

การตรวจหาและติดตามใบหน้าที่ด้วยกล้องวีดิทัศน์ 2 ตัวแบบสายกัมเมยและซุม



นายพิชัย อำนวยกาญจนสิน

สถาบันวิทยบริการ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า

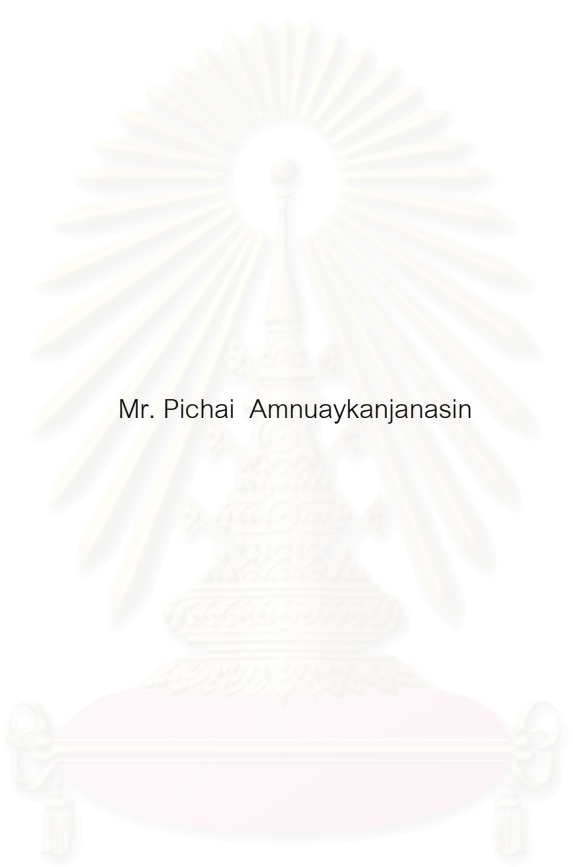
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2548

ISBN 974-53-2858-8

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

FACE DETECTION AND TRACKING USING TWO PAN / TILT / ZOOM VIDEO CAMERAS



Mr. Pichai Amnuaykanjanasin

สถาบันวิทยบริการ  
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 2005

ISBN 974-53-2858-8



พิชัย อำนวยการกาญจนสิน : การตรวจหาและติดตามใบหน้าด้วยกล้องวีดิทัศน์ 2 ตัวแบบสายก้มเงย และซูม. (FACE DETECTION AND TRACKING USING TWO PAN / TILT / ZOOM VIDEO CAMERAS) อ. ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร. สุภาวดี อร่ามวิทย์, อ. ที่ปรึกษาร่วม : ผู้ช่วยศาสตราจารย์ ดร. ธนารัตน์ ชลิดาพงศ์, 129 หน้า. ISBN 974-53-2858-8.

วิทยานิพนธ์นี้นำเสนอขั้นตอนวิธีการตรวจหา ติดตาม และรู้จำใบหน้าเวลาจริง โดยใช้กล้องวีดิทัศน์สาย ก้มเงย และซูม (PTZ) 2 ตัว ทำงานร่วมกัน โดยแบ่งการทำงานร่วมกันออกได้เป็น 2 ส่วน คือ ส่วนการรู้จำใบหน้าและติดตามบุคคลเป้าหมาย ในส่วนของการรู้จำใบหน้าร่วมกันจะอาศัยผลการรู้จำที่ได้จากทั้งสองกล้อง ข้อดีของกระบวนการนี้ คือ เป็นการเพิ่มโอกาสที่จะรับภาพใบหน้าได้ในมุมมองที่ดีทั้งในด้านขนาดและตำแหน่งมุมมอง นั่นคือ เป็นการเพิ่มความแม่นยำของการรู้จำใบหน้า โดยการใช้ผลของการรู้จำใบหน้าจากกล้องทั้งสองมาพิจารณาและวิเคราะห์ร่วมกัน ทั้งนี้สำหรับกล้องวีดิทัศน์แต่ละตัว ใบหน้ามนุษย์จะถูกตรวจหาและแยกส่วนจากฉากวีดิทัศน์โดยใช้การตรวจหาการเคลื่อนที่และสีผิว จากนั้นจะนำภาพใบหน้าที่ได้ไปใช้สำหรับการรู้จำ ซึ่งจะใช้กระบวนการจับคู่ใบหน้าด้วยการวัดความคล้ายของการแจกแจงของสีระหว่างภาพใบหน้าที่รับเข้ากับภาพใบหน้าในฐานข้อมูลด้วยระยะ Bhattacharyya ส่วนการทำงานร่วมกันแบบที่สอง คือ การทำงานร่วมกันสำหรับติดตามบุคคลเป้าหมายอย่างต่อเนื่องระหว่างกล้องวีดิทัศน์ทั้งสองตัว โดยใช้ลักษณะสีและทิศทางของการเคลื่อนที่สำหรับการระบุบุคคล ผลการทดลอง แสดงให้เห็นว่าขั้นตอนวิธีที่นำเสนอสามารถตรวจหา ติดตาม และรู้จำใบหน้ามนุษย์ได้ด้วยอัตราความแม่นยำที่สมเหตุสมผล ภายใต้สภาพแวดล้อมที่กำหนดมุมมองของมนุษย์ นอกเหนือจากนี้ยังมีการนำเสนอผลการทดลองของการส่งต่อข้อมูลระหว่างกล้องในส่วนของการติดตามบุคคลเป้าหมายอย่างต่อเนื่องร่วมกันระหว่างกล้องสองตัวในบริเวณกว้าง

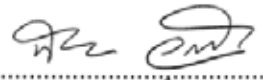
ภาควิชา.....วิศวกรรมไฟฟ้า.....ลายมือชื่อนิสิต.....  
 สาขาวิชา.....วิศวกรรมไฟฟ้า.....ลายมือชื่ออาจารย์ที่ปรึกษา.....  
 ปีการศึกษา.....2548.....ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

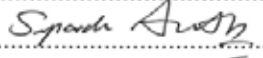
## 4670411421 : MAJOR ELECTRICAL ENGINEERING


KEY WORD: COOPERATIVE / MULTIPLE CAMERA / CAMERA CONTROL / TRACKING / DETECTION

PICHAJ AMNUAYKANJANASIN : FACE DETECTION AND TRACKING USING TWO  
PAN / TILT / ZOOM VIDEO CAMERAS. THESIS ADVISOR : ASST. PROF.  
SUPAVADEE ARAMVITH, THESIS COADVISOR : ASST. PROF. THANARAT  
CHALIDABHONGSE, Ph.D., 129 pp. ISBN 974-53-2858-8.

In this thesis, we present a method for real-time face detection, tracking, and recognizing using two cooperative pan-tilt-zoom (PTZ) cameras. The cooperative tasks are divided into two parts – face recognition and target human tracking. The cooperative face recognition identifies face based on results from both cameras. The advantages of this approach include increasing chances that we can obtain good face images (both in size and pose), as well as increasing recognition accuracy by combining the recognition results from multiple cameras. For each camera, the human face is detected and segmented from the video scene using motion and skin color cues. Then, the recognition is performed using a face matching method based on Bhattacharyya distance to measure a similarity between color-distribution of face images. The other part of the multiple camera collaborations is the cooperative in continuous tracking a target person between both cameras. In tracking, to identify person, we use color and motion as features. Experimental results show our proposed method can detect, track, and recognize human face under unconstrained human pose environment with reasonable accuracy rate. In addition, the results of the camera hand-off in continuous cooperative tracking of a target person under wide area are also presented.

Department..... Electrical Engineering..... Student's signature..... 

Field of study..... Electrical Engineering..... Advisor's signature..... 

Academic year..... 2005..... Co-advisor's signature..... 

## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จได้ด้วยคามกรุณาและความช่วยเหลืออย่างดียิ่งจาก ผศ. ดร. สุภาวดี อร่ามวิทย์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ และ ผศ. ดร. ธนวัฒน์ ชลิดาพงศ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ที่ได้สละเวลาในการให้คำปรึกษา คำแนะนำอันมีประโยชน์ และ ประสพการณ์ด้านการทำงานอันมีค่า ตลอดจนช่วยตรวจทานแก้ไขวิทยานิพนธ์และบทความตีพิมพ์ต่าง ๆ จนวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี

ขอขอบคุณ รศ. ดร.สมชาย จิตะพันธ์กุล และ อ. สุวิทย์ นาคพิระยुทธ ซึ่งเป็น กรรมการสอบวิทยานิพนธ์

ขอขอบคุณ พี่พรเทพ ที่คอยให้คำแนะนำในการเรียนปริญญาโทเป็นอย่างดี พี่ถิรพิรุฬห์ ที่คอยให้คำปรึกษาและคำแนะนำอย่างดีเยี่ยมต่อการทำวิทยานิพนธ์ฉบับนี้ในทุก ๆ เรื่องที่ไม่เข้าใจ พี่นรุต ที่ช่วยให้คำแนะนำความเข้าใจเกี่ยวกับขั้นตอนวิธีที่นำเสนอในวิทยานิพนธ์ อย่างดี มากตลอดมา พี่วิสูตรและพี่สมภพ ที่ช่วยให้คำแนะนำและสั่งสอนเกี่ยวกับการเขียนโปรแกรม

ขอขอบคุณพี่ ๆ (พี่แจจ พี่โอะ พี่อาร์ม พี่ตุ๊ก) เพื่อน ๆ (เนษฐ์ ต้น ชัน ตูน) กลุ่มวิจัยทางการประมวลผลสัญญาณดิจิทัลศน์ของ ผศ. ดร. สุภาวดี อร่ามวิทย์ ที่คอยให้คำปรึกษา และ วิเคราะห์สิ่งต่าง ๆ ให้มีวิสัยทัศน์ที่กว้างไกลมากขึ้นกว่าคิดคนเดียว รวมทั้งให้ความร่วมมืออย่างดียิ่งในความช่วยเหลือสำหรับการทดลอง

ขอขอบคุณ คุณพิชัย วัฒนระภาดร ที่ช่วยเหลือในด้านการทดลองของข้าพเจ้า อย่างดีตลอดมาตั้งแต่วันแรกที่ได้พบกัน และขอบคุณ จัง อี๊ด เต้ ขวัญ โฉก ชาย กู้ก ที่เป็นเพื่อนที่ดี ตลอดมา ให้ความช่วยเหลือในการทดลองของข้าพเจ้าอย่างดีตลอดมา และขอบคุณ พี่ ๆ เพื่อน ๆ และน้อง ๆ ในห้องปฏิบัติการวิจัยกรรมวิธีสัญญาณดิจิทัล ที่ช่วยทำให้เป็นห้องวิจัยที่เพียบพร้อมที่สุด

สุดท้ายนี้ ขอกราบขอบพระคุณน้ำพล น้ำแก้ว ที่ช่วยเลี้ยงดูให้แก่ปเป็นแก่ปในวันนี้ได้ จนข้าพเจ้าสามารถสำเร็จการศึกษาได้สูงถึงระดับวิศวกรรมศาสตรมหาบัณฑิตได้ในวันนี้ แม้อาจไม่มีคำพูดหวาน ๆ หรือ ซึ้งๆ แต่ในใจแลการกระทำจะบ่งบอกความรู้สึกต่อท่านได้เป็น อย่างดีที่สุด และที่จะลืมไม่ได้ คือ พี่กอล์ฟ ถ้าไม่มีเขาคนนี้ ข้าพเจ้าคงทำอะไรไม่เป็น ไม่รู้จะ ดำเนินชีวิตไปทางไหน ช่วยสั่งสอนในทุก ๆ เรื่องที่เขาสามารถจะช่วยให้ ข้าพเจ้าอยากจะทำว่า ขอขอบคุณพี่กอล์ฟมากนะครับ





บทที่	หน้า
2.2.1 การดำเนินการตรวจกับรูปภาพฐานสอง.....	19
2.2.2 การทำไดเลชัน (Dilation) และ การทำอีโรชัน (Erosion).....	20
2.2.3 ตัวดำเนินการแบบเปิดและปิด.....	22
2.2.4 รูปแบบโครงสร้างส่วนย่อย.....	23
2.2.5 การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน.....	23
2.2.5.1 การวิเคราะห์รูปร่างของกลุ่มบริเวณที่เชื่อมต่อ.....	24
2.3 แบบจำลองกล้อง (Camera model) และแบบจำลองการเคลื่อนที่ (Motion model).....	26
2.3.1 แบบจำลองกล้อง.....	26
2.3.2 การเคลื่อนที่ของกล้อง (Camera Motion).....	27
2.3.3 แบบจำลองการเคลื่อนที่.....	29
2.4 การตรวจหาและติดตามใบหน้า (Face Detection and Tracking) .....	30
2.4.1 การตรวจหาการเคลื่อนที่ (Motion Detection).....	30
2.4.2 กระบวนการลบฉากหลังพื้นฐาน (Basic background Subtraction).....	32
2.4.3 การตรวจหาดัวยสี (color detection).....	33
2.4.4 กระบวนการติดตามใบหน้า.....	37
2.4.5 การวัดประสิทธิภาพเชิงปรวิสัย (Objective Performance Evaluation).....	39
2.4.5.1 อัตราการตรวจหา (Detection rate : DR).....	41
2.4.5.2 อัตราการฟ้องความผิดพลาด (False Alarm Rate : FAR).....	41
2.4.5.3 อัตราการปฏิเสธความผิดพลาด (False Reject Rate : FRR).....	41
2.5 การควบคุมกล้องวิดีโอ PTZ.....	42
2.6 การจับคู่ใบหน้า (Face matching).....	45
2.6.1 การเปรียบเทียบฮิสโทแกรมด้วยระยะ Bhattacharyya.....	45
3 โครงสร้างและระบบตรวจหาติดตามที่นำเสนอ.....	47
3.1 รูปแบบการทำงานของระบบที่นำเสนอ.....	48
3.1.1 กล้องวิดีโอ PTZ.....	49
3.1.2 ลักษณะโครงสร้างการจัดวางตำแหน่งกล้องวิดีโอ PTZ ทั้ง 2 ตัว.....	49
3.1.3 ไดอะแกรมการทำงานโดยรวม.....	51
3.2 ขั้นตอนวิธีในการตรวจหาและติดตามใบหน้า.....	52



3.2.1	กระบวนการตรวจหาด้วยสีผิว.....	53
3.2.2	กระบวนการตรวจหาการเคลื่อนที่ที่นำเสนอ.....	54
3.2.3	กระบวนการตรวจหาโดยใช้การเคลื่อนที่และสีผิว.....	57
3.2.4	การจำแนกเฟรมภาพในการประมวลผล.....	58
3.2.5	กระบวนการลดสัญญาณรบกวน.....	59
3.2.6	การจำแนกส่วนบริเวณใบหน้า.....	60
3.2.7	การกำหนดขอบเขตให้กับบริเวณใบหน้า.....	60
3.2.8	การจำแนกประเภทใบหน้าตามลักษณะมุมมอง (Face alignment).....	62
3.3	การควบคุมการส่าย ก้มเงย และขมุกขมองของกล้องวิดีโอ PTZ เพื่อติดตามใบหน้า บุคคลเป้าหมาย.....	63
3.3.1	การคำนวณมุมในการเปลี่ยนมุมมองและอัตราการขมุก.....	63
3.3.2	การชดเชยผลกระทบเนื่องจากการการขมุก.....	64
3.4	ขั้นตอนวิธีการเปรียบเทียบใบหน้ากับใบหน้าที่รวมในฐานข้อมูล.....	65
3.4.1	กระบวนการวัดความคล้าย.....	65
3.5	การทำงานร่วมกันระหว่างกล้อง.....	65
3.5.1	การประมวลผลร่วมกันสำหรับระบุบุคคลโดยการจับคู่ใบหน้า.....	65
3.5.2	การทำงานร่วมกันสำหรับตรวจหาติดตามใบหน้าอย่างต่อเนื่อง.....	67
3.5.2.1	การแบ่งบริเวณบุคคลในส่วนของบริเวณที่ตรวจหา.....	67
3.5.2.2	การพิจารณาทิศทางเคลื่อนที่ของบุคคลที่สนใจ.....	68
4	ผลการจำลองระบบ.....	70
4.1	ผลการจำลองระบบในส่วนของการตรวจหาใบหน้า.....	71
4.1.1	ผลการตรวจหาด้วยสีผิว.....	71
4.1.1.1	การหาแบบจำลองวงรี.....	71
4.1.1.2	การเปรียบเทียบความแตกต่างของแบบจำลองวงรี.....	76
4.1.1.3	ผลการตรวจหาด้วยแบบจำลองวงรีที่กำหนดและปัญหาที่ เกิดขึ้นและการวิเคราะห์ผล.....	78
4.1.2	ผลการตรวจหาการเคลื่อนที่.....	80
4.1.2.1	ผลกระทบของการเลือกช่วงห่างระหว่างเฟรมในการประมวลผล หาผลต่างระหว่างเฟรมภาพ.....	81

4.1.2.2 การวิเคราะห์ผลการตรวจหาการเคลื่อนที่และผลการเติม จุดภาพ.....	82
4.1.3 ผลการตรวจหาใบหน้าด้วยการตรวจหาการเคลื่อนที่ร่วมกับการ ตรวจหาด้วยสีผิว.....	84
4.1.3.1 การวัดประสิทธิภาพของผลการตรวจหา.....	84
4.1.3.2 การแยกกลุ่มบริเวณใบหน้า.....	86
4.1.4 กระบวนการกำหนดขอบเขตและตัดส่วนภาพใบหน้า.....	87
4.1.5 ผลการจำแนกตำแหน่งมุมมองใบหน้า.....	89
4.2 ผลการจำลองระบบในส่วนการควบคุมการส่าย ก้มเงยและหมุนติดตามบุคคล.....	90
4.3 ผลการจำลองระบบในส่วนการทำงานร่วมกันระหว่างกล้องวีดิทัศน์.....	92
4.3.1 การประมวลผลร่วมกันสำหรับจับคู่ใบหน้า.....	93
4.3.1.1 ผลการควบคุมการส่ายและการก้มเงยของกล้องวีดิทัศน์ ติดตามใบหน้าจาก 2 มุมมอง.....	93
4.3.1.2 มุมมองภาพใบหน้าที่รับจากกล้องทั้งสอง.....	94
4.3.1.3 การฝึกสอนในกระบวนการจับคู่ใบหน้า.....	97
4.3.1.4 ผลการรู้จำใบหน้าที่ใช้วิธีการจับคู่ใบหน้า.....	99
4.3.2 การทำงานร่วมกันสำหรับการติดตามบุคคลต่อเนื่องระหว่างกล้อง.....	100
4.3.2.1 ผลการแยกบริเวณเสื้อผ้า สีผิว และผม ของบุคคล.....	100
4.3.2.2 การทดลองเปรียบเทียบความคล้ายของแต่ละกลุ่มบริเวณ.....	101
4.3.2.3 ผลการระบุบุคคลระหว่างกล้องวีดิทัศน์ PTZ 2 กล้อง.....	104
5 สรุปผลการวิจัยและข้อเสนอแนะ.....	112
5.1 สรุปผลการวิจัย.....	112
5.2 อภิปรายผลการวิจัย.....	112
5.3 ข้อเสนอแนะ.....	114
รายการอ้างอิง.....	115
ภาคผนวก.....	119
บทความทางวิชาการที่ได้รับการเผยแพร่.....	120
ประวัติผู้เขียนวิทยานิพนธ์.....	129

## สารบัญตาราง

ตารางประกอบ	หน้า
ตารางที่ 2.1 การดำเนินการตรรกะพื้นฐาน 4 ตัวดำเนินการ.....	20
ตารางที่ 2.2 ขนาดพื้นที่ได้กราฟของกราฟการแจกแจงปกติเมื่อพิจารณาเป็นจำนวนเท่า ของค่าเบี่ยงเบนมาตรฐาน.....	37
ตารางที่ 4.1 ขนาดของวงรีของค่าตัดสินใจ $n = 1, 2, 3, 4, 5$ .....	73
ตารางที่ 4.2 ประสิทธิภาพของผลการตรวจหาสีผิวด้วยแบบจำลองวงรีเทียบกับผลถูกต้อง พื้นฐาน.....	74
ตารางที่ 4.3 เปรียบเทียบประสิทธิภาพการตรวจหาของแบบจำลองวงรี.....	77
ตารางที่ 4.4 เปรียบเทียบประสิทธิภาพของผลการตรวจหา.....	85
ตารางที่ 4.5 เปรียบเทียบผลการจับคู่ใบหน้าระหว่างผลของกล่องแต่ละตัวกับผลที่ได้จาก การประมวลผลร่วมกัน.....	99
ตารางที่ 4.6(ก) ผลการเปรียบเทียบความคล้ายของบุคคลระหว่างกล่องของกลุ่มบริเวณสี เสื้อผ้า.....	103
ตารางที่ 4.6(ข) ผลการเปรียบเทียบความคล้ายของบุคคลระหว่างกล่องของกลุ่มบริเวณสีผิว.....	103
ตารางที่ 4.6(ค) ผลการเปรียบเทียบความคล้ายของบุคคลระหว่างกล่องของกลุ่มบริเวณสีผม.....	103

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

## สารบัญภาพ

ภาพประกอบ	หน้า
รูปที่ 1.1	การประยุกต์ใช้กับระบบสอดส่องรักษาความปลอดภัย ..... 4
รูปที่ 1.2	งานประยุกต์กล้องวีดิทัศน์ PTZ สำหรับงานการเรียนการสอนทางไกล ..... 4
รูปที่ 1.3	ระบบตรวจหาและติดตามใบหน้าสำหรับจัดเก็บเป็นฐานข้อมูลใบหน้า ..... 5
รูปที่ 1.4	งานประยุกต์ใช้สำหรับการรู้จำกิริยาท่าทางโดยใช้กล้องวีดิทัศน์ PTZ เคลื่อนที่ และค้นหาใบหน้าและมือ ..... 6
รูปที่ 1.5	ภาพฉากหลังที่ถูกบันทึกไว้ ณ ตำแหน่งการเคลื่อนที่ของกล้อง ..... 8
รูปที่ 1.6	ภาพฉากหลังแบบโมเสก ..... 8
รูปที่ 1.7	การติดตามผู้เล่นโดยใช้ ระยะเวลา Bhattacharyya ..... 10
รูปที่ 1.8	ระบบห้องฉลาดเพื่อเลือกมุมมองที่ดีที่สุดในการรับภาพ ..... 11
รูปที่ 1.9	ตัวอย่างระบบการรู้จำใบหน้าที่ใช้ในเวลาจริง ..... 11
รูปที่ 1.10	ผลการรู้จำใบหน้า ..... 12
รูปที่ 2.1	ปริภูมิสีแบบ RGB (สีแดง เขียว และน้ำเงิน) ..... 16
รูปที่ 2.2	(ก) ปริภูมิสีแบบ YUV (ข) ปริภูมิสีแบบ YCbCr ..... 19
รูปที่ 2.3	ตัวอย่างของการดำเนินการตรรกะระหว่างรูปภาพฐานสอง ..... 20
รูปที่ 2.4	ตัวอย่างการทำไคเลชัน ..... 21
รูปที่ 2.5	การทำอีโรชัน ..... 22
รูปที่ 2.6	ตัวดำเนินการแบบเปิด ..... 22
รูปที่ 2.7	ตัวดำเนินการแบบปิด ..... 23
รูปที่ 2.8	ตัวอย่างโครงสร้างส่วนย่อยในการทำไคเลชันและการทำอีโรชัน ..... 23
รูปที่ 2.9	การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกันระหว่างจุดภาพใกล้เคียง ..... 24
รูปที่ 2.10	การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกันด้วยรหัสแบบ 4 จุดเชื่อมต่อ ..... 24
รูปที่ 2.11	การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกันด้วยรหัสแบบ 8 จุดเชื่อมต่อ ..... 24
รูปที่ 2.12	ตัวอย่างภาพฐานสองที่นำมาวิเคราะห์หาบริเวณกลุ่มใบหน้า ..... 25
รูปที่ 2.13	การฉายภาพแบบทัศนมิติ (perspective projection) ของแบบจำลองกล้อง รูเข็ม ..... 26
รูปที่ 2.14	การเคลื่อนที่ของกล้อง ..... 27
รูปที่ 2.15	ลำดับภาพที่รับได้จากการเคลื่อนที่ของกล้องแบบการส่าย ..... 28
รูปที่ 2.16	ลำดับภาพที่รับได้จากการเคลื่อนที่ของกล้องแบบการก้มเงย ..... 28
รูปที่ 2.17	ลำดับภาพที่รับได้จากการเคลื่อนที่ของกล้องแบบการหมุน ..... 28

ภาพประกอบ	หน้า
รูปที่ 2.18	สนามการเคลื่อนที่ของการส่ายไปทางซ้าย..... 29
รูปที่ 2.19	สนามการเคลื่อนที่ของการก้มเงยลง..... 29
รูปที่ 2.20	สนามการเคลื่อนที่ของการซุ่ม..... 30
รูปที่ 2.21	ตัวอย่างผลการตรวจหาการเคลื่อนที่..... 31
รูปที่ 2.22	ปัญหาของการตรวจหาการเคลื่อนที่..... 32
รูปที่ 2.23	ตัวอย่างการตรวจหาการเคลื่อนที่..... 32
รูปที่ 2.24	ผลการตรวจหาด้วยขั้นตอนวิธีการลบฉากหลัง..... 33
รูปที่ 2.25	การกระจายตัวของค่าองค์ประกอบสี..... 35
รูปที่ 2.26	การแปลงพิกัดจากพิกัด $(Cb, Cr)$ ไปเป็นพิกัด $(\alpha, \beta)$ ..... 35
รูปที่ 2.27	กราฟการแจกแจงข้อมูลปกติและแนวขอบเขตพื้นที่ได้กราฟที่เป็นจำนวนเท่า ของค่าเบี่ยงเบนมาตรฐาน $(\sigma)$ ..... 36
รูปที่ 2.28	ผลการตรวจหาในงานวิจัย [16]..... 37
รูปที่ 2.29	ผลการตรวจหาสีผิว..... 37
รูปที่ 2.30	การกำหนดขอบเขตในการตรวจหาวัตถุเฟรมถัดไป..... 39
รูปที่ 2.31	การกำหนดบริเวณวัตถุที่สนใจ..... 39
รูปที่ 2.32	เรขาคณิตของกล้อง (มุมมองด้านบน)..... 42
รูปที่ 3.1	กล้องวีดิทัศน์ในการทดลอง..... 49
รูปที่ 3.2	การจัดวางกล้องเพื่อการประมวลผลจับคู่ใบหน้าร่วมกันแบบที่หนึ่ง..... 50
รูปที่ 3.3	การจัดวางกล้องเพื่อการประมวลผลจับคู่ใบหน้าร่วมกันแบบที่สอง..... 50
รูปที่ 3.4	การจัดวางกล้องเพื่อการประมวลผลตรวจหาาร่วมกัน..... 51
รูปที่ 3.5	ไดอะแกรมโดยรวมสำหรับวิธีการประมวลผลจับคู่ใบหน้าร่วมกัน..... 52
รูปที่ 3.6	ไดอะแกรมโดยรวมสำหรับวิธีการประมวลผลตรวจหาาร่วมกัน..... 52
รูปที่ 3.7	รูปแบบจำลองวงรีที่กำหนดขึ้นจากการฝึกสอน..... 53
รูปที่ 3.8	ปัญหาของการตรวจหาสีผิว..... 54
รูปที่ 3.9	ไดอะแกรมการทำงานในส่วนการต่อเชื่อมจุดภาพ..... 55
รูปที่ 3.10	ผลการตรวจหาการเคลื่อนที่..... 56
รูปที่ 3.11	ผลการเติมจุดภาพที่มีปัญหา..... 56
รูปที่ 3.12	ตัวอย่างผลการตรวจหาด้วยการตรวจหาการเคลื่อนที่และตรวจหาสีผิว..... 57

ภาพประกอบ	หน้า
รูปที่ 3.13 ลำดับวิถีทัศนที่เกิดจากการส่าย ก้มเงย และชুমของกล้องวิดีโอจากหยุดนิ่งเคลื่อนที่ และหยุดนิ่ง.....	58
รูปที่ 3.14 ผลการตรวจหาการเคลื่อนที่ ณ เฟรมภาพที่กล้องกำลังเปลี่ยนมุมมองภาพในลักษณะส่ายและก้มเงย.....	59
รูปที่ 3.15 ผลการตรวจหาการเคลื่อนที่ ณ เฟรมภาพที่กล้องกำลังเปลี่ยนมุมมองในลักษณะชুম.....	59
รูปที่ 3.16 โครงสร้างส่วนย่อยในการทำอิโรชันและไดเลชัน.....	59
รูปที่ 3.17 ผลการตรวจหาในอุดมคติ.....	61
รูปที่ 3.18 ขั้นตอนการกำหนดขอบเขตใบหน้า.....	62
รูปที่ 3.19 ลักษณะการจำแนกใบหน้าเป็น 3 ประเภท.....	63
รูปที่ 3.20 ขอบเขตการส่งให้กล้องเคลื่อนที่.....	64
รูปที่ 3.21 ปัญหาของการเปลี่ยนมุมมองกล้องวิดีโอเมื่อมีการชুম.....	65
รูปที่ 3.22 การแบ่งบริเวณโครงร่างบุคคลที่ได้จากการตรวจหา.....	68
รูปที่ 3.23 การพิจารณาทิศทางการเคลื่อนที่ของบุคคลจากตำแหน่งการเปลี่ยนมุมมองของกล้อง.....	69
รูปที่ 4.1 ภาพใบหน้าที่นำมาใช้หาขอบเขตแบบจำลองวงรี.....	71
รูปที่ 4.2 ตัวอย่างการกระจายตัวของค่าองค์ประกอบสีในปริภูมิย่อย CbCr ของภาพใบหน้า.....	71
รูปที่ 4.2(ต่อ) ตัวอย่างการกระจายตัวของค่าองค์ประกอบสีในปริภูมิย่อย CbCr ของภาพใบหน้า.....	72
รูปที่ 4.3 รูปการกระจายตัวของค่าจุดภาพใบหน้าที่สีผิวในปริภูมิค่าองค์ประกอบสี CbCr ของทั้ง 20 ภาพใบหน้า.....	73
รูปที่ 4.4 แบบจำลองวงรีที่ค่าตัดสีผิว $\eta = 1\ 2\ 3\ 4$ และ 5.....	74
รูปที่ 4.5 ผลการตรวจหาด้วยสีผิว ที่ใช้แบบจำลองวงรีที่ค่าตัดสีผิวต่างกัน.....	75
รูปที่ 4.6 การหาแบบจำลองวงรีด้วยภาพใบหน้าเฉพาะบุคคล.....	76
รูปที่ 4.7 เปรียบเทียบผลการตรวจหาสีผิวของบุคคล ก.....	76
รูปที่ 4.8 เปรียบเทียบผลการตรวจหาสีผิวของบุคคล ข.....	77
รูปที่ 4.9 เปรียบเทียบผลการตรวจหาสีผิวของบุคคล ค.....	77



ภาพประกอบ	หน้า
รูปที่ 4.10 ผลการตรวจหาสีผิวด้วยแบบจำลองวงรีที่กำหนด.....	79
รูปที่ 4.11 การวิเคราะห์ความผิดพลาดที่เกิดจากการตรวจหาสีผิวตามแบบจำลองวงรี.....	80
รูปที่ 4.12 ผลการตรวจหาการเคลื่อนที่ เมื่อเลือกช่วงห่างระหว่างเฟรมต่างกัน.....	81
รูปที่ 4.13 ผลการหาตำแหน่งจุดภาพเริ่มต้นและจุดภาพตำแหน่งสุดท้ายของภาพฐานสอง ที่ได้จากการหาความต่างระหว่างเฟรม.....	82
รูปที่ 4.13(ต่อ) ผลการหาตำแหน่งจุดภาพเริ่มต้นและจุดภาพตำแหน่งสุดท้ายของภาพฐานสอง ที่ได้จากการหาความต่างระหว่างเฟรม.....	83
รูปที่ 4.14 ผลการเติมจุดภาพระหว่างตำแหน่งจุดภาพเริ่มต้นและจุดภาพตำแหน่งสุดท้าย..	83
รูปที่ 4.14(ต่อ) ผลการเติมจุดภาพระหว่างตำแหน่งจุดภาพเริ่มต้นและจุดภาพตำแหน่งสุดท้าย..	84
รูปที่ 4.15 ผลการตรวจหาใบหน้าหลังจากผ่านการตรวจหาการเคลื่อนที่ร่วมกับตรวจหา สีผิว.....	84
รูปที่ 4.16 ผลการตรวจหาส่วนสีผิว.....	85
รูปที่ 4.17 ผลการแยกบริเวณใบหน้า.....	86
รูปที่ 4.18 ผลการกำหนดขอบเขตใบหน้าในลักษณะสีเหลี่ยม.....	87
รูปที่ 4.19 ข้อผิดพลาดในการกำหนดขอบเขตด้านซ้ายและขวาของกรอบสีเหลี่ยม.....	88
รูปที่ 4.20 ผลการกำหนดขอบเขตและตัดส่วนภาพใบหน้าในแบบวงรีหรือวงกลม.....	89
รูปที่ 4.21 ภาพผลการตรวจหาเฉพาะบริเวณใบหน้า.....	89
รูปที่ 4.21(ต่อ) ภาพผลการตรวจหาเฉพาะบริเวณใบหน้า.....	90
รูปที่ 4.22 การจำแนกกลุ่มของใบหน้า.....	90
รูปที่ 4.23 ผลการควบคุมกล้องเพื่อเปลี่ยนมุมมองติดตามใบหน้า.....	91
รูปที่ 4.24 ปัญหาในการถ่าย และก้มเงยกล้องวีดิทัศน์ ติดตามบุคคลที่มีการเคลื่อนที่.....	91
รูปที่ 4.25 ผลการเปลี่ยนมุมมองกล้องเมื่อกล้องอยู่ในสถานะมุมมองที่ถูกข่มอยู่.....	92
รูปที่ 4.26 ภาพตัวอย่างจากลำดับวีดิทัศน์การถ่ายก้มเงยและซูมกล้องวีดิทัศน์เพื่อติดตาม ใบหน้าบุคคลที่สนใจ ที่รับจากกล้องวีดิทัศน์ PTZ 2 ตัวในตำแหน่งมุมมอง ต่างกัน สำหรับประยุกต์ใช้ในการจับคู่ใบหน้า.....	93
รูปที่ 4.26(ต่อ) ภาพตัวอย่างจากลำดับวีดิทัศน์การถ่ายก้มเงยและซูมกล้องวีดิทัศน์เพื่อติดตาม ใบหน้าบุคคลที่สนใจ ที่รับจากกล้องวีดิทัศน์ PTZ 2 ตัวในตำแหน่งมุมมอง ต่างกัน สำหรับประยุกต์ใช้ในการจับคู่ใบหน้า.....	94



ภาพประกอบ	หน้า
รูปที่ 4.27 ผลการตัดส่วนภาพใบหน้าจากทั้งสองกล้องและผลการจำแนกประเภทของ มูมมองใบหน้า.....	95
รูปที่ 4.27(ต่อ) ผลการตัดส่วนภาพใบหน้าจากทั้งสองกล้องและผลการจำแนกประเภทของ มูมมองใบหน้า.....	96
รูปที่ 4.28 ภาพตัวอย่างในฐานข้อมูล จำแนกตามประเภทของมูมมองใบหน้า.....	97
รูปที่ 4.28(ต่อ) ภาพตัวอย่างในฐานข้อมูล จำแนกตามประเภทของมูมมองใบหน้า.....	98
รูปที่ 4.29 ตัวอย่างผลการจับคู่ใบหน้า.....	99
รูปที่ 4.30 ผลการทดลองจำแนกบริเวณที่ได้จากการตรวจหา.....	101
รูปที่ 4.31 ลำดับภาพบุคคลที่นำมาใช้ในการเปรียบเทียบความคล้าย.....	102
รูปที่ 4.32 กราฟเปรียบเทียบการวัดความคล้ายโดยใช้กลุ่มบริเวณสีเสื้อผ้า สีผิว และสีผม.....	104
รูปที่ 4.33 ลำดับภาพผลการติดตามและระบุบุคคลระหว่างกล้องวีดิทัศน์ตัวที่ 1 และ ตัวที่ 2 ของตัวอย่างสถานการณ์ที่ 1.....	106
รูปที่ 4.33(ต่อ) ลำดับภาพผลการติดตามและระบุบุคคลระหว่างกล้องวีดิทัศน์ตัวที่ 1 และ ตัวที่ 2 ของตัวอย่างสถานการณ์ที่ 1.....	107
รูปที่ 4.34 ลำดับภาพผลการติดตามและระบุบุคคลระหว่างกล้องวีดิทัศน์ตัวที่ 1 และ ตัวที่ 2 ของตัวอย่างสถานการณ์ที่ 2.....	108
รูปที่ 4.34(ต่อ) ลำดับภาพผลการติดตามและระบุบุคคลระหว่างกล้องวีดิทัศน์ตัวที่ 1 และ ตัวที่ 2 ของตัวอย่างสถานการณ์ที่ 2.....	109
รูปที่ 4.35 ลำดับภาพผลการติดตามและระบุบุคคลระหว่างกล้องวีดิทัศน์ตัวที่ 1 และ ตัวที่ 2 ของตัวอย่างสถานการณ์ที่ 3.....	110
รูปที่ 4.35(ต่อ) ลำดับภาพผลการติดตามและระบุบุคคลระหว่างกล้องวีดิทัศน์ตัวที่ 1 และ ตัวที่ 2 ของตัวอย่างสถานการณ์ที่ 3.....	111

## ดัชนีคำศัพท์

algorithm	ขั้นตอนวิธี
affine transform	การแปลงสัมพรรค
analog	แอนะล็อก
background subtraction	ลบฉากหลัง
bhattacharyya coefficient	สัมประสิทธิ์ Bhattacharyya
bhattacharyya distance	ระยะ Bhattacharyya
cooperative	ทำงานร่วมกัน
center of gravity	จุดศูนย์กลางถ่วง
camera control	การควบคุมกล้อง
chrominance component	องค์ประกอบสี
color bin	ถังสี
cropping	การตัดส่วนภาพ
camera motion	การเคลื่อนที่ของกล้อง
camera model	แบบจำลองกล้อง
center of projection	ศูนย์กลางการฉายภาพ
chrominance	สัญญาณสี
closing operator	ตัวดำเนินการแบบปิด
computer vision	คอมพิวเตอร์วิทัศน์
connected component analysis	การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน
column	สดมภ์
dimension	มิติ
detection rate	อัตราการตรวจหา
detection	การตรวจหา
dilation	ไดเลชัน
erosion	อีโรชัน
elliptical model	แบบจำลองวงรี
elliptical boundary model	แบบจำลองขอบเขตวงรี
fixed camera	กล้องอยู่กับที่

field of view	ขอบเขตภาพ
frame differencing	ผลต่างระหว่างเฟรม
frame rate	อัตราเฟรม
focal length	ความยาวโฟกัส
face matching	การจับคู่ใบหน้า
face recognition	การรู้จำใบหน้า
false alarm rate	อัตราการฟ้องค่าผิดพลาด
false negative	ความผิดพลาดแบบลบ
false positive	ความผิดพลาดแบบบวก
ground truth	ผลถูกต้องพื้นฐาน
histogram	ฮิสโทแกรม
homography	โฮโมกราฟี
human vision system	ระบบการมองเห็นของมนุษย์
interaction	การโต้ตอบ
input	รับเข้า
luminance component	องค์ประกอบแสง
labeling	การทำป้าย
monitor	เฝ้าสังเกต
multi-camera system	ระบบกล้องหลายตัว
matching score	คะแนนการจับคู่
mean	ค่าเฉลี่ย
morphological image processing	การประมวลผลภาพเชิงสัญญาณ
motion vector	เวกเตอร์การเคลื่อนที่
major axis	แกนเอก
minor axis	แกนโท
motion detection	การตรวจหาการเคลื่อนที่
normalize	นอร์มัลไลซ์
non parametric model	แบบจำลองชนิดไม่มีตัวแปร

opening operator	ตัวดำเนินการแบบเปิด
PTZ video camera	กล้องวีดิทัศน์ PTZ
Person identify	ระบุบุคคล
person silhouette	โครงร่างบุคคล
program	โปรแกรม
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
similarity measure	การวัดความคล้าย
spatial	เชิงพื้นที่
structuring element	โครงสร้างส่วนย่อย
skin color model	แบบจำลองสีผิว
tilt	ก้มเงย
threshold	ขีดเริ่มเปลี่ยน
true negative	ความถูกต้องแบบลบ
true positive	ความถูกต้องแบบบวก
temporal	เชิงเวลา
variance	ความแปรปรวน
YCbCr color space	ปริภูมิสี YCbCr
zoom	ซูม

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันเทคโนโลยีการประมวลผลภาพและวีดิทัศน์ได้เข้ามามีบทบาทต่อการดำเนินชีวิตของมนุษย์มากยิ่งขึ้น เนื่องด้วยความก้าวหน้าและพัฒนาอย่างรวดเร็วของทั้งตัวประมวลผลและเทคนิควิธีต่าง ๆ ที่ถูกนำมาประยุกต์ใช้ รวมถึงความต้องการของมนุษย์ที่ต้องการนำเทคโนโลยีมาช่วยแบ่งเบาและอำนวยความสะดวกให้แก่มนุษย์ ซึ่งบางอย่างมนุษย์ไม่อาจสามารถกระทำตัวเอง หรือสามารถกระทำได้ แต่ด้วยข้อจำกัดบางประการทำให้ประสิทธิภาพในการทำงานลดลง จนอาจทำให้เกิดความผิดพลาดหรือความเสียหายเกิดขึ้นได้

ระบบสอดส่อง (surveillance system) รักษาความปลอดภัย เป็นรูปแบบงานที่มีการนำเทคโนโลยีการประมวลผลภาพวีดิทัศน์มาประยุกต์ใช้กันอย่างแพร่หลาย โดยรูปแบบของระบบจะประกอบด้วยกล้องวีดิทัศน์ซึ่งถูกติดตั้ง ณ บริเวณสถานที่ต่าง ๆ ในมุมมองและตำแหน่งที่เหมาะสมต่อการสอดส่อง จากนั้นสัญญาณวีดิทัศน์จะถูกส่งผ่านไปยังส่วนควบคุม เพื่อทำการบันทึกสัญญาณวีดิทัศน์ รวมทั้งประมวลผลในรูปแบบต่าง ๆ สำหรับเป็นหลักฐานในกรณีที่ต้องการตรวจสอบสิ่งผิดปกติที่เกิดขึ้นในภายหลัง ในระยะเริ่มต้นของการประยุกต์ใช้งาน จะใช้มนุษย์เป็นผู้สังเกตเห็นจอคอมพิวเตอร์ว่ามีสิ่งผิดปกติ หรือมีบุคคลที่เข้าข่ายผู้ต้องสงสัยมีท่าทีพิรุธ อยู่ในขอบเขตภาพของกล้องวีดิทัศน์ตัวใดตัวหนึ่งในระบบหรือไม่ การใช้มนุษย์เป็นผู้สอดส่อง มีข้อเสียเนื่องมาจากขีดจำกัดทางกายภาพของมนุษย์ เมื่อต้องใช้สายตานาน ๆ จะเกิดอาการอ่อนล้า ทำให้ประสิทธิภาพของการสอดส่องลดลงและอาจใช้เวลานานในการตรวจหา หากวีดิทัศน์มีจำนวนมาก ดังนั้นในระยะต่อมาเทคโนโลยีการประมวลผลภาพวีดิทัศน์ในลักษณะต่าง ๆ ได้ถูกนำมาประยุกต์ใช้ เช่น การตรวจหาลักษณะการเคลื่อนที่ที่ผิดปกติภายในขอบเขตภาพ เป็นต้น และเมื่อพิจารณาที่กระบวนการทำงาน จะเห็นว่าแม้ในระบบสอดส่องรักษาความปลอดภัยที่ผ่านมาจะมีกล้องวีดิทัศน์ภายในระบบหลายตัว แต่ไม่มีการประมวลผลร่วมกันเลย กล่าวคือ กล้องวีดิทัศน์ตัวหนึ่งก็จะถูกเฝ้าสังเกตผ่านจอคอมพิวเตอร์หนึ่งเครื่อง ดังนั้นจึงเกิดเป็นแนววิจัยและลักษณะงานในรูปแบบใหม่ที่มีการประมวลผลภาพร่วมกัน (co-operate) หรือใช้ความสัมพันธ์ร่วมกันระหว่างกล้องตั้งแต่ 2 ตัวขึ้นไป ให้เกิดประโยชน์ เช่น ใช้ตำแหน่งของขอบเขตการรับภาพของกล้องแต่ละตัวที่เหลื่อมกันอยู่ มาช่วยเพิ่มประสิทธิภาพในการตรวจหาวัตถุเป้าหมาย โดยสามารถทำนายล่วงหน้าได้ว่า จะมีวัตถุเคลื่อนที่เข้าไปยังมุมมองกล้องตัวใด และในทิศทางใด เป็นต้น เรียกกระบวนการที่ทำงานในลักษณะนี้ว่า ระบบกล้องหลายตัว (Multi-camera system) ประโยชน์ที่ได้

จากการทำงานร่วมกันในลักษณะรูปแบบของระบบนี้ ได้แก่ การขยายอาณาเขตครอบคลุมของการรับภาพของบริเวณที่สนใจ ประโยชน์ทางการคำนวณหาความลึกของวัตถุเป้าหมาย ประโยชน์ทางการติดตามวัตถุเป้าหมายระหว่างกล้อง และการเพิ่มโอกาสในการเลือกมุมมองการรับภาพวัตถุเป้าหมายในมุมมองหรือตำแหน่งที่ดีและเหมาะสมกว่า เป็นต้น ทำให้ในปัจจุบัน มีงานวิจัยที่มุ่งพัฒนาทั้งในส่วนขั้นตอนวิธีและลักษณะการทำงานของระบบให้มีประสิทธิภาพและประโยชน์การใช้งานในรูปแบบต่าง ๆ ให้เพิ่มมากขึ้น

รูปแบบการทำงานของกล้องหลายตัวโดยทั่วไปจะใช้กล้องวีดิทัศน์อยู่กับที่ (fixed/stationary camera) (กล้องวีดิทัศน์ที่มีทั้งตำแหน่งมุมมองการรับภาพและอัตราซูมที่แน่นอน) ซึ่งเป็นข้อจำกัดของระบบในกรณีที่หากวัตถุเป้าหมายเคลื่อนที่ออกจากมุมมองการรับภาพ ทำให้ไม่สามารถติดตามวัตถุเป้าหมายต่อเนื่องได้ และในกรณีที่ต้องการซูมภาพก็ไม่สามารถทำได้ รวมทั้งด้วยความก้าวหน้าทางเทคโนโลยีทางด้านกล้องวีดิทัศน์ ทำให้เริ่มมีการนำกล้องวีดิทัศน์ PTZ (PTZ camera) (กล้องวีดิทัศน์ที่ตำแหน่งมุมมองการรับภาพสามารถเปลี่ยนแปลงได้ในลักษณะการเคลื่อนที่แนวสาย (pan) และแนวก้มเงย (tilt) และสามารถซูม (zoom) รายละเอียดของวัตถุเป้าหมายได้ โดยลักษณะทั้งหมดสามารถควบคุมผ่านคอมพิวเตอร์ส่วนบุคคล) มาประยุกต์ใช้งานเพื่อลดข้อจำกัดดังที่ได้กล่าวมา ทำให้เกิดรูปแบบงานที่หลากหลายและประโยชน์ที่เพิ่มมากขึ้น ดังนั้นวิทยานิพนธ์นี้จึงมีแนวความคิดในการจำลองระบบกล้องหลายตัวสำหรับการตรวจหาและติดตามใบหน้าบุคคลที่สนใจ เพื่อนำไปประยุกต์ใช้งานในด้านการรู้จำบุคคล โดยใช้กล้องวีดิทัศน์ PTZ ในระบบ 2 ตัว ทำงานร่วมกันในลักษณะการประมวลผลร่วมกันและแลกเปลี่ยนข้อมูลระหว่างกัน

เมื่อพิจารณางานประยุกต์กล้องหลายตัว โดยเฉพาะงานทางการตรวจหาและติดตามบุคคลโดยใช้กล้องวีดิทัศน์ PTZ ส่วนใหญ่จะเป็นการทำงานร่วมกันระหว่างกล้องวีดิทัศน์อยู่กับที่กับกล้องวีดิทัศน์ PTZ กล้องวีดิทัศน์อยู่กับที่จะถูกกำหนดให้เป็นกล้องตัวหลัก (master camera) สำหรับประมวลผลตรวจหาและส่งคำสั่งหรือพารามิเตอร์ต่าง ๆ ไปยังกล้องวีดิทัศน์ PTZ เพื่อให้กล้องวีดิทัศน์ PTZ เปลี่ยนมุมมองการรับภาพติดตามบุคคลที่สนใจ และซูมเข้า (zoom in) ไปยังบุคคลเป้าหมาย กล่าวคือ กล้องวีดิทัศน์ PTZ ไม่ต้องประมวลผลภาพ ซึ่งเป็นข้อดีของระบบในลักษณะแบบนี้ เพราะว่าขั้นตอนวิธีการตรวจหาและติดตามบุคคลที่จะนำไปใช้กับกล้องวีดิทัศน์ PTZ นั้นมีความยุ่งยากมากกว่าและประสิทธิภาพที่น้อยกว่ากล้องวีดิทัศน์อยู่กับที่ แต่ก็มีข้อเสียที่ ต้องใช้กล้องวีดิทัศน์เพิ่มขึ้นในระบบอย่างน้อย 1 ตัว โดยงานประยุกต์ในงานวิจัยที่ผ่านมาจะใช้สำหรับระบบสอดส่องรักษาความปลอดภัยในมุมมองสูง [1, 2] เช่น บริเวณลานจอดรถ เป็นต้น



เมื่อพิจารณางานประยุกต์ทางการรู้จำใบหน้าบุคคล (face recognition) สิ่งที่ต้องพิจารณาคือ การทำงานในเวลาจริง (real time processing) ซึ่งส่วนใหญ่ในงานวิจัยที่ผ่านมาจะมีขั้นตอนวิธีที่ซับซ้อน และใช้ทรัพยากรในการประมวลผลสูง ทำให้ขั้นตอนวิธีที่ใช้ไม่เหมาะสมกับการนำมาประยุกต์ใช้ในสถานการณ์จริง แต่ในงานวิจัย [3, 4] ได้นำเสนอวิธีการรู้จำใบหน้าบุคคลที่ใช้ในเวลาจริง โดยพิจารณา รูปแบบตำแหน่งใบหน้าบุคคลที่รับมาจากกล้องวิดีโอทัศน์เพื่อใช้ในการประมวลผล ต้องเป็นใบหน้าในลักษณะหน้าตรงเท่านั้น และกล้องวิดีโอทัศน์ที่รับภาพเป็นกล้องวิดีโอทัศน์อยู่กับที่ ทำให้สามารถเลือกใช้ขั้นตอนวิธีที่เหมาะสมต่อการประมวลผลในเวลาจริง และมีประสิทธิภาพต่อการประมวลผลตรวจหาบริเวณใบหน้า มากกว่างานประยุกต์ที่ต้องใช้กล้องวิดีโอทัศน์ PTZ

วิทยานิพนธ์นี้จึงเป็นเสมือนจุดเริ่มต้นในการนำข้อจำกัดของงานวิจัยทางด้านนี้ มาปรับปรุง เสนอความคิด และพัฒนาให้ระบบสอดส่องรักษาความปลอดภัยที่ฉลาดขึ้นสามารถทำงานได้ในเวลาจริง ทั้งกระบวนการตรวจหา กระบวนการติดตาม และกระบวนการระบุบุคคล และมีความเป็นไปได้มากขึ้นในการนำไปประยุกต์ใช้จริงกับงานประยุกต์ตามสถานที่ต่าง ๆ ได้อย่างมีประสิทธิภาพ

## 1.2 ภาพรวมของงานวิจัยที่เกี่ยวข้อง

ในหัวข้อนี้จะกล่าวถึงภาพรวมของงานวิจัยที่เกี่ยวข้อง ซึ่งจะแบ่งเป็น 2 ส่วนหลัก คือ ส่วนแรกจะกล่าวถึงงานวิจัยที่นำเสนอระบบกล้องหลายตัวในรูปแบบต่าง ๆ โดยจะอธิบายถึงลักษณะ รูปแบบ และการนำไปใช้งานของแต่ละงานวิจัย และส่วนที่สอง จะกล่าวถึงงานวิจัยที่นำเสนอขั้นตอนวิธีแบบต่าง ๆ ของการตรวจหาติดตามสำหรับประยุกต์ใช้กับกล้องวิดีโอทัศน์ PTZ รวมทั้งรายละเอียดเกี่ยวกับงานวิจัยอ้างอิงในขั้นตอนวิธีการตรวจหาด้วยสีผิว

### 1.2.1 งานวิจัยที่นำเสนอระบบกล้องหลายตัวในรูปแบบต่าง ๆ

ในส่วนนี้จะกล่าวถึงเฉพาะระบบกล้องหลายตัวที่มีการประยุกต์ใช้กล้องวิดีโอทัศน์ PTZ เท่านั้น ซึ่งรูปแบบงานประยุกต์ใช้งาน ได้แก่ ระบบสอดส่องรักษาความปลอดภัย ระบบสำหรับรู้จำกิริยาท่าทาง (Gesture recognition) ระบบการเรียนการสอนทางไกล (Distance lecture) เป็นต้น

Ser-Nam Lim และคณะ [1,2] เสนอระบบสอดส่องรักษาความปลอดภัยอัตโนมัติ โดยใช้กล้องในระบบ 3 ตัว ประกอบด้วยกล้องวิดีโอทัศน์อยู่กับที่ 1 ตัว ทำหน้าที่เป็นกล้องตัวหลัก และกล้องวิดีโอทัศน์ PTZ อีก 2 ตัว ทำหน้าที่เปลี่ยนมุมมองติดตามบุคคลที่สนใจในแนวสาย และก้ม



เงย รวมทั้งชุมชนขยายรายละเอียดของบุคคลเป้าหมายที่สนใจ โดยจะประมวลผลเฉพาะภาพที่รับจากกล้องวีดิทัศน์อยู่กับที่ (กล้องตัวหลัก) เท่านั้น ซึ่งงานวิจัยนี้จะใช้ขั้นตอนวิธีกระบวนการลบฉากหลังแบบปรับตัวได้ (มีกระบวนการปรับค่าความเข้มสีของฉากหลังที่ถูกเก็บไว้ก่อนหน้า ตามการเปลี่ยนแปลงของฉากหลัง เช่น ในกรณีมีการเปลี่ยนแปลงความเข้มของแสงสว่าง หรือ มีการเคลื่อนไหวของฉากหลังที่ไม่ใช่วัตถุเป้าหมาย เป็นต้น) ซึ่งทำให้เพิ่มประสิทธิภาพในผลของการลบฉากหลังได้ เมื่อได้ตำแหน่งของบุคคลเป้าหมายแล้วจะอ้างอิงไปยังพิกัดของกล้องวีดิทัศน์ PTZ เพื่อคำนวณพารามิเตอร์การเคลื่อนที่ของแต่ละกล้องที่ต้องใช้สำหรับเคลื่อนที่ไปหาเป้าหมาย ดังนั้นข้อดีของระบบนี้ คือ ลดขั้นตอนในการประมวลผลภาพที่รับมาจากกล้องวีดิทัศน์ PTZ

S. Tsuruoka และคณะ [5] ประยุกต์ใช้กล้องวีดิทัศน์ PTZ ในการแก้ปัญหา ภาพด้านปลายทางไม่ชัดเจนเพียงพอต่อการเรียนการสอนทางไกล โดยได้เสนอระบบที่ประกอบด้วยกล้องวีดิทัศน์ 2 ตัว คือ กล้องวีดิทัศน์อยู่กับที่และกล้องวีดิทัศน์ PTZ กล้องวีดิทัศน์อยู่กับที่ ทำหน้าที่ตรวจหาใบหน้า มือ และลำตัว ของผู้สอน จากนั้นใช้กระบวนการแยกลักษณะพฤติกรรมของผู้สอน (Lecturer's behavior) เพื่อใช้เป็นตัวกำหนดตำแหน่งการเปลี่ยนมุมมองการรับภาพของกล้องวีดิทัศน์ PTZ เช่น หากผู้สอนกำลังอธิบายสมการที่กระดานดำ กล้องวีดิทัศน์ PTZ ก็จะต้องส่าย ก้มเงยและซูมไปยังตำแหน่งกระดานดำ เพื่อขยายรายละเอียดให้ด้านผู้เรียนเห็นรายละเอียดชัดเจนขึ้น ดังรูปที่ 1.2 ทั้งนี้กล้องวีดิทัศน์ PTZ ไม่ต้องประมวลผลเพื่อหาตำแหน่งวัตถุเอง



