

การตรวจหาและติดตามบุคคลอย่างต่อเนืองด้วยกล้องวีดิทัศน์หลายตัว
โดยใช้สีและการเคลื่อนไหวที่

นายนนทรรัฐ บำรุงเกียรติ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2551
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

CONTINUOUS PERSON TRACKING ACROSS MULTIPLE CAMERAS
USING COLOR AND MOTION

Mr.Nontarat Bumrungrat

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Electrical Engineering

Department of Electrical Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2008

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การตรวจหาและติดตามบุคคลอย่างต่อเนืองด้วยกล้องวิดีโอทัศนหลายตัวโดยใช้สีและการเคลื่อนที่
โดย	นายธนรัฐ บำรุงเกียรติ
สาขาวิชา	วิศวกรรมไฟฟ้า
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.สุภาวดี อร่ามวิทย์
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	ผู้ช่วยศาสตราจารย์ ดร.ธนารัตน์ ชลิดาพงศ์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

..... คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร.บุญสม เลิศหิรัญวงศ์)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(รองศาสตราจารย์ ดร. เฉษฐา ชินรุ่งเรือง)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร. สุภาวดี อร่ามวิทย์)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(ผู้ช่วยศาสตราจารย์ ดร. ธนารัตน์ ชลิดาพงศ์)

..... กรรมการ
(อาจารย์ ดร. สุรีย์ พุมรินทร์)

..... กรรมการภายนอกมหาวิทยาลัย
(ดร. ศุภกร สิทธิไชย)

นันทรัฐ บำรุงเกียรติ : การตรวจหาและติดตามบุคคลอย่างต่อเนื่องด้วยกล้องวิดีโอทัศน
หลายตัวโดยใช้สีและการเคลื่อนที่. (CONTINUOUS PERSON TRACKING ACROSS
MULTIPLE CAMERAS USING COLOR AND MOTION) อ.ที่ปรึกษาวิทยานิพนธ์หลัก
: ผศ.ดร.สุภาวดี อร่ามวิทย์, อ. ที่ปรึกษาวิทยานิพนธ์ร่วม : ผศ.ดร.ธนารัตน์ ชลิตาพงศ์,
114 หน้า.

การติดตามวัตถุในเวลาจริงนั้นถูกนำมาประยุกต์ใช้ในงานร่วมกับวิดีโอทัศนในหลากหลาย
ประเภท เช่น ใช้เพื่อประโยชน์ในการติดตามบุคคลที่สนใจของระบบรักษาความปลอดภัย ใช้ใน
การติดตามวัตถุต่างๆด้วยหุ่นยนต์ติดตาม หรือใช้ติดตามรถยนต์ในระบบขนส่งจราจรอัจฉริยะ เป็น
ต้น ซึ่งงานประยุกต์เกี่ยวกับการติดตามวัตถุในเวลาจริงดังกล่าวนี้ ต้องการขั้นตอนวิธีการของการ
ติดตามที่มีประสิทธิภาพ แม่นยำ โดยเฉพาะอย่างยิ่งเมื่อถูกนำมาใช้งานในข้อจำกัดที่ยากต่อการ
ติดตามหรือในสถานการณ์จำกัดต่างๆ เช่น เมื่อมุมมองของวิดีโอทัศนมีการเปลี่ยนแปลง ความสว่าง
ของบริเวณที่ต้องการติดตามมีการเปลี่ยนแปลง หรือเมื่อมีการใช้งานส่งผ่านข้อมูลระหว่างกล้อง
วิดีโอทัศนหลายตัวให้ทำงานร่วมกัน เป็นต้น วิทยานิพนธ์ฉบับนี้นำเสนอขั้นตอนวิธีการติดตามบุคคล
อย่างต่อเนื่องในเวลาจริง โดยใช้กล้องวิดีโอทัศนแบบสาย ก้มเงย และซูม (Pan Tilt Zoom Camera;
PTZ) 2 ตัว ทำงานร่วมกัน โดยอาศัยขั้นตอนวิธีการย้ายตามค่าเฉลี่ย ซึ่งส่วนใหญ่แล้ว
คุณลักษณะสำคัญที่ใช้ในการติดตามวัตถุหรือบุคคลนั้น จะอาศัยเพียงค่าสีเท่านั้น ซึ่งจะเกิด
ปัญหาเป็นอย่างมากเมื่อความสว่างหรือมุมมองของกล้องวิดีโอทัศนมีการเปลี่ยนแปลง ซึ่ง
วิทยานิพนธ์เล่มนี้จะใช้คุณลักษณะสำคัญสองประการคือ สี เช่น สีผิวหรือสีเสื้อผ้าและขอบของ
ร่างกายของบุคคล โดยที่กล้องวิดีโอทัศนตัวที่หนึ่งในระบบจะทำหน้าที่เก็บข้อมูลคุณลักษณะสำคัญ
ของบุคคลที่สนใจและทำการติดตามด้วยขั้นตอนวิธีดังกล่าวข้างต้น และเมื่อบุคคลที่สนใจนั้นออก
จากมุมมองของกล้องวิดีโอทัศนตัวที่หนึ่งมายังมุมมองของกล้องวิดีโอทัศนตัวที่สอง กล้องวิดีโอทัศนตัวที่
สองในระบบจะรับค่าคุณลักษณะสำคัญนั้นๆจากกล้องวิดีโอทัศนตัวที่หนึ่ง และทำการติดตามบุคคล
ที่สนใจนั้นอย่างต่อเนื่องต่อไป นอกจากนี้ยังอาศัยการเทียบมาตรฐานทางสีระหว่างกล้องวิดีโอทัศน
เพื่อให้การส่งข้อมูลสีระหว่างกล้องวิดีโอทัศนทั้งสองตัวในระบบสามารถติดตามได้อย่างถูกต้อง

ภาควิชา.....วิศวกรรมไฟฟ้า.....ลายมือชื่อนิสิต.....
สาขาวิชา.....วิศวกรรมไฟฟ้า.....ลายมือชื่ออาจารย์ที่ปรึกษา.....
ปีการศึกษา.....2551.....ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4970369021 : MAJOR ELECTRICAL ENGINEERING

KEYWORDS: REAL TIME OBJECT TRACKING / MEAN SHIFT ALGORITHM / CAMERA CONTROL

NONTARAT BUMRUNGKIAT: CONTINUOUS PERSON TRACKING ACROSS MULTIPLE CAMERAS USING COLOR AND MOTION. ADVISOR: ASST. PROF. SUPAVADEE ARAMVITH, Ph.D., CO-ADVISOR : ASST. PROF. THANARAT CHALIDABHONGSE, Ph.D., 114 pp.

The real time object tracking is used in many applications such as video surveillance, human machine interfaces, robot tracking, and intelligent transportation systems. These applications need good object detection and tracking methods. However, there are still some difficulties in tracking such as the cases where there are changes in background, view point, or illumination especially when using multiple cameras. This thesis proposed a framework for handover method in continuously tracking a person of interest across cooperative pan-tilt-zoom cameras (PTZ Camera). The algorithm here is called the mean shift algorithm. Most tracking algorithms use only one cue (such as color). The color features are not always discriminative enough for target localization because illumination or viewpoints tend to change. Moreover the background may be of a color similar to that of the target. The system, the continuous person tracking across cooperative PTZ cameras by mean shift tracking that using color and shape histogram to be feature distributions. Color and shape distributions of interested person are used to register the target person across cameras. For first camera, we select interested person for tracking using skin color, cloth color and boundary of body. To handover tracking process among two cameras, the second camera receives color and shape cues of a target person from the first camera and using linear color calibration to help with handover process. Our experimental results demonstrate color and shape feature in mean shift algorithm is capable for continuously and accurately track the target person across cameras.

Department : Electrical Engineering Student's Signature

Field of Study : Electrical Engineering Advisor's Signature

Academic Year : 2008 Co-Advisor's Signature

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความกรุณาและความช่วยเหลืออย่างดียิ่งจาก ผู้ช่วยศาสตราจารย์ ดร. สุภาวดี อร่ามวิทย์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ และ ผู้ช่วยศาสตราจารย์ ดร. ธนารัตน์ ชลิตาพงศ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ที่ได้สละเวลาในการให้คำปรึกษา และคำแนะนำอันมีประโยชน์ ตลอดจนช่วยตรวจทานแก้ไขวิทยานิพนธ์และบทความตีพิมพ์ต่าง ๆ จนวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี

กราบขอบพระคุณบิดา มารดา และครอบครัว ที่ให้โอกาสและสนับสนุนทางการศึกษาดำเนินการศึกษาดำเนินการดีตลอดมา รวมถึงให้กำลังใจและเอาใจใส่ในทุกๆด้าน ซึ่งทำให้ผู้วิจัยได้รับแต่สิ่งดีๆ เสมอมาและปลุกฝังให้เติบโตขึ้นมาเป็นคนดีในสังคมเพื่อให้เป็นกำลังสำคัญในการทำประโยชน์แก่ประเทศได้อย่างเต็มที่

กราบขอบพระคุณบรรพคณาจารย์ทุกท่าน ตั้งแต่อดีตจนถึงปัจจุบันที่ได้ประสิทธิ์ประสาทความรู้ทางวิชาการและจริยธรรม รวมถึงปลุกฝังและสร้างแรงบันดาลใจในความเป็นวิศวกรที่ดีให้กับผู้วิจัย

ขอขอบคุณคุณนงลักษณ์ คุณณัฐชัย คุณพิชัย และคุณจันทนา รวมถึงสมาชิกในกลุ่มวิจัยเทคโนโลยีทางวิศวกรรมและกลุ่มวิจัยเทคโนโลยีทางโทรคมนาคมทุกคนที่คอยให้คำปรึกษา และช่วยแนะนำวิเคราะห์สิ่งต่าง ๆ ให้มีวิสัยทัศน์ที่กว้างไกลมากขึ้น รวมทั้งให้ความร่วมมืออย่างดียิ่งในความช่วยเหลือสำหรับการทดลองต่างๆ

ขอขอบคุณเพื่อนๆ คณะวิศวกรรมโทรคมนาคม สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง และเพื่อนๆ ม.6/1 โรงเรียนสิงห์บุรี ที่คอยสร้างเสียงหัวเราะ ให้ความช่วยเหลือในสิ่งที่ผู้วิจัยร้องขอ และเป็นทีปรึกษาปัญหาในการดำเนินชีวิตต่างๆ ได้อย่างดี

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ญ
สารบัญภาพ.....	ฎ
1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 ภาพรวมของงานวิจัยที่เกี่ยวข้อง.....	7
1.2.1 งานวิจัยที่นำเสนอเกี่ยวกับระบบกล้องวีดิทัศน์ในรูปแบบต่างๆ.....	7
1.2.2 งานวิจัยที่นำเสนอเกี่ยวกับการเทียบมาตรฐานกล้องวีดิทัศน์.....	8
1.2.3 งานวิจัยที่นำเสนอเกี่ยวกับขั้นตอนวิธีการติดตามบุคคล.....	10
1.3 แนวทางวิทยานิพนธ์.....	12
1.4 เป้าหมายและขอบเขตของงานวิจัย.....	13
1.5 ขั้นตอนและวิธีการดำเนินงานวิจัย.....	13
1.6 ประโยชน์ที่คาดว่าจะได้รับ.....	13
2 ความรู้พื้นฐานและทฤษฎีที่เกี่ยวข้อง.....	14
2.1 ปริภูมิสีและเวกเตอร์ค่าสี.....	14
2.1.1 ปริภูมิสี RGB.....	14
2.1.2 ปริภูมิสี YUV และ YCbCr.....	16
2.1.3 ปริภูมิสี HSV.....	19
2.2 การหาขอบภาพ.....	20
2.2.1 ตัวดำเนินการสำหรับหาขอบภาพ.....	21
2.2.2 ตัวอย่างวิธีการหาขอบภาพ.....	24
2.2.2.1 การหาขอบภาพโดยใช้อนุพันธ์อันดับที่หนึ่งทัศน์.....	24
2.2.2.2 การหาขอบภาพโดยใช้อนุพันธ์อันดับที่สองทัศน์.....	25
2.3 แบบจำลองกล้องวีดิทัศน์และแบบจำลองการเคลื่อนที่กล้องวีดิทัศน์.....	26
2.3.1 แบบจำลองกล้องวีดิทัศน์แบบรูเข็ม.....	27

2.3.2 การเคลื่อนที่ของกล้องวิดีโอ	27
2.4 การเทียบมาตรฐานทางของกล้องวิดีโอ	29
2.5 พื้นฐานการติดตามวัตถุหรือบุคคลในลำดับภาพวิดีโอ	30
2.6 การประมาณความหนาแน่น	32
2.6.1 การประมาณความหนาแน่นด้วยฮิสโทแกรม	33
2.6.2 ตัวประมาณความหนาแน่นแบบง่าย	34
2.6.3 การประมาณความหนาแน่นแบบเคอร์เนล	35
2.7 ขั้นตอนวิธีการย้ายตามค่าเฉลี่ย	37
2.8 การวัดประสิทธิภาพด้วยฟังก์ชันความเหมือนของ Bhattacharyya	41
3. โครงสร้างของระบบการติดตามบุคคลอย่างต่อเนื่องที่นำเสนอ	43
3.1 รูปแบบของการทำงานของระบบที่นำเสนอ	44
3.1.1 กล้องวิดีโอ PTZ	45
3.1.2 ลักษณะโครงสร้างการจัดวางตำแหน่งกล้องวิดีโอ PTZ ทั้ง 2 ตัว	45
3.1.3 ไดอะแกรมการทำงานโดยรวม	46
3.2 การเทียบมาตรฐานทางสีของกล้องวิดีโอ	47
3.3 ขั้นตอนวิธีการติดตามบุคคลที่นำเสนอ	52
3.3.1 ขั้นตอนการดึงคุณลักษณะสำคัญและการประมาณฟังก์ชันความหนาแน่นของบุคคลเป้าหมาย	52
3.3.1.1 การดึงคุณลักษณะสำคัญทางสีเพื่อประมาณฟังก์ชันความหนาแน่น	55
3.3.1.2 การดึงคุณลักษณะสำคัญทางขอบเพื่อประมาณฟังก์ชันความหนาแน่น	57
3.3.2 ขั้นตอนวิธีการติดตามบุคคลด้วยการย้ายตามค่าเฉลี่ย	63
3.4 การควบคุมการส่าย ก้มเงย และซูมของกล้องวิดีโอ PTZ เพื่อติดตามบุคคลเป้าหมาย	67
3.4.1 การคำนวณมุมในการเปลี่ยนมุมมองของกล้องวิดีโอ PTZ	67
4. ผลการจำลองระบบ	69
4.1 ผลการจำลองระบบในส่วนของขั้นตอนวิธีการเทียบมาตรฐานทางสีที่นำเสนอ	70

4.1.1 ผลการทดลองในเชิงปริมาณในส่วนของขั้นตอนวิธีการเทียบมาตรฐาน ทางสีที่นำเสนอ.....	70
4.1.2 ผลการทดลองในเชิงคุณภาพในส่วนของขั้นตอนวิธีการเทียบมาตรฐาน ทางสีที่นำเสนอ.....	75
4.2 ผลการจำลองระบบในส่วนของขั้นตอนวิธีการติดตามบุคคล.....	76
4.2.1 ผลการทดสอบเมื่อใช้คุณลักษณะสำคัญทางสีที่ปริภูมิสีต่างกัน.....	76
4.2.2 ผลการทดสอบเมื่อใช้คุณลักษณะสำคัญทางสีและขอบ.....	78
4.3 ผลการจำลองระบบในส่วนการทำงานร่วมกันระหว่างกล้องวีดิทัศน์ PTZ.....	80
4.3.1 ผลการทดลองในเชิงปริมาณในส่วนการทำงานร่วมกันระหว่าง กล้องวีดิทัศน์ PTZ.....	81
4.3.2 ผลการทดลองในเชิงคุณภาพในส่วนการทำงานร่วมกันระหว่าง กล้องวีดิทัศน์ PTZ.....	82
5. สรุปผลการวิจัยและข้อเสนอแนะ.....	94
5.1 สรุปผลการวิจัย.....	94
5.2 อภิปรายผลการวิจัย.....	95
5.2 ข้อเสนอแนะ.....	96
รายการอ้างอิง.....	97
ภาคผนวก.....	100
ประวัติผู้เขียนวิทยานิพนธ์.....	114

สารบัญตาราง

	หน้า
ตารางที่ 2.1	แสดงตัวอย่างการคำนวณการประมาณความหนาแน่นแบบเคอร์เนล.....36
ตารางที่ 3.1	แสดงสีและข้อมูลสี ในปริภูมิสี RGB ของตารางสีอ้างอิงที่ใช้ในการทดลอง.....49
ตารางที่ 4.1	แสดงค่าเฉลี่ยของค่าความผิดพลาดกำลังสอง ในแต่ละสีของตารางสีอ้างอิงที่ รับได้จากกล้องวิดีโอที่สแกนตัวที่สองก่อนและหลังทำการเทียบมาตรฐานทางสี.....71
ตารางที่ 4.2	แสดงการเปรียบเทียบค่าเฉลี่ยของค่าความผิดพลาดกำลังสอง (MSE) ในแต่ละสีของตารางสีอ้างอิงจากวิธีการเทียบมาตรฐานทางสีในรูปแบบต่างๆ.....73
ตารางที่ 4.3	แสดงค่าสี R, G และ B ของภาพบุคคลก่อนหลังเทียบมาตรฐานทางสี.....74
ตารางที่ 4.4	แสดงค่าค่าเฉลี่ยของค่าความผิดพลาดกำลังสองของสีบุคคลทดสอบ จากกล้องวิดีโอที่สแกนก่อนและหลังทำการเทียบมาตรฐานทางสี.....74
ตารางที่ 4.5	แสดงลำดับภาพตัวอย่างการทดลองก่อนหลังทำการเทียบมาตรฐานทางสี.....75
ตารางที่ 4.6	ผลการเปรียบเทียบความคล้ายของฟังก์ชันความหนาแน่นของตัวแทนบุคคล เป้าหมายที่กล้องวิดีโอที่สแกนตัวที่ 1 และตัวที่ 2 ก่อนและหลังทำการเทียบ มาตรฐานทางสี.....81
ตารางที่ 4.7	แสดงลำดับภาพวิดีโอที่สแกนตัวอย่างเปรียบเทียบการทดลองการติดตามบุคคล เป้าหมายก่อนและหลังทำการเทียบมาตรฐานทางสีให้กับกล้องวิดีโอที่สแกน.....84
ตารางที่ 4.8	แสดงลำดับภาพวิดีโอที่สแกนตัวอย่างการส่าย ก้มเงยเพื่อติดตามบุคคล เป้าหมายอย่างต่อเนื่องด้วยกล้องวิดีโอที่สแกนทั้งสองตัวในระบบด้วยสีในปริภูมิ ระดับสีเทา.....85
ตารางที่ 4.9	แสดงลำดับภาพวิดีโอที่สแกนตัวอย่างการส่าย ก้มเงยเพื่อติดตามบุคคลเป้าหมาย อย่างต่อเนื่องด้วยกล้องวิดีโอที่สแกนทั้งสองตัวในระบบด้วยสีในแซลแนล H ของปริภูมิสี HSV.....86
ตารางที่ 4.10	แสดงลำดับภาพวิดีโอที่สแกนตัวอย่างการส่าย ก้มเงยเพื่อติดตามบุคคลเป้าหมาย อย่างต่อเนื่องด้วยกล้องวิดีโอที่สแกนทั้งสองตัวในระบบด้วยสีในแซลแนล H ของปริภูมิสี HSV และขอบ.....88
ตารางที่ 4.11	แสดงสถานการณ์ที่บุคคลเป้าหมายเคลื่อนที่เข้ากล้องวิดีโอที่สแกนตัวที่ 1 แต่ไม่เคลื่อนที่เข้ากล้องวิดีโอที่สแกนตัวที่ 2.....89
ตารางที่ 4.12	แสดงสถานการณ์ที่บุคคลเป้าหมายเคลื่อนที่เข้ากล้องวิดีโอที่สแกนตัวที่ 1 และมี บุคคลที่ไม่ใช่เป้าหมายเคลื่อนที่เข้ากล้องวิดีโอที่สแกนตัวที่ 2 ในทิศทางเดียวกัน.....91

สารบัญภาพ

	หน้า
รูปที่ 1.1	แสดงตัวอย่างระบบสอดส่องรักษาความปลอดภัยด้วยระบบกล้องวีดิทัศน์..... 1
รูปที่ 1.2	แสดงภาพที่รับได้จากกล้องวีดิทัศน์สามตัวในแต่ละมุมมอง..... 2
รูปที่ 1.3	แสดงภาพของกล้องวีดิทัศน์แต่ละตัวที่เกิดจากการบิดเบือนของเลนส์..... 8
รูปที่ 1.4	แสดงภาพวัตถุที่ใช้ในการเปรียบเทียบมาตรฐานทางสี..... 9
รูปที่ 1.5	แสดงตำแหน่งการติดตั้งกล้องที่มีข้อจำกัดทางด้านความสว่างในแต่ละบริเวณ..... 9
รูปที่ 2.1	ปริภูมิสีแบบ RGB (สีแดง เขียว และน้ำเงิน)..... 15
รูปที่ 2.2	ปริภูมิสีแบบ YUV และปริภูมิสีแบบ YCbCr..... 18
รูปที่ 2.3	ปริภูมิสี HSV..... 19
รูปที่ 2.4	แสดงแบบจำลองของขอบ..... 21
รูปที่ 2.5	แสดงวิธีการพื้นฐานในการหาขอบภาพ..... 23
รูปที่ 2.6	แสดงตัวดำเนินการการหาขอบภาพด้วยวิธีโซเบล..... 24
รูปที่ 2.7	แสดงตัวอย่างการหาขอบภาพด้วยวิธีโซเบล..... 24
รูปที่ 2.8	แสดงตัวดำเนินการต่างๆที่ใช้ในการคำนวณหาขอบภาพ..... 25
รูปที่ 2.9	แสดงการหาขอบภาพด้วยตัวดำเนินการดีกรีสอง..... 26
รูปที่ 2.10	แสดงตัวอย่างตัวดำเนินการดีกรีสองสำหรับการหาขอบภาพ..... 26
รูปที่ 2.11	การฉายภาพแบบทัศนมิติของแบบจำลองกล้อง..... 27
รูปที่ 2.12	การเคลื่อนที่ของกล้องวีดิทัศน์..... 28
รูปที่ 2.13	ลำดับภาพวีดิทัศน์ตัวอย่างที่ได้จากการเคลื่อนที่ของกล้องแบบส่าย..... 28
รูปที่ 2.14	ลำดับภาพวีดิทัศน์ตัวอย่างที่ได้จากการเคลื่อนที่ของกล้องแบบก้มเงย..... 28
รูปที่ 2.15	ลำดับภาพวีดิทัศน์ตัวอย่างที่ได้จากการเคลื่อนที่ของกล้องแบบหมุน..... 29
รูปที่ 2.16	สนามการเคลื่อนที่ของการส่ายและก้มเงย..... 29
รูปที่ 2.17	ตัวอย่างภาพวัตถุที่ใช้ในการเทียบมาตรฐานกล้องวีดิทัศน์..... 30
รูปที่ 2.18	ตัวอย่างลำดับภาพวีดิทัศน์ที่มีวัตถุเคลื่อนที่..... 31
รูปที่ 2.19	แสดงทิศทางการย้ายตำแหน่งจุดศูนย์กลางการคำนวณ..... 39
รูปที่ 2.20	แสดงแผนภาพแสดงขั้นตอนวิธีการย้ายตามค่าเฉลี่ย..... 41
รูปที่ 3.1	กล้องวีดิทัศน์ในการทดลอง..... 45
รูปที่ 3.2	แสดงการติดตั้งระบบการติดตามบุคคลที่สนใจด้วยกล้องวีดิทัศน์ 2 ตัว..... 46
รูปที่ 3.3	ไดอะแกรมโดยรวมสำหรับการติดตามบุคคลอย่างต่อเนื่องที่นำเสนอ..... 47

รูปที่ 3.4	แสดงรูปตารางอ้างอิงการเทียบมาตรฐานสีที่ใช้ในเทียบมาตรฐานสี.....	48
รูปที่ 3.5	แสดงการหาฟังก์ชันความหนาแน่นสำหรับการติดตามบุคคลที่นำเสนอ.....	55
รูปที่ 3.6	แสดงภาพตัวอย่างการคำนวณการแปลงค่าสีจากปริภูมิ RGB เป็นปริภูมิสี HSV ในแซลแนล H.....	56
รูปที่ 3.7	แสดงภาพจำลองการคำนวณฟังก์ชันความหนาแน่นของข้อมูลสีด้วยวิธีการประมาณความหนาแน่นด้วยฟังก์ชันหลัก.....	57
รูปที่ 3.8	แสดงฟังก์ชันความหนาแน่นของข้อมูลสีบริเวณบุคคลเป้าหมาย.....	57
รูปที่ 3.9	ตัวกระทำของการหาขอบภาพด้วยวิธี Sobel.....	59
รูปที่ 3.10	วิธีการคอนโวลูชัน.....	59
รูปที่ 3.11	แสดงภาพตัวอย่างการคำนวณการแปลงค่าสีจากปริภูมิ RGB เป็นปริภูมิระดับสีเทา.....	60
รูปที่ 3.12	แสดงภาพจำลองการคำนวณฟังก์ชันความหนาแน่นของข้อมูลขอบด้วยวิธีการประมาณความหนาแน่นด้วยฟังก์ชันหลัก.....	61
รูปที่ 3.13	แสดงฟังก์ชันความหนาแน่นของข้อมูลขอบบริเวณบุคคลเป้าหมาย.....	61
รูปที่ 3.14	แสดงฟังก์ชันความหนาแน่นที่จะนำไปใช้ในการติดตามบุคคล.....	62
รูปที่ 3.15	แสดงแผนภาพการหาฟังก์ชันความหนาแน่นที่จะนำไปใช้ในการติดตาม.....	62
รูปที่ 3.16	แสดงการหาความหนาแน่นของความน่าจะเป็นสำหรับการติดตามบุคคล.....	63
รูปที่ 3.17	แสดงตัวอย่างฟังก์ชันความหนาแน่นที่จะนำไปใช้อ้างอิงเป็นบุคคลเป้าหมาย.....	64
รูปที่ 3.18	แสดงตัวอย่างฟังก์ชันความหนาแน่นเป้าหมายที่ต้องการติดตามที่ตำแหน่ง.....	64
รูปที่ 3.19	รูปแบบการควบคุมกล้องวิดีโอ PTZ.....	68
รูปที่ 4.1	แสดงตำแหน่งการเก็บภาพจากกล้องวิดีโอสองตัวในระบบ.....	70
รูปที่ 4.2	แสดงภาพที่ได้จากกล้องวิดีโอสองตัวเพื่อเป็นข้อมูลในการเทียบมาตรฐานทางสี.....	70
รูปที่ 4.3	แสดงภาพที่ได้จากกล้องวิดีโอสองตัวที่หลังจากมาตรฐานทางสีที่นำเสนอ.....	71
รูปที่ 4.4	ลำดับภาพตัวอย่างเปรียบเทียบของผลการติดตามบุคคลเป้าหมายเมื่อใช้คุณลักษณะสำคัญทางสีที่ปริภูมิสีต่างกัน.....	77
รูปที่ 4.5	ลำดับภาพวิดีโอตัวอย่างเปรียบเทียบของผลการติดตามบุคคลเป้าหมายเมื่อความส่องสว่างมีการเปลี่ยนแปลงอย่างกะทันหัน.....	79

	หน้า
รูปที่ 4.6	ลำดับภาพวีดิทัศน์ตัวอย่างเปรียบเทียบของผลการติดตามบุคคลเป้าหมาย เมื่อบุคคลเป้าหมายมีการเปลี่ยนทิศทางการเดิน..... 79
รูปที่ 4.7	ลำดับภาพวีดิทัศน์ตัวอย่างเปรียบเทียบของผลการติดตามบุคคลเป้าหมาย เมื่อบุคคลเป้าหมายมีการเปลี่ยนทิศทางการเดิน..... 79
รูปที่ 4.8	แสดงผลการควบคุมกล้องวีดิทัศน์เพื่อเปลี่ยนมุมมองเพื่อติดตามบุคคล เป้าหมายที่สนใจ..... 83
รูปที่ 4.9	แสดงทิศทางการเดินของบุคคลเป้าหมาย..... 87
รูปที่ 4.10	แสดงทิศทางการเดินของบุคคลเป้าหมายและบุคคลที่ไม่ใช่เป้าหมาย เมื่อบุคคลเป้าหมายไม่เดินทางเข้าสู่มุมมองของกล้องวีดิทัศน์ตัวที่ 2..... 89
รูปที่ 4.11	แสดงทิศทางการเดินของบุคคลเป้าหมายและบุคคลที่ไม่ใช่เป้าหมาย เมื่อมีทิศทางการเดินเดียวกัน..... 91
รูปที่ 4.12	แสดงทิศทางการเดินของบุคคลเป้าหมายและบุคคลที่ไม่ใช่เป้าหมาย เมื่อมีทิศทางการเดินตรงข้ามกัน..... 91

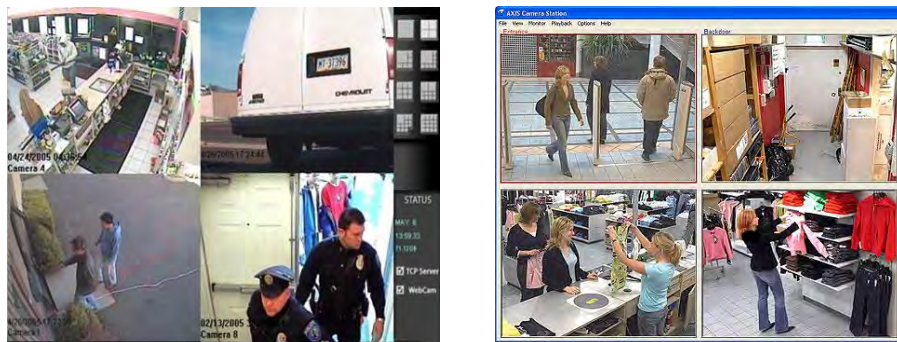
บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

อันตรายที่เกิดจากการกระทำผิดของมนุษย์ในปัจจุบันนั้น ทราบกันดีว่ายิ่งโลกพัฒนาไปมากเท่าใด อันตรายที่เกิดจากการกระทำผิดของมนุษย์ที่มีความรุนแรงเพิ่มมากขึ้นเท่านั้น อีกทั้งยังพบได้ในหลายหลากรูปแบบที่ลึ้นยากที่จะป้องกันและแก้ไข แต่ด้วยความก้าวหน้าทางเทคโนโลยีที่มีขีดความสามารถเพิ่มขึ้นอย่างไม่หยุดนิ่งเช่นเดียวกัน ทำให้ปัญหาดังกล่าวอาจจะลดความรุนแรงไปได้บ้างเมื่อมีระบบรักษาความปลอดภัยที่ดี อีกทั้งเนื่องจากข้อจำกัดของมนุษย์บางประการทำให้การป้องกันอันตรายจากความผิดของมนุษย์ด้วยตนเองดังกล่าวนั้นอาจจะไม่เพียงพอ มนุษย์จึงนำเทคโนโลยีในด้านต่างๆ เข้าร่วมเพื่อป้องกันปัญหาต่างๆ กล้องวีดิทัศน์เป็นเทคโนโลยีหนึ่งที่มนุษย์นำมาใช้และได้ถูกนำมาประยุกต์ใช้เพื่อเพิ่มขีดความสามารถให้กับมนุษย์ในหลายๆ ด้าน โดยเฉพาะอย่างยิ่งงานที่เกี่ยวข้องกับระบบสอดส่องรักษาความปลอดภัย

ระบบสอดส่องรักษาความปลอดภัย (Surveillance system) เป็นระบบการทำงานชนิดหนึ่งที่สามารถช่วยเหลือ และแก้ไขข้อจำกัดทางการทำงานของมนุษย์ได้เป็นอย่างดี และเพื่อเพิ่มขีดความสามารถให้กับระบบสอดส่องรักษาความปลอดภัย ระบบกล้องวีดิทัศน์จึงเป็นส่วนหนึ่งของระบบรักษาความปลอดภัยอยู่เสมอ ดังรูปที่ 1.1 แสดงตัวอย่างระบบสอดส่องรักษาความปลอดภัยทั้งภายในและภายนอกอาคารซึ่งส่วนใหญ่แล้วจะอาศัยการทำงานร่วมกันของกล้องวีดิทัศน์หลายๆตัว



รูปที่ 1.1 แสดงตัวอย่างระบบสอดส่องรักษาความปลอดภัยด้วยระบบกล้องวีดิทัศน์

ระบบการทำงานร่วมกันระหว่างกล้องวิดีโอทัศนหลายตัว หรือที่เรียกว่า ระบบกล้องวิดีโอทัศน (Camera system) นั้น เป็นรูปแบบของการประยุกต์ใช้งานกล้องวิดีโอทัศนมากกว่าหนึ่งตัวให้ทำงานร่วมกันไม่ว่าจะด้านการประมวลผลภาพร่วมกัน การควบคุมหรือรับส่งข้อมูลร่วมกัน ซึ่งส่วนใหญ่แล้วระบบดังกล่าว มีวัตถุประสงค์หลักเพื่อทำหน้าที่ในการตรวจหา ติดตาม ระบุหรือรู้จำบุคคลที่สนใจ ซึ่งในงานวิจัยที่ผ่านมา เห็นได้ชัดเจนว่า ระบบกล้องวิดีโอทัศนหลายตัวนี้ ส่วนใหญ่จะถูกนำไปใช้งานเกี่ยวกับระบบรักษาความปลอดภัย เช่น ในงานวิจัยของ S.Lim และคณะ [1,2] จะมีรูปแบบการทำงานของระบบรักษาความปลอดภัยภายในอาคารที่ระบบประกอบด้วยกล้องวิดีโอทัศนอยู่กับที่หนึ่งตัว ทำหน้าที่ตรวจหาและติดตามบุคคล จากนั้นเมื่อได้ตำแหน่งของวัตถุก็จะส่งตำแหน่งนี้ไปให้กล้องวิดีโอทัศนภายในโครงข่ายกล้องทั้งหมด เพื่อเลือกกล้องวิดีโอทัศนตัวใดตัวหนึ่งใน การติดตามวัตถุต่อโดยใช้ตำแหน่งที่กล้องวิดีโอทัศนตัวหลักส่งมาให้

ไม่เพียงแต่นำไปใช้ในงานรักษาความปลอดภัยเท่านั้น ระบบกล้องวิดีโอทัศนยังสามารถนำไปประยุกต์ใช้งานแบบอื่นได้ เช่น งานวิจัยของ H.Hongo และคณะ [3] ที่นำไปประยุกต์ใช้กับการเรียนการสอนทางไกล (Distance lecture system) เป็นระบบที่ใช้รับภาพผู้สอนจากกล้องวิดีโอทัศนอยู่กับที่ และหาตำแหน่งการเคลื่อนที่ของผู้สอนด้วยกล้องวิดีโอทัศนตัวอื่นๆในระบบ

เมื่อกล่าวถึงวิธีการในการประมวลผลภาพและทำงานร่วมกันของกล้องวิดีโอทัศนหลายตัวในระบบนั้น ข้อมูลสีถือเป็นข้อมูลที่สำคัญในการประมวลผลภาพทางวิดีโอทัศน เนื่องมาจากข้อมูลสีเป็นข้อมูลที่เข้าถึงได้ง่ายที่สุด และเหมาะสำหรับการประยุกต์ใช้งานที่มนุษย์สามารถเข้าใจได้ง่ายที่สุด จึงเห็นได้ว่า งานวิจัยที่ผ่านมา ข้อมูลสีนั้นจะถูกนำมาใช้งานทางด้านคอมพิวเตอร์วิทัศน์มากที่สุด เช่น ในงานวิจัยของ P. Amnuaykanchanasin และคณะ [4] อาศัยการตรวจหาสีผิวของมนุษย์ (Skin color) ร่วมกันภายในระบบกล้องวิดีโอทัศนและงานวิจัยของ N. Soontranon และคณะ [5] จะใช้ข้อมูลสีผิวในการตรวจหาเพื่ออ่านภาษามือ ซึ่งสิ่งหนึ่งที่ต้องคำนึงถึงเมื่อจะใช้คุณลักษณะทางสี ก็คือ ปริภูมิสี (Color space) ที่ใช้ในการประมวลผลต้องสอดคล้องกับสถานการณ์การทำงานและสีของภาพที่รับได้จากกล้องวิดีโอทัศนแต่ละตัวในระบบต้องไม่มีความผิดเพี้ยนออกจากกัน เพื่อป้องกันการประมวลผลร่วมกันที่ผิดพลาด

สิ่งสำคัญสำหรับระบบกล้องวิดีโอทัศนที่อาศัยสีในการประมวลผล นั่นคือ ถ้าการประมวลผลทางสีของภาพจากกล้องวิดีโอทัศนแต่ละตัวมีความผิดเพี้ยนไป อันเนื่องมาจากสาเหตุบางประการ เช่น คุณสมบัติของตัวจับภาพของกล้องวิดีโอทัศนแต่ละตัวต่างกัน สภาพแวดล้อมการติดตั้งกล้องวิดีโอทัศนแต่ละตัวต่างกัน การผลิตจากโรงงานอุตสาหกรรมที่ต่างกัน เป็นต้น ดังนั้น ในการทำงานของระบบการทำงานร่วมกันระหว่างกล้องวิดีโอทัศนหลายตัวนั้น จึงต้องมีวิธีการของการเทียบมาตรฐาน

ทางสีของกล้องวิดีโอแต่ละตัวให้เหมือนกันก่อน ดังเช่นงานวิจัยของ G.Unal และ A. Yezzi [8] นำเสนอปัญหาของการเทียบมาตรฐานพารามิเตอร์ของกล้องวิดีโอการเทียบมาตรฐานกล้องวิดีโอด้วยสี ซึ่งหลักการ คือ นำภาพที่ได้จากกล้องวิดีโอแต่ละตัว มาคำนวณหาค่าสีภายในภาพของกล้องวิดีโอแต่ละตัว แล้วเปรียบเทียบกับวัตถุที่ใช้ในการเปรียบเทียบมาตรฐานทางสี (Object color checker) เช่น ลูกบาศก์สี (Color cube) เป็นต้น จุดประสงค์เพื่อให้ระบบกล้องวิดีโอหลายตัวนี้ ทำงานเปรียบเทียบได้ว่าเป็นกล้องวิดีโอตัวเดียวกัน มองภาพและประมวลผลทางสีของภาพได้อย่างถูกต้องแม่นยำมากยิ่งขึ้น ในทำนองเดียวกันกับ งานวิจัยของ F. Porikli และคณะ [8,9] นำเสนอวิธีการเทียบมาตรฐานกล้องวิดีโอเพื่อใช้ในระบบการตรวจหาบุคคลด้วยกล้องวิดีโอ งานวิจัยนี้ได้ให้ความสำคัญกับ การสร้างให้ระบบกล้องวิดีโอทำงานร่วมกันได้อย่างมีประสิทธิภาพ โดยหาค่าสหสัมพันธ์ของค่าสีระหว่างกันของกล้องวิดีโอแต่ละคู่ภายในระบบ เพื่อให้ระบบสามารถตรวจหาบุคคลหรือวัตถุที่สนใจได้ ในสถานการณ์ที่กล้องวิดีโอแต่ละตัวถูกติดตั้งภายใต้แสงสว่างที่ต่างกัน แต่ปัญหาคือต้องมีการคำนวณที่ซับซ้อนอาจจะมีปัญหาเมื่อต้องการใช้งานในเวลาจริงและระบบประกอบด้วยกล้องวิดีโอหลายๆตัว ตัวอย่างภาพที่รับได้จากกล้องวิดีโอสามตัวในแต่ละมุมมองแสดงดังรูปที่ 1.2



รูปที่ 1.2 แสดงภาพที่รับได้จากกล้องวิดีโอสามตัวในแต่ละมุมมอง

ดังนั้นสิ่งที่วิทยานิพนธ์นี้ให้ความสำคัญสิ่งแรกในการสร้างระบบกล้องวิดีโอ คือ การมองถึงปัญหาการเทียบมาตรฐานสีของกล้องวิดีโอ จากงานวิจัยที่ผ่านมาดังกล่าว ส่วนใหญ่จะกำหนดให้แสงสว่างบริเวณที่กล้องแต่ละตัวติดตั้งนั้น มีค่าเท่ากันหรือประมาณให้มีค่าใกล้เคียงกัน แต่ในความเป็นจริงแล้วปัญหาความสว่างของบริเวณที่ติดตั้งกล้องวิดีโอแต่ละตัวที่ต่างกันนั้น ไม่สามารถควบคุมได้ ซึ่งปัญหาดังกล่าวอาจจะมีผลผิดพลาดบ้างเมื่อถูกนำมาใช้งานเกี่ยวกับการตรวจหาบุคคล แต่จะเป็นปัญหาใหญ่สำหรับการทำงานที่เกี่ยวข้องกับการติดตามบุคคลที่สนใจ เนื่องจากถ้าข้อมูลผิดพลาด การส่งข้อมูลเพื่อติดตามบุคคลเป้าหมายเดียวกันในระบบอาจจะผิดพลาดได้

นอกจากปัญหาเรื่องความแตกต่างของสิ่งที่รับได้จากกล้องวิดีโอทั้งสองตัวแล้ว ถ้าใช้งานกล้องวิดีโอประเภทอยู่กับที่ (Stationary camera) เพื่อการติดตามบุคคลอย่างต่อเนื่องนั้น (กล้องวิดีโอที่มีที่ตั้งตำแหน่งมุมมองการรับภาพและอัตราซูมที่แน่นอน) จะเกิดข้อจำกัดของระบบขึ้นในกรณีที่หากวัตถุเป้าหมายเคลื่อนที่ออกจากมุมมองการรับภาพ ทำให้ไม่สามารถติดตามวัตถุเป้าหมายต่อเนื่องได้ ด้วยความก้าวหน้าทางเทคโนโลยีทางด้านกล้องวิดีโอในระยะหลัง ทำให้เริ่มมีการนำกล้องวิดีโอประเภทสาย ก้มเงยและซูม (PTZ camera) (กล้องวิดีโอที่ตำแหน่งมุมมองการรับภาพสามารถเปลี่ยนแปลงได้ในลักษณะการเคลื่อนที่แนวสาย (Pan) และแนวก้มเงย (Tilt) และสามารถซูม (Zoom) รายละเอียดของวัตถุเป้าหมายได้ โดยลักษณะทั้งหมดสามารถควบคุมผ่านคอมพิวเตอร์ส่วนบุคคล) มาประยุกต์ใช้งาน เพื่อลดข้อจำกัดดังที่ได้กล่าวมา ทำให้เกิดรูปแบบงานที่หลากหลายและประโยชน์ที่เพิ่มมากขึ้น ดังนั้นวิทยานิพนธ์ฉบับนี้จึงมีแนวความคิดในการจำลองระบบกล้องวิดีโอหลายตัวสำหรับการติดตามบุคคลที่สนใจอย่างต่อเนื่อง โดยใช้กล้องวิดีโอประเภท PTZ ในระบบ 2 ตัว ทำงานร่วมกันในลักษณะการประมวลผลและแลกเปลี่ยนข้อมูลระหว่างกัน

ในส่วนของขั้นตอนวิธีที่ใช้ในการติดตามบุคคลในเวลาจริงนั้น เนื่องจากการติดตามบุคคลหรือวัตถุใดๆ ในลำดับภาพวิดีโอที่นำมาประยุกต์ใช้กับงานทางด้านคอมพิวเตอร์วิทัศน์ในหลากหลายประเภท เช่น ระบบสอดส่องรักษาความปลอดภัย [10,11] หรือ ห้องอัจฉริยะ [12] เป็นต้น ซึ่งจากงานวิจัยในอดีตเกี่ยวกับการติดตามบุคคลในเวลาจริงนั้น จะแบ่งออกเป็น 2 หลักการใหญ่ คือ กระบวนการติดตามจากล่างขึ้นบน (A bottom-up tracking process) และกระบวนการติดตามจากบนลงล่าง (A top-down tracking process) ในส่วนของกระบวนการติดตามจากล่างขึ้นบนนั้น หมายถึง การนำเพียงข้อมูลบางส่วนเป็นตัวแทนของเป้าหมายเพื่อใช้ข้อมูลบางส่วนนี้ช่วยในการติดตาม ซึ่งตามหลักการแล้วจะประกอบไปด้วยสองส่วนสำคัญคือ การกำหนดตัวแทนเป้าหมาย (Target representation) และการหาตำแหน่งเป้าหมาย (Target localization) ซึ่งเหมาะกับงานต้องการติดตามเป้าหมายที่เป้าหมายอาจจะมีการเปลี่ยนแปลงภายในลำดับภาพและเหมาะสมกับงานที่ไม่ซับซ้อน ส่วนกระบวนการติดตามจากบนลงล่างนั้น หมายถึง การมองภาพรวมทั้งหมดของข้อมูลก่อนแล้วจึงนำข้อมูลที่วัดได้นำมาทำนายทิศทางการเคลื่อนที่ของเป้าหมายต่อไป ซึ่งจะอาศัยหลักการของตัวกรองและความสัมพันธ์ระหว่างข้อมูล (Filtering and Data association) ซึ่งมักจะใช้งานในสถานการณ์ที่ซับซ้อน จำเป็นต้องมีการเรียนสถานการณ์ในลำดับภาพก่อนหน้าเพื่อประมาณตำแหน่งเป้าหมายในลำดับภาพถัดไปภายใต้สมมุติฐานตามที่กำหนด ซึ่งการนำหลักการใดมาใช้ขึ้น ขึ้นอยู่กับหลักกรนั้นๆ มาสามารถนำประยุกต์ใช้กับงาน

ประเภทใดได้ดี เช่น ในการติดตามใบหน้าที่บุคคลนั้นจะเหมาะสมกับการกำหนดตัวแทนเป้าหมายเป็นใบหน้ามากกว่าการหาตัวแปรที่เปลี่ยนแปลงของเป้าหมาย เช่น ความเร็วของการเคลื่อนที่ใบหน้า ในขณะที่งานเกี่ยวกับการสอดส่องตรวจตราทางอากาศ (Aerial video surveillance) ซึ่งการเคลื่อนที่ของเป้าหมายจะเป็นองค์ประกอบด้านความเร็วที่สามารถประมาณได้ จึงเหมาะสมกับหลักการของตัวกรองและความสัมพันธ์ระหว่างข้อมูลมากกว่า เพื่อประมาณตำแหน่งเป้าหมายในลำดับภาพต่อไปด้วยความเร็ว เป็นต้น

ในหลักการของตัวกรองและความสัมพันธ์ระหว่างข้อมูลจะมองการทำงานเป็นลักษณะของสเตท (The state space approach) โดยเป้าหมายนั้นจะแทนด้วยลำดับสเตท $\{x_k\}_{k=0,1,\dots}$ ที่มีฟังก์ชันที่เปลี่ยนแปลงตามเวลาและขึ้นอยู่กับสเตทก่อนหน้าเป็น $x_k = f_k(x_{k-1}, v_k)$ และลักษณะของสเตทที่วัดได้ $\{z_k\}_{k=0,1,\dots}$ ที่มีฟังก์ชันที่เปลี่ยนแปลงตามเวลาที่ได้จากการวัดในสเตทปัจจุบันเป็น $z_k = h_k(x_k, n_k)$ โดยที่ $\{v_k\}_{k=0,1,\dots}$ และ $\{n_k\}_{k=0,1,\dots}$ จะเป็นสัญญาณรบกวนที่ถูกกำหนดขึ้น

จุดประสงค์หลักของการติดตามแบบอาศัยหลักการของตัวกรองและความสัมพันธ์ระหว่างข้อมูลนั้น คือการประมาณสเตท x_k ด้วยข้อมูลที่ได้จากการวัดในสเตทก่อนหน้า z_{k-1} ซึ่งสเตทนั้นโดยทั่วไปแล้วอาจจะแทนด้วยฟังก์ชันความน่าจะเป็นของความหนาแน่น (Probability Density function; pdf) ถ้ากำหนดสัญญาณรบกวนเป็นฟังก์ชันแบบเกาส์เซียน (Gaussian function) และฟังก์ชันของ f_k และ h_k เป็นแบบเชิงเส้น (Linear function) การแก้ปัญหาจะอาศัยหลักการของตัวกรองแบบคาร์มาล (Kalman filter) [14] แต่ถ้าฟังก์ชันของ f_k และ h_k เป็นแบบไม่เชิงเส้น (Non-Linear function) แล้ว ตัวกรองแบบเอ็กซ์เทนเดดคาร์มาล (Extended kalman filter; EKF) [15] จะถูกนำมาใช้แก้ปัญหาแทน หรือถ้าทราบจำนวนสเตทที่แน่นอน อาจจะใช้ตัวกรองแบบฮิดเดนมาร์คอฟ (Hidden markov model filter; HMM) มาช่วยในการแก้ปัญหาได้ เป็นต้น ซึ่งตัวอย่างดังกล่าวมานี้จำเป็นต้องทราบพารามิเตอร์ของฟังก์ชันความน่าจะเป็นของความหนาแน่น เช่น ค่าเฉลี่ย หรือ ค่าโคแวนเรียนซ์ของฟังก์ชัน (Parametric density function) ซึ่งจะมีการคำนวณที่ซับซ้อนมากเมื่อนำมาใช้ในงานในเวลาจริง เนื่องจากข้อมูลฟังก์ชันความน่าจะเป็นของความหนาแน่นนั้นมีมาก ในงานวิจัยระยะหลังๆ จึงเริ่มมีการคิดค้นการนำเพียงตัวอย่างสุ่มของฟังก์ชันความน่าจะเป็นของความหนาแน่นนั้นมาเป็นข้อมูลเพื่อทำการติดตามเป้าหมาย โดยอาศัยเทคนิคของตัวกรองแบบพาทิเคิล (Particle filter) เป็นต้น

