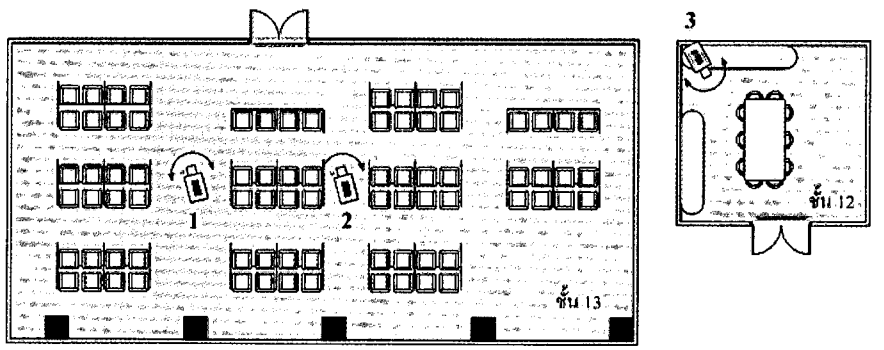
บทที่ 3
โครงสร้างของระบบการติดตามบุคคลที่นำเสนอ

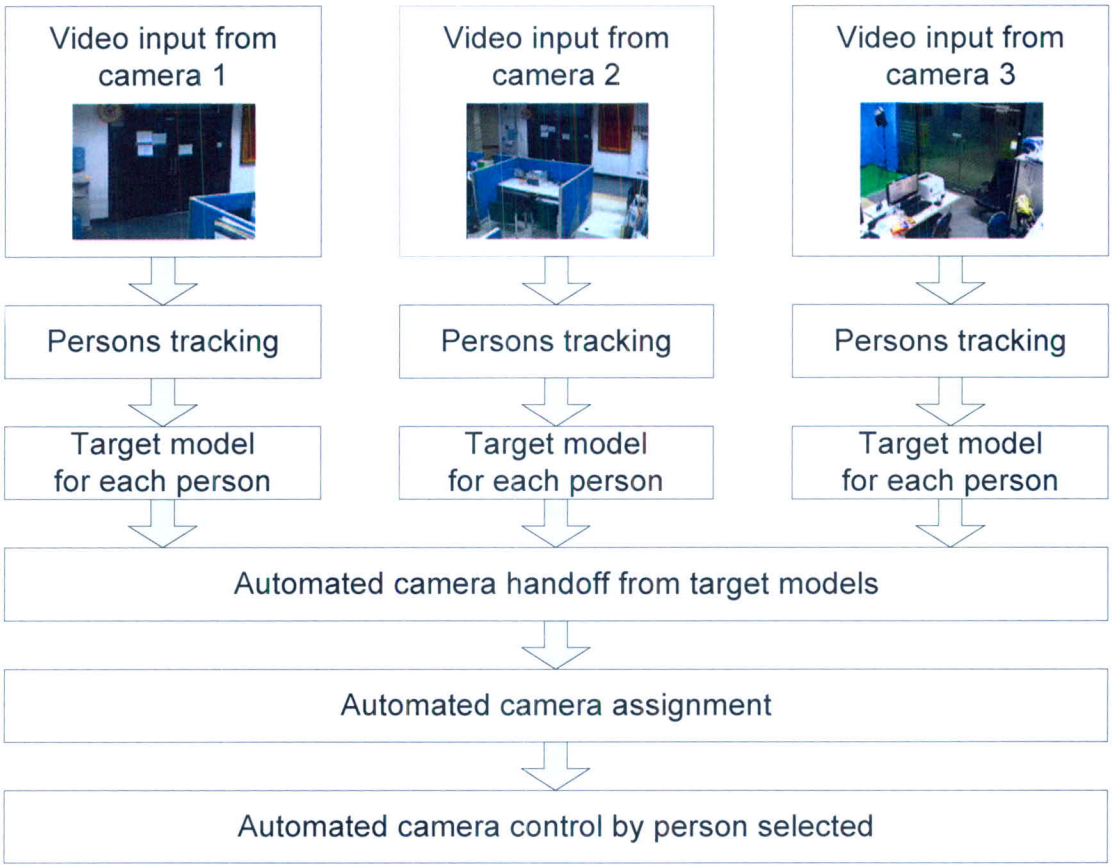
3.1 รูปแบบการทำงานของระบบการติดตามบุคคลที่นำเสนอ

ระบบการติดตามบุคคลที่นำเสนอเป็นระบบที่ประกอบไปด้วยกล้องวีดิทัศน์แบบ PTZ รุ่น SONY SNC-RZ50P จำนวน 3 ตัว โดยกล้องวีดิทัศน์ตัวที่ 1 และ 2 จะถูกติดตั้งไว้ที่ห้องปฏิบัติการวิจัยกรรมวิธีสัญญาณดิจิทัล ชั้น 13 และกล้องวีดิทัศน์ตัวที่ 3 จะถูกติดตั้งไว้ที่ห้องปฏิบัติการวิจัยวีดิทัศน์ ชั้น 12 ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย แสดงแผนผังการติดตั้งกล้องวีดิทัศน์ดังรูปที่ 3.1 โดยมีรูปแบบการประมวลผลภาพวีดิทัศน์ที่ได้จากกล้องวีดิทัศน์ทั้งสามเป็นแบบรวมศูนย์ (Centralized video processing) ที่นำเอาข้อมูลภาพวีดิทัศน์ทั้งหมดมาประมวลผลที่เครื่องคอมพิวเตอร์รุ่น Dell Vastoro 3300 เพียงเครื่องเดียว



รูปที่ 3.1 แผนผังการติดตั้งกล้องวีดิทัศน์

จากบทที่ 1 ได้กล่าวถึงปัญหาที่เกิดจากข้อมูลที่เป็นตัวแทนของบุคคลที่ไม่เพียงพอต่อการส่งต่อบุคคลจากกล้องวีดิทัศน์หนึ่งสู่กล้องวีดิทัศน์อีกตัวหนึ่ง ซึ่งจะส่งผลให้เกิดความผิดพลาดในการติดตามบุคคลระหว่างกล้องวีดิทัศน์ วิทยานิพนธ์นี้จึงนำเสนอรูปแบบการติดตามบุคคลระหว่างกล้องวีดิทัศน์ที่อาศัยการเก็บสะสมข้อมูลซึ่งเป็นตัวแทนของบุคคล อีกทั้งยังทำการปรับปรุงขั้นตอนวิธีการติดตามบุคคลให้มีประสิทธิภาพยิ่งขึ้น แสดงภาพรวมของระบบการติดตามบุคคลที่นำเสนอดังรูปที่ 3.2