รูปที่ 1.1 การประยุกต์ใช้กับระบบสอดส่องรักษาความปลอดภัย [1]



รูปที่ 1.2 งานประยุกต์กล้องวีดิทัศน์ PTZ สำหรับงานการเรียนการสอนทางไกล [5]

J. Kang และคณะ [6] นำเสนอระบบติดตามวัตถุเป้าหมายระหว่างกล้องวีดิทัศน์อยู่กับที่และกล้องวีดิทัศน์ PTZ โดยมีการส่งต่อข้อมูลของวัตถุเป้าหมายระหว่างกล้องทั้งสอง โดยลักษณะสีจะถูกส่งต่อจากกล้องวีดิทัศน์หนึ่งตัวไปยังอีกตัวหนึ่งเพื่อระบุเป้าหมายที่ต้องการจากกล้องแต่ละตัว สำหรับกระบวนการตรวจหาสำหรับกล้องวีดิทัศน์ PTZ จะเสนอกระบวนการปรับฉากหลัง โดยใช้การแปลงสัมพรรค (affine transform) เพื่อคำนวณการเคลื่อนที่ของแต่ละจุดภาพที่สัมพันธ์กับพารามิเตอร์ในการถ่ายและกัมเมยของกล้องวีดิทัศน์ PTZ เพื่อพิจารณาจุดภาพว่าเป็นการเคลื่อนที่อันเนื่องจากการถ่ายและกัมเมยของกล้องวีดิทัศน์หรือการเคลื่อนที่ของวัตถุ ส่วนในการติดตามบุคคลเป้าหมายระหว่างกล้อง นอกจากจะใช้ลักษณะสีแล้ว ยังจะใช้ลักษณะการแปลงโฮโมกราฟี (homography transform) มาอ้างอิงเป้าหมายระหว่างกล้องวีดิทัศน์เพราะว่ากล้องวีดิทัศน์ทั้งสองมีมุมมองการรับภาพ (field of view) ที่ซ้อนเหลื่อมกันอยู่ (overlapping view)

A. Hampapur และคณะ [7] สร้างระบบเพื่อจัดเก็บข้อมูลใบหน้า (face cataloger) เพื่อสำหรับนำไปใช้ในกระบวนการรู้จำใบหน้าต่อไป ระบบที่นำเสนอประกอบด้วยกล้องวีดิทัศน์ทั้งหมด 4 ตัว กล้องวีดิทัศน์ 2 ตัวแรกจะเป็นกล้องวีดิทัศน์อยู่กับที่ จะประมวลผลสเตอริโอ (stereo) ร่วมกันเพื่อคำนวณหาตำแหน่งวัตถุในพิกัด 3 มิติ ซึ่งจะใช้ตำแหน่งนี้ในการสั่งให้กล้องวีดิทัศน์อีก 2 ตัวที่เป็นกล้องวีดิทัศน์ PTZ ให้ถ่ายและกัมเมยไปหาเป้าหมาย และซูมไปที่ใบหน้าบุคคล โดยจะเลือกส่งกล้องวีดิทัศน์ PTZ ที่มีตำแหน่งเหมาะสมต่อการถ่าย และกัมเมยหาใบหน้ามากที่สุด ดังตัวอย่างในรูปที่ 1.3 ข้อดีของระบบ คือ ได้ข้อมูลเกี่ยวกับตำแหน่ง ระยะทาง และแนวทางการเคลื่อนที่ (motion trajectory) ของบุคคล แต่ในรูปแบบงานประยุกต์ที่ไม่ต้องการข้อมูลในส่วนนี้ จะต้องใช้กล้องวีดิทัศน์อย่างน้อย 3 ตัวในการติดตามบุคคลเพียงคนเดียว ทำให้ต้องเปลืองจำนวนกล้องวีดิทัศน์ที่ต้องใช้กับระบบ



รูปที่ 1.3 ระบบตรวจหาและติดตามใบหน้าสำหรับจัดเก็บเป็นฐานข้อมูลใบหน้า [7]

H. Hongo และคณะ [8] เสนอระบบที่มีรูปแบบและจำนวนกล้องที่คล้ายกับงานของ A. Hampapur แต่นำไปใช้ในงานประยุกต์การรู้จำการทำทางของมนุษย์ โดยใช้กล้องวีดิทัศน์ทั้งสองตัวถ่าย และกัมเมยติดตามใบหน้าหรือมือ รวมทั้งซูมเพื่อขยายรายละเอียดของทั้งใบ

หน้าและมือให้เหมาะต่อการนำไปใช้ในกระบวนการรู้จำ ดังนั้นระบบสามารถรองรับการติดตามวัตถุได้ 2 อย่างพร้อมกันในเวลาเดียวกัน เท่ากับจำนวนกล้องวิดีโอ PTZ ที่มีในระบบ รูปที่ 1.4 แสดงตัวอย่างการใช้งานของระบบนี้



รูปที่ 1.4 งานประยุกต์ใช้สำหรับการรู้จำกิจกรรมท่าทางโดยใช้กล้องวิดีโอ PTZ เคลื่อนที่และซูมหาใบหน้าและมือ [8]

จากงานวิจัยที่กล่าวมา จะเห็นว่าส่วนใหญ่ระบบจะใช้กล้องวิดีโออยู่กับที่ในการประมวลผลภาพและวิดีโอ เพื่อตรวจหาบริเวณวัตถุเป้าหมาย และคำนวณหามุมในการส่ายและการก้มเงยเพื่อส่งคำสั่งไปให้กับกล้องวิดีโอ PTZ เพื่อติดตามวัตถุเป้าหมาย ซึ่งเป็นการสิ้นเปลืองทรัพยากรกล้องวิดีโอที่ต้องใช้ในระบบ แต่จะมีข้อดีที่การตรวจหาจะมีความถูกต้องมากกว่าการตรวจหาโดยใช้กล้องวิดีโอ PTZ

## 1.2.2 งานวิจัยที่เกี่ยวข้องกับขั้นตอนวิธี

### 1.2.2.1 งานวิจัยที่นำเสนอขั้นตอนวิธีการตรวจหาติดตามสำหรับประยุกต์ใช้กับกล้องวิดีโอ PTZ

ขั้นตอนวิธีการตรวจหาและติดตามวัตถุมีการนำมาใช้กับภาพที่รับจากกล้องวิดีโออยู่กับที่ เช่น ขั้นตอนวิธีการลบฉากหลังเบื้องต้น (basic background subtraction) เป็นต้น ไม่สามารถนำมาประยุกต์ใช้กับภาพที่รับจากกล้องวิดีโอ PTZ ได้โดยตรง เนื่องจากว่าฉากหลังของภาพที่รับจากกล้องจะเปลี่ยนแปลงภาพไปทุกครั้ง ดังนั้นจึงมีงานวิจัยต่าง ๆ ที่เสนอวิธีในการตรวจหาและติดตามวัตถุเป้าหมายโดยประยุกต์ใช้สำหรับกล้องวิดีโอ PTZ สามารถจำแนกตามขั้นตอนวิธีได้ดังนี้

- **ขั้นตอนวิธีเก็บฉากหลัง (recorded background) ตามตำแหน่งการถ่ายและกัมเมยของกล้องวิดีโอไว้ล่วงหน้า**

งานวิจัย [9, 10, 11, 12] ได้เสนอขั้นตอนวิธีการตรวจหาตำแหน่งวัตถุที่สนใจ โดยเลือกใช้ขั้นตอนวิธีการลบฉากหลังเบื้องต้นเพื่อแยกบุคคลที่เป็นฉากหน้า และเลือกใช้วิธีการเก็บฉากหลังทุกตำแหน่งที่ต้องการให้กล้องวิดีโอ PTZ เคลื่อนที่ได้ เพราะฉะนั้นข้อเสียของขั้นตอนวิธีนี้ คือ การเคลื่อนที่ของกล้องจะไม่เป็นอิสระ จะต้องเคลื่อนที่ในตำแหน่งที่กำหนดไว้เท่านั้น แต่ข้อดี คือ ประสิทธิภาพในการตรวจหาวัตถุ จะมีความถูกต้องและน่าเชื่อถือมากกว่าขั้นตอนวิธีแบบอื่น (เช่น การตรวจหาด้วยลักษณะการเคลื่อนที่)

R. T. Collins และคณะ [12] เสนอระบบวิดีโอที่สอดคล้องและรักษาความปลอดภัย โดยเน้นที่งานประยุกต์ใช้บริเวณลานจอดรถ ซึ่งจะทำให้การเก็บภาพฐานข้อมูลฉากหลังในตำแหน่งการถ่ายและกัมเมยของกล้องวิดีโอ PTZ ดังรูปที่ 1.5 ในการประมวลผลจะเลือกฉากหลังที่มีตำแหน่งใกล้เคียงต่อตำแหน่งของกล้องวิดีโอ ณ มุมถ่ายและกัมเมยขณะนั้นมากที่สุด แล้วใช้กระบวนการชดเชยการเคลื่อนที่เพื่อปรับให้ภาพรับเข้าที่จะนำมาเทียบกับภาพฉากหลังในฐานข้อมูลมีความใกล้เคียงในแต่ละตำแหน่งจุดภาพมากขึ้น

Yu-Wen Huang และคณะ [11] เสนอระบบการติดตามบุคคลโดยใช้กล้องวิดีโอ PTZ ขั้นตอนวิธีในการเก็บฉากหลังจะทำการแบ่งฉากหลังทั้งหมดที่กล้องวิดีโอ PTZ สามารถถ่ายและกัมเมยไปได้ออกเป็นกรอบสี่เหลี่ยม ซึ่งแต่ละกรอบสี่เหลี่ยมจะแทนฉากหลังหนึ่งตำแหน่งที่ไม่มีการซ้อนทับกันระหว่างฉากหลังแต่ละตำแหน่ง โดยระบบจะทำการลบฉากหลังเพื่อตรวจหาบุคคล ในกรณีที่กล้องวิดีโอ PTZ ถ่ายและกัมเมยมายังตำแหน่งกรอบสี่เหลี่ยมที่มีการเก็บฉากหลังไว้เท่านั้น

Y. Ye และคณะ [10] เสนอวิธีการใช้กล้องวิดีโอ PTZ ในการติดตามบุคคล โดยใช้ขั้นตอนวิธีลบฉากหลัง เพื่อตรวจหาบริเวณบุคคลที่สนใจ โดยกำหนดเป็นสถานะการเคลื่อนที่ ซึ่งแต่ละสถานะการเคลื่อนที่จะมีการกำหนดภาพฉากหลังสำหรับกระบวนการลบฉากหลังที่แน่นอน กล่าวคือ สมมติปัจจุบันระบบอยู่ที่สถานะ ก เมื่อบุคคลมีการเคลื่อนที่ ระบบจะคำนวณพารามิเตอร์ที่กล้องวิดีโอ PTZ จะต้องใช้ในการถ่าย และกัมเมยเพื่อติดตามวัตถุ จากนั้นจะนำตำแหน่งพารามิเตอร์นี้ไปอ้างอิงกับสถานะที่ถูกกำหนดไว้แน่นอนแล้วว่า สถานะการเคลื่อนที่ที่กล้องวิดีโอ PTZ จะต้องเคลื่อนที่คือตำแหน่งใด

T. WADA และคณะ [9] ได้เสนอแบบจำลองวงกลมสภาพปรากฏ (appearance sphere model) ที่ใช้กับกล้องวิดีโอแบบมุมมอง 360 องศา ในการเก็บภาพฉากหลังในลักษณะ

วงกลมรอบทิศ และใช้พารามิเตอร์ของการส่ายและก้มเงยของกล้องวิดีโอเพื่อคำนวณและสร้างฉากหลังขึ้นมาสำหรับกระบวนการลบฉากหลัง



รูปที่ 1.5 ภาพฉากหลังที่ถูกบันทึกไว้ ณ ตำแหน่งการเคลื่อนที่ของกล้อง [12]

นอกจากขั้นตอนวิธีการเก็บฉากหลังตามตำแหน่งการส่ายและก้มเงยของกล้องวิดีโอที่ค้นไว้ล่วงหน้า จะทำให้การส่ายและก้มเงยของกล้องวิดีโอ PTZ ไม่เป็นอิสระ ต้องเปลี่ยนมุมมองการรับภาพตามตำแหน่งที่กำหนดไว้แล้ว ยังต้องเปลืองเนื้อที่หน่วยความจำจำนวนมากในการเก็บฐานข้อมูลฉากหลังในแต่ละตำแหน่งการเคลื่อนที่ของกล้องวิดีโอ PTZ และมีความยุ่งยากในขั้นตอนการเก็บภาพฉากหลัง

#### - ขั้นตอนวิธีสร้างฉากหลังจากภาพโมเสก (mosaic image)

I. Haritaoglu และคณะ [13] สร้างระบบรักษาความปลอดภัยภายนอกอาคาร โดยใช้กล้องวิดีโอ PTZ เพื่อส่าย ก้มเงย และซูมหาบุคคลที่เคลื่อนที่ภายในอาณาบริเวณที่กล้องสามารถส่ายและก้มเงยไปได้ โดยวิธีการในการแยกฉากหน้าที่เป็นบุคคลออกจากฉากหลังนั้น ระบบจะเก็บฉากหลังทั้งหมดที่กล้องสามารถส่ายและก้มเงยไปได้ให้อยู่ในรูปของภาพโมเสก ซึ่งในการนำไปใช้งานนั้น ฉากหลัง ณ ตำแหน่งดังกล่าวจะถูกสร้างขึ้นโดยอ้างอิงกับพิกัดภาพโมเสก และต้องสัมพันธ์กับมุมส่ายและมุมก้มเงยที่กล้องได้เคลื่อนที่ไปแล้ว เมื่อได้ฉากหลังแล้ว ก็จะใช้ขั้นตอนวิธีการลบฉากหลังเพื่อหาส่วนฉากหน้า รูปที่ 1.6 แสดงตัวอย่างของภาพโมเสกฉากหลังที่ถูกสร้างขึ้น



รูปที่ 1.6 ภาพฉากหลังแบบโมเสก [13]



### - ขั้นตอนวิธีการตรวจหาการเคลื่อนที่

D. Murray และคณะ [14] เสนอการติดตามบุคคลโดยใช้ลักษณะการเคลื่อนที่ร่วมกับลักษณะขอบ เพื่อหาบริเวณเค้าโครงบุคคลเป้าหมาย โดยภาพที่ใช้ในการประมวลผลจะรับมาจากกล้องวิดีโออยู่กับที่ที่ถูวางอยู่บนฐานที่สามารถหมุนได้ ดังนั้นจึงสามารถถ่ายและกัมเงยติดตามบุคคลเป้าหมาย ซึ่งได้เสนอวิธีแก้ปัญหาในกรณีที่กล้องมีการเคลื่อนที่ โดยใช้การชดเชยฉากหลัง (background compensation) เพื่อให้สามารถใช้รูปแบบการตรวจหาการเคลื่อนที่และขอบ ที่มีใช้อยู่แล้ว ให้ยังสามารถทำงานได้ตามปกติ

#### 1.2.2.2 งานวิจัยที่นำเสนอขั้นตอนวิธีการตรวจหาด้วยลักษณะสีผิว

ในวิทยานิพนธ์นี้ได้พิจารณาใช้ลักษณะสีผิวเป็นลักษณะหนึ่งในการตรวจหาตำแหน่งใบหน้า ดังนั้นจึงเหมาะสมกับการนำมาประยุกต์ใช้สำหรับการตรวจหาที่ประยุกต์ใช้กับกล้องวิดีโอ PTZ ทั้งนี้เพราะลักษณะสีไม่ขึ้นกับตำแหน่งของฉากหลังที่เปลี่ยนไป รวมทั้งยังสามารถรองรับการตรวจหาวัตถุที่มีการเปลี่ยนแปลงรูปร่างได้ และสามารถรองรับในกรณีที่วัตถุมีการเปลี่ยนแปลงขนาดอันเนื่องมาจากการซูมของกล้องวิดีโอ

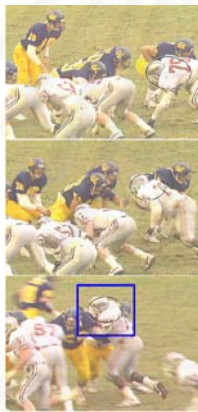
Rein-Lien Hsu และคณะ [15] เสนอขั้นตอนวิธีการตรวจหาใบหน้า โดยใช้ปริภูมิสี YCbCr จำลองการกระจายตัวขององค์ประกอบ Cb และ Cr เป็นรูปวงรีและสร้างเป็นขอบเขตวงรีขึ้นมาเพื่อแยกจุดสีที่มีความน่าจะเป็นส่วนผิวเนื้อมนุษย์ออกมา รวมทั้งยังเสนอวิธีการชดเชยแสง (lighting compensation) เพื่อลดผลกระทบของแสงที่มีผลต่อประสิทธิภาพในการตรวจหา ส่วนวิธีในการแยกใบหน้าออกจากบริเวณอื่นที่เป็นการตรวจหาผิดพลาด จะใช้โครงสร้างของตำแหน่งตาและปากเพื่อเลือกบริเวณใบหน้าที่แท้จริงออกมา

N. Soontranon และคณะ [16] ได้เสนอขั้นตอนวิธีการตรวจหาและติดตามใบหน้าและมือโดยใช้ลักษณะสีผิว สำหรับงานประยุกต์ทางด้านการแปลความหมายภาษามือ โดยปริภูมิสีที่เลือกใช้ คือ ปริภูมิสี YCbCr และเลือกใช้เฉพาะข้อมูลองค์ประกอบ Cb และ Cr เท่านั้น เพื่อลดความซับซ้อนในการคำนวณและประมวลผล ขั้นตอนวิธีตรวจหาที่นำเสนอจะสร้างแบบจำลองวงรี (Elliptical boundary) เพื่อใช้เป็นตัวกำหนดขอบเขตค่าจุดสีในพิกัด CbCr ว่าขอบเขตใดเป็นส่วนสีผิวและส่วนที่ไม่ใช่สีผิว ซึ่งแบบจำลองวงรีจะถูกคำนวณขึ้นจากตัวอย่างสีผิวใบหน้าของกลุ่มบุคคลต้นแบบ นอกจากนั้นยังมีการประมวลผลภาพเชิงสัณฐาน (morphological image processing) กับภาพฐานสองที่ได้จากการตรวจหา เพื่อลดสัญญาณรบกวนและเพิ่มประสิทธิภาพในการวิเคราะห์บริเวณวัตถุที่สนใจ

### 1.2.2.3 งานวิจัยที่ประยุกต์ใช้ระยะ Bhattacharyya

วิทยานิพนธ์นี้เลือกใช้ระยะ Bhattacharyya เพื่อใช้ในการจับคู่ภาพใบหน้ารับเข้า กับภาพใบหน้าในฐานะข้อมูลว่ามีความใกล้เคียงกับภาพใบหน้าฐานข้อมูลใดมากที่สุด เพื่อใช้ระบุ บุคคลที่กล้องวีดิทัศน์รับภาพได้ ณ ขณะนั้น

D. Comaniciu และคณะ [17] ใช้ระยะ Bhattacharyya มาประยุกต์ใช้สำหรับ สร้างระบบการติดตามวัตถุที่มีประสิทธิภาพและความเร็วในการประมวลผล โดยประมวลผลกับ ลำดับวีดิทัศน์ (video sequence) ทั้งนี้ลักษณะการทำงานเริ่มต้นจะต้องมีการกำหนดตำแหน่ง ของขอบเขตวัตถุที่สนใจก่อน (กำหนดขอบเขตในรูปแบบวงรี) ดังตัวอย่างในรูป 1.7 จากนั้นจะ คำนวณหาบริเวณที่มีระยะใกล้เคียงกับบริเวณขอบเขตต้นแบบมากที่สุด โดยรูปที่ 1.7(ก) ประมวลผลกับลำดับวีดิทัศน์กีฬา และรูปที่ 1.7(ข) ประมวลผลกับลำดับวีดิทัศน์ที่ถูกบันทึกจาก กล้องวีดิทัศน์อยู่กับที่บริเวณสถานีรถไฟ



(ก)



(ข)

รูปที่ 1.7 การติดตามบริเวณที่สนใจโดยใช้ ระยะ Bhattacharyya (ก) ในลำดับวีดิทัศน์กีฬา (ข) วีดิทัศน์รักษาความปลอดภัยในสถานีรถไฟ [17]

K. Nummiaro และคณะ [18] ได้เสนอระบบห้องฉลาด (smart room) ที่ ประกอบด้วยกล้องวีดิทัศน์อยู่กับที่ 3 ตัว ระบบจะเลือกกล้องวีดิทัศน์ที่มีมุมมองการรับภาพใบหน้า ที่ใกล้เคียงกับภาพใบหน้าต้นแบบมากที่สุดเพื่อเลือกเก็บเฉพาะข้อมูลที่ดีที่สุดสำหรับบุคคลที่สนใจ เริ่มต้นระบบจะทำการเก็บค่าฮิสโทแกรมสีของบริเวณใบหน้าด้านตรงไว้ดังรูปที่ 1.8 จากนั้นในการ ประมวลผลจะใช้ระยะ Bhattacharyya เพื่อคำนวณหาระยะเปรียบเทียบของบริเวณใบหน้าของ แต่ละกล้องวีดิทัศน์เทียบกับภาพใบหน้าอ้างอิง ดังผลการเปรียบเทียบรูปที่ 1.8





รูปที่ 1.8 ระบบห้องฉลาดเพื่อเลือกมุมมองที่ดีที่สุดในการรับภาพ [18]

### 1.2.3 งานวิจัยที่เกี่ยวกับระบบการรู้จำใบหน้าที่ใช้ในเวลาจริง

ที่ผ่านมา มีงานวิจัยมากมายที่ให้ความสนใจศึกษาเกี่ยวกับการพัฒนาขั้นตอนวิธีในการรู้จำใบหน้า ซึ่งส่วนมากจะประมวลผลกับภาพที่มีอยู่ในฐานข้อมูลกลาง ดังที่มีการสำรวจอยู่ในงานวิจัย [19] ซึ่งเมื่อพิจารณาในกรณีที่ต้องนำมาประยุกต์ใช้กับระบบในเวลาจริง ปัญหาด้านความซับซ้อนของขั้นตอนวิธีจะมีผลอย่างมากต่อเวลาในการประมวลผล แต่ก็ม้งานวิจัยบางงานที่เสนอระบบที่สามารถประยุกต์ใช้ในเวลาจริง ดังในงานวิจัย [3, 4]

M.J. Seow และคณะ [3] เสนอรูปแบบการรู้จำใบหน้าในเวลาจริง โดยใช้ขั้นตอนวิธี Composite Principal Component Analysis (CPCA) ซึ่งรับภาพจากกล้องวีดิทัศน์อยู่กับที่สามารถประมวลผลได้ที่อัตราภาพ 10 ภาพต่อวินาที ที่ขนาดภาพกว้าง 320 จุดภาพ ยาว 240 จุดภาพ และสามารถรองรับการเปลี่ยนมุมมองเฉียงจากใบหน้าตรงได้เล็กน้อย ดังตัวอย่างฐานข้อมูลและผลการรู้จำ ดังรูปที่ 1.10

F. Zuo และคณะ [4] ได้เสนอระบบบ้านฉลาด (Smart home) ซึ่งประกอบด้วยกระบวนการตรวจหา ติดตาม และรู้จำใบหน้าในเวลาจริง ที่รับภาพจากกล้องวีดิทัศน์อยู่กับที่ สามารถประมวลผลได้ที่อัตราภาพ 3-4 ภาพต่อวินาที ใช้การแยกส่วนประกอบของใบหน้า เช่น ปาก และตา เพื่อเป็นลักษณะ (feature) ในการจับคู่ใบหน้าที่ฐานข้อมูลใบหน้า สามารถรู้จำได้เฉพาะภาพใบหน้าตรง หรือใบหน้าที่มีมุมเอียงไม่เกิน  $\pm 20^\circ$  องศา เท่านั้น รูปที่ 1.9 แสดงตัวอย่างผลการตรวจหาใบหน้าในสถานที่จริง



รูปที่ 1.9 ตัวอย่างระบบการรู้จำใบหน้าที่ใช้ในเวลาจริง [4]



รูปที่ 1.10 ผลการรู้จำใบหน้า [3]

### 1.3 แนวทางวิทยานิพนธ์

วิทยานิพนธ์นี้นำเสนอระบบกล้องหลายตัวที่สามารถนำไปประยุกต์ใช้งานสำหรับระบบสอดส่องรักษาความปลอดภัยภายในอาคาร โดยระบบจะประกอบด้วยกล้องวีดิทัศน์ PTZ 2 ตัวซึ่งจะมีการประมวลผลร่วมกันเพื่อประโยชน์ในด้านการตรวจหาติดตาม และส่วนของการระบุบุคคล โดยใช้วิธีการจับคู่ใบหน้า วิทยานิพนธ์นี้จะใช้ประโยชน์ของระบบกล้องหลายตัวและความสามารถของกล้องวีดิทัศน์ PTZ กล่าวคือ ความสามารถในการส่ายและก้มเงยเปลี่ยนมุมมอง การรับภาพและการซูมของกล้องวีดิทัศน์ PTZ เพื่อให้ได้ตำแหน่งใบหน้าที่ดีที่สุด ประโยชน์ของระบบกล้องหลายตัว ที่จะมีมุมมองการรับภาพที่กว้างกว่า ทำให้เพิ่มโอกาสในการประมวลผลกับภาพที่มีความเหมาะสมมากที่สุด ประโยชน์ของการส่งต่อพารามิเตอร์ระหว่างกล้องเพื่อช่วยเพิ่มประสิทธิภาพในการประมวลผล

วิทยานิพนธ์นี้จึงจะเสนอขั้นตอนวิธีในการทำงานร่วมกันระหว่างกล้องวีดิทัศน์ PTZ ทั้งสองตัว โดยแบ่งการทำงานออกเป็น 2 ลักษณะ คือ

1. การประมวลผลร่วมกันสำหรับการระบุบุคคล โดยใช้วิธีการจับคู่ภาพใบหน้า การทำงานของทั้งสองกล้องจะทำงานในลักษณะขนาน(parallel) โดยจะประกอบด้วยกระบวนการตรวจหา ติดตาม และการจับคู่ใบหน้า ซึ่งจะแบ่งการจับคู่ใบหน้าเป็น 3 กลุ่ม ตามการเรียงตัวของมุมมองใบหน้า (หน้าตรง หน้าเอียงด้านซ้าย หน้าเอียงด้านขวา) โดยระบบจะคัดเลือกภาพใบหน้าที่อยู่ในตำแหน่งที่เหมาะสมกว่า ได้แก่ ใบหน้าตรงหรือค่อนข้างตรง ต่อการนำไปจับคู่ใบหน้าจากกล้องวีดิทัศน์ 2 ตัว ซึ่งในวิทยานิพนธ์นี้เลือกใช้การจับคู่ใบหน้าด้วยลักษณะสีสำหรับจำแนกบุคคล ซึ่งต้องมีข้อมูลภาพใบหน้าอยู่ภายในฐานข้อมูล

2. การทำงานร่วมกันสำหรับการตรวจหาติดตามใบหน้าอย่างต่อเนื่องระหว่างกล้องวีดิทัศน์ 2 ตัว โดยใช้ลักษณะสี แบ่งเป็นส่วนสีเสื้อ ส่วนสีผิว และส่วนสีผม และลักษณะทิศ

ทางการเคลื่อนที่ สำหรับระบุบุคคลระหว่างกล้องวีดิทัศน์ ในกรณีที่บุคคลเป้าหมายเคลื่อนที่พ้นขอบเขตมุมมองของกล้องวีดิทัศน์ตัวแรก กล้องตัวที่สอง จะทำการตรวจหาบุคคลและต้องระบุได้ว่าเป็นบุคคลเป้าหมายเดิม จึงจะทำการถ่าย ก้มเงย ติดตามและซูมหาเป้าหมาย

สำหรับขั้นตอนวิธีสำหรับตรวจหาใบหน้าที่ต้องประยุกต์ใช้กับกล้องวีดิทัศน์ PTZ ซึ่งเป็นส่วนสำคัญที่ต้องมีการพิจารณาในกรณีที่ต้องทำงานกับกล้องวีดิทัศน์ PTZ วิทยานิพนธ์นี้ได้เสนอกระบวนการตรวจหา โดยเลือกใช้ใช้ลักษณะสีผิวและลักษณะการเคลื่อนที่ ลักษณะสีผิวเป็นลักษณะหลักในการตรวจหาบริเวณที่มีความน่าจะเป็นส่วนสีผิว และใช้ลักษณะทางการเคลื่อนที่เป็นลักษณะเสริมที่ช่วยเพิ่มประสิทธิภาพในการตรวจหา รวมทั้งลดสัญญาณรบกวนที่เกิดจากบริเวณที่มีสีคล้ายส่วนสีผิว นอกจากนี้ยังได้ใช้กระบวนการประมวลผลภาพเชิงสัญญาณเกี่ยวกับการทำโดเลขันและอีโรชัน กับภาพฐานสองที่ได้จากการตรวจหา เพื่อประโยชน์ในการวิเคราะห์หาตำแหน่งบริเวณใบหน้าที่ต้องการ ซึ่งทำให้ระบบที่นำเสนอสามารถควบคุมการถ่าย และก้มเงยของกล้องวีดิทัศน์ PTZ ได้อย่างเป็นอิสระ และลดความซับซ้อนในการประมวลผล

#### 1.4 เป้าหมายและขอบเขตของวิทยานิพนธ์

- 1.4.1 ออกแบบและสร้างระบบจำลองการติดตามใบหน้าโดยใช้กล้องวีดิทัศน์ PTZ ทำงานร่วมกัน 2 ตัว
- 1.4.2 ระบบจะติดตามซูมขยายใบหน้าและหามุมมองของใบหน้าที่เหมาะสมต่อการนำไปประยุกต์ใช้ในการรู้จำใบหน้ามนุษย์ โดยมีบุคคลเพียงบุคคลเดียวภายในมุมมองการรับภาพ
- 1.4.3 วิธีในการตรวจหาส่วนสีผิวจะใช้ได้กับสภาพแวดล้อมภายในห้องปฏิบัติการ วิทยกรรมวิธีสัญญาณดิจิทัลที่ได้กำหนดตำแหน่งไว้เท่านั้น

#### 1.5 ขั้นตอนและวิธีการดำเนินงาน

- 1.5.1 ศึกษารูปแบบและลักษณะการนำไปใช้งานของระบบกล้องหลายตัวและหาแนวทางในการประยุกต์ใช้กับวิทยานิพนธ์
- 1.5.2 ศึกษาขั้นตอนวิธี และแบบจำลองที่ใช้ในการตรวจหาบริเวณสีผิวของมนุษย์
- 1.5.3 ศึกษาขั้นตอนวิธีในการตรวจหาโดยใช้ลักษณะการเคลื่อนที่
- 1.5.4 ออกแบบขั้นตอนวิธีในการประยุกต์ใช้การตรวจหาด้วยสีผิวทำงานร่วมกับการตรวจหาด้วยการเคลื่อนที่

- 1.5.5 ศึกษาและออกแบบขั้นตอนวิธีในการจัดการเกี่ยวกับการตัดส่วนภาพใบหน้าจากภาพรับเข้าให้เหมาะกับการประยุกต์ใช้รู้จำใบหน้า
- 1.5.6 ศึกษาขั้นตอนวิธีในกระบวนการจับคู่ใบหน้าโดยการวัดความคล้าย
- 1.5.7 ประยุกต์และเลือกใช้ขั้นตอนวิธีในการจับคู่ใบหน้าให้เหมาะสมกับระบบที่นำเสนอ
- 1.5.8 ศึกษาและเขียนโปรแกรมเพื่อส่งคำสั่งไปยังกล้องวีดิทัศน์ PTZ เพื่อให้กล้องตอบสนองในรูปแบบการทำงานตามที่ต้องการ
- 1.5.9 ศึกษาและออกแบบขั้นตอนวิธีในการควบคุมการเคลื่อนที่ของกล้องวีดิทัศน์ PTZ และหาพารามิเตอร์การเคลื่อนที่
- 1.5.10 ศึกษาและออกแบบลักษณะการทำงานร่วมกันระหว่างกล้อง 2 ตัว ในรูปแบบของการรู้จำใบหน้า และการติดตามวัตถุอย่างต่อเนื่องระหว่างกล้อง
- 1.5.11 เปรียบเทียบและวิเคราะห์ผลการจำลองระบบ
- 1.5.12 จัดทำรายงานฉบับสมบูรณ์

## 1.6 ประโยชน์ที่คาดว่าจะได้รับ

- 1.6.1 เข้าใจถึงความรู้พื้นฐานและการทำงานของระบบกล้องหลายตัว
- 1.6.2 เข้าใจถึงความรู้พื้นฐานเกี่ยวกับการตรวจหา และติดตามวัตถุที่ประยุกต์ใช้กับกล้องวีดิทัศน์ PTZ
- 1.6.3 สร้างระบบจำลองการตรวจหาและติดตามใบหน้าโดยใช้กล้องวีดิทัศน์ PTZ
- 1.6.4 สร้างระบบจำลองการทำงานร่วมกันระหว่างกล้อง สำหรับงานประยุกต์เกี่ยวกับการรู้จำใบหน้า และการติดตามบุคคลต่อเนื่องระหว่างกล้อง
- 1.6.5 เพื่อเป็นจุดเริ่มต้นของการรวมระบบทั้ง การตรวจหา ติดตามและการรู้จำเบื้องต้น ให้สามารถประมวลผลร่วมกันในเวลาจริง
- 1.6.6 สามารถนำไปประยุกต์ใช้กับงานรูปแบบต่าง ๆ ไม่ว่าจะเป็น งานด้านการรักษาความปลอดภัย หรืองานเกี่ยวกับการโต้ตอบระหว่างมนุษย์กับคอมพิวเตอร์

## บทที่ 2

### ความรู้พื้นฐานและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงความรู้และทฤษฎีพื้นฐานที่จำเป็นสำหรับวิทยานิพนธ์ โดยจะแบ่งเนื้อหาออกเป็น 5 หัวข้อหลัก คือ ปฏิภูมิของสี การประมวลผลภาพเชิงสัญญาณและการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน กระบวนการตรวจหาและติดตามใบหน้าและการวัดประสิทธิภาพการตรวจหา การควบคุมกล้องวิดีโอ PTZ และการจับคู่ใบหน้า

#### 2.1 ปฏิภูมิของสี

ค่าสี (Color) ถือเป็นลักษณะหลักในระบบการประมวลผลทางดิจิทัลของภาพและวิดีโอ โดยเฉพาะอย่างยิ่ง การทำงานเกี่ยวกับการตรวจหาใบหน้า งานวิจัยส่วนมากจะอาศัยข้อมูลของค่าสีเป็นสำคัญขั้นต้นในการตรวจหาบริเวณใบหน้า ค่าสีที่เกิดขึ้นจากการรับภาพวิดีโอจากกล้องวิดีโอ ซึ่งอาจอยู่ในรูปแบบสัญญาณดิจิทัล หรือแอนะล็อก โดยสัญญาณแบบแอนะล็อกจะถูกนำไปผ่านการสุ่ม (Sampling) และการแจกหน่วย (Quantization) เป็นสัญญาณดิจิทัลเพื่อการประมวลผล ซึ่งลักษณะของสัญญาณวิดีโอที่รับมาจากกล้องวิดีโอจะขึ้นอยู่กับสายสัญญาณที่เชื่อมต่อระหว่างกล้องกับตัวประมวลผล จากนั้นข้อมูลสีจะถูกเก็บในรูปแบบค่าสีโดยเป็นผลรวมขององค์ประกอบฐานสี (Color Basis) เข้าเป็นสีต่างๆตามภาพที่รับได้ในหนึ่งจุดภาพ (Pixel) ดังนั้นจึงสามารถแทนค่าสีด้วยเวกเตอร์ค่าสีในปฏิภูมิสี (Color Space) แบบต่าง ๆ ได้ ซึ่งโดยทั่วไปในงานประมวลผลภาพดิจิทัลมักนิยมใช้ปฏิภูมิสีแบบ RGB YCbCr รวมทั้งแบบ HSI แต่จะกล่าวถึงในรายละเอียดและความเหมาะสมกับงานเฉพาะปฏิภูมิสีที่ถูกนำไปประยุกต์ใช้ในวิทยานิพนธ์นี้เท่านั้น คือ ปฏิภูมิสี RGB และปฏิภูมิสี YCbCr ว่ามีความแตกต่างกันอย่างไร ต่อไป

##### 2.1.1 ปฏิภูมิสีแบบ RGB (RGB Color Space)

ปฏิภูมิสีแบบ RGB เป็นปฏิภูมิสีที่ถูกนำมาใช้มากที่สุดในงานกราฟิกคอมพิวเตอร์และการประมวลผลภาพดิจิทัล เนื่องจากเป็นองค์ประกอบสีหลักในอุปกรณ์แสดงผล (Display) ทั่วๆ ไปที่อาศัยการแผ่รังสีในย่านคลื่นแสงที่มองเห็นอันประกอบด้วยฐานสี สีแดง เขียว และน้ำเงิน ดังนั้นจึงเป็นการง่ายในการจัดการและออกแบบระบบเพื่อทำงานกับปฏิภูมิสี RGB ดังสังเกตได้จากการแปลงปฏิภูมิสีต่างๆ มัก แปลงโดยเริ่มต้นจากปฏิภูมิสี RGB



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

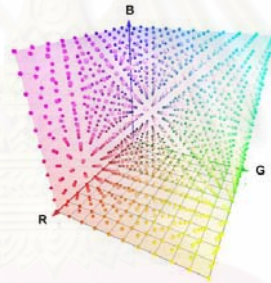
$$\mathbf{X}_{i,j}[n] = (X_{i,j}^R[n], X_{i,j}^G[n], X_{i,j}^B[n]) \quad (2.1)$$

โดยที่  $X_{i,j}^R[n]$  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 แถบสีไม่ว่าจะใช้ข้อมูลในการประมวลผลจริงก็แถบสี เพราะฉะนั้นปริภูมิสีแบบ RGB จึงไม่ถือเป็นปริภูมิที่มีประสิทธิภาพมากที่สุด

### 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) แต่เดิมสัญญาณวิดีโอโทรทัศน์จะเป็นสัญญาณภาพขาว-ดำ เมื่อผู้ให้บริการเปลี่ยนมาใช้ระบบส่งสัญญาณภาพสี ปริภูมิสีแบบ YUV ได้เข้ามามีบทบาทในแง่ของการแยกองค์ประกอบความสว่าง (Y) ซึ่งใช้ในสัญญาณภาพขาว-ดำออกจากองค์ประกอบทางสี (U และ V) ทำให้สามารถส่งสัญญาณภาพรวมสี โดยที่ไม่ต้องทำการปรับเปลี่ยนระบบรับภาพขาวดำที่มีอยู่

สมการการแปลงพื้นฐานระหว่างปริภูมิ RGB และ YUV อาศัยการแปลงเพื่อหาองค์ประกอบความสว่าง (Luminance) Y ณ จุดภาพที่พิจารณา ก่อน 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

$$\mathbf{w}_{i,j}[n] = \begin{bmatrix} W_{i,j}^Y[n] \\ W_{i,j}^U[n] \\ W_{i,j}^V[n] \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} X_{i,j}^R \\ X_{i,j}^G \\ X_{i,j}^B \end{bmatrix} \quad (2.2)$$

และ

$$\mathbf{x}_{i,j}[n] = \begin{bmatrix} X_{i,j}^R[n] \\ X_{i,j}^G[n] \\ X_{i,j}^B[n] \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.140 \\ 1 & -0.395 & -0.581 \\ 1 & 0.032 & 0 \end{bmatrix} \begin{bmatrix} W_{i,j}^Y \\ W_{i,j}^U \\ W_{i,j}^V \end{bmatrix} \quad (2.3)$$

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 เนื่องจากองค์ประกอบสี



อาจมีค่าเป็นลบรวมทั้งขนาดของหน่วยความจำที่ใช้เก็บไม่เท่ากัน จึงไม่เหมาะสมนำมาใช้ในระบบภาพดิจิทัล ดังนั้นจึงจำเป็นต้องมีการพัฒนาปริภูมิสี YUV ใช้งานได้ดีขึ้นเพื่อระบบสัญญาณวิดีโอที่สีดิจิทัลอันเป็นจุดกำเนิดของปริภูมิสี YCbCr

ปริภูมิสี YCbCr ถูกพัฒนาขึ้นตามมาตรฐาน ITU-R BT.601 เพื่อใช้ในการกำหนดรูปแบบข้อมูลวิดีโอที่สีดิจิทัล ซึ่งปริภูมิ YCbCr มีรูปแบบที่ถูกพัฒนาต่อจากรูปแบบการเก็บข้อมูลสีแบบ YUV เป็นการขยายแถบ (Band Expansion) และเลื่อนช่วงข้อมูล (Offset) ของปริภูมิ YUV เดิม (สมการที่ (2.2)) ให้อยู่ในช่วงองค์ประกอบข้อมูลละ 8 บิต (มีค่าระหว่าง 0 ถึง 255) และมีค่าเป็นบวก

โดยวิทยานิพนธ์นี้พิจารณาใช้การแปลงเวกเตอร์ค่าสีระหว่างปริภูมิ RGB  $\mathbf{X}_{i,j}[n]$  และเวกเตอร์ค่าสีในปริภูมิ YCbCr  $\mathbf{Q}_{i,j}[n]$  ตามมาตรฐาน ITU-R BT.709 โดยมีสมการการแปลงดังนี้

$$\mathbf{Q}_{i,j}[n] = \begin{bmatrix} Q_{i,j}^Y[n] \\ Q_{i,j}^{Cb}[n] \\ Q_{i,j}^{Cr}[n] \end{bmatrix} = \begin{bmatrix} 0.183 & 0.614 & 0.062 \\ -0.101 & -0.338 & 0.439 \\ 0.439 & -0.399 & -0.040 \end{bmatrix} \begin{bmatrix} X_{i,j}^R \\ X_{i,j}^G \\ X_{i,j}^B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (2.4)$$

และ

$$\mathbf{X}_{i,j}[n] = \begin{bmatrix} X_{i,j}^R[n] \\ X_{i,j}^G[n] \\ X_{i,j}^B[n] \end{bmatrix} = \begin{bmatrix} 1.164 & 0 & 1.793 \\ 1.164 & -0.534 & -0.213 \\ 1.164 & 2.115 & 0 \end{bmatrix} \left( \begin{bmatrix} Q_{i,j}^Y \\ Q_{i,j}^{Cb} \\ Q_{i,j}^{Cr} \end{bmatrix} - \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \right) \quad (2.5)$$

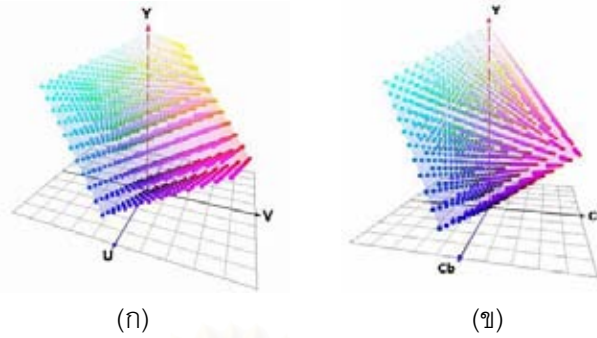
โดยที่  $Q_{i,j}^Y[n]$  เป็นค่าองค์ประกอบความสว่าง และ  $16 \leq Q_{i,j}^Y[n] \leq 235$

$Q_{i,j}^{Cb}[n]$  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)



รูปที่ 2.2 (ก) ปริภูมิสีแบบ YUV (ข) ปริภูมิสีแบบ YCbCr

ข้อดีในการใช้ปริภูมิสีแบบ YCbCr คือ องค์ประกอบในปริภูมิ มีการแยกองค์ประกอบเชิงความสว่างออกจากองค์ประกอบเชิงสีออกอย่างชัดเจน เป็นการง่ายที่จะตรวจพบความเปลี่ยนแปลงของแต่ละองค์ประกอบอย่างเป็นอิสระต่อกัน

## 2.2 การประมวลผลภาพเชิงสัณฐาน (Morphological Image Processing) และการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน (Connected component analysis)

เป็นกระบวนการสำคัญหนึ่งในการประมวลผลภาพ เพื่อวิเคราะห์ความหมายของภาพฐานสองที่ได้จากการประมวลผลในรูปแบบต่าง ๆ เช่น การตรวจหาขอบ การตรวจหาสีผิว เป็นต้น โดยกระบวนการทั่วไปจะเป็นการประมวลผลทางคณิตศาสตร์เชิงสัณฐานระหว่างภาพฐานสอง กับ โครงสร้างวัตถุในรูปแบบทางเรขาคณิตอย่างง่าย (วงกลม สี่เหลี่ยม) ซึ่งจะเรียกโครงสร้างนี้ว่า โครงสร้างส่วนย่อย (Structuring element) ผลลัพธ์เชิงความหมายของบริเวณวัตถุที่สนใจที่ได้จากการวิเคราะห์ในส่วนนี้ เช่น ขนาด (size) รูปร่าง (shape) ลักษณะการเชื่อมต่อ (connectivity) ความเรียบของวัตถุ (smoothness) และ การเรียงตัว (orientation) โดยการประมวลผลเชิงสัณฐานจะมีหลายรูปแบบ แต่ในวิทยานิพนธ์นี้สนใจการประมวลผลเชิงสัณฐานที่เกี่ยวข้องกับการลดสัญญาณรบกวนที่เกิดขึ้นในภาพฐานสองของผลการตรวจหา คือ กระบวนการไถเลขนและอีโรชัน ซึ่งมีรายละเอียดในแต่ละวิธีการดังนี้

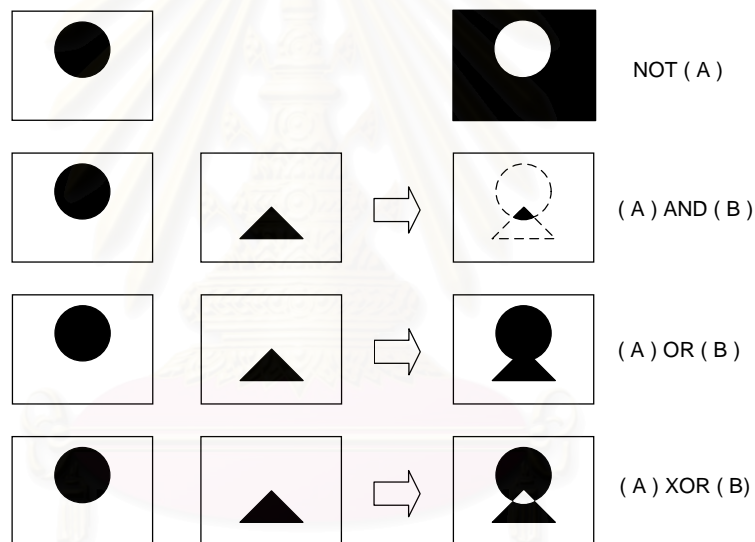
### 2.2.1 การดำเนินการตรรกะกับรูปภาพฐานสอง

ในเบื้องต้นจะอธิบายการดำเนินการตรรกะกับรูปภาพฐานสอง ซึ่งเป็นขั้นตอนดำเนินการก่อนกระบวนการไถเลขนและอีโรชัน การดำเนินการตรรกะเป็นพื้นฐานทางการประมวลผลที่สำคัญและสามารถนำไปประยุกต์ใช้หรือเป็นพื้นฐานแนวความคิดเกี่ยวกับขั้นตอนวิธีได้อย่างมีประสิทธิภาพ หลักของการดำเนินการตรรกะที่นำมาใช้ในการประมวลผลภาพ คือ AND, OR, XOR และ NOT ซึ่งจะกระทำบนฐานหลักของจุดภาพระหว่าง 2 รูปภาพหรือมากกว่านั้น (ยกเว้น

NOT ที่เป็นตัวดำเนินการของรูปภาพเพียงรูปเดียว) สำหรับการประมวลผลที่ซับซ้อนกว่านี้ เมื่อพิจารณาอย่างถ่องแท้แล้วจะพบว่าเกิดจากตัวดำเนินการหลัก 4 ตัวนี้เรียงลำดับอยู่ในขั้นตอนเช่นกัน

ตารางที่ 2.1 การดำเนินการตรรกะพื้นฐาน 4 ตัวดำเนินการ

A	B	A AND B	A OR B	A XOR B	NOT(A)
0	0	0	0	0	1
0	1	0	1	1	1
1	0	0	1	1	0
1	1	1	1	0	0



รูปที่ 2.3 ตัวอย่างของการดำเนินการตรรกะระหว่างรูปภาพฐานสอง

### 2.2.2 การทำไดเลชัน (Dilation) และ การทำอีโรชัน (Erosion)

รูปภาพฐานสอง (Binary image) คือ กลุ่มของจุดภาพสีดำและขาว การทำไดเลชันและอีโรชัน ถือเป็นตัวดำเนินการการเชิงสัญญาณเบื้องต้น และเป็นพื้นฐานในการนำไปใช้กับกระบวนการที่มีความซับซ้อนมากขึ้น ได้แก่ ตัวดำเนินการแบบเปิด (Opening operator) ตัวดำเนินการแบบปิด (Closing operator) และการแยกรูปร่าง (Shape decomposition)

### การทำไดเลชัน

นิยาม คือ เซตของจุด  $p$  ที่เกิดจาก  $x+d$  โดยที่  $x$  เป็นสมาชิกของ  $X$  และ  $d$  เป็นสมาชิกของ  $D$  ดังสมการที่ (2.6)

$$X \oplus D = \{p \in \mathcal{E}^2 : p = x + d, (x \in X) \& (d \in D)\} \quad (2.6)$$

โดยที่  $X$  คือ รูปภาพที่ทำการไดเลชัน

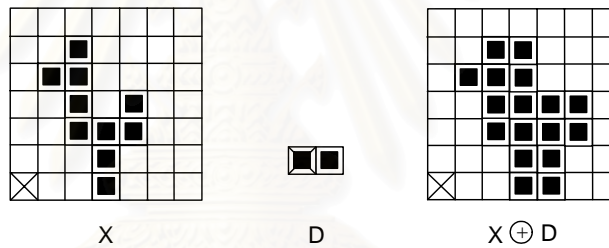
$D$  คือ โครงสร้างส่วนย่อยในการทำไดเลชัน

**ตัวอย่าง** การทำไดเลชัน รูปที่ 2.4

กำหนดให้  $X = \{(0,3), (1,3), (2,2), (2,3), (2,4), (3,2), (3,4), (4,1), (4,2), (5,2)\}$

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

$$X \oplus D = \{(0,3), (0,4), (1,3), (1,4), (2,2), (2,3), (2,4), (2,5), (3,2), (3,3), (3,4), (3,5), (4,1), (4,2), (4,3), (5,2), (5,3)\}$$



รูปที่ 2.4 ตัวอย่างการทำไดเลชัน

### การทำอีโรชัน

นิยาม คือ เซตของจุด  $p$  ทุกจุดจากรูปภาพทดสอบ ผลลัพธ์ของการทำอีโรชันคือ จุด  $p$  เหล่านั้นที่ทำให้  $p+e$  เป็นสมาชิกของ  $X$  สำหรับทุกๆ  $e$  ที่เป็นสมาชิกของ  $E$  แสดงในสมการที่ (2.7)

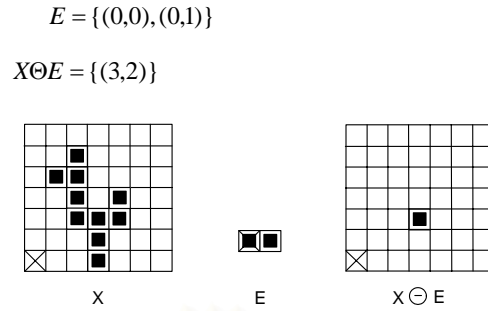
$$X \ominus E = \{p \in \mathcal{E}^2 : p + e \in X \text{ for every } e \in E\} \quad (2.7)$$

โดยที่  $X$  คือ รูปภาพที่ทำการอีโรชัน

$E$  คือ โครงสร้างส่วนย่อยในการทำอีโรชัน

**ตัวอย่าง** การทำอีโรชัน รูปที่ 2.5

กำหนดให้  $X = \{(0,3), (1,3), (2,2), (2,3), (2,4), (3,2), (3,4), (4,1), (4,2), (5,2)\}$



รูปที่ 2.5 การทำอีโรชัน

### 2.2.3 ตัวดำเนินการแบบเปิดและปิด

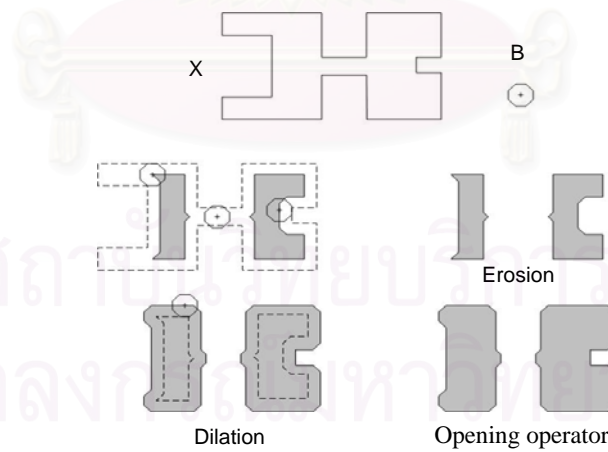
การทำอีโรชันและไดเลชันไม่ถือเป็นการแปลงผกผัน (Inverse transformation) ระหว่างกัน ดังนั้น รูปภาพที่นำมาผ่านการอีโรชันและไดเลชัน ไม่จำเป็นต้องเท่ากับรูปภาพดั้งเดิม แต่จะได้ผลลัพธ์ซึ่งมีรายละเอียดซับซ้อนน้อยกว่ารูปภาพดั้งเดิมแทน

กระบวนการอีโรชันแล้วตามด้วยไดเลชัน เรียกว่า ตัวดำเนินการแบบเปิด

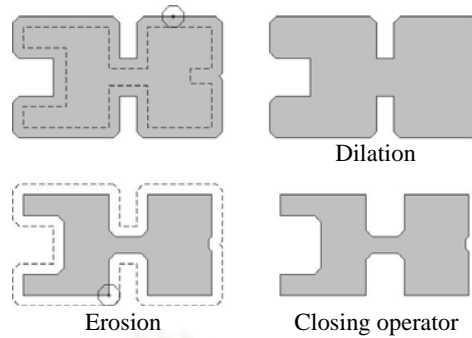
$$X \circ B = (X \ominus B) \oplus B \tag{2.8}$$

กระบวนการไดเลชันแล้วตามด้วยอีโรชัน เรียกว่า ตัวดำเนินการแบบปิด

$$X \bullet B = (X \oplus B) \ominus B \tag{2.9}$$



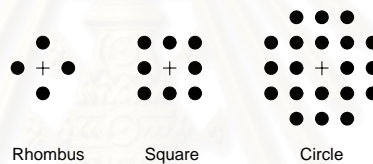
รูปที่ 2.6 ตัวดำเนินการแบบเปิด [34]



รูปที่ 2.7 ตัวดำเนินการแบบปิด [34]

### 2.2.4 รูปแบบโครงสร้างส่วนย่อย

รูปแบบโครงสร้างส่วนย่อยในการทำโดเลชันและการทำอีโรชันจะมีรูปแบบต่าง ๆ ดังตัวอย่างรูปที่ 2.8 การเลือกนำไปใช้งานจะขึ้นอยู่กับลักษณะงานประยุกต์ว่าต้องการผลที่ได้จากการประมวลผลที่ได้ในแบบใด



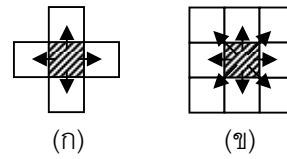
รูปที่ 2.8 ตัวอย่างโครงสร้างส่วนย่อยในการทำโดเลชันและการทำอีโรชัน