อีกกระบวนการหนึ่ง คือ กระบวนการติดตามจากล่างขึ้นบนซึ่งประกอบด้วยการกำหนดตัวแทนเป้าหมายและหาตำแหน่งของเป้าหมายนั้นจำเป็นต้องมีการระบุเจาะจงเป้าหมายลงในลำดับภาพ ส่วนสำคัญของกระบวนการนี้คือต้องมีขั้นตอนการตรวจสอบความถูกต้อง

(Registration method) เพื่อให้แน่ใจได้ว่าตำแหน่งของเป้าหมายนั้นในลำดับภาพต่อไปนั้นให้ความถูกต้องที่มากที่สุด ซึ่งเหมาะสำหรับสมมุติฐานที่ตำแหน่งของเป้าหมายที่ลำดับภาพที่ต่อเนื่องกันไม่มีการเปลี่ยนแปลงตำแหน่งมากนัก การคำนวณจึงไม่จำเป็นต้องมีการทำนายด้วยหลักการที่ซับซ้อน เพียงมีขั้นตอนตรวจสอบความถูกต้องก็เพียงพอ ในงานวิจัยของ [16] นำเสนอขั้นตอนวิธีการติดตามบุคคลโดยอาศัยฟังก์ชันหลัก (Kernel based object tracking) ซึ่งเป็นกระบวนการติดตามแบบล่างขึ้นบน และมีประสิทธิภาพในการติดตามวัตถุแบบไม่แข็งเกร็ง (Non rigid object) ได้อย่างรวดเร็ว เหมาะสำหรับการนำมาใช้งานในเวลาจริง

ซึ่งแต่ละวิธีที่กล่าวมานี้ต่างมีจุดเด่นจุดด้อยต่างกันไปหรืออีกนัยหนึ่งคือ ขั้นตอนวิธีแต่ละวิธีนั้นต่างเหมาะสมกับการทำงานในสถานการณ์ที่ต่างกัน สำหรับการทำงานในเวลาจริงนั้น สิ่งสำคัญอันดับแรกของการทำงานในเวลาจริงนั้นคือ ความรวดเร็ว ไม่ซับซ้อนในการคำนวณ ใช้ทรัพยากรในการประมวลผลน้อย แต่ยังคงประสิทธิภาพการทำงานที่ดีอยู่ งานวิจัยของ D.Comanicu และคณะ [17] จึงเสนอขั้นตอนวิธีการย้ายตามค่าเฉลี่ยขึ้นซึ่งเป็นหนึ่งในกระบวนการติดตามแบบล่างขึ้นบน ซึ่งเหมาะสำหรับการทำงานในเวลาจริง และไม่จำเป็นต้องทราบพารามิเตอร์ของฟังก์ชันความน่าจะเป็นของความหนาแน่นของเป้าหมายก่อน หรือที่เรียกว่า (Non-parametric density function) ซึ่งเป็นหลักการที่มีการคำนวณที่ไม่ซับซ้อนและสามารถติดตามบุคคลได้เป็นอย่างดี แต่ยังมีข้อจำกัดอยู่บ้างในเรื่องของข้อมูลสีของวัตถุที่สนใจ คล้ายกับฉากหลัง วิทยานิพนธ์ฉบับนี้จึงให้ความสำคัญกับอีกปัญหาหนึ่ง นั่นคือ การพัฒนาปรับปรุงความบกพร่องของขั้นตอนวิธีการย้ายตามค่าเฉลี่ยในเรื่องของข้อมูลสีดังกล่าวด้วยการเพิ่มข้อมูลอื่นๆ ให้กับขั้นตอนการย้ายค่าเฉลี่ย เพื่อให้ขั้นตอนการย้ายค่าเฉลี่ยนี้สามารถทำงานได้อย่างมีประสิทธิภาพมากขึ้น

สรุปได้ว่าวิทยานิพนธ์ฉบับนี้นำเสนอรูปแบบของระบบกล้องหลายตัว โดยนำมาประยุกต์ใช้กับการติดตามการเคลื่อนที่ของบุคคลอย่างต่อเนื่อง นำเสนอด้วยระบบกล้องวิดิทัศน์สองตัวที่ถูกติดตั้งในต่างมุมมองกันภายในอาคาร และอาศัยเทคนิคการเทียบมาตรฐานทางสีให้กับกล้องวิดิทัศน์ทั้งสองตัวในระบบเพื่อลดปัญหาความเพี้ยนทางสี หลังจากทำการเทียบมาตรฐานทางสีให้กับกล้องวิดิทัศน์ทั้งสองตัวในระบบแล้ว จะกำหนดให้กล้องทั้งสองตัวในระบบประมวลผลแบบขนานหรือทำงานร่วมกันในการติดตามบุคคลที่สนใจตลอดเวลา พร้อมทั้งส่งพารามิเตอร์ของการส่ายกัมเมยและซุมเพื่อควบคุมการทำงานของกล้องวิดิทัศน์แต่ละตัวให้มีการติดตามบุคคลที่สนใจอย่างต่อเนื่องด้วย

1.2 ภาพรวมของงานวิจัยที่เกี่ยวข้อง

ในหัวข้อนี้ขอกล่าวถึงงานวิจัยที่เกี่ยวข้องกับวิทยานิพนธ์ฉบับนี้ ซึ่งจะแบ่งออกเป็น 3 หัวข้อสำคัญ คือ นำเสนอภาพรวมของงานวิจัยที่ผ่านมาบางส่วนเกี่ยวกับการทำงานของระบบกล้องวิดีโอทัศนียภาพในรูปแบบต่างๆ จากนั้นนำเสนองานวิจัยที่เกี่ยวข้องกับการเทียบมาตรฐานทางให้กับระบบกล้องวิดีโอทัศนียภาพ และในที่สุดท้ายจะกล่าวถึงงานวิจัยเกี่ยวกับขั้นตอนวิธีการติดตามบุคคลที่เคยถูกนำเสนอ เพื่อนำมาเป็นแนวความคิดเพื่อสร้างระบบการติดตามบุคคลอย่างต่อเนื่องในเวลาจริงด้วยระบบกล้องวิดีโอทัศนียภาพให้สามารถทำงานได้เป็นอย่างดี

1.2.1 งานวิจัยที่นำเสนอเกี่ยวกับระบบกล้องวิดีโอทัศนียภาพในรูปแบบต่าง ๆ

เนื่องจากวิทยานิพนธ์ฉบับนี้เป็นการศึกษาเกี่ยวกับการทำงานของระบบกล้องวิดีโอทัศนียภาพหรือเรียกอีกอย่างหนึ่งว่า การมองเห็นการประยุกต์ใช้และปัญหาที่จะทำให้เกิดกล้องวิดีโอทัศนียภาพสามารถทำงานร่วมกันได้มากกว่า 1 ตัว ดังนั้นจึงขอกล่าวถึงงานวิจัยที่มีการนำระบบกล้องวิดีโอทัศนียภาพหลายตัวในรูปแบบต่างๆ พอสังเขป

S.Lim [1,2] เสนอระบบสอดส่องรักษาความปลอดภัยอัตโนมัติ โดยใช้กล้องวิดีโอทัศนียภาพอยู่กับที่ 1 ตัวทำหน้าที่เป็นกล้องตัวหลัก ร่วมกับกล้องวิดีโอทัศนียภาพ PTZ อีก 2 ตัว ทำหน้าที่เปลี่ยนมุมมองติดตามบุคคลที่สนใจในแนวสาย และก้มเงย รวมทั้งซูมขยายรายละเอียดของบุคคลเป้าหมายที่สนใจ โดยจะประมวลผลเฉพาะภาพที่รับจากกล้องวิดีโอทัศนียภาพอยู่กับที่เท่านั้น ซึ่งงานวิจัยนี้จะใช้ขั้นตอนวิธีการระบวงการลบฉากหลัง เมื่อได้ตำแหน่งของบุคคลเป้าหมายแล้วจะอ้างอิงไปยังพิกัดของกล้องวิดีโอทัศนียภาพ PTZ เพื่อคำนวณพารามิเตอร์การเคลื่อนที่ของแต่ละกล้องที่ต้องใช้สำหรับเคลื่อนที่ไปหาเป้าหมาย ดังนั้นข้อดีของระบบนี้ คือ ลดขั้นตอนในการประมวลผลภาพที่รับมาจากกล้องวิดีโอทัศนียภาพ PTZ

J. Kang และคณะ [6] นำเสนอระบบติดตามวัตถุเป้าหมายระหว่างกล้องวิดีโอทัศนียภาพอยู่กับที่และกล้องวิดีโอทัศนียภาพ PTZ โดยมีการส่งต่อข้อมูลของวัตถุเป้าหมายระหว่างกล้องทั้งสอง โดยลักษณะสีจะถูกส่งต่อจากกล้องวิดีโอทัศนียภาพหนึ่งตัวไปยังอีกตัวหนึ่งเพื่อระบุเป้าหมายที่ต้องการจากกล้องคนแต่ละตัว สำหรับกระบวนการตรวจหาสำหรับกล้องวิดีโอทัศนียภาพ PTZ จะเสนอกระบวนการปรับฉากหลัง โดยใช้การแปลงสัมพรรค (Affine transform) เพื่อคำนวณการเคลื่อนที่ของแต่ละจุดภาพที่สัมพันธ์กับพารามิเตอร์ในการสายและก้มเงยของกล้องวิดีโอทัศนียภาพ PTZ เพื่อพิจารณาจุดภาพว่าเป็นการเคลื่อนที่อันเนื่องจากการสายและก้มเงยของกล้องวิดีโอทัศนียภาพหรือการเคลื่อนที่ของวัตถุ ส่วนในการติดตามบุคคลเป้าหมายระหว่างกล้อง นอกจากจะใช้ลักษณะสีแล้ว ยังจะใช้

ลักษณะการแปลงโฮโมกราฟี (Homography transform) มาอ้างอิงเป้าหมายระหว่างกล้องวิดีโอทัศน์เพราะว่ากล้องวิดีโอทัศน์ทั้งสองมีมุมมองการรับภาพ (Field of view) ที่ซ้อนเหลื่อมกันอยู่ (Overlapping view)

H. Hongo และคณะ [3] เสนอระบบที่นำไปใช้ในงานประยุกต์การรู้จำกริยาท่าทางของมนุษย์ โดยใช้กล้องวิดีโอทัศน์ทั้งสองตัวถ่าย และกัมเมยติดตามใบหน้าหรือมือ รวมทั้งซูมเพื่อขยายรายละเอียดของทั้งใบ หน้าและมือ

1.2.2 งานวิจัยที่นำเสนอเกี่ยวกับการเทียบมาตรฐานกล้องวิดีโอทัศน์

จากงานวิจัยที่ผ่านมาส่วนมากแล้วจะให้ความสำคัญกับการเทียบมาตรฐานทางตำแหน่งเนื่องจากสามารถนำไปประยุกต์กับงานทางด้าน การตรวจหาได้อย่างมีประสิทธิภาพ เมื่อมีเทคนิคการเทียบมาตรฐานที่ดี แต่สำหรับงานทางด้าน การติดตามนั้น จำเป็นต้องอาศัยขั้นตอนวิธีในการติดตามอยู่แล้ว ดังนั้นประสิทธิภาพการติดตามจะดีหรือไม่ขึ้นอยู่ขึ้นอยู่กับข้อมูลที่ใช้ในขั้นตอนวิธีของการติดตาม เช่น ข้อมูลสี วิชยานิพนธ์ฉบับนี้จึงให้ความสำคัญกับข้อมูลสีมากกว่าข้อมูลในส่วนอื่นๆ

จากงานวิจัยของ G.Unal และ A.Yezzi [7] นำเสนอปัญหาของการเทียบมาตรฐานพารามิเตอร์ของกล้องวิดีโอทัศน์ ได้แก่ การเทียบมาตรฐานกล้องวิดีโอทัศน์ด้วยความบิดเบือนของเลนส์ จากรูปที่ 1.3 แสดงมุมมองของภาพที่รับมาจากกล้องวิดีโอทัศน์สามตัว ซึ่งพบว่าวัตถุเป็นชนิดเดียวกันแต่เมื่อรับภาพจากมุมมองของกล้องที่ต่างกัน ภาพที่ได้จะมีลักษณะที่ต่างกันด้วย ดังนั้นจึงต้องทำการเทียบมาตรฐานกล้องวิดีโอทัศน์ด้วยความบิดเบือนของเลนส์ เพื่อให้ระบบการทำงานร่วมกันระหว่างกล้องวิดีโอทัศน์หลายตัว ทราบว่าวัตถุนี้เป็นวัตถุชนิดเดียวกัน เนื่องจากงานวิจัยนี้เป็นงานวิจัยเกี่ยวกับกระบวนการตรวจหา ปัญหาการบิดเบือนของเลนส์จึงจำเป็นต้องแก้ไข



รูปที่ 1.3 แสดงภาพของกล้องวิดีโอทัศน์แต่ละตัวที่เกิดจากการบิดเบือนของเลนส์ [7]

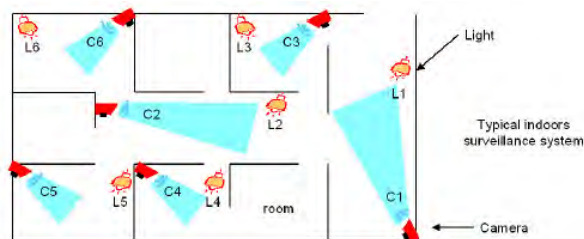
แต่สำหรับงานที่เกี่ยวกับการติดตามบุคคลนั้น ปัญหาของระบบการทำงานร่วมกันระหว่างกล้องวิดีโอทัศน์หลายตัวนั้น คือ การประมวลผลทางสีของภาพจากกล้องแต่ละตัวมี

ความผิดพลาดอื่น ๆ อันเนื่องมาจากสาเหตุบางประการ เช่น คุณสมบัติของตัวจับภาพของกล้องแต่ละตัวต่างกัน สภาพแวดล้อมการติดตั้งกล้องแต่ละตัวต่างกัน สภาพแสงต่างกัน การผลิตจากโรงงานอุตสาหกรรมที่ต่างกัน เป็นต้น ดังนั้น ในการทำงานของระบบการทำงานร่วมกันระหว่างกล้องวิดีโอหลายตัวนั้น จึงต้องมีวิธีการของการเทียบมาตรฐานทางสีของกล้องแต่ละตัว ให้กล้องวิดีโอแต่ละตัวมองเห็นเป็นสีเดียวกันมีข้อมูลเดียวกันหรือใกล้เคียงกัน นำเสนอการเทียบมาตรฐานกล้องวิดีโอด้วยสี ซึ่งหลักการ คือ นำภาพที่ได้จากกล้องวิดีโอแต่ละตัว มาคำนวณหาค่าสีภายในภาพของกล้องแต่ละตัว แล้วเปรียบเทียบกับวัตถุที่ใช้ในการเปรียบเทียบมาตรฐานทางสีดังรูปที่ 1.4 จุดประสงค์เพื่อให้ระบบกล้องวิดีโอหลายตัวนี้ ทำงานเปรียบเทียบได้ว่าเป็นกล้องวิดีโอตัวเดียวกัน มองภาพและประมวลผลทางสีของภาพได้อย่างถูกต้องแม่นยำมากยิ่งขึ้น



รูปที่ 1.4 แสดงภาพวัตถุที่ใช้ในการเปรียบเทียบมาตรฐานทางสี (Color Calibration Object) [7]

จากงานวิจัยของ F. Porikli และคณะ [8,9] นำเสนอวิธีการเทียบมาตรฐานกล้องวิดีโอเพื่อใช้ในระบบการตรวจหาและติดตามบุคคลด้วยกล้องวิดีโอ งานวิจัยนี้ได้ให้ความสำคัญกับการสร้างให้ระบบกล้องวิดีโอทำงานร่วมกันอย่างอัตโนมัติ มีการเปรียบเทียบมาตรฐานกล้องวิดีโออย่างถูกต้องแม่นยำ และสามารถตรวจจับและติดตามบุคคลหรือวัตถุที่สนใจได้อย่างถูกต้อง สิ่งในงานวิจัยนี้ให้ความสำคัญมากที่สุด คือ การมองถึงปัญหาการเทียบมาตรฐานสีของกล้องวิดีโอ จากงานวิจัยที่ผ่านมา จะกำหนดให้แสงสว่างบริเวณที่กล้องแต่ละตัวติดตั้งนั้น มีค่าเท่ากันหรือประมาณให้มีค่าใกล้เคียงกัน ซึ่งเป็นสิ่งที่ทำให้ผลการตรวจหาและติดตามบุคคลที่สนใจมีความผิดพลาด แต่ในความเป็นจริงแล้วปัญหาความสว่างของบริเวณที่ติดตั้งกล้องวิดีโอแต่ละตัวที่ต่างกันนั้น ไม่สามารถควบคุมได้ จากรูปที่ 1.5 แสดงสถานการณ์การติดตั้งระบบกล้องวิดีโอที่ใช้ในงานวิจัยนี้



รูปที่ 1.5 แสดงตำแหน่งการติดตั้งกล้องที่มีข้อจำกัดทางด้านความสว่างในแต่ละบริเวณ [8,9]

ซึ่งงานวิจัยเกี่ยวกับการเทียบมาตรฐานทางสีนั้นยังมีให้เห็นไม่มากนัก ส่วนใหญ่แล้วจะแก้ปัญหาด้วยการชดเชยสีที่เพี้ยนไปอย่างคร่าวๆเท่านั้น ในวิทยานิพนธ์ฉบับนี้จึงนำเสนอการเทียบมาตรฐานทางสีด้วยหลักการของสมการถดถอยเชิงเส้น (Linear Regression) ซึ่งจะกล่าวต่อไปในบทที่ 3

1.2.3 งานวิจัยที่นำเสนอเกี่ยวกับขั้นตอนวิธีการติดตามบุคคล

งานวิจัยที่อาศัยเทคนิคของตัวกรองและความสัมพันธ์ระหว่างข้อมูลนั้น เช่น งานวิจัยของ Y, Boykov และ D.Huttenlocher [13] ได้นำเสนอระบบจรรยาจรอัจฉริยะเพื่อติดตามการเคลื่อนที่ของยานพาหนะโดยอาศัยตัวกรองแบบคาร์มาล และงานวิจัยของ R.Rosales และ S.Sclaroff [14] นำเสนอการประมาณทิศทางการเคลื่อนที่ของวัตถุในสามมิติ จากภาพการเคลื่อนที่ของวัตถุในสองมิติโดยอาศัยเทคนิคของตัวกรองแบบเอ็กเทนเดตคาร์มาล นอกจากนี้ยังมีงานวิจัยของ M.Isard และ A.Blake [15] ซึ่งเป็นงานวิจัยแรกที่น่าเสนอการใช้ตัวอย่างสุ่มที่ได้มาจากฟังก์ชันความน่าจะเป็นของความหนาแน่นของเป้าหมายมาใช้โดยอาศัยเทคนิคของตัวกรองแบบพาทิเคิล

เนื่องจากวิทยานิพนธ์ฉบับนี้ให้ความสำคัญกับเทคนิคการกำหนดตัวแทนเป้าหมายและหาตำแหน่งของเป้าหมาย จึงขอเสนองานวิจัยที่ผ่านมาเกี่ยวกับการกำหนดตัวแทนเป้าหมายและหาตำแหน่งของเป้าหมายดังกล่าวดังนี้

จากงานวิจัยของ Fukunaga และ Hosteler [17] กล่าวว่า การประมาณความหนาแน่นของข้อมูลขึ้นอยู่กับฟังก์ชันความหนาแน่นที่มีความต่อเนื่อง ซึ่งสามารถประมาณความหนาแน่นจุดใดๆ ได้โดยการสำรวจข้อมูลตัวอย่างภายในบริเวณเล็กๆ รอบจุดนั้น (นั่นคือการประมาณฟังก์ชันความหนาแน่นด้วยการใช้ฟังก์ชันหลัก) ปัญหาของการศึกษาและวิจัยส่วนใหญ่ของงานเกี่ยวกับคอมพิวเตอร์วิทัศน์ คือ การใช้ฟังก์ชันความหนาแน่นของข้อมูล จึงจำเป็นต้องหา

วิธีการหรือเทคนิคเพื่อให้ทราบถึงรูปแบบความหนาแน่นของข้อมูล รวมถึงการกำหนดตัวแปร (Parameters) ที่จำเป็นต้องใช้ในวิธีการหรือเทคนิคนั้นๆ ควรกำหนดอยู่ในพื้นฐานของข้อมูลที่ใช้ ดังนั้น เทคนิคการประมาณความหนาแน่นแบบไม่กำหนดตัวแปร (Non-parametric density estimation) จึงเหมาะสมที่จะนำมาใช้แก้ปัญหาดังกล่าวจะเป็นจุดเริ่มต้นให้เกิดขั้นตอนวิธีการย้ายตามค่าเฉลี่ย (Mean-shift algorithm)

Cheng [18] ได้ศึกษาขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเพิ่มเติมและพิสูจน์ให้เห็นถึงความเป็นไปได้ที่จะนำมาประยุกต์ใช้ในงานเกี่ยวกับคอมพิวเตอร์วิทัศน์บางประเภท เช่น การจัดกลุ่มข้อมูล (Clustering) และพิสูจน์ให้เห็นว่าขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเป็นลักษณะที่เรียกว่า Mode-seeking process กระบวนการนี้นิยมใช้กันอย่างกว้างขวางในงานเกี่ยวกับภาพไม่ว่าจะเป็นการแบ่งส่วนภาพ (Image segmentation) การทำให้ภาพเรียบที่ยังคงความไม่ต่อเนื่องของภาพไว้ (Discontinuity-persevering smoothing) รวมถึงการติดตามภาพ (Tracking) เป็นต้น

งานวิจัยที่เกี่ยวกับการติดตามภาพโดยอาศัยขั้นตอนวิธีการย้ายตามค่าเฉลี่ยนั้น ในระยะแรก Gray R. Bradski [19] ได้ทำการศึกษาขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเพื่อติดตามบริเวณของใบหน้า (Face tracking) จากภาพเคลื่อนไหว แต่ปัญหาที่พบคือหากภาพพื้นหลังหรือองค์ประกอบภายในภาพมีสีที่ใกล้เคียงกับสีใบหน้าจะทำให้การติดตามผิดพลาด

และงานวิจัยของ Z.Wen และ Z.Cai [20] พิสูจน์สมการของขั้นตอนวิธีการย้ายตามค่าเฉลี่ย ซึ่งแสดงให้เห็นว่าการใช้ฟังก์ชันหลักเป็นฟังก์ชันแบบเกาส์เซียนนั้นสามารถนำมาประยุกต์ใช้ในงานเกี่ยวกับการติดตามบุคคลได้จริงและมีประสิทธิภาพ

อีกทั้งในงานวิจัยของ K.Bai และ W.Liu [21] ยังให้ความสำคัญกับขั้นตอนวิธีการย้ายตามค่าเฉลี่ยโดยการนำขั้นตอนวิธีการย้ายตามค่าเฉลี่ยนี้นี้มาประยุกต์เข้ากับขั้นตอนวิธีของตัวกรองแบบพาทิเคิล ซึ่งจากงานวิจัยนี้ได้พิสูจน์ให้เห็นว่าเมื่อนำขั้นตอนวิธีย้ายตามค่าเฉลี่ยมารวมกับขั้นตอนวิธีของจรวดตัวกรองแบบพาทิเคิลแล้ว จะทำให้การนำขั้นตอนวิธีของตัวกรองแบบพาทิเคิลที่ถูกปรับปรุงแล้วมาติดตามวัตถุได้อย่างมีประสิทธิภาพเพิ่มมากขึ้นด้วย

ในส่วนของคุณลักษณะสำคัญที่ใช้ในขั้นตอนวิธีการย้ายตามค่าเฉลี่ยนั้น งานวิจัยของ H.Liu และคณะ [22] นำเสนอการเพิ่มคุณลักษณะทางการเคลื่อนที่ของเป้าหมายร่วมกับคุณลักษณะทางสีเพื่อแก้ปัญหาการบดบังการติดตามของเป้าหมายที่สนใจ

จากงานวิจัยที่เกี่ยวข้องทั้งหมดจะเห็นว่าขั้นตอนวิธีการย้ายตามค่าเฉลี่ยสามารถประยุกต์ใช้งานได้หลากหลายรวมทั้งการติดตามภาพได้อย่างมีประสิทธิภาพและรวดเร็ว

วิทยานิพนธ์ฉบับนี้มีความจำเป็นต้องการขั้นตอนวิธีการติดตามบุคคลได้อย่างถูกต้องและรวดเร็ว ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยดังกล่าวจึงถูกนำมาใช้เป็นแนวคิดหลักในวิทยานิพนธ์ฉบับนี้ด้วย

1.3 แนวทางวิทยานิพนธ์

วิทยานิพนธ์นี้นำเสนอระบบกล้องหลายตัวที่สามารถนำไปประยุกต์ใช้งานสำหรับระบบสอดส่องรักษาความปลอดภัยภายในตัวอาคาร โดยระบบจะประกอบด้วยกล้องวีดิทัศน์ PTZ 2 ตัว ซึ่งจะมีการประมวลผลร่วมกันเพื่อประโยชน์ในส่วนการติดตามบุคคลที่สนใจอย่างต่อเนื่อง โดยใช้ขั้นตอนวิธีการย้ายตามค่าเฉลี่ย วิทยานิพนธ์นี้จะใช้ประโยชน์จากระบบกล้องหลายตัวและความสามารถของกล้องวีดิทัศน์ PTZ คือ ความสามารถในการส่ายและก้มเงยเปลี่ยนมุมมองการรับภาพและการซูมของกล้องวีดิทัศน์ PTZ เพื่อให้สามารถติดตามบุคคลได้อย่างมีประสิทธิภาพ เนื่องจากประโยชน์ของระบบกล้องหลายตัว ที่จะมีมุมมองการรับภาพที่กว้างกว่า ทำให้เพิ่มประสิทธิภาพให้การติดตามได้มากขึ้น ติดตามบุคคลได้ไกลขึ้น ทำให้ได้รับประโยชน์จากการใช้งานจริงอย่างเต็มประสิทธิภาพมากขึ้น

วิทยานิพนธ์นี้จึงจะเสนอขั้นตอนวิธีในการทำงานร่วมกันระหว่างกล้องวีดิทัศน์ PTZ ทั้งสองตัว เพื่อการติดตามบุคคลอย่างต่อเนื่อง โดยให้ความสำคัญใน 2 สิ่ง คือ

1. ขั้นตอนการเทียบมาตรฐานทางสีระหว่างกล้องวีดิทัศน์ ดังที่กล่าวมาแล้วว่า เมื่อสร้างระบบกล้องวีดิทัศน์ที่ถูกติดตั้งต่างบริเวณกันนั้น เมื่อแสงสว่างในสองบริเวณหรือคุณลักษณะทางกายภาพของกล้องวีดิทัศน์แต่ละตัวในระบบต่างกัน เมื่อต้องการใช้ข้อมูลสีจากกล้องวีดิทัศน์แต่ละตัวร่วมกัน จึงจำเป็นต้องหาวิธีเพื่อทำการเทียบค่าสีที่รับได้ จากกล้องวีดิทัศน์แต่ละตัวให้สื่อความหมายเป็นค่าเดียวกันมากที่สุด โดยเฉพาะอย่างยิ่งเมื่อนำมาใช้งานเกี่ยวกับการติดตามบุคคล อีกทั้งยังมีการส่งข้อมูลระหว่างกล้องวีดิทัศน์ด้วย ดังเช่นวิทยานิพนธ์ฉบับนี้ จึงจำเป็นต้องทำให้กล้องวีดิทัศน์แต่ละตัวรู้ค่าสีเดียวกันได้ได้ตรงกัน

2. ขั้นตอนวิธีการติดตามบุคคล การทำงานร่วมกันสำหรับการติดตามบุคคลอย่างต่อเนื่องระหว่างกล้องวีดิทัศน์ 2 ตัว โดยใช้คุณลักษณะสำคัญของสีและรูปร่างซึ่งแทนด้วยขอบของบุคคลเป้าหมาย สำหรับระบุเป้าหมายบุคคลจะเลือกใช้สีส่วนที่สามารถเห็นได้ชัดเจน เช่น สีเสื้อ หรือสีกางเกง เป็นต้น ซึ่งการทำงานเพื่อติดตามบุคคลและควบคุมกล้องวีดิทัศน์ PTZ นั้นทำได้โดยเมื่อบุคคลเป้าหมายเข้ามายังขอบเขตมุมมองของกล้องวีดิทัศน์ตัวที่หนึ่งแล้ว กล้องวีดิทัศน์ตัวที่หนึ่งจะทำการติดตามบุคคลเป้าหมายพร้อมทั้งส่ายก้มเงยตามบุคคลเป้าหมายไปเรื่อยๆ จนเมื่อบุคคลเป้าหมายจะเคลื่อนที่พ้นขอบเขตมุมมองของกล้องวีดิทัศน์ตัวแรก กล้องวีดิทัศน์ตัวที่สอง จะทำ

การตรวจหาคุณลักษณะสำคัญของบุคคลเป้าหมายที่ส่งต่อมาจากกล้องวิดีโอตัวที่หนึ่ง และระบุได้ว่าเป็นบุคคลเป้าหมายเดิม จึงจะทำการสาย กัมเมย เพื่อติดตามเป้าหมายบุคคลเดิมต่อไป

สำหรับขั้นตอนวิธีสำหรับการติดตามบุคคลอย่างต่อเนื่อง ที่ต้องประยุกต์ใช้กับกล้องวิดีโอ PTZ ซึ่งเป็นส่วนสำคัญที่ต้องมีการพิจารณาในกรณีที่ต้องทำงานกับกล้องวิดีโอ PTZ วิทยานิพนธ์นี้ได้เสนอขั้นตอนวิธีการย้ายตามค่าเฉลี่ย โดยเลือกใช้คุณลักษณะสำคัญของประการคือ คุณลักษณะสำคัญทางสีผิวและคุณลักษณะสำคัญทางสีขอบของบุคคลเป้าหมาย

1.4 เป้าหมายและขอบเขตของงานวิจัย

- 1.4.1 ออกแบบและสร้างรูปแบบจำลองการติดตามบุคคลโดยใช้กล้องวิดีโอหลายตัวทำงานร่วมกันโดยอาศัยสีและการเคลื่อนไหวที่
- 1.4.2 นำวิธีการเทียบมาตรฐานที่นำเสนอมาประยุกต์ใช้งานร่วมกับการติดตามบุคคลอย่างต่อเนื่องได้อย่างมีประสิทธิภาพ

1.5 ขั้นตอนและวิธีการดำเนินงานวิจัย

- 1.5.1 ศึกษารูปแบบและลักษณะการนำไปใช้งานของระบบกล้องหลายตัวและหาแนวทางในการประยุกต์ใช้กับวิทยานิพนธ์
- 1.5.2 ศึกษาระเบียบวิธีกระบวนการเทียบมาตรฐานทางสีของกล้องวิดีโอเพื่อประยุกต์ใช้กับวิทยานิพนธ์
- 1.5.3 เลือกคุณลักษณะและระเบียบวิธีที่เหมาะสมมาใช้ในวิธีการเทียบมาตรฐานทางสี
- 1.5.4 ศึกษาระเบียบวิธีการติดตามบุคคลและการควบคุมกล้องวิดีโอ
- 1.5.5 เลือกคุณลักษณะและระเบียบวิธีที่เหมาะสมมาใช้ในการติดตามบุคคลและการควบคุมกล้องวิดีโอ
- 1.5.6 เปรียบเทียบและวิเคราะห์ผลการจำลองระบบ
- 1.5.7 จัดทำรายงานฉบับสมบูรณ์

1.6 ประโยชน์ที่คาดว่าจะได้รับ

- 1.6.1 ทราบถึงความรู้พื้นฐานและการทำงานร่วมกันของกล้องวิดีโอในระบบกล้องวิดีโอหลายตัวเพื่อติดตามบุคคลอย่างต่อเนื่องได้อย่างถูกต้องและมีประสิทธิภาพ
- 1.6.2 สามารถนำไปประยุกต์และพัฒนาระบบการติดตามบุคคลด้วยสีในระบบกล้องวิดีโอหลายตัวสำหรับระบบเฝ้าระวังให้มีประสิทธิภาพดีขึ้น

บทที่ 2

ความรู้พื้นฐานและทฤษฎีที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงความรู้พื้นฐานและทฤษฎีที่จำเป็นสำหรับวิทยานิพนธ์ โดยจะแบ่งเนื้อหาออกเป็น 5 หัวข้อหลัก คือ ปริภูมิสีและเวกเตอร์ค่าสีรวมถึงขั้นตอนการแปลงปริภูมิสีที่จำเป็นในงานวิจัย แบบจำลองกล้องวิดีโอทัศน์และแบบจำลองการเคลื่อนที่ของกล้องวิดีโอทัศน์ การเทียบมาตรฐานระหว่างกล้องวิดีโอทัศน์ ขั้นตอนวิธีการติดตามบุคคล และขั้นตอนการควบคุมกล้องวิดีโอทัศน์ประเภทสาย กัมเมยและซูม

2.1 ปริภูมิของสีและเวกเตอร์ค่าสี

ค่าสี (Color) ถือเป็นลักษณะหลักสำคัญในระบบการประมวลผลทางดิจิทัลของภาพและวิดีโอ งานวิจัยส่วนมากจะอาศัยข้อมูลของค่าสีเป็นสำคัญเนื่องจากง่ายต่อการนำมาใช้ประมวลผล ซึ่งค่าสีที่ได้เกิดจากการรับภาพวิดีโอทัศน์จากกล้องวิดีโอทัศน์นั้น อาจอยู่ในรูปแบบสัญญาณดิจิทัล หรือแอนะล็อก โดยสัญญาณแบบแอนะล็อกจะถูกนำไปผ่านการสุ่ม (Sampling) และการแจกหน่วย (Quantization) เป็นสัญญาณดิจิทัลเพื่อการประมวลผล ซึ่งลักษณะของสัญญาณวิดีโอทัศน์ที่รับมาจากกล้องวิดีโอทัศน์จะขึ้นอยู่กับสายสัญญาณที่เชื่อมต่อกับตัวประมวลผล จากนั้นข้อมูลสีจะถูกเก็บในรูปแบบค่าสีโดยเป็นผลรวมขององค์ประกอบฐานสี (Color basis) เข้าเป็นสีต่างๆตามภาพที่รับได้ในหนึ่งจุดภาพ (Pixel) ดังนั้นจึงสามารถแทนค่าสีด้วยเวกเตอร์ค่าสีในปริภูมิสี (Color space) แบบต่างๆได้หลายรูปแบบ

เนื่องจากระบบการมองเห็นของมนุษย์ (Human vision system) นั้นมนุษย์มีความสามารถในการรับรู้ทางความสว่างและสีได้ดีไม่เท่ากัน โดยสายตามนุษย์จะมีความไวต่อความส่องสว่างมากกว่าสี จึงเกิดปริภูมิสีที่มีการแยกแยะระหว่างความส่องสว่าง (Luminance) และสี (Chrominance) ขึ้น ปริภูมิสีที่สำคัญต่อการประมวลผลทางดิจิทัลของภาพและวิดีโอทัศน์มีมากมายหลากหลายประเภท ซึ่งโดยทั่วไปมักนิยมใช้ปริภูมิสีแบบ RGB, YCbCr และ HSV และถูกใช้งานในวิทยานิพนธ์ฉบับนี้ด้วย

2.1.1 ปริภูมิสี RGB (RGB color space)

ปริภูมิสีแบบ RGB เป็นปริภูมิสีที่ถูกนำมาใช้มากที่สุดในงานการประมวลผลภาพดิจิทัล เนื่องจากเป็นองค์ประกอบสีหลักในอุปกรณ์แสดงผล (Display) ทั่วๆ ไปที่อาศัยการแผ่รังสีในย่านคลื่นแสงที่มองเห็นอันประกอบด้วยฐานสี สีแดง สีเขียว และสีน้ำเงิน ดังนั้นจึงเป็นการง่าย

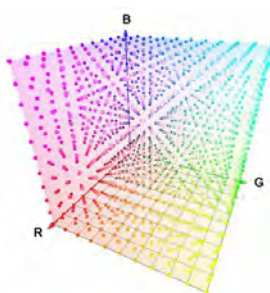
ในการจัดการและออกแบบระบบเพื่อทำงานกับปริภูมิสี RGB ดังสังเกตได้จากการแปลงปริภูมิสีต่างๆ มักแปลงโดยเริ่มต้นจากปริภูมิสี RGB

เมื่อพิจารณาเวกเตอร์ค่าสีในปริภูมิ RGB จะได้เวกเตอร์ซึ่งเป็นผลรวมเชิงเส้นขององค์ประกอบสีแดง เขียว และน้ำเงิน เวกเตอร์ค่าสี ณ จุดภาพหลักที่ i แถวที่ j ของเฟรมที่ n แสดงในสมการที่ (2.1)

$$\mathbf{X}_{i,j}[n] = (X_{i,j}^R[n], X_{i,j}^G[n], X_{i,j}^B[n]) \quad (2.1)$$

โดย $X_{i,j}^R[n]$ เป็นค่าความเข้มขององค์ประกอบสีแดง และ $0 \leq X_{i,j}^R[n] \leq D^R$
 $X_{i,j}^G[n]$ เป็นค่าความเข้มขององค์ประกอบสีเขียว และ $0 \leq X_{i,j}^G[n] \leq D^G$
 $X_{i,j}^B[n]$ เป็นค่าความเข้มขององค์ประกอบสีน้ำเงิน และ $0 \leq X_{i,j}^B[n] \leq D^B$

โดย D^R D^G และ D^B เป็นค่าสีสูงสุดในองค์ประกอบสีแดง เขียว น้ำเงิน ตามลำดับ ขึ้นอยู่กับความลึกของสี (Color Depth) จากกระบวนการการแจกหน่วย โดยที่นิยมใช้ในปัจจุบันจะอยู่ในความลึกสี 24 บิตต่อจุดภาพ อันประกอบด้วยองค์ประกอบสีละ 8 บิตต่อจุดภาพ ซึ่งมีค่า D^R D^G และ D^B เป็น 255 ทั้ง 3 ค่า ปริภูมิ RGB แสดงได้ดังรูปที่ 2.1



รูปที่ 2.1 ปริภูมิสีแบบ RGB (สีแดง เขียว และน้ำเงิน)

อย่างไรก็ตามปริภูมิสีแบบ RGB ไม่ยืดหยุ่นในการใช้งานจริง เนื่องด้วยเหตุผลที่อุปกรณ์รับรู้สี (Color Sensor) มีความไว (Sensitivity) ต่อแสงสีย่านสีแดง เขียว และน้ำเงินแตกต่างกัน แต่รูปแบบการเก็บข้อมูลสีในปริภูมิ RGB จะต้องมีความละเอียดของระดับสีเท่ากันทั้ง 3 แถบสัญญาณสี ดังนั้นจึงจำเป็นที่จะต้องทำการปรับสมดุลค่าสี (Color Equalization) อันทำให้สูญเสียข้อมูลในแถบสัญญาณสีที่มีความละเอียดสูง นอกจากนั้นในการประมวลผลทางดิจิทัล

การอ่านค่าสีในรูปแบบ RGB จะทำการอ่านค่าสีทั้ง 3 แฉบสีไม่ว่าจะใช้ข้อมูลในการประมวลผลจริงก็แถบสี เพราะฉะนั้นปริภูมิสีแบบ RGB จึงไม่ถือเป็นปริภูมิที่มีประสิทธิภาพมากที่สุด

ปัญหาของปริภูมิสี RGB คือ การรวมกันของความสว่างและสี จึงเกิดเป็นปริภูมิสีนอร์มอลไลซ์ RGB (rgb color Space) ซึ่งมีวัตถุประสงค์เพื่อตัดความสนใจขององค์ประกอบความสว่างทิ้งไป โดยหากค่าจุดภาพในปริภูมิสี RGB นั้นมีสีเดียวกันแต่ความสว่างต่างกัน จะถูกกำหนดเป็นค่าเดียวกันในปริภูมิสีนอร์มอลไลซ์ RGB และแทนเป็นสัญลักษณ์ "rgb" แต่ละค่าสามารถหาได้จากสมการที่ 2.2- 2.4

$$r = \frac{R}{(R+G+B)} \quad (2.2)$$

$$g = \frac{G}{(R+G+B)} \quad (2.3)$$

$$b = \frac{B}{(R+G+B)} \quad (2.4)$$

ถ้าค่าในปริภูมิสี RGB (255,0,0) ทำการแปลงไปสู่ rgb จะได้ (1,0,0) และอีกค่าสีหนึ่งในปริภูมิสี RGB (100,0,0) ทำการแปลงไปสู่ rgb จะได้(1,0,0) ซึ่งทั้งสองค่านี้ในปริภูมิสี RGB จะมีเฉพาะองค์ประกอบของสีแดง ที่มีค่าความเข้มของแสงแตกต่างกัน ส่วนในปริภูมิสี rgb องค์ประกอบทางแสงจะถูกทิ้งไปทำให้ 2 ค่านี้ถือเป็นค่าเดียวกัน

ข้อด้อยของวิธีนี้ คือ การรบกวนกันในปริภูมิสีนอร์มอลไลซ์ RGB เกิดขึ้นได้ง่ายกว่าปริภูมิสีอื่นๆ โดยเฉพาะบริเวณตำแหน่งจุดภาพที่มีค่าใกล้เคียงกับสีดำ จะมีความแปรปรวนได้ง่าย และค่าของจุดภาพจำพวกนี้มีโอกาสที่จะเป็นค่าใดก็ได้ตั้งแต่ค่าต่ำสุดจนถึงสูงสุดนในปริภูมิสีนอร์มอลไลซ์ RGB ทำให้ค่าสีที่นำไปใช้อาจผิดพลาดได้

2.1.2 ปริภูมิสีแบบ YUV และ YCbCr (YUV and YCbCr Color Space)

ปริภูมิสีแบบ YUV ถูกพัฒนาขึ้นเพื่อนำมาใช้ในระบบโทรทัศน์แบบ PAL (Phase Alternation Line) NTSC (National Television System Committee) และ SECAM (Sequentiel Couleur Avec Mémoire or Sequential Color with Memory) ซึ่งมีลักษณะเป็นสัญญาณภาพรวม (Composite Video) แต่เดิมสัญญาณวีดิทัศน์จะเป็นสัญญาณภาพขาว-ดำ เมื่อผู้ให้บริการเปลี่ยนมาใช้ระบบส่งสัญญาณภาพสี ปริภูมิสีแบบ YUV ได้เข้ามามีบทบาทในแง่ของการ

แยกองค์ประกอบความสว่าง (Y) ซึ่งใช้ในสัญญาณภาพขาว-ดำออกจากองค์ประกอบทางสี (U และ V) ทำให้สามารถส่งสัญญาณภาพรวมสี โดยที่ไม่ต้องทำการปรับเปลี่ยนระบบรับภาพขาวดำที่มีอยู่

สมการการแปลงพื้นฐานระหว่างปริภูมิ RGB และ YUV อาศัยการแปลงเพื่อหาองค์ประกอบความสว่าง (Luminance) Y ณ จุดภาพที่พิจารณาก่อน จากนั้นจึงหาองค์ประกอบเชิงสี U ซึ่งแปรผันโดยตรงกับผลต่างต่างระหว่างองค์ประกอบสีน้ำเงิน (B) ในปริภูมิ RGB กับองค์ประกอบความสว่าง (Y) และองค์ประกอบเชิงสี V ซึ่งแปรผันโดยตรงกับผลต่างต่างระหว่างองค์ประกอบสีแดง (R) ในปริภูมิ RGB กับองค์ประกอบความสว่าง (Y) ดังสมการการแปลงเวกเตอร์ค่าสีในปริภูมิ RGB $\mathbf{X}_{i,j}[n]$ และเวกเตอร์ค่าสีในปริภูมิ YUV $\mathbf{W}_{i,j}[n]$ ดังนี้

$$\mathbf{W}_{i,j}[n] = \begin{bmatrix} W_{i,j}^Y[n] \\ W_{i,j}^U[n] \\ W_{i,j}^V[n] \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} X_{i,j}^R \\ X_{i,j}^G \\ X_{i,j}^B \end{bmatrix} \quad (2.5)$$

และ

$$\mathbf{X}_{i,j}[n] = \begin{bmatrix} X_{i,j}^R[n] \\ X_{i,j}^G[n] \\ X_{i,j}^B[n] \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.140 \\ 1 & -0.395 & -0.581 \\ 1 & 0.032 & 0 \end{bmatrix} \begin{bmatrix} W_{i,j}^Y \\ W_{i,j}^U \\ W_{i,j}^V \end{bmatrix} \quad (2.6)$$

โดยที่ $W_{i,j}^Y[n]$ เป็นค่าองค์ประกอบความสว่าง Y และ $0 \leq W_{i,j}^Y[n] \leq D^Y$

$W_{i,j}^U[n]$ เป็นค่าองค์ประกอบเชิงสี U และ $-D^U \leq W_{i,j}^U[n] \leq D^U$

$W_{i,j}^V[n]$ เป็นค่าองค์ประกอบเชิงสี V และ $-D^V \leq W_{i,j}^V[n] \leq D^V$

ค่าสีในปริภูมิ RGB 24 บิตต่อจุดภาพจะมีค่าตั้งแต่ 0 ถึง 255 ในแต่ละองค์ประกอบสี ทำให้ค่าองค์ประกอบความสว่าง Y ในปริภูมิ YUV มีค่าระหว่าง 0 ถึง 255 (D^Y) องค์ประกอบเชิงสี U มีค่าในช่วง -112 ($-D^U$) ถึง 112 (D^U) และองค์ประกอบเชิงสี V มีค่าในช่วง -157 ($-D^V$) ถึง 157 (D^V) จะเห็นได้ว่าค่าที่ได้ไม่เหมาะสมในการใช้งาน เนื่องจากองค์ประกอบสีอาจมีค่าเป็นลบรวมทั้งขนาดของหน่วยความจำที่ใช้เก็บไม่เท่ากัน จึงไม่เหมาะสมนำมาใช้ในระบบภาพดิจิทัล ดังนั้นจึงจำเป็นต้องมีการพัฒนาปริภูมิสี YUV ใช้งานได้ดีขึ้นเพื่อระบบสัญญาณวิดีโอทัศน์ดิจิทัลอันเป็นจุดกำเนิดของปริภูมิสี YCbCr

ปริภูมิสี YCbCr เป็นปริภูมิสีดิจิทัล และมีการกำหนดเป็นมาตรฐานโดย BT.601 ซึ่งจะแบ่งข้อมูลของสีออกในลักษณะที่สอดคล้องกับการรับรู้ของมนุษย์เช่นเดียวกับปริภูมิสี HSI

โดยมีค่า Y เป็นองค์ประกอบความสว่างมีค่าตั้งแต่ 16-235 ส่วนค่า Cb และ Cr คือองค์ประกอบของสี ซึ่งเป็นค่าความแตกต่างของ สีน้ำเงินกับค่าความสว่าง ($B-Y$) และ สีแดงกับค่าความสว่าง ($R-Y$) ตามลำดับ โดยที่ Cb และ Cr มีค่าตั้งแต่ 16-240 โดยวิทยานิพนธ์นี้พิจารณาใช้การแปลงเวกเตอร์ค่าสีระหว่างปริภูมิ RGB และเวกเตอร์ค่าสีในปริภูมิ YCbCr ตามมาตรฐาน ITU-R BT.709 โดยมีสมการการแปลงดังนี้

$$\mathbf{Q}_{i,j}[n] = \begin{bmatrix} Q_{i,j}^Y[n] \\ Q_{i,j}^{Cb}[n] \\ Q_{i,j}^{Cr}[n] \end{bmatrix} = \begin{bmatrix} 0.183 & 0.614 & 0.062 \\ -0.101 & -0.338 & 0.439 \\ 0.439 & -0.399 & -0.040 \end{bmatrix} \begin{bmatrix} X_{i,j}^R \\ X_{i,j}^G \\ X_{i,j}^B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (2.7)$$

และ

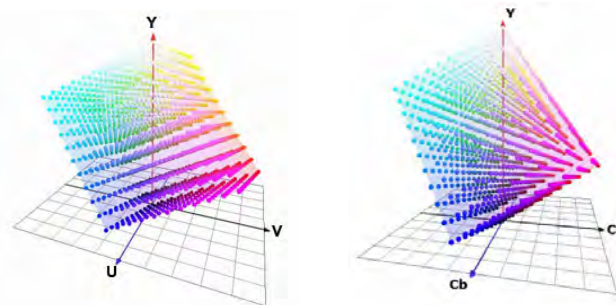
$$\mathbf{X}_{i,j}[n] = \begin{bmatrix} X_{i,j}^R[n] \\ X_{i,j}^G[n] \\ X_{i,j}^B[n] \end{bmatrix} = \begin{bmatrix} 1.164 & 0 & 1.793 \\ 1.164 & -0.534 & -0.213 \\ 1.164 & 2.115 & 0 \end{bmatrix} \left(\begin{bmatrix} Q_{i,j}^Y \\ Q_{i,j}^{Cb} \\ Q_{i,j}^{Cr} \end{bmatrix} - \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \right) \quad (2.8)$$

โดย $Q_{i,j}^Y[n]$ เป็นค่าองค์ประกอบความสว่าง และ $16 \leq Q_{i,j}^Y[n] \leq 235$

$Q_{i,j}^{Cb}[n]$ เป็นค่าองค์ประกอบ Cb และ $16 \leq Q_{i,j}^{Cb}[n] \leq 240$

$Q_{i,j}^{Cr}[n]$ เป็นค่าองค์ประกอบ Cr และ $16 \leq Q_{i,j}^{Cr}[n] \leq 240$

สาเหตุในการเลือกใช้การแปลงตามมาตรฐาน ITU-R BT.709 เนื่องจากเมื่อทำการแปลงองค์ประกอบกลับมาอยู่ในปริภูมิ RGB องค์ประกอบค่าสี แดง เขียว และน้ำเงินจะมีค่าตั้งแต่ 0 ถึง 255 ต่างจากการแปลงตามมาตรฐาน ITU-R BT.601 เมื่อแปลงกลับมาอยู่ในปริภูมิ RGB แต่ละองค์ประกอบจะมีค่าตั้งแต่ 16 ถึง 235 ซึ่งทำให้เสียความแตกต่างของข้อมูลในช่วงค่าสี 0-15 และ 236-255 จึงเหมาะสมน้อยกว่าที่จะใช้ในระบบคอมพิวเตอร์ จากข้างต้นสามารถแสดงค่าสีเพื่อเปรียบเทียบระหว่างปริภูมิ YUV ดังในรูปที่ 2.2 (ก) และ ปริภูมิ YCbCr ในรูปที่ 2.2(ข)



