



รายงานวิจัยฉบับสมบูรณ์

โครงการ การตรวจจับการหกล้มของผู้สูงอายุในห้องพักในบ้านพักคนชรา (Fall Detection System for Monitoring an Elderly Person in Elderly Care Center)

คณะผู้วิจัย

นายภูสิต กุลเกษม	หัวหน้าโครงการวิจัย
นางสาวสุวรรณา รัศมีขวัญ	ผู้ร่วมวิจัย
นางสาวเบญจภรณ์ จันทรวงกุล	ผู้ร่วมวิจัย
นายกฤษณะ ชินสาร	ผู้ร่วมวิจัย
นายอัครณัฐพันธ์ รอดทุกข์	ผู้ร่วมวิจัย
นายชิตชนก เหลือสินทรัพย์	ผู้ร่วมวิจัย
นายอภิเชษฐ์ ยาใจ	ผู้ช่วยนักวิจัย

โครงการวิจัยประเภทงบประมาณเงินรายได้
จากเงินอุดหนุนรัฐบาล (งบประมาณแผ่นดิน)
ปีงบประมาณ พ.ศ. ๒๕๕๗
มหาวิทยาลัยบูรพา

รหัสโครงการ 2557A10802244

เลขที่สัญญา 32/2557

รายงานวิจัยฉบับสมบูรณ์
โครงการ การตรวจจับการหกล้มของผู้สูงอายุในห้องพักในบ้านพักคนชรา
(Fall Detection System for Monitoring an Elderly Person in Elderly
Care Center)

คณะผู้วิจัย

นายภูสิต กุลเกษม	หัวหน้าโครงการวิจัย
นางสาวสุวรรณา รัชมีขวัญ	ผู้ร่วมวิจัย
นางสาวเบญจภรณ์ จันทรวงกุล	ผู้ร่วมวิจัย
นายกฤษณะ ชินสาร	ผู้ร่วมวิจัย
นายอรรถนุพันธ์ รอดทุกข์	ผู้ร่วมวิจัย
นายชิตชนก เหลือสินทรัพย์	ผู้ร่วมวิจัย
นายอภิเชษฐ์ ยาใจ	ผู้ช่วยนักวิจัย

คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา

กันยายน 2558

กิจกรรมประกาศ

งานวิจัยนี้ได้รับการสนับสนุนการวิจัยจากงบประมาณเงินรายได้จากเงินอุดหนุนรัฐบาล (งบประมาณแผ่นดิน) ประจำปีงบประมาณ พ.ศ. 2557 มหาวิทยาลัยบูรพา ผ่านสำนักงานคณะกรรมการการวิจัยแห่งชาติ เลขที่สัญญา 32/2557

คณะผู้วิจัย
กันยายน 2558

บทคัดย่อ

การล้มถือเป็นปัญหาด้านสุขภาพที่มีความสำคัญอย่างยิ่งในปัจจุบัน ตลอดระยะเวลาที่ผ่านมา นักวิจัยต่างพัฒนาระบบการตรวจจับการล้มในรูปแบบของการวิเคราะห์ผลจากภาพ โดยจะระบุสถานะของการล้มได้เมื่อผู้ล้มได้ล้มลงไปกับพื้นแล้ว งานวิจัยชิ้นนี้นำเสนอเทคนิคใหม่สำหรับตรวจจับการล้ม ชื่อ “กล่องขอบเขตแบบทิศทาง” หรือ Directional Bounding Box (DBB) ที่สามารถตรวจจับการล้มในกรณีที่ทิศทางของการล้มขนานกับมุมมองของกล้องได้เป็นอย่างดี เนื่องจากกล่องขอบเขตถูกสร้างขึ้นมาจากลักษณะเสมือนมองเข้าด้านข้างของทิศทางการล้มในเชิงลึกเสมอ นอกจากนี้ได้มีการนำเสนอค่าอัตราส่วนของจุดศูนย์กลางมวล (Center of Gravity Ratio) ที่สามารถนำมาใช้สนับสนุนติดตามการเคลื่อนไหวของการล้มได้ โดยทำการทดสอบและประเมินผลกับภาพเคลื่อนไหวที่ได้จากอุปกรณ์รับภาพที่มีคุณสมบัติแสดงข้อมูลเชิง 3 มิติได้ สำหรับผลการทดสอบพบว่าเทคนิคที่นำเสนอสามารถเพิ่มความถูกต้องและลดเวลาที่ใช้สำหรับตอบสนองการล้มได้ดีกว่าเทคนิคการใช้กล่องขอบเขตในงานวิจัยก่อนหน้า

Abstract

Falls are significant public health problem. In the last few years, several researches based on computer vision system have been developed to detect a person who has fallen to the ground. This paper presents a novel fall detection technique namely the directional bounding box (DBB) to detect a falls event especially a situation of fall direction paralleling the line of camera's sight. The DBB is constructed with perspective side view transformation of depth information. Moreover, a new aspect ratio namely the center of gravity point (COG) is proposed to monitor human movement. The proposed technique was evaluated with the video data set gathering from a RGB-D sensor. The experimental result of the proposed technique was better both accuracy and response times than previous works.

สารบัญ

บทที่ 1 บทนำ.....	1
1.1 ที่มาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์ของโครงการวิจัย.....	2
1.3 ขอบเขตของโครงการวิจัย	2
1.4 วิธีการดำเนินงานวิจัยโดยสรุป.....	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	5
2.1 การประมวลผลภาพ.....	5
2.2 กล้อง Kinect.....	6
2.3 การสร้างตัวแบบสำหรับภาพพื้นหลัง.....	7
2.4 การสกัดคุณลักษณะ (Feature Extraction).....	8
2.5 การรู้จำท่าทางมนุษย์	9
2.5.1 วิธีการรู้จำจากภาพสองมิติ (2D Appearance-based Methods).....	9
2.5.2 วิธีการรู้จำจากภาพสามมิติ (3D Appearance-based Methods).....	9
2.6 การสกัดคุณลักษณะ (Feature Extraction).....	10
2.7 กล้องขอบเขต	11
2.8 จุดศูนย์กลางมวลของร่างกายมนุษย์.....	12
2.9 เครื่องมือในการวัดประสิทธิภาพ	14
2.10 การทบทวนวรรณกรรมที่เกี่ยวข้อง.....	15
บทที่ 3 วิธีดำเนินการวิจัย.....	21
3.1 กล้องขอบเขตแบบทิศทาง (Directional Bounding Box: DBB).....	21
3.1.1 การสร้างเวกเตอร์ของร่างกาย.....	21

3.1.2	การคำนวณค่ามุมสำหรับการหมุนจุดร่างกาย	23
3.1.3	การหมุนกลุ่มจุดสำคัญของร่างกายให้อยู่ในแนวตั้งฉากกับมุมมอง	24
3.1.4	การระบุตำแหน่งและขนาดของกล่องขอบเขต.....	24
3.2	พีเจียร์หลักและพีเจียร์เสริมของกล่องขอบเขตแบบทิศทาง (Directional Bounding Box Feature).....	25
3.3	เฟรมเวิร์คสำหรับการตรวจจับการล้ม.....	27
3.3.1	Joint Point Provider.....	27
3.3.2	DBB Extractor.....	28
3.3.3	DBB Analyzer	28
3.4	การจัดสภาพแวดล้อมและชุดข้อมูลสำหรับการทดสอบระบบ.....	31
3.4.1	การจัดสภาพแวดล้อม	31
3.4.2	การจัดชุดข้อมูลสำหรับทดสอบ	32
3.4.3	รูปแบบสำหรับการทดสอบ	33
บทที่ 4	ผลการทดลอง	35
4.1	ผลการทดสอบความถูกต้องในการจำแนกเหตุการณ์ล้มหรือไม่ล้ม	35
4.2	ผลการทดสอบด้านเวลาที่ใช้ในการตอบสนองต่อการล้มในแต่ละทิศทาง	40
4.3	ผลการทดสอบด้านเวลาที่ใช้ในการตอบสนองต่อการล้มที่เกิดจากผลกระทบจากการแกว่งแขนหรือการก้าวเดิน	42
บทที่ 5	สรุปผลการทดลอง.....	45
5.1	สรุปผลการทดลอง	45
5.1.1	ผลการทดสอบความถูกต้องในการจำแนกเหตุการณ์ล้มหรือไม่ล้ม	45
5.1.2	ผลการวัดประสิทธิภาพด้านเวลาที่ใช้ในการตอบสนอง	46
5.1.3	ผลการทดสอบด้านเวลาที่ใช้ในการตอบสนองต่อการล้มที่เกิดจากผลกระทบจากการแกว่งแขนหรือการก้าวเดิน	46
5.2	ปัญหาและข้อเสนอแนะ	47

บทที่ 1 บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ในปัจจุบันประชากรผู้สูงอายุในประเทศไทยเพิ่มขึ้นเป็นอย่างมาก และสถานะภาพปัจจุบันของสังคมไทยคนส่วนใหญ่เป็นวัยทำงานซึ่งต้องทำงานนอกบ้านเพื่อหาเลี้ยงครอบครัว ทำให้ในครอบครัวซึ่งเป็นผู้สูงวัยจำเป็นต้องอยู่บ้านเพียงลำพังโดยไม่มีคนดูแล ดังนั้น บุตรหลานจึงนิยมขอรับบริการการดูแลผู้สูงวัยจากหน่วยงานที่รับดูแลผู้สูงอายุ และก็เป็นที่น่าเสียดายว่าในขณะที่จำนวนผู้สูงอายุที่เข้าใช้บริการบ้านพักคนชรา มากขึ้นเรื่อย ๆ แต่สถานที่ดังกล่าวกลับมีจำนวนเจ้าหน้าที่ผู้ดูแลจำนวนจำกัดไม่เพียงพอในการสอดส่องดูแลเอาใจใส่การใช้ชีวิตในแต่ละวัน ทำให้บางครั้งเมื่อเกิดอุบัติเหตุหรือการเป็นลมหรือหกล้มในขณะที่ผู้สูงวัยท่านนั้นอยู่เพียงลำพังจะทำให้เจ้าหน้าที่ไม่สามารถให้ความช่วยเหลือได้ทันกาล และอาจจะเกิดการสูญเสียได้

การเฝ้าดูกิจกรรมหรือการใช้ชีวิตของผู้สูงวัยที่มีความผิดปกติในภาพวิดีโอหรือระบบเทคโนโลยีสารสนเทศจะก่อให้เกิดความรู้สึกที่ตึงเครียดทั้งผู้สูงวัยและเจ้าหน้าที่ผู้ดูแล กล่าวคือ ผู้สูงวัยจะรู้สึกได้ถึงความเป็นส่วนตัวและไม่รู้สึกว่าถูกเฝ้าดูจากเจ้าหน้าที่ ส่วนเจ้าหน้าที่ผู้ปฏิบัติงานก็จะสามารถทำงานอื่นๆ ไปพร้อมกันในขณะที่เฝ้าดูและสังเกตการณ์การใช้ชีวิตของผู้สูงวัยจากภาพวิดีโอ จึงทำให้เจ้าหน้าที่สามารถดูแลผู้สูงวัยได้หลาย ๆ คน ไปพร้อม ๆ กันได้

กล่องขอบเขต (Bounding Box) เป็นเทคนิคที่ถูกนำมาใช้อย่างแพร่หลายในงานวิจัยที่ทำการประมวลผลกับภาพสำหรับการตรวจจับการล้มของผู้สูงอายุ โดยที่เทคนิคการสร้างกล่องขอบเขตรวมถึงการนำคุณสมบัติของข้อมูลที่ได้จากกล่องขอบเขตในแบบเดิมมาใช้ นั้น พบว่ามีข้อจำกัดในการตรวจจับการล้มที่มีลักษณะที่ทิศทางการล้มเกิดขึ้นในแนวเดียวกับมุมมองของกล้องและในบางสถานการณ์การเคลื่อนไหวบางกิจกรรมในชีวิตประจำวันส่งผลทำให้รูปร่างของกล่องมีความแปรปรวนทำให้ไม่สามารถที่จะตรวจจับการล้มได้อย่างถูกต้องแม่นยำได้

งานวิจัยนี้เสนอกล่องขอบเขตแบบใหม่สำหรับตรวจจับการล้มที่มีชื่อว่า กล่องขอบเขตแบบทิศทางหรือ Directional Bounding Box (DBB) ที่สามารถตรวจจับการล้มในกรณีทิศทางของการล้มขนานกับมุมมองของกล้องได้เป็นอย่างดี เนื่องจากกล่องขอบเขตที่สร้างขึ้นใหม่นั้นถูกออกแบบมาโดยสามารถแสดงรูปร่างที่สัมพันธ์กับทิศทางของการล้มในมุมมองที่แท้จริง นอกจากนี้ได้นำเสนอค่าอัตราส่วนของจุดศูนย์กลางมวล

(Center of Gravity Ratio) สำหรับสนับสนุนติดตามการเคลื่อนไหวของการล้ม โดยทำการทดสอบและประเมินผลกับภาพเคลื่อนไหวที่ได้จากอุปกรณ์รับภาพที่มีคุณสมบัติแสดงข้อมูลเชิง 3 มิติ ผลการทดสอบพบว่าเทคนิคที่นำเสนอสามารถเพิ่มความถูกต้องและลดเวลาที่ใช้สำหรับตอบสนองการล้มได้ดีกว่าเทคนิคการใช้กล้องขอบเขตในงานวิจัยก่อนหน้านี้

1.2 วัตถุประสงค์ของโครงการวิจัย

1. เพื่อศึกษาเทคนิคการวิเคราะห์ท่าทางมนุษย์ สำหรับใช้ในการสกัดคุณลักษณะของข้อมูลสำหรับการนำเข้าเพื่อการรู้จำ
2. เพื่อศึกษาและพัฒนาขั้นตอนวิธีเพื่อการรู้จำท่าทางมนุษย์ด้วยวิธีการเรียนรู้แบบผสม
3. เพื่อพัฒนาระบบเทคโนโลยีสารสนเทศสำหรับการเฝ้าติดตามการใช้ชีวิตของผู้สูงอายุในบ้านพักคนชรา เพื่อแจ้งเตือนเมื่อมีการหกล้มเกิดขึ้น
4. เพื่อให้ผู้ที่สนใจสามารถนำแนวความคิดที่นำเสนอ ไปศึกษาเพื่อทำการพัฒนาหรือประยุกต์ใช้ในงานวิจัยของตนเองต่อไป

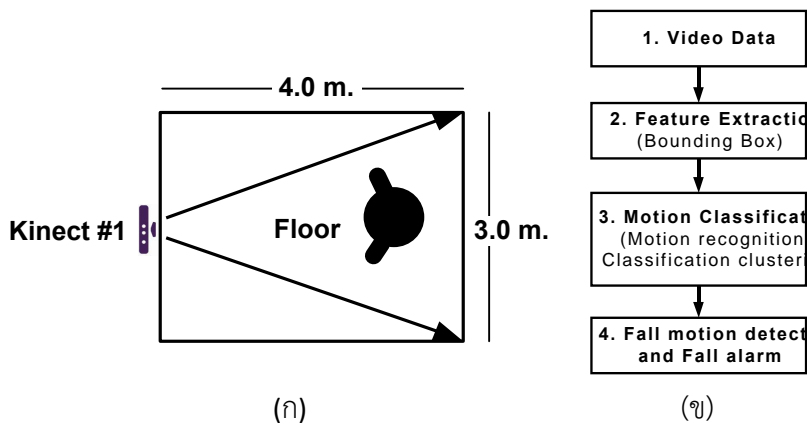
1.3 ขอบเขตของโครงการวิจัย

การวิจัยครั้งนี้มุ่งที่จะศึกษาและพัฒนาการนำเทคนิคการเรียนรู้แบบผสมเพื่อการรู้จำท่าทางการหกล้มของผู้สูงอายุในบ้านพักคนชรา โดยมีขอบเขตดังต่อไปนี้

1. ข้อมูลภาพที่ใช้เป็นภาพที่ได้จากกล้อง Infrared จำนวน 1 ตัว
2. ภาพวิดีโอที่ใช้ต้องมี frame rate อย่างน้อย 30 fps
3. ขนาดภาพขนาด 640 x 480 จุดภาพ
4. ระยะห่างของวัตถุ (คน) กับกล้องอยู่ระหว่าง 1.3 - 3 เมตร

1.4 วิธีการดำเนินงานวิจัยโดยสรุป

ในงานวิจัยนี้นำเสนอการทดลองโดยจำลองการล้มที่เกิดขึ้นในห้องโดยมีการจัดตำแหน่งกล้องดังภาพที่ 1-1 (ก) ซึ่งมีการกำหนดขั้นตอนที่สำคัญสำหรับการทำงานวิจัยไว้ 4 ขั้นตอนหลัก คือ ขั้นตอนแรกคือการรับภาพวิดีโอจำลองการล้ม จากนั้นในขั้นตอนที่ 2 จะนำภาพที่ได้ไปสกัดรูปร่างขอบเขตของตัวคนออกจากภาพพื้นหลัง ขั้นตอนที่ 3 จะทำการตรวจจับการเคลื่อนไหวโดยการบูรณาการเทคนิคการเรียนรู้จำและจำแนกการล้ม และ ขั้นตอนที่ 4 ตรวจจับและแจ้งเตือนการล้ม ซึ่งรายละเอียดย่อยของขั้นตอนหลักจะถูกนำเสนอ ดังนี้



รูปที่ 1-1 กรอบแนวคิดในการทำงานวิจัยที่นำเสนอ

จากรูปที่ 1-1 ในขั้นตอนที่ 1 ผู้วิจัยจะเลือกใช้ข้อมูลที่น่าเข้าเป็นข้อมูลวิดีโอจากกล้อง Kinect เนื่องจากกล้องประเภทนี้จะให้ทั้งภาพ และ สัญญาณอินฟราเรด (Infrared) ซึ่งมีส่วนสำคัญที่จะทำให้ผลการวิจัยมีประสิทธิภาพและสามารถนำไปใช้งานได้จริงในทุกสภาพวะของแสง

จากนั้น ในขั้นตอนที่สอง ขั้นตอนการสกัดคุณลักษณะ ในงานวิจัยทั่วไปการใช้กล่องขอบเขตเพื่อทำการตรวจจับนั้นส่วนใหญ่จะพบข้อเสียอันเกิดขึ้นจากการสร้างกล่องขอบเขตที่ไม่สามารถแสดงคุณลักษณะของร่างกายมนุษย์ได้อย่างถูกต้อง งานวิจัยนี้จึงพัฒนารูปแบบของพีเจอร์ดังกล่าวให้ยืดหยุ่นแสดงลักษณะของการเคลื่อนไหวของมนุษย์ให้ดียิ่งขึ้น

ในขั้นตอนที่สาม จะทำการรู้จำเพื่อการตรวจจับการเคลื่อนที่ของมนุษย์โดยเทคนิคการเรียนรู้แบบผสมโดยอาศัยพีเจอร์ดสนับสนุนเข้าด้วยกันผ่านการชุดจำลอง

สถานการณ์การล้มและไม่ล้ม เพื่อนำผลการเรียนรู้ที่ได้ไปใช้ในขั้นตอนสุดท้ายในการพยากรณ์ว่าผู้สูงอายุท่านนั้นว่าล้ม ก้มนั่งหรือเก็บของ เพราะเหตุการณ์ที่จำลองสถานการณ์ดังกล่าวเป็นเหตุการณ์ที่รูปร่างกายของผู้สูงอายุใกล้เคียงกับการนอนลงไปบนพื้นเช่นกัน

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ขั้นตอนวิธีการหาลักษณะเด่นของท่าทางมนุษย์เพื่อการรู้จำ โดยอาศัยเทคนิคการสกัดคุณลักษณะและการรู้จำแบบผสม
2. สามารถนำไปพัฒนาระบบสารสนเทศช่วยตัดสินใจสำหรับการเฝ้าดูการหกล้มของผู้สูงอายุในบ้านพักคนชรา ซึ่งเกิดขึ้นเป็นประจำ
3. ขั้นตอนวิธีที่นำเสนอจะสามารถใช้เป็นต้นแบบในการศึกษาขั้นสูงต่อไป

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การประมวลผลภาพ

นิยาม การประมวลผลข้อมูลภาพ หรือ การประมวลผลภาพ (Image Processing) หมายถึง การจัดการและการวิเคราะห์สารสนเทศของภาพโดยใช้คอมพิวเตอร์ โดยมีวัตถุประสงค์เพื่อ

- ปรับปรุงหรือเพิ่มสารสนเทศของรูปเพื่อการแปลความหมายของมนุษย์ที่ดีขึ้น
- ทำให้รูปนั้นๆ เปลี่ยนแปลงในทางที่ดีขึ้นเพื่อให้เหมาะสมกับการประมวลผล หรือ การรับรู้ด้วยเครื่องจักรอย่างอัตโนมัติ (Autonomous machine perception)

ตัวอย่างขั้นตอนวิธีของการจัดการและการวิเคราะห์สารสนเทศของภาพ เช่น การแปลงภาพ (Image Transformation) การนิยามภาพ (Image Description) การกรองภาพ (Image Filters) การคืนคืนภาพ (Image Restoration) การปรับปรุงคุณภาพของภาพ (Image Enhancement) การแบ่งภาพและการหาขอบวัตถุในภาพ (Image Segmentation and Edge Detection) และ การบีบอัดข้อมูลภาพ (Image Compression) เป็นต้น

จากที่ได้กล่าวมาข้างต้น การประมวลผลภาพเชิงตัวเลข หรือเรียกสั้นๆ ว่า การประมวลผลภาพ (Digital Image Processing or Image Processing) คือ การประมวลผลของคอมพิวเตอร์เพื่อเปลี่ยนแปลงธรรมชาติของภาพเชิงตัวเลข (Digital Image) ดังนั้น การประมวลผลภาพถูกพัฒนาขึ้นมาเพื่อตอบปัญหาหลักๆ เช่น

- การทำดิจิทัลไจซ์ และ การเข้ารหัสข้อมูลภาพเพื่อการส่ง การพิมพ์ และ การจัดเก็บอย่างมีประสิทธิภาพ ปัญหาที่ เนื่องจากรูปต้นฉบับ (Original picture) นั้นเป็นข้อมูลเชิงอนาล็อก (Analog picture) ซึ่งเป็นข้อมูลที่มีความต่อเนื่อง ยังไม่สามารถประมวลผลได้ด้วยคอมพิวเตอร์ จึงจำเป็นต้องใช้การทำดิจิทัลไจซ์เพื่อทำการสุ่มข้อมูล และ แปลงข้อมูลดังกล่าวนั้น ให้เป็นข้อมูลเชิงตัวเลข (Discrete or Digital image) เพื่อการประมวลผลสำหรับเครื่องคอมพิวเตอร์ต่อไป

- การปรับปรุงคุณภาพของภาพ เช่น การเพิ่มความคมชัดให้กับขอบของวัตถุในภาพทำให้ภาพดูสว่างมากขึ้น ดังตัวอย่างในรูปที่ 2 โดยภาพทาง

ซ้ายมือ คือ ภาพต้นฉบับ และ เมื่อผ่านขั้นตอนการปรับปรุงภาพโดยการเพิ่มความสว่าง ทำให้ภาพมีความคมชัดมากขึ้น

