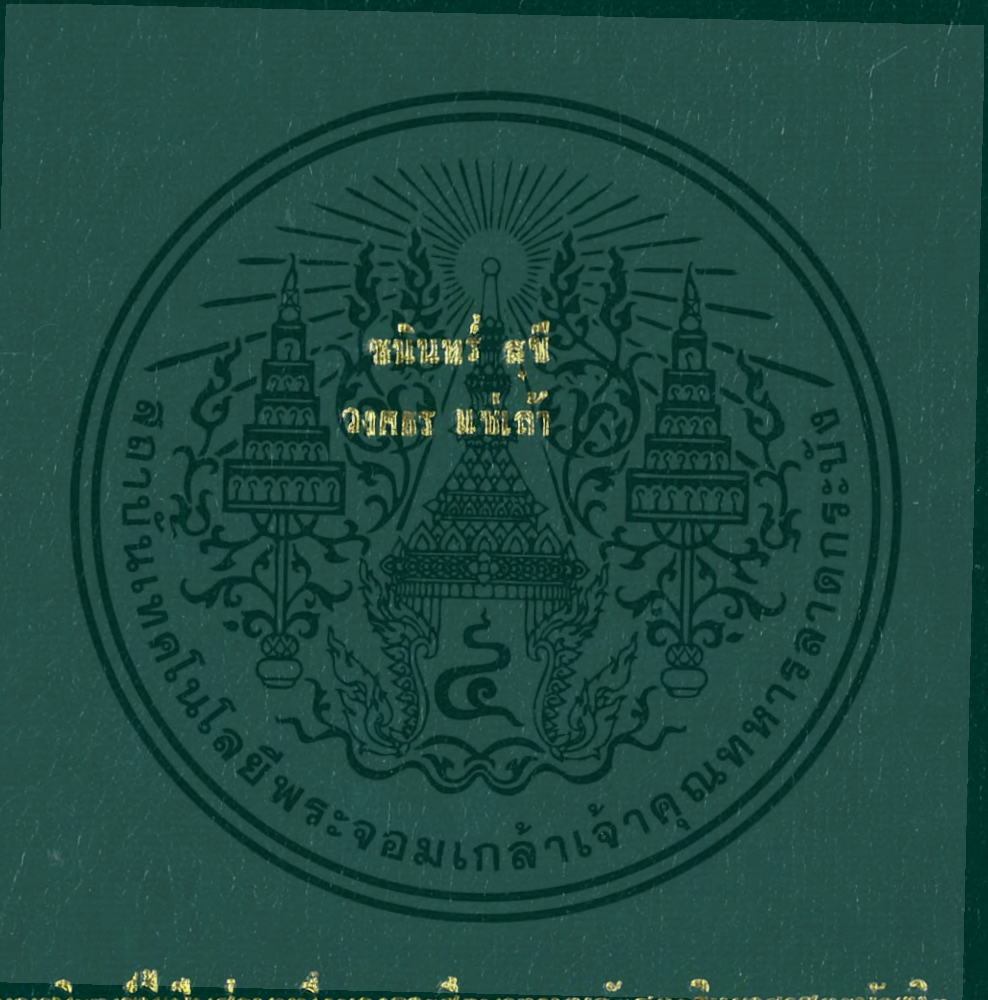
การตรวจจับการเข้าและออกของคนโดยการเข้าติดตาม

TRACKING-BASED HUMAN ENTRY/EXIT DETECTION



ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคเรียนที่ 2 ปีการศึกษา 2558

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การตรวจจับการเข้าและออกของคนโดยการเฝ้าติดตาม

TRACKING-BASED HUMAN ENTRY/EXIT DETECTION

T146242

ชนินทร์ สุขี

CHANIN SUKEE

วงศธร แซ่เล้า

WONGSATORN SAELAO

อาจารย์ที่ปรึกษา

ผู้ช่วยศาสตราจารย์ ดร. สมเกียรติ วังศิริพิทักษ์

เลขทะเบียน 146242
วันเดือนปี 25 เมย 2560

b.12841675
i.

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคเรียนที่ 2 ปีการศึกษา 2558

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

TRACKING-BASED HUMAN ENTRY/EXIT DETECTION

CHANIN SUKEE

WONGSATORN SAELAO

A PROJECT SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR THE DEGREE OF

BACHELOR OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY

FACULTY OF INFORMATION TECNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2/2015

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2016

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใบรับรองปริญญานิพนธ์ ประจำปีการศึกษา 2558

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การตรวจจับการเข้าและออกของคนโดยการเฝ้าติดตาม

TRACKING-BASED HUMAN ENTRY/EXIT DETECTION

ผู้จัดทำ

1. นายชนินทร์ สุขี รหัสนักศึกษา 55070025
2. นายวงศธร แซ่เล้า รหัสนักศึกษา 55070098

..... อาจารย์ที่ปรึกษา

(ผู้ช่วยศาสตราจารย์ ดร. สมเกียรติ วังศิริพิทักษ์)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อโครงงาน	การตรวจจับการเข้าและออกของคนโดยการเฝ้าติดตาม	
นักศึกษา	นายชนินทร์ สุขี	**รหัสนักศึกษา** 55070025
	นายวงศธร แซ่เล้า	**รหัสนักศึกษา** 55070098
ปริญญา	วิทยาศาสตรบัณฑิต	
สาขาวิชา	เทคโนโลยีสารสนเทศ	
ปีการศึกษา	2558	
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร. สมเกียรติ วังศิริพิทักษ์	

บทคัดย่อ

ระบบตรวจจับคนเข้าและออกโดยการเฝ้าติดตามการเคลื่อนไหว พัฒนาขึ้นเพื่อใช้กับกล้องวงจรปิดที่กำลังได้รับความนิยม สามารถแจ้งเตือนได้ทันทีเมื่อตรวจพบการเคลื่อนไหวโดยไม่จำเป็นต้องติดตั้งอุปกรณ์เพิ่มเติม การทำงานจะใช้การสกัดจุดสำคัญ (Feature point) บนภาพวิดีโอด้วยวิธี Feature from Accelerated Segment Test (FAST) จุดสำคัญจากหลายเฟรมก่อนหน้าจะถูกนำมาสร้างเป็น Background Feature Model (BFM) เพื่อใช้เปรียบเทียบหาจุดสำคัญที่เกิดใหม่ (Foreground Feature Points) ในเฟรมปัจจุบัน ระบบจะเฝ้าติดตามจุดสำคัญที่เกิดใหม่โดยเปรียบเทียบหาจุดที่มีความคล้ายคลึงกันในเฟรมถัด ๆ ไป และคำนวณหาระยะการเคลื่อนไหวบนพื้นระนาบ เพื่อวิเคราะห์ว่ามีบุคคลเข้าหรือออกจากพื้นที่เฝ้าระวังหรือไม่ ระบบที่พัฒนาขึ้นสามารถทำงานได้อย่างแม่นยำ ด้วยความเร็วไม่น้อยกว่า 30 เฟรมต่อวินาที

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Project Title	Tracking-Based Human Entry/Exit Detection
Student	Mr. Chanin Sukee **Student ID** 55070025
	Mr. Wongsatorn Saelao **Student ID** 55070098
Degree	Bachelor of Science
Program	Information Technology
Academic	Year 2015
Advisor	Asst. Prof. Dr. Somkiat Wangsiripitak

Abstract

Tracking-based human entry/ exit detection system is developed for a CCTV camera, which is now popular in Thailand. The proposed system is capable of triggering an alarm as soon as motion is detected – without extra equipment. The system first extracts FAST Feature points from image frames. Many points from some previous frames constitute a Background Feature Model (BFM), which is used in finding new Feature points in current image frame. Those new points are tracked by matching with the most similar ones in next frame. A moving distance (of each Feature) on ground plane is then utilized in determination of human entry/exit. The proposed system runs at 30 fps or faster without false alarm and motion detection error.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

โครงงาน "การตรวจจับการเข้าและออกของคนโดยการเฝ้าติดตาม" สำเร็จลุล่วงได้ด้วยความช่วยเหลือของอาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. สมเกียรติ วังศิริพิทักษ์ ที่ให้แนวคิดและอุปกรณ์สำหรับการริเริ่มดำเนินงาน และยังเป็นแรงผลักดันที่ได้เสนอแนะให้คิด วิเคราะห์และทดลองทำโครงงานเรื่องนี้ รวมไปถึงให้คำแนะนำและความรู้ในการพัฒนาระบบ ทางผู้จัดทำขอขอบพระคุณเป็นอย่างยิ่ง ในความเอาใจใส่ดูแล ที่ทำให้โครงงานนี้สามารถพัฒนาขึ้นมาได้ด้วยดี

ชนินทร์ สุขี

วงศธร แซ่เล้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย I

บทคัดย่อภาษาอังกฤษ II

กิตติกรรมประกาศ III

สารบัญ IV

สารบัญ (ต่อ) V

สารบัญ (ต่อ) VI

สารบัญรูป VII

สารบัญรูป (ต่อ) VIII

บทที่ 1. บทนำ 1

1.1 ความเป็นมาและความสำคัญของปัญหา 1

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา 1

1.3 ขอบเขตการวิจัย 2

1.4 ขั้นตอนของการศึกษา 2

1.5 ประโยชน์ที่คาดว่าจะได้รับ 3

1.6 โครงสร้างของวิทยานิพนธ์ 3

บทที่ 2. หลักการและทฤษฎีที่เกี่ยวข้อง 4

2.1 การตรวจหาวัตถุ 4

2.1.1การตรวจหาวัตถุด้วยวิธี Background Subtraction 4

2.1.2 Feature Base Detection 5

2.2 ระบบการติดตามวัตถุต่าง ๆ (Tracking Algorithms) 5

2.2.1 กระบวนการติดตามส่วนแกนกลางวัตถุ (Kernel Tracking) 6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.2.2 กระบวนการติดตามส่วนพื้นผิวของวัตถุ (Silhouette Tracking) 6
2.2.3 กระบวนการติดตามจุดในภาพ (Point Tracking) 6
2.3 ระบบการตรวจจับที่เกี่ยวข้อง 7
2.4 การหา Feature ด้วยวิธี FAST (FAST Features Detection) 8
2.5 Speeded-Up Robust Features (SURF) 9
2.6 Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration 9
2.7 A Flexible New Technique for Camera Calibration 10

บทที่ 3. การดำเนินงานวิจัย 11
3.1 ขั้นตอนการติดตั้งการทำ Camera Calibration 11
3.2 เค้าโครงร่างของกระบวนการตรวจสอบ (System Framework) 14
3.3 Background Feature Model (BFM) 16
3.4 การหา Foreground Feature (Foreground Feature Extraction) 17
3.5 การตรวจจับ Moving Feature (Moving Feature Detection) 18
3.6 รายละเอียดการเลือกและติดตาม Foreground Feature 22
3.7 การตรวจสอบคนเข้าและออก (Detection of Human Entry/Exit) 23
3.8 Window Size of Buffer Zone 24
3.9 Feature matching สำหรับการเปรียบเทียบ Moving Foreground Feature 24

บทที่ 4. การทดลองและผลลัพธ์ 26
4.1 ความไม่เสถียรของ FAST Point Detector 26
4.2 ตัวแปรที่ส่งผลต่อการตรวจจับคนเข้าและออก 28
4.3 Recall Rate อัตราความถูกต้องของเฟรมที่มีคนเข้า/ออก 30

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

4.4 การเปรียบเทียบการทดลองกับระบบ 3 มิติที่พัฒนาเพิ่มขึ้น 32
4.4.1 กราฟแสดงจำนวนเฟรมที่ถูกต้องขณะมีคนในเฟรม 32
4.4.2 กราฟแสดงค่าเฉลี่ยของจำนวน MFF ขณะที่มีคนอยู่ในเฟรม 34
4.4.3 กราฟแสดงค่าเฉลี่ยความยาว MFF สูงสุดขณะมีคนอยู่ในเฟรม 35
4.4.4 กราฟแสดงความถูกต้องของจำนวนครั้งที่ระบบแจ้งเตือน 36
4.5 Screenshots 36

บทที่ 5. บทสรุป 39
5.1 สรุปผลการทดลอง 39
บรรณานุกรม 40
ประวัติผู้เขียน ..... 43.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่ หน้า

3.1 การทำ camera calibration 11

3.2 camera calibration บนมุมภาพที่เปลี่ยนไป 11

3.3 จำลองการสร้างสมการ ground plane 12

3.4 ตัวอย่างภาพสำหรับการสร้างสมการground plane 13

3.5 System Framework 14

3.6 An Image Sequence 16

3.7 เฟรมสำหรับเปรียบเทียบเพื่อหา Foreground Feature 17

3.8 การตรวจจับ Moving Feature (D) 18

3.9 การตรวจจับ Moving Feature (R) 19

3.10 การลากเส้นตรงตัดผ่านสมการ ground plane 21

3.11 ระยะการเลือก Foreground Feature ในเฟรมถัดไป 22

4.1 ตัวอย่างภาพในวิดีโอที่ใช้เก็บผลการทดลอง 26

4.2 จำนวน FAST Feature ใน Resolution ที่ต่างกัน 27

4.2.1 $C_{in(min)}= 0$, $C_{out(min)}= 0$ 28

4.2.2 $C_{in(min)}= 2$, $C_{out(min)}= 2$ 29

4.2.3 $C_{in(min)}= 4$, $C_{out(min)}= 3$ 29

4.3.1 Negative In 30

4.3.2 Positive In 31

4.3.3 Negative Out 31

4.3.4 Positive Out 32

4.4.1 กราฟแสดงจำนวนเฟรมที่ถูกต้องขณะมีคนในเฟรม 33

4.4.2 กราฟแสดงค่าเฉลี่ยของจำนวน MFF ขณะที่มีคนอยู่ในเฟรม 34

4.4.3 กราฟแสดงค่าเฉลี่ยความยาว MFF สูงสุดขณะมีคนอยู่ในเฟรม 35

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่ หน้า

4.4.4 กราฟแสดงความถูกต้องของจำนวนครั้งที่ระบบแจ้งเตือน 36

4.5.1 การติดตาม Feature เมื่อกล้องจับภาพในมุมมอง 3 มิติ 37

4.5.2 ภาพตัวอย่างการทดลองในความละเอียดที่ต่างกัน 38

4.5.3 การเปรียบเทียบการติดตาม Foreground Feature 38

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบัน การใช้กล้องวงจรปิดเพื่อการรักษาความปลอดภัยได้ถูกใช้กันในวงกว้างทั้งในด้านการทดลองเฝ้าสังเกตการณ์และการดูแลรักษาความปลอดภัย โดยการตรวจสอบช่วงเวลาที่วัตถุเริ่มเคลื่อนไหว และทำการคัดแยกช่วงเวลานั้นออกมาเพื่อนำไปประยุกต์ใช้กับงานต่าง ๆ เช่นการตรวจสอบและติดตามลักษณะสิ่งที่เกิดขึ้นขณะวัตถุดังกล่าวเคลื่อนไหว เป็นต้น

ระบบตรวจจับและติดตามที่ถูกพัฒนาขึ้นมามากมาย ได้แสดงให้เห็นถึงความสามารถ การนำไปใช้ และการประยุกต์ที่แตกต่างกันมากมาย ซึ่งโดยส่วนมากแล้วจะประยุกต์ระบบมาจากการทำ Background Subtraction[1] Histogram of Oriented Gradient (HOG) และ Kalman Filter [2] แต่แม้ว่าจะเป็นระบบที่ถูกใช้อย่างแพร่หลาย ก็ยังคงมีปัญหาอยู่ เช่น การทำ Background Subtraction ที่ยังมีปัญหาเรื่องการเปลี่ยนแปลงของแสง และ HOG กับ Kalman Filter ที่ไม่สามารถประมวลผลได้ถึง 30 fps

โดยระบบที่ทำการพัฒนาขึ้นมานี้จะใช้เทคนิค FAST Feature Detector ซึ่งเป็นระบบที่ใช้ในการตรวจหา Feature หรือจุดที่สนใจในภาพเพราะเทคนิคดังกล่าวมีความเร็ว และเหมาะสมสำหรับการประมวลผลแบบ real-time video กว่าการหามุมของเทคนิคอื่น ๆ และทำการเก็บ Feature ที่พบไว้และคอยอัพเดต ซึ่งจะประยุกต์ใช้ร่วมกับการวัดระยะทางในมุมมองแบบ 3 มิติ โดยใช้พิกัดแบบ global coordinate ในการติดตามวัตถุ ซึ่งจากทั้งหมดที่กล่าวมานั้นจะสามารถทำการตรวจจับวัตถุได้อย่างมีประสิทธิภาพและแม่นยำในทุก ๆ มุมภาพแม้แต่ในมุมลึก

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

เนื่องมาจากการใช้กล้องวงจรปิดมีประโยชน์ทั้งในด้านการรักษาความปลอดภัยและการเฝ้าสังเกตการณ์อื่น ๆ ซึ่งโดยส่วนใหญ่เพื่อใช้เป็นหลักฐานเมื่อเกิดเหตุการณ์ต่าง ๆ รวมถึงเพื่อการแจ้งเตือนต่าง ๆ เช่นกัน ทั้งนี้เพื่อให้กล้องสามารถระบุและแจ้งเตือนว่ามีวัตถุหรือบุคคลเข้ามาภายในระยะกล้องนั้นจำเป็นต้องทำการค้นคว้าดังต่อไปนี้

1. ศึกษาและวิจัยวิธีการตรวจจับคนด้วยกล้องด้วยวิธีการที่รวดเร็วและแม่นยำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ศึกษาปัจจัยต่าง ๆ ที่มีผลกระทบต่อระบบตรวจจับคนผ่านกล้อง
3. พัฒนาระบบกล้องให้มีความสามารถในการตรวจจับและติดตามที่ถูกต้องแม่นยำยิ่งขึ้น

1.3 ขอบเขตการวิจัย

ในการกำหนดขอบเขตการศึกษาวิจัยนี้ ผู้วิจัยได้ทำการศึกษาและทดลองโดยอยู่ภายใต้ข้อจำกัดดังต่อไปนี้

1. การวิจัยนี้ใช้ทดลองด้วย กล้อง ที่มีระบบ Fixed Focus และ Fixed White Balanced เพื่อเป็นการจำกัดค่าแสงให้ง่ายต่อการตรวจจับภาพ
2. การวิจัยนี้ทดสอบบนกล้องเว็บแคมที่ต่อผ่าน USB
3. ระบบการตรวจจับนี้ใช้ตรวจจับเฉพาะภายในอาคาร(ไม่มีการเปลี่ยนแปลงของแสงมาก) เท่านั้น
4. ระบบการตรวจจับนี้จะแจ้งเตือนว่ามีวัตถุผ่านเข้ามา เมื่อมีวัตถุเคลื่อนไหวมากกว่าหรือเท่ากับ 1 สิ่ง
5. ภาพที่ทำการทดลองถ่ายวิดีโอจะถูกทดลองจากในมุมสูงเอียง

1.4 ขั้นตอนของการศึกษา

ในการพัฒนาระบบนั้น จำเป็นต้องมีข้อมูลที่จะใช้ในการทดสอบซึ่งเป็นข้อมูลวิดีโอและมีการเก็บข้อมูลบันทึกผลเพื่อหาผลลัพธ์ที่ดีกว่าต่อไป รวมไปถึงการกระยุกต์ใช้รูปแบบการตรวจจับการติดตาม และการคำนวณระยะทางเพื่อหาสมการต่าง ๆ ดังนั้นแล้วขั้นตอนในการดำเนินงานจะมีทั้งหมด ดังนี้

1. รวบรวมข้อมูลภาพวิดีโอในรูปแบบต่าง ๆ เช่น วีดีโอที่ไม่มีคน วีดีโอที่มีคนเดินผ่าน
2. เริ่มต้นการดำเนินการด้วยเขียนโปรแกรมพัฒนาระบบ นำระบบต่าง ๆ มาประยุกต์กับโปรแกรมและวิดีโอ เพื่อทำการเก็บข้อมูลผลลัพธ์
3. สร้างสมการ ตัวแปร หรือ เพิ่มพารามิเตอร์ บางอย่างเพื่อให้ผลลัพธ์มีประสิทธิภาพและแม่นยำมากขึ้น
4. วิเคราะห์ผลลัพธ์และประเมินผลการวิจัย
5. สรุปผลการวิจัย และแก้ไขจุดบกพร่องที่เกิดขึ้น จากนั้นจึงวนทำข้อ 2. อีกครั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 ประโยชน์ที่คาดว่าจะได้รับ

สำหรับการพัฒนาระบบการตรวจจับคนเข้าและออกนี้ ผู้วิจัยมีความคาดหวังว่าระบบการตรวจจับคนเข้าออกที่นำเสนอนี้จะเป็นประโยชน์ และสามารถสร้างความรู้ความเข้าใจในเรื่องระบบประมวลผลภาพและวิดีโอ รวมไปถึงการแก้ไขปัจจัยที่เป็นผลกระทบต่าง ๆ ได้ ดังนี้

1. ได้เรียนรู้แนวทางการวิเคราะห์และปฏิบัติการประมวลผลภาพวีดีโอ
2. ได้รับทราบถึงบทความงานวิจัยที่เกี่ยวข้องกับงานวิจัยนี้
3. สามารถพัฒนาระบบกล้องวงจรปิดให้ตรวจจับคนเข้า/ออกได้อย่างทันที
4. ช่วยในการวิเคราะห์และหาแนวทางแก้ไขปัญหาที่อาจเกิดขึ้นจากตัวแปรในด้านต่าง ๆ
5. สามารถพัฒนาและประยุกต์ใช้ระบบนี้กับงานอื่น ๆ ที่เกี่ยวข้องให้ได้ประสิทธิภาพสูงสุด

1.6 โครงสร้างของวิทยานิพนธ์

วิทยานิพนธ์ฉบับนี้ประกอบด้วยรายละเอียดของเนื้อหาเป็นบทต่าง ๆ ดังนี้

บทที่ 1 บทนำ กล่าวถึงความเป็นมาของงานวิจัย วัตถุประสงค์ของการศึกษา ขอบเขตการวิจัย ขั้นตอนของการศึกษา และประโยชน์ที่คาดว่าจะได้รับ

บทที่ 2 กล่าวถึงทฤษฎีและงานวิจัยอื่น ที่เกี่ยวข้อง

บทที่ 3 กล่าวถึงการดำเนินการของระบบตรวจจับคนเข้า/ออก โดยใช้หลักทฤษฎีที่เกี่ยวข้อง การวัดระยะและคำนวณพื้นที่ในระดับ 3D (global coordinate) และระบบการติดตามด้วย Feature matching

บทที่ 4 กล่าวถึงการทดลองและผลการทดลอง

บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ บทนี้จะสรุปภาพรวมทั้งหมดที่ได้จากการวิจัย ปัญหาที่พบจากการวิจัย พร้อมทั้งได้เสนอแนวทางในการทำวิจัยต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
หลักการและทฤษฎีที่เกี่ยวข้อง

ปัจจุบันได้มีงานวิจัยที่ศึกษาเกี่ยวกับระบบตรวจจับบนวิดีโอเป็นจำนวนมากซึ่งจากข้อมูลทั้งหมดคืองานวิจัยและทฤษฎีที่เกี่ยวข้องที่ได้ถูกนำมาปรับปรุงใช้งานในระบบการตรวจจับวัตถุเข้าและออกซึ่งมีอยู่ดังนี้

2.1 การตรวจหาวัตถุ

ในขณะที่มีการถ่ายภาพวิดีโอตามปกติ ในมุมมองของคอมพิวเตอร์ (computer vision) จะไม่สามารถแยกแยะความแตกต่างของแต่ละสิ่งในภาพได้ จนกว่าจะมีการใช้วิธีการต่าง ๆ ซึ่งถูกกำหนดโดยมนุษย์เพื่อให้คอมพิวเตอร์สามารถคำนวณแยกแยะและตรวจหาวัตถุภายในภาพตามที่ได้มีการออกแบบกำหนดได้

โดยปกติการตรวจหาวัตถุที่ระบบการตรวจจับส่วนใหญ่นิยมใช้กันอย่างกว้างขวางและได้ผลลัพธ์ที่มีประสิทธิภาพจะมีดังต่อไปนี้

2.1.1 การตรวจหาวัตถุด้วยวิธี Background Subtraction

การตรวจจับการเคลื่อนไหวโดยใช้ การหักล้างกับภาพพื้นหลัง (Background Subtraction) [16] [20] โดยขั้นตอนเบื้องต้นนี้จะต้องมีการเก็บภาพพื้นหลัง (Background) ไว้สำหรับใช้ในการเปรียบเทียบกับเฟรมภาพปัจจุบัน เพื่อให้ได้วัตถุที่ไม่ใช่พื้นหลัง (Foreground) และจึงทำการการลดสิ่งรบกวน (Noise) บนผลลัพธ์ที่ได้หลังจากนั้นทำการตรวจสอบว่าวัตถุนั้นมีการเคลื่อนไหวหรือไม่ และต้องทำการคัดแยกสิ่งรบกวน (Noise) พิจารณาจากผลลัพธ์ที่ได้ เนื่องจากอาจมีสิ่งรบกวนที่มีการเคลื่อนไหวอยู่ด้วย แล้วจึงแสดงผลออกมา

- วิธีติดตามวัตถุแบบ Background Subtraction-Based Multiple Object Tracking Using Particle Filter

การใช้ Background Subtraction และการติดตามวัตถุ (tracking) โดยนำ Particle Filter มาประยุกต์ [21] ซึ่งเป็นการติดตาม Feature ที่ใช้ข้อมูลของสี RGB ทั้ง 3 channel ในการประมวลผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประกอบไปด้วย รูปทรงเรขาคณิตแบบดั้งเดิม (สี่เหลี่ยม วงกลม วงรี) ดังนั้นการติดตามตรวจจับจึงถูกดำเนินการได้โดยการเพิ่มความคล้ายคลึงกันของรูปร่างวัตถุภายใน 2 เฟรมภาพ (ภาพเฟรมก่อนหน้าจะถูกทำให้เป็น Appearance Model) และจึงเริ่มประมาณการ การเคลื่อนไหวของวัตถุ เมื่อการเคลื่อนไหวของแบบจำลองถูกสร้างขึ้น ตำแหน่งของวัตถุในเฟรมภาพถัดไป ก็จะสามารถถูกคาดการณ์ล่วงหน้าได้ แต่ ข้อจำกัดของวิธีการดังกล่าวนี้คือความไม่เหมาะสมต่อการ ติดตาม (track) วัตถุที่ไม่มีรูปทรงเรขาคณิตหรือสิ่งมีชีวิตโดยทั่วไป

2.2.2 กระบวนการติดตามส่วนพื้นผิวของวัตถุ (Silhouette Tracking)

Silhouette Tracking หรือการตรวจจับภาพตามส่วนพื้นผิวของวัตถุ โดยปกติแล้วจะติดตามพื้นผิวโดยรวมของวัตถุ ซึ่งมักจะใช้กับวัตถุที่ไม่เป็นรูปทรงและจะมีประโยชน์เมื่อต้องการติดตามวัตถุที่เปลี่ยนรูปร่าง โดยการประมาณการความเปลี่ยนแปลงในรูปร่างหรือเส้นแสดงรูปร่างของวัตถุและปรับปรุงพื้นที่ของวัตถุในแต่ละเฟรมภาพ เพื่อให้การติดตามวัตถุที่ไม่มีรูปร่างแน่ชัดและเปลี่ยนไปเรื่อย ๆ สามารถคาดการณ์และติดตามได้ (ข้อมูลอ้างอิงของ Silhouette Tracking อยู่ใน [10], [11], และ [12])

2.2.3 กระบวนการติดตามจุดในภาพ (Point Tracking)

เมื่อวัตถุถูกตรวจพบโดยวิธีการหา Feature ในแบบต่าง ๆ และตัว Feature ถูกแสดงขึ้นมา วิธีการ Point Tracking ([13], [14], [15]) ก็จะถูกใช้สำหรับติดตามวัตถุ วิธีการนี้จะเป็นประโยชน์ในการติดตามวัตถุที่มีขนาดเล็กมาก (มี 2-3 Feature) อย่างไรก็ตาม สำหรับวัตถุขนาดใหญ่ (ที่มีหลายจุดจากหลายวัตถุ) การจัดกลุ่มจุดในแต่ละวัตถุ ว่าจุดใดเป็นของวัตถุใด จำเป็นต้องใช้การติดตามแบบเฉพาะเจาะจง การเข้ามาและออกไป ของแต่ละวัตถุ สามารถถูกตรวจพบได้โดยการยืนยันการคงอยู่ (การเข้ามา) และการหายไป (ออกไป) ของวัตถุ(หรือ Feature point ของวัตถุ)

แต่จุดที่มีการซ้ำซ้อน หรือตัดกัน จะทำให้เกิดข้อผิดพลาดในการติดตาม ดังนั้น การเพิ่มข้อจำกัดในวิธีการต่าง ๆ ให้สอดคล้องกับวัตถุประสงค์ หรือ ใช้ความไม่แน่นอนในวิธีการทางสถิติมาคำนวณจะสามารถช่วยปรับปรุงการติดตามเช่นนี้ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 ระบบการตรวจจับที่เกี่ยวข้อง

เมื่อสามารถทำการติดตามการเคลื่อนไหวของวัตถุในภาพวิดีโอได้แล้ว ต่อมาจึงสามารถนำรูปแบบการติดตามชนิดต่าง ๆ มาใช้ในการตรวจจับวัตถุที่ต้องการได้ ซึ่งเป็นหลักการและทฤษฎีที่เกี่ยวข้องต่อการวิจัยนี้เป็นอย่างมาก โดยจะมีระบบต่าง ๆ เช่นการตรวจจับของ Harris[18], M.Donoser and H.Bischof [3], E. Rublee, V. Rabaud, K. Konolige, and G. Bradski [4], J.Liu and X. Liang. [5] โดยจะมีการอธิบายระบบอื่น ๆ คร่าว ๆ ดังนี้

- Harris Corner Detector (C. Harris and M.J. Stephens)

ในการมองหาจุดที่สนใจในภาพ การใช้มุมของวัตถุถือเป็นการแก้ปัญหาที่ดี เพราะเป็นจุดเด่น ๆ ที่สามารถมองหาได้ง่ายและมีจำนวนมากในวัตถุที่สร้างโดยมนุษย์ เช่น ประตู โต๊ะ กำแพง ซึ่งระบบตรวจจับ Harris Corner Detector [18] นี้เองที่เป็นต้นแบบของระบบการตรวจจับอื่น ๆ เพิ่มเติมขึ้นมาโดยใช้หลักการตรวจจับมุมของวัตถุเป็นพื้นฐาน โดยวิธีการคำนวณเป็นสมการทางคณิตศาสตร์หาค่าทิศทางในแต่ละทิศทางในพื้นที่รอบ ๆ จุดที่สนใจ โดยใช้สมการของ Taylor ในการหาทิศทาง

- SIFT & SURF (SIFT: D. Lowe & SURF: H. Bay, A. Ess, T. Tuytelaars and L. Van Gool)

เวลาที่ต้องการจับคู่ Feature ระหว่างภาพที่แตกต่างกัน มักจะมีปัญหาเรื่องขนาดภาพที่เปลี่ยนไป นั่นคือภาพที่แตกต่างกันถูกเปรียบเทียบจากระยะทาง (Scale) ที่แตกต่างกัน แต่เมื่อผู้วิจัยพยายามเปรียบเทียบ Feature จาก 2 ภาพนี้โดยใช้ขนาดของพื้นที่รอบข้าง (พื้นหลัง) ที่ไม่เปลี่ยนขนาดไป รูปแบบความเข้มของภาพก็จะไม่เข้าคู่กันได้อยู่ดี เพื่อที่จะแก้ไขปัญหานี้ จึงมีการใช้แนวความคิด Scale-invariant Features เพื่อให้ใช้ตัวแปรที่ควบคุมเรื่องขนาดที่เปลี่ยนไปในแต่ละจุด Feature ที่ถูกตรวจจับได้ ซึ่งในเวลาต่อมาวิธีการ SIFT [19] ก็ได้ถูกพัฒนาให้เป็นวิธี (Speeded Up Robust Features) SURF [20] ซึ่งเป็นวิธีที่มีประสิทธิภาพและรวดเร็วกว่า ซึ่งจะอธิบายในหัวข้อ 2.5

- FAST Features Detection (E. Rosten, R. Porter, and T. Drummond)

ทฤษฎีที่ใช้วิธีการ Harris Corner Detector ในข้อ 2.3.1 เป็นพื้นฐานในการพัฒนาระบบการตรวจจับที่รวดเร็วและมีประสิทธิภาพมากนั่นคือ FAST Feature Detection [2] โดยวิธีการหลักคือการดู 16 จุด รอบ ๆ ว่ามีจุดไหนที่มืดกว่าหรือสว่างกว่าจุดกึ่งกลางแล้วนำมาดูความต่อเนื่องของจุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เหล่านั้นว่าเป็นมุมหรือไม่ โดยความสามารถตรวจจับได้รวดเร็วถึงขั้นสามารถใช้ได้ ณ เวลาขณะที่กล้องกำลังจับภาพในทันที ซึ่งด้วยคุณสมบัติเหล่านี้ จึงเป็นเหตุผลให้ใช้ FAST Feature ในการสร้างระบบการตรวจจับคนโดยการติดตามนี้ซึ่งจะอธิบายต่อในข้อถัดไป

2.4 การหา Feature ด้วยวิธี Features from accelerated segment test (FAST) Features Detection

FAST (Features from accelerated segment test) คือคุณลักษณะที่ได้จากการทดลองแบบแบ่งส่วนหรือก็คือ เป็นวิธีการตรวจจับมุมของสิ่งที่สนใจในภาพ (Feature) ซึ่งสามารถใช้ในการจำแนก Feature points (จุดที่สนใจ) ออกมา อีกทั้งวิธีการนี้มีจุดเด่นทางด้านประสิทธิภาพในการคำนวณ และความรวดเร็วที่มีมากกว่าวิธีการจำแนก Feature ในแบบอื่น ๆ อีกด้วย โดยผู้วิจัยจะเรียกวิธีการนี้ว่า FAST corner detector ซึ่งจากคุณสมบัติทั้งหมดที่กล่าวมา ทำให้วิธีการนี้เป็นวิธีที่เหมาะสมที่สุดในการตรวจจับภาพแบบ Real-time

วิธีการ Detect FAST Feature นั้น จากที่ได้กล่าวว่าเป็นวิธีการตรวจจับมุมภายในภาพ นั่นหมายถึงการตรวจเช็คพิกเซล (pixels) 16 จุด จากจุดกึ่งกลางของพิกเซลที่สนใจ ซึ่งหากพิกเซลที่จุด 1 ถึง 16 มีค่าความมืด/สว่าง มากเกินค่า ความมืด/สว่างของจุด p+threshold หรือ p-threshold (ขึ้นอยู่กับว่า พิกเซลที่จุด 1-16 ดังกล่าวสว่างกว่าหรือมืดกว่า ถ้าสว่างกว่าให้เทียบกับค่า p+threshold และถ้ามืดกว่าก็ให้เทียบกับค่า p-threshold) เมื่อเปรียบเทียบกับที่จุดกึ่งกลาง p เอง แล้วได้ผลลัพธ์ว่าเกินกว่าค่า p +/- threshold มากกว่าหรือเท่ากับ 9 จุด จึงจะสามารถยืนยันได้ว่า จุด p ที่สนใจคือ Feature point ทั้งนี้เหตุที่ผู้วิจัยสังเกตที่ 9 จุด เพราะใช้ FAST ในรุ่น FAST-9 นั่นหมายความว่าหากเป็น FAST-12 ก็จะเทียบที่ 12 จุด

และด้วยการที่ FAST Feature ต้องการตรวจจับ Feature points ให้ได้อย่างรวดเร็วและมีประสิทธิภาพ จึงใช้วิธีการตรวจจับเริ่มจาก 4 จุดแรกคือ จุดที่ 1 5 9 13 และจึงเปรียบเทียบกับจุด p โดยเริ่มการเปรียบเทียบที่จุด 1 9 5 และ 13 ตามลำดับ ซึ่งหากมีไม่ถึง 3 ใน 4 จุดที่มีค่าเกินกว่า p+/-threshold จุด p ก็จะไม่ถือเป็น corner และหากตรวจเช็ค 3 ใน4จุดดังกล่าวได้ว่ามีค่าเกินกว่าจุด p แล้ว จึงจะสามรถตรวจไล่จากจุด 2 3 4 ... จนกว่าจะครบ 16 จุดเพื่อตัดสินว่าจุด p ดังกล่าวเป็น Feature หรือไม่ ซึ่งวิธีนี้จะทำให้สามารถคัดจุดที่เป็น Feature point ได้อย่างรวดเร็วยิ่งขึ้นเพราะ จะสามารถระบุได้ทันทีว่าจุด p ที่ตรวจสอบอยู่ไม่ใช่ Feature point แล้วจึงทำการตรวจสอบจุด p อื่น ๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต่อไป โดยที่เมื่อได้ Feature point ต่าง ๆ เข้ามาในภาพแล้วจะช่วยให้สามารถดำเนินการต่าง ๆ ได้อย่างมีประสิทธิภาพมากขึ้น

2.5 Speeded-Up Robust Features (SURF)

SURF คือระบบที่ได้รับการพัฒนาต่อยอดมากจาก SIFT (ซึ่งได้กล่าวถึงไปก่อนหน้านี้) โดยเพิ่มเติมการนำ Haar Wavelet เข้ามา ซึ่งระบบ SURF นี้จะมีหน้าที่ช่วยในการทำ Object Recognition หรือ 3D reconstruction แต่ตัวระบบจะมีความเร็วสูงกว่า และรับมือกับการเปลี่ยนแปลงของภาพได้ดีกว่า SIFT มาก

ในการตรวจหาจุดที่สนใจ SURF จะใช้การประมาณค่าจำนวนเต็มของ Hessian blob detector โดย Feature descriptor ที่ได้มานั้นจะมาจากผลรวมของค่า Haar Wavelet ที่ตอบสนองอยู่รอบ ๆ จุดที่สนใจ (Feature) ซึ่งนั่นทำให้ SURF descriptor สามารถระบุตำแหน่งและจดจำ object คน หรือ ใบหน้า เพื่อใช้ในการ สร้างภาพ 3 มิติ เพื่อติดตามวัตถุ และ เพื่อคัดแยกจุดที่สนใจออกมา

จากความสามารถที่ SURF ทำได้ จึงทำให้ผู้วิจัยนำมาใช้ในการช่วยติดตาม descriptor ของ Feature ที่เคลื่อนไหว เพราะการตรวจจับนี้จะอยู่ในมุมมอง 3 มิติ เมื่อติดตาม Feature ที่เคลื่อนไหวในมุมมอง 3 มิติ ได้แล้วจะช่วยเพิ่มความสามารถของระบบให้มีความแม่นยำในการติดตามที่ถูกต้องได้แม่นยำมาก ทั้งนี้ การจะสามารถเริ่มกระบวนการติดตามแบบ 3 มิติได้นั้น อาจจำเป็นต้องมีการปรับการตั้งค่าบางอย่างกับกล้องที่ใช้ นั่นคือการทำ Camera Calibration ซึ่งเป็นวิธีการที่จำเป็นในกระบวนการของ 3D computer vision เพื่อใช้ในการคัดแยกข้อมูลตัวชี้วัดต่าง ๆ จากภาพ 2 มิติ ซึ่งจะกล่าวถึงในหัวข้อย่อยต่อไป

2.6 Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration (FANN)

FANN [23] คือวิธีการทำ Feature matching ที่ใช้สำหรับเปรียบเทียบ Feature ในขณะเคลื่อนไหว เพราะเมื่อคนหรือวัตถุทำการเดินย่อมต้องมีการหมุน/เอียง/หัน ทำให้ Feature ที่ติดตามอยู่เกิดการเปลี่ยนตำแหน่ง หากใช้การติดตามใน Feature ที่ใกล้ที่สุดจะทำให้การติดตามดังกล่าวไม่แม่นยำพอเพราะ Feature ที่ใกล้ที่สุดที่ติดตามต่อไปนั้นไม่ใช่ Feature เดิมใน frame ก่อนหน้า ดังนั้นแล้วการนำ FANN มาช่วยในระบบนี้จะทำให้การเปรียบเทียบ Feature มีความถูกต้องมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เนื่องจาก FANN สามารถปรับรูปแบบการใช้งานได้ 2 วิธีตาม dataset และข้อมูลที่ต้องการ คือ อัลกอริทึมการค้นหาลำดับชั้นของ k-means trees (searching hierarchical k-means trees) และวิธีที่ 2 ก็คือการใช้การสุ่ม kd-trees (multiple randomized kd-trees) ซึ่งในการใช้วิธีการทั้ง 2 นี้สามารถครอบคลุมความสามารถในการหา Feature ที่มีลักษณะคล้ายกัน ในข้อมูลที่มีหลากหลายมิติ

FANN ได้มีการจัด library ไว้สำหรับใช้กับ dataset ต่าง ๆ ซึ่งจะมีชื่อว่า Fast Library for Approximate Nearest Neighbors (FLANN) ซึ่งสามารถนำไปใช้กับภาษา Programing ที่หลากหลายได้ ซึ่งได้ถูกนำมาปรับใช้กับระบบการตรวจจับคนเข้า/ออกเพื่อเพิ่มประสิทธิภาพความแม่นยำในการติดตามในมุมมอง 3 มิติ ซึ่งจะกล่าวถึงขั้นตอนดังกล่าวในบทถัดไป

2.7 A Flexible New Technique for Camera Calibration

กระบวนการที่ถูกคิดค้นโดย Zhang [24] โดยเป็นกระบวนการที่มีไว้สำหรับการปรับกล้องให้สามารถคัดแยกข้อมูล 3 มิติจากภาพที่กล้องเห็น (ซึ่งเป็น 2 มิติ) วิธีการดังกล่าวนี้จะต้องให้กล้องทำการจับภาพไปยังภาพแนวระนาบ (อาจเป็นรูปสี่เหลี่ยมหรือวงกลมที่มีลักษณะเฉพาะเหมือนกัน) โดยต้องมีภาพดังกล่าวอย่างน้อยที่สุด 2 ภาพ จากนั้นจึงทำการสร้างแบบจำลองเลนส์ที่ปรับรัศมีภาพ ซึ่งจะใช้การปรับแต่งที่ไม่เป็นเชิงเส้น (nonlinear) บนความเป็นไปได้ที่มากที่สุด โดยทั้งการจำลองทางคอมพิวเตอร์และข้อมูลจริงจะถูกใช้ทดสอบเพื่อดูการวัดระยะของกล้องกับวัตถุจากการทำ calibration ว่ามีระยะทางที่ถูกต้อง ซึ่งจะทำให้กล้องสามารถจับภาพในมุมมองระดับ 3 มิติได้ วิธีนี้จึงถือได้ว่าเป็นวิธีที่ยืดหยุ่นและเหมาะสมกับระบบตรวจจับคนเข้าออกมากที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การดำเนินงานวิจัย

3.1 ขั้นตอนการติดตั้ง

- ในการทำ camera calibration ผู้วิจัยได้นำรูปแบบวิธีการมาจาก Zhang [24] ซึ่งวิธีดังกล่าวนี้เป็นวิธีการที่สะดวกและยืดหยุ่น โดยการนำ pattern ภาพที่เป็นรูปทรงเดียวกันอย่างน้อย 2 ภาพขึ้นไปมาใช้กับกล้อง โดยจะกล้องจะมีการจำจุดมุม ของ pattern ของภาพที่วางไว้ เมื่อกล้องสามารถปรับการทำงานตามภาพที่เห็นในเลนส์ โดยขั้นตอนการทำ camera calibration จะสังเกตได้ดังรูปที่ 3.1 และ 3.2

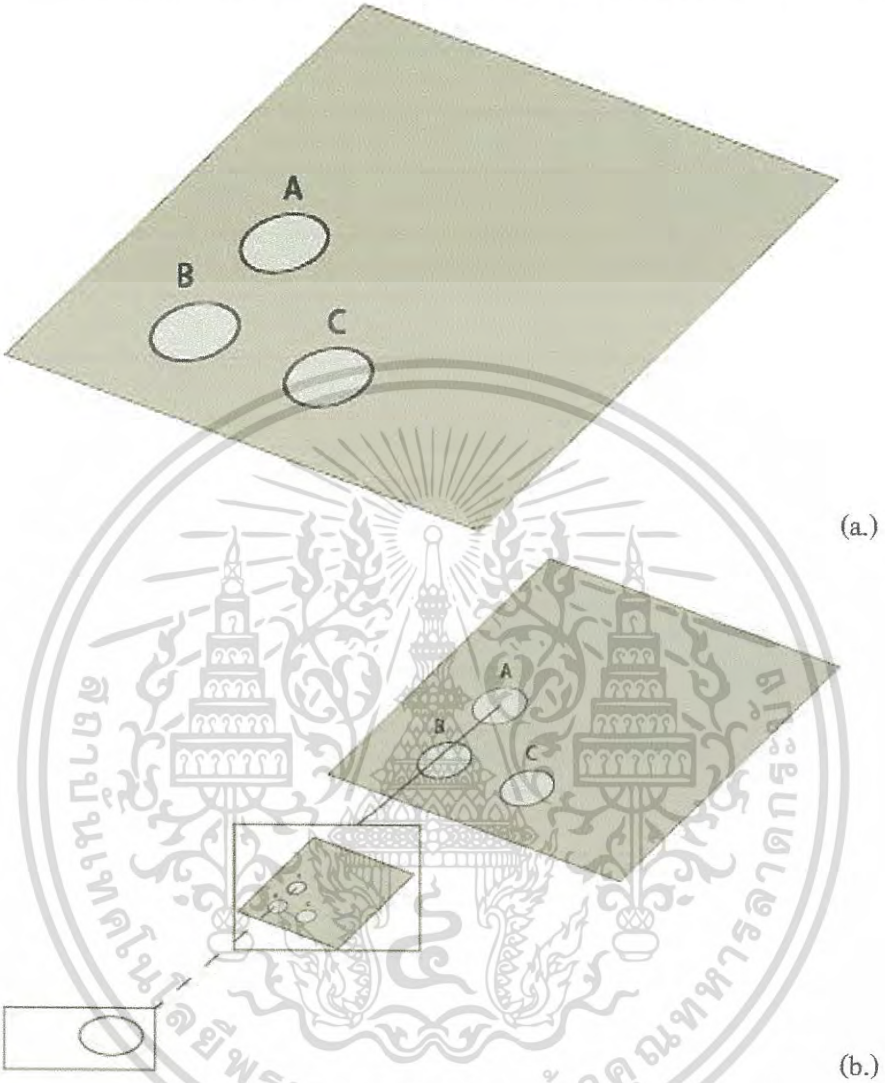
รูปที่ 3.1 การทำ camera calibration โดยกล้องจะจับมุมของแพทเทิร์นสี่เหลี่ยมในรูปตารางหมากรุก

รูปที่ 3.2 เมื่อมีการทำ camera calibration จุดที่อยู่บนตารางหมากรุกจะยังคงเดิมในมุมที่เปลี่ยนไป

- ติดตั้งวงกลมสำหรับการวัดระยะจริง(สำหรับการติดตามแบบ3d)โดยระบบจะสามารถทำการคำนวณระยะทางต่าง ๆ ของวัตถุจากกล้อง โดยเริ่มจากการหาสมการระนาบพื้นผิว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(ground plane) ซึ่งวัดระยะพิกัดแบบ global coordinate โดยใช้จุดตั้งต้น (0, 0, 0)เดียวกับ camera coordinate แทนที่การวัดระยะในระดับ pixel โดยมีวิธีการต่าง ๆ ดังต่อไปนี้



รูปที่ 3.3 จำลองการสร้างสมการ ground plane

ในการทำการ calibrate กล้องนั้น ขั้นตอนแรกจะเริ่มจากการวางแผ่นกระดาษรูปวงกลมอย่างน้อยที่สุด 3 แผ่น บนพื้นโดยแต่ละแผ่นมีระยะห่างเท่า ๆ กัน ต่อมาคือการเริ่มจับภาพที่สามารถเห็นแผ่นวงกลมที่นำไปติดตั้งในฉากได้ แล้วจึงคำนวณลักษณะพื้นผิวจากเฟรมภาพที่ทำการบันทึกด้วยพารามิเตอร์ในกล้องและลักษณะบนแผ่นวงกลมเทียบกับพื้นผิว โดยผู้วิจัยจะใช้แผ่นวงกลม 6 แผ่น วางในรูปแบบเมทริก 2x3 ที่มีความกว้างเท่ากัน โดยให้จุดกึ่งกลางและแกนของกล้องงเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวกำหนดพิกัดบนภาพ กำหนดให้จุด ๆ หนึ่งบนพื้นผิวจริงมีค่า $X = [x, y, z]$ $X_p =$ จุดที่ทราบตำแหน่งบนพื้นผิวแล้ว และ $n =$ เวคเตอร์ n ของพื้นผิวจริง

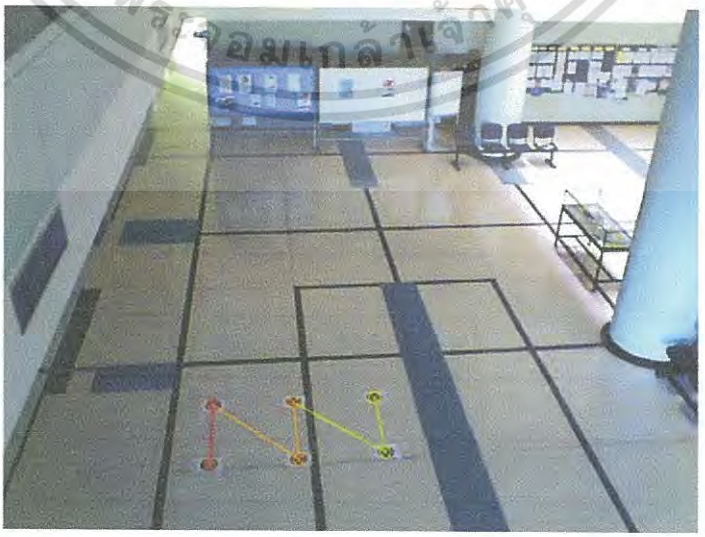
$$(\mathrm{X} - X_p) \cdot \mathrm{n} = 0 \qquad (3.1)$$

โดยจากสมการที่ 3.1 • หมายถึงการทำ dot product โดยพิกัดของกล้องจะถูกใช้เสมือนเป็นพิกัดจริงบนโลกเนื่องจากกล้องได้รับการปรับปรุงจากการทำ calibration พิกัดทั้งหมดในกล้องจะถูกปรับตามพื้นผิวที่กล้องจับภาพอยู่

เมื่อได้สมการพื้นผิวจริง แต่ละตำแหน่งพิกเซล 2 มิติจะสามารถถูกเปลี่ยนเป็นตำแหน่งในพื้นผิว 3 มิติได้โดยการหาจุดตัดระหว่างเส้นที่แสงกระทบจากจุดกึ่งกลางของกล้องไปยังตำแหน่ง 3 มิติ ของจุด Feature กับเส้นพื้นผิวจริง กำหนดให้เส้นแสงกระทบดังกล่าวเป็น เวคเตอร์ L และ Xl เป็นจุด ๆ หนึ่งบนเส้นดังกล่าว ตำแหน่งที่เกิดการตัดกันจึงมีค่าดังสมการที่ 3.2

$$X' = kL + X_l \text{ โดย } \mathrm{k} = \frac{(X_p - X_l)\cdot \mathrm{n}}{L \cdot n} \text{ ซึ่ง } (L \cdot n) \neq 0 \qquad (3.2)$$

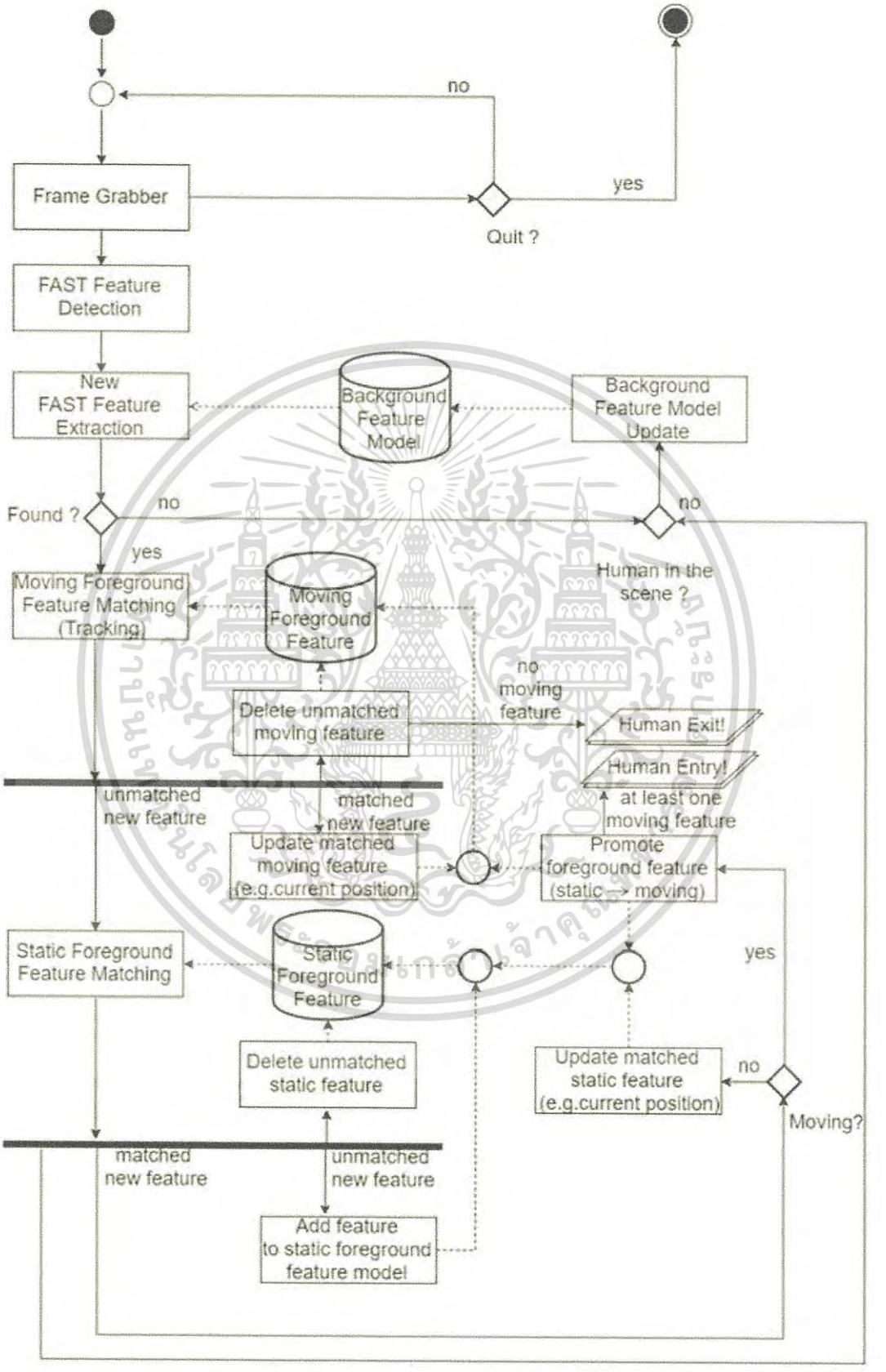
ระบบตรวจจับคนเข้า/ออกนี้จะใช้สมการที่ 2 ในการหาตำแหน่ง 3 มิติของแต่ละ Feature ที่ซึ่งต่อมาจะถูกใช้ในการคำนวณหาระยะทาง 3 มิติเพื่อใช้ในการจำกัดระยะทางการค้นหา Feature R ซึ่งจะกล่าวถึงในหัวข้อย่อยที่ 3.5



รูปที่ 3.4 ตัวอย่างภาพสำหรับการสร้างสมการground plane(ใช้เพียง3จุด)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 เค้าโครงร่างของกระบวนการตรวจสอบ (System Framework)



รูปที่ 3.5 System Framework

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

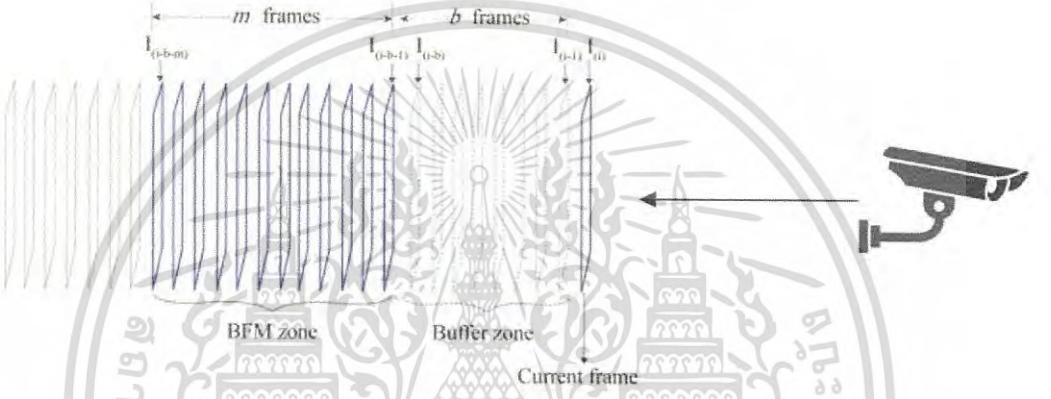
สมมุติให้มนุษย์เป็นวัตถุเดียวที่เคลื่อนไหวในฉาก การจะตรวจพบการเข้า/ออกของมนุษย์จะสามารถทำได้โดยการติดตามจุดที่สนใจในภาพ (tracking point Features) แล้วจึงทำการวิเคราะห์ว่าจุดเหล่านั้นมีการเคลื่อนไหวหรือไม่ ซึ่งจุดที่สนใจนั้นจะใช้วิธีการ FAST ในการตรวจหาจุดดังกล่าวและเก็บค่าตำแหน่งไว้ภายใน Background และเมื่อมี Feature ใหม่เกิดขึ้น ก็จะทำการติดตาม โดยระบบการตรวจจับจะปล่อยสัญญาณบอก “คนเข้า” (human entry) เมื่อ มีจุดที่เคลื่อนไหว (Moving Feature) และจึงทำการติดตามจุดดังกล่าว และเมื่อจุดดังกล่าวได้หายไปจากฉาก ระบบจึงจะปล่อยสัญญาณบอก “คนออก” รายละเอียดของเค้าโครงระบบได้ถูกแสดงไว้ในรูปที่ 3.5

จากในภาพ FAST Features ที่ตรวจพบในฉากซึ่งเป็น Feature ที่หยุดนิ่งไม่มีการเคลื่อนไหวจะถูกเก็บไว้ใน Background โดยจะเรียก Feature ที่ถูกเก็บไว้นี้ว่า Background Feature Model (BFM) ซึ่งสามารถทำการอัพเดตค่าของ Feature ต่าง ๆ ในตัวเองได้อย่างต่อเนื่อง (Feature มีการเคลื่อนไหวตลอดเวลาเนื่องมาจาก noise ในภาพ และการเปลี่ยนแปลงของแสง) และ Feature ใหม่ที่ปรากฏขึ้นจะถูกติดตามการเคลื่อนไหว แล้วจึงถูกนำไปเปรียบเทียบกับ Feature ใน BFM ว่ามีการคลาดเคลื่อนเกินกว่าระยะที่กำหนดหรือไม่ เพื่อใช้หา FAST Features อันใหม่ที่พึ่งปรากฏขึ้นใหม่ ในฉากซึ่งเรียกว่า Foreground Feature แต่ในกรณีที่ไม่มี Foreground Feature เฟรมดังกล่าวจะถูกใช้ อับเดต BFM ต่อไป และถ้า Foreground Feature ถูกพบ มันก็เป็นไปได้ที่จะเป็น (1) หนึ่งใน Moving Foreground Features (MFFs) หรือ (2) หนึ่งใน Static Foreground Features (SFFs) หรือ (3) Foreground Feature อันใหม่ ในกรณีของข้อ (1) ระบบจะแจ้งเตือนว่ามีคนเข้า และ ระบบจะทำการติดตาม Feature ต่อไป แต่ในกรณีของ (3) ระบบจะเพิ่มพิกัดของ Foreground Feature เข้าไปใน List ของ Static Foreground Features (SFFs) เพื่อทำการติดตาม Foreground Feature เหล่านี้จนกระทั่งมันเริ่มทำการเคลื่อนไหว การสังเกตการณ์นี้คือกรณีของข้อ (2)ซึ่งเป็นการตรวจสอบเพิ่มเติมว่า Static Foreground Features (SFFs) ได้เคลื่อนออกจากตำแหน่งดั้งเดิมหรือไม่ มันจะถูกเปลี่ยนให้เป็น Moving Foreground Features (MFFs) ถ้าเริ่มเคลื่อนไหวเกินขอบเขตที่ระบบกำหนดไว้ ระบบจะแจ้งเตือนว่ามีคนเข้า และ ระบบจะแจ้งเตือนว่าคนออกก็ต่อเมื่อ Moving Foreground Features (MFFs) ทั้งหมดที่ถูกติดตามอยู่ ได้หายไปจากฉากแล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 Background Feature Model (BFM)

Background Feature Model (BFM) คือโมเดลที่จัดเก็บพิกัดของ FAST Feature ในแต่ละเฟรมซึ่งในโมเดลนี้จะเก็บจำนวนเฟรมเป็นจำนวน m เฟรม ตามที่ได้เห็นในรูปที่ 3.6 โดย BFM จะเลื่อนไปตามเฟรมปัจจุบัน และจะมีการอัพเดตเรื่อย ๆ ตามขั้นตอนของระบบในรูปที่ 3.5 โดยการอัพเดตจะถูกทำก็ต่อเมื่อ ไม่มีคนเข้ามาในฉาก เพื่อที่ BFM จะเก็บแค่ภาพที่มีเพียง FAST Features ของภาพที่ไม่มีคนเท่านั้น



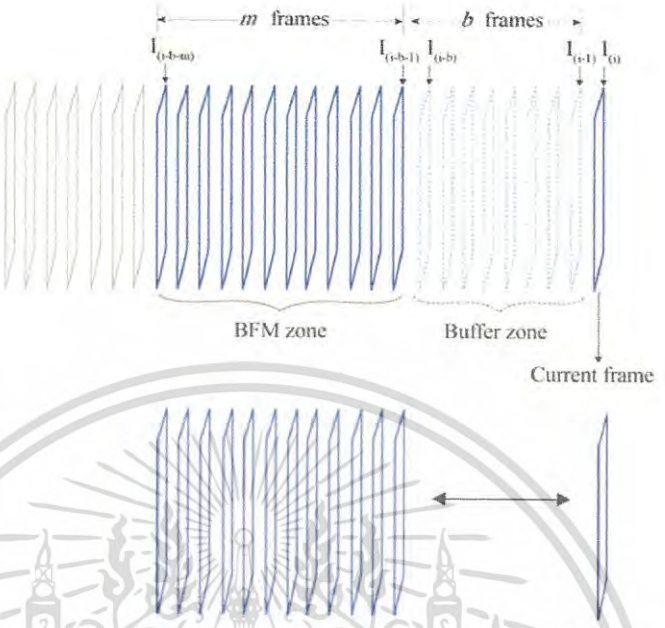
รูปที่ 3.6 An image sequence

ในตอนนี้ ถ้าค่า $I_{(i)}$ เป็นเฟรมปัจจุบัน BFM จะประกอบไปด้วยข้อมูลของ FAST Features ที่อยู่ในเฟรมก่อนหน้า คือ $I_{(i-b-1)}$, $I_{(i-b-2)}$ ไปเรื่อย ๆ จนถึง $I_{(i-b-m)}$ โดยค่า b คือจำนวนของเฟรมภาพที่อยู่ใน Buffer Zone เพราะ การที่ระบบจะตรวจจับคนเข้าได้นั้นต้องใช้ระยะเวลาประมาณหนึ่ง เนื่องจากระบบต้องตรวจสอบการเคลื่อนไหวเป็นระยะทางระยะหนึ่ง (D) ซึ่งจะทำให้ BFM สามารถเก็บเฟรมที่คนกำลังเข้ามาลงไปได้ โดยจำนวนของ b จะอธิบายในบทที่ 3.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

3.4 การหาForeground Feature (Foreground Feature Extraction)



รูปที่ 3.7 แสดงเฟรมที่ถูกนำมาเปรียบเทียบเพื่อค้นหา Foreground Feature

การจำแนก Foreground Feature ในเฟรมปัจจุบัน ถูกทำได้โดยการหา FAST Features ใหม่ โดยการเปรียบเทียบแต่ละพิกัดของ FAST Features ระหว่างเฟรมปัจจุบัน กับ BFM อย่างไรก็ตาม ปัญหาข้อหนึ่งของ การติดตามแบบจุด (Point Tracking) ตามที่ได้กล่าวไว้ก่อนหน้านี้ คือ ค่าผิดปกติ (FAST Features ใหม่ที่เกิดจากความผิดพลาด) ซึ่งจะปรากฏขึ้นเป็นบางครั้ง เมื่อมีการเปลี่ยนแปลงในสภาพแสง หรือระบบตรวจจับภาพเกิดความไม่เสถียร ปัญหานี้โดยปกติแล้วจะเกิดขึ้นเมื่อภาพมีความละเอียด (Resolution) ที่สูงขึ้น

ดังนั้นจำนวนของ Background Frames (ค่า *m*) ใน BFM จะต้องถูกตัดสินใจตามค่าความละเอียดของภาพ ตัวอย่างเช่น จำนวนของ Background Frames จะต้องมากขึ้นใน ความละเอียดที่สูงขึ้น การทำแบบนี้จะช่วยให้รวมค่าผิดปกติที่ไม่สม่ำเสมอไว้ใน BFM ซึ่งจะช่วยลดการตรวจพบที่ไม่ถูกต้องซึ่งไปพบบางจุดที่เกิดจากความไม่เสถียรนี้ว่าเป็น Foreground Feature ในเฟรมปัจจุบัน และใช้ในการติดตามและตรวจจับคนในระบบ ซึ่งผลลัพธ์ของการทดลองที่จะได้อธิบายในบทถัดไปสามารถยืนยันข้อสมมติฐานนี้ได้

Foreground Features ใหม่ ที่พึ่งถูกจำแนกในเฟรมปัจจุบัน จะถูกเก็บค่าไว้ใน Static Foreground Feature Model (SFFM) โดยแบ่งเป็น 2 กรณี คือ (1.) ยังไม่มีคนในฉาก และ (2.) มีคน

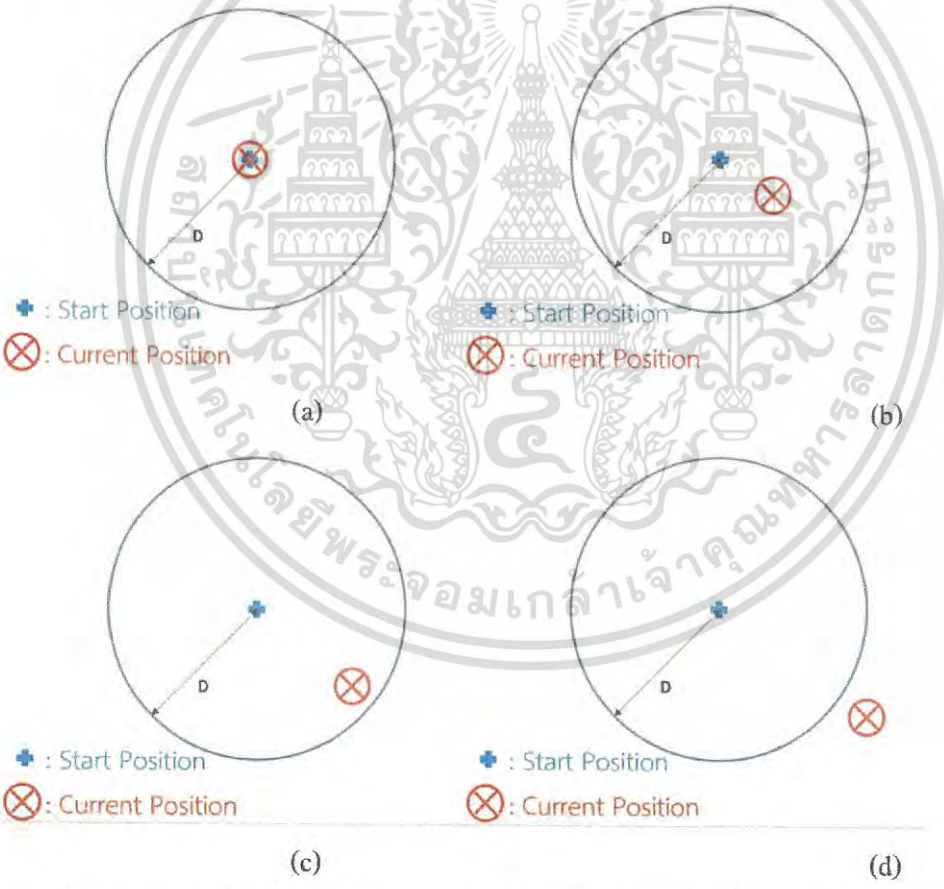
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
146242

ในฉาก และ/หรือ มี Foreground Features ใหม่ที่ไม่อยู่ใน SFFM ซึ่งจุดเหล่านี้จะสามารถเป็น Static Foreground Feature จุดใหม่ หรือ อาจกำลังถูกติดตามจนกลายเป็น Moving Foreground Feature

3.5 การตรวจจับ Moving Feature (Moving Feature Detection)

เมื่อได้ Foreground Feature ใหม่มาแล้ว ผู้วิจัยจะทำการติดตามจุดเหล่านั้นไปเรื่อย ๆ ในแต่ละเฟรมโดยการติดตามจากจุดที่ใกล้ที่สุดไปเรื่อย ๆ (Nearest Neighbor Matching)โดยการติดตามนั้นจะติดตาม Foreground Feature ที่อยู่ในทั้ง Static Foreground Feature Model (SFFM) และ Moving Foreground Feature Model (MFFM) โดยขั้นตอนทั้งหมดนี้จะมีการวัดระยะในระดับ pixel และจะอธิบายการวัดระยะในระดับความยาวจริงบน ground plane ต่อไป

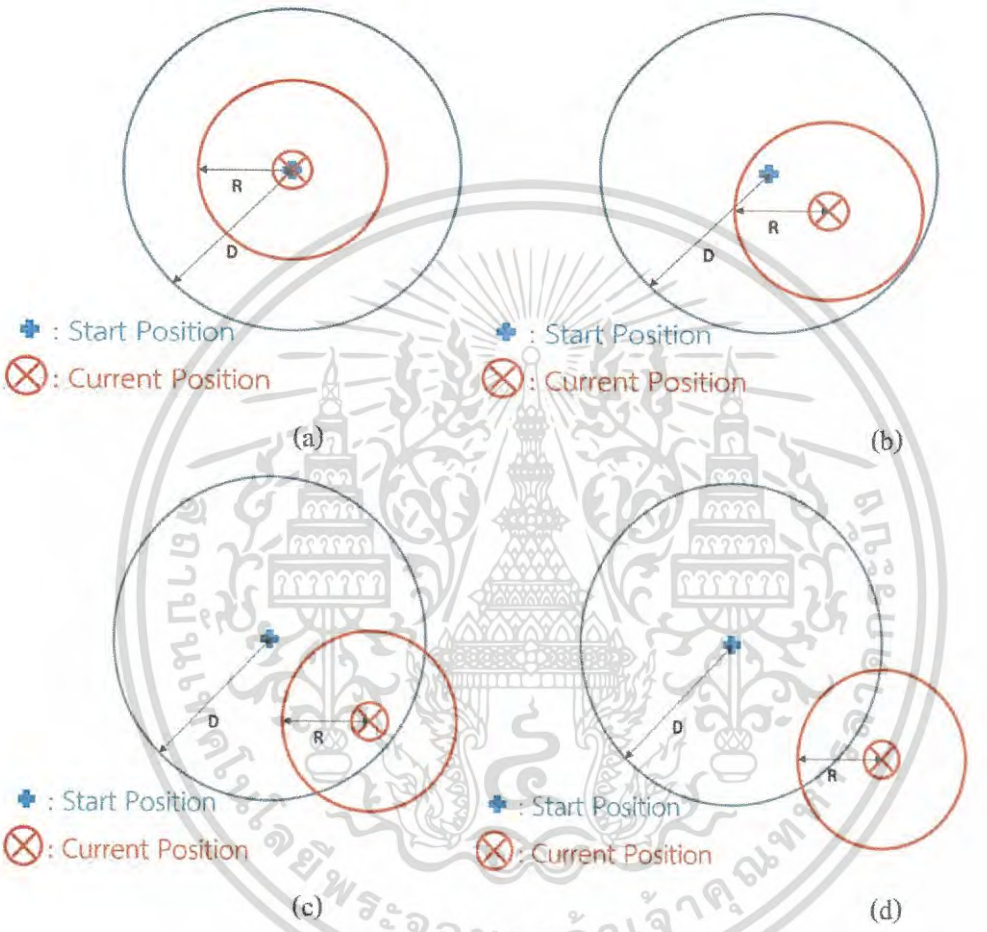
- การติดตามเพื่อตรวจจับ Feature ในระดับ Pixel-Based



รูปที่ 3.8 การตรวจจับ Moving Feature โดย D คือค่าระยะทางที่ไว้คัดแยก Moving Feature + คือตำแหน่งที่ Foreground Features ถูกพบเป็นครั้งแรก Ⓧคือตำแหน่งพิกัดเฟรมปัจจุบัน (Current Frame)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การตรวจจับ Moving Feature ทำได้โดยการติดตาม Foreground Feature ได้เรื่อย ๆ จนกว่าจะเลื่อนที่ออกจากจุดตั้งต้นเป็นระยะทางเกินระยะ D ซึ่งจากรูปที่ (3.8a, 3.8b, 3.8c) คือการเคลื่อนที่อยู่ภาพในระยะ D โดยทั้ง 3 รูปนี้ ระบบจะยังเห็นว่าจุดนั้นเป็น Static Foreground Feature แต่เมื่อเคลื่อนที่ออกจากระยะ D (ดังรูปที่ (3.8d))



รูปที่ 3.9 การตรวจจับ Moving Feature

กำหนดให้ R คือรัศมีสำหรับการค้นหาและติดตาม Foreground Feature D คือค่าระยะทางที่ไว้คัดแยก Moving Feature + คือตำแหน่งที่ Foreground Features ถูกพบเป็นครั้งแรก ⊗ คือตำแหน่งพิกัดเฟรมปัจจุบัน (Current Frame) ระยะค้นหารัศมี R ถูกใช้ค้นหา คู่ที่ตรงกัน ตามที่เห็นในรูปที่ 3.9 เพื่อเพิ่มความเร็วในกระบวนการและลดความเป็นไปได้ของการจับคู่ที่ผิดพลาดและรวดเร็วต่อการค้นหาจุด Foreground รอบ ๆ เพื่อการติดตามโดยที่ ค่าของ R จึงถูกตัดสินใจโดยการนำค่าเข้าสู่สูตร (1.) ค่าความเร็วที่สูงที่สุดที่มนุษย์จะเดินได้ (v_{max}) (282.4 cm/s ซึ่งได้ถูกสำรวจไว้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แล้วใน [6]) (2.) ระยะโฟกัส (f) และ frame rate (r_{camera}) ของกล้อง และ (3.) ระยะห่างระหว่างกล้องกับฉากภาพ (d_{scene}) ซึ่งในทางปฏิบัติ ค่าของ r ในหน่วยของพิกเซล จะเปลี่ยนไปโดยขึ้นอยู่กับ ความละเอียดของกล้อง

$$R = \frac{(V_{max})(f)}{(d_{scene})(r_{camera})} \tag{3.3}$$

โดยสรุปเมื่อมี Foreground Feature ใหม่เข้ามาในระบบ จะทำการจับคู่กับ SFFM โดยที่ SFFM จะเลือกพิกัดที่ใกล้กับตัวเองที่สุดและเรียกจุดนั้นว่าจุดปัจจุบันของตนเอง (Current Position) และเมื่อทำการเลือกจุดต่อไปเรื่อย ๆ ในเฟรมต่อ ๆ ไปจนจุดแรกที่พบ (Start Position) มีระยะห่างมากกว่า D ซึ่งจะเปลี่ยน Static Foreground Feature จุดนั้นให้กลายเป็น Moving Foreground Feature และทำการย้ายพิกัดตั้งต้น (Start Position) และ พิกัดปัจจุบัน (Current Position) ไปเก็บไว้ใน Moving Foreground Feature Model (MFFM) (สังเกตได้จากรูปที่ 3.9c)

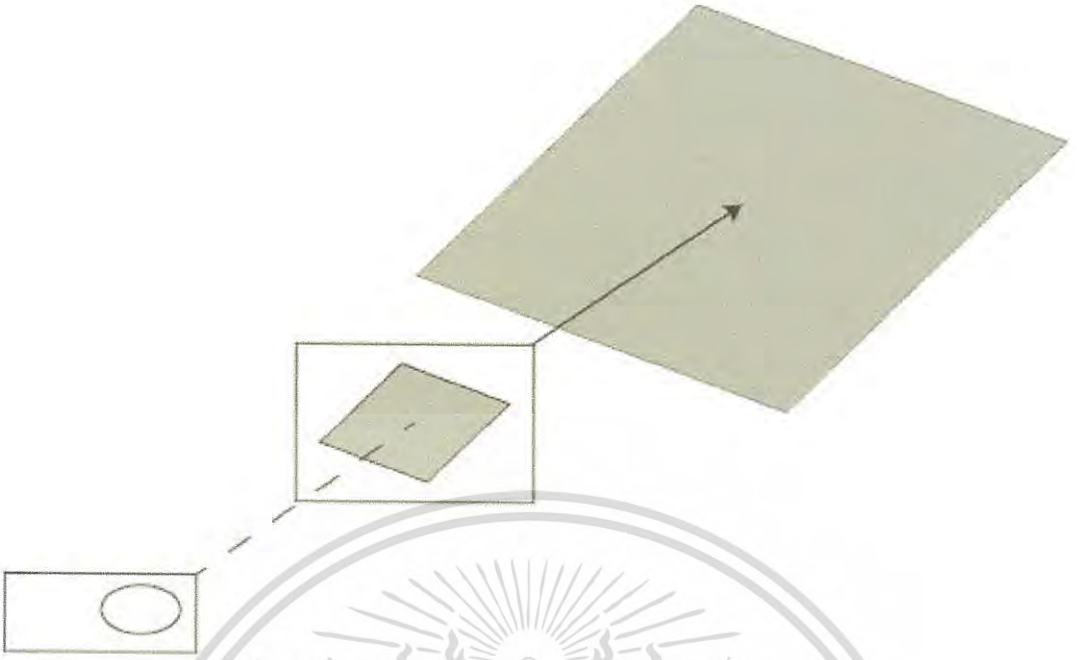
*ระยะห่างเริ่มต้น D สำหรับ Moving Feature จะต้องถูกเลือกอย่างระมัดระวัง เพื่อที่จะหลีกเลี่ยงการตรวจจับที่ผิดพลาด และสามารถรับรู้การเคลื่อนที่ของ Feature ในทันทีที่เป็นไปได้ ผู้วิจัยจึงได้กำหนดค่า D=2R

แต่ทั้งนี้เนื่องจากการวัดระยะทั้งหมดถูกกระทำในระดับ pixel แต่ระบบที่ได้ทำการ camera calibration ให้สามารถหาสมการพื้นผิว ground plane เพื่อใช้ในการวัดในระยะจริงที่ให้ความถูกต้องแม่นยำที่สูงกว่าในมุมมอง 3 มิติ ซึ่งจะอธิบายในหัวข้อด้านล่าง

- การติดตาม Feature ในการวัดระยะจริง (3D)

หลังจากที่ได้พิกัดบน image coordinate ผู้วิจัยจะสามารถสร้างเส้นตรงตัดผ่านระหว่างจุด (0, 0, 0) ผ่านพิกัดบน image coordinate ไปยังเส้นสมการ ground plane ซึ่งจะทำให้ระบบรู้ถึงพิกัด global coordinate ของ Foreground Feature เพื่อใช้ในการวัดระยะทาง 3 มิติได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.10 จำลองการลากเส้นตรงตัดผ่าน สมการ ground plane

โดยตำแหน่งต่าง ๆ บน ground plane ในระยะจริงที่ได้จากการคำนวณและการนำ global coordinate ที่นำมาใช้เหล่านี้จะยังถือว่าเป็นค่าที่ได้จากการประมาณการ แต่ถือว่ามีความใกล้เคียงมากที่สุด ณ ขณะนี้

จากการประยุกต์ใช้ global coordinate และการ calibrate camera ทำให้เกิดการปรับปรุงในระบบให้กล้องสามารถดำเนินการระยะทางในมุมมอง 3 มิติ เพื่อใช้ติดตาม Feature ให้มีความถูกต้องแม่นยำมากขึ้นได้ ทำให้ต้องมีการเปลี่ยนวิธีการติดตาม Feature จากเดิม (ในการติดตามระดับ pixel หรือ 2 มิติ) ที่ใช้วิธีติดตามเฉพาะแต่ละ Feature เฉพาะจุดที่อยู่ใกล้ที่สุดไปเรื่อย ๆ ซึ่งให้ผลลัพธ์ที่ดีในการตรวจจับการเคลื่อนไหวในแนวนอนแต่ประสิทธิภาพของระบบจะลดลงเมื่อมีคนเข้ามาในมุมลึก ดังนั้นการจะสามารถติดตาม Foreground Feature ในมุมมอง 3 มิติได้ จึงต้องนำ global coordinate เข้ามาช่วยในการวัดระยะ จากนั้นจึงจะสามารถเริ่มขั้นตอนการตรวจสอบคนเข้าออกได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.6 รายละเอียดการเลือกและติดตาม Foreground Feature

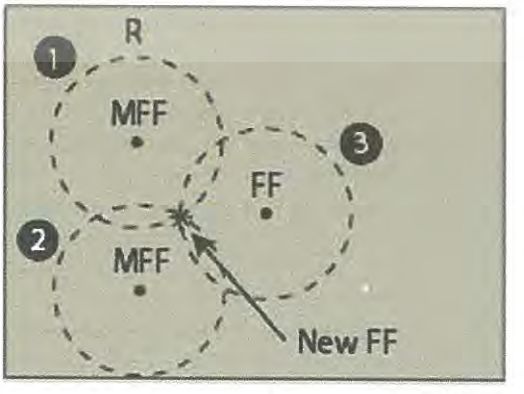
การติดตาม Foreground Feature จะทำการเลือกจากค้นหา Foreground Feature ที่เกิดขึ้นในระยะ R โดยเลือกพิกัดของ Foreground ใหม่ที่ใกล้ที่สุดกับตำแหน่งปัจจุบัน ไปเรื่อย ๆ ดังรูปที่ 3.9 โดยมีโอกาสเกิดขึ้น 3 กรณี

1. เกิด Foreground Feature ใหม่ 1จุดที่เดิมกับเฟรมก่อนหน้า : ระบบจะเก็บพิกัด (Current Position)เป็นพิกัดเดิม
2. เกิด Foreground Feature ใหม่2จุดหรือมากกว่า โดยมีระยะห่างจากจุดเดิมของเฟรมก่อนหน้าไม่เท่ากัน : ระบบจะเก็บพิกัด(Current Position)เป็นพิกัดที่ใกล้ที่สุด และ พิกัดอื่น ๆ จะถือว่าเป็น Static Foreground Feature ตัวใหม่(เก็บพิกัดลงใน SFFM)
3. เกิด Foreground Feature ใหม่2จุดหรือมากกว่า โดยมีระยะห่างจากจุดเดิมของเฟรมก่อนหน้าเท่ากัน : ปกติระบบจะทำการค้นหาจากซ้ายบนลงไปยังขวาล่าง โดยไล่ดูในแนวตั้ง

   ดังนั้นจุดที่ถูกพบจุดแรก หรือ จุดที่อยู่ด้านซ้ายและสูงที่สุด จะได้รับการเลือก และเก็บพิกัด(Current Position) ส่วนจุดที่เหลือ จะถือว่าเป็น Static Foreground Feature ตัวใหม่(เก็บพิกัดลงใน SFFM)

โดยหลังจากที่ระบบสามารถทำการติดตาม Foreground Feature ได้แล้ว ต่อไปจะเป็นการอธิบายถึงขั้นตอนการเลือก Feature ที่จะทำการติดตาม

- ลำดับการเลือก Foreground Feature ในเฟรมเพื่อใช้ในการติดตาม



รูปที่ 3.11 ระยะการเลือก Foreground Feature ในเฟรมถัดไปเพื่อการ matching และติดตาม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการติดตาม Foreground Feature ต้องเข้าใจก่อนว่า ผู้วิจัยไม่สามารถเลือก Foreground Feature ในเฟรมปัจจุบันที่ซ้ำกันได้เนื่องจากการใช้ Moving Foreground Feature เป็นตัวบ่งบอกการเข้าออกของ object ในระบบ ผู้วิจัยจึงจัดลำดับสิทธิในการเลือก Foreground Feature ไว้โดยที่ให้สิทธิกับ Foreground Feature ที่เป็น Moving Foreground Feature แล้วในการเลือก Foreground Feature ในเฟรมปัจจุบันก่อน แล้วจึงค่อยทำการเลือกในส่วนของ Foreground Feature ที่เหลือ โดยการเลือกทั้งหมดจะทำการไล่จากด้านบนซ้ายไปยังขวาล่างไล่ไปตามแนวแกนนอนดังตัวอย่างในรูปภาพที่ 3.11 Foreground Feature ที่ 1 2 และ 3 มีระยะห่างจาก Foreground Feature ในเฟรมปัจจุบันเท่า ๆ กันแต่เบอร์ 1 และ 2 เป็น Moving Foreground Feature ทั้งคู่จึงมีสิทธิในการเลือกก่อนและเนื่องจากเบอร์ 1 อยู่บนเบอร์ 2 ทำให้ได้สิทธิในการเลือกก่อนเบอร์ 2 ส่วนเบอร์ 3 นั้นจะได้เลือกจับคู่ท้ายสุด

3.7 การตรวจสอบคนเข้าและออก (Detection of Human Entry/Exit)

การเข้า/ออกของคน จะถูกตรวจพบโดยการตรวจสอบจำนวนของ Moving Feature ใน MFFM ซึ่งโดยปกติแล้ว ถ้ามีอย่างน้อย 1 Moving Feature ในระบบจะแจ้งเตือนว่ามีคนเข้า แต่ถ้าไม่มี Moving Feature ใน MFFM เลย (ทั้งหมดหายไปหรือหลุดออกจากฉาก) จะแจ้งเตือนว่ามีคนออก อย่างไรก็ตาม FAST Feature อาจหายไปและปรากฏขึ้นใหม่ได้ในบางครั้ง เนื่องจากความไม่เสถียรของการรับสัญญาณภาพ (image sensor) การเปลี่ยนแปลงของแสง ฯลฯ ดังนั้นจึงต้องมี 2 ข้อจำกัดเพื่อควบคุมความไม่แน่นอนเหล่านี้

(1) ข้อจำกัดคนเข้า: ข้อจำกัดแรกคือ จำนวนที่น้อยที่สุด (minimum number: $C_{in(min)}$) ของ frame ที่ต่อเนื่องกันก่อนจะแจ้งเตือนว่ามีคนเข้า โดยเมื่อระบบมีการตรวจพบ Moving Feature อย่างน้อย 1 ตัว(จากการเคลื่อนที่ออกห่างเป็นระยะมากกว่า D) จะยังไม่แจ้งเตือนว่ามีคนเข้า ทันทีโดยจะแจ้งเตือนเมื่อมี Moving Feature อยู่เป็นระยะเวลา $C_{in(min)}$ เฟรมก่อน (เพื่อป้องกันการกระพริบของ FAST Feature ที่เกิดขึ้นจากความไม่เสถียรของเซ็นเซอร์ หรือ การเปลี่ยนแปลงของแสง จนทำให้เกิด Moving Feature ที่ไม่ถูกต้อง)

(2) ข้อจำกัดคนออก: ข้อจำกัดที่สองถูกประยุกต์ใช้เพื่อลดปัญหาที่เกิดขึ้นมาจากการหายไปแบบเป็นครั้งคราวของ Moving Feature (อันเนื่องมาจากความไม่เสถียรของเซ็นเซอร์ หรือการเปลี่ยนแปลงของแสง) เมื่อคนอยู่ห่างไกลจากกล้อง จนมี Moving Feature ให้ตรวจจับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ได้น้อย ถ้าอยู่ ๆ Moving Feature ทั้งหมดหายไปแบบชั่วคราวทั้งหมดในเวลาเดียวกัน ระบบก็จะส่งสัญญาณ “คนออก” อย่างผิดพลาด ดังนั้น ค่าจำนวนที่น้อยที่สุด (minimum number: $C_{out(min)}$) ของเฟรมที่ต่อเนื่องกันซึ่ง ค่านี้คือจำนวนเฟรมที่ระบบยอมรับการหายไปของ Moving Feature (หายไปเป็นจำนวนเฟรมไม่เกิน$C_{out(min)}$จะไม่ทำการลบออกจาก MFFM) และ เมื่อMoving Feature มีการหายไปเป็นจำนวนมากกว่า $C_{out(min)}$ ระบบจึงทำการลบออกจาก MFFM

3.8 Window Size of Buffer Zone

ขนาดของ Buffer Zone มาจากจำนวนเฟรมที่คาดคิดว่าคนจะเคลื่อนที่เกินจากระยะ D เป็นเวลาช้าที่สุดกี่เฟรม (ในสมการที่ 2) โดยผู้วิจัยได้ใช้ความเร็วในการเดินต่ำสุด (v_{min}) (106.1cm/s [6]) ระยะ D และข้อจำกัดคนเข้า (จำนวน $C_{in(min)}$ เฟรม ในบทที่3.5) เพื่อคำนวณจำนวนเฟรมใน Buffer Zone ดังสมการ(3.4)

$$b \geq \frac{D}{V_{min}}(r_{camera}) + C_{in(min)} + 1 \qquad (3.4)$$

ในขณะที่ 1 frame ที่เพิ่มเติมขึ้นมาในสมการ (+1) มีไว้สำหรับเฟรมเริ่มต้น ตอนที่ Foreground Feature ถูกตรวจพบเป็นครั้งแรก

3.9 Feature matching สำหรับการเปรียบเทียบ Moving Foreground Feature (MFF)

ปกติการเลือก Feature ในระดับ pixel-level นั้นจะทำการเลือกติดตาม Feature ที่เกิดใกล้กับ Feature ในเฟรมก่อนหน้าที่สุดในระยะรัศมี R ซึ่งทำให้อาจเกิดการเลือกติดตามที่ไม่ถูกต้อง ที่สุดผู้วิจัยจึงนำการ tracking ในแบบที่เลือก Feature ที่มีลักษณะคล้ายที่สุดเข้ามาใช้ โดยมีรายละเอียดดังนี้

1. การใช้ Speeded-Up Robust Features (SURF) บ่งบอกรายละเอียดของ Feature

ระบบจะทำการคัดแยก Foreground Feature ออกมาจาก Feature ที่เกิดขึ้นนอกเหนือจากในพิกัดบน BFM แล้วจากนั้นจึงจะนำ Feature ใหม่เหล่านั้นไปทำให้ Descriptor บ่งบอกรายละเอียด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของแต่ละ Feature และนำไปใช้ในขั้นตอน matching ต่อไป ซึ่ง Descriptor ที่ผู้วิจัยเลือกใช้นั้นคือ SURF[10] เพราะ ความรวดเร็ว และความสามารถในการบ่งบอกรายละเอียดได้อย่างถูกต้อง โดยเมื่อได้รายละเอียดข้อมูลของ Foreground Feature แล้ว ขั้นตอนต่อไปคือ การจับคู่ Foreground Feature ในเฟรมปัจจุบันกับเฟรมก่อนหน้าโดยอิงตามข้อมูลที่ได้จาก SURF

2. การใช้ Feature Matching ในการจับคู่ Feature

เมื่อระบบมีรายละเอียดของแต่ละ Foreground Feature ที่ได้มาจากวิธีการ SURF แล้วนั้น ต่อมาระบบจะทำการเลือก Feature ที่อยู่ในรัศมี R และทำการคัดเลือก Feature ที่มีเวคเตอร์การเคลื่อนที่ และระยะทาง (ซึ่งเป็นข้อมูลของตัว Feature นั้น ๆ เช่น การกระจัด ทิศทาง หรือระยะทาง) คล้ายกับใน Feature ของเฟรมก่อนหน้ามากที่สุด แทนที่จากเดิมที่ผู้วิจัยจะเลือกจับคู่แบบ Nearest Matching ซึ่งเป็นการจับคู่ Feature ที่อยู่ใกล้มากที่สุด โดยจะวัดความคล้ายกันเหล่านั้นโดยใช้ FLANN [23] ซึ่งเป็น Matcher ที่มีความเร็วสูง เหมาะสมกับการนำมาในการประมวลผลแบบ real-time

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

การทดลองและผลลัพธ์

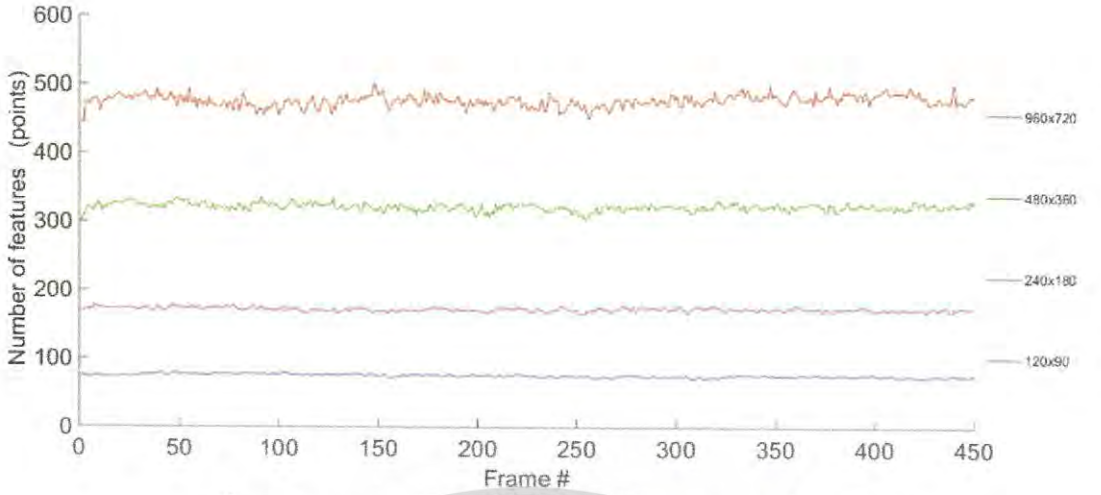
ระบบการตรวจจับคนเข้าออกนี้ถูกใช้กับคอมพิวเตอร์รุ่น Intel Core I7-4700HQ @ 2.40GHz ใช้ภาพ Grey level ที่มีขนาดความละเอียด (resolution) ที่ 120×90, 240×180, 480×360, และ 960×720 พิกเซล โดยได้ระดับความถี่ของวิดีโอ ที่ประมาณ 30 fps ด้วยกล้อง CMOS ผ่านทางสาย USB โดยลำดับภาพทั้งหมดที่ถูกใช้ในการทดลองนี้ถูกเก็บจากภายในอาคาร ถึงแม้ว่าระยะเวลาการดำเนินการจะแปรผันตามขนาด resolution นั่นคือ ยิ่ง resolution มากก็ยิ่งกินเวลามาก แต่ในทุก ๆ resolution ระบบจะสามารถดำเนินการที่ความถี่ประมาณ 30 fps

4.1 ความไม่เสถียรของ FAST Point Detector

รูปที่ 4.1 ตัวอย่างภาพในวีดีโอที่ใช้ทำการเก็บผลการทดลอง

ความไม่เสถียรของ FAST Feature Detector บน Video Resolution ที่หลากหลายจะถูกตรวจสอบในหัวข้อนี้ โดยการตรวจจับ FAST Feature ในภาพที่มีแต่ Background (ไม่มีการเคลื่อนไหวในภาพ ดังรูปที่ 4.1) ภาพจำนวนทั้งสิ้น 450 เฟรม จะถูกตรวจสอบใน แต่ละ Resolution ตามรูปที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.2 จำนวนของ FAST Feature ใน Resolution ที่แตกต่างกัน

ตามทฤษฎีแล้ว จำนวนของ FAST Feature ควรจะมีเท่าเดิม แต่ในความเป็นจริงแล้วมันจะเกิดการเปลี่ยนแปลงอันเนื่องมาจาก แสงที่เปลี่ยนไป หรือ ความไม่เสถียรของ Image Sensor ฯลฯ ตามที่เห็นในภาพว่าจำนวนของ FAST Feature และ การเบี่ยงเบนของมันจะเพิ่มขึ้นที่ Resolution ที่มีค่าสูงขึ้น

ในการทดลองนี้ ส่วนเบี่ยงเบนมาตรฐานของจำนวน FAST Feature ใน Video Resolution ที่ 120x90, 240x180, 480x360, 960x720 คือ 1.63, 2.75, 5.23 และ 8.98 ตามลำดับ

เนื่องจากว่าแต่ละ พิกเซลในภาพที่มีความละเอียดต่ำจะมีความสอดคล้องกับพิกเซลจำนวนมากในภาพ High-Resolution ซึ่งภาพแบบ Low-Resolution ค่อนข้างจะช่วยแก้ปัญหาในวิธีการแบบเดียวกับที่ ได้นำ Averaging Filter มาใช้ ซึ่งจากที่ได้กล่าวไว้ข้างต้นนั้นจะดูสมเหตุผลต่อการตรวจจับ Feature ที่มีความเสถียรมากกว่า ในระดับ Video Resolution ที่ต่ำกว่าลงไป

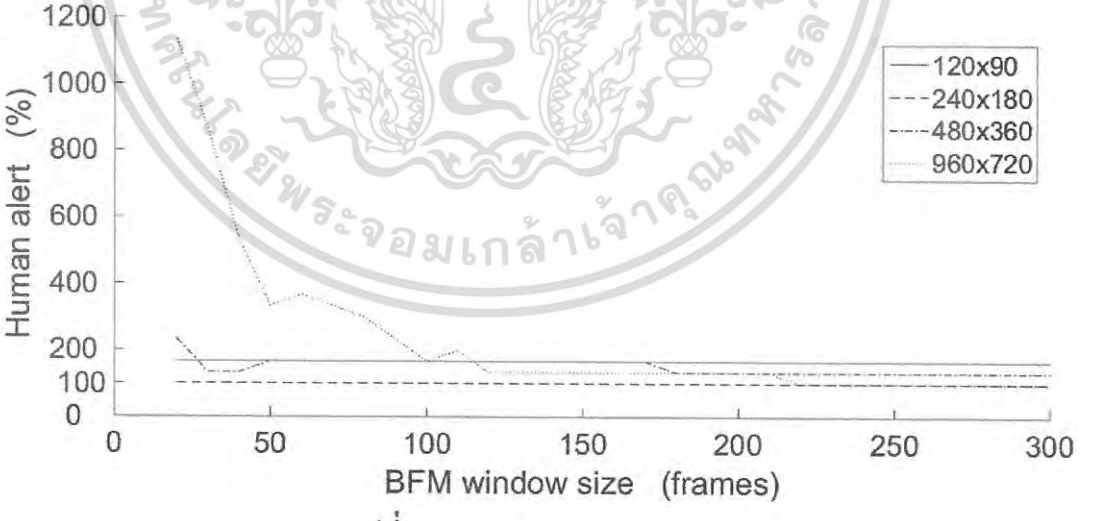
FAST Feature ที่ไม่สม่ำเสมอเหล่านี้จะมีผลกระทบต่อประสิทธิภาพของการติดตามเพราะมันจะก่อให้เกิดการตรวจจับ Moving Feature สำหรับคนเข้าและคนออกที่ผิดพลาด (ตรวจพบคน ๆ เดียวแต่ระบบเกิดแจ้งการเข้าออกในภาพเดิมหลายรอบ) และบางครั้งก็ดีเลย์ในการตรวจพบคนเข้า ในทางทฤษฎีแล้วการเพิ่มขนาดของ BFM จะช่วยให้ระบบรับรู้ถึงการมีอยู่ของ FAST Feature ที่กระพริบและจัดการแก้ไขปัญหาที่ระบุไว้ข้างต้นได้ การทดลองและผลลัพธ์เพื่อการพิสูจน์สมมติฐานนี้จะถูกอธิบายในหัวข้อถัดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ตัวแปรที่ส่งผลต่อการตรวจจับคนเข้าและออก (Parameter Effect on Human Entry/Exit)

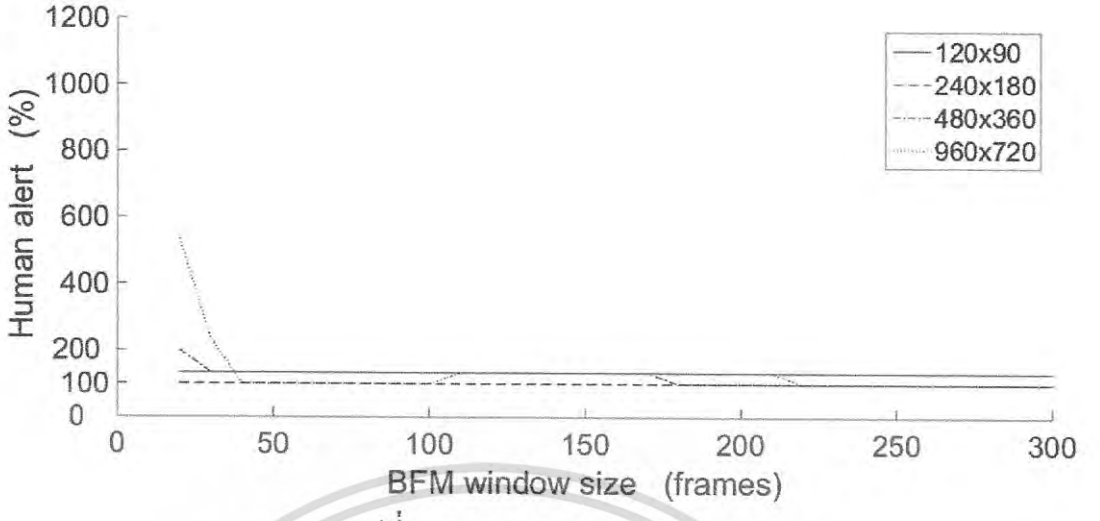
ตัวแปรที่ส่งผลต่อประสิทธิภาพของการตรวจจับคนเข้าและออก จะถูกพิจารณาในหัวข้อนี้ โดยมี Parameter 2 ตัวคือ $C_{in(min)}$ และ $C_{out(min)}$ ที่จะถูกทดสอบและผันแปรไปตาม Video Resolution ต่าง ๆ (คือ 120x90, 240x180, 480x360, 960x720 pixels) จะสังเกตได้ว่า $C_{in(min)}$ คือจำนวนที่น้อยที่สุดของเฟรมต่อเนื่องที่เกิด Moving Feature ก่อนที่ระบบจะทำการแจ้งเตือนว่ามีคนเข้า โดยดูให้แน่ใจได้ว่า Foreground Feature จะถูกเลื่อนขั้นจาก Static Foreground Feature ไปเป็น Moving Foreground Feature ไม่ใช่เกิดจากข้อผิดพลาด และ $C_{out(min)}$ ก็เป็นจำนวนที่น้อยที่สุดของเฟรมต่อเนื่องที่ Moving Foreground Feature ไม่สามารถหาพิกัดต่อไปได้หากไม่สามารถหาได้เป็นระยะเวลามากกว่าหรือเท่ากับ $C_{out(min)}$ เฟรมระบบถึงจะทำการลบจุด Moving Foreground Feature นี้ออกจาก MFFM

โดย รูปที่ 4.2.1 – 4.2.3 ทดลองโดยใช้วีดีโอที่มีคนเดินเข้า และ ออกเพียง1ครั้งในจำนวน 5 ตัวอย่างวีดีโอ แล้วนำมาหาค่าเฉลี่ยซึ่งหากมีการแจ้งเตือนเป็นจำนวนที่ถูกต้อง Human Alert (แกน Y) ค่าจะอยู่ที่ 100% แต่ถ้าหากมีการแจ้งเตือนมากกว่าจำนวนที่ถูกต้อง Human Alert(แกน Y) ก็จะมากกว่า 100%



รูปที่ 4.2.1 $C_{in(min)}= 0, C_{out(min)}= 0$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.2.2 $C_{in(min)}= 2$, $C_{out(min)}= 2$

รูปที่ 4.2.3 $C_{in(min)}= 4$, $C_{out(min)}= 3$

โดยสรุปจากรูปที่ 4.2.1-4.2.3 จะเห็นได้ว่าความละเอียดที่สูงขึ้นทำให้เกิดข้อผิดพลาดที่มากขึ้นในจำนวนของ BFM ที่น้อย และจะค่อย ๆ มีผลลัพธ์ที่ดีขึ้นเมื่อ BFM มีจำนวนมากขึ้นเรื่อย ๆ ซึ่งตัวแปร $C_{in(min)}$, $C_{out(min)}$มีผลต่อความถูกต้องของระบบด้วย ดังรูปที่ 4.2.1 จะเห็นได้ว่าหากไม่ได้ใช้ค่าตัวแปร $C_{in(min)}$, $C_{out(min)}$เลยเมื่อ BFM มีจำนวนมากขึ้นก็ไม่สามารถลดความผิดพลาดจะเป็นปกติได้ ต่อไปคือรูปที่ 4.2.2 คือทดลองด้วยค่า $C_{in(min)}$, $C_{out(min)}$ โดยให้ค่าเท่ากับ 2 จะเห็นได้ว่าผลลัพธ์ดีขึ้นกว่าการที่ไม่ได้ใช้ตัวแปรเหล่านั้นแต่ก็ยังไม่สามารถลดข้อผิดพลาดได้ทั้งหมด ผู้วิจัยจึงทำการหาค่าที่เหมาะสมดังรูปที่ 4.2.3 โดยทดลองด้วยค่า $C_{in(min)}= 4$, $C_{out(min)}= 3$จะเห็นได้ว่า เมื่อจำนวนของ BFM มากกว่าหรือเท่ากับ220 ระบบจะทำงานได้อย่างไม่มีผิดพลาด ในทุก ๆ Resolution

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

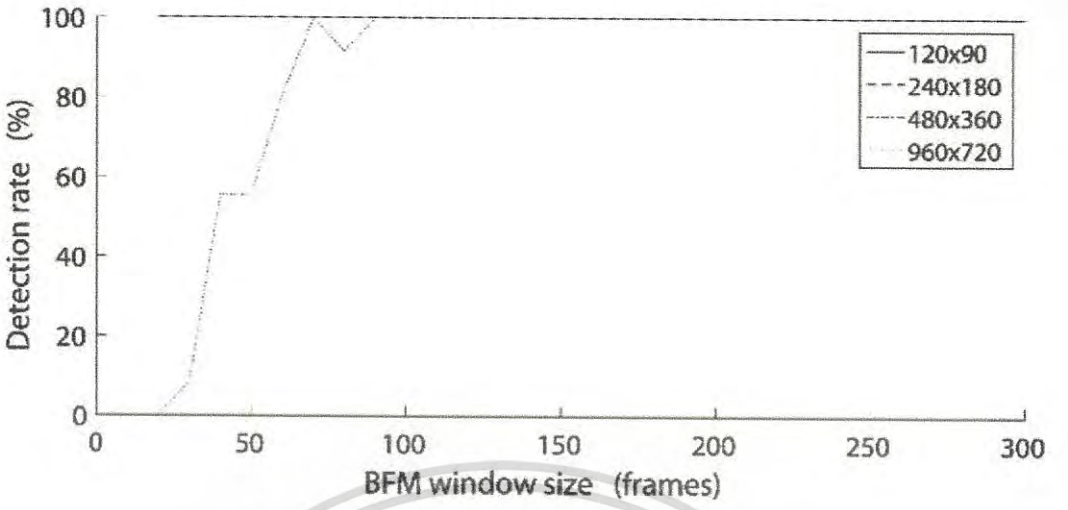
4.3 Recall Rate อัตราความถูกต้องของเฟรมที่มีคนเข้า/ออก

ทดลองโดยใช้วีดีโอที่มีการเดินเข้าและออกหลายครั้งจำนวน5วีดีโอซึ่งทดลองโดยวัดจากเปอร์เซ็นต์ความถูกต้องของระบบ เปรียบเทียบกับความจริง(ทำการ Mark เฟรมด้วยมือ)โดยแบ่งเป็น 4 กลุ่มดังนี้

1. Negative In: เปอร์เซ็นต์ความถูกต้องของเฟรมก่อนมีคนเข้า
2. Positive In : เปอร์เซ็นต์ความถูกต้องของเฟรมที่ระบบต้องเริ่มตรวจจับว่ามีคนเข้า โดยให้ระยะเวลา 45 เฟรมหลังจากคนเข้าจริง
3. Negative Out: เปอร์เซ็นต์ความถูกต้องของเฟรมก่อนมีคนออก หรือ เฟรมที่มีคนอยู่
4. Positive Out: เปอร์เซ็นต์ความถูกต้องของเฟรมที่ระบบต้องตรวจจับว่ามีคนออกจากเฟรม โดยให้ระยะเวลา 45เฟรม หลังมีคนออกจริง

รูปที่ 4.3.1 Negative In

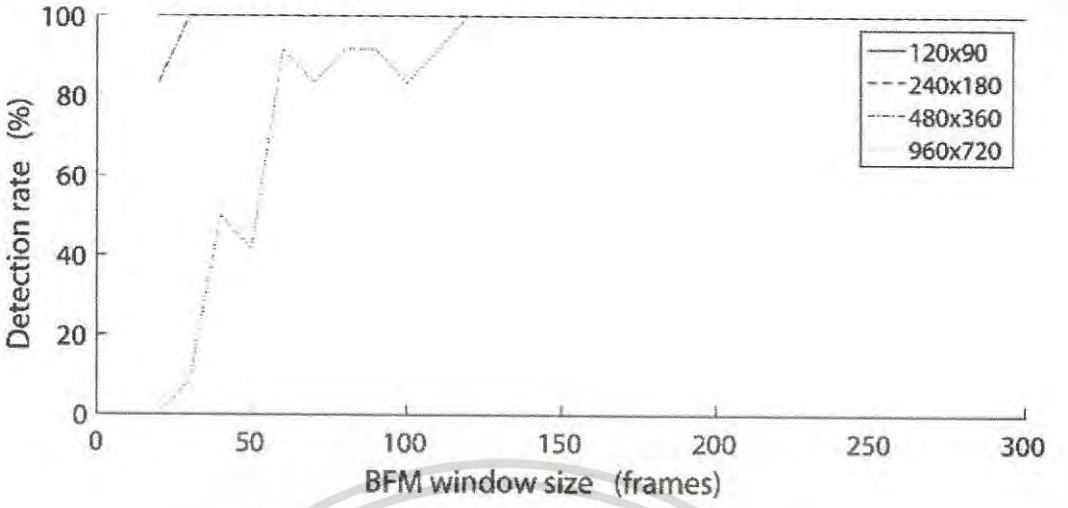
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.3.2 Positive In

รูปที่ 4.3.3 Negative Out

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.3.4 Positive Out

หากว่า $C_{in(min)} = 4$ และ $C_{out(min)} = 3$ อัตราค่าความถูกต้องของจำนวนเฟรมที่มีคนเข้า กับ ที่มีคนออกสำหรับทั้ง 4 video resolution จะถูกแสดงในรูปที่ 4.3.1 – 4.3.4 ผลลัพธ์จะแสดงให้เห็นว่าภาพที่มี Resolution ต่ำ จะมีประสิทธิภาพดีกว่าภาพที่มี Resolution สูง อย่างไรก็ดี ณ ตอนนี้ยังไม่มีข้อผิดพลาดสำหรับ ทุก ๆ ความละเอียดภาพ แต่ก็ต่อเมื่อ ขนาดของ BFM มีค่ามากเพียงพอ อย่างน้อย 120 เฟรม(แต่เพื่อผลลัพธ์ที่ดีที่สุดควรใช้ 220 เฟรม)

4.4 การเปรียบเทียบการทดลองกับระบบ 3 มิติที่พัฒนาเพิ่มขึ้น

เพื่อให้ประสิทธิภาพในการตรวจจับคนเข้าและออก สามารถทำได้อย่างถูกต้องและมีประสิทธิภาพมากที่สุดจึงได้มีการนำวิธีการติดตาม Feature ในมุมมอง 3 มิติเข้ามาใช้ ทำให้เกิดการเปลี่ยนแปลงต่าง ๆ ดังต่อไปนี้

4.4.1 การทดลองความถูกต้องของจำนวนเฟรมขณะที่มีคนอยู่ในกล้อง

สำหรับในรูปที่ 4.4.1 ถึง รูปที่ 4.4.4 จะมีการอธิบายดังต่อไปนี้

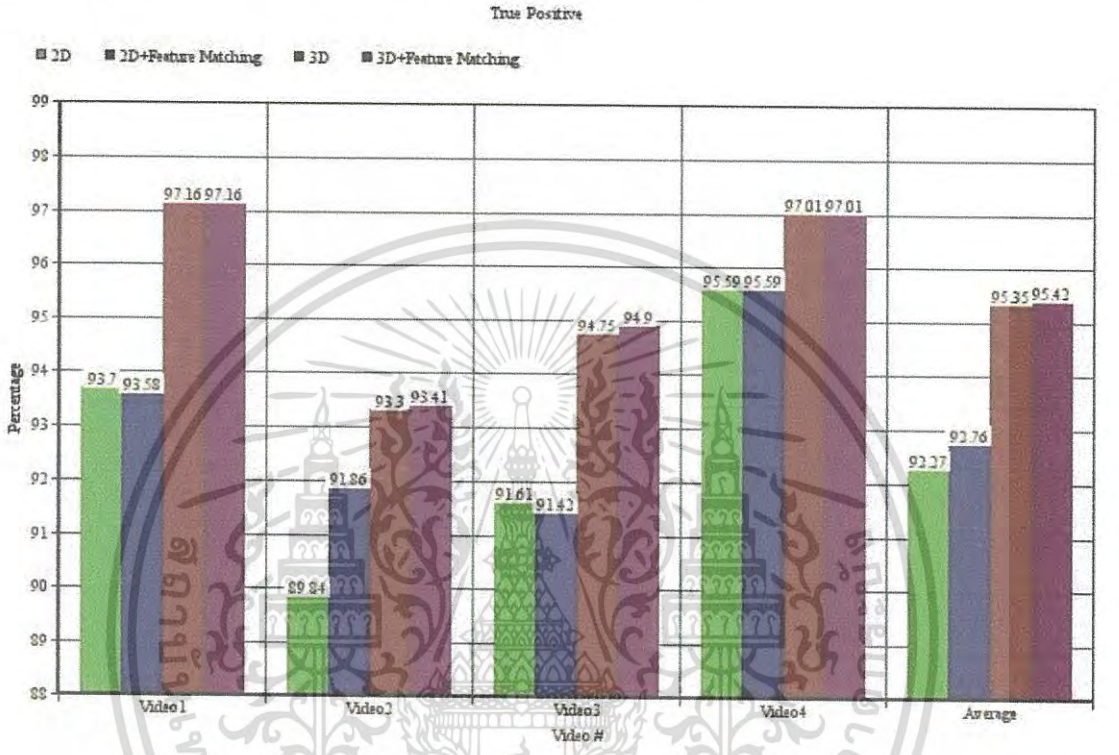
1. ให้กราฟแท่งสีเขียวหมายถึงภาพวิดีโอที่ใช้ระบบการตรวจจับคนเข้าและออกในการใช้มาตรวัดระดับ pixel-level ด้วยวิธีการ track กับ Feature เฟรมก่อนหน้าที่ใกล้ที่สุดในรัศมี R

2. ให้กราฟแท่งสีน้ำเงินหมายถึงภาพวิดีโอที่มีการตรวจจับคนเข้าและออกในการใช้มาตรวัดระดับ pixel-level ด้วยวิธีการ track แบบ Feature matching

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. ให้กราฟแท่งสีแดงหมายถึงภาพวิดีโอที่มีการตรวจจับคนเข้าและออกในการใช้มาตรวัดตามจริง (global coordinate) ใช้วิธี track กับ Feature เฟรมก่อนหน้าที่ใกล้ที่สุดในรัศมี R

4. ให้กราฟแท่งสีม่วงหมายถึงภาพวิดีโอที่มีการตรวจจับคนเข้าและออกในการใช้มาตรวัดตามจริง (global coordinate) ใช้วิธี track แบบ Feature matching

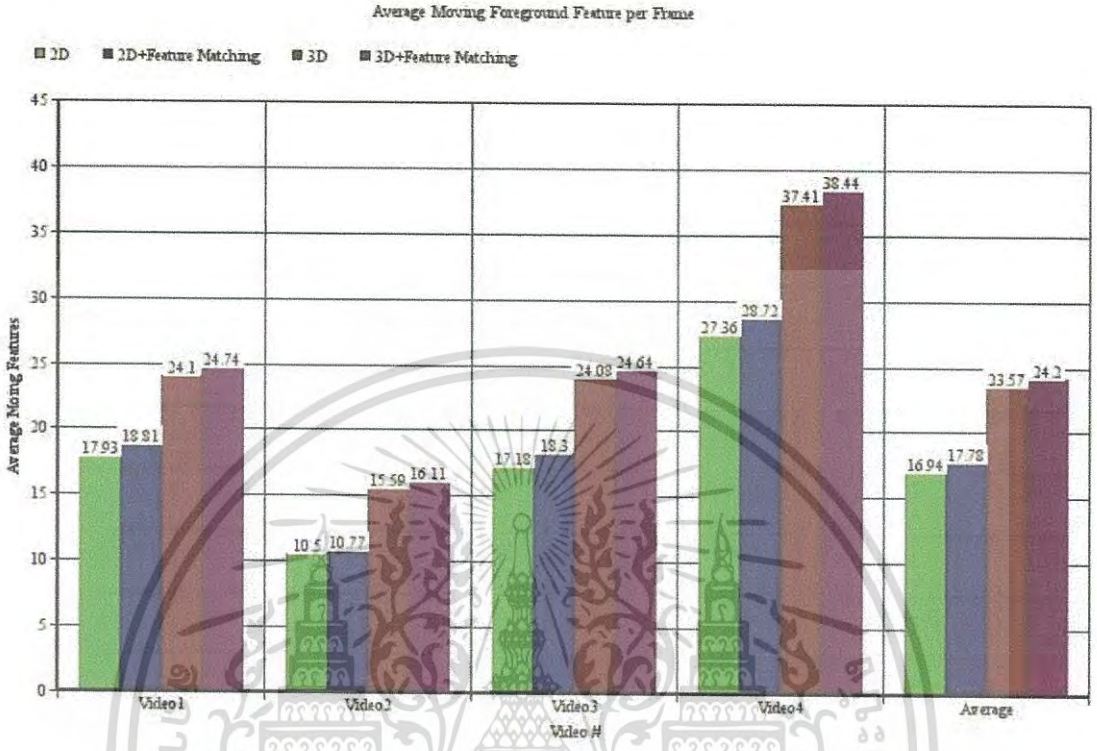


รูปที่ 4.4.1 กราฟแสดงจำนวนเฟรมที่ถูกต้องขณะมีคนในเฟรม

ในรูปที่ 4.4.1 เมื่อนำการใช้ global coordinate หรือการวัดระยะในระดับ3มิติเข้ามาช่วยแล้วนั้นจะมีประสิทธิภาพเพิ่มขึ้นอย่างเห็นได้ชัด และ การใช้ Feature matching ก็จะช่วยส่งเสริมความถูกต้องขึ้นไปในระดับหนึ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4.2 การทดลองความถูกต้องของจำนวน moving Foreground Feature ที่เกิดขึ้นขณะมีคนอยู่ในกล้อง

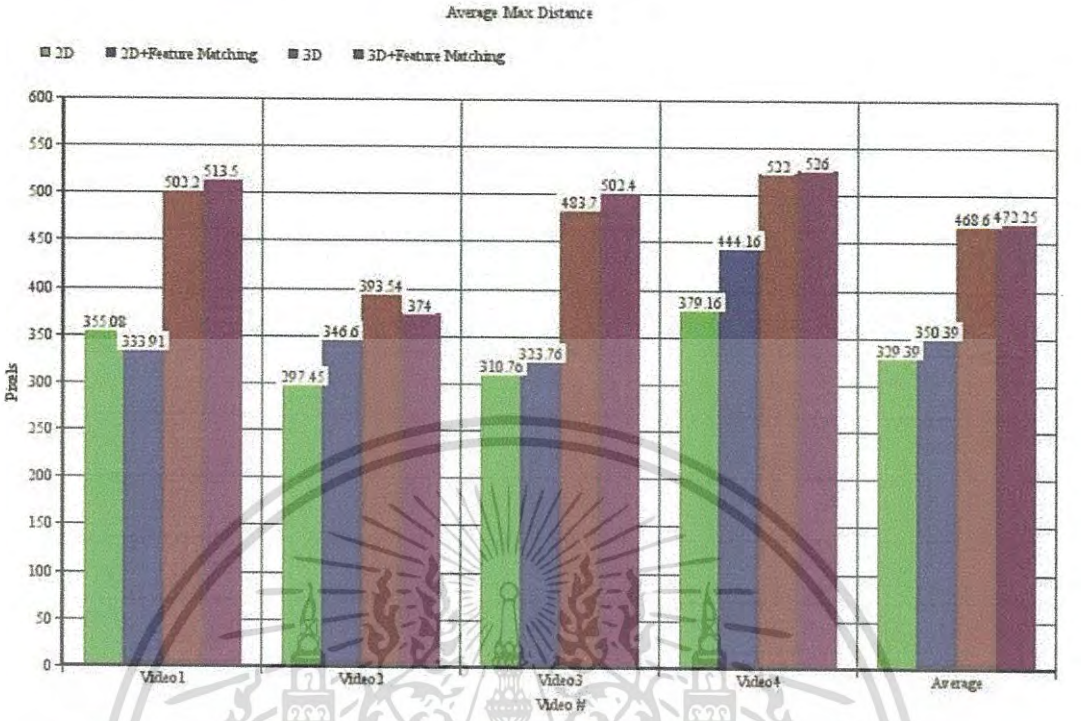


รูปที่ 4.4.2 กราฟแสดงค่าเฉลี่ยของจำนวน MFF ขณะที่มีคนอยู่ในเฟรม

จากรูปที่ 4.4.2 จะเห็นได้ว่าจำนวนของ moving Foreground Feature จะสามารถถูกตรวจพบได้มากขึ้นอย่างเห็นได้ชัดเมื่อนำการใช้ global coordinate และ Feature matching เข้ามาช่วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4.3 การทดสอบความถูกต้องบนค่าเฉลี่ยความยาวสูงสุดของ moving Foreground Feature

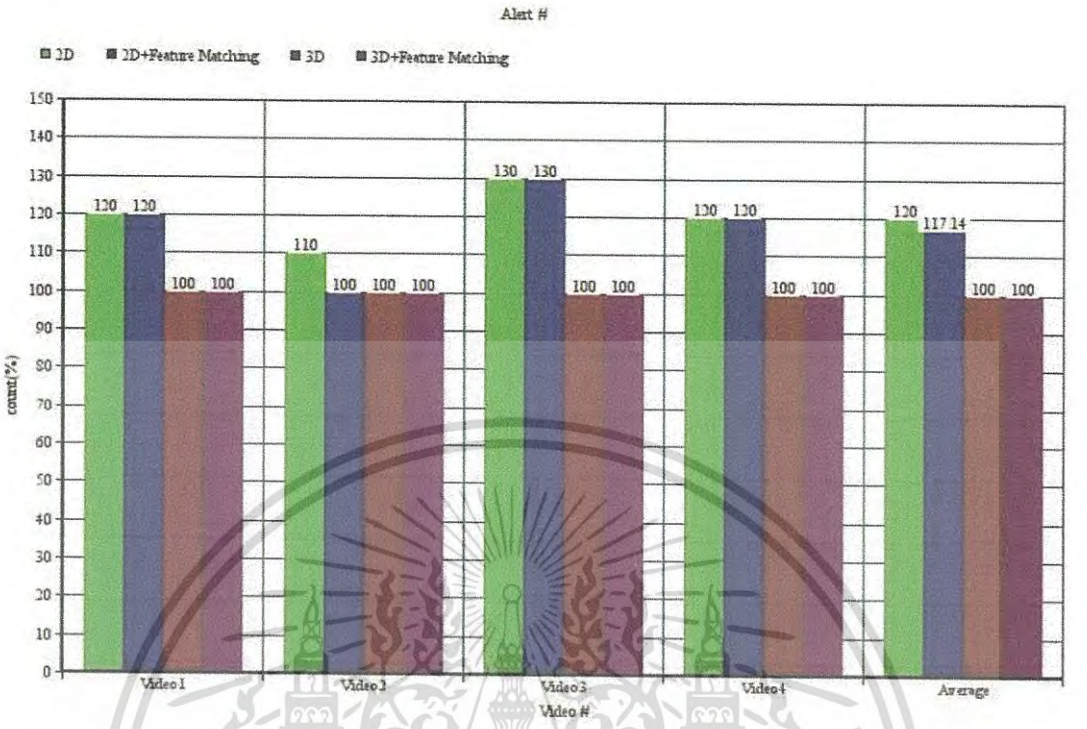


รูปที่ 4.4.3 กราฟแสดงค่าเฉลี่ยความยาว MFF สูงสุดขณะมีคนอยู่ในเฟรม

การนำ global coordinate มาประยุกต์ใช้กับการทำ Feature matching ช่วยให้สามารถตรวจจับ MFF ได้แม่นยำมากขึ้น โดยระยะทางการติดตามจะมีค่าความยาวสูงสุดมากขึ้นดังในรูปที่ 4.4.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4.4 การทดสอบความถูกต้องในจำนวนครั้งที่มีการแจ้งเตือนคนเข้าและออกผิดพลาด (error)



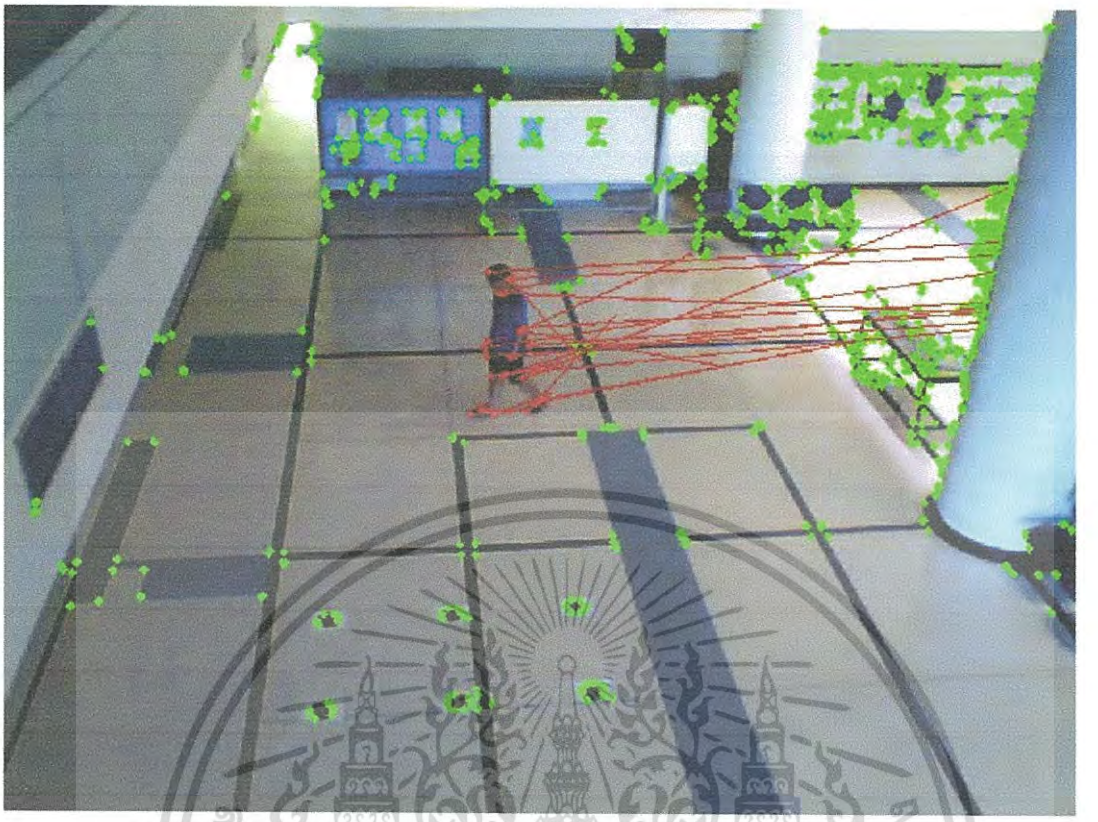
รูปที่ 4.4.4 กราฟแสดงความถูกต้องของจำนวนครั้งที่ระบบแจ้งเตือน

ในการติดตาม Feature ในมุมมอง 3 มิติ ของกล้องด้วยการนำ Feature matching มาใช้ และการนำ global coordinate มาใช้ในการวัดระยะ ช่วยให้ข้อผิดพลาดของการแจ้งเตือนคนเข้าและออกลดน้อยลงซึ่งส่วนใหญ่มาจาก การเข้าและออกในมุมลึก

4.5 Screenshots

ภาพตัวอย่างจากโปรแกรมที่พัฒนา (รูปที่ 4.5.1) โดยมีการนำ Matlab และ Qt creator มาใช้ทำเป็นโปรแกรมสำหรับเป็นหน้าต่างคำนวณและแสดงผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

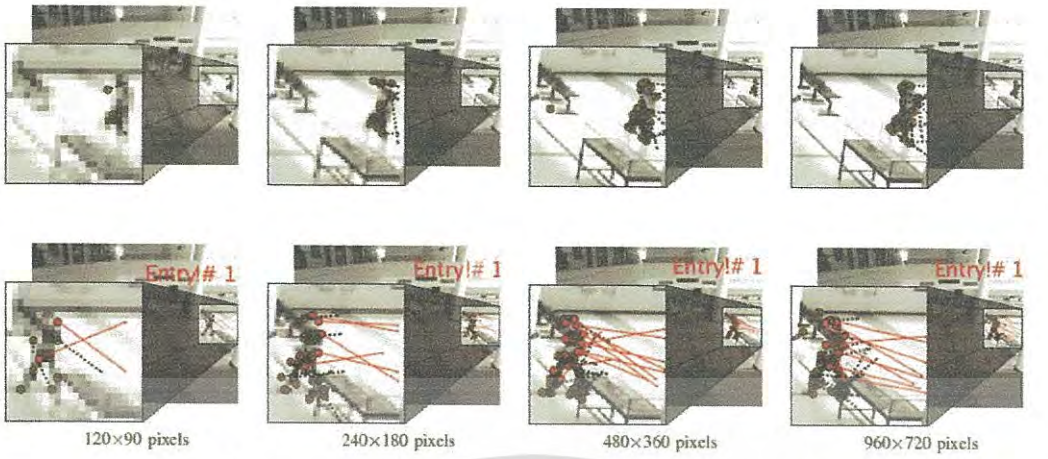


รูปที่ 4.5.1 การติดตาม Feature เมื่อกล้องจับภาพในมุมมอง 3 มิติ

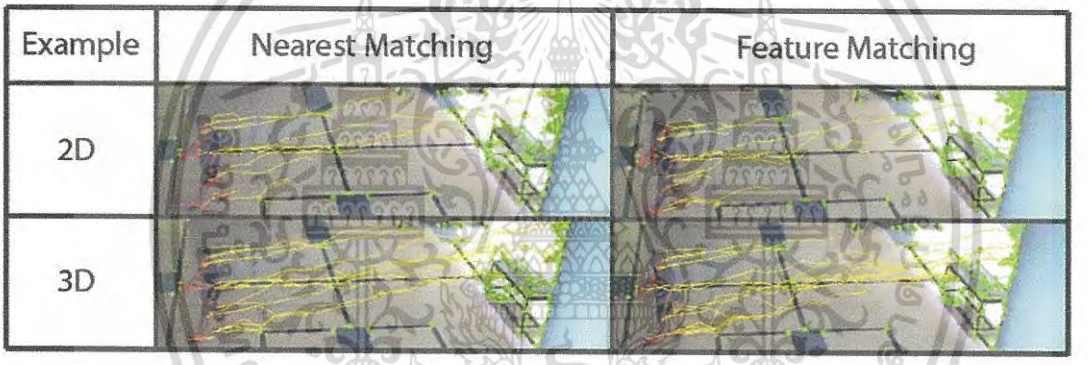
จากรูปที่ 4.5.1 ภาพวงกลมสีดำ 6 ภาพถูกวางในตำแหน่งช่วงด้านล่าง เมื่อทำการ calibrate camera และดำเนินการ tracking โดยกล้องรับรู้ถึงภาพ 3 มิติ การติดตาม Feature มีระยะทางที่ยาวมากขึ้น ซึ่งหมายความว่า Feature ที่ถูกติดตามอยู่นั้นเป็น Feature ตัวเดียวกันกับในเฟรมก่อนหน้าหมายความว่าการติดตามมีความถูกต้องสมบูรณ์มากขึ้นทั้งในมุมแนวนอนและในแกนลึก

ภาพถ่ายจากการทดสอบบนระบบจะถูกแสดงใน รูปที่ 4.5.2 ผลลัพธ์จากทั้ง 4 ความละเอียดวิดีโอภาพนั้น; เริ่มจากฝั่งซ้ายสุดคือ 120x90 พิกเซล 240x180 พิกเซล 480x360 พิกเซล และ 960x720 พิกเซล ที่รูปขวาสุด โดยภาพที่แถวด้านบนจะแสดงถึงผลลัพธ์ของ Static Foreground Features (Foreground Features ที่คงที่) โดยที่ภาพในแถวด้านล่างจะแสดง Moving Foreground Feature (Foreground Features ที่เคลื่อนที่) ตำแหน่งปัจจุบันของ Moving Feature จะถูกแสดงเป็น วงกลมสีแดงใหญ่ที่มีเส้นสีแดงลากเชื่อมไปถึงตำแหน่งต้นกำเนิดของ Feature และ ตำแหน่งของ Static Foreground Feature จะถูกแสดงเป็นรูป ⓧ สีส้มที่มีเส้นประลากโยงไปถึงตำแหน่งต้นกำเนิดของ Feature

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.5.2 ภาพตัวอย่างการทดลองในความละเอียดที่ต่างกัน



รูปที่ 4.5.3 การเปรียบเทียบการติดตาม Foreground Feature

ในรูปที่ 4.5.3 จะเป็นการเปรียบเทียบระหว่างระบบตรวจจับคนเข้าออกที่ใช้การวัดระยะของกล้องในระดับ pixel-level (2D) กับ ระบบที่ใช้การวัดระยะตามจริง (3D) โดยมีรูปแบบการติดตาม 2 แบบ คือ 1. การติดตามด้วยการเทียบจากระยะห่าง Feature ในเฟรมปัจจุบันกับ Feature ในเฟรมก่อนหน้า (Nearest Matching) และ 2. การติดตามโดยเทียบจากความคล้ายกันของลักษณะ เช่นการกระจัด ของ Feature ในเฟรมปัจจุบันกับ เฟรมก่อนหน้า (Feature Matching)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
บทสรุป

5.1 สรุปผลการทดลอง

วิธีการตรวจพบคนเข้าออกได้มีการนำเสนอโดยใช้ FAST Features การดำเนินงานวิจัยถึงตัวแปรต่าง ๆ ที่มีผลต่อความละเอียดภาพวิดีโอที่ต่างกันเกิดขึ้นเพื่อใช้ในการตัดสินใจค่าที่ดีที่สุดของตัวแปรนั้น ๆ ระบบที่พัฒนาขึ้นสามารถตรวจพบการเข้าออกของคนได้อย่างถูกต้องภายใต้สภาพแวดล้อมและข้อจำกัดที่กำหนดโดยปราศจากการเตือนที่ผิดพลาดและปราศจากการตรวจพบที่คลาดเคลื่อน เมื่อขนาดของ Background Feature Model อยู่ที่ 220 เฟรม หรือมากกว่า ระบบนี้จะสามารถดำเนินการได้ที่ 30 fps (เฟรมภาพต่อวินาที) หากความละเอียดของภาพอยู่ที่ 960x720 พิกเซล และจะยิ่งมากกว่า 30 fps ที่ความละเอียดภาพวิดีโอที่ต่ำลง แต่การตรวจจับดังกล่าวนี้ยังไม่ครอบคลุมถึงในกรณีที่มีการเข้า/ออกของคนในมุมลึก เนื่องจากการติดตาม Feature ของคนที่เคลื่อนไหวในมุมลึกจะเกิดการผิดเพี้ยนเพราะการเดินเข้ามาในมุมลึกจะเป็นการเคลื่อนที่ในระดับ pixel ที่น้อยกว่ารัศมีจริงทำให้ผู้วิจัยนำการวัดระยะในระดับ global coordinate และ Feature matching เข้ามาช่วยแก้ไขในจุดนี้ ทำให้การนำ global coordinate และ Feature matching เข้ามาช่วยนั้นส่งผลต่อการทำงานของระบบให้มีประสิทธิภาพมากขึ้นโดยที่อาจจะต้องแลกมาด้วยการต้องเข้าไปติดตั้งก่อน และ ระยะการประมวลผลที่ลดลงหากมีจำนวน Foreground Feature ที่มากขึ้น แต่ก็สามารถลดปัญหาที่เกิดขึ้นได้อย่างดี

ระบบตรวจจับคนเข้า/ออกโดยการเฝ้าติดตามนั้น สามารถประมวลผลได้แบบ real-time (30fps) เพราะว่ากระบวนการทำงานไม่ได้ใช้ pixel ทั้งหมดในการประมวลผล แต่เป็นการใช้เพียงจุดสนใจ (Feature) ที่ได้มาจากกระบวนการของ FAST ซึ่งมีความเร็วในการสกัดจุดสนใจออกมาจากภาพที่รวดเร็ว นอกจากนั้นการติดตาม Foreground Feature โดยมาตรฐานแล้ว ผู้วิจัยจะใช้การเลือกจุดสนใจที่ใกล้ที่สุดและวัดระยะทางในระดับ pixel ซึ่งจะทำให้ความเร็วในการประมวลผลดีที่สุด ซึ่งผู้วิจัยก็ได้นำเสนอถึงการเพิ่มประสิทธิภาพ โดยการใช้ global coordinate (วัดระยะจริง) โดยต้องทำกระบวนการติดตั้งแผ่นวงกลมให้ระบบรู้จักพื้นที่จริง (camera calibration) และการใช้ Feature matching ในการเลือกจุดสนใจแทนการเลือกตัวที่ใกล้ที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] E. Rosten and T. Drummond. "**Machine learning for high-speed corner detection**." in Computer Vision – ECCV 2006, A. Leonardis, H. Bischof, and A. Pinz, Eds. Springer Berlin Heidelberg, 2006, pp. 430–443.

[2] E. Rosten, R. Porter, and T. Drummond. "**Faster and better: A machine learning approach to corner detection**." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 1, pp. 105–119, Jan. 2010.

[3] M. Donoser and H. Bischof. "**Efficient maximally stable extremal region (MSER) tracking**." in 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, Jun. 2006, pp. 553– 560.

[4] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. "**ORB: An efficient alternative to SIFT or SURF**." in 2011 IEEE International Conference on Computer Vision (ICCV), Nov. 2011, pp. 2564–2571.

[5] J.Liu and X. Liang. "**I-BRIEF: A fast Feature point descriptor with more robust Features**." in 2011 Seventh International Conference on Signal Image Technology and Internet-Based Systems (SITIS), Nov. 2011, pp. 322–328.

[6] R. W. Bohannon, "**Comfortable and maximum walking speed of adults aged 20—79 years: reference values and determinants**." Age and Ageing, vol. 26, no. 1, pp. 15–19, Jan. 1997.

[7] H. Tao, H. Sawhney, and R. Kumar, "**Object tracking with bayesian estimation of dynamic layer representations**." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 1, pp. 75–89, Jan. 2002.

[8] D. Comaniciu, V. Ramesh, and P. Meer, "**Kernel-based object tracking**." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 5, pp. 564–577, May 2003.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม (ต่อ)

[9] A. Jepson, D. Fleet, and T. El-Maraghi, "**Robust online appearance models for visual tracking**." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 10, pp. 1296–1311, Oct. 2003.

[10] Y. Chen, Y. Rui, and T. Huang, "**JPDAF based HMM for real-time contour tracking**." in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001, vol. 1, 2001, pp. I–543–I–550 vol.1.

[11] K. Sato and J. K. Aggarwal, "**Temporal spatio-velocity transform and its application to tracking and interaction**." Computer Vision and Image Understanding, vol. 96, no. 2, pp. 100–128, Nov. 2004.

[12] A. Yilmaz, X. Li, and M. Shah, "**Contour-based object tracking with occlusion handling in video acquired using mobile cameras**." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 11, pp. 1531–1536, Nov. 2004.

[13] I. Cox and S. Hingorani. "**An efficient implementation of reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking**." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 2, pp. 138–150, Feb. 1996.

[14] C. Veenman, M. Reinders, and E. Backer. "**Resolving motion correspondence for densely moving points**." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 1, pp. 54–72, Jan. 2001.

[15] K. Shafique and M. Shah. "**A non-iterative greedy algorithm for multi-frame point correspondence**." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 1, pp. 51–65, Jan. 2005.

[16] Zhang, Lijing, and Yingli Liang. "**Motion Human Detection Based on Background Subtraction**." In 2010 Second International Workshop on ETCS, 1:284–87, 2010. doi:10.1109/ETCS.2010.440.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม (ต่อ)

[17] Li, Changyan, Lijun Guo, and Yichen Hu. "**A New Method Combining HOG and Kalman Filter for Video-Based Human Detection and Tracking**." In 2010 CISP, 1:290–93, 2010. doi:10.1109/CISP.2010.5648239.

[18] C. Harris and M.J. Stephens. "**A combined corner and edge detector**." Alvey Vision Conference pp. 147–152, 1988 The article by J. Shi and C. Tomasi, Good Features to track, Int. Conference on Computer Vision and Pattern Recognition, pp. 593-600, 1994.

[19] D. Lowe. "**Distinctive Image Features from Scale Invariant Features**." International Journal of Computer Vision, Vol. 60, No. 2, 2004, pp. 91-110.

[20] A. Ess, T. Tuytelaars and L. Van Gool. "**Speeded Up Robust Features by H. Bay**." in Computer Vision and Image Understanding, vol. 110, No. 3, pp. 346-359, 2008.

[21] Intaek Kim, Tayyab Wahab Awan and Youngsung Soh. "**Background Subtraction-Based Multiple Object Tracking Using Particle Filter**." IWSSIP 2014, 21st International Conference on Systems, Signals and Image Processing, 12-15 May 2014

[22] Gang XU, Dong ZHAO, Qi ZHOU, Ding HUANG. "**Moving Target Tracking based on Adaptive Background Subtraction and Improved Camshift Algorithm**." Audio, Language and Image Processing (ICALIP), 2012 International Conference, 16-18 July 2012.

[23] Marius Muja, David G. Lowe. "**Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration**." VISAPP (1), page 331-340. INSTICC Press, 2009.

[24] Z. Zhang, "**A flexible new technique for camera calibration**", IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330–1334, 2000.

[25] Donald W. Marquardt, "**An Algorithm for Least-Squares Estimation of Nonlinear Parameters**." Journal of the Society for Industrial and Applied Mathematics, Vol. 11.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ-นามสกุล	นายชนินทร์ สุขี
วัน เดือน ปีเกิด	1 สิงหาคม 2536 ที่กรุงเทพมหานคร
ที่อยู่	8/81 ถนนชัยพฤกษ์ แขวงตลิ่งชัน เขตตลิ่งชัน กรุงเทพมหานคร 10170
โทร	0850642111
อีเมล์	peet0007@gmail.com
ประวัติการศึกษา	
2558	วิทยาศาสตรบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ชื่อ-นามสกุล	นายวงศธร แซ่เล้า
วัน เดือน ปีเกิด	14 พฤษภาคม 2537 ที่กรุงเทพมหานคร
ที่อยู่	55-55/1 ซอย14 ถนนสาธุประดิษฐ์ แขวงบางโคล่ เขตบางคอแหลม กรุงเทพมหานคร 10120
โทร	0841119378
อีเมล์	himhhm@gmail.com
ประวัติการศึกษา	
2558	วิทยาศาสตรบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การตรวจจับการเข้าและออกของคนโดยการติดตาม

วงศธร แซ่เล้า และ ชนินทร์ สุขี

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

Emails: himhhm@gmail.com, peet0007@gmail.com

บทคัดย่อ

ปัจจุบันกล้องวงจรปิดเป็นที่นิยมใช้กันอย่างแพร่หลายทั่วโลกและการตรวจจับการเข้าและออกของคนเป็นสิ่งที่น่าสนใจในการวิจัยและพัฒนา กระบวนการตรวจจับคนเข้าและออกโดยส่วนใหญ่จะใช้การหักล้างกับภาพพื้นหลัง background subtraction หรือการตรวจจับคนบนภาพโดยใช้ Histograms of Oriented Gradients (HOG) เพื่อนำไปใช้ในการตรวจจับและติดตามคนเข้าและออก โดยหลักการติดตามก็มีหลากหลายวิธีการออกไป ซึ่งระบบของเราจะใช้การประมวลผลบนจุดที่สนใจ (feature level) โดยเราทำการหาจุดสนใจบนภาพโดยใช้วิธี Feature from Accelerated Segment Test (FAST) ซึ่งมีความเร็วในการค้นหา โดยนำมาเก็บเป็น background feature model แล้วสกัด foreground feature ออกมาเพื่อทำการติดตามจุดสนใจเหล่านั้นซึ่งการติดตามก็จะมีตัวแปรมากมายที่เข้ามาเกี่ยวข้องโดยงานวิจัยเราจะศึกษากระบวนการดังกล่าวและค่าตัวแปรต่างๆในการตรวจจับคนเข้าและออก

คำสำคัญ – Human Detection; Feature Tracking; Human Entry Tracking; Real-Time Detection

1. บทนำ

ปัจจุบัน การใช้กล้องวงจรปิดเพื่อการรักษาความปลอดภัยได้ถูกใช้กันในวงกว้างทั้งในด้านการทดลองเฝ้าสังเกตการณ์และการดูแลรักษาความปลอดภัย การตรวจสอบช่วงเวลาที่วัตถุเริ่มเคลื่อนไหวและดำเนินการการคัดแยกช่วงเวลานั้นออกมาเพื่อนำไปประยุกต์ใช้กับงานต่างๆ เช่นการตรวจสอบและติดตามลักษณะสิ่งที่เกิดขึ้นขณะวัตถุดังกล่าวเคลื่อนไหว เป็นต้น

ระบบตรวจจับและติดตามที่ถูกพัฒนาขึ้นมามากมาย ได้แสดงให้เห็นถึงความสามารถ การนำไปใช้ และการประยุกต์ที่แตกต่างกันมากมาย ซึ่งโดยส่วนมากแล้วจะประยุกต์ระบบมาจากการทำ background subtraction[1], Histogram of Oriented Gradient (HOG) กับ Kalman filter [2] แต่แม้ว่าจะเป็นระบบที่ถูกใช้งานอย่างแพร่หลาย ก็ยังคงมีปัญหาบางอย่างอยู่ เช่น การทำ background subtraction ที่ยังมีปัญหาเรื่องการเปลี่ยนแปลงของแสง และ HOG กับ Kalman filter ที่ไม่สามารถประมวลผลได้ถึง 30 fps

โดยระบบที่ทำการพัฒนาขึ้นมานี้จะใช้เทคนิค FAST feature detector ซึ่งเป็นระบบที่ใช้ในการตรวจหา feature ที่สนใจในภาพเพราะเทคนิคดังกล่าวมีความเร็วและเหมาะสมสำหรับการประมวลผลแบบ real-time video กว่าการหา feature ในเทคนิคอื่น ๆ โดยจะทำการเก็บ feature ที่พบไว้และคอยอัพเดต ซึ่งจะมีการประยุกต์ใช้ร่วมกับการวัดระยะทางในมุมมอง 3 มิติ โดยใช้พิกัดแบบ global coordinate ในการติดตามวัตถุ ซึ่งจากทั้งหมดที่กล่าวมานั้นจะสามารถทำการตรวจจับวัตถุได้อย่างมีประสิทธิภาพและแม่นยำในทุกๆมุมภาพแม้แต่ในมุมลึก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. บทความที่เกี่ยวข้อง

ปัจจุบันระบบที่ใช้ในการตรวจจับและติดตามการเคลื่อนไหวของคนได้ถูกตีพิมพ์นำเสนอออกมาเป็นจำนวนมาก ดังนั้นในหัวข้อนี้จึงได้รวบรวมข้อมูลสิ่งตีพิมพ์และบทความจากแหล่งที่มาต่างๆที่หลากหลายและมีความเกี่ยวข้องกับระบบของเรา และทำการเสนอตามลำดับการใช้งานภายในระบบการตรวจจับคนเข้าและออก โดยจะมีบทความดังต่อไปนี้

2.1 การตรวจหาวัตถุ

ในขณะที่มีการถ่ายภาพวิดีโอตามปกติ ในมุมมองของคอมพิวเตอร์ (computer vision) จะไม่สามารถแยกแยะความแตกต่างของแต่ละสิ่งในภาพได้ จนกว่าจะมีการใช้วิธีการต่างๆซึ่งถูกกำหนดโดยมนุษย์เพื่อให้คอมพิวเตอร์สามารถคำนวณแยกแยะและตรวจหาวัตถุภายในภาพตามที่ได้มีการออกแบบกำหนดได้ โดยปกติการตรวจหาวัตถุที่ระบบการตรวจจับส่วนใหญ่นิยมใช้กันอย่างกว้างขวางและได้ผลลัพธ์ที่มีประสิทธิภาพจะแบ่งออกเป็น 2 วิธีการดังนี้

การตรวจหาวัตถุด้วยวิธี background subtraction ซึ่งเป็นการหักล้างภาพพื้นหลังออกไปเพื่อให้ได้วัตถุหรือสิ่งที่สนใจ (feature) จากนั้นภายในภาพวิดีโอที่หักล้างพื้นหลังออกไปจะยังคงมี สิ่งรบกวน (noise) อยู่จึงต้องทำการลดมันลงไปเพื่อให้ได้วัตถุในภาพที่สมบูรณ์ โดยการตรวจหาวัตถุด้วยวิธี background subtraction จะมีการใช้งานร่วมกันกับ Particle Filter และ Camshift Algorithm โดยมีรายละเอียดดังนี้ 1. Background Subtraction-Based Multiple Object Tracking Using Particle Filter [3] จะเป็นการตรวจหา feature โดยใช้ particle filter คือ ข้อมูลของสี RGB ทั้ง 3 channel ทำการประมวลผล feature ที่สกัดออกมาจาก background โดยจะไม่สามารถทำแบบ real-time ได้หากมี feature เป็นจำนวนมาก และ 2. Moving Target Tracking based on Adaptive Background Subtraction and Improved Camshift Algorithm [4] เป็นการประยุกต์ Kalman filter มาใช้ร่วมกันกับการทำ background subtraction โดยมี Camshift Algorithm ที่ใช้บนข้อมูลสีแบบ HSV ในการประมวลผล ซึ่งทำให้ใช้ระยะเวลามากกว่าในการทำงาน real-time

การตรวจจับวัตถุที่ Feature โดยตรง จะเป็นการตรวจไปที่ feature ในวิดีโอและทำการติดตาม feature ที่ตรวจจับได้อย่างต่อเนื่อง โดยมีวิธีการที่น่าสนใจอยู่อย่างหนึ่งคือ Histogram of Oriented Gradient (HOG) คู่กับ Kalman Filter [2] โดย HOG จะใช้ Kalman Filter ในการกรอง Noise หรือสิ่งรบกวนในภาพออก แล้วจึงติดตามทิศทางกับการกระจัดของ feature ซึ่งจะใช้เวลาคำนวณผลนานถึง 1.6 วินาที ทำให้ล่าช้าเมื่อนำไปใช้กับวิดีโอที่ต้องการผลลัพธ์ในทันที (real-time)

2.2 ระบบการตรวจจับ feature

ก่อนที่จะสามารถติดตาม feature ได้นั้น จำเป็นต้องทำการตรวจหา feature ให้เจอก่อน ซึ่งในวิธีการที่จะใช้ตรวจหา feature ให้เจอได้นั้นจะมี 2 วิธีที่น่าสนใจดังต่อไปนี้

1. SIFT & SURF คือการตรวจจับโดยดึงคุณลักษณะของ feature (เช่น การกระจัด ทิศทาง และระยะทางของ feature) มาใช้เปรียบเทียบด้วยกันกับตัว feature ตัวเดียวกันแต่ในมุมมอง หรือ ระยะมองที่ต่างกัน ซึ่งจะมีปัญหาในกรณีของ SIFT คือเมื่อเปรียบเทียบ feature ในภาพที่มีแสงต่างกันจะทำให้ feature ไม่สามารถเข้าคู่กันได้ จึงได้มีการพัฒนา SURF [5] ที่ลบจุดบกพร่องของ SIFT และพัฒนาให้เร็วขึ้น

2. Feature from Accelerated Segment Test (FAST) Detection เป็นวิธีที่ถูกคิดค้นโดย E. Rosten, R. Porter, T. Drummond [6] เกิดจากการนำวิธีการ Harris Corner Detector มาประยุกต์และพัฒนา โดยวิธีการนี้สามารถคัดแยก feature ได้รวดเร็วที่สุดจึงทำให้เป็นวิธีที่เหมาะสมในระบบการตรวจจับคนเข้าและออกเพื่อให้ระบบสามารถแสดงผลได้ในทันที (real-time)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. การดำเนินงานวิจัย

การดำเนินงานวิจัยเริ่มต้นจากการติดตั้งโดยการทำ camera calibration ด้วยวิธีการของ Zhang [8] โดยเราได้ทำการติดตั้งภาพวงกลมสีดำ 6 วง บนพื้นระนาบ และใช้กล้องถ่ายวิดีโอลงมาจากมุมสูงเอียง ต่อมาคือการเริ่มจับภาพที่สามารถเห็นแผ่นวงกลมที่นำไปติดตั้งในฉากได้ แล้วจึงคำนวณลักษณะพื้นผิวจากเฟรมภาพที่ทำการบันทึกด้วยพารามิเตอร์ในกล้องและลักษณะบนแผ่นวงกลมเทียบกับพื้นผิวโดยใช้แผ่นวงกลม 6 แผ่น วางในรูปแบบเมทริก 2x3 ที่มีความกว้างเท่ากัน โดยให้จุดกึ่งกลางและแกนของกล้องเป็นตัวกำหนดพิกัดบนภาพ กำหนดให้จุด ๆ หนึ่งบนพื้นผิวจริงมีค่า X = [x,y,z] Xp= จุดที่ทราบตำแหน่งบนพื้นผิวแล้ว และ n= เวคเตอร์ norm ของพื้นผิวจริง

$$(X - X_p)^{*}n = 0 \qquad (1)$$

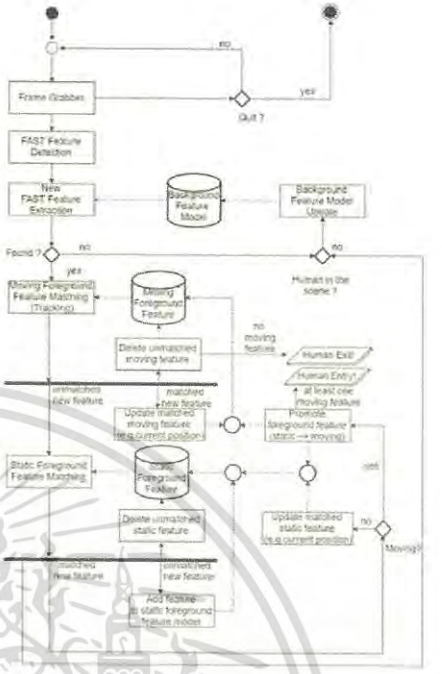
จะทำให้ได้สมการพื้นผิวจริง (3D) เพื่อใช้หาพื้นผิวระนาบ 3 มิติ บนหน้าจอกล้องได้ เมื่อได้สมการพื้นผิวจริง แต่ละตำแหน่งพิกเซล 2 มิติจะสามารถถูกเปลี่ยนเป็นตำแหน่งในพื้นผิว 3 มิติได้โดยการหาจุดตัดระหว่างเส้นที่แสงกระทบจากจุดกึ่งกลางของกล้องไปยังตำแหน่ง 3 มิติ ของจุด Feature กับเส้นพื้นผิวจริง กำหนดให้เส้นแสงกระทบดังกล่าวเป็น เวคเตอร์ L และ Xl เป็นจุด ๆ หนึ่งบนเส้นดังกล่าว ตำแหน่งที่เกิดการตัดกันจึงมีค่าดังสมการที่ 3.2

$$X' = kL + X_l \qquad (2)$$

$$\text{โดย } k = \frac{(X_p - X_l)\cdot n}{L \cdot n} \text{ ซึ่ง } (L \cdot n) \neq 0$$

ระบบตรวจจับคนเข้า/ออกนี้จะใช้สมการที่ 2 ในการหาตำแหน่ง 3 มิติของแต่ละ Feature ที่ซึ่งต่อมาจะถูกใช้ในการคำนวณหาระยะทาง 3 มิติเพื่อใช้ในการจำกัดระยะทางการค้นหา Feature R

3.1 โครงร่างและการทำงานของระบบ



รูปที่ 1. System Framework

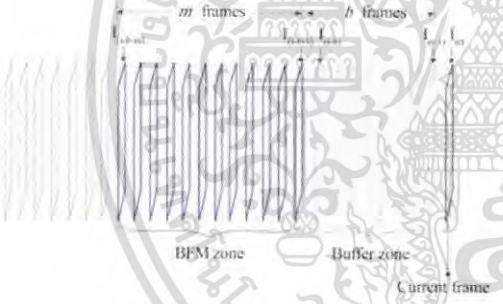
FAST features ที่ตรวจพบในฉากซึ่งเป็น feature ที่หยุดนิ่งไม่มีการเคลื่อนไหวจะถูกเก็บไว้ใน background โดยจะเรียก feature ที่ถูกเก็บไว้นี้ว่า Background Feature Model (BFM) ซึ่งสามารถทำการอัพเดตค่าของ feature ต่างๆในตัวเองได้อย่างต่อเนื่อง และ feature ใหม่ที่ปรากฏขึ้นจะถูกติดตามการเคลื่อนไหว แล้วจึงถูกนำไปเปรียบเทียบกับ feature ใน BFM ว่ามีการคลาดเคลื่อนเกินกว่าระยะที่กำหนดหรือไม่ เพื่อใช้หา FAST Features อันใหม่ที่พึ่งปรากฏขึ้นใหม่ ในฉากซึ่งเรียกว่า Foreground Feature(FF) แต่ในกรณีที่ไม่มี FF เฟรมดังกล่าวจะถูกใช้อัปเดต BFM ต่อไป และถ้า FF ถูกพบ มันก็เป็นไปได้ที่จะเป็น (1) หนึ่งใน Moving Foreground Features (MFFs) หรือ (2) หนึ่งใน Static Foreground Features (SFFs) หรือ (3) Foreground Feature อันใหม่ ในกรณีของ (1) ระบบจะแจ้งเตือนว่ามีคนเข้า และทำการติดตาม feature ต่อ แต่ในกรณีของ (3) ระบบจะเพิ่มพิกัดของ FF เข้าไปใน List ของ SFFs เพื่อทำการติดตาม FF เหล่านี้จนกระทั่งมันเริ่มทำการเคลื่อนไหว การสังเกตการณ์นี้คือกรณีของข้อ (2) ซึ่งเป็นการตรวจสอบเพิ่มเติมว่า SFFs ได้เคลื่อนออกจาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตำแหน่งดั้งเดิมหรือไม่ มันจะถูกเปลี่ยนให้เป็น MFFs ถ้าเริ่มเคลื่อนไหวเกินขอบเขตที่ระบบกำหนดไว้ ระบบจะแจ้งเตือนว่ามีคนเข้าโดยจะมีการเตือนว่าคนออกก็ต่อเมื่อ MFFs ทั้งหมดที่ถูกติดตามอยู่ ได้หายไปจากฉากแล้ว

3.2 Background Feature Model

Background Feature Model (BFM) ทำหน้าที่เก็บเฟรมจำนวน m เฟรม โดยเลื่อนและอับเดตตามเฟรมปัจจุบันซึ่งจะมีการอับเดตนี้ต่อเมื่อไม่มีคนเข้ามาในฉาก เพื่อให้ BFM เก็บตำแหน่ง feature ในเฟรมที่ไม่มีคนเท่านั้น (เสมือนเป็นที่เก็บ feature ของ background) แล้วจึงนำ feature ขณะที่มีคนเข้ามาเทียบกับ feature ใน BFM เพื่อให้ระบบสามารถรับรู้ได้ว่ามีคนเข้า เพราะตำแหน่ง feature ที่เพิ่มมาเปลี่ยนไป โดย BFM จะมีขั้นตอนเป็นไปดังรูปที่ 2

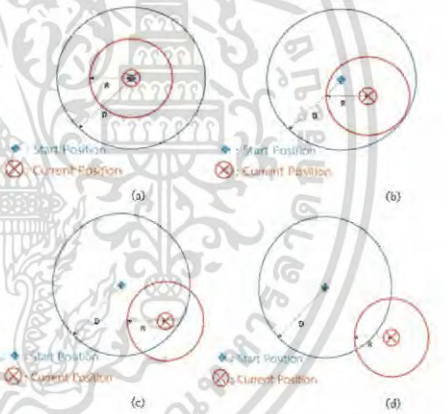


รูปที่ 2. การทำงานของ BFM

ในตอนนี้ ถ้าค่า $I_{(i)}$ เป็นเฟรมปัจจุบัน BFM จะประกอบไปด้วยข้อมูลของ FAST Features ที่อยู่ในเฟรมก่อนหน้า คือ $I_{(i-b-1)}$, $I_{(i-b-2)}$ ไปเรื่อย ๆจนถึง $I_{(i-b-m)}$ โดยค่า b คือจำนวนของเฟรมภาพที่อยู่ใน Buffer Zone เพราะการที่ระบบจะตรวจจับคนเข้าได้นั้นต้องใช้ระยะเวลาประมาณหนึ่งเนื่องจากระบบต้องตรวจสอบการเคลื่อนไหวเป็นระยะทางระยะหนึ่ง (D) ซึ่งจะทำให้ BFM สามารถเก็บเฟรมที่คนกำลังเข้ามาลงไปได้

ทั้งนี้ในการจะคัดแยก FF จากเฟรมปัจจุบันออกมาด้วยการเปรียบเทียบพิกัด FAST feature ในเฟรมปัจจุบันกับใน BFM จะยังไม่สามารถทำได้อย่างถูกต้องเนื่องจาก การเปลี่ยนแปลงของแสงและความไม่เสถียรของระบบ ซึ่งมีที่มาจาก ค่าความละเอียดของภาพที่มีสูงขึ้น ดังนั้นจึงมีการกำหนดขั้นตอนต่างๆที่ใช้สำหรับป้องกันความผิดพลาดต่างๆดังนี้

1. สร้างระยะค้นหารัศมี R ไว้ใช้ค้นหา คู่ feature ที่ตรงกันตามที่เห็นในรูปที่ 3 เพื่อเพิ่มความเร็วในกระบวนการและลดความเป็นไปได้ของการจับคู่ที่ผิดพลาดและรวดเร็วต่อการค้นหาจุด Foreground รอบๆเพื่อการติดตามโดยที่ ค่าของ R จึงถูกตัดสินใจโดยการนำค่าเข้าสู่สูตร (1.) ค่าความเร็วที่สูงที่สุดที่มนุษย์จะเดินได้ (v_{max}) (282.4 cm/s ซึ่งได้ถูกสำรวจไว้แล้วใน [10]) (2.) ระยะโฟกัส (f) และ frame rate (r_{camera}) ของกล้อง และ (3.) ระยะห่างระหว่างกล้องกับฉากภาพ (d_{scene}) ซึ่งในทางปฏิบัติ ค่าของ r ในหน่วยของพิกเซล จะเปลี่ยนไปโดยขึ้นอยู่กับ ความละเอียดของกล้อง



รูปที่ 3. การตรวจจับ MFF

จากรูประยะ R คือรัศมีที่ใช้ค้นหาและติดตาม FF และระยะ D คือค่าระยะทางที่ใช้วัดเพื่อแยกว่าเป็น MFF หากมีการเคลื่อนที่มากกว่าระยะ D นี้ + คือตำแหน่งที่ FF ถูกตรวจพบและติดตามในครั้งแรก Ⓧ คือตำแหน่งพิกัดเฟรมปัจจุบัน (Current Frame) โดยจะมีการคำนวณตามสมการที่ 3

$$R = \frac{(v_{max})(f)}{(d_{scene})(r_{camera})} \qquad (3)$$

ค่าของตัวแปรต่างๆจะมี 1. ค่าความเร็วที่สูงสุดที่มนุษย์จะเดินได้ (v_{max} = 282.4 cm/s) 2. ระยะโฟกัส (f) และ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

frame rate (r_{camera}) ของกล้อง และ 3. ระยะห่างระหว่างกล้องกับฉากภาพ (d_{scene}) โดยระยะห่างเริ่มต้น D สำหรับ Moving Feature จะต้องถูกเลือกอย่างระมัดระวัง เพื่อที่จะหลีกเลี่ยงการตรวจจับที่ผิดพลาด และสามารถรับรู้การเคลื่อนที่ของ feature ในทันทีที่เป็นไปได้ เราจึงต้องกำหนดค่า D=2R

3.3 การติดตาม FF ในการวัดระยะจริง (3D)

เนื่องจากการวัดระยะทั้งหมดถูกกระทำในระดับ pixel แต่ระบบที่ได้ทำการ camera calibration ให้สามารถหาสมการพื้นผิว ground plane เพื่อใช้ในการวัดในระยะจริงที่ให้ความถูกต้องแม่นยำที่สูงกว่าในมุมมอง 3 มิติ

โดยหลังจากที่ได้พิกัดบน image coordinate เราจะสามารถสร้างเส้นตรงตัดผ่านระหว่างจุด (0, 0, 0) ผ่านพิกัดบน image coordinate ไปยังเส้นสมการ ground plane ซึ่งจะทำให้ระบบรู้ถึงพิกัด global coordinate ของ Foreground feature เพื่อใช้ในการวัดระยะทาง 3 มิติได้ ซึ่งตำแหน่งต่างๆบน ground plane ในระยะจริงที่ได้จากการคำนวณและการนำ global coordinate ที่นำมาใช้เหล่านี้จะยังถือว่าเป็นค่าที่ได้จากการประมาณการ แต่ถือว่ามีความใกล้เคียงมากที่สุด ณ ขณะนี้

จากการประยุกต์ใช้ global coordinate และการ calibrate camera ทำให้เกิดการปรับปรุงในระบบให้กล้องสามารถดำเนินการระยะทางในมุมมอง 3 มิติ เพื่อใช้ติดตาม feature ให้มีความถูกต้องแม่นยำมากขึ้นได้ ทำให้ต้องมีการเปลี่ยนวิธีการติดตาม feature จากเดิม (ในการติดตามระดับ pixel หรือ 2 มิติ) ที่ใช้วิธีติดตามเฉพาะแต่ละ feature เฉพาะจุดที่อยู่ใกล้ที่สุดไปเรื่อยๆ ซึ่งให้ผลลัพธ์ที่ดีในการตรวจจับการเคลื่อนไหวในแนวนอนแต่ประสิทธิภาพของระบบจะลดลงเมื่อมีคน

เข้ามาในมุมลึก ดังนั้นการจะสามารถติดตาม Foreground feature ในมุมมอง 3 มิติได้ จึงต้องนำ global coordinate เข้ามาช่วยในการวัดระยะ จากนั้นจึงจะสามารถเริ่มขั้นตอนการตรวจสอบคนเข้าออกได้

3.4 ลำดับการเลือก FF เพื่อการติดตาม feature

ในการติดตาม FF ต้องเข้าใจก่อนว่า เราไม่สามารถเลือก FF ในเฟรมปัจจุบันที่ซ้ำกันได้เนื่องจากเราใช้ MFF เป็นตัวบ่งบอกการเข้าออกของ object ในระบบเราจึงจัดลำดับสิทธิในการเลือก FF ไว้โดยที่ให้สิทธิกับ FF ที่เป็น MFF ก่อนแล้วในการเลือก foreground feature ในเฟรมปัจจุบันก่อน แล้วจึงค่อยทำการเลือก FF ที่เหลือ

3.5 Feature Matching

ในการติดตาม FF โดยปกติแล้วจะติดตามแบบ nearest matching คือติดตาม feature ที่เกิดขึ้นใกล้ที่สุดในระยะ R แต่วิธีการดังกล่าวจะมีประสิทธิภาพน้อยกว่าเมื่อต้องมีการติดตาม MFF ของคนที่กำลังเดินออกไปหรือเดินเข้ามาในแนวลึก เพราะเมื่อมองในมุมมอง 2 มิติ feature จะไม่ขยับออกมาเกินระยะ D ทำให้เกิดการตรวจจับผิดพลาดคือทั้งที่มี คนหรือ วัตถุเข้ามา แต่ระบบมองไม่เห็น ดังนั้นการใช้ Feature Matching โดยการนำวิธี SURF มาหาข้อมูลของ feature (การกระจัด ทิศทาง ฯลฯ) เพื่อเปรียบเทียบข้อมูลดังกล่าวระหว่าง feature ในเฟรมปัจจุบัน กับใน BFM ภายในระยะ R ด้วยวิธี [8]

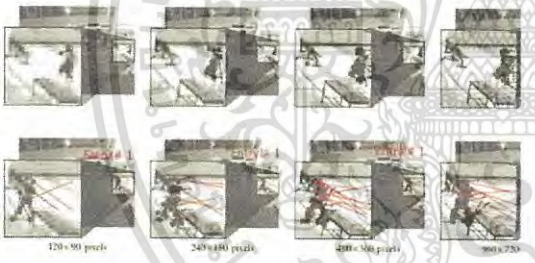
4. การทดลองและผลลัพธ์

ระบบการตรวจจับคนเข้าออกนี้ถูกใช้กับคอมพิวเตอร์รุ่น Intel Core I7-4700HQ @ 2.40GHz ใช้ภาพ Grey level ที่มีขนาดความละเอียด (resolution) ที่ 120×90, 240×180, 480×360, and 960×720 พิกเซล ด้วยกล้อง CMOS ผ่านทางสาย USB โดยลำดับภาพทั้งหมดที่ถูกใช้ในการทดลองนี้ถูกเก็บจากภายในอาคาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1 ผลลัพธ์การทดลองความละเอียด

มี Parameter 2 ตัวคือ $C_{in(min)}$ และ $C_{out(min)}$ ที่จะถูกทดสอบและผันแปรไปตาม Video Resolution ต่างๆ (คือ 120x90, 240x180, 480x360, 960x720 pixels) จะสังเกตได้ว่า $C_{in(min)}$ คือจำนวนที่น้อยที่สุดของเฟรมต่อเนื่องที่เกิด Moving Feature ก่อนที่ระบบจะทำการแจ้งเตือนว่ามีคนเข้า โดยดูให้แน่ใจได้ว่า FF จะถูกเลื่อนขั้นจาก SFF ไปเป็น MFFไม่ใช่เกิดจากข้อผิดพลาด และ $C_{out(min)}$ ก็เป็นจำนวนที่น้อยที่สุดของเฟรมต่อเนื่องที่ MFF ไม่สามารถหาพิกัดต่อไปได้หากไม่สามารถหาได้เป็นระยะเวลามากกว่าหรือเท่ากับ $C_{out(min)}$ เฟรม ระบบถึงจะทำการลบจุด MFF นี้ ออก โดยจากการทดลองสามารถสรุปได้ว่า ยิ่งความละเอียดมีมากขึ้นก็จะสามารถติดตาม feature และตรวจพบคนเข้าได้ง่ายขึ้น โดยค่า parameter ที่เหมาะสมที่สุดคือ $C_{in(min)} = 4$ และ $C_{out(min)} = 3$ และเก็บ BFMไว้มากกว่าหรือเท่ากับ 220 เฟรมจะทำให้ระบบทำงานได้มีประสิทธิภาพที่สุด



รูปที่ 4. การทดลองบนความละเอียดที่ต่างกัน

4.2 การเปรียบเทียบกับการทดลองในมุมมอง 3 มิติ

ในการติดตาม feature ในมุมมอง 3 มิติ ของกล้องด้วยการนำ global coordinate มาใช้ในการวัดระยะ (3D) และการนำ feature matching มาใช้เปรียบเทียบลักษณะของ FF เฟรมปัจจุบันและเฟรมใน BFM ช่วยให้การตรวจพบ MFF มีความแม่นยำและถูกต้องมากขึ้น ระบบจะมีประสิทธิภาพมากขึ้นเมื่อเทียบกับ การจับคู่แบบระยะใกล้สุดและการวัดระยะระดับ pixel (2D) ซึ่งผลลัพธ์ที่กล่าวมาจะสังเกตได้ในรูปที่ 5.

รูปที่ 5. การเปรียบเทียบรูปแบบการติดตาม FF

จากรูปเป็นการเทียบระหว่างระบบตรวจจับคนเข้าและออกที่ใช้การวัดระยะของกล้องในระดับ pixel-level (2D) กับระบบที่ใช้การวัดระยะตามจริง (3D) โดยมีรูปแบบการติดตาม 2 แบบ คือ 1. การติดตามด้วยการเทียบจากระยะห่าง feature (Nearest Matching) และ 2. การติดตามโดยเทียบจากความคล้ายกันของลักษณะ เช่นการกระจัด (Feature Matching)

5. สรุปผลการทดลอง

ระบบของเรานั้นสามารถประมวลผลได้แบบ real-time (30fps) เพราะว่ากระบวนการทำงานของเราไม่ได้ใช้ pixel ทั้งหมดในการประมวลผล แต่เป็นการใช้เพียงจุดสนใจ (feature) ที่ได้มาจากกระบวนการของ FAST ซึ่งมีความเร็วในการสกัดจุดสนใจออกมาจากภาพที่รวดเร็ว นอกจากนั้นการติดตาม foreground feature โดยมาตรฐานเราใช้การเลือกจุดสนใจที่ใกล้ที่สุดและวัดระยะทางในระดับ pixel ซึ่งจะทำให้ความเร็วในการประมวลผลดีที่สุด ซึ่งเราก็นำเสนอถึงการเพิ่มประสิทธิภาพ โดยการใช้ global coordinate(วัดระยะจริง) โดยต้องทำกระบวนการติดตั้งแผ่นวงกลมให้ระบบรู้จักพื้นที่จริง และการใช้ feature matching ในการเลือกจุดสนใจแทนการเลือกตัวที่ใกล้ที่สุด

เอกสารอ้างอิง

[1] Zhang, Lijing, and Yingli Liang. "Motion Human Detection Based on Background Subtraction." In 2010 ETCS, 1:284–87

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[2] Li, Changyan, Lijun Guo, and Yichen Hu. "A New Method Combining HOG and Kalman Filter for Video-Based Human Detection and Tracking." In 2010 3rd CISP, 1:290–93.

[3] Intaek Kim, Tayyab Wahab Awan and Youngsung Soh. "Background Subtraction-Based Multiple Object Tracking Using Particle Filter." IWSSIP 2014.

[4] Gang XU, Dong ZHAO, Qi ZHOU, Ding HUANG. "Moving Target Tracking based on Adaptive Background Subtraction and Improved Camshift Algorithm." ICALIP.

[5] A. Ess, T. Tuytelaars and L. Van Gool. "Speeded Up Robust Features by H. Bay." Computer Vision and Image Understanding, 2008.

[6] E. Rosten, R. Porter, and T. Drummond. "Faster and better: A machine learning approach to corner detection." IEEE, 2010.

[7] Z. Zhang, "A flexible new technique for camera calibration", IEEE, 2000.

[8] Marius Muja, David G. Lowe. "Fast Approximate Nearest Neighbors." INSTICC, 2009.

[9] Donald W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters." Journal of the Society for Industrial and Applied Mathematics.

[10] R. W. Bohannon, "Comfortable and maximum walking speed of adults aged 20—79 years: reference values and determinants." Age and Ageing, Jan. 1997.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้