การทำให้ท่าทางการเคลื่อนไหวและเสียงเพลงสอดคล้องกันโดยใช้โมชันกราฟ

นายอรรถวุฒิ หลายชูไทย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2553
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

SYNCHRONIZATION BETWEEN MOTION AND MUSIC USING MOTION GRAPH

Mr. Auttawut Laichuthai

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Computer Engineering

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2010

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การทำให้ท่าทางการเคลื่อนไหวและเสียงเพลงสอดคล้องกันโดยใช้โมชันกราฟ
โดย	นายอรรถวุฒิ หลายชูไทย
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร. พิษณุ คนองชัยยศ

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร.บุญสม เลิศหิรัญวงศ์)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.อดิวงศ์ สุชาโต)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร. พิษณุ คนองชัยยศ)

.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.โชติรัตน์ รัตนามหัทธนะ)

.....กรรมการภายนอกมหาวิทยาลัย
(ดร.ชาคริต วัชโรภาส)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

อรรถวุฒิ หลายชูไทย : การทำให้ท่าทางการเคลื่อนไหวและเสียงเพลงสอดคล้องกันโดยใช้โมชันกราฟ. (SYNCHRONIZATION BETWEEN MOTION AND MUSIC USING MOTION GRAPH) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ. ดร.พิษณุ คนองชัยยศ, 36หน้า.

การสังเคราะห์ท่าทางการเคลื่อนไหวใหม่ขึ้นมาโดยใช้ฐานข้อมูลการเคลื่อนไหวเดิมที่มีอยู่แล้วนั้นโดยที่จะไม่ต้องไปจับภาพการเคลื่อนไหวใหม่อีกครั้งหนึ่ง วิธีการนำเอาข้อมูลที่เรามีอยู่แล้วมาใช้ใหม่อีกครั้งนั้นประกอบไปด้วยหลายวิธี อาทิเช่น วิธีปรับแต่งข้อมูลของการเคลื่อนไหว วิธีการนำเอาข้อมูลการการเคลื่อนไหวมาประยุกต์ใช้กับตัวละครประเภทต่างๆ วิธีการสังเคราะห์ท่าทางการเคลื่อนไหวใหม่จากฐานข้อมูลเดิม ซึ่งการนำเอาข้อมูลเดิมมาใช้ใหม่นั้นมีข้อดีอยู่หลายข้อ โดยด้านที่สนใจจะเป็นการนำเอาวิธีการสังเคราะห์ท่าทางการเคลื่อนไหวใหม่จากฐานข้อมูลเดิม มาใช้เพื่อสร้างท่าทางการเคลื่อนไหวของตัวละครและนำไปผสมผสานระหว่างการวิเคราะห์เสียงเพลงแล้วนำทั้งสองอย่างมาเชื่อมเข้าด้วยกันเพื่อสร้างเป็นรูปแบบการเคลื่อนไหวที่ประกอบกับเสียงดนตรีขึ้นมา โดยอาจจะมีการเพิ่มคุณลักษณะพิเศษอย่างอื่นเข้าไป เช่น สามารถกำหนดท่าทางการเคลื่อนไหวในเพลงเองได้

งานวิจัยฉบับนี้จึงจะนำเสนอวิธีการที่จะสร้างรูปแบบการเคลื่อนไหวชุดใหม่ขึ้นมาจากฐานข้อมูลของการเคลื่อนไหวเดิมโดยจะใช้วิธี การสังเคราะห์การเคลื่อนไหว และเมื่อมีการใส่ดนตรีเข้ามาระบบก็จะทำการปรับแต่งทั้งสองส่วนให้สอดคล้องกัน โดยข้อดีของงานวิจัยชิ้นนี้คือระบบจะทำหน้าที่ในการทำให้ทั้งสองส่วนสอดคล้องกันโดยอัตโนมัติ และจากผลการทดลองแสดงให้เห็นว่าวิธีการที่นำเสนอสามารถปรับแต่งการเคลื่อนไหวให้มีการแสดงออกที่สอดคล้องกับเสียงดนตรี

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา...วิศวกรรมคอมพิวเตอร์.....ลายมือชื่อนิสิต.....อรรถวุฒิ หลายชูไทย.....
สาขาวิชา...วิศวกรรมคอมพิวเตอร์..... ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....
ปีการศึกษา.....2553.....

5270571021 : MAJOR COMPUTER ENGINEERING

KEYWORDS : MOTION SYNTHESIS , BEAT TRACKING

AUTTAWUT LAICHUTHAI : SYNCHRONIZATION BETWEEN MOTION AND MUSIC USING MOTION GRAPH. ADVISOR : ASST.PROF. PIZZANU KANONGCHAIYOS, Ph.D., 36 pp.

To synthesis new motion from an existing motion data have three ways, First is motion editing, Second is Motion retargeting and Motion synthesis is the third. We are focusing in motion synthesis and use this method to synthesized new motion to synchronize with an input music

Synchronization between motion and music is an interesting area in computer graphics however this part of work is difficult for generating automatically therefore it is usually done by an animators. We present a scheme for synchronizing music with synthesized motion by mapping both feature together. Feature points of music file and motion data are extracted. Then, we pair and synchronize. We also use motion graph, which is a well-known method for novice user, to generate long motions and fully automatic process of motion synthesis. We synthesize a new motion which matches to the music the most.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Department : Computer Engineering Student's Signature อรรฏวุฒิ ไลชุธัย

Field of Study : Computer Engineering Advisor's Signature

Academic Year : 2010

กิตติกรรมประกาศ

งานวิจัยนี้ (AS585A) ได้รับการสนับสนุนบางส่วนจาก โครงการส่งเสริมการวิจัยในอุดมศึกษาและการพัฒนามหาวิทยาลัยวิจัยแห่งชาติ ของสำนักงานคณะกรรมการการอุดมศึกษา

ขอบพระคุณ ผศ.ดร. พิษณุ คนองชัยยศ ที่ปรึกษาของวิทยานิพนธ์นี้ ขอบคุณอาจารย์ที่ช่วยชี้แนะ สั่งสอน ตลอดมา

ขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ฉบับนี้ ที่ให้ข้อคิดและข้อเสนอแนะต่าง ๆ อันเป็นประโยชน์อย่างยิ่งในการพัฒนาคุณภาพของวิทยานิพนธ์ฉบับนี้

ขอขอบคุณเพื่อน ๆ พี่ ๆ ที่ทั้งในห้องวิจัยคอมพิวเตอร์กราฟิก และเพื่อน ๆ ทุกคนที่แบ่งปันความรู้ ช่วยเหลือกัน และเป็นกำลังใจให้กันเสมอมา

สุดท้ายนี้ขอขอบคุณมารดาและบิดาของข้าพเจ้าที่เลี้ยงดู สนับสนุน ทำให้ข้าพเจ้ามีโอกาสได้เข้ามาศึกษาที่นี้จนสามารถสร้างสรรค์วิทยานิพนธ์ฉบับนี้ขึ้นมาได้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ฉ

สารบัญ ..... ช

สารบัญตาราง ..... ญ

สารบัญภาพ ..... ฏ

บทที่ 1 บทนำ 1

1.1 ที่มาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์ของการวิจัย 3

1.3 ขอบเขตของการวิจัย 3

1.4 ประโยชน์ที่ได้รับ 3

1.5 วิธีดำเนินการวิจัย 4

1.6 ผลงานตีพิมพ์จากงานวิจัย 5

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 6

2.1 ทฤษฎีที่เกี่ยวข้อง 6

2.1.1 การสังเคราะห์การเคลื่อนไหว 6

2.1.2 พิกัดออยเลอร์และพิกัดควอเตอร์เนียน 7

2.1.2.1 พิกัดออยเลอร์ 7

2.1.2.2 พิกัดควอเตอร์เนียน 8

2.1.3 การวัดความเหมือนของการเคลื่อนไหวด้วยความเหมือนของโคซายน์ 10

2.1.4 การประมาณในช่วงเชิงเส้นรอบวงกลม 10

2.1.5 การประมาณค่าในช่วงเชิงเส้น 11

2.1.6 การวิเคราะห์หาจังหวะของเสียงดนตรี 12

หน้า

2.1.6.1 การตรวจจับองค์ประกอบของเสียงดนตรี 12
2.1.6.2 การประมาณหาจังหวะเริ่มต้นและช่วงเวลาของจังหวะ 14
2.1.6.3 การวิเคราะห์หาจังหวะของเสียงดนตรี 15
2.2 งานวิจัยที่เกี่ยวข้อง 15
2.2.1 การสังเคราะห์การเคลื่อนไหวและการหาลักษณะเด่น 16
2.2.2 การวิเคราะห์หาจังหวะของเสียงดนตรี 17
2.2.3 การทำให้การเคลื่อนไหวและเสียงดนตรีเคลื่อนไหวแบบสอดคล้องกัน 17
บทที่ 3 ขั้นตอนการดำเนินงานวิจัย 18
3.1 ภาพรวมของระบบ 18
3.2 การวิเคราะห์หาจังหวะของเสียงดนตรี 18
3.2.1 การคำนวณหาตำแหน่งของจังหวะดนตรี 18
3.3 การสังเคราะห์การเคลื่อนไหวโดยใช้โมชันกราฟ 20
3.3.1 การนำเสนอข้อมูลการเคลื่อนไหว 20
3.3.2 การแปลงพิกัดของข้อมูลการเคลื่อนไหวเพื่อนำไปวิเคราะห์ 21
3.3.3 การหาตำแหน่งที่จะเป็นการเชื่อมต่อระหว่างคลิป 22
3.3.4 การเลือกตำแหน่งที่สามารถเปลี่ยนแปลงได้ 22
3.3.5 การกำจัดตำแหน่งที่ไม่สามารถนำมาเป็นตำแหน่งของการเชื่อมต่อ 23
3.3.6 การสร้างการเคลื่อนไหวเพื่อเชื่อมต่อระหว่างคลิปของการเคลื่อนไหว 24
3.3.7 การแปลงพิกัดของข้อมูลการเคลื่อนไหวเพื่อนำไปใช้ประกอบเสียงดนตรี 24
3.3.8 การวิเคราะห์หาลักษณะเด่นของการเคลื่อนไหว 25
3.4 การจับคู่ระหว่างลักษณะเด่นของการเคลื่อนไหวและเสียงดนตรี 26
บทที่ 4 การทดลองและวิเคราะห์ผล 29
4.2 ผลของการสังเคราะห์การเคลื่อนไหว 29
4.3 การทดสอบความถูกต้องของการวิเคราะห์เสียงดนตรี 31
4.3.1 วิเคราะห์โดยใช้ Beat Detection Error Rate 31

หน้า

4.4 ผลการทดสอบความถูกต้องของของการวิเคราะห์หาจังหวะของเสียงดนตรี 32
4.5 การทดสอบความถูกต้องของการเคลื่อนไหวประกอบเสียงดนตรี 32
4.5.1 การเปรียบเทียบเวลาของจังหวะและท่าทางการเคลื่อนไหว 32
4.6 อภิปรายผลการทดลอง (Discussion) 32
บทที่ 5 สรุปผลการวิจัย และข้อเสนอแนะ 34
5.1 สรุปผลการวิจัย (Conclusion) 34
5.2 ข้อเสนอแนะ (Future Work) 34
รายการอ้างอิง 34
ประวัติผู้เขียนวิทยานิพนธ์ 36

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

หน้า

ตารางที่ 4.1 เกณฑ์การวัดความผิดพลาดทั้งสองชนิด 32
ตารางที่ 4.2 ตารางเปรียบเทียบเวลาของจังหวะและท่าทางการเคลื่อนไหว 32

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

หน้า

รูปที่ 1.1 แสดงการเต้นประกอบจังหวะ 1
รูปที่ 2.1 รูปแบบของโมชันกราฟ 6
รูปที่ 2.2 พิกัดออยเลอร์ 7
รูปที่ 2.3 พิกัดควอเตอร์เนียน 8
รูปที่ 2.4 Cosine Angle (θ) 10
รูปที่ 2.5 การประมาณในช่วงเชิงเส้นรอบวงกลม 11
รูปที่ 2.6 การประมาณค่าในช่วงเชิงเส้น 11
รูปที่ 2.7 Onset Time ณ ช่วงของความถี่ต่างๆ 13
รูปที่ 2.8 การประมาณหาช่วงเวลาของจังหวะ 14
รูปที่ 2.9 การวิเคราะห์หาจังหวะของเสียงดนตรี 15
รูปที่ 3.1 ภาพรวมของขั้นตอนการทำงานวิจัย 18
รูปที่ 3.2 ตำแหน่งของจังหวะดนตรี (จุดสีแดง) 19
รูปที่ 3.3 ตำแหน่งของจังหวะดนตรีในรูปของฟังก์ชันที่เป็นลักษณะเด่น 20
รูปที่ 3.4 ข้อมูลที่ถูกเก็บอยู่ในไฟล์ เอเอสอฟ 21
รูปที่ 3.5 กลุ่มจุดที่แสดงตามค่าของมุมโคซายน์ 22
รูปที่ 3.6 แสดงตำแหน่งของจุดเปลี่ยนแปลง 23
รูปที่ 3.7 โมชันกราฟก่อนกำจัดจุดที่เป็นไปไม่ได้ 23
รูปที่ 3.8 ลักษณะเด่นของการเคลื่อนไหวของข้อต่อ 25
รูปที่ 3.9 กราฟที่แสดงลักษณะเด่นของข้อต่อที่เราสนใจ 26
รูปที่ 3.10 ลักษณะเด่นของการเคลื่อนไหวและจังหวะของเสียงดนตรี 27
รูปที่ 3.11 การจับคู่ระหว่างการเคลื่อนไหวและเสียงดนตรี 27
รูปที่ 3.12 ผลลัพธ์ของการจับคู่ระหว่างการเคลื่อนไหวและเสียงเพลง 28
รูปที่ 4.1 ตำแหน่งที่สามารถเชื่อมต่อกันได้ เมื่อ X = 50 30
รูปที่ 4.2 ตำแหน่งที่สามารถเชื่อมต่อกันได้ เมื่อ X = 90 31

