การปรับปรุงวิธีการเพื่อแก้ไขปัญหาการรวมตัวของผู้เล่นฟุตบอล

วิทยานิพนธ์

ของ

ปิยะ แก้วบัวดี

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

ธันวาคม 2562

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

พหูนํ ปณฺฑิโต ชีเว

การปรับปรุงวิธีการเพื่อแก้ไขปัญหาการรวมตัวของผู้เล่นฟุตบอล

วิทยานิพนธ์

ของ

ปิยะ แก้วบัวดี

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

ธันวาคม 2562

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

IMPROVING OCCLUSION OF FOOTBALL PLAYERS

Piya Kaewbuadee

A Thesis Submitted in Partial Fulfillment of Requirements

for Doctor of Philosophy (Computer Science)

December 2019

Copyright of Mahasarakham University

คณะกรรมการสอบวิทยานิพนธ์ ได้พิจารณาวิทยานิพนธ์ของนายปิยะ แก้วบัวดี แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ของมหาวิทยาลัยมหาสารคาม

คณะกรรมการสอบวิทยานิพนธ์

....ประธานกรรมการ
(ผศ. ดร. วรารัตน์ สงฆ์แป้น)

....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผศ. ดร. พัฒนพงษ์ ชมภูวิเศษ)

....อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(ผศ. ดร. พนิดา ทรงรัมย์)

....กรรมการ
(ผศ. ดร. ฉัตรเกล้า เจริญผล)

....กรรมการ
(ผศ. ดร. สุชาติ คุ้มมะณี)

มหาวิทยาลัยอนุมัติให้รับวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา ปรัชญาดุษฎีบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ของมหาวิทยาลัยมหาสารคาม

....	....
(ผศ. ศศิธร แก้วมั่น)	(ผศ. ดร. กริสน์ ชัยมูล)
คณบดีคณะวิทยาการสารสนเทศ	คณบดีบัณฑิตวิทยาลัย

พหูนํ ปญฺญิโต ชีเว

ชื่อเรื่อง การปรับปรุงวิธีการเพื่อแก้ไขปัญหาการรวมตัวของผู้เล่นฟุตบอล
ผู้วิจัย ปิยะ แก้วบัวดี
อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. พัฒนพงษ์ ชมภูวิเศษ
ผู้ช่วยศาสตราจารย์ ดร. พนิดา ทรงรัมย์
ปริญญา ปรัชญาดุษฎีบัณฑิต **สาขาวิชา** วิทยาการคอมพิวเตอร์
มหาวิทยาลัย มหาวิทยาลัยมหาสารคาม **ปีที่พิมพ์** 2562

บทคัดย่อ

งานวิจัยนี้ มุ่งเน้นการปรับปรุงวิธีการเพื่อแก้ไขปัญหาการการบดบังซึ่งเกิดจากการรวมตัวของผู้เล่นในกีฬาฟุตบอล ประกอบด้วย 4 กระบวนการหลัก ได้แก่ 1) การตรวจจับสนาม ประยุกต์ใช้โมเดลผสมเกาส์เซียน (GMM) ร่วมกับเทคนิคการปรับแต่งค่าโดยใช้มาร์คอฟแรนดอมฟิวส์ (MRF) ในการจำแนกพิกเซลสนาม การเรียนรู้ของระบบจะอยู่บนพื้นฐานของโมเดลการทำนายที่ถูกนำเสนอซึ่งถูกสร้างขึ้นโดยใช้อัลกอริทึมนาอีฟเบย์โดยอาศัยข้อมูลตำแหน่งของสนามที่ปรากฏภายในภาพ 2) การตรวจจับและติดตามผู้เล่นในมุมมองภาพ ประยุกต์ใช้อัลกอริทึมคอนเดนเซชันร่วมกับเทคนิคการสุ่มแซมเปิลบนพื้นฐานของพื้นที่ที่ถูกนำเสนอ 3) การติดตามผู้เล่นในมุมมองเสมือน ประยุกต์ใช้คาร์มาลฟิลเตอร์สำหรับการติดตาม การผสมข้อมูลและการประมาณตำแหน่งได้นำเสนอโมเดลรูปร่างของข้อมูลบนพื้นฐานทางกายภาพร่วมกับเทคนิคการประมาณตำแหน่งบนพื้นฐานของสมการเส้นตรง และกระบวนการสุดท้าย 4) การแก้ปัญหาการบดบังจากการรวมตัวของผู้เล่น ประยุกต์ใช้อัลกอริทึมคอนเดนเซชันสำหรับการติดตามและแก้ปัญหาการบดบังในเบื้องต้น และอาศัยข้อมูลที่ได้จากกระบวนการผสมข้อมูลตำแหน่งผู้เล่นจากหลายกล้องร่วมกับผลจากการทำนายตำแหน่งผู้เล่นในมุมมองเสมือนสำหรับการประมาณตำแหน่งผู้เล่นที่ถูกบดบังภายในกลุ่มซึ่งไม่สามารถมองเห็นได้จากกล้องใดๆ ผลการทดลองพบว่า การใช้เทคนิคการปรับแต่งค่าด้วยมาร์คอฟแรนดอมฟิวส์ช่วยให้การจำแนกข้อมูลพิกเซลสนามมีประสิทธิภาพดีกว่าการใช้เทคนิคมีเดียนฟิลเตอร์ (MF) การตรวจจับและติดตามผู้เล่นในมุมมองภาพ พบว่าวิธีที่เสนอให้ผลในภาพรวมดีกว่า 2 เทคนิคเดิม คือ การใช้อัลกอริทึมคอนเดนเซชันเพียงอย่างเดียวและการใช้คาลมาลฟิลเตอร์ (KF) นอกจากนี้ยังพบว่าวิธีการสุ่มแซมเปิลบนพื้นฐานของพื้นที่ให้ผลที่ดีกว่าวิธีการสุ่มแซมเปิลบนพื้นฐานของบราวดิ้งบล็อก การติดตามผู้เล่นในมุมมองเสมือน พบว่ากระบวนการผสมข้อมูลที่เสนอในภาพรวมให้ผลการประมาณตำแหน่งบนมุมมองเสมือนที่ใกล้เคียงกับตำแหน่งจริงมากกว่าวิธีเดิมคือการประมาณตำแหน่งโดยใช้ค่าเฉลี่ย สุดท้าย การแก้ปัญหาการบดบัง พบว่าการประยุกต์ใช้อัลกอริทึมคอนเดนเซชันร่วมกับเทคนิคการสุ่มตัวอย่างบนพื้นฐานของพื้นที่ยังสามารถติดตามผู้เล่นที่ถูกบดบังในเบื้องต้นได้

ดี นอกจากนี้การใช้ข้อมูลจากการผสมตำแหน่งผู้เล่นจากหลายกล้องร่วมกับผลการทำนายตำแหน่งในมุมมองเสมือนสามารถช่วยในการประมาณตำแหน่งผู้เล่นที่ถูกบดบังแบบสมบูรณ์ได้

คำสำคัญ : การรวมตัวของผู้เล่น, โมเดลผสมแบบเกาเซียน, มาร์คอฟแรนดอมฟิวส์, อัลกอริทึมคอนเดนเซชัน, เทคนิคการสุ่มแซมเปิลบนพื้นฐานของพื้นที่, โมเดลรูปร่างของข้อมูลบนพื้นฐานทางกายภาพ, การประมาณตำแหน่งบนพื้นฐานของสมการเส้นตรง

TITLE IMPROVING OCCLUSION OF FOOTBALL PLAYERS

AUTHOR Piya Kaewbuadee

ADVISORS Assistant Professor Phatthanaphong Chompoowises , Ph.D.

Assistant Professor Panida Songram , Ph.D.

DEGREE Doctor of Philosophy MAJOR Computer Science

UNIVERSITY Mahasarakham University YEAR 2019

ABSTRACT

This research focuses on the improvement of a solution to the occlusion problem due to the integration of players in football. There are 4 main processes to solve the problem: 1) playfield detection, applying a Gaussian Mixed Model (GMM) with a data refinement technique using Markov Random Field (MRF) for playfield's pixel classification. Prior information is used to provide the clue of playfield's pixels in the images. The prior information is created using the Naive Bayes algorithm based on the playfield pixel's position appearing in the image. 2) The player detection and tracking on image views, applying the condensation algorithm with the sampling technique based on the regions being presented. 3) The player tracking in virtual views, applying the Kalman Filter (KF) to perform tracking. The data fusion and position estimation, offered the physical based data shape model and the linear equations based position estimation techniques. And the final process, 4) solving player occlusion from the aggregation of players. This performs by applying the condensation algorithms for tracking and solves the occlusion problem. The technique is based on the player position information obtained from the data fusion process from multiple cameras. In addition, the results of the predicted player position in the virtual views (estimating an obscured player's position within a group of playing which is not visible from any camera) is used in this process. The results showed that using the data refinement techniques, MRF, can help to classify the playfield's pixel data more effectively than using the Median Filter (MF) technique. The player detection and tracking on image view provide good results in the overall

better than the two original techniques, using only condensation algorithms and KF. This study also found that the sampling method based on region yielded better results than the sampling method based on the Bounding Box. The player tracking in virtual view gave the promising results, by estimating the position of players in a virtual view close to the actual position than the original technique that uses an average distance-based method. In the occlusion problem, this work showed that applying the condensation algorithms in conjunction with the region based sampling technique can track the obstructed players well. In addition, using the positioning data from multi-camera in conjunction with predictive position in a virtual view can effectively help to estimate the completed obstructed position of players.

Keyword : Player's Integration, Gaussian Mixture Models (GMM), Markov Random Filed (MRF), Condensation Algorithm, Region based Sampling Technique, Physical based Data Shape Model, Straight Line Equation based Position Estimation

พหูน ปณฺ ฑิโต ชีเว

กิตติกรรมประกาศ

งานวิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาระดับปริญญาเอก หลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม การดำเนินการวิทยานิพนธ์ฉบับนี้ สำเร็จลุล่วงได้ด้วยดี ด้วยความอนุเคราะห์อย่างสูงยิ่งจากที่ปรึกษาหลัก ผู้ช่วยศาสตราจารย์ ดร.พัฒนพงษ์ ชมภูวิเศษ ที่ปรึกษาร่วม ผู้ช่วยศาสตราจารย์ ดร.พนิดา ทรงรัมย์ และคณะกรรมการผู้ทรงคุณวุฒิทั้งจากภายในและภายนอกสาขาวิชา ผู้ช่วยศาสตราจารย์ ดร.ฉัตรเกล้า เจริญผล ผู้ช่วยศาสตราจารย์ ดร.สุชาติ คุ้มมะณี ผู้ช่วยศาสตราจารย์ ดร.มนัสวี แก่นอำพรพันธ์ ผู้ช่วยศาสตราจารย์ ดร.จันทิมา พลพินิจ ผู้ช่วยศาสตราจารย์ ดร.รพีพร ช่ำชอง ดร.สมนึก พ่วงพรพิทักษ์ ผู้ช่วยศาสตราจารย์ ดร.วรารัตน์ สงฆ์แป้น และ ดร.นวภัค เอื้ออนันต์ ที่ได้แนะนำและให้คำปรึกษาแนวทางที่เป็นประโยชน์อย่างยิ่งต่อการแก้ไขข้อบกพร่องต่างๆ ด้วยความเอาใจใส่เป็นอย่างดีตลอดมาตั้งแต่ต้นจนสำเร็จเรียบร้อย ผู้วิจัยขอกราบขอบพระคุณไว้เป็นอย่างสูง

คุณค่าและประโยชน์จากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นเครื่องบูชาคุณบิดา มารดา ที่เคารพยิ่ง ตลอดจนบูรพาจารย์ และผู้มีพระคุณที่ให้การอบรมสั่งสอนให้ผู้วิจัยประสบความสำเร็จในปัจจุบันและเจริญก้าวหน้าในอนาคต

ปิยะ แก้วบัวดี

พหูน ปณฺฑิโต ชีเว

สารบัญ

หน้า

บทคัดย่อภาษาไทย .... ง

บทคัดย่อภาษาอังกฤษ .... ฉ

กิตติกรรมประกาศ .... ซ

สารบัญ .... ฌ

สารบัญตาราง .... ฎ

สารบัญภาพ .... ฐ

บทที่ 1 บทนำ (Introduction) 1

1.1 ปัญหาและความเป็นมา (Statement of the Problems) 1

1.2 วัตถุประสงค์ของการวิจัย (Research Objectives) 3

1.3 ขอบเขตการวิจัย (Scope of Research) 3

1.4 ความสำคัญของการวิจัย (Significance of the Research) 4

1.5 นิยามศัพท์ (Definitions) 4

บทที่ 2 ปริทัศน์เอกสารข้อมูล (Literature Review) 6

2.1 การเรียนรู้ของเครื่องจักร (Machine Learning) 6

2.2 การประมวลผลภาพ (Image Processing) 8

2.3 ทฤษฎีคอมพิวเตอร์วิชัน (Computer Vision Theory) 12

2.4 การแปลงมุมมอง (Perspective Transformation) 23

2.5 การสอบเทียบกล้อง (Camera Calibration) 25

2.6 การวัดประสิทธิภาพ (Performance Evaluation) 28

2.7 งานวิจัยที่เกี่ยวข้อง (Related Research) 32

บทที่ 3 วิธีดำเนินงานวิจัย (Research Methodology) 68

พหุน ปญฺญา ทีเว

3.1 กระบวนการตรวจจับสนาม (Playfield Detection Process) 69

3.2 กระบวนการตรวจจับและติดตามผู้เล่นในมุมมองภาพ (Player Detection & Tracking on Image View Process).. 71

3.3 กระบวนการติดตามผู้เล่นในมุมมองเสมือน (Player Tracking on Virtual View Process) 72

บทที่ 4 การตรวจจับสนาม (Playfield Detection)... 74

4.1 วิธีดำเนินการวิจัย (Research Methodology) 74

1. ขั้นตอนการจัดเตรียมข้อมูลก่อนการประมวลผล (Data Pre-Processing).................... 76

2. ขั้นตอนการสกัดเอกลักษณ์ (Feature Extraction)... 80

3. ขั้นตอนการจำแนกข้อมูล (Data Classification).. 82

4. ขั้นตอนการปรับแต่งข้อมูล (Data Refinement).. 84

4.2 กระบวนการทดลองและผลการทดลอง (Experimental Process and Results).............. 87

4.3 สรุปผลการดำเนินงาน (Summary Operating Results).. 93

บทที่ 5 การตรวจจับและติดตามผู้เล่น (Player Detection and Tracking).............................. 94

5.1 วิธีดำเนินการวิจัย (Research Methodology) 94

1. ขั้นตอนการตรวจจับผู้เล่น (Player Detection) 95

2. ขั้นตอนการจำแนกผู้เล่น (Player Classification)... 105

3. ขั้นตอนการแสดงผลผู้เล่น (Player Representation) 107

4. ขั้นตอนการติดตามผู้เล่น (Player Tracking) 107

5. ขั้นตอนการแสดงเส้นทางผู้เล่นในมุมมองภาพ (Image View based Player Track Representation).. 111

5.2 กระบวนการทดลองและผลการทดลอง (Experimental Process and Results)............ 112

5.3 สรุปผลการดำเนินงาน (Summary Operating Results).. 119

บทที่ 6 การติดตามผู้เล่นบนมุมมองเสมือน (Player Tracking on Virtual View)...................... 121

6.1 วิธีดำเนินการวิจัย (Research Methodology) 121

พหุน ปณฺ ศฺ ถโต ชีเว

1. เทคนิคการติดตามผู้เล่นบนมุมมองเสมือน (Player Tracking Technique on Virtual View)....121
1.1 ขั้นตอนการจัดเตรียมข้อมูลก่อนการประมวลผล (Data Pre-Processing)....122
1.2 ขั้นตอนการสกัดเอกลักษณ์ (Feature Extraction)....130
1.3 ขั้นตอนการแปลงมุมมอง (Perspective Transform)....131
1.4 ขั้นตอนการผสมข้อมูล (Data Fusion)....138
1.5 ขั้นตอนการเชื่อมโยงข้อมูล (Data Association)....153
1.6 ขั้นตอนการเลือกเป้าหมาย (Target Selection)....154
1.7 ขั้นตอนการระบุตัวตนผู้เล่น (Player Identification)....154
1.8 ขั้นตอนการแสดงผลผู้เล่น (Player Representation)....155
1.9 ขั้นตอนการติดตามผู้เล่น (Player Tracking)....155
2. แนวทางการแก้ปัญหาการบดบัง (Occlusion Solution)....155
6.2 กระบวนการทดลองและผลการทดลอง (Experimental Process and Results)....156
6.3 สรุปผลการดำเนินงาน (Summary Operating Results)....160
บทที่ 7 สรุปผลและอภิปรายผล (Conclusions and Discussions)....162
7.1 สรุปผล (Conclusions)....162
7.2 อภิปรายผล (Discussions)....163
7.3 เทคนิคที่เสนอในงานวิจัย (Techniques Proposed in the Research)....164
7.4 ปัญหาที่พบ (Problems Encountered)....165
7.5 ข้อเสนอแนะ (Suggestions)....167
บรรณานุกรม....170
ประวัติผู้เขียน....183

พหูน ปณ ฑิโต ชีเว

สารบัญตาราง

หน้า

ตารางที่ 2.1 ตาราง Confusion Matrix 29

ตารางที่ 4.1 เอกลักษณ์ของ Haralick และคณะ [89] 81

ตารางที่ 4.2 ผลลัพธ์ของการจำแนกข้อมูลโดยใช้พื้นที่ที่มีขนาดแตกต่างกัน 88

ตารางที่ 4.3 ผลลัพธ์ของการจำแนกข้อมูลและการปรับปรุงข้อมูล 90

ตารางที่ 5.1 ผลการประเมินประสิทธิภาพการติดตาม 112

ตารางที่ 5.2 การประเมินประสิทธิภาพด้านเวลา 116

สารบัญภาพ

หน้า

ภาพที่ 2.1 ตัวอย่างการแสดงภาพขาวดำ (ที่มา: [36]) ... 8

ภาพที่ 2.2 ตัวอย่างการแสดงภาพสีเทา (ที่มา: [36]) ... 9

ภาพที่ 2.3 ตัวอย่างการแสดงภาพสี (ที่มา: [36]) ... 9

ภาพที่ 2.4 ตัวอย่างการแสดงภาพแบบดัชนี (ที่มา: [36]) ... 10

ภาพที่ 2.5 โมเดลสีแบบ RGB (ที่มา: [36]) ... 10

ภาพที่ 2.6 โมเดลสีแบบ YCbCr (ที่มา: [36]) ... 11

ภาพที่ 2.7 โมเดลสีแบบ HSI (ที่มา: [36]) ... 11

ภาพที่ 2.8 โมเดลสีแบบ HSV (ที่มา: [36]) ... 12

ภาพที่ 2.9 ขั้นตอนการลบแบ็คกราวนด์ (ที่มา: [61]) ... 15

ภาพที่ 2.10 การแสดงผลวัตถุ (Object Representations) (ที่มา: Yilmaz และคณะ [7]) ... 17

ภาพที่ 2.11 การแสดงผลวัตถุแบบ 3 มิติ (Object Representations in 3D) (ที่มา: Khan [19]) . 17

ภาพที่ 2.12 ขั้นตอนการติดตามวัตถุในวิดีโอ (ที่มา: Dash และคณะ [16]) ... 18

ภาพที่ 2.13 ขั้นตอนการประมวลผลของตัวกรองคาร์ลมาล (ที่มา:กฤษฎา แสงเพ็ชร์ส่อง [66]) ... 20

ภาพที่ 2.14 โครงสร้างกราฟ ... 22

ภาพที่ 2.15 แสดงความสัมพันธ์ระหว่างจุดบนพิกัดภาพ (P) และจุดบนพิกัดจริง (P’) ... 24

ภาพที่ 2.16 ตารางขนาด 30 x 30 mm. สำหรับสอบเทียบ (ที่มา:สุเทพ อาษาและคณะ [57]) ... 25

ภาพที่ 2.17 ตัวอย่างขั้นตอนและผลลัพธ์ที่ได้จากการสอบเทียบ (ที่มา:สุเทพ อาษาและคณะ [57]) 26

ภาพที่ 2.18 โมเดลของกล้องรูเข็ม, (ที่มา: Z. Zhang [63]) ... 26

ภาพที่ 2.19 การแบ่งกลุ่มประเภทงานวิจัย ... 32

ภาพที่ 2.20 ระบบพิกัดโลก (a) และระบบพิกัดจอ (b) (ที่มา: Tahar และ Imed [37]) ... 54

ภาพที่ 2.21 การวางแนวกล้อง 3 ตัวสำหรับการสอบเทียบเบื้องต้น (ที่มา: Tahar และ Imed [37])56

ภาพที่ 2.22 การแก้ไขการบิดเบือนแบบไม่เป็นเชิงเส้น (ที่มา: Tahar และ Imed [37]) 58
ภาพที่ 2.23 จุดอ้างอิงบนสนาม (ที่มา: Tahar และ Imed [37]).. 58
ภาพที่ 2.24 จุดสำหรับการสอบเทียบ (ที่มา: Nishino และคณะ [41]) 59
ภาพที่ 3.1 การบดบังที่เกิดจากการรวมกลุ่ม... 68
ภาพที่ 3.2 กระบวนการดำเนินงานวิจัย... 69
ภาพที่ 3.3 พื้นที่บริเวณสนาม... 70
ภาพที่ 4.1 กระบวนการตรวจจับสนาม 75
ภาพที่ 4.2 ตัวอย่างการแปลงภาพ... 78
ภาพที่ 4.3 การดึงเอกลักษณ์จากภาพต้นฉบับ 79
ภาพที่ 4.4 ภาพต้นฉบับที่กำหนดลาเบล... 80
ภาพที่ 4.5 การแบ่งส่วนภาพต้นฉบับที่กำหนดลาเบล... 80
ภาพที่ 4.6 การแสดงพิกเซลด้วยกราฟและการแสดงเทอมพลังงานในรูปแบบ unary, binary 85
ภาพที่ 4.7 ตัวอย่างผลลัพธ์ที่ได้จากการจำแนกข้อมูล... 88
ภาพที่ 4.8 ผลการจำแนกข้อมูลโดยใช้เพียงเอกลักษณ์เปรียบเทียบกับการใช้กระบวนการปรับปรุงข้อมูล 89
ภาพที่ 4.9 การเปรียบเทียบผลการจำแนกข้อมูล.. 91
ภาพที่ 4.10 ชุดภาพสำหรับสอนและทดสอบระบบ 92
ภาพที่ 4.11 ผลการตรวจจับสนามภาพเฟรมแรกของ ISSIA Dataset [82]... 92
ภาพที่ 5.1 กระบวนการตรวจจับและติดตามผู้เล่น.. 94
ภาพที่ 5.2 ขอบเขตภายในภาพจากกล้อง C1 (ISSIA Dataset [82]) 96
ภาพที่ 5.3 กระบวนการตรวจจับผู้เล่น.. 97
ภาพที่ 5.4 หลักการประมาณตำแหน่งผู้เล่น 105
ภาพที่ 5.5 การแสดงผลบราวดิ้งบล็อกผู้เล่น... 107
ภาพที่ 5.6 แซมเปิลรอบตัวอย่างนอกขอบเขตการตรวจจับ 109

พหูน ปณ ฑิโต ชีเว

ภาพที่ 5.7 วิธีการสุ่มบนพื้นฐานของบราวดิ้งบล็อกและพื้นที่ 111
ภาพที่ 5.8 ตัวอย่างการแสดงเส้นทางผู้เล่นในมุมมองภาพ 111
ภาพที่ 5.9 Distance Cost Average 113
ภาพที่ 5.10 Absolute Error Average 114
ภาพที่ 5.11 Bounding Box Error Average 114
ภาพที่ 5.12 Object Tracking Error Average 115
ภาพที่ 5.13 Root Mean Square Error Average 115
ภาพที่ 5.14 ตัวกรองคาร์ลมาลฟิลเตอร์: Frame 1186 116
ภาพที่ 5.15 อัลกอริทึมคอนเดนเซชัน: Frame 1186 117
ภาพที่ 5.16 เทคนิคที่เสนอ: Frame 1186 117
ภาพที่ 5.17 ตัวกรองคาร์ลมาลฟิลเตอร์: Frame 1263 117
ภาพที่ 5.18 อัลกอริทึมคอนเดนเซชัน: Frame 1263 118
ภาพที่ 5.19 เทคนิคที่เสนอ: Frame 1263 118
ภาพที่ 5.20 รายการจุดมุมทั้ง 4 มุมจากเฟรมแรกของ ISSIA Dataset [82] 119
ภาพที่ 6.1 กระบวนการติดตามผู้เล่นบนมุมมองเสมือน 122
ภาพที่ 6.2 การจำลองพื้นที่สนาม 123
ภาพที่ 6.3 แผนผังตำแหน่งกล้องทั้ง 9 ตัว 124
ภาพที่ 6.4 กล้อง C1 124
ภาพที่ 6.5 กล้อง C2 125
ภาพที่ 6.6 กล้อง C3 125
ภาพที่ 6.7 กล้อง C4 126
ภาพที่ 6.8 กล้อง C5 126
ภาพที่ 6.9 กล้อง C6 127
ภาพที่ 6.10 กล้อง C7 127

พหุน ปณฺ ติโต ชีเว

ภาพที่ 6.11 กล้อง C8 128
ภาพที่ 6.12 กล้อง C9 (Top View) 128
ภาพที่ 6.13 ตัวอย่างภาพทดสอบ (กล้อง C1) 129
ภาพที่ 6.14 ความกว้างของลำตัวผู้เล่น 130
ภาพที่ 6.15 ตัวอย่างภาพมุมมองเสมือน (Virtual Plane/ Top View) 131
ภาพที่ 6.16 ตัวอย่างการกำหนดจุดคอนโทรลพอยต์ 132
ภาพที่ 6.17 ตัวอย่าง FOV ประเภท 1 และ 2 133
ภาพที่ 6.18 ตัวอย่าง FOV ประเภท 1, 2 และ 3 133
ภาพที่ 6.19 ตัวอย่างตำแหน่งผู้เล่นในมุมมองภาพ 136
ภาพที่ 6.20 ตัวอย่างตำแหน่งผู้เล่นในมุมมองเสมือน 137
ภาพที่ 6.21 ตัวอย่างเส้นทางผู้เล่นในมุมมองภาพ 137
ภาพที่ 6.22 ตัวอย่างเส้นทางผู้เล่นในมุมมองเสมือน 138
ภาพที่ 6.23 วิธีกำหนดตำแหน่งวัตถุและผลการประมาณตำแหน่งบนมุมมองเสมือน 139
ภาพที่ 6.24 ตำแหน่งวัตถุที่ได้จากมุมมองกล้องที่ต่างกัน 140
ภาพที่ 6.25 ตำแหน่งของวัตถุที่อยู่ใกล้กัน 141
ภาพที่ 6.26 ผลจากตำแหน่งและทิศทางกล้องร่วมกับตำแหน่งทางกายภาพของวัตถุ 142
ภาพที่ 6.27 ลักษณะคอนทัวร์ที่ไม่สมบูรณ์ 143
ภาพที่ 6.28 หลักการของเข็มทิศสำหรับหาทิศทางของกล้อง 145
ภาพที่ 6.29 การใช้ตำแหน่งศีรษะผู้เล่นและลักษณะมุมกล้องที่ต่างกัน 145
ภาพที่ 6.30 การหาค่าเฉลี่ยจาก 2 ตำแหน่ง 146
ภาพที่ 6.31 ลักษณะโครงสร้างของตำแหน่งของข้อมูล 147
ภาพที่ 6.32 โมเดลรูปร่างของข้อมูล (Data Shape Model) 148
ภาพที่ 6.33 โมเดลรูปร่างข้อมูลจากกล้อง 9 ตัว 150
ภาพที่ 6.34 โมเดลรูปร่างข้อมูลที่ต่างกัน 151

พหูน ปณฺ ฑิโต ชีเว

ภาพที่ 6.35 การผสมข้อมูลจากกล้อง C1,C5 156
ภาพที่ 6.36 การผสมข้อมูลจากกล้อง C5,C6 157
ภาพที่ 6.37 การผสมข้อมูลจากกล้อง C5,C7 157
ภาพที่ 6.38 การผสมข้อมูลจากกล้อง C5,C8 158
ภาพที่ 6.39 Average Distance Error 159
ภาพที่ 7.1 ผลการประมาณตำแหน่งกล้อง C1 - C5 โดยการกำหนดเงื่อนไข 166
ภาพที่ 7.2 Average Distance Error หลังกำหนดเงื่อนไข 166
ภาพที่ 7.3 แนวคิดการแก้ปัญหาการรวมตัวของผู้เล่น 167

บทที่ 1

บทนำ (Introduction)

1.1 ปัญหาและความเป็นมา (Statement of the Problems)

ปัจจุบันรายการบันเทิงที่ผู้ชมจำนวนมากนิยมดูรายการหนึ่ง คือ รายการแข่งขันกีฬาประเภทต่างๆ เช่น ฟุตบอล บาสเกตบอลและกีฬาประเภทอื่นๆ เป็นต้น รายการแข่งขันกีฬาฟุตบอลเป็นหนึ่งในหลายๆ รายการที่เป็นที่นิยมดูกันทั่วโลกซึ่งพบเห็นได้บ่อยในรายการโทรทัศน์ตามช่องต่างๆ รวมถึงในปัจจุบันได้มีการถ่ายทอดสดผ่านเครือข่ายอินเตอร์เน็ตแล้วซึ่งในระหว่างการแข่งขันหรือหลังจากการแข่งขันจะมีกลุ่มคนต่างๆ เช่น ผู้ชม ผู้บรรยายการแข่งขัน โดยเฉพาะผู้ควบคุมทีมในการแข่งขันซึ่งจะวิเคราะห์และวิจารณ์การแข่งขันของแต่ละคู่ซึ่งในการวิเคราะห์รูปแบบและข้อผิดพลาดในการเล่นของทีมตนเองรวมถึงรูปแบบการเล่นของฝั่งตรงข้ามนั้นเป็นสิ่งที่ผู้ควบคุมทีมจำเป็นต้องรู้ วิธีการหนึ่งที่ผู้ควบคุมทีมนิยมใช้ในการดำเนินการนี้ คือ การเฝ้าดูในระหว่างการแข่งขันและการนำบันทึกวิดีโอการแข่งขันมาดูย้อนหลังอีกครั้งเพื่อพิจารณาหรือศึกษาข้อมูลต่างๆ ที่ตนเองสนใจ ส่วนผู้ชมกีฬาฟุตบอลโดยส่วนใหญ่จะเลือกชมเหตุการณ์หรือติดตามผู้เล่นที่ตนเองชื่นชอบและจะติดตามชมวิดีโอการแข่งขันย้อนหลังในรายการหรือแหล่งบริการต่างๆ เช่น อินเตอร์เน็ต เป็นต้น ซึ่งปัญหาที่ผู้ควบคุมทีมหรือผู้ชมโดยส่วนใหญ่มักประสบ คือ กรณีเกิดเหตุการณ์ที่ภาพในวิดีโอบางช่วงไม่ชัดเจน ผู้เล่นหันหลังให้กับกล้อง ภาพที่แสดงนั้นค่อนข้างไกลทำให้มองเห็นผู้เล่นได้ไม่ชัดเจน รวมถึงการบดบังระหว่างผู้เล่น ซึ่งปัญหาเหล่านี้ทำให้ผู้ชมไม่สามารถรับรู้ได้ว่าบุคคลในวิดีโอในฉากต่างๆ นั้นเป็นใครบ้างและมีใครอยู่ในเหตุการณ์ที่สำคัญที่เกิดขึ้นบ้างในการแข่งขันแต่ละครั้ง

จากความนิยมของผู้คนในการชมกีฬาฟุตบอลและการทำธุรกิจเกี่ยวกับกีฬาดังที่กล่าวไว้ในงานวิจัยของ Wang [1] รวมถึงปัญหาความต้องการของผู้ชมในการชมวิดีโอจึงเป็นสาเหตุหนึ่งที่ทำให้นักวิจัยจำนวนมากสนใจที่จะพัฒนาวิธีการหรือเครื่องมือที่จะช่วยอำนวยความสะดวกแก่ผู้ชมในการเลือกชมหรือค้นหาข้อมูลเกี่ยวกับกีฬาฟุตบอล เช่น งานวิจัยของ Wang [1] เสนอวิธีการสืบค้นเหตุการณ์ที่สำคัญในวิดีโอการแข่งขันฟุตบอล เช่น การสืบค้นเหตุการณ์ที่มีการยิงประตู การค้นหาผู้เล่นที่ครองบอลดังที่เสนอในงานวิจัยของ Yu และคณะ [25] การค้นหาลูกบอลในวิดีโอการแข่งขันฟุตบอลดังที่เสนอในงานวิจัยของ Ali และคณะ [21] Liu และคณะ [22] Huang และคณะ [26] และการค้นหาและติดตามบุคคลในวิดีโอการแข่งขันฟุตบอล เช่น ในงานวิจัยของ Ali และคณะ [21] Liu และคณะ [22] Utsumi และคณะ [24] Huang และคณะ [26] เป็นต้น จากงานวิจัยดังที่กล่าวมาข้างต้นสามารถแบ่งลักษณะงานวิจัยได้เป็น 2 กลุ่มหลัก คือ กลุ่มที่เน้นการค้นหาเหตุการณ์สำคัญในเกมส์การแข่งขันกีฬาและกลุ่มที่เน้นการตรวจจับและติดตามวัตถุที่เป็นบุคคลหรือลูกบอลโดยงานวิจัย

ในกลุ่มแรกจะเน้นเพียงการค้นหาเหตุการณ์ที่สำคัญในเกมส์การแข่งขัน เช่น เหตุการณ์ยิงประตู โดยไม่มุ่งเน้นในเรื่องของการตรวจจับและติดตามผู้เล่นแต่ยังต้องอาศัยข้อมูลจากการตรวจจับและติดตามผู้เล่นและลูกบอลเป็นพื้นฐานในการทำงานสำหรับการวิเคราะห์เหตุการณ์ต่างๆ ซึ่งกระบวนการตรวจจับและติดตามผู้เล่นและลูกบอลที่ถูกนำเสนอในหลายงานวิจัยนั้นส่วนใหญ่เพื่อจัดการกับประเด็นปัญหาที่สำคัญต่างๆ เช่น ความถูกต้องในการตรวจจับและติดตามผู้เล่น การระบุตัวตนของผู้เล่นเนื่องจากปัญหาการสลับตัวตนระหว่างผู้เล่น (Identity Switch) สิ่งรบกวนต่างๆ เช่น แสงและเงา การบดบังหรือการจับกลุ่มและความเหมือนระหว่างผู้เล่น เป็นต้น

สำหรับข้อมูลเส้นทางการเคลื่อนที่ของผู้เล่นถือเป็นข้อมูลที่สำคัญสำหรับใช้ในการวิเคราะห์เกมส์การแข่งขันและการค้นหาเหตการณ์สำคัญต่างๆ เช่น การบุก-รับ การยิงประตูและการล้ำหน้า เป็นต้น ปัญหาที่มักพบได้บ่อยในการติดตามผู้เล่น ได้แก่ การบดบังระหว่างผู้เล่น การสลับตัวตนระหว่างผู้เล่น การจับกลุ่มหรือการรวมตัวของผู้เล่นหลายๆ คนและความเหมือนของชุดนักฟุตบอลซึ่งประเด็นปัญหาเหล่านี้จะส่งผลต่อความถูกต้องในการติดตามและการระบุตัวตนผู้เล่น

ปัญหาการบดบังที่เกิดจากการจับกลุ่มของผู้เล่นเป็นปัญหาที่พบได้บ่อยครั้งในเกมส์การแข่งขันของกีฬาประเภททีมโดยเฉพาะกีฬาฟุตบอลซึ่งจะส่งผลกระทบต่อประสิทธิภาพของการตรวจจับ การติดตามและการระบุตัวตนของผู้เล่นอย่างมากโดยมีหลายแนวทางที่ถูกนำเสนอเพื่อแก้ไขปัญหานี้ เช่น การใช้กราฟสำหรับการติดตามและแก้ปัญหาการบดบัง เช่น ในงานวิจัยของ Mei Han และคณะ [40] การใช้ข้อมูลจากหลายมุมมองโดยใช้กล้องจำนวนหลายตัวเพื่อแก้ปัญหาการบดบัง เช่น ในงานวิจัยของ Iwase และ Saito [42] ซึ่งใช้กล้องจำนวน 15 ตัวในการติดตามผู้เล่น งานวิจัยของ Cheihk และคณะ [43] ซึ่งใช้ข้อมูลจากกล้อง 2 ตัวเพื่อการติดตามบุคคลในกิจกรรมต่างๆ เช่น การเดิน การจับมือและการบดบัง งานวิจัยของ Park และคณะ [44] เสนอวิธีการติดตามผู้เล่นโดยใช้กล้องจำนวน 4 ตัวและแก้ปัญหาการบดบังโดยการใช้ข้อมูลจากกล้องที่สามารถมองเห็นผู้เล่นที่ถูกบดบังได้ชัดเจน เป็นต้น ประเด็นที่สังเกตพบ คือ งานวิจัยส่วนใหญ่เน้นแก้ปัญหาการบดบังระหว่างผู้เล่นเพียง 2-3 คน แต่ยังไม่มีงานวิจัยใดที่กล่าวถึงปัญหาการจับกลุ่มหรือการรวมตัวของผู้เล่นที่มีจำนวนมากกว่า 2 หรือ 3 คนขึ้นไป โดยเหตุการณ์นี้สามารถเกิดขึ้นได้ในเกมส์การแข่งขันฟุตบอลซึ่งจะมีปัญหาที่สำคัญตามมา คือ ปัญหาการบดบัง การสลับตัวตน การติดตามเส้นทางที่ถูกต้อง เป็นต้น

จากประเด็นปัญหาดังกล่าวผู้วิจัยจึงได้ทำการศึกษาและพัฒนาวิธีการเพื่อแก้ไขปัญหาการบดบัง ซึ่งเกิดจากการรวมตัวของผู้เล่นในวิดีโอการแข่งขันฟุตบอลโดยอาศัยความสัมพันธ์ทางกายภาพของตำแหน่งของผู้เล่นในมุมมองเสมือนสำหรับจัดการปัญหาการบดบังของผู้เล่นและผสมผสานเทคนิคอัลกอริทึมคอนเดนเซชันร่วมกับเทคนิคการสุ่มแซมเปิลบนพื้นฐานของพื้นที่ (Region based Sampling Technique) ในการติดตามผู้เล่นในมุมมองภาพสำหรับแก้ปัญหาการบดบังในเบื้องต้น นอกจากนี้ยังเสนอโมเดลรูปร่างของข้อมูลบนพื้นฐานทางกายภาพร่วมกับเทคนิคการประมาณ

ตำแหน่งบนพื้นฐานของสมการเส้นตรงสำหรับกระบวนการผสมข้อมูลและการประมาณตำแหน่งในมุมมองเสมือน ซึ่งวิธีที่เสนอนี้สามารถช่วยจัดการกับปัญหาการบดบังได้ดีขึ้นและยังคงสามารถติดตามเส้นทางของผู้เล่นได้อย่างถูกต้อง

1.2 วัตถุประสงค์ของการวิจัย (Research Objectives)

เพื่อปรับปรุงวิธีการเพื่อแก้ปัญหาการบดบังจากการรวมตัวของผู้เล่นฟุตบอลในวิดีโอการแข่งขันกีฬาฟุตบอล

1.3 ขอบเขตการวิจัย (Scope of Research)

การดำเนินการวิจัยในครั้งนี้มีขอบเขตของการวิจัยดังนี้

1. มุ่งเน้นการปรับปรุงวิธีการเพื่อแก้ไขปัญหาการบดบังซึ่งเกิดจากการจับกลุ่มหรือการรวมตัวของผู้เล่นในกีฬาฟุตบอลเนื่องจากเป็นสาเหตุของปัญหาอื่นๆ เช่น การระบุตัวตนและการติดตามผู้เล่น เป็นต้น

2. การตรวจจับสนามฟุตบอลให้ความสนใจกับเอกลักษณ์ (Feature) พื้นฐานและวิธีการที่จะนำมาใช้ในการตรวจจับสนามเพื่อหาเอกลักษณ์และเทคนิคที่เหมาะสมซึ่งจะถูกใช้ตลอดงานวิจัยนี้

3. ประเด็นการแก้ไขปัญหาการรวมตัวของผู้เล่นในงานวิจัยนี้พิจารณาเฉพาะเหตุการณ์การรวมตัวของผู้เล่นที่ผู้เล่นทั้งหมดในกลุ่มอยู่ในลักษณะท่ายืนโดยไม่พิจารณาเหตุการณ์การรวมตัวของผู้เล่นที่มีผู้เล่นบางคนในกลุ่มอยู่ในลักษณะท่านั่ง ท่านอนหรือซ้อนทับกัน

4. ข้อมูลสำหรับทดสอบระบบที่พัฒนา ได้แก่ รายการภาพจากแหล่งต่างๆ ในอินเตอร์เน็ตสำหรับการดำเนินการตรวจจับสนาม และวิดีโอการแข่งขันกีฬาฟุตบอลสำหรับการดำเนินการตรวจจับและติดตามผู้เล่นจากฐานข้อมูล ISSEA Dataset [82] และหรือ VS-PETS’2003 [113] เป็นต้น โดยภาพวิดีโอมีรายละเอียดดังนี้

4.1 เป็นข้อมูลภาพวิดีโอที่สามารถตรวจจับพื้นที่สนามได้ครอบคลุมทั้งหมดหรือเกือบทั้งหมดเนื่องจากงานวิจัยนี้มีการใช้ข้อมูลจำนวนผู้เล่นเป็นเงื่อนไขในการเลือกเป้าหมายโดยฐานข้อมูล ISSEA Dataset [82] มีจำนวนกล้องที่ใช้ 6 ตัวและชุดข้อมูลสำหรับการทดสอบกระบวนการผสมข้อมูลใช้กล้องจำนวน 9 ตัว

4.2 ในภาพวิดีโอควรต้องมีช่วงเหตุการณ์ที่ผู้เล่นยืนประจำตำแหน่งก่อนการเริ่มเกมส์การแข่งขัน (ถ้ามี) เนื่องจากงานวิจัยนี้มีการใช้ข้อมูลที่เกี่ยวข้องในเหตุการณ์นี้สำหรับกระบวนการตรวจจับและระบุตัวตน

4.3 มุมมองของภาพเป็นแบบ Long-View เนื่องจากงานวิจัยนี้เน้นการตรวจจับผู้เล่นทั้งหมดในสนามในลักษณะภาพรวมเป็นหลัก

4.4 ภาพวิดีโอที่ได้จากกล้องแต่ละตัวเป็นภาพที่ถูกบันทึกด้วยกล้องที่ติดตั้งอยู่กับที่โดยตัวกล้องไม่มีการเคลื่อนที่และไม่มีการปรับเปลี่ยนมุมมองการถ่ายภาพ เช่น การเอียง (Tilt) การแพน (Pan) เป็นต้น

1.4 ความสำคัญของการวิจัย (Significance of the Research)

ผลจากการดำเนินการวิจัยในครั้งนี้ส่งผลให้เกิดประโยชน์ในด้านต่างๆ ดังนี้

1. เทคนิคการผสมข้อมูลที่นำเสนอช่วยให้กระบวนการติดตามผู้เล่นในลักษณะหลายมุมมองนั้นมีความถูกต้องมากขึ้นรวมถึงสามารถลดปริมาณของข้อมูลในการคำนวณและสามารถช่วยในการแก้ปัญหาการบดบังได้

2. การติดตามผู้เล่นบนพื้นฐานของอัลกอริทึมคอนเดนเซชัน (Condensation Algorithm) ช่วยแก้ปัญหาการบดบังและการทำนายตำแหน่งของผู้เล่นในเบื้องต้นได้ รวมถึงการใช้เทคนิคการสุ่มบนพื้นฐานของพื้นที่ที่เสนอสามารถช่วยลดปริมาณข้อมูลที่ใช้คำนวณและได้แซมเปิล (Sample) จริงของวัตถุมากขึ้น

3. เทคนิคการตรวจจับสนามโดยใช้วิธีการเรียนรู้บนพื้นฐานความน่าจะเป็นที่เสนอเป็นวิธีที่ง่าย ไม่ซับซ้อน นอกจากนี้การประยุกต์ใช้เทคนิคการปรับปรุงข้อมูลโดยใช้มาร์คอฟแรนดอมฟิลด์ (Markov Random Field: MRF) สามารถช่วยปรับปรุงประสิทธิภาพของกระบวนการจำแนกข้อมูลให้มีความถูกต้องมากยิ่งขึ้น

พหุน ปณ ฑิโต ชีเว

1.5 นิยามศัพท์ (Definitions)

1. การรวมตัวของผู้เล่น หมายถึง ผู้เล่นมากกว่า 2 คนขึ้นไปเคลื่อนที่เข้าหากันหรือยืนอยู่ใกล้กันในลักษณะของการจับกลุ่มซึ่งจะก่อให้เกิดปัญหาการบดบังระหว่างผู้เล่น

2. Field of View (FOV) คือ มุมมองหรือขอบเขตของพื้นที่บนสนามที่กล้องสามารถจับภาพได้

3. Long View คือ มุมมองภาพที่ได้จากการถ่ายภาพในระยะไกลซึ่งสามารถมองเห็นผู้เล่นในลักษณะเต็มตัวและสามารถมองเห็นขอบเขตของสนามฟุตบอลเป็นบริเวณกว้าง

4. การบดบัง หมายถึง เหตุการณ์ที่ในมุมมองของกล้องใดๆ มีผู้เล่นคนหนึ่งยืนอยู่ด้านหน้าของผู้เล่นอีกคนหนึ่งทำให้ส่วนของร่างกายของผู้เล่นที่ยืนอยู่ด้านหลังถูกบังทำให้กล้องนั้นไม่สามารถจับภาพร่างกายผู้เล่นคนดังกล่าวได้ครบถ้วน

5. การตรวจจับผู้เล่น หมายถึง กระบวนการค้นหาและสกัดบล็อบ (Blob) หรือกลุ่มของพื้นที่ที่คาดว่าจะเป็นส่วนหนึ่งของร่างกายของนักฟุตบอลซึ่งปรากฏอยู่ภายในภาพอินพุต (Input)

บทที่ 2
ปริทัศน์เอกสารข้อมูล (Literature Review)

วัตถุประสงค์ของการวิจัยดังที่เสนอในบทที่ 1 เกี่ยวข้องกับเรื่องของการตรวจจับและการติดตามวัตถุซึ่งในที่นี้ คือ ผู้เล่น ซึ่งกระบวนการต่างๆ เหล่านี้ต้องอาศัยทฤษฎีและความรู้พื้นฐานต่างๆ รวมถึงเทคนิคใหม่ๆ ที่นักวิจัยหลายท่านได้นำเสนอไว้ เพื่อให้การดำเนินการวิจัยในครั้งนี้สามารถบรรลุผลตามวัตถุประสงค์ที่ตั้งไว้ได้อย่างครบถ้วนซึ่งทฤษฎีและงานวิจัยที่เกี่ยวข้องกับงานวิจัยนี้มีรายละเอียดดังต่อไปนี้

2.1 การเรียนรู้ของเครื่องจักร (Machine Learning)

การเรียนรู้ของเครื่องจักร [38] เป็นสาขาหนึ่งของปัญญาประดิษฐ์ซึ่งเกี่ยวข้องกับการพัฒนาเทคนิคหรือวิธีการเพื่อให้คอมพิวเตอร์สามารถเรียนรู้โดยเน้นที่วิธีการเพื่อสร้างโปรแกรมคอมพิวเตอร์จากการวิเคราะห์ชุดข้อมูลการเรียนรู้ของเครื่องซึ่งเกี่ยวข้องกับงานทางด้านสถิติเนื่องจากทั้งสองสาขาศึกษาการวิเคราะห์ข้อมูลเช่นเดียวกัน สำหรับขั้นตอนวิธีการเรียนรู้ของเครื่องสามารถแบ่งได้เป็น 3 แบบ คือ

1. การเรียนรู้แบบมีผู้สอน (Supervised Learning)

การเรียนรู้แบบมีผู้สอน [38] เทคนิคการเรียนรู้ประเภทนี้ต้องใช้การเรียนรู้จากข้อมูลฝึกที่มีการใส่ลาเบล (Label) ให้กับข้อมูลฝึกไว้แล้วเพื่อให้คอมพิวเตอร์สามารถเข้าใจรูปแบบและได้สมมติฐานเพื่อทำงานกับข้อมูลในภายหน้าได้ ตัวอย่างเทคนิคประเภทนี้ เช่น การเรียนรู้แบบตัวจำแนกแบบเบย์อย่างง่ายและการเรียนรู้แบบต้นไม้ตัดสินใจ (Decision Tree) ขั้นตอนวิธีการเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbour: KNN) โครงข่ายประสาทเทียม (Neural Network) และซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines: SVM) เป็นต้น

ทฤษฎีความน่าจะเป็นของเบย์ (Bayesian Probability Theory) ตามที่อธิบายไว้ในงานวิจัยของ Olshausen [32] เป็นสมการทางคณิตศาสตร์ที่ถูกนำมาใช้ในการอนุมาน (Inference) หรือการให้เหตุผลโดยใช้พื้นฐานของความน่าจะเป็น ปัจจุบันถูกนำมาประยุกต์ใช้ในงานต่างๆ ทางด้านวิทยาศาสตร์ โดยทฤษฎีของเบย์นี้บ่อยครั้งถูกนำไปใช้ในการปรับความถูกต้องของสมมติฐานหรือใช้ในการปรับพารามิเตอร์ของโมเดล การให้เหตุผลของเบย์นั้นถูกนำไปใช้เพื่อทำการตัดสินใจและการสรุปทางสถิติที่เกี่ยวกับการอนุมานความน่าจะเป็น (Probability Inference) โดยทฤษฎีของเบย์นั้นจะใช้ความรู้จากเหตุการณ์ที่รู้อยู่ก่อนแล้วมาใช้เพื่อการทำนายเหตุการณ์ในอนาคต

สำหรับความน่าจะเป็นร่วมกันของสองเหตุการณ์ เช่น เหตุการณ์ A และเหตุการณ์ B สามารถเขียนเป็นสมการอธิบายความสัมพันธ์ของสองเหตุการณ์ในรูปกฎของเบย์ได้ดังสมการที่ 2.1

$$P(AB) = P(A|B)P(B) = P(B|A)P(A) \quad (2.1)$$

โดยที่ P คือฟังก์ชันของความน่าจะเป็น โดยที่ส่วนหนึ่งของเหตุการณ์นี้ คือ สมมติฐาน ซึ่งในที่นี้ คือ H และส่วนอื่นๆ คือ ข้อมูล ซึ่งในที่นี้ คือ D ในการตัดสินความจริงเชิงสัมพันธ์ของสมมติฐานโดยมีข้อมูลเบื้องต้นให้นั้นตามกฎของเบย์สามารถกระทำผ่านความสัมพันธ์ดังสมการที่ 2.2

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)} \quad (2.2)$$

โดยที่ $P(D|H)$ คือ ฟังก์ชันความเป็นไปได้ (Likelihood) และฟังก์ชันนี้จะประมวลผลความน่าจะเป็นของข้อมูลที่ได้จากการสังเกตที่เกิดจากสมมติฐาน โดยปกติมักจะได้มาจากการทดลอง ซึ่งก็คือ ค่าความน่าจะเป็นของข้อมูล D ที่ได้จากสมมติฐาน H ส่วน $P(H)$ เรียกว่า ไพออร์ (Prior) ซึ่งจะให้ค่าของความรู้ที่มีอยู่เดิมก่อนที่ข้อมูลจะถูกตัดสินใจซึ่งก็คือ ค่าความน่าจะเป็นเดิมของสมมติฐาน H ส่วน $P(D)$ เรียกว่า อีวิเดนซ์ (Evidence) คือ ค่าผลรวมของ $P(D|H)P(H)$ ทั้งหมดของสมมติฐาน H โดยมักจะเป็นค่าคงที่ที่มีการทำให้อยู่ในรูปทั่วไป (Normalizing) ซึ่งก็คือ ค่าความน่าจะเป็นเดิมของข้อมูล D และส่วนของ $P(H|D)$ เรียกว่า โพสเตอริเออร์ (Posterior) โดยเป็นค่าความน่าจะเป็นของสมมติฐานหลังจากทำการพิจารณาข้อมูลต่างๆ ที่เกี่ยวข้องแล้วซึ่งก็คือ ค่าความน่าจะเป็นของสมมติฐาน H ที่ได้จากข้อมูล D

2. การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning)

การเรียนรู้แบบไม่มีผู้สอน [38] ใช้ข้อมูลฝึกหรือชุดตัวอย่างที่ไม่มีการใส่ลาเบลให้กับข้อมูลและเรียนรู้โดยการนำข้อมูลไปผ่านกระบวนการหาความคล้ายคลึงของตัวอย่างจนกระทั่งได้กลุ่มตัวอย่างที่จัดเป็นประเภทอย่างเหมาะสม เทคนิคประเภทนี้ ได้แก่ เคมีน (K-Mean) การหาค่าคาดหมายสูงสุด (Expectation–Maximization: EM) ขั้นตอนวิธีการเพื่อนบ้านใกล้ที่สุด (KNN) และฟัซซีเคมีน (Fuzzy k-Mean) เป็นต้น

3. การเรียนรู้แบบกึ่งมีผู้สอน (Semi Supervised Learning)

การเรียนรู้แบบกึ่งมีผู้สอน [38] ใช้ข้อมูลฝึกที่มีการใส่ลาเบลเพียงบางส่วนของข้อมูลฝึกทั้งหมด การเรียนรู้ลักษณะนี้มีความคล้ายคลึงกับการเรียนรู้แบบมีผู้สอนแต่ใช้ข้อมูลที่ไม่มีลาเบลประกอบเพื่อช่วยให้เรียนรู้ได้ดีขึ้น สิ่งที่ได้จากการเรียนรู้แบบกึ่งมีผู้สอนจะเหมือนกับการเรียนรู้แบบมีผู้สอน คือ ได้แบบจำลอง

2.2 การประมวลผลภาพ (Image Processing)

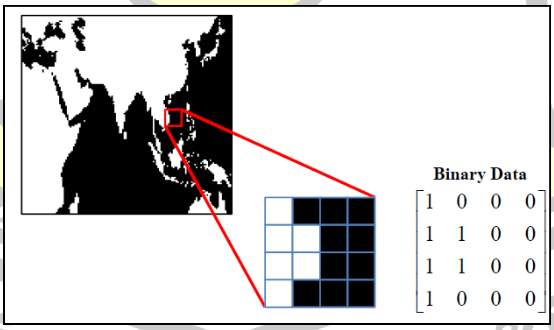
การประมวลผลภาพ คือ การประยุกต์ใช้งานการประมวลผลสัญญาณบนสัญญาณ 2 มิติ เช่น ภาพนิ่งหรือวิดีโอและยังรวมถึงสัญญาณ 2 มิติอื่น ๆ ที่ไม่ใช่ภาพด้วยโดยลักษณะพื้นฐานของภาพมีรายละเอียดดังนี้

1. เอกลักษณ์ของภาพ (Image Features)

โดยทั่วไปสามารถแบ่งรูปภาพที่ปรากฏและใช้งานบนเครื่องคอมพิวเตอร์ออกเป็น 2 ประเภท [36] คือ ภาพแบบบิตแมป (Bitmap Image) และภาพแบบเวกเตอร์ (Vector Image) โดยรูปภาพแบบบิตแมปจะพิจารณาตัวรูปภาพซึ่งถูกแบ่งออกเป็นส่วนย่อยๆ หรือที่เรียกว่า พิกเซล (Pixels) ที่ถูกนำมารวมกันและใช้แสดงผลภาพ ส่วนรูปภาพแบบเวกเตอร์จะประกอบด้วยเส้นลายต่างๆ ที่สร้างขึ้นจากการคำนวณทางคณิตศาสตร์ของลักษณะทางเรขาคณิตเพื่อสร้างรูปทรงต่างๆ เรียกว่า เวกเตอร์ (Vectors) สามารถแบ่งประเภทของภาพบิตแมปตามคุณสมบัติของสีออกได้เป็น 4 ประเภท ได้แก่ ภาพขาวดำ (Binary Image) ภาพระดับสีเทา (Gray Scale Image) ภาพสี (Color Image) และภาพแบบดัชนี (Index Image) โดยมีรายละเอียดดังนี้

1.1 ภาพขาวดำ (Binary Image)

ลักษณะของภาพขาวดำ [36] คือ ในแต่ละพิกเซลจะแสดงด้วยค่าแบบไบนารี (Binary) คือ มี 1 บิต ซึ่งประกอบไปด้วยค่า 1 และ 0 ดังภาพที่ 2.1 โดยที่ค่า 1 หมายถึงจุดภาพสีขาว และค่า 0 หมายถึงจุดภาพสีดำ ภาพประเภทนี้เหมาะสำหรับภาพที่เกี่ยวกับตัวอักษร (Text) ภาพลายนิ้วมือ เป็นต้น

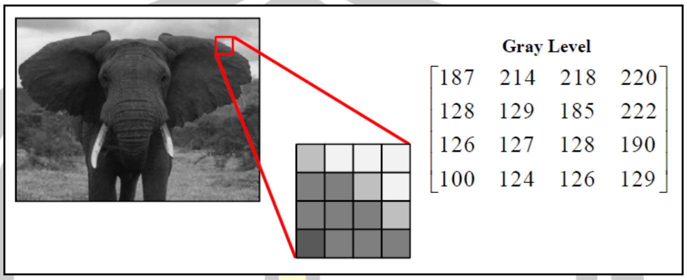


ภาพที่ 2.1 ตัวอย่างการแสดงภาพขาวดำ (ที่มา: [36])

1.2 ภาพระดับสีเทา (Gray Scale Image)

ภาพระดับสีเทา [36] ในแต่ละพิกเซลจะมีค่าความเข้มของแสงในแต่ละระดับที่แตกต่างกันไปตั้งแต่สีขาวไปจนถึงสีดำ สามารถกำหนดระดับความเข้มของแสงนั้นไดโดยใช้ค่าระดับ

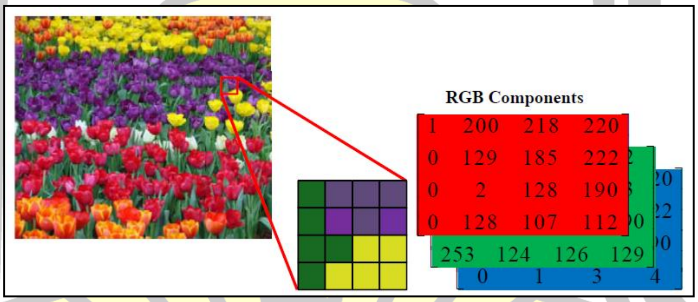
เทา ภาพแบบระดับสีเทาจะมีความละเอียด (Resolution) เท่ากับ 8 บิต ซึ่งภาพจะมีค่าระดับความเข้มแสงของสีดำเท่ากับ 0 ส่วนค่าระดับความเข้มของสีขาวจะมีค่าเท่ากับ 255 ดังภาพที่ 2.2



ภาพที่ 2.2 ตัวอย่างการแสดงภาพสีเทา (ที่มา: [36])

1.3 ภาพสี (Color Image)

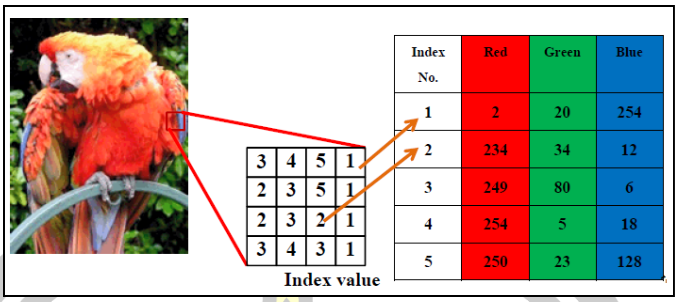
ในแต่ละพิกเซลของภาพสี [36] เก็บค่าระดับความเข้มแสงแต่ละแถบแสงของแม่สีหลัก 3 สีที่ซ้อนกัน คือ สีแดง (Red) สีเขียว (Green) และสีน้ำเงิน (Blue) ซึ่งในแต่ละพิกเซลนั้นๆ จะแสดงผลของค่าสีของแต่ละพิกเซลตามระดับความเข้มในแต่ละแถบแสงนั้นดังแสดงในภาพที่ 2.3



ภาพที่ 2.3 ตัวอย่างการแสดงภาพสี (ที่มา: [36])

1.4 ภาพแบบดัชนี (Index Image)

ภาพแบบดัชนี [36] ในแต่ละพิกเซลของภาพเก็บค่าดัชนี (Index Number) ซึ่งเป็นตัวเลขจำนวนเต็มซึ่งค่าดัชนีดังกล่าวใช้เทียบกับตารางสี (Color Table) ซึ่งเป็นตารางแสดงค่าแสงสีแดง เขียวและน้ำเงิน ซึ่งค่าดัชนีนี้จะเป็นตัวชี้ให้เห็นว่าภาพในตำแหน่งพิกเซลนั้นๆ มีค่าอัตราส่วนของแม่สีแสง 3 สีในอัตราส่วนละเท่าไร ดังแสดงในภาพที่ 2.4



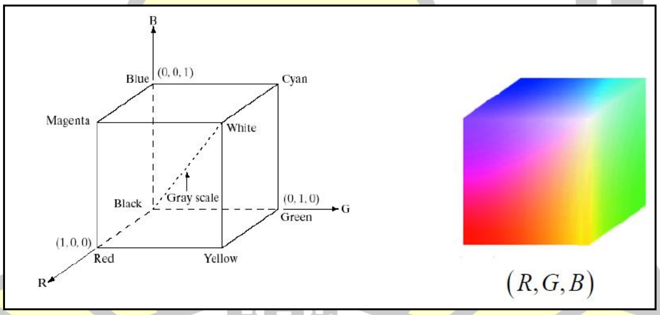
ภาพที่ 2.4 ตัวอย่างการแสดงภาพแบบดัชนี (ที่มา: [36])

2. เอกลักษณ์ของสี (Color Features)

แบบจำลองหรือโมเดลสี [36] เป็นการกำหนดรูปแบบของการแสดงสีที่เป็นมาตรฐานซึ่งโมเดลสีที่นิยมใช้ในระบบภาพดิจิตอล เช่น Red Green Blue (RGB), YCbCr, Hue Saturation Value (HSV) และ Hue Saturation Intensity (HSI) เป็นต้น ซึ่งมีรายละเอียดดังนี้

2.1 โมเดลสีแบบ Red Green Blue (RGB)

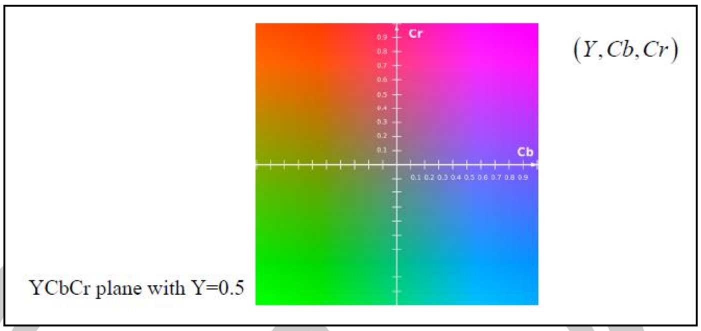
โมเดลสี RGB ประกอบด้วยแม่สีหลัก 3 สี คือ สีแดง เขียวและน้ำเงินซึ่งผลจากการผสมแม่สีหลักนี้จะได้สีอื่นๆ เพิ่มขึ้นโดยมีโครงสร้างของการผสมสีดังภาพที่ 2.5



ภาพที่ 2.5 โมเดลสีแบบ RGB (ที่มา: [36])

2.2 โมเดลสีแบบ YCbCr

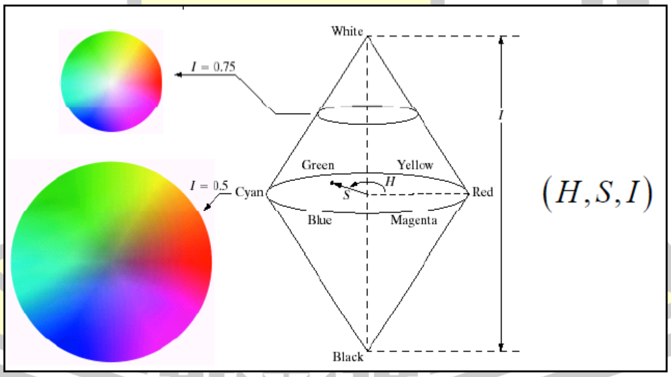
โมเดลสี YCbCr นิยมใช้ในสัญญาณวิดีโอดิจิตอลมีการประยุกต์ใช้ในการลดขนาดของภาพและขนาดของวิดีโอโดย Y เก็บข้อมูลความเข้มแสงหรือค่าความส่องสว่าง (Luminance) Cb และ Cr เก็บข้อมูลของค่าสีน้ำเงินและสีแดงที่ตัดความสว่างออกไป โดยมีโครงสร้างของโมเดลแสดงดังภาพที่ 2.6



ภาพที่ 2.6 โมเดลสีแบบ YCbCr (ที่มา: [36])

2.3 โมเดลสีแบบ Hue Saturation Intensity (HSI)

โมเดลสีแบบ HSI หรือเรียกได้ในอีกหลายชื่อ เช่น Hue Saturation Lightness (HSL), HLS, HIS และ HSB [112] เป็นโมเดลสีที่มีความใกล้เคียงกับโมเดลสีแบบ RGB ในด้านการรับรู้สี คุณสมบัติการรับรู้สีจะแบ่งเป็น ค่าความสว่าง (Brightness) ค่าโทนสี (Hue) และค่าความอิ่มตัว (Saturation) โดยค่าความสว่างเป็นการรับรู้ค่าความส่องสว่าง (Luminance) ค่าโทนสีของสีหมายถึงความเขียว เช่น เขียวแก่ ความแดง เช่น แดงอ่อน ส่วนค่าความอิ่มตัวจะเกี่ยวกับความแรงของแสงสีขาวโดยโครงสร้างของโมเดลสีแบบ HSI แสดงดังภาพที่ 2.7

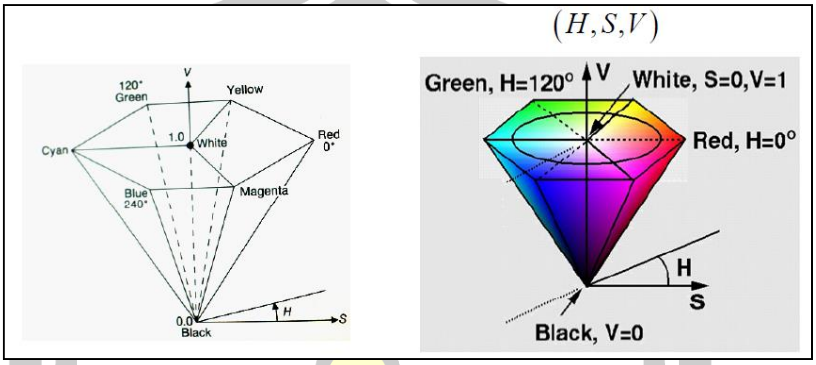


ภาพที่ 2.7 โมเดลสีแบบ HSI (ที่มา: [36])

2.4 โมเดลสีแบบ Hue Saturation Value (HSV)

โมเดลสีแบบ HSV สร้างมาจากลูกบาศก์สีของ RGB ตามแนวแกนเทา (Gray Axis) ซึ่งเป็นแกนที่เชื่อมระหว่างจุดมุมสีขาวและดำ โดยโทนสี H คือ มุมในรูปหกเหลี่ยมโดยแกนสีแดงมีมุม

เป็นศูนย์ องค์ประกอบ V จะวัดตามแนวดิ่ง ค่าความอิ่มตัวจะวัดจากระยะทางจากแกน V ซึ่งถือว่าระบบ HSV เป็นระบบในพิกัดทรงกระบอกซึ่งมีลักษณะดังภาพที่ 2.8



ภาพที่ 2.8 โมเดลสีแบบ HSV (ที่มา: [36])

3. ขั้นตอนการประมวลผลภาพ

ขั้นตอนการประมวลผลภาพประกอบด้วย 3 ขั้นตอนหลัก ได้แก่

1. ขั้นตอนการนำข้อมูลภาพเข้าสู่เครื่องคอมพิวเตอร์
2. ขั้นตอนการพัฒนาอัลกอริทึมเพื่อประมวลผลภาพ
3. ขั้นตอนการแสดงผลภาพ

ก่อนการประมวลผลภาพต้องจัดเตรียมข้อมูลภาพซึ่งมีขั้นตอนหลัก ดังนี้

1. ปรับปรุงข้อมูลภาพในส่วนที่ไม่คมชัด
2. กำจัดสัญญาณรบกวน (Noise)
3. การปรับคอนทราสต์ (Contrast) หรือปรับปรุงคุณสมบัติของการมองเห็น เช่น การปรับค่าความเข้มของจุดภาพและการแก้ไขตำแหน่งของจุดภาพ เป็นต้น
4. การแบ่งแยกข้อมูลภาพออกจากสีพื้นโดยใช้ค่าเทรชโฮลด์ (Threshold)
5. การแปลงข้อมูลภาพในทางเรขาคณิต เช่น การหมุนภาพ การเปลี่ยนแปลงขนาดของภาพ เป็นต้น
6. การแก้ไขข้อมูลภาพในส่วนที่มีการผิดเพี้ยน

2.3 ทฤษฎีคอมพิวเตอร์วิชัน (Computer Vision Theory)

คอมพิวเตอร์วิชันหรือคอมพิวเตอร์วิทัศน์ (Computer Vision) [39] เป็นสาขาหนึ่งของวิทยาการคอมพิวเตอร์ซึ่งเกี่ยวข้องกับการดึงสารสนเทศจากรูปภาพหรือวิดีโอ เครื่องมือที่ใช้ใน

คอมพิวเตอร์วิทัศน์ ได้แก่ คณิตศาสตร์ เช่น เรขาคณิต พีชคณิตเชิงเส้น สถิติและการวิเคราะห์เชิงฟังก์ชัน เป็นต้น โดยเครื่องมือเหล่านี้ถูกใช้ในการสร้างขั้นตอนวิธีหรือขั้นตอนวิธีในการแยกส่วนภาพและการจัดกลุ่มภาพเพื่อให้คอมพิวเตอร์สามารถเข้าใจทัศนียภาพหรือเอกลักษณ์ (Feature) ต่างๆ ในภาพซึ่งเป้าหมายโดยทั่วไปของคอมพิวเตอร์วิทัศน์ ได้แก่ การตรวจจับ การตัดแบ่งขอบเขต การระบุตำแหน่งและการรู้จำวัตถุที่ต้องการในภาพ การประเมินผลสำหรับการตัดแบ่งขอบเขตวัตถุในภาพ การติดตามวัตถุในภาพที่ต่อเนื่อง การประมาณท่าทางต่างๆ ของมนุษย์และส่วนต่างๆ ของร่างกายมนุษย์ การค้นหารูปภาพด้วยเนื้อหาของภาพในฐานข้อมูลภาพขนาดใหญ่ เป็นต้น เพื่อที่จะบรรลุเป้าหมายเหล่านี้ระบบคอมพิวเตอร์วิทัศน์จะต้องใช้กระบวนการต่างๆ เช่น การรู้จำแบบ การเรียนรู้เชิงสถิติ การประมวลผลภาพ ทฤษฎีกราฟและอื่นๆ ในการดำเนินการ ในที่นี้จะมุ่งเน้นในเรื่องของการตรวจจับวัตถุ การแสดงผลวัตถุและการติดตามวัตถุโดยมีรายละเอียดดังนี้

1. การตรวจจับวัตถุ (Object Detection)

การตรวจจับวัตถุมีหลายวิธีซึ่งอยู่บนพื้นฐานของเอกลักษณ์พื้นฐานต่างๆ เช่น สี (Color) พื้นผิว (Texture) การเคลื่อนไหว (Motion) และรูปร่าง (Shape) การค้นหาวัตถุเป็นขั้นตอนแรกที่ต้องดำเนินการเพื่อช่วยให้สามารถกำหนดตำแหน่งของวัตถุที่สนใจเพื่อที่จะสามารถดำเนินการในขั้นตอนอื่นๆ ได้สะดวกขึ้น เช่น ขั้นตอนการแสดงผล (Representation) และขั้นตอนการติดตามวัตถุ (Object Tracking) ในงานวิจัยของ Yilmaz และคณะ [7] ได้แบ่งวิธีการตรวจจับวัตถุเป็น 4 ประเภท คือ 1) ตัวตรวจจับจุด (Point Detector) เป็นลักษณะของการค้นหาจุดที่สนใจในภาพ 2) แบ็คกราวนด์ซับแทรคชัน (Background Subtraction) เป็นลักษณะของการเปรียบเทียบความคลาดเคลื่อนระหว่างภาพอินพุตกับโมเดลพื้นหลังของภาพที่กำหนด 3) การแบ่งส่วนเซกเมนต์ (Segmentation) เป็นลักษณะของการแบ่งภาพออกเป็นพื้นที่เพื่อแยกพื้นที่ที่สนใจ และ 4) การเรียนรู้แบบมีผู้สอน (Supervised Learning) เป็นวิธีการเรียนรู้จากชุดข้อมูลตัวอย่างโดยสร้างฟังก์ชันสำหรับการแมพ (Maps) ข้อมูลอินพุตกับข้อมูลเอาท์พุต (Output) เทคนิคในแต่ละประเภทสามารถอธิบายได้ดังนี้

1.1 โมเดลผสมแบบเกาเซียน (Gaussian Mixture Model: GMM)

GMM เป็นวิธีการตรวจจับวัตถุบนพื้นฐานของการเรียนรู้จากชุดข้อมูลสอน เป็นโมเดลทางสถิติที่ถูกนำมาประยุกต์ใช้ในการจำแนกหรือจัดกลุ่มของพิกเซลในภาพโดยโมเดลผสมแบบเกาเซียน G นี้เป็นการรวมกันของฟังก์ชันเกาเซียนหลายๆ ฟังก์ชัน $G_1, G_2, \dots, G_k$ เพื่อใช้ในการจัดการกับข้อมูลที่มีความซับซ้อนมากขึ้นซึ่งมีรูปของสมการดังสมการที่ 2.3 ถึงสมการที่ 2.5 ซึ่งอธิบายไว้ในงานวิจัยของ Liu และคณะ [28]

$$G = \sum_{i=1}^{k} \pi_i G_i \tag{2.3}$$

โดยที่

$$G_i(x,\theta_i) = \frac{1}{2\pi^{d/2}\left|\sum_i\right|^{1/2}} \exp^{-\frac{1}{2}}\left(x-\mu_i\right)^T\left(\sum_i\right)^{-1}\left(x-\mu_i\right) \quad (2.4)$$

และ

$$\sum_{i=1}^{k}\pi_i = 1 \quad (2.5)$$

โมเดลเกาเซียนนี้มีการกำหนดพารามิเตอร์ คือ $\boldsymbol{\theta}$ ซึ่งประกอบด้วยเวกเตอร์ของค่าเฉลี่ย (Mean Vector) μ_i และเมทริกซ์ความแปรปรวนร่วม (Covariance Matrix) $\sum_i$ ส่วนตัวแปร d คือ มิติ (Dimension) ของข้อมูลแซมเปิล (Sample Data) x คือ ข้อมูลที่สนใจ k คือ จำนวนเกาเซียน ส่วนเซตของ $\{\pi_i,\theta_i\}$ เป็นพารามิเตอร์ที่ไม่รู้ค่าซึ่งเป็นของสเปซพารามิเตอร์ (Parameter Space) โดยทั่วไปจะถูกประมาณค่าโดยใช้อัลกอริทึมการหาค่าคาดหวังสูงสุด

1.2 อัลกอริทึมโซเบลเกรเดียน

อัลกอริทึมโซเบลเกรเดียน (Sobel Gradient) [21] ถูกนำมาใช้สำหรับการตรวจจับการเปลี่ยนแปลงของวัตถุโดยคำนวณหาค่าอนุพันธ์ของความหนาแน่นซึ่งค่าอนุพันธ์ที่มากที่สุดจะเป็นจุดที่บอกถึงการเปลี่ยนแปลงนี้ สำหรับเกรเดียน (Gradient) คือ เวกเตอร์ โดยคอมโพเนนท์จะถูกวัดในทิศทางของ x และ y โดยใช้สมการที่ 2.6 และสมการที่ 2.7 ดังนี้

$$\frac{\partial f(x,y)}{\partial x} = \Delta x = \frac{f(x+dx,y)-f(x,y)}{dx} \quad (2.6)$$

$$\frac{\partial f(x,y)}{\partial y} = \Delta y = \frac{f(x,y+dy)-f(x,y)}{dy} \quad (2.7)$$

โดยที่ f คือ ฟังก์ชันความหนาแน่น ในการตรวจจับการมีอยู่ของเกรเดียนที่ไม่ต่อเนื่องค่าการเปลี่ยนแปลงในเกรเดียนที่ (x, y) จะถูกคำนวณ ส่วนค่าขนาด (Magnitude) (M) และทิศทางของเกรเดียน (Gradient Direction) $(\vec{\theta})$ จะถูกคำนวณโดยใช้สมการที่ 2.8 และสมการที่ 2.9 ดังนี้

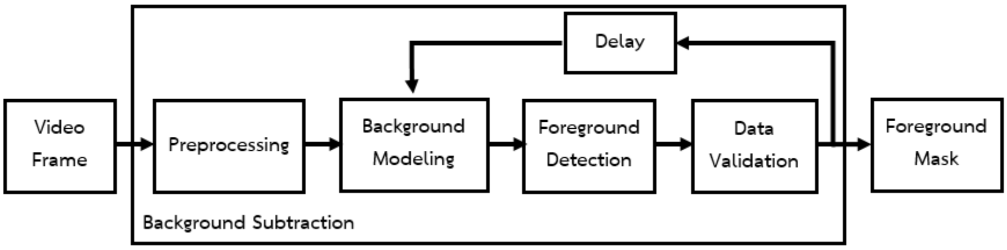
$$M = \sqrt{\Delta x^2 + \Delta y^2} \quad (2.8)$$

$$\vec{\theta} = \tan^{-1}\left[\frac{\Delta y}{\Delta x}\right] \quad (2.9)$$

พหูนำ ปณ ศิต ชีเว

1.3 การลบแบ็คกราวนด์หรือแบ็คกราวนด์ซับแทรคชัน

แบ็คกราวนด์ซับแทรคชัน [61] ใช้พื้นฐานของข้อมูลในเฟรมก่อนหน้าในการตรวจจับวัตถุซึ่งอัลกอริทึมการลบแบ็คกราวนด์โดยทั่วไปมี 4 ขั้นตอนดังที่อธิบายไว้ในงานของ Cheung และ Kamath [61] แสดงดังภาพที่ 2.9



ภาพที่ 2.9 ขั้นตอนการลบแบ็คกราวนด์ (ที่มา: [61])

1) Preprocessing เป็นขั้นตอนการเตรียมข้อมูลก่อนการประมวลผลเพื่อช่วยลดสิ่งรบกวนและเพิ่มความรวดเร็วและความถูกต้องในการประมวลผล

2) Background Modeling เป็นขั้นตอนการโมเดลแบ็คกราวนด์ (Background) เริ่มต้นเพื่อใช้ในการเปรียบเทียบกับภาพที่ต้องการตรวจจับวัตถุโดยใช้ข้อมูลจากเฟรมวิดีโอช่วงสั้นๆ ในการสร้างโมเดล

3) Foreground Detection เป็นขั้นตอนการตรวจจับวัตถุโดยใช้โมเดลแบ็คกราวด์ที่ได้ทำการเปรียบเทียบกับภาพอินพุตแล้วระบุพิกเซลโฟร์กราวนด์ (Foreground) จากภาพอินพุตนั้นโดยพิจารณาจากความแตกต่างระหว่างพิกเซล

4) Data Validation เป็นขั้นตอนของการปรับปรุงโฟร์กราวนด์มาร์ค (Foreground Mask) ที่ได้โดยอาศัยข้อมูลที่ได้จากภายนอกโมเดลแบ็คกราวนด์ซึ่งโฟร์กราวนด์มาร์คที่ได้อาจไม่ถูกต้องสมบูรณ์ เช่น การปรากฏพิกเซลที่แยกอยู่โดยไม่รวมกับพื้นที่ใดๆ พื้นที่ของวัตถุเดียวกันแต่ปรากฏแยกกัน เป็นต้น วิธีที่ใช้ในการปรับปรุงโฟร์กราวนด์ เช่น การกรองในลักษณะทางสัณฐานวิทยา (Morphological Filtering)

1.4 การแบ่งส่วนเซกเมนต์ภาพ

การแบ่งส่วนเซกเมนต์ภาพ [64] เป็นกระบวนการของการแบ่งภาพออกเป็นหลายเซกเมนต์ จุดประสงค์ของการเซกเมนต์ คือ การแบ่งภาพออกเป็นส่วนๆ เพื่อการวิเคราะห์ข้อมูลและเพื่อการเปลี่ยนแปลงการแสดงผลภาพไปเป็นรูปแบบอื่นที่มีความหมายและง่ายต่อการวิเคราะห์ โดยทั่วไปมักใช้เพื่อการกำหนดตำแหน่งวัตถุและขอบเขต (Boundary) เช่น เส้นตรงและเส้นโค้ง

ภายในภาพ อีกนัยหนึ่งการเซกเมนต์ภาพเป็นกระบวนการกำหนดลาเบลให้กับแต่ละพิกเซลที่มีเอกลักษณ์ที่คล้ายกันภายในภาพ ผลจากการเซกเมนต์ คือ เซตของเซกเมนต์หรือเซตของคอนทัวร์ (Contour) โดยทั่วไปมี 3 วิธี ได้แก่ การเซกเมนต์โดยอาศัยการกำหนดค่าเทรชโฮลด์ การเซกเมนต์บนพื้นฐานของขอบ (Edge-Based) และการเซกเมนต์บนพื้นฐานของพื้นที่ (Region-Based) เป็นต้น

1.5 การตรวจจับจุดที่สนใจ

การตรวจจับจุดที่สนใจ Ahsan และ Mohamad [67] ได้อธิบายไว้ว่า เป็นการค้นหาจุดที่แสดงถึงการเปลี่ยนแปลงของความหนาแน่นทั้งในแนวนอนและแนวตั้งซึ่งจุดเหล่านี้มักถูกเรียกว่าจุดที่สำคัญ (Keypoints) หรือจุดที่สนใจ (Interest Points) จุดนี้จะมีความคงที่ต่อการเปลี่ยนแปลงในเรื่องของการเปลี่ยนแปลงและการส่องแสง (Transformation and Illumination) ตัวตรวจจับที่มีพื้นฐานบนจุดที่สนใจ เช่น ตัวตรวจจับจุดที่สนใจแบบแฮร์ริส (Harris Interest Point Detector) การแปลงเอกลักษณ์แบบสเกลคงที่หรือเรียกว่า ซิฟท์ (Scale Invariant Feature Transform (SIFT)) เอกลักษณ์ที่แข็งแรงที่เร็วขึ้นหรือเรียกว่า เซิร์ฟ (Speed-Up Robust Features (SURF)) โดยทั่วไปซิฟท์และเซิร์ฟจะคงทนต่อการเปลี่ยนแปลงของแสง การหมุนและสเกล ส่วนแฮร์ริสจะไม่คงทนต่อสเกลแต่จะทำงานได้เร็วกว่าสองวิธีแรกขณะเดียวกันจะมีความถูกต้องน้อยกว่าสองวิธีแรกและได้สรุปผลการเปรียบเทียบระหว่างซิฟท์และเซิร์ฟในประเด็นของการหมุน การเปลี่ยนแปลงสเกล สิ่งรบกวนภาพ การเปลี่ยนแปลงของแสงและการเปลี่ยนแปลงด้านมุมมองจากงานวิจัยอ้างอิงว่า ซิฟท์ทำงานได้ดีกว่าเซิร์ฟเล็กน้อยแต่ช้ากว่าและมีการคำนวณที่ซับซ้อนมากกว่า

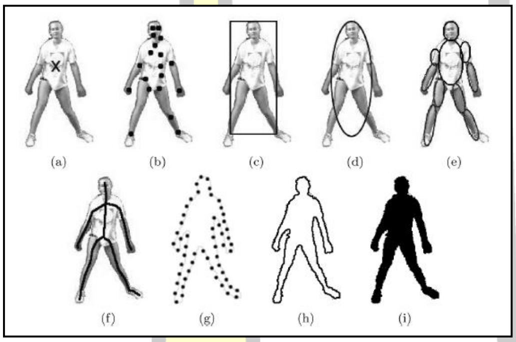
1.6 ฮิสโตแกรมแบ็คโปรเจ็คชัน

วิธีฮิสโตแกรมแบ็คโปรเจ็คชัน (Histogram Backprojection Method) [74, 60] เป็นวิธีการที่ประยุกต์ใช้สำหรับค้นหาวัตถุในภาพหรือการแบ่งส่วนเซกเมนต์ภาพโดยอาศัยข้อมูลฮิสโตแกรม (Histogram) ของภาพ การทำแบ็คโปรเจ็คชัน (Backprojection) จะทำการแทนที่ทุกๆ พิกเซลในภาพด้วยความน่าจะเป็นที่เกิดขึ้นในภาพของตัวมันเองโดยมีกระบวนการ คือ การคำนวณฮิสโตแกรมสี (Color Histogram) ของวัตถุที่ต้องการค้นหาและภาพที่ใช้สำหรับค้นหา จากนั้นคำนวณค่าเรโช (Ratio) แล้วใช้ค่าเรโชนี้เป็นจานสี (Palette) และสร้างภาพใหม่ด้วยทุกๆ พิกเซลที่ตรงกับความน่าจะเป็นของเป้าหมายแล้วหาค่าที่น้อยสุดในแต่ละพิกเซลของภาพใหม่ที่ได้ ขั้นตอนถัดไป คือ การทำคอนโวลูชัน (Convolution) แบบเซอร์คูล่าดิส (Circular Disc) ซึ่งตำแหน่งที่มีค่าความหนาแน่นมากที่สุดจะเป็นตำแหน่งของวัตถุที่ต้องการค้นหา

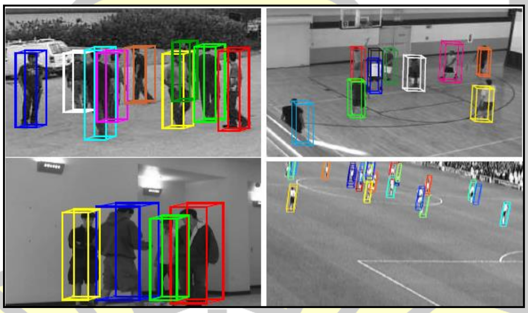
2. การแสดงผลวัตถุ (Object Representation)

การแสดงผลวัตถุที่ได้จากกระบวนการตรวจจับโดยส่วนใหญ่สามารถแสดงผลวัตถุนั้นได้ 2 ลักษณะ คือ แบบรูปร่าง (Shape) และแบบลักษณะการปรากฏ (Appearances) ของวัตถุดังที่กล่าวไว้ในงานวิจัยของ Yilmaz และคณะ [7] วิธีที่นิยมใช้ในการแสดงผลแบบรูปร่าง เช่น จุด โครง

ร่างของจุด กรอบสี่เหลี่ยม วงรีและแบบคอนทัวร์ ดังในงานวิจัยของ Yilmaz และคณะ [7] งานวิจัยของ Jalal และ Singh [17] ดังภาพที่ 2.10 เป็นต้น และในปัจจุบันมีวิธีการแสดงผลแบบกล่องสี่เหลี่ยมสามมิติดังที่เสนอในงานวิจัยของ Khan [19] ดังภาพที่ 2.11 เป็นต้น แต่ละวิธีมีทั้งข้อดีข้อเสียแตกต่างกัน เช่น ในด้านความเร็วในการทำงานและพื้นที่ในการจัดเก็บข้อมูล ส่วนวิธีการแสดงผลแบบลักษณะการปรากฏ ได้อธิบายไว้ในงานวิจัยของ Yilmaz และคณะ [7] ซึ่งโมเดลการแสดงผลเหล่านี้สามารถนำไปใช้สำหรับการติดตามวัตถุได้อีกด้วย



ภาพที่ 2.10 การแสดงผลวัตถุ (Object Representations) (ที่มา: Yilmaz และคณะ [7])



