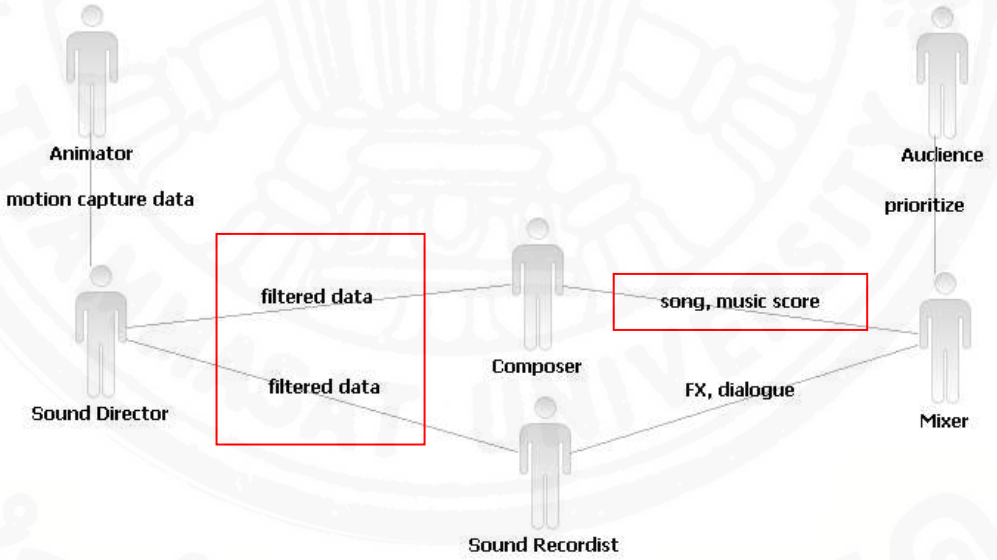
บทที่ 3

วิธีการดำเนินงานวิทยานิพนธ์

3.1 องค์ประกอบที่ใช้ในวิทยานิพนธ์

ภาพรวมของวิทยานิพนธ์นี้อาจกล่าวง่ายๆว่าเป็นการประมวลผลจากภาพเคลื่อนไหวไปสู่ดนตรี แต่ในขั้นตอนการทำงานจริงด้วยมนุษย์แล้วเกิดจากการร่วมมือทำงานจากหลายบุคลากร มีโปรแกรมหลากหลายที่นำมาใช้ในการประมวลผลนี้ซึ่งจำเพาะเจาะจงในแต่ละสายงาน ในแผนภาพต่อไปนี้แสดงถึงขั้นตอนการทำงานและบทบาทของแต่ละคน

ภาพที่ 3.1
บทบาทและขั้นตอนการทำงานในการสร้างเสียงประกอบ



เมื่อ Animator ส่งข้อมูล Motion Capture Data ให้ Sound Director ในการสื่อสารให้เข้าใจกันระหว่างทั้งสองสายงานจำเป็นต้องมีเอกสารหรือระเบียบการปฏิบัติเพื่อให้ได้การประมวลผลที่ถูกต้อง จากจุดนี้ Sound Director พิจารณาถึงจุดที่ควรใส่เสียงหรือดนตรีลงไปโดยส่งงานให้ทาง Sound Recodist หรือ Composer สร้างเสียงมาให้ นอกจากจุดที่ควรใส่เสียงลงไป Sound Director ทำหน้าที่กำหนดอารมณ์ของตัวละคร สถานการณ์และสภาพแวดล้อมของ

ภาพยนตร์นั้น ซึ่งในวิทยานิพนธ์นี้จำเป็นต้องกรองเอาค่าต่างเหล่านั้นออกมาและทำการแปลงข้อมูลเหล่านั้นให้เป็นดนตรี

ในกรณีที่มีเสียงหรือเสียงเครื่องดนตรีเกิดขึ้นมากเกินไป เราไม่สามารถรับฟังหรือรับรู้เสียงได้ทั้งหมดในเวลาเดียวกัน เราจึงจำเป็นต้องมี Mixer เพื่อรักษาสมดุลของเสียงเหล่านั้น โดยอาจต้องเพิ่มหรือลดบางเสียงลงตามความจำเป็นก่อนถึงส่วนของผู้ฟัง (Audience) การกรองข้อมูลที่เกิดขึ้นนั้นเราต้องทราบถึงความสำคัญของแต่ละเสียงเช่นเดียวกับการกรองข้อมูลออกมาจากการจับความเคลื่อนไหวในบางครั้งข้อมูลที่มีความสำคัญน้อยกว่าอาจถูกตัดทิ้งไปเลย

จากภาพรวมดังที่กล่าวมาข้างต้นนี้ สามารถแยกงานส่วนที่สำคัญออกมาได้ถึง 3 งานได้แก่ Motion Capture Data-Filtering, Music Composition และ Synchronization Mapping ซึ่งในการทำงานร่วมกันของทั้ง 3 งานนี้เรายังต้องคำนึงถึงการผ่านค่าพารามิเตอร์ต่างๆ ระหว่างกันด้วย

3.2 ขั้นตอนการกรองข้อมูลจากการเคลื่อนไหว (Motion Capture Data-Filtering)