- การกำจัดสัญญาณรบกวน เมื่อนำภาพดังกล่าวไปผ่านกระบวนการการกำจัดสัญญาณรบกวน ทำให้ได้คืนภาพที่มีความสมบูรณ์เกือบเหมือนจริงดังภาพขวามือ เป็นต้น ซึ่งภาพที่คืนมาได้นั้น จะทำให้การแปลความต่อสารสนเทศที่อยู่บนข้อมูลภาพในปัญหาต่างๆ ทำได้ง่ายขึ้น

การแบ่งภาพ และการนิยามภาพ ถือเป็นขั้นตอนสำคัญสำหรับระบบการเรียนรู้ และ มองเห็นด้วยเครื่องจักร (Machine Learning and Vision) เพราะขั้นตอนวิธีนี้จะช่วยให้คอมพิวเตอร์หรือเครื่องจักรสามารถประมวลผลกับส่วนต่างๆ ที่แบ่งออกมาได้ง่ายขึ้น

2.2 กล้อง Kinect

กล้อง Kinect เป็นอุปกรณ์จับความเคลื่อนไหวของไมโครซอฟต์ได้รับการพัฒนาและวางจำหน่ายเมื่อวันที่ 4 พฤศจิกายน พ.ศ. 2553 ในฐานะอุปกรณ์เสริมของเครื่องเกมส์ Xbox 360 เป็นกล้องที่สามารถจดจำผู้เล่น (Facial Recognition) และให้ผู้เล่นควบคุมเกมส์ผ่านทางท่าเคลื่อนไหวของร่างกายของผู้เล่นได้โดยตรง (3D Motion Recognition) และสามารถจดจำเสียงของผู้เล่นได้ (Voice Recognition) ตัวกล้องมีลักษณะเป็นแท่งแบน ๆ วางอยู่ในแนวระดับที่เชื่อมต่อกับฐานขนาดเล็ก สามารถหมุนปรับมุมกล้องได้ด้วยมอเตอร์ ภายในแท่งประกอบด้วยกล้อง CMOS RGB 1 ตัว กล้อง IR สำหรับตรวจวัดความลึกของภาพ 1 ตัว และไมโครโฟน 4 ตัว เอาท์พุทของกล้องได้แก่ภาพวิดีโอที่มีอัตราการแสดงภาพเคลื่อนไหว 30 เฟรมต่อวินาที ขนาดภาพเท่ากับ 640x480 พิกเซล โดยที่ 1 พิกเซลมีขนาด 8 บิต สำหรับภาพขาวดำที่เกิดจากกล้อง IR เป็นภาพขนาด 640x480 พิกเซลที่สามารถถ่ายได้ในทุกสภาวะแสง สามารถแสดงระดับความลึกของภาพได้สูงถึง 2048 ระดับ หรือ 11 บิตต่อพิกเซล มุมกล้องสามารถปรับในแนวนอนได้ถึง 57 องศา แนวตั้ง 43 องศา นอกจากนี้ใช้ฟังก์ชันสำหรับการเล่นเกมแล้ว ปัจจุบันไดร์เวอร์สำหรับระบบปฏิบัติการ Linux และ Windows ได้ถูกพัฒนาขึ้นทำให้ผู้ใช้งานสามารถนำกล้อง Kinect มาประยุกต์ใช้งานได้หลากหลายยิ่งขึ้น ไม่ว่าจะนำมาใช้เพื่อความบันเทิงหรือใช้สำหรับควบคุมการเคลื่อนไหวของหุ่นยนต์



รูปที่ 2-1 กล้อง Kinect

2.3 การสร้างตัวแบบสำหรับภาพพื้นหลัง

การสกัดวัตถุออกจากภาพพื้นหลังเมื่อใช้กล้องที่ติดตั้งอยู่กับที่นักวิจัยส่วนใหญ่จะใช้ background subtraction ด้วยแบบจำลองทางคณิตศาสตร์ของสภาพแวดล้อมและการสังเกตการณ์ โดยพยายามตรวจจับสิ่งที่เปลี่ยนไปในเฟรมปัจจุบัน ซึ่งวิธีนี้จะเหมาะสมกับกรณีที่ภาพพื้นหลังและสภาพแสงในพื้นที่บริเวณนั้นไม่มีการเปลี่ยนแปลง แต่ถ้าส่วนภาพพื้นหลังหรือสภาพแสงในภาพนั้นมีการเปลี่ยนแปลงไปจากเดิมตลอดเวลา จำเป็นจะต้องมีขั้นตอนวิธีที่สามารถปรับปรุงภาพพื้นหลังตลอดเวลาเช่นกัน ทั้งนี้ เพื่อให้ผลของการสกัดภาพวัตถุออกมีความถูกต้องสูงสุด การสร้างตัวแบบสำหรับภาพพื้นหลังเพื่อการสกัดวัตถุ จะประกอบด้วย ขั้นตอนย่อย 3 ขั้นตอน คือ

1. Background Model: จำลองแต่ละจุดภาพของพื้นหลังโดย single Gaussian distribution
2. Update of Background Model: เนื่องจากในระบบจริงภาพพื้นหลังจะไม่หยุดนิ่ง ดังนั้น แบบจำลองต้องยืดหยุ่นเพียงพอที่จะปรับอัตโนมัติต่อการเปลี่ยนแปลงต่าง ๆ ในสภาพแวดล้อม เพื่อจัดการกับการเปลี่ยนแปลงของแสงสว่างเช่นการเปลี่ยนไปของแสงตาม และ
3. Postprocessing: วัตถุที่ถูกตรวจจับโดยวิธีนี้เข้ากันได้เป็นอย่างดีกับพื้นที่จำกัดที่มีขอบเรียบ การตรวจหาที่ผิดจะเป็นไปทั่วบริเวณของภาพและจะตกอยู่เป็นกลุ่มเล็ก ๆ ของจุดภาพที่ถูกแยกตัวออกมา เราจะใช้การดำเนินการต่อรูปทรงเพื่อขจัดจุดภาพที่ถูกแยกตัวออกมาและเพื่อเติม

ช่องว่างในภาพ foreground แล้ว foreground pixels จะถูกจัดกลุ่มเป็นส่วนประกอบที่เชื่อมโยงกัน

2.4 การสกัดคุณลักษณะ (Feature Extraction)

การสกัดคุณลักษณะ หรือ Feature Extraction เป็นขั้นตอนสำคัญในการสกัดเอาลักษณะเฉพาะของแต่ละข้อมูลนำเข้าเพื่อนำไปใช้เป็นข้อมูลนำเข้าในการฝึกฝนระบบและทดสอบระบบ

การสกัดคุณลักษณะ เป็นเทคโนโลยีตัวหนึ่ง ที่ได้เข้ามามีบทบาทร่วมกับงานวิจัยหลายๆ ด้าน และสามารถนำมาประยุกต์ใช้งานในด้านต่างๆ มากมาย สำหรับงานด้านการประมวลผลภาพ โดยเฉพาะที่เกี่ยวข้องกับการรู้จำตัวแบบ การสกัดคุณลักษณะมักนำมาใช้ในการลดมิติข้อมูล เมื่อข้อมูลนำเข้าของขั้นตอนวิธี (Algorithm) ใดๆ เกิดมีขนาดใหญ่เกินกว่าจะสามารถประมวลผลและเสี่ยงต่อการเกิดความซ้ำซ้อนของข้อมูล ข้อมูลมีมาก แต่ที่สามารถนำไปใช้ได้จริงมีน้อย ดังนั้นข้อมูลนำเข้าจะถูกนำไปแปลงให้มีเฉพาะกลุ่มที่เป็นตัวแทนของคุณลักษณะเท่านั้น เรียกว่าเป็น "เวกเตอร์คุณลักษณะ" หรือ "Features Vector" การแปลงข้อมูลนำเข้าให้อยู่ในรูปกลุ่มของคุณลักษณะนี้เรียกว่า "การสกัดคุณลักษณะ"

แต่ละคุณลักษณะที่ถูกสกัดได้ จะถูกเลือกไปใช้อย่างระมัดระวัง เนื่องจากคุณลักษณะนั้นๆ จะได้รับการคาดหวังว่าเป็นกลุ่มคุณลักษณะที่สกัดเอาแต่ส่วนข้อมูลสำคัญจากข้อมูลนำเข้ามาแล้ว เพื่อที่จะสามารถดำเนินงานให้บรรลุจุดประสงค์ได้โดยใช้กลุ่มตัวแทนนี้ แทนที่จะส่งข้อมูลนำเข้าทั้งหมดไปประมวลผล

2.5 การรู้จำท่าทางมนุษย์

วิธีการรู้จำท่าทางของมนุษย์แบ่งออกเป็น 2 ประเภท คือ

2.5.1 วิธีการรู้จำจากภาพสองมิติ (2D Appearance-based Methods)

แบ่งตัวแบบออกเป็นประเภทที่เป็นวิธีการรู้จำแบบตรง โดยเริ่มต้นด้วยการตรวจสอบส่วนที่เป็นร่างกายมนุษย์ในภาพ และหาองค์ประกอบหลักของร่างกายมนุษย์ ซึ่งได้แก่ส่วนที่เป็นศีรษะ มือ และเท้า จากนั้นจึงทำการตรวจสอบส่วนของร่างกายที่เป็นองค์ประกอบรองคือส่วนที่เป็นข้อต่อได้แก่ หัวไหล่ ข้อศอก และ หัวเข่า จากนั้นจึงทำการวิเคราะห์ท่าทางของมนุษย์โดยการวิเคราะห์เส้นขอบของภาพร่างกายมนุษย์ที่ได้ มีข้อดีคือใช้เวลาในการประมวลผลน้อย แต่มีข้อเสียคือหากมีการแยกส่วนของร่างกายผิด จะทำให้เกิดการรู้จำที่ผิดพลาดตามไปด้วย ตัวแบบประเภทที่สองเป็นวิธีการรู้จำเชิงสถิติ โดยการคำนวณค่าโปรเจกต์ชันแนวตั้งและแนวนอนจากภาพท่าทางของมนุษย์เพื่อใช้หาท่าทางหลัก(ท่ายืน ท่านั่ง ท่าก้มตัวคลาน และท่านอน) และใช้พิจารณาหาทิศทางของร่างกายว่าเป็นมุมมองที่หันด้านใด(ด้านหน้า ด้านซ้าย หรือด้านขวา) ซึ่งการรู้จำท่าทางของมนุษย์นั้นจะทำการคำนวณค่าโปรเจกต์ชันของท่าทางมนุษย์ในภาพปัจจุบันเทียบกับค่าโปรเจกต์ชันของท่าทางมนุษย์ในชุดภาพที่นิยามไว้ก่อนหน้า จากนั้นจึงทำการหาส่วนต่างๆของร่างกายโดยวิเคราะห์จากเส้นขอบของภาพท่าทางมนุษย์ วิธีนี้มีข้อดีคือใช้เวลาในการประมวลผลน้อย ข้อเสียคือผลที่ได้ขึ้นอยู่กับมุมมองของภาพ

2.5.2 วิธีการรู้จำจากภาพสามมิติ (3D Appearance-based Methods)

แบบจำลองสามมิติถูกสร้างขึ้นจากวัตถุทรงเรขาคณิตรูปทรงต่างๆ เช่น วัตถุทรงกระบอก ทรงกลม หรือรูปกรวย โดยแบบจำลองนี้ประกอบด้วยตัวแปรที่ใช้ในการกำหนดความสัมพันธ์ระหว่างวัตถุทรงเรขาคณิตที่เป็นส่วนประกอบต่างๆ แบ่งตัวแบบออกเป็น 2 ประเภท คือ

1. วิธีการรู้จำจากภาพสามมิติที่ใช้กล่องตัวเดียว ทำได้โดยวิธีการรู้จำแบบตรง ซึ่งเป็นแบบจำลองที่เกิดจากการเชื่อมต่อของรูปทรงเรขาคณิตชนิดต่างๆ โดยจะต้องอาศัยความรู้เดิมเกี่ยวกับโครงสร้างของแบบจำลองมนุษย์และข้อจำกัดต่างๆที่เกี่ยวข้อง

กับโครงสร้างนั้นๆก่อน หรืออีกวิธีหนึ่งคือโดยวิธีการรู้จำแบบใช้วิธีการเรียนรู้ควบคู่ไปด้วย วิธีนี้ไม่จำเป็นต้องใช้แบบจำลองสามมิติเหมือนแบบตรง แต่จะทำการเก็บภาพท่าทางสามมิติไว้ในฐานข้อมูลภาพ เพื่อใช้การการเปรียบเทียบ ข้อดีคือผลที่ได้ไม่ขึ้นอยู่กับมุมมองของภาพ ข้อเสียคือใช้เวลาในการประมวลผลมาก และจำเป็นต้องมีฐานข้อมูลที่สามารถเก็บรายละเอียดข้อมูลได้ครบถ้วน

2. วิธีการรู้จำจากภาพสามมิติที่ใช้กล้องหลายตัว สามารถรู้จำท่าทางของมนุษย์ได้ในกรณีภาพอาจถูกบังเป็นบางส่วน เช่น มีเฉพาะร่างกายส่วนบนในภาพ เป็นต้น ทำได้โดยวิธีการรู้จำโดยใช้แบบจำลอง ซึ่งจะทำการสร้างเป็นแบบจำลองสามมิติขึ้นพร้อมทั้งคำนวณค่าตัวแปรต่างๆที่เกี่ยวข้อง และอีกวิธีหนึ่งคือวิธีการรู้จำแบบใช้การเรียนรู้ ซึ่งจะทำการเรียนรู้จากภาพท่าทางของมนุษย์ที่ทำการเรียนรู้และทราบก่อนหน้าแล้ว ข้อดีคือผลที่ได้ไม่ขึ้นอยู่กับมุมมองของภาพ ข้อเสียคือใช้เวลาในการประมวลผลมาก

2.6 การสกัดคุณลักษณะ (Feature Extraction)

การสกัดคุณลักษณะ หรือ Feature Extraction เป็นขั้นตอนสำคัญในการสกัดเอาลักษณะเฉพาะของแต่ละข้อมูลนำเข้าเพื่อนำไปใช้เป็นข้อมูลนำเข้าในการฝึกฝนระบบและทดสอบระบบ

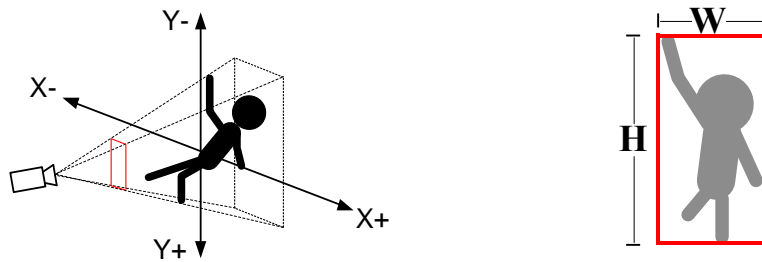
การสกัดคุณลักษณะ เป็นเทคโนโลยีตัวหนึ่ง ที่ได้เข้ามามีบทบาทร่วมกับงานวิจัยหลายๆ ด้านและสามารถนำมาประยุกต์ใช้งานในด้านต่างๆ มากมาย สำหรับงานด้านการประมวลผลภาพ โดยเฉพาะที่เกี่ยวข้องกับการรู้จำตัวแบบ การสกัดคุณลักษณะมักนำมาใช้ในการลดมิติข้อมูล เมื่อข้อมูลนำเข้าของขั้นตอนวิธี (Algorithm) ใดๆ เกิดมีขนาดใหญ่เกินกว่าจะสามารถประมวลผลและเสี่ยงต่อการเกิดความซ้ำซ้อนของข้อมูล ข้อมูลมีมาก แต่ที่สามารถนำไปใช้ได้จริงมีน้อย ดังนั้นข้อมูลนำเข้าจะถูกนำไปแปลงให้มีเฉพาะกลุ่มที่เป็นตัวแทนของคุณลักษณะเท่านั้น เรียกว่าเป็น "เวกเตอร์คุณลักษณะ" หรือ "Features Vector" การแปลงข้อมูลนำเข้าให้อยู่ในรูปกลุ่มของคุณลักษณะนี้เรียกว่า "การสกัดคุณลักษณะ"

แต่ละคุณลักษณะที่ถูกสกัดได้ จะถูกเลือกไปใช้อย่างระมัดระวัง เนื่องจากคุณลักษณะนั้นๆ จะได้รับการคาดหวังว่าเป็นกลุ่มคุณลักษณะที่สกัดเอาแต่ส่วนข้อมูลสำคัญจากข้อมูลนำเข้ามาแล้ว เพื่อที่จะสามารถดำเนินงานให้บรรลุจุดประสงค์ได้โดยใช้กลุ่มตัวแทนนี้ แทนที่จะส่งข้อมูลนำเข้าทั้งหมดไปประมวลผล

2.7 กล่องขอบเขต

กล่องขอบเขต (Bounding Box) หรือรูปสี่เหลี่ยมขอบเขตเล็กสุด (Minimum Bounding Box :MBR) เป็นกล่องที่มีลักษณะสี่เหลี่ยมที่อยู่บริเวณโดยรอบของวัตถุหรือร่างกายมนุษย์ที่ต้องการพิจารณา ด้วยคุณสมบัติที่สามารถแสดงได้ถึงความแตกต่างระหว่างการยื่นและการล้มลงได้อย่างชัดเจน นักวิจัยจึงนิยมนำค่าทางคณิตศาสตร์ที่เกิดขึ้นในกล่องขอบเขตนี้มาใช้ในการติดตามท่าทางการเคลื่อนไหวของมนุษย์ในงานวิจัยของตนเอง โดยที่กล่องขอบเขตนี้สามารถจำแนกออกเป็น 2 รูปแบบดังนี้

1. กล่องขอบเขตแบบสองมิติ (2D Bounding Box: 2DBB) เป็นกล่องขอบเขตที่สามารถสร้างได้จากภาพที่ได้จากกล่องทั่วไปที่สามารถแสดงผลได้เพียงสองมิติเท่านั้น กล่องขอบเขตชนิดนี้ถูกใช้ในงานวิจัยอย่างแพร่หลายเพราะอุปกรณ์รับภาพที่ใช้สามารถจัดหาได้ง่ายทั่วไปดังแสดงในรูปที่ 2-2



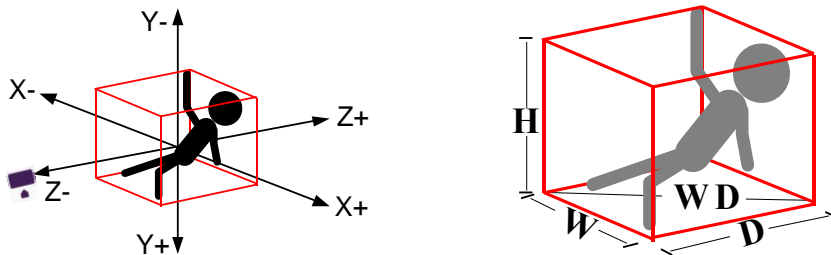
รูปที่ 2-2 กล่องขอบเขตแบบสองมิติ

โดยที่ค่าอัตราส่วนระหว่างค่าความสูง (H) และค่าความกว้าง (W) ของกล่องขอบเขตนี้ ถูกนำมาใช้เป็นตัวแทนสถานะของการล้มของตัวบุคคล ซึ่งหาค่ามีค่าน้อยกว่า 1 หรือค่าความสูงน้อยกว่าความกว้าง นั้นหมายความว่ามีความเป็นไปได้ที่จะล้ม เกิด โดยค่าอัตราส่วนนี้คือ R_{HW} สามารถคำนวณหาได้จากสมการที่ (2.1)

$$R_{HW} = \frac{H}{W} \quad (2.1)$$

2. กล่องขอบเขตแบบสามมิติ (3D Bounding Box: 3DBB) เป็นกล่องขอบเขตที่ถูกยกระดับขึ้นมาอันเป็นผลมาจากการพัฒนาของอุปกรณ์รับภาพที่สามารถแสดงตำแหน่งของจุดภาพในแนวแกน Z หรือในเชิงความลึกได้ ซึ่งตัวอุปกรณ์รับภาพจะใช้

เทคโนโลยีของอินฟราเรดมาใช้ในการสร้างชุดข้อมูลที่มีมิติเหล่านี้ได้ นักวิจัยจึงนำประโยชน์ในมิติข้อมูลดังกล่าวมาประยุกต์ใช้สำหรับการตรวจจับการล้มที่มีทิศทางเข้าและออกจากตัวเซนเซอร์รับภาพได้ ดังแสดงดังภาพที่ 2-3



รูปที่ 2-3 กล่องขอบเขตแบบสามมิติ

ค่าทางคณิตศาสตร์ที่นักวิจัยนิยมใช้ในการตรวจจับการล้ม คือ ความยาวของเส้น WD หรือ WD length มาใช้ โดยที่ค่าความยาวของเส้น WD ยังมีค่ามาเท่าไรยิ่งสอดคล้องกับสถานะของการล้มลงมากเท่านั้น โดยค่าของความยาวของเส้น WD จะสามารถคำนวณได้จากสมการที่ (2.2)

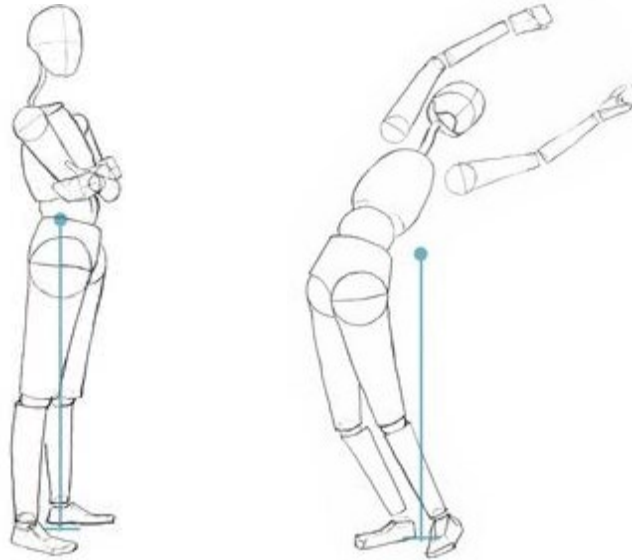
$$WD = \sqrt{W^2 + D^2} \quad (2.2)$$

และเนื่องจากค่าของความยาวของเส้น WD มีความแปรปรวนมากหากแขนหรือการก้าวเดิน แม้ว่านักวิจัยจะมีการตรวจสอบยืนยันโดยใช้ค่าความสูงของกล่องมารวมกันวิเคราะห์ แต่ผลการตรวจจับยังเกิดผลที่อยู่ในรูปแบบ false negative หรือมีการแจ้งว่าไม่ล้มแต่ตัวบุคคลได้ล้มลงไปที่พื้นแล้วเกิดขึ้น