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

รูปที่ 1.1 แสดงการเต้นประกอบจังหวะ

ในปัจจุบันคอมพิวเตอร์กราฟิกได้มีบทบาทสำคัญมากขึ้น สังเกตได้จากงานด้านต่าง ๆ ได้มีการนำคอมพิวเตอร์กราฟิกเข้าไปใช้ในการแสดงผลมากขึ้น ดังเช่น งานทางด้านโฆษณา การทำแอนิเมชัน การทำหนัง การทำเกมส์ และยังมีอีกมากมายที่ไม่ได้กล่าวถึงแต่ก็ได้มีการนำเอางานทางด้านคอมพิวเตอร์กราฟิกไปใช้ไม่มากก็น้อย ซึ่งวิธีการหนึ่งที่มีความสำคัญในอุตสาหกรรมภาพยนตร์และแอนิเมชันคือวิธีการที่เรียกว่า โมชันแคปเจอร์ (Motion Capture) ซึ่งเป็นวิธีที่ได้รับความนิยมเป็นอย่างมากในขณะนี้ เนื่องจากผลที่ได้จากการที่ใช้วิธีนี้จะทำให้สามารถจับภาพการเคลื่อนไหวของมนุษย์ได้อย่างเหมือนจริงและมีรายละเอียดมากที่สุด โดยการจะทำการจับภาพการเคลื่อนไหวเราจะใช้อุปกรณ์กล้องชนิดพิเศษรวมกับตัวส่งสัญญาณ โดยข้อมูลที่เราได้มาจะอยู่ในรูปของตำแหน่งของการเคลื่อนไหวเช่น การหมุน การเหวี่ยง ของข้อต่อต่าง ๆ ตามร่างกายของคนที่เราจับภาพ แล้วนำค่าที่ได้มาจำลองการเคลื่อนไหวของตัวละครที่เราสร้างขึ้นมา แต่ว่าวิธีการดังกล่าวก็ยังมีข้อจำกัดอยู่ก็คือ จำเป็นจะต้องมีผู้ใช้เป็นผู้กำหนดหรือควบคุมในการที่จะให้ตัวละครที่เราสร้างขึ้นมาสามารถเชื่อมต่อกับ

ภาพที่เราจับมาได้โดยตัวระบบเองยังไม่ได้เป็นแบบอัตโนมัติ และข้อจำกัดอีกข้อหนึ่งที่สำคัญก็คือข้อมูลของการเคลื่อนไหวที่เราได้มานั้นไม่สามารถดัดแปลงแก้ไขให้เป็นท่าทางการเคลื่อนไหวแบบอื่น ๆ ได้ หมายความว่าถ้าเราได้จับท่าทางการเดินของคนแบบเดินตรงและคนที่วิ่งมา เราก็ไม่สามารถนำไปแปลงเป็นท่าการเคลื่อนไหวแบบที่มีการแสดงออกทั้งการเดินและวิ่งแบบต่อเนื่องกันได้

จากข้อจำกัดดังกล่าวทำให้มีงานวิจัยที่เสนอการสังเคราะห์ท่าทางการเคลื่อนไหวใหม่ขึ้นมาโดยใช้ฐานข้อมูลการเคลื่อนไหวเดิมที่เรามีอยู่แล้วเพื่อที่จะได้ไม่ต้องไปจับภาพการเคลื่อนไหวใหม่อีกครั้งหนึ่ง วิธีการจะนำเอาข้อมูลที่เรามีอยู่แล้วมาใช้ใหม่อีกครั้งนั้นประกอบไปด้วยหลายวิธี อาทิเช่น วิธีปรับแต่งข้อมูลของการเคลื่อนไหว วิธีการนำเอาข้อมูลการการเคลื่อนไหวมาประยุกต์ใช้กับตัวละครประเภทต่างๆ วิธีการสังเคราะห์ท่าทางการเคลื่อนไหวใหม่จากฐานข้อมูลเดิม ซึ่งการนำเอาข้อมูลเดิมมาใช้ใหม่นั้นมีข้อดีอยู่หลายข้อ เช่น ช่วยประหยัดเวลาและค่าใช้จ่ายในการจับภาพใหม่ สามารถสร้างท่าทางการเคลื่อนไหวที่ซับซ้อนซึ่งถ้าเป็นมนุษย์ไม่สามารถทำได้ สามารถนำไปใช้กับงานประเภทการจำลองการเคลื่อนไหวของกลุ่มคนจำนวนมากได้ โดยด้านที่สนใจจะเป็นการนำเอาวิธีการสังเคราะห์ท่าทางการเคลื่อนไหวใหม่จากฐานข้อมูลเดิม มาใช้เพื่อสร้างท่าทางการเคลื่อนไหวของตัวละครและนำไปผสมผสานระหว่างการวิเคราะห์เสียงเพลงแล้วนำทั้งสองอย่างมาเชื่อมเข้าด้วยกันเพื่อสร้างเป็นรูปแบบการเคลื่อนไหวที่ประกอบกับเสียงดนตรีขึ้นมา โดยอาจจะมีการเพิ่มคุณลักษณะพิเศษอย่างอื่นเข้าไปเช่น สามารถกำหนดท่าทางการเคลื่อนไหวในเพลงเองได้ สาเหตุที่สนใจงานวิจัยทางด้านนี้เพราะว่าเนื่องจากในปัจจุบันไม่ว่าจะในหนัง แอนิเมชัน หรือแม้กระทั่งเกม ทุก ๆ การเคลื่อนไหวของตัวละครจะมีเสียงดนตรีประกอบไปด้วย แต่ขั้นตอนในการจะทำแบบนั้นจะเป็นกระบวนการที่แยกส่วนกัน กล่าวคือ จับภาพการเคลื่อนไหวก่อนแล้วค่อยนำดนตรีมาประกอบเข้าไปด้วยกัน ทำให้บางครั้งท่าทางการเคลื่อนไหวไม่สอดคล้องกับจังหวะของเสียงดนตรีในขณะนั้น ซึ่งในส่วนนี้ที่ทำให้ต้องมี ผู้ใช้เข้ามามีส่วนช่วยในการปรับแต่งให้ทั้งสองส่วนสามารถแสดงออกมารูปแบบที่สอดคล้องกัน

งานวิจัยฉบับนี้จึงจะนำเสนอวิธีการที่จะสร้างรูปแบบการเคลื่อนไหวชุดใหม่ขึ้นมาจากฐานข้อมูลของการเคลื่อนไหวเดิมโดยจะใช้วิธีการสังเคราะห์ท่าทาง(Motion Synthesis) และเมื่อมีการใส่ดนตรีเข้ามา ระบบก็จะทำการปรับแต่งทั้งสองส่วนให้สอดคล้องกัน โดยข้อดีของงานวิจัยชิ้นนี้คือระบบจะทำหน้าที่ในการทำให้ทั้งสองส่วนสอดคล้องกันโดยอัตโนมัติ จากที่ผ่าน ๆ มาจะต้องให้ผู้ใช้เข้ามามีส่วนในการปรับแต่ง และผู้ใช้อาจจะมีส่วนร่วมในการเลือกท่าที่ต้องการให้มีในเพลงได้

1.2 วัตถุประสงค์ของการวิจัย

เพื่อที่จะเสนอวิธีการที่จะนำเอาท่าทางการเคลื่อนไหวและเสียงดนตรีมาเชื่อมต่อกันแบบอัตโนมัติโดยไม่ต้องพึ่งพาผู้ใช้ในการปรับอีก ที่ต้องมีการปรับให้ท่าทางการเคลื่อนไหวสอดคล้องไปกับเสียงเพลงนั้นก็เพราะว่าจะได้เพิ่มความเหมือนจริงให้กับภาพที่แสดงออกมา และยังสามารถที่จะนำไปใช้ในการทำเกมและฉากแอนิเมชันต่างๆ เพราะว่าในปัจจุบันการทำเกมหรือแอนิเมชัน ทุก ๆ การเคลื่อนไหวจะมีเสียงดนตรีประกอบอยู่ด้วยเสมอ

1.3 ขอบเขตของการวิจัย

1 รูปแบบข้อมูลที่นำเข้าของเสียงเพลงจะเป็นไฟล์แบบ MIDI

2 รูปแบบข้อมูลการเคลื่อนไหวที่นำมาใช้ในงานจะเป็นไฟล์ ASF/AMC โดยจะเก็บเป็นข้อมูลการเคลื่อนไหวของข้อต่อต่าง ๆ ในร่างกาย

3 การแสดงผลจะอยู่ในรูปแบบของสามมิติ ที่เป็นการแสดงท่าเต้นประกอบกับเสียงดนตรี

4 โปรแกรมที่ใช้สำหรับการทำการทดลองคือโปรแกรม MATLAB ใช้สำหรับวิเคราะห์ข้อมูลเพื่อนำไปใช้ในการสังเคราะห์ข้อมูลการเคลื่อนไหว และในส่วนของเสียงดนตรีจะใช้สำหรับวิเคราะห์หาจังหวะของเสียงดนตรี และอีกโปรแกรมคือ โปรแกรม MAYA ใช้สำหรับแสดงผล โดยทั้งสองโปรแกรมจะใช้งานบนระบบปฏิบัติการวินโดว์

1.4 ประโยชน์ที่ได้รับ

สามารถลดเวลาในการทำให้ภาพและเสียงมีการแสดงออกที่สอดคล้องกันในการทำหนังแอนิเมชัน และสามารถสร้างท่าทางการเคลื่อนไหวใหม่ให้สอดคล้องกับเพลงได้โดยไม่ต้องไปเก็บข้อมูลแสดงภาพการเคลื่อนไหวใหม่

1.5 วิธีดำเนินการวิจัย

	เดือนเริ่มต้น	ระยะเวลา (เดือน)	มิ.ย.-52	ก.ค.-52	ส.ค.-52	ก.ย.-52	ต.ค.-52	พ.ย.-52	ธ.ค.-52	ม.ค.-53	ก.พ.-53	มี.ค.-53	เม.ย.-53	พ.ค.-53	มิ.ย.-53	ก.ค.-53	ส.ค.-53	ก.ย.-53	ต.ค.-53	พ.ย.-53	ธ.ค.-53	ม.ค.-54	ก.พ.-54	มี.ค.-54
ศึกษางานวิจัย และทฤษฎีที่เกี่ยวข้อง	ก.ค.-52	6																						
วิเคราะห์ และออกแบบอัลกอริทึม	ก.ย.-52	7																						
ออกแบบโปรแกรมประยุกต์ที่ใช้ทดสอบ	ม.ค.-52	5																						
พัฒนา และทดสอบโปรแกรมประยุกต์	พ.ค.-53	8																						
ทดลองและประเมินผล	ต.ค.-53	5																						
สรุปผลการดำเนินงาน	ม.ค.-54	2																						
จัดทำวิทยานิพนธ์	ก.พ.-54	2																						

รูปที่ 1.2 แผนภาพแกนท์แสดงระยะเวลาการดำเนินการวิจัย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

1.6 ผลงานตีพิมพ์จากงานวิจัย

งานวิจัยนี้ได้รับการตีพิมพ์เป็นบทความทางวิชาการดังนี้

- “SYNCHRONIZATION BETWEEN MOTION AND MUSIC USING MOTION GRAPH” โดย อรรถวุฒิ หลายชูไทย และ พิษณุ คนองชัยยศ ในงานประชุมวิชาการ “ THE EIGHTH ANNUAL INTERNATIONAL CONFERENCE ORGANIZED BY ELECTRICAL ENGINEERING/ ELECTRONICS, COMPUTER TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY ”: ซึ่งจัดขึ้นที่ ขอนแก่น ประเทศไทย ในระหว่างวันที่ 17-19 พฤษภาคม 2554

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

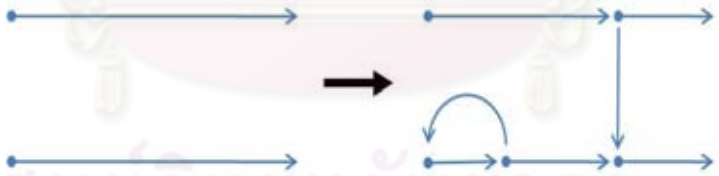
บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 การสังเคราะห์การเคลื่อนไหว

การนำเอาข้อมูลการเคลื่อนไหวที่มีอยู่แล้วกลับมาใช้ใหม่นั้นสามารถแบ่งออกเป็น 3 วิธีใหญ่ๆ คือ การปรับแต่งข้อมูลเดิม (Motion Editing) การนำเอาไปใช้ในงานประเภทอื่น (Motion Retargeting) และการสังเคราะห์การเคลื่อนไหว (Motion Synthesis) โดยในงานวิจัยนี้ได้ใช้วิธีการสังเคราะห์การเคลื่อนไหว เนื่องจากวิธีการอื่นๆที่ได้นำเสนอมานั้นไม่เหมาะกับการนำมาใช้ในงานวิจัยนี้ สาเหตุเพราะว่าทั้งวิธีการปรับแต่งข้อมูลเดิมและการนำเอาไปใช้ในงานประเภทอื่นนั้นไม่สามารถปรับแต่งระยะเวลา หรือว่านำเอาข้อมูลการเคลื่อนไหวสองชุดมาทำให้แสดงออกแบบต่อเนื่องกันได้ วิธีการสังเคราะห์การเคลื่อนไหวที่เราใช้มีชื่อเรียกว่า “โมชันกราฟ (Motion Graph)”

โมชันกราฟ (Motion Graph) คือ กราฟเส้นตรงที่แสดงถึงกลุ่มของท่าทางการเคลื่อนไหว โดยวิธีการของโมชันกราฟจะเป็นการนำเอาข้อมูลท่าทางการเคลื่อนไหวที่เรามีในฐานข้อมูล มาวิเคราะห์เพื่อสร้างรูปแบบการเคลื่อนไหวแบบใหม่ ที่ไม่เคยมีจากในฐานข้อมูล



รูปที่ 2.1 รูปแบบของโมชันกราฟ

ในตัวอย่างของกราฟที่แสดงให้เห็นในรูปที่ 2.1 เส้นแต่ละเส้นที่อยู่ระหว่างจุดจะแสดงถึงการเคลื่อนไหว 1 ครั้ง และจุดจะแสดงถึงเป็นตัวเชื่อมระหว่างการเคลื่อนไหวแต่ละครั้ง โดยเราสามารถเพิ่มจุดเข้าไประหว่างเส้นได้ ถ้าเส้นการเคลื่อนที่นั้นสามารถแตกออกเป็น 2 การเคลื่อนที่ได้ หรือแม้กระทั่งจะเป็นการกระโดดข้ามเส้นของการเคลื่อนที่ก็ได้ ถ้าท่าทางการเคลื่อนที่ของเส้นเดิมสามารถไปเชื่อมต่อกับท่าทางการเคลื่อนที่ของเส้นใหม่ได้ แต่อย่างไรก็ตามการที่เราจะทำการเชื่อมต่อการเคลื่อนไหว 2 ท่าทางเข้าด้วยกันก็ทำได้ยาก เช่น การเปลี่ยนแปลงท่าทางจากการวิ่งไปเป็นท่ากระโดดกลับตัวไปข้างหลัง ในโลกของความเป็นจริงก็ยังต้องการเวลาเป็น วินาทีจึงจะสามารถกระทำได้อย่างที่กล่าวมา ทำให้ในการที่จะเชื่อมท่าทาง