รูปที่ 3.2 ภาพรวมของระบบการติดตามบุคคลที่นำเสนอ

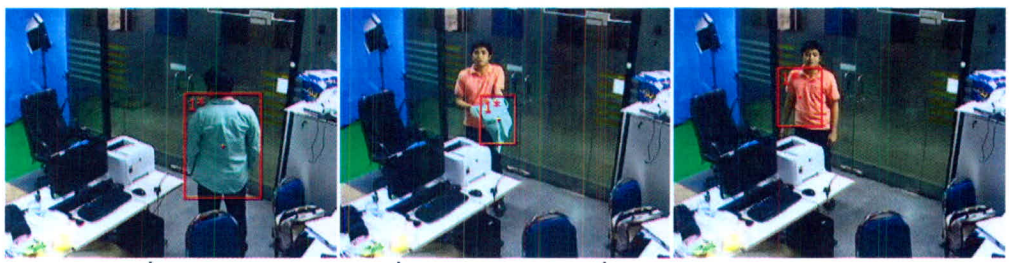
วิทยานิพนธ์นี้ได้แบ่งส่วนของโครงสร้างการติดตามบุคคลที่นำเสนอออกตามส่วนต่าง ๆ ตามวิธีทำงานของระบบ โดยส่วนแรกจะกล่าวถึงการติดตามบุคคลด้วยการใช้ข้อมูลสีบริเวณเสื้อของบุคคลในการติดตาม ซึ่งได้มีการปรับปรุงในขั้นตอนการติดตามบางส่วนให้สามารถทำงานได้ดีขึ้น ข้อมูลสีบริเวณเสื้อของบุคคลจะถูกใช้ในการติดตามทุก ๆ ลำดับภาพ วิทยานิพนธ์นี้จึงได้นำข้อมูลดังกล่าวมาใช้ในการสร้างแบบจำลองลักษณะของบุคคล เพื่อเพิ่มโอกาสการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์อย่างอัตโนมัติให้มีความแม่นยำมากขึ้น และหากพบว่ามีบุคคลเดียวกันปรากฏอยู่ในมุมมองการรับภาพของกล้องวีดิทัศน์มากกว่าหนึ่งตัว ระบบจะทำการคัดเลือกกล้องวีดิทัศน์ที่มีมุมมองต่อบุคคลดังกล่าวที่ดีที่สุดในหัวข้อที่ 3.5 บุคคลที่ถูกติดตามด้วยกล้องวีดิทัศน์ทุกตัวในระบบสามารถถูกคัดเลือกให้เป็นบุคคลที่สนใจจากการเลือกด้วยมือได้ (Manual selection) ซึ่งภายหลังจากระบบทราบว่ามีบุคคลที่สนใจ จะกำหนดให้กล้องวีดิทัศน์ที่มีมุมมองต่อบุคคลที่สนใจดังกล่าวที่ดีที่สุดทำการเคลื่อนที่ตามด้วยการส่ายและก้มเงยอย่างอัตโนมัติในหัวข้อที่ 3.6 สำหรับ

การวัดประสิทธิภาพของระบบการติดตามที่นำเสนอนี้จะใช้วิธีการนับบุคคลซึ่งกล่าวไว้ในหัวข้อที่ 3.7

3.2 การติดตามบุคคลด้วยวิธีที่นำเสนอ

วิทยานิพนธ์นี้ได้เลือกใช้ขั้นตอนวิธีการติดตามบุคคลแบบการย้ายตามค่าเฉลี่ยแบบปรับตัวได้อย่างต่อเนื่อง (Continuously Adaptive Mean-Shift, CAMShift) สำหรับหลักการทำงานทั่วไปได้กล่าวไว้แล้วในหัวข้อที่ 2.3.2 CAMShift มีข้อดี คือ สามารถทำการติดตามวัตถุที่มีการเปลี่ยนแปลงทางรูปร่างและขนาดได้ดี โดยการทำงานของ CAMShift สำหรับการประมวลผลภาพวีดิทัศน์จะเริ่มต้นจากการกำหนดบริเวณของข้อมูลของบุคคลที่ได้จากขั้นตอนการตรวจจับบุคคลจากการเคลื่อนที่ ข้อมูลที่อยู่ในบริเวณดังกล่าวนี้จะถูกนำมาใช้ในขั้นตอนการทำนายตำแหน่งของบุคคลในลำดับภาพถัดไป การพิจารณาข้อมูลของบุคคลแบบมีขอบเขตนี้ทำให้ลักษณะของข้อมูลอื่นที่อยู่นอกบริเวณการพิจารณาไม่ส่งผลกระทบต่อการติดตามบุคคล วิทยานิพนธ์นี้ได้เลือกใช้ข้อมูลสีของบุคคลในรูปแบบของ Hue ฮิสโทแกรมจากบริเวณที่พิจารณาทั้งหมดมาใช้ในการทำนายตำแหน่งของบุคคลในลำดับภาพถัดไป เนื่องจาก Hue ฮิสโทแกรมเป็นข้อมูลสีที่ได้จากปริภูมิสี HSL ซึ่งเป็นปริภูมิสีที่แยกค่าเนื้อสีและแสงสว่างออกจากกัน การดึงเอาเฉพาะข้อมูล Hue ออกมาใช้ในการติดตามวัตถุ จึงทำให้ผลลัพธ์ของการติดตามคงทนต่อการเปลี่ยนแปลงทางด้านแสงสว่างได้ สำหรับ Hue ฮิสโทแกรมที่นำมาใช้เป็นลักษณะตัวแทนของบุคคลนี้ประกอบด้วยจำนวน 360 ถังข้อมูล เมื่อได้ข้อมูลซึ่งเป็นตัวแทนของบุคคลแล้วจึงทำการคำนวณหาค่าเฉลี่ยของข้อมูลภายในขอบเขตการพิจารณาในแนวแกนตั้งและแกนนอน ผลลัพธ์จากการคำนวณนี้จะทำให้ได้ตำแหน่งใหม่เพื่อใช้เป็นจุดศูนย์กลางของขอบเขตการพิจารณาในลำดับภาพถัดไป นอกจากนี้ CAMShift ยังได้มีการคำนวณขนาดของขอบเขตการพิจารณาใหม่ทุกครั้งที่มีการเปลี่ยนแปลงตำแหน่ง โดยให้สามารถย่อหรือขยายได้ตามปริมาณของข้อมูลสีส่วนใหญ่จากค่าเฉลี่ยของข้อมูลที่คำนวณได้ ผลลัพธ์ของการติดตามที่ได้จะถูกแสดงด้วยขอบเขตของการพิจารณาข้อมูลในแต่ละลำดับภาพซึ่งเรียกว่ากรอบการติดตาม วิทยานิพนธ์นี้ได้เลือกใช้กรอบการติดตามรูปสี่เหลี่ยมที่มีตัวเลขซึ่งเป็นลำดับของบุคคลที่ติดตามได้กำกับไว้ การคำนวณหาตำแหน่งใหม่และการปรับเปลี่ยนขนาดกรอบการติดตามของ CAMShift ได้กล่าวไว้แล้วในหัวข้อที่ 2.3.2

วิทยานิพนธ์นี้ได้ทำการปรับปรุง CAMShift ในส่วนของการสร้างเงื่อนไขในการปรับเปลี่ยนขนาดกรอบการติดตามบุคคล ซึ่งโดยปกติแล้ว CAMShift จะทำการปรับเปลี่ยนขนาดของกรอบการติดตามนี้ตามค่าเฉลี่ยของการแจกแจงความน่าจะเป็นที่ได้จากลักษณะที่ใช้เป็นตัวแทนของบุคคลหรือ Hue ฮิสโทแกรม ดังนั้นเมื่อลักษณะสีส่วนใหญ่ของบุคคลเปลี่ยนแปลงไปทั้งจากปัจจัยด้านแสงสว่าง และลักษณะการเคลื่อนที่ของบุคคลที่ส่งผลให้ลักษณะสีที่กล้องวีดิทัศน์สังเกตได้เปลี่ยนแปลงไป หรือการถูกบดบังจากสิ่งกีดขวาง สิ่งเหล่านี้ล้วนทำให้แนวโน้มของการปรับเปลี่ยนกรอบการติดตามมีขนาดเล็กลง จนในที่สุดอาจทำให้การติดตามบุคคลล้มเหลวได้ วิทยานิพนธ์นี้จึงทำการสร้างเงื่อนไขของการปรับเปลี่ยนขนาดของกรอบการติดตาม ให้ทำการขยายขนาดตามการเคลื่อนที่ซึ่งอยู่รอบข้างเมื่อพบว่าขนาดของกรอบการติดตามมีความกว้างหรือความยาวน้อยกว่า 30 จุดภาพ ผลจากการสร้างเงื่อนไขดังกล่าวทำให้ CAMShift สามารถติดตามบุคคลมีการเปลี่ยนสีเสื้ออย่างฉับพลันได้แสดงดังรูปที่ 3.3



