	122	123	124	125	126	127	55	55		

ตารางที่ ก-1 รหัสค่าเสียงตัวโน้ตเมื่อเทียบเป็นเลขฐานสิบที่ Octave ต่างๆ

มาตรฐานของ MIDI จะกำหนดให้ค่าเสียงตัวโน้ต 60 เป็น C มาตรฐาน (Middle C) ซึ่งรหัสค่าเสียงตัวโน้ตอื่นๆ จะได้มาจากการเทียบกับค่า C มาตรฐานนั่นเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข.
หมายเลขของเครื่องดนตรีแต่ละชนิด

รูปแบบของมิดี้ไฟล์มีเสียงของเครื่องดนตรีอยู่ทั้งหมด 128 เสียงด้วยกัน ซึ่งเราสามารถที่จะเลือกได้โดยการกำหนดให้แทร็กที่จะทำการเล่นนั้น มีหมายเลขเครื่องดนตรีเป็นเสียงของเครื่องดนตรีชิ้นที่เราต้องการ โดยเครื่องดนตรีแต่ละชิ้นมีหมายเลขดังต่อไปนี้

หมายเลข	เครื่องดนตรี
PIANO	
1	Acoustic Grand
2	Bright Acoustic
3	Electric Grand
4	Honky-Tonk
5	Electric Piano 1
6	Electric Piano 2
7	Harpsichord
8	Clavinet
CHROMATIC PERCUSSION	
9	Celesta
10	Glockenspiel
11	Music Box
12	Vibraphone
13	Marimba
14	Xylophone
15	Tubular Bells
16	Dulcimer
ORGAN	
17	Drawbar Organ
18	Percussive Organ
19	Rock Organ
20	Church Organ
21	Reed Organ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

22	Accoridan
23	Harmonica
24	Tango Accordian
GUITAR	
25	Nylon String Guitar
26	Steel String Guitar
27	Electric Jazz Guitar
28	Electric Clean Guitar
29	Electric Muted Guitar
30	Overdriven Guitar
31	Distortion Guitar
32	Guitar Harmonics
BASS	
33	Acoustic Bass
34	Electric Bass(finger)
35	Electric Bass(pick)
36	Fretless Bass
37	Slap Bass 1
38	Slap Bass 2
39	Synth Bass 1
40	Synth Bass 2
SOLO STRINGS	
41	Violin
42	Viola
43	Cello
44	Contrabass
45	Tremolo Strings
46	Pizzicato Strings
47	Orchestral Strings
48	Timpani
ENSEMBLE	
49	String Ensemble 1
50	String Ensemble 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

51	SynthStrings 1
52	SynthStrings 2
53	Choir Aahs
54	Voice Oohs
55	Synth Voice
56	Orchestra Hit
BRASS	
57	Trumpet
58	Trombone
59	Tuba
60	Muted Trumpet
61	French Horn
62	Brass Section
63	SynthBrass 1
64	SynthBrass 2
REED	
65	Soprano Sax
66	Alto Sax
67	Tenor Sax
68	Baritone Sax
69	Oboe
70	English Horn
71	Bassoon
72	Clarinet
PIPE	
73	Piccolo
74	Flute
75	Recorder
76	Pan Flute
77	Blown Bottle
78	Skakuhachi
79	Whistle
80	Ocarina

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

SYNTH LEAD	
81	Lead 1 (square)
82	Lead 2 (sawtooth)
83	Lead 3 (calliope)
84	Lead 4 (chiff)
85	Lead 5 (charang)
86	Lead 6 (voice)
87	Lead 7 (fifths)
88	Lead 8 (bass+lead)
SYNTH PAD	
89	Pad 1 (new age)
90	Pad 2 (warm)
91	Pad 3 (polysynth)
92	Pad 4 (choir)
93	Pad 5 (bowed)
94	Pad 6 (metallic)
95	Pad 7 (halo)
96	Pad 8 (sweep)
SYNTH EFFECTS	
97	FX 1 (rain)
98	FX 2 (soundtrack)
99	FX 3 (crystal)
100	FX 4 (atmosphere)
101	FX 5 (brightness)
102	FX 6 (goblins)
103	FX 7 (echoes)
104	FX 8 (sci-fi)
ETHNIC	
105	Sitar
106	Banjo
107	Shamisen
108	Koto
109	Kalimba

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

110	Bagpipe
111	Fiddle
112	Shanai
PERCUSSIVE	
113	Tinkle Bell
114	Agogo
115	Steel Drums
116	Woodblock
117	Taiko Drum
118	Melodic Tom
119	Synth Drum
120	Reverse Cymbal
SOUND EFFECTS	
121	Guitar Fret Noise
122	Breath Noise
123	Seashore
124	Bird Tweet
125	Telephone Ring
126	Helicopter
127	Applause
128	Gunshot

ตารางที่ ข-1 แสดงหมายเลขของเครื่องดนตรีแต่ละชนิด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] Rafael C. Gonzalez, Richard E. Woods: *"Digital Image Processing"*, Second Edition , Prentice Hall 2002.

[2] Gregory A. Baxes: *"Digital Image Processing"*, Principles and Applications, Wiley 1994.

[3] J. R. Parker: "*Algorithms For Image Processing And Computer Vision*", Wiley 1997.

[4] George Hansper. "An Introduction to MIDI."[Online].
Available : http://www.crystal.apana.org.au/ghansper/midi_introduction/contents.html 1998.

[5] Mark Clay. "MIDI File Format." [Online].
Available : http://ourworld.compuserve.com/homepages/mark_clay/midi.htm. 1996.

[6] Harmony Central ,Inc. All rights reserved. "MIDI Documentation." [Online].
Available : http://www.hamony-central.com/MIDI/Doc/doc.html. 1995-97

[7] สมนึก อุ่นแก้ว: "*ทฤษฎีดนตรี แนวปฏิบัติ*", พิมพ์ครั้งที่ 2, Musico 2537.

[8] เล็ก วงศ์สว่าง: "*Picking*", พิมพ์ครั้งที่ 2, วงศ์สว่างการพิมพ์ 2523.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้