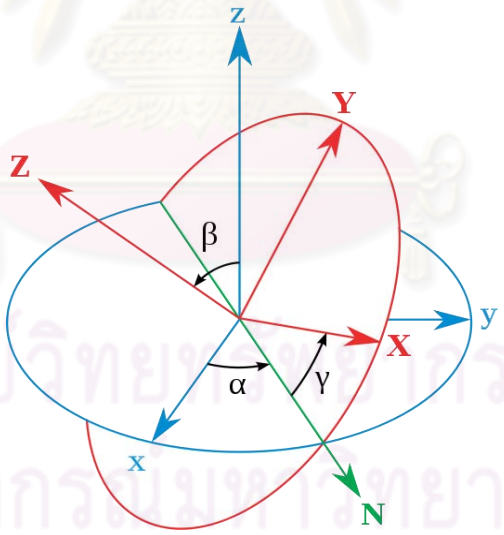
ที่มีการเปลี่ยนแปลงมาก ๆ เข้าด้วยกันเป็นไปได้ยากและผลที่ออกมาก็จะได้ไม่เหมือนจริง แต่อย่างไรก็ตามการที่จะสร้างการเคลื่อนไหวที่เหมือนจริงมาก ๆ กับการเชื่อมการเคลื่อนไหว 2 การเคลื่อนไหวเข้าด้วยกันก็เป็นงานที่ยากพอ ๆ กัน

การสร้างโมชันกราฟนั้นจะเราจะทำการวิเคราะห์ข้อมูลการเคลื่อนไหวที่เรามีอยู่ในฐานข้อมูลและสร้างโมชันกราฟขึ้นมาตั้งแต่แรก แล้วหลังจากนั้นก็จะเก็บกราฟที่สร้างมาได้ให้อยู่รูปของไฟล์นามสกุลเอเอ็มซี (.amc)

2.1.2 พิกัดออยเลอร์และพิกัดควอเตอร์เนียน

เนื่องจากในข้อมูลการเคลื่อนไหวที่จะนำมาใช้วิเคราะห์ในงานวิจัยนี้ จะถูกเก็บอยู่ในรูปแบบของพิกัดออยเลอร์และพิกัดควอเตอร์เนียน ทำให้เราต้องมีการแปลงข้อมูลดังกล่าวให้อยู่ในรูปที่สามารถนำไปใช้การคำนวณได้ง่าย และในบางขั้นตอนของการวิเคราะห์การเคลื่อนไหวอาจจะมีการแปลงพิกัดออยเลอร์ไปพิกัดควอเตอร์เนียน หรือจากพิกัดควอเตอร์เนียนไปยังพิกัดออยเลอร์

2.1.2.1 พิกัดออยเลอร์



รูปที่ 2.2 พิกัดออยเลอร์

มุมออยเลอร์จะประกอบไปด้วยมุมการหมุนตามแนวแกน 3 แกน (X,Y,Z) ดังรูปที่ 2.2 ซึ่งมุมออยเลอร์มีข้อดีคือ สามารถเขียนให้ในอยู่รูปของเวกเตอร์โดยการใช้โครงสร้างข้อมูลในลักษณะเวกเตอร์ได้ ลักษณะการหมุนของมุมออยเลอร์จะหมุนในรูปแบบของ การหมุน

ตัว (Roll), การปักเงย (Pitch) และ การเลี้ยว (Yaw) ซึ่งเป็นการหมุนที่เครื่องบินใช้ในการบังคับการบิน โดยให้เราจินตนาการว่า แกน Z จะอยู่ในแนวลำเครื่องบินจากหัวเครื่องบินไปยังท้ายเครื่องบิน แกน Y จะอยู่ในแนวที่ชี้พุ่งขึ้นมาจากพื้นดิน และแกน X จะอยู่ในแนวของปีกเครื่องบิน เพราะฉะนั้นการหมุนตัว คือการหมุนในแนวแกน Z , การเลี้ยว คือการหมุนในแนวแกน Y และการปักเงย คือการหมุนในแนวแกน X

มุมออยเลอร์มีรูปแบบการเขียนสมการดังนี้

$$\boldsymbol{Euler's\ angle} = [x, y, z]$$

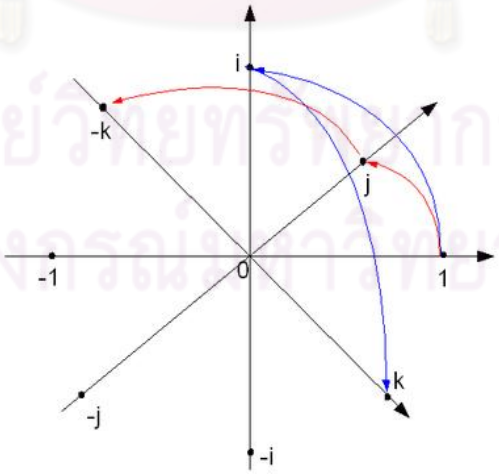
วิธีการแปลงสมการจากมุมควอเตอร์เนียนเป็นมุมออยเลอร์

$$x = \frac{arctan2(q_0 q_1 + q_2 q_3)}{1 - 2(q_1^2 + q_2^2)}$$

$$y = arcsin(2(q_0 q_1 - q_3 q_1))$$

$$z = \frac{arctan2(q_0 q_3 + q_1 q_2)}{1 - 2(q_2^2 + q_3^2)}$$

2.1.2.2 พิกัดควอเตอร์เนียน



รูปที่ 2.3 พิกัดควอเตอร์เนียน

โดยทั่วไปการหมุนวัตถุจะนิยมใช้มุมออยเลอร์เนื่องจากค่าของตัวแปรในการกำหนดเป็นค่าที่มนุษย์สามารถนำไปใช้ได้ง่ายกว่าเพราะเป็นค่ามุมบนระนาบ 3 มิติ และมุมควอเตอร์เนียนก็เป็นทางเลือกของการหมุนเช่นเดียวกันโดยในงานของเราจะใช้การหมุนโดยใช้มุมควอเตอร์เนียนในการคำนวณการหมุนกระดูกของโมเดล ซึ่งมุมควอเตอร์เนียนมีพื้นฐานในการคิดมาจากจำนวนเชิงซ้อน และมีข้อดีคือสามารถสร้างและปรับค่าของการให้มีความลื่นไหลมากกว่าวิธีการใช้การหมุนด้วยการปรับค่าบนมุมออยเลอร์ การใช้มุมควอเตอร์เนียนสามารถประยุกต์นำมาใช้ในการคำนวณเมตริกซ์ขนาด 4x4 ซึ่งเป็นเมตริกซ์ที่นิยมใช้ในการคำนวณการหมุนของวัตถุได้ง่ายกว่า

มุมควอเตอร์เนียนมีรูปแบบการเขียนสมการดังนี้

$$Quarternion's\ angle = [a, b, c, d]$$

สมการแปลงมุมออยเลอร์ ให้เป็นมุมควอเตอร์เนียนโดยกำหนดให้ตัวแปรในมุมออยเลอร์เป็น(X,Y,Z) และตัวแปรในมุมควอเตอร์เนียนเป็น (a,b,c,d) ตามลำดับ จะได้วิธีการแปลงสมการจากมุมออยเลอร์เป็นมุมควอเตอร์เนียนดังนี้

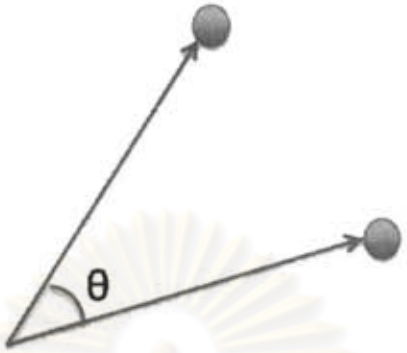
$$a = \left(cos\left(\frac{x}{2}\right) \times cos\left(\frac{y}{2}\right) \times cos\left(\frac{z}{2}\right)\right) + \left(sin\left(\frac{x}{2}\right) \times sin\left(\frac{y}{2}\right) \times sin\left(\frac{z}{2}\right)\right)$$

$$b = \left(sin\left(\frac{x}{2}\right) \times cos\left(\frac{y}{2}\right) \times cos\left(\frac{z}{2}\right)\right) + \left(cos\left(\frac{x}{2}\right) \times sin\left(\frac{y}{2}\right) \times sin\left(\frac{z}{2}\right)\right)$$

$$c = \left(cos\left(\frac{x}{2}\right) \times sin\left(\frac{y}{2}\right) \times cos\left(\frac{z}{2}\right)\right) + \left(sin\left(\frac{x}{2}\right) \times sin\left(\frac{y}{2}\right) \times sin\left(\frac{z}{2}\right)\right)$$

$$d = \left(cos\left(\frac{x}{2}\right) \times cos\left(\frac{y}{2}\right) \times sin\left(\frac{z}{2}\right)\right) + \left(sin\left(\frac{x}{2}\right) \times sin\left(\frac{y}{2}\right) \times cos\left(\frac{z}{2}\right)\right)$$

2.1.3 การวัดความเหมือนของการเคลื่อนไหวด้วยความเหมือนของโคซายน์



รูปที่ 2.4 Cosine Angle (θ)

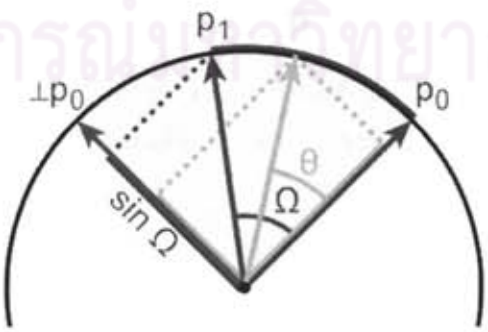
ความเหมือนของมุมโคซายน์เป็นการวัดความเหมือนระหว่างเวกเตอร์สองเวกเตอร์โดยวัดจากการหามุมโคซายน์ระหว่างเวกเตอร์ทั้งสอง ดังรูปที่ 2.4 ผลลัพธ์ของการใช้การวัดแบบความเหมือนของมุมโคซายน์จะเท่ากับ 1 เมื่อมุมเป็น 0 และจะน้อยกว่า 1 เมื่อมุมมีค่าอื่นๆ การคำนวณหามุมโคซายน์ระหว่างสองเวกเตอร์จึงจะเป็นตัวกำหนดว่าเวกเตอร์ทั้งสองมีทิศทางไปทิศเดียวกันหรือไม่

การหามุมโคซายน์ของสองเวกเตอร์จะหาได้จากผลคูณเชิงสเกลาร์ของยูคลิเดียน (Euclidean Dot Product)

$$\boldsymbol{a} \bullet \boldsymbol{b} = \|\boldsymbol{a}\|\|\boldsymbol{b}\| \cos \boldsymbol{\theta}$$

$$\boldsymbol{similarity} = \cos(\boldsymbol{\theta}) = \frac{\boldsymbol{A} \bullet \boldsymbol{B}}{\|\boldsymbol{A}\|\|\boldsymbol{B}\|}$$

2.1.4 การประมาณในช่วงเชิงเส้นรอบวงกลม



รูปที่ 2.5 การประมาณในช่วงเชิงเส้นรอบวงกลม

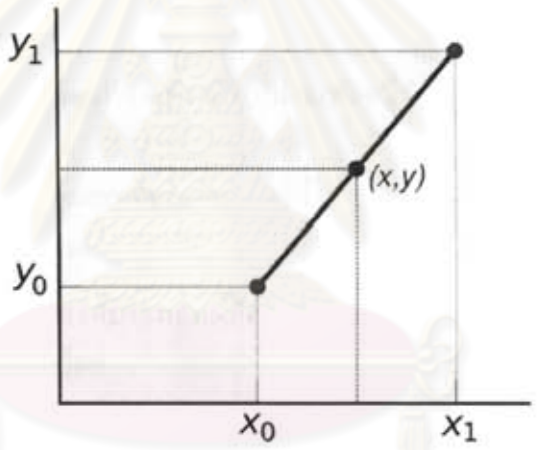
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การประมาณค่าในช่วงเชิงเส้นรอบวงกลม มีคำย่อว่า "SLERP" เป็นวิธีที่ใช้สำหรับประมาณค่าการเคลื่อนไหวที่มีความเร็วคงที่ตามส่วนโค้งของวงกลม โดยที่มีการประมาณค่าพารามิเตอร์ที่กำหนดระหว่าง 0 และ 1

สูตรทางเรขาคณิตของสเลิป (Slerp) จะประกอบไปด้วยการรวมกันเชิงเส้น (Linear Combination) ของปลาย P_0 และ P_1 โดยมี t เป็นพารามิเตอร์โดยที่มีค่าอยู่ระหว่าง 0 ถึง 1 และมี Ω เป็นมุมระหว่างเวกเตอร์ $\cos\Omega = p_0 \bullet p_1$

$$Slerp(p_0, p_1; t) = \frac{sin[(1-t)\Omega]}{sin\,\Omega} p_0 + \frac{sin[t\Omega]}{sin\,\Omega} p_1$$

2.1.5 การประมาณค่าในช่วงเชิงเส้น



รูปที่ 2.6 การประมาณค่าในช่วงเชิงเส้น

การประมาณค่าในช่วงเชิงเส้นเป็นวิธีการประมาณหาค่าส่วนโค้งที่เหมาะสมโดยใช้พหุนามเชิงเส้นในการวิเคราะห์

ถ้าเรารู้พิกัดของจุดสองจุด (x_0, y_0) และ (x_1, y_1) เส้นที่เราประมาณมาจากการใช้วิธีการประมาณค่าในช่วงเชิงเส้นคือแนวเส้นตรงระหว่างจุดสองจุดที่เราสนใจ โดยมีสมการดังนี้

$$\frac{y - y_0}{x - x_0} = \frac{y_1 - y_0}{x_1 - x_0}$$

$$y = y_0 + (x - x_0)\frac{y_1 - y_0}{x_1 - x_0} = \frac{(x - x_0)y_1 + (x_1 - x)y_0}{x_1 - x_0}$$

สูตรนี้ยังสามารถเข้าใจเป็นค่าเฉลี่ยถ่วงน้ำหนัก โดยที่การเฉลี่ยน้ำหนักมีความสัมพันธ์แบบแปรผกผันกับระยะทางจากจุดสิ้นสุดไปยังจุดใดๆ และจุดที่ยิ่งอยู่ใกล้ก็จะยิ่งมีอิทธิพลมากขึ้น

2.1.6 การวิเคราะห์หาจังหวะของเสียงดนตรี

วิธีการวิเคราะห์หาจังหวะของเสียงดนตรีเป็นวิธีการที่ไว้ใช้สำหรับตรวจจับจังหวะของเสียงเพลงสากล โดยการจะใช้วิธีการนี้จะพิจารณาองค์ประกอบในเพลงอยู่ 3 องค์ประกอบ คือ