2.8 จุดศูนย์กลางมวลของร่างกายมนุษย์

จุดศูนย์กลางมวลของร่างกายมนุษย์ (Center of Mass of a Human Body) (Arthur Chapman, et.al 2008) ใช้หลักการเดียวกันกับจุดศูนย์กลางของแรงโน้มถ่วง (Center of gravity) ที่สามารถบ่งบอกได้ถึงสมดุลของวัตถุได้ โดยที่จุดศูนย์กลางมวลของร่างกายมนุษย์ขึ้นอยู่กับท่าทางของมนุษย์ที่แสดงออก ณ ขณะนั้นตัวอย่างเช่นในสถานการณ์ที่มีการยืนตรงจะทำให้จุดศูนย์กลางมวลจะอยู่ภายในร่างกายบริเวณ

เหนือท้องขึ้นมาเล็กน้อย แต่ถ้าหากในสถานการณ์ที่ร่างกายมีการโน้มเอียงไปยังข้างใดข้างหนึ่งจะมีผลทำให้จุดศูนย์กลางมวลอยู่บริเวณนอกร่างกายดังรูปที่ 2-4

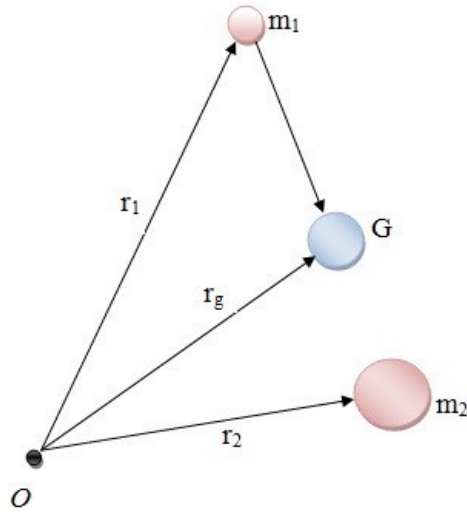


รูปที่ 2-4 แสดงจุดศูนย์กลางมวลอยู่ภายในและนอกริเวณของร่างกาย

การคำนวณจุดศูนย์กลางของมวลในวัตถุหนึ่งชิ้นนั้นสามารถทำได้แม้ว่าวัตถุนั้นจะมีทิศทางการวางตัวไปในทิศทางใดก็ตาม เนื่องมาจากจุดศูนย์กลางดังกล่าวจะไม่มี การเลื่อนตำแหน่งแม้ว่าวัตถุนั้นจะถูกหมุนไปในทิศทางใด โดยจุดศูนย์กลางของมวลจะอยู่ที่ r_{cm} ดังสมการ (2.3) ด้านล่าง

$$r_{cm} = \sum_{i=1}^n \frac{r_i m_i}{m_{tot}} \quad (2.3)$$

โดยที่ r_i คือ ตำแหน่งเวกเตอร์ของมวลที่ i
 m_i คือ มวลของวัตถุชิ้นที่ i
 m_{tot} คือ ผลรวมของมวลทั้งหมดในระบบ
 n คือ จำนวนของจุดมวลทั้งหมดในระบบ



รูปที่ 2-5 แสดงตำแหน่งจุดศูนย์กลางมวลระหว่างมวลสองตำแหน่ง

จากรูปที่ 2-5 จะแสดงถึงระบบที่มีมวล 2 ตำแหน่งได้แก่ m_1 และ m_2 โดยที่มวลของ m_1 มีขนาดน้อยกว่า m_2 จึงทำให้ตำแหน่งของจุดศูนย์กลางมวล G ใกล้ไปทางมวล m_2 ที่มีขนาดของมวลที่มากกว่า ในขณะที่เวกเตอร์ r_1 และ r_2 ต่างเป็นเวกเตอร์ที่มีจุดเริ่มต้นอ้างอิงในเกณฑ์เดียวกันในระบบ

2.9 เครื่องมือในการวัดประสิทธิภาพ

วิธีการวิเคราะห์ความถูกต้อง ในงานวิจัยนี้จะวัดประสิทธิภาพของผลการทดลองโดยพิจารณาจากค่า accuracy specificity และ sensitivity (Tom Fawcett, 2006) โดยค่าดังกล่าวจะคำนวณได้จากสมการ (2.4) - (2.5) ต่อไปนี้

ค่า Accuracy

$$accuracy = \frac{TP + FN}{TP + TN + FP + FN} \quad (2.4)$$

ค่า Specificity

$$specificity = \frac{TN}{TN + FP} \quad (2.5)$$

ค่า Sensitivity

$$sensitivity = \frac{TP}{TP + FN} \quad (2.6)$$

โดยที่ TP = True Positive,
 TN = True Negative,
 FN = False Negative, และ
 FP = False Positive

2.10 การทบทวนวรรณกรรมที่เกี่ยวข้อง

Nor Surayahani Suriani และคณะ (Nor Surayahani, et.al 2012) ได้ตั้งสมมติฐานว่าสามารถใช้การกระจายแบบเรขาคณิตของการเคลื่อนไหว ในการระบุความผิดปกติในเชิง spatial และใช้การวิเคราะห์การเคลื่อนที่ในการระบุความผิดปกติในเชิงtemporal ได้ โดยทำการทดลองกับชุดข้อมูล Multiple camera fall dataset ซึ่งประกอบด้วยกิจกรรมในชีวิตประจำวันได้แก่ การเดิน การก้มตัว และการล้มลักษณะต่าง ๆ เช่น ล้มไปด้านหน้า ล้มไปด้านหลัง และล้มไปด้านข้าง ในงานวิจัยดังกล่าวนำเสนอวิธีการโดยแบ่งออกเป็น 3 ขั้นตอน โดยขั้นตอนแรกจะใช้ Motion History Histogram เพื่อติดตามพฤติกรรมเคลื่อนไหว ในการสกัดการเปลี่ยนแปลงของร่างกาย ซึ่งมีความเร็วในการเคลื่อนไหวในขณะเดินและขณะล้มแตกต่างกัน ขั้นที่สองทำการคำนวณค่า Motion Geometric Distribution (MGD) เพื่อใช้ลดความเร็วในการเคลื่อนไหวกว้างภาพ ทำให้สามารถตรวจพบลักษณะการล้มได้ง่ายยิ่งขึ้น และขั้นตอนสุดท้ายใช้ Biological inspired Feed-forward network ในการแบ่งกลุ่มกิจกรรมการเดินและการล้มออกจากกัน ซึ่งผลการทดลองพบว่าวิธีการที่ใช้สามารถแบ่งกลุ่มสถานะการเดินและการล้มของมนุษย์ได้อย่างถูกต้อง

Jia Luen Chua และคณะ (Jia Luen Chua, et.al 2012) ได้นำเสนอเทคนิคใหม่ในการตรวจจับการล้มโดยขึ้นกับการเปลี่ยนแปลงของรูปร่างมนุษย์ เทคนิคนี้ใช้จุดสามจุดเพื่อแทนบุคคลแทนที่จะใช้ ellipse หรือ bounding box คุณลักษณะที่ถูกสกัดจากเส้นที่เกิดจากจุดสามจุดนี้ได้ถูกใช้ในการวิเคราะห์การเปลี่ยนแปลงของรูปร่างเพื่อตรวจจับการล้ม เทคนิคการใช้จุดสามจุดนี้ไม่เพียงแต่เพิ่มอัตราการตรวจจับการล้มได้เพิ่มขึ้นแต่ยังลดค่า computational complexity อีกด้วย สำหรับวิธีการที่ใช้

ในงานวิจัยนี้ ได้แก่การสกัดเพื่อแยกรูปร่างมนุษย์ (Human Shape Extraction) ซึ่งจะทำการลบภาพพื้นหลังก่อน เพื่อตรวจหาตัวบุคคลในชุดภาพวิดีโอ เมื่อพบแล้วก็คำนวณหาจุดสามจุดที่ใช้แทนบริเวณต่าง ๆ ของร่างกายมนุษย์ได้แก่ ศีรษะ ลำตัว และขา ลักษณะอื่น ๆ จะได้ตามมา เช่น การเปลี่ยนมุม ผลรวมของความสูง และอัตราของความสูง และนำไปใช้วิเคราะห์รูปร่างของมนุษย์ ขั้นตอนถัดไปเป็นการใช้ระบบการตรวจจับการล้ม ซึ่งประกอบด้วย 2 ส่วนได้แก่ ส่วนที่หนึ่งเป็นการรู้จำการล้ม (fall recognition) ซึ่งเป็นส่วนที่ขึ้นอยู่กับการวิเคราะห์รูปร่างมนุษย์ และ ส่วนที่สองเป็นการยืนยันการล้ม (fall confirmation) ซึ่งจะขึ้นอยู่กับช่วงเวลาที่คุณคนนั่งไป ผลการทดลองจากงานวิจัยนี้ ให้ความถูกต้องในการตรวจจับการล้ม 90.5% และมีอัตราการเตือนผิด (false alarm rate) อยู่ที่ 10% วิธีนี้สามารถให้ข้อมูลเกี่ยวกับการเปลี่ยนแปลงในส่วนบนและส่วนล่างของร่างกายมนุษย์ ผลการทดลองบ่งชี้ว่าเทคนิคที่ใช้ในงานวิจัยนี้สามารถใช้ได้ดีกับการตรวจจับการล้มในชุดภาพวิดีโอแบบเวลาจริง (real-time video sequences)

Miao Yu และคณะ (Miao Yu, et.al 2012) ได้ทำการวิจัยโดยเริ่มจากการลบภาพพื้นหลัง (background subtraction) เพื่อสกัดภาพเบื้องหน้า (foreground) ที่เป็นตัวคนออกมา จากนั้นนำผลลัพธ์ที่ได้ ไปปรับปรุงโดยใช้วิธีการ post-processing ที่เหมาะสม โดยข้อมูลจาก ellipse fitting และ projection histogram ตามแกนของ ellipse ถูกใช้เป็นคุณลักษณะ (features) สำหรับแยกแยะท่าทางต่าง ๆ ของมนุษย์ และคุณลักษณะเหล่านี้จะถูกนำมาป้อนเข้า directed acyclic graph support vector machine (DAGSVM) เพื่อแยกแยะท่าทาง ผลลัพธ์ที่ได้จะนำมาพร้อมกับข้อมูลของพื้นเพื่อตรวจจับการล้ม ผลการทดลองจากข้อมูลของคน 15 คน ระบบนี้สามารถตรวจจับการล้มได้ในอัตราสูงถึง 97.08% และให้ false detection rate ที่ต่ำเพียง 0.8% ภายในสภาพบ้านจำลอง

Damien Brulin และคณะ (Damien Brulin, et.al 2012) ได้นำเสนอวิธีการตรวจจับการล้ม ออกเป็น 2 ส่วนหลักได้แก่ ส่วนที่หนึ่งเป็นการตรวจจับโครงร่างที่เป็นมนุษย์ ภายใต้สภาพแวดล้อมที่แตกต่างไปจากเดิมได้ เช่น ความสว่างของแสงเปลี่ยนไปหรือวัตถุรอบข้างเปลี่ยนตำแหน่งไปจากเดิม เป็นต้น ซึ่งแบ่งออกเป็น 3 ขั้นตอน ได้แก่ ขั้นตอนที่ 1. การตรวจจับการเปลี่ยนแปลง โดยมีสมมติฐานว่าสิ่งที่เปลี่ยนแปลงจากเฟรมภาพปัจจุบันมีความเป็นไปได้สูงที่สิ่งนั้นจะเป็นมนุษย์ จึงทำการลบภาพพื้นหลังด้วยการแทนแต่ละจุดภาพพื้นหลังด้วยวิธี Single Gaussian distribution และวัดความแตกต่างของภาพปัจจุบันกับภาพพื้นหลังด้วยระยะทาง Mahalanobis ขั้นตอนนี้

2. การติดตาม ผลลัพธ์ที่ได้หลังจากการลบภาพพื้นหลังจะเป็นองค์ประกอบที่มีการเชื่อมโยงซึ่งอาจจะเป็นมนุษย์หรือวัตถุอื่นๆที่เคลื่อนไหวได้ ซึ่งหากพิจารณาการแทนที่ขององค์ประกอบที่มีการเชื่อมโยงเหล่านั้นในภาพที่ผ่านมาในอดีตจะทำให้สามารถกำหนดคุณลักษณะของวัตถุที่เราติดตามได้ โดยกำหนดเป็นกลุ่มของจุดที่สนใจ และทำการติดตามกลุ่มของจุดเราสนใจเหล่านี้ไปตลอดทุกเฟรมภาพ จากนั้นจึงนำรายการองค์ประกอบที่เชื่อมโยงที่เป็นผลลัพธ์จากขั้นตอนก่อนหน้าและรายการวัตถุที่เราติดตามการเคลื่อนไหว มาทำการจับคู่กันเพื่อให้การตรวจจับหรือรู้จำสิ่งที่เป็มนุษย์ในเฟรมภาพนั้นมีความแม่นยำมากยิ่งขึ้น และ ขั้นตอนที่ 3. การแบ่งกลุ่ม โดยใช้วิธีของ Viola and Jones โดยเริ่มด้วยการกรองข้อมูลด้วย Adaboost และ Haar-like filters และการแบ่งกลุ่มโครงร่างของมนุษย์ที่อาจมีลักษณะโครงร่างเหมือนมนุษย์ทั้งหมด หรือมีเพียงบางส่วน เช่น มีโครงร่างเหมือนมนุษย์ส่วนที่เป็นครึ่งบนของร่างกาย (ด้านหน้า-หลัง-ซ้าย-ขวา) ซึ่งตัวที่ใช้แบ่งกลุ่มข้อมูลนั้นจะทำการเรียนรู้โดยใช้ชุดข้อมูลมนุษย์ที่เป็นที่รู้จักกันดีชื่อ INRIA

หลังจากทำการตรวจจับโครงร่างที่เป็นมนุษย์แล้ว ส่วนที่สองจะเป็นส่วนของการรู้จำลักษณะท่าทาง เมื่อตรวจพบมนุษย์ในฉากของเฟรมภาพแล้ว ขั้นตอนต่อไปจะเป็นการรู้จำลักษณะท่าทาง โดยแบ่งออกเป็น 2 ขั้นตอนคือ 1. สกัดคุณลักษณะ โดยนำภาพที่ทำการลบพื้นหลังแล้วมาทำการคำนวณหาแกนหลักของร่างกายมนุษย์ โดยใช้วิธีการวิเคราะห์ Principal Component Analysis (PCA) เพื่อหาแนวโปรเจ็คชันที่เป็นตัวแทนที่ดีที่สุดของข้อมูลภาพที่ตรวจพบว่าเป็นมนุษย์นั้น 2. ใช้ Fuzzy Logic System (FLS) ในการจับคู่ข้อมูลเข้ากับผลลัพธ์ที่ต้องการ ซึ่ง FLS ที่ใช้ในงานวิจัยนี้ประกอบไปด้วย 4 ส่วน ได้แก่ Fuzzifier, Rules, Inference system, Defuzzifier ผลการทดลองของวิธีการที่นำเสนอในงานวิจัยนี้พบว่าในส่วนของ การตรวจจับโครงร่างที่เป็นมนุษย์มีอัตราความถูกต้องอยู่ที่ 97% และมีอัตราการตรวจจับผิดพลาดอยู่ที่ 3% และผลการทดลองในส่วนของการรู้จำท่าทางมนุษย์โดยใช้วิธี FLS เมื่อเทียบกับวิธี 1-NN พบว่าวิธี FLS มีความแม่นยำในการรู้จำท่าทางของมนุษย์ได้ดีกว่าไม่ว่าจะเป็นท่าทางห้อยขาหรือท่าทางแบบชันเข่าก็ตาม

Caroline Rougier และคณะฯ (Caroline Rougier, et.al 2011) ได้นำเสนอวิธีการโดยใช้ระยะที่วัดจากระดับพื้นถึงจุดศูนย์กลาง (centroid height) ของคนในการตรวจจับการล้ม โดยแบ่งออกเป็น 3 ขั้นตอน คือ ขั้นตอนที่ 1. การตรวจหาระนาบพื้น (Ground plane detection) โดยคำนวณหาระนาบพื้นของห้องอย่างอัตโนมัติโดยวิธี V-disparity ขั้นตอนที่ 2. การติดตามตัวบุคคล(Person tracking) และ Localization

โดยการหาตัวบุคคลและตำแหน่งในรูปภาพ ขั้นตอนที่ 3. การตรวจจับการล้ม(Fall detection) โดยวิเคราะห์วิถีการเดินแบบสามมิติของคน (person 3D trajectory) โดยแยกความแตกต่างการล้มจากกิจกรรมทั่วไป เพื่อให้ผลการทดลองเกิดความแม่นยำมากขึ้น ระยะที่วัดจากระดับพื้นถึงจุดสูงถ่วงของคนถูกนำมาใช้ในการตรวจจับการล้ม โดยจะให้ผลการตรวจจับที่ดีก็ต่อเมื่อไม่มีวัตถุใดๆมาบังหรือกีดขวางการล้ม หากมีเฟอร์นิเจอร์หรือวัตถุใด ๆ กีดขวางไม่ให้กล้องจับภาพคนที่ล้มได้ อัตราความเร็วในการเคลื่อนที่ของร่างกายก่อนถูกวัตถุค้ำได้ถูกนำมาใช้ในการตรวจจับการล้ม ในการทดลองของงานวิจัยนี้ได้มีการทำชุดข้อมูลสำหรับการสอน(training data set) จากสภาพการเดิน นั่ง ก้มลงอย่างเป็นปกติ เพื่อคำนวณหาค่า thresholds ว่าสถานการณ์ใดเป็นสถานการณ์ปกติ หรือ ผิดปกติ มีการคำนวณระยะที่วัดจากระดับพื้นถึงจุดศูนย์ถ่วงของคน (Dtrain) และความเร็วในการเคลื่อนที่ของคนจากภาพเคลื่อนไหว (Vtrain) และนำสองค่านี้มาใช้ในการหาค่า thresholds ต่อไป ซึ่งผลการทดลองของวิธีที่นำเสนอในงานวิจัยนี้พบว่ามีการตรวจจับการล้มที่ดีโดยอัตราความถูกต้องอยู่ที่ 98.7%

Shaou-Gang Miaou และคณะฯ (Shaou-Gang Miaou, et.al 2006) ได้ศึกษาวิจัยการตรวจจับการล้มของผู้สูงอายุโดยใช้อุปกรณ์รับภาพชื่อว่า MapCam เป็นกล้องชนิด Omni camera ที่สามารถรับภาพวิดีโอมุมมองได้ 360 องศาได้ โดยตรวจจับการล้มได้โดยใช้เทคนิคการประมวลผลภาพแยกตัวบุคคลออกจากภาพในการสร้างกล่องขอบเขตขึ้นมาเพื่อใช้พีเจอร์หลักคือ อัตราระหว่างความยาวและความสูงของคนร่วมกับข้อมูลส่วนตัวของผู้ล้มได้แก่ BMI (Body Mass Index) และประวัติการรักษาโรค เพื่อประกอบใช้ในอัลกอริธึมสำหรับตรวจจับการล้มโดยการปรับค่าพารามิเตอร์ให้มีความเหมาะสมในแต่ละกรณีเพื่อหลีกเลี่ยงการแจ้งเตือนที่ไม่จำเป็นและเป็นการเพิ่มการดูแลเป็นพิเศษสำหรับกรณีและผู้ล้มต้องการรับการดูแลมากกว่าปกติ ผลการทดสอบของวิธีที่นำเสนอพบว่าความถูกต้องสำหรับกรณีที่ไม่ใช้ข้อมูลส่วนบุคคลและใช้ข้อมูลส่วนบุคคลคือ 79.8% และ 69% ตามลำดับ

Peng Shen Ong และคณะฯ (Peng Shen Ong, et.al 2014) นำเสนอวิธีการตรวจจับการล้มที่มีประสิทธิภาพที่สามารถทำงานประมวลผลภาพได้แบบ Real time โดยเลือกใช้อุปกรณ์ลอจิกแบบโปรแกรมได้ (Programmable Gate Array: FPGA) สำหรับประมวลผลภาพ ภาพขนาด 640 x 480 พิกเซล ที่ ซึ่งมีอัลกอริธึมของลำดับการตรวจจับการล้มคือ 1. การตรวจจับการเคลื่อนไหวของวัตถุที่ใช้หลักการของการลบออกของพื้นหลังเพื่อสกัดวัตถุที่เคลื่อนไหวออกมาจากภาพ 2. การลดในส่วนของ

เงาโดยการพิจารณาโดเมนของสีที่อยู่ในรูปแบบของ YCrCb และ 3. การตรวจจับการล้มนำโดยกำหนดกล่องขอบเขตสำหรับเป็นตัวแทนของความสูงและความกว้างของรูปร่างมนุษย์ โดยพิจารณาการล้มนั้นจะเกิดขึ้นเมื่อค่าความสูงมีความน้อยกว่าความกว้าง แต่ไม่สามารถทำงานได้หากการล้มนั้นเกิดขึ้นในแนวขนานกับเส้นมุมมองของตัวกล้อง ซึ่งวิธีการที่นำเสนอนี้สามารถลดอัตราการใช้พลังงานลงได้ถึง 33.33% และทำความเร็วในการประมวลผลได้ถึง 58.36 เฟรมต่อวินาที

Behçet Uğur Töreyn และคณะ^๗ (Behçet Uğur Töreyn, et.al 2005) ได้วิจัยการตรวจจับการล้มนที่ผสมผสานระหว่างการประมวลผลภาพวิดีโอที่ได้จากกล้องวงจรปิดร่วมกับข้อมูลเสียงที่ถูกบันทึกโดยใช้ไมโครโฟน โดยมีลำดับขั้นตอนการตรวจจับตามชนิดของข้อมูลทั้ง 2 ดังนี้

1. การวิเคราะห์ข้อมูลวิดีโอ ผู้วิจัยใช้วิธีการตรวจจับวัตถุที่เคลื่อนไหวโดยใช้การเปรียบเทียบเฟรมก่อนหน้าและเฟรมปัจจุบัน จากนั้นจึงสร้างกล่องขอบเขตเพื่อระบุขนาดรูปร่างโดยรวมของมนุษย์ในภาพขึ้น ซึ่งในขั้นตอนนี้จะทำให้ค่าอัตราส่วนระหว่างความสูงและความกว้างของกล่องเกิดขึ้น ผู้วิจัยได้ใช้ Wavelet transform ทำให้ค่ามีความแม่นยำขึ้น จากนั้นจึงนำไปสู่การจำแนกโดยใช้ Hidden Markov Models (HMM)

2. การวิเคราะห์ข้อมูลเสียง ผู้วิจัยใช้เทคนิค Wavelet transform ที่มีลักษณะเป็น High-pass filter (HPF) กับสัญญาณเสียงเพื่อให้คุณลักษณะของสัญญาณที่มีความสำคัญปรากฏชัดเจนยิ่งขึ้น จากนั้นจึงจำแนกเหตุการณ์โดยใช้ HMM เช่นเดียวกันกับที่ถูกใช้ในการวิเคราะห์ภาพ โดยผลการทดสอบพบว่าการใช้เสียงร่วมในการพิจารณาการล้มนสามารถทำได้เป็นอย่างดี สามารถลดการแจ้งเตือนในรูปแบบ false positive หรือการที่แจ้งเตือนว่าล้มนแต่ไม่เกิดการล้มนจริงได้