### 2.2.5 การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน

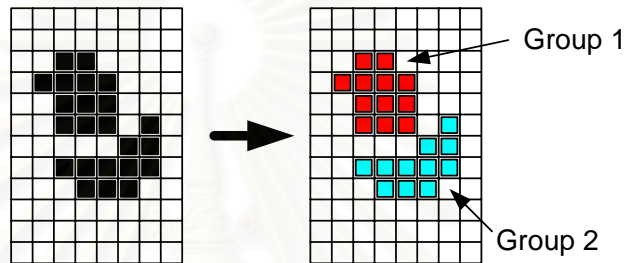
หลังจากที่ได้รูปภาพฐานสองที่ผ่านกระบวนการประมวลผลภาพเชิงสัญญาณแล้ว ก็จะได้ถึงขั้นตอนการวิเคราะห์ผลเพื่อใช้บ่งบอกความหมายจากผลการประมวลผลภาพ เช่น ต้องการพิจารณาว่าบริเวณใดในรูปภาพฐานสองที่ได้จากกระบวนการตรวจจับสีผิวเป็นกลุ่มบริเวณใบหน้า เป็นต้น กระบวนการเบื้องต้น คือ การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน เพื่อจัดจำแนกบริเวณเป็นกลุ่ม ๆ ทำให้ทราบถึงลักษณะรูปร่างของส่วนประกอบ โดยจะสามารถพิจารณาได้ว่ามีจุดภาพกลุ่มใดบ้างที่เป็นวัตถุเดียวกันจากความสัมพันธ์ของรหัส ซึ่งเป็นตัวกำหนดรูปแบบการเชื่อมต่อกันของจุดภาพที่อยู่ในบริเวณใกล้เคียงกัน การกำหนดรหัสจะมีทั้งแบบพิจารณาจุดภาพข้างเคียง 4 จุดเชื่อมต่อ และ แบบ 8 จุดเชื่อมต่อ ดังรูปที่ 2.9 ข้อแตกต่างระหว่าง 2 ลักษณะนี้สามารถแสดงได้ดังรูปที่ 2.10 และ 2.11 จะเห็นว่าข้อแตกต่างระหว่างรหัสทั้งสองแบบนี้ คือ ทิศทางตำแหน่งจุดภาพใกล้เคียงที่ใช้ในการวิเคราะห์ ในแบบ 8 จุดเชื่อมต่อจะมีการพิจารณาในแนวเฉียงเพิ่มเข้ามา ซึ่งส่งผลโดยตรงต่อขนาดและจำนวนของกลุ่มที่วิเคราะห์ได้ จากรูปที่ 2.10 วิเคราะห์ว่ารูปภาพฐานสองนี้มีผลแบ่งเป็น 2 กลุ่ม แต่เมื่อวิเคราะห์แบบ 8 จุดเชื่อมต่อ จะเหลือแค่บริเวณกลุ่มเดียว ดังนั้นขึ้นอยู่กับความต้องการของนักพัฒนาและนักวิจัยที่จะเลือกใช้ให้เหมาะกับรูปแบบงานด้านการ



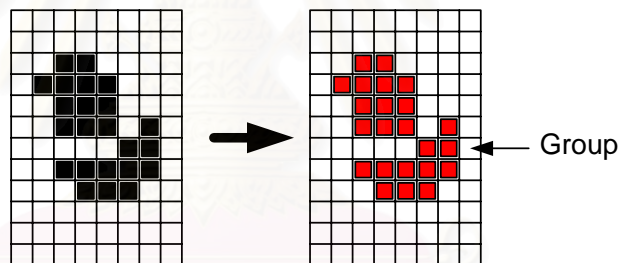
ประมวลผลภาพมากที่สุด ว่าต้องการความสัมพันธ์ของตำแหน่งจุดรอบข้างที่ใกล้เคียงกันมากน้อยเท่าไร



รูปที่ 2.9 การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกันระหว่างจุดภาพใกล้เคียง (ก) แบบ 4 จุดเชื่อมต่อ (ข) แบบ 8 จุดเชื่อมต่อ



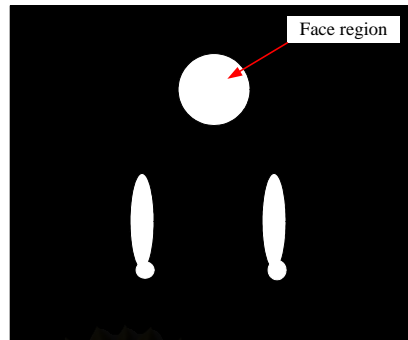
รูปที่ 2.10 การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกันด้วยรหัสแบบ 4 จุดเชื่อมต่อ



รูปที่ 2.11 การวิเคราะห์ส่วนประกอบที่เชื่อมต่อกันด้วยรหัสแบบ 8 จุดเชื่อมต่อ

### 2.2.5.1 การวิเคราะห์รูปร่างของกลุ่มบริเวณที่เชื่อมต่อ

รูปร่างเป็นลักษณะที่สำคัญเพื่อจำแนกความหมายของกลุ่มบริเวณที่เชื่อมต่อ ที่ได้จากการตรวจหาว่าเป็นกลุ่มบริเวณที่สนใจหรือไม่ ดังรูปตัวอย่างที่ 2.12 ลักษณะที่เลือกนำมาพิจารณาในที่นี้ คือ ลักษณะความแน่น (Compactness) ลักษณะโซลิดิตี (Solidity) และ ลักษณะทิศทาง (Orientation) โดยพารามิเตอร์ที่นำมาใช้ในการพิจารณาจะประกอบด้วย พื้นที่ (area) ( $A$ ) เส้นรอบรูป (perimeter) ( $P$ ) ความยาวจากจุดภาพวัตถุในตำแหน่งขอบล่างถึงขอบบน (lower-upper length) ในแนวแกนอน ( $D_x$ ) และแกนตั้ง ( $D_y$ )



รูปที่ 2.12 ตัวอย่างภาพฐานสองที่นำมาวิเคราะห์หาบริเวณกลุ่มใบหน้า

**ลักษณะความแน่น** เป็นค่าอัตราส่วนระหว่างขนาดพื้นที่จริงของกลุ่มบริเวณที่สนใจและเส้นรอบรูปกำลังสอง ดังสมการที่ (2.11)

$$Compactness = \frac{A}{P^2} \quad (2.11)$$

สมการดังกล่าวบ่งบอกถึงความหนาแน่นของจุดภาพที่อยู่ภายในขอบเขตทั้งหมดของวัตถุ ซึ่งรูปร่างวัตถุแต่ละแบบก็จะมีค่านี้แตกต่างกัน ดังนั้นค่าขีดเริ่มเปลี่ยนที่ใช้ในการจำแนกต้องขึ้นอยู่กับลักษณะรูปร่างที่ต้องการ

**ลักษณะโซลิดิตี** เป็นค่าอัตราส่วนระหว่างขนาดพื้นที่จริงของกลุ่มบริเวณที่สนใจและพื้นที่สี่เหลี่ยมที่เกิดจาก  $D_x \times D_y$  ดังสมการที่ (2.12)

$$Solidity = \frac{A}{D_x D_y} \quad ; \quad 0 < Solidity \leq 1 \quad (2.12)$$

สมการดังกล่าวบ่งบอกถึงค่าประสิทธิภาพการใช้พื้นที่ของกลุ่มบริเวณที่สนใจ กล่าวคือ ถ้าพิจารณาวัตถุรูปร่างวงกลมที่มีรัศมีเท่ากัน แต่รูปแรกเป็นทรงกลมตันและรูปที่สองเป็นทรงกลมกลวง ค่าโซลิดิตีของรูปแรกจะมีค่าเข้าใกล้ 1 มากกว่ารูปที่สอง

**ลักษณะทิศทาง** เป็นค่าอัตราส่วนระหว่างความกว้างในแนวแกนตั้งกับความกว้างในแนวแกนนอน ดังสมการที่ (2.13)

$$Orientation = \frac{D_y}{D_x} \quad ; \quad \frac{1}{W} \leq Orientation \leq H \quad (2.13)$$

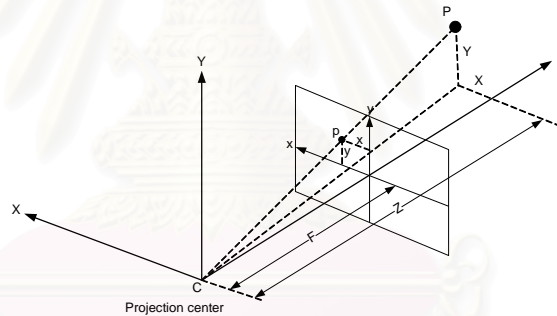
โดยที่  $W$  และ  $H$  เป็นความกว้างและสูงของขนาดภาพที่ใช้ ตามลำดับ และกลุ่มของบริเวณที่สนใจ ต้องมีขนาดความกว้างอย่างน้อย 1 จุดภาพ

สมการดังกล่าวบ่งบอกถึงลักษณะของรูปร่างในเชิงทิศทาง ซึ่งจำแนกตามความกว้างยาวของกลุ่มบริเวณเชื่อมต่อ เช่น รูปร่างวงกลม ค่านี้จะมีค่าเข้าใกล้ 1 ส่วนรูปร่างทรงกระบอก จะได้ค่านี้น้อยมากเนื่องจากความต่างระหว่างความกว้างและความยาว

## 2.3 แบบจำลองกล้อง (Camera model) และแบบจำลองการเคลื่อนที่ (Motion model)

### 2.3.1 แบบจำลองกล้อง [33]

ในวิทยานิพนธ์นี้ต้องมีการทำงานร่วมกับกล้องวีดิทัศน์โดยมีการควบคุมกล้องวีดิทัศน์ให้เคลื่อนที่ไปยังตำแหน่งที่ต้องการ ดังนั้นทฤษฎีการทำงานเกี่ยวกับแบบจำลองกล้องจึงมีความสำคัญในด้านการมโนภาพ (imagine) เพื่อออกแบบการทำงานของระบบ โดยแบบจำลองกล้องเป็นการจำลองรูปแบบการสร้างภาพที่ปรากฏอยู่บนจอภาพในพิกัดภาพ (image coordinate) 2 มิติ จากวัตถุในฉากจริง (real scene) ในพิกัด 3 มิติ มีแบบจำลองหลายรูปแบบที่ใช้บรรยายการสร้างภาพ เช่น แบบจำลองกล้องรูเข็ม (pinhole camera model) แบบจำลองกล้อง CAHV (CAHV camera model) เป็นต้น แต่ในที่นี้จะพิจารณาเฉพาะแบบจำลองกล้องรูเข็มเท่านั้น



รูปที่ 2.13 การฉายภาพแบบทัศนมิติ (perspective projection) ของแบบจำลองกล้องรูเข็ม [33]

แบบจำลองกล้องรูเข็ม เป็นแบบจำลองที่นิยมใช้กันอย่างมากมายในการวิเคราะห์การสร้างภาพและงานประยุกต์ใน 3 มิติ แสดงได้ดังรูปที่ 2.13 โดยจะอาศัยพื้นฐานการฉายภาพแบบทัศนมิติ ในการสร้างภาพในพิกัด 2 มิติจากวัตถุในพิกัด 3 มิติ พิจารณาที่จุดพิกัดวัตถุ P (X, Y, Z) ในพิกัด 3 มิติ สามารถฉายลงมายังจุด p (x, y) ในระนาบภาพ (imaging plane) 2 มิติ ได้ดังสมการที่ (2.14) โดยมีจุดตำแหน่ง C เป็นจุดศูนย์กลางในการฉายภาพ (center of projection)

$$x = F \frac{X}{Z}, \quad y = F \frac{Y}{Z} \quad (2.14)$$

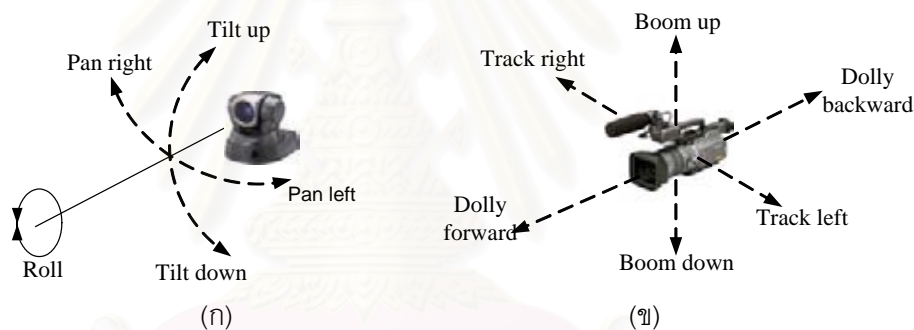
โดยที่ F คือ ความยาวโฟกัส (Focal length)

### 2.3.2 การเคลื่อนที่ของกล้อง (Camera Motion) [33]

การเคลื่อนที่ของกล้องโดยทั่วไปจะแบ่งออกได้เป็น 2 รูปแบบ คือ

1. การเคลื่อนที่ของกล้องที่ตำแหน่งศูนย์กลางการฉายภาพไม่เปลี่ยนแปลง ดังรูปที่ 2.14 (ก) ประกอบด้วย การส่าย (pan) เป็นลักษณะการเคลื่อนที่กล้องโดยยึดแกนตั้ง (vertical axis) เป็นแกนหมุนรอบ การก้มเงย (tilt) เป็นลักษณะการเคลื่อนที่กล้องโดยยึดแกนนอน (horizontal axis) เป็นแกนหมุนรอบ และ การหมุน (roll) เป็นลักษณะการเคลื่อนที่กล้องโดยยึดแกนเชิงแสง (optical axis) เป็นแกนหมุนรอบ

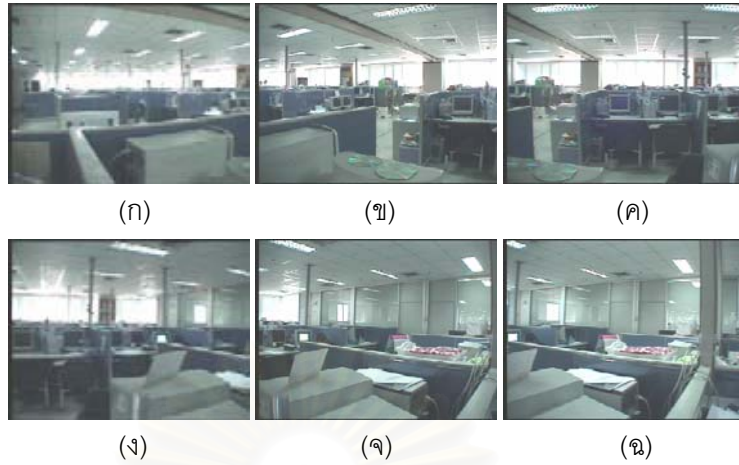
2. การเคลื่อนที่ของกล้องที่ตำแหน่งศูนย์กลางการฉายภาพมีการเปลี่ยนแปลง ตำแหน่ง ดังรูปที่ 2.14 (ข) ประกอบด้วย การติดตาม (track) เป็นลักษณะการเคลื่อนที่กล้องในลักษณะการเลื่อนขนาน (translation) ตามแนวแกนนอนของระนาบภาพ การยก (boom) เป็นลักษณะการเคลื่อนที่กล้องในลักษณะการเลื่อนขนานตามแนวแกนตั้งของระนาบภาพ และ การดอลลี (dolly) เป็นลักษณะการเคลื่อนที่กล้องในลักษณะการเลื่อนขนานตามแนวแกนเชิงแสง



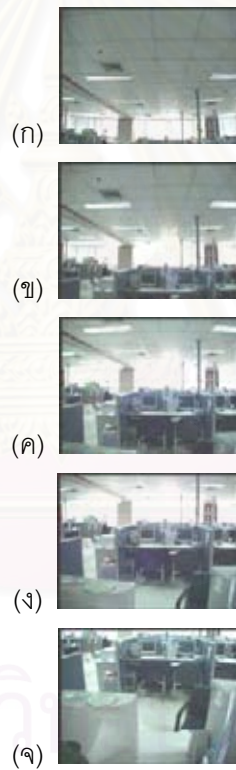
รูปที่ 2.14 การเคลื่อนที่ของกล้อง

ในวิทยานิพนธ์นี้จะประยุกต์ใช้ลักษณะการเคลื่อนที่ดังรูปแบบที่ 1 คือ การส่าย และการก้มเงย รวมทั้งมีการซูม (zoom) ซึ่งเป็นรูปแบบการเปลี่ยนแปลงขนาดวัตถุที่รับได้บนระนาบภาพ โดยเกิดจากการเปลี่ยนแปลงความยาวโฟกัส ตัวอย่างเฟรมภาพที่รับได้จากการเคลื่อนที่ในลักษณะดังกล่าว แสดงได้ดังรูปที่ 2.15 2.16 และ 2.17

จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.15 ลำดับภาพวีดิทัศน์ที่รับได้จากการเปลี่ยนมุมมองของกล้องแบบการถ่าย



รูปที่ 2.16 ลำดับภาพวีดิทัศน์ที่รับได้จากการเปลี่ยนมุมมองของกล้องแบบการก้มเงย



รูปที่ 2.17 ลำดับภาพวีดิทัศน์ที่รับได้จากการเปลี่ยนมุมมองของกล้องแบบการซูม

### 2.3.3 แบบจำลองการเคลื่อนที่ [33]

ในที่นี้จะพิจารณาเฉพาะการเคลื่อนที่ในแนว 2 มิติเท่านั้น ดังนี้

**การส่ายและก้มเงย** เป็นการหมุนกลิ้งรอบแกนตั้ง ( $Y$ ) และแกนนอน ( $X$ ) ตามลำดับ ถ้าให้  $\theta_y$  และ  $\theta_x$  เป็นมุมในการหมุน จะได้พิกัดกล้องใหม่หลังการหมุนที่สัมพันธ์กับพิกัดอันเก่าดังสมการที่ (2.15)

$$X' = [R_x][R_y]X \quad (2.15)$$

โดยที่  $[R_x]$  และ  $[R_y]$  เป็นเมทริกซ์การหมุนที่เกิดจากการส่ายและการก้มเงย ถ้าพิจารณาที่มุมการหมุนที่เล็ก จะได้

$$[R_x][R_y] = \begin{bmatrix} 1 & 0 & \theta_y \\ 0 & 1 & -\theta_x \\ -\theta_y & \theta_x & 1 \end{bmatrix} \quad (2.16)$$

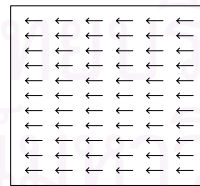
ถ้า  $Y\theta_x \ll Z$  และ  $X\theta_y \ll Z$  แล้ว  $Z' \approx Z$  จากสมการที่ (2.14) จะได้เป็น

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \theta_y F \\ -\theta_x F \end{bmatrix} \quad (2.17)$$

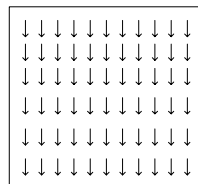
หรือ

$$\begin{bmatrix} d_x(x, y) \\ d_y(x, y) \end{bmatrix} = \begin{bmatrix} \theta_y F \\ -\theta_x F \end{bmatrix} \quad (2.18)$$

สนามการเคลื่อนที่ (motion field) ของการส่ายและการก้มเงยสามารถแสดงได้ดังรูปที่ 2.18 และ 2.19

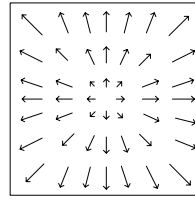


รูปที่ 2.18 สนามการเคลื่อนที่ของการส่ายไปทางซ้าย



รูปที่ 2.19 สนามการเคลื่อนที่ของการก้มเงยลง





รูปที่ 2.20 สนามการเคลื่อนที่ของการซูม

การซูม ให้  $F$  และ  $F'$  เป็นความยาวโฟกัสก่อนและหลังการซูม จะได้สมการการเคลื่อนที่ที่เกิดจากการซูมดังสมการที่ (2.19)

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \rho \cdot x \\ \rho \cdot y \end{bmatrix} \quad (2.19)$$

โดยที่  $\rho = \frac{F'}{F}$  คือ ตัวประกอบการซูม (zoom factor) ดังนั้นได้สมการสนามการเคลื่อนที่ดังสมการที่ (2.20) และแสดงได้ดังรูปที่ 2.20

$$\begin{bmatrix} d_x(x, y) \\ d_y(x, y) \end{bmatrix} = \begin{bmatrix} (1-\rho)x \\ (1-\rho)y \end{bmatrix} \quad (2.20)$$

## 2.4 การตรวจหาและติดตามใบหน้า (Face Detection and Tracking)

การตรวจหา (Detection) เป็นกระบวนการทำงานขั้นแรกที่สำคัญของระบบทางด้านการประมวลผลภาพและวีดิทัศน์ เป้าหมายของกระบวนการนี้ คือ การตรวจหาตำแหน่งหรือบริเวณของวัตถุเป้าหมายภายในเฟรมภาพ สำหรับนำไปวิเคราะห์หรือประมวลผลภาพในขั้นตอนอื่นต่อไป ขั้นตอนวิธีที่ถูกนำมาใช้ในการตรวจหานั้นขึ้นอยู่กับรูปแบบของวัตถุที่ต้องการตรวจหา เช่น ในการตรวจหาใบหน้าบุคคล ลักษณะที่นิยมนำมาใช้ ได้แก่ ลักษณะสี ถ้าเป็นการตรวจหารถยนต์ ลักษณะการเคลื่อนที่ จะเป็นลักษณะหลักที่ถูกเลือกนำมาใช้ ในส่วนนี้จะพิจารณาการตรวจหาใน 4 ลักษณะ คือ การตรวจหาการเคลื่อนที่ การตรวจหาด้วยวิธีกระบวนการลบฉากหลัง การตรวจหาขอบ และการตรวจหาด้วยสี

### 2.4.1 การตรวจหาการเคลื่อนที่ (Motion Detection)

ลักษณะการเคลื่อนที่เป็นลักษณะที่ถูกนำมาใช้ในการตรวจหาวัตถุที่มีการเคลื่อนที่ โดยจะอาศัยการคำนวณการเปลี่ยนแปลงค่าความเข้มสี (color intensity) ของจุดภาพแต่ละจุดภาพในเฟรมภาพ ว่ามีการเปลี่ยนแปลงเกินกว่าค่าขีดเริ่มเปลี่ยน (threshold) ที่ได้กำหนดไว้หรือไม่ ภาพรวมของหลักการจะคล้ายกับกระบวนการตรวจหาด้วยขั้นตอนวิธีลบฉากหลัง (ดังจะกล่าวในหัวข้อถัดไป) แต่ไม่จำเป็นต้องมีการเก็บภาพฉากหลังไว้ เนื่องจากจะใช้การคำนวณหา

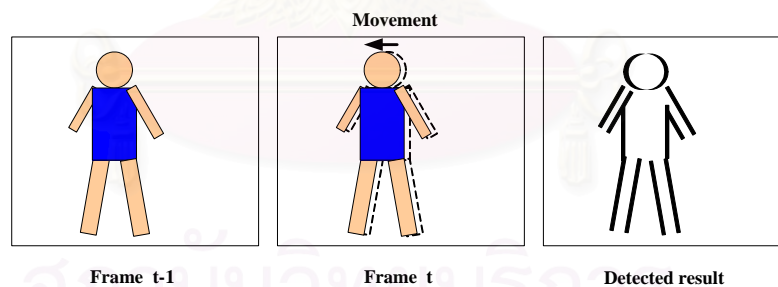
ความแตกต่างระหว่างเฟรมภาพ ที่เรียกว่า ขั้นตอนวิธีการหาผลต่างระหว่างเฟรม (frame differencing)

กำหนดให้  $I_t$  และ  $I_{t-1}$  เป็นเฟรมภาพสองเฟรมที่ต่อเนื่องกัน ดังนั้นผลของการตรวจหาจะแสดงได้ดังสมการที่ (2.21)

$$M_{motion}^{i,j} = \begin{cases} 1 & , |I_t^{i,j} - I_{t-1}^{i,j}| \geq \tau_{motion} \\ 0 & , |I_t^{i,j} - I_{t-1}^{i,j}| < \tau_{motion} \end{cases} \quad (2.21)$$

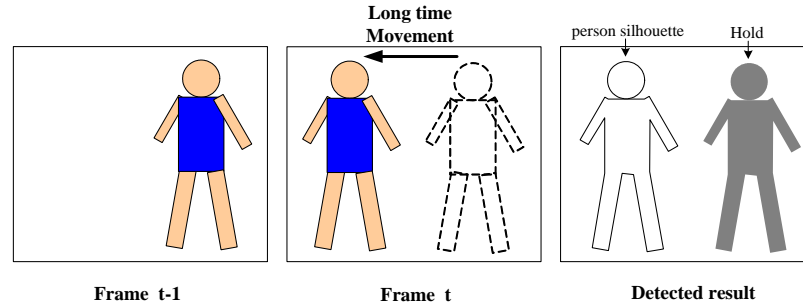
เมื่อ  $M_{motion}^{i,j}$  คือผลการตรวจหาการเคลื่อนที่ในรูปของภาพฐานสอง ณ ตำแหน่งจุดภาพที่  $(i, j)$  โดยมีค่าเป็น 1 เมื่อตีความว่ามีการเปลี่ยนแปลงความเข้มสีอันเนื่องจากการเคลื่อนที่ของวัตถุ เป็น 0 เมื่อตีความว่าไม่มีการเปลี่ยนแปลงความเข้มสีแสดงว่าไม่มีการเคลื่อนที่เกิดขึ้น  $I_t^{i,j}$  และ  $I_{t-1}^{i,j}$  เป็นเฟรมภาพสองเฟรมที่ต่อเนื่องกัน ณ ตำแหน่งจุดภาพที่  $(i, j)$  และ  $\tau_{motion}$  คือ ค่าขีดเริ่มเปลี่ยนของการตัดสินใจการเปลี่ยนแปลงค่าความเข้มของจุดภาพทั้งเฟรม

โดยช่วงระยะห่างระหว่างเฟรมสองเฟรมที่นำมาใช้คำนวณหาการเปลี่ยนแปลงนั้น จะขึ้นอยู่กับอัตราเฟรม (frame rate) ของสัญญาณวิดีโอที่นำมาใช้ในการประมวลผล ซึ่งโดยส่วนใหญ่ก็จะเลือกใช้เฟรมที่ติดกัน มาใช้ในการตรวจหาการเคลื่อนที่ ดังรูปที่ 2.21 จะเห็นว่าผลจากการตรวจหาการเคลื่อนที่จะได้เป็นบริเวณขอบที่วัตถุนั้นเคลื่อนที่ไป ซึ่งจะแตกต่างกับการใช้กระบวนการลบฉากหลังที่ผลการตรวจหาจะได้เป็นบริเวณทั้งหมดของวัตถุ



รูปที่ 2.21 ตัวอย่างผลการตรวจหาการเคลื่อนที่

ในกรณีที่เลือกช่วงห่างระหว่างเฟรมสองเฟรมที่นำมาใช้ที่ห่างกันมาก จะทำให้เกิดปัญหาขึ้นดังรูปที่ 2.22 กล่าวคือ จะไม่ได้ผลการตรวจหาที่เป็นบริเวณขอบการเคลื่อนที่ แต่จะได้เป็นบริเวณสองบริเวณ หนึ่งคือบริเวณวัตถุที่ต้องการจะตรวจหา และอีกส่วนจะได้เป็นบริเวณฉากหลังที่มีรูปร่างคล้ายวัตถุ หรือเรียกว่า หลุม (Hole) ซึ่งจะส่งผลโดยตรงต่อกระบวนการวิเคราะห์ส่วนประกอบที่เชื่อมต่อกัน



รูปที่ 2.22 ปัญหาของการตรวจหาการเคลื่อนที่

ข้อดีของขั้นตอนวิธีการตรวจหาการเคลื่อนที่ คือ ไม่ต้องอ้างอิงกับฉากหลังในการตรวจหา ดังนั้นจึงสามารถประยุกต์ใช้ได้กับลักษณะงานประยุกต์ที่ฉากหลังจะต้องมีการเปลี่ยนแปลง แต่ก็ต้องเพิ่มความยุ่งยากในส่วนของทฤษฎีการวิเคราะห์บริเวณวัตถุที่สนใจ



รูปที่ 2.23 ตัวอย่างการตรวจหาการเคลื่อนที่

#### 2.4.2 กระบวนการลบฉากหลังพื้นฐาน (Basic background Subtraction)

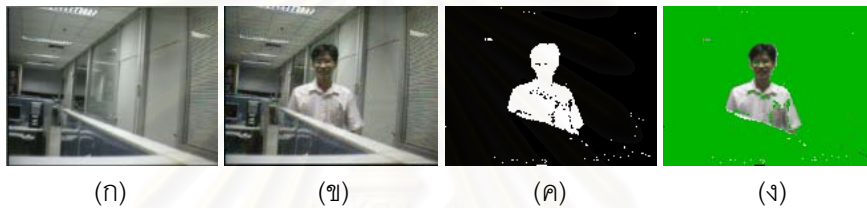
เป็นขั้นตอนวิธีในการแยกฉากหน้า (foreground) ออกจากฉากหลัง ซึ่งเป็นขั้นตอนวิธีที่นิยมนำไปใช้กับงานประยุกต์มากมาย เนื่องจากง่ายต่อการคำนวณและออกแบบ โดยหลักการเบื้องต้น คือ จะเก็บภาพฉากหลัง ณ ตำแหน่งมุมมองภาพที่ต้องการไว้เป็นภาพฉากหลังอ้างอิง ก่อนที่จะประมวลผล จากนั้นจะตรวจหาฉากหน้าโดยลบภาพรับเข้ากับภาพฉากหลังอ้างอิง

กำหนดให้  $I_{in}$  เป็นภาพรับเข้า และ  $B_{ref}$  เป็นภาพฉากหลังอ้างอิงที่ถูกจัดเก็บไว้ ดังนั้นผลของการตรวจหาด้วยวิธีลบฉากหลังจะแสดงได้ดังสมการที่ (2.22)

$$M_{bgs}^{i,j} = \begin{cases} 1 & , |I_{in}^{i,j} - B_{ref}^{i,j}| \geq \tau_{bgs} \\ 0 & , |I_{in}^{i,j} - B_{ref}^{i,j}| < \tau_{bgs} \end{cases} \quad (2.22)$$

เมื่อ  $M_{bgs}^{i,j}$  คือผลการลบฉากหลัง ณ ตำแหน่งจุดภาพ  $(i, j)$  แสดงในรูปภาพฐานสอง โดยเป็น 1 เมื่อตีความว่าจุดภาพนั้นเป็นฉากหน้าและเป็น 0 เมื่อตีความว่าจุดภาพนั้นเป็นฉากหลัง  $I_{in}^{i,j}$  เป็นภาพรับเข้า ณ ตำแหน่งจุดภาพ  $(i, j)$  และ  $B_{ref}^{i,j}$  เป็นเฟรมภาพฉากหลังอ้างอิง ณ ตำแหน่งจุดภาพที่  $(i, j)$  และ  $\tau_{bgs}$  คือ ค่าขีดเริ่มเปลี่ยนของการตัดสินใจว่าจุดภาพนั้นเป็นฉากหลังหรือฉากหน้า

รูปที่ 2.24 แสดงตัวอย่างรูปฉากหลังอ้างอิง และผลที่ได้จากการลบฉากหลัง เนื่องจากว่ากระบวนการลบฉากหลังจะขึ้นกับเฟรมภาพฉากหลังอ้างอิงที่เก็บไว้ ดังนั้นปัญหาของฉากหลังก็เป็นส่วนหนึ่งที่ส่งผลกระทบต่อผลของการตรวจหา เช่น ปัญหาเรื่องแสง ปัญหาเรื่องสิ่งแวดล้อมที่มีการเคลื่อนที่ตลอดเวลา และในกรณีที่มีการเก็บภาพฉากหลังไว้ล่วงหน้านานเกินไป ทำให้สิ่งแวดล้อม ณ ขณะที่ประมวลผลมีการเปลี่ยนแปลง ดังนั้นฉากหลังจึงมีความสำคัญมากในการประมวลผล ถ้ากล้องวิดีโอที่เข้ารับภาพในการประมวลผลมีการเปลี่ยนแปลงมุมมอง การรับภาพ จะต้องมีการวนการที่ต้องปรับภาพฉากหลังอ้างอิงให้เข้ากับตำแหน่งมุมมองกล้องนั้นด้วย ซึ่งเป็นข้อจำกัดอย่างหนึ่งสำหรับการประยุกต์ใช้กับกล้องวิดีโอที่ ต้องมีการเปลี่ยนแปลงมุมมองการรับภาพ



รูปที่ 2.24 ผลการตรวจหาด้วยขั้นตอนวิธีการลบฉากหลัง [31] (ก) ภาพฉากหลัง (ข) ภาพรับเข้า (ค) ผลการลบฉากหลัง (ง) ภาพฉากหน้าที่แยกออกมาได้

### 2.4.3 การตรวจหาด้วยสี (color detection)

ลักษณะสี เป็นอีกทางเลือกหนึ่งในการตรวจหาวัตถุที่สนใจ โดยใช้ค่าความเข้มสีในการตรวจหาวัตถุที่สนใจที่มีสีเป็นจุดเด่นที่สามารถนำมาใช้แยกบริเวณ ในกรณีการตรวจหาใบหน้าก็เป็นอีกงานประยุกต์ที่นักวิจัยส่วนมากเลือกลักษณะสีผิว (skin color) เป็นลักษณะการตรวจหาหลัก เนื่องจากเป็นลักษณะที่สามารถรองรับการเปลี่ยนแปลงรูปร่างของใบหน้าได้ สำหรับการตรวจหาใบหน้าของบุคคลที่กำลังเคลื่อนที่ ซึ่งทำให้ลักษณะทางรูปร่างไม่สามารถรองรับผลกระทบตรงนี้ได้

สิ่งที่ต้องคำนึงถึงเมื่อเลือกใช้การตรวจหาด้วยสีผิว คือ ปริภูมิสีที่จะเลือกใช้ และการเลือกแบบจำลองที่จะนำมาใช้ในการตรวจหาที่เหมาะสมกับการกระจายตัวของสีผิวในแต่ละ ปริภูมิสีที่เลือกใช้ เพราะในแต่ละปริภูมิสีการกระจายตัวขององค์ประกอบสีก็จะต่างกันไป ปริภูมิสีที่นิยมนำมาใช้ เช่น ปริภูมิสี RGB ซึ่งเป็นปริภูมิสีเบื้องต้นของกระบวนการประมวลผลภาพและวิดีโอทัศน์ ปริภูมิสี YCbCr ปริภูมิสี RGB นอร์มัลไลซ์ ปริภูมิสี HSI เป็นต้น (ดังได้กล่าวรายละเอียดของแต่ละปริภูมิสีในส่วนต้นของบทนี้)

V. Vezhnevets และคณะ [21] เสนองานวิจัยเกี่ยวกับการสำรวจงานวิจัยต่าง ๆ ที่ใช้คุณลักษณะสีผิวในการตรวจหาใบหน้า ซึ่งได้แบ่งวิธีการตัดสินใจสำหรับพิจารณาการตรวจหาว่าเป็นส่วนบริเวณสีผิวและที่ไม่ใช่สีผิว ออกเป็น 4 ลักษณะ ดังนี้

-แบบวิธีการกำหนดช่วงบริเวณที่แน่นอน (Explicitly defined skin region) จะมีการกำหนดเป็นช่วงขอบเขตบริเวณที่แน่นอนในแต่ละองค์ประกอบสีในแต่ละปริภูมิสีที่เลือกใช้ ดังตัวอย่างในสมการที่ (2.23) ซึ่งใช้กับปริภูมิสี RGB มีข้อดีที่ง่ายต่อการออกแบบและสามารถประยุกต์ใช้กับงานที่ต้องการความเร็วในการประมวลผล แต่ประสิทธิภาพจะขึ้นกับการเลือกปริภูมิสีและเงื่อนไขการตัดสินใจที่ตรงมีความเหมาะสมเพียงพอต่อภาพหรือวิดีโอที่นำมาใช้ในการตรวจหา

$$\begin{aligned} (R, G, B) \text{ is classified as skin color if :} \\ R > 95 \text{ and } G > 40 \text{ and } B > 20 \text{ and} \\ \max\{R, G, B\} - \min\{R, G, B\} > 15 \text{ and} \\ |R - G| > 15 \text{ and } R > G \text{ and } R > B \end{aligned} \quad (2.23)$$

#### -วิธีการจำลองการกระจายตัวของส่วนสีผิวด้วยแบบจำลองชนิดไม่มีตัวแปร

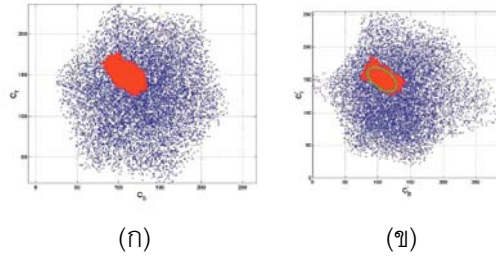
(Non-parametric model) จะใช้วิธีประมาณการกระจายตัวของสีผิวจากข้อมูลฝึกสอนโดยตรง โดยไม่ต้องมีการสร้างแบบจำลองสีมารองรับ เช่น วิธีการเปรียบเทียบค่าจากตาราง (Lookup Table Method) และ วิธีของเบย์ (Bayesian Method) เป็นต้น ซึ่งมีข้อดีที่ความเร็วในการฝึกสอนและไม่ขึ้นอยู่กับการกระจายตัวขององค์ประกอบสีในแต่ละปริภูมิสี แต่ข้อเสียก็จะต้องใช้เนื้อที่ในการเก็บข้อมูลฝึกสอนที่จะนำมาเปรียบเทียบที่ค่อนข้างมาก

#### -วิธีการจำลองการกระจายตัวของส่วนสีผิวด้วยแบบจำลองชนิดมีตัวแปร

(Parametric model) วิธีนี้จะหาแบบจำลองในการจำลองการกระจายตัวขององค์ประกอบสีของส่วนสีผิวที่เหมาะสมต่อการกระจายตัวในแต่ละปริภูมิสี ซึ่งส่วนใหญ่ในงานวิจัย [21, 22] จะพิจารณาการกระจายตัวของเฉพาะองค์ประกอบสี (chrominance component) เท่านั้น โดยจะตัดองค์ประกอบทางแสง (luminance component) ออก เพื่อลดผลกระทบทางแสงในขณะรับภาพมาประมวลผลและลดเวลาในการคำนวณ

ในวิทยานิพนธ์นี้จะสนใจการประยุกต์ใช้การตรวจหาสีผิวโดยใช้ขั้นตอนวิธีสร้างแบบจำลองวงรี จึงขอกล่าวเกี่ยวกับรายละเอียดบางส่วนที่อ้างอิงจากงานวิจัย [15, 16] จากรายละเอียดใน [20] ได้แสดงให้เห็นว่าการแทนการกระจายตัวขององค์ประกอบสีด้วยแบบจำลองแบบเกาส์ไม่สามารถแทนได้อย่างมีประสิทธิภาพ เนื่องจากความไม่สมมาตรของกลุ่มสีผิว และแสดงให้เห็นว่าการกระจายตัวมีโครงสร้างรูปร่างคล้ายวงรี จึงเสนอการตรวจหาสีผิวด้วยแบบจำลองวงรี (Elliptical boundary model)





รูปที่ 2.25 การกระจายตัวของค่าองค์ประกอบสี (ก)การกระจายตัวขององค์ประกอบสี Cb และ Cr ของส่วนบริเวณสีผิว (ข)ขอบเขตวงรี [16]

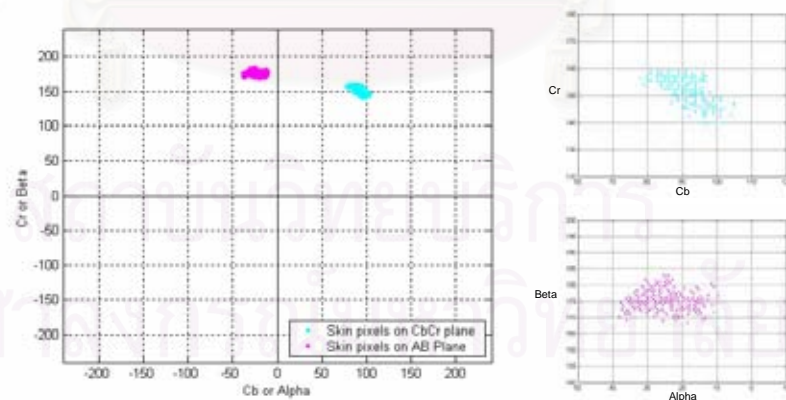
สมการวงรี ที่นำมาใช้ในการแยกบริเวณสีผิว คือ

$$\frac{(x-x_0)^2}{a^2} + \frac{(y-y_0)^2}{b^2} = 1 \quad (2.24)$$

ถ้าพิจารณาจากรูปที่ 2.25(ข) จะเห็นว่าวงรีของขอบเขตสีผิวมีแกนเอก (major axis) และแกนโท (minor axis) ไม่ขนานกับแกนปกติ (Cb,Cr) งานวิจัย [15] จึงเสนอวิธีการในการแปลงคู่พิกัดให้ไปอยู่ในแนวแกนใหม่ ซึ่งกำหนดให้แปลงเป็นคู่พิกัด  $\alpha$  และ  $\beta$  มีสมการการแปลงคู่อันดับ (Cb,Cr) ดังสมการที่ (2.25)

$$\begin{bmatrix} \alpha(Cb, Cr) \\ \beta(Cb, Cr) \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} Cb \\ Cr \end{bmatrix} \quad (2.25)$$

โดยค่า  $\theta$  เป็นมุมของแกนเอก (major axis) และแกนโท (minor axis) ของวงรี ส่วนบริเวณสีผิวทำกับแกน (Cb,Cr) ปกติ



รูปที่ 2.26 การแปลงพิกัดจากพิกัด (Cb,Cr) ไปเป็นพิกัด  $(\alpha, \beta)$  [15]

จากนั้นหลังจากผ่านกระบวนการแปลงแนวแกนใหม่แล้ว ก็หาค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของพิกัด  $\alpha$  และ  $\beta$  เพื่อนำมากำหนดขนาดวงรีที่จะนำไปใช้ในการแยกส่วนสี



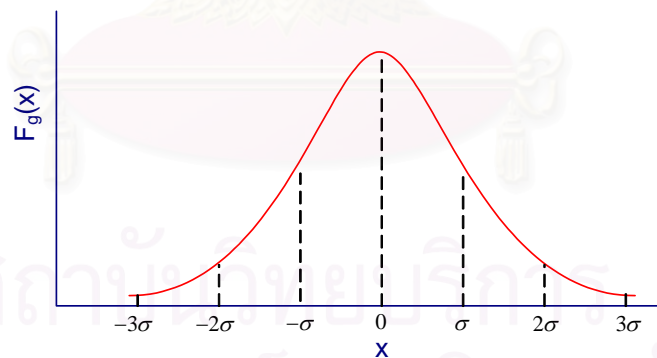
ผิว สมการในการหาจุดยอดของแกนเอกและโทของวงรีในแกน  $(\alpha, \beta)$  มีดังสมการที่ (2.26) และ (2.27) รูปที่ 2.26 แสดงรูปการแปลงแนวแกนจาก  $(Cb, Cr)$  เป็น  $(\alpha, \beta)$

$$V_{\pm\alpha} = \mu_{\alpha} \pm \eta * std(\alpha) \quad (2.26)$$

$$V_{\pm\beta} = \mu_{\beta} \pm \eta * std(\beta) \quad (2.27)$$

โดยที่  $\mu_{\alpha}$  คือ ค่าเฉลี่ยของ  $\alpha$   
 $\mu_{\beta}$  คือ ค่าเฉลี่ยของ  $\beta$   
 $std(\alpha)$  คือ ค่าเบี่ยงเบนมาตรฐานของ  $\alpha$   
 $std(\beta)$  คือ ค่าเบี่ยงเบนมาตรฐานของ  $\beta$

ค่า  $V_{\pm\alpha}$  และ  $V_{\pm\beta}$  เป็นค่าจุดยอดของแกนเอกและแกนโทตามลำดับ ส่วนค่า  $\eta$  เป็นค่าตัดสินสำหรับเลือกขนาดของวงรีที่จะนำไปใช้ ทั้งในแนวแกนเอกและแกนโท พิจารณาจากพื้นที่ใต้กราฟการแจกแจงปกติ (normal distribution) ทั้งในด้านบวกและด้านลบ โดยเป็นการกำหนดขอบเขตของข้อมูลที่ต้องการในการระบุจุดภาพสี่ผิว ซึ่งในวิทยานิพนธ์นี้จะเลือกพิจารณาที่ค่า  $\eta$  เป็นจำนวนเต็ม หรือก็คือ เป็นจำนวนเท่าของค่าเบี่ยงเบนมาตรฐานของกลุ่มข้อมูล โดยจะเลือกใช้ค่า  $\eta$  มีค่า 1 2 3 4 และ 5 โดยแต่ละค่าจำนวนเท่าจะสามารถแสดงถึงจำนวนพื้นที่ใต้กราฟดังรูปที่ 2.27 ได้ดังตารางที่ 2.2



รูปที่ 2.27 กราฟการแจกแจงข้อมูลปกติและแนวขอบเขตพื้นที่ใต้กราฟที่เป็นจำนวนเท่าของค่าเบี่ยงเบนมาตรฐาน ( $\sigma$ )

ตารางที่ 2.2 ขนาดพื้นที่ได้กราฟของกราฟการแจกแจงปกติเมื่อพิจารณาเป็นจำนวนเท่าของค่าเบี่ยงเบนมาตรฐาน [36]

* จำนวนเท่าของค่าเบี่ยงเบนมาตรฐาน ( $\sigma$ )	† ขนาดพื้นที่ได้กราฟ	ขนาดพื้นที่ได้กราฟ (ร้อยละ)
1	0.68270	68.27
2	0.95450	95.45
3	0.99730	99.73
4	0.99994	99.99
5	0.99999	99.99

\* ใช้กำหนดขอบเขตพื้นที่ได้กราฟทั้งในด้านบวกและด้านลบ

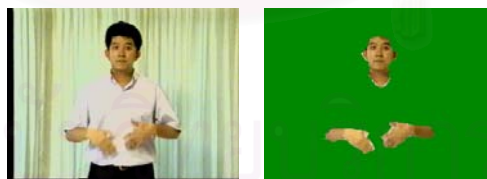
† เมื่อพื้นที่ได้กราฟทั้งหมดมีค่าเท่ากับ 1

รูปที่ 2.28 และรูปที่ 2.29 แสดงตัวอย่างผลการตรวจหาบริเวณสีผิวด้วยแบบจำลองวงรี



รูปที่ 2.28 ผลการตรวจหาในงานวิจัย [16] (ก) ภาพรับเข้า (ข) ภาพหลังผ่าน

กระบวนการขีดเซตแสง (ค) ผลการตรวจหาสีผิวของภาพ (ก) (ง) ผลการตรวจหาสีผิวของภาพ (ข)



รูปที่ 2.29 ผลการตรวจหาสีผิว [15]

#### 2.4.4 กระบวนการติดตามใบหน้า

ในกระบวนการติดตามวัตถุ จุดประสงค์หลักของกระบวนการนี้ คือ การลดเวลาในการประมวลผลของลำดับภาพวิดีโอที่ค้นในแต่ละเฟรม แนวทางการประมวลผลของลำดับภาพวิดีโอที่เริ่มเริ่มต้นจากในเฟรมแรกจะตรวจหาสีผิวและทำกระบวนการอื่น(การลดสัญญาณรบกวนและการวิเคราะห์จุดเชื่อมต่อ) ทุกจุดภาพภายในภาพ ซึ่งใช้เวลาประมวลผลมาก ถ้านำกระบวนการทั้งหมดไปใช้สำหรับทุกลำดับภาพวิดีโอที่ค้นในเวลาจริง จะต้องใช้เวลาในการประมวลผลมาก การ

ลดเวลาในการประมวลผลจึงเป็นสิ่งจำเป็น พิจารณาผลการตรวจหาที่ได้จากเฟรมแรก คือ ตำแหน่งจุดศูนย์กลางของวัตถุเป้าหมาย และเมื่อพิจารณาลักษณะของสัญญาณวิดีโอที่ค้นในเฟรมที่ต่อเนื่องกัน พบว่าภาพทั้งสองเฟรมมีความแตกต่างกันไม่มาก ดังนั้นถ้าต้องการลดเวลาในการประมวลผล แนวทางหนึ่งที่ทำได้ คือ การลดขอบเขตในการตรวจหาลงเมื่อเปรียบเทียบกับเฟรมเริ่มต้น หลักการเบื้องต้นคือ ใช้ขอบเขตของบริเวณวัตถุในเฟรมที่ผ่านมา รวมกับระยะจุดภาพที่ต้องการตรวจหาสี่ผิวนำไปให้ครอบคลุมกับตำแหน่งวัตถุที่จะเกิดขึ้นในเฟรมถัดไป ดังรูปที่ 2.30 และแสดงได้ดังสมการที่ (2.28)

$$\begin{bmatrix} x_{Search}(t+1) \\ y_{Search}(t+1) \end{bmatrix} = \begin{bmatrix} x(t) \pm W \\ y(t) \pm W \end{bmatrix} \quad (2.28)$$

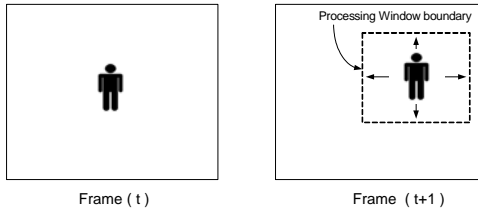
เมื่อ  $x_{Search}(t+1)$  และ  $y_{Search}(t+1)$  คือ ขอบเขตการประมวลผลภาพในแนวแกนนอนและแกนตั้งของการตรวจหา ณ เฟรมภาพถัดไป  $x(t)$  และ  $y(t)$  คือ ขอบเขตของบริเวณวัตถุในแนวแกนนอนและแกนตั้งที่ตรวจหาได้ ณ เฟรมภาพปัจจุบัน และ  $W$  คือ ขอบเขตในการค้นหาเพิ่ม

ระยะจุดภาพที่เพิ่มเข้าไปในส่วนขอบเขตของการตรวจหาสี่ผิวนี้ ถ้ากำหนดเป็นค่าคงที่ไว้ก็จะง่ายต่อการออกแบบระบบ แต่จะมีข้อเสีย คือ ถ้าวัตถุมีการเคลื่อนที่เร็ว กล่าวคือ ภาพสองภาพมีความแตกต่างกันมากขึ้น วัตถุดังกล่าวอาจเคลื่อนที่ออกนอกขอบเขตที่กำหนด ดังนั้นจึงต้องพิจารณาการออกแบบขอบเขตในการตรวจหาแบบปรับเปลี่ยนได้ พารามิเตอร์เกี่ยวกับการเคลื่อนที่ของวัตถุที่ถูกนำมาใช้ คือ ความเร็วและความเร่งในการเคลื่อนที่ของวัตถุ ซึ่งหาได้จากตำแหน่งจุดภาพของวัตถุในเฟรมปัจจุบันเทียบกับเฟรมที่ผ่านมา ดังนั้นสามารถประมาณตำแหน่งวัตถุในเฟรมถัดไปได้ดังสมการที่ (2.29)

$$\begin{bmatrix} x(t+1) \\ y(t+1) \end{bmatrix} = \begin{bmatrix} x(t) + V_x(t) \\ y(t) + V_y(t) \end{bmatrix} + \frac{1}{2} \begin{bmatrix} A_x(t) \\ A_y(t) \end{bmatrix} \quad (2.29)$$

เมื่อ  $(x(t+1), y(t+1))$  และ  $(x(t), y(t))$  คือตำแหน่งจุดภาพของวัตถุในเฟรมถัดไป และเฟรมปัจจุบัน ตามลำดับ  $V_x(t)$  และ  $V_y(t)$  คือค่าความเร็วของวัตถุที่เฟรมปัจจุบันในแนวแกนนอนและแกนตั้ง  $A_x(t)$  และ  $A_y(t)$  คือค่าความเร่งของวัตถุที่เฟรมปัจจุบันในแนวแกนนอนและแกนตั้ง

แต่เนื่องจากในวิทยานิพนธ์นี้ ต้องนำไปประยุกต์ใช้ประมวลผลกับสัญญาณวิดีโอที่รับจากกล้องวิดีโอ PTZ ที่เปลี่ยนมุมมองการรับภาพได้ จึงจะเกิดปัญหาขึ้นในกรณีที่ขณะกล้องกำลังเปลี่ยนมุมมอง ทำให้อาจเกิดความผิดพลาดในขั้นตอนการติดตาม จึงไม่เลือกนำกระบวนการนี้มาประยุกต์ใช้ในวิทยานิพนธ์นี้



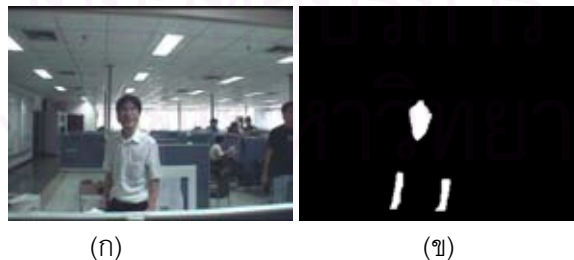
รูปที่ 2.30 การกำหนดขอบเขตในการตรวจหาวัตถุเฟรมถัดไป

#### 2.4.5 การวัดประสิทธิภาพเชิงปริวิสัย (Objective Performance Evaluation)

การวัดประสิทธิภาพการตรวจหาเชิงปริวิสัยมักกระทำโดยการเปรียบเทียบผลการตรวจหาที่ได้ออกมาจากการตรวจหาอ้างอิงหรือผลถูกต้องพื้นฐาน (Ground Truth) อันถือว่าเป็นผลการตรวจหาในอุดมคติ (Ideal Result) และผลการทดลองตามอุดมคตินี้ จะถือว่ามีความใกล้เคียงกับกระบวนการแยกส่วนภาพด้วยตาของมนุษย์มากที่สุด ตามปกติการสร้างผลการทดลองอ้างอิงนี้ มักให้มนุษย์เป็นผู้จำแนกกลุ่มจุดภาพจากภาพรับเข้า ซึ่งในการทดลองในวิทยานิพนธ์นี้ทำโดยการลากเส้นแบ่งขอบเขตในโปรแกรมประยุกต์ด้านการประมวลผลภาพทั่วไป

วิธีการตรวจหาตามปกติจะยึดจุดภาพวัตถุที่สนใจเป็นบริเวณฉากหน้า (Foreground Region) หรือบริเวณบวก (Positive Region) และบริเวณอื่น ๆ ที่นอกเหนือจากบริเวณวัตถุ ให้เป็นบริเวณฉากหลัง (Background Region) หรือบริเวณลบ (Negative Region) โดยการให้สัญลักษณ์จะให้บริเวณบวกเป็นสีขาวและบริเวณลบเป็นสีดำ ดังตัวอย่างรูปที่ 2.31 ซึ่งกำหนดให้บริเวณวัตถุที่สนใจ คือส่วนสีผิวทั้งหมด (ใบหน้าและมือรวมท่อนแขนทั้งสองข้าง ดังนั้นในกรณีของการจำแนกของผลถูกต้องพื้นฐานจะมีเขตของจุดภาพอยู่ 2 ประเภทได้แก่

- $F_T$  เป็นเขตจุดภาพของผลถูกต้องพื้นฐานซึ่งถูกจำแนกให้เป็นฉากหน้า (Truth Foreground)
- $B_T$  เป็นเขตจุดภาพของผลถูกต้องพื้นฐานซึ่งถูกจำแนกให้เป็นฉากหลัง (Truth Background)



รูปที่ 2.31 การกำหนดบริเวณวัตถุที่สนใจ (ก) ภาพรับเข้า (ข) ผลถูกต้องพื้นฐาน

เมื่อได้ผลถูกต้องพื้นฐาน (Ground Truth) แล้ว จะนำผลการทดลองจากขั้นตอนวิธีต่าง ๆ มาเปรียบเทียบความเหมือน-ความต่างกับผลถูกต้องพื้นฐาน เช่นเดียวกับกับกรณีผล

ถูกต้องพื้นฐาน ผลการทดลองที่นำมาเปรียบเทียบจะมีการจำแนกเขตของจุดภาพออกเป็น 2 ประเภทได้แก่

**F<sub>C</sub>** เป็นเขตจุดภาพของผลการทดลองที่พิจารณาซึ่งถูกจำแนกให้เป็นฉากหน้า (Considered Foreground )

**B<sub>C</sub>** เป็นเขตจุดภาพของผลการทดลองที่พิจารณาซึ่งถูกจำแนกให้เป็นฉากหลัง (Considered Background )

ขั้นตอนการเปรียบเทียบจะถือว่าจุดภาพของผลการทดลองที่พิจารณาที่ให้ผลเหมือนกับผลถูกต้องพื้นฐานจะเป็นจุดภาพที่ถูก "จำแนกถูก" และที่ให้ผลต่างกับผลถูกต้องพื้นฐานเป็นจุดภาพที่ถูก "จำแนกผิด" ดังนั้น เมื่อทำการเปรียบเทียบแล้วจะมีจุดภาพอยู่ 4 ประเภท ดังนี้