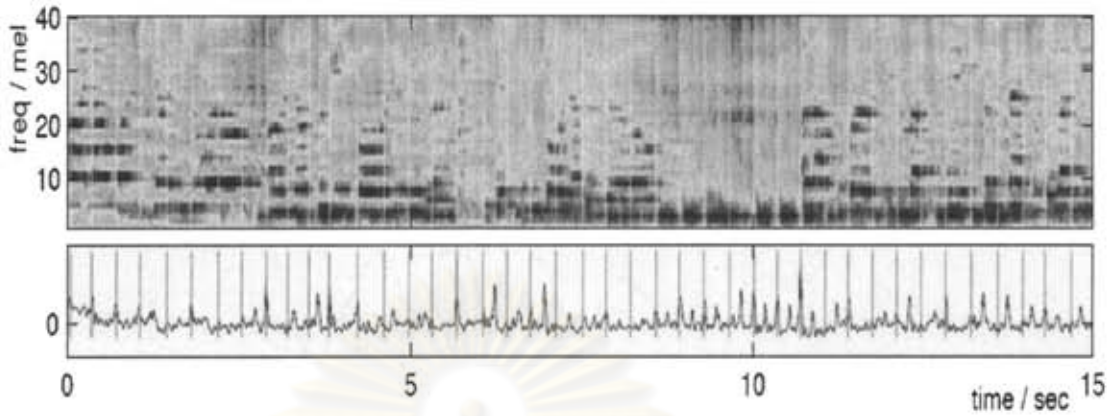
1. Onset Time : เป็นองค์ประกอบที่ดีอันหนึ่งในการใช้สำหรับวิธีการ Beat Tracking
2. Chord Changes : เนื่องจากความถี่ของแต่ละคอร์ดจะไม่เหมือนกัน แต่ว่าถ้าเป็นในเพลงญี่ปุ่นก็จะไม่ค่อยมีการเปลี่ยนคอร์ดมากนัก
3. Percussion Patterns : ใช้สำหรับแยกแยะพวกเสียงกลอง หรือเสียงเบส

เนื่องจากสามองค์ประกอบนี้ก็เพียงพอสำหรับการแยกแยะเพลงสากลทั่วไป อย่างเช่น เพลงป๊อป หรือเพลงร็อค

2.1.6.1 การตรวจจับองค์ประกอบของเสียงดนตรี

องค์ประกอบที่เราจะตรวจจับเรียกว่า "Onset Time" ซึ่งจะหมายถึงเมื่อใดก็ตามที่มีเสียงเกิดขึ้นพลังงานที่มีความสัมพันธ์กับความถี่ของเสียงที่เกิดขึ้นนั้นก็จะสูงตามไปด้วย โดยที่จะนำเสนออยู่ในรูปแบบของเวกเตอร์ที่มีขนาดขึ้นอยู่กับความถี่ในช่วงต่าง ๆ ดังรูปที่ 2.7

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.7 Onset Time ณ ช่วงของความถี่ต่างๆ

ค่าของการเพิ่มขึ้นของพลังงานเมื่อมีการเกิดเสียงหาได้จากสมการนี้

$$d(t,f)=\begin{cases}\max\left(p(t,f),p(t+1,f)\right)-\Pr evPow \\ \quad\left(\min\left(p(t,f),p(t+1,f)\right)\geq \Pr evPow\right), \\ 0 \quad (otherwise),\end{cases}$$

$d(t,f)$ คือค่าของพลังงานที่เพิ่มขึ้นเมื่อมีเสียงเกิดขึ้นที่เวลา t และที่ความถี่ f

$$\boldsymbol{PrevPow = max(p(t,f-1),p(t-1,f),p(t-1,f-1))}$$

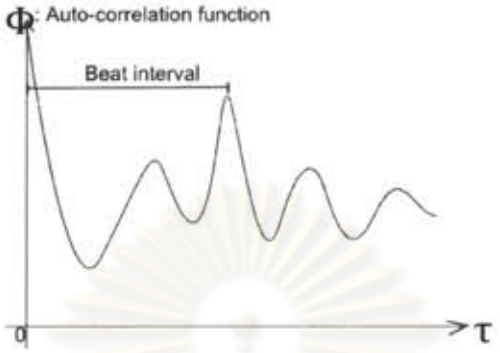
$p(t,f)$ คือค่าของพลังงานที่เวลา t และที่ความถี่ f

โดยค่าของ $PrevPow$ จะพิจารณาพลังงานที่ความถี่ระหว่าง $f-1,\ f, f+1$ ที่เวลา $t-1$ เพราะฉะนั้นความผิดพลาดจากช่วงของ Onset Time เนื่องมาจากเสียงรบกวน หรือว่าความสามารถของผู้เล่น จะไม่นำมาคิด ต่อไปก็จะหาช่วงความถี่โดยผู้ใช้เป็นผู้เลือก และจะคำนวณหา $D(t)$ ได้โดย

$$\boldsymbol{D(t)=\sum_{ChosenFreqBands} d(t,f)}$$

โดยช่วงของความถี่แบ่งออกเป็น 7 ช่วง $(0-125Hz, 125-250Hz, 250-500Hz, 500-1000Hz, 1k-2kHz, 2k-4kHz, \gg 4kHz)$ ซึ่งแต่ละช่วงจะมีความสัมพันธ์กับความสามารถในการฟังที่ประสาทหูของมนุษย์รับรู้ได้

2.1.6.2 การประมาณหาจังหวะเริ่มต้นและช่วงเวลาของจังหวะ



รูปที่ 2.8 การประมาณหาช่วงเวลาของจังหวะ

ในขั้นตอนนี้จะอธิบายถึงวิธีการกะประมาณจุดเริ่มต้นของจังหวะและค่าเฉลี่ยของช่วงจังหวะดังรูปที่ 2.8 โดยขั้นแรกจะคำนวณหาฟังก์ชันความสัมพันธ์แบบอัตโนมัติ $\left(Auto-correlation\ function\left(R_{DD}(\tau)\right)\right)$ ของ $D(t)$ และค่าที่เราคำนวณมาได้จะแสดงถึงการเป็นช่วงของ $D(t)$ และมีวิธีการคำนวณดังนี้

$$R_{DD}(\tau) = \frac{1}{T}\sum_{t=1}^{T} D(t)\, D(t+T)$$

โดยที่ T คือจำนวนของข้อมูล

ถ้าค่าของ $R_{DD}(\tau)$ มีค่าสูงสุดที่ตำแหน่ง τ_{max}, $D(t)$ มีค่าใกล้เคียงกับ $D(t+\tau_{max})$ และ τ_{max} คือค่าเฉลี่ยของช่วงจังหวะ (τ_{rhythm})

หลังจากนั้นจะคำนวณหาฟังก์ชันความสัมพันธ์แบบข้าม $\left(cross-correlation\ function\left(R_{DP}(\tau)\right)\right)$ ของ $D(t)$ และ $P(t)$ ตามสมการดังนี้

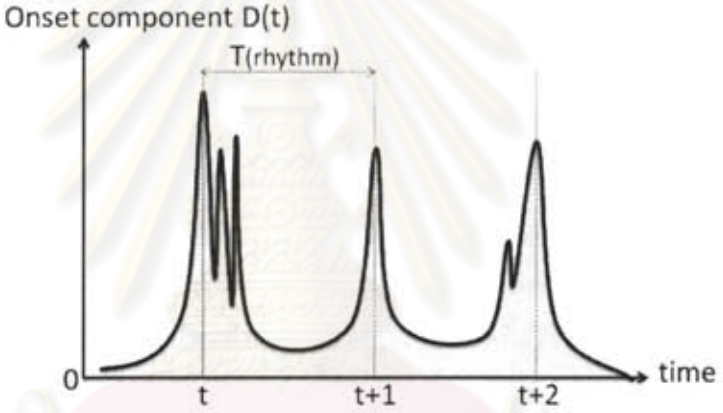
$$R_{DP}(\tau) = \frac{1}{T}\sum_{t=1}^{T} D(t)P(t+\tau)$$

ที่เวลาเมื่อปรากฏค่าของ R_{DP} สูงสุดที่เวลานั้นคือ เวลาเริ่มต้นของจังหวะ $beat\ start(t_{st})$

2.1.6.3 การวิเคราะห์หาจังหวะของเสียงดนตรี

ถึงแม้ว่าจังหวะของเสียงเพลงจะไม่มีการเปลี่ยนแปลงจังหวะหรือบางครั้งอาจจะเปลี่ยนแค่เพียงเล็กน้อยเท่านั้น สาเหตุอาจเนื่องมาจากความสามารถของผู้เล่น และจากการที่เพลงมีการเปลี่ยนแปลงจังหวะแค่เพียงเล็กน้อยทำให้การตรวจจับจังหวะทำได้ยากขึ้น ปัญหานี้สามารถแก้ได้โดยทำตามขั้นตอนต่อไปนี้

1. สำหรับเวลาของจังวะเพลงช่วงหนึ่ง ๆ(t_i)ระบบจะหาจุดที่มี Onset ที่มากที่สุดในรอบ ๆช่วงเวลา $t_i + t_{rhythm}\,(t_0 = t_{st})$ หลังจากนั้นเวลาช่วงต่อไป $t_i + 1$ ก็จะถือว่าเป็นจุดที่มี Onset สูงสุด
2. เมื่อทำกระบวนการแรกเสร็จสิ้น ก็ให้ทำซ้ำไปเรื่อย ๆ จนจบเพลงดังรูปที่ 2.9



รูปที่ 2.9 การวิเคราะห์หาจังหวะของเสียงดนตรี

2.2 งานวิจัยที่เกี่ยวข้อง

ในการทำงานวิจัยทางด้านนี้จะสามารถแบ่งออกเป็น สองวิธีใหญ่ ๆ โดยแบ่งตามการสังเคราะห์ กล่าวคือสังเคราะห์ข้อมูลการเคลื่อนไหวเพื่อนำไปทำให้สอดคล้องกับเสียงดนตรี [1],[2] กับอีกวิธีคือทำการสังเคราะห์เสียงดนตรีเพื่อนำไปใช้กับข้อมูลการเคลื่อนไหว [3] งานวิจัยของราจะมุ่งเน้นไปที่ด้านของการสังเคราะห์ข้อมูลการเคลื่อนไหวที่มีอยู่เพื่อนำไปใช้กับเสียงดนตรีที่รับเข้ามา

2.2.1 การสังเคราะห์การเคลื่อนไหวและการหาลักษณะเด่น

ในงานวิจัยทางด้านการสังเคราะห์การเคลื่อนไหวมีการนำเสนอหลากหลายวิธี และแต่ละวิธีก็มีจุดประสงค์ในการใช้แตกต่างกันออกไป โดยสามารถแบ่งออกเป็นสองประเภท

ใหญ่ ๆ ประเภทแรกคือการจำลองการเคลื่อนไหวของข้อต่อต่าง ๆในร่างกาย เนื่องจากว่าการแสดงออกของมนุษย์นั้นซับซ้อนมากและยากต่อการสร้างข้อมูลการเคลื่อนไหวดังกล่าวในแต่ละข้อต่อ อีกประเภทหนึ่งคือการสังเคราะห์การเคลื่อนไหวจากฐานข้อมูลที่มีอยู่แต่เดิม แต่วิธีการดังกล่าวก็ยังมีข้อจำกัดอยู่ตรงที่จำนวนของการเคลื่อนไหวรูปแบบใหม่จะขึ้นอยู่กับข้อมูลของการเคลื่อนไหวเดิมที่เรามีอยู่ในฐานข้อมูล กล่าวคือการจะสร้างการเคลื่อนไหวให้มีท่าทางหลากหลายจะขึ้นอยู่กับจำนวนท่าทางการเคลื่อนไหวที่เรามีอยู่ โดยวิธีการนี้คือวิธีที่เรานำมาใช้งานวิจัยฉบับนี้

ในงานวิจัยของ Xiao et al. [4] ได้นำเสนอกลุ่มของวิธีการหาลักษณะเด่นที่ใช้เป็นตัวแทนของท่าทางเคลื่อนไหว การแบ่งส่วนลำดับของการเคลื่อนไหว และการหาเฟรมสำคัญในลำดับของการเคลื่อนไหวซึ่งเป็นส่วนที่สำคัญมากสำหรับการสังเคราะห์การเคลื่อนไหว เพราะว่าวิธีการดังกล่าวสามารถแทนคุณลักษณะทางเรขาคณิตได้ และยังมีงานวิจัยอื่น ๆ อีกมากที่นำเสนอวิธีการสังเคราะห์การเคลื่อนไหว ตัวอย่างเช่น Arikan et al. [5] ได้นำเสนอวิธีการสังเคราะห์การเคลื่อนไหวโดยใช้วิธีกำกับชื่อให้กับทุก ๆข้อมูลการเคลื่อนไหวที่มีอยู่ในฐานข้อมูล โดยการกำกับจะกำกับได้ทั้งชื่อของการเคลื่อนไหวนั้น การเชื่อมต่อระหว่างข้อมูลต่าง ๆ เพื่อใช้ในการเรียกข้อมูลดังกล่าวมาใช้ แต่ข้อจำกัดของวิธีนี้คือการที่จะกำกับชื่อให้กับทุกเคลื่อนไหวนั้นจะต้องทำโดยผู้ใช้เอง J.K. Hodgkins et. al [6] ได้นำเอา state machine และ inverse kinematic มาใช้โดยจะใช้ SM ในการวิเคราะห์หาตำแหน่งที่ใช้ในการเปลี่ยนการเคลื่อนไหวและใช้ IK ในการเชื่อมต่อการเคลื่อนไหวดังกล่าวเข้าด้วยกัน อีกหนึ่งวิธีการที่ได้รับความนิยมมากอีกวิธีหนึ่งคือ วิธีที่เรียกว่า "Motion Graph" ซึ่งนำเสนอโดย Kovar et al. [7] ซึ่งก็เป็นกราฟโครงสร้างคล้ายกับของ move tree โดยที่โมชันกราฟจะเป็นกราฟเส้นตรงที่ประกอบไปด้วยชุดของการเคลื่อนไหวเดิมและการเคลื่อนไหวชุดใหม่ที่สร้างขึ้นมา

แต่ว่าข้อดีของขั้นตอนวิธีนี้คือระบบเป็นแบบอัตโนมัติ ในปี 2009 Liming et al. [8] ได้เสนอวิธีใหม่ในการสังเคราะห์การเคลื่อนไหวโดยพัฒนามาจากวิธีของโมชันกราฟ โดยผลลัพธ์ของที่ได้จะเป็นการเคลื่อนไหวชุดใหม่ที่มีความสมจริงมากขึ้นและ เวลาในการประมวลผลลดลง

2.2.2 การวิเคราะห์หาจังหวะของเสียงดนตรี

โดยทั่วไปแล้วจุดสำคัญระหว่างเสียงดนตรีและท่าเต้นคือการเข้าจังหวะระหว่างเสียงดนตรีและท่าเต้น ยิ่งไปกว่านั้นในงานทางด้านการออกแบบท่าเต้นจะใช้จังหวะเพื่อทำการเชื่อมโยงเพลงและการเคลื่อนไหวเข้าด้วยกัน อย่างไรก็ตามการวิเคราะห์หาโครงสร้างทางจังหวะจากเสียงดนตรีก็ยังถือเป็นงานที่ยาก แต่ก็ยังมีงานวิจัยที่ผ่าน ๆ มาได้ทำการเสนอวิธีการ