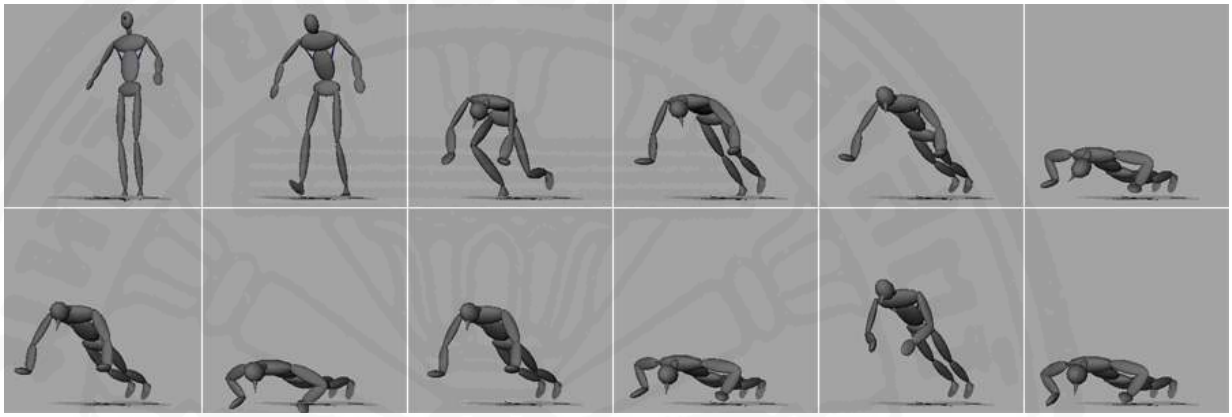
สำหรับการกรองข้อมูลของที่ได้จากการจับการเคลื่อนไหวของมนุษย์ คำถามที่เกิดขึ้นคือ

3.2.1 ขั้นตอนการปรับข้อมูล

“เราจะต้องปรับเปลี่ยนข้อมูลให้เป็นแบบใด จึงจะสามารถใช้งานได้?” การใช้งานข้อมูลโดยตรงไปตรงมาเพียงอย่างเดียวไม่สามารถให้ข้อมูลได้มากนักจึงจำเป็นต้องปรับเปลี่ยนเป็นรูปแบบอื่น เช่นเปลี่ยนแปลงจากองศาในการหมุน (Rotation) มาเป็นตำแหน่งที่เกิดขึ้นจริงในสภาพแวดล้อมนั้น (World Co-ordinate)

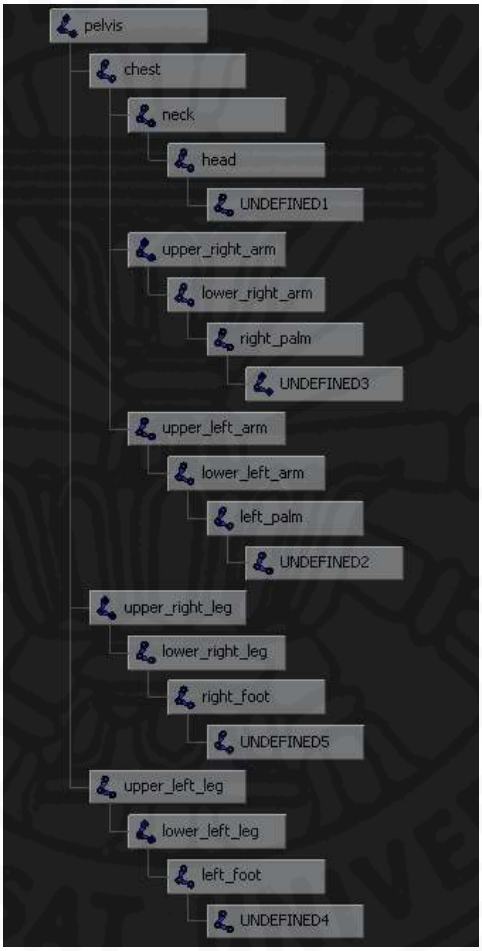
เพื่ออธิบายถึงขั้นตอนเบื้องต้นในการกรองข้อมูล ขอยกตัวอย่างจากข้อมูลของการเคลื่อนไหวในการวิดพื้นดังรูปต่อไปนี้

ภาพที่ 3.2
การเคลื่อนไหวที่เกิดขึ้นในแต่ละเฟรม



เมื่อเริ่มต้นจากข้อมูลที่ได้จากเครื่องจับความเคลื่อนไหว ข้อมูลที่ได้อยู่ในรูปแบบของลำดับชั้นของแชนเนลดังรูป

