สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

ระบบช่วยฝึกสอนการออกกำลังกายโดยการประยุกต์ใช้เทคนิคการตรวจจับการเคลื่อนไหวของมนุษย์จากวีดิทัศน์

EXERCISE TRAINING SYSTEM USING VISION-BASED MOTION CAPTURE

โดย

ชานน ศรีวิเศษสม

ธนศร คันธวัฒน์

อาจารย์ที่ปรึกษา

ผศ.ดร.ธนารัตน์ ชลิดาพงศ์

เลขหมู่.....
เลขทะเบียน 73073
วัน,เดือน,ปี - 2 ก.ค. 2550

b 11779937
i.....

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคการศึกษาที่ 2 ปีการศึกษา 2549

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

EXERCISE TRAINING SYSTEM USING VISION-BASED MOTION CAPTURE

CHANON SRIVESESSOM

THANASORN KANTAWAT

A PROJECT SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR THE DEGREE OF

BACHELOR OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY

FACULTY OF INFORMATION TECNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2/2006

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2007

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใบรับรองปริญญานิพนธ์ ประจำปีการศึกษา 2549

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง ระบบช่วยฝึกสอนการออกกำลังกายโดยการประยุกต์ใช้เทคนิคการตรวจจับการเคลื่อนไหวของมนุษย์จากวีดิทัศน์

Exercise Training System using Vision-based Motion Capture

ผู้จัดทำ

1. นายชานน ศรีวิเศษสม รหัสประจำตัว 46060010

2. นายธนศร คันธวัฒน์ รหัสประจำตัว 46000016

.....อาจารย์ที่ปรึกษา

(ผศ. ดร.ธนารัตน์ ชลิดาพงศ์)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อ	ระบบช่วยฝึกสอนการออกกำลังกายโดยการประยุกต์ใช้เทคนิคการตรวจจับ การเคลื่อนไหวของมนุษย์จากวีดิทัศน์
นักศึกษา	นายชานน ศรีวิเศษสม รหัสนักศึกษา 46060010 นายธนศร คันธวัฒน์ รหัสนักศึกษา 46060016
ปริญญา	วิทยาศาสตรบัณฑิต (เทคโนโลยีสารสนเทศ)
สาขาวิชา	เทคโนโลยีสารสนเทศ
ปีการศึกษา	2549
อาจารย์ที่ปรึกษา	ผศ.ดร.ธนารัตน์ ชลิดาพงศ์

บทคัดย่อ

ระบบช่วยฝึกการออกกำลังกายเบื้องต้นโดยการประยุกต์ใช้เทคนิคการตรวจจับ การเคลื่อนไหวของมนุษย์จากวีดิทัศน์เป็นระบบที่นำเอาทฤษฎีซึ่งเกี่ยวข้องกับคอมพิวเตอร์วิชั่นมาประยุกต์ใช้ โดยแนวคิดที่จะพัฒนาระบบนี้เกิดขึ้นเนื่องจากได้พบว่าระบบหรือสื่อในระบบช่วยฝึกการออกกำลังกายในปัจจุบันนั้นไม่สามารถโต้ตอบกับผู้ใช้ได้ ถึงแม้ว่าในบางสื่อจะมีการพูดคุยกับผู้ใช้แต่สื่อช่วยสอนในการทำงานใช้การคาดเดาว่าผู้ใช้จะทำตามที่ผู้ฝึกสอนแนะนำซึ่งในความเป็นจริงแล้วผู้ใช้อาจจะไม่ได้ทำตามหรือทำท่าทางที่ไม่ตรงกับที่ผู้สอนสอนอยู่ จากเหตุผลดังกล่าวจึงมีแนวคิดที่จะพัฒนาระบบที่สามารถโต้ตอบกับผู้ใช้ได้จริง โดยนำเอาแนวคิดของการวิเคราะห์การเคลื่อนไหวมาใช้ร่วมกันกับการฝึกการออกกำลังกายโดยใช้วีดิทัศน์จับภาพผู้ใช้และนำข้อมูลที่ได้มาวิเคราะห์ว่าการเคลื่อนไหวของผู้ใช้เพื่อตรวจสอบความถูกต้องของท่าทางของผู้ใช้ โดยนำผลที่ได้จากการวิเคราะห์ไปสร้างการโต้ตอบกับผู้ใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Title	Exercise Training System using Vision-based Motion Capture
Student	Mr. CHANON SRIVESESSOM Student ID.46060010
	Mr. THANASORN KANTAWAT Student ID.46060016
Degree	Bachelor of Science (Information Technology)
Programme	Information Technology
Academic Year	2006
Advisor	Asst. Prof. Dr. Thanarat Chalidabhongse

ABSTRACT

This report describes a development of Exercise Training System that can detect and track user's motion from video. The system is developed by employing computer vision technology and techniques. The motivation behind this project is that the existing media and systems for exercise training typically are one-way communication. The trainer instructs the trainee and never knows whether the trainee follows the instructions. With this reason, we have designed an exercise training system that can interact with the trainee or user. To do this, we employ image processing and computer vision techniques to capture user's motion from web-cam that is connected to the PC running our system. Then, the system compares the user's motion with the animated trainer. The motion comparison result is used to control the sequences of the animated instruction.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

โครงงานนี้ประสบความสำเร็จลุล่วงไปได้ด้วยดี ด้วยความช่วยเหลือจากบุคคลหลายท่าน โดยเฉพาะความกรุณาจากอาจารย์ที่ปรึกษาโครงงาน ผศ.ดร.ธนารัตน์ ชลิดาพงศ์ ผู้ซึ่งช่วยเหลือให้คำแนะนำ ชี้แนะแนวทางที่เป็นประโยชน์ในทุกๆด้านในโครงงานนี้

ขอขอบคุณบิดา มารดา และครอบครัว ซึ่งเป็นที่รักและเคารพยิ่ง ที่คอยให้กำลังใจ คอยเป็นห่วงเป็นใยอยู่ตลอดเวลา และให้ความอนุเคราะห์ทางด้านทุนทรัพย์

ขอขอบคุณ คุณดนุวัฒน์ แสวงผล(พี่เด่น) พี่แบงค์ พี่วามม์ พี่สน รวมถึงทุกคนใน PIC LAB ที่คอยให้คำปรึกษาและให้ข้อแนะนำต่างๆ เกี่ยวกับการใช้งานทางด้าน Image รวมทั้งเพื่อนๆทุกคน ที่คอยให้กำลังใจคณะผู้จัดทำในการพัฒนาโครงงานเสมอมา

สุดท้ายนี้ขอขอบคุณอาจารย์ทุกท่านที่ได้ให้วิชาความรู้และถ่ายทอดประสบการณ์ที่ดีให้

ชานน ศรีวิเศษสม

ธนศร คันธวัฒน์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....I

บทคัดย่อภาษาอังกฤษ.....II

กิตติกรรมประกาศ.....III

สารบัญ.....IV

สารบัญภาพ.....VII

บทที่ 1 บทนำ..... 1

1.1 ความสำคัญและความเป็นมาของปัญหา.....1

1.2 วัตถุประสงค์ของการพัฒนาระบบงาน.....1

1.3 แผนการดำเนินงาน.....2

1.4 ขอบเขตของการพัฒนาระบบงาน.....2

1.5 ประโยชน์ที่ได้รับ.....2

บทที่ 2 หลักการณ์และทฤษฎีที่เกี่ยวข้อง.....3

2.1 การตรวจจับและติดตามการเคลื่อนไหว.....3

2.2 การแยกส่วนที่สนใจออกจากสิ่งแวดล้อม.....3

2.3 การทำเครื่องหมายส่วนของร่างกาย.....5

2.3.1 การตรวจจับส่วนของร่างกายโดยใช้โครงร่าง (Silhouette)..... 5

2.3.1.1 การระบุท่าทางต่างๆ.....5

2.3.1.2 แบบจำลองร่างกาย 2 มิติ (2D Body Model).....6

2.3.1.3 การให้เครื่องหมายส่วนของร่างกาย (Body Part Labeling)..... 8

2.3.2 การตรวจจับและติดตามส่วนของร่างกายโดยใช้ Cardboard Model ใน W4.....9

2.3.2.1 การที่วัตถุปรากฏขึ้น (Appearing Objects)..... 10

2.3.2.2 การติดตามการเคลื่อนไหว (Tracking)..... 10

2.3.2.3 การแบ่งแยกบริเวณ (Region splitting)..... 11

2.3.2.4 การรวมบริเวณ (Region merging).....12

2.3.2.5 การติดตามการเคลื่อนไหวของส่วนต่างๆของร่างกาย.....13

2.4 พื้นฐานการประมวลผลภาพ..... 15

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.4.1 มอร์โฟโลยี (Morphology)..... 15

2.4.1.1 การทำไดเลชัน (Dilation)..... 15

2.4.1.2 การทำอีโรชัน (Erosion)..... 16

2.4.1.3 การทำโอเพนนิง (Opening)..... 17

2.4.1.4 การทำโคลสซิง (Closing)..... 17

2.4.2 การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน (Connected Component Analysis)..... 18

2.4.3 การหาเส้นรอบรูป (Contour Finding)..... 19

2.5 การเขียนโปรแกรมเชิงวิชั่น (Vision Programming)..... 19

2.5.1 หลักการเขียนภาษาเชิงวัตถุ และการสร้างโปรแกรมด้วย MS Visual C++ และ MFC..... 19

2.5.2 หลักการทำงานของ OpenCV..... 22

2.5.3 การเลือกภาษาในการพัฒนาโปรแกรม..... 23

2.5.4 ความสัมพันธ์ระหว่าง OpenCV กับ ไลบรารีอื่นๆ..... 23

2.6 แอนนิเมชั่น (Animation)..... 23

บทที่ 3 การออกแบบ และการพัฒนาระบบ..... 25

3.1 ความต้องการระบบ..... 25

3.2 การออกแบบระบบ..... 25

3.3 การสร้างระบบ..... 26

3.3.1 เครื่องมือที่ใช้..... 26

3.3.2 ด้านการทำแอนนิเมชั่น..... 27

3.3.3 การทำการประมวลผลภาพ..... 29

บทที่ 4 วิธีทดลองและผลการทดลอง..... 41

4.1 อุปกรณ์ที่ใช้ในการทดลอง..... 41

4.2 ขั้นตอนการทดลอง..... 41

4.2.1 การทดลองการตัดแยกผู้ใช้ออกจากฉากหลัง..... 41

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

4.2.1 การทดลองการตัดแยกผู้ใช้ออกจากฉากหลัง..... 42
4.2.2 การทดลองการระบุตำแหน่งส่วนของร่างกาย.....42
4.2.3 การวิเคราะห์ท่าทาง.....43
4.3 สรุปผลการทดลอง..... 47

บทที่ 5 สรุปผลโครงงานและข้อเสนอแนะ.....49
5.1 สรุปผลโครงงาน.....49
5.2 ปัญหาและข้อจำกัดในการทำงานของระบบ.....49
5.3 ข้อเสนอแนะในการพัฒนาระบบต่อไป..... 50

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

ภาพที่ หน้า

2.1 แสดงการลบฉากหลังออกจากภาพ..... 4

2.2 แสดงท่าทางต่างๆเมื่อแสดงให้เห็นมุมมองที่ต่างกัน..... 5

2.3 ทำการแสดง Sequence ของการเคลื่อนไหว..... 6

2.4 การแสดงลำดับของตำแหน่งของส่วนต่างๆ..... 7

2.5 แสดงท่ายืน Input Image (a) convex-hull vertices (b) เส้นเชื่อมระหว่าง Convex-hull vertices (c) convex และ concave hull vertices (d) และ เส้นเชื่อมจุด Convex และ concave hull vertices (e)..... 8

2.6 การระบุส่วนของร่างกาย (ส่วนหัว)..... 8

2.7 แสดงการคาดเดาตำแหน่งของศีรษะ ด้วยวิธี Ghost 9

2.8 การคาดเดาการเคลื่อนไหวของร่างกายโดยใช้โครงร่างจับคู่กันระหว่าง 1 เฟรมที่เลือกมา ภาพที่เข้ามา (a) การตรวจจับโครงร่าง (b) alignment ของขอบโครงร่างโดยใช้ความแตกต่าง ของ Median (c) alignment สุดท้ายหลังจากทำความสัมพันธ์โครงร่าง (d)..... 10

2.9 Cardboard model ที่ใช้ใน W4 (a) และ แบบจำลองที่เคลื่อนไหวที่ใช้สำหรับศีรษะ (b) และ มือ(c)..... 11

2.10 เป็น โครงร่างอย่างการใช้ Cardboard Model แสดงการบอกตำแหน่งส่วนต่างๆ ของร่างกาย (a) แสดงค่าเริ่มต้นของกรอบบอกตำแหน่งขอบเขตของโครงร่าง (b) ใช้ Cardboard Model วิเคราะห์ตำแหน่งของส่วนของร่างกาย (c) แสดงการระบุตำแหน่งของ ร่างกาย (ศีรษะ, ลำตัว, มือทั้งสองข้าง และ เท้าทั้งสองข้าง)..... 14

2.11 แสดงอัตราส่วนในการแยกส่วนประกอบของร่างกาย..... 14

2.12 แสดงการทำไดเลชัน 16

2.13 แสดงการทำอีโรชัน..... 17

2.14 แสดงการทำโอเพนนิง..... 17

2.15 แสดงการทำโคลสซิง..... 18

2.16 การกำหนดรหัสแบบ (ก) 4 จุดเชื่อมต่อ (ข) 8 จุดเชื่อมต่อ..... 18

2.17 แสดงการเข้ารหัสลูกโซ่แบบ 4 จุดและการแสดงการเข้ารหัส..... 19

3.1 แสดงโครงสร้างของโครงงาน..... 25

3.2 แผนภาพแสดงลักษณะการทำงานของระบบ..... 26

3.3 แสดงฉากของแอนนิเมชั่น..... 27

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ (ต่อ)

ภาพที่ หน้า

3.4 แสดงภาพตัวการ์ตูนแอนนิเมชั่นท่ากางแขนโดยแยกเป็นวัตถุ....28

3.5 แสดงภาพตัวการ์ตูนแอนนิเมชั่นท่างอข้อศอก....28

3.6 แสดงภาพตัวการ์ตูนแอนนิเมชั่นท่าย่อตัว....28

3.7 แสดงภาพการ์ตูนแอนนิเมชั่นพร้อมฉากหลัง....29

3.8 ฉากหลังที่ได้จากการสร้างฉากหลัง....29

3.9 การตัดแยกผู้ใช้ออกจากฉากหลัง (a) แสดงข้อมูลที่ได้จากวีดีทัศน์ (b) แสดงผลลัพธ์ที่ได้จากการตัดแยกผู้ใช้ออกจากฉากหลัง....30

3.10 แสดงกระบวนการตัดแยกผู้ใช้ออกจากฉากหลัง....30

3.11 การกำจัดสัญญาณรบกวน (a) แสดงภาพก่อนการกำจัดสัญญาณรบกวน (b) แสดงผลลัพธ์ที่ได้จากการกำจัดสัญญาณรบกวน....31

3.12 การหาองค์ประกอบภายในภาพ (a) แสดงภาพก่อนการหาองค์ประกอบภายในภาพ (b) แสดงผลลัพธ์ที่ได้ จากการหาองค์ประกอบภายในภาพ....31

3.13 การหาเส้นรอบรูปภายในภาพ (a) แสดงภาพก่อนการหาเส้นรอบรูปภายในภาพ (b) แสดงผลลัพธ์ที่ได้ จากการหาเส้นรอบรูปภายในภาพ....32

3.14 แสดงการระบุตำแหน่งส่วนของร่างกาย....33

3.15 แสดงขั้นตอนการวิเคราะห์การเคลื่อนไหว....33

3.16 แสดงกราฟข้อมูลตำแหน่งของแขน....34

3.17 แสดงกราฟที่ผ่านการทำการปรับค่าข้อมูล....35

3.18 แสดงกราฟที่ผ่านการหาค่าความแตกต่าง....35

3.19 แสดงกราฟที่ผ่านการหาค่าความแตกต่างโดยผ่านการหาค่าขีดแบ่ง....36

3.20 แสดงทิศทางการเคลื่อนไหวที่ใช้ในการเปรียบเทียบ....36

3.21 แสดงผลที่ได้จากการทำงานในแต่ละขั้นตอน....37

3.22 หน้าจอเริ่มต้นโปรแกรม....38

3.23 หน้าจอแสดงการสร้างฉากหลัง....39

3.24 หน้าจอแสดงการทำงานของโปรแกรม....40

4.1 ผลการทดลองการการตัดแยกผู้ใช้ออกจากฉากหลัง....42

4.2 ผลการทดลองการการตัดแยกผู้ใช้ออกจากฉากหลังเมื่อลดค่าความสว่างลง....42

4.3 ผลการทดลองระบุตำแหน่งส่วนของร่างกายที่ไม่เกิดสัญญาณรบกวน....43

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ (ต่อ)

ภาพที่ หน้า

4.4 ผลการทดลองระบุตำแหน่งส่วนของร่างกายที่มีสัญญาณรบกวนเกิดขึ้น..... 43
4.5 แสดงการ์ตูนแอนนิเมชั่นท่ายกแขน2ข้าง..... 44
4.6 แสดงการวิเคราะห์ท่าทางการออกกำลังกายท่ายกแขน2ข้าง..... 44
4.7 แสดงการ์ตูนแอนนิเมชั่นท่ายกแขนซ้าย..... 45
4.8 แสดงการวิเคราะห์ท่าทางการออกกำลังกายท่ายกแขนซ้าย..... 45
4.9 แสดงการ์ตูนแอนนิเมชั่นท่ายกแขนขวา..... 46
4.10 แสดงการวิเคราะห์ท่าทางการออกกำลังกายท่ายกแขนขวา..... 47
4.11 แสดงการ์ตูนแอนนิเมชั่นท่าย่อเข่า..... 47
4.12 แสดงการวิเคราะห์ท่าทางการออกกำลังกายท่าย่อเข่า..... 47

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความสำคัญและความเป็นมาของปัญหา

ในโครงงานนี้เป็นการออกแบบและพัฒนาระบบช่วยฝึกการออกกำลังกายเบื้องต้นโดยการประยุกต์ใช้เทคนิคการตรวจจับการเคลื่อนไหวของมนุษย์จากวีดิทัศน์ เป็นระบบที่นำเอาทฤษฎีซึ่งเกี่ยวข้องกับคอมพิวเตอร์วิชั่นมาประยุกต์ใช้ โดยแนวคิดที่จะพัฒนาระบบนี้เกิดขึ้นเนื่องจากได้พบว่า ระบบหรือสื่อในระบบช่วยฝึกการออกกำลังกายในปัจจุบันนั้นไม่สามารถโต้ตอบกับผู้ใช้ได้ ถึงแม้ว่าในบางสื่อจะมีการพูดคุยกับผู้ใช้ แต่สื่อช่วยสอนในการทำงานใช้การคาดเดาว่าผู้ใช้จะทำตามที่ผู้ฝึกสอนแนะนำ ซึ่งในความเป็นจริงแล้วผู้ใช้อาจจะไม่ได้ทำตามหรือทำท่าทางที่ไม่ตรงกับที่ผู้สอนสอนอยู่ จากเหตุผลดังกล่าวจึงมีแนวคิดที่จะพัฒนาระบบที่สามารถโต้ตอบกับผู้ใช้ได้จริง โดยนำเอาแนวคิดของการวิเคราะห์การเคลื่อนไหวมาใช้ร่วมกันกับการฝึกการออกกำลังกายโดยใช้วีดิทัศน์จับภาพผู้ใช้และนำข้อมูลที่ได้มาวิเคราะห์ว่าการเคลื่อนไหวของผู้ใช้เพื่อตรวจสอบความถูกต้องของท่าทางของผู้ใช้ โดยนำผลที่ได้จากการวิเคราะห์ไปสร้างการโต้ตอบกับผู้ใช้ เช่น บอกชมเมื่อทำท่าทางถูกต้อง และเปลี่ยนท่าไปทำท่าอื่นแทน ซึ่งจะทำให้ระบบสามารถวิเคราะห์ท่าทางของผู้ใช้ว่าทำท่าทางเดียวกันกับผู้ฝึกสอนหรือไม่ และสามารถโต้ตอบกับผู้ใช้ได้โดยใช้การ์ตูนในการโต้ตอบกับผู้ใช้ ซึ่งจะช่วยให้ผู้ใช้เกิดความเพลิดเพลินและปฏิบัติท่าทางได้ถูกต้องตามผู้ฝึกสอนซึ่งจะมีการสอนการออกกำลังกายเบื้องต้น

1.2 วัตถุประสงค์ของการพัฒนาระบบงาน

ในการพัฒนาการวิเคราะห์และออกแบบระบบช่วยฝึกการออกกำลังกายเบื้องต้นโดยการประยุกต์ใช้เทคนิคการตรวจจับการเคลื่อนไหวของมนุษย์จากวีดิทัศน์มีวัตถุประสงค์เพื่อ

1. เพื่อศึกษาหลักการและทฤษฎีทางด้านคอมพิวเตอร์วิชั่นในเรื่องเกี่ยวกับจับภาพการเคลื่อนไหว การแยกส่วนและการวิเคราะห์หาส่วนต่างๆของร่างกายออกจากวีดิทัศน์

2. เพื่อศึกษาเครื่องมือที่ใช้ในการพัฒนาระบบคอมพิวเตอร์วิชัน เช่น C/C++ Programming, OpenCV เป็นต้น

3. เพื่อสร้างโปรแกรมประยุกต์ที่มีความสามารถในการตรวจจับและวิเคราะห์ท่าทางของผู้ใช้งานระบบช่วยฝึกการออกกำลังกาย โดยโต้ตอบกับผู้ใช้ระบบได้แบบเวลาจริง (Real-Time)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3 แผนการดำเนินงาน

ภาคเรียนที่ 1

- ศึกษาทฤษฎีและหลักการของการจับภาพเคลื่อนไหว (Motion Capture)
- ศึกษาทฤษฎีและหลักการของการแยกส่วนที่สนใจออกจากสิ่งแวดล้อม (Segmentation)
- ศึกษาทฤษฎีและหลักการการวิเคราะห์ท่าทาง (Posture Analysis)
- ศึกษาหลักการเขียนภาษาเชิงวัตถุและการสร้างโปรแกรมด้วย MS Visual C++ และ MFC
- ศึกษาการทำงานของ OpenCV
- พัฒนาโปรแกรมในส่วนการจับภาพเบื้องต้นโดยใช้ฟังก์ชั่นของ OpenCV
- พัฒนาโปรแกรมในส่วนการแบ่งส่วน

ภาคเรียนที่ 2

- พัฒนาโปรแกรมในส่วนวิเคราะห์ท่าทาง (Posture Analysis)
- พัฒนาโปรแกรมในส่วนภาพการ์ตูนเคลื่อนไหว (Animation)
- พัฒนาโปรแกรมในส่วนการเชื่อมต่อระหว่างการวิเคราะห์ท่าทาง กับส่วนอนิเมชั่น

1.4 ขอบเขตของการพัฒนาระบบงาน

ในการพัฒนาระบบช่วยฝึกการออกกำลังกายโดยการประยุกต์ใช้เทคนิคการตรวจจับการเคลื่อนไหวของมนุษย์จากวีดิทัศน์มีขอบเขตของการศึกษาและพัฒนาระบบ ดังต่อไปนี้ในการวางตำแหน่งของกล้องตายตัว ไม่มีการเคลื่อนไหวกล้อง ณ เวลาที่ระบบเริ่มทำงาน ท่า แอโรบิคที่ใช้ในการพัฒนาระบบมีจำนวนจำกัด ลักษณะท่าทางของผู้ใช้ต้องอยู่ในท่าทางการยืนเสมอ และจำนวนผู้ใช้เพียงคนเดียว