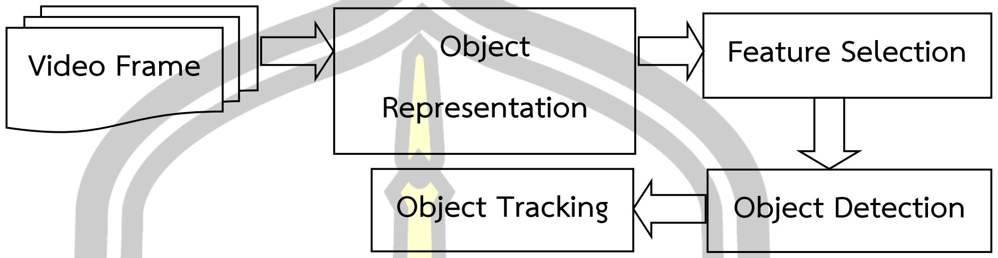
ภาพที่ 2.11 การแสดงผลวัตถุแบบ 3 มิติ (Object Representations in 3D) (ที่มา: Khan [19])

3. การติดตามวัตถุ (Object Tracking)

จุดประสงค์ของการติดตามวัตถุ (Object Tracking) [7] คือ การจัดเก็บข้อมูลของการเคลื่อนที่ของวัตถุที่สนใจโดยการกำหนดตำแหน่งของวัตถุในทุกเฟรมวิดีโอซึ่งข้อมูลที่ได้นั้นจะถูกนำไปประยุกต์ใช้ในงานด้านต่างๆ เช่น การทำนายการเคลื่อนที่ของวัตถุจนถึงการประยุกต์ใช้กับการวิเคราะห์พฤติกรรมของวัตถุ เช่น ระบบการตรวจสอบที่ชาญฉลาด เป็นต้น สำหรับขั้นตอนพื้นฐาน

พหุน ปณ ศิต ชีเว

ของกระบวนการติดตามวัตถุในวิดีโอดังที่อธิบายไว้ในงานวิจัยของ Dash และคณะ [16] ประกอบด้วย 4 ขั้นตอนหลัก ดังภาพที่ 2.12



ภาพที่ 2.12 ขั้นตอนการติดตามวัตถุในวิดีโอ (ที่มา: Dash และคณะ [16])

ขั้นตอนของการติดตามวัตถุนั้นเริ่มจากการนำวิดีโอมาแปลงเป็นชุดของเฟรมภาพ จากนั้นทำการแสดงผลวัตถุที่สนใจ หลังจากกำหนดวิธีการแสดงผลวัตถุแล้วจะเป็นขั้นตอนการเลือกเอกลักษณ์ซึ่งเอกลักษณ์นี้ขึ้นอยู่กับการเลือกใช้งานของนักวิจัยแต่ละท่านจากนั้นจึงเป็นขั้นตอนของการตรวจจับวัตถุ ส่วนขั้นตอนสุดท้าย เป็นขั้นตอนของการติดตามวัตถุที่สนใจซึ่งเทคนิคของการติดตามวัตถุมีนักวิจัยหลายท่านได้นำเสนอไว้ เช่น ในงานวิจัยของ Beyan และ Temizel [2] ได้เสนอวิธีการติดตามวัตถุโดยใช้อัลกอริทึมมีนชีฟท์ (Mean-Shift Algorithm) งานวิจัยของ Comaniciu และคณะ [5] ได้เสนอวิธีการติดตามวัตถุบนพื้นฐานของเคอร์เนล (Kernel-Based Object Tracking) งานวิจัยของ Huang และคณะ [9] เสนอวิธีการติดตามวัตถุโดยใช้ความสัมพันธ์ของผลตอบสนองของการตรวจจับวัตถุ โดยงานวิจัยบางกลุ่มจะใช้เทคนิคพื้นฐานที่มีอยู่แล้วและในงานวิจัยบางกลุ่มได้ทำการปรับปรุงเทคนิคขึ้นใหม่เพื่อให้ผลการติดตามวัตถุนั้นมีประสิทธิภาพมากขึ้น สำหรับตัวอย่างของอัลกอริทึมที่ใช้ในการทำติดตามวัตถุ เช่น การติดตามวัตถุบนพื้นฐานของเคอร์เนล (Kernel-Based Object Tracking) โดยใช้ฮิสโตแกรมของสี (Color Histogram) และการติดตามวัตถุโดยใช้ตัวติดตามความแปรปรวน (Covariance Tracker) เช่น ในงานวิจัยของ Dash และคณะ [16] เป็นต้น

อัลกอริทึมสำหรับการติดตามวัตถุในวิดีโอที่ดีควรจะมีลักษณะดังที่ได้กล่าวไว้ในงานวิจัยของ Jalal และ Singh [17] ซึ่งมีรายละเอียดดังนี้

1) อัลกอริทึมนั้นควรจะสามารถตรวจจับวัตถุทั้งหมดที่เข้าและออกจากฉาก

2) อัลกอริทึมนั้นควรแยกความแตกต่างของแต่ละวัตถุที่อยู่ในฉากเวลาเดียวกันได้

3) เมื่อมีการติดตามและดึงร่องรอยของวัตถุทั้งหมด ลาเบล (Label) ที่กำหนดให้กับวัตถุนั้นควรที่จะยังคงเหมือนเดิม

4) กรณีการเคลื่อนไหวหรือขาดการเคลื่อนไหวของวัตถุจะต้องไม่เกิดการเปลี่ยนแปลงลาเบลที่กำหนดให้แต่ละวัตถุนั้น

5) อัลกอริทึมนั้นควรจัดการกับปัญหากรณีมีการบดบังและการปรากฏขึ้นมาใหม่โดยไม่มีการเปลี่ยนแปลงลาเบลที่กำหนดไว้

การจัดการกับปัญหาของการติดตามโดยทั่วไปมี 2 วิธี ดังที่กล่าวไว้ในงานวิจัยของ Jalal และ Singh [17] คือ แบบบนลงล่าง (Top Down) หรือการติดตามแบบไปข้างหน้า (Forward Tracking) และแบบล่างขึ้นบน (Bottom-Up) หรือการติดตามแบบถอยหลัง (Back-Tracking) ส่วนวิธีในการติดตามวัตถุดังที่อธิบายไว้ในงานวิจัยของ Bunyak [12] นั้นมีอยู่หลายวิธี เช่น การติดตามบนพื้นฐานของพื้นที่ (Region-Based Tracking Methods) การติดตามบนพื้นฐานของคอนทัวร์ (Contour-Based Tracking Methods) การติดตามบนพื้นฐานของเอกลักษณ์ (Feature-Based Tracking Methods) การติดตามบนพื้นฐานของโมเดล (Model-Based Tracking Methods) การตรวจสอบแบบหลายมุมมอง (Multi-View Monitoring) การติดตามแบบหลายคิว (Multi-cue Tracking) และการผสมข้อมูล (Data Fusion) เป็นต้น

สำหรับการเลือกเอกลักษณ์เพื่อใช้ในการติดตามวัตถุ Yilmaz และคณะ [7] ได้อธิบายไว้ว่าในหลายงานวิจัยนิยมใช้หลายเอกลักษณ์ร่วมกันซึ่งโดยทั่วไปเอกลักษณ์ที่นิยมใช้ เช่น สี ขอบ (Edge) ออปติคอลโฟลว์ (Optical Flow) และพื้นผิว (Texture) เป็นต้น

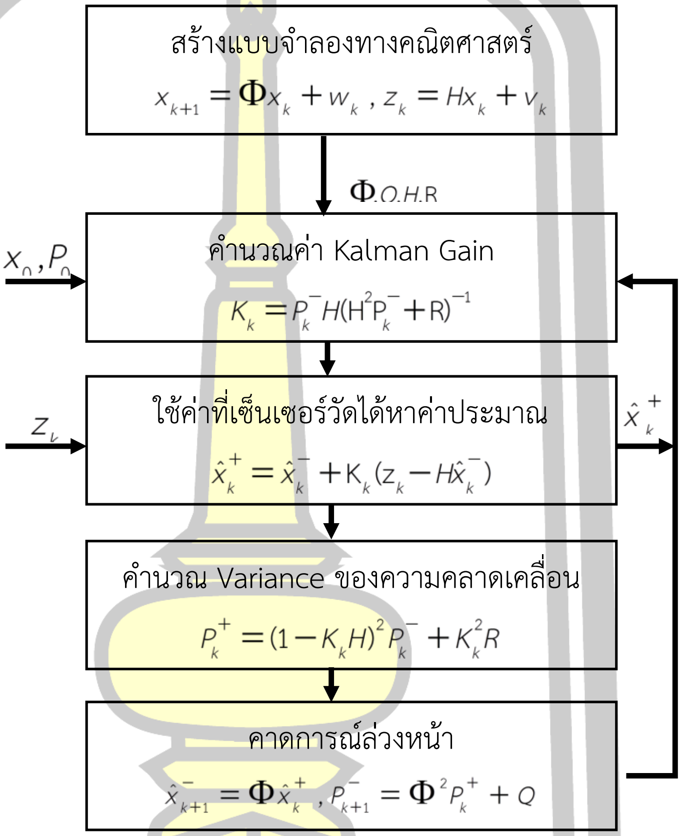
ในที่นี้ได้แบ่งกระบวนการติดตามวัตถุออกเป็น 2 กระบวนการหลัก คือ การติดตามผู้เล่นในมุมมองภาพ (Single/Image View) และการติดตามผู้เล่นในมุมมองเสมือนหรือมุมมองพื้น (Virtual/Ground View) โดยมีรายละเอียดดังนี้

3.1 การติดตามผู้เล่นในแบบมุมมองภาพ (Single/Image-View Player Tracking)

วิธีการที่ใช้ในการติดตามวัตถุในแต่ละมุมมองภาพ (2D) มีหลายวิธี เช่น ตัวกรองคาร์ลมาล การติดตามบนพื้นฐานของโมเดลการกระจายตัวของจุด (Point Distribution Models (PDM)) การใช้ตัวกรองอนุภาค (Particle Filter) การติดตามบนพื้นฐานของโมเดลกราฟ (Graph Model) และการใช้วิธีอัลกอริทึมคอนเดนเซชัน (Condensation Algorithm) เป็นต้น ซึ่งมีรายละเอียดดังนี้

1) ตัวกรองคาร์ลมาล งานของกฤษฎา แสงเพ็ชร์ส่อง [66] ได้อธิบายถึงตัวกรองคาร์ลมาลว่าเป็นวิธีการเพื่อประมาณสถานะของระบบมีลักษณะแบบวนซ้ำ ขั้นตอนการประมวลผลของตัวกรองคาร์ลมาล มีอยู่ 5 ขั้นตอน ดังภาพที่ 2.13 ในขั้นตอนแรก คือ การสร้างโมเดลของระบบประกอบด้วย โมเดลโปรเซส (Process Model) และโมเดลการวัด (Measurement Model) ขั้นตอนที่สอง หาค่าประมาณเริ่มต้นของสถานะระบบ (x_0) และความแปรปรวนของความคลาดเคลื่อน (P_0) จากนั้น คาดการณ์สถานะของระบบและคำนวณความแปรปรวนของความคลาดเคลื่อนของการคาดการณ์ สุดท้าย คำนวณค่า คาร์ลมาลเกน (Kalman Gain (K_1)) ที่ดีที่สุดจากข้อมูลสถานะ ความแปรปรวน ขั้นตอนที่สาม หาค่าประมาณจากค่าเกน (Gain) และค่าที่เซนเซอร์วัดได้ (z_k) เพื่อแก้ไข

ค่าที่คาดการณ์ไว้ล่วงหน้า ขั้นตอนที่สี่ คำนวณค่าความแปรปรวนของความคลาดเคลื่อนของการประมาณและขั้นตอนสุดท้าย คาดการณ์สถานะของระบบล่วงหน้าและคำนวณความแปรปรวนของความคลาดเคลื่อนของการคาดการณ์ กระบวนการของตัวกรองคาร์ลมาลจะกระทำซ้ำในขั้นตอนที่ 2 ถึงขั้นตอนที่ 5 ไปเรื่อยๆ ตามลำดับของเวลา



ภาพที่ 2.13 ขั้นตอนการประมวลผลของตัวกรองคาร์ลมาล (ที่มา:กฤษฎา แสงเพ็ชร์ส่อง [66])

2) โมเดลการกระจายตัวของจุด [69] เป็นเทคนิคที่ถูกนำมาใช้สำหรับการอธิบายรูปร่างของวัตถุด้วยเซตของจุดโดยอาศัยการเรียนรู้จากชุดข้อมูลสอนหรือชุดของรูปร่าง อินพุตของโมเดลการกระจายตัวของจุด คือ เซตของจุดแลนมาร์ค (Landmarks Points) ซึ่งเป็นตัวแทนของชุดสอน การสร้างโมเดลมีขั้นตอนหลัก คือ ขั้นตอนแรก ชุดข้อมูลสอนจะถูกจัดแนว (Aligned) ไปเป็นเฟรมของพิกัดมาตรฐาน (รูปร่างเฉลี่ย) โดยใช้อัลกอริทึมการวิเคราะห์โปรครัสต์ (Procrustes Analysis Algorithm) ขั้นตอนที่สอง ใช้หลักการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA) เพื่อวิเคราะห์พิกัดของจุดแลนมาร์คซึ่งจะได้ตำแหน่งเฉลี่ยของแต่ละแลนมาร์คและคำอธิบายของเส้นทางหลักของแต่ละจุดเหล่านี้ที่จะเคลื่อนที่ไปด้วยกัน ข้อมูลทางสถิติ

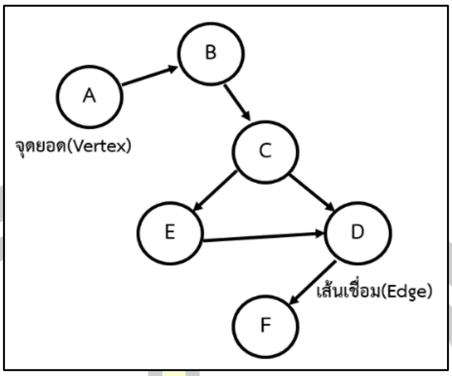
ที่ได้เกี่ยวกับค่าเฉลี่ยและความแปรปรวนของค่าเกรย์สเกล (Grayscale) ของพิกเซลบริเวณรอบแต่ละจุดแลนมาร์คจะถูกนำไปใช้ในวิธีของโมเดลการกระจายตัวของจุดต่างๆ เช่น การจับคู่ (Match) ระหว่างจุดแลนมาร์คในโมเดล (Active Shape Model : ASM) และการสร้างโมเดลลักษณะการปรากฏในโมเดล (Active Appearance Model: AAM) เป็นต้น โมเดลของจุดที่ได้นี้จะถูกใช้เป็นเอกลักษณ์สำหรับกระบวนการติดตามต่อไป

3) การวิเคราะห์โปรครัสต์ [80, 81] คือ การวิเคราะห์รูปร่างทางสถิติซึ่งถูกใช้เพื่อวิเคราะห์การกระจายตัวของเซตของรูปร่าง เป็นวิธีการจัดรูปแบบของเซตของจุดให้อยู่ในรูปร่างมาตรฐานซึ่งการวิเคราะห์โปรครัสต์แบบทั่วไป (Generalized Procrustes Analysis) มี 4 ขั้นตอนดังนี้ ขั้นตอนแรก เลือกการประมาณเริ่มต้นของรูปร่างเฉลี่ย เช่น รูปร่างแรกในเซต ขั้นตอนที่สอง จัดตำแหน่งรูปร่างที่เหลือให้เข้ากับรูปร่างเฉลี่ยที่กำหนด ขั้นตอนที่สาม คำนวณการประมาณใหม่ในส่วนของค่าเฉลี่ยจากรูปร่างที่ทำการจัดแนวตำแหน่งนั้น และขั้นตอนที่สี่ ถ้าค่าเฉลี่ยที่ประมาณมีการเปลี่ยนแปลงให้ดำเนินการในขั้นตอนที่สองใหม่

4) ตัวกรองอนุภาค คือ ตัวกรองแบบวนซ้ำโดยใช้พื้นฐานของวิธี มอนติคาร์โล (Monte Carlo) [68] ซึ่งจะแสดงผลการประมาณเป็นลักษณะของเซตของกลุ่มตัวอย่างที่สุ่มของฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability Density Function : PDF) ซึ่งโดยทั่วไปประกอบด้วย 2 โมเดลทางคณิตศาสตร์ คือ โมเดลระบบ (State Dynamics Equation) และโมเดลการวัด (Measurement Equation) กระบวนการทำงานประกอบด้วย 2 ขั้นตอนหลัก คือ ขั้นตอนการทำนายกลุ่มตัวอย่างก่อนหน้า (Prior) จากกลุ่มตัวอย่างเดิมและขั้นตอนการอัพเดทค่าน้ำหนักของกลุ่มตัวอย่างก่อนหน้า

5) โมเดลกราฟ [55] เป็นโมเดลข้อมูลที่ถูกนำมาใช้แสดงความสัมพันธ์ของข้อมูลในลักษณะต่างๆ ซึ่งประกอบด้วย 2 ส่วน คือ เซตของโหนด (Vertices) และเซตของเส้นเชื่อม (Edges) หรือเส้นกราฟโดยโหนดจะแสดงถึงข้อมูลต่างๆ ส่วนเส้นกราฟจะแสดงความสัมพันธ์ระหว่างข้อมูลนั้นๆ โดยทั่วไปจะแสดงกราฟด้วยรูปภาพโดยให้จุดแทนจุดยอดและเส้นแทนเส้นเชื่อมดังภาพที่ 2.14 ในงานที่เกี่ยวข้องกับการติดตามวัตถุได้ประยุกต์ใช้กราฟเพื่อแสดงความสัมพันธ์ของข้อมูลในลักษณะของพื้นที่และเวลาโดยในแต่ละช่วงเวลา แต่ละข้อมูลหรือวัตถุจะถูกกำหนดด้วยโหนดของกราฟซึ่งจะมีข้อมูลพื้นฐานต่างๆ เช่น ตำแหน่งและเอกลักษณ์อื่นๆ ส่วนระยะทางระหว่างข้อมูลหรือโหนดจะถูกกำหนดด้วยเส้นของกราฟซึ่งข้อมูลจากกราฟจะถูกใช้ในการติดตามวัตถุ โมเดลการแสดงผลนี้สามารถช่วยจัดการปัญหาของความสอดคล้องของวัตถุซึ่งช่วยในการแยกบล็อบและการติดตามในลำดับวิดีโอได้ การติดตามจะกระทำโดยการค้นหาเส้นทางที่ดีที่สุดในกราฟ

พหุน ปณฺ ตา โต ชีวา



ภาพที่ 2.14 โครงสร้างกราฟ

การจำแนกชนิดของกราฟ [55] สามารถจำแนกได้ 2 ลักษณะ คือ การจำแนกตามลักษณะของข้อมูลที่เก็บและการจำแนกตามการเชื่อมโยง การจำแนกตามลักษณะของข้อมูลที่เก็บสามารถแบ่งออกเป็น 2 ประเภทได้แก่ 1) กราฟแบบมีทิศทาง (Directed Graph) และกราฟแบบไม่มีทิศทาง (Undirected Graph) 2) กราฟแบบมีน้ำหนัก (Weighted Graph) และกราฟแบบไม่มีน้ำหนัก (Unweighted Graph) ส่วนการจำแนกตามการเชื่อมโยงสามารถแบ่งออกเป็น 4 ประเภทได้แก่ 1) กราฟสมบูรณ์ (Complete Graph) 2) กราฟต่อเนื่อง (Connected Graph) 3) กราฟไม่ต่อเนื่อง (Unconnected Graph) และ 4) ต้นไม้ (Tree) เป็นต้น

6) อัลกอริทึมคอนเดนเซชัน [65] เป็นอัลกอริทึมที่ใช้อธิบายรูปร่างของวัตถุที่อธิบายด้วยเวกเตอร์ มีลักษณะการทำงานบนพื้นฐานของการสุ่มเชิงปัจจัย (Factor Sampling) เอาท์พุตของอัลกอริทึมจะเป็นความหนาแน่นของความน่าจะเป็นเชิงเงื่อนไข (Conditional Probability Density) ซึ่งจะแสดงความน่าจะเป็นของรูปร่างที่เป็นไปได้ของวัตถุบนพื้นฐานของรูปร่างและการวัดก่อนหน้า อัลกอริทึมนี้ถูกปรับให้ทำงานแบบวนซ้ำเพื่อดำเนินการกับรายการภาพในวิดีโอได้ กระบวนการทำงานในแต่ละช่วงเวลาจะประกอบด้วย N รอบเพื่อสร้างเซตของกลุ่มตัวอย่างใหม่และแต่ละรอบจะมี 3 ขั้นตอน คือ ขั้นตอนแรก การเลือกกลุ่มตัวอย่างใหม่จากกลุ่มตัวอย่างเดิมโดยใช้ค่าน้ำหนักของความน่าจะเป็นในการพิจารณา ขั้นตอนที่สอง การทำนายโดยการสุ่มตัวอย่างจากความหนาแน่นเชิงเงื่อนไขเพื่อสร้างเซตของกลุ่มตัวอย่างใหม่ และขั้นตอนที่สาม ค่าการวัด (Observation Measurement) ที่ได้ โดยทำการกำหนดค่าน้ำหนักของตำแหน่งใหม่สำหรับกลุ่มตัวอย่างใหม่จากเอกลักษณ์ที่วัดได้โดยแต่ละค่าน้ำหนักจะประมาณจากฟังก์ชันความหนาแน่น เอาท์พุตของอัลกอริทึมสามารถใช้คำนวณตำแหน่งเฉลี่ยของวัตถุที่ถูกติดตามได้โดยตรง รวมถึงงานด้านอื่นๆ ของวัตถุที่ถูกติดตาม ข้อดีของอัลกอริทึมนี้ คือ สามารถติดตามวัตถุทั้งแบบมีรูปร่างตายตัว (Rigid) และไม่ตายตัว (Non-Rigid) และทำงานได้ดีกว่าตัวกรองคาร์ลมาล

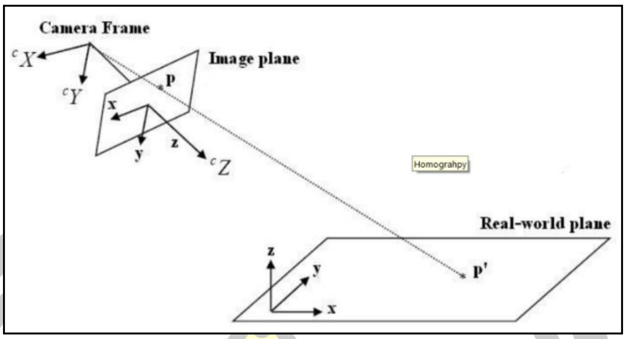
7) การสุ่มเชิงปัจจัย [65] เป็นวิธีการสุ่มเลือกตัวอย่างจากกลุ่มตัวอย่างที่มีอยู่โดยใช้หลักการของความน่าจะเป็นของเบย์ กระบวนการทำงานมี 2 ขั้นตอนหลัก คือ ขั้นตอนการทำนายข้อมูลใหม่จากข้อมูลก่อนหน้า (Prior) และขั้นตอนการรวบรวมน้ำหนักของกลุ่มตัวอย่างจากความเป็นไปได้ทั้งหมด (Likelihood)

3.2 การติดตามผู้เล่นในมุมมองเสมือน (Virtual/Ground-View Player Tracking)

กระบวนการติดตามวัตถุในลักษณะหลายมุมมอง โดยทั่วไปจะเป็นการนำข้อมูลการตรวจจับและติดตามของแต่ละวัตถุจากแต่ละกล้องมารวมกันแล้วแสดงเป็นวัตถุเดียวในมุมมองเสมือน (Virtual / Top View) การแสดงตำแหน่งของวัตถุในมุมมองเสมือนโดยส่วนใหญ่จะใช้หลักการโปรเจ็คชัน (Projection) โดยการแปลงตำแหน่งจากมุมมองภาพไปเป็นมุมมองพื้น การแสดงเส้นทางของวัตถุในมุมมองเสมือนอาจใช้ข้อมูลตำแหน่งจากมุมมองภาพในการอัพเดท (Update) โดยตรงประยุกต์ใช้กราฟในการแสดงเส้นทางหรือใช้เทคนิคการประมาณ เช่น ตัวกรองคาร์ลมาลสำหรับการประมาณตำแหน่งวัตถุ เป็นต้น โดยสรุปกระบวนการที่เกี่ยวข้องในการติดตามวัตถุในลักษณะหลายมุมมองโดยทั่วไปจะมี 3 กระบวนการหลัก คือ การโปรเจ็คชัน และหรือการประมาณตำแหน่งวัตถุเป็นการกำหนดตำแหน่งของวัตถุบนมุมมองเสมือน การกำหนดเป้าหมาย (Target Assignment) หรือการเชื่อมโยงข้อมูล (Data Association) เป็นการเชื่อมโยงหรือสร้างความสัมพันธ์ระหว่างเป้าหมายในโลคอลกับเป้าหมายในมุมมองเสมือนและการผสมข้อมูล (Data Fusion) เป็นการรวมข้อมูลของวัตถุเดียวกันที่ได้จากหลายกล้องให้เป็นวัตถุเพียงอันเดียวเพื่อแสดงบนมุมมองเสมือน

2.4 การแปลงมุมมอง (Perspective Transformation)

เนื่องจากวัตถุที่ปรากฏในภาพที่ได้จากกล้องจะมีมุมมองที่แตกต่างจากมุมมองในโลกจริงกระบวนการแปลงมุมมองระหว่างมุมมองของภาพหรือระนาบภาพกับมุมมองของโลกหรือระนาบพื้นจึงถูกนำมาใช้เพื่อการแสดงผลวัตถุในมุมมองที่เหมาะสมกับงานนั้นๆ เช่น การแสดงตำแหน่งวัตถุในมุมมอง 3 มิติ เนื่องจากภาพที่ได้จากกล้องจะเป็นมุมมองแบบ 2 มิติ จึงต้องมีการแปลงตำแหน่งของวัตถุในภาพให้อยู่ในลักษณะ 3 มิติ ซึ่งเป็นตำแหน่งจริง เป็นต้น วิธีที่ใช้ในการแปลงระยะทางระหว่างมุมมอง คือ ระบบการวัดบนระนาบหรือโฮโมกราฟฟีทรานส์ฟอร์ม (Homography Transform) เป็นวิธีการวัดในระบบเรขาคณิตซึ่งถูกนำมาใช้แปลงระยะทางในพิกัดภาพไปเป็นระยะทางจริง คือ พิกัดพื้นโดยเทียบกับพิกัดอ้างอิงในระบบ 2 มิติ เมทริกซ์โฮโมกราฟฟีจะถูกใช้เป็นตัวแปลงระหว่างจุดที่เหมือนกันหรือจุดโฮโมจีเนียส (Homogeneous Point) และพิกัดที่เหมือนกัน (Homogeneous Coordinate) ภาพที่ 2.15 แสดงความสัมพันธ์ระหว่างจุดบนพิกัด



ภาพที่ 2.15 แสดงความสัมพันธ์ระหว่างจุดบนพิกัดภาพ (P) และจุดบนพิกัดจริง (P’)

จากความสัมพันธ์ระหว่างระบบพิกัดภาพ (*P*) และจุดบนพิกัดจริง (*P’*) สามารถเขียนเป็นความสัมพันธ์ได้ดังสมการที่ 2.10

$$P' = HP \tag{2.10}$$

โดย P' และ P เป็นพิกัดโฮโมจีเนียส ที่ร่วมกันของตำแหน่ง P (พิกัดภาพ) และ P' (พิกัดจริง) ซึ่งมี H เป็นเมทริกซ์โฮโมกราฟฟี่ ดังสมการที่ 2.11

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix} \tag{2.11}$$

จากสมการที่ 2.11 กำหนดตัวแปร *w* และ w' เพิ่มในสมการ ในที่นี้กำหนดให้มีค่าเท่ากับ 1 เมื่อ $\frac{x'}{w'}$ และ $\frac{y'}{w'}$ โดยที่ค่า *w* และ w' เท่ากับ 1 ซึ่งสามารถจัดรูปแบบของสมการใหม่ได้ดังสมการที่ 2.12

$$x' = \left[\frac{h_1x + h_2y + h_3}{h_7x + h_8y + h_9}\right], y' = \left[\frac{h_4x + h_5y + h_6}{h_7x + h_8y + h_9}\right] \tag{2.12}$$

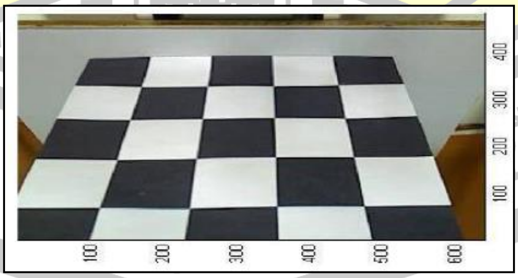
จากสมการที่ 2.12 สามารถจัดให้อยู่ในรูปแบบของ โฮโมจีเนียส ได้ดังสมการที่ 2.13 โดยตัวแปรทั้ง 9 ตัวภายในเมทริกซ์ *H* สามารถแปลความหมายเป็นอัตราการขยายหรือตัวแปลงค่าจุดระหว่างพิกัด โดยในการหาค่า เมทริกซ์โฮโมกราฟฟี่ ต้องทราบค่าจุดบนพิกัดอย่างน้อย 4 จุด

$$
\begin{bmatrix}
x_1 & y_1 & 1 & 0 & 0 & 0 & x'_1x'_1 & x'_1y'_1 & x'_1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_4 & y_4 & 1 & 0 & 0 & 0 & x'_4x_4 & x'_4y_4 & x'_4 \\
0 & 0 & 0 & x_1 & y_1 & 1 & y'_1x_1 & y'_1y_1 & y'_1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & x_4 & y_4 & 1 & y'_4x_4 & y'_4y_4 & y'_4
\end{bmatrix}
\begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \\ h_9 \end{bmatrix}
=
\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\quad (2.13)
$$

โดยเมทริกซ์โฮโมกราฟฟีนี้สามารถคำนวณได้โดยใช้ระบบสมการเชิงเส้น Singular Value Decomposition (SVD) ซึ่งเป็นวิธีการประมาณค่าแบบกำลังสองน้อยสุด (Least Squares: LS) เมื่อทราบค่าของ H จะสามารถนำไปคำนวณหาค่าจริงของตำแหน่งบนภาพได้โดยการคูณกันระหว่าง H และ P ดังสมการที่ 2.10 ข้างต้น

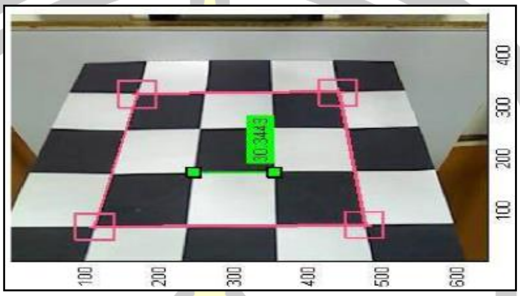
2.5 การสอบเทียบกล้อง (Camera Calibration)

จุดประสงค์ของการสอบเทียบเพื่อคำนวณหาค่าตำแหน่งหรือพารามิเตอร์กล้องซึ่งค่าที่ได้นี้จะนำไปใช้แทนค่าในเมทริกซ์โฮโมกราฟฟี่ ตัวอย่างวิธีการสอบเทียบ เช่น งานของสุเทพ อาษา และคณะ [57] ใช้ตารางที่ทราบขนาดเพื่อใช้เป็นตัวกำหนดตำแหน่ง 4 จุดของพิกัดรูปภาพเทียบกับขนาดจริงในพิกัดจริงแสดงดังภาพที่ 2.16



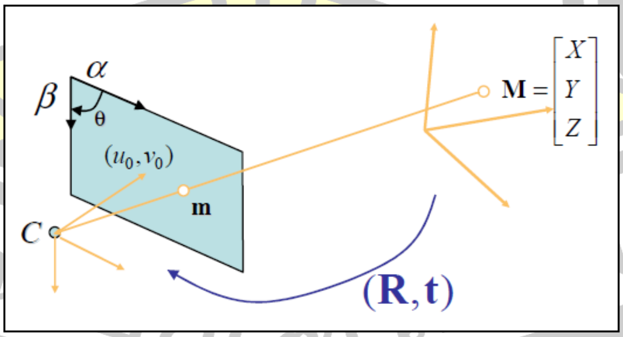
ภาพที่ 2.16 ตารางขนาด 30 x 30 mm. สำหรับสอบเทียบ (ที่มา:สุเทพ อาษาและคณะ [57])

โดยทำการกำหนดจุดแต่ละจุดและนำค่าที่ได้ไปคำนวณตามสมการที่ 2.11 โดยใช้ฟังก์ชัน SVD เช่น ในโปรแกรม Matlab ในการคำนวณหาค่า *H* ผลลัพธ์ที่ได้จะเป็นเมทริกซ์ขนาด 3 x 3 ซึ่งสามารถนำค่า *H* ไปแทนในสมการที่ 2.10 ได้ ผลลัพธ์การสอบเทียบแสดงดังภาพที่ 2.17



ภาพที่ 2.17 ตัวอย่างขั้นตอนและผลลัพธ์ที่ได้จากการสอบเทียบ (ที่มา:สุเทพ อาษาและคณะ [57])

การสอบเทียบกล้องเป็นกระบวนการหนึ่งที่ใช้หาข้อมูลเมทริกซ์จากภาพ 2 มิติ เพื่อหาพารามิเตอร์กล้องจริง โดยทั่วไปพารามิเตอร์กล้องจะแทนด้วยเมทริกซ์ขนาด 3 x 4 เรียกว่า เมทริกซ์กล้อง (Camera / Projection Matrix) การสอบเทียบกล้องเพื่อให้ได้ภาพที่ไม่บิดเบือนเนื่องจากปัญหาของเลนส์กล้องที่มีการบิดเบือนเกิดขึ้นซึ่งมี 2 ลักษณะ คือ การบิดเบือนที่เกิดจากรูปทรงของเลนส์ (Radial Distortion) และการบิดเบือนที่เกิดจากข้อบกพร่องของกระบวนการผลิต (Tangential Distortion) สัญลักษณ์ที่ใช้ในการแสดงความสัมพันธ์ระหว่างตำแหน่งต่างๆ ของโมเดลกล้องสามารถแสดงได้ดังภาพที่ 2.18



ภาพที่ 2.18 โมเดลของกล้องรูเข็ม, (ที่มา: Z. Zhang [63])

โดยที่จุด 2D คือ $m=[u,v]^T$ และจุด 3D คือ $M=[X,Y,Z]^T$ และให้ $\tilde{x}$ เป็นเวกเตอร์ที่ขยายเพิ่มโดยเพิ่มค่า 1 ในอิลีเม้นท์สุดท้ายของเวกเตอร์จะได้เป็น $m=[u,v,1]^T$ และ $M=[X,Y,Z,1]^T$ และให้ C เป็นศูนย์กลางสายตา (Optical Center) ดังนั้น ทั้ง 3 จุด คือ M, m และ C จะอยู่ในแนวเดียวกันเป็นลักษณะเส้นตรง ส่วนระนาบภาพจะอยู่ระหว่างจุดของฉาก (Scene Point) และศูนย์กลางสายตา โดยความสัมพันธ์ระหว่าง จุด M และ m สามารถแสดงได้ดังสมการที่ 2.14 ถึงสมการที่ 2.16

$$s\tilde{m}=A\begin{bmatrix} R & t \end{bmatrix}\tilde{M}\equiv P\tilde{M} \tag{2.14}$$

โดยที่

$$A=\begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 0 \end{bmatrix} \tag{2.15}$$

และ

$$P=A\begin{bmatrix} R & t \end{bmatrix} \tag{2.16}$$

โดย s คือ ปัจจัยสเกล (Scale Factor) ส่วน R และ t เป็นพารามิเตอร์ภายนอก คือ การหมุน (Rotation) และการแปลง (Translate) ซึ่งใช้เทียบเคียงระบบพิกัดโลกกับระบบพิกัดกล้อง ตัวแปร A เป็นพารามิเตอร์ภายใน ซึ่งมี (u_0,v_0) เป็นพิกัดของจุดหลัก (Principal Point) α และ β เป็น ปัจจัยสเกล ในแกน u และ v ส่วน γ เป็นพารามิเตอร์ที่อธิบายการเอียง (Skew) ของ 2 แกนในภาพ เมทริกซ์ P ขนาด 3 x 4 เรียกว่า โปรเจ็คชันเมทริกซ์ (Projection Matrix) ซึ่งจะรวมพารามิเตอร์ภายในและภายนอกเข้าด้วยกัน องศาระหว่าง 2 แกนในภาพจะแสดงด้วย θ และสามารถคำนวณค่า γ จากสมการ $\gamma=\alpha\cot\theta$ ถ้าพิกเซลเป็นรูปสี่เหลี่ยมผืนผ้า ดังนั้น $\theta=90^\circ$ และ $\gamma=0$ โดยสรุปการสอบเทียบกล้องเป็นการคำนวณหาพารามิเตอร์ของการแปลง (Transformation) ระหว่าง 3D สเปซและภาพ 2D ที่ได้จากกล้องจากข้อมูลเสมือนซึ่งการแปลงนี้จะประกอบด้วยพารามิเตอร์ภายในและภายนอกโดยเมทริกซ์การหมุน (Rotation Matrix) เนื่องจากมีทั้งหมด 9 อิลีเม้นท์ แต่มีเพียง 3 อิลีเม้นท์ที่เปลี่ยนแปลงได้ ส่วนเวกเตอร์การแปลง (Translation Vector) จะมี 3 พารามิเตอร์ ดังนั้น พารามิเตอร์ภายนอกที่ต้องคำนวณจะมีทั้งหมด 6 พารามิเตอร์ ส่วนพารามิเตอร์ภายในจะมี 5 พารามิเตอร์ รวมทั้งหมดเป็น 11 พารามิเตอร์ โดยทั่วไปจะใช้ตัวย่อเป็น A^{-T} สำหรับ $(A^{-1})^T$ หรือ $(A^T)^{-1}$ เป็นต้น

พหุน ปณฺ ชีว

อัลกอริทึมที่นิยมใช้ในการสอบเทียบกล้อง เช่น อัลกอริทึมของไซด์ [56] เป็นอัลกอริทึมที่ใช้สำหรับการสอบเทียบพารามิเตอร์ของโมเดลกล้องได้แก่ พารามิเตอร์เมทริกซ์การหมุน (Rotation Matrix: R) เวกเตอร์การแปลง (Translation Vector: T) ความยาวโฟกัสของกล้อง (Focal length of camera: f) สัมประสิทธิ์การบิดเบือนรัศมีเลนส์ (Radial lens distortion coefficient: k_1) และค่าสเกลแฟคเตอร์ (Scale factor: s_x) ประกอบด้วย 2 ขั้นตอนได้แก่ ขั้นตอนแรก การคำนวณพารามิเตอร์ภายนอกได้แก่ s_x และ R และ 2 คอมโพเนนท์แรกของเวกเตอร์การแปลง ได้แก่ T_x และ T_y รวมถึงความยาวโฟกัส f และคอมโพเนนท์ z ของเวกเตอร์การแปลง T_z ด้วยโดยใช้หลักการของระบบสมการเชิงเส้นซึ่งอินพุต คือ พิกัดของจุดในแพทเทิร์นการสอบเทียบ (Calibration Pattern) ของพิกัดภาพและพิกัดโลก ขั้นตอนที่สอง การค้นหาจุดกำเนิดของความลาดชัน (Steepest Descent) เพื่อใช้ในการคำนวณแฟคเตอร์การบิดเบือนรัศมี (Radial Distortion Factor) k_1 ซึ่งไม่สามารถคำนวณได้จากแพทเทิร์นการสอบเทียบ ส่วนค่า f และ T_z จะถูกปรับในระหว่างกระบวนการการค้นหาด้วย

2.6 การวัดประสิทธิภาพ (Performance Evaluation)

การประเมินประสิทธิภาพของระบบหรืออัลกอริทึมที่พัฒนาขึ้นมีหลายวิธียกตัวอย่าง เช่น การวัดประสิทธิภาพของโมเดลการเรียนรู้ การวัดประสิทธิภาพของความแม่นยำในการค้นหาพิกเซล และการวัดประสิทธิภาพของอัลกอริทึมการตรวจจับและติดตามวัตถุ โดยมีรายละเอียดดังนี้

1. การวัดประสิทธิภาพของโมเดลการเรียนรู้

ตัวอย่างการวัดประสิทธิภาพของโมเดลการเรียนรู้โครงข่ายประสาทเทียม [70] มี 2 ขั้นตอนย่อย คือ การใช้เทคนิค Cross-Validation และการใช้มาตรวัดประสิทธิภาพ Confusion Matrix โดยมีรายละเอียดดังนี้

1.1 การใช้เทคนิค Cross-Validation

การนำชุดข้อมูลมาใช้สำหรับสอนและทดสอบโดยแบ่งข้อมูลส่วนหนึ่งไว้สำหรับสอนและอีกส่วนไว้สำหรับทดสอบโดยการทำ Cross-Validation มีหลายวิธี เช่น Holdout Method และ K-fold Cross Validation ดังนี้

1) วิธี Holdout Method มีวิธีการ คือ ชุดข้อมูลจะถูกแบ่งออกเป็น 2 ชุด คือ ชุดข้อมูลสอน (Training Set) และชุดข้อมูลทดสอบ (Testing Set) ข้อดี คือ สามารถประเมินชุดข้อมูลจำนวนมากได้ในเวลาที่ไม่นาน

2) วิธี K-Fold Cross Validation เป็นวิธีที่ปรับปรุงมาจากวิธีแรกโดยข้อมูลจะถูกแบ่งออกเป็นส่วนย่อยแล้วกำหนดส่วนย่อยจำนวนหนึ่งเป็นชุดข้อมูลสอนและมีส่วนย่อยเพียงส่วนเดียวที่กำหนดเป็นชุดข้อมูลทดสอบตามหลักการของวิธีแรก ทำการกำหนดชุดข้อมูลสอนและชุด

ข้อมูลทดสอบในลักษณะเดียวกันแต่สลับกลุ่มไปเรื่อยๆ จนชุดข้อมูลทดสอบนั้นครอบคลุมในทุกกลุ่มของข้อมูลเป็นจำนวน k ครั้ง ข้อเสีย คือ ในการสอนต้องประมวลผลใหม่ทั้งหมด k ครั้งตามจำนวนที่แบ่งกลุ่ม

1.2 การใช้มาตรวัดประสิทธิภาพของโมเดลจำแนกประเภทข้อมูล (Confusion Matrix)

การวัดประสิทธิภาพของโมเดลจำแนกประเภทข้อมูลนั้นจะอาศัย Confusion Matrix [71, 72] ในการเก็บข้อมูลจำนวนแถวที่จำแนกจากกลุ่มข้อมูลจริงและกลุ่มข้อมูลจากการทำนาย ซึ่งตารางนั้นจะมีขนาด $m \times m$ โดยที่ m คือ จำนวนของกลุ่ม โดยเป็นรูปแบบการแบ่งแยกและเปรียบเทียบผลของการแบ่งแยกข้อมูลจริงกับข้อมูลที่เกิดจากการทำนายด้วยระบบการจำแนก (Classification System) เช่น การหาค่าข้อมูลในลักษณะของตารางเมทริกซ์ดังตารางที่ 2.1

ตารางที่ 2.1 ตาราง Confusion Matrix

		ค่าทำนาย (Prediction)	
		ผู้เล่น (Positive)	**ไม่ใช่ผู้เล่น (Negative)**
ค่าความจริง (Actual)	**ผู้เล่น (Positive)**	TP	FN
	ไม่ใช่ผู้เล่น (Negative)	FP	TN

ตามตารางที่ 2.1 มีความหมาย คือ ค่า TP และ TN เป็นจำนวนข้อมูลที่ทำนายถูกต้อง ส่วนค่า FP และ FN เป็นจำนวนข้อมูลที่ทำนายไม่ถูกต้องซึ่งจากตารางนี้สามารถคำนวณหาค่าความระลึก ค่าความแม่นยำและค่าความถูกต้องของโมเดลระบบได้โดยในงานวิจัยนี้แต่ละค่าในตารางมีความหมายดังนี้ คือ True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่าเป็นผู้เล่นและกราวนด์ทรูธ (Ground Truth) ระบุว่าเป็นผู้เล่น True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่าไม่ใช่ผู้เล่นและกราวนด์ทรูธระบุว่าไม่ใช่ผู้เล่น False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่าเป็นผู้เล่นแต่กราวนด์ทรูธระบุว่าไม่ใช่ผู้เล่น False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่าไม่ใช่ผู้เล่นแต่กราวนด์ทรูธระบุว่าเป็นผู้เล่น สำหรับค่าความถูกต้อง (Accuracy) คือ สัดส่วนของจำนวนการทำนายทั้งหมดที่ถูกต้องหาได้จากสมการที่ 2.17 ดังนี้

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \tag{2.17}$$

ค่าความระลึก (Recall) ของ TP (True Positive Rate: TPR) คือ สัดส่วนของกรณีการทำนายถูกต้องว่าเป็นผู้เล่น หาได้จากสมการที่ 2.18 ดังนี้

$$TPR = TP / (TP + FN) \tag{2.18}$$

ค่า True Negative Rate (TNR) คือ สัดส่วนของกรณีที่ทำนายถูกต้องว่าไม่ใช่ผู้เล่น หาได้จากสมการที่ 2.19 ดังนี้

$$TNR = TN / (TN + FP) \tag{2.19}$$

ค่า False Positive Rate (FPR) คือ สัดส่วนของกรณีที่ทำนายผิดว่าเป็นผู้เล่น หาได้จากสมการที่ 2.20 ดังนี้

$$FPR = FP / (FP + TN) \tag{2.20}$$

ค่า False Negative Rate (FNR) คือ สัดส่วนของกรณีที่ทำนายผิดว่าไม่ใช่ผู้เล่น หาได้จากสมการที่ 2.21 ดังนี้

$$FNR = FN / (FN + TP) \tag{2.21}$$

และค่าความแม่นยำ (Precision) P คือ สัดส่วนของกรณีที่ทำนายว่าใช่ซึ่งทำนายถูกต้อง หาได้จากสมการที่ 2.22 ดังนี้

$$Precision\ (P) = TP / (TP + FP) \tag{2.22}$$

ตัวอย่างงานวิจัยที่มีการประยุกต์ใช้ค่าความระลึกและค่าความแม่นยำในการวัดประสิทธิภาพ เช่น งานวิจัยของ Wang และคณะ [1] Utsumi และคณะ [24] และของ Yang Liu และคณะ [22] เป็นต้น

2. การวัดประสิทธิภาพของความแม่นยำในการค้นหาพิกเซล

ตัวอย่างการวัดประสิทธิภาพของความแม่นยำในการค้นหาพิกเซลของระบบโดยทั่วไปจะพิจารณาจากอัตราส่วนระหว่างจำนวนพิกเซลที่ค้นหาได้ถูกต้องกับจำนวนพิกเซลทั้งหมดดังที่เสนอไว้ในงานวิจัยของ Ngo และคณะ [27]

3. การวัดประสิทธิภาพของอัลกอริทึมการตรวจจับและติดตามวัตถุ

ตัวอย่างของการวัดประสิทธิภาพของอัลกอริทึมการตรวจจับและติดตามวัตถุ ยกตัวอย่าง เช่น การใช้เมทริกซ์ประสิทธิภาพ (Performance Matrix) ซึ่งมีรายละเอียดดังนี้

เมทริกซ์ประสิทธิภาพ [83] เป็นวิธีหนึ่งที่ถูกนำมาใช้สำหรับการประเมินประสิทธิภาพของอัลกอริทึมการตรวจจับและติดตามโดยมีพารามิเตอร์สำหรับการประเมินดังนี้

3.1 Absolute error (AE)

AE เป็นขนาดของค่าความแตกต่างระหว่างค่าจริง (True Values) กับค่าที่ถูกติดตาม (Tracked Values) สามารถคำนวณได้จากสมการที่ 2.23 ดังนี้

$$\varepsilon = \left| O_{true} - O_{tracked} \right| \tag{2.23}$$

โดยที่ O_{true} คือ ค่าจริงของพารามิเตอร์ของวัตถุส่วน $O_{tracked}$ คือ ค่าที่ถูกติดตามของพารามิเตอร์ของวัตถุโดยที่ค่า ε น้อยหมายความว่าระบบมีประสิทธิภาพที่ดี

3.2 Root Mean Square Error (RMS)

RMSE จะวัดความแตกต่างระหว่างค่าที่ถูกทำนาย (Predicted) และค่าที่เก็บได้จริง (Observed) สามารถคำนวณได้จากสมการที่ 2.24 ดังนี้

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(0_{predicted} - 0_{observed} \right)}{n}} \tag{2.24}$$

โดยที่ $0_{predicted}$ คือ การประมาณพารามิเตอร์ของวัตถุที่ถูกติดตาม ส่วน $0_{observed}$ คือ การประมาณจริงของวัตถุและ n คือ จำนวนของเฟรมที่ปรากฏวัตถุนั้น โดยที่ค่า RMSE น้อยหมายความว่าระบบมีประสิทธิภาพที่ดี

3.3 Object tracking error (OTE)

OTE คือ ความคลาดเคลื่อนเฉลี่ยของจุดศูนย์กลางของวัตถุที่ถูกติดตามกับค่าจริงของวัตถุ สามารถคำนวณได้จากสมการที่ 2.25 ดังนี้

$$OTE = \sqrt{\frac{\sum \left(x_{true} - x_{tracked} \right)^2 + \left(y_{true} - y_{tracked} \right)^2}{n}} \tag{2.25}$$

โดยที่ x_{true} และ y_{true} คือ พิกัด *x* และ *y* จริงของวัตถุ ส่วน $X_{tracked}$ และ $y_{tracked}$ คือ พิกัด *x* และ *y* ที่ถูกติดตามของวัตถุโดยที่ค่า OTE น้อยหมายความว่าระบบมีประสิทธิภาพที่ดี

3.4 Tracking Detection Rate (TDR)

TDR เป็นการตรวจสอบจำนวนของความสำเร็จในการตรวจจับวัตถุ สามารถคำนวณได้จากสมการที่ 2.26 ดังนี้

$$TDR = \frac{no_of_frame_object_detected}{no_of_frames_with_object_present} * 100 \tag{2.26}$$

โดยที่ค่า TDR สูงแสดงว่าระบบมีประสิทธิภาพที่ดี

พหุ บณ ศิต ชีเว

3.5 Bounding Box Error (BBE)

BBE เป็นการคำนวณหาพื้นที่ที่ซ้อนทับกันของวัตถุที่ถูกติดตามกับวัตถุจริง ถ้า BBE มีค่าเป็น 1 หมายความว่าการติดตามมีความถูกต้องทั้งในตำแหน่งและขนาด กรณี BBE มีค่าเป็น 0 หมายความว่ามีประสิทธิภาพที่ไม่ดี เช่น ไม่มีวัตถุที่ถูกติดตาม เป็นต้น สามารถคำนวณได้จากสมการที่ 2.27 ดังนี้

$$BBE = \frac{Area(A \cup T)}{Area(A \cap T)} \qquad (2.27)$$

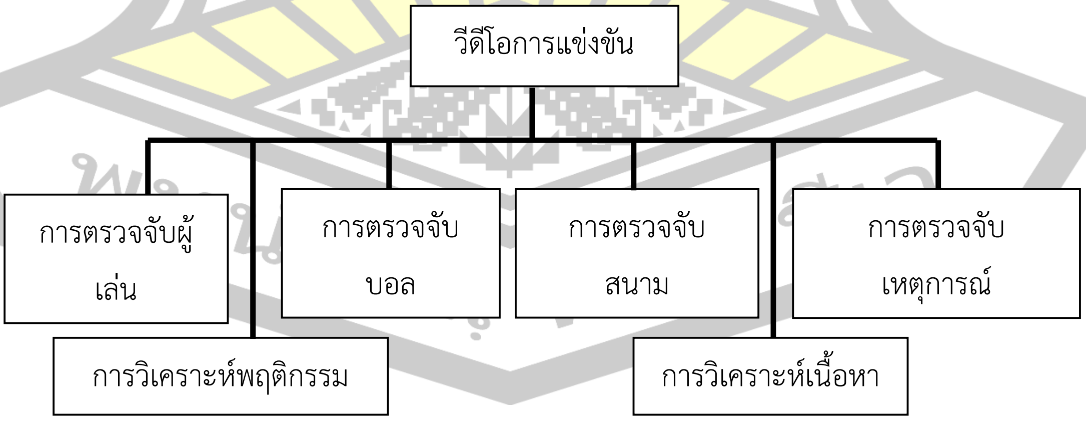
โดยที่ *A* คือ พื้นที่ (Area) จริงของวัตถุ ส่วน *T* คือ พื้นที่ของวัตถุที่ถูกติดตาม

3.6 Time Analysis

Time Analysis เป็นการประเมินประสิทธิภาพของความเร็วในการประมวลผลของระบบโดยพิจารณาจากเวลาที่ใช้ในการติดตามวัตถุในแต่ละเฟรม

2.7 งานวิจัยที่เกี่ยวข้อง (Related Research)

จากงานวิจัยที่ทำการศึกษาสามารถแบ่งกลุ่มงานวิจัยที่เกี่ยวข้องกับวิดีโอการแข่งขันกีฬาฟุตบอลออกเป็นกลุ่มหลักๆ คือ 1. กลุ่มการตรวจจับและติดตามผู้เล่น 2. กลุ่มการตรวจจับและติดตามลูกบอล 3. กลุ่มการตรวจจับสนาม 4. กลุ่มการตรวจจับเหตุการณ์ต่างๆ เช่น การตรวจจับการล้ำหน้า 5. กลุ่มที่เน้นวิเคราะห์พฤติกรรม เช่น การวิเคราะห์เทคนิคการเล่นของทีม และ 6. กลุ่มที่เกี่ยวกับการวิเคราะห์เนื้อหาข้อมูลในวิดีโอ เช่น การสรุปเนื้อหาวิดีโอ ซึ่งภาพรวมของกลุ่มงานวิจัยแสดงดังภาพที่ 2.19



ภาพที่ 2.19 การแบ่งกลุ่มประเภทงานวิจัย

ในงานนี้มุ่งเน้นในประเด็นการตรวจจับสนามและการตรวจจับและติดตามผู้เล่นในกีฬาฟุตบอลเป็นหลัก เนื่องจากการตรวจจับสนามเป็นขั้นตอนที่สำคัญขั้นตอนแรกที่จะช่วยจำกัดขอบเขตการทำงานในขั้นตอนอื่นๆ เช่น การตรวจจับผู้เล่นและช่วยลดปัญหาต่างๆ เช่น วัตถุที่ไม่ถูกพิจารณาที่อยู่บริเวณรอบสนาม เช่น ผู้ชม เป็นต้น ส่วนการตรวจจับและติดตามผู้เล่น ในงานนี้พิจารณาว่าเอาท์พุตที่ได้จากกระบวนการนี้เป็นอินพุตหลักที่สำคัญต่อกระบวนการทำงานของกลุ่มงานวิจัยอื่นๆ เช่น การตรวจจับเหตุการณ์ การวิเคราะห์พฤติกรรมและการวิเคราะห์เนื้อหาวิดีโอ เป็นต้น ดังนั้นเอาท์พุตจากกระบวนการตรวจจับสนามและการตรวจจับและติดตามผู้เล่นจึงมีความสำคัญมาก หากเอาท์พุตที่ได้ไม่ถูกต้องน่าเชื่อถืออาจจะส่งผลต่อประสิทธิภาพการทำงานในกลุ่มงานวิจัยที่ต้องใช้ข้อมูลนี้เป็นอินพุตในการดำเนินการ กระบวนการตรวจจับสนามและกระบวนการตรวจจับและติดตามผู้เล่นมีประเด็นสำคัญที่เกี่ยวข้องในการดำเนินการโดยมีรายละเอียดดังนี้

1. การตรวจจับสนาม (Playfield Detection)

Ali และคณะ [21] ประยุกต์ใช้อัลกอริทึมการตรวจจับพื้นสนามแบบอัตโนมัติเพื่อขจัดพื้นสนาม เอกลักษณ์สีที่ใช้ คือ RGB โดยมีพื้นฐานอยู่บนกฎ G>R>B ซึ่งจะถูกใช้เป็นเกณฑ์สำหรับตรวจจับสนามจากนั้นทำการแปลงภาพเป็นแบบไบนารีโดยพิกเซลที่ไม่ใช่สนามจะถูกตรวจจับตามกฎนี้และถูกทำเครื่องหมาย (Mask) โดยใช้สมการที่ 2.28 ดังนี้

$$Ground(x,y)=\begin{cases}0,\left(g(x,y)\rangle r(x,y)\rangle b(x,y)\right)\\1,otherwise\end{cases} \tag{2.28}$$

Utsumi และคณะ [24] ทำการตรวจจับพื้นสนามโดยมี 2 ขั้นตอนหลัก คือ 1) การแยกพื้นที่สนามที่อยู่นอกสนาม (Extra-Field Region) ออกโดยตรวจจับพื้นที่สีเขียวโดยใช้ค่าสี H (Hue) ที่นิยามในโมเดลรูปกรวยหกเหลี่ยมของสมิธ (Smith's Hexagonal Cone Model) และนิยามช่วงของสีที่ $\frac{1}{3}\pi \le H \le \frac{5}{6}\pi$ สำหรับพิกเซลที่เป็นสีเขียว พื้นที่ที่มีความต่อเนื่องและมีขนาดใหญ่จะถูกพิจารณาว่าเป็นพื้นสนามและส่วนพื้นที่สีเขียวที่มีขนาดเล็กจะเป็นส่วนที่อยู่นอกสนาม

Huang และคณะ [26] ใช้ตัวค้นหาสนามฟุตบอลโดยใช้พื้นฐานของฮิสโตแกรมของสีและการวิเคราะห์องค์ประกอบที่เชื่อมต่อกัน (Connected Component) ซึ่งเป็นเทคนิคที่ง่ายในการค้นหาสนามโดยใช้โมเดลสีสำหรับนำมาหาค่าฮิสโตแกรม เพื่อใช้พิจารณาว่าส่วนใด คือ สนาม โดยนำค่าฮิสโตแกรมที่ได้นี้ไปทำการเรียนรู้จากชุดข้อมูลสอนซึ่งได้ทำการสอนให้กับระบบไว้แล้ว

Liu และคณะ [28] ใช้โมเดลผสมแบบเกาเซียนในการตรวจจับสนามโดยใช้โมเดลผสมแบบเกาเซียนร่วมกับการใช้อัลกอริทึมการหาค่าคาดหมายสูงสุดแบบเพิ่มค่า (Incremental Expectation Maximization (IEM) Algorithms) สำหรับการปรับค่าพารามิเตอร์ของโมเดลและใช้ชุดแซมเปิลสำหรับสอน (Training Samples) ในการสอนให้กับโมเดลผสมแบบเกาเซียนโดยดำเนินการสอนผ่านระบบออนไลน์และทำการสอนเพียงครั้งเดียวโดยงานวิจัยนี้ให้เหตุผลว่าจะช่วยประหยัดเนื้อที่หน่วยความจำ (Buffer) สำหรับเก็บข้อมูลสำหรับสอน สำหรับผลจากการตรวจจับสนามจะถูกนำไปใช้สำหรับการรู้จำโซน (Zone) ที่สำคัญของสนามที่ปรากฏในวิดีโอ ณ ปัจจุบัน สำหรับอัลกอริทึมสำหรับตรวจจับสนามจะพิจารณาจากค่าฮิสโตแกรมของเอกลักษณ์สี คือ CbCr ในการจัดกลุ่มโดยให้เหตุผลว่าจากผลการทดลองนั้นเอกลักษณ์สีแบบ CbCr จะมีประสิทธิภาพดีกว่าเอกลักษณ์สีแบบ HSI และถ้าขอบเขตของสนามมีขนาดใหญ่พอ ค่าฮิสโตแกรมของเฟรมก็จะมีลักษณะของขอบเขตของกลุ่มที่ชัดเจนด้วย

Barnard และ Odobez [29] เสนอวิธีการแยกสนามฟุตบอลโดยใช้อัลกอริทึมการหาค่าสูงสุดของความน่าจะเป็นครั้งหลัง (Maximum a Posteriori (MAP) Adaptation) ที่มีการปรับปรุงโดยใช้โมเดลทางสถิติ คือ โมเดลผสมแบบเกาเซียนร่วมกับอัลกอริทึมการหาค่าสูงสุดของความน่าจะเป็นครั้งหลัง (MAP) ซึ่งวิธีที่เสนอนี้เป็นแบบไม่มีผู้สอนและมีความแข็งแรงต่อสิ่งรบกวน (Noise) จำนวนมากของภาพโดยกำหนดโมเดลสีด้วยโมเดลผสมแบบเกาเซียนและทำการปรับโมเดลนี้ด้วยอัลกอริทึมการหาค่าสูงสุดของความน่าจะเป็นครั้งหลัง (MAP) สำหรับโมเดลผสมแบบเกาเซียนจะถูกสอนโดยใช้ภาพสนามจากภาพการแข่งขันต่างๆ เพื่อสร้างโมเดลของสนาม ทำการสอนโมเดลสิ่งรบกวนจากภาพที่ไม่มีสนามโดยเอกลักษณ์ที่ใช้ในการสอน คือ เอกลักษณ์รงค์ของสี r และ g (Chromatic r และ g) ซึ่งคำนวณได้จาก r=R/(R+G+B), g=G/(R+G+B) และ b=B/(R+G+B) ส่วนชุดสำหรับตรวจสอบ (Validation Set) จะถูกใช้เป็นไฮเปอร์พารามิเตอร์ (Hyperparameters) ของอัลกอริทึม โมเดลที่สร้างจะถูกทดสอบด้วยภาพวิดีโอการแข่งขันประเภทต่างๆ เช่น รักบี้ ฮ็อกกี้และฟุตบอล โดยไม่ใช้วิดีโอที่ใช้ในการสอนระหว่างการทดสอบ โมเดลสนามจะถูกปรับโดยใช้อัลกอริทึมการหาค่าสูงสุดของความน่าจะเป็นครั้งหลัง (MAP) เพื่อให้ได้สีของสนามในแต่ละการแข่งขันที่ทำการทดสอบ สำหรับโมเดลผสมแบบเกาเซียนจะใช้ 2 มิติ (Dimension) ในการกำหนดโมเดลสีของสนาม x=(r, g) โดยความเป็นไปได้ (Likelihood) ของข้อมูล x จะกำหนดด้วยสมการที่ 2.29

$$p(x)=\sum_{i=1}^{N} w_i \bullet N\left(x_i,u_i,\Sigma_i\right) \tag{2.29}$$

โดยที่พารามิเตอร์ W คือเซตของค่าน้ำหนัก Σ คือ เซตของเมทริกซ์ของความแปรปรวนร่วม (Covariance) แบบทแยงมุม (Diagonal Metrics) และ μ คือ เซตของค่าเฉลี่ย

(Mean) สำหรับวิธีการสอนโมเดลผสมแบบเกาเซียนจะใช้อัลกอริทึมการหาค่าคาดหมายสูงสุด (EM) ส่วนการหาค่าสูงสุดของความน่าจะเป็นครั้งหลัง (MAP) จะมีรูปแบบดังสมการที่ 2.30

$$\hat{\theta} = \operatorname{argmax}_{\theta} P(\theta \mid X) = \operatorname{argmax}_{\theta} p(X \mid \theta) \bullet p(\theta) \tag{2.30}$$

การปรับค่าของพารามิเตอร์ *i* จะถูกปรับโดยใช้ชุดของสมการการอัพเดทค่าจากงานวิจัยของ Gauvain และ Lee [108] และงานวิจัยของ Mariethoz และ Bengio [109]

ข้อมูลพิกเซลของภาพที่ถูกใช้สำหรับการปรับนั้นจะได้มาจากสองขั้นตอนหลัก คือ 1. ขั้นตอนการเลือกสี มีวิธีการ คือ ทำการรู้จำพิกเซลของสนามโดยการปรับค่าเกณฑ์ของอัตราส่วน (Ratio) ของความน่าจะเป็นของเอกลักษณ์พิกเซลโดยพิกเซลจะถูกกำหนดลาเบล (Labeled) เป็นสนามถ้าเป็นไปตามสมการที่ 2.31 โดยที่ P_c คือ เอกลักษณ์ (r, g) ของพิกเซล *p* และ *d* คือ ค่าเกณฑ์ที่ถูกกำหนดโดยชุดข้อมูลสำหรับตรวจสอบ และ 2. ขั้นตอนการกรองค่าความเป็นไปได้ครั้งก่อน (Prior) และลักษณะรูปร่าง (Morphological) มีวิธีการ คือ เฟรมภาพแบบไบนารี (Binary Image) จะถูกสร้างโดยใช้สมการที่ 2.30 โดยความน่าจะเป็นของสิ่งรบกวนในส่วนของครึ่งบนของภาพมักจะเป็น สเตเดียม การโฆษณาและต้นไม้ ในงานวิจัยนี้พิจารณาเฉพาะครึ่งล่างของภาพโดยภาพไบนารีจะถูกทำการลดและการขยายโดยมีการแบ่งเป็นโครงสร้างสีเหลี่ยมขนาด 5 x 5 โดยการกรองนี้จะช่วยขจัดพื้นที่ขนาดเล็กของสิ่งรบกวนในภาพเพื่อให้เหลือเพียงพื้นที่ขนาดใหญ่ในการทำการปรับปรุง (Adaptation) สำหรับการอัพเดทพารามิเตอร์ของโมเดลจะกระทำทุกๆ 30 วินาที

$$\frac{p\left(c_p \mid \theta_k^{grass}\right)}{p\left(c_p \mid \theta^{non-grass}\right)} > d \tag{2.31}$$

ในส่วนของการทดลอง ทำการแบ่งข้อมูลเป็น 3 ชุด คือ ชุดข้อมูลสำหรับสอนจาก 6 การแข่งขันของฟุตบอล ชุดตรวจสอบจาก 4 การแข่งขันของฟุตบอลและชุดทดสอบจาก 8 การแข่งขันของรักบี้ 2 การแข่งขันของฟุตบอลและ 1 การแข่งขันของฮ็อกกี้ โมเดลสีของสนามจะถูกสอนโดยภาพสนามจำนวน 163 ภาพและโมเดลของสิ่งรบกวนจะถูกสอนโดยใช้ภาพที่ไม่มีสนามจำนวน 87 ภาพ โดยในงานวิจัยนี้ได้เปรียบเทียบประสิทธิภาพกับงานวิจัยของ Utsumi และคณะ [110] ที่ใช้วิธีกำหนดเกณฑ์ค่าสีแบบง่าย (Simple Hue Thresholding) และวิธีของ Xu และคณะ [111] ที่ใช้วิธีกำหนดเกณฑ์ค่าสีทางสถิติ (Statistical Hue Thresholding) โดยหาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าสีที่เด่น (Dominant) ซึ่งผลการทดลองระบุว่ามีประสิทธิภาพที่ดีกว่าทั้งสองวิธีโดยในชุดข้อมูลทดสอบจำนวน 5 ชุดมีค่าประมาณ 0.9 ทั้ง 5 ชุดทดสอบ

Ngo และคณะ [27] เสนอวิธีการตรวจจับสนามโดยใช้วิธีที่เรียกว่า พื้นที่ครอบคลุม (Area-of-Coverage: AoC) โดยประเด็นปัญหา คือ ความแม่นยำในการตรวจจับสนามของวิธีแบบเก่ายังมีความหลากหลายซึ่งขึ้นอยู่กับรูปแบบของหญ้า (Grass Pattern) ของสนามในงานวิจัยนี้จึงเสนอวิธีการค้นหาพื้นที่ของสนาม (Playfield Region) โดยใช้แนวคิดที่เรียกว่า พื้นที่ครอบคลุม (AoC) ในการเพิ่มความแม่นยำของการตรวจจับสนามโดยใช้ข้อมูลของสีและข้อมูลของโฮโมกราฟฟี่ (Homography Information) ร่วมกัน สำหรับระบบที่เสนอประกอบด้วย 3 ส่วนหลัก คือ 1. การตรวจจับสนาม (Field Detection) 2. การประมาณโฮโมกราฟฟี่ (Homography Estimation) และ 3. การแบ่งแยกพื้นที่ครอบคลุม (AoC) ในส่วนของการตรวจจับสนามจะใช้คิวของสี (Color Cue) เพื่อทำการประมาณแบบหยาบ (Rough Estimation) พื้นที่สนามและในส่วนของการประเมินโฮโมกราฟฟี่และส่วนของการแบ่งแยกพื้นที่ครอบคลุม (AoC) จะเป็นส่วนที่ใช้ปรับผลลัพธ์โดยใช้ข้อมูลของโฮโมกราฟฟี่ ส่วนของขั้นตอนการประมาณสนาม (Field Estimation) มี 3 ขั้นตอนย่อย คือ 1. การตรวจจับขอบเขตของสีที่สำคัญ (Dominant Color Region Detection) 2. การประเมินโมเดลผสมแบบเกาเซียน (GMM Estimation) และ 3. การปรับค่ารีเจียนโกร์ววิ่ง (Region-Growing Refinement) โดยทำการตรวจจับขอบเขตของสีที่สำคัญเพื่อกรองพิกเซลที่อยู่นอกเหนือจากสีที่สำคัญและกำหนด (Fit) เป็นอินพุตให้กับโมเดลผสมแบบเกาเซียน ในขั้นตอนการประเมินโมเดลผสมแบบเกาเซียนจะทำการวิเคราะห์พิกเซลเหล่านี้และทำการจัดกลุ่มต่อไป ในงานวิจัยนี้ระบุว่าสนามที่ค้นพบมักจะมีเศษเล็กๆ หรือแฟรกเมนท์ (Fragment) ระหว่างพิกเซลของหญ้ามากเนื่องจากสิ่งรบกวนหรือการจัดกลุ่มที่ไม่ถูกต้อง (Miss-Classification) ซึ่งในขั้นตอนของการปรับรีเจียนโกร์ววิ่งจะถูกนำมาใช้เพื่อขจัดเศษเล็กๆ เหล่านี้เพื่อให้ได้ผลลัพธ์ที่สะอาดมากขึ้น ในขั้นตอนการตรวจจับขอบเขตของสีที่สำคัญจะใช้ค่าฮิสโตแกรมของสีเพื่อทำการวิเคราะห์ข้อมูลทางสถิติของสีของพิกเซลโดยทำการแปลงภาพเป็นแบบ YCbCr ซึ่งงานวิจัยนี้ระบุว่าพื้นที่สีนี้จะอ่อนไหวต่อสิ่งรบกวนน้อยกว่าพื้นที่สีแบบอื่น เช่น RGB หรือ HSV โดยขอบเขตของสีที่สำคัญจะถูกคำนวณจากฮิสโตแกรมของภาพในส่วนของถังสี (Bin) ที่เชื่อมต่อกัน (Connected Bins) ที่ใหญ่ที่สุดหรือที่มีจำนวนของพิกเซลที่สูงสุด ในขั้นตอนการประเมินโมเดลผสมแบบเกาเซียน หลังจากกำหนดขอบเขตของสีที่สำคัญในฮิสโตแกรมแล้วจะใช้โมเดลผสมแบบเกาเซียนที่มีการปรับปรุงมาใช้ในการกรองพิกเซลภายในถังสีออกเป็นกลุ่มของสีที่สัมพันธ์กัน (Color-Related Cluster) โดยเฟรมแรกจะใช้อัลกอริทึมการหาค่าคาดหมายสูงสุด (EM) ในการกำหนดค่าเริ่มต้นให้กับพารามิเตอร์ของโมเดลผสมแบบเกาเซียนและในเฟรมถัดไปจะใช้อัลกอริทึมการหาค่าคาดหมายสูงสุด (EM) ในการอัพเดทโมเดลผสมแบบเกาเซียนโดยวิธีการดำเนินการนี้ได้ประยุกต์มาจากงานวิจัยของ Liu และคณะ [28] สำหรับกระบวนการค้นหาในงานวิจัยนี้ใช้คอมโพเนนท์ของเกาเซียน (Gaussian Component) จำนวน 3 คอมโพเนนท์ สำหรับโมเดลผสมแบบเกาเซียนโดย 2 คอมโพเนนท์จะใช้สำหรับกำหนดโมเดลของคอมโพเนนท์ที่เป็นหญ้า

ของสนาม (Field Grass Component) อีก 1 คอมโพเนนท์จะใช้สำหรับกำหนดโมเดลของคอมโพเนนท์ที่เป็นหญ้าที่ไม่ใช่ของสนาม (Non-Grass Component) ที่อยู่ภายในขอบเขตของสนาม จุดประสงค์เพื่อแยกแยะส่วนประกอบ (Element) ของหญ้า (Grass) และที่ไม่ใช่หญ้า (Non-Grass) ให้ชัดเจนขึ้นและในเวลาเดียวกันก็สามารถจัดการกับความแปรปรวนของรูปแบบของหญ้า (Grass Pattern) ได้ด้วย ในงานวิจัยนี้ระบุว่าพิกเซลที่ถูกจัดกลุ่มอยู่ในคอมโพเนนท์ของเกาเซียนที่มีขนาดใหญ่สุดทั้งสองคอมโพเนนท์จะเป็นพิกเซลของสนาม (Field Pixel) ส่วนพิกเซลที่ถูกจัดกลุ่มอยู่ในคอมโพเนนท์ของเกาเซียนที่มีขนาดเล็กสุดจะเป็นพิกเซลของส่วนที่ไม่ใช่สนาม (Non-Field Pixel) ในขั้นตอนการปรับค่ารีเจียนโกร์ววิ่ง ซึ่งเป็นขั้นตอนสุดท้ายของกระบวนการประเมินสนามจะนำผลลัพธ์ที่ได้จากขั้นตอนของการประเมินโมเดลผสมแบบเกาเซียนมาปรับปรุงโดยการกรองเอาขอบเขตที่มีขนาดเล็กที่อยู่แยกจากกลุ่มออกไปซึ่งขอบเขตเหล่านี้อาจเป็นสิ่งรบกวนที่อยู่ภายในพื้นที่สนามหรือเป็นพิกเซลที่ไม่ใช่ของสนาม (Field’s Alike Pixel) ที่อยู่ภายนอกขอบเขตของสนาม สำหรับผลการทดลองในงานวิจัยของ Ngo และคณะ [27] นี้ได้ทำการเปรียบเทียบกับงานวิจัยของ Ekinand และ Tekalp (2003) ที่ใช้วิธีการแยกสีที่สำคัญ (Dominant Color Segmentation) และงานวิจัยของ Liu และคณะ [28] ที่ใช้วิธีการแยกโดยใช้โมเดลผสมแบบเกาเซียนที่มีการปรับปรุง (Adaptive GMM Segmentation) ซึ่งเป็นวิธีการแบ่งแยกสนามแบบไม่มีพารามิเตอร์และมีพารามิเตอร์ตามลำดับ (Non-Parametric and Parametric Playfield Segmentation) โดยงานวิจัยนี้ระบุว่าวิธีที่เสนอมีความถูกต้องมากกว่าอีกสองวิธีดังกล่าว โดยระบุว่าเนื่องจากระบบสามารถขจัดพื้นที่หญ้าที่อยู่นอกสนามได้โดยใช้พื้นฐานของข้อมูลโฮโมกราฟฟี่ โดยให้ข้อสังเกตว่าในการแบ่งแยกสีที่สำคัญ (Dominant Color Segmentation) ให้ผลที่ดีเมื่อสีของสนามมีลักษณะรูปแบบที่เหมือนกัน (Homogenous) แต่จะมีความผิดพลาดในจัดการกับความแปรปรวนเมื่อเกิดเฉดสี (Shades) ในสนามซึ่งโมเดลผสมแบบเกาเซียนจะจัดการปัญหานี้ได้ดีโดยใช้แบบหลายเกาเซียน (Multiple Gaussian) เพื่อโมเดลรูปแบบหญ้าของสนาม ในการวัดคุณภาพของความแม่นยำของการค้นหาในงานวิจัยนี้ใช้วิธีคำนวณบนพื้นฐานของอัตราส่วน (Ratio) ของจำนวนของพิกเซลทั้งหมดที่ถูกค้นพบลบด้วยจำนวนของพิกเซลที่ผิดพลาดบนจำนวนของพิกเซลสนามที่เป็นค่าจริง (Ground-Truth) โดยทำการกำหนดสัญลักษณ์ด้วยมือ (Manually Annotated)

Hung และคณะ [34] เสนอวิธีการเซกเมนต์สนามทั่วไปของวิดีโอกีฬาโดยใช้เอกลักษณ์สีร่วมกับการใช้ฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability Density Function (PDF)) ในการประมาณคอมโพเนนท์สีของเฟรมภาพและเสนอแบบแผนฮิลไคลมิง (Hill-Climbing Schemes) 2 แบบแผน คือ ใช้ฟังก์ชันความหนาแน่นความน่าจะเป็นแบบ 2 มิติ และฟังก์ชันความหนาแน่นความน่าจะเป็นแบบ 1 มิติ สำหรับการจัดกลุ่ม และได้พัฒนาอัลกอริทึมใหม่ที่ใช้ความรู้หลัก (Domain Knowledge) ของสนามกีฬาเพื่อผสานกลุ่มเหล่านั้นเข้าเป็น 4 คลาสสี คือ สีแดง สีเขียว สีน้ำเงินและ

สีเทา สุดท้ายแบบแผนนี้จะรวมพื้นที่ขนาดเล็กเข้ากับพื้นที่ขนาดใหญ่ที่อยู่ติดกันกับมันเพื่อให้ได้ผลของการเซกเมนต์ เหตุผลของงานวิจัยนี้ คือ เนื่องจากงานวิจัยส่วนใหญ่ที่ผ่านมามักจะเน้นที่การเซกเมนต์สนามเพียงประเภทเดียว หลายงานวิจัยใช้วิธีการกำหนดค่าเทรชโฮลด์ ซึ่งมักจะเกิดข้อผิดพลาดเพราะสีที่แตกต่างกันเมื่อมีการเปลี่ยนแปลงของแสง ส่วนวิธีการที่ซับซ้อนในการกำหนดโมเดลพื้นหลัง เช่น โมเดลผสมแบบเกาเซียนซึ่งใช้อธิบายการกระจายตัวของสนามในเอกลักษณ์สเปซ ซึ่งมีข้อเสีย คือ มีความซับซ้อนในการคำนวณและมักใช้ตรวจสอบสนามที่เป็นชนิดหญ้าเพียงอย่างเดียว โดยงานวิจัยส่วนใหญ่มุ่งเน้นการแก้ปัญหาความแปรปรวนของแสงซึ่งส่วนใหญ่ออกแบบสำหรับเฉพาะกีฬาหรือชนิดของสนามเพียงชนิดเดียว บางงานวิจัย เช่น งานวิจัยของ Jiang และคณะ [35] ได้เสนอวิธีการตรวจจับสนามโดยใช้โมเดลผสมแบบเกาเซียน จากนั้นใช้วิธีที่เรียกว่า รีเจียนโกร์ววิ่ง เพื่อเชื่อมพิกเซลสนามที่ไม่ต่อเนื่องเข้ากับพื้นที่ ทำการขจัดสิ่งรบกวนและทำการปรับค่าความสมูธ (Smooth) ขอบเขต (Boundaries) ของวัตถุ เป็นต้น ในงานวิจัยนี้ได้ให้เหตุผลไว้ว่าโดยปกติสนามแต่ละที่จะมีลักษณะที่แตกต่างกัน เช่น สนามภายนอกและภายในร่มจะใช้วัสดุที่แตกต่างกัน เช่น หญ้าและไม้ เป็นต้น ดังนั้นเอกลักษณ์สีจึงมักถูกนำมาใช้สำหรับการเซกเมนต์สนามในวิดีโอกีฬา ดังนั้นในงานวิจัยนี้จึงเสนอวิธีการที่สามารถเซกเมนต์สนามได้ทั่วไปโดยใช้ฟังก์ชันความหนาแน่นความน่าจะเป็นและใช้วิธีของอัลกอริทึมฮิลไคลมิง (Hill Climbing Algorithm) เพื่อสร้างกลุ่ม จากนั้นใช้วิธีการผสานกลุ่มสีทั้งสี่กลุ่มเพื่อเซกเมนต์ภาพออกเป็นพื้นที่และใช้แบบแผนการหลอมพื้นที่ในการผสานพื้นที่ขนาดเล็กเข้ากับพื้นที่ขนาดใหญ่ ผลการทดลองระบุว่าวิธีที่เสนอดีกว่าวิธีที่อยู่บนพื้นฐานของโมเดลผสมแบบเกาเซียนทั้งในเรื่องความถูกต้องในการเซกเมนต์และประสิทธิภาพในการคำนวณและมีการคำนวณที่น้อยกว่าวิธีบนพื้นฐานของมีนชิฟท์ (Mean-Shift) โดยได้ทดลองกับวิดีโอกีฬาประเภทต่างๆ เช่น บาสเก็ตบอล เบสบอล ฟุตบอล เทนนิสและวอลเลย์บอล ในการประเมินประสิทธิภาพวัดจากความถูกต้องของการเซกเมนต์โดยคำนวณได้จากจำนวนพิกเซลสนามจริงบวกกับจำนวนพิกเซลที่ไม่ใช่สนามจริงหารด้วยจำนวนพิกเซลทั้งหมด

Jiang และคณะ [35] เสนอวิธีการเซกเมนต์สนามและการประยุกต์ใช้กับการวิเคราะห์การแข่งขันในวิดีโอการแข่งขันกีฬา โดยงานวิจัยนี้ให้เหตุผลไว้ 3 ประเด็นคือ 1. วิดีโอกีฬาเป็นสิ่งที่สนใจจากผู้ชมจำนวนมาก เนื้อหาวิดีโอแบบดิจิตอลมีเพิ่มมากขึ้นและผู้ใช้ต้องการเข้าถึงเนื้อหา (Content) ผ่านวิธีการทางเครือข่ายต่างๆ ด้วยอุปกรณ์ทางด้านดิจิตอลต่างๆ การดึงข้อมูล (Information) ที่เป็นประโยชน์แบบอัตโนมัติจากวิดีโอช่วยอำนวยความสะดวกในการค้นคืนและการจัดระบบซึ่งถือเป็นปัญหาที่สำคัญ 2. สนามเป็นส่วนที่สำคัญพื้นฐานในการวิเคราะห์แบบอัตโนมัติในหลายๆ โปรแกรมด้านกีฬา เบาะแสเชิงความหมายต่างๆ มักถูกสรุปมาจากการเซกเมนต์สนาม (Playfield Segmentation) และ 3. จากลักษณะการปรากฏของสนามที่แตกต่างกันออกไปซึ่งเกิดจากชนิดของกีฬาที่ต่างกัน สนามกีฬา (Stadium) ที่ต่างกัน เงื่อนไขทางด้านแสงและสภาพอากาศ

ภายในสนามกีฬารวมถึงลักษณะของพื้นสนามที่ไม่ได้มีลักษณะที่เหมือนกันตลอดทั้งสนามโดยอาจมีรูปแบบบางอย่าง เช่น เป็นลักษณะลายตาราง ดังนั้น การเลือกวิธีเพื่อประมาณพื้นที่สนามนั้นต้องมีลักษณะที่ทั่วๆ ไปที่เพียงพอเพื่อโมเดลการกระจายตัวของสีที่หลากหลายได้ กระบวนการเซกเมนต์สนามจะใช้เทคนิคโมเดลผสมแบบเกาเซียนและอัลกอริทึมรีเจียนโกร์ววิ่งโดยใช้ชุดพิกเซลสำหรับสอน (Training Pixel) ซึ่งถูกเลือกเป็นกลุ่มตัวอย่างที่ได้มาจากเฟรมแบบอัตโนมัติ จากนั้นได้สมมติว่าพิกเซลสนามเป็นคอมโพเนนท์ที่เด่นเป็นส่วนใหญ่ในวิดีโอ ทำการสร้างโมเดลผสมแบบเกาเซียนของพิกเซลสนามและใช้โมเดลนี้ในการตรวจจับพิกเซลสนาม สุดท้ายใช้วิธีที่เรียกว่ารีเจียนโกร์ววิ่งในการแบ่งส่วนพื้นที่สนามออกจากพื้นหลัง การประมาณค่าความหนาแน่นจะพิจารณาจากค่าฮิสโตแกรมโดยใช้โมเดลผสมแบบเกาเซียนด้วยวิธีแบบกึ่งพารามิเตอร์ซึ่งโมเดลผสมแบบเกาเซียนเป็นวิธีทั่วๆ ไปที่เพียงพอต่อการโมเดลการกระจายตัวที่เป็นแบบไม่เชิงเส้นและมีความซับซ้อนที่สูง ภายหลังการตรวจจับพิกเซลสนามที่ไม่ต่อเนื่อง สิ่งรบกวนอาจยังมีอยู่ทำให้มันไม่มีความเป็นรูปแบบของพื้นที่จึงใช้วิธีรีเจียนโกร์ววิ่งในการเชื่อมพิกเซลสนามเข้ากับพื้นที่นั้นๆ ขจัดสิ่งรบกวนและปรับค่าความสมูธและขอบเขตให้เรียบและสุดท้ายสร้างผลของการเซกเมนต์เพื่อใช้สำหรับการวิเคราะห์เนื้อหาวิดีโอ โดยได้กล่าวถึงข้อดีของวิธีการเซกเมนต์สนามที่เสนออยู่ 2 ประเด็น คือ 1. มีความเป็นอิสระที่แข็งแรงต่อชนิดของกีฬาและลักษณะการปรากฏ (Appearance) แม้จะมีเงื่อนไขของหญ้าที่แย่มากๆ และ 2. ผลการแบ่งส่วนจะได้ข้อมูลพื้นที่สีที่เด่น (Dominant Color) และเป็นส่วนที่กักเก็บวัตถุที่อยู่ในสนามไว้ซึ่งจะอำนวยความสะดวกต่อการวิเคราะห์การแข่งขันกีฬาได้ การเซกเมนต์สนามจะใช้ชุดข้อมูลสอนแบบสุ่มจำนวน 100 เฟรมโดยพิกเซลจะถูกสุ่มมาอย่างสม่ำเสมอจากแต่ละเฟรม การกระจายตัวของสีของพิกเซลจะถูกโมเดลรูปแบบของค่าความสว่างของสี (Hue-Luminance Space) เพราะสี Hue จะสะท้อนให้เห็นถึงสีของสนามและความสว่างจะสอดคล้องกับความสว่างของสีที่สะท้อนให้เห็นถึงเงื่อนไขของแสงของสนามกีฬา จากการทดลอง โมเดลสี HL สามารถเทียบได้กับ HSL และมีประสิทธิภาพดีกว่าวิธีการแสดงสีแบบอื่นๆ ในการประมาณค่าพารามิเตอร์ทั้ง 3 ตัว คือ เวกเตอร์ค่าเฉลี่ย เมทริกซ์ความแปรปรวน และค่าน้ำหนักของโมเดล โมเดลผสมแบบเกาเซียนจะใช้อัลกอริทึมการหาค่าคาดหมายสูงสุด (EM) โดยหลังขั้นตอนการสอนจะได้คอมโพเนนท์ของโมเดลผสมแบบเกาเซียนจำนวน 4 คอมโพเนนท์ สำหรับการทำเซกเมนต์ งานวิจัยนี้ได้เสนออัลกอริทึมรีเจียนโกร์ววิ่งวิธีใหม่ด้วยซึ่งจากการทดลองได้สรุปว่าการใช้โมเดลผสมแบบเกาเซียนในการเซกเมนต์จะให้ผลที่ดีกว่ากรณีใช้ฮิสโตแกรมและในกรณีการใช้อัลกอริทึมรีเจียนโกร์ววิ่ง (Region Growing Algorithm) หลังจากใช้โมเดลผสมแบบเกาเซียนได้สรุปไว้ว่าสามารถช่วยเพิ่มประสิทธิภาพในการเซกเมนต์ได้

จากการศึกษางานวิจัยข้างต้นมีประเด็นของการตรวจจับสนามสามารถสรุปได้ดังนี้ คือ เนื่องจากการตรวจจับสนาม (Playfield Detection) ในวิดีโอการแข่งขันกีฬาโดยเฉพาะกีฬาฟุตบอลซึ่งเป็นกีฬาที่ได้รับความนิยมสูงในปัจจุบัน เป็นกระบวนการแรกที่สำคัญซึ่งจะเป็นพื้นฐานต่อ