Alex Edgcomb และคณะ^๘ (Alex Edgcomb, et.al 2012) นำเสนอการเปรียบเทียบการตรวจจับการล้มนกับภาพวิดีโอที่ถูกปรับปรุงภาพเพื่อความเป็นส่วนตัวประกอบไปด้วยวิดีโอทั้ง 5 ชนิด ได้แก่ วิดีโอต้นฉบับ, วิดีโอที่มีการเบลอตัวบุคคล, วิดีโอที่มีการใช้เงาโครงร่างตัวคน, วิดีโอที่มีการสร้างวงรีแทนโครงร่างตัวคน และวิดีโอที่มีการใช้กล่องขอบเขตล้อมรอบตัวคน งานวิจัยนี้ใช้พีเจอร์ทหลักคือค่าอัตราส่วนระหว่างความสูงและความกว้าง และส่วนกลับของค่าดังกล่าวมาใช้วิเคราะห์เหตุการณ์การล้มนแบบ Binary tree classification เพื่อพิจารณาค่า Dynamic Time Warping (DTW) จากการประมวลผลสัญญาณค่าอัตราส่วนเพื่อใช้สำหรับเปรียบเทียบกราฟที่เกิดขึ้นเพื่อจำแนกเหตุการณ์ของการล้มนได้ ซึ่งผลของการเปรียบเทียบพบว่าชนิดของวิดีโอที่มีการ

สร้างวงรีแทนโครงร่างตัวคนนั้นสามารถทำค่าความถูกต้องการจำแนกเหตุการณ์การล้ม Sensitivity และ Specificity ที่ 91% และ 92% ตามลำดับ

Ji Tao และคณะฯ (Ji Tao, et.al 2005) เสนองานวิจัยเกี่ยวกับการตรวจจับการล้มในสภาพแวดล้อมภายในอาคารสำนักงาน โดยวิธีการตรวจจับถูกแบ่งออกเป็น 2 ส่วนคือ ส่วนแรกคือการตรวจจับและติดตามตัวบุคคลที่ปรากฏอยู่ในภาพวิดีโอโดยใช้การหาความต่างของภาพในกรอบเฟรมเวลา เพื่อที่จะทำการสกัดพีเจอร์ค่าอัตราส่วนของความสูงและความกว้างของกล่องขอบเขตที่ล้อมตัวบุคคลออกมาแสดง ส่วนที่สองคือการอนุมานเหตุการณ์โดยใช้ค่าอัตราส่วนที่ได้จากกระบวนการแรกมาจำแนกเหตุการณ์โดยใช้ Gaussian Mixture Model (GMM) ร่วมกับการกำหนดอัลกอริธึมที่มีคุณสมบัติการตรวจสอบสมมติฐานในรูปแบบ Finite State Machine (FSM) โดยสามารถทำการตรวจจับการล้มได้ในระยะที่ไกลเป็นอย่างดี แต่รูปแบบชนิดของการล้มส่วนใหญ่ที่มีอัตราค่าความถูกต้องสูงจะอยู่ในกลุ่มของชุดข้อมูลทดสอบที่อาสาสมัครล้มด้านข้างของมุมมองกล้องในส่วนใหญ่

Georgios Mastorakis และคณะฯ (Georgios Mastorakis, et.al 2012) และ Vitoantonio Bevilacqua และคณะฯ (Vitoantonio Bevilacqua, et.al 2014) ต่างนำเสนอระบบการตรวจจับการล้มโดยใช้กล้อง Microsoft Kinect ที่มีคุณสมบัติรับภาพวัตถุในมิติของความลึกแบบ 3 มิติได้ ทำให้สามารถแสดงคุณลักษณะของร่างกายคนในเชิงลึกเหนือกว่ากล้องทั่วไป ผู้วิจัยได้อาศัยข้อได้เปรียบในจุดนี้ทำการสร้างตัวกล่องขอบเขตแบบ 3 มิติที่บรรจุล้อมร่างกายของคนขึ้นทำให้สามารถตรวจจับการล้มในหลายทิศทางได้สะดวก โดยทั้งสองงานได้นำเสนอพีเจอร์หลักที่ใช้สำหรับพิจารณาการล้ม คือ ความยาวของเส้นทแยงมุมหรือ WD length ที่เกิดขึ้นระหว่างความกว้างและความลึกของกล่องขอบเขตชนิดนี้ในบริเวณฐานล่างของร่างกาย จากนั้นมีการแปลงสัญญาณข้อมูลโดยใช้อนุพันธ์อันดับที่ 1 ของค่าความยาวเส้นที่เกิดขึ้นร่วมกับการใช้ค่าของความสูงของตำแหน่งศีรษะคน โดยนำค่าที่ได้ทั้งสองค่านี้นำเข้าสู่อัลกอริธึมสำหรับตรวจจับที่มีลักษณะเป็นกฎสำหรับการจำแนกการล้มพิจารณาร่วมกับค่าพารามิเตอร์ที่ถูกเรียนรู้ไว้เรียบร้อยแล้ว ซึ่งสามารถจำแนกเหตุการณ์ได้ดี แต่มีบางเหตุการณ์ที่มีผลกระทบต่อความถูกต้องในการตรวจจับได้ คือ เมื่อจังหวะของการล้มนั้นถูกรบกวนด้วยการเดินหรือการแกว่งแขนของตัวบุคคลนั้นมีผลทำให้กล่องขอบเขตที่มีลักษณะเป็น 3 มิติที่ถูกเปลี่ยนแปลงไปโดยตลอดเวลา ทำให้ผลของการตรวจจับออกมาเป็น false negative หรือมีการแจ้งว่าไม่ล้มแต่ตัวบุคคลได้ล้มลงไปที่พื้นแล้ว

บทที่ 3 วิธีดำเนินการวิจัย

งานวิจัยนี้ได้ทำการศึกษาปัญหาของกล่องขอบเขตที่ถูกนำมาใช้ในงานวิจัยที่เกี่ยวข้องกับการตรวจจับการล้มที่มีอยู่เดิม จึงได้นำเสนอกกล่องขอบเขตแบบใหม่ที่สามารถเป็นตัวแทนของรูปร่างของคนที่ปรากฏขึ้นในตำแหน่งที่เหมาะสมสามารถนำค่าทางคณิตศาสตร์มาใช้เพื่อประยุกต์เข้ากับกรอบงานของระบบการตรวจจับการล้มได้

3.1 กล่องขอบเขตแบบทิศทาง (Directional Bounding Box: DBB)

กล่องขอบเขตแบบทิศทาง (Directional Bounding Box: DBB) คือกล่องที่มีลักษณะที่ครอบคลุมทุกส่วนบนร่างกายโดยจะเกิดขึ้นไปในแนวเดียวกับทิศทางของการล้ม โดยที่กล่องชนิดนี้จะใช้ข้อมูลที่ได้จากอุปกรณ์รับภาพที่เป็นอินฟราเรดเช่นเดียวกับกล่องขอบเขตแบบสามมิติแต่ผลลัพธ์สุดท้ายจะมีเพียงกล่องที่อยู่ในรูปแบบสองมิติเท่านั้น จึงทำให้เหมือนเป็นการสร้างกล่องขอบเขตแบบสองมิติที่เกิดขึ้นจากการมองไปที่ด้านข้างของการล้มอยู่ตลอดเวลา โดยขั้นตอนการสร้างประกอบไปด้วย 4 ขั้นตอนดังนี้

3.1.1 การสร้างเวกเตอร์ของร่างกาย

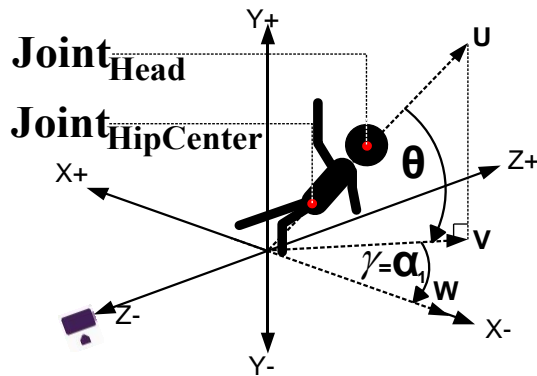
ขั้นตอนนี้จะเป็นขั้นตอนแรกสำหรับการกำหนดทิศทางการเอียงของร่างกายเพื่อที่จะให้ขั้นตอนต่อไปได้รู้ว่าการจัดเรียงตัวของกล่องขอบเขตแบบทิศทางจะอยู่ในแนวใด โดยที่ขั้นตอนนี้จะทำการรับข้อมูลจุดสำคัญบนร่างกาย (Joint point) ที่สกัดออกมาจากเซ็นเซอร์รับภาพที่มีลักษณะเป็นจุดที่อยู่ในพิกัด 3 มิติ ที่มีทั้งแกน X, Y และ Z มาพิจารณาเพียง 2 จุด ได้แก่ head ($Joint_{Head}$) และ hip ($Joint_{HipCenter}$) ซึ่งเหตุผลของการเลือกจุดสองจุดนี้ เพราะว่าเส้นเวกเตอร์ที่เกิดขึ้นขณะล้มนั้นแนวแกนเวกเตอร์ดังกล่าวค่อนข้างที่จะมีการจัดเรียงตัวที่คงเหมาะสม โดยที่เวกเตอร์ของร่างกายจะแทนด้วย U ซึ่งสามารถคำนวณได้จากสมการที่ (3.1) – (3.3)

$$X_Q = X_{head} - X_{hip} \quad (3.1)$$

$$Z_Q = Z_{head} - Z_{hip} \quad (3.2)$$

$$U = (X_Q, Z_Q) \quad (3.3)$$

โดยที่	X_{head}	คือ	ค่าพิกัด X ของจุดบริเวณศีรษะ
	X_{hip}	คือ	ค่าพิกัด X ของจุดบริเวณกึ่งกลางสะโพก
	Z_{head}	คือ	ค่าพิกัด Z ของจุดบริเวณศีรษะ
	Z_{hip}	คือ	ค่าพิกัด Z ของจุดบริเวณกึ่งกลางสะโพก



รูปที่ 3-1 แสดงเวกเตอร์ร่างกาย

จากรูปที่ 3-1 จะเห็นได้ว่าเวกเตอร์ร่างกาย U จะสร้างมุม θ ที่กระทำกับเวกเตอร์ V ที่เป็นเวกเตอร์ที่เกิดจากเวกเตอร์ U ที่ถูกสร้างขึ้นเป็นตัวแทนในแนวระนาบ โดยที่มุมดังกล่าวสามารถคำนวณได้ดังสมการที่ 3.4

$$\theta = \arccos\left(\frac{U \cdot V}{\|U\| \|V\|}\right) \quad (3.4)$$

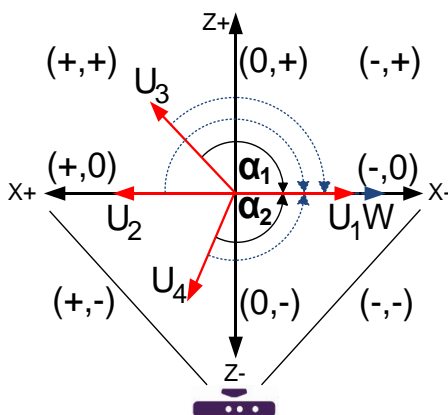
โดยผลลัพธ์ของกระบวนการนี้จะมีการพิจารณาขนาดของมุม θ ที่เกิดขึ้นว่ามีค่าน้อยกว่าที่กำหนดไว้หรือไม่ ซึ่งโดยปกติแล้วแนวแกนร่างกายของคนปกติจะอยู่ที่ประมาณ 90 องศา ซึ่งในงานวิจัยนี้ได้กำหนดเงื่อนไขมุมองศาที่ 80 องศาคือจุดเริ่มต้นที่จะสามารถทำการสร้างกล่องขอบเขตนี้ได้ โดยถ้าตรงเงื่อนไขกระบวนการในขั้นตอนต่อไปจึงจะเริ่มทำงาน

3.1.2 การคำนวณค่ามุมสำหรับการหมุนจุดร่างกาย

เนื่องจากการที่จะทำให้กล่องขอบเขตที่ต้องการอยู่ที่ตำแหน่งเสมือนการมองข้างของการล้มเสมอ การหมุนจุดต่างๆของร่างกายจึงถูกนำมาใช้เพื่อให้จุดเข้าสู่ระบบพิกัดที่สามารถระบบตัวกล่องขอบเขตได้ง่าย ซึ่งในการหมุนนั้นจำเป็นที่จะต้องรู้ค่าขนาดของมุมที่จะหมุน ในขั้นตอนนี้จึงมีหน้าที่สำคัญคือการคำนวณหาค่าขนาดของมุมที่ใช้สำหรับหมุน γ ที่ใช้ในการหมุนในมุมมองจากด้านบนโดยใช้โดยพิจารณาจากควอดรนต์ที่เวกเตอร์ U กระทำไว้จากมุมมองที่ตำแหน่งด้านบนโดยสามารถคำนวณได้จากสมการที่ 3.5 ซึ่งสอดคล้องกับรูปที่ 3-2

$$\gamma = \begin{cases} 0 & \text{if } (X_Q < 0) \text{ and } (Z_Q = 0) \\ -180 & \text{if } (X_Q > 0) \text{ and } (Z_Q = 0) \\ -\alpha_1 & \text{if } (Z_Q > 0) \\ \alpha_2 & \text{if } (Z_Q < 0) \end{cases} \quad (3.5)$$

โดยที่ X_Q คือ ค่าพิกัด X ของบริเวณควอดรนต์
 Z_Q คือ ค่าพิกัด Z ของบริเวณควอดรนต์



รูปที่ 3-2 แสดงไดอะแกรมสำหรับคำนวณมุมสำหรับการหมุนจากเวกเตอร์ U

ซึ่งหลังจากที่ค่าขนาดของมุมที่ใช้สำหรับการหมุนถูกคำนวณเรียบร้อยแล้ว ค่าของมุมดังกล่าวจะถูกนำไปใช้ในกระบวนการหมุนแกนในขั้นตอนต่อไป

3.1.3 การหมุนกลุ่มจุดสำคัญของร่างกายให้อยู่ในแนวตั้งฉากกับมุมมอง

หลังจากได้ค่ามุม γ สำหรับใช้หมุนแล้ว จึงทำการหมุนจุดสำคัญของร่างกายทุกจุดเหมือนการหมุนเวกเตอร์ร่างกาย V รอบแกน Y โดยได้มีการทำการปรับปรุงเมตริกซ์ที่ใช้สำหรับการหมุนในแกนตั้งในสมการ 3.6 ให้สามารถใช้จุดศูนย์กลางมวลเป็นจุดหมุน (Center of gravity) ที่ประกอบด้วยค่าจุดพิกัด 3 ค่าคือ $(x_{COG}, y_{COG}, z_{COG})$ ให้มีรูปแบบใหม่ดังสมการที่ 3.7 เพื่อให้ได้ผลลัพธ์เวกเตอร์ W ที่ตั้งฉากกับมุมมองของตัวเซนเซอร์รับภาพซึ่งสามารถคำนวณได้จากสมการที่ 3.8

$$R_y(\gamma) = \begin{bmatrix} \cos(\gamma) & 0 & \sin(\gamma) \\ 0 & 1 & 0 \\ -\sin(\gamma) & 0 & \cos(\gamma) \end{bmatrix} \quad (3.6)$$

$$R_{COG}(\gamma) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -x_{COG} & -y_{COG} & -z_{COG} & 1 \end{bmatrix} \begin{bmatrix} \cos(\gamma) & 0 & \sin(\gamma) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\gamma) & 0 & \cos(\gamma) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.7)$$

$$W = V * R_{COG}(\gamma) \quad (3.8)$$

การหมุนดังกล่าวสามารถอธิบายในมุมมองของจุดบนร่างกายในระบบโดยใช้สมการ 3.9 โดยที่ Pr_i คือจุดสำคัญบนร่างกายที่ถูกหมุนแล้วตัวที่ i ซึ่ง i จะมีค่าตั้งแต่ 1 ถึงจุดที่ n ตัว

$$Pr_i = P_i * R_{COG}(\gamma) \quad (3.9)$$

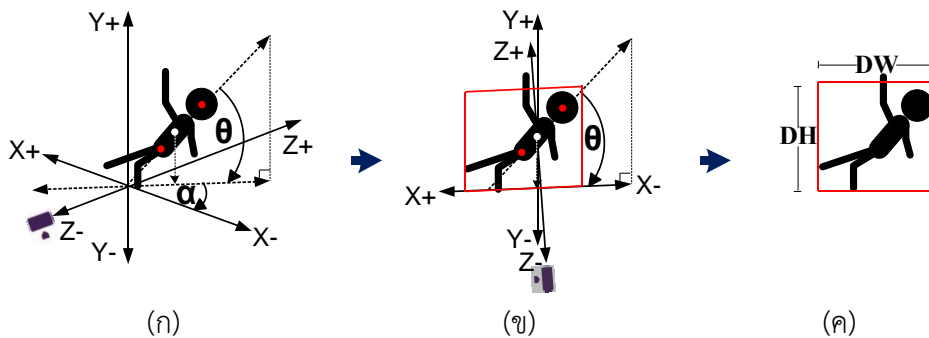
3.1.4 การระบุตำแหน่งและขนาดของกล่องขอบเขต

หลังจากที่จุดสำคัญบนร่างกายถูกหมุนให้อยู่ในแนวระนาบที่ตั้งฉากกับมุมมองแล้ว ต่อไปคือการระบุขอบเขตของกล่องที่ล้อมรอบร่างกายของตัวคนได้ครบทุกจุด โดยใช้การพิจารณาเพียงจุดพิกัด X และ Y เพื่อสร้างจุดบนขวาสุด P_{TR} และจุดล่างซ้ายสุด

P_{BL} ดังสมการที่ 3.10 และ 3.11 เพื่อให้สามารถสร้างกรอบของกล่องขอบเขตแบบทิศทางทางนี้ได้ดังปรากฏดังภาพที่ 3-3 (ค)

$$P_{TR} = [MAX(Prx_1, Prx_2, \dots, Prx_{20}), MAX(Pry_1, Pry_2, \dots, Pry_{20})] \quad (3.10)$$

$$P_{BL} = [Min(Prx_1, Prx_2, \dots, Prx_{20}), MIN(Pry_1, Pry_2, \dots, Pry_{20})] \quad (3.11)$$



รูปที่ 3-3 แสดงภาพรวมของการสร้างกล่องขอบเขตแบบทิศทาง

จากรูปที่ 3-3 จะแสดงให้เห็นถึงภาพรวมสำหรับการสร้างกล่องขอบเขตที่นักวิจัยนำเสนอ จากภาพที่ 3-3 (ก) จะแสดงถึงจุดเริ่มต้นของการสร้างกล่องขอบเขตโดยพิจารณาตำแหน่งของการเอียงของร่างกายมนุษย์ที่กระบวนการนี้จะทำให้รู้ว่าต้องมีการปรับมุมมองไปในทิศทางใดและจะต้องหมุนจุดร่างกายไปในขนาดมุมเท่าไร จากนั้นมาในภาพที่ 3-3 (ข) จะเห็นได้ว่าแกนเวกเตอร์ร่างกายบนแนวระนาบนั้นจะถูกทดแทนด้วยแนวแกน X อันเป็นผลมาจากกลุ่มจุดของร่างกายได้ถูกหมุนมาตั้งฉากกับมุมมองเรียบร้อยแล้วและยังสามารถหากล่องขอบเขตแบบทิศทาง(กรอบสีแดง)ได้ง่ายเพราะค่าอยู่ในระนาบที่เหมาะสมในการคำนวณกรอบขอบเขตดังภาพที่ 3-3 (ค)

3.2 พีเจอร์ทิศทางและพีเจอร์เสริมของกล่องขอบเขตแบบทิศทาง (Directional Bounding Box Feature)

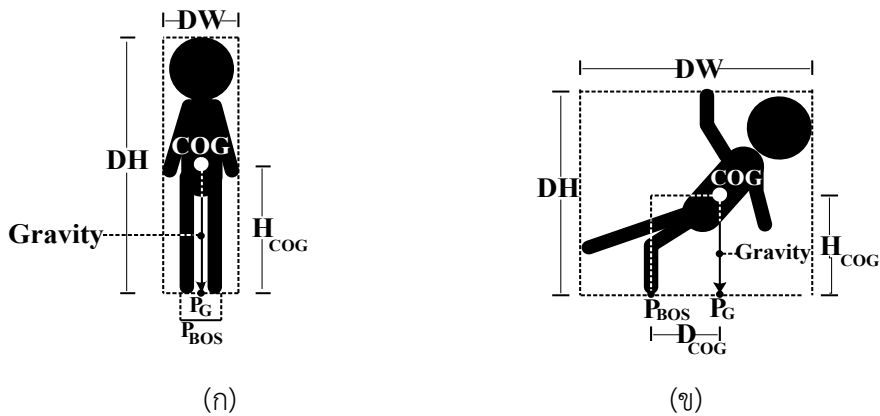
งานวิจัยนี้ใช้พีเจอร์ทิศทางจากกล่องขอบเขตแบบทิศทางคือ ค่าอัตราส่วนระหว่างความสูงและความกว้างของกล่อง (R_{DHW}) เช่นเดียวกับงานวิจัยที่มีการใช้กล่องขอบเขตแบบสองมิติ จากที่แสดงดังภาพที่ 3-3 (ค) โดยที่ค่าอัตราส่วนนี้สามารถ

คำนวณหาได้จากสมการที่ 3.12 โดยที่ DH และ DW คือความสูงและความกว้างของกล่องขอบเขตแบบทศทางตามลำดับ

$$R_{DHW} = \frac{DH}{DW} \quad (3.12)$$

นอกจากฟีเจอร์หลัก (Main feature) ที่นำเสนอแล้วในงานวิจัยนี้ยังนำเสนอฟีเจอร์เสริม (Additional feature) เพื่อช่วยสนับสนุนการตรวจจับการล้มอีกทางหนึ่งเช่นกัน โดยฟีเจอร์ดังกล่าวได้อาศัยการพิจารณาจุดศูนย์กลางของมวลมาประยุกต์ใช้ร่วมกับท่าทางการเคลื่อนไหวของร่างกาย

โดยปกติแล้วการที่วัตถุจะเกิดความสมดุลไม่เกิดการเอียงล้มจากแรงที่ถูกระงับโดยแรงดึงดูดของโลกนั้น จุดศูนย์กลางของมวลจำเป็นต้องอยู่เหนือภายในบริเวณที่เรียกว่าฐานสนับสนุน (Base of support: BOS) เช่นแสดงดังรูปที่ 3-4 (ก) โดยที่จุด P_G เป็นจุดร่วมที่เกิดจากเส้นโปรเจกชันของจุดศูนย์กลางมวลของร่างกายลงมาถึงพื้นในแนวตั้งฉาก โดยมีระยะห่างจากพื้นคือ H_{COG} ค่าจุดดังกล่าวนี้จะถูกนำมาใช้พิจารณาสถานะความสมดุลของร่างกายว่าอยู่ในบริเวณฐานสนับสนุนหรือไม่ ถ้าหากร่างกายมีความสมดุลจะอยู่ในบริเวณของฐานสนับสนุน P_{BOS} ซึ่งจะเป็นบริเวณพื้นที่ของเท้าที่สัมผัสที่พื้น ณ ขณะเวลานั้น



รูปที่ 3-4 แสดงตำแหน่งของจุดศูนย์กลางมวลของร่างกาย

(ก) อยู่เหนือบริเวณฐานสนับสนุน และ (ข) อยู่นอกบริเวณฐานสนับสนุน

สำหรับกรณีที่ร่างกายเกิดการเสียสมดุลดังแสดงในรูปที่ 3-4 (ข) จุด P_G จะอยู่นอกบริเวณฐานสนับสนุน P_{BOS} ตำแหน่งดังกล่าวนี้จะทำให้เกิดค่าระยะทางที่เกิดขึ้น

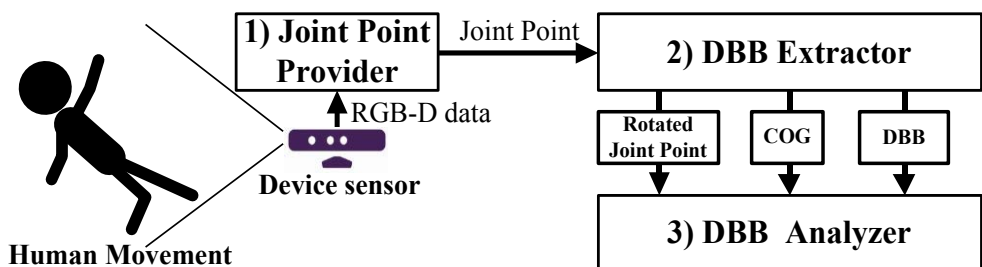
ระหว่างจุด P_G และ P_{BOS} ที่เรียกว่า D_{COG} ที่มีความสัมพันธ์กับโอกาสของการล้ม ซึ่งถ้าหากค่าดังกล่าวมีค่ามากเท่าไร โอกาสที่จะเกิดการล้มจะมีสูงมากเท่านั้น

ฟีเจอร์เสริม (Additional feature) ที่งานวิจัยนี้แนะนำเสนอคือ ค่าอัตราส่วนระหว่างความสูงของจุดศูนย์กลางมวลเทียบกับพื้นและระยะทางระหว่างโปรเจกชันของจุดศูนย์กลางมวลเทียบกับฐานสนับสนุน (R_{COG}) ดังสามารถคำนวณได้จากสมการที่ 3.13 โดยเหตุผลที่เลือกใช้เป็นค่าของอัตราส่วนก็เพราะค่า D_{COG} ที่เกิดขึ้นในแต่ละตัวบุคคลที่มีความสูงต่างกันจะมีค่าที่ไม่เท่ากัน ค่าระยะห่างค่าหนึ่งของคนที่สูงจะใช้เป็นตัวแทนสำหรับคนที่ความสูงน้อยกว่าไม่ได้ ดังนั้นการใช้ค่าสัดส่วนจะทำให้ปรับเข้ากับความเสี่ยงของบุคคลที่มีความแตกต่างกันได้เป็นอย่างดี

$$R_{COG} = \frac{D_{COG}}{H_{COG}} \quad (3.13)$$

3.3 เฟรมเวิร์คสำหรับการตรวจจับการล้ม

ผู้วิจัยได้ออกแบบกรอบงานที่ใช้สำหรับจัดลำดับขั้นกระบวนการตรวจจับการล้มออกเป็นทั้ง 3 กระบวนการหลักที่สำคัญด้วยกัน ดังแสดงในรูปที่ 3-5 โดยแต่ละกระบวนการจะมีหน้าที่ดังต่อไปนี้

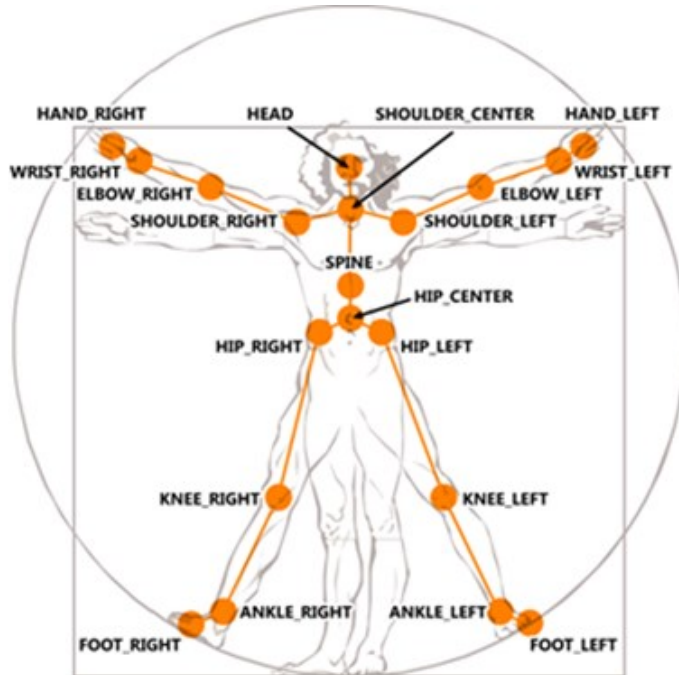


รูปที่ 3-5 แสดงกรอบงานของระบบการตรวจจับการล้มที่ผู้วิจัยนำเสนอ

3.3.1 Joint Point Provider

บทบาทของกระบวนการนี้คือ การผลิตจุดสำคัญของร่างกาย (Joint point) จากภาพที่มีลักษณะที่เป็นภาพเชิงลึก (Depth image) จากกล้อง Microsoft Kinect ชนิด RGB-D camera โดยใช้ไลบรารีมาตรฐานที่ใช้สำหรับการพัฒนาโปรแกรม

ประยุกต์เพื่อจะสกัดจุดดังกล่าวทั้งหมด 20 จุดดังภาพที่ 3-6 เพื่อส่งให้แก่กระบวนการถัดไป



รูปที่ 3-6 แสดงจุดสำคัญบนร่างกายที่ถูกสร้างเมื่อผ่านกระบวนการ Joint point provider

3.3.2 DBB Extractor

กระบวนการนี้มีหน้าที่ในการสร้างกล่องขอบเขตแบบทิศทาง ที่งานวิจัยนี้ได้ใช้เป็นเทคนิคหลักสำหรับประยุกต์เข้ากับการตรวจจับการล้ม โดยที่รายละเอียดสำหรับการสร้างกล่องดังกล่าวถูกอธิบายแล้วในหัวข้อที่ 3.1

3.3.3 DBB Analyzer

การที่ระบบการตรวจจับการล้มจะสามารถทำการตัดสินใจได้ว่าการเคลื่อนไหวแบบใดที่ถูกจำแนกได้ว่าล้มหรือไม่ล้มนั้นจะได้คำตอบที่กระบวนการในส่วนนี้ ซึ่งในงานวิจัยชิ้นนี้ได้ออกแบบการตรวจจับการล้มโดยใช้ Finite state machine (FSM) โดยจะทำการพิจารณาค่าพีเจอร์หลักและพีเจอร์เสริมให้ตรงกับเงื่อนไขถูกออกแบบไว้ดังรูปภาพที่ 3-7

งานวิจัยนี้ได้อาศัยรูปแบบการรู้จำค่าพารามิเตอร์เทอร์สโสลต์ที่ใช้สำหรับการตรวจสอบเงื่อนไขการเลื่อนสถานะปัจจุบันของเครื่องสถานะจำกัด ผู้วิจัยจึงได้ทำการฝึกทดลองเพื่อเก็บค่าที่เหมาะสม โดยที่งานวิจัยนี้ได้ให้ความสนใจไปที่การแจ้งเตือนก่อนที่จะล้มลงไปบนพื้น จากการเก็บข้อมูลสำหรับทดสอบในเบื้องต้นพบว่าเวลาที่ใช้ในการล้มอยู่ที่ 51.24 เฟรมโดยเฉลี่ยหรือ 1.7 วินาที สำหรับค่าพารามิเตอร์ที่นำไปใช้ในอัลกอริธึม ผู้วิจัยใช้ค่าเทอร์สโสลต์พารามิเตอร์ทั้งหมด 4 ค่า ดังนี้

1. The Critical frame number: CFN = 22 เฟรมหรือ 0.73 วินาที

ค่าจำนวนเฟรมวิกฤตนี้ใช้สำหรับการกำหนดขอบเขตของเวลาที่ใช้สำหรับการตรวจจับการล้มซึ่งถ้าเงื่อนไขที่กำหนดนั้นผ่านตรงตามเงื่อนไขของตนเองและอยู่ในเวลาดังกล่าวนี้ถือว่าการล้มโดยสมบูรณ์

2. The upper bound of R_{DHW} : TU = 3.21 หน่วย

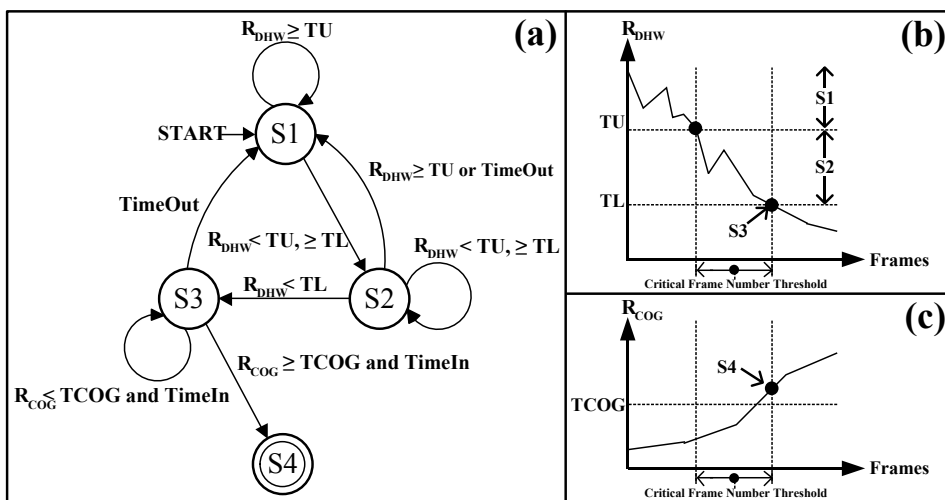
ค่าขอบเขตบนของ R_{DHW} : ค่านี้จะเป็นค่าที่ใช้พิจารณาว่าร่างกายของมนุษย์เริ่มที่จะมีแนวโน้มของการล้มเกิดขึ้นแล้ว

3. The lower bound of R_{DHW} : TL = 1.38 หน่วย

ค่าขอบเขตล่างของ R_{DHW} ค่านี้จะเป็นค่าที่ใช้พิจารณาว่าร่างกายมนุษย์ใกล้ถึงจุดที่จะเกิดการล้มเกิดขึ้นแล้ว

4. The threshold of R_{COG} : TCOG = 0.27

ค่าเทอร์สโสลต์ของอัตราส่วนของตำแหน่งจุดศูนย์กลางมวล ค่านี้จะถูกนำมาใช้เพื่อยืนยันอีกครั้งว่าการล้มเกิดขึ้นจริงหรือไม่



รูปที่ 3-7 (a) เครื่องสถานะจำกัดของอัลกอริธึมสำหรับการตรวจจับการล้ม

- (b) กราฟแสดงค่าของ R_{DHW} ที่มีการเลื่อนจากสถาน S1 ไปยัง S2 และ S3
 (c) กราฟแสดงค่าของ R_{COG} ที่มีค่าตรงตามเงื่อนไขและถูกเลื่อนไปยังสถานะ S4

ไดอะแกรมเครื่องจักรสถานะจำกัดสำหรับการจำแนกการล้มประกอบด้วยทั้งหมด 4 สถานะด้วยกันได้แก่ S1, S2, S3 และ S4 ดังแสดงในภาพที่ 3-7 (a)

1. S1 เป็นสถานะที่ร่างกายมนุษย์ยังไม่เกิดการล้ม
2. S2 เป็นสถานะที่ร่างกายเริ่มมีแนวโน้มว่าเป็นไปได้ว่าจะล้ม
3. S3 เป็นสถานะที่ร่างกายถึงจุดที่ใกล้จะล้มที่สุดแล้ว
4. S4 เป็นสถานะที่ร่างกายจะเกิดการล้มขึ้นแน่นอนแล้ว

โดยการที่ระบบจะแจ้งเตือนการล้มที่เกิดขึ้นได้จำเป็นต้องมีการเลื่อนสถานะปัจจุบันของเครื่องจักรสถานะจำกัดจาก S1 ไปยัง S4 ให้ได้โดยมีการพิจารณาค่าพีเจอร์ 2 ตัวได้แก่ R_{DHW} และ R_{COG} โดยผ่านสถานะตามลำดับดังนี้

1. S1 ในส่วนแรกนั้นจะยึดการพิจารณาค่าของ R_{DHW} เป็นหลักโดย
 - ถ้าหากมีค่าที่เหนือกว่าหรือเท่ากับค่า TU ให้คงสถานะเดิม
 - ถ้าหากมีค่าน้อยกว่า TU แล้วให้เลื่อนสถานะปัจจุบันไปที่สถานะที่ S2 ซึ่งในขั้นตอนนี้ระบบจะเริ่มมีการจับเวลาเริ่มต้นของเหตุการณ์ (Accumulating time)
2. S2. ในสถานะนี้ยังคงจะพิจารณาค่าของ R_{DHW} อยู่โดย
 - ถ้าหากมีค่าที่สูงกว่า TU หรือ accumulating time มีค่ามากกว่า CFN ให้เลื่อนสถานะปัจจุบันไปที่ S1
 - ถ้าหากมีค่าน้อยกว่า TU แต่มากกว่าหรือเท่ากับ TL และเวลาที่ใช้อย่างไม่เกิน CFN ให้คงที่อยู่ในสถานะเดิม
 - ถ้าหากมีค่าน้อยกว่า TL และเวลาที่ใช้อย่างไม่เกิน CFN ให้เลื่อนสถานะปัจจุบันไปที่ S3 ซึ่งเป็นสถานะสุดท้ายของการพิจารณาค่านี้
3. S3 ในสถานะนี้จะพิจารณาค่าของ R_{COG} โดย
 - ถ้า accumulating time มากกว่า CFN ให้เลื่อนไปยัง S1
 - ถ้าหากมีค่าน้อยกว่า TCOG และ accumulating time น้อยกว่าหรือเท่ากับ CFN ให้คงสถานะปัจจุบันที่ตำแหน่งเดิม

- ถ้าหากมีค่าที่สูงกว่าหรือเท่ากับ TCOG และ accumulating time น้อยกว่าหรือเท่ากับ CFN ให้เลื่อนสถานะปัจจุบันไปที่ S4 ซึ่งเป็นสถานะที่ชี้ได้ว่าการล้มขึ้น

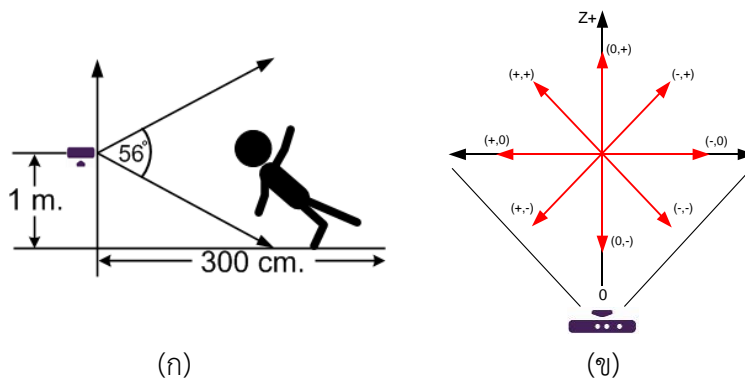
จะเห็นได้ว่าอัลกอริทึมที่นำเสนอทำงานได้สะดวกและง่ายโดยการพิจารณาค่าทั้งสองพีเจอร์เทียบกับค่าเทรสโฮลด์ที่ได้ถูกเรียนรู้ไว้โดยการล้มจะเกิดขึ้นก็ต่อเมื่อทั้งพีเจอร์ผ่านตามเงื่อนไขตามกำหนดของค่าและกรอบของระยะเวลาที่กำหนด

3.4 การจัดสภาพแวดล้อมและชุดข้อมูลสำหรับการทดสอบระบบ

ผู้วิจัยได้มีการจัดสภาพแวดล้อม การสร้างชุดข้อมูล และการกำหนดรูปแบบการทดสอบ สำหรับใช้ในการทดสอบระบบการตรวจจับการล้ม เพื่อเปรียบเทียบผลของการใช้กล่องขอบเขตแบบทิศทางกับงานวิจัยเดิมที่มีการใช้เทคนิคการสร้างกล่องเขตแบบสองมิติและสามมิติดังนี้

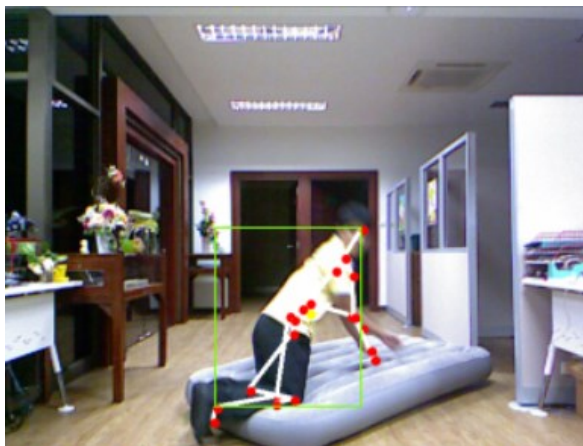
3.4.1 การจัดสภาพแวดล้อม

ผู้วิจัยได้จัดวางกล่องไว้ชิดด้านผนังของห้องที่ใช้สำหรับทดสอบ โดยกำหนดระยะห่างจากพื้นที่ 1 เมตรและฉายออกไปในแนวระดับสายตา ขณะที่ร่างกายของมนุษย์ควรรออยู่ในระยะที่ 1.5 ถึง 3 เมตร ซึ่งตำแหน่งดังกล่าวมีความใกล้เคียงกับงานวิจัยที่นำมาใช้สำหรับเปรียบเทียบและมีเหมาะสมที่จะติดตามการเคลื่อนไหวของร่างกาย โดยเฉพาะการตรวจจับสถานะของร่างกายขณะที่เกิดการล้มได้เป็นอย่างดี ดังรูปที่ 3-8 (ก)



รูปที่ 3-8 (ก) การจัดวางตำแหน่งของอุปกรณ์รับภาพและตำแหน่งของผู้ล้ม
(ข) ทิศทางการล้มทั้ง 8 ทิศที่อาสาสมัครทำการจำลองเพื่อสร้างชุดข้อมูลทดสอบ

นอกจากนี้เพื่อเกิดความปลอดภัยที่เกิดจากการทดสอบการล้มลงของผู้อาสาสมัครผู้วิจัยได้มีการจัดเตรียมเบาะรองรับที่มีคุณสมบัติรองรับการล้มได้เป็นอย่างดีจัดไว้ในบริเวณตำแหน่งที่จะมีการล้มเกิดขึ้นดังภาพที่ 3-9



รูปที่ 3-9 ภาพแสดงเบาะรองรับที่ใช้สำหรับการจำลองการล้ม

3.4.2 การจัดชุดข้อมูลสำหรับทดสอบ

ผู้วิจัยได้สร้างชุดทดสอบระบบโดยใช้อาสาสมัครจำนวน 5 คน เพื่อสร้างกลุ่มทดสอบ โดยแบ่งออกเป็น 2 กลุ่มคือ กลุ่มที่เกิดการล้ม และกลุ่มที่เกิดการโน้มเอียงแต่

ไม่เกิดการล้มนำทำการจำลองการล้มนำใน 8 ทิศทางดังรูปที่ 3-8 (ข) โดยที่แต่ละคนจะทำการล้มนำหรือไม่นำเอียงในแต่ละทิศทางจำนวน 5 ครั้งดังรายละเอียดดังตารางที่ 3-1

ตารางที่ 3-1 รายละเอียดของชุดข้อมูลสำหรับการทดลอง

อาสาสมัคร	เพศ	ส่วนสูง	ล้มนำ	ไม่ล้มนำ
1	ชาย	178	40	40
2	หญิง	167	40	40
3	ชาย	160	40	40
4	หญิง	153	40	40
5	ชาย	150	40	40
รวม			400	

3.4.3 รูปแบบสำหรับการทดสอบ

สำหรับรูปแบบการทดสอบที่ผู้วิจัยได้ให้ความสนใจสามารถจำแนกได้เป็น 3 มุมมองดังต่อไปนี้

3.4.3.1 ด้านของความถูกต้องในการจำแนกเหตุการณ์ล้มนำหรือไม่ล้มนำ นั้นการที่ระบบจะสามารถจำแนกได้ว่าการเคลื่อนไหวแบบใดคือการล้มนำหรือไม่ล้มนำถือว่าเป็นตัวชี้วัดสำคัญว่าระบบที่ถูกปรับปรุงโดยเทคนิคที่นำเสนอขึ้นทำงานได้แม่นยำเพียงใด หากเทียบกับเทคนิคอื่นที่นำเสนอมาก่อนหน้า งานวิจัยนี้ได้ผลการทดสอบโดยใช้ 3 ตัวชี้วัดสำคัญได้แก่ ค่าความถูกต้อง = Accuracy, อัตราการทายล้มนำได้ถูก = Sensitivity และ อัตราการทายล้มนำได้ถูก = Specificity ในการเปรียบเทียบ

3.4.3.2 ด้านของเวลาที่ใช้ในการตอบสนองต่อการล้มนำในแต่ละทิศทาง การวิจัยนี้ได้ให้ความสำคัญกับประสิทธิภาพของการตอบสนองของระบบที่สามารถตรวจจับการล้มนำในทิศทางของการล้มนำที่ต่างกัน ซึ่งถ้าหากระบบสามารถตรวจจับการล้มนำมากเท่าไรจะมีส่วนช่วยที่จะให้ผู้ล้มนำนั้นได้รับการช่วยเหลือได้รวดเร็วมากขึ้นเท่านั้น การทดสอบในมุมมองนี้จะใช้กลุ่มผลการทดสอบของชุดข้อมูลที่มีผลแบบ true positive (แจ้งเตือนว่าล้มนำได้ถูกต้อง) มาใช้เพื่อทำการเปรียบเทียบในแต่ละเทคนิคที่นำเสนอไว้ก่อนหน้า

3.4.3.3 ด้านของเวลาที่ใช้ในการตอบสนองต่อการล้มที่เกิดจากผลกระทบจากการแกว่งแขนหรือการก้าวเดิน การทดสอบเปรียบเทียบอีกแง่ที่นำเสนอซึ่งเป็นการเปรียบเทียบว่าการแกว่งแขนหรือการเดินมีผลต่อการตอบสนองการตรวจจับได้เฉลี่ยช้าที่สุดที่เวลาใด นั้นหมายความว่าขณะที่เกิดการล้มนั้นผลกระทบของการแกว่งแขนหรือการเดินเกิดขึ้นที่เวลาเท่าใดจึงจะทำให้ยังสามารถตรวจจับการล้มได้อยู่ การทดสอบนี้จะจำแนกโดยใช้กลุ่มชุดข้อมูลที่มีผลแบบ true positive (แจ้งเตือนว่าล้มได้ถูกต้อง) โดยแยกเป็นเหตุการณ์ 2 เหตุการณ์คือ กลุ่มที่มีผลกระทบของแขนหรือการก้าวเดินเกิดขึ้นระหว่างเกิดการล้มและกลุ่มที่เกิดการล้มโดยที่แขนหรือการก้าวเดินไม่เกิดขึ้นระหว่างเกิดการล้ม

บทที่ 4 ผลการทดลอง

4.1 ผลการทดสอบความถูกต้องในการจำแนกเหตุการณ์ ล้มหรือไม่ล้ม

การทดลองจะทำการเปรียบเทียบประสิทธิภาพของการนำกล่องขอบเขตชนิด 2 มิติ และ 3 มิติมาใช้ตรวจจับการล้มโดยทำการเลือกตัวอย่างงานวิจัยที่เลือกเป็นตัวแทนในแต่ละรูปแบบชนิดของกล่องขอบเขตที่มีการใช้พีเจอร์หลักและพีเจอร์เสริมเพื่อนำเข้าประมวลผลในอัลกอริธึม ซึ่งงานวิจัยทั้งสองแบบนี้ต่างมีวิธีการตรวจจับในรูปแบบที่มีการเรียนรู้ก่อนในกลุ่มข้อมูลที่เตรียมไว้ในทุกลักษณะของทิศทางการล้ม

ตาราง 4-1 แสดงผลการตรวจจับการล้มของกล่องขอบเขตแบบสองมิติ พบว่าการแจ้งเตือนว่าไม่ล้มในขณะที่เกิดการล้ม(FN) จะเกิดในกรณีที่มีทิศการล้มออกจากตัวกล่อง สำหรับผลการตรวจว่าล้มแต่ไม่เกิดการล้มจริงได้ผิดพลาดนั้นส่วนใหญ่เกิดขึ้นในทิศทางการล้มด้านข้างที่มีทิศทางมาทางด้านหน้า (FP)

ตารางที่ 4-1 ผลการทดสอบความถูกต้องสำหรับงานวิจัยที่ใช้เทคนิคกล่องขอบเขตแบบสองมิติ

ทิศทางการล้ม	ชุดข้อมูลทดสอบที่ไม่ล้ม			ชุดข้อมูลทดสอบที่ล้ม		
	จำนวน	TN	FP	จำนวน	TP	FN
หน้า	25	20	5	25	25	0
หลัง	25	20	5	25	8	17
ซ้าย	25	2	23	25	25	0
ขวา	25	4	21	25	25	0
หน้าขวา	25	2	23	25	25	0
หน้าซ้าย	25	1	24	25	25	0
หลังซ้าย	25	20	5	25	25	0
หลังขวา	25	20	5	25	25	0
รวม	200	89	111	200	183	17



รูปที่ 4-1 (ก) แสดงทิศการล้มของคนที่มีทิศทางออกจากตัวกล่อง
 (ข) แสดงการโน้มตัวของตัวคนมีทิศทางเฉียงทางด้านหน้า

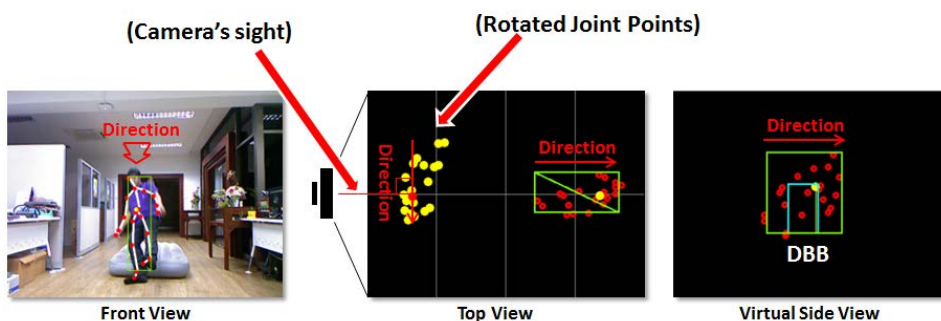
จะเห็นได้ว่าในรูปที่ 4-1 (ก) แสดงถึงรูปร่างของกล่องขอบเขตที่เป็นชนิดแบบสองมิตินั้นจะไม่สามารถแสดงได้ถึงลักษณะของการล้มลงได้ เพราะข้อจำกัดของตัวกล่องจึงทำให้การแจ้งเตือนการล้มจะมีผลแจ้งว่าไม่เกิดการล้มขึ้นในขณะที่ร่างกายของอาสาสมัครได้ล้มลงไปนอนกับพื้นแล้ว ในขณะที่ภาพ (ข) แสดงให้เห็นว่าในเหตุการณ์ที่ร่างกายได้มีการโน้มตัวไปทางด้านข้างหรือเฉียงทางด้านหน้านั้น เทคนิคการใช้กล่องขอบเขตแบบสองมิตินี้จะมีการแจ้งเตือนว่าเกิดการล้มขึ้นซึ่งเป็นการแจ้งเตือนที่ไม่ถูกต้อง

ตาราง 4-2 แสดงผลการตรวจจับการล้มของกล่องขอบเขตแบบสามมิติ ซึ่งสามารถทำการจำแนกเหตุการณ์ของการล้มได้เป็นอย่างดี ผลของการทดลองแสดงปรากฏว่าผลของค่า FP และ FN อยู่ในจำนวนที่เล็กน้อยเท่านั้น

ตารางที่ 4-2 ผลการทดสอบความถูกต้องสำหรับงานวิจัยที่ใช้เทคนิคกล่องขอบเขตแบบสามมิติ

ทิศทางการล้ม	ชุดข้อมูลทดสอบที่ไม่ล้ม			ชุดข้อมูลทดสอบที่ล้ม		
	จำนวน	TN	FP	จำนวน	TP	FN
หน้า	25	24	1	25	24	1
หลัง	25	24	1	25	23	2
ซ้าย	25	24	1	25	23	2
ขวา	25	25	0	25	23	2
หน้าขวา	25	25	0	25	24	1
หน้าซ้าย	25	24	1	25	23	2
หลังซ้าย	25	24	1	25	23	2
หลังขวา	25	24	1	25	24	1
รวม	200	194	6	200	187	13

ภาพรวมของการใช้เทคนิคการใช้อ็กล่องขอบเขตแบบมีทิศทางสามารถแสดงได้แบบสังเขปดังภาพที่ 4-2 จากภาพแสดงขณะที่กำลังจะเกิดการล้มขึ้นที่มีทิศทางเอียงร่างกายออกจากตัวกล้อง โดยใน Front View จะเห็นได้ว่าหากใช้อ็กล่องขอบเขตแบบสองมิติจะเห็นได้ว่าไม่สามารถบอกถึงสถานะที่แท้จริงได้



รูปที่ 4-2 ภาพแสดงการใช้งานกล่องขอบเขตแบบทิศทาง

สำหรับภาพ Top View แสดงการล้มในมุมมองแบบมุมสูง ซึ่งจะแสดงให้เห็นถึงเทคนิคการสร้างกล่องขอบเขตแบบทิศทางที่มีการหมุนกลุ่มจุดร่างกายสีแดงทั้งหมด

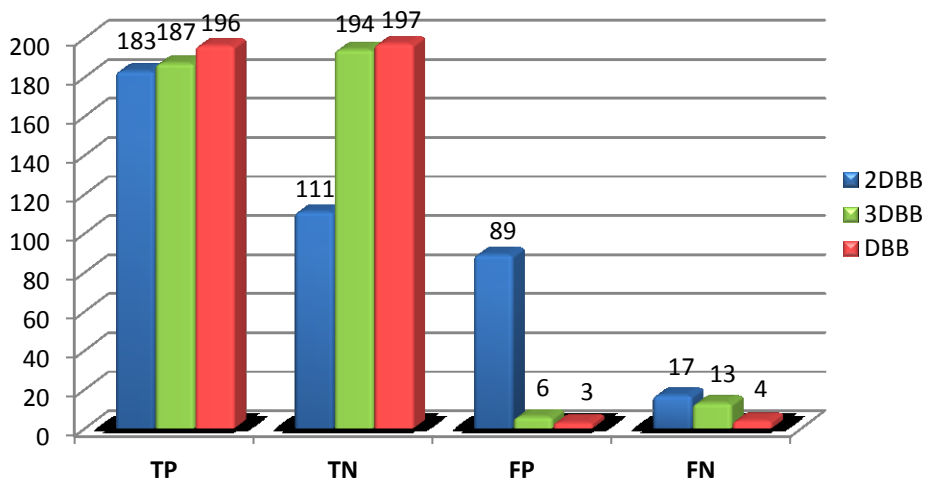
ให้อยู่ในแนวระนาบตั้งฉากกับมุมมองของกล้อง (Camera's sight) ซึ่งแสดงเป็นกลุ่มจุดสีเหลือง ซึ่งประโยชน์ของการหมุนนี้จะทำให้สามารถหากล่องขอบเขตแบบทิศทางได้สะดวกขึ้นดังแสดงในในกรอบสีเขียวดังแสดงในภาพ Virtual Side View จะเห็นว่านอกจากกล่องขอบเขตที่นำเสนอแล้วจะมีกล่องขอบเขตสีฟ้าที่นำเสนอตำแหน่งของจุดศูนย์กลางมวลร่างกายที่ใช้เป็นค่าพีเจอร์เสริมสำหรับพิจารณาการล้มอีกด้วย

ตารางที่ 4-3 ผลการทดสอบความถูกต้องสำหรับงานวิจัยที่ใช้เทคนิคกล่องขอบเขตแบบทิศทาง

ทิศทางการล้ม	ชุดข้อมูลทดสอบที่ไม่ล้ม			ชุดข้อมูลทดสอบที่ล้ม		
	จำนวน	TN	FP	จำนวน	TP	FN
หน้า	25	25	0	25	25	0
หลัง	25	25	0	25	23	2
ซ้าย	25	24	1	25	25	0
ขวา	25	24	1	25	25	0
หน้าขวา	25	25	0	25	25	0
หน้าซ้าย	25	24	1	25	25	0
หลังซ้าย	25	25	0	25	25	0
หลังขวา	25	25	0	25	23	2
รวม	200	197	3	200	196	4

ตาราง 4-3 แสดงผลการตรวจจับการล้มของกล่องขอบเขตแบบทิศทางสามารถทำการจำแนกการล้มได้เป็นอย่างดีเยี่ยมเนื่องจากผลการทดสอบแบบ FP และ FN มีค่าน้อยมาก ส่วนผลสรุปของการเปรียบเทียบค่าเป็นไปได้อิงกรณีที่เป็นไปได้ในการตรวจจับการล้มกล่องขอบเขตทั้งหมดถูกแสดงในตารางที่ 4-4 และรูปที่ 4-3 ตารางที่ 4-4 ผลการเปรียบเทียบค่าของกรณีที่เป็นไปได้อิงการตรวจจับการล้มของแต่ละชนิดของกล่องขอบเขต

		2DBB	3DBB	DBB
ท่ายูก	TP ทายล้มผลล้ม	183	187	196
	TN ทายไม่ล้มผลไม่ล้ม	111	194	197
ท่ายผิด	FP ทายล้มผลไม่ล้ม	89	6	3
	FN ทายไม่ล้มผลล้ม	17	13	4

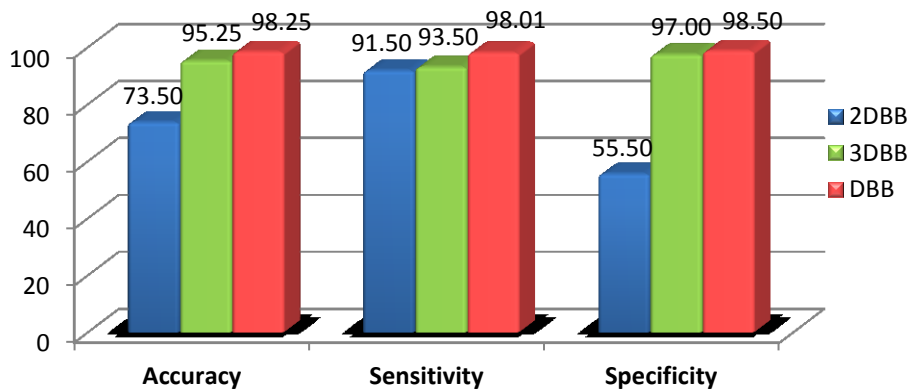


รูปที่ 4-3 แสดงกราฟแสดงการเปรียบเทียบค่าของกรณี
ที่เป็นไปได้ในการตรวจจับการล้ม

ตารางที่ 4-5 และรูปที่ 4-4 แสดงค่าความถูกต้องของการตรวจจับของแต่ละชนิดของกล่องขอบเขต จะเห็นได้ว่าเทคนิคการใช้กล่องขอบเขตแบบสองมิติทำค่าความถูกต้องได้ต่ำสุดที่ 73.50% ในขณะที่ค่าความถูกต้องของกล่องขอบเขตแบบสามมิติและกล่องขอบเขตแบบทิศทางมีค่าสูงใกล้เคียงกัน แต่เทคนิคกล่องที่ผู้วิจัยได้สูงกว่าเล็กน้อยที่ 3%

ตารางที่ 4-5 แสดงค่าความถูกต้องของการตรวจจับของแต่ละชนิดของกล่องขอบเขต

Approaches	Accuracy (%)	Sensitivity (%)	Specificity (%)
2DBB	73.50	91.50	55.50
3DBB	95.25	93.50	97.00
DBB	98.25	98.01	98.50



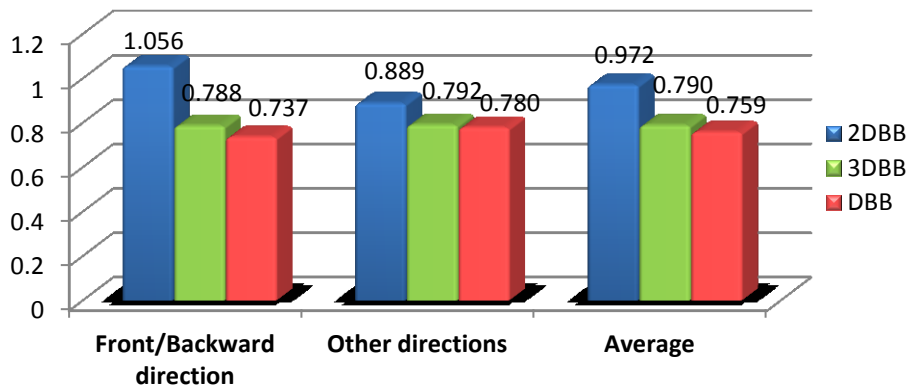
รูปที่ 4-4 แสดงกราฟเปรียบเทียบค่าความถูกต้องของกล่องขอบเขตทั้ง 3 แบบ

4.2 ผลการทดสอบด้านเวลาที่ใช้ในการตอบสนองต่อการล้มในแต่ละทิศทาง

งานวิจัยนี้ได้วิเคราะห์การตอบสนองต่อการล้มของกล่องขอบเขตแต่ละชนิด โดยแสดงในตารางที่ 4-6 และรูปที่ 4-5 ซึ่งพบว่าเวลาการตอบสนองเฉลี่ยของกล่องขอบเขตแบบสองมิติอยู่ที่ 0.972 วินาที ซึ่งมีค่าสูงกว่าเทคนิคการใช้กล่องขอบเขตแบบสามมิติและแบบทิศทางที่ 0.790 วินาที และ 0.759 วินาทีตามลำดับ

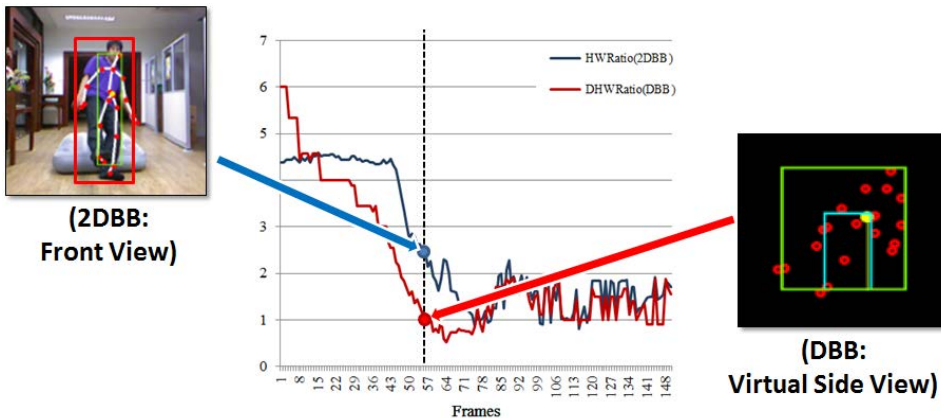
ตารางที่ 4-6 ผลการเปรียบเทียบเวลาที่ใช้สามารถระบุการล้มของของกล่องขอบเขตแต่ ละชนิด

CASE	Fall detection frame and Response time (second)		
	2DBB	3DBB	DBB
<i>Front/Backward direction</i>	31.69 (1.056 s.)	23.65 (0.788 s.)	22.13 (0.737 s.)
<i>Other directions</i>	26.67 (0.889 s.)	23.76 (0.792 s.)	23.41 (0.780 s.)
<i>Average</i>	29.18 (0.972 s.)	23.70 (0.790 s.)	22.77 (0.759 s.)



รูปที่ 4-5 แสดงกราฟเปรียบเทียบเวลาที่ใช้ในการระบุการล้มของกล่องขอบเขตทั้ง 3 ชนิด

เมื่อพิจารณาพีเจอร์ที่ได้จากกล่องขอบเขตแบบสองมิติเมื่อเทียบกับกล่องขอบเขตแบบทิศทางแล้วจะพบว่า ค่าที่จะนำไปใช้พิจารณานั้นแสดงได้ถึงลักษณะสถานะของการล้มได้ช้ากว่ากล่องขอบเขตแบบทิศทางดังแสดงในภาพที่ 4-6 ซึ่งจะเห็นได้ว่าค่าพีเจอร์ของกล่องขอบเขตแบบทิศทางมีคุณสมบัติเข้าถึงค่าเงื่อนไขที่กำหนดได้เร็วกว่าค่าพีเจอร์ที่ได้จากกล่องขอบเขตแบบสองมิติโดยเฉลี่ย



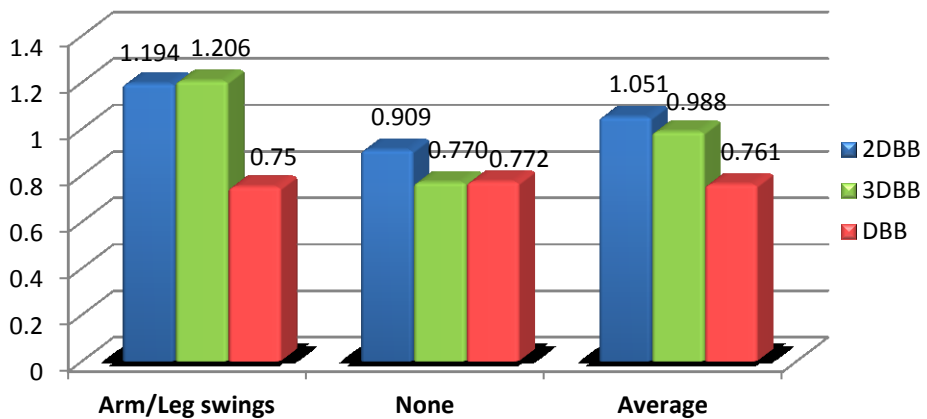
รูปที่ 4-6 แสดงการเปรียบเทียบค่าของพีเจอร์ระหว่างกล่องขอบเขตแบบสองมิติ และแบบทิศทางในช่วงระยะเวลาที่กำลังเกิดการล้ม

4.3 ผลการทดสอบด้านเวลาที่ใช้ในการตอบสนองต่อการล้มที่เกิดจากผลกระทบจากการแกว่งแขนหรือการก้าวเดิน

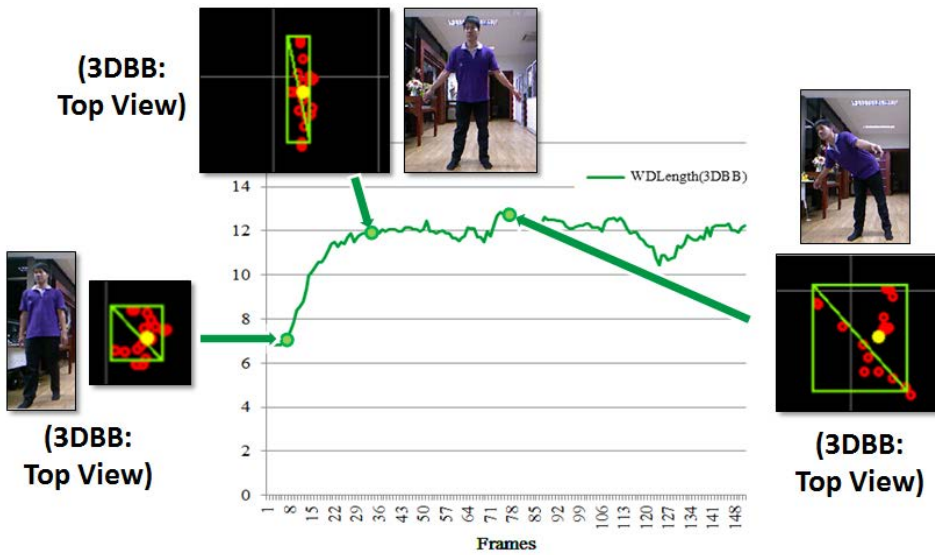
ผลของการทดลองผลกระทบของการเคลื่อนไหวแขนและขาตอบสนองการตรวจจับการล้มจะแสดงดังตารางที่ 4-7 และภาพที่ 4-7 ซึ่งจะเห็นได้ว่าเทคนิคการใช้กล่องขอบเขตแบบสองมิตียังคงมีค่าเฉลี่ยของการตอบสนองต่ำสุดที่ 1.051 วินาที สำหรับทุกกรณี ส่วนเทคนิคการใช้กล่องขอบเขตแบบสามมิติจะได้รับผลกระทบมากในกรณีที่เกิดการแกว่งแขนหรือการก้าวเดินระหว่างการล้มโดยมีค่าเฉลี่ยของการตอบสนองอยู่ที่ 0.988 วินาที สุดท้ายคือค่าเฉลี่ยการตอบสนองของกล่องขอบเขตแบบทิศทางที่มีค่าที่ค่อนข้างคงที่โดยเฉลี่ยคือ 0.761 วินาที ในทุกกรณี

ตารางที่ 4-7 แสดงการเปรียบเทียบเวลาที่ใช้ตอบสนองการล้มที่เกิดจากผลกระทบจากการแกว่งแขนหรือการก้าวเดิน

CASE	Detection frame and Response time (second)		
	2DBB	3DBB	DBB
Arm/Leg swings	35.82 (1.194 s.)	36.18 (1.206 s.)	22.50 (0.750 s.)
None	27.29 (0.909 s.)	23.12 (0.770 s.)	23.17 (0.772 s.)
Average	31.55 (1.051 s.)	29.65 (0.988 s.)	22.83 (0.761 s.)



รูปที่ 4-7 แสดงกราฟเปรียบเทียบเวลาที่ใช้ตอบสนองการล้มที่เกิดจากผลกระทบจากการแกว่งแขนหรือการก้าวเดินของกล่องขอบเขตทั้ง 3 ชนิด



รูปที่ 4-8 แสดงความแปรปรวนของค่าพีเจอรซ์ของกล่องขอบเขตแบบสามมิติ
ที่เกิดจากผลกระทบของการเคลื่อนไหวแขนหรือขา

ผลกระทบของการเคลื่อนไหวแขนและขาในบางสถานการณ์นั้นมีผลกระทบต่อเวลาที่ตอบสนองของกล่องขอบเขตแบบสามมิติดังแสดงในรูปที่ 4-8 จะเห็นได้ว่าตัวอย่างของค่าพีเจอรซ์ที่มีค่าเพิ่มขึ้นสูงในขณะที่ไม่มีการล้มเกิดขึ้นในการจำลองสถานการณ์ อันเนื่องจากระบบวิธีการตรวจจับการล้มที่มีการใช้เทคนิคกล่องขอบเขตแบบสามมิตินั้นค่อนข้างมีความแปรปรวนเมื่อร่างกายเกิดการเคลื่อนไหวซึ่งมีผลกระทบเข้าไปเสริมหรือขัดเซยค่าพีทเจอรซ์หลักจึงทำให้ค่าพีเจอรซ์หลักดังกล่าวไม่สามารถนำไปใช้งานเพื่อพิจารณาสำหรับตรวจจับการล้มได้

บทที่ 5 สรุปผลการทดลอง

5.1 สรุปผลการทดลอง

งานวิจัยนี้ได้นำเสนอกล่องขอบเขตชนิดใหม่ที่ใช้สำหรับปรับปรุงประสิทธิภาพของการตรวจจับการลัมของผู้สูงอายุที่เป็นปัญหาทางสุขภาพที่สำคัญในปัจจุบัน กล่องขอบเขตของเราถูกสร้างขึ้นมาบนพื้นฐานของการนำเสนอรูปร่างของกล่องให้แสดงถึงแนวการลัมที่เกิดขึ้นจริงเสมือนการทำให้มุมมองส่องไปในทิศทางตั้งฉากกับการลัมอยู่ตลอดเวลา โดยมีผลสรุปการทดลองดังต่อไปนี้

5.1.1 ผลการทดสอบความถูกต้องในการจำแนกเหตุการณ์ลัมหรือไม่ลัม

ผลจากการทดลองนั้นพบว่า ค่าเปอร์เซ็นต์ของค่าความถูกต้องของการใช้กล่องขอบเขตแบบสองมิติ นั้นมีค่าต่ำสุดทุกรูปแบบของกล่อง ปัจจัยหลักที่ทำให้ค่าความถูกต้องลดลงมากนั้นเกิดจากการแจ้งเตือนชนิด false positive (แจ้งเตือนว่าลัมแต่ไม่เกิดการลัมขึ้นจริง) เกิดจากโน้มตัวของอาสาสมัครที่มีทิศทางการโน้มไปด้านหน้าเข้าหากล่อง ออกจากกล่อง และออกด้านข้างที่มีลักษณะโน้มตัวไปข้างหน้าต่างทำให้ค่าพีเจอร์หลักแสดงผลได้ไม่ถูกต้อง

นอกจากนี้ความผิดพลาดของกล่องขอบเขตแบบสองมิติ นั้นยังเกิดขึ้นในส่วนของการทดสอบที่เป็นการลัม โดยจะเกิดขึ้นในลักษณะการลัมที่เป็นการนอนลัมลงไปบนพื้นที่มีทิศทางขนานกับเส้นสายตาของตัวกล่อง มีผลทำให้ค่าของพีเจอร์ไม่มีโอกาสถึงค่าเทรชโฮลด์ที่กำหนดไว้ได้ จึงทำให้เกิดการแจ้งเตือนชนิด false negative

เปอร์เซ็นต์ของค่าความถูกต้องของการใช้กล่องขอบเขตแบบสามมิติสามารถทำค่าได้สูง ซึ่งมีความมากกว่าแบบการใช้กล่องขอบเขตแบบสองมิติอันเป็นผลมาจากมิติของข้อมูลที่มากกว่าสามารถที่จะตรวจจับการลัมในทางลึกลงภาพได้ แต่ผลเสียของการใช้กล่องในลักษณะเช่นนี้นั้นคือหากมีการแกว่งแขนหรือการก้าวเดินระหว่างการลัมจะทำให้ลักษณะรูปร่างและขนาดของกล่องมีการเปลี่ยนแปลงอยู่ตลอดเวลาซึ่งในบางครั้งอาจเกิดขึ้นในขณะที่เกิดหรือไม่เกิดการลัมจึงทำให้การตรวจจับผิดพลาด ซึ่งผลที่ตามมาคือ การเกิด false negative (แจ้งเตือนว่าไม่ลัมแต่เกิดการลัมขึ้นจริง) ขึ้น

สำหรับการใช้กล่องขอบเขตแบบทิศทางที่ได้นำเสนอมีค่าความถูกต้องมากกว่าเทคนิคการใช้กล่องทั้งสองชนิดในงานวิจัยก่อนหน้า นั้นหมายความว่ากล่องดังกล่าวสามารถสร้างค่าพีเจอร์ที่สอดคล้องกับการตรวจจับการล้มได้ในมุมมองจริง

5.1.2 ผลการวัดประสิทธิภาพด้านเวลาที่ใช้ในการตอบสนอง

แสดงให้เห็นว่าเวลาที่ใช้ตอบสนองของเทคนิคการใช้กล่องขอบเขตแบบสองมิติจะค่อนข้างมีการตอบสนองช้าโดยมีค่าความเร็วเฉลี่ยที่ช้ากว่าการใช้กล่องขอบเขตอื่นซึ่งใช้เวลาในการตรวจจับได้ช้าสำหรับกรณีที่มีทิศทางการล้มเข้าและออกจากตัวกล่อง

ส่วนเวลาการตรวจจับของกล่องขอบเขตแบบสามมิติจะมีค่ามากกว่ากล่องขอบเขตแบบทิศทาง เป็นผลมาจากในขณะที่มีการล้มเกิดขึ้นนั้นลักษณะท่าทางของผู้ล้มมีผลต่อค่าพีเจอร์หลักที่ใช้พิจารณา สำหรับความเร็วของเทคนิคการใช้กล่องขอบเขตแบบทิศทางที่ผู้วิจัยนำเสนอมีความเร็วสูงกว่าเทคนิคการใช้กล่องขอบเขตอื่น

5.1.3 ผลการทดสอบด้านเวลาที่ใช้ในการตอบสนองต่อการล้มที่เกิดจากผลกระทบจากการแกว่งแขนหรือการก้าวเดิน

ผลการทดลองที่เกิดขึ้นในงานวิจัยนี้สามารถแสดงได้ว่า อิทธิพลของการแกว่งแขนหรือการก้าวเดินมีผลทำให้การตรวจจับทำได้ช้าลงกว่าในงานวิจัยที่ใช้กล่องขอบเขตแบบสองมิติและแบบสามมิติแบบเดิม

สำหรับกรณีที่ไม่เกิดการแกว่งแขนหรือการก้าวเดินนั้น ระบบการตรวจจับการล้มที่มีการใช้กล่องขอบเขตแบบทิศทางและแบบสามมิติมีเวลาที่สามารถตรวจจับการล้มได้เร็วที่สุดมีค่าใกล้เคียงกัน แต่สำหรับกรณีที่การแกว่งแขนหรือการก้าวเดินนั้นกล่องขอบเขตแบบทิศทางนั้นสามารถความเร็วในการตอบสนองการตรวจจับการล้มได้ดีกว่าทุกรูปแบบ โดยสามารถทำความเร็วเร็วกว่าทั้งการใช้กล่องแบบสองมิติและแบบสามมิติ

จะเห็นได้ว่าผลการทดลองของระบบการตรวจจับการล้มที่คุณสมบัติของขอบเขตแบบทิศทาง สามารถทำงานได้แบบ Real time ที่สามารถทำการวิเคราะห์การ

ลัมรวดเร็ว โดยที่ไม่จำเป็นต้องมีการปรับกรองค่าสัญญาณข้อมูลซึ่งอาจมีผลทำให้ความเร็วในการตอบสนองลดลงไปได้ กล้องขอบเขตที่นำเสนอสามารถให้ค่าพีเจอร์ที่มีคุณสมบัติสำหรับใช้ในการพิจารณาเหตุการณ์ของการลัมได้เป็นอย่างดี ซึ่งสามารถทำการแจ้งเตือนได้แม้ว่าร่างกายของมนุษย์ยังไม่ได้ลัมลงบนพื้นโดยสมบูรณ์และกรณีที่มีการแกว่งแขนหรือการก้าวเดินระหว่างการลัมได้

5.2 ปัญหาและข้อเสนอแนะ

งานวิจัยนี้ได้นำเสนอรูปแบบของการสร้างพีเจอร์ที่เกิดจากการออกแบบกล้องขอบเขตที่สร้างขึ้นใหม่โดยทำการทดสอบในชุดข้อมูลที่จำลองขึ้นมาเพื่อทดสอบระบบ ซึ่งผลที่ได้พบว่าทำงานได้ดีก็ต่อเมื่อสามารถข้อมูลภาพเก็บทุกส่วนของร่างกายของอาสาสมัครที่จำลองการลัมได้เท่านั้น เนื่องจากระบบที่ได้เสนอนั้นทำงานอยู่บนพื้นฐานของ Joint Point ที่ SDK ของ Microsoft Kinect ได้บริการข้อมูลกลุ่มของจุดสำคัญร่างกายทั้งหมด 20 จุดให้นำมาใช้อย่างอัตโนมัติ โดยจะมีข้อเสียอยู่ที่หาวิถีวัตถุขวางกั้นบางส่วนหรือมีบางส่วนจากร่างกายที่อยู่นอกขอบเขตของกล้องจะทำให้ไม่สามารถตรวจจับได้เต็มประสิทธิภาพ อันเกิดมาจากการแสดงค่าของตำแหน่งกลุ่มจุดดังกล่าวอยู่ในตำแหน่งที่ไม่ถูกต้อง ส่งผลทำให้ลักษณะรูปร่างและขนาดกล้องขอบเขตแบบทิศทางนั้นมีความแปรปรวน

ดังนั้นแนวทางในการแก้ไขปัญหาก็สามารถทำให้ขนาดของกล้องขอบเขตมีขนาดที่แปรผันไม่แน่นอนนั้น การใช้ข้อมูลในส่วนของ Depth Image เข้ามาช่วยในการปรับตำแหน่งของ Joint Point จะเป็นทางเลือกที่ดีที่สุด เพราะข้อมูลที่เป็น Cloud point เหล่านี้จะสามารถบอกตำแหน่งของวัตถุได้โดยตรงไม่ต้องผ่านตัวแทน Joint point ที่มีความไม่แน่นอนสำหรับในกรณีที่บางส่วนถูกบดบัง ซึ่งการนำ Depth Image มาใช้นั้นอาจใช้รูปแบบวิธีการประมาณการ Interpolation เพื่อเติมเต็มบริเวณที่ขาดหายไปแล้วทำการวิเคราะห์ใหม่อีกครั้ง

สำหรับงานวิจัยที่น่าสนใจในอนาคตที่ทีมงานวิจัยมองเห็นว่าการใช้ประโยชน์จากกล้องที่สามารถแสดงผลได้ 3 มิติได้ แม้ว่าจะทำให้เราสามารถเห็นข้อมูลที่กล้องทั่วไปไม่สามารถทำได้ แต่การเพิ่มจำนวนกล้องที่ใช้ในการตรวจจับถือว่าเป็นเพิ่มภาระในการจัดหาเครื่องมือที่ใช้ในการประมวลผลที่มีประสิทธิภาพสูงเพื่อรองรับการเชื่อมต่อของกล้องเหล่านี้ นอกจากนี้หากมีการพัฒนาการตรวจจับที่สามารถติดตั้งแล้วมีคุณสมบัติที่ปรับรูปแบบการตรวจจับให้เข้ากับสภาพแวดล้อมได้นั้น จะถือว่าทำให้เกิดความสะดวกในการติดตั้งเพื่อตรวจจับการลัมได้

บรรณานุกรม

- Arthur Chapman, Biomechanical Analysis of Fundamental Human Movements, HUMAN KINETICS, 2008.
- Tom Fawcett. “An Introduction to ROC Analysis”. Pattern Recognition Letters, Vol. 27, 2006, pp. 861-874, 2.
- Nor Surayahani Suriani, and Aini Hussain, “Sudden fall classification using motion features”, Signal Processing and its Applications (CSPA), 2012, pp. 519 – 524.
- Jia Luen Chua, Yoong Choon Chang, and Wee Keong Lim, “Intelligent Visual Based Fall Detection Technique for Home Surveillance”, Computer, Consumer and Control (IS3C) 2012, pp. 183 – 187.
- Miao Yu, Adel Rhuma, Syed Mohsen Naqvi, Liang Wang and Jonathon Chambers, “Posture Recognition Based Fall Detection System For Monitoring An Elderly Person In A Smart Home Environment”, Information Technology in Biomedicine, IEEE Transactions on Issue: 99, 2012, pp. 1– 11.
- Damien Brulin, Yannick Benezeth, and Estelle Courtial, “Posture Recognition Based on Fuzzy Logic for Home Monitoring of the Elderly”, Information Technology in Biomedicine, IEEE Transactions Volume: 16, Issue: 5, 2012, pp. 974 – 982.
- Caroline Rougier, Edouard Auvinet, Jacqueline Rousseau, Max Mignotte and Jean Meunier, “Fall Detection from Depth Map Video Sequences”, 9th International Conference on Smart Homes and Health Telematics (ICOST), 2011, pp. 121 – 128.
- Shaou-Gang Miaou, Pei-Hsu Sung, and Chia-Yuan Huang, “A Customized Human Fall Detection System Using Omni-Camera Images and Personal Information”, IEEE International Conference on Distributed Diagnosis and Home Healthcare (D2H2) Conference, 2006, pp.103-107.

- Peng Shen Ong, et al, "An FPGA-Based Hardware Implementation of Visual based Fall Detection," IEEE International Conference on Region 10 Symposium, 2014, pp. 397-402.
- Behçet Uğur Töreyn, Yiğithan Dedeoğlu and A. Enis Çetin, "HMM based falling person detection using both audio and video", IEEE International Workshop on Human-Computer Interaction, Vol. 3766, 2005, pp. 211-220.
- Alex Edgcomb and Frank Vahid, "Automated Fall Detection on Privacy-Enhanced Video", IEEE International Conference on Engineering in Medicine and Biology Society (EMBC), 2012, pp. 252-255.
- Ji Tao, et al, "Fall Incidents Detection for Intelligent Video Surveillance," IEEE International Conference on Information, Communications and Signal Processing, 2005, pp. 1590 - 1594.
- Georgios Mastorakis and Dimitrios Makris, "Fall detection system using Kinect's infrared sensor", Journal of Real-Time Image Processing, Vol. 9, Issue 4, 2012, pp. 635-646.
- Vitoantonio Bevilacqua, et al, "Fall Detection in indoor environment with Kinect sensor," IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA), 2014, pp. 319-324.

ภาคผนวก

Fall Detection using Directional Bounding Box

Apichet Yajai¹, Annupan Rodtook², Krisana Chinnasarn³ and Suwanna Rasmeequan⁴

1,3,4 Department of Computer Science, Faculty of Informatics, Burapha University, Chonburi, Thailand

2 Department of Computer Science, Faculty of Science, Ramkhamhaeng University, Bangkok, Thailand

1 apichet.yaj@stou.ac.th

Abstract— Falls are significant public health problem. In the last few years, several researches based on computer vision system have been developed to detect a person who has fallen to the ground. This paper presents a novel fall detection technique namely the directional bounding box (DBB) to detect a falls event especially a situation of fall direction paralleling the line of camera's sight. The DBB is constructed with perspective side view transformation of depth information. Moreover, a new aspect ratio namely the center of gravity point (COG) is proposed to monitor human movement. The proposed technique was evaluated with the video data set gathering from a RGB-D sensor. The experimental result of the proposed technique was better both accuracy and response times than previous works.

Keywords—Fall Detection; Depth image; Bounding box; Center of mass; Aspect ratio

I. INTRODUCTION

In the present moment, falls are the major causes (the main courses or a leading cause) of fatal injury (affecting the elderly who live alone) for elderly citizen who living alone in the residence all around the world.

The impact of falls leads to the severe physical injuries including traumatic brain. In a particular report of World Health Organization (WHO) [1], it has been revealed that the third cause of chronic disability is falls. Both intrinsic and extrinsic factors are the main risk factors for falls which falling rate in elderly over than 65 years is around 28-35 % per year [2]. Especially, the number of older Thai citizens has been growing rapidly which the estimation report of the National Statistical Office [3] revealed that the rating of elderly population will be increased from 11.9% in 2010 up to 25.1% within 2030.

The increasing of elderly is not just impact to increase the costs of the related health care, but the problem of standby staff has also occurred. The consequences of falls among elderly are common and often devastating problem. The effects of the fall are sometimes numerous leading to the death especially the rapid treatment of medical staffs and treatment facilities were inefficient and ineffective. The aforementioned problem leads to academic research and product development on fall detection and prediction which save many human lives. A fall detection system can be assistive device to alert to medical staffs when a human fallen.

In the past few years, much research on the topic of fall detection was being widely conducted. The proposed method can be categorized into three types: wearable sensor, image vision and fusion/ambient [4]. Wearable sensor-based have used the portable sensor such as accelerometer and gyroscope sensor to detect the body movement. These devices were

attached on significant joint points of body. Although the detection accuracy was high, it had some disadvantages in case of elderly often forgot to wear and it was uncomfortable in daily life. [5]. Vision based is the most favorite approach of researcher to monitor and detect fall. They analyzed event via the body shapes from video streaming by using image processing method [6]. The third method in [7] showed the combination of multi device sensors for detection fall. Hence, the computation time and cost were high.

Bounding box has been widely used in several key aspects of computer vision to monitor human movement. The box can be categorized into two general classes onto the 2D and 3D frame systems. Several numerical data are extracted from these virtual boxes. Firstly, the 2D data description is obtained from general camera. Human physical appearances as height and width are considered as ratio-based feature. Researcher in [8] presented the embedded vision-based system to increased speed and reduced the power consumption by using FPGA. They detected fall event through considering the value of HW ratio when value was less than threshold value. HW ratio in [9] was the main feature in their system that used Omni-camera. To detect fall, they combined the personal information to adjust the individual of threshold each person. In [10], they processed with visual data from CCTV and audio data. HW ratio was evaluated to reach final fall event by using Hidden Markov Model. Research as in [11] used HW ratio and the reverse one features from bounding box in various type of private video picture. They applied Dynamic Time Warping (DTW) to separates fall event. The study in [12] extracted HW ratio from projection the subtractive image. They classified the kind ratio by using Gaussian Mixture Model (GMM) and applied with FSM to detect fall.

The 3D bounding box is gathered from RGB-D camera. The depth information of RGB-D sensors is represented in z-axis of 3D frame system of box. The key of depth information is benefit for improving the quality of bounding box data to monitor human body movement. This feature appeared in [13, 14], they proposed the WD line length feature from the 3D bounding box. The value of WD length was the agent of human fall activity. To detect falling, they computed the velocity of WD line length combine with height's value of position head through rule base algorithm.

Although entire bounding box above are good for detecting falls of human, but there are drawback in some situation of each virtual box. The weak point for using 2D bounding box occurred in case of fall direction parallels the line of camera's sight that made the false positive result, because the size of bounding box cannot represents the shape of the characteristic of falls correctly. Using 3D bounding box, the cuboids' size has been change all the time when people walk or swing their

arms. Their shape sizes are uncertain and high variant. For this cause, it shows the false negative from the system that cannot detect when human has fallen down.

The main objective of this research is to propose a bounding box framework based on computer vision for higher efficiency of fall detection system. The novel bounding box called the directional bounding box (DBB) achieves robust design which the direction of a cloud of 3D joint points obtained by RGB-D sensor is considered and these points are mapped using perspective transform. Moreover, this research proposes a new aspect ratio namely the center of gravity point (COG) to monitor human movement. Experimental results show that the DBB provide high performance for accuracy and speed of fall detection.

The paper is organized as follows. Section 2 gives an overview of fundamental concepts of three frameworks of the bounding box. An outline of the proposed algorithm is described in section 3. Section 4 presents experimental results of fall detection based on a variety of bounding box frameworks. Discussion and briefly conclusion is given in the last section.

II. THE THEORY OF BOUNDINB BOX

A bounding box also known as a minimum bounding rectangle (MBR) is a small rectangle containing objects or human body [10]. Numerical features of human-movement within a bounding box are extracted and analyzed. Some criterions were designed among numerical features to recognize or detect fall events which the position of human skeleton are changed. The different skeleton positions provide the different bounding box shape and extracted features. Moreover, the dimensions of features describing human movement are various depending on the outline of bounding box. Now, there are two types of bounding box that are a rectangle in 2D (see Fig. 1) and a rectangular cuboid in 3D (see Fig. 2).

A. 2D Bounding Box

Many researchers [8-12] have used 2D axis bounding box (2DBB) provided by a regular camera to monitor and extract human-movement features. A regular camera is easy to get cheap. However, the information subjected by this device is two dimensional data shown in Fig. 1. The bounding box attributes as height and width values are projected by the human body silhouette. Finally, a ratio-based feature of height and width (see eq.1) is considered to represent human physical appearances.

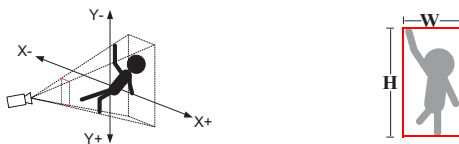


Fig. 1. 2D bounding box.

The aspect ratio of a bounding box (HW ratio) is the ratio of the box's width to the box's height defined as below equation. The low aspect ratios (height < width) are positively associated with a risk of fall event. It means that the plane of human skeleton seem to be parallel to the floor.

$$R_{HW} = \frac{H}{W} \quad (1)$$

HW ratio is effective to trace the different poses of falling along the sides [8,10-12]. However, some limited view of this box type is a drawback to indicate positions in real space of human especially human's fall in a parallel direction to the line of camera's sight. In such situation, it is really hard or impossible to work with a 2D bounding box.

B. 3D Bounding Box

3D bounding box (3DBB) is a box that encloses all the geometry of 3D objects (see Fig. 2). The depth information of RGB-D sensors is represented in z-axis of this box type. It is benefit to detect the depth from motion when a human moves toward in the line of camera's sight.

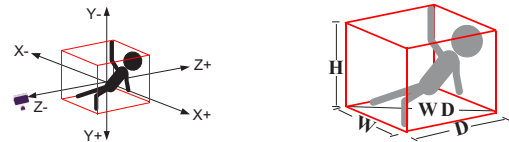


Fig. 2. 3D bounding box.

In this kind of box, a combination of width and depth is considered to create additional numerical feature as WD length (see eq.2). WD length is a line presenting a state of human fallen in depth. The higher WD length is also associated with a fallen event. It shows that the plane of human skeleton seem to be close to the floor.

$$WD = \sqrt{W^2 + D^2} \quad (2)$$

WD length is more sensitive to noise induced by body movement. Arm or leg swings often enhance some noise to the length of WD line which leads to false negative result of detection. Therefore, some additional features are the alternative option of researchers which data speed is one of them. Velocity analysis of WD length and the box's height are used to trace the fallen state [13,14].

C. Directional Bounding Box

The contribution of this research is the directional bounding box (DBB) designed to break down the drawback of 2D/3D bounding box by applying the direction of human fall to redefine a new bounding box. The numerical features extracted from this bounding box type are flexible and efficient to decide the human state of fall or non-fall. The concept of DBB will be described in details below.

The main concept of this research is to create the virtual bounding box at side view of human. The fall direction gathered from RGB-D sensor is essential for generating virtual view from arbitrary perspectives.

1. The falls can be traced actually from the angle between body parts. Firstly, to create the human body vector (see Fig. 3). Two joint points of head ($Joint_{Head}$) and hip ($Joint_{HipCenter}$) are used to generate a body vector $U = (X_Q, Z_Q)$ which $X_Q = X_{head} - X_{hip}$ and $Z_Q = Z_{head} - Z_{hip}$. U is considered as tangent vector presenting the direction of the fall. The vector projection of U onto xz-plane is also known as

vector V . The angle theta (θ) between two vectors U and V is given by equation (3) as below.

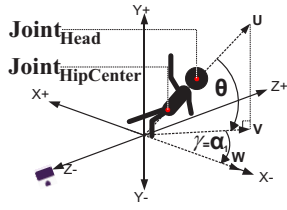


Fig. 3. Main body vector.

$$\theta = \arccos\left(\frac{U \cdot V}{\|U\| \|V\|}\right) \quad (3)$$

The theta is used as detector to monitor the starting point of falls when its value is less than the threshold.

2. To determine the quadrant of a human body vector. The quadrant describes four zones of the coordinate plane. The zone information is a usage for calculating the angle of the rotation of main body vector. The rotation angle (γ) is specified as the following criteria.

$$\gamma = \begin{cases} 0 & \text{if } (X_Q < 0) \text{ and } (Z_Q = 0) \\ -180 & \text{if } (X_Q > 0) \text{ and } (Z_Q = 0) \\ -\alpha_1 & \text{if } (Z_Q > 0) \\ \alpha_2 & \text{if } (Z_Q < 0) \end{cases} \quad (4)$$

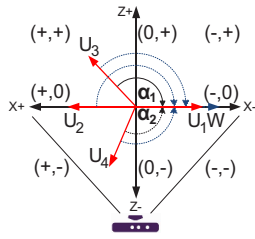


Fig. 4. Diagram for calculating the rotational angle in each the quadrant of vector U .

3. To rotate all joint points of human skeleton around the COG [15]

Firstly, to define a rotation matrix (see eq. 5) in 3 dimensional space that rotate vector V by γ which $W = V * R_y(\gamma)$.

$$R_y(\gamma) = \begin{bmatrix} \cos(\gamma) & 0 & \sin(\gamma) \\ 0 & 1 & 0 \\ -\sin(\gamma) & 0 & \cos(\gamma) \end{bmatrix} \quad (5)$$

Then, a matrix $R_y(\gamma)$ is applied to rotate all joint points of body around the COG point instead of foot as pivoting point which the matrix composition of such transformation in homogeneous space is given by the following matrix (6).

$$R_{COG}(\gamma) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -x_{COG} & -y_{COG} & -z_{COG} & 1 \end{bmatrix} \begin{bmatrix} \cos(\gamma) & 0 & \sin(\gamma) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\gamma) & 0 & \cos(\gamma) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

$P_{r_i} = P_i * R_{COG}(\gamma); i = 1, \dots, n$ where $P_i = (x_i, y_i, z_i, 1)$ is the i^{th} joint point and $P_{r_i} = (x_{r_i}, y_{r_i}, z_{r_i}, 1)$ is the i^{th} rotated joint point. The COG point is denoted by $(x_{COG}, y_{COG}, z_{COG})$ and n represents the joint points number including head hip and COG points.

4. To locate the DBB. All joint points have been rotated onto virtual side view, then a rectangle namely "directional bounding box", which cover all points but excludes hand, wrist and elbow joint points. It is defined by considering only x and y coordinates of the top-right and bottom-left position (see Fig.5).

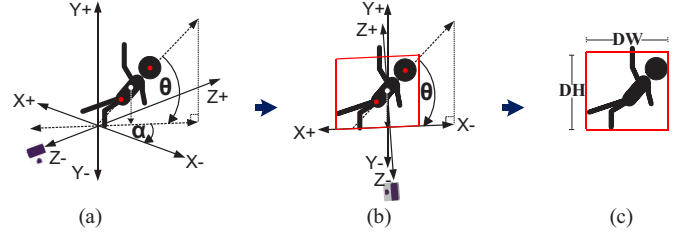


Fig. 5. Overview of the DBB extractor, (a) The position of sensor and human in real world coordinate, (b) All joint points are rotated around the COG and (c) DBB.

The main feature of this bounding box is DHW ratio. It can be calculated as (7)

$$R_{DHW} = \frac{DH}{DW} \quad (7)$$

III. THE PROPOSED APPROACH

This part is to describe a process sequence of fall detection algorithm which is based on DBB (see a diagram in Fig. 6).

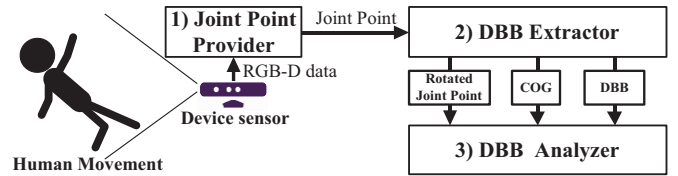


Fig. 6. Fall detection system framework.

1. Joint Point Provider

An RGB-D sensor such as the Microsoft Kinect providing the depth information is used to produce joint points of human skeleton. This sensor device can return the twenty tracked joint points information. Set of all points is considered and analyzed in the process of DDB extractor.

2. DBB Extractor

The main process is to define the bounding box through virtual side view of human body by considering the direction of falls. Then, all rotated joint points in framework of DDB are passed along to the next process.

3. DBB Analyzer

Numerical features are extracted from DDB as follows:

3.1 Aspect ratios of DHW and COG

The main feature of proposed algorithm is DHW ratio (R_{DHW}) which is ratio of DW to DH as described in section II. Moreover, this research proposes a new aspect ratio namely COG ratio (R_{COG}) to support DHW ratio in monitoring of human movement. This feature is calculated from the rotated joint points and COG of the DBB. The equation of COG ratio is given by eq. 8.

$$R_{COG} = \frac{D_{COG}}{H_{COG}} \quad (8)$$

The COG is a geometric property of any object specifying a status of the balance [16]. For a human body to balance, we have to have the center of gravity directly above the base of support (BOS) which is area between feet touch on the ground (see Fig. 7 (a)). P_G is intersection point of a line drawn down to the ground from the COG and H_{COG} is the length of such line.

The COG of a human body sometimes may be located outside the base of support (see Fig.7 (b)). In this situation, it generates a distance namely D_{COG} which is a measure of how far a P_G point has travelled from its P_{BOS} point (the rightmost point of BOS). The high aspect ratios ($H_{COG} < D_{COG}$) are positively associated with a risk of fall event.

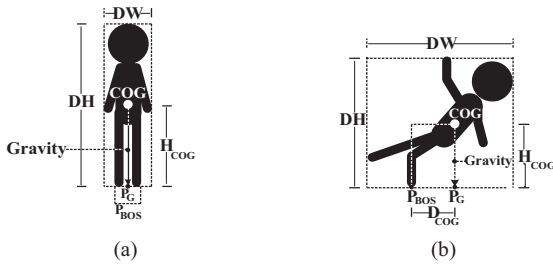


Fig. 7. The position of COG in difference situation on DBB .

3.2 Detection Algorithm

To choose appropriate parameters for fall detection algorithm is more important. Thus, the parameters of proposed algorithm are set up from the training set selection. In the training set, we can get the time and speed at the end of human fall. Frame Count is defined as the number of frames elapsed from the first frame of fall situation which the frame number was 51 and the times of falling bodies were 1.7 sec. This research is to focus on the completion time of fall because we prefer to detect fall at the early state (before to be in "lying down") with very good accuracy.

In the training section, four appropriate parameter values of algorithm used as decision criterion were discovered as below.

1. The critical frame number was 22 frames or 0.73 second of falling body time.
2. The upper bound of DHW ratio (TU) was 3.20
3. The lower bound of DHW ratio (TL) was 1.38
4. The threshold of COG ratio (TCOG) was 0.27.

A finite state machine diagram of the detection algorithm consisting for 4 major states as S1, S2, S3 and S4 is shown in Fig. 8 (a). S1 describe a state of the regular acting of human as

running, walking and standing. In finite state machine, the machine is in only one state at a time. Thus, it will move from state S1 to state S2 when the R_{DHW} is a range [TU TL]. The machine will be shifted to next state (S3) when the R_{DHW} is less than TL (see Fig. 8 (b)). The final state is S4 which satisfies the criteria of human fall. R_{COG} is applied to detect whether human falls or not. If R_{COG} is greater than TCOG (see Fig. 8 (c)) and the time travelling along states is in the critical frame number range, the current machine will be located at the final state (S4)).

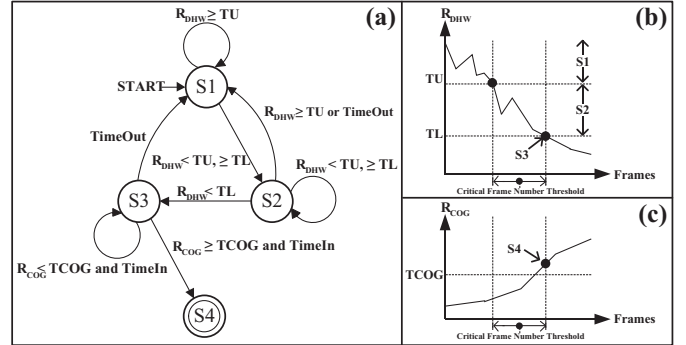


Fig. 8. The finite state machine diagram of detection algorithm.

IV. EXPERIMENTAL AND RESULT

We tested our proposed methods with unseen moving image sequence. The sequence was classified into two categories which consists of fall and non-fall image sequence. Experimental results were compared with the results from the literatures. 2DBB and 3DBB environment has been used for this experiment. Details of related content for experimental implementation describe as follows:

A. Environment

We placed a Microsoft Kinect sensor from the side of testing room at one meter height from ground level. Distance between Kinect sensor and human should be 1.5 to 3 meters. This position is similar to the existing works that we will compare the results with [8][13]. In addition, this position is suitable and convenient for collecting and tracing the movement of human body, especially in case of human is starting to fall. In other words, if the position of Kinect sensor is too high, the some joint points data of human body will be disappeared from the sensor. Sensor will collect only head or top part of body. Then, we will be able to calculate or predict of the falling.

B. Datasets

We have tested with the datasets from five volunteers. Collecting data were classified into two subgroups, fall and non-fall classes. Movement in each class consisted of eight moving directions which were moving backward, backward-left, backward-right, forward, forward-left, forward-right, sideways-left and sideways-right. The samples acted five times in each direction. Then, total of moving data, used in this experiment, are 400 movements/patterns.

C. Performance measurement

The performance of the system was measured by confusion matrix. Then, accuracy performance values including accuracy (9), specificity (10) and sensitivity (11) were computed in order to compare the accuracy performance with the existing algorithms in the literature, detailed in Fig. 9 [17]. This confusion matrix represents the instances of predicted and actual classes. As mention above, fall and non-fall classes were considering in this research. Therefore, evaluation was categorized into four possibilities. Moreover, response times of the system were used for measuring performance in term of the detecting times.

		PREDICTED (System)	
		NEG (NonFall)	POS (Fall)
ACTUAL (Ground Truth)	NEG (NonFall)	TN	FP
	POS (Fall)	FN	TP

Fig. 9. Confusion Matrix

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

$$True\ Negative\ Rate\ (Specificity) = \frac{TN}{TN+FP} \quad (10)$$

$$True\ Positive\ Rate\ (Sensitivity) = \frac{TP}{TP+FN} \quad (11)$$

D. Results

We compared the result of our proposed with two existing approaches in [8] and [13]. The existing approaches used 2D/3D bounding box for acquiring the learning and testing features. In [8] used 2DBB for creating HW ratio of human body and the system alarmed when the ratio value was less than one. On the other hand, in [13] used WD length creating from 3DBB. The proposed system checked the alarm point when the velocity of WD length reached the given threshold value. As described in Section 4.3, accuracy performance and response time of detection have been used for evaluating the proposed method with two existing approaches.

D.1 Accuracy performance

Accuracy performance comparing between proposed methods, DBB with two existing methods, 2DBB and 3DBB, was illustrated in TABLE 1.

TABLE 1. The efficiency each approach.

Approaches	Accuracy (10)	Specificity (11)	Recall (12)
2DBB	73.75	55.62	91.87
3DBB	95.31	97.5	93.12
DBB	98.43	98.75	98.12

The percentage accuracy of HW ratio from 2DBB was the lowest comparing with the others. It was at 73.75%. Accuracy ratio from the 2DBB was low because it provided on 2D-plane data. In other word, there is non-fall data from HW ratio. Then, false positive detection was occurred when the volunteer slope their body in the direction of side away left/right and front left/right from the sensor. The percentage

of sensitivity is decreased by the direction of fall in the fall dataset that slope out from the camera sensor. It makes the value of HW ratio is always not less than value one.

The percentage accuracy of WD length from 3DBB was at 95.31%. The value was higher than 2D bounding box because this technique created 3D coordinate data. Therefore, it was able to detect fall in any direction. But, there were some drawback in some situations of arm swinging or leg walking movement. These movements affected directly to the value of WD length because the shape of 3D bounding box features were changed any time, not only depend on slope of main body. Then, false negative was happened when the movements compensate the velocity value of WD length.

Accuracy performance of the proposed method, called DHW ratio from DBB, was up to 98.43%. This type of bounding box endured with any fall direction and movements because the value of DHW ratio corresponds with the shape of human movement in real vision.

D.2 Response times

The response times of fall detection is another factor which we should concern. Shorter response time is better system. In other word, we can easily to help elderly people in case that some incidents are happening to them. Streaming data used in the proposed system were collected from the Microsoft Kinect sensor at frame rate at 30 frames per second. We selected test set from true positive results for comparing the detection response time. The measurements of response time each approach was shown in Table 2.

TABLE 2. Falling detection frame and Response time for detecting falling directions

CASE	Falling detection frame and Response time (s)		
	2DBB	3DBB	DBB
Front/Backward direction	31.69 (1.056 s.)	23.65 (0.788 s.)	22.13 (0.737 s.)
Other directions	26.67 (0.889 s.)	23.76 (0.792 s.)	23.41 (0.780 s.)

Table 2 showed image frame and response time in detecting falling people. From TP values, we calculated an average of detecting frame and response time. We found that our proposed method detected the front/backward direction at frame 22.13 or 0.737s. In other word, experimental results showed that our proposed method was the fastest in detecting falling direction of human body. In the others directions, our proposed method provided an optimal response times. It was better than the 3DBB algorithm.

TABLE 3. Detection frame and Response time for detecting Arm/Leg swings

CASE	Detection frame and Response time (s)		
	2DBB	3DBB	DBB
Arm/Leg swings	35.82 (1.194 s.)	36.18 (1.206 s.)	22.50 (0.750 s.)
None	27.29 (0.909 s.)	23.12 (0.770 s.)	23.17 (0.772 s.)

In case of Arm/Leg swings, detection frame and response time were illustrated in Table 3. If the human were swing their legs and arms, our proposed algorithm provided a significant

results. It was about 0.444s. and 0.456s. which is faster than the 2DBB and 3DBB algorithm, respectively.

Fig. 10 has shown the trend value of three features of all bounding box. Fig. 10 (a). depicted the stability of DHW ratio that reaches to the threshold value before HW ratio. For Fig. 10 (b)), WD length takes up to high value even though the falls have not been occurred.

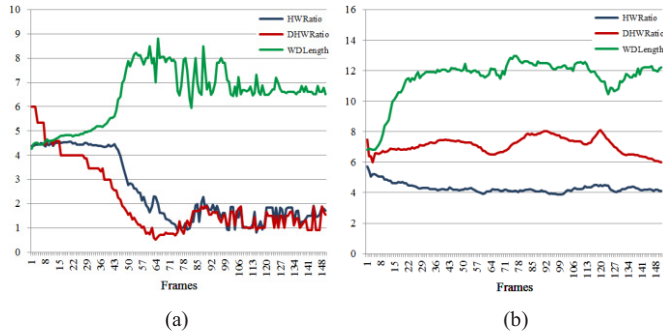


Fig. 10. (a) Graph depicts three feature when human will fall.
(b) Graph depicts three feature when the human walk or swing arms.

Because DHW ratio has a stable value, to add feature to improve detection is comfortable way. We integrated the COG of body with DHW ratio feature to detect falls. This technique can improve the responds time shorter for detection at DHW ratio is 1.38. For this reason, our proposed method is run faster than others in the literatures.

To track human during falling by using DBB technique was illustrated in Fig. 11. In Fig.11 (a) presented the human movement from front view of the vision sensor. In Fig.11 (b) depicted the top view of falling that presented all rotated joint points in yellow color. Fig.11 (c) revealed DBB and COG that presented the starting falling state. Both features provided the correct answer for detecting falling of human body.

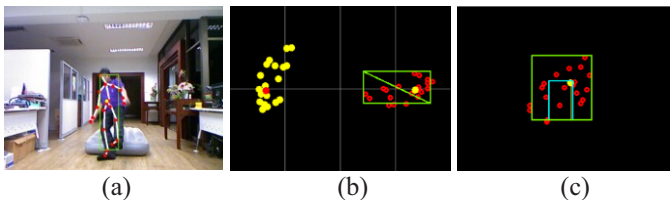


Fig. 11. (a) Front view of falling, (b) Top view of falling. (c) DBB and COG

V. CONCLUSION AND FUTUREWORK

Nowadays, effect of high growing rate on elderly population is an important issue. People life is longer than the past. Then, living problem of elderly people should be concerned. Significant problem of elderly people is falling. Hence, it is important to do a research for detecting of falling in early stage or starting to fall. In this paper, we have presented the DBB for improving the efficacy of fall detection system. The previous works have applied the 2D/3D bounding box to their research, but there are the weak points in limit of perspective view and some movements. Those reasons could make some noise of signal value then decreased the accuracy of detection. DBB is created based on the direction of body fall with 3D coordinate data. This operation is simple and takes a little computation time. The new

bounding box is brought to generate the DHW ratio and COG ratio features.

The result of our experiment shows the real time fall detection system. Our approach is faster and more accuracy than the others works. Our bounding box provided the stable features and suitable for monitoring in early state of fall. Moreover, it doesn't have to use any filter with data signal that can reduce the computation time. After this work, we intend to focus on the others fall environments.

REFERENCES

- [1] C. J. L. Murray and A. D. Lopez, "Global and regional descriptive epidemiology of disability: Incidence, prevalence, health expectancies and years lived with disability", *Global Burden Disease*, Vol. 1, Pages: 201–246, 1996.
- [2] U. Laessoe, H.C. Hoeck, O. Simonsen, Thomas Sinkjaer, Michael Voigt, "Fall risk in an active elderly population can it be assessed?", *Journal of Negative Results in BioMedicine*, vol. 6, 2007.
- [3] S. Siriboon. and C. Milinthrakoon (2008, August). Population statistics of Thai Elderly [Online]. Available: http://www.cps.chula.ac.th/cps/research_division/article_ageing/ageing_001.html.
- [4] M. Mubashir, L. Shao and L. Seed, "A survey on fall detection: Principles and approaches", *Neurocomputing*, Vol. 100, Pages: 144–152, 2013.
- [5] C.F.Lai, et al, "Detection of cognitive injured body region using multipletri-axial accelerometers for elderly falling", *IEEE Sensors Journal*, Vol. 11, Pages:763–770, 2011.
- [6] H.Foroughi, B.S.Aski and H.Pourreza, "Intelligent Video Surveillance for Monitoring Fall Detection of Elderly in Home Environments", 11th IEEE International Conference on Computer and Information Technology (ICCIIT), Pages: 219–224, 2008.
- [7] C.Doukas, et al, "Patient Fall Detection Using Support Vector Machines", *IFIP The International Federation for Information Processing*, SpringerLink, Vol. 247, Pages: 147–156, 2007.
- [8] P. S. Ong, et al, "An FPGA-Based Hardware Implementation of Visual based Fall Detection," *IEEE International Conference on Region 10 Symposium*, Pages: 397-402, 2014.
- [9] S. G. Miaou, P. H. Sung, and C. Y. Huang, "A Customized Human Fall Detection System Using Omni-Camera Images and Personal Information", *IEEE International Conference on Distributed Diagnosis and Home Healthcare (D2H2) Conference*, Pages: 103-107, 2006
- [10] B.U. Toreyin, Y. Dedeoglu and A.E. Cetin, "HMM based falling person detection using both audio and video", *IEEE International Workshop on Human-Computer Interaction*, Vol. 3766, Pages: 211–220, 2005.
- [11] A. Edgcomb, and F. Vahid, "Automated Fall Detection on Privacy-Enhanced Video", *IEEE International Conference on Engineering in Medicine and Biology Society (EMBC)*, Pages: 252-255, 2012.
- [12] J. Tao, et al, "Fall Incidents Detection for Intelligent Video Surveillance," *IEEE International Conference on Information, Communications and Signal Processing*, Pages: 1590 - 1594, 2005.
- [13] G. Mastorakis and D. Makris, "Fall detection system using Kinect's infrared sensor", *Journal of Real-Time Image Processing*, Vol. 9, Issue 4, Pages: 635-646, 2012.
- [14] V. Bevilacqua, et al, "Fall Detection in indoor environment with Kinect sensor," *IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA)*, Pages: 319-324, 2014.
- [15] R. Drillis, R. Contini, and M. Bluestein, "Body segment parameters: A survey of measurement techniques", *Artificial Limbs*, Vol.8, Pages: 44-66, 1964.
- [16] A. Chapman, *Biomechanical Analysis of Fundamental Human Movements*, HUMAN KINETICS, 2008.
- [17] F. Tom. "An Introduction to ROC Analysis". *Pattern Recognition Letters*, Vol. 27, Pages: 861-874, 2006. 2