**T<sub>P</sub>** เป็นเขตจุดภาพที่ถูกจำแนกถูก และเป็นบริเวณฉากหน้าหรือบริเวณบวก

**T<sub>N</sub>** เป็นเขตจุดภาพที่ถูกจำแนกถูก และเป็นบริเวณฉากหลังหรือบริเวณลบ

**F<sub>P</sub>** เป็นเขตจุดภาพที่ถูกจำแนกผิดเนื่องจากการทดลองที่พิจารณาจำแนกให้เป็นฉากหน้าหรือพื้นที่บวก แต่ผลถูกต้องพื้นฐานจำแนกให้เป็นฉากหลังหรือพื้นที่ลบ

**F<sub>N</sub>** เป็นเขตจุดภาพที่ถูกจำแนกผิดเนื่องจากการทดลองที่พิจารณาจำแนกให้เป็นฉากหลังหรือพื้นที่ลบ แต่ผลถูกต้องพื้นฐานจำแนกให้เป็นฉากหน้าหรือพื้นที่บวก

เมื่อสามารถแบ่งจุดภาพออกได้เป็น 4 กลุ่มและจะทำการหาจำนวนของจุดภาพในแต่ละกลุ่มเป็นจำนวน  $n(\mathbf{T}_P)$   $n(\mathbf{T}_N)$   $n(\mathbf{F}_P)$  และ  $n(\mathbf{F}_N)$  จุดภาพตามลำดับ เมื่อ  $n(\cdot)$  เป็นฟังก์ชันการหาสมาชิกในเซตที่พิจารณา จากนั้นจะหาอัตราส่วนของจำนวนจุดภาพในแต่ละกลุ่มโดยการนอร์มัลไลซ์ด้วยขนาดภาพกว้าง  $W$  สูง  $H$  ดังนี้

$$N(\mathbf{T}_P) = \frac{n(\mathbf{T}_P)}{W \times H} \quad (2.30)$$

$$N(\mathbf{T}_N) = \frac{n(\mathbf{T}_N)}{W \times H} \quad (2.31)$$

$$N(\mathbf{F}_P) = \frac{n(\mathbf{F}_P)}{W \times H} \quad (2.32)$$

$$N(\mathbf{F}_N) = \frac{n(\mathbf{F}_N)}{W \times H} \quad (2.33)$$

โดยที่  $N(\mathbf{T}_P)$  เป็นอัตราส่วนจำแนกถูกแบบบวก

$N(\mathbf{T}_N)$  เป็นอัตราส่วนจำแนกถูกแบบลบ

$N(\mathbf{F}_P)$  เป็นอัตราส่วนจำแนกผิดแบบบวก

$N(\mathbf{F}_N)$  เป็นอัตราส่วนจำแนกผิดแบบลบ

#### 2.4.5.1 อัตราการตรวจหา (Detection rate : DR)

อัตราการตรวจหา คือ จำนวนสมาชิกในเซตของจุดภาพฉากหน้าและระบบจำแนกให้เป็นฉากหน้า เมื่อเทียบกับจำนวนสมาชิกในเซตของจุดภาพฉากหน้าทั้งหมด จะได้ค่าอัตราการตรวจหา ดังนี้

$$DR = \frac{N(\mathbf{T}_P)}{N(\mathbf{T}_P) + N(\mathbf{F}_N)} \quad (2.34)$$

โดยที่ อัตราการตรวจหา สามารถบ่งชี้ประสิทธิภาพของการตรวจหาฉากหน้า ว่าผลการทดลองจากวิธีที่พิจารณาสามารถตรวจหาฉากหน้าได้ ครบถ้วนเพียงใด ถ้า อัตราการตรวจหา มีค่ามาก บ่งชี้การตรวจหาฉากหน้าได้อย่างครบถ้วน

#### 2.4.5.2 อัตราการฟ้องความผิดพลาด (False Alarm Rate: FAR)

การใช้อัตราการตรวจหาเป็นตัวชี้วัดเพียงอย่างเดียว ไม่สามารถชี้ประสิทธิภาพได้อย่างสมบูรณ์ เนื่องจากผลการทดลองที่มีพื้นที่การจำแนกเป็นบวก ที่ซ้อนทับกับผลถูกต้องพื้นฐานได้ จะมีค่าอัตราการตรวจหาเท่ากัน แต่ตามมนุษย์จะสามารถสังเกตเห็นส่วนเกินที่ไม่ใช่พื้นที่ที่ต้องการได้ ซึ่งถือเป็นบริเวณจุดภาพที่จำแนกผิด

อัตราการฟ้องความผิดพลาด คือ จำนวนสมาชิกในเซตของจุดภาพฉากหลังและถูกจำแนกผิดให้เป็นฉากหน้า เมื่อเทียบกับ จำนวนสมาชิกในเซตของจุดภาพที่ระบบจำแนกให้เป็นฉากหน้าหรือเป็นบวกทั้งหมด จะได้ค่าอัตราการฟ้องความผิดพลาด ดังนี้

$$FAR = \frac{N(\mathbf{F}_P)}{N(\mathbf{T}_P) + N(\mathbf{F}_P)} \quad (2.35)$$

โดยที่ อัตราการฟ้องความผิดพลาด สามารถบ่งชี้ประสิทธิภาพของการตรวจหาฉากหลัง ว่าผลการทดลองจากวิธีที่พิจารณาสามารถตรวจหาฉากหลังได้ครบถ้วนหรือเกิดการจำแนกผิดแบบบวกเพียงใด ถ้าอัตราการฟ้องความผิดพลาดมีค่ามาก บ่งชี้ระบบจำแนกฉากหลังส่วนเกินให้เป็นฉากหน้ามาก จะถือว่าระบบมีประสิทธิภาพการจำแนกต่ำ

#### 2.4.5.3 อัตราการปฏิเสธความผิดพลาด (False Reject Rate: FRR)

อัตราการปฏิเสธความผิดพลาด คือ จำนวนสมาชิกในเซตของจุดภาพฉากหน้าและระบบจำแนกให้เป็นฉากหลัง เมื่อเทียบกับจำนวนสมาชิกในเซตของจุดภาพฉากหน้าทั้งหมด จะได้ค่าอัตราการปฏิเสธความผิดพลาด ดังนี้

$$FRR = \frac{N(\mathbf{F}_N)}{N(\mathbf{T}_P) + N(\mathbf{F}_N)} \quad (2.36)$$



## 2.5 การควบคุมกล้องวิดีโอ PTZ

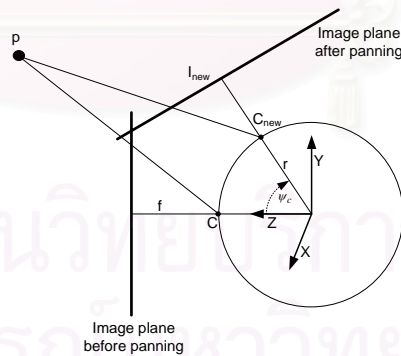
การควบคุมการถ่าย ก้มเงย และซูมของกล้องวิดีโอ PTZ เป็นกระบวนการสำคัญ เป้าหมายหลัก คือ การควบคุมตำแหน่งมุมมองการรับภาพของกล้องวิดีโอ PTZ ให้วัตถุที่สนใจอยู่บริเวณกึ่งกลางของมุมมองภาพ ซึ่งในการเคลื่อนที่จะประกอบด้วยพารามิเตอร์การเคลื่อนที่ที่สำคัญ 3 ตัว คือ มุมถ่าย ( $\psi$ ) มุมก้มเงย ( $\phi$ ) และตัวประกอบการซูม (zoom factor) ( $\kappa$ )

จากงานวิจัยที่ผ่านมาสามารถแบ่งกระบวนการหาพารามิเตอร์ได้เป็นสองแบบ คือ การอ้างอิงจากพิกัดใน 3 มิติ และการอ้างอิงจากพิกัดภาพ 2 มิติ ซึ่งจะต่างกันตรงที่ในแบบแรก จะมีการอ้างอิงจุดภาพจากพิกัดภาพ (image plane) ไปยังพิกัดกล้องที่เป็นพิกัด 3 มิติก่อน เพื่อใช้คำนวณหามุมในการเคลื่อนที่ แต่แบบที่สองจะคำนวณโดยตรงจากพิกัดจุดภาพใน 2 มิติ

**แบบที่ 1 อ้างอิงจากพิกัดใน 3 มิติ** [1] ในการคำนวณมุมของการเคลื่อนที่นั้น จะอาศัยทฤษฎีเกี่ยวกับการฉายภาพ (projection) ของกล้อง ซึ่งต้องมีการนำเรขาคณิตของกล้อง (Camera Geometry) มาประยุกต์ใช้ประกอบการคำนวณ เริ่มต้นสมมติวัตถุจุดในพิกัดโลก

(world coordinate) 3 มิติ อยู่ที่ตำแหน่ง  $\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$

ให้เริ่มต้นพิจารณาที่การเปลี่ยนแปลงมุมถ่าย ( $\psi_c$ ) ที่ 0 องศา และที่มุมก้มเงย ( $\phi$ ) ใดๆ จะได้เรขาคณิตของกล้อง ดังรูปที่ 2.32



รูปที่ 2.32 เรขาคณิตของกล้อง (มุมมองด้านบน) [1]

เมื่อกำหนดให้  $f$  คือ ความยาวโฟกัสของกล้อง และ  $r$  คือ รัศมีการหมุน จะคำนวณค่าตำแหน่งศูนย์กลางการฉายภาพ (center of projection) ของกล้อง ( $C$ ) ได้ดังสมการที่ (2.37)

$$C = \begin{bmatrix} C_x \\ C_y \\ C_z \end{bmatrix} = \begin{bmatrix} 0 \\ -r \cos(\phi) \\ r \sin(\phi) \end{bmatrix} \quad (2.37)$$

ถ้ามีการเปลี่ยนแปลงมุมสาย ( $\psi_c$ ) จะทำให้เกิดการเปลี่ยนแปลงมุมโดยสามารถแสดงความสัมพันธ์ของมุมที่เปลี่ยนแปลงนี้ได้ในรูปแบบเมทริกซ์การหมุน (rotation matrix) โดยเมทริกซ์นี้จะมีรูปแบบขึ้นอยู่กับว่าใช้แกนอะไรเป็นแกนหมุน ในที่นี้ให้แกนหมุนเป็น แกน y (y-axis) จะได้เมทริกซ์การหมุน ( $R$ ) ดังแสดงในสมการที่ (2.38)

$$R = \begin{bmatrix} \cos(\psi_c) & 0 & -\sin(\psi_c) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(\psi_c) & 0 & \cos(\psi_c) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.38)$$

เมื่อมีการหมุนหรือการย้ายพิกัดภาพ จะทำให้ตำแหน่งศูนย์กลางการฉายภาพของกล้องเปลี่ยนไปด้วย จะได้ตำแหน่งใหม่  $C_{new}$  ดังแสดงในสมการที่ (2.39)

$$C_{new} = \begin{bmatrix} -r \sin(\phi) \sin(\psi_c) \\ -r \cos(\phi) \\ r \sin(\phi) \cos(\psi_c) \end{bmatrix} \quad (2.39)$$

เมื่อมีการเคลื่อนที่ของกล้อง ตำแหน่งจุดในพิกัดโลกยังคงตำแหน่งเดิม แต่ที่เปลี่ยนไป คือ ค่าจุดที่มองจากระบบพิกัดของกล้องจะเปลี่ยนไปจากเดิม โดยมีเปลี่ยนแปลงของตำแหน่งศูนย์กลางการฉายภาพในทิศตรงกันข้ามและเมทริกซ์การหมุนจะกลับทิศทางกัน ดังนั้นตำแหน่งของจุดหลังจากการเคลื่อนที่จะถูกแสดงในสมการที่ (2.40)

$$\begin{bmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \\ 1 \end{bmatrix} = \begin{bmatrix} X \cos(\psi_c) + Z \sin(\psi_c) \\ Y + r \cos(\phi) \\ -X \sin(\psi_c) + Z \cos(\psi_c) - r \sin(\phi) \\ 1 \end{bmatrix} \quad (2.40)$$

จากสมการที่ (2.40) ( $X_{cam}, Y_{cam}, Z_{cam}$ ) เป็นจุดตำแหน่งวัตถุในระบบพิกัดกล้อง ซึ่งจะเห็นว่าถ้าต้องการจะแปลงให้อยู่ในระบบพิกัดภาพ จะต้องใช้กระบวนการสร้างภาพดังที่กล่าวไว้ตอนต้น จะได้ตำแหน่งของระบบพิกัดภาพใหม่  $I_{new} (I_x, I_y)$  หลังจากจากกล้องเคลื่อนที่ ดังแสดงในสมการที่ (2.41) ซึ่งแสดงความสัมพันธ์เมื่อมีการกำหนดมุมสายและมุมก้มเงย จะสามารถคำนวณหาตำแหน่งของระบบพิกัดภาพใหม่ได้ ในทางกลับกันถ้าทราบระบบพิกัดอันใหม่ ก็ทำให้สามารถคำนวณหามุมสายและมุมก้มเงยได้เพื่อนำไปใช้ในการควบคุมกล้องต่อไป

$$I_{new} = \begin{bmatrix} I_x \\ I_y \end{bmatrix} = \begin{bmatrix} \alpha_u \frac{X \cos \psi_c + Z \sin \psi_c}{-X \sin \psi_c + Z \cos \psi_c - r \sin \phi} + u_0 \\ \alpha_v \frac{Y + r \cos \phi}{-X \sin \psi_c + Z \cos \psi_c - r \sin \phi} + v_0 \end{bmatrix} \quad (2.41)$$

โดยที่  $(u_0, v_0)$  คือ จุดमुखสำคัญ (principle point) ของภาพ และ  $\alpha_u, \alpha_v$  คือ ตัวประกอบมาตรา (scale factor) ทางด้านภาพในแนวแกนขนานและแกนตั้งตามลำดับ

**แบบที่ 2 อ้างอิงจากพิกัดภาพ 2 มิติ** [13] ในการคำนวณมุมของการเคลื่อนที่และอัตราการซูมนั้น จะอาศัยค่าตำแหน่งของจุดพิกัดวัตถุในพิกัดภาพเท่านั้น กล่าวคือเมื่อผ่านกระบวนการตรวจหาแล้ว ตำแหน่งวัตถุซึ่งอาจจะเป็นจุดศูนย์กลางถ่วงของวัตถุ จะถูกนำมาใช้สำหรับคำนวณหามุมของการเปลี่ยนมุมมองของกล้องและอัตราการซูมโดยตรง ดังสมการที่ (2.42)

$$\begin{bmatrix} \psi_n \\ \phi_n \\ \kappa_n \end{bmatrix} = \begin{bmatrix} \psi_c + \mu(\kappa_c) * (x_{t+1}^p - C_x) \\ \phi_c + \mu(\kappa_c) * (y_{t+1}^p - C_y) \\ \kappa_c \end{bmatrix} \quad (2.42)$$

เมื่อ  $(\psi_n, \phi_n, \kappa_n)$  เป็นพารามิเตอร์การเคลื่อนที่ที่คำนวณได้ (มุมสาย มุมก้มเงย และตัวประกอบการซูม ตามลำดับ)  $(\psi_c, \phi_c, \kappa_c)$  เป็นพารามิเตอร์การเคลื่อนที่ ณ ตำแหน่งปัจจุบัน  $\mu(\kappa_c)$  เป็นตัวประกอบของอัตราขยาย (magnification factor)  $(C_x, C_y)$  เป็นตำแหน่งจุดศูนย์กลางกลางของภาพ และ  $(x_{t+1}^p, y_{t+1}^p)$  คือตำแหน่งประมาณของวัตถุในเฟรมถัดไป โดยในงานวิจัยนี้ จะสร้างส่วนตัดสินใจเพื่อเช็คค่าตำแหน่งประมาณของวัตถุในเฟรมถัดไปหลุดจากมุมมองการรับภาพของกล้อง จึงจะสั่งการให้กล้องเคลื่อนที่ไปโดยใช้พารามิเตอร์ที่คำนวณได้ดังสมการที่ (2.42) สำหรับตัวประกอบการซูมจำคำนวณได้ดังสมการที่ (2.43)

$$\kappa_n = \frac{s_n * \mu(\kappa_c)}{s_c} \quad (2.43)$$

เมื่อ  $s_n$  และ  $s_c$  คือ ขนาดของวัตถุหรือฉากหน้าที่ทำการตรวจหา ณ เฟรมปัจจุบัน และขนาดของวัตถุที่ต้องการ ตามลำดับ

เมื่อพิจารณาการคำนวณหาพารามิเตอร์ทั้งสองแบบ จะเห็นว่าแบบแรกที่ใช้การคำนวณอ้างอิงจากพิกัด 3 มิติ มีการคำนวณที่ค่อนข้างซับซ้อน แต่เมื่อพิจารณาที่ผลลัพธ์จะได้ตำแหน่งการเคลื่อนที่ที่แม่นยำกว่าแบบที่สอง เพราะคำนวณจากแกนพิกัดในการเคลื่อนที่ของกล้อง แต่สำหรับแบบที่สองซึ่งอ้างอิงจากพิกัดภาพ 2 มิติ จะเป็นการประมาณมุมของการเปลี่ยนมุมมองกล้องและอัตราการซูมเพื่อต้องการให้วัตถุมาอยู่กึ่งกลางภาพเท่านั้น อาจมีความผิดพลาดในตำแหน่งการเคลื่อนที่ แต่มีข้อดีที่ง่ายต่อการคำนวณและใช้เวลาน้อยในการประมวลผล

## 2.6 การจับคู่ใบหน้า (Face matching)

การจับคู่ใบหน้า เป็นกระบวนการสำหรับการเปรียบเทียบภาพใบหน้าที่รับเข้ากับภาพใบหน้าทั้งหมดที่ถูกเก็บในฐานข้อมูล ว่ามีความใกล้เคียงกับภาพใบหน้าที่ใดในฐานข้อมูลมากที่สุด โดยจะต้องเลือกลักษณะ (feature) ที่จะนำมาเป็นตัวเปรียบเทียบ เช่น ลักษณะสี รูปร่าง และขอบ เป็นต้น โดยวิธีการหนึ่งที่เป็นที่นิยมใช้สำหรับการจับคู่ คือ กระบวนการวัดความคล้าย (Similarity measure) โดยอาศัยวิธีการวัดระยะทางระหว่างกลุ่มการกระจายตัวของข้อมูล 2 ชุดที่สนใจ ซึ่งมีวิธีการในการหาระยะทางนี้ได้หลายแบบ ในวิทยานิพนธ์นี้จะมุ่งสนใจเฉพาะการเปรียบเทียบฮิสโทแกรมด้วยระยะ Bhattacharyya ซึ่งมีรายละเอียดดังนี้

### 2.6.1 การเปรียบเทียบฮิสโทแกรมด้วยระยะ Bhattacharyya [17, 23]

ระยะ Bhattacharyya เป็นรูปแบบการเปรียบเทียบลักษณะหนึ่งในกระบวนการวัดความคล้าย ระหว่างชุดการกระจายตัวของข้อมูล 2 ชุด ซึ่งได้ถูกนำมาประยุกต์ใช้ในกระบวนการทางการแพทย์สำหรับการเปรียบเทียบฮิสโทแกรมของสองกลุ่มข้อมูลที่ไม่จำเป็นต้องเป็นค่าฮิสโทแกรมของทั้งภาพ อาจเป็นแค่บริเวณใดบริเวณหนึ่งที่สนใจเท่านั้น โดยไม่จำเป็นต้องเป็นทั้งสองบริเวณจะต้องมีขนาดเท่ากัน (ในหน่วยจุดภาพ) จึงเป็นข้อดีที่สามารถรองรับการเปรียบเทียบวัตถุที่มีการเปลี่ยนแปลงรูปร่างได้ทั้งในลักษณะการเคลื่อนที่ที่เปลี่ยนมุมมองและการเปลี่ยนแปลงขนาด

สมมติให้  $\hat{p}$  เป็นฮิสโทแกรมสีของแบบจำลองเป้าหมายและ  $\hat{q}$  เป็นฮิสโทแกรมสีของแบบจำลองรับเข้าที่ต้องการนำมาเปรียบเทียบ โดย  $\hat{p} = \{\hat{p}_u\}_{u=1, \dots, m}$  (เมื่อ  $\sum_{u=1}^m \hat{p}_u = 1$ ) และ  $\hat{q} = \{\hat{q}_u\}_{u=1, \dots, m}$  (เมื่อ  $\sum_{u=1}^m \hat{q}_u = 1$ ) เมื่อ  $m$  เป็นจำนวนถังสีของฮิสโทแกรม (bin histogram) เปรียบเทียบ เช่น ถ้าในกรณีเลือกใช้ขนาดขององค์ประกอบสีแต่ละองค์ประกอบมีขนาด 8 บิต ฉะนั้นจะได้ขนาดถังสีเท่ากับ 256 ถัง ดังนั้นสามารถคำนวณหาสัมประสิทธิ์ Bhattacharyya ได้ดังสมการที่ (2.44)

$$\hat{p} = \sum_{u=1}^m \sqrt{\hat{p}_u \times \hat{q}_u} \quad (2.44)$$

อธิบายความหมายทางเรขาคณิตตามสมการที่ (2.44) จะแสดงถึงค่าโคไซน์ (cosine) ของมุมระหว่างค่าถังสีทั้ง  $m$  มิติในรูปเวกเตอร์หนึ่งหน่วย  $(\sqrt{\hat{p}_1}, \dots, \sqrt{\hat{p}_m})^T$  และ  $(\sqrt{\hat{q}_1}, \dots, \sqrt{\hat{q}_m})^T$

จากสมการที่ (2.44) จะสามารถหาค่าระยะ Bhattacharyya ที่เป็นระยะระหว่างสองกลุ่มของการกระจายตัวของข้อมูล ( $d_{Bh}$ ) ดังสมการที่ (2.45)

$$d_{Bh} = \sqrt{1 - \rho[\hat{p}, \hat{q}]} \quad (2.45)$$

ข้อดีของการเปรียบเทียบวัดความคล้ายด้วยระยะ Bhattacharyya คือ สามารถรองรับการเปรียบเทียบชุดข้อมูล 2 ชุดที่ไม่จำเป็นต้องมีขนาดหรือจำนวนข้อมูลที่เท่ากันก็ได้ ซึ่งนอกจากการเปรียบเทียบด้วยระยะ Bhattacharyya แล้ว อีกรูปแบบที่มีการนำมาใช้ คือ การเปรียบเทียบวัดความคล้ายด้วยระยะ Mahalanobis แต่เมื่อพิจารณาจากงานวิจัยที่ [24] ก็จะเห็นว่าจากผลการทดลองแสดงให้เห็นว่าการวัดความคล้ายด้วยระยะ Bhattacharyya มีประสิทธิภาพที่ดีกว่าการวัดความคล้ายด้วยระยะ Mahalanobis



สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

## บทที่ 3

### โครงสร้างและระบบตรวจหาติดตามที่นำเสนอ

ในบทที่ 3 นี้จะกล่าวถึงโครงสร้างและระบบการตรวจหาติดตามใบหน้าที่น่าเสนอ จากบทที่ 1 จะเห็นว่าระบบกล้องหลายตัวส่วนใหญ่ [1, 2, 5, 7, 8] ประยุกต์ใช้กล้องวีดิทัศน์ PTZ ในรูปแบบของการรับคำสั่งพารามิเตอร์การถ่าย ก้มเงย และซูม จากกล้องวีดิทัศน์อยู่กับที่ซึ่งเป็น กล้องวีดิทัศน์ตัวหลักในระบบ ทำให้กล้องวีดิทัศน์ PTZ ไม่ต้องการประมวลผลภาพ แต่ทำให้ต้อง เปลี่ยนการใช้กล้องวีดิทัศน์เพิ่มขึ้นอย่างน้อยอีกหนึ่งตัวเข้ามาในระบบ และเมื่อพิจารณาที่ขั้นตอน วิธีการตรวจหาที่ประยุกต์ใช้กับกล้องวีดิทัศน์ PTZ ส่วนใหญ่จะเลือกกระบวนการลบฉากหลังมา ประยุกต์ใช้งาน ซึ่งมีข้อเสียที่ ต้องเปลี่ยนเนื้อที่หน่วยความจำมากในการเก็บฐานข้อมูลฉากหลังในแต่ละตำแหน่งการเคลื่อนที่ของกล้องวีดิทัศน์ PTZ เพิ่มความยุ่งยากในขั้นตอนการเก็บฐานข้อมูล เมื่อพิจารณาในการใช้งาน อาจมีความคลาดเคลื่อนในการเลือกฉากหลังที่เหมาะสมต่อการลบฉาก หลังทั้งในเรื่องของตำแหน่งและเรื่องผลกระทบของแสง และตำแหน่งการเคลื่อนที่ของกล้องไม่เป็น อิสระ

วิทยานิพนธ์นี้เสนอระบบการตรวจหาและติดตามใบหน้าโดยใช้กล้องวีดิทัศน์ PTZ 2 ตัว ประมวลผลร่วมกัน สำหรับนำไปประยุกต์ใช้งานด้านการจับคู่ใบหน้าเพื่อระบุบุคคล ซึ่ง ช่วยเพิ่มประสิทธิภาพทั้งในขั้นตอนการจับคู่ใบหน้าและขั้นตอนการตรวจหา ระบบที่นำเสนอจะ แบ่งการทำงานร่วมกันออกเป็น 2 แบบ คือ 1. การทำงานร่วมกันในการประมวลผลสำหรับการระบุ บุคคลด้วยภาพใบหน้า 2. การทำงานร่วมกันในการประมวลผลการตรวจหา ขั้นตอนวิธีการ ตรวจหาใบหน้าที่น่าเสนอจะประยุกต์ใช้การตรวจหาด้วยสีผิว (ดังวิธีที่กล่าวในบทที่ 2) ซึ่งสามารถ รองรับการตรวจหาในกรณีกล้องวีดิทัศน์มีการถ่าย และก้มเงยเปลี่ยนมุมมองได้ และนำการ ตรวจหาการเคลื่อนที่มาใช้งานร่วมด้วย เพื่อเพิ่มประสิทธิภาพในการตรวจหา และลดผลกระทบ ของสัญญาณรบกวนที่เกิดขึ้นเนื่องจากสีของฉากหลังที่คล้ายกับสีผิว ทำให้ไม่เกิดปัญหาในกรณีที่ ฉากหลังมีการเปลี่ยนแปลง ส่งผลให้การเปลี่ยนมุมมองการรับภาพของกล้องวีดิทัศน์สามารถ เปลี่ยนตำแหน่งได้อย่างอิสระ และได้นำกระบวนการประมวลผลเชิงสัญญาณ คือ ตัวดำเนินการ แบบเปิด ที่จะทำกระบวนการอีโรชันแล้วตามด้วยการทำไดเลชัน เพื่อลดสัญญาณรบกวนเล็กน้อย ที่เกิดจากกระบวนการตรวจหาให้ถ่ายทอดการวิเคราะห์ส่วนบริเวณใบหน้า สำหรับการระบุบุคคล สิ่งที่ต้องคำนึงถึง คือ การทำงานในเวลาจริงและตำแหน่งของใบหน้าที่รับภาพได้ เนื่องจากไม่ได้ กำหนดเส้นทางการเดินของบุคคล ดังนั้นจึงเลือกใช้กระบวนการจับคู่ใบหน้าสำหรับการระบุบุคคล ด้วยกระบวนการวัดความคล้ายด้วยระยะ Bhattacharyya ที่สามารถรองรับการทำงานในเวลาจริง



โดยได้นำเสนอวิธีการการจำแนกภาพประเภทของใบหน้าที่จับภาพได้ ทำให้สามารถจัดหมวดหมู่ของการจับคู่ใบหน้าได้ดีขึ้น โดยเป้าหมายของกระบวนการจับคู่นี้ไม่ได้อยู่ที่การพัฒนาขั้นตอนวิธีการจับคู่ใบหน้า แต่จะนำเสนอประโยชน์ที่ได้จากการทำงานร่วมกันของกล้อง

### 3.1 รูปแบบการทำงานของระบบที่นำเสนอ

รูปแบบการทำงานของระบบที่นำเสนอจะแบ่งออกเป็น 2 รูปแบบ คือ

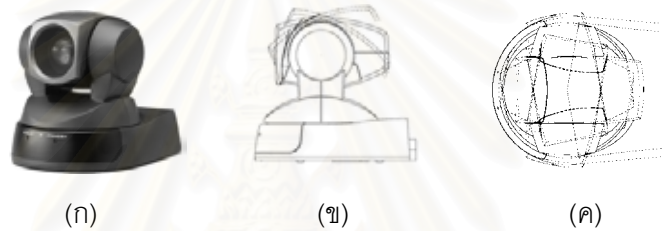
รูปแบบที่ 1 จะนำเสนอวิธีการประมวลผลจับคู่ใบหน้าเพื่อระบุบุคคลร่วมกันของกล้องวิดีโอ PTZ 2 ตัว เวลาจริง แนวความคิดหลักของรูปแบบการทำงานนี้ คือ ในการประมวลผลสำหรับตรวจหา ติดตาม และจับคู่ใบหน้า บุคคลเคลื่อนที่ ในเวลาจริง ตำแหน่งของภาพใบหน้าที่กล้องรับภาพได้จะมีหลายรูปแบบ บางครั้งอาจจับภาพได้ในแบบหน้าตรง หรือบางครั้งอาจจับภาพได้แบบหน้าเอียง ซึ่งผลกระทบต่อประสิทธิภาพในการจับคู่ใบหน้าอย่างมาก เพราะเมื่อพิจารณาจากงานวิจัยที่ผ่านมา [3, 4, 19] ในด้านการรู้จำใบหน้า ระบบจะรู้จำภาพใบหน้าตรงหรือค่อนข้างตรงได้ดีกว่า ด้วยข้อจำกัดดังกล่าว ในการออกแบบระบบจึงแก้ปัญหา โดยตั้งกล้องวิดีโอ PTZ 2 ตัว ในตำแหน่งที่ต่างกัน แต่ต้องมีมุมมองการรับภาพที่เหลื่อมกันอยู่ ทำให้มีความน่าจะเป็นที่กล้องตัวใดตัวหนึ่งในสองตัวจะจับภาพใบหน้าได้ในลักษณะหน้าตรงหรือค่อนข้างตรง ณ เวลาเดียวกัน ซึ่งเป็นข้อได้เปรียบกว่าระบบที่มีกล้องวิดีโอตัวเดียว เพราะการที่มีกล้องวิดีโอเพียงตัวเดียว อาจมีความเป็นไปได้สูงที่จะจับภาพใบหน้าได้เฉพาะด้านข้าง

รูปแบบที่ 2 นำเสนอลักษณะการทำงานร่วมกันในกระบวนการตรวจหาบุคคลระหว่างกล้องวิดีโอ PTZ 2 ตัว รูปแบบการทำงานจะเป็นการตรวจหาและเคลื่อนที่ติดตามบุคคลเป้าหมายคนเดียวอย่างต่อเนื่องจากกล้องวิดีโอตัวหนึ่งไปยังกล้องวิดีโอตัวที่สอง จะทำให้ระบบสอดคล้องและรักษาความปลอดภัยสามารถติดตามการเคลื่อนที่ของบุคคลเป้าหมายที่สนใจได้อย่างต่อเนื่องและมีประสิทธิภาพ กล่าวคือ เมื่อบุคคลเป้าหมายเดินเข้ามาในระบบภายในอาณาบริเวณการรับภาพของกล้องวิดีโอตัวหนึ่ง จะทำการเก็บข้อมูลของบุคคลนั้นในรูปแบบของข้อมูลลักษณะสี จากนั้นเมื่อบุคคลเป้าหมายเคลื่อนที่พ้นจากขอบเขตของกล้องวิดีโอตัวที่หนึ่ง จะส่งข้อมูลนี้ไปยังกล้องวิดีโออีกตัว ทำให้กล้องวิดีโอตัวที่สองนี้จะเปลี่ยนสถานะจากตรวจหาอย่างเดียวเป็นตรวจหาและทำการเปรียบเทียบเพื่อระบุว่าเป็นบุคคลเป้าหมายคนเดิม แนวความคิดหลัก คือ การส่งต่อพารามิเตอร์ระหว่างกล้องวิดีโอ PTZ ทั้งสอง ซึ่งเป็นข้อได้เปรียบกว่ากล้องวิดีโอที่ทำงานเพียงตัวเดียว รวมถึงระบบกล้องหลายตัวที่ไม่มีการใช้ข้อมูลร่วมกันระหว่างกล้อง

จากรูปแบบการทำงานทั้งสองแบบที่นำเสนอ แม้ว่าจะอาศัยหลักการของทำงานและประมวลผลร่วมกันระหว่างกล้องวิดีโอ 2 ตัว แต่มีแนวความคิดการประยุกต์ใช้ข้อดีของระบบกล้องหลายตัวในแนวทางที่ต่างกัน ดังนั้นลักษณะรูปแบบการจัดวางกล้อง และสมมติฐานของการทดลองที่จะกล่าวในบทที่ 4 ก็จะต่างกันด้วย ซึ่งจะกล่าวในรายละเอียดของแต่ละรูปแบบการจัดวางในส่วนหัวข้อย่อยถัดไป

### 3.1.1 กล้องวิดีโอ PTZ

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



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

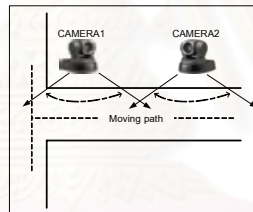
- ในการควบคุมการเคลื่อนที่ของกล้อง ต้องส่งคำสั่ง (command) ควบคุมผ่านทางสายสัญญาณ RS-232
- สามารถซูมเชิงแสง (Optical zoom) ได้สูงสุด 10 เท่า
- แนวการส่ายสูงสุด  $\pm 100$  องศา
- แนวการก้มเงยสูงสุด  $\pm 25$  องศา
- ความเร็วสูงสุดในการเคลื่อนที่แบบส่าย 300 องศา/วินาที
- ความเร็วสูงสุดในการเคลื่อนที่แบบก้มเงย 125 องศา/วินาที
- รองรับการเชื่อมต่อสัญญาณระหว่างกล้องกับส่วนประมวลผลได้ 2 รูปแบบ คือแบบสายสัญญาณภาพรวม (composite video cable) และสายสัญญาณวิดีโอแบบเอสวิดีโอ (S-video cable)

### 3.1.2 ลักษณะโครงสร้างการจัดวางตำแหน่งกล้องวิดีโอ PTZ ทั้ง 2 ตัว

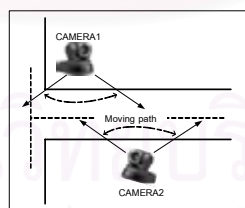
ในหัวข้อย่อยนี้จะอธิบายถึงโครงสร้างการจัดวางกล้องในแต่ละรูปแบบการทำงาน ดังนี้

รูปแบบที่ 1 (วิธีการประมวลผลสำหรับจับคู่ใบหน้าร่วมกัน) เพื่อให้รองรับแนวความคิดดังกล่าวไว้ในตอนต้น การจัดวางกล้องแบบนี้ ต้องเน้นที่มุมมองการรับภาพของทั้งสองกล้องวิดีโอที่ติดตั้งต้องเหลื่อมกันในมุมที่เหมาะสมและเกิดประโยชน์ต่อการรับภาพใบหน้ามาประมวลผลมากที่สุด การจัดวางกล้องวิดีโอที่ติดตั้งสามารถทำได้ 2 รูปแบบ ดังแสดงในรูปที่ 3.2 และ 3.3 แต่ในวิทยานิพนธ์นี้จะเลือกรูปแบบตามรูปที่ 3.2 เนื่องจากเป็นรูปแบบที่กล้องทั้งสองสามารถรับภาพได้ในตำแหน่งที่เหมาะสมต่อการนำไปประมวลผลจับคู่ สมมติฐานของสถานการณ์ในการจัดวางแบบนี้ ได้แก่ บริเวณจุดแสดงผลงานหรือสินค้าที่ต้องการความปลอดภัย บริเวณจุดติดต่อบริษัทสัมพันธ์ เป็นต้น ส่วนการจัดวางดังรูป 3.3 มีข้อบกพร่องเกี่ยวกับมุมมอง เช่น ในกรณีบุคคลหันไปทางกล้องวิดีโอตัวเดียว ทำให้กล้องวิดีโออีกตัวอาจจับภาพได้เป็นส่วนศีรษะด้านหลังและในกรณีถ้าบุคคลเดินโดยไม่ได้หันเข้ากล้องวิดีโอตัวหนึ่ง ทำให้กล้องวิดีโอทั้งสองรับภาพใบหน้าได้เฉพาะด้านข้าง จึงไม่ได้ภาพใบหน้าตรงเลย

ตำแหน่งการวางกล้องวิดีโอในรูปแบบที่ 1 คือ สูงจากพื้น 1.4 เมตร เท่ากันทั้งสองตัว และ กล้องวิดีโอทั้งสองอยู่ห่างกัน 1.5 เมตร



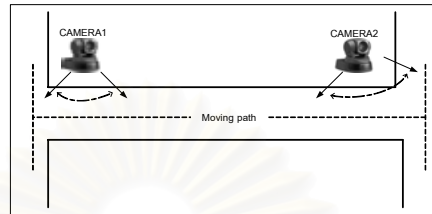
รูปที่ 3.2 การจัดวางกล้องเพื่อการประมวลผลจับคู่ใบหน้าร่วมกันแบบที่หนึ่ง



รูปที่ 3.3 การจัดวางกล้องเพื่อการประมวลผลจับคู่ใบหน้าร่วมกันแบบที่สอง

รูปแบบที่ 2 (วิธีการประมวลผลตรวจหาาร่วมกัน) การจัดวางกล้องวิดีโอที่ติดตั้งจะเป็นดังรูปที่ 3.4 จะเน้นที่กล้องวิดีโอ PTZ ทั้งสองควรจะอยู่ห่างกัน และไม่มีกรเหลื่อมกันในมุมมองการรับภาพ ดังนั้นการจัดวางกล้องจึงไม่ต้องมีการกำหนดในตำแหน่งมาก (แต่เนื่องจากข้อจำกัดในการทดลอง จึงจำเป็นต้องจัดวางในลักษณะดังรูป) อาจจะวางในมุมตรงข้าม หรืออาจจะอยู่ต่างบริเวณกันก็ได้

ตำแหน่งการวางกล้องวิดีโอทัศนในรูปแบบที่ 2 คือ สูงจากพื้น 1.4 เมตร เท่ากันทั้งสองตัว และ กล้องวิดีโอทัศนทั้งสองอยู่ห่างกัน 4.9 เมตร



รูปที่ 3.4 การจัดวางกล้องวิดีโอทัศนเพื่อการประมวลผลตรวจหาาร่วมกัน

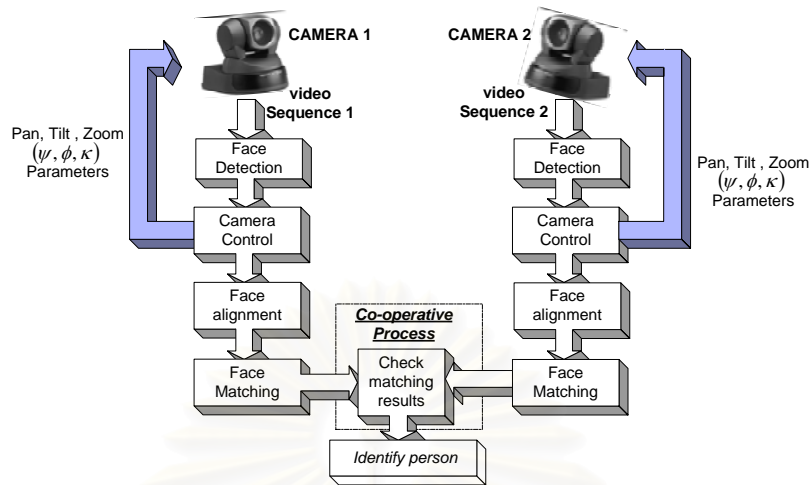
### 3.1.3 ไตอะแกรมการทำงานโดยรวม

**ไตอะแกรมการทำงานของรูปแบบที่ 1** (วิธีการประมวลผลจับคู่ใบหน้าร่วมกัน)

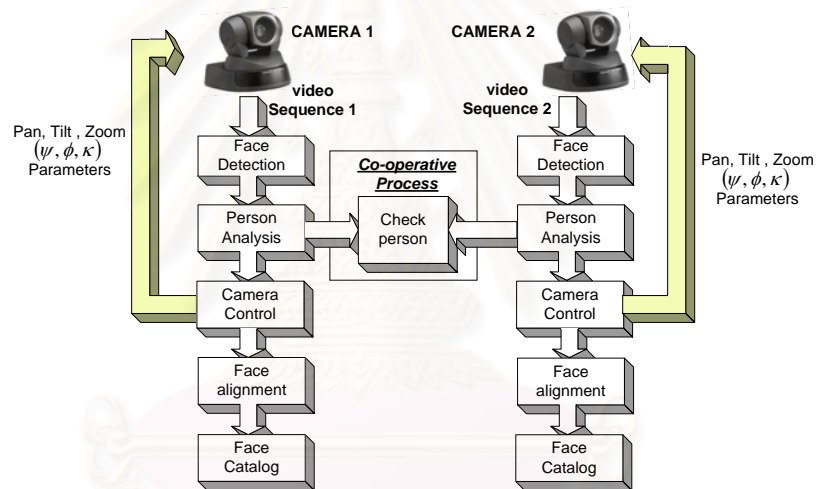
กล้องวิดีโอทัศน PTZ ทั้งสองตัวจะทำงานประมวลผลแบบขนาน (parallel processing) ดังรูปที่ 3.5 กล่าวคือ กล้องวิดีโอทัศนทั้งสองจะประมวลผลในขั้นตอนวิธีที่เหมือนกันและพร้อมกันตลอดเวลา แต่จะมีส่วนการประมวลผลร่วมกันในการวิเคราะห์ผลการจับคู่ใบหน้าที่ได้จากกล้องวิดีโอทัศนแต่ละตัวเพื่อสรุปเป็นผลการจับคู่ใบหน้าโดยรวมของระบบ การประมวลผลของกล้องวิดีโอทัศนแต่ละตัวจะเริ่มจากการตรวจหาใบหน้า ซึ่งจะได้ตำแหน่งจุดศูนย์กลางถ่วงของใบหน้าบุคคลที่สนใจ จากนั้นนำตำแหน่งนี้ไปคำนวณหาค่ามุมในการส่ายและก้มเงยของกล้องวิดีโอทัศน ส่วนภาพบริเวณใบหน้าที่ตรวจหาได้ จะนำไปผ่านกระบวนการหามุมมองใบหน้า (face alignment) เพื่อพิจารณาว่าใบหน้าที่รับได้ ณ ขณะนั้นเป็นใบหน้าที่ด้านตรงหรือเอียง สุดท้ายเป็นกระบวนการจับคู่ใบหน้า

**ไตอะแกรมการทำงานของรูปแบบที่ 2** (วิธีการประมวลผลตรวจหาาร่วมกัน)

การประมวลผลของกล้องวิดีโอทัศน PTZ จะคล้ายกับกับวิธีการที่กล่าวในส่วนไตอะแกรมการทำงานของรูปแบบวิธีการประมวลผลจับคู่ใบหน้าร่วมกัน ดังรูปที่ 3.6 แต่จะมีส่วนของการตรวจสอบบุคคลเพิ่มเข้ามาก่อนที่จะมีการส่ายและก้มเงยกล้องวิดีโอทัศนติดตามบุคคลเป้าหมาย กล่าวคือ ในกรณีที่กล้องวิดีโอทัศนทั้งสองยังตรวจหาไม่พบบุคคล ก็จะทำงานในลักษณะปกติ แต่เมื่อกำหนดวิดีโอทัศนตัวหนึ่งตรวจหาพบบุคคลแล้ว กล้องวิดีโอทัศนอีกตัวจะต้องเปลี่ยนสถานะการประมวลผล โดยต้องตรวจสอบและระบุว่าบุคคลคนเดิมเท่านั้น จึงจะสั่งให้กล้องทำการส่ายและก้มเงยติดตาม ขั้นตอนสุดท้ายของการประมวลผล ก็จะทำการเก็บภาพใบหน้าไว้เป็นฐานข้อมูลสำหรับการใช้งานต่อไป



รูปที่ 3.5 ไดอะแกรมโดยรวมสำหรับวิธีการประมวลผลจับคู่ใบหน้าร่วมกัน



รูปที่ 3.6 ไดอะแกรมโดยรวมสำหรับวิธีการประมวลผลตรวจหาใบหน้าร่วมกัน

### 3.2 ขั้นตอนวิธีในการตรวจหาและติดตามใบหน้า

ในหัวข้อนี้จะอธิบายถึงรายละเอียดขั้นตอนวิธีในการตรวจหาและติดตามใบหน้าด้วยกล้องวิดีโอ PTZ แนวความคิดหลักของการตรวจหาที่นำเสนอ คือ ต้องสามารถประมวลผลตรวจหาได้ในกรณีที่กล้องวิดีโอ PTZ มีการส่ายและก้มเงยเปลี่ยนมุมมองการรับภาพ รวมทั้งแนวการส่ายและก้มเงยของกล้องวิดีโอ PTZ ต้องเป็นไปอย่างอิสระทั้งในสองแนว เมื่อพิจารณาที่ขั้นตอนวิธีที่ผ่านมา [9, 10, 11, 12] ส่วนมากจะต้องมีการเก็บฉากหลังในแต่ละตำแหน่งการส่ายและก้มเงยของกล้อง ทำให้เกิดปัญหาและข้อจำกัด ดังที่กล่าวในบทที่ 2 ดังนั้นวิทยานิพนธ์นี้จึงได้นำเสนอการประยุกต์ใช้งานร่วมกันระหว่างการตรวจหาด้วยสีผิวและการตรวจหาด้วยการเคลื่อนที่



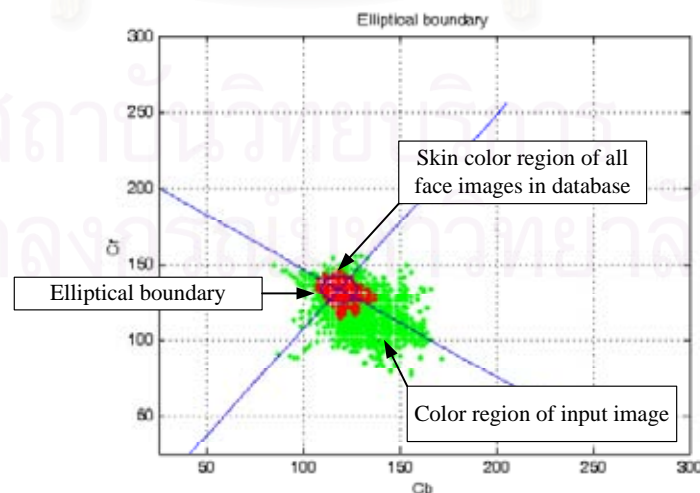
การตรวจหาด้วยสีผิวและการตรวจหาด้วยการเคลื่อนที่สามารถประมวลตรวจหาได้โดยไม่ขึ้นกับฉากหลัง ดังนั้นจึงสามารถประยุกต์ใช้กับกล้องวีดิทัศน์ PTZ ได้อย่างมีประสิทธิภาพ โดยไม่จำกัดตำแหน่งในการส่ายและก้มเงยของกล้องวีดิทัศน์ และสามารถตรวจหาวัตถุที่มีการเปลี่ยนแปลงรูปร่างได้ เหมาะกับการตรวจหาใบหน้าบุคคล

กระบวนการตรวจหาจะเริ่มต้นด้วยการตรวจหาตำแหน่งการเคลื่อนที่ของบุคคลสิ่งที่ได้ คือ คำโครงโดยรวมของบุคคลที่สนใจ จากนั้นใช้กระบวนการตรวจหาสีผิวเพื่อระบุตำแหน่งของใบหน้าที่อยู่ภายในคำโครงบุคคล เหตุผลที่นำการตรวจหาการเคลื่อนที่มาประยุกต์ใช้ร่วมก็เพื่อลดสัญญาณรบกวนที่เกิดจากวัตถุที่มีสีคล้ายสีผิวในกระบวนการตรวจหาสีผิว รวมทั้งลดบริเวณในการตรวจหาสีผิวให้เหลือแค่บริเวณคำโครงบุคคลที่ตรวจหาได้

รายละเอียดในหัวข้อย่อยที่จะกล่าวถึงในหัวข้อนี้ คือ ขั้นตอนวิธีการตรวจหาสีผิวที่ประยุกต์มาจากงานวิจัย [15] กระบวนการตรวจหาการเคลื่อนที่ที่ปรับปรุงจากขั้นตอนวิธีทั่วไปให้เหมาะกับรูปแบบงานที่น่าเสนอ วิธีการกำหนดและหาขอบเขตของใบหน้าเพื่อสามารถนำไปประยุกต์ใช้ในการจับคู่ใบหน้าได้อย่างมีประสิทธิภาพ และวิธีการจำแนกประเภทของมุมมองใบหน้าของภาพที่รับได้

### 3.2.1 กระบวนการตรวจหาด้วยสีผิว

กระบวนการตรวจหาสีผิวที่เลือกใช้จะอ้างอิงขั้นตอนวิธีจาก [15, 16] ดังรายละเอียดที่กล่าวในหัวข้อ 2.3.4 ในบทที่ 2 ซึ่งจะได้ขอบเขตวงรีที่นำมาใช้ในระบบดังรูปที่ 3.7 โดยจะได้กล่าวในรายละเอียดของการหาแบบจำลองวงรีในบทที่ 4 ต่อไป



รูปที่ 3.7 รูปแบบจำลองวงรีที่กำหนดขึ้นจากการฝึกสอน



ปัญหาที่เกิดขึ้นเมื่อใช้แบบจำลองวงรี ในการตรวจหาส่วนสีผิว คือ ปัญหาเกี่ยวกับสิ่งแวดล้อมจากหลังที่มีสีที่คล้ายกับส่วนสีผิวดังรูปที่ 3.8 (ส่วนมากวัตถุที่มีผลกระทบจะมีองค์ประกอบสีหลักเป็นสีแดง) ซึ่งทำให้เกิดความยุ่งยากในการวิเคราะห์บริเวณ ทำให้จำเป็นต้องมีขั้นตอนการเลือกขนาดของแบบจำลองวงรีที่เหมาะสมต่อการตรวจหามากที่สุด รวมทั้งในวิทยานิพนธ์นี้จะมีการประยุกต์ใช้การตรวจหาด้วยลักษณะการเคลื่อนที่ร่วมด้วย ซึ่งจะสามารถลดผลกระทบของปัญหาดังได้กล่าวมาได้อย่างมาก



รูปที่ 3.8 ปัญหาของการตรวจหาสีผิว

### 3.2.2 กระบวนการตรวจหาการเคลื่อนที่ที่น่าเสนอ

กระบวนการเบื้องต้นของการตรวจหาการเคลื่อนที่โดยทั่วไป คือ การหาความแตกต่างระหว่างเฟรม (frame differencing) ซึ่งจะได้ผลการตรวจหาดังรูปที่ 3.10 จะเห็นว่าผลที่ได้จะเป็นบริเวณขอบของการเคลื่อนที่และบริเวณบางส่วนภายในขอบเขตการเคลื่อนที่ที่มีค่าความต่างของความเข้มสีอย่างชัดเจน ซึ่งผลที่ได้จะไม่ได้เป็นโครงร่างโดยรวมของบุคคล ทำให้มีความยุ่งยากในขั้นตอนการวิเคราะห์ผลการตรวจหา จึงได้เสนอวิธีการเติมจุดภาพลงในเส้นภาพตามแนวนอน ซึ่งต่างจากวิธีเติมเต็มลงในบริเวณที่ถูกล้อมรอบ เนื่องจากในขั้นตอนการตรวจหาการเคลื่อนที่ที่ไม่สามารถรับประกันได้ว่า ผลการตรวจหาการเคลื่อนที่ที่เป็นขอบจะมีการเชื่อมต่อกันตลอดโครงร่างบุคคล ดังรูปที่ 3.10(ค) ผลการตรวจหาการเคลื่อนที่ในรูปแบบไปนารีแสดงได้ดังสมการที่ (3.1)

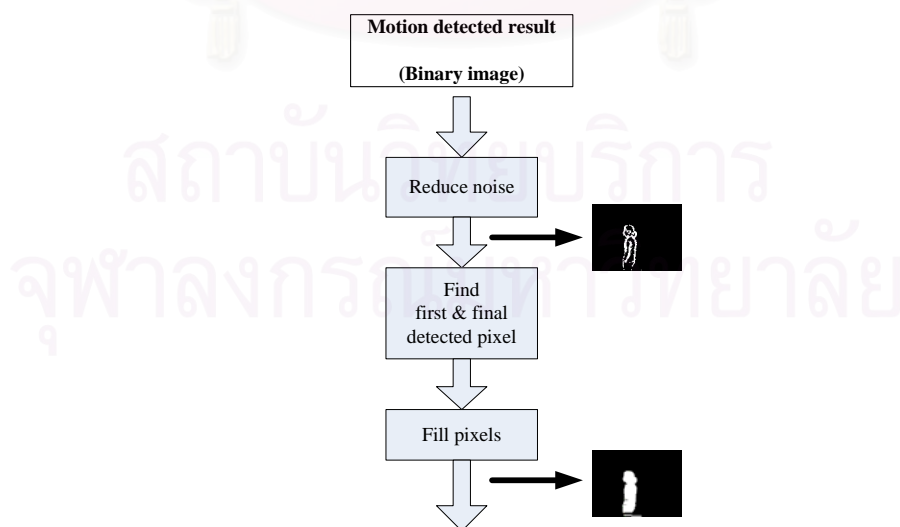
$$x(i, j) = \begin{cases} 1, & \text{motion pixel} \\ 0, & \text{non motion pixel} \end{cases} \quad (3.1)$$

เมื่อ  $x(i, j)$  คือ ผลการตรวจหาในรูปแบบฐานสอง  $i = 0, \dots, \text{Height} - 1$  คือ ตำแหน่งแถวในภาพ  $j = 0, \dots, \text{Width} - 1$  คือ ตำแหน่งสดมภ์ในภาพ

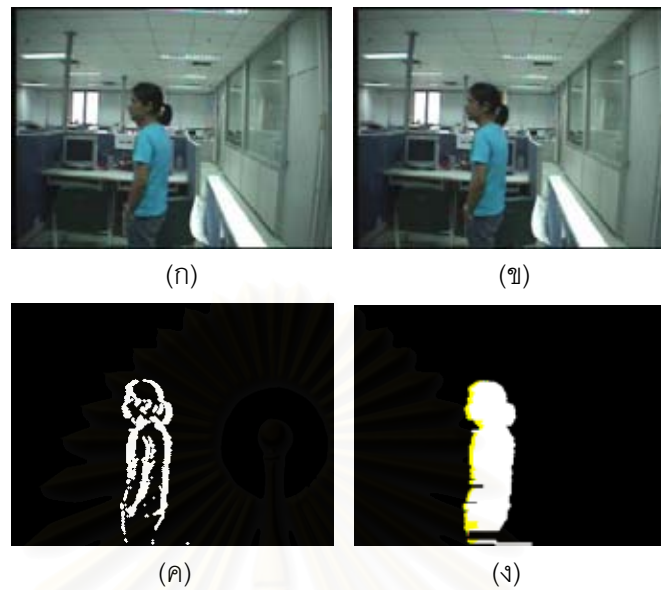
ไดอะแกรมการทำงานแสดงได้ดังรูปที่ 3.9 และรูปที่ 3.10(ง) แสดงตัวอย่างผลการเติมจุดภาพ รายละเอียดของการทำงานจะอธิบายได้ดังนี้