กระบวนการวิเคราะห์ความหมายในระดับที่สูงกว่าอื่นๆ ดังที่กล่าวไว้ในงานวิจัยของ Ngo และคณะ [27] นอกจากนี้การตรวจจับและการทำความเข้าใจผู้เล่นในเหตุการณ์ต่างๆ ที่เกิดขึ้นในเกมส์การแข่งขันถือเป็นงานที่ท้าทายโดยปัญหาที่มักจะพบได้บ่อย คือ การเคลื่อนที่ของวัตถุในลักษณะที่ไม่สัมพันธ์กัน (Non-Related) การบดบังระหว่างวัตถุ สัญญาณรบกวนอื่นๆ และการเปลี่ยนแปลงของแสง เป็นต้น ดังนั้นเพื่อลดความซับซ้อนของงานการตรวจจับและติดตาม เทคนิคที่นิยมนำมาใช้ คือ การเซกเมนต์สนาม ซึ่งเป็นวิธีที่ถูกนำมาใช้ในหลายงานวิจัยทางด้านคอมพิวเตอร์วิชัน (Computer Vision) ซึ่งถูกพิจารณาว่าเป็นขั้นตอนแรกที่จำเป็นต่อกระบวนการวิเคราะห์ข้อมูลในระดับสูงในกระบวนการอื่นๆ เช่น การตรวจจับและการติดตามผู้เล่นและลูกบอลรวมถึงการระบุตัวตนของผู้เล่น [3, 8, 13] ดังนั้นในงานวิจัยนี้จะดำเนินการตรวจจับสนามเป็นขั้นตอนแรก ผลจากการตรวจจับสนามและเส้นสนามจะถูกใช้เป็นความรู้พื้นฐานเกี่ยวกับโครงสร้างสนามซึ่งจะใช้ประกอบสำหรับการตรวจจับและการจำแนกตัวตนของผู้เล่นและกรรมการในขั้นตอนถัดไป

เทคนิคหรือวิธีการหนึ่งที่นิยมนำมาใช้ในการตรวจจับสนามในภาพของเกมส์การแข่งขันฟุตบอลนั้น คือ การใช้โมเดลทางสถิติ (Model Statistical) ซึ่งเทคนิคนี้จะแบ่งพิกเซลของภาพออกเป็น 2 กลุ่ม (Classes) ได้แก่ พิกเซลที่เป็นสนาม (Playfield Pixels) และพิกเซลที่ไม่ใช่สนาม (Non-Playfield Pixels) โดยโมเดลทางสถิติที่นิยมใช้โมเดลหนึ่ง คือ โมเดลผสมแบบเกาเซียน โดยเป็นโมเดลที่นำมาประยุกต์ใช้ในการเซกเมนต์ พื้นที่ของสนามออกจากส่วนอื่นๆ ของภาพ ตัวอย่างของงานวิจัยที่นำโมเดลดังกล่าวมาประยุกต์ใช้ในการตรวจจับหรือเซกเมนต์สนาม เช่น งานวิจัยของ Liu และคณะ [28] ซึ่งทำการจัดกลุ่มพิกเซลออกเป็นกลุ่มของพิกเซลสนามและกลุ่มของพิกเซลที่ไม่ใช่สนามและใช้อัลกอริทึมการหาค่าคาดหมายสูงสุดแบบเพิ่มค่า (IEM) ในการปรับค่าพารามิเตอร์ของโมเดลโดยใช้วิธีที่เรียกว่า ลีสสแควร์เมธอด (Least Squares Method) ในงานวิจัยนี้ได้ทำการทดลองกับวิดีโอการแข่งขันกีฬาฟุตบอลและบาสเก็ตบอลซึ่งผลการทดลองพบว่าบรรลุผลตามที่คาดหวังไว้ ในงานที่คล้ายกัน คือ งานวิจัยของ Ngo และคณะ [27] ซึ่งใช้เกาเซียนในการระบุพื้นที่ของสนามในภาพเอกลักษณ์ที่ใช้ คือ ฮิสโตแกรมของสี ส่วนการจำแนกข้อมูล (Classification) จะดำเนินการโดยใช้อัลกอริทึมการหาค่าคาดหมายสูงสุด (EM) และใช้เทคนิคที่เรียกว่า พื้นที่ครอบคลุม (AoC) ในการปรับค่าพารามิเตอร์ การประเมินระบบได้ทดลองกับวิดีโอการแข่งขันฟุตบอลซึ่งพบว่าให้ผลลัพธ์ที่ดีและสามารถนำเทคนิคดังกล่าวไปประยุกต์ใช้กับงานด้านอื่นๆ ได้ เช่น ระบบเฝ้าระวัง (Surveillance System) หรือการรีจีสเตรชันภาพแบบ 3 มิติ (3D Image Registration) งานวิจัยของ Lin และคณะ [22] ได้ใช้เทคนิคบนพื้นฐานของโมเดลผสมแบบเกาเซียนร่วมกับเอกลักษณ์ฮิสโตแกรมของสีโดยได้ทดสอบกับวิดีโอการแข่งขันฟุตบอลที่ทำการออกอากาศ (Broadcast) ในงานวิจัยของ Barnard and Odobez [29] ได้ใช้โมเดลผสมแบบเกาเซียนร่วมกับเทคนิคที่เรียกว่าอัลกอริทึมการหาค่าสูงสุดของความน่าจะเป็นครั้งหลัง (MAP) สำหรับการเซกเมนต์สนาม โดยได้ทำการทดสอบกับวิดีโอการแข่งขัน

กีฬาหลายประเภท เช่น ฟุตบอล รักบี้และฮ๊อกกี้ เป็นต้น นอกจากนี้ Bin และคณะ [30] ได้เสนอเทคนิคที่เรียกว่า AGMM ซึ่งได้ทำการปรับปรุงมาจากเทคนิคโมเดลผสมแบบเกาเซียนและยังใช้อัลกอริทึมการหาค่าคาดหมายสูงสุง ในการปรับค่าพารามิเตอร์ของโมเดล เทคนิคที่เสนอนี้จะทำการแยกพิกเซลออกเป็นกลุ่มของพิกเซลสนามและกลุ่มของพิกเซลที่ไม่ใช่สนาม นอกจากนี้ในงานวิจัยนี้ยังทำการจัดกลุ่มของฉากตามเหตุการณ์ที่เกิดขึ้นในเกมส์การแข่งขันด้วย การทดลองได้ทดสอบกับวิดีโอการแข่งขันกีฬาหลายประเภทได้แก่ ฟุตบอล แบตมินตัน เทนนิสและวิดีโอที่มีคุณภาพที่แตกต่างกัน ผลการทดลองพบว่าเทคนิคที่เสนอสามารถตรวจจับสนามได้อย่างมีประสิทธิภาพเมื่อเปรียบเทียบกับวิธีเดิม คือ การใช้โมเดลผสมแบบเกาเซียนร่วมกับอัลกอริทึมการหาค่าคาดหมายสูงสุง นอกจากนี้วิธีการจำแนกสถานการณ์ (Scenario Classification) ที่เสนอยังสามารถใช้ในการประมาณการกระจายตัว (Distribution) ของพื้นที่สนามภายในภาพได้อย่างมีประสิทธิภาพ

ข้อดีของโมเดลผสมแบบเกาเซียน คือ โมเดลนี้ใช้ฟังก์ชันเกาเซียนแบบ 1 ฟังก์ชัน (Single Gaussian Function) รวมกันหลายๆ ฟังก์ชันในการทำงานเป็นลักษณะแบบหลายฟังก์ชัน (Multiple Single Gaussian Function) จึงสามารถอธิบายกลุ่มข้อมูลหลายๆ กลุ่มได้ สามารถอธิบายการเปลี่ยนแปลงของค่าของพิกเซลในฉากที่ซับซ้อนได้ สามารถอธิบายการกระจายตัวของโมเดลพื้นหลังในสถานการณ์ที่ซับซ้อนได้ (Complex Scenarios) และให้ผลที่ดีในการประมวลผลพื้นหลังที่มีการเปลี่ยนแปลงรูปแบบ (Mutation Background) การรบกวนพื้นหลัง (Background Disturbance) รวมถึงสถานการณ์อื่นๆ ด้วยซึ่งข้อดีเหล่านี้ได้อธิบายไว้ในงานวิจัยของ Bin และคณะ [30] ส่วนในงานวิจัยของ Liu และคณะ และ Reynolds และคณะ [28, 75] ได้ให้เหตุผลที่คล้ายกันในการเลือกใช้โมเดลผสมแบบเกาเซียนว่าเป็นโมเดลที่สามารถอธิบายความหนาแน่นของการกระจายตัวได้ดี (Probability Density) และนอกจากนี้ในบางสนามกีฬา สีของสนามอาจจะมีลักษณะที่ไม่เหมือนกันรวมถึงโมเดลผสมแบบเกาเซียนนั้นยังสามารถทำการปรับค่าพารามิเตอร์ของโมเดลได้อัตโนมัติซึ่งทำให้ได้ค่าที่เหมาะสมที่สุดในแต่ละสถานการณ์หรือในแต่ละปัญหาด้วย อย่างไรก็ตามเนื่องจากเกาเซียนจำเป็นต้องใช้ข้อมูลที่มีจำนวนไดเมนชันที่สูง (High Dimensional Data) ซึ่งถือเป็นปัญหาที่สำคัญปัญหาหนึ่งของเทคนิคนี้

นอกจากการใช้เกาเซียนในการเซกเมนต์สนามแล้วยังมีวิธีการจำแนกข้อมูลโดยใช้อัตราส่วนของสี (Color Ratio) และโลคอลเอนโทรปี (Local Entropy) ซึ่งถูกนำเสนอในงานวิจัยของ Yu และคณะ [84] วิธีที่เรียกว่า Generalized Lloyd Algorithm (GLA) เป็นวิธีที่มีพื้นฐานอยู่บนการใช้เอกลักษณ์สีซึ่งถูกนำมาใช้ในการตรวจจับบริเวณของหญ้า (Grass Region) ที่ปรากฏภายในภาพโดยถูกนำเสนอในงานวิจัยของ Bai และคณะ [85] รวมถึงการจำแนกข้อมูลโดยมีพื้นฐานอยู่บนเอกลักษณ์ฮิสโตแกรมแบบทั่วไปซึ่งถูกนำเสนออยู่ในงานวิจัยของ Ekin และคณะ [86] และในงานวิจัยของ Zameni และคณะ [87] เป็นต้น

ในงานวิจัยของ Barnard และ Odobez [29] ได้เสนอวิธีการค้นหาสนามของกีฬาประเภทต่างๆ โดยใช้เทคนิคโมเดลผสมแบบเกาเซียนและอัลกอริทึมการหาค่าสูงสุดของความน่าจะเป็นครั้งหลัง (MAP) ในการเรียนรู้และปรับปรุงค่าความน่าจะเป็นของพารามิเตอร์ต่างๆ ซึ่งข้อดีของวิธีนี้ คือ สามารถรองรับกับกีฬาได้หลากหลายมากขึ้นซึ่งในงานวิจัยส่วนใหญ่จะเน้นที่กีฬาประเภทใดประเภทหนึ่งเพียงอย่างเดียว สำหรับวิธีการตรวจจับสนามโดยพิจารณาจากความน่าจะเป็นของพิกเซลนั้นมีลักษณะแนวคิดคล้ายคลึงกับงานวิจัยของผู้วิจัยที่ใช้พื้นฐานของความน่าจะเป็นมาใช้ช่วยเพิ่มประสิทธิภาพของการค้นหาพิกเซลที่เป็นสนามแต่จะแตกต่างกันตรงที่ในงานวิจัยของผู้วิจัยจะใช้เอกลักษณ์สี เช่น RGB และเอกลักษณ์อื่นๆ อีกจำนวน 6 เอกลักษณ์รวมทั้งหมด 9 เอกลักษณ์ ส่วนของค่าเฉลี่ยและค่าความแปรปรวนจะมีขนาด 9 ค่าและ 9 x 9 เมทริกซ์ตามลำดับซึ่งในงานวิจัยของ Barnard และ Odobez จะใช้เมทริกซ์ของค่าความแปรปรวนแบบทแยงมุม ส่วนการแบ่งภาพไบนารีนั้นจะแบ่งภาพเป็น 2 ส่วนเพื่อใช้ในการหาความน่าจะเป็นแต่ในงานของผู้วิจัยเองจะแบ่งเป็น 3 ส่วนโดยสันนิษฐานว่าบริเวณตอนบนของภาพส่วนใหญ่มักจะเป็นส่วนอื่นๆ ที่อยู่นอกสนาม เช่น สเตเดียมและต้นไม้ ส่วนตอนกลางของภาพนั้นมีโอกาสที่จะเป็นสนามและไม่ใช่สนามเท่าๆ กันและส่วนตอนล่างของภาพน่าจะเป็นพื้นที่ของสนามมากที่สุด ข้อแตกต่างอีกประเด็น คือ ในงานวิจัยของ Barnard และ Odobez จะใช้ชุดข้อมูลภาพที่ใช้สอนโมเดลที่เป็นสนามและโมเดลที่ไม่ใช่สนามแยกชุดกัน คือ ชุดของภาพที่มีสนามและชุดของภาพที่ไม่มีสนามแต่ของผู้วิจัยเองจะใช้ภาพเดียวกันเพียงแต่จะใช้กลุ่มของพิกเซลในภาพนั้นคนละกลุ่มในการสอนโมเดล คือ กลุ่มของพิกเซลที่เป็นสนามและกลุ่มของพิกเซลที่ไม่ใช่สนาม

ในงานวิจัยของ Ngo และคณะ [27] พบว่ามีประเด็นที่น่าจะเป็นประโยชน์ต่องานวิจัยของผู้วิจัย คือ การใช้ค่าสีแบบ YCbCr โดยงานวิจัยนี้ให้เหตุผลว่าค่าสีนี้อ่อนไหวต่อสิ่งรบกวนน้อยกว่าค่าสีแบบ RGB และ HSV ซึ่งผู้วิจัยจะนำไปทดลองกับงานของตนเองเพื่อประเมินความเหมาะสมต่อไป ประเด็นที่สอง คือ ในงานวิจัยนี้มีการเสนอวิธีการตัดพื้นที่ที่เป็นหญ้าแต่ไม่ใช่พื้นที่ของสนามออกด้วยและมีการใช้โมเดลผสมแบบเกาเซียนแบบ 3 คอมโพเนนท์ซึ่งใช้ในการแยกกลุ่มของสนามกลุ่มที่อยู่นอกสนามและกลุ่มที่อยู่ในสนามแต่ไม่ใช่ส่วนของสนามซึ่งผู้วิจัยเห็นว่าวิธีดังกล่าวนี้น่าจะช่วยเพิ่มประสิทธิภาพในการตรวจจับสนามในงานของตนเองได้และประเด็นสุดท้าย คือ หลักการวัดคุณภาพของความแม่นยำในการตรวจจับสนามที่ใช้วิธีการพิจารณาจากอัตราส่วนของจำนวนพิกเซลที่ค้นหาได้ถูกต้องและไม่ถูกต้อง ซึ่งในงานวิจัยนี้จะนำวิธีการวัดคุณภาพของการตรวจจับสนามนี้ไปปรับใช้กับงานวิจัยของตนเองในขั้นตอนของการตรวจจับสนามซึ่งเป็นกระบวนการแรกของงานวิจัย

ในทำนองเดียวกันงานวิจัยของ Liu และคณะ [28] ได้ให้ข้อสังเกตหลายประเด็นซึ่งเป็นประโยชน์ในการใช้เป็นข้อมูลพื้นฐานในการดำเนินงานวิจัยของผู้วิจัยได้เนื่องจากเป็นเนื้อหาที่คล้ายคลึงกัน เช่น ข้อสังเกตเกี่ยวกับเส้นของสนามที่ไม่ชัดเจนและมุมมองของภาพบางมุมมองที่ไม่

จำเป็นต้องพิจารณา คือ มุมมองที่อยู่นอกสนามหรือมุมมองที่มีระยะใกล้ (Close-Up View) ซึ่งมีโอกาสทำให้เกิดข้อผิดพลาดในการตรวจจับสนามได้ ในงานวิจัยของผู้วิจัยจึงจะไม่พิจารณามุมมองภาพที่อยู่นอกสนามนี้ เช่น ภาพที่ถ่ายเฉพาะผู้ชม เนื่องจากเป็นพื้นที่ที่ไม่ถูกพิจารณา ส่วนเอกลักษณ์สี CbCr ที่งานวิจัยนี้ระบุว่าให้ผลที่ดีกว่าเอกลักษณ์สีแบบ HIS ซึ่งสอดคล้องกับงานวิจัยของ Ngo และคณะ [27]

ในงานวิจัยนี้ได้นำเสนอเทคนิคการตรวจจับบริเวณของสนาม (Playfield Regions) ที่ปรากฏอยู่ภายในภาพของเกมส์การแข่งขันฟุตบอลโดยใช้เทคนิคบนพื้นฐานของการจำแนกข้อมูล (Classification-based) ซึ่งกระบวนการจำแนกนี้จะดำเนินการในระดับพิกเซล (Pixel-Level) โดยที่พิกเซลที่อยู่ภายในภาพจะถูกจำแนกออกเป็น 2 กลุ่ม คือ พิกเซลที่เป็นสนามและพิกเซลที่ไม่ใช่สนาม ในการจำแนกนั้นจะใช้ตัวอธิบาย (Descriptor) 2 ตัว ได้แก่ ความหนาแน่นของพิกเซล (Pixel Intensity) และพื้นผิว (Texture) บริเวณรอบข้างของพิกเซล ในการสร้างโมเดลสำหรับการจำแนกข้อมูลในงานวิจัยนี้จะใช้ข้อมูลความน่าจะเป็นก่อนหน้า (Prior Information) เกี่ยวกับตำแหน่งของบริเวณสนามเป็นข้อมูลพื้นฐานในการสร้างโมเดล หลังจากกระบวนการจำแนกข้อมูลเสร็จสิ้นจะใช้วิธีการปรับแต่งค่า (Refinement Method) เพื่อปรับปรุงประสิทธิภาพผลลัพธ์ของการจำแนกอีกครั้ง โดยจะใช้ประโยชน์จากข้อมูลเชิงบริบท (Contextual Information) ของความสัมพันธ์เชิงพื้นที่ (Spatial Relationship) ของพิกเซลที่ปรากฏภายในภาพ ในการวัดประสิทธิภาพของการตรวจจับสนามจะพิจารณาจากอัตราส่วนระหว่างพิกเซลที่ตรวจจับได้ถูกต้อง พิกเซลที่ตรวจจับไม่ถูกต้องและจำนวนพิกเซลทั้งหมด

2. การตรวจจับผู้เล่น (Player Detection)

Ali และคณะ [21] ได้ใช้วิธี Sobel Gradient Method สำหรับตรวจจับผู้เล่นและลูกบอลที่เป็นไปได้ ผลการทดลองพบว่าในส่วนของการค้นหาผู้เล่นนั้นในงานวิจัยนี้ยังมีบางจุดที่บกพร่อง โดยในบางฉากระบบไม่สามารถแสดงผลของผู้เล่นได้ครบทุกคนและมีการแสดงผลในส่วนของกรรมการด้วยและยังเกิดข้อผิดพลาดในการตรวจจับผู้เล่นกรณีที่อยู่ติดกันจำนวนมาก

Huang และคณะ [26] ได้เสนอรูปแบบในการตรวจจับและกำหนดตำแหน่งผู้เล่นโดยใช้วิธีการวิเคราะห์รูปร่าง (Shape Analysis) เพื่อระบุผู้เล่นและลูกบอลจากโฟร์กราวนด์ที่ถูกดึงออกมาแบบคร่าวๆ โดยใช้ตัวค้นหาสนามฟุตบอลโดยใช้พื้นฐานของฮิสโตแกรมของสีและการวิเคราะห์องค์ประกอบที่เชื่อมต่อกัน (Connected Component) ใช้ Euclidean Distance Transform ในการดึงโครงร่าง (Skeletons) จากโฟร์กราวนด์ จากนั้นทำการวิเคราะห์รูปร่างเพื่อขจัดการแจ้งผลที่ผิดพลาด (False Alarms)

Liu และคณะ [33] เสนอวิธีการตรวจจับผู้เล่นหลายคนแบบอัตโนมัติโดยการทำแบ็คกราวนด์ซับแทรคชันบนสีที่สำคัญ (Dominant Color) และตัวตรวจจับที่มีการบูต (Boosting

Detector) ด้วยเอกลักษณ์ฮาร์ (Haar Feature) จากนั้นรวบรวมตัวอย่างผู้เล่น 100 ตัวอย่างด้วยตัวตรวจจับผู้เล่น (Player Detector) และเรียนรู้โมเดลลักษณะการปรากฏของผู้เล่น (Player Appearance Model) บนฐานของโค๊ดบุ๊ก (Codebook) โดยอัลกอริทึมการจัดกลุ่มแบบไม่มีผู้สอน (Unsupervised Clustering Algorithm) ผู้เล่นสามารถถูกกำหนดลาเบลออกเป็น 4 ชนิด คือ 2 ทีม กรรมการ (Referee) หรือ ผู้ที่ไม่เกี่ยวข้อง (Outlier) ความสามารถในการเรียนรู้นี้ทำให้วิธีนี้ใช้กับวิดีโอทั่วไปได้โดยปราศจากการเริ่มต้นด้วยมือ

Han และคณะ [40] เสนอวิธีการตรวจจับวัตถุบนพื้นฐานบนโครงข่ายประสาทเทียม วัตถุที่ตรวจจับ คือ คนเดินเท้า การตรวจจับวัตถุจะใช้วิธีการโมเดลแบ็คกราวนด์บนพื้นฐานของโมเดลผสมแบบเกาเซียนสำหรับสร้างภาพแบบไบนารีโฟร์กราวนด์มาร์ค โมดูลการตรวจจับจะใช้พิกเซลโฟร์กราวนด์นี้เป็นอินพุตและเอาท์พุตความน่าจะเป็นของการตรวจจับวัตถุ โดยค้นหาพิกเซลโฟร์กราวนด์และกำหนดความน่าจะเป็นให้แต่ละตำแหน่งที่มีสเกล (Scale) ที่แน่นอนซึ่งแต่ละบล็อบโฟร์กราวนด์นี้ คือ ภาพของบุคคลที่อาจเกิดขึ้น ส่วนโครงข่ายประสาทเทียมจะใช้ตำแหน่งของพิกเซลในการประมวลผลโดยสร้างคะแนนหรือความน่าจะเป็นที่บ่งบอกถึงความเป็นไปได้ที่บล็อบรอบๆ พิกเซลนั้นจะแสดงถึงคน ศูนย์กลางบริเวณด้านบนของศีรษะจะใช้แสดงตำแหน่งของวัตถุซึ่งแสดงเป็นจุดสีซึ่งเป็นจุดที่มีค่าคะแนนการตรวจจับที่สูงที่สุด

Ali และคณะ [21] เสนออัลกอริทึมในการตรวจจับผู้เล่นและลูกบอลโดยหลังจากขจัดพื้นสนามโดยใช้พื้นฐานของสีแล้วจะใช้วิธีโซเบลเกรเดียนในการสกัดผู้เล่นและลูกบอลและเส้นของสนามและใช้ค่าเทรชโฮลด์ในการขจัดวัตถุที่ไม่ต้องการ ใช้วิธี Dilation and Erosion Methods ในการเชื่อมเส้นสนามที่ไม่ต่อเนื่องกัน เส้นตรงในสนามจะถูกตรวจจับโดยใช้เทคนิค Kernel based Hough Transformation และถูกขจัดออกไป ป้ายบอกคะแนนและผู้ชมจะถูกขจัดออกโดยใช้ข้อมูลของขนาดเป็นเกณฑ์ ทำการทดลองกับวิดีโอการแข่งขันฟุตบอลในรายการ 2002 FIFA World cup ระหว่างทีมบราซิลและเยอรมนี ผลการทดลองพบว่าสามารถตรวจจับลูกบอลที่อยู่ในบริเวณเส้นของสนามได้แต่ในกรณีที่ผู้เล่นอยู่ในบริเวณที่ใกล้กันมากพบว่าระบบไม่สามารถตรวจจับผู้เล่นได้ถูกต้อง เนื่องจากมีขนาดของวัตถุที่ใหญ่ขึ้นทำให้ระบบตรวจจับเป็นผู้เล่นเพียงคนเดียว

Iwase และ Saito [42] ใช้วิธีการลบภาพแบ็คกราวนด์ในการสกัดพื้นที่ผู้เล่นโดยภาพไบนารีจะถูกกระทำโดยการกำหนดค่าเทรชโฮลด์ของค่าความหนาแน่น (Intensity) และพื้นที่ (Area) ส่วนงานวิจัยของ Figueroa และคณะ [59] ใช้วิธีการลบภาพแบ็คกราวนด์ในการเซกเมนต์ผู้เล่น ซึ่งวิธีนี้ส่วนใหญ่ต้องมีการอัพเดทโมเดลแบ็คกราวนด์ งานวิจัยของ C J Needham & R D Boyle [73] ทำการเซกเมนต์ภาพโดยใช้โมเดลสเปซของสี (Color Space Model) โดยโมเดลโฟร์กราวนด์จะใช้สเปซ HIS (HIS Space) งานวิจัยของ Ishii และคณะ [47] ทำการตรวจจับบอลด้วยวิธีการลบเฟรมภาพที่ต่อเนื่องกันซึ่งจะได้พื้นที่ของวัตถุที่เคลื่อนไหว ปัญหาที่พบ คือ จะได้วัตถุจากเฟรมก่อนหน้าด้วย

งานวิจัยของ Hayet และคณะ [49] ทำการสกัดจุดที่สนใจและสร้างโมเดลการกระจายตัวของจุดจากการเรียนรู้ข้อมูลวิดีโอสั้นๆ เพื่อใช้ในการจับคู่กับจุดในภาพปัจจุบันซึ่งระบุว่าสามารถใช้แก้ปัญหาเรื่องการบดบังได้ด้วยรวมถึงการลดคอสต์ของการทำงานได้ งานวิจัยของ Ren และคณะ [46] ใช้วิธีการหาความแตกต่างของภาพในการตรวจจับบอล ส่วน Xu และคณะ [58] ใช้วิธีเดียวกันในการตรวจจับผู้เล่นซึ่งวิธีการหาความแตกต่างของภาพจะต้องมีการอัพเดทภาพเริ่มต้นอย่างต่อเนื่อง Nishino และคณะ [41] ทำการตรวจจับผู้เล่นโดยใช้ตัวจำแนกบนพื้นฐานของ SVM โดยเรียนรู้จากชุดข้อมูลสอน ปัญหาที่พบ คือ กรณีมีผู้เล่นหลายคนเดินผ่านกันจะไม่สามารถตรวจจับได้ถูกต้อง ส่วนในงานของ Swain และ Ballard [60] ได้ประยุกต์ใช้วิธีฮิสโตแกรมแบ็คโปรเจ็คชัน (Histogram Backprojection Method) ในการตรวจจับผู้เล่น เนื่องจากวิธีนี้มีพื้นฐานอยู่บนข้อมูลสี ในกรณีที่ผู้เล่นมีสีของเสื้อผ้าที่คล้ายกับสีของสนามอาจจะเกิดปัญหาการตรวจจับที่ผิดพลาดได้ งานวิจัยของ Hashimoto และ Ozawa [52] ใช้วิธีการลบแบ็คกราวนด์จากภาพอินพุตเพื่อสกัดพื้นที่ผู้เล่นและงานวิจัยของ Spagnolo และคณะ [53] ทำการตรวจจับผู้เล่นโดยใช้วิธีลบแบ็คกราวนด์ เช่น เดียวกัน เป็นต้น

จากการศึกษางานวิจัยข้างต้นสามารถสรุปประเด็นเกี่ยวกับการตรวจจับวัตถุได้ดังนี้ คือ การตรวจจับวัตถุ (Object Detection) ตามที่งานวิจัยของ Mohanty และ Shantaiya [90] ได้อธิบายไว้นั้นสามารถสรุปได้ว่าเป็นประเด็นที่สำคัญที่หลายงานวิจัยให้ความสนใจเนื่องจากสามารถประยุกต์ใช้ข้อมูลดังกล่าวกับโปรแกรมประยุกต์ทางด้านคอมพิวเตอร์วิชันซึ่งมีประโยชน์มากต่อการแยกชุดวิดีโอออกเป็นเนื้อหาที่สำคัญต่างๆ เช่น เนื้อหาของสิ่งที่เคลื่อนไหวและเนื้อหาของแบ็คกราวนด์ นอกจากนี้ยังสามารถนำไปประยุกต์ใช้ในงานทางด้านการรู้จำ (Recognition) การจำแนก (Classification) และการตรวจสอบกิจกรรมต่างๆ (Activity Scrutiny) เป็นต้น นอกจากนี้งานในการตรวจจับวัตถุยังกลายเป็นงานพื้นฐานที่สำคัญต่อการพัฒนาระบบเฝ้าระวังซึ่งในองค์กรสมัยใหม่ได้นำมาประยุกต์ใช้เพื่อควบคุมดูแลความปลอดภัยในองค์กรของตน ดังที่ได้กล่าวไว้ในงานวิจัยของ Risha และ Kumar [91]

เทคนิคหรือวิธีการในปัจจุบันที่นิยมนำมาใช้ในการตรวจจับผู้เล่นตามที่ได้อธิบายไว้ในงานวิจัยของ Mohanty และ Shantaiya [90] งานวิจัยของ Risha และ Kumar [91] งานวิจัยของ Jing และ Chaojian [92] งานวิจัยของ Prajapati และ Galiyawala [93] งานวิจัยของ Dixit และคณะ [94] งานวิจัยของ Parekh และคณะ [95] และงานวิจัยของ Patil และ Bhagat [96] มีหลายวิธียกตัวอย่าง เช่น แบ็คกราวนด์ซับแทรคชัน (Background Subtraction) เฟรมดิฟเฟอเรน (Frame Differencing) และออปติคอลโฟลว์ (Optical Flow) โดยงานวิจัยของ Jing และ Chaojian [92] ได้อธิบายถึงวิธีแบ็คกราวนด์ซับแทรคชันไว้ว่าเป็นวิธีที่ง่าย การคำนวณสำหรับการสร้างแบ็คกราวนด์น้อย ใช้กับงานแบบเรียลไทม์ได้ดีและเหมาะสมกับภาพแบ็คกราวนด์แบบสแตติก (Static Background) แต่จะไม่เหมาะสมกับภาพแบ็คกราวนด์แบบไดนามิก (Dynamic Background) ซึ่งจะ

เกิดการตรวจจับที่ผิดพลาดหรือตรวจจับวัตถุไม่พบได้ (False or Missing Detection) ส่วนวิธีเฟรมดิฟเฟอเรนมีข้อดี คือ สามารถประยุกต์ใช้กับสภาพแวดล้อมแบบไดนามิกได้ดี (Dynamic Environment) เป็นอัลกอริทึมที่ง่าย เหมาะสมกับระบบแบบเรียลไทม์ ไม่ต้องการการอัพเดทข้อมูลแบ็คกราวนด์ ข้อเสีย คือ วิธีนี้จะสกัดได้เพียงประวัติ (Profile) ของวัตถุที่เคลื่อนที่ ถ้าระยะทางในการเคลื่อนที่มีระยะทางที่น้อยจะส่งผลต่อความผิดพลาดในการตรวจจับโดยจะเกิดหลุม (Hole) ขนาดใหญ่ได้และถ้าวัตถุเคลื่อนที่ช้าหรืออยู่กับที่จะส่งผลให้ไม่สามารถตรวจจับวัตถุนั้นพบดังนั้นวิธีนี้จึงต้องกำหนดช่วงของเฟรม (Interval) ที่เหมาะสม นอกจากนี้งานวิจัยของ Jing และ Chaojian [92] ยังสรุปไว้ว่าวิธีแบ็คกราวนด์ซับแทรคชันยังให้ผลลัพธ์ที่ดีกว่าวิธีเฟรมดิฟเฟอเรนในกรณีที่เป็นการตรวจจับวัตถุในหลายวิดีโอ (Multi-Video) แบบเรียลไทม์ ในขณะที่งานวิจัยของ Mohanty และ Shantaiya [90] ได้อธิบายถึงข้อดีและข้อเสียของวิธีออปติคอลโฟลร์ไว้ว่าเป็นวิธีที่ให้ความถูกต้องได้ดีแต่มีความซับซ้อนที่สูงและค่อนข้างไม่เหมาะสมในการนำไปปรับใช้กับระบบงานแบบเรียลไทม์ ส่วนวิธีเทมโพลอลดิฟเฟอเรน (Temporal Differencing Method) นั้นมีความซับซ้อนในการคำนวณน้อยและสามารถปรับใช้ได้กับวิดีโอที่มีการเปลี่ยนแปลงแบบไดนามิก (Dynamic Variation) แต่จะมีความอ่อนไหวได้ง่ายต่อค่าเทรชโฮลด์ เมื่อมีการกำหนดการเปลี่ยนแปลงภายในวิดีโอเฟรมที่แตกต่างกันและต้องใช้อัลกอริทึมในการสนับสนุนที่แตกต่างกันในการตรวจจับวัตถุที่อยู่คงที่ (Stationary Object) สำหรับวิธีแบ็คกราวนด์ซับแทรคชันนั้นเป็นวิธีที่ทำงานได้เร็วและมีความแม่นยำแต่ยังคงอ่อนแอต่อการเปลี่ยนแปลงของแสงและภาพแบ็คกราวนด์สำหรับภาพที่มีการจัดการกับค่าความหนาแน่น (Intensity Value) ที่มีระดับที่สูง (Large Degree) ซึ่งจะส่งผลต่อประสิทธิภาพโดยรวม สุดท้ายในงานวิจัยของ Risha และ Kumar [91] ได้อธิบายความสามารถของเทคนิคในการตรวจจับวัตถุไว้ดังนี้ วิธีแบ็คกราวนด์ซับแทรคชันแบบไม่มีการวนซ้ำ (Non-Recursive Techniques) มีความสามารถปรับตัวได้สูง (Highly Adaptive) ไม่ต้องอาศัยข้อมูลประวัติของเฟรมที่ถูกจัดเก็บไว้ในบัฟเฟอร์แต่ไม่เหมาะสำหรับภาพแบ็คกราวนด์แบบไดนามิก และภาพที่มีการเปลี่ยนแปลงของแสงหรือภาพที่มีเงาเป็นต้น ส่วนวิธีเฟรมดิฟเฟอเรนจะมีข้อจำกัดเรื่องการได้ข้อมูลคอนทัวร์ที่สมบูรณ์ของวัตถุ วิธีเทมโพลอลดิฟเฟอเรน จะมีความยืดหยุ่นต่อการเปลี่ยนแปลงแบบไดนามิกภายในฉากแต่วัตถุที่เคลื่อนที่ช้าจะไม่สามารถตรวจจับได้ พื้นที่ที่ติดกัน (Converging Region) จะถูกมองเป็นวัตถุที่เคลื่อนที่ เอาท์พุตจะมีสัญญาณรบกวน เนื่องจากความผิดพลาดจากการประมาณการเคลื่อนไหว สัญญาณรบกวนในพิกเซลแบ็คกราวนด์จะถูกพิจารณาเป็นพิกเซลที่เคลื่อนไหวซึ่งความผิดพลาดที่สูงส่งผลให้เอาท์พุตมีสัญญาณรบกวนที่มากขึ้นด้วยและวิธีออปติคอลโฟลว์ซึ่งวิธีนี้จะได้ความรู้ (Knowledge) ที่สมบูรณ์เกี่ยวกับการเคลื่อนที่ของวัตถุและเป็นข้อมูลที่เป็นประโยชน์ต่อการคำนวณเป้าหมายที่เคลื่อนที่ออกจากแบ็คกราวนด์ ส่วนข้อเสียของวิธีนี้ คือ ต้องการจำนวนของข้อมูลในการคำนวณค่อนข้างมาก

จากการศึกษาเทคนิคดังกล่าวข้างต้นในงานวิจัยนี้ได้ประยุกต์ใช้เทคนิคการตรวจจับผู้เล่นโดยประยุกต์ใช้เทคนิคการตรวจจับบนพื้นฐานของวิธีแบ็คกราวนด์ซับแทรคชันเนื่องจากเป็นวิธีที่ง่าย มีการคำนวณน้อยและเหมาะสมกับระบบงานที่มีภาพแบ็คกราวนด์เป็นแบบสแตติกและเป็นแบบเรียลไทม์และสามารถทำงานได้กับภาพวิดีโอหลายชุดจึงมีความเหมาะสมกับระบบงานในงานวิจัยนี้ซึ่งมีการดำเนินการกับชุดภาพวิดีโอหลายชุดพร้อมๆ กัน ภาพแบ็คกราวนด์เป็นแบบสแตติกและในบางสถานการณ์ผู้เล่นอาจมีการหยุดอยู่กับที่ซึ่งวิธีแบบออปติคอลโฟลว์จะเกิดปัญหาเกี่ยวกับการปรากฏหลุมจากการตรวจจับได้ง่าย นอกจากนี้ในกระบวนการของการตรวจจับสนามของงานวิจัยนี้ได้ประยุกต์ใช้โมเดลผสมแบบเกาเซียนในการจำแนกสนามด้วยซึ่งทำให้ง่ายต่อการดำเนินการในกระบวนการตรวจจับผู้เล่น

3. การติดตามผู้เล่นในแบบมุมมองภาพ (Single View Player Tracking)

Beyan และ Temizel [2] นำเสนอเทคนิคสำหรับการติดตามหลายวัตถุแบบอัตโนมัติโดยใช้อัลกอริทึมมีนชีฟท์ (Mean-Shift Algorithms) ที่มีการปรับปรุงเทคนิคเพิ่มเติมเพื่อให้การติดตามนั้นมีประสิทธิภาพเพิ่มมากขึ้นโดยข้อมูลที่ใช้ในการทดลอง คือ วิดีโอที่มีคนเดินผ่านในรูปแบบต่างๆ ผลการทดลองพบว่าระบบที่พัฒนามีประสิทธิภาพดีกว่าแบบเก่า คือ ระบบที่ใช้วิธีอัลกอริทึมมีนชีฟท์แบบเดิมในการติดตาม

Liu และคณะ [22] ใช้ตัวกรองคาร์ลมาล (Kalman Filter) ในการติดตามลูกบอล งานวิจัยของ Ishii และคณะ [47] ใช้ตัวกรองคาร์ลมาลในการประมาณตำแหน่งลูกบอลแบบ 3D ในระบบแบบ 2 กล้อง งานวิจัยของ Ren และคณะ [62] ใช้ตัวกรองคาร์ลมาลในการกรอง มีชัวร์เมนท์ (Measurement) และแยกวัตถุที่ถูกผสานเข้าด้วยกัน (Merged) สำหรับการติดตามลูกบอลในระบบแบบหลายกล้อง งานวิจัยของ Xu และคณะ [58] ใช้ตัวกรองคาร์ลมาลในการติดตามผู้เล่นในระบบแบบหลายกล้อง งานวิจัยของ Liu และคณะ [33] ทำการติดตามผู้เล่นหลายคนด้วยความสัมพันธ์ของข้อมูลแบบมาร์คอฟเชนมอนท์คาร์โล (Markov Chain Monte Carlo (MCMC) Data Association) งานวิจัยของ Needham และ Boyle [73] ใช้วิธีอัลกอริทึมคอนเดนเซชันในการติดตามและใช้ตัวกรองคาร์ลมาลในการประมาณตำแหน่งผู้เล่นและแก้ปัญหากรณีวัตถุอยู่ไกลจากกลุ่มเดียวกันในขั้นตอนการทำนายของอัลกอริทึมคอนเดนเซชันซึ่งวิธีอัลกอริทึมคอนเดนเซชันสามารถใช้ติดตามวัตถุได้หลายวัตถุพร้อมๆ กัน งานวิจัยของ Nishino และคณะ [41] ใช้พาติเคิลฟิลด์เตอร์แบบ 3D ในการติดตามผู้เล่นในวิดีโอแบบกล้องตาเดียว (Monocular) บนพื้นฐานของตัวตรวจจับที่มีความเชื่อมั่น งานวิจัยของ Hayet และคณะ [49] ใช้เอกลักษณ์บนพื้นฐานของโมเดลการกระจายตัวของจุดสำหรับการติดตามผู้เล่นในระบบแบบหลายกล้องและมีบางกล้องที่สามารถซูมหรือแพนได้ ข้อดีของวิธีนี้ คือ สามารถติดตามวัตถุผ่านการบดบังแบบบางส่วนได้แม้จะเป็นทีมเดียวกัน ส่วนข้อจำกัดของวิธีนี้ คือ การเปลี่ยนแปลงขนาดของจุดที่สนใจซึ่งอาจเกิดจากการย้ายหรือการซูมเนื่องจากวิธีนี้มีพื้นฐานบนจุด

งานวิจัยของ Choi และคณะ [45] ใช้พาติเคิลฟิลด์เตอร์ในการประมาณเส้นทางของลูกบอล งานวิจัยของ Sato และ Aggarwal [50] เสนอวิธีการติดตามนักฟุตบอลโดยใช้การแปลงเชิงพื้นที่-ความเร็วแบบชั่วคราว (Temporal Spatio-Velocity (TSV) Transform) งานวิจัยนี้ระบุว่าการแปลง TSV แบบโลคอลมีข้อดี คือ เนื่องจากมันจำกัดแรงค์ของการคำนวณรอบบริเวณของวัตถุที่ถูกติดตามทำให้มีข้อดี คือ การปรับปรุงประสิทธิภาพในการคำนวณและการติดตามโดยตรงโดยปราศจากกระบวนการในการเชื่อมโยงบล็อบ (Blob-Association) ซึ่งการแก้ไขนี้จะช่วยให้สามารถทำการติดตามได้มั่นคงในสถานการณ์ที่มีการบดบัง งานวิจัยนี้ได้ระบุถึงจุดเด่นของระบบที่เสนอไว้ 3 ประเด็น คือ ประเด็นแรกการติดตามที่มั่นคงในสถานการณ์ที่มีการบดบังนั้นมีความเป็นไปได้โดยใช้การแปลง TSV บนภาพแบบเกรย์สเกลและภาพสี ประเด็นที่สอง คือ การติดตามโดยปราศจากการเชื่อมโยงบล็อบซึ่งถูกจัดการโดยการแปลง TSV แบบโลคอลและประเด็นสุดท้าย คือ โหลดในการคำนวณที่ลดลงเนื่องจากการแปลง TSV แบบโลคอล งานวิจัยของ Mazzeo และคณะ [54] เสนอวิธีการติดตามผู้เล่นในการแข่งขันฟุตบอลโดยใช้วิธีการสุ่ม (Stochastic Approach) บนพื้นฐานของการประเมินความน่าจะเป็นครั้งหลังสูงสุด (Maximum a Posteriori Probability (MAP)) โดยจะทำการประเมินข้อมูลทางเรขาคณิตบนบล็อบที่เหลื่อมซ้อนกันจากนั้นใช้การจำแนกเอกลักษณ์สีเพื่อติดตามผู้เล่นและแก้ปัญหาในสถานการณ์ของการผสานบล็อบ งานวิจัยของ Han และคณะ [40] เสนอวิธีแบบหลายสมมติฐาน (Multiple Hypothesis Method) ในการติดตามวัตถุจำนวนหลายวัตถุบนพื้นฐานของการตรวจจับวัตถุ รวบรวมผลการตรวจจับไว้ในรูปของโครงสร้างกราฟและเก็บสมมติฐานของเส้นทางของแต่ละวัตถุไว้ขณะเดียวกันจะให้ผลตอบกลับ (Feedbacks) ซึ่งเป็นการทำนายตำแหน่งของวัตถุไปยังโมดูลการตรวจจับ และงานวิจัยของ Figueroa และคณะ [59] ใช้ข้อมูลจากโมเดลกราฟในการติดตามผู้เล่นในระบบแบบหลายกล้องโดยระบุถึงข้อดีของการใช้โมเดลกราฟ คือ ดำเนินการได้ง่าย สามารถใช้แก้ไขปัญหาการบดบังเบื้องต้นได้รวมถึงใช้ในการแยกบล็อบที่ผสานเข้าด้วยกันได้ การติดตามกระทำโดยค้นหาเส้นทางที่ดีที่สุดในกราฟ เป็นต้น

จากการศึกษางานวิจัยดังกล่าวข้างต้นสามารถสรุปประเด็นเกี่ยวกับการติดตามผู้เล่นได้ดังนี้ คือ การติดตามผู้เล่น (Player Tracking) ในวิดีโอการแข่งขันกีฬาฟุตบอลเป็นกระบวนการที่สำคัญซึ่งจะเป็นข้อมูลพื้นฐานต่อกระบวนการวิเคราะห์ความหมายในระดับที่สูงกว่าอื่นๆ เช่น การวิเคราะห์กลยุทธ์การแข่งขัน (Strategy Analysis) การกู้คืนฉาก (Scene Recovery) การทำดัชนีฉากแบบอัตโนมัติ (Automatic Indexing of Soccer Scene) และการคัดเลือกกีฬาที่ชาญฉลาดแบบอัตโนมัติ (Automatic Intelligent Sports Casting) ดังที่กล่าวไว้ในงานวิจัยของ Iwase และ Saito [42] ส่วนในงานวิจัยของ Yamada และคณะ [98] ได้กล่าวไว้ในทำนองเดียวกันว่าข้อมูลการเคลื่อนที่ของผู้เล่นเป็นข้อมูลที่จำเป็นต่อกระบวนการรู้จำ (Recognize) และการค้นคืน (Retrieve) ฉากในเกมส์การแข่งขันฟุตบอล นอกจากนี้ในงานวิจัยของ Sterning และคณะ [99] ยังกล่าวถึงความสำคัญ

ของกระบวนการติดตามผู้เล่นที่คล้ายกับงานวิจัยของ Iwase และ Saito [42] ไว้ว่าการติดตามวัตถุเป็นงานที่สำคัญในงานทางด้านคอมพิวเตอร์วิชันและถูกใช้เป็นขั้นตอนแรกสำหรับการวิเคราะห์วิดีโอเกี่ยวกับการเฝ้าระวัง (Surveillance) กีฬาและโปรแกรมทางด้านอุตสาหกรรม (Industrial Applications) ส่วนงานวิจัยของ Prajapati และ Galiyawala [93] ได้อธิบายในทำนองเดียวกันว่าการตรวจจับและติดตามวัตถุกลายเป็นหัวข้อที่สำคัญในงานวิจัยทางด้านการประมวลผลภาพและคอมพิวเตอร์วิชันซึ่งมีการประยุกต์ใช้ในโปรแกรมประยุกต์ต่างๆ เช่น ระบบการเฝ้าระวังที่มีการติดตามวัตถุแบบเรียลไทม์และระบบการตรวจสอบการจราจร (Traffic Monitoring) เป็นต้น

เทคนิคพื้นฐานในการติดตามวัตถุมีหลายวิธี เช่น การติดตามบนพื้นฐานของจุด (Point –based Tracking) การติดตามบนพื้นฐานของเคอร์เนล (Kernel-based Tracking) และการติดตามบนพื้นฐานของภาพเงา (Silhouette-based Tracking) ซึ่งได้อธิบายไว้ในงานวิจัยของ Prajapati และ Galiyawala [93] งานวิจัยของ Nevhal และ Ghotkar [100] งานวิจัยของ Sindhuja และ Devi [97] งานวิจัยของ Dixit และคณะ [94] งานวิจัยของ Parekh และคณะ [95] และงานวิจัยของ Patil และ Bhagat [96] เป็นต้น ตัวอย่างวิธีการติดตามวัตถุบนพื้นฐานของเทคนิคการติดตามแบบจุด เช่น ตัวกรองคาร์ลมาลฟิลเตอร์ ตัวกรองอนุภาคและการติดตามแบบหลายสมมติฐาน (Multiple Hypothesis Tracking) วิธีการติดตามวัตถุบนพื้นฐานของเทคนิคการติดตามแบบเคอร์เนล เช่น ซิมเปิลเทมเพลตแมทชิ่ง (Simple Template Matching) อัลกอริทึมมีนชีฟท์ อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) และการติดตามบนพื้นฐานของเลเยอร์ (Layering Based Tracking) และวิธีการติดตามบนพื้นฐานของเทคนิคการติดตามแบบภาพเงา เช่น คอนทัวร์แมทชิง (Contour Matching) และเชฟแมทชิง (Shape Matching) เป็นต้น นอกจากนี้ยังมีเทคนิคบนพื้นฐานของการสุ่มเชิงปัจจัย (Factored Sampling) เช่น อัลกอริทึมคอนเดนเซชันซึ่งถูกนำเสนอโดย Isard และ Blake [65] โดยถูกนำมาประยุกต์ใช้ในงานด้านการติดตามวัตถุซึ่งวิธีดังกล่าวนี้มีความแข็งแรงต่อปัญหาการบดบังและการเปลี่ยนแปลงรูปร่างของวัตถุ

จากการศึกษาข้อมูลในเบื้องต้นในงานวิจัยนี้ได้ประยุกต์ใช้เทคนิคพื้นฐานทั่วไปในการติดตามวัตถุโดยเลือกใช้เทคนิคการติดตามบนพื้นฐานของจุดเนื่องจากเป็นการประมวลผลในระดับโลคอล (Local) ซึ่งจะใช้ข้อมูลในการคำนวณน้อยลงเมื่อเทียบกับแบบโกลบอล (Global) ซึ่งจะต้องมีการสร้างโมเดลต่างๆ ในระดับที่สูงขึ้นที่เกี่ยวข้องเพิ่มเติม อีกทั้งเทคนิคการติดตามบนพื้นฐานของจุดเป็นวิธีที่ง่าย นอกจากนี้วิธีที่มีพื้นฐานของการติดตามแบบจุดนี้โดยส่วนใหญ่ไม่ต้องการการกำหนดโมเดลและกระบวนการสอนระบบมากนักเมื่อเทียบกับเทคนิคบนพื้นฐานของเคอร์เนลและเทคนิคบนพื้นฐานของภาพเงา ในการเพิ่มความแข็งแรงต่อกระบวนการตรวจจับ การติดตามและปัญหาการบดบังในเบื้องต้นในงานวิจัยนี้ทำการประยุกต์ใช้วิธีการสุ่มแซมเปิลบนพื้นฐานของพื้นที่ร่วมกับอัลกอริทึมคอนเดนเซชัน

4. การจำแนกตัวตน (Identification)

การจำแนกหรือระบุตัวตนของผู้เล่น โดยทั่วไปมักใช้ข้อมูลเอกลักษณ์ที่มีอยู่ของวัตถุ เช่น ข้อมูลสี ค่าความหนาแน่น เป็นพื้นฐานในการทำงาน เช่น ในงานวิจัยของ Liu และคณะ [33] ทำการเรียนรู้โมเดลลักษณะการปรากฏของผู้เล่น (Player Appearance Model) บนฐานของโค๊ดบุ๊ก (Codebook) โดยอัลกอริทึมการจัดกลุ่มแบบไม่มีผู้สอน (Unsupervised Clustering Algorithm) ผู้เล่นสามารถถูกกำหนดลาเบลออกเป็น 4 ชนิด คือ ผู้เล่น 2 ทีม กรรมการ (Referee) และผู้ที่ไม่เกี่ยวข้อง (Outlier) งานวิจัยของ Choi และคณะ [56] ใช้ข้อมูลสีของชุดเสื้อผ้าของผู้เล่นในการแยกแยะตัวตนของแต่ละทีมโดยใช้วิธีฮิสโตแกรมแบ็คโปรเจ็คชันในการระบุทีมผู้เล่น แต่ในกรณีที่กล้องนั้นติดตั้งอยู่ไกลทำให้ข้อมูลสีไม่ชัดเจนจึงยากต่อการแยกแยะสี งานวิจัยของ Hashimoto และ Ozawa [52] ทำการจำแนกพื้นที่ผู้เล่นที่ไม่ถูกบดบังเป็น 5 กลุ่ม คือ ผู้เล่น 2 ทีม ผู้รักษาประตูและกรรมการ จากนั้นเปรียบเทียบระยะทาง (Distance) ระหว่างค่าเฉลี่ยของสถิติสี (R, G, B, H, S, L) ของพื้นที่ผู้เล่นกับแต่ละกลุ่มนั้น การวัดระยะทางของสีจะใช้วิธีของมหาลาโนบิส (Mahalanobis Distance) โดยพื้นที่ที่มีค่าสีนั้นจะเป็นพื้นที่ของเสื้อบนพื้นฐานของโมเดลมนุษย์ งานวิจัยของ Mazzeo และคณะ [54] ใช้เอกลักษณ์ค่าฮิสโตแกรมของสี RGB ในการจัดกลุ่มและใช้วิธีวัดระยะทางแบบแมนฮัทตันในการประมาณความใกล้ชิด (Proximity) ของสองฮิสโตแกรมซึ่งวิธีแมนฮัทตันนี้มีความง่ายและเร็วกว่าวิธีอื่น เช่น ยูคลีเดียน [53] งานวิจัยของ Figuerou และคณะ [59] จึงแก้ปัญหาโดยใช้ข้อมูลความหนาแน่น (Intensity) ของสีของชุดเสื้อผ้าผู้เล่น

กระบวนการจำแนกโดยอาศัยการกระจายตัวของความหนาแน่นของสี ในงานวิจัยของ Figuerou และคณะ [59] อธิบายไว้ดังนี้ โดยทั่วไปผู้เล่นสามารถโมเดลได้เป็นกลุ่มของพื้นที่ซึ่งแต่ละพื้นที่จะมีสีที่เด่นปรากฏอยู่ ในงานวิจัยนี้จึงแบ่งโมเดลของผู้เล่นออกเป็น 2 พื้นที่หรือมากกว่า แต่ละพื้นที่จะแสดงถึงส่วนต่างๆ ของชุดผู้เล่น เช่น เสื้อ กางเกง และถุงเท้า เป็นต้น แต่ละพื้นที่จะทำการนิยามตัวกรองบนพื้นฐานของการกระจายตัวของความหนาแน่นในแนวตั้ง (Vertical Intensity Distribution) ของบล็อบซึ่งจะได้ค่าที่น้อยสุด (T1) และมากสุด (T2) เป็นค่าเฉลี่ยของการกระจายตัวในแนวตั้งซึ่ง 2 ค่านี้จะถูกใช้เป็นเกณฑ์ในการคำนวณค่าความหนาแน่นของแต่ละพื้นที่ ในแต่ละพื้นที่ R_i จะนับจำนวนพิกเซล p ของบล็อบที่อยู่นอกเกณฑ์ของค่าเทรชโฮลด์ คือ T1 และ T2 ซึ่งนิยามได้ดังสมการที่ 2.32

$$S1_{R_i} = \#\{p\} \forall p \in R_i \wedge R_i(p) < T1$$
$$S2_{R_i} = \#\{p\} \forall p \in R_i \wedge R_i(p) < T2 \quad (2.32)$$

โดยทั้ง 2 ค่านี้จะเกี่ยวข้องกับค่าความหนาแน่นของแต่ละพื้นที่บนพื้นฐานของการกระจายตัวในแนวตั้งของบล็อบ ดังนั้นสีหรือความหนาแน่นของแต่ละพื้นที่ R_i จะถูกนิยามได้ดังสมการที่ 2.33

$$C_{R_i} = \begin{cases} 1, if \dfrac{S1_{R_i}}{S1_{R_i} + S2_{R_i}} > 0.6 \\ 2, if \dfrac{S1_{R_i}}{S1_{R_i} + S2_{R_i}} < 0.4 \\ 0, otherwise \end{cases} \qquad (2.33)$$

โดยที่สีของบล็อบที่ใช้แสดงตัวตนของทีมสามารถนิยามได้จากข้อมูลสีของพื้นที่หรือการรวมสีของหลายๆ พื้นที่เข้าด้วยกัน

5. ประเด็นการบดบัง (Occlusion Issues)

Sato และ Aggarwal [50] ใช้การแปลง TSV แบบโลคอล (Local TSV Transform) ซึ่งระบุว่าการแปลง TSV แบบโลคอลมีข้อดี คือ เนื่องจากมันจำกัดแรงค์ของการคำนวณรอบบริเวณของวัตถุที่ถูกติดตามทำให้มีข้อดี คือ การปรับปรุงประสิทธิภาพในการคำนวณและการติดตามโดยตรงโดยปราศจากกระบวนการในการเชื่อมโยงบล็อบ (Blob-Association) ซึ่งการแก้ไขนี้จะช่วยให้สามารถทำการติดตามได้มั่นคงในสถานการณ์ที่มีการบดบัง

Hashimoto และ Ozawa [52] ได้กล่าวไว้ว่า ถ้าพื้นที่ที่มีอยู่นั้นมีอยู่ในกล้องอย่างน้อย 1 กล้องการติดตามจะมีความเป็นได้โดยการใช้หลายกล้องในการทำงานจะช่วยให้มีความแข็งแรงต่อการบดบังมากกว่าการใช้กล้องเพียงตัวเดียว

Figuerou และคณะ [59] แก้ปัญหาการบดบังโดยใช้ข้อมูลจากกราฟโดยเน้นการบังแบบ 2 คนเป็นหลักซึ่งจะพิจารณาจากเส้นทางของวัตถุว่าเป็นเส้นทางเดียวกันหรือไม่ ในการรักษาเส้นทางที่ถูกต้องระหว่างการติดตามเมื่อผู้เล่นแยกออกจากกันจะพิจารณาจากข้อมูลสีและระยะทางระหว่างบล็อบ (Blob) จากกราฟ ส่วนระหว่างการบดบังจะพิจารณาจากทิศทางของเส้นทางวัตถุเพื่อคำนวณหาเส้นทางที่ถูกต้องโดยในการโฟกัสผู้เล่นจากกล้องที่เห็นชัดเจนที่สุดจะพิจารณาจากข้อมูล FOV และพารามิเตอร์ของกล้องซึ่งจะรู้ได้ว่าผู้เล่นเห็นได้ชัดเจนในกล้องใด การแยกบล็อบจะใช้ข้อมูลสีโดยใช้ตัวดำเนินการทางสัณฐานวิทยา (Morphological Operator)

Xu และคณะ [58] แก้ปัญหาการบดบังโดยพิจารณาจากความสูงและความกว้างของวัตถุซึ่งสันนิษฐานว่ามีขนาดที่คงที่โดยพิจารณาจากบราวดิ้งเอดจ์ (Bounding Edge) ที่เหลือที่ยัง

สามารถใช้ได้ในการคำนวณโดยการประมาณจะถูกอัพเดทโดยใช้ มีชัวร์เมนท์ บางส่วนเมื่อไรก็ตามที่มันสามารถใช้งานได้ซึ่งวิธีนี้ให้ผลที่ถูกต้องและแข็งแรงกว่ากรณีใช้วิธีการทำนายเพียงอย่างเดียว

จากการศึกษางานวิจัยดังกล่าวข้างต้นสามารถสรุปประเด็นเกี่ยวกับการติดตามผู้เล่นได้ดังนี้ คือ ปัญหาที่สำคัญที่พบได้บ่อยในงานทางด้านการติดตามวัตถุ คือ การบดบัง การเปลี่ยนแปลงของแสง การเปลี่ยนแปลงของสภาพแวดล้อม การเปลี่ยนแปลงรูปร่างและการเคลื่อนที่ของวัตถุที่รวดเร็วและสัญญาณรบกวนซึ่งปัญหาเหล่านี้ที่ถูกกล่าวถึงอยู่ในหลายงานวิจัย เช่น งานวิจัยของ Martin และ Martínez [101] งานวิจัยของ Patil และ Bhagat [96] และงานวิจัยของ Nagalakshmi และคณะ [102] เป็นต้น โดยเฉพาะปัญหาการบดบังซึ่งมีโอกาสเกิดขึ้นได้บ่อยในวิดีโอการแข่งขันกีฬาประเภททีม เช่น บาสเก็ตบอลและฟุตบอล เป็นต้น ปัญหาการบดบังจึงถูกกล่าวถึงในหลายงานวิจัย เช่น งานวิจัยของ Martin และ Martínez [101] ซึ่งได้กล่าวไว้ว่าการบดบังนั้นเป็นปัญหาต่อกระบวนการติดตาม เช่น การสูญหายของผู้เล่น (Losing Player) หรือเกิดการผสมบล็อบจากบล็อบของผู้เล่นหลายคนรวมถึงเกิดความสับสนในการติดตามในกรณีที่ผู้เล่นใส่เสื้อผ้าที่มีสีคล้ายกัน เป็นต้น

ปัญหาการบดบัง เป็นปัญหาที่พบได้บ่อยในเกมส์การแข่งขันของกีฬาประเภททีม การบดบังสามารถแบ่งออกเป็น 2 ประเภท คือ การบดบังเพียงบางส่วนและการบดบังแบบสมบูรณ์ ซึ่งแนวทางแก้ปัญหามีผู้เสนอไว้หลากหลายวิธี เช่น งานวิจัยของ Han และคณะ [40] ทำการรวบรวมผลการตรวจจับไว้ในรูปของโครงสร้างกราฟและเก็บสมมติฐานของเส้นทางของแต่ละวัตถุไว้ซึ่งระบุว่าวิธีการติดตามที่เสนอสามารถจัดการกับปัญหาการบดบังได้ งานวิจัยของ Hayet และคณะ [49] ใช้ข้อดีของการติดตามบนพื้นฐานของเอกลักษณ์ทางโลคอลซึ่งการใช้ลักษณะการปรากฏทางโลคอล (Local Appearance) สามารถช่วยแก้ปัญหาการบดบังได้ซึ่งในกรณีที่ใช้ลักษณะเฉพาะทางโกลบอล (Global Characterization) อาจมีปัญหากับการบดบังบางส่วนได้ วิธีการ คือ การจับคู่จุดของโมเดลการกระจายตัวของจุดกับจุดในภาพและการปรับพารามิเตอร์โมเดล คุณสมบัตินี้ช่วยให้ตัวติดตามมีความแข็งแรงต่อการบดบังแบบบางส่วนได้

ในปัจจุบันกรณีของการใช้ข้อมูลจากกล้องเพียงตัวเดียวยังไม่สามารถแก้ปัญหาการบดบังนี้ได้จึงมีเทคนิคหนึ่งที่หลายงานวิจัยให้ความสนใจ คือ การอาศัยข้อมูลจากกล้องหลายตัวเพื่อแก้ปัญหาการบดบัง ดังเช่น งานวิจัยของ Iwase และ Saito [42] ทำการติดตามผู้เล่นฟุตบอลโดยการรวมตำแหน่งที่ตรวจจับได้ในหลายมุมมองซึ่งผลการติดตามมีความแข็งแรงต่อการบดบังแม้ในบางภาพจะเกิดการบดบังเนื่องจากในภาพจากกล้องอื่นสามารถตรวจจับตำแหน่งผู้เล่นได้ งานวิจัยของ Park และคณะ [44] แก้ปัญหาการบดบังโดยถ้าบล็อบผู้เล่นจากอินพุตของมุมมองหนึ่งไม่ชัดเจนจะทำการเช็คว่ามีมุมมองอื่นที่ชัดเจนหรือไม่ซึ่งการเช็คนี้สามารถใช้วิธีการแปลงโฮกราฟฟีแล้วใช้ข้อมูลจากมุมมองที่มองเห็นได้ชัดเจน งานวิจัยของ Martin และ Martínez [101] ที่เสนอวิธีตรวจจับและติดตามผู้เล่นโดยใช้ข้อมูลจากหลายกล้องและเสนอวิธีการผสมแทร็ค (Tracks) ของบล็อบที่ถูก

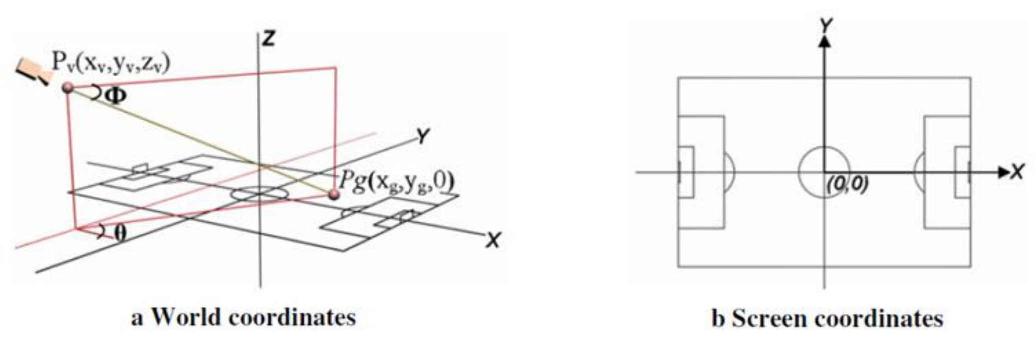
ตรวจจับเพื่อทำการจับคู่แทร็คข้ามกล้อง งานวิจัยของ Ren และคณะ [103] เสนอเทคนิคการประมาณเส้นทางของลูกบอลโดยใช้ข้อมูลจากหลายกล้องและใช้เทคนิคการติดตามบนพื้นฐานของตัวกรองคาร์ลมาลฟิลเตอร์ นอกจากนี้ยังมีกลุ่มที่ประยุกต์ใช้วิธีการสุ่มแซมเปิลเพื่อเพิ่มความแข็งแรงต่อกระบวนการตรวจจับ การติดตามและปัญหาการบดบังโดยมีพื้นฐานอยู่บนอัลกอริทึมคอนเดนเซชัน เช่น งานวิจัยของ Isard และ Blake [65] และงานวิจัยของ Needham และ Boyle [73] เป็นต้น

บางงานวิจัยได้ประยุกต์ใช้บราวดิ้งบล็อก (Bounding Box) ในการติดตามและแก้ไขปัญหาการบดบังในเบื้องต้น ยกตัวอย่างงานวิจัย เช่น งานวิจัยของ Xu และคณะ [58] ได้ประยุกต์ใช้ข้อมูลของบราวดิ้งบล็อกเป็นอินพุตของตัวติดตามตัวกรองคาร์ลมาลฟิลเตอร์และใช้บราวดิ้งเอดจ์ในการประมาณตำแหน่งในกรณีที่มีการบดบังแทนการใช้วิธีการทำนายซึ่งมีความน่าเชื่อถือน้อยและงานวิจัยของ Ren และคณะ [103] ที่ประยุกต์ใช้ข้อมูลบราวดิ้งบล็อกเป็นอินพุตของตัวติดตามตัวกรองคาร์ลมาลฟิลเตอร์โดยได้อธิบายไว้ว่า กรณีที่บราวดิ้งบล็อกของลูกบอลซึ่งมีขนาดเล็กถูกซ้อนทับโดยบราวดิ้งบล็อกของผู้เล่นซึ่งมีขนาดใหญ่กว่าจะถูกพิจารณาว่าลูกบอลนั้นหายไป (Missing) เป็นต้น การใช้บราวดิ้งบล็อกนั้นมีข้อดี คือ เป็นวิธีที่ง่ายและเร็ว [18] และใช้ข้อมูลในการประมวลผลที่ค่อนข้างน้อยแต่วิธีนี้จะเหมาะสมกับกรณีที่เป็นการบดบังเพียงบางส่วนเท่านั้นเพราะเมื่อวัตถุถูกบดบังแบบสมบูรณ์โดยปกติบราวดิ้งบล็อกก็จะถูกบังแบบสมบูรณ์ด้วยโดยเฉพาะระบบที่มีการกำหนดขนาดของบราวดิ้งบล็อกแบบคงที่ นอกจากนี้ในกรณีที่เกิดการรวมกลุ่มของผู้เล่นมากขึ้นบราวดิ้งบล็อกก็จะซ้อนทับกันมากขึ้นส่งผลให้กระบวนการจำแนกตัวตนเป็นไปได้ยากมากขึ้นและจากงานวิจัยของ Ren และคณะ [103] มีข้อสังเกต คือ ในกรณีที่บราวดิ้งบล็อกของลูกบอลถูกซ้อนทับโดยบราวดิ้งบล็อกของผู้เล่นที่ใหญ่กว่าจะถูกพิจารณาว่าไม่พบนั้นอาจจะไม่เป็นจริงเสมอไปซึ่งในความเป็นจริงลูกบอลอาจจะไม่ถูกบล็อบของผู้เล่นบดบัง เพียงแต่อยู่ในบริเวณของบราวดิ้งบล็อกของผู้เล่นเท่านั้น นอกจากนี้ในงานวิจัยของ Belagiannis และคณะ [104] ยังกล่าวถึงปัญหาของการใช้บราวดิ้งบล็อกในลักษณะที่คล้ายกันว่ากรณีการใช้บราวดิ้งบล็อกเป็นพื้นฐานในการติดตามจะเกิดปัญหาการสะสมของแซมเปิลอนุภาค (Particle Sample) ของส่วนที่เป็นพื้นหลังปะปนมากับแซมเปิลอนุภาคของวัตถุซึ่งจะส่งผลต่อกระบวนการการติดตามที่ผิดพลาดได้โดยเฉพาะกับกรณีการใช้เทคนิคการติดตามบนพื้นฐานของตัวกรองอนุภาค

ในงานวิจัยนี้ได้ทำการประยุกต์ใช้อัลกอริทึมคอนเดนเซชันในการติดตามผู้เล่นซึ่งจะช่วยแก้ปัญหาการบดบังแบบบางส่วนได้ในเบื้องต้นและประยุกต์ใช้ความรู้พื้นฐานของตำแหน่งที่ได้จากการผสมข้อมูลและการทำนายเส้นทางผู้เล่นในมุมมองเสมือน (Top View) ในการระบุตำแหน่งของผู้เล่นในมุมมองภาพในกรณีที่มีการบดบังซึ่งเกิดจากการรวมกลุ่มของผู้เล่นซึ่งเทคนิคที่นำเสนอนี้สามารถช่วยในการแก้ปัญหาการบดบังซึ่งเกิดจากการรวมกลุ่มโดยยังสามารถติดตามเส้นทางของผู้เล่นที่ถูกบดบังได้

6. การแปลงมุมมอง (Perspective Transformation)

Tahar และ Imed [37] อธิบายกระบวนการแปลงมุมมองไว้ดังนี้ การแปลงนี้เป็นส่วนหนึ่งของการแปลงเชิงการฉายภาพ (Perspective Transformation) ระบบพิกัดโลก (World Coordinates (x, y, z)) ดังภาพที่ 2.20 (a) และระบบพิกัดจอ (Screen Coordinates (X, Y)) ดังภาพที่ 2.20 (b) ถ้ากล้องถูกติดตั้งที่ $P_v(x_v, y_v, z_v)$ และเลนส์ถูกชี้ (Pointed) ที่ $P_g(x_g, y_g, z_g)$ การแปลงจากระบบพิกัดโลก (*x, y, z*) ไปยังระบบพิกัดจอ (*X, Y*) กระทำได้ดังสมการที่ 2.34 ถึงสมการที่ 2.37



ภาพที่ 2.20 ระบบพิกัดโลก (a) และระบบพิกัดจอ (b) (ที่มา: Tahar และ Imed [37])

$$X = -f\frac{x_1}{z_1}, Y = -f\frac{y_1}{z_1} \tag{2.34}$$

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} -\sin\theta & \cos\theta & 0 \\ -\sin\varnothing\cos\theta & -\sin\varnothing\cos\theta & \cos\varnothing \\ \cos\varnothing\cos\theta & \cos\varnothing\sin\theta & \sin\varnothing \end{bmatrix} \bullet \begin{bmatrix} x - x_v \\ y - y_v \\ z - z_v \end{bmatrix} \tag{2.35}$$

$$\theta = \tan^{-1}\left(\frac{y_v - y_g}{x_v - x_g}\right) \tag{2.36}$$

$$\theta = \tan^{-1}\left(\frac{z_v - z_g}{\sqrt{(x_v - x_g)^2 + (y_v - y_g)^2}}\right) \tag{2.37}$$

โดยที่ x_v, y_v, z_v เป็นระบบพิกัดของตำแหน่งกล้อง ส่วนพารามิเตอร์ f คือกำลังขยาย (Magnification) ของกล้องและ x_g, y_g คือ พารามิเตอร์ควบคุมมุมกล้อง (Camera Angle) การแปลงนี้จะฉายภาพหรือโปรเจ็คในระบบพิกัดโลกไปเป็นระบบพิกัดจอซึ่งพารามิเตอร์เหล่านี้จะถูกเรียกเป็นพารามิเตอร์กล้อง โดยวัตถุในพิกัดโลก ดังภาพที่ 2.20 (a) จะถูกแปลงไปเป็นพิกัดจอโดยกล้อง ดังภาพที่ 2.20 (b)

Iwase และ Saito [42] ใช้หลักการของโฮโมกราฟฟี่ในการหาความสัมพันธ์ระหว่าง 2 กล้องซึ่งอยู่ในรูปแบบเมทริกซ์ขนาด 3 x 3 โดยใช้คู่ของจุดระหว่าง 2 กล้องมากกว่า 4 คู่ในการคำนวณ ในลักษณะความสัมพันธ์แบบหนึ่งต่อหนึ่งซึ่งจะได้ภาพเสมือนของโมเดลสนาม เมทริกซ์โฮโมกราฟฟี่ระหว่างภาพเสมือนกับภาพจากกล้องจะถูกคำนวณไว้ล่วงหน้าโดยใช้จำนวน 10 ถึง 20 จุด เช่น มุมของเส้นสนามและตำแหน่งเท้าผู้เล่น จากนั้นตำแหน่งผู้เล่นจะถูกโปรเจ็คลงบนภาพเสมือนโดยอาศัยความสัมพันธ์ของโฮโมกราฟฟี่นี้ จากนั้นทำการจัดกลุ่มเพื่อหาจุดที่ถูกโปรเจ็คเพื่อจัดเรียงตำแหน่งที่ถูกโปรเจ็คเข้ากับตำแหน่งที่ถูกประมาณจากเส้นทางในเฟรมก่อนหน้าโดยตัวกรองคาร์ลมาล ตำแหน่งผู้เล่นในเฟรมแรกบนภาพเสมือนจะถูกกำหนดด้วยมือสำหรับเป็นตำแหน่งเริ่มต้นในการติดตามและใช้ข้อมูลในการสอบเทียบกล้องน้อย

Park และคณะ [44] ใช้หลักการของโฮโมกราฟฟีโดยใช้เมทริกซ์ขนาด 3 x 3 ในการแปลงมุมมองภาพโดยในส่วนของตำแหน่งผู้เล่นที่ถูกโปรเจ็คบนภาพเสมือนจะทำการจัดกลุ่มแบบอัตโนมัติเพื่อแก้ปัญหาสัญญาณรบกวนจากการประมวลผลภาพและกรณีผู้เล่นอยู่ติดกันซึ่งทำให้ได้บล็อบเพียงบล็อบเดียว

Hayet และคณะ [49] ได้กล่าวไว้ว่าการปรับปรุงภาพให้ถูกต้อง (Rectification) ถือเป็นงานสำคัญต่อระบบเพื่อให้สามารถรวม (Merge) ข้อมูลที่ได้จากแต่ละกล้องซึ่งข้อมูลที่ได้จะถูกแสดงในเฟรมอ้างอิงทั่วไป เช่น เฟรมพื้น (Ground Frame) และใช้เมทริกซ์โฮโมกราฟฟี่ในการแมพจุดบนระนาบภาพไปเป็นจุดบนระนาบพื้นโดยมี 2 ส่วน คือ การใช้เทคนิคบนพื้นฐานของเส้นและการใช้เทคนิคบนพื้นฐานของจุด ในส่วนแรก จะใช้เส้นสนามที่ตรวจจับได้กำหนดให้กับโมเดลและติดตามเส้นสนามนี้เพื่ออัพเดทโฮโมกราฟฟี่ ในส่วนที่สอง จะใช้ข้อมูลจุดเพื่อประมาณโฮโมกราฟฟี่ เมื่อเทคนิคแรกไม่สามารถใช้ได้เนื่องจากไม่สามารถตรวจจับเส้นสนามได้ การอัพเดทค่าการประมาณของ โฮโมกราฟฟี่เมทริกซ์ของแต่ละกล้อง (k) สำหรับการแปลงตำแหน่งเป้าหมาย t_j^k ที่ถูกติดตามในระนาบ 2D ไปยังการประมาณ T_j^k ในระนาบพื้นแบบ 2D จะใช้สมการที่ 2.38 ดังนี้

$$\begin{pmatrix} T_j^k \\ 1 \end{pmatrix} \sim H^k \begin{pmatrix} t_j^k \\ 1 \end{pmatrix} \tag{2.38}$$

7. การสอบเทียบกล้อง (Camera Calibration)

Tahar และ Imed [37] เสนอวิธีการสอบเทียบกล้องหรือการประมาณพารามิเตอร์กล้องโดยใช้โมเดลสนามสำหรับฟุตบอลแบบเรียลไทม์โดยนำเสนอกระบวนการประมาณ ตำแหน่ง (Position) การวางแนว (Orientation) และความยาวโฟกัส (Focal Length) ของกล้องในสนามฟุตบอล พารามิเตอร์เหล่านี้จะถูกใช้ในการซ้อนทับภาพกราฟิกบนสนามฟุตบอลแบบเรียลไทม์ วิธีนี้จะใช้โมเดลสนามซึ่งประกอบด้วยเส้นโค้ง (Arcs) และเส้นตรง (Lines) โดยมีหลักการดำเนินการดังนี้

การสอบเทียบกล้องเริ่มต้น (Initial Camera Calibration) โดยทั่วไปตำแหน่งกล้องจะอยู่คงที่ระหว่างการแพร่ภาพการแข่งขัน ตำแหน่งนี้มักจะไม่เปลี่ยนแปลงระหว่างการแข่งขันที่แตกต่างกัน ณ พื้นสนามเดียวกัน ซึ่งจุดการติดตั้งกล้องโดยปกติจะติดตั้งคงที่กับโครงสร้างของสนามกีฬา ซึ่งสิ่งนี้จะช่วยลดพารามิเตอร์กล้องได้ระหว่างกระบวนการติดตาม การประมาณตำแหน่งกล้องจากภาพที่ตรงกันเป็นปัญหาที่มีข้อจำกัดที่ไม่ดีโดยเฉพาะถ้าความยาวโฟกัสต้องถูกนำมาประมาณด้วยเพื่อปรับปรุงความถูกต้องและลดองศาที่อิสระ (Degrees of Freedom) จึงใช้หลายภาพที่มีค่าความยาวโฟกัสเดียวกันและแก้ไขค่าตำแหน่งกล้องโดยคำนวณตำแหน่งกล้อง การวางแนวและมุมมองสนาม (Field-of-View) สำหรับการวางแนวกล้องที่ต่างกัน คือ กลาง ซ้ายและขวา ดังภาพที่ 2.21 (a), (b) และ (c) ตามลำดับ

ภาพที่ 2.21 การวางแนวกล้อง 3 ตัวสำหรับการสอบเทียบเบื้องต้น (ที่มา: Tahar และ Imed [37])

พารามิเตอร์เหล่านี้จะถูกคำนวณพร้อมกันทุกภาพและตำแหน่งจะถูกจำกัดเป็นค่าทั่วไปซึ่งจะช่วยลดความกำกวมระหว่างระยะทางของกล้องจากเอกลักษณ์ที่อ้างอิงและความยาวโฟกัสที่มีอยู่ ในวิธีการสอบเทียบกล้องส่วนใหญ่ พิกเซลในเส้นตรงที่ถูกตรวจจับบนสนามจะถูกใช้เพื่อขจัดความเพี้ยนหรือการบิดเบือน (Distortion) ของกล้องและใช้ถูกเพื่อคำนวณการแปลงมุมมองโดยใช้อัลกอริทึม Levenberg-Marquardt Nonlinear Least Squares และวิธีจาโคเบียน (Jacobian)

โดยการรวมกันของความแตกต่างที่แน่นอน (Finite Differences) และการอัพเดทเส้นตรงที่ตัดวงกลม (Secant)

การขจัดการบิดเบือนของกล้อง (Removing Camera Distortion) ระบบฉายภาพส่วนใหญ่จะมีเลนส์ซึ่งจะมีการบิดเบือนภาพแม้กล้องออกอากาศ (Broadcast) จะมีการบิดเบือนที่น้อย แต่ยังคงต้องโมเดลการบิดเบือนของรัศมีและเส้นสัมผัสวง (Radial and Tangential) และเพื่อให้การวัดสำหรับการแปลงมุมมองมีความถูกต้องมากขึ้น การบิดเบือนที่เกิดจากเลนส์จะมี 2 องค์ประกอบ คือ รัศมีและเส้นสัมผัสวงซึ่งแต่ละส่วนจะถูกโมเดลด้วยข้อกำหนดที่ไม่จำกัด (Infinite Terms) รูปแบบที่แพร่หลายของการบิดเบือน คือ บาร์เรล (Barrel) และการโค้งเข้า (Pincushion) โดยปกติเลนส์ที่มีมุมกว้างจะมีกำลังการขยายภาพที่สูงในบริเวณกลางภาพมากกว่าบริเวณขอบนอก ทำให้ขอบ (Edges) ของภาพจะหดตัวลงรอบศูนย์กลางและมีรูปร่างแบบทรงกระบอก (Barrel) การบิดเบือนการโค้งเข้าจะมีลักษณะที่ตรงข้ามเมื่อขอบถูกขยายมากขึ้น โมเดลการบิดเบือนจะสร้างความสัมพันธ์ระหว่างพิกเซลที่มีการบิดเบือนและพิกเซลที่ไม่มีการบิดเบือนโดยแสดงดังสมการที่ 2.39

$$
\begin{aligned}
x &= x_d\left(1+k_1r^2+k_2r^4+k_5r^6\right)+d_x \\
y &= y_d\left(1+k_1r^2+k_2r^4+k_5r^6\right)+d_y
\end{aligned}
\qquad (2.39)
$$

โดยที่ $\left(x_d, y_d\right)$ เป็นพิกัดพิกเซลที่มีการบิดเบือน $r=\sqrt{x_d^2+y_d^2}$ เป็นพารามิเตอร์กล้องที่โมเดลการบิดเบือน x_d, y_d เป็นพารามิเตอร์การบิดเบือนเส้นสัมผัสวง (Tangential Distortion Parameters) ดังสมการที่ 2.40

$$
\begin{aligned}
d_x &= 2k_3xy+k_4\left(r^2+2x^2\right) \\
d_y &= 2k_4xy+k_3\left(r^2+2y^2\right)
\end{aligned}
\qquad (2.40)
$$

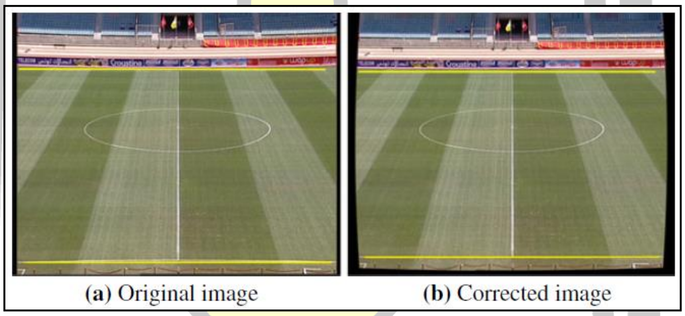
การประมาณการบิดเบือนจะเป็นลักษณะพหุนาม (Polynomial) ของดีกรีที่ไม่จำกัด (Infinite Degree) ดังสมการที่ 2.41

$$
\begin{aligned}
x &= x_d\left(1+k_1r^2\right)+d_x \\
y &= y_d\left(1+k_1r^2\right)+d_y
\end{aligned}
\qquad (2.41)
$$

เพื่อโมเดลเลนส์มุมกว้าง สัมประสิทธิ์ 2 ตัวจะถูกกำหนดดังสมการที่ 2.42

$$x = x_d\left(1 + k_1 r^2 + k_2 r^4\right) + d_x$$
$$y = y_d\left(1 + k_1 r^2 + k_2 r^4\right) + d_y \quad (2.42)$$

โดยที่ k_1 จะควบคุมพฤติกรรมทั่วไปของการบิดเบือนและ k_2 จะถูกปรับถ้าการบิดเบือนมีความรุนแรง เนื่องจากกายภาพขององค์ประกอบแสง (Optical Elements) การบิดเบือนของรัศมีและเส้นสัมผัสวง (Radial and Tangential Distortions) เป็นลักษณะแปรผัน (Proportional) กับกำลังของ r และดังนั้นจึงมีเพียงกำลังของ r ที่ปรากฏในโมเดลการบิดเบือน พารามิเตอร์ k_1 จะกำหนดชนิดของการบิดเบือนเลนส์รัศมีเมื่อ $k_1 > 0$ การบิดเบือนทรงกระบอก (Barrel Distortion) จะได้รับเมื่อ $k_1 < 0$ การบิดเบือนเส้นโค้ง (Pincushion Distortion) จะปรากฏ ฟังก์ชันการบิดเบือนที่ไม่สามารถย้อนกลับได้จึงใช้วิธีวนซ้ำเพื่อทำให้ภาพที่ถูกฉายนั้นถูกต้องและดังนั้นในระนาบโฟกัส (Focal Plane) ของกล้อง เซกเมนต์ของเส้นตรงจะแทนที่เส้นโค้ง (Curves) แสดงดังภาพที่ 2.22



ภาพที่ 2.22 การแก้ไขการบิดเบือนแบบไม่เป็นเชิงเส้น (ที่มา: Tahar และ Imed [37])

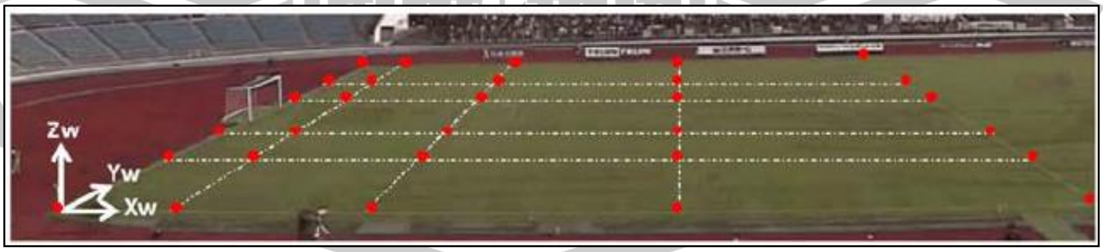
ภาพที่ 2.23 จุดอ้างอิงบนสนาม (ที่มา: Tahar และ Imed [37])

การประมาณเบื้องต้นของตำแหน่งกล้อง (Initial Estimation of Camera Positions) ในการคำนวณตำแหน่งกล้องและความยาวโฟกัสของแต่ละมุมมอง (View) ดังภาพที่ 2.21 โดยใช้สมการการแปลงมุมมองซึ่งจะได้เซกเมนต์ของเส้นตรงและจุดตัดที่เป็นลักษณะเฉพาะของสนามฟุตบอลซึ่งจะใช้ 35 จุดที่ชัดเจนที่สามารถระบุบนสนามฟุตบอล ดังภาพที่ 2.23 โดยสามารถใช้จุดตัดเสมือน (Virtual Intersection Points) อื่นๆ ที่สามารถปรากฏออกมาได้ถ้าเซกเมนต์ของเส้นตรงทั้งหมดถูกยืดออกมาเป็นเส้นตรงที่ไม่จำกัด จุดตัด (3-12) กับ (1, 16) หรือ (2, 13) กับ (19, 20) เป็นตัวอย่างของจุดเสมือน (Virtual Points) ซึ่งทำให้สามารถระบุจุดอื่นๆ ได้ โดยจำนวนจุดที่สามารถระบุได้สูงสุด คือ 96 จุดในมุมมองของกล้องเดียว

Nishino และคณะ [41] ทำการสอบเทียบกล้องโดยใช้ข้อมูลจุดบนสนามฟุตบอลในการกำหนดพิกัดโลกโดยใช้สมการในการแปลงมุมมองคล้ายกับของงานวิจัยของสุเทพ อาษาและคณะ [57] ซึ่งมีหลักการสอบเทียบกล้อง คือ มุมมองการฉายภาพหรือการโปรเจ็คจากพิกัดโลกไปยังพิกัดภาพจะใช้สมการที่ 2.43 โดย *A* คือ พารามิเตอร์ภายใน *R* คือ เมทริกซ์การหมุน *t* คือ เวกเตอร์การแปลง โดย *R* และ *t* คือ พารามิเตอร์ภายนอก

$$\begin{bmatrix} x_f \\ y_f \\ 1 \end{bmatrix} = A\left[R \mid t\right] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \qquad (2.43)$$

การดำเนินการจะใช้รายละเอียดของสนามบอล คือ จุดต่างๆ บนสนาม ดังภาพที่ 2.24 ในการกำหนดพิกัดโลกโดยจะดำเนินการโดยใช้คน จากนั้นพารามิเตอร์กล้องจะถูกสอบเทียบด้วยวิธีอัลกอริทึมของไซด์



ภาพที่ 2.24 จุดสำหรับการสอบเทียบ (ที่มา: Nishino และคณะ [41])

ตัวอย่างของงานวิจัยที่ทำการสอบเทียบกล้องโดยใช้อัลกอริทึมของไซด์ เช่น งานวิจัยของ Ren และคณะ [46] เป็นต้น

8. การกำหนดเป้าหมาย (Target Assignment/Data Association)

Xu และคณะ [58] ทำการเชื่อมโยงข้อมูลโดยประกอบด้วย 2 ขั้นตอน คือ การเชื่อมโยงเอกลักษณ์กับเป้าหมายและการเชื่อมโยงเอกลักษณ์กับเอกลักษณ์ ในขั้นตอนของการเชื่อมโยงเอกลักษณ์กับเป้าหมายมี 3 ขั้นตอนย่อย คือ 1. การโมเดลผู้เล่นเป็นเป้าหมายที่แสดงด้วยสเตท (State) 2. การอัพเดท State ด้วยมีชัวร์เมนท์ที่วัดได้ และ 3. กรณีที่ไม่มีมีชัวร์เมนท์ใดใช้ได้จะอัพเดทสเตทโดยใช้ผลการประมาณก่อนหน้า ส่วนขั้นตอนของการเชื่อมโยงเอกลักษณ์กับเอกลักษณ์จะเกิดขึ้นในกรณีที่มีชัวร์เมนท์ไม่จับคู่กับแทรค (Track) ที่มีอยู่ มี 2 ขั้นตอนย่อย คือ 1. การเปรียบเทียบระหว่างมีชัวร์เมนท์ในแต่ละกล้องเพื่อค้นหาเป้าหมายใหม่และ 2. 2 มีชัวร์เมนท์จากแต่ละกล้องที่เหมือนกันจะถูกเชื่อมโยงเข้าด้วยกันเพื่อสร้างเป้าหมายใหม่โดยพิจารณาจากระยะทาง

Hayet และคณะ [49] ทำการกำหนดเป้าหมายโดยใช้มีชัวร์เมนท์จากโลคอลแทรคเกอร์ (Local Tracker) ในการอัพเดทสเตทเวกเตอร์ (State Vector) ของเป้าหมายในโกลบอลแทรคเกอร์ (Global Tracker) ซึ่งประกอบด้วยตำแหน่งและความเร็ว การกำหนดเป้าหมายจะใช้หลักการของคอสต์เมทริกซ์ (Cost Matrix) โดยพิจารณาจากระยะทางระหว่างตำแหน่งของเป้าหมายในโลคอลกับตำแหน่งที่ทำนายของเป้าหมายในโกลบอลด้วยวิธีมหาลาโนบิส โดยคอสต์เมทริกซ์นี้ได้ปรับใช้กับวัตถุที่ไม่ปรากฏและการตรวจจับปลอมด้วย สุดท้ายการกำหนด (Assignment) ที่ถูกต้องจะมีเพียงรายการเดียวในเมทริกซ์ซึ่งหมายถึงมีความสอดคล้องเดียวในแถวหรือคอลัมน์ใดๆ ในเมทริกซ์

จากการศึกษางานวิจัยดังกล่าวข้างต้นสรุปได้ว่าการเชื่อมโยงข้อมูลเป็นกระบวนการของการสร้างความสัมพันธ์ระหว่างเป้าหมายในมุมมองภาพ (Local Target) กับเป้าหมายในมุมมองเสมือน (World Target) เพื่อกำหนดเป้าหมายและเส้นทางการเคลื่อนที่ที่ถูกต้องระหว่างมุมมองภาพและมุมมองเสมือน

9. การผสมข้อมูล (Data Fusion)

Zhao และคณะ [48] ทำการผสมข้อมูลจากหลายกล้องโดยจะรวมแทร็คของมนุษย์คนเดียวกันในระดับโลคอลไปเป็นแทรคระดับโกลบอล ส่วนงานวิจัยของ Xu และคณะ [58] มีหลักการสร้างมีชัวร์เมนท์ที่ถูกผสม คือ เซตของผู้เล่นจะถูกเชื่อมโยงเข้ากับเซตของมีชัวร์เมนท์ จากแต่ละกล้อง ผลที่ได้ คือ เมทริกซ์ความสัมพันธ์ (Association Matrix) โดยวิธีการเชื่อมโยงแบบง่ายจะใช้หลักการของเพื่อนบ้านที่ใกล้ที่สุด (Nearest Neighbour) เช่น ใช้วิธีมหาลาโนบิสในการหาระยะทางระหว่างมีชัวร์เมนท์กับเป้าหมายที่ทำนาย (Target Prediction) ดังนั้น มีชัวร์เมนท์เพียงอันเดียวของแต่ละเป้าหมายจะเป็นการรวมมีชัวร์เมนท์มาจากแต่ละกล้องซึ่งแต่ละเป้าหมายจะถูกอัพเดทโดยใช้มีชัวร์เมนท์ที่ถูกผสมนี้ในขั้นตอนของการเชื่อมโยงข้อมูล

Hayet และคณะ [49] ทำการผสมข้อมูลโดยพิจารณาจากความสอดคล้องที่พบในคอสต์เมทริกซ์ ซึ่ง คือ พิกัดเป้าหมายที่วัดได้จากแต่ละกล้อง โดยพิกัดที่ได้เหล่านี้จะถูกผสมเข้าด้วยกันเพื่อสร้างเป็นเป้าหมายเพียงอันเดียวในมุมมองโกลบอล ในการผสมข้อมูลสามารถกระทำโดยใช้สมการที่ 2.44

$$Z_n = \left(\sum_k \left(R^k_{\lambda^k_n} \right)^{-1} \right)^{-1} \sum_k \left(R^k_{\lambda^k_n} \right)^{-1} T^k_{\lambda_n} \tag{2.44}$$

จากสมการจะได้มีชัวร์เมนท์ที่ถูกผสม Z_n ซึ่งจะถูกนำไปใช้ในตัวกรองคาร์ลมาล สำหรับอัพเดทสเตทของเป้าหมายในมุมมองโกลบอล โดยที่ T คือ พิกัดของวัตถุ k คือ รายการวิดีโอ λ คือ ดัชนีและ R คือ ความแปรปรวน เพื่อให้ได้สมการนี้งานวิจัยนี้ได้สันนิษฐานว่าสัญญาณรบกวนบนพิกัดพื้นที่มาจากชุดวิดีโอต่างๆ เป็นอิสระต่อกัน

Cheikh และคณะ [43] ทำการผสมข้อมูลจากหลายกล้องโดยประยุกต์ใช้อัลกอริทึมเกลแชปลีย์ (Gale-Shapley Algorithm: GSA) ทำการเชื่อมโยงเป้าหมายจากกล้องต่างๆ เข้าด้วยกันโดยใช้ข้อมูลสี ความสูงและความกว้างของวัตถุจาก 2 มุมมอง เป็นต้น

10. การเลือกเป้าหมาย (Target Selection)

งานวิจัยของ Xu และคณะ [58] เสนอวิธีการตรวจจับและติดตามผู้เล่นโดยใช้หลายกล้องเพื่อช่วยแก้ปัญหาการบดบัง ใช้วิธีการหาความแตกต่างของภาพในการตรวจจับผู้เล่นและใช้ตัวกรองคาร์ลมาลในการติดตามผู้เล่นในระบบแบบหลายกล้อง การเลือกเป้าหมายจะใช้จำนวนของประชากรทั้งหมดในสนามในการเลือกรายแทรคที่ดีที่สุดจำนวน 25 แทรค จากรายการแทรคทั้งหมดที่มี เป็นต้น

จากการศึกษางานวิจัยดังกล่าวข้างต้นสรุปได้ว่าจุดประสงค์ของการเลือกเป้าหมายเนื่องมาจากรายการวัตถุหรือแทรคของวัตถุที่ได้จากการตรวจจับและการโปรเจ็คลงบนภาพเสมือนนั้นมาจากข้อมูลของกล้องหลายตัวส่งผลให้จำนวนของวัตถุหรือแทรคนั้นมีจำนวนมากและอาจเกิดความซ้ำซ้อนจึงต้องมีการคัดเลือกวัตถุหรือแทรคที่เหมาะสมที่สุด วิธีหนึ่งในการกำหนดจำนวนวัตถุหรือแทรคที่ต้องการคัดเลือก คือ พิจารณาจากความรู้ที่มีอยู่เกี่ยวกับวัตถุที่เกี่ยวข้อง เช่น ในการแข่งขันฟุตบอลโดยปกติจะมีบุคคลที่ปรากฏในสนามมากที่สุด 25 คน คือ ผู้เล่นทีมละ 11 คน รวมเป็น 22 คน กรรมการ 3 คน รวมทั้งสิ้น 25 คน เป็นต้น

11. การติดตามผู้เล่นในแบบหลายมุมมอง (Multi-View Player Tracking)

Iwase และ Saito [42] ทำการรวมตำแหน่งเท้าของผู้เล่นจากแต่ละกล้องโดยประยุกต์ใช้วิธีโฮโมกราฟี่แล้วทำการโปรเจ็คตำแหน่งผู้เล่นลงบนภาพพื้นเสมือน (Virtual Ground

Image) ในมุมมองด้านบน (Top-View) ในทุกๆ เฟรมแล้วทำการจัดกลุ่มตำแหน่งที่ถูกฉายบนภาพพื้นเสมือนโดยจัดเรียงตำแหน่งที่ถูกฉายเข้าสู่ตำแหน่งผู้เล่นที่ถูกประมาณจากเส้นทางในเฟรมก่อนหน้าด้วยวิธีตัวกรองคาร์ลมาลฟิลด์เตอร์

Park และคณะ [44] ทำการแก้ไขสี (Color Correction) ก่อนการดำเนินการติดตามเพื่อปรับข้อมูลสีของภาพจากแต่ละกล้องให้สอดคล้องกัน การติดตามจะใช้ตำแหน่งเท้าผู้เล่นจากแต่ละกล้อง ตำแหน่งที่ถูกโปรเจ็คจะถูกจับคู่เข้ากับรายการแทรคโดยพิจารณาจากความแตกต่างระหว่างตำแหน่งที่ทำนายกับตำแหน่งที่วัดได้เทียบกับค่าเทรชโฮลด์ เป็นต้น

Ren และคณะ [46] ทำการติดตามผู้เล่นโดยใช้เมทริกซ์ความสัมพันธ์เพื่อแสดงความสัมพันธ์ระหว่างผู้เล่นกับมีชัวร์เมนท์ การหาความสัมพันธ์จะใช้วิธีหาระยะทางแบบมหาลาโนบิส (Mahalanobis Distance)

จากการศึกษาประเด็นการแสดงผลตำแหน่งผู้เล่นในมุมมองเสมือนสามารถสรุปได้ดังนี้ คือ การแสดงผลตำแหน่งและเส้นทางของผู้เล่นในวิดีโอการแข่งขันกีฬาฟุตบอลที่นิยมในปัจจุบันมี 2 แบบ คือ การแสดงผลในมุมมองภาพ (Image View Representation) และการแสดงผลในมุมมองเสมือน (Virtual Plan / Top View Representation) ซึ่งวิธีการแสดงผลแบบที่สองนั้นมักจะใช้ในกรณีที่ข้อมูลภาพจากกล้องเพียงกล้องเดียวไม่สามารถครอบคลุมพื้นที่ทั้งสนามได้จึงนำข้อมูลจากกล้องหลายๆ ตัวมารวมกันแล้วแสดงในหน้าจอเสมือนจริงเพื่อให้เห็นภาพรวมทั้งหมดภายในสนามซึ่งการแสดงผลวิธีนี้มีข้อดี คือ ผู้ชมสามารถเห็นตำแหน่งและการเคลื่อนที่รวมถึงรูปแบบของเกมส์การแข่งขันได้ซึ่งช่วยเพิ่มอรรถรสในการชมกีฬา นอกจากนี้ผู้ควบคุมทีมหรือโค๊ชยังสามารถใช้ข้อมูลเหล่านี้ในการวางแผนการเล่นรวมถึงการวิเคราะห์ประสิทธิภาพการเล่นของทีมและผู้เล่นแต่ละคนได้อีกด้วย ในปัจจุบันมีหลายงานวิจัยที่ประยุกต์ใช้วิธีการแสดงผลในมุมมองเสมือน เช่น งานวิจัยของ Iwase และ Saito [42] งานวิจัยของ Xu และคณะ [58] และงานวิจัยของ Yamada และคณะ [98] เป็นต้น

ประเด็นปัญหาแรกที่จะพบจากการแสดงผลในมุมมองเสมือนที่สำคัญ คือ การแปลงตำแหน่งจากมุมมองภาพไปยังมุมมองเสมือนเนื่องจากมีความแตกต่างกันในเรื่องของมุมมองและลักษณะทางกายภาพซึ่งทำให้ตำแหน่งของวัตถุที่ได้จะแตกต่างกันระหว่างมุมมองภาพกับมุมมองเสมือนและระหว่างมุมมองภาพจากแต่ละกล้องซึ่งปัญหานี้จะส่งผลต่อความถูกต้องในการแปลงตำแหน่งจากมุมมองภาพไปยังมุมมองเสมือนและส่งผลต่อกระบวนการผสมข้อมูลระหว่างกล้องซึ่งผู้เล่นคนเดียวกันอาจถูกแปลงไปยังมุมมองเสมือนในตำแหน่งที่ต่างกันและทำให้การผสมข้อมูลมีความคลาดเคลื่อนไปจากความเป็นจริงได้

วิธีพื้นฐานในการแปลงตำแหน่งระหว่างมุมมองภาพและมุมมองเสมือน คือ การใช้หลักการของการแปลงมุมมอง ซึ่งจะช่วยให้ตำแหน่งระหว่างมุมมองภาพและมุมมองเสมือนมีความสอดคล้องกันตามหลักของความเป็นจริงในทางกายภาพซึ่งเป็นวิธีที่งานวิจัยส่วนใหญ่นิยมใช้

ยกตัวอย่างงานวิจัย เช่น งานวิจัยของ Iwase และ Saito [42] งานวิจัยของ Yamada และคณะ [98] งานวิจัยของ Sterning และคณะ [99] งานวิจัยของ Rodrigues และคณะ [105] และงานวิจัยของ Martin และ Martínez [101] เป็นต้น

ประเด็นปัญหาต่อมาที่จะพบจากการแสดงผลในมุมมองเสมือน คือ การผสมข้อมูลจากแต่ละกล้องให้เป็นข้อมูลชุดเดียวกันเพื่อแสดงผลบนมุมมองเสมือน เนื่องจากในแต่ละกล้องข้อมูลผู้เล่นคนเดียวกันอาจมีลักษณะที่ไม่เหมือนกันโดยสมบูรณ์ เช่น กล้องหนึ่งเห็นด้านหน้าส่วนอีกกล้องหนึ่งเห็นด้านหลัง กล้องหนึ่งเห็นวัตถุอยู่ในระยะใกล้แต่อีกกล้องหนึ่งเห็นวัตถุอยู่ในระยะไกลทำให้ลักษณะของวัตถุที่ได้อาจแตกต่างกัน เช่น รูปร่าง สีและขนาด เป็นต้น จึงเกิดคำถามที่ว่า เราจะรู้ได้อย่างไรว่าผู้เล่นจากกล้องหนึ่ง คือ คนเดียวกันกับผู้เล่นที่ปรากฏอยู่อีกกล้องหนึ่ง ดังนั้นปัญหาการจำแนกและการระบุตัวตนของผู้เล่นระหว่างกล้องจึงกลายเป็นปัญหาหลักที่สำคัญของกระบวนการผสมข้อมูล (Data Fusion) คำถามที่ตามมา คือ ในเมื่อแต่ละกล้องได้ข้อมูลผู้เล่นที่อาจแตกต่างกัน เช่น มีขนาดที่ต่างกันเราจะประมาณตำแหน่งของผู้เล่นบนมุมมองเสมือนอย่างไรให้ใกล้เคียงกับตำแหน่งจริงมากที่สุดในเมื่อแต่ละกล้องอาจมีผลการประมาณตำแหน่งผู้เล่นที่ไม่เหมือนกัน เป็นต้น

เทคนิคพื้นฐานหนึ่งที่ถูกนำเสนอเพื่อแก้ปัญหาการผสมข้อมูลระหว่างกล้อง คือ การใช้ข้อมูลระยะทางระหว่างตำแหน่งที่ถูกโปรเจ็คบนมุมมองเสมือนเป็นเกณฑ์ในการพิจารณาว่าเป็นผู้เล่นคนเดียวกันหรือไม่หากอยู่ในค่าเทรชโฮลด์ที่กำหนดจะถูกพิจารณาว่า 2 ตำแหน่งนั้น คือ วัตถุคนเดียวกัน ตัวอย่างงานวิจัยที่ประยุกต์ใช้เทคนิคนี้ เช่น งานวิจัยของ Martin และ Martínez [101] ซึ่งมีหลักการพิจารณา 2 บล็อบจาก 2 กล้องว่าเป็นวัตถุเดียวกันโดยใช้ Mean Square Distance หากค่า Score อยู่ในเกณฑ์ของค่าเทรชโฮลด์ จะพิจารณาว่าเป็นวัตถุเดียวกัน ส่วนงานวิจัยของ Iwase และ Saito [42] จะใช้หลักการพิจารณาที่คล้ายกัน คือ การพิจารณาจากระยะทาง (Distance) ระหว่าง 2 จุดบนมุมมองเสมือนหากอยู่ในเกณฑ์ของค่าเทรชโฮลด์จะถูกพิจารณาว่าเป็นวัตถุเดียวกัน เป็นต้น การใช้เทคนิคนี้มีข้อดี คือ การประมวลผลไม่ซับซ้อน ง่ายและรวดเร็ว แต่อาจมีปัญหากรณีที่วัตถุที่ได้จากแต่ละกล้องมีขนาดที่ต่างกันทำให้ตำแหน่งที่ได้จากการประมาณแตกต่างกันเมื่อถูกโปรเจ็คลงบนมุมมองเสมือนจะได้ตำแหน่งที่แตกต่างกันซึ่งอาจจะอยู่ห่างกันมากเกินกว่าค่าเทรชโฮลด์ที่กำหนดส่งผลให้ระบบพิจารณาว่าเป็นคนละคนกันได้ นอกจากนี้เนื่องจากมุมมองของกล้องที่แตกต่างกันก็จะส่งผลให้ตำแหน่งที่ได้ในมุมมองเสมือนแตกต่างกันด้วยซึ่งอาจส่งผลต่อเทคนิคนี้ได้เช่นกัน อีกทั้งการกำหนดค่าเทรชโฮลด์หากไม่สอดคล้องหรือเหมาะสมกับลักษณะของข้อมูลก็อาจจะส่งผลให้ระบบทำงานผิดพลาดได้ด้วย

อีกเทคนิคหนึ่งที่นำมาประยุกต์ใช้ คือ การใช้เอกลักษณ์รูปร่าง (Shape) ของวัตถุในการโปรเจ็คลงบนมุมมองเสมือนแล้วคำนวณหาพื้นที่ที่ซ้อนทับกันโดยอาจพิจารณาเอกลักษณ์อื่นๆ ร่วมด้วย เช่น เอกลักษณ์สี รายการจุดในรูปร่างและข้อมูลระยะทางซึ่งวิธีนี้จะช่วยเพิ่มความถูกต้องในการ

ผสมข้อมูลและสามารถแก้ปัญหาของเทคนิคแรกที่ใช้ข้อมูลระยะทางเพียงอย่างเดียวดังที่อธิบายไว้ข้างต้น แต่เทคนิคนี้ยังสามารถเกิดปัญหาในลักษณะที่คล้ายกันได้ในกรณีที่วัตถุที่ได้จากแต่ละกล้องมีลักษณะและขนาด รวมถึงมุมมองที่แตกต่างกันมากส่งผลทำให้รูปร่างของวัตถุและตำแหน่งที่ได้ในมุมมองเสมือนมีความแตกต่างกันมากซึ่งในบางครั้งรูปร่างของวัตถุคนเดียวกันไม่ได้ซ้อนทับกันจึงส่งผลให้ระบบจะพิจารณาว่าเป็นคนละคนกันได้เช่นกัน นอกจากนี้ยังต้องใช้เอกลักษณ์ในการประมวลผลเพิ่มมากขึ้นด้วย

ประเด็นปัญหาสุดท้ายซึ่งส่งผลต่อการประมาณตำแหน่งบนมุมมองเสมือน คือ ลักษณะของมุมมองกล้องที่แตกต่างกันและวิธีการกำหนดตำแหน่งของวัตถุในมุมมองภาพ เมื่อกล้องแต่ละตัวมีลักษณะของมุมมองกล้องที่แตกต่างกันทำให้ตำแหน่งของวัตถุที่ได้แตกต่างกันเมื่อโปรเจ็คลงบนภาพในมุมมองเสมือนก็จะได้ตำแหน่งที่แตกต่างกันด้วย ส่วนวิธีการกำหนดตำแหน่งของวัตถุในมุมมองภาพโดยทั่วไปมี 3 วิธี คือ การกำหนดตำแหน่งบริเวณศีรษะของผู้เล่น การกำหนดตำแหน่งบริเวณจุดศูนย์กลางของบราวดิ้งบล็อกและการกำหนดตำแหน่งบริเวณเท้าของผู้เล่นซึ่งแต่ละวิธีมีข้อดีและข้อเสียที่แตกต่างกันไปโดยวิธีการกำหนดตำแหน่งบริเวณเท้าผู้เล่นจะมีข้อดี คือ ตำแหน่งที่ได้จะใกล้เคียงกับตำแหน่งจริงในมุมมองโลกเพราะเป็นบริเวณที่อยู่ใกล้หรือสัมผัสกับพื้นดินมากที่สุด แต่วิธีนี้จะมีข้อเสียในกรณีที่วัตถุมีขนาดเล็กส่งผลให้การตรวจจับเท้าของผู้เล่นอาจผิดพลาดส่งผลให้ตำแหน่งที่ได้มีความคลาดเคลื่อนได้เช่นกัน

ในงานวิจัยนี้ได้นำเสนอเทคนิคการผสมข้อมูลระหว่างกล้องบนมุมมองเสมือนโดยใช้เทคนิคที่เรียกว่า โมเดลรูปร่างของข้อมูลบนพื้นฐานทางกายภาพ (Physical based Data Shape Model) และใช้หลักการพื้นฐานของการแปลงมุมมองในการแปลงตำแหน่งระหว่างมุมมองภาพและมุมมองเสมือน การประมาณตำแหน่งสำหรับการผสมข้อมูลในมุมมองเสมือนในงานวิจัยนี้เสนอเทคนิคที่เรียกว่า การประมาณตำแหน่งบนพื้นฐานของสมการเส้นตรง (Straight Line Equation based Position Estimation) ซึ่งมีแนวคิดมาจากหลักการของเข็มทิศและมีพื้นฐานอยู่บนสมการเส้นตรงทางคณิตศาสตร์ ข้อดีของเทคนิคที่เสนอนี้ คือ เป็นวิธีที่ง่าย ไม่ซับซ้อน ใช้ข้อมูลสำหรับการประมวลผลที่น้อยและสามารถตอบโจทย์ 2 ประเด็นหลักของกระบวนการผสมข้อมูล ได้แก่ 1. การตัดสินว่าผู้เล่นในแต่ละกล้องเป็นคนเดียวกันหรือไม่และ 2. การประมาณตำแหน่งของผู้เล่นที่ใกล้เคียงกับตำแหน่งจริงแม้ว่าขนาดของวัตถุและมุมมองกล้องจะแตกต่างกันและสามารถรองรับกับวิธีการกำหนดตำแหน่งของวัตถุได้หลายแบบ เช่น การกำหนดตำแหน่งวัตถุในบริเวณศีรษะ จุดศูนย์กลางบล็อกและบริเวณเท้า เป็นต้น หลังจากกระบวนการโปรเจ็คตำแหน่งที่ได้จากการผสมข้อมูลลงบนมุมมองเสมือนแล้วจะใช้เทคนิคการติดตามตำแหน่งโดยใช้วิธีพื้นฐานทั่วไป คือ ตัวกรองคาร์ลมาลฟิลเตอร์สำหรับการติดตามเส้นทางการเคลื่อนที่ของผู้เล่นในมุมมองเสมือนซึ่งข้อมูลตำแหน่งและเส้นทางของผู้เล่นนี้จะถูกนำไปประยุกต์ใช้สำหรับการทำนายและแก้ปัญหาการบดบังซึ่งเกิดจากการรวมกลุ่มต่อไป

12. ประเด็นปัญหาการรวมตัวของผู้เล่น (Player's Aggregation Problem Issues)

ตัวอย่างงานวิจัยที่ประสบปัญหาจากการจับกลุ่มของคนอย่างหนาแน่น คือ งานวิจัยของ Ren และคณะ [46] ซึ่งได้เสนอระบบเฝ้าระวังแบบใช้หลายกล้องสำหรับการวิเคราะห์แบบเรียลไทม์และการสร้างเกมฟุตบอลใหม่ โดยเสนอเทคนิคและการดำเนินการระบบสำหรับการติดตามผู้เล่นและลูกบอลแบบใช้หลายกล้องซึ่งติดตั้งคงที่ ผลการทดลองพบว่ามียังความผิดพลาดในการติดตามเนื่องจากสถานการณ์ที่มีความหนาแน่นมาก เช่น ผู้เล่นจับคู่กันอย่างหนาแน่นในระหว่างการเตะมุม งานวิจัยของ Ali และคณะ [21] ได้ประยุกต์ใช้วิธีโซเบลเกรเดียนสำหรับสกัดผู้เล่นและลูกบอลที่เป็นไปได้ ผลการทดลองพบว่าในส่วนของการค้นหาผู้เล่นนั้นในบางฉากระบบไม่สามารถแสดงผลของผู้เล่นได้ครบทุกคนและมีการแสดงผลในส่วนของกรรมการด้วยและยังเกิดข้อผิดพลาดในการตรวจจับผู้เล่นกรณีที่ผู้เล่นอยู่ติดกันจำนวนมาก งานวิจัยของ Ali และคณะ [21] ทำการตรวจจับผู้เล่นและลูกบอลโดยหลังจากขจัดพื้นสนามแล้วจะใช้วิธีโซเบลเกรเดียนในการสกัดผู้เล่นและลูกบอลและใช้ค่าเทรชโฮลด์ในการขจัดวัตถุที่ไม่ต้องการ ผลนั้นพบว่ายังมีข้อผิดพลาดในการตรวจจับผู้เล่นที่อยู่ติดกันโดยระบบจะพิจารณาเป็นวัตถุเดียวกัน โดยงานวิจัยของ Hayet และคณะ [49] ได้ใช้โมเดลการกระจายตัวของจุดแก้ปัญหาการบดบังแบบบางส่วน ส่วนงานวิจัยของ Figuerou และคณะ [59] แก้ปัญหาการบดบังโดยใช้ข้อมูลจากกราฟโดยเน้นการบดบังแบบ 2 คนเป็นหลักซึ่งจะพิจารณาจากเส้นทางของวัตถุว่าเป็นเส้นทางเดียวกันหรือไม่ การรักษาเส้นทางที่ถูกต้องระหว่างการติดตามเมื่อผู้เล่นแยกออกจากกันจะพิจารณาจากข้อมูลสีและระยะทางระหว่างบล็อบจากกราฟ ส่วนในช่วงระหว่างการบดบังจะพิจารณาจากทิศทางของเส้นทางวัตถุเพื่อคำนวณหาเส้นทางที่ถูกต้องโดยจะโฟกัสผู้เล่นจากกล้องที่เห็นชัดเจนที่สุด ส่วนวิธีที่กำลังได้รับความนิยม คือ การใช้ข้อมูลจากหลายกล้องในการดำเนินการเพื่อแก้ปัญหาต่างๆ เช่น การบดบัง และการติดตามเส้นทาง เป็นต้น

จากการศึกษางานวิจัยดังกล่าวข้างต้นพบว่าส่วนใหญ่มุ่งเน้นการแก้ปัญหาการบดบัง การระบุตัวตนและการขาดหายของเส้นทางซึ่งลักษณะของการบดบังมักจะเป็นการบดบังระหว่างผู้เล่น 2-3 คน อีกทั้งผู้เล่นมักอยู่แบบกระจายเป็นส่วนใหญ่ แต่ยังไม่มีงานวิจัยใดที่กล่าวถึงปัญหาการจับกลุ่มมากนักและยังไม่มีงานวิจัยใดที่มุ่งเน้นการแก้ไขปัญหาการจับกลุ่มซึ่งเป็นปัญหาหลักที่สามารถก่อให้เกิดปัญหาอื่นๆ ดังกล่าวข้างต้น นอกจากนี้วิธีแก้ปัญหาการบดบังที่หลายงานวิจัยได้เสนอข้างต้นแม้จะสามารถแก้ปัญหาการบดบังได้แต่ก็เป็นเพียงระดับหนึ่งเท่านั้นซึ่งส่วนใหญ่จะใช้ได้กับกรณีเป็นปัญหาการบดบังระหว่างผู้เล่นเพียง 2-3 คนเท่านั้น หากเกิดกรณีเหตุการณ์ที่มีการรวมตัวของผู้เล่นเป็นจำนวนมากจะเกิดปัญหาการบดบังระหว่างผู้เล่นในลักษณะเป็นกลุ่มขึ้นซึ่งส่งผลให้มีความยากต่อการแก้ปัญหามากกว่ากรณีการบดบังระหว่างผู้เล่นเพียง 2-3 คน ดังที่งานวิจัยส่วนใหญ่ให้ความสนใจ ตัวอย่างงานวิจัยที่ประสบกับปัญหาการบดบังแบบกลุ่มนี้ เช่น Ren และคณะ [46] และ Ali และคณะ [21] เป็นต้น สำหรับวิธีแก้ปัญหาการบดบังที่ใช้โมเดลการกระจายตัวของจุดในงานวิจัยของ Hayet

และคณะ [49] จะสามารถแก้ปัญหาการบดบังแบบบางส่วนได้แต่กรณีที่มีการบดบังแบบสมบูรณ์ซึ่งสามารถเกิดขึ้นได้ในเหตุการณ์การรวมตัวของผู้เล่นทำให้วิธีนี้ค่อนข้างจะไม่สามารถใช้ได้ อีกทั้งการจับกลุ่มของผู้เล่นอาจส่งผลต่อการเปลี่ยนแปลงรูปร่างของจุดซึ่งอาจส่งผลต่อโมเดลการกระจายตัวของจุดได้เนื่องจากโมเดลนี้มีพื้นฐานบนข้อมูลจุด ส่วนการใช้กราฟแก้ปัญหาการบดบังในงานวิจัยของ Figuerou และคณะ [59] สามารถจัดการกับการบดบังระหว่างคนเพียง 2-3 คนเท่านั้น เมื่อเกิดเหตุการณ์การจับกลุ่มของผู้เล่นหลายคนขึ้น การอาศัยเพียงข้อมูลพื้นฐานจากกราฟ เช่น เส้นทางการเคลื่อนที่ ทิศทาง ขนาด สีและระยะทางระหว่างบล็อบ ในการทำนายเส้นทางและระบุตัวตนอาจจะไม่เพียงพอ เนื่องจากสภาพแวดล้อมของเหตุการณ์การจับกลุ่มจะส่งผลต่อทิศทางการเคลื่อนที่ของผู้เล่นซึ่งอาจจะไม่เป็นไปในทิศทางเดิมโดยอาจเปลี่ยนไปในทิศทางอื่นหรือตรงกันข้ามได้ อีกทั้งผลจากการบดบังที่หนาแน่นนั้นจะส่งผลต่อประสิทธิภาพในการตรวจจับและทำนายเส้นทางด้วย เนื่องจากระยะทางหรือพื้นที่ของการบดบังที่มีลักษณะกว้างมากขึ้นและอาจมีระยะเวลาที่นานมากขึ้นซึ่งทำให้วิธีการตรวจจับ การติดตามและทำนายเส้นทางแบบเดิมไม่สามารถจัดการได้

วิดีโอการแข่งขันกีฬาประเภททีมโดยเฉพาะกีฬาฟุตบอลโดยส่วนใหญ่จะมีการเคลื่อนที่ไปมาอยู่ตลอดเวลาจึงมีโอกาสที่จะเกิดปัญหาการบดบังกันระหว่างผู้เล่นทำให้ผู้เล่นที่ถูกบังไม่สามารถตรวจจับและติดตามได้ในบางช่วงเวลาซึ่งมีหลายงานวิจัยทางด้านการตรวจจับและติดตามที่พยามยามเสนอวิธีในการแก้ปัญหา เช่น การใช้ข้อมูลจากกล้องหลายตัวแทนการใช้กล้องเพียงตัวเดียว เช่น ในงานวิจัยของ Xu และคณะ [58] และงานวิจัยของ Iwase และ Saito [42] ซึ่งวิธีนี้ช่วยแก้ปัญหาการบดบังของผู้เล่นจากบางกล้องโดยใช้ข้อมูลจากกล้องตัวอื่นที่สามารถมองเห็นได้ ส่วนงานวิจัยของ Needham และ Boyle [73] ได้เสนอวิธีบนพื้นฐานของอัลกอริทึมคอนเดนเซชันซึ่งช่วยเพิ่มประสิทธิภาพในการติดตามผ่านการบดบังได้

ในบางช่วงเวลาของเกมส์การแข่งขันฟุตบอลจะมีโอกาสที่จะเกิดการรวมกลุ่มของผู้เล่นจำนวนมากซึ่งจะส่งผลให้เกิดการบดบังที่สมบูรณ์มากขึ้นทำให้มีผู้เล่นบางคนหรือหลายคนที่ไม่สามารถมองเห็นได้จากกล้องใดๆ ในงานวิจัยของ Iwase และ Saito [42] ได้เสนอแนวทางแก้ปัญหาการบดบังกรณีที่เกิดจากการรวมกลุ่มของผู้เล่นจำนวนมากโดยสันนิษฐานว่าการบดบังนั้นจะผ่านไปเมื่อผู้เล่นเคลื่อนที่ผ่านบริเวณที่ถูกบดบังซึ่งตำแหน่งและเส้นทางก็จะปรากฏขึ้นอีกครั้ง

ในงานวิจัยทางด้านการติดตามกลุ่มคน (Crowd) เช่น งานวิจัยของ Rodriguez และคณะ [106] ได้แบ่งลักษณะของกลุ่มคนที่มีการเคลื่อนที่ออกเป็น 2 ประเภท คือ แบบมีโครงสร้างและแบบไม่มีโครงสร้าง การเคลื่อนที่แบบมีโครงสร้างจะมีลักษณะของทิศทางและพฤติกรรมทั่วไปที่มีทิศทางเดียวกัน ส่วนแบบไม่มีโครงสร้างทิศทางการเคลื่อนที่จะมีลักษณะที่หลากหลายซึ่งการทำนายเส้นทางหรือพฤติกรรมของกลุ่มจะค่อนข้างยากกว่าแบบแรก รูปแบบของการโมเดลพฤติกรรมกลุ่มคนงานวิจัยของ Stuart [107] ได้อธิบายไว้ว่ามี 3 รูปแบบ แบบแรก คือ Microscopic โดยจะโมเดลการ

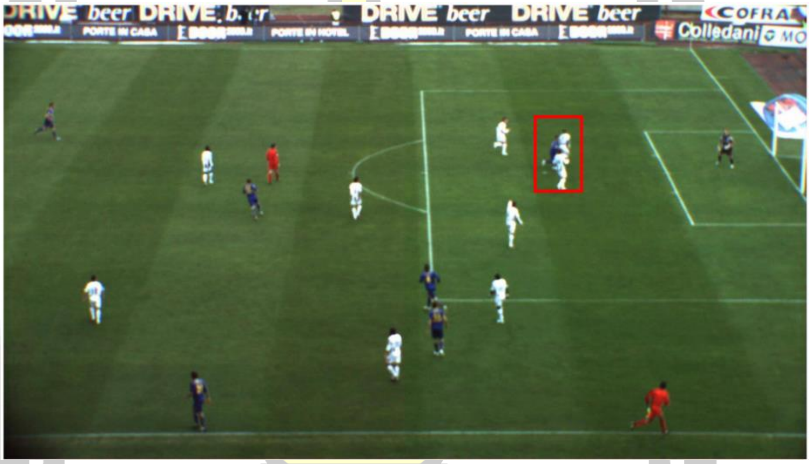
กระทำ (Action) การโต้ตอบกับผู้อื่น การโต้ตอบกับสภาพแวดล้อมของตัวแทน (Agent) ของแต่ละคนในกลุ่ม วิธีนี้จะมีความถูกต้องมากที่สุด ทำงานได้ดีสำหรับการโมเดลพฤติกรรมที่แตกต่างกันรวมถึงพฤติกรรมที่แสดงในการโต้ตอบหรือปฏิสัมพันธ์ (Interactions) ขนาดเล็กแต่จะมีความซับซ้อนในการคำนวณ แบบที่สอง คือ Mesoscopic ตัวแทนจะไม่ถูกมองเป็นแต่ละคนในระยะยาวแต่จะมองเป็นลักษณะของคลาส (Class) หรือกลุ่ม (Group) ด้วยพฤติกรรมที่ถูกแสดงในลักษณะกระจาย เช่น ความเร็ว (Velocity) หรือกระแส/การไหล (Flow) และแบบสุดท้าย คือ Macroscopic ตัวแทนทั้งหมดจะถูกมองเป็นก้อนเดียวกัน (Homogeneously) และแนวคิดของการเคลื่อนที่ (Movement) จะถูกกำหนดเป็นลักษณะเฉพาะในเรื่องของกระแส (Flow) ความหนาแน่น (Density) และความเร็วเฉลี่ย (Average Velocity) วิธีนี้พยายามรวมข้อดีของสองแบบแรก คือ ให้ได้พฤติกรรมและลักษณะเฉพาะของระดับ Microscopic ที่มีความง่ายและพฤติกรรมกลุ่มของระดับ Macroscopic โมเดลหนึ่งที่นิยมใช้ในการโมเดลพฤติกรรมของกลุ่มคน คือ โซเชียลฟอกซ์โมเดล (Social Force Model) ซึ่งมักใช้ในการโมเดลกลุ่มคนแบบ Microscopic โดยมีพื้นฐานอยู่บนผลรวมของ 3 แรง คือ ทิศทางการเคลื่อนที่ที่ต้องการ ปฏิสัมพันธ์ของคนกับคนอื่นๆ ที่อยู่รอบตนเองและปฏิกิริยาที่มีต่อสภาพแวดล้อมรอบตนเอง

จากการศึกษาข้อมูลพื้นฐานในเบื้องต้นในงานวิจัยนี้จึงให้ความสนใจกับปัญหาการบดบังที่เกิดจากการจับกลุ่มหรือการรวมกลุ่มของผู้เล่นเป็นสำคัญโดยเชื่อว่าถ้าสามารถจัดการปัญหาการจับกลุ่มหรือการรวมกลุ่มของผู้เล่นได้จะส่งผลต่อประสิทธิภาพของกระบวนการจัดการปัญหาการบดบังได้ โดยในงานวิจัยนี้มีแนวคิดที่จะอาศัยข้อมูลพื้นฐานที่ได้จากมุมมองเสมือนมาช่วยในการแก้ปัญหาการบดบัง ซึ่งสันนิษฐานว่าข้อมูลพื้นฐานที่ได้จากการผสมข้อมูลจากหลายกล้องรวมถึงผลจากการทำนายเส้นทางในมุมมองเสมือนจะสามารถช่วยในการประมาณตำแหน่งของผู้เล่นที่ถูกบดบังภายในกลุ่มได้

บทที่ 3

วิธีดำเนินงานวิจัย (Research Methodology)

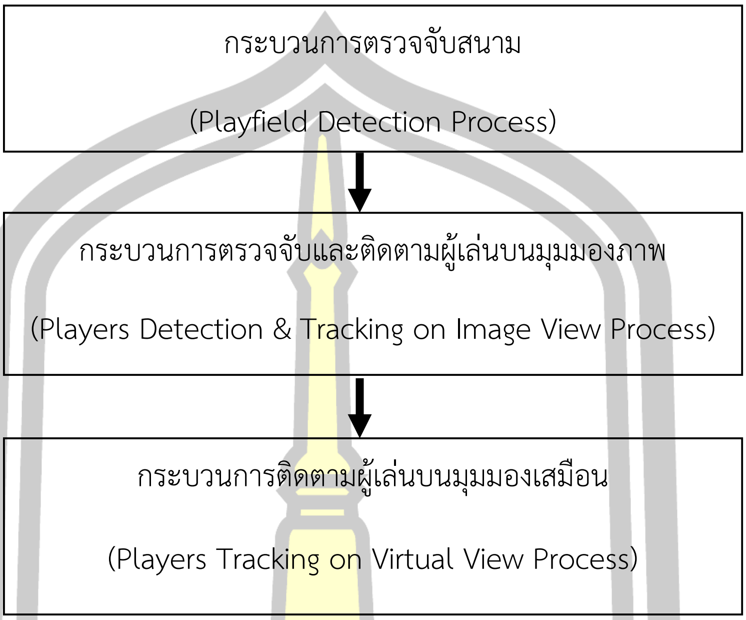
การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อแก้ปัญหาการบดบังซึ่งเกิดจากเหตุการณ์การรวมกลุ่มของนักฟุตบอลในวิดีโอการแข่งขันฟุตบอลตัวอย่างแสดงดังภาพที่ 3.1



ภาพที่ 3.1 การบดบังที่เกิดจากการรวมกลุ่ม

ตามภาพที่ 3.1 เป็นตัวอย่างของการบดบังระหว่างผู้เล่น ซึ่งเกิดจากเหตุการณ์การรวมกลุ่มของผู้เล่น ณ บริเวณพื้นที่กรอบสี่เหลี่ยมสีแดง ดังนั้นเพื่อให้บรรลุวัตถุประสงค์ของงานวิจัยที่ตั้งไว้ดังกล่าว คณะผู้วิจัยจึงเสนอกระบวนการดำเนินงานวิจัยโดยแสดงดังภาพที่ 3.2 ซึ่งประกอบด้วย 3 กระบวนการหลัก ได้แก่ 1. กระบวนการตรวจจับสนาม (Playfield Detection Process) 2. กระบวนการตรวจจับและติดตามผู้เล่นบนมุมมองภาพ (Players Detection & Tracking on Image View Process) และ 3. กระบวนการติดตามผู้เล่นบนมุมมองเสมือน (Players Tracking on Virtual View Process)

พฤาน ปณฺ ฑิโต ชีเว



ภาพที่ 3.2 กระบวนการดำเนินงานวิจัย

3.1 กระบวนการตรวจจับสนาม (Playfield Detection Process)

จุดประสงค์ของการตรวจจับสนาม เพื่อค้นหาพื้นที่ของบริเวณสนามที่คาดว่าจะปรากฏภายในภาพอินพุต ตัวอย่างพื้นที่บริเวณสนามที่สนใจแสดงดังภาพที่ 3.3 ในบริเวณพื้นที่กรอบสี่เหลี่ยมสีแดง จากนั้นจะใช้ข้อมูลขอบเขตของพื้นที่สนามที่ได้นี้ในการจำกัดขอบเขตหรือพื้นที่ในการตรวจจับและค้นหาผู้เล่นในภาพอินพุตซึ่งอยู่ในกระบวนการที่สอง นอกจากนี้การใช้ข้อมูลขอบเขตของพื้นที่สนามเป็นเกณฑ์ในการตรวจจับผู้เล่นภายในภาพอินพุตยังช่วยให้สามารถลดข้อผิดพลาดของกระบวนการตรวจจับผู้เล่นได้ในเบื้องต้นด้วย เพราะการกำหนดขอบเขตการตรวจจับผู้เล่นให้อยู่เฉพาะในพื้นที่ของสนามจะช่วยลดปัญหาการตรวจจับวัตถุอื่นๆ ที่ไม่เกี่ยวข้องภายในภาพอินพุตออกไปได้ เช่น ป้ายโฆษณา ผู้ชมและต้นไม้ เป็นต้น ดังนั้นวัตถุที่ถูกตรวจจับได้ในกระบวนการตรวจจับผู้เล่นจึงมีความเป็นไปได้สูงที่จะเป็นผู้เล่นมากกว่าวัตถุอื่นๆ ซึ่งส่งผลให้ช่วยลดปัญหาในการจำแนกผู้เล่นในขั้นตอนถัดไปได้อีกด้วย

ภาพที่ 3.3 พื้นที่บริเวณสนาม

กระบวนการตรวจจับสนาม ประกอบด้วย 4 ขั้นตอนย่อยดังนี้

1. ขั้นตอนจัดเตรียมข้อมูลก่อนการประมวลผล (Data Pre-Processing)

2. ขั้นตอนการสกัดเอกลักษณ์ (Feature Extraction)

3. ขั้นตอนการจำแนกข้อมูล (Data Classification)

4. ขั้นตอนการปรับแต่งข้อมูล (Data Refinement)

การดำเนินการในการตรวจจับสนามมีหลักการสำคัญ คือ ทำการนอร์มัลไลเซชัน (Normalization) ชุดภาพที่ใช้สอนให้มีรูปแบบของลักษณะสีที่เป็นรูปแบบเดียวกัน เนื่องจากภาพอาจจะมีความหลากหลายซึ่งเกิดจากเรื่องของแสงและสี จากนั้นใช้โมเดลผสมแบบเกาเซียนในการจำแนกข้อมูลซึ่งจะดำเนินการในระดับพิกเซล (Pixel-Level) โดยที่พิกเซลที่อยู่ภายในภาพจะถูกจำแนกออกเป็น 2 กลุ่ม คือ พิกเซลที่เป็นสนามและพิกเซลที่ไม่ใช่สนาม การจำแนกนั้นใช้ตัวอธิบาย หรือ เอกลักษณ์ (Descriptors or Features) 2 แบบ ได้แก่ ความหนาแน่นของสีของพิกเซล (Color Intensity) และพื้นผิว (Texture) บริเวณรอบข้างของพิกเซล การสร้างโมเดลสำหรับการจำแนกข้อมูล ในงานวิจัยนี้ใช้ข้อมูลความน่าจะเป็นก่อนหน้า (Prior Information) เกี่ยวกับตำแหน่งของบริเวณสนามที่ปรากฏภายในภาพเป็นข้อมูลพื้นฐานในการสร้างโมเดล หลังจากกระบวนการจำแนกข้อมูลเสร็จสิ้นจึงใช้วิธีการปรับปรุงการจำแนก (Refinement Method) เพื่อเพิ่มประสิทธิภาพผลลัพธ์ของการจำแนกอีกครั้ง โดยจะใช้ประโยชน์จากข้อมูลเชิงบริบท (Contextual Information) ของความสัมพันธ์เชิงพื้นที่ (Spatial Relationship) ของพิกเซลที่ปรากฏภายในภาพ

ข้อมูลที่ใช้ในการทดลองประกอบด้วย ชุดภาพนิ่งสำหรับใช้ในการสอนระบบ ชุดภาพนิ่งสำหรับการทดสอบและประเมินประสิทธิภาพของระบบที่พัฒนา ภาพที่ใช้สอนและทดสอบระบบนี้เป็นชุดภาพที่ได้จากแหล่งต่างๆ ในอินเตอร์เน็ตซึ่งลักษณะของภาพจะเป็นภาพที่มีพื้นที่ของสนามฟุตบอลปรากฏอยู่ในตำแหน่งต่างๆ ของภาพ หลังจากพัฒนาระบบเสร็จสิ้นจะนำภาพจากเฟรมแรกของวิดีโอการแข่งขันฟุตบอลจากฐานข้อมูล ISSIA Dataset [82] จากกล้องแต่ละตัวโดยชุดวิดีโอนี้มีการใช้กล้องจำนวน 6 ตัวในการบันทึกภาพการแข่งขันซึ่งจะได้ภาพจำนวน 6 ภาพ มาทำการตรวจจับสนาม ผลจากการตรวจจับสนามในแต่ละภาพนี้จะได้ข้อมูลพื้นที่ของบริเวณสนาม จากนั้นคำนวณหาขอบเขตคร่าวๆ ของสนามซึ่งเอาท์พุตที่ได้ คือ รายการพิกัดตำแหน่งทั้ง 4 มุมของพื้นที่สี่เหลี่ยมรอบพื้นที่สนาม เอาท์พุตนี้จะถูกนำไปใช้ในการกำหนดขอบเขตของพื้นที่ภายในเฟรมภาพอินพุตของวิดีโอสำหรับการตรวจจับผู้เล่นในกระบวนการตรวจจับและติดตามผู้เล่นในมุมมองภาพต่อไป รายละเอียดของกระบวนการตรวจจับสนามจะกล่าวถึงในบทที่ 4

3.2 กระบวนการตรวจจับและติดตามผู้เล่นในมุมมองภาพ (Player Detection & Tracking on Image View Process)

จุดประสงค์ของการตรวจจับผู้เล่น เพื่อค้นหาวัตถุที่ปรากฏภายในภาพที่คาดว่าจะเป็นผู้เล่นซึ่งเป็นวัตถุที่ให้ความสนใจและใช้ข้อมูลพื้นฐานต่างๆ ของวัตถุที่ตรวจจับได้นี้ในการดำเนินการต่างๆ เช่น การติดตามเส้นทางการเคลื่อนที่ การจำแนกตัวตนของแต่ละวัตถุ เป็นต้น ส่วนการติดตามผู้เล่นมีวัตถุประสงค์เพื่อจัดเก็บข้อมูลเส้นทางที่วัตถุเคลื่อนที่เพื่อใช้ในการวิเคราะห์เส้นทางการเคลื่อนที่ของแต่ละวัตถุและใช้ในการทำนายตำแหน่งถัดไปที่คาดว่าวัตถุจะเคลื่อนที่ไป นอกจากนี้ยังสามารถใช้เป็นข้อมูลพื้นฐานในการทำนายเส้นทางการเคลื่อนที่เมื่อเกิดเหตุการณ์การบดบังระหว่างผู้เล่นได้ กระบวนการตรวจจับและติดตามผู้เล่นบนมุมมองภาพ ประกอบด้วย 5 ขั้นตอนย่อยดังนี้

1. ขั้นตอนการตรวจจับผู้เล่น (Player Detection)
2. ขั้นตอนการจำแนกผู้เล่น (Player Classification)
3. ขั้นตอนการแสดงผลผู้เล่น (Player Representation)
4. ขั้นตอนการติดตามผู้เล่น (Player Tracking)
5. ขั้นตอนการแสดงเส้นทางผู้เล่นในมุมมองภาพ (Image View based Player Track Presentation)

การดำเนินการในการตรวจจับและติดตามผู้เล่นมีหลักการ คือ หลังจากการตรวจจับสนามในเฟรมแรกของชุดวิดีโอการแข่งขันฟุตบอลจะได้ข้อมูลของขอบเขตของพื้นที่สนามภายในภาพอินพุต

พหูน ชนฺ ฅาโต ชีเว

ข้อมูลขอบเขตของพื้นที่สนามนี้จะถูกนำมาใช้เป็นพารามิเตอร์สำหรับการครอป (Crop) พื้นที่ภายในภาพอินพุตซึ่งเป็นภาพเฟรมแรกของวิดีโอ เอาท์พุตที่ได้ คือ ภาพที่มีพื้นที่ของสนามปรากฏอยู่ ภาพเอาท์พุตนี้จะถูกนำไปใช้ในการทำแบ็คกราวนด์ซับแทรคชันในเฟรมถัดไป เพื่อตรวจจับผู้เล่นที่ปรากฏอยู่ภายในภาพอินพุต เอาท์พุตที่ได้หลังจากการทำแบ็คกราวนด์ซับแทรคชัน คือ รายการของบล็อบวัตถุ จากนั้นทำการจำแนกกลุ่มของวัตถุที่ได้นี้ออกเป็น 2 กลุ่มหลัก คือ กลุ่มที่คาดว่าจะเป็นผู้เล่นและกรรมการและกลุ่มที่คาดว่าจะไม่ใช่ผู้เล่น โดยใช้เอกลักษณ์สี ขนาดและการเคลื่อนที่เป็นเกณฑ์ในการพิจารณา โดยในงานวิจัยนี้จะพิจารณาเพียง 23 คน คือ ผู้เล่นทั้ง 2 ทีมและกรรมการตัดสิน ดังนั้นในขั้นตอนการจำแนกนี้จะคัดเลือกเฉพาะ 23 วัตถุแรกที่มีความน่าจะเป็นว่าจะเป็นผู้เล่นและกรรมการตัดสิน ส่วนวัตถุอื่นๆ ที่เหลือจะไม่ถูกนำมาพิจารณา สุดท้ายรายการวัตถุที่ได้นี้จะถูกนำไปใช้สำหรับขั้นตอนการติดตามต่อไป โดยในงานนี้จะประยุกต์ใช้อัลกอริทึมคอนเดนเซชันร่วมกับเอกลักษณ์สีและเทคนิคการสุ่มแซมเปิลบนพื้นฐานของพื้นที่สำหรับการติดตามผู้เล่น

ข้อมูลที่ใช้ในการทดลองประกอบด้วย ชุดภาพวิดีโอการแข่งขันฟุตบอลจาก ISSIA Dataset [82] จากกล้องตัวหนึ่งจากทั้งหมด 6 ตัว โดยเอาท์พุตที่ได้จากขั้นตอนการตรวจจับผู้เล่น คือ รายการของบล็อบวัตถุที่คาดว่าจะเป็นผู้เล่นและกรรมการตัดสิน ส่วนเอาท์พุตที่ได้จากขั้นตอนการติดตามผู้เล่น คือ ข้อมูลตำแหน่งและเส้นทางการเคลื่อนที่ของผู้เล่นแต่ละคน โดยข้อมูลทั้ง 2 ส่วนนี้จะถูกนำไปใช้ในกระบวนการติดตามผู้เล่นในมุมมองเสมือนต่อไป รายละเอียดของกระบวนการตรวจจับและติดตามผู้เล่นในมุมมองภาพจะกล่าวถึงในบทที่ 5

3.3 กระบวนการติดตามผู้เล่นในมุมมองเสมือน (Player Tracking on Virtual View Process)

จุดประสงค์ของการติดตามผู้เล่นในมุมมองเสมือน เพื่อแสดงข้อมูลพื้นฐานของผู้เล่นได้แก่ ตำแหน่งและเส้นทางการเคลื่อนที่ของผู้เล่นในภาพรวมทั้งสนาม เนื่องจากข้อมูลจากกล้องเพียงตัวเดียวไม่สามารถแสดงข้อมูลของผู้เล่นทุกคนที่ปรากฏอยู่ภายในสนามได้และการใช้กล้องหลายตัวในการบันทึกภาพผู้เล่นจะเกิดปัญหาในการจำแนกผู้เล่นระหว่างกล้อง ในการแสดงผลจึงจำเป็นต้องทำการรวมข้อมูลจากทุกกล้องเป็นข้อมูลเดียว นอกจากนี้การแสดงผลในมุมมองเสมือนยังสามารถช่วยให้ผู้ชมสามารถเห็นภาพรวมของการแข่งขันได้ อีกทั้งผู้ควบคุมทีมยังสามารถใช้ข้อมูลพื้นฐานนี้ในการวิเคราะห์เทคนิคการเล่นของผู้เล่นแต่ละคนและของทีม รวมถึงการวางแผนการเล่นของทีมได้อีกด้วย กระบวนการติดตามผู้เล่นบนมุมมองเสมือน ประกอบด้วย 9 ขั้นตอนย่อยดังนี้

1. ขั้นตอนการจัดเตรียมข้อมูลก่อนการประมวลผล (Data Pre-Processing)

2. ขั้นตอนการสกัดเอกลักษณ์ (Feature Extraction)

3. ขั้นตอนการแปลงมุมมอง (Perspective Transform)

4. ขั้นตอนการผสมข้อมูล (Data Fusion)

5. ขั้นตอนการเชื่อมโยงข้อมูล (Data Association)

6. ขั้นตอนการเลือกเป้าหมาย (Target Selection)

7. ขั้นตอนการระบุตัวตนผู้เล่น (Player Identification)

8. ขั้นตอนการแสดงผลผู้เล่น (Player Representation)

9. ขั้นตอนการติดตามผู้เล่น (Player Tracking)

การดำเนินการในการติดตามผู้เล่นในมุมมองเสมือนมีหลักการ คือ หลังจากกระบวนการตรวจจับและติดตามผู้เล่นในมุมมองภาพแล้ว เอาท์พุตที่ได้ คือ ตำแหน่งและเส้นทางการเคลื่อนที่ของผู้เล่นแต่ละคนจากกล้องแต่ละตัว ข้อมูลตำแหน่งนี้จะถูกโปรเจ็คลงบนภาพเสมือนตามหลักการของการแปลงมุมมอง ขั้นตอนต่อมา คือ การรวมข้อมูลตำแหน่งที่โปรเจ็คจากทุกกล้องนี้เป็นข้อมูลเดียวกันซึ่งในงานวิจัยนี้ได้เสนอเทคนิคการผสมข้อมูลโดยใช้โมเดลรูปร่างของข้อมูลบนพื้นฐานทางกายภาพ ส่วนการประมาณตำแหน่งผู้เล่นบนมุมมองเสมือนได้เสนอเทคนิคการประมาณตำแหน่งโดยอาศัยพื้นฐานของสมการเส้นตรง ในขั้นตอนสุดท้ายเอาท์พุตตำแหน่งผู้เล่นที่ได้จากการผสมข้อมูลและการประมาณตำแหน่งนี้จะถูกนำมาใช้กับคาร์มาลฟิลเตอร์ในการติดตามผู้เล่นบนมุมมองภาพเสมือน ข้อมูลตำแหน่งและเส้นทางบนมุมมองเสมือนนี้จะถูกนำไปใช้ในการแก้ปัญหากรณีเกิดเหตุการณ์การบดบังที่เกิดจากการรวมกลุ่มของผู้เล่นในมุมมองภาพของแต่ละกล้องโดยข้อมูลตำแหน่งจากกล้องที่สามารถมองเห็นผู้เล่นได้จะช่วยเพิ่มความถูกต้องในการประมาณตำแหน่งผู้เล่นภายในกล้องที่เกิดเหตุการณ์การบดบังได้

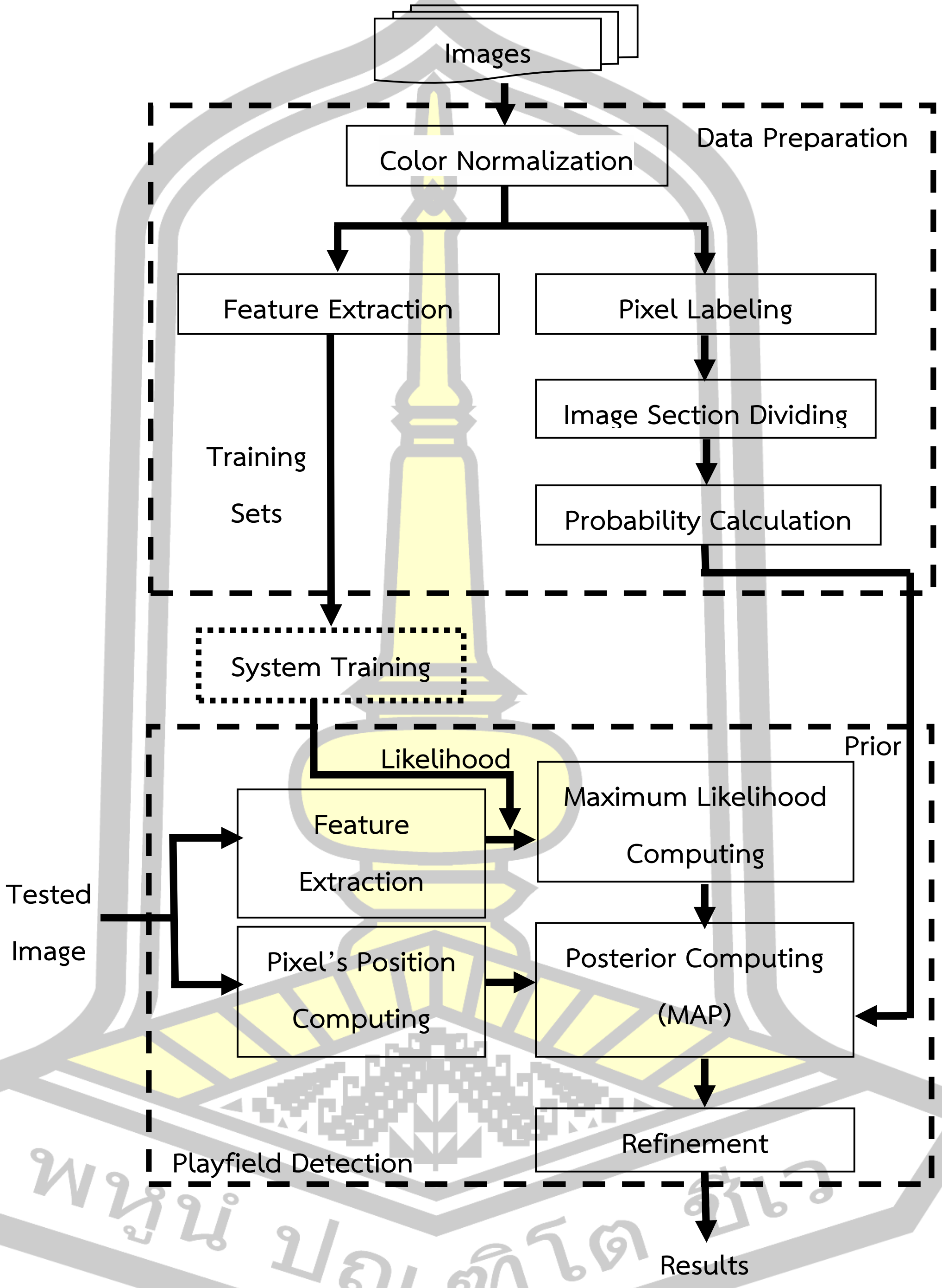
ข้อมูลที่ใช้ในการทดลองประกอบด้วย ชุดภาพนิ่งที่ปรากฏตำแหน่งของวัตถุจำนวนหนึ่ง โดยชุดภาพนิ่งนี้ได้มาจากการถ่ายภาพในมุมมองต่างๆ จำนวน 9 มุมมองตามทิศต่างๆ ที่ดำเนินการด้วยมือ โดยมุมมองที่ 9 กล้องจะอยู่ในตำแหน่งด้านบนเหนือวัตถุซึ่งข้อมูลตำแหน่งของกล้องตัวที่ 9 นี้จะถูกใช้เป็นกราวนด์ทรูธสำหรับการประเมินประสิทธิภาพของเทคนิคการประมาณตำแหน่งข้อมูลบนภาพเสมือน ส่วนชุดข้อมูลภาพจากวิดีโอ ISSIA Dataset [82] จะถูกใช้เพื่อประเมินประสิทธิภาพของเทคนิคการผสมข้อมูลที่เสนอเพื่อศึกษาว่าเทคนิคที่เสนอสามารถผสมข้อมูลจากแต่ละกล้องได้ถูกต้องหรือไม่ รายละเอียดของกระบวนการติดตามผู้เล่นในมุมมองเสมือนจะกล่าวถึงในบทที่ 6

บทที่ 4

การตรวจจับสนาม (Playfield Detection)

4.1 วิธีดำเนินการวิจัย (Research Methodology)

ในบทนี้ได้นำเสนอเทคนิคการตรวจจับสนามฟุตบอลโดยใช้เทคนิคของการเรียนรู้ของเครื่องและเทคนิคการประมวลผลภาพ (Image Processing and Machine Learning Techniqjues) ซึ่งจะประกอบด้วย 4 ขั้นตอนหลัก ได้แก่ ขั้นตอนแรก คือ ขั้นตอนจัดเตรียมข้อมูลก่อนการประมวลผล (Data Pre-Processing) ประกอบด้วย 5 ขั้นตอนย่อย ได้แก่ การทำการนอร์มัลไลซ์ภาพ (Color Normalization) การกำหนดลาเบลพิกเซล (Pixel Labeling) การแบ่งส่วนของภาพ (Image Section Dividing) การสอนระบบ (System Training) และการคำนวณค่าความน่าจะเป็น (Probability Calculation) ขั้นตอนที่สอง คือ ขั้นตอนการสกัดเอกลักษณ์ (Feature Extraction) ขั้นตอนที่สาม คือ ขั้นตอนการจำแนกข้อมูล (Data Classification) ประกอบด้วย 3 ขั้นตอนย่อย ได้แก่ การคำนวณตำแหน่งของพิกเซล (Pixel’s Position Computing) การคำนวณค่าความน่าจะเป็นสูงสุด (Maximum Likelihood Computing) และการคำนวณค่าความน่าจะเป็นครั้งหลัง (Posterior Computing) และ ขั้นตอนสุดท้าย คือ ขั้นตอนการปรับแต่งข้อมูล (Data Refinement) โดยขั้นตอนทั้งหมดแสดงดังภาพที่ 4.1 ในขั้นตอนแรกจะทำการแปลงค่าสี (Color Transformation) ของภาพเพื่อลดความแปรปรวนของสี (Color Variation) ที่เกิดขึ้นในภาพ จากนั้นจึงทำการสกัดเอกลักษณ์ (Features) โดยในงานนี้ใช้ 2 เอกลักษณ์หลัก ได้แก่ ความเข้มของค่าสีของสีและพื้นผิวบริเวณรอบพิกเซล ในขั้นตอนการจำแนกข้อมูล โมเดลการทำนายจะถูกสร้างขึ้นโดยใช้อัลกอริทึมนาอีฟเบย์ (Naive Bayes Algorithm) โมเดลนี้จะใช้ข้อมูลความน่าจะเป็นก่อนหน้า (Prior Information) เกี่ยวกับตำแหน่งของบริเวณสนามที่ปรากฏภายในภาพในการดำเนินการ ภายหลังขั้นตอนการจำแนกข้อมูลเสร็จสิ้น ในขั้นตอนสุดท้ายจะใช้วิธีการปรับแต่งข้อมูลเพื่อปรับปรุงประสิทธิภาพของผลลัพธ์ที่ได้จากขั้นตอนการจำแนกข้อมูลอีกครั้งโดยใช้ประโยชน์จากข้อมูลเชิงบริบทของความสัมพันธ์เชิงพื้นที่ของพิกเซลที่ปรากฏภายในภาพในการดำเนินการซึ่งในขั้นตอนนี้จะใช้เทคนิคที่เรียกว่า มาร์คอฟ แรนดอมฟิวส์ (Markov Random Filed Technique) และเทคนิคที่เรียกว่า มีเดียนฟิลเตอร์ (Median Filter) ในการทดลอง จุดประสงค์เพื่อหาเทคนิคที่เหมาะสมสำหรับใช้ในงานวิจัย รายละเอียดของการดำเนินงานในแต่ละขั้นตอนประกอบด้วยหัวข้อดังต่อไปนี้



ภาพที่ 4.1 กระบวนการตรวจจับสนาม

1. ขั้นตอนการจัดเตรียมข้อมูลก่อนการประมวลผล (Data Pre-Processing)

ในขั้นตอนนี้จะดำเนินการจัดเตรียมข้อมูลภาพเพื่อใช้สำหรับการสอนและการทดสอบ โดยในงานนี้จะประยุกต์ใช้โมเดลผสมแบบเกาเซียนร่วมกับทฤษฎีความน่าจะเป็นในการตรวจจับสนาม โมเดลความน่าจะเป็นของพิกเซลสนามจะดำเนินการโดยแบ่งภาพออกเป็น 3 ส่วน ตามแนวนอน แสดงดังภาพที่ 4.5 โดยสันนิษฐานว่าภาพแบบ Long View นั้นบริเวณที่เป็นสนามมักจะอยู่ในบริเวณส่วนล่างสุดของภาพเป็นส่วนใหญ่รองลงมาจะเป็นส่วนกลางภาพและส่วนบนสุดของภาพตามลำดับ ดังนั้นพิกเซลใดที่อยู่ในส่วนล่างสุดของภาพจะให้น้ำหนักว่าเป็นสนามมากกว่าค่าน้ำหนักที่ไม่ใช่สนาม ส่วนพิกเซลที่อยู่ในบริเวณส่วนบนสุดของภาพจะให้น้ำหนักว่าเป็นสนามน้อยกว่าค่าน้ำหนักที่ไม่ใช่สนาม ค่าน้ำหนักนี้ คือ ค่าความน่าจะเป็นของพิกเซลทั้งที่เป็นสนามและไม่ใช่สนามซึ่งหาได้จากชุดข้อมูลสอนที่ถูกกำหนดลาเบล โดยค่าน้ำหนักนี้จะถูกใช้ร่วมกับโมเดลผสมแบบเกาเซียนเพื่อเพิ่มความแม่นยำในการตรวจจับพิกเซลที่เป็นสนาม

การจัดเตรียมข้อมูลภาพเพื่อใช้สำหรับการสอนและการทดสอบระบบมีรายละเอียดของการดำเนินการดังนี้

1.1 ทำการจัดเตรียมข้อมูลภาพจำนวน 2 ชุด ได้แก่ ชุดข้อมูลภาพสำหรับใช้ในการสอนระบบและชุดข้อมูลภาพสำหรับใช้ในการทดสอบระบบ

1.2 ทำการนอร์มัลไลซ์ (Normalized) รายการภาพทั้งหมด จุดประสงค์เพื่อลดความแปรปรวนของสีที่เกิดขึ้นในภาพ ซึ่งโดยปกติความแปรปรวนนี้จะเกิดขึ้นเมื่อมีการเปลี่ยนแปลง (Permutation) เอกลักษณ์บางอย่างของภาพ หลักการแปลงภาพนี้จะประยุกต์มาจากงานวิจัยของ Reinhard และคณะ [88] โดยการทำนอร์มัลไลซ์ภาพมีหลักการดังนี้ คือ ภาพทั้งหมดซึ่งในที่นี้จะแทนด้วยสัญลักษณ์ I จะถูกแปลง (Transformation) ไปเป็นภาพที่มีลักษณะเฉพาะที่เหมาะสมกับงานวิจัยโดยในที่นี้ คือ การปรับคุณภาพของสีโดยเฉพาะสีของสนามฟุตบอลเพื่อให้เอกลักษณ์ของสีของพิกเซลมีความชัดเจนและมีความเหมาะสมต่อการดำเนินการ โดยภาพที่ได้นี้จะมีเอกลักษณ์ที่คล้ายกับภาพที่เรียกว่า ภาพแม่แบบหรือเทมเพลต (Template Image) ซึ่งในที่นี้จะแทนด้วยสัญลักษณ์ T โดยภาพเทมเพลตนี้ผู้วิจัยได้ทำการเลือกโดยใช้คนในการพิจารณาซึ่งในที่นี้จะใช้เพียง 1 ภาพ วิธีการ คือ จะใช้ฟังก์ชันแมพปิ้ง (Mapping Function) ในการเปลี่ยนแปลงคุณสมบัติของภาพอินพุตไปเป็นภาพเป้าหมายหรือภาพเทมเพลต โดยที่คุณสมบัติของภาพในที่นี้ คือ ค่าความเข้มของค่าสีของสีของพิกเซลใน RGB Color Space โดยกระบวนการเปลี่ยนแปลงคุณสมบัตินี้จะมี 3 ขั้นตอนหลักๆ คือ

1) การปรับมาตรฐาน (Standardizing) ค่าของพิกเซลในช่องสี RGB ของภาพอินพุตโดยการลบค่าในแต่ละช่องสีนี้ด้วยค่าเฉลี่ยของแต่ละช่องสีตามสมการที่ 4.1 ดังนี้

$$
\begin{aligned}
I'_r &= I_r - \bar{I}_r \\
I'_g &= I_g - \bar{I}_g \\
I'_b &= I_b - \bar{I}_b
\end{aligned}
\tag{4.1}
$$

โดยที่ I_r, I_g, I_b คือค่าของพิกเซลภาพอินพุตในช่องสีแดง เขียวและน้ำเงินตามลำดับ ส่วน $\bar{I}_r, \bar{I}_g, \bar{I}_b$ คือค่าเฉลี่ยของภาพอินพุตในช่องสีแดง เขียวและน้ำเงินตามลำดับ

2) การปรับขนาด (Scaling) ของข้อมูลที่ผ่านการปรับมาตรฐานโดยใช้อัตราส่วน (Ratio) ของความแปรปรวน (Variance) ระหว่างภาพอินพุตกับภาพเทมเพลตตามสมการที่ 4.2 ดังนี้

$$
\begin{aligned}
I''_r &= \frac{\sigma_r^T}{\sigma_r^I} \bullet I'_r \\
I''_g &= \frac{\sigma_g^T}{\sigma_g^I} \bullet I'_g \\
I''_b &= \frac{\sigma_b^T}{\sigma_b^I} \bullet I'_b
\end{aligned}
\tag{4.2}
$$

โดยที่ I'_r, I'_g, I'_b คือภาพที่ได้จากการปรับมาตรฐานในช่องสีแดง เขียวและน้ำเงินตามลำดับ ส่วน σ_r^T, σ_g^T, σ_b^T คือค่าความแปรปรวนของภาพเทมเพลตในช่องสีแดง เขียวและน้ำเงินตามลำดับ และ σ_r^I, σ_g^I, σ_b^I คือค่าความแปรปรวนของภาพอินพุตในช่องสีแดง เขียวและน้ำเงินตามลำดับ

3) ทำการโปรเจ็ค (Projecting) ข้อมูลกลับไปยังเทมเพลตสเปซ (Template Space) โดยการบวกค่าเฉลี่ยของเทมเพลตเข้ากับภาพที่ถูกปรับขนาดตามสมการที่ 4.3 ดังนี้

$$
\begin{aligned}
\hat{I}_r &= I''_r + \bar{T}_r \\
\hat{I}_g &= I''_g + \bar{T}_g \\
\hat{I}_b &= I''_b + \bar{T}_b
\end{aligned}
\tag{4.3}
$$

พหูน ปณฺฑิโต ชีเว

โดยที่ I''_r, I''_g, I''_b คือภาพที่ได้จากการปรับขนาดในช่องสีแดง เขียวและน้ำเงินตามลำดับ ส่วน $\overline{T}_r, \overline{T}_g, \overline{T}_b$ คือค่าเฉลี่ยของภาพเทมเพลตในช่องสีแดง เขียวและน้ำเงินตามลำดับ

ตัวอย่างของการแปลงภาพที่ได้จากขั้นตอนดังกล่าวข้างต้นแสดงได้ดังภาพที่ 4.2

ภาพที่ 4.2 ตัวอย่างการแปลงภาพ

ตามภาพที่ 4.2 ภาพในคอลัมน์ซ้าย คือ ภาพเทมเพลตซึ่งถูกใช้เป็นภาพต้นแบบสำหรับการแปลงภาพ ส่วนภาพในคอลัมน์กลาง คือ ภาพอินพุตที่จะถูกทำการแปลงภาพให้มีคุณสมบัติเหมือนกับภาพเทมเพลต และภาพในคอลัมน์ขวา คือ ตัวอย่างภาพผลลัพธ์ที่ได้จากกระบวนการแปลงภาพ

1.3 การจัดเตรียมชุดข้อมูลสอนและประเมินระบบ ประกอบด้วยชุดข้อมูลภาพจำนวน 2 กลุ่ม คือ

1) ชุดข้อมูลภาพต้นฉบับ ใช้สำหรับการดึงค่าเอกลักษณ์พื้นฐานต่างๆ ของแต่ละพิกเซลในแต่ละภาพ ซึ่งในงานวิจัยนี้ใช้จำนวน 9 เอกลักษณ์ ได้แก่ ค่าเฉลี่ยของสีแดง ค่าเฉลี่ยของสีเขียว ค่าเฉลี่ยของสีน้ำเงิน ค่าความแปรปรวนของสีแดง ค่าความแปรปรวนของสีเขียว ค่าความแปรปรวนของสีน้ำเงิน ค่าสีแดง ค่าสีเขียวและค่าสีน้ำเงิน โดยกลุ่มพิกเซลที่เป็นสนามและกลุ่มพิกเซลที่ไม่ใช่สนามของแต่ละภาพจะถูกสุ่มเลือกเพื่อนำไปใช้สำหรับสอนระบบโดยจะคำนวณหาค่าเฉลี่ย $(\overline{\mu})$ และค่าความแปรปรวน (Σ) รวมของชุดภาพสำหรับสอนของทั้งสองกลุ่มซึ่งจะถูกนำไปใช้

เป็นโมเดลของระบบต่อไป ลักษณะของการดึงเอกลักษณ์ของภาพแสดงดังภาพที่ 4.3 ซึ่งแต่ละภาพจะได้รายการเอาท์พุตมา 2 ชุด โดยชุดแรกเป็นของกลุ่มพิกเซลสนามซึ่งประกอบด้วย รายการค่าเฉลี่ยของ 9 เอกลักษณ์และรายการค่าความแปรปรวนของ 9 เอกลักษณ์และชุดที่สองเป็นของกลุ่มพิกเซลไม่ใช่สนามซึ่งประกอบด้วย รายการค่าเฉลี่ยของ 9 เอกลักษณ์และรายการค่าความแปรปรวนของ 9 เอกลักษณ์ เป็นต้น

รายการพิกเซลสนาม				
Num.1	Value.1	Value.2	…	Value.9
Num.2	…	…	…	…
…	…	…	…	…
Num.N	Value.1	Value.2	…	Value.9

Mean Value	Value.1	Value.2	…	Value.9
Variance Value	Value.1	Value.2	…	Value.9

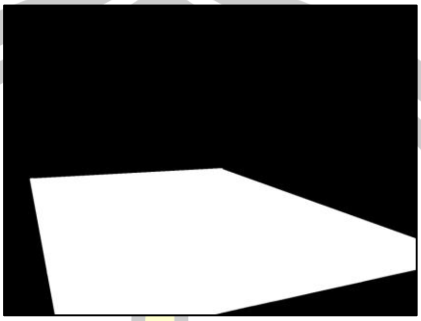
รายการพิกเซลไม่ใช่สนาม				
Num.1	Value.1	Value.2	…	Value.9
Num.2	…	…	…	…
…	…	…	…	…
Num.N	Value.1	Value.2	…	Value.9

Mean Value	Value.1	Value.2	…	Value.9
Variance Value	Value.1	Value.2	…	Value.9

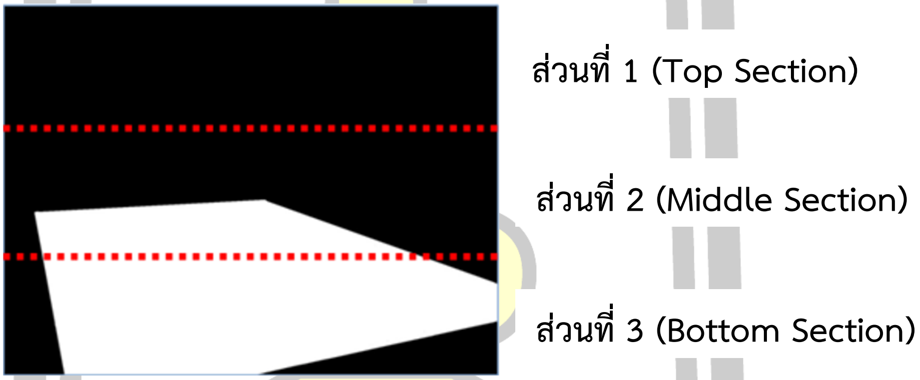
ภาพที่ 4.3 การดึงเอกลักษณ์จากภาพต้นฉบับ

2) ชุดข้อมูลภาพที่กำหนดลาเบล (Label) เป็นภาพต้นฉบับที่ทำการกำหนดลาเบลให้กับแต่ละพิกเซลเพื่อจัดกลุ่มของพิกเซลออกเป็น 2 กลุ่ม คือ กลุ่มที่เป็นสนามและกลุ่มที่ไม่ใช่สนาม วิธีการกำหนดลาเบลจะดำเนินการโดยใช้คนพิจารณา ซึ่งภาพกลุ่มนี้เป็นภาพผลเฉลย (Ground Truth) โดยจะใช้สำหรับการสอนระบบและใช้ในการประเมินระบบด้วย ตัวอย่างการกำหนดลาเบลแสดงดังภาพที่ 4.4 โดยพิกเซลในกลุ่มที่เป็นสนามจะกำหนดให้เป็นสีขาวและพิกเซลในกลุ่มที่ไม่ใช่สนามจะกำหนดให้เป็นสีดำ จากนั้นแต่ละภาพจะถูกแบ่งออกเป็น 3 ส่วนหลักๆ ตามแนวนอนของภาพแสดงดังภาพที่ 4.5 โดยในแต่ละส่วนจะใช้คำนวณหาค่าความน่าจะเป็นของกลุ่มพิกเซลที่เป็นสนามและกลุ่มพิกเซลที่ไม่ใช่สนาม ดังนั้นในแต่ละภาพจะได้ข้อมูลความน่าจะเป็นของพิกเซลกลุ่มที่เป็นสนามจำนวน 3 กลุ่มและข้อมูลความน่าจะเป็นของพิกเซลกลุ่มที่ไม่ใช่สนามจำนวน 3 กลุ่ม รวมทั้งสิ้น 6 กลุ่ม ข้อมูลความน่าจะเป็นทั้งหมดนี้ คือ ค่าความน่าจะเป็นจากเงื่อนไข (Prior) ซึ่งจะถูก

นำไปใช้สำหรับเพิ่มประสิทธิภาพของระบบในการทำนายพิกเซลสำหรับการค้นหาสนามในขั้นตอนการจำแนกข้อมูลต่อไป



ภาพที่ 4.4 ภาพต้นฉบับที่กำหนดลาเบล



ภาพที่ 4.5 การแบ่งส่วนภาพต้นฉบับที่กำหนดลาเบล

1.4 ทำการสอนระบบ โดยใช้ชุดข้อมูลของภาพจากข้อที่ 1.3 และการทดสอบระบบจะใช้วิธีของ Cross Validation

2. ขั้นตอนการสกัดเอกลักษณ์ (Feature Extraction)

ในขั้นตอนนี้ภาพจะถูกแบ่งออกเป็นพื้นที่ (Windows) ขนาด n x n พิกเซล ซึ่งพื้นที่เหล่านี้จะถูกใช้เพื่อสกัดเซตของเอกลักษณ์ซึ่งเป็นเอกลักษณ์ของพิกเซลที่อยู่ ณ ตำแหน่งศูนย์กลางของพื้นที่นั้นๆ โดยในงานวิจัยนี้ใช้ตัวอธิบาย (Descriptor) จำนวน 2 ตัวสำหรับใช้ในการจำแนกข้อมูลในแต่ละพิกเซล (Per-Pixel Classification) ได้แก่ เอกลักษณ์ความเข้มของค่าสีของสี (Color Intensity-based Feature) และเอกลักษณ์พื้นผิว (Texture-based Feature) ซึ่งมีรายละเอียดดังต่อไปนี้

2.1 เอกลักษณ์ความเข้มของค่าสีของสี (Color Intensity Feature)

ในแต่ละพื้นที่ภาพ (Image Window) จะถูกแปลงค่าสีไปเป็นระบบสี YCrCb Color-Space จากนั้นในแต่ละช่องสี คือ Cr และ Cb จะทำการคำนวณหาค่าเฉลี่ย (Average) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ของพิกเซลทั้งหมดในพื้นที่ภาพซึ่งผลลัพธ์ที่ได้จะถูกใช้เป็นตัวอธิบายข้อมูลของพิกเซลที่อยู่ ณ ตำแหน่งศูนย์กลางของพื้นที่นั้นๆ สำหรับช่องสี Y ที่เหลือในงานวิจัยนี้ไม่นำมาใช้ในการประมวลผลเนื่องจากเพื่อหลีกเลี่ยงความแปรปรวนของแสงซึ่งอาจจะส่งผลต่อประสิทธิภาพของระบบได้

2.2 เอกลักษณ์พื้นผิว (Texture Feature)

ตารางที่ 4.1 เอกลักษณ์ของ Haralick และคณะ [89]

เอกลักษณ์ (Features)	**คำอธิบาย (Description)**
ค่าเอนโทรปี (Entropy)	$-\sum_i\sum_j G_d(i,j)\log G_d(i,j)$
ค่าคอนทรัสท์ (Contrast)	$-\sum_i\sum_j (i-j)^2 G_d(i,j)$
โมเมนต์ความแตกต่างย้อนกลับ (Inverse Different Moment)	$\sum_i\sum_j \frac{G_d(i,j)}{1+(i,j)^2}$
โมเมนต์อันดับสองเชิงมุม (Angular Second Moment)	$\sum_i\sum_j G_d(i,j)^2$
ค่าเฉลี่ยเกรย์สเกล (Mean Gray Value)	$\frac{\sum_{i\in\Omega} g_i}{\lvert\Omega\rvert}$
ผลรวมความแปรปรวน (Sum Variance)	$\sum_{i=2}^{2N_g}(i-\sigma)^2 G_{i+j}(i)$
ผลรวมเอนโทรปี (Sum Entropy)	$-\sum_{i=2}^{2N_g} G_{i+j}(i)\log\{G_{i+j}(i)\}$
ผลต่างความแปรปรวน (Different Variance)	$\sum_{i=0}^{N_g-1} i^2 G_{i-j}(i)$
ผลต่างเอนโทรปี (Different Entropy)	$-\sum_{i=0}^{N_g-1} G_{i-j}(i)\log\{G_{i-j}(i)\}$

ในเอกลักษณ์พื้นผิวนี้จะใช้เมทริกซ์การแจกแจงร่วมระดับเกรย์สเกล (Gray Level Co-Occurrence Matrix: GLCM) โดยค่าเมทริกซ์การแจกแจงร่วมในที่นี้จะแทนด้วยสัญลักษณ์ G ของพิกเซลซึ่งในที่นี้จะแทนด้วยสัญลักษณ์ g ในแต่ละพื้นที่ภาพซึ่งในที่นี้จะแทนด้วยสัญลักษณ์ Ω จะถูกสร้างขึ้น แต่ละอีลีเมนท์ (Element) ในเมทริกซ์นี้จะแสดงถึงการเกิดร่วมกัน (Occurrence) ของค่าในพิกเซล ซึ่งกำหนดสัญลักษณ์ด้วย 0...N โดยที่ N คือ ตัวเลขของระดับค่า (Value Levels) และพิกเซลเพื่อนบ้าน (Neighboring Pixel) ในพื้นที่ภาพที่มีค่าออฟเซ็ต (Offset) ที่เฉพาะ (t=1)

และทิศทาง (Direction) (d=0°,45°,90°,135°) ดังนั้นจะได้เมทริกซ์ทั้งหมด 4 เมทริกซ์ หลังจากนั้นแต่ละเมทริกซ์การแจกแจงร่วมซึ่งในที่นี้จะแทนด้วยสัญลักษณ์ G_d จะถูกใช้สำหรับการสร้างคุณสมบัติทางสถิติ (Statistic Properties) ซึ่งจะถูกสกัดมาจากเมทริกซ์เหล่านี้ โดยคุณสมบัติทางสถิติที่ใช้ในงานวิจัยนี้จะแสดงดังตารางที่ 4.1 ขั้นตอนสุดท้าย เอกลักษณ์พื้นผิวจะได้มาโดยการจัดเรียง (Arranging) คุณสมบัติทางสถิติทั้งหมดของทิศทางทั้งหมด ซึ่งจะได้ค่าเอกลักษณ์ทั้งหมด 36 เอกลักษณ์สำหรับแต่ละพื้นที่ภาพ (พิกเซล)

3. ขั้นตอนการจำแนกข้อมูล (Data Classification)

หลังจากขั้นตอนการสกัดเอกลักษณ์เสร็จสิ้น เอกลักษณ์เหล่านี้จะถูกนำมาใช้ในการจำแนกข้อมูลเพื่อทำการจำแนกแต่ละพิกเซลออกเป็นกลุ่มของพิกเซลของสนามและกลุ่มของพิกเซลที่ไม่ใช่สนามโดยจะมีรายละเอียดดังนี้

3.1 หลักการตรวจจับสนาม

การตรวจจับสนามดำเนินการโดยใช้พื้นฐานของความน่าจะเป็นตามทฤษฎีของเบย์ (Bayes’s Threory) และใช้โมเดลผสมแบบเกาเซียนในการจำแนกกลุ่มของพิกเซล ทำการพิจารณาความน่าจะเป็นของแต่ละพิกเซลที่จะเป็นสนามและไม่ใช่สนาม โดยกำหนดให้ X เป็นเซตของพิกเซล $X=\{x_1,x_2,\dots,x_N\}$ กำหนดให้ C เป็นเซตของลาเบลคลาส $C=\{c_f,c_{nf}\}$ ซึ่งมี 2 คลาส คือ คลาสที่เป็นสนามและคลาสที่ไม่ใช่สนามและ F คือ ข้อมูลที่เป็นเซตของเอกลักษณ์ซึ่งในงานนี้ใช้ 9 เอกลักษณ์ คือ ค่าเฉลี่ยสีแดง (M_R) ค่าเฉลี่ยสีเขียว (M_G) ค่าเฉลี่ยสีน้ำเงิน (M_B) ค่าความแปรปรวนสีแดง (Std_R) ค่าความแปรปรวนสีเขียว (Std_G) ค่าความแปรปรวนสีน้ำเงิน (Std_B) ค่าสีแดง (R) ค่าสีเขียว (G) และค่าสีน้ำเงิน (B) เป็นต้น โดยกำหนดเป็นเซตของตัวแปรเอกลักษณ์ ดังสมการที่ 4.4

$$F=\{M_R,M_G,M_B,Std_R,Std_G,Std_B,R,G,B\} \quad (4.4)$$

โดยแทนด้วยสัญลักษณ์ $F=\{t_1,t_2,\dots,t_N\}$ โดยในที่นี้จะใช้อัลกอริทึมเบย์เป็นฟังก์ชันแมพปิ้งหรือตัวจำแนก (Classifier) ซึ่งจะทำหน้าที่แมพอินพุตพิกเซลเข้ากับลาเบลคลาสดังสมการที่ 4.5

$$f=F\rightarrow C \quad (4.5)$$

สำหรับข้อมูลพื้นฐานในส่วนของค่าเฉลี่ยและค่าความแปรปรวนของชุดข้อมูลสำหรับสอนนั้นมีขั้นตอนการดำเนินการดังนี้

พหูน ปญฺญา จิโต ชิโต

1) สุ่มเลือกพิกเซลในกลุ่มที่เป็นสนามและกลุ่มที่ไม่ใช่สนามในแต่ละภาพโดยมีจำนวนเท่ากับ 2 ใน 3 ของแต่ละกลุ่ม

2) แต่ละพิกเซลจะถูกคำนวณหาค่าของแต่ละเอกลักษณ์จำนวน 9 เอกลักษณ์ คือ ค่าเฉลี่ยสีแดง ค่าเฉลี่ยสีเขียว ค่าเฉลี่ยสีน้ำเงิน ค่าความแปรปรวนสีแดง ค่าความแปรปรวนสีเขียว ค่าความแปรปรวนสีน้ำเงิน ค่าสีแดง ค่าสีเขียวและค่าสีน้ำเงิน โดยที่ค่าเฉลี่ยและค่าความแปรปรวนนั้นจะถูกคำนวณได้จากค่าของพิกเซลใกล้เคียงซึ่งในที่นี้จะใช้พิกเซลใกล้เคียงที่มีขนาดของเมทริกซ์เท่ากับ 5 x 5, 7 x 7, 15 x 15 และ 25 x 25 พิกเซล ตามลำดับ

3) นำข้อมูลเอกลักษณ์ของพิกเซลในแต่ละกลุ่มของทุกภาพในชุดสอนนั้นรวมกัน ซึ่งจะได้กลุ่มของเอกลักษณ์ 2 กลุ่ม คือ เอกลักษณ์ของกลุ่มพิกเซลที่เป็นสนามและเอกลักษณ์ของกลุ่มพิกเซลที่ไม่ใช่สนาม

4) คำนวณหาค่าเฉลี่ยของแต่ละเอกลักษณ์จะได้เป็นเวกเตอร์ของค่าเฉลี่ย $\vec{\mu}$ และค่าความแปรปรวนทั้งหมดจะได้เป็นเมทริกซ์ของค่าความแปรปรวน Σ ซึ่งมีขนาด 9 x 9 เมทริกซ์

5) จะได้เวกเตอร์ของค่าเฉลี่ยและเมทริกซ์ค่าความแปรปรวนจำนวน 2 ชุด คือ ชุดของกลุ่มพิกเซลที่เป็นสนามและชุดของกลุ่มพิกเซลที่ไม่ใช่สนามซึ่งจะนำไปใช้แทนในสมการของเกาเซียนต่อไปซึ่งมีรูปแบบดังสมการที่ 4.6 ดังนี้

$$p(t|\theta_k) = \frac{1}{\theta_k\sqrt{2\pi}} e^{\frac{(t-\mu_k)^2}{2\sigma^2}} \tag{4.6}$$

โดยที่ $\theta_k = \{\mu_k, \sigma_k\}$ คือ พารามิเตอร์ของคลาส C ซึ่งคำนวณมาจากขั้นตอนการสอนระบบ

3.2 หลักการคำนวณความน่าจะเป็น

สำหรับค่าความน่าจะเป็นสูงสุด (Maximum Likelihood: ML) ของพิกเซล x ใดๆ จะคำนวณจากผลของฟังก์ชันเกาเซียนทั้ง 2 ค่าโดยใช้สมการที่ 4.7 ของ ML ดังนี้

$$ML = \arg\max P\left(F, x = c_f, c_{nf}\right) \tag{4.7}$$

สำหรับค่าความน่าจะเป็นก่อนหน้าหรือค่าไพออร์ (Prior) จะคำนวณได้จากความน่าจะเป็นของกลุ่มพิกเซลที่เป็นสนามและไม่ใช่สนามของภาพที่ได้จากการสุ่มมาจากข้อมูลชุดสอนซึ่งเป็นชุดของภาพที่มีการกำหนดลาเบลไว้แล้ว โดยค่าความน่าจะเป็นที่จะนำมาใช้ในสมการจะมีวิธีใน

การพิจารณา คือ พิจารณาจากตำแหน่งของพิกเซลของภาพทดสอบว่าอยู่ในส่วนใดของภาพจากเกณฑ์การแบ่งภาพออกเป็น 3 ส่วนเท่าๆ กันดังภาพที่ 4.5 เช่น ถ้าพิกเซลอยู่ในส่วนตอนบนสุดของภาพก็จะนำค่าความน่าจะเป็นครั้งก่อนหรือค่าไพออร์ในส่วนบนสุดของภาพมาคำนวณ เป็นต้น ส่วนค่าความน่าจะเป็นครั้งหลัง (Posterior) สามารถคำนวณได้โดยใช้รูปแบบดังสมการที่ 4.8 ดังนี้

$$MAP = \arg\max P\left(F \mid x = c_f, c_{nf}\right) P\left(x = c_f, c_{nf}\right) \tag{4.8}$$

สำหรับค่าของอีวิเดนส์ (Evidence) ในส่วนของ $\sum_C P(F)P(C)$ สามารถคำนวณได้จากความน่าจะเป็นรวมของทั้งสองกลุ่มซึ่งโดยปกติแล้วจะเป็นค่าคงที่

ในลำดับสุดท้ายจะได้ค่าความน่าจะเป็นของพิกเซล x ใดๆ ซึ่งถ้าค่าของความน่าจะเป็นของสนามมากกว่าค่าของความน่าจะเป็นที่ไม่ใช่สนามพิกเซล x จะถูกพิจารณาเป็นสนาม มิฉะนั้นจะถูกพิจารณาเป็นพิกเซลที่ไม่ใช่สนาม หลังจากพิจารณาครบทุกพิเซลของภาพทดสอบแล้วระบบจะทำการกำหนดลาเบลของพิกเซลในแต่ละกลุ่มและกำหนดเป็นเอาท์พุตผลลัพธ์ซึ่งจะถูกใช้ในขั้นตอนการปรับแต่งข้อมูลต่อไป

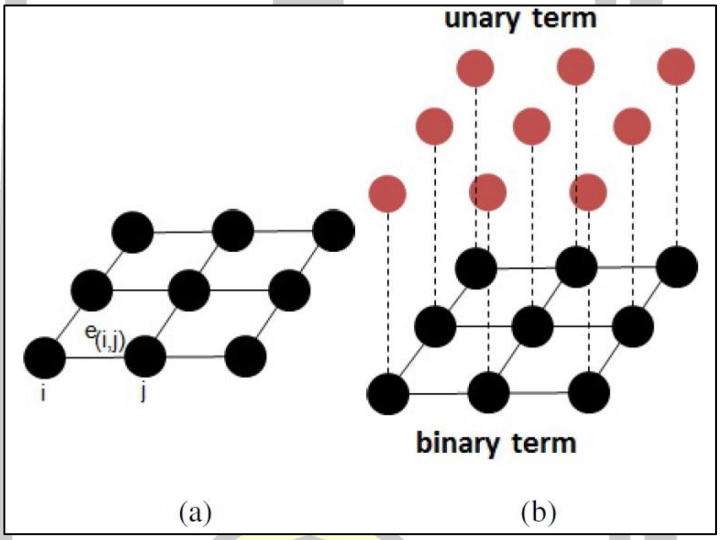
4. ขั้นตอนการปรับแต่งข้อมูล (Data Refinement)

ภายหลังจากขั้นตอนการจำแนกข้อมูลเสร็จสิ้น ผลลัพธ์ที่ได้จากขั้นตอนการจำแนกข้อมูลนั้นอาจยังไม่มีประสิทธิภาพเท่าที่ควร เนื่องจากอาจเกิดปัญหามีกลุ่มของพิกเซลที่ถูกจัดเข้ากลุ่มที่ไม่ถูกต้อง ซึ่งกลุ่มของพิกเซลเหล่านี้โดยทั่วไปมักจะปรากฏอยู่บริเวณขอบระหว่าง 2 พื้นที่ของสนามและพื้นที่ที่ไม่ใช่สนามแสดงดังภาพที่ 4.7 ส่งผลให้กระบวนการจำแนกข้อมูลไม่สามารถแยกแยะได้ถูกต้องว่าพิกเซลเหล่านี้ควรถูกจัดเข้าอยู่ในกลุ่มใดระหว่างกลุ่มที่เป็นสนามและกลุ่มที่ไม่ใช่สนามส่งผลให้พิกเซลเหล่าจะถูกจัดเข้ากลุ่มที่ไม่ถูกต้องตามความเป็นจริง นอกจากนี้โดยทั่วไปแล้วภายในภาพมักจะประกอบด้วยพื้นที่สนามและพื้นที่ที่ไม่ใช่สนามปะปนกันหลายพื้นที่ นั่นหมายความว่ายิ่งมีรอยต่อระหว่าง 2 พื้นที่เหล่านี้มากเท่าไรก็จะมีกลุ่มของพิกเซลที่มีปัญหาเหล่านี้ปรากฏอยู่มากขึ้นเท่านั้นซึ่งส่งผลให้ประสิทธิภาพของกระบวนการจำแนกข้อมูลลดลงตามไปด้วย

จากปัญหาดังกล่าวนี้ในงานวิจัยนี้จึงแก้ไขโดยการเพิ่มวิธีการหลังการประมวลผล (Post-Processing Approach) โดยการเพิ่มขั้นตอนการปรับปรุงข้อมูลเข้ามาในระบบซึ่งมีจุดประสงค์เพื่อการปรับปรุงประสิทธิภาพของกระบวนการจำแนกข้อมูล วิธีการที่เสนอนี้จะพยายามปรับปรุงผลลัพธ์ที่ได้จากขั้นตอนการจำแนกข้อมูลโดยการพิจารณาถึงปฏิสัมพันธ์ (Interaction) ของความสัมพันธ์เชิงพื้นที่ (Spatial Relationships) ของพิกเซลภายในภาพ โดยในงานนี้ประยุกต์ใช้เทคนิคที่เรียกว่า มาคอฟแรนดอมฟิวส์ (MRF) สำหรับการคำนวณความสัมพันธ์ระหว่างพิกเซลที่พิจารณาและพิกเซลเพื่อนบ้านหรือพิเซลที่อยู่รอบบริเวณของพิกเซลที่พิจารณา โดยมีรายละเอียดดังนี้

กำหนดเซตของพิกเซลด้วยสัญลักษณ์ $S=\{s_1,s_2,...,s_N\}$ โดยที่ N คือ จำนวนของพิกเซลภายในภาพ ซึ่งเซตของพิกเซลนี้จะสัมพันธ์กับความน่าจะเป็น $p(c|\text{p})$ ที่มีการกำหนดคลาสให้กับพิกเซลไว้แล้วซึ่งความน่าจะเป็นนี้ได้มาจากความน่าจะเป็นครั้งหลัง (Post Probability) ด้วยการใช้อัลกอริทึมตัวจำแนกแบบเบย์ (Naïve Bayes Classifier) ที่ได้อธิบายไว้ในขั้นตอนก่อนหน้านี้ ความน่าจะเป็นที่เกี่ยวข้องนี้จะถูกใช้เป็นตัวเริ่มต้นลาเบลคลาสของพิกเซล วิธีการดำเนินการของขั้นตอนการปรับแต่งข้อมูลนี้มีรายละเอียดดังนี้

กำหนดให้พิกเซลแทนด้วยสัญลักษณ์ S ทำการสร้างกราฟตาข่าย (Lattice Graph) ซึ่งในที่นี้แทนด้วยสัญลักษณ์ P โดย เอดจ์ (Edge: e) จะเป็นตัวกำหนดความต่อเนื่องของคู่ของพิกเซล i และ j ใดๆ ในกราฟ ซึ่งแสดงตัวอย่างดังภาพที่ 4.6



ภาพที่ 4.6 การแสดงพิกเซลด้วยกราฟและการแสดงเทอมพลังงานในรูปแบบ unary, binary

ตามภาพที่ 4.6 แสดงตัวอย่างการแสดงรายการพิกเซลในรูปแบบของกราฟแบบตาข่ายดังภาพที่ 4.6 (a) และการแสดงเทอมของพลังงาน (Energy Term) ในรูปของ unary และ binary ดังภาพที่ 4.6 (b)

เพื่อให้บรรลุผลตามวัตถุประสงค์ในงานวิจัยนี้จึงใช้อัลกอริทึมที่เรียกว่า ออฟติมอลอินเฟอร์เรนซ์ (Optimal Inference Algorithm) ในการอนุมานข้อมูลบนกราฟโดยใช้เทคนิคที่เรียกว่า กราฟคัท (Graph Cut Technique) เพื่อให้ได้ผลลัพธ์สุดท้ายในการจำแนกข้อมูลซึ่งเป็นเทคนิคที่ถูกนำเสนออยู่ในงานวิจัยของ Zabih และคณะ [14] โดยมีรูปแบบดังสมการที่ 4.9 ดังนี้

$$E(P) = E_{smooth}(P) + E_{data}(P) \tag{4.9}$$

โดยที่ $E_{smooth}(P)$ คือ ฟังก์ชันในการคำนวณความสมูธ (Smoothness) โดยการรักษาความไม่ต่อเนื่อง (Discontinuity) ณ บริเวณขอบระหว่างคลาส โดยฟังก์ชัน $E_{smooth}(P)$ สามารถนิยามได้ดังสมการที่ 4.10 ดังนี้

$$E_{smooth}(P) = \sum_{\{i,j\} \in N} V(s_i, s_j) \tag{4.10}$$

สำหรับ $V(s_i, s_j)$ จะทำการคำนวณโดยมีหลักการทำงานอยู่บนฟังก์ชันที่มีพื้นฐานอยู่บนรัศมี (Radial-based Function) โดยมีการนิยามไว้ตามสมการที่ 4.11 ดังนี้

$$V(s_j, s_j) = e^{-\frac{1}{2\sigma^2}\|s_i - s_j\|^2} \tag{4.11}$$

สำหรับ σ จะเป็นตัวกำหนดออฟเซ็ตของความสัมพันธ์ระหว่างพิกเซล โดยเทอมของข้อมูล (Data Term) ในงานวิจัยนี้จะนิยามโดยการใช้ความน่าจะเป็นครั้งหลังที่ได้จากขั้นตอนการจำแนกข้อมูลซึ่งกระทำโดยตัวจำแนกแบบเบย์ โดยมีการนิยามไว้ในสมการที่ 4.12 ดังนี้

$$E_{data}(P) = \sum_{s \in S} D(s) \tag{4.12}$$

ฟังก์ชันพลังงาน (Energy Function) ที่นิยามในสมการที่ 4.9 จะลดจำนวนในส่วนของลาเบลคลาส C เพื่อให้ได้ผลลัพธ์การจำแนกที่ดีที่สุด (Optimal Classification Results) ดังนั้นเทอมของข้อมูลจะนิยามได้ตามสมการที่ 4.13 ดังนี้

$$D(s) = -\lambda \log(p(c \mid s)) \tag{4.13}$$

หลังจากกระบวนการปรับปรุงข้อมูลเสร็จสิ้นจะได้ผลลัพธ์สุดท้ายของการจำแนกข้อมูลซึ่งจะประกอบด้วยพิกเซล 2 กลุ่ม คือ กลุ่มที่เป็นสนามและกลุ่มที่ไม่ใช่สนาม โดยผลจากการตรวจจับสนามนี้ในกลุ่มพิกเซลที่ไม่ใช่สนามโดยส่วนใหญ่มักจะเป็นส่วนของวัตถุซึ่งอาจเป็นผู้เล่น กรรมการ ลูกบอลและวัตถุอื่นๆ เช่น เส้นสนามและป้ายโฆษณา เป็นต้น ซึ่งในงานวิจัยนี้จะทำการจัดเก็บข้อมูลเหล่านี้ไว้ในรายการด้วยเพื่อนำไปใช้ในขั้นตอนการตรวจจับผู้เล่นต่อไป

4.2 กระบวนการทดลองและผลการทดลอง (Experimental Process and Results)

ในงานวิจัยนี้ได้เสนอเทคนิคสำหรับการจำแนกส่วนของพื้นที่ (Region) ของสนามฟุตบอลที่ปรากฏภายในภาพ ซึ่งเทคนิคที่เสนอจะมีการทำงานอยู่บนพื้นฐานของ 2 เอกลักษณ์ คือ ความเข้มของค่าสีและพื้นผิว สำหรับกระบวนการจำแนกพิกเซลในงานนี้จะใช้ข้อมูลความน่าจะเป็นของพื้นที่สนามเป็นพื้นฐานในการทำงาน หลังจากขั้นตอนการจำแนกเสร็จสิ้น ในงานนี้ได้เสนอวิธีการปรับปรุงข้อมูลเพื่อเพิ่มประสิทธิภาพผลลัพธ์ที่ได้จากขั้นตอนการจำแนกข้อมูล ส่วนการทดสอบระบบจะใช้วิธี Cross Validation

กระบวนการทดลองระบบใช้ชุดข้อมูลภาพสำหรับการสอนระบบจำนวน 120 ภาพ และภาพสำหรับทดสอบระบบจำนวน 120 ภาพ รวม 240 ภาพ ในแต่ละภาพจะคำนวณหาค่าความน่าจะเป็นของแต่ละพิกเซลที่จะเป็นสนามและไม่ใช่สนามโดยใช้พื้นฐานของความน่าจะเป็นตามทฤษฎีของเบย์ แล้วทำการเปรียบเทียบค่าความน่าจะเป็นทั้ง 2 ค่าเพื่อเลือกค่าความน่าจะเป็นที่มากที่สุดเป็นคำตอบของพิกเซลนั้น จากนั้นตรวจสอบว่าพิกเซลนั้นอยู่ในส่วนใดของภาพตามเกณฑ์การแบ่งภาพออกเป็น 3 ส่วนดังภาพที่ 4.5 แล้วคำนวณหาความน่าจะเป็นที่ดีที่สุดอีกครั้งโดยใช้ค่าความน่าจะเป็นที่ได้จากขั้นตอนที่ 1. การจัดเตรียมข้อมูลก่อนการประมวลผล ในหัวข้อที่ 4.1 วิธีดำเนินการวิจัยมาร่วมพิจารณาด้วยซึ่งจะช่วยเพิ่มความถูกต้องในการค้นหามากขึ้น

การประเมินประสิทธิภาพของระบบโดยภาพรวมดำเนินการโดยใช้ตาราง Confusion Matrix โดยหาค่าความถูกต้อง ค่าความระลึกและค่าความแม่นยำของผลการดำเนินการในแต่ละภาพ

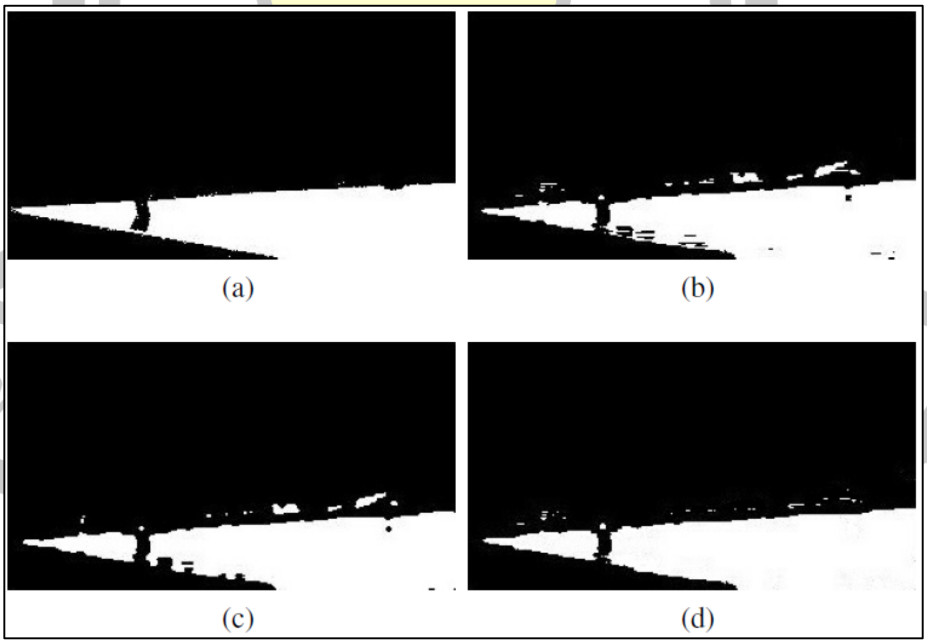
การประเมินประสิทธิภาพของกระบวนการปรับปรุงข้อมูลดำเนินการโดยการเปรียบเทียบระหว่างผลก่อนการปรับปรุงและผลที่ได้หลังจากขั้นตอนการปรับปรุงซึ่งในงานนี้ได้ใช้ 2 เทคนิค คือ วิธีมีเดียนฟิลด์เตอร์ (MF) และวิธีมาคอฟแรนดอมฟิวส์ (MRF) ส่วนข้อมูลภาพที่ใช้ทดสอบจะมีจำนวน 120 ภาพ ซึ่งในแต่ละภาพจะมีพื้นที่ของสนามปรากฏอยู่ ส่วนภาพกราวนด์ทรูธจะดำเนินการสร้างขึ้นโดยใช้คนพิจารณา

ภาพอินพุตที่ใช้ทดสอบจะถูกแบ่งออกเป็นพื้นที่ที่มีลักษณะซ้อนทับกัน (Over-Lapping Window) ซึ่งขนาดของพื้นที่จะถูกกำหนดไว้ล่วงหน้า จากนั้นแต่ละพื้นที่จะถูกนำมาสกัดเอกลักษณ์ดังที่อธิบายไว้ในขั้นตอนที่ 2. การสกัดเอกลักษณ์ ในหัวข้อที่ 4.1 วิธีดำเนินการวิจัย การจำแนกข้อมูลในงานนี้ดำเนินการกับพื้นที่ที่มีขนาดที่แตกต่างกัน จุดประสงค์เพื่อศึกษาและสังเกตถึงแนวโน้มของประสิทธิภาพในการจำแนกข้อมูลซึ่งตัวอย่างของผลลัพธ์จากกระบวนการจำแนกในแต่ละ พื้นที่แสดงดังตารางที่ 4.2 ดังนี้

ตารางที่ 4.2 ผลลัพธ์ของการจำแนกข้อมูลโดยใช้พื้นที่ที่มีขนาดแตกต่างกัน

เอกลักษณ์	5x5 Pixel	7x7 Pixel	15x15 Pixel	25x25 Pixel
ความเข้มของค่าสี (Color Intensity)	0.944	0.945	0.935	0.926
พื้นผิว (Texture)	0.702	0.703	0.715	0.723
ความเข้มของค่าสีร่วมกับพื้นผิว (Intensity+Texture)	0.813	0.787	0.776	0.795

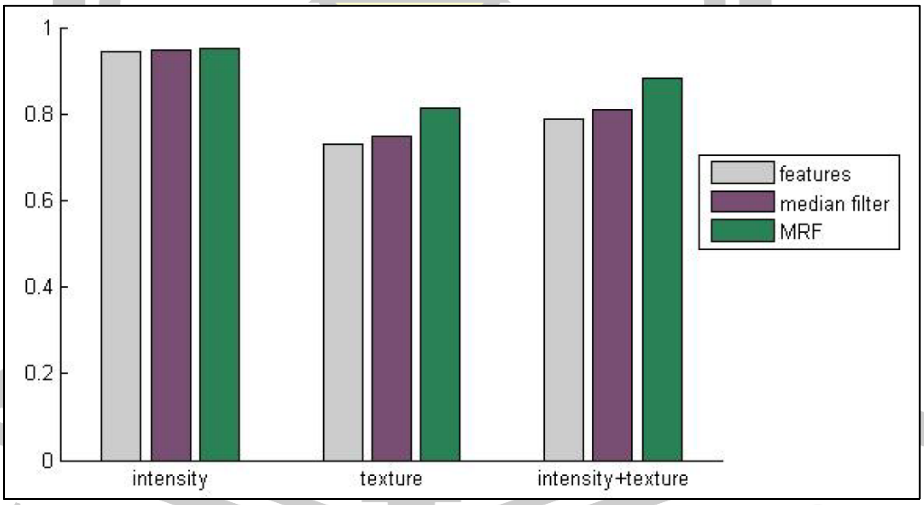
ตามตารางที่ 4.2 แสดงตัวอย่างของผลลัพธ์ที่ได้จากการจำแนกข้อมูลซึ่งจะเห็นว่าการแบ่งภาพออกเป็นพื้นที่แพตช์ (Window Patch) ขนาด 7 x 7 พิกเซล เอกลักษณ์ความเข้มของค่าสีจะให้ผลลัพธ์ที่ดีกว่า 2 เอกลักษณ์ที่เหลือโดยมีความถูกต้องอยู่ที่ 94.5% นอกจากนี้จากผลลัพธ์แสดงให้เห็นว่าการใช้เอกลักษณ์ความเข้มของค่าสีจะมีประสิทธิภาพที่ดีกว่ากรณีใช้เอกลักษณ์พื้นผิวและกรณีใช้เอกลักษณ์ความเข้มของค่าสีร่วมกับพื้นผิว กรณีเปรียบเทียบระหว่างกรณีการใช้เอกลักษณ์ความเข้มของค่าสีกับกรณีการใช้เอกลักษณ์ความเข้มของค่าสีร่วมกับเอกลักษณ์พื้นผิวโดยใช้พื้นที่แพตช์ขนาด 7 x 7 พิกเซล พบว่าเอกลักษณ์ความเข้มของค่าสีให้ผลลัพธ์ที่ดีกว่า โดยมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (Statistically Significant) ที่ p=0.014



ภาพที่ 4.7 ตัวอย่างผลลัพธ์ที่ได้จากการจำแนกข้อมูล

ตามภาพที่ 4.7 แสดงตัวอย่างของผลลัพธ์ที่ได้จากกระบวนการจำแนกโดยใช้พื้นที่แพตช์ขนาด 7 x 7 พิกเซล โดยภาพที่ 4.7 (a) คือ ภาพกราวนด์ทรูธ ส่วนภาพที่ 4.7 (b) คือ ผลจากการจำแนกข้อมูลแบบปกติที่ไม่มีการใช้ข้อมูลบริบทใดๆ มาเกี่ยวข้อง (Non-Context Classification) ซึ่งจะเห็นว่ามีกลุ่มของพิกเซลที่ถูกจัดกลุ่มไม่ถูกต้องปรากฏอยู่บริเวณขอบระหว่างพื้นที่สนามและพื้นที่ที่ไม่ใช่สนาม ภาพที่ 4.7 (c) คือ ผลจากการปรับปรุงข้อมูลโดยใช้วิธีมีเดียนฟิวส์เตอร์และภาพที่ 4.7 (d) คือ ผลจากการปรับปรุงข้อมูลโดยใช้วิธีมาคอฟแรนดอมฟิวส์ เมื่อพิจารณาจะเห็นว่าการเพิ่มกระบวนการปรับปรุงข้อมูลโดยใช้วิธีมีเดียนฟิวส์เตอร์และวิธีมาคอฟแรนดอมฟิวส์สามารถปรับปรุงประสิทธิภาพของผลลัพธ์ที่ได้จากกระบวนการจำแนกข้อมูลแบบปกติที่ไม่ใช้ข้อมูลบริบทมาเกี่ยวข้อง นอกจากนี้ยังพบว่ากรณีการใช้วิธีมาคอฟแรนดอมฟิวส์นั้นจะให้ผลลัพธ์ที่ดีกว่ากรณีการใช้วิธีมีเดียนฟิวส์เตอร์

เมื่อพิจารณาถึงขั้นตอนการจำแนกข้อมูลดังที่อธิบายไว้ก่อนหน้าจะเห็นว่าในงานนี้ได้ประยุกต์ใช้ 2 เทคนิคได้แก่ มีเดียนฟิวส์เตอร์และมาคอฟแรนดอมฟิวส์ ซึ่งถูกนำมาใช้ในการปรับปรุงผลลัพธ์ของกระบวนการจำแนกข้อมูล โดยผลลัพธ์ของกระบวนการปรับปรุงข้อมูลโดยใช้ 2 เทคนิคดังกล่าวจะแสดงดังภาพที่ 4.8 ดังนี้



ภาพที่ 4.8 ผลการจำแนกข้อมูลโดยใช้เพียงเอกลักษณ์เปรียบเทียบกับการใช้กระบวนการปรับปรุงข้อมูล

ตามภาพที่ 4.8 เป็นการแสดงการเปรียบเทียบระหว่างผลลัพธ์ที่ได้จากกระบวนการจำแนกข้อมูลโดยใช้เอกลักษณ์เพียงอย่างเดียวซึ่งมี 3 เอกลักษณ์ คือ ความเข้มของค่าสี พื้นผิวและความเข้มของค่าสีร่วมกับพื้นผิวและกรณีการเพิ่มขั้นตอนการปรับปรุงข้อมูลซึ่งใช้ 2 เทคนิค คือ วิธีมีเดียนฟิวส์

เตอร์และวิธีมาคอฟแรนดอมฟิวส์ พบว่ากรณีการใช้เอกลักษณ์ความเข้มของค่าสีนั้นทั้ง 3 วิธีจะให้ผลที่ใกล้เคียงกันโดยเรียงจากน้อยไปหามาก คือ การใช้เอกลักษณ์เพียงอย่างเดียว การใช้วิธีมีเดียนฟิวส์เตอร์และการใช้วิธีมาคอฟแรนดอมฟิวส์ ตามลำดับ นอกจากนี้ยังพบว่าการใช้เอกลักษณ์ความเข้มของค่าสีนี้ยังให้ผลลัพธ์ที่ดีกว่า 2 เอกลักษณ์ที่เหลือ คือ เอกลักษณ์พื้นผิวและเอกลักษณ์ความเข้มของค่าสีร่วมกับเอกลักษณ์พื้นผิว ในกรณีการใช้เอกลักษณ์พื้นผิวจะพบว่าการใช้วิธีมีเดียนฟิวส์เตอร์และวิธีมาคอฟแรนดออมฟิวส์ให้ผลที่ดีกว่ากรณีการใช้เอกลักษณ์เพียงอย่างเดียว นอกจากนี้ยังพบว่ากรณีการใช้วิธีมาคอฟแรนดออมฟิวส์ยังให้ผลที่ดีกว่ากรณีการใช้มีเดียนฟิวส์เตอร์อีกด้วยและในกรณีการใช้เอกลักษณ์ความเข้มของค่าสีร่วมกับเอกลักษณ์พื้นผิวนั้นพบว่าให้ผลลัพธ์ในลักษณะเดียวกับกรณีการใช้เอกลักษณ์พื้นผิวเพียงอย่างเดียว

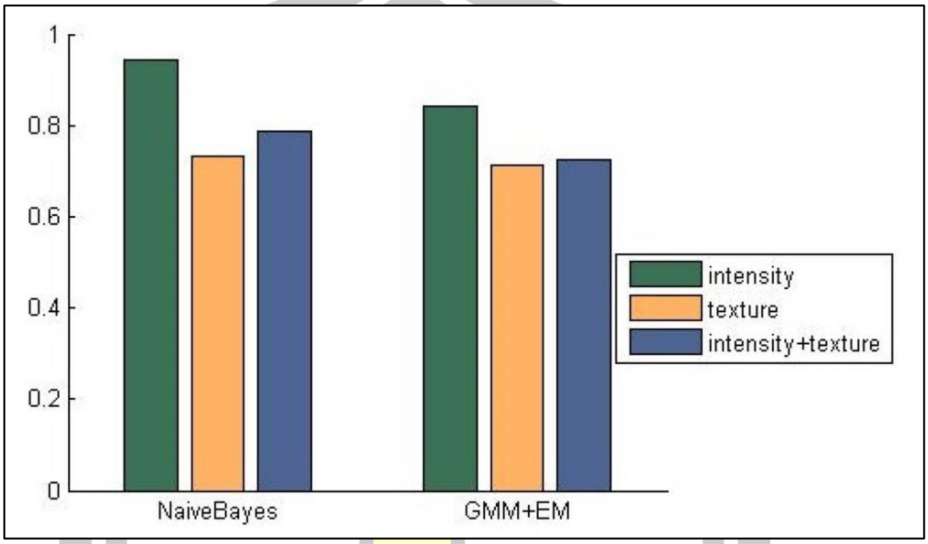
ผลการเปรียบเทียบระหว่างกรณีการใช้เอกลักษณ์เพียงอย่างเดียวและกรณีการใช้ขั้นตอนการปรับปรุงข้อมูลแสดงดังตารางที่ 4.3 ดังนี้

ตารางที่ 4.3 ผลลัพธ์ของการจำแนกข้อมูลและการปรับปรุงข้อมูล

	Feature Alone	Refinement	
		Median Filter	MRF
ความเข้มของค่าสี (Intensity)	0.945±0.016	0.946±0.014	0.951±0.024
พื้นผิว (Texture)	0.731±0.124	0.748±0.098	0.812±0.540
ความเข้มของค่าสีร่วมกับพื้นผิว (Intensity+Texture)	0.787±0.045	0.811±0.015	0.884±0.075

ตามตารางที่ 4.3 แสดงการเปรียบเทียบระหว่างกรณีของการจำแนกข้อมูลแบบปกติและกรณีของการใช้ขั้นตอนการปรับปรุงข้อมูล ในที่นี้เป็นตัวอย่างผลลัพธ์กรณีใช้พื้นที่แพตช์ขนาด 7 x 7 พิกเซล พบว่ากรณีการใช้เอกลักษณ์ความเข้มของค่าสีนั้นจะเห็นว่ากรณีการใช้วิธีแบบเดิมอย่างเดียวให้ความถูกต้องอยู่ที่ 0.945 ± 0.016 คิดเป็น 94.5% ขณะที่กรณีการใช้มีเดียนฟิวส์เตอร์และมาคอฟแรนดอมฟิวส์จะให้ผลที่ดีกว่า โดยมีความถูกต้องอยู่ที่ 0.946 ± 0.014 คิดเป็น 94.6% และ 0.951 ± 0.024 คิดเป็น 95.1% ตามลำดับ ซึ่งตามภาพที่ 4.8 แสดงให้เห็นว่าเทคนิคการปรับปรุงข้อมูลที่เสนอให้ผลที่ดีกว่าการใช้เอกลักษณ์เพียงอย่างเดียวในทุกเอกลักษณ์ เมื่อพิจารณาที่เอกลักษณ์ความเข้มของค่าสีจะเห็นว่ากระบวนการปรับปรุงข้อมูลโดยใช้มาคอฟแรนดอมฟิวส์สามารถปรับปรุงประสิทธิภาพของกระบวนการจำแนกข้อมูลแบบเดิมที่ใช้เอกลักษณ์เพียงอย่างเดียว จาก 94.5% เป็น 95.1% โดยมีค่านัยสำคัญทางสถิติอยู่ที่ p=0.023

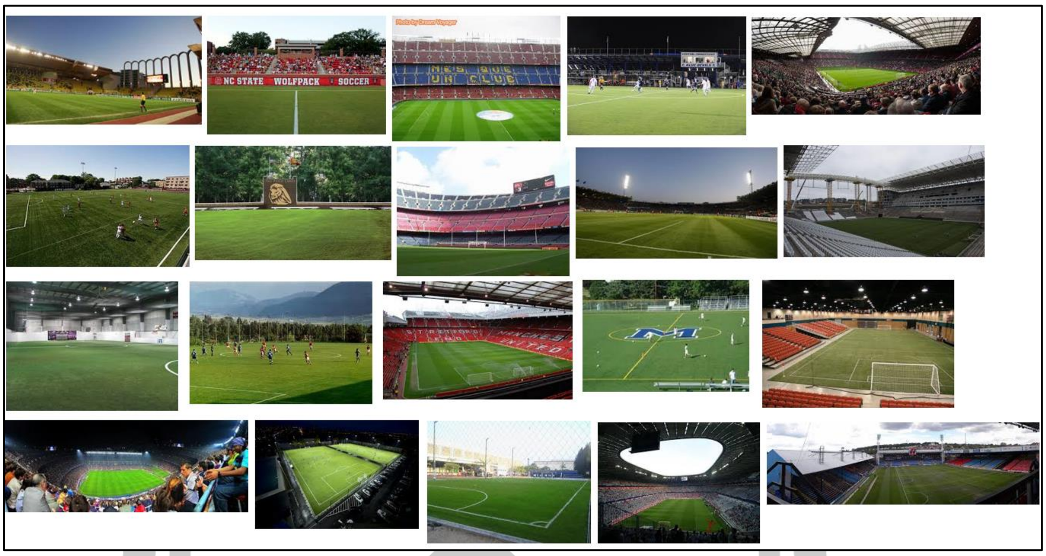
ผลการเปรียบเทียบระหว่างการจำแนกข้อมูลโดยใช้เทคนิคที่เสนอและกรณีการใช้โมเดลผสมแบบเกาเซียนร่วมกับอัลกอริทึม EM แสดงได้ดังภาพที่ 4.9 ดังนี้



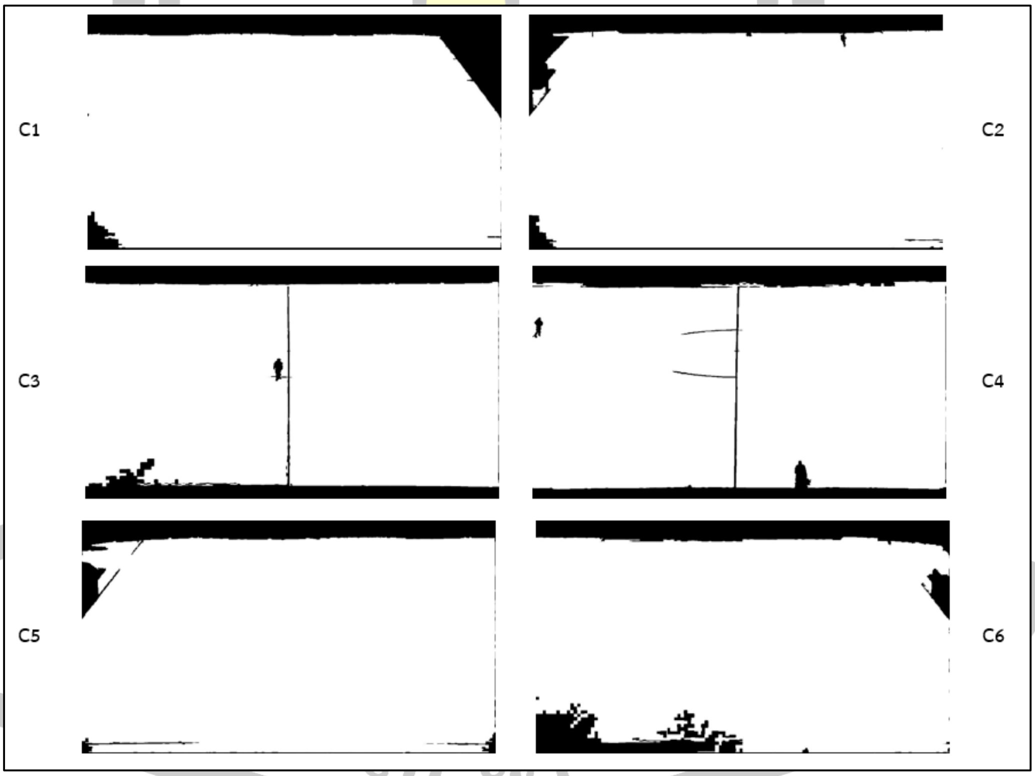
ภาพที่ 4.9 การเปรียบเทียบผลการจำแนกข้อมูล

ตามภาพที่ 4.9 เป็นการเปรียบเทียบระหว่างผลลัพธ์จากการจำแนกข้อมูลโดยใช้เทคนิคที่เสนอและกรณีการใช้โมเดลผสมแบบเกาเซียนร่วมกับอัลกอริทึม EM ในการใช้เอกลักษณ์ทั้ง 3 แบบ จะเห็นว่าเทคนิคที่เสนอซึ่งใช้โมเดลผสมแบบเกาเซียนร่วมกับอัลกอริทึมแบบเบย์และเทคนิคการปรับปรุงข้อมูลจะให้ผลลัพธ์ที่ดีกว่ากรณีการใช้เทคนิคแบบเดิมที่ใช้โมเดลผสมแบบเกาเซียนร่วมกับอัลกอริทึม EM โดยจะเห็นว่าเทคนิคที่เสนอให้ผลลัพธ์ที่ดีกว่าในทุกกรณีของการใช้เอกลักษณ์

ตัวอย่างชุดภาพสำหรับใช้ในการสอนระบบและทดสอบระบบแสดงดังภาพที่ 4.10 ซึ่งเป็นภาพประเภท JPEG ขนาดภาพ (Dimension) อยู่ในช่วง 170 x 106 ถึง 1280 x 757 ส่วนผลการตรวจจับสนามจากเฟรมภาพแรกของวิดีโอการแข่งขันฟุตบอลจากฐานข้อมูล ISSIA Dataset [82] แสดงดังภาพที่ 4.11



ภาพที่ 4.10 ชุดภาพสำหรับสอนและทดสอบระบบ



ภาพที่ 4.11 ผลการตรวจจับสนามภาพเฟรมแรกของ ISSIA Dataset [82]

4.3 สรุปผลการดำเนินงาน (Summary Operating Results)

ในงานวิจัยนี้ได้นำเสนอเทคนิคในการตรวจจับสนามภายในภาพของเกมส์การแข่งขันฟุตบอล ซึ่งกระบวนการจำแนกข้อมูลนั้นจะดำเนินการบนพื้นฐานของเอกลักษณ์ในระดับพิกเซลโดยใช้เอกลักษณ์หลัก 2 ตัว คือ เอกลักษณ์ความเข้มของค่าสีและเอกลักษณ์พื้นผิว นอกจากนี้ในงานวิจัยนี้ยังประยุกต์ใช้ตัวจำแนกข้อมูลแบบเบย์มาปรับใช้ในกระบวนการจำแนกข้อมูลด้วย ซึ่งข้อมูลความน่าจะเป็นหรือไพรออร์จะได้มาจากข้อมูลของพื้นที่ของสนาม สำหรับการสกัดเอกลักษณ์จะใช้พื้นที่แพตช์ในขนาดต่างๆ เพื่อศึกษาถึงแนวโน้มของผลลัพธ์ซึ่งพบว่าพื้นที่แพตช์ขนาด 7 x 7 พิกเซล ให้ผลลัพธ์ที่เหมาะสมที่สุดในทุกเอกลักษณ์ หลังจากกระบวนการจำแนกเสร็จสิ้น ในงานนี้ได้ประยุกต์ใช้เทคนิคการปรับปรุงข้อมูล 2 ตัว คือ มีเดียนฟิวส์เตอร์และมาคอฟแรนดอมฟิวส์ จุดประสงค์เพื่อปรับปรุงประสิทธิภาพของผลลัพธ์ที่ได้จากขั้นตอนการจำแนกข้อมูล ซึ่งผลการทดลองพบว่า การใช้วิธีการปรับปรุงข้อมูลแบบมาคอฟแรนดอมฟิวส์ให้ผลที่ดีกว่าแบบมีเดียนฟิวส์เตอร์ นอกจากนี้ยังพบว่าเทคนิคการปรับปรุงข้อมูลที่เสนอสามารถช่วยปรับปรุงผลลัพธ์ของกระบวนการจำแนกข้อมูลได้ ซึ่งส่งผลให้การจำแนกข้อมูลมีประสิทธิภาพเพิ่มมากขึ้น

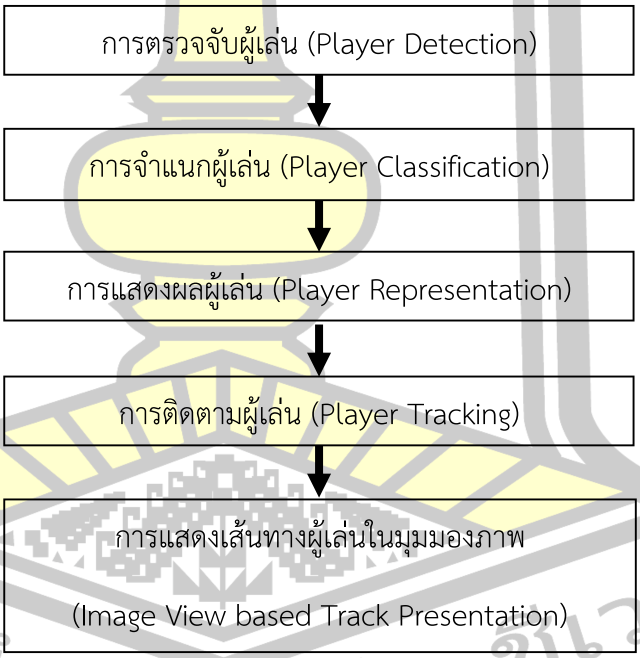
พหูนํ ปณฺฑิโต ชีเว

บทที่ 5

การตรวจจับและติดตามผู้เล่น (Player Detection and Tracking)

ในงานนี้มีวัตถุประสงค์เพื่อตรวจจับและติดตามผู้เล่นในมุมมองภาพ การตรวจจับมีจุดประสงค์เพื่อค้นหาวัตถุที่สนใจและสกัดเอกลักษณ์ของวัตถุเพื่อใช้ประกอบการติดตาม ส่วนการติดตามผู้เล่นมีจุดประสงค์เพื่อจัดเก็บข้อมูลเส้นทางการเคลื่อนที่ของวัตถุ ซึ่งข้อมูลที่ได้จากการตรวจจับและติดตามนี้จะถูกใช้สำหรับการทำนายตำแหน่งเส้นทางของผู้เล่นที่ถูกบดบังรวมถึงสามารถใช้ประกอบการวิเคราะห์พฤติกรรมต่างๆ เช่น รูปแบบการเคลื่อนที่ เทคนิคการเล่น เป็นต้น โดยรายละเอียดในการดำเนินการมีดังนี้

5.1 วิธีดำเนินการวิจัย (Research Methodology)



ภาพที่ 5.1 กระบวนการตรวจจับและติดตามผู้เล่น

ในงานนี้ได้นำเสนอเทคนิคการตรวจจับและติดตามผู้เล่นฟุตบอลในมุมมองภาพ ประกอบด้วย 5 ขั้นตอนหลัก แสดงดังภาพที่ 5.1 ได้แก่ 1. ขั้นตอนการตรวจจับผู้เล่น (Player Detection) 2. ขั้นตอนการจำแนกผู้เล่น (Player Classification) 3. ขั้นตอนการแสดงผลผู้เล่น (Player Representation) 4. ขั้นตอนการติดตามผู้เล่น (Player Tracking) และ 5. ขั้นตอนการแสดงเส้นทางผู้เล่นในมุมมองภาพ (Image View based Track Presentation) ตามลำดับ

ในขั้นตอนแรกจะทำการตรวจจับผู้เล่นโดยใช้เทคนิคพื้นฐานทั่วไป คือ วิธีการหาส่วนต่างของพื้นหลังหรือวิธีแบ็คกราวนด์ซับแทรคชัน จากนั้นจะทำการจำแนกวัตถุที่ตรวจจับได้ออกเป็นกลุ่มของผู้เล่น คือ นักฟุตบอลและกรรมการ และวัตถุอื่นๆ เช่น ป้ายโฆษณา เส้นสนาม เป็นต้น หลังจากได้กลุ่มของวัตถุที่เป็นผู้เล่นแล้วจะแสดงผลการตรวจจับผู้เล่นโดยในขั้นตอนนี้จะใช้วิธีการแสดงผลแบบกรอบสี่เหลี่ยมรอบวัตถุหรือบราวดิ้งบล็อก ขั้นตอนต่อมาจะเป็นการติดตามเส้นทางการเคลื่อนที่ของผู้เล่นโดยในขั้นตอนนี้จะประยุกต์ใช้เทคนิคที่เรียกว่า คาลมานฟิลเตอร์และเทคนิคที่เรียกว่าอัลกอริทึมคอนเดนเซชันในการทดลอง ซึ่งเป็นเทคนิคที่ได้รับความนิยมในการประยุกต์ใช้สำหรับงานการติดตามวัตถุในวิดีโอต่างๆ จุดประสงค์เพื่อหาเทคนิคที่เหมาะสมสำหรับใช้ในงานวิจัยต่อไป ส่วนขั้นตอนสุดท้ายจะเป็นการแสดงผลเส้นทางของผู้เล่นในมุมมองภาพ ซึ่งจะใช้วิธีการเขียนโปรแกรมแสดงผลทั่วไป รายละเอียดของการดำเนินงานในแต่ละขั้นตอนจะประกอบด้วยหัวข้อดังต่อไปนี้

1. ขั้นตอนการตรวจจับผู้เล่น (Player Detection)

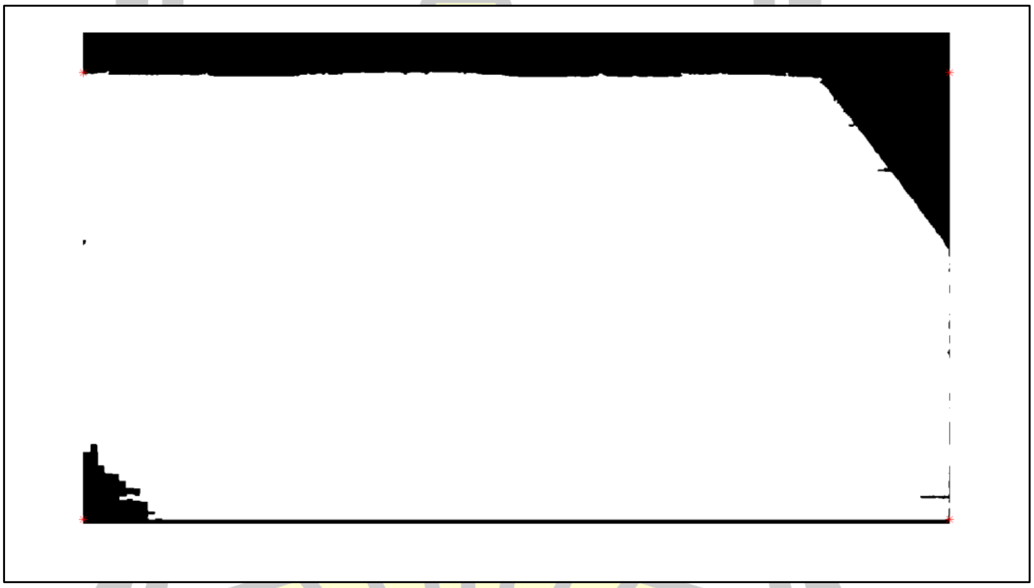
ในขั้นตอนนี้จะดำเนินการตรวจจับผู้เล่นโดยประยุกต์ใช้เทคนิคการตรวจจับวัตถุบนพื้นฐานของวิธีแบ็คกราวนด์ซับแทรคชันและความรู้พื้นฐานที่มีอยู่ซึ่งได้มาจากขั้นตอนการตรวจจับสนาม โดยมีรายละเอียดของการดำเนินการดังนี้

จากขั้นตอนของการตรวจจับสนามโดยใช้โมเดลผสมแบบเกาเซียน ข้อมูลสำคัญที่ได้จากขั้นตอนนี้ได้แก่ ตำแหน่งและขอบเขตของสนามที่ปรากฏภายในภาพ แสดงดังภาพที่ 5.2 ซึ่งขอบเขตของสนามนี้จะถูกใช้เป็นความรู้พื้นฐานสำหรับช่วยจำกัดพื้นที่การตรวจจับผู้เล่นและยังสามารถใช้แบ่งแยกพื้นที่ภายในภาพที่ไม่เกี่ยวข้องกับผู้เล่นออกไปด้วย ซึ่งช่วยให้ลดปัญหาการตรวจจับวัตถุอื่นๆที่ไม่เกี่ยวข้องออกไปได้ในระดับหนึ่ง เช่น ป้ายโฆษณา ผู้ชม อัฒจันทร์และต้นไม้ เป็นต้น ทำให้ได้วัตถุที่เกี่ยวข้องมากที่สุด ส่งผลให้วัตถุที่ใช้จำแนกจะเป็นวัตถุที่ต้องการมากที่สุดและในขั้นตอนการติดตามยังช่วยควบคุมพื้นที่ให้สามารถติดตามได้ในขอบเขตสนามเท่านั้น วิธีการหาขอบเขตสนามจะใช้วิธีพื้นฐานทั่วไป คือ จากภาพไบนารี ทำการค้นหาพิกเซลสนามที่อยู่ในตำแหน่งขอบของแต่ละด้านของพื้นที่สนามซึ่งโดยปกติจะมี 4 ด้าน คือ ด้านบน ด้านซ้าย ด้านขวาและด้านล่าง ดังสมการที่ 5.1 ผลลัพธ์จะได้พิกัดพิกเซลที่เป็นสนามมา 4 พิกัด ตัวอย่างแสดงด้วยจุดสีแดงในภาพที่ 5.2 จากนั้นจะใช้พิกัดทั้ง 4 ตำแหน่งนี้ในการกำหนดพิกัดมุมทั้งสี่ของพื้นที่สี่เหลี่ยมซึ่งเป็นพิกัดขอบเขตของสนาม โดย

ข้อมูลพิกัดมุมทั้งสี่มุมนี้จะถูกใช้สำหรับกำหนดขอบเขตพื้นที่สนามคร่าวๆ สำหรับใช้ในการดำเนินการแบ็คกราวนด์ซับแทรคชันในกระบวนการตรวจจับผู้เล่นต่อไป

$$corner_point_1^N = \left\{ \begin{array}{l} \text{find}\left(\left(x_{\min}, y_{\max}\right), \left(x_{\max}, y_{\max}\right), \left(x_{\max}, y_{\min}\right), \left(x_{\min}, y_{\min}\right)\right) \\ \left| x, y \in \max\left(region_{field}\right)\right. \end{array} \right\} \quad (5.1)$$

โดยที่ *corner_point* คือ รายการพิกัด (*x, y*) ของมุมรอบบริเวณสนาม *N* คือ จำนวนพิกัดจุดทั้ง 4 จุด

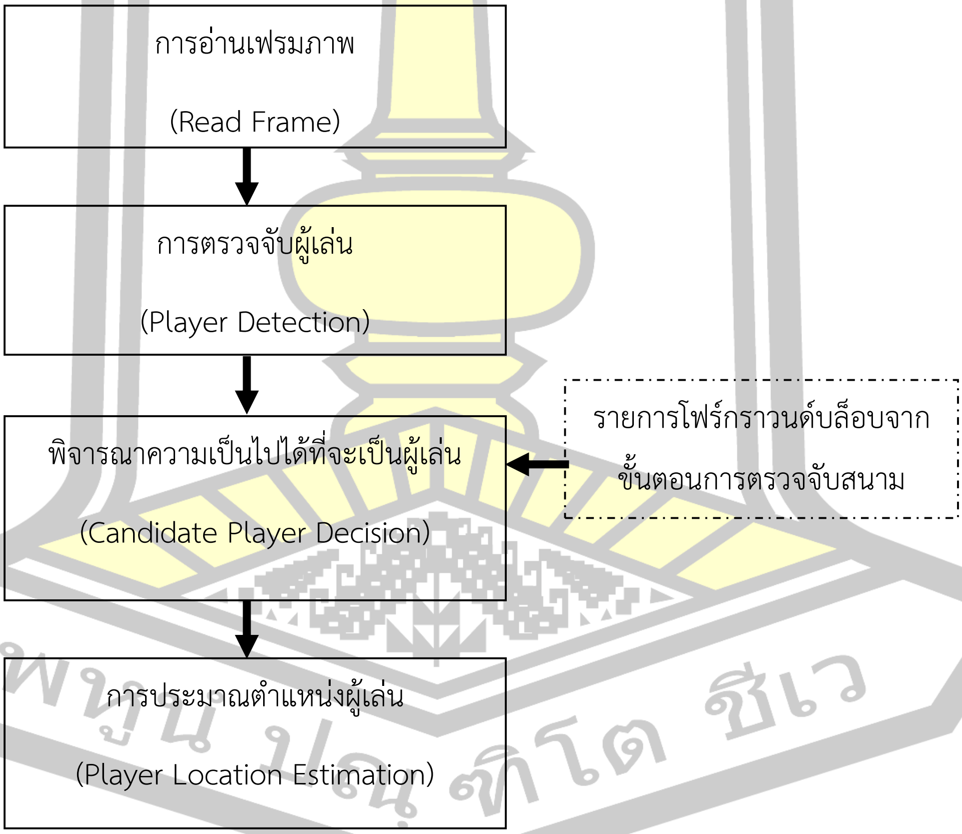


ภาพที่ 5.2 ขอบเขตภายในภาพจากกล้อง C1 (ISSIA Dataset [82])

หลังจากขั้นตอนการตรวจจับสนามในเฟรมแรกจะได้รายการความน่าจะเป็นของวัตถุมาจำนวนหนึ่ง ซึ่งประกอบด้วยผู้เล่นทั้ง 2 ทีม กรรมการ ลูกบอลและวัตถุอื่นๆ จะทำการจำแนกประเภทออกเป็น 4 กลุ่ม คือ ผู้เล่นทีม A ผู้เล่นทีม B กรรมการและวัตถุอื่นๆ รายละเอียดของวิธีการจำแนกตัวตนจะอธิบายอยู่ในหัวข้อที่ 2. ขั้นตอนการจำแนกผู้เล่น ส่วนในเฟรมที่ 2 จะใช้วิธีแบ็คกราวนด์ซับแทรคชันในการตรวจจับผู้เล่น เช่น วิธีบนพื้นฐานของการเคลื่อนไหว (Motion-based) เป็นต้น ซึ่งจะได้รายการวัตถุมาจำนวนหนึ่ง รายการวัตถุที่ได้จากเฟรมที่ 2 นี้จะถูกจำแนกออกเป็น 4 กลุ่ม เช่นเดียวกัน จากนั้นข้อมูลจาก 2 ส่วนนี้จะถูกนำมาพิจารณาร่วมกันอีกครั้งโดยใช้เอกลักษณ์

ต่างๆ เช่น ขนาด การเคลื่อนที่และสี เป็นต้น โดยข้อมูลรายการวัตถุที่ได้จากขั้นตอนการตรวจจับสนามและจากขั้นตอนการทำแบ็คกราวนด์ซับแทรคชันนี้จะเลือกใช้เฉพาะ 3 กลุ่ม คือ ผู้เล่นทั้ง 2 ทีมและกรรมการกลางและจะถูกนำมาใช้คำนวณร่วมกันในช่วงของเฟรมที่ 2 เท่านั้นจากนั้นในเฟรมถัดไปจะใช้วิธีบนพื้นฐานของแบ็คกราวด์ซับแทรคชันในการตรวจจับผู้เล่นเพียงอย่างเดียว โดยใช้วิธีบนพื้นฐานของการเคลื่อนไหว เนื่องจากอาจจะยังมีผู้เล่นบางคนที่ยังไม่มีการเคลื่อนไหวในช่วงเฟรมแรก โดยจะจำกัดพื้นที่การค้นหาเฉพาะในขอบเขตพื้นที่สนามเท่านั้น สำหรับกรรมการในงานนี้จะพิจาณาเฉพาะกรรมการกลางเท่านั้นโดยจะไม่พิจารณาในส่วนของกรรมการข้างสนาม 2 คน เนื่องจากงานวิจัยนี้จะพิจารณาเฉพาะบุคคลที่อยู่ภายในสนามเท่านั้น

กระบวนการตรวจจับผู้เล่นฟุตบอลจะแสดงดังภาพที่ 5.3 ประกอบด้วย 4 ขั้นตอนหลัก คือ 1. การอ่านเฟรมภาพ 2. การตรวจจับผู้เล่น 3. การพิจารณาความเป็นไปได้ที่จะเป็นผู้เล่นและ 4. การประมาณตำแหน่งของผู้เล่น



ภาพที่ 5.3 กระบวนการตรวจจับผู้เล่น

กระบวนการทั้ง 4 ขั้นตอนของการตรวจจับผู้เล่นมีรายละเอียดดังนี้

1.1 อ่านเฟรมภาพ (Read Frame)

ทำการอ่านเฟรมภาพอินพุตจากชุดวิดีโอทีละเฟรมภาพและทำการครอป (Crop) ภาพเฉพาะพื้นที่บริเวณสนามโดยใช้รายการพิกัดมุมที่ได้จากขั้นตอนการตรวจจับสนามเป็นพารามิเตอร์ในการดำเนินการดังสมการที่ 5.2 ซึ่งจะได้ภาพเอาท์พุตที่เป็นบริเวณพื้นที่สนาม

$$crop_image = crop(\mathrm{I} \mid (\mathrm{x},\mathrm{y}) \in \mathrm{corner_point}_1^N) \qquad (5.2)$$

โดยที่ *I* คือ ภาพอินพุตและ *crop_image* คือ ภาพเอาท์พุตที่ได้จากการ crop ภาพ

1.2 การตรวจจับผู้เล่น (Player Detection)

ภาพเอาท์พุตบริเวณสนามที่ได้จากขั้นตอนแรกจะถูกนำมาใช้สำหรับทำการตรวจจับผู้เล่นซึ่งในเฟรมแรกจะได้รายการวัตถุที่เป็นผู้เล่นจากขั้นตอนการตรวจจับสนามอยู่แล้ว ดังนั้นในกระบวนการนี้จะดำเนินการในเฟรมที่สองโดยใช้วิธีแบ็คกราวนด์ซับแทรคชันบนพื้นฐานของการเคลื่อนไหว (Motion based) หากจำนวนผู้เล่นและกรรมการยังไม่ครบตามจำนวนที่กำหนด คือ 23 คน ระบบจะทำการตรวจจับผู้เล่นเพิ่มเติมในเฟรมถัดไปจนกว่าจะได้ครบตามจำนวนที่กำหนดระบบจึงจะหยุดดำเนินการตรวจจับผู้เล่น ในงานนี้จะพิจารณาเฉพาะบุคคลที่อยู่ภายในสนามเท่านั้นซึ่งประกอบด้วย ผู้เล่นทั้ง 2 ทีมและกรรมการกลาง ส่วนกรรมการข้างสนามจำนวน 2 คน จะอยู่นอกขอบเขตที่พิจารณา ดังนั้นจำนวนรวมของบุคคลที่พิจารณาจึงมีจำนวนเท่ากับ 23 คน นอกจากนี้กระบวนการตรวจจับผู้เล่นนี้จะดำเนินการโดยอาศัยความรู้พื้นฐานเกี่ยวกับขอบเขตของสนาม ซึ่งข้อมูลขอบเขตสนามนี้ได้มาจากขั้นตอนการตรวจจับสนามและการโมเดลสนาม สุดท้ายผลของการทำแบ็คกราวด์ซับแทรคชันนี้จะได้รายการวัตถุที่มีความน่าจะเป็นที่จะเป็นผู้เล่นและกรรมการกลางมาจำนวนหนึ่ง ดังสมการที่ 5.3

$$forground_blob_1^N = \left\{ I_{new_frame}^{crop_image} - I_{blackground_frame}^{crop_image} > Th \right\} \qquad (5.3)$$

โดยที่ *forground_blob* คือ เอาท์พุตรายการบล็อบที่ได้จากกระบวนการทำแบ็คกราวนด์ซับแทรคชันและ *Th* คือ ค่าเทรชโฮลด์ ซึ่งถูกใช้เป็นเกณฑ์ในการจำแนกระหว่างโฟร์กราวนด์และแบ็คกราวนด์

1.3 พิจารณาความเป็นได้ที่จะเป็นผู้เล่น (Candidate Player Decision)

ในขั้นตอนการพิจารณาความเป็นไปได้ที่จะเป็นผู้เล่นฟุตบอลนี้จะถูกแบ่งออกเป็น 5 ขั้นตอนย่อย คือ

1.3.1 นำรายการโฟร์กราวนด์บล็อบที่ได้จากขั้นตอนการตรวจจับสนามในเฟรมแรกมาพิจารณาถึงความเป็นไปได้ที่จะเป็นผู้เล่นและกรรมการกลาง โดยทำการขจัดสัญญาณรบกวนด้วยวิธีการดำเนินการในลักษณะทางสัณฐานวิทยา (Morphological Operations) จากนั้นทำการพิจารณาความเป็นไปได้ที่จะเป็นผู้เล่นจากข้อมูลเอกลักษณ์ สีและขนาด เป็นต้น สำหรับเกณฑ์การพิจารณามี 3 ข้อ คือ

1) ใช้ขอบเขตสนามในการกำหนดพื้นที่การค้นหาดังนั้นสันนิษฐานในเบื้องต้นว่าวัตถุนั้น คือ ผู้เล่นหรือกรรมการ ส่วนวัตถุที่อยู่นอกขอบเขตสนามสันนิษฐานว่าไม่ใช่ผู้เล่นหรือกรรมการจะไม่นำมาพิจารณา ดังสมการที่ 5.4

$$pf_player_1^N = gm_X_1^M(x) \mid x(x,y) \in corner_point_1^4 \tag{5.4}$$

2) ถ้าวัตถุมีขนาดที่ใกล้เคียงกันจำนวนตั้งแต่ 2 ขึ้นไป สันนิษฐานว่าวัตถุนั้น คือ ผู้เล่นหรือกรรมการ ดังสมการที่ 5.5

$$ps_player_1^N = gm_X_1^M(x^i) \mid \mathrm{count}(x_{size}^i = x_{size}^{1,\ldots,n}) >= 2 \tag{5.5}$$

3) ถ้าวัตถุมีสีที่ใกล้เคียงกันจำนวนตั้งแต่ 2 ขึ้นไป สันนิษฐานว่าวัตถุนั้น คือ ผู้เล่นหรือกรรมการ ดังสมการที่ 5.6

$$pc_player_1^N = gm_X_1^M(x^i) \mid \mathrm{count}(x_{color}^i = x_{color}^{1,\ldots,n}) >= 2 \tag{5.6}$$

โดยที่ gm_X คือ รายการวัตถุที่ได้จากการตรวจจับสนาม N และ M คือจำนวนของวัตถุ

1.3.2 นำรายการโฟร์กราวนด์บล็อบที่ได้จากขั้นตอนการตรวจจับผู้เล่นโดยใช้วิธีแบ็คกราวนด์ซับแทรคชันในเฟรมที่สองมาพิจารณาถึงความเป็นไปได้ที่จะเป็นผู้เล่นและกรรมการกลาง โดยทำการขจัดสัญญาณรบกวนด้วยวิธีการดำเนินการในลักษณะทางสัณฐานวิทยา จากนั้นทำการพิจารณาความเป็นไปได้ที่จะเป็นผู้เล่นจากข้อมูลเอกลักษณ์ สีและขนาด เป็นต้น สำหรับเกณฑ์การพิจารณามี 3 ข้อ คือ

1) ใช้ขอบเขตสนามในการกำหนดพื้นที่การค้นหาดังนั้นสันนิษฐานในเบื้องต้นว่าวัตถุนั้น คือ ผู้เล่นหรือกรรมการ ส่วนวัตถุที่อยู่นอกขอบเขตสนามสันนิษฐานว่าไม่ใช่ผู้เล่นหรือกรรมการจึงไม่นำมาพิจารณา ดังสมการที่ 5.7

$$bf_player_1^N = bg_X_1^M(\mathrm{x}) \mid x(x,y) \in corner_po\mathrm{int}_1^4 \tag{5.7}$$

2) ถ้าวัตถุมีขนาดที่ใกล้เคียงกันจำนวนตั้งแต่ 2 ขึ้นไป สันนิษฐานว่าวัตถุนั้น คือ ผู้เล่นหรือกรรมการ ดังสมการที่ 5.8

$$bs_player_1^N = bg_X_1^M(\mathrm{x}^i) \mid \mathrm{count}(x^i_{size} = x^{1,\dots,n}_{size}) >= 2 \tag{5.8}$$

3) ถ้าวัตถุมีสีที่ใกล้เคียงกันจำนวนตั้งแต่ 2 ขึ้นไป สันนิษฐานว่าวัตถุนั้น คือ ผู้เล่นหรือกรรมการ ดังสมการที่ 5.9

$$bc_player_1^N = bg_X_1^M(\mathrm{x}^i) \mid \mathrm{count}(x^i_{color} = x^{1,\dots,n}_{color}) >= 2 \tag{5.9}$$

โดยที่ bg_X คือ รายการวัตถุที่ได้จากการตรวจจับผู้เล่น N คือ จำนวนของวัตถุที่ผ่านเงื่อนไขและ M คือ จำนวนของวัตถุทั้งหมดที่ตรวจจับได้

เหตุผลที่กำหนดจำนวนขั้นต่ำเนื่องจากอาจมีวัตถุอื่นที่เข้าเงื่อนไขในข้อแรก คือ อยู่ในสนาม เช่น ลูกบอล แต่โดยปกติในเกมส์การแข่งขันลูกบอลนี้จะมีจำนวนเพียง 1 เท่านั้น จึงแก้ปัญหาโดยกำหนดจำนวนขั้นต่ำ คือ 2 วัตถุที่เหมือนกัน เพราะโดยปกติผู้เล่นและกรรมการจะมีมากกว่า 1 คน แต่ในกรณีที่เป็นสิ่งมีชีวิตชนิดอื่น เช่น นก ที่บินผ่านอาจมีมากกว่า 1 ตัว การใช้จำนวนขั้นต่ำเป็น 2 วัตถุที่เหมือนกันจะไม่สามารถขจัดวัตถุเหล่านี้ได้ แต่จะแก้ปัญหาโดยการจัดเรียงลำดับของกลุ่มตามขนาด แล้วพิจารณาจากขนาดหากวัตถุกลุ่มใดมีขนาดเล็กกว่าวัตถุกลุ่มอื่นๆ จะสันนิษฐานว่าวัตถุในกลุ่มนั้นไม่ใช่ผู้เล่นและกรรมการ เพราะโดยปกติแล้วผู้เล่นจะเป็นกลุ่มที่มีขนาดใหญ่ที่สุดในสนาม รองลงมาจะเป็นวัตถุอื่นๆ เช่น ลูกบอล และเส้นสนาม เป็นต้น

1.3.3 จากเกณฑ์ที่กำหนดในหัวข้อ 1.3.1) และ 1.3.2) นี้ จะถูกนำมาใช้เป็นเงื่อนไขในการคำนวณหาค่าความน่าจะเป็นของแต่ละวัตถุที่ตรวจจับได้ โดยในแต่ละเกณฑ์นั้นวัตถุจะถูกคำนวณหาค่าความน่าจะเป็นที่จะเป็นผู้เล่นและกรรมการ ถ้าตรงตามเงื่อนไขนั้นจะถูกกำหนดค่าให้เป็น หนึ่ง ซึ่งหมายถึงเป็นผู้เล่นหรือกรรมการ มิฉะนั้นจะถูกกำหนดค่าให้เป็น ศูนย์ ซึ่งหมายถึงวัตถุอื่นๆ ดังสมการที่ 5.10 ถึงสมการที่ 5.15

$$\text{pro_pf_}player_1^N = \begin{cases} 1, when\ gm_X_1^N(x) \in pf_player \\ 0, otherwise \end{cases} \tag{5.10}$$

$$\text{pro_ps_}player_1^N = \begin{cases} 1, when\ gm_X_1^N(x) \in ps_player \\ 0, otherwise \end{cases} \tag{5.11}$$

$$\text{pro_pc_}player_1^N = \begin{cases} 1, when\ gm_X_1^N(x) \in pc_player \\ 0, otherwise \end{cases} \tag{5.12}$$

$$\text{pro_bf_}player_1^N = \begin{cases} 1, when\ bg_X_1^N(x) \in bf_player \\ 0, otherwise \end{cases} \tag{5.13}$$

$$\text{pro_bs_}player_1^N = \begin{cases} 1, when\ bg_X_1^N(x) \in bs_player \\ 0, otherwise \end{cases} \tag{5.14}$$

$$\text{pro_bc_}player_1^N = \begin{cases} 1, when\ bg_X_1^N(x) \in bc_player \\ 0, otherwise \end{cases} \tag{5.15}$$

เมื่อพิจารณาครบทุกเกณฑ์แล้วจะทำการคำนวณหาผลรวมของความน่าจะเป็นทั้งหมด จากนั้นจะทำการจับกลุ่มตามขนาดของวัตถุจำนวน 3 กลุ่มโดยใช้อัลกอริทึมเคมีน และเรียงลำดับกลุ่มจากกลุ่มที่มีวัตถุที่มีขนาดโดยเฉลี่ยใหญ่ที่สุดไปหากลุ่มที่มีวัตถุที่มีขนาดโดยเฉลี่ยเล็กที่สุดและในแต่ละกลุ่มจะจัดเรียงลำดับความน่าจะเป็นของทุกวัตถุตามค่าความน่าจะเป็นที่ได้โดยเรียงจากมากไปหาน้อย ดังสมการที่ 5.16 ถึงสมการที่ 5.19

$$gm_player(weight)_1^N = \frac{\sum_1^N \left(pro_pf_player_1^N, pro_ps_player_1^N, pro_pc_player_1^N \right)}{3} \tag{5.16}$$

$$\text{bg_}player(weight)_1^N = \frac{\sum_1^N \left(pro_\text{bf}_player_1^N, pro_\text{bs}_player_1^N, pro_\text{bc}_player_1^N \right)}{3} \tag{5.17}$$

$$\text{gm_X}_1^N = \text{sort}_{by\ object\ size(mean)}^{all\ group}\left(sort_{by\ gm_player(weight)}^{each\ group}\left(group_{by\ object\ size}^{1,2,3}\left(\text{gm_X}_1^N\right)\right)\right) \tag{5.18}$$

$$bg_X_1^N = sort_{by\ object\ size(mean)}^{all\ group}(sort_{by\ bg_player(weight)}^{each\ group}(group_{by\ object\ size}^{1,2,3}(bg_X_1^N))) \quad (5.19)$$

จากการเรียงลำดับตามความน่าจะเป็นนี้จะมีกรรมการกลางเพียงคนเดียวที่จะมีโอกาสถูกจัดลำดับอยู่ในลำดับสุดท้ายในกลุ่ม เนื่องจากอาจมีเอกลักษณ์สีที่ต่างจากผู้เล่น สุดท้ายหลังจากการจัดเรียงลำดับตามขนาดของวัตถุและตามความน่าจะเป็นที่จะเป็นผู้เล่นแล้วจึงทำการเลือกรายการวัตถุจำนวน 23 รายการแรก โดยเลือกรายการวัตถุที่อยู่ในกลุ่มที่มีขนาดใหญ่ที่สุดก่อน โดยสันนิษฐานว่าผู้เล่นจะมีขนาดใหญ่ที่สุดในสนาม หากจำนวนยังไม่ครบ 23 คน จึงจะเลือกจากกลุ่มที่มีขนาดรองลงมาตามลำดับ โดยสันนิษฐานว่าผู้เล่นอาจถูกจับกลุ่มได้มากกว่า 1 กลุ่มเนื่องจากผู้เล่นแต่ละคนอาจปรากฏอยู่ในภาพในตำแหน่งที่ต่างกันและอยู่ห่างจากกล้องไม่เท่ากันทำให้มีขนาดที่ต่างกัน สำหรับในกรณีที่เฟรมภาพนั้นมีผู้เล่นปรากฏอยู่ไม่ถึง 23 คน อาจเกิดการผิดพลาดซึ่งระบบจะทำการเลือกวัตถุจากกลุ่มที่มีขนาดเล็กรองลงมาซึ่งอาจเป็นกลุ่มของวัตถุอื่นๆ ที่ไม่ใช่ผู้เล่น ในงานนี้แก้ปัญหาโดยสันนิษฐานว่าวัตถุที่เป็นผู้เล่นจะมีขนาดและรูปร่างที่ไม่แตกต่างกันมากนัก ดังนั้นกลุ่มใดที่มีขนาดของวัตถุโดยเฉลี่ยแล้วต่างจากกลุ่มแรกเกิน 50% จะถือว่าไม่ใช้ผู้เล่น ดังนั้นกลุ่มของวัตถุกลุ่มนี้จะไม่ถูกนำมาพิจารณา

1.3.4 รายการโฟร์กราวนด์บล็อบที่ได้จากขั้นตอนการตรวจจับสนามในเฟรมแรกและจากขั้นตอนการตรวจจับผู้เล่นในเฟรมที่สองโดยใช้วิธีแบ็คกราวนด์ซับแทรคชันจะถูกนำมาพิจารณาร่วมกันอีกครั้ง จุดประสงค์เพื่อเพิ่มความแม่นยำในการตรวจจับผู้เล่น เพราะถ้าการตรวจจับผู้เล่นมีความผิดพลาดจะส่งผลต่อกระบวนการติดตามด้วยเช่นกัน ดังนั้นในงานนี้จึงอาศัยข้อมูลที่ได้จากวิธีการตรวจจับสนามโดยใช้โมเดลผสมแบบเกาเซียนและข้อมูลที่ได้จากการตรวจจับผู้เล่นโดยใช้วิธีแบ็คกราวนด์ซับแทรคชันมาทวนสอบกันเพื่อให้ผลการตรวจจับผู้เล่นมีความแม่นยำมากขึ้น โดยพิจารณาจากเอกลักษณ์ สี ขนาดและการเคลื่อนไหวและมีเกณฑ์การพิจารณา คือ

1) ถ้าโฟร์กราวนด์บล็อบทั้ง 2 กลุ่มนี้มีวัตถุที่มีขนาดใกล้เคียงกันอย่างน้อย 2 วัตถุ สันนิษฐานว่าวัตถุที่มีขนาดเดียวกันนี้เป็นผู้เล่นหรือกรรมการ ดังสมการที่ 5.20

$$pro_s_gm(weight)_1^N, pro_s_bg(weight)_1^M = \begin{cases} 1\ if\ count\left(gm_X_1^N(x)_{size} = bg_X_1^N(x)_{size}\right) >= 2 \\ 0, otherwise \end{cases} \quad (5.20)$$

2) ถ้าโฟร์กราวนด์บล็อบทั้ง 2 กลุ่มนี้มีวัตถุที่มีสีใกล้เคียงกันอย่างน้อย 2 วัตถุ สันนิษฐานว่าวัตถุที่มีสีลักษณะเดียวกันนี้เป็นผู้เล่นหรือกรรมการ ดังสมการที่ 5.21

$$pro_c_gm(weight)_1^N, pro_c_bg(weight)_1^M = \begin{cases} 1 \ if \ count\left(gm_X_1^N(x)_{color} = bg_X_1^N(x)_{color}\right) >= 2 \\ 0, otherwise \end{cases} \quad (5.21)$$

3) ถ้าโฟร์กราวนด์บล็อบทั้ง 2 กลุ่มนี้มีวัตถุที่มีการเคลื่อนไหวอย่างน้อย 1 วัตถุ สันนิษฐานว่าวัตถุที่มีเอกลักษณ์เดียวกันนี้เป็นผู้เล่นหรือกรรมการ ดังสมการที่ 5.22

$$pro_m_gm(weight)_1^N, pro_m_bg(weight)_1^M = \begin{cases} 1 \ if \ count\left(gm_X_1^N(x)_{size,color} = bg_X_1^N(x)_{size,color}\right) \\ \& \ distance\left(gm_X_1^N(x)_{(x,y)}, bg_X_1^N(x)_{(x,y)}\right) < 5pixel \\ 0, otherwise \end{cases} \quad (5.22)$$

คำนวณหาผลรวมของความน่าจะเป็น ดังสมการที่ 5.23 ถึงสมการที่ 5.26

$$new_gm_player(weight)_1^N = \frac{\sum_1^N \left(pro_s_gm_1^N, pro_c_gm_1^N, pro_m_gm_1^N\right)}{3} \quad (5.23)$$

$$new_bg_player(weight)_1^N = \frac{\sum_1^N \left(pro_s_bg_1^N, pro_c_bg_1^N, pro_m_bg_1^N\right)}{3} \quad (5.24)$$

$$new_gm_X_1^N = sort_{by\ new_gm_player(weight)}(gm_X_1^N) \quad (5.25)$$

$$new_bg_X_1^N = sort_{by\ new_bg_player(weight)}(bg_X_1^N) \quad (5.26)$$

1.3.5 ทำการจับคู่วัตถุจากทั้งสองกลุ่มเข้าด้วยกัน เนื่องจากในเฟรมแรกวัตถุยังไม่มีการเคลื่อนที่แต่ในเฟรมที่สองบางวัตถุอาจมีการเคลื่อนที่ออกจากตำแหน่งเดิม ดังนั้นจึงต้องทำการจับคู่วัตถุที่ได้จากสองเฟรมนี้ที่เป็นวัตถุหรือผู้เล่นคนเดียวกันเข้าด้วยกัน ซึ่งจะได้จำนวนวัตถุที่เป็นจริง มิฉะนั้นถ้าใช้รายการโฟร์กราวนด์บล็อบจากทั้ง 2 กลุ่มเลยจะได้จำนวนที่เกินจากความเป็นจริง โดยหลังจากพิจารณาความเหมือนในเรื่องของสีและขนาดแล้ว ในขั้นตอนนี้จะพิจารณาจากข้อมูล

ระยะทางที่น้อยที่สุดระหว่างตำแหน่งของวัตถุที่ปรากฏในเฟรมแรกกับตำแหน่งของวัตถุที่ปรากฏในเฟรมที่สองซึ่งสันนิษฐานว่าเป็นวัตถุเดียวกัน ดังสมการที่ 5.27

$$newplayer_1^N = match\left(new_gm_X_1^N, new_bg_X_1^N\right)_{by\ distance}^{by\ size,color} \quad (5.27)$$

รายการวัตถุที่ได้นี้จะถูกนำไปใช้สำหรับกระบวนการติดตามต่อไปโดยสันนิษฐานว่าเป็นรายการที่มีความน่าจะเป็นที่จะเป็นผู้เล่นหรือกรรมการมากที่สุด

1.4 การประมาณตำแหน่งผู้เล่น (Player Location Estimation)

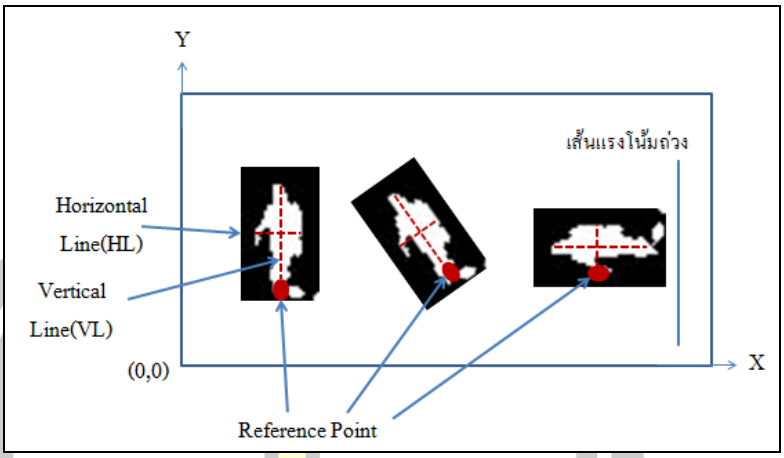
การประมาณตำแหน่งของผู้เล่นจะใช้จุดบริเวณตำแหน่งที่อยู่ล่างสุดของพื้นที่ผู้เล่นเป็นตำแหน่งอ้างอิงโดยปกติ คือ ตำแหน่งเท้าของผู้เล่น เหตุผลที่ใช้ตำแหน่งล่างสุดของวัตถุเป็นตำแหน่งอ้างอิงเนื่องจากในความเป็นจริงนั้นตำแหน่งล่างสุดของวัตถุนี้จะเป็นบริเวณที่อยู่ใกล้หรือสัมผัสกับพื้นดินมากที่สุด ซึ่งเป็นตำแหน่งที่ใกล้กับตำแหน่งจริงในระบบพิกัดของมุมมองโลกมากที่สุด นอกจากนี้เนื่องจากโดยทั่วไปผู้เล่นจะอยู่ในลักษณะท่ายืน นั่งและนอน ดังนั้นในงานนี้จึงใช้ตำแหน่งของตัวผู้เล่นที่อยู่ในตำแหน่งล่างสุดและอยู่ตรงกลางวัตถุเป็นเกณฑ์ในการกำหนดตำแหน่งอ้างอิง เพราะในกรณีที่ผู้เล่นอยู่ในท่ายืนแต่กางขาและอยู่ในท่านั่งและท่านอน เท้าของผู้เล่นจะไม่ใช่จุดที่เป็นศูนย์กลางของตัวผู้เล่นหรือวัตถุ ดังนั้นการใช้ตำแหน่งเท้าเพียงอย่างเดียวจะส่งผลให้ตำแหน่งอ้างอิงมีความคลาดเคลื่อนไปจากตำแหน่งจริงในระบบพิกัดโลก ดังนั้นงานนี้จึงดำเนินการโดยใช้ตำแหน่งในแนวกลางของวัตถุร่วมกับตำแหน่งล่างสุดของวัตถุในการกำหนดตำแหน่งอ้างอิง โดยมีหลักการคำนวณ คือ ทำการหาเส้นแกนกลางในแนวตั้งและแนวนอนของพื้นที่จากค่าเฉลี่ยหรือค่ากลางของความกว้างและของความยาวของพื้นที่ ดังสมการที่ 5.28 และสมการที่ 5.29

$$ver_center(x,y)_{top,bottom} = (x,y)_1^N \in regionblob(\text{width_size})_{newplayer_1^N} / 2 \quad (5.28)$$

$$hor_center(x,y)_{left,right} = (x,y)_1^N\ regionblob(\text{high_size})_{newplayer_1^N} / 2 \quad (5.29)$$

โดยที่ *ver_center* คือ รายการตำแหน่งในแนวตั้งซึ่งมี 2 ตำแหน่ง คือ บริเวณด้านบนและด้านล่างของบล็อบพื้นที่และ *hor_center* คือ รายการตำแหน่งในแนวนอนซึ่งมี 2 ตำแหน่ง คือ บริเวณด้านซ้ายและด้านขวาของบล็อบพื้นที่ จากนั้นทำการค้นหาตำแหน่งที่อยู่ต่ำที่สุดในพื้นที่ซึ่งจะอยู่ใกล้ระนาบพื้นมากที่สุดโดยใช้ลักษณะทางกายภาพของระบบพิกัดภาพเป็นเกณฑ์ ดังสมการที่ 5.30 ซึ่งจะใช้ตำแหน่งดังกล่าวนี้เป็นจุดอ้างอิงตำแหน่งของวัตถุ แสดงดังภาพที่ 5.4

$$ref_position(x,y) = find(\max_y([\text{vertical_center}(x,y)\ \text{horizontal_center}(x,y)])) \quad (5.30)$$



ภาพที่ 5.4 หลักการประมาณตำแหน่งผู้เล่น

ตามภาพที่ 5.4 ในกรณีที่ผู้เล่นอยู่ในท่ายืนตามปกติ ตำแหน่งอ้างอิงที่คำนวณได้จะอยู่ตรงบริเวณเท้าของผู้เล่น ส่วนในกรณีที่ผู้เล่นอยู่ในลักษณะเอียงตัว ตำแหน่งอ้างอิงโดยส่วนใหญ่จะยังเป็นตำแหน่งตรงบริเวณเท้าเนื่องจากบริเวณเท้ายังเป็นบริเวณที่อยู่ใกล้หรือสัมผัสกับพื้นดินมากที่สุด แต่ในกรณีที่ผู้เล่นอยู่ในท่านั่งหรือท่านอนจะมีส่วนของร่างกายผู้เล่นที่อยู่ใกล้หรือสัมผัสกับพื้นดินหลายแห่ง ดังนั้นโดยธรรมชาติเราจะยึดตำแหน่งศูนย์กลางของตัวผู้เล่นซึ่งเป็นตำแหน่งของจุดศูนย์ถ่วงของวัตถุเป็นเกณฑ์ ดังนั้นจึงเป็นเหตุผลที่งานนี้ใช้เส้นกึ่งกลางของวัตถุเป็นเกณฑ์ในการกำหนดตำแหน่งอ้างอิงซึ่งในกรณีนี้ คือ เส้นกึ่งกลางตามแนวนอนหรือความกว้างและเส้นกึ่งกลางตามแนวตั้งหรือความสูงของวัตถุ

2. ขั้นตอนการจำแนกผู้เล่น (Player Classification)

หลังจากขั้นตอนการพิจารณาความเป็นได้ที่จะเป็นผู้เล่นจะได้รายการวัตถุที่เป็นผู้เล่นและกรรมการจำนวน 23 รายการ $player(X)_1^{23} = regionblob_{newplayer_1^N}$ ในขั้นตอนนี้จะทำการจำแนกตัวตน (Identification) ของผู้เล่นออกเป็นกลุ่มต่างๆ คือ การจำแนกทีมผู้เล่น การจำแนกกรรมการกลางและการจำแนกผู้เล่นรายบุคคล ซึ่งมีหลักการพิจารณาดังนี้

เอกลักษณ์และความรู้พื้นฐานที่ใช้สำหรับการจำแนก คือ เอกลักษณ์สีและข้อมูลความรู้เกี่ยวกับตำแหน่งการยืนของผู้เล่น และข้อมูลฝั่งของสนามซึ่งแบ่งเป็นฝั่งซ้ายและฝั่งขวาโดยใช้ข้อมูลพิกัดบนเส้นครึ่งสนามจากภาพเสมือนเป็นเกณฑ์ $playfield_side = (x, y)_{on\ virtualview}$ ซึ่งจะดำเนินการด้วยมือ รวมถึงข้อมูลจำนวนของวัตถุที่มีเอกลักษณ์ที่เหมือนกัน ส่วนเกณฑ์เบื้องต้นในการพิจารณามีอยู่ 2 เกณฑ์ คือ

1) ถ้าวัตถุมีลักษณะของสีที่คล้ายกันเกิน 2 คน สันนิษฐานว่าวัตถุนั้น คือ ผู้เล่น หากน้อยกว่า 2 คน สันนิษฐานว่าวัตถุนั้น คือ กรรมการกลาง ดังสมการที่ 5.31

$$\text{player_ref}(X)_1^N = \begin{cases} player, if\ count(\text{player}(x)_i^{color} = \text{player}(x)_{1,...,n}^{color}) >= 2 \\ reference, otherwise \end{cases} \quad (5.31)$$

2) ถ้าวัตถุที่มีลักษณะของสีคล้ายกันเกิน 2 คน มี 2 กลุ่ม โดยแยกตามเอกลักษณ์ของสี สันนิษฐานว่าวัตถุนั้น คือ ผู้เล่นทีม A และผู้เล่นทีม B ดังสมการที่ 5.32

$$\text{player_AB}(X)_1^N = \begin{cases} team\ A, if\ count(\text{player}(x)_i^{color} = \text{player}(x)_{1,...,n}^{color}) >= 2 \\ team\ B, if\ count(\text{player}(x)_i^{color} = \text{player}(x)_{1,...,n}^{color}) >= 2 \end{cases} by\ A^{color} \neq B^{color} \quad (5.32)$$

หลักการพิจารณาในการจำแนกตัวตนจะมีพื้นฐานบนเอกลักษณ์ของความหนาแน่นของสี โดยกระบวนการจำแนกจะอาศัยการกระจายตัวของความหนาแน่นของสีซึ่งจะประยุกต์ใช้หลักการของ Figuerou และคณะ [59] โดยในงานนี้จะพิจารณาเฉพาะภายในพื้นที่ของบล็อบวัตถุซึ่งมีอยู่ 3 แนวทาง โดยมีรายละเอียดดังนี้

1) การจำแนกทีมผู้เล่นและกรรมการจะพิจารณาจากเอกลักษณ์ของความหนาแน่นของสี โดยสันนิษฐานว่ากรรมการกลางจะมีเอกลักษณ์ของสีที่ต่างจากผู้เล่นแม้จะมีลักษณะของสีเสื้อที่คล้ายกัน

2) การจำแนกระหว่าง 2 ทีมผู้เล่นจะใช้เอกลักษณ์ของความหนาแน่นของสีร่วมกับความรู้พื้นฐานเกี่ยวกับโครงสร้างสนามและแผนผังการยืนของผู้เล่น ซึ่งปกติสนามจะแบ่งออกเป็น 2 ฝั่งและก่อนการเริ่มเกมส์แต่ละทีมจะยืนประจำตำแหน่งในฝั่งของตนเอง จากข้อมูลที่มีอยู่นี้สามารถใช้จำแนกทีมผู้เล่นออกเป็น 2 ทีมได้ แม้ผู้เล่นของทั้งสองทีมจะสวมเสื้อผ้าสีคล้ายกัน ดังสมการที่ 5.33

$$\text{player_AB}(X)_1^N = \begin{cases} team\ A, if\ player_1^N(x,y) \mid \text{x} < \text{playfield_side(x, y)} \\ \text{team B}, if\ player_1^N(x,y) \mid \text{x} > \text{playfield_side(x, y)} \end{cases} \quad (5.33)$$

สำหรับการจำแนกผู้รักษาประตูออกจากผู้เล่นคนอื่นๆ ในทีมจะอาศัยความรู้จากตำแหน่งการยืนของผู้เล่นซึ่งปกติก่อนการเริ่มเกมส์ในช่วงเฟรมแรกผู้รักษาประตูมักจะยืนอยู่ในตำแหน่งท้ายสุดของฝั่งตนเองหรืออยู่หน้าประตูเป็นส่วนใหญ่ ดังสมการที่ 5.34 และสมการที่ 5.35

$$goal_keeper(x,y)_{team\ A} = player(X)_1^N(x,y) \mid \text{find(min(x) \& x} < \text{playfield_side(x, y))} \quad (5.34)$$

$$goal_keeper(x,y)_{team\ B} = player(X)_1^N(x,y) \mid \mathrm{find}(\max(x)\ \&\ x > \mathrm{playfield_side}(x,y)) \quad (5.35)$$

3) ขณะที่มีการบดบังระหว่างผู้เล่นจะใช้ข้อมูลของเอดจ์ (Edge) ของบล็อกที่เหลือในการประมาณตำแหน่งและจำแนกตัวบุคคล กรณีที่ไม่มีเอดจ์ใดสามารถใช้ได้จะประยุกต์ใช้ข้อมูลตำแหน่งและบราวดิ้งบล็อกที่ได้จากมุมมองเสมือนในการประมาณตำแหน่งและจำแนกตัวตน

3. ขั้นตอนการแสดงผลผู้เล่น (Player Representation)

หลังจากขั้นตอนการจำแนกผู้เล่นและการประมาณตำแหน่งผู้เล่น ข้อมูลของพื้นที่บล็อบจะถูกใช้ในการกำหนดเป็นบราวดิ้งบล็อกสำหรับแสดงผลในลักษณะ 2 มิติในมุมมองภาพ ตัวอย่างแสดงดังภาพที่ 5.5

ภาพที่ 5.5 การแสดงผลบราวดิ้งบล็อกผู้เล่น

ข้อมูลของบราวดิ้งบล็อกจะเก็บข้อมูลพิกัดตำแหน่งมุมบนซ้าย ความกว้างและความยาวของบล็อก ดังสมการที่ 5.36

$$bounding_box = \{(x,y), widght, height\} \quad (5.36)$$

โดยในที่นี้จะกำหนดขนาดความกว้างและความยาวของบล็อกตามขนาดของพื้นที่บล็อบที่ตรวจจับได้

4. ขั้นตอนการติดตามผู้เล่น (Player Tracking)

ขั้นตอนการติดตามผู้เล่นในแบบมุมมองภาพ (Image View) ประยุกต์ใช้อัลกอริทึมการติดตาม 2 เทคนิค คือ ตัวกรองคาร์มาลฟิลเตอร์และอัลกอริทึมคอนเดนเซชันโดยทำการทดลองใช้ 2 วิธีนี้เป็นพื้นฐานในการศึกษาเพื่อหาแนวทางในการพัฒนาเทคนิคใหม่เพื่อเพิ่มประสิทธิภาพของระบบการตรวจจับและติดตามสำหรับใช้ในงานวิจัยต่อไป ซึ่งมี 2 ขั้นตอนย่อย โดยมีรายละเอียดดังนี้

4.1 การศึกษาประสิทธิภาพของเทคนิคการติดตาม

การศึกษาเทคนิคการติดตามโดยใช้ 2 อัลกอริทึมพื้นฐาน คือ ตัวกรองคาร์มาลฟิลเตอร์และอัลกอริทึมคอนเดนเซชันในการติดตามผู้เล่น โดยข้อมูลที่ใช้ทดสอบ คือ วิดีโอการแข่งขันฟุตบอลของฐานข้อมูล ISSIA Dataset [82] เป็นภาพจากกล้องตัวที่ 1 จากทั้งหมด 6 ตัว โดยใช้เฟรมภาพในการประเมินประสิทธิภาพจำนวน 108 เฟรม ทำการวัดประสิทธิภาพเฉพาะเฟรมที่ 1186 ถึงเฟรมที่ 1294 ในกรณีนี้ใช้ตัวตรวจจับโฟร์กราวนด์ (Foreground Detector) บนพื้นฐานของการเคลื่อนไหว (Motion Based) ในโปรแกรม Matlab สำหรับการตรวจจับผู้เล่น และทำการประเมินประสิทธิภาพของการตรวจจับและติดตามโดยประยุกต์ใช้ 2 พารามิเตอร์จาก 6 พารามิเตอร์ [83] ของเมทริกซ์ประสิทธิภาพ คือ 1. OTE Error และ 2. Time Analysis ผลการทดสอบประสิทธิภาพในการติดตามของตัวกรองคาร์มาลฟิลเตอร์และอัลกอริทึมคอนเดนเซชันและวิธีที่พัฒนาเพิ่มเติม คือการใช้เทคนิคของอัลกอริทึมคอนเดนเซชันร่วมกับเอกลักษณ์สี จากผลการทดสอบเบื้องต้นพบว่าทั้ง 3 วิธีให้ผลที่ไม่แตกต่างกันมากนัก ดังนั้นผู้วิจัยจึงให้ความสนใจเทคนิคของอัลกอริทึมคอนเดนเซชันเนื่องจากในทางทฤษฎีแล้วจะสามารถแก้ปัญหาการบดบังได้ดีกว่าแบบตัวกรองคาร์มาลฟิลเตอร์รวมถึงสามารถรองรับโมเดลข้อมูลได้มากกว่า 1 โมเดลและนอกจากนี้ค่า Object Tracking Error (OTE Error) และ Average Time (AVG Time) ของทั้ง 2 วิธีนั้นไม่แตกต่างกัน

4.2 การพัฒนาเทคนิคการติดตาม

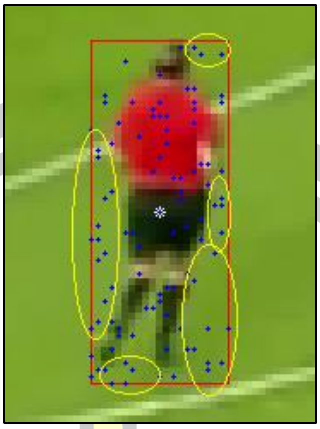
จากการทดสอบเทคนิคการติดตามผู้เล่นโดยใช้ตัวกรองคาร์มาลฟิลเตอร์และเทคนิคที่เสนอ คือ อัลกอริทึมคอนเดนเซชันร่วมกับเอกลักษณ์สี ดังที่มีการศึกษาในหัวข้อที่ 4.1 การศึกษาประสิทธิภาพของเทคนิคการติดตาม ในเบื้องต้นนั้นเนื่องจากทั้งตัวกรองคาร์มาลฟิลเตอร์และอัลกอริทึมคอนเดนเซชันจะใช้ตำแหน่งศูนย์กลางของบราวดิ้งบล็อกในการประมาณตำแหน่งของผู้เล่นจากการสังเกตของผู้วิจัยพบ 2 ประเด็นปัญหาดังนี้

4.2.1 ตำแหน่งศูนย์กลางของบราวดิ้งบล็อกมีการเปลี่ยนแปลงตามลักษณะรูปร่าง

กรณีนี้จะส่งผลให้ตำแหน่งศูนย์กลางของบล็อกไม่อยู่บนพื้นที่หรือบนพื้นที่ของวัตถุจริงแต่จะปรากฎอยู่ในบริเวณพื้นหลังหรืออยู่บริเวณพื้นที่ที่ไม่สำคัญภายในพื้นที่ที่พิจารณา จากสาเหตุนี้สันนิษฐานได้ 2 ประเด็น ดังนี้

1) อาจทำให้การประมาณตำแหน่งโดยใช้ศูนย์กลางบล็อกของตัวกรองคาร์มาลฟิลเตอร์และอัลกอริทึมคอนเดนเซชันคลาดเคลื่อนแม้ว่าจะมีประสิทธิภาพที่ดีเมื่อเปรียบเทียบกับค่ากราวนด์ทรูธ เพราะกราวนด์ทรูธนั้นใช้ศูนย์กลางบล็อกเป็นเกณฑ์ แต่ในความเป็นจริงตำแหน่งที่ได้นั้นในบางช่วงเวลาอาจไม่ใช่ตำแหน่งจริงของวัตถุ

2) อาจส่งผลต่อการใช้เอกลักษณ์สีในการพิจารณาความเหมือนของอัลกอริทึมคอนเดนเซชัน เช่น กรณีที่ศูนย์กลางบล็อกอยู่บริเวณพื้นหลังซึ่งมีลักษณะสีที่แตกต่างจากวัตถุ ดังนั้นเอกลักษณ์สีที่ได้จึงไม่ใช่เอกลักษณ์สีจริงของวัตถุ



ภาพที่ 5.6 แซมเปิลรอบตัวอย่างนอกขอบเขตการตรวจจับ

4.2.2 วิธีการสุ่มแซมเปิลของอัลกอริทึมคอนเดนเซชัน

สันนิษฐานว่า อัลกอริทึมคอนเดนเซชันใช้วิธีการสุ่มบนพื้นฐานของบราวดิ้งบล็อก ซึ่งพบข้อสังเกตอยู่ 2 ประเด็น ดังนี้

1) ตำแหน่งของแซมเปิลที่สุ่มได้บางกลุ่มอาจอยู่นอกพื้นที่ แสดงดังภาพที่ 5.6 ซึ่งสามารถสันนิษฐานได้ 2 เหตุผล ดังนี้

1.1) แม้ตำแหน่งที่ประมาณจะอยู่ใกล้กับศูนย์กลางบล็อกของกราวนด์ทรูธ แต่อาจจะไม่ใช่ตำแหน่งที่อยู่บนพื้นที่จริงของวัตถุและเอกลักษณ์สีที่ได้ก็อาจไม่ใช่เอกลักษณ์สีจริงของวัตถุ ซึ่งอาจเกิดขึ้นกรณีที่มีการเปลี่ยนแปลงรูปร่างของวัตถุทำให้ศูนย์กลางบล็อกอยู่นอกพื้นที่

1.2) เนื่องจากเอกลักษณ์สีของแซมเปิลอยู่นอกพื้นที่ซึ่งไม่ใช่เอกลักษณ์สีจริงของวัตถุจึงไม่สามารถนำไปใช้จำแนกตัวตนระหว่างผู้เล่นได้ โดยเฉพาะในกรณีที่ผู้เล่นอยู่ใกล้กัน

2) ตำแหน่งของแซมเปิลที่สุ่มได้อาจอยู่ห่างจากจุดศูนย์กลางบล็อกหรืออยู่บริเวณใดบริเวณหนึ่งของบล็อกทำให้ผลการประมาณจะอยู่ห่างจากจุดศูนย์กลางเสมอและแซมเปิลส่วนใหญ่อาจอยู่นอกพื้นที่ซึ่งไม่อาจใช้เอกลักษณ์สีในการจำแนกตัวตนได้ ทำให้ไม่เกิดประโยชน์ เปลืองเนื้อที่ในการจัดเก็บและเสียเวลาในการประมวลผล

จากประเด็นปัญหาดังกล่าวผู้วิจัยจึงมีแนวคิดในการปรับปรุงประสิทธิภาพของอัลกอริทึมคอนเดนเซชัน เพื่อแก้ปัญหาตำแหน่งและเอกลักษณ์สีของแซมเปิลที่ไม่อยู่ในพื้นที่จริงของวัตถุ ซึ่งเทคนิคที่เสนอจะช่วยให้สามารถได้ตำแหน่งจริงของวัตถุและสามารถใช้เอกลักษณ์สีจริงของวัตถุในการจำแนกตัวตนได้ อีกทั้งยังช่วยลดแซมเปิลที่ไม่สามารถใช้ในการจำแนกตัวตนได้ทำให้ไม่ต้องเสียเวลาในการประมวลผลแซมเปิลเหล่านี้ นอกจากนี้ยังช่วยให้มั่นใจได้ว่าแซมเปิลส่วนใหญ่จะเป็น

แซมเปิลจริงของวัตถุ เทคนิคที่เสนอ คือ การสุ่มแซมเปิลบนพื้นฐานของพื้นที่ (Region based Sampling) โดยมีหลักการ คือ แซมเปิลที่ใช้ในการสุ่มจะอยู่ในขอบเขตของพื้นที่แทนที่จะอยู่ในขอบเขตของบราวดิ้งบล็อกแบบเดิม

กระบวนการทำงานของอัลกอริทึมคอนเดนเซชันและการสุ่มแซมเปิลบนพื้นฐานของพื้นที่มีดังนี้ คือ จากรายการแซมเปิล ณ เวลาก่อนหน้า ดังสมการที่ 5.37

$$S_{t-1}=\left\{s_{t-1}^{(i)},\pi_{t-1}^{(i)},i=1,\ldots,N\right\} \qquad (5.37)$$

โดยที่ π คือ ค่าน้ำหนักความน่าจะเป็นเริ่มต้นของแต่ละแซมเปิล ทำการสร้างรายการแซมเปิลใหม่ S_t โดยการเลือกแซมเปิล $s_t^{'(n)}$ จากรายการแซมเปิลเดิม S_{t-1} โดยพิจารณาจากค่าน้ำหนักความน่าจะเป็น π_{t-1} จากนั้นทำการทำนาย s_t^{n} โดยการสุ่มแซมเปิลจากสมการที่ 5.38

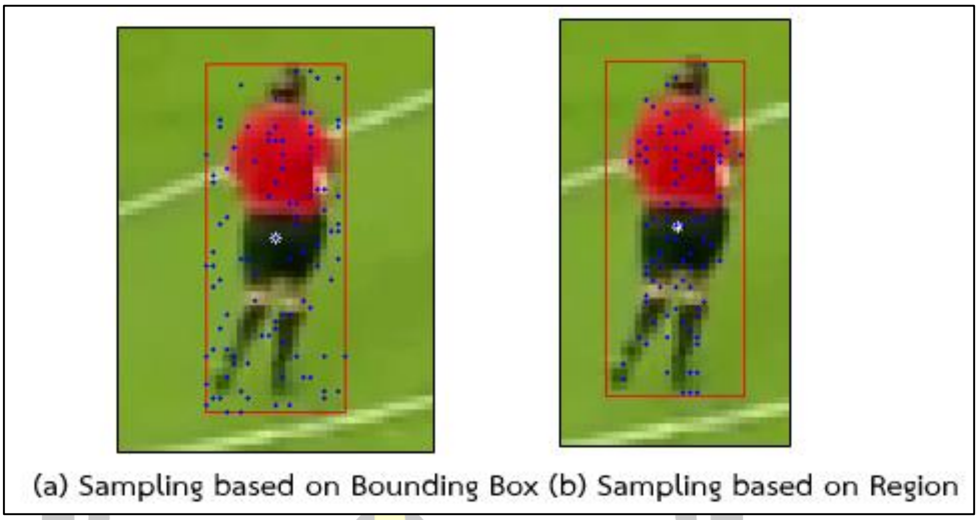
$$p(x_t \mid x_{t-1}=s_t^{'(n)} \; and \; x_{t-1} \in \text{region_blob}) \qquad (5.38)$$

ในขั้นตอนสุดท้ายทำการวัดและให้ค่าน้ำหนักความน่าจะเป็นตามเอกลักษณ์ Z_t แสดงดังสมการที่ 5.39

$$\pi_t^{(n)}=p\left(z_t \mid x_t=s_t^{(n)}\right) \qquad (5.39)$$

ผลลัพธ์สุดท้าย คือ การแจกแจงความน่าจะเป็น $p\left(x_t \mid z_1,\ldots,z_t\right)$ ซึ่งสามารถนำไปใช้ในการประมาณตำแหน่งของวัตถุ

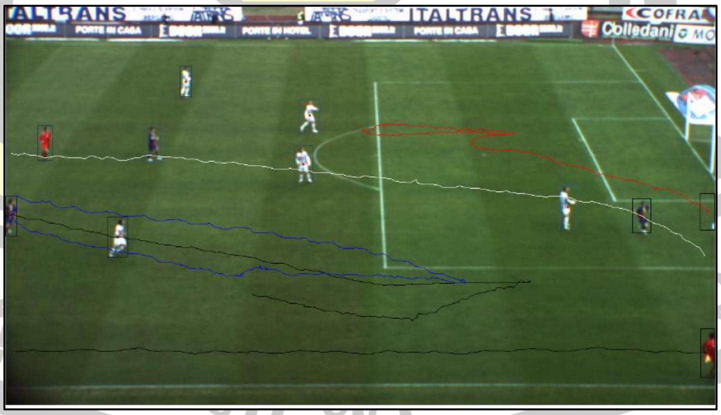
ตัวอย่างการสุ่มแซมเปิลบนพื้นฐานของบราวดิ้งบล็อกและบนพื้นฐานของพื้นที่แสดงดังภาพที่ 5.7 จะพบว่าแซมเปิลที่ได้จากวิธีการสุ่มบนพื้นฐานของบราวดิ้งบล็อกจะมีบางส่วนอยู่นอกพื้นที่และอาจอยู่ห่างจากจุดศูนย์กลางบล็อก ส่วนแซมเปิลที่ได้จากวิธีการสุ่มบนพื้นฐานของพื้นที่จะเป็นแซมเปิลจริงของวัตถุทั้งหมด ซึ่งทำให้ได้เอกลักษณ์ตำแหน่งและเอกลักษณ์สีจริงของวัตถุ



ภาพที่ 5.7 วิธีการสุ่มบนพื้นฐานของบราวดิ้งบล็อกและพื้นที่

5. ขั้นตอนการแสดงเส้นทางผู้เล่นในมุมมองภาพ (Image View based Player Track Representation)

หลังจากกระบวนการตรวจจับและติดตามตำแหน่งผู้เล่นจะได้ข้อมูลตำแหน่งของผู้เล่นแต่ละคนในแต่ละเฟรมภาพ ซึ่งสามารถนำข้อมูลตำแหน่งนี้มาใช้แสดงเส้นทางการเคลื่อนที่ของผู้เล่นแต่ละคนได้โดยใช้หลักการจัดเก็บข้อมูลและการแสดงเส้นทางด้วยเส้นที่มีสีต่างๆ ตามหลักของการเขียนโปรแกรมทั่วไป โดยในส่วนนี้จะเป็นตัวอย่างการแสดงเส้นทางในมุมมองภาพแสดงดังภาพที่ 5.8



ภาพที่ 5.8 ตัวอย่างการแสดงเส้นทางผู้เล่นในมุมมองภาพ

5.2 กระบวนการทดลองและผลการทดลอง (Experimental Process and Results)

ในกระบวนการนี้จะประกอบด้วย 3 ส่วนหลัก คือ การวัดประสิทธิภาพการติดตาม ตัวอย่างผลการติดตาม และตัวอย่างผลการคำนวณขอบเขตสนาม โดยมีรายละเอียดดังนี้

1. การวัดประสิทธิภาพการติดตาม

ในการทดลองครั้งนี้ทำการประเมินประสิทธิภาพของการติดตามโดยใช้ 4 พารามิเตอร์จาก 6 พารามิเตอร์ [83] ของเมทริกซ์ประสิทธิภาพเพื่อประเมินระบบในเบื้องต้น จากข้อมูลของกราวนด์ทรูธจะทำการแปลงพิกัดของบล็อกไปเป็นพิกัดศูนย์กลางของบล็อกจะได้ข้อมูลกราวนด์ทรูธสำหรับประเมินระบบ 2 ส่วน คือ พิกัดศูนย์กลางและพิกัดของบล็อคซึ่งจะถูกใช้ในสมการในแต่ละพารามิเตอร์ดังกล่าว อัลกอริทึมที่ใช้ในการทดลองประกอบด้วย ตัวกรองคาร์มาลฟิลเตอร์ อัลกอริทึมคอนเดนเซชันแบบเดิมและเทคนิคที่เสนอ คือ อัลกอริทึมคอนเดนเซชันร่วมกับเอกลักษณ์สีโดยใช้วิธีการสุ่มบนพื้นฐานของพื้นที่ ข้อมูลทดสอบจากชุดวิดีโอ ISSIA Dataset [82] จำนวนเฟรมในวิดีโอคือ 3003 เฟรม ขนาดเฟรมภาพ 1920 x 1088 พิกเซล อัตราเฟรม 25 เฟรมต่อวินาที ความยาว 2 นาที ในที่นี้ใช้ข้อมูลจากกล้องตัวที่ 1 จากทั้งหมด 6 ตัว โดยทำการทดสอบเพียง 1300 เฟรมและทำการวัดประสิทธิภาพเฉพาะช่วงเฟรมที่ 1186 ถึง 1294 ผลการทดลองสรุปได้ดังตารางที่ 5.1 และตารางที่ 5.2 ดังนี้

ตารางที่ 5.1 ผลการประเมินประสิทธิภาพการติดตาม

Method/Parameter	Distance (Average)	AE (Average)	BBE (Average)	OTE Error (Average)	RMSE Error (Average)
Kalman	92.914	713.04	0	8.0097e+05	8.9699e+05
Condensation	90.347	782.81	0	8.4006e+05	9.1697e+05
Proposed Method	92.971	705.83	0.0022881	8.179e+05	6.0149e+05

ตามตารางที่ 5.1 แสดงผลการประเมินประสิทธิภาพโดยพิจาณาจากค่าเฉลี่ยรวมจากทุกเฟรมโดยในแต่ละเฟรมทำการหาค่าเฉลี่ยรวมของผู้เล่นทุกคนที่ตรวจจับได้ พารามิเตอร์หลักที่ใช้ประเมินประกอบด้วย Distance, Absolute Error (AE), Object Tracking Error (OTE) และ Root Mean Square Error (RMSE) ซึ่งเป็นค่าความผิดพลาดระหว่างค่าที่ระบบประมาณได้กับค่าของกราวนด์ทรูธโดยค่าน้อยแสดงว่าประสิทธิภาพดีและค่า Bounding Box Error (BBE) ซึ่งเป็นการเปรียบเทียบพื้นที่ที่ซ้อนทับกันระหว่างบราวดิ้งบล็อกของระบบที่คำนวณได้กับบราวดิ้งบล็อกของกราวนด์ทรูธโดยค่าเข้าใกล้ 1 แสดงว่าระบบมีประสิทธิภาพดี จากข้อมูลที่ได้จะพบข้อสังเกต 5 ประเด็น ดังนี้

1) ค่า Distance พบว่าวิธีอัลกอริทึมคอนเดนเซชันเดิมนั้นให้ผลดีกว่าวิธีตัวกรองคาร์มาลฟิลเตอร์และวิธีที่เสนอ คือ อัลกอริทึมคอนเดนเซชันร่วมกับเอกลักษณ์สีและเทคนิคการสุ่มบนพื้นฐานพื้นที่

2) ค่า Absolute Error พบว่าวิธีที่เสนอให้ผลที่ดีกว่าวิธีที่ 1 และ 2 สันนิษฐานว่าเนื่องจากการใช้เอกลักษณ์สีร่วมในการพิจารณาซึ่งทำให้ได้พื้นที่ใกล้เคียงกับพื้นที่ของวัตถุมากขึ้น

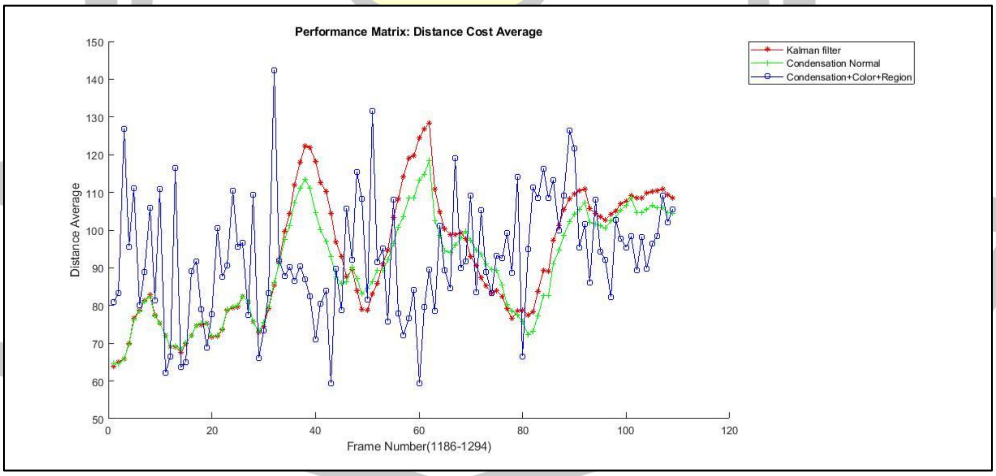
3) ค่า Bounding Box Error พบว่าวิธีที่เสนอให้ผลที่ดีกว่าวิธีที่ 1 และ 2 สันนิษฐานว่าการใช้เอกลักษณ์สีและการสุ่มบนพื้นฐานพื้นที่ช่วยให้ได้พื้นที่จริงของวัตถุมากขึ้นทำให้ค่า BBE เข้าใกล้ 1 มากกว่า 2 วิธีแรก

4) ค่า Object Tracking Error พบว่าวิธีที่ 1 ให้ผลที่ดีกว่าวิธีที่ 2 และวิธีที่เสนอ

5) ค่า Root Mean Square Error พบว่าวิธีที่เสนอให้ผลที่ดีกว่าวิธีที่ 1 และ 2

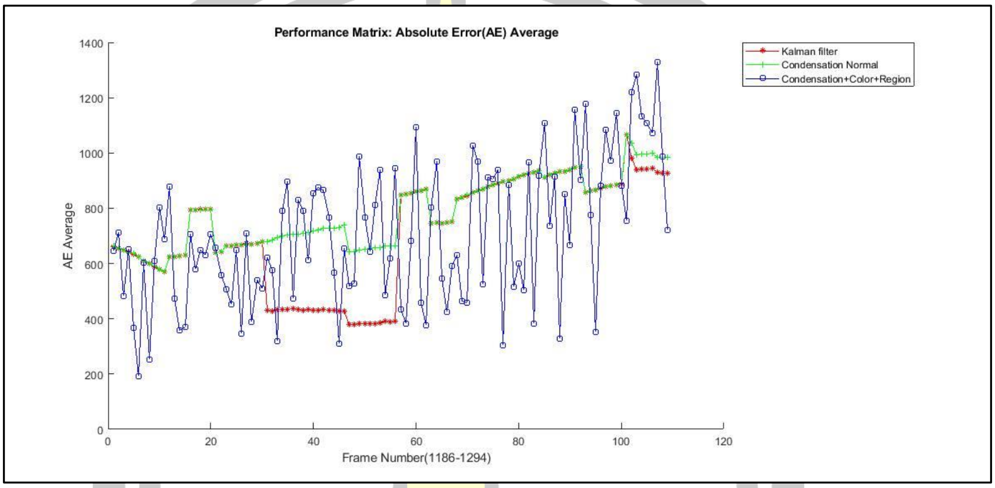
จากผลการทดลองสรุปได้ว่าการใช้เอกลักษณ์สีและการใช้วิธีการสุ่มบนพื้นฐานของพื้นที่ช่วยให้ค่า AE, BBE, OTE และ RMSE ดีกว่ากรณีใช้พื้นฐานของบราวดิ้งบล็อกแต่การใช้แซมเปิลบนพื้นฐานของพื้นที่อาจมีผลกระทบ คือ เนื่องจากแซมเปิลที่ได้นั้นอยู่บนพื้นที่จึงอาจทำให้ตำแหน่งที่ได้จากการประมาณอยู่ห่างจากศูนย์กลางบล็อกของกราวนด์ทรูธจึงทำให้ค่า Distance สูงกว่าแบบการประมาณบนพื้นฐานบราวดิ้งบล็อก

ตัวอย่างผลการวัดประสิทธิภาพโดยพิจารณาจากค่าเฉลี่ยในแต่ละเฟรมภาพของแต่ละพารามิเตอร์แสดงดังภาพที่ 5.9 - ภาพที่ 5.13 โดยเป็นผลการทดลองในช่วงเฟรมที่ 1186 ถึง 1294 ของกล้องตัวที่ 1



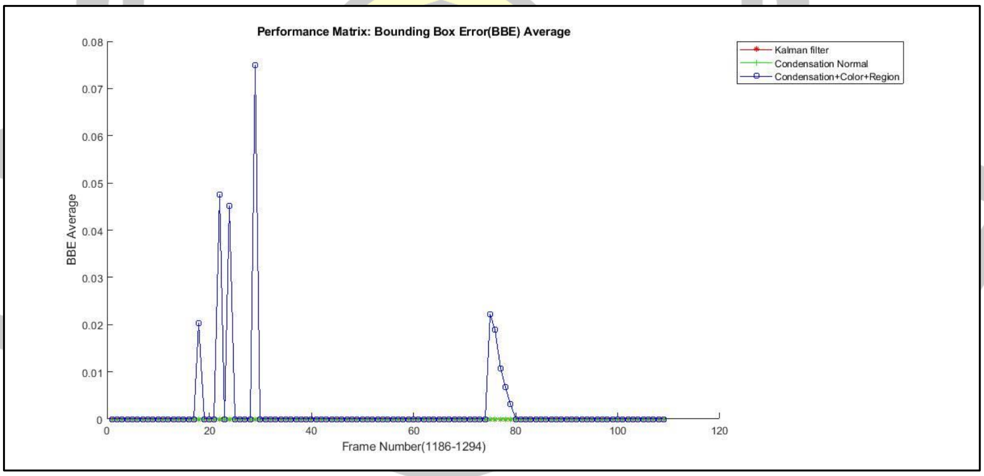
ภาพที่ 5.9 Distance Cost Average

ตามภาพที่ 5.9 เป็นการเปรียบเทียบผลการประเมินของค่า Distance Cost Average ระหว่าง 3 วิธี ซึ่งพบว่าวิธีที่เสนอให้ผลที่ดีกว่า 2 วิธีเดิมในบางช่วงของเฟรมวิดีโอ แต่ในบางช่วงของเฟรมวิดีโอให้ผลที่แย่กว่า



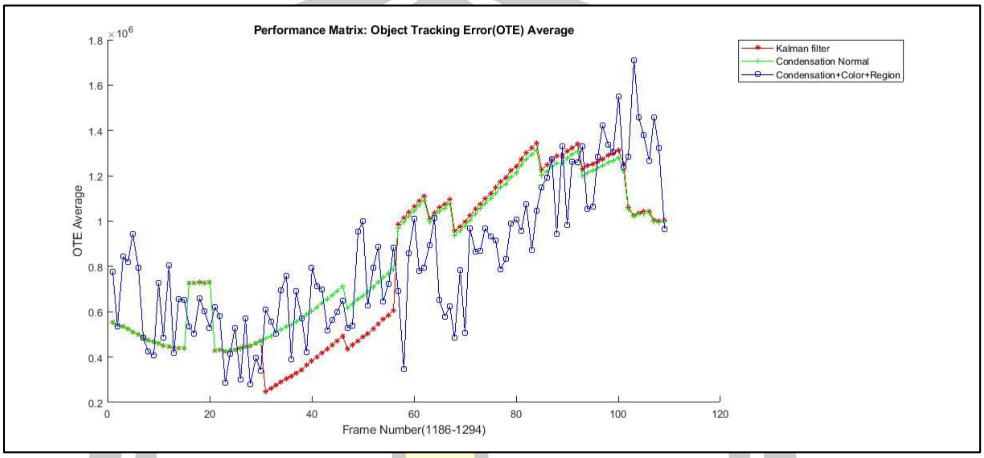
ภาพที่ 5.10 Absolute Error Average

ตามภาพที่ 5.10 เป็นการเปรียบเทียบผลการประเมินของค่า Absolute Error Average ระหว่าง 3 วิธี ซึ่งพบว่าวิธีที่เสนอให้ผลในภาพรวมที่ดีกว่า 2 วิธีเดิมในส่วนใหญ่ของช่วงเฟรมวิดีโอ



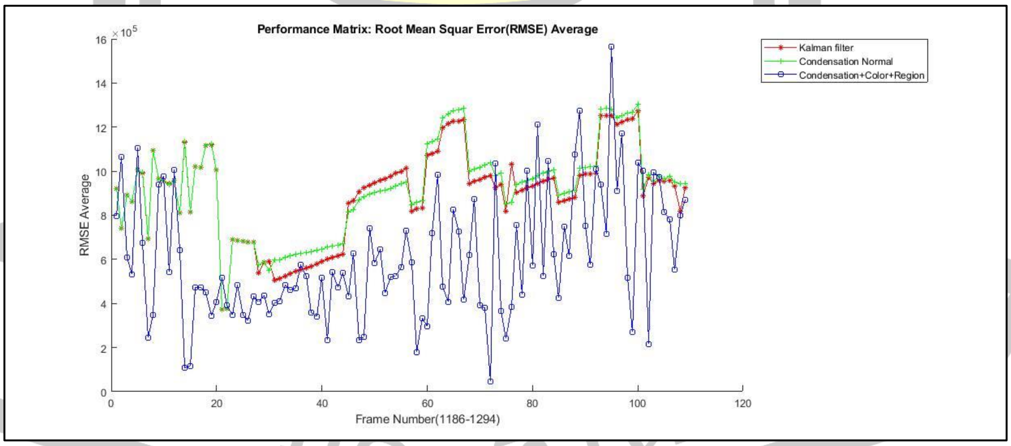
ภาพที่ 5.11 Bounding Box Error Average

ตามภาพที่ 5.11 เป็นการเปรียบเทียบผลการประเมินของค่า Bounding Box Error Average ระหว่าง 3 วิธี ซึ่งพบว่าวิธีที่เสนอให้ผลในบางช่วงของเฟรมวิดีโอที่ดีกว่า 2 วิธีเดิม



ภาพที่ 5.12 Object Tracking Error Average

ตามภาพที่ 5.12 เป็นการเปรียบเทียบผลการประเมินของค่า Object Tracking Error Average ระหว่าง 3 วิธี ซึ่งพบว่าวิธีที่เสนอให้ผลในบางช่วงของเฟรมวิดีโอที่ดีกว่า 2 วิธีเดิม แต่ยังมีบางช่วงของเฟรมวิดีโอที่ให้ผลที่แย่กว่าอีก 2 วิธี



ภาพที่ 5.13 Root Mean Square Error Average

ตามภาพที่ 5.13 เป็นการเปรียบเทียบผลการประเมินของค่า Root Mean Square Error Average ระหว่าง 3 วิธี ซึ่งพบว่าวิธีที่เสนอให้ผลในส่วนใหญ่ของเฟรมวิดีโอที่ดีกว่า 2 วิธีเดิม

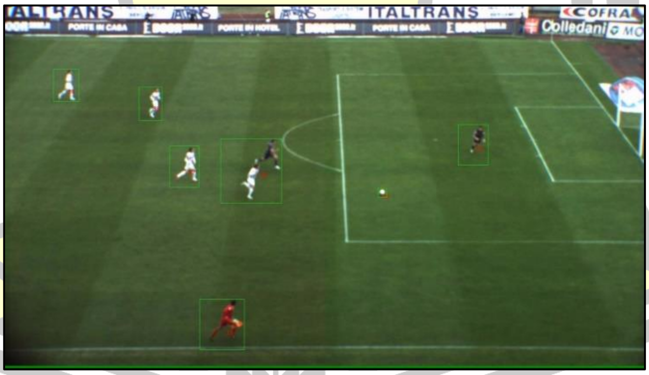
ตารางที่ 5.2 การประเมินประสิทธิภาพด้านเวลา

Average Time (second)			
Frame Total	Kalman Filter	Condensation Algorithm	Proposed Method
3003	15.914	6.8685	7.1833

ตามตารางที่ 5.2 พบว่าวิธีอัลกอริทึมคอนเดนเซชันแบบเดิมจะใช้เวลาในการประมวลผลน้อยที่สุด รองลงมา คือ เทคนิคที่เสนอ คือ การใช้อัลกอริทึมคอนเดนเซชันบนพื้นฐานของเอกลักษณ์ของสีร่วมกับเทคนิคการสุ่มแซมเปิลบนพื้นฐานของพื้นที่ และตัวกรองคาร์ลมาลฟิลเตอร์ตามลำดับ ข้อสังเกตที่พบ คือ วิธีอัลกอริทึมคอนเดนเซชันจะใช้เวลาน้อยกว่าตัวกรองคาร์ลมาลฟิลเตอร์และการใช้เอกลักษณ์สีส่งผลให้ใช้เวลาในการประมวลผลมากขึ้น

2. ตัวอย่างผลการติดตาม

ตัวอย่างผลการติดตามของวิธีตัวกรองคาร์ลมาลฟิลเตอร์ อัลกอริทึมคอนเดนเซชันและเทคนิคที่เสนอ คือ การใช้อัลกอริทึมคอนเดนเซชันบนพื้นฐานของเอกลักษณ์ของสีร่วมกับเทคนิคการสุ่มแซมเปิลบนพื้นฐานของพื้นที่ แสดงดังภาพที่ 5.14 - ภาพที่ 5.19 ดังต่อไปนี้



ภาพที่ 5.14 ตัวกรองคาร์ลมาลฟิลเตอร์: Frame 1186

พหุ่นํ ปณฺ ฑิโต ชีเว

ภาพที่ 5.15 อัลกอริทึมคอนเดนเซชัน: Frame 1186

ภาพที่ 5.16 เทคนิคที่เสนอ: Frame 1186

ภาพที่ 5.17 ตัวกรองคาร์ลมาลฟิลเตอร์: Frame 1263

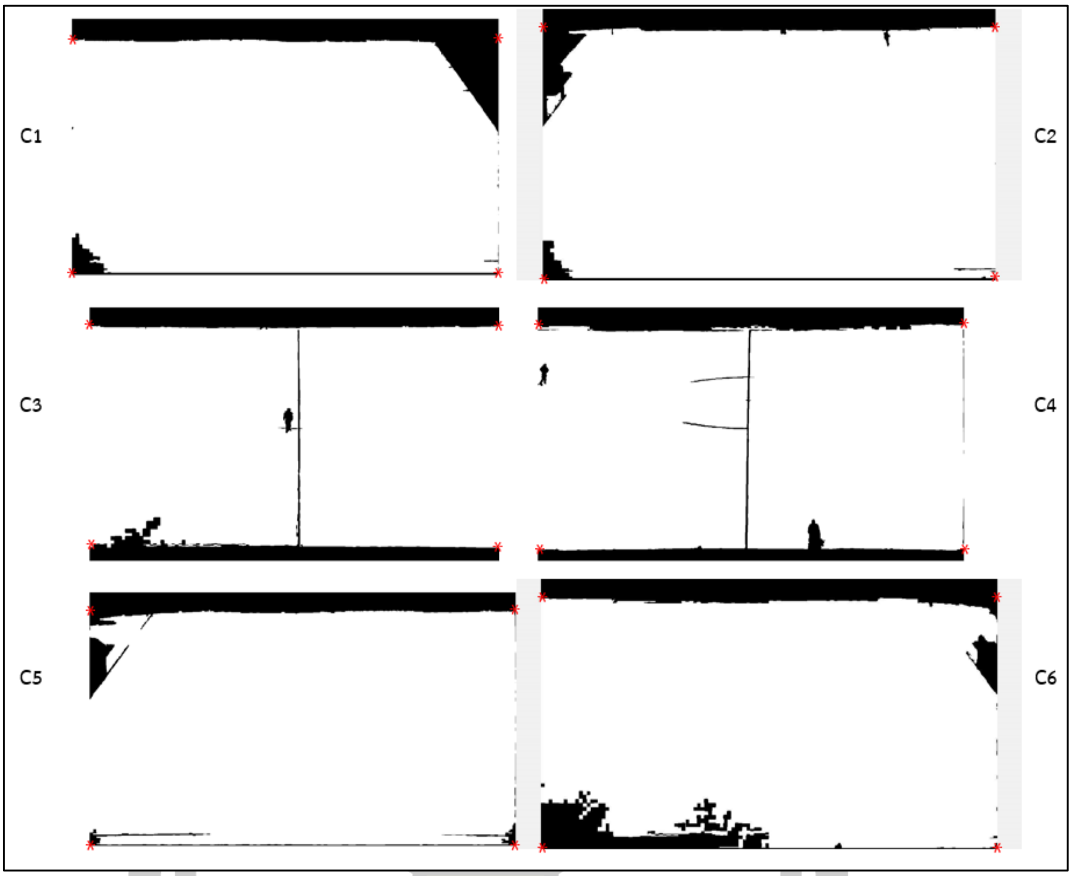
ภาพที่ 5.18 อัลกอริทึมคอนเดนเซชัน: Frame 1263

ภาพที่ 5.19 เทคนิคที่เสนอ: Frame 1263

3. ตัวอย่างผลการคำนวณขอบเขตสนาม

ตัวอย่างผลการคำนวณจุดมุมทั้ง 4 มุมรอบพื้นที่สนามแสดงดังภาพที่ 5.20 ซึ่งเป็นภาพเฟรมแรกจากกล้องทั้ง 6 ตัวของ ISSIA Dataset [82]

พหูนํ ปณฺฑิโต ชีเว



ภาพที่ 5.20 รายการจุดมุมทั้ง 4 มุมจากเฟรมแรกของ ISSIA Dataset [82]

ตามภาพที่ 5.20 เป็นตัวอย่างผลการคำนวณหาจุดมุมทั้ง 4 มุมรอบพื้นที่บริเวณของสนามโดยแสดงสัญลักษณ์ด้วยเครื่องหมายดอกจันสีแดง

5.3 สรุปผลการดำเนินงาน (Summary Operating Results)

ในกระบวนการวิจัยนี้ได้ประยุกต์ใช้เทคนิคการติดตามพื้นฐาน 2 ตัว คือ ตัวกรองคาร์ลมาลฟิลเตอร์และอัลกอริทึมคอนเดนเซชันสำหรับการศึกษาเพื่อหาวิธีการเพิ่มประสิทธิภาพในกระบวนการติดตาม ซึ่งในกระบวนการติดตามนี้ได้นำเสนอเทคนิคในการติดตามโดยใช้วิธีอัลกอริทึมคอนเดนเซชันบนพื้นฐานของสีร่วมกับการใช้เทคนิคการสุ่มแซมเปิลบนพื้นฐานของพื้นที่ เหตุผลที่ใช้วิธีอัลกอริทึมคอนเดนเซชันเนื่องจากผลการวัดประสิทธิภาพการติดตามในเบื้องต้นพบว่าไม่แตกต่างกันมากนัก นอกจากนี้อัลกอริทึมคอนเดนเซชันยังสามารถประยุกต์ใช้เพื่อแก้ปัญหาการบดบังบังได้ดีกว่าวิธีตัวกรองคาร์ลมาลฟิลเตอร์ สำหรับวิธีใหม่ที่เสนอนี้ผู้วิจัยใช้ข้อมูลเอกลักษณ์สีซึ่งช่วยให้การประมาณตำแหน่งมีประสิทธิภาพดีกว่าการใช้ข้อมูลตำแหน่งของแซมเปิลเพียงอย่างเดียว นอกจากนี้การใช้

วิธีการสุ่มแซมเปิลบนพื้นฐานของพื้นที่ยังช่วยให้การประมาณตำแหน่งวัตถุเข้าใกล้ตำแหน่งจริงของวัตถุมากขึ้นกว่ากรณีการใช้วิธีการสุ่มแซมเปิลบนพื้นฐานของบราวดิ้งบล็อก สาเหตุเนื่องมาจากภายในบราวดิ้งบล็อกจะได้แซมเปิลที่เป็นพื้นหลังผสมมาด้วยทำให้การประมาณตำแหน่งอาจคลาดเคลื่อนได้ ซึ่งสอดคล้องกับงานวิจัยของ Belagiannis และคณะ [104] ที่ใช้วิธีการสุ่มแซมเปิลบนพื้นฐานของการเซกเมนท์เตชัน ซึ่งให้เหตุผลที่คล้ายกันว่าการใช้บราวดิ้งบล็อกเป็นพื้นฐานในการสุ่มจะเกิดปัญหาโดยจะได้แซมเปิลเป็นพื้นหลังปะปนมาด้วยซึ่งจะส่งผลต่อกระบวนการติดตามได้

จากการทดลองพบว่าวิธีที่เสนอโดยกรณีใช้เอกลักษณ์สีและวิธีการสุ่มแซมเปิลบนพื้นฐานของพื้นที่ให้ผล AE, BBE, OTE และ RMSE ได้ดีกว่ากรณีไม่ใช้เอกลักษณ์สีและกรณีการสุ่มแซมเปิลบนพื้นฐานของบราวดิ้งบล็อก สำหรับผลการวัดประสิทธิภาพด้านเวลาพบว่าวิธีอัลกอริทึมคอนเดนเซชันแบบเดิมใช้เวลาน้อยที่สุด รองลงมา คือ วิธีอัลกอริทึมคอนเดนเซชันบนพื้นฐานของสีร่วมกับการสุ่มแซมเปิลบนพื้นฐานของพื้นที่และวิธีตัวกรองคาร์ลมาลฟิลเตอร์ ตามลำดับ โดยที่ 2 วิธีแรกนั้นใช้เวลาไม่แตกต่างกันมากนักแต่ในส่วนของวิธีสุดท้ายจะใช้เวลาค่อนข้างมากและมีค่าที่ห่างจาก 2 วิธีแรกค่อนข้างชัดเจน

บทที่ 6

การติดตามผู้เล่นบนมุมมองเสมือน (Player Tracking on Virtual View)

ในงานนี้มีวัตถุประสงค์เพื่อการติดตามผู้เล่นบนมุมมองภาพเสมือนเพื่อการแสดงผลในภาพรวมของเกมส์การแข่งขันรวมถึงการนำข้อมูลที่ได้จากการติดตามไปใช้ประกอบการแก้ปัญหาการบดบังซึ่งเกิดจากเหตุการณ์การรวมกลุ่มระหว่างผู้เล่น โดยมีรายละเอียดการดำเนินงานดังนี้

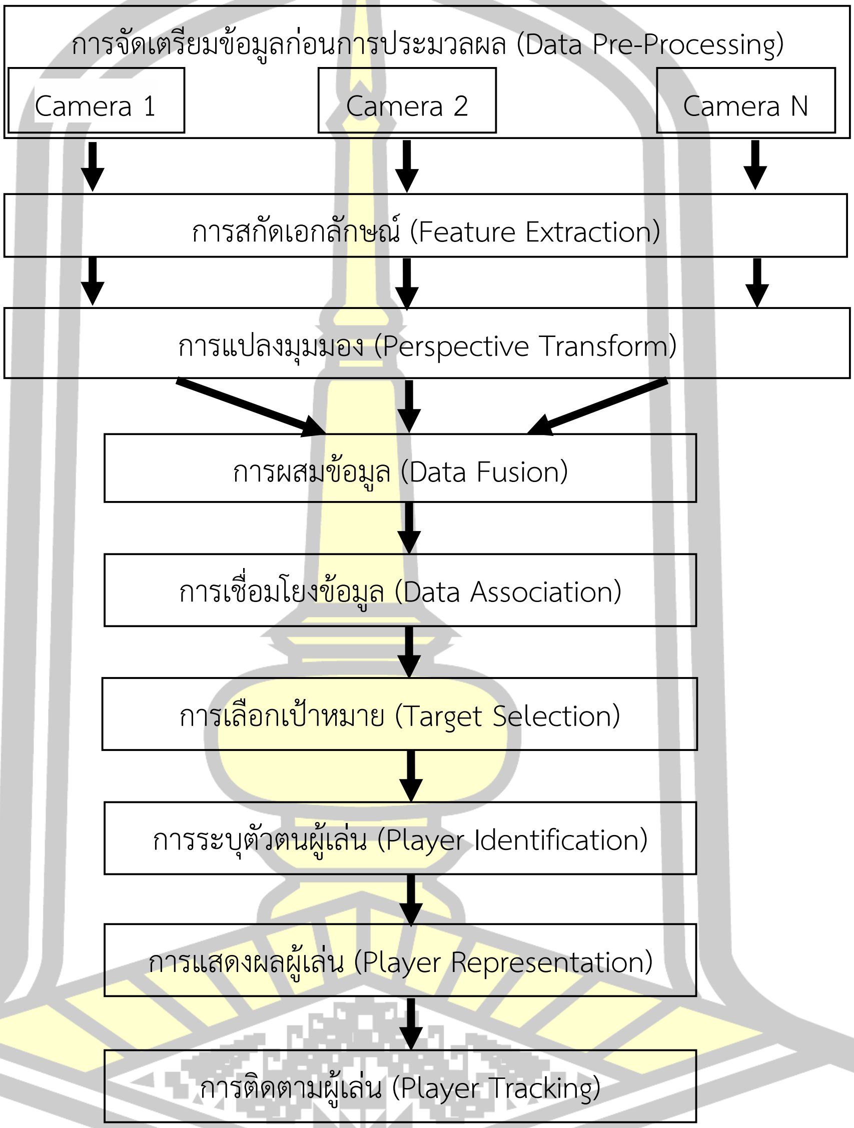
6.1 วิธีดำเนินการวิจัย (Research Methodology)

งานในส่วนนี้ประกอบด้วย 2 ส่วนหลัก คือ ส่วนแรก เป็นการนำเสนอเทคนิคการติดตามผู้เล่นบนมุมมองเสมือน ประกอบด้วย 9 ขั้นตอนหลัก แสดงดังภาพที่ 6.1 ได้แก่ 1. ขั้นตอนการจัดเตรียมข้อมูลก่อนการประมวลผล (Data Pre-Processing) 2. ขั้นตอนการสกัดเอกลักษณ์ (Feature Extraction) 3. ขั้นตอนการแปลงมุมมอง (Perspective Transform) 4. ขั้นตอนการผสมข้อมูล (Data Fusion) 5. ขั้นตอนการเชื่อมโยงข้อมูล (Data Association) 6. ขั้นตอนการเลือกเป้าหมาย (Target Selection) 7. ขั้นตอนการระบุตัวตนผู้เล่น (Player Identification) 8. ขั้นตอนการแสดงผลผู้เล่น (Player Representation) และ 9. ขั้นตอนการติดตามผู้เล่น (Player Tracking) และในส่วนที่สอง เป็นแนวทางการแก้ไขปัญหาการบดบังระหว่างผู้เล่น โดยมีรายละเอียดดังนี้

1. เทคนิคการติดตามผู้เล่นบนมุมมองเสมือน (Player Tracking Technique on Virtual View)

ในขั้นตอนแรก เป็นการจัดเตรียมสภาพแวดล้อมสำหรับการทดลอง เช่น การกำหนดแผนผังของโมเดลสนาม แผนผังตำแหน่งของกล้อง การวางตำแหน่งโมเดลตุ๊กตาจำลอง เป็นต้น ในขั้นตอนที่สอง รายการบล็อบวัตถุที่ได้มาจากกระบวนการตรวจจับวัตถุจะถูกนำมาสกัดเอกลักษณ์ โดยมี 3 เอกลักษณ์หลัก ได้แก่ สี ตำแหน่งและความกว้างของบล็อบวัตถุ ในขั้นตอนการแปลงมุมมอง จะทำการแปลงตำแหน่งของผู้เล่นที่ได้จากมุมมองภาพของแต่ละกล้องไปยังมุมมองเสมือน จากนั้นทำการผสมข้อมูลที่ได้จากกล้องทั้งหมดเป็นข้อมูลเดียวกัน จากนั้นทำการเชื่อมโยงเป้าหมายจากทุกมุมมองเข้าด้วยกันเพื่อใช้สำหรับการติดตาม การจำแนกตัวตน การแก้ปัญหาการบดบัง เป็นต้น เสร็จแล้วทำการเลือกเป้าหมายที่ต้องการติดตามตามจำนวนที่กำหนด จากนั้นทำการระบุตัวตนด้วย ID เพื่อใช้ประกอบการติดตามและแสดงผล ขั้นตอนต่อมา คือ การแสดงผลตำแหน่งและ ID ของผู้เล่นที่ได้จากขั้นตอนการระบุตัวตนลงบนภาพในมุมมองเสมือน ในขั้นตอนสุดท้าย เป็นการติดตามการเคลื่อนที่ของผู้เล่นในมุมมองเสมือนซึ่งในงานนี้ใช้เทคนิคการติดตามพื้นฐาน คือ ตัวกรองคาร์ลมาลฟิลเตอร์เนื่องจากข้อมูลที่ใช้สำหรับการติดตามมีเพียงตำแหน่งของผู้เล่นเท่านั้นจึงไม่จำเป็นต้องใช้

เทคนิคการติดตามที่ซับซ้อน รายละเอียดของการดำเนินงานในแต่ละขั้นตอนประกอบด้วย 9 หัวข้อ ดังต่อไปนี้



ภาพที่ 6.1 กระบวนการติดตามผู้เล่นบนมุมมองเสมือน

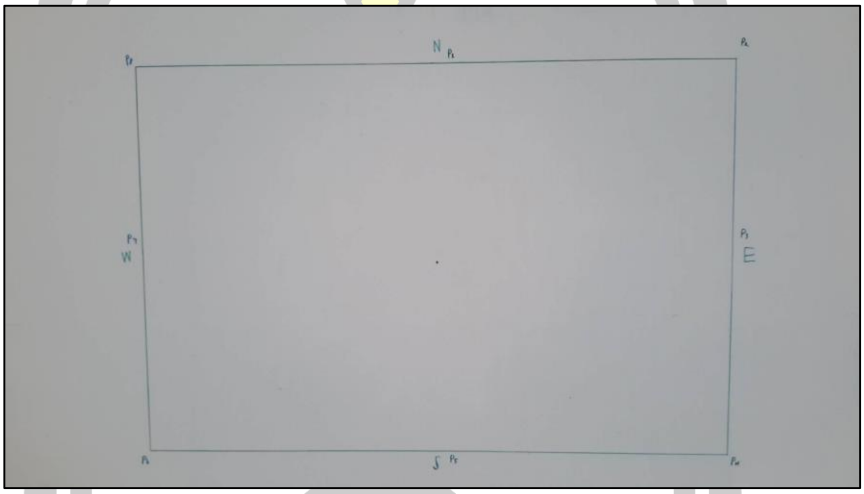
1.1 ขั้นตอนการจัดเตรียมข้อมูลก่อนการประมวลผล (Data Pre-Processing)

กระบวนการทดลองการแปลงมุมมองโดยใช้ตำแหน่งต่างๆ ของผู้เล่น เช่น กรณีการใช้ตำแหน่งศีรษะ ตำแหน่งศูนย์กลางบล็อกและตำแหน่งเท้าของผู้เล่น รวมถึงกระบวนการผสมข้อมูล

ในงานนี้ทำการจำลองพื้นที่สนามขึ้นมาและใช้โมเดลตุ๊กตาเป็นตัวแทนของผู้เล่น จุดประสงค์เพื่อให้เห็นภาพรวมของตำแหน่งวัตถุและพื้นที่สนามทั้งหมด โดยมีรายละเอียดดังนี้

1) การจำลองโมเดลสนาม

โมเดลสนามที่จำลองขึ้นมานี้จะใช้วิธีการวาดเส้นขอบสนามด้วยมือบนกระดานที่มีพื้นสีขาวแสดงดังภาพที่ 6.2

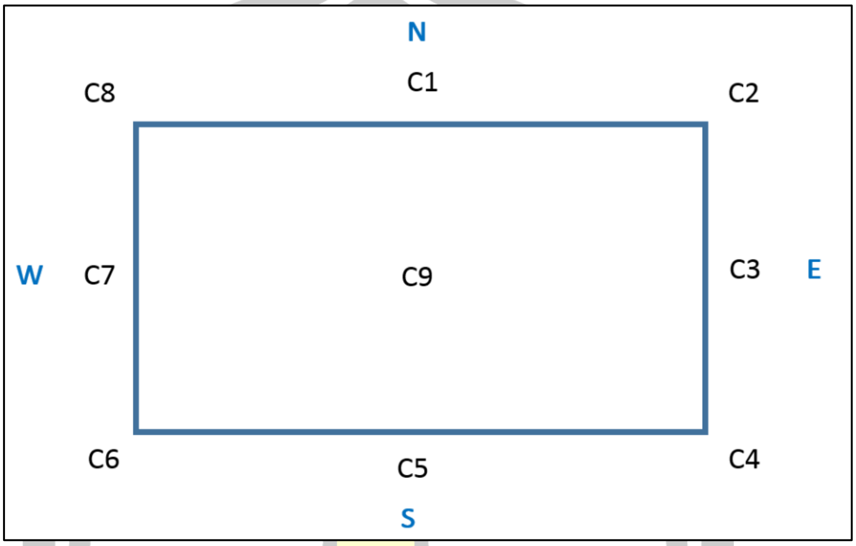


ภาพที่ 6.2 การจำลองพื้นที่สนาม

2) การจำลองตำแหน่งของกล้อง

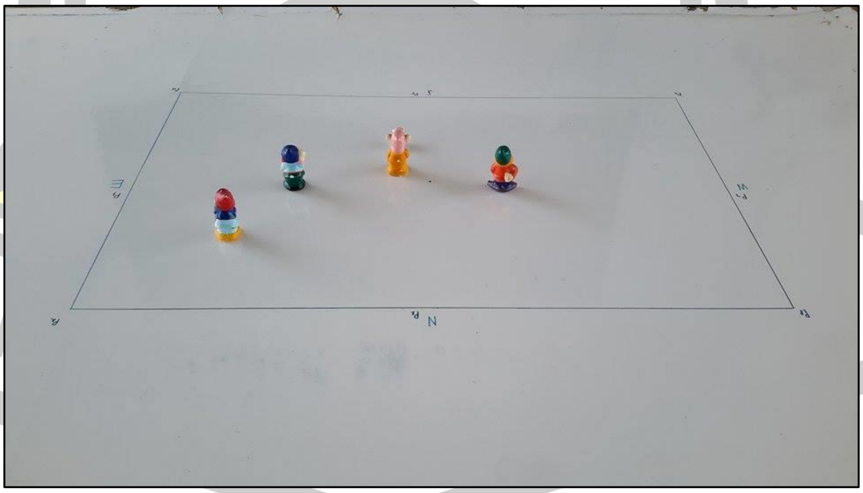
การจำลองตำแหน่งของกล้อง ลักษณะของกล้องที่ใช้ในการทดลองนี้มีทั้งหมด 9 ตัว โดยมีมุมมองตามทิศต่างๆ ทั้ง 8 ทิศและอยู่ในตำแหน่งเหนือศีรษะ (Top View) อีก 1 ตัว โดยตำแหน่งต่างๆ ของกล้องจะแสดงดังภาพที่ 6.3 จุดประสงค์ที่ใช้กล้องจำนวน 9 ตัว เพื่อ 2 เหตุผลดังนี้ คือ 1. เพื่อใช้กล้องในตำแหน่งเหนือศีรษะสำหรับเป็นเกณฑ์ในการวัดประสิทธิภาพของระบบสำหรับกล้องที่เหลือทั้ง 8 ตัว เนื่องจากตำแหน่งวัตถุในมุมมองเหนือศีรษะจะเป็นตำแหน่งที่ตรงกับตำแหน่งจริงในมุมมองโลกในทางกายภาพและ 2. การใช้กล้อง 8 ตัวตามทิศต่างๆ ทั้ง 8 ทิศ เพื่อทดสอบสมมติฐานของคณะผู้วิจัยที่สันนิษฐานไว้ว่า สมมติฐานแรก คือ ตำแหน่งของวัตถุที่ได้จากกล้องในทิศทางต่างๆ นั้นจะเป็นตำแหน่งเดียวกันเสมอในโลกความเป็นจริงดังนั้นเมื่อโปรเจ็คตำแหน่งจากกล้องต่างๆ เหล่านี้ลงบนภาพในมุมมองเสมือนแล้วก็ควรที่จะได้ตำแหน่งเดียวกันเสมอ โดยในที่นี้ไม่พิจารณาปัญหาที่เกิดจากกระบวนการตรวจจับวัตถุ สมมติฐานที่สอง คือ แม้กล้องจะมีมุมมองต่างกันและวิธีการกำหนดตำแหน่งของวัตถุที่แตกต่างกันซึ่งจะได้ผลการประมาณตำแหน่งที่แตกต่างกันนั้นแต่

เราสามารถใช้หลักการของสมการเส้นตรงทางคณิตศาสตร์บนพื้นฐานทางกายภาพในการแก้ปัญหาได้ โดยมีแนวคิดอยู่บนสมมติฐานแรก เป็นต้น

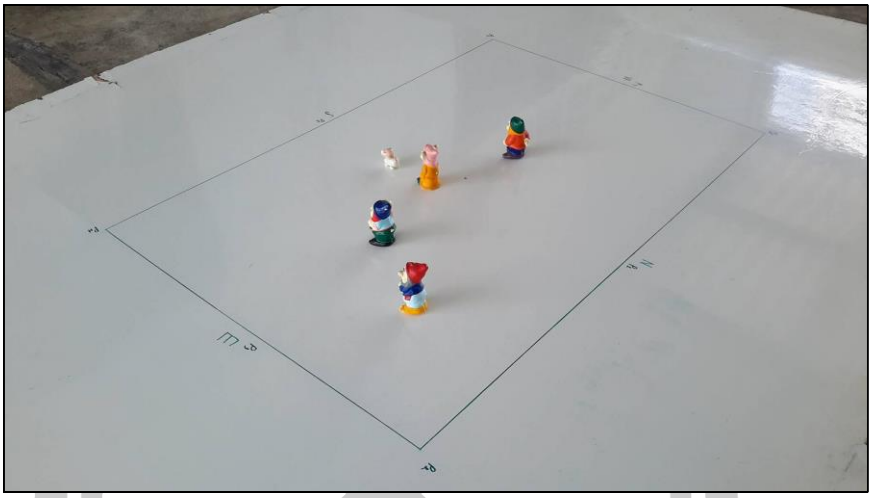


ภาพที่ 6.3 แผนผังตำแหน่งกล้องทั้ง 9 ตัว

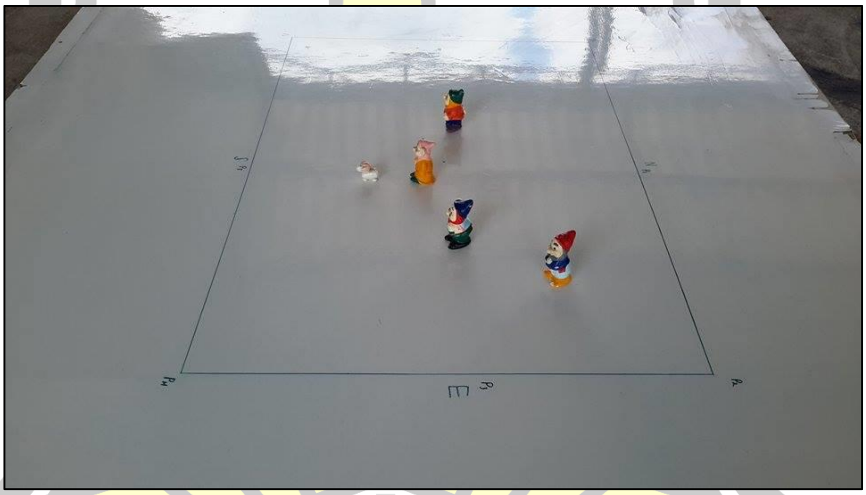
ตัวอย่างของข้อมูลภาพที่ได้จากกล้องในแต่ละตัวในแต่ละทิศทางทั้ง 9 ทิศทาง แสดงดังภาพที่ 6.4 - ภาพที่ 6.12 ดังนี้



ภาพที่ 6.4 กล้อง C1

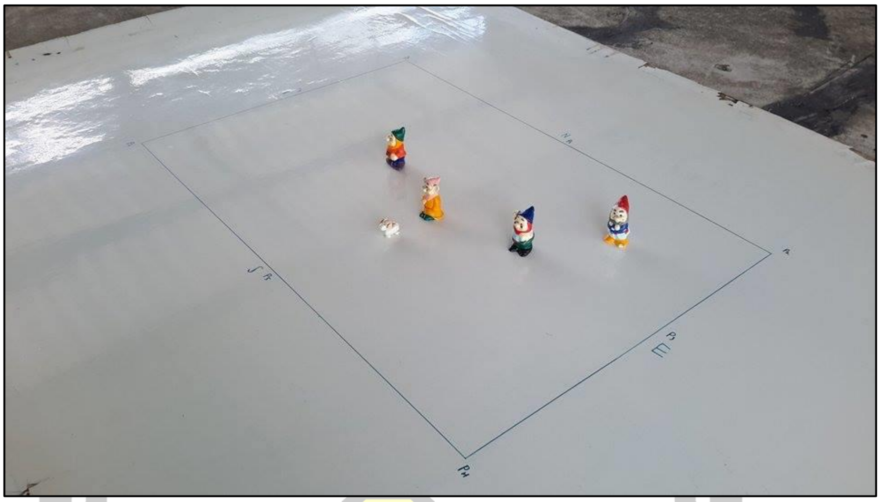


ภาพที่ 6.5 กล้อง C2

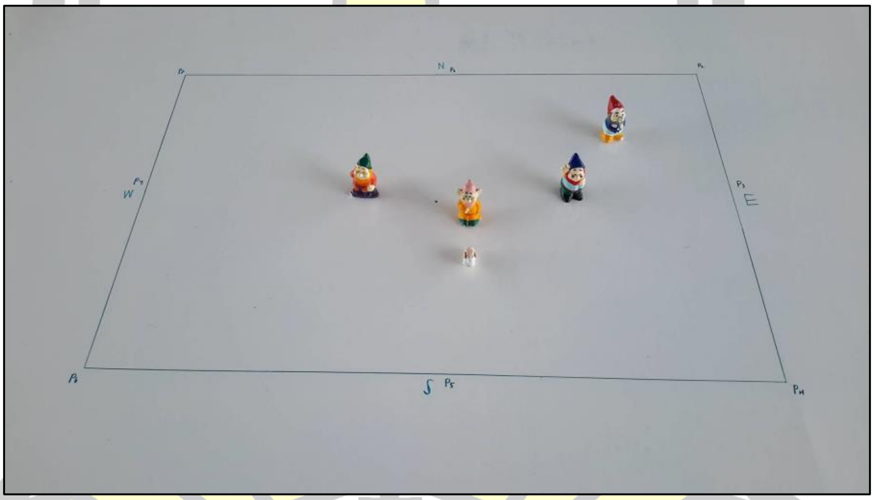


ภาพที่ 6.6 กล้อง C3

พหูนํ ปณฺฑิโต ชีเว

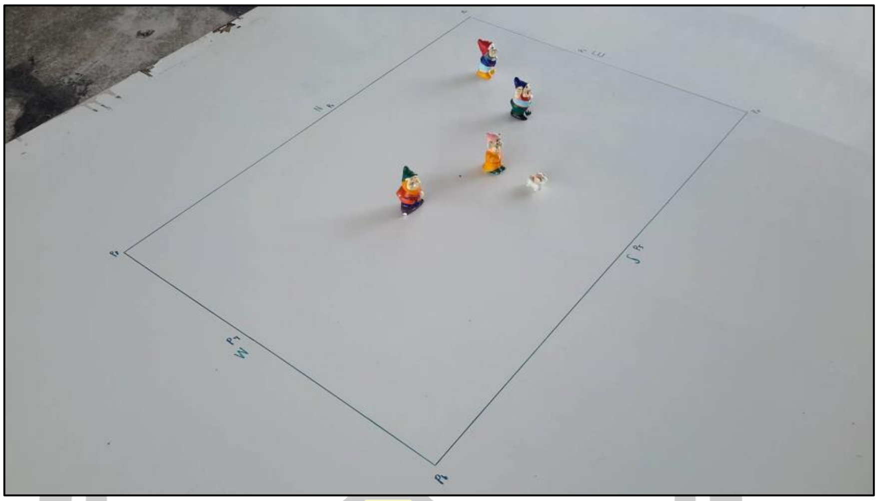


ภาพที่ 6.7 กล้อง C4

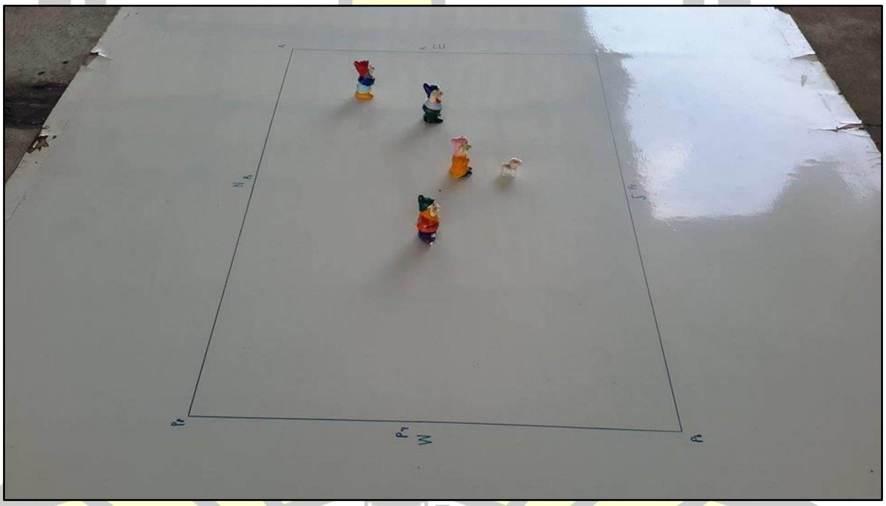


ภาพที่ 6.8 กล้อง C5

พหูน ปณฺ ฑิโต ชีเว

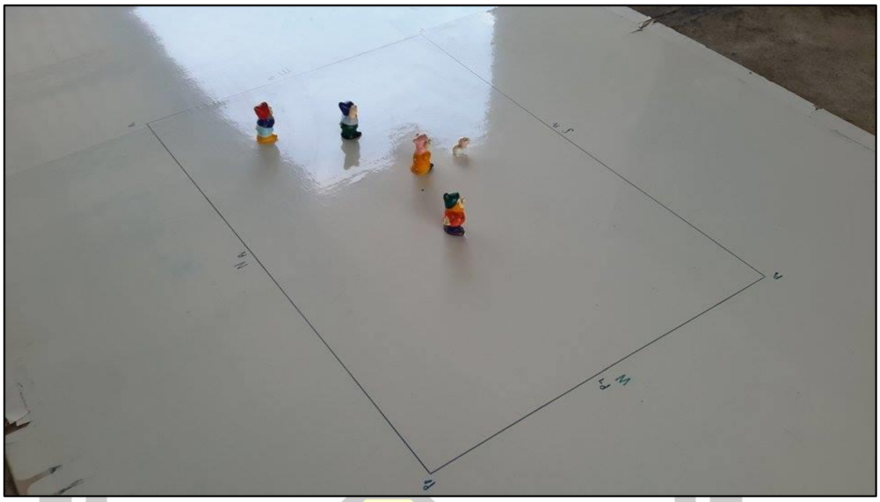


ภาพที่ 6.9 กล้อง C6

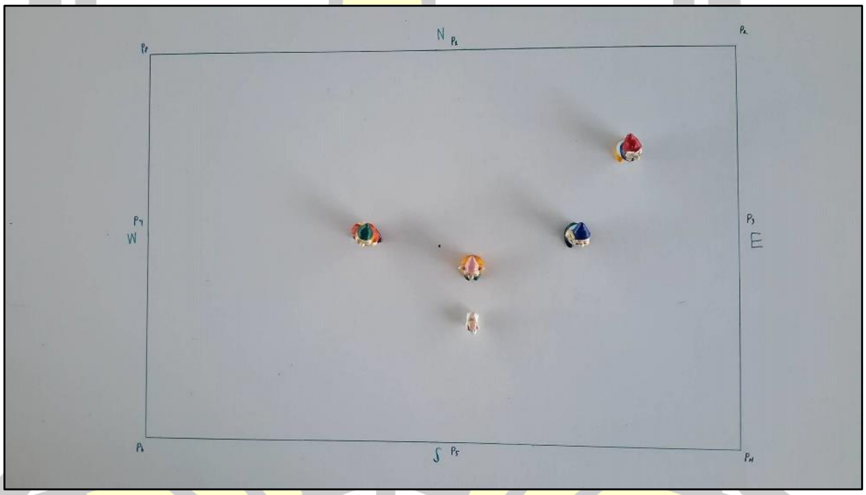


ภาพที่ 6.10 กล้อง C7

พหูน ปณฺ ฑิโต ชีเว



ภาพที่ 6.11 กล้อง C8



ภาพที่ 6.12 กล้อง C9 (Top View)

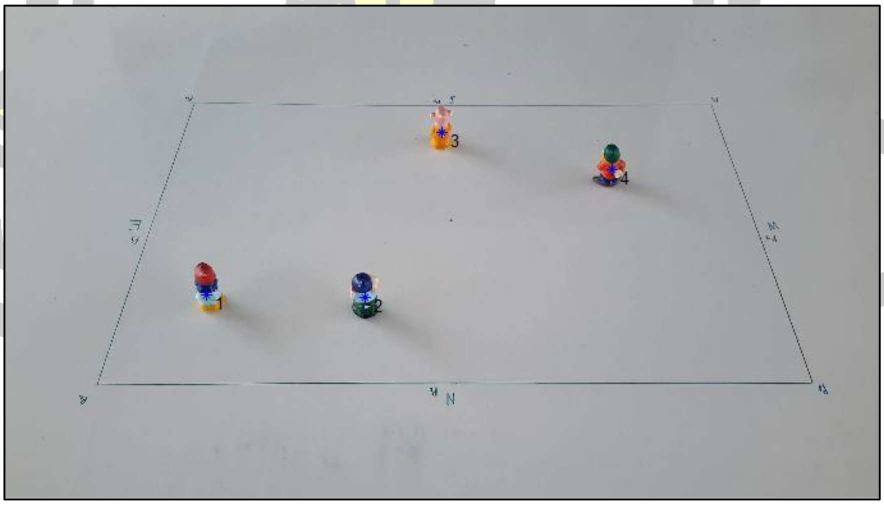
3) การกำหนดค่าพื้นฐานของกล้องและวิธีการบันทึกภาพ

อุปกรณ์ที่ใช้บันทึกภาพในการทดลองนี้ คือ Tablet ยี่ห้อ Galaxy Tab S ระบบปฏิบัติการ Android 6.0.1 จำนวน 1 เครื่อง ลักษณะภาพที่ได้ คือ ภาพประเภท JPEG ค่า Resolution ในแนวตั้งและแนวนอน คือ 96 dpi ค่า Bit depth คือ 24 สำหรับขนาดความกว้างและความสูงของภาพในแต่ละมุมกล้องจะแตกต่างกันโดยจะประกอบด้วย ขนาดความกว้าง 960 พิกเซล ขนาดความสูง 540 พิกเซล และขนาดความกว้าง 2048 พิกเซล ขนาดความสูง 1152 พิกเซล เป็นต้น

การกำหนดค่าตำแหน่งพื้นฐานทางกายภาพของกล้อง ในงานนี้ไม่มีการกำหนดตำแหน่งการติดตั้งกล้องว่ามีระยะห่างจากขอบสนามเท่าไร ไม่มีการกำหนดระยะความสูงของกล้องว่าสูงจากพื้นเท่าไรและไม่มีการกำหนดมุมการเอียงของกล้องในการบันทึกภาพ จุดประสงค์เพื่อการทดลองหาคำตอบของสมมติฐาน คือ ระบบจะสามารถประมาณตำแหน่งของผู้เล่นจากกล้องที่อยู่ในทิศทางหรือมุมมองที่แตกต่างกันลงบนมุมมองภาพแบบเสมือนได้โดยไม่ต้องอาศัยข้อมูลพื้นฐานของกล้องได้หรือไม่ สืบเนื่องจากในหลายงานวิจัยโดยส่วนใหญ่ในกระบวนการของการแปลงมุมมองระหว่างมุมมองภาพและมุมมองเสมือนจะต้องกำหนดข้อมูลพื้นฐานของกล้อง เช่น ตำแหน่งกล้อง องศาการเอียงของกล้องและการบิดเบือนของเลนส์ เป็นต้น และในหลายงานวิจัยมักจะต้องทำการสอบเทียบกล้องก่อน ซึ่งคณะผู้วิจัยเห็นว่ากระบวนการสอบเทียบกล้องและการคำนวณหาค่าพื้นฐานของกล้องเหล่านี้เป็นส่วนที่เพิ่มความยุ่งยากในกระบวนการแปลงมุมมองโดยเฉพาะกับระบบที่ต้องทำงานแบบเรียลไทม์

การบันทึกภาพดำเนินการโดยผู้วิจัย โดยวิธีการบันทึกภาพ การประมาณระยะห่าง ความสูงและมุมเอียงของกล้องจะดำเนินการด้วยมือโดยเน้นให้ภาพที่ได้ในแต่ละกล้องหรือในแต่ละทิศทางนั้นครอบคลุมพื้นที่สนามทั้งหมด

การทดสอบในที่นี้ใช้ชุดทดสอบที่ดำเนินการด้วยมือ โดยจำลองตำแหน่งของข้อมูลบนภาพจำลอง ซึ่งใช้ข้อมูลภาพจากกล้องทั้งหมด 9 ตัว ข้อมูลตำแหน่งที่ใช้ทดสอบจำนวน 4 ตำแหน่ง จากนั้นดำเนินการกำหนดค่าต่างๆ เช่น ตำแหน่งจุดควบคุม (Control Point) สำหรับการแปลงมุมมองระหว่างมุมมองภาพและมุมมองเสมือนด้วยมือ ตัวอย่างภาพในชุดทดลองแสดงดังภาพที่ 6.13 ซึ่งเป็นภาพที่ได้จากกล้อง C1 จากทั้งหมด 9 ตัว



ภาพที่ 6.13 ตัวอย่างภาพทดสอบ (กล้อง C1)

1.2 ขั้นตอนการสกัดเอกลักษณ์ (Feature Extraction)

ในขั้นตอนนี้แต่ละบล็อบของวัตถุที่ได้มาจากกระบวนการตรวจจับผู้เล่นดังที่อธิบายไว้ในบทที่ 5 จะถูกนำมาสกัดเอกลักษณ์พื้นฐาน ได้แก่ ตำแหน่งและความกว้างของบล็อบดังสมการที่ 6.1 โดยตำแหน่งของบล็อบหรือผู้เล่นนั้นสามารถดึงข้อมูลนี้มาจากเอาท์พุตของกระบวนการประมาณตำแหน่งผู้เล่น ซึ่งได้อธิบายไว้ในบทที่ 5 ส่วนความกว้างของบล็อบนั้นสามารถคำนวณได้จากความกว้างในแนวนอนของบล็อบหารด้วย 2 ซึ่งจะได้ขนาดความกว้างของลำตัวผู้เล่นด้านหนึ่งมาดังสมการที่ 6.2 โดยในงานนี้มีหลักการพิจารณาว่า ตำแหน่งของผู้เล่นในขณะที่ยืนอยู่นั้น คือ ตำแหน่งศูนย์กลางลำตัวของผู้เล่น ส่วนความกว้างของผู้เล่นจะพิจารณาเพียงครึ่งหนึ่งของความกว้างลำตัวทั้งหมดแสดงดังภาพที่ 6.14 เนื่องจากเวลาที่คน 2 คนยืนชิดกันหัวไหล่ของทั้ง 2 คนจะชนกันพอดี ซึ่งในงานนี้จะพิจารณาว่าเข้าเกณฑ์ของเหตุการณ์การรวมกลุ่ม สำหรับขนาดความกว้างที่ได้จากความกว้างของวัตถุทั้งหมดที่ตรวจจับได้ในเฟรมเริ่มต้น จะทำการหาค่าเฉลี่ย ความกว้างเฉลี่ยนี้จะถูกใช้สำหรับกำหนดเป็นค่ารัศมีของวงกลม ซึ่งจะถูกแสดงบนภาพในมุมมองเสมือน โดยในที่นี้กำหนดสัญลักษณ์เป็น *r* เพื่อใช้สำหรับการแสดงขนาดความกว้างลำตัวของผู้เล่นและใช้สำหรับกระบวนการวิเคราะห์เหตุการณ์การรวมกลุ่มด้วย

ภาพที่ 6.14 ความกว้างของลำตัวผู้เล่น

$$Feature_X = \{position(x.y), width_X\} \tag{6.1}$$

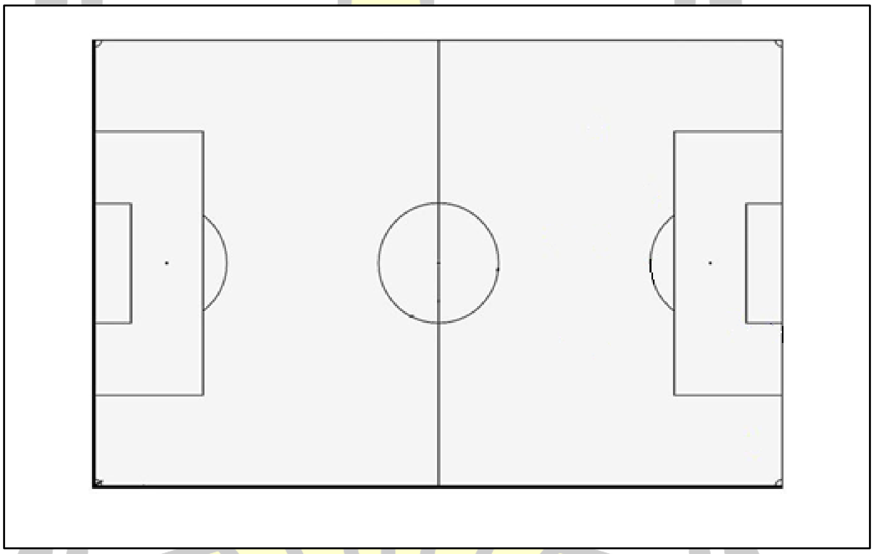
$$width_X = mean\left(\frac{blob_1^N(size)}{2}\right) \tag{6.2}$$

1.3 ขั้นตอนการแปลงมุมมอง (Perspective Transform)

หลังจากขั้นตอนการสกัดเอกลักษณ์เสร็จสิ้น เอกลักษณ์ตำแหน่งของผู้เล่นที่ได้จากแต่ละกล้องจะถูกนำมาโปรเจ็คลงบนภาพในมุมมองเสมือนโดยใช้หลักการพื้นฐานของการแปลงมุมมอง โดยมีรายละเอียดดังนี้

1) การกำหนดภาพโมเดลสนามในมุมมองเสมือน (Virtual Playfield Image Model)

การแสดงผลตำแหน่งผู้เล่นในมุมมองเสมือนในกระบวนการสุดท้ายของงานนี้จะใช้ภาพของสนามฟุตบอลที่แสดงผลในลักษณะมุมมองเหนือศีรษะ ตัวอย่างแสดงดังภาพที่ 6.15 โดยเป็นภาพประเภท Bit Map ขนาดความกว้าง 516 พิเซลและขนาดความสูง 323 พิกเซล ค่า Bit Depth คือ 24 ซึ่งภาพจะมีโครงสร้างคล้ายกับสนามจริง



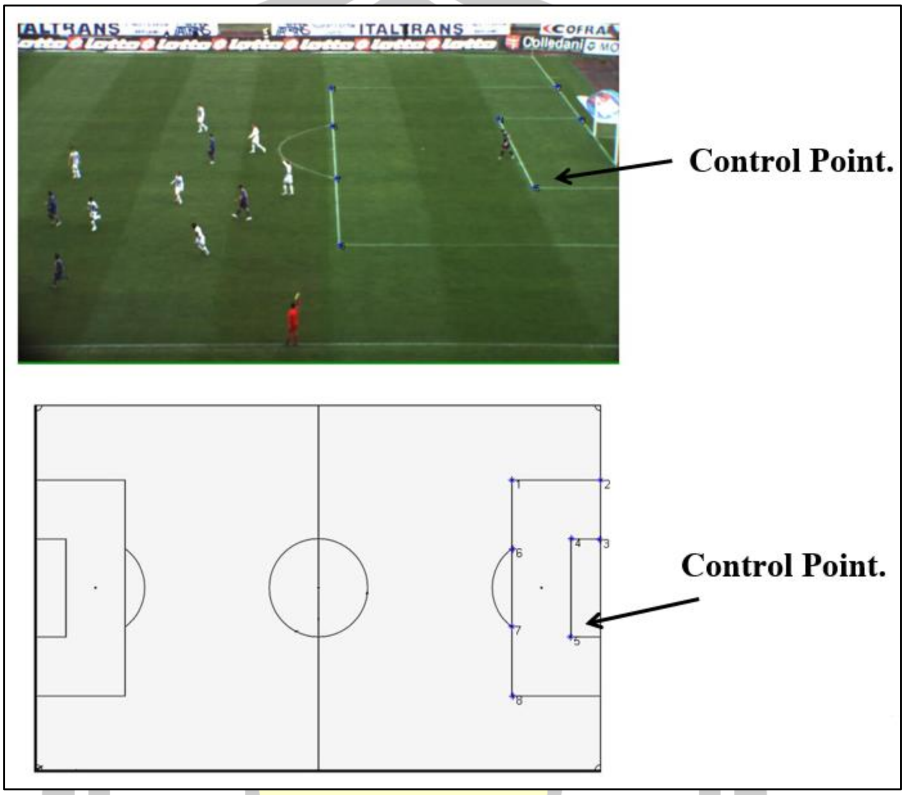
ภาพที่ 6.15 ตัวอย่างภาพมุมมองเสมือน (Virtual Plane/ Top View)

สำหรับภาพมุมมองเสมือนที่ใช้ในการทดสอบในงานนี้จะใช้ภาพจำลองโมเดลสนามที่สร้างด้วยมือแสดงดังภาพที่ 6.2 จุดประสงค์เพื่อให้สามารถมองเห็นข้อมูลตำแหน่งผู้เล่นได้ในภาพรวมของสนามทั้งหมด

2) การกำหนดจุดควบคุม (Control Point)

การกำหนดจุดควบคุมหรือที่เรียกว่า คอนโทรลพอยต์ (Control Point) ในการทดลองของงานนี้ดำเนินการกำหนดจุดคอนโทรลพอยต์ด้วยมือ โดยเลือกบริเวณตำแหน่งมุมของเส้นสนามที่เป็นจุดสำคัญจำนวนทั้งหมด 4 - 8 จุด ขึ้นอยู่กับลักษณะของภาพที่ได้ในแต่ละกล้อง ตัวอย่าง

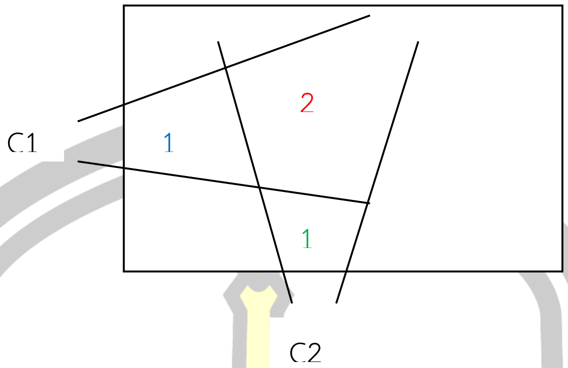
การกำหนดจุดคอนโทรลพอยต์ระหว่างภาพมุมมองเสมือนกับภาพจากกล้องตัวหนึ่งจากทั้งหมด 6 ตัว ของชุดวิดีโอทดสอบจาก ISSIA Dataset [82] แสดงดังภาพที่ 6.16



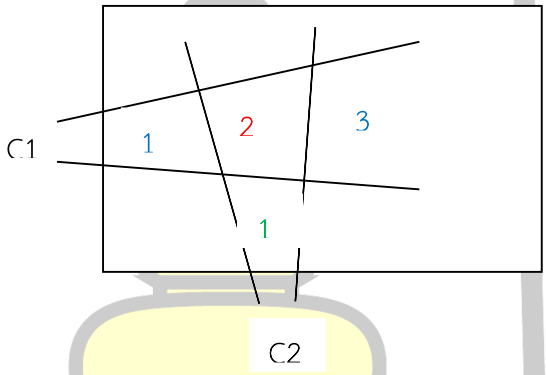
ภาพที่ 6.16 ตัวอย่างการกำหนดจุดคอนโทรลพอยต์

3) การคำนวณ Field of View (FOV) (FOV Calculation)

FOV คือ มุมมองหรือขอบเขตของพื้นที่บนสนามที่กล้องสามารถจับภาพได้ซึ่งจะแสดงขอบเขตด้วยเส้นตรง 2 เส้น ดังภาพที่ 6.17 และภาพที่ 6.18 บนภาพเสมือน (Virtual View) ขนาดพื้นที่ของ FOV ในแต่ละกล้องจะแตกต่างกันขึ้นอยู่กับปัจจัยต่างๆ เช่น ระยะทางของกล้องจากสนาม ความสูงและองศาของกล้อง เป็นต้น ดังนั้นกล้องจะสามารถจับภาพวัตถุได้ก็ต่อเมื่อวัตถุนั้นผ่านเข้ามาในบริเวณ FOV ของตนเอง



ภาพที่ 6.17 ตัวอย่าง FOV ประเภท 1 และ 2



ภาพที่ 6.18 ตัวอย่าง FOV ประเภท 1, 2 และ 3

เกณฑ์การพิจารณา FOV จะมี 3 ประเภท คือ 1. กรณีที่ 2 กล้องไม่มีพื้นที่ FOV ร่วมกัน 2. กรณีที่ 2 กล้องมีพื้นที่ FOV ซ้อนทับกันและ 3. ในกรณีที่กล้องใดๆ ถูกแบ่งพื้นที่ FOV ออกจากกันเนื่องจากมีพื้นที่ FOV ของกล้องอื่นตัดขวาง

การจัดการปัญหาเรื่องของสถานะการปรากฏของผู้เล่นในแต่ละกล้องและปัญหาการเชื่อมโยงข้อมูลระหว่างเป้าหมายในระดับโลคอลกับเป้าหมายในระดับโกลบอล ดำเนินการโดยทำการค้นหา FOV ของแต่ละกล้องบนมุมมองเสมือน จากนั้นแบ่งประเภท FOV นี้เป็น 2 กลุ่ม คือ FOV ที่ซ้อนทับกันและ FOV ที่ไม่ซ้อนทับกับพื้นที่ใด จุดประสงค์เพื่อใช้ข้อมูลประเภทกลุ่มของ FOV ในการจัดการประชากรของวัตถุที่จะทำการผสมข้อมูล ซึ่งเดิมจะโปรเจ็คต์วัตถุทั้งหมดจากทุกกล้องแล้วมาจัดกลุ่มและทำการคำนวณหาว่าเป็นวัตถุเดียวกันหรือไม่ โดยวิธีที่เสนอนี้จะช่วยลดจำนวนวัตถุที่ไม่เกี่ยวข้องในการเปรียบเทียบความเหมือนได้ในระดับหนึ่ง

วิธีการคำนวณหา FOV ของแต่ละกล้องบนมุมมองเสมือน ในงานนี้ดำเนินการโดยคำนวณหาจากตำแหน่งของมุมทั้ง 4 ของภาพอินพุต โดยหลังจากกำหนดจุดคอนโทรลพอยต์ระหว่างมุมมองภาพและมุมมองเสมือนแล้วจะกำหนดตำแหน่งบริเวณมุมของภาพทั้ง 4 มุมของภาพในแต่ละกล้องจากนั้นทำการแปลงตำแหน่งนี้ลงบนภาพมุมมองเสมือนซึ่งเราจะได้ตำแหน่งบนมุมมองเสมือนทั้งหมด 4 จุด ซึ่งข้อมูลตำแหน่งนี้จะถูกเก็บไว้สำหรับการกำหนดพื้นที่ของ FOV ของกล้องแต่ละตัว และใช้สำหรับกระบวนการผสมข้อมูลต่อไป

ข้อมูลกล้องสามารถแทนด้วยสัญลักษณ์ ดังสมการที่ 6.3

$$Camera_List(CL) = \{c_1, c_2, ..., c_N\} \tag{6.3}$$

เมื่อ N คือ จำนวนของกล้อง รายการจุดมุมภาพของแต่ละกล้องแทนด้วยสัญลักษณ์ดังสมการที่ 6.4

$$Image_Point_List(IPL) = IPL_1^N = \{mp_1, mp_2, mp_3, mp_4\} \tag{6.4}$$

เมื่อ $mp_1 - mp_4$ คือ รายการพิกัด (x, y) ของจุดมุมภาพทั้ง 4 จุด (Corner Point Image) โดยเรียงลำดับจากตำแหน่งมุมบนซ้าย มุมบนขวา มุมล่างขวาและมุมล่างซ้ายและ N คือ จำนวนกล้อง หลังจากการแปลงตำแหน่งมุมภาพนี้ลงบนภาพมุมมองเสมือนจะได้รายการตำแหน่งมุมภาพบนภาพเสมือนซึ่งแทนด้วยสัญลักษณ์ ดังสมการที่ 6.5

$$Top_Point_List(TPL) = TPL_1^N = \{tp_1, tp_2, tp_3, tp_4\} \tag{6.5}$$

เมื่อ $tp_1 - tp_4$ คือ รายการพิกัด (x, y) ของทั้ง 4 จุด (Point Top) ของมุมมองภาพบนภาพมุมมองเสมือนของแต่ละกล้องและ N คือ จำนวนของกล้อง ในเบื้องต้นจะกำหนดรายการพื้นที่ FOV ทั้งหมดบนมุมมองเสมือนด้วยสัญลักษณ์ ดังสมการที่ 6.6

$$Area_List(AL) = \{A_1, A_2, ..., A_N\} = AL_1^N \tag{6.6}$$

เมื่อ A คือ รายการพื้นที่ FOV และ N คือ จำนวนพื้นที่ FOV บนมุมมองเสมือนซึ่งเป็นพื้นที่ที่ได้จากแต่ละกล้อง กำหนดชนิดของพื้นที่ที่ซ้อนทับกันด้วยสัญลักษณ์ OA (Overlap Area: OA) และชนิดของพื้นที่ที่ไม่ซ้อนทับกันด้วยสัญลักษณ์ NOA (Non-Overlap Area: NOA) จากนั้นทำการคำนวณหาพื้นที่ที่ซ้อนทับกันระหว่างพื้นที่ของแต่ละกล้องโดยใช้หลักการคำนวณพื้นที่ตามวิธีพื้นฐานทางคณิตศาสตร์ซึ่งมีขั้นตอนดังนี้ ขั้นตอนแรกในแต่ละพื้นที่ของ AL ทำการสร้างพื้นที่จากรายการจุดทั้ง 4 จุด จากรายการจุดใน TPL ซึ่งแทนด้วยสัญลักษณ์ ดังสมการที่ 6.7

พหุนปณฺ ศิต ชีเว

$$Area_Scope(AS) = AS_{AL_1^M} = TPL_1^N \tag{6.7}$$

เมื่อ M คือ จำนวนรายการพื้นที่ทั้งหมดบนมุมมองเสมือน N คือ จำนวนของรายการจุดมุมทั้ง 4 จุดของแต่ละกล้อง ขั้นตอนต่อมาทำการคำนวณหาพื้นที่ที่เกิดการซ้อนทับกันระหว่าง FOV โดยใช้หลักการของเซต (Set Logic) โดยในงานนี้จะใช้พื้นที่ทั้งหมดของภาพเสมือนในการคำนวณ โดยมีเหตุผลว่าในบางกล้องอาจมีบางพื้นที่ที่อยู่นอกขอบเขตสนามและในบางกรณีผู้เล่นอาจเคลื่อนที่ออกนอกสนามได้ ซึ่งการใช้พื้นที่ภาพทั้งหมดนี้จะช่วยลดปัญหากรณีการติดตามผู้เล่นที่ออกนอกขอบเขตสนามได้รวมถึงสามารถใช้ติดตามกรรมการข้างสนามได้อีกด้วย สุดท้ายจะได้รายการพื้นที่บนภาพเสมือนทั้งหมดมาโดยแทนด้วยสัญลักษณ์ ดังสมการที่ 6.8

$$New_Area_List(NAL) = \left\{ NA_1^K \right\} \tag{6.8}$$

เมื่อ K คือ จำนวนพื้นที่ทั้งหมดและ NA คือ รายการพื้นที่ ซึ่งประกอบด้วยรายการของตำแหน่งจุดมุมรอบพื้นที่แทนด้วยสัญลักษณ์ ดังสมการที่ 6.9

$$NA_1^K = \left\{ pc_1^M \right\} \tag{6.9}$$

โดย K คือ จำนวนรายการพื้นที่ pc คือ พิกัดจุดมุม (x, y) และ M คือ จำนวนจุดทั้งหมดของแต่ละพื้นที่ ซึ่งข้อมูลพื้นที่นี้จะถูกใช้สำหรับการตรวจสอบว่าตำแหน่งผู้เล่นอยู่ในพื้นที่ใดสำหรับประกอบการตัดสินใจในการผสมข้อมูลและการโปรเจ็คตำแหน่ง ขั้นตอนต่อมา คือ การกำหนดชนิดของแต่ละพื้นที่โดยแทนด้วยสัญลักษณ์ ดังสมการที่ 6.10

$$Area_Type(AT) = AT_1^K \left\{ NA_1^K, \left(OA \mid NOA \right) \right\} \tag{6.10}$$

โดยที่แต่ละพื้นที่จะถูกกำหนดชนิดเป็นพื้นที่ที่มีการซ้อนทับกัน (Overlap Area (OA)) หรือเป็นพื้นที่ที่ไม่มีการซ้อนทับกัน (Non-overlap Area (NOA)) ซึ่งข้อมูลของชนิดพื้นที่นี้จะถูกใช้สำหรับการตรวจสอบว่าพื้นที่ใดเป็นพื้นที่ที่มีการซ้อนทับระหว่างกล้องและพื้นที่ใดที่ไม่มีการซ้อนทับกันระหว่างกล้องสำหรับกระบวนการผสมข้อมูลและการโปรเจ็คตำแหน่งต่อไป

4) การแปลงมุมมอง (Perspective Transform)

การแปลงมุมมองในงานวิจัยนี้ดำเนินการโดยใช้หลักการของโฮโมกราฟฟี (Homography) หรือเพอสเปคทีพทรานส์ฟอร์ม (Perspective Transform) โดยในที่นี้จะใช้จุดควบคุมคอนโทรลพอยต์ (Control Point) ในการสร้างเมทริกซ์ความสัมพันธ์ โดยที่เมทริกซ์โฮโม

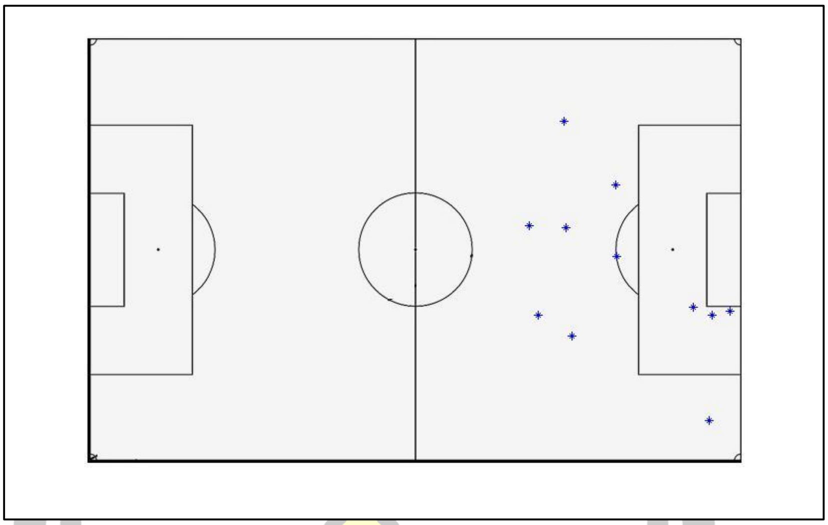
กราฟฟี่ (Homography Matrix: H) จะมีขนาด 3 x 3 แสดงดังสมการที่ 6.11 [1] เพื่อใช้คำนวณหาความสัมพันธ์ระหว่างมุมมองภาพและมุมมองเสมือนซึ่งแสดงดังสมการที่ 6.12 [1] ซึ่งผู้วิจัยได้ทำการเลือกจุดควบคุมนี้ด้วยมือทั้งหมดจำนวน 4 - 8 จุด ดังอธิบายไว้ในหัวข้อ 2) การกำหนดจุดควบคุม โดยในการแปลงมุมมองแบบเพอสเปคทีพทรานส์ฟอร์มจะใช้จุดควบคุมอย่างน้อย 4 จุดในการดำเนินการผู้วิจัยจึงเลือก 4 จุดแรกจากรายการสำหรับการคำนวณ

$$\begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix} \tag{6.11}$$

$$\begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix} \cong \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} \tag{6.12}$$

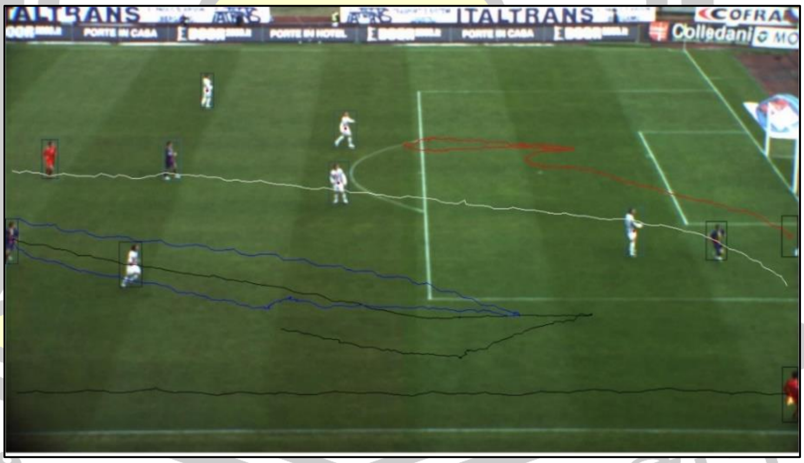
หลังจากได้จุดควบคุมคอนโทรลพอยต์แล้วข้อมูลนี้จะถูกใช้ในการสร้างเมทริกซ์ความสัมพันธ์ระหว่างมุมมอง จากนั้นเมทริกซ์ความสัมพันธ์ที่ได้จะถูกใช้ในการคำนวณตำแหน่งในมุมมองเสมือน (x_2, y_2) จากตำแหน่งอินพุตในมุมมองภาพ (x_1, y_1) ตัวอย่างการแปลงตำแหน่งผู้เล่นจากมุมมองภาพไปเป็นมุมมองเสมือนแสดงดังภาพที่ 6.19 และภาพที่ 6.20 ตามลำดับ

ภาพที่ 6.19 ตัวอย่างตำแหน่งผู้เล่นในมุมมองภาพ

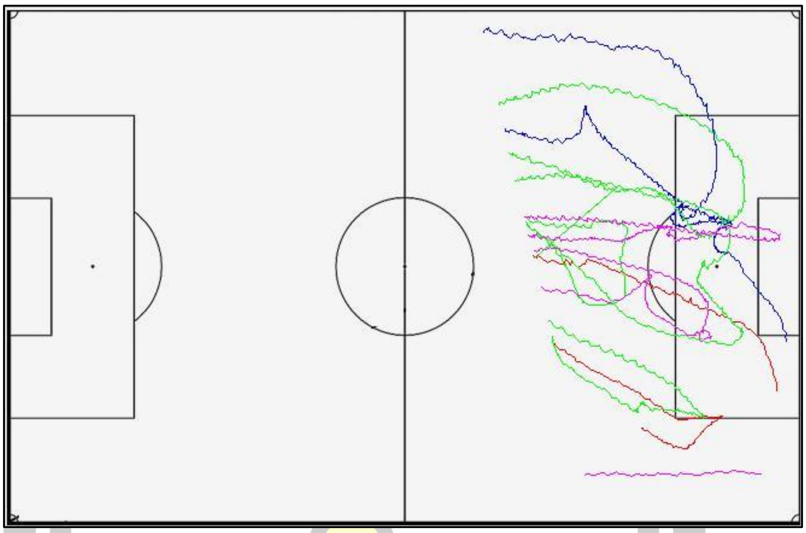


ภาพที่ 6.20 ตัวอย่างตำแหน่งผู้เล่นในมุมมองเสมือน

ข้อมูลเส้นทางของผู้เล่นในมุมมองภาพสามารถแสดงในมุมมองเสมือนได้โดยใช้เมทริกซ์ความสัมพันธ์นี้ซึ่งตัวอย่างการแสดงเส้นทางผู้เล่นในมุมมองภาพบนมุมมองเสมือนแสดงดังภาพที่ 6.21 และภาพที่ 6.22 ตามลำดับ



ภาพที่ 6.21 ตัวอย่างเส้นทางผู้เล่นในมุมมองภาพ



ภาพที่ 6.22 ตัวอย่างเส้นทางผู้เล่นในมุมมองเสมือน

1.4 ขั้นตอนการผสมข้อมูล (Data Fusion)

ภายหลังจากขั้นตอนการแปลงมุมมองของตำแหน่งผู้เล่นจากมุมมองภาพในแต่ละกล้องไปยังภาพในมุมมองเสมือนแล้ว ในขั้นตอนนี้จะเป็นการผสมข้อมูลตำแหน่งผู้เล่นจากแต่ละกล้องเข้าด้วยกันเพื่อแสดงผลตำแหน่งสุดท้ายบนภาพมุมมองเสมือนโดยมีรายละเอียดดังนี้

1) ปัจจัยที่ส่งผลต่อกระบวนการผสมข้อมูล

กระบวนการผสมข้อมูลตำแหน่งผู้เล่นจากแต่ละมุมมองกล้องเป็นตำแหน่งเดียวเพื่อแสดงผลในมุมมองเสมือนจะมีความถูกต้องมากน้อยขึ้นอยู่กับปัจจัยหลายด้าน ซึ่งจากการสังเกตและทดสอบระบบพบว่ามีอยู่ 3 ปัจจัยหลักที่สำคัญ ได้แก่

1.1) กระบวนการตรวจจับวัตถุ

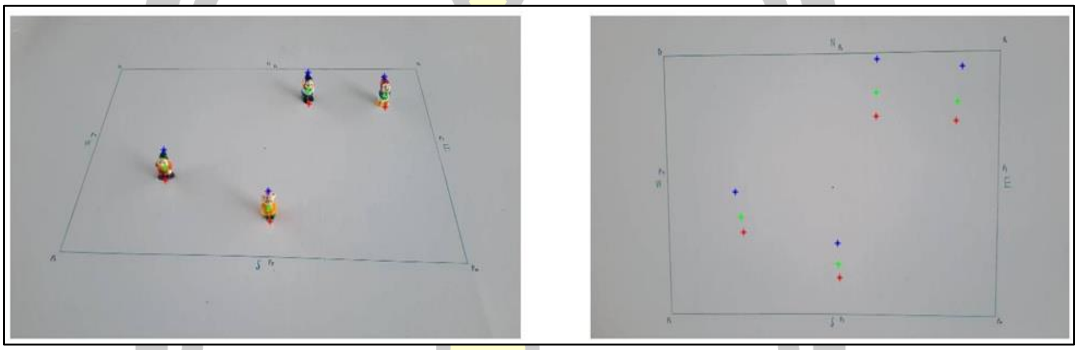
กระบวนการตรวจจับวัตถุอาจมีความผิดพลาดเกิดขึ้นได้ เช่น วัตถุมีขนาดเล็กหรือเกิดจากปัญหาการบดบัง เป็นต้น ซึ่งอาจทำให้ได้ตำแหน่งที่คลาดเคลื่อนจากตำแหน่งจริง

1.2) วิธีกำหนดตำแหน่งวัตถุที่เลือกใช้

การกำหนดตำแหน่งของวัตถุสำหรับการติดตามและแสดงผลนั้นมี 3 แบบ ที่นิยม ได้แก่ การใช้ตำแหน่งเท้า การใช้จุดศูนย์กลางของบล็อกและการใช้ตำแหน่งศรีษะ ตัวอย่างแสดงดังภาพที่ 6.23 ภาพด้านซ้าย ซึ่งทั้ง 3 แบบนี้มีข้อดีข้อเสียที่แตกต่างกัน เช่น การใช้ตำแหน่งเท้าจะได้ตำแหน่งที่ใกล้เคียงกับตำแหน่งจริงมากที่สุดแต่ถ้ากระบวนการตรวจจับเกิดผิดพลาดเนื่องจากวัตถุมีขนาดเล็กอาจทำให้ได้ตำแหน่งเท้าที่คลาดเคลื่อนไปจากตำแหน่งจริงได้เช่นกัน ในขณะที่การใช้จุดศูนย์กลางบล็อกจะดำเนินการได้ง่ายแต่ตำแหน่งอาจเปลี่ยนแปลงไปมาเนื่องจากมีการเปลี่ยนแปลง

พหุนปณสิต ชีเว

รูปร่างของวัตถุ เช่น การกางแขนหรือมีการเคลื่อนที่เกิดขึ้น ส่วนการใช้ตำแหน่งศีรษะจะมีปัญหาคล้ายกับกรณีการใช้ตำแหน่งเท้า คือ เมื่อวัตถุมีขนาดเล็กจะส่งผลให้ตรวจจับและประมาณตำแหน่งได้ยากทำให้อาจได้ตำแหน่งที่คลาดเคลื่อนไป เป็นต้น นอกจากนี้การใช้จุดศูนย์กลางบล็อกและศรีษะจะได้ตำแหน่งที่ห่างออกไปจากตำแหน่งจริงของวัตถุเพิ่มมากขึ้นตามความสูงของวัตถุนั้นๆ ตัวอย่างแสดงดังภาพที่ 6.23 ภาพด้านขวา

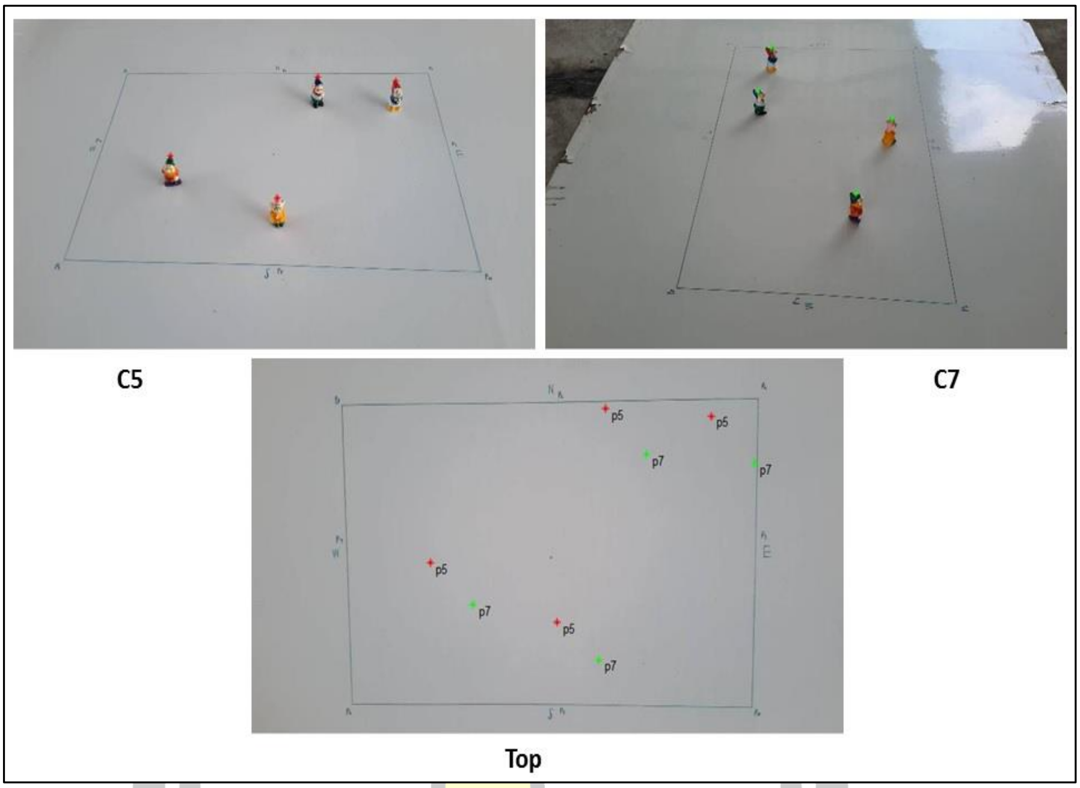


ภาพที่ 6.23 วิธีกำหนดตำแหน่งวัตถุและผลการประมาณตำแหน่งบนมุมมองเสมือน

ตามภาพที่ 6.23 ภาพด้านซ้ายมือเป็นตัวอย่างการกำหนดตำแหน่งวัตถุทั้ง 3 แบบโดยสัญลักษณ์ดอกจันสีแดง คือ การประมาณตำแหน่งบริเวณเท้า ดอกจันสีเขียว คือ การประมาณตำแหน่งบริเวณศูนย์กลางบล็อกและดอกจันสีน้ำเงิน คือ การประมาณตำแหน่งบริเวณศีรษะ ส่วนภาพที่ 6.23 ภาพด้านขวามือเป็นตัวอย่างผลที่ได้จากการประมาณตำแหน่งของทั้ง 3 วิธีบนภาพมุมมองเสมือนซึ่งจะเห็นว่าตำแหน่งที่ได้จะแตกต่างกันโดยวิธีกำหนดตำแหน่งโดยใช้ศูนย์กลางบล็อกและศีรษะจะได้ผลที่ไกลจากตำแหน่งจริงเพิ่มมากขึ้นตามลำดับ

1.3) ตำแหน่งและทิศทางของกล้อง

ตำแหน่งและทิศทางของกล้อง เป็นปัจจัยหนึ่งที่สำคัญต่อประสิทธิภาพของกระบวนการผสมข้อมูล เนื่องจากตำแหน่งกล้องที่ต่างกันจะส่งผลต่อตำแหน่งของวัตถุที่ต่างกัน อีกทั้งทิศทางของกล้องก็ส่งผลต่อตำแหน่งที่ได้ในทิศทางที่ต่างกันด้วย ตัวอย่างแสดงดังภาพที่ 6.24



ภาพที่ 6.24 ตำแหน่งวัตถุที่ได้จากมุมมองกล้องที่ต่างกัน

ตามภาพที่ 6.24 แสดงตัวอย่างการใช้ข้อมูลภาพจาก 2 กล้อง คือ C5 และ C7 ซึ่งอยู่ในตำแหน่งและมุมมองที่ต่างกันซึ่งในที่นี้ทั้ง 2 กล้องจะทำมุมกันประมาณ 90 องศา จะเห็นว่าเมื่อกล้องอยู่ในตำแหน่งและทิศทางที่ต่างกันตำแหน่งของวัตถุที่ได้ก็แตกต่างกันด้วย โดยตำแหน่งบนภาพมุมมองเสมือนที่ได้จากกล้อง C7 จะแทนด้วยสัญลักษณ์ดอกจันสีเขียว ส่วนตำแหน่งที่ได้จากกล้อง C5 จะแทนด้วยดอกจันสีแดง

2) ปัญหาของกระบวนการผสมข้อมูล

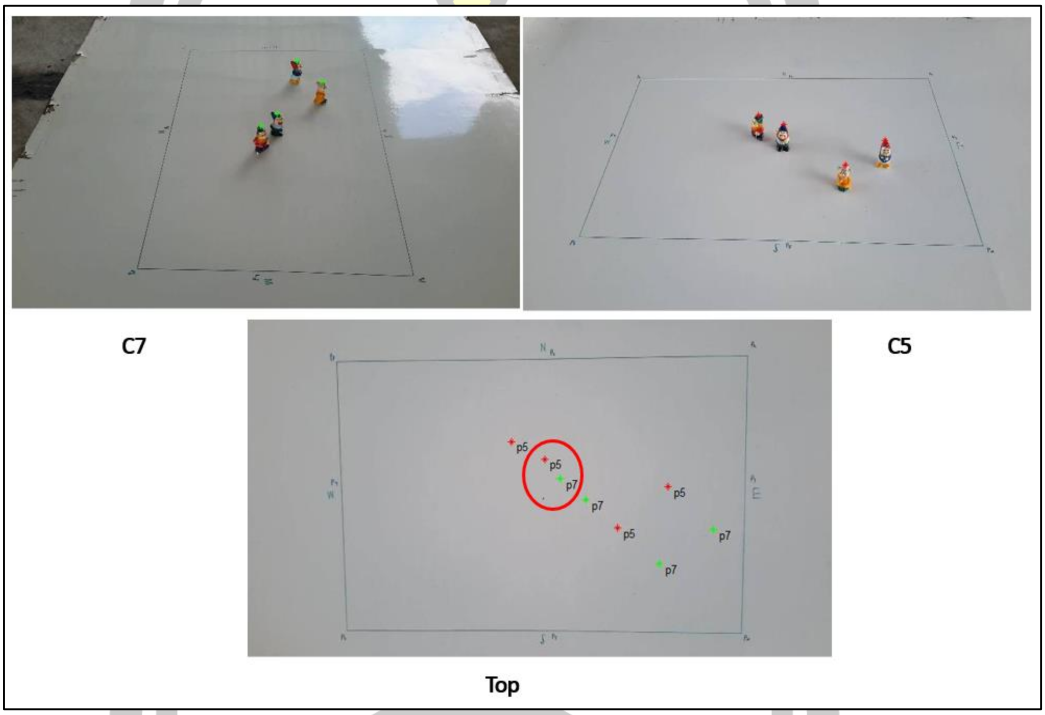
จากปัจจัยต่างๆ ที่ส่งผลต่อกระบวนการผสมข้อมูลตามที่ได้อธิบายไว้ข้างต้นนั้น สามารถสรุปประเด็นปัญหาที่พบได้ดังนี้

2.1) เมื่อวัตถุมีขนาดเล็ก ผลการตรวจจับตำแหน่งของวัตถุอาจคลาดเคลื่อนได้

2.2) การกำหนดตำแหน่งวัตถุ กรณีใช้ตำแหน่งเท้าจะได้ตำแหน่งที่ใกล้เคียงกับตำแหน่งจริงมากที่สุดแต่อาจคลาดเคลื่อนเมื่อการตรวจจับผิดพลาด กรณีใช้จุดศูนย์กลางบล็อกจะได้ตำแหน่งที่ห่างออกจากตำแหน่งจริงมากกว่าการใช้ตำแหน่งเท้าและกรณีใช้ตำแหน่งศรีษะจะได้ตำแหน่งที่ห่างออกจากตำแหน่งจริงมากกว่าการใช้จุดศูนย์กลางบล็อกตามลำดับ

2.3) ตำแหน่งและทิศทางของกล้อง เมื่อกล้องอยู่ในตำแหน่งและทิศทางที่ต่างกันจะส่งผลต่อตำแหน่งวัตถุที่ต่างกันด้วย

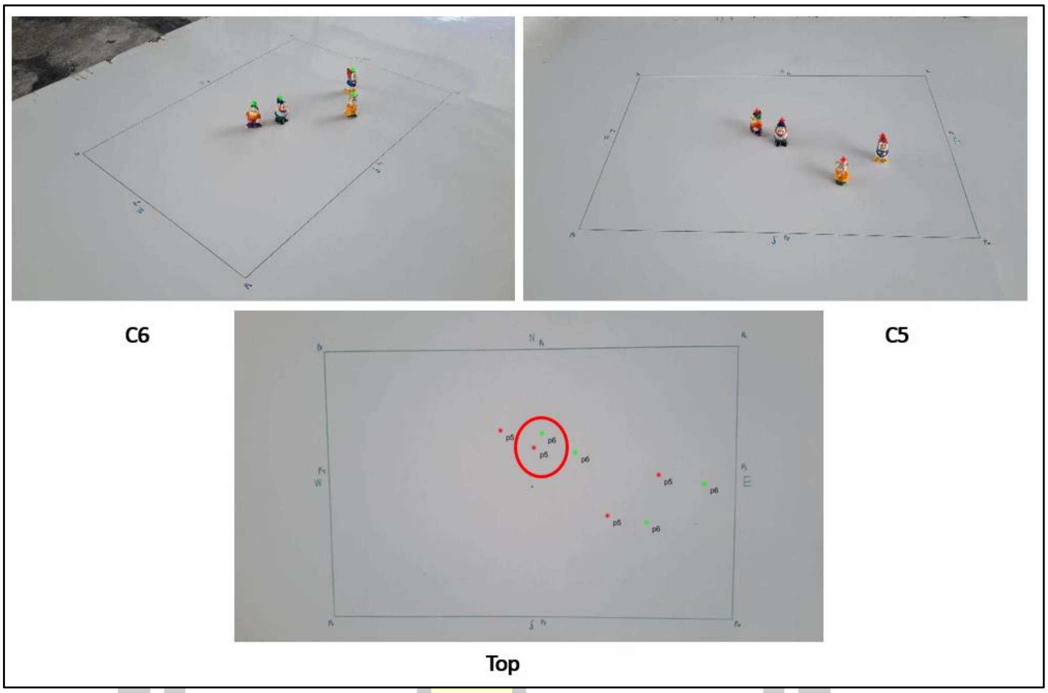
2.4) วิธีการกำหนดตำแหน่งวัตถุและตำแหน่งทางกายภาพของวัตถุ กรณีที่มีวัตถุที่อยู่ใกล้กันและเลือกกำหนดตำแหน่งวัตถุที่เป็นศูนย์กลางบล็อกหรือศรีษะ การผสมข้อมูลอาจผิดพลาดเนื่องจากตำแหน่งที่ได้ของ 2 วัตถุอยู่ใกล้กัน ตัวอย่างแสดงดังภาพที่ 6.25



ภาพที่ 6.25 ตำแหน่งของวัตถุที่อยู่ใกล้กัน

ตามภาพที่ 6.25 ตำแหน่งบนภาพมุมมองเสมือนจากกล้อง C7 จะแทนด้วยสัญลักษณ์ดอกจันสีเขียวส่วนตำแหน่งที่ได้จากกล้อง C5 จะแทนด้วยสัญลักษณ์ดอกจันสีแดง สังเกตว่า 2 วัตถุทางซ้ายมือในบริเวณวงกลมสีแดงมีตำแหน่งทางกายภาพที่ใกล้กันและในตัวอย่างนี้เลือกกำหนดตำแหน่งวัตถุโดยใช้ศีรษะส่งผลให้ตำแหน่งที่ได้ของ 2 วัตถุในวงกลมสีแดงอาจถูกรวมเป็นตำแหน่งเดียวกันหากใช้วิธีแบบเดิม คือ การพิจารณาจากระยะทาง ทั้งที่ความจริงเป็นคนละวัตถุกัน

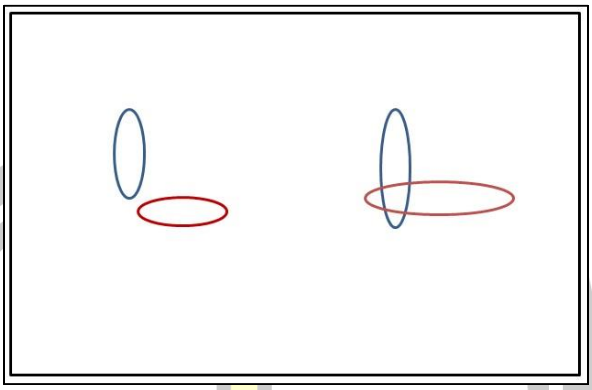
2.5) ตำแหน่งและทิศทางกล้องและตำแหน่งทางกายภาพของวัตถุ กรณีที่มีวัตถุที่อยู่ใกล้กันและตำแหน่งและทิศทางของกล้องเหมาะสม อาจส่งผลให้การผสมข้อมูลเกิดผิดพลาดเนื่องจากตำแหน่งที่ได้ของ 2 วัตถุอยู่ใกล้กัน ตัวอย่างแสดงดังภาพที่ 6.26 ซึ่งกล้อง C5 และ C6 มีลักษณะทำมุมที่เฉียงกันทำให้ตำแหน่งของ 2 วัตถุนั้นตกอยู่ในตำแหน่งที่ใกล้กัน



ภาพที่ 6.26 ผลจากตำแหน่งและทิศทางกล้องร่วมกับตำแหน่งทางกายภาพของวัตถุ

3) วิธีการผสมข้อมูลแบบเดิม

กระบวนการผสมข้อมูลโดยทั่วไปมีวิธีพื้นฐานสำหรับการตรวจสอบความเหมือนของ 2 วัตถุจากกล้องที่ต่างกันอยู่ 2 วิธี คือ การวัดจากค่าระยะทางระหว่าง 2 ตำแหน่งในมุมมองเสมือนและการตรวจสอบการซ้อนทับกันของ 2 คอนทัวร์ในมุมมองเสมือน ซึ่งทั้ง 2 วิธีมีข้อดีข้อเสียแตกต่างกัน เช่น การใช้ค่าระยะทางจะดำเนินการได้ง่ายใช้เอกลักษณ์สำหรับการคำนวณน้อยแต่อาจเกิดความผิดพลาดได้ง่ายเมื่อมี 3 ปัจจัยดังกล่าวข้างต้นเกิดขึ้น นอกจากนี้วิธีนี้จะต้องทำการเปรียบเทียบค่าระยะทางกับทุกๆ วัตถุที่มีซึ่งถ้ามีวัตถุจำนวนมากส่งผลให้ต้องประมวลผลมากขึ้น อีกทั้งการรวมข้อมูลระหว่าง 2 วัตถุที่เหมือนกันจะใช้วิธีหาตำแหน่งเฉลี่ยจาก 2 ตำแหน่งซึ่งอาจได้ตำแหน่งที่คลาดเคลื่อนไปจากตำแหน่งจริง ซึ่งเกิดจากการใช้ตำแหน่งวัตถุทั้ง 3 แบบและทิศทางของกล้องที่ต่างกัน ส่วนการใช้คอนทัวร์จะตรวจสอบความเหมือนระหว่างวัตถุได้ถูกต้องมากกว่าแบบแรกยกเว้นเมื่อกระบวนการตรวจจับที่ผิดพลาดซึ่งทำให้ได้คอนทัวร์ที่ไม่สมบูรณ์ส่งผลให้การตรวจสอบการซ้อนทับกันของพื้นที่คอนทัวร์จะผิดพลาดตามไปด้วย อีกทั้งต้องใช้เอกลักษณ์สำหรับการคำนวณที่มากขึ้นและนอกจากนี้กระบวนการตรวจสอบความเหมือนต้องซับซ้อนมากขึ้นเมื่อมีวัตถุอยู่ใกล้กันจำนวนมาก ตัวอย่างคอนทัวร์ที่ไม่สมบูรณ์แสดงดังภาพที่ 6.27 ภาพด้านซ้าย



ภาพที่ 6.27 ลักษณะคอนทัวร์ที่ไม่สมบูรณ์

4) แนวคิดเทคนิคการผสมข้อมูลที่เสนอ

เทคนิคของการผสมข้อมูลที่เสนอในงานนี้สามารถแก้ปัญหาที่เกิดขึ้นกับวิธีแบบเดิมได้ อีกทั้งยังมีความแข็งแรงต่อปัญหาที่เกิดจาก 3 ปัจจัยดังกล่าวข้างต้นได้ค่อนข้างดี นอกจากนี้ยังเป็นวิธีที่ง่ายและใช้เอกลักษณ์ในการประมวณผลน้อย วิธีที่เสนอมีแนวคิดและหลักการทำงานดังนี้

แนวคิดของวิธีที่เสนอประยุกต์มาจากหลักการของเงาที่ตกกระทบพื้นซึ่งโดยธรรมชาตินั้นเงาจะอยู่ในทิศทางตรงกันข้ามกับตำแหน่งของดวงอาทิตย์ ร่วมกับความรู้พื้นฐาน (Knowledge) ที่เกี่ยวข้องกับตำแหน่งทางกายภาพของวัตถุที่ปรากฏในภาพและโครงสร้างทางกายภาพของตัวภาพเอง ซึ่งเราสามารถอนุมานได้ว่าตำแหน่งและทิศทางของกล้องจะอยู่ตรงกันข้ามกับวัตถุที่ปรากฏในภาพเสมอซึ่งเปรียบได้กับดวงอาทิตย์ที่จะอยู่ตรงกันข้ามกับเงา คำถาม คือ ระบบจะหาตำแหน่งและทิศทางของกล้องโดยอาศัยข้อมูลของความรู้พื้นฐานทางกายภาพของวัตถุและโครงสร้างทางกายภาพของตัวภาพเองได้อย่างไร

5) การหาทิศทางของกล้อง

แนวคิดในการหาตำแหน่งและทิศทางของกล้องประยุกต์มาจากหลักการของเข็มทิศซึ่งเข็มทิศจะประกอบด้วย 2 ส่วนหลัก คือ แกนหมุนซึ่งเป็นจุดศูนย์กลางของเข็มทิศและแกนของเข็มทิศที่ใช้ชี้บอกตำแหน่งของทิศต่างๆ ซึ่งเราจะสังเกตเห็นว่าทิศเหนือจะอยู่ตรงกันข้ามกับทิศใต้เสมอโดยมีจุดศูนย์กลางของเข็มทิศหรือแกนหมุนอยู่ตรงกลาง จากหลักการของเข็มทิศเราสามารถนำมาประยุกต์ใช้ในการหาตำแหน่งและทิศทางของกล้องได้โดยใช้อัลกอริทึมที่ไม่ซับซ้อน คือ

<u>อัลกอริทึม :</u> การหาทิศทางของกล้องด้วยเข็มทิศ (Camera's Direction Finding based on Compass (CDFC))

Input: ตำแหน่งของวัตถุ

Output: ทิศทางของกล้อง

Start:

1) กำหนดตำแหน่งของวัตถุอินพุตเป็นฐาน เปรียบเสมือนจุดศูนย์กลางของเข็มทิศ ในที่นี้เรียกว่า ตำแหน่งฐาน (Base_Position)

2) ทำการสุ่มตำแหน่ง 1 ตำแหน่ง โดยมีระยะห่างจากตำแหน่งฐานในขนาดที่คงที่ *sx* ใดๆ และต้องอยู่ในตำแหน่งที่ตั้งฉากกับตำแหน่งฐานในแนวตั้งในมุมมองภาพ ในที่นี้เรียกว่า ตำแหน่งสุ่ม (Sampling_Position) ซึ่งเปรียบได้กับปลายของเข็มทิศที่ชี้ไปยังทิศเหนือ ซึ่งรายการพิกัดใหม่ที่ได้จะมี 2 พิกัดต่อ 1 รายการของตำแหน่งผู้เล่น ดังสมการที่ 6.13

$$NewX_1^N\left(base_point, sampling_point\right) = \left\{X_1^M(x,y), X_1^M(x,y) + (0, sx)\right\} \qquad (6.13)$$

3) ทำการแปลงมุมมอง 2 พิกัดนี้ลงบนภาพมุมมองเสมือน

4) ใช้สมการเส้นตรงในการสร้างเส้นตรงจาก 2 จุดดังกล่าว ดังสมการที่ 6.14

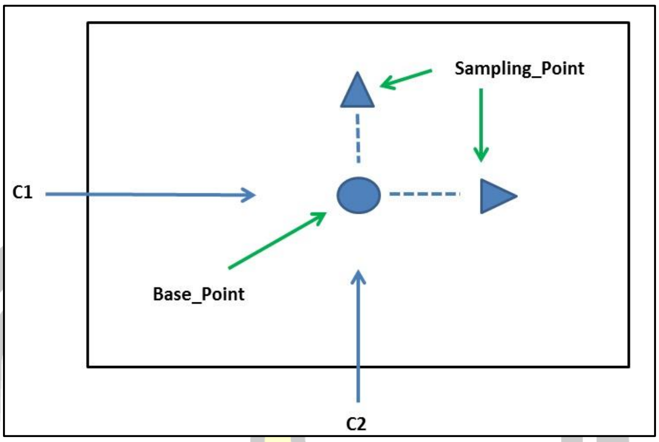
$$L = line\left(x_1^2, y_1^2\right) \qquad (6.14)$$

5) ทิศทางของกล้องจะอยู่ในทิศทางตรงกันข้ามกับตำแหน่งสุ่ม Sampling_Position ซึ่งเปรียบได้กับทิศใต้

End

หลังจากได้ตำแหน่งของวัตถุทั้ง 2 ตำแหน่งและทำการโปรเจ็คลงในภาพมุมมองเสมือนและใช้สมการเส้นตรงในการวาดเส้นตรงผ่าน 2 ตำแหน่งนี้แล้วจะสามารถทราบทิศทางของกล้องได้ว่าอยู่บริเวณใดของภาพในมุมมองเสมือน ตัวอย่างหลักการทำงานแสดงดังภาพที่ 6.28

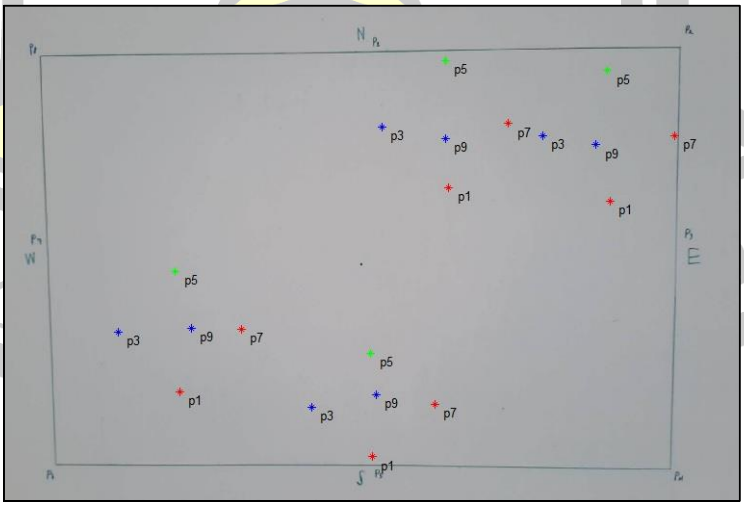
พหูน ปณ ฑิโต ชีโว



ภาพที่ 6.28 หลักการของเข็มทิศสำหรับหาทิศทางของกล้อง

ตามภาพที่ 6.28 เป็นตัวอย่างกรณีใช้ 2 กล้อง คือ C1 และ C2 ซึ่งอยู่ในทิศทางที่ต่างกันหลังจากกำหนดตำแหน่งฐานและตำแหน่งสุ่มและทำการโปรเจ็คลงบนมุมมองเสมือนแล้ว เมื่อใช้พื้นฐานของสมการเส้นตรงในการวาดเส้นตรงตัดผ่าน 2 ตำแหน่งดังกล่าวจะสามารถทราบทิศทางของตำแหน่งกล้องเหล่านี้ได้

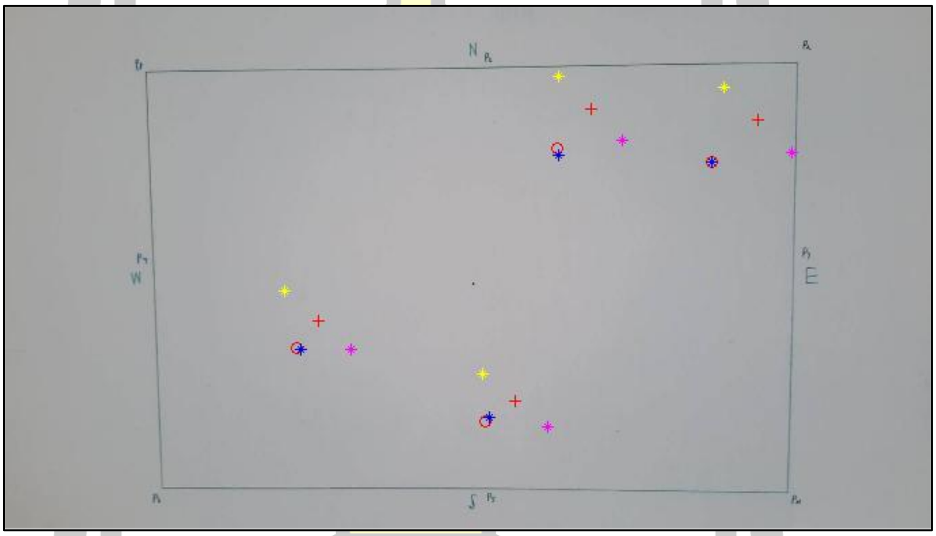
เหตุผลในการหาทิศทางของตำแหน่งกล้องเนื่องมาจากปัญหาของวิธีการกำหนดตำแหน่งของวัตถุทั้ง 3 แบบและทิศทางของกล้องที่ต่างกันจะส่งผลให้ตำแหน่งของวัตถุที่ปรากฏบนมุมมองเสมือนจะต่างกันไปด้วยซึ่งตำแหน่งที่ได้จะห่างจากตำแหน่งจริงมากขึ้นตามประเภทของวิธีกำหนดตำแหน่งที่ใช้และระยะของมุมกล้อง ตัวอย่างแสดงดังภาพที่ 6.29



ภาพที่ 6.29 การใช้ตำแหน่งศีรษะผู้เล่นและลักษณะมุมกล้องที่ต่างกัน

ตามภาพที่ 6.29 เป็นตัวอย่างการใช้มุมมองกล้องทั้งหมด 5 ตัว ซึ่งวางอยู่ในทิศทางที่แตกต่างกัน คือ ทิศเหนือ ทิศตะวันออก ทิศใต้ ทิศตะวันตกและกล้องตัวสุดท้ายจะอยู่ด้านบนเหนือศีรษะของวัตถุในตำแหน่ง p9 สำหรับใช้เป็นกราวนด์ทรูธ เนื่องจากเป็นตำแหน่งที่ใกล้เคียงกับตำแหน่งจริงมากที่สุด ซึ่งผลการทดสอบปรากฏว่าในแต่ละมุมกล้องจะได้ตำแหน่งที่ต่างกันและมีทิศทางที่ต่างกัน

นอกจากนี้ในกรณีที่มีมุมกล้องที่ต่างกันจะส่งผลให้วิธีการผสมข้อมูลแบบเดิม คือ วิธีวัดจากค่าระยะทาง ซึ่งทำการผสมข้อมูลด้วยการคำนวณหาค่าเฉลี่ยพิกัดระหว่าง 2 ตำแหน่งนั้นมีความคลาดเคลื่อนได้ ตัวอย่างแสดงดังภาพที่ 6.30



ภาพที่ 6.30 การหาค่าเฉลี่ยจาก 2 ตำแหน่ง

ตามภาพที่ 6.30 สัญลักษณ์ดอกจันสีเหลืองเป็นตำแหน่งที่ได้จากกล้อง C5 ซึ่งอยู่ในทิศใต้ สัญลักษณ์ดอกจันสีชมพูเป็นตำแหน่งที่ได้จากกล้อง C7 ซึ่งอยู่ในทิศตะวันตก สัญลักษณ์ดอกจันสีน้ำเงินเป็นตำแหน่งจริงของวัตถุซึ่ง คือ ตำแหน่งเท้าโดยได้จากกล้อง C9 ที่อยู่ในตำแหน่งด้านบนเหนือศีรษะ ส่วนเครื่องหมายบวกสีแดง คือ ตำแหน่งที่ได้จากการคำนวณค่าเฉลี่ยระหว่าง 2 ตำแหน่ง คือ สัญลักษณ์ดอกจันสีเหลืองและสัญลักษณ์ดอกจันสีชมพู ซึ่งจะเห็นว่ามีความคลาดเคลื่อนไปจากตำแหน่งจริงค่อนข้างมากและถ้าใช้วิธีกำหนดตำแหน่งวัตถุแบบจุดศูนย์กลางบล็อกและศรีษะผลที่ได้นั้นก็จะคลาดเคลื่อนมากขึ้น ส่วนสัญลักษณ์วงกลมสีแดง คือ ผลที่ได้จากวิธีที่เสนอซึ่งจะอธิบายในหัวข้อถัดไป

6) โมเดลรูปร่างของข้อมูลบนพื้นฐานทางกายภาพ (Physical based Data Shape Model)

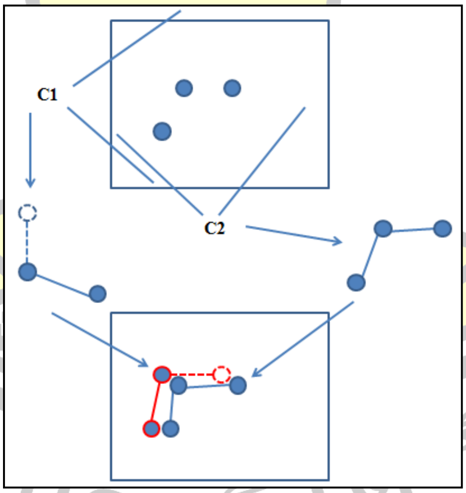
เทคนิคของการผสมข้อมูลที่เสนอในงานนี้อาศัยความรู้พื้นฐานเกี่ยวกับตำแหน่งทางกายภาพของวัตถุที่ปรากฏในภาพและตำแหน่งทางกายภาพของวัตถุบนโลกของความเป็นจริง โดยมีข้อสันนิษฐาน 3 ข้อดังนี้

1) ตำแหน่งของวัตถุที่ปรากฏในแต่ละมุมมองกล้องจะมีตำแหน่งบนโลกจริงที่เป็นตำแหน่งเดียวกันเสมอ

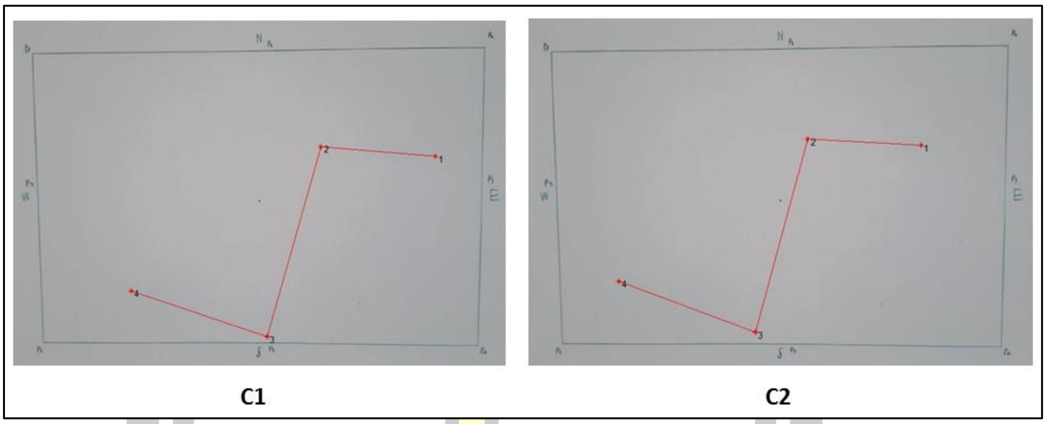
2) ตำแหน่งวัตถุที่ปรากฏในภาพของแต่ละมุมมองกล้องจะมีโครงสร้างที่เหมือนกันเสมอเมื่อมองในมุมมองเสมือนจริง เนื่องจากข้อสันนิษฐานข้อที่ 1

3) จำนวนของวัตถุที่ปรากฏในภาพ ในบริเวณของพื้นที่ FOV ที่ซ้อนทับกัน (Overlap Area) ระหว่างมุมมองกล้องจะมีจำนวนที่เท่ากันเสมอ แม้จะมีบางวัตถุที่ถูกบดบัง

จากข้อสันนิษฐานดังกล่าวเราสามารถอนุมานได้ว่า จำนวนวัตถุในมุมมองกล้องที่มีพื้นที่ FOV ที่ซ้อนทับกันจะมีจำนวนเท่ากัน นอกจากนี้ วัตถุที่ปรากฏในมุมมองเสมือนจะมีรูปร่างของโครงสร้างของตำแหน่งวัตถุที่เหมือนกัน ตัวอย่างแสดงดังภาพที่ 6.31 ซึ่งจากข้อสันนิษฐานนี้สามารถโมเดลรูปร่างของโครงสร้างของข้อมูล (Data Shape Model) ตำแหน่งวัตถุทั้งหมดที่ปรากฏในมุมมองเสมือนได้โดยอาศัยความรู้พื้นฐานทางกายภาพ ตัวอย่างดังภาพที่ 6.32



ภาพที่ 6.31 ลักษณะโครงสร้างของตำแหน่งของข้อมูล



ภาพที่ 6.32 โมเดลรูปร่างของข้อมูล (Data Shape Model)

ตามภาพที่ 6.32 เป็นตัวอย่างตำแหน่งของวัตถุจำนวน 4 วัตถุที่ได้จากกล้อง C1 และ C2 ซึ่งจะเห็นว่าจะมีรูปร่างของโครงสร้างตำแหน่งข้อมูลที่เหมือนกันเพียงแต่มีตำแหน่งที่แตกต่างกันเท่านั้น

สำหรับผู้เล่นที่ถูกบังจากผู้เล่นคนอื่นในงานนี้จะใช้วิธีพื้นฐานทั่วไปในการประมาณตำแหน่ง คือ การใช้หลักการทำนายตำแหน่งจากข้อมูลเดิมที่มีอยู่ โดยในงานนี้สันนิษฐานว่าแม้จะเป็นตำแหน่งที่ไม่แน่นอนแต่ยังสามารถใช้ข้อมูลผู้เล่นคนเดียวกันนี้จากกล้องตัวอื่นที่สามารถมองเห็นได้ชัดเจนเป็นข้อมูลช่วยให้การผสมข้อมูลมีความถูกต้องมากยิ่งขึ้นได้

7) หลักการสร้างโมเดลรูปร่างของโครงสร้างข้อมูล

หลักการสร้างโมเดลรูปร่างของข้อมูลมีเกณฑ์การพิจารณาจากประเภทของพื้นที่ FOV ทั้ง 3 ประเภทซึ่งได้อธิบายไว้ในหัวข้อ 1.3 ขั้นตอนการแปลงมุมมอง ดังนั้นการสร้างโมเดลรูปร่างของข้อมูลจะมี 2 กลุ่มหลัก คือ โมเดลรูปร่างของข้อมูลที่อยู่ในกลุ่มของพื้นที่ FOV ที่ซ้อนทับกันและโมเดลรูปร่างของข้อมูลที่อยู่ในกลุ่มของพื้นที่ FOV ที่ไม่ซ้อนทับกัน จากข้อมูลตำแหน่งผู้เล่นในแต่ละกล้องถ้าผู้เล่นอยู่ในพื้นที่ของ FOV ประเภทที่ซ้อนทับกันพื้นที่ใด ข้อมูลตำแหน่งผู้เล่นจะถูกรวมอยู่ในโมเดลรูปร่างของข้อมูลชุดเดียวกันนั้น ซึ่งจำนวนของโมเดลรูปร่างของข้อมูลจะเท่ากับจำนวนพื้นที่ที่มีการซ้อนทับกัน ในกรณีที่ผู้เล่นอยู่ในกลุ่มของ FOV ที่ไม่ซ้อนทับกับพื้นที่ใดจะถูกสร้างเป็นรายการไว้โดยไม่มีการสร้างเป็นโมเดลรูปร่างของข้อมูล จากนั้นข้อมูลโมเดลรูปร่างของข้อมูลทั้งหมดนี้จะถูกโปรเจ็คลงบนภาพเสมือนเพื่อใช้ในการประมวลผลต่อไป การสร้างโมเดลรูปร่างของโครงสร้างข้อมูลของแต่ละมุมมองกล้องจะดำเนินการเฉพาะตำแหน่งข้อมูลที่อยู่ภายในขอบเขตพื้นที่ที่มีการซ้อนทับกันเท่านั้น ซึ่งสามารถตรวจเช็คว่าตำแหน่งที่โปรเจ็คนั้นอยู่ในพื้นที่ที่ซ้อนทับหรือไม่โดยใช้วิธีการพื้นฐานทางคณิตศาสตร์โดยเปรียบเทียบตำแหน่งวัตถุที่โปรเจ็คกับข้อมูลพื้นที่จาก

New_Area_List (*NAL*) กรณีตำแหน่งของวัตถุที่ถูกโปรเจ็คอยู่ในพื้นที่ใดซึ่งเป็นพื้นที่ที่ซ้อนทับกันซึ่งถูกกำหนดไว้ในรายการชนิดพื้นที่ *Area_Type* (*AT*) ตำแหน่งดังกล่าวนี้จะถูกนำไปดำเนินการสร้างโมเดลรูปร่างของข้อมูลต่อไป ส่วนตำแหน่งของวัตถุใดที่อยู่ในพื้นที่ FOV ที่ไม่ซ้อนทับกับพื้นที่ของ FOV ใดๆ สามารถโปรเจ็คลงบนภาพมุมมองเสมือนได้โดยไม่ต้องผ่านกระบวนการผสมข้อมูล เนื่องจากไม่มีข้อมูลจากกล้องอื่นมาเกี่ยวข้อง ดังสมการที่ 6.15 ถึงสมการที่ 6.17

$$Model_list_1^N = X_1^M(x) \mid x \in NAL(AT = OA) \quad (6.15)$$

$$non_Model_1^N = X_1^M(x) \mid x \in NAL(AT = NOA) \quad (6.16)$$

$$project_list_1^K = \{Model_list_1^N; non_Model_1^N\} \quad (6.17)$$

หลักการสร้างโมเดลรูปร่างของข้อมูลสามารถดำเนินการได้โดยมี 3 ขั้นตอนดังนี้

1) นำตำแหน่งที่ตรวจจับได้ในมุมมองของแต่ละกล้องมาโปรเจ็คลงบนภาพมุมมองเสมือน

2) คำนวณหาระยะทาง (Distance) ของแต่ละตำแหน่งเทียบกับจุด Origin (O) ของภาพเสมือน ซึ่งโดยทั่วไปจะมีค่าคงที่ คือ (1, 1) ดังสมการที่ 6.18 และสมการที่ 6.19

$$O = (1,1) \quad (6.18)$$

$$Dist_1^N = pdist\left(X_1^M, O, \text{'euclidean'}\right) \quad (6.19)$$

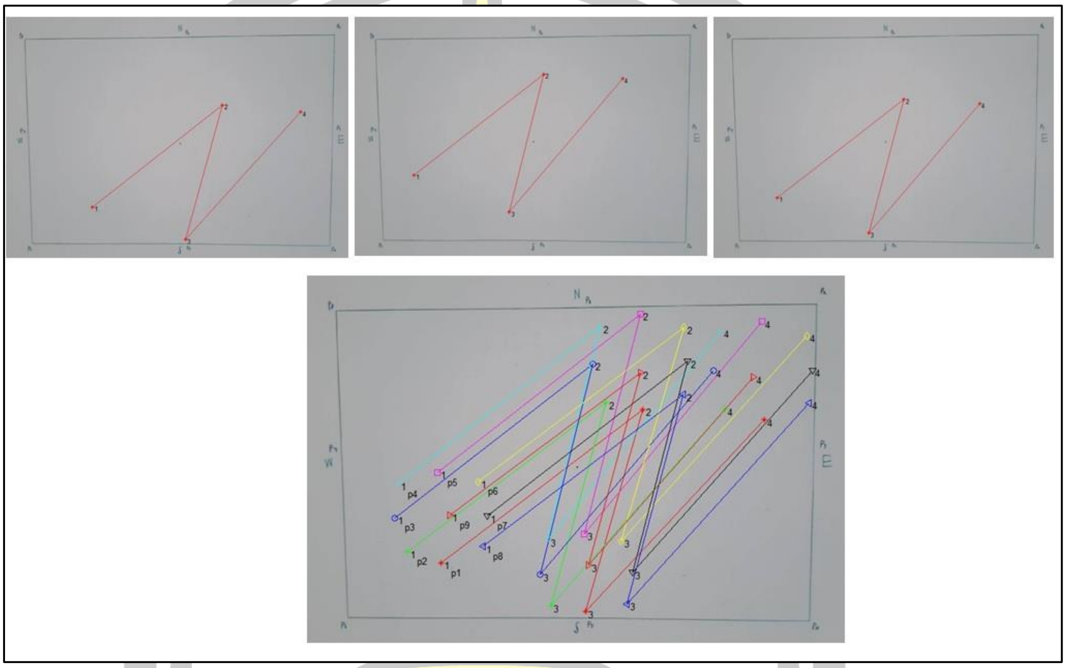
เมื่อ X_1^M คือ รายการพิกัดตำแหน่งผู้เล่นบนภาพเสมือนและ *O* คือ ค่าพิกัดมุม Origin (1,1) ของภาพเสมือน

3) ทำการจัดเรียงตำแหน่งใหม่โดยเรียงจากน้อยไปมากหรือจากมากไปน้อย โดยต้องใช้วิธีเดียวกันในทุกมุมมองกล้องซึ่งในที่นี้จะเลือกค่าของพิกัดที่น้อยกว่าก่อน กรณีที่ค่าระยะทางเท่ากันให้พิจารณาจากค่า *x* หรือค่า *y* อย่างใดอย่างหนึ่งโดยใช้วิธีเดียวกันในทุกกล้อง โดยในงานนี้ไม่พิจารณากรณีที่ค่าพิกัด (*x, y*) ทั้งสองค่าเท่ากัน โดยสันนิษฐานว่าในโลกของความเป็นจริงผู้เล่นจะไม่อยู่ในตำแหน่งเดียวกันในลักษณะที่ซ้อนทับกัน หลักการจัดเรียงตำแหน่งแสดงดังสมการที่ 6.20

$$S = sort\left(X_1^M\left(Dist_1^N\right)\right) \quad (6.20)$$

หลังจากดำเนินการทั้ง 3 ขั้นตอนแล้วจะได้โมเดลรูปร่างของโครงสร้างตำแหน่งข้อมูลซึ่งในงานนี้จะเรียกว่า โมเดลรูปร่างของข้อมูลทางกายภาพ (Physical based Data Shape

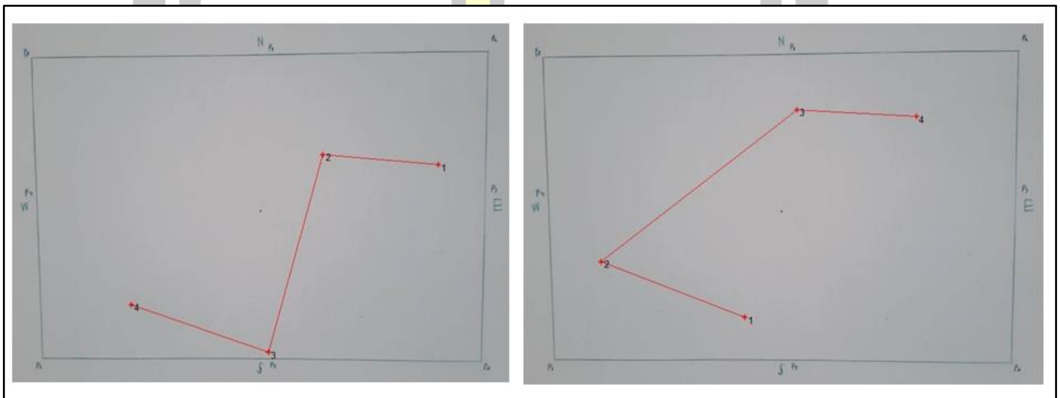
Model) ของแต่ละกล้องซึ่งจะมีรูปร่างและลำดับของตำแหน่งที่เหมือนกันในทุกมุมมองกล้อง ซึ่งจะอยู่ในลักษณะโมเดลเวกเตอร์ขนาด n x 2 โดยคอลัมน์แรก คือ พิกัด x ส่วนคอลัมน์ที่ 2 คือ พิกัด y ตัวอย่างโมเดลรูปร่างข้อมูลแสดงดังภาพที่ 6.33



ภาพที่ 6.33 โมเดลรูปร่างข้อมูลจากกล้อง 9 ตัว

ตามภาพที่ 6.33 ภาพในแถวบนเป็นตัวอย่างของข้อมูลอินพุตจำนวน 4 ตำแหน่งจากกล้องจำนวน 3 ตัวซึ่งผ่านกระบวนการสร้างโมเดลรูปร่างของข้อมูลเรียบร้อยแล้วซึ่งจะเห็นว่าทั้ง 3 กล้องมีลักษณะของรูปร่างข้อมูลที่เหมือนกันจะแตกต่างกันที่ตำแหน่งเท่านั้น ส่วนภาพในแถวล่างสุดเป็นตัวอย่างของข้อมูลอินพุตจำนวน 4 ตำแหน่งที่ได้จากกล้องจำนวน 9 ตัวในตำแหน่งและทิศทางต่างๆ ได้แก่ 8 ทิศรอบสนามและอีก 1 ตัวอยู่ในตำแหน่งด้านบนเหนือศีรษะ จากตัวอย่างจะเห็นว่าผลที่ได้นั้นตำแหน่งทั้ง 4 ตำแหน่งอยู่ในตำแหน่งและทิศทางที่ต่างกันในแต่ละกล้องสืบเนื่องจากปัจจัยหลักดังที่อธิบายไว้ในหัวข้อที่ผ่านมา แต่เมื่อผ่านกระบวนการสร้างโมเดลรูปร่างข้อมูลด้วยเทคนิคที่เสนอในงานวิจัยนี้จะเห็นว่าข้อมูลในแต่ละกล้องจะมีรูปร่างของข้อมูลที่เหมือนกัน จากข้อมูลโมเดลรูปร่างที่ได้นี้สามารถทำการผสมข้อมูลจากทุกมุมมองกล้องเป็นข้อมูลเดียวกันได้โดยทำการคำนวณตามลำดับรายการ โดยไม่ต้องทำการคำนวณระยะทางระหว่างวัตถุและเปรียบเทียบเอกลักษณ์อื่นๆ เช่น สี เหมือนกับวิธีเดิม

ปัญหาในการสร้างโมเดลรูปร่างของโครงสร้างตำแหน่งข้อมูล คือ ในแต่ละกล้องอาจตรวจจับจำนวนวัตถุที่ได้ไม่เหมือนกัน นอกจากนี้วิธีการเรียงข้อมูลในมุมมองเสมือนหากใช้วิธีการสุ่มตำแหน่งเริ่มต้นแล้วทำการคำนวณหาระยะทางระหว่างวัตถุด้วยกันเองจะส่งผลให้ลำดับรายการวัตถุที่ได้ไม่เหมือนกันในแต่ละกล้อง ตัวอย่างแสดงดังภาพที่ 6.34 ส่งผลให้การผสมข้อมูลไม่ถูกต้อง ซึ่งปัญหาดังกล่าวสามารถแก้ไขได้โดยการปฏิบัติตามขั้นตอนของการสร้างโมเดลรูปร่างข้อมูลดังที่อธิบายไว้ข้างต้น



ภาพที่ 6.34 โมเดลรูปร่างข้อมูลที่ต่างกัน

8) หลักการผสมข้อมูล (Data Fusion Principal)

หลักการผสมข้อมูลของเทคนิคการผสมข้อมูลที่เสนอในงานนี้อาศัยความรู้พื้นฐานที่ได้จากกระบวนการหาทิศทางของกล้องและโมเดลรูปร่างของโครงสร้างข้อมูลตำแหน่งที่คำนวณได้ร่วมกับเทคนิคการคำนวณหาจุดตัดระหว่างเส้นตรงสองเส้นบนพื้นฐานของสมการเส้นตรง (Straight Line Equation based Point Cross Finding Technique) โดยมี 7 ขั้นตอนดังนี้

1) ทำการผสมข้อมูลทีละคู่ระหว่างโมเดลเวกเตอร์ของแต่ละกล้องที่อยู่ในพื้นที่ซ้อนทับเดียวกัน โดยเรียงตามลำดับรายการในเวกเตอร์

2) ในแต่ละคู่ที่คำนวณ แต่ละวัตถุในรายการเวกเตอร์จะมีพิกัดของวัตถุจำนวน 2 พิกัด คือ พิกัดตำแหน่งฐานในที่นี้กำหนดเป็น (x_1, y_1) และพิกัดตำแหน่งสุ่มในที่นี้กำหนดเป็น (x_2, y_2) ซึ่งได้มาจากกระบวนการหาทิศทางของกล้อง

3) ในแต่ละวัตถุ ทำการคำนวณหาความชัน (Slope) ของเส้นตรงที่ตัดผ่าน 2 จุดดังกล่าว ดังสมการที่ 6.21

$$m = \frac{(y_2 - y_1)}{(x_2 - x_1)} \tag{6.21}$$

4) คำนวณหาจุดตัดระหว่างเส้นตรง 2 เส้นที่ได้จาก 2 วัตถุ ซึ่งจุดตัดนี้ คือผลลัพธ์ของการผสมข้อมูล ดังสมการที่ 6.22

$$\begin{aligned} new_x &= \frac{\left(\left(\left((x_1 * y_2) - (y_1 * x_2)\right) * (x_3 - x_4)\right) - \left((x_1 - x_2) * \left((x_3 * y_4) - (y_3 * x_4)\right)\right)\right)}{\left(\left((x_1 - x_2) * (y_3 - y_4)\right) - \left((y_1 - y_2) * (x_3 - x_4)\right)\right)} \\ new_y &= \frac{\left(\left(\left((x_1 * y_2) - (y_1 * x_2)\right) * (y_3 - y_4)\right) - \left((y_1 - y_2) * \left((x_3 * y_4) - (y_3 * x_4)\right)\right)\right)}{\left(\left((x_1 - x_2) * (y_3 - y_4)\right) - \left((y_1 - y_2) * (x_3 - x_4)\right)\right)} \end{aligned} \tag{6.22}$$

โดยที่ $(x_3, y_3), (x_4, y_4)$ คือ พิกัดทั้ง 2 จุดของเส้นตรงที่สองในส่วนของพิกัดตำแหน่งฐานและพิกัดตำแหน่งสุ่มตามลำดับ

5) กรณีเส้นตรง 2 เส้นนั้นขนานกันให้ใช้วิธีคำนวณค่าเฉลี่ยระหว่าง 2 ตำแหน่งฐานตามปกติ ดังสมการที่ 6.23 และสมการที่ 6.24

$$parallel = (m_1 = m_2)\begin{cases} Inf, -Inf \\ any\ value \end{cases} \tag{6.23}$$

$$(new_x, new_y) = \left(mean(x_1, x_3), \mathrm{mean}(y_1, y_3)\right) \tag{6.24}$$

6) กรณีที่เส้นตรง 2 เส้นนี้ไม่ขนานกันแต่มีลักษณะเกือบจะขนานกัน ในที่นี้จะดำเนินการโดยใช้วิธีหาค่าเฉลี่ย เช่น เดียวกับในข้อ 5) เนื่องจากตำแหน่งของจุดตัดระหว่าง 2 เส้นตรงจะอยู่ห่างกันมากขึ้นตามองศาที่น้อยลงระหว่าง 2 เส้นตรงนี้ การกำหนดเงื่อนไของศาระหว่าง 2 เส้นตรงคำนวณได้ดังสมการที่ 6.25

$$\theta = \left|\left(\tan^{-1}(m_1)\right) - \left(\tan^{-1}(m_2)\right)\right| \tag{6.25}$$

โดยในงานนี้กำหนดเงื่อนไขของค่า θ ที่ 0.1 กรณีค่าองศาระหว่าง 2 เส้นตรงน้อยกว่าหรือเท่ากับค่าที่กำหนดจะทำการผสมข้อมูลโดยใช้วิธีหาค่าเฉลี่ยตามข้อ 5) กรณีมีค่ามากกว่าจะใช้วิธีที่เสนอตามข้อ 4)

7) จะได้ตำแหน่งใหม่ที่เกิดจากการผสมข้อมูลระหว่าง 2 วัตถุใดๆ

จากขั้นตอนการผสมข้อมูลข้างต้นจะได้ตำแหน่งใหม่จากการผสมข้อมูลระหว่าง 2 ตำแหน่งจาก 2 กล้องใดๆ ในกรณีที่จำนวนกล้องมากกว่า 2 ตัว สามารถดำเนินการผสมข้อมูลได้ 2 แนวทาง คือ แนวทางแรกใช้เพียง 2 กล้องแรกในการผสมข้อมูลซึ่งในงานนี้สันนิษฐานว่าข้อมูลจาก 2 กล้องน่าจะเพียงพอในการประมาณตำแหน่งที่ใกล้เคียงกับตำแหน่งจริงแล้ว ส่วนแนวทางที่สอง คือ ทำการผสมข้อมูลระหว่าง 2 กล้องทีละคู่จนครบ เช่น มีกล้องทั้งหมด 3 ตัว ให้ทำการจับคู่ระหว่างกล้อง C1 กับ C2 และระหว่าง C2 กับ C3 ซึ่งจะได้ผลการผสมข้อมูลมา 2 ชุด จากนั้นหาค่าเฉลี่ยในแต่ละตำแหน่งจากรายการของผลการผสมข้อมูลทั้งหมด สุดท้ายจะได้ผลการผสมข้อมูลจากทุกกล้อง เป็นต้น จากกระบวนการผสมข้อมูลที่เสนอดังกล่าวนี้จะมั่นใจได้ว่าจะได้ตำแหน่งที่ใกล้เคียงกับตำแหน่งจริงมากที่สุดซึ่งให้ผลดีกว่าแบบเดิมที่หาจากค่าเฉลี่ยเพียงอย่างเดียว นอกจากนี้ยังมั่นใจได้ว่าวิธีที่เสนอนี้มีความแข็งแรงต่อปัญหาที่เกิดจาก 3 ปัจจัยหลักดังที่ได้อธิบายไปแล้วข้างต้น

1.5 ขั้นตอนการเชื่อมโยงข้อมูล (Data Association)

ภายหลังจากขั้นตอนการผสมข้อมูลเสร็จสิ้น ในขั้นตอนนี้จะทำการเชื่อมโยงข้อมูลจากแต่ละมุมมองกล้องเข้าด้วยกันเพื่อใช้ประกอบสำหรับกระบวนการติดตาม การระบุตัวตนและการแก้ปัญหาการบดบังในกระบวนการถัดไป โดยการเชื่อมโยงข้อมูลมีรายละเอียดดังนี้

จากกระบวนการผสมข้อมูลซึ่งมีรายการกล้อง Camera_List (CL) กำหนดให้ตำแหน่งผู้เล่นบนมุมมองเสมือนแทนด้วยสัญลักษณ์ ดังสมการที่ 6.26

$$Top_Player(TP) = \left\{ tp_1^N \right\} \qquad (6.26)$$

โดย N คือ จำนวนตำแหน่งผู้เล่นทั้งหมดบนมุมมองเสมือนที่ได้จากกระบวนการผสมข้อมูลและ tp คือ พิกัด (x, y) โดยกำหนดให้ตำแหน่งผู้เล่นในแต่ละกล้องแทนด้วยสัญลักษณ์ ดังสมการที่ 6.27

$$Camera_Player(CP) = CP_1^N = \left\{ cp_1^M \right\} \qquad (6.27)$$

โดย cp คือ พิกัด (x, y) ส่วน M คือ จำนวนตำแหน่งผู้เล่นทั้งหมดของแต่ละกล้อง และ N คือ รายการกล้องทั้งหมด ในที่นี้สามารถกำหนดความสัมพันธ์ของผู้เล่นระหว่างมุมมองเสมือนกับมุมมองกล้องต่างๆ ดังสมการที่ 6.28

$$View_Data_Association(VDA) = VDA_1^K = \left\{ TP_1^N, CP_1^M, CL_1^J \right\} \qquad (6.28)$$

โดยที่ K คือ จำนวนความสัมพันธ์ของข้อมูล N คือ จำนวนตำแหน่งผู้เล่นบนมุมมองเสมือน M คือ จำนวนตำแหน่งผู้เล่นบนมุมมองภาพและ J คือ รายการกล้องทั้งหมด

1.6 ขั้นตอนการเลือกเป้าหมาย (Target Selection)

ภายหลังจากขั้นตอนการผสมข้อมูลและการเชื่อมโยงข้อมูล ในขั้นตอนนี้จะทำการเลือกเป้าหมายผู้เล่นเพื่อใช้สำหรับการติดตามซึ่งในงานนี้จะพิจารณาเพียง 23 คน ได้แก่ ผู้เล่นทั้ง 2 ทีมและกรรมการกลาง 1 คน โดยไม่พิจารณากรรมการข้างสนาม 2 คน เนื่องจากอยู่นอกเหนือขอบเขตของการวิจัย ซึ่งในกระบวนการตรวจจับผู้เล่นในมุมมองภาพนั้นได้มีการคัดกรองวัตถุที่คาดว่าจะเป็นผู้เล่นอยู่แล้วในเบื้องต้น แต่ในบางกรณีการตรวจจับวัตถุในบางกล้องอาจผิดพลาด เช่น ได้จำนวนผู้เล่นที่ไม่ครบ มีสัญญาณรบกวน หรือมีวัตถุอื่นที่คล้ายกันอยู่บริเวณใกล้เคียง เป็นต้น ในขั้นตอนนี้จะเป็นการคัดกรองเป้าหมายเพิ่มเติมเพื่อความถูกต้องในการเลือกเป้าหมายที่ต้องการติดตาม โดยพิจารณาจากข้อมูลสี ขนาดและจำนวนของวัตถุ หากวัตถุที่เป็นผลจากการผสมนั้นได้มาจากวัตถุในหลายกล้องที่มีลักษณะของสี และขนาดที่คล้ายกันและมีจำนวนของกล้องที่ตรวจจับพบมากกว่า 1 กล้อง ในพื้นที่ FOV ที่ซ้อนทับกันจะถูกพิจารณาว่าเป็นผู้เล่นหรือกรรมการ ซึ่งจะถูกนับจำนวนเข้ารายการเป้าหมายเป็นลำดับแรก ในกรณีวัตถุที่ได้จากการผสมนั้นได้มาจากข้อมูลกล้องเพียงตัวเดียวในพื้นที่ FOV ที่ไม่ซ้อนทับกัน แต่มีลักษณะของเอกลักษณ์สีและขนาดที่เหมือนกับวัตถุอื่นในรายการวัตถุที่ผ่านการผสมข้อมูลแล้วและถูกนับเข้ารายการเป้าหมายแล้วจะพิจารณาว่าเป็นผู้เล่นหรือกรรมการ ซึ่งจะถูกนับจำนวนเข้ารายการเป้าหมายเป็นลำดับรองลงมา โดยจะไล่พิจารณารายการวัตถุที่ผ่านการผสมแล้วนั้นทีละรายการจนครบ 23 คน โดยในงานนี้สันนิษฐานว่า แม้ว่าจำนวนผู้เล่นที่ปรากฏในแต่ละกล้องจะไม่ครบจำนวน 23 คนแต่เมื่อโปรเจ็คข้อมูลจากทุกกล้องลงบนมุมมองเสมือนแล้ว ในมุมมองเสมือนนั้นควรที่จะมีข้อมูลผู้เล่นครบจำนวน 23 คน ในกรณีที่เฟรมแรกกรรมการกลางถูกตรวจจับพบเพียงกล้องเดียวซึ่งจะไม่มีรายการวัตถุเดียวกันนี้ในรายการเป้าหมาย ในกรณีนี้จะถูกพิจารณาเป็นลำดับสุดท้ายซึ่งสันนิษฐานว่าในความเป็นจริงถ้าไล่นับจำนวนผู้เล่นครบทั้ง 2 ทีม แล้วซึ่งจะได้จำนวนเป้าหมายจำนวน 22 พอดีซึ่งจะเหลือ 1 เป้าหมายสุดท้ายสำหรับกรรมการกลางพอดีด้วย หลักการเลือกเป้าหมายแสดงดังสมการที่ 6.29

$$target_1^N = blob_1^M \begin{cases} 1)\ count(x_{size,color} \in \mathrm{NAL}(\mathrm{AT} = OA)) > 1 \\ 2)\ x_{size,color} \in \mathrm{NAL}(\mathrm{AT} = NOA) = x_{size,color} \in \mathrm{NAL}(\mathrm{AT} = OA) \end{cases} \qquad (6.29)$$

1.7 ขั้นตอนการระบุตัวตนผู้เล่น (Player Identification)

หลังจากขั้นตอนการเลือกเป้าหมายในขั้นตอนนี้จะทำการระบุตัวตนของผู้เล่นและกรรมการ โดยมีรายละเอียดดังนี้ จากรายการเป้าหมายที่ได้จำนวน 23 เป้าหมาย ทำการพิจารณาวัตถุเป้าหมายที่มีเอกลักษณ์สีที่เหมือนกันจำนวนเกิน 2 คนเป็นทีม A ขณะเดียวกันวัตถุเป้าหมายที่มีสี

ที่ต่างกันและมีจำนวนเกิน 2 คนจะถูกพิจารณาเป็นทีม B ส่วนวัตถุเป้าหมายที่มีเอกลักษณ์สีที่ไม่เข้ากับ 2 กลุ่มแรกหรือมีเพียง 1 คนจะถูกพิจารณาเป็นกรรมการกลาง กรณีทั้ง 2 ทีมมีเอกลักษณ์สีที่เหมือนกันจะแก้ปัญหาโดยพิจารณาจากตำแหน่งการยืนในสนามในเฟรมแรกซึ่งเป็นเฟรมก่อนเริ่มเกมส์การแข่งขัน (ถ้ามี) หากเป้าหมายอยู่ในฝั่งสนามเดียวกันจะถูกพิจารณาเป็นทีมเดียวกัน หลังจากจำแนกทีมและกรรมการกลางแล้วจะทำการกำหนด ID ประจำแต่ละคนโดยแทนด้วยสัญลักษณ์ดังนี้ ทีม $A=\{a_1^N\}$ ทีม $B=\{b_1^N\}$ และกรรมการกลาง คือ C ดังสมการที่ 6.30

$$identity_1^N = \begin{cases} team\ A \mid \text{count}(x^{color}_{\in target_1^N}) > 1 \\ team\ B \mid count(x^{color}_{\in target_1^N}) > 1 \\ \text{referee C, otherwise} \end{cases} \tag{6.30}$$

1.8 ขั้นตอนการแสดงผลผู้เล่น (Player Representation)

หลังจากขั้นตอนการระบุตัวตน ข้อมูล ID ของผู้เล่นและกรรมการกลางจะถูกนำมาใช้แสดงบนมุมมองเสมือนและมุมมองภาพในแต่ละกล้อง โดยในมุมมองเสมือนจะแสดงลาเบลตำแหน่งผู้เล่นด้วยสัญลักษณ์วงกลมพร้อมด้วย ID ประจำตัวซึ่งทั้ง 2 ทีมและกรรมการจะมีสีของวงกลมที่ต่างกัน ขนาดของวงกลมได้มาจากขนาดรัสมี r ที่เป็นความกว้างของบล็อบที่ได้จากหัวข้อ 1.2 ขั้นตอนการสกัดเอกลักษณ์ ซึ่งผ่านการคำนวณหาตำแหน่งในมุมมองภาพและทำการแปลงมุมมองแล้วโดยจะกำหนดเป็นค่าคงที่

1.9 ขั้นตอนการติดตามผู้เล่น (Player Tracking)

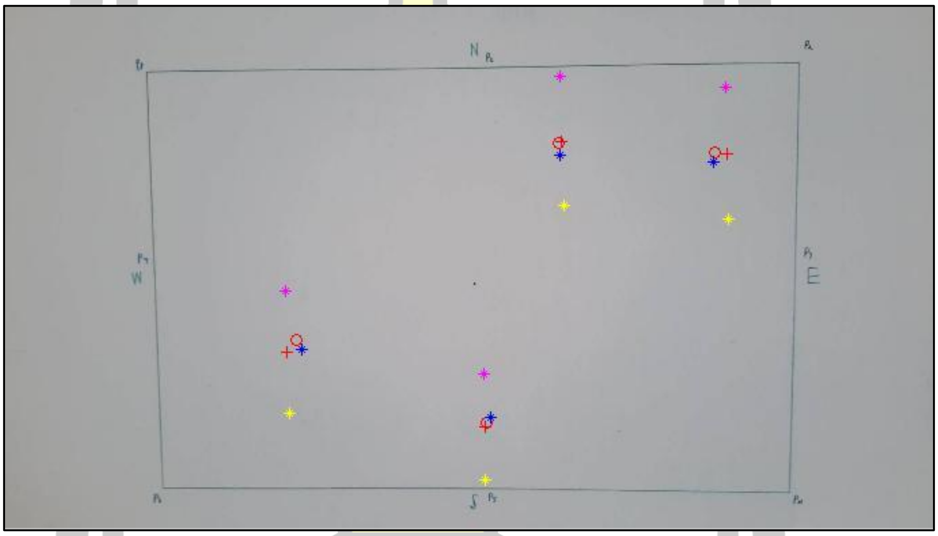
หลังจากกำหนดลาเบลให้กับผู้เล่นและกรรมการกลางแล้ว กระบวนการสุดท้ายจะเป็นการติดตามผู้เล่นในมุมมองเสมือนโดยในงานนี้ใช้วิธีพื้นฐาน คือ ใช้ตัวกรองคาร์ลมาลฟิลเตอร์ในการติดตามตำแหน่ง เอกลักษณ์ที่ใช้ คือ ตำแหน่งพิกัดของผู้เล่นบนมุมมองเสมือน

2. แนวทางการแก้ปัญหาการบดบัง (Occlusion Solution)

จากกระบวนการผสมข้อมูลจากหลายกล้อง สามารถช่วยในการแก้ปัญหาการบดบังได้ โดยในกรณีที่เกิดการบดบังในบางกล้องระบบจะยังสามารถใช้ข้อมูลตำแหน่งจากกล้องอื่นที่สามารถมองเห็นได้มาช่วยในการติดตามและเพิ่มความถูกต้องในการทำนายตำแหน่งของผู้เล่นในกล้องที่เกิดการบดบังได้ ในกรณีที่เกิดการบดบังในทุกกล้องระบบจะยังสามารถใช้ข้อมูลที่ได้จากการผสมข้อมูลจากการทำนายในแต่ละกล้องร่วมกับข้อมูลการทำนายเส้นทางที่ได้จากมุมมองเสมือนได้ซึ่งจะยังช่วยให้สามารถประมาณตำแหน่งผู้เล่นได้ใกล้เคียงกับตำแหน่งจริงมากกว่ากรณีการใช้ข้อมูลการทำนายจากกล้องเพียงตัวเดียว

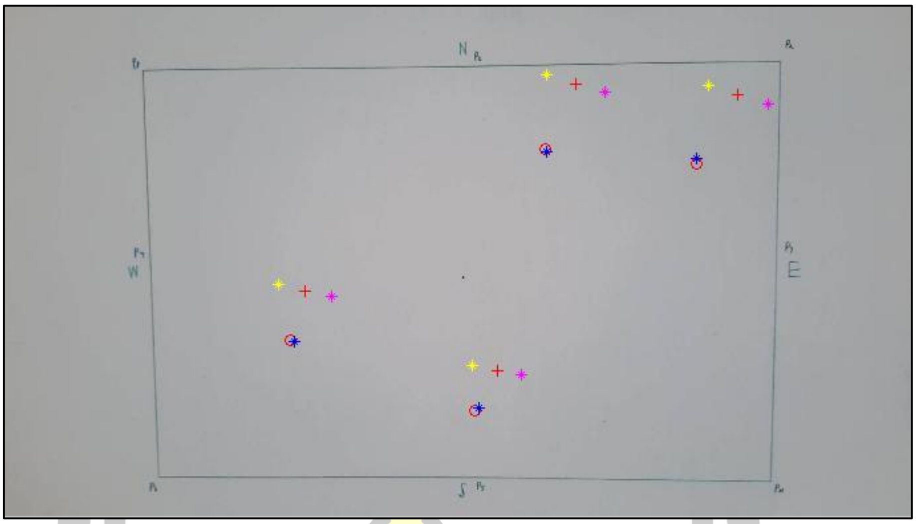
6.2 กระบวนการทดลองและผลการทดลอง (Experimental Process and Results)

กระบวนการผสมข้อมูล การทดลองนี้ทำการเปรียบเทียบประสิทธิภาพของกระบวนการผสมข้อมูลระหว่างวิธีเดิมที่คำนวณหาตำแหน่งจากค่าเฉลี่ยของ 2 ตำแหน่งใดๆ กับวิธีใหม่ที่เสนอซึ่งใช้อัลกอริทึมในการหาทิศทางของกล้องร่วมกับความรู้พื้นฐานทางกายภาพและการประมาณตำแหน่งโดยมีพื้นฐานอยู่บนสมการเส้นตรง โดยชุดข้อมูลทดสอบประกอบด้วยตำแหน่งวัตถุจำนวน 4 วัตถุโดยกำหนดขึ้นด้วยมือ จำนวนกล้องที่ใช้บันทึกภาพจำนวน 9 ตัว จากนั้นทำการเปรียบเทียบการผสมข้อมูลระหว่าง 2 กล้องใดๆ ในที่นี้ทดลองผสมข้อมูลระหว่าง กล้อง C1-C5, C5-C6, C5-C7 และ C5-C8 ตัวอย่างผลการทดลองแสดงดังภาพที่ 6.35 - ภาพที่ 6.38 ดังนี้



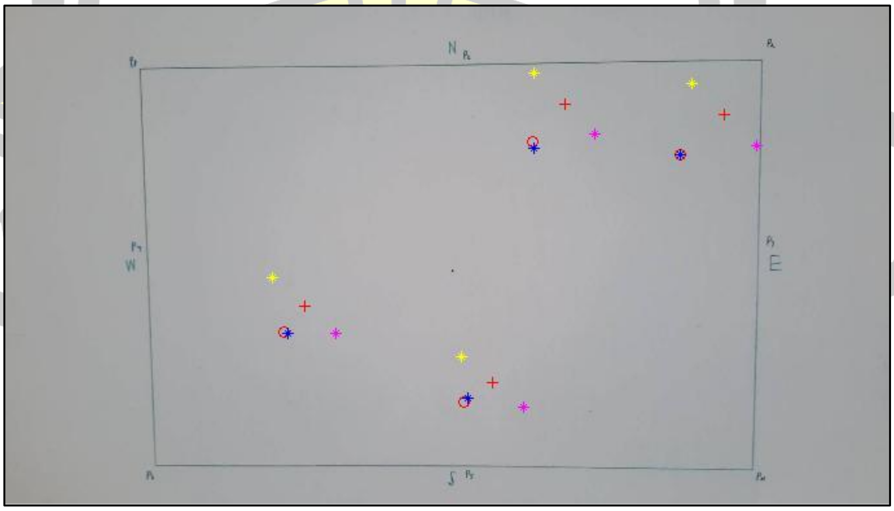
ภาพที่ 6.35 การผสมข้อมูลจากกล้อง C1,C5

ตามภาพที่ 6.35 เป็นตัวอย่างการผสมข้อมูลตำแหน่งจำนวน 4 วัตถุ จากกล้อง C1 และ C5 โดยตำแหน่งเครื่องหมายดอกจันสีเหลืองได้จากกล้อง C1 ตำแหน่งเครื่องหมายดอกจันสีชมพูได้จากกล้อง C5 เครื่องหมายบวกสีแดงเป็นผลจากการผสมข้อมูลโดยวิธีการหาค่าเฉลี่ยแบบเดิมและวงกลมสีแดงเป็นผลจากการผสมข้อมูลโดยใช้วิธีที่เสนอ สุดท้ายตำแหน่งเครื่องหมายดอกจันสีน้ำเงิน คือตำแหน่งจริงของวัตถุซึ่งได้จากมุมมองด้านบนจากกล้อง C9 ซึ่งถูกใช้เป็นกราวนด์ทรูธโดยจะสังเกตเห็นว่ากรณีนี้กล้อง C1 และ C5 ทำมุมกันในลักษณะขนานกันวิธีที่เสนอจะใช้วิธีการหาค่าเฉลี่ยตามปกติซึ่งผลลัพธ์นั้นค่อนข้างใกล้เคียงกันกับวิธีเดิม



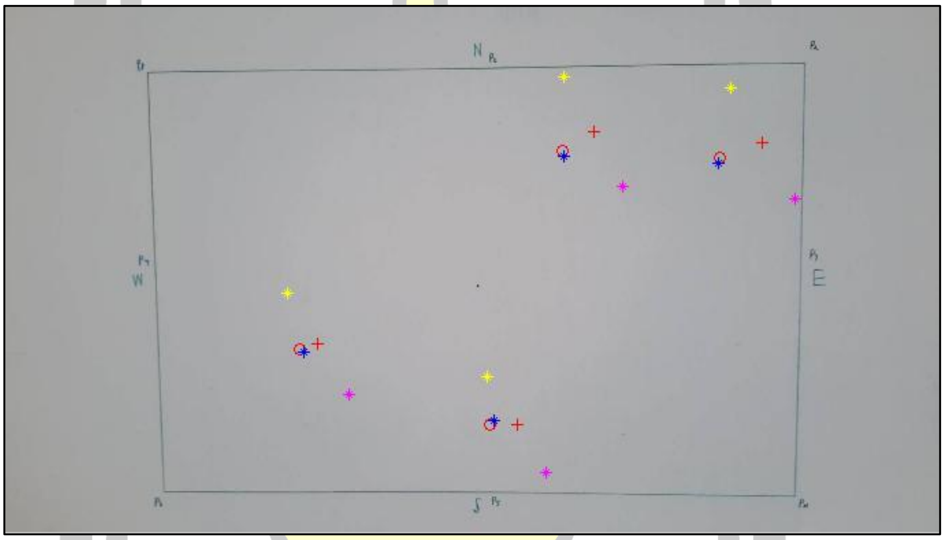
ภาพที่ 6.36 การผสมข้อมูลจากกล้อง C5,C6

ตามภาพที่ 6.36 เป็นตัวอย่างการผสมข้อมูลตำแหน่งจำนวน 4 วัตถุ จากกล้อง C5 และ C6 โดยตำแหน่งเครื่องหมายดอกจันสีเหลืองได้จากกล้อง C5 ตำแหน่งเครื่องหมายดอกจันสีชมพูได้จากกล้อง C6 เครื่องหมายบวกสีแดงเป็นผลจากการผสมข้อมูลโดยวิธีการหาค่าเฉลี่ยแบบเดิมและวงกลมสีแดงเป็นผลจากการผสมข้อมูลโดยใช้วิธีที่เสนอ สุดท้ายตำแหน่งเครื่องหมายดอกจันสีน้ำเงิน คือตำแหน่งจริงของวัตถุซึ่งได้จากมุมมองด้านบนจากกล้อง C9 ซึ่งถูกใช้เป็นกราวนด์ทรูธโดยจะสังเกตเห็นว่ากรณีนี้กล้อง C5 และ C6 ทำมุมกันในลักษณะเฉียงกันประมาณ 45 องศา ซึ่งผลลัพธ์ของวิธีเดิมนั้นจะห่างจากตำแหน่งจริงมากส่วนวิธีที่เสนออยู่ใกล้ตำแหน่งจริงมากกว่า



ภาพที่ 6.37 การผสมข้อมูลจากกล้อง C5,C7

ตามภาพที่ 6.37 เป็นตัวอย่างการผสมข้อมูลตำแหน่งจำนวน 4 วัตถุ จากกล้อง C5 และ C7 โดยตำแหน่งเครื่องหมายดอกจันสีเหลืองได้จากกล้อง C5 ตำแหน่งเครื่องหมายดอกจันสีชมพูได้จากกล้อง C7 เครื่องหมายบวกสีแดงเป็นผลจากการผสมข้อมูลโดยวิธีการหาค่าเฉลี่ยแบบเดิมและวงกลมสีแดงเป็นผลจากการผสมข้อมูลโดยใช้วิธีที่เสนอ สุดท้ายตำแหน่งเครื่องหมายดอกจันสีน้ำเงิน คือ ตำแหน่งจริงของวัตถุซึ่งได้จากมุมมองด้านบนจากกล้อง C9 ซึ่งถูกใช้เป็นกราวนด์ทรูธโดยจะสังเกตเห็นว่ากรณีนี้กล้อง C5 และ C7 ทำมุมกันในลักษณะเฉียงกันประมาณ 90 องศา ซึ่งผลลัพธ์ของวิธีเดิมจะห่างจากตำแหน่งจริงมากส่วนวิธีที่เสนออยู่ใกล้ตำแหน่งจริงมากกว่า



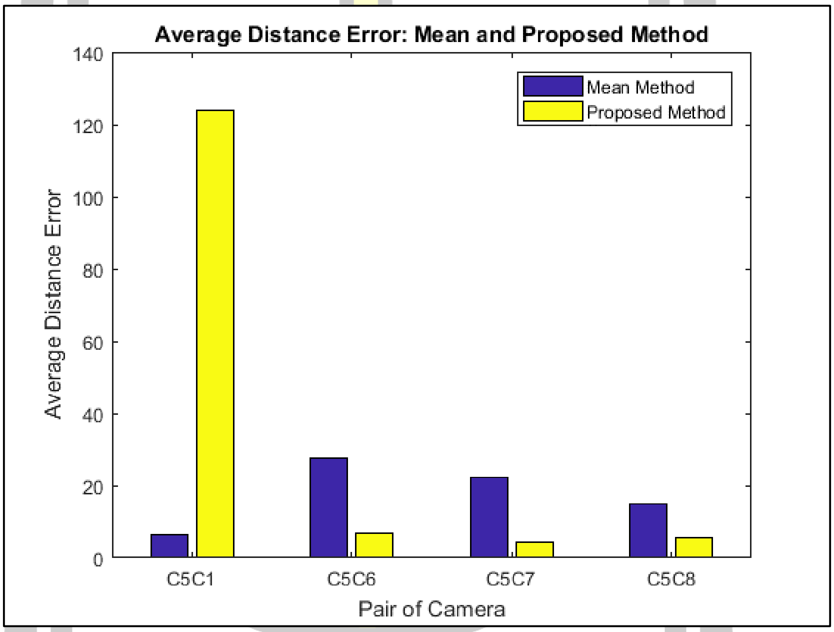
ภาพที่ 6.38 การผสมข้อมูลจากกล้อง C5,C8

ตามภาพที่ 6.38 เป็นตัวอย่างการผสมข้อมูลตำแหน่งจำนวน 4 วัตถุ จากกล้อง C5 และ C8 โดยตำแหน่งเครื่องหมายดอกจันสีเหลืองได้จากกล้อง C5 ตำแหน่งเครื่องหมายดอกจันสีชมพูได้จากกล้อง C8 เครื่องหมายบวกสีแดงเป็นผลจากการผสมข้อมูลโดยวิธีการหาค่าเฉลี่ยและวงกลมสีแดงเป็นผลจากการผสมข้อมูลโดยใช้วิธีที่เสนอ สุดท้ายตำแหน่งเครื่องหมายดอกจันสีน้ำเงิน คือ ตำแหน่งจริงของวัตถุซึ่งได้จากมุมมองด้านบนจากกล้อง C9 ซึ่งถูกใช้เป็นกราวนด์ทรูธโดยจะสังเกตเห็นว่ากรณีนี้กล้อง C5 และ C8 ทำมุมกันในลักษณะเฉียงกันมากกว่า 90 องศา ซึ่งผลลัพธ์ของวิธีเดิมจะห่างจากตำแหน่งจริงมากส่วนวิธีที่เสนออยู่ใกล้ตำแหน่งจริงมากกว่า

ในการทดลองครั้งนี้ได้ทดลองการผสมข้อมูลระหว่าง 2 กล้องที่ทำมุมกันในลักษณะมุมมองต่างๆ โดยข้อมูลภาพจาก 2 กล้องมีแนวเส้นตรงของมุมมองที่ตัดกันที่องศาต่างๆ เช่น ขนานกัน (กล้อง C1-C5) ตัดกันที่ 45 องศา (กล้อง C5-C6), 90 องศา (กล้อง C5-C7) และมากกว่า 90 องศา

(กล้อง C5-C8) เป็นต้น ซึ่งผลการทดลองแสดงให้เห็นว่าโดยส่วนใหญ่แม้แต่ละกล้องจะทำมุมกันในลักษณะใดระบบที่เสนอนั้นยังสามารถประมาณตำแหน่งผู้เล่นได้ใกล้เคียงกับตำแหน่งจริงมากกว่าวิธีเดิมที่ใช้ค่าเฉลี่ยของระยะทางเพียงอย่างเดียว

จากการทดลองระหว่าง 2 วิธี คือ วิธีใช้ค่าเฉลี่ยกับวิธีใหม่ที่เสนอโดยใช้ข้อมูลทดสอบจำนวน 4 วัตถุ ทำการทดลองผสมข้อมูลระหว่าง 2 กล้องใดๆ พบว่าประสิทธิภาพในการประมาณตำแหน่งนั้นค่อนข้างแตกต่างกันแสดงดังภาพที่ 6.39



ภาพที่ 6.39 Average Distance Error

ตามภาพที่ 6.39 เป็นผลการวัดประสิทธิภาพในด้านความผิดพลาดของค่าระยะทางเฉลี่ย (ADE: Average Distance Error) โดยหลังจากโปรเจ็คตำแหน่งของวัตถุทั้ง 4 ตำแหน่งลงบนภาพเสมือนแล้วจะทำการวัดค่าระยะทางระหว่างตำแหน่งที่ประมาณกับตำแหน่งกราวนด์ทรูธที่ได้จากกล้อง C9 ซึ่งจะได้เป็นค่าของระยะทางที่ผิดพลาด จากนั้นทำการหาค่าเฉลี่ยจากรายการทั้งหมดในภาพรวม ผลการประเมินพบว่าในกรณีของการผสมข้อมูลระหว่างกล้อง C1 กับ C5 วิธีที่เสนอมีค่าความผิดพลาดที่สูงกว่าแบบเดิมที่ใช้วิธีหาค่าเฉลี่ย ส่วนการผสมข้อมูลระหว่าง 2 กล้องในกรณีที่เหลือ คือ กล้อง C5 กับ C6 กล้อง C5 กับ C7 และกล้อง C5 กับ C8 พบว่าวิธีที่เสนอให้ผลที่ดีกว่าวิธีเดิมที่ใช้ค่าเฉลี่ย

6.3 สรุปผลการดำเนินงาน (Summary Operating Results)

ในงานนี้ได้นำเสนอเทคนิคการผสมข้อมูลในมุมมองเสมือนซึ่งได้มาจากข้อมูลของแต่ละกล้องโดยใช้เทคนิคที่เรียกว่า โมเดลรูปร่างของข้อมูลทางกายภาพ (Physical based Data Shape Model) สำหรับการจัดการโครงสร้างของข้อมูลของแต่ละกล้องให้อยู่ในรูปแบบเดียวกันซึ่งข้อดีของเทคนิคที่เสนอนี้ คือ จะช่วยให้การผสมข้อมูลนั้นง่ายยิ่งขึ้น ใช้ข้อมูลในการประมวลผลที่น้อย การดำเนินการไม่ซับซ้อน ซึ่งเป็นเทคนิคที่ใช้วิธีการจัดเรียงข้อมูลแบบง่ายโดยอาศัยข้อมูลพื้นฐานทางกายภาพของตัวภาพในมุมมองเสมือนเป็นเกณฑ์

นอกจากนี้ในงานนี้ยังได้เสนอเทคนิคการประมาณตำแหน่งบนมุมมองเสมือนจากข้อมูลในแต่ละกล้องโดยเสนอเทคนิคที่เรียกว่า การประมาณตำแหน่งบนพื้นฐานของสมการเส้นตรง (Straight Line Equation based Position Evaluation) ซึ่งมีข้อดี คือ เป็นวิธีที่ง่าย ไม่ซับซ้อนและได้ผลการประมาณที่ใกล้เคียงกับตำแหน่งจริงมากกว่าวิธีการหาค่าเฉลี่ยแบบเดิม ซึ่งมีแนวคิดพื้นฐานมากจากการตกกระทบของเงา เนื่องจากโดยธรรมชาติตำแหน่งของเงาจะอยู่ตรงกันข้ามกับดวงอาทิตย์เสมอ ดังนั้นตำแหน่งของกล้องจึงควรจะอยู่ตรงกันข้ามกับวัตถุภายในภาพเช่นกัน สำหรับการหาตำแหน่งและทิศทางของกล้องจากข้อมูลพื้นฐานทางกายภาพของภาพที่พิจารณานั้นในงานนี้เสนอวิธีการหาทิศทางของกล้องโดยอาศัยแนวคิดจากหลักการทำงานของเข็มทิศซึ่งแกนหมุนของเข็มทิศจะเปรียบได้กับตำแหน่งวัตถุ ปลายเข็มทิศที่ชี้ไปยังทิศเหนือเปรียบได้กับตำแหน่งตรงกันข้ามกับกล้องหรือตัวของภาพและปลายของเข็มทิศที่ชี้ไปยังทิศใต้จะเป็นตำแหน่งและทิศทางของกล้องที่ต้องการค้นหา โดยในงานนี้ได้เสนออัลกอริทึมสำหรับการหาทิศทางของกล้องที่เรียกว่า การหาทิศทางของกล้องด้วยเข็มทิศ (Camera’s Direction Finding based on Compass (CDFC)) โดยเป็นอัลกอริทึมที่ง่าย ไม่ซับซ้อน ผลจากอัลกอริทึมจะได้ข้อมูลตำแหน่งของวัตถุเพิ่มขึ้นมาอีกหนึ่งตำแหน่งซึ่งจะถูกใช้สำหรับการสร้างเส้นตรงสำหรับกระบวนการประมาณตำแหน่งผู้เล่นในมุมมองเสมือน

การใช้เทคนิคของโมเดลรูปร่างข้อมูลทางกายภาพร่วมกับเทคนิคการประมาณตำแหน่งบนพื้นฐานของสมการเส้นตรงยังส่งผลให้ระบบมีความแข็งแรงต่อปัจจัยหลักทั้ง 3 ปัจจัยที่ส่งผลต่อกระบวนการผสมข้อมูลซึ่งได้แก่ 1. กระบวนการตรวจจับวัตถุ 2. วิธีกำหนดตำแหน่งวัตถุที่เลือกใช้ และ 3. ตำแหน่งและทิศทางของกล้อง ซึ่งในงานนี้สันนิษฐานว่าเทคนิคที่เสนอนี้ทำให้ระบบสามารถรองรับกับการทำงานในการบันทึกภาพการแข่งขันกีฬาและการจัดการข้อมูลจากหลายกล้องได้สะดวกและง่ายยิ่งขึ้นซึ่งระบบที่พัฒนานั้นจะอาศัยข้อมูลพื้นฐานของกล้องน้อยลง สามารถลดกระบวนการสอบเทียบกล้องได้และสามารถรองรับกับเงื่อนไขต่างๆ ทางกายภาพของกล้อง เช่น ตำแหน่งกล้อง มุมเอียงของกล้อง ความสูงของกล้อง รวมถึงการใช้กล้องต่างยี่ห้อ ต่างรุ่นกันในการบันทึกภาพ เป็นต้น ทำให้สามารถลดขั้นตอนการจัดเตรียมและติดตั้งกล้องในการบันทึกภาพโดยสามารถบันทึกภาพได้ทันทีเพียงแต่ผู้ใช้ทำการกำหนดจุดคอนโทรลพอยต์เท่านั้นโดยไม่ต้องเสียเวลาในการคำนวณ

ตำแหน่ง ความสูงและค่าพื้นฐานของกล้อง นอกจากนี้ผลจากอัลกอริทึมการหาทิศทางของกล้องนั้นยังทราบถึงทิศทางที่กล้องนั้นอยู่ด้วยซึ่งข้อมูลพื้นฐานนี้สามารถนำไปประยุกต์ใช้ในงานด้านอื่นๆ ได้อีกด้วย เช่น การหาแผนผังของตำแหน่งและทิศทางของกล้องทั้งหมดรอบสนาม เป็นต้น

กระบวนการทดลองสำหรับการผสมข้อมูล ในงานนี้ทำการทดลองโดยใช้ข้อมูลตำแหน่งวัตถุจำนวน 4 วัตถุ จากข้อมูลกล้องจำนวน 9 ตัว โดยแต่ละกล้องที่นำมาผสมข้อมูลกันนั้นจะอยู่ในตำแหน่งและทิศทางที่แตกต่างกัน ซึ่งจะมีแนวเส้นตรงของมุมมองที่ตัดกันที่องศาต่างๆ เช่น ขนานกัน ตัดกันที่ 45, 90 และมากกว่า 90 องศา เป็นต้น ในการทดลองนี้ได้ทำการบันทึกภาพแบบง่ายโดยไม่มีการกำหนดตำแหน่งกล้อง ความสูงของกล้องและมุมเอียงของกล้อง แต่ใช้วิธีการประมาณด้วยมือ จุดประสงค์เพื่อการหาข้อพิสูจน์ว่าวิธีที่เสนอนี้สามารถตอบโจทย์ของข้อสันนิษฐานที่ว่า ระบบสามารถทำการแปลงมุมมองของข้อมูลและทำการผสมข้อมูลได้โดยไม่ต้องอาศัยข้อมูลพื้นฐานของกล้องเหล่านี้ได้หรือไม่ จากผลการทดลองแสดงให้เห็นว่าแม้ว่าแต่ละกล้องจะทำมุมกันในลักษณะใด วิธีการเลือกตำแหน่งวัตถุจะใช้แบบไหน เช่น ใช้ตำแหน่งเท้า ตำแหน่งศูนย์กลางบล็อกและตำแหน่งศีรษะของวัตถุ และผลการตรวจจับวัตถุในบางกล้องอาจจะไม่สมบูรณ์ ระบบที่เสนอนั้นยังสามารถทำการผสมข้อมูลและประมาณตำแหน่งผู้เล่นได้ใกล้เคียงกับตำแหน่งจริง โดยปัญหาที่ยังพบของวิธีที่เสนอ คือ ในกรณีที่ 2 กล้องทำมุมกันในลักษณะขนานกันอาจจะส่งผลให้เส้นตรงที่ได้จากกระบวนการหาทิศทางกล้องมีลักษณะที่ขนานกันหรือเกือบจะขนานกันส่งผลให้ตำแหน่งที่ได้จากการประมาณจะอยู่ห่างจากตำแหน่งจริง ในงานนี้แก้ปัญหาในเบื้องต้นโดยถ้าเป็นกรณีที่ทั้ง 2 เส้นตรงนั้นขนานกันจะใช้วิธีการประมาณตำแหน่งจากค่าเฉลี่ย ส่วนกรณีที่ทั้ง 2 เส้นตรงอยู่ในลักษณะที่เกือบขนานกันจะแก้ไขโดยกำหนดค่าเงื่อนไขขององศาระหว่าง 2 เส้นตรง โดยถ้ามีค่าองศาต่ำกว่าเกณฑ์ที่กำหนดจะประมาณตำแหน่งโดยใช้ค่าเฉลี่ยแต่ถ้ามากกว่าเกณฑ์ที่กำหนดจะใช้วิธีที่เสนอ ผลจากการปรับปรุงพบว่าให้ผลที่ดีกว่าเดิมแต่ยังให้ผลที่แย่กว่าวิธีเดิมที่ใช้ค่าเฉลี่ย

บทที่ 7

สรุปผลและอภิปรายผล (Conclusions and Discussions)

7.1 สรุปผล (Conclusions)

การดำเนินงานวิจัยในครั้งนี้มีจุดประสงค์เพื่อศึกษาและพัฒนาวิธีการแก้ไขปัญหาการบดบังของนักฟุตบอลในวิดีโอการแข่งขันกีฬาฟุตบอลโดยมีกระบวนการดำเนินงานอยู่ 3 กระบวนการหลัก ได้แก่ 1. กระบวนการตรวจจับสนาม 2. กระบวนการตรวจจับและติดตามผู้เล่นในมุมมองภาพ และ 3. กระบวนการติดตามผู้เล่นในมุมมองเสมือนและการแก้ไขปัญหาการบดบังของผู้เล่น สำหรับกระบวนการตรวจจับสนามนั้นในงานนี้ได้ประยุกต์ใช้เทคนิคโมเดลผสมแบบเกาเซียนร่วมกับหลักการความน่าจะเป็นของเบย์โดยระบบจะทำการเรียนรู้ลักษณะของสนามจากข้อมูลชุดสอนซึ่งเป็นชุดภาพที่มีสนามปรากฏอยู่ในลักษณะต่างๆ เพื่อใช้เป็นความรู้พื้นฐานในการตรวจจับสนามในชุดภาพทดสอบ สำหรับกระบวนการตรวจจับและติดตามผู้เล่นในมุมมองภาพ ได้ประยุกต์ใช้เทคนิคของแบ็คกราวนด์ซับแทรคชันสำหรับการตรวจจับบล็อบของผู้เล่น โดยในการพิจารณาบล็อบที่เป็นได้ว่าจะเป็นผู้เล่นนั้นจะพิจารณาจากรายการบล็อบสองส่วน คือ รายการบล็อบที่ตรวจจับได้จากขั้นตอนการตรวจจับสนามและรายการบล็อบที่ตรวจจับได้จากขั้นตอนการตรวจจับผู้เล่นด้วยเทคนิคแบ็คกราวนด์ซับแทรคชันโดยเอกลักษณ์ที่ใช้ คือ ขนาด สีและการเคลื่อนที่ โดยจะพิจารณาเฉพาะวัตถุที่อยู่ในขอบเขตของสนามเท่านั้น ส่วนการติดตามผู้เล่นนั้นประยุกต์ใช้เทคนิคของอัลกอริทึมคอนเดนเซชันซึ่งมีพื้นฐานอยู่บนเอกลักษณ์สี ร่วมกับเทคนิคการสุ่มที่เสนอคือการสุ่มแซมเปิลบนพื้นฐานของพื้นที่ สำหรับกระบวนการติดตามผู้เล่นในมุมมองเสมือน ในงานนี้ใช้วิธีการแปลงมุมมองพื้นฐานทั่วไป คือ ใช้หลักการของเพอสเปคทีพทรานส์ฟอร์ม ส่วนวิธีการผสมข้อมูลนั้นในงานนี้ได้เสนอเทคนิคการผสมข้อมูลโดยอาศัยพื้นฐานความรู้ของโครงสร้างข้อมูลทางกายภาพ โดยเสนอโมเดลรูปร่างของข้อมูลบนพื้นฐานทางกายภาพสำหรับการจัดการข้อมูลจากกล้องต่างๆ ให้อยู่ในรูปแบบเดียวกันและเสนอเทคนิคการประมาณตำแหน่งบนพื้นฐานของสมการเส้นตรง โดยการโปรเจ็คตำแหน่งผู้เล่นจากมุมมองภาพมายังมุมมองเสมือนนั้นจะมีเงื่อนไขอยู่ว่าผู้เล่นนั้นต้องอยู่ในพื้นที่ของ FOV เดียวกันจึงจะทำการผสมข้อมูลได้ในกรณีที่ผู้เล่นไม่อยู่ในพื้นที่ FOV ที่ซ้อนทับกับพื้นที่ใดจะทำการโปรเจ็คข้อมูลตำแหน่งนั้นตามปกติโดยไม่ผ่านขั้นตอนการผสมข้อมูล และสุดท้ายการแก้ไขปัญหาการบดบังซึ่งเกิดจากการรวมตัวของผู้เล่น ในงานนี้ได้เสนอเทคนิคต่างๆ สำหรับแก้ปัญหาเริ่มตั้งแต่กระบวนการตรวจจับและติดตามผู้เล่นในมุมมองภาพซึ่งได้ประยุกต์ใช้อัลกอริทึมคอนเดนเซชันซึ่งเป็นเทคนิคหนึ่งที่ช่วยแก้ปัญหาการบดบังแบบบางส่วนได้ในเบื้องต้น ในกรณีที่ผู้เล่นถูกบดบังแบบสมบูรณ์ในแต่ละกล้องจะใช้วิธีการทำนายแบบพื้นฐานร่วมกับการใช้ข้อมูลที่ได้จากกระบวนการผสมข้อมูลเข้าช่วยแก้ปัญหา

สำหรับผู้เล่นที่ถูกบดบังจากกล้องหนึ่งโดยใช้ข้อมูลที่ได้จากกล้องตัวอื่นมาช่วยในการประมาณตำแหน่งผู้เล่นคนดังกล่าว และสุดท้ายในกรณีเกิดการรวมตัวของผู้เล่นจำนวนมากซึ่งอาจส่งผลให้มีผู้เล่นบางคนไม่สามารถมองเห็นได้จากกล้องใดๆ จะใช้ข้อมูลที่ได้จากขั้นตอนการผสมข้อมูลซึ่งใช้ข้อมูลอินพุตจากผลการทำนายในแต่ละกล้องร่วมกับผลการทำนายเส้นทางในมุมมองเสมือนในการทำนายตำแหน่งผู้เล่นที่ถูกบดบังซึ่งระบบจะยังสามารถทำการติดตามผู้เล่นต่อไปได้

7.2 อภิปรายผล (Discussions)

ในการดำเนินงานและจากผลการทดลองสามารถสรุปแยกตามแต่ละกระบวนการได้ดังนี้ กระบวนการตรวจจับสนาม พบว่าระบบสามารถตรวจจับพื้นที่ที่เป็นสนามภายในภาพทดสอบได้ถูกต้องแต่ในบางกรณีของภาพทดสอบพบว่าจะมีบางส่วนของพิกเซลภาพที่ระบบพิจารณาไม่ถูกต้องโดยเฉพาะบริเวณขอบสนามซึ่งจะเกิดการผิดพลาดในการตรวจจับโดยมีบางพิกเซลที่ถูกพิจารณาว่าเป็นสนามทั้งที่ในความเป็นจริงนั้นไม่ใช่พิกเซลสนามและในทางกลับกันพิกเซลที่เป็นสนามถูกพิจารณาว่าไม่ใช่สนาม จากปัญหานี้ในงานวิจัยนี้จึงได้ประยุกต์ใช้เทคนิคของมาร์คอฟแรนดอมฟิวส์และเทคนิคที่เรียกว่า มีเดียนฟิลเตอร์มาทดลองใช้ในการปรับปรุงประสิทธิภาพของการจำแนกข้อมูลของระบบซึ่งพบว่าสามารถช่วยเพิ่มความถูกต้องในการจำแนกข้อมูลได้ดีขึ้นและนอกจากนี้ยังพบว่าวิธีของมาร์คอฟแรนดอมฟิวส์จะให้ผลที่ดีกว่าวิธีของมีเดียนฟิลเตอร์

สำหรับกระบวนการตรวจจับและติดตามผู้เล่นในมุมมองภาพ ในขั้นตอนการติดตามในงานนี้ได้ทดลองใช้ 3 วิธี คือ 1. คาลมานฟิลเตอร์ 2. อัลกอริทึมคอนเดนเซชัน และ 3. อัลกอริทึมคอนเดนเซชันบนพื้นฐานของสีร่วมกับวิธีการสุ่มแซมเปิลบนพื้นฐานของพื้นที่ ซึ่งจากผลการทดลองพบว่าค่าเฉลี่ยรวมของพารามิเตอร์ต่างๆ ประกอบด้วย ค่าของระยะทาง วิธีที่ 2 จะให้ผลที่ดีกว่าวิธีแรกและวิธีที่เสนอ ค่าของ Absolute Error (AE) วิธีที่เสนอให้ผลที่ดีกว่า 2 วิธีเดิม ค่าของ Bounding Box Error (BBE) พบว่าวิธีที่เสนอให้ผลที่เข้าใกล้ค่า 1 มากกว่า 2 วิธีเดิม ค่าของ Object Tracking Error (OTE) และ Root Mean Square Error (RMSE) วิธีที่เสนอให้ผลที่ดีกว่า 2 วิธีเดิม ซึ่งจากผลการทดลองสรุปได้ว่าการใช้เอกลักษณ์สีและเทคนิคการสุ่มบนพื้นฐานของพื้นที่ให้ผลที่ดีกว่ากรณีไม่ใช้เอกลักษณ์สีและเทคนิคการสุ่มบนพื้นฐานของบราวดิ้งบล็อก นอกจากนี้เวลาในการประมวณผลในภาพรวมนั้นพบว่าวิธีที่เสนอใช้เวลาน้อยกว่าวิธีแรกแต่มากกว่าวิธีที่ 2 เล็กน้อย

กระบวนการติดตามผู้เล่นในมุมมองเสมือน ในงานนี้ได้นำเสนอเทคนิคในการผสมข้อมูลในมุมมองเสมือนซึ่งได้ข้อมูลตำแหน่งมาจากข้อมูลของแต่ละกล้อง โดยใช้เทคนิคที่เรียกว่า โมเดลรูปร่างของข้อมูลทางกายภาพ ซึ่งข้อดีของเทคนิคที่เสนอนี้ คือ ช่วยให้การผสมข้อมูลนั้นง่ายยิ่งขึ้น ใช้ข้อมูลในการประมวลผลที่น้อย การดำเนินการไม่ซับซ้อน ซึ่งเป็นเทคนิคที่ใช้วิธีการจัดเรียงข้อมูลแบบง่าย นอกจากนี้ยังได้เสนอเทคนิคการประมาณตำแหน่งบนมุมมองเสมือนจากข้อมูลในแต่ละกล้อง โดย

เสนอเทคนิคที่เรียกว่า การประมาณตำแหน่งบนพื้นฐานของสมการเส้นตรง ซึ่งมีข้อดี คือ เป็นวิธีที่ง่าย ไม่ซับซ้อนและได้ผลการประมาณที่ใกล้เคียงกับตำแหน่งจริงมากกว่าวิธีการหาค่าเฉลี่ยแบบเดิม กระบวนการทดลองสำหรับการผสมข้อมูล ในงานนี้ทำการทดลองโดยใช้ข้อมูลตำแหน่งวัตถุจำนวน 4 วัตถุ จากข้อมูลกล้องจำนวน 9 ตัว โดยแต่ละกล้องที่นำมาผสมข้อมูลกันนั้นจะอยู่ในตำแหน่งและทิศทางที่แตกต่างกัน ซึ่งจะมีแนวเส้นตรงของมุมมองที่ตัดกันที่องศาต่างๆ เช่น ขนานกัน ตัดกันที่ 45, 90 และมากกว่า 90 องศา เป็นต้น จากผลการทดลองแสดงให้เห็นว่าแม้ว่าแต่ละกล้องจะทำมุมกันในลักษณะใด วิธีการเลือกตำแหน่งวัตถุจะใช้แบบไหน เช่น ใช้ตำแหน่งเท้า ตำแหน่งศูนย์กลางบล็อกและตำแหน่งศีรษะและผลของการตรวจจับวัตถุในบางกล้องจะคลาดเคลื่อน แต่ผลในภาพรวมนั้นระบบที่เสนอยังสามารถทำการผสมข้อมูลและประมาณตำแหน่งผู้เล่นได้ใกล้เคียงกับตำแหน่งจริง

สรุปกระบวนการแก้ไขปัญหาการบดบังซึ่งเกิดจากการรวมตัวของผู้เล่นฟุตบอล ในงานนี้ประกอบด้วย การประยุกต์ใช้อัลกอริทึมคอนเดนเซชันในการติดตามผู้เล่นในมุมมองภาพซึ่งสามารถช่วยแก้ปัญหาการบดบังแบบบางส่วนได้ ขั้นต่อมา คือ เสนอเทคนิคการผสมข้อมูลจากแต่ละกล้องเป็นข้อมูลชุดเดียวและแสดงผลบนภาพเสมือนซึ่งกระบวนการผสมข้อมูลนี้สามารถช่วยแก้ปัญหาผู้เล่นที่ถูกบดบังซึ่งไม่สามารถมองเห็นจากกล้องใดๆ ได้โดยอาศัยข้อมูลจากกล้องตัวอื่นที่ยังสามารถมองเห็นได้มาช่วยในการประมาณตำแหน่งและในกรณีเกิดการรวมกลุ่มของผู้เล่นเป็นจำนวนมากจนมีผู้เล่นบางคนไม่สามารถมองเห็นได้จากทุกกล้อง ระบบยังสามารถใช้ผลการประมาณตำแหน่งที่ได้จากการผสมข้อมูลของผลการทำนายจากทุกกล้องร่วมกับผลการทำนายเส้นทางในมุมมองเสมือนในการทำนายตำแหน่งผู้เล่นที่ถูกบดบังได้ส่งผลให้ระบบยังสามารถประมาณตำแหน่งและติดตามการเคลื่อนที่ของผู้เล่นที่ถูกบดบังภายในกลุ่มได้

7.3 เทคนิคที่เสนอในงานวิจัย (Techniques Proposed in the Research)

ในการดำเนินงานวิจัยในครั้งนี้ได้นำเสนอเทคนิคต่างๆ สำหรับกระบวนการตรวจจับและติดตามผู้เล่นในวิดีโอการแข่งขันฟุตบอลประกอบด้วย 5 เทคนิค ดังนี้

1. เทคนิคการตรวจจับสนามบนพื้นฐานของความน่าจะเป็นร่วมกับเทคนิคการปรับปรุงข้อมูลโดยใช้เทคนิคของมาร์คอฟแรนดอมฟิวส์ ซึ่งข้อดี คือ กระบวนการเรียนรู้บนพื้นฐานของความน่าจะเป็นสามารถดำเนินการได้ง่ายเป็นวิธีที่ไม่ซับซ้อน นอกจากนี้การประยุกต์ใช้เทคนิคมาร์คอฟแรนดอมฟิวส์ยังสามารถช่วยเพิ่มประสิทธิภาพของกระบวนการจำแนกพิกเซลสนามและพิกเซลที่ไม่ใช่สนามได้ดียิ่งขึ้น

2. การติดตามผู้เล่นโดยประยุกต์ใช้อัลกอริทึมคอนเดนเซชันร่วมกับเอกลักษณ์สีและเทคนิคการสุ่มบนพื้นฐานของพื้นที่ ซึ่งมีข้อดี คือ เอกลักษณ์สีและการสุ่มบนพื้นฐานของพื้นที่ช่วยให้กระบวนการตรวจจับและติดตามมีความถูกต้องมากขึ้น แซมเปิลที่สุ่มได้จะเป็นแซมเปิลจริงของวัตถุ

มากกว่าวิธีของอัลกอริทึมคอนเดนเซชันเดิมซึ่งสันนิษฐานว่าเป็นการสุ่มบนพื้นฐานของบราวดิ้งบล็อก ซึ่งจะได้แซมเปิลที่เป็นพื้นหลังผสมมาด้วย ดังนั้นแซมเปิลที่ได้นี้จะสามารถใช้สำหรับการจำแนกตัวตนของผู้เล่นได้ถูกต้องมากขึ้น

3. เทคนิคการผสมข้อมูลในมุมมองเสมือน โดยใช้โมเดลรูปร่างของข้อมูลทางกายภาพ ซึ่งข้อดีของเทคนิคที่เสนอนี้ คือ ช่วยให้การผสมข้อมูลนั้นง่ายยิ่งขึ้น ใช้ข้อมูลในการประมวลผลที่น้อย การดำเนินการไม่ซับซ้อน เป็นเทคนิคที่ใช้วิธีการจัดเรียงข้อมูลแบบง่าย

4. เทคนิคการประมาณตำแหน่งบนมุมมองเสมือนที่เรียกว่า การประมาณตำแหน่งบนพื้นฐานของสมการเส้นตรง ซึ่งมีข้อดี คือ เป็นวิธีที่ง่าย ไม่ซับซ้อนและได้ผลการประมาณที่ใกล้เคียงกับตำแหน่งจริงมากกว่าวิธีการหาค่าเฉลี่ยแบบเดิม นอกจากนี้ยังมีความแข็งแรงต่อ 3 ปัจจัยหลักที่มีผลต่อความถูกต้องในการประมาณตำแหน่งวัตถุ ได้แก่ รูปแบบของการกำหนดตำแหน่งวัตถุ ทิศทางของกล้องและผลจากการตรวจจับวัตถุ

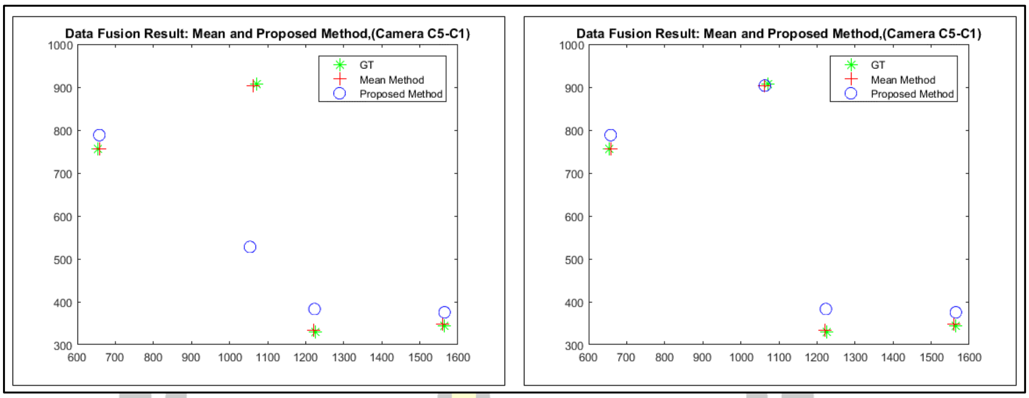
5. เทคนิคการค้นหาทิศทางของกล้องที่เสนอ สามารถนำไปประยุกต์ใช้ในงานด้านอื่นๆ ได้ เช่น การกำหนดแผนผังของตำแหน่งกล้องรอบบริเวณสนามฟุตบอล เป็นต้น

7.4 ปัญหาที่พบ (Problems Encountered)

จากผลการดำเนินงานในกระบวนการผสมข้อมูลพบ 3 ประเด็นปัญหาดังนี้

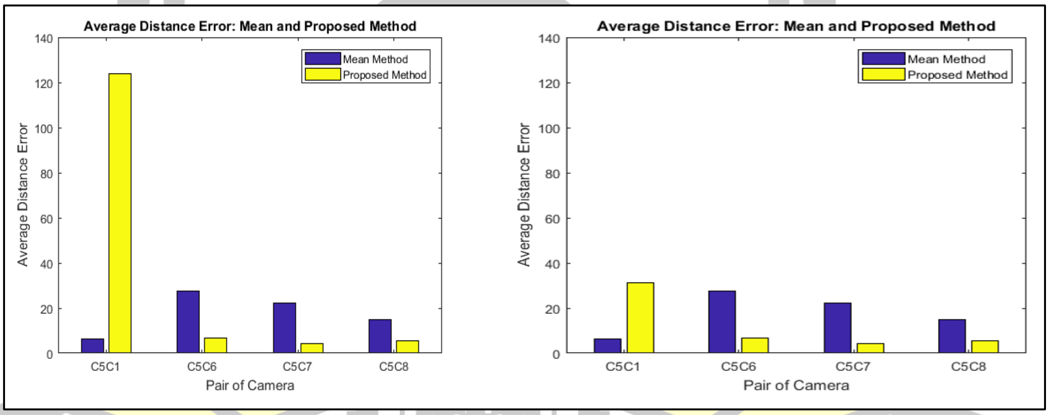
1. ในกระบวนการหาทิศทางของกล้อง อาจเกิดข้อผิดพลาดได้ในกรณีที่ทำการสุ่มเลือกตำแหน่งสุ่ม (Sampling_Position) ซึ่งมีตำแหน่งที่เกินขอบเขตของภาพและอาจส่งผลให้ไม่ปรากฏตำแหน่งนี้ในมุมมองเสมือนด้วยทำให้กระบวนการผสมข้อมูลไม่ถูกต้อง ดังนั้นต้องแน่ใจว่าตำแหน่งดังกล่าวนี้ไม่อยู่นอกขอบเขตของภาพเสมือน

2. ในการประมาณตำแหน่งในมุมมองเสมือน ข้อมูลจากกระบวนการผสมข้อมูลระหว่าง 2 กล้องใดๆ จากการทดลองพบว่าในกรณีที่ 2 กล้องอยู่ในลักษณะที่เกือบจะขนานกัน คือ มีองศาของมุมระหว่าง 2 เส้นตรงที่น้อยจะส่งผลให้การประมาณตำแหน่งของวิธีที่เสนอค่อนข้างแตกต่างจากตำแหน่งจริงมาก โดยถ้ามีองศาระหว่าง 2 เส้นตรงที่น้อยซึ่งมีลักษณะที่ขนานกัน ผลการประมาณนั้นจะอยู่ห่างจากตำแหน่งจริงมากขึ้น ในงานวิจัยนี้แก้ปัญหาในเบื้องต้นโดยกำหนดเงื่อนไขของขนาดองศาระหว่างเส้นตรง 2 เส้นอยู่ที่ 0.1 กรณีค่าที่ได้น้อยกว่าค่าที่กำหนดนี้จะใช้วิธีการประมาณโดยใช้ค่าเฉลี่ย และกรณีมีค่ามากกว่าค่าที่กำหนดนี้จะใช้วิธีที่เสนอ ซึ่งผลการทดลองให้ผลดีกว่ากรณีการไม่กำหนดเงื่อนไขแต่ยังให้ผลที่แย่กว่าแบบเดิมที่ประมาณโดยใช้ค่าเฉลี่ย ตัวอย่างผลการปรับปรุงประสิทธิภาพแสดงดังภาพที่ 7.1 และ ภาพที่ 7.2 ดังนี้



ภาพที่ 7.1 ผลการประมาณตำแหน่งกล้อง C1 - C5 โดยการกำหนดเงื่อนไข

ตามภาพที่ 7.1 ผลจากการกำหนดเงื่อนไขขององศาระหว่าง 2 เส้นตรงในการผสมข้อมูลระหว่างกล้อง C1 กับ C5 พบว่าผลการประมาณโดยใช้วิธีกำหนดเงื่อนไขดังภาพที่ 7.1 ภาพด้านขวาให้ผลที่ดีขึ้นกว่ากรณีไม่มีการกำหนดเงื่อนไขดังภาพที่ 7.1 ภาพด้ายซ้าย



ภาพที่ 7.2 Average Distance Error หลังกำหนดเงื่อนไข

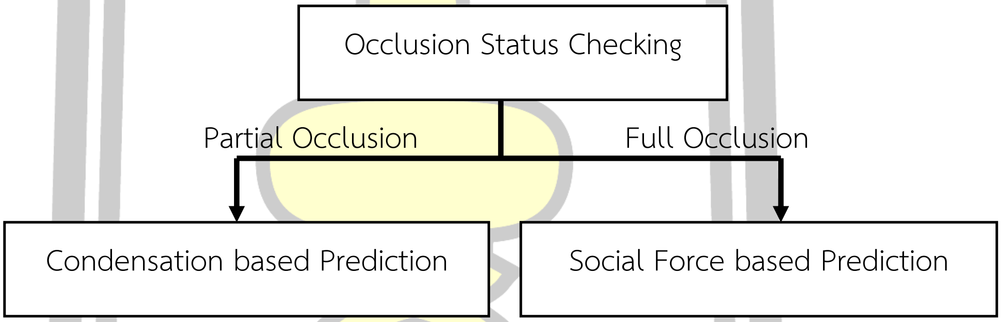
ตามภาพที่ 7.2 หลังจากการกำหนดเงื่อนไขขององศาระหว่าง 2 เส้นตรง พบว่าการผสมข้อมูลระหว่างกล้อง C1 กับ C5 ให้ผลในภาพรวมที่ดีขึ้นกว่ากรณีที่ไม่มีการกำหนดเงื่อนไขในการผสมข้อมูลซึ่งแสดงด้วยกราฟสีเหลืองแต่ยังให้ผลในภาพรวมที่แย่กว่าแบบเดิมที่ใช้ค่าเฉลี่ยซึ่งแสดงด้วยกราฟสีน้ำเงิน

3. กระบวนการผสมข้อมูล ข้อมูลตำแหน่งของผู้เล่นจากกล้องใดๆ ที่จะโปรเจ็คลงบนมุมมองเสมือนในพื้นที่ FOV ที่ซ้อนทับกันต้องมีจำนวนที่เท่ากันวิธีที่เสนอจึงจะสามารถทำการผสมข้อมูลได้ ซึ่งจากการทดลองกับชุดวิดีโอทดสอบของ ISSIA Dataset [82] พบว่าการผสมข้อมูลระหว่างกล้อง C1 กับ C2 มีปัญหาไม่สามารถทำการผสมข้อมูลได้ เนื่องจากทั้ง 2 กล้องมีลักษณะของของมุมมองที่ต่างกันค่อนข้างมากทำให้ผลการประมาณตำแหน่งที่ได้ของผู้เล่นที่อยู่บริเวณขอบของพื้นที่ FOV จะอยู่นอกพื้นที่ของ FOV ส่งผลให้จำนวนตำแหน่งผู้เล่นที่ได้ของทั้ง 2 กล้องในพื้นที่ FOV ที่ซ้อนทับกันไม่เท่ากัน ส่วนการผสมข้อมูลระหว่างกล้อง C4 และ C5 นั้นสามารถทำการผสมข้อมูลได้ เนื่องจากมีลักษณะของมุมมองกล้องที่ค่อนข้างใกล้เคียงกัน เป็นต้น

7.5 ข้อเสนอแนะ (Suggestions)

1. แนวทางการแก้ปัญหาการรวมตัวของผู้เล่น

กรอบแนวคิดในการแก้ปัญหาการรวมตัวของผู้เล่นในกรณีที่ผู้เล่นบางคนไม่สามารถมองเห็นได้จากกล้องใดๆ มีกระบวนการหลักแสดงดังภาพที่ 7.3 ดังนี้



ภาพที่ 7.3 แนวคิดการแก้ปัญหาการรวมตัวของผู้เล่น

กระบวนการแก้ปัญหาการรวมตัวของผู้เล่นมี 3 ขั้นตอนหลัก คือ การตรวจสอบสถานะการบดบังของผู้เล่น การทำนายตำแหน่งผู้เล่นบนพื้นฐานของข้อมูลเอาท์พุตจากอัลกอริทึมคอนเดนเซชันและการทำนายตำแหน่งผู้เล่นบนพื้นฐานของโมเดลโซเชียลฟอร์ด (Social Force Model) ในแต่ละขั้นตอนมีแนวคิดดังนี้

1.1 กระบวนการตรวจสอบสถานะการบดบังของผู้เล่น

จากกระบวนการตรวจจับและติดตามผู้เล่นในมุมมองภาพสามารถตรวจสอบได้ว่าในแต่ละกล้องมีผู้เล่นคนใดบ้างที่ถูกบดบัง โดยพิจารณาจากตำแหน่งและขนาดของพื้นที่ของแต่ละบล็อบ

ในกรณีที่ถูกบดบังเพียงบางส่วนจะดำเนินการต่อในกระบวนการที่ 1.2 ในกรณีที่ถูกบดบังโดยสมบูรณ์จะดำเนินการต่อในกระบวนการที่ 1.3 ดังสมการที่ 7.1

$$occlusion_status = \begin{cases} Full, if \left(blob^{position} \; near \; blob^{position}_{neighbor} \right) and \left(blob^{size}_{new_frame} = 0 \right) \\ partial, if \left(blob^{position} \; near \; blob^{position}_{neighbor} \right) and \left(0 < blob^{size}_{new_frame} < blob^{size}_{old_frame} \right) \end{cases} \quad (7.1)$$

1.2 กระบวนการทำนายตำแหน่งผู้เล่นบนพื้นฐานของอัลกอริทึมคอนเดนเซชัน

กรณีที่พบว่าผู้เล่นถูกบดบังแบบเพียงบางส่วนในกระบวนการตรวจจับและติดตามผู้เล่นในมุมมองภาพซึ่งอธิบายอยู่ในบทที่ 5 จากการใช้อัลกอริทึมคอนเดนเซชันในการติดตามวัตถุโดยจากการสุ่มเลือกแซมเปิลสำหรับการประมาณตำแหน่งนั้นจะได้เอาท์พุตของอนุภาคที่เกิดจากการสุ่มเลือกแซมเปิลในพื้นที่มาจำนวนหนึ่ง ซึ่งแม้วัตถุจะถูกบดบังระบบนั้นยังสามารถตรวจจับแซมเปิลที่เหลือของวัตถุได้ จากข้อมูลแซมเปิลที่เหลือนี้สามารถนำมาใช้เป็นข้อมูลพื้นฐานในการประมาณตำแหน่งของวัตถุได้ เนื่องจากโดยทั่วไปมักจะกำหนดตำแหน่งของวัตถุที่ศูนย์กลางของวัตถุ แต่โดยปกติการบดบังมักจะเริ่มจากบริเวณด้านข้างของพื้นที่จากด้านหนึ่งไปยังอีกด้านหนึ่งดังนั้นแซมเปิลที่เหลือจึงมักเป็นกลุ่มที่อยู่บริเวณด้านข้างของพื้นที่ทำให้การประมาณตำแหน่งจริงของวัตถุซึ่งจะอยู่ตรงกลางของพื้นที่อาจคลาดเคลื่อนไป ในที่นี้มีแนวคิดโดยการสร้างความสัมพันธ์ระหว่างแต่ละแซมเปิลกับตำแหน่งจริงที่ประมาณของแต่ละเฟรมโดยใช้สถิติค่าระยะทางเฉลี่ยจากเฟรมก่อนหน้าในการประมาณตำแหน่งจริงใหม่ ดังสมการที่ 7.2

$$position_{new} = mean\left(pdist\left(sample_1^N, position_{old} \right) \right) \quad (7.2)$$

1.3 กระบวนการทำนายตำแหน่งผู้เล่นบนพื้นฐานของโมเดลโซเชียลฟอร์ด

กรณีที่พบว่าผู้เล่นถูกบดบังแบบสมบูรณ์ แนวคิดในการแก้ปัญหา คือ โดยทั่วไปผู้เล่นที่อยู่บริเวณรอบนอกของกลุ่มจะยังสามารถมองเห็นได้จากกล้องที่อยู่ในทิศทางเดียวกัน ข้อมูลของผู้เล่นเหล่านี้จะถูกนำมาใช้เป็นข้อมูลพื้นฐานในการทำนายผู้เล่นภายในกลุ่มที่ถูกบดบัง โดยใช้โมเดลโซเชียลฟอร์ดในการทำนาย ซึ่งจะดำเนินการในมุมมองภาพเสมือน หลักการดำเนินงานมีอยู่ 6 ขั้นตอนหลัก คือ ขั้นตอนแรก กำหนดค่าเริ่มต้นให้กับแต่ละวัตถุ เช่น ขนาดของ Mass (น้ำหนักเป็นกิโลกรัม) รัศมีความกว้างของวงกลมรอบวัตถุซึ่งในที่นี้ได้มาจากกระบวนการสกัดเอกลักษณ์ในบทที่ 5 ขั้นตอนที่สอง ทำการคำนวณหาข้อมูลตำแหน่ง ทิศทางการเคลื่อนที่และความเร็วของผู้เล่นในกลุ่มที่ยังสามารถมองเห็นได้จากกระบวนการตรวจจับและติดตามในบทที่ 5 ขั้นตอนที่สาม ทำการคำนวณหาตำแหน่งทิศทางและความเร็วของผู้เล่นที่ถูกบดบังเพื่อทำนายตำแหน่งใหม่ที่คาดว่าจะปรากฏภายในพื้นที่ของ

การรวมกลุ่ม โดยใช้หลักของการทำนายทั่วไปของเทคนิคการติดตาม ขั้นตอนที่สี่ ทำการคำนวณหาทิศทางและความเร็วเฉลี่ยของกลุ่มเพื่อกำหนดตำแหน่งเป้าหมายที่กลุ่มจะเคลื่อนที่ไป ขั้นตอนที่ห้า การกำหนดเป้าหมายให้กับผู้เล่นที่ถูกบดบังในโมเดลโซเชียลฟอร์ด ขั้นตอนที่หก การใช้โมเดลโซเชียลฟอร์ดเพื่อทำการทำนายตำแหน่ง ทิศทางและความเร็วของผู้ที่ถูกบดบัง โดยในที่นี้จะทำนายเฉพาะผู้ที่ถูกบดบังเท่านั้น ส่วนผู้เล่นคนอื่นในกลุ่มจะใช้โมเดลโซเชียลฟอร์ดในการหาค่าแรงต่างๆ โดยอาศัยข้อมูลพื้นฐานที่ได้จากการตรวจจับและติดตาม เช่น ตำแหน่ง ทิศทางและความเร็ว จุดประสงค์เพื่อใช้ประกอบการทำนายระยะห่างของตำแหน่งระหว่างผู้เล่นที่ถูกบดบังกับผู้เล่นคนอื่นๆ ภายในกลุ่ม

2. แนวทางการปรับปรุงกระบวนการผสมข้อมูล

ในอนาคตจะทำการศึกษาและทดลองในการผสมข้อมูลในมุมมองเสมือนโดยการประมาณตำแหน่งจากข้อมูลกล้องหลายตัวพร้อมกันโดยไม่ต้องดำเนินการประมาณตำแหน่งข้อมูลทีละคู่ระหว่าง 2 กล้อง ซึ่งคาดว่าจะช่วยให้ระบบมีประสิทธิภาพมากขึ้นและลดเวลาในการประมวลผลได้ดีขึ้น นอกจากนี้ข้อมูลพื้นฐานของทิศทางของกล้องที่ได้จากอัลกอริทึมการหาทิศทางกล้องด้วยเข็มทิศ ในอนาคตอาจนำไปประยุกต์สำหรับการสร้างเครือข่ายของกล้องรอบสนามสำหรับการประยุกต์ใช้ในระบบการติดตามผู้เล่นในการแก้ปัญหาการบดบังต่อไป ส่วนวิธีประมาณตำแหน่งโดยใช้พื้นทางของเส้นตรง ในอนาคตจะทดลองใช้หลักการของเส้นโค้ง (Curve) มาประยุกต์ใช้ในการประมาณตำแหน่งซึ่งคาดว่าจะช่วยลดปัญหาในการผสมข้อมูลของวิธีที่เสนอได้

พหูน ปณฺฑิโต ชีเว

บรรณานุกรม

พหูน ปณฺ ฑิโต ชีเว

บรรณานุกรม

[1] Wang และคณะ. “An Approach of Exciting Events Detection in Soccer Sports Videos”. International Journal of Information Acquisition. Vol. 7, No. 1. pp. 15-30, March 2010.

[2] Beyan และ Temizel. “Adaptive Mean-Shift for Automated Multi Object Tracking”. IET Computer Vision. Vol. 6, Issue 1. pp. 1–12, June 2011.

[3] Zivkovic. “Improved Adaptive Gaussian Mixture Model for Background Subtraction”. Proc. 17th Int. Conf. on Pattern Recognition (ICPR), Cambridge, UK, pp. 28–33, August 2004.

[4] Cucchiara และคณะ. “Detecting Objects, Shadows and Ghosts in Video Streams by Exploiting Color and Motion Information”. Proc. 11th Int. Conf. on Image Analysis and Processing (ICIAP), Palermo, Italy, pp. 360–365, 2001.

[5] Comaniciu และคณะ. “Kernel-Based Object Tracking”, IEEE Trans. Pattern Anal. Mach. Intell., pp. 564–577, 2003.

[6] Dedeoglu. “Moving Object Detection, Tracking and Classification for Smart Video Surveillance”. Master’s thesis, Bilkent University, Department of Computer Engineering, Turkey, August 2004.

[7] Yilmaz และคณะ. “Object Tracking: A Survey”. ACM Computing Survey. Vol.38, No.4, Article 13, December 2006.

[8] Nassar และคณะ. “A Novel Feature Extraction Scheme for Human Gait Recognition”. International Journal of Image and Graphics. Vol. 10, No. 4. pp. 575–587, 2010.

[9] Huang และคณะ. “Robust Object Tracking by Hierarchical Association of Detection Responses”. Springer-Verlag Berlin Heidelberg. pp. 788–801, 2008.

พหุนปณฑิโต ชีเว

[10] Wu. “Part Based Object Detection, Segmentation, and Tracking by Boosting Simple Shape Feature Based Weak Classifiers”. Faculty of the Graduate School University of Southern California. pp. 1–216, August 2008.

[11] Ye และคณะ. “Human Detection Based on Motion Object Extraction and Head–Shoulder Feature”. Optik. pp. 1-6, 2013.

[12] Bunyak. “Moving Object Detection and Tracking for Event-Based Video Analysis”. University of Missouri-Rolla. pp. 10, 2005.

[13] Cheng. “Multi-Object Tracking via Particle Sampling and Enhanced Probabilistic Data Association for Event Detection in Intelligent Video Systems”. University of Washington. pp. 8, 2008.

[14] Yu และคณะ. “Robust Object Tracking with Occlusion Handle”. Neural Comput & Applic. pp. 1027-1034, 2011.

[15] Banerjee และ Sengupta. “Human Motion Detection and Tracking for Video Surveillance”. Department of Electronics and Electrical Communication Engineering Indian Institute of Technology, India. pp.: 1-5, 2008.

[16] Dash และคณะ. “Kernel Based Object Tracking using Color Histogram Technique”. International Journal of Electronics and Electrical Engineering, Volume 2, Issue 4. pp. 28-35, April 2012.

[17] Jalal และ Singh. “The State-of-the-Art in Visual Object Tracking”. Informatica. pp. 227-248, 2012.

[18] Ding และคณะ. “A Survey on Feature Extraction for Pattern Recognition”. Artif Intell Rev. pp. 169–180, 2012.

[19] Khan. “Multi-view Approaches to Tracking, 3D Reconstruction and Object Class Detection”. A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, University of Central Florida. pp. 5, 2008.

พหูน ปณฺ ฑิโต ชีเว

[20] Sood และคณะ. “Motion Human Detection & Tracking Based On Background Subtraction”. International Journal of Engineering Inventions. Volume 2, Issue 6. pp. 34-37, April 2013.

[21] Ali และคณะ. “An Efficient Algorithm for Detection of Soccer Ball and Players”. Department of Industrial and Management Engineering, Hankuk University of Foreign Studies, REPUBLIC of KOREA. pp. 39-46, 2012.

[22] Lin และคณะ. “Extracting 3D Information from Broadcast Soccer Video”. Image and Vision Computing. 2006.

[23] Kong และคณะ. “Learning Group Activity in Soccer Videos from Local Motion”. In: Zha H., Taniguchi R., Maybank S. (eds) Computer Vision – ACCV 2009. ACCV 2009. Lecture Notes in Computer Science, vol 5994. Springer, Berlin, Heidelberg. 2010.

[24] Utsumi และคณะ. “An Object Detection Method for Describing Soccer Games from Video”. Graduate School of Engineering, The University of Tokyo, Graduate School of Information Science and Technology, The University of Tokyo, National Institute of Informatics, Japan. ICME. 2002.

[25] Yu และคณะ. “A Player-Possession Acquisition System for Broadcast Soccer Video”. IEEE International Conference on Multimedia and Expo. 2005.

[26] Huang และคณะ. “Players and Ball Detection in Soccer Videos Based on Color Segmentation and Shape Analysis”. Springer-Verlag Berlin Heidelberg. pp. 416–425, 2007.

[27] Ngo และคณะ, “Accurate Playfield Detection Using Area-of-Coverage”, IEEE, pp. 3441-3444, 2010.

[28] Liu และคณะ, “Playfield Detection using Adaptive GMM and Its Application”, In Proceeding of International Conference on Acoustic, Speech and Signal Processing, ICASSP 2005.

พหุนฺ ปญฺญา ชีเว

[29] Barnard และ Odobez, “Robust Playfield Segmentation using MAP Adaptation”, Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, pp.610-613, Vol.3, 2004.

[30] Bin และคณะ, “Playfield Detection Algorithm Based on AGMM in Sports Video”, Journal of Networks, Vol. 8, No. 8, August 2013.

[31] Ekin และ Tekalp, “Shot Type Classification by Dominant Color for Sports Video Segmentation and Summarization”, ICASSP, 2003.

[32] Olshausen, “Bayesian Probability Theory”, March 1, 2004.

[33] Liu และคณะ, “Automatic Player Detection, Labeling and Tracking in Broadcast Soccer Video”, Pattern Recognition Letters, ELSEVIER, Available online 29 February 2008, 2009.

[34] Hung และคณะ, “Generalized Playfield Segmentation of Sport Videos using Color Features”, Pattern Recognition Letters, ELSEVIER, Available online 5 March 2011, 2011.

[35] Jiang และคณะ, “A New Method to Segment Playfield and Its Applications in Match Analysis in Sports Video”, ACM, pp. 292-295, 2004.

[36] โสภณ ผู้มีจรรยา, “การประมวลผลภาพดิจิตอล”, ที่มา: https://sites.google.com/site/dip7123512/sheet, เข้าถึงเมื่อ กุมภาพันธ์ 2557.

[37] Tahar และ Imed, “Camera Calibration using Court Models for Real-Time Augmenting Soccer Scenes”, Springer, 2011.

[38] T. Mitchell, “Machine Learning”, McGraw Hill, 1997.

[39] Wikipedia, “คอมพิวเตอร์วิทัศน์”, ที่มา: http://th.wikipedia.org/wiki/คอมพิวเตอร์วิทัศน์, เข้าถึงเมื่อ กุมภาพันธ์ 2557.

[40] Han และคณะ, “A Detection-Based Multiple Object Tracking Method”, 2004 International Conference on Image Processing, ICIP '04, 2004.

[41] Nishino และคณะ, “Tracking of Multiple Soccer Player Using a 3D Particle Filter Based on Detector Confidence”, Advances in Computer Science and Engineering, Volume 6, pp. 93-104, 2011.

[42] Iwase และ Saito, “Parallel Tracking of All Soccer Players by Integrating Detected Positions in Multiple View Images”, IEEE, 2004.

[43] Cheikh และคณะ, “Multi-People Tracking Across Multiple Cameras”, International Journal on New Computer Architectures and Their Applications (IJNCAA), 2012.

[44] Park และคณะ, “Multi-Camera Multi-Player Tracking”, International Conference on Intelligent Computing. August 23-26, HeFei, China, pp. 873-878, 2005.

[45] Choi และคณะ, “Tracking the Ball and Players from Multiple Football Videos”, International Journal of Information Acquisition, pp.1 – 8, 2006.

[46] Ren และคณะ, “Multi-Camera Video Surveillance for Real-Time Analysis and Reconstruction of Soccer Games”, Springer-Verlag, 2009. Machine Vision and Applications, pp. 855–863, 2010.

[47] Ishii และคณะ, “3D Tracking of a Soccer Ball Using Two Synchronized Cameras”, PCM, pp. 196-205, 2007.

[48] Zhao และคณะ, “Real-time Wide Area Multi-Camera Stereo Tracking”, Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), IEEE, 2005.

[49] Hayet และคณะ, “A Modular Multi-Camera Framework for Team Sports Tracking”, IEEE, pp. 493-498, 2005.

[50] Sato และ Aggarwal, “Tracking Soccer Players Using Broadcast TV Images”, IEEE, pp. 546-551, 2005.

[51] Morariu และ Camps, “Modeling Correspondences for Multi-Camera Tracking Using Nonlinear Manifold Learning and Target Dynamics”, Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06), IEEE, 2006.

พหุนปณฺ ฑิโต ชีเว

[52] Hashimoto และ Ozawa, “A System for Automatic Judgment of Offsides in Soccer Games”, IEEE, pp. 1889-1892, 2006.

[53] Spagnolo และคณะ, “An Unsupervised Approach for Segmentation and Clustering of Soccer Players”, International Machine Vision and Image Processing Conference, IEEE, pp. 133-140, 2007.

[54] Mazzeo และคณะ, “Visual Players Detection and Tracking in Soccer Matches”, IEEE Fifth International Conference on Advanced Video and Signal Based Surveillance, IEEE, pp. 326-333, 2008.

[55] Wikipedia, “ทฤษฎีกราฟ”, ที่มา: http://th.wikipedia.org/wiki/ทฤษฎีกราฟ, เข้าถึงเมื่อ กรกฎาคม 2558.

[56] Gupta และคณะ, “Camera Calibration Technique Using Tsai’s Algorithm”, International Journal of Enterprise Computing and Business Systems, ISSN (Online), Vol.1 Issue 2, 2011.

[57] สุเทพ อาษา และคณะ, “วิธีการที่มีประสิทธิผลในการสร้างรหัสสำหรับเครื่องกัดชิ้นงานด้วยคอมพิวเตอร์วิชัน”, การประชุมวิชาการเครือข่ายวิศวกรรมเครื่องกลแห่งประเทศไทย ครั้งที่ 28, pp. 369-378, 2557.

[58] Xu และคณะ, “Architecture and Algorithms for Tracking Football Players with Multiple Cameras”, Intelligent Distributed Surveilliance Systems, IEE, pp. 51 – 55, 2004.

[59] Figueroa และคณะ, “Tracking Soccer Players using the Graph Representation”, Proceedings of the 17th International Conference on Pattern Recognition, ICPR.2004, IEEE, pp. 787-790, 2004.

[60] Swain และ Ballard, “Color Indexing”, International Journal of Computer Vision, pp. 11-32, 1991.

[61] Cheung และ Kamath, “Robust Techniques for Background Subtraction in Urban Traffic Video”, Visual Communications and Image Processing 2004, 2004.

พหุปญฺญา ธีโต ชีเว

[62] Ren และคณะ, “Tracking the Soccer Ball using Multiple Fixed Cameras”, Computer Vision and Image Understanding, pp. 633–642, 2009.

[63] Zhang, “Camera Calibration”, Emerging Topics in Computer Vision, Prentice Hall Professional Technical Reference, Chapter 2, pp. 4-43, 2004.

[64] Wikipedia, “Image Segmentation”, ที่มา: http://th.wikipedia.org/wiki/Image Segmentation, เข้าถึงเมื่อ ตุลาคม 2558.

[65] Isard และ Blake, “CONDENSATION—Conditional Density Propagation for Visual Tracking”, International Journal of Computer Vision, 29(1), pp. 5-28, 1998.

[66] กฤษฎา แสงเพ็ชร์ส่อง, “แนะนำ Kalman Filter Algorithm”, เซมิคอนดักเตอร์อิเล็กทรอนิกส์. ฉบับที่ 371, pp. 126-137, 2555.

[67] Ahsan และ Mohamad, “Features Extraction for Object Detection Based on Interest Point”, TELKOMNIKA, Vol. 11, No. 5, pp.2716–2722, 2013.

[68] Salmond และ Gordon, “An Introduction to Particle Filters”, INCOMPLETE DRAFT, pp. 1–27, 2005.

[69] Petersen และคณะ, “Conditional Point Distribution Models”, Lecture Notes in Computer Science: Authors’ Instructions, pp. 2–10.

[70] wbi.msu.ac.th, “การวัดประสิทธิภาพ”, Slide Presentation, ที่มา: www.wbi.msu.ac.th/file/648/doc_31.ppt, เข้าถึงเมื่อ สิงหาคม 2558.

[71] Wikipedia, “Confusion Matrix”, ที่มา: https://en.wikipedia.org/wiki/Confusion_matrix, เข้าถึงเมื่อ สิงหาคม 2558.

[72] cs.uregina.ca, “Confusion Matrix”, ที่มา: http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html, เข้าถึงเมื่อ สิงหาคม 2558.

[73] Chris J. Needham และ Roger D. Boyle, “Tracking Multiple Sports Players through Occlusion, Congestion and Scale”, BMVC, pp. 93–102, 2001.

พหุนปณฺ ติ โต ชีเว

[74] OpenCV-Python Tutorials, “Histogram Backprojection”, ที่มา: http://opencv-python-tutroals.readthedocs.org/en/latest/py_tutorials/py_imgproc/py_histograms/py_histogram_backprojection/py_histogram_backprojection.html#histogram-backprojection, เข้าถึงเมื่อ กรกฎาคม 2558.

[75] Reynolds และคณะ, “Speaker Verification Using Adapted Gaussian Mixture Models”, Digital Signal Processing 10, pp. 1941, 2000.

[76] Martín และ Martínez, “A Semi-Supervised System for Players Detection and Tracking in Multi-Camera Soccer Videos”, Multimed Tools Appl, Springer Science + Business Media New York, 2013.

[77] Jin และ Bhanu, “Integrating Crowd Simulation for Pedestrian Tracking in a Multi-Camera System”, Center for Research in Intelligent Systems University of California, Riverside California 92521.

[78] Bernardin และ Stiefelhagen, “Evaluating Multiple Object Tracking Performance:The CLEAR MOT Metrics”, EURASIP Journal on Image and Video Processing, pp. 1-10, 2008.

[79] Kasturi และคณะ, “Performance Evaluation Protocol for Face, Person and Vehicle Detection & Tracking in Video Analysis and Content Extraction (VACE-II) Clear - Classification of Events, Activities and Relationships”, Advanced Research and Development Activity, pp. 1-24, 2006.

[80] Stegmann และ Gomez, “A Brief Introduction to Statistical Shape Analysis”, http://www.imm.dtu.dk/˜mbs/, pp. 1-15, 2002.

[81] Wikipedia, “Procrustes Analysis”, ที่มา: https://en.wikipedia.org/wiki/Procrustes_analysis, เข้าถึงเมื่อ ธันวาคม 2558.

[82] Orazio และคณะ, “A Semi-Automatic System for Ground Truth Generation of Soccer Video Sequences”, 6th IEEE International Conference on Advanced Video and Signal Surveillance, Genoa, Italy September, 2-4, 2009.

พหุน ปณฺ ทิโต ชีเว

[83] Jatoth และคณะ, “Performance Analysis of Alpha Beta Filter, Kalman Filter and Mean shift for Object Tracking in Video Sequences”, I.J. Image, Graphics and Signal Processing, 3, 24-30, 2015.

[84] Yu และคณะ, “Playfield and Ball Detection in Soccer Video”, Chapter Playfield and Ball Detection in Soccer Video, Advances in Visual Computing: Third International Symposium, ISVC, Volume 4842 of the series Lecture Notes in Computer Science, pp. 387-396, 2007.

[85] Bai และคณะ, “Playfield Detection Using Color Ratio and Local Entropy”, Conference: Seventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIHMSP 2011, Dalian, China, 14-16, October 2011.

[86] Ekin และคณะ, “Automatic Soccer Video Analysis and Summarization”, IEEE Transactions on Image Processing, vol. 12, no. 7, pp. 796-807, 2003.

[87] Zameni และคณะ, “A Low Cost Algorithm for Expected Goal Events Detection in Broadcast Soccer Video”, International Journal of Digital Content Technology and its Applications Volume 4, Number 8, November 2010.

[88] Reinhard และคณะ, “Color Transfer between Images”, Computer Graphics and Applications, IEEE, VOL. 21, NO. 5, pp. 34 - 41, 2001.

[89] Haralick และคณะ, “Textural Features for Image Classification”, IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-3, pp. 610-621, 1973.

[90] Mohanty และ Shantaiya, “A Survey on Moving Object Detection using Background Subtraction Methods in Video”, International Journal of Computer Applications (0975 – 8887National Conference on Knowledge, Innovation in Technology and Engineering (NCKITE 2015), pp. 5-10, 2015.

[91] Risha และ hempak, “Novel Method of Detecting Moving Object in Video”, Procedia Technology 24, International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015), ScienceDirect, pp.1055-1060, 2016.

[92] Jing และ Chaojian, “Survey on Methods of Moving Object Video Detection in Marine Environment”, International Conference on Computer Science and Information Technology (ICCSIT 2011), IPCSIT vol. 51 (2012), IACSIT Press, Singapore, pp.437-440, 2011.

[93] Prajapati และ Galiyawala, “A Review on Moving Object Detection and Tracking”, International Journal of Computer Application, Volume 5– No. 3, pp. 168-175, 2015.

[94] Dixit และคณะ, “Survey on Video Object Detection & Tracking”, International Journal of Current Trends in Engineering & Technology, Volume: 02, Issue: 02, pp. 264-268, 2016.

[95] Parekh และคณะ, “A Survey on Object Detection and Tracking Methods”, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 2, pp. 2970-2978, 2014.

[96] Patil และ Bhagat, “Detection and Tracking of Moving Object: A Survey”, Int. Journal of Engineering Research and Applications, Vol. 5, Issue 11, (Part - 5), pp.138-142, 2015.

[97] Sindhuja และ Devi, “A Survey on Detection and Tracking of Objects in Video Sequence”, International Journal of Engineering Research and General Science Volume 3, Issue 2, Part 2, pp.418-426, 2015.

[98] Yamada และคณะ, “Tracking Players and a Ball in Video Image Sequence and Estimating Camera Parameters for 3D Interpretation of Soccer Games”, Conference: Pattern Recognition. Proceedings. 16th International Conference on, Volume: 1, February 2002.

[99] Sterning และคณะ, “Multi-Camera Multi-Object Tracking by Robust Hough-based Homography Projections”, Computer Vision Workshops (ICCV Workshops), IEEE International Conference on, 2011.

พหูน ปณฺฑิโต ชีเว

[100] Nevhal และ Ghotkar, "A Survey on Object Detection and Tracking Algorithms", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 5, pp.9311-9318, 2017.

[101] Martin และ Martinez, "A Semi-Supervised System for Players Detection and Tracking in Multi-Camera Soccer Videos", Springer Science + Business Media New York, 2013.

[102] Nagalakshmi และคณะ, "Object Detection and Tracking in Videos: A Review", International Journal of Engineering and Computer Science, Volume 3 Issue 5, pp.5905-5912, 2014.

[103] Ren และคณะ, "Tracking the Soccer Ball using Multiple Fixed Cameras", Computer Vision and Image Understanding, 113 (5). pp. 633-642, 2009.

[104] Belagiannis และคณะ, "Segmentation Based Particle Filtering for Real-Time 2D Object Tracking", Computer Vision – ECCV 2012. ECCV 2012. Lecture Notes in Computer Science, vol 7575. Springer, Berlin, Heidelberg, pp.1-14, 2012.

[105] Rodrigues และคณะ, "A Computer Vision based Web Application for Tracking Soccer Players", work was partly supported by the Portuguese Foundation for Science and Technology (FCT), project LARSyS PEst-OE/EEI/LA0009/2013 and project FootData QREN I&DT, n. 23119, pp.1-12, 2013.

[106] Rodriguez และคณะ, "Tracking in Unstructured Crowded Scenes", ICCV.2009, IEEE International Conference on Computer, 2009.

[107] Stuart, "Microscopic Modeling of Crowds Involving Individuals with Physical Disability: Exploring Social Force Interaction", All Graduate Theses and Dissertations, 4696, Utah State University, 2015.

[108] Gauvain และ Lee, "Maximum a Posteriori Estimation for Multivariate Gaussian Mixture Observations of Markov Chains", IEEE Transactions on Speech Audio Processing, 2:291–298, April 1994.

[109] Mariethoz และ Bengio, “A Comparative Study of Adaptation Methods for Speaker Verification”, In ICSLP, Denver, USA, September 2002.

[110] Utsumi และคณะ, “An Object Detection Method for Describing Soccer Games from Video”, In ICME2002, pp. 45–48, Aug 2002.

[111] Xu และคณะ, “Algorithms and System for Segmentation and Structure Analysis in Soccer Video”, In Proc. ICME, Tokyo, Japan, Aug 22-25 2001.

[112] Nishad และ Chezian, “Various Colour Spaces and Colour Space Conversion Algorithms”, Volume 4, No. 1, pp.44-48, January 2013.

[113] IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (PETS), 2003.

พหุน ปณ ฑิโต ชีเว

ประวัติผู้เขียน

ชื่อ	นายปิยะ แก้วบัวดี
วันเกิด	วันที่ 24 ตุลาคม พ.ศ. 2522
สถานที่เกิด	ตำบลท่างิ้ว อำเภอบรรพตพิสัย จังหวัดนครสวรรค์
สถานที่อยู่ปัจจุบัน	บ้านเลขที่ 355/6 หมู่ 2 ตำบลท่างิ้ว อำเภอบรรพตพิสัย จังหวัดนครสวรรค์ รหัสไปรษณีย์ 60180
ตำแหน่งหน้าที่การงาน	อาจารย์
สถานที่ทำงานปัจจุบัน	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์ เลขที่ 145 หมู่ 15 ถนนสุรินทร์-ปราสาท ตำบลนอกเมือง อำเภอเมืองสุรินทร์ จังหวัดสุรินทร์ รหัสไปรษณีย์ 32000
ประวัติการศึกษา	พ.ศ. 2547 ประกาศนียบัตรวิชาภาษาอังกฤษเฉพาะอาชีพคอมพิวเตอร์ (ป.อ.) มหาวิทยาลัยสุโขทัยธรรมาธิราช พ.ศ. 2545 วิทยาศาสตรบัณฑิต (วท.บ.) สาขาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเจ้าพระยา พ.ศ. 2550 วิทยาศาสตรมหาบัณฑิต (วท.ม.) สาขาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยนเรศวร พ.ศ. 2562 ปริญญาปรัชญาดุษฎีบัณฑิต (ปร.ด.) สาขาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยมหาสารคาม

พหูน ปณฺ ฑิโต ชีเว