ภาพที่ 3.3
ลำดับชั้นในการเชื่อมต่อของข้อต่อแต่ละจุด

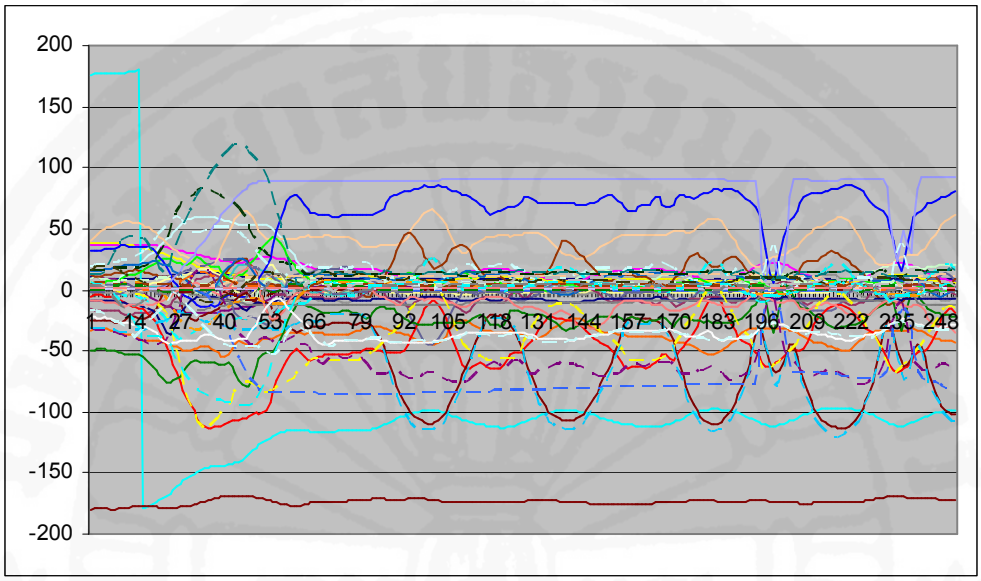


ซึ่งแต่ละแชนเนลเก็บมุมที่ใช้หมุนในแต่ละจุด หากเราเปลี่ยนข้อมูลเหล่านั้นเป็นรูปแบบของตำแหน่งที่เกิดขึ้นจริงในสภาพแวดล้อมแล้ว ข้อมูลหมายถึงพิกัด X และ Y ของแต่ละจุดแทน ในการทดลองเบื้องต้นเราสามารถปรับให้ Maya แปลงค่าอินพุทจากมุมที่ใช้หมุนเป็นตำแหน่งที่เกิดขึ้นจริงในสภาพแวดล้อมได้ ซึ่งในแต่ละจุดที่เกิดขึ้นนั้นสิ่งที่เราต้องการคือจุดที่บ่งบอกความสำคัญของการเคลื่อนไหวนั้นๆ เราสามารถปรับเปลี่ยนได้ทั้งวิธีคำนวณและรูปแบบขั้นตอนของการคำนวณเพื่อหาค่านั้นได้

3.2.1.1 **ข้อมูลจากการหมุน** (Rotation)

ภาพที่ 3.4

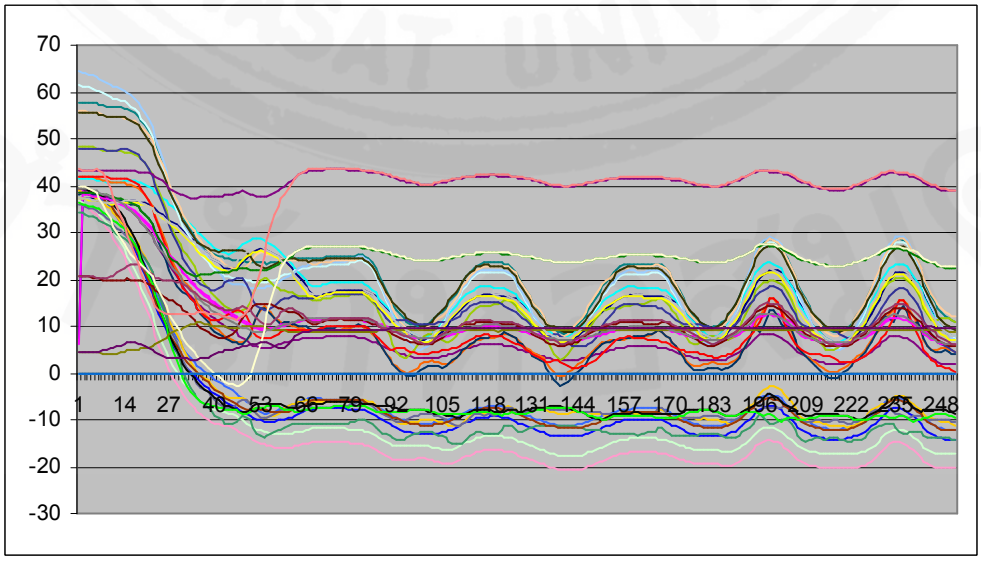
ข้อมูลแสดงกราฟที่ได้จากการหมุนของทุกแชนเนล และมุมของ X,Y และ Z



3.2.1.2 **ข้อมูลจากตำแหน่งที่เกิดขึ้นจริงในสภาพแวดล้อม** (World Coordinate)

ภาพที่ 3.5

ข้อมูลแสดงกราฟที่ได้จากตำแหน่งที่เกิดขึ้นจริงในสภาพแวดล้อมของทุกแชนเนล และพิกัดของ X,Y และ Z



3.2.2 ขั้นตอนการหาจุดเชื่อมต่อ

“เราจะกรองเอาจุดสนใจหรือ Spotting Session ออกมาได้อย่างไร?” ลักษณะของการเคลื่อนไหวที่เป็นจุดที่เข้ากับจังหวะนั้นหาได้จากหลายๆวิธีเช่น การหาจุดสูงสุดหรือต่ำสุดจาก Motion Graph, การหาระยะห่างจากจุดศูนย์กลางหรือตำแหน่งที่เกิดความเร่งสูงสุด เป็นต้น

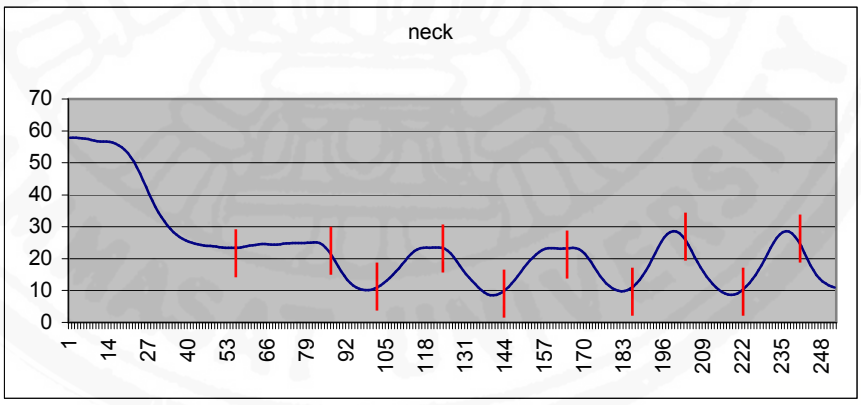
ในการเคลื่อนไหวของแต่ละจุดนั้น เราสามารถตรวจสอบการเคลื่อนไหวเหล่านี้เพื่อกำหนดให้เป็นจุดสนใจได้จาก