1.5 ประโยชน์ที่ได้รับ

1. ได้ฝึกฝนการใช้ภาษาเชิงวัตถุในการพัฒนาโปรแกรมประยุกต์ที่สามารถนำไปใช้งานได้
2. ได้นำทฤษฎีด้านคอมพิวเตอร์วิชั่นไปใช้ในการสร้างโปรแกรมประยุกต์ที่สามารถทำงานได้จริง และเป็นแนวทางในการพัฒนาโปรแกรมประยุกต์ในด้านอื่นๆต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
หลักการและทฤษฎีที่เกี่ยวข้อง

ในหัวข้อนี้จะกล่าวถึงทฤษฎีพื้นฐานต่างๆ ที่เกี่ยวข้องในการวิจัย และพื้นฐานของระบบการออกกำลังกายเบื้องต้นโดยการประยุกต์ใช้เทคนิคการตรวจจับการเคลื่อนไหวของมนุษย์จากวีดิทัศน์ ซึ่งเนื้อหาในบทนี้จะกล่าวถึงเทคนิคการแยกส่วนที่สนใจออกจากสิ่งแวดล้อม การตรวจจับและวิเคราะห์ท่าทางของผู้ใช้งาน รวมถึงครื่องมือที่ใช้ในการพัฒนาระบบ

2.1 การตรวจจับและติดตามการเคลื่อนไหวของมนุษย์

ขั้นตอนในการตรวจจับและติดตามการเคลื่อนไหวของมนุษย์จากวีดิทัศน์ นี้ จะต้องมีการเตรียมอุปกรณ์ ต่างๆในการทำงานโดยเฉพาะ โดยระบบกล้องจะอยู่กับที่ในช่วงเวลาที่ทำงาน จากนั้นมีขั้นตอนต่างๆดังนี้

- การแยกส่วนที่สนใจออกจากสิ่งแวดล้อม (Target segmentation) เทคนิคที่เป็นที่นิยมใช้ในการแยกส่วนของร่างกายออกจากฉากหลังคือ เทคนิคการลบฉากหลัง (Background Subtraction) เทคนิคนี้เหมาะสำหรับกรณีที่ฉากหนังในวิดีโอไม่มีการเปลี่ยนแปลง หรือมีการเปลี่ยนแปลงน้อยเท่านั้น ในกรณีที่ฉากหลังมีการเคลื่อนไหวมาก จำเป็นต้องใช้เทคนิคอื่นในการแยกเป้าหมาย
- หลังจากได้ส่วนที่เป็นเป้าหมาย ซึ่งในที่นี้คือ ส่วนที่เป็นตัวโครงร่างมนุษย์ (Silhouette) ออกจากรูปแล้วก็จะทำการวิเคราะห์หาส่วนต่างๆของของร่างกายต่อไป (Body Part Localization) หรือใช้วิธีหา Motion Template เพื่อใช้ในการวิเคราะห์การเคลื่อนไหวแทนการหาตำแหน่งของส่วนของร่างกาย

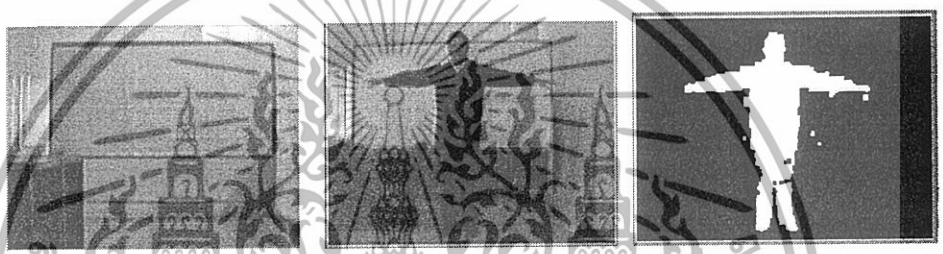
2.2 การแยกส่วนที่สนใจออกจากสิ่งแวดล้อม

ในการแยกส่วนที่เราสนใจออกจากสิ่งแวดล้อมนั้นจะช่วยให้การคำนวณและการวิเคราะห์ในขั้นตอนต่อไปสามารถทำได้ง่ายขึ้นโดยเทคนิคที่เป็นที่นิยมและเหมาะสมกับในส่วนที่ศึกษาคือ เทคนิคการลบฉากหลัง (Background Subtraction) เป็นเทคนิคที่ง่าย และมีประสิทธิภาพ ในการแยกโครงร่าง (Silhouette) ของเป้าหมายหรือวัตถุที่เราสนใจออกจากฉากหลังในภาพ ซึ่งความสามารถในการแยกเป้าหมายออกจากฉากหลังหรือสิ่งแวดล้อมนั้น มีความสำคัญอย่างมากต่อขั้นตอนการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิเคราะห์อื่นๆ ในลำดับต่อไป ดังนั้นหากการลบฉากหลังทำได้ไม่ดีการวิเคราะห์โครงร่างเพื่อค้นหาอวัยวะส่วนต่างๆก็จะผิดตามไปด้วย

หลักการในการลบฉากหลังนั้น เริ่มจากการคำนวณเพื่อสร้างรูปแบบจำลองของฉากหลัง (background model) จาก วิดีโอ วิธีการนี้เหมาะสม ในกรณีที่กล้องตั้งอยู่กับที่ (stationary camera) โดยฉากหลังไม่มีการเคลื่อนไหว หรือ เคลื่อนไหวน้อย เช่น ใบไม้หรือต้นไม้ที่โยกไปมาเนื่องจากแรงลม หลังจากสร้างรูปแบบจำลองของฉากหลัง (background model) แล้ว ในการแยกแยะเป้าหมาย (ในที่นี้คือร่างกายมนุษย์) จากวิดีโอ ให้นำแต่ละเฟรมของวิดีโอลบออกจากรูปแบบจำลองฉากหลัง ผลลัพธ์ที่ได้คือ โครงร่างของเป้าหมาย (target's silhouette) ดังแสดงในภาพที่ 2.1 โดยโครงร่างที่ได้จะนำไปใช้ในการวิเคราะห์ คำนวณหาตำแหน่งอวัยวะที่สำคัญของร่างกายในขั้นตอนต่อไป



(a) (b) (c)

ภาพที่ 2.1 แสดงการลบฉากหลังออกจากภาพ

จากภาพที่ 2.1 (a) แสดงแบบจำลองของฉากหลัง (Background model) ที่ถูกสร้างขึ้นจากช่วงเวลาที่แต่ละเฟรมที่ไม่มีเป้าหมาย (ในที่นี้คือร่างกายมนุษย์) นำช่วงเวลาดังกล่าวแต่ละเฟรมมาหาค่าเฉลี่ยในแต่ละ Pixel สร้างเป็นแบบจำลองของฉากหลัง (Background model)

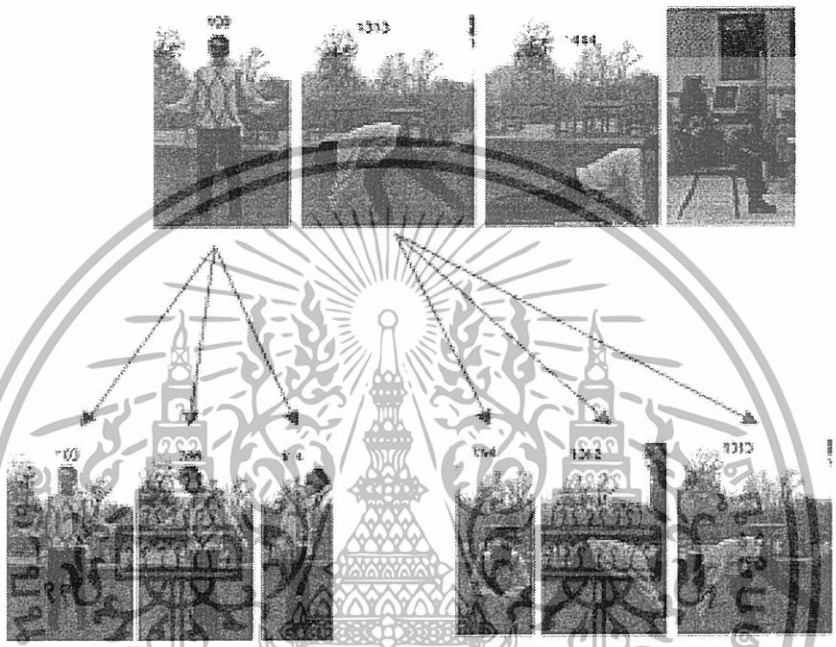
ภาพที่ 2.1 (b) เป็นการแสดงภาพ วีดีโอที่มีเป้าหมายเข้าสู่ระบบเรียกว่า Input Image ข้อมูลดังกล่าวไปลบกับแบบจำลองของฉากหลัง (Background model) จะได้ส่วนที่แตกต่างซึ่งในการลบกัน จะตรวจสอบถึงสีและแสงของข้อมูลด้วย หากมีสีเดียวกัน แต่แสงน้อยกว่าก็อาจคาดเดาได้ว่าเป็นเงาที่เกิดขึ้น หรือหากมีสีเดียวกัน แต่แสงมากกว่าก็อาจคาดเดาได้ว่าเป็นแสงสะท้อนที่เกิดขึ้นได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 การทำเครื่องหมายส่วนของร่างกาย

2.3.1 การตรวจจับส่วนของร่างกายโดยใช้โครงร่าง (Silhouette)

ท่าทางของคนเราสามารถแบ่งได้เป็นท่าหลักๆได้ 4 ท่า คือท่า ยืน, นั่ง, คลาน (ท่าโก้งโค้ง) และ ท่านอน (หมอบ) และในแต่ละท่าทางสามารถแบ่งมุมมองได้อีก คือ หน้า, หลัง, ซ้าย และ ขวา

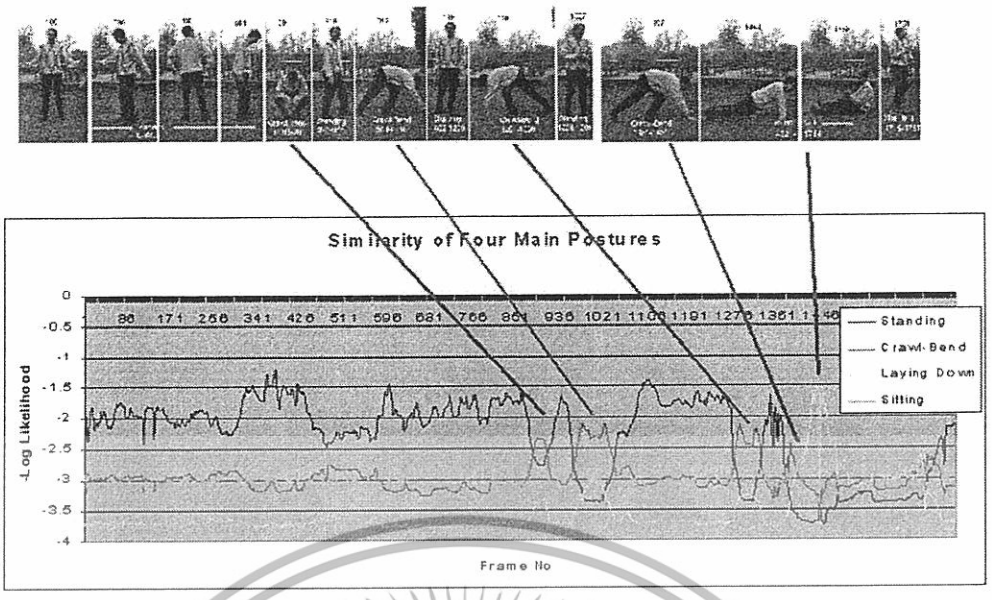


ภาพที่ 2.2 แสดงท่าทางต่างๆเมื่อแสดงให้เห็นมุมมองที่ต่างกัน

2.3.1.1 การระบุท่าทางต่างๆ

ในการระบุท่าทาง เราจะดูได้จากการทำการคำนวณหาความคล้ายคลึงในภาพทางแนวนอนและแนวตั้ง (Horizontal and Vertical Projection) แล้วทำการทำให้อยู่ในรูปปกติ แล้วทำการเปรียบเทียบกับท่าหลักทั้ง 4 ท่า คือ ท่า ยืน นั่ง คลาน นอน ท่าทางของร่างกายที่ได้ผลคล้ายคลึงที่สุดที่วัดได้ก็จะสามารถทำการคาดเดาได้ว่าเป็นท่าทางนั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.3 ทำการแสดง Sequence ของการเคลื่อนไหว

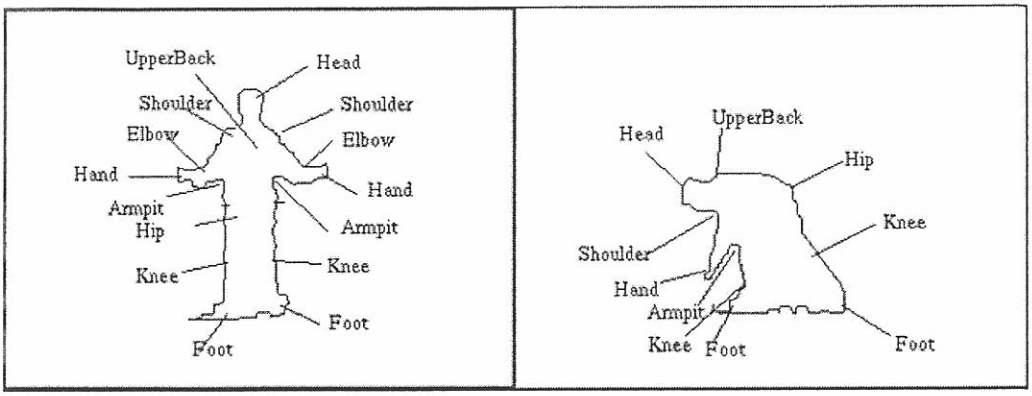
การวิเคราะห์ท่าทางมนุษย์แบบลำดับชั้น (Hierarchical Body Posture) คือ การนำโครงร่างมาคำนวณหาความคล้ายคลึงกับ การฉาย ของ พิกเซล ในภาพตามแนวนอนและแนวตั้ง (Horizontal and vertical Projection) เพื่อใช้ตรวจจับโครงร่างและท่าหลัก ท่าทางของร่างกายที่ได้ผลคล้ายคลึงที่สุดที่วัดได้จะสามารถคาดเดาได้ว่าเป็นท่าทางนั้น

2.3.1.2 แบบจำลองร่างกาย 2 มิติ (2D Body Model)

โครงสร้างแบบจำลอง 2 มิติ ใช้โดยการสร้างแบบจำลองของรางกายโดยใช้ ส่วนหลักๆ 6 ส่วน คือหัว, มือ, เท้า, ลำตัว และส่วนประกอบย่อยอีก 10 ส่วน คือ ข้อศอก, หัวเข่า, ไหล่, สะโพก, รักแร้ , หลังส่วนบน

ความสัมพันธ์ของตำแหน่งของร่างกายมนุษย์ในแต่ละท่าทางที่คล้ายคลึงกัน หรือ ท่าทางท่าเดียวกัน ตำแหน่งของ หัว, ลำตัว, มือ และ เท้า ทั้งสองบนเส้นขอบเขตโครงร่าง (Silhouette Boundary)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



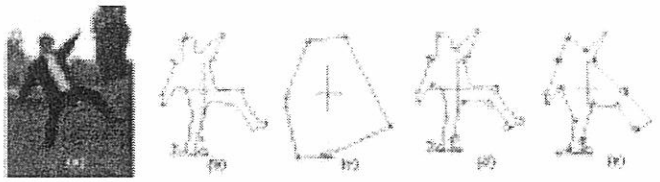
ภาพที่ 2.4 การแสดงลำดับของตำแหน่งของส่วนต่างๆ

การคาดเดาท่าทาง มนุษย์ สามารถมีท่าทางได้หลายท่าทางในการกระทำในแต่ละการกระทำ และในแต่ละท่าทางยังสามารถได้หลายมุมมองขึ้นกับองศาที่เห็น ซึ่งในการทำงานของระบบนี้จะใช้การเก็บตัวอย่างท่าทางจาก ระยะไกลจากโครงร่างมนุษย์ (Silhouette) เพื่อค้นหาลำดับของส่วนของร่างกายในท่าทางหลัก (ยืน, นั่ง, คลาน (ท่าโก้งโค้ง), ท่านอน (หมอบ) โดยแบ่งแยกประเภทโดยสร้างแบบโครงร่างแบบลำดับชั้น ซึ่งจะได้ท่าทาง 4 หลักและในแต่ละท่าหลักจะสามารถแบ่งมุมมองได้อีกคือ หน้าตรง ด้านซ้าย ด้านขวา

ท่าทางต่างๆจะสามารถแสดงได้ในรูป ของพิกเซล ในภาพตามแนวนอนและแนวตั้ง (Horizontal and vertical Projection) ที่ทำการให้อยู่ในรูปปกติแล้ว ในการตรวจจับภาพจากวีดิทัศน์ ก็จะนำแต่ละเฟรมไปคำนวณหาค่าของพิกเซลในภาพตามแนวนอนและแนวตั้ง (Horizontal and vertical Projection) แล้วนำมาเทียบกับกราฟเดิมที่มีอยู่ว่ามีค่าใกล้เคียงกับกราฟท่าทางใด ก็จะคาดเดาว่าเป็นท่าทางนั้น

การตรวจสอบ ส่วนที่ยื่นออกมา (Convex hulls) ส่วนที่เว้าเข้าข้างใน (Concave hulls) ของโครงร่างมนุษย์ (Silhouettes) โดยใช้ Graham scan convex-hull algorithm มาปรับปรุงสามารถที่จะใช้ประโยชน์ของคุณสมบัติของโครงร่างและการคำนวณให้เร็วขึ้น ซึ่งการใช้การคำนวณส่วนที่ยื่นออกมา (Convex hulls) โดยใช้วิธีเดียวไม่พอที่จะใช้ในการตรวจสอบโครงร่างได้ จึงต้องใช้หลายวิธีการเพื่อที่จะตรวจสอบได้ดีขึ้น โดยลำดับแรกหาส่วนที่ยื่นออกมาสูงสุด และ เชื่อมจุดสูงสุดนั้นโดยใช้ อัลกอริทึม ที่พัฒนาขึ้นเพื่อนำไปหาส่วนเว้าเข้าข้างใน (Concave hulls) และทำซ้ำอีก

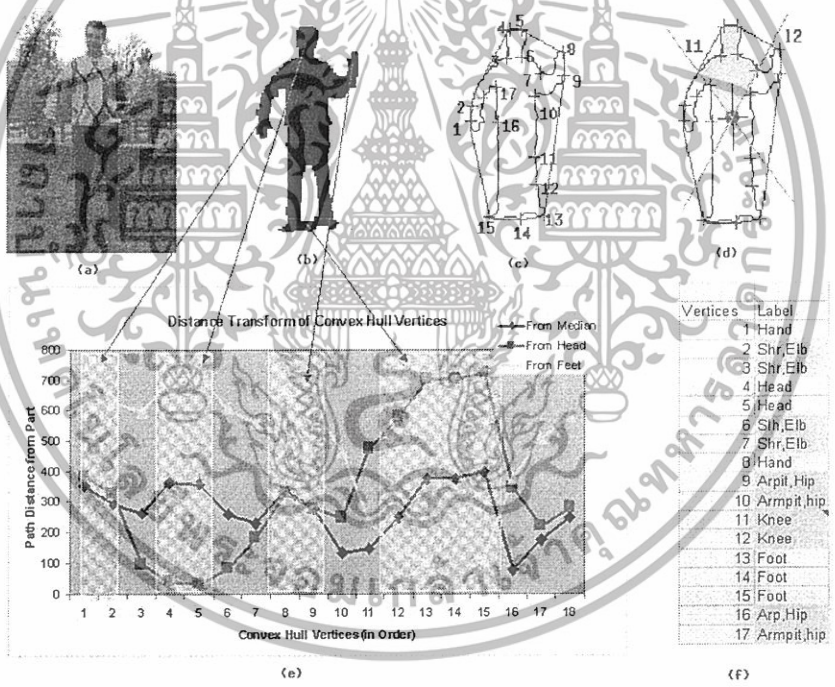
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.5 แสดงท่ายืน Input Image (a) convex-hull vertices (b) เส้นเชื่อมระหว่าง convex-hull vertices (c) convex และ concave hull vertices (d) และ เส้นเชื่อมจุด convex และ concave hull vertices (e)

2.3.1.3 การให้เครื่องหมายส่วนของร่างกาย (Body Part Labeling)

การใช้หัวเป็นจุดในการอ้างอิงในลำดับแรก และส่วนอื่นๆ เป็นลำดับต่อไป จะใช้การประเมินรวมกันระว่าง แกน, มัธยฐาน และ ท่าทาง ในการหาส่วนโครงร่างโดยการตรวจสอบส่วนที่ยื่นออกมาและส่วนที่เว้าเข้าไป (Convex hulls &Concave hulls)

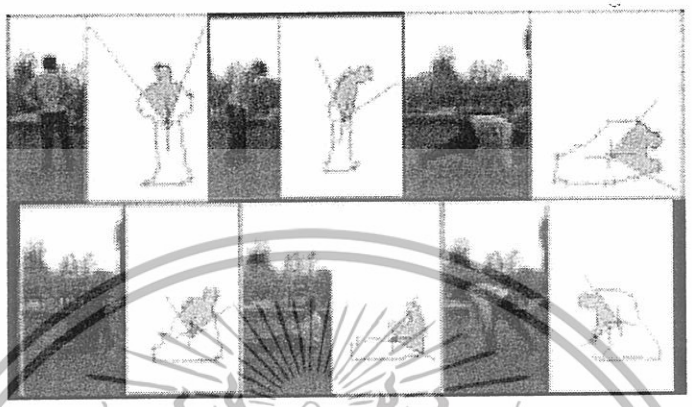


ภาพที่ 2.6 การระบุส่วนของร่างกาย (ส่วนหัว)

การคาดเดาส่วนที่เป็นศีรษะ ใช้หลักการของความสัมพันธ์ระหว่างตำแหน่งของร่างกายจากโครงร่างที่ได้ทำการคาดเดาท่าทางแล้วกับแกนหลักของโครงร่าง และ ข้อมูลของระนาบพื้นเพื่อใช้ในการคาดเดาส่วนย่อยของโครงร่างที่อยู่ตรงข้ามกับศีรษะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การคาดเดาตำแหน่งของ เท้า มือ และ ลำตัว หลังจากที่ทำการคาดเดาตำแหน่งของศีรษะแล้วก็จะใช้การคาดเดาส่วนอื่นๆจากลำดับของตำแหน่งของส่วนต่างๆในแต่ละท่าทางที่ทำเอาไว้ โดยใช้ข้อกำหนดของลำดับของตำแหน่งของส่วนของร่างกาย และ ความสัมพันธ์ของระยะระหว่างส่วนต่างๆ



ภาพที่ 2.7 แสดงการคาดเดาตำแหน่งของศีรษะ ด้วยวิธี Ghost

2.3.2 การตรวจจับและติดตามส่วนของร่างกายโดยใช้ Cardboard Model ใน W4

W4 เป็นระบบสายตะเวนเพื่อใช้ในการตรวจจับและติดตามการเคลื่อนไหวของมนุษย์จากวีดิทัศน์ พัฒนาโดย UNIVERSITY OF MARYLAND เมื่อปี 1998