- ภาพฐานสองที่ได้จากการตรวจหาการเคลื่อนที่โดยวิธีทั่วไป จะถูกผ่านขั้นตอนการลดสัญญาณรบกวน (กล่าวถึงในหัวข้อย่อย 3.2.5)
- หาตำแหน่งจุดภาพเริ่มต้นและจุดภาพสุดท้ายในแต่ละแถวของภาพเพื่อใช้เป็นขอบเขตในการเติมจุดภาพ
- ในการหาตำแหน่งจุดภาพเริ่มต้นจะเริ่มพิจารณาตรวจหาจากจุดภาพทางซ้ายสุดของแต่ละแถว จนเมื่อพบจุดภาพตำแหน่งแรกที่ถูกกำหนดว่าเป็นจุดภาพเคลื่อนที่ (motion pixel) จะกำหนดตำแหน่งจุดภาพนี้เป็น ตำแหน่งเริ่มต้นในการเติมจุดภาพ
- ในการหาตำแหน่งจุดภาพสุดท้าย จะเริ่มพิจารณาตรวจหาจากตำแหน่งจุดภาพทางขวาสุดของแต่ละแถว จนเมื่อพบจุดภาพตำแหน่งแรกที่ถูกกำหนดว่าเป็นจุดภาพเคลื่อนที่ (motion pixel) จะกำหนดตำแหน่งจุดภาพนี้เป็น ตำแหน่งสุดท้ายในการเติมจุดภาพ
- ขั้นตอนการเติมจุดภาพจะทำการเปลี่ยนแปลงค่าของภาพฐานสองของจุดภาพที่อยู่ระหว่างตำแหน่งจุดภาพเริ่มต้นและจุดภาพสุดท้ายจากค่า 0 เป็น 1 หรือจาก 1 เป็น 1 ดังสมการที่ (3.2) (เมื่อพิจารณาที่แถวภาพที่  $i$ )

$$x_i(j) = 1 \quad ; \text{start point} < j < \text{stop point} \quad (3.2)$$



รูปที่ 3.9 ไดอะแกรมการทำงานในส่วนการต่อเชื่อมจุดภาพ



รูปที่ 3.10 ผลการตรวจหาการเคลื่อนที่ (ก) เฟรมภาพรับเข้าที่เวลา  $t-1$  (ข) เฟรมภาพรับเข้าที่เวลา  $t$  (ค) ผลการตรวจหาเบื้องต้นในรูปแบบภาพฐานสอง (ง) ผลการตรวจหาเมื่อผ่านกระบวนการเติมจุดภาพ

ปัญหาของการเติมจุดภาพดังที่กล่าวมา คือ ในกรณีที่ภาพฐานสองของผลการตรวจหาเบื้องต้นไม่เป็นไปตามผลในอุดมคติที่ต้องการ เพราะยังเหลือมีสัญญาณรบกวนที่เกิดจากสิ่งแวดล้อม เช่น ผลกระทบของหน้าจอกอมพิวเตอร์ (ดังรูปที่ 3.11) เป็นต้น แม้จะผ่านกระบวนการลดสัญญาณรบกวนแล้วก็ตาม จึงต้องมีกระบวนการตรวจสอบว่าหลังจากผ่านกระบวนการเติมจุดภาพ สามารถนำภาพฐานสองนี้ไปวิเคราะห์ต่อไปได้หรือไม่



รูปที่ 3.11 ผลการเติมจุดภาพที่มีปัญหา (ก) เฟรมภาพรับเข้าที่เวลา  $t-1$  (ข) เฟรมภาพรับเข้าที่เวลา  $t$  (ค) ผลการตรวจหาเบื้องต้นในรูปแบบภาพฐานสอง (ง) ผลการเติมจุดภาพที่ผิดพลาด

ดังที่กล่าวไว้ในเนื้อหาบทที่ 2 จะเห็นว่าสิ่งที่ต้องพิจารณา คือ ช่วงห่างระหว่างเฟรมที่จะนำมาตรวจหาการเคลื่อนที่ด้วยกระบวนการลบหาความแตกต่าง ในวิทยานิพนธ์นี้ เลือกใช้เฟรมที่ติดกันเพื่อนำมาใช้ประมวลผล (อัตราเฟรมของสัญญาณวีดิทัศน์ในการประมวลผล คือ 15 เฟรมต่อวินาที) เพราะจะทำให้ได้ขอบเขตของโครงร่างบุคคลที่ใกล้เคียงกับขอบเขตจริง ซึ่งถ้าเลือกช่วงห่างมากเกินไปจะผลการตรวจหาที่มีลักษณะคล้ายการซ้อนเหลื่อมของบุคคล ทำให้ไม่เหมาะในการประยุกต์ใช้กับขั้นตอนวิธีการเติมจุดภาพที่ได้กล่าวไว้ (ซึ่งจะได้อธิบายและแสดงรูปผลกระทบนี้นในบทที่ 4)

### 3.2.3 กระบวนการตรวจหาโดยใช้การเคลื่อนที่และสีผิว

กระบวนการตรวจหาโดยรวมของระบบ จะนำการตรวจหาทั้งสองส่วนดังที่กล่าวมาในหัวข้อ 3.2.1 และ 3.2.2 มาตรวจหาร่วมกัน โดยผลการตรวจหาจะได้ดังสมการที่ (3.3)

$$x(i, j) = \begin{cases} 1 & ;(\text{motion pixel}) \text{ and } (\text{skin pixel}) \\ 0 & ; \text{otherwise} \end{cases} \quad (3.3)$$

ลำดับการประมวลผล คือ เริ่มต้นใช้กระบวนการตรวจหาการเคลื่อนที่ เพื่อหาโครงร่างของบุคคลที่เกิดจากการเคลื่อนที่ จากนั้นตรวจหาสีผิวในบริเวณโครงร่างดังกล่าว ซึ่งจะช่วยให้ลดจำนวนในการประมวลผลต่อจุดภาพและเพิ่มประสิทธิภาพในการตรวจหา โดยจะลดสัญญาณรบกวนที่เกิดจากวัตถุที่มีสีคล้ายสีผิวมนุษย์ ดังตัวอย่างสัญญาณรบกวนรูปที่ 3.8 และประเด็นสำคัญคือ สามารถนำไปประยุกต์ตรวจหากับสัญญาณวีดิทัศน์ที่รับจากกล้องวีดิทัศน์ PTZ ได้โดยไม่ต้องคำนึงถึงผลกระทบของการเปลี่ยนฉากหลัง รูปที่ 3.12 แสดงผลลัพธ์ในการตรวจหาเมื่อรวมทั้งสองวิธี



รูปที่ 3.12 ตัวอย่างผลการตรวจหาด้วยการตรวจหาการเคลื่อนที่และตรวจหาสีผิว  
(ก) ภาพรับเข้า (ข) ผลการตรวจหา

### 3.2.4 การจำแนกเฟรมภาพในการประมวลผล

ในการประมวลผลภาพเพื่อตรวจหาการเคลื่อนที่จากสัญญาณวิดีโอที่รับภาพจากกล้องวิดีโอ PTZ สิ่งหนึ่งที่ต้องพิจารณา คือ ลักษณะของเฟรมที่ขึ้นอยู่กับเคลื่อนที่ของกล้อง ซึ่งสามารถแบ่งได้เป็น

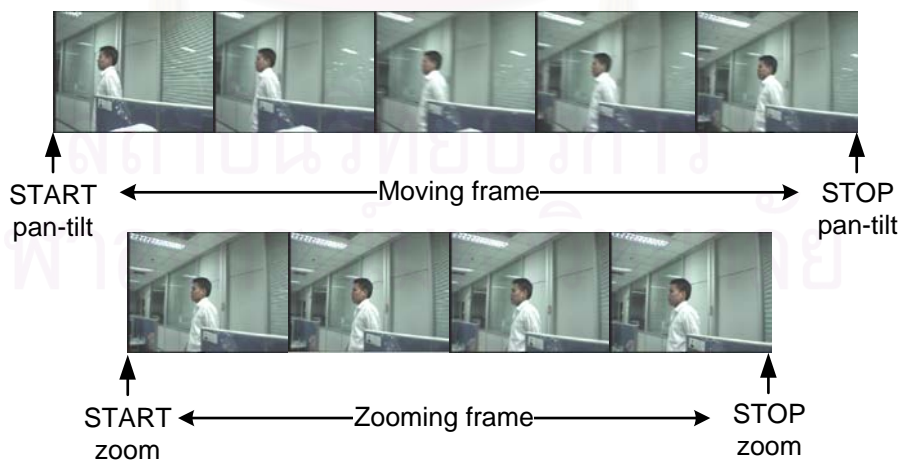
- **เฟรมภาพปกติ (normal frame)** เป็นเฟรมภาพทั่วไป ที่กล้องวิดีโอไม่มีการส่าย ก้มเงย และการซูม เพราะฉะนั้นสามารถประมวลผลตรวจหาใบหน้าได้ตามปกติ

- **เฟรมภาพขณะกล้องวิดีโอ PTZ กำลังเปลี่ยนมุมมองการรับภาพ (เฟรมภาพเคลื่อนที่ (moving frame))** เป็นเฟรมภาพที่ไม่สามารถนำมาวิเคราะห์การตรวจหาได้ เนื่องจากผลกระทบของการส่าย ก้มเงย (แสดงตัวอย่างผลกระทบดังรูปที่ 3.14) และการซูม (แสดงตัวอย่างผลกระทบดังรูปที่ 3.15) ดังนั้นจะไม่มีกรตรวจหาใบหน้าถ้าเป็นเฟรมภาพแบบนี้ รูปที่ 3.13 จะแสดงให้เห็นถึงขั้นตอนการเปลี่ยนแปลงในลำดับภาพวิดีโอตั้งแต่เฟรมปกติ กล้องวิดีโอเริ่มเปลี่ยนมุมมอง จนเฟรมภาพกลับมาอยู่ในสถานะปกติ

ในการจำแนกเฟรมภาพจะตรวจจากจำนวนจุดภาพเคลื่อนที่ทั้งหมด ดังสมการที่ (3.4)

$$\text{if } N_{\text{motion}} \begin{cases} \geq \tau_f & ;= \text{normal frame} \\ < \tau_f & ;= \text{moving frame} \end{cases} \quad (3.4)$$

เมื่อ  $N_{\text{motion}}$  คือ จำนวนจุดภาพทั้งหมดในเฟรมภาพที่ถูกพิจารณาว่าเป็นจุดภาพเคลื่อนที่ และ  $\tau_f$  คือ ค่าขีดเริ่มเปลี่ยนในการแบ่งเฟรมภาพปกติกับเฟรมภาพเคลื่อนที่

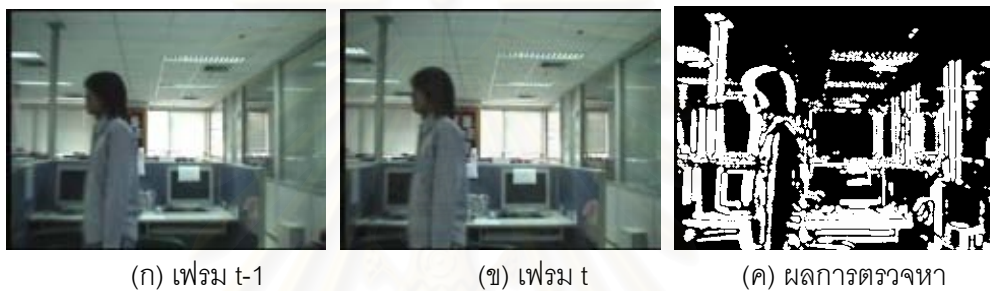


รูปที่ 3.13 ลำดับวิดีโอที่เกิดจากการส่าย ก้มเงย และซูมของกล้องวิดีโอ จากหยุดนิ่ง เคลื่อนที่ และหยุดนิ่ง





รูปที่ 3.14 ผลการตรวจหาการเคลื่อนที่ ณ เฟรมภาพที่กำลังเปลี่ยนมุมมอง  
ภาพในลักษณะสายและกัมเมย



รูปที่ 3.15 ผลการตรวจหาการเคลื่อนที่ ณ เฟรมภาพที่กำลังเปลี่ยนมุมมองในลักษณะชুম

### 3.2.5 กระบวนการลดสัญญาณรบกวน

หลังจากผ่านกระบวนการตรวจหา ทั้งการตรวจหาด้วยสีผิวและการตรวจหาการเคลื่อนที่ แม้สัญญาณรบกวนที่เกิดจากฉากหลังที่มีสีคล้ายกับสีผิว จะถูกตัดทิ้งไป เหลือเพียงแค่บริเวณสีผิวแท้จริง แต่ก็ยังไม่เหมาะสมต่อการนำมาวิเคราะห์ผล เนื่องจากเหตุผล คือ ยังมีสัญญาณรบกวนในระดับจุดภาพเล็กน้อย ซึ่งมีผลต่อการวิเคราะห์บริเวณที่เชื่อมต่อกัน ดังนั้นจึงเลือกใช้ตัวดำเนินการแบบเปิด ในรูปแบบการประมวลผลเชิงสัญญาณ ซึ่งทำกระบวนการอีโรชันแล้วตามด้วยการทำไดเลชันอย่างละหนึ่งรอบ กล่าวคือ หลังจากผ่านกระบวนการ กลุ่มของจุดภาพที่ได้จากการตรวจหาจะมีการเชื่อมต่อกันมากขึ้นและสามารถตัดสัญญาณรบกวน นอกจากนั้นยังช่วยให้การวิเคราะห์บริเวณที่เชื่อมต่อกันมีประสิทธิภาพมากขึ้นด้วย โดยโครงสร้างส่วนย่อยที่เลือกใช้สำหรับกระบวนการไดเลชันและอีโรชัน จะเป็นดังรูปที่ 3.16

0	1	0
1	1	1
0	1	0

รูปที่ 3.16 โครงสร้างส่วนย่อยในการทำอีโรชันและไดเลชัน



### 3.2.6 การจำแนกส่วนบริเวณใบหน้า

ในการจำแนกส่วนบริเวณใบหน้าจากผลการตรวจหา จะพิจารณาที่ผลการตรวจหาด้วยสีผิว ที่ผ่านการลดสัญญาณรบกวนด้วยการตรวจหาด้วยการเคลื่อนที่แล้ว ซึ่งในโครงร่างบุคคล จะประกอบด้วยบริเวณสีผิวที่สามารถตรวจหาได้ 3 ส่วนหลัก คือ ใบหน้า และมือ(รวมแขน) ทั้งสองด้าน ลักษณะที่ใช้แยกส่วนบริเวณใบหน้าที่เลือกใช้ในวิทยานิพนธ์นี้ คือ ลักษณะทางขนาด ตำแหน่ง และรูปร่าง

- ลักษณะขนาด ใช้กรอบบริเวณที่มีขนาดเล็กมากจนไม่มีความน่าจะเป็นที่จะเป็นบริเวณใบหน้า และบริเวณที่มีขนาดใหญ่จนเกินไป
- ลักษณะตำแหน่ง สมมติบุคคลมีรูปแบบการเคลื่อนที่และการแสดงท่าทางที่เป็นปกติ คือ เดินในท่าทางปกติ จึงสมมติให้ตำแหน่งความน่าจะเป็นของบริเวณใบหน้าที่มีตำแหน่งสูงสุดในตัวเลือกรูปร่างทั้งหมด
- ลักษณะรูปร่าง รูปร่างโดยทั่วไปของใบหน้าบุคคล จะค่อนข้างกลม จึงเป็นลักษณะหลักที่นำมาใช้ในการจำแนก

ในกรณีพิจารณาการใส่เสื้อของบุคคลเป้าหมายในการตรวจหา

- สวมเสื้อแขนยาว ลักษณะที่จำเป็นในการแยก คือ ขนาดและตำแหน่ง
- สวมเสื้อแขนสั้น ลักษณะทางรูปร่างจะมีความจำเป็นมาก เนื่องจากเมื่อพิจารณาที่รูปร่าง ท่อนแขน จะมีลักษณะคล้ายทรงกระบอก ซึ่งจะต่างกับใบหน้าที่เป็นทรงกลมอย่างชัดเจน

### 3.2.7 การกำหนดขอบเขตให้กับบริเวณใบหน้า

เป้าหมายหลักของการนำไปประยุกต์ใช้ของระบบในวิทยานิพนธ์นี้ คือ การตรวจหาและติดตามใบหน้าเพื่อนำภาพใบหน้าที่ได้ไปใช้สำหรับการจับคู่ใบหน้า ในสิ่งแวดล้อมจริง ซึ่งมีปัจจัยที่สำคัญที่ควรนำมาพิจารณา 2 ประการ คือ ผลกระทบของฉากหลังในบริเวณใกล้เคียงกับใบหน้าที่จะติดอยู่ในภาพใบหน้า และตำแหน่งมุมมองของใบหน้า (จะกล่าวรายละเอียดในหัวข้อย่อย 3.2.7

โดยทั่วไปงานทางด้านการรู้จำภาพใบหน้า [3, 4, 19] จะใช้การตัดส่วนภาพ (cropping) ใบหน้าเป็นลักษณะภาพสี่เหลี่ยมจากภาพรับเข้า และนำภาพใบหน้าที่ได้นี้ไปใช้สำหรับกระบวนการรู้จำเลย และส่วนใหญ่จะกำหนดลักษณะพื้นฉากหลังเป็นพื้นฉากหลังสีเรียบ ดังนั้นผลกระทบของฉากหลังจึงไม่ส่งผลต่อประสิทธิภาพการรู้จำใบหน้ามากนัก แต่ในวิทยานิพนธ์

นี้ได้ทำการทดลองในห้องปฏิบัติการวิจัยกรรมวิธีสัญญาณดิจิทัลที่ไม่ได้กำหนดลักษณะฉากหลัง รวมทั้งกล้องวีดิทัศน์มีการเปลี่ยนมุมมองการรับภาพเพื่อติดตามบุคคลทำให้ฉากหลังเปลี่ยนแปลงไป ดังนั้นในกรณีระบบที่น่าเสนอ ถ้าใช้การตัดส่วนภาพใบหน้าในลักษณะสีเหลืองเพื่อนำมาใช้ในการจับคู่ใบหน้า จะส่งผลกระทบต่อประสิทธิภาพการจับคู่ใบหน้า เพราะฉากหลังของภาพใบหน้าที่เก็บไว้ในฐานข้อมูลจะต่างจากฉากหลังที่รับภาพได้ในขณะนั้น จึงต้องมีการตัดส่วนภาพใบหน้าในลักษณะทางเรขาคณิตรูปแบบอื่นเพื่อลดผลกระทบของฉากหลังให้น้อยที่สุด เมื่อพิจารณาที่โครงร่างของใบหน้าบุคคลโดยทั่วไป จะเห็นว่ารูปทรงจะคล้ายรูปวงรีหรือวงกลม จึงมีความเหมาะสมที่จะประยุกต์ใช้การตัดส่วนภาพใบหน้าในลักษณะภาพวงรีหรือวงกลมมากที่สุด

ดังนั้นวิทยานิพนธ์นี้จึงได้เลือกลักษณะการตัดส่วนภาพใบหน้าในลักษณะวงรีหรือวงกลม ขั้นตอนการตัดส่วนภาพจะแบ่งออกเป็น 2 ชั้น คือ ชั้นที่ 1 การกำหนดขอบเขตอย่างคร่าวของบริเวณใบหน้าในลักษณะสีเหลือง และชั้นที่ 2 การตัดส่วนภาพในลักษณะวงรีหรือวงกลม พิจารณาจากภาพผลการตรวจหาในอุดมคติในรูปที่ 3.17 เมื่อผลกระบวนการทำป้าย (labeling) จะได้บริเวณสีผิวทั้งหมด 3 กลุ่ม ขั้นตอนแรกในการกำหนดขอบเขตแบบสีเหลือง จะกำหนดได้ดังนี้

**ขอบบน** เลือกจากตำแหน่งสูงสุดของผลการตรวจหา

**ขอบล่าง** เลือกจากตำแหน่งจุดภาพต่ำสุดของบริเวณกลุ่มใบหน้าจากผลการตรวจหาสีผิว

**ขอบซ้าย** เลือกจากตำแหน่งจุดภาพที่อยู่ตำแหน่งซ้ายสุด โดยเลือกเฉพาะเฉพาะภาพระหว่างขอบบนกับขอบล่าง เท่านั้น

**ขอบขวา** เลือกจากตำแหน่งจุดภาพที่อยู่ตำแหน่งขวาสุด โดยเลือกเฉพาะเฉพาะภาพระหว่างขอบบนกับขอบล่าง เท่านั้น



รูปที่ 3.17 ผลการตรวจหาในอุดมคติ



รูปที่ 3.18 ขั้นตอนการกำหนดขอบเขตใบหน้า (ก) กำหนดขอบเขตในลักษณะสีเหลือง (ข) กำหนดขอบเขตในลักษณะวงรี

จะเห็นว่าการกำหนดขอบเขตสีเหลืองเป็นการกำหนดบริเวณอย่างคร่าว ๆ (ดังตัวอย่างรูปที่ 3.18(ก) เพื่อใช้เป็นบริเวณสำหรับการหาขอบเขตของวงรี (ดังตัวอย่างรูปที่ 3.18(ข)) และตำแหน่งที่เหมาะสมที่สุด (position optimization) การหาขนาดของวงรีจะหาได้ดังนี้

**ขนาดแกนเอก** (แกนตั้ง) หาจากขนาดความสูงของขอบเขตสีเหลือง

**ขนาดแกนโท** (แกนนอน) หาจากขอบเขตที่อยู่ทางซ้ายสุดและทางขวาสุดของกลุ่มบริเวณที่ตรวจหาได้ แต่ต้องอยู่ภายในขอบเขตสีเหลืองที่ได้จากขั้นตอนแรก โดยจะตัดแถวที่มีความยาวมากกว่าและน้อยกว่าค่าขีดเริ่มเปลี่ยนในการตัดสินใจที่กำหนดไว้ เมื่อเปรียบเทียบกับขนาดความยาวโดยเฉลี่ย

### 3.2.8 การจำแนกประเภทใบหน้าตามลักษณะมุมมอง (Face alignment)

เมื่อกำหนดขอบเขตสำหรับการตัดส่วนภาพใบหน้าแล้ว สิ่งที่ต้องวิเคราะห์ คือ เป็นภาพใบหน้าในมุมมองใด ซึ่งไม่จำกัดเฉพาะใบหน้าตรงเท่านั้น เหตุผลที่ต้องมีการวิเคราะห์ในหัวข้อนี้ เนื่องจากเป็นระบบสำหรับนำภาพใบหน้าไปใช้จับคู่ใบหน้า มุมมองของใบหน้ามีส่วนสำคัญอย่างมากในประสิทธิภาพของการจับคู่

โดยวิทยานิพนธ์นี้ได้จำแนกประเภทมุมมองของใบหน้าเป็น 3 ประเภท คือ 1. ใบหน้าตรง 2. ใบหน้าด้านข้างซ้าย 3. ใบหน้าด้านข้างขวา ดังรูปที่ 3.19 วิธีการจำแนกจะวิเคราะห์จากผลการตรวจหาสีผิวเทียบกับผลการตรวจหาการเคลื่อนไหวที่ ดังนี้

กำหนดให้กลุ่มสีผิวที่ตรวจหาได้ (บริเวณสีแดงในรูปที่ 3.19) มีตำแหน่งจุดศูนย์กลาง (center of gravity) ที่แถว  $i_s$  และคอลัมน์ที่  $j_s$  และให้บริเวณที่ตรวจหาได้จากการตรวจหาสีผิวและการเคลื่อนไหว (บริเวณสีแดงรวมกับสีเขียว) เฉพาะขอบเขตวงรีบริเวณใบหน้า(ดัง

รายละเอียดในหัวข้อย่อย 3.2.7) มีตำแหน่งจุดศูนย์กลางที่แถว  $i_h$  และคอลัมน์ที่  $j_h$  จะสามารถจำแนกประเภทใบหน้าตามลักษณะมุมมองได้เป็น

- กลุ่มใบหน้าตรง เมื่อ  $|j_s - j_h| < \tau_{front}$
- กลุ่มใบหน้าเอียงด้านซ้ายซ้าย เมื่อ  $j_s - j_h < \tau_{left}$
- กลุ่มใบหน้าเอียงด้านซ้ายขวา เมื่อ  $j_s - j_h < \tau_{right}$

เมื่อ  $\tau_{front}$  คือ ค่าขีดเริ่มเปลี่ยนของการตัดสินใจเป็นลักษณะใบหน้าตรง  $\tau_{left}$  คือ ค่าขีดเริ่มเปลี่ยนของการตัดสินใจเป็นลักษณะใบหน้าเอียงด้านซ้าย(มีค่าติดลบ)  $\tau_{right}$  คือ ค่าขีดเริ่มเปลี่ยนของการตัดสินใจเป็นลักษณะใบหน้าเอียงด้านขวา(มีค่าเป็นบวก)



รูปที่ 3.19 ลักษณะการจำแนกใบหน้าเป็น 3 ประเภท

### 3.3 การควบคุมการส่าย ก้มเงย และหมุนของกล้องวิดีโอ PTZ เพื่อติดตามใบหน้าบุคคลเป้าหมาย

เป็นกระบวนการสำหรับคำนวณหามุมและอัตราการหมุนในการเปลี่ยนมุมมองของกล้องวิดีโอ โดยเป้าหมาย คือ พยายามเปลี่ยนมุมมองการรับภาพของกล้องให้ใบหน้าบุคคลเป้าหมายมาอยู่บริเวณกึ่งกลางภาพ ขั้นตอนการคำนวณพารามิเตอร์ต่าง ๆ จะเลือกใช้เฉพาะข้อมูลตำแหน่งจุดภาพในพิกัด 2 มิติเท่านั้น โดยสมมติว่ามุมในการเปลี่ยนมุมมองแต่ละขั้น (step) ไม่มาก ซึ่งประมาณได้ว่าการคำนวณหามุมด้วยพิกัดตำแหน่งใน 2 มิติ จะมีค่าความผิดพลาดไม่มากเมื่อเปรียบเทียบกับการคำนวณด้วยพิกัดตำแหน่ง 3 มิติ

#### 3.3.1 การคำนวณมุมในการเปลี่ยนมุมมองและอัตราการหมุน

กำหนดให้  $\psi$  คือ มุมส่าย  $\phi$  คือ มุมก้มเงย และ  $\kappa$  คือ ตัวประกอบการหมุน ดังนั้นสามารถหาค่าพารามิเตอร์ทั้งสามได้ ดังนี้

- มุมส่าย (Pan angle)

$$\psi = \frac{Diff_x}{\beta_x} = \frac{x_f - x_{center}}{\beta_x} \quad (3.5)$$

- มุมก้มเงย (Tilt angle)

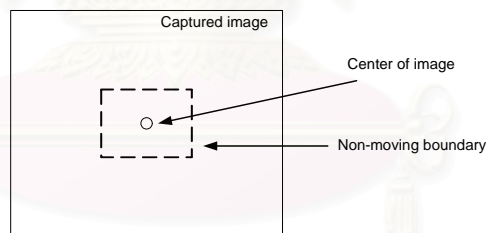
$$\phi = \frac{Diff_y}{\beta_y} = \frac{y_f - y_{center}}{\beta_y} \quad (3.6)$$

- ตัวประกอบการซูม (Zoom factor)

$$\kappa = \frac{Q_{target}}{Q_{detected}} \quad (3.7)$$

เมื่อ  $Diff_x$  และ  $Diff_y$  คือ ระยะห่างในแนวนอนและแนวตั้งระหว่างตำแหน่งจุดศูนย์กลางของใบหน้า  $(x_f, y_f)$  กับตำแหน่งจุดศูนย์กลางภาพ  $(x_{center}, y_{center})$   $\beta_x$  และ  $\beta_y$  คือ อัตราของจำนวนจุดภาพต่อการเคลื่อนที่ของกล้องไป 1 องศา  $Q_{target}$  คือ ขนาดของบริเวณใบหน้าเป้าหมาย และ  $Q_{detected}$  คือ ขนาดของบริเวณใบหน้าที่ตรวจหาได้ (ในหน่วยจุดภาพ)

ในกรณีที่ใบหน้าเป้าหมายอยู่บริเวณใกล้กับกึ่งกลางภาพอยู่แล้ว ไม่จำเป็นต้องสั่งให้กล้องวิดิทัศน์เปลี่ยนมุมมอง ดังนั้นจึงได้กำหนดช่วงขอบเขตในภาพเพื่อใช้เป็นขอบเขตในการเคลื่อนที่ดังรูปที่ 3.20 ถ้าตรวจหาได้ว่าใบหน้าที่ยังอยู่ในช่วงที่กำหนดที่ไม่ต้องมีการเปลี่ยนมุมมอง (Non-moving boundary) ก็จะไม่สั่งกล้องให้เคลื่อนที่ จนกว่าใบหน้าเป้าหมายจะอยู่นอกขอบเขต ก็จะสั่งกล้องให้เคลื่อนที่



รูปที่ 3.20 ขอบเขตการสั่งให้กล้องเคลื่อนที่

### 3.3.2 การขจัดเซชผลกระทบนี้อาจจากการการซูม

ในกระบวนการคำนวณพารามิเตอร์การส่ายและก้มเงย เมื่อมุมมองการรับภาพถูกซูมขยายแล้ว จะเกิดปัญหาขึ้นเมื่อคำนวณค่าตามหัวข้อ 3.3.1 ดังตัวอย่างปัญหาในรูปที่ 3.21 เนื่องจากเมื่อกล้องมีการซูมตำแหน่งของจุดภาพจะถูกสเกลด้วยอัตราส่วนการซูมดังสมการที่ (3.8) ดังนั้นตำแหน่ง  $(x_f, y_f)$  ที่ใช้คำนวณในสมการที่ (3.5) และ (3.6) จะต้องถูกสเกลกลับมายังพิกัดปกติก่อน โดยคำนวณย้อนกลับจากสมการที่ (3.8)

$$X_z = \frac{f'}{f} \begin{bmatrix} x_n \\ y_n \end{bmatrix} \quad (3.8)$$



เมื่อ  $x_c$  คือ ตำแหน่งจุดภาพที่ถูกสเกลด้วยตัวประกอบ  $f'/f(x_n, y_n)$  คือ ตำแหน่งจุดภาพ ณ มุมมองปกติ  $f'$  คือ ความยาวโฟกัสหลังการซูม และ  $f$  คือ ความยาวโฟกัสก่อนการซูม



รูปที่ 3.21 ปัญหาของการเปลี่ยนมุมมองกล้องวิดีโอทัศน์เมื่อมีการซูม

### 3.4 ขั้นตอนวิธีการเปรียบเทียบใบหน้ากับใบหน้ารวมในฐานข้อมูล

#### 3.4.1 กระบวนการวัดความคล้าย

ในกระบวนการวัดความคล้ายจะประยุกต์ใช้การวัดความคล้ายด้วยระยะ Bhattacharyya (ดังได้กล่าวในรายละเอียดหัวข้อย่อย 2.5.1) มาใช้สำหรับเป็นค่าตัดสินในการเปรียบเทียบภาพใบหน้าที่รับภาพและตัดส่วนภาพได้กับภาพใบหน้าที่ถูกเก็บไว้ในฐานข้อมูล เนื่องจากในระบบที่นำเสนอ กล้องวิดีโอทัศน์ PTZ จะเปลี่ยนมุมมองการรับภาพตลอดเวลา ทำให้ภาพใบหน้าที่รับได้มีหลายมุมมอง รวมทั้งมีการสเกลที่เกิดจากการซูมทำให้ขนาดภาพใบหน้าที่มีขนาดเปลี่ยนแปลงอยู่เสมอ ซึ่งเป็นข้อดีของการเปรียบเทียบโดยใช้ระยะ Bhattacharyya เพราะสามารถเปรียบเทียบระหว่างชุดข้อมูลที่มีขนาดไม่เท่ากัน รวมทั้งมีงานวิจัยต่าง ๆ [17, 18, 24] ที่นำระยะนี้ไปประยุกต์ใช้ ในลักษณะการประยุกต์ที่ใกล้เคียงกับระบบที่นำเสนอ

โดยในการเปรียบเทียบภาพใบหน้า จะใช้ข้อมูลค่าความเข้มสีของภาพใบหน้า เฉพาะตำแหน่งจุดภาพที่อยู่ในบริเวณขอบเขตวงรี (หรือวงกลม) เท่านั้น เนื่องจากต้องการลดผลกระทบของฉากหลัง ที่จะส่งผลต่อประสิทธิภาพในการจับคู่ใบหน้าที่ถูกต้อง รายละเอียดของการทดลองของการจับคู่ใบหน้าจะกล่าวในหัวข้อย่อย 4.2.1

### 3.5 การทำงานร่วมกันระหว่างกล้อง

เป็นส่วนการประมวลผลที่เพิ่มขึ้นจากระบบกล้องวิดีโอทัศน์ตัวเดียว ซึ่งเป็นข้อได้เปรียบของระบบกล้องหลายตัว โดยในวิทยานิพนธ์นี้ได้เสนอรูปแบบการทำงานร่วมกันระหว่างกล้องวิดีโอทัศน์ PTZ 2 ตัว ใน 2 รูปแบบ ดังนี้



### 3.5.1 การประมวลผลร่วมกันสำหรับระบุบุคคลโดยการจับคู่ใบหน้า

เป็นการทำงานร่วมกันในลักษณะประมวลผลร่วมกัน เพื่อสรุปผลการจับคู่ใบหน้า โดยใช้ผลการจับคู่ที่ได้จากภาพใบหน้าจากทั้งสองกล้องที่ต้องรับภาพใบหน้าในคนละมุมมอง ซึ่งสมมติฐานเบื้องต้นของมุมมองการรับภาพทั้งสองกล้อง คือ ควรจะมีกล้องวีดิทัศน์ตัวใดตัวหนึ่งที่สามารถจับภาพใบหน้าได้ในตำแหน่งหน้าตรงหรือ ค่อนข้างตรง และเมื่อพิจารณาที่งานวิจัยที่ผ่านมาทางด้านกรู้อจำ [19] ตำแหน่งใบหน้าที่มีประสิทธิภาพในการจับคู่ใบหน้านามากกว่า จะเป็นมุมมองใบหน้าแนวตรง มากกว่าในมุมมองอื่น ดังนั้นจึงจะให้น้ำหนักในการสรุปผลจากกล้องวีดิทัศน์ PTZ ตัวที่สามารถรับภาพใบหน้าด้านตรงมากกว่ากล้องวีดิทัศน์อีกตัวหนึ่ง

กำหนดให้คะแนนการจับคู่ใบหน้าที่ระหว่างภาพใบหน้าที่รับเข้ากับภาพใบหน้าในฐานข้อมูลทั้งหมด  $N$  ภาพใบหน้าจากบุคคลต้นแบบ  $M$  คน ของกล้องแต่ละตัวมีค่าเป็น  $\gamma_i$  เมื่อ  $i=1, \dots, N$  เป็นค่าอันดับความคล้ายของภาพใบหน้าเรียงลำดับจากค่าที่มีความคล้ายมากที่สุดได้ลง โดยค่า  $\gamma_i$  หาได้จากสมการที่ (2.44) ในการพิจารณาผลของการจับคู่ใบหน้าที่จะไม่พิจารณาเฉพาะค่าคะแนนการจับคู่ที่สูงสุดเท่านั้น แต่จะพิจารณาจากอันดับความคล้าย  $P$  ลำดับสูงสุด ซึ่งใน  $P$  ลำดับนี้จะสามารถคำนวณค่าคะแนนการจับคู่ใบหน้าที่ของแต่ละบุคคลต้นแบบได้ดังสมการที่ (3.11)

$$\mu^i = \frac{1}{n^i} \sum_{j=1}^P \gamma_j^i \quad ; i=1, \dots, M \quad (3.11)$$

เมื่อ  $\mu^i$  คือ ค่าเฉลี่ยของคะแนนการจับคู่ใบหน้าที่ที่ระบุเป็นบุคคลต้นแบบ  $i$  ในฐานข้อมูล  $n^i$  คือ จำนวนความถี่ใน  $P$  ครั้งที่จะบระบุว่าเป็นบุคคลต้นแบบ  $i$  ฉะนั้นจะได้คะแนนการจับคู่ใบหน้าที่ในการระบุบุคคลแต่ละบุคคลในฐานข้อมูลต้นแบบ  $ID^i$  ดังสมการที่ 3.12

$$ID^i = \frac{n^i}{P} \mu^i \quad (3.12)$$

การที่ไม่ใช้ค่า  $\mu^i$  ซึ่งเป็นค่าเฉลี่ยของคะแนนโดยตรง ก็เนื่องจากต้องมีการพิจารณาความได้เปรียบในเรื่องความถี่ในการระบุบุคคล เช่น ในกรณีมีบุคคลต้นแบบในฐานข้อมูล 5 คน คือ บุคคล ก บุคคล ข บุคคล ค บุคคล ง บุคคล จ พิจารณาใน 20 อันดับสูงสุดระบุเป็นบุคคล ก 10 ครั้ง บุคคล ข 3 ครั้ง บุคคล ค 1 ครั้ง บุคคล ง 4 ครั้ง และบุคคล จ 2 ครั้ง ดังนั้นจะต้องให้ค่าความได้เปรียบกับบุคคล ก มากที่สุด

โดยทั้งสองกล้องวีดิทัศน์จะประมวลผลในรูปแบบเดียวกันดังที่กล่าวมา ดังนั้นสามารถสรุปผลของการจับคู่ใบหน้าที่เพื่อระบุบุคคลโดยรวมจากทั้งสองกล้องได้ดังสมการที่ 3.13

$$ID_{final}^i = \eta \cdot (ID_1^i) + (1 - \eta) \cdot (ID_2^i) \quad (3.13)$$

เมื่อ  $ID_{final}^i$  คือ คะแนนการจับคู่ใบหน้าที่ระบุเป็นบุคคลต้นแบบที่  $i$  โดยสรุปผลมาจากทั้งสองตัว  $ID_1^i$  และ  $ID_2^i$  คือ คะแนนการจับคู่ใบหน้าที่ระบุเป็นบุคคลต้นแบบที่  $i$  จากกล้องตัวที่ 1 และตัวที่ 2 ตามลำดับ ส่วนค่า  $\eta$  เป็นค่าถ่วงน้ำหนักระหว่างทั้งสองตัว โดย

$\eta = 0.75$  เมื่อกล้องตัวใดตัวหนึ่งสามารถรับภาพใบหน้าได้ในมุมมองหน้าตรง และอีกกล้องหนึ่งเป็นภาพใบหน้ามุมมองด้านข้าง (ให้น้ำหนักกับภาพใบหน้าตรงมากกว่า)

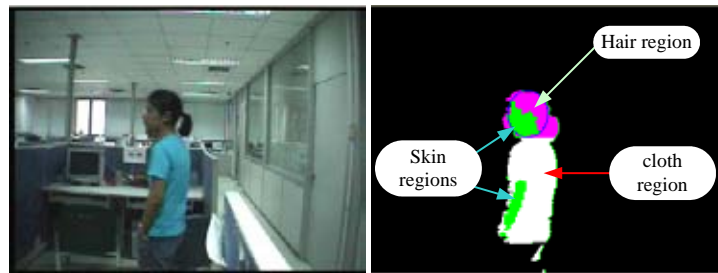
$\eta = 0.50$  เมื่อกล้องทั้งสองตัวรับภาพใบหน้าได้ในมุมมองเดียวกัน อาจจะเป็นหน้าตรงหรือหน้ามุมมองด้านข้างทั้งสอง (ให้น้ำหนักเท่ากัน)

### 3.5.2 การทำงานร่วมกันสำหรับตรวจหาติดตามใบหน้าอย่างต่อเนื่อง

เป็นการทำงานร่วมกัน ในลักษณะการตรวจสอบผลการตรวจหาระหว่างกล้องวีดิทัศน์เพื่อระบุความเป็นบุคคลเดียวกัน ซึ่งมีเป้าหมาย คือ การตรวจหาและติดตามอย่างต่อเนื่องระหว่างกล้องวีดิทัศน์ PTZ 2 ตัว เช่น ในสถานการณ์เมื่อกล้องวีดิทัศน์ตัวแรกสามารถตรวจหาพบบุคคล ก็จะเก็บข้อมูลลักษณะ (feature data) ของบุคคลนี้ไว้ ซึ่งข้อมูลชุดนี้จะถูกส่งไปให้กับกล้องวีดิทัศน์ตัวที่สอง กล้องวีดิทัศน์ตัวที่สองนี้จะนำข้อมูลนี้เพื่อใช้เปรียบเทียบบุคคลที่เข้ามาในมุมมองกล้องว่าเป็นบุคคลเดียวกันหรือไม่ ถ้าใช่ก็จะทำการตรวจหาและทำการส่าย คัมเมย และซูมกล้องวีดิทัศน์เพื่อติดตามบุคคลต่อเนื่อง เพื่อสามารถจับภาพบุคคลที่สนใจได้อย่างต่อเนื่อง จากกล้องวีดิทัศน์ตัวหนึ่งไปยังอีกตัวหนึ่ง โดยในวิทยานิพนธ์นี้เลือกใช้ลักษณะสีและลักษณะทิศทางเคลื่อนที่ของบุคคลเป็นลักษณะในการใช้เปรียบเทียบบุคคล สำหรับลักษณะสีจะแบ่งออกเป็น 3 ส่วนหลักจากผลการตรวจหาได้ดังนี้

#### 3.5.2.1 การแบ่งบริเวณบุคคลในส่วนบริเวณที่ตรวจหา

ข้อมูลลักษณะสีที่เลือกใช้ คือ ฮิสโทแกรมของสีในบริเวณ RGB แล้วใช้การเปรียบเทียบด้วยระยะ Bhattacharyya ดังรายละเอียดที่กล่าวในบทที่ 2 โดยจากผลการตรวจหาที่ได้เป็นโครงร่างบุคคลนั้น ถ้าเลือกใช้เฉพาะข้อมูลสีบริเวณใบหน้า จะทำให้ไม่มีประสิทธิภาพในการระบุบุคคลระหว่างกล้องทั้งสอง ดังนั้นจะเห็นว่าถ้านำข้อมูลสีบริเวณเสื้อผ้าของบุคคลมาใช้ในการเปรียบเทียบด้วย ก็น่าจะเพิ่มความมั่นใจในการเปรียบเทียบ และสมมติฐานเบื้องต้นของบุคคลในวิทยานิพนธ์นี้ คือ มีความแตกต่างของสีเสื้อผ้าระหว่างบุคคล



รูปที่ 3.22 การแบ่งบริเวณโครงร่างบุคคลที่ได้จากการตรวจหา

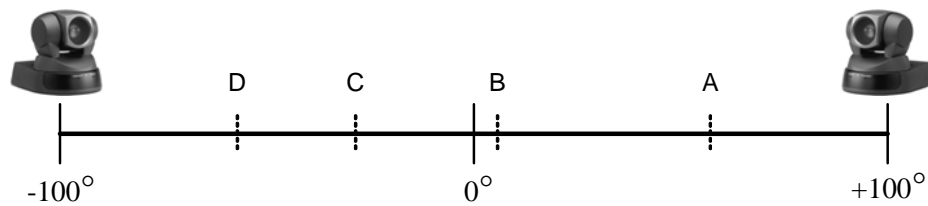
โดยได้แยกบริเวณของโครงร่างบุคคลที่ตรวจหาได้ออกเป็น 3 บริเวณ คือ บริเวณเสื้อผ้า บริเวณสีผิว และบริเวณสีผม ดังรูปที่ 3.22 จะได้ฮิสโทแกรมของทั้งสามส่วนเป็น  $H_{cloths}$ ,  $H_{skin}$  และ  $H_{hair}$  ดังนั้นเมื่อเปรียบเทียบกับระยะ Bhattacharyya จะได้คะแนนของการเปรียบเทียบซึ่งจะมีการกำหนดค่าขีดเริ่มเปลี่ยนเพื่อใช้ตัดสินบุคคลว่าเป็นบุคคลคนเดิมหรือไม่ ดังสมการที่ 3.14

$$Bh_{score} < \tau_c \quad (3.14)$$

โดยที่  $Bh_{score}$  คือ ค่าสัมประสิทธิ์ Bhattacharyya ที่ได้จากการเปรียบเทียบฮิสโทแกรมของแต่ละส่วนบริเวณระหว่างฮิสโทแกรมสีที่ถูกส่งมาจากกล้องวีดิทัศน์ตัวที่ 1 กับฮิสโทแกรมสี ณ ขณะนั้นที่รับได้จากกล้องวีดิทัศน์ตัวที่ 2 และ  $\tau_c$  คือ ค่าขีดเริ่มเปลี่ยนในการตัดสินใจสำหรับระบุบุคคล

### 3.5.2.2 การพิจารณาทิศทางการเคลื่อนที่ของบุคคลที่สนใจ

อีกลักษณะหนึ่งที่จะถูกนำมาพิจารณาระบุบุคคลในวิทยานิพนธ์นี้ คือ ลักษณะทิศทางการเคลื่อนที่ของบุคคลเป้าหมายที่สนใจ เมื่อสมมติฐานการเคลื่อนที่ของบุคคล น่าจะมีเป้าหมายที่แน่นอน เช่น บุคคล ก ต้องการเดินไปยังห้องเรียนตามทางเดินระเบียง แล้วกล้องวีดิทัศน์ที่รับภาพทั้งสองก็วางอยู่ในตำแหน่งต้นและปลายของทางเดิน ดังนั้นสมมติว่ากล้องวีดิทัศน์ตัวที่ 1 ระบุได้ว่า บุคคลมีทิศทางการเคลื่อนที่ไปทางซ้าย ระบบก็จะส่งข้อมูลทิศทางนี้ไปยังกล้องวีดิทัศน์ตัวที่ 2 เพื่อใช้ในการพิจารณาระบุบุคคลเป้าหมาย โดยแบ่งรูปแบบการพิจารณาการเคลื่อนที่ออกเป็น 2 ลักษณะ คือ



รูปที่ 3.23 การพิจารณาทิศทางการเคลื่อนที่ของบุคคลจากตำแหน่งการเปลี่ยนมุมมองของกล้อง

1. สำหรับกล้องวิดีโอที่ตัวแรกที่ตรวจพบบุคคลเป้าหมายก่อน บุคคลแรกที่เดินเข้ามาในระบบตรวจหา ไม่ว่าจะเดินเข้ามายังมุมมองกล้องวิดีโอที่ตัวที่ 1 หรือ ตัวที่ 2 จะถูกกำหนดให้เป็นบุคคลเป้าหมาย และกล้องจะทำการถ่ายและก้มเงยติดตามบุคคลนี้ไปจนสุดขอบเขตที่กล้องวิดีโอจะสามารถเปลี่ยนมุมมองได้ การพิจารณาทิศทางการเคลื่อนที่ที่จะพิจารณาจากแนวทางตำแหน่งการเปลี่ยนมุมมองการรับภาพของกล้องวิดีโอ โดยคำนวณจากการเปลี่ยนตำแหน่งจากตำแหน่งเริ่มต้นของการตรวจหาไปจนถึงตำแหน่งสุดท้าย จากรูปที่ 3.23 กล้องวิดีโอสามารถเปลี่ยนมุมมองในแนวแกนสายได้ระหว่างมุม  $-100$  องศา ถึง  $+100$  องศา สมมติถ้าตำแหน่งเริ่มต้นของการเปลี่ยนมุมมองกล้องวิดีโออยู่ที่ตำแหน่ง A และตำแหน่งการเปลี่ยนมุมมองสุดท้ายของกล้องวิดีโออยู่ที่ D จะพิจารณาว่ามีทิศทางการเคลื่อนที่ไปทางซ้าย (left direction) ของกล้องวิดีโอ ถ้าตำแหน่งเริ่มต้นของการเปลี่ยนมุมมองกล้องวิดีโออยู่ที่ตำแหน่ง C และตำแหน่งการเปลี่ยนมุมมองสุดท้ายของกล้องวิดีโออยู่ที่ B จะพิจารณาว่ามีทิศทางการเคลื่อนที่ไปทางขวา (right direction) ของกล้องวิดีโอ

2. สำหรับกล้องวิดีโอที่สองที่ต้องทำการตรวจสอบบุคคล ในกรณีนี้ กล้องวิดีโอจะยังไม่เคลื่อนที่จนกว่าจะระบุได้ว่าเป็นบุคคลเป้าหมายที่เคลื่อนที่ผ่านกล้องวิดีโอตัวแรกมา ดังนั้นจึงใช้วิธีการคำนวณการติดตามเบื้องต้นได้ คือ พิจารณาจากตำแหน่งวัตถุก่อนหน้านี้กับตำแหน่งวัตถุปัจจุบัน ดังนี้

$$\text{if } \begin{cases} X_t^j - X_{t-1}^j > 0 & ; \text{ right direction} \\ X_t^j - X_{t-1}^j < 0 & ; \text{ left direction} \\ X_t^j - X_{t-1}^j = 0 & ; \text{ no move} \end{cases} \quad (3.15)$$

โดยที่  $X_t^{\text{column}}$  คือตำแหน่งใบหน้าของบุคคลเป้าหมาย ณ ปัจจุบัน  $X_{t-1}^{\text{column}}$  คือตำแหน่งใบหน้าของบุคคลเป้าหมายที่ตรวจหาได้ก่อนหน้านี้ (จะพิจารณาเฉพาะตำแหน่งสดมภ์ (column) ของภาพเท่านั้น โดยสดมภ์ที่ 0 จะอยู่ทางซ้ายสุดของภาพ)

## บทที่ 4

### ผลการจำลองระบบ

ในบทนี้จะเป็นกล่าวถึงรายละเอียดของผลการจำลองระบบ ซึ่งนำเสนอผลการทดลองทั้งในรูปแบบของปรวิสัย (Objective) และ อัจติวิสัย (Subjective) แต่เนื่องจากลักษณะรูปแบบของวิทยานิพนธ์นี้เป็นในลักษณะงานวิจัยเชิงประยุกต์ ทำให้ผลการทดลองส่วนใหญ่ มีการวิเคราะห์ผลในรูปแบบอัจติวิสัย โดยแบ่งการทดลองออกเป็นส่วนใหญ่ 3 ส่วน ตามโครงสร้างของระบบที่นำเสนอในบทที่ 3 คือ ส่วนการตรวจหาใบหน้า ส่วนควบคุมการสายและกัมเยของกล้องวีดิทัศน์ PTZ ติดตามบุคคล และส่วนการทำงานร่วมกันระหว่างกล้องทั้งสองตัว

- ส่วนการตรวจหาใบหน้า จะแบ่งเป็นการทดลองเกี่ยวกับการตรวจหาด้วยสีผิว และการตรวจหาด้วยการเคลื่อนไหวที่ เกี่ยวกับการตรวจหาด้วยสีผิว จะแสดงขั้นตอนวิธีในกระบวนการหาแบบจำลองวงรีที่เหมาะสมที่จะนำมาประยุกต์ใช้กับระบบที่นำเสนอ รวมทั้งปัญหาที่เกิดขึ้น ส่วนเกี่ยวกับการตรวจหาด้วยการเคลื่อนไหวที่ ประกอบด้วย การทดลองเกี่ยวกับขั้นตอนในการประยุกต์การตรวจหาการเคลื่อนไหวที่เบื้องต้น ให้สามารถประยุกต์ใช้กับระบบได้อย่างมีประสิทธิภาพและเหมาะสมมากที่สุด จากนั้นจะเป็นการทดลองเมื่อใช้การตรวจหาสีผิวร่วมกับการตรวจหาการเคลื่อนไหวที่ จะแสดงถึงประสิทธิภาพของการตรวจหาส่วนสีผิวที่มากขึ้นเมื่อใช้การตรวจหาในลักษณะที่นำเสนอเมื่อเทียบกับการตรวจหาด้วยสีผิวเพียงอย่างเดียว ในส่วนสุดท้ายจะกล่าวถึงขั้นตอนวิธีในการกำหนดขอบเขตเพื่อตัดส่วนภาพใบหน้า ให้เหมาะสมต่อการนำไปประยุกต์สำหรับรู้จำใบหน้า

- ส่วนควบคุมการสายและกัมเยของกล้องวีดิทัศน์ PTZ ติดตามบุคคล จะทำการทดลองเพื่อแสดงถึงผลของการควบคุมกล้องวีดิทัศน์ PTZ สาย กัมเย และซูมติดตามบุคคล โดยให้ความสนใจที่ใบหน้าของบุคคล ซึ่งผลการทดลองยังแสดงถึงประสิทธิภาพในการประยุกต์ใช้การตรวจหาดังที่กล่าวในส่วนแรก มาประยุกต์ใช้กับกล้องวีดิทัศน์ PTZ โดยเป็นการเคลื่อนที่อย่างอิสระ ไม่มีการกำหนดตำแหน่งของการสายและกัมเยของกล้องวีดิทัศน์ที่แน่นอน

- ส่วนการทำงานร่วมกันระหว่างกล้องวีดิทัศน์ทั้งสองตัว ในการทดลองแรก จะแสดงถึงการประมวลผลร่วมกันในการจับคู่ใบหน้าบุคคล โดยมีแนวความคิดเพื่อทดลองให้เห็นว่าการที่สามารถรับภาพใบหน้าบุคคลจาก 2 มุมมองกล้องในตำแหน่งที่เหมาะสม จะช่วยเพิ่มประสิทธิภาพของความถูกต้องในการจับคู่ใบหน้ามากขึ้นกว่าการจับคู่ใบหน้าโดยใช้ภาพใบหน้าจากกล้องแต่ละตัว ซึ่งวิทยานิพนธ์นี้ได้เลือกใช้การจับคู่ใบหน้าสำหรับการรู้จำใบหน้า โดยการวัดความคล้ายด้วยระยะ Bhattacharyya ซึ่งสามารถประยุกต์ใช้กับระบบที่ประมวลผลในเวลาจริง



และข้อดีของการจับคู่ด้วยระยะ Bhattacharyya คือ สามารถจับคู่ระหว่างกลุ่มบริเวณที่มีขนาดต่างกันได้อย่างมีประสิทธิภาพ นอกจากนี้เหตุผลในการเลือกใช้การจับคู่ในลักษณะนี้ เพราะสามารถนำไปประยุกต์ใช้กับรูปแบบงานอื่นได้ง่าย ดังเช่นในการทดลองที่สอง ซึ่งเป็นส่วนการทำงานร่วมกันระหว่างกล้องวีดิทัศน์เพื่อตรวจหาและเคลื่อนที่ติดตามบุคคลอย่างต่อเนื่อง การทำงานจะใช้การส่งข้อมูลระหว่างกล้องวีดิทัศน์เพื่อระบุว่าเป็นบุคคลเดียวกัน โดยใช้ลักษณะสีและการเคลื่อนที่มาเป็นตัวตัดสินใจ

#### 4.1 ผลการจำลองระบบในส่วนของการตรวจหาใบหน้า

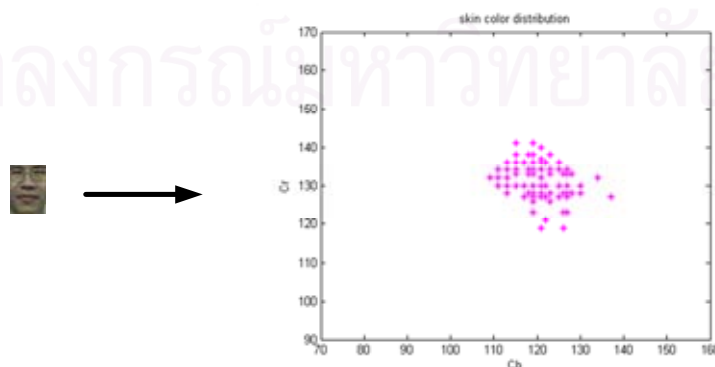
##### 4.1.1 ผลการตรวจหาด้วยสีผิว

###### 4.1.1.1 การหาแบบจำลองวงรี

ในการหาแบบจำลองวงรีที่นำมาใช้เป็นขอบเขตของการจำแนกส่วนบริเวณสีผิวใบหน้า จะใช้ภาพบริเวณใบหน้าจาก 20 บุคคลตัวอย่างเพื่อนำมาหาขอบเขตของวงรีที่เหมาะสมกับการตรวจหาด้วยสีผิวมากที่สุด โดยภาพใบหน้าที่นำมาใช้ในการหาขอบเขตแบบจำลองวงรีสามารถแสดงได้ดังรูปที่ 4.1 ซึ่งจะต้องเป็นภาพใบหน้าที่ถูกตัดส่วนภาพ (cropping) มาจากเฟรมภาพที่รับมาจากกล้องวีดิทัศน์ SONY รุ่น EVID-100 ที่ถูกใช้ในการจำลองระบบ เพื่อให้แบบจำลองที่คำนวณได้มีประสิทธิภาพในการตรวจหากับกล้องวีดิทัศน์ที่ใช้มากที่สุด