วิเคราะห์หาจังหวะของเสียงดนตรีเอาไว้ และเราก็จะอ้างอิงถึงวิธีการดังกล่าวแล้วนำมาแก้ไขปรับปรุงเพื่อนำมาใช้กับเสียงดนตรีในงานนี้

ในงานวิจัยของ Hainsworth และ Macleod [9] ได้เสนอขั้นตอนการวิเคราะห์หาจังหวะของเสียงเพลงโดยหาจุดเปลี่ยนของเสียงดนตรีจากขั้นตอนวิธีการกรอง ในงานวิจัยของ Goto [10] ได้เสนอว่าการตรวจหาจุดสำคัญของข้อมูลทางเสียงดนตรีนั้นสามารถทำได้ในการประมวลผลแบบออฟไลน์ และได้ทำการพัฒนาขั้นตอนวิธีการวิเคราะห์หาจังหวะออกเป็นสองแบบ คือแบบเพลงที่ประกอบด้วยเสียงกลองและไม่มีเสียงกลอง นอกจากนี้ยังมีงานวิจัยของ Shiratori et al. [11],[12] ที่ใช้วิธีการวิเคราะห์หาจังหวะของเสียงดนตรีในการหาคุณลักษณะของการเคลื่อนไหวแบบเข้าจังหวะแล้วแปลงให้อยู่ในรูปของท่าทางการเคลื่อนไหวแบบพื้นฐานตามจังหวะเสียงดนตรี

2.2.3 การทำให้การเคลื่อนไหวและเสียงดนตรีเคลื่อนไหวแบบสอดคล้องกัน

งานวิจัยหลาย ๆ งานได้นำเสนอวิธีการประสานเวลาระหว่างการเคลื่อนไหวและเสียงดนตรีเข้าด้วยกัน โดยสามารถจำแนกออกเป็นสองประเภทใหญ่ขึ้นอยู่กับว่าจะเลือกเอาส่วนไหนระหว่างส่วนของข้อมูลการเคลื่อนไหวหรือว่าจะนำเอาเสียงดนตรีมาใช้

ประเภทแรกจะเป็นการนำเอาเสียงดนตรีมาเป็นตัวกำหนดท่าทางการเคลื่อนไหวดังเช่นในงานของ Cardle et. a.l [1] ที่ได้นำเอาตัวโน๊ตมาควบคุมการเคลื่อนไหวของวัตถุในสภาพแวดล้อมจำลองแบบสามมิติ H.Lee et. al. [13] ได้นำเสนอวิธีการสังเคราะห์เสียงดนตรีขึ้นมาใหม่เพื่อนำไปใช้กับข้อมูลการเคลื่อนไหวและยังทำการปรับแต่งข้อมูลของการเคลื่อนไหว

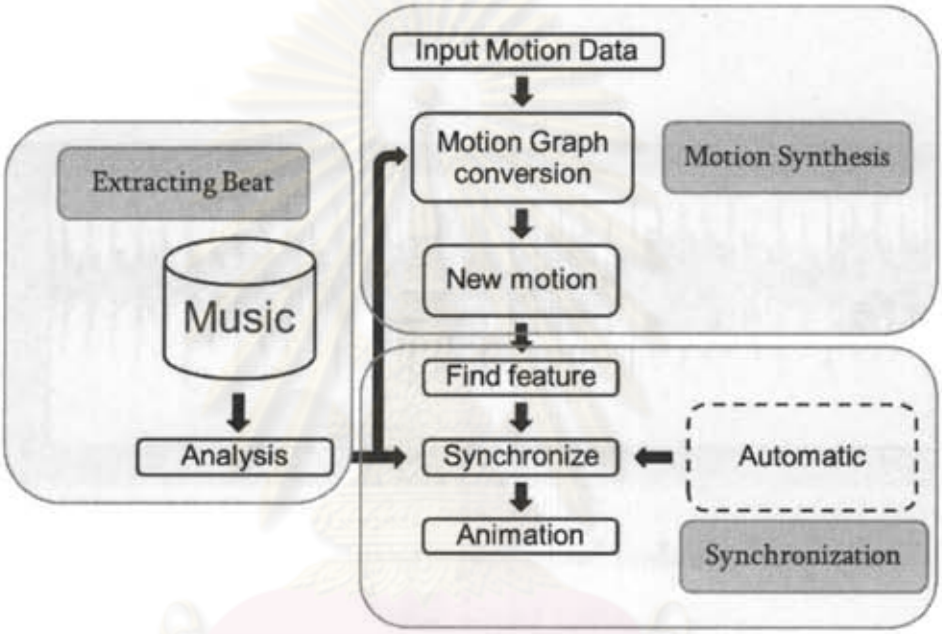
Kim et al. [14] ได้นำเสนอวิธีการสำหรับสกัดหาท่าทางพื้นฐานโดยประมาณเอาจากจังหวะของท่าทางการเคลื่อนไหว แล้วนำไปใช้สำหรับสร้างการเคลื่อนไหวชุดใหม่ แต่ว่าในความเป็นจริงนั้นท่าทางการเคลื่อนไหวของมนุษย์นั้นจะขึ้นอยู่กับจังหวะของเสียงดนตรี เพราะว่าโดยทั่วไปแล้วนั้นคนเราจะเต้นตามจังหวะของเสียงดนตรี

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3
ขั้นตอนการดำเนินงานวิจัย

3.1 ภาพรวมของระบบ

ภาพรวมของขั้นตอนการทำงานวิจัย มีลักษณะดังนี้



รูปที่ 3.1 ภาพรวมของขั้นตอนการทำงานวิจัย

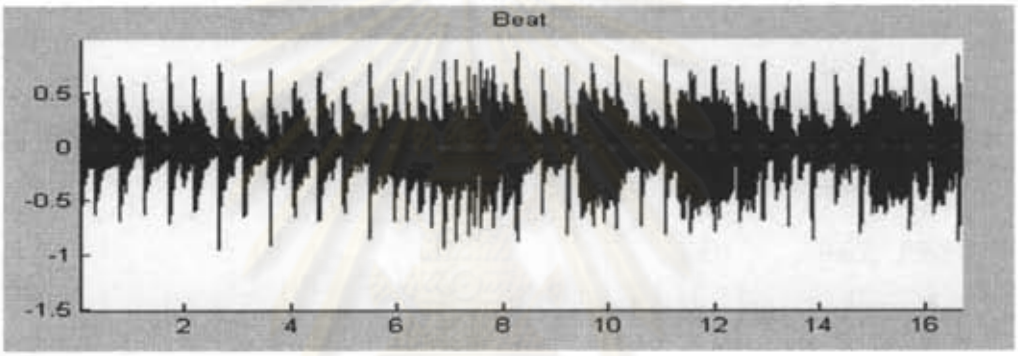
3.2 การวิเคราะห์หาจังหวะของเสียงดนตรี

นิยามของคำว่าจังหวะของเสียงดนตรีหรือว่า "Beat" คือ รูปแบบที่เกิดขึ้นแบบซ้ำ ๆ ของจังหวะในเสียงเพลง โดยที่บางทีจะถูกเรียกว่า "Tempo" ซึ่งมีหน่วยเป็น Beat Per Minute (BPM) และจังหวะของเสียงเพลงนั้นจะไม่มีการเปลี่ยนแปลง หมายความว่าถ้าเรามีจังหวะของเสียงเพลงเป็น 60 BPM ก็จะแปลว่าทุก ๆ 1 วินาทีจะเป็นจังหวะของเสียงดนตรี

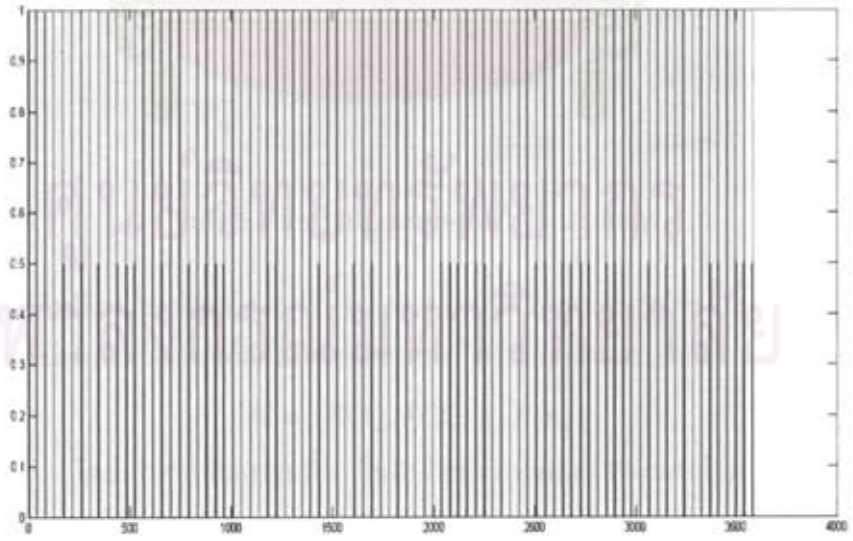
3.2.1 การคำนวณหาตำแหน่งของจังหวะดนตรี

ในขั้นตอนนี้เราจะใช้วิธีการที่เรียกว่า "Beat Tracking" ซึ่งได้นำเสนอวิธีการไปแล้วในบทที่ 2 เพื่อนำมาคำนวณหาตำแหน่งที่เป็นจังหวะของเสียงดนตรี โดยเราจะถือว่าตำแหน่งที่เป็นตำแหน่งของจังหวะนั้นคือตำแหน่งที่มีลักษณะเด่นของเสียงดนตรี (ดังรูปที่ 3.2)

หลังจากนั้นเราจะทำการแปลงตำแหน่งที่เป็นตำแหน่งของจังหวะเสียงเพลงให้อยู่ในรูปของกราฟที่แสดงถึงลักษณะเด่น (ดังรูปที่ 3.3 โดยที่แกน X จะเป็นแกนเวลาที่แสดงออกมาในรูปแบบของเฟรม ที่ต้องแปลงให้อยู่ในรูปแบบนี้เพราะว่าละเอียดกว่า และสะดวกเวลานำไปใช้จับคู่กับลักษณะเด่นของการเคลื่อนไหวในขั้นตอนที่ 3.4) สาเหตุสำคัญที่เราจำเป็นต้องคำนวณหาตำแหน่งที่เป็นจังหวะของเสียงดนตรี นั้นก็เพราะว่าเราจะทำการเชื่อมโยงเสียงดนตรีให้เข้ากับการเคลื่อนไหวที่เราได้สร้างขึ้นมา และอีกประการหนึ่งคือเราจะนำเอาความยาวของเสียงเพลงไปเป็นตัวกำหนดว่าเราจะต้องสังเคราะห์ชุดการเคลื่อนไหวให้มีความยาวเท่ากันโดยจะนำไปใช้ในขั้นตอนที่ 3.3



รูปที่ 3.2 ตำแหน่งของจังหวะดนตรี (จุดสีแดง)

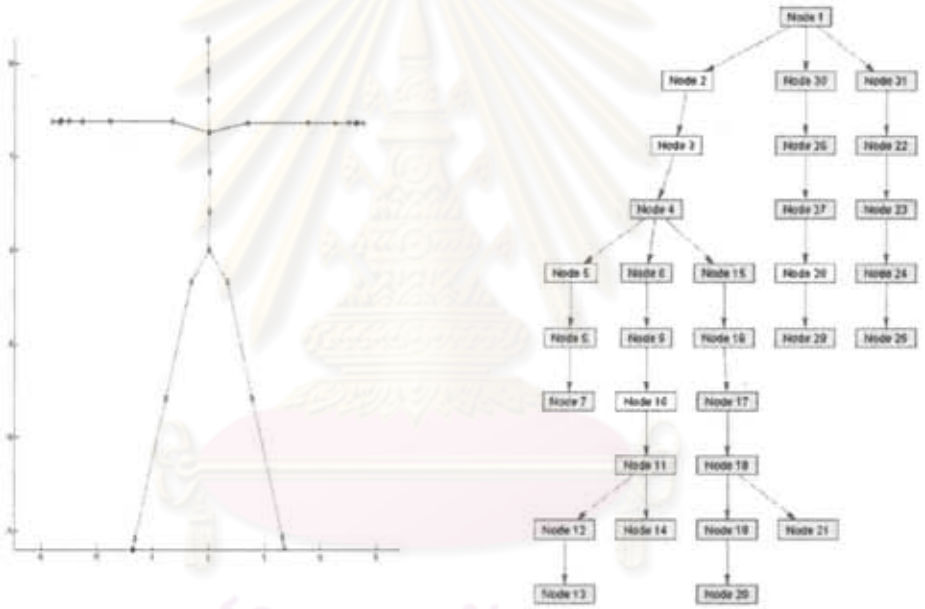


รูปที่ 3.3 ตำแหน่งของจังหวะดนตรีในรูปของฟังก์ชันที่เป็นลักษณะเด่น

3.3 การสังเคราะห์การเคลื่อนไหวโดยใช้โมชันกราฟ

3.3.1 การนำเสนอข้อมูลการเคลื่อนไหว

ข้อมูลของการเคลื่อนไหวที่เรานำมาใช้จะเป็นไฟล์ชนิดแอคเคลมไฟล์ (Acclaim File) ซึ่งประกอบไปด้วยสองส่วนอยู่ภายในส่วนแรกจะเรียกว่า "เอเอสเอฟ" (ASF) และส่วนที่สองเรียกว่า "เอเอมซี" (AMC) โดยในส่วนที่เรียกว่า เอเอสเอฟจะเก็บข้อมูลเกี่ยวกับโครงสร้างของโครงกระดูกร่างกายมนุษย์ ความสัมพันธ์ของข้อต่อแต่ละข้อต่อ และค่าของตัวแปรต่าง ๆ ในท่าตั้งต้น ดังรูปที่ 3.3 ทางซ้ายมือจะแสดงถึงโครงสร้างของโครงกระดูกมนุษย์ และทางขวามือจะแสดงถึงความสัมพันธ์ของข้อต่อแต่ละข้อต่อ โดยที่แต่ละโหนดจะหมายถึงข้อต่อ 1 ข้อ



รูปที่ 3.4 ข้อมูลที่ถูกเก็บอยู่ในไฟล์ เอเอสอฟ

รูปขององศาที่เปลี่ยนไป ข้อมูลของท่าทางการเคลื่อนไหวที่ถูกเก็บอยู่ในเอเอมซีจะแสดงอยู่ในรูปของลำดับของเฟรมได้ โดยที่แต่ละเฟรมจะแสดงถึงท่าของร่างกาย ดังสมการ

$$\boldsymbol{M} = \{\boldsymbol{F(1)}, \boldsymbol{F(2)}, \ldots, \boldsymbol{F(n)}\}$$

โดย $F(t)$ คือเฟรมที่ t ในลำดับของเฟรม M และภายในจะประกอบไปด้วยพิกัดของข้อต่อทุกข้อในร่างกาย

$$\boldsymbol{F(t)} = \left(\boldsymbol{T_{root}(t)}, \boldsymbol{R_{root}(t)}, \boldsymbol{R_1(t)}, \boldsymbol{R_2(t)} \ldots \boldsymbol{R_n(t)}\right)$$