หลักการการทำงานของ W4 กล่าวคือ W4 จะทำการตรวจจับและติดตามวัตถุที่เคลื่อนไหวในฉาก โดยทำการตรวจจับและติดตามโครงร่างโดยรวม (Whole silhouette region) ของวัตถุนั้นๆ ในขณะเดียวกันก็จะทำการตรวจจับและติดตามส่วนประกอบของร่างกายด้วยถ้าวัตถุที่ตรวจพบนั้นเป็นมนุษย์

ก่อนอื่นจะกล่าวถึงหลักการในการตรวจจับและติดตามโครงร่างโดยรวม (Whole silhouette region) ของแต่ละวัตถุที่เคลื่อนไหวในฉากก่อนหน้า จากนั้นจะกล่าวถึงเทคนิคที่ W4 ใช้ในการตรวจจับและติดตามส่วนของร่างกายในกรณีที่ ภาพโครงสร้างนั้นเป็นร่างกายมนุษย์โดยใช้ Card Board Model ในการตรวจจับและติดตามโครงร่างที่เคลื่อนไหวที่ปรากฏในฉากนั้นมีขั้นตอนดังนี้

1. หาว่าวัตถุใหม่เข้ามาในมุมมองของระบบเมื่อไหร่ และเริ่มสร้างตัวต้นแบบการเคลื่อนไหวสำหรับการติดตามการเคลื่อนไหววัตถุนั้น

2. คำนวณความสัมพันธ์ระหว่างบริเวณโครงร่างที่ตรวจพบโดยการลบพื้นหลัง และ วัตถุที่ถูกติดตามการเคลื่อนไหวโดย W4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. ใช้กระบวนทำงานของการติดตามการเคลื่อนไหวเพื่อประมาณตำแหน่ง (ของร่างกาย) ของวัตถุแต่ละโครงร่าง และทำการอัพเดทตัวต้นแบบการเคลื่อนไหวที่ใช้ในการติดตามการเคลื่อนไหว W4 ใช้โครงร่างต้นแบบการเคลื่อนไหวแบบอนุพันธ์อันดับสอง (รวมไปถึง ความเร็ว และ ค่าความเร่ง ที่อาจเป็นศูนย์) ในการสร้างโครงร่างต้นแบบของการเคลื่อนไหวทั้งโครงร่างและการเคลื่อนไหวของส่วนต่างๆของร่างกาย

W4 จะติดตามการเคลื่อนไหววัตถุต่อไปถึงแม้ว่าความสามารถในการตรวจจับจะอยู่ในระดับต่ำไม่สามารถแบ่งผู้คนเป็นวัตถุจากโครงร่างโดยรวม (Whole silhouette region) ได้ ซึ่งอาจจะเกิดขึ้นเนื่องจากวัตถุถูกบดบังชั่วคราวในแต่ละเฟรม หรือ วัตถุแยกเป็นหลายชิ้น (อาจจะเนื่องมาจากผู้คนเอาถือวัตถุเข้ามาในฉากใน หรือ คนถูกปิดบังโดยวัตถุชิ้นเล็กๆ) ในที่สุด วัตถุที่ระบบติดตามการเคลื่อนไหวที่แยกกันอยู่ก็อาจจะมารวมเป็นชิ้นเดียวกันเนื่องจากการปฏิสัมพันธ์กันระหว่างบุคคล ภายใต้เงื่อนไขนี้การวิเคราะห์รูปร่างโดยรวมและกระบวนการทำงานการติดตามการเคลื่อนไหวโดย W4 อาจจะเกิดข้อผิดพลาด ระบบก็จะใช้เทคนิคความสัมพันธ์ในการพยายามที่จะติดตามการเคลื่อนไหวของวัตถุที่ปฏิสัมพันธ์กัน

อันดับแรก W4 จะจับคู่วัตถุกับบริเวณโครงร่างโดยการหาบริเวณที่ซ้อนทับกันระหว่างกรอบที่ประมาณค่าโดยใช้ต้นแบบการเคลื่อนไหวโดยรวมของวัตถุ และ กรอบของบริเวณโครงร่างจากเฟรมปัจจุบัน สำหรับวัตถุแต่ละโครงร่าง บริเวณโครงร่างทั้งหมด ซึ่งกรอบซ้อนทับกันมากพอที่จะเหมาะนำมาจับคู่กับวัตถุนั้น ในทางอุดมคติ การจับคู่แบบหนึ่งต่อหนึ่ง (การติดตามการเคลื่อนไหว) จะถูกพบในระหว่างการติดตามการเคลื่อนไหวหนึ่งๆ อย่างไรก็ตาม การจับคู่แบบหนึ่งต่อจำนวนมาก (วัตถุที่ถูกติดตามการเคลื่อนไหวถูกแบ่งแยกเป็นบริเวณโครงร่างหลายอัน) การจับคู่แบบจำนวนมากกับหนึ่ง (วัตถุที่ถูกติดตามการเคลื่อนไหวรวมโครงร่างกันเป็นบริเวณโครงร่างเพียงหนึ่งบริเวณ) การจับคู่แบบหนึ่งกับศูนย์ (การหายไป) และ การจับคู่แบบศูนย์กับหนึ่ง (การปรากฏขึ้น) อาจจะเกิดขึ้นอยู่บ่อยๆ W4 จะใช้การติดตามการเคลื่อนไหวโดยใช้วิธีที่ต่างกันสำหรับการจับคู่แต่ละแบบ

2.3.2.1 การที่วัตถุปรากฏขึ้น (Appearing Objects)

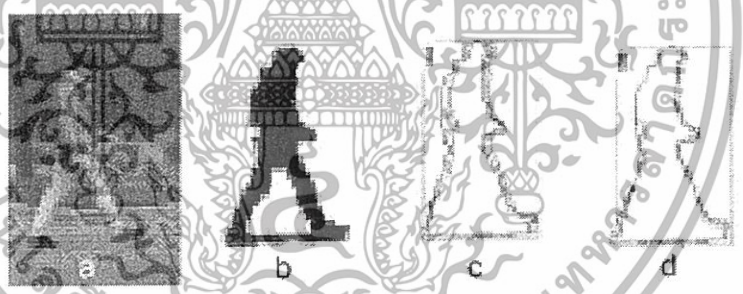
เมื่อบริเวณโครงร่าง ถูกตรวจพบ เราไม่สามารถบอกได้เลยทีเดียวว่ามันเป็นวัตถุจริงหรือแค่สัญญาณรบกวนเท่านั้น ถ้าบริเวณนั้นสามารถถูกติดตามการเคลื่อนไหวได้ในหลายๆเฟรม มันก็จะถูกเพิ่มเข้าไปในรายการวัตถุที่จะต้องถูกติดตามการเคลื่อนไหวด้วย

2.3.2.2 การติดตามการเคลื่อนไหว (Tracking)

ส่วนนี้ เราจะพิจารณาสถานการณ์ที่วัตถุถูกติดตามการเคลื่อนไหวที่เป็นโครงร่างเพียงโครงร่างเดียว W4 ใช้รูปแบบการเคลื่อนไหวแบบอนุพันธ์อันดับสองสำหรับวัตถุแต่ละโครงร่างในการประมาณตำแหน่งในเฟรมต่อมา การพยากรณ์จากโครงร่างต้นแบบนี้ถูกนำมาใช้ในการประมาณหา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตำแหน่งของกรอบสำหรับวัตถุแต่ละโครงร่าง ตำแหน่งของกรอบที่ถูกพยากรณ์ก็จะถูกนำมาเปรียบเทียบกับกรอบจริงของบริเวณโครงร่างที่ตรวจจับได้ กำหนดให้วัตถุถูกจับคู่กับบริเวณโครงร่างโครงร่างเดียว (และขนาดของบริเวณเหล่านั้นก็เกือบเท่ากัน) W4 ต้องหาตำแหน่งปัจจุบันของวัตถุเพื่อที่จะทำการอัพเดตโครงร่างการเคลื่อนไหวต้นแบบ แม้ว่าการเคลื่อนไหวทั้งหมดของวัตถุจะค่อนข้างน้อยในแต่ละเฟรม สำหรับโครงร่างที่มีเปลี่ยนแปลงมาก(เคลื่อนไหวมาก)จะใช้เทคนิคง่ายๆ เช่น การติดตามจุดศูนย์กลางรวมของบริเวณโครงร่างผิดพลาดได้ แทนที่จะให้เป็นแบบนั้น W4 ใช้วิธีจับคู่แบบสองชั้นเพื่อที่จะอัพเดตการประมาณตำแหน่งโดยรวมของวัตถุ การประมาณตำแหน่งตอนเริ่มต้นถูกคำนวณเป็นการเคลื่อนไหวของจุดพิกัดมัธยฐาน ของวัตถุ ค่าพิกัดมัธยฐานนี้เป็นการประมาณตำแหน่งของวัตถุที่มีประสิทธิภาพมากกว่า และไม่ถูกกระทบโดยการเคลื่อนไหวมากๆ (ซึ่งมีแนวโน้มที่จะกระทบจุดศูนย์กลางรวมอย่างมาก) มันทำให้เราบีบการค้นหาที่ว่างสำหรับการเคลื่อนไหวของวัตถุ อย่างไรก็ตาม การประมาณนี้ไม่เพียงพอสำหรับการติดตามการเคลื่อนไหวในระยะยาว ดังนั้นหลังจากที่ย้ายภาพตัดขวางของวัตถุจากเฟรมที่ผ่านมาโดยใช้การประมาณโดยใช้ median-base เราใช้ความสัมพันธ์ของขอบแบบฐานสองระหว่างโครงร่างปัจจุบันและโครงร่างก่อนหน้านี้ ความสัมพันธ์นี้ถูกคำนวณโดยใช้ชุดของจุดตำแหน่ง 5x3 ความสัมพันธ์จะถูกครอบครองโดยขอบของร่างกายและหัว ซึ่งรูปร่างของมันจะเปลี่ยนไปอย่างช้าๆ จากเฟรมหนึ่งไปยังอีกเฟรมหนึ่ง กระบวนการติดตามความเคลื่อนไหวได้แสดงดัง ภาพที่ 2.8



ภาพที่ 2.8 การคาดเดาการเคลื่อนไหวของร่างกายโดยใช้โครงร่างจับคู่กันระหว่าง 1 เฟรมที่เลือกมาภาพที่เข้ามา (a) การตรวจจับโครงร่าง (b) alignment ของขอบโครงร่างโดยใช้ความแตกต่างของ Median (c) alignment สุดท้ายหลังจากทำความสัมพันธ์โครงร่าง (d)

2.3.2.3 การแบ่งแยกบริเวณ (Region splitting)

วัตถุที่ถูกติดตามการเคลื่อนไหวอาจจะแบ่งเป็นบริเวณโครงร่างหลายๆบริเวณ ไม่ว่าจะเป็นการบดบังเพียงบางส่วน หรือ เพราะว่าคนได้นำวัตถุเข้ามาในฉากด้วย ในกรณีนี้ โครงร่างหนึ่งจะถูกจับคู่กับบริเวณสองบริเวณหรือมากกว่านั้น W4 จะหาว่าการแบ่งแยกนี้เป็นการแบ่งแยกของจริงหรือหลอก (เนื่องมาจากสัญญาณรบกวน) โดยการสังเกตดูเฟรมต่อมา ในขณะที่ติดตามดูวัตถุที่ถูก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบ่งแยกออกมาเป็นวัตถุแต่ละโครงร่าง ถ้า W4 สามารถติดตามดูวัตถุย่อยๆได้ในหลายๆเฟรม แล้วเราสามารถสรุปได้ว่าพวกมันเป็นวัตถุแต่ละโครงร่าง และเริ่มติดตามดูความเคลื่อนไหวของมันแต่ละโครงร่าง

2.3.2.4 การรวมบริเวณ (Region merging)

ในส่วนนี้จะพิจารณาสถานการณ์ที่วัตถุถูกติดตามการเคลื่อนไหวที่เป็นโครงร่าง 2 โครงร่างหรือ คน 2 คน เมื่อคนสองคนพบกัน (เดินสวนกัน) พวกเขาจะถูกแบ่งเป็นบริเวณโครงร่างเพียงหนึ่งโครงร่างโดยกระบวนการทำงานการลบพื้นหลัง W4 จำได้ว่านี่มันเกิดขึ้นโดยมีพื้นฐานมาจากการวิเคราะห์อย่างง่ายของกรอบที่ถูกพยากรณ์ของวัตถุที่เราติดตามการเคลื่อนไหวและกรอบของบริเวณโครงร่างที่รวมแล้ว บริเวณที่รวมแล้วจะถูกจับตาดูความเคลื่อนไหวจนกระทั่งมันแยกกลับเป็นวัตถุย่อย เนื่องจากภาพตัดขวางของบริเวณมีแนวโน้มที่จะเปลี่ยนรูปอย่างรวดเร็วและไม่สามารถคาดเดาได้ W4 จึงใช้วิธีเทียบค่าอย่างง่ายในการสร้างโครงร่างต้นแบบการพยากรณ์การเคลื่อนไหวสำหรับวัตถุที่ได้รวมมานี้ ปัญหาที่เกิดขึ้นเมื่อบริเวณที่รวมกันได้แบ่งแยกออก และ คนปรากฏโครงร่างขึ้น ก็คือการหาความสัมพันธ์ระหว่างคนที่ถูกติดตามการเคลื่อนไหวก่อนที่จะมีการปฏิสัมพันธ์และคนที่ได้มารวมกันจากการปฏิสัมพันธ์ ในการแก้ปัญหานี้ W4 ใช้โครงร่างต้นแบบการปรากฏสองชนิดที่สร้างขึ้นในขณะที่มันกำลังติดตามดูการเคลื่อนไหวของคนที่แยกออกมา

W4 ได้สร้างแผ่นเทมเพลทเชิงเวลา หรือเรียกว่า เทมเพิลรอลเท็กชั่ลเทมเพลท (temporal texture template) เป็นโครงร่างต้นแบบการเคลื่อนไหวเทมเพิลรอลเท็กชั่ลเทมเพลทสำหรับวัตถุสามารถหาได้จากสูตร

$$\Psi^{t}(x,y)=\frac{I(x,y)+w^{t-1}(x,y)\times\Psi^{t-1}(x,y)}{w^{t-1}(x,y)+1} \tag{2.2}$$

ในที่นี้ ขออ้างถึงบริเวณโครงร่างที่ถูกตรวจพบในระหว่างการติดตามการเคลื่อนไหววัตถุและจุดพิกัดก็ถูกแสดงโดยสัมพันธ์กับค่า มัธยฐาน ของแผ่นเทมเพลทหรือบริเวณโครงร่าง น้ำหนักจากสูตร ก็คือความถี่ที่ ค่าพิกเซล ใน Ψ ถูกตรวจพบเป็นพิกเซลของโครงร่างในระหว่างการติดตามการเคลื่อนไหว ค่าน้ำหนักเริ่มต้น $w^{t}(x,y)$ ของ Ψ มีค่าเป็นศูนย์และจะเพิ่มขึ้นในแต่ละครั้งที่ตำแหน่งที่สัมพันธ์กับค่าพิกัดของค่ามัธยฐาน ของเทมเพลทถูกตรวจพบในลักษณะของพิกเซลของโครงร่างในภาพของ ส่วนที่รับเข้ามา

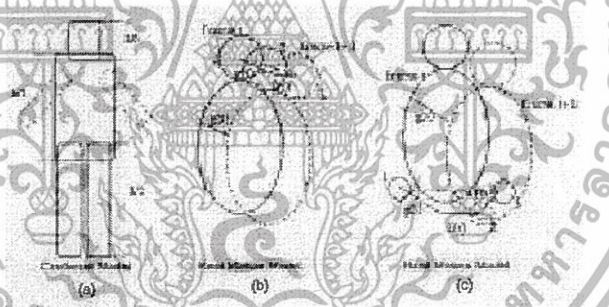
หลังจากแยกออกจากกัน วัตถุย่อยก็ถูกจับคู่กับวัตถุที่แยกออกจากกันโดยทำให้ เทมเพิลรอลเทมเพลท ของพวกมันสัมพันธ์กัน เนื่องจาก เทมเพิลรอลเท็กชั่ลเทมเพลทได้มาจากการดู มันอาจจะไม่สามารถจับคู่ได้ถ้ามันมีการเปลี่ยนแปลงการขยับของวัตถุมากในระหว่างเหตุการณ์บดบัง ดังนั้น วิธีที่ไม่ได้ใช้การดูซึ่งใช้การแสดงผลวัตถุเป็นสัญลักษณ์ ก็ได้นำมาใช้ในการวิเคราะห์การบด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บังนี้ด้วย ยกโครงร่างอย่าง เช่น ถ้า เทมเพิลรอลเท็กชั่ลเทมเพลทไม่สามารถให้ค่าความสัมพันธ์ที่สูงพอ ดังนั้นเมื่อเราจับคู่วัตถุโดยใช้ความเข้มเฉลี่ยในส่วนบน ล่าง และกลาง ในการพยายามที่จะระบุวัตถุเมื่อมันแยกออกไปแล้ว

2.3.2.5 การติดตามการเคลื่อนไหวของส่วนต่างๆของร่างกาย

ในส่วนเพิ่มเติมของการติดตามการเคลื่อนไหวของคนทั้งหมด เราต้องการที่จะระบุตำแหน่งของส่วนต่างๆของร่างกาย เช่น ศีรษะ ร่างกาย ขา และ เท้า และ ติดตามการเคลื่อนไหวของมันเพื่อที่จะเข้าใจการกระทำ W4 ใช้การวิเคราะห์ร่างกายร่วมกับการจับคู่เท็มเพลทเพื่อที่จะดูการเคลื่อนไหวของส่วนต่างๆนี้ (เมื่อคนถูกบัง และรูปร่างของมันไม่สามารถพยากรณ์ได้ง่ายนัก เราจะใช้แค่การจับคู่เท็มเพลทในการดูการเคลื่อนไหวของส่วนต่างๆของร่างกาย) โครงร่างต้นแบบของรูปร่างนำมาใช้งานโดยการใช้ Cardboard Model ซึ่งแสดงถึงตำแหน่งและขนาดสัมพัทธ์ของร่างกายและส่วนต่างๆของมัน เมื่อใช้งานร่วมกับโครงร่างต้นแบบการพยากรณ์การเคลื่อนไหวแบบอนุพันธ์อันดับสอง Cardboard Model สามารถนำมาใช้ในการพยากรณ์หาตำแหน่งของส่วนต่างๆของร่างกายในแต่ละเฟรมได้ ภาพที่ 2.9 ได้แสดงถึงโครงร่างต้นแบบการเคลื่อนไหวที่ใช้สำหรับมือและศีรษะ ตำแหน่งเหล่านี้ถูกระบุและกลั่นกรองโดยใช้การจับคู่เท็มเพลทที่คงที่จาก เทมเพิลรอลเท็กชั่ลเทมเพลท ของส่วนของร่างกายที่เราสังเกต

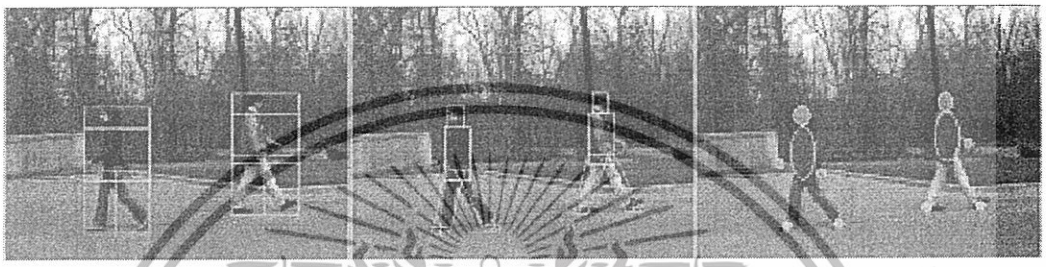


ภาพที่ 2.9 Cardboard model ที่ใช้ใน W^4 (a) และ แบบจำลองที่เคลื่อนไหวที่ใช้สำหรับศีรษะ (b) และ มือ(c)

Cardboard model เป็นแบบจำลองร่างกายมนุษย์ในท่ายืน ดังภาพ 2.9(a) มันถูกนำมาใช้ในการพยากรณ์ค่าตำแหน่งของส่วนต่างๆของร่างกาย (ศีรษะ ร่างกาย เท้า มือ ขา) ความสูงของกรอบของวัตถุก็คือความสูงของ cardboard model จากนั้นเราก็ใช้ค่าขนาดที่คงที่ในแนวตั้งการหาตำแหน่งเริ่มต้น (กรอบ) ของส่วนของร่างกายแต่ละส่วน ดังภาพที่ ความยาวของกรอบตอนเริ่มต้นของศีรษะ ร่างกายและขาจะถูกคำนวณเป็น 1/5, 1/2 และ 1/2 ของความยาวของกรอบของวัตถุตามลำดับ ความกว้างของกรอบของศีรษะ ร่างกายและขาถูกคำนวณโดยการหาความกว้างมัธยฐานภายในกรอบตอนเริ่ม นอกจากการหาขนาดและตำแหน่ง ค่าโมเมนต์ของค่าพิกเซลของโครงร่าง

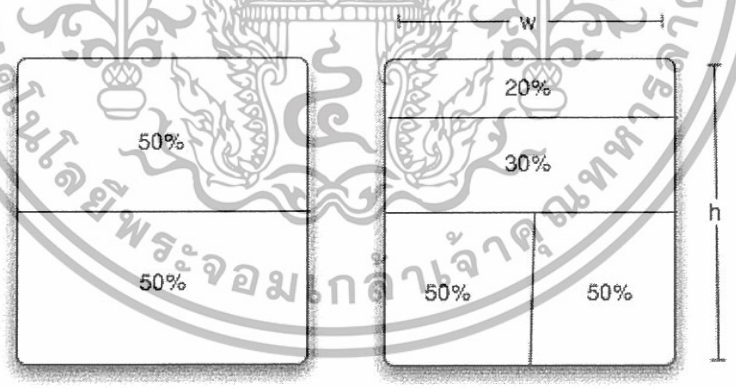
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภายในกรอบก็ต้องถูกคำนวณเพื่อประมาณค่าแกนหลักของมัน แกนหลักนี้จะให้ข้อมูลเกี่ยวกับท่าทางของส่วนต่างๆของร่างกาย ศีรษะจะถูกหาก่อน ตามมาด้วย ร่างกายและ ขา มือจะถูกหาหลังจากร่างกายโดยการหาบริเวณที่เข้มข้นที่เชื่อมต่อกับร่างกายและอยู่ข้างนอกร่างกาย เท้าก็จะถูกพบในบริเวณเข้มในทิศทางของแกนหลักของขา ภาพที่ 2.10 ได้แสดงถึงโครงร่างอย่างของวิธีนำ cardboard model ไปใช้ในการพยากรณ์ตำแหน่งของส่วนต่างๆของร่างกายในสองขั้นตอน (การประมาณตำแหน่งตอนเริ่มและตำแหน่งสุดท้ายที่ถูกประมาณ) และแสดงมันในแบบที่ละเอาไว้



(a) (b) (c)

ภาพที่ 2.10 เป็นโครงร่างอย่างการใช้ Cardboard Model แสดงการบอกตำแหน่งส่วนต่างๆของร่างกาย (a) แสดงค่าเริ่มต้นของกรอบบอกตำแหน่งขอบเขตของโครงร่าง (b) ใช้ Cardboard Model วิเคราะห์ตำแหน่งของส่วนของร่างกาย (c) แสดงการระบุตำแหน่งของร่างกาย (ศีรษะ, ลำตัว, มือทั้งสองข่าง และ เท้าทั้งสองข้าง)



(a) (b)