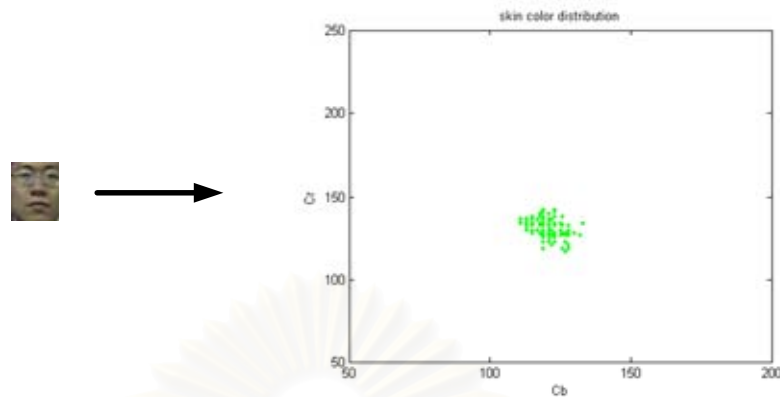


รูปที่ 4.1 ภาพใบหน้าที่นำมาใช้หาขอบเขตแบบจำลองวงรี



รูปที่ 4.2 ตัวอย่างการกระจายตัวของค่าองค์ประกอบสีในปริภูมิย่อย CbCr ของภาพใบหน้า





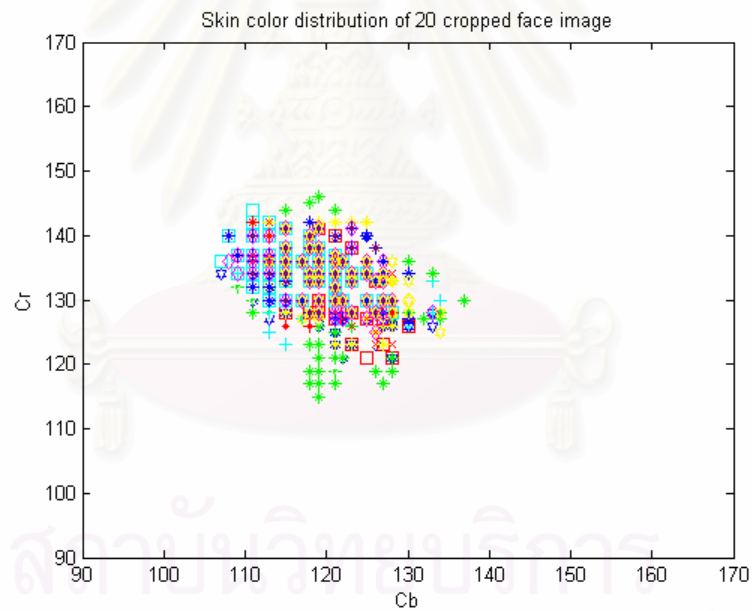
รูปที่ 4.2(ต่อ) ตัวอย่างการกระจายตัวของค่าองค์ประกอบสีในปริภูมิย่อย CbCr ของภาพใบหน้า

จากรูปที่ 4.2 เป็นรูปตัวอย่างของการกระจายตัวของค่าจุดภาพในปริภูมิย่อย CbCr โดยวาดจากภาพใบหน้าที่ถูกตัดส่วนภาพด้วยมือ เพื่อให้สำหรับการกำหนดขอบเขตแบบจำลองวงรีโดยเฉพาะ โดยตัวอย่างการกระจายตัวของค่าจุดภาพทั้งสองภาพ จะวาดในสเกลที่ต่างกันเพื่อให้เห็นถึงรูปร่างของการกระจายตัวที่มีลักษณะคล้ายวงรีเฉียง

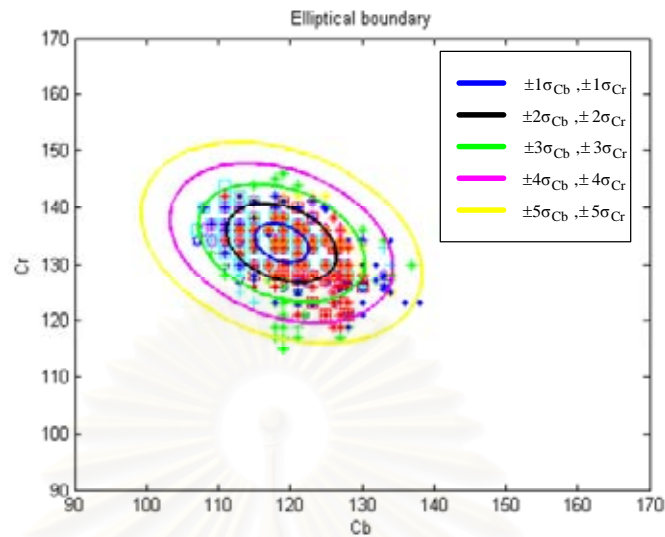
พารามิเตอร์หนึ่งที่สำคัญในกระบวนการหาขอบเขตของแบบจำลองวงรี คือ ค่าตัดสินสำหรับเลือกขนาดของวงรี ( $\eta$ ) ดังสมการที่ 2.26 - 2.27 ซึ่งมีผลโดยตรงต่อบริเวณสีผิวที่ตรวจหาได้ รวมทั้งสัญญาณรบกวนของวัตถุที่มีสีเหมือนสีผิวด้วย โดยเป็นการกำหนดขอบเขตของข้อมูลที่ต้องการในการระบุจุดภาพสีผิวจากพื้นที่ได้กราฟการแจกแจงข้อมูลปกติ โดยกำหนดขอบเขตเป็นจำนวนเท่าของค่าเบี่ยงเบนมาตรฐานของกลุ่มข้อมูล (แบ่งเป็นกลุ่มข้อมูลองค์ประกอบ Cb และ Cr) ซึ่งจะเลือกใช้ค่าตัดสิน  $\eta$  มีค่า 1 2 3 4 และ 5 เท่าของค่าเบี่ยงเบนมาตรฐานในการทดลอง (พื้นที่ได้กราฟที่ได้ของแต่ละค่าตัดสินใจคิดเป็นร้อยละ แสดงในตารางที่ 2.2) รูปที่ 4.3 แสดงผลการวาดการกระจายตัวของค่าองค์ประกอบสี CbCr แต่ละจุดภาพของภาพใบหน้าทั้งหมดที่ถูกนำมาใช้ทั้ง 20 ภาพ ในกราฟเดียวกัน ขั้นตอนในการเลือกแบบจำลองวงรีที่เหมาะสม จะพิจารณาจากค่าอัตราการตรวจหา (Detection Rate : DR) และค่าอัตราการฟ้องค่าผิดพลาด (False Alarm Rate : FAR) และค่าอัตราการปฏิเสธความผิดพลาด (False Reject Rate : FRR) โดยเปรียบเทียบที่ค่า  $\eta = 1, 2, 3, 4, 5$  ซึ่งสามารถแสดงค่าของขนาดวงรีดังตารางที่ 4.1 และรูปที่ 4.4 แสดงถึงขนาดของวงรีที่ได้จากค่าตัดสินแต่ละค่าดังที่กล่าวมา และทดลองหาประสิทธิภาพของแบบจำลองวงรีแต่ละอันจากภาพตัวอย่างรับเข้ารูปที่ 4.5 นำมาหาค่าอัตราการตรวจหา ค่าอัตราการฟ้องค่าผิดพลาด และค่าอัตราการปฏิเสธความผิดพลาด ได้ดังตารางที่ 4.2

ตารางที่ 4.1 ขนาดของวงรีของค่าตัดสีผิว  $\eta = 1, 2, 3, 4, 5$

ค่าตัดสีผิว( $\eta$ ) (จำนวนเท่าของค่า เบี่ยงเบนมาตรฐาน)	ขนาดของแกนเอก (a)	ขนาดของแกนโท (b)
1	7.14	3.72
2	14.35	7.49
3	21.64	11.31
4	29.00	15.19
5	36.43	19.11



รูปที่ 4.3 รูปการกระจายตัวของค่าจุดภาพใบหน้าสีผิวในปริภูมิค่าองค์ประกอบสี CbCr ของทั้ง 20 ภาพใบหน้า



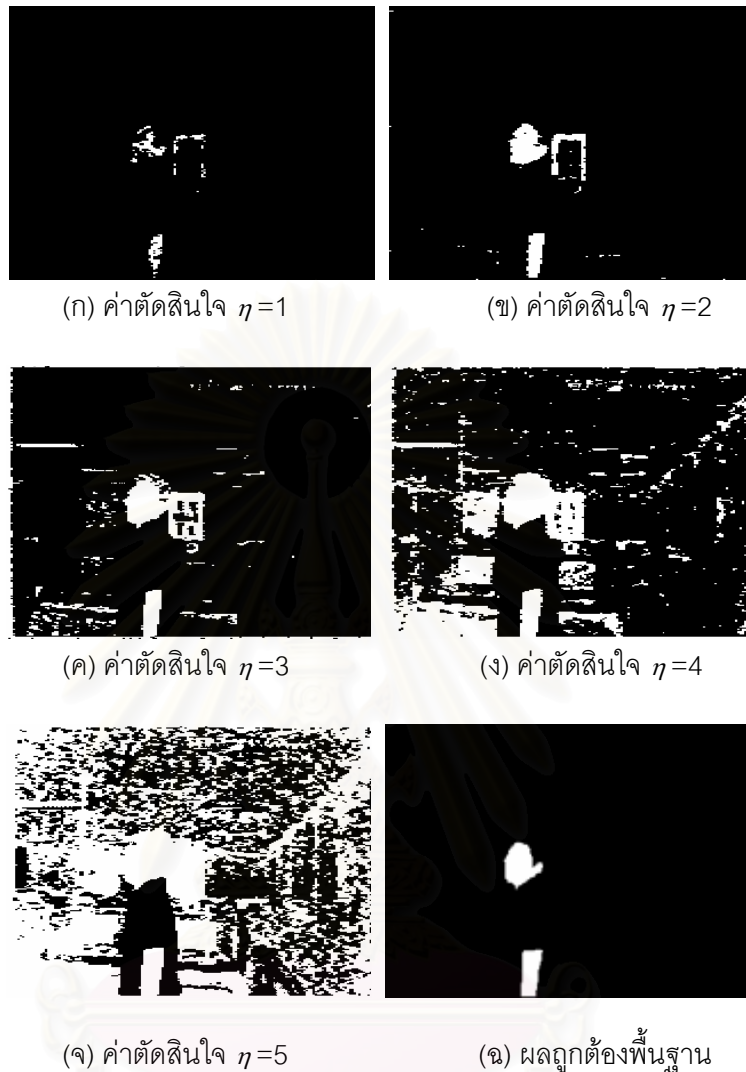
รูปที่ 4.4 แบบจำลองวงรีที่ค่าตัดสินใจ  $\eta = 1, 2, 3, 4$  และ 5

โดยที่  $\sigma_{Cb}$  และ  $\sigma_{Cr}$  คือ ค่าเบี่ยงเบนมาตรฐานของกลุ่มข้อมูลองค์ประกอบ  $C_b$  และองค์ประกอบ  $C_r$  ตามลำดับ

ตารางที่ 4.2 ประสิทธิภาพของผลการตรวจสอบสี่มิติด้วยแบบจำลองวงรีเทียบกับผลถูกต้องพื้นฐาน

แบบจำลองวงรีที่ค่าตัดสินใจ	DR	FAR	FRR
$\eta=1$	0.3371	0.2348	0.6629
$\eta=2$	0.8332	0.3303	0.1668
$\eta=3$	0.9444	0.7913	0.0556
$\eta=4$	0.9944	0.9404	0.0056
$\eta=5$	0.9986	0.9672	0.0014

จากตารางที่ 4.2 จะเห็นว่าค่าตัดสินใจในการเลือกแบบจำลองวงรีที่เหมาะสมน่าจะอยู่ที่ค่าตัดสินใจ  $\eta=2$  หรือ  $\eta=3$  เพราะถ้าที่ค่า  $\eta = 1$  อัตราการตรวจหาจะต่ำมากแม้ว่าอัตราการฟองค่าผิดพลาดจะมีค่าต่ำก็ตาม ส่วนที่ค่าตัดสินใจ  $\eta = 4$  และ 5 อัตราการฟองค่าผิดพลาดจะสูงมาก แม้ว่าจะสามารถตรวจหาด้วยอัตราการตรวจหาที่ค่าเข้าใกล้ 1 หรือแสดงว่าสามารถตรวจหาได้เกือบทั้งหมดก็ตาม

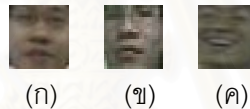


รูปที่ 4.5 ผลการตรวจหาด้วยวิธีที่ใช้แบบจำลองวงรีที่ค่าตัดสินใจต่างกัน

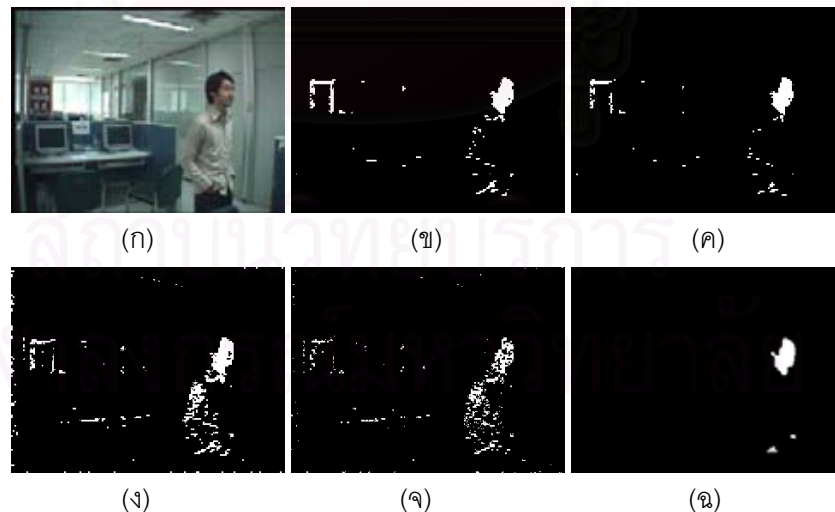
ในการหาแบบจำลองวงรีในลักษณะนี้ อาจจะส่งผลกระทบต่อประสิทธิภาพในเรื่องของอัตราการตรวจหา (Detection rate) เมื่อเทียบกับการสร้างแบบจำลองวงรีโดยใช้ภาพใบหน้าของแต่ละบุคคลตรวจหาเฉพาะบุคคลนั้น เนื่องจากต้องหาแบบจำลองที่สามารถเข้ากันได้กับทุก ๆ การกระจายตัวขององค์ประกอบสีของแต่ละใบหน้า แต่เนื่องจากการจำลองระบบ ไม่ได้กำหนดบุคคลในการทดลองที่แน่นอน ถ้าต้องมีการเปลี่ยนแบบจำลองวงรีทุกครั้งตามแต่ละบุคคล ก็จะเป็นการยุ่งยาก รวมทั้งทำให้ไม่สามารถนำไปประยุกต์ใช้งานจริงได้ การทดลองในส่วนถัดไป จะแสดงถึงผลกระทบดังที่กล่าวมา

#### 4.1.1.2 การเปรียบเทียบความแตกต่างของแบบจำลองวงรี

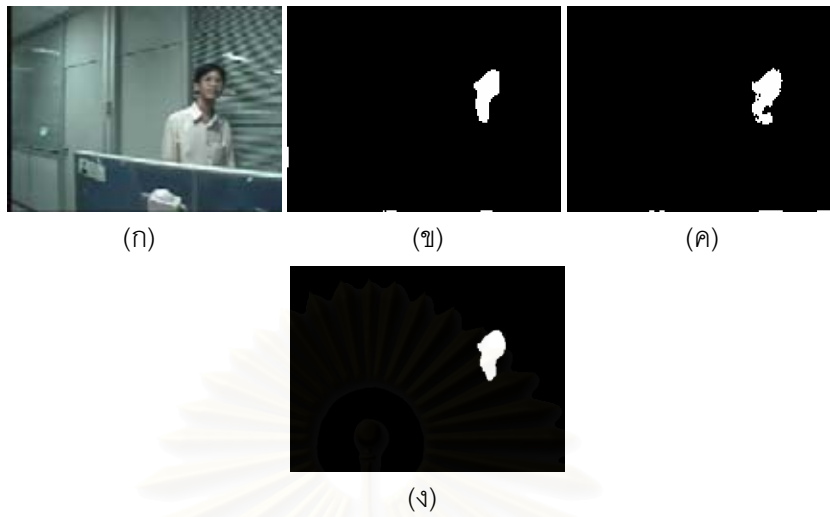
ในหัวข้อย่อยนี้ จะนำเสนอการทดลองเกี่ยวกับรูปแบบการหาแบบจำลองวงรีที่นำมาใช้ตรวจหาสีผิว โดยเปรียบเทียบข้อแตกต่างระหว่าง แบบที่ 1 การหาแบบจำลองวงรีด้วยภาพใบหน้าเฉพาะแต่ละบุคคล ซึ่งจะต้องมีการหาแบบจำลองของบุคคลนั้นก่อนการทดลองก่อน และแบบที่ 2 การหาแบบจำลองวงรีด้วยกลุ่มภาพใบหน้าฐานข้อมูลที่ได้จากในหัวข้อย่อย 4.1.1.1 โดยทดสอบกับบุคคลตัวอย่าง 3 คน รูปที่ 4.6 แสดงภาพใบหน้าบุคคลทั้งสามที่ใช้สำหรับหาขอบเขตแบบจำลองวงรีของแต่ละคน โดยพิจารณาที่ค่าตัดสีผิวเดียวกัน คือ  $\eta = 2$  ในการวัดผลการตรวจหา จะนำแบบจำลองทั้งสองแบบมาตรวจหาโดยใช้ภาพรับเข้าเดียวกัน เปรียบเทียบระหว่างการใช้แบบจำลองวงรีจากที่ได้ในหัวข้อย่อย 4.1.1.1 กับการใช้แบบจำลองวงรีของบุคคลนั้น และเปรียบเทียบโดยใช้แบบจำลองวงรีของบุคคล ข และ ค นำไปตรวจหากับภาพบุคคล ก รูปที่ 4.7-4.9 แสดงถึงผลการตรวจหาของแต่ละรูปแบบ และตารางที่ 4.2 แสดงผลการเปรียบเทียบวัดเป็นค่าอัตราการตรวจหาและอัตราการฟุ้งค่าผิดพลาด



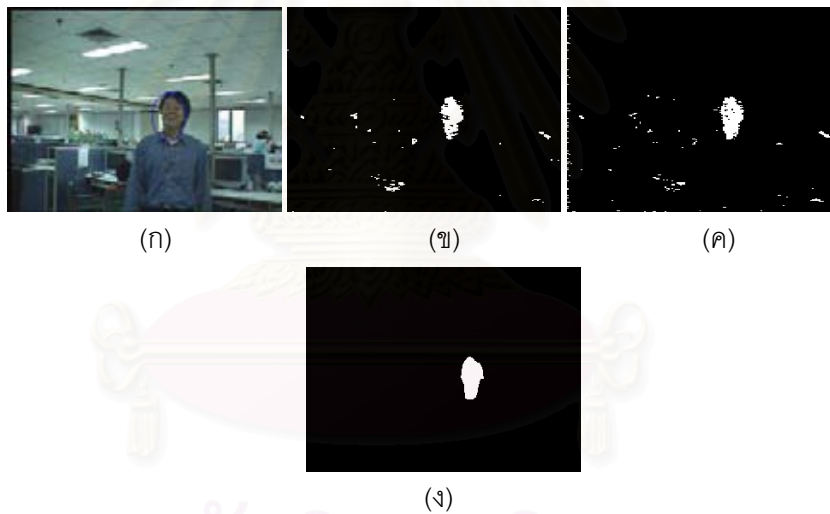
รูปที่ 4.6 การหาแบบจำลองวงรีด้วยภาพใบหน้าเฉพาะบุคคล (ก) ตัดส่วนภาพใบหน้าบุคคล ก (ข) ตัดส่วนภาพใบหน้าบุคคล ข (ค) ตัดส่วนภาพใบหน้าบุคคล ค



รูปที่ 4.7 เปรียบเทียบผลการตรวจหาสีผิวของบุคคล ก (ก) ภาพรับเข้า เมื่อใช้แบบจำลองวงรีของ(ข) แบบจำลองวงรีรวม (ค) แบบจำลองวงรีเฉพาะ (ง) แบบจำลองวงรีของบุคคล ข (จ) แบบจำลองวงรีของบุคคล ค (ฉ) ผลถูกต้องพื้นฐาน



รูปที่ 4.8 เปรียบเทียบผลการตรวจหาสี่ผิวนของบุคคล ข (ก) ภาพรับเข้า เมื่อใช้แบบจำลองวงรี(ข) แบบจำลองวงรีรวม (ค) แบบจำลองวงรีเฉพาะ (ง) ผลถูกต้องพื้นฐาน



รูปที่ 4.9 เปรียบเทียบผลการตรวจหาสี่ผิวนของบุคคล ค (ก) ภาพรับเข้า เมื่อใช้แบบจำลองวงรี(ข) แบบจำลองวงรีรวม (ค) แบบจำลองวงรีเฉพาะ (ง) ผลถูกต้องพื้นฐาน

ตารางที่ 4.3 เปรียบเทียบประสิทธิภาพการตรวจหาของแบบจำลองวงรี

	เมื่อใช้แบบจำลองวงรีโดยรวม			เมื่อใช้แบบจำลองเฉพาะ		
	DR	FAR	FRR	DR	FAR	FRR
ภาพบุคคล ก	0.8745	0.5479	0.1255	0.8431	0.4571	0.1569
ภาพบุคคล ข	0.8740	0.1767	0.1260	0.8825	0.3090	0.1175
ภาพบุคคล ค	0.8249	0.3109	0.1751	0.8987	0.3814	0.1013



จากตารางที่ 4.3 จะเห็นว่าค่าอัตราการตรวจหาและอัตราการฟุ้งค่าผิดพลาดของแต่ละภาพบุคคลเมื่อใช้แบบจำลองวงรีเฉพาะบุคคลและแบบจำลองวงรีรวมจะมีค่าค่อนข้างใกล้เคียงกัน โดยเฉพาะค่าอัตราการตรวจหาที่ใกล้เคียงกัน ส่วนค่าความต่างระหว่างอัตราการฟุ้งค่าผิดพลาด จะขึ้นอยู่กับผลกระทบที่เกิดจากการตัดส่วนภาพใบหน้าเพื่อมาใช้ในการสร้างแบบจำลองด้วย และเมื่อใช้การเปรียบเทียบแบบอัตโนมัติจากภาพผลการตรวจหา ก็จะเห็นว่าผลการตรวจหาไม่ต่างกัน และข้อดีของการใช้แบบจำลองวงรีแบบโดยรวม คือ สามารถประยุกต์ใช้กับการตรวจหาบุคคลได้กับทุกบุคคลภายในฐานข้อมูลที่ใช้ในการหาแบบจำลอง หรืออาจจะรวมถึงบุคคลอื่นที่ไม่ได้อยู่ในฐานข้อมูลด้วย นอกจากนี้ทดลองเปรียบเทียบระหว่างแบบจำลองโดยรวมกับแบบจำลองเฉพาะบุคคลแล้ว ได้ทดลองเพื่อพิจารณาว่าถ้านำแบบจำลองเฉพาะบุคคลไปใช้กับบุคคลอื่นจะมีผลเป็นดังนี้

- ตรวจหาบุคคล ก ด้วยแบบจำลองวงรีของบุคคล ข ได้ ผลการตรวจหาดังรูปที่ 4.7(จ) ได้ผลการตรวจหา คือ อัตราการตรวจหา 0.5738 และอัตราการฟุ้งค่าผิดพลาด 0.7256

- ตรวจหาบุคคล ก ด้วยแบบจำลองวงรีของบุคคล ค ได้ ผลการตรวจหาดังรูปที่ 4.7(ง) ได้ผลการตรวจหา คือ อัตราการตรวจหา 0.8535 และอัตราการฟุ้งค่าผิดพลาด 0.7255

จะเห็นว่าทั้งสองรูปแบบจะมีค่าอัตราการฟุ้งค่าผิดพลาดค่อนข้างสูงมาก ส่วนอัตราการตรวจหา ก็จะขึ้นอยู่กับสีผิวของบุคคลนั้นว่าใกล้เคียงกับของแบบจำลองนั้นๆ หรือไม่ ดังผลการทดลองจะเห็นว่า การตรวจหาบุคคล ก ด้วยแบบจำลองวงรีของบุคคล ข มีอัตราการตรวจหาที่ต่ำ แต่การตรวจหาบุคคล ก ด้วยแบบจำลองวงรีของบุคคล ค มีอัตราตรวจหาที่สูง ทำให้วิธีการตรวจหาด้วยแบบจำลองเฉพาะบุคคลไม่เหมาะที่นำมาใช้ในการตรวจหาบุคคลทั่วไป

#### 4.1.1.3 ผลการตรวจหาด้วยแบบจำลองวงรีที่กำหนดและปัญหาที่เกิดขึ้นและการวิเคราะห์ผล

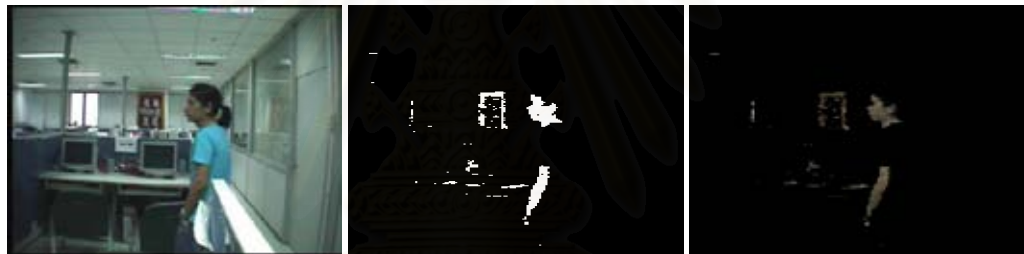
จากแบบจำลองที่เลือกใช้จากหัวข้อย่อยที่ผ่านมา นำมาทดลองตรวจหาบริเวณสีผิวของบุคคล (ใบหน้าและมือทั้งสองข้าง) ได้ผลการตรวจหาดังรูปที่ 4.10 ซึ่งจากผลการทดลองจะเห็นว่าสัญญาณรบกวนที่เกิดขึ้น ถ้ามีขนาดเล็กจะสามารถถูกลดผลได้ด้วยกระบวนการลดสัญญาณรบกวนดังที่กล่าวไว้ในบทที่ 2 แต่ถ้าเป็นกลุ่มบริเวณที่ไม่สามารถลดผลกระทบได้ จะเป็นปัญหาที่เกิดขึ้น ซึ่งมีผลต่อขั้นตอนในการวิเคราะห์หาบริเวณใบหน้า ดังนั้นจึงได้นำขั้นตอนวิธีการตรวจหาการเคลื่อนที่มาช่วยลดผลกระทบในส่วนนี้ ดังที่ได้กล่าวไว้ในบทที่ 3 และจะแสดงผลการทดลองในหัวข้อถัดไป



(ก)



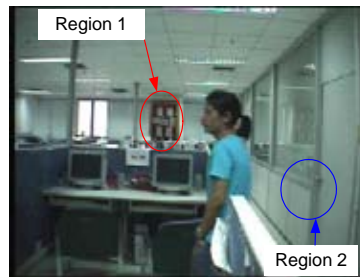
(ข)



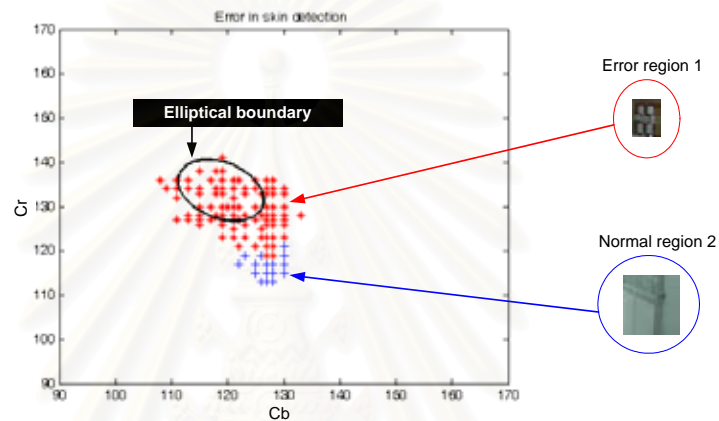
(ค)

รูปที่ 4.10 ผลการตรวจหาสี่ผิวด้วยแบบจำลองวงรีที่กำหนด

ในการวิเคราะห์ความผิดพลาดในการตรวจหาสี่ผิวด จะพิจารณาโดยนำบริเวณพื้นที่ปัญหา(เป็นบริเวณที่ไม่ใช่ใบหน้าและมือ แต่ระบบตีความว่าเป็นส่วนสี่ผิวด) มาพล็อตค่าแต่ละจุดภาพลงในปริภูมิย่อย CbCr จากรูปที่ 4.11(ข) จะเห็นว่าการกระจายตัวขององค์ประกอบสีของบริเวณที่มีปัญหามีค่าตกอยู่ในตำแหน่งขอบเขตของแบบจำลองวงรีที่สร้างขึ้น เนื่องจากคุณสมบัติของปริภูมิสี YCbCr ที่จะแยกเป็นองค์ประกอบความสว่าง และองค์ประกอบสี ทำให้ถ้าสังเกตค่าความเข้มของบริเวณสี่ผิวดที่สนใจ อาจจะไม่เห็นว่าไม่น่าจะเป็นส่วนสี่ผิวด แต่เมื่อตัดองค์ประกอบของความสว่างออกไป จะได้ช่วงของค่าองค์ประกอบสีที่ใกล้เคียงกัน ดังผลการตรวจหาในรูปที่ 4.11



(ก) ภาพรับเข้า และบริเวณที่ใช้ในการวิเคราะห์



(ข)

รูปที่ 4.11 การวิเคราะห์ความผิดพลาดที่เกิดจากการตรวจหาสีผิวตามแบบจำลองวงรี

#### 4.1.2 ผลการตรวจหาการเคลื่อนไหว

จากผลการตรวจหาสีผิว จะเห็นว่าปัญหาหลักที่เกิดขึ้น คือ สัญญาณรบกวนที่เกิดจากวัตถุในฉากหลังมีค่าความเข้มขององค์ประกอบสีที่ใกล้เคียงกับสีผิว ดังนั้นคุณสมบัติของการเคลื่อนไหวจึงได้ถูกนำมาประยุกต์ใช้ เพื่อใช้แยกบริเวณส่วนที่มีการเคลื่อนไหวที่เท่านั้น ออกจากฉากหลัง ซึ่งจะตัดผลกระทบที่เกิดจากการตรวจหาสีผิวได้

ในหัวข้อนี้จะกล่าวถึงการทดลองและวิเคราะห์ผลเกี่ยวกับการตรวจหาการเคลื่อนไหว ที่ คือ การทดลองเปรียบเทียบความแตกต่างของผลกระทบเมื่อเลือกใช้ช่วงห่างระหว่างเฟรมภาพที่นำมาประมวลผลที่ต่างกัน และ ผลการทดลองในกระบวนการเติมจุดภาพระหว่างแฉภาพในแนวนอน

#### 4.1.2.1 ผลกระทบของการเลือกช่วงห่างระหว่างเฟรมในกระบวนการหาผลต่างระหว่างเฟรมภาพ

ทดลองหาผลกระทบของช่วงห่างระหว่างสองเฟรมภาพที่นำมาใช้ประมวลผลในกระบวนการหาผลต่างระหว่างเฟรม ซึ่งช่วงห่างที่เลือกใช้ในการทดลอง คือ เฟรมที่ติดกันเลย ช่วงห่าง 5 และ 10 เฟรม ซึ่งสิ่งที่ต้องคำนึงถึงในการทดลอง จะประกอบด้วย อัตราเฟรม และ ความเร็วในการเคลื่อนที่ของบุคคลที่สนใจ อัตราเฟรมที่เลือกใช้ในการทดลองนี้และการทดลองทั้งหมดในระบบจะมีค่า 15 เฟรมต่อวินาที ซึ่งเป็นอัตราภาพทั่วไปที่ต้องการสำหรับระบบที่ต้องการทำงานในเวลาจริง (10-30 เฟรมต่อวินาที) ส่วนสมมติฐานการเคลื่อนที่ของบุคคลทดลองจะเป็นความเร็วปกติโดยทั่วไป รูปที่ 4.12 จะแสดงผลกระทบที่เกิดขึ้น



(ก) ช่วงห่างเฟรมที่ติดกัน



(ข) ช่วงห่าง 5 เฟรม



(ค) ช่วงห่าง 10 เฟรม

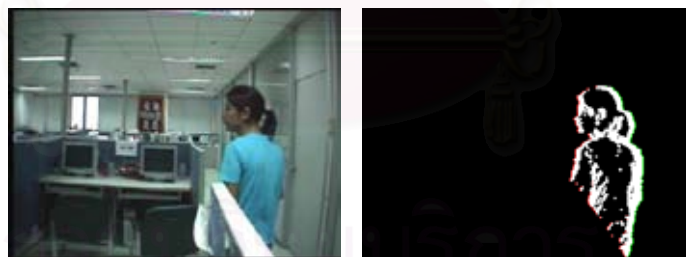
รูปที่ 4.12 ผลการตรวจหาการเคลื่อนที่ เมื่อเลือกช่วงห่างระหว่างเฟรมต่างกัน

จากรูปจะเห็นว่า ยิ่งช่วงห่างของสองเฟรมภาพที่นำมาใช้ประมวลผลเพิ่มมากขึ้น (ดังรูปที่ 4.12(ข) และ 4.12(ค)) จะได้ผลการตรวจหาการเคลื่อนที่เป็นบริเวณขอบการเคลื่อนที่

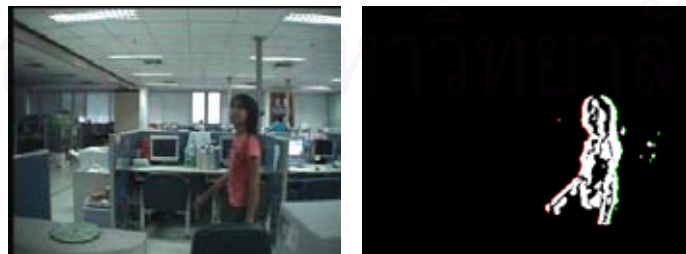
ที่มากขึ้น จนได้ผลการตรวจหาที่คล้ายกับลักษณะของโครงร่างบุคคล 2 คนซ้อนเหลื่อมกันอยู่ ซึ่งผลในลักษณะแบบนี้ น่าจะเหมาะที่จะนำไปใช้ร่วมกับกระบวนการตรวจหาขอบมากกว่า แต่วิทยานิพนธ์นี้เลือกประยุกต์ใช้กับการตรวจหาสีผิว ดังนั้นผลการตรวจหาที่เหมาะสมน่าจะเป็นเค้าโครงของตัวบุคคลเลย ซึ่งจากผลการทดลอง จึงเลือกใช้เฟรมที่ติดกันเลยสำหรับใช้หาผลต่างระหว่างเฟรมภาพ

#### 4.1.2.2 การวิเคราะห์ผลการตรวจหาการเคลื่อนที่และผลการเติมจุดภาพ

ดังผลการตรวจหาการเคลื่อนที่ในหัวข้อที่ผ่านมา ก็จะทำให้เห็นว่าผลการตรวจหาจะได้เฉพาะขอบการเคลื่อนที่และบางบริเวณเท่านั้น จึงต้องมีกระบวนการเพื่อเติมจุดภาพระหว่างแถวให้เต็มเพื่อจะได้เป็นโครงร่างของบุคคล ที่จะนำไปตรวจหาด้วยสีผิวอีกครั้ง รูปที่ 4.13 แสดงตัวอย่างการหาจุดภาพตำแหน่งเริ่มต้นและจุดภาพตำแหน่งสุดท้ายของแต่ละแถวในภาพฐานสองที่ผ่านการตรวจหาผลต่างระหว่างเฟรม ปัญหาของการวิเคราะห์จุดภาพเริ่มต้นและจุดภาพสุดท้ายคือ ผลกระทบที่เกิดขึ้นเมื่อฉากหลังประกอบด้วยจอภาพมอนิเตอร์ของคอมพิวเตอร์ ซึ่งบริเวณดังกล่าวมีการเปลี่ยนแปลงค่าความเข้มสีตลอดเวลา ทำให้ถูกต้องกว่าเป็นส่วนของวัตถุที่มีการเคลื่อนที่และทำให้การกำหนดจุดเริ่มต้นและสุดท้ายผิดพลาดดังรูปที่ 4.13(ข) และ 4.13(ค) รูปที่ 4.14 จะแสดงผลการเติมจุดภาพในแต่ละแถวที่อยู่ระหว่างจุดภาพเริ่มต้นและจุดภาพสุดท้าย โดยในรูปที่ 4.14(ค) จะแสดงปัญหาดังที่ได้กล่าวมา คือ จะเติมจุดภาพเกินขอบเขตของโครงร่างบุคคล



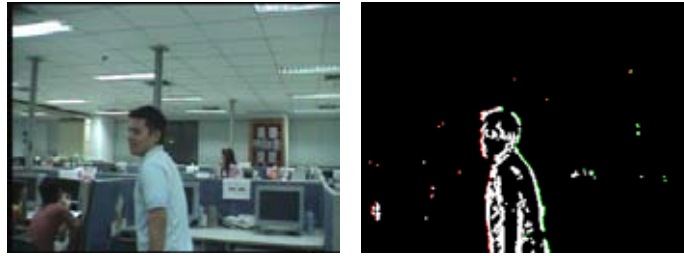
(ก)



(ข)

รูปที่ 4.13 ผลการหาตำแหน่งจุดภาพเริ่มต้นและจุดภาพตำแหน่งสุดท้ายของภาพฐานสองที่ได้จากการหาความต่างระหว่างเฟรม





(ค)

รูปที่ 4.13(ต่อ) ผลการหาตำแหน่งจุดภาพเริ่มต้นและจุดภาพตำแหน่งสุดท้ายของภาพฐานสองที่ได้จากการหาความต่างระหว่างเฟรม

การแก้ปัญหาในส่วนนี้สามารถทำได้โดยนำค่าตำแหน่งจุดภาพเริ่มต้นและจุดภาพสุดท้ายของแต่ละแถวและเป็นจุดภาพเคลื่อนที่มาหาค่าเฉลี่ย และค่าความแปรปรวน 2 กลุ่มประกอบด้วย ค่าของกลุ่มจุดภาพเริ่มต้น และค่าของกลุ่มจุดภาพสุดท้าย เพื่อตรวจสอบว่าค่าความผิดพลาดที่จะเกิดจากการเติมจุดภาพมากเกินไปที่กำหนดหรือไม่ โดยมีทางเลือกการแก้ปัญหา 2 แบบ คือ ไม่นำเฟรมภาพนี้ไปประมวลผลหาตำแหน่งใบหน้าต่อ หรือ ตัดแถวที่เกิดข้อผิดพลาดทิ้ง ซึ่งในกรณีแบบที่ 2 นี้ ไม่ได้เลือกนำมาใช้ เนื่องจากจะมีปัญหาถ้าบริเวณส่วนที่ถูกตัดทิ้งเป็นส่วนใบหน้า แต่จะไม่มีปัญหาถ้าเป็นส่วนของเสื้อผ้า จึงเลือกใช้วิธีแก้ปัญหาในแบบแรก ซึ่งจะมีบางเฟรมภาพที่ยังมีข้อผิดพลาดบ้าง (ดังรูปที่ 4.14(ข)) แต่ไม่ถูกตัดเฟรมภาพทิ้ง ซึ่งผลในส่วนนี้จะถูกแก้ปัญหาอีกครั้งในส่วนการกำหนดขอบเขตใบหน้า



(ก) ผลการเติมจุดภาพของรูปผลการตรวจหารูปที่ 4.13(ค)



(ข) ผลการเติมจุดภาพของรูปผลการตรวจหารูปที่ 4.13(ข)

รูปที่ 4.14 ผลการเติมจุดภาพระหว่างตำแหน่งจุดภาพเริ่มต้นและจุดภาพตำแหน่งสุดท้าย



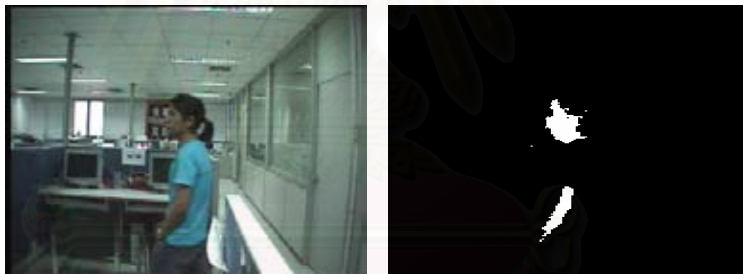


(ค) ผลการเติมจุดภาพของรูปผลการตรวจหารูปที่ 4.13(ค)

รูปที่ 4.14(ต่อ) ผลการเติมจุดภาพระหว่างตำแหน่งจุดภาพเริ่มต้นและจุดภาพตำแหน่งสุดท้าย

#### 4.1.3 ผลการตรวจหาใบหน้าด้วยการตรวจหาการเคลื่อนที่ร่วมกับการตรวจหาด้วยสีผิว

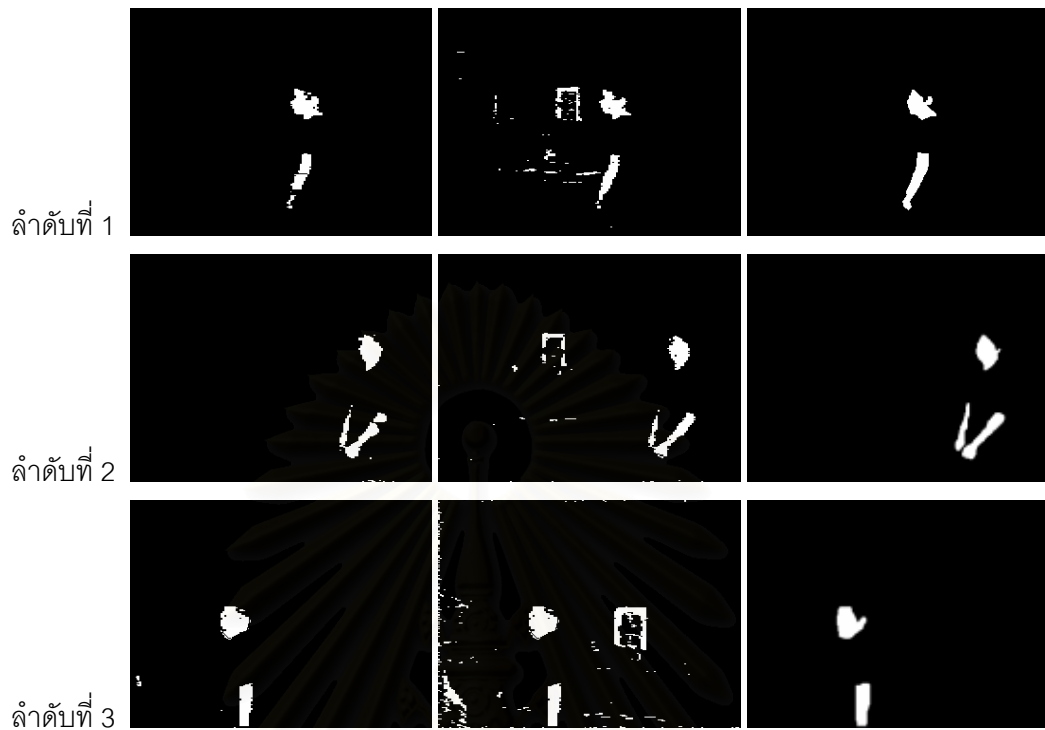
ขั้นตอนการตรวจหาใบหน้าในวิทยานิพนธ์นี้ จะใช้การตรวจหาการเคลื่อนที่เพื่อลดบริเวณที่จะนำมาใช้ในการตรวจหาสีผิวเพื่อหาบริเวณใบหน้า และ เมื่อ ได้อย่างมีประสิทธิภาพมากขึ้น ผลการตรวจหาใบหน้า แสดงได้ดังรูปที่ 4.15



รูปที่ 4.15 ผลการตรวจหาใบหน้าหลังจากผ่านการตรวจหาการเคลื่อนที่ร่วมกับตรวจหาสีผิว

##### 4.1.3.1 การวัดประสิทธิภาพของผลการตรวจหา

เป็นการวัดอัตราการตรวจหาส่วนบริเวณสีผิวระหว่างผลการตรวจหาที่ได้จากการตรวจหาการเคลื่อนที่ร่วมกับตรวจหาสีผิว เทียบกับผลการตรวจหาที่ใช้การตรวจหาสีผิวอย่างเดียว โดยใช้ค่าค่าอัตราการตรวจหาและค่าอัตราการฟองค่าผิดพลาด ในการเปรียบเทียบและผลการตรวจหาทั้งสองต้องนำมาเปรียบเทียบกับผลถูกต้องพื้นฐาน (Ground Truth) (ซึ่งถูกกำหนดด้วยมือ) รูปที่ 4.16 แสดงตัวอย่างของภาพฐานสองของผลการตรวจหาเปรียบเทียบระหว่าง 2 แบบ และตารางที่ 4.4 เป็นผลการวัดประสิทธิภาพในแต่ละแบบ



รูปที่ 4.16 ผลการตรวจหาส่วนสีผิว (ก) ตรวจหาด้วยการเคลื่อนที่ร่วมกับตรวจหาด้วยสีผิว (ข) ตรวจหาด้วยสีผิวอย่างเดียว (ค) ผลถูกต้องพื้นฐาน

ในการวัดประสิทธิภาพของการตรวจหา จะเปรียบเทียบกลุ่มบริเวณสีผิว (บริเวณใบหน้า รวมทั้งบริเวณท่อนแขนและมือทั้งสองข้างด้วย) ที่ถูกตรวจหาได้ กับบริเวณสีผิวของผลถูกต้องพื้นฐาน (ซึ่งถูกกำหนดขึ้นเอง)

ตารางที่ 4.4 เปรียบเทียบประสิทธิภาพของผลการตรวจหา

	ตรวจหาด้วยการเคลื่อนที่และสีผิว		ตรวจหาด้วยสีผิวอย่างเดียว	
	DR	FAR	DR	FAR
ลำดับที่ 1	0.7623	0.1494	0.7406	0.4357
ลำดับที่ 2	0.8143	0.1007	0.7508	0.2817
ลำดับที่ 3	0.8743	0.1107	0.8743	0.6153

จากผลการวัดประสิทธิภาพ จะเห็นว่าค่าอัตราการตรวจหาระหว่างสองวิธีมีความแตกต่างกันไม่มาก หรือเกือบเท่ากัน แต่สิ่งที่เป็นข้อดีของการตรวจหาด้วยการเคลื่อนที่ร่วมกับสีผิว คือ ค่าอัตราการฟ้องค่าผิดพลาดน้อยมากเมื่อเทียบกับการตรวจหาด้วยสีผิวอย่างเดียว

#### 4.1.3.2 การแยกกลุ่มบริเวณใบหน้า

ผลการตรวจหาสีผิวที่ได้ จะประกอบด้วยส่วนบริเวณ 3 ส่วนหลัก ประกอบด้วย ส่วนใบหน้า และส่วนท่อนแขนรวมมือทั้งสองข้าง ซึ่งสำคัญในการพิจารณา คือ ต้องเลือกบริเวณ ใบหน้าออกจากกลุ่มบริเวณท่อนแขน ซึ่งปัญหาในการแยกบริเวณใบหน้าจะเกิดขึ้น เมื่อบุคคลที่สนใจสวมใส่เสื้อแขนสั้นดังรูปที่ 4.17 (กรณีที่ 1 และ 2)



กรณีที่ 1 เสื้อแขนสั้น

กรณีที่ 2 เสื้อแขนสั้น

กรณีที่ 3 เสื้อแขนยาว

(ก) ผลการตรวจหา

(ข) ผลการแยกบริเวณใบหน้า

รูปที่ 4.17 ผลการแยกบริเวณใบหน้า

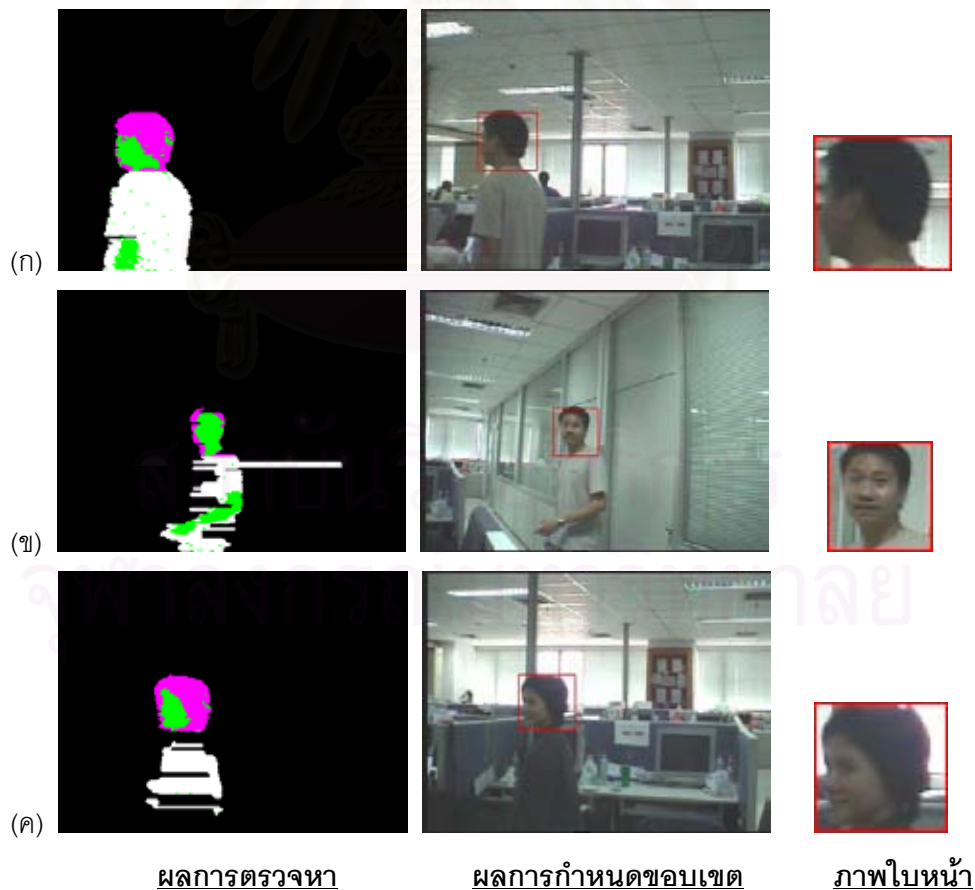
ในกรณีบุคคลสวมใส่เสื้อแขนยาว เมื่อพิจารณาผลการตรวจหา (รูปที่ 4.17 (กรณี ที่ 3)) จะเห็นว่าใช้คุณลักษณะด้านขนาดก็สามารถแยกบริเวณใบหน้าได้ แต่ในกรณีสวมเสื้อแขน สั้น (รูปที่ 4.17 (กรณีที่ 1 และ 2)) ถ้าใช้ลักษณะขนาดอย่างเดียว จะไม่สามารถแยกบริเวณใบหน้า ได้อย่างถูกต้อง เนื่องจากขนาดของบริเวณท่อนแขนค่อนข้างใกล้เคียงกับบริเวณใบหน้า ดังนั้นรูป ที่ 4.17(กรณีที่ 1 และ 2) สามารถแยกบริเวณใบหน้าได้โดยใช้คุณลักษณะเกี่ยวกับรูปร่าง ดังที่

อธิบายไว้สมการที่ (2.12) และ (2.13) และสมมติฐานในการวิเคราะห์ในส่วนที่กล่าวมานี้ คือบุคคลต้องเดินด้วยท่าปกติ (ไม่มีการยกแขน)

#### 4.1.4 กระบวนการกำหนดขอบเขตและตัดส่วนภาพใบหน้า

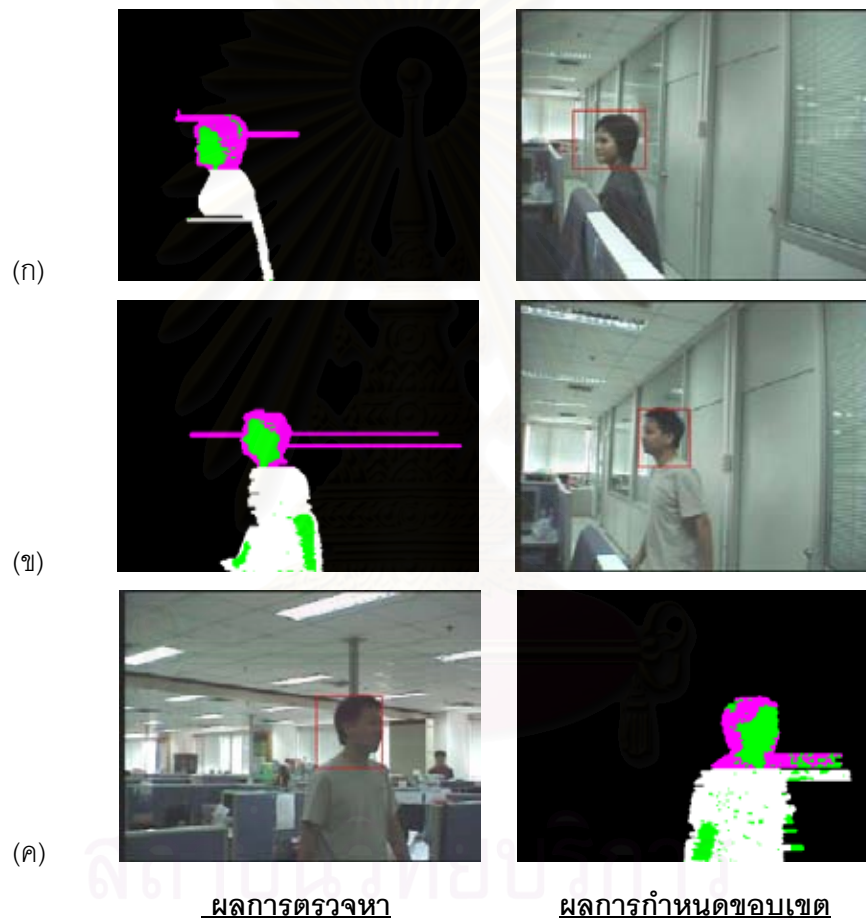
จากหัวข้อการทดลองที่ผ่านมา บริเวณใบหน้าจะถูกแยกออกมาและผ่านการจัดกลุ่มเรียบร้อยแล้ว ซึ่งกระบวนการในหัวข้อนี้ถือว่าเป็นส่วนสำคัญ ที่จะกำหนดขอบเขตของใบหน้า เพื่อนำมาใช้เป็นขอบเขตในการตัดส่วนภาพใบหน้า เนื่องจากเป้าหมายของวิทยานิพนธ์นี้ต้องการนำภาพใบหน้าเพื่อไปใช้ในกระบวนการจับคู่ใบหน้า

ขั้นตอนที่ 1 เป็นการกำหนดขอบเขตอย่างคร่าว ในลักษณะสี่เหลี่ยม ดังรูปที่ 4.18 ซึ่งจะเกิดปัญหาขึ้นที่บริเวณส่วนลำคอของบุคคล ที่จะใช้เป็นขอบล่างของขอบเขตสี่เหลี่ยม เพราะขอบเขตสี่เหลี่ยมจะถูกกำหนด โดยครอบคลุมถึงบริเวณหัวไหล่ของบุคคลด้วยดังรูปที่ 4.18(ข) ดังนั้นขอบเขตในลักษณะนี้จึงยังไม่เหมาะต่อการนำไปใช้สำหรับจับคู่ใบหน้า เนื่องจากยังติดบริเวณที่เป็นฉากหลังมาก



รูปที่ 4.18 ผลการกำหนดขอบเขตใบหน้าในลักษณะสี่เหลี่ยม

ปัญหาที่เกิดขึ้นอีกอย่างในส่วนของการกำหนดขอบเขตแบบสีเหลี่ยม จะเป็นผลกระทบที่มาจากส่วนการเติมจุดภาพ (ดังที่กล่าวไว้ในหัวข้อย่อย 4.1.2.2) และจะเกิดข้อผิดพลาดในการหาขอบเขตซ้ายและขวาของกรอบสี่เหลี่ยม ดังตัวอย่างรูปที่ 4.19(ข) จึงต้องมีการตรวจสอบตำแหน่งของขอบเขตซ้ายและขวาอีกครั้ง โดยจะตัดแถวที่มีความยาวของจำนวนจุดภาพเกินกว่าค่าขีดเริ่มเปลี่ยนที่กำหนดออก ส่วนที่มีความยาวน้อย จะไม่ส่งผลกระทบต่อการหาขอบซ้ายและขอบขวา