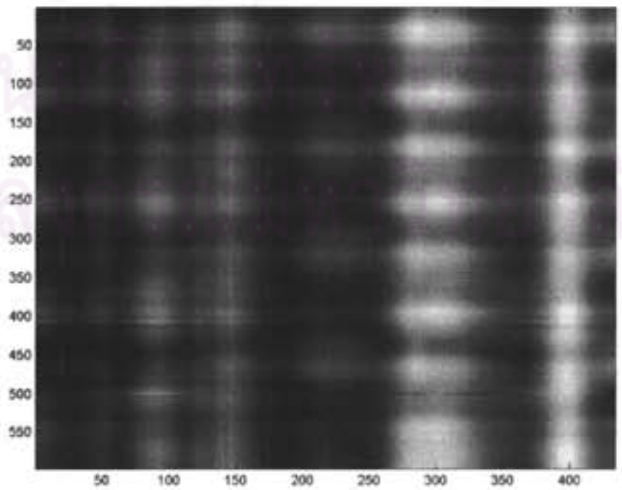
โดย $T_{root}(t)$ และ $R_{root}(t)$ คือการเลื่อนและเวกเตอร์ชี้ทิศทางที่เปลี่ยนไปของข้อต่อ ณ เวลา t และ $R_i(t)$ คือเวกเตอร์ชี้ทิศทางของข้อต่อ i เมื่อเทียบกับข้อต่อที่อยู่เหนือมันขึ้นไป ณ เวลา t

3.3.2 การแปลงพิกัดของข้อมูลการเคลื่อนไหวเพื่อนำไปวิเคราะห์

จากข้อมูลของการเคลื่อนไหวที่เรานำมาใช้จากฐานข้อมูลที่เรามีนั้น เนื่องจากมุมของข้อต่อต่างๆนั้นถูกเก็บอยู่ในพิกัดของออยเลอร์ ดังนั้นเราจะทำการแปลงพิกัดดังกล่าวให้ไปอยู่ในรูปพิกัดควอเตอร์เนียน สาเหตุที่เราต้องทำดังนี้เพราะว่าในพิกัดของออยเลอร์นั้นมีข้อจำกัดอยู่ตรงที่ค่าของมุมออยเลอร์หนึ่งค่าไม่ได้หมายถึงมุมออยเลอร์เพียงมุมเดียว ทำให้เวลาเรานำไปคำนวณ จะก่อให้เกิดความผิดพลาดขึ้นได้ในกระบวนการที่เราจะหาตำแหน่งที่จะเป็นตำแหน่งของการเชื่อมต่อระหว่างคลิปการเคลื่อนไหว

3.3.3 การหาตำแหน่งที่จะเป็นการเชื่อมต่อระหว่างคลิป

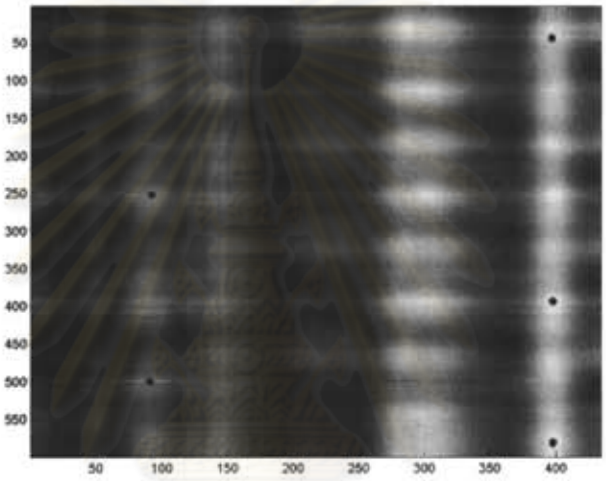
ในขั้นตอนนี้เราจะทำการหาตำแหน่งของเฟรมที่สามารถเปลี่ยนแปลงจากคลิปตั้งต้นไปยังคลิปถัดไปได้ โดยในการหาตำแหน่งดังกล่าวนั้นเราจะทำโดยหาความเหมือนของมุมโคซายน์ระหว่างเฟรมทุกเฟรมของคลิปทั้งสอง แล้วนำค่ามุมที่เราคำนวณออกมาได้นั้นมาพิจารณาในเทอมของกลุ่มจุด โดยคู่เฟรมใดได้ค่าของมุมออกมาน้อยแปลว่า ณ ตำแหน่งนั้นอาจจะสามารถนำไปใช้เป็นตำแหน่งของการเปลี่ยนคลิปได้ โดยสีที่แสดงออกมาก็จะเป็นสีขาว แต่ถ้าได้ค่าของมุมออกมามากก็จะแปลว่า ณ ตำแหน่งนั้นไม่เหมาะสมที่นำไปใช้เป็นตำแหน่งของการเปลี่ยนแปลง โดยสีที่แสดงออกมาก็จะใช้เป็นสีดำ ดังรูปที่ 3.4 แกน Y จะหมายถึงคลิปที่หนึ่ง และแกน X จะหมายถึงคลิปที่สอง



รูปที่ 3.5 กลุ่มจุดที่แสดงตามค่าของมุมโคซายน์

3.3.4 การเลือกตำแหน่งที่สามารถเปลี่ยนแปลงได้

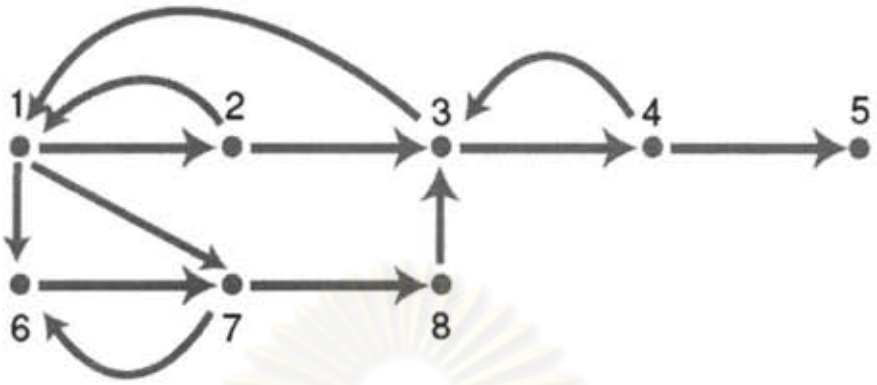
ในการที่เราจะหาจุดที่สามารถนำมาใช้เป็นจุดที่เปลี่ยนแปลงได้นั้น เราจะใช้ขั้นตอนวิธีที่เรียกว่า "Local Minima" ในการคำนวณ เนื่องจากเมื่อเราพิจารณาในเทอมของกลุ่มจุด ค่าของมุมโคซายน์ที่เราคิดมาได้จากขั้นตอนที่ 3.2.3 นั้นจะมีค่าที่ใกล้เคียงกันอยู่มาก และกระจายอยู่ทั่วทั้งกราฟดังจะสังเกตได้ดังรูปที่ 3.4 ดังนั้นเราจึงต้องใช้วิธีนี้ในการหาค่าที่ต่ำที่สุดรอบ ๆ บริเวณที่เราพิจารณา ผลที่จะแสดงดังรูปที่ 3.5 โดยจุดสีดำที่เห็นอยู่ในภาพจะแสดงถึงตำแหน่งที่สามารถเปลี่ยนแปลงไปยังคลิปต่อไปได้



รูปที่ 3.6 แสดงตำแหน่งของจุดเปลี่ยนแปลง

เป็นขันตอนในการปรับแต่งเอาเส้นทางการเคลื่อนไหวที่เป็นไปไม่ได้หรือไม่ต้องการออกจากกราฟของการเคลื่อนไหว (ดังรูปที่ 3.6) สาเหตุที่ต้องทำการกำจัดบางตำแหน่งออกเพราะว่า

1. จุดที่หามาได้อาจจะเป็นจุดที่ไม่สามารถสร้างเส้นทางสำหรับเชื่อมต่อไปยังคลิปการเคลื่อนไหวอื่นๆได้
2. สามารถปรับปรุงคุณภาพลำดับของการเคลื่อนไหวได้โดยการกำจัดจุดเปลี่ยนแปลงที่ไม่เหมาะสมออกไป
3. ลดความซับซ้อนของโมชันกราฟ
4. การจัดเก็บโมชันกราฟก็ไม่กินเนื้อที่ เนื่องจากโมชันกราฟที่ได้จะเป็นโมชันกราฟที่ผ่านการกำจัดจุดที่ไม่สามารถใช้ได้ออกไปแล้ว



รูปที่ 3.7 โมชันกราฟก่อนกำจัดจุดที่เป็นไปไม่ได้

ในรูปที่ 3.5 จะแสดงให้เห็นว่าจุดที่ [1, 2, 3, 6, 7, 8] เป็นจุดที่มีการถูกเลือกมาใช้มากที่สุดจึงจัดว่าเป็นส่วนประกอบของการเคลื่อนไหวที่เชื่อมต่อกันมากที่สุด และที่ตำแหน่ง 5 จะเรียกว่าเป็นจุดสิ้นสุดของการเคลื่อนไหวซึ่งเราจะกำจัดจุดนี้ออก สาเหตุที่เราต้องกำจัดตำแหน่งที่ 5 เพราะว่าถ้าระบบของเราเลือกตำแหน่งนี้ไปใช้ในการเชื่อมต่อคลิป จะทำให้คลิปนี้เล่นจนจบถึงแค่ตำแหน่งที่ 5 ไม่สามารถไปต่อได้

3.3.6 การสร้างการเคลื่อนไหวเพื่อเชื่อมต่อระหว่างคลิปของการเคลื่อนไหว

ในขั้นตอนนี้จะแบ่งออกเป็นสองขั้นตอนย่อย ขั้นตอนแรกคือเราจะทำการกำหนดก่อนว่าชุดการเคลื่อนไหวที่จะทำการสังเคราะห์ขึ้นมานั้นต้องมีความยาวเท่าใด โดยความยาวของชุดการเคลื่อนไหวนั้นจะทำการกำหนดโดยความยาวของเสียงเพลงที่เราได้ทำการวิเคราะห์ไปแล้วในขั้นตอนที่ 3.2 ในขั้นตอนต่อไประบบจะทำการเลือกคู่ของชุดการเคลื่อนไหวที่จะนำมาทำการเชื่อมต่อกัน ณ ตำแหน่งที่เป็นจุดของการเปลี่ยนแปลงด้วยวิธีการประมาณค่าในช่วงเชิงเส้นรอบวงกลม สาเหตุที่เราใช้วิธีการนี้ เพราะว่าพิกัดของข้อมูลการเคลื่อนไหวของเรานั้นเก็บอยู่ในรูปของพิกัดควอเตอร์เนียน และอีกสาเหตุหนึ่งคือวิธีการนี้สามารถสร้างการเชื่อมต่อระหว่างคลิปการเคลื่อนไหวได้อย่างราบรื่น และดูสมจริง

เราจะดำเนินการซ้ำในขั้นตอนที่สองไปเรื่อยๆจนกว่าเราจะสังเคราะห์ชุดของการเคลื่อนไหวที่มีความยาวเท่ากับความยาวของเสียงเพลง หลักเกณฑ์ที่เราใช้ในการเลือกว่าจะเอาชุดการเคลื่อนไหวไหนมาต่อกับชุดการเคลื่อนไหวในรอบถัดๆไปจะประกอบไปด้วยสามชนิด ชนิดแรกคือความสมจริงของตำแหน่งที่จะนำมาใช้เชื่อมต่อกัน ชนิดที่สองคือความยาวของการเคลื่อนไหวที่จะนำมาใช้ และชนิดสุดท้ายจะเป็นค่าการสุ่มเพื่อเพิ่มความหลากหลายในการ

เลือกใช้ชุดของการเคลื่อนไหว โดยเราจะให้ความสำคัญที่สองชนิดแรกมากกว่าชนิดสุดท้าย เนื่องจากสิ่งที่สำคัญที่สุดในการสังเคราะห์ชุดของการเคลื่อนไหวนั้นคือความสมจริงและราบรื่น

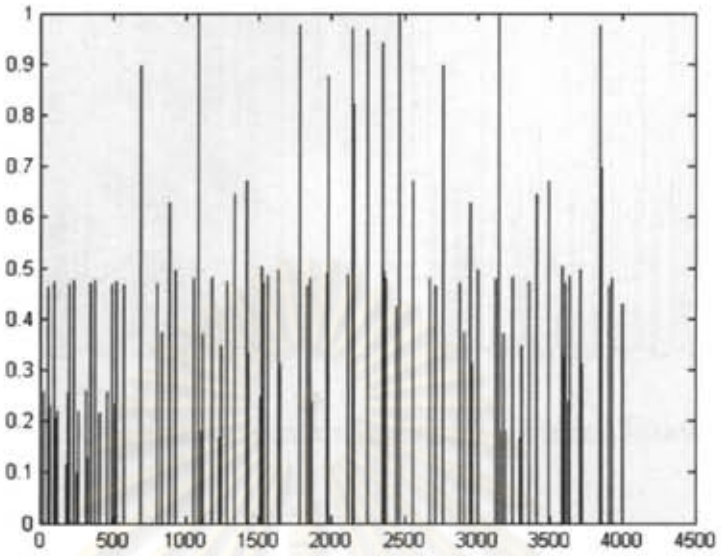
3.3.7 การแปลงพิกัดของข้อมูลการเคลื่อนไหวเพื่อนำไปใช้ประกอบเสียงดนตรี

จากขั้นตอนที่แล้วหลังจากที่เราได้คลิปการเคลื่อนไหวที่ทำการเชื่อมต่อเสร็จเรียบร้อยมาแล้วนั้น เราจะทำการแปลงพิกัดจากเดิมที่อยู่ในรูปของพิกัดควอเตอร์เนียน ให้กลับไปอยู่ในรูปของพิกัดออยเลอร์เพื่อจะได้นำไปใช้ในการหาลักษณะเด่นในขั้นตอนต่อไป อีกหนึ่งสาเหตุที่เราจำเป็นต้องแปลงพิกัดให้กลับไปอยู่ในรูปของพิกัดออยเลอร์ นั้นเพราะว่าในไฟล์เอเอมซีนั้น มุมของข้อต่อต่าง ๆ นั้นจะเก็บค่าในรูปแบบของพิกัดออยเลอร์ ดังนั้นหลังจากที่เราทำการเชื่อมต่อเสร็จเรียบร้อยเราจึงต้องทำการแปลงกลับ