ภาพที่ 2.11 แสดงอัตราส่วนในการแยกส่วนประกอบของร่างกาย

หลังจากที่พยากรณ์ค่าตำแหน่งของศีรษะและมือโดยใช้ cardboard model แล้ว , ตำแหน่งของพวกมันก็จะถูกระบุและกลั่นกรองโดยใช้ เทมเพิลรอลเท็กชั่ลเทมเพลท, เทมเพิลรอลเท็กชั่ลเทมเพลทเหล่านี้ก็จะถูกอัพเดทดังที่ได้อธิบายไปก่อนหน้านี้ เว้นเสียแต่ว่ามันถูกพบในรูปด้านของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของร่างกาย ในกรณีนี้ ค่าพิกเซลที่เกี่ยวข้องกับศีรษะและมือจะถูกฝังอยู่ในส่วนประกอบที่ใหญ่กว่าที่เกี่ยวข้องกับร่างกาย นี้ทำให้มันยากในการที่ประมาณตำแหน่งมัธยฐาน ของแต่ละส่วนได้อย่างถูกต้องหรือหาว่าพิกเซลไหนที่อยู่ในร่างกายที่เป็นพิกเซลของส่วนอื่น ในกรณีเหล่านี้ ส่วนต่างๆ ของร่างกายก็จะถูกติดตามการเคลื่อนไหวโดยใช้การสัมพันธ์กัน แต่แผ่นเทมเพลทจะไม่ถูกอัพเดท

สรุปหลักการการทำงานของ W4 กล่าวคือ W4 จะทำการตรวจจับและติดตามวัตถุที่เคลื่อนไหวในฉากโดยทำการตรวจจับและติดตามโครงร่างโดยรวม (Whole silhouette region) ของวัตถุนั้นๆ ในขณะเดียวกันก็จะทำการตรวจจับและติดตามส่วนประกอบของร่างกายด้วยถ้าวัตถุที่ตรวจพบนั้นเป็นมนุษย์ใช้หลักการ จับคู่เท็มเพลทในการติดตามส่วนต่างๆและใช้แบบจำลอง Cardboard model ในการพยากรณ์ค่าตำแหน่งของส่วนต่างๆของร่างกาย (ศีรษะ ร่างกาย เท้า มือ และ ขา) โดยความสูงของกรอบของวัตถุก็คือความสูงของ Cardboard model จากนั้นเราก็ใช้ค่าขนาดที่คงที่ในแนวตั้ง การหาตำแหน่งเริ่มต้นของส่วนของร่างกายแต่ละส่วน โดยความยาวของกรอบ ศีรษะ ร่างกายและขาจะถูกคำนวณเป็น 1/5, 1/2 และ1/2 ของความยาวของกรอบของวัตถุตามลำดับ วิธีนี้เป็นวิธีการที่มีความเร็วในการตรวจจับและติดตามสูง และสามารถที่จะตรวจจับมนุษย์ได้มากกว่า 1 คน แต่มีข้อจำกัดที่ใช้งานได้เฉพาะในท่ายืน

2.4 พื้นฐานการประมวลผลภาพ

2.4.1 มอร์โฟโลยี (Morphology)

มอร์โฟโลยีเป็นการกำจัดสัญญาณรบกวน (Noise) ที่เกิดขึ้นออกไปก่อนที่จะนำภาพไปประมวลผลต่อ ซึ่งจะทำการพิจารณากลุ่มสีที่เป็นวัตถุในภาพตามที่กำหนดไว้ นั้นคือ พิจารณาจุดขาวซึ่งอยู่บนพื้นหลังดำ ซึ่งจะมีการกำหนดStructure element ขึ้นมาในการพิจารณา โดยการทำงานพื้นฐานของมอร์โฟโลยีจะประกอบไปด้วย อีรอชัน, ไดเลชัน

รูปภาพหรือกลุ่มของจุดภาพสีดำและขาว ซึ่งมีลักษณะเป็นรูปภาพฐานสอง (Binary Image) การทำไดเลชันและอีโรชันถือเป็น ตัวดำเนินการเชิงสัณฐานเบื้องต้น และเป็นพื้นฐานในการนำไปใช้กับกระบวนการที่มีความซับซ้อนอันได้แก่ ตัวดำเนินการแบบเปิด (Opening Operator), ตัวดำเนินการแบบปิด (Closing Operator)

2.4.1.1 การทำไดเลชัน (Dilation)

นิยาม คือ เซตของจุด p ที่เกิดจาก x+b โดยที่ x เป็นสมาชิกของ x และ b เป็นสมาชิกของ B

$$X \oplus B = \left\{ p \in \varepsilon^2 : p = x + b, x \in X \,\&\, b \in B \right\} \tag{2.3}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ X คือ รูปภาพที่ทำการไดเลชัน

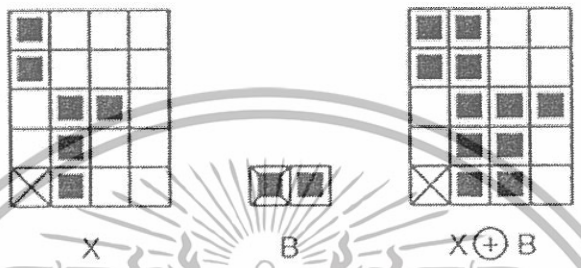
B คือ โครงสร้างส่วนย่อยในการทำไดเลชัน

ตัวอย่าง การทำไดเลชันกำหนดให้

$$X = \{(1,0),(1,1),(1,2),(2,2),(0,3),(0,4)\}$$

$$B = \{(0,0),(1,0)\}$$

$$X \oplus B = \{(1,0),(1,1),(1,2),(0,3),(0,4),(2,0),(2,1),(2,2),(3,2),(1,3),(1,4)\}$$



ภาพที่ 2.12 แสดงการทำไดเลชัน

2.4.1.2 การทำอีโรชัน (Erosion)

นิยามคือ เซตของจุด p ทุกจุดจากรูปภาพทดสอบ ผลลัพธ์ของการทำอีโรชันคือ จุด p เหล่านั้นที่ทำให้ p+b เป็นสมาชิกของ X สำหรับทุกๆ b ที่เป็นสมาชิกของ B

$$X \ominus B = \{p \in \varepsilon^2 : p + b \in X \ for \quad every \ b \in B\} \tag{2.4}$$

โดยที่ X คือ รูปภาพที่ทำการอีโรชัน

B คือ โครงสร้างส่วนย่อยในการทำอีโรชัน

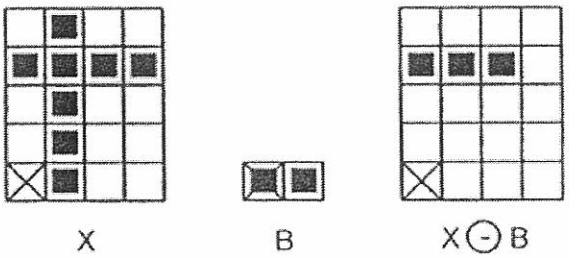
ตัวอย่าง การทำอีโรชันกำหนดให้

$$X = \{(1,0),(1,1),(1,2),(0,3),(1,3),(2,3),(3,3),(1,4)\}$$

$$B = \{(0,0),(1,0)\}$$

$$X \ominus B = \{(0,3),(1,3),(2,3)\}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

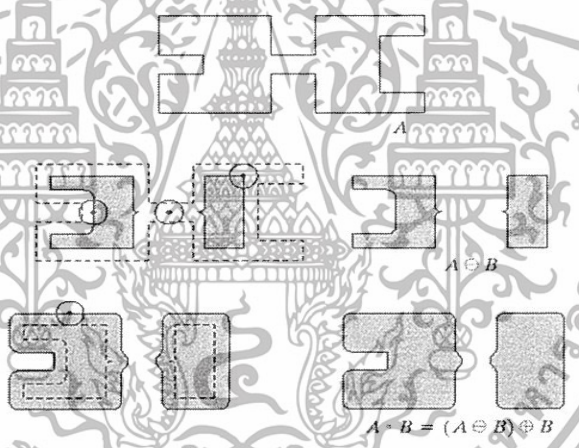


ภาพที่ 2.13 แสดงการทำอีโรชัน

2.4.1.3 การทำโอเพนนิง (Opening)

เป็นการทำงานร่วมกันระหว่างการทำอีรอชันและไดเลชันตามลำดับ คือจะเริ่มทำจากการทำอีรอชันโดยลดจำนวนจุดภาพขาว แล้วตามด้วยการทำไดเลชันโดยเพิ่มจำนวนจุดขาวลงไปในภาพให้เท่ากับจำนวนรอบของการทำอีรอชัน เพื่อทำให้ลดสัญญาณรบกวนขนาดเล็กในรูปภาพหายไป

$$A \circ B = (A \ominus B) \oplus B \tag{2.5}$$



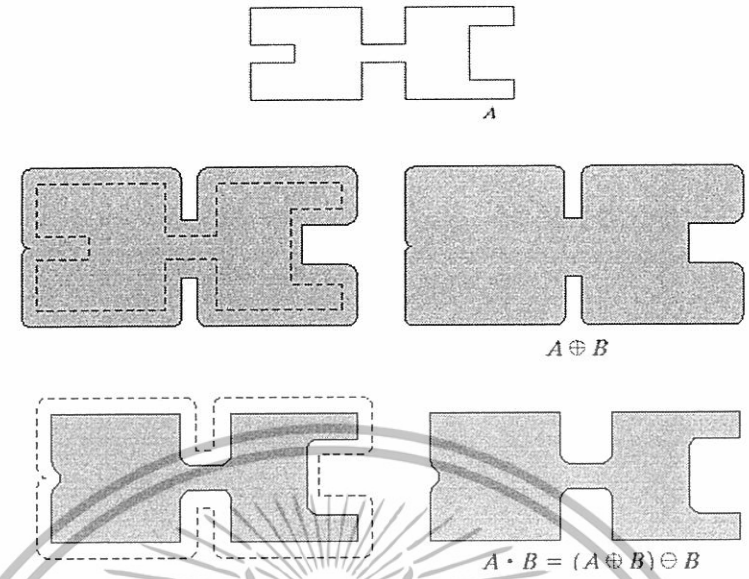
ภาพที่ 2.14 แสดงการทำโอเพนนิง

2.4.1.4 การทำโคลสซิง (Closing)

เป็นการลดสัญญาณรบกวนประเภทหลุมขนาดเล็กสามารถทำได้โดยผ่านกระบวนการประมวลผลภาพดิจิทัล คือ การทำไดเลชันและอีโรชัน หากนำ 2 กระบวนการนี้มาใช้ร่วมกัน ในลำดับการทำงานของไดเลชันและตามด้วยอีโรชัน จะเรียกว่า ตัวดำเนินการแบบปิด เพื่อทำให้ช่องว่าง หรือ ช่องว่างเล็ก ๆ ของ บริเวณที่เชื่อมต่อกัน หมดไป

$$A \bullet B = (A \oplus B) \ \ominus \ B \tag{2.6}$$

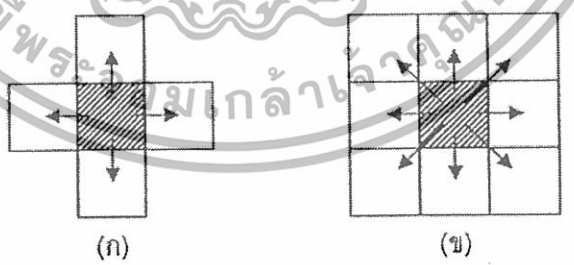
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
73073



ภาพที่ 2.15 แสดงการทำโคลสซิง

2.4.2 การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน (Connected Component Analysis)

ในการกำหนดขอบเขตและอธิบายรูปร่างของแต่ละวัตถุในรูปภาพฐานสอง วิธีหนึ่งที่ใช้กัน คือ วิธีแบบรหัสลูกโซ่ (Chain Code) โดยจะมีขั้นตอนการทำคือจะมีการกำหนดรหัสไว้ก่อนล่วงหน้า โดยแต่ละรหัสจะแสดงถึงการเปลี่ยนแปลงทิศทางของจุดภาพในบริเวณที่อยู่ติดกัน ซึ่งวิธีนี้เมื่อทำการเข้ารหัสรูปภาพฐานสองแล้วจะสามารถพิจารณาว่าในภาพนั้นมีจุดภาพกลุ่มใดบ้าง การกำหนดมีทั้งการกำหนดรหัสแบบ 4 จุดเชื่อมต่อและแบบ 8 จุดเชื่อมต่อ

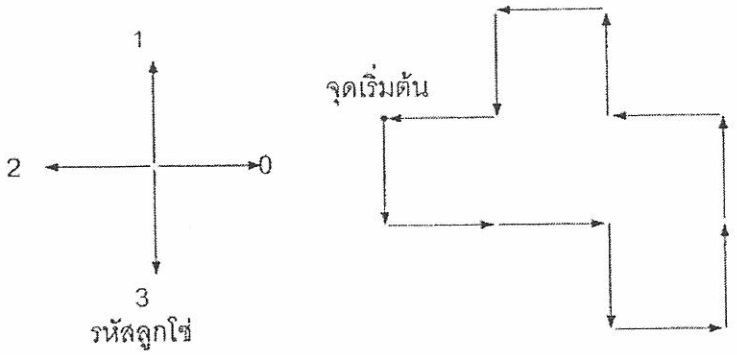


ภาพที่ 2.16 การกำหนดรหัสแบบ (ก) 4 จุดเชื่อมต่อ (ข) 8 จุดเชื่อมต่อ

โดยในการอธิบายรูปร่างของวัตถุ จะใช้ค่าอนุพันธ์มาใช้ในการบ่งบอกซึ่งค่าอนุพันธ์นี้หาได้จากการเปลี่ยนแปลงของลูกศรเป็นทิศ 90 หรือ 45 องศาและตามเข็ม หรือทวนเข็มตามนาฬิกา

<u>ตัวอย่าง</u> การเข้ารหัสลูกโซ่แบบ 4 จุดเชื่อมต่อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.17 แสดงการเข้ารหัสลูกโซ่แบบ 4 จุดและการแสดงการเข้ารหัส

จากตัวอย่างการเข้ารหัสลูกโซ่แบบ 4จุดเชื่อมต่อ โดยมีรหัสดังนี้คือ 3,0,0,3,0,1,1,2,1,2,3,2 และค่าอนุพันธ์คือ 1,0,3,1,1,0,1,3,1,1,3,1

2.4.3 การหาเส้นรอบรูป (Contour Finding)

การติดตามเส้นขอบเป็นการติดตามเส้นรอบรูปวัตถุในภาพ ซึ่งเป็นการไต่ไล่ไปตามเส้นรอบรูป โดยใช้ความแตกต่างของสีจุดภาพระหว่างวัตถุกับพื้นหลัง ซึ่งโดยทั่วไปแล้ววิธีการนี้จะใช้กับภาพไบนารีเช่น แทนจุดภาพดำเป็นส่วนของวัตถุ และแทนจุดภาพขาวเป็นส่วนของพื้นหลัง ดังนั้นในการทำงานจะตรวจกวาดไปทุกๆ จุดภาพ โดยเริ่มจากจุดภาพแรกที่มุมซ้ายบนของภาพ และตรวจกวาดไปในทิศทางจากซ้ายไปขวาและเลื่อนจากบนลงล่าง เมื่อตรวจกวาดมาพบจุดภาพดำจุดแรกซึ่งเป็นส่วนของวัตถุ จะเปลี่ยนลักษณะการเคลื่อนที่ไปยังจุดภาพถัดไปใหม่ โดยมีเงื่อนไขการเคลื่อนที่และเก็บลงหน่วยความจำ โดยใช้รูปแบบรหัสลูกโซ่ เช่นเดียวกับ การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน

2.5 การเขียนโปรแกรมเชิงวิชั่น (Vision Programming)

2.5.1 หลักการเขียนภาษาเชิงวัตถุ และการสร้างโปรแกรมด้วย MS Visual C++ และ MFC

การพัฒนาโปรแกรมเริ่มจากการใช้หลักการเขียนโปรแกรมแบบโครงสร้าง ที่มีการจำแนกปัญหาและเขียนโปรแกรมแยกเป็นส่วนๆ โดยเรียกว่าโมดูล (Module) แต่วิธีการนี้พบว่ามีปัญหาในการพัฒนาอยู่หลายประการ เช่น ถ้ามีการแก้ไขบางส่วนในโมดูลจะทำไม่ได้ จะต้องเขียนขึ้นมาใหม่ ทำให้เสียเวลาในการแก้ไขโปรแกรมมาก ซึ่งจะมีผลกระทบกับโมดูลอื่นๆด้วย ดังนั้นจึงได้มีวิธีการพัฒนาโปรแกรมแนวใหม่ขึ้น คือ Object Oriented Programming (OOP)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การเขียนโปรแกรมแบบ OOP หรือออบเจ็กต์โอเรียนเท็ดโปรแกรมมิ่ง เป็นแนวคิดในการเขียนโปรแกรมแบบหนึ่ง โดยจะไม่มุ่งเน้นไปในการเขียนโปรแกรมเพียงอย่างเดียว แต่เป็นการมองภาพรวมของสิ่งที่อยู่รอบๆ ตัวเราโดยสามารถให้คำจำกัดความได้ดังนี้ “วัตถุแต่ละอย่างต่างก็มีลักษณะ และวิธีการใช้งานเป็นของตัวมันเอง” กล่าวคือ วัตถุแต่ละชนิดแต่ละอัน ต่างก็มีรูปร่างลักษณะ และการใช้งานที่แตกต่างกันออกไป เราจะเรียกคุณลักษณะของวัตถุนี้ว่าคุณลักษณะ (Attribute) และเราจะเรียกวิธีการใช้งานวัตถุว่า วิธีการ (Method) ยกตัวอย่างเช่น จักรยานคันหนึ่งมีสีแดง สามารถใช้ขับขี่ได้ สีแดงจัดว่าเป็น คุณลักษณะ ของจักรยาน ในขณะที่การใช้งานขับขี่ถือเป็น วิธีการเป็นต้น

จากที่กล่าวมาข้างต้นจะเห็นได้ว่าการเขียนโปรแกรมแบบ OOP เป็นแนวคิดที่คล้ายธรรมชาติของสิ่งของสิ่งหนึ่งซึ่งเราสามารถแบ่งแยกสิ่งต่างๆ ออกเป็นประเภทๆ ได้ ถ้าเราได้นำเอาแนวคิดของ OOP มาใช้ในการเขียนโปรแกรม และการจัดการข้อมูล เราจะพบว่าโปรแกรม หรือฟังก์ชั่น จะมีความเป็นอิสระต่อกันอย่างเห็นได้ชัด อธิบายง่ายๆ ก็คือ โปรแกรม หรือฟังก์ชั่นแต่ละตัวถึงแม้จะมาจากที่เดียวกัน แต่สามารถทำงานในคนละหน้าที่หรือเก็บข้อมูลคนละค่าได้

คลาสใน OOP คือกลุ่มของวัตถุที่คุณลักษณะพื้นฐาน และมีฟังก์ชั่นพื้นฐานเหมือนกัน เช่น คลาสคน กล่าวคือไม่ว่าคนนั้นจะมีชื่อใด มีอาชีพใด หรือมีส่วนสูงเท่าไหร่ ก็จัดอยู่ในคลาสคน ซึ่งจะมีคุณลักษณะ และฟังก์ชั่นพื้นฐานเหมือนกัน โดยผู้เขียนโปรแกรมไม่สามารถนำคลาสไปใช้งานได้โดยตรง ต้องมีการสร้าง ออปเจค (หรือ Instance) ขึ้นมาก่อน ซึ่ง ออปเจค ที่สร้างจากคลาสใดจะมีคุณลักษณะพื้นฐานมาจากคลาสนั้นด้วย เพราะฉะนั้นการสร้างโปรแกรมด้วยวิธีการแบบ OOP จะทำให้การสร้างโปรแกรมมีความยืดหยุ่น และมีประสิทธิภาพสูง

คุณสมบัติพิเศษอีกอย่างหนึ่งที่ทำให้การเขียนโปรแกรมในแบบ OOP คือคุณสมบัติ การสืบทอด (Inheritance) การที่ OOP นำแนวคิดนี้มาใช้ทำให้ผู้เขียนโปรแกรมสามารถนำเอา Source Code (อาจเป็นคลาส) ที่เขียนไว้แล้วสามารถนำกลับมาใช้ใหม่ (Code Reuse) หรือมาเปลี่ยนแปลงแก้ไขให้เหมาะสมกับงานเฉพาะอย่าง หมายว่า คลาสแต่ละคลาส (คลาสแม่) สามารถที่จะสืบทอดออกมาเป็นคลาสใหม่ได้ (คลาสลูก) โดยคลาสใหม่ที่สืบทอดออกมานี้ จะยังคงมีคุณสมบัติเหมือนกับคลาสแม่ทุกประการ นอกจากนี้เรายังสามารถที่จะเพิ่มเติมคุณลักษณะใหม่ที่ไม่มีอยู่ในคลาสแม่ลงไปในคลาสลูกได้ ซึ่งทำให้การเขียนโปรแกรมเป็นไปได้อย่างรวดเร็วและเป็นการใช้โค้ด ที่มีอยู่ให้เกิดประโยชน์สูงสุด

OOP มีการกำหนดระดับการเข้าถึงข้อมูลภายในคลาส เพื่อความปลอดภัยในการเข้าถึงข้อมูลภายในคลาส ซึ่งการกำหนดระดับการเข้าถึงข้อมูลภายในคลาสนี้เป็นคุณสมบัติอีกอย่างที่มีในแนวคิดแบบ OOP การเข้าถึงข้อมูลภายใน (Member Variable) หรือฟังก์ชั่นภายใน (Member Function หรืออาจเรียกว่า Method) แบ่งออกได้ 3 ระดับ ดังนี้คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. ระดับ Private เป็นระดับที่จะถูกกำหนดไว้อัตโนมัติ (Default) หากผู้เขียนโปรแกรมไม่กำหนด ระดับนี้เป็นระดับการป้องกันสูงสุด กล่าวคือเป็นการป้องกันไม่ให้กระบวนการใดๆ ที่อยู่นอกคลาสเรียกใช้สิ่งที่อยู่ในคลาสไม่ได้ การจะเข้าถึงข้อมูลภายในจะต้องผ่านจากฟังก์ชั่นที่เป็นของคลาสนั้นอนุญาตให้ใช้ได้

2. ระดับ Public เป็นระดับที่ตรงข้ามกับระดับ Private กล่าวคือกระบวนการใดๆ ที่อยู่นอกคลาสนั้นสามารถเข้าถึงข้อมูลภายใน หรือฟังก์ชั่นภายในได้อย่างอิสระ ไม่จำเป็นต้องได้รับอนุญาตก่อนการเข้าใช้งาน หรือเรียกอีกอย่างนึงว่า การประกาศสมาชิกแบบสาธารณะ

3. ระดับ Protected คล้ายกับในระดับ Private แต่มีข้อแตกต่างคือระดับ Protected นั้น คลาสลูกที่สืบทอดคุณลักษณะจากคลาสแม่ จะสามารถเข้าถึงข้อมูลภายในคลาสแม่ได้โดยตรง

ความสามารถในการกำหนดระดับการเข้าถึงข้อมูลภายในคลาสอาจเรียกว่า คุณสมบัติการห่อหุ้มวัตถุ (Data Hiding)

การเขียนโปรแกรมแบบ OOP ผู้เขียนโปรแกรมสามารถใช้คุณสมบัติของฟังก์ชั่นคอนสตรัคเตอร์ (Constructor) และ ดีสตรัคเตอร์ (Destructor)