รูปที่ 4.19 ข้อผิดพลาดในการกำหนดขอบเขตด้านซ้ายและขวาของกรอบสี่เหลี่ยม

ขั้นตอนที่ 2 เป็นการกำหนดขอบเขตแบบวงรีหรือวงกลม ซึ่งจะตัดผลกระทบของฉากหลังออกได้มาก เมื่อพิจารณาตามรูปร่างโดยคร่าวของใบหน้า และเป็นลักษณะที่มีความเหมาะสมที่สุดสำหรับการประมวลผลจับคู่ใบหน้าในเวลาจริง ที่ไม่มีการกำหนดฉากหลังที่แน่นอน จากรูปที่ 4.20 จะเห็นว่าเมื่อพิจารณาที่รูปร่างแล้ว ในกรณีของบุคคลอยู่ในท่าทางที่หน้าตรงหรือหน้าค่อนข้างตรงจะมีประสิทธิภาพของการกำหนดขอบเขตได้ดีที่สุด



รูปที่ 4.20 ผลการกำหนดขอบเขตและตัดส่วนภาพใบหน้าในแบบวงรีหรือวงกลม

#### 4.1.5 ผลการจำแนกตำแหน่งมุมมองใบหน้า

รูปที่ 4.21 แสดงภาพใบหน้าที่เป็นภาพที่ผ่านการตรวจหา โดยพิจารณาที่ตำแหน่งของกลุ่มบริเวณสีผิวใบหน้ากว้างตัวอยู่ในขอบเขตวงรี เพื่อให้แยกประเภทของใบหน้าตามมุมมอง (จะเห็นว่าขนาดภาพใบหน้าแต่ละกลุ่มจะไม่เท่ากัน ขึ้นอยู่กับขนาดของใบหน้า ณ เวลาที่รับภาพได้) และรูปที่ 4.22 แสดงผลการจำแนกที่แยกภาพใบหน้า ออกเป็นภาพใบหน้ากลุ่มหน้าตรง กลุ่มหน้าเอียงด้านซ้าย และกลุ่มหน้าเอียงด้านขวา



รูปที่ 4.21 ภาพผลการตรวจหาเฉพาะบริเวณใบหน้า

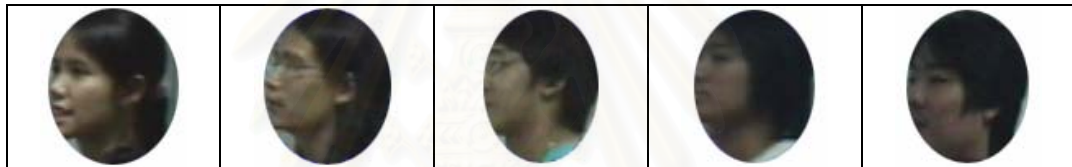




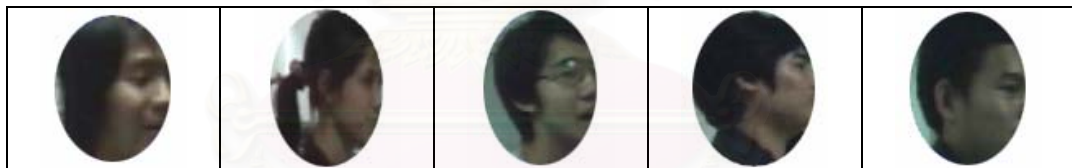
รูปที่ 4.21 (ต่อ) ภาพผลการตรวจหาเฉพาะบริเวณใบหน้า



(ก)



(ข)



(ค)

รูปที่ 4.22 การจำแนกกลุ่มของใบหน้า (ก) ใบหน้าตรง (ข) ใบหน้าเอียงซ้าย (ค) ใบหน้าเอียงขวา

#### 4.2 ผลการจำลองระบบในสถานการณ์ควบคุมการสาย ก้มเงยและซูมติดตามบุคคล

รูปที่ 4.23 แสดงตัวอย่างของการก้มเงย สาย และซูม เพื่อเป้าหมายที่พยายามให้บุคคลอยู่ในบริเวณกึ่งกลางภาพ ซึ่งปัญหาที่เกิดขึ้นกับการควบคุมตำแหน่งการเปลี่ยนมุมมองของกล้องวิดีโอคือ กรณีที่บุคคลมีการเคลื่อนที่ไปด้วย ขณะที่กำลังสั่งให้กล้องเปลี่ยนมุมมอง จะเกิดผลกระทบเกี่ยวกับตำแหน่งของใบหน้าบุคคลที่ไม่อยู่ในบริเวณกึ่งกลางภาพ ดังที่คาดหวัง แต่จะมีการเอียงตำแหน่งบ้าง ดังรูปที่ 4.24(ก) จะเห็นว่าแม้จะตำแหน่งจะไม่อยู่กึ่งกลางแต่ก็ยังอยู่ในกรอบภาพ และอยู่ในมุมมองที่ดีขึ้นกว่ามุมมองเดิมก่อนการเปลี่ยนมุมมองกล้องวิดีโอ

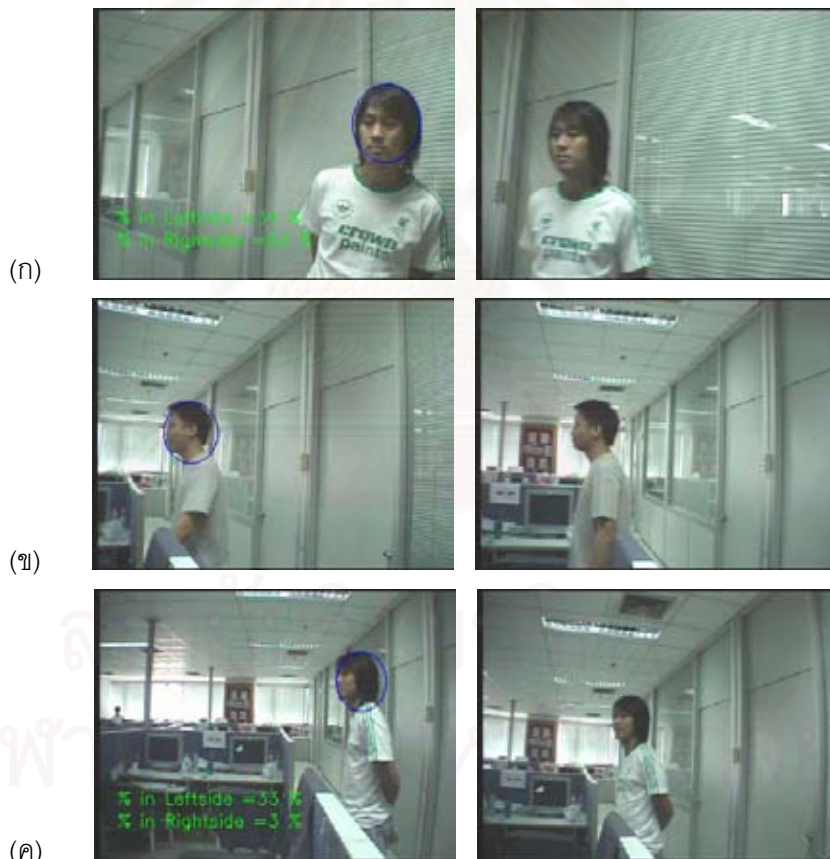


ลำดับวีดิทัศน์ที่ 1



ลำดับวีดิทัศน์ที่ 2

รูปที่ 4.23 ผลการควบคุมกล้องเพื่อเปลี่ยนมุมมองติดตามใบหน้า



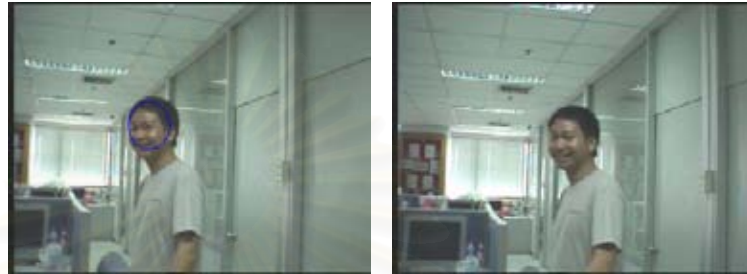
ก่อนเคลื่อนที่

หลังเคลื่อนที่

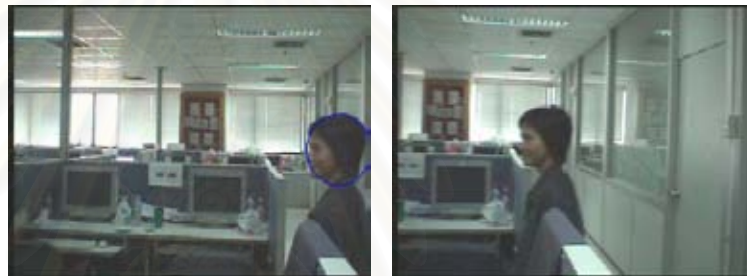
รูปที่ 4.24 ปัญหาในการถ่าย และก้มเงยกล้องวีดิทัศน์ ติดตามบุคคลที่มีการเคลื่อนที่ (ก) เมื่อบุคคลเคลื่อนที่เปลี่ยนทิศทาง (ข) เมื่อบุคคลกำลังเคลื่อนที่ด้วยความเร็ว (ค) เมื่อบุคคลหยุดนิ่ง

และในกรณีของการควบคุมการส่ายและการก้มเงย ณ ขณะที่กำลังอยู่ในสถานะ มุมมองที่ถูกซูมอยู่ จากปัญหาที่ได้กล่าวไปให้เห็นในส่วนหัวข้อย่อย 3.3.2 หลังจากผ่าน กระบวนการชดเชยตัวประกอบการซูมแล้ว จะสามารถลดปัญหาในส่วนนี้ได้ ดังผลการทดลองรูปที่ 4.25

ตัวอย่างที่ 1



ตัวอย่างที่ 2



รูปที่ 4.25 ผลการเปลี่ยนมุมมองกล้องเมื่อกำลังอยู่ในสถานะมุมมองที่ถูกซูมอยู่

#### 4.3 ผลการจำลองระบบในส่วนการทำงานร่วมกันระหว่างกล้องวิดีโอทัศน์

การจำลองระบบในส่วนการทำงานร่วมกันระหว่างกล้องวิดีโอทัศน์ จะทำการ ทดลองแยกกัน แบ่งเป็น การทดลองเพื่อแสดงผลการประมวลผลร่วมกันสำหรับจับคู่ใบหน้า ซึ่ง เป้าหมายหลักของการทดลองนี้เพื่อแสดงให้เห็นถึงประโยชน์และข้อได้เปรียบของการใช้กล้องวิดีโอ ทัศน์ PTZ 2 ตัว และการทดลองสำหรับการทำงานร่วมกันในการตรวจหาติดตามบุคคลเป้าหมายที่ สนใจอย่างต่อเนื่องระหว่างกล้องวิดีโอทัศน์ตัวหนึ่งไปยังกล้องวิดีโอทัศน์อีกตัวหนึ่ง โดยทั้งสองกระบวน การนี้จะใช้ขั้นตอนวิธีการจับคู่ด้วยการวัดความคล้ายที่เหมือนกัน ทำให้ง่ายต่อการออกแบบและ การนำไปประยุกต์ใช้งาน







การจำลองระบบจะประมวลผลในเวลาจริง ขนาดภาพที่ใช้ในการประมวลผลมี ขนาดยาว 320 จุดภาพ และกว้าง 240 จุดภาพ โดยโปรแกรมกระบวนการทำงานทุกอย่างด้วย ภาษาซี (C programming) และนำคลังโปรแกรม (program library) ของอินเทลที่เรียกว่า OpenCV มาใช้เป็นเครื่องมือหลักในการโปรแกรมทางการเชื่อมต่อกับกล้องวิดีโอทัศน์ โดยคุณสมบัติ

ของเครื่องคอมพิวเตอร์ที่ใช้ในการประมวลผล คือ เครื่องคอมพิวเตอร์รุ่น Pentium 4 ความเร็วในการประมวลผล 2.8 GHz ความจุแรม 512 MB บริเวณที่ใช้สำหรับการทดลองจะเป็นบริเวณห้องปฏิบัติการวิจัยกรรมวิธีสัญญาณดิจิทัล และมีสมมติฐานที่ว่า มีบุคคลเคลื่อนที่อยู่ภายในมุมมองกล้องวิดีโอที่คนเพียงบุคคลเดียวเท่านั้น

#### 4.3.1 การประมวลผลร่วมกันสำหรับจับคู่ใบหน้า

##### 4.3.1.1 ผลการควบคุมการส่ายและการก้มเงยของกล้องวิดีโอที่ติดตามใบหน้าบุคคลจาก 2 มุมมอง

รูปที่ 4.26 แสดงลำดับภาพจากสัญญาณวิดีโอที่รับภาพจากกล้องวิดีโอ PTZ ทั้งสองกล้อง ซึ่งจะเห็นว่ากล้องวิดีโอทั้งสองตัวจะประมวลผลอย่างเป็นอิสระ และเคลื่อนที่ไปยังตำแหน่งที่เป็นอิสระ โดยจะทำงานขนานกัน สังเกตที่มุมมองของใบหน้าที่ได้รับได้ จะต้องมียังน้อยกล้องหนึ่งตัวที่รับภาพใบหน้าในลักษณะที่ค่อนข้างตรงที่หน้าจะมีความน่าจะเป็นในการจับคู่ใบหน้าได้ดีกว่ามุมมองอื่น ซึ่งเป็นข้อดีของการทำงานและจัดวางตำแหน่งกล้องดังที่กล่าวไว้ในบทที่ 3

Frame No.	98	322	508
Camera 1			
Camera 2			

(ก) ตัวอย่างที่ 1

รูปที่ 4.26 ภาพตัวอย่างจากลำดับวิดีโอการส่าย ก้มเงยและซูมกล้องวิดีโอเพื่อติดตามใบหน้าบุคคลที่สนใจ ที่รับจากกล้องวิดีโอ PTZ 2 ตัวในตำแหน่งมุมมองต่างกัน สำหรับประยุกต์ใช้ในการจับคู่ใบหน้า









(ข) ตัวอย่างที่ 2

รูปที่ 4.26(ต่อ) ภาพตัวอย่างจากลำดับวีดิทัศน์การสาย กัมเมยและชุมก้องวีดิทัศน์เพื่อติดตามใบหน้าบุคคลที่สนใจ ที่รับจากกล้องวีดิทัศน์ PTZ 2 ตัวในตำแหน่งมุมมองต่างกัน สำหรับประยุกต์ใช้ในการจับคู่ใบหน้า

#### 4.3.1.2 มุมมองภาพใบหน้าที่รับจากกล้องทั้งสอง


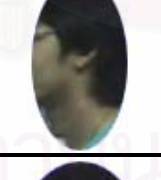


จากผลการทดลองในหัวข้อย่อย 4.3.1.1 เป็นภาพที่ได้จากการติดตามใบหน้า ซึ่งเมื่อผ่านกระบวนการกำหนดขอบเขตใบหน้า กระบวนการตัดส่วนภาพใบหน้า และกระบวนการจำแนกมุมมองของใบหน้าแล้ว ผลที่ได้จะเป็นภาพเฉพาะบริเวณใบหน้าที่มีรูปร่างวงรีหรือวงกลมที่จะตัดส่วนบริเวณฉากหลัง และทราบประเภทมุมมองของใบหน้านั้น ๆ ว่าถูกจัดอยู่ในกลุ่มมุมมองใด (ใบหน้าที่ตรง ใบหน้าด้านซ้าย และใบหน้าที่ด้านขวา) ดังผลการทดลองรูปที่ 4.27 และพร้อมที่จะนำไปประมวลต่อไปในส่วนของการจับคู่ใบหน้า

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

ลำดับภาพที่ 1	กล้องวีดิทัศน์ PTZ ตัวที่ 1		กล้องวีดิทัศน์ PTZ ตัวที่ 2	
	ภาพใบหน้า	มุมมองใบหน้า	ภาพใบหน้า	มุมมองใบหน้า
88		ซ้าย		ซ้าย
94		ตรง		ซ้าย
122		ซ้าย		ซ้าย
149		ตรง		ตรง
174		ขวา		ตรง
235		ตรง		ตรง
501		ขวา		ตรง

รูปที่ 4.27 ผลการตัดส่วนภาพใบหน้าจากทั้งสองกล้องและผลการจำแนกประเภทของมุมมองใบหน้า



ลำดับภาพที่ 2	กล้องวีดิทัศน์ PTZ ตัวที่ 1		กล้องวีดิทัศน์ PTZ ตัวที่ 2	
	ภาพใบหน้า	มุมมองใบหน้า	ภาพใบหน้า	มุมมองใบหน้า
102		ตรง		ขวา
119		ตรง		ตรง
136		ตรง		ขวา
177		ตรง		ขวา
255		ตรง		ตรง
367		ซ้าย		ตรง
440		ซ้าย		ตรง

รูปที่ 4.27 (ต่อ) ผลการตัดส่วนภาพใบหน้าจากทั้งสองกล้องและผลการจำแนกประเภทของมุมมองใบหน้า

#### 4.3.1.3 การฝึกสอนในกระบวนการจับคู่ใบหน้า

ในการจับคู่ภาพใบหน้าโดยใช้หลักการเปรียบเทียบค่าการกระจายตัวของลักษณะสีของภาพใบหน้าที่รับเข้าและภาพใบหน้าในฐานข้อมูล โดยใช้สัมประสิทธิ์ Bhattacharyya จำเป็นต้องมีกระบวนการในการฝึกสอนเพื่อหาสัมประสิทธิ์ Bhattacharyya ของแต่ละใบหน้าในฐานข้อมูลทั้งหมด

โดยฐานข้อมูลใบหน้าจะแยกออกเป็น 3 กลุ่ม คือ หน้าตรง หน้าด้านข้างซ้าย และ หน้าด้านข้างขวา ดังตัวอย่างภาพใบหน้าในฐานข้อมูลรูปที่ 4.28 และสิ่งหนึ่งที่สำคัญที่ต้องพิจารณา คือ ที่มาของภาพฐานข้อมูลทั้งหมด โดยในวิทยานิพนธ์นี้เลือกผลของการตัดส่วนภาพใบหน้าที่ได้จากการทดลองประมวลผลระบบจริง มาเป็นฐานข้อมูล เพื่อให้มีประสิทธิภาพในการจับคู่ภาพใบหน้ามากกว่ารูปแบบที่ใช้การตัดส่วนภาพใบหน้าด้วยมือ

ในการทดลอง จำนวนบุคคลในฐานข้อมูล มีทั้งหมด 5 คน ประกอบด้วย บุคคลชาย 4 คนและบุคคลหญิง 1 คน โดยจำนวนภาพใบหน้าของฐานข้อมูลที่ใช้ในการทดลอง คือ ภาพใบหน้าตรงทั้งหมด 200 ภาพ แบ่งเป็นของแต่ละบุคคลละ 40 ภาพใบหน้า ภาพใบหน้าเอียงซ้ายทั้งหมด 100 ภาพ แบ่งเป็นของแต่ละบุคคลละ 20 ภาพใบหน้า และภาพใบหน้าเอียงขวาทั้งหมด 100 ภาพ แบ่งเป็นของแต่ละบุคคลละ 20 ภาพใบหน้า ดังตัวอย่างใบหน้าในฐานข้อมูลรูปที่ 4.28



(ก) ฐานข้อมูลมุมมองใบหน้าเอียงซ้าย

รูปที่ 4.28 ภาพตัวอย่างในฐานข้อมูล จำแนกตามประเภทของมุมมองใบหน้า



(ข)ฐานข้อมูลมุมมองใบหน้าตรง



(ค)ฐานข้อมูลมุมมองใบหน้าเฉียงขวา

รูปที่ 4.28 (ต่อ) ภาพตัวอย่างในฐานข้อมูล จำแนกตามประเภทของมุมมองใบหน้า

#### 4.3.1.4 ผลการรู้จำใบหน้าด้วยวิธีการจับคู่ใบหน้า

ในการวัดประสิทธิภาพในการจับคู่ใบหน้าจะทำการเปรียบเทียบเปอร์เซ็นต์การจับคู่ใบหน้าที่ต้องต้องระหว่างเปอร์เซ็นต์ของกล้องวิดีโอทีละตัว และเปอร์เซ็นต์ที่ได้จากการประมวลผลจับคู่ใบหน้าที่รวมกัน ซึ่งแสดงได้ดังตารางที่ 4.5 และตัวอย่างผลการจับคู่ใบหน้าของภาพใบหน้ารับเข้าและภาพใบหน้าในฐานข้อมูลแสดงดังรูปที่ 4.29

ตารางที่ 4.5 เปรียบเทียบผลการจับคู่ใบหน้าที่ระหว่างผลของกล้องแต่ละตัวกับผลที่ได้จากการประมวลผลรวมกัน

ผลการรู้จำโดยการจับคู่ใบหน้า	กล้องวิดีโอ PTZ ตัวที่ 1	กล้องวิดีโอ PTZ ตัวที่ 2	กล้องวิดีโอ PTZ 2 ตัวประมวลผลรวมกัน
เปอร์เซ็นต์การจับคู่ (ร้อยละ)	52.08	66.67	72.91



(ก)



(ข)

รูปที่ 4.29 ตัวอย่างผลการจับคู่ใบหน้า (ก) ถูกต้อง (ข) ไม่ถูกต้อง



จากตารางที่ 4.4 จะเห็นว่าเมื่อใช้การประมวลผลร่วมกันสำหรับสรุปผลการจับคู่ใบหน้าบุคคลที่สนใจนั้น เปอร์เซ็นต์การจับคู่ใบหน้าจะเพิ่มมากขึ้นกว่าเปอร์เซ็นต์การจับคู่ของกล้องแต่ละตัว เมื่อพิจารณาเทียบกับผลการจับคู่ใบหน้าจากกล้องตัวที่ 2 อาจจะเห็นว่าเปอร์เซ็นต์มีอัตราเพิ่มไม่มาก แต่เมื่อเปรียบเทียบกับผลของกล้องตัวที่ 1 ก็จะทำให้มีอัตราเปอร์เซ็นต์การจับคู่ใบหน้าเพิ่มขึ้นถึงประมาณร้อยละ 20

ข้อเสียของการจับคู่ คือ จำเป็นต้องมีฐานข้อมูลของใบหน้าของบุคคลที่สนใจในปริมาณที่มาก และต้องมีหลายลักษณะมุมมองของใบหน้า แต่การเรียงตัวของใบหน้าในภาพไม่ต้องการกำหนดตำแหน่งที่แน่นอน กล่าวคือ ในการรู้จำที่มีข้อมูลในงานวิจัยทางด้านการรู้จำใบหน้า สมมติในฐานข้อมูลเป็นภาพใบหน้าตำแหน่งใบหน้าอยู่ในตำแหน่งประมาณกึ่งกลางของภาพใบหน้า ดังนั้นภาพรับเข้าที่จะใช้ประมวลผลก็จำเป็นต้องมีการเรียงตัวอยู่กึ่งกลางภาพด้วยเช่นกัน ซึ่งประสิทธิภาพจะลดลงเมื่อตำแหน่งของใบหน้ามีการเอียงไปทางซ้ายหรือทางขวาของภาพ และข้อดีอีกประการ คือ ใช้เวลาในการประมวลผลไม่มาก เหมาะสำหรับการประมวลผลในเวลาจริง

จากรูปที่ 4.29(ข) จะเห็นว่ากรการจับคู่ที่ผิดพลาด เกิดจากภาพใบหน้าที่รับเข้ามีความเข้มของแสงที่ใกล้เคียงกับภาพใบหน้าในฐานข้อมูล และยังมีลักษณะรูปแบบการตัดส่วนภาพใบหน้าที่คล้ายกันด้วย

#### 4.3.2 การทำงานร่วมกันสำหรับการติดตามบุคคลต่อเนื่องระหว่างกล้อง

ในหัวข้อนี้จะแสดงถึงการทดสอบระบบในการติดตามบุคคลอย่างต่อเนื่องระหว่างกล้องวีดิทัศน์ โดยลักษณะหลักที่เลือกใช้ คือ การวัดความคล้ายของฮิสโทแกรมสีด้วยสัมประสิทธิ์ Bhattacharyya โดยจะแสดงถึงขั้นตอนวิธีรวมทั้งปัญหาที่เกิดขึ้น ซึ่งทำให้ต้องนำลักษณะที่ศทางการเคลื่อนที่ของบุคคลมาประยุกต์ร่วมด้วยเพื่อความถูกต้องของการระบุบุคคลที่เพิ่มมากขึ้น

##### 4.3.2.1 ผลการแยกบริเวณเสื้อผ้า สีผิว และผม ของบุคคล

รูปที่ 4.30 แสดงผลการแยกและจัดกลุ่มของบริเวณส่วนที่ได้จากการตรวจหาการเคลื่อนที่ร่วมกับการตรวจหาสีผิว โดยแบ่งบริเวณออกเป็น 3 ส่วนหลัก คือ ส่วนที่เป็นเสื้อผ้าหรือเครื่องสวมใส่ ส่วนบริเวณสีผิว และส่วนบริเวณผม ซึ่งเหตุผลในการแยกบริเวณเพื่อให้การเปรียบเทียบฮิสโทแกรมระหว่างกล้องตัวแรกไปยังกล้องตัวที่สองมีประสิทธิภาพมากยิ่งขึ้นดังผลการทดลองในหัวข้อย่อยถัดไป



รูปที่ 4.30 ผลการทดลองจำแนกบริเวณที่ได้จากการตรวจหา

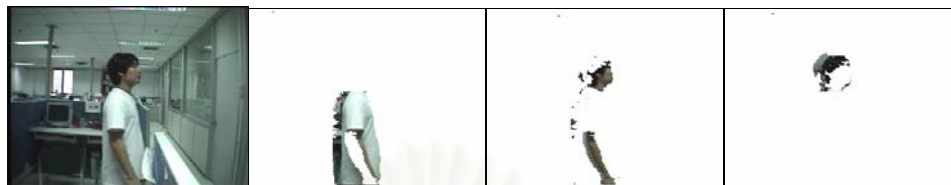
ปัญหาที่เกิดขึ้นในการจำแนกบริเวณ จะเป็นผลมาจากผลการตรวจหา คือ ในกรณีผลการตรวจหาการเคลื่อนที่มีขอบเขตเกินกว่าขอบเขตโครงร่างบุคคล เนื่องจากความเร็วของการเคลื่อนที่และปัญหาที่เกิดจากการกำหนดขอบเขตบริเวณใบหน้าเลยมาถึงช่วงไหล่ของบุคคล ผลกระทบต่อบริเวณส่วนที่เป็นเสื้อผ้าหรือเครื่องสวมใส่จะมีไม่มาก อาจส่งผลแค่จำนวนจุดภาพ อาจจะลดลง แต่จะส่งผลกระทบต่อค่าฮิสโทแกรมของบริเวณส่วนผม เพราะจะมีค่าสีในช่วงที่น่าจะมีเกิดขึ้น(โดยธรรมชาติสีผมของบุคคลที่สนใจจะมีโทนสีในช่วงสีดำเป็นส่วนใหญ่ และค่อนข้างมีฮิสโทแกรมที่เหมือนกันในทุกคน)

#### 4.3.2.2 การทดลองเปรียบเทียบความคล้ายของแต่ละกลุ่มบริเวณ

หัวข้อนี้จะกล่าวถึงการทดลองเปรียบเทียบความคล้ายระหว่างภาพบริเวณเค้าโครงบุคคลที่รับจากกล้องวีดิทัศน์ตัวที่หนึ่ง และตัวที่สอง ทั้งสามกลุ่มบริเวณ โดยแสดงผลเป็นค่าสัมประสิทธิ์ Bhattacharyya ดังตารางที่ 4.6 (ค่าสัมประสิทธิ์ Bhattacharyya นี้จะหาทั้งสามองค์ประกอบสีในปริภูมิสี RGB แล้วหาค่าระยะเฉลี่ยจากทั้งสามองค์ประกอบ) โดยภาพที่ใช้ในการ



ทดลองจะเป็นดังรูปที่ 4.31 และใช้ภาพบุคคล ก ที่กล้องวีดิทัศน์ตัวที่ 1 เป็นบุคคลหลักในการเปรียบเทียบ



ภาพที่ 1 บุคคล ก กล้องวีดิทัศน์ตัวที่ 1



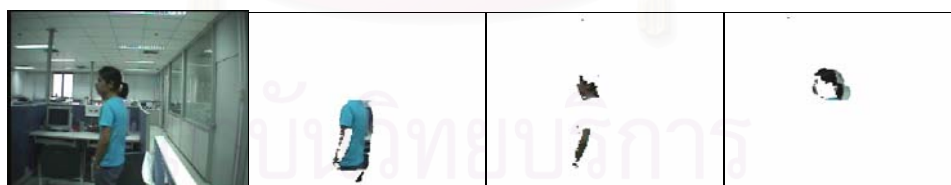
ภาพที่ 2 บุคคล ก กล้องวีดิทัศน์ตัวที่ 2



ภาพที่ 3 บุคคล ก กล้องวีดิทัศน์ตัวที่ 1



ภาพที่ 4 บุคคล ข กล้องวีดิทัศน์ตัวที่ 1



ภาพที่ 5 บุคคล ค กล้องวีดิทัศน์ตัวที่ 1



ภาพที่ 6 บุคคล ข กล้องวีดิทัศน์ตัวที่ 2

รูปที่ 4.31 ลำดับภาพบุคคลที่นำมาใช้ในการเปรียบเทียบความคล้าย

ตารางที่ 4.6 (ก) ผลการเปรียบเทียบความคล้ายของบุคคลระหว่างกล้องของกลุ่มบริเวณสีเสื้อผ้า

	ค่าสัมประสิทธิ์ Bhattacharyya			
	องค์ประกอบ สีแดง	องค์ประกอบ สีเขียว	องค์ประกอบ สีน้ำเงิน	รวม 3 องค์ประกอบ
ภาพที่ 1 กับ ภาพที่ 2	0.8777	0.8546	0.8315	0.8546
ภาพที่ 1 กับ ภาพที่ 3	0.8790	0.8665	0.8641	0.8699
ภาพที่ 1 กับ ภาพที่ 4	0.8289	0.5821	0.5810	0.6640
ภาพที่ 1 กับ ภาพที่ 5	0.5943	0.7725	0.8081	0.7250
ภาพที่ 1 กับ ภาพที่ 6	0.7186	0.6386	0.6621	0.6731

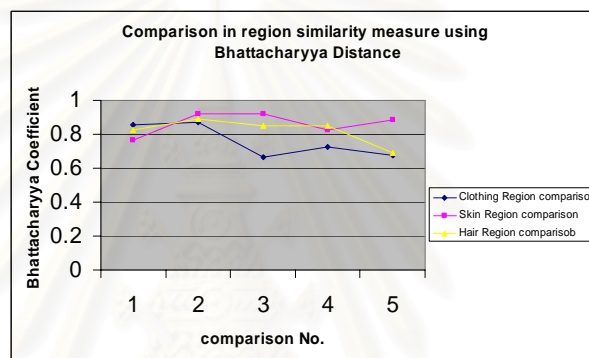
ตารางที่ 4.6 (ข) ผลการเปรียบเทียบความคล้ายของบุคคลระหว่างกล้องของกลุ่มบริเวณสีผิว

	ค่าสัมประสิทธิ์ Bhattacharyya			
	องค์ประกอบ สีแดง	องค์ประกอบ สีเขียว	องค์ประกอบ สีน้ำเงิน	รวม 3 องค์ประกอบ
ภาพที่ 1 กับ ภาพที่ 2	0.7448	0.7374	0.8082	0.7635
ภาพที่ 1 กับ ภาพที่ 3	0.9097	0.9218	0.9343	0.9219
ภาพที่ 1 กับ ภาพที่ 4	0.9085	0.9313	0.9244	0.9214
ภาพที่ 1 กับ ภาพที่ 5	0.8060	0.8102	0.8520	0.8227
ภาพที่ 1 กับ ภาพที่ 6	0.8869	0.8585	0.9060	0.8838

ตารางที่ 4.6 (ค) ผลการเปรียบเทียบความคล้ายของบุคคลระหว่างกล้องของกลุ่มบริเวณสีผม

	ค่าสัมประสิทธิ์ Bhattacharyya			
	องค์ประกอบ สีแดง	องค์ประกอบ สีเขียว	องค์ประกอบ สีน้ำเงิน	รวม 3 องค์ประกอบ
ภาพที่ 1 กับ ภาพที่ 2	0.7968	0.8360	0.8440	0.8256
ภาพที่ 1 กับ ภาพที่ 3	0.8958	0.8781	0.9016	0.8918
ภาพที่ 1 กับ ภาพที่ 4	0.8680	0.8404	0.8443	0.8509
ภาพที่ 1 กับ ภาพที่ 5	0.8688	0.8354	0.8487	0.8510
ภาพที่ 1 กับ ภาพที่ 6	0.6940	0.7250	0.6570	0.6920

จากผลการทดลองจะเห็นว่าการเปรียบเทียบฮิสโทแกรมด้วยกลุ่มบริเวณสีผิวและสีผมระหว่างภาพจากกล้องทั้งสองตัว ดังผลในตารางที่ 4.6(ข) และ (ค) ให้ค่าสัมประสิทธิ์ Bhattacharyya ที่ไม่ต่างกันมากระหว่างการเปรียบเทียบบุคคลเดียวกันกับเปรียบเทียบบุคคลอื่น ทำให้ยากต่อการกำหนดค่าขีดเริ่มเปลี่ยนในการตัดสินใจในการจำแนก แต่เมื่อพิจารณาที่ผลการเปรียบเทียบบริเวณเสื้อผ้าในตารางที่ 4.6(ง) จะมีความแตกต่างอย่างชัดเจนระหว่างการเปรียบเทียบบุคคลเดียวกันกับการเปรียบเทียบบุคคลอื่น โดยสามารถแสดงเป็นกราฟดังรูปที่ 4.32 ดังนั้นในการจำลองระบบจริง จึงเลือกใช้เฉพาะบริเวณเสื้อผ้าเท่านั้นในการใช้เปรียบเทียบฮิสโทแกรมระหว่างกล้อง



รูปที่ 4.32 กราฟเปรียบเทียบการวัดความคล้ายโดยใช้กลุ่มบริเวณสีเสื้อผ้า สีผิว และสีผม

จากรูปกราฟที่ 4.32

- การเปรียบเทียบลำดับที่ 1 คือ เปรียบเทียบภาพที่ 1 กับภาพที่ 2
- การเปรียบเทียบลำดับที่ 2 คือ เปรียบเทียบภาพที่ 1 กับภาพที่ 3
- การเปรียบเทียบลำดับที่ 3 คือ เปรียบเทียบภาพที่ 1 กับภาพที่ 4
- การเปรียบเทียบลำดับที่ 4 คือ เปรียบเทียบภาพที่ 1 กับภาพที่ 5
- การเปรียบเทียบลำดับที่ 5 คือ เปรียบเทียบภาพที่ 1 กับภาพที่ 6

#### 4.3.2.3 ผลการระบุบุคคลระหว่างกล้องวิดีโอทัศน์ PTZ 2 กล้อง

จะกำหนดการจำลองระบบเพื่อทดลองเป็นดังนี้ คือ มีบุคคลแรกที่เป็นบุคคลที่สนใจเดินเข้ามาในระบบ ซึ่งอาจจะเป็นกล้องวิดีโอทัศน์ตัวใดก็ได้ในสองตัว กล้องวิดีโอทัศน์นี้จะทำการถ่าย ก้มเงย และซูมติดตามพร้อมทั้งเก็บค่าข้อมูลสีของบุคคลที่สนใจนี้ไว้ โดยเลือกเก็บเป็นฐานข้อมูลฮิสโทแกรมสี 10 ค่า ที่มีช่วงเวลาการเก็บที่ต่างกัน จากนั้นจะทดลองการระบุบุคคลโดยให้บุคคลอื่นเดินเข้าไปยังมุมมองกล้องวิดีโอทัศน์ที่สอง กล้องวิดีโอทัศน์ที่สองจะทำการตรวจสอบบุคคลเมื่อระบุว่าไม่ใช่บุคคลเดิมจะไม่สั่งให้กล้องวิดีโอทัศน์ติดตาม จนเมื่อบุคคลที่สนใจเดินเข้า

มายังกล่องวิดิทัศน์ตัวที่สองเมื่อผ่านการตรวจสอบบุคคล โปรแกรมจะสั่งให้กล่องวิดิทัศน์สาย ก้มเงย และหมุนเพื่อติดตามบุคคลต่อไป โดยในการระบุบุคคลยังใช้ทิศทางการเคลื่อนที่มาพิจารณาด้วย ซึ่งในสมมติฐานของการเคลื่อนที่ คือ กล่องวิดิทัศน์ตัวแรก จะเก็บทิศทางการเคลื่อนที่ของบุคคล ทำให้ในการระบุบุคคลของกล่องวิดิทัศน์ตัวที่สอง จะพิจารณาว่าถ้าบุคคลนั้นเคลื่อนที่กลับทิศทาง ทำให้มีความเป็นไปได้มาก ที่จะไม่ใช้บุคคลเดิม


สถานการณ์สมมติของการทดสอบมี 3 แบบ คือ

- ในกรณีมีบุคคลเดียวเคลื่อนที่ระหว่างกล่องวิดิทัศน์ทั้งสองตัว กรณีนี้จะใช้ทั้งข้อมูลลักษณะสีและลักษณะทิศทางการเดินของบุคคลเป็นลักษณะตรวจสอบ ดังลำดับภาพผลการทดสอบรูปที่ 4.33

- ในกรณีมีบุคคลอื่นเคลื่อนที่เข้ามามองกล่องวิดิทัศน์ตัวที่ 2 ก่อนบุคคลเป้าหมายแต่มีทิศทางการเคลื่อนที่ตรงกันข้ามกับทิศทางการเคลื่อนที่ของบุคคลเป้าหมาย สถานการณ์แบบนี้จะเห็นว่า ถ้าบุคคลอื่นมีข้อมูลลักษณะสีที่ใกล้เคียงกับข้อมูลลักษณะสีของบุคคลเป้าหมายที่เก็บข้อมูลได้จากกล่องวิดิทัศน์ตัวที่ 1 จะผ่านการตรวจสอบข้อมูลลักษณะสี แต่เมื่อตรวจสอบลักษณะทิศทางการเคลื่อนที่ จะไม่ผ่านการตรวจสอบ เนื่องจากมีทิศทางการเดินตรงกันข้าม ฉะนั้นจะระบุว่าบุคคลนี้ไม่ใช่บุคคลเป้าหมาย กล่องวิดิทัศน์ตัวที่ 2 จะไม่ติดตามและเก็บข้อมูล ดังลำดับภาพผลการทดสอบรูปที่ 4.34

- ในกรณีมีบุคคลอื่นเคลื่อนที่เข้ามามองกล่องวิดิทัศน์ตัวที่ 2 ก่อนบุคคลเป้าหมายแต่มีทิศทางการเคลื่อนที่ทั้งตรงกันข้ามและทิศทางเดียวกันกับทิศทางการเคลื่อนที่ของบุคคลเป้าหมาย สถานการณ์แบบนี้จะเห็นว่าเมื่อบุคคลอื่นเคลื่อนที่ในลักษณะตรงข้าม จะไม่ผ่านการตรวจสอบทิศทางการเคลื่อนที่ แต่เมื่อบุคคลอื่นเคลื่อนที่ในทิศทางเดียวกับบุคคลเป้าหมาย จะผ่านการตรวจสอบลักษณะทิศทางการเคลื่อนที่ ดังนั้นจึงต้องใช้ข้อมูลลักษณะสีในการระบุบุคคลเป้าหมาย (ซึ่งในการทดสอบกำหนดให้เครื่องแต่งกายของบุคคลเป้าหมายและบุคคลอื่น มีสีค่อนข้างแตกต่างกัน) ดังลำดับภาพผลการทดสอบรูปที่ 4.35

จุฬาลงกรณ์มหาวิทยาลัย

	มุมมองกล้องวีดิทัศน์ตัวที่ 1	มุมมองกล้องวีดิทัศน์ตัวที่ 2
เฟรมภาพปกติเริ่มต้น		
บุคคลเป้าหมายเดินเข้ามามุมมองกล้องวีดิทัศน์ตัวที่ 1		--
กล้องวีดิทัศน์ตัวที่ 1 สาย กัมเมย และซูมติดตามเป้าหมายและเก็บข้อมูลลักษณะสีและข้อมูลทิศทางการเคลื่อนที่ไปทางซ้าย		--
สุดขอบเขตการเปลี่ยนมุมมองเพื่อติดตามของกล้องวีดิทัศน์ตัวที่ 1		--
บุคคลเป้าหมายเคลื่อนเข้ามามุมมองกล้องวีดิทัศน์ที่ 2 และเริ่มทำการตรวจสอบข้อมูลลักษณะสีและข้อมูลทิศทางการเคลื่อนที่	--	
การเคลื่อนที่เป็นไปในทิศทางซ้าย เช่นเดียวกับข้อมูลทิศทางที่ส่งมา จากกล้องวีดิทัศน์ตัวที่ 1	--	











รูปที่ 4.33 ลำดับภาพผลการติดตามและระบุบุคคลระหว่างกล้องวีดิทัศน์ตัวที่ 1 และ ตัวที่ 2 ของตัวอย่างสถานการณ์ที่ 1







<p>กล้องวีดิทัศน์ตัวที่ 2 เริ่มทำการถ่าย ก้มเงย และซูมติดตามบุคคล เป้าหมาย</p>	<p>--</p>	
<p>กล้องวีดิทัศน์ตัวที่ 2 ยังเปลี่ยนมุมมองติดตาม ถ้าบุคคลเป้าหมายยัง อยู่ในอาณาบริเวณที่สามารถถ่าย และก้มเงยไปได้</p>	<p>--</p>	

รูปที่ 4.33(ต่อ) ลำดับภาพผลการติดตามและระบุบุคคลระหว่างกล้องวีดิทัศน์ตัวที่ 1 และ ตัวที่ 2 ของตัวอย่างสถานการณ์ที่ 1

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย








	มุมมองกล้องวีดิทัศน์ตัวที่ 1	มุมมองกล้องวีดิทัศน์ตัวที่ 2
เฟรมภาพปกติเริ่มต้น		
บุคคลเป้าหมายเดินเข้ามุมมองกล้องวีดิทัศน์ตัวที่ 1		
กล้องวีดิทัศน์ตัวที่ 1 ทำการส่าย ก้มเงย และซูมติดตามบุคคลเป้าหมาย และเก็บข้อมูลลักษณะสีและข้อมูลทิศทางการเคลื่อนที่ไปทางซ้าย		
สุดขอบเขตการเปลี่ยนมุมมองติดตามของกล้องวีดิทัศน์ตัวที่ 1		
บุคคลอื่นเคลื่อนเข้ามุมมองกล้องวีดิทัศน์ตัวที่ 2 ก่อนบุคคลเป้าหมาย	--	
แต่การเคลื่อนที่เป็นไปในทิศทางตรงข้ามกับบุคคลเป้าหมาย (ทิศทางขวา)	--	

รูปที่ 4.34 ลำดับภาพผลการติดตามและระบุบุคคลระหว่างกล้องวีดิทัศน์ตัวที่ 1 และ ตัวที่ 2 ของตัวอย่างสถานการณ์ที่ 2

บุคคลเป้าหมายเริ่มเคลื่อนที่เข้าไป ยังมุมมองกล้องวีดิทัศน์ตัวที่ 2 เริ่มทำการระบุบุคคลด้วยข้อมูล ลักษณะสีและทิศทางการเคลื่อนที่	--	
บุคคลมีทิศทางการเคลื่อนที่ไป ทางซ้าย สามารถระบุได้ว่าเป็น บุคคลเดิม	--	
กล้องวีดิทัศน์ตัวที่ 2 จะเริ่มส่าย ก้ม เงย และซูมติดตาม	--	
กล้องตัวที่ 2 ทำการส่าย ก้มเงย และซูมติดตามต่อเนื่อง	--	







รูปที่ 4.34(ต่อ) ลำดับภาพผลการติดตามและระบุบุคคลระหว่างกล้องวีดิทัศน์ตัว  
ที่ 1 และ ตัวที่ 2 ของตัวอย่างสถานการณ์ที่ 2

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

	มุมมองกล้องวีดิทัศน์ตัวที่ 1	มุมมองกล้องวีดิทัศน์ตัวที่ 2
เฟรมภาพปกติเริ่มต้น		
บุคคลเป้าหมายเดินเข้ามุมมองกล้องวีดิทัศน์ตัวที่ 1		--
กล้องวีดิทัศน์ตัวที่ 1 ทำการส่าย ก้ม เงย และซูมติดตามเป้าหมายและเก็บข้อมูลลักษณะสีและข้อมูลทิศทาง การเคลื่อนที่ไปทางซ้าย		--
สุดขอบเขตการเปลี่ยนมุมมองติดตามของกล้องวีดิทัศน์ตัวที่ 1		--
บุคคลอื่นเคลื่อนเข้ามุมมองกล้องวีดิทัศน์ตัวที่ 2 ก่อนบุคคลเป้าหมาย	--	
แต่การเคลื่อนที่เป็นไปในทิศทาง (เคลื่อนที่ไปทางขวา) ตรงข้ามกับบุคคลเป้าหมาย	--	

รูปที่ 4.35 ลำดับภาพผลการติดตามและระบุบุคคลระหว่างกล้องวีดิทัศน์ตัวที่ 1 และ ตัวที่ 2 ของตัวอย่างสถานการณ์ที่ 3



บุคคลอื่นเคลื่อนที่กลับในทิศทางซ้าย เช่นเดียวกับทิศทางของบุคคลเป้าหมายที่ส่งมาจากกล้องวีดิทัศน์ตัวที่ 1	--	
แต่ไม่ผ่านการตรวจสอบด้วยข้อมูลลักษณะสี ระบุว่าไม่ใช่บุคคลเป้าหมาย และกล้องวีดิทัศน์ตัวที่ 2 ไม่เปลี่ยนมุมมองติดตาม	--	
บุคคลเป้าหมายเริ่มเคลื่อนที่เข้าไปยังมุมมองกล้องวีดิทัศน์ตัวที่ 2 เริ่มทำการระบุบุคคลด้วยข้อมูลลักษณะสีและทิศทางการเคลื่อนที่	--	
บุคคลมีการเคลื่อนที่ไปทางซ้าย และผ่านการตรวจสอบข้อมูลลักษณะสี ระบุว่าบุคคลเป้าหมายเดิม	--	
กล้องวีดิทัศน์ตัวที่ 2 เริ่มทำการส่ายก้มเงย และซูมติดตาม	--	
กล้องวีดิทัศน์ตัวที่ 2 ส่าย ก้มเงย และซูมติดตามบุคคลต่อเนื่อง	--	

รูปที่ 4.35(ต่อ) ลำดับภาพผลการติดตามและระบุบุคคลระหว่างกล้องวีดิทัศน์ตัวที่ 1 และ ตัวที่ 2 ของตัวอย่างสถานการณ์ที่ 3



## บทที่ 5

### สรุปผลการวิจัย และข้อเสนอแนะ

#### 5.1 สรุปผลการวิจัย

วิทยานิพนธ์นี้ได้นำเสนอระบบตรวจหาและติดตามใบหน้าโดยใช้กล้องวีดิทัศน์ PTZ 2 ตัว ทำงานร่วมกัน ซึ่งสามารถแบ่งการทำงานร่วมกันระหว่างกล้องได้เป็น 2 ลักษณะ คือ

- ลักษณะการทำงานร่วมกันแบบที่ 1 การประมวลผลร่วมกันเพื่อสรุปผลการระบุบุคคลจากทั้งสองกล้องวีดิทัศน์โดยใช้วิธีการจับคู่ใบหน้า ด้วยวิธีการวัดความคล้ายด้วยระยะ Bhattacharyya สำหรับจับคู่ภาพใบหน้าที่รับภาพได้กับภาพใบหน้าทั้งหมดในฐานข้อมูล โดยเลือกภาพใบหน้าในฐานข้อมูลที่มีความใกล้เคียงมากที่สุดในการระบุบุคคล ซึ่งจะใช้ประโยชน์ในเรื่องตำแหน่งมุมมองการรับภาพของกล้องวีดิทัศน์ กล่าวคือ สามารถรับภาพใบหน้าบุคคลที่กำลังเคลื่อนที่ที่สนใจได้จากทั้งสองกล้องวีดิทัศน์ที่มีตำแหน่งการวางกล้องวีดิทัศน์ที่ต่างกัน ทำให้มีโอกาสในการรับภาพใบหน้าในมุมมองที่เหมาะสม (มุมมองใบหน้าตรง) ต่อการนำไปใช้ในการจับคู่ใบหน้าอย่างต่อเนื่อง มากกว่าการรับภาพจากกล้องวีดิทัศน์ตัวเดียว นอกจากนี้การที่ประยุกต์ใช้กล้องวีดิทัศน์ PTZ ทำให้สามารถเพิ่มขอบเขตในการตรวจหา มีโอกาสได้ภาพในตำแหน่งมุมมองที่ดีขึ้น รวมทั้งยังสามารถเพิ่มรายละเอียดของใบหน้าได้ด้วยการซูม

- ลักษณะการทำงานร่วมกันแบบที่ 2 การทำงานร่วมกันเพื่อติดตามบุคคลที่เคลื่อนที่ที่สนใจอย่างต่อเนื่องระหว่างกล้องวีดิทัศน์ PTZ ทั้ง 2 ตัว โดยใช้ลักษณะสีของเสื้อผ้าบุคคลที่สนใจและลักษณะทิศทางการเคลื่อนที่ของบุคคลมาใช้ระบุความเป็นบุคคลคนเดียวกัน จะใช้การส่งผ่านข้อมูลจากกล้องวีดิทัศน์ตัวที่ 1 ไปยังกล้องวีดิทัศน์ตัวที่ 2 และมีการเก็บข้อมูลของบุคคลที่สนใจเป็นฐานข้อมูล ประกอบด้วย ภาพใบหน้าและข้อมูลลักษณะสีของเสื้อผ้า เพื่อการจัดเก็บข้อมูลอย่างมีประสิทธิภาพและสามารถนำมาตรวจสอบได้อย่างรวดเร็ว ประโยชน์ที่ได้จากการใช้กล้องวีดิทัศน์ PTZ คือ เพิ่มอาณาบริเวณในการตรวจหาและสอดส่องมากขึ้นกว่ากล้องวีดิทัศน์อยู่กับที่ ซึ่งส่งผลให้สามารถลดจำนวนกล้องวีดิทัศน์ที่ต้องใช้ในระบบ ถ้าจะนำไปประยุกต์ใช้งานจริง

#### 5.2 อภิปรายผลการวิจัย

**ส่วนการตรวจหาใบหน้า** ในการประยุกต์ใช้การตรวจหาด้วยสีผิวร่วมกับการตรวจหาด้วยลักษณะการเคลื่อนที่ ทำให้สามารถลดสัญญาณรบกวนที่เกิดจากวัตถุที่มีสีคล้ายกับสีผิวได้อย่างมาก รวมทั้งการตรวจหาทั้งสองลักษณะสามารถประยุกต์ใช้กับภาพที่รับจากกล้องวีดิทัศน์

ทัศน์ PTZ ได้ โดยไม่ต้องมีขั้นตอนจัดการใด ๆ ก่อนการประมวลผล และมีการประมวลผลที่ไม่ซับซ้อนเหมาะกับงานประยุกต์ในเวลาจริง แต่ข้อเสีย คือ ความไม่แน่นอนของเค้าโครงบุคคลที่ถูกตรวจหาได้ที่เกิดจากข้อจำกัดของการตรวจหาด้วยการเคลื่อนที่ เนื่องจากต้องขึ้นอยู่กับลักษณะการเคลื่อนที่ของบุคคล ณ ขณะนั้น

**ส่วนการถ่ายและกัมเมยกล้องวีดิทัศน์เพื่อติดตามบุคคล** เนื่องจากไม่ต้องประมวลผลขึ้นกับฉากหลัง เพราะใช้การตรวจหาด้วยการเคลื่อนที่และด้วยสีผิว ทำให้สามารถควบคุมและสั่งกล้องวีดิทัศน์ให้ถ่ายและกัมเมยได้อย่างอิสระ ตามตำแหน่งที่ต้องการ โดยความเร็วในการถ่ายและกัมเมยกล้องวีดิทัศน์จะต้องขึ้นอยู่กับผลตอบสนองของคำสั่งของกล้องวีดิทัศน์ที่ประยุกต์ใช้ แต่กล้องวีดิทัศน์ SONY EVID-100 ที่นำมาใช้ในงานวิจัยนี้ ไม่มีปัญหาในส่วนผลกระทบนี้

**ส่วนการตัดส่วนภาพใบหน้าแบบวงรี** ข้อดีของการตัดส่วนภาพใบหน้าแบบนี้ คือ จะลดผลกระทบของฉากหลังได้อย่างมาก สำหรับการนำภาพใบหน้าไปใช้ในการรู้จำ แต่จะเหมาะกับลักษณะใบหน้ามุมมองหน้าตรงมากกว่าหน้าเอียง เนื่องด้วยธรรมชาติของรูปร่างใบหน้า

**ส่วนการจับคู่ใบหน้า** เนื่องจากต้องใช้ภาพใบหน้าที่ได้รับภาพจากกล้องในเวลาจริงและเป็นการรับภาพบุคคลที่กำลังเคลื่อนที่ในตำแหน่งที่ไม่มีกำหนด ทำให้การประยุกต์ใช้ขั้นตอนวิธีการรู้จำใบหน้าทั่วไปไม่เหมาะสม อันเนื่องมาจากตำแหน่งของใบหน้า รวมทั้งผลกระทบด้านแสง ดังนั้นดังที่กล่าวไว้ข้างต้น วิทยานิพนธ์นี้จึงเลือกที่จะใช้การจับคู่ใบหน้าเพื่อหาใบหน้าในฐานข้อมูลที่ใกล้เคียงที่สุด โดยใช้การวัดความคล้ายด้วยระยะ Bhattacharyya ทำให้สามารถเปรียบเทียบใบหน้าได้ในกรณีที่ขนาดไม่เท่ากันอันเนื่องจากการสเกลภาพที่เกิดจากการซูม และตำแหน่งของใบหน้าในภาพไม่จำเป็นต้องอยู่ในตำแหน่งที่กำหนดเท่านั้น แต่ข้อจำกัด คือ ประสิทธิภาพของการจับคู่ที่ถูกต้องจะเพิ่มขึ้น เมื่อเพิ่มจำนวนฐานข้อมูลใบหน้า และจำเป็นต้องเก็บข้อมูลภาพใบหน้าในฐานข้อมูล ณ บริเวณตำแหน่งที่กล้องวีดิทัศน์ตั้งอยู่จริง เพราะจะได้รับภาพใบหน้าที่ความเข้มแสงที่ใกล้เคียงกับความเป็นจริงมากที่สุด

**ส่วนการทำงานร่วมกัน** ในการประมวลผลร่วมกันเพื่อสรุปผลการระบุบุคคล จะเห็นว่าประสิทธิภาพของการจับคู่ใบหน้าจะเพิ่มขึ้นกว่าประสิทธิภาพของกล้องวีดิทัศน์แต่ละตัว เนื่องจากในกรณีบางเฟรมภาพที่กล้องวีดิทัศน์ตัวหนึ่งรับภาพได้ในมุมมองหน้าเอียง ความน่าจะเป็นในการจับคู่ถูกต้องค่อนข้างน้อย แต่ในเวลาเดียวกันกล้องวีดิทัศน์อีกตัวสามารถรับภาพได้ในมุมมองหน้าตรง ซึ่งมีโอกาสในการจับคู่ถูกต้องมากกว่า ทำให้สามารถเพิ่มประสิทธิภาพในการจับคู่ใบหน้า และน้ำหนักในการตัดสินใจผลการจับคู่ก็จะเอนไปทางด้านภาพใบหน้าที่มุมมองหน้าตรงมากกว่า ส่วนในการทำงานร่วมกันเพื่อติดตามบุคคลอย่างต่อเนื่อง การใช้ข้อมูลลักษณะสีอย่าง

เดียว ส่งผลต่อการตัดสินใจในการระบุบุคคล นอกจากผลกระทบทางด้านแสงแล้ว ผลการตรวจหาด้วยการเคลื่อนที่ก็ส่งผลต่อการระบุบุคคลด้วย เนื่องจากค่าสีบริเวณขอบการเคลื่อนที่(ซึ่งเป็นฉากหลังจะรวมอยู่ในบริเวณที่สนใจด้วย ทำให้การเปรียบเทียบข้อมูลลักษณะสีไม่ถูกต้อง) ดังนั้นการใช้ลักษณะทิศทางของการเคลื่อนที่จึงสามารถช่วยเพิ่มค่าอัตราการตัดสินใจได้ ดังผลการทดสอบการติดตามบุคคลที่แสดงในบทที่ 4

**ระบบโดยรวม** ระบบที่นำเสนอในวิทยานิพนธ์นี้ ถูกทดสอบระบบที่การประมวลผลอัตราเฟรม 15 เฟรมต่อวินาที ซึ่งอยู่ในเกณฑ์มาตรฐานของการประมวลผลในเวลาจริง เนื่องจากเลือกใช้วิธีการจับคู่ใบหน้า โดยใช้ลักษณะสี จึงไม่มีปัญหาในด้านเวลาที่ต้องใช้ในการประมวลผล แต่ประสิทธิภาพในการระบุบุคคลไม่มาก เมื่อเปรียบเทียบกับขั้นตอนวิธีการระบุบุคคลวิธีอื่น ยังขึ้นอยู่กับ บริเวณหรือสถานที่ที่ใช้ในการทดสอบระบบ กล่าวคือ เรื่องแสงของภาพใบหน้าในฐานะข้อมูลกับภาพใบหน้าที่รับเข้าต้องมีความใกล้เคียงกัน สำหรับปัญหาในการประยุกต์ใช้ระบบในด้านกล้องวิดีโอ PTZ คือ ต้นทุนด้านราคากล้องวิดีโอที่ค่อนข้างสูง ซึ่งแนวทางในการแก้ปัญหา จำเป็นต้องมีการออกแบบส่วนควบคุมการส่ายและก้มเงยที่สร้างขึ้นมาจะช่วยลดต้นทุนได้อย่างมาก

### 5.3 ข้อเสนอแนะ

- ในการตรวจหาด้วยการเคลื่อนที่จำเป็นต้องมีขั้นตอนวิธีการประมวลผลภาพในการเพิ่มความมั่นใจในเค้าโครงของบุคคลที่จะตรวจหาได้ เพื่อลดผลกระทบเกี่ยวกับขอบของบริเวณการเคลื่อนที่ ซึ่งจะส่งผลให้การเติมจุดภาพระหว่างแถวในภาพมีประสิทธิภาพมากขึ้นด้วย
- ในกระบวนการจับคู่ใบหน้า ถ้าเป็นการจับคู่ใบหน้าที่เลือกใช้ในวิทยานิพนธ์นี้ ควรจะพิจารณาในเรื่องผลกระทบของแสงที่มีผลต่อภาพใบหน้าที่รับภาพได้ รวมทั้งต้องเลือกการตัดส่วนภาพในรูปแบบที่มีประสิทธิภาพมากขึ้น ในกรณีของใบหน้าในมุมมองหน้าเอียง แต่ต้องสามารถลดผลกระทบของฉากหลังได้ด้วย
- ในส่วนการระบุบุคคลระหว่างกล้อง จะขึ้นอยู่กับประสิทธิภาพในการตรวจหาและลักษณะทางแสง ดังนั้นถ้าสามารถพัฒนาในส่วนการตรวจหาและสามารถแยกบริเวณต่าง ๆ ของบุคคลได้อย่างถูกต้องมากขึ้น ก็จะทำให้การระบุบุคคลระหว่างกล้องวิดีโอมีประสิทธิภาพมากขึ้น
- ส่วนการทำงานร่วมกัน สามารถพัฒนารูปแบบของการทำงานร่วมกันในลักษณะงานประยุกต์อื่น ๆ ได้ เช่น ใช้ตำแหน่งการเคลื่อนที่ของกล้องวิดีโอแต่ละตัวมาใช้ประโยชน์ในการระบุตำแหน่งบุคคลเพื่อให้ความถูกต้องในการตรวจหาตำแหน่งบุคคลมากขึ้น

## รายการอ้างอิง

1. Ser-Nam Lim, Ahmed Elgammal, and Larry S. Davis. Image-Based Pan-Tilt Camera Control in a Multi-Camera surveillance. Image-Based Pan-Tilt Camera Control in a Multi-Camera surveillance. Proceedings of the IEEE International Conference on Multimedia & Expo (July 2003) : 645-648.
2. Ser-Nam Lim, Larry S. Davis, and Ahmed Elgammal. 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. M. J. Seow, R. Gottumukkal, D. Valaparla, and K. V. Asari. A Robust Face Recognition System for Real Time Surveillance. Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04) (2004) : 631-635.
4. F. Zuo, and P. H. N. de With. Real-time Embedded Face Recognition for Smart Home. IEEE Transaction on Consumer Electronics (February 2005) : 183-190.
5. Shinji Tsuruoka, Toru Yamaguchi, Kenji Kato, Tomohiro Yoshikawa, and Tsuyoshi Shinogi. A Camera Control Based on Fuzzy Behavior Recognition of Lecturer for Distance Lecture. Proceedings of the IEEE International Fuzzy Systems Conference (2001) : 940-943.
6. Jinman Kang, Isaac Cohen, and Gerard Medioni. Continuous Tracking Within and Across Camera Streams. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (June 2003) : 267-272.
7. Arun Hampapur, Sharat Pankanti, Andrew Senior, Ying-Li Tian, Lisa Brown, and Ruud Bolle. Face Cataloger : Multi-Scale Imaging for Relating Identity to Location. Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance (July 2003) : 849-853.
8. Hitoshi Hongo, Mitsunori Ohya, Mamoru Yasumoto, Yoshinori Niwa, and Kazuhiko 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.