รูปที่ 3.3 ผลจากการสร้างเงื่อนไขการปรับเปลี่ยนขนาดกรอบการติดตาม

ผลลัพธ์จากการติดตามข้างต้นทำให้ระบบการติดตามบุคคลที่ 1 สามารถทำการติดตามบุคคลได้จากข้อมูลสีเขียวของเสื้อแขนยาวและข้อมูลสีส้มจากเสื้อแขนสั้น ดังนั้นหากระบบการติดตามสามารถนำข้อมูลที่ได้จากการติดตามบุคคลนี้ไปใช้ในการส่งต่อบุคคลจากกล้องวีดิทัศน์ตัวหนึ่งสู่กล้องวีดิทัศน์อีกตัวหนึ่งได้ ก็อาจทำให้โอกาสในการส่งต่อบุคคลมีความแม่นยำมากขึ้น ซึ่งวิทยานิพนธ์นี้ได้ใช้เทคนิคการสะสมข้อมูล Hue ฮิสโทแกรมของบุคคลจากขณะทำการติดตามนี้ มาใช้ในการสร้างแบบจำลองบุคคลเพื่อเพิ่มโอกาสในการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์ด้วย ดังจะกล่าวในหัวข้อถัดไป

3.3 การสร้างแบบจำลองลักษณะของบุคคลที่นำเสนอ (Target modeling)

เป้าหมายของการสร้างแบบจำลองลักษณะของบุคคล คือ การสะสมข้อมูลซึ่งเป็นตัวแทนของบุคคลให้เพียงพอต่อการส่งต่อบุคคลจากกล้องวีดิทัศน์หนึ่งสู่กล้องวีดิทัศน์อีกตัวหนึ่ง หากสมมติฐานของการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์ที่แม่นยำที่สุด คือ การสะสมข้อมูลของบุคคลให้ได้มากที่สุด การสร้างแบบจำลองลักษณะบุคคลอาจทำได้จากการเก็บสะสมข้อมูล Hue ฮิสโทแกรมของบุคคลขณะทำการติดตามในทุก ๆ ลำดับภาพ แต่ทั้งนี้ก็จะทำให้ระบบการติดตามมีข้อมูลซึ่งเป็นตัวแทนของแต่ละบุคคลเป็นจำนวนมากทำให้สิ้นเปลืองทรัพยากรด้านหน่วยความจำ อีกทั้งยังอาจทำให้เกิดความล่าช้าในกระบวนการค้นหาลักษณะของบุคคลที่เหมือนกันเพื่อใช้ในการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์อีกด้วย ดังนั้นข้อมูลที่จะถูกเก็บสะสมเพื่อใช้เป็นตัวแทนของแต่ละบุคคลนี้ควรถูกคัดกรองแบบมีเงื่อนไข วิทยานิพนธ์นี้ได้สร้างเงื่อนไขในการสะสมข้อมูลเพื่อสร้างแบบจำลองของบุคคลจากความแตกต่างของข้อมูล Hue ฮิสโทรแกรมของแต่ละบุคคลที่ได้จากขั้นตอนการติดตาม กล่าวคือ ขณะทำการติดตามบุคคล ณ ลำดับภาพหนึ่ง ๆ หากพบว่าข้อมูลสีของบุคคลไม่แตกต่างจากข้อมูลที่ถูกจัดเก็บไว้ ก็ไม่มีความจำเป็นต้องเก็บสะสมข้อมูล Hue ฮิสโทแกรม ณ ลำดับภาพดังกล่าว แต่หากพบว่าข้อมูล Hue ฮิสโทแกรมมีความแตกต่างกันก็ควรจัดเก็บไว้เพื่อใช้เป็นตัวแทนของบุคคลในขั้นตอนการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์ต่อไป ยกตัวอย่างข้อมูลสีของบุคคลขณะทำการติดตามที่อาจมีความแตกต่างกันในแต่ละลำดับภาพ เช่น เมื่อบุคคลเคลื่อนที่ไปยังบริเวณที่แสงสว่างแตกต่างกันส่งผลให้ค่าสี Hue ฮิสโทแกรมเปลี่ยนแปลงไป หรือเกิดจากการเคลื่อนที่ของบุคคลซึ่งสวมเสื้อที่มีสีไม่เป็นเนื้อเดียวกัน การสะพายกระเป๋า หรือการสวมผ้าคลุมศีรษะ สิ่งเหล่านี้อาจทำให้ค่าสี Hue ฮิสโทแกรมของบุคคลเปลี่ยนแปลงไปขณะทำการติดตามได้ทั้งสิ้น

การสร้างแบบจำลองลักษณะบุคคลที่นำเสนอได้แบ่งขั้นตอนการคัดกรองข้อมูลออกเป็น 2 ส่วน คือ การคัดกรองเพื่อจัดเก็บข้อมูลในหน่วยความจำระยะสั้น (Shot-term memory, ST) และการคัดกรองเพื่อจัดเก็บข้อมูลในหน่วยความจำระยะยาว (Long-term memory, LT) หน้าที่ของหน่วยความจำทั้งสองคือการคัดกรองข้อมูลให้เพียงพอต่อการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์ ทั้งนี้ระบบการติดตามบุคคลจะใช้เพียงข้อมูลที่อยู่ใน LT เท่านั้นในการส่งต่อบุคคล เนื่องจากหน้าที่

ของ ST คือ การคัดกรองข้อมูลของบุคคลอย่างหยาบ กล่าวคือ Hue ฮิสโทแกรมที่ได้จากการติดตามแต่ละบุคคลอาจถูกกำหนดไว้ด้วยคาบของลำดับภาพที่ต้องการพิจารณา เช่น ต้องการพิจารณา Hue ฮิสโทแกรมของบุคคลที่ 1 ทุก ๆ 20 ลำดับภาพ หรือสามารถกำหนดจากคาบเวลาที่ต้องการได้เช่นกัน ซึ่งวิทยานิพนธ์ฉบับนี้ได้กำหนดให้ ST พิจารณาทุก ๆ ลำดับภาพวีดิทัศน์เนื่องจากระบบการติดตามบุคคลที่นำเสนอเป็นระบบการติดตามบุคคลภายในอาคารที่มุมมองการรับภาพของกล้องวีดิทัศน์แต่ละตัวมีไม่มากนัก จึงสามารถสังเกตเห็นบุคคลเคลื่อนที่ได้ในช่วงระยะเวลาสั้น ๆ เท่านั้น ทุก ๆ ลำดับภาพจึงมีความสำคัญทั้งหมด เมื่อ Hue ฮิสโทแกรมของบุคคลในลำดับภาพถัดไปเข้ามาจะถูกนำไปเปรียบเทียบกับข้อมูล Hue ฮิสโทแกรมของบุคคลดังกล่าวที่มีอยู่เดิมใน LT หากพบว่าข้อมูล Hue ฮิสโทแกรมทั้งสองมีความแตกต่างกันเกินค่าขีดเริ่มเปลี่ยนของระยะ Bhattacharyya สำหรับการสร้างแบบจำลองลักษณะบุคคล (BH_{Th}^{M}) ก็จะบันทึกให้ข้อมูล Hue ฮิสโทแกรมที่อยู่ใน ST ปัจจุบันนี้เป็นข้อมูลใน LT ด้วยเช่นกัน ขั้นตอนการการสร้างแบบจำลองลักษณะของบุคคลที่นำเสนอสามารถอธิบายได้ดังนี้

1. ให้ C แทนกล้องวีดิทัศน์ที่ใช้ในระบบการติดตามบุคคลที่นำเสนอ และ P แทนบุคคลที่ระบบสามารถทำการติดตามได้ดังสมการที่ (3.1) และ (3.2)

$$C = \{C_1, C_2, ..., C_n\} \tag{3.1}$$

$$P = \{P_1, P_2, ..., P_i\} \tag{3.2}$$

เมื่อ n คือ จำนวนกล้องวีดิทัศน์ทั้งหมดในระบบการติดตาม

i คือ จำนวนบุคคลทั้งหมดที่ระบบสามารถติดตามได้

2. ให้ ST เป็นหน่วยความจำระยะสั้นซึ่งมีหน้าที่ในการคัดกรองข้อมูลอย่างหยาบ และ LT เป็นหน่วยความจำระยะยาว ซึ่งข้อมูลทั้งหมดใน LT ของแต่ละบุคคลจะถูกใช้เป็นแบบจำลองของบุคคลดังกล่าวในการนำไปพิจารณาสำหรับกระบวนการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์ โดย ST และ LT เป็นหน่วยความจำของแต่ละบุคคลที่ได้จากแต่ละกล้องวีดิทัศน์ ดังนั้นจึงสามารถอธิบายเซตของข้อมูลใน ST และ LT ทั้งหมดได้จากสมการที่ (3.3) และ (3.4) ตามลำดับ

$$ST_P^C = \{ST_P^{C,1}, ST_P^{C,2}, ..., ST_P^{C,Index_{ST}}\} \tag{3.3}$$

$$LT_P^C = \{LT_P^{C,1}, LT_P^{C,2}, ..., LT_P^{C,Index_{LT}}\} \quad (3.4)$$

เมื่อ $Index_{ST}$ คือ ลำดับของข้อมูลใน ST

$Index_{LT}$ คือ ลำดับของข้อมูลใน LT

แต่เนื่องจากวิทยานิพนธ์นี้ให้ความสำคัญกับทุก ๆ ลำดับภาพ ดังนั้นข้อมูล Hue ฮิสโทแกรมของแต่ละบุคคลที่ติดตามจึงถูกนำมาบันทึกลงใน ST ดังสมการที่ (3.5)

$$ST_P^C = hist_P^C \quad (3.5)$$

เมื่อ $hist_P^C$ คือ Hue ฮิสโทแกรมของบุคคล P ที่ได้จากการติดตามในลำดับภาพปัจจุบันของกล้องวีดิทัศน์ C นั่นหมายความว่าข้อมูลใน ST จะถูกปรับเปลี่ยนใหม่เสมอ

3. ข้อมูลใน ST จะถูกบันทึกลง LT ก็ต่อเมื่อ Hue ฮิสโทแกรมของ ST มีความแตกต่างจาก LT ที่มีอยู่เดิม ยกเว้นเป็นการติดตามบุคคลดังกล่าวครั้งแรกซึ่งข้อมูลใน ST จะถูกบันทึกลงใน LT ทันที การหาความแตกต่างของฮิสโทแกรมสามารถทำได้จากการเปรียบเทียบความเหมือนของสองฮิสโทแกรมซึ่งคำนวณโดยการหาระยะห่าง Bhattacharyya หากพบว่าค่าระยะห่างที่ได้มีค่ามากกว่าขีดเริ่มเปลี่ยนของระยะความเหมือน Bhattacharyya สำหรับการสร้างแบบจำลอง (BH_{Th}^{Model}) ก็จะตัดสินให้ข้อมูลที่อยู่ใน ST ปัจจุบันบันทึกลงใน LT สามารถอธิบายการบันทึกข้อมูล Hue ฮิสโทแกรมลงใน LT ได้ดังสมการที่ (3.6) และ (3.7)

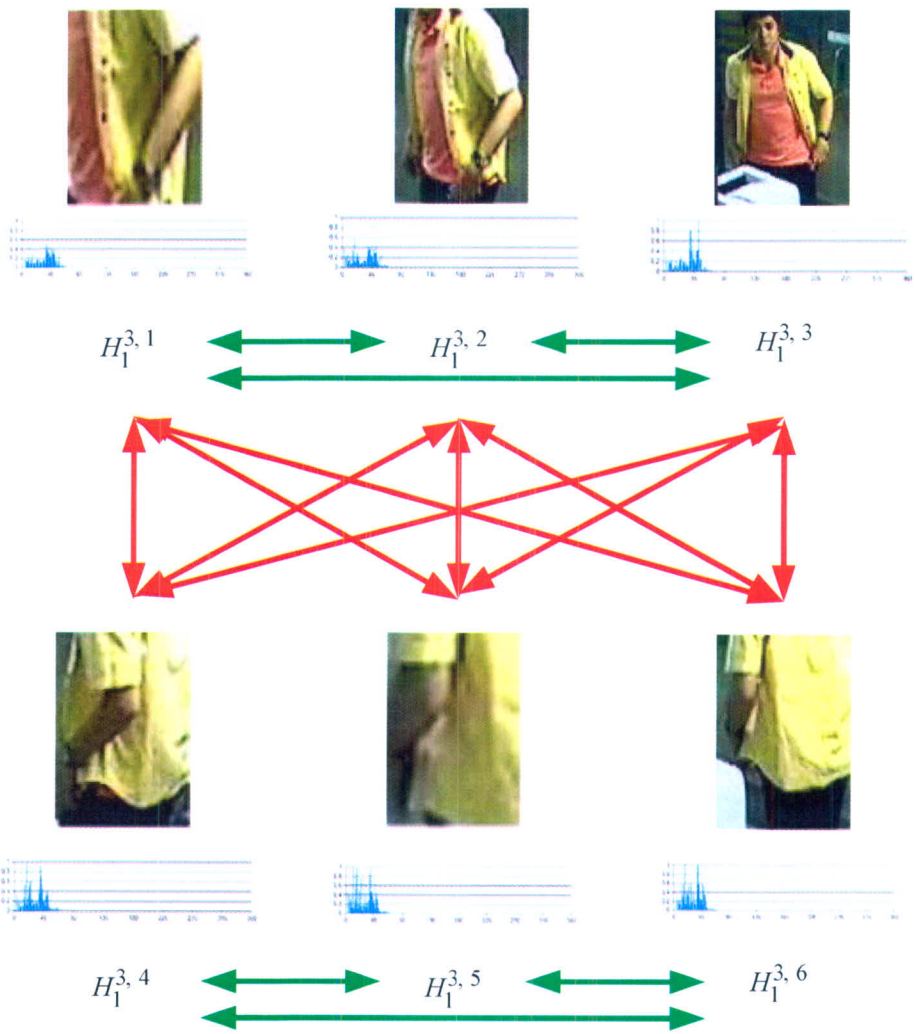
$$LT_P^{C,1} = ST_P^C \quad (3.6)$$

$$LT_P^{C,Index_{LT}+1} = ST_P^C, \quad BH(LT_P^{C,Index_{LT}}, ST_P^C) > BH_{Th}^{Model} \quad (3.7)$$

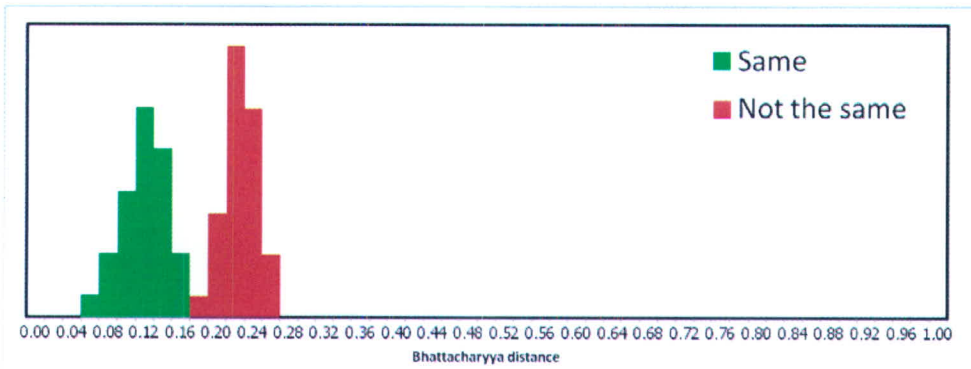
เมื่อ $BH(LT_P^{C,Index_{LT}}, ST_P^C)$ คือ ระยะห่าง Bhattacharyya ที่ได้จากข้อมูล Hue ฮิสโทแกรมใน $LT_P^{C,Index_{LT}}$ และ ST_P^C

การคำนวณหาค่าขีดเริ่มเปลี่ยนของระยะ Bhattacharyya สำหรับการสร้างแบบจำลอง (BH_{Th}^{Model}) ที่เหมาะสมทำได้จากการเปรียบเทียบ Hue ฮิสโทแกรมของบุคคลเดียวกันขณะทำการติดตามด้วยกล้องวีดิทัศน์แต่ละตัว ณ ลำดับภาพที่ให้ความสนใจ เช่น เมื่อบุคคลเคลื่อนที่ไปยังบริเวณที่แสงสว่างแตกต่างกัน หรือทำให้มุมมองของกล้องวีดิทัศน์ที่มีต่อบุคคลเปลี่ยนไป รูปที่ 3.4

แสดงตัวอย่างการเปรียบเทียบ Hue ฮิสโทแกรมของบุคคลเดียวกันภายในกล้องวีดิทัศน์เดียวกัน กำหนดให้ $H_{P_i}^{C_x,Index_h}$ คือ Hue ฮิสโทแกรมที่ได้จากขั้นตอนการติดตามบุคคลที่ P_i จากกล้องวีดิทัศน์ตัวที่ C_x ลำดับที่ $Index_h$ จากรูปบุคคลที่ถูกติดตาม คือ บุคคลที่ 1 ซึ่งถูกติดตามด้วยกล้องวีดิทัศน์ตัวที่ 3 การคำนวณหา BH_{Th}^{Model} จึงทำได้โดยการเปรียบเทียบจากลักษณะสีเสื้อของบุคคลที่ใกล้เคียงกันและแตกต่างกันนี้ ค่าเฉลี่ยของระยะห่าง Bhattacharyya ที่ได้จากการเปรียบเทียบลักษณะสีบริเวณเสื้อที่แตกต่างกันจะถูกนำมาใช้เป็นค่าขีดเริ่มเปลี่ยนของระยะ Bhattacharyya (BH_{Th}^{Model}) นั่นหมายความว่าระบบได้ยอมรับให้ค่าระยะห่าง Bhattacharyya ส่วนใหญ่ที่ได้จากการเปรียบเทียบลักษณะสีบริเวณเสื้อที่แตกต่างกัน เป็นค่าขีดแบ่งที่กำหนดการบันทึกข้อมูลลงใน LT เพื่อใช้สำหรับการสร้างแบบจำลองบุคคลนั่นเอง แสดงรูปผลการกระจายตัวของข้อมูลระยะ Bhattacharyya ของบุคคลเดียวกันภายในกล้องวีดิทัศน์ที่ลักษณะสีเสื้อของบุคคลที่ใกล้เคียงกันและแตกต่างกันดังรูปที่ 3.5



รูปที่ 3.4 ตัวอย่างการเปรียบเทียบ Hue ฮิสโทแกรมของบุคคลเดียวกันภายในกล้องวีดิทัศน์ที่ลักษณะสีเสื้อของบุคคลที่ใกล้เคียงกันและแตกต่างกัน



รูปที่ 3.5 การกระจายตัวของข้อมูลระยะ Bhattacharyya ของบุคคลเดียวกันภายในกล้องวีดิทัศน์ที่ลักษณะสีเสื้อของบุคคลที่ใกล้เคียงกันและแตกต่างกัน

จากรูปที่ 3.4 กำหนดให้ลูกศรเส้นสีแดงแสดงการเชื่อมโยงข้อมูล Hue ฮิสโทแกรมของบุคคลเดียวกันที่ได้จากการเปรียบเทียบลักษณะสีบริเวณเสื้อที่แตกต่างกัน ซึ่งสัมพันธ์กับสีของรูปการกระจายตัวข้อมูลในรูปที่ 3.5 ซึ่งพบว่าค่าเฉลี่ยของระยะ Bhattacharyya ที่ได้ คือ 0.25 ซึ่งวิทยานิพนธ์นี้ได้นำค่าระยะ Bhattacharyya ดังกล่าวไปใช้ในการกำหนดเป็นค่าขีดเริ่มเปลี่ยนของระยะ Bhattacharyya สำหรับการสร้างแบบจำลองลักษณะบุคคล (BH_{Th}^{Model})

3.4 การส่งต่อบุคคลระหว่างกล้องวีดิทัศน์ในระบบที่นำเสนอ

โดยทั่วไปในการส่งต่อบุคคลจากกล้องวีดิทัศน์สู่กล้องวีดิทัศน์ทำได้โดยการนำข้อมูลซึ่งเป็นตัวแทนของบุคคลมาทำการเปรียบเทียบ หากพบว่าข้อมูลดังกล่าวมีลักษณะที่คล้ายคลึงกันก็จะทำการส่งต่อบุคคลนั้นทันที การแสดงผลลัพธ์ของการส่งต่อบุคคลอาจทำได้โดยการระบุตัวที่ใช้แทนตัวบุคคลให้เป็นเลขเดียวกันหรือกำหนดสีของกรอบการติดตามบุคคลให้เป็นสีเดียวกัน จากบทที่ 1 ที่ได้กล่าวถึงปัญหาของระบบการติดตามบุคคลระหว่างกล้องวีดิทัศน์ที่เป็นผลกระทบจากมุมมองของกล้องวีดิทัศน์ในระบบที่เป็นอิสระจากลักษณะของบุคคลที่ปรากฏ การสะสมข้อมูลที่ใช้เป็นตัวแทนของบุคคลที่เพียงพอจึงเป็นทางแก้ปัญหาที่เกิดจากผลกระทบดังกล่าว สำหรับขั้นตอนการสะสมข้อมูลเพื่อใช้เป็นแบบจำลองลักษณะของบุคคลได้กล่าวไว้แล้วในหัวข้อที่ 0

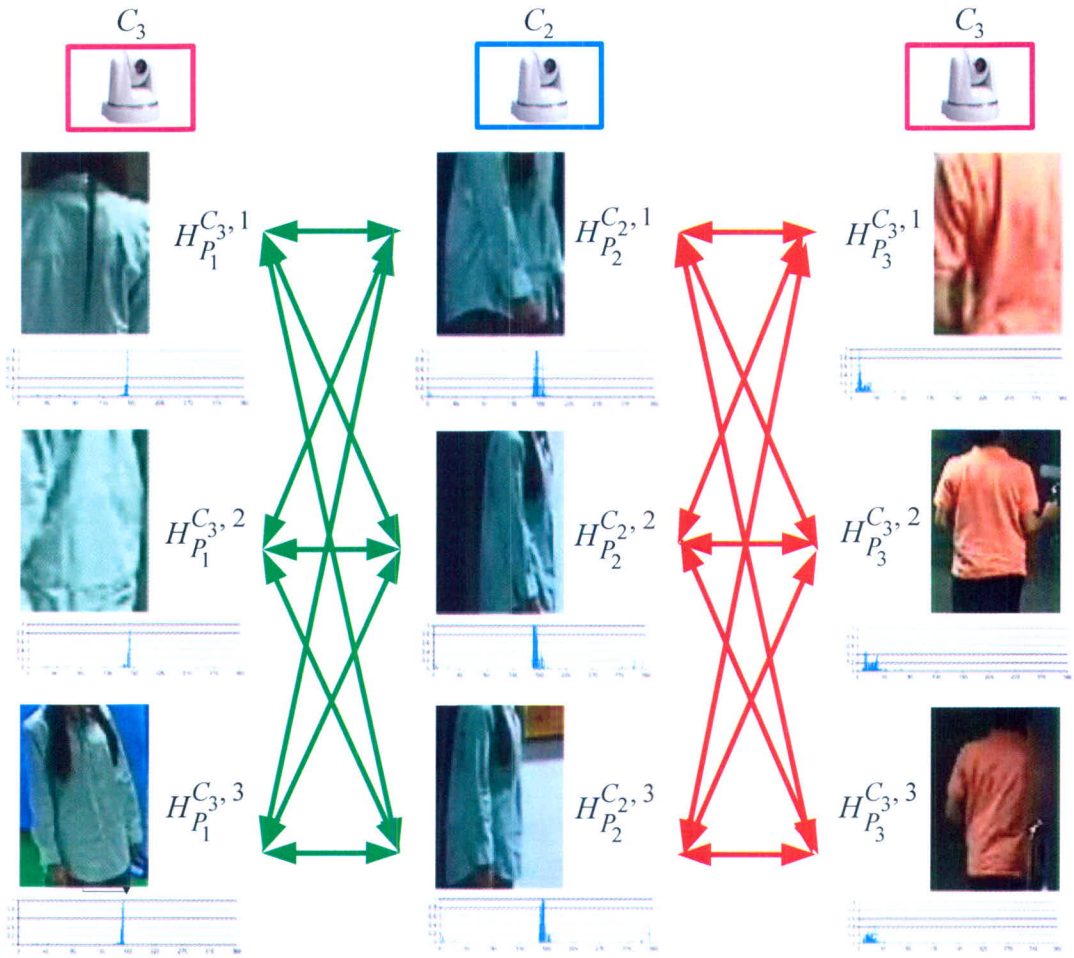
สำหรับการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์ในระบบที่นำเสนอจะเป็นการนำข้อมูลในหน่วยความจำระยะยาว (LT) ซึ่งประกอบไปด้วยข้อมูล Hue ฮิสโทแกรมของแต่ละบุคคลที่ได้จากกล้องวีดิทัศน์แต่ละตัว มาทำการเปรียบเทียบความเหมือนของคู่ฮิสโทแกรมระหว่างข้อมูลใน LT ที่ระบบมีอยู่ ยกตัวอย่างเช่น หาก $LT_P^C = \{LT_1^{1,1}, LT_1^{1,2}, LT_2^{2,1}\}$ นั่นหมายความว่าระบบการติดตามมีข้อมูลที่ใช้เป็นลักษณะของบุคคลที่ 1 ซึ่งได้จากการติดตามที่กล้องวีดิทัศน์ตัวที่ 1 จำนวนสองแบบจำลอง และมีข้อมูลที่ใช้เป็นลักษณะของบุคคลที่ 2 ซึ่งได้จากการติดตามที่กล้องวีดิทัศน์ตัวที่ 2 จำนวนหนึ่งแบบจำลอง ดังนั้นระบบการติดตามจะทำการเปรียบเทียบความเหมือนของฮิสโทแกรมทั้งหมด 2 คู่ คือ คู่ฮิสโทแกรมของ $LT_1^{1,1}$ กับ $LT_2^{2,1}$ และคู่ฮิสโทแกรมของ $LT_1^{1,2}$ กับ $LT_2^{2,1}$ เป็นต้น หากพบว่าฮิสโทแกรมคู่ใดมีความคล้ายคลึงกันก็จะทำการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์ สำหรับการตัดสินความคล้ายคลึงกันนั้นสามารถคำนวณได้จากค่าระยะห่าง Bhattacharyya ระหว่างคู่ของฮิสโทแกรมที่ต้องการเปรียบเทียบ ซึ่งหากพบกว่าค่าระยะห่าง Bhattacharyya ที่ได้มี

ค่าน้อยกว่าขีดเริ่มเปลี่ยนของระยะความเหมือน Bhattacharyya สำหรับการส่งต่อบุคคล ($BH_{Th}^{Handoff}$) ก็จะกำหนดให้คู่ของฮิสโทแกรมนั้นเป็นข้อมูลชุดเดียวกัน และทำการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์ทันที แสดงตัวอย่างการคำนวณดังสมการที่ (3.8)

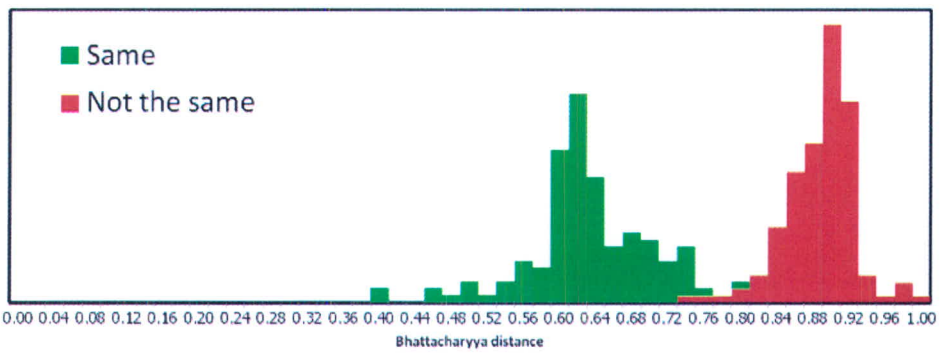
$$LT_{P_i}^{C_x} = LT_{P_j}^{C_y}, \quad BH(LT_{P_i}^{C_x}, LT_{P_j}^{C_y}) < BH_{Th}^{Handoff} \tag{3.8}$$

เมื่อ $BH(LT_{P_i}^{C_x}, LT_{P_j}^{C_y})$ คือ ระยะห่าง Bhattacharyya ที่ได้จากข้อมูล Hue ฮิสโทแกรมใน $LT_{P_i}^{C_x}$ และ $LT_{P_j}^{C_y}$

การคำนวณหาค่าขีดเริ่มเปลี่ยนของระยะ Bhattacharyya สำหรับการส่งต่อบุคคล ($BH_{Th}^{Handoff}$) ที่เหมาะสมทำได้จากการคำนวณระยะ Bhattacharyya ของคู่ Hue ฮิสโทแกรมที่ได้จากการติดตามของบุคคลเดียวกันระหว่างกล้องวีดิทัศน์และต่างบุคคลระหว่างกล้องวีดิทัศน์ ผลจากการคำนวณจะทำให้ทราบถึงแนวโน้มของค่าความเหมือนและความแตกต่างของบุคคลเดียวกันที่ติดตามได้จากกล้องวีดิทัศน์แต่ละตัว แสดงตัวอย่างการเปรียบเทียบ Hue ฮิสโทแกรมดังรูปที่ 3.6 ซึ่งประกอบด้วยกล้องวีดิทัศน์จำนวนสองตัว คือ กล้องวีดิทัศน์ตัวที่ 2 และ 3 ผลจากการเปรียบเทียบระยะห่าง Bhattacharyya ทำให้ทราบการกระจายตัวข้อมูลของบุคคลเดียวกันที่ถูกติดตามได้จากต่างกล้องวีดิทัศน์ ซึ่งพบว่าที่ค่าเบี่ยงเบนมาตรฐานเท่ากับ 3 ($3SD$) จากค่าเฉลี่ย จะให้ความมั่นใจได้ว่าระบบการติดตามจะสามารถส่งต่อบุคคลระหว่างกล้องวีดิทัศน์ซึ่งเป็นบุคคลเดียวกันได้ถึง 99.7 เปอร์เซ็นต์ ซึ่งตรงกับค่าระยะห่าง Bhattacharyya ที่เท่ากับ 0.75 วิทยานิพนธ์ได้นำค่าระยะห่าง Bhattacharyya ดังกล่าวไปใช้สำหรับการกำหนดค่าขีดเริ่มเปลี่ยนของระยะ Bhattacharyya สำหรับการส่งต่อบุคคล ($BH_{Th}^{Handoff}$) แสดงผลการกระจายตัวของข้อมูลระยะ Bhattacharyya ของบุคคลเดียวกันระหว่างกล้องวีดิทัศน์และต่างบุคคลระหว่างกล้องวีดิทัศน์ได้ดังรูปที่ 3.7



รูปที่ 3.6 ตัวอย่างการเปรียบเทียบ Hue ฮิสโทแกรมของบุคคลเดียวกันระหว่างกล้องวีดิทัศน์และต่างบุคคลระหว่างกล้องวีดิทัศน์



รูปที่ 3.7 การกระจายตัวของข้อมูลระยะ Bhattacharyya ของบุคคลเดียวกันระหว่างกล้องวีดิทัศน์และต่างบุคคลระหว่างกล้องวีดิทัศน์

จากรูปที่ 3.6 กำหนดให้ลูกศรเส้นสีเขียวแสดงการเชื่อมโยงข้อมูล Hue ฮิสโทแกรมที่ได้จากการติดตามบุคคลเดียวกันจากต่างกล้องวีดิทัศน์ ซึ่งสัมพันธ์กับสีของรูปการกระจายตัวข้อมูลในรูปที่ 3.7

ทั้งนี้หากระบบการติดตามพบว่ามีบุคคลเดียวกันถูกติดตามด้วยกล้องวีดิทัศน์มากกว่าหนึ่งตัว ระบบจะทำการคัดเลือกกล้องวีดิทัศน์ที่มีมุมมองต่อบุคคลดังกล่าวที่ดีที่สุดโดยใช้ทฤษฎีการคัดเลือกกล้องวีดิทัศน์ [20] ซึ่งวิทยานิพนธ์นี้ได้นำผลลัพธ์จากการเลือกกล้องวีดิทัศน์มาใช้ในการเคลื่อนที่ตามบุคคลซึ่งอยู่ในความสนใจอีกด้วย

3.5 การเลือกกล้องวีดิทัศน์ที่มีมุมมองการรับภาพต่อบุคคลที่ดีที่สุด

วิทยานิพนธ์นี้ได้ประยุกต์ใช้ทฤษฎีเกมที่มีความสำคัญในการวิเคราะห์และการตัดสินใจสำหรับสถานการณ์ที่มีความซับซ้อนในการเลือกกล้องวีดิทัศน์ที่มีมุมมองต่อบุคคลที่ดีที่สุด [20] ผลเฉลยของเกมจะขึ้นอยู่กับการนิยามปัญหาและการเลือกใช้ฟังก์ชันอรรถประโยชน์ต่าง ๆ ซึ่งเป็นเครื่องมือในการวิเคราะห์เกม ดังนี้

3.5.1 การนิยามปัญหา

ทฤษฎีเกมสามารถนำมาประยุกต์สำหรับการวิเคราะห์ปฏิสัมพันธ์หรือความขัดแย้งระหว่างผู้เล่นในเกมได้ ปัญหาในการเลือกกล้องวีดิทัศน์ที่มีมุมมองต่อบุคคลที่ดีที่สุดจัดเป็นกิจกรรมอย่างหนึ่งของระบบกล้องวีดิทัศน์ ที่มีความจำเป็นในการแลกเปลี่ยนข้อมูลรวมถึงการตัดสินใจภายใต้ข้อตกลงเดียวกัน กล้องวีดิทัศน์ที่มีอยู่ในระบบจะถูกนิยามให้เป็นผู้เล่นในเกม ซึ่งมีความพยายามที่จะสร้างผลตอบแทนหรือค่าอรรถประโยชน์จากบุคคลที่ต้องการติดตามให้กับตัวเองมากที่สุด ดังนั้นเพื่อให้ได้ผลเฉลยของเกมตามลักษณะดังกล่าว เกมนี้จึงเป็นเกมสำหรับผู้เล่นหลายคน (Multi-player game) ทั้งนี้หากมีมากกว่าหนึ่งบุคคลที่ต้องการติดตามในระบบ จะเรียกเกมดังกล่าวว่าเป็นเกมสำหรับผู้เล่นหลายคนที่มีหลายเกม (Multiple of multi-player game) ในรูปแบบเกมที่เล่นพร้อม ๆ กัน

การวิเคราะห์ปัญหาในงานวิจัย [20] พบว่าเป็นเกมสำหรับผู้เล่นหลายคน ที่มีจุดประสงค์ในการเลือกกล้องวีดิทัศน์ที่เหมาะสมที่สุดจากข้อมูลของมุมมองบุคคลที่ต้องการติดตาม ฟังก์ชันอรรถประโยชน์ต่าง ๆ ถูกนิยามขึ้นมาเพื่อใช้เป็นกลยุทธ์ในการเปรียบเทียบค่าความพึงพอใจที่กล้องวีดิทัศน์แต่ละตัวมีต่อบุคคลที่ต้องการติดตาม ผลเฉลยของเกมจะขึ้นอยู่กับค่าอรรถประโยชน์ 3 อย่างด้วยกัน คือ

1. ค่าอรรถประโยชน์ทั้งหมด (Global utility) เป็นค่าที่แสดงปริมาณความพึงพอใจทั้งหมดของประสิทธิภาพในการติดตาม

2. ค่าอรรถประโยชน์ของกล้องวีดิทัศน์ (Camera utility) เป็น และ ค่าอรรถประโยชน์ของกล้องวีดิทัศน์นี้จะขึ้นอยู่กับเกณฑ์การประเมินมุมมองของกล้องวีดิทัศน์ที่มีต่อบุคคล ซึ่งเป็นข้อตกลงเดียวกันที่ใช้กับกล้องวีดิทัศน์ทุกตัว

3. ค่าอรรถประโยชน์ของบุคคล (Person utility) เป็นค่าที่แสดงปริมาณเหมาะสมที่แต่ละบุคคลสมควรถูกติดตามโดยกล้องวีดิทัศน์บางตัวในระบบ ดังนั้นค่าอรรถประโยชน์ของบุคคลจึงเป็นข้อมูลสำคัญที่นำไปสู่การแก้ปัญหาในการเลือกกล้องวีดิทัศน์ที่มีมุมมองที่ดีที่สุดต่อแต่ละบุคคลนั่นเอง

เป้าหมายของผู้เล่นหรือกล้องวีดิทัศน์แต่ละตัวในเกม คือ การทำให้ค่าอรรถประโยชน์ทั้งหมดที่เกิดจากการคำนวณฟังก์ชันอรรถประโยชน์ต่าง ๆ มีค่าสูงที่สุด ซึ่งจะเป็นการบ่งบอกว่ากล้องวีดิทัศน์ตัวใดจะเป็นผู้ชนะในเกมรอบนั้น นั่นหมายถึงผลเฉลยของเกมซึ่งกล้องวีดิทัศน์ดังกล่าวจะเป็นกล้องวีดิทัศน์ที่มีมุมมองที่ดีที่สุดต่อบุคคลที่ต้องการติดตาม สำหรับการหาผลเฉลยของเกมในแต่ละรอบจะใช้กลไกการต่อรองจากเซตของความน่าจะเป็นในการเลือกกล้องวีดิทัศน์ หากผลเฉลยในรอบถัดไปมีการเปลี่ยนแปลงซึ่งส่งผลให้ต้องเปลี่ยนกล้องวีดิทัศน์เพื่อทำการติดตามบุคคล กระบวนการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์ก็จะเกิดขึ้นนั่นเอง

3.5.2 การกำหนดฟังก์ชันอรรถประโยชน์

เมื่อเลือกกล้องวีดิทัศน์ที่มีมุมมองที่ดีที่สุดได้แล้วจึงทำการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์ ดังนั้นการกำหนดฟังก์ชันอรรถประโยชน์จึงนิยามจากค่าความพึงพอใจของกล้องวีดิทัศน์ที่มีต่อบุคคลเป็นหลัก ซึ่งสามารถคำนวณได้จากหัวข้อที่ 2.6

ข้อเสียของการเลือกกล้องวีดิทัศน์จาก [20] คือ ต้องการกระบวนการส่งต่อบุคคลจากกล้องวีดิทัศน์สู่กล้องวีดิทัศน์ที่มีความแม่นยำมาก การสร้างแบบจำลองลักษณะของบุคคลที่นำเสนอสามารถช่วยเพิ่มประสิทธิภาพให้กับขั้นตอนการเลือกกล้องวีดิทัศน์ [20] ได้ วิทยานิพนธ์นี้จึงได้ทำการทดลองเพื่อเปรียบเทียบผลของการติดตามจากวิธีที่นำเสนอ และวิธีอ้างอิง [20] ซึ่งได้แสดงผลการทดลองไว้ในหัวข้อที่ 4.2

3.6 การควบคุมการส่าย และก้มเงยของกล้องวีดิทัศน์แบบ PTZ

วิธีการหนึ่งที่เพิ่มประสิทธิภาพการติดตามบุคคลเป้าหมายให้มีประสิทธิภาพมากยิ่งขึ้น นั้นคือ การนำประโยชน์ทางด้านความสามารถในการส่าย และก้มเงยของกล้องวีดิทัศน์แบบ PTZ มาประยุกต์ใช้กับงาน ดังนั้นกระบวนการหนึ่งที่วิทยานิพนธ์ฉบับนี้ให้ความสำคัญเพื่อเพิ่มประสิทธิภาพให้กับระบบการติดตามบุคคลด้วยการทำงานร่วมกันของกล้องวีดิทัศน์ คือ การส่งพารามิเตอร์เพื่อควบคุมกล้องวีดิทัศน์ PTZ เพื่อปรับเปลี่ยนมุมมองของกล้องวีดิทัศน์ในการส่ายและก้มเงยเพื่อเคลื่อนที่ตามบุคคล โดยกล้องวีดิทัศน์จะทำการเคลื่อนที่ตามบุคคลที่สนใจซึ่งได้จากการคัดเลือกบุคคลด้วยมือ (Manual person selection) ซึ่งหากพบว่ามีกล้องวีดิทัศน์มากกว่าหนึ่งตัวสามารถทำการติดตามบุคคลที่สนใจได้พร้อมกัน ระบบจะกำหนดให้กล้องวีดิทัศน์ที่มีมุมมองต่อบุคคลที่สนใจที่ดีที่สุดในการเคลื่อนที่ตามอย่างอัตโนมัติ ทั้งนี้ลักษณะในการเคลื่อนที่ของกล้องวีดิทัศน์ คือ การทำให้บริเวณบุคคลเป้าหมายนั้นมาอยู่ตำแหน่งบริเวณกึ่งกลางภาพวีดิทัศน์ ขั้นตอนการคำนวณพารามิเตอร์ต่าง ๆ จะเลือกใช้เฉพาะข้อมูลตำแหน่งจุดภาพในพิกัด 2 มิติเท่านั้น

พารามิเตอร์ที่ใช้ในการควบคุมกล้องวีดิทัศน์ PTZ นั้น ในวิทยานิพนธ์ฉบับนี้ประกอบไปด้วยสองส่วนคือ พารามิเตอร์ของการส่าย (ψ) หรือ มุมของการส่าย (Pan angle) และพารามิเตอร์ของการก้มเงย (ϕ) หรือมุมของการก้มเงย (Tilt angle) สามารถหาค่าพารามิเตอร์ทั้งสองนี้ได้ ดัง

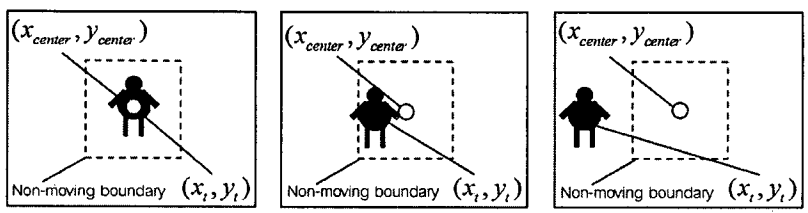
สมการ (3.9) และ (3.10) ตามลำดับ ส่วนพารามิเตอร์ของการซูมนั้น ในวิทยานิพนธ์ฉบับนี้ไม่ได้ให้ความสำคัญไว้ เนื่องจากการติดตามบุคคลเป้าหมายนั้นไม่จำเป็นต้องเก็บรายละเอียดของบุคคลมากนัก อีกทั้งถ้ากล้องวีดิทัศน์ PTZ มีการซูมเกิดขึ้นจะทำให้มุมมองในการติดตามบุคคลแคบลงอีกด้วย

$$\psi = \frac{Diff_x}{\beta_x} = \frac{x_t - x_{center}}{\beta_x} \tag{3.9}$$

$$\phi = \frac{Diff_y}{\beta_y} = \frac{y_t - y_{center}}{\beta_y} \tag{3.10}$$

เมื่อกำหนดให้ $Diff_x$ และ $Diff_y$ คือ ระยะห่างในแนวนอนและแนวตั้งระหว่างตำแหน่งกึ่งกลางบุคคลเป้าหมาย (x_t, y_t) กับตำแหน่งจุดศูนย์กลางภาพวีดิทัศน์ (x_{center}, y_{center}) ส่วน β_x และ β_y คือ อัตราของจำนวนจุดภาพต่อการเคลื่อนที่ของกล้องไป 1 องศาในแนวแกนนอนและแกนตั้งตามลำดับ

ในกรณีที่บุคคลเป้าหมายอยู่บริเวณกึ่งกลางภาพอยู่แล้ว ดังรูปที่ 3.8(ก) หรืออยู่ใกล้กับบริเวณกึ่งกลางภาพแต่อยู่ภายในขอบเขตการเปลี่ยนมุมมองกล้องวีดิทัศน์ (Non-moving boundary) ดังรูปที่ 3.8(ข) นั้น ระบบจะไม่มีการส่งพารามิเตอร์คำสั่งควบคุมให้กล้องวีดิทัศน์เพื่อเปลี่ยนมุมมอง แต่ถ้าในลำดับภาพวีดิทัศน์ใดกึ่งกลางบุคคลเป้าหมายได้เคลื่อนที่ออกจากขอบเขตวีดิทัศน์ที่ไม่ต้องเปลี่ยนมุมมอง จึงจะคำนวณตำแหน่งจุดกึ่งกลางบุคคลเป้าหมายและส่งพารามิเตอร์เพื่อให้กล้องวีดิทัศน์ PTZ เคลื่อนที่ ผลลัพธ์ที่ได้ คือ ตำแหน่งของบุคคลเป้าหมายนั้นจะอยู่กึ่งกลางภาพวีดิทัศน์หรือภายในบริเวณขอบเขตวีดิทัศน์ที่ไม่ต้องเปลี่ยนมุมมองตลอดเวลา



(ก) (ข) (ค)

รูปที่ 3.8 การควบคุมการเคลื่อนที่ของกล้องวีดิทัศน์แบบ PTZ

3.7 การนับบุคคล

ในหัวข้อนี้จะนำเสนอวิธีการที่ใช้สำหรับการประเมินค่าระบบการติดตามบุคคลที่นำเสนอ โดยจะเริ่มทำการนับบุคคลเมื่อระบบสามารถทำการตรวจจับบุคคลที่เคลื่อนไหวได้ดังที่กล่าวไว้แล้วในหัวข้อที่ 2.3.1 ดังนั้นหากมีบุคคลถูกตรวจจับได้จากการเคลื่อนไหวนี้จะนับเป็นหนึ่งบุคคล โดยมีแนวคิดในการยอมรับผลการติดตามที่สำเร็จ คือ เมื่อมีบุคคลถูกตรวจจับในลำดับภาพวีดิทัศน์ใด ๆ แล้ว ระบบการติดตามจะต้องแสดงผลลัพธ์ของการติดตามอย่างถูกต้องได้ด้วย แสดงภาพตัวอย่างการนับบุคคลดังรูปที่ 3.9 (ก) ซึ่งสามารถนับได้สองบุคคลและผลลัพธ์จากการติดตามสามารถบ่งชี้ได้ว่านับได้สองบุคคลเช่นกัน แต่สำหรับรูปที่ 3.9 (ข) สามารถนับบุคคลได้เพียงหนึ่งบุคคลแต่ระบบการติดตามสามารถนับได้สองบุคคลเนื่องจากสัญญาณภาพรบกวนเข้ามาเกี่ยวข้อง

(ก) (ข)

รูปที่ 3.9 ภาพตัวอย่างการนับบุคคล

ผลจากการนับบุคคลนี้จะถูกนำไปใช้ต่อในการวัดประสิทธิภาพของระบบการติดตามในบทถัดไป