- ฟังก์ชั่นคอนสตรัคเตอร์ เป็นฟังก์ชั่นที่จะถูกเรียกให้ทำงานเมื่อมีการสร้างออบเจ็กต์ของคลาสเกิดขึ้น
- ฟังก์ชั่นดีสตรัคเตอร์ เป็นฟังก์ชั่นที่จะถูกเรียกให้ทำงาน เมื่อสิ้นสุดการทำงานของออบเจ็กต์ในคลาสนั้นๆ เช่น จบโปรแกรม หรือเมื่อมีการลบออบเจ็กต์ออกจากหน่วยความจำ

เพื่อเพิ่มความยืดหยุ่นให้กับโปรแกรมได้ในการสร้างคลาสแต่ละครั้งเราไม่จำเป็นต้องสร้างคอนสตรัคเตอร์และดีสตรัคเตอร์ทุกครั้งไป การสร้างคอนสตรัคเตอร์และดีสตรัคเตอร์นั้นขึ้นอยู่กับความจำเป็นในแต่ละสถานการณ์และการดำเนินงานของคลาส และในคลาสหนึ่งสามารถมีคอนสตรัคเตอร์ได้หลายตัวโดยฟังก์ชั่นคอนสตรัคเตอร์ต้องมีชื่อฟังก์ชั่นเหมือนชื่อของคลาส และ ไม่มีการคืนค่าใดๆออกมาเลย ฟังก์ชั่นคอนสตรัคเตอร์จะถูกเรียกโดยอัตโนมัติเมื่อมีการสร้างวัตถุและการสร้าง Object สามารถผ่านพารามิเตอร์ (หรือไม่ผ่านก็ได้เป็น default) ซึ่ง คอมไพเลอร์จะเรียกฟังก์ชั่นคอนสตรัคเตอร์ได้ตรงกัน คุณสมบัตินี้เรียกว่า สามารถโอเวอร์โหลดฟังก์ชั่นได้

โพลีมอฟีซึ่ม (Polymorphism) คือ การที่ตัวแปรอ้างอิงนึง สามาถอ้างอิงกับออบเจ็กต์ชนิดอื่นได้และสามารถเรียกเมทธอดที่เข้ากับชนิดออบเจ็กต์ได้อย่างถูกต้องโดยอัตโนมัติในระหว่างที่โปรแกรมทำงาน

การเขียนโปรแกรมแบบ OOP ก็มีลักษณะเช่นเดียวกับการเขียนโปรแกรมแบบโครงสร้างโดยใช้หลักการที่คล้ายกัน แต่ OOP นั้น จะมีประสิทธิภาพที่สูงกว่า เพราะเป็นการนำเอา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คุณลักษณะ (Attribute) และวิธีการ (Method) เช่น ข้อมูล, ตัวแปร หรือฟังก์ชั่นของวัตถุนั้นมารวมไว้ในกลุ่มๆเดียวกัน ที่เรียกว่า คลาส (Class)

MFC (Microsoft Foundation Class) เป็นกลุ่มของคลาสที่ออกแบบเพื่อช่วยให้การเขียนโปรแกรมทำได้ง่ายขึ้น คลาสภายใน MFC ก็จะมีการสืบทอดกันมา หรือ เป็นคลาสที่ไม่ได้สืบทอดจากคลาสใดเลยก็มี MFC นั้นถือว่าเป็นส่วนนึงของ Visual C++ ก็ว่าได้ เพราะ Visual C++ เป็นเครื่องมือในการพัฒนาโปรแกรมบนวินโดวส์ที่มี MFC เป็นไลบรารี่ (library) ที่ช่วยอำนวยความสะดวก

ข้อดีของการเขียนโปรแกรมด้วย Visual C++ และ MFC คือ ความสามารถในการพอร์ตเทเบิล (Portability) หมายว่า สามารถที่จะทำการคอมไพล์และลิงค์กับ Visual C++ ในเวอร์ชั่นที่ต่ำกว่ากับตัว Visual C++ เวอร์ชั่นที่สูงกว่าได้

นอกจากนี้ยังมีคุณสมบัติ Dynamic Binding (Late Binding) ซึ่งเป็นคุณสมบัติพิเศษที่ทำให้การเขียนโปรแกรมแบบ OOP แตกต่างจากการเขียนแบบโครงสร้าง ซึ่งจะไม่ขอกล่าวในรายงานฉบับนี้

2.5.2 หลักการทำงานของ OpenCV

Open Source Computer Vision Library หรือที่รู้จักกันว่า OpenCV นั้นเป็นไลบรารีสำหรับใช้งานเรื่องการประมวลผลภาพ (image processing) ซึ่งสามารถค้นหาข้อมูลและนำมาใช้งาน เพื่อพัฒนาโปรแกรมทางด้าน Image Processing, การประมวลผลแบบเวลาจริง (Real-time Image Processing) และ คอมพิวเตอร์วิชั่น (Computer Vision) บนระบบปฏิบัติการ Microsoft Windows ได้โดยไม่เสียค่าใช้จ่าย จากบริษัท Intel ซึ่งเป็นผู้พัฒนาขึ้น ซึ่งก่อนหน้าที่จะพัฒนา OpenCV ไลบรารีก็ได้พัฒนา VisualSDK แต่ไม่เป็นที่นิยมมากนัก ความสามารถของ OpenCV ในไลบรารีที่นิยมใช้กันมาก คือ การทำเบลอให้กับไฟล์ภาพ การอุดช่องว่างภายในรูป (threshold) ค้นหาขอบภาพ ตรวจสอบการเคลื่อนไหว การทำการตัดแยกรูปภาพ นอกจากนั้น OpenCV ยังสามารถทำงานได้กับไฟล์วิดิทัศน์ที่ ซึ่งไลบรารี่ของโปรแกรมฟังก์ชั่นหลักที่ใช้เพื่อมุ่งหลักไปที่การประมวลผลในแบบเวลาจริง

OpenCV มีคลาสและฟังก์ชั่นในภาษา C++ ซึ่งใช้สำหรับการทำงานที่เกี่ยวข้องกับงานด้านภาพ ผู้เขียนโปรแกรมที่ติดตั้ง OpenCV สามารถพัฒนาโปรแกรม โดยใช้ MS Visual C++ ในชุด Visual Studio ได้อย่างอิสระ โดยไม่ขึ้นกับอุปกรณ์ต่อเชื่อมที่นำสัญญาณเข้าเครื่องคอมพิวเตอร์

ปัจจุบันทาง บริษัท Intel ได้ทำการเผยแพร่ออกมา OpenCV beta5 ซึ่งคุณสมบัติที่เพิ่มมากขึ้นจากเดิม คือมี Python binding มาให้ด้วย ซึ่งปกติทาง OpenCV จะเป็นเพียงชุดของคำสั่งเท่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การใช้งานต้องเขียนโปรแกรมไปเรียกคำสั่งในไลบรารีที่มีมาให้ โดยสามารถแยกได้เป็น 2 ภาษา และมีข้อดีข้อเสียดังต่อไปนี้

- ข้อดีของการพัฒนากับภาษา C++

  ใช้เวลาในการประมวลผลเร็วมากและเป็นที่นิยมในหมู่นักพัฒนาด้านการประมวลผลภาพ (Image processing)

- ข้อดีของการพัฒนากับ ภาษา Python

  เป็นภาษาที่ใช้พัฒนาได้ในเวลาไม่นาน เมื่อเปรียบเทียบกับภาษา C++ การเขียนคำสั่งจะน้อยกว่าประมาณ 5 เท่า

2.5.3 การเลือกภาษาในการพัฒนาโปรแกรม

การพัฒนาโปรแกรมทางด้านการประมวลผลภาพ หากพบว่าโปรแกรมที่จะดำเนินการจัดทำนั้นเป็นโปรแกรมที่มีการทำงานไม่ซับซ้อนมากและไม่คำนึงถึงเวลาในการประมวลผลมากนัก ภาษาที่เหมาะสมในการพัฒนาคือ ภาษา Python แต่ถ้าหากโปรแกรมที่ต้องการพัฒนามีการทำงานที่ค่อนข้าง มาก และต้องการใช้เวลาเวลาในการประมวลผลที่สั้น ภาษาที่เหมาะสมในการพัฒนาคือ ภาษา C++

2.5.4 ความสัมพันธ์ระหว่าง OpenCV กับ ไลบรารีอื่นๆ

OpenCV ออกแบบมาให้ใช้งานร่วมกับอินเทล IPL (Image Processing Library) และยังสามารถทำการขยายการทำงานไปยังรูปภาพ และ รูปแบบการวิเคราะห์ (pattern analysis) ฉะนั้น OpenCV จึงสามารถใช้งานร่วมกับรูปแบบรูปภาพอื่นๆที่ใกล้เคียงกันโดยใช้งาน IPL

OpenCV ทำงานบนคอมพิวเตอร์ส่วนบุคคลซึ่งต้องอยู่บนพื้นฐานของหน่วยประมวลผลสถาปัตยกรรมอินเทลและยังสามารถทำงานบนไมโครซอฟวินโดว์ 95, 98, 2000, NT หรือ XP และสามารถใช้งานบนโปรแกรมประยุกต์ MS Visual C++ ซึ่งจะมีคำอธิบายฟังก์ชัน (Function description) ที่จะเรียกใช้

2.6 แอนนิเมชั่น (Animation)

แอนนิเมชั่นคือการเคลื่อนที่ของภาพจากภาพหนึ่งไปเป็นอีกภาพหนึ่งด้วยความเร็วในระดับหนึ่ง เรียกภาพแต่ละภาพที่ถูกเปลี่ยนไปนั้นว่าเฟรม การเคลื่อนไหวของแอนนิเมชั่นเพียง 1 วินาทีนั้น อาจจะเกิดจากภาพหลายๆ ภาพก็เป็นได้ เช่น ภาพยนตร์ที่เราเห็น (ที่ฉายจากเครื่องฉายประเภทฟิล์ม) เกิดจากการเปลี่ยนของภาพทั้งหมด 24 ภาพ ใน 1 วินาที หรือเรียกว่า 24 เฟรมต่อวินาที และยิ่งจำนวน เฟรมต่อวินาทียิ่งมากขึ้น การเคลื่อนไหวที่เห็นก็จะยิ่งดูลื่นไหลมากขึ้นเท่านั้น และภาพที่สามารถเป็นอนิเมชั่นที่ดีได้ต้องมีลักษณะภาพในแต่ละเฟรม มีการเคลื่อนที่ใกล้เคียงกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยจำนวนเฟรมที่เหมาะสมที่สุดในการนำมาทำแอนนิเมชั่นที่นิยมในปัจจุบัน คือ 30 เฟรมต่อ 1 วินาที ซึ่งจำนวนเฟรมที่น้อยที่สุดต่อวินาทีที่ทำให้ สายตาของมนุษย์มองภาพแล้วเกิดการเคลื่อนไหวอย่างต่อเนื่อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การออกแบบ และการพัฒนาระบบ

3.1 ความต้องการระบบ

โครงสร้างของโครงงานระบบช่วยฝึกการออกกำลังกายเบื้องต้นโดยการประยุกต์ใช้เทคนิคการตรวจจับการเคลื่อนไหวของมนุษย์จากวีดิทัศน์ จะประกอบด้วย กล้องวีดิทัศน์ซึ่งใช้ในการจับภาพการเคลื่อนไหวของผู้ใช้ โดยผู้ใช้จะทำท่าทางตามการ์ตูนแอนนิเมชั่น ซึ่งตัวการ์ตูนจะเป็นผู้แสดงท่าทางการออกกำลังกายและมีหน้าที่โต้ตอบกับผู้ใช้โดยใช้คำพูด เช่นการกล่าวชมเมื่อผู้ใช้ทำท่าทางถูกต้องตามการ์ตูนแอนนิเมชั่น และเปลี่ยนท่าทางเมื่อผู้ใช้ทำท่าทางถูกต้องดังภาพที่ 3.1

ภาพที่ 3.1 แสดงโครงสร้างของโครงงาน

3.2 การออกแบบระบบ

จากการวิเคราะห์ความต้องการระบบดังกล่าว สามารถออกแบบขั้นตอนการทำงานของระบบได้ ดังภาพที่ 3.2 กล่าวคือ เมื่อเปิดโปรแกรมระบบจะเริ่มจากหน้าจอเริ่มต้น โดยเมื่อผู้ใช้กดปุ่ม Start ระบบจะทำการเรียนรู้ฉากหลัง (Learn Background) และเริ่มตรวจจับการเคลื่อนไหว เมื่อผู้ใช้เข้าในส่วนที่สามารถตรวจจับภาพได้ ระบบก็จะเริ่มวิเคราะห์ท่าทางการเคลื่อนไหวของผู้ใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หากการเคลื่อนไหวถูกต้องระบบจะกล่าวชมและเปลี่ยนท่าการเต้น แต่หากผู้ใช้ทำท่าทางไม่ตรงกับการ์ตูนแอนนิเมชั่นที่นำเต้นระบบจะให้ทำท่าทางนั้นซ้ำอีกจนกว่าผู้ใช้จะทำถูก เมื่อทำครบทุกท่าทางที่มีอยู่ในระบบหลังจากนั้นระบบจะสิ้นสุดการนำเต้น

ระบบจะรับข้อมูลการเคลื่อนไหวของผู้ใช้จากวีดิทัศน์และนำไปประมวลผลโดยการตัดผู้ใช้ออกจากฉากหลังด้วยวิธีลบฉากหลังออก หลังจากนั้นก็จะนำไปวิเคราะห์ท่าทาง ซึ่งผลที่ได้จากการวิเคราะห์จะใช้ในการทำโต้ตอบ (Feedback) กับผู้ใช้หากผู้ใช้ทำท่าทางที่ถูกต้อง จะแสดงการการสื่อสารออก ทางหน้าจอ (Monitor) เป็นแบบการ์ตูนแอนนิเมชั่น (Animation) ผ่านส่วนควบคุม (Controller) ที่รับภาพวีดิทัศน์ (Video Camera) แล้วนำมาทำการเลือกข้อมูลพื้นที่ที่สนใจ (Motion Segmentation) และวิเคราะห์ท่าทาง (Motion analysis)

ภาพที่ 3.2 แผนภาพแสดงลักษณะการทำงานของระบบ

3.3 การสร้างระบบ

3.3.1 เครื่องมือที่ใช้

ประกอบด้วยส่วนต่างๆซึ่งมีรายละเอียดดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- การแสดงภาพการ์ตูนแอนนิเมชั่น และส่วนการโต้ตอบกับผู้ใช้อาศัย ActiveX Control Shockwave Flash Object ในการแสดงผลไฟล์ Flash

- การรับข้อมูลวีดิทัศน์ โดยอาศัยไลบารี่ OpenCV ทำหน้าที่เป็นตัวกลางในการติดต่อระหว่างโปรแกรม กับ อุปกรณ์ต่อเชื่อม

- การตัดผู้ใช้ออกจากฉากหลังใช้เทคนิคลบฉากหลัง (Horprasert, 1999) และทำการลบข้อมูลที่ไม่ใช่ส่วนที่ต้องการ โดยใช้เทคนิคประมวลผลภาพ โดยการทำมอฟลอโลจี (Morphology) เพื่อลบส่วนที่เป็นสัญญาณรบกวน (noise) และการหา เส้นรอบรูป (Contour) เพื่อเลือกโครงร่างส่วนหลักที่ติดกันมาประมวลผล

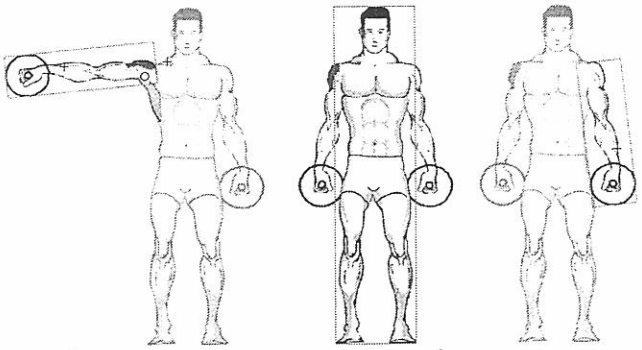
3.3.2 ด้านการทำการ์ตูนแอนนิเมชั่น

ทำการสร้างไฟล์แอนนิเมชั่นโดยใช้โปรแกรม Macromedia Flash โดยในการวาดการ์ตูนจะวาดเป็นภาพลายเส้น (vector) เนื่องจากเมื่อมีการขยายภาพแล้วจะทำให้แอนนิเมชั่นที่ต้องการใช้งานภาพไม่แตก ในส่วนของการสร้างฉากหลัง จะเลือกสีฉากหลังแบบเรียบง่ายเพื่อดึงดูดจุดสนใจของผู้ใช้งานไปที่ตัวการ์ตูนแอนนิเมชั่นที่จะสร้างขึ้น ดังภาพที่ 3.3

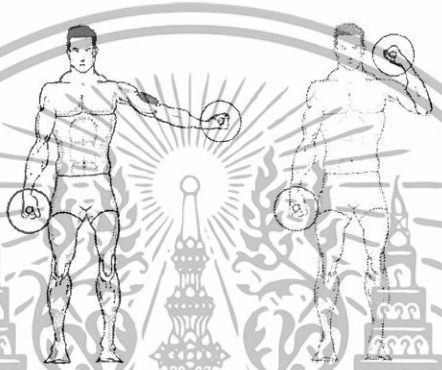
ภาพที่ 3.3 แสดงฉากของแอนนิเมชั่น

ในการทำการ์ตูนแอนนิเมชั่นจะใช้หลักการแยกส่วนอวัยวะที่จะทำการเคลื่อนไหวออกเป็นวัตถุ ตัวอย่างเช่น การทำให้การ์ตูนแอนนิเมชั่น กางแขน จะแยกเป็นทั้งหมดสามส่วนคือ ส่วนของร่างกายทั้งหมด ตั้งแต่ส่วนหัวของตัวการ์ตูนจนถึงเท้าของตัวการ์ตูน แขนข้างขวา และแขนข้างซ้าย ดังภาพที่ 3.4

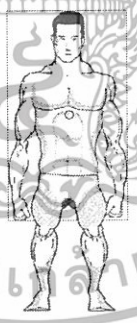
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.4 แสดงภาพตัวการ์ตูนแอนนิเมชั่นท่ากางแขนโดยแยกเป็นวัตถุ



ภาพที่ 3.5 แสดงภาพตัวการ์ตูนแอนนิเมชั่นท่างอข้อศอก

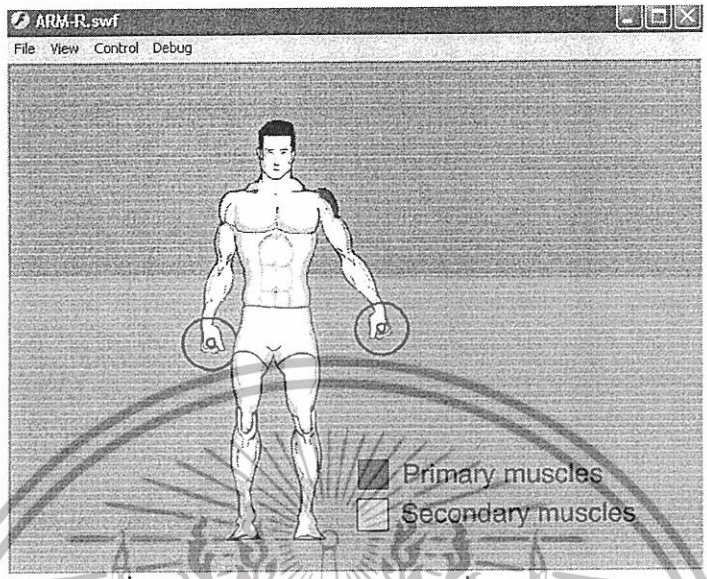


ภาพที่ 3.6 แสดงภาพตัวการ์ตูนแอนนิเมชั่นท่าย่อตัว

จากภาพที่ 3.5 และภาพที่ 3.6 จะเห็นได้ว่าการแยกอวัยวะส่วนต่างๆของร่างกายจะขึ้นอยู่กับท่าทางที่แสดง โดยภาพที่ 3.5 จะแยกเป็นทั้งลำตัวรวมถึงท่อนแขนด้านบน และท่อนแขนด้านล่าง ส่วนภาพที่ 3.6 จะแยกส่วนของอวัยวะเป็นส่วนตั้งแต่หัวเข่าขึ้นไป และส่วนของหัวเข่าลงมาจนถึงเท้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

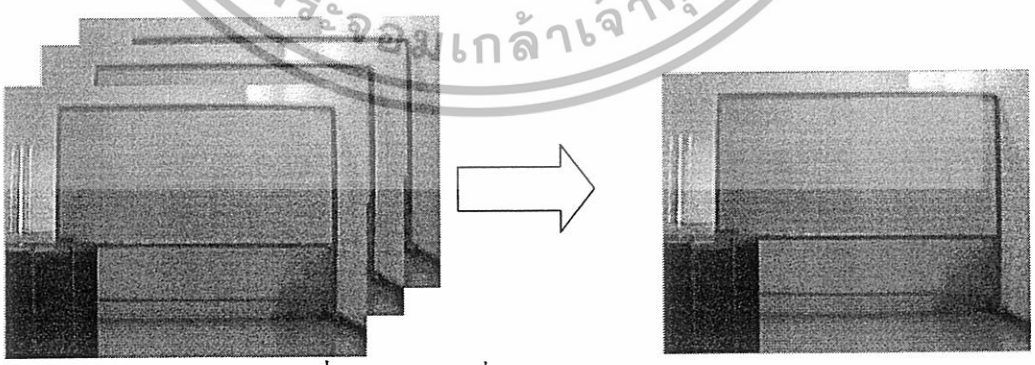
การรวมฉากหลังและการ์ตูนแอนนิเมชั่น จะนำทั้งสองส่วนมาอยู่ในไฟล์เดียวกัน โดยจะทำการแยกระหว่างฉากหลังและตัวการ์ตูนแอนนิเมชั่นออกเป็นชั้น (Layer) ดังภาพที่ 3.7



ภาพที่ 3.7 แสดงภาพการ์ตูนแอนนิเมชั่นพร้อมฉากหลัง

3.3.3 การทำการประมวลผลภาพ

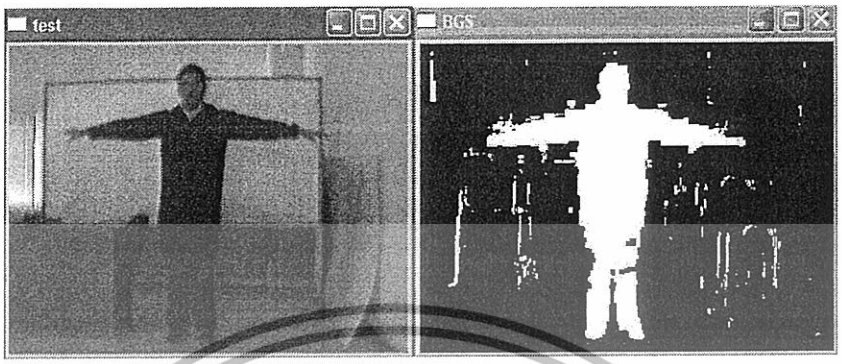
ในการสร้าง ฉากหลัง (Background Model) เมื่อเริ่มต้นโปรแกรมจะแสดงภาพที่ได้จากวีดีทัศน์ หรือเป็นข้อมูลเข้าซึ่งส่วนนี้จะแสดงออกมาแบบเวลาจริง ในส่วนการสร้างฉากหลังจะทำโดยเริ่มจากการสร้าง แบบของฉากหลัง (Background Model) ขึ้นมา โดยทำการเรียนรู้จากจำนวนเฟรมที่ทำการเรียนรู้ (Learn) โดยจะใช้การหาค่าเฉลี่ยของแต่ละจุดภาพจากภาพที่รับเข้าจำนวนหนึ่ง (ในระบบนี้ตั้งค่าไว้ 50 ภาพ) เพื่อนำมาสร้างเป็นตัวแบบของฉากหลัง เมื่อสร้างเสร็จจะแสดงผลฉากหลัง ที่ถูกสร้าง ณ เวลาที่ไม่มีผู้ใช้อยู่ในจอภาพ ดังภาพที่ 3.8



ภาพที่ 3.8 ฉากหลังที่ได้จากการสร้างฉากหลัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

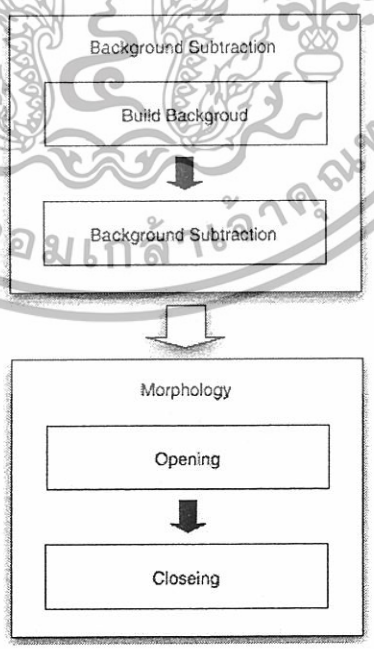
ในการแยกส่วนที่ตัดแยกผู้ใช้ออกจากฉากหลัง (Background Subtraction) นั้นจะทำการนำเอาภาพที่เป็นข้อมูลเข้าที่ได้จากวีดีทัศน์ มาทำการเปรียบเทียบกับส่วนของฉากหลังที่ได้สร้างไว้แล้ว โดยจะนำมาเปรียบเทียบและลบส่วนที่ซ้ำกันออกไป ดังภาพที่ 3.9



(a) (b)

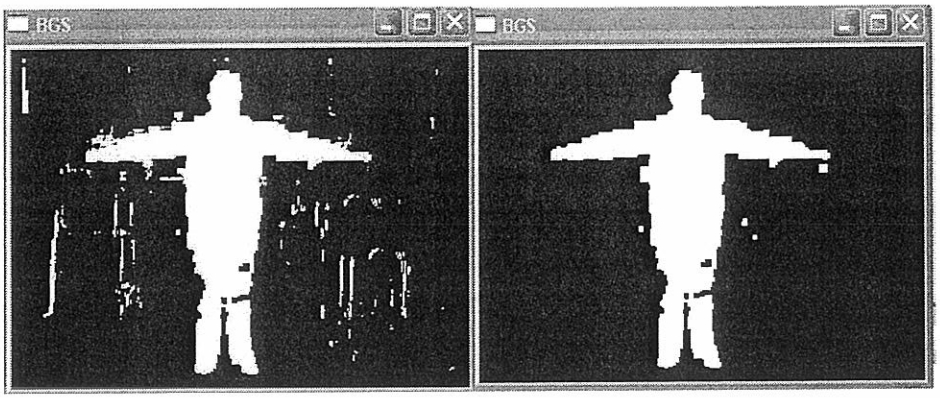
ภาพที่ 3.9 การตัดแยกผู้ใช้ออกจากฉากหลัง (a) แสดงข้อมูลที่ได้จากวีดีทัศน์ (b) แสดงผลลัพธ์ที่ได้จากการตัดแยกผู้ใช้ออกจากฉากหลัง

ผลลัพธ์ที่ได้จากการตัดแยกผู้ใช้ออกจากฉากหลังนั้นอาจจะมีสัญญาณรบกวน (Noise) เกิดขึ้นเป็นจำนวนมาก ดังนั้นจึงต้องทำการกำจัดสัญญาณรบกวน ที่เกิดขึ้นออกไปก่อนที่จะนำภาพไปประมวลผลต่อ โดยใช้วิธีการมอร์ฟอโลจี (Morphology) แบบโอเพนนิง (Opening) และโคลสซิง (Closing) ในการกำจัดสัญญาณรบกวนให้หมดไปหรือให้เหลือน้อยที่สุด ดังภาพที่ 3.10



ภาพที่ 3.10 แสดงกระบวนการตัดแยกผู้ใช้ออกจากฉากหลัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(a) (b)

ภาพที่ 3.11 การกำจัดสัญญาณรบกวน (a) แสดงภาพก่อนการกำจัดสัญญาณรบกวน (b) แสดงผลลัพธ์ที่ได้จากการกำจัดสัญญาณรบกวน

วิธีในการหาตำแหน่งของตัวบุคคล มีเทคนิคอยู่ 2 วิธีคือ เทคนิคการหาจุดเชื่อมต่อ และเทคนิคการหาเส้นรอบรูปภาย ซึ่งจะให้ผลออกมาดังนี้

เทคนิคการหาจุดเชื่อมต่อ (Connected Component) เพื่อวิเคราะห์หาองค์ประกอบทั้งหมดที่มีอยู่ในรูปภาพ ดังภาพที่ 3.12

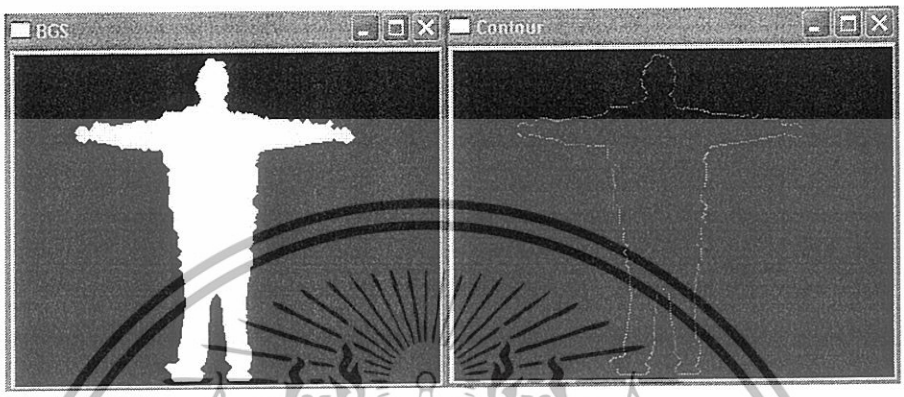
(a) (b)

ภาพที่ 3.12 การหาองค์ประกอบภายในภาพ (a) แสดงภาพก่อนการหาองค์ประกอบภายในภาพ (b) แสดงผลลัพธ์ที่ได้ จากการหาองค์ประกอบภายในภาพ

จะเห็นว่าจากภาพอินพุตเข้าจากภาพที่ 3.12 (a) นั้นนอกจากจะทำการตัดตัวผู้ใช้ออกมาแล้ว นั้นยังได้สัญญาณรบกวนที่กำจัดไม่หมดติดมาด้วย ดังนั้นจึงใช้วิธีการหาองค์ประกอบภายในภาพเข้ามาช่วยในการเลือกเฉพาะส่วนที่เราสนใจในภาพเพื่อไปทำการวิเคราะห์ ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เทคนิคการหาเส้นรอบรูป (Contour Finding) นั้นจะนำผลลัพธ์ที่ได้จากการกำจัดสัญญาณรบกวนออกจากภาพแล้ว มาทำการติดตามหาเส้นขอบเป็นการติดตามรอยขอบของวัตถุในภาพ ซึ่งเป็นการไต่ไล่ไปตามขอบ โดยใช้ความแตกต่างของสีจุดภาพระหว่างวัตถุกับพื้นหลัง เพื่อระบุโครงร่างทั้งหมดที่เกิดขึ้น จากนั้นจึงทำการเลือกโครงร่างที่สนใจออกมาเพื่อใช้ในกระบวนการต่อไป ดังภาพที่ 3.13



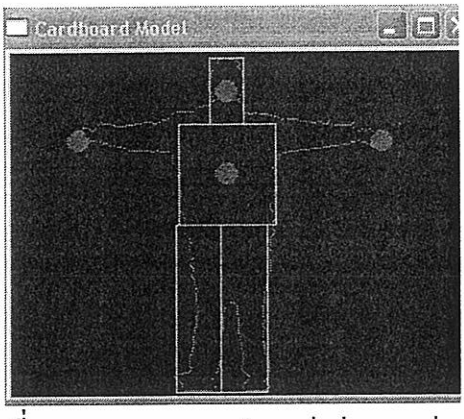
(a) (b)

ภาพที่ 3.13 การหาเส้นรอบรูปภายในภาพ (a) แสดงภาพก่อนการหาเส้นรอบรูปภายในภาพ (b) แสดงผลลัพธ์ที่ได้ จากการหาเส้นรอบรูปภายในภาพ

ในโครงงานนี้ได้เลือกใช้เทคนิคการหาเส้นรอบรูปในการใช้งาน โดยที่เลือกใช้เทคนิคนี้เพราะเป็นวิธีที่สามารถหาตำแหน่งของตัวบุคคลได้อย่างถูกต้องแม่นยำ

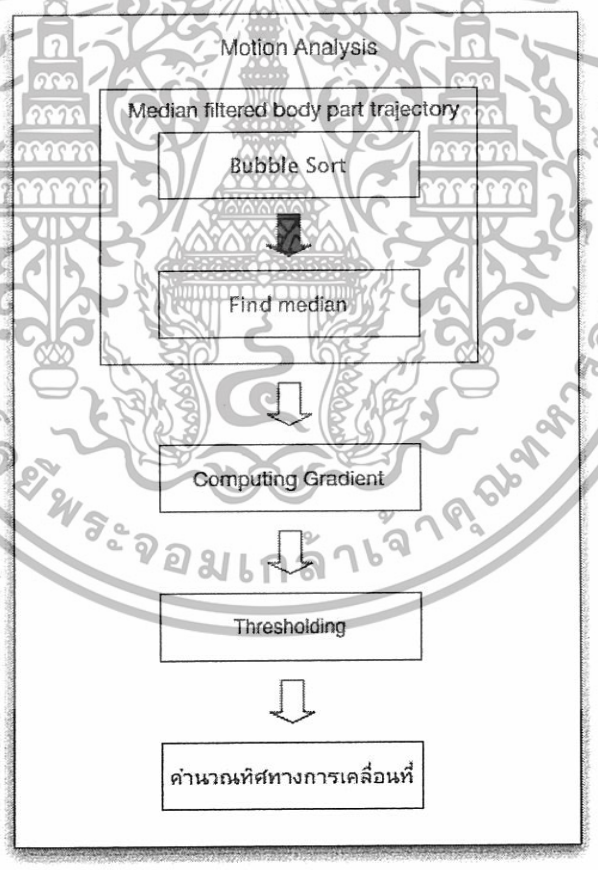
ในส่วนการระบุตำแหน่งส่วนของร่างกาย (Body part localization) นั้นจะทำการเอาโครงร่างที่ได้มาทำการระบุส่วนของร่างกายที่เราจะนำมาใช้ในการวิเคราะห์ลงไป โดยนำเทคนิคแบบจำลองมนุษย์คาร์ดบอร์ด (Cardboard Model) มาใช้ คือ ทำการระบุส่วนศรีษะ, ลำตัว, แขนซ้าย และแขนขวา เพื่อจะนำมาใช้สร้างตำแหน่งไว้อ้างอิงถึงตำแหน่งพิเซลของส่วนต่างๆของร่างกายได้เพื่อสะดวกที่จะนำไปใช้ในการวิเคราะห์ ดังภาพที่ 3.14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.14 แสดงการระบุตำแหน่งส่วนของร่างกาย

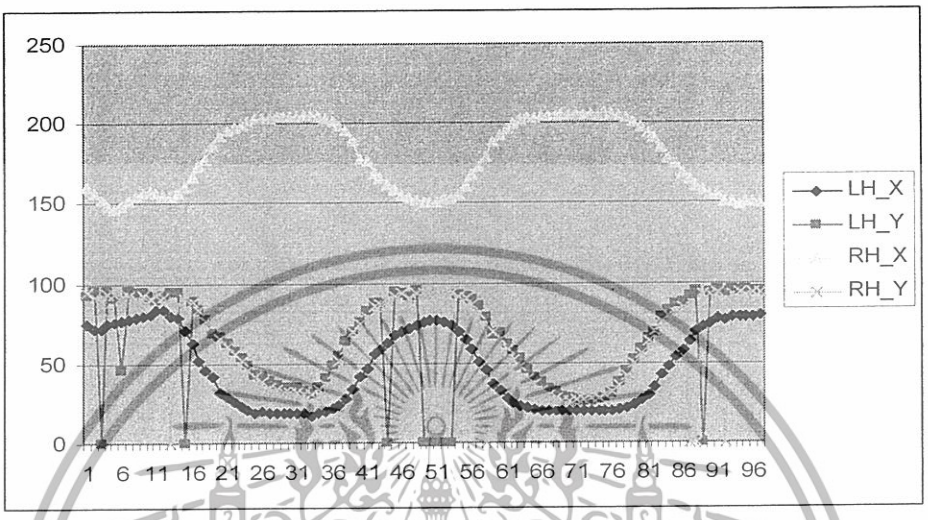
ในส่วนการวิเคราะห์การเคลื่อนไหว (Motion analysis) จะนำเอาค่าตำแหน่งของแต่ละส่วนของร่างกายที่ได้มาจากขั้นตอนการระบุตำแหน่งมาทำการวิเคราะห์ว่าส่วนนั้นมีการเคลื่อนไหวในลักษณะใด



ภาพที่ 3.15 แสดงขั้นตอนการวิเคราะห์การเคลื่อนไหว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

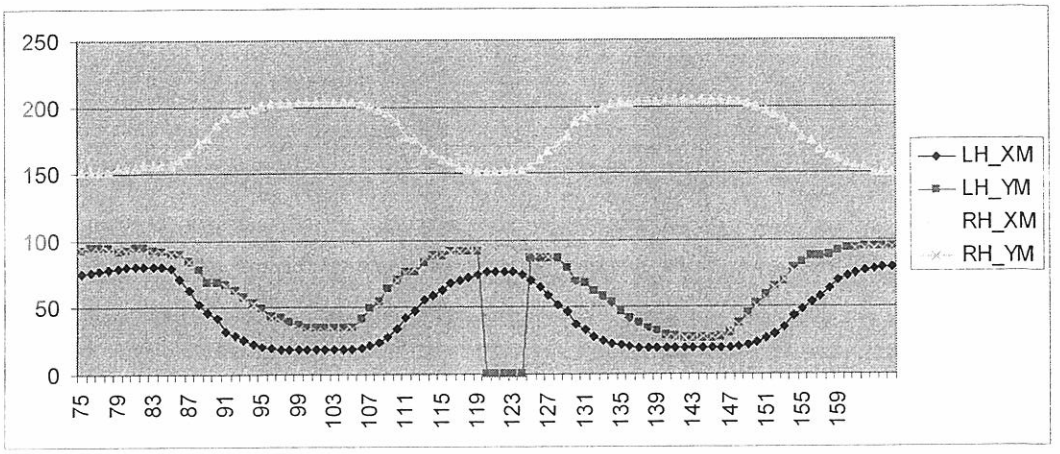
ซึ่งหลักการวิเคราะห์จะทำโดยนำข้อมูลตำแหน่งของส่วนต่างๆของร่างกายที่ได้มาจากแต่ละเฟรม มาทำการคำนวณผลต่างของตำแหน่งอวัยวะในเฟรมต่างๆ และนำผลที่ได้มาทำการคำนวณหามุมเพื่อใช้ดูทิศทางของการเคลื่อนไหวของอวัยวะนั้นๆ จะแสดงกราฟข้อมูลตำแหน่งพิกัด x, y ของแขนซ้ายและขวาที่ได้มาจากขั้นตอนการระบุตำแหน่งของร่างกาย ดังภาพที่ 3.16



ภาพที่ 3.16 แสดงกราฟข้อมูลตำแหน่งของแขน

จากรูปที่ 3.16 จะสังเกตุได้ว่าพิกัดที่ได้มานั้นเส้นกราฟจะไม่เรียบเนื่องจากจะมีข้อมูลที่ผิดพลาดติดมาด้วย จะเห็นได้ว่ากราฟของข้อมูล LH_Y, RX_Y นั้นจะมีการกระโดดของข้อมูลเป็นจำนวนมาก ทำให้กราฟที่ออกมาไม่เรียบ เราจึงต้องนำข้อมูลนี้มาทำการปรับค่าของข้อมูล (Smoothing data) ก่อนเป็นการกำจัดค่าข้อมูลที่ผิดพลาดที่ติดมาออกไปก่อน โดยหลักการที่ใช้ในการปรับข้อมูลจะใช้การทำการปรับค่าทุกๆ 7 เฟรม ซึ่งจะเอาค่าพิกัดในเฟรมต่อๆกันมา 7 ค่ามาทำการเรียงข้อมูลจากน้อยไปมากโดยใช้การเรียงชุดข้อมูลฟองสบู่ (Bubble Sort) แล้วทำการเลือกค่ามัธยฐาน (Median) ออกมา เมื่อได้ค่ามัธยฐานออกมาแล้วก็จะทำการเอาค่าข้อมูลตัวแรกออกไปเพื่อที่จะได้นำข้อมูลเฟรมต่อไปเข้ามาเพื่อทำการจัดเรียงและหาค่ามัธยฐานต่อไป เพื่อให้ข้อมูลที่ได้มีความต่อเนื่อง ดังภาพที่ 3.17

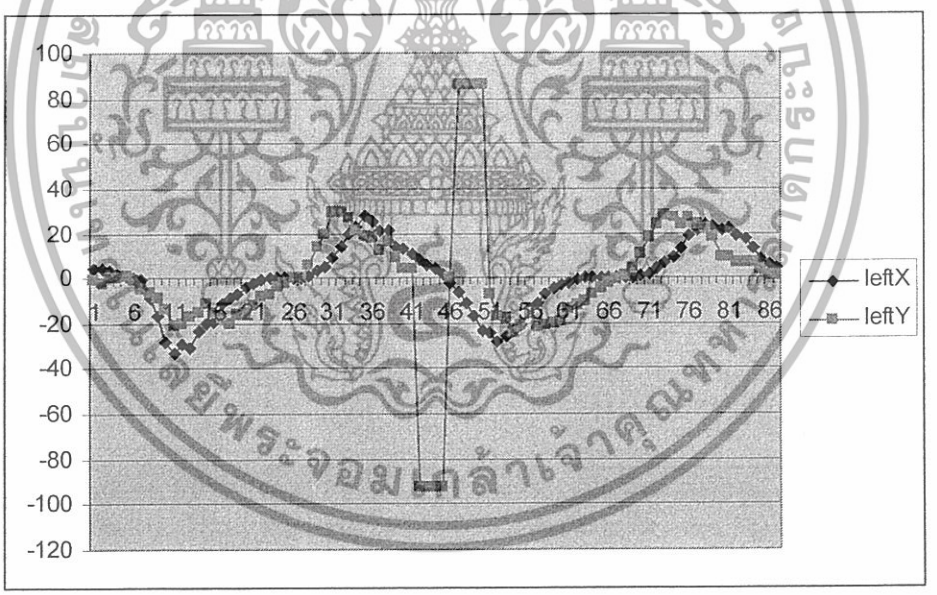
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.17 แสดงกราฟที่ผ่านการทำการปรับค่าข้อมูล

จากรูปที่ 3.17 จะเป็นการแสดง ข้อมูล LH_Y, RX_Y ที่ผ่านการปรับค่าของข้อมูลแล้ว จะเห็นว่าข้อมูล LH_YM, RX_YM ที่ได้มาจะมีการกระโดดของข้อมูลน้อยลง

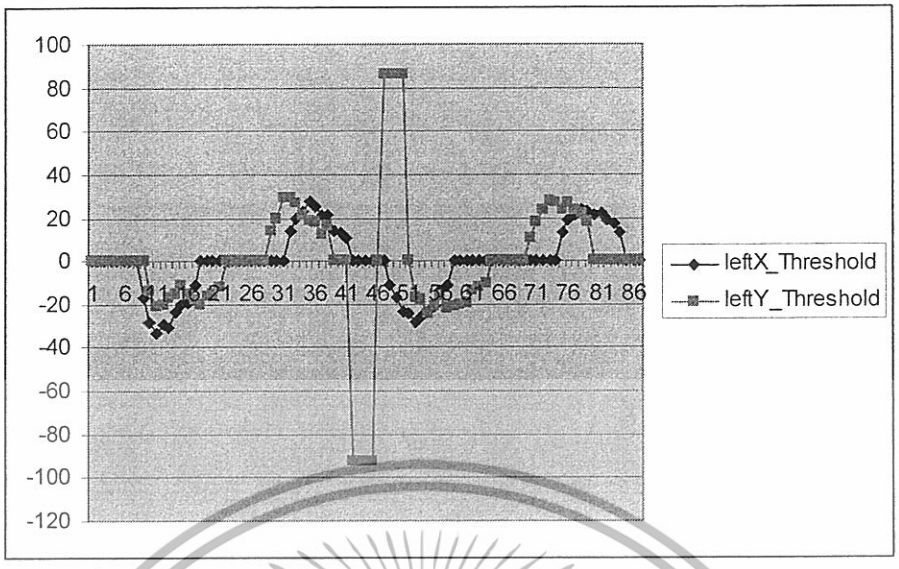
หลังจากที่ทำการปรับค่าข้อมูลแล้วก็จะนำเอาข้อมูลที่ได้มาดูความแตกต่างของแต่ละเฟรม มาหาผลต่างระหว่างเฟรมโดยจะนำทุกๆเฟรมที่ 1 และเฟรมที่ 5 มาหาผลต่าง ดังภาพที่ 3.18



ภาพที่ 3.18 แสดงกราฟที่ผ่านการหาค่าความแตกต่าง

หลังจากที่ได้หาความแตกต่างมา จะมีการใช้เทคนิคการทำหาค่าขีดแบ่ง (Thresholding) เทคนิคนี้คือ ค่าที่เราทำการตั้งไว้เพื่อใช้ในการแยกแยะที่จะระบุค่าความแตกต่าง ซึ่งถ้าผลต่างที่หามาได้ มีค่าน้อยกว่า 10 จะให้ค่าเป็นศูนย์ โดยจะถือว่าไม่มีความแตกต่าง ส่วนค่าที่ได้นอกเหนือจากนั้นจะให้ค่าเป็นค่าเดิมออกมา ถือเป็นค่าความแตกต่างที่ได้ออกมา ดังภาพที่ 3.19

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.19 แสดงกราฟที่ผ่านการหาค่าความแตกต่างโดยผ่านการหาค่าขีดแบ่ง

สุดท้ายเราจะเอาพิกัด x, y ของการเคลื่อนที่แต่ละส่วนของร่างกายมาหาค่ามุมโดยการนำ x, y มาทำการหา Arc tan เพื่อนำเอาค่าออกมา ซึ่งเราจะเอาค่ามุมที่ได้มาทำการเทียบค่ากับทิศทางการเคลื่อนไหวโดยเราจะแบ่งทิศทางในการเคลื่อนไหวเป็น 8 ทิศทาง ดังภาพที่ 3.20 หลังจากที่ได้ทิศทางในการเคลื่อนไหวแล้วก็จะนำทิศที่ได้มาทำการวิเคราะห์เปรียบเทียบดูกับท่าการเคลื่อนไหวของเราว่ามีการเคลื่อนไหวตรงกันหรือไม่

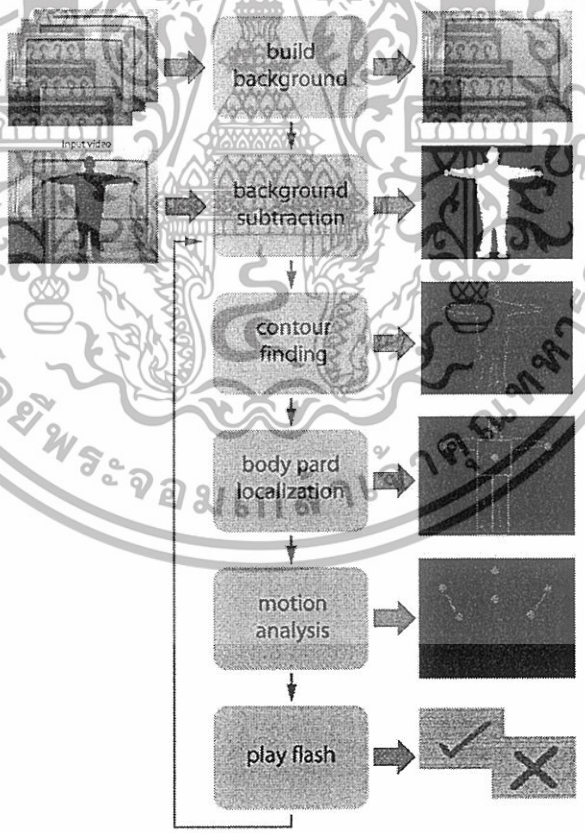
ภาพที่ 3.20 แสดงทิศทางการเคลื่อนไหวที่ใช้ในการเปรียบเทียบ

จากภาพที่ 3.20 จะแสดงทิศทางการเคลื่อนไหวที่เราใช้นำมาเทียบกับค่ามุมที่ได้มา เพื่อหาการเคลื่อนที่ที่เกิดขึ้น โดยที่ จะกำหนดดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ค่าเท่ากับ 1 เมื่อมุมอยู่ระหว่าง 0-21 และ 340-360 องศา หมายถึงมีการเคลื่อนที่ทางขวา
- ค่าเท่ากับ 2 เมื่อมุมอยู่ระหว่าง 22-68 องศา หมายถึงมีการเคลื่อนที่ไปทางขวาบน
- ค่าเท่ากับ 3 เมื่อมุมอยู่ระหว่าง 69-113 องศา หมายถึงมีการเคลื่อนที่ขึ้นบน
- ค่าเท่ากับ 4 เมื่อมุมอยู่ระหว่าง 114-159 องศา หมายถึงมีการเคลื่อนที่ไปทางบนซ้าย
- ค่าเท่ากับ 5 เมื่อมุมอยู่ระหว่าง 160-203 องศา หมายถึงมีการเคลื่อนที่ทางซ้าย
- ค่าเท่ากับ 6 เมื่อมุมอยู่ระหว่าง 204-249 องศา หมายถึงมีการเคลื่อนที่ไปทางล่างซ้าย
- ค่าเท่ากับ 7 เมื่อมุมอยู่ระหว่าง 250-293 องศา หมายถึงมีการเคลื่อนที่ลง
- ค่าเท่ากับ 8 เมื่อมุมอยู่ระหว่าง 294-339 องศา หมายถึงมีการเคลื่อนที่ไปทางล่างขวา

ส่วนการแสดงผลการ์ตูนแอนนิเมชั่น (Flash) จะแสดงผลของการวิเคราะห์ท่าทางการเคลื่อนไหวของผู้ใช้งานว่าถูกต้องตามตัวแอนนิเมชั่นที่แสดงอยู่หรือไม่ โดยที่การแสดงผลนั้นจะแสดงในรูปแบบแอนนิเมชั่นซึ่งถ้าระบบวิเคราะห์ว่าผู้เล่นทำท่าทางได้ถูกต้องระบบจะกล่าวชมและทำการเปลี่ยนท่าการเต้น แต่หากวิเคราะห์ได้ว่าผู้ใช้ทำท่าทางที่ไม่ตรงกับการ์ตูนแอนนิเมชั่นที่นำเต้นอยู่ระบบจะทำการแจ้งว่าผู้เล่นทำผิดท่าและจะให้ทำท่าทางนั้นซ้ำไปเรื่อยๆจนกว่าผู้เล่นจะทำท่าได้ถูกต้อง



ภาพที่ 3.21 แสดงผลที่ได้จากการทำงานในแต่ละขั้นตอน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

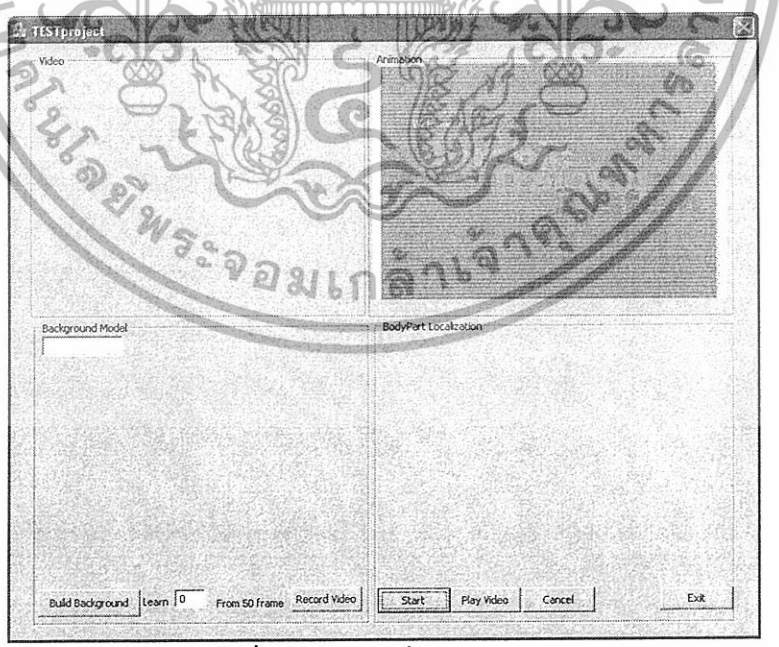
จากภาพที่3.21 จะแสดงถึงภาพรวมของขั้นตอนการทำงานด้านการประมวลผลภาพ พร้อมทั้งผลลัพธ์ที่ได้จากการทำงานในแต่ละขั้นตอน

3.4 ส่วนการออกแบบส่วนติดต่อกับผู้ใช้

การทำงานของโปรแกรมนี้ จะมีขั้นตอนในการใช้งานโปรแกรมที่ง่าย ทำให้ผู้ใช้สามารถเข้าใจในการใช้งานและใช้งานได้อย่างถูกต้อง โดยส่วนที่ใช้ติดต่อกับผู้ใช้ (User Interface) ของระบบช่วยฝึกการออกกำลังกายแบบแอโรบิคเบื้องต้นโดยคอมพิวเตอร์วิชั่นจะเป็นการแสดงผลของหน้าจอที่ใช้กล้องวีดีโอ, หน้าจอแสดงผลการ์ตูนแอนนิเมชั่น, หน้าจอแสดงแบบจำลองฉากหลัง, หน้าจอแสดงการระบุส่วนของร่างกาย ซึ่งในการออกแบบส่วนติดต่อกับผู้ใช้มี ส่วนหลักๆ ดังนี้

1. หน้าจอหลักโปรแกรม
2. หน้าจอแสดงภาพที่รับมาจากกล้อง
3. หน้าจอแสดงผลการ์ตูนแอนนิเมชั่น
4. หน้าจอฉากหลัง
5. หน้าจอการระบุส่วนของร่างกาย

หน้าจอหลักของโปรแกรมเป็นหน้าต่างของโปรแกรมหลักโดยจะพบเมื่อทำการเรียกใช้โปรแกรมนี้ขึ้นมาโดยหน้าจอนี้จะเป็นหน้าจอที่ใช้ควบคุมการทำงานส่วนต่างๆ ดังภาพที่3.22

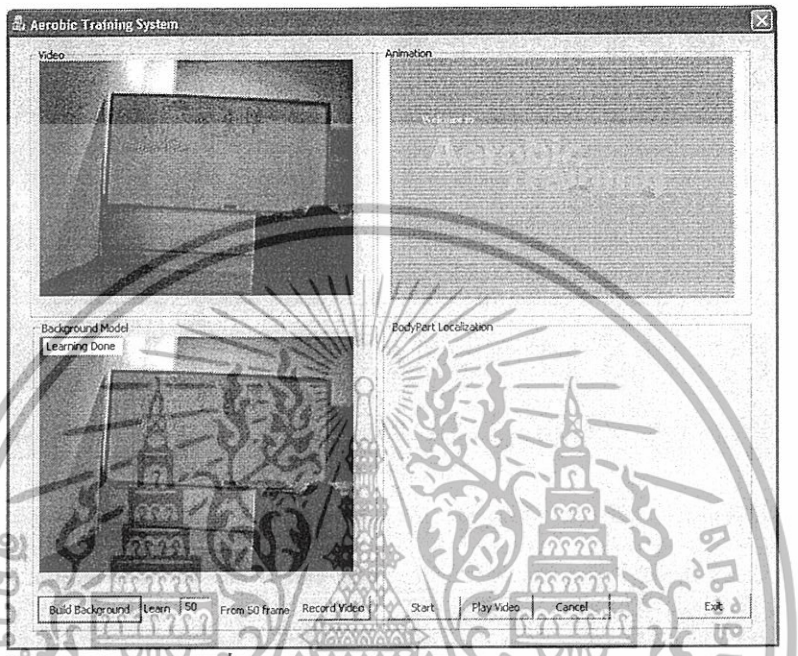


ภาพที่3.22 หน้าจอเริ่มต้นโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากภาพที่3.22 หน้าจอของโปแกรมในภาพประกอบด้วย 4 ส่วนหลักคือ มุมขวาบนจะแสดงภาพการ์ตูนกล่าวคำต้อนรับ มุมซ้ายบนเป็นภาพที่ได้จากวีดีทัศน์ หรือเป็นข้อมูลเข้า มุมซ้ายล่างจะแสดงภาพแบบจำลองฉากหลังที่สร้างขึ้น มุมขวาล่างจะแสดงการระบุส่วนของร่างกาย

สำหรับหน้าจอของการแบบจำลองฉากหลังนั้น เป็นส่วนที่มีความสำคัญที่จะทำให้ระบบทำงานได้อย่างถูกต้อง โดยจะทำการสร้างฉากหลังขึ้น ณ เวลาที่ไม่มีผู้ใช้อยู่ในจอภาพ

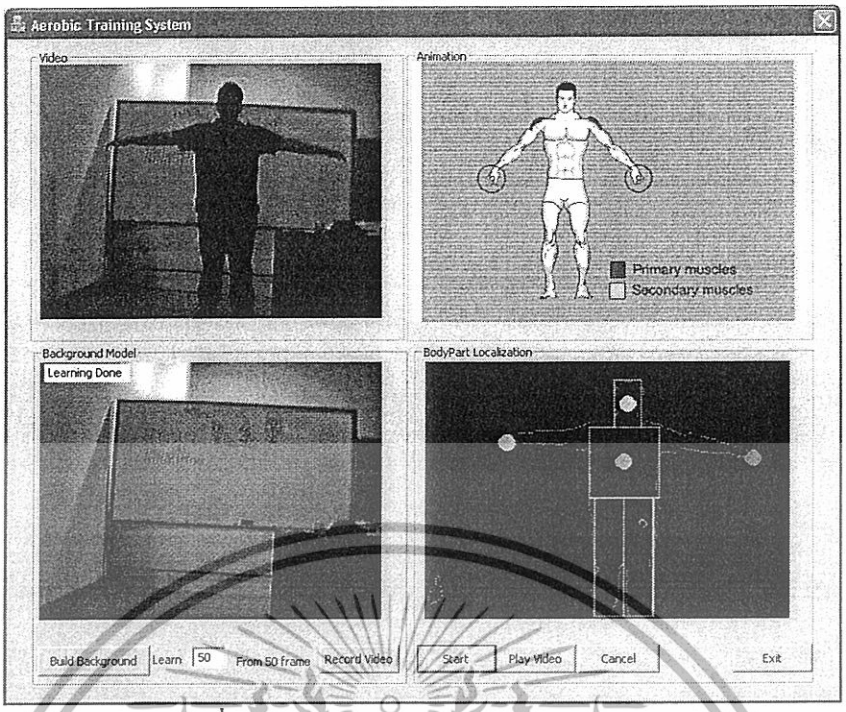


ภาพที่3.23 หน้าจอแสดงการสร้างฉากหลัง

จากภาพที่ 3.23 เมื่อกดปุ่ม Build Background ระบบก็จะเริ่มสร้าง Background Model โดยแสดงจำนวนเฟรมที่ทำการ Learn และแสดงข้อความ "Learning..." เมื่อสร้างเสร็จจะแสดงผล ภาพฉากหลัง (Background Model) ภาพซ้ายล่าง โดยที่มุมซ้ายบนเป็นภาพที่ได้จากวีดีทัศน์

เมื่อทำการสร้างฉากหลังเสร็จแล้วในจอภาพ จะแสดงผลขึ้นว่า Learning done แล้วจะทำการเริ่มเล่นได้ต่อเมื่อกดปุ่ม Start เพื่อเริ่มการตรวจจับและวิเคราะห์การเคลื่อนไหวเมื่อมีผู้ใช้เข้าสู่ระบบ ดังภาพที่ 2.24

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่3.24 หน้าจอแสดงการทำงานของโปรแกรม

จากภาพที่ 3.24 แสดงตัวอย่างการระบุตำแหน่งของศีรษะ และมือทั้งสองข้างจากท่ากางแขนทั้ง 2 ข้าง ซึ่งจะสังเกตว่าระบบสามารถหาตำแหน่งของส่วนทั้งสองได้หลังจากนั้น ระบบก็จะเริ่มทำการวิเคราะห์ท่าทางว่าตรงกับท่าทางของการ์ตูนหรือไม่และทำการโต้ตอบกับผู้ใช้โดยใช้เสียงรวมกับเครื่องหมายถูก และเครื่องหมายผิดในการบอกว่าทำท่าทางผิดหรือไม่ตรงกันที่การ์ตูนแอนนิเมชั่นแสดงอยู่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
วิธีทดลองและผลการทดลอง

ในบทนี้จะเป็นการกล่าวถึงรายละเอียดของผลการทดลองของระบบช่วยฝึกการออกกำลังกายแบบแอโรบิคโดยการประยุกต์ใช้เทคนิคการตรวจจับการเคลื่อนไหวของมนุษย์จากวีดิทัศน์โดยการทดลองนั้น จะทดลองในส่วนที่สำคัญได้แก่ การตัดแยกผู้ใช้ออกจากฉากหลัง การระบุตำแหน่งส่วนของร่างกาย รวมทั้งความถูกต้องในการวิเคราะห์ท่าทาง

4.1 อุปกรณ์ที่ใช้ในการทดลอง

- กล้องวีดิทัศน์ ใช้ในการส่งภาพวีดิโอที่จับภาพได้เข้าสู่เครื่องคอมพิวเตอร์เพื่อใช้ในการประมวลผลภาพ โดยการทดสอบนี้จะใช้เว็บแคม (Web Cam) รุ่น D-Link DSB-C310 ซึ่งเป็นเว็บแคมระดับปานกลาง จากการพัฒนาระบบซึ่งได้ทำการทดลองใช้งานระบบ สามารถสรุปผลการทำงานของ ระบบช่วยฝึกการออกกำลังกายแบบแอโรบิคโดยการประยุกต์ใช้เทคนิคการตรวจจับการเคลื่อนไหวของมนุษย์จากวีดิทัศน์ได้ดี

- เครื่องคอมพิวเตอร์ ซีพียู อินเทลเพนเทียม 4 1.6 (CPU Pentium(R)4 1.6 GHz) หน่วยความจำหลัก 512 MB

4.2 ขั้นตอนการทดลอง

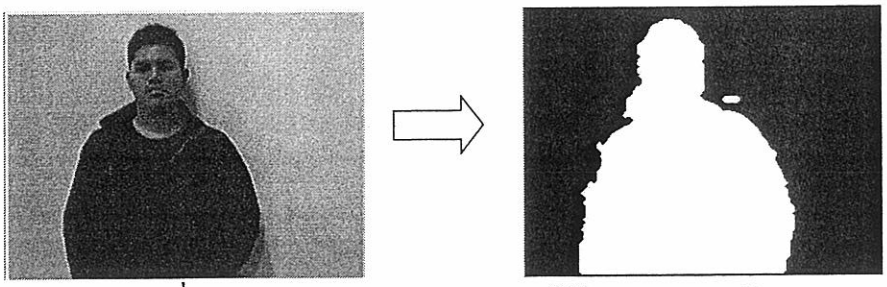
4.2.1 การทดลองการตัดแยกผู้ใช้ออกจากฉากหลัง

ในการทดสอบการตัดแยกผู้ใช้ออกจากฉากหลัง การที่จะทำให้ฉากหลังหายไปจากภาพโดยสมบูรณ์ตลอดเวลานั้นเป็นเรื่องที่ยากอย่างเช่น เมื่อแสงมีการเปลี่ยนไป หรือมีสิ่งของบางอย่างหายไปจากฉากหลัง ซึ่งถ้ากระบวนการลบฉากหลังไม่ดีเพียงพอก็จะทำให้กระบวนการขั้นตอนต่อไปทำได้ยาก ในที่นี้จะต้องทำการสร้างฉากหลังขึ้นมาก่อนโดยในการสร้างฉากหลังขึ้นมานั้นทำการทดลองกับสภาพแวดล้อม ณ เวลาที่ไม่มีผู้ใช้อยู่ในขอบเขตภาพ

สำหรับการทดลองนี้ จะทดสอบผลของค่าแสงที่อยู่ในสภาพแวดล้อมเดียวกัน ซึ่งจะทำการทดลองสภาวะแสง 2 ลักษณะคือ สภาพที่มีแสงสว่างเหมาะสม และแสงที่มืดลงซึ่งการทดลองของระบบการตัดแยกผู้ใช้ออกจากฉากหลังในสภาวะค่าแสงที่ต่างกัน ผลที่ได้จากสภาวะต่างๆ มีดังต่อไปนี้

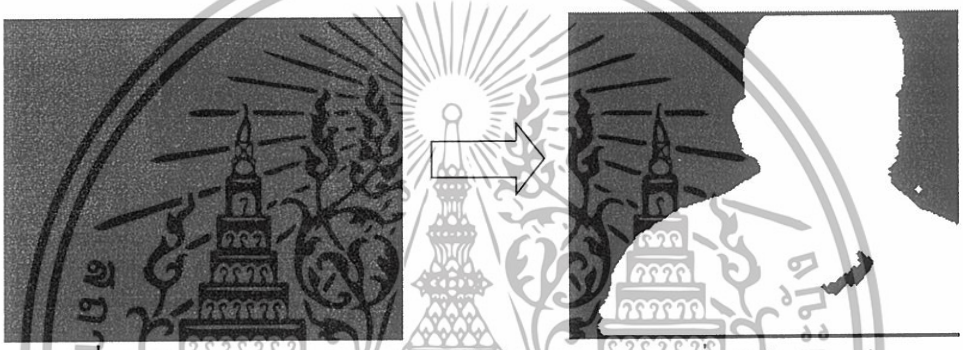
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ผลการทดลองการตรวจสีผิวในสภาวะแสงที่สว่าง



ภาพที่ 4.1 ผลการทดลองการการตัดแยกผู้ใช้ออกจากฉากหลัง

- ผลการทดลองเมื่อลดค่าความสว่างลง



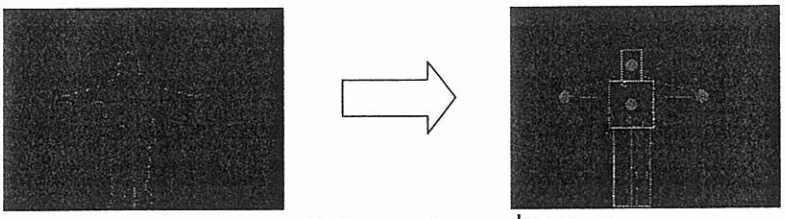
ภาพที่ 4.2 ผลการทดลองการการตัดแยกผู้ใช้ออกจากฉากหลังเมื่อลดค่าความสว่างลง

ผลที่ได้หลังจากลดค่าของแสงลงส่วนที่มีพื้นผิวสีเข้มได้หายไปบางส่วน เนื่องจากมีความใกล้เคียงกับสีของฉากหลัง แต่แบบจำลองการการตัดแยกผู้ใช้ออกจากฉากหลังนั้นก็ยังคงสามารถทำงานได้

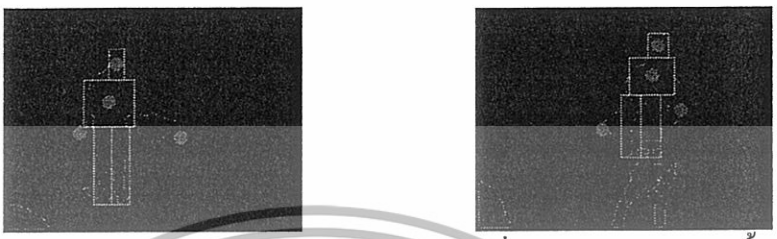
4.2.2 การทดลองการระบุตำแหน่งส่วนของร่างกาย

การระบุตำแหน่งส่วนของร่างกายนั้นจะทำการหาตำแหน่งและแสดงจุดที่หาได้โดยจุดที่ทำการตรวจจับการเคลื่อนไหวนั่นคือศีรษะ ลำตัว และมือทั้งสองข้าง สำหรับการทดลองนี้ จะทดสอบผลกระทบของสัญญาณรบกวนที่เกิดขึ้น ซึ่งจะทำการทดลองใน 2 ภาวะ คือ ภาวะที่มีสัญญาณรบกวนและภาวะที่ไม่มีสัญญาณรบกวน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 4.3 ผลการทดลองระบุตำแหน่งส่วนของร่างกายที่ไม่เกิดสัญญาณรบกวน



ภาพที่ 4.4 ผลการทดลองระบุตำแหน่งส่วนของร่างกายที่มีสัญญาณรบกวนเกิดขึ้น

จากรูปที่ 4.4 จะเห็นว่าผลการทดลองการระบุตำแหน่งส่วนของร่างกาย ที่มีสัญญาณรบกวนเกิดขึ้นนั้น มีการระบุตำแหน่งของแขนที่ผิดพลาดซึ่งไม่สามารถนำภาพที่ได้ในลักษณะเช่นนี้ไปใช้ ในการระบุตำแหน่ง แขนซ้าย และแขนขวาได้

ผลที่ได้จากการทดลองที่ไม่มีสัญญาณรบกวนสามารถทำการระบุตำแหน่งของร่างกายได้อย่างถูกต้อง แต่ถ้ารูปที่ได้มาจากขั้นตอนที่มีสัญญาณรบกวนซึ่งในที่นี้คือ มีเงาเกิดขึ้น ก็อาจจะมีความผิดพลาดในการระบุตำแหน่งส่วนต่างๆของร่างกาย

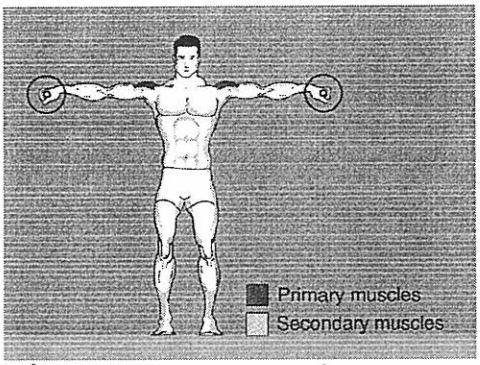
4.2.3 การวิเคราะห์ท่าทาง

การวิเคราะห์ท่าทาง เป็นขั้นตอนถัดจากการระบุตำแหน่งส่วนของร่างกาย ในขั้นตอนนี้จะทำการวิเคราะห์ความถูกต้องของท่าทางโดยนำตำแหน่งของร่างกายที่ได้ มาเปรียบเทียบความแตกต่างและทิศทางในการเคลื่อนที่ เพื่อมาเปรียบเทียบว่าทำท่าได้ถูกต้องตามแอนนิเมชั่นต้นแบบหรือไม่

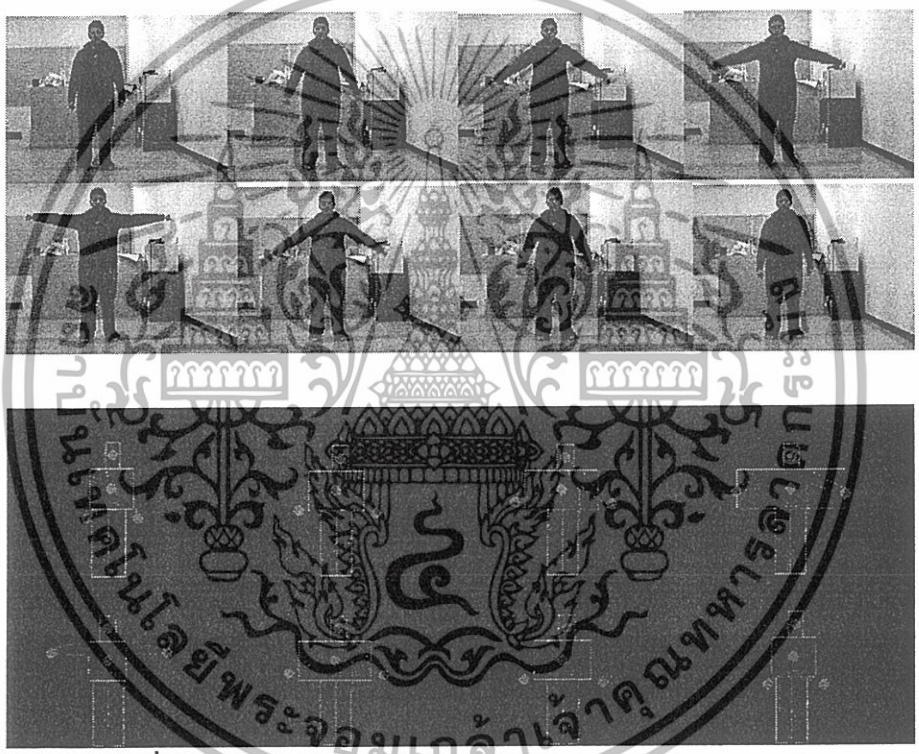
สำหรับการทดลองนี้ จะทำการทดสอบทำท่าทางให้ตรงกับตัวแอนนิเมชั่นแล้วดูว่าระบบสามารถ วิเคราะห์ข้อมูลออกมาได้ถูกต้องหรือไม่ โดยในการทดลองจะทดสอบกับท่าการออกกำลังกาย 4 ท่า ได้แก่ ท่ายกแขนทั้ง 2 ข้าง ท่ายกแขนซ้าย ท่ายกแขนขวา และท่าย่อเข่า ซึ่งผลการทดสอบมีดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ผลการทดลองกับท่ายกแขน2ข้าง



ภาพที่ 4.5 แสดงการ์ตูนแอนนิเมชั่นท่ายกแขน2ข้าง

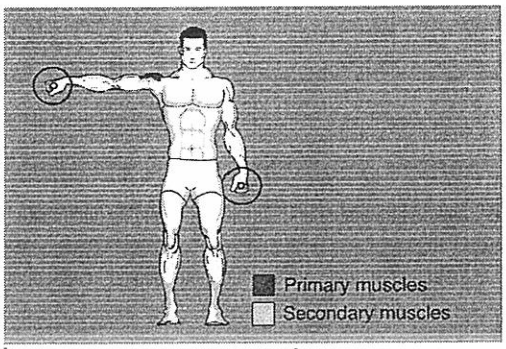


ภาพที่ 4.6 แสดงการวิเคราะห์ท่าทางการออกกำลังกายท่ายกแขน2ข้าง

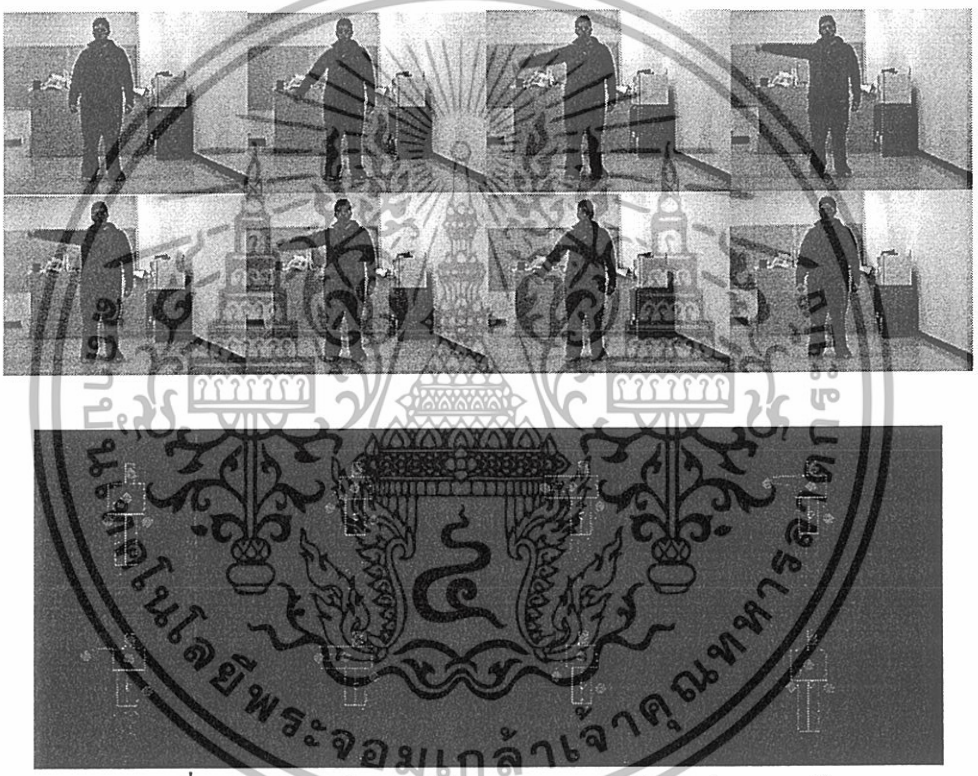
จากภาพที่ 4.6 จะเห็นว่าเมื่อเราทำท่าตามที่ตัวการ์ตูนแอนนิเมชั่นแสดง คือ ยกแขนทั้ง 2 ข้างขึ้นลงพร้อมกัน ระบบสามารถที่จะทำการวิเคราะห์ท่าทางออกมาได้อย่างถูกต้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ผลการทดลองกับท่ายกแขนซ้าย



ภาพที่ 4.7 แสดงการ์ตูนแอนนิเมชั่นท่ายกแขนซ้าย

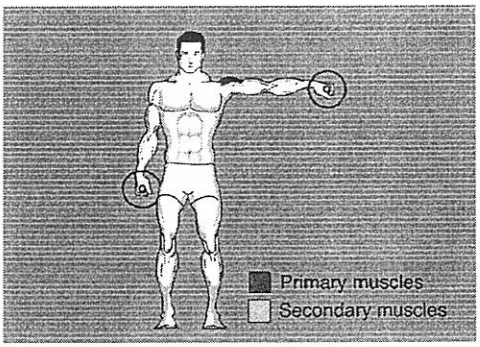


ภาพที่ 4.8 แสดงการวิเคราะห์ท่าทางการออกกำลังกายท่ายกแขนซ้าย

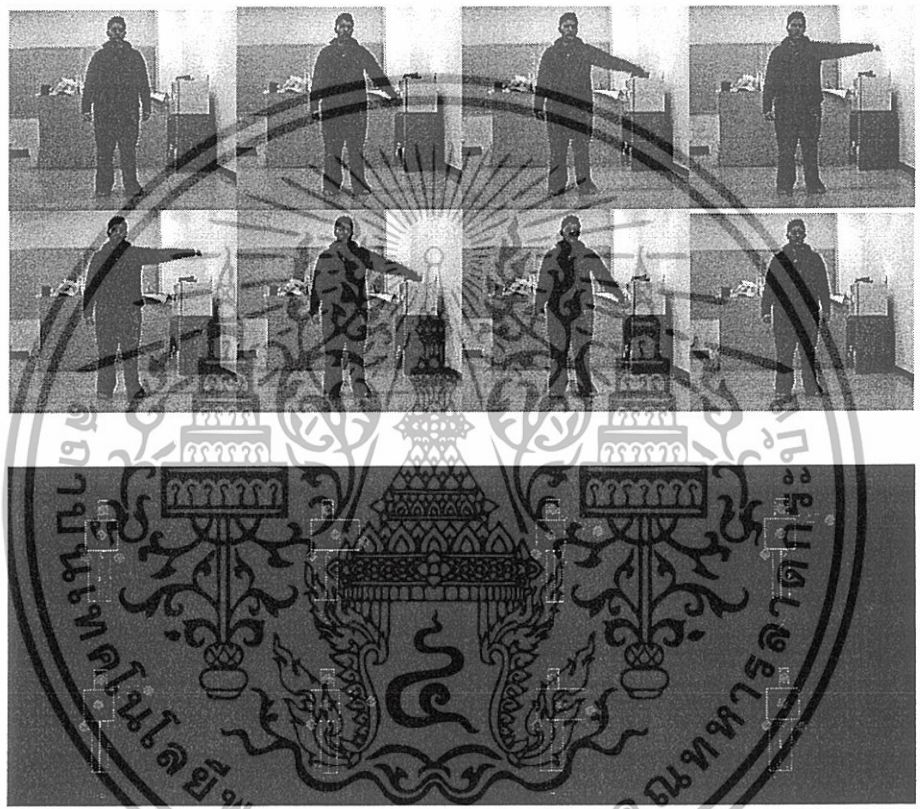
จากภาพที่ 4.8 จะเห็นว่าเมื่อเราทำท่าตามที่ตัวการ์ตูนแอนนิเมชั่นแสดง คือ ยกแขนซ้ายขึ้นลง ระบบสามารถที่จะทำการวิเคราะห์ท่าทางออกมาได้อย่างถูกต้อง

- ผลการทดลองกับท่ายกแขนขวา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 4.9 แสดงการ์ตูนแอนนิเมชั่นท่ายกแขนขวา

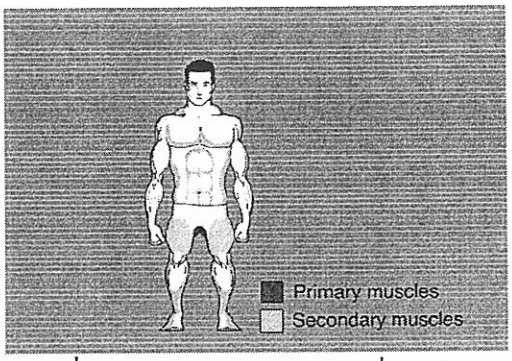


ภาพที่ 4.10 แสดงการวิเคราะห์ท่าทางการออกกำลังกายท่ายกแขนขวา

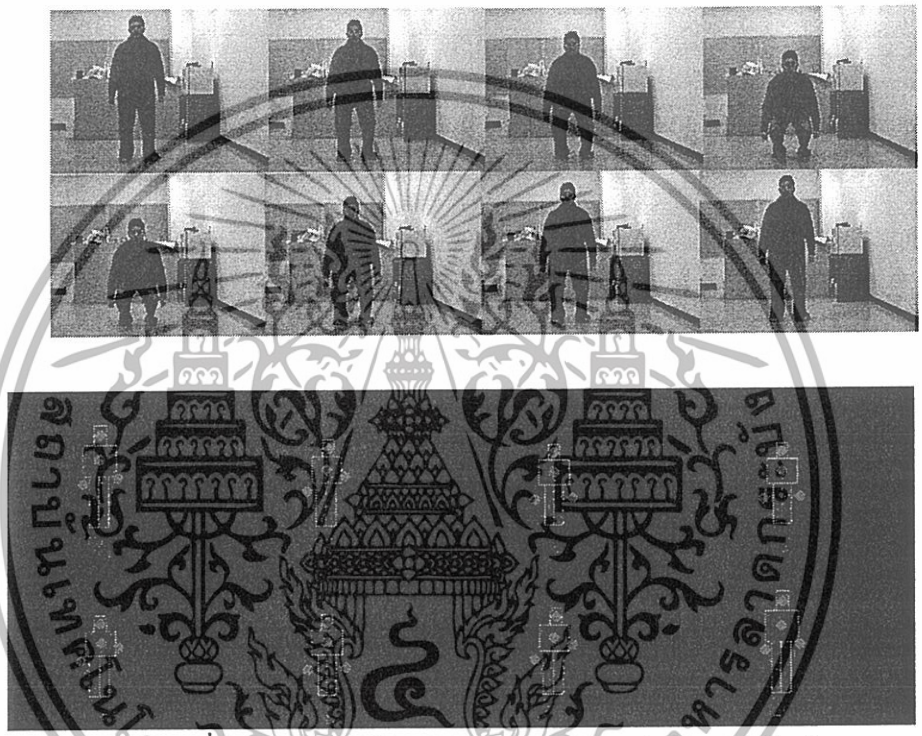
จากภาพที่ 4.10 จะเห็นว่า เมื่อเราทำท่าตามที่ตัวการ์ตูนแอนนิเมชั่นแสดง คือ ยกแขนขวาขึ้นลง ระบบสามารถที่จะทำการวิเคราะห์ท่าทางออกมาได้อย่างถูกต้อง

- ผลการทดลองกับท่าย่อเข่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 4.11 แสดงการ์ตูนแอนนิเมชั่นท่าย่อเข่า



ภาพที่ 4.12 แสดงการวิเคราะห์ท่าทางการออกกำลังกายท่าย่อเข่า

จากภาพที่ 4.12 จะเห็นว่า เมื่อเราทำท่าตามที่ตัวการ์ตูนแอนนิเมชั่นแสดง คือ ย่อเข่าขึ้นลง ระบบสามารถที่จะทำการวิเคราะห์ท่าทางออกมาได้อย่างถูกต้อง

4.3 สรุปผลการทดลอง

ระบบสามารถทำงานได้อย่างถูกต้องเมื่อฉากหลังมีการเปลี่ยนแปลงของแสงและเงาน้อย และในขั้นตอนการวิเคราะห์ท่าทางในการเคลื่อนไหว จะขึ้นอยู่กับคุณภาพของภาพที่ได้มาจากขั้นตอนการลบฉากหลังออก ซึ่งจะส่งผลต่อการหาตำแหน่งของร่างกาย นอกจากนี้หากในบาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ท่าทางผู้ใช้เคลื่อนไหวน้อยไป เช่น มีการย่อเข่าน้อยไปกว่าระยะการย่อที่ระบบกำหนดเอาไว้ ระบบจะวิเคราะห์ว่าทำท่าทางไม่ถูกต้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลโครงงานและข้อเสนอแนะ

5.1 สรุปผลโครงงาน

โครงงานพัฒนาระบบช่วยฝึกการออกกำลังกายโดยการประยุกต์ใช้เทคนิคการตรวจจับการเคลื่อนไหวของมนุษย์จากวีดิทัศน์เป็นการศึกษาทฤษฎีที่เกี่ยวข้องกับการพัฒนาระบบในส่วนต่างๆ และเป็นการสร้างระบบให้สามารถทำงานได้ตามที่กำหนดไว้ในขอบเขตของการพัฒนาระบบ โดยระบบประกอบไปด้วยส่วนหลักๆ คือ การนำข้อมูลจากวีดิทัศน์เข้าสู่ระบบ การตัดส่วนที่สนใจออกจากฉากหลัง การตรวจจับการเคลื่อนไหวของมนุษย์ การระบุตำแหน่งของร่างกาย การวิคราะห์ท่าทาง และการโต้ตอบกับผู้ใช้

จากที่ผ่านมาการทดสอบตัวระบบ พบว่าระบบมีความแม่นยำในการตรวจจับ และการวิเคราะห์ท่าทาง คิดเป็น 87% จากการทดลองกับท่าเต้นทั้งหมด 5 ท่า นอกจากนี้ยังพบว่าในส่วนการตัดส่วนที่สนใจออกจากฉากหลัง ระบบไม่สามารถที่จะทำการตรวจจับได้ถูกต้อง ถ้าตัวผู้ใช้ใส่ชุดที่มีสีกลมกลืนกับฉากหลังที่ได้สร้างไว้ โดยระบบจะทำการระบุตำแหน่งต่างๆของร่างกายไม่ตรงกับความเป็นจริง ในการพัฒนาระบบยังพบปัญหาและข้อจำกัดในการพัฒนาระบบหลายอย่าง ดังจะได้กล่าวในหัวข้อถัดไป

5.2 ปัญหาและข้อจำกัดในการทำงานของระบบ

ในการพัฒนาระบบช่วยฝึกการออกกำลังกายโดยการประยุกต์ใช้เทคนิคการตรวจจับการเคลื่อนไหวของมนุษย์จากวีดิทัศน์ มีปัญหาและข้อจำกัดในการพัฒนาหลายอย่าง เนื่องจากระบบมีปัจจัยและข้อจำกัดที่ส่งผลต่อการทำงานของระบบ ดังต่อไปนี้

1. กล้องวีดิทัศน์ ที่ใช้เป็นกล้อง web cam ซึ่งมักเป็นกล้องที่มีการปรับค่า (gain) อัตโนมัติ ทำให้ความสว่างของภาพปรับเปลี่ยนไปส่งผลต่อการทำงานของการลบฉากหลังที่ใช้ซึ่งเป็นอัลกอริทึมแบบไม่มีการปรับตัว จึงทำให้ผลการลบฉากหลังประกอบด้วยสัญญาณรบกวนจำนวนหนึ่ง

2. ในส่วนของการนำเข้าข้อมูลจากวีดีทัศน์ยังพบปัญหาว่าไม่สามารถนำข้อมูลจากวีดิทัศน์เข้าในระบบได้ทุกยี่ห้อทุกรุ่น จากการสังเกตพบว่าระบบต้องการวีดิทัศน์ที่มี Driver Support การแสดงผลภาพแบบ RGB

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. ในส่วนการสร้างตัวแบบของฉากหลังขณะที่ทำการสร้างตัวแบบ ฉากหลังต้องไม่มีการเคลื่อนไหว

4. ในส่วนของการตัดส่วนที่สนใจออกจากฉากหลังทำงานได้ดีเมื่อแสงและเงามีการเปลี่ยนแปลงน้อย ไม่เหมาะกับพื้นผิวที่มีการสะท้อนแสงเช่นกระจก หรือวัสดุที่มีความมันวาว และในส่วนของการประมวลผลโดยการหา เส้นรอบรูป ทำให้ระบบต้องประมวลผลมากทำให้ไม่สามารถประมวลผลได้ทุกๆเฟรมที่เข้ามาในระบบได้ซึ่งมีผลกระทบในข้อต่อไป

5. การตรวจจับการเคลื่อนไหวของมนุษย์ที่ต้องนำข้อมูลจากการตัดส่วนที่สนใจออกจากฉากหลังที่มีการประมวลผล หากผู้ใช้เคลื่อนที่ด้วยความเร็วระบบจะไม่สามารถทำการเคลื่อนไหวอย่างต่อเนื่องได้ เนื่องจากบางเฟรมหายไป ซึ่งมีสาเหตุมาจาก 2 ส่วนคือ วีดิทัศน์ไม่สามารถจับภาพได้ และการประมวลผลการลบฉากหลังทำได้ไม่ครบทุกเฟรม

6. การตรวจจับการเคลื่อนไหวของมนุษย์ในโครงงานนี้จะตรวจจับได้เฉพาะการเคลื่อนไหวที่ไม่ซับซ้อนมากโดยจะไม่สามารถตรวจจับท่าทางที่ใช้มือชูเหนือระดับศีรษะ เพราะระบบใช้การคาดเดาว่าส่วนที่อยู่สูงที่สุดคือส่วนของศีรษะและใช้การคาดเดาตำแหน่งจากสัดส่วนที่ควรจะเป็นในส่วนอื่นๆต่อไป

7. การตรวจจับการเคลื่อนไหวของมนุษย์ไม่สามารถตรวจจับความเคลื่อนไหวของมนุษย์ได้มากกว่า 1 คน

8. ระบบต้องการทำงานแบบตอบสนองแบบเวลาจริง(Real-time) ทำให้มีการประมวลสูง ใช้ทรัพยากรต่างๆ สูง จึงจำเป็นต้องใช้เครื่องที่มีประสิทธิภาพ และมีหน่วยความจำสูงในการประมวลผลระบบ

5.3 ข้อเสนอแนะในการพัฒนาระบบต่อไป

1. พัฒนาวิธีการในการประมวลผลให้สามารถตอบสนองได้เร็วหรือสามารถวิเคราะห์ข้อมูลที่ได้ในทุกเฟรมเพื่อความถูกต้อง

2. ปรับปรุงวิธีในการวิเคราะห์การเคลื่อนไหวที่สามารถวิเคราะห์ท่าทางที่ซับซ้อนได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

นิรุธ อำนวยศิลป์. 2548. **เขียนโปรแกรมบนวินโดวส์ด้วย Visual C++ และ MFC**.พิมพ์ครั้งที่1. กรุงเทพ: ด่านสุทธาการพิมพ์

Haritaoglu, I. Harwood, D. and Davis, L.S. 1998 "Ghost: A Human Body Part Labeling System Using Silhouettes" **14th International Conference on Patten Recognition.**

Haritaoglu, I. Harwood, D. and Davis, L.S. 1998 "W4: Who? When? Where? What? A Real time System for Detection and tracking People" **International Conference on Face and Gesture Recognition**.

Horprasert, T. Harwood, D. and Davis, L.S. 1999 "A Statistical Approach for Real-time Robust Background Subtraction and Shadow Detection", **Proc. IEEE ICCV'99 FRAME-RATE Workshop, Kerkyra, Greece.**

R.C. Gonzales and R.E. Woods.1992. **Digital Image Processing.** Addision Wesley.

Shapiro,L.G. and Stockman,G.C.2001.**Computer Vision.** New Jersey :Prentice Hall

Silhouette-based Method for Object Classification and Human Action Recognition in Video Yigithan Dedeoglu1, B. Ugur Töreyin2, Ugur Güdükbay1, A. Enis Çetin2 , **Bilkent University Department of Computer Engineering1**

Sourceforge.2006.**OpenCV Library**.[Online] Available: http://opencvlibrary.sourceforge.net/

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ-นามสกุล	นายชานน ศรีวิเศษสม
วัน-เดือน-ปีเกิด	26 กรกฎาคม 2528
ที่อยู่	218 ซ.พัฒนาการ 53 แขวงสวนหลวง เขตสวนหลวง กทม.10250
ประวัติการศึกษา	โรงเรียนเตรียมอุดมศึกษาพัฒนาการ

ชื่อ-นามสกุล	ธนศร คันธวัฒน์
วัน-เดือน-ปีเกิด	28 กรกฎาคม 2527
ที่อยู่	49/64 ซ.วิภาวดี64 ถนนวิภาวดี-รังสิต หมู่ที่3 แขวงตลาดบางเขน เขตหลักสี่ กทม. 10210
ประวัติการศึกษา	โรงเรียนหอวัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้