9. T. WADA, and T. MATSUYAMA. Appearance Sphere : Background Model for Pan-Tilt-Zoom Camera. Proceedings of the 13th International Conference on Pattern Recognition (August 1996) : 25-29.
10. Y. Ye, J. K. Tsotsos, K. Bennet, and E. Harley. Tracking a Person with Pre-recorded Image Database and a Pan, Tilt, and Zoom Camera. Proceedings of Third Asian Conference on Computer Vision(ACCV'98) (January 1998) : 567-574.
11. Y-Wen Huang, B-Yu Hsieh, S-Yi Chien, and L-Gee Chen. Simple and Effective Algorithm for Automatic Tracking of a Single Object using a Pan-Tilt-Zoom Camera. Proceedings of the IEEE International Conference on Multimedia & Expo(ICME2002) (2002) : 789-792.
12. R. T. Collins, A. J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, and L. Wixson. A System for Video Surveillance and Monitoring. VSAM Final Report CMU-RI-TR-00-12 (2000) : 1-68.
13. Ismail Haritaoglu, David Harwood, and Larry S. Davis. Active Outdoor Surveillance. Proceedings of the IEEE 10th International Conference on Image Analysis and Processing (1999) : 1096-1099.
14. D. Murray, and A. Basu. Active Tracking. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (July 1993) : 1021-1028.
15. R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain. Face Detection in Color Images. IEEE Transaction on Pattern Analysis and Machine Intelligence (May 2002) : 696-706.
16. N. Soontranon, S. Aramvith, and T. H. Chalidabhongse. Face and Hands Localization and Tracking for Sign Language Recognition. Proceedings of International Symposium on Communication and Information Technologies (ISCIT'04) (October 2004) : 1246-1251.
17. D. Comaniciu, V. Ramesh, and P. Meer. Real-Time Tracking of Non-Rigid Objects using Mean Shift. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (June 2000) : 673-678.
18. K. Nummiaro, E. Koller-Meier, T. Svoboda, D. Roth, and L. Van Gool. Color-Based Object Tracking in Multi-Camera Environments. Proceedings of 25<sup>th</sup> DAGM Symposium on Pattern Recognition (September 2003) : 591-599.
19. W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face Recognition : A



- Literature Survey. UMD CfAR Technical Report CAR-TR-948 (2000).
20. Lee, J. Y., and Yoo. S. I.. An elliptical boundary model for skin color detection. Proceedings of the 2002 International Conference on Imaging Science, Systems, and Technology (2002)
  21. V. Vezhnevets, V. Sazonov, and A. Andreeva. A Survey on Pixel-Based Skin Color Detection Techniques. Proceedings of International Conference on Computer Graphics & Vision (GraphiCon 2003) (September 2003) : 85-92.
  22. J.C. Terrillon, M.N. Shirazi, H. Fukamachi, and S. Akamatsu. Comparative Performance of Different Skin Chrominance Models and Chrominance Spaces for the Automatic Detection of Human Faces in color Images. Proceedings of IEEE Int'l Conference Face and Gesture Recognition (2000) : 54-61.
  23. N. A. Thacker, F. J. Aherne, and P. I. Rockett. The Bhattacharyya Metric as an Absolute Similarity Measure for Frequency Coded Data. Kybernetika (1997) : 363-368.
  24. K. Xu, B. Georgescu, D. Comaniciu, and P. Meer. Performance Analysis in Content-based Retrieval with Textures. Proceedings of 15th International Conference on Computer Vision and Pattern Recognition (September 2000) : 275-278.
  25. M. Bertini, A. Del Bimbo, and W. Nunziati. Player Identification in Soccer Videos. Proceedings of Multimedia Information Retrieval (MIR) (November 2005) : 25-32.
  26. R.C.K Hua, L.C. De Silva, and P. Vadakkepat. Detection and Tracking of Faces in Real-Time Environments. Proceedings of International Conference on Imaging Science, Systems and Technology(CISST'2002) (June 2002).
  27. K. Yachi, T. Wada, and T. Matsuyama. Human Head Tracking using Adaptive Appearance Models with a Fixed-View point Pan-Tilt-Zoom Camera. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000 (2000).
  28. I. Haritaoglu, D. Harwood, and L. S. Davis. W4 : Who? When? Where? What? A Real Time System for Detecting and Tracking People. International Conference on Face and Gesture Recognition (April 1998) : 222-227.
  29. S. Birchfield. Elliptical Head Tracking Using Intensity Gradients and color

- Histograms. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (June 1998) : 232-237.
30. A. Suganuma, S. Tsuruta, and R. Taniguchi. An Automatic Camera System for Distant Lecturing System . IPA97 Conference Publication (July 1997) : 566-570.
31. ถิรพิรุฬห์ ทองคำวิฑูรย์. การลบฉากหลังทางสถิติเชิงพาราเมตริกแบบปรับตัวได้สำหรับการแยกส่วนภาพวีดิทัศน์. วิทยานิพนธ์ระดับมหาบัณฑิตสาขาวิศวกรรมไฟฟ้าคณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2547.
32. ภาสวรรณ สกุลสุจิราภา. การพัฒนาเทคนิคการค้นคืนภาพดิจิทัลโดยการใช้ขอโตคอรีโกลแกรมและการแปลงเวฟเล็ต. วิทยานิพนธ์ระดับมหาบัณฑิตสาขาวิศวกรรมไฟฟ้าคณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2547.
33. Y. Wang, J. Ostermann, and Ya-Qin Zhang. VIDEO PROCESSING AND COMMUNICATION. Prentice-Hall, Inc (2002).
34. R.C. Gonzalez and R. E. Woods. Digital Image Processing. 2 nd ed. New Jersey, United States of America : Prentice hall, 2001.
35. I. Pitas. DIGITAL IMAGE PROCESSING ALGORITHMS AND APPLICATIONS. A Wiley-Interscience Publication (2000).
36. [www.mathworld.wolfram.com](http://www.mathworld.wolfram.com).

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก

สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

### บทความทางวิชาการที่ได้รับการเผยแพร่

1. P. Amnuaykanjanasin, S. Aramvith, and T.H. Chalidabhongse. FACE TRACKING USING TWO COOPERATIVE STATIC AND MOVING CAMERAS, The International Conference on Multimedia and Expo, (July 2005): 1158-1161.
2. พิชัย อำนวยกาญจนสิน, สุภาวดี อรามวิทย์, และ ธนารัตน์ ชลิตาพงศ์ การติดตามใบหน้าโดยใช้กล้องอยู่กับที่และกล้องเคลื่อนที่ทำงานร่วมกัน. การประชุมวิชาการทางวิศวกรรมไฟฟ้าครั้งที่ 28 จังหวัดภูเก็ต, 2 (2548) : 977-980.



สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย

## FACE TRACKING USING TWO COOPORATIVE STATIC AND MOVING CAMERAS

*P. Amnuaykanjanasin and S. Aramvith<sup>\*</sup>, T.H. Chalidabhongse<sup>†</sup>*

<sup>\*</sup>Department of Electrical Engineering  
Chulalongkorn University  
Bangkok 10330 Thailand  
Tel: +66-2218-6909  
E-mail: [Supavadee.A@chula.ac.th](mailto:Supavadee.A@chula.ac.th)

<sup>†</sup>Faculty of Information Technology  
King Mongkut's Institute of Technology  
Ladkrabang  
Bangkok 10520 Thailand  
Tel: +66-2737-2551 Ext.526  
E-mail: [thanarat@it.kmitl.ac.th](mailto:thanarat@it.kmitl.ac.th)

### ABSTRACT

In this paper, we present a new stereo approach for tracking human face by using only two cameras in system. One pan-tilt camera is used for tracking person focused on face. One static camera cooperate with pan-tilt camera are used as a stereo system to estimate face 3D position. We propose to update relative position between cameras to reflect camera moving and the change of relative position. Experimental results shows that our proposed system is able to track one person in camera viewing and can estimate the 3D moving path of interesting person.

### 1. INTRODUCTION

Face detection and tracking are widely interesting research topic and can be applied to many applications such as human surveillance, facial gesture recognition, sign language translation, and human-computer interaction. There have been two main approaches in solving this problem. Early approaches concern with 2D segmentation and tracking; not many features beside 2D shape can be used. This limits the capability of visual modeling from images. Later, many approaches presently more interest in working in 3D which provides more visual cues for detecting and tracking position of the interesting targets.

As the human tracking using visual cues has been interested by many researchers, the problem is extended to ability to collect and analyze the human motion and use it to detect and recognize some interesting events. Temporal and details information of the target is needed. Currently, many researches apply active camera system such as moving and pan-tilt-zoom (PTZ) camera in detecting and tracking interesting target to follow and deliver a fine view of it. Ser-Nam Lim et al. [1] present a wide area surveillance system using multiple co-operative cameras that can be zoomed in to and follow a target. S. Tsuruoka

et al. [2] work with two active camera systems trying to understand what is on the lecture board as well as understand the lecturer gesture in distance learning system. S. Bahadori et al. [3] present a surveillance system based on 3D reconstruction of a museum environment. The system signal an alarm when detecting a person located close to certain museum areas.

There exists a few works which have addressed the issue of 3D in PTZ camera systems. Among these, it includes A. Hampapur et al. [4] which propose the person identification system using multi-scale imaging. H. Hongo et al. [5] apply a set of cooperative fixed cameras and PTZ cameras for face and hand gesture recognition. Both systems consist of four cameras, two fixed stereo cameras for depth recovering and two PTZ for 3D tracking of two objects.

Our work is focused on building a system that can do both estimating 3D position of a person's face and tracking its 3D motion using two PTZ cameras.

The organization of this paper is as follows: Section 2 describes the overall architecture of the proposed system. Section 3 presents the stereo matching technique used in this system. Section 4 proposes a method of managing a moving stereo camera. Section 5 shows the process for camera control. The experimental results are shown in Section 6. Section 7 concludes the paper and addresses future works.

### 2. SYSTEM OVERVIEW

Figure 1 shows the camera setup of the two Sony EVI-D100 cameras for the proposed system. The two cameras have overlapping fields of view and are used for stereo triangulation to estimate 3D position of person. First camera is assigned as a fixed static camera. It is used to detect the person's face and set as a reference face blob to match with candidate face blobs that can detect in the second camera. The second camera is set as a pan-tilt camera that follows the target face while tracking. Both



cameras are calibrated using Bouguet's camera calibration toolbox [8].

The overall architecture shows in Figure 2. In stereo, there are two key steps; stereo matching and triangulation. The first step needs to find correspondence between cameras by match feature points between the left and right images. Our work use human's skin color as a cue for matching. In addition, for static camera, we also use background subtraction to accurately detect the real face blob. The second step, which is the triangulation, is the process of computing to obtain the estimate of 3D position of the person's face. The 3D face position is then used by the active camera control to assign the right pan-tilt parameters to the camera in order to follow the target.

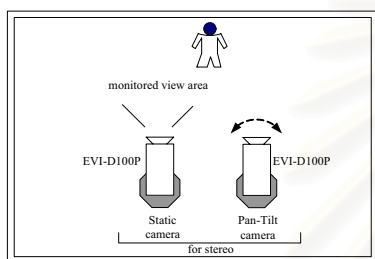


Figure 1. System configuration

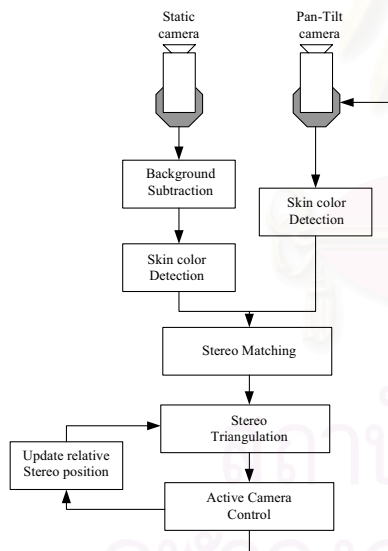


Figure 2. System architecture

### 3. STEREO MATCHING

To find a corresponding point for 3D position estimation, we use color cue to detect face region from the stereo camera images. Various methods, that are used to detect skin, usually differ in reference color space. Experimental study [6, 7] has suggested that the YCbCr color space is a suitable choice for skin color modeling. The YCbCr color space consists of the brightness component (Y) and color

values component (Cb and Cr) so it is capable of detect skin region in changeable brightness environment. We employ method proposed in [6,7] to extract skin color region. The method models skin color with elliptical function on CbCr space. The Y component is discarded to make it invariance to brightness. Figure 3 shows the CbCr color distribution of the human skin color that fits elliptical model.

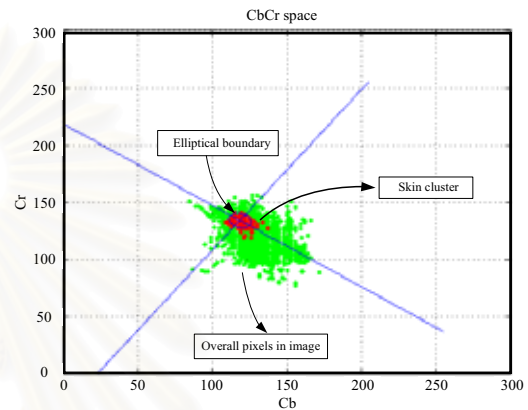


Figure 3. Elliptical model and color distribution

#### 3.1. Face detection in static camera

In static camera, in addition to the skin color, we also apply the background subtraction technique to locate face region. We first subtract image from the background model to discard any background including the skin-color-like background. In this work, we assume that the human's clothes are different from skin color. Figure 4 (a) shows the input image. Figure 4 (b) shows the result of background subtraction and Figure 4 (c) is the final result after applying the skin color detection on Figure 4 (b).

#### 3.2. Blobs candidate search in active camera

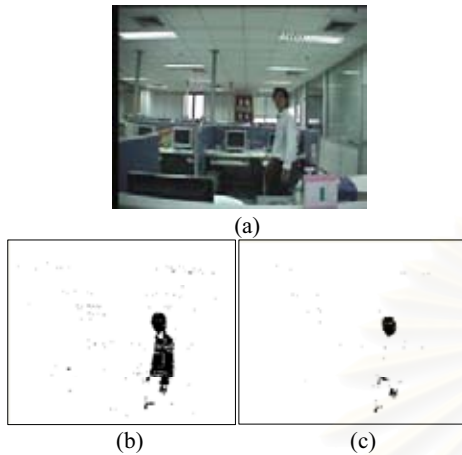
In active camera, the camera's field of view instantly changes by the pan and tilt motion. Basic background subtraction does not work, so we use only skin color detection to detect the face. The result of face detect still include many noises. Thus, to reduce such noises, we apply the affine transform in Eq. (1-2) to calculate the approximate disparity between camera both horizontal and vertical direction when camera pan and tilt.

$$u' = a_0u + b_0v + c_0 + du \quad (1)$$

$$v' = a_1u + b_1v + c_1 + dv \quad (2)$$

where  $(u',v')$  is the estimate position of face blob in active camera.  $(u,v)$  is the center of gravity of the face blob in static camera.  $a_0, a_1, b_0, b_1, c_0$  and  $c_1$  are affine parameters that can pre-calculate from both image in the

initial position of two camera before active camera will moved.  $du$  and  $dv$  are estimate shift pixels that appropriate to pan and tilt angle of each frame.



**Figure 4.** (a) input image (b) foreground result of background subtraction (c) detected face after add skin color detection

#### 4. STEREO SYSTEM MANAGEMENT

The second step in stereo is stereo triangulation. Normally most stereo applications use two static cameras. The relative position between cameras is fixed. This is not the case in our proposed system, since we use one active camera cooperate with the other one static camera. So the relative position between cameras will be changed according to the camera movement. Thus we have to update the cameras' matrix to reflect the camera moving and the change of relative position.

##### 4.1. Update relative matrix between cameras

In the calibrated environment, the two cameras coordinate systems are related by a rotation matrix ( $R$ ) and a translation matrix ( $T$ ). Eq. 3 shows the relation when using the static camera as a reference.

$$X' = R^T X + T \quad (3)$$

$X'$ : active camera coordinate systems  
 $X$ : static camera coordinate systems

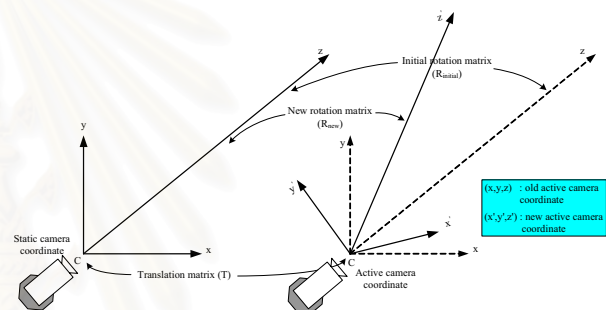
We assume that when active camera changes the field of view, the projection center does not change position so a translation matrix has the same value as in the initial position. A rotation matrix will be changed and need update procedure to find new relative position after active camera moved. Figure 5 shows camera coordinate relation when active camera coordinate changed by pan angle ( $\phi$ ) and tilt angle ( $\psi$ ). Rotation matrix can be changed in 2

cases. First, rotation matrix for pan direction ( $R_{pan}$ ) is updated by Eq. 4. Second, rotation matrix for tilt direction ( $R_{tilt}$ ) is updated by Eq. 5. So the new relative position matrix between cameras ( $R_{new}(t+1)$ ) can update as shown in Eq. 6 and ready to calculate the object depth in next time.

$$R_{pan} = \begin{bmatrix} \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 1 & 0 \\ \sin(\phi) & 0 & \cos(\phi) \end{bmatrix} \quad (4)$$

$$R_{tilt} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi) & -\sin(\psi) \\ 0 & \sin(\psi) & \cos(\psi) \end{bmatrix} \quad (5)$$

$$R_{new}(t+1) = R(t) \cdot R_{pan} \cdot R_{tilt} \quad (6)$$



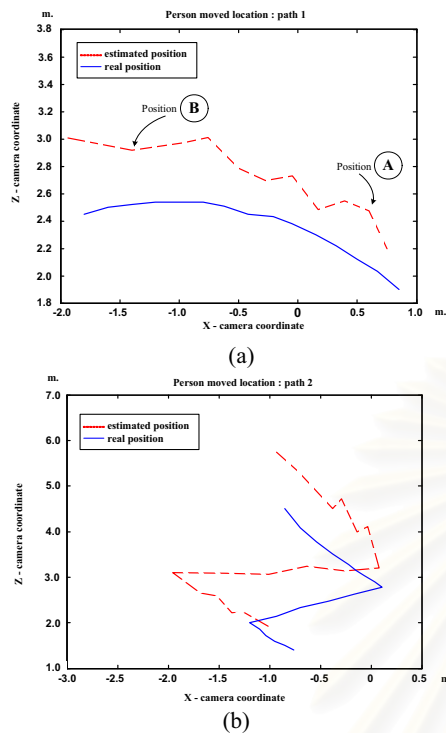
**Figure 5.** Relative position between cameras when active camera moved

#### 5. ACTIVE CAMERA CONTROL

This module is in charge of controlling the pan and tilt parameters of the active camera. The objective of camera control is to maintain the person's face being tracked within center of its camera view. We use 3D object position that get from triangulation for steering the pan-tilt active camera to position the detected face location at the center of the active camera image. The pan and tilt angle are calculated from 3D estimate location

#### 6. EXPERIMENTAL RESULTS

We first modeled the elliptical skin color in CbCr space using a set of 20 person's face image blob. The cameras were then calibrated to obtain camera matrix as well as the relative location between cameras. We marked two trajectory paths on the ground and let the subject walk along. One path was a curve; while the other was a zigzag. Figure 6 shows trajectories of tracked subject (red) versus the groundtruth (blue) plotted on XZ plane (top view). Figure 7 shows the corresponding images of two different frames captured while the subject was moving.



**Figure 6.** Trajectories of tracked moving subject (red) compared to the groundtruth (blue) plotted on XZ plane for both curve path (a) and zigzag path (b).

The 3D Euclidean distance error for the curve path is 0.369 meter while for the zigzag path is 0.836 meter. The results show some location errors however the patterns of the path are quite corresponding to the groundtruth.

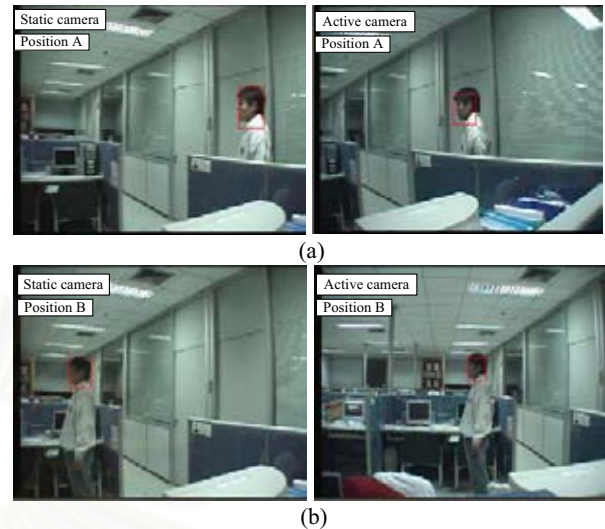
## 7. CONCLUSIONS AND FUTURE WORK

We have presented a new system that can track a person focus on face and can estimate location of a person by using two stereo cameras, one for static camera and another for active camera. Simulation results shows that our system can track person for fine view and can evaluate the 3D person position that has precise 3D coordinate direction but has some error in the same direction.

Currently, we are working on camera zoom that can deliver a multi-scale face image and can be useful to many applications such as face gesture recognition. The multiple cues approach will also be pursued to improve the corresponding points matching.

## 8. ACKNOWLEDGEMENT

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



**Figure 7.** Corresponding images of two different frames (a) and (b) captured while the subject was moving. Left images are taken from static camera. Right images are from the moving camera.

## 9. 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. I-645-8, Jul 2003.
- [2] S. Tsuruoka, T. Yamaguchi, K. Kato, T. Yoshikawa and T. Sninogi, "A Camera Control Based on Fuzzy Behavior Recognition of Lecturer for Distance Lecture", in *IEEE International Conference on Fuzzy Systems*, Hawaii, pp. 940-943, Jan 2001.
- [3] S. Bahadori, and L. Iocchi, "A Stereo Vision System for 3D Reconstruction and Semi-Automatic Surveillance of Museum Areas", in *Workshop on Intelligenza Artificiale per I Beni Culturali*, Pisa, 2003.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face Detection in Color Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 696-706, May 2002.
- [8] Camera Calibration Toolbox for Matlab. Available: [http://www.vision.caltech.edu/bouguetj/calib\\_doc/](http://www.vision.caltech.edu/bouguetj/calib_doc/)

## การติดตามใบหน้าโดยใช้กล้องอยู่กับที่และกล้องเคลื่อนที่ทำงานร่วมกัน

### Human Face Tracking using Two Cooperative Stationary and Moving Cameras

พิชัย อำนวยกาญจนสิน<sup>1</sup> สุภาวดี อร่ามวิทย์<sup>1</sup> และ ธนารัตน์ ชลิตาพงศ์<sup>2</sup>

<sup>1</sup>ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย

254 ถนนพญาไท เขตปทุมวัน กรุงเทพฯ 10330

โทร. 0-2218-6909 E-mail: [Supavadee.A@chula.ac.th](mailto:Supavadee.A@chula.ac.th)

<sup>2</sup>คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

3 หมู่ 2 ถนนฉลองกรุง เขตลาดกระบัง กรุงเทพฯ 10520

โทร. 0-2737-2551 ต่อ 526 E-mail: [thanasarat@it.kmitl.ac.th](mailto:thanasarat@it.kmitl.ac.th)

#### บทคัดย่อ

บทความฉบับนี้นำเสนอ ระบบติดตามใบหน้ามนุษย์โดยใช้กล้องทำงานร่วมกัน 2 ตัว และยังสามารถประยุกต์ใช้เข้ากับหลักการสเตอริโอเพื่อคำนวณหาตำแหน่งใบหน้าในพิกัด 3 มิติ กล้องตัวแรกจะทำหน้าที่เคลื่อนที่ติดตามใบหน้าในลักษณะสายและก้มเงย ส่วนกล้องตัวที่สองจะเป็นกล้องที่มีมุมมองการรับภาพอยู่กับที่ ซึ่งจะต้องทำการประมวลผลสเตอริโอร่วมกับกล้องตัวแรก โดยให้นำเสนอขั้นตอนวิธีการปรับเมตริกซ์ความสัมพันธ์ระหว่างกล้องทั้งสองตัวที่มีความจำเป็นสำหรับการประมวลผลสเตอริโอ ในกรณีที่กล้องตัวแรกมีการเคลื่อนที่เปลี่ยนมุมมองการรับภาพจากตำแหน่งเดิม ซึ่งในส่วนผลการทดลองแสดงให้เห็นว่าระบบสามารถติดตามใบหน้าของบุคคลเดี่ยวภายในมุมมองการรับภาพได้และสามารถประมาณตำแหน่งของใบหน้าใน 3 มิติได้

#### Abstract

In this paper, we present a new stereo approach for tracking human face by using only two cameras in system. One pan-tilt camera is used for tracking person focused on face. One static camera cooperate with pan-tilt camera are used as a stereo system to estimate face 3D position. We propose to update relative position between cameras to reflect camera moving and the change of relative position. Experimental results shows that our proposed system is able to track one person in camera viewing and can estimate the 3D moving path of interesting person.

**Keywords:** stereo, camera control, skin-color, tracking

#### 1. คำนำ

ในปัจจุบันเทคโนโลยีด้านการประมวลผลวิทัศน์ หรือด้านการประมวลผลภาพได้เข้ามามีบทบาทต่อการดำเนินชีวิตของผู้คนมากขึ้น ไม่ว่าจะเป็น ทางด้านระบบรักษาความปลอดภัย ระบบแปลความภาษามือหรือจะเป็นในลักษณะของการโต้ตอบกันระหว่างมนุษย์และคอมพิวเตอร์รวมทั้งงานประยุกต์อื่นๆที่เป็นประโยชน์ กระบวนการหนึ่งที่มีความสำคัญอย่างมาก ก็คือ กระบวนการตรวจหาและติดตามวัตถุที่สนใจซึ่ง

ส่วนใหญ่วัตถุที่สนใจก็จะเป็นใบหน้าของบุคคลที่สนใจ เพื่อนำไปใช้ประมวลผลในขั้นตอนถัด ๆ ไป ซึ่งก็มีนักวิจัยมากมายที่ให้ความสนใจในหัวข้อนี้เพื่อพัฒนาและหาแนวทางการทำงานในรูปแบบใหม่ ในระยะเริ่มต้นการประมวลผลจะมุ่งเน้นพิจารณาจากภาพในพิกัด 2 มิติเป็นส่วนใหญ่ ลักษณะสำคัญที่ถูกนำมาใช้ในการตรวจหาและติดตามจะเป็นลักษณะรูปร่าง ซึ่งอาจจะไม่เพียงพอในลักษณะรูปแบบงานเฉพาะบางอย่าง ทำให้เริ่มมีการทำวิจัยที่มุ่งเน้นไปที่การประมวลผลกับระบบพิกัด 3 มิติ ข้อมูลที่เพิ่มขึ้นมาอีกมิตินี้จะสามารถช่วยเพิ่มประสิทธิภาพทั้งในการตรวจหาและติดตาม รวมทั้งขยายรูปแบบงานประยุกต์ใช้

ซึ่งในการทำงานกับระบบพิกัด 3 มิตินี้ ไม่สามารถประมวลผลโดยใช้กล้องเพียงตัวเดียวได้ จำเป็นต้องประกอบด้วยกล้องที่ทำงานร่วมกันอย่างน้อย 2 ตัว โดยใช้หลักการที่เรียกว่า สเตอริโอ แต่รูปแบบการทำงานที่มีใช้อยู่ในปัจจุบันจะเป็นในลักษณะที่จะรับภาพจากกล้อง 2 ตัวที่มีมุมมองการรับภาพอยู่กับที่ ซึ่งจุดมุ่งหมายจะเน้นเพียงการคำนวณหาตำแหน่งจุดวัตถุที่สนใจในพิกัด 3 มิติ แต่ไม่เน้นที่รายละเอียดของวัตถุ เช่น ในงานวิจัย [1] ทำงานในงานประยุกต์ใช้ในพิพิธภัณฑ์ ประมวลผลสเตอริโอร่วมกันเพื่อหาตำแหน่งบุคคลที่สนใจในพิกัด 3 มิติ สำหรับใช้ตำแหน่งนี้พิจารณาแจ้งเตือนว่ามีการเดินเข้าไปในเขตห้ามหรือไม่ ซึ่งถ้ารูปแบบงานบางอย่างต้องการทั้งสองอย่าง เช่น ต้องการทราบตำแหน่งบุคคลที่สนใจ และต้องการขยายรายละเอียดของใบหน้าให้มากขึ้นเพื่อประโยชน์ในการนำไปประมวลผลต่อไปด้วย ดังนั้นลักษณะรูปแบบการทำงานที่สามารถรองรับได้ คือ ต้องใช้กล้องที่สามารถเคลื่อนที่เปลี่ยนมุมมองการรับภาพได้เคลื่อนที่ติดตามวัตถุ เพื่อให้ได้ตำแหน่งการซูมขยายที่เหมาะสมกว่าตำแหน่งมุมมองกล้องเริ่มต้น

สำหรับในงานวิจัย [2, 3] สามารถทำงานทั้งในด้านการหาพิกัด 3 มิติของบุคคลและสามารถเคลื่อนที่และซูมกล้องไปยังตำแหน่งที่สนใจได้ แต่ระบบต้องประกอบด้วยกล้อง 4 ตัว เป็นกล้องอยู่กับที่ 2 ตัว ทำหน้าที่ประมวลผลสเตอริโอหาพิกัดบุคคลใน 3 มิติและมีกล้องเคลื่อนที่อีก 2 ตัว เพื่อใช้เคลื่อนที่กล้องตามบุคคลโดยใช้พิกัด 3 มิติที่คำนวณได้จากกล้อง 2 ตัวแรก โดยนำไปใช้ในงานประยุกต์สำหรับการระบุบุคคลและการรู้จำท่าทาง ตามลำดับ

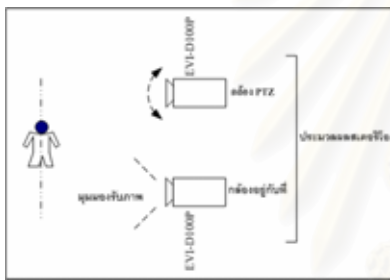
บทความที่นำเสนอนี้ จะมีจุดมุ่งหมายเพื่อสร้างระบบเบื้องต้นที่สามารถประมาณตำแหน่งพิกัดของใบหน้าบุคคลใน 3 มิติ และยังสามารถ



เคลื่อนที่เปลี่ยนมุมมองกล้องติดตามใบหน้าโดยใช้พิกัด 3 มิติที่คำนวณได้ โดยระบบจะประกอบด้วยกล้องเพียง 2 ตัว

**2. ภาพรวมของระบบ**

รูปที่ 1 แสดงรูปแบบของการจัดวางกล้องที่ใช้ในระบบที่นำเสนอ กล้องที่นำมาใช้งาน คือ กล้อง SONY EVI-D100P จำนวน 2 ตัว กล้องทั้งสองตัวจะมีมุมมองการรับภาพเริ่มต้นร่วมกัน โดยกล้องตัวแรกจะถูกกำหนดการทำงานให้เป็นกล้องอยู่กับที่ มีหน้าที่ตรวจหาใบหน้าบุคคลที่สนใจ และใช้ผลลัพท์ของใบหน้าที่ตรวจหาได้เพื่อเป็นล๊อบอ้างอิงสำหรับกล้องตัวที่สอง โดยกล้องตัวที่สองถูกกำหนดให้เป็นกล้องเคลื่อนที่ซึ่งมีหน้าที่ตรวจหาและเคลื่อนที่ติดตามใบหน้า จากนั้นกล้องทั้งสองตัวจะทำการประมวลผลร่วมกันโดยใช้หลักการสามเหลี่ยมสเตอริโอ ซึ่งจะได้เป็นตำแหน่งประมาณของใบหน้าใน 3 มิติ กล้องทั้งสองตัวจะต้องผ่านกระบวนการเทียบมาตรฐานกล้องถ่ายภาพ [6] ก่อนเพื่อให้ทราบถึงความสัมพันธ์ระหว่างทั้งสองกล้อง



รูปที่ 1 รูปแบบของระบบ

กระบวนการทำงานของระบบสามารถแสดงได้ดังรูปที่ 2 สำหรับกระบวนการสเตอริโอจะแบ่งการทำงานออกได้เป็น 2 ส่วนหลัก คือ ส่วนที่ 1. ส่วนของการจับคู่ความสัมพันธ์ซึ่งต้องมีการหาค่าตำแหน่งจุดภาพของวัตถุเดียวกันแต่อยู่ในมุมมองของกล้องแต่ละตัว ลักษณะที่เลือกมาใช้จะเป็นคุณสมบัติทางสีหรือก็คือ ส่วนการตรวจหาด้วยสีผิว และในกรณีของกล้องอยู่กับที่ได้มีการเพิ่มกระบวนการลบฉากหลังเพื่อความถูกต้องในการตรวจหามากยิ่งขึ้น และส่วนการทำงานที่ 2. ส่วนกระบวนการสามเหลี่ยมสเตอริโอใช้ตำแหน่งที่ได้จากส่วนแรกมาคำนวณหาตำแหน่ง 3 มิติของใบหน้า จากนั้นตำแหน่ง 3 มิตินี้ก็จะถูกส่งต่อไปยังส่วนสุดท้ายเพื่อใช้กล้องเคลื่อนที่เคลื่อนที่ไปหาตำแหน่งใบหน้า

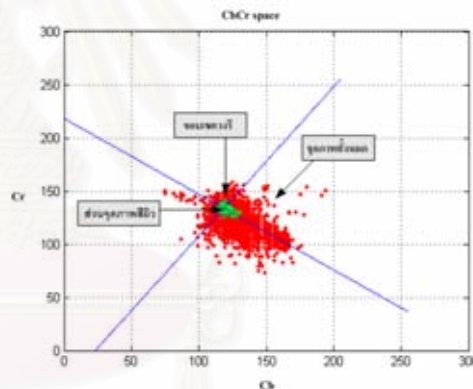
**3. ส่วนการจับคู่ความสัมพันธ์สเตอริโอ**

ในการหาความสัมพันธ์ของตำแหน่งใบหน้าระหว่างกล้องทั้งสองบทความนี้จะเลือกใช้การตรวจหาสีผิวเพื่อใช้หาตำแหน่งใบหน้า ซึ่งเมื่อเลือกใช้คุณสมบัติทางสี สิ่งที่ต้องคำนึงถึง คือ ปริภูมิสีที่เลือกใช้ อ้างอิงงานวิจัย [4] ซึ่งได้ทดลองและสรุปว่า ปริภูมิสี YCbCr เหมาะที่จะนำมาใช้ตรวจหาสีผิว และเมื่อพิจารณาจากองค์ประกอบ ซึ่งปริภูมิสี YCbCr จะแยกองค์ประกอบทางแสงและสีออกจากกัน ทำให้สามารถ

ช่วยลดผลกระทบเกี่ยวกับความสว่างได้ แบบจำลองการตรวจหาที่เลือกใช้จะอ้างอิงตาม [4, 5] ซึ่งจำลองการกระจายตัวขององค์ประกอบสี CbCr เป็นแบบจำลองวงรี ที่ใช้จำแนกว่าจุดภาพนั้นๆ ควรจะเป็นส่วนสีผิวหรือไม่ ดังรูปตัวอย่างแบบจำลองวงรีรูปที่ 3



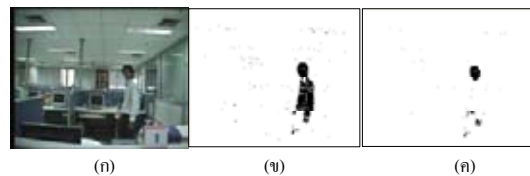
รูปที่ 2 กระบวนการทำงานของทั้งสองกล้อง



รูปที่ 3 แบบจำลองวงรีสำหรับจำแนกส่วนสีผิว

**3.1 การตรวจหาใบหน้าสำหรับกล้องอยู่กับที่**

สำหรับกรณีของกล้องอยู่กับที่ สิ่งต่างจากกล้องเคลื่อนที่ก็คือพื้นฉากหลังที่อยู่กับที่ ดังนั้นนอกจากการตรวจหาส่วนใบหน้าโดยใช้แบบจำลองวงรีตรวจหาแล้ว จะมีการเพิ่มกระบวนการลบฉากหลังเพิ่มเข้ามา ซึ่งทำให้ความถูกต้องที่จะตรวจหาส่วนสีผิวที่เป็นใบหน้าจริงมีมากขึ้น ดังผลการตรวจหารูปที่ 4



รูปที่ 4 (ก) ภาพขาเข้า (ข) ผลการตรวจหาโดยการลบฉากหลัง (ค) ผลการตรวจหาด้วยสีผิวและการลบฉากหลัง



### 3.2 การตรวจหาใบหน้าสำหรับกล้องเคลื่อนที่

กรณีของกล้องเคลื่อนที่ ซึ่งสามารถเปลี่ยนมุมมองการรับภาพได้ การใช้กระบวนการลบฉากหลังจำเป็นต้องมีกระบวนการที่ซับซ้อนมากขึ้น ดังนั้นจึงเลือกใช้เฉพาะการตรวจหาสีผิวด้วยแบบจำลองวงรีอย่างเดียวเท่านั้น แต่เนื่องจากการที่ระบบประกอบด้วยกล้องอีกตัวหนึ่งซึ่งเป็นที่ตั้งกล้องอยู่กับที่ และมีความน่าจะเป็นในการตรวจหาใบหน้าถูกต้องมากกว่า ทำให้สามารถตรวจสอบความถูกต้องของกล้องเคลื่อนที่ได้โดยเปรียบเทียบตำแหน่งใบหน้าที่ตรวจหาได้จริง กับตำแหน่งประมาณที่คำนวณได้จากการแปลงสัมพรรคตามสมการที่ 1 และ 2 ซึ่งอ้างอิงมาจากตำแหน่งลบบที่ตรวจหาได้จากกล้องอยู่กับที่

$$u' = a_0u + b_0v + c_0 + du \quad (1)$$

$$v' = a_1u + b_1v + c_1 + dv \quad (2)$$

เมื่อ  $(u', v')$  คือตำแหน่งจุดศูนย์กลางของลบบใบหน้าประมาณในมุมมองกล้องเคลื่อนที่  $(u, v)$  คือตำแหน่งจุดศูนย์กลางของลบบใบหน้าที่ตรวจหาได้จากกล้องอยู่กับที่ และพารามิเตอร์  $a_0, b_0, c_0, a_1, b_1$  และ  $c_1$  เป็นพารามิเตอร์ของการแปลงสัมพรรค ซึ่งสามารถหาได้จากความสัมพันธ์ของกล้องก่อนการประมวลผล ส่วน  $du$  และ  $dv$  จะหาได้จากมุมสายและมุมก้มเงยของกล้องเมื่อเทียบเป็นหน่วยจุลาภาพ

### 4. กระบวนการสเตอริโอในกรณีที่มีการเคลื่อนที่

กระบวนการสามเหลี่ยมสเตอริโอโดยทั่วไปจะประมวลโดยใช้ข้อมูลจากกล้อง 2 ตัว โดยที่กล้องทั้งสองจะต้องมีมุมมองการรับภาพอยู่กับที่เดิมตลอด แต่ในระบบที่นำเสนอนี้กล้องเคลื่อนที่สามารถเคลื่อนที่เปลี่ยนมุมมองเพื่อติดตามใบหน้าได้ ดังนั้นจึงได้นำเสนอกระบวนการเพื่อปรับเมตริกซ์ความสัมพันธ์ระหว่างกล้องทั้งสองตัว ทุกขณะที่กล้องเคลื่อนที่มีการเคลื่อนที่เปลี่ยนตำแหน่งมุมมองไปจากเดิม

#### 4.1 การปรับเมตริกซ์ความสัมพันธ์ระหว่างกล้องสองตัว

จากกระบวนการเทียบมาตรฐานกล้องถ่ายภาพ จะได้ความสัมพันธ์ระหว่างกล้องสองตัวอยู่ในรูปเมตริกซ์ 2 ตัว คือ เมตริกซ์การหมุน ( $R$ ) และเมตริกซ์การเลื่อน ( $T$ ) ได้ความสัมพันธ์ระหว่างกล้องทั้งสองตัวเป็นดังสมการที่ 3

$$X' = R^T X + T \quad (3)$$

$X'$  เป็นระบบพิกัดกล้องเคลื่อนที่

$X$  เป็นระบบพิกัดกล้องอยู่กับที่

โดยจากสมมติฐานที่ว่าเมื่อกล้องเคลื่อนที่มีการเคลื่อนที่เปลี่ยนมุมมอง แต่ตำแหน่งจุดศูนย์กลางของการฉายภาพ ( $C$ ) ยังคงไม่เปลี่ยนแปลง ดังนั้นเมตริกซ์การเลื่อนจะไม่มีเปลี่ยนแปลง จึงเหลือเพียงเมตริกซ์การหมุน ซึ่งจำเป็นต้องผ่านกระบวนการปรับความสัมพันธ์ทุกครั้งที่มีมุมมองการรับภาพเปลี่ยน เมื่อพิจารณาจากรูปที่ 5 ที่แสดงถึงพิกัดของกล้องทั้งสองตัว ถ้ากล้องเคลื่อนที่ในลักษณะการส่ายด้วยมุม  $\phi$

และเคลื่อนที่ในลักษณะก้มเงยด้วยมุม  $\psi$  ดังนั้นเมตริกซ์การหมุนที่เกิดจากการส่ายจะเป็นดังสมการที่ 4 และเมตริกซ์การหมุนที่เกิดจากการก้มเงยจะเป็นดังสมการที่ 5 จะได้เป็นเมตริกซ์ความสัมพันธ์ระหว่างกล้องทั้งสองตัวใหม่หลังจากที่กล้องเคลื่อนที่เคลื่อนที่ เป็นดังสมการที่ 6

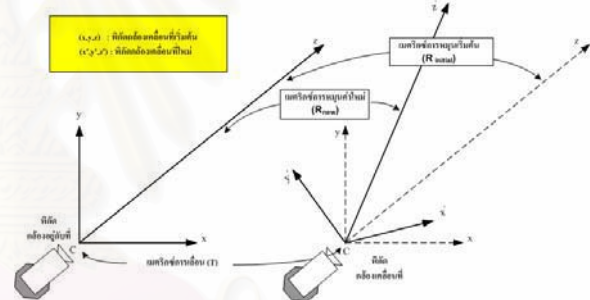
$$R_{pan} = \begin{bmatrix} \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 1 & 0 \\ \sin(\phi) & 0 & \cos(\phi) \end{bmatrix} \quad (4)$$

$$R_{tilt} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi) & -\sin(\psi) \\ 0 & \sin(\psi) & \cos(\psi) \end{bmatrix} \quad (5)$$

$$R_{new}(t+1) = R(t) \cdot R_{pan} \cdot R_{tilt} \quad (6)$$

### 5. การควบคุมกล้องเคลื่อนที่

แนวความคิดหลักของส่วนนี้ก็คือ พยายามเคลื่อนที่กล้องเคลื่อนที่เพื่อให้ได้ตำแหน่งของใบหน้าของบุคคลที่สนใจให้มาอยู่บริเวณกึ่งกลางของภาพที่รับได้ โดยระบบจะใช้ตำแหน่งพิกัดของใบหน้า 3 มิติ ที่ได้จากกระบวนการสเตอริโอมาคำนวณหาตำแหน่งสายและมุมก้มเงยที่ใช้ติดตามใบหน้า



รูปที่ 5 ความสัมพันธ์ระหว่างกล้อง 2 ตัว

### 6. ผลการทดลอง

ในส่วนของการกำหนดขอบเขตของแบบจำลองวงรี จะใช้ตัวอย่างสีผิวของบุคคล 20 คน เพื่อหาขอบเขตของวงรีที่เหมาะสมที่สุดที่ใช้ตรวจหาบริเวณใบหน้า สำหรับการทดสอบระบบ จะกำหนดรูปแบบการเดินของบุคคลที่สนใจเป็น 2 ลักษณะ คือ แบบแรกจะให้เดินในลักษณะแนวคดเคี้ยวและแบบที่สองจะให้เดินเป็นแนวเส้นโค้ง รูปที่ 6.1 และ 6.2 แสดงตัวอย่างของเฟรมภาพ 2 ตำแหน่งจากกล้องอยู่กับที่และกล้องเคลื่อนที่ที่รับภาพขณะบุคคลเคลื่อนที่ ส่วนรูปที่ 7 จะแสดงผลของการทดลองโดยเปรียบเทียบค่าตำแหน่งพิกัด 3 มิติของบุคคลที่คำนวณได้จากกระบวนการสเตอริโอกับค่าที่วัดจากตำแหน่งการเดินจริงของบุคคล

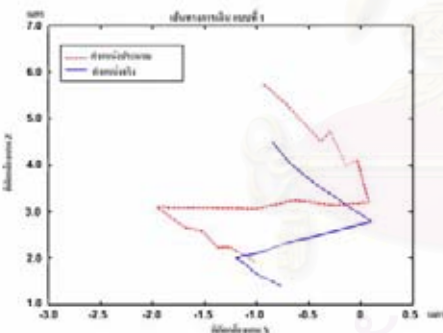
ค่าความผิดพลาดยุคคิดใน 3 มิติที่คำนวณได้จากทั้งสองรูปแบบคือ ในกรณีที่บุคคลเดินคดเคี้ยว มีค่าความผิดพลาด 0.836 เมตร และกรณีที่บุคคลเดินเป็นแนวโค้ง มีค่าความผิดพลาด 0.369 เมตร ซึ่งจะเห็นว่าพิกัด 3 มิติที่คำนวณได้ยังมีค่าความผิดพลาด แต่เมื่อพิจารณาที่เส้นกราฟทางเดินของบุคคล ก็จะเห็นว่ายังคงเป็นไปในทิศทางทางเดินที่ถูกต้อง



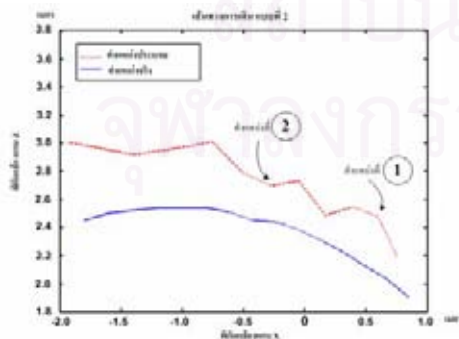
รูปที่ 6.1 ตำแหน่งที่ 1 (ก) ภาพจากกล้องอยู่กับที่ (ข) ภาพจากกล้องเคลื่อนที่ก่อนเคลื่อนกล้อง (ค) ภาพจากกล้องเคลื่อนที่หลังเคลื่อนกล้อง



รูปที่ 6.2 ตำแหน่งที่ 2 (ก) ภาพจากกล้องอยู่กับที่ (ข) ภาพจากกล้องเคลื่อนที่ก่อนเคลื่อนกล้อง (ค) ภาพจากกล้องเคลื่อนที่หลังเคลื่อนกล้อง



(ก)



(ข)

รูปที่ 7 วิธีการเคลื่อนที่ของบุคคลที่สนใจโดยเปรียบเทียบระหว่างตำแหน่งจริง (เส้นตรง) กับตำแหน่งที่คำนวณได้จากระบบ (เส้นประ) (ก) แนวคเคียว (ข) แนวโค้ง

จะเห็นว่าค่าความผิดพลาดที่เกิดขึ้นนั้น เกิดจากผลกระทบในส่วนของการจับคู่ความสัมพันธ์สเตอริโอ ซึ่งในบทความนี้ได้นำเสนอแนวความคิดเบื้องต้น ที่ใช้หลักการตรวจหาสีผิวมาประยุกต์ใช้ในการจับคู่ความสัมพันธ์ โดยไม่ได้มีการนำวิธีการอื่น ๆ มาใช้ร่วม เพื่อให้ง่ายและประหยัดเวลาในการคำนวณ ซึ่งสามารถประมาณหาความสัมพันธ์ระหว่างกล้องได้ความถูกต้องในระดับหนึ่งเท่านั้น ส่งผลต่อค่าความผิดพลาดของระยะทางที่เกิดขึ้นดังแสดงในผลการทดลอง อีกปัจจัยหนึ่งคือ ระยะห่างของบุคคลจากกล้อง เมื่อพิจารณาหลักการสเตอริโอ จะเห็นว่าถ้าบุคคลอยู่ในระยะห่างไกลจากกล้องมากเท่าไร ก็จะส่งผลให้ค่าความผิดพลาดมากขึ้นเท่านั้น ในผลการทดลองกรณีที่บุคคลเดินคเคียว ระยะห่างของบุคคลนั้นห่างจากกล้องมาก ทำให้มีค่าความผิดพลาดมากกว่ากรณีที่บุคคลเดินเป็นแนวโค้ง

### 7. สรุป

บทความนี้เสนอ รูปแบบของระบบที่ประกอบด้วยกล้องอยู่กับที่กับกล้องเคลื่อนที่ทำงานในรูปแบบสเตอริโอร่วมกัน โดยที่ระบบสามารถเคลื่อนที่กล้องติดตามใบหน้า พร้อมทั้งสามารถประมาณตำแหน่งของใบหน้าในพิกัด 3 มิติ ซึ่งจากผลการทดลองก็แสดงให้เห็นว่าระบบสามารถติดตามใบหน้าและคำนวณหาพิกัด 3 มิติของใบหน้าได้

### 8. กิตติกรรมประกาศ

บทความนี้ได้รับการสนับสนุนจาก โครงการเสริมสร้างความเชื่อมโยงระหว่างภาควิศวกรรมไฟฟ้าและภาคเอกชนทางด้านการวิจัยและพัฒนา

### เอกสารอ้างอิง

- [1] S. Bahadori, and L. Iocchi, "A Stereo Vision System for 3D Reconstruction and Semi-Automatic Surveillance of Museum Areas", in *Workshop on Intelligenza Artificiale per I Beni Culturali*, Pisa, 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] 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.
- [5] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face Detection in Color Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 696-706, May 2002.
- [6] Camera Calibration Toolbox for Matlab. Available: [http://www.vision.caltech.edu/bouquetj/calib\\_doc/](http://www.vision.caltech.edu/bouquetj/calib_doc/)

## ประวัติผู้เขียนวิทยานิพนธ์

นายพิชัย อำนวยกาญจนสิน เกิดวันที่ 23 ธันวาคม พ.ศ. 2524 ที่จังหวัด กรุงเทพมหานคร สำเร็จการศึกษาวิศวกรรมศาสตรบัณฑิต (เกียรตินิยมอันดับหนึ่ง) สาขาวิชา วิศวกรรมไฟฟ้าสื่อสาร จากภาควิชาวิศวกรรมไฟฟ้า สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ ในปีการศึกษา 2545 และ เข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชา วิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อปี การศึกษา 2546



สถาบันวิทยบริการ  
จุฬาลงกรณ์มหาวิทยาลัย