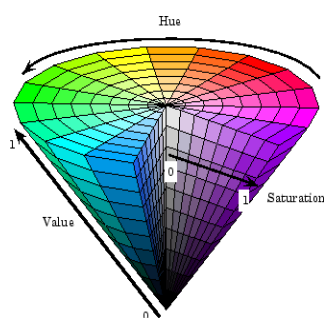
(ก) ปริภูมิสีแบบ YUV (ข) ปริภูมิสีแบบ YCbCr

รูปที่ 2.2 ปริภูมิสีแบบ YUV และปริภูมิสีแบบ YCbCr

ข้อดีในการใช้ปริภูมิสีแบบ YCbCr คือ องค์ประกอบในปริภูมิ มีการแยกองค์ประกอบเชิงความสว่างออกจากองค์ประกอบเชิงสีออกอย่างชัดเจน เป็นการถ่ายที่จะตรวจพบความเปลี่ยนแปลงของแต่ละองค์ประกอบอย่างเป็นอิสระต่อกัน

2.1.3 ปริภูมิสีแบบ HSV (HSV Color Space)

ค่าในปริภูมิสี HSV แยกเป็นค่า H(Hue) และ S(Saturate) คือ องค์ประกอบของสี ส่วนค่า V(Value) คือ ความเข้มของแสง เป็นปริภูมิสีที่สร้างขึ้นตามพื้นฐานการมองเห็นสีด้วยสายตาของมนุษย์ โดยตัว H เป็นสีที่สะท้อนมาจากสีของวัตถุ ซึ่งแตกต่างกันตามความยาวของคลื่นแสงที่มากกระทบวัตถุและสะท้อนกลับมาจากตาเรา โดยวัดเป็นมุม คือ 0 – 360 องศา ซึ่งสีแดง สีเหลืองและสีเขียวจะมีค่าต่างกันสีละ 60 องศา ตัว S คือค่าความเข้มของเนื้อสี หรือค่าความบริสุทธิ์ของสี มีค่าตั้งแต่ 0 – 100 โดยสีจะมีความเข้มมากขึ้นเรื่อยๆ เมื่อค่า Saturation มีค่าเพิ่มขึ้น หรืออีกนัยหนึ่งคือเป็นค่าความอิ่มตัวสี นั่นคือ สัดส่วนที่มีอยู่ในสีเทา วัดเป็นค่าเปอร์เซ็นต์คือจาก 0 (สีเทา) ถึง 100 (สีมีความอิ่มตัวเต็มที่) และ ตัว V หรือ Brightness คือ ความสว่างของสี โดยภาพจะสว่างมากขึ้นเรื่อยๆ เมื่อ Brightness มีค่าเพิ่มขึ้นเรื่อยๆ หรืออีกนัยหนึ่งคือค่าความเข้มของแสงที่วัดเป็นค่าเปอร์เซ็นต์ คือจาก 0 (ความมืด) ถึง 100 (ความสว่าง) แสดงรูปปริภูมิสี HSV ดังรูปที่ 2.3 และแสดงการแปลงข้อมูลสีจากปริภูมิ RGB ไปเป็นปริภูมิ HSV ได้ตามขั้นตอนดังนี้



รูปที่ 2.3 ปริภูมิสี HSV

1) นำข้อมูลสี R, G และ B จากปริภูมิสี RGB แล้วเปรียบเทียบเพื่อหาค่าสูงสุด $\text{Max}(R,G,B)$ ต่ำสุด $\text{Min}(R,G,B)$ ระหว่างค่าข้อมูลสีทั้งสาม และหาค่าความต่างของค่าสูงสุดต่ำสุด $\text{Delta} = \text{Max}(R,G,B) - \text{Min}(R,G,B)$

2) หาค่า H, S และ V ได้จากสมการ 2.9 -2.11 ตามลำดับ

$$H = \begin{cases} 0, \text{Max}(R, G, B) = \text{Min}(R, G, B) \\ (60^\circ \times \frac{G-B}{\text{Delta}} + 360^\circ) \bmod 360^\circ, \text{Max} = R \\ 60^\circ \times \frac{B-R}{\text{Delta}} + 120^\circ, \text{Max} = G \\ 60^\circ \times \frac{R-G}{\text{Delta}} + 240^\circ, \text{Max} = B \end{cases} \quad (2.9)$$

$$S = \frac{\text{Max}(R, G, B) - \text{Min}(R, G, B)}{\text{Max}(R, G, B)} \times 100 \quad (2.10)$$

$$V = \text{Max}(R, G, B) \times 100 \quad (2.11)$$

จากลักษณะโมเดลของระบบ Hue พบว่าจะมีค่าอย่างน้อยหนึ่งค่าที่จะเท่ากับ 0 แต่ถ้ามีสองค่าเท่ากับ 0 แล้ว hue จะเป็นมุมของค่าสีที่มีค่าเป็นไปตามสีที่สามและถ้าทั้งสามสีมีค่าเท่ากับ 0 แล้วจะทำให้ไม่มีค่าของ Hue หรือสีที่ได้จะมีค่าเท่ากับสีขาวนั่นเอง ตัวอย่างเช่น จอภาพขาว-ดำ ถ้าเกิดมีสีใดสีหนึ่งมีค่าเท่ากับ 0 จะทำให้ค่าสีที่ได้เป็นไปตามสีที่เหลือ การให้นำหนักในการพิจารณาเมื่อสีแดงมีค่าเท่ากับ 0

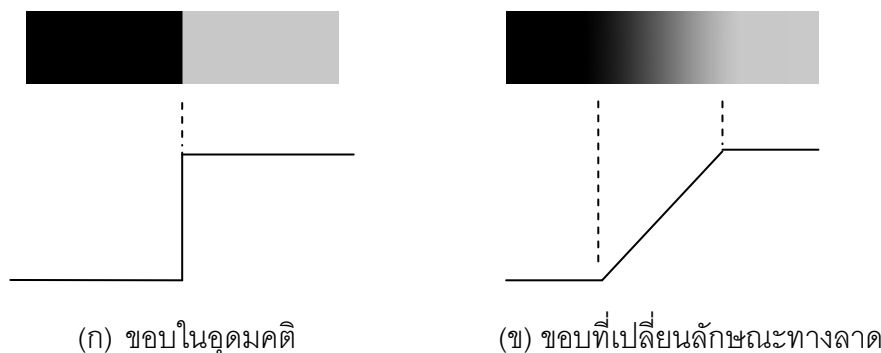
2.2 การหาขอบภาพ (Edge Detection)

การหาขอบภาพ เป็นกระบวนการหาเส้นรอบวัตถุที่อยู่ในภาพ เมื่อทราบเส้นรอบของวัตถุแล้ว เราจะสามารถคำนวณหาพื้นที่ (ขนาด) รู้จำนวนของวัตถุ หรือเป็นข้อมูลเพื่อเป็นตัวแทนวัตถุนั้นๆ ได้ อย่างไรก็ตาม การหาขอบภาพที่ถูกต้องสมบูรณ์ไม่ใช่เป็นเรื่องที่ง่าย โดยเฉพาะอย่างยิ่ง การหาขอบของภาพที่มีคุณภาพต่ำ มีความแตกต่างระหว่างพื้นหน้ากับพื้นหลังน้อย หรือมีความสว่างไม่สม่ำเสมอทั่วภาพขอบภาพเกิดจากความแตกต่างของความเข้มแสงจากจุดหนึ่งไปยังอีกจุดหนึ่ง หากความแตกต่างนี้มีค่ามาก ขอบภาพก็จะเห็นได้ชัด ถ้าความแตกต่างมีค่าน้อย ขอบภาพก็จะไม่ชัดเจนในหัวข้อนี้จะกล่าวถึงเทคนิคเบื้องต้นในการหาขอบภาพ

การตรวจหาขอบ ถูกนำมาใช้ในการวิเคราะห์รูปภาพเพื่อกำหนดขอบเขตของบริเวณที่ต้องการ ซึ่งกระบวนการนี้จะกระทำโดยตรวจสอบว่า จุดภาพจุดใดที่มีการเปลี่ยนแปลงขององค์ประกอบทางความสว่างอย่างทันทีทันใดเมื่อเปรียบเทียบกับบริเวณใกล้เคียงจะกำหนดจุดภาพนั้นเป็นขอบ

วิธีการที่ใช้หาการเปลี่ยนแปลงของฟังก์ชันต่อเนื่อง คือ การหาอนุพันธ์ แต่ในฟังก์ชันของรูปภาพจะขึ้นกับ 2 ตัวแปรในระบบพิกัดระนาบรูปภาพ ดังนั้นวิธีการที่นำมาใช้ คือ การหาอนุพันธ์ย่อย โดยการเปลี่ยนแปลงของฟังก์ชันรูปภาพสามารถอธิบายด้วยค่าเกรเดียนต์ในทิศทางสูงสุด

และเนื่องจาก ขอบ คือ คุณสมบัติที่เกี่ยวกับจุดภาพใดจุดภาพหนึ่งซึ่งนำไปคำนวณร่วมกับพฤติกรรมของจุดภาพรอบข้าง และประกอบไปด้วยตัวแปรชนิดเวกเตอร์ 2 ตัวแปร คือ ขนาดและทิศทาง ขนาดของขอบสามารถหาได้จากขนาดของการเกรเดียนต์ และทิศทางของขอบถูกหมุนในส่วนทิศเกรเดียนต์ ซึ่งมีค่าสูงสุด จากจุดภาพสีดำ $f(i,j) = 0$ ไปจุดภาพสีขาว $f(i,j) = 255$



รูปที่ 2.4 แสดงแบบจำลองของขอบ

2.2.1 ตัวดำเนินการสำหรับหาขอบภาพ

อนุพันธ์อันดับหนึ่งของรูปภาพดิจิทัล จะหาได้จากการประมาณของเกรเดียนต์ใน 2 มิติ ค่าเกรเดียนต์ของรูปภาพ $f(x,y)$ ที่ตำแหน่ง (x,y) ถูกกำหนดด้วยเวกเตอร์ดังนี้

$$\nabla \mathbf{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \tag{2.12}$$

ตัวอย่างเช่น ถ้าต้องการหาขอบภาพในแนวนอนอย่างง่าย วิธีการก็คือหาผลต่างระหว่างจุดหนึ่งกับจุดที่อยู่ข้างล่าง (หรือข้างบน) ของจุดนั้น การหาขอบภาพในลักษณะนี้เรียกว่า การหาขอบภาพในแนวนอน ดังสมการที่ 2.13

$$Y_{diff}(x, y) = f(x, y) - f(x, y+1) \tag{2.13}$$

โดยที่ Y_{diff} คือค่าความแตกต่างในแนวแกนตั้ง
 $I(x,y)$ คือค่าความเข้มแสงของจุดภาพที่ตำแหน่ง (x,y)

การหาขอบภาพในแนวตั้งก็สามารถหาได้เช่นเดียวกันคือสมการที่ 2.14

$$X_{diff}(x, y) = f(x, y) - f(x-1, y) \quad (2.14)$$

โดยที่ X_{diff} คือค่าความแตกต่างในแนวนอน
 $I(x,y)$ คือค่าความเข้มแสงของจุดภาพที่ตำแหน่ง (x,y)

จากสมการที่ (2.13) และ (2.14) มีค่าเทียบเท่ากับการคอนโวลูชันภาพด้วยตัวดำเนินการ

1
-1

 และ

1	-1
---	----

 ตามลำดับ

บางครั้งเราต้องการรวมผลต่างของค่าความแตกต่างในแนวแกนนอน และแกนตั้งเข้าด้วยกัน เพื่อที่จะได้มีตัววัดความแรงของขอบภาพ (Gradient magnitude) เพียงตัวเดียว เนื่องจากค่าความแตกต่างอาจมีค่าเป็นบวกหรือลบ ดังนั้น การบวกค่าความแตกต่างของทั้งสองแกนอาจทำให้ขอบภาพเกิดการหักล้างกันเอง ในทางปฏิบัติ เราจะต้องนำ ค่าสัมบูรณ์ (Absolute value) หรือค่ากำลังสอง (Squared value) ของค่าความแตกต่างของทั้งสองแกนมาบวกกันแทน นอกจากหาความแรงของขอบภาพแล้ว การหาทิศทางของขอบภาพ (Gradient direction) ก็มีประโยชน์เช่นกัน การหาทิศทางของขอบภาพสามสามารถทำได้โดยการใช้สมการที่ 2.15

$$GD(x, y) = \tan^{-1} \left[\frac{Y_{diff}(x, y)}{X_{diff}(x, y)} \right] \quad (2.15)$$

โดยที่ $GD(x,y)$ คือทิศทางของขอบภาพในตำแหน่ง (x,y)

ตัวอย่างการคำนวณหาขอบ ความแรงของขอบ และทิศทางของขอบภาพของภาพแสดงได้ดังนี้

1	1	1	1	1
1	1	1	1	1
1	1	1	0	0
1	1	0	0	0
1	1	0	0	0

(ก) ค่าข้อมูลภาพตัวอย่าง

$$X_{diff} = \begin{matrix} \begin{matrix} 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & * \\ 0 & 0 & -1 & 0 & * \\ 0 & -1 & 0 & 0 & * \\ 0 & -1 & 0 & 0 & * \end{matrix} \end{matrix}$$

$$Y_{diff} = \begin{matrix} \begin{matrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ * & * & * & * & * \end{matrix} \end{matrix}$$

(ข) ค่าความแตกต่างในแนวแกนนอน

(ค) ค่าความแตกต่างในแนวแกนตั้ง

รูปที่ 2.5 แสดงวิธีการพื้นฐานในการหาขอบภาพ

หาผลต่างตามแนวแกนนอนด้วยการคอนโวลูชันภาพกับตัวดำเนินการ

1	-1
---	----

 และหาผลต่างตามแนวแกนตั้งด้วยการคอนโวลูชันภาพกับตัวดำเนินการ

1
-1

 ดังผลลัพธ์ดังนี้
 ความแรงของขอบหาได้จาก $|X_{diff}| + |Y_{diff}|$ และหาทิศทางของขอบดังสมการที่

2.15 ได้ผลลัพธ์ดังนี้

$$\text{ความแรงของภาพ} = \begin{matrix} \begin{matrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{matrix} \end{matrix}$$

$$\text{ทิศทางของขอบภาพ} = \begin{matrix} \begin{matrix} * & * & * & * & * \\ * & * & * & \updownarrow & \updownarrow \\ * & * & \swarrow & * & * \\ * & \leftrightarrow & * & * & * \\ * & \leftrightarrow & * & * & * \end{matrix} \end{matrix}$$

วิธีการหาขอบภาพที่กล่าวมามีข้อเสียคือ การหาขอบภาพไม่ได้นำ ค่าความเข้มแสงของจุดรอบข้างมาใช้ ซึ่งในความเป็นจริงแล้ว ภาพที่ได้มาส่วนใหญ่จะไม่มีขอบภาพชัดเจนดังเช่นที่แสดงในตัวอย่าง ขอบเขตของพื้นที่มักมีการเปลี่ยนแปลงอย่างซ้ำๆ วิธีที่กล่าวมาจะไม่สามารถหาขอบภาพได้อย่างแม่นยำ กับภาพลักษณะนี้ การใช้ตัวดำเนินการที่มีขนาดใหญ่หรือเป็นรูปแบบอื่นๆ จะช่วยแก้ปัญหานี้ได้

2.2.2 ตัวอย่างวิธีการหาขอบภาพ

2.2.2.1 การหาขอบภาพโดยใช้อนุพันธ์อันดับที่หนึ่ง

ยกตัวอย่างการหาขอบภาพโดยใช้อนุพันธ์อันดับที่หนึ่งด้วยตัวดำเนินการโซเบล การหาขอบภาพโดยวิธีโซเบล (Sobel Edge Detection) นั้น เป็นการหาขอบภาพโดยใช้ตัวดำเนินการขนาด 3x3 สองตัว โดยตัวดำเนินการแรกจะใช้หาค่าความแตกต่างในแนวนอน (X_{diff}) และค่าความแตกต่างในแนวตั้ง (Y_{diff}) ดังรูปที่ 2.6

$$X_{diff} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \qquad Y_{diff} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

(ก) ตัวดำเนินการโซเบลในแนวแกนตั้ง (ข) ตัวดำเนินการโซเบลในแนวแกนนอน

รูปที่ 2.6 แสดงตัวดำเนินการการหาขอบภาพด้วยวิธีโซเบล

รูป 2.7 (ก) และ (ข) แสดงตัวอย่างการหาขอบภาพด้วยวิธีโซเบล รูป 2.7 (ค) ได้จากการทำเทรซโฮลด์ด้วยค่า 12 แล้วปรับความเข้มแสงให้อยู่ในช่วง 0-1 ซึ่งสามารถสร้างตัวดำเนินการโซเบลที่มีขนาดใหญ่กว่า 3x3 เพื่อที่จะให้มีการครอบคลุมพื้นที่มากขึ้นได้

0	0	0	0	0	0	2	0	3	3
0	0	0	1	0	0	0	2	4	2
0	0	2	0	2	4	3	3	2	3
0	0	1	3	3	4	3	3	3	3
0	1	0	4	3	3	2	4	3	2
0	0	1	2	3	3	4	4	4	3

4	6	4	10	14	12	14	4
6	8	10	20	16	12	6	0
4	10	14	10	2	4	2	4
2	12	12	2	2	4	6	8

0	0	0	0	1	1	1	0
0	0	0	1	1	1	0	0
0	0	1	0	0	0	0	0
0	1	1	0	0	0	0	0

(ก) ค่าข้อมูลภาพตัวอย่าง (ข) คอนโวลูชันด้วยตัวดำเนินการ (ค) ผลกำหนดขีดเริ่มเปลี่ยน

รูปที่ 2.7 แสดงตัวอย่างการหาขอบภาพด้วยวิธีโซเบล

นอกจากนี้ตัวดำเนินการแบบโซเบลแล้วการหาขอบภาพด้วยอนุพันธ์อันดับหนึ่งที่จุดภาพใดๆ นั้น ยังสามารถหาได้ด้วยตัวดำเนินการที่ต่าง ๆ กัน ได้ดังต่อไปนี้

- ตัวดำเนินการไขว้เกรเดียนต์แบบโรเบิร์ต (Roberts cross-gradient operators)

$$G_x = (z_9 - z_5) \quad (2.16)$$

$$G_y = (z_8 - z_6) \quad (2.17)$$

- ตัวดำเนินการแบบพรีวิตต์ (Prewitt operators)

$$G_x = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3) \quad (2.18)$$

$$G_y = (z_3 + z_6 + z_9) - (z_1 + z_4 + z_7) \quad (2.19)$$

- ตัวดำเนินการแบบโซเบล (Sobel operators)

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3) \quad (2.20)$$

$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7) \quad (2.21)$$

$$G_x = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \quad G_y = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

(ก) ตัวดำเนินการไขว้เกรเดียนต์แบบโรเบิร์ต

$$G_x = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad G_y = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

(ข) ตัวดำเนินการแบบพรีวิตต์

$$G_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad G_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

(ค) ตัวดำเนินการแบบโซเบล

รูปที่ 2.8 แสดงตัวดำเนินการต่างๆที่ใช้ในการคำนวณหาขอบภาพ

2.2.2.2 การหาขอบภาพด้วยตัวดำเนินการดีกรีสอง

หากภาพที่ต้องการหาขอบมีการไล่ระดับความเข้มแสงแบบเป็นเชิงเส้น การใช้ตัวดำเนินการโซเบลในการหาขอบภาพจะไม่สามารถทำได้ ตัวอย่างเช่น เมื่อหาขอบภาพตามแกนตั้งด้วยตัวดำเนินการ

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \text{ จะได้ผลลัพธ์ดังนี้}$$

1	2	3	4	5	6	7	8	9
1	2	3	4	5	6	7	8	9
1	2	3	4	5	6	7	8	9
1	2	3	4	5	6	7	8	9
1	2	3	4	5	6	7	8	9

(ก) ค่าข้อมูลภาพตัวอย่าง

*	*	*	*	*	*	*	*	*
*	8	8	8	8	8	8	8	*
*	8	8	8	8	8	8	8	*
*	8	8	8	8	8	8	8	*
*	*	*	*	*	*	*	*	*

*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*
*	0	0	0	0	0	0	0	*
*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*

(ข) คอนโวลูชันด้วยตัวดำเนินการครั้งที่ 1 (ค) คอนโวลูชันด้วยตัวดำเนินการครั้งที่ 2

รูปที่ 2.9 แสดงการหาขอบภาพด้วยตัวดำเนินการดิกรีสอง

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

รูปที่ 2.10 แสดงตัวอย่างตัวดำเนินการดิกรีสองสำหรับการหาขอบภาพ (ลาปลาเซียน)

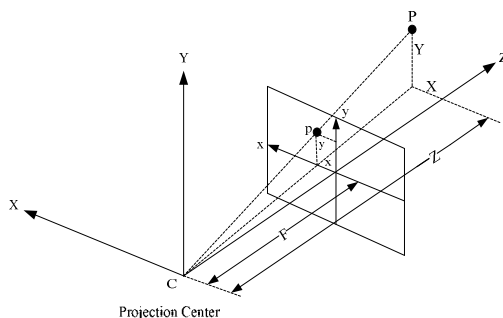
ตัวดำเนินการลาปลาเซียนเป็นที่นิยมใช้กันมาก เพราะสามารถกำจัดการไล่โทนสีหรือโทนความเข้มแสงได้ มันสามารถเน้นความเปลี่ยนแปลงได้ดีกว่าวิธีโซเบล แต่มันไม่ให้ข้อมูลเกี่ยวกับทิศทางของขอบ

2.3 แบบจำลองกล้องวิดิทัศน์และแบบจำลองการเคลื่อนที่ของกล้องวิดิทัศน์

เนื่องจากในวิทยานิพนธ์ฉบับนี้มีการทำงานร่วมกับกล้องวิดิทัศน์โดยมีการควบคุมกล้องวิดิทัศน์ให้เคลื่อนที่ไปยังตำแหน่งที่ต้องการ ดังนั้นทฤษฎีการทำงานเกี่ยวกับแบบจำลองกล้องจึงมีความสำคัญในด้านการมโนภาพ (Imagine) หรือเพื่อให้เข้าใจหลักการเกี่ยวกับการควบคุมกล้องวิดิทัศน์เพิ่มมากขึ้น โดยแบบจำลองกล้องวิดิทัศน์นั้น เป็นการจำลองรูปแบบการสร้างภาพที่ปรากฏอยู่บนจอภาพในพิกัดภาพ (Image coordinate) 2 มิติ จากวัตถุในฉากจริง (Real scene) ในพิกัด 3 มิติ มีแบบจำลองหลายรูปแบบที่ใช้บรรยายการสร้างภาพ เช่น แบบจำลองกล้องรูเข็ม (Pinhole camera model) แบบจำลองกล้อง CAHV (CAHV Camera model) เป็นต้น แต่

แบบจำลองกล้องทัศนศาสตร์สามารถแสดงความเข้าใจพื้นฐานเกี่ยวกับกล้องทัศนศาสตร์ได้อย่างดี จึงขอยกตัวอย่างแบบจำลองของกล้องทัศนศาสตร์แบบรูเข็มในวิทยานิพนธ์ฉบับนี้

2.3.1 แบบจำลองกล้องทัศนศาสตร์แบบรูเข็ม



รูปที่ 2.11 การฉายภาพแบบทัศนมิติ (Perspective projection) ของแบบจำลองกล้องรูเข็ม

แบบจำลองกล้องรูเข็ม เป็นแบบจำลองที่นิยมใช้กันอย่างมากมายในการวิเคราะห์การสร้างภาพและงานประยุกต์ใน 3 มิติ แสดงได้ดังรูปที่ 2.6 โดยจะอาศัยพื้นฐานการฉายภาพแบบทัศนมิติ ในการสร้างภาพในพิกัด 2 มิติจากวัตถุในพิกัด 3 มิติ พิจารณาที่จุดพิกัดวัตถุ $P(X, Y, Z)$ ในพิกัด 3 มิติ สามารถฉายลงมายังจุด $P(x, y)$ ในระนาบภาพ (Imaging plane) 2 มิติ ได้ดังสมการที่ 2.22 โดยมีจุดตำแหน่ง C เป็นจุดศูนย์กลางในการฉายภาพ (Center of projection)

$$x = F \frac{X}{Z} \quad \text{และ} \quad y = F \frac{Y}{Z} \quad (2.22)$$

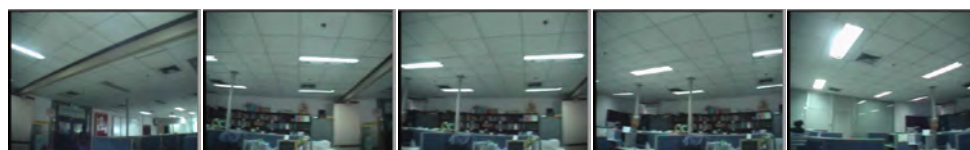
โดยที่ F คือ ความยาวโฟกัส (Focal length)

2.3.2 การเคลื่อนที่ของกล้องทัศนศาสตร์

การเคลื่อนที่ของกล้องโดยทั่วไปจะแบ่งออกได้เป็น 2 รูปแบบ คือ

2.3.2.1 การเคลื่อนที่ของกล้องที่ตำแหน่งศูนย์กลางการฉายภาพไม่เปลี่ยนแปลง ดังรูปที่ 2.11 (ก) ประกอบด้วย การส่าย (Pan) เป็นลักษณะการเคลื่อนที่ที่กล้องโดยยึดแกนตั้ง (Vertical axis) เป็นแกนหมุนรอบ การก้มเงย (Tilt) เป็นลักษณะการเคลื่อนที่ที่กล้องโดยยึดแกนนอน (Horizontal axis) เป็นแกนหมุนรอบ และ การหมุน (Roll) เป็นลักษณะการเคลื่อนที่ที่กล้องโดยยึดแกนเชิงแสง (Optical axis) เป็นแกนหมุนรอบ

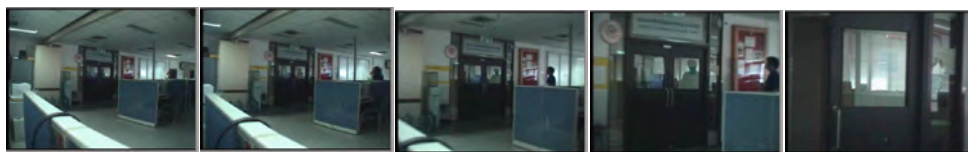
2.3.2.2 การเคลื่อนที่ของกล้องที่ตำแหน่งศูนย์กลางการฉายภาพมีการเปลี่ยนแปลงตำแหน่ง ดังรูปที่ 2.11 (ข) ประกอบด้วย การติดตาม (Track) เป็นลักษณะการเคลื่อนที่กล้องในลักษณะการเลื่อนขนาน (Translation) ตามแนวแกนของระนาบภาพ การยก (Boom) เป็นลักษณะการเคลื่อนที่กล้องในลักษณะการเลื่อนขนานตามแนวแกนตั้งของระนาบภาพ และการดอลลี (Dolly) เป็นลักษณะการเคลื่อนที่กล้องในลักษณะการเลื่อนขนานตามแนวแกนเชิงแสง



รูปที่ 2.13 ลำดับภาพวีดิทัศน์ตัวอย่างที่ได้จากการเคลื่อนที่ของกล้องแบบการถ่าย

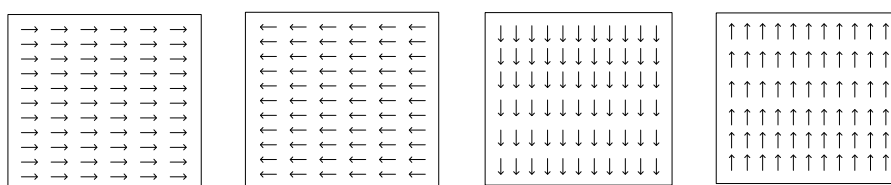


รูปที่ 2.14 ลำดับภาพวีดิทัศน์ตัวอย่างที่ได้จากการเคลื่อนที่ของกล้องแบบการก้มเงย



รูปที่ 2.15 ลำดับภาพวิดีโอที่บันทึกตัวอย่างที่ได้จากการเคลื่อนที่ของกล้องแบบซูม

ในวิทยานิพนธ์ฉบับนี้จะประยุกต์ใช้ลักษณะการเคลื่อนที่ดังหัวข้อที่ 2.2.2.1 คือ การส่ายและการก้มเงย ส่วนในลักษณะการซูมนั้น วิทยานิพนธ์ฉบับนี้ไม่ได้ทำการทดสอบไว้ เนื่องจากการติดตามบุคคลนั้นไม่จำเป็นต้องเก็บรายละเอียดของขนาดของบุคคลเป้าหมาย การส่ายและก้มเงย เป็นการหมุนกล้องรอบแกนตั้ง (Y) เป็นการหมุนกล้องรอบแกนตั้ง (X) ตามลำดับสนามการเคลื่อนที่ (Motion field) ของการส่ายและการก้มเงยสามารถแสดงได้ดังรูปที่ 2.15



รูปที่ 2.16 สนามการเคลื่อนที่ของการส่ายและก้มเงย

2.4 การเทียบมาตรฐานของกล้องวิดีโอ (Camera Calibration)

วิธีการเทียบมาตรฐานของกล้องวิดีโอ เป็นวิธีที่นิยมใช้กันมากในกระบวนการที่ในการประมวลผลภาพวิดีโอร่วมกันระหว่างกล้องวิดีโอหลายๆตัว ประโยชน์หลักของกระบวนการนี้คือการปรับภาพที่จะถูกแสดงผ่านกล้องวิดีโอแต่ละตัวในระบบให้มีความถูกต้อง เพราะโดยทั่วไปกล้องวิดีโอทุกตัวที่ถูกติดตั้งให้ทำงานร่วมกันนั้น จะให้ภาพที่แสดงออกมาไม่ตรงกับลักษณะของวัตถุที่จริง ไม่ว่าจะเป็นขนาด ความกว้าง ความยาวของวัตถุ เช่น เมื่อกล้องวิดีโอจับภาพของกรอบสี่เหลี่ยม แต่ได้ภาพออกมามีลักษณะโค้งเล็กน้อย เป็นต้น หรือเรื่องของข้อมูลสี ความสว่างของบริเวณที่ติดตั้งกล้องวิดีโออาจจะทำให้ค่าสีที่อ่านได้ต่างกัน ผลที่ได้จากการเทียบมาตรฐานกล้องวิดีโอทั้งสองแบบดังกล่าวนั้นก็คือ จะได้ค่าพารามิเตอร์ต่างๆ สำหรับสร้างความสัมพันธ์ระหว่างกัน ซึ่งอาจแทนด้วยเมทริกซ์ใดเมทริกซ์หนึ่ง ในวิทยานิพนธ์ฉบับนี้ขอยกตัวอย่างการเทียบมาตรฐานของกล้องวิดีโอสองรูปแบบคือ การเทียบมาตรฐานกล้องวิดีโอทางตำแหน่งและการเทียบมาตรฐานกล้องวิดีโอทางสี

วิธีการเทียบมาตรฐานกล้องวิดีโอทางตำแหน่ง จะมีหลักการ คือ ให้กล้องวิดีโอที่ต้องการเทียบมาตรฐานนั้นรับภาพหรือถ่ายภาพวัตถุที่ทราบขนาดกว้างยาวที่แน่นอน ดังรูปที่ 2.16 (ก) ปกติจะใช้เป็นแผ่นตารางหมากรุกที่ทราบความกว้างยาวของแต่ละช่องของตารางที่แน่นอนอยู่

แล้ว จากนั้นนำไปประมวลผลโดยพิจารณาลักษณะ รูปร่าง ขนาด และอัตราส่วนของช่องตารางที่กล้องวิดีโอตัวนั้นรับภาพได้เทียบกับขนาดจริงที่ควรจะเป็น ซึ่งจะขอไม่ลงรายละเอียดในขั้นตอนระเบียบวิธีการหาค่าพารามิเตอร์ต่างๆ เนื่องจากวิทยานิพนธ์ฉบับนี้ไม่ได้ใช้หลักการนี้

ส่วนวิธีการเทียบมาตรฐานกล้องวิดีโอทางสีนั้น หลักการจะเช่นเดียวกับ วิธีการเทียบมาตรฐานกล้องวิดีโอทางตำแหน่ง คือ ให้กล้องวิดีโอที่ต้องการเทียบมาตรฐานนั้นรับภาพหรือถ่ายภาพวัตถุที่ทราบข้อมูลสีที่แน่นอน ดังรูปที่ 2.16 (ข) ปกติแล้วงานที่มักให้ความสำคัญกับสี เช่น งานการพิมพ์ภาพ หรือถ่ายภาพ ในโรงพิมพ์ จะใช้ตารางสีอ้างอิงมาตรฐาน (Color checker board) เป็นอุปกรณ์ในการเทียบ ขั้นตอนวิธีการเทียบมาตรฐานทางสีที่นำเสนอจะอธิบายในบทที่ 3 ต่อไป



(ก) วัตถุที่ทราบความกว้างยาวของแต่ละช่อง



(ข) วัตถุที่ทราบค่าข้อมูลสีแต่ละช่อง

รูปที่ 2.17 ตัวอย่างภาพวัตถุที่ใช้ในการเทียบมาตรฐานกล้องวิดีโอ

2.5 พื้นฐานการติดตามวัตถุหรือบุคคลในลำดับภาพวิดีโอ

การติดตามวัตถุหรือบุคคลที่เคลื่อนที่ภายในลำดับภาพวิดีโอที่ต่อเนื่องนั้น คือ การพยายามที่จะระบุตำแหน่งของบุคคลหรือวัตถุในลำดับภาพวิดีโอให้ถูกต้องและต่อเนื่องไปในทุกๆลำดับภาพ หรือทำนายว่าทิศทางการเคลื่อนที่ของวัตถุหรือบุคคลนั้นควรจะเป็นอย่างไร ตัวอย่างเช่นแสดงดังรูปที่ 2.17 เมื่อจักรยานมีการเคลื่อนที่ภายในลำดับภาพ วิธีการติดตามนั้นจะต้องพยายามระบุตำแหน่งและทำนายว่าจักรยาน ควรจะอยู่ ณ ตำแหน่งใดในลำดับภาพต่อไป นอกจากนี้การติดตามวัตถุหรือบุคคลในระดับการทำงานทางคอมพิวเตอร์วิทัศน์ที่สูงขึ้น อาจแสดงถึงการกระทำของวัตถุหรือบุคคล ซึ่งอาจจะนำไปสู่การรู้จำได้อีกด้วย นักวิจัยหลายท่านได้นำเสนอการติดตามวัตถุหรือบุคคลที่เคลื่อนที่ ซึ่งโดยทั่วไปแล้วสามารถแบ่งออกได้เป็น 4 ประเภทคือ การติดตามวัตถุหรือบุคคลด้วยโมเดลรูปร่าง (Model-based tracking) การติดตามวัตถุหรือบุคคลด้วยพื้นที่ (Region-based tracking) การติดตามวัตถุหรือบุคคลด้วยขอบ (Contour-based tracking) และการติดตามวัตถุหรือบุคคลโดยอาศัยคุณลักษณะเฉพาะ (Feature-based tracking)



รูปที่ 2.18 ตัวอย่างลำดับภาพวิถีทัศนที่มีวัตถุเคลื่อนที่

- การติดตามวัตถุหรือบุคคลด้วยโมเดลรูปร่าง เป็นวิธีที่จะติดตามวัตถุที่มีรูปร่างลักษณะเข้ากับโมเดลรูปร่างที่เตรียมไว้ ซึ่งการติดตามวัตถุหรือบุคคลด้วยวิธีนี้ จำเป็นต้องสร้างโมเดลรูปร่างที่สนใจไว้ก่อน ดังนั้นวิธีการนี้จึงมีข้อจำกัด คือ ไม่สามารถทำงานกับวัตถุที่ไม่ทราบรูปร่างล่วงหน้าได้

- การติดตามวัตถุหรือบุคคลด้วยพื้นที่ เป็นการดึงข้อมูลสำคัญออกมาเป็นลักษณะของพื้นที่ที่ประกอบไปด้วยข้อมูลใดๆ เช่น ข้อมูลของสี ข้อมูลลวดลาย เป็นต้น จากนั้นจะทำการติดตามพื้นที่เหล่านั้นโดยหาข้อมูลของสิ่งที่ดึงออกมาใช้นั่นเอง

- การติดตามวัตถุหรือบุคคลด้วยขอบ วิธีนี้จะขึ้นอยู่กับขอบของวัตถุหรือบุคคลที่มีการเคลื่อนที่ ไม่ได้ขึ้นกับตัววัตถุทั้งชิ้น ข้อมูลทางขอบวัตถุจะถูกดึงออกมาถูกปรับแต่งในแต่ละลำดับภาพที่เกิดขึ้น ดังนั้นการติดตามวัตถุที่เคลื่อนที่ได้ผลดีหรือไม่ขึ้นกับขอบวัตถุเริ่มแรกที่ถูกดึงออกมาได้ ซึ่งวิธีนี้ยากต่อการนำมาใช้ในระบบตรวจจับเพื่อรักษาความปลอดภัย

- การติดตามวัตถุหรือบุคคลโดยอาศัยคุณลักษณะเฉพาะ มีจุดประสงค์ที่จะค้นหาและติดตามวัตถุที่เคลื่อนที่โดยใช้คุณลักษณะเฉพาะบางอย่างที่ได้จากการคำนวณในระดับสูงขึ้นไป เช่น การประมาณความเร็วของวัตถุ การหาค่าการกระจายตัวของข้อมูลของวัตถุ การประมาณความหนาแน่นของข้อมูลวัตถุ เป็นต้น ซึ่งวิธีการประเภทนี้จะมีข้อดี คือ วัตถุหรือบุคคลที่จะติดตามนั้นจะมีข้อมูลแยกออกมาจากสิ่งแวดล้อมได้อย่างดีกว่าวิธีอื่นๆ งานวิจัยที่เกี่ยวกับ การติดตามวัตถุหรือบุคคลโดยอาศัยคุณลักษณะเฉพาะนี้ เช่น การติดตามวัตถุด้วยกล้องวิถีทัศนหลายตัว โดยใช้เทคนิคการเปรียบเทียบข้อมูลคุณลักษณะเฉพาะของวัตถุในทุกๆลำดับภาพ การติดตามวัตถุขนาดเล็กจากกล้องวิถีทัศนบริเวณอากาศยาน โดยใช้การคำนวณเคลื่อนที่ของตัวอากาศยานเองร่วมกับทิศทางกรไหลของคลื่นในอากาศ เป็นต้น

2.6 การประมาณความหนาแน่น (Density Estimation)

ฟังก์ชันความหนาแน่น เป็นพื้นฐานที่สำคัญของการวิเคราะห์งานเกี่ยวกับคอมพิวเตอร์วิทัศน์ ทั้งนี้เพราะเมื่อทราบฟังก์ชันความหนาแน่นจะทำให้ทราบรูปร่างของการแจกแจงของข้อมูล และหาความน่าจะเป็นของการเกิดค่าต่างๆ ของข้อมูลที่สนใจได้ การประมาณความหนาแน่นจึงมีวัตถุประสงค์หลักเพื่อพยายามที่จะประมาณรูปภาพของฟังก์ชันความหนาแน่นของข้อมูลที่เรารู้ความสนใจ เพื่อนำไปใช้งานในด้านต่างๆ

โดยทั่วไปแล้วการประมาณความหนาแน่นนั้นอาจแบ่งได้เป็น 2 แบบด้วยกัน คือ การประมาณความหนาแน่นแบบกำหนดตัวแปร (Parametric density estimation) และการประมาณความหนาแน่นแบบไม่กำหนดตัวแปร (Non-parametric density estimation) โดยการประมาณความหนาแน่นแบบกำหนดตัวแปร ทำได้ด้วยการสมมติว่าตัวอย่างถูกสุ่มมาจากข้อมูลในกลุ่มของการแจกแจงที่ทราบ ตัวอย่างเช่น สมมติว่า ตัวอย่างถูกสุ่มมาจากข้อมูลที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย μ และความแปรปรวน σ^2 ดังนั้นฟังก์ชันความหนาแน่น $f(x)$ สามารถประมาณได้โดยการหาค่าประมาณของ μ และ σ^2 จากตัวอย่าง และแทนค่าประมาณดังกล่าวลงในสูตรฟังก์ชันความหนาแน่นแบบปกติ ในวิทยานิพนธ์ฉบับนี้จะไม่พิจารณาการประมาณความหนาแน่นแบบกำหนดตัวแปร แต่เราจะพิจารณาการประมาณความหนาแน่นแบบไม่กำหนดตัวแปรเพียงอย่างเดียวเนื่องจากวิธีการประมาณความหนาแน่นแบบไม่กำหนดตัวแปรนี้ มีข้อสมมติเกี่ยวกับการแจกแจงของข้อมูลที่ตัวอย่างถูกสุ่มมาที่มีความยืดหยุ่นมากกว่า และข้อมูลตัวอย่างควรจะถูกใช้ในการกำหนดค่าประมาณความหนาแน่นมากกว่าที่จะสมมติฟังก์ชันความหนาแน่น $f(x)$ ให้มีการแจกแจงอย่างใดอย่างหนึ่ง

นักสถิติหลายท่านมีความพยายามที่จะประมาณความหนาแน่นของประชากรจากชุดของตัวอย่างสุ่มที่สุ่มมาจากข้อมูล ซึ่งการประมาณความหนาแน่นที่จะกล่าวถึงนี้ถูกเสนอขึ้นครั้งแรกเพื่อที่จะลดข้อสมมติเกี่ยวกับการแจกแจงข้อมูล โดยทั่วไปการประมาณความหนาแน่นใช้เพื่อการตรวจสอบคุณสมบัติของเซตข้อมูลซึ่งสามารถบอกถึงลักษณะเด่นๆ ของข้อมูลที่เป็นไปได้ เช่น ความโด่ง ความลาดเอียง หรือการมีหลายจุดสูงสุด (Mode) ในข้อมูล และจะแสดงให้เห็นลักษณะของข้อมูลจริง

วิธีการประมาณความหนาแน่นดังกล่าวนี้ เพื่อแสดงความเข้าใจในการอธิบายตัวแปรให้ตรงกันและชัดเจน จึงกำหนดสัญลักษณ์ที่ใช้เป็นมาตรฐาน โดยสมมติว่า x_1, x_2, \dots, x_n เป็นตัวอย่างสุ่มที่มาจากข้อมูลใดๆ ในแฟรมภาพวีดิทัศน์ที่เราสนใจซึ่งมีฟังก์ชันความหนาแน่น $f(x)$ และ

กำหนดให้ $\hat{f}(x)$ เป็นตัวประมาณความหนาแน่น เราจะเริ่มโดยการแนะนำวิธีการประมาณความหนาแน่นที่ง่ายและถูกใช้กันมานาน นั่นคือการประมาณโดยใช้ฮิสโตแกรม

2.6.1 การประมาณความหนาแน่นด้วยฮิสโตแกรม (Histogram)

วิธีการประมาณความหนาแน่นที่มีมานานและนิยมใช้กันอย่างแพร่หลาย คือ ฮิสโตแกรม การประมาณความหนาแน่นด้วยวิธีฮิสโตแกรมนี้จำเป็นต้องกำหนดจุดกำเนิดเริ่มต้นและค่าช่วงของความกว้างเสียก่อน สมมติว่ามี x_0 เป็นจุดกำเนิดเริ่มต้น และ h เป็นขนาดความกว้างของช่วง (Bin width) สามารถสร้างช่วงความกว้างต่างๆ ของฮิสโตแกรมที่มีขนาด h ได้จาก $[(x_0 + mh) , (x_0 + (m+1)h))$ สำหรับจำนวนเต็มบวกและลบ m ช่วงความกว้างถูกเลือกให้เป็นช่วงปิดทางซ้าย และช่วงเปิดทางขวา และค่าประมาณความหนาแน่น $f(x)$ ที่จุด x จากฮิสโตแกรมดังสมการ 2.23

$$\hat{f}(x) = \frac{1}{nh} \cdot (\text{จำนวนค่าสังเกตที่ตกอยู่ในช่วงเดียวกันกับ } x) \quad (2.23)$$

เมื่อ n เป็น ขนาดตัวอย่าง สังเกตว่าการสร้างฮิสโตแกรมจะต้องเลือกจุดเริ่มต้นและความกว้างของช่วง ซึ่งความกว้างของช่วงจะเป็นตัวควบคุมความราบเรียบของรูปฮิสโตแกรมแนวคิดของการประมาณความหนาแน่นแบบนี้ คือ การตรึงจุดกำเนิดและความกว้างของช่วงต่างๆ ไว้ ความสูงของฮิสโตแกรมจะได้จากจำนวน x_i ที่ตกอยู่ในช่วงต่างๆ

สำหรับการแสดงและการตรวจสอบข้อมูล ฮิสโตแกรมถือเป็นเครื่องมือที่มีประโยชน์อย่างมากในการประมาณความหนาแน่นในกรณีของข้อมูลตัวแปรเดี่ยว อย่างไรก็ตามเมื่อพิจารณาจากการประมาณความหนาแน่นด้วยวิธีฮิสโตแกรม พบว่าลักษณะรูปร่างของการแจกแจงโดยการใช้ฮิสโตแกรมมีความแตกต่างกันขึ้นอยู่กับการกำหนดจุดกำเนิดและความกว้างของช่วง นอกจากนี้ไม่สามารถหาอนุพันธ์ของฟังก์ชันประมาณความหนาแน่นที่จุดต่างๆ ได้ และสำหรับการแสดงรูปร่างของฮิสโตแกรมในข้อมูลตั้งแต่ 2 ตัวแปรขึ้นไปนั้นการสร้างยังมีความลำบากขึ้น

จากจุดบกพร่องต่างๆ ของการประมาณความหนาแน่นด้วยวิธีฮิสโตแกรมจึงมีการพัฒนาวิธีการประมาณความหนาแน่นให้ดีขึ้น เช่น การประมาณความหนาแน่นแบบง่าย (Naive estimator) ซึ่งเป็นพื้นฐานสำหรับการประมาณความหนาแน่นแบบเคอร์เนลซึ่งเป็นวิธีการประมาณความหนาแน่นสำหรับใช้ในวิทยานิพนธ์ฉบับนี้

2.6.2 ตัวประมาณความหนาแน่นแบบง่าย (The naïve estimator)

วิธีการใช้ตัวประมาณความหนาแน่นแบบง่ายนี้ ใช้หลักการที่ว่าถ้าตัวแปรสุ่ม x มีฟังก์ชันความหนาแน่น $f(x)$ ดังสมการ 2.24

$$f(x) = \lim_{x \rightarrow \infty} \frac{1}{2h} P(x-h < X < x+h) \quad (2.24)$$

สำหรับค่า h ที่กำหนดใด ๆ เราสามารถประมาณ $P(x-h < X < x+h)$ โดยใช้สัดส่วนของตัวอย่างที่อยู่ในช่วง $(x-h < X < x+h)$ ดังนั้นตัวประมาณความหนาแน่น $\hat{f}(x)$ ของ $f(x)$ ที่เป็นธรรมชาติ คือการเลือกค่า h เล็กๆ และกำหนดสมการ 2.25

$$\hat{f}(x) = \frac{1}{2nh} \cdot (\text{จำนวนค่าสังเกตที่ตกอยู่ในช่วง}(x-h, x+h)) \quad (2.25)$$

ซึ่งเรียกรูปวิธีการประมาณความหนาแน่นด้วยวิธีการแบบนี้ว่าการประมาณความหนาแน่นแบบง่าย โดยที่ h เป็นความกว้างช่วงเพื่อความเข้าใจวิธีของการประมาณความหนาแน่นแบบง่ายนี้ ในการประมาณฟังก์ชันความหนาแน่น $f(x)$ มากขึ้นจะสมมติให้มีฟังก์ชันถ่วงน้ำหนัก w

โดยที่

$$w(x) = \begin{cases} \frac{1}{2} & ; |x| < 1 \\ 0 & ; \text{otherwise} \end{cases} \quad (2.26)$$

ดังนั้นวิธีการประมาณความหนาแน่นแบบง่าย สามารถเขียนได้ในรูปของฟังก์ชันถ่วงน้ำหนัก w ดังสมการที่ 2.27

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} w\left(\frac{x-x_i}{h}\right) \quad (2.27)$$

ซึ่งเปรียบเสมือนกับการนำกล่องที่มีความกว้าง $2h$ และสูง $(2nh)^{-1}$ วางลงที่ค่าสังเกตแต่ละค่า และค่าประมาณความหนาแน่นที่จุด x ใดๆ จะได้จากส่วนของพื้นที่ซึ่งซ้อนทับกันขึ้นเป็น

ความสูงของค่าประมาณความหนาแน่นที่จุด x แต่การประมาณความหนาแน่นแบบง่ายนี้ยังคงมีปัญหาของการเลือกความกว้างช่วง h อยู่ ซึ่งจะมีผลต่อการประมาณความหนาแน่นเนื่องจากช่วงกว้างดังกล่าวนี้จะเป็นตัวควบคุมความราบเรียบของตัวประมาณความหนาแน่นดังนั้นการเลือกความกว้างช่วงที่ต่างกันจะทำให้รูปร่างของการแจกแจงที่ประมาณแตกต่างกันไปด้วย

จากแนวความคิดพื้นฐานของ 2 วิธีการที่กล่าวมาแล้ว ปัญหาที่สำคัญของการประมาณความหนาแน่น คือ ความไม่ต่อเนื่องและไม่สามารถหาอนุพันธ์ของฟังก์ชันที่ประมาณได้ ดังนั้นการประมาณความหนาแน่นแบบเคอร์เนลจึงได้รับการพัฒนาขึ้นเพื่อแก้ปัญหาดังกล่าว

2.6.3 การประมาณความหนาแน่นแบบเคอร์เนล (Kernel density estimation)

จากวิธีการประมาณความหนาแน่นแบบง่ายถ้าแทนฟังก์ชันถ่วงน้ำหนัก $w\left(\frac{x-x_i}{h}\right)$ ด้วยฟังก์ชันเคอร์เนล (Kernel function; K) ซึ่งสอดคล้องกับเงื่อนไขดังนี้

$$\int_{-\infty}^{\infty} K(x)dx = 1 \quad \text{หรือ} \quad \sum_{i=1}^n K(x_i) = 1$$

ซึ่งโดยทั่วไป $K(x)$ มักจะเป็นฟังก์ชันความหนาแน่นที่สมมาตร และจะได้วิธีการประมาณความหนาแน่นที่เรียกว่าวิธีการประมาณความหนาแน่นแบบเคอร์เนลซึ่งนิยามได้ดังสมการที่ 2.28

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (2.28)$$

เมื่อ $K(x)$ เป็นฟังก์ชันเคอร์เนล และ h เป็นความกว้างวินโดว์ (Window width) แสดงตัวอย่างการคำนวณการประมาณความหนาแน่นด้วยฟังก์ชันเคอร์เนลแสดงได้ดังนี้

สมมติมีข้อมูลดังนี้ 0.0 0.8 1.0 1.2 1.3 1.3 1.4 1.8 2.4 4.6 ต้องการประมาณความหนาแน่นที่ $x = 4$ เมื่อกำหนดฟังก์ชันเคอร์เนลให้เป็นแบบการแจกแจงปกติมาตรฐาน และกำหนดความกว้างวินโดว์ เป็น $\frac{2s}{n^{1/5}}$ โดยที่ s คือ ส่วนเบี่ยงเบนมาตรฐาน ซึ่งสมมติเป็น 1.51

เพราะฉะนั้นความกว้างวินโดว์จึงเท่ากับ $\frac{(2 \times 1.51)}{10^{1/5}} = 1.9$ และแสดงการคำนวณค่าของ

$K\left(\frac{x-x_i}{h}\right)$ ที่ $x = 4$ ได้ดังตารางที่ 2.1

ตารางที่ 2.1 แสดงตัวอย่างการคำนวณการประมาณความหนาแน่นแบบเคอร์เนล

ค่าสังเกต (x_i)	$\left(\frac{x-x_i}{h}\right)$	$K\left(\frac{x-x_i}{h}\right)$
0.0	2.10526	0.043500
0.8	1.68421	0.096596
1.0	1.57895	0.114695
1.2	1.47368	0.134686
1.3	1.42105	0.145347
1.3	1.42105	0.145347
1.4	1.36842	0.156417
1.8	1.15789	0.204069
2.4	0.84211	0.27848
4.6	0.31579	0.379538
รวม		1.70004

จะได้ $\sum_{i=1}^{10} K\left(\frac{4-x_i}{1.9}\right) = 1.70004$ จากนิยามขั้นตอนค่าประมาณความหนาแน่นที่ตำแหน่ง $x = 4$ คำนวณได้ดังนี้

$$\hat{f}(4) = \frac{1}{10 \times 1.9} [1.70004] = 0.08947$$

ดังนั้นค่าประมาณความหนาแน่นที่ตำแหน่ง $x = 4$ มีค่าเป็น 0.08947

ในขณะที่วิธีการประมาณความหนาแน่นแบบง่ายประมาณความหนาแน่นที่จุด x โดยเปรียบเสมือนการใช้กล่องสี่เหลี่ยมที่มีความกว้าง $2h$ และสูง $(2nh)^{-1}$ วางลงที่ค่าสังเกตแต่ละค่า และค่าประมาณความหนาแน่นที่จุด x ใดๆ จะได้จากส่วนของพื้นที่ซึ่งซ้อนทับกันขึ้นเป็นความสูงของค่าประมาณความหนาแน่นที่จุด x หรือวิธีการประมาณความหนาแน่นแบบเคอร์เนลนี้ จะเปรียบเสมือนกับการนำโค้ง (ตามรูปของเคอร์เนลที่ใช้) วางลงที่ค่าสังเกตแต่ละค่าและค่าประมาณความหนาแน่นที่จุด x ได้จากผลรวมของโค้งที่จุด x ดังกล่าว

ค่าของ h ในวิธีการประมาณความหนาแน่นแบบเคอร์เนล จะเรียกว่าความกว้างวินโดว์ หรือแบนด์วิธ (Bandwidth) หรือบางครั้งเรียกว่า Smoothing parameter ทั้งนี้เพราะค่า h จะเป็นตัวกำหนดความราบเรียบของฟังก์ชันที่จะประมาณ วิธีประมาณความหนาแน่นแบบเคอร์เนลมีข้อดีหลายประการ คือ ฟังก์ชันที่ประมาณจะมีลักษณะต่อเนื่องและหาอนุพันธ์ได้ที่ทุกจุดแต่อย่างไรก็ตามไม่อาจกล่าว่ว่าวิธีประมาณความหนาแน่นแบบเคอร์เนลดีที่สุดที่สุดในบรรดาวิธีการประมาณความหนาแน่นแบบต่างๆ แต่เป็นวิธีที่นิยมใช้กันแพร่หลายมากที่สุดและได้มีผู้ศึกษาคุณสมบัติของวิธีประมาณความหนาแน่นไว้เป็นจำนวนมาก ดังนั้นผู้วิจัยจึงจะใช้วิธีการประมาณความหนาแน่นแบบเคอร์เนล เพื่อประมาณความหนาแน่นของประชากรสำหรับใช้ในงานวิจัยครั้งนี้ เนื่องจากวิธีประมาณความหนาแน่นแบบเคอร์เนลเป็นวิธีการที่จะใช้เป็นพื้นฐานของการติดตามบุคคลในวิทยานิพนธ์ฉบับนี้ จึงขอกล่าวถึงคุณสมบัติต่างๆของวิธีการนี้พอสังเขป

2.7 ขั้นตอนวิธีการย้ายตามค่าเฉลี่ย (Mean Shift Algorithm)

จากบทที่ 1 ได้กล่าวถึงที่มาของขั้นตอนวิธีการย้ายตามค่าเฉลี่ยที่ได้ถูกคิดค้นขึ้นในอดีต ซึ่งจะเห็นได้ว่า ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยนั้นสามารถนำมาประยุกต์ใช้งานได้หลากหลาย เช่น การจัดกลุ่มและแยกส่วนภาพ การทำภาพให้เรียบ การรู้จำภาพ รวมถึงการติดตามวัตถุหรือบุคคล ดังที่วิทยานิพนธ์ฉบับนี้สนใจด้วย และอีกทั้งจากหัวข้อ 2.6 ที่ผ่านมา เนื่องจากขั้นตอนวิธีการย้ายตามค่าเฉลี่ยนั้นมีพื้นฐานมาจากการประมาณความหนาแน่นแบบเคอร์เนล จึงจำเป็นต้องใช้ความรู้เกี่ยวกับขั้นตอนของการประมาณความหนาแน่นแบบเคอร์เนลมาประยุกต์ใช้ด้วย แสดงหลักการของขั้นตอนวิธีการย้ายตามค่าเฉลี่ยได้ดังนี้

กำหนด x_1, x_2, \dots, x_n เป็นเซตของข้อมูลที่สนใจ n ข้อมูล จากงานวิจัยของ Fugunaga [] ได้เสนอขั้นตอนวิธีการย้ายตามค่าเฉลี่ยที่มีการคำนวณอย่างไม่ซับซ้อน แต่สามารถรับประกันได้ว่าทิศทางการเปลี่ยนแปลงตำแหน่งของการคำนวณตามขั้นตอนวิธี จะมีทิศเข้าหาบริเวณที่มีความหนาแน่นมากเสมอ จึงนำเสนอวิธีการประมาณความหนาแน่น กำหนดโดยสมการที่ 2.29

$$P_n(x) = \frac{1}{nh} \sum_{s=x_1}^{x_n} k\left(\frac{x-s}{h}\right) \quad (2.29)$$

$$\text{โดยที่ } \int_{-\infty}^{\infty} k(x)dx = 1 \quad \text{หรือ} \quad \sum_{i=1}^n k(x_i) = 1$$

เพื่อต้องการหาจุดสูงสุดหรือโหนดของฟังก์ชันของความหนาแน่นดังกล่าว จึงเสนอการประมาณความหนาแน่นกราดิเยน สามารถคำนวณได้ดังสมการ 2.30

$$\nabla P_n(x) = \frac{1}{nh} \sum_{s=x_1}^{x_n} \nabla k\left(\frac{x-s}{h}\right) \quad (2.30)$$

โดยที่ $\nabla k(x) = \left(\frac{\partial k(x)}{\partial x_1}, \frac{\partial k(x)}{\partial x_2}, \dots, \frac{\partial k(x)}{\partial x_n} \right)^T$ และ $k(x)$ คือ ฟังก์ชันเคอร์เนล

เมื่อพิจารณาฟังก์ชันเคอร์เนลที่เป็นฟังก์ชันเกาส์เซียน (Gaussian kernel function) ซึ่งมีคุณลักษณะของฟังก์ชันเคอร์เนล (Profile) เป็นดังสมการ 2.31

$$k(x) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}x^T x} \quad (2.31)$$

สามารถประมาณความหนาแน่นของข้อมูลและประมาณความหนาแน่นกราดิเยนของข้อมูลในบริเวณที่พิจารณา (Search window: S) ดังสมการที่ 2.32 และ 2.33 ตามลำดับ

$$\begin{aligned} P_n(x) &= \frac{1}{nh} \sum_{s \in S} k\left(\frac{x-s}{h}\right) \\ &= \frac{1}{nh} \sum_{s \in S} \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2h^2}(x-s)^T(x-s)} \end{aligned} \quad (2.32)$$

$$\begin{aligned} \nabla P_n(x) &= \frac{1}{nh} \sum_{s \in S} \frac{(s-x)}{(2\pi)^{n/2}} e^{-\frac{1}{2h^2}(x-s)^T(x-s)} \\ &= \frac{1}{nh} \sum_{s \in S} (s-x) k\left(\frac{x-s}{h}\right) \end{aligned} \quad (2.33)$$

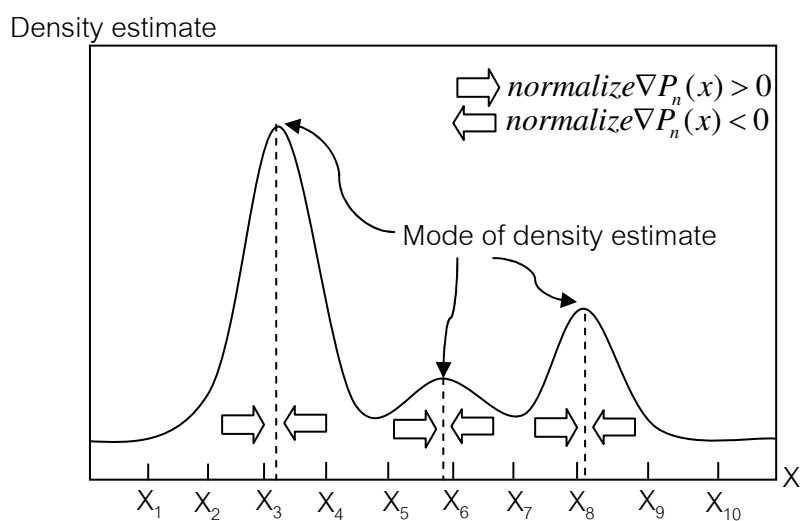
จากนั้นจะทำการนอร์มอลไลซ์ฟังก์ชันความหนาแน่นกราดิเยน $normalize \nabla P_n(x)$ ดังกล่าวได้ดังสมการที่ 2.34

$$\begin{aligned}
 \text{normalize} \nabla P_n(x) &= \frac{\nabla P_n(x)}{P_n(x)} \\
 &= \frac{1}{h^2} \frac{\sum_{s \in S} (s-x) k\left(\frac{x-s}{h}\right)}{\sum_{s \in S} k\left(\frac{x-s}{h}\right)} \\
 &= -\frac{1}{h^2} \left[x - \frac{\sum_{s \in S} s k\left(\frac{x-s}{h}\right)}{\sum_{s \in S} k\left(\frac{x-s}{h}\right)} \right]
 \end{aligned} \tag{2.34}$$

ซึ่งพจน์สุดท้ายของสมการจะเรียกว่า ค่าเฉลี่ยตัวอย่าง (Sample mean) ของบริเวณที่พิจารณา

$$m(x) = \frac{\sum_{s \in S} s k\left(\frac{x-s}{h}\right)}{\sum_{s \in S} k\left(\frac{x-s}{h}\right)} \tag{2.35}$$

ถ้าความหนาแน่นกราฟเดียวนั้นในบริเวณที่พิจารณามีค่าเป็นศูนย์ ค่าเฉลี่ยในบริเวณนั้นจะมีค่าเป็นศูนย์ด้วย จะสามารถจัดกลุ่มข้อมูล โดยพิจารณาการเปลี่ยนตำแหน่งดังสมการและแสดงทิศทางการย้ายตำแหน่งจุดศูนย์กลางการคำนวณได้ดังรูป 2.18



รูปที่ 2.19 แสดงทิศทางการย้ายตำแหน่งจุดศูนย์กลางการคำนวณ

พิจารณาจากรูปที่ 2.18 การเปลี่ยนตำแหน่งของการคำนวณในแต่ละรอบจะมีทิศทางเปลี่ยนตามทิศของลูกศรพิจารณา x_i ค่า $normalize \nabla P_n(x) > 0$ จากนั้นในรอบการคำนวณต่อไป x_i จะมีค่าเพิ่มขึ้นซึ่งทำให้การคำนวณในแต่ละรอบจะย้ายตำแหน่งมาใกล้กับจุดยอด (mode) หรือจุดศูนย์กลางของกลุ่มข้อมูลมากขึ้น

อีกผลหนึ่งที่ใช้ $normalize \nabla P_n(x)$ ในการย้ายตำแหน่ง เนื่องจากจากในบริเวณที่มีข้อมูลมีความหนาแน่นน้อย ($P_n(x)$ น้อย) $normalize \nabla P_n(x) = \frac{\nabla P_n(x)}{P_n(x)} > \nabla P_n(x)$ ทำให้ขนาดของการเปลี่ยนแปลงตำแหน่งมีค่ามากขึ้น การเคลื่อนที่เข้าหาจุดศูนย์กลางจะเร็วกว่าการใช้ $\nabla P_n(x)$ เพียงอย่างเดียวและในกรณีที่ไม่พิจารณา h จะได้ว่า $x_n^{t+1} = x_n^t + (m(x) - x_n^t) = m(x)$ ซึ่งเป็นที่มาของขั้นตอนวิธีการย้ายตามค่าเฉลี่ยดังกล่าว

และจากงานวิจัยของ chang ได้ศึกษาขั้นตอนวิธีการย้ายตามค่าเฉลี่ย เพิ่มเติมทำให้เป็นที่รู้จักและมีการนำมาใช้งานอย่างกว้างขวางมากขึ้น สมมุติมีข้อมูลจำนวน n ชุด ประกอบด้วย x_1, x_2, \dots, x_n ค่าเฉลี่ยของข้อมูลรอบจุด x ใดๆจะสามารถแทนได้ด้วยสมการที่ 2.36 ซึ่งอ้างอิงมาจากค่าเฉลี่ยตัวอย่างดังสมการที่ 2.35

$$m(x) = \frac{\sum_{s \in S} k(s-x)w(s)s}{\sum_{s \in S} k(s-x)w(s)} \quad (2.36)$$

โดยที่ $m(x)$ คือ ค่าเฉลี่ยของข้อมูล (Sample mean) ณ จุด x
 $w(s)$ คือ ค่าน้ำหนัก (Weight) ณ ตำแหน่ง s
 $k(x)$ คือ ฟังก์ชันเคอร์เนลที่ใช้ในการคำนวณ

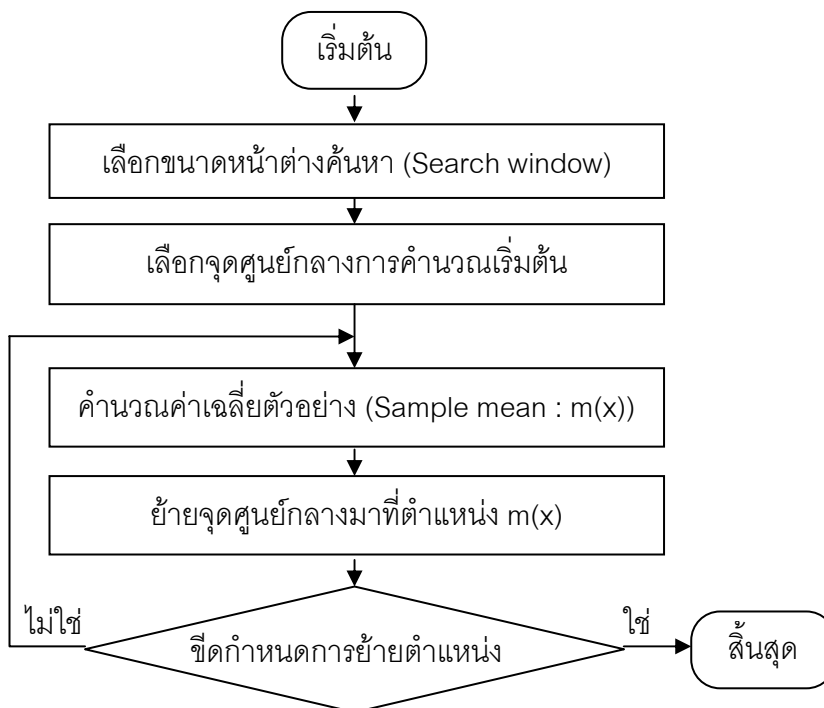
ผลต่างระหว่าง $m(x)-x$ จะพจน์ที่จะถูกเรียกว่า mean-shift และการคำนวณค่าเฉลี่ยตัวอย่าง แล้วย้ายจุด x มายังจุดค่าเฉลี่ยตัวอย่าง จนกระทั่งไม่มีการย้ายค่า x อีก ซึ่งเรียกว่าขั้นตอนการย้ายตามค่าเฉลี่ยนั่นเอง

ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยสามารถสรุปขั้นตอนการดำเนินการได้ดังต่อไปนี้

1. เลือกจุดศูนย์กลางการคำนวณเริ่มต้นที่ตำแหน่ง x
2. คำนวณหา $m(x)$
3. กำหนดให้ $m(x)$ เป็นจุดศูนย์กลางการคำนวณใหม่ที่ตำแหน่ง x^{t+1}

4. จากนั้นทำการคำนวณตามขั้นตอนที่ 2 และ 3 จนจุดศูนย์กลางการคำนวณไม่ย้ายหรือมีการย้ายน้อยกว่าขีดเริ่มเปลี่ยนที่กำหนด (Threshold)

แสดงแผนภาพแสดงขั้นตอนวิธีการย้ายตามค่าเฉลี่ยได้ดังรูป 2.19



รูปที่ 2.20 แสดงแผนภาพแสดงขั้นตอนวิธีการย้ายตามค่าเฉลี่ย

2.8 การวัดประสิทธิภาพด้วยฟังก์ชันความเหมือนของ Bhattacharyya

การประมวลผลร่วมกันของระบบกล้องวิทัศน์เพื่อติดตามบุคคลคนเดียวกันอย่างต่อเนื่อง นั้น จำเป็นต้องมีการระบวนการวัดความเหมือนของข้อมูลคุณลักษณะสำคัญว่ามีความใกล้เคียงหรือเหมือนกันหรือไม่ ไม่ว่าคุณลักษณะ (Feature) ใด ที่จะนำมาเป็นตัวเปรียบเทียบ เช่น ลักษณะสี รูปร่าง และขอบ เป็นต้น โดยวิธีการหนึ่งที่เป็นที่นิยมใช้สำหรับการเปรียบเทียบความเหมือน คือ กระบวนการวัดความคล้าย (Similarity measure) โดยอาศัยวิธีการวัดระยะทางระหว่างกลุ่มการกระจายตัวของข้อมูล 2 ชุดที่สนใจ ซึ่งมีวิธีการในการหาระยะทางนี้ได้หลายแบบ ในวิทยานิพนธ์นี้จะอาศัยเทคนิคการเปรียบเทียบความเหมือนฮิสโทแกรมด้วยฟังก์ชันความเหมือนของ Bhattacharyya (Bhattacharyya similarity function) ซึ่งประกอบไปด้วยตัวแปรสองส่วน ได้แก่ ระยะ Bhattacharyya (Bhattacharyya distance) และสัมประสิทธิ์ของ Bhattacharyya (Bhattacharyya coefficient)

ฟังก์ชันความเหมือนของ Bhattacharyya เป็นรูปแบบการเปรียบเทียบลักษณะหนึ่งในกระบวนการวัดความคล้าย ระหว่างชุดการกระจายตัวของข้อมูล 2 ชุด ซึ่งได้ถูกนำมาประยุกต์ใช้ในกระบวนการทางการประมวลผลภาพสำหรับการเปรียบเทียบฮิสโทแกรมของสองกลุ่มข้อมูลที่ไม่จำเป็นต้องเป็นค่าฮิสโทแกรมของทั้งภาพ อาจเป็นแค่บริเวณใดบริเวณหนึ่งที่สนใจเท่านั้น และไม่จำเป็นต้องทั้งสองบริเวณจะต้องมีขนาดเท่ากัน (ในหน่วยจุดภาพ) จึงเป็นข้อดีที่สามารถรองรับการเปรียบเทียบวัตถุที่มีการเปลี่ยนแปลงรูปร่างได้ในลักษณะการเคลื่อนที่เปลี่ยนมุมมองและการเปลี่ยนแปลงขนาด สมมติให้ \hat{p} เป็นฮิสโทแกรมสีของแบบจำลองเป้าหมายและ \hat{q} เป็นฮิสโทแกรมสีของแบบจำลองรับเข้าที่ต้องการนำมาเปรียบเทียบ โดยที่ $\hat{p}_u = \{p_u\}_{u=1,2,\dots,m}$ (เมื่อ $\sum_{u=1}^m p_u = 1$) และ $\hat{q}_u = \{q_u\}_{u=1,2,\dots,m}$ (เมื่อ $\sum_{u=1}^m q_u = 1$) เมื่อ m เป็นจำนวนถังข้อมูลของฮิสโทแกรม (Bin histogram) เปรียบเทียบ เช่น ถ้าในกรณีเลือกใช้ขนาดขององค์ประกอบสีแต่ละองค์ประกอบมีขนาด 4 บิต ฉะนั้นจะได้ขนาดถังสีเท่ากับ 16 ถัง ดังนั้นสามารถคำนวณหาสัมประสิทธิ์ Bhattacharyya (ρ) ได้ดังสมการที่ 2.37

$$\rho[\hat{p}, \hat{q}] = \sum_{u=1}^m \sqrt{p_u \times q_u} \quad (2.37)$$

อธิบายความหมายทางเรขาคณิตตามสมการที่ 2.37 จะแสดงถึงค่าโคไซน์ (Cosine) ของมุมระหว่างค่าถังข้อมูลทั้ง m ถังมิติในรูปเวกเตอร์หนึ่งหน่วย $(\sqrt{p_1}, \dots, \sqrt{p_m})^T$ และ $(\sqrt{q_1}, \dots, \sqrt{q_m})^T$ และจากสมการที่ 2.37 จะสามารถหาค่าระยะ Bhattacharyya ที่เป็นระยะระหว่างสองกลุ่มของการกระจายตัวของข้อมูล (d_{BH}) ดังสมการที่ 2.38

$$d_{BH} = \sqrt{1 - \rho[\hat{p}, \hat{q}]} \quad (2.38)$$

ข้อดีของการเปรียบเทียบวัดความคล้ายด้วยฟังก์ชันความเหมือนของ Bhattacharyya คือ สามารถรองรับการเปรียบเทียบชุดข้อมูล 2 ชุดที่ไม่จำเป็นต้องมีขนาดหรือจำนวนข้อมูลที่เท่ากันก็ได้

บทที่ 3

โครงสร้างและระบบติดตามบุคคลอย่างต่อเนื่องที่นำเสนอ

ในบทที่ 3 นี้จะกล่าวถึงโครงสร้างและระบบติดตามบุคคลอย่างต่อเนื่องที่นำเสนอ โดยอาศัยพื้นฐานของขั้นตอนวิธีการย้ายตามค่าเฉลี่ย ด้วยกล้องวิดิทัศน์สองตัวที่ถูกติดตั้งอยู่ในลักษณะมุมมองที่ต่างกันให้ทำงานร่วมกันเพื่อติดตามบุคคลเป้าหมาย ซึ่งจากบทที่ 1 ที่กล่าวมาแล้วนั้นจะเห็นได้ว่าการทำงานของระบบกล้องวิดิทัศน์ส่วนใหญ่ในอดีตจะใช้งานเพียงกล้องวิดิทัศน์ประเภทอยู่กับที่ [1,2] ซึ่งมีข้อจำกัดมากในการทำงานในเวลาจริงหรือการใช้งานได้อย่างเต็มประสิทธิภาพ ในระยะหลังๆ กล้องวิดิทัศน์ประเภท PTZ จึงได้ถูกนำมาประยุกต์เข้ากับการทำงานในระบบกล้องวิดิทัศน์หลายตัว และเนื่องจากกล้องวิดิทัศน์ประเภท PTZ นั้นจะสามารถรองรับรูปแบบของการรับคำสั่งพารามิเตอร์ของการส่าย ก้มเงย และซูมได้ จึงเหมาะสมกับการนำมาประยุกต์ใช้งานร่วมกับงานทางด้านคอมพิวเตอร์วิทัศน์หลากหลายประเภท เช่น การตรวจหาใบหน้าบุคคลในระบบรักษาความปลอดภัยเพื่อให้ได้รายละเอียดของใบหน้าบุคคลที่สนใจด้วยการส่าย ก้มเงยและซูม หรือใช้ในการติดตามบุคคลที่สนใจได้ในเวลาจริงเพื่อให้ติดตามบุคคลได้อย่างต่อเนื่องและมีประสิทธิภาพดังที่วิทยานิพนธ์ฉบับนี้ได้นำเสนอ เป็นต้น

วิทยานิพนธ์เล่มนี้จึงขอเสนอระบบการติดตามบุคคลที่สนใจอย่างต่อเนื่องโดยใช้กล้องวิดิทัศน์ PTZ 2 ตัว ในระบบให้ประมวลผลร่วมกัน สำหรับนำไปประยุกต์ใช้งานด้านการรักษาความปลอดภัยในบริเวณภายในอาคาร ห้องปฏิบัติการหรือสถานที่ต่างๆ ที่เสี่ยงต่อการเกิดอาชญากรรม เป็นต้น ขั้นตอนวิธีการติดตามบุคคลในอดีตได้ถูกนำเสนอไว้หลากหลายวิธี ซึ่งสามารถแบ่งออกเป็นสองประเภท คือ ขั้นตอนวิธีการติดตามด้วยวิธีการทางความน่าจะเป็น (Probability Method) และขั้นตอนวิธีการติดตามด้วยวิธีการตรวจหา (Deterministic Method) ซึ่งจากงานวิจัย [19] กล่าวว่าขั้นตอนวิธีที่อาศัยทฤษฎีความน่าจะเป็นนั้นจะมีการคำนวณที่ซับซ้อนเป็นอย่างมากเมื่อถูกนำมาใช้ในสถานการณ์ เช่น เมื่อภาพมีสัญญาณรบกวนมาก แสงสว่างมีการเปลี่ยนแปลง หรือวัตถุที่ถูกติดตามอยู่นั้นมีการเคลื่อนที่ที่ผิดปกติ เช่น เมื่อใช้ขั้นตอนวิธีติดตามด้วยตัวกรองแบบคาร์มาล ซึ่งจะอาศัยพารามิเตอร์ทางด้านความเร็ว ถ้าวัตถุสองวัตถุอยู่ใกล้หรือไกลกล้องวิดิทัศน์ที่ต่างกัน ความเร็วของวัตถุที่ถูกติดตามจะต่างกันทำให้ในการติดตามวัตถุ วัตถุเดียวมีความผิดพลาดได้ เป็นต้น จากงานวิจัยของ [22] จึงให้ความสำคัญกับขั้นตอนวิธีการสำรวจ เนื่องจากการคำนวณที่ไม่ซับซ้อน อาศัยเพียงคุณลักษณะสำคัญบางอย่างที่สามารถนำมาใช้ประมวลผลอย่างง่าย เช่น ข้อมูลค่าสีของวัตถุ ฮิสโตแกรมของข้อมูลสี ความหนาแน่นของ

ข้อมูล เป็นต้น จึงสามารถทำงานได้อย่างรวดเร็ว อีกทั้งในบางสถานการณ์ยังสามารถแก้ปัญหาการบังคับบุคคลหรือวัตถุที่สนใจได้แม่นยำถูกต้องมากกว่าการใช้ขั้นตอนวิธีการใช้ความน่าจะเป็น

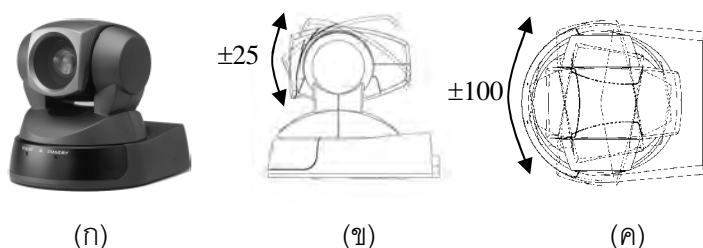
ขั้นตอนวิธีการติดตามบุคคลที่นำเสนอ นั้น คือ ขั้นตอนวิธีการย้ายตามค่าเฉลี่ย [17] ซึ่งในแต่เดิมจะใช้เพียงคุณลักษณะสำคัญทางสีในปริภูมิระดับสีเทาเพียงอย่างเดียวเท่านั้น ซึ่งไม่เหมาะกับการนำมาใช้ในระบบกล้องวิดีโอที่ที่มีการติดตั้งในมุมมองที่ต่างกัน และการใช้งานจริงวิทยานิพนธ์นี้จึงประยุกต์ใช้คุณลักษณะสำคัญสองประการ คือ ข้อมูลสีในปริภูมิสี HSV (แซลแนล H) และข้อมูลรูปร่างโดยดูตามเส้นขอบ อีกทั้งเพิ่มประสิทธิภาพในการส่งข้อมูลสีระหว่างกล้องวิดีโอทั้งสองในระบบที่มีการติดตั้งต่างมุมมองกันด้วยวิธีการเทียบมาตรฐานทางสี

3.1 รูปแบบการทำงานของระบบที่นำเสนอ

รูปแบบการทำงานของระบบที่นำเสนอ นั้นจะเป็นการทำงานร่วมกันเพื่อติดตามบุคคลที่สนใจด้วยกล้องวิดีโอ PTZ 2 ตัวอย่างต่อเนื่อง ซึ่งจะเป็นการติดตามบุคคลเป้าหมายคนเดียวจากกล้องวิดีโอตัวหนึ่งไปยังกล้องวิดีโออีกตัวหนึ่ง จะทำให้ระบบสอดคล้องและรักษาความปลอดภัยสามารถติดตามการเคลื่อนที่ของบุคคลเป้าหมายที่สนใจได้อย่างต่อเนื่องและมีประสิทธิภาพ กล่าวคือ เมื่อบุคคลเป้าหมายเดินเข้ามาในระบบภายในอาณาบริเวณการรับภาพของกล้องวิดีโอตัวหนึ่ง จะทำการเก็บข้อมูลของบุคคลนั้นในรูปแบบของข้อมูลคุณลักษณะของสีและข้อมูลคุณลักษณะของขอบ จากนั้นเมื่อบุคคลเป้าหมายเคลื่อนที่พ้นจากขอบเขตของกล้องวิดีโอตัวที่หนึ่งแล้ว จะส่งข้อมูลทั้งสองนี้ไปยังกล้องวิดีโออีกตัว จากนั้นกล้องวิดีโอตัวที่สองนี้จะทำหน้าที่ติดตามบุคคลเป้าหมายบุคคลเดิมอย่างต่อเนื่อง ซึ่งแนวความคิดต่อไป คือ การส่งต่อพารามิเตอร์ตำแหน่งของบุคคลเป้าหมายผ่านกล้องวิดีโอ PTZ ทั้งสองตัว เพื่อควบคุมให้กล้องวิดีโอทั้งสองตัวเคลื่อนที่ตามบุคคลเป้าหมายได้ แต่ปัญหาหนึ่งที่เกิดขึ้นจากระบบที่นำเสนอ คือ ระบบที่นำเสนอจะติดตั้งกล้องวิดีโอทั้งสองตัวในบริเวณที่ต่างกัน ปัญหาอันเนื่องมาจากความสว่างหรือคุณลักษณะทางกายภาพของกล้องวิดีโอที่ต่างกัน อาจจะทำให้ข้อมูลสีเดียวกันที่รับได้จากกล้องวิดีโอทั้งสองตัวนั้นอ่านค่าได้ต่างกัน ซึ่งมีผลต่อการส่งข้อมูลสีไปยังกล้องวิดีโอตัวที่สองเพื่อทำหน้าที่รับการติดตามบุคคลต่อ วิทยานิพนธ์เล่มนี้จึงเพิ่มวิธีการเทียบมาตรฐานทางสีให้กับกล้องวิดีโอทั้งสองตัวในระบบ เพื่อแก้ปัญหาความแตกต่างของข้อมูลสีดังกล่าวนี้ด้วย

3.1.1 กล้องวิดีโอ PTZ

กล้องวิดีโอ PTZ ที่ใช้ในระบบ คือ กล้องวิดีโอ SONY รุ่น EVID-100P ดังรูปที่ 3.1 มีคุณสมบัติที่สำคัญเกี่ยวกับการสาย การก้มเงยและการซูม ดังนี้



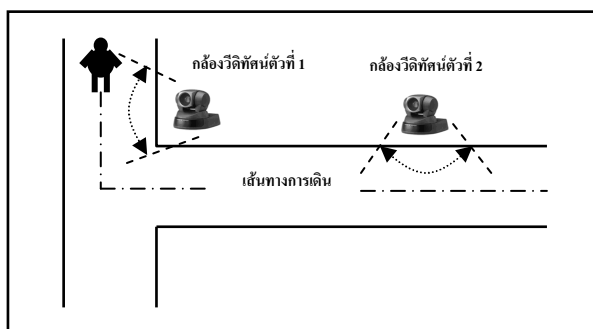
รูปที่ 3.1 กล้องวิดีโอในการทดลอง(ก) กล้องวิดีโอ EVID-100 (ข) ลักษณะการก้มเงย
(ค) ลักษณะการสาย

- ในการควบคุมการเคลื่อนที่ของกล้อง ต้องส่งคำสั่ง (Command) ควบคุมผ่านทางสายสัญญาณ RS-232
- สามารถซูมเชิงแสง (Optical zoom) ได้สูงสุด 10 เท่า
- แนวการสายสูงสุด ± 100 องศา
- แนวการก้มเงยสูงสุด ± 25 องศา
- ความเร็วสูงสุดในการเคลื่อนที่แบบสาย 300 องศา/วินาที
- ความเร็วสูงสุดในการเคลื่อนที่แบบก้มเงย 125 องศา/วินาที
- รองรับการเชื่อมต่อสัญญาณระหว่างกล้องกับส่วนประมวลผลได้ 2 รูปแบบ คือ แบบสายสัญญาณภาพรวม (Composite video cable) และสายสัญญาณวิดีโอแบบเอสวีดีโอ (S-video cable)

3.1.2 ลักษณะโครงสร้างการจัดวางตำแหน่งกล้องวิดีโอ PTZ ทั้ง 2 ตัว

การจัดวางระบบกล้องวิดีโอที่นำเสนอจะเป็นดังรูปที่ 3.2 จะเน้นที่กล้องวิดีโอ PTZ ทั้งสองควรจะอยู่ห่างกันและต่างมุมมองการรับภาพกัน หรือเรียกอย่างง่ายว่าไม่มีมุมมองการรับภาพที่เหลื่อมกัน (Non-overlapped view) ดังนั้นการจัดวางกล้องวิดีโอจึงไม่ต้องการกำหนดตำแหน่งในการติดตั้งมาก (แต่เนื่องจากข้อจำกัดในการทดลอง ในบริเวณห้องปฏิบัติการจึงจำเป็นต้องจัดวางในลักษณะดังรูปที่ 3.2 อาจะวางในมุมตรงข้าม หรืออาจอยู่ต่างบริเวณ

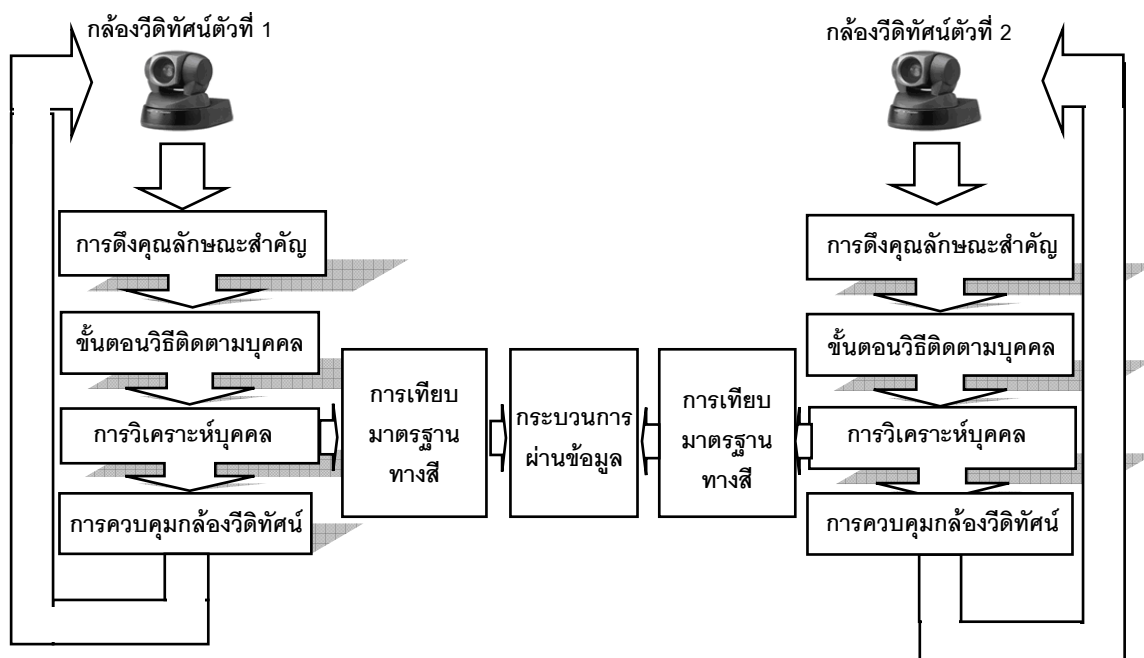
กันก็ได้ ตำแหน่งการวางกล้องวีดิทัศน์ในการทดลองนี้จะสูงจากพื้นประมาณ 1.5 เมตร เท่ากันทั้งสองตัว และ กล้องวีดิทัศน์ทั้งสองอยู่ห่างกัน ประมาณ 2-3 เมตร



รูปที่ 3.2 แสดงการติดตั้งระบบการติดตามบุคคลที่สนใจด้วยกล้องวีดิทัศน์ 2 ตัว

3.1.3 ไคอะแกรมการทำงานโดยรวม

สถาปัตยกรรมโดยรวมของโครงร่างวิทยานิพนธ์ที่นำเสนอนี้แสดงดังรูปที่ 3.3 ซึ่งประกอบด้วยกล้องวีดิทัศน์ประเภท PTZ จำนวนสองตัวในระบบ ซึ่งมีวัตถุประสงค์เพื่อใช้กล้องวีดิทัศน์ทั้งสองตัวนี้ในการติดตามบุคคลที่มีการเคลื่อนที่ผ่านระบบกล้องวีดิทัศน์ โดยระบบที่เสนอจะแบ่งการทำงานร่วมกันออกเป็น 4 ขั้นตอนใหญ่ๆ ได้แก่ ส่วนของการดึงคุณลักษณะสำคัญของบุคคลที่สนใจจากกล้องวีดิทัศน์ ส่วนของขั้นตอนวิธีในการติดตามบุคคล ส่วนของการวิเคราะห์บุคคล ซึ่งในส่วนของกระบวนการระบุบุคคลนี้จะประกอบไปด้วยสองส่วนย่อยคือ ขั้นตอนการเทียบมาตรฐานทางสีและการส่งต่อข้อมูลระหว่างกล้องวีดิทัศน์ และในที่สุดท้ายคือส่วนของการควบคุมกล้องวีดิทัศน์ให้ติดตามบุคคลเป้าหมาย



รูปที่ 3.3 ไดอะแกรมโดยรวมสำหรับการติดตามบุคคลอย่างต่อเนื่องที่นำเสนอ

การประมวลผลของกล้องวิดีโอ PTZ ตัวที่หนึ่งนั้น เมื่อบุคคลเป้าหมายเข้ามายังมุมมองกล้องวิดีโอตัวที่หนึ่งแล้ว จะทำการดึงคุณลักษณะสำคัญและคำนวณฟังก์ชันความหนาแน่นของข้อมูลคุณลักษณะสำคัญ (จะกล่าวในหัวข้อ 3.3) และส่งต่อค่าฟังก์ชันความหนาแน่นของข้อมูลคุณลักษณะสำคัญนี้ ไปยังกล้องวิดีโอตัวที่สองทันที จึงทำให้ขณะที่บุคคลเป้าหมายยังไม่เคลื่อนเข้าสู่มุมมองของกล้องตัวที่สอง กล้องวิดีโอตัวที่สองจะคงสถานะการเคลื่อนที่ไว้ก่อน นั่นคือไม่มีการส่าย ก้มเงยและซูม จนกระทั่งมีบุคคลเป้าหมายเข้ามาในมุมมองแล้วสามารถติดตามบุคคลดังกล่าวได้ จึงสั่งให้กล้องวิดีโอตัวที่สองมีการเคลื่อนที่เพื่อติดตามบุคคลเป้าหมายคนเดิมอย่างต่อเนื่องต่อไป

3.2 วิธีการเทียบมาตรฐานทางสีที่นำเสนอ (Camera Color Calibration)

สิ่งสำคัญของระบบการทำงานร่วมกันระหว่างกล้องวิดีโอตั้งแต่สองตัวขึ้นไป คือ ขั้นตอนของการประมวลผลภาพจากกล้องวิดีโอแต่ละตัวและการส่งต่อข้อมูลระหว่างกันให้ถูกต้องแม่นยำที่สุด โดยระบบการติดตามบุคคลด้วยระบบกล้องวิดีโอที่นำเสนอนี้ มีการติดตั้งกล้องวิดีโอแต่ละตัวในบริเวณที่ต่างกัน หรือคุณสมบัติการรับภาพและแสดงภาพของกล้องวิดีโอแต่ละตัวในระบบให้ผลที่ต่างกัน จึงทำให้การนำข้อมูลสีมาใช้ในการประมวลผลระหว่างกันมีความผิดพลาดตามไปด้วย ดังนั้นในวิทยานิพนธ์ฉบับนี้ จึงมีแนวความคิดที่จะแก้ปัญหาคือ

แตกต่างของค่าสีที่แสดงผลจากกล้องวิดีโอทีละตัวดังกล่าวนั้น ด้วยการอาศัยวิธีการเทียบมาตรฐานทางสี ให้กับกล้องวิดีโอทีละตัว ก่อนที่จะนำข้อมูลหลังจากการเทียบมาตรฐานสีนั้นไปประมวลผลเพื่อติดตามบุคคลอย่างต่อเนื่องต่อไป

ขั้นตอนของวิธีการเทียบมาตรฐานทางสีที่น่าเสนอ มีแนวคิดโดยการหาเมทริกซ์ที่เหมาะสมเมทริกซ์หนึ่งเป็นเมทริกซ์ในการเทียบมาตรฐานทางสี (Color Calibration matrix) ให้กับกล้องวิดีโอทีละตัวในระบบ ซึ่งอ้างอิงมาจากกล้องวิดีโอตัวใดตัวหนึ่งในระบบ และวิทยานิพนธ์นี้ได้นำเสนอวิธีการหาค่าเมทริกซ์ของการเทียบมาตรฐานทางสีของกล้องวิดีโอทีละตัวนี้ ด้วยวิธีการของสมการเชิงเส้นแบบถดถอย (Linear Regression Equations) และอาศัยข้อมูลสีที่ใช้อ้างอิง จากตารางค่าสีอ้างอิงของการเทียบมาตรฐานสี (Color checker board) ดังรูปที่ 3.4

1	2	3	4	5	6
7	8	9	10	11	1
1	1	1	1	1	1
1	2	2	2	2	2

รูปที่ 3.4 แสดงรูปตารางอ้างอิงการเทียบมาตรฐานสีที่ใช้เทียบมาตรฐานสี

จากตารางค่าสีอ้างอิงของการเทียบมาตรฐานสี ดังรูปที่ 3.4 ประกอบด้วย 24 สีอ้างอิง ซึ่งเป็นสีมาตรฐานและเป็นสีที่จำเป็นในการทำงานส่วนใหญ่ (อ้างอิงค่าสีมาจากตารางตรวจสอบค่าสีจากโรงพิมพ์หรือร้านถ่ายภาพ) ในปริภูมิสี RGB ซึ่ง 24 สีที่นำมาอ้างอิงนี้ มีข้อมูลค่าสีแสดงดังตารางที่ 3.1

วิธีการเตรียมระบบในการเทียบมาตรฐานทางสีนั้นทำได้โดย เก็บภาพตารางค่าสีอ้างอิงของการเทียบมาตรฐานสี จากกล้องวิดีโอทีละตัวในระบบและนำภาพที่ได้เข้าขั้นตอนของวิธีการเทียบมาตรฐานทางสีที่น่าเสนอ โดยใช้สมมุติฐานที่ว่ากล้องแต่ละตัวในระบบมีความสัมพันธ์แบบเชิงเส้นต่อกันในระบบ

ตารางที่ 3.1 แสดงสีและข้อมูลสี ในปริภูมิสี RGB ของตารางสีอ้างอิงที่ใช้ในการทดลอง

สีจากตารางอ้างอิง	ข้อมูลสี		
	R	G	B
1.DARK SKIN	131	69	56
2.LIGHT SKIN	241	185	170
3.BLUE SKY	119	161	211
4.FOLIAGE	91	140	48
5.BLUE FLOWER	176	168	227
6.BLUISH GREEN	172	250	234
7.ORANGE	237	127	12
8.PURPLISH BLUE	71	107	217
9.MODERATE RED	248	85	132
10.PURPLE	101	58	140
11.YELLOW GREEN	196	238	61
12.ORANGE YELLOW	252	215	12
13.RED	255	0	0
14.GREEN	0	255	0
15.BLUE	0	0	255
16.YELLOW	255	255	0
17.MAGENTA	255	0	255
18.CYAN	0	255	255
19.WHITE	255	255	255
20.NEUTRAL 8	200	200	200
21.NEUTRAL 6.5	160	160	160
22.NEUTRAL 5	120	120	120
23.NEUTRAL 3.5	80	80	80
24.BLACK	0	0	0

จึงนำวิธีการของสมการเชิงเส้นถดถอยมาประยุกต์ใช้ นั่นคือต้องหาค่าความสัมพันธ์ระหว่างกันระหว่างกล่องวัดทัศนสองตัว จึงสามารถตั้งสมการเชิงเส้นเริ่มต้นได้ดังสมการที่ 3.1

$$C_1 = W C_2 \quad (3.1)$$

โดยที่ C_2 คือ ข้อมูลทางสีของกล้องตัวที่ 1
 W คือ เวทพารามิเตอร์ระหว่างกล้องทั้งสองตัว
 C_1 คือ ข้อมูลทางสีของกล้องตัวที่ 2

เมื่อรับภาพจากกล้องวีดิทัศน์ทั้งสองตัวแล้ว สามารถตั้งสมการของแต่ละสีข้างอิง ด้วยสมการถดถอยเชิงเส้น แยกตามแชนแนลสี R, G และ B ดังสมการที่ 3.2, 3.3 และ 3.4 ตามลำดับ โดยที่ i มีค่าตั้งแต่ 1 ถึง 24 ซึ่งเป็นจำนวนของสีที่ใช้ข้างอิงในตารางค่าสีข้างอิงของการเทียบมาตรฐานสี

$$R_1^i = \beta_{0R} + \beta_{1R}R_2^i + \beta_{2R}G_2^i + \beta_{3R}B_2^i \quad (3.2)$$

$$G_1^i = \beta_{0G} + \beta_{1G}R_2^i + \beta_{2G}G_2^i + \beta_{3G}B_2^i \quad (3.3)$$

$$B_1^i = \beta_{0B} + \beta_{1B}R_2^i + \beta_{2B}G_2^i + \beta_{3B}B_2^i \quad (3.4)$$

หรือสามารถเขียนในรูปสมการทั้งสามดังกล่าวในรูปของเมทริกซ์ได้ดังสมการที่ 3.5

$$\begin{bmatrix} R_1^i \\ G_1^i \\ B_1^i \\ 1 \end{bmatrix} = \begin{bmatrix} \beta_{1R} & \beta_{2R} & \beta_{3R} & \beta_{0R} \\ \beta_{1G} & \beta_{2G} & \beta_{3G} & \beta_{0G} \\ \beta_{1B} & \beta_{2B} & \beta_{3B} & \beta_{0B} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_2^i \\ G_2^i \\ B_2^i \\ 1 \end{bmatrix} \quad (3.5)$$

โดยที่ $\begin{bmatrix} R_1^i \\ G_1^i \\ B_1^i \\ 1 \end{bmatrix}$ คือ เมทริกซ์ข้อมูลสีของสีที่ i ในตารางค่าสีข้างอิงของการเทียบมาตรฐานสีจากกล้องวีดิทัศน์ตัวที่ 1

$\begin{bmatrix} R_2^i \\ G_2^i \\ B_2^j \\ 1 \end{bmatrix}$ คือ เมทริกซ์ข้อมูลสีของสีที่ i ในตารางค่าสีข้างอิงของการเทียบมาตรฐานสีจากกล้องวีดิทัศน์ตัวที่ 2

และ

$\begin{bmatrix} \beta_{1R} & \beta_{2R} & \beta_{3R} & \beta_{0R} \\ \beta_{1G} & \beta_{2G} & \beta_{3G} & \beta_{0G} \\ \beta_{1B} & \beta_{2B} & \beta_{3B} & \beta_{0B} \\ 0 & 0 & 0 & 1 \end{bmatrix}$ คือ เมทริกซ์ของการเทียบมาตรฐานทางสีระหว่างกล้องวีดิทัศน์สองตัว

จากตารางค่าสีอ้างอิงของการเทียบมาตรฐานสีที่ใช้ในการทดลอง มีทั้งหมด 24 สี ดังนั้น จึงได้ทั้งหมด 24 สมการเมทริกซ์ สามารถนำมาจัดเรียงให้อยู่ในรูปแบบเมทริกซ์เดียวได้ดังสมการที่ 3.6

$$\begin{bmatrix} R_1^1 & R_1^2 & \dots & R_1^{24} \\ G_1^1 & G_1^2 & \dots & G_1^{24} \\ B_1^1 & B_1^2 & \dots & B_1^{24} \\ 1 & 1 & \dots & 1 \end{bmatrix}_{4 \times 24} = \begin{bmatrix} \beta_{1R} & \beta_{2R} & \beta_{3R} & \beta_{0R} \\ \beta_{1G} & \beta_{2G} & \beta_{3G} & \beta_{0G} \\ \beta_{1B} & \beta_{2B} & \beta_{3B} & \beta_{0B} \\ 0 & 0 & 0 & 1 \end{bmatrix}_{4 \times 4} \begin{bmatrix} R_2^1 & R_2^2 & \dots & R_2^{24} \\ G_2^1 & G_2^2 & \dots & G_2^{24} \\ B_2^1 & B_2^2 & \dots & B_2^{24} \\ 1 & 1 & \dots & 1 \end{bmatrix}_{4 \times 24} \quad (3.6)$$

และจากสมการที่ 3.6 กำหนดให้

$$\begin{aligned} C_1 &= \begin{bmatrix} R_1^1 & R_1^2 & \dots & R_1^{24} \\ G_1^1 & G_1^2 & \dots & G_1^{24} \\ B_1^1 & B_1^2 & \dots & B_1^{24} \\ 1 & 1 & \dots & 1 \end{bmatrix}_{4 \times 24} \\ C_2 &= \begin{bmatrix} R_1^1 & R_1^2 & \dots & R_1^{24} \\ G_1^1 & G_1^2 & \dots & G_1^{24} \\ B_1^1 & B_1^2 & \dots & B_1^{24} \\ 1 & 1 & \dots & 1 \end{bmatrix}_{4 \times 24} \\ \text{และ} \\ W &= \begin{bmatrix} \beta_{1R} & \beta_{2R} & \beta_{3R} & \beta_{0R} \\ \beta_{1G} & \beta_{2G} & \beta_{3G} & \beta_{0G} \\ \beta_{1B} & \beta_{2B} & \beta_{3B} & \beta_{0B} \\ 0 & 0 & 0 & 1 \end{bmatrix}_{4 \times 4} \end{aligned}$$

ดังนั้น สามารถหาค่า W เพื่อใช้แปลงข้อมูลสีจากกล้องวีดิทัศน์ตัวหนึ่งไปยังกล้องวีดิทัศน์อีกตัวหนึ่งได้ได้จากการหาอินเวอร์สเมทริกซ์ดังสมการที่ 3.7

$$\begin{aligned} C_1 &= WC_2 \\ C_1(C_2)^T &= WC_2(C_2)^T \\ W &= (C_1(C_2)^T)(C_2(C_2)^T)^{-1} \end{aligned} \quad (3.7)$$

3 ขั้นตอนวิธีการติดตามบุคคลที่น่าเสนอ

การประมาณความหนาแน่นของข้อมูล มีพื้นฐานมาจากฟังก์ชันความหนาแน่นที่มีความต่อเนื่อง สามารถประมาณได้โดยสำรวจข้อมูลตัวอย่าง ภายในบริเวณเล็กๆ รอบจุดๆนั้น ปัญหาโดยส่วนใหญ่ของรูปแบบการวิจัยเกี่ยวกับงานคอมพิวเตอร์วิทัศน์ คือ ความน่าจะเป็นของความหนาแน่นของข้อมูล จึงจำเป็นต้องหาวิธีการหรือเทคนิคเพื่อให้ทราบถึงรูปแบบความหนาแน่นของข้อมูลหรือฟังก์ชันความหนาแน่นของข้อมูล และการกำหนดตัวแปรที่จำเป็นต้องใช้ในวิธีการหรือเทคนิคนั้นๆ ควรกำหนดอยู่บนพื้นฐานของข้อมูลที่ให้ ดังนั้นจึงเกิดเทคนิคการประมาณแบบไม่กำหนดตัวแปร [19] และเป็นจุดเริ่มต้นของขั้นตอนวิธีการย้ายตามค่าเฉลี่ย และได้กล่าวถึงขั้นตอนวิธีการย้ายตามค่าเฉลี่ยพื้นฐานไว้แล้วในบทที่ 2

ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยถูกนำมาใช้งานด้านคอมพิวเตอร์วิทัศน์ในหลายๆด้านรวมทั้งด้านการติดตามบุคคลด้วย ในหัวข้อนี้จึงขอเสนอการติดตามบุคคลด้วยหลักการของการย้ายตามค่าเฉลี่ย ซึ่งสิ่งสำคัญของการติดตามบุคคลนั้นคือ ข้อมูลทางคุณลักษณะสำคัญของเป้าหมายบุคคลที่ต้องการติดตาม เนื่องจากคุณลักษณะสำคัญถือเป็นตัวแปรหลักในการทำให้ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยดำเนินไปได้ ซึ่งจะนิยมใช้ข้อมูลคุณลักษณะสำคัญในรูปแบบของฟังก์ชันความหนาแน่น ในระยะแรกมีการนำฟังก์ชันความหนาแน่นของข้อมูลสีในระดับสีเทามาเป็นตัวแทนของบุคคลเป้าหมาย หลังจากเริ่มมีการวิจัยมากขึ้นในปัจจุบัน จึงได้มีการพัฒนาหาฟังก์ชันความหนาแน่นที่ได้จากคุณลักษณะสำคัญอื่นๆ เข้ามาช่วยในขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเพื่อให้การติดตามบุคคลนี้ อย่างมีประสิทธิภาพมากขึ้น ขั้นตอนวิธีการติดตามบุคคลด้วยหลักการย้ายตามค่าเฉลี่ยที่น่าเสนอนี้ประกอบไปด้วยสองขั้นตอนสำคัญ คือ ขั้นตอนการดึงคุณลักษณะสำคัญและการประมาณฟังก์ชันความหนาแน่นของบุคคลเป้าหมาย และขั้นตอนการติดตามบุคคลด้วยขั้นตอนวิธีการย้ายตามค่าเฉลี่ย

3.3.1 ขั้นตอนการดึงคุณลักษณะสำคัญและการประมาณฟังก์ชันความหนาแน่น

ของบุคคลเป้าหมาย

ในหัวข้อนี้จะแสดงถึงขั้นตอนวิธีการดึงคุณลักษณะสำคัญของบุคคลเป้าหมายเพื่อใช้ในการติดตามบุคคลภายในลำดับภาพวิทัศน์ ซึ่งจะทำการดึงข้อมูลแสดงบุคคลเป้าหมายในสองกรณี ได้แก่ การดึงคุณลักษณะสำคัญเพื่อนำไปประมาณฟังก์ชันความหนาแน่นอ้างอิงเป็นบุคคลเป้าหมาย (Target model) และการดึงคุณลักษณะสำคัญเพื่อนำไปประมาณเป็นฟังก์ชันความ

หนาแน่นของบุคคลเป้าหมายที่ต้องการติดตาม (Target candidate) วิทยานิพนธ์ฉบับนี้จะทำการประมาณฟังก์ชันความหนาแน่นจากข้อมูลทางคุณลักษณะสำคัญสองประเภท คือ คุณลักษณะสำคัญทางสีและคุณลักษณะสำคัญทางรูปร่างโดยพิจารณาจากขอบบุคคลเป้าหมาย โดยมีการคำนวณหาฟังก์ชันความหนาแน่นเช่นเดียวกันในสองกรณี และจะถูกแสดงเป็นสมการฟังก์ชันความหนาแน่นอ้างอิงเป็นบุคคลเป้าหมาย และฟังก์ชันความหนาแน่นของบุคคลเป้าหมายที่ต้องการติดตาม ได้ตั้งสมการที่ 3.8 และ 3.9 ตามลำดับ

$$q_u = C_q \sum_{i=1}^n k(\|x_i\|^2) \delta[b(x_i) - u] \quad (3.8)$$

โดยที่

$$C_q = \frac{1}{\sum_{i=1}^n k(\|x_i\|^2)}$$

เมื่อ q_u คือ ฟังก์ชันความหนาแน่นที่อ้างอิงเป็นบุคคลเป้าหมายจำนวน u ถึงข้อมูล

x_i คือ ตำแหน่งข้อมูลที่จุดภาพ i ในบริเวณที่พิจารณาทั้งหมด n จุดภาพ

$k(x)$ คือ ฟังก์ชันเคอร์เนลที่ใช้ในการประมาณความหนาแน่น ณ จุดภาพ x

$b(x_i)$ คือ ค่าข้อมูลคุณลักษณะสำคัญที่ตำแหน่งข้อมูลที่จุดภาพ i

$\delta[x]$ คือ อิมพัลส์ฟังก์ชัน ซึ่งมีเงื่อนไขดังนี้ $\delta[x] = \begin{cases} 1, & x = 0 \\ 0, & \text{otherwise} \end{cases}$

C_q คือ ค่านอมอลไลส์ที่ทำให้ฟังก์ชันความหนาแน่นมีค่าเท่ากับหนึ่ง

$$p_u(y) = C_p \sum_{i=1}^n k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \delta[b(x_i) - u] \quad (3.9)$$

โดยที่

$$C_p = \frac{1}{\sum_{i=1}^n k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)}$$

เมื่อ $p_u(y)$ คือ ฟังก์ชันความหนาแน่นที่อ้างอิงเป็นบุคคลเป้าหมายที่ต้องการติดตาม

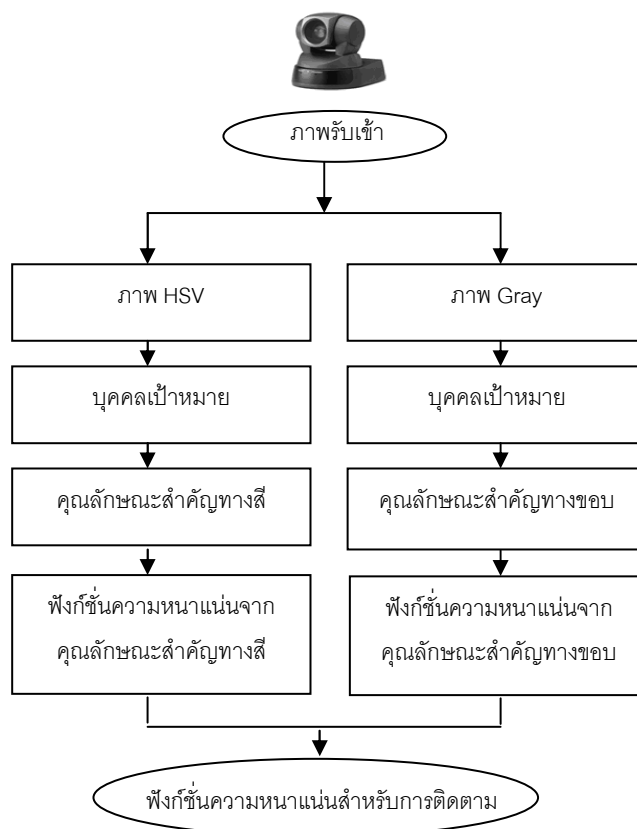
จำนวน u ถึงข้อมูล ณ ตำแหน่งจุดภาพ y

- x_i คือ ตำแหน่งข้อมูลที่จุดภาพ i ในบริเวณที่พิจารณาทั้งหมด n จุดภาพ
- y คือ ตำแหน่งข้อมูลที่จุดภาพ y
- $k(x)$ คือ ฟังก์ชันเคอร์เนลที่ใช้ในการประมาณความหนาแน่น ณ จุดภาพ x
- $b(x_i)$ คือ ค่าข้อมูลคุณลักษณะสำคัญที่ตำแหน่งข้อมูลจุดภาพ i
- $\delta[x]$ คือ อิมพัลส์ฟังก์ชัน ซึ่งมีเงื่อนไขดังนี้ $\delta[x] = \begin{cases} 1, x = 0 \\ 0, otherwise \end{cases}$
- C_p คือ ค่านอมอลไลส์ที่ทำให้ฟังก์ชันความหนาแน่นมีค่าเท่ากับหนึ่ง

จากสมการที่ 3.8 และ 3.9 นั้นจะถูกคำนวณเพื่อประมาณฟังก์ชันความหนาแน่น หลังจากดึงคุณลักษณะสำคัญทางสีและทางขอบออกมาเรียบร้อยแล้ว ไดอะแกรมแสดงการหาฟังก์ชันความหนาแน่นสำหรับการติดตามบุคคลที่น่าเสนอ แสดงดังรูปที่ 3.5 เมื่อพิจารณาลำดับภาพวีดิทัศน์หนึ่งภาพ ระบบจะทำการแปลงปริภูมิสีของภาพเป็นปริภูมิสี HSV และปริภูมิระดับสีเทา เพื่อทำการดึงข้อมูลคุณลักษณะสำคัญทางสีและทางขอบออกมาตามลำดับ

การประมาณฟังก์ชันความหนาแน่นตามสมการที่ 3.8 และ 3.9 นั้นจะเป็นการนำข้อมูลคุณลักษณะสำคัญทางสีและทางขอบที่ใช้ ซึ่งสรุปได้ว่า ในขั้นตอนการประมาณความหนาแน่นที่น่าเสนอนี้ จะประกอบไปด้วยสองส่วนที่ทำงานขนานกัน คือ การดึงคุณลักษณะสำคัญทางสีเพื่อนำมาประมาณฟังก์ชันความหนาแน่น และการดึงคุณลักษณะสำคัญทางขอบเพื่อนำมาประมาณฟังก์ชันความหนาแน่น หลังจากนั้นจะนำฟังก์ชันความหนาแน่นทั้งสองมารวมกันเพื่อสร้างเป็นฟังก์ชันความหนาแน่นของบุคคลเป้าหมายใช้ในกระบวนการติดตามบุคคลเป้าหมายต่อไป

เพื่อความเข้าใจที่ดีขึ้น ขอแสดงตัวอย่างการคำนวณการประมาณความหนาแน่นทั้งสองส่วนได้ดังหัวข้อต่อไป



รูปที่ 3.5 แสดงการหาฟังก์ชันความหนาแน่นสำหรับการติดตามบุคคลที่น่าสงสัย

3.3.1.1 การดึงคุณสมบัติสำคัญทางสีเพื่อประมาณฟังก์ชันความหนาแน่น จากบทที่ 2 หัวข้อเรื่องปริภูมิสี เนื่องจากปริภูมิสี HSV ค่า H หรือ Hue นั้นแสดงถึงค่าของสีที่สะท้อนมาจากสีของวัตถุ ซึ่งแตกต่างกันตามความยาวของคลื่นแสงที่มากกระทบวัตถุ และสะท้อนกลับมาที่ตาเรา ซึ่งจะถูกแทนด้วยค่า 0 ถึง 360 องศา (สีแดง = 0° , สีเขียว = 120° และสีน้ำเงิน = 240° .) สามารถแปลงค่าสีจากปริภูมิสี RGB เป็นปริภูมิสี HSV ได้ดังนี้:

หาค่าสูงสุดระหว่างข้อมูลค่าสี R, G และ B ($Max(R, G, B)$), ค่าต่ำสุดระหว่างข้อมูลค่าสี R, G และ B ($Min(R, G, B)$) และ ค่าความแตกต่างระหว่างข้อมูลค่าสีสูงสุดและต่ำสุด ($Delta = Max(R, G, B) - Min(R, G, B)$) คำนวณหาค่า H ได้จากสมการที่ 3.10

$$H = \begin{cases} 0, & \text{if } Max(R, G, B) = Min(R, G, B) \\ (60^\circ \times \frac{G - B}{Delta}) \bmod 360, & \text{if } Max(R, G, B) = R \\ 60^\circ \times \frac{B - R}{Delta} + 120^\circ, & \text{if } Max(R, G, B) = G \\ 60^\circ \times \frac{R - G}{Delta} + 240^\circ, & \text{if } Max(R, G, B) = B \end{cases} \quad (3.10)$$

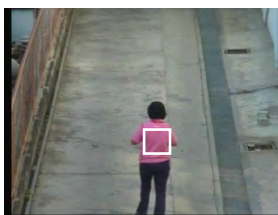
ตัวอย่างการคำนวณการแปลงค่าสีจากปริภูมิ RGB ไปเป็นค่า H ในปริภูมิสี HSV เช่น ถ้าค่าสีในปริภูมิสี R, G และ B เป็น 255, 255 และ 0 ตามลำดับ

ค่าสูงสุดระหว่างข้อมูลค่าสี R, G และ B คือ 255, ค่าต่ำสุดระหว่างค่าสี R, G และ B คือ 0 และค่าความแตกต่างระหว่างค่าสูงสุดและต่ำสุดระหว่างข้อมูลค่าสี R, G และ B คือ 255 แสดงการคำนวณหาค่า H ได้จากสมการ และแสดงภาพตัวอย่างการคำนวณการแปลงค่าสีจากปริภูมิ RGB เป็นแซลแนล H ได้จากรูปที่ 3.6

$$H = (60^\circ \times \frac{255-0}{255} + 360^\circ) \bmod 360 = 60^\circ$$

หรือ

$$H = (60^\circ \times \frac{0-255}{255} + 120^\circ) = 60^\circ$$



(ก) บริเวณบุคคลเป้าหมาย

255	255	255	255	255	255	255	255
255	255	2	255	255	255	255	255
255	255	2	255	255	255	255	255
255	255	2	255	255	2	0	0
255	255	2	255	255	2	0	0
255	255	2	255	255	2	0	0
255	255	2	255	255	2	0	0
255	255	2	255	255	2	0	0
255	255	2	255	255	2	0	0
255	255	2	255	255	2	0	0

R

255	255	255	255	255	255	255	255
255	255	2	255	255	255	255	255
255	255	2	255	255	255	255	255
255	255	2	255	255	2	0	0
255	255	2	255	255	2	0	0
255	255	2	255	255	2	0	0
255	255	2	255	255	2	0	0
255	255	2	255	255	2	0	0
255	255	2	255	255	2	0	0
255	255	2	255	255	2	0	0

G

60	60	60	60	60	60	60	60
60	60	60	60	60	60	60	60
60	60	60	60	60	60	60	60
60	60	60	60	60	60	60	60
60	60	60	60	60	60	60	60
60	60	60	60	60	60	60	60
60	60	60	60	60	60	60	60
60	60	60	60	60	60	60	60
60	60	60	60	60	60	60	60
60	60	60	60	60	60	60	60

B

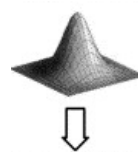
(ข) ค่าสีบริเวณบุคคลในปริภูมิสี RGB (ค) ค่าสีบริเวณบุคคลปริภูมิสี HSV ในแซลแนล H
รูปที่ 3.6 แสดงภาพตัวอย่างการคำนวณการแปลงค่าสีจากปริภูมิ RGB เป็นปริภูมิสี HSV ในแซลแนล H

จากรูปที่ 3.6 (ค) เนื่องจากข้อมูลค่าสีในแซลแนล H จะถูกแทนค่าด้วยค่า 0-360 องศา วิทยานิพนธ์ฉบับนี้จะทำการแจกแจงข้อมูล 0-360 องศาตั้งกล่าวให้อยู่ใน 16 ถึงข้อมูล

เนื่องจากเพื่อต้องการลดความซับซ้อนในการคำนวณเมื่อเข้าสู่ขั้นตอนวิธีการติดตามบุคคล เพราะฉะนั้นเมื่อได้โครงสร้างของ H แล้วจะทำการคำนวณด้วยการตั้งหาร 22.5 เพื่อแบ่งออกเป็น 16 ถึงข้อมูลดังรูปที่ 3.7 (ก)

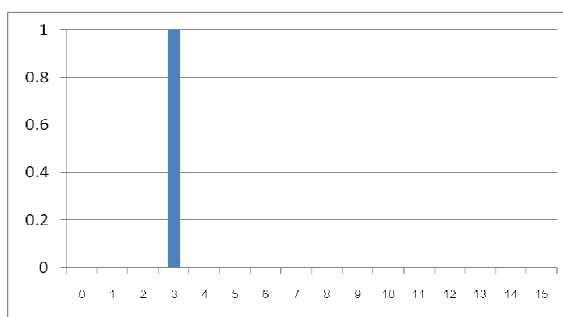
จากหลักการของการประมาณค่าความหนาแน่นจากบทที่ 2 สามารถแสดงตัวอย่างขั้นตอนการประมาณความหนาแน่นของสี่รอบจุดศูนย์กลางของบริเวณบุคคลเป้าหมายได้ด้วยฟังก์ชันหลักซึ่งมีโปรไฟล์เป็นแบบเกาส์เซียนได้ดังรูปที่ 3.7 (ข) และแสดงฟังก์ชันความหนาแน่นที่ได้จากการคำนวณตามสมการที่ 3.8 ของสี่บริเวณบุคคลเป้าหมายได้ดังรูปที่ 3.8

3	3	3	3	3	3	3
3	3	3	3	3	3	3
3	3	3	3	3	3	3
3	3	3	3	3	3	3
3	3	3	3	3	3	3
3	3	3	3	3	3	3
3	3	3	3	3	3	3



0.25	0.50	0.25
0.50	1.00	0.50
0.25	0.50	0.25

(ก) ค่าสี่บริเวณบุคคลหลังจากควอนไทล์เป็น 16 ถึงข้อมูล (ข) ฟังก์ชันเกาส์เซียน 3x3
รูปที่ 3.7 แสดงภาพจำลองการคำนวณฟังก์ชันความหนาแน่นของข้อมูลสี่
ด้วยวิธีการประมาณความหนาแน่นด้วยฟังก์ชันหลัก



รูปที่ 3.8 แสดงฟังก์ชันความหนาแน่นของข้อมูลสี่บริเวณบุคคลเป้าหมาย

3.3.1.2 การตั้งคุณลักษณะสำคัญทางขอบเพื่อประมาณฟังก์ชันความหนาแน่น และจากบทที่ 2 การหาขอบภาพ คือ การตรวจสอบว่าเส้นขอบลากผ่านหรือใกล้เคียงกับจุดใด โดยวัดจากการเปลี่ยนแปลงของความเข้มในตำแหน่งที่ใกล้เคียงกับจุดดังกล่าว ซึ่งวิธีการหาขอบนั้นมีด้วยหลายวิธี สามารถแบ่งได้เป็น 2 กลุ่มคือการหาขอบภาพด้วยอนุพันธ์อันดับที่หนึ่งและการหาขอบภาพด้วยอนุพันธ์อันดับที่สอง สำหรับการหาขอบภาพด้วยอนุพันธ์

อันดับที่หนึ่งนั้น จะหาขอบภาพโดยการหาจุดต่ำสุดและจุดสูงสุดในรูปของอนุพันธ์อันดับหนึ่งของภาพ เช่น โซเบล, โรเบิร์ต, ฟรีวิตต์ หรือแคนนี่ เป็นต้น การหาขอบภาพด้วยอนุพันธ์อันดับที่สองจะหาขอบ เช่น วิธีของ Laplacian และ วิธีของ Marrs-Hildreth เป็นต้น

สำหรับวิทยานิพนธ์ฉบับนี้เลือกใช้การหาขอบภาพด้วยวิธีโซเบล เนื่องจากการหาขอบภาพของบุคคลเป้าหมายนั้นไม่จำเป็นที่การหาขอบภาพจะต้องได้รายละเอียดครบ เช่น รายละเอียดของใบหน้า เป็นต้น แต่จะต้องสามารถหาขอบของภาพให้ได้ผลลัพธ์ในระดับดี และมีการคำนวณที่ไม่ยุ่งยากซับซ้อน ขั้นตอนการทำงานของวิธีโซเบล จะหาค่าขอบของภาพโดยใช้หลักการหาค่าอนุพันธ์อันดับหนึ่งกับภาพที่มีลักษณะเป็นภาพ 2 มิติ ค่าขอบของภาพจะถูกแทนด้วย $\nabla f(x, y)$ ณ ตำแหน่ง (x, y) ซึ่งถูกกำหนดด้วยเวกเตอร์ดังสมการที่ 3.11

$$\nabla f(x, y) = \left[\frac{\partial G_x}{\partial x} \quad \frac{\partial G_y}{\partial y} \right] \quad (3.11)$$

สิ่งที่ได้มาจากวิเคราะห์เวกเตอร์ทำให้ได้ค่า $\nabla f(x, y)$ ที่เกิดจากการเปลี่ยนแปลงในระดับมากที่สุดของขอบ ในทิศทางต่างๆ ที่ตำแหน่ง (x, y) และที่สำคัญที่สุดในเชิงปริมาณในการหาเส้นขอบ (Magnitude of Edge) นั่นก็คือค่า ∇f ดังสมการที่ 3.12

$$\nabla f = \text{magnitude}(\nabla f(x, y)) = \sqrt{G_x^2 + G_y^2} \quad (3.12)$$

ทิศทางของขอบ (Orientation of Edge) สามารถคำนวณโดยให้ $\theta(x, y)$ แทนด้วยทิศทางมุมของขอบ (x, y) โดยใช้สมการ 3.13

$$\theta(x, y) = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad (3.13)$$

เมื่อมุมคือระยะทางของแนวแกน x ทิศทางของขอบที่ตำแหน่ง (x, y) คือเส้นตั้งฉากไปยังทิศทางของขอบภาพที่จุดนั้นๆ การทำงานของวิธีโซเบล อาศัยวิธีการหาค่าขอบภาพด้วยตัวดำเนินการที่มีขนาด 3×3 สำหรับเป็นตัวกระทำการในขั้นตอนวิธี แสดงตัวกระทำการของการหาขอบภาพด้วยวิธีโซเบลดังรูปที่ 3.9 (ก) ตัวกระทำการของวิธีหา ขอบภาพแบบ Sobel ในแนว G_x และตัวกระทำการของวิธีหา ขอบภาพแบบ Sobel ในแนว G_y ดังรูปที่ 3.9 (ข)

-1	-2	-1
0	0	0
1	2	1

(ก) ตัวกระทำการของวิธีหา ขอบภาพแบบ Sobel ในแนว G_x

-1	0	1
-2	0	2
-1	0	1

(ข) ตัวกระทำการของวิธีหา ขอบภาพแบบ Sobel ในแนว G_y

x ₁	x ₂	x ₃
x ₄	x ₅	x ₆
x ₇	x ₈	x ₉

(ค) การแทนที่ต่างๆของตัวกระทำการ

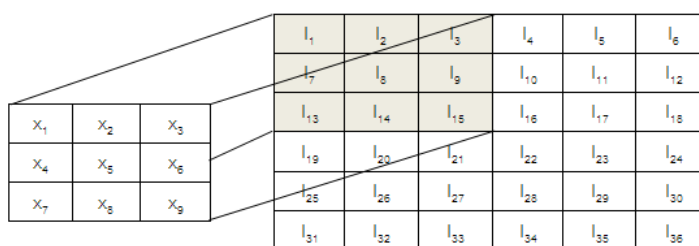
รูปที่ 3.9 ตัวกระทำการของการหาขอบภาพด้วยวิธี Sobel

เมื่อได้ตัวดำเนินการของโซเบลที่มีขนาด 3 x 3 เรียบร้อยแล้วก็จะนำไปแทนที่ในสมการ โดยค่า G_x และ G_y สามารถคำนวณได้จากสมการที่ 3.14 และ 3.15 ตามลำดับ

$$G_x = (x_7 + 2x_8 + x_9) - (x_1 + 2x_2 + x_3) \quad (3.14)$$

$$G_y = (x_3 + 2x_6 + x_9) - (x_1 + 2x_4 + x_7) \quad (3.15)$$

หรือแสดงขั้นตอนการหาขอบภาพด้วยวิธีการทำคอนโวลูชัน นั่นคือ การนำค่า G_x และ G_y ของตัวดำเนินการแบบโซเบลมากระทำการกันด้วยการคูณกับภาพ แสดงดังรูปที่ 3.10



รูปที่ 3.10 วิธีการคอนโวลูชัน

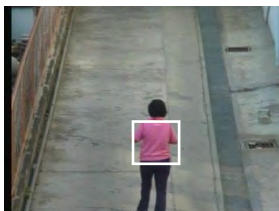
แสดงตัวอย่างขั้นตอนการหาคุณลักษณะสำคัญทางขอบของเป้าหมายบุคคลที่สนใจ ก่อนนำไปประมวลความหนาแน่น สามารถทำตามขั้นตอนได้ดังนี้

ก) แปลงภาพวีดิทัศน์รับเข้าให้เป็นภาพในปริภูมิระดับสีเทา

ข) กำหนดบริเวณบุคคลเป้าหมายเพื่อหาค่าขอบภาพด้วยวิธีของโซเบลทั้งขอบภาพในแนวแกน x และ ขอบภาพในแนวแกน y

ค) หาทิศทางของขอบได้จากสมการ 3.13 แต่เนื่องจากค่าของ $\arctan(\theta)$ ใดๆ นั้นจะมีค่าอยู่ในช่วง $(-\frac{\pi}{2}, \frac{\pi}{2})$ หรือมีค่าอยู่ในช่วงประมาณ 1.57 ถึง -1.57 เนื่องจากค่าลบนั้นไม่สามารถนำมาใช้งานจริงได้ จึงทำการนอมอลไลส์ ค่าดังกล่าวให้มีค่าบวก โดยบวกค่า $\frac{\pi}{2}$ ทุกจุดภาพ และแปลงข้อมูลหลังให้อยู่ในรูปองศา (0° - 180°) องศาอีกครั้งหนึ่ง จากนั้นจึงนำข้อมูลที่ได้นี้ มาหาฟังก์ชันความหนาแน่นเช่นเดียวกับการประมวลความหนาแน่นของข้อมูลสี แสดงการหาข้อมูลทิศทางของขอบหลังจากนอมอลไลส์ดังกล่าว ได้ดังสมการที่ 3.16 และแสดงตัวอย่างขั้นตอนการประมวลฟังก์ชันความหนาแน่นของข้อมูลคุณลักษณะสำคัญทางขอบได้ดังนี้

$$\theta(x, y) = \left[\tan^{-1} \left(\frac{D_y(x, y)}{D_x(x, y)} \right) + \frac{\pi}{2} \right] \times \frac{180^\circ}{\pi} \quad (3.16)$$



(ก) บริเวณบุคคลเป้าหมาย

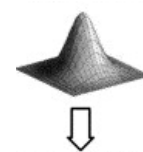
2	3	255	255	1	2	0
3	255	255	255	255	255	1
255	255	2	2	14	255	255
255	1	2	2	11	2	255
255	4	5	12	10	1	255
255	10	11	11	255	255	255
255	18	20	13	255	255	255

*	*	*	*	*	*	*
*	45	90	72	90	109	*
*	45	63	92	117	134	*
*	180	44	28	134	152	*
*	178	63	46	74	46	*
*	178	84	20	46	72	*
*	*	*	*	*	*	*

(ข) ค่าสีบริเวณบุคคลในปริภูมิสีระดับสีเทา (ค) ข้อมูลขอบบริเวณบุคคลคำนวณสมการที่ 3.16 รูปที่ 3.11 แสดงภาพตัวอย่างการคำนวณการแปลงค่าสีจากปริภูมิ RGB เป็นปริภูมิระดับสีเทา

วิทยานิพนธ์ฉบับนี้จะทำการแจกแจงข้อมูลของขอบดังกล่าวให้อยู่ใน 8 ถึงข้อมูล ดังรูปที่ 3.12 (ก) เนื่องจากเพื่อต้องการลดความซับซ้อนในการคำนวณเมื่อเข้าสู่ขั้นตอนวิธีการ ติดตามบุคคล สามารถแสดงตัวอย่างขั้นตอนการประมาณความหนาแน่นของขอบรอบจุด ศูนย์กลางของบริเวณบุคคลเป้าหมายได้ด้วยฟังก์ชันหลักซึ่งมีโปรไฟล์เป็นแบบเกาส์เซียนได้ดังรูปที่ 3.7 (ข) และแสดงฟังก์ชันความหนาแน่นที่ได้จากการคำนวณตามสมการที่ 3.8 ของขอบบริเวณ บุคคลเป้าหมายได้ดังรูปที่ 3.13

*	*	*	*	*	*	*
*	2	4	3	4	4	*
*	2	3	4	5	5	*
*	7	2	1	5	6	*
*	7	3	2	3	2	*
*	7	3	1	2	3	*
*	*	*	*	*	*	*

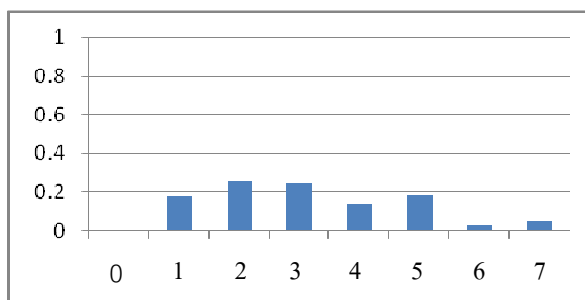


0.25	0.50	0.25
0.50	1.00	0.50
0.25	0.50	0.25

(ก) ค่าสีบริเวณบุคคลหลังจากควอนไทส์เป็น 8 ถึงข้อมูล

(ข) ฟังก์ชันเกาส์เซียน 3x3

รูปที่ 3.12 แสดงภาพจำลองการคำนวณฟังก์ชันความหนาแน่นของข้อมูลขอบ ด้วยวิธีการประมาณความหนาแน่นด้วยฟังก์ชันหลัก



รูปที่ 3.13 แสดงฟังก์ชันความหนาแน่นของข้อมูลขอบบริเวณบุคคลเป้าหมาย

จากรูปที่ 3.13 แสดงฟังก์ชันความหนาแน่นของข้อมูลขอบบริเวณบุคคลเป้าหมาย และเนื่องจากวิทยานิพนธ์ฉบับนี้ได้นำเอาสองคุณลักษณะสำคัญสองประเภท ได้แก่ สี และขอบ มาประมาณฟังก์ชันความหนาแน่น ดังนั้นเพื่อสร้างฟังก์ชันความหนาแน่นสำหรับการติดตามบุคคลในขั้นตอนวิธีการย้ายตามค่าเฉลี่ยใหม่ แสดงได้ดังสมการที่ 3.17 และแสดงฟังก์ชันความหนาแน่นสุดท้ายเพื่อใช้ในการติดตามบุคคลได้ดังรูปที่ 3.14

$$f_u = C_t \left(\sum_{u=0}^{15} A_u + \sum_{u=16}^{23} B_{u-16} \right) \quad (3.17)$$

โดยที่

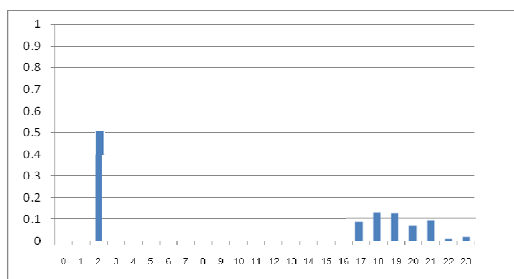
$$C_t = \frac{1}{\sum_{u=0}^{15} A_u + \sum_{u=16}^{23} B_{u-16}}$$

เมื่อ f_u คือ ฟังก์ชันความหนาแน่นที่อ้างอิงเป็นบุคคลเป้าหมายจากข้อมูลสีและขอบ
ต้องการติดตาม จำนวน u ถึงข้อมูล

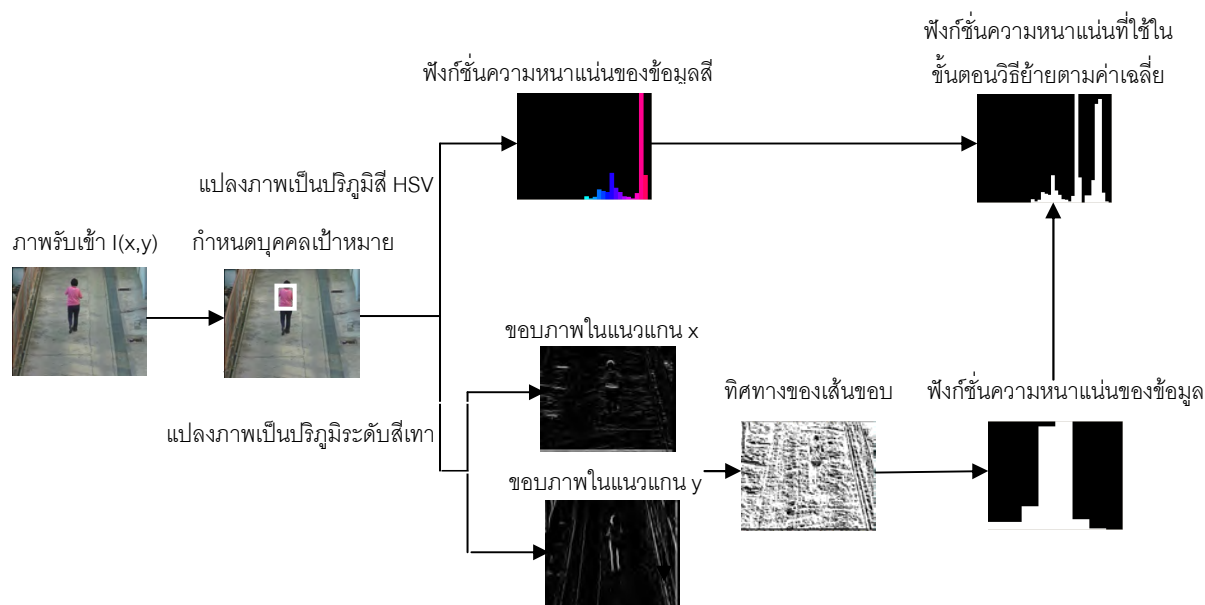
A_{16} คือ ฟังก์ชันความหนาแน่นของข้อมูลขอบ จำนวน 16 ถึงข้อมูล

B_8 คือ ฟังก์ชันความหนาแน่นของข้อมูลขอบ จำนวน 8 ถึงข้อมูล

C_t คือ ค่านอมอลไลส์ที่ทำให้ฟังก์ชันความหนาแน่นมีค่าเท่ากับหนึ่ง



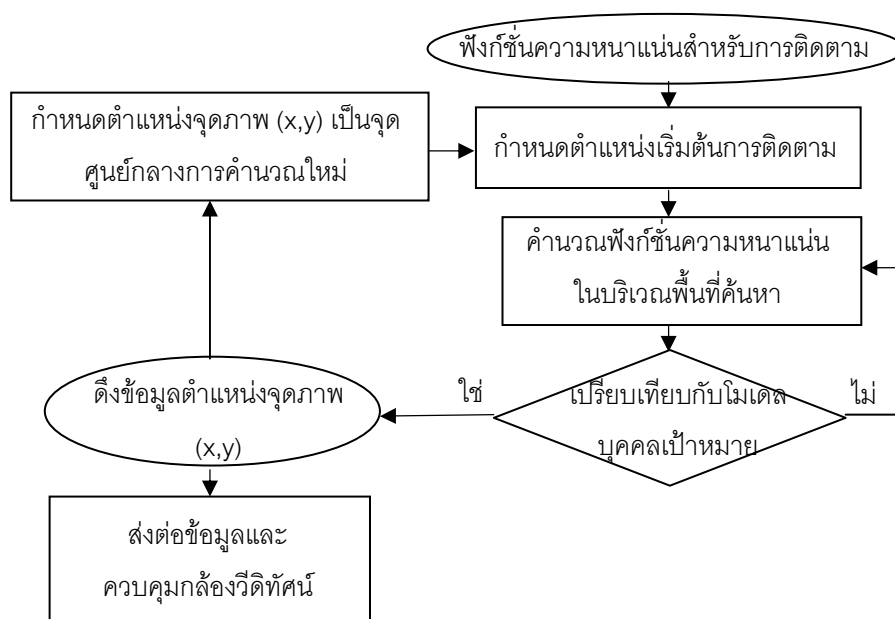
รูปที่ 3.14 แสดงฟังก์ชันความหนาแน่นที่จะนำไปใช้ในการติดตามบุคคล



รูปที่ 3.15 แสดงแผนภาพการหาฟังก์ชันความหนาแน่นที่จะนำไปใช้ในการติดตามบุคคล

3.3.2 ขั้นตอนวิธีการติดตามบุคคลด้วยการย้ายตามค่าเฉลี่ย

จากบทที่ 2 ได้กล่าวถึงที่มาของขั้นตอนวิธีย้ายตามค่าเฉลี่ยที่ได้ถูกคิดค้นขึ้นในอดีต ซึ่งจะเห็นได้ว่า ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยนั้นสามารถนำมาประยุกต์ใช้งานได้อย่างหลากหลาย เช่น การจัดกลุ่มและแยกส่วนภาพ การทำภาพให้เรียบ การรู้จำภาพ รวมถึงการติดตามวัตถุหรือบุคคล ดังที่วิทยานิพนธ์ฉบับนี้สนใจด้วย ในหัวข้อนี้จึงขออธิบายถึงการนำขั้นตอนวิธีการย้ายตามค่าเฉลี่ยมาใช้เพื่อติดตามบุคคล โดยมีพื้นฐานมาจากการประมาณฟังก์ชันความหนาแน่นด้วยฟังก์ชันเคอร์เนลที่ถูกกล่าวไว้ในหัวข้อที่ 3.3.1 ไว้แล้ว หลังจากได้ฟังก์ชันความหนาแน่นมาแล้ว ในขั้นตอนต่อไปคือการนำฟังก์ชันความหนาแน่นนี้ มาใช้ในขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเพื่อติดตามบุคคลได้อย่างไร แสดงบล็อกไดอะแกรมการนำฟังก์ชันความหนาแน่นที่ได้มาใช้สำหรับการติดตามบุคคลด้วยขั้นตอนวิธีการย้ายตามค่าเฉลี่ยที่น่าเสนอ ดังรูปที่ 3.16

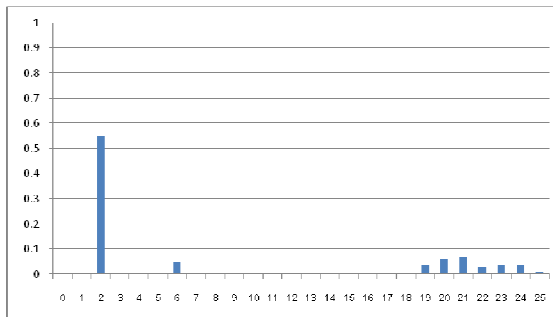


รูปที่ 3.16 แสดงการหาความหนาแน่นของความน่าจะเป็นสำหรับการติดตามบุคคลที่น่าเสนอ

การนำขั้นตอนวิธีการย้ายตามค่าเฉลี่ยมาใช้เพื่อติดตามบุคคลนั้นสามารถอธิบายเป็นขั้นตอนการคำนวณได้ดังนี้

ก) ขั้นตอนแรกเมื่อบุคคลเป้าหมายเข้าสู่มุมมองกล้องวิดีโอตัวที่หนึ่ง จะทำการดึงคุณลักษณะสำคัญเพื่อนำไปประมาณฟังก์ชันความหนาแน่นอ้างอิงเป็นบุคคลเป้าหมาย ด้วย

สมการที่ 3.8 เพื่อให้ทำความเข้าใจได้ง่ายขอยกตัวอย่างการคำนวณด้วยการสมมุติข้อมูลและภาพต่างๆ เป็นดังข้อมูลทีกล่าวไว้ในหัวข้อ 3.3.1



รูปที่ 3.17 แสดงตัวอย่างฟังก์ชันความหนาแน่นที่จะนำไปใช้อ้างอิงเป็นบุคคลเป้าหมาย

ข) ในลำดับภาพวีดิทัศน์ต่อไปจะทำการประมาณเป็นฟังก์ชันความหนาแน่นของบุคคลเป้าหมายที่ต้องการติดตามที่ตำแหน่ง y_0 ($p(y_0)$) ดังสมการที่ 3.9 และหาค่าสัมประสิทธิ์ Bhattacharyya ของฟังก์ชันความหนาแน่นที่ประมาณขึ้น ณ ตำแหน่งลำดับภาพปัจจุบัน y_0 กับ ฟังก์ชันความหนาแน่นอ้างอิงเป็นบุคคลเป้าหมาย $\rho(q_u, p_u(y_0))$ ซึ่งคำนวณได้จากสมการที่ 3.18

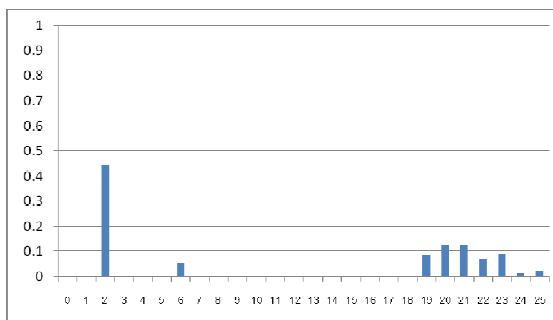
$$\rho(q_u, p_u(y_0)) = \sum_{u=0}^m \sqrt{q_u \times p_u(y_0)} \quad (3.18)$$

เมื่อ $\rho(q_u, p_u(y_0))$ คือ ค่าสัมประสิทธิ์ Bhattacharyya ระหว่าง q_u และ $p_u(y_0)$

q_u คือ ฟังก์ชันความหนาแน่นที่จะนำไปใช้อ้างอิงเป็นบุคคลเป้าหมาย

$p_u(y_0)$ คือ ฟังก์ชันความหนาแน่นของบุคคลเป้าหมายที่ต้องการติดตามตำแหน่ง y_0

m คือ จำนวนถึงข้อมูลทั้งหมด



รูปที่ 3.18 แสดงตัวอย่างฟังก์ชันความหนาแน่นเป้าหมายที่ต้องการติดตามที่ตำแหน่ง y_0

จากตัวอย่างฟังก์ชันที่อธิบายให้เข้าใจอย่างง่ายนี้ สามารถหาค่า สัมประสิทธิ์ Bhattacharyya ระหว่างฟังก์ชันความหนาแน่นที่จะนำไปใช้อ้างอิงเป็นบุคคลเป้าหมาย กับ ฟังก์ชันความหนาแน่นเป้าหมายที่ต้องการติดตามที่ตำแหน่ง y_0 ได้ดังนี้

$$\rho(q_u, (p_u(y_0))) = \sum_{u=1}^{m(bin)} \sqrt{q_u \times p_u(y_0)} = \sqrt{q_0 \times p_0(y_0)} + \sqrt{q_1 \times p_1(y_0)} + \dots + \sqrt{q_{25} \times p_{25}(y_0)} \approx 0.971$$

ค) ขั้นตอนต่อไปคือการประมาณฟังก์ชันความหนาแน่นภายในหน้าต่างค้นหาที่กำหนดขึ้น โดยตั้งสมมุติฐานที่ว่า ต้องการตำแหน่งภายในหน้าต่างค้นหา ที่มีค่าฟังก์ชันความหนาแน่นที่ประมาณขึ้นนั้นเหมือนกับฟังก์ชันความหนาแน่นที่จะนำไปใช้อ้างอิงเป็นบุคคลเป้าหมายมากที่สุด โดยอ้างอิงความเหมือนของฟังก์ชันทั้งสองด้วยฟังก์ชันความเหมือนของ Bhattacharyya (Bhattacharyya similarity function) มีขั้นตอนการคำนวณดังนี้

จากสมการของสัมประสิทธิ์ Bhattacharyya ที่ 3.18 สามารถประมาณด้วยอนุกรมเทเลอร์ดังสมการที่ 3.19

$$f(y) \approx \frac{1}{2} \sum_{u=1}^{m(bin)} \sqrt{p_u(y_0) \times q_u} + \frac{1}{2} \sum_{u=1}^{m(bin)} p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}} \quad (3.19)$$

เมื่อ y_0 คือ ตำแหน่งบุคคลเป้าหมายในลำดับภาพปัจจุบัน (Model target location)

y คือ ตำแหน่งบุคคลเป้าหมายในลำดับภาพถัดไป (Candidate target location)

เนื่องจากเป้าหมายในการคำนวณเพื่อหาตำแหน่งบุคคลเป้าหมายในลำดับภาพต่อไปนั้น ฟังก์ชันความหนาแน่นของบุคคลเป้าหมายที่ต้องติดตามต่อไปนั้น จะต้องมีความเหมือนกับฟังก์ชันความหนาแน่นที่จะนำไปใช้อ้างอิงเป็นบุคคลเป้าหมาย ดังนั้นเมื่อพิจารณาจาก

สมการที่ 3.19 พจน์ $\frac{1}{2} \sum_{u=1}^{m(bin)} \sqrt{p_u(y_0) \times q_u}$ นี้สังเกตได้ว่าจะเป็นอิสระจากตำแหน่งบุคคล

เป้าหมายในลำดับภาพถัดไป ความเหมือนหรือความต่างของฟังก์ชันความเหมือนนี้ จึงขึ้นอยู่กับ

พจน์ $\frac{1}{2} \sum_{u=1}^{m(bin)} p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}}$ เพียงเท่านั้น จึงประมาณฟังก์ชันความเหมือนได้ดังสมการที่ 3.20

$$f(y) \approx \frac{1}{2} \sum_{u=1}^{m(bin)} p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}} \quad (3.20)$$

โดยที่

$$p_u(y) = C \sum_{i=1}^n k \left(\left\| \frac{y - x_i}{h} \right\|^2 \right) \delta[b(x_i) - u]$$

และ

$$w_i = \sum_{u=1}^{m(\text{bin})} \sqrt{\frac{q_u}{p_u(y_0)}} \delta[b(x_i) - u]$$

เมื่อ $f(y)$ คือ ฟังก์ชันความหนาแน่นที่ประมาณขึ้นในตำแหน่ง y ของลำดับภาพถัดไป
 $p_u(y)$ คือ ฟังก์ชันความหนาแน่นที่อ้างอิงเป็นบุคคลเป้าหมายที่ต้องการติดตาม
 จำนวน u ถึงข้อมูล ณ ตำแหน่ง y ของลำดับภาพถัดไป

$p_u(y_0)$ คือ ฟังก์ชันความหนาแน่นที่อ้างอิงเป็นบุคคลเป้าหมายที่ต้องการติดตาม
 จำนวน u ถึงข้อมูล ณ ตำแหน่ง y ของลำดับภาพปัจจุบัน

$k(x)$ คือ ฟังก์ชันเคอร์เนลที่ใช้ในการประมาณความหนาแน่น ณ จุดภาพ x

$b(x_i)$ คือ ค่าข้อมูลคุณลักษณะสำคัญที่ตำแหน่งข้อมูลที่จุดภาพ i

$\delta[x]$ คือ อิมพัลส์ฟังก์ชัน ซึ่งมีเงื่อนไขดังนี้ $\delta[x] = \begin{cases} 1, & x = 0 \\ 0, & \text{otherwise} \end{cases}$

ตามหลักการพื้นฐานของการย้ายตามค่าเฉลี่ยที่กล่าวไว้ในบทที่ 2 ในการย้ายตำแหน่งเพื่อหาตำแหน่งที่มีฟังก์ชันความหนาแน่นที่มีความเหมือนกับฟังก์ชันความหนาแน่นที่จะนำไปใช้อ้างอิงเป็นบุคคลเป้าหมายมากที่สุดนั้น จำเป็นต้องมีการนอลมอลไลส์ด้วยฟังก์ชันความหนาแน่นที่อ้างอิงเป็นบุคคลเป้าหมายที่ต้องการติดตามด้วย เนื่องมาจากในบริเวณที่มีข้อมูลมีความหนาแน่นน้อย ทำให้ขนาดของการเปลี่ยนแปลงตำแหน่งมีค่ามากขึ้น การเคลื่อนที่เข้าหาจุดศูนย์กลางจะเร็วกว่าการใช้เพียงฟังก์ชันความหนาแน่นที่ประมาณขึ้นในตำแหน่งของลำดับภาพถัดไป และจากหลักการของการย้ายตามค่าเฉลี่ยในบทที่ 2 นี้สามารถคำนวณตำแหน่งใหม่ y ได้ ดังสมการที่ 3.21

$$f(y) = \frac{\sum_{i=1}^n x_i w_i k \left(\left\| \frac{y_0 - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^n w_i k \left(\left\| \frac{y_0 - x_i}{h} \right\|^2 \right)} \quad (3.21)$$

ง) หลังจากได้ตำแหน่งของฟังก์ชันความหนาแน่นในลำดับภาพถัดไปแล้วจะทำการคำนวณฟังก์ชันความหนาแน่นที่อ้างอิงเป็นบุคคลเป้าหมายที่ต้องการติดตามของลำดับภาพปัจจุบันนี้ $p_u(y)$ และคำนวณค่าสัมประสิทธิ์ Bhattacharyya ระหว่างฟังก์ชันความหนาแน่นที่อ้างอิงเป็นบุคคลเป้าหมาย q_u และ $p_u(y)$ ที่ตำแหน่ง y คือ $\rho(q_u, p_u(y))$

จ) ถ้า $\rho(q_u, p_u(y_0))$ มีค่าน้อยกว่า $\rho(q_u, p_u(y))$ หรือมีค่าอยู่ในช่วง ขีดเริ่มเปลี่ยนที่กำหนด จะย้ายตำแหน่งจุดศูนย์กลางการคำนวณของหน้าต่างค้นหาไปยังตำแหน่ง y

ฉ) จากนั้นจึงกำหนดค่าเริ่มต้นตำแหน่ง y เป็นตำแหน่ง y_0 แล้วเริ่มต้นทำการคำนวณตามขั้นตอน ข) อีกครั้งไปเรื่อยๆ

3.4 การควบคุมการส่าย ก้มเงย และซูมของกล้องวิดีโอ PTZ เพื่อติดตามใบหน้าบุคคลเป้าหมาย

วิธีการหนึ่งที่จะเพิ่มประสิทธิภาพการติดตามบุคคลเป้าหมายให้มีประสิทธิภาพมากยิ่งขึ้นนั้น คือ การนำประโยชน์ทางด้านความสามารถในการส่าย และก้มเงยของกล้องวิดีโอประเภท PTZ มาประยุกต์ใช้กับงาน ด้วยการส่งค่าพารามิเตอร์ในการควบคุม ดังนั้นกระบวนการหนึ่งที่วิทยานิพนธ์ฉบับนี้ให้ความสำคัญเพื่อเพิ่มประสิทธิภาพให้กับระบบการติดตามบุคคลอย่างต่อเนื่องที่นำเสนอ คือ การส่งพารามิเตอร์เพื่อควบคุมกล้องวิดีโอ PTZ เพื่อเปลี่ยนมุมมองของกล้องวิดีโอทั้งส่าย และก้มเงย ติดตามบุคคลเป้าหมาย โดยเป้าหมายของการคำนวณหาพารามิเตอร์ของตำแหน่งบุคคลเป้าหมาย คือ ความต้องการที่จะเปลี่ยนมุมมองการรับภาพของกล้องวิดีโอ PTZ ให้ทำงานเต็มประสิทธิภาพ โดยให้บริเวณบุคคลเป้าหมายนั้นมาอยู่ตำแหน่งบริเวณกึ่งกลางภาพวิดีโอ ขั้นตอนการคำนวณพารามิเตอร์ต่าง ๆ จะเลือกใช้เฉพาะข้อมูลตำแหน่งจุดภาพในพิกัด 2 มิติเท่านั้น

3.4.1 การคำนวณมุมในการเปลี่ยนมุมมองของกล้องวิดีโอ PTZ

พารามิเตอร์ที่ใช้ในการควบคุมกล้องวิดีโอ PTZ นั้น ในวิทยานิพนธ์ฉบับนี้ประกอบไปด้วยสองส่วนคือ พารามิเตอร์ของการส่าย (ψ) หรือเรียกง่าย ๆ ว่า มุมของการส่าย (Pan angle) และพารามิเตอร์ของการก้มเงย (ϕ) หรือ มุมของการก้มเงย (Tilt angle) สามารถหาค่าพารามิเตอร์ทั้งสองนี้ได้ ดังสมการ 3.5 และ 3.6 ตามลำดับ ส่วนพารามิเตอร์ของการซูมนั้น ในวิทยานิพนธ์ฉบับนี้ไม่ได้ให้ความสำคัญไว้ เนื่องจากการติดตามบุคคลเป้าหมายนั้นไม่จำเป็นต้องเก็บรายละเอียดของบุคคลมากนัก อีกทั้งถ้ากล้องวิดีโอ PTZ มีการซูมเกิดขึ้นจะทำให้มุมมองใน

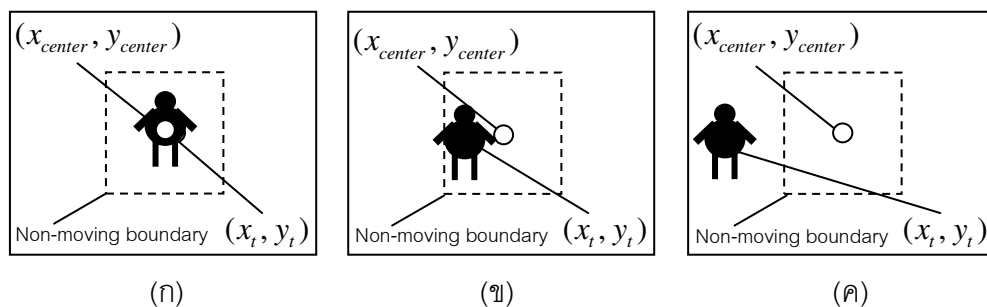
การติดตามแคบลง การใช้งานกล้องวิดีโอเพื่อส่าย ก้มเงยติดตามบุคคลเป้าหมายนั้นๆ จะถูกใช้งานอย่างไม่เต็มที่

$$\psi = \frac{Diff_x}{\beta_x} = \frac{x_t - x_{center}}{\beta_x} \tag{3.5}$$

$$\phi = \frac{Diff_y}{\beta_y} = \frac{y_t - y_{center}}{\beta_y} \tag{3.6}$$

เมื่อกำหนดให้ $Diff_x$ และ $Diff_y$ คือ ระยะห่างในแนวนอนและแนวตั้งระหว่างตำแหน่งกึ่งกลางบุคคลเป้าหมาย (x_t, y_t) กับตำแหน่งจุดศูนย์กลางภาพวิดีโอ (x_{center}, y_{center}) β_x และ β_y คือ อัตราของจำนวนจุดภาพต่อการเคลื่อนที่ของกล้องไป 1 องศาในแนวแกนแนวนอนและแกนตั้ง ตามลำดับ

ในกรณีที่บุคคลเป้าหมายอยู่บริเวณกึ่งกลางภาพอยู่แล้ว ดังรูปที่ 3.20 (ก) หรืออยู่ใกล้กับบริเวณกึ่งกลางภาพแต่อยู่ภายในขอบเขตวิดีโอที่ไม่ต้องเปลี่ยนมุมมอง (Non-moving boundary) ดังรูปที่ 3.20 (ข) นั้น ระบบจะไม่มีคำสั่งควบคุมให้กล้องวิดีโอเพื่อเปลี่ยนมุมมอง แต่ถ้าในลำดับภาพวิดีโอใดที่บุคคลเป้าหมายได้เคลื่อนที่ออกจากขอบเขตวิดีโอที่ไม่ต้องเปลี่ยนมุมมอง จึงจะคำนวณตำแหน่งจุดกึ่งกลางบุคคลเป้าหมายและส่งพารามิเตอร์เพื่อให้กล้องวิดีโอ PTZ เคลื่อนที่ ผลลัพธ์ที่ได้ คือ ตำแหน่งของบุคคลเป้าหมายนั้นจะอยู่กึ่งกลางภาพวิดีโอหรือภายในบริเวณขอบเขตวิดีโอที่ไม่ต้องเปลี่ยนมุมมองตลอดเวลา



รูปที่ 3.19 รูปแบบการควบคุมกล้องวิดีโอ PTZ

บทที่ 4

ผลการจำลองระบบ

ในบทนี้จะกล่าวถึงผลการจำลองระบบการติดตามบุคคลอย่างต่อเนื่องที่นำเสนอ และแบ่งการแสดงผลการทดลองออกเป็นสองรูปแบบ คือ การแสดงผลการทดลองในเชิงปริมาณหรือผลการทดลองที่ถูกแสดงด้วยค่าตัวเลข (Quantitative) และการแสดงผลการทดลองในเชิงคุณภาพ (Qualitative) จะถูกแสดงผลการทดลองด้วยลำดับภาพวีดิทัศน์ตัวอย่าง แต่เนื่องจากลักษณะรูปแบบของวิทยานิพนธ์นี้เป็นในลักษณะงานวิจัยเชิงประยุกต์กับการใช้งานในเวลาจริง ทำให้ผลการทดลองส่วนใหญ่จะ มีการวิเคราะห์ผลในเชิงคุณภาพเป็นส่วนใหญ่ โดยแบ่งการทดลองออกเป็นส่วนใหญ่ 3 ส่วน ตามโครงสร้างของระบบที่นำเสนอในบทที่ 3 คือ ส่วนของขั้นตอนวิธีการเทียบมาตรฐานทางสีที่นำเสนอ ขั้นตอนวิธีการติดตามบุคคลที่นำเสนอ และส่วนสุดท้ายคือส่วนของการควบคุมการถ่ายและกัมเมยของกล้องวีดิทัศน์ PTZ ทั้งสองตัวและส่งข้อมูลระหว่างกัน เพื่อให้สามารถติดตามบุคคลเป้าหมายได้อย่างต่อเนื่อง ซึ่งจะแสดงรายละเอียดเบื้องต้นของการทดลองระบบในแต่ละส่วน แยกเป็นหัวข้อได้ดังนี้

1) ส่วนของขั้นตอนวิธีการเทียบมาตรฐานทางสีที่นำเสนอ นั้น แบ่งการทดลองออกเป็น 2 รูปแบบ คือ ส่วนแรกเป็นการแสดงผลการทดลองในเชิงปริมาณ ซึ่งจะแสดงผลการทดลองเปรียบเทียบค่าสีของภาพจากกล้องวีดิทัศน์สองตัวในระบบที่ถูกติดตั้งต่างบริเวณกัน ทั้งก่อนและหลังทำการเทียบมาตรฐานทางสีให้กับระบบ และในส่วนที่สอง คือ การแสดงผลการทดลองในเชิงคุณภาพ นั่นคือ แสดงลำดับภาพวีดิทัศน์ตัวอย่างที่ได้จากการติดตามบุคคล ทั้งก่อนและหลังทำการเทียบมาตรฐานทางสี

2) ส่วนของขั้นตอนวิธีการติดตามบุคคลที่นำเสนอ จะทำการทดลองเพื่อเปรียบเทียบประสิทธิภาพของขั้นตอนวิธีที่นำมาใช้ในรูปแบบการใช้คุณลักษณะสำคัญเพื่อการติดตามบุคคลเป้าหมายที่แตกต่างกัน รวมถึงเปรียบเทียบคุณลักษณะสำคัญของบุคคลเป้าหมายที่ได้นำเสนอ ซึ่งการทดลองในส่วนนี้จะขอแสดงเพียงผลการทดลองเชิงคุณภาพ ซึ่งแสดงด้วยลำดับภาพวีดิทัศน์ตัวอย่างของการติดตามเป้าหมายเท่านั้น

3) ส่วนของการทำงานร่วมกันระหว่างกล้องวีดิทัศน์ PTZ เพื่อถ่ายกัมเมยติดตามบุคคลเป้าหมายอย่างต่อเนื่องและส่งต่อข้อมูลระหว่างกันนั้น จะแบ่งการทดลองออกเป็นสองส่วน คือ ผลการจำลองระบบในเชิงปริมาณจะเปรียบเทียบความคล้ายของคุณลักษณะสำคัญทางสีจากกล้องวีดิทัศน์ทั้งสองตัวในระบบ ซึ่งแสดงการเปรียบเทียบด้วยค่าสัมประสิทธิ์ Bhattacharyya ทั้งก่อน

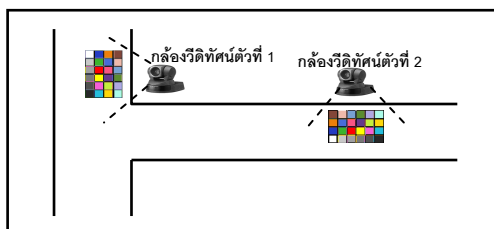
และหลังเทียบมาตรฐานทางสี และในส่วนที่สองจะแสดงผลการจำลองระบบในเชิงคุณภาพ จะแสดงลำดับภาพวีดิทัศน์ตัวอย่างที่ได้จากการติดตามบุคคลจากกล้องวีดิทัศน์ตัวหนึ่งไปอีกตัวหนึ่ง

4.1 ผลการจำลองระบบในส่วนของการขั้นตอนวิธีการเทียบมาตรฐานทางสีที่นำเสนอ

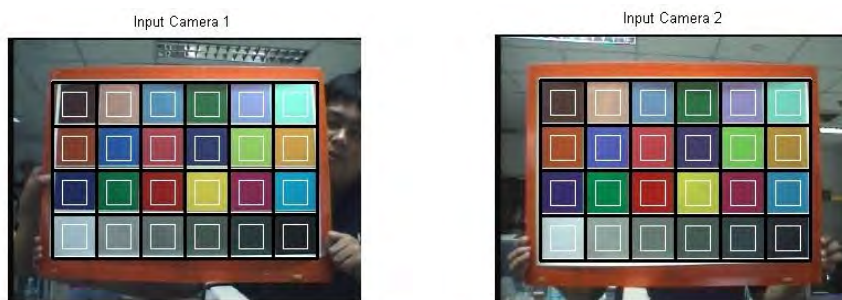
เนื่องจากจุดประสงค์ของการเทียบมาตรฐานทางสีนั้น เพื่อให้ข้อมูลสีเดียวกันที่รับได้จากกล้องวีดิทัศน์ที่ถูกติดตั้งต่างบริเวณกันมีค่าข้อมูลสีได้เหมือนกันหรือใกล้เคียงกันมากที่สุด วิธีการเทียบมาตรฐานทางสีที่นำเสนอนี้ จะใช้เทคนิคการประมาณด้วยสมการถดถอยแบบเชิงเส้นเข้าช่วย โดยอาศัยความสัมพันธ์ระหว่างค่าสีจากภาพตารางสีอ้างอิงของกล้องวีดิทัศน์ทั้งสองตัวในระบบ ซึ่งถูกแทนด้วยเมทริกซ์ ที่เรียกว่า เมทริกซ์ของการเทียบมาตรฐานทางสี ซึ่งแสดงการคำนวณหาค่าเมทริกซ์ของการเทียบมาตรฐานทางสี ได้จากบทที่ 3

4.1.1 ผลการทดลองในเชิงปริมาณในส่วนของการขั้นตอนวิธีการเทียบมาตรฐานทางสีที่นำเสนอ

ขั้นตอนการทดลองนั้น ขั้นแรกจะทำการเก็บภาพตารางสีอ้างอิงจากกล้องวีดิทัศน์ทั้งสองตัวในระบบ ตามลักษณะการเก็บภาพ ดังรูปที่ 4.1 และแสดงภาพที่ได้จากกล้องวีดิทัศน์ตัวที่หนึ่งและกล้องวีดิทัศน์ตัวที่สองดังรูปที่ 4.2(ก) และ 4.2(ข) ตามลำดับ



รูปที่ 4.1 แสดงตำแหน่งการเก็บภาพจากกล้องวีดิทัศน์สองตัวในระบบ



(ก) แสดงภาพจากกล้องวีดิทัศน์ตัวที่ 1 (ข) แสดงภาพจากกล้องวีดิทัศน์ตัวที่ 2

รูปที่ 4.2 แสดงภาพที่ได้จากกล้องวีดิทัศน์สองตัวเพื่อเป็นข้อมูลในการเทียบมาตรฐานทางสี

จะเก็บข้อมูลสีของจุดภาพในบริเวณกรอบสี่เหลี่ยมสีขาวเล็กๆในรูปที่ 4.2(ก) และรูปที่ 4.2(ข) ซึ่งมีทั้งหมด 400 จุดภาพต่อหนึ่งกรอบสี่เหลี่ยม และนำข้อมูลค่าสีที่ได้ทั้งหมดเป็นข้อมูลสำหรับนำไปเข้าสู่วิธีการเทียบมาตรฐานทางสี เพื่อหาเมทริกซ์ของการเทียบมาตรฐานทางสีต่อไป

เมื่อได้เมทริกซ์ของการเทียบมาตรฐานทางสีแล้ว จะนำเมทริกซ์ของการเทียบมาตรฐานทางสีนี้คูณเข้ากับข้อมูลสี RGB ของกล้องวิดีโอที่ตัวที่สอง เพื่อปรับให้ข้อมูลสี RGB ของกล้องวิดีโอที่ตัวที่สอง มีค่าใกล้เคียงกับกล้องวิดีโอที่ตัวที่หนึ่งมากที่สุด และจากนั้นจะทำการหาค่าเฉลี่ยของค่าความผิดพลาดกำลังสอง (Mean Square Error, MSE) ของสีต่างๆ ในภาพตารางสีข้างอิงที่รับได้จากกล้องวิดีโอทั้งสองตัว ทั้งก่อนและหลังทำการเทียบมาตรฐานทางสีจากรูปที่ 4.3 แสดงภาพที่ได้จากกล้องวิดีโอที่ตัวที่สองที่ได้หลังจากทำการเทียบมาตรฐานทางสีด้วยวิธีที่นำเสนอ



รูปที่ 4.3 แสดงภาพที่ได้จากกล้องวิดีโอที่ตัวที่สองหลังจากมาตรฐานทางสีที่นำเสนอ

ตารางที่ 4.1 แสดงค่าเฉลี่ยของค่าความผิดพลาดกำลังสอง (MSE) ในแต่ละสีของตารางสีข้างอิงที่รับได้จากกล้องวิดีโอที่ตัวที่สองก่อนและหลังทำการเทียบมาตรฐานทางสี

สีข้างอิง	ค่า MSE ก่อนเทียบมาตรฐานทางสี	ค่า MSE หลังเทียบมาตรฐานทางสี
1.DARK SKIN	16.823	21.641
2.LIGHT SKIN	38.206	32.280
3.BLUE SKY	30.670	25.364
4.FOLIAGE	10.017	10.149
5.BLUE FLOWER	28.694	27.356
6.BLUIISH GREEN	36.336	30.133
7.ORANGE	25.826	26.652

สีอ้างอิง (ต่อ)	ค่า MSE ก่อนเทียบมาตรฐานทางสี (ต่อ)	ค่า MSE หลังเทียบมาตรฐานทางสี (ต่อ)
8.PURPLISH BLUE	26.839	35.464
9.MODERATE RED	14.154	9.129
10.PURPLE	8.869	13.565
11.YELLOW GREEN	19.502	9.037
12.ORANGE YELLOW	23.072	12.234
13.BLUE	15.875	15.716
14.GREEN	17.234	25.859
15.RED	13.379	4.933
16.YELLOW	9.983	13.916
17.MAGENTA	4.967	4.899
18.CYAN	24.338	37.956
19.WHITE	46.719	43.707
20.NEUTRAL 8	29.132	31.964
21.NEUTRAL 6.5	15.000	14.617
22.NEUTRAL 5	12.179	6.928
23.NEUTRAL 3.5	17.926	16.553
24.BLACK	7.439	7.165

ตารางที่ 4.1 แสดงค่าเฉลี่ยของค่าความผิดพลาดกำลังสองทั้ง 24 สีอ้างอิงจากตารางสีอ้างอิงเปรียบเทียบก่อนและหลังทำการเทียบมาตรฐานทางสีซึ่งผลการทดลองพบว่าหลังจากทำการเทียบมาตรฐานทางสีให้กับกล้องวิดีโอที่ศน์ตัวที่สองในระบบแล้ว ค่าสีอ้างอิงของภาพจากกล้องวิดีโอที่ศน์ตัวที่สองส่วนใหญ่นั้น จะมีค่าใกล้เคียงกับกล้องวิดีโอที่ศน์ตัวที่หนึ่ง ซึ่งสังเกตได้จากค่าค่าเฉลี่ยของค่าความผิดพลาดกำลังสองที่ลดลงนั่นเอง

และเพื่อยืนยันผลการทดลองว่าการเทียบมาตรฐานทางสีด้วยวิธีการประมาณด้วยสมการถดถอยเชิงเส้นมีประสิทธิภาพ จึงได้ทำการทดลองเปรียบเทียบวิธีการเทียบมาตรฐานทางสีในแบบต่างเพิ่มเติม ได้แก่ การเทียบมาตรฐานทางสีด้วยวิธีการประมาณด้วยสมการถดถอยเชิงเส้นโดยใช้ข้อมูลสีแซลแนล H ของปริภูมิสี HSV และการเทียบมาตรฐานทางสีด้วยวิธีการประมาณด้วยสมการถดถอยแบบไม่เชิงเส้นโดยใช้ข้อมูลสีแซลแนล H ของปริภูมิสี HSV แสดงผลการเปรียบเทียบค่า MSE ดังตารางที่ 4.2

ตารางที่ 4.2 แสดงการเปรียบเทียบค่าเฉลี่ยของค่าความผิดพลาดกำลังสอง (MSE) ในแต่ละสีของ ตารางสีอ้างอิงจากวิธีการเทียบมาตรฐานทางสีในรูปแบบต่างๆ

สีอ้างอิง	ค่า MSE ก่อนเทียบมาตรฐานทางสี	ค่า MSE หลังเทียบมาตรฐานทางสี		
		RGB-Linear	H-Linear	H-Nonlinear
1.DARK SKIN	16.823	21.641	16.125	21.315
2.LIGHT SKIN	38.206	32.280	25.443	39.421
3.BLUE SKY	30.670	25.364	36.815	20.478
4.FOLIAGE	10.017	10.149	21.810	16.543
5.BLUE FLOWER	28.694	27.356	35.242	23.584
6.BLUISH GREEN	36.336	30.133	66.317	52.351
7.ORANGE	25.826	26.652	22.942	55.764
8.PURPLISH BLUE	26.839	35.464	55.468	32.182
9.MODERATE RED	14.154	9.129	77.242	105.263
10.PURPLE	8.869	13.565	24.338	7.371
11.YELLOW GREEN	19.502	9.037	18.166	12.275
12.ORANGE YELLOW	23.072	12.234	32.259	24.940
13.BLUE	15.875	16.716	34.814	12.517
14.GREEN	17.234	25.859	37.417	20.873
15.RED	13.379	4.933	7.895	54.997
16.YELLOW	9.983	13.916	16.941	19.900
17.MAGENTA	4.967	4.899	56.024	89.536
18.CYAN	24.338	37.956	68.598	25.732
19.WHITE	46.719	43.707	29.827	30.116
20.NEUTRAL 8	29.132	31.964	22.672	20.042
21.NEUTRAL 6.5	15.000	14.617	15.155	11.860
22.NEUTRAL 5	12.179	6.928	8.505	4.690
23.NEUTRAL 3.5	17.926	16.553	12.193	12.583
24.BLACK	7.439	7.165	7.767	7.348
Average	21.549	19.925	31.248	30.081

นอกจากนี้ยังทดสอบวิธีการเทียบมาตรฐานทางสีที่นำเสนออีกวิธีหนึ่ง คือ การทดสอบจากค่าสีที่ได้จากภาพในบริเวณของบุคคลทดสอบภายในลำดับภาพจริงของกล้องวิดีโอที่คนทั้งสองตัว ซึ่งการทดลองนี้ใช้ข้อมูลสีในบริเวณทดสอบสามส่วนคือ ส่วนของสีผม สีผิวและสีเสื้อ จากบุคคลทดสอบจำนวนสามคน โดยแสดงค่าสี R, G และ B ของภาพบริเวณบุคคลทดสอบก่อนและหลังทำการเทียบมาตรฐานทางสีแสดงดังตารางที่ 4.3 และตารางที่ 4.4 แสดงค่าเฉลี่ยของค่าความผิดพลาดกำลังสองของสีบุคคลทดสอบจากกล้องวิดีโอทั้งก่อนและหลังทำการเทียบมาตรฐานทางสีเช่นเดียวกัน

ตารางที่ 4.3 แสดงค่าสี R, G, B และ H ของภาพบุคคลทดสอบก่อนและหลังเทียบมาตรฐานทางสี

บุคคลทดสอบ	กล้องตัวที่ 1	กล้องตัวที่ 2				กล้องตัวที่ 2							
		R	G	B	H	ก่อนเทียบมาตรฐานทางสี				หลังเทียบมาตรฐานทางสี			
1	สีเสื้อ (สีแดง)	130	42	46	2.7°	175	52	47	0.7°	111	33	32	2.4°
	สีผิว	113	99	90	23.4°	98	87	73	36.3°	105	90	71	33.5°
	สีผม	29	33	33	0°	30	33	36	0°	32	30	31	0°
2	สีเสื้อ (สีชมพู)	133	142	148	204°	201	175	177	300°	172	172	176	240°
	สีผิว	108	92	85	18.2°	89	75	67	21.8°	135	107	90	22.6°
	สีผม	47	52	53	190°	48	47	54	254°	54	49	46	277°
3	สีเสื้อ (สีส้ม)	230	198	105	44.6°	198	140	67	33.4°	246	180	110	33.4°
	สีผิว	139	123	108	49.6°	120	110	94	36.9°	134	108	92	39.5°
	สีผม	47	44	40	34.2°	46	48	46	120°	54	51	48	30°
4	สีเสื้อ (สีน้ำเงิน)	52	55	107	236°	32	50	90	221°	42	47	96	234°
	สีผิว	103	95	79	40°	91	85	78	32.3°	106	94	81	31.2°
	สีผม	35	36	38	220°	31	33	33	180°	42	38	37	288°

ตารางที่ 4.4 แสดงค่าค่าเฉลี่ยของค่าความผิดพลาดกำลังสองของสีบุคคลทดสอบจากกล้องวิดีโอทั้งก่อนและหลังทำการเทียบมาตรฐานทางสี

บุคคลทดสอบ	ค่า MSE ก่อนเทียบมาตรฐานทางสี	ค่า MSE หลังเทียบมาตรฐานทางสี	
1	สีเสื้อ (สีแดง)	26.621	14.583
	สีผิว	14.810	12.987
	สีผม	1.826	2.708

บุคคลทดสอบ (ต่อ)	ค่า MSE ก่อนเทียบมาตรฐานทางสี (ต่อ)	ค่า MSE หลังเทียบมาตรฐานทางสี (ต่อ)	
2	สีเสื้อ (สีชมพู)	46.740	32.685
	สีผิว	18.019	18.065
	สีผม	3.000	5.972
3	สีเสื้อ (สีส้ม)	44.091	14.201
	สีผิว	15.556	12.987
	สีผม	4.203	7.348
4	สีเสื้อ (สีน้ำเงิน)	15.427	9.747
	สีผิว	9.037	2.160
	สีผม	4.082	4.243
ค่าเฉลี่ย	16.951	11.473	

4.1.2 ผลการทดลองในเชิงคุณภาพในส่วนของขั้นตอนวิธีการเทียบมาตรฐานทาง

สีที่นำเสนอ

ตารางที่ 4.5 แสดงลำดับภาพตัวอย่างการทดลองก่อนและหลังทำการเทียบมาตรฐานทางสี

บุคคล	กล้องตัวที่ 1	กล้องตัวที่ 2 ก่อนเทียบมาตรฐานทางสี	กล้องตัวที่ 2 หลังเทียบมาตรฐานทางสี
บุคคลที่ 1 (เสื้อแดง)			
บุคคลที่ 2 (เสื้อชมพู)			
บุคคลที่ 3 (เสื้อส้ม)			

บุคคล (ต่อ)	กล้องตัวที่ 1	กล้องตัวที่ 2 ก่อนเทียบมาตรฐานทางสี	กล้องตัวที่ 2 หลังเทียบมาตรฐานทางสี
บุคคลที่ 4 (เสื้อน้ำเงิน)			

จากตารางที่ 4.5 แสดงลำดับภาพวีดิทัศน์ตัวอย่างที่ได้จากทดลองรับภาพบุคคลจากสถานการณ์จริงจากกล้องตัวที่หนึ่งและตัวที่สองทั้งก่อนและหลังทำการเทียบมาตรฐานทางสี ซึ่งจากผลของลำดับภาพวีดิทัศน์ตัวอย่างจากกล้องวีดิทัศน์ตัวสองที่แสดงในตารางที่ 4.5 แสดงให้เห็นว่า สีที่มองด้วยตาเปล่านั้นอาจจะมีความแตกต่างกันอย่างเห็นได้ไม่ชัดเจน จึงจำเป็นต้องเก็บค่าข้อมูลสีจากแต่ละจุดภาพมาทำการเปรียบเทียบ ซึ่งจะได้ข้อมูลที่ถูกต้องแม่นยำมากกว่า

4. 2 ผลการจำลองระบบในส่วนของขั้นตอนวิธีการติดตามบุคคล

ดังที่กล่าวมาแล้วในบทที่ 3 ขั้นตอนการติดตามบุคคลเป้าหมายที่วิทยานิพนธ์ฉบับนี้ให้ความสำคัญคือ ขั้นตอนวิธีการย้ายตามค่าเฉลี่ย ซึ่งมีข้อดีคือการคำนวณที่ไม่ซับซ้อน เหมาะสำหรับการใช้งานประมวลผลในเวลาจริง แต่ด้วยข้อจำกัดบางประการการติดตามบุคคลโดยอาศัยขั้นตอนวิธีการย้ายตามค่าเฉลี่ยในอดีตนั้น ยังไม่สามารถรองรับการติดตามบุคคลได้ดีในหลากหลายสถานการณ์ เช่น เมื่อความสว่างของบริเวณที่บุคคลเป้าหมายอยู่นั้นมีการเปลี่ยนแปลงอย่างกะทันหัน หรือเมื่อบุคคลเป้าหมายถูกบังด้วยข้อมูลสีที่เหมือนกันกับข้อมูลสีของบุคคลเป้าหมายที่ใช้ในการติดตาม ทำให้การติดตามนั้นผิดพลาดได้

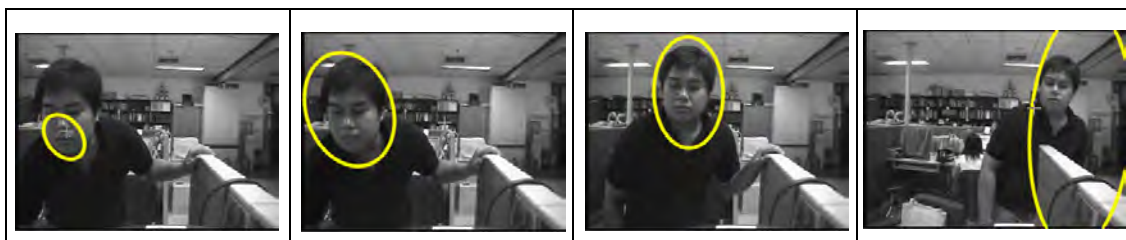
ในหัวข้อนี้ จึงขอแสดงผลการจำลองระบบการติดตามบุคคลด้วยขั้นตอนวิธีการย้ายตามค่าเฉลี่ยในหลายสถานการณ์ เพื่อรับรองได้ว่าขั้นตอนวิธีการย้ายตามค่าเฉลี่ยที่นำมาใช้ในวิทยานิพนธ์ฉบับนี้รองรับกับสถานการณ์ใดๆ ได้ดี และจำเป็นต้องมีการปรับปรุงเพิ่มเติมในส่วนของคุณลักษณะสำคัญใดให้เหมาะกับการทำงานติดตามบุคคลที่มีประสิทธิภาพมากยิ่งขึ้น

4.2.1 ผลการทดสอบเมื่อใช้คุณลักษณะสำคัญทางสีที่ปริภูมิสีต่างกัน

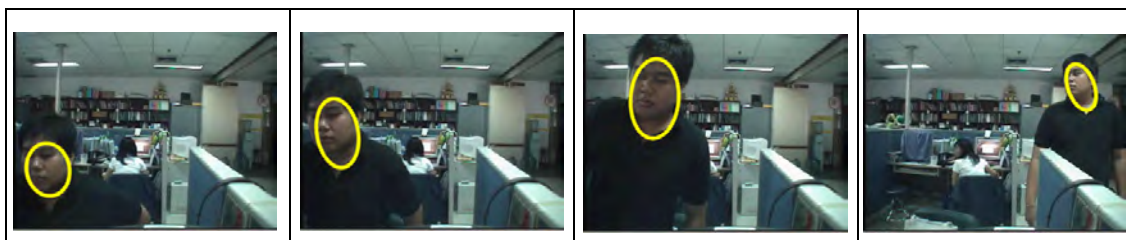
ในอดีตการติดตามวัตถุด้วยขั้นตอนวิธีย้ายตามค่าเฉลี่ยนั้นจะอาศัยเพียงคุณลักษณะสำคัญทางสีในปริภูมิระดับสีเทาเท่านั้น เนื่องจากฟังก์ชันความหนาแน่นที่ถูก

ประมาณขึ้นนั้นอยู่ใน 1 มิติ แต่ด้วยข้อจำกัดในการใช้งานจริงที่เป็นค่าสีที่มองเห็นได้จริง ข้อมูลของปริภูมิระดับสีเทาจึงไม่เหมาะสมกับการนำมาใช้งานจริง แสดงลำดับภาพตัวอย่างเปรียบเทียบของผลของการติดตามบุคคลเป้าหมายเมื่อใช้คุณลักษณะสำคัญทางสีที่ระดับสีเทาและระดับสีในปริภูมิ HSV (แซลแนล H) ได้ดังรูปที่ 4.4

จากการเปรียบเทียบรูปที่ 4.4 (ก) เนื่องจากต้องการติดตามบุคคลโดยใช้สีผิว บริเวณใบหน้าของบุคคลเป้าหมายเป็นคุณลักษณะสำคัญ จะเห็นได้ว่าเนื่องจากสีในโทนสีผิวนั้น เมื่อถูกแปลงให้อยู่ในระดับสีเทาแล้วจะมีค่าข้อมูลที่ใกล้เคียงกับสีขาวหรือบริเวณที่มีความสว่าง ซึ่งอาจจะคล้ายกับสีของสิ่งแวดล้อม ทำให้การติดตามมีความผิดพลาด เส้นกรอบของบริเวณที่ติดตามจึงเคลื่อนที่เข้าหาบริเวณที่มีความสว่างมากภายในลำดับภาพวิดีโอ



(ก) คุณลักษณะสำคัญทางสีในปริภูมิระดับสีเทา



(ข) คุณลักษณะสำคัญทางสีในปริภูมิ HSV (เฉพาะแซลแนล H)

รูปที่ 4.4 ลำดับภาพตัวอย่างเปรียบเทียบของผลการติดตามบุคคลเป้าหมาย
เมื่อใช้คุณลักษณะสำคัญทางสีที่ปริภูมิสีต่างกัน

และจากรูปที่ 4.4 (ข) เมื่อทดลองเปลี่ยนคุณลักษณะสำคัญของบุคคลเป้าหมายที่ใช้ในการติดตามเป็นข้อมูลสีแซลแนล H ในปริภูมิสี HSV พบว่าการติดตามมีอิสระต่อสีของสิ่งแวดล้อมมากขึ้นเนื่องจากระดับสีในแซลแนล H เป็นข้อมูลดึงเฉพาะสีออกมาวิเคราะห์เท่านั้น อีกทั้งระดับของข้อมูลสียังมีความแตกต่างกันมากกว่าระดับข้อมูลสีในปริภูมิระดับสีเทา (ค่าสีในแซลแนล H แบ่งเป็นระดับออกเป็นองศา $0-360^{\circ}$ จึงมีความแตกต่างของข้อมูลสีถึง 360 ข้อมูล) จึงทำให้การติดตามมีประสิทธิภาพมากกว่าการใช้ปริภูมิระดับสีเทาเป็นคุณลักษณะสำคัญ

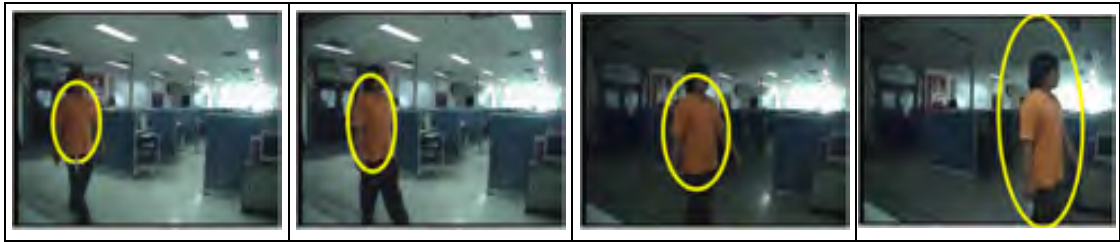
4.2.2 ผลการทดสอบเมื่อใช้คุณลักษณะสำคัญทางสีและขอบ

จากการทดลองในหัวข้อที่ 4.2.1 จะเห็นได้ว่าเมื่อเปรียบเทียบการใช้คุณลักษณะสำคัญทางสีในปริภูมิระดับสีเทาและปริภูมิสี HSV สำหรับบุคคลเป้าหมายแล้ว พบว่าคุณลักษณะสำคัญทางสีในปริภูมิสี HSV จะมีความผิดพลาดในการติดตามที่น้อยกว่า จึงเหมาะสมกับการนำมาใช้งาน แต่ด้วยข้อจำกัดของการทำงานร่วมกันระหว่างกล้องวีดิทัศน์ตามที่วิทยานิพนธ์ฉบับนี้นำเสนอ นั่นคือ ความสว่างจะมีผลต่อค่าสีที่กล้องวีดิทัศน์ทั้งสองตัวในระบบเพื่อนำไปติดตามบุคคลร่วมกันแม้ว่าจะใช้ข้อมูลสีในแชนแนล H ในปริภูมิสี HSV ก็ตาม

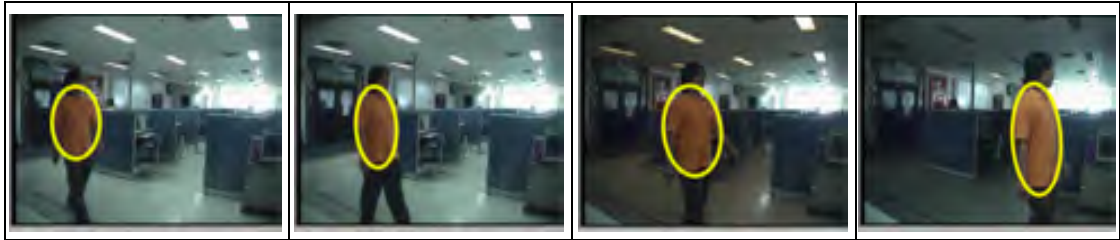
วิธีแก้ไขในขั้นตอนแรก คือ การเทียบมาตรฐานทางสีให้กับกล้องวีดิทัศน์ ดังที่กล่าวไว้ในหัวข้อ แต่เนื่องด้วยถ้าภายในกล้องวีดิทัศน์ตัวใดตัวหนึ่ง ความสว่างได้ถูกเปลี่ยนแปลงอย่างกะทันหันภายในมุมมองของมันเองแล้ว จะไม่สามารถนำการเทียบมาตรฐานทางสีดังกล่าวมาใช้งานได้ จึงจำเป็นต้องเพิ่มคุณลักษณะสำคัญบางประการเพื่อเพิ่มประสิทธิภาพและแก้ปัญหาการติดตามบุคคลในกรณีที่เกิดความผิดพลาดนี้ นั่นคือ การเพิ่มข้อมูลทางขอบเป็นคุณลักษณะสำคัญของบุคคลเป้าหมายด้วย ดังที่วิทยานิพนธ์ฉบับนี้เสนอ

รูปที่ 4.5 (ก) แสดงการติดตามบุคคลเป้าหมายด้วยคุณลักษณะสำคัญทางสีในแชนแนล H ของปริภูมิสี HSV ในสถานการณ์ความสว่างภายในบริเวณที่ติดตั้งกล้องวีดิทัศน์มีการเปลี่ยนแปลงอย่างกะทันหัน เปรียบเทียบกับรูปที่ 4.5 (ข) แสดงการติดตามบุคคลเป้าหมายด้วยคุณลักษณะสำคัญทางสีในแชนแนล H ของปริภูมิสี ร่วมกับคุณลักษณะสำคัญทางขอบ ซึ่งผลการทดสอบเห็นว่า เมื่อนำคุณลักษณะทางขอบเข้ามาช่วยในการติดตามบุคคลเป้าหมายแล้ว แม้มีการเปลี่ยนแปลงของค่าสีของบุคคลเป้าหมายอันเนื่องมาจากความสว่างที่เปลี่ยนแปลงอย่างกะทันหันแล้ว ผลการติดตามบุคคลก็ยังสามารถทำงานได้เป็นปกติ

นอกจากการทดสอบความสว่างที่เปลี่ยนแปลงอย่างกะทันหันแล้ว ในกรณีที่บุคคลเป้าหมายมีการเปลี่ยนทิศทางการเดินทำให้ข้อมูลทางขอบขอบบุคคลเป้าหมายอาจมีการเปลี่ยนแปลงไปด้วย ซึ่งอาจจะมีผลต่อการติดตามอย่างหลีกเลี่ยงไม่ได้ แต่ในวิทยานิพนธ์ฉบับนี้ให้ความสำคัญกับข้อมูลทางขอบน้อยกว่าความสำคัญของข้อมูลทางสี (ฟังก์ชันความหนาแน่นที่ใช้เป็นตัวแทนบุคคลเป้าหมายจะประกอบจาก ฟังก์ชันความหนาแน่นของข้อมูลสี 16 ถึงข้อมูล และฟังก์ชันความหนาแน่นของข้อมูลทางขอบ 8 ถึงข้อมูล) ผลการทดสอบสถานการณ์ดังกล่าว แสดงดังรูปที่ 4.6 (ก) และ (ข) เมื่อใช้คุณลักษณะสำคัญทางสีในแชนแนล H ของปริภูมิสี HSV อย่างเดียว และใช้คุณลักษณะสำคัญทางสีในแชนแนล H ของปริภูมิสี HSV ร่วมกับคุณลักษณะสำคัญทางขอบ ตามลำดับ



(ก) คุณลักษณะสำคัญทางสีแซลแนล H เพียงอย่างเดียว

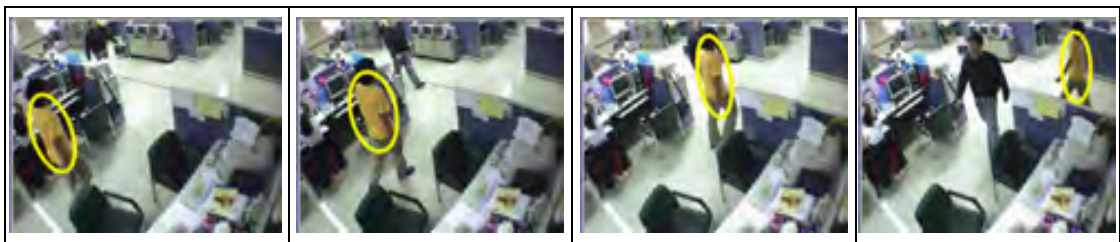


(ข) คุณลักษณะสำคัญทางสีแซลแนล H และขอบ

รูปที่ 4.5 ลำดับภาพวีดิทัศน์ตัวอย่างเปรียบเทียบของผลการติดตามบุคคลเป้าหมาย
เมื่อความส่องสว่างมีการเปลี่ยนแปลงอย่างกะทันหัน

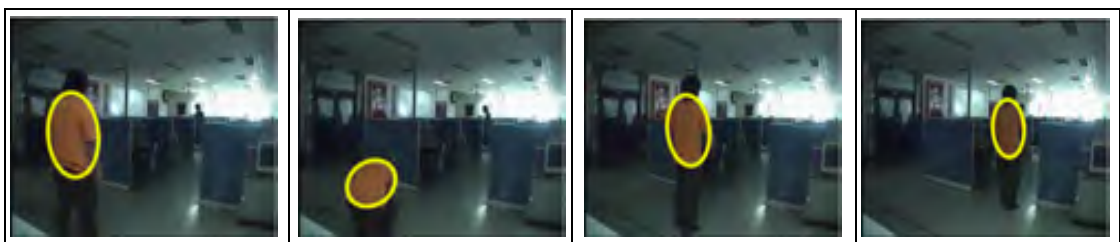


(ก) คุณลักษณะสำคัญทางสีเพียงอย่างเดียว



(ข) คุณลักษณะสำคัญทางสีและขอบ

รูปที่ 4.6 ลำดับภาพวีดิทัศน์ตัวอย่างเปรียบเทียบของผลการติดตามบุคคลเป้าหมาย
เมื่อบุคคลเป้าหมายมีการเปลี่ยนทิศทางการเดิน



(ก) คุณลักษณะสำคัญทางสีเพียงอย่างเดียว



(ข) คุณลักษณะสำคัญทางสีและขอบ

รูปที่ 4.7 ลำดับภาพตัวอย่างเปรียบเทียบของผลการติดตามบุคคลเป้าหมาย
เมื่อบุคคลเป้าหมายมีการเปลี่ยนรูปร่างอย่างกะทันหัน

จากรูปที่ 4.7 ทำการทดสอบสถานการณ์ที่บุคคลเป้าหมายมีการเปลี่ยนรูปร่างอย่างกะทันหัน ซึ่งผลการทดสอบนั้นพบว่า การติดตามบุคคลเป้าหมายนั้นยังคงดำเนินต่อไปได้ ทั้งการใช้คุณลักษณะสำคัญทางสีในแชนแนล H ของปริภูมิสี HSV อย่างเดียว และใช้คุณลักษณะสำคัญทางสีในแชนแนล H ของปริภูมิสี HSV ร่วมกับคุณลักษณะสำคัญทางขอบ ดังรูปที่ 4.7 (ก) และ รูปที่ 4.7 (ข) ตามลำดับ

4.3 ผลการจำลองระบบในส่วนการทำงานร่วมกันระหว่างกล้องวีดิทัศน์ PTZ

การจำลองระบบในส่วนการทำงานร่วมกันระหว่างกล้องวีดิทัศน์ PTZ นั้นจะทำการทดลอง โดยแบ่งออกเป็นสองส่วน ได้แก่ การทดลองเพื่อแสดงผลการประมวลผลร่วมกันเพื่อเปรียบเทียบฟังก์ชันความหนาแน่นที่ได้จากคุณลักษณะสำคัญทางสีที่ได้ถูกประมวลขึ้น ซึ่งเป้าหมายหลักของการทดลองนี้เพื่อแสดงให้เห็นว่า เมื่อมีการส่งต่อข้อมูลระหว่างกล้องวีดิทัศน์ระหว่างกันแล้ว ข้อมูลนั้นๆ ควรมีความถูกต้องเพียงใด ทั้งก่อนและหลังทำการเทียบมาตรฐานทางสี และในส่วนที่สองจะทำการทดสอบการทำงานร่วมกันของกล้องวีดิทัศน์ทั้งสองตัว เพื่อทำหน้าที่ส่าย ก้มเงยและซูมเพื่อติดตามเป้าหมายบุคคลที่ต้องการให้อยู่ตำแหน่งกึ่งกลางจอภาพ

ซึ่งในการจำลองระบบการประมวลผลในเวลาจริง ขนาดภาพวีดิทัศน์ที่ใช้ในการประมวลผลนั้นมีขนาดความยาว 320 จุดภาพ และความกว้าง 240 จุดภาพ โดยโปรแกรมกระบวนการทำงานทุกอย่างด้วยภาษาซี (C programming) (ยกเว้นการคำนวณในส่วนของการเทียบมาตรฐานทางสีจะโปรแกรมการทำงานด้วย MATLAB) และนำคลังโปรแกรม (Program library) ของอินเทลที่เรียกว่า OpenCV มาใช้เป็นเครื่องมือหลักในการโปรแกรมทางการเชื่อมต่อกับกล้องวีดิทัศน์ โดยคุณสมบัติของเครื่องคอมพิวเตอร์ที่ใช้ในการประมวลผล คือ เครื่องคอมพิวเตอร์รุ่น Pentium 4 ความเร็วในการประมวลผล 2.8 GHz ความจุแรม 512 MB บริเวณที่

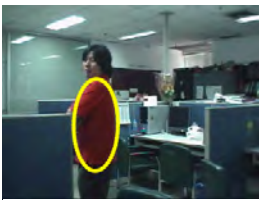
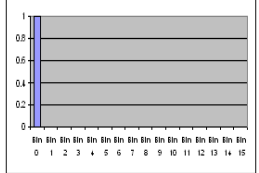

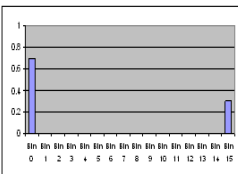
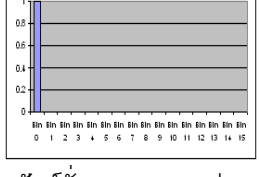
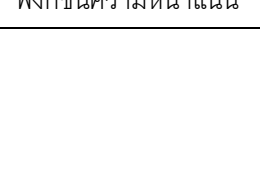

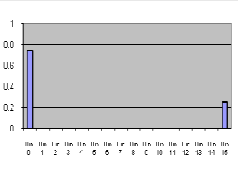
ใช้สำหรับการทดลองจะเป็นบริเวณภายในห้องปฏิบัติการวิจัยกรรมวิธีสัญญาณดิจิทัล ภายในสมมุติฐานที่มีการติดตั้งกล้องวีดิทัศน์ต่างบริเวณกันและไม่มีมุมมองของกล้องวีดิทัศน์ทั้งสองที่ไม่ซ้อนทับกัน

4.3.1 ผลการทดลองในเชิงปริมาณในส่วนการทำงานร่วมกันระหว่างกล้องวีดิทัศน์

PTZ

เนื่องจากคุณลักษณะสำคัญที่ใช้ในการแสดงเป็นตัวแทนของบุคคลเป้าหมาย คือ ฟังก์ชันความหนาแน่นที่ถูกประมาณขึ้นด้วยข้อมูลทางสีและข้อมูลทางขอบ ดังนั้นเมื่อจำเป็นต้องมีการส่งข้อมูลระหว่างกล้องวีดิทัศน์ทั้งสองตัวในระบบ ให้ทำงานร่วมกันแล้ว จึงจำเป็นต้องให้ความสำคัญกับความถูกต้องของข้อมูล ซึ่งในกรณีนี้คือ ความคล้ายกันของฟังก์ชันความหนาแน่นที่เป็นตัวแทนของบุคคลเป้าหมายที่กล้องวีดิทัศน์ตัวที่ 1 และตัวที่ 2 ขั้นตอนวิธีการวัดความคล้ายของฟังก์ชันความหนาแน่นนั้นในวิทยานิพนธ์ฉบับนี้จะถูกแสดงด้วยค่าสัมประสิทธิ์ Bhattacharyya ตัวอย่างเช่น เมื่อข้อมูลค่าสัมประสิทธิ์นี้มีค่าน้อยแสดงว่าความคล้ายของฟังก์ชันความหนาแน่นระหว่างสองฟังก์ชันมีค่าน้อย (ทฤษฎีและการคำนวณในบทที่ 2 และ 3 ตามลำดับ)

ตารางที่ 4.6 ผลการเปรียบเทียบความคล้ายของฟังก์ชันความหนาแน่นของตัวแทนบุคคลเป้าหมายที่กล้องวีดิทัศน์ตัวที่ 1 และตัวที่ 2 ก่อนและหลังทำการเทียบมาตรฐานทางสี

	ตัวอย่างลำดับภาพจากสถานการณ์ต่างๆ		ฟังก์ชันความหนาแน่นของค่าสี (แชนแนล H) ในบริเวณบุคคลเป้าหมาย	ค่าสัมประสิทธิ์ Bhattacharyya ระหว่างกล้องวีดิทัศน์ 2 ตัว
1. เปรียบมาตรฐานจากกล้องวีดิทัศน์ตัวที่ 2 เข้าหา	  <p>ฟังก์ชันความหนาแน่น</p>	 <p>ก่อนเทียบมาตรฐานทางสี</p>		0.836988
ตัวที่ 1	  <p>ฟังก์ชันความหนาแน่น</p>	 <p>หลังเทียบมาตรฐานทางสี</p>		0.860973

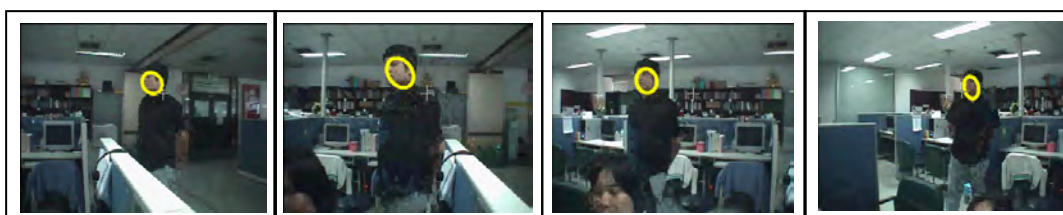
ตัวอย่างลำดับภาพจากสถานการณ์ต่างๆ (ต่อ)	ฟังก์ชันความหนาแน่นของค่าสี (แชนแนล H) ในบริเวณบุคคลเป้าหมาย (ต่อ)			ค่าสัมประสิทธิ์ Bhattacharyya ระหว่างกล้องวิดีโอ 2 ตัว (ต่อ)
2. เทียบมาตรฐานจากกล้องวิดีโอตัวที่ 1 เข้าหากล้องวิดีโอตัวที่ 2	  <p>ฟังก์ชันความหนาแน่น</p>	 <p>ก่อนเทียบมาตรฐานทางสี</p>  <p>หลังเทียบมาตรฐานทางสี</p>	 	0.749449
3. เทียบมาตรฐานจากกล้องวิดีโอสองตัวเมื่อความสว่างต่างกัน	  <p>ฟังก์ชันความหนาแน่น</p>	 <p>ก่อนเทียบมาตรฐานทางสี</p>  <p>หลังเทียบมาตรฐานทางสี</p>	 	0.949111

ตารางที่ 4.6 ผลการเปรียบเทียบความคล้ายของฟังก์ชันความหนาแน่นของตัวแทนบุคคลเป้าหมายที่กล้องวิดีโอตัวที่ 1 และตัวที่ 2 ด้วยค่าสัมประสิทธิ์ Bhattacharyya ก่อนและหลังทำการเทียบมาตรฐานทางสี

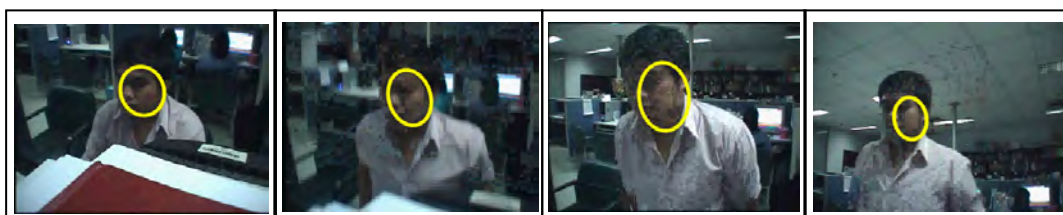
4.3.2 ผลการทดลองในเชิงคุณภาพในส่วนการทำงานร่วมกันระหว่างกล้องวิดีโอ 2 ตัว PTZ

รูปที่ 4.8 แสดงตัวอย่างของการส่าย ก้มเงย ของกล้องวิดีโอ เพื่อให้บุคคลเป้าหมายที่สนใจอยู่บริเวณกึ่งกลางภาพ ซึ่งปัญหาที่เกิดขึ้นกับการควบคุมตำแหน่งการเปลี่ยน

มุมมองของกล้องวิดีโอ ก็ยังมีอยู่บ้าง นั่นคือ เมื่อบุคคลมีการเคลื่อนที่อยู่ตลอดเวลา ขณะที่กำลังสั่งให้กล้องเปลี่ยนมุมมอง จะเกิดผลกระทบเกี่ยวกับตำแหน่งของบุคคลที่ไม่อยู่ในบริเวณกึ่งกลางภาพ ดังที่คาดหวัง แต่จะมีการเยื้องตำแหน่งบ้าง ดังรูปที่ 4.8(ก) และ 4.8 (ข) จะเห็นว่าแม้จะตำแหน่งจะไม่อยู่กึ่งกลางแต่ก็ยังอยู่ในกรอบภาพวิดีโอและยังคงสามารถติดตามเป้าหมายบุคคลได้ อีกทั้งยังอยู่ในมุมมองที่ดีขึ้นกว่ามุมมองเดิมก่อนการเปลี่ยนมุมมองกล้องวิดีโออีกด้วย



(ก) แสดงภาพจากกล้องวิดีโอเมื่อสายมุมมองเพื่อติดตามบุคคลเป้าหมาย



(ข) แสดงภาพจากกล้องวิดีโอเมื่อก้มเงยมุมมองเพื่อติดตามบุคคลเป้าหมาย

รูปที่ 4.8 แสดงผลการควบคุมกล้องวิดีโอเพื่อเปลี่ยนมุมมองเพื่อติดตามบุคคลเป้าหมายที่สนใจ

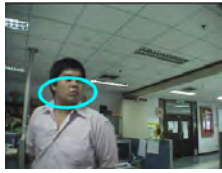
นอกจากการทดสอบการควบคุมการถ่าย ก้มเงยของกล้องวิดีโอให้ติดตามบุคคลเป้าหมายแล้ว ยังมีการทดสอบการทำงานร่วมกันระหว่างกล้องวิดีโอ นั่นคือ การทดสอบเมื่อบุคคลเป้าหมายเดินเข้าสู่มุมมองของกล้องวิดีโอตัวที่หนึ่งแล้ว กล้องวิดีโอตัวที่หนึ่งจะทำการถ่าย ก้มเงยตามบุคคลเป้าหมายนั้นๆ จนหมดขอบเขตการถ่าย ก้มเงยของกล้องวิดีโอ หลังจากนั้นเมื่อบุคคลเป้าหมายคนเดิมเริ่มเดินเข้าสู่มุมมองของกล้องวิดีโอตัวที่สอง กล้องวิดีโอตัวที่สองจะทำการถ่าย ก้มเงยติดตามบุคคลเป้าหมายคนเดิมอย่างต่อเนื่องต่อไป

ซึ่งการทดสอบในหัวข้อนี้จะแบ่งการทดสอบดังกล่าวโดยแบ่งออกเป็นสี่หัวข้อดังนี้ ทดสอบเปรียบเทียบการติดตามบุคคลเป้าหมายด้วยคุณลักษณะสำคัญทางสีในแซลแนล H ของปริภูมิสี HSV เพียงอย่างเดียว ทั้งก่อนและหลังทำการเทียบมาตรฐานทางสี แสดงผลการทดสอบด้วยลำดับภาพวิดีโอตัวอย่างดังตารางที่ 4.7 การทดสอบในส่วนที่สองจะทดสอบการถ่าย ก้มเงยของกล้องวิดีโอทั้งสองตัวเพื่อติดตามบุคคลเป้าหมายอย่างต่อเนื่อง โดยใช้คุณลักษณะสำคัญทางสีในปริภูมิระดับสีเทา แสดงผลการทดสอบด้วยลำดับภาพวิดีโอตัวอย่างดังตารางที่







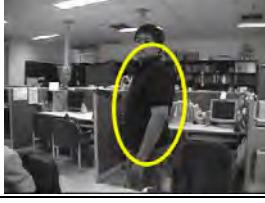

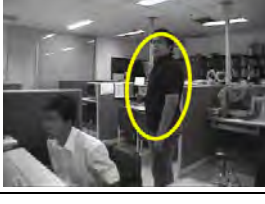
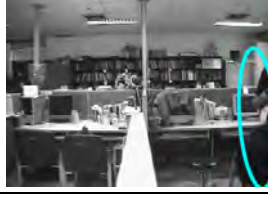
4.8 และการทดสอบในส่วนที่สามจะทดสอบแสดงการส่าย ก้มเงยของกล้องวีดิทัศน์ทั้งสองตัวเพื่อติดตามบุคคลเป้าหมายอย่างต่อเนื่อง โดยใช้คุณลักษณะสำคัญทางสีในแซลแนล H ของปริภูมิสี HSV เพียงอย่างเดียว แสดงผลการทดสอบนี้ด้วยลำดับภาพวีดิทัศน์ตัวอย่าง แสดงดังตารางที่ 4.9 และการทดสอบในส่วนสุดท้ายจะทดสอบแสดงการส่าย ก้มเงยของกล้องวีดิทัศน์ทั้งสองตัวเพื่อติดตามบุคคลเป้าหมายอย่างต่อเนื่อง โดยใช้คุณลักษณะสำคัญทางสีในแซลแนล H ของปริภูมิสี HSV ร่วมกับการใช้คุณลักษณะสำคัญทางขอบ แสดงผลการทดสอบดังตารางที่ 4.10

ตารางที่ 4.7 แสดงลำดับภาพวีดิทัศน์ตัวอย่างเปรียบเทียบการทดลองการติดตามบุคคลเป้าหมายก่อนและหลังทำการเทียบมาตรฐานทางสีให้กับกล้องวีดิทัศน์

สถานการณ์	กล้องวีดิทัศน์ตัวที่ 1	กล้องวีดิทัศน์ตัวที่ 2 ก่อนเทียบ มาตรฐานทางสี	กล้องวีดิทัศน์ตัวที่ 2 หลังเทียบ มาตรฐานทางสี
มุมมองภาพปกติ			
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้ามุมมองของกล้องวีดิทัศน์ตัวที่ 1			
บุคคลเป้าหมายเคลื่อนที่ภายในมุมมองของกล้องวีดิทัศน์ตัวที่ 1			
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้ามุมมองของกล้องวีดิทัศน์ตัวที่ 2			
กล้องวีดิทัศน์ตัวที่ 2 ทำการส่าย ก้มเงย ติดตามบุคคลเป้าหมาย เพื่อให้บุคคลเป้าหมาย			

สถานการณ์ (ต่อ)	กล้องวีดิทัศน์ตัวที่ 1 (ต่อ)	กล้องวีดิทัศน์ตัวที่ 2 ก่อนเทียบ มาตรฐานทางสี(ต่อ)	กล้องวีดิทัศน์ตัวที่ 2 หลังเทียบ มาตรฐานทางสี(ต่อ)
กล้องวีดิทัศน์ทั้งสองตัวทำ การติดตามบุคคลเป้าหมาย อย่างต่อเนื่อง			

ตารางที่ 4.8 แสดงลำดับภาพวีดิทัศน์ตัวอย่างการถ่าย ก้มเงยเพื่อติดตามบุคคลเป้าหมายอย่างต่อเนื่องด้วยกล้องวีดิทัศน์ทั้งสองตัวในระบบด้วยสีในปริภูมิระดับสีเทา

สถานการณ์	กล้องวีดิทัศน์ตัวที่ 1	กล้องวีดิทัศน์ตัวที่ 2
มุมมองภาพปกติ		
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้ามามุมมอง ของกล้องวีดิทัศน์ตัวที่ 1		
บุคคลเป้าหมายเคลื่อนที่ภายในมุมมอง ของกล้องวีดิทัศน์ตัวที่ 1		
กล้องวีดิทัศน์ตัวที่ 1 ทำการถ่าย ก้มเงย ติดตามบุคคลเป้าหมาย เพื่อให้บุคคล เป้าหมายบริเวณกึ่งกลาง		
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้ามามุมมอง ของกล้องวีดิทัศน์ตัวที่ 2		

สถานการณ์ (ต่อ)	กล้องวีดิทัศน์ตัวที่ 1 (ต่อ)	กล้องวีดิทัศน์ตัวที่ 2 (ต่อ)
กล้องวีดิทัศน์ตัวที่ 2 ทำการส่าย ก้มเงย ติดตามบุคคลเป้าหมาย		
กล้องวีดิทัศน์ทั้งสองตัวทำการติดตามบุคคลเป้าหมายอย่างต่อเนื่อง		
กล้องวีดิทัศน์ทั้งสองตัวทำการติดตามบุคคลเป้าหมาย		

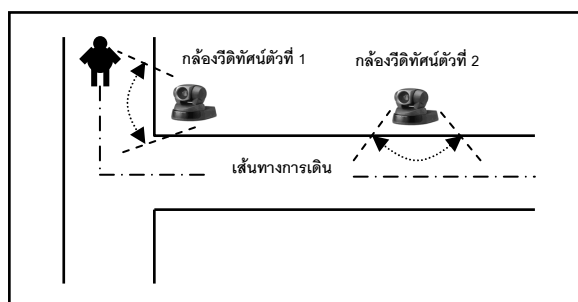
ตารางที่ 4.9 แสดงลำดับภาพวีดิทัศน์ตัวอย่างการส่าย ก้มเงยเพื่อติดตามบุคคลเป้าหมายอย่างต่อเนื่องด้วยกล้องวีดิทัศน์ทั้งสองตัวในระบบด้วยสีในแซลแนล H ของปริภูมิสี HSV

สถานการณ์	กล้องวีดิทัศน์ตัวที่ 1	กล้องวีดิทัศน์ตัวที่ 2
มุมมองภาพปกติ		
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้ามามุมมองของกล้องวีดิทัศน์ตัวที่ 1		
บุคคลเป้าหมายเคลื่อนที่ภายในมุมมองของกล้องวีดิทัศน์ตัวที่ 1		
กล้องวีดิทัศน์ตัวที่ 1 ทำการส่าย ก้มเงย ติดตามบุคคลเป้าหมาย เพื่อให้บุคคลเป้าหมายบริเวณกึ่งกลาง		

สถานการณ์ (ต่อ)	กล้องวิดิทัศน์ตัวที่ 1 (ต่อ)	กล้องวิดิทัศน์ตัวที่ 2 (ต่อ)
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้ามามุมมองของกล้องวิดิทัศน์ตัวที่ 2		
กล้องวิดิทัศน์ตัวที่ 2 ทำการส่าย ก้มเงย ติดตามบุคคลเป้าหมาย เพื่อให้บุคคลเป้าหมายอยู่บริเวณกึ่งกลาง		
กล้องวิดิทัศน์ทั้งสองตัวทำการติดตามบุคคลเป้าหมายอย่างต่อเนื่อง		

จากการทดลองเกี่ยวกับการติดตามบุคคลเป้าหมายอย่างต่อเนื่องระหว่างกล้องวิดิทัศน์สองตัวในระบบนั้น ได้แบ่งการทดลองออกเป็นสถานการณ์ต่างๆ เพื่อยืนยันว่า การติดตามบุคคลเป้าหมายนั้นจะมีความถูกต้องแม่นยำแน่นอน ตัวอย่างสถานการณ์ที่นำมาทดสอบได้แก่ เมื่อมีบุคคลเป้าหมายคนเดียวเดินเข้ามาในระบบ แสดงดังรูปที่ 4.9 และตารางที่ 4.10 เมื่อบุคคลเป้าหมายเดินเข้าสู่มุมมองกล้องวิดิทัศน์ตัวที่ 1 แต่ไม่เดินเข้าสู่มุมมองกล้องวิดิทัศน์ตัวที่ 2 แสดงดังรูปที่ 4.10 และตารางที่ 4.11 เมื่อบุคคลเป้าหมายเดินเข้าสู่ระบบพร้อมกับบุคคลที่ไม่ใช่บุคคลเป้าหมายในทิศทางเดียวกันและตรงกันข้ามกัน แสดงดังรูปและตารางที่ 4.12 และ 4.13 ตามลำดับ

ในสถานการณ์ทดลองที่ 1 บุคคลเป้าหมายคนเดียวเดินเข้ามาในระบบ แสดงดังรูปที่ 4.9 และตารางที่ 4.10



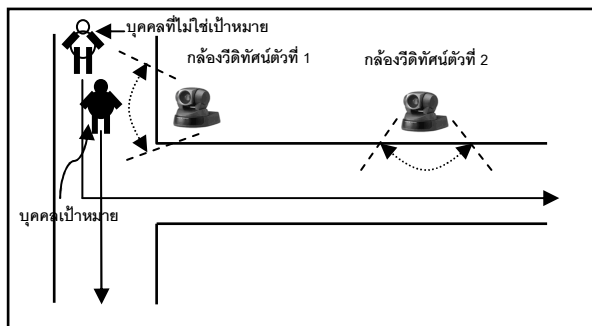
รูปที่ 4.9 แสดงทิศทางการเดินของบุคคลเป้าหมาย

ตารางที่ 4.10 แสดงลำดับภาพวีดิทัศน์ตัวอย่างการส่าย ก้มเงยเพื่อติดตามบุคคลเป้าหมายอย่างต่อเนื่องด้วยกล้องวีดิทัศน์ทั้งสองตัวในระบบด้วยสีในแซดแนล H และขอบ

สถานการณ์	กล้องวีดิทัศน์ตัวที่ 1	กล้องวีดิทัศน์ตัวที่ 2
มุมมองภาพปกติ		
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้ามามุมมองของกล้องวีดิทัศน์ตัวที่ 1		
บุคคลเป้าหมายเคลื่อนที่ภายในมุมมองของกล้องวีดิทัศน์ตัวที่ 1		
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้ามามุมมองของกล้องวีดิทัศน์ตัวที่ 2		
กล้องวีดิทัศน์ตัวที่ 2 ทำการส่าย ก้มเงยติดตามบุคคลเป้าหมาย เพื่อให้บุคคลเป้าหมายอยู่บริเวณกึ่งกลาง		
กล้องวีดิทัศน์ทั้งสองตัวทำการติดตามบุคคลเป้าหมายอย่างต่อเนื่อง		

สถานการณ์ (ต่อ)	กล้องวีดิทัศน์ตัวที่ 1 (ต่อ)	กล้องวีดิทัศน์ตัวที่ 2 (ต่อ)
กล้องวีดิทัศน์ทั้งสองตัวทำการติดตามบุคคลเป้าหมายอย่างต่อเนื่อง		

ทดลองในสถานการณ์ที่ 2 เมื่อบุคคลเป้าหมายเดินเข้าสู่มุมมองกล้องวีดิทัศน์ตัวที่ 1 แต่ไม่เดินเข้าสู่มุมมองกล้องวีดิทัศน์ตัวที่ 2 และมีบุคคลที่ไม่ใช่เป้าหมายเดินเข้าสู่มุมมองกล้องวีดิทัศน์ตัวที่ 2 แทน แสดงดังรูปที่ 4.10 และตารางที่ 4.11



รูปที่ 4.10 แสดงทิศทางการเดินของบุคคลเป้าหมายและบุคคลที่ไม่ใช่เป้าหมาย เมื่อบุคคลเป้าหมายไม่เดินทางเข้าสู่มุมมองของกล้องวีดิทัศน์ตัวที่ 2

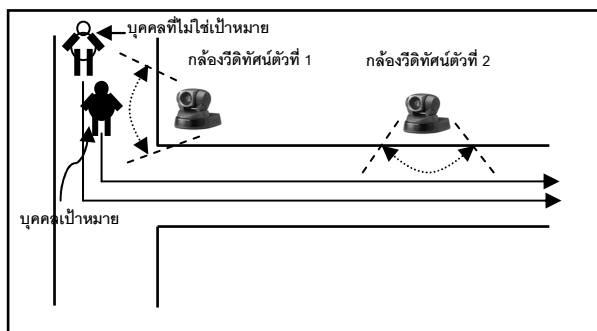
ตารางที่ 4.11 แสดงสถานการณ์ที่บุคคลเป้าหมายเคลื่อนที่เข้ากล้องวีดิทัศน์ตัวที่ 1 แต่ไม่เคลื่อนที่เข้ากล้องวีดิทัศน์ตัวที่ 2

สถานการณ์	กล้องวีดิทัศน์ตัวที่ 1	กล้องวีดิทัศน์ตัวที่ 2
มุมมองภาพปกติ		
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้าสู่มุมมองของกล้องวีดิทัศน์ตัวที่ 1		

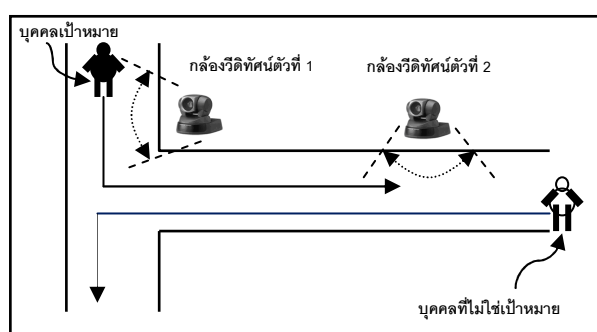
สถานการณ์ (ต่อ)	กล้องวีดิทัศน์ตัวที่ 1 (ต่อ)	กล้องวีดิทัศน์ตัวที่ 2 (ต่อ)
บุคคลเป้าหมายเคลื่อนที่ภายในมุมมองของกล้องวีดิทัศน์ตัวที่ 1		
บุคคลที่ไม่ใช่บุคคลเป้าหมายเคลื่อนที่เข้ามามุมมองของกล้องวีดิทัศน์ตัวที่ 2		
กล้องวีดิทัศน์ตัวที่ 2 ไม่มีการส่ายก้มเงย เพื่อติดตามบุคคลที่ไม่ใช่บุคคลเป้าหมาย		
กล้องวีดิทัศน์ตัวที่ 2 ไม่มีการส่ายก้มเงย เพื่อติดตามบุคคลที่ไม่ใช่บุคคลเป้าหมาย		

จากตารางที่ 4.11 แสดงให้เห็นได้ว่าเมื่อมีบุคคลที่ไม่ใช่เป้าหมายเดินเข้าสู่กล้องวีดิทัศน์ตัวที่ 2 กล้องวีดิทัศน์ตัวที่ 2 จะไม่ทำการติดตามและส่าย ก้มเงย เพื่อติดตามบุคคลที่ไม่ใช่เป้าหมาย แสดงให้เห็นว่าระบบสามารถทำงานได้อย่างถูกต้องติดตามบุคคลเป้าหมายเท่านั้น

ในสถานการณ์การทดลองที่ 3 เมื่อบุคคลเป้าหมายเดินเข้าสู่ระบบพร้อมกับบุคคลที่ไม่ใช่บุคคลเป้าหมายในทิศทางเดียวกัน แสดงดังรูปที่ 4.11 และตารางที่ 4.12 และในสถานการณ์ทดลองที่ 4 เมื่อบุคคลเป้าหมายเดินเข้าสู่ระบบพร้อมกับบุคคลที่ไม่ใช่บุคคลเป้าหมายในทิศทางตรงกันข้ามกัน แสดงดังรูปที่ 4.12 และตารางที่ 4.13



รูปที่ 4.11 แสดงทิศทางการเดินของบุคคลเป้าหมายและบุคคลที่ไม่ใช่เป้าหมาย
เมื่อมีทิศทางการเดินเดียวกัน



รูปที่ 4.12 แสดงทิศทางการเดินของบุคคลเป้าหมายและบุคคลที่ไม่ใช่เป้าหมาย
เมื่อมีทิศทางการเดินตรงข้ามกัน

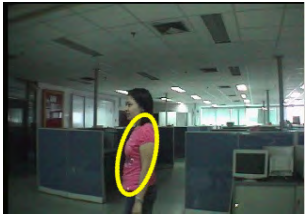

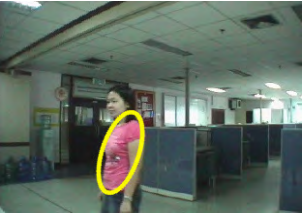

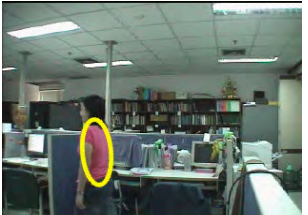
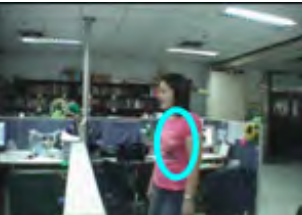
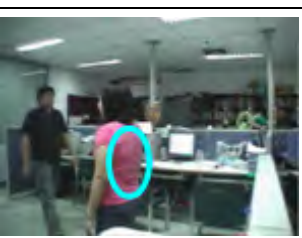
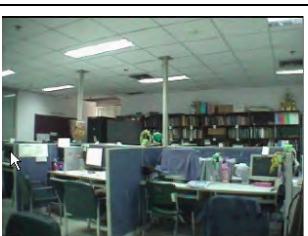
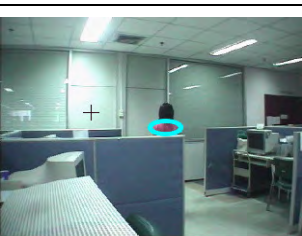
ตารางที่ 4.12 แสดงสถานการณ์ที่บุคคลเป้าหมายเคลื่อนที่เข้ากล้องวีดิทัศน์ตัวที่ 1 และมีและไม่มีบุคคลที่ไม่ใช่เป้าหมายเคลื่อนที่เข้ากล้องวีดิทัศน์ตัวที่ 2 ในทิศทางเดียวกัน

สถานการณ์	กล้องวีดิทัศน์ตัวที่ 1	กล้องวีดิทัศน์ตัวที่ 2
มุมมองภาพปกติ		
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้ามุมมองของกล้องวีดิทัศน์ตัวที่ 1		
บุคคลเป้าหมายเคลื่อนที่ภายในมุมมองของกล้องวีดิทัศน์ตัวที่ 1		

สถานการณ์ (ต่อ)	กล้องวีดิทัศน์ตัวที่ 1 (ต่อ)	กล้องวีดิทัศน์ตัวที่ 2 (ต่อ)
บุคคลที่ไม่ใช่บุคคลเป้าหมาย เคลื่อนที่เข้ามามุมมองของกล้องวีดิ ทัศน์ตัวที่ 1 พร้อมบุคคลเป้าหมาย		
บุคคลที่ไม่ใช่บุคคลเป้าหมาย เคลื่อนที่เข้ามามุมมองของกล้องวีดิ ทัศน์ตัวที่ 2 พร้อมบุคคลเป้าหมาย		
กล้องวีดิทัศน์ตัวที่ 2 ทำการติดตาม บุคคลเป้าหมายอย่างต่อเนื่อง		
กล้องวีดิทัศน์ตัวที่ 2 ทำการติดตาม บุคคลเป้าหมายอย่างต่อเนื่อง		

ตารางที่ 4.13 แสดงสถานการณ์ที่บุคคลเป้าหมายเคลื่อนที่เข้ากล้องวีดิทัศน์ตัวที่ 1 และมีบุคคลที่ไม่ใช่เป้าหมายเคลื่อนที่เข้ากล้องวีดิทัศน์ตัวที่ 2 ในทิศทางตรงกันข้าม

สถานการณ์	กล้องวีดิทัศน์ตัวที่ 1	กล้องวีดิทัศน์ตัวที่ 2
มุมมองภาพปกติ		
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้า มุมมองของกล้องวีดิทัศน์ตัวที่ 1		

สถานการณ์ (ต่อ)	กล้องวีดิทัศน์ตัวที่ 1 (ต่อ)	กล้องวีดิทัศน์ตัวที่ 2 (ต่อ)
บุคคลเป้าหมายเคลื่อนที่ภายในมุมมองของกล้องวีดิทัศน์ตัวที่ 1		
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้ามุมมองของกล้องวีดิทัศน์ตัวที่ 2		
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้ามุมมองของกล้องวีดิทัศน์ตัวที่ 2		
กล้องวีดิทัศน์ทั้งสองตัวทำการติดตามบุคคลเป้าหมายอย่างต่อเนื่อง		
กล้องวีดิทัศน์ตัวที่ 2 ติดตามบุคคลเป้าหมายอย่างต่อเนื่อง		

จากการทดลองเกี่ยวกับการติดตามบุคคลอย่างต่อเนื่องระหว่างกล้องวีดิทัศน์ทั้งสองตัวพบว่า ทุกสถานการณ์ที่ทำการทดลองยืนยันได้ว่ากล้องวีดิทัศน์ตัวที่ 2 ในระบบจะทำการติดตามเฉพาะบุคคลเป้าหมายที่ถูกต้องเท่านั้น แม้ในสถานการณ์ที่มีการรบกวนจากบุคคลที่ไม่ใช่บุคคลเป้าหมายเดินเข้าสู่ระบบ

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

วิทยานิพนธ์ฉบับนี้ได้นำเสนอระบบการติดตามบุคคลอย่างต่อเนื่องด้วยการใช้กล้องวิดีโอทัศนประเภท PTZ 2 ตัวทำงานร่วมกัน ซึ่งที่วิทยานิพนธ์ฉบับนี้ให้ความสนใจในการศึกษาแบ่งออกเป็นสามส่วน คือ การเทียบมาตรฐานทางสีให้กับกล้องวิดีโอทัศน ขั้นตอนวิธีการติดตามบุคคล และการส่งต่อข้อมูลระหว่างกล้องวิดีโอทัศนรวมถึงควบคุมกล้องวิดีโอทัศนเพื่อให้ติดตามบุคคลได้อย่างต่อเนื่อง สรุปผลการวิจัยทั้งสามส่วนนี้ได้ดังนี้

สรุปผลการวิจัยเกี่ยวกับการเทียบมาตรฐานทางสีให้กับกล้องวิดีโอทัศน ขั้นตอนวิธีการเทียบมาตรฐานทางสีที่นำเสนอนี้ เป็นการประมาณความสัมพันธ์ทางสีระหว่างกล้องวิดีโอทัศนสองตัวในลักษณะเชิงเส้น ซึ่งถูกแทนด้วยเมทริกซ์ของการเทียบมาตรฐานทางสี เมื่อมีการแปลงข้อมูลทางสีจากกล้องวิดีโอทัศนตัวหนึ่งไปหาอีกกล้องวิดีโอทัศนอีกตัวหนึ่งแล้วด้วยเมทริกซ์ของการเทียบมาตรฐานทางสีแล้วนั้น จากผลการทดลองพบว่าข้อมูลสีจากกล้องวิดีโอทัศนทั้งสอง จะมีค่าข้อมูลสีเดียวกันที่มีความใกล้เคียงมากขึ้น สังเกตได้จากผลการทดลองที่ค่าเฉลี่ยความผิดพลาดกำลังสองที่ลดลง แต่เนื่องจากข้อมูลสีในปริภูมิ RGB มีความสัมพันธ์ไม่เป็นเชิงเส้นกัน จึงทำให้ผลการทดสอบยังไม่ได้มากขึ้นเท่าที่พอใจ แต่ยังมีใจได้ว่า การทำการเทียบมาตรฐานทางสีที่นำเสนอนี้ ทำให้ข้อมูลค่าสีระหว่างกล้องวิดีโอทัศนสองตัวที่ติดตั้งต่างบริเวณกันรับค่าสีเดียวกันแล้วได้ข้อมูลที่ใกล้เคียงกันมากขึ้น

สรุปผลการวิจัยเกี่ยวกับขั้นตอนวิธีการติดตามบุคคลที่นำเสนอ ขั้นตอนวิธีสำหรับการติดตามบุคคลที่วิทยานิพนธ์นำเสนอ คือ ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยซึ่งมีพื้นฐานมาจากการประมาณฟังก์ชันความหนาแน่นของข้อมูลคุณลักษณะสำคัญ โดยเลือกใช้คุณลักษณะสำคัญในสองประการคือ ข้อมูลคุณลักษณะสำคัญทางสีและทางขอบของบุคคลเป้าหมาย ซึ่งจากผลการทดลองพบว่า ในสถานการณ์ต่างๆที่ให้ความสนใจ เช่น เมื่อความสว่างมีการเปลี่ยนแปลงอย่างกะทันหัน บุคคลเป้าหมายมีการเปลี่ยนทิศทางการเดิน หรือบุคคลเป้าหมายมีการเปลี่ยนรูปร่าง เป็นต้นนี้ พบว่าการนำคุณลักษณะสำคัญทั้งสองที่กล่าวมาแล้วมาประมาณเป็นฟังก์ชันความหนาแน่น สามารถรองรับการติดตามบุคคลด้วยขั้นตอนวิธีการย้ายค่าเฉลี่ยได้ดี เมื่อเทียบกับการใช้คุณลักษณะสำคัญด้วยสีเพียงอย่างเดียวดังที่มีงานวิจัยในอดีต อีกทั้งเนื่องจากข้อดีของปริภูมิสี

HSV ในแชลแนล H คือ ข้อมูลของสีจะแยกออกจากข้อมูลทางความสว่างอย่างเป็นอิสระ จึงทำให้ การเลือกข้อมูลสีปริภูมิสี HSV ในแชลแนล H ของวิทยานิพนธ์ฉบับนี้มีประสิทธิภาพมากยิ่งขึ้นด้วย

ส่วนของการส่งต่อข้อมูลระหว่างกล้องวิดีโอทัศนวิสัยรวมถึงควบคุมกล้องวิดีโอทัศนวิสัยเพื่อให้ติดตาม บุคคลได้อย่างต่อเนื่องนั้น เป้าหมายของการส่งต่อข้อมูลระหว่างกันจำเป็นอย่างยิ่งยิ่งที่ข้อมูล คุณลักษณะสำคัญที่จะนำมาเป็นตัวแทนบุคคลเป้าหมายนั้นควรมีความคล้ายกันระหว่างกล้องวิดีโอ ทัศนวิสัยสองตัว วิธีการเทียบมาตรฐานทางสีดังที่กล่าวมาแล้วจึงเป็นวิธีการหนึ่งที่แก้ปัญหาได้ อีกทั้ง แม้ในสถานการณ์ที่ข้อมูลคุณลักษณะทางสีทำให้เป้าหมายบุคคลมีความผิดพลาด ก็จะสามารถ นำคุณลักษณะทางขอบของบุคคลเป้าหมายมาช่วยแก้ปัญหาได้เช่นเดียวกัน และในส่วนของการ ควบคุมกล้องวิดีโอทัศนวิสัยนั้น สรุปได้ว่า สามารถควบคุมให้กล้องวิดีโอทัศนวิสัยทั้งสองตัวทำงานเพื่อให้ บุคคลเป้าหมายอยู่กึ่งกลางลำดับภาพวิดีโอทัศนวิสัยหรือภายในขอบเขตที่ไม่มีเคลื่อนที่ของกล้องวิดีโอ ทัศนวิสัยได้จริง

5.2 อภิปรายผลการวิจัย

ในส่วนของการเทียบมาตรฐานทางสี เพื่อเพิ่มประสิทธิภาพระบบการติดตามบุคคลใน สถานการณ์ใด ทำให้สามารถติดตามบุคคลลดสัญญาณรบกวนที่เกิดจากวัตถุที่มีสีคล้ายกับสีผิวได้ อย่างมาก รวมทั้งการตรวจหาทั้งสองลักษณะสามารถประยุกต์ใช้กับภาพที่รับจากกล้องวิดีโอทัศนวิสัย PTZ ได้ โดยไม่ต้องมีขั้นตอนจัดการใด ๆ ก่อนการประมวลผล และมีการประมวลผลที่ไม่ซับซ้อน เหมาะกับงานประยุกต์ในเวลาจริง แต่ข้อเสีย คือ ความไม่แน่นอนของเค้าโครงบุคคลที่ถูกตรวจหา ได้ที่เกิดจากข้อจำกัดของการตรวจหาด้วยการเคลื่อนที่ เนื่องจากต้องขึ้นอยู่กับลักษณะการ เคลื่อนที่ของบุคคล ณ ขณะนั้น

ในส่วน of ขั้นตอนวิธีการติดตามบุคคลที่นำเสนอ นั้น ขั้นตอนวิธีใดหรือคุณลักษณะ สำคัญใดที่จะเหมาะสมกับการทำงานนั้น จะขึ้นอยู่กับสถานการณ์ที่จะใช้ติดตามบุคคล ขั้นตอนวิธี หนึ่งอาจจะเหมาะสมกับสถานการณ์หนึ่ง หรือคุณลักษณะสำคัญหนึ่งอาจจะไม่เหมาะสมกับการ ทำงานในอีกสถานการณ์หนึ่งเป็นต้น สำหรับวิทยานิพนธ์ฉบับนี้จึงเลือกขั้นตอนวิธีและ คุณลักษณะสำคัญที่เหมาะสมกับสมมติฐานของสถานการณ์ของการติดตั้งกล้องวิดีโอทัศนวิสัยสองตัวที่ ต่างกัน ซึ่งผลการทดลองรับรองได้ว่าสามารถติดตามบุคคลเป้าหมายได้จริง

และส่วนสุดท้ายคือส่วนของการถ่ายและกัมเมยงกล้องวิดีโอทัศนวิสัยเพื่อติดตามบุคคล เนื่องจาก ผลลัพธ์ของการติดตามบุคคลเป้าหมายนั้น คือ ตำแหน่งจุดภาพของบุคคลเป้าหมายในลำดับภาพ วิดีโอทัศนวิสัย ทำให้สามารถควบคุมและสั่งกล้องวิดีโอทัศนวิสัยให้ถ่ายและกัมเมยงได้ด้วยตำแหน่งจุดภาพของ

ภาพวิดีโอที่บันทึก โดยความเร็วในการถ่ายและกัมเมกัล้องวิดีโอที่บันทึกจะขึ้นอยู่กับผลตอบสนองของคำสั่งของกล้องวิดีโอที่ประยุกต์ใช้ แต่กล้องวิดีโอที่บันทึก SONY EVID-100 ที่นำมาใช้ในงานวิจัยนี้ไม่มีปัญหาในส่วนผลกระทบนี้

ดังนั้นระบบโดยรวมที่นำเสนอในวิทยานิพนธ์ฉบับนี้ ถูกทดสอบระบบที่การประมวลผลอัตราลำดับภาพ 15 ลำดับภาพต่อวินาที ซึ่งเป็นอัตราลำดับภาพมาตรฐานของการประมวลผลในเวลาจริง เนื่องจากการใช้คุณลักษณะสำคัญเป็นข้อมูลสี่และชอบสามารถดึงออกมาใช้งานได้ง่าย จึงไม่มีปัญหาในด้านเวลาที่ต้องใช้ในการประมวลผลมากนัก อีกทั้งถึงจะมีการแปลงข้อมูลคุณลักษณะสำคัญเป็นฟังก์ชันความหนาแน่นก็ตาม แต่เวลาที่ต้องใช้ในการประมวลผลก็ยังคงเป็นที่น่าพอใจสำหรับการนำมาใช้งานในเวลาจริง สำหรับปัญหาในการประยุกต์ใช้ระบบในด้านกล้องวิดีโอ PTZ คือ ต้นทุนด้านราคากล้องวิดีโอที่ค่อนข้างสูง ซึ่งแนวทางการแก้ปัญหาจำเป็นต้องมีการออกแบบส่วนควบคุมการถ่ายและกัมเมกัล้องที่สร้างขึ้นเอง จะช่วยลดต้นทุนได้อย่างมาก

5.3 ข้อเสนอแนะ

ในการเทียบมาตรฐานให้กับกล้องวิดีโอที่บันทึกนั้น ถ้าเป็นระบบการทำงานที่ใช้จำนวนกล้องวิดีโอที่บันทึกมากขึ้น อาจจะต้องมีการเพิ่มการเทียบมาตรฐานทางตำแหน่ง อาจจะทำให้ตำแหน่งเป้าหมายบุคคลที่ต้องการติดตามนั้นแม่นยำมากขึ้น และส่วนการทำงานร่วมกันระหว่างกล้องวิดีโอที่บันทึก สามารถพัฒนารูปแบบของการทำงานร่วมกันในลักษณะงานประยุกต์อื่น ๆ ได้ เช่น ใช้ตำแหน่งการเคลื่อนที่ของกล้องวิดีโอแต่ละตัวมาใช้ประโยชน์ในการระบุตำแหน่งบุคคลเพื่อให้ความถูกต้องในการตรวจหาตำแหน่งบุคคลมากขึ้น

รายการอ้างอิง

- [1] S.Lim, A.Elgammal, and S.Davis. Image-Based Pan-Tilt Camera Control in a Multi-Camera Surveillance Environment. Proceedings of the IEEE International Conference on Multimedia & Expo (ICME'03). (July 2003) : 645-648.
- [2] S.Lim, A.Elgammal, and S.Davis. A Scalable Image-Based Multi-Camera Visual Surveillance System. Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance. (July 2003) : 205-212.
- [3] H.Hongo, M.Ohya, M.Yasumoto, Y.Niwa, and K.Yamamoto. Focus of Attention for Face and Hand Gesture Recognition using Multiple Cameras. Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition. (March 2000): 156-161
- [4] P.Amnuaykanchanasin, S.Aramvith, and T.H.Chalidabhongse. Real-Time Face Identification using Two Cooperative Active Cameras. The 9th International Conference on Control, Automation, Robotics and Vision, 2006(ICARCV '06). (December 2006): 1-6
- [5] N.Soontranon, S.Aramvith, and T.H.Chalidabhongse. Face and Hand Localization and Tracking for Sign Language Recognition. International Symposium on Communication and Information Technologies (ISCIT'04). (October 2004): 1246 - 1251
- [6] J.Kang, I.Cohen, and G.Medioni. Continuous Tracking Within and Across Camera Streams. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04). (June 2003): 267-272
- [7] G.Unal, A.Yezz. A Variational Approach to Problems in Calibration of Multiple Cameras. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04). (July 2004): 172-178
- [8] F.Porikli. Inter Cameras Color Calibration by Correlation Model Function. Proceedings of the IEEE International Conference on Image Processing (ICIP'03). 133-136

- [9] F.Porikli, and A.Divakaran. Multi-Camera Calibration Object Tracking and Query Generation. Proceedings of the IEEE International Conference on Multimedia & Expo (ICME'03): 653-656
- [10] R.Collins, A.Lipton, H.Fujiyashi, and T.Kanade. Algorithms for Cooperative Multisensor Surveillance. Proceeding of the IEEE. (October 2001):1456-1477
- [11] M.Greifenhagen, D.Comaniciu, and V.Ramesh. Design Analysis and Engineering of Video Monitoring System. Proceeding of the IEEE. (October 2001):1498-1517
- [12] C.Wren, A.Azarbayejani, T.Darrell, and A.Pentland. PFinder Real time Tracking of human body. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI'97). (July 1997): 780-785
- [13] Y.Boykov, and D.Huttenlocher. Adaptive Bayesian recognition in tracking rigid objects. IEEE Conference on Conference on Computer Vision and Pattern Recognition (CVPR'00). 697-704
- [14] R.Rosales, and S.Sclarof. 3D Trajectory Recovery for Tracking Multiple Objects and Trajectory Guided Recognition of Action. IEEE Conference on Computer vision and Pattern Recognition (CVPR'99). (June 1999): 117-123
- [15] M.Isard, and A.Blake. Condensation-Conditional Density propagation for visual tracking. International Journal of Computer vision. (August 1998)
- [16] D.Comaniciu, V.Ramesh, and P.Meer. Kernel Based Object Tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence. (May 2003) 564-577
- [17] K.Fukunaga, and L.Hosteler. The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Transactions on Information Theory (January 1975): 32-40
- [18] Y.Cheng. Mean shift, mode seeking, and clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence. (August 1995): 790-799
- [19] G. R. Bradski. Computer vision face tracking as a component of a perceptual user interface. Proceeding of The IEEE Workshop on Applications of Computer Vision. (October 1998): 214-219.

- [20] Z.Wen and Z.Cai. Mean Shift algorithm and its Application in Tracking of Objects. Proceeding of The 5th International Conference on Machine Learning and Cybernetics. (August 2005)
- [21] K.Bai and W.Liu. Improve Object Tracking with Particle Filter and Mean Shift Algorithm. Proceeding of IEEE International Conference on Automation and Logistic. (August 2007): 431-435
- [22] H.Liu, Z.Yu and H.Zha. Robust Mean Shift Tracking Based on Multi-cue Integration.IEEE International Conference on Systems Man and Cybernetics(SMC'06). (October 2006): 5160-5166

ภาคผนวก

ดัชนีคำศัพท์

algorithm	ขั้นตอนวิธี
affine transform	การแปลงสัมพรรค
analog	แอนะล็อก
background subtraction	ลบฉากหลัง
bhattacharyya coefficient	สัมประสิทธิ์ Bhattacharyya
bhattacharyya distance	ระยะ Bhattacharyya
cooperative	ทำงานร่วมกัน
camera control	การควบคุมกล้อง
color bin	ถังสี
camera motion	การเคลื่อนที่ของกล้อง
camera model	แบบจำลองกล้อง
computer vision	คอมพิวเตอร์วิทัศน์
column	สดมภ์
detection	การตรวจหา
fixed camera	กล้องอยู่กับที่
field of view	ขอบเขตภาพ
frame differencing	ผลต่างระหว่างเฟรม
frame rate	อัตราเฟรม
focal length	ความยาวโฟกัส
face matching	การจับคู่ใบหน้า
face recognition	การรู้จำใบหน้า
false negative	ความผิดพลาดแบบลบ
false positive	ความผิดพลาดแบบบวก
ground truth	ผลถูกต้องพื้นฐาน
histogram	ฮิสโตแกรม
human vision system	ระบบการมองเห็นของมนุษย์
input	รับเข้า

luminance component	องค์ประกอบแสง
monitor	เฝ้าสังเกต
multi-camera system	ระบบกล้องหลายตัว
mean	ค่าเฉลี่ย
morphological image processing	การประมวลผลภาพเชิงสัญญาณ
motion vector	เวกเตอร์การเคลื่อนที่
motion detection	การตรวจหาการเคลื่อนที่
normalize	นอร์มัลไลซ์
non parametric model	แบบจำลองชนิดไม่มีตัวแปร
opening operator	ตัวดำเนินการแบบเปิด
PTZ video camera	กล้องวีดิทัศน์ PTZ
Person identify	ระบุบุคคล
person silhouette	โครงร่างบุคคล
pixel	จุดภาพ
pan	ถ่าย
real time	เวลาจริง
recognition	การรู้จำ
row	แถว
RGB color space	ปริภูมิสี RGB
RGB	แดง เขียว น้ำเงิน
surveillance	การสอดส่องดูแล
skin color	สีผิว
skin color detection	การตรวจหาสีผิว
spatial	เชิงพื้นที่
skin color model	แบบจำลองสีผิว
tilt	ก้มเงย
threshold	ขีดเริ่มเปลี่ยน
YCbCr color space	ปริภูมิสี YCbCr
zoom	ซูม

บทความทางวิชาการที่ได้รับการเผยแพร่

1. N. Bumrungrat, S. Aramvith, and T.H. Chalidabhongse. CONTINUOUS PERSON TRACKING ACROSS MULTIPLE CAMERAS WITH COLOR CALIBRATION, International Workshop on Advanced Image Technology (IWAIT'08), Taiwan, (January 2008)
2. N. Bumrungrat, S. Aramvith, and T.H. Chalidabhongse, CONTINUOUS PERSON TRACKING ACROSS MULTIPLE ACTIVE CAMERAS USING SHAPE AND COLOR CUES. International Workshop on Advanced Image Technology (IWAIT'09), Korea, (January 2009)

CONTINUOUS PERSON TRACKING ACROSS MULTIPLE CAMERAS WITH COLOR CALIBRATION

N. Bumrungrat^{}, S. Aramvith^{*}, and T.H. Chalidabhongse[†]*

^{*}Department of Electrical Engineering
Chulalongkorn University
Bangkok 10330 Thailand
Tel: +66-2218-6909

[†]Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok 10520 Thailand
Tel: +66-2737-2551 Ext.526

ABSTRACT

This paper proposes a framework for handover method with camera color calibration for continuously detecting and tracking a person of interest across cooperative pan-tilt-zoom (PTZ) cameras. Multiple Color distribution of person is used to register the target person across cameras. For each camera, the human face is detected, segmented, and tracked using motion and skin color cues. To handover the detection and tracking process among two cameras, the second camera receives color cues of a target person from the first camera. To achieve the color matching across the different radiometric characteristic cameras with different viewpoints, linear color calibration is proposed to help with handover process. Our experimental results demonstrate color calibration helps reduce color mismatch between two cameras and the system is capable for continuously and accurately track the target person across cameras.

Keywords: Color Calibration, Multi-Camera system, tracking, handover.

1. INTRODUCTION

With increasing processor power and dropping cost of cameras, more attention has been given to develop a real time smart surveillance system. In general, processes in surveillance system can be divided into three main functions: human detection, human tracking and human identification or recognition. In the past, most of the systems were based on a static single camera. Recently, more researchers have been interested in employing multiple active cameras for detecting, tracking, and recognizing people such as the works presented in [1,2,3]. By using active or PTZ cameras, we are able to enlarge the coverage area of the surveillance as well as zoom into the detected target to obtain finer resolution for further analysis. The work presented in [1] was a wide area surveillance system using three cooperative cameras that can detect, zoom, and track two human targets which the system process on only stationary camera and using background subtraction method to find position parameters and then

send its to controlling the PTZ camera. The works in [2] and [3] used four cameras in which two are assigned as fixed stereo camera to estimate face positions and the other two are assigned as active camera for detection and tracking.

Some surveillance system such as the works presented in [4, 5] continuously track person across multiple cameras by handover the tracked person from one camera to another camera. The work in [4] proposed a continuous tracking within and across stationary and PTZ cameras. They use motion and color models for registering the object between overlapping camera view. The works in [5] use a homographic relation mapping between overlapping static cameras to transfer a target object point from one camera to another. This work also can track multiple objects within camera view.

The work in [6,7] propose tracking multiple people in a multi-camera system which can be divided into two part: the first part is tracking multiple people in a single camera they using motion detection and ellipse algorithm to detect a new person in surveillance area. And then in the second part is tracking multiple people in a multi - cameras they use template matching and ellipse matching to track the person.

The color-based approaches base the matching essentially on the color of the tracks, as in [8] where a color space invariant to illumination changes is proposed and using histogram-based information at regions and blobs are exploited to solve match tracks.

The major problem of such systems that need to handover the tracked objects between cameras is the color mismatch due to dissimilar radiometric characteristics of the cameras. Recently the work in [9-10] presented the approaches for color calibration problem using the pixel mapping techniques. They found the pixel mapping function by minimum cost path of the correlation matrix of color histogram between 2 cameras under various lighting.

And the work in [11] proposed color calibration techniques for multi-cameras system using variation parameters, This variation parameters purpose the differences in absolute colors measured in the response of each camera are modeled by a simple multiplicative factor in each of color RGB channel measurements and an

additive offset parameter.

The aim of our work is to develop a real-time surveillance system for uncontrolled indoor environments such as in typical offices, labs, apartments, etc. In this paper, we extend our work [12] by proposing a framework for handover method with camera color calibration for continuously detecting and tracking a person of interest across cooperative PTZ cameras. We employ color and motion information of the tracked person in doing handover between the two neighbor cameras. To achieve the color matching across the different radiometric characteristic cameras with different viewpoints, we proposed a method that employs linear color calibration. Color correlation transform matrix between a pair of color sets from the two cameras is computed using linear regression method. The experimental result showed that the proposed system succeeds in person detection and tracking in uncontrolled background, and the system is capable of continuously track the target person across cameras.

This paper is organized as follows. In Section 2, the overall system is proposed and described. Section 3 presents our proposed algorithm for person detection and tracking. In Section 4, the algorithms for person analysis, color calibration, and cooperative person tracking across cameras are presented. Experimental results are presented in Section 5. Section 6 concludes the paper.

2. SYSTEM OVERVIEW

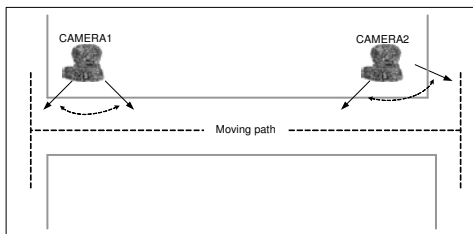


Figure 1 Cameras set up.

Fig. 1 shows our system setup. We connect two SONY EVI-D100 PTZ cameras to a PC and use them to detect and track the target person from different viewpoints. The two cameras are distantly separated and have non-overlapped viewpoint.

The overall architecture is shown in Figure 2. Both cameras simultaneously process the same algorithm. First, the face blob is detected and located using motion and skin color cues. The location of the facial blob is used to calculate pan, tilt and zoom variables for the camera to follow the moving target. In person analysis section, color features of interested target person, i.e., cloth, skin, and hair, are extracted. The processing of each camera must depend on a status of the system that can be divided into 2 cases. First case: If no person is detected in both cameras, both cameras still continuously perform normal detection

and tracking. Second case: if the system detects a target person in any camera, the characteristics of that person will be collected and sent to another camera for the purpose of continuously detect and track the same target person. To enhance the tracking process across cameras, color calibration technique which is further described in Section 4 is integrated into the system. Thus, our system can track and follow a target person effectively throughout all reachable cameras view.

3. PERSON DETECTION AND TRACKING

3.1 Person detection

In order to detect and segment the face blob from video obtaining from a moving camera, the simple background subtraction techniques do not work. Thus, we have to find other visual cues that can detect and localize face with the changing viewpoints. For this research, we employ motion detection to detect moving target silhouette. Then, we apply skin color detection over the detected region to obtain face blob. We then segment the head silhouette by fitting an ellipse.

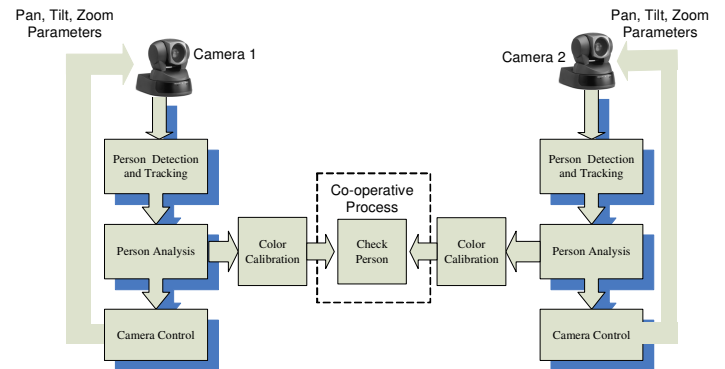


Figure 2 System diagram.

Motion detection technique that we employ in this research is frame differencing due to its computational efficiency. The technique detects any moving pixels by comparing colors of the pixel between consecutive frames. However, the simple frame differencing yields only the boundaries, not the silhouette, of the moving regions. To fix this drawback, we add pixels that horizontally lie between two moving boundaries. The final moving target's segmented silhouette is shown in Fig. 3.

To model skin color, we employ method similar to the ones proposed in [13] which model the skin color in YCbCr color space. During learning period, the skin color is modeled with elliptical function on CbCr space, see Eq. (1). The Y component is discarded to allow the system invariant to brightness variation,

$$\frac{(x-c_x)^2}{a^2} + \frac{(y-c_y)^2}{b^2} = 1 \quad (1)$$

,where x and y are transformed Cb and Cr components, c_x and c_y are shifted center of ellipse, a and b are major and minor axis of the ellipse. We perform skin color detection over the moving target blob that detected from the previous step only. This could reduce a lot of non-skin background noise.

To segment the head region, we fit an elliptical boundary to the segmented blob. The output segmentation includes face and hair regions.

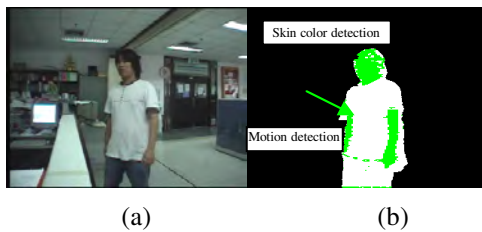


Figure 3 Result of face detection (a) input image (b) motion and skin detected result

3.2 Person tracking

To track the target face, the face detection step is performed over one frame and then is used as a seed to predict face position in the next frame. Because the camera can move and that changes viewpoint, we divide tracking algorithm into 2 cases. First is the case of normal tracking where camera orientation is not changed. The face boundary of the previous frame is expanded both horizontally and vertically and serves as a searching window for the current frame.

The second is the case when camera moves and the viewpoint changes. For this case, base on our camera control algorithm that tries to align the tracked face area in the center of the picture, the new searching window will be set at the center of image.

3.3 Camera control

To keep the moving target in the camera's field of view, pan, tilt and zoom variables are computed and used as parameters to control the motion of the camera. To find pan angle (ψ) and tilt angle (ϕ), we calculate those angles using ratio of pixel-based distance change and degree-based distance change, see Eq.(2) and Eq.(3),

$$\psi = \frac{Diff_x}{\beta_x} = \frac{x_f - x_{center}}{\beta_x} \quad (2)$$

$$\phi = \frac{Diff_y}{\beta_y} = \frac{y_f - y_{center}}{\beta_y} \quad (3)$$

,where $Diff_x$ and $Diff_y$ are distance between center of the face blob (x_f, y_f) and the center of image

(x_{center}, y_{center}) in horizontal and vertical direction respectively. β_x and β_y are values of pixel distance change when camera moves by 1 degree in horizontal (pan) and vertical (tilt) directions.

4. COOPERATIVE PERSON TRACKING ACROSS CAMERAS

After blob of target person has been identified, color features of interested target person, i.e., cloth, skin, and hair, are extracted. The characteristics of that person thus will be collected and sent to another camera for the purpose of continuously detect and track the same target person. To enhance the tracking process across cameras as there exists color mismatch problem across cameras, simple color calibration technique is proposed and is integrated into the cooperative person tracking system.

4.1 Person analysis

To hand-off the detection and tracking process between two cameras, the second camera needs to receive some information of a target person from the first camera. We choose color characteristic of the target human to be a cue for hand-off process because it provides a prominent feature for a non-rigid object tracking.

Color characteristic of our interested target person can be divided in 3 main regions: cloth region, skin region (include face and hands) and hair region. Thus, we propose an appearance-based model using a multiple color distribution for representing any target person.

4.2 Color calibration

To achieve the color matching across the different radiometric characteristic cameras with different viewpoints, we proposed a method that employs linear color calibration. Color correlation transform matrix between a pair of color sets from the two cameras is computed using linear regression method.

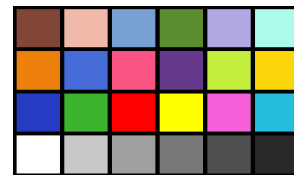


Figure 4 Color checker board used in color calibration

To calibrate the cameras, we use a color checker board shown in Fig.4. A color checker board has 24 reference colors covering the major color palettes in RGB color space. The setup is shown in Fig.5. We adjust the cameras' positions and put the color checker board so that the checker board appears in both cameras' field of views. The images of the checker board are captured by both cameras. In this work, we use the RGB color space due to its being widely accepted as simple and useful probabilistic

models. Then the color correlation transformation matrix between the two cameras is computed according to Eqs. (4) – (9).

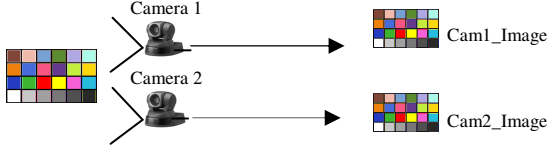


Figure 5 Color calibration setup

Let C^1 (Eq. 4) and C^2 (Eq. 5) be the 24x4 color matrices of the 24 RGB colors taken from checker board images of the camera 1 and camera 2 respectively.

$$C^1 = \begin{bmatrix} C_R^1 \\ C_G^1 \\ C_B^1 \\ 1 \end{bmatrix} = \begin{bmatrix} R_1^1 & R_2^1 & R_3^1 \dots & R_{24}^1 \\ G_1^1 & G_2^1 & G_3^1 \dots & G_{24}^1 \\ B_1^1 & B_2^1 & B_3^1 \dots & B_{24}^1 \\ 1 & 1 & 1 \dots & 1 \end{bmatrix} \quad (4)$$

$$C^2 = \begin{bmatrix} C_R^2 \\ C_G^2 \\ C_B^2 \\ 1 \end{bmatrix} = \begin{bmatrix} R_1^2 & R_2^2 & R_3^2 \dots & R_{24}^2 \\ G_1^2 & G_2^2 & G_3^2 \dots & G_{24}^2 \\ B_1^2 & B_2^2 & B_3^2 \dots & B_{24}^2 \\ 1 & 1 & 1 \dots & 1 \end{bmatrix} \quad (5)$$

where R_i^c , G_i^c , B_i^c are the red, green, and blue matrices of the color region i for camera c . Each matrix is $4 \times n$ where n is the number of pixels in the color region i of the checker board.

$$R_1^1 = \begin{bmatrix} R_{Color1_Camera1_pixel_1} \\ R_{Color1_Camera1_pixel_2} \\ \vdots \\ R_{Color1_Camera1_pixel_n} \end{bmatrix}_{n \times 1}^T \quad (6)$$

$$R_2^1 = \begin{bmatrix} R_{Color2_Camera1_pixel_1} \\ R_{Color2_Camera1_pixel_2} \\ \vdots \\ R_{Color2_Camera1_pixel_n} \end{bmatrix}_{n \times 1}^T \quad (7)$$

Then we can apply linear regression to obtain a color correlation transformation matrix that can be used to convert color of one camera to the color of the other camera. Let W is the corresponding regression coefficient of the color of camera 2 to the color of camera 1. We obtain

$$C^1 = WC^2 \quad (8)$$

Then, the W can be obtained from inversed matrix analysis.

$$W = (C^1 C^{2T}) (C^2 C^{2T})^{-1} \quad (9)$$

The color correlation transformation matrix, W , will be used to transform color from one camera to the color of another camera while we are performing the tracked person handover across the cameras.

4.3 Cooperative Person Tracking Across Cameras using Color Calibration

For tracking the target person across two cooperative cameras, we obtain the color characteristic described in the previous section to identify the target. For the first time that the target person moves into one camera viewpoint, the color characteristic of that successfully detected target person will be collected and accumulated in our system profile. Those characteristics are the mean and variance of cloth, skin and hair color distributions, i.e., $[E, \sigma]_{cloth}$, $[E, \sigma]_{skin}$ and $[E, \sigma]_{hair}$, respectively. If the target person moves out of this camera, the overall mean and variance of each color characteristics are then calculated and sent to another camera. For registering the previous target to the second camera, we compare the color characteristics of an incoming person to those that are sent from the first camera, as shown in Eqs.(10)-(12),

$$| [E_{in}]_{cloth} - [E_{ref}]_{cloth} | < \tau_c \quad (10)$$

$$| [E_{in}]_{skin} - [E_{ref}]_{skin} | < \tau_s \quad (11)$$

$$| [E_{in}]_{hair} - [E_{ref}]_{hair} | < \tau_h \quad (12)$$

,where $E(E^R, E^G, E^B)$ represents the mean of color vectors in RGB color space, E_{in} is a mean of color vector of the incoming frame at the second camera, E_{ref} is an average mean of color vector which calculated from the reference camera (first detected camera). The comparison process is taken in all defined color regions. Thus, there are 3 thresholds, i.e., τ_c , τ_s and τ_h for cloth, skin and hair color regions, respectively, to determine the target person at the second camera.

5. EXPERIMENTAL RESULTS

We test our proposed system on a Pentium 4 2.8 GHz PC running Windows XP. The two SONY EVI-D100 PTZ cameras are connected to the system. Distance between both cameras is about 1 meter. The experiments are done in two parts. The first one is to verify the effectiveness of applying color calibration to enhance color matching between two cameras. The second one is to show the tracking performance improvement of our proposed handover system integrated with color calibration technique.

For the first experimental case, the color checker board was used as an input image to camera 1 and 2 viewpoint, respectively, as stated in Section 4.2. The input images are shown in Fig. 7(a), and 7(b), respectively, for camera 1 and 2. Color mismatch among two cameras is obvious when observed closely. We collect pixel values in RGB color space from each 24-color region. Through the process explained in Section 4.2, we can determine the color correlation transform matrix, as shown in Eq. (9), which will be used to transform the corresponding colors from camera 2 to match with the colors from camera 1. Fig. 7 shows the calibrated images from camera 2. Table 1 shows the value of mean squared error (MSE) of the tested image from camera 1 and 2, before and after applying color calibration. Note that the MSE after color calibration is significantly decreased yielding the closer color matching for camera 1 and 2.

For the second experimental case, we employed the color calibration and color correction as described in Section 4.2 into our person tracking system. This is very crucial to the system because the cameras handover the tracked person using color and motion correlations. The experimental result shown in Table 2 demonstrates that with the color calibration/correction, the target color correlation can be improved. Fig. 9 shows the some snapshots of the videos from both cameras while the system was tracking a target.

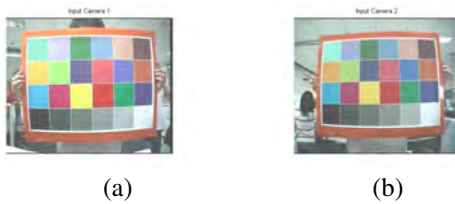


Figure 7 (a) input image from camera 1 (b) show input from camera 2



Figure 8 Calibrated image from camera 2










Table 1 MSE reduction of test image before and after color calibration

COLOR	MSE (Before Calibrate)	MSE (After Calibrate)
1. DARK SKIN	8.67	1.67
2. LIGHT SKIN	305.67	1.67
3. BLUE SKY	993.67	1.67
4. FOLIAGE	463.33	1.67
5. BLUE FLOWER	2127.33	0.67
6. BLUISH GREEN	1.67	0.67
7. ORANGE	37.67	0.67
8. PURPLISH BLUE	279	0.33
9. MODERATE RED	425	0.33
10. PURPLE	283.67	0.67
11. YELLOWGREEN	530.67	1
12. ORANGE YELOW	1165.33	0.67
13. RED	353.67	0
14. GREEN	142	0.67
15. BLUE	75.67	0.67
16. YELLOW	439.33	0.67
17. MAGENTA	339.67	0.67
18. CYAN	630.33	0.33
19. WHITE	50.67	0.33
20. NEUTRAL 8	138	0
21. NEUTRAL 6.5	217	12.33
22. NEUTRAL 5	46	2
23. NEUTRAL 3.5	7.33	0.33
24. BLACK	0.67	0.67

Table 2 MSE reduction of tracking images before and after color calibration

Object	MSE (before calibration)	MSE (after calibration)
Yellow Cloth	22.0	14.4
White Cloth	17.96	2.17
Skin Color	6.26	6.11
Hair Color	50.9	39.9

Figure 9 Some snapshots of the videos from the two cameras showing the handover moment.

	Camera 1	Camera 2
(a) Normal view		
(b) The target person enters the first camera's view.		-
(c) First camera tracks a target.		-
(d) Another person walks into the second camera's view.	-	
(e) After some analysis, the second camera will stop tracking the person that is not target.	-	
(f) The target now enters the second camera's view.	-	
(g) The second camera starts tracking as it is a target person from the first camera.	-	
(e) Continuous tracking for the target.	-	

5. CONCLUSIONS

We have presented a real-time system for continuously detect and track a person across cooperative pan-tilt-zoom active cameras. To handover the detection and tracking process among two cameras, the second camera receives color cues of a target person from the first camera. To achieve the color matching across the different radiometric characteristic cameras with different viewpoints, linear color calibration is proposed to help with handover process. The experimental results confirm color calibration helps reduce color mismatch between two cameras and the system is capable for continuously and accurately track the target person across cameras.

6. ACKNOWLEDGEMENT

This research is in part supported by the Cooperation Project between Department of Electrical Engineering and Private Sector for Research and Development,

Chulalongkorn University.

7. REFERENCES

- [1] S.-N. Lim, A. Elgammal, and L. S. Davis, "Image-Based Pan-Tilt Camera Control in a Multi-Camera Surveillance", in *Proceedings of IEEE International Conference on Multimedia and Expo*, Maryland, pp. 1-645-8, Jul 2003.
- [2] A. Hampapur, S. Pankanti, A. Senior, Y.-L. Tian, L. Brown, and R. Bolle, "Face Cataloger : Multi-Scale Imaging for Relating Identity to Location", in *Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance*, Florida, pp. 13-20, Jul 2003.
- [3] H. Hongo, M. Ohya, M. Yasumoto, Y. Niwa, and K. Yamamoto, "Focus of Attention for Face and Hand Gesture Recognition Using Multiple Cameras", in *Proceedings of Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, Grenoble, pp. 156-161, Mar 2000.
- [4] J. Kang, I. Cohen, and G. Medioni, "Continuous Tracking Within and Across Camera Streams", in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03)*, Wisconsin, Jun 2003.
- [5] J. Black, T. Ellis, and P. Rosin, "Multi View Image Surveillance and Tracking", in *Proceedings of the IEEE Workshop on Motion and Video Computing (MOTION'02)*, Florida, Dec 2002.
- [6] Y. Shan, C. Huang, L. Chen, "Multiple People Visual Tracking in a Multi-Camera System for Cluttered Environments" *Intelligent Robots and Systems*, 2006
- [7] R. Cucchiara1, A. Prati2, R. Vezzani1, "Posture Classification in a Multi-Camera Indoor Environment", *Proceeding of ICIIP 2005*.
- [8] J. Li, C.S. Chua, Y.K. Ho, "Color Based Multiple People Tracking", in *Proc. of IEEE Intl Conf. on Control, Automation, Robotics and Vision*, 1, pp. 309-314, 2002.
- [9] F. Porikli, "Inter Cameras Color Calibration by Correlation Model Function," *Image Processing*, 2003. *ICIP'03*, Vol. 2, pp II - 133-6, 2003.
- [10] F. Porikli, and A. Divakaran., "Multi-Camera Calibration, Object Tracking and Query Generation," *Multimedia and Expo*, 2003. *ICME'03*. Vol. 1, pp I - 653-6, 2003.
- [11] U. Gozde, Y. Anthony, "A Variational Approach to Problems in Calibration of Multiple Cameras", *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on* Volume 29, Issue 8, Aug. 2007
- [12] P. Amnuaykanchanasin, S. Aramvith, T.H. Chalidabhongse, "Continuous person tracking using two cooperative active cameras" ,*Proceeding of IJWISA 2007*, Korea, Feb 2007.
- [13] N. Soontranon, S. Aramvith, and T. H. Chalidabhongse, "Face and Hand Localization and Tracking for Sign Language Recognition", in *International Symposium on Communication and Information Technologies (ISCIT'04)*, Sapporo, Oct 2004.

CONTINUOUS PERSON TRACKING ACROSS MULTIPLE ACTIVE CAMERAS USING SHAPE AND COLOR CUES

N. Bumrungrat^{}, S. Aramvith^{*}, and T.H. Chalidabhongse[†]*

^{*}Department of Electrical Engineering
Chulalongkorn University
Bangkok 10330 Thailand
Tel: +66-2218-6909

[†]Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok 10520 Thailand
Tel: +66-2723-4977

ABSTRACT

This paper proposed a framework for handover method in continuously tracking a person of interest across cooperative pan-tilt-zoom (PTZ) cameras. The algorithm here is based on a robust non-parametric technique for climbing density gradients to find the peak of probability distributions called the mean shift algorithm. Most tracking algorithms use only one cue (such as color). The color features are not always discriminative enough for target localization because illumination or viewpoints tend to change. Moreover the background may be of a color similar to that of the target. In our proposed system, the continuous person tracking across cooperative PTZ cameras by mean shift tracking that using color and shape histogram to be feature distributions. Color and shape distributions of interested person are used to register the target person across cameras. For first camera, we select interested person for tracking using skin color, cloth color and boundary of body. To handover tracking process among two cameras, the second camera receives color and shape cues of a target person from the first camera and using linear color calibration to help with handover process. Our experimental results demonstrate color and shape feature in mean shift algorithm is capable for continuously and accurately track the target person across cameras.

Keywords: Tracking, Handover, Mean shift algorithm, color calibration.

1. INTRODUCTION

The real time object tracking is used in many applications such as video surveillance, human machine interfaces, robot tracking, and intelligent transportation systems. These applications need good object detection and tracking methods which have been progressed in recent years. However, there are still some difficulties in tracking such as the [cases](#) where there are changes in background, view point, or illumination especially when using multiple cameras.

The object tracking in dynamic scene is a typical non-rigid vision tracking problem. The difficulty of the problem with moving cameras is that the environment is changing and the illumination may vary, which will cause the change of the color in the camera view. Appearance-base methods have been investigated for solving this tracking problem. Most appearance-base tracking approaches are based on some kind of representations of image appearance. Target and localization are robust and more efficient for tracker. In this paper, we focus on the problem of tracking an individual person with cooperative PTZ cameras when viewpoint of camera during a long tracking tends to change.

Some surveillance system such as the works presented in [1, 2] continuously track person across multiple cameras by handover the tracked person from one camera to another camera. The work in [1] proposed a continuous tracking within and across stationary and PTZ cameras. They use motion and color models for registering the object between overlapping camera view. The works in [2] use a homographic relation mapping between overlapping static cameras to transfer a target object point from one camera to another. This work also can track multiple objects within camera view. Some systems employ two cameras with one providing a fixed wide view and second Pan-Tilt-Zoom to acquire large face image for tracking such as [3]. The work in [4] proposed the process of building a surveillance system with multiple PTZ cameras installed within an indoor laboratory. The system demonstrated multiple PTZ cameras can hand over tracked target to each other.

Most tracking algorithms fall into two categories. The first category is probabilistic methods. These methods view the tracking algorithm as a state solving problem under the Bayesian framework, model uncertainty and propagate the conditional densities through the tracking process. The representative methods are Kalman filter and its derivatives, Condensation [5], particle filter [6], etc. The second category is deterministic methods. These methods compare a model with current frame and find out the most probable region. Mean Shift [7][8] fall into this category.

The mean shift algorithm has success in object tracking due to its simplicity and robustness. It finds local

minimum of a similarity measure between the color histogram of model and the candidates in the image. The mean shift algorithm is based on appearance model. Color histogram is the most popular feature that has been used to represent the appearance of the target. But color of an object depends on illumination, viewpoint of camera during a long tracking (important when use PTZ cameras system). Thus the single histogram feature is not always discriminative enough. The work in [9] presents object tracking from multiple stationary and moving cameras using color distribution as the main cue for tracking object across views. Since color information can be easily biased by several factor such as illumination, shadow, and appearance change or difference camera setup, color cue may be not very reliable for tracking moving object. Some other works such as the work in [10] proposed multi-feature tracking approach for adaptive real-time object tracking using a stationary camera. In this paper, we propose a tracking method that employ both color and shape features to improve accuracy of the tracker especially when tracking objects under changes in illumination and viewpoints.

The paper is organized as follows. Section 2 we introduce our system overview. In Section 3, we explain how to extract the color and shape features to represent the target. Section 4 discusses the person tracking method. And then discusses person analysis in cooperative process and control both camera in Section 5 and 6 respectively. We experiment the performance of the proposed method in Section 7. This paper is concluded in Section 8.

2. SYSTEM OVERVIEW

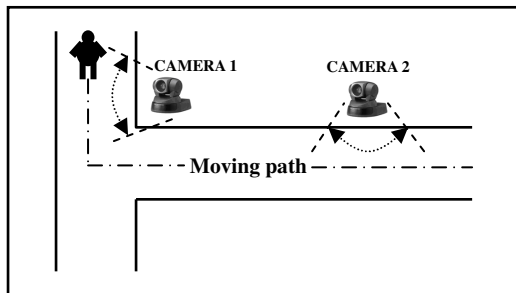


Figure 1 Cameras set up

Fig. 1 shows our system setup. We connect two SONY EVI-D100 PTZ cameras to a PC and use them to track the person from different viewpoints. The two cameras are distantly separated and have non-overlapped viewpoint.

Overall architecture is shown in Fig. 2. Both cameras simultaneously run the same algorithm. First, for each video frame, the raw input image is processed to construct a color probability distribution image via color histogram model. Also, the derivatives in x and y directions of the image is calculated to find shape cue that is represented by orientation histogram of the target person being tracked.

In person tracking step, we employ mean shift algorithm using color and shape features of target person. The location of the target person is use for continuously tracking and also used to calculate pan-tilt variables for the camera control in

following the moving target. In order to continuously track the same target person across different cameras, we also calibrate color using color calibration matrix between two cameras to solve the color mismatch problem. This will be described in details in Section 5. By using color and shape features for tracking, our system can track and follow the target effectively throughout all reachable cameras' view.

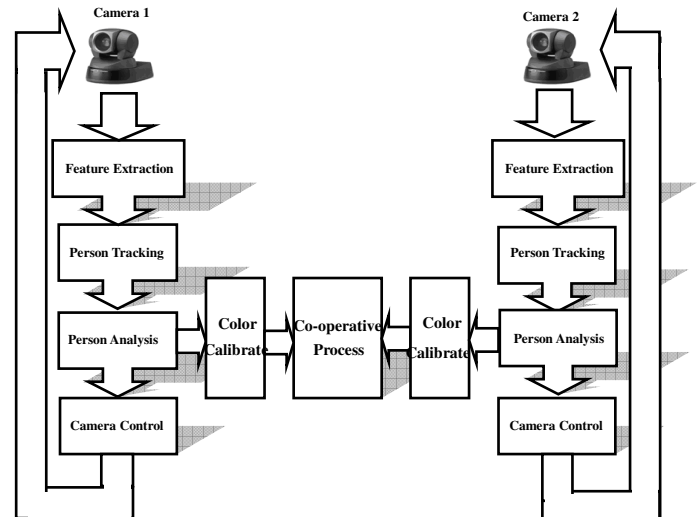


Figure 2 System diagram

3. FEATURE EXTRACTION

Fig. 3 show block diagram of feature extraction process which is composed of two sub-processes, that is finding color and shape features for using in mean shift tracking process.

3.1 Color Probability Distribution

In many person tracking processes, color distributions are used as target representation because of its independence from non-rigidity, scaling and partial occlusions. In order to track color objects in a video scene, a probability distribution image of the desired color (cloth and/or skin color) in the video scene must be created. In this paper, we use hue histogram to represent color of the target because of its brightness invariance. The color distributions of the target and candidate are respectively calculated. The evaluating the similarity of two color distributions is done using the Bhattacharyya distance.

3.2 The Orientation Histogram of Shape

In this work, we represent shape feature of the target using an orientation histogram. It is done by calculating image pixel derivatives in x and y directions. We use a standard Sobel masks (S_x and S_y), as shown in Eq.(1)-(3), because of its simplicity and efficiently computed for real time applications.

$$M_x(x, y) = S_x * I(x, y) \quad (1)$$

$$M_y(x, y) = S_y * I(x, y)$$

(2)

The Orientation of edge is

$$\theta(x, y) = \arctan\left(\frac{M_x(x, y)}{M_y(x, y)}\right) \quad (3)$$

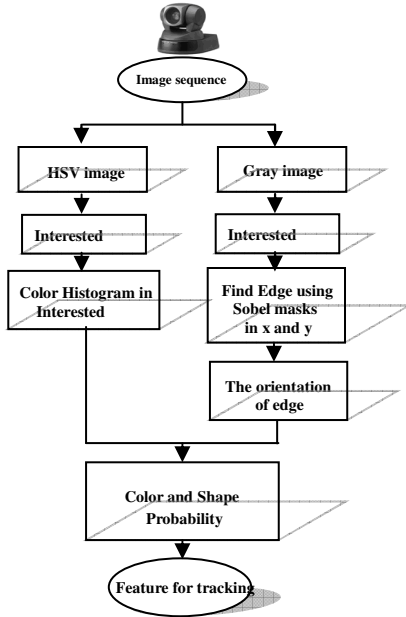


Figure 3 Block diagram of feature extraction process

3. PERSON TRACKING

The mean shift algorithm operates on probability distributions. To track interested person in video frame sequences, the color image data has to be represent as a probability distribution; we use color and shape histogram to accomplish feature distributions derived from video image sequences changed over time. Block diagram of mean shift algorithm is shown in Fig. 4.

The mean shift algorithm is a tracking method that finds the mode of the probability distributions of sample data without any assumptions about the priori distribution. This algorithm avoids choosing a distribution of model and estimating its distribution parameters.

There are several researches have been working on object tracking using mean shift algorithm such [8-9]. The basic mean shift tracking algorithm assumes that the target object has to separate sufficiently from background, but this assumption is not always true especially when tracking is carried out in dynamic backgrounds, e.g., surveillance with a moving camera.

The mean shift tracking finds the location corresponding to the target in the current frame based on the

appearance of the target. Therefore, a similarity measure is needed between the distributions of an interested region in the current frame and the target model. A popular measure between two distributions is the Bhattacharyya distance considering discrete densities such as two feature (color and shape) histogram and the coefficient is calculated by

$$\rho[p, q] = \sum_{bin=1}^m \sqrt{p^{(bin)} q^{(bin)}} \quad (4)$$

The larger ρ is, the more similar the distributions are. For two identical histograms we obtain $\rho = 1$, indicating a perfect match. As distance between two distributions, the measure can be defined as

$$d = \sqrt{1 - \rho[p, q]} \quad (5)$$

which d is the Bhattacharyya distance. The tracking algorithm is to recursively compute of an offset value from the current location y_0 to a new location y_1 according to the mean shift vector y_1 is calculated by using

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i g\left(\frac{y_0 - x_i}{h}\right)}{\sum_{i=1}^{n_h} w_i g\left(\frac{y_0 - x_i}{h}\right)} \quad (6)$$

when

$$w_i = \sum_{u=1}^m \sqrt{\frac{q^{(u)}}{p^{(u)}(y_0)}} \delta[h(x_i) - b_{in}] \quad (7)$$

and

$$g(x) = -k'(x) \quad (8)$$

The person tracking algorithm is shown in Fig 4.

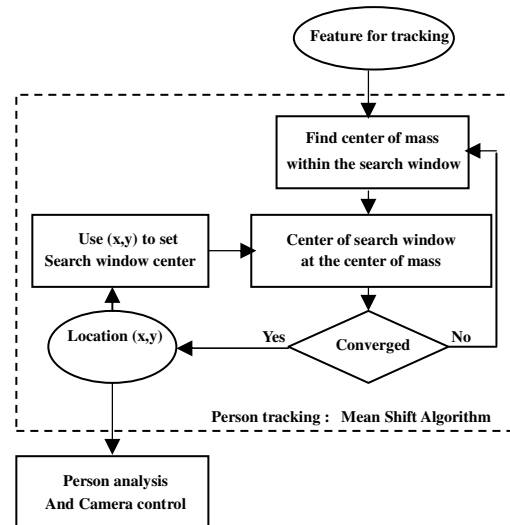


Figure 4 Block diagram of Person tracking with mean shift algorithm

$$p = \{p^{(u)}\}_{u=1}^m \quad q = \{q^{(u)}\}_{u=1}^m$$

5. PERSON MODEL AND COOPERATIVE TRACKING ACROSS CAMERAS

5.1 Person model

In tracking handover between two cameras, the second camera needs to receive some information of the tracked target from the first camera. As we use color and shape distributions of the target person in tracking, we must handover these distributions in hand-off process.

The color characteristic of tracked target person can be divided into 2 main regions: cloth and skin regions, while the shape characteristic is represented by x and y gradients in tracked region.

5.2 Camera color calibration

To achieve the color matching across the different radiometric characteristic cameras with different viewpoints, we proposed a method that employs linear color calibration. Color correlation transform matrix between a pair of color sets from the two cameras is computed using linear regression method.

To calibrate the cameras, we use a color checker board. A color checker board has 24 reference colors covering the major color palettes in RGB color space. The setup is shown in Fig.5. We adjust the cameras' positions and put the color checker board so that the checker board appears in both cameras' field of views. The images of the checker board are captured by both cameras. In this work, we use the RGB color space due to its being widely accepted as simple and useful probabilistic models. Then the color correlation transformation matrix between the two cameras is computed according to Eqs. (10) – (12).

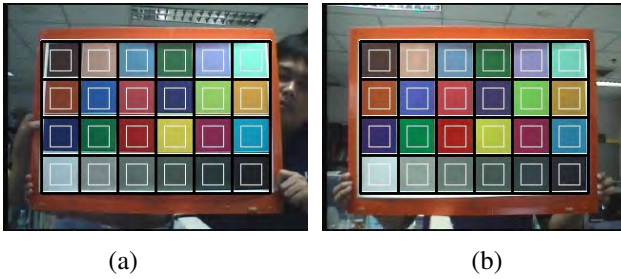


Figure 5 (a) show input image from camera 1 (b) show input image from camera 2

$$R_N^C = \frac{\sum_{k=1}^{400} I_R(k)}{400} \quad (10)$$

$$G_N^C = \frac{\sum_{k=1}^{400} I_G(k)}{400} \quad (11)$$

$$B_N^C = \frac{\sum_{k=1}^{400} I_B(k)}{400} \quad (12)$$

Let R_N^C (Eq. 4), G_N^C (Eq. 5) and B_N^C (Eq. 5) be average color information from each camera in R, G and B channel respectively. And superscript C be number of camera in system, subscript N be number of reference color from checker board and k be pixel of the image.

$$C^1 = \begin{bmatrix} R_1^1 & R_2^1 & \dots & R_{24}^1 \\ G_1^1 & G_2^1 & \dots & G_{24}^1 \\ B_1^1 & B_2^1 & \dots & B_{24}^1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (13)$$

$$C^2 = \begin{bmatrix} R_1^2 & R_2^2 & \dots & R_{24}^2 \\ G_1^2 & G_2^2 & \dots & G_{24}^2 \\ B_1^2 & B_2^2 & \dots & B_{24}^2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (14)$$

Let C^1 (Eq. 7) and C^2 (Eq. 8) be the 24x4 color matrices of the 24 RGB colors taken from checker board images of the camera 1 and camera 2 respectively. Then we can apply linear regression to obtain a color correlation transformation matrix that can be used to convert color of one camera to the color of the other camera. Let W is the corresponding regression coefficient of the color of camera 2 to the color of camera 1. We obtain

$$C^1 = WC^2 \quad (15)$$

Then, the W can be obtained from inversed matrix analysis. The color correlation transformation matrix, W , will be used to transform color from one camera to the color of another camera while we are performing the tracked person handover across the cameras.

6. CAMERA CONTROL

To keep the moving target in the camera's field of view, pan, tilt and zoom variables are computed and used as parameters to control the motion of the camera. To find pan angle (ψ) and tilt angle (ϕ), we calculate those angles using ratio of pixel-based distance change and degree-based distance change, see Eq.(16) and Eq.(17),

$$\psi = \frac{Diff_x}{\beta_x} = \frac{x_f - x_{center}}{\beta_x} \quad (16)$$

$$\phi = \frac{Diff_y}{\beta_y} = \frac{y_f - y_{center}}{\beta_y} \quad (17)$$

where $Diff_x$ and $Diff_y$ are distance between center of the interested region (x_f, y_f) and the center of image in (x_{center}, y_{center}) horizontal and vertical direction respectively. β_x and β_y are values of pixel distance change when camera moves by 1 degree in horizontal (pan) and vertical (tilt) directions.

7. EXPERIMENTAL RESULTS

We tested our proposed system on a Pentium 4 2.8 GHz PC running Windows XP. The two SONY EVI-D100 PTZ cameras are connected to the system. Distance between both cameras is about 1 meter. The experiment is to show the tracking performance of our proposed handover system when integrating shape cue is feature in mean shift tracking algorithm.

Fig. 6 shows person tracking result when using color and shape mean shift algorithm. Fig.6(a) is an input image, and Fig.6(b) is the tracked result. Histograms shown in Fig.6(c)-(e) are color histogram, shape histogram, and the concatenation of them respectively.

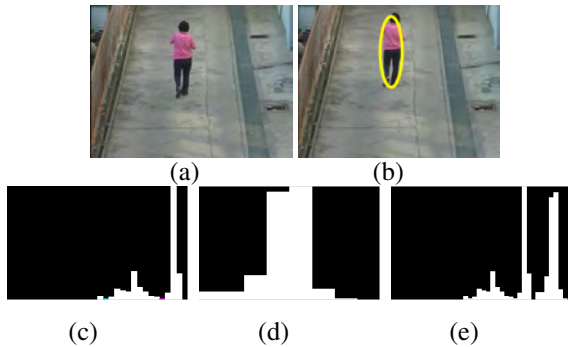


Figure 6 (a) Input image (b) Tracked result (c) 32-bin color histogram (d) 8-bin shape histogram (e) Concatenate histogram

Fig. 7 - 9 demonstrate tracking comparisons of mean shift tracking using color only versus using color and shape in various scenarios. Fig.7 shows an outdoor scene where a person walking in a straight trajectory under constant illumination condition, both trackers yielded good results. In Fig.8, a person is walking indoor and there is a brightness change happened in the image sequence. Under illumination change condition, it apparently shows that the color and shape tracker performed better than the color only tracker. For the sequence in Fig.9, we tried some scenarios where there is change in shape of the target. Fig.9(a) shows a sequence when a person change walking direction, and Fig.9(b) shows a sequence when a person abruptly sits down while walking. We found that both trackers work well. These assure us that by incorporating shape information to the color information helps improving accuracy of tracking.

Fig. 10 shows some snapshots of the tracked videos from the two cameras. It illustrates the handover result using color and shape cues in mean shift tracking algorithm.

8. CONCLUSIONS

We have presented a real-time system for continuously track a person across cooperative pan-tilt-zoom cameras. To handover the tracking process among two cameras, the second camera receives color cue and shape cue of a target person from the first camera. To achieve the mean shift algorithm tracking across the different radiometric characteristic cameras with different view point, our system uses color calibration to help with handover process. The experimental results confirm using color and shape cues helps reduce color mismatch between two cameras and the system is capable for continuously and accurately track the target person across cameras successfully.

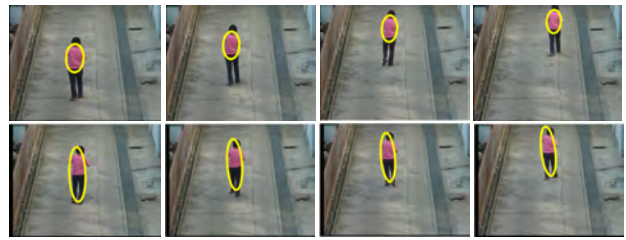


Figure 7 Tracking result using only color cue (top) versus using both color and shape cues (bottom) on an outdoor sequence with constant motion and illumination.

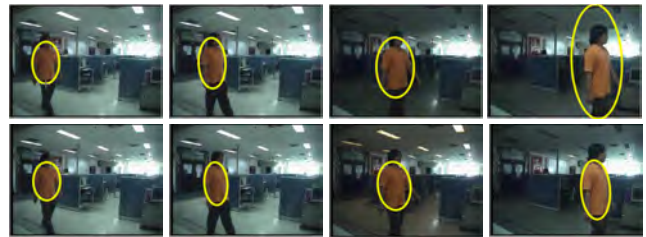


Figure 8 Tracking result using only color cue (top) versus using both color and shape cues (bottom) on an indoor sequence with changing illumination.



(a)



(b)

Figure 9 Tracking results using only color cue (top) versus using both color and shape cues (bottom) on two indoor sequences with changing motion direction (which makes target's silhouette shape changed) under constant

illumination.

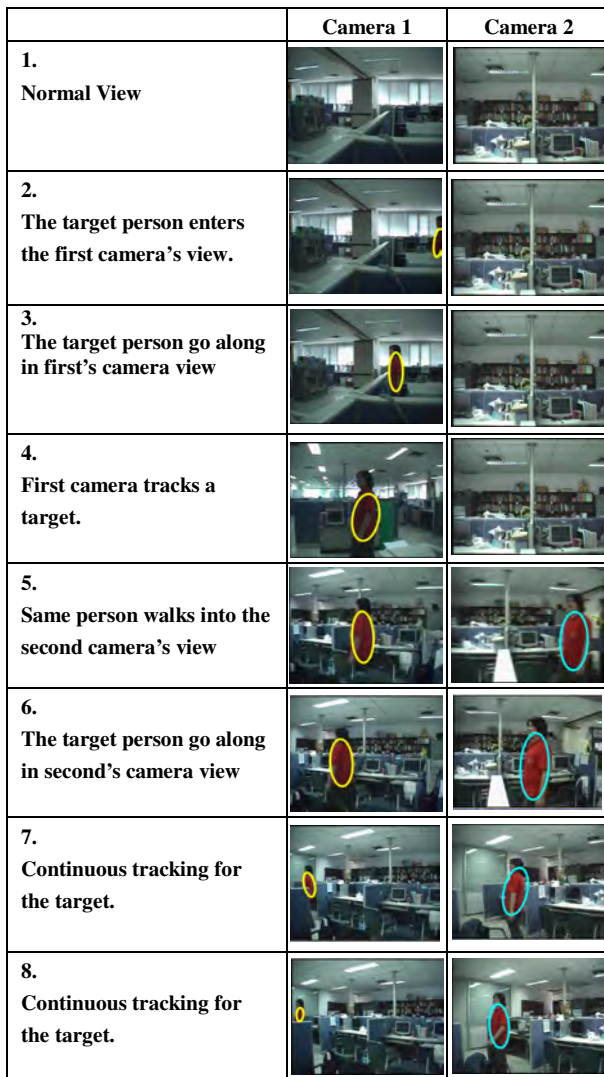


Figure 11 Some snapshots of the videos from the two cameras showing the handover moment. Using color and shape cue

9. ACKNOWLEDGEMENT

This research is in part supported by the Cooperation Project between Department of Electrical Engineering and Private Sector for Research and Development, Chulalongkorn University.

10. REFERENCES

- [1] J. Kang, I. Cohen, and G. Medioni, "Continuous Tracking Within and Across Camera Streams", in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR'03)*, Wisconsin, Jun 2003.
- [2] J. Black, T. Ellis, and P. Rosin, "Multi View Image Surveillance and Tracking", in *Proceedings of the IEEE Workshop on Motion and Video Computing(MOTION'02)*, Florida, Dec 2002.
- [3] L.Machestori, S.Piva and C.Raguzzoni, "Cooperative Multiple system for Realtime Face Detection and Tracking in Uncontrolled Condition",

SPIE Image and Video communication and processing,2005

- [4] L.Everts, N.Sebe and G.A.John,"Cooperative Object Tracking with Multiple PTZ Cameras", in *International conference On Image Analysis and Processing (ICIAP'07)*, 2007
- [5] M. Isard and A. Blake, "CONDENSATION - conditional density propagation for visual tracking", *International. Journal of Computer Vision*, pp.5-28, 1998.
- [6] K. Nummiaro, E. Koller-Meier and L. Van Gool, "Object tracking with an adaptative color-based Particle Filter", *Image and Vision Computing*, pp.99-111, 2002
- [7] P. Perez, C. Hue, J. Vermaak and M. Gangnet, "Color-based probabilistic tracking", *European conference on Computer Vision*, pp 661-675, 2002.
- [8] G.R. Bradski, "Computer vision face tracking for use in a perceptual user interface", *IEEE Workshop on Applications of Computer Vision*, pp.214-219, 1998.
- [9] Jinman Kang, Issaac Cohen, and Gerard Medioni, "Tracking object from multiple stationary and moving cameras", in *International Symposium on Communication and Information Technologies(ISCIT'04)*, Sapporo, Oct 2005
- [10] Junqiu Wang, and Yasushi Yagi, "Intrgrated Shape and Color for Real time Object Tracking ", in *International Conference on Robotics and Biomimetics*, Kunming, China, Dec 2006
- [11] P. Amnuaykanchanasin, S. Aramvith, T.H. Chalidabhongse, "Continuous person tracking using two cooperative active cameras," *Proceeding of IJWISA 2007*, Korea, Feb 2007.
- [12] N. Soontranon, S. Aramvith, and T. H. Chalidabhongse, "Face and Hand Localization and Tracking for Sign Language Recognition", in *International Symposium on Communication and Information Technologies(ISCIT'04)*, Sapporo, Oct 2004.

ประวัติผู้เขียนวิทยานิพนธ์

นายณนทรัฐ บำรุงเกียรติ เกิดเมื่อวันที่ 16 มกราคม 2527 เข้ารับการศึกษาในหลักสูตรวิศวกรรมศาสตร์บัณฑิต คณะวิศวกรรมศาสตร์ สาขาวิชาวิศวกรรมโทรคมนาคม สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ในปีการศึกษา 2545 และสำเร็จการศึกษาปริญญาวิศวกรรมศาสตร์บัณฑิต สาขาวิชาวิศวกรรมโทรคมนาคม ในปีการศึกษา 2548 และเข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สังกัดห้องปฏิบัติการกรรมวิธีสัญญาณดิจิทัล ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2549