3.2.2.1 **จุดวกกลับของการเคลื่อนไหว** (Point of Returning)

ได้แก่จุดที่กราฟขึ้นไปจนสุดหรือลงมาจนสุดแล้ววกกลับไปอีกทางหนึ่ง หากเราเลือกกราฟมาเพียงเส้นเดียวหรือสามารถรวมจากทุกจุดเข้าด้วยกันแล้ว จึงได้ตำแหน่งของจุดสนใจดังภาพ 3.6

ภาพที่ 3.6

แสดงตำแหน่งของจุดสนใจ



3.2.2.2 **ความเร็วของการเคลื่อนไหว** (Velocity)

หรือความเร่งของการเคลื่อนไหว (Acceleration) หากเรานำเอาค่าระหว่างแต่ละจุดมาทำการหาความเร็วหรือความเร่ง หรืออาจเป็นค่าเฉลี่ย เราสามารถกำหนดถึงค่าใดๆ ที่เกิดขึ้นพร้อมๆกับจุดสนใจนั้นๆได้เช่นจากตัวอย่างด้านบน เราอาจหาค่าจากเฟรมที่ 57-85 หรือ 85-99 ได้โดยหาความเร็วเฉลี่ยจากพิกัด

ตารางที่ 3.1

ค่าของตำแหน่งต่างๆที่เกิดขึ้นในช่วงเวลา นำมาหาค่าเฉลี่ยของความเร็ว

Keyframe 57-85	
Position	*Velocity*
23.67892	
23.87441	0.210751
24.08516	0.111313
24.19648	0.106548
24.30302	0.170663
24.47369	0.13169
24.60538	-0.0014
24.60398	-0.10925
24.49473	-0.1146
24.38013	-0.02192
24.3582	0.081297
24.4395	0.112862
24.55236	0.114024
24.66639	0.091352
24.75774	0.073627
24.83137	0.050802
24.88217	0.027848
24.91001	0.010529
24.92054	-0.00192
24.91862	0.009219
24.92784	0.045024
24.97287	0.081031
25.0539	0.083936
25.13784	-0.00481
25.13302	-0.2501

Keyframe 85-99	
Position	*Velocity*
23.4461	
22.27668	-1.16942
20.86145	-1.41523
19.26584	-1.59561
17.62332	-1.64252
16.04677	-1.57655
14.62064	-1.42613
13.4013	-1.21934
12.37698	-1.02432
11.56584	-0.81114
10.95355	-0.61229
10.52773	-0.42582
10.28119	-0.24654
10.14624	-0.13495
10.13256	-0.01368
Ave. Vel.	**-0.95097**

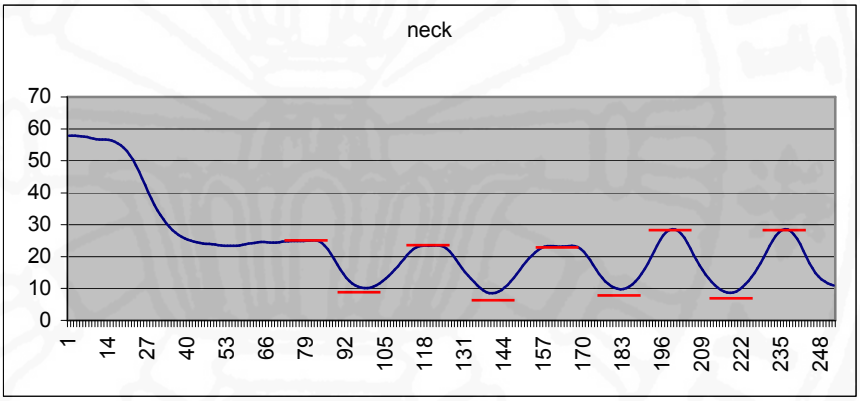
24.88292	-0.56373
24.31919	-0.87309
23.4461	-0.00862
Ave. Vel.	**0.195497**

3.2.2.3 **ความสูงของกราฟ** (Amplitude)

เราสามารถดึงเอาความสูงของกราฟมาโดยตรงเช่นกัน เพื่อหาค่าใดๆ ณ จุดสนใจนั้น

ภาพที่ 3.7

แสดงตำแหน่งของแอมปลิจูด

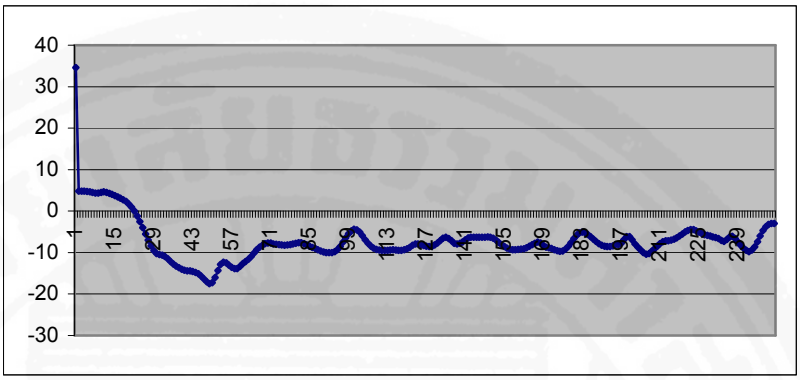


3.2.2.4 **ความห่างระหว่างจุด** (Distance)

ในการเคลื่อนไหวใดๆของมนุษย์ ส่วนปลายของร่างกายเช่นแขนหรือขา มักต้องเคลื่อนไหวมากกว่าส่วนอื่น เราอาจดึงเอาความห่างระหว่างแต่ละจุดมาสนใจ ยกตัวอย่างเช่น การเทียบระหว่างปลายแขนและจุดกึ่งกลางร่างกาย สามารถคำนวณผลต่างและนำเสนอออกมาดังกราฟ