3.3.8 การวิเคราะห์หาลักษณะเด่นของการเคลื่อนไหว

การหาลักษณะเด่นของท่าทางการเคลื่อนไหวทำโดยพิจารณาจากตำแหน่งการเคลื่อนไหวที่เปลี่ยนแปลงของข้อต่อที่เราสนใจ ในที่นี้คือข้อต่อของมือซ้าย มือขวา เท้าซ้าย และเท้าขวา สาเหตุที่เราใช้แค่สี่ตำแหน่งนี้นั้นก็เพราะว่าข้อต่อทั้งสี่ตำแหน่งนั้นเมื่อเกิดการเปลี่ยนขึ้นมาแล้วจะเป็นจุดสังเกตที่ชัดเจนที่สุดในร่างกายของมนุษย์ ยกตัวอย่างเช่น การยกเท้าและวางเท้าลงบนพื้น การเหวี่ยงแขน เป็นต้น ในส่วนของการคำนวณนั้นเราจะการคำนวณจากการหาจุดบนกราฟการเคลื่อนไหวของแต่ละข้อต่อที่มีค่าของความชันเป็นศูนย์ หรือก็คือจุดที่เป็นจุดเปลี่ยนความชัน จากนั้นก็จะแปลงให้ไปอยู่ในรูปของฟังก์ชันที่แสดงลักษณะเด่น (ดังรูปที่ 3.7 ในรูปทางซ้ายมือคือกราฟการเคลื่อนไหวของข้อต่อเพียงข้อเดียว ส่วนกราฟทางขวามือคือกราฟที่แสดงลักษณะเด่นของข้อต่อนั้นๆ) จากนั้นก็จะนำเอาฟังก์ชันที่แสดงถึงลักษณะเด่นของข้อต่อทั้งสี่ข้อมารวมกันในกราฟเดียว (ดังรูปที่ 3.8 ทางแกน Y จะแสดงถึงตำแหน่งของการเคลื่อนไหว ส่วนแกน X จะหมายถึงเฟรมการเคลื่อนไหว) เพื่อจะได้นำไปใช้ในขั้นตอนต่อไป คือขั้นตอนของการจับคู่เข้ากับจังหวะของเสียงเพลง

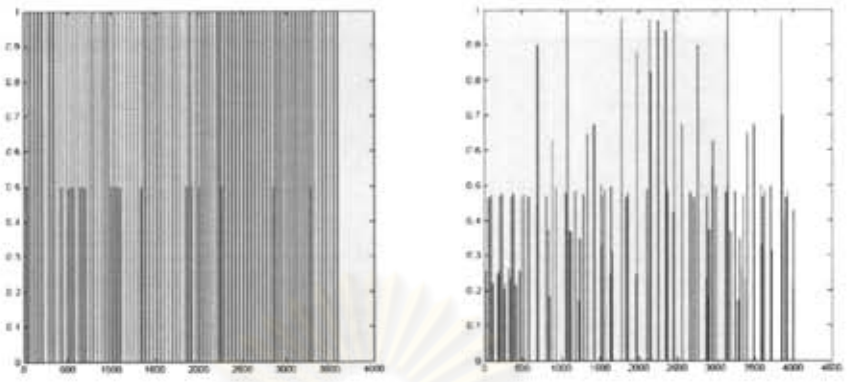
รูปที่ 3.8 ลักษณะเด่นของการเคลื่อนไหวของข้อต่อ



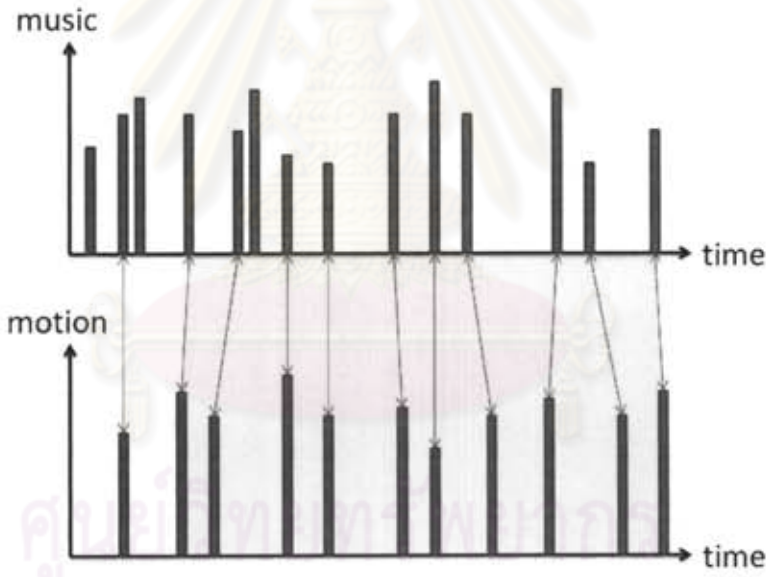
รูปที่ 3.9 กราฟที่แสดงลักษณะเด่นของข้อต่อที่เราสนใจ

3.4 การจับคู่ระหว่างลักษณะเด่นของการเคลื่อนไหวและเสียงดนตรี

ในขั้นตอนนี้จะเป็นขั้นตอนที่นำเอาลักษณะเด่นของการเคลื่อนไหวและจังหวะของเสียงเพลงมาจับคู่กัน โดยข้อมูลดังกล่าวเราได้หามาแล้วจากขั้นตอนที่ 3.2 และ 3.3 (ดังรูปที่ 3.10) โดยใช้วิธีจับคู่ โดยนำเอาลักษณะที่เด่นของท่าทางการเคลื่อนไหวมาจับคู่กับตำแหน่งที่เป็นจังหวะของเสียงเพลง (ดังรูปที่ 3.11) โดยเกณฑ์ที่ใช้ในการพิจารณาหาว่าตำแหน่งไหนถึงจะเหมาะสมในการนำมาจับคู่กัน คือความใกล้ไกลระหว่างตำแหน่งของลักษณะเด่นของการเคลื่อนไหวกับจังหวะของเสียงเพลง (ดังรูปที่ 3.12 แกน X จะแสดงถึงจังหวะของเสียงเพลง และแกน Y จะแสดงถึงตำแหน่งที่เป็นลักษณะเด่นของชุดการเคลื่อนไหวที่เราทำการสังเคราะห์ขึ้นมา ส่วนเส้นสีแดงคือเส้นที่บ่งบอกถึงตำแหน่งที่เกิดจากการจับคู่กันระหว่างลักษณะเด่นของการเคลื่อนไหวและจังหวะของเสียงเพลง) หลังจากนั้นเราจะทำการยืดหรือหดการเคลื่อนไหวตามตำแหน่งที่เป็นตำแหน่งของลักษณะเด่นของการเคลื่อนไหวเพื่อให้ได้ตำแหน่งเดียวกันกับจังหวะของเสียงเพลง โดยต้องคำนึงถึงว่าการปรับจะต้องไม่ทำให้ลักษณะท่าทางการเคลื่อนไหวเปลี่ยนแปลงไปจนทำให้ไม่สมจริง และข้อกำหนดของการจับคู่ก็คือ ทุก ๆ ลักษณะเด่นของการเคลื่อนไหวที่ได้ทำการสังเคราะห์ขึ้นมานั้นจะต้องถูกนำไปจับคู่กับลักษณะเด่นของเสียงเพลง

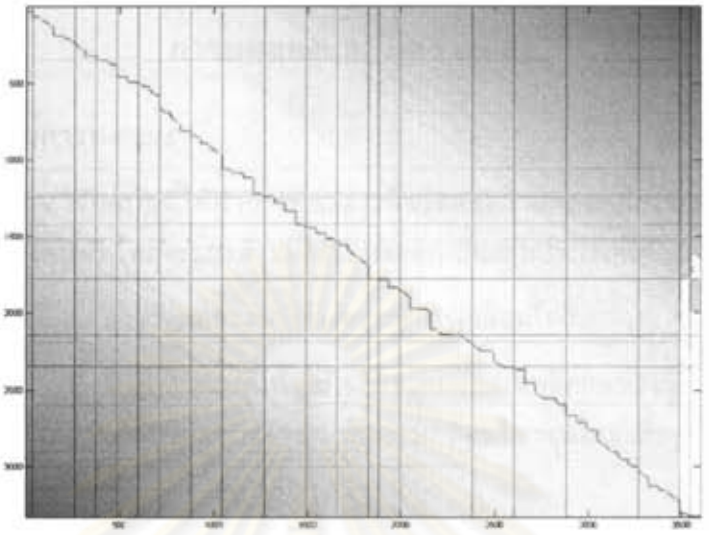


รูปที่ 3.10 ลักษณะเด่นของการเคลื่อนไหวและจังหวะของเสียงดนตรี



รูปที่ 3.11 การจับคู่ระหว่างการเคลื่อนไหวและเสียงดนตรี

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.12 ผลลัพธ์ของการจับคู่ระหว่างการเคลื่อนไหวและเสียงเพลง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4
การทดลองและวิเคราะห์ผล

4.1 ภาพรวมของการทดลอง

งานวิจัยฉบับนี้ได้ทำการทดลองเพื่อวัดผลความสอดคล้องระหว่างท่าทางการเคลื่อนไหวกับเสียงดนตรี โดยจะแบ่งด้านที่ใช้วัดผลออกเป็นสามประเภทได้แก่

1. การวัดผลของการสังเคราะห์การเคลื่อนไหว

มีวัตถุประสงค์เพื่อใช้ในการวัดผลประสิทธิภาพของวิธีการที่นำมาใช้ในงานวิจัย โดยทำการทดสอบออกมาเป็นการเคลื่อนไหวชุดใหม่เพื่อดูว่ามีความสมจริงมากน้อยเพียงใด

2. การทดสอบความถูกต้องของการวิเคราะห์เสียงดนตรี

มีวัตถุประสงค์เพื่อใช้ในการวัดความถูกต้องของตำแหน่งที่เป็นจังหวะของเสียงดนตรีแล้วนำไปเทียบกับตำแหน่งที่เป็นจังหวะของเสียงดนตรีที่แท้จริง

3. การทดสอบความถูกต้องของการเคลื่อนไหวประกอบเสียงดนตรี

มีวัตถุประสงค์เพื่อใช้ในการวัดความสมจริงของการเคลื่อนไหวประกอบเสียงดนตรีแบบเข้าจังหวะหลังจากที่มีการทำให้การเคลื่อนไหวและเสียงเพลงมีการแสดงออกแบบสอดคล้องกัน

4.2 ผลของการสังเคราะห์การเคลื่อนไหว

ในฐานข้อมูลของการเคลื่อนไหวที่เรานำมาใช้จะมีแต่กลุ่มของการเคลื่อนไหวแบบท่าเต้นเท่านั้น ซึ่งจะมีทั้งหมด 20 คลิปการเคลื่อนไหว และแต่ละคลิปการเคลื่อนไหวก็จะมีระยะเวลาในการแสดงออกไม่เท่ากัน โมชันกราฟจะถูกสร้างขึ้นจากฐานข้อมูลนี้ จากคลิปการเคลื่อนไหวแบบท่าเต้นทั้ง 20 คลิป เมื่อนำมาสังเคราะห์จะได้จุดเปลี่ยนแปลงทั้งหมดประมาณ 500 จุดที่สามารถเปลี่ยนแปลงได้

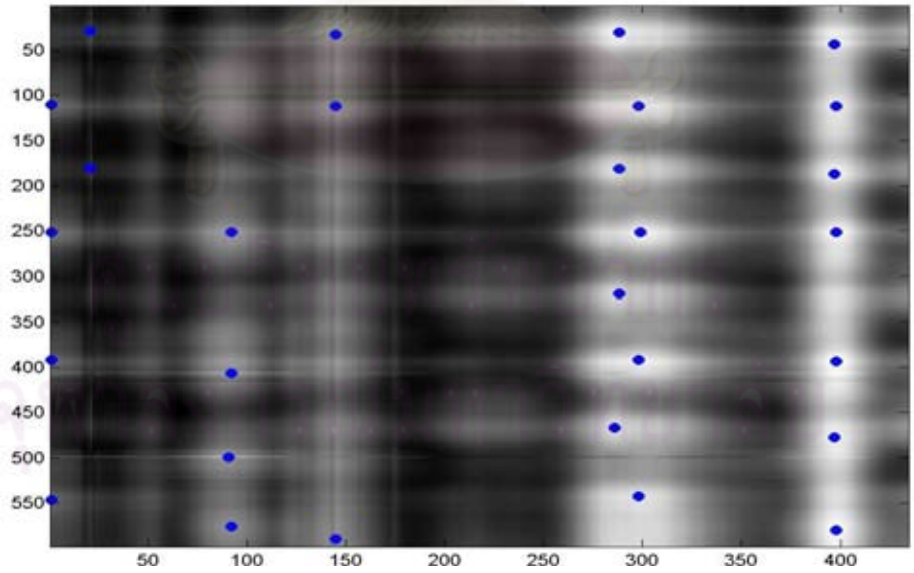
จากผลการทดลองเรื่องการสังเคราะห์การเคลื่อนไหวทำให้ทราบว่าการจะคำนวณหาจุดที่สามารถเปลี่ยนแปลงไปยังคลิปการเคลื่อนไหวอื่นนั้นเป็นงานที่ยากและใช้เวลานานในการคำนวณค่อนข้างมากตั้งแต่การปรับเตรียมข้อมูลในฐานข้อมูล การแปลงพิกัด การคำนวณหามุมระหว่างสองเวกเตอร์ จนกระทั่งถึงการเชื่อมต่อระหว่างคลิปการเคลื่อนไหว งานที่ระยะเวลานานที่สุดจะเป็นในขั้นตอนของการสร้างโมชันกราฟ เพราะว่าการจะสร้างโมชันกราฟได้นั้นจะต้องทำตั้งแต่การเตรียมฐานข้อมูลการเคลื่อนไหว การเก็บค่าของการเคลื่อนไหว

ของทุกคลิปการเคลื่อนไหว การแปลงค่าที่ได้เก็บมาให้อยู่ในรูปที่สามารถนำไปใช้สำหรับการเชื่อมต่อระหว่างคลิปการเคลื่อนไหวได้ การคำนวณหาตำแหน่งที่จะเป็นการเชื่อมต่อระหว่างคลิป ดังแสดงตารางที่ 4.1

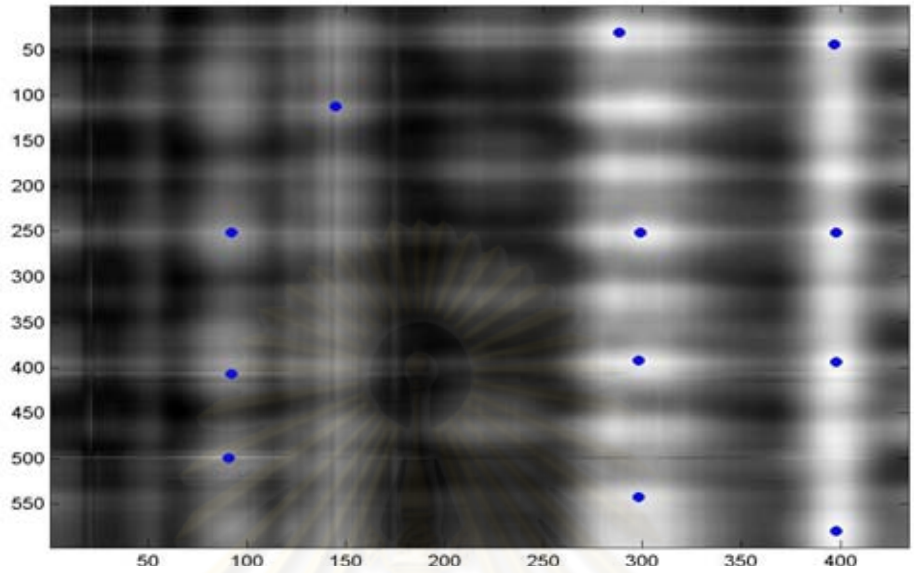
Process	Construct Motion graph	Motion synthesis	Extracting Motion Feature
Time(min)	20	0.1	2.2

ตาราง 4.1 เวลาที่ใช้ในการประมวลผลการสังเคราะห์การเคลื่อนไหว

การสังเคราะห์การเคลื่อนไหวจะมีตัวแปรที่สำคัญคือค่าที่ใช้ในการหา Local Minima ซึ่งตัวแปรตัวนี้จะเป็นตัวกำหนดว่าเราต้องการจุดเปลี่ยนที่มีความสมจริงมากแค่ไหน ถ้าเรายิ่งกำหนดค่าของตัวแปรนี้มาก จุดเปลี่ยนที่ได้ก็จะยิ่งมีจำนวนน้อยนั้นหมายความว่า จุดเปลี่ยนที่ให้ความสมจริงนั้นมีน้อย ในทางตรงกันข้ามถ้าเรากำหนดค่าของตัวแปรนี้น้อย ผลที่ได้ออกมาก็คือจุดเปลี่ยนจะมากขึ้นแต่ว่าจุดเปลี่ยนที่ได้เพิ่มนั้นจะมีความสมจริงน้อยลง สมมติให้ตัวแปรดังกล่าวคือตัวแปร X ดังรูป 4.1 และ 4.2



รูปที่ 4.1 ตำแหน่งที่สามารถเชื่อมต่อกันได้ เมื่อ X = 50



รูปที่ 4.2 ตำแหน่งที่สามารถเชื่อมต่อกันได้ เมื่อ X = 90

4.3.1 วิเคราะห์โดยใช้ Beat Detection Error Rate

ในการวัดความถูกต้องของการวิเคราะห์หาจังหวะของเสียงดนตรี ขั้นแรกเราจะทำเครื่องหมายไว้ทุกตำแหน่งที่เป็นจังหวะของเสียงดนตรี และเรียกว่า “Reference-beats” หลังจากนั้นจะทำการเปรียบเทียบกับจังหวะที่เราตรวจจับได้จากวิธีการของเรา และเรียกมันว่า “Caught-beats” เกณฑ์การวัดความผิดพลาดมีอยู่ 2 แบบดังต่อไปนี้

$$\boldsymbol{Redundant\ rate} = \frac{(\boldsymbol{No.of\ redundant\ beats})}{(\boldsymbol{No.of\ caught\ beats})}$$

โดยที่ $No.of\ redundant\ beats$ คือ ตำแหน่งของจังหวะที่ไม่มีอยู่ใน $Reference - beats$

$$\boldsymbol{Missing\ rate} = \frac{(\boldsymbol{No.of\ missed\ beats})}{(\boldsymbol{No.of\ reference\ beats})}$$

โดยที่ $No.of\ missed\ beats$ คือ ตำแหน่งของจังหวะที่ตรวจจับไม่พบแต่ว่ามีอยู่ใน $Reference - beats$

	Pop music	Rock music	Mixed music
Redundant rate	0.07	0.01	0.41
Missing rate	0.1	0.01	0.03

ตารางที่ 4.1 เกณฑ์การวัดความผิดพลาดทั้งสองชนิด

4.4 ผลการทดสอบความถูกต้องของของการวิเคราะห์หาจังหวะของเสียงดนตรี

จากผลการทดลองที่ได้ออกมาดังตารางที่ 4.1 สามารถสรุปได้ว่า โดยส่วนมากนั้นจังหวะสามารถตรวจจับได้โดยวิธีการที่ได้นำเสนอไป แต่ว่าก็ยังมีตำแหน่งที่มีการตรวจจับผิดพลาดอยู่บ้าง และในเพลงประเภทที่มีเสียงกลองชัดเจนอย่างเช่นเพลงแนวร็อคนั้นจะมีผลการทดลองที่ดีกว่าเพลงแนวอื่น สาเหตุก็เพราะว่าเสียงกลองนั้นจะง่ายต่อการตรวจจับ

4.5 การทดสอบความถูกต้องของการเคลื่อนไหวประกอบเสียงดนตรี

4.5.1 การเปรียบเทียบเวลาของจังหวะและท่าทางการเคลื่อนไหว

	Length of Data (sec)		Number of feature points		Computational time (sec)	
	Music	Motion	Music	Motion	DP matching	DTW
1st example	30	33	84	104	12.0472	9.3297
2nd example	30	31	57	102	11.5900	9.7720

ตารางที่ 4.2 ตารางเปรียบเทียบเวลาของจังหวะและท่าทางการเคลื่อนไหว

4.6 อภิปรายผลการทดลอง (Discussion)

จากผลการทดลองข้างต้น ทำให้สรุปได้ว่าการสังเคราะห์การเคลื่อนไหวแล้วนำมาทำให้มีการเคลื่อนไหวแบบเข้าจังหวะกับเสียงดนตรีนั้นจะเสียเวลาส่วนมากไปกับการเตรียมฐานข้อมูล จัดการเบื้องต้นกับข้อมูลการเคลื่อนไหวที่มีทั้งหมดในฐานข้อมูล และอีกส่วน

หนึ่งที่ใช้เวลารองลงมาก็คือการวิเคราะห์หาตำแหน่งที่เป็นลักษณะเด่นของการเคลื่อนไหวที่ได้ทำการสังเคราะห์ขึ้นมาเพื่อนำไปใช้ในการจับคู่กับจังหวะของเสียงเพลง

ในส่วนของผลลัพธ์ที่ได้ยังมีส่วนที่ไม่ดีอยู่คือ ชุดการเคลื่อนไหวใหม่ที่ได้สังเคราะห์ขึ้นมาหลังจากโดนการยืด หรือหดไปแล้วนั้น บางครั้งทำให้มีการแสดงออกที่ไม่สมจริง บางท่าการเคลื่อนไหวก็ถูกทำให้มีการแสดงออกที่เร็วเกินไป หรือบางครั้งก็ถูกทำให้ยืดจนมีการแสดงออกที่ดูช้าเกินกว่าปกติ สาเหตุที่ทำให้เป็นแบบนั้นเพราะว่าในบางครั้งระยะระหว่างช่วงของการเคลื่อนไหวและช่วงของเสียงเพลงมีความต่างกันมากเกินไป ซึ่งในงานวิจัยนี้ยังไม่ได้พิจารณาในจุดนี้ ดังนั้นในงานวิจัยต่อไปจึงจะมีการนำเอาปัญหาในส่วนนี้มาแก้ไขโดยอาจจะเพิ่มข้อบังคับในการยืดหรือหดชุดของการเคลื่อนไหวที่จะนำไปจับคู่กับเสียงเพลงว่าความต่างทางเวลาของข้อมูลทั้งสองชนิดนั้นควรจะมีค่าประมาณเท่าใด ถ้านอกเหนือจากค่าที่ได้กำหนดไว้อาจจะยอมให้ข้ามช่วงของเพลงในช่วงนั้นไปได้ แล้วพิจารณาช่วงของเพลงในช่วงถัดไปแทน หรือในกรณีที่ไม่ได้จริง ๆ อาจจะทำการประมาณค่าช่วงของเสียงเพลงใหม่ขึ้นมาเพื่อนำไปใช้ในการจับคู่กับชุดของการเคลื่อนไหวที่ได้ทำการสังเคราะห์มา

ในการวัดผลการทดลองจะใช้วิธีที่เรียกว่า “Hallway Testing” ซึ่งเป็นวิธีที่ใช้สำหรับทดสอบว่าผลการทดลองที่ได้มานั้น มีข้อบกพร่องตรงไหน และควรจะแก้ไขในส่วนใด โดยจะทำการทดสอบโดยเปิดให้ผู้คนทั่วไปชมวีดีโอที่เป็นผลลัพธ์ของงานวิจัย พร้อมกับให้ตอบคำถามตามที่เราได้ตั้งไว้ให้ จำนวนของคนที่ทำการทดสอบจะอยู่ที่ 10 คน โดยแบ่งออกเป็น 2 กลุ่ม กลุ่มละ 5 คน โดยกลุ่มแรกจะเป็นกลุ่มคนที่มีความรู้ความเข้าใจในงานวิจัยประเภทนี้ ส่วนอีกกลุ่มจะเป็นกลุ่มคนที่ไม่มีความรู้ทางด้านนี้ สาเหตุที่ใช้คนแค่กลุ่มละ 5 คนเพราะว่าได้มีงานวิจัยให้ผลสรุปออกมาแล้วว่าในการจะทำการทดสอบประเภทนี้ใช้คนแค่ 5 คนก็เพียงพอ ในส่วนของคำถามที่ใช้ถามผู้ที่เข้ามาทดสอบจะมีอยู่ 2 ข้อ

1. หลังจากดูแล้วคิดว่าควรจะปรับปรุงในด้านไหนบ้าง
    a. ความเร็วของการเคลื่อนไหว
    b. การเข้าจังหวะระหว่างการเคลื่อนไหวกับเสียงเพลง
    c. ความหลากหลายของท่าเต้น
2. หลังจากดูแล้วคิดว่ามีความน่าสนใจ หรือว่ามีความสมจริงมากน้อยเพียงใด

ข้อมูลที่ได้มาจากการทำการทดสอบจะนำไปใช้ในการปรับปรุงแก้ไข และพัฒนางานวิจัยต่อไปในอนาคต

บทที่ 5
สรุปผลการวิจัย และข้อเสนอแนะ

5.1 สรุปผลการวิจัย (Conclusion)

ในงานวิจัยนี้ได้นำเสนอวิธีการสังเคราะห์ชุดของการเคลื่อนไหวแบบใหม่ขึ้นมาได้จากฐานข้อมูลที่มีอยู่เดิม ในขณะเดียวกันยังสามารถที่จะทำให้ท่าทางการเคลื่อนไหวและเสียงดนตรีมีการแสดงออกแบบสอดคล้องกันโดยใช้ลักษณะเด่นของทั้งข้อมูลการเคลื่อนไหวและเสียงดนตรี วิธีการนี้เป็นแบบอัตโนมัติ

5.2 ข้อเสนอแนะ (Future Work)

ในงานถัดไปได้มีการวางแผนว่าจะเพิ่มเติมในส่วนติดต่อกับผู้ใช้ ซึ่งจะสามารถให้ผู้ใช้เลือกได้ว่าในการสังเคราะห์การเคลื่อนไหวชุดใหม่นั้น ต้องการให้มีการเคลื่อนไหวรูปแบบใดประกอบบ้าง อาจจะเลือกแค่บางท่า หรืออาจจะกำหนดไปเลยว่าให้มีได้แค่ท่าบางประเภทเท่านั้น ซึ่งในการจะทำให้ได้ผลลัพธ์ดังกล่าว อาจจะต้องมีการทำการให้คำนิยามกับท่าทางการเคลื่อนไหวทุก ๆ ท่า ที่เรามีในฐานข้อมูลเพื่อที่จะได้เลือกนำมาใช้ในการสังเคราะห์ได้ถูกประเภท เนื่องจากว่าในงานวิจัยนี้ไม่ได้มีการกำหนดเรื่องท่าทางการเคลื่อนไหว แต่จะพิจารณาถึงความเป็นไปได้ของการสร้างจุดการเปลี่ยนแปลงจากทุก ๆ คลิปที่เรามีอยู่ เป็นการสังเคราะห์การเคลื่อนไหวแบบสุ่ม

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รายการอ้างอิง

[1] Cardle, M., Barthe, L., Brooks, S., and Robinson, P., Music-driven motion editing: local motion transformations guided by music analysis, Proceedings 20th Eurographics UK Conference, (2002): 38-44.

[2] Kim, T.-hoon, Park, S. I., and Shin, S. Y., Rhythmic-motion synthesis based on motion-beat analysis, ACM Transactions on Graphics, 22 (July 2003): 392.

[3] Hahn, J. K., Fouad, H., Gritz, L., and Lee, J. W., Integrating Sounds and Motions in Virtual Environments, Presence: Teleoperators and Virtual Environments, 7 (February 1998): 67-77.

[4] Xiao, J., Zhuang, Y., Wu, F., Guo, T., and Liang, Z., A group of novel approaches and a toolkit for motion capture data reusing, Multimedia Tools and Applications, 47 (July 2009): 379-408.

[5] Arikan, O., Forsyth, D. a, and O'Brien, J. F., Motion synthesis from annotations, ACM Transactions on Graphics, 22 (July 2003): 402.

[6] Safonova, A. and Hodgins, J. K., Construction and optimal search of interpolated motion graphs, ACM Transactions on Graphics, 26 (July 2007): 106.

[7] Kovar, L., Gleicher, M., and Pighin, F., Motion graphs, ACM Transactions on Graphics, 21 (July 2002): 1-10.

[8] Zhao, L. and Safonova, A., Achieving good connectivity in motion graphs, Graphical Models, 71 (July 2009): 139-152.

[9] Hainsworth, S. and Macleod, M., Beat tracking with particle filtering algorithms, Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop on., IEEE, , p. 91–94, 2003.

[10] Goto, M., An Audio-based Real-time Beat Tracking System for Music With or Without Drum-sounds, Journal of New Music Research, 30 (June 2001): 159-171.

[11] Shiratori, T., Nakazawa, a, and Ikeuchi, K., Rhythmic motion analysis using motion capture and musical information, Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, MFI2003., (2003): 89-94.

[12] Shiratori, T., Nakazawa, A., and Ikeuchi, K., Dancing-to-Music Character Animation, Computer Graphics Forum, 25 (September 2006): 449-458.

[13] Lee, H.-C. and Lee, I.-K., Automatic Synchronization of Background Music and Motion in Computer Animation, Computer Graphics Forum, 24 (September 2005): 353-361.

[14] Kim, G., Wang, Y., and Seo, H., Motion Control of a Dancing Character with Music, 6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007), (2007): 930-936.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นาย อรรถวุฒิ หลายชูไทย เกิดวันที่ 17 กันยายน พ.ศ. 2529 สำเร็จการศึกษาระดับมัธยมศึกษาที่โรงเรียนสาธิตมหาวิทยาลัยรามคำแหง จากนั้นจึงเข้าศึกษาต่อที่คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ในปีการศึกษา 2551 และเข้าศึกษาต่อในหลักสูตรวิศวกรรมศาตรมหาบัณฑิตต่อในปีการศึกษา 2552 ในภาควิชาคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย