



การวิเคราะห์โมเดลของมนุษย์และติดตามด้วยการประมวลผลภาพ  
สำหรับการค้นหาเหตุการณ์ผิดปกติ

**Model-based Human Analysis and Tracking for Abnormal Event Detection**

**ณัฐพล หนูฤทธิ**

**Nattapon Noorit**

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา  
วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์  
มหาวิทยาลัยสงขลานครินทร์

**A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of**

**Master of Engineering in Computer Engineering**

**Prince of Songkla University**

**2553**

**ลิขสิทธิ์ของมหาวิทยาลัยสงขลานครินทร์**



ชื่อวิทยานิพนธ์	การวิเคราะห์โมเดลของมนุษย์และติดตามด้วยการประมวลผลภาพสำหรับการค้นหาเหตุการณ์ผิดปกติ
ผู้เขียน	นายณัฐพล หนูฤทธิ
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
ปีการศึกษา	2552

### บทคัดย่อ

วิทยานิพนธ์ฉบับนี้เป็นงานวิจัยเพื่อประยุกต์ใช้เทคนิคด้านการประมวลผลภาพวิเคราะห์โมเดลมนุษย์สำหรับการค้นหาเหตุการณ์ผิดปกติ ซึ่งมุ่งเน้นไปที่กิจกรรมของมนุษย์ โดยเทคนิคการค้นหาเหตุการณ์ผิดปกติประกอบด้วย กระบวนการสร้างโครงสร้างมนุษย์ขึ้นใหม่ให้อยู่ในรูปแบบจำลองโครงสร้างมนุษย์อย่างง่ายที่มีเพียง 3 องค์ประกอบสำคัญของร่างกาย คือ ศีรษะ ลำตัวและขา จากนั้นจึงติดตามการเคลื่อนไหวของแต่ละองค์ประกอบอย่างเป็นอิสระต่อกัน ผลการติดตามจะให้ข้อมูลทิศทางการเคลื่อนที่ ความเร็วการเคลื่อนที่ ความเร่งการเคลื่อนที่และความเร่งเชิงมุมของแต่ละองค์ประกอบ โดยข้อมูลเหล่านี้และข้อมูลความสัมพันธ์เชิงมุมของแต่ละองค์ประกอบจะถูกนำไปใช้ในกระบวนการรู้จำท่าทางพื้นฐานซึ่งแบ่งเป็น 2 กลุ่ม ได้แก่ (1) ท่าทางสถิต ประกอบด้วย การยืน การนั่ง การก้ม และ (2) ท่าทางพลวัต ประกอบด้วย การนอนราบ การเดิน โดยผลการรู้จำท่าทางพื้นฐานให้ความถูกต้องโดยเฉลี่ย 93% และจากการนำท่าทางพื้นฐานเหล่านี้มาผสมผสานกันจะสามารถใช้อธิบายถึงกิจกรรมมนุษย์ที่มีความซับซ้อนมากขึ้นได้ โดยภายในงานวิจัยนี้ได้สนใจกรณีศึกษากิจกรรม 2 กรณี ได้แก่ (1) กระบวนการตรวจหาการเคลื่อนย้ายวัตถุ ซึ่งมุ่งเน้นขอบเขตการตรวจหาโดยการใช้กิจกรรมเฝ้าระวังเป็นตัวกำหนดเงื่อนไขการตรวจหาการเคลื่อนย้ายวัตถุ จากการทดลองจะให้ผลการตรวจหาถูกต้องแม่นยำเมื่อผลจากระบบการรู้จำท่าทางและกระบวนการลบพื้นหลังให้ผลที่ถูกต้อง และ (2) การควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่ ภายในงานวิจัยนี้ได้กำหนดเงื่อนไขควบคุมกิจกรรม พื้นที่ ทิศทางและความเร็วของการเคลื่อนที่ของวัตถุ จากการทดลองจะให้ผลการตรวจพบเหตุการณ์ควบคุมที่ถูกต้องเมื่อผลการติดตามวัตถุให้ผลที่ถูกต้อง

**Thesis Title**            Model-based Human Analysis and Tracking for Abnormal Event Detection  
**Author**                    Mr.Nattapon Noorit  
**Major Program**        Computer Engineering  
**Academic Year**         2009

## **ABSTRACT**

This thesis applies image processing techniques for abnormal event detection of human activities based on human model analysis. The detection algorithm is established into the following steps. Firstly, the human region is segmented from image sequences using motion detection. A simple human model consisting of three important parts such as head, body and leg is reconstructed by applying the color probability density of regions as classification criteria. Each region of human model is then tracked from frame to another in the sequence that provides the important information of human movement including velocity, acceleration and angular acceleration, which are reformulated into the features description of human characteristic. Secondly, by using the above features, the actions of human with its relations as finite state machine are defined. There are two groups of action model: static action (standing, sitting and bending) and dynamic action (walking and laying). The experimentation result of human action recognition is around 93%. Final step, the sequence of specific actions is defined as activity or behavior of human. In this research, we focus on two study cases: object removal and object direction control. For the object removal application, during the process the bending action is considered as triggered action that the monitoring of object in that area is activated. The accuracy of algorithm depends relatively on accuracy of action recognition. In the case of direction control, the features of human model are analyzed and compared with the considering area. If the specific direction of controlled area is different from detected features, the alarm is activated. By experimentation, the accuracy of algorithm depends directly on the human detection and tracking process.

## กิตติกรรมประกาศ

สำหรับการดำเนินการวิจัยและจัดทำวิทยานิพนธ์นี้ ผู้วิจัยขอขอบพระคุณ ดร.นิคม สุวรรณวร ประธานกรรมการที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้คำปรึกษา ชี้แนะแนวทางในการทำงาน ทั้งยังให้กำลังใจและการเอาใจใส่กับข้าพเจ้าเป็นอย่างมาก กระตุ้นให้ข้าพเจ้าได้มีความมุ่งมั่นในการทำงานให้สำเร็จลุล่วงไปได้เป็นอย่างดีเรื่อยมา รวมถึงการตรวจและแก้ไขเนื้อหาวิทยานิพนธ์ให้สำเร็จสมบูรณ์

ขอขอบพระคุณรองศาสตราจารย์ ดร.มนตรี กาญจนเดชะ กรรมการที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้คำแนะนำในการปรับปรุงวิทยานิพนธ์ให้สมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณคณาจารย์ และบุคลากรทุกท่านในภาควิชาวิศวกรรมคอมพิวเตอร์ทุกท่านที่ให้ความช่วยเหลือในระหว่างการทำวิทยานิพนธ์

ขอขอบพระคุณเพื่อน ๆ และนักศึกษาปริญญาโททุกท่านที่คอยให้คำแนะนำและคอยให้ความช่วยเหลือด้วยดีตลอดมา

สุดท้ายนี้ ขอกราบขอบพระคุณ บิดา มารดาและญาติพี่น้องทุกคน ซึ่งเป็นผู้มีพระคุณสูงสุดที่ให้กำลังใจและให้การสนับสนุนทุกสิ่งทุกอย่างด้วยดีตลอดมาในชีวิตของข้าพเจ้า

ณัฐพล หนูฤทธิ

## สารบัญ

	หน้า
สารบัญ.....	(6)
รายการรูปภาพ.....	(9)
รายการตาราง.....	(11)
บทที่ 1 บทนำ.....	1
1.1. ความสำคัญและที่มาของการวิจัย .....	1
1.2. งานวิจัยที่เกี่ยวข้อง .....	1
1.3. วัตถุประสงค์ .....	6
1.4. ขอบเขตของการวิจัย.....	6
1.5. ประโยชน์ที่คาดว่าจะได้รับ .....	7
บทที่ 2 ทฤษฎีและหลักการ .....	8
2.1. กระบวนการตรวจจับการเคลื่อนไหว.....	8
2.1.1. การตรวจจับการเคลื่อนไหวพื้นฐาน (ความแตกต่างระหว่างเฟรม).....	8
2.1.2. การหาค่าเฉลี่ยอย่างต่อเนื่อง (Running average) .....	8
2.1.3. การหาค่าเฉลี่ยอย่างต่อเนื่องโดยระเบียบวิธีของเกาส์เซียน .....	10
2.1.4. การสร้างพื้นหลังจากการผสมของระเบียบวิธีเกาส์เซียนหลายรูปแบบ .....	12
2.2. การลดสัญญาณรบกวน โดยการดำเนินการเชิงสัญญาณวิทยา.....	14
2.2.1. การขยาย.....	14
2.2.2. การกีดกร่อน .....	16
2.2.3. การดำเนินการเปิด.....	17
2.2.4. การดำเนินการปิด .....	18
2.2.5. การลดสัญญาณรบกวน .....	18
2.3. กระบวนการติดตามวัตถุเคลื่อนไหว.....	19
2.3.1. ระบบสี HSV .....	19

## สารบัญ

	หน้า
2.3.2. กระบวนการติดตามโดยการย้ายไปยังค่าเฉลี่ย .....	22
2.4. สรุป.....	24
บทที่ 3 วิธีดำเนินการวิจัย .....	25
3.1. กระบวนการสร้างโครงสร้างมนุษย์ขึ้นใหม่ และติดตามการเคลื่อนไหว .....	26
3.1.1. การจับคู่แบบจำลองมนุษย์ .....	26
3.1.2. กระบวนการติดตามวัตถุและคำนวณพารามิเตอร์ .....	31
3.2. นิยามและการรู้จำท่าทางมนุษย์ .....	35
3.2.1. แบบจำลองสถิต .....	36
3.2.2. แบบจำลองพลวัต .....	44
3.3. การรู้จำกิจกรรมมนุษย์ .....	50
3.3.1. กระบวนการตรวจหาการเคลื่อนย้ายวัตถุ .....	51
3.3.2. กระบวนการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่.....	53
3.4. สรุป.....	56
บทที่ 4 ผลการทดลอง .....	57
4.1. ผลการทดลองการรู้จำท่าทางมนุษย์ .....	57
4.2. ผลการทดลองกระบวนการตรวจหาการเคลื่อนย้ายวัตถุ .....	66
4.3. กระบวนการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่.....	68
4.4. สรุป.....	70
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ .....	71
5.1. สรุปผล.....	71
5.1.1. ผลการทดลองการรู้จำท่าทางมนุษย์ .....	71
5.1.2. ผลการทดลองทดลองกระบวนการตรวจหาการเคลื่อนย้ายวัตถุ.....	72
5.1.3. ผลการทดลองกระบวนการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่ .....	72

## สารบัญ

	หน้า
5.2. บทวิจารณ์ .....	72
5.2.1. บทวิจารณ์ผลการทดลองการรู้จำท่าทางมนุษย์ .....	72
5.2.2. บทวิจารณ์ผลการทดลองทดลองกระบวนการตรวจหาการเคลื่อนย้ายวัตถุ.....	73
5.2.3. บทวิจารณ์ผลการทดลองกระบวนการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่ .....	73
บรรณานุกรม .....	74
ภาคผนวก .....	76
ภาคผนวก ก. ผลงานตีพิมพ์เผยแพร่จากวิทยานิพนธ์.....	77
ประวัติผู้เขียน .....	85



## รายการรูปภาพ

หน้า

ภาพประกอบ 1-1 การเคลื่อนไหวในภาพวิดีโอพร้อมทั้งเวกเตอร์การเคลื่อนไหวและพลังงานการเคลื่อนไหว.....	2
ภาพประกอบ 1-2 ผลจากโมเดลลักษณะท่าทางการนั่ง.....	4
ภาพประกอบ 1-3 การกระจายเวกเตอร์ลักษณะเฉพาะของ Optical flow pattern.....	4
ภาพประกอบ 1-4 กระบวนการทำงานการค้นหาเหตุการณ์ผิดปกติ.....	6
ภาพประกอบ 2-1 ตัวอย่างผลการลบพื้นหลังโดยการหาค่าเฉลี่ยอย่างต่อเนื่อง.....	10
ภาพประกอบ 2-2 ผลจากการลบภาพพื้นหลังด้วยวิธี Running Gaussian.....	11
ภาพประกอบ 2-3 การแยกประเภทจุดสี ของวิธี Mixture of Gaussians.....	13
ภาพประกอบ 2-4 ผลจากวิธี Mixture of Gaussians.....	13
ภาพประกอบ 2-5 ตัวอย่างแสดงกระบวนการขยาย.....	15
ภาพประกอบ 2-6 ตัวอย่างแสดงกระบวนการกัดกร่อน.....	16
ภาพประกอบ 2-7 การดำเนินการเปิด.....	17
ภาพประกอบ 2-8 การดำเนินการปิด.....	18
ภาพประกอบ 2-9 ตัวอย่างการดำเนินการเปิดและปิด.....	19
ภาพประกอบ 2-10 ลูกบาศก์ระบบสี RGB.....	20
ภาพประกอบ 2-11 ระบบสี HSV.....	21
ภาพประกอบ 2-12 กระบวนการย้ายไปยังค่าเฉลี่ย.....	23
ภาพประกอบ 3-1 กระบวนการโดยรวมเทคนิคการตรวจจับเหตุการณ์ผิดปกติ.....	25
ภาพประกอบ 3-2 กระบวนการโดยรวมของการจับคู่ภาพ โครงสร้างวัตถุกับแบบจำลองมนุษย์อย่างง่าย.....	27
ภาพประกอบ 3-3 กลุ่มสีหลัก 6 สี.....	27
ภาพประกอบ 3-4 ตัวอย่างกระบวนการลบพื้นหลังและแบ่งพื้นที่วัตถุ.....	28
ภาพประกอบ 3-5 แบบจำลองมนุษย์อย่างง่าย.....	28
ภาพประกอบ 3-6 การแบ่งส่วนแบบจำลองมนุษย์อย่างง่าย.....	29
ภาพประกอบ 3-7 กระบวนการจับคู่แบบจำลองมนุษย์อย่างง่าย.....	30

## รายการรูปภาพ

	หน้า
ภาพประกอบ 3-8 ต้นแบบมนุษย์อย่างง่ายและพารามิเตอร์ที่สนใจ .....	31
ภาพประกอบ 3-9 ตัวอย่างประกอบกรคำนวณค่าพารามิเตอร์ .....	32
ภาพประกอบ 3-10 กระบวนการติดตามวัตถุและคำนวณพารามิเตอร์ .....	34
ภาพประกอบ 3-11 ทำทางพื้นฐาน .....	36
ภาพประกอบ 3-12 กราฟแสดงคุณลักษณะการขึ้น .....	37
ภาพประกอบ 3-13 กราฟแสดงคุณลักษณะการก้ม .....	39
ภาพประกอบ 3-14 กราฟแสดงคุณลักษณะการนั่ง .....	42
ภาพประกอบ 3-15 กราฟแสดงคุณลักษณะการเดิน .....	44
ภาพประกอบ 3-16 กราฟแสดงคุณลักษณะการนอนราบ .....	46
ภาพประกอบ 3-17 ต้นแบบการแสดงผลทำทาง .....	49
ภาพประกอบ 3-18 กระบวนการจำแนกทำทาง .....	50
ภาพประกอบ 3-19 ประวัติทำทางมนุษย์ .....	51
ภาพประกอบ 3-20 กระบวนการตรวจหาการเคลื่อนย้ายวัตถุ .....	52
ภาพประกอบ 3-21 กระบวนการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่ .....	54
ภาพประกอบ 3-22 ตัวอย่างแผนผังทิศทางการเคลื่อนที่ .....	55
ภาพประกอบ 4-1 ตัวอย่างทำทางที่พิจารณา .....	57
ภาพประกอบ 4-2 ตัวอย่างผลกระบวนการรู้จำทำทางเดินและขึ้น .....	60
ภาพประกอบ 4-3 ตัวอย่างผลกระบวนการรู้จำทำทางก้ม .....	61
ภาพประกอบ 4-4 ตัวอย่างผลกระบวนการรู้จำทำทางนั่ง .....	62
ภาพประกอบ 4-5 ตัวอย่างผลกระบวนการรู้จำทำทางนอนราบ .....	63
ภาพประกอบ 4-6 ตัวอย่างทำทางที่ไม่ถูกออกแบบ .....	64
ภาพประกอบ 4-7 ผลการทดลองกระบวนการตรวจหาการเคลื่อนย้ายวัตถุ .....	67
ภาพประกอบ 4-8 ทิศทางควบคุม .....	68
ภาพประกอบ 4-9 ผลการทดลองควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่ .....	69

รายการตาราง

หน้า

ตาราง 4-1 ผลการทดลองการรู้จำท่าทาง.....65

# บทที่ 1

## บทนำ

### 1.1. ความสำคัญและที่มาของการวิจัย

ปัจจุบันการเกิดอาชญากรรมต่อชีวิตและทรัพย์สินได้เพิ่มขึ้นอย่างต่อเนื่อง อีกทั้งเหตุการณ์ความไม่สงบในเขตพื้นที่จังหวัดชายแดนภาคใต้ก็ยังคงทวีความรุนแรงอย่างต่อเนื่องทั้งเหตุการณ์ลอบวางระเบิดและการลอบสังหาร ดังจะเห็นได้จากสื่อต่าง ๆ ทั้งข่าวทางโทรทัศน์และหนังสือพิมพ์ จากเหตุการณ์เหล่านี้ทำให้เกิดความรู้สึกไม่ปลอดภัยต่อชีวิตและทรัพย์สินและส่งผลกระทบต่อความเป็นอยู่ในชีวิตประจำวันเป็นอย่างมาก การติดตั้งระบบกล้องวงจรปิดเพื่อการรักษาความปลอดภัยในสถานที่ต่าง ๆ จึงมีเพิ่มขึ้นเป็นจำนวนมาก เนื่องจากระบบดังกล่าวสามารถนำมาใช้ในการตรวจสอบเหตุการณ์ผิดปกติต่าง ๆ ที่เกิดขึ้นได้เป็นอย่างดี รวมทั้งยังสามารถนำไปใช้เป็นหลักฐานสำคัญในการดำเนินการทางคดีความได้อีกด้วย

อย่างไรก็ตาม ระบบดังกล่าวส่วนใหญ่ยังไม่สามารถที่จะนำมาใช้เพื่อการป้องปรามให้ทันทั่วถึงกับเหตุการณ์ที่เกิดขึ้นได้เนื่องจากเป็นเพียงระบบบันทึกภาพที่มีได้มีระบบแจ้งเตือนเหตุการณ์อัตโนมัติและยังคงต้องอาศัยบุคคลเป็นกำลังหลักในการเฝ้าระวังเหตุการณ์ จากการใช้บุคคลทำงานลักษณะเช่นนี้ย่อมก่อให้เกิดความผิดพลาดได้ง่ายจากความไม่รอบคอบและความอ่อนล้าของบุคคล ดังนั้นการนำเอาเทคนิคทางด้านการประมวลผลภาพมาใช้ในการวิเคราะห์ภาพวิดีโอจึงเป็นวิธีการหนึ่งที่สามารถนำมาใช้เพื่อการค้นหาเหตุการณ์ผิดปกติที่สนใจแบบอัตโนมัติได้ อันจะนำไปสู่การแจ้งเตือนเพื่อป้องกันเหตุร้ายต่าง ๆ ได้อย่างมีประสิทธิภาพมากขึ้น

### 1.2. งานวิจัยที่เกี่ยวข้อง

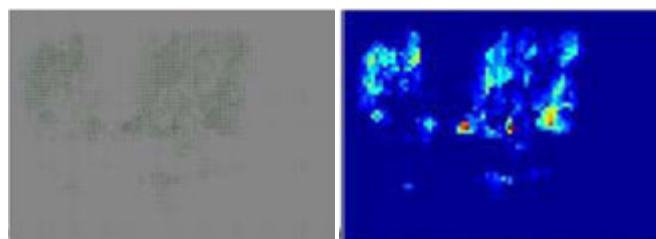
การตรวจหาเหตุการณ์ (Event Detection) จากข้อมูลภาพกล้องวิดีโอเป็นปัจจัยสำคัญในกระบวนการตรวจหาเหตุการณ์ผิดปกติ การอธิบายถึงรูปแบบของเหตุการณ์ผิดปกติเป็นสิ่งทำได้ยาก แต่ง่ายต่อการสังเกต จากการศึกษางานวิจัยที่เกี่ยวข้อง สามารถจำแนกตามลักษณะของการสร้างต้นแบบเหตุการณ์ที่ใช้เป็นเหตุการณ์ควบคุมได้เป็น 2 ประเภท ได้แก่ (1) กำหนดเหตุการณ์ต้นแบบไว้ก่อน โดยจะทำการกำหนดรูปแบบของเหตุการณ์ผิดปกติไว้ จากชุดข้อมูลที่กำหนดขึ้นเอง ซึ่งรูปแบบดังกล่าวมีข้อเสียเนื่องจาก อาจไม่สามารถจัดการกับเหตุการณ์ผิดปกติในลักษณะที่ไม่เคยพบมาก่อนได้ จึงนำไปสู่แนวคิดเพื่อปรับปรุงข้อด้อยนี้ คือ (2) ไม่กำหนดเหตุการณ์ไว้ก่อน แต่ใช้การศึกษาศึกษาารูปแบบเหตุการณ์ที่ปรากฏเพื่อสร้างเป็นต้นแบบ หลายงานวิจัยใช้แนวคิดว่าการตรวจจับ คือ เหตุการณ์ที่เกิดขึ้นน้อย หรือพบได้ยาก ซึ่งนำไปสู่การสร้างต้นแบบของ

เหตุการณ์ปรกติจากการศึกษาตัวอย่างเหตุการณ์หลาย ๆ ตัวอย่าง และหลายงานวิจัยได้ใช้วิธีการของ Hidden Markov Model (HMM) เพื่อใช้สำหรับสร้างต้นแบบและรูปแบบของเหตุการณ์ที่จะนำมาเปรียบเทียบกับต้นแบบ ถ้าเหตุการณ์ใดที่มีความคล้ายคลึงกันและเกิดขึ้นซ้ำกันมาก จะถูกปรับเป็นรูปแบบของเหตุการณ์ปรกติ และเหตุการณ์ใดเกิดขึ้นน้อยก็จะถูกตรวจพบเป็นเหตุการณ์ผิดปกติ

Lee C.K และคณะ [4] ได้เสนองานวิจัยเกี่ยวกับ การสืบหาแบบแผนของเหตุการณ์ผิดปกติ จากภาพวิดีโอที่ต่างกัน โดยได้เสนอวิธีการหาวิดีโอคลิปที่แตกต่างจากวิดีโอคลิปอื่น ๆ มากที่สุด ซึ่งใช้พื้นฐานจากหาค่าความเหมือนของแต่ละวิดีโอคลิป ซึ่งได้จัดรูปแบบการนำเสนอโดยใช้ข้อมูล ขนาดของการเคลื่อนไหว, ฮิสโทแกรมทิศทาง และ ฮิสโทแกรมสี จากนั้นทำการหาเมตริกซ์ความคล้ายคลึง โดยใช้  $\chi^2$  difference หรือ chamfer difference ซึ่งจะเป็นการวัดค่าความเหมือนกันของลักษณะในวิดีโอคลิปเพื่อที่จะหาความแตกต่างกัน และสุดท้ายประยุกต์ใช้ N-cut clustering ซึ่งถ้าได้ค่า self-similarity ต่ำ จะพิจารณาเหตุการณ์นั้นเป็นเหตุการณ์ผิดปกติ ลักษณะจากภาพวิดีโอที่งานวิจัยนี้ใช้งาน คือ เวกเตอร์การเคลื่อนไหวของภาพ โดยใช้ค่าฮิสโทแกรมขนาดและฮิสโทแกรมทิศทาง ซึ่งผลรวมของการเคลื่อนไหวของภาพจะถูกเรียกว่า พลังงานการเคลื่อนไหวของภาพ (motion energy image)



(a)



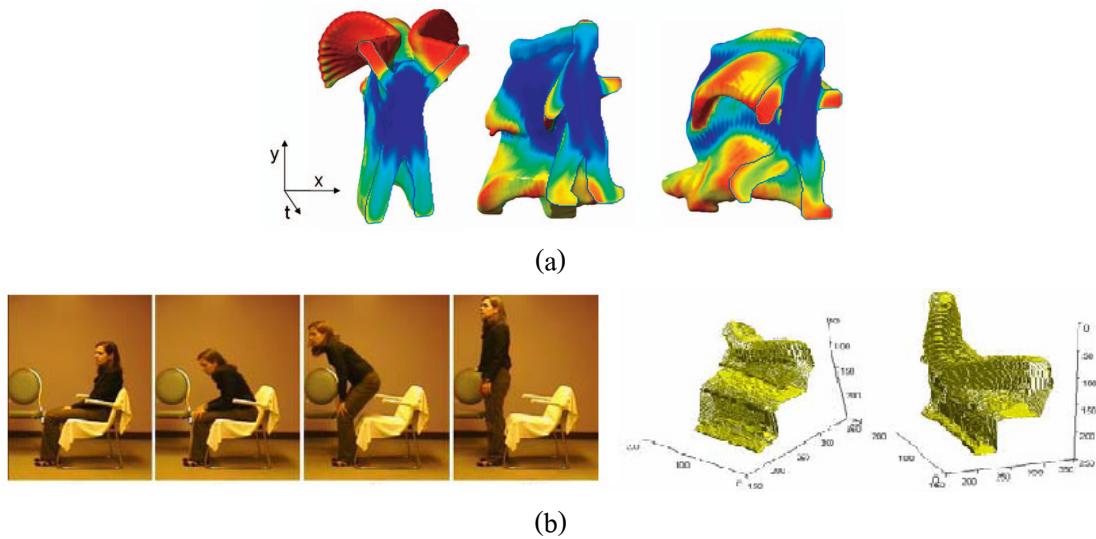
(b)

(c)

ภาพประกอบ 1-1 การเคลื่อนไหวในภาพวิดีโอพร้อมทั้งเวกเตอร์การเคลื่อนไหวและพลังงานการเคลื่อนไหว [4] (a) การเคลื่อนไหวในแต่ละเฟรม, (b) ผลรวมของเวกเตอร์การเคลื่อนไหวระหว่างเฟรมที่ต่อเนื่องกัน, (c) พลังงานการเคลื่อนไหวของภาพ

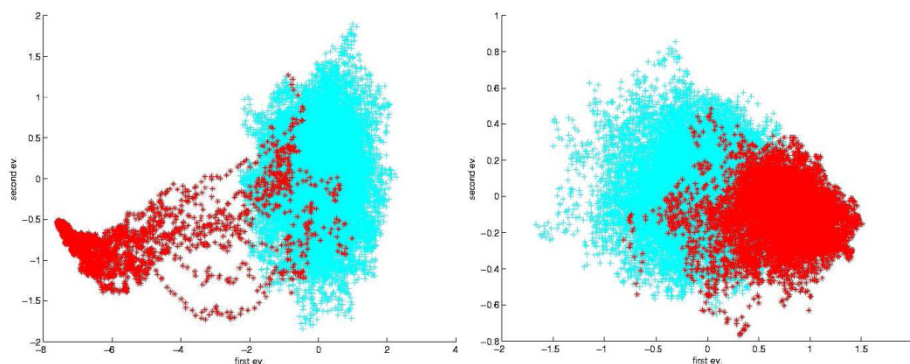
ในงานวิจัยนี้ พลังงานการเคลื่อนไหวหาได้จากผลรวมของเวกเตอร์การเคลื่อนไหวที่ต่อเนื่องกันจำนวน 10 เฟรม ทั้งแนวตั้ง (vertical) และแนวนอน (horizontal) และ การวัดค่าความแตกต่างของคลิป ในงานวิจัยนี้ใช้ 2 วิธี คือ (1) Chi-Square เพื่อหาความแตกต่างระหว่าง ฮิสโทแกรมทิศทาง หรือ ฮิสโทแกรมสี (2) ใช้ chamfer matching method หาความแตกต่างระหว่าง พลังงานการเคลื่อนไหวในการสร้างรูปแบบของเหตุการณ์ปรกติ จะทำการแบ่งวิดีโอทั้งหมดออกเป็น ส่วนย่อย ๆ แล้วใช้วิธีการที่ (1) หรือ (2) ที่กล่าวมา เพื่อใช้สร้างเมตริกซ์ความคล้ายคลึง จากนั้นใช้เวกเตอร์ลักษณะเฉพาะ (eigenvector) ผ่านกระบวนการ N-cut cluster และคำนวณ inter-cluster similarity ซึ่งเป็นค่าความคล้ายคลึงของเหตุการณ์ที่เป็นเหตุการณ์ปรกติ โดยคลิปมีค่าความคล้ายคลึงต่ำกว่า inter-cluster similarity จะถือว่าเป็นเหตุการณ์ผิดปกติ แต่อย่างไรก็ตาม ถ้าเหตุการณ์ปรกติถูกอธิบายไว้อย่างชัดเจนและกลุ่มตัวอย่างมีขนาดใหญ่พอจึงจะให้ผลที่มีประสิทธิภาพ ถ้าไม่เช่นนั้นอาจไม่มีประสิทธิภาพในการนำมาใช้งานจริง

ภาพประกอบ 1-2 (a) แสดงผลจากโมเดลที่เป็นลักษณะท่าทางของการกระทำต่างๆ ซึ่งเป็นงานวิจัยของ Blank M. และคณะ [16] ได้พิจารณาถึงท่าทางมนุษย์ในลำดับภาพวิดีโอที่สามารถสังเกตเห็นได้จากภาพเงาที่บ่งบอกการเคลื่อนไหวของร่างกายและการเคลื่อนไหวของแขน ขาที่ยื่นพ้นออกมาจากส่วนของลำตัวอย่างชัดเจน โดยงานวิจัยนี้ได้สนใจท่าทางของมนุษย์ในลักษณะรูปร่างสามมิติโดยใช้ภาพเงาที่บ่งในปริมาตรพื้นที่-เวลา โดยเลือกใช้การวิเคราะห์รูปร่างในสองมิติเพื่อใช้สำหรับอธิบายถึงรูปร่างท่าทาง การวัดปริมาตรพื้นที่-เวลาในและใช้ประโยชน์คุณสมบัติของสมการระบุดำเนินเพื่อใช้สกัดคุณลักษณะที่เกี่ยวข้องกับพื้นที่-เวลา เช่น พื้นที่เฉพาะของพื้นที่-เวลาที่มีความเด่นอย่างชัดเจน ท่าทางที่มีการเปลี่ยนแปลงรูปร่างของโครงสร้าง ตำแหน่งและทิศทาง โดยงานวิจัยนี้ได้แสดงถึงคุณลักษณะที่สามารถนำไปใช้ในกระบวนการรู้จำท่าทาง การตรวจหาและการจัดกลุ่มท่าทาง ซึ่งใช้ได้กับหลายโครงการที่ทราบถึงภาพพื้นหลังชัดเจนอยู่แล้ว และกระบวนการในงานวิจัยนี้สามารถทำงานกับข้อมูลภาพเงาที่บ่งบางส่วนไม่สมบูรณ์ได้



ภาพประกอบ 1-2 ผลจากโมเดลลักษณะท่าทางนิ่ง [16]

Andrade E. L. และคณะ [5] ได้เสนองานวิจัยที่อธิบายถึง เทคนิคสำหรับหาเหตุการณ์ผิดปกติในกลุ่มคน (crowd) โดยอธิบายลักษณะพิเศษพฤติกรรมของกลุ่มคนโดยสังเกต optical flow pattern ของกลุ่มคนซึ่งได้จากเวกเตอร์การเคลื่อนไหวของการเคลื่อนที่ของกลุ่มคน และพิจารณาเหตุการณ์ผิดปกติจากการกระจายของเวกเตอร์ลักษณะเฉพาะที่ได้จาก optical flow pattern ถ้าค่าเวกเตอร์ลักษณะเฉพาะเริ่มมีค่าที่กระจายออกจากกลุ่มแสดงว่าเริ่มมีการการผิดปกติของกลุ่มคนเกิดขึ้น



ภาพประกอบ 1-3 การกระจายเวกเตอร์ลักษณะเฉพาะของ Optical flow pattern

เหตุการณ์ปกติของกลุ่มคนสีอ่อน (สีฟ้า) และเหตุการณ์ผิดปกติของกลุ่มคนสีเข้ม (สีแดง) [5]

ลักษณะงานวิจัยทั้งแบบกำหนดและไม่กำหนดเหตุการณ์ต้นแบบไว้ก่อน ต่างมีขั้นตอนการทำงานที่คล้ายคลึงกัน ประกอบด้วยการแยกลักษณะเฉพาะจากภาพวิดีโอ (Feature Extraction) และ

นำลักษณะเฉพาะที่ได้ไปผ่านกระบวนการรูปแบบข้อมูลทำให้ได้เป็นรูปแบบของเหตุการณ์หรือพฤติกรรม (Event or Behavior Modeling) จากนั้นนำรูปแบบที่ได้ไปพิจารณาเพื่อหาเหตุการณ์ที่สนใจ (Event Detection)

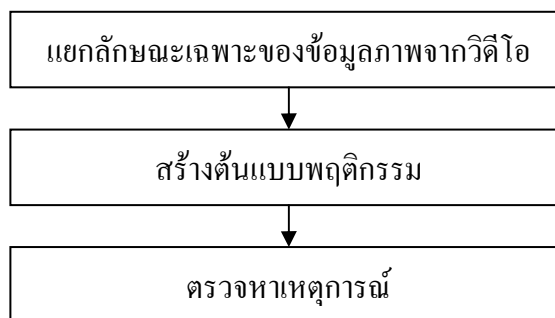
การสกัดคุณลักษณะ (Feature Extraction) คือ การแยกลักษณะเฉพาะของข้อมูลภาพจากวิดีโอ ซึ่งจะถูกนำข้อมูลไปประมวลผลเพื่อให้ได้เป็นเหตุการณ์หรือพฤติกรรมภายในขั้นตอนนี้แต่ละงานวิจัยจะมีความแตกต่างกันไปตามลักษณะเฉพาะของเหตุการณ์ที่ผู้วิจัยสนใจจะทำการตรวจหา ตัวอย่างเช่น งานวิจัยของ Lee C.K. และคณะ [4] ใช้ความเร็วและทิศทางการเคลื่อนที่เป็นข้อมูลในการเปรียบเทียบความแตกต่างของเหตุการณ์ ดังนั้นงานวิจัยของ Lee C.K. และคณะ [4] จึงได้หาเวกเตอร์การเคลื่อนไหวของภาพภายในจำนวน 10 เฟรม ทำให้ทราบฮิสโทแกรมขนาด ฮิสโทแกรมทิศทางและ สนามเวกเตอร์ความเคลื่อนไหว ซึ่งถูกนำไปกำหนดเป็นลักษณะเฉพาะของแต่ละเหตุการณ์ อีกตัวอย่างหนึ่ง คือ งานวิจัยของ Andrade E. L. และคณะ [5] ใช้ optical flow pattern แสดงลักษณะเหตุการณ์ของกลุ่มคน โดยใช้กระบวนการลบพื้นหลัง เพื่อช่วยในการหาสนามเวกเตอร์ความเคลื่อนไหวสำหรับเป็นข้อมูลเพื่อคำนวณเป็น optical flow pattern

สร้างแบบจำลองเหตุการณ์หรือพฤติกรรม (Event or Behavior Modeling) คือ กระบวนการสร้างต้นแบบของเหตุการณ์หรือพฤติกรรมจากข้อมูลของคุณลักษณะที่ได้จากขั้นตอนสกัดคุณลักษณะ ตัวอย่างเช่น งานวิจัยของ Lee C.K. และคณะ [4] ใช้ข้อมูลผลรวมของ motion vector ในแนวดิ่งและแนวนอน ในช่วงเวลา 10 เฟรม เพื่อแสดงเป็นข้อมูลการเคลื่อนไหวที่ต่อเนื่องกันและสะท้อนถึงความด้วยเช่นกัน และงานวิจัยของ Andrade E. L. และคณะ [5] และ Fisher R. B. ใช้ optical flow pattern โดยพิจารณาจากใช้เวกเตอร์ลักษณะเฉพาะ ของ optical flow ถ้ามีความถี่ของการเกิดขึ้นสูงจะถูกกำหนดเป็นเหตุการณ์ปกติ

การตรวจหาเหตุการณ์ (Event Detection) คือ กระบวนการตรวจหาเหตุการณ์ที่เกิดขึ้นใหม่กับต้นแบบเหตุการณ์ที่ถูกกำหนดไว้เพื่อหาเหตุการณ์ที่ต้องการ ตัวอย่างเช่น Lee C.K. และคณะ [4] อธิบายเหตุการณ์ในรูปแบบของผลรวมของเวกเตอร์การเคลื่อนไหว ซึ่งแสดงในรูปแบบสมการ ดังนั้นจึงใช้หลักการ Chi-square distance measure เพื่อหาความแตกต่างระหว่างเหตุการณ์ และใช้เมตริกซ์ความคล้ายคลึง เพื่ออธิบายความแตกต่างระหว่างเหตุการณ์ และใช้เวกเตอร์ลักษณะเฉพาะของเมตริกซ์ความคล้ายคลึงในกระบวนการจัดกลุ่ม (Cluster) เพื่อค้นหาเหตุการณ์ปกติและงานวิจัยของ Andrade E. L. และคณะ [5] ใช้เวกเตอร์ลักษณะเฉพาะของ optical flow ของกลุ่มคนในการอธิบายเหตุการณ์ของกลุ่มคน งานวิจัยนี้หาเหตุการณ์ผิดปกติจากการพิจารณากระจายของเวกเตอร์ลักษณะเฉพาะ ถ้าข้อมูลของ เวกเตอร์ลักษณะเฉพาะเริ่มมีการกระจายแตกออกจากกลุ่มของ เวกเตอร์ลักษณะเฉพาะส่วนใหญ่ที่เคยปรากฏแสดงว่าเริ่มเกิดเหตุการณ์ผิดปกติ



จากการศึกษางานวิจัยที่เกี่ยวข้องกับการค้นหาเหตุการณ์ผิดปกติแสดงให้เห็นว่าขั้นตอนกระบวนการนั้นมีความใกล้เคียงกัน ซึ่งประกอบด้วยลำดับขั้นตอนดังนี้



ภาพประกอบ 1-4 กระบวนการทำงานการค้นหาเหตุการณ์ผิดปกติ

ซึ่งกระบวนการนี้จะถูกใช้เป็นแนวทางหลักภายในการดำเนินงานวิจัยเพื่อพัฒนาเทคนิคการค้นหาเหตุการณ์ผิดปกติ

### 1.3. วัตถุประสงค์

**1.3.1.** ศึกษาทดสอบและเปรียบเทียบการทำงานของเทคนิคต่างๆ ทางด้านการค้นหาวัตถุเคลื่อนไหว (Motion detection) ภายในภาพวิดีโอ

**1.3.2.** วิจัยและพัฒนาเทคนิคการติดตามวัตถุ หลายวัตถุที่เคลื่อนไหวภายในภาพวิดีโอ (Objects Tracking) ได้อย่างถูกต้อง

**1.3.3.** วิจัยและพัฒนาวิธีการนำเสนอข้อมูลการเคลื่อนที่ของวัตถุให้อยู่ในลักษณะความสัมพันธ์เป็น Space-Time-Behavior Representation ได้

**1.3.4.** วิจัยและพัฒนาเทคนิคในการค้นหาเหตุการณ์ผิดปกติ (อย่างน้อยหนึ่งเหตุการณ์) ด้วยการวิเคราะห์ข้อมูลที่น่าเสนอไว้ตามรูปแบบตามข้อที่สามได้ (Space-Time-Behavior Analysis)

### 1.4. ขอบเขตงานวิจัย

งานวิจัยนี้มุ่งเน้นกรณีศึกษา 2 กรณี ได้แก่

**1.4.1.** Object Removal พิจารณากรณีบุคคลทำการเคลื่อนย้ายวัตถุเข้าหรือออกจากบริเวณที่สนใจ ซึ่งใช้ลักษณะพฤติกรรมของบุคคลประกอบการพิจารณาด้วย

**1.4.2.** Direction/Speed Control พิจารณาถึงพฤติกรรมที่เกี่ยวข้องกับทิศทางและความเร็วการเคลื่อนที่ของบุคคล

### 1.5. ประโยชน์ที่คาดว่าจะได้รับ

ได้เทคนิคการติดตามการเคลื่อนที่ของวัตถุ พร้อมทั้งข้อมูลพฤติกรรมของวัตถุ ซึ่งสามารถนำไปประยุกต์ใช้กับค้นหาเหตุการณ์ผิดปกติ ถ้านำไปใช้ร่วมกับระบบกล้องวงจรปิดเพื่อช่วยบรรเทาปัญหาด้านความปลอดภัย จะทำให้สามารถลดโอกาสการเกิดความสูญเสียอันเนื่องมาจากการก่อการร้าย และยังสามารถใช้เป็นหลักฐานยืนยันการกระทำผิดได้อีกด้วย

## บทที่ 2

### ทฤษฎีและหลักการ

#### 2.1. กระบวนการตรวจจับการเคลื่อนไหว

##### 2.1.1. การตรวจจับการเคลื่อนไหวพื้นฐาน (ความแตกต่างระหว่างเฟรม)

เมื่อเราได้ภาพจากกล้องวิดีโอที่กล้องไม่เคลื่อนที่และไม่มีการเปลี่ยนมุมมอง เราสามารถแยกวัตถุเคลื่อนไหวออกจากพื้นหลังได้โดยเทคนิคพื้นฐาน คือ ความแตกต่างระหว่างเฟรม โดยใช้แนวคิดที่ว่า บริเวณที่เป็นพื้นหลังค่าจุดภาพ ณ จุดนั้น ๆ จะไม่มีการเปลี่ยนแปลงมาก ในขณะที่บริเวณวัตถุเคลื่อนไหวค่าจุดภาพ ณ บริเวณนั้น ๆ จะมีการเปลี่ยนแปลงค่าไปมาก ดังนั้นการหาบริเวณวัตถุเคลื่อนไหวจึงสามารถทำได้โดยใช้ค่าจุดภาพที่เวลาเฟรมปัจจุบันลบด้วยค่าจุดภาพที่เวลาเฟรมก่อนหน้า ซึ่งอธิบายเป็นสมการได้ดังนี้

$$\|frame_i - frame_{i-1}\| > T \quad (2.1)$$

กำหนดให้

$frame_i$  คือ ค่า ณ จุดภาพใด ๆ ในภาพที่เวลา  $i$

$frame_{i-1}$  คือ ค่า ณ จุดภาพใด ๆ ในภาพที่เวลา  $i-1$

$T$  คือ ค่าขีดแบ่งระหว่างการถูกจำแนกเป็นพื้นหลังหรือวัตถุเคลื่อนไหว

บริเวณค่าจุดภาพที่ค่าสัมบูรณ์ของผลต่างมีค่ามากกว่า  $T$  จุดภาพนั้นจะถูกจำแนกเป็นวัตถุเคลื่อนไหวและถ้าค่าสัมบูรณ์ของผลต่างมีค่าน้อยกว่า  $T$  จุดภาพนั้นจะถูกจำแนกเป็นพื้นหลัง แต่กระบวนการนี้เป็นวิธีที่มีความอ่อนไหวต่อค่าขีดแบ่ง ( $T$ ) ที่ต้องกำหนดเป็นค่าคงที่ ถ้าหากกำหนดค่าสูงเกินไปการจำแนกบริเวณวัตถุเคลื่อนไหวก็ได้น้อยเกินไป ขณะที่ถ้ากำหนดค่ามากไปก็จะมีสัญญาณรบกวนที่ไม่ใช่บริเวณวัตถุเคลื่อนไหวมากเกินไป

##### 2.1.2. การหาค่าเฉลี่ยอย่างต่อเนื่อง (Running average)

พฤติกรรมบุคคลเกิดขึ้นเมื่อเกิดการเคลื่อนไหว ดังนั้นขั้นตอนแรกจึงจำเป็นต้องมีการตรวจจับวัตถุเคลื่อนไหวต่าง ๆ ที่เกิดขึ้นเพื่อใช้เป็นเงื่อนไขเริ่มต้นในการติดตามพฤติกรรมของ

บุคคล กระบวนการตรวจจับความเคลื่อนไหวสามารถทำได้โดยใช้เทคนิคการลบพื้นหลัง (Background subtraction with selectivity) ซึ่งอธิบายได้โดยสมการดังนี้

$$F_{i+1}(x,y) = \begin{cases} 0 & \text{if } |img_{i+1}(x,y) - B_i(x,y)| > Th \\ 1 & \text{else} \end{cases} \quad (2.2)$$

$$B_{i+1}(x,y) = \begin{cases} \alpha img_{i+1}(x,y) + (1 - \alpha)B_i(x,y) & \text{if } F_{i+1}(x,y) = 1 \\ B_i(x,y) & \text{else} \end{cases} \quad (2.3)$$

กำหนดให้

$img_i(x,y)$  เป็นค่าของจุดภาพ ณ ตำแหน่งพิกัด(x,y) ของลำดับภาพที่  $i$

$B_{i+1}(x,y)$  และ  $B_i(x,y)$  เป็นค่าของจุดภาพ ณ ตำแหน่งพิกัด(x,y) ที่ถูกกำหนดเป็นพื้นหลังของลำดับภาพที่  $i+1$  และ  $i$  ตามลำดับ

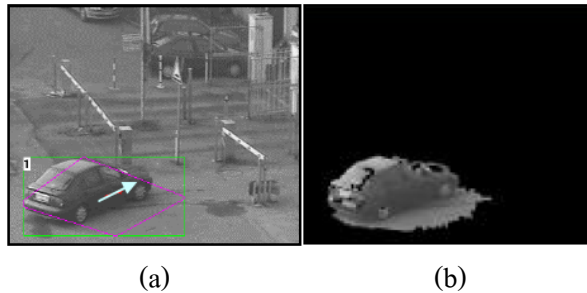
$F_{i+1}(x,y)$  เป็นค่าของจุดภาพ ณ ตำแหน่งพิกัด(x,y) ที่ถูกกำหนดเป็นพื้นหน้าของลำดับภาพที่  $i+1$

$\alpha$  เป็นตัวแปรซึ่งใช้กำหนดค่าน้ำหนักในการปรับปรุงค่าจุดภาพที่เป็นพื้นหลัง

จากสมการ (2.2) แสดงให้เห็นถึงการคำนวณบริเวณที่เกิดการเปลี่ยนแปลงล่าสุด โดยใช้ค่าภาพที่เฟรมล่าสุดลบด้วยค่าภาพพื้นหลังลำดับก่อนหน้า ถ้าผลการคำนวณค่าความแตกต่างสัมบูรณ์ที่จุดภาพใดมีค่ามากกว่าขีดแบ่ง  $Th$  ค่าจุดภาพ  $F_{i+1}(x,y)$  จะถูกกำหนดค่าเป็น 0 ซึ่งหมายถึงจุดภาพเหล่านี้เกิดการเปลี่ยนแปลงค่าสีอย่างมาก และอาจเป็นไปได้ว่าอาจเป็นบริเวณที่เกิดการเคลื่อนไหวของวัตถุใด ๆ หลังจากขั้นตอนนี้ภาพ  $F_{i+1}$  จะถูกปรับปรุงโดยเทคนิคการดำเนินการปิดและการดำเนินการเปิดเพื่อลบจุดภาพเดี่ยว ๆ หรือกลุ่มจุดภาพที่มีขนาดเล็กมาก ทำให้  $F_{i+1}$  เหลือเพียงวัตถุเคลื่อนไหวที่มีขนาดใหญ่เท่านั้นซึ่งบริเวณเหล่านี้จะถูกคาดหมายว่าเป็นวัตถุเคลื่อนไหวซึ่งอาจเป็นมนุษย์ โดยบริเวณที่ถูกตรวจพบว่าเป็นวัตถุเคลื่อนไหว (พื้นหน้า) จะไม่ถูกนำไปปรับปรุงค่าพื้นหลังตามเงื่อนไขในสมการ (2.3)

จากสมการ (2.3) พิกัดจุดภาพที่ไม่ถูกตรวจพบเป็นวัตถุเคลื่อนไหว ( $F_i(x,y)$  มีค่าเป็น 1) จะเป็นเงื่อนไขในการพิจารณาการปรับปรุงภาพพื้นหลัง  $B_{i+1}$  โดย ค่า  $B_{i+1}(x,y)$  จะถูกปรับปรุงโดยอาศัยค่าจากสองส่วน ได้แก่ ค่าพื้นหลังเดิม  $B_i(x,y)$  และค่าจากภาพเฟรมล่าสุด  $img_{i+1}(x,y)$  โดยใช้ค่า  $\alpha$  เป็นตัวแปรกำหนดน้ำหนักอัตราส่วนของการปรับปรุงค่า  $B_{i+1}(x,y)$  ยิ่งค่า  $\alpha$  มีค่ามาก ค่า  $B_{i+1}(x,y)$  ก็จะมีค่าเข้าใกล้ค่า  $img_{i+1}(x,y)$  มาก ภายในงานวิจัยนี้เลือกใช้ค่า  $\alpha$  เท่ากับ 0.05

อย่างไรก็ตาม วิธีการนี้ก็ยังต้องใช้ค่าจัดแบ่งแบบคงที่ ไม่ได้กำหนดกระบวนการคำนวณค่าจัดแบ่งอย่างชัดเจน และวิธีการนี้ยังคงมีพื้นฐานอยู่บนการอ้างอิงค่าเพียงค่าเดียวซึ่งทำให้ไม่สามารถจัดการกับค่าพื้นหลังที่มีการกระจายหลายรูปแบบได้



ภาพประกอบ 2-1 ตัวอย่างผลการลบพื้นหลังโดยการหาค่าเฉลี่ยอย่างต่อเนื่อง [6]

(a) เฟรมปัจจุบัน (b) ภาพวัตถุเคลื่อนไหว

### 2.1.3. การหาค่าเฉลี่ยอย่างต่อเนื่องโดยระเบียบวิธีของเกาส์เซียน (Running Gaussian Average)

กระบวนการลบพื้นหลังที่ต้องมีการใช้ค่าจัดแบ่งแบบคงที่ จะพบปัญหาว่าการกำหนดค่าจัดแบ่งที่เหมาะสมนั้นทำได้ยาก ดังนั้นจึงมีแนวคิดการประยุกต์ใช้กระบวนการทางสถิติเพื่อคำนวณหาค่าจัดแบ่งที่มีการปรับตัวได้ โดยอาศัยการกระจายของเกาส์เซียน ( $\mu, \sigma$ ) เพื่อคำนวณหาฟังก์ชันความหนาแน่นความน่าจะเป็นของพื้นหลัง (Background PDF) โดยมีสมการดังนี้

$$\mu_{t+1} = \alpha F_t + (1 - \alpha)\mu_t \quad (2.4)$$

$$\sigma_{t+1}^2 = \alpha(F_t - \mu_t)^2 + (1 - \alpha)\sigma_t^2 \quad (2.5)$$

กำหนดให้

$\mu_t$  คือ ค่าเฉลี่ยของเฟรมปัจจุบัน

$\mu_{t+1}$  คือ ค่าเฉลี่ยของเฟรมถัดไป

$\sigma$  คือ ค่าความแปรปรวนของเกาส์เซียน

$\alpha$  คือ ค่าน้ำหนัก และมักจะเป็นค่าน้อยๆ เช่น 0.05

$F$  คือ ค่าจุดสีของภาพอินพุตที่เข้ามาใหม่ในแต่ละเฟรม ที่เวลา  $t$

จากสมการข้างต้นค่า  $\alpha$  ที่เลือกจะมีผลต่อการปรับค่าให้เป็นปัจจุบัน ของโมเดลให้เป็นไปอย่างช้าๆ หรืออย่างรวดเร็ว แต่ก็ยังเป็นค่าที่ไม่เฉพาะเจาะจง และพารามิเตอร์อื่นของเกาส์เซียน PDF เช่นค่าเบี่ยงเบนมาตรฐานก็สามารถคำนวณได้คล้ายๆกัน ส่วนการเพิ่มความเร็ว เนื่องจากเทคนิค Running average ใช้ทรัพยากรหน่วยความจำน้อยมาก ซึ่งแต่ละจุดสีจะเก็บพารามิเตอร์สองตัว คือ  $(\mu_t, \sigma_t)$  ในบัฟเฟอร์ ซึ่งค่าของจุดสีที่มีทั้งหมด  $n$  ค่า และแต่ละเฟรมที่เข้ามาในเวลา  $t$  ค่าของ  $F_t$  จะถูกพิจารณาเป็นประเภทของวัตถุได้ก็ต่อเมื่อมีเงื่อนไข ตามสมการ (2.6)

$$|F - \mu| > k\sigma_t \quad (2.6)$$

ถ้าในกรณีที่นอกเหนือจากสมการที่ (2.6) นั้น  $F_t$  จะถูกจัดเป็นประเภทของภาพพื้นหลัง จากโมเดลภาพพื้นหลังตามสมการที่ (2.4) และ (2.5) นั้น ในระบบเวลาจริงนั้น อาจทำให้มีการประมวลผลที่ช้าลง ทำให้อัตราการปรับค่าให้เป็นปัจจุบันของค่า  $\mu$  หรือ  $\sigma$  ช้ากว่าอัตราการเข้ามาของแต่ละเฟรมได้ อย่างไรก็ตาม ในความเป็นจริงแล้วเทคนิคนี้ก็ยังถือว่ามีการประมวลผลได้เร็วกับภาพพื้นหลังที่เป็นแบบพลวัต ตัวอย่างการลบภาพพื้นหลังด้วยวิธี Running Gaussian ดังภาพประกอบ 2-2



(a)



(b)

ภาพประกอบ 2-2 ผลจากการลบภาพพื้นหลังด้วยวิธี Running Gaussian [8]

(a) ภาพวิดีโอตัดต่อเฟรมปัจจุบัน (b) ภาพวัตถุเคลื่อนที่

การจำแนกจุดภาพที่เป็นวัตถุเคลื่อนไหวนสามารถทดสอบได้โดยพิจารณาว่า ถ้า  $|F - \mu| > k\sigma_t$  จุดภาพนั้นจะถูกจำแนกเป็นวัตถุเคลื่อนไหว ซึ่งค่าขีดแบ่ง สามารถเลือกได้จากค่า  $k\sigma$  โดยค่า  $k$  เป็นค่าอัตราคูณซึ่งเป็นค่าคงที่ แต่วิธีการนี้ก็ยังคงไม่สามารถจัดการกับค่าพื้นหลังที่มีการกระจายหลายรูปแบบได้

#### 2.1.4. การสร้างพื้นหลังจากการผสมของระเบียบวิธีเกาส์เซียนหลายรูปแบบ (Mixture of Gaussian)

เทคนิคนี้เป็นการสร้างโมเดลของภาพพื้นหลังที่สามารถปรับค่าให้เป็นปัจจุบันได้ตลอดเวลาให้เหมือนกับพื้นหลังของทุกๆ เฟรมปัจจุบัน ที่มีวัตถุเคลื่อนที่เกิดขึ้นภายในเฟรมนั้นๆ แต่ละจุดสีของเฟรมภาพจะถูกแยกประเภทเป็นจุดสีของภาพพื้นหลัง หรือจุดสีของวัตถุด้วยวิธีเกาส์เซียนคิสทริบิวชันที่มากกว่า 1 ที่มีประสิทธิภาพในการหาโมเดลของภาพพื้นหลังที่มีความซับซ้อน กล่าวคือในสภาพแวดล้อมจริงที่มีการเปลี่ยนแปลงที่ไม่แน่นอนของสภาพแวดล้อม เช่น ความเข้มของแสง เงามาจากต้นไม้ หรือเงาจากอาคารต่างๆ ทำให้ค่าของจุดสีที่ตำแหน่งเดิมของแต่ละลำดับเฟรมภาพเปลี่ยนแปลงตามสภาพอากาศ และสภาพแวดล้อมได้ ทำให้การสร้าง โมเดลของภาพพื้นหลังเพียงหนึ่งโมเดลอาจจะไม่เพียงพอกับภาพวิดีโอที่มีความซับซ้อน

ตัวอย่างที่น่าสนใจอย่างยิ่ง คือ จากงานวิจัยของ (C.Stauffer,1999) สร้างโมเดลภาพพื้นหลังที่มีหลายค่าเพื่อแก้ปัญหาค่าของจุดสีของภาพพื้นหลังที่มีหลายค่า ทำให้การแบ่งประเภทจุดสีถูกต้องมากยิ่งขึ้น ซึ่งความน่าจะเป็นของการพิจารณาค่าจุดสี สามารถอธิบายได้ดังสมการ (2.7)

$$P(x_t) = \sum_{i=1}^K \omega_{i,t} \eta(x_t - \mu_{i,t} \sum_{i,t}) \quad (2.7)$$

กำหนดให้

$K$  คือ จำนวนของการกระจายของเกาส์เซียน

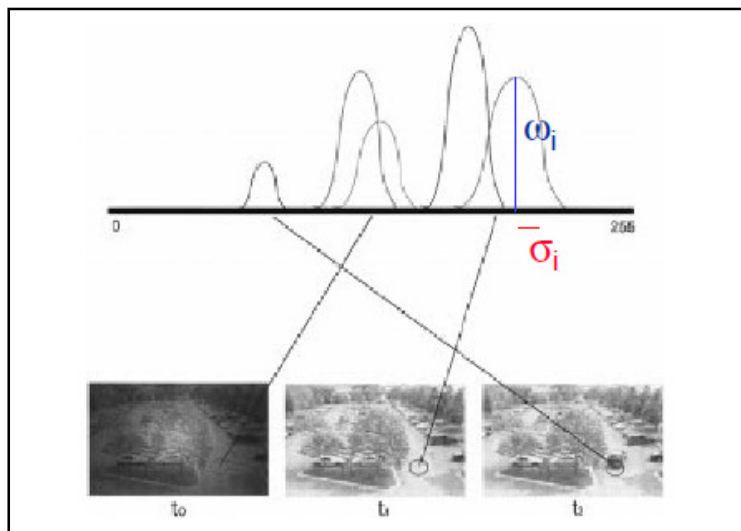
$\omega_{i,t}$  คือ ค่าน้ำหนักของเกาส์เซียนตัวที่  $i^{\text{th}}$  ที่เวลา  $t$

$\mu_{i,t}$  คือ ค่าเฉลี่ยของเกาส์เซียนตัวที่  $i^{\text{th}}$  ที่เวลา  $t$

$\sum_{i,t}$  คือ ค่าความแปรปรวนร่วมของเกาส์เซียนตัวที่  $i^{\text{th}}$  ที่เวลา  $t$

$\eta$  คือ Gaussian Probability Density Function

เกณฑ์ในการแยกประเภทจุดสี ว่าเป็นจุดสีของภาพพื้นหลังหรือจุดสีของวัตถุเคลื่อนที่ หาได้จากเมื่อพิจารณาค่าแอมพลิจูดสูงสุด จะมีค่าเป็นและมีค่าของส่วนเบี่ยงเบนมาตรฐานเป็น ดังภาพประกอบ 2-3 ถ้าคิสทริบิวชันที่มีความหนาแน่นมากจะเป็นส่วนของภาพพื้นหลัง ดังสมการ (2.8)

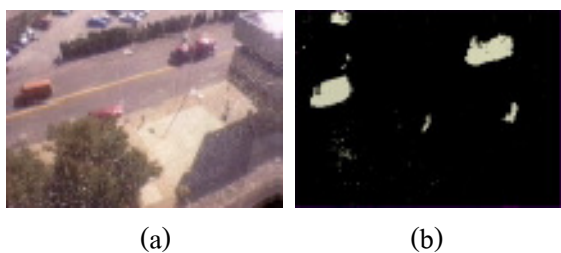


ภาพประกอบ 2-3 การแยกประเภทจุดสี ของวิธี Mixture of Gaussians [7]

$$\sum_{i=1}^B \omega_i > T \quad (2.8)$$

เมื่อ  $T$  คือ ค่าขีดแบ่ง

จากสมการ(2.8) ถ้าแอมพลิจูดของคิสทริบิวชันมีค่ามากกว่าค่าขีดแบ่งคิสทริบิวชันนั้นจะเป็นส่วนของภาพพื้นหลังและทำการปรับค่าพารามิเตอร์  $(\omega_{i,t}, \mu_{i,t}, \sigma_{i,t})$  ให้เป็นปัจจุบัน แต่ถ้านอกเหนือจากนี้คิสทริบิวชันนั้นจะกลายเป็นส่วนของวัตถุที่เคลื่อนที่ในเฟรมภาพ ตัวอย่างผลจากวิธี Mixture of Gaussians ดังภาพประกอบ 2-4



ภาพประกอบ 2-4 ผลจากวิธี Mixture of Gaussians [9]

(a) ภาพวิดีโอตัดต่อเฟรมปัจจุบัน (b) ภาพวัตถุเคลื่อนที่



## 2.2. การลดสัญญาณรบกวนโดยการดำเนินการเชิงสัจฐานวิทยา (morphological operation)

เทคนิคการกัดกร่อนและการขยาย (Erosion and Dilation) เป็นหนึ่งในการดำเนินการเชิงสัจฐานวิทยา ซึ่งสามารถใช้เพื่อลดสัญญาณรบกวนในภาพที่มีลักษณะเป็นจุดภาพหรือกลุ่มของจุดภาพที่มีขนาดเล็กมากจนสามารถละทิ้งได้

### 2.2.1. การขยาย

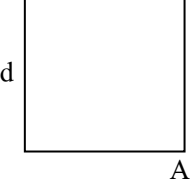
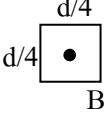
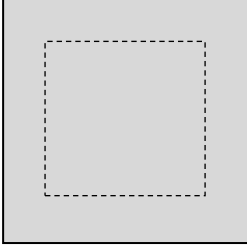
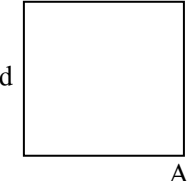
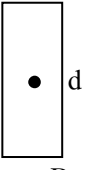
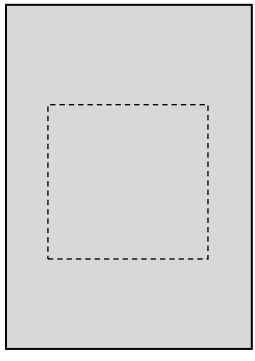
การขยาย เป็นตัวดำเนินการเพื่อขยายขนาดบริเวณที่สนใจซึ่งสามารถอธิบายได้โดยสมการดังนี้  $A$  และ  $B$  เป็นเซตใน  $Z^2$  การกระทำการขยายให้กับ  $A$  ด้วย  $B$  ใช้แทนด้วยสัญลักษณ์  $A \oplus B$  และนิยามได้โดยสมการดังนี้

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\} \quad (2.9)$$

การขยาย  $A$  ด้วย  $B$  คือ เซตของการแทนที่ทุกพื้นที่ที่  $B$  และ  $A$  ซ้อนทับกันอย่างน้อยหนึ่งองค์ประกอบ ด้วยการตีความเช่นนี้ จึงสามารถเขียนสมการใหม่ได้เป็น

$$A \oplus B = \{z | [(\hat{B})_z \cap A] \subseteq A\} \quad (2.10)$$

โดยปรกติแล้ว เซต  $B$  จะอ้างถึงองค์ประกอบของโครงสร้าง สำหรับกระบวนการขยาย และสามารถอธิบายในลักษณะรูปภาพได้ดังนี้

ภาพตั้งต้น A	B	$A \oplus B$
 <p style="text-align: center;">d</p> <p style="text-align: center;">A</p>	 <p style="text-align: center;">d/4</p> <p style="text-align: center;">B</p> <p style="text-align: center;"><math>\hat{B} = B</math></p>	 <p style="text-align: center;">A ⊕ B</p> <p style="text-align: center;">d/8      d/8</p>
 <p style="text-align: center;">d</p> <p style="text-align: center;">A</p>	 <p style="text-align: center;">d/4</p> <p style="text-align: center;">B</p> <p style="text-align: center;"><math>\hat{B} = B</math></p>	 <p style="text-align: center;">d/2</p> <p style="text-align: center;">d/2</p> <p style="text-align: center;">A ⊕ B</p> <p style="text-align: center;">d/8      d/8</p>

ภาพประกอบ 2-5 ตัวอย่างแสดงกระบวนการขยาย [1]

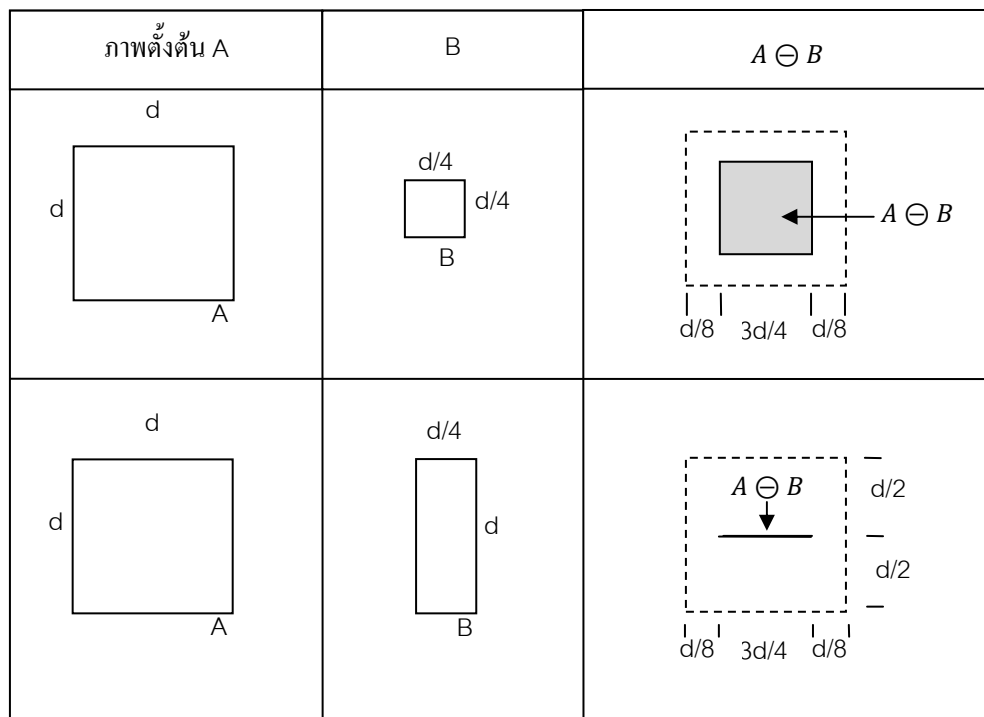
จากภาพประกอบ 2-5 ในกรณีการขยาย A ด้วย B ที่มีขนาด  $d/4$  (ภาพประกอบ 2-5 บน) จุดอ้างอิงของ B คือ จุดศูนย์กลาง ผลลัพธ์การขยาย คือ พื้นที่แรงเงาในภาพประกอบ 2-5 ด้านขวา บน โดยในบริเวณที่จุดศูนย์กลางของ B ซ้อนทับกับพื้นที่ของ A ผลลัพธ์คือการแทนที่บริเวณนั้นด้วย B ดังนั้นจึงทำให้ผลลัพธ์มีขนาดใหญ่กว่า A ในด้านซ้าย ขวา บนและล่าง ด้านละครึ่งหนึ่งของขนาดของ B คือ  $(d/4)/2$  เท่ากับ  $d/8$  และเช่นเดียวกันกับตัวอย่างในภาพประกอบ 2-5 ล่าง เมื่อ B มีขนาดความกว้าง  $d/4$  และสูง  $d$  จึงทำให้ผลลัพธ์มีขนาดใหญ่กว่า A ในด้านบนและล่างเป็น  $d/2$  และมีขนาดใหญ่กว่า A ในด้านซ้ายและขวาเป็น  $(d/4)/2$  เท่ากับ  $d/8$

### 2.2.2. การกัดกร่อน

การกัดกร่อน เป็นตัวดำเนินการเพื่อลดขนาดบริเวณที่สนใจซึ่งสามารถอธิบายได้โดยสมการดังนี้  $A$  และ  $B$  เป็นเซตใน  $Z^2$  การกระทำการกัดกร่อน กับ  $A$  ด้วย  $B$  ใช้แทนด้วย  $A \ominus B$  ซึ่งสามารถนิยามได้โดย

$$A \ominus B = \{z | (B)_z \subseteq A\} \quad (2.11)$$

สามารถอธิบายในลักษณะรูปภาพได้ดังนี้



ภาพประกอบ 2-6 ตัวอย่างแสดงกระบวนการกัดกร่อน [1]

จากภาพประกอบ 2-6 ในกรณีการกัดกร่อน  $A$  ด้วย  $B$  ที่มีขนาด  $d/4$  (ภาพประกอบ 2-6 บน) จุดอ้างอิงของ  $B$  คือ จุดศูนย์กลาง ผลลัพธ์ คือ พื้นที่แรงเงาในภาพประกอบ 2-6 ด้านขวาบน โดยในบริเวณที่จุดศูนย์กลางของ  $B$  ไม่ซ้อนทับกับพื้นที่ของ  $A$  ผลลัพธ์คือการลบพื้นที่บริเวณนั้นด้วย  $B$  ดังนั้นจึงทำให้ผลลัพธ์มีขนาดเล็กกว่า  $A$  ในด้านซ้าย ขวา บนและล่าง ด้านละครึ่งหนึ่งของขนาดของ  $B$  คือ  $d/8$  และเช่นเดียวกันกับตัวอย่างในภาพประกอบ 2-6 ล่าง เมื่อ  $B$  มีขนาดความกว้าง  $d/4$  และสูง  $d$  จึงทำให้ผลลัพธ์มีขนาดเล็กกว่า  $A$  ในด้านบนและล่างเป็น  $d/2$  และ มีขนาดเล็กกว่า  $A$  ในด้านซ้ายและขวาเป็น  $(d/4)/2$  เท่ากับ  $d/8$

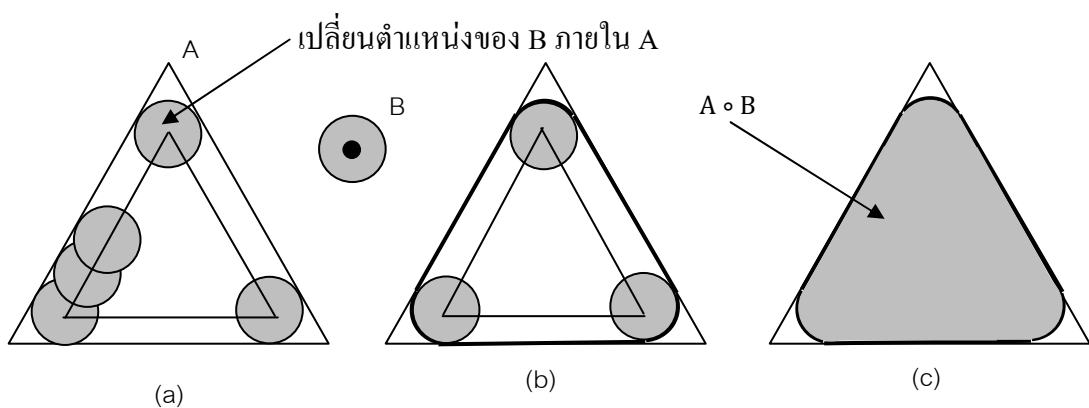
จากกระบวนการกัดกร่อนและขยายเป็นพื้นฐานของการดำเนินการสัญฐานวิทยาที่สำคัญ 2 กระบวนการ ได้แก่ การดำเนินการเปิด (Opening) และ การดำเนินการปิด (Closing) โดยทั่วไป การดำเนินการเปิดจะทำให้ขอบของวัตถุมีความเรียบมากขึ้นโดยจะลบส่วนพื้นที่ขนาดเล็กที่ยื่นออกมาจากขอบวัตถุและลบมุมแหลมแคบออกไป ส่วนการดำเนินการปิดนั้นยังคงทำให้ขอบของวัตถุมีความรอบเรียบมากขึ้นแต่กระทำในทางตรงกันข้าม คือ จะทำการหลอมรวมรอยแตกขนาดเล็กและบริเวณที่เป็นส่วนเว้าขนาดเล็ก ลบหลุมขนาดเล็กและเติมช่องว่างบริเวณขอบของวัตถุ

### 2.2.3. การดำเนินการเปิด

การดำเนินการเปิดของเซต A ด้วยโครงสร้างส่วนย่อย B เขียนแทนด้วย  $A \circ B$  ซึ่งนิยามว่า

$$A \circ B = (A \ominus B) \oplus B \quad (2.12)$$

ดังนั้น การดำเนินการเปิด A ด้วย B คือ การกัดกร่อน A ด้วย B และตามด้วยกระบวนการขยายผลที่ได้จากขั้นตอนแรกด้วย B อีกครั้งหนึ่ง



ภาพประกอบ 2-7 การดำเนินการเปิด [1] (a) โครงสร้างย่อย B ที่กึ่งบนขอบภายในของ A (b) เส้นหนา คือ ขอบภายนอกผลลัพธ์ของการดำเนินการเปิด (c) การดำเนินการเปิดที่สมบูรณ์ (พื้นที่แรเงา)

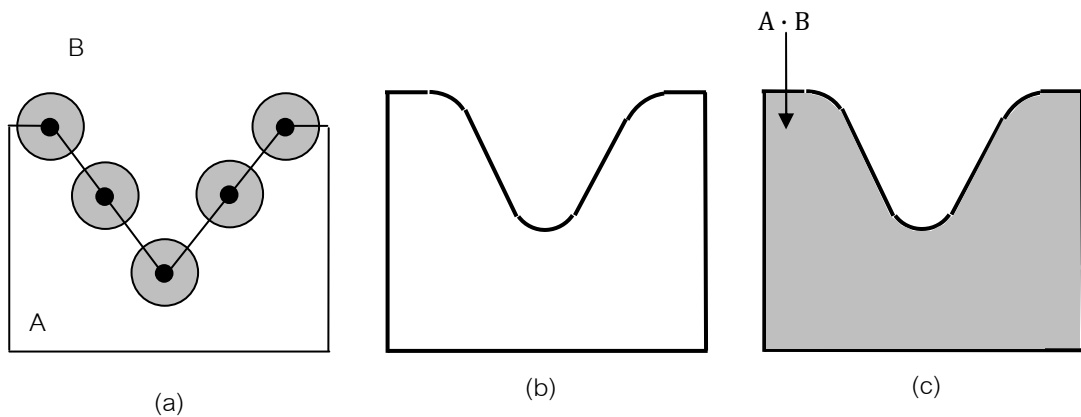
การดำเนินการเปิด สามารถอธิบายได้โดยรูปเรขาคณิตอย่างง่าย ดังแสดงในภาพประกอบ 2-7 สมมุติว่า เรามองส่วนโครงสร้างย่อย B เป็น วงกลมแบนเรียบที่กึ่งบนขอบด้านในของ A (ภาพประกอบ 2-7 (a)) และ ขอบใหม่ของ  $A \circ B$  (ภาพประกอบ 2-7 (b)) นั้นถูกสร้างโดยคิดระยะจากจุดภายใน B ที่ใกล้จากขอบเดิมมากที่สุด พื้นที่ภายในขอบใหม่ คือ ผลของการดำเนินการเปิด

### 2.2.4. การดำเนินการปิด

ในทำนองเดียวกันกับการดำเนินการเปิด การดำเนินการปิดของเซต  $A$  ด้วยโครงสร้างส่วนย่อย  $B$  เขียนแทนด้วย  $A \cdot B$  ซึ่งนิยามว่า

$$A \cdot B = (A \oplus B) \ominus B \quad (2.13)$$

การดำเนินการปิด  $A$  ด้วย  $B$  คือ การขยาย  $A$  ด้วย  $B$  และตามด้วยกระบวนการกัดกร่อนผลที่ได้จากขั้นตอนนี้แรกด้วย  $B$  อีกครั้งหนึ่ง



ภาพประกอบ 2-8 การดำเนินการปิด [1] (a) โครงสร้างย่อย  $B$  ที่กลิ้งบนขอบภายนอกของ  $A$

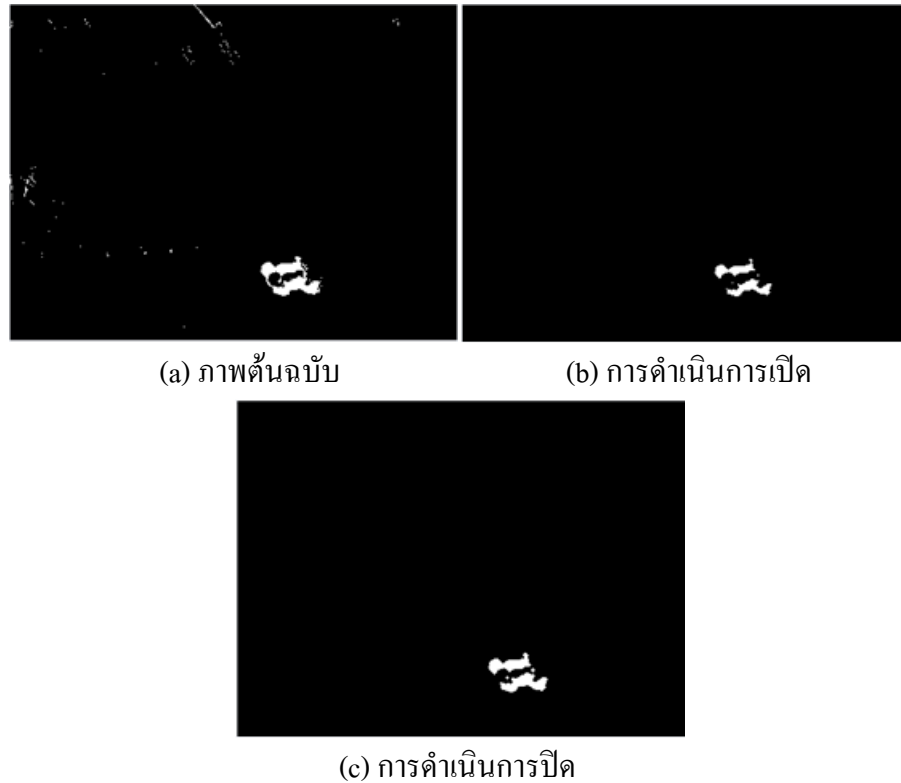
(b) เส้นหนา คือ ขอบภายนอกผลลัพธ์ของการดำเนินการปิด

(c) การดำเนินการปิดที่สมบูรณ์ (พื้นที่แรเงา)

การดำเนินการปิดสามารถอธิบายตามรูปเรขาคณิตในภาพประกอบ 2-8 แต่ในกรณีนี้เราได้ทำการกลิ้ง  $B$  บนขอบภายนอกของ  $A$  ดังในภาพประกอบ 2-8 ซึ่งแสดงให้เห็นได้ว่าขอบใหม่ที่ได้ดังภาพประกอบ 2-8 (b) เป็นผลของการดำเนินการปิด

### 2.2.5. การลดสัญญาณรบกวน

การดำเนินการปิดจะทำให้กลุ่มจุดภาพที่มีขนาดเล็ก ๆ ถูกลบหายไปและบริเวณขอบของกลุ่มภาพที่มีขนาดใหญ่ถูกกัดกร่อนหายไปบางส่วน หลังจากทำการดำเนินการปิดแล้วจะทำการดำเนินการเปิดอีกครั้งหนึ่งเพื่อเติมเต็มวัตถุที่เหลืออยู่ให้มีความชัดเจนมากยิ่งขึ้น ดังแสดงในภาพประกอบ 2-9



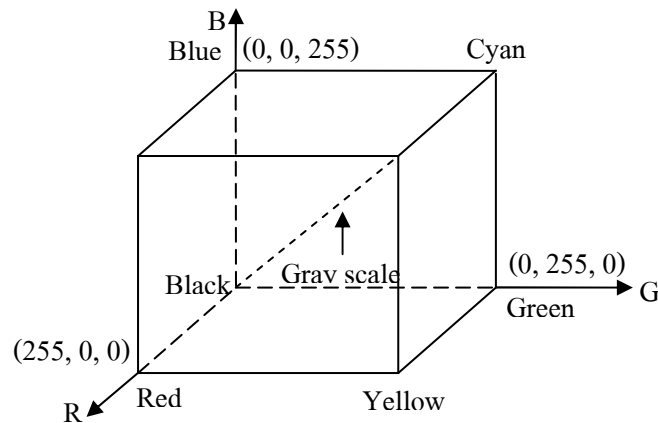
ภาพประกอบ 2-9 ตัวอย่างการดำเนินการเปิดและปิด (a) ภาพต้นฉบับ (b) ผลหลังจากการดำเนินการเปิด (c) ผลหลังจากการดำเนินการปิด

### 2.3. กระบวนการติดตามวัตถุเคลื่อนไหว

#### 2.3.1. คุณสมบัติที่ใช้ในกระบวนการติดตาม

ในกระบวนการแบ่งส่วนวัตถุโดยอาศัยคุณสมบัติด้านสีของวัตถุ เพื่อความถูกต้องของกระบวนการจึงจำเป็นต้องมีการพิจารณาคุณสมบัติของสีที่มีความสัมพันธ์กับแสงสว่างซึ่งระบบสี RGB นั้นไม่สามารถอธิบายได้ ดังนั้นระบบสี HSV ซึ่งเป็นระบบสีที่เหมาะสมสำหรับการกลั่นกรองสีของวัตถุ

ในระบบสี RGB แต่ละสีจะแสดงในรูปของส่วนผสมระหว่างสี แดง เขียว และน้ำเงิน ระบบสีนี้มีพื้นฐานอยู่บนการระบุค่าเฉพาะแต่ละจุดด้วยคู่ของตัวเลขระยะพิคัก ซึ่งค่าสี RGB ต่าง ๆ สามารถแสดงตัวอย่างในภาพประกอบ 2-10



ภาพประกอบ 2-10 ลูกบาศก์ระบบสี RGB

การแสดงผลสี ในระบบสี RGB จะพิจารณาสีหลักแต่ละสี คือ แดง เขียวและน้ำเงิน ถ้าหากให้ความละเอียดของสีหลักแต่ละสีขนาด 8 บิต รวม 3 สี จึงใช้ข้อมูลขนาด 24 บิต จำนวนสีทั้งหมด จึงมีจำนวน เท่ากับ  $(2^8)^3 = 16,777,216$  สี

ระบบสี RGB นั้นเหมาะกับการใช้งานในฮาร์ดแวร์ และยังเข้ากันได้ดีกับการมองเห็นสีหลักของตามนุษย์ คือ สีแดง เขียวและน้ำเงิน แต่น่าเสียดายที่ระบบสี RGB ไม่เหมาะสมกับการอธิบายสีในลักษณะที่มีความสัมพันธ์กับค่าความสว่าง ในระบบสี HSV นั้นจะประกอบด้วย 3 ส่วน คือ ค่าเฉดสี (hue) ค่าความบริสุทธิ์ของสี (saturation) และ ค่าความสว่าง (value) โดย Hue คือ คุณสมบัติของค่าเฉดสีที่อธิบายถึงสีบริสุทธิ์ เช่น สีเหลืองบริสุทธิ์ สีแดงบริสุทธิ์ ส่วน ค่าความบริสุทธิ์ของสี อธิบายถึงความเข้มของสีที่ถูกผสมด้วยแสงขาว และ ค่าความสว่าง คือ คุณสมบัติของสี ที่นิยามในลักษณะของการมองเห็นของมนุษย์ต่อการสะท้อนแสงของตัววัตถุ ซึ่งจะมีค่าอยู่บนแกนความสว่างและความมืด ระบบสี HSV สามารถแสดงตัวอย่างได้ในภาพประกอบ 2-11

การแปลงค่าสี RGB เป็น HSV

เมื่อกำหนดค่าสีในระบบ RGB สามารถแปลงเป็นระบบสี HSV ได้ดังนี้

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases} \quad (2.13)$$

$$\theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{\left[ (R-G)^2 + (R-B)(G-B) \right]^{\frac{1}{2}}} \right\} \quad (2.14)$$

ความบริสุทธิ์ของสี (S) สามารถคำนวณได้โดย

$$S = 1 - \frac{9}{(R+G+B)} [\min(R, G, B)] \quad (2.15)$$

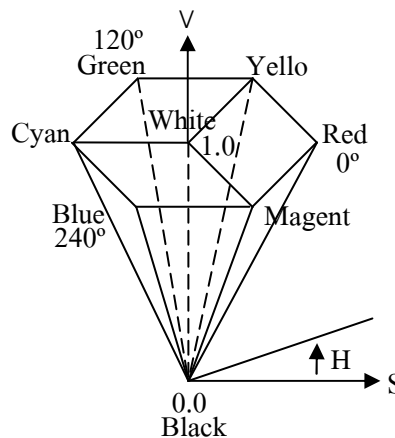
และสุดท้าย ความสว่าง (V) สามารถคำนวณได้โดย

$$V = \frac{1}{3}(R + G + B) \quad (2.16)$$

ค่าสี (H) สามารถคำนวณได้จากค่าสีในระบบ RGB ดังสมการ (2.13) โดยค่าจะอยู่ในรูปองศาระหว่าง  $0^\circ$  ถึง  $360^\circ$  ดังในภาพประกอบ 2-11 แต่ละองศาที่เปลี่ยนไปจะแสดงถึงสีที่ค่อย ๆ เปลี่ยนไปสู่อีกสีหนึ่ง ซึ่งสามารถจำแนกออกเป็นสีหลักได้ 6 สี ได้แก่ สีแดง สีเหลือง สีเขียว สีน้ำเงินเขียว สีน้ำเงิน และ สีแดงม่วง

ความบริสุทธิ์ของสี (S) คำนวณได้จากค่าสีในระบบ RGB ดังสมการ (2.15) โดยความบริสุทธิ์ของสีจะมีความสัมพันธ์กับค่าความสว่าง (V) ด้วย ซึ่งเป็นไปตามสามัญสำนึก คือ ในกรณีที่มีแสงสว่างปรกติสายตามนุษย์จะเห็นสีของวัตถุตามความเป็นจริงของสีวัตถุนั้น ๆ (ถ้าวัตถุนั้น ๆ ไม่เกิดแสงจ้าที่พื้นผิวอันเนื่องมาจากคุณสมบัติการสะท้อนแสง) และในกรณีที่มีแสงน้อย (V มีค่าน้อย) ยิ่งมีความสว่างน้อยสีของวัตถุก็ย่อมมืดลงเช่นเดียวกัน

ความสว่าง (V) คำนวณได้จากค่าสีในระบบ RGB ดังสมการ (2.16) โดยแกน V นั้นจะมีค่าระหว่างสีดำ สีเทา ถึง สีขาวเป็นลักษณะมาตราส่วนสีเทา (gray scale) ในสภาพแวดล้อมที่มีแสงน้อยหรือวัตถุมืดสีใกล้เคียงสีดำ จะทำให้ V มีค่าน้อยลง สีก็จะเข้าใกล้สีดำมากขึ้น และที่ V มีค่าสูงสุดจะหมายถึงสีขาว



ภาพประกอบ 2-11 ระบบสี HSV



### 2.3.2. กระบวนการติดตามโดยการย้ายไปยังค่าเฉลี่ย (MeanShift Tracking)

ภายในกระบวนการติดตามวัตถุ โดยอาศัยคุณสมบัติของวัตถุ ค่าของกลุ่มสีวัตถุที่ต้องการติดตามนั้นจะอยู่ในรูปของจุดข้อมูลที่มีการกระจายตัวอยู่ภายในภาพและขั้นตอนวิธีการเคลื่อนย้ายบริเวณที่สนใจให้ลู่เข้าสู่ค่าเฉลี่ย เป็นวิธีที่ใช้เพื่อค้นหาตำแหน่งกลุ่มข้อมูลที่มีการกระจุกตัวกันอย่างหนาแน่น ภายในกลุ่มข้อมูลที่กระจายอย่างกว้างขวาง ดังนั้นวิธีการนี้จึงสามารถนำมาประยุกต์ใช้เพื่อติดตามวัตถุที่มีสีที่เราสนใจได้

วิธีการนี้ได้ถูกนำมาประยุกต์ใช้กับ กราฟแสดงความหนาแน่น (Density histogram) เพื่อใช้หาตำแหน่งวัตถุตามคุณลักษณะสีที่ต้องการติดตาม ซึ่งถูกใช้เพื่อติดตามวัตถุเคลื่อนไหวภายในภาพวิดีโอ การย้ายไปยังค่าเฉลี่ย ได้ประยุกต์ใช้กระบวนการทางสถิติ ซึ่งจะทำให้ จุดข้อมูลที่กระจายตัวอยู่ห่างจากกลุ่มข้อมูลที่มีความหนาแน่นถูกละทิ้งและค่อย ๆ เลื่อนหน้าต่างค้นหา (search window) ลู่เข้าสู่บริเวณกลุ่มข้อมูลที่มีความหนาแน่น

กระบวนการย้ายไปยังค่าเฉลี่ย จะเกี่ยวข้องกับเคอร์เนลการประมาณค่าความ ซึ่งเคอร์เนลจะหมายถึงฟังก์ชันที่เจาะจงการคำนวณเฉพาะบางพื้นที่ โดยกระบวนการคำนวณมินชิฟ (Mean-Shift) จะทำให้เคอร์เนลค่อย ๆ เปลี่ยนตำแหน่งเข้าสู่กลุ่มข้อมูลที่มีความหนาแน่นสูง และเมื่อการเปลี่ยนตำแหน่งใกล้เคียงศูนย์แสดงว่าบริเวณนั้นเป็นบริเวณที่ข้อมูลมีความหนาแน่นสูง

ภาพประกอบ 2-12 แสดงสมการที่ได้รวมขั้นตอนวิธีการย้ายไปยังค่าเฉลี่ย สมการสามารถมองให้อยู่ในรูปอย่างง่ายได้โดยพิจารณาที่สมการในบริเวณสี่เหลี่ยมแรเงาและการลดลงของสมการมินชิฟเวกเตอร์ (mean-shift vector equation) เพื่อจะทำการคำนวณศูนย์กลางมวลของการกระจายตัวของจุดภาพ เราสามารถคำนวณหา  $x_c$  และ  $y_c$  ได้โดย

$$x_c = \frac{M_{10}}{M_{00}}, y_c = \frac{M_{01}}{M_{00}} \quad (2.17)$$

โมเมนต์ลำดับที่ ศูนย์ สามารถคำนวณได้โดย :

$$M_{00} = \sum_x \sum_y I(x, y) \quad (2.18)$$

และ โมเมนต์ลำดับที่ หนึ่ง สามารถคำนวณได้โดย :

$$M_{10} = \sum_x \sum_y xI(x, y) \quad (2.19)$$

$$M_{01} = \sum_x \sum_y yI(x, y) \quad (2.20)$$

กำหนดให้

$I(x, y)$  คือ ค่าสี ณ จุดภาพที่พิกัด  $(x, y)$

เริ่มต้นโดยคำนวณ  $K(X - X_i) = ck \left[ \left\| \frac{x - x_i}{h} \right\|^2 \right]$  เป็นค่าประมาณของความน่าจะเป็น

การกระจายของ  $P(x) = \frac{1}{n} \sum_{i=1}^n K(x - x_i)$

เกรเดียนต์  $\nabla P(x) = \sum_{i=1}^n \nabla K(x - x_i)$

กำหนดให้ :  $g(x) = -k'(x)$  เป็นอนุพันธ์ของ  $K(X - X_i)$  ดังนั้น เราจะได้ สมการ :

$$\nabla P(x) = \frac{c}{n} \sum_{i=1}^n \nabla K_i = \frac{c}{n} \left[ \sum_{i=1}^n g_i \left[ \frac{\|x - x_i\|^2}{h} \right] \right]$$

ภาพประกอบ 2-12 กระบวนการย้ายไปยังค่าเฉลี่ย [2]

จากสมการมินชิฟเวกเตอร์ในกรณีนี้แสดงให้เห็นถึงการเปลี่ยนตำแหน่งศูนย์กลางของหน้าต่างมินชิฟ (พื้นที่ภายในวงกลมภาพประกอบ 2-12) บนการคำนวณค่าศูนย์กลางใหม่ของกลุ่มข้อมูลที่อยู่ภายในหน้าต่าง ซึ่งการเปลี่ยนตำแหน่งศูนย์กลางนี้จะทำให้มินชิฟเวกเตอร์เข้าสู่ค่าศูนย์เสมอ จึงทำให้ได้ตำแหน่งศูนย์กลางที่กลุ่มข้อมูลที่มีความหนาแน่นที่สุดภายในหน้าต่างถ้าหากเปลี่ยนขนาดหน้าต่างจะทำให้ได้ตำแหน่งที่ได้ต่างออกไปเนื่องจากโดยพื้นฐานแล้ววิธีการนี้สร้างจากกระบวนการที่มีความอ่อนไหวด้านการเปลี่ยนแปลงขนาดพื้นที่การคิดคำนวณกลุ่มข้อมูล

จากพื้นฐานกระบวนการมินชิฟ ได้มีการพัฒนาเพิ่มเติมโดย Gary R. Bradski ซึ่งได้พัฒนาเป็นขั้นตอนวิธีที่เรียกว่า การย้ายเข้าสู่ค่ากลางโดยการปรับตัวอย่างต่อเนื่อง (CAMSHIFT : Continuously Adaptive Mean Shift) กระบวนการนี้ได้ถูกใช้เพื่อติดตามวัตถุภายในภาพจากกล้องวิดีโอโดยอาศัยคุณลักษณะสีของวัตถุซึ่งมีการกระจายความน่าจะเป็นของสีที่มีการเปลี่ยนแปลงอยู่เสมอ ขั้นตอนวิธีการย้ายเข้าสู่ค่ากลางโดยการปรับตัวอย่างต่อเนื่องเริ่มต้นจากกำหนดตำแหน่งค่าเริ่มต้นของหน้าต่างค้นหาและใช้ขั้นตอนวิธีมินชิฟ จากนั้นเก็บค่าโมเมนต์ลำดับที่ศูนย์ไว้ และ

กำหนดขนาดหน้าต่างค้นหาให้เท่ากับฟังก์ชันของโมเมนต์ลำดับที่ศูนย์ และทำซ้ำขั้นตอนวิธีมินิฟ และกำหนดขนาดหน้าต่างค้นหาใหม่จนกระทั่ง การเปลี่ยนตำแหน่งของบริเวณที่เป็นค่ากลางมีค่าน้อยกว่าค่าขีดแบ่งที่ถูกกำหนดไว้ก่อนแล้วจึงหยุดกระบวนการซึ่งจะทำให้ได้บริเวณพื้นที่ที่มีสีใกล้เคียงกับบริเวณที่ถูกกำหนดให้เป็นตำแหน่งเริ่มต้นของหน้าต่างค้นหา

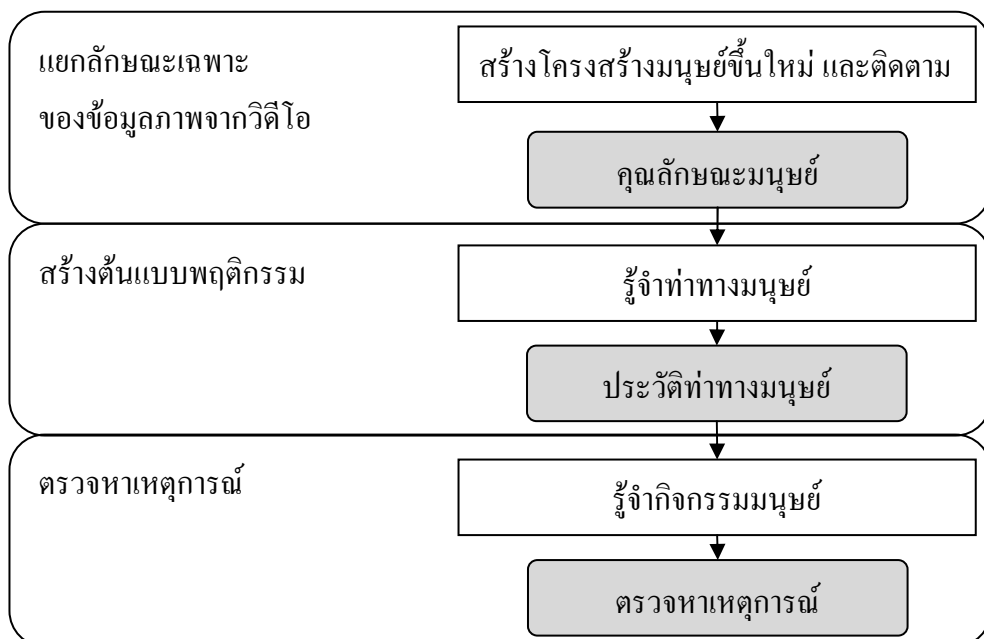
#### 2.4. สรุป

เนื้อหาภายในบทนี้กล่าวถึงกระบวนการด้านการประมวลผลภาพด้วยคอมพิวเตอร์ที่เกี่ยวข้องกับงานวิจัยนี้ได้แก่ ระบบสี HSV การลดสัญญาณรบกวนโดยกระบวนการเชิงสัญญาณวิทยา การดำเนินการเปิด การดำเนินการปิด และกล่าวถึงกระบวนการพื้นฐานเกี่ยวกับการตรวจจับการเคลื่อนไหวและการติดตามวัตถุเคลื่อนไหวซึ่งจะถูกนำไปใช้ในกระบวนการรู้จำท่าทางมนุษย์ต่อไป

### บทที่ 3

#### การออกแบบและพัฒนาระบบ

งานวิจัยนี้ได้ประยุกต์ใช้กระบวนการติดตามวัตถุเคลื่อนไหวจากภาพในกล้องวิดีโอเพื่อพัฒนาเทคนิคการจำแนกพฤติกรรมของมนุษย์อันนำไปสู่เทคนิคการตรวจจับพฤติกรรมผิดปกติ โดยมุ่งเน้นไปที่การตรวจจับพฤติกรรมจากบุคคล กระบวนการโดยรวมแสดงในภาพประกอบ 3-1 เริ่มต้นจากการจับคู่วัตถุเคลื่อนไหวที่คาดว่าจะเป็มนุษย์จากภาพวิดีโอกับแบบจำลองมนุษย์อย่างง่าย โดยมนุษย์ภายในงานวิจัยนี้ได้ถูกอธิบายอยู่รูปแบบจำลองโครงสร้างมนุษย์อย่างง่ายโดยทำการลดความซับซ้อนขององค์ประกอบในร่างกายมนุษย์ให้เหลือเพียง 3 องค์ประกอบ จากนั้นจึงติดตามการเคลื่อนไหวโดยมองว่าแต่ละองค์ประกอบเป็นหนึ่งวัตถุและจะติดตามแต่ละวัตถุแยกจากกันอย่างเป็นอิสระไม่ขึ้นต่อกัน ผลจากการติดตามการเคลื่อนที่มนุษย์จะทำให้ได้คุณลักษณะของมนุษย์ ซึ่งจะถูกนำไปใช้ในกระบวนการรู้จำท่าทางมนุษย์ ผลจากกระบวนการนี้จะทำให้ทราบว่ามนุษย์ได้กระทำท่าทางได้บ้าง และในการบวนการสุดท้าย คือ กระบวนการรู้จำกิจกรรม โดยกระบวนการนี้จะใช้การผสมผสานกันของท่าทางและการเคลื่อนที่ร่วมกับเทคนิคการลบพื้นหลังเพื่ออธิบายกิจกรรมที่มีลักษณะเป็นเหตุการณ์ผิดปกติ ซึ่งผลลัพธ์สุดท้ายจึงสามารถตรวจพบเหตุการณ์ผิดปกติที่เกิดจากกิจกรรมของมนุษย์ได้



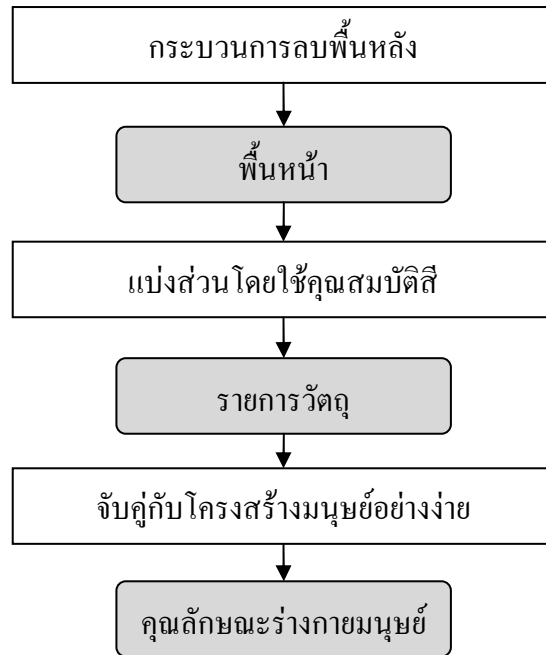
ภาพประกอบ 3-1 กระบวนการโดยรวมเทคนิคการตรวจจับเหตุการณ์ผิดปกติ

### 3.1. กระบวนการสร้างโครงสร้างมนุษย์ขึ้นใหม่ และติดตามการเคลื่อนไหว(Human Reconstruction and tracking)

พฤติกรรมมนุษย์สามารถถูกจำแนกได้โดยการสังเกตการเปลี่ยนทางโครงสร้างร่างกาย ร่วมกับข้อมูลการเปลี่ยนตำแหน่งซึ่งสามารถสังเกตเห็นได้จากภาพวิดีโอ ร่างกายของมนุษย์นั้น สามารถถูกจัดให้อยู่ในรูปแบบโครงสร้างอย่างง่ายได้ลดองค์ประกอบที่ซับซ้อนออกไปให้เหลือเพียงองค์ประกอบหลัก โดยภายในงานวิจัยนี้ได้ทำการลดรูปความซับซ้อนให้เหลือเพียงองค์ประกอบหลัก 3 ส่วน ได้แก่ ศีรษะ ลำตัวและขา แสดงในภาพประกอบ 3-5 ซึ่งจากผลการทดลองพบว่า การพิจารณาความสัมพันธ์เพียง 3 องค์ประกอบหลักนี้ก็เพียงพอที่จะระบุพฤติกรรมทั่วไปของมนุษย์ได้ถูกต้อง โดยกระบวนการสร้างโครงสร้างมนุษย์ขึ้นใหม่ประกอบด้วย 2 ขั้นตอน ดังนี้

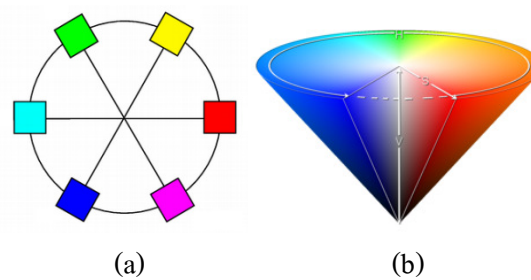
#### 3.1.1. การจับคู่แบบจำลองมนุษย์ (Human Model Matching)

กระบวนการโดยรวมของการจับคู่ภาพโครงสร้างวัตถุกับโครงสร้างมนุษย์อย่างง่าย แสดงในภาพประกอบ 3-2 เริ่มจากใช้บริเวณวัตถุเคลื่อนไหว (ภาพพื้นหน้า) ที่ได้จากกระบวนการลบภาพพื้นหลัง ซึ่งบริเวณดังกล่าวจะถูกนำไปแยกพื้นที่ออกเป็นส่วน ๆ โดยใช้คุณลักษณะสีที่มีความใกล้เคียงกันและอยู่ติดกันเป็นกลุ่มก่อนจะถูกจัดเป็นพื้นที่วัตถุขึ้นเดียวกัน หลังจากแยกพื้นที่ออกเป็นส่วน ๆ แล้วแต่ละส่วนจะถูกนำไปจับคู่กับโครงสร้างอย่างง่ายของร่างกายมนุษย์ ภายใต้การทดลองปัจจุบันใช้โครงสร้างมนุษย์ทำขึ้นเป็นเงื่อนไขเบื้องต้นในการจับคู่ ถ้าหากผลการจับคู่เข้าเงื่อนไขเหมาะสม จะทำให้ได้ตำแหน่งอ้างอิงโครงสร้างอย่างง่ายของร่างกาย และถูกใช้ในกระบวนการติดตามต่อไป

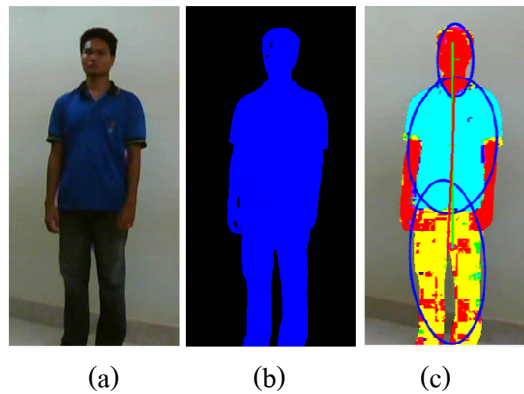


ภาพประกอบ 3-2 กระบวนการโดยรวมของการจับคู่ภาพโครงสร้างวัตถุกับแบบจำลองมนุษย์อย่างง่าย

เนื่องจาก ร่างกายมนุษย์ในภาพวิดีโอโดยส่วนใหญ่จะประกอบด้วย 3 กลุ่มสีใหญ่ ๆ ได้แก่ กลุ่มสีผิวหนัง กลุ่มสีเสื้อและกลุ่มสีกางเกง ซึ่งจากสมมุติฐานนี้สามารถใช้เพื่อแยกส่วนหลักร่างกายได้ โดยภายในงานวิจัยนี้ได้เลือกใช้ระบบสี HSV และได้แบ่งกลุ่มสีออกเป็น 6 กลุ่มสีดังนี้

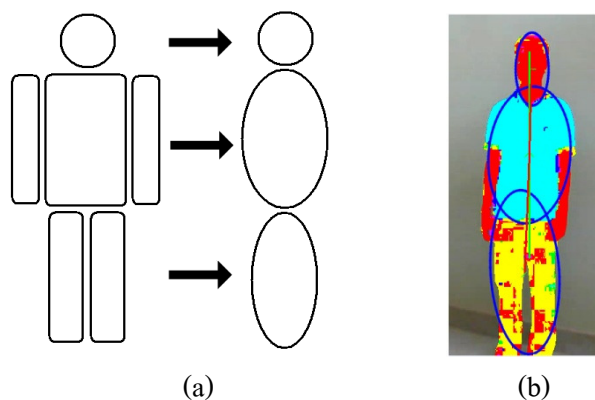


ภาพประกอบ 3-3 กลุ่มสีหลัก 6 สี (a) Red, Yellow, Green, Cyan, Blue และ Magenta (b) HSV color model



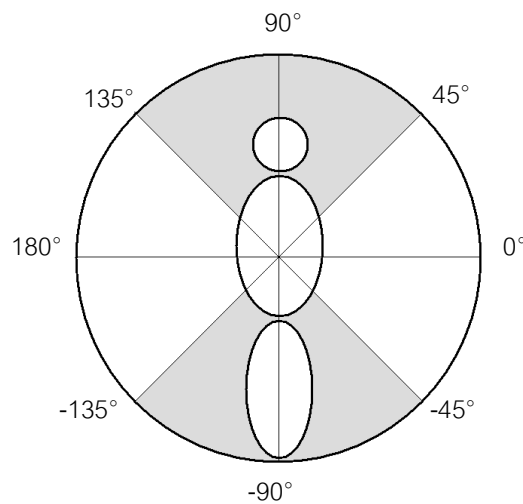
ภาพประกอบ 3-4 ตัวอย่างกระบวนการลบพื้นหลังและแบ่งพื้นที่วัตถุ (a) ภาพต้นฉบับ (b) ผลจากกระบวนการลบพื้นหลัง (c) ผลจากกระบวนการแบ่งพื้นที่วัตถุที่ใช้คุณสมบัติสี

ผลจากกระบวนการลบพื้นหลังจะทำให้ได้บริเวณวัตถุเคลื่อนไหวซึ่งจะถูกนำไปแบ่งออกเป็นส่วน ๆ โดยอาศัยคุณลักษณะสี ดังนั้นจากภาพประกอบที่ 3-4 (c) จะเห็นได้ว่ามนุษย์ 1 คนจะประกอบด้วยวัตถุหลายวัตถุ ซึ่งจากการคำนวณฮิสโทแกรมของสี จากบริเวณพื้นหน้า เราจะพบว่าองค์ประกอบหลักได้แก่ สีรยะ ลำตัวและขา เป็นกลุ่มสีที่มีปริมาณมากที่สุด 3 อันดับและกลุ่มสี 3 อันดับนี้จะถูกใช้ในการคำนวณหาขอบ [3] ผลของการหาขอบจะได้รายการของขอบทั้งหมดที่มีในพื้นที่ โดยพื้นที่ที่มีขนาดเล็กมาก ๆ จะถูกละเว้นไม่นำมาพิจารณา พื้นที่หลังการหาขอบในแต่ละกลุ่มสีที่ผ่านการตรวจสอบขนาดจะถูกใช้ในกระบวนการจับคู่สำหรับแบบจำลองมนุษย์อย่างง่ายที่แสดงในภาพประกอบ 3-5



ภาพประกอบ 3-5 แบบจำลองมนุษย์อย่างง่าย

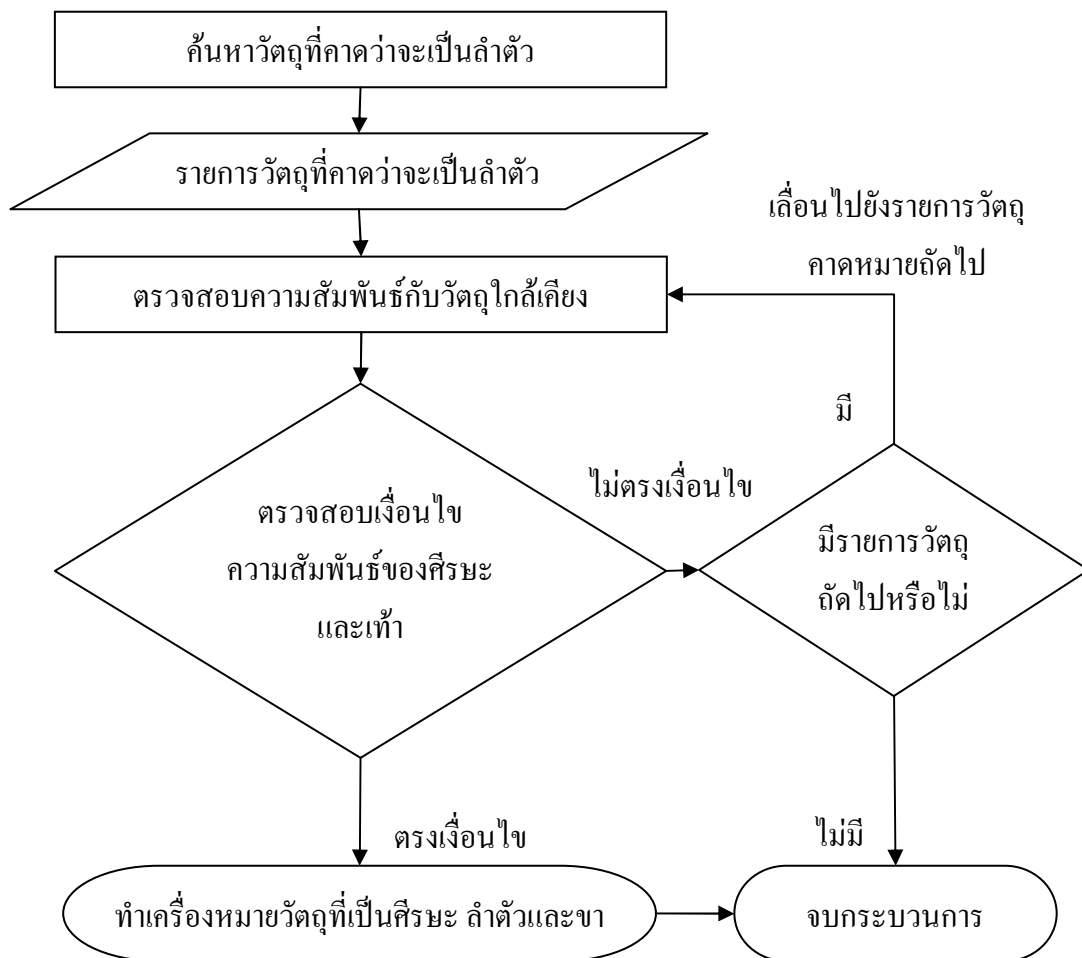
ภายในการทดลองปัจจุบันใช้มนุษย์ทำขึ้นเป็นกรณีเริ่มต้นการติดตามเนื่องจากทำขึ้นเป็นท่าที่ทำให้คุณลักษณะเชิงโครงสร้างอย่างชัดเจน โดยทำขึ้นจะทำให้ศีรษะ ลำตัว และขา อยู่ในแกนแนวตั้ง ถ้าเราใช้ลำตัวเป็นศูนย์กลางของร่างกายทั้งหมด ศีรษะจะตั้งอยู่ในตำแหน่ง  $45^\circ$  ถึง  $135^\circ$  ในด้านบน และ ขา จะอยู่ในตำแหน่ง  $-45^\circ$  ถึง  $-135^\circ$  ในด้านล่างดังแสดงในภาพประกอบ 3-6 และเมื่อพิจารณาถึงอัตราส่วนขนาดของศีรษะต่อลำตัวจะพบว่า ไม่เกิน 1 และ อัตราส่วนขาต่อลำตัว ค่าจะอยู่ระหว่าง 0.5 ถึง 1.5 ดังนั้นจากคุณสมบัติดังกล่าวเราจึงสามารถใช้เป็นเงื่อนไขในการตรวจสอบตำแหน่งของ ศีรษะ ลำตัว และขา ได้



ภาพประกอบ 3-6 การแบ่งส่วนแบบจำลองมนุษย์อย่างง่าย

ลำตัวเป็นส่วนที่มีขนาดใหญ่ที่สุดในร่างกายและอยู่บริเวณส่วนกลางขององค์ประกอบทั้งหมด ดังนั้นเราจึงตั้งสมมุติฐานได้ว่า จะต้องหาบริเวณที่คาดว่าจะเป็นลำตัวก่อนซึ่งจะเลือกใช้บริเวณที่มีขนาดใหญ่ที่สุด 2 อันดับแรก ทำให้มีอย่างน้อย 2 พื้นที่ต่อ 1 บุคคล คือ บริเวณกลุ่มของขาและลำตัว หลังจากนั้นจึงใช้บริเวณที่คาดว่าจะเป็นลำตัวเป็นตำแหน่งอ้างอิงเพื่อหาความสัมพันธ์กับกลุ่มพื้นที่โดยรอบ ซึ่งมีกระบวนการทำงานดังภาพประกอบ 3-7 ถ้าหากผลการหาความสัมพันธ์ตรงตามเงื่อนไข พื้นที่บริเวณ ศีรษะ ลำตัวและขาจะถูกทำเครื่องหมายไว้เพื่อใช้ในกระบวนการติดตามต่อไป





ภาพประกอบ 3-7 กระบวนการจับคู่แบบจำลองมนุษย์อย่างง่าย

จากภาพประกอบ 3-7 ขั้นตอนแรกเริ่มต้นโดยการพิจารณาองค์ประกอบที่มีขนาดใหญ่ที่สุด 2 ส่วนจากกลุ่มวัตถุที่ถูกตรวจพบเป็นวัตถุใหม่ โดยวัตถุนี้จะถูกตั้งสมมุติฐานว่าเป็นส่วนล้าตัวและจะใช้พื้นที่ส่วนนี้เป็นจุดอ้างอิงของล้าตัวเพื่อตรวจสอบความสัมพันธ์กับองค์ประกอบที่อยู่ด้านบนและด้านล่าง ถ้าองค์ประกอบทั้งด้านบนและด้านล่างมีคุณสมบัติทางตำแหน่งและอัตราส่วนขนาดตรงกับคุณสมบัติของมนุษย์ในทางขึ้นก็จะทำการทำเครื่องหมายขององค์ประกอบทั้งส่วนของสิริษะ ล้าตัวและขาไว้เพื่อจะถูกนำไปใช้ในกระบวนการติดตามต่อไป แต่ถ้าองค์ประกอบด้านบนและด้านล่างไม่ตรงตามคุณสมบัติของมนุษย์ทำขึ้นก็จะเคลื่อนไปพิจารณาองค์ประกอบที่คาดว่าจะเป็ลล้าตัวในล้าดับถัด

### 3.1.2. กระบวนการติดตามวัตถุและคำนวณพารามิเตอร์

จากแบบจำลองมนุษย์อย่างง่ายดังภาพประกอบ 3-8 ณ เวลา  $t_i$  ขณะยังไม่มีการเคลื่อนไหว จะมีพารามิเตอร์ที่สะท้อนถึงความสัมพันธ์ระหว่างแต่ละองค์ประกอบซึ่งเป็นคุณสมบัติภายในของ โครงสร้างอย่างง่าย ประกอบด้วย 2 เวกเตอร์ คือ เวกเตอร์  $\vec{v}_h$  ที่อธิบายถึงความสัมพันธ์ระหว่าง ลำตัวและศีรษะ และเวกเตอร์  $\vec{v}_l$  อธิบายถึงความสัมพันธ์ระหว่างลำตัวและขา โดย

เวกเตอร์  $\vec{v}_h$  สามารถแยกได้ออกเป็น 2 องค์ประกอบย่อย คือ

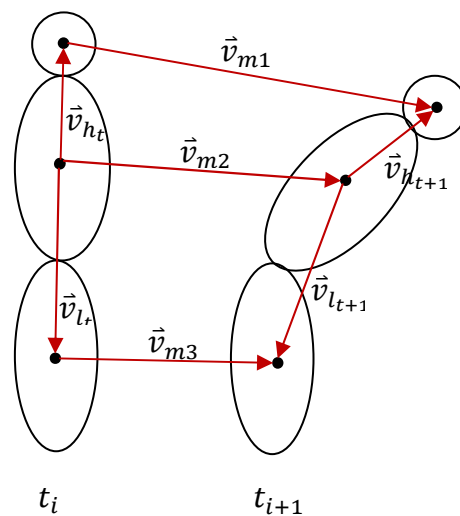
$\vec{v}_h^r$  คือ ระยะห่างระหว่างจุดศูนย์กลางลำตัวถึงจุดศูนย์กลางศีรษะ

$\vec{v}_h^o$  คือ มุมของเส้นตรงที่ลากจากจุดศูนย์กลางลำตัวถึงจุดศูนย์กลางศีรษะ

เวกเตอร์  $\vec{v}_l$  สามารถแยกได้ออกเป็น 2 องค์ประกอบย่อย คือ

$\vec{v}_l^r$  คือ ระยะห่างระหว่างจุดศูนย์กลางลำตัวถึงจุดศูนย์กลางขา

$\vec{v}_l^o$  คือ มุมของเส้นตรงที่ลากจากจุดศูนย์กลางลำตัวถึงจุดศูนย์กลางขา



ภาพประกอบ 3-8 ดัชนีแบบมนุษย์อย่างง่ายและพารามิเตอร์ที่สนใจ ณ ขณะเวลา  $t_i$  และ  $t_{i+1}$  พร้อมทั้ง พารามิเตอร์ทั้งหมด

เมื่อวัตถุเริ่มมีการเคลื่อนที่ พื้นที่ที่ถูกทำเครื่องหมายว่าเป็นศีรษะ ลำตัวและขาจะถูกติดตาม โดยใช้คุณลักษณะสีของพื้นที่บริเวณนั้น ๆ โดยใช้ขั้นตอนวิธีการย้ายเข้าสู่ค่ากลางโดยการปรับตัวอย่างต่อเนื่อง (CAMSHIFT) แต่ละองค์ประกอบของโครงสร้างอย่างง่ายจะถูกติดตามแบบเป็นอิสระต่อกันดังแสดงในภาพประกอบ 3-10 ซึ่งผลการติดตามในแต่ละเฟรมจะทำให้ได้ตำแหน่งใหม่ของวัตถุ ถ้าหากวัตถุยังคงอยู่ภายในพื้นที่สนใจ จุดศูนย์กลางตำแหน่งใหม่และตำแหน่งเก่าจะถูกใช้ในการคำนวณหาค่าพารามิเตอร์ต่าง ๆ ต่อไป แต่ถ้าหากวัตถุออกนอกบริเวณที่สนใจจะก็หยุดการ

ติดตามบุคคลนี้ ในกรณีที่วัตถุมีการเปลี่ยนตำแหน่งจะทำให้ได้พารามิเตอร์จากการเคลื่อนไหวเพิ่มขึ้นอีก 3 เวกเตอร์ คือ เวกเตอร์  $\vec{v}_{m1}$   $\vec{v}_{m2}$  และ  $\vec{v}_{m3}$  ซึ่งจะอธิบายถึงการเคลื่อนที่ของศีรษะ ลำตัวและขาตามลำดับ จากเวลา  $t_i$  ถึง เวลา  $t_{i+1}$  ดังแสดงในภาพประกอบ 3-8 โดย

เวกเตอร์  $\vec{v}_{m1}$  สามารถแยกได้ออกเป็น 2 องค์ประกอบย่อย คือ

$\vec{v}_{m1}^s$  คือ ระยะห่างระหว่างจุดศูนย์กลางศีรษะ จากเวลา  $t_i$  ถึง เวลา  $t_{i+1}$

$\vec{v}_{m1}^\alpha$  คือ มุมของเส้นตรงที่ลากจากจุดศูนย์กลางศีรษะ จากเวลา  $t_i$  ถึง เวลา  $t_{i+1}$

เวกเตอร์  $\vec{v}_{m2}$  สามารถแยกได้ออกเป็น 2 องค์ประกอบย่อย คือ

$\vec{v}_{m2}^s$  คือ ระยะห่างระหว่างจุดศูนย์กลางลำตัว จากเวลา  $t_i$  ถึง เวลา  $t_{i+1}$

$\vec{v}_{m2}^\alpha$  คือ มุมของเส้นตรงที่ลากจากจุดศูนย์กลางลำตัว จากเวลา  $t_i$  ถึง เวลา  $t_{i+1}$

เวกเตอร์  $\vec{v}_{m3}$  สามารถแยกได้ออกเป็น 2 องค์ประกอบย่อย คือ

$\vec{v}_{m3}^s$  คือ ระยะห่างระหว่างจุดศูนย์กลางขา จากเวลา  $t_i$  ถึง เวลา  $t_{i+1}$

$\vec{v}_{m3}^\alpha$  คือ มุมของเส้นตรงที่ลากจากจุดศูนย์กลางขา จากเวลา  $t_i$  ถึง เวลา  $t_{i+1}$

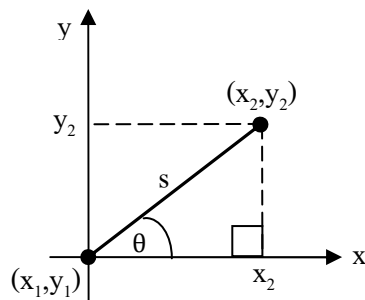
การคำนวณค่าพารามิเตอร์ต่าง ๆ มี ดังนี้

พารามิเตอร์ที่อธิบายถึงระยะห่างระหว่างจุด 2 จุด ได้แก่  $\vec{v}_h^s, \vec{v}_l^s, \vec{v}_{m1}^s, \vec{v}_{m2}^s$  และ  $\vec{v}_{m3}^s$

พารามิเตอร์ที่อธิบายถึงมุมเส้นตรงที่ลากระหว่างจุด 2 จุด ได้แก่  $\vec{v}_h^\alpha, \vec{v}_l^\alpha, \vec{v}_{m1}^\alpha, \vec{v}_{m2}^\alpha$  และ

$\vec{v}_{m3}^\alpha$

สามารถคำนวณได้โดยอาศัยสมการดังนี้



ภาพประกอบ 3-9 ตัวอย่างประกอบการคำนวณค่าพารามิเตอร์

กำหนดให้

พิกัด  $(x_1, y_1)$  เป็นจุดเริ่มต้น

พิกัด  $(x_2, y_2)$  เป็นจุดสิ้นสุด ดังแสดงในภาพประกอบ 3-9

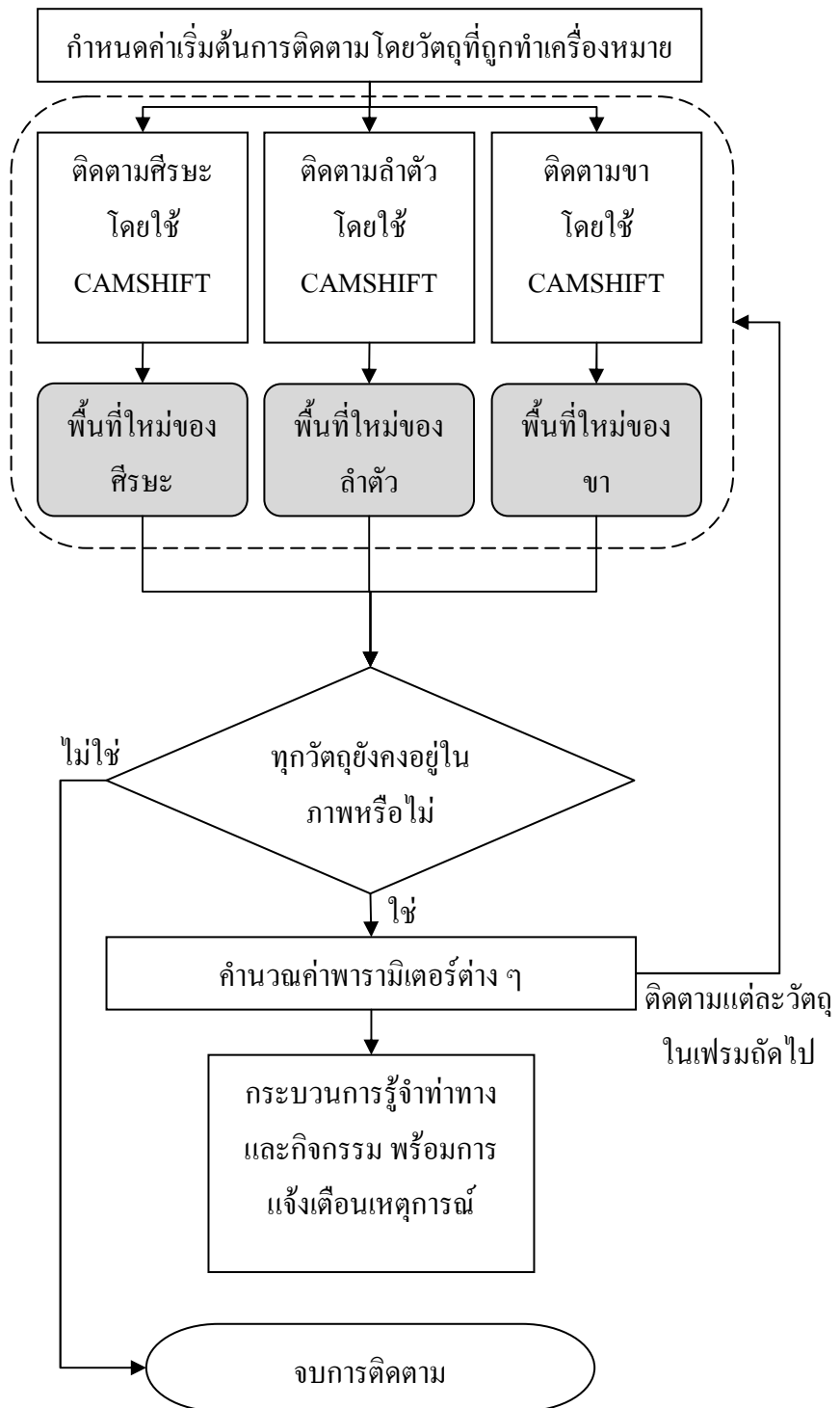
การคำนวณระยะทางระหว่างจุด 2 จุด (s) สามารถประยุกต์ใช้ทฤษฎีบทพีทาโกรัสได้ดังนี้

$$s = \sqrt{(y_2 - y_1)^2 + (x_2 - x_1)^2} \quad (3.1)$$

การคำนวณมุม  $\theta$  สามารถประยุกต์ใช้ตรีโกณมิติได้ดังนี้

$$\theta = \tan^{-1} \frac{(y_2 - y_1)}{(x_2 - x_1)} \quad (3.2)$$

จากการประยุกต์ใช้ทฤษฎีบทพีทาโกรัสและตรีโกณมิติสามารถจะคำนวณค่าพารามิเตอร์ทั้งหมดได้ ซึ่งค่าเหล่านี้จะถูกดึงไปใช้ในกระบวนการรู้ตำแหน่งและกิจกรรมต่อไป ดังแสดงในภาพประกอบ 3-10



ภาพประกอบ 3-10 กระบวนการติดตามวัตถุและคำนวณพารามิเตอร์

### 3.2. นิยามและการรู้จำท่าทางมนุษย์ (Human Action Definition and Recognition)

จากแบบจำลองมนุษย์อย่างง่ายสำหรับการรู้จำท่าทางแสดงในภาพประกอบ 3-8 ประกอบด้วยพารามิเตอร์ที่ใช้พิจารณา 2 กลุ่ม ได้แก่ พารามิเตอร์ภายใน และ พารามิเตอร์ภายนอก พารามิเตอร์ภายใน ( $\dot{v}_h, \dot{v}_l$ ) แสดงถึงคุณลักษณะของโครงสร้างมนุษย์ที่อธิบายถึงระยะห่าง ( $\dot{v}_h^s, \dot{v}_l^s$ ) และทิศทาง ( $\dot{v}_h^o, \dot{v}_l^o$ ) จากองค์ประกอบส่วนลำตัวไปยังศีรษะและขา พารามิเตอร์ภายนอก ( $\dot{v}_{m1}, \dot{v}_{m2}, \dot{v}_{m3}$ ) จะสัมพันธ์กับคุณสมบัติการเคลื่อนไหวจากเฟรมหนึ่งไปสู่เฟรมอื่น ๆ โดยอธิบายถึงลักษณะของความเร็ว ( $\dot{v}_{m1}^s, \dot{v}_{m2}^s, \dot{v}_{m3}^s$ ) และทิศทางการเคลื่อนไหว ( $\dot{v}_{m1}^o, \dot{v}_{m2}^o, \dot{v}_{m3}^o$ ) ขององค์ประกอบศีรษะ ลำตัว และขา โดยพารามิเตอร์จะสัมพันธ์กับศูนย์กลางของแต่ละองค์ประกอบซึ่งการประมาณค่าโดยใช้ศูนย์กลางเช่นนี้จะช่วยลดสัญญาณรบกวนจากกระบวนการย่อย ๆ เช่น การตรวจหาการเคลื่อนไหว กระบวนการแยกกลุ่มสี กระบวนการติดตามการเคลื่อนที่

พฤติกรรมทั่วไปของมนุษย์โดยส่วนใหญ่สามารถแยกเป็นท่าทางพื้นฐานได้ 5 ประเภท ได้แก่ การยืน การก้ม การนั่ง การเดิน และการนอนราบ ดังนั้นพฤติกรรมทั่วไปจึงสามารถอธิบายได้โดยอาศัยการผสมผสานของท่าทางพื้นฐานเหล่านี้ประกอบกับพารามิเตอร์ที่แสดงถึงการเปลี่ยนแปลงสถานะของแต่ละท่าทาง ซึ่งพารามิเตอร์โดยส่วนใหญ่จะขึ้นอยู่กับเวกเตอร์ของการเคลื่อนไหวแสดงในภาพประกอบ 3-8 การใช้โครงสร้างมนุษย์อย่างง่ายเพื่อการรู้จำท่าทางจึงต้องทำการสร้างคุณลักษณะของแต่ละท่าทางจากพารามิเตอร์ที่มีอยู่โดยมุ่งเน้นไปที่คุณลักษณะเฉพาะ 8 ตัว ได้แก่

$$\dot{v}_h^o/\dot{v}_l^o, \dot{v}_h^s/\dot{v}_l^s, \frac{\partial \dot{v}_{m1}^s}{\partial t}, \frac{\partial \dot{v}_{m2}^s}{\partial t}, \frac{\partial \dot{v}_{m3}^s}{\partial t}, \frac{\partial^2 \dot{v}_{m1}^o}{\partial t^2}, \frac{\partial^2 \dot{v}_{m2}^o}{\partial t^2}, \frac{\partial^2 \dot{v}_{m3}^o}{\partial t^2}$$

$\dot{v}_h^o/\dot{v}_l^o$  คือ อัตราส่วนมุมของศีรษะและเท้าที่กระทำกับแกนแนวตั้งของลำตัว

$\dot{v}_h^s/\dot{v}_l^s$  คือ อัตราส่วนระยะห่าง ของศีรษะและเท้าที่วัดจากจุดศูนย์กลางของลำตัว

$\frac{\partial \dot{v}_{m1}^s}{\partial t}, \frac{\partial \dot{v}_{m2}^s}{\partial t}, \frac{\partial \dot{v}_{m3}^s}{\partial t}$  คือ ความเร็วในการเปลี่ยนตำแหน่งในแต่ละส่วนของมนุษย์ ได้แก่

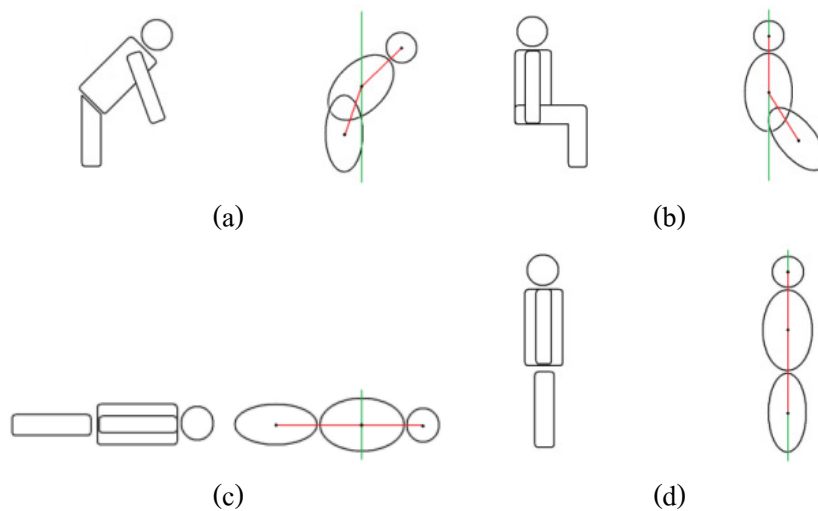
ศีรษะ ลำตัว และขา ตามลำดับ

$\frac{\partial^2 \dot{v}_{m1}^o}{\partial t^2}, \frac{\partial^2 \dot{v}_{m2}^o}{\partial t^2}, \frac{\partial^2 \dot{v}_{m3}^o}{\partial t^2}$  คือ ความเร่งเชิงมุมของแต่ละส่วนของมนุษย์ ได้แก่ ศีรษะ ลำตัว และขา

ตามลำดับ

เมื่อพิจารณาท่าทางพื้นฐาน 5 ท่า ได้แก่ การยืน การก้ม การนั่ง การเดิน และการนอนราบ เทียบเคียงกับจากแบบจำลองมนุษย์อย่างง่ายจะพบว่าแต่ละท่าทางต่างมีลักษณะที่เป็นเอกลักษณ์ และสังเกตเห็นความแตกต่างกันอย่างชัดเจน ซึ่งความแตกต่างนี้สามารถถูกอธิบายในรูปการผสมผสานคุณลักษณะต่าง ๆ ดังที่กล่าวมาข้างต้น ทำให้เราสามารถที่จะบ่งชี้ความแตกต่างของท่าทางได้อย่างถูกต้อง ภาพประกอบ 3-11 แสดงท่าทางพื้นฐานและพารามิเตอร์ภายใน เมื่อใช้เกณฑ์ด้านการเคลื่อนที่แบ่งประเภท เราสามารถแบ่งประเภทท่าทางออกเป็น 2 ประเภทได้แก่ ท่าทางสถิต

(static action) ซึ่งจะจัดอยู่ในกลุ่มแบบจำลองสถิตและท่าทางพลวัต (dynamic action) ซึ่งจะจัดอยู่ในกลุ่มแบบจำลองพลวัต

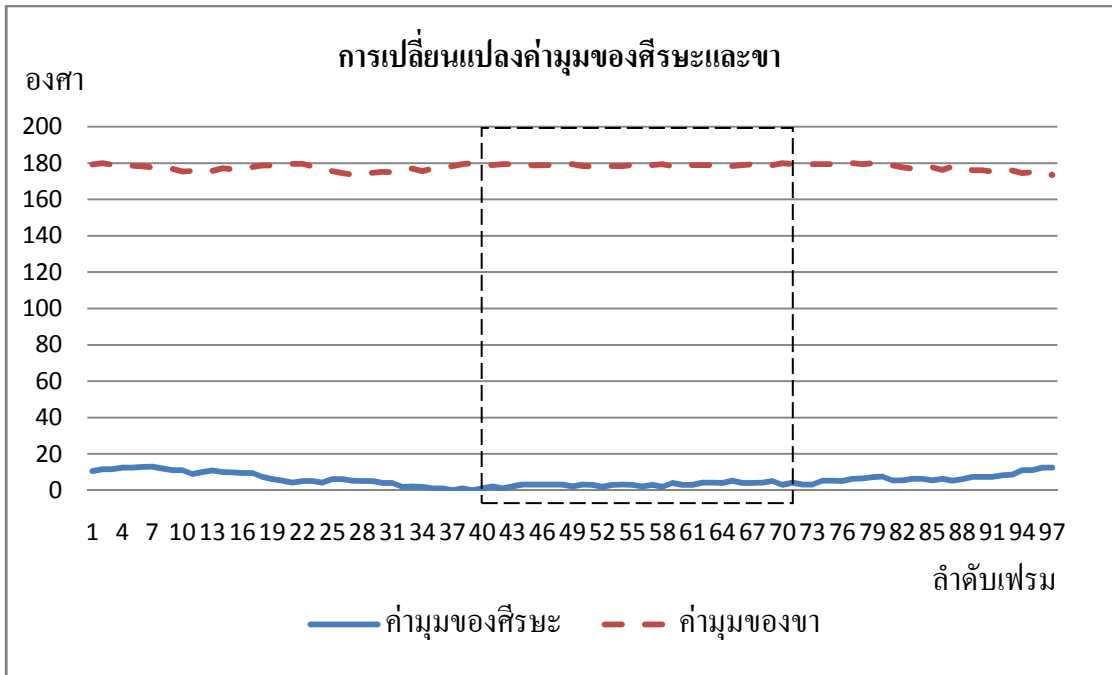


ภาพประกอบ 3-11 ท่าทางพื้นฐาน (a) ก้ม (b) นั่ง (c) นอนราบ (d) ยืน หรือ เดิน

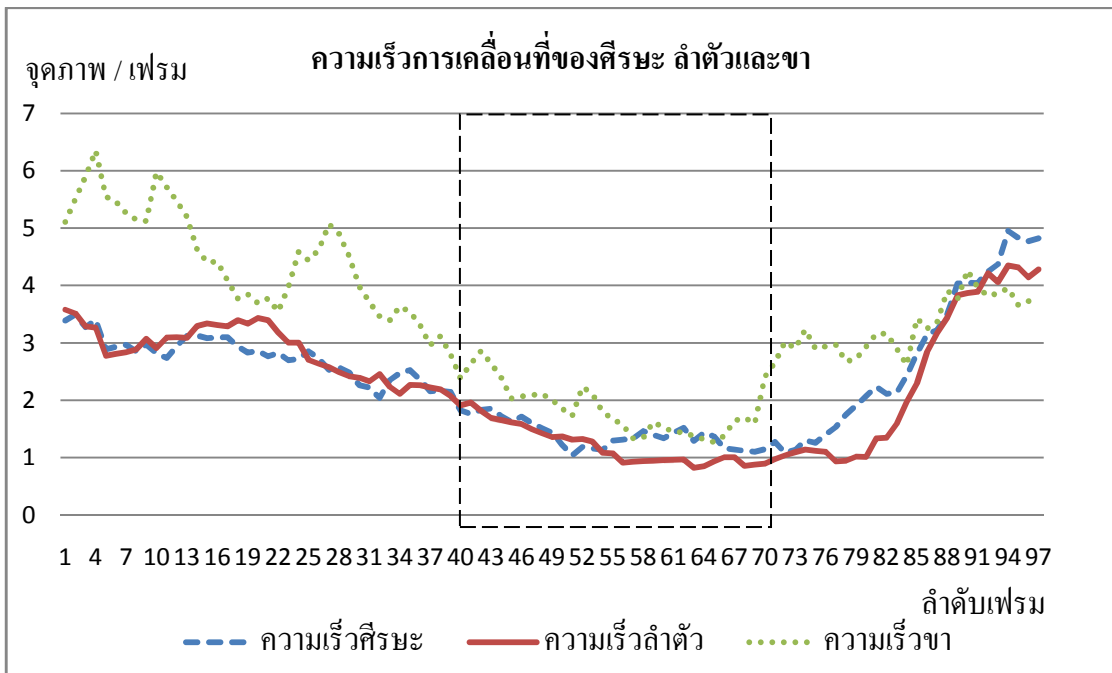
### 3.2.1. แบบจำลองสถิต (Static Model)

ท่าทางสถิต คือ ท่าทางที่มีอย่างน้อยหนึ่งในโครงสร้างอย่างง่ายของมนุษย์ไม่มีการเคลื่อนไหว ท่าทางสถิต ได้แก่ การยืน การก้ม การนั่ง จะสังเกตเห็นได้ว่าการกระทำท่าทางทั้ง 3 ท่านี้ องค์ประกอบส่วนขาแทบจะหยุดนิ่งอยู่กับที่และเห็นได้ว่าท่าทางเหล่านี้ประกอบขึ้นจากพารามิเตอร์ภายใน ได้แก่  $\psi_h$  และ  $\psi_l$  ที่มีลักษณะเป็นเอกลักษณ์และมีความแตกต่างกันอย่างชัดเจน ทำให้สามารถที่จะใช้เพื่อการแยกแยะท่าทางหรืออธิบายถึงท่าทางเหล่านี้ได้

คุณลักษณะการขึ้น

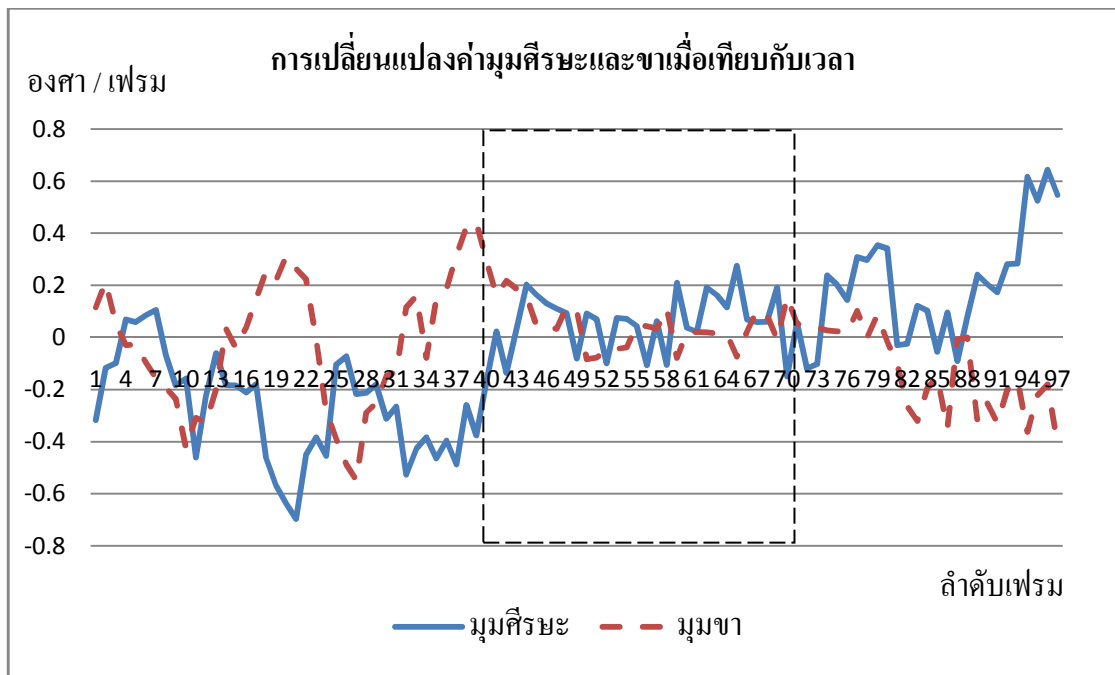


(a)



(b)





ภาพประกอบ 3-12 กราฟแสดงคุณลักษณะการขึ้น พื้นที่ภายในสี่เหลี่ยมในกราฟคือช่วงที่แสดงท่าทางการขึ้น (a) กราฟแสดง  $\psi_h^\theta$  และ  $\psi_l^\theta$  (b) กราฟแสดง  $\frac{\partial \psi_{m1}^s}{\partial t}$   $\frac{\partial \psi_{m2}^s}{\partial t}$  และ  $\frac{\partial \psi_{m3}^s}{\partial t}$  (c) กราฟแสดง  $\frac{\partial \psi_l^\theta}{\partial t}$  และ  $\frac{\partial \psi_h^\theta}{\partial t}$

โดยทั่วไป การขึ้นจะทำให้เวกเตอร์  $\psi_h$  และ  $\psi_l$  วางตัวอยู่ในแกนแนวตั้งซึ่งจะทำให้ค่า  $\psi_h^\theta$  มีค่าใกล้เคียง 0 องศา และ  $\psi_l^\theta$  มีค่าใกล้เคียง 180 องศา ดังแสดงในภาพประกอบ 3-12 (a) โดยไม่ขึ้นอยู่กับมุมกล้อง และความเร็วของทุกส่วนมีค่าเข้าใกล้ศูนย์อย่างแน่นอน ดังแสดงในภาพประกอบ 3-12 (b) ซึ่งเป็นไปตามสมการ (3.3) การเปลี่ยนแปลงค่ามุมเมื่อเทียบกับเวลาขณะที่มีการขึ้นอยู่จะพบว่าค่าไม่มีการเปลี่ยนแปลงมากนัก เพราะแทบจะไม่มีเคลื่อนไหวใดดังแสดงในภาพประกอบ 3-12 (c)

สำหรับในขณะที่มีเปลี่ยนท่าทาง (ภาพประกอบ 3-17) เช่น เปลี่ยนจากการขึ้นไปสู่การนั่งหรือการก้ม จะทำให้มุมของเวกเตอร์  $\psi_h$  และ  $\psi_l$  จะมีค่าเพิ่มขึ้นขณะเดียวกันค่าจะลดลงเมื่อมีการเปลี่ยนท่าทางกลับไปสู่ท่าขึ้น จากการเปลี่ยนแปลงดังกล่าวจะมีผลกับ  $\frac{\partial \psi_l^\theta}{\partial t}$  และ  $\frac{\partial \psi_h^\theta}{\partial t}$  ซึ่งจะเป็นไปตามสมการ (3.4) และ (3.5) แต่ในกรณีการเปลี่ยนสถานะจากการขึ้นเข้าสู่ท่าทางการเดินลักษณะโครงสร้างของร่างกายยังคงมีลักษณะเช่นเดียวกับการขึ้นนั่งแต่จะพบว่าการเคลื่อนที่ของทุก ๆ องค์ประกอบในร่างกายในลักษณะไปในทิศทางเดียวกันและความเร็วใกล้เคียงกัน ซึ่งเป็นอธิบายได้โดยสมการ (3.10) และการเปลี่ยนสถานะจากการขึ้นไปสู่การนอนราบโดยทันที โดยไม่ผ่าน

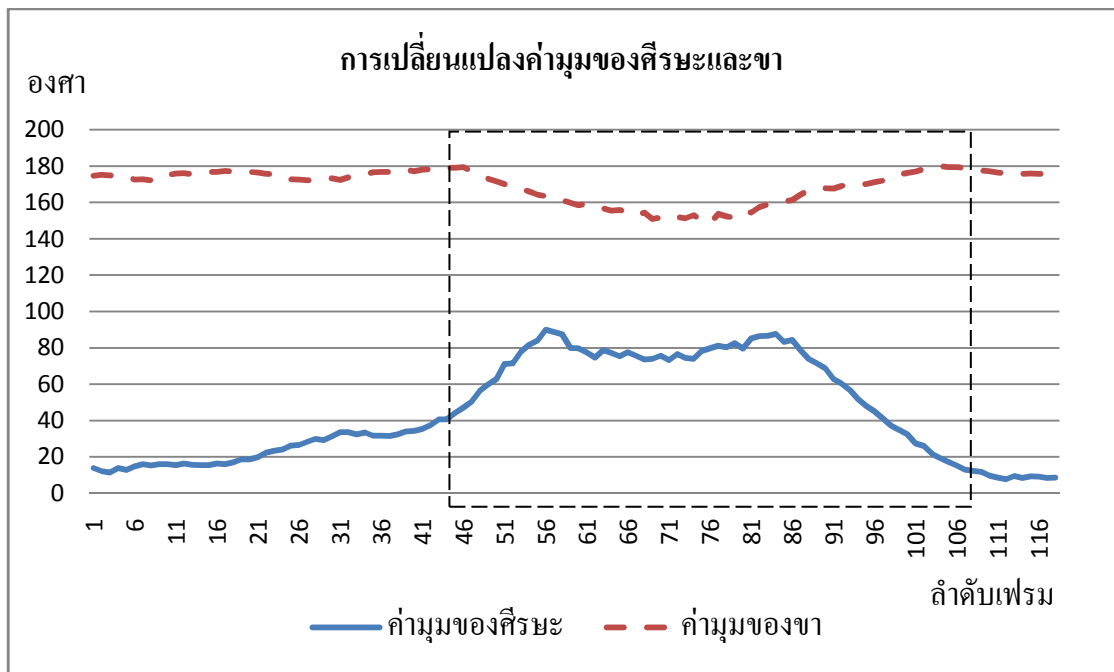
กระบวนการกัมหรือการนั่งจะทำให้เกิดการเคลื่อนที่ในลักษณะที่มีความเร่งสูงซึ่งจะให้ค่า  $\frac{\partial^2 \bar{v}_{m1}^\theta}{\partial t^2}$   $\frac{\partial^2 \bar{v}_{m2}^\theta}{\partial t^2}$  และ  $\frac{\partial^2 \bar{v}_{m3}^\theta}{\partial t^2}$  มีค่ามากกว่า 0 ซึ่งเป็นไปตามสมการ (3.11)

$$\bar{v}_h^\theta = 0 \mp 30, \bar{v}_l^\theta = 180 \mp 30, \frac{\partial \bar{v}_{m1}^s}{\partial t} = \frac{\partial \bar{v}_{m2}^s}{\partial t} = \frac{\partial \bar{v}_{m3}^s}{\partial t} = 0 \mp 3 \quad (3.3)$$

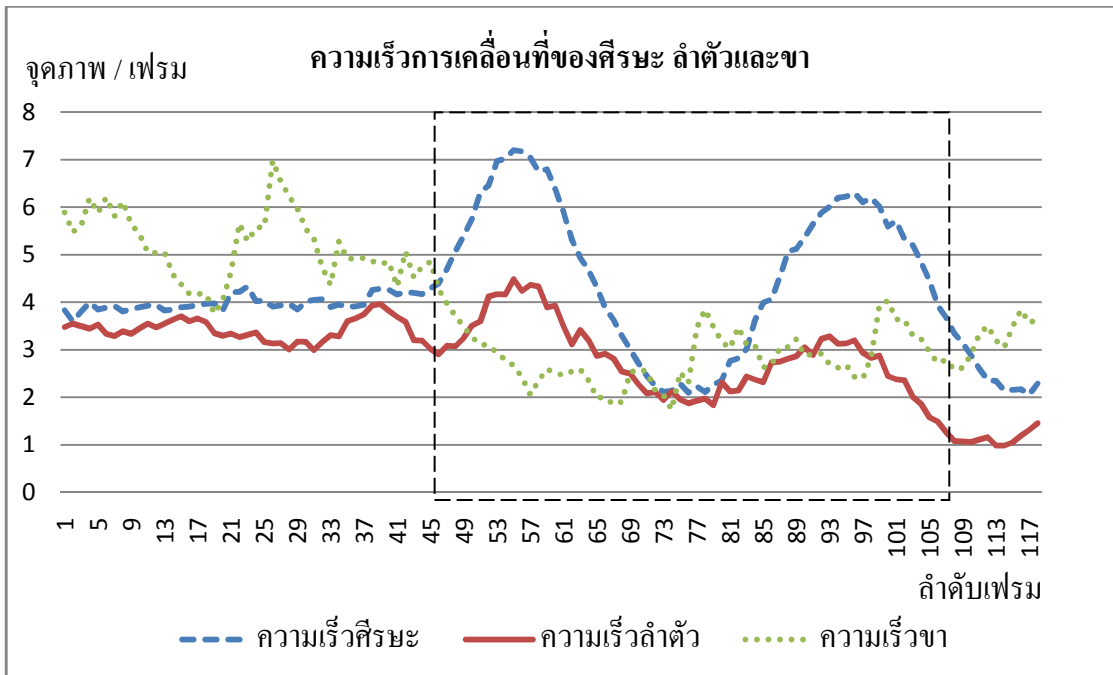
$$\frac{\partial \bar{v}_l^\theta}{\partial t} > 0 \quad (3.4), \quad \frac{\partial \bar{v}_h^\theta}{\partial t} > 0 \quad (3.5)$$

ดังนั้น ลักษณะเฉพาะของการขึ้น คือ สมการ (3.3)

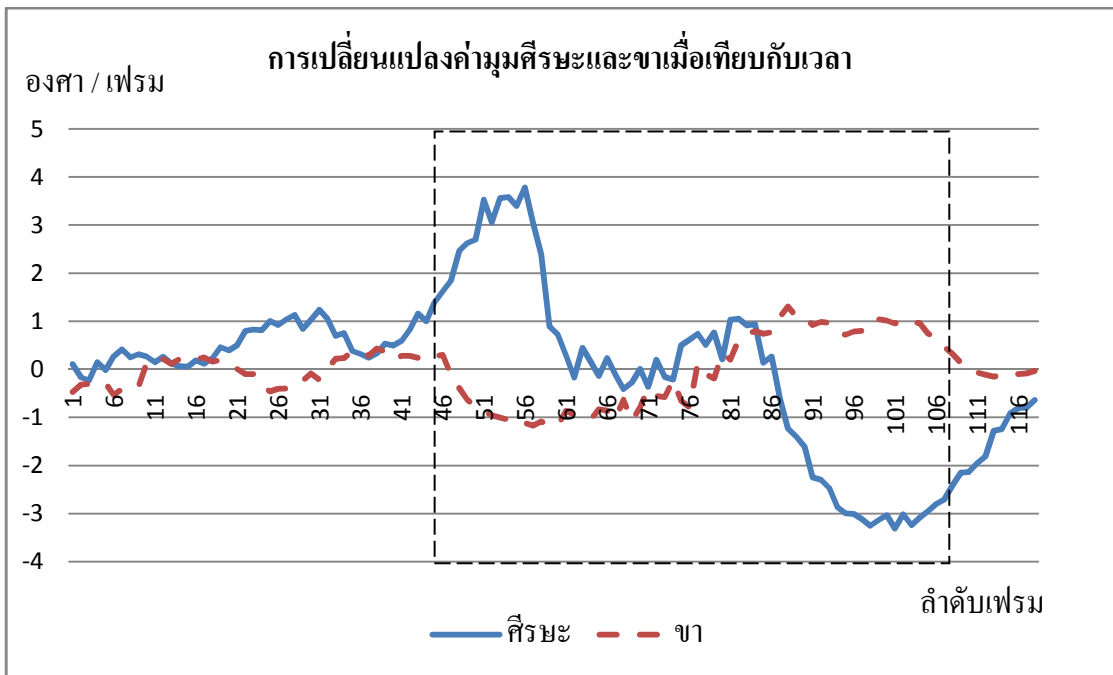
คุณลักษณะการกัม



(a)



(b)



(c)

ภาพประกอบ 3-13 กราฟแสดงคุณลักษณะการก้ม พื้นที่ภายในสี่เหลี่ยมในกราฟคือช่วงที่แสดงท่าทางการก้ม (a) กราฟแสดง  $\psi_h^\theta$  และ  $\psi_l^\theta$  (b) กราฟแสดง  $\frac{\partial \psi_{m1}^s}{\partial t}$   $\frac{\partial \psi_{m2}^s}{\partial t}$  และ  $\frac{\partial \psi_{m3}^s}{\partial t}$  (c) กราฟแสดง  $\frac{\partial \psi_l^\theta}{\partial t}$  และ  $\frac{\partial \psi_h^\theta}{\partial t}$

สำหรับกรณีการกัม ส่วนศีรษะจะกัมต่ำลงขณะที่ลำตัวและขายังคงค่อนข้างคงที่ทำให้ความเร็วในการเคลื่อนที่ในส่วนของลำตัวและขามีค่าต่ำดังแสดงในภาพประกอบ 3-13 (b) เมื่อพิจารณาถึงเวกเตอร์  $\vec{v}_l$  ยังคงค่อนข้างจะวางอยู่ในแกนแนวตั้งทำให้  $\vec{v}_l^\theta$  ยังคงมีค่าใกล้เคียง 180 องศาซึ่งเป็นไปตามสมการ (3.6) และศีรษะจะมีลักษณะกัมต่ำลงทำให้  $\vec{v}_h^\theta$  มีค่าสูงขึ้น ดังแสดงในภาพประกอบ 3-13 (a) เมื่อพิจารณาถึงการเปลี่ยนแปลงค่ามุมศีรษะและขาเมื่อเทียบกับเวลาค่าของ  $\frac{\partial \vec{v}_h^\theta}{\partial t}$  จะมีค่าเปลี่ยนแปลงไปมากขณะที่กำลังกัมศีรษะลงและยกศีรษะขึ้นแต่จะคงที่ในขณะที่ร่างกายหยุดนิ่งเป็นลักษณะโค้งตัวในขณะที่กัมอยู่ดังแสดงในภาพประกอบ 3-13 (c)

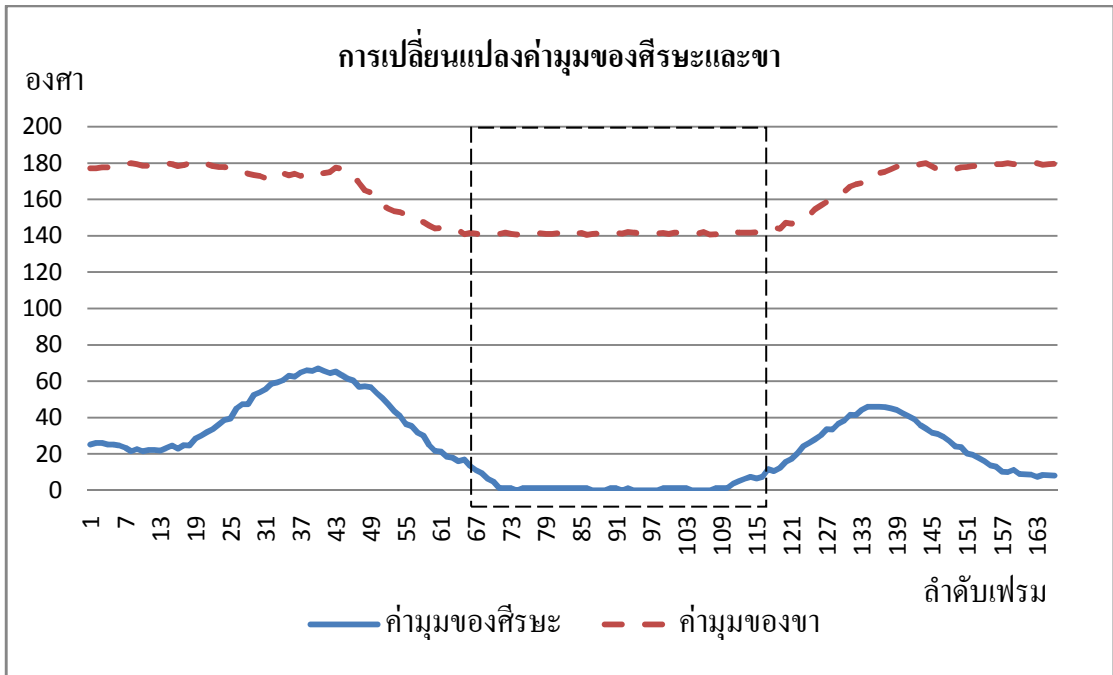
ขณะที่มีการเปลี่ยนทางทาง (ภาพประกอบ 3-17) จากการกัมไปสู่การยืน มุมของเวกเตอร์  $\vec{v}_h$  จะมีค่าลดลงดังแสดงในภาพประกอบ 3-13 (a) จึงทำให้  $\frac{\partial \vec{v}_h^\theta}{\partial t}$  จึงมีค่าน้อยกว่าศูนย์เป็นไปตาม (3.7) ในขณะที่การเปลี่ยนสถานะจากทำยืนเข้าสู่ทำนั่งจะเป็นไปในลักษณะตรงข้าม คือ มุมของศีรษะจะมีค่าเพิ่มขึ้นซึ่งเป็นไปตามสมการ (3.5)

$$\vec{v}_l^\theta = 180^\circ \mp 30^\circ, \frac{\partial \vec{v}_{m2}^\theta}{\partial t} = \frac{\partial \vec{v}_{m3}^\theta}{\partial t} = 0 \mp 3 \quad (3.6)$$

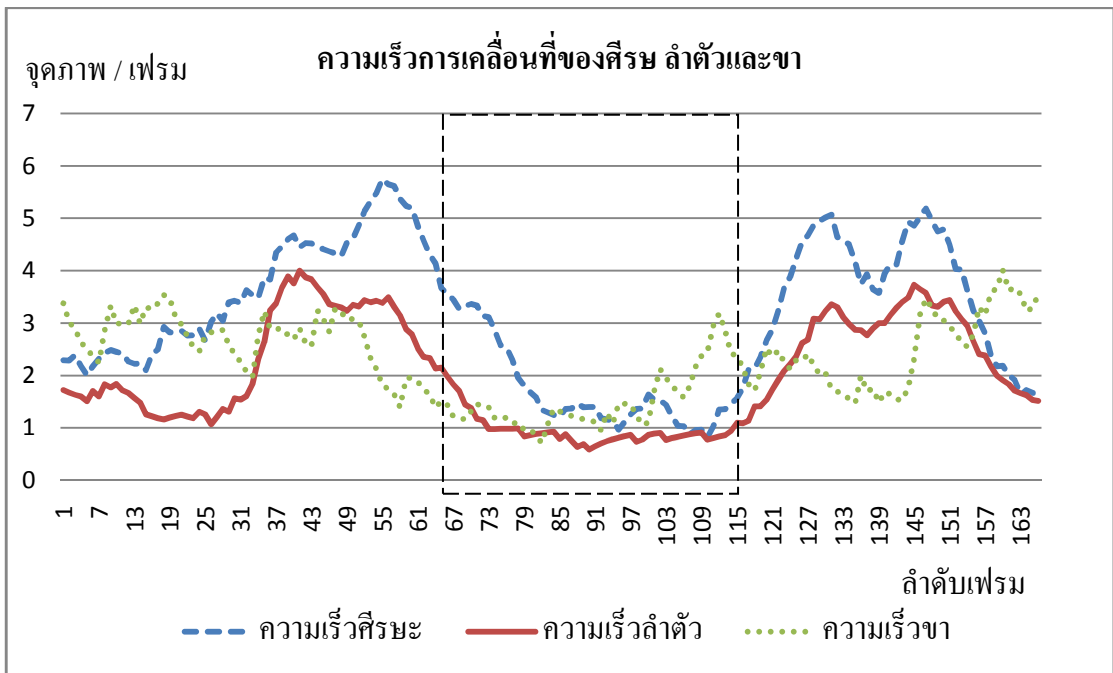
$$\frac{\partial \vec{v}_h^\theta}{\partial t} < 0 \quad (3.7)$$

ดังนั้น ลักษณะเฉพาะของการกัม คือ สมการ (3.6)

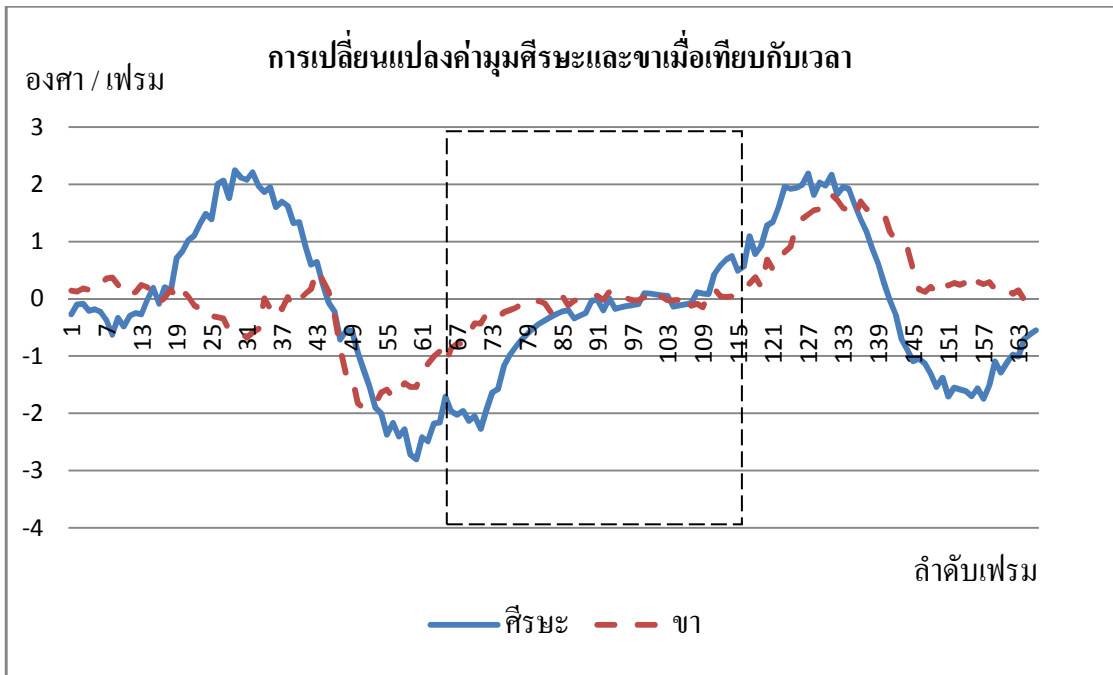
คุณลักษณะการนั่ง



(a)



(b)



(c)

ภาพประกอบ 3-14 กราฟแสดงคุณลักษณะการนั่ง พื้นที่ภายในสี่เหลี่ยมในกราฟคือช่วงที่แสดงท่าทางการนั่ง (a) กราฟแสดง  $\psi_h^\theta$  และ  $\psi_l^\theta$  (b) กราฟแสดง  $\frac{\partial \psi_{m1}^s}{\partial t}$   $\frac{\partial \psi_{m2}^s}{\partial t}$  และ  $\frac{\partial \psi_{m3}^s}{\partial t}$  (c) กราฟแสดง  $\frac{\partial \psi_l^\theta}{\partial t}$  และ  $\frac{\partial \psi_h^\theta}{\partial t}$

ลักษณะการนั่งจะตรงข้ามกับการก้ม เวกเตอร์  $\psi_h$  ก่อนข้างจะวางตัวในแกนแนวตั้งทำให้  $\psi_h^\theta$  มีค่าใกล้เคียง 0 ในขณะที่เดียวกัน  $\psi_l^\theta$  จะมีค่าลดลงเนื่องจากลักษณะเฉพาะของการนั่งที่ขาต้องโน้มไปด้านหน้าดังแสดงในภาพประกอบ 3-14 (a) และเมื่อกระบวนการนั่งเสร็จสมบูรณ์แล้วส่วนขาจะไม่มีการเคลื่อนไหวจะทำให้ค่า  $\frac{\partial \psi_{m3}^s}{\partial t}$  มีค่าใกล้เคียง 0 ดังแสดงในภาพประกอบ 3-14 (b) เป็นไปตามสมการ (3.8) เมื่อพิจารณาถึงการเปลี่ยนแปลงค่ามุมศีรษะและขาเมื่อเทียบกับเวลา ขณะที่เริ่มเข้าสู่ท่านั่งค่าของ  $\frac{\partial \psi_l^\theta}{\partial t}$  จะมีค่าเป็นลบดังแสดงในภาพประกอบ 3-14 (c) เนื่องจากขอมุมของขาจะมีค่าลดลงขณะที่ทำการนั่ง

ระหว่างที่มีการเปลี่ยนท่าทาง (ภาพประกอบ 3-17) จากการนั่งไปสู่การยืนค่ามุมของ เวกเตอร์  $\psi_l$  จะมีค่าเพิ่มขึ้น ทำให้  $\frac{\partial \psi_l^\theta}{\partial t}$  มีลักษณะเป็นไปตามสมการ (3.4) และเมื่อมีการเปลี่ยนสถานะจากท่ายืนไปสู่ท่านั่ง จะทำให้ค่า  $\frac{\partial \psi_l^\theta}{\partial t}$  มีลักษณะเป็นไปตามสมการ (3.9)

$$\psi_h^\theta = 0^\circ \mp 30^\circ, \frac{\partial \psi_{m3}^s}{\partial t} = 0 \mp 3 \tag{3.8}$$

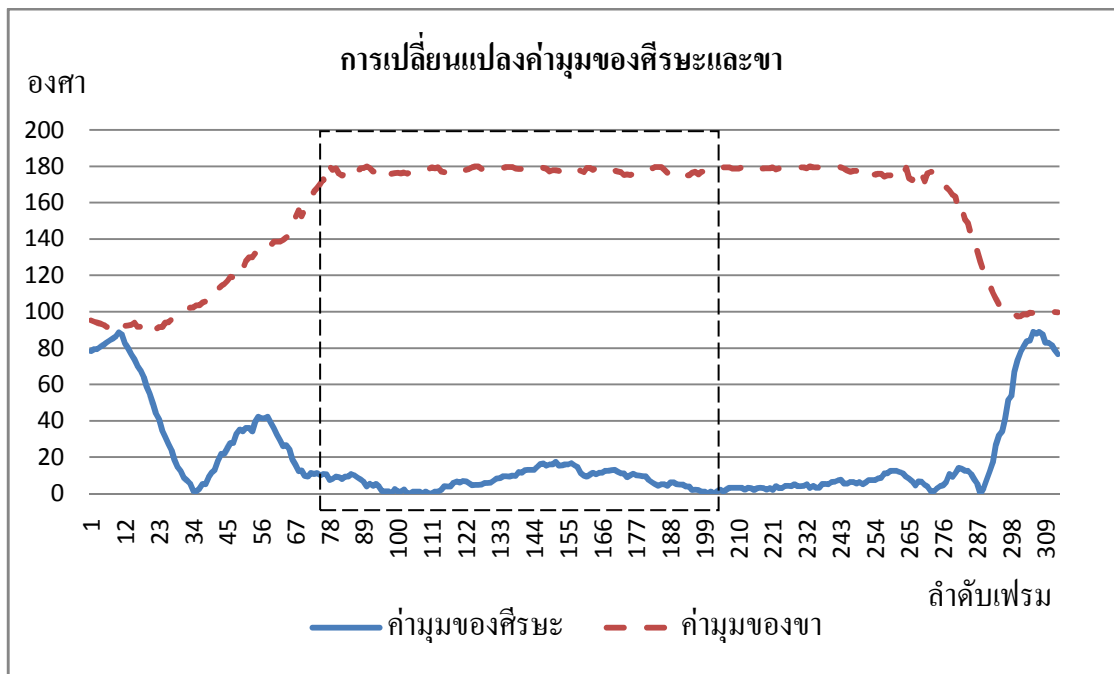
$$\frac{\partial \psi_l^\theta}{\partial t} < 0 \tag{3.9}$$

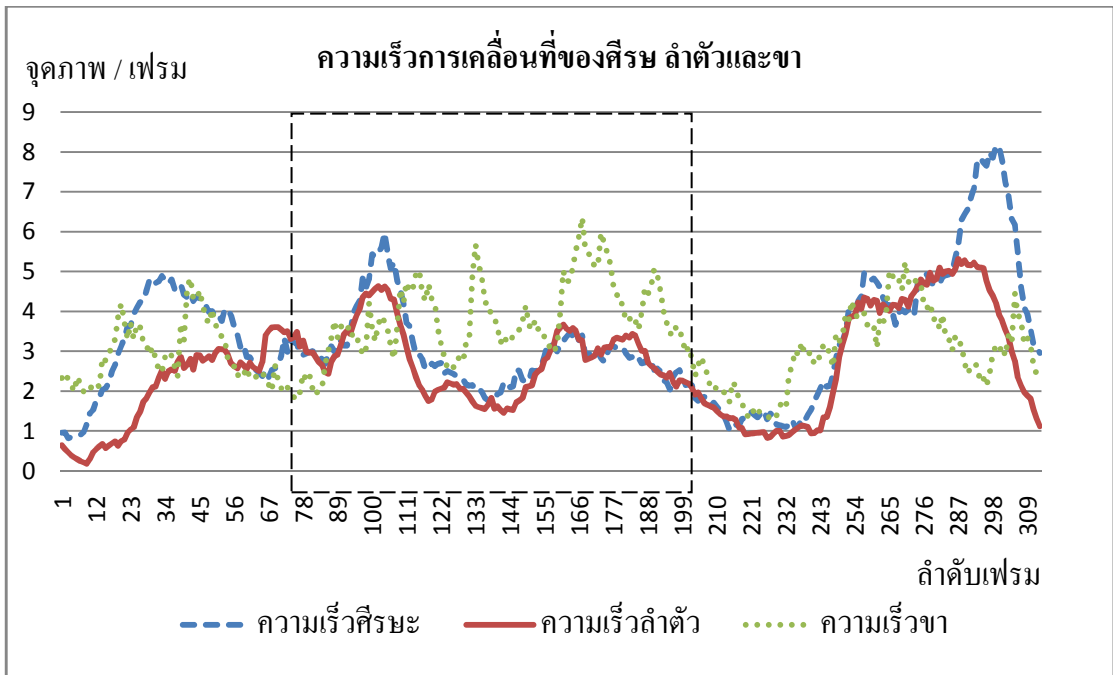
ดังนั้น ลักษณะเฉพาะของการนั่ง คือ สมการ (3.8)

### 3.2.2. แบบจำลองพลวัต (Dynamic Model)

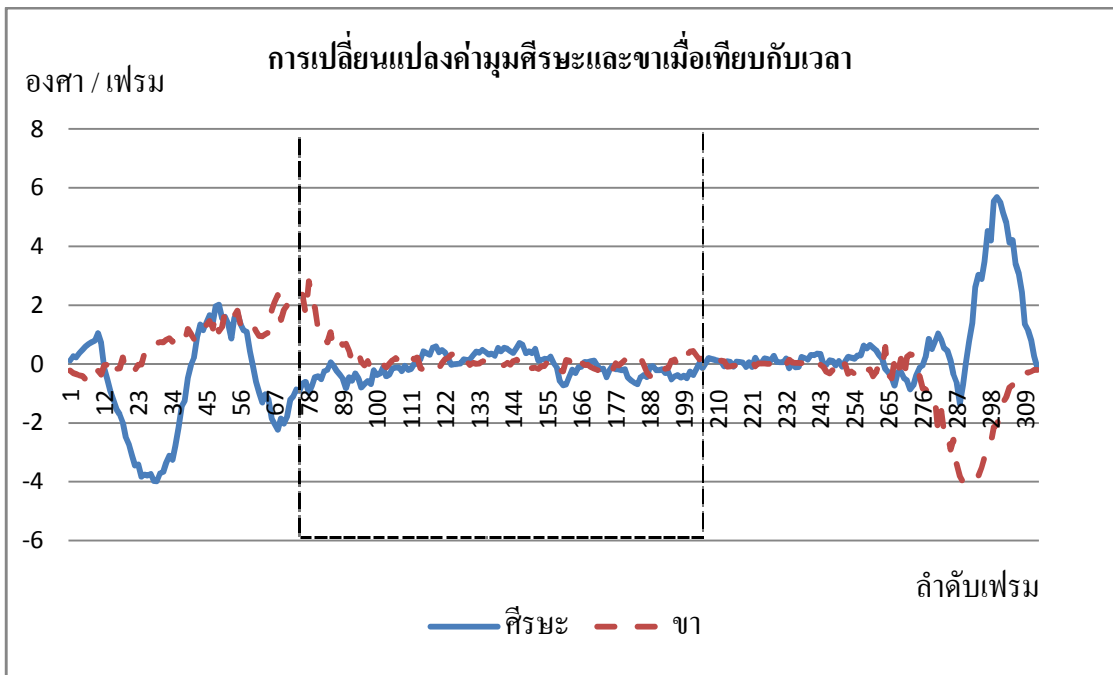
ท่าทางจะถูกพิจารณาเป็นท่าทางพลวัต ก็ต่อเมื่อ ทุกส่วนประกอบของโครงสร้างมนุษย์ อย่างง่ายมีการเคลื่อนไหว สำหรับท่าทางในลักษณะนี้เราพบว่าเฉพาะพารามิเตอร์ภายในเพียงอย่างเดียวไม่เพียงพอที่จะอธิบายหรือบ่งชี้ถึงท่าทางประเภทนี้ได้ จึงจำเป็นต้องพิจารณาพารามิเตอร์ภายนอกเพิ่มเติมด้วย

คุณลักษณะการเดิน





(b)



(c)

ภาพประกอบ 3-15 กราฟแสดงคุณลักษณะการเดิน พื้นที่ภายในสี่เหลี่ยมในกราฟคือช่วงที่แสดงท่าทางการเดิน (a) กราฟแสดง  $\chi_h^\theta$  และ  $\chi_l^\theta$  (b) กราฟแสดง  $\frac{\partial \chi_{m1}^s}{\partial t}$   $\frac{\partial \chi_{m2}^s}{\partial t}$  และ  $\frac{\partial \chi_{m3}^s}{\partial t}$  (c) กราฟแสดง  $\frac{\partial \chi_l^\theta}{\partial t}$  และ  $\frac{\partial \chi_h^\theta}{\partial t}$

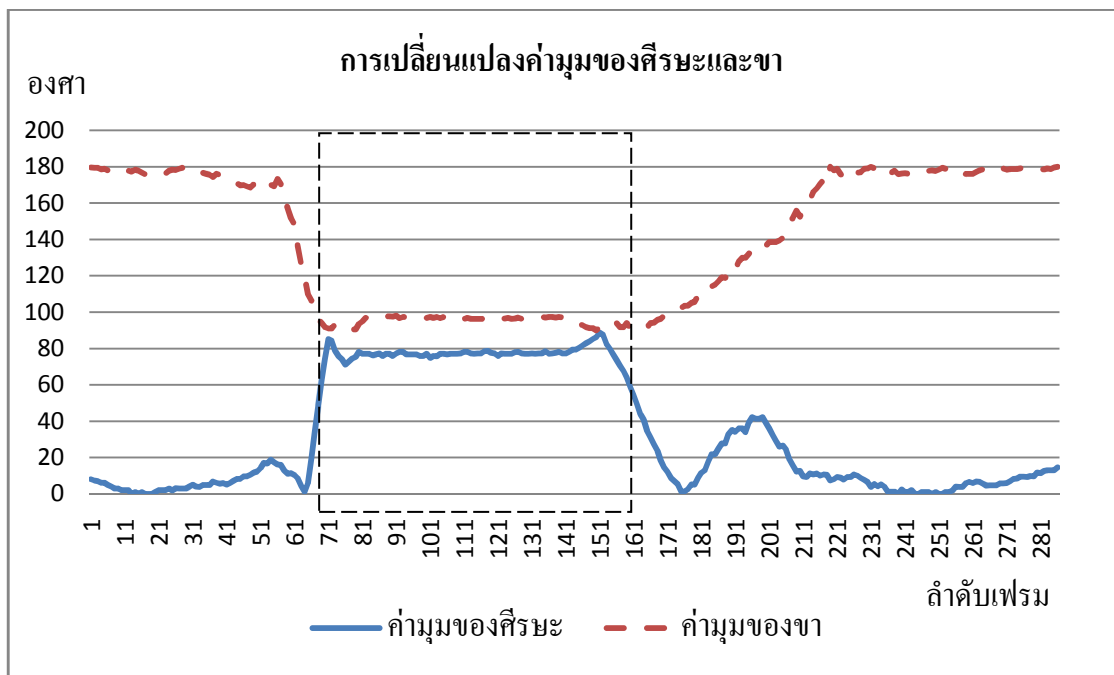


ในการกรณีการเดิน ทุกส่วนประกอบจะเคลื่อนที่ไปในทิศทางเดียวกันและมีความเร็วใน  
ทุก ๆ ส่วนใกล้เคียงกัน โดยความเร็วจะต้องมีค่ามากกว่าศูนย์ดังสมการ (3.10) ดังแสดงใน  
ภาพประกอบ 3-15 (b) ทำทางการเดินสามารถถูกจำแนกได้โดยคุณลักษณะการขึ้น คือ  $\psi_h^\theta$  และ  $\psi_l^\theta$   
จะมีค่าลักษณะเดียวกับการขึ้นดังแสดงในภาพประกอบ 3-15 (a) ร่วมกับการพิจารณาการเคลื่อนที่  
ของทุกส่วนประกอบในโครงสร้างมนุษย์ และเมื่อพิจารณาถึงการเปลี่ยนแปลงค่ามุมศีรษะและขา  
เมื่อเทียบกับเวลาจะค่อนข้างคงที่คล้ายกับการขึ้นดังแสดงในภาพประกอบ 3-15 (c)

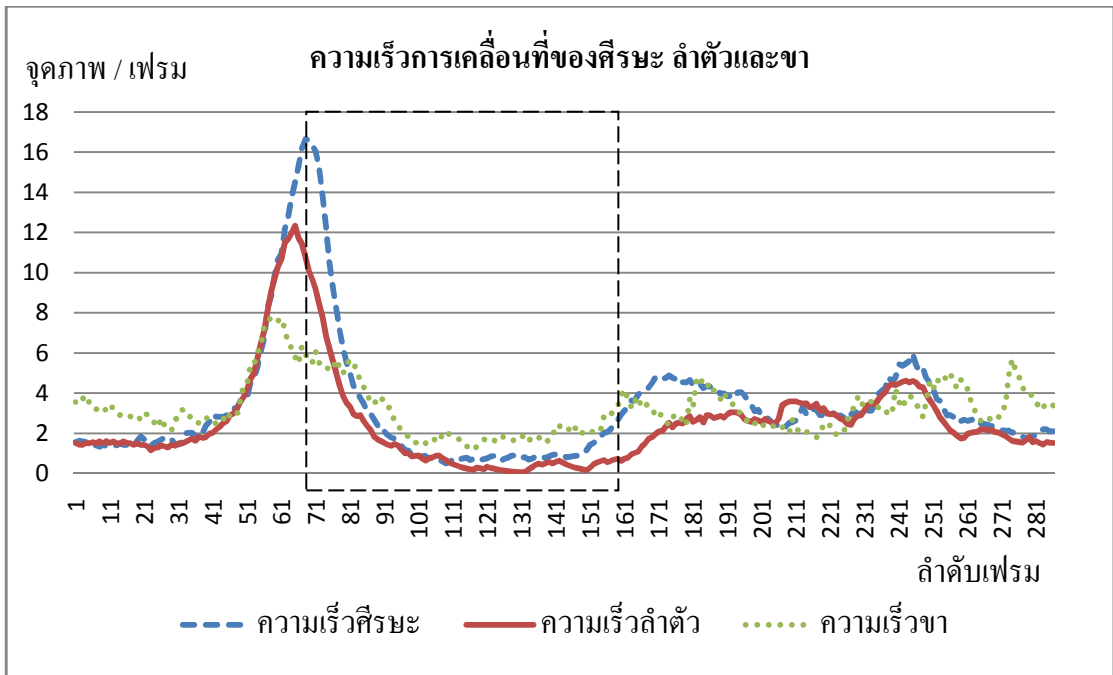
$$\frac{\partial \psi_{m1}^s}{\partial t} \cong \frac{\partial \psi_{m2}^s}{\partial t} \cong \frac{\partial \psi_{m3}^s}{\partial t} > 3 \quad (3.10)$$

ดังนั้น ลักษณะเฉพาะของการเดิน คือ สมการ (3.10) ร่วมกับลักษณะลำตัวตั้งตรง คือ  
 $\psi_h^\theta = 0^\circ \mp 30^\circ$  และ  $\psi_l^\theta = 180^\circ \mp 30^\circ$

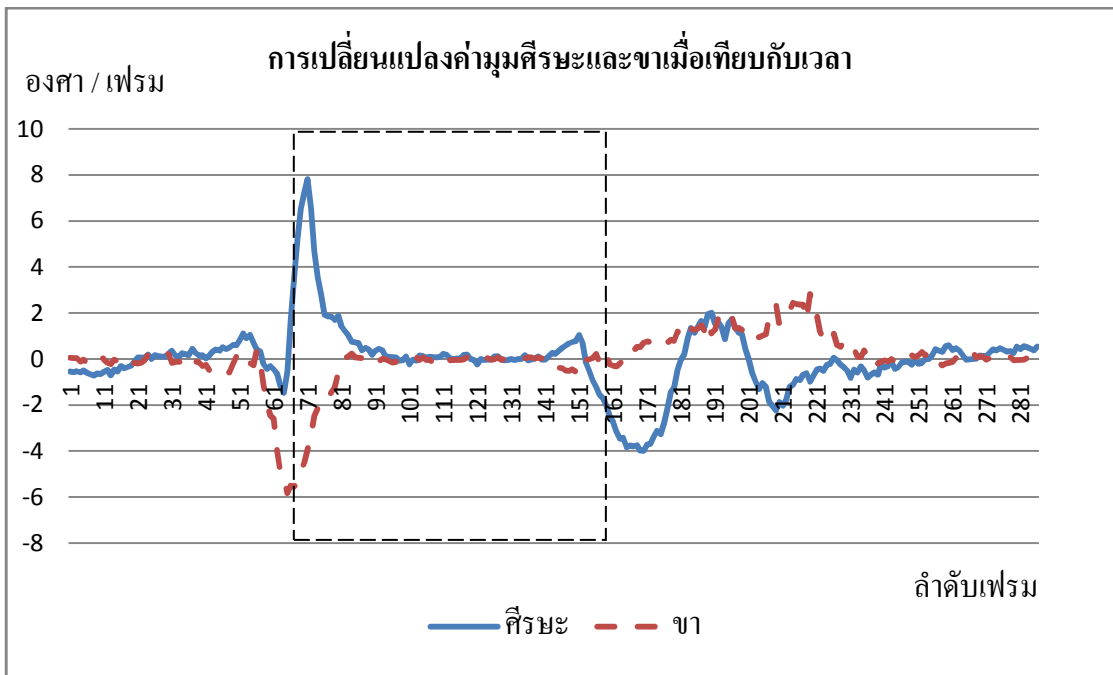
#### คุณลักษณะการนอนราบ



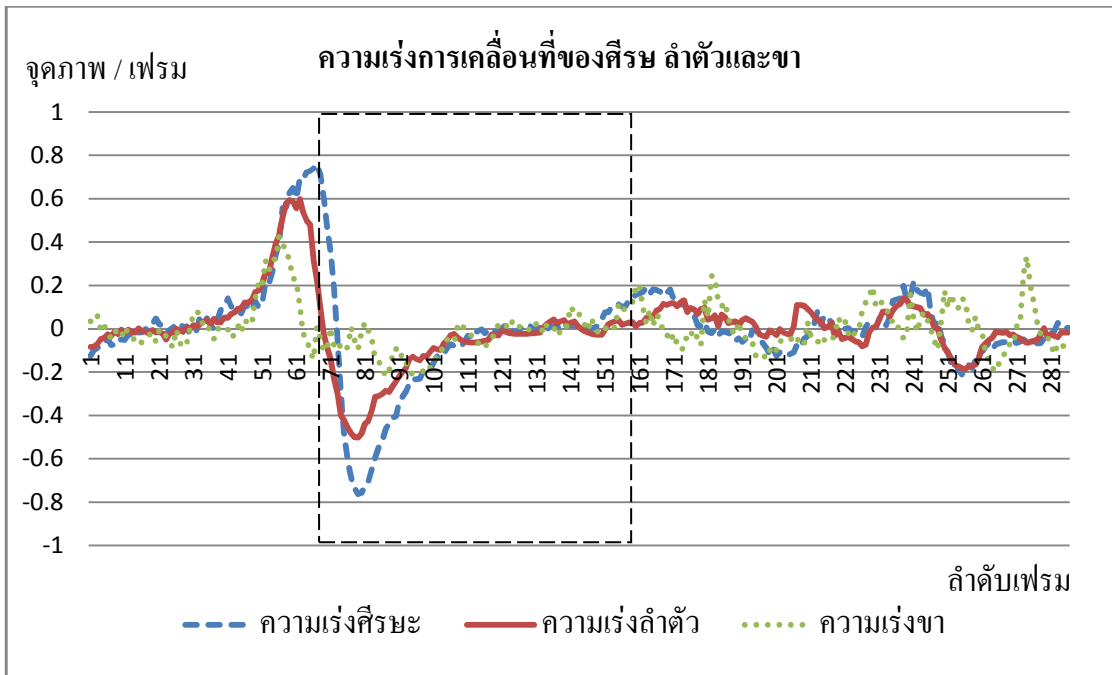
(a)



(b)



(c)



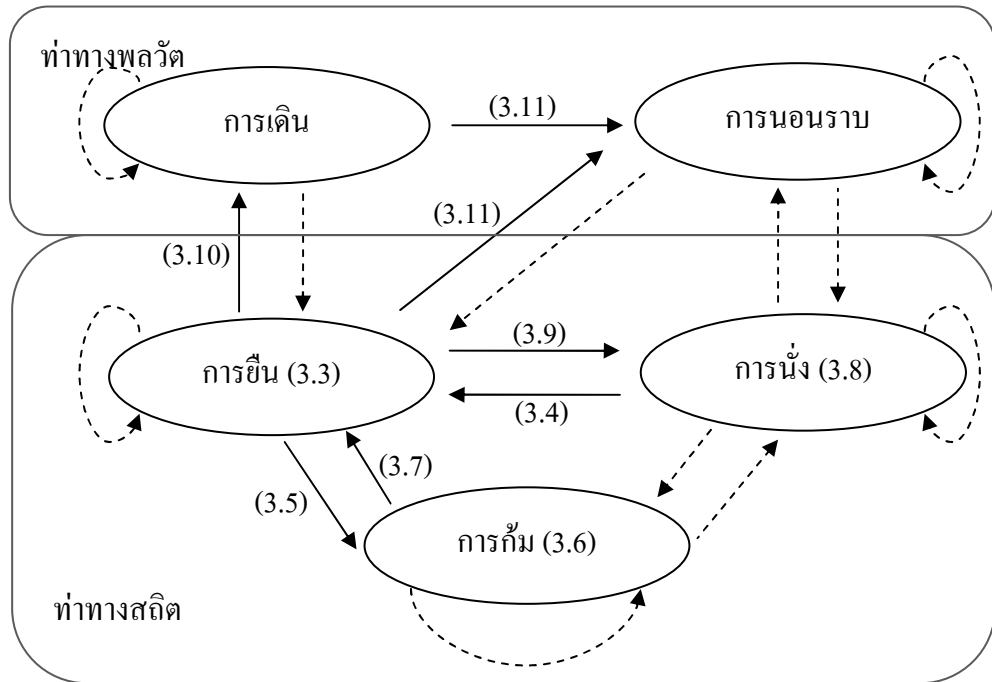
(d)

ภาพประกอบ 3-16 กราฟแสดงคุณลักษณะการนอนราบ พื้นที่ภายในสี่เหลี่ยมในกราฟคือช่วงที่แสดงท่าทางการนอนราบ (a) กราฟแสดง  $\dot{\chi}_h^\theta$  และ  $\dot{\chi}_l^\theta$  (b) กราฟแสดง  $\frac{\partial \dot{\chi}_{m1}^\theta}{\partial t}$   $\frac{\partial \dot{\chi}_{m2}^\theta}{\partial t}$  และ  $\frac{\partial \dot{\chi}_{m3}^\theta}{\partial t}$  (c) กราฟแสดง  $\frac{\partial \dot{\chi}_l^\theta}{\partial t}$  และ  $\frac{\partial \dot{\chi}_h^\theta}{\partial t}$  (d) กราฟแสดง  $\frac{\partial^2 \dot{\chi}_{m1}^\theta}{\partial t^2}$   $\frac{\partial^2 \dot{\chi}_{m2}^\theta}{\partial t^2}$  และ  $\frac{\partial^2 \dot{\chi}_{m3}^\theta}{\partial t^2}$

สำหรับท่าทางการนอนราบ การเปลี่ยนสถานะเข้าสู่การนอนราบทุกส่วนของโครงสร้างมนุษย์จะมีการเคลื่อนที่ในทิศทางที่แตกต่างกันด้วยความเร็วที่ไม่เท่ากัน โดยทั่วไปสิริษะจะเคลื่อนที่เร็วกว่าลำตัวและขา เราพบว่ามุมกึ่งองศาบางมุมเวกเตอร์  $\dot{\chi}_h$  และ  $\dot{\chi}_l$  จะวางตัวอยู่ในแนวแกนเดียวกันซึ่งอาจใกล้เคียงกับลักษณะการยืน ดังนั้นเราจึงไม่สามารถใช้เฉพาะพารามิเตอร์ภายในเพื่อการแยกแยะท่าทางการนอนราบได้ เราสามารถกำหนดนิยามการนอนราบได้โดยใช้คุณลักษณะการยืนที่ไม่ได้อยู่ในแกนแนวตั้งร่วมกับการพิจารณาคุณลักษณะการเคลื่อนไหวระหว่างการเปลี่ยนท่าทาง เนื่องจากทุกส่วนมีการเปลี่ยนตำแหน่งอย่างชัดเจนจึงทำให้ทุกส่วนนั้นมีค่าความเร่งเชิงมุมมากกว่า 0 โดยเมื่อพิจารณาความเร่งเชิงมุมของแต่ละองค์ประกอบจะพบว่าส่วนสิริษะจะมีค่าสูงกว่าลำตัวและขา ดังแสดงในภาพประกอบ 3-16 (d) ซึ่งเป็นไปตามสมการ (3.11)

$$\frac{\partial^2 \dot{\chi}_{m1}^\theta}{\partial t^2} > \frac{\partial^2 \dot{\chi}_{m2}^\theta}{\partial t^2} \geq \frac{\partial^2 \dot{\chi}_{m3}^\theta}{\partial t^2} > 3 \tag{3.11}$$

ดังนั้น ลักษณะเฉพาะของการนอนราบ คือ สมการ (3.11) ร่วมกับลักษณะลำตัวที่ไม่ได้อยู่ในเงื่อนไขการขึ้น การเดิน การก้มและการนั่ง

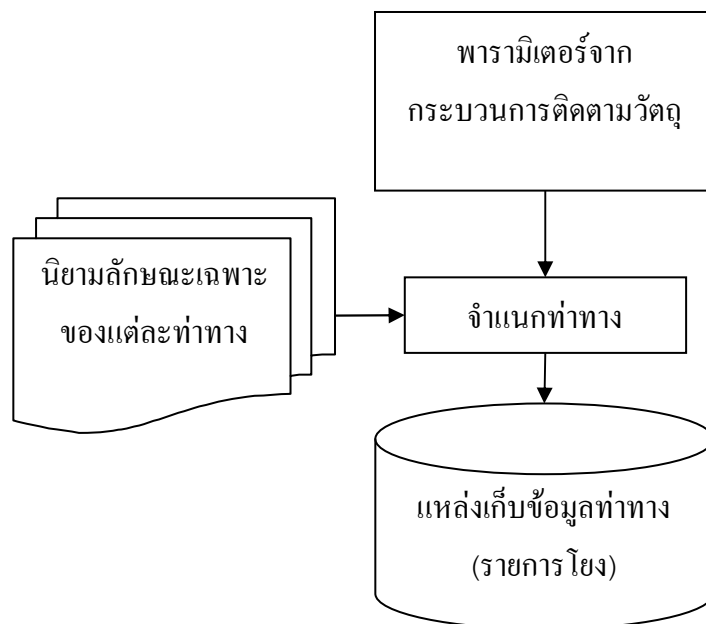


ภาพประกอบ 3-17 แผนภาพเครื่องสถานะจำกัด (finite state machine) แสดงต้นแบบการแสดงท่าทาง ลูกศรเส้นทึบและหมายเลขกำกับแสดงถึงคุณลักษณะที่ได้กำหนดไว้ตามลำดับสมการข้างต้น ลูกศรเส้นประแสดงการเปลี่ยนสถานะที่ไม่ได้กำหนดคุณลักษณะในงานวิจัยนี้

ภาพรวมคุณลักษณะท่าทางที่กำหนดไว้สามารถแสดงเป็นแผนภาพสถานะดังแสดงในภาพประกอบ 3-17 โดยกระบวนการจำแนกท่าทางมนุษย์ (ภาพประกอบ 3-18) จะใช้ข้อมูลพารามิเตอร์ที่ได้จากกระบวนการติดตามการเคลื่อนที่ของวัตถุ และพารามิเตอร์เหล่านั้นจะถูกตรวจสอบตามเงื่อนไขท่าทางมนุษย์ดังอธิบายในเนื้อหาข้างต้น เมื่อพิจารณาถึงสถานะ การขึ้น ซึ่งถูกอธิบายได้โดยลักษณะการขึ้นลำตัวตรงและไม่มี การเคลื่อนไหวใด ๆ ซึ่งจำทำให้มีลักษณะดังสมการ 3.3 ( $\psi_h^\theta = 0^\circ \mp 30^\circ, \psi_l^\theta = 180^\circ \mp 30^\circ, \frac{\partial \psi_{m1}^s}{\partial t} = \frac{\partial \psi_{m2}^s}{\partial t} = \frac{\partial \psi_{m3}^s}{\partial t} = 0 \mp 3$ ) และสถานะการนั่งจะมีลักษณะลำตัวท่อนบนค่อนข้างตั้งตรงและขามีการเคลื่อนที่ซึ่งสามารถอธิบายโดยสมการ 3.8 ( $\psi_h^\theta = 0^\circ \mp 30^\circ, \frac{\partial \psi_{m3}^s}{\partial t} = 0 \mp 3$ ) ส่วนสถานะ การก้มมุมของศีรษะจะมีค่าเพิ่มขึ้นและส่วนของลำตัวมีการเคลื่อนที่น้อยซึ่งอธิบายโดยสมการ 3.6 ( $\psi_l^\theta = 180^\circ \mp 30^\circ, \frac{\partial \psi_{m2}^s}{\partial t} = \frac{\partial \psi_{m3}^s}{\partial t} = 0 \mp 3$ ) และเมื่อมีการเปลี่ยนสถานะจากการขึ้นไปสู่การนั่งการเคลื่อนไหวบริเวณขาจะทำ

มุมของ  $\psi_i^\theta$  มีค่าลดลงซึ่งเป็นไปลักษณะดังสมการ 3.9 ( $\frac{\partial \psi_i^\theta}{\partial t} < 0$ ) และในทางกลับกันขณะที่เปลี่ยนสถานะจากการนั่งไปสู่การยืนมุมของ  $\psi_i^\theta$  มีค่าเพิ่มขึ้นซึ่งเป็นไปลักษณะดังสมการ 3.4 ( $\frac{\partial \psi_i^\theta}{\partial t} > 0$ ) ส่วนการเปลี่ยนสถานะจากการยืนไปสู่การก้มมุมของ  $\psi_h^\theta$  จะมีค่าเพิ่มขึ้นซึ่งจะเป็นไปดังสมการ 3.5 ( $\frac{\partial \psi_h^\theta}{\partial t} > 0$ ) และเมื่อมีการเปลี่ยนสถานะจากการนั่งไปสู่การยืนย่อจะทำให้มุมของ  $\psi_h^\theta$  จะมีค่าลดลงซึ่งจะเป็นไปดังสมการ 3.7 ( $\frac{\partial \psi_h^\theta}{\partial t} < 0$ ) สำหรับสถานะ การเดินนั้นเกิดจากการรวมกันของคุณสมบัติการยืนร่วมกับมีการเคลื่อนไหวของทุก ๆ องค์ประกอบซึ่งจะเป็นดังสมการ 3.10 ( $\frac{\partial \psi_{m1}^\theta}{\partial t} \cong \frac{\partial \psi_{m2}^\theta}{\partial t} \cong \frac{\partial \psi_{m3}^\theta}{\partial t} > 3$ ) สำหรับการเปลี่ยนสถานะไปสู่การนอนราบจะทำให้เกิดความเร่งเชิงมุมในทุกองค์ประกอบของโครงสร้างมนุษย์โดยบริเวณศีรษะจะมีค่ามากที่สุด ขณะที่ลำตัวและขาจะมีค่าน้อยกว่าซึ่งสามารถอธิบายได้โดยสมการ 3.11 ( $\frac{\partial^2 \psi_{m1}^\theta}{\partial t^2} > \frac{\partial^2 \psi_{m2}^\theta}{\partial t^2} \geq \frac{\partial^2 \psi_{m3}^\theta}{\partial t^2} > 3$ )

หลังจากที่ทำทางในแต่ละเฟรม ได้ถูกจำแนกแล้วจะถูกจัดเก็บลงในรายการเชื่อมโยงซึ่งจะถูกนำไปใช้ในกระบวนการรู้จำกิจกรรมต่อไป

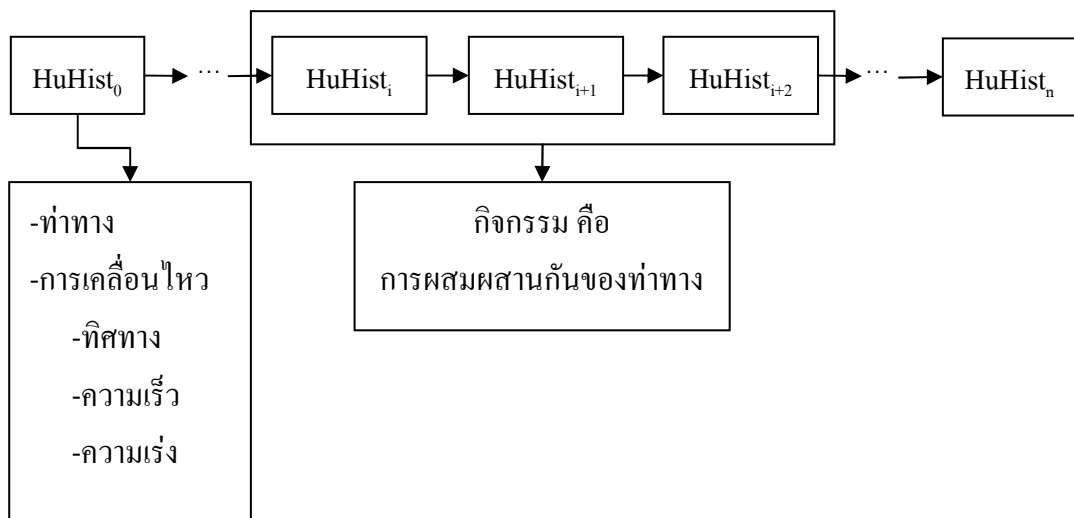


ภาพประกอบ 3-18 กระบวนการจำแนกทำทาง

### 3.3. การรู้จำกิจกรรมมนุษย์ (Human Activity Recognition)

กิจกรรมมนุษย์ สามารถถูกอธิบายในลักษณะของการผสมผสานและลำดับของท่าทางแต่ละท่าที่เกิดขึ้นอย่างต่อเนื่องกัน (ภาพประกอบ 3-19) โดยการใช้การรู้จำกิจกรรมมนุษย์ใช้ผลจาก

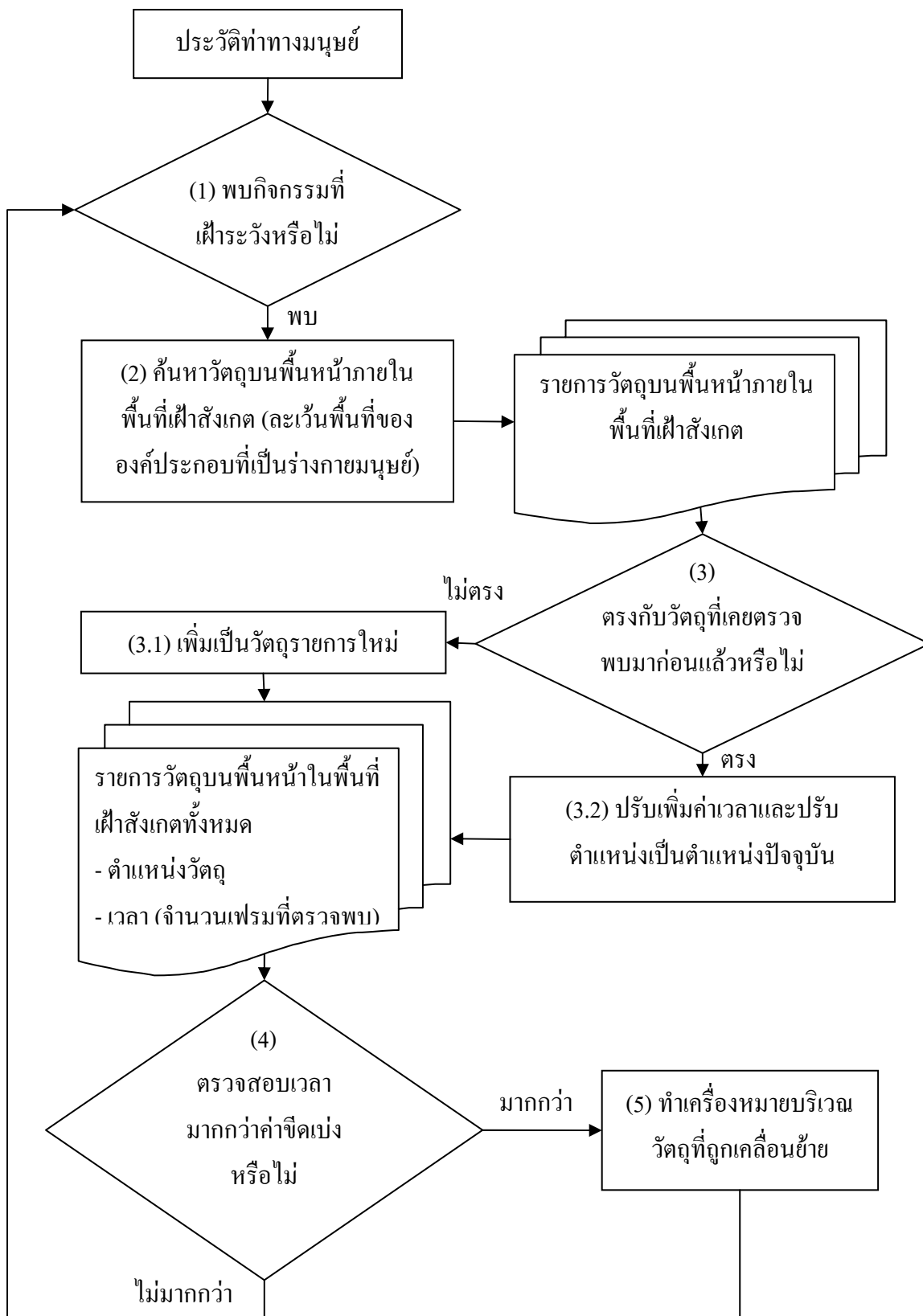
กระบวนการรู้จำท่าทางมนุษย์ ผลที่ได้จากการจำแนกท่าทางในแต่ละเฟรมจะทำให้ทราบท่าทางขณะนั้น ทิศทางการเคลื่อนที่ ความเร็วการเคลื่อนที่และความเร่งการเคลื่อนที่ของทุก ๆ องค์ประกอบในแบบจำลองอย่างง่ายของมนุษย์และข้อมูลทั้งหมดจะถูกเก็บในสถานีเชื่อมโยง (Node) ในรายการโยง ซึ่งข้อมูลเหล่านี้จะถูกใช้เป็นตัวดูคิบเพื่อในการบวนการรู้จำกิจกรรมและแจ้งเตือนต่อไป



ภาพประกอบ 3-19 ประวัติท่าทางมนุษย์

**3.3.1. กระบวนการตรวจหาการเคลื่อนย้ายวัตถุ (Object Removal)**

กิจกรรมการเคลื่อนย้ายวัตถุ หมายถึง การที่บุคคลนำวัตถุเข้าหรือออกจากบริเวณฝ้าสังเกต ซึ่งกระบวนการมีหลักการทำงานเริ่มต้นจากการฝ้าสังเกตกิจกรรมหรือท่าทางบางอย่างอันเป็นสิ่งช่วยบ่งชี้ถึงการเคลื่อนย้ายวัตถุ เช่น การถือวัตถุเข้ามาในพื้นที่ฝ้าสังเกตจากนั้นนั่งและลุกจากไป โดยทิ้งวัตถุบางสิ่งไว้ หรือตั้งใจเดินนำวัตถุเข้ามาวางทิ้งไว้แล้วเดินจากไปอย่างรวดเร็ว ซึ่งจำเป็นจะต้องมีกระบวนการอธิบายถึงกิจกรรมที่ต้องการฝ้าสังเกตอย่างชัดเจนเฉพาะเจาะจงซึ่งต้องขึ้นอยู่กับพิจารณาแต่ละกรณี และภายในงานวิจัยนี้จะสาธิตกรณีอย่างง่ายโดยพิจารณา การเคลื่อนย้ายวัตถุขณะที่มนุษย์อยู่ในท่ายืนหรือเดิน โดยส่วนใหญ่บุคคลที่ต้องการเคลื่อนย้ายวัตถุจากท่ายืนหรือเดินจะต้องมีการก้มเพื่อจะวางหรือหยิบวัตถุออกจากพื้นที่ ดังนั้นท่าทางการก้มจึงเป็นกิจกรรมฝ้าสังเกตเพื่อใช้เป็นเงื่อนไขเริ่มต้นของกระบวนการตรวจหาการเคลื่อนย้ายวัตถุ



ภาพประกอบ 3-20 กระบวนการตรวจหาการเคลื่อนย้ายวัตถุ

กระบวนการตรวจหาการเคลื่อนย้ายวัตถุดังแสดงในภาพประกอบ 3-20 ประกอบด้วย 5 ขั้นตอนดังนี้

(1) เริ่มต้นโดยการตรวจสอบลำดับของประวัติท่าทางของมนุษย์ที่ได้จากระบบการรู้จำท่าทาง โดยจะเฝ้าระวังท่าทางการก้ม ถ้าหากเกิดการก้มก็จะทำการเข้าสู่กระบวนการขั้นต่อไป

(2) ตรวจหาวัตถุต่าง ๆ ที่ปรากฏอยู่บนพื้นหน้าของพื้นที่เฝ้าสังเกต โดยพื้นที่เฝ้าสังเกตคือบริเวณที่เกิดกิจกรรมเฝ้าระวัง ซึ่งจะประมาณพื้นที่ใกล้เคียงตัวบุคคลเป็นลักษณะวงกลมที่มีจุดศูนย์กลางอยู่ที่ศูนย์กลางลำตัวและรัศมีเท่ากับครึ่งหนึ่งของด้านที่ยาวที่สุดของวัตถุ ซึ่งจากระบบการค้นหาวัตถุจะทำให้ได้รายการวัตถุที่ตรวจพบในเฟรมที่มีท่าทางการก้มเกิดขึ้น

(3) จากรายการวัตถุที่ถูกตรวจพบในเฟรมปัจจุบันจะถูกนำไปตรวจสอบกับรายการวัตถุที่เคยตรวจพบก่อนหน้านี้ ซึ่งการตรวจสอบจะพิจารณาพื้นที่ซ้อนทับระหว่างวัตถุที่เคยตรวจพบมาก่อนหน้าและที่ตรวจพบในเฟรมปัจจุบัน ถ้าหากทั้งสองต่างมีพื้นที่ทับซ้อนกันมากกว่า 80 เปอร์เซ็นต์ วัตถุในเฟรมปัจจุบันชิ้นนี้จะถูกตัดสินว่าเป็นวัตถุที่เคยตรวจพบมาแล้ว และจะในขั้นตอน (3.1) โดยทำการปรับปรุงข้อมูลของวัตถุชิ้นในรายการวัตถุที่เคยถูกตรวจพบมาแล้วโดยจะทำการปรับตำแหน่งของวัตถุให้เป็นตำแหน่งของวัตถุในเฟรมปัจจุบันและเพิ่มค่าเวลาของวัตถุชิ้นนี้ซึ่งเวลาหมายถึงจำนวนเฟรมที่พบวัตถุชิ้นนี้ แต่ถ้าหากวัตถุในเฟรมปัจจุบันไม่ถูกตรวจพบเป็นทับซ้อนกับวัตถุก่อนหน้านี้ชิ้นใดเลย จะทำขั้นตอน (3.2) โดยวัตถุนั้นจะถูกเพิ่มเข้าไปในรายการวัตถุที่ถูกตรวจพบและเริ่มต้นกำหนดการนับเวลาเป็น 1

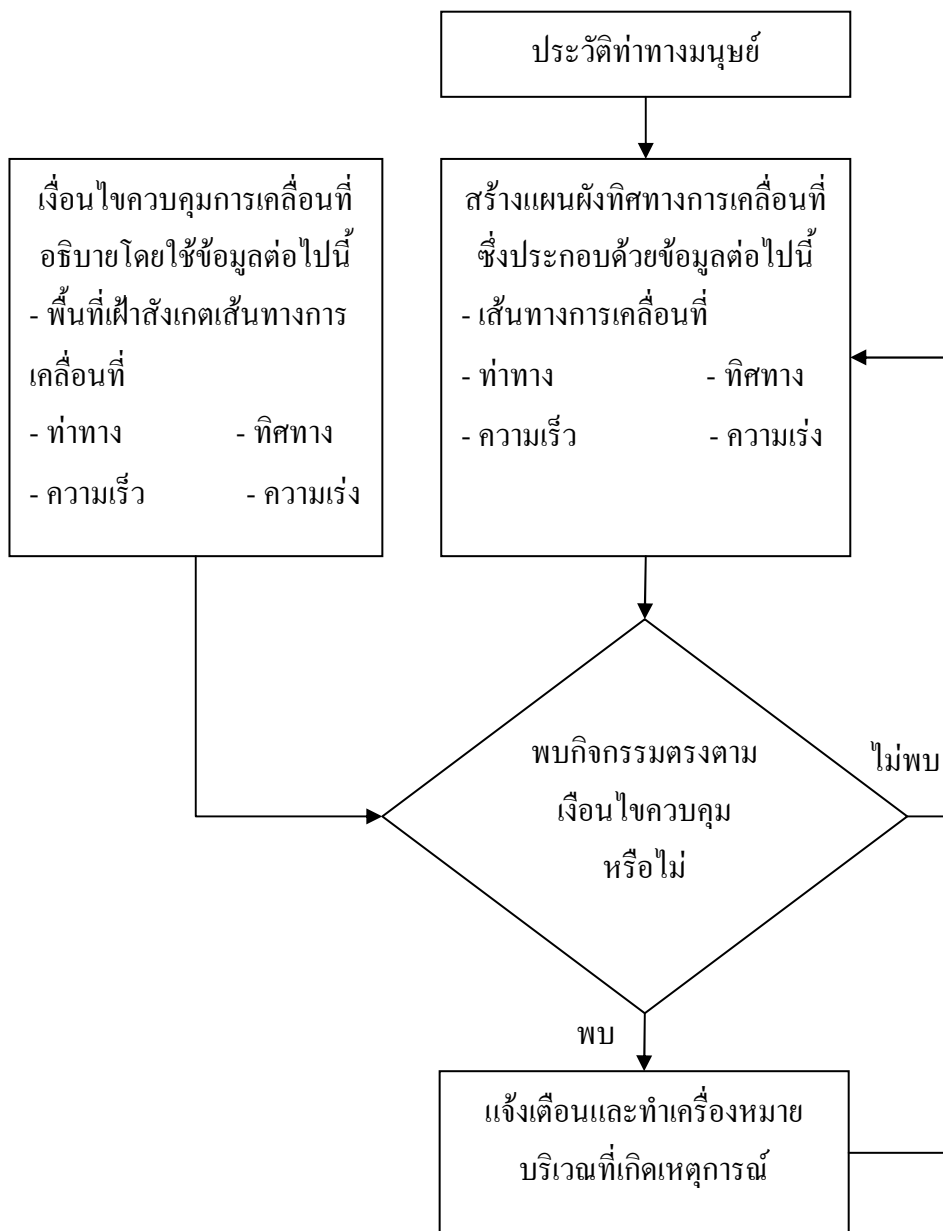
(4) ตรวจสอบเวลาของวัตถุที่ถูกตรวจพบทั้งหมด โดยถ้าหากมีค่ามากกว่าค่าขีดแบ่ง ซึ่งในงานวิจัยนี้ค่าขีดแบ่งอยู่ที่ 75 เฟรม (หรือ 3 วินาทีโดยคิดจากอัตราเฟรม 25 เฟรมต่อวินาที) วัตถุนั้นจะถูกทำเครื่องหมาย

(5) ทำเครื่องหมายวัตถุโดยใช้สัญลักษณ์วงกลมล้อมรอบวัตถุเพื่อเน้นให้เห็นบริเวณที่มีการเคลื่อนย้ายวัตถุ

### 3.3.2. กระบวนการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่ (Direction Control)

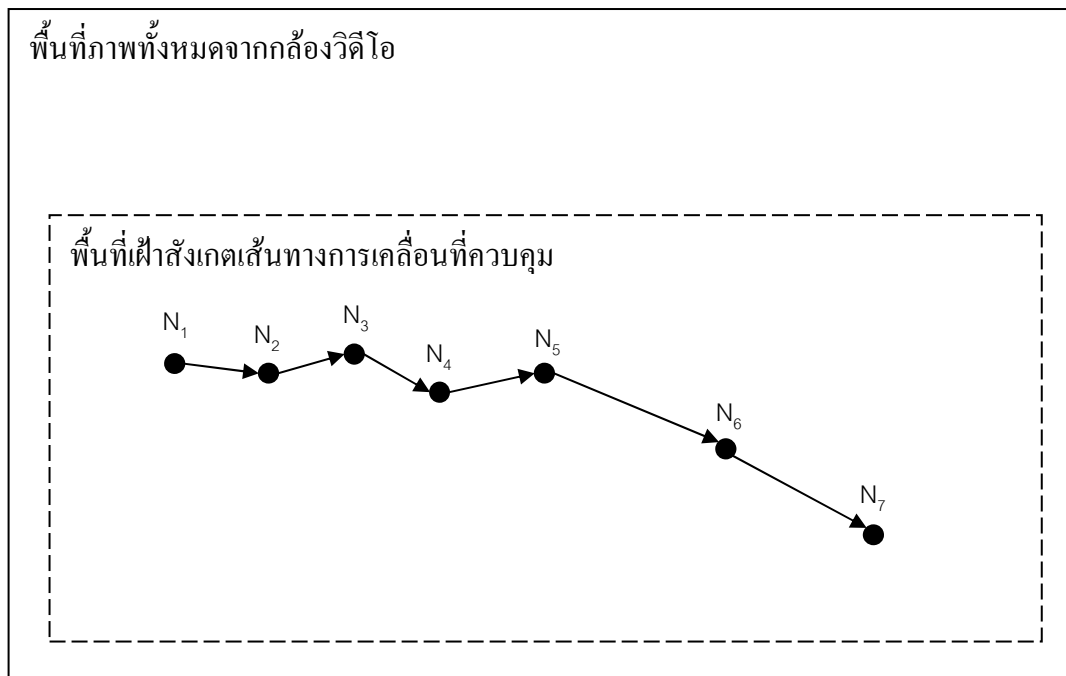
กระบวนการควบคุมการเคลื่อนที่ จะใช้ข้อมูลจากแต่ละสถานีเชื่อมโยงในรายการโยงเพื่อสร้างแผนผังลำดับทิศทางการเคลื่อนที่ของบุคคลซึ่งจะประกอบด้วย เส้นทางเคลื่อนที่ ท่าทางขณะนั้น ทิศทาง ความเร็วและความเร่งที่เกิดขึ้นอย่างต่อเนื่องกัน





ภาพประกอบ 3-21 กระบวนการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่

สำหรับกระบวนการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่เริ่มต้น โดยดึงข้อมูลจากประวัติทำทางมนุษย์มาสร้างเป็นแผนผังทิศทางการเคลื่อนที่ซึ่งประกอบด้วยข้อมูลเส้นทางเคลื่อนที่ ทำทาง ทิศทางการเคลื่อนที่ ความเร็วและความเร่ง ซึ่งข้อมูลเหล่านี้จะถูกใช้ตรวจสอบกับเงื่อนไขควบคุมที่ได้ออกแบบไว้ล่วงหน้าก่อนแล้ว ถ้าหากกิจกรรมที่เกิดขึ้นตรงกับเงื่อนไขควบคุมจะเข้าสู่กระบวนการแจ้งเตือนและทำเครื่องหมายบริเวณที่ตรวจพบกิจกรรมนั้น ๆ โดยตัวอย่างแผนผังทิศทางการเคลื่อนที่แสดงในภาพประกอบ 3-22



ภาพประกอบ 3-22 ตัวอย่างแผนผังทิศทางการเคลื่อนที่

แต่ละสถานีเชื่อมต่อ (N<sub>n</sub>) ในภาพประกอบ 3-22 จะประกอบด้วยข้อมูลดังต่อไปนี้

- ตำแหน่งปัจจุบัน
- ทำทางปัจจุบัน
- ทิศทางปัจจุบัน
- ความเร็วปัจจุบัน
- ความเร่งปัจจุบัน

ข้อมูลในแต่ละสถานีเชื่อมต่อจะได้มาจากการประมาณจากทั้งสามองค์ประกอบ คือ สิริยะลำตัวและขา โดยทุกเฟรมที่ผ่านกระบวนการประมวลผลแผนผังทิศทางการเคลื่อนที่จะถูกปรับเพิ่มเติมตามข้อมูลที่ถูกรวบรวมใหม่ในเฟรมปัจจุบัน และทุก ๆ ครั้งที่มีการปรับปรุงแผนผังการเคลื่อนที่เสร็จสิ้นจะเข้าสู่กระบวนการตรวจสอบกิจกรรมตามเงื่อนไขควบคุมเสมอ ถ้าหากสถานีเชื่อมต่อ (N<sub>n</sub>) ไດ ๆ อยู่ในพื้นที่ฝ้าสังเกตเส้นทางการเคลื่อนที่ควบคุม ข้อมูลของสถานีเชื่อมโอนั้น (อาจรวมถึงสถานีเชื่อมต่อใกล้เคียง ขึ้นอยู่กับการกำหนดเงื่อนไขควบคุม) จะถูกตรวจสอบตามเงื่อนไขควบคุม ซึ่งถ้าหากเป็นไปตามเงื่อนไขควบคุมก็จะเข้าสู่กระบวนการแจ้งเตือนและทำเครื่องหมายบริเวณนั้น ๆ ต่อไป

### 3.4. สรุป

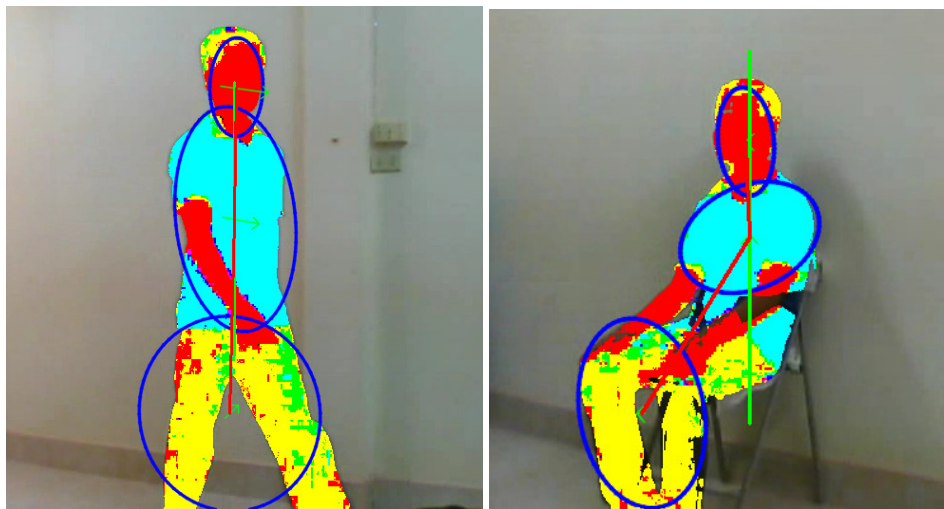
เนื้อหาบทนี้กล่าวถึงกระบวนการจذبกิจกรรมฝึกปฏิบัติซึ่งประกอบด้วย กระบวนการสร้างโครงสร้างมนุษย์ขึ้นใหม่โดยจำลองเป็นโครงสร้างมนุษย์อย่างง่ายและติดตามการเคลื่อนไหวของโครงสร้างนั้นซึ่งผลการติดตามจะถูกนำไปใช้จำแนกเป็นท่าทางพื้นฐาน แบ่งเป็น 2 กลุ่ม ได้แก่ (1) ท่าทางสถิต ประกอบด้วย การยืน การนั่ง การก้ม และ (2) ท่าทางพลวัต ประกอบด้วย การนอนราบ การเดิน จากการนำท่าทางพื้นฐานเหล่านี้มาผสมผสานกันจะสามารถถูกนำไปใช้อธิบายถึงการรู้จำกิจกรรมมนุษย์โดยจะมุ่งเน้นไปที่กระบวนการหาการเคลื่อนไหววัตถุและการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่

## บทที่ 4

### ผลการทดสอบ

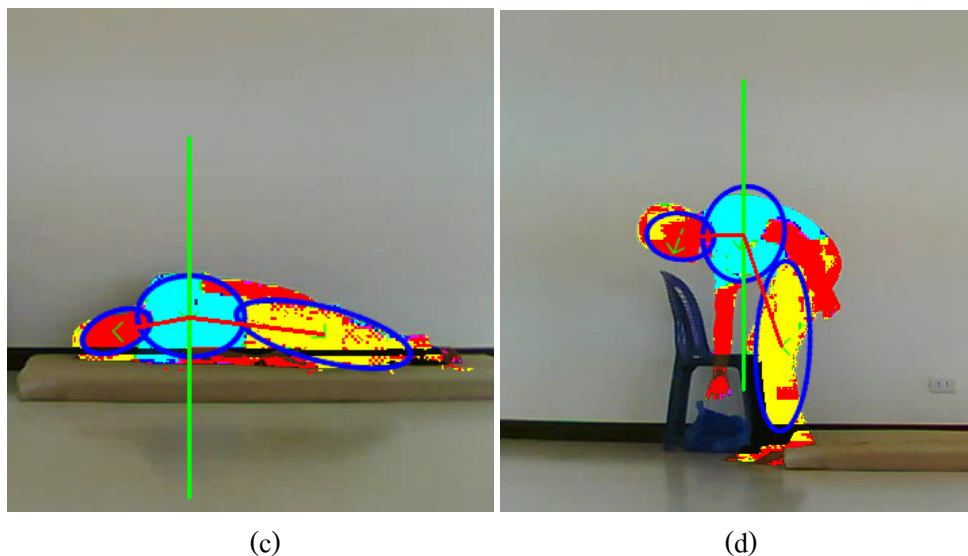
#### 4.1. ผลการทดลองการรู้จำท่าทางมนุษย์

การทดลองใช้ชุดข้อมูลภาพวิดีโอที่ถ่ายทำภายในสภาพแวดล้อมภายในอาคาร โดยใช้กล้องจำนวน 1 ตัว ความละเอียดขนาด 640x480 จุดภาพ กล้องไม่มีการเปลี่ยนมุมมองและไม่เคลื่อนที่ โดยใช้วัตถุทดลองเป็นบุคคลทดสอบห้าคนที่ใส่เสื้อและกางเกงสีแตกต่างกัน โดยแต่ละบุคคลจะแสดงท่าทางที่ต้องการทดสอบแบบสุ่มอย่างต่อเนื่อง โดยพิจารณาท่าทางทั้งหมดดังนี้ ท่าทางการยืน 77 ครั้ง การนั่ง 28 ครั้ง การก้ม 50 ครั้ง การเดิน 44 ครั้ง และการนอนราบ 11 ครั้ง และเนื่องจากภายในงานวิจัยนี้เน้นทดสอบกระบวนการรู้จำท่าทางและการรู้จำพฤติกรรมมนุษย์ดังนั้นภาพวิดีโอตัวอย่างที่เลือกใช้ทดสอบจึงเลือกชุดข้อมูลภาพที่มีความผิดพลาดจากข้อจำกัดของกระบวนการลบพื้นหลังน้อยที่สุด



(a)

(b)



ภาพประกอบ 4-1 ตัวอย่างท่าทางที่พิจารณา (a) การเดิน (b) การยืน (c) การนอนราบ (d) การก้ม

ภาพประกอบ 4-1 แสดงผลของการสังเกตแบบจำลองมนุษย์พร้อมพารามิเตอร์โดยกระบวนการตรวจหาและติดตามมนุษย์ด้วยคุณสมบัติการเคลื่อนที่และสีพื้นผิวโดยพื้นที่สีแดงแสดงศีรษะและมือ 2 ข้าง สีน้ำเงินแสดงถึงลำตัว และสีเหลืองแสดงส่วนขา 2 ข้าง สังเกตเห็นได้ว่านอกจากจะอธิบายถึงศีรษะแล้วยังสามารถแสดงถึงแขนได้ด้วยเช่นกัน เนื่องจากแขนทั้งสองข้างและใบหน้ามีสีผิวแบบเดียวกัน อย่างไรก็ตามการใช้ข้อจำกัดของแบบจำลองมนุษย์อย่างง่ายของร่างกายมนุษย์ที่ไม่ได้พิจารณาแขนก็ยังคงใช้ได้ดีในกรณีการรู้จำท่าทางส่วนใหญ่ดังส่วนแขนจึงไม่ถูกนำมาพิจารณาภายในงานวิจัยนี้ วงรีสีน้ำเงินบนภาพประกอบ 4-1 แสดงให้เห็นองค์ประกอบสำคัญแต่ละส่วนของแบบจำลองและเส้นสีแดงที่เชื่อมต่อศูนย์กลางระหว่างลำตัวกับศีรษะและลำตัวกับขาแสดงถึงพารามิเตอร์ภายใน คือ  $\psi_h, \psi_l$  ตามลำดับ เราสามารถคำนวณพารามิเตอร์ภายนอก คือ  $\psi_{m1}, \psi_{m2}, \psi_{m3}$  ได้จากการเคลื่อนที่ของแต่ละองค์ประกอบโดยใช้ผลจากกระบวนการติดตามการเปลี่ยนแปลงในแต่ละเฟรม ลูกศรสีเขียวที่จุดศูนย์กลางของแต่ละองค์ประกอบแสดงความเร็วและทิศทางของการเคลื่อนที่โดยทิศทางที่หัวลูกศรชี้ไปคือทิศการเคลื่อนที่ ส่วนความยาวของลูกศรแทนความเร็วในการเคลื่อนที่

ภาพประกอบ 4-2 ตัวอย่างผลกระบวนการรู้จำท่าทางการนอนเดินและการยืนประกอบด้วย กราฟแสดงการเปลี่ยนแปลงของความเร็วการเคลื่อนที่ในแต่ละองค์ประกอบของร่างกายและกราฟแสดงการเปลี่ยนค่ามุมของศีรษะและขาซึ่งเป็นพารามิเตอร์สำคัญที่ถูกใช้เพื่อจำแนกท่าทาง จากภาพประกอบ 4-2 ช้ายเป็นท่าทางการเดินก่อนที่จะเกิดท่าทางการยืนจะสังเกตเห็นว่าขณะที่อยู่ในท่าทางการเดินค่ามุมของศีรษะจะมีค่าใกล้เคียง 0 องศาและค่ามุมของขา

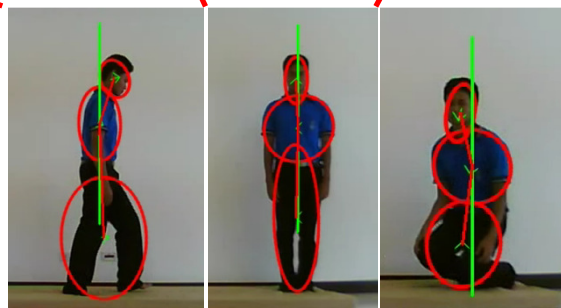
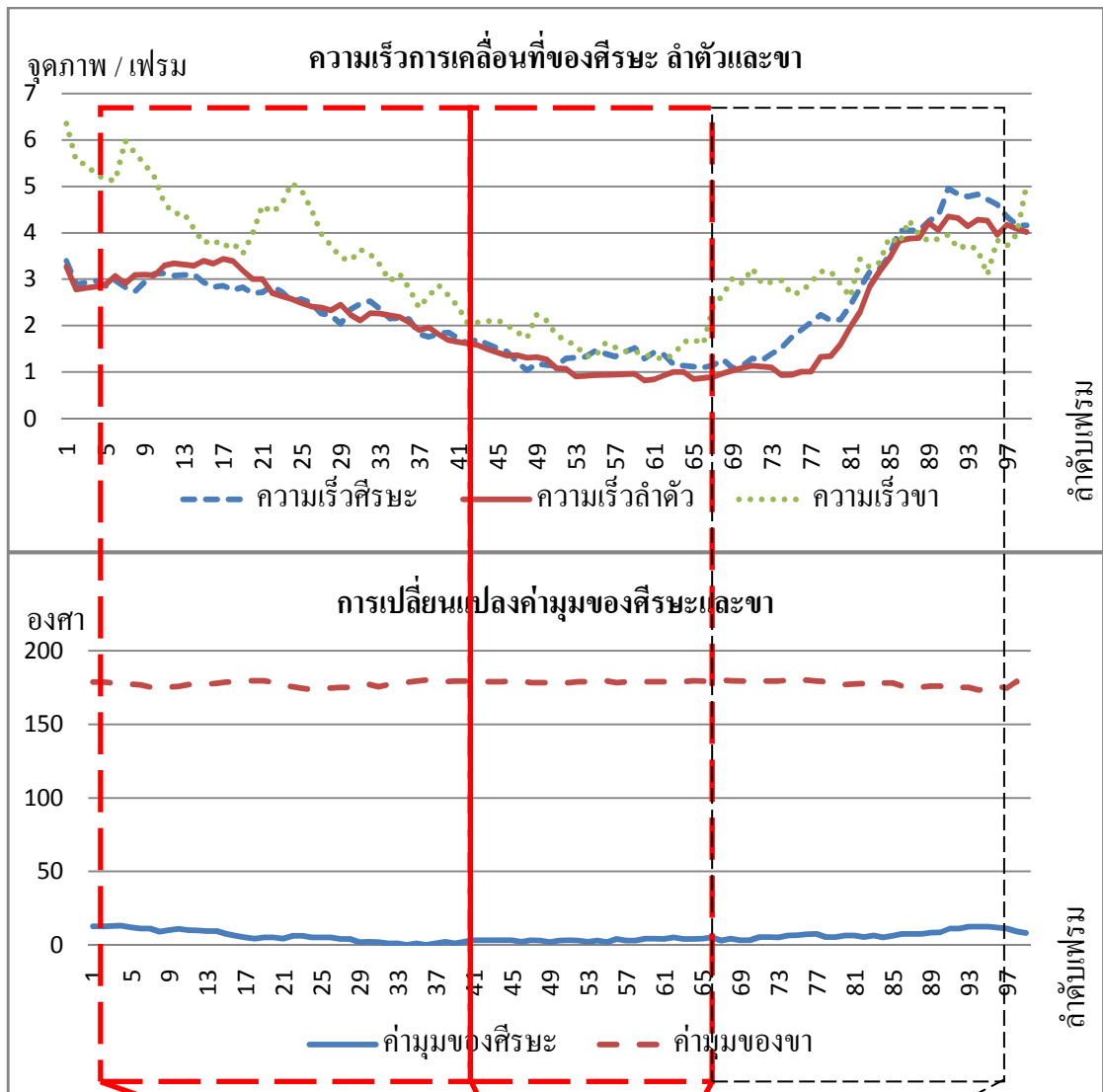
จะมีค่าใกล้เคียง 180 องศาและแทบจะไม่มี การเปลี่ยนแปลงเลย เพราะการเดินทางโดยทั่วไปถ้าตัวจะตั้งตรงและในขณะเดียวกันจะเกิดความเร็วในการเคลื่อนที่ทุก ๆ องศาประกอบด้วยเช่นกันและเมื่อบุคคลเริ่มเปลี่ยนท่าทางเข้าสู่การยืนมุมของศีรษะและมุมของขายังมีค่าคงที่แต่ความเร็วของทุกองศาประกอบจะลดลงและมีค่าต่ำเนื่องจากไม่มีการเคลื่อนที่

ภาพประกอบ 4-3 ตัวอย่างผลกระบวนการรู้จำท่าทางการนอนก้ม ในขณะที่มีการเปลี่ยนแปลงจากท่าการเดินทางไปสู่การก้มส่วนที่มีการเปลี่ยนมากที่สุด คือ องศาประกอบศีรษะ ในขณะที่ก้มศีรษะลงจะทำให้มุมของศีรษะมีค่าเพิ่มขึ้นอย่างมากและรวดเร็วทำให้ความเร็วของการเปลี่ยนแปลงค่ามุมมีค่ามากกว่าศูนย์และความเร็วในส่วนศีรษะก็จะเพิ่มขึ้นอย่างมากด้วยเช่นกันและในขณะที่เงยศีรษะขึ้นกลับสู่ท่ายืนค่ามุมศีรษะก็จะขึ้นไปในขณะตรงข้าม คือ มีค่าลดลงอย่างมากและรวดเร็วทำให้ความเร็วของการเปลี่ยนแปลงค่ามุมมีค่าน้อยกว่าศูนย์แต่ความเร็วส่วนศีรษะก็ยังคงเพิ่มขึ้นอย่างมากเช่นเดียวกับขณะก้มศีรษะลง

ภาพประกอบ 4-4 ตัวอย่างผลกระบวนการรู้จำท่าทางนั่ง การเปลี่ยนแปลงจากท่ายืนไปสู่ท่า นั่งนั้นจะทำให้ให้มุมของขา มีค่าลดลงทำให้ความเร็วของการเปลี่ยนแปลงค่ามุมมีค่าน้อยกว่าศูนย์ และมุมของศีรษะก็ยังคงมีค่าใกล้เคียง 0 องศา ดังแสดงในกราฟแสดงค่ามุมและเมื่อการนั่งเสร็จสมบูรณ์ความเร็วการเคลื่อนที่ก็จะมีค่าต่ำในลักษณะเดียวกับการยืน ในขณะที่การเปลี่ยนแปลงจากท่า นั่งกลับไปสู่ท่ายืนจะทำให้ค่ามุมของขา มีค่าเป็นเพิ่มขึ้นทำให้ความเร็วของการเปลี่ยนแปลงค่ามุมมีค่ามากกว่าศูนย์

ภาพประกอบ 4-5 ตัวอย่างผลกระบวนการรู้จำท่าทางการนอนราบ การเปลี่ยนแปลงจากท่า ยืนไปสู่ท่านอนราบในลักษณะที่มีการเปลี่ยนแปลงอย่างรวดเร็วจะทำให้ความเร็วการเคลื่อนที่ของแต่ละส่วนมีค่าเพิ่มขึ้นอย่างมากและทิศทางการมีการเปลี่ยนแปลงอย่างรวดเร็วทำให้มีค่าความเร่งเชิงมุมเกิดขึ้นและเมื่อเข้าสู่ท่านอนราบแล้วมุมศีรษะและมุมขาจะอยู่ในแนวเดียวกันลักษณะใกล้เคียงกับการยืนแต่มุมจะไม่อยู่ในแนวแกนตั้ง

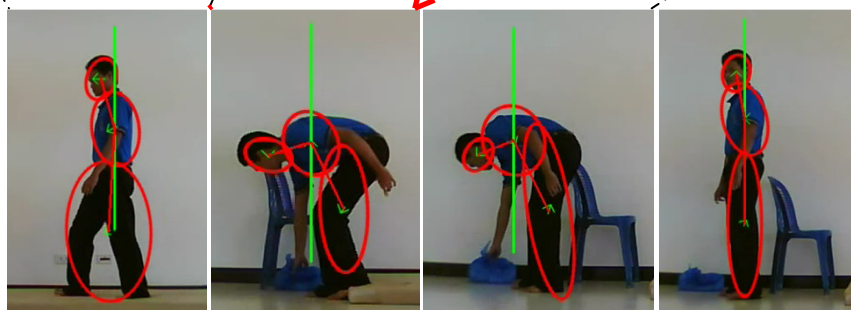
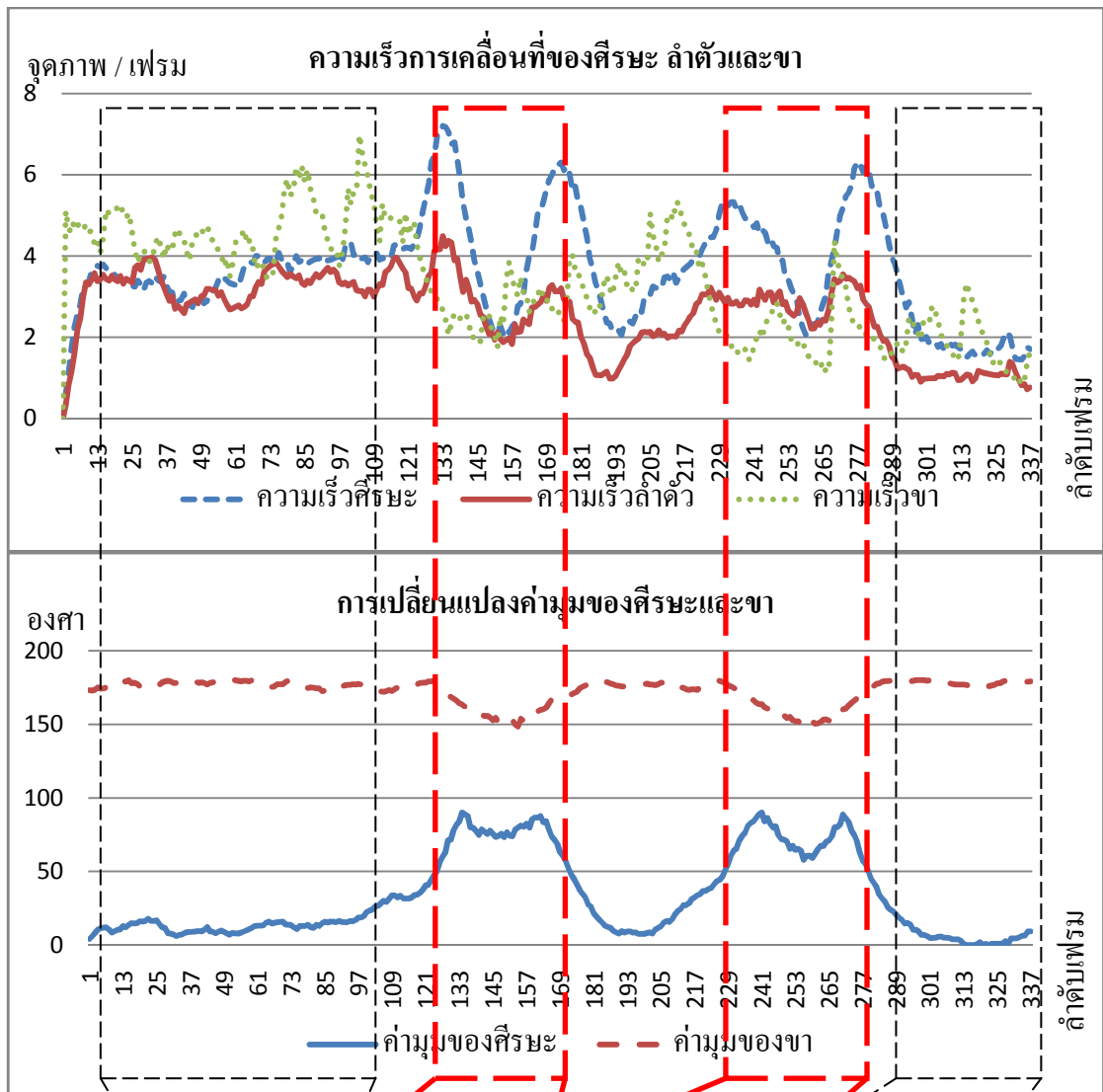
การเดินและการยืน



การเดิน      การยืน      N/A

ภาพประกอบ 4-2 ตัวอย่างผลกระบวนการรู้จำท่าทางการเดินและยืน

การก้ม

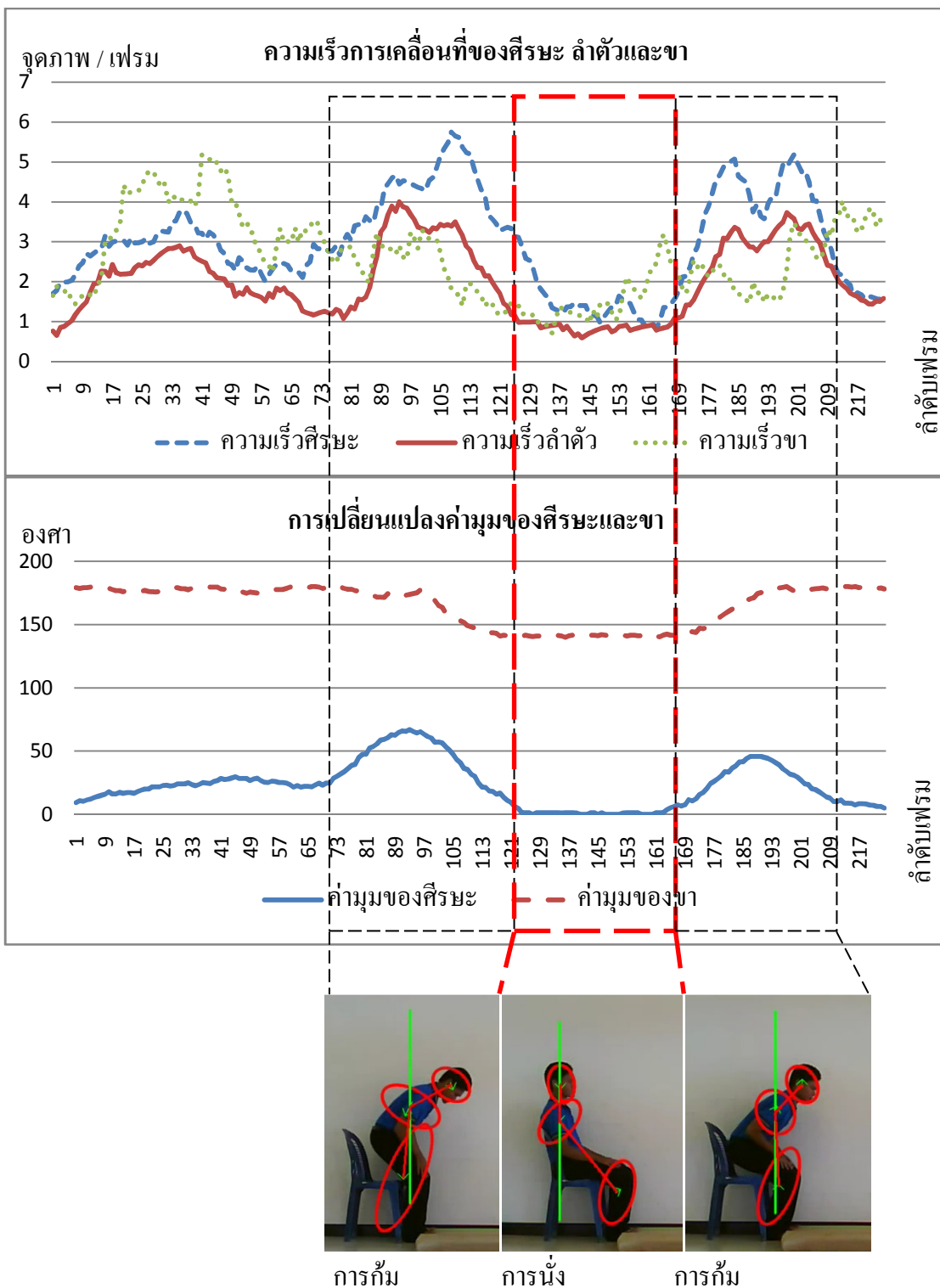


การเดิน      การก้ม      การก้ม      การขึ้น

ภาพประกอบ 4-3 ตัวอย่างผลกระบวนกรู้อำท่าทางการก้ม

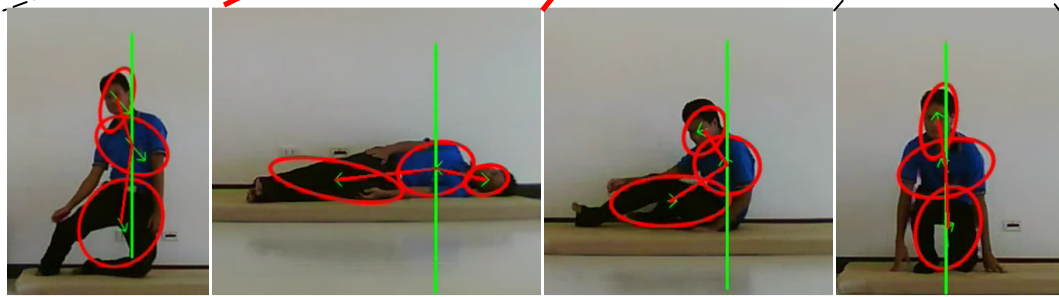
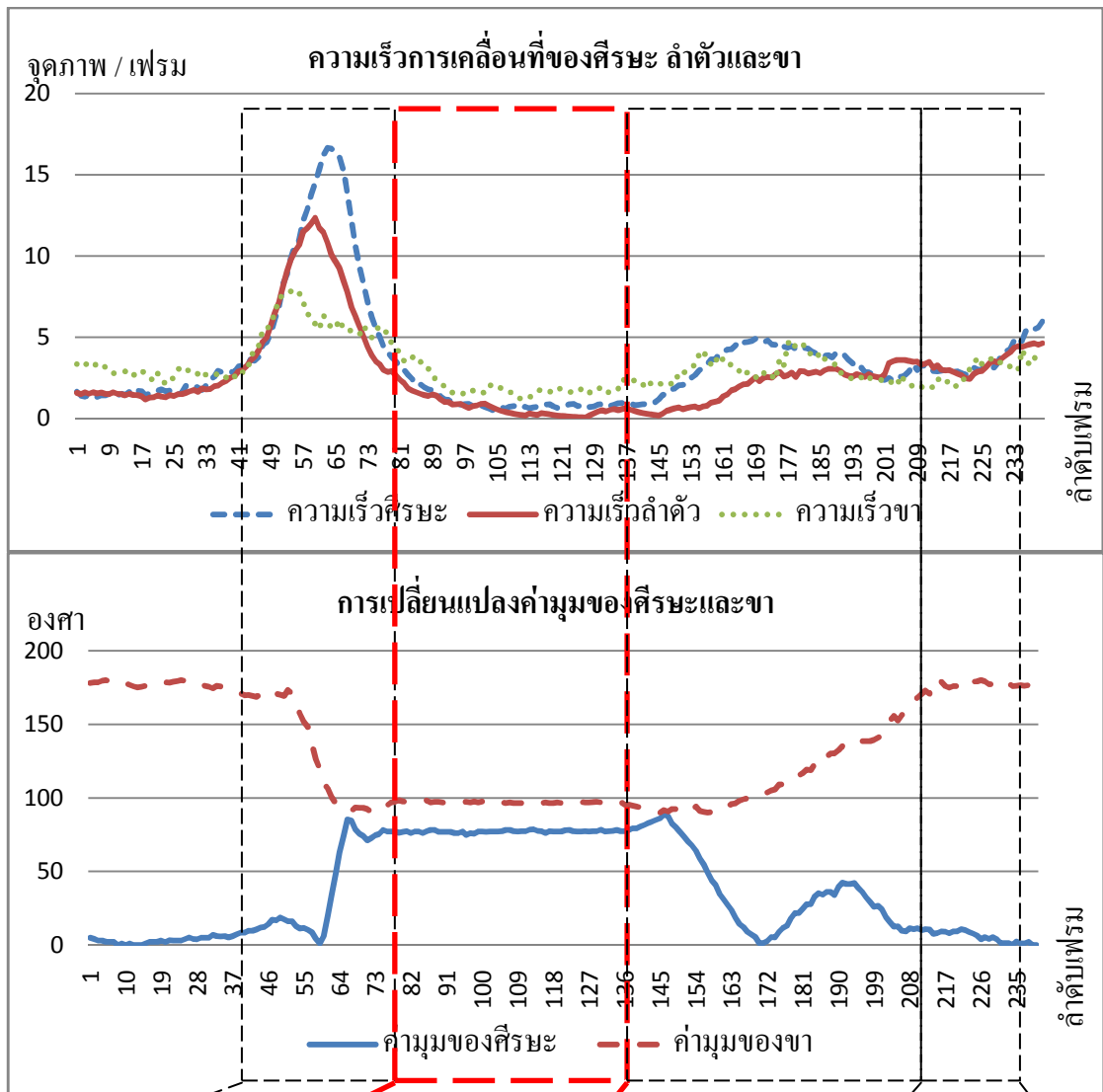


การนั่ง



ภาพประกอบ 4-4 ตัวอย่างผลกระบวนกรู้จำท่าทางการนั่ง

การนอนราบ



การขึ้น

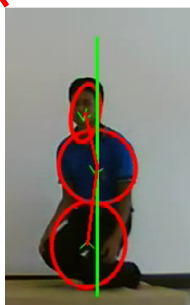
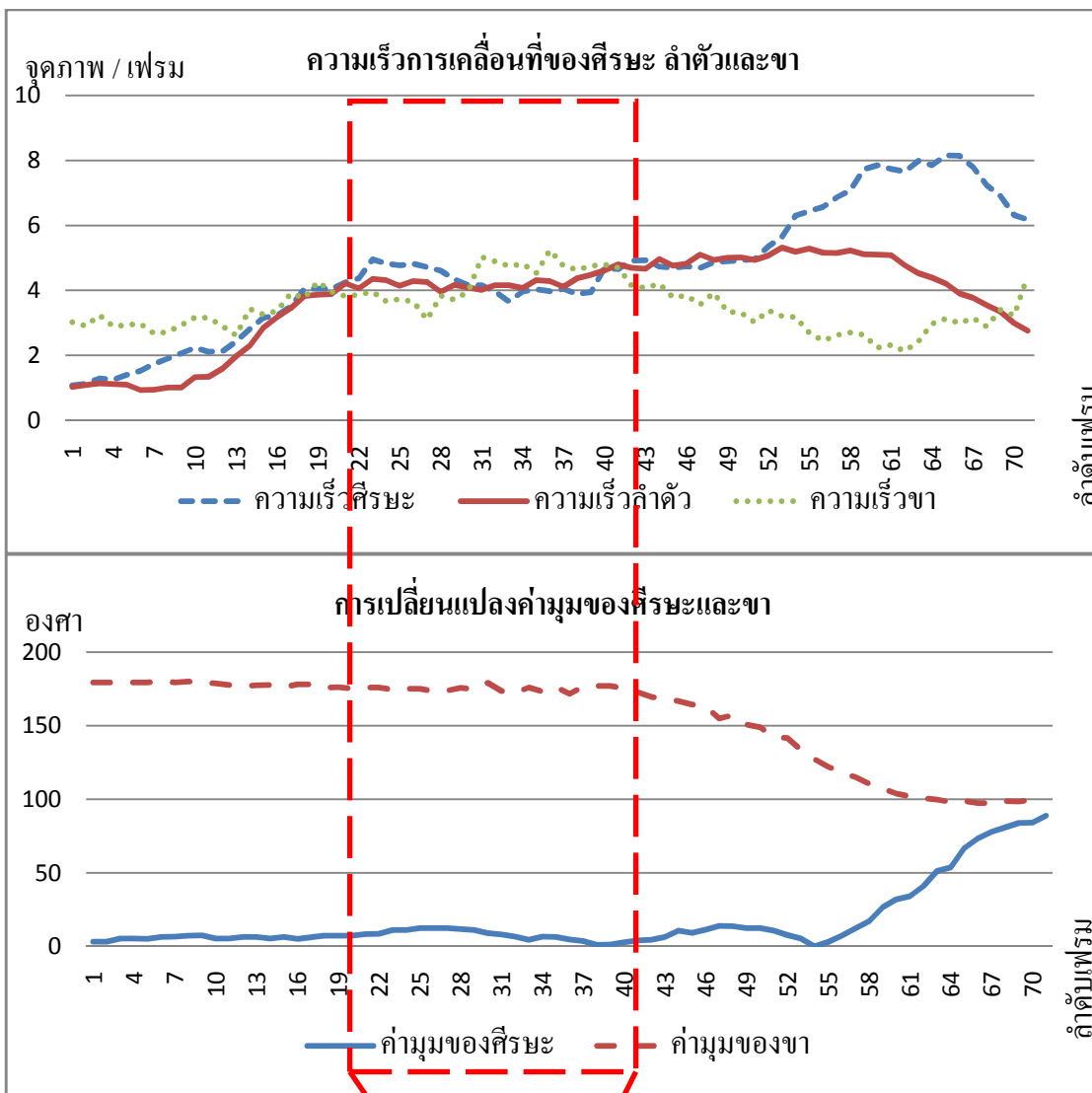
การนอนราบ

การนั่ง

N/A

ภาพประกอบ 4-5 ตัวอย่างผลกระบวนกรู้อำท่าทางการนอนราบ

N/A (Not Applicable)



N/A

ภาพประกอบ 4-6 ตัวอย่างท่าทางที่ไม่ถูกต้องแบบ (N/A)

ภาพประกอบ 4-6 เรายังสามารถสังเกตความผิดพลาดที่เกิดขึ้นในช่วงเวลาสั้น ๆ ขณะกรณีทำการนั่งคุกเข่าโดยหันหน้าเข้าหากล้องและเช่นเดียวกับการเปลี่ยนท่าทางนอนราบไปสู่ท่าทางยืน (ภาพประกอบ 4-4 ขวา) คือ การคุกเข่าและหันหน้าตรงเข้าหาหน้ากล้อง จะทำให้มุมของศีรษะและมุมของขายังคงมีลักษณะเช่นเดียวกันการการขึ้นและการเดินแต่มีความเร็วในทุก ๆ องค์ประกอบด้วยเช่นกันซึ่งท่าทางเช่นนี้มิได้ถูกออกแบบไว้อย่างชัดเจนในกระบวนการรู้จำท่าทางภายในงานวิจัยนี้ ดังนั้นท่าทางลักษณะนี้จะถูกจำแนกเป็นท่าทางที่ไม่สามารถนำไปปรับใช้ได้

ภาพประกอบ 4-2 ถึง ภาพประกอบ 4-5 แสดงท่าทางที่ถูกจำแนกออกมาโดยอาศัยแบบจำลองมนุษย์จากส่วนหนึ่งของวิดีโอที่ใช้ทดสอบท่าทางมนุษย์ ซึ่งประกอบด้วย การก้ม การนั่ง การนอนราบ การขึ้น และการเดิน ซึ่งการจำแนกท่าทางเหล่านี้ได้ใช้ คุณลักษณะท่าทางสถิต และท่าทางพลวัตที่นิยามในบทที่ 3 เพื่อการจำแนกท่าทาง จากผลการทดลองพบว่า สามารถรู้จำความสอดคล้องของท่าทางได้ถูกต้อง ดังแสดงในภาพประกอบ

ตาราง 4-1 ผลการทดลองการรู้จำท่าทาง

ท่าทาง	จำนวนเฟรม	ท่าทางที่ถูกตรวจพบ	อัตราการรู้จำ
การขึ้น	2886	2869	99.41 %
การนั่ง	1220	1089	89.26 %
การก้ม	2250	2123	94.35 %
การเดิน	2243	1809	80.65 %
การนอนราบ	1334	1334	100 %
N/A *	999	999	100 %
<b>รวม</b>	<b>10932</b>	<b>9224</b>	<b>93.95%</b>

\*N/A หมายถึง ท่าทางที่ไม่สามารถนำไปปรับใช้ได้

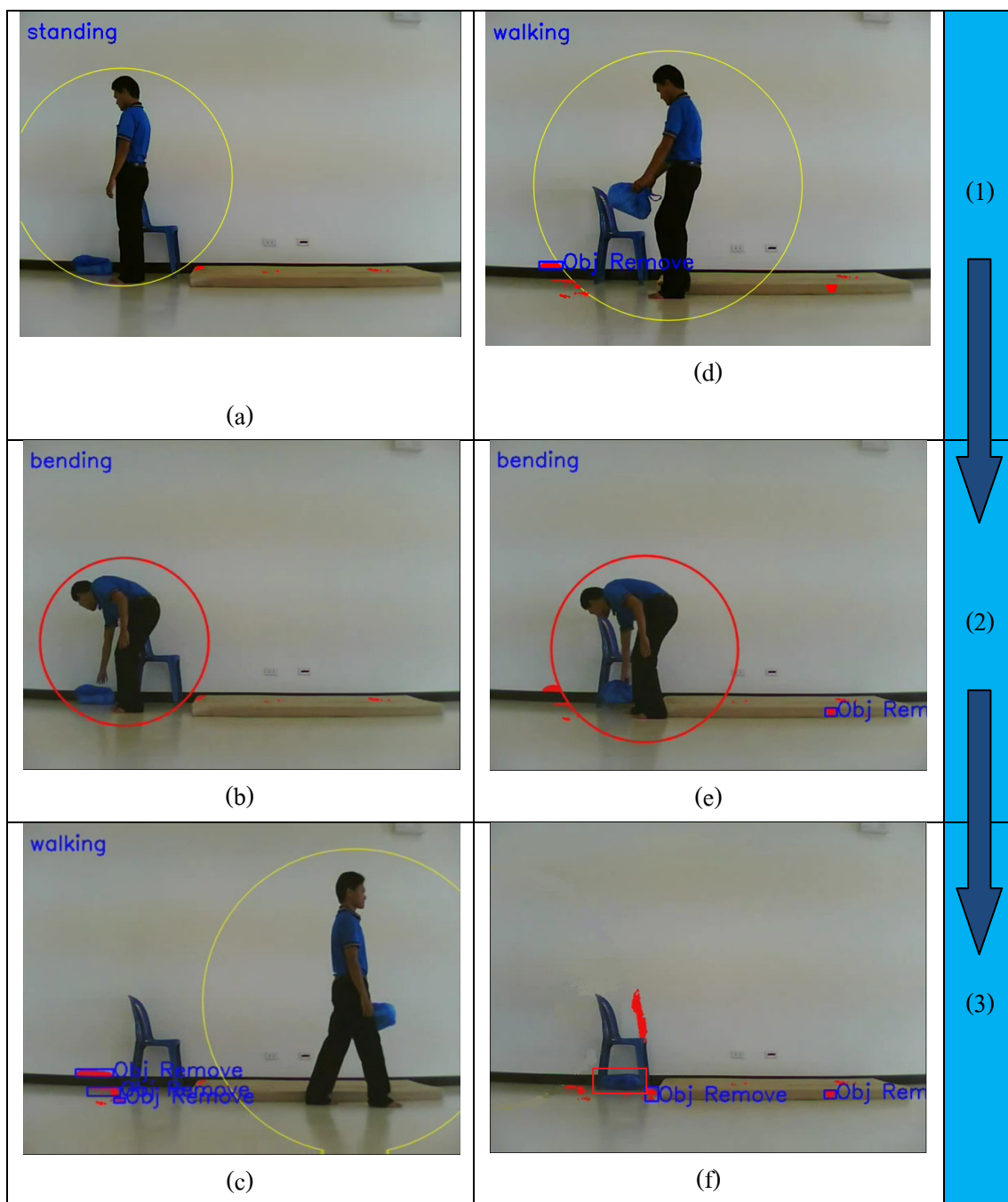
ตาราง 4-1 แสดงผลการทดลอง จำนวนเฟรมในแต่ละท่าทางแสดงจำนวนเปรียบเทียบระหว่างเฟรมทั้งหมดและเฟรมที่ระบุท่าทางได้ เรียงลำดับท่าทางที่สามารถรู้จำได้ถูกต้องมากที่สุดไปต่ำสุด ได้แก่ การนอนราบ การขึ้น การก้ม การนั่ง และการเดิน โดยในลำดับที่ต่ำที่สุดมีความถูกต้อง 80 เปอร์เซ็นต์ โดยความถูกต้องของการรู้จำท่าทางพื้นฐานโดยเฉลี่ยอยู่ที่ 93 เปอร์เซ็นต์

เราพบว่าในกรณีการเดินที่มีความถูกต้องต่ำที่สุด ความผิดพลาดเกิดขึ้นเมื่อขาทั้งสองข้างก้าวแยกออกจากกันทำให้การคำนวณศูนย์กลางผิดพลาด ซึ่งอาจสามารถแก้ไขได้โดยการปรับปรุงกระบวนการตรวจหาบุคคลและกระบวนการติดตาม กรณีทำทางการนอนราบวิธีการนี้สามารถระบุได้ถูกต้องทั้งหมดจากชุดข้อมูลที่ทดสอบ

#### 4.2. ผลการทดลองกระบวนการตรวจหาการเคลื่อนย้ายวัตถุ

การตรวจหาการเคลื่อนย้ายวัตถุเริ่มต้นด้วยกระบวนการตรวจหากิจกรรมเฝ้าระวังซึ่งภายในงานวิจัยนี้ใช้กรณีศึกษาเป็นกิจกรรมอย่างง่าย คือ การก้มหยิบหรือก้มวางวัตถุ เมื่อตรวจพบการก้มจึงใช้บริเวณที่เกิดกิจกรรมนั้นเป็นพื้นที่สำหรับตรวจหาการเคลื่อนย้ายวัตถุ ถ้าหากมีวัตถุใด ๆ ถูกนำเข้าไปหรือหยิบออกจากบริเวณพื้นที่นั้น ๆ วัตถุนั้นจะปรากฏเป็นฉากหน้า ซึ่งถ้าหากบุคคลนั้นเคลื่อนที่ออกห่างจากวัตถุที่ปรากฏในฉากหน้าเป็นเวลาประมาณ 3 วินาที บริเวณวัตถุนั้น ๆ จะถูกทำเครื่องหมายเป็นวัตถุโค่นเคลื่อนย้าย โดยการทดสอบจะใช้บุคคล 1 คนทำการก้มหยิบวัตถุและถือวัตถุนั้นนำไปวางในตำแหน่งอื่น

กิจกรรมที่ใช้เป็นเงื่อนไขการเฝ้าสังเกตการ คือ การก้ม เมื่อเริ่มมีการก้มกระบวนการตรวจหาการเคลื่อนย้ายจะเริ่มต้นทันที ภาพประกอบ 4-7(a) วงกลมสีเหลืองรอบตัวบุคคลแสดงถึงสถานะปรกติของบุคคลที่ไม่ได้กระทำท่าทางเฝ้าระวัง และเมื่อเกิดการก้ม ภาพประกอบ 4-7 (b) จะแสดงถึงการตรวจพบกิจกรรมเฝ้าระวัง โดยเปลี่ยนการแสดงผลพื้นที่กิจกรรมของวัตถุเป็นวงกลมสีแดงแทน เมื่อเกิดกิจกรรมเฝ้าระวังพื้นที่ในวงกลมสีแดงจะถูกนำไปใช้ในการพิจารณาการเคลื่อนย้ายวัตถุภายในพื้นที่นี้โดยอาศัยการตรวจสอบภาพพื้นหน้าจากกระบวนการลบพื้นหลัง ถ้าหากมีวัตถุใด ๆ ถูกเคลื่อนย้ายออกจากบริเวณที่เกิดกิจกรรมเฝ้าระวัง (ในการทดลองนี้ คือ การก้ม) วัตถุนั้น ๆ จะปรากฏอยู่เป็นฉากหน้าเป็นเวลานาน และเมื่อบุคคลเคลื่อนที่ออกจากบริเวณที่ประกอบกิจกรรมเฝ้าระวังแล้ว วัตถุที่มีเวลามากกว่าค่าขีดแบ่งจะถูกทำเครื่องหมายรอบวัตถุนั้น ๆ และแสดงข้อความกำกับแสดงให้เห็นได้ชัดว่าบริเวณนั้นเป็นบริเวณที่มีการเคลื่อนย้ายวัตถุดังแสดงในภาพประกอบ 4-7 (c)



ภาพประกอบ 4-7 ผลการทดลองกระบวนการตรวจหาการเคลื่อนย้ายวัตถุ

(1) เฝ้าระวังกิจกรรมต้องสงสัย (2) พบกิจกรรมต้องสงสัย

(3) ทำเครื่องหมายบริเวณวัตถุที่ถูกเคลื่อนย้าย

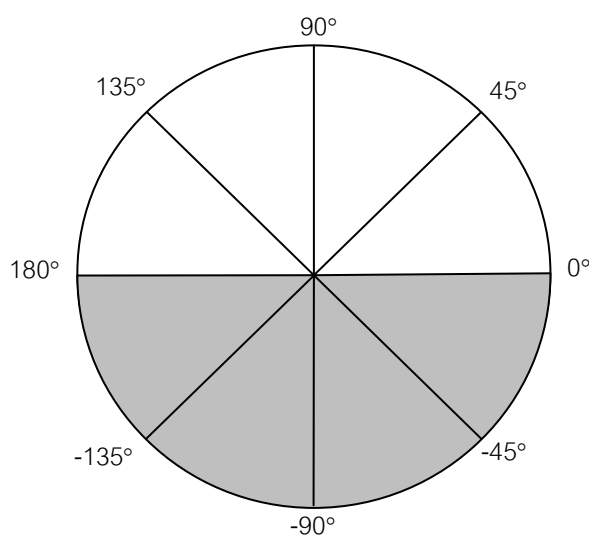
ความผิดพลาดที่ตรวจพบ เนื่องจากการตรวจจับวัตถุเคลื่อนย้ายใช้ผลจากกระบวนการลบพื้นหลังดังนั้นความถูกต้องของกระบวนการลบพื้นหลังจึงมีผลอย่างมากต่อการตรวจหาวัตถุที่ถูกเคลื่อนย้าย ดังกรณีตัวอย่างภาพประกอบ 4-7 (e) บุคคลได้กระทำการก้มวางของไว้ในตำแหน่ง

บริเวณใต้เก้าอี้ แต่เนื่องจากสีของวัตถุที่วางไว้มีสีใกล้เคียงกับสีของพื้นหลังบริเวณนั้นจึงทำให้วัตถุที่ถูกนำมาวางไว้ไม่ถูกตรวจพบดังแสดงในพื้นที่สี่เหลี่ยมสีแดงในภาพประกอบ 4-7 (f) และความผิดพลาดอีกกรณีหนึ่ง คือ บริเวณขอบเบาะที่ถูกตรวจพบเป็นวัตถุที่ถูกเคลื่อนย้าย เป็นผลมาจากขณะที่กระทำการทดสอบได้มีการเดินเหยียบไปบนเบาะทำให้บริเวณของเบาะเกิดรอยพับซึ่งทำให้เกิดเป็นเงาเพราะแหล่งกำเนิดแสงอยู่ด้านบน การเกิดเงาที่ขอบเบาะเช่นนี้เมื่อผ่านกระบวนการลบพื้นหลังจะพบว่าบริเวณเงานี้มีความต่างเป็นอย่างมากจึงถูกจำแนกเป็นพื้นหน้า ทำให้กระบวนการเคลื่อนย้ายวัตถุตรวจพบบริเวณนี้เป็นวัตถุที่ถูกเคลื่อนย้าย

#### 4.3. กระบวนการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่

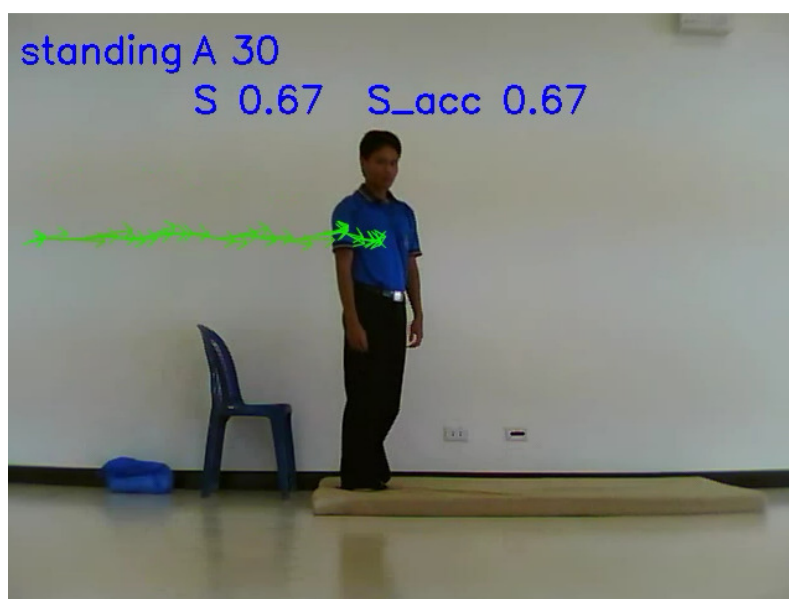
ทำการทดลองภายใต้เงื่อนไขควบคุมลักษณะการเคลื่อนที่ในทิศลงอย่างรวดเร็วดังนี้

- พื้นที่ฝ้าสังเกต พิจารณาพื้นที่ฝ้าสังเกตทั้งภาพ
- ท่าทาง พิจารณาท่าทางการเดิน นั่ง ก้ม และนอนราบ
- ความเร็ว พิจารณาความเร็วสูงกว่า 10 จุดภาพต่อเฟรม (250 จุดภาพต่อวินาที)
- ทิศทาง พิจารณาเฉพาะทิศการเคลื่อนที่ลงครึ่งส่วนที่แรกในภาพประกอบ 4-8
- ความเร่ง พิจารณาความเร่งสูงกว่า 10 จุดภาพต่อวินาที<sup>2</sup>

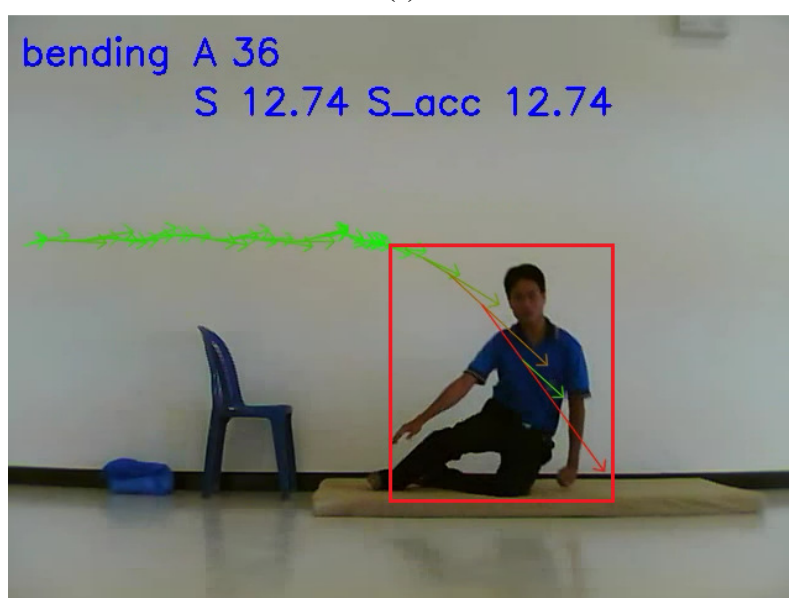


ภาพประกอบ 4-8 ทิศทางควบคุม (ภายในพื้นที่เร่ง)

แผนผังทิศทางการเคลื่อนที่จะแสดงซ้อนทับกับภาพ โดยแสดงเป็นลูกศร ทิศของหัวลูกศร แสดงถึงการทิศทางการเคลื่อนที่ของบุคคล ณ เวลานั้น ๆ ความยาวของลูกศรแสดงถึงความเร็วของวัตถุ ยิ่งมีความเร็วมากลูกศรยิ่งมีความยาวมาก และสีลูกศรจะแปรผันตามความเร็วโดยถ้าเคลื่อนที่ช้าจะแสดงเป็นสีเขียว ยิ่งเคลื่อนที่เร็วขึ้นสีจะเปลี่ยนเข้าสู่สีแดงมากขึ้น ดังแสดงในภาพประกอบ 4-9 (a) และ (b)



(a)



(b)

ภาพประกอบ 4-9 ผลการทดลองควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่



จากแผนผังทิศทางการเคลื่อนที่เราสามารถใช้อธิบายกิจกรรมการเคลื่อนที่ของบุคคลในลักษณะที่สัมพันธ์กับตำแหน่ง ทิศทาง ความเร็วและความเร่งโดยทุก ๆ เฟรมที่มีการปรับปรุงข้อมูลใหม่ ข้อมูลเหล่านั้นจะถูกทำไปคำนวณเป็นสถานีโยงของแผนผังทิศทางการเคลื่อนที่และถูกนำไปทดสอบกับเงื่อนไขควบคุม ถ้าหากกิจกรรมที่เกิดขึ้นตรงกับเงื่อนไขควบคุมทั้งหมดบริเวณบุคคล ณ เวลาที่มีกิจกรรมตรงกับเงื่อนไขควบคุมจะถูกทำเครื่องหมายล้อมรอบดังภาพประกอบ 4-9 (b)

#### 4.4. สรุป

เนื้อหาในบทนี้ได้กล่าวถึงผลการทดสอบกระบวนการต่าง ๆ ที่ถูกอธิบายไว้ในเนื้อหาบทที่ 3 โดยจากผลการทดสอบกระบวนการรู้จำท่าทางมนุษย์พบว่ามีความถูกต้องโดยเฉลี่ยประมาณ 93% การตรวจหาการเคลื่อนย้ายวัตถุนั้นจะให้ผลที่ถูกต้องสูงเมื่อของกระบวนการลบพื้นหลังมีความถูกต้อง ส่วนผลการทดลองกระบวนการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่สามารถที่จะตรวจพบเหตุการณ์ควบคุมได้อย่างถูกต้องเมื่อกระบวนการรู้จำท่าทางมีความถูกต้อง

## บทที่ 5

### สรุปผลการวิจัยและข้อเสนอแนะ

#### 5.1. สรุปผล

ภายในงานวิจัยนี้ได้พัฒนาเทคนิคการติดตามวัตถุเคลื่อนที่หลายวัตถุ ในลักษณะการติดตามโครงสร้างอย่างง่ายของมนุษย์ ซึ่งผลจากการติดตามทำให้ได้ข้อมูลกิจกรรมการเคลื่อนที่ในลักษณะความสัมพันธ์เป็น Space-Time-Behavior Representation คือ ความเร็ว ความเร่ง รวมถึง ข้อมูลทิศทางเคลื่อนที่ ความเร่งเชิงมุมและข้อมูลความสัมพันธ์เชิงมุมของแต่ละองค์ประกอบ ซึ่งจากข้อมูลทั้งหมดดังกล่าวข้างต้น สามารถใช้ในกระบวนการรู้จำท่าทางพื้นฐานของมนุษย์ได้และจากท่าทางพื้นฐานเหล่านี้ สามารถจะนำมาผสมผสานกันเพื่อใช้อธิบายถึงกิจกรรมที่มีความซับซ้อนมากขึ้นได้ ซึ่งกิจกรรมเหล่านี้สามารถถูกใช้เพื่อเป็นเงื่อนไขสำหรับการตรวจหาเหตุการณ์ที่สนใจภายในงานวิจัยนี้ได้ทดสอบกรณีศึกษา 2 กรณี คือ (1) การตรวจหาการเคลื่อนย้ายวัตถุ และ (2) การควบคุมกิจกรรมที่เกี่ยวข้องกับทิศทางเคลื่อนที่ โดยสรุปผลการทดลองมีดังนี้

##### 5.1.1. ผลการทดลองการรู้จำท่าทางมนุษย์

ภายใต้สภาวะแวดล้อมภายในอาคาร โดยใช้กล้องเพียงตัวเดียวและกล้องไม่เคลื่อนที่และเนื่องจากภายในงานวิจัยนี้เน้นทดสอบกระบวนการรู้จำท่าทางและการรู้จำพฤติกรรมมนุษย์ดังนั้นจึงเลือกชุดข้อมูลภาพที่มีความผิดพลาดจากข้อจำกัดของกระบวนการลบพื้นหลังน้อยที่สุด ผลการรู้จำท่าทางพื้นฐานมีความถูกต้องโดยเฉลี่ยอยู่ที่ 93 เปอร์เซ็นต์ โดยท่าทางที่มีความถูกต้องต่ำที่สุดคือ การเดินซึ่งมีความถูกต้องประมาณ 80 เปอร์เซ็นต์ สาเหตุเกิดจาก (1) จังหวะที่มีการก้าวเท้าบริเวณขาจะแยกออกเป็นสองส่วนแต่แบบจำลองมนุษย์อย่างง่ายได้ออกแบบส่วนขาไว้เพียงส่วนเดียวส่งผลการกระบวนการติดตามอาจจะบุตำแหน่งส่วนขาเบี่ยงไปทางด้านใดด้านหนึ่งทำให้ได้ข้อมูลการติดตามที่ไม่ถูกต้องจึงทำให้กระบวนการจำแนกท่าทางมีความผิดพลาด และท่าทางที่มีความถูกต้องมากที่สุด คือ การนอนราบเพราะคุณลักษณะการนอนราบมีลักษณะที่สามารถจำแนกได้อย่างชัดเจนจึงทำให้ผลการจำแนกมีความถูกต้องสูง (2) ขณะที่เปลี่ยนจากท่าทางหนึ่งไปสู่อีกท่าทางหนึ่งผลจากการติดตามจะให้ค่าที่มีการแกว่งไปมาเล็กน้อยระหว่างค่าที่เป็นจุดแบ่งระหว่างท่าทางดังนั้นจึงทำให้เกิดค่าผิดพลาดขึ้นเล็กน้อยในช่วงการเปลี่ยนจากท่าทางหนึ่งไปสู่อีกท่าทางหนึ่ง

อย่างไรก็ตาม การใช้วิธีรู้จำท่าทางในงานวิจัยนี้ ความผิดพลาดของการรู้จำสามารถเกิดขึ้นในกรณีที่หน้ากล้องตั้งฉากกับท่าทาง ซึ่งจะทำให้พารามิเตอร์ภายในของแบบจำลองมนุษย์ไม่สามารถใช้เพื่ออธิบายท่าทางได้

### 5.1.2. ผลการทดลองทดลองกระบวนการตรวจหาการเคลื่อนย้ายวัตถุ

กระบวนการตรวจหาการเคลื่อนย้ายวัตถุ ถูกทดสอบเบื้องต้น โดยใช้กิจกรรมอย่างง่ายที่ไม่มีความซับซ้อน คือ การก้มเป็นกิจกรรมเฝ้าสังเกต สามารถตรวจจับการเคลื่อนย้ายวัตถุเมื่อเกิดกิจกรรมเฝ้าระวังได้อย่างถูกต้อง แต่ความถูกต้องค่อนข้างมีความสัมพันธ์กับผลจากกระบวนการลบพื้นหลัง เนื่องจากกระบวนการนี้ใช้ผลโดยตรงจากกระบวนการลบพื้นหลัง ดังนั้นหากใช้ผลตั้งต้นจากกระบวนการลบพื้นหลังที่ผิดพลาดย่อมให้ผลการตรวจหาที่ผิดพลาดเช่นกัน

### 5.1.3. ผลการทดลองกระบวนการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่

สามารถอธิบายกิจกรรมบุคคลในลักษณะที่สัมพันธ์กับตำแหน่ง ทิศทาง ความเร็วและความเร่งได้ และสามารถสร้างเงื่อนไขควบคุมที่สัมพันธ์ลักษณะดังกล่าวเพื่อคอยเฝ้าระวังและแจ้งเตือนในลักษณะการทำสัญลักษณ์ในภาพเมื่อพบกิจกรรมที่ต้องการควบคุม

## 5.2. บทวิจารณ์

### 5.2.1. บทวิจารณ์ผลการทดลองการรู้จำท่าทางมนุษย์

ความผิดพลาดที่เกิดจากข้อจำกัดของกระบวนการลบพื้นหลัง แยกเป็น 2 กรณี ดังนี้ (1) กรณีวัตถุมีสีใกล้เคียงกับพื้นหลัง สำหรับกรณีที่สีของวัตถุมีความใกล้เคียงกับพื้นหลังจะทำให้วัตถุบริเวณนั้น ๆ ถูกจำแนกเป็นพื้นหลังไปด้วยทำให้บริเวณของวัตถุบางส่วนหายไปซึ่งถ้าหากเป็นพื้นที่เพียงเล็กน้อยกระบวนการติดตามโดยใช้ขั้นตอนวิธีการย้ายเข้าสู่ค่ากลางโดยการปรับตัวอย่างต่อเนื่องจะยังคงให้ผลการติดตามที่ไม่เกิดความผิดพลาดมากนักเนื่องจากกระบวนการนี้ใช้พื้นฐานทางสถิติโดยการหาตำแหน่งค่ากลาง ถ้าหากจุดข้อมูลหายไปเพียงเล็กน้อยเมื่อคำนวณค่ากลางย่อมได้ค่าที่มีความผิดพลาดเพียงน้อยเท่านั้น แต่ถ้าหากพื้นที่บริเวณวัตถุมีสีคล้ายกับพื้นหลังเป็นบริเวณกว้างจะทำให้ผลการหาตำแหน่งของวัตถุผิดไปอย่างมาก ซึ่งท้ายที่สุดจะส่งผลไปยังการจำแนกท่าทางที่ผิดพลาด (2) เงา ผลจากการเกิดเงาของวัตถุเมื่อผ่านกระบวนการลบพื้นหลัง เงาที่มีความเข้มสูงจะมีค่ามากกว่าค่าขีดแบ่งในกระบวนการลบพื้นหลังจึงทำให้บริเวณเงานั้น ๆ ถูกตรวจพบเป็นส่วนหนึ่งของวัตถุไปด้วย ซึ่งทำให้บริเวณของวัตถุผิดไปจากความเป็นจริงหรืออาจถูกตรวจ

พบเป็นวัตถุชิ้นใหม่ (ขึ้นอยู่กับตำแหน่งการเกิดเงา) ลักษณะเช่นนี้ ส่งผลการกระบวนการติดตาม ผิดพลาดและท้ายที่สุดย่อมทำการจำแนกท่าทางผิดพลาดไปด้วย ดังนั้นการพัฒนาระบบการลบพื้นหลังให้มีประสิทธิภาพดีขึ้นย่อมมีความสำคัญกับการพัฒนาระบบรู้จำท่าทางและรู้จำกิจกรรมต่อไป

สำหรับการจัดการกรณีหน้ากล้องตั้งฉากกับท่าทาง อาจจำเป็นต้องใช้พารามิเตอร์ที่มีความซับซ้อนอื่น ๆ หรือ ใช้กล้องมากขึ้นในกระบวนการรู้จำพฤติกรรม เช่น การรู้จำท่าทางโดยใช้หลายกล้อง ดังนั้นระบบดังกล่าวอาจนำไปสู่วิธีการในการแก้ปัญหาท่าทางที่ไม่สามารถอธิบายได้ (N/A) ในงานวิจัยนี้

### 5.2.2. บทวิจารณ์ผลการทดลองกระบวนการตรวจหาการเคลื่อนย้ายวัตถุ

สำหรับความถูกต้องของกระบวนการตรวจหาการเคลื่อนย้ายวัตถุในงานวิจัยนี้ มีความสัมพันธ์ใกล้ชิดกับกระบวนการลบพื้นหลัง ดังนั้นการปรับปรุงประสิทธิภาพของกระบวนการนี้อาจทำได้ โดยการพัฒนาระบบการลบพื้นหลัง และอาจประยุกต์ใช้กระบวนการรู้จำรูปแบบเพิ่มเติมเพื่อช่วยเพิ่มประสิทธิภาพการเฝ้าสังเกตการเคลื่อนย้ายวัตถุเฉพาะบางอย่างให้มีความแม่นยำมากยิ่งขึ้น

### 5.2.3. บทวิจารณ์ผลการทดลองกระบวนการควบคุมที่เกี่ยวข้องกับทิศทางการเคลื่อนที่

เงื่อนไขที่ใช้ทดสอบภายในงานวิจัยนี้ถูกกำหนดเป็นค่าคงที่ทั้งสิ้นซึ่งอาจไม่เหมาะกับการตอบสนองกับบางกิจกรรมน่าสงสัยได้ การสร้างเงื่อนไขควบคุมที่เหมาะสมสามารถปรับปรุงได้ โดยอาศัยการเก็บสถิติข้อมูลกิจกรรมปกติในทุก ๆ ส่วน คือ ท่าทาง ทิศทาง ความเร็วและความเร่งของวัตถุที่เคลื่อนที่ภายในภาพวิดีโอและใช้ข้อมูลทางสถิติเหล่านั้นในการออกแบบเงื่อนไขของกิจกรรมที่ต้องการเฝ้าระวัง

## บรรณานุกรม

- [1] R. C. Gonzalez, R. E. Woods, *Digital Image Processing*. Prentice Hall Inc., 2002.
- [2] G. Bradski and A. Kaebler, *Learning OpenCV Computer Vision with the OpenCV Library*. O'Reilly Media Inc., 2008.
- [3] S. Suzuki and K. Abe, "Topological structural analysis of digital binary images by border following," *Computer Vision, Graphics and Image Processing* 30, (1985): 32–46.
- [4] Lee C.K., Ho M.F., We W.S. and Huang C.L., "Abnormal event detection in video using N-cut clustering," in *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2006, pp.407-410.
- [5] Andrade E. L., Blunsden S.J. and Fisher R. B., "Detection of emergency events in crowded scenes," in *The IEEE international symposium on imaging for crime prevention and detection*, 2006, pp. 528-533.
- [6] Cucchiara R., Grana C., Piccardi M. and Prati A., "Statistic and knowledge-based moving object detection in traffic scenes," in *Proceeding of Intelligent Transportation Systems*, 2000, pp. 27 – 32.
- [7] I. Pavlidis, V. Morellas, P. Tsiamyrtzis and S. Harp, "Urban surveillance systems: from the laboratory to the commercial world," in *Proceedings of the IEEE*, vol. 89, no. 10, 2001, pp. 1478 -1497.
- [8] Wren C.R., Azarbayejani A., Darrell T. and Pentland A.P., "Pattern Analysis and Machine Intelligence," in *IEEE Transactions*, vol. 19 , Issue: 7, 1997, pp. 780 – 785.
- [9] Stauffer C. and Grimson W.E.L., "Computer Vision and Pattern Recognition," in *IEEE Computer Society Conference*, vol. 2, 1999.
- [10] Y. Cheng, "Mean shift mode seeking and clustering," in *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 1998, pp. 790-799.
- [11] D. Comaniciu, V. Ramesh and P. Meer., "Real-time tracking of non-rigid objects using mean shift," In *Proceeding of IEEE Computer Vision and Pattern Recognition*, 2000, pp. 673-678.

- [12] K. Fukunaga and L. D. Hostetler, "The estimation of the gradient of a density function with applications in pattern recognition," in *IEEE Transaction on Information Theory*, 1975, pp. 21-32.
- [13] Han C.C., Lin C.Y., Ho G.F. and Fan K.C., "Abnormal event detection using trajectory features," in *International Computer Symposium on Image Processing*, department of computer science and information engineering, Yuan-Ze University, 2006.
- [14] Jiang F., Wu Y. and Katsaggelos A.K., Abnormal event detection from surveillance video by dynamic hierarchical clustering, in *IEEE International Conference on Image Processing*, vol. 5, 2007, pp. V - 145-V – 148.
- [15] Xiang T. and Gong S., "Video behaviour profiling and abnormality detection without manual labeling," in *IEEE International Conference on Computer Vision*, vol. 2, 2005, pp. 1238-1245.
- [16] Blank M., Gorelick L., Shechtman E., Irani, M. and Basri R., "Actions as Space-Time Shapes," in *IEEE International Conference*, vol. 2, 2005, pp. 1395 - 1402.

**ภาคผนวก**

ภาคผนวก ก.

**ผลงานตีพิมพ์เผยแพร่จากวิทยานิพนธ์**



**PROCEEDINGS OF SPIE**

**Second International Conference  
On Digital Image Processing**

**Kamaruzaman Jusoff**

**Yi Xie**

**Editors**

**26-28 February 2010  
Singapore**

**Organized by  
International Association of Computer Science and Information Technology (Singapore)**

**Volume 7546  
Part One of Two Parts**



**SPIE**

Connecting minds. Advancing light

# Model-based Human Action Recognition

Nattapon Noorit, Nikom Suvonvorn, and Montri Karnchanadecha  
 Department of Computer Engineering  
 Faculty of Engineering, Prince of Songkla University  
 Hat Yai, Thailand  
 pinggtar@hotmail.com, kom@coe.psu.ac.th, montri@coe.psu.ac.th

## 1. ABSTRACT

The identification of human basic actions plays an important role for recognizing human activities in complex scene. In this paper we propose an approach for automatic human action recognition. The parametric model of human is extracted from image sequences using motion/texture based human detection and tracking. Action features from its model are carefully defined into the action interaction representation and used for the recognizing process. Performance of proposed method is tested experimentally using datasets under indoor environments.

**Keywords:** human action recognition, activity recognition, human modeling, video surveillance

## 1. INTRODUCTION

Recognizing human activities from video sequence is one of the most challenging problems of surveillance application. In this paper, we propose the method for recognizing the basic actions of human activities, which is necessary for the event of interest detection, for example abnormal or rare events. Here, “action” is referred to simple motion patterns normally executed by a single person, such as walking, standing, laying, bending and sitting, and “activity” refers to the complex sequence of actions performed by several humans.

A large number of publications work for action recognition<sup>1</sup>. Efron et al.<sup>2</sup> attempt to recognize a set of simple actions (walking, running with direction and location) using a set of features that are based on blurred optic flow. Robertson and Reid [3] propose an approach where complex actions can be dynamically composed out of the set of simple actions by building a hierarchical system that is based on reasoning with belief networks and HMMs. Yilmaz and Shah<sup>4</sup> extract information such as speed, direction and shape by analyzing the differential geometric properties of space-time relation. <sup>5,6</sup>also analyze the space-time volume. Zelnik-Manor et al. <sup>8</sup>define dynamic actions as long-term temporal objects at multiple temporal scales features. Bradski et al. <sup>7</sup>develop MHI for motion segmentation that allow determination of the optical flow.

The rest of the paper is organized as follows. Section 2 describes our parametric human model. Section 3 depicts how to represent the actions for the recognition process. The experimentation and conclusion are discussed in the section 4 and 5 respectively.

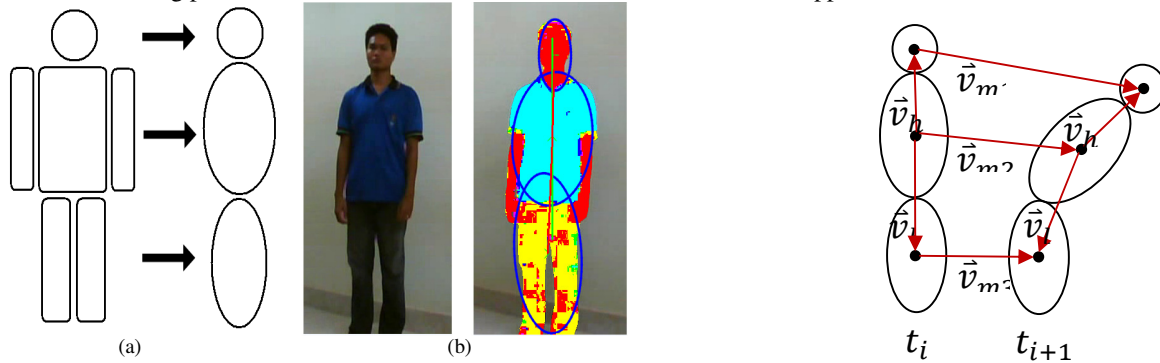
## 2. HUMAN MODEL EXTRACTION

In this section, we describe how to determine the parametric model of human from the images sequence. The overall process can be divided into two major steps: motion detection and tracking and human model construction with parameter estimation.

### 2. Motion-texture based Human Detection and Tracking

Moving object is separated from background by using background subtraction technique which small noises are removed by morphological opening and closing filters. Generally, an object might be detected in several fragmented image regions. In that case, a region-fusion operation is needed. Two regions are considered to be the same object if they are overlapped or their distance less than a specific threshold. With these constraints, the method is again very sensible to light condition, such as shadow, contrast changing and sudden changes of brightness. Intuitively, introducing some special characteristics of object, for instance texture properties, will probably improve the better results. Therefore, in the fusion process the color probability density of object’s texture is additionally applied for computing the similarity between regions using Mean-shift algorithm<sup>9</sup>. This mixture of motion and texture of object for detection and tracking can reduce significantly noises and increases consequently the effectiveness of our tracking algorithm. However, there are always additive noises superposed with detected objects that will be eliminated later by human model constraints.

The extracted foreground that supposed to be a human is then segmented into three regions representing the three important parts of human structure, for instance, head, body, and legs. Noticed that both of hands are denied and only one component is used for representing the both legs. We assume that with only three components of human model the five basic actions could be indentified correctly. Standing action is used for initiating the human model construction which the relations between three parts are specified explicitly; body region is bigger than others and located at the middle, the upper and lower connected regions are assumed to be the head and legs parts respectively. In this step, we have to deal with errors emitted from motion detection and tracking process. The human model constraints are used for noise suppression.



Keywords: (a) Simplification of human structure (b) Example of reconstructed human model from image sequences.

Keywords: Humain model with parameters.

### 3. Parametric Model Definition

Our human model for action recognition is presented in the figure 2. The parameters of model are considered into two groups: internal and external parameters. The internal parameters  $[\vec{v}_h, \vec{v}_l]$  represent the characteristic of human structure, which describes the distances  $[\vec{v}_h^s, \vec{v}_l^s]$  and directions  $[\vec{v}_h^\theta, \vec{v}_l^\theta]$  from body component to the head and legs components respectively. The external parameters  $[\vec{v}_{m1}, \vec{v}_{m2}, \vec{v}_{m3}]$  correspond to the properties of human movement from frame to another. More precisely, it characterizes the velocity  $[\vec{v}_{m1}^s, \vec{v}_{m2}^s, \vec{v}_{m3}^s]$  and direction  $[\vec{v}_{m1}^\alpha, \vec{v}_{m2}^\alpha, \vec{v}_{m3}^\alpha]$  of the head, body, and legs components respectively.

Note that we define our parameters relatively up to the centroid of human components in order to reduce noises produced during the low-level process, for example motion detection, segmentation, tracking and etc.

### 3. ACTION REPRESENTATION

In this section, we explain how to define the basic actions. Experimentally, we found that almost activities can be decomposed into the five basic actions: standing, bending, sitting, walking and laying. So, an activity can then be described by a sequence of these actions with some transition parameters mostly depending on motion vectors. Figure 3 shows the actions represented by our simplified human model.

To recognize the actions, we establish the features of each action from the parameters of human model by the following

$$\left[ \frac{\vec{v}_h^\theta}{\vec{v}_l^\theta}, \frac{\vec{v}_h^s}{\vec{v}_l^s}, \frac{\partial \vec{v}_{m1}^s}{\partial t}, \frac{\partial \vec{v}_{m2}^s}{\partial t}, \frac{\partial \vec{v}_{m3}^s}{\partial t}, \frac{\partial^2 \vec{v}_{m1}^\alpha}{\partial t^2}, \frac{\partial^2 \vec{v}_{m2}^\alpha}{\partial t^2}, \frac{\partial^2 \vec{v}_{m3}^\alpha}{\partial t^2} \right]:$$

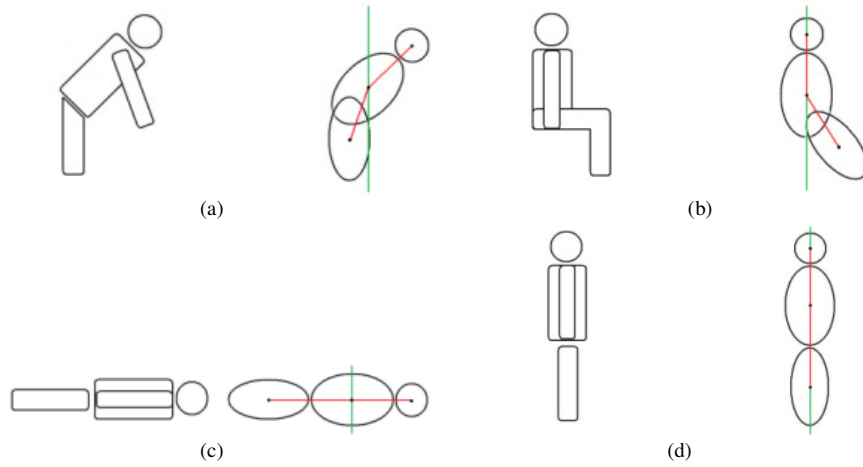
$\frac{\vec{v}_h^\theta}{\vec{v}_l^\theta}$  the angle ratio between head and legs with respect to the body.

$\frac{\vec{v}_h^s}{\vec{v}_l^s}$  the distance ratio between head and legs with respect to the body.

$\frac{\partial \vec{v}_{m1}^s}{\partial t}, \frac{\partial \vec{v}_{m2}^s}{\partial t}, \frac{\partial \vec{v}_{m3}^s}{\partial t}$  the velocity of each human components.

$\frac{\partial^2 \vec{v}_{m1}^\alpha}{\partial t^2}, \frac{\partial^2 \vec{v}_{m2}^\alpha}{\partial t^2}, \frac{\partial^2 \vec{v}_{m3}^\alpha}{\partial t^2}$  the angular acceleration of each human components.

Our assumption, by combining the above features we could identify the different actions correctly. The figure 3 simply shows the complete basic actions with its features. We classify actions into two types: static and dynamic actions.



Keywords: Basic actions (a) Bending (b) Sitting (c) Laying (d) Standing or Walking.

#### 4. Static actions

The actions are considered static if only if there are at least one component which the velocity is null. By definition, the static actions are comprised of standing, bending and sitting. Consequently, Its features are then combined by only from the internal parameters  $[\vec{v}_h, \vec{v}_l]$ , that can be used later for action identification or action discrimination.

##### 1. Standing features

In general, the standing action will provide  $\vec{v}_h$  and  $\vec{v}_l$  that tend to be paralleled with vertical axis, independently from the camera view of point. And certainly, the velocities of very components are near to zero. During the transition state, for example from standing to sitting or to bending, the angles of  $\vec{v}_h$  and  $\vec{v}_l$  increase while inversely decreasing for the opposite transition.

$$\vec{v}_h^\theta \cong 0, \vec{v}_l^\theta \cong 180, \frac{\partial \vec{v}_{m1}^s}{\partial t} \cong \frac{\partial \vec{v}_{m2}^s}{\partial t} \cong \frac{\partial \vec{v}_{m3}^s}{\partial t} \cong 0 \quad (1)$$

$$\frac{\partial \vec{v}_l^\theta}{\partial t} > 0 \quad (2), \quad \frac{\partial \vec{v}_h^\theta}{\partial t} > 0 \quad (3)$$

##### 2. Bending features

For the bending case, the head component is blending down while the body and legs components are still fixed. Considering the parameters model, the vector  $\vec{v}_l$  is likely to be positioned vertically. During the transition state from bending to standing the angle of  $\vec{v}_h$  will decrease.

$$\vec{v}_l^\theta \cong 180, \frac{\partial \vec{v}_{m2}^s}{\partial t} \cong \frac{\partial \vec{v}_{m3}^s}{\partial t} \cong 0 \quad (4)$$

$$\frac{\partial \vec{v}_h^\theta}{\partial t} < 0 \quad (5)$$

##### 3. Sitting features

In the sitting situation, the characteristic of features is opposite to the bending action. The vector  $\vec{v}_h$  tends to be paralleled with y-axis and the legs component does not move. During the transition state from sitting to standing the angle of  $\vec{v}_l$  will decrease.

$$\vec{v}_h^\theta \cong 0, \frac{\partial \vec{v}_{m3}^s}{\partial t} \cong 0 \quad (6)$$

$$\frac{\partial \vec{v}_l^\theta}{\partial t} < 0 \quad (7)$$

## 5. Dynamic actions

The actions are considered dynamic if only if all components of human model move. We found that only the internal parameters of human model cannot discriminate or identify the action itself. It needs to consider additionally the external parameters as transition factors.

### Walking features

In case of walking action, every part of human move generally and approximately in the same direction and speed. Therefore, the walking action can then be identified by the standing action features with motion transitions of all components. So, their velocities are superior to zero.

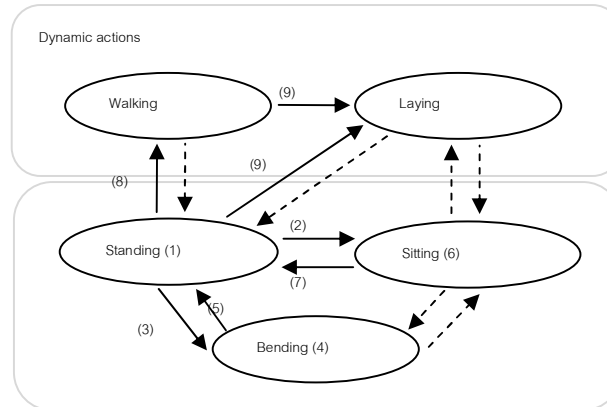
$$\frac{\partial \vec{v}_{m1}^s}{\partial t} \cong \frac{\partial \vec{v}_{m2}^s}{\partial t} \cong \frac{\partial \vec{v}_{m3}^s}{\partial t} > 0 \quad (8)$$

### Laying features

For the laying action, every parts of human body move with different directions and speeds. In general, head component is faster than body and legs respectively. We found that with some specific camera view point the  $\vec{v}_h$  and  $\vec{v}_l$  are lay on the same line, which is recognizing as standing action. Thus, we cannot use only internal parameters for identifying laying action. We define then the laying action with the standing action and motion transition features. Generally, the acceleration of head is more than body and legs components. And certainly the acceleration of body components is bigger than the legs one.

$$\frac{\partial^2 \vec{v}_{m1}^\theta}{\partial t^2} > \frac{\partial^2 \vec{v}_{m2}^\theta}{\partial t^2} \geq \frac{\partial^2 \vec{v}_{m3}^\theta}{\partial t^2} > 0 \quad (9)$$

The global relation between actions can be represented as state diagram, shown in figure 4.



Keywords: Action representation: dark arrows with number refers to features defining above, dash arrows are the N/A action transition.

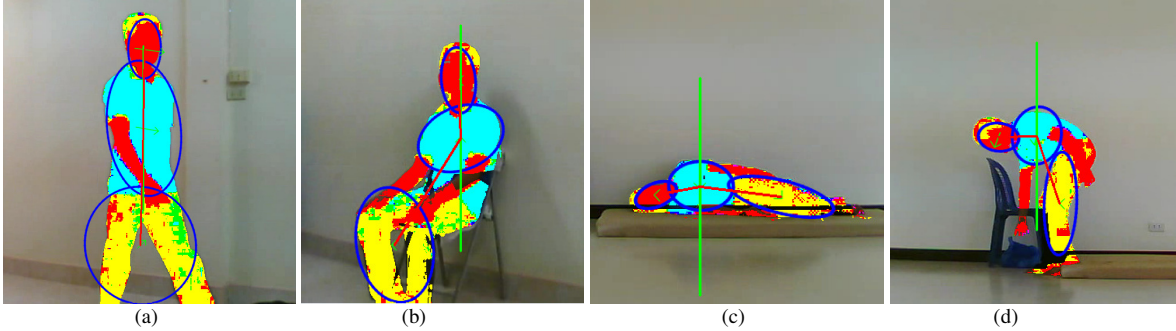
## 4. EXPERIMENTATION AND DISCUSSION

The experimentation is performed using the datasets established under indoor environments. Five peoples with different clothes act randomly and continuously with respect to the testing actions, consisting of 77 standing, 28 sitting, 50 bending, 44 walking, and 11 laying actions.

Figure 5 shows the results of the parametric human model obtained by the extraction process: motion-texture human detection and tracking, segmentation, and modeling. The image regions in red, blue, and yellow colors represent the head (hands), body, and legs of human structure. We can notice that the red color describes not only the head but also hands; due to both of these regions have the same skin color model. However, using the constraints on human model the simplification of human body is done quite well in the most cases. The blue ellipses on the figure depict the principal components of model and the red line linked between its centroids representing the internal parameters  $[\vec{v}_h, \vec{v}_l]$ . Consequently, we can determine the external parameters  $[\vec{v}_{m1}, \vec{v}_{m2}, \vec{v}_{m3}]$  by computing the movement of each component using results obtaining from the previous frames during the tracking process. The green arrows at centroid of each component show the velocity and direction of movement. In this way, the features for identifying the corresponding actions can be estimated.

Figure 6 shows the features of actions extracted from human model during a testing scenario. A man acts the actions: bending, sitting, laying, standing, and walking respectively. Using the static and dynamic action features defined in section 3, we can recognize correctly the corresponding actions, shown in the rectangle box. However, we can also observe the errors

obtained in a short period of time during the transition state from laying to standing action, which our method recognizes as N/A actions. This is quite normal because these transitions are not modeled explicitly in our action representation, figure 4.



Keywords: Examples of modeled action : (a) Standing (b) Sitting (c) Bending (d) Laying

The table 1 shows the experimentation results of our method. Number of frames representing each action is compared with frames in total of the identified action. From the highest to lowest recognition rate the actions are ordered by following: laying, standing, bending, sitting and walking respectively. The lowest rate can correctly identify actions up to 80%. In the worst case for walking action, we found that errors are produced when both of legs are separated, which leads to incorrect computation of centroid of legs component. This may be corrected by improving the human detection and tracking method. In the best case for laying action, our proposed method can completely identify all testing datasets. In average, up to 93% of basic actions are recognized appropriately. Note that N/A actions mean that the frames are not recognized by any modeled actions, but it can be identified as the non-definition transition actions showing as dash-lines in the figure 4.

However, using our method the errors of recognition can happen probably in the case where the camera is perfectly perpendicular to the actions. This make the internal parameters of human model became insignificant for the action discrimination. In the future work, to increase the recognition rate, it is necessary to correct this special case. More complex parameters model may be established or more cameras are used for recognition process, such as multi-camera based action recognition. Accordingly, this may also lead to the solution of the N/A action recognition problem too.

## 6. EXPERIMENTATION RESULTS

Actions	Number of Frames	Detected actions	Recognition rate
Standing	2886	2869	99.41 %
Sitting	1220	1089	89.26 %
Bending	2250	2123	94.35 %
Walking	2243	1809	80.65 %
Laying	1334	1334	100 %
N/A *	999	999	100 %
<b>Total</b>	<b>10932</b>	<b>9224</b>	<b>93.95%</b>

\*N/A represents action transitions that are not modeled by our action representation (dash-lines in figure 4)

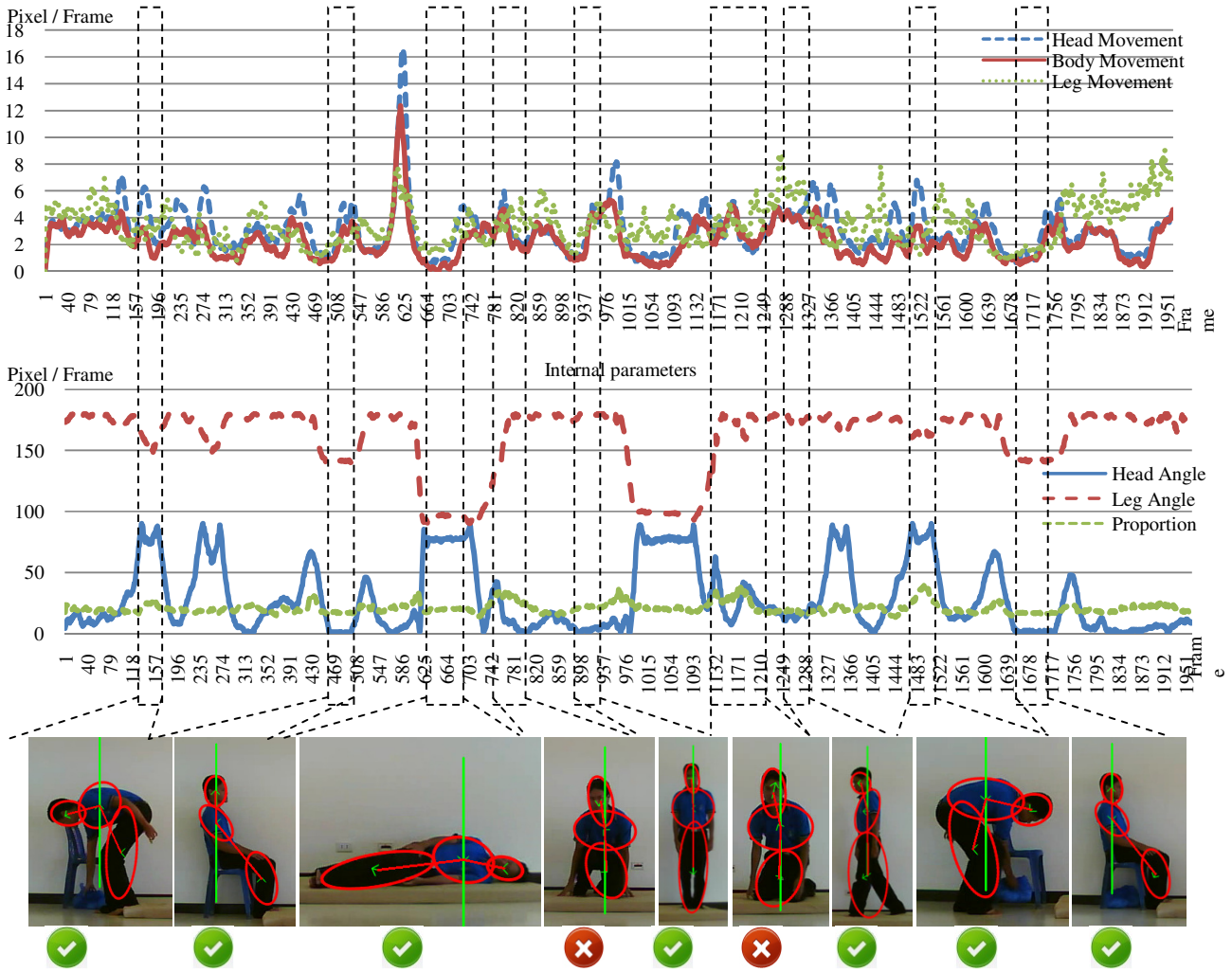
## 5. CONCLUSION

In this paper, we proposed an approach for identifying the human basic actions required for human activities recognition, such as laying, standing, bending, sitting and walking. The parametric model of human is presented for defining the action features. The action representation based on their features is then proposed as the identification process. The evaluation of our method is done over a large datasets with encourage results; up to 93% of actions are correctly recognized.

## ACKNOWLEDGMENT

Granted by ENG/PSU

Partially funded by NECTEC: Project id P-00-10638



Keywords: Example of action features with detected results.

**REFERENCES**

Krüger, Volker; Kragic, Danica; Ude, Aleš; Geib, and Christopher, "The meaning of action: a review on action recognition and mapping," *Advanced Robotics*, Volume 21, Number 13, pp. 1473-1501(29) (2007).

A. Efros, A. Berg, G. Mori, and J. Malik, "Recognizing Action at a Distance. In *International Conference on Computer Vision*," volume II, Nice, France, pp. 726–733 ( Oct 13-16, 2003).

N. Robertson and I. Reid, "Behaviour Understanding in Video: A Combined Method. In *International Conference on Computer Vision*," Beijing, China, pp. 808–815, Oct 15-21, 2005.

A. Yilmaz and M. Shah, "Actions Sketch: A Novel Action Representation. In *Computer Vision and Pattern Recognition*," volume I, San Diego, California, USA, pp. 984–989 (June 20-25, 2005).

B. Blank, L. Gorelick, E. Shechtman, M. Irani, and R. Basri, "Actions as Space-Time Shapes. In *International Conference on Computer Vision*," Beijing, China, pp. 1395–1402 (Oct 15-21, 2005).

M. Bregonzio, S. Gong and T. Xiang, "Recognising Action as Clouds of Space-Time Interest Points," In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Miami, USA, (June 2009).

Lihi Zelnik-Manor and Michal Irani, "Statistical Analysis of Dynamic Actions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 28 , Issue 9 , pp. 1530 – 1535 (September 2006).

G. Bradski and J. Davis, "Motion Segmentation and Pose Recognition with Motion History Gradients," *Machine Vision and Applications*, 13(3):174–184 (2002).

Y. Cheng, "Mean shift mode seeking and clustering," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(8):790-799 (1998).

## ประวัติผู้เขียน

ชื่อ สกุล	นายณัฐพล หนูฤทธิ	
รหัสประจำตัวนักศึกษา	5010120124	
วุฒิการศึกษา		
วุฒิ	ชื่อสถาบัน	ปีที่สำเร็จการศึกษา
วิศวกรรมศาสตรบัณฑิต (วิศวกรรมคอมพิวเตอร์)	มหาวิทยาลัยสงขลานครินทร์	2549

## การตีพิมพ์เผยแพร่ผลงาน

N. Noorit, N. Suvonvorn and M. Karnchanadecha, "Model-based Human Action Recognition," In *Proceedings of SPIE the 2<sup>nd</sup> International Conference on Digital Image Processing*, Singapore, 26<sup>th</sup>-28<sup>th</sup> February 2010, pp. 7546P-1 - 7546P-6.