ภาพที่ 3.8

แสดงตำแหน่งจากการหาค่าความห่างระหว่างจุด



3.2.3 ขั้นตอนการหาจุดข้อต่อที่สำคัญ

“แชนเนลไหนที่มีความสำคัญ?” ในการเคลื่อนไหวที่ต่างกัน แต่ละแชนเนลหรือกลุ่มของหลายๆแชนเนลรวมกันสามารถสะท้อนได้ถึงการเคลื่อนไหวหนึ่งๆ ยกตัวอย่างการเดินหรือการชกมวยนั้นเราต้องแยกออกมาว่ามือหรือเท้านั้นเป็นส่วนที่บ่งชี้ถึงการลักษณะเด่นของเคลื่อนไหวและสามารถบอกได้ว่าเป็นการเคลื่อนไหวอะไร

แต่ปัญหาที่เกิดขึ้นนั้น เราไม่สามารถเลือกออกมาเพียงจุดใดจุดหนึ่งเพื่อแสดงถึงการเคลื่อนไหวได้ จึงจำเป็นที่ต้องรวมข้อมูลเข้ากับแชนเนลอื่นๆด้วยโดยเราอาจลำดับขั้นตอนในการรวมให้แตกต่างกันออกไปเพื่อให้ได้ผลลัพธ์ออกมาเป็นที่น่าพอใจที่สุด โดยคุณสมบัติที่เราสามารถสลับขั้นตอนได้ ได้แก่

3.2.3.1 เริ่มจาก “พิกัดของ X และ Y”

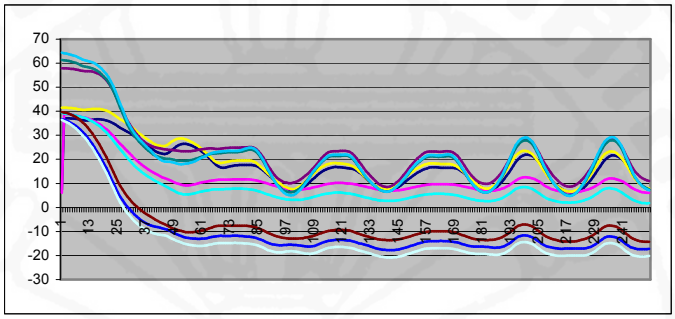
หากเราเริ่มรวมจากส่วนนี้ก่อน นั่นหมายความว่าในพิกัดของ X หรือ Y ถูกรวมกันก่อนเพื่อแสดงถึงพิกัดรวมค่าเดียวที่แสดงถึงลักษณะเด่นของการเคลื่อนไหวที่จุดนั้นๆ ซึ่งในบางครั้งการเคลื่อนไหวอาจเกิดขึ้นเพียงในแกนใดแกนหนึ่ง การที่เรารวมคุณสมบัตินี้ก่อนก็สามารถรองรับได้ในหลายๆการเคลื่อนไหว แต่ในทางกลับกันแล้วหากการรวมทั้งสองพิกัดเข้าด้วยกันแล้วให้ผลในทางหักล้างกัน เราจำเป็นต้องปรับค่าใดค่าหนึ่งไปในทางตรงข้ามแทนเพื่อให้ผลลัพท์ออกมาชัดเจนกว่า

3.2.3.2 เริ่มจาก “แชนเนลภายใต้ Parent เดียวกัน”

เพื่อให้การรวมเกิดขึ้นในทางที่สอดคล้องกัน หากพิจารณาจากลำดับชั้น แล้ว การรวมจากแชนเนลต่างๆใต้ Parent เดียวกันเช่น รวมจากสายของแขนหรือขาก่อนดังรูปที่แสดงการแยกกลุ่มของแชนเนลออกมา

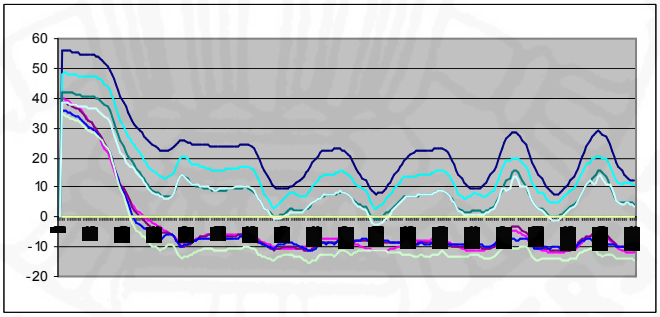
ภาพที่ 3.9

แชนเนลรวมส่วนแกนกลางลำตัว



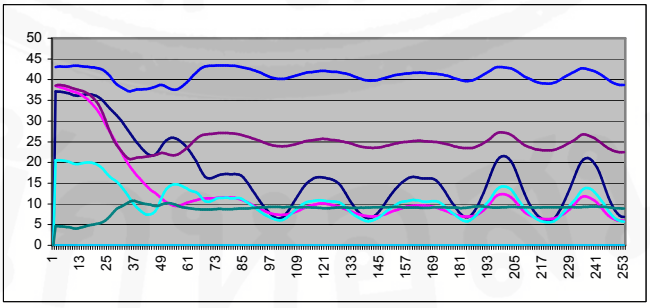
ภาพที่ 3.10

แชนเนลรวมส่วนแขนขวา



ภาพที่ 3.11

แชนเนลรวมส่วนขาขวา



3.2.3.3 เริ่มจาก “เลือกแชนเนลแบบไม่อิง Parent เดียวกัน”

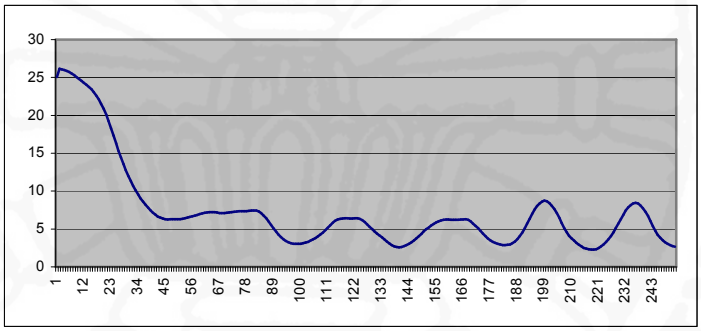
อาจให้ผลแตกต่างออกไป จากการที่แต่ละแชนเนลที่เกิดขึ้นขาดความสอดคล้องกันอย่างสิ้นเชิง

3.2.3.4 เริ่มจาก “รวมทุกแชนเนลเข้าพร้อม ๆกัน”

อาจเป็นการง่ายที่สุดหากรวมทุก แชนเนลข้าด้วยกันทั้งหมดแล้ว ข้อมูลที่เกิดขึ้นนั้นโดดเด่นเพียงพอหรือหมายความว่าสามารถบ่งชี้จุดสำคัญๆของการเคลื่อนไหวได้ จากภาพแสดงถึงการรวมโดยใช้การหาค่าเฉลี่ย

ภาพที่ 3.12

การรวมโดยใช้การหาค่าเฉลี่ย



3.2.3.5 ลำดับในการรวม

เมื่อถึงจุดนี้เห็นได้ว่าจากตัวอย่างที่ยกมานั้น การคำนวณหาจุดสนใจสำหรับการเคลื่อนไหววิดพื้นกันนั้นให้ผลที่ใกล้เคียงกันมาก ไม่ว่าเป็นการดึงออกมาเพียงจุดเดียวหรือทั้งหมด แต่หากเราต้องการทำให้ผลลัพธ์ที่ได้มีการเปลี่ยนแปลงผิดเพี้ยน (Distortion) ไปจากเดิมนั้น เราสามารถนำคุณสมบัติที่กล่าวมานี้มาสลับขั้นตอนกัน ยกตัวอย่างเช่น

- A->B->D รวมพิกัด X และ Y ก่อนแล้วจึงรวมแชนเนลใต้ Parent เดียวกัน และนำแต่ละกลุ่มมารวมกันท้ายสุด
- D->A รวมทุกแชนเนลเป็นอย่างแรกโดยแยกระหว่าง X และ Y และนำทั้งคู่นั้นมารวมกันท้ายสุด
- C->A->D สุ่มเลือกแชนเนลมาต่างๆกันไป โดยแยกเป็นกลุ่มของ X และ Y จากนั้นจึงค่อยนำมารวมกัน

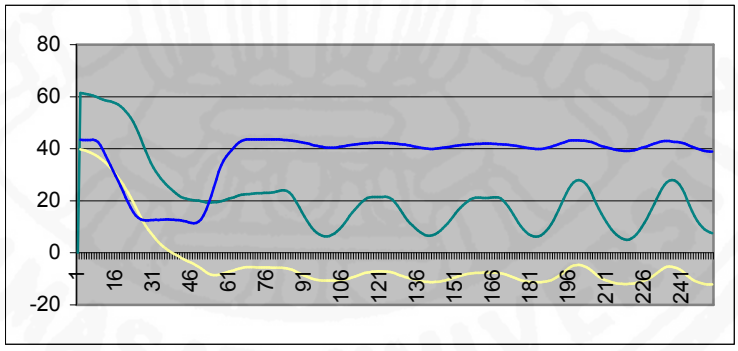
3.2.3.6 **ฟังก์ชั่นที่ใช้รวม**

นอกจากการเปลี่ยนแปลงถึงขั้นตอนแล้ว เรายังสามารถปรับฟังก์ชั่นเข้าไปใช้ในแต่ละขั้นตอนของการรวมก่อนส่งต่อไปถึงขั้นต่อไปได้อีก โดยฟังก์ชั่นที่ใช้รวมนั้นสามารถกำหนดได้ดังนี้

- คำนวณหาค่าเฉลี่ย (mean) ซึ่งเป็นพื้นฐานที่สุดในการหาค่ากลางระหว่างสองค่าขึ้นไป
- คำนวณหาค่ามาก (frequently) เราอาจสนใจแต่เฉพาะค่าที่เกิดขึ้นบ่อยๆ เท่านั้นโดยที่อาจตัดค่าที่เกิดขึ้นเป็นครั้งคราวทิ้งออกไป โดยอาจกำหนดเป็นช่วงๆของค่าเป็นต้น ยกตัวอย่างเช่นในการรวมทุกแซนเนลในครั้งเดียวนั้น เราอาจกรองเอาแต่ค่าที่เกิดขึ้นบ่อยสุดหากเปรียบเทียบกันในแต่ละแซนเนลเท่านั้น อาจเป็น 3 อันดับแรกออกมาไม่ใช่รวมสรุปออกมาเป็นเพียงค่าเดียว จากรูปนั้นเลือกออกมาแค่ 3 ค่าเท่านั้นโดยอาจกำหนดเป็นช่วงๆไปแล้วแต่ขอบเขตของข้อมูล

ภาพที่ 3.13

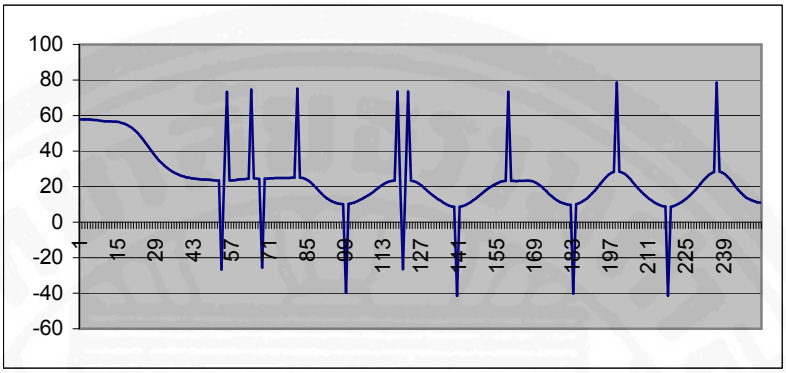
การรวมโดยเลือกเพียง 3 กลุ่ม



- เพิ่มความชัดเจน (Stress) ของจุดสนใจเพื่อเพิ่มความชัดเจนของการเคลื่อนไหว เราอาจให้น้ำหนัก (Weight) กับค่าๆหนึ่งให้โดดเด่นออกมาเลย เพื่อให้ง่ายต่อการค้นหา ยกตัวอย่างเช่น หากเราเน้นถึงจุดวกกลับของการเคลื่อนไหว (Point of Returning) เป็นพิเศษ ก็สามารถ ให้น้ำหนักกับค่านั้นๆให้ชัดเจนดังรูป ได้ให้น้ำหนักเพิ่มไปอีก 50 สำหรับจุดวกกลับเป็นต้น นอกจากนั้นหากเราเพิ่มเติมการเน้นไปอีกในช่วงต่างๆระหว่างจุดวกกลับนั้น เช่นจุดที่มีความเร่งชัดเจนเป็นต้น ยิ่งเพิ่มเติมส่วนที่บ่งชี้ถึงจุดสนใจ ได้อีกหากแต่ว่าการพิจารณาถึงความหนาแน่นก็เป็นสิ่งจำเป็น เราอาจกำหนดขอบเขต Tolerance ไว้ไม่ให้ถี่เกินไป

ภาพที่ 3.14

การเพิ่มความชัดเจน



3.3 ขั้นตอนการประพันธ์ดนตรี (Music Composition)

การประพันธ์ดนตรีในวิทยานิพนธ์นี้คือกระบวนการในการสร้างเอาท์พุทออกมาจากระบบ วิธีในการประพันธ์ดนตรีขึ้นอยู่กับวิธีของแต่ละบุคคล โดยพื้นฐานแล้วดนตรีประกอบด้วย จังหวะ (Rhythm) ทำนอง (Melody) และเสียงประสาน (Harmony) ในการสร้างจังหวะนั้น เราสามารถเชื่อมต่อกับจังหวะของการเคลื่อนไหวได้ทันที เมื่อเชื่อมต่อแล้วหากเราต้องการดนตรีที่มีจังหวะเร็วขึ้นก็เพียงแต่เพิ่มโน้ตเข้าไปเท่านั้น ขั้นตอนในการสร้างดนตรีเริ่มขึ้นได้หลายวิธีไม่มีกฎตายตัว แต่ในวิทยานิพนธ์ชิ้นนี้หากต้องการให้ดนตรีที่ได้มีความสอดคล้องกับการเคลื่อนไหว สิ่งแรกที่ต้องการก็คือเริ่มต้นจากจังหวะ

3.3.1 กำหนดตำแหน่งของโน้ตที่เกิดขึ้น (Timing) โดยที่ตำแหน่งนี้หมายถึงตำแหน่งของโน้ตตัวหลัก (Leading Note) ที่ทำการเชื่อมต่อที่เวลาเดียวกันกับการเคลื่อนไหวที่ได้ทำการกรองออกมา รวมทั้งกำหนดค่าของความถี่เสียง (Pitch) ความยาวของโน้ต (Duration) และความหนักแน่นของโน้ต (Velocity) ลงไปด้วย

3.3.2 เมื่อกำหนดตัวโน้ตในทุกๆตำแหน่งที่ทำการเชื่อมต่อกับการเคลื่อนไหวแล้ว จึงได้โน้ตจำนวนหนึ่งมาใช้ในการคำนวณถึงคีย์หลัก (Key Signature) ที่ใช้บังคับโน้ตที่ปรากฏ ซึ่งมีประโยชน์ในการกำหนดโน้ตที่ใส่ลงไปเพิ่มเติมในภายหลัง (Passing Note) ที่ใช้ผ่านไปสู่โน้ตตัวหลัก (Leading Note)

3.3.3 ส่วนที่เป็นโน้ตที่ใส่ลงไปเพิ่มเติมนั้น หากเราใช้คีย์หลักที่เรามีอยู่เป็นหลัก โน้ตที่เพิ่มเติมขึ้นมาสามารถสร้างขึ้นโดยทฤษฎีทางของโน้ตที่เป็นไปได้ (Scale) หรือการที่กำหนดว่าเริ่มเล่นตัวไหน (Mode)

3.3.4 จากหลักการในสามข้อแรกนั้นสามารถใช้แตกต่างกันตามชนิดของเครื่องดนตรี เราอาจแบ่งชนิดของเครื่องดนตรีตามเงื่อนไขของการสร้างโน้ตได้ดังต่อไปนี้

a. เครื่องดนตรีที่สร้างโน้ตทีละตัวในหนึ่งช่วงเวลา เช่น เบส แซคโซโฟน ซึ่งสามารถสร้างได้จากหลักการในสามข้อแรก

b. เครื่องดนตรีที่สร้างโน้ตหลายตัวในหนึ่งช่วงเวลา เช่น กลอง กีต้าร์ เปียโน ซึ่งต้องการเสียงประสานเพิ่ม จึงจำเป็นต้องมีตัวโน้ตที่สามารถเล่นพร้อมๆกันได้ โดยการที่เราเลือกเล่นโน้ตตั้งแต่สองตัวขึ้นไปในสเกลหนึ่งๆเรียกว่าคอร์ด (Chord)

3.3.5 เราสามารถกำหนดความหนาแน่นของโน้ต (Density) ให้มากหรือน้อยได้เช่นเดียวกัน ซึ่งส่งผลให้ความเร็วหรืออารมณ์เปลี่ยนไปเป็นอย่างมาก นอกเหนือจากนั้นในจุดที่การเคลื่อนไหวมีอารมณ์ที่ชัดเจนที่สุดนั้นยังสามารถเชื่อมต่อมาเป็นจุด (Peak) ของเพลงได้ โดยเราพิจารณาว่าดนตรีเปรียบได้กับการเล่าเรื่อง ย่อมต้องมีส่วนที่ เริ่มเรื่อง ตื่นเต้น และบทสรุป

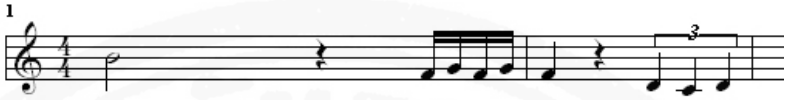
3.4 ขั้นตอนการเชื่อมต่อระหว่างการเคลื่อนไหวและดนตรี (Motion and Music -Synchronization)

จุดสนใจที่ได้มาจากการจับความเคลื่อนไหวนั้นอาจเปรียบได้กับจังหวะที่เกิดขึ้นในดนตรีหากเราผนวกจุดนั้นเข้ากับค่าใดๆที่เกิดขึ้น ณ จุดนั้นหรือจากรอบๆจุดนั้นด้วย เช่นค่าของความเร่ง หรือระยะห่างจากจุดศูนย์กลาง เราสามารถแมพเข้าเป็นตัวโน้ตสำหรับเอาท์พุทที่เป็นดนตรีได้ทันที ยกตัวอย่างเช่นเมื่อได้ตำแหน่งพร้อมค่าใดๆต่อไปนี้ ซึ่งสามารถแสดงออกมาในรูปแบบของโน้ตดนตรีได้ดังภาพ 3.15

ตารางที่ 3.2
พารามิเตอร์ที่นำมาเชื่อมกันระหว่างการเคลื่อนไหวและดนตรี

Mapping	*From Animation*	*Spotting Session*	*Amplitude*	*Velocity*
	To Music	*Timing*	*Pitch*	*Duration*
		1	B	2
		4	F	¼
		5	F	1
		7	D	1/3

ภาพที่ 3.15

โน้ตที่ได้มาจากการเชื่อมต่อ

โดยระดับล่างสุดของดนตรีที่ประกอบด้วย จังหวะ, ทำนอง และเสียงประสาน ยังสามารถขยายออกไปเป็นดนตรีระดับที่สูงขึ้นอีกโดยเราต้องอาศัยทฤษฎีทางดนตรีนำมาประกอบกับอินพุทจากการจับความเคลื่อนไหวที่เรามีอยู่ โดยโมเดลต่อไปนี้คือระดับของการพัฒนาให้ดนตรีมีความซับซ้อนขึ้นไปอีกจากการสร้างโดยวิธีการคำนวณทางคอมพิวเตอร์ กล่าวโดยคร่าวๆ ดังต่อไปนี้

ตารางที่ 3.3

ระดับขั้นในการเพิ่มความซับซ้อนของดนตรี

	ระดับ	*ส่วนที่เกี่ยวข้อง*	*คำอธิบาย*
ส่วนที่เป็นเนื้อหาของดนตรี	Lv. 7	Artist	ระดับที่มีรูปแบบเป็นตัวของตัวเอง แตกต่างอย่างชัดเจน
	Lv. 6	Song Part	การกำหนดสิ่งต่างๆข้างใต้นี้ในรูปแบบเช่นการเล่าเรื่อง มีเริ่ม มีจุดที่น่าจดจำ มีบทสรุป
	Lv. 5	Motif	รูปแบบการเล่นเฉพาะตัว
	Lv. 4	Peak	จุดที่สำคัญที่สุดในเพลงหรือส่วนช่วงหนึ่งของเพลง
ส่วนที่เป็นเนื้อหาของดนตรีที่ใช้ในการเชื่อมต่อ	Lv. 3	Chord Progression	รูปแบบขั้นตอนของเสียงที่เล่นประสานขึ้นในเวลาเดียวกัน
	Lv. 2	Key Signature, Scale, Mode, Chord	เป็นนิยามในระดับที่สูงขึ้น ที่บ่งบอกควบคุมถึงกลุ่มของระดับเสียงที่สามารถเกิดขึ้นได้ โดยประกอบด้วย ระดับเสียงต่างๆที่ใช้ได้ในรูปแบบเดียวกัน ระดับเสียงต่างๆที่ถูกบังคับให้ใช้ในรูปแบบที่กำหนด เสียงที่ต้องการขึ้นต้นในรูปแบบนั้น และเสียงที่เล่นประสานขึ้นในเวลาเดียวกัน
	Lv. 1	Pitch, Duration, Timing, Velocity	เป็นรูปแบบข้อมูลซึ่งนำเสนอดนตรีในหน่วยพื้นฐานที่สุด ประกอบด้วย ความถี่ของเสียงที่บ่งบอกถึงระดับเสียงทางดนตรี, เวลาที่เสียงนั้นคงค้างอยู่, เวลาที่มีเสียงนั้นเกิดขึ้น และความดังหรือหนักแน่นของเสียงนั้นๆ

หากมองจากลักษณะที่กำหนดไว้ในโน้ตดนตรีแล้ว เราระบุสิ่งต่างๆที่กล่าวไว้ข้างต้นได้ดังต่อไปนี้ [D99]

ภาพที่ 3.15

แสดงโน้ตดนตรี และอธิบายส่วนประกอบที